Zusammenfassung der Vorlesung

Analysis 2

Herbert Koch
Franz Gmeineder
Universität Bonn
Sommersemester 2018

Dies ist eine gekürzte Zusammenfassung und *kein* vollständiges Skript der Vorlesung. Deshalb kann diese Zusammenfassung ein Lehrbuch *nicht* ersetzen. Wie in der Vorlesung besprochen, werden folgende Bücher empfohlen:

Fo2 Forster, Analysis II, Springer

Hi2 Kaballo, Einführung in die Analysis 2,

Ta2 Terence Tao, Analysis 2, Springer 2009.

Die ersten zwei Bücher decken alle den Stoff der Vorlesung im Wesentlichen ab. Das Buch von Tao verwendet mehr Zeit für die Grundlagen, was als Ergänzung interessant sein kann.

Tippfehler und Korrekturen bitte an

`koch@math.uni-bonn.de` oder
`fgmeined@math.uni-bonn.de`

oder in den jeweiligen Sprechstunden.

Diese Zusammenfassung ist nur für Hörer der Vorlesung V2G1 Analysis 2 an der Universität Bonn, Sommersemester 2018, bestimmt.

Die zentralen Themen der Vorlesung sind metrische Räume, Differentialrechung im $\mathbb{R}^d$ und gewöhnliche Differentialgleichungen. Die Vorlesung ist in drei große Abschnitte mit diesen Titeln gegliedert. Sie setzt die Analysis 1 aus dem Wintersemester fort und baut auf der Analysis 1 und der linearen Algebra 1 auf.

Inhaltsverzeichnis

[09.04.2018]

1 Einführung in metrische Räume

Definition 1.1 (Metrischer Raum)**.** *Sei X eine Menge. Eine Metrik auf X ist eine Abbildung $d : X \times X \to [0, \infty)$ mit den Eigenschaften*

1. *$d(x, y) = 0$ genau dann wenn $x = y$.*
2. *(Symmetrie) $d(x, y) = d(y, x)$ für alle $x, y \in X$.*
3. *(Dreiecksungleichung) $d(x, z) \leq d(x, y) + d(y, z)$ für alle $x, y, z \in X$.*

Wir nennen das Paar (X, d) einen metrischen Raum.

Beispiel 1.1. $X = \mathbb{R}$, $d(x, y) = |x - y|$.

Beispiel 1.2. $X = \mathbb{C}$, $d(z, w) = |z - w|$.

Beispiel 1.3. Teilmengen metrischer Räume sind wieder metrische Räume. Sei (X, d) ein metrischer Raum und $A \subset X$. Dann definiert die Einschränkung von d auf $A \times A$, $d|_{A \times A}$, eine Metrik auf A.

Beispiel 1.4. Sei X die Menge der Städte in Deutschland, $d(x, y)$ die Entfernung auf der Straße.

Beispiel 1.5. Sei X eine Menge. Die diskrete Metrik auf X ist definiert durch $d(x, y) = 0$ falls $x = y$ und $d(x, y) = 1$ sonst.

1.1 Normierte Vektorräume

Definition 1.2 (Normierter Raum)**.** *Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Eine Norm ist eine Abbildung $\|.\| : V \times V \to [0, \infty)$ mit den Eigenschaften*

1. *$\|x\| = 0$ genau dann wenn $x = 0$.*
2. *$\|\lambda x\| = |\lambda| \, \|x\|$ für alle $\lambda \in \mathbb{K}$ and $x \in V$.*
3. *(Dreiecksungleichung) $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in V$.*

Das Paar $(V, \|.\|)$ wird normierter Vektorraum genannt.

Satz 1.1. *Sei $(V, \|.\|)$ ein normierter Vektorraum. Dann definiert*

$$d(x, y) = \|x - y\|$$

eine Metrik auf V.

Zum Beweis verifiziert man die drei Bedingungen für die Norm.

Beispiel 1.6. $(\mathbb{R}^d, |.|)$ and $(\mathbb{C}^d, |.|)$.

1.2 Offene und abgeschlossene Mengen

Definition 1.3 (offener Ball)**.** *Sei (X,d) ein metrischer Raum, $r > 0$ und $a \in X$. Die Menge $B_r(a) = \{x \in X : d(x,a) < r\}$ heißt Ball mit Mittelpunkt a und Radius r.*

Definition 1.4 (offene, abgeschlossene Mengen)**.** *$U \subset X$ wird offen genannt, wenn zu jedem $x \in U$ ein $r > 0$ existiert, so dass $B_r(x) \subset U$. $A \subset X$ heißt abgeschlossen, wenn das Komplement $X \backslash A$ offen ist.*

Beispiel 1.7. Die Bälle $B_r(a)$ sind offen.

Beispiel 1.8. Die Bälle $\overline{B_r(x)} = \{y \in X : d(x,y) \leq r\}$ sind abgeschlossen.

Beispiel 1.9. Die leere Menge und der ganze Raum sind offen und abgeschlossen.

Rechenregeln:

1. Endliche Schnitte und beliebige Vereinigungen offener Mengen sind offen.
2. Beliebige Schnitte und endliche Vereinigungen abgeschlossener Mengen sind abgeschlossen.
3. $X\backslash(A \cup B) = (X\backslash A) \cap (X\backslash B)$ und $X\backslash(\bigcup_{A\in\mathcal{A}} A) = \bigcap_{A\in\mathcal{A}}(X\backslash A)$ wobei $\mathcal{A}$ eine Menge von Teilmengen von X ist.
4. $X\backslash(A \cap B) = (X\backslash A) \cup (X\backslash B)$ und $X\backslash(\bigcap_{A\in\mathcal{A}} A) = \bigcup_{A\in\mathcal{A}}(X\backslash A)$ wobei $\mathcal{A}$ eine Menge von Teilmengen von X ist.
5. Ist $A \subset B$ so gilt $B\backslash A = (X\backslash A)\backslash(X\backslash B)$.

Definition 1.5 (Inneres)**.** *Sei (X,d) ein metrischer Raum, $A \subset X$. Das Innere $\mathring{A}$ ist die größte in A enthaltenden Teilmenge. Der Abschluß $\bar{A}$ ist die kleinste abgeschlossene Menge, die A enthält. Der Rand ist $\bar{A}\backslash(\mathring{A})$.*

Bemerkung: Das Innere von A ist die Vereinigung aller in A enthaltenen offenen Mengen. Der Abschluss von A ist der Schnitt aller A enthaltenden abgeschlossenen Mengen. Daraus ergibt sich die Wohldefiniertheit.

Lemma. *Es gilt $\mathring{A} \cap \overline{X\backslash A} = \varnothing$ und $\mathring{A} \cup \overline{X\backslash A} = X$ und $\partial A = \overline{X} \cap \overline{A}$.*

Beweis. Es gilt

$$x \in \mathring{A} \Longleftrightarrow \exists r > 0 : B_r(x) \subset A \Longleftrightarrow$$
$$\exists r > 0 : \overline{X\backslash A} \subset X\backslash B_r(x) \longleftrightarrow x \notin \overline{X\backslash A}$$

woraus die erste Formel folgt. Damit folgt auch

$$X\backslash\mathring{A} = \overline{X\backslash A}$$

und, durch Übergang zu Komplementen

$$X \backslash \overline{A} = X \overset{\circ}{\backslash} A$$

Es folgt

$$\partial A = \bar{A} \backslash \mathring{A} = \overline{X \backslash A} \backslash (X \overset{\circ}{\backslash} A) = (X \backslash \mathring{A}) \backslash (X \backslash \bar{A})$$

und daher $\partial A \subset \bar{A} \cap \overline{X \backslash A}$ und aufgrund der Definition

$$\partial A \cap \mathring{A} = \varnothing = \partial A \cap (X \backslash \bar{A}).$$

Der Beweis ist vollständig. □

[09.04.2018]

[12.04.2018]

Satz 1.2. *Es gilt immer*

1. *$x \in \mathring{A}$ genau dann, wenn x ein innerer Punkt ist, d.h. falls $r > 0$ existiert mit $B_r(x) \subset A$.*
2. *$x \in \bar{A}$ genau dann, x ein Berührpunkt ist, d.h. wenn jeder Ball um x sowohl A wie auch $X \backslash A$ schneidet, d.h., wenn für alle $r > 0$ $B_r(x) \cap A \neq \varnothing$.*
3. *$x \in \partial A$ genau dann, wenn x ein Berührpunkt von A und $X \backslash A$ ist, d.h. wenn für alle $r > 0$ weder $B_r(x) \cap A$ noch $B_r(x) \cap (X \backslash A)$ die leere Menge ist.*

Beweis. Die erste Aussage ist die Definition. Nun ist $x \in \bar{A}$ genau dann wenn x nicht in $X \backslash \bar{A}$ liegt, was genau dann der Fall ist, wenn kein $r > 0$ existiert so dass $B_r(x) \subset X \backslash A$. Also schneidet jeder Ball dieser Form A. Da $x \in \partial A$ genau dann gilt wenn $x \in \bar{A} \cap \overline{X \backslash A}$ folgt die letzte Aussage aus der zweiten Aussage. □

1.3 Konvergenz von Folgen

Definition 1.6 (Folgenkonvergenz). *Ein Folge in einem metrischen Raum ist eine Abbildung $\mathbb{N} \to X$. Wir schreiben sie wieder als $(x_n)_{n \in \mathbb{N}}$. Wir sagen, die Folge konvergiert gegen $x \in X$ falls für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass*

$$d(x_n, x) < \varepsilon$$

für alle $n \geq N$. Wir sagen, die Folge konvergiert, wenn sie gegen ein x konvergiert. Wir nennen $(x_n)_{n \in \mathbb{N}}$ Cauchyfolge, wenn für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass

$$d(x_n, x_m) < \varepsilon \qquad \text{für } n, m \geq N$$

Ein Element $y \in X$ heißt Häufungswert der Folge, wenn eine Teilfolge gegen y konvergiert.

Bemerkungen.

1. Der Grenzwert ist eindeutig wenn er existiert.

2. Jede konvergente Folge ist eine Cauchyfolge. Jede Cauchyfolge ist beschränkt.

3. Es gibt metrische Räume, in denen *nicht* jede Cauchyfolge konvergiert.

Satz 1.3. *A ist genau dann abgeschlossen, wenn aus $x_n \in A$, $x_n \to y$ folgt $y \in A$.*

Beweis. Sei A abgeschlossen, $x_n \in A$, $x_n \to y$. Dann ist $x \notin \overset{\circ}{(X \backslash A)}$ (sonst erhält man leicht einen Widerspruch) und damit $x \in \bar{A}$. Nun ist aber $A = \bar{A}$.

Nun habe jede konvergente Folge in A einen Limes in A. Wir zeigen: Dann ist $X \backslash A$ offen. Genauer zeigen wir: Ist $y \in X \backslash A$ so existert $r > 0$ so dass $B_r(y) \subset X \backslash A$. Falls nicht existiert eine Folge in A, die gegen y konvergiert. Das kann nicht sein, da nach Annahme dann $y \in A$. □

Definition 1.7 (Vollständiger metrischer Raum)**.** *Ein metrischer Raum heißt* vollständig*, wenn jede Cauchyfolge konvergiert.*

Satz 1.4. *Der normierte Vektorraum $(\mathbb{R}^d, |.|)$ ist vollständig.*

Beweis. Sei $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge. Dann ist die Folge der Koordinaten $(x_{n,j})_{n\in\mathbb{N}}$ eine reelle Cauchfolge da

$$|x_{n,j} - x_{n,j}| \leq |x_n - x_m|.$$

Wir definieren $x_j = \lim_{n\to\infty} x_{n,j}$ für alle $j = 1, ..., d$. Dann ist

$$|x - x_n| \leq \sum_{j=1}^{d} |x_j - x_{n,j}| \to 0$$

da die endliche Summe von Nullfolgen eine Nullfolge ist. □

Bemerkung: $x_n \to x$ in $\mathbb{R}^d$ genau dann wenn jede Koordinate konvergiert.

1.4 Stetige Abbildungen

Definition 1.8 (Stetigkeit). *Seien (X, d_X) und (Y, d_Y) metrische Räume. Eine Abbildung $f : X \to Y$ heißt stetig in $a \in X$ falls für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass*

$$d_Y(f(x), f(a)) < \varepsilon \qquad \textit{falls} \qquad d_X(x, a) < \delta.$$

f heißt stetig, wenn alle Urbilder offener Mengen offen sind, d.h. wenn aus $V \subset Y$ offen folgt

$$f^{-1}(V) = \{x \in X : f(x) \in V\}$$

offen ist.

Satz 1.5. *f ist genau dann stetig, wenn f in jedem Punkt stetig ist.*

Beweis. $f : X \to Y$ sei stetig in jedem Punkt, $V \subset Y$ offen, $f(a) \in V$. Da V offen ist existiert $\varepsilon > 0$ so dass $B^Y_\varepsilon(f(a)) \subset V$. Da f in A stetig ist, existiert $\delta > 0$ so dass $f(x) \in B^Y_\varepsilon(f(a))$ falls $x \in B_\delta(a)$. Damit ist f stetig.

Umgekehrt: Sei f stetig, $a \in X$, $V \subset Y$ offen, $B^Y_\varepsilon(f(a))$ wie oben. Dann ist

$$U = f^{-1}(B^Y_\varepsilon(f(a)) = \{x \in X : d_y(f(x), f(a)) < \varepsilon\}$$

offen in x und es existiert $\delta > 0$ so dass $B_\delta(a) \subset U$. Damit ist f im Punkt a stetig. □

Beispiel 1.10. Sei (X, d) ein metrischer Raum. Dann ist

$$d(x, A) := \inf\{d(x, y) : y \in A\}$$

stetig.

Beispiel 1.11. Sei $(V, \|.\|)$ ein normierter Vektorraum. Dann ist

$$x \to \|x\|$$

stetig.

Das Kartesische Produkt zweier metrischer Räume $X \times Y$ ist ein metrischer Raum mit der Metrik

$$d((x_1, y_1), (x_2, y_2)) = \max\{d_X(x_1, x_2), d_y(y_1, y_2)\}$$

ist wieder ein metrischer Raum.

Beispiel 1.12. Die Abbildung add : $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$, die zwei Zahlen auf ihre Summe abbildet ist stetig. Genauso auch die Summe in normierten Vektorräumen, die Multiplikation, und die Division mit Nenner in $\mathbb{R}\backslash\{0\}$. Die gleichen Aussagen gelten über den komplexen Zahlen.

Beispiel 1.13. Sei V ein normierter Vektorraum. Dann ist

$$\mathbb{K} \times V \ni (\lambda, x) \to \lambda x \in V$$

stetig.

Beispiel 1.14. Produkte und Quotienten stetiger Funktionen sind stetig (falls der Nenner nie Null ist).

Beispiel 1.15. Monome und Polynome in $\mathbb{R}^d$ sind stetig. Die Abbildung auf die Determinante ist stetig.

Satz 1.6. *Die Komposition stetiger Abbildungen ist stetig.*

Beweis: Wie in Analysis I.

Lemma. *Seien (X, d_X) und (Y, d_Y) metrische Räume. Dann ist $f : X \to Y$ genau dann stetig in $a \in X$, wenn für jede Folge $x_n \to a$ folgt $f(x_n) \to f(a)$.*

Der Beweis ist identisch zur analogen Aussage in Analysis I.

Satz 1.7. *Gleichmäßige Limiten stetiger Abbildungen sind stetig.*

Beweis. Seien (X, d_X) und (Y, d_Y) metrische Räume und $f_n : X \to Y$ eine Folge stetiger Abbildungen, die gleichmäßig gegen $f : X \to Y$ konvergieren (d.h. für jedes $\varepsilon > 0$ existiert $N \in \mathbb{N}$ so dass

$$d_Y(f_n(x), f(x)) < \varepsilon \qquad \text{für } x \in X \text{ und } n \geq N.$$

Sei jetzt $a \in X$, $\varepsilon > 0$ und $N \in \mathbb{N}$ so dass

$$d_Y(f_n(x), f(x)) < \varepsilon/3 \quad \text{für } x \in X, n \geq N.$$

Da f_N stetig ist, existiert $\delta > 0$ so dass

$$d_Y(f_N(x), f_N(a)) < \varepsilon/3 \quad \text{für } d_X(x, a) < \delta.$$

Für diese x folgt nun mit mehrfacher Anwendung der Dreiecksungleichung

$$\begin{aligned} d_Y(f(x), f(a)) \leq & d_Y(f_N(x), f(x)) + d_Y(f_N(x), f_N(a)) + d_Y(f_N(a), f(a)) \\ < & \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon. \end{aligned}$$

Damit ist f stetig in a. □

Satz 1.8 (Fixpunktsatz von Banach)**.** *Sei (X, d) ein vollständiger metrischer Raum, $f : X \to X$ eine Kontraktion, d.h. es existiert ein $\theta < 1$ so dass*

$$d(f(x), f(y)) \leq \theta d(x, y)$$

für alle $x, y \in X$ gilt. Dann gibt es genau ein $x \in X$ so dass

$$x = f(x)$$

gilt.

Beweis. Wir konstruieren rekursiv eine Folge durch $x_0 \in X$,

$$x_{n+1} = f(x_n).$$

Es folgt

$$d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq \theta d(x_n, x_{n-1}) \leq \theta^n d(x_1, x_0)$$

und, für $m \geq n$,

$$d(x_{m+1}, x_n) \leq \sum_{j=n}^{m} d(x_{j+1}, x_j) \leq \sum_{j=0}^{m-n} \theta^j d(x_{n+1}, x_n) \leq \frac{\theta^n}{1-\theta} d(x_1, x_0),$$

woraus folgt, dass (x_n) eine Cauchyfolge ist. Wir definieren $x = \lim_{n\to\infty} x_n$. Die Abbildung f ist stetig (wähle $\delta = \epsilon$) und daher gilt

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n) = f(x).$$

Es gelte nun auch $y = f(y)$. Daraus folgt

$$d(x, y) = d(f(x), f(y)) \leq \theta d(x, y)$$

was nur wahr sein kann, wenn $d(x, y) = 0$ und damit $x = y$. Der Beweis ist vollständig.

□

[12.04.2018]

[16.04.2018]

2 Euklidsche und normierte Vektorräume

2.1 Euklidsche Vektorräume

Definition 2.1 (Euklidische Vektorräume)**.** *Ein Euklidscher Vektorraum ist ein $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ -Vektorraum V mit einer Abbildung $\langle ., . \rangle : V \times V \to \mathbb{K}$ mit den Eigenschaften*

1. *Für alle $y \in V$ ist $x \to \langle x, y \rangle$ eine lineare Abbildung.*
2. *Es gilt immer $\langle x, y \rangle = \overline{\langle y, x \rangle}$*
3. *Es gilt immer $\langle x, x \rangle \geq 0$. Es gilt und $\langle x, x \rangle = 0$ genau dann wenn $x = 0$.*

Wir definieren dann $\|x\| = \sqrt{\langle x, x \rangle}$. $\langle ., . \rangle$ heißt inneres Produkt.

Manchmal nennt man $\mathbb{C}$-Vektorräume mit den obigen Eigenschaften auch *unitäre Vektorräume.*

Satz 2.1 (Cauchy-Schwarz-Ungleichung)**.** $\|.\|$ *ist eine Norm und es gilt für alle* $x, y \in V$ *die Cauchy-Schwarz-Ungleichung*

$$|\langle x, y\rangle| \leq \|x\|\|y\|.$$

Beweis. Aus $\|x\| = 0$ folgt $\langle x, x\rangle = 0$ und nach der dritten Forderung $x = 0$. Für $\lambda \in \mathbb{K}$ und $x \in V$ gilt

$$\|\lambda x\|^2 = \langle \lambda x, \lambda x\rangle = \lambda\langle x, \lambda x\rangle = \lambda\overline{\langle \lambda x, x\rangle} = \lambda\overline{\lambda\langle x, x\rangle} = |\lambda|^2\|x\|^2$$

Wir zeigen die Cauchy-Schwarzungleichung. Seien $x, y \in V$, $y \neq 0$. Wir definieren

$$z = x - \frac{\langle x, y\rangle}{\langle y, y\rangle}y$$

Man sieht leicht: $\langle z, y\rangle = 0 = \langle y, z\rangle$. Es folgt

$$\|x\|^2 = \|z + \frac{\langle x, y\rangle}{\langle y, y\rangle}y\|^2 = \|z\|^2 + \frac{|\langle x, y\rangle|^2}{\|y\|^2}$$

woraus

$$|\langle x, y\rangle|^2 \leq \|x\|^2\|y\|^2$$

und damit die Cauchy-Schwarzungleichung folgt.
Zuletzt verifizieren wir die Dreiecksungleichung:

$$\begin{aligned}\|x + y\|^2 =& \langle x + y, x + y\rangle \\ =& \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ \leq& \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ =& (\|x\| + \|y\|)^2\end{aligned}$$

□

Definition 2.2. *Wir nennen* x *und* y *orthogonal, falls* $\langle x, y\rangle = 0$.

2.2 Banachräume

Definition 2.3 (Banachraum)**.** *Ein Banachraum ist ein normierter Vektorraum, der als metrischer Raum vollständig ist.*

Beispiel 2.1.

$\mathbb{R}^d$ and $\mathbb{C}^d$ sind Banachräume.

Sei M eine Menge. Die beschränkten Funktionen $\mathcal{B}(M) = \{f : M \to \mathbb{R}, f \text{ beschränkt } \}$ sind ein Banachraum mit der Suppremumsnorm

$$\|f\|_M = \sup\{|f(x)| : x \in M\}$$

Satz 2.2. *Auf einem metrischen Raum (X, d) betrachten wir den Raum der stetigen beschränkten Funktionen $C_b(X)$ mit der Supremumsnorm. $C_b(X)$ ist ein Banachraum.*

Beweis. Da Summen und Vielfache stetiger Funktionen stetig sind, und da $C_b(X) \subset \mathcal{B}(X)$ ist $C_b(X)$ ein Vektorraum und die Supremumsnorm eine Norm. Seien f_n eine Cauchyfolge. Dann ist für jedes $x \in X$ auch $f_n(x)$ eine reelle Cauchyfolge. Dies ist auch eine Cauchyfolge in $\mathcal{B}(X)$ und f_n konvergiert gleichmässig gegen die Limesfunktion $f \in \mathcal{B}(X)$. Nach Satz 1.7 ist f stetig. Damit konvergiert f_n gegen f in $C_b(X)$. □

Bemerkung: Wir können $\mathbb{R}$ ohne wesentliche Änderungen im Beweis durch einen Banachraum ersetzen.

[16.04.2018]

[19.04.2018]

2.3 Beschränkte lineare Abbildungen

Es seien $(V, \|.\|_V)$ und $(W, \|.\|_W)$ normierte Vektorräume über $\mathbb{R}$ oder $\mathbb{C}$.

Definition 2.4 (Beschränkte lineare Abbildungen)**.** *Wir nennen eine lineare Abbildung $A : V \to W$ beschränkt, wenn*

$$\|A\|_{V\to W} = \|A\|_{L(V,W)} = \sup\{\|f(x)\|_W : \|f\|_X \le 1\} = \|A|_{\overline{B_1^X(0)}}\|_{\overline{B_1^X(0)}} < \infty$$

Wir bezeichen den Raum der beschränkten linearen Abbildungen mit $L(V, W)$.

Bemerkungen:

(i) Es genügt, das Supremum über den Rand der Einheitskugel zu nehmen. Da für $x \neq 0$

$$\|Ax\|_W = \|x\|_V \|A(\frac{1}{\|x\|_V} x)\|_W \le \|A\|_{V\to W} \|x\|_X$$

können wir $\|A\|_{V\to W}$ auch als die kleinste Konstante definieren, für die diese Ungleichung gilt.

(ii) Eine lineare Abbildung $A : V \to W$ ist genau dann stetig, wenn sie beschränkt ist.

Satz 2.3. *Die beschränkten linearen Abbildungen sind ein normierter Vektorraum mit der angegebenen Norm. Ist W ein Banachraum, so ist auch $L(V,W)$ ein Banachraum.*

Beweis. Vielfache beschränkter linearer Abbildungen sind beschränkte lineare Abbildungen und genauso Produkte. Damit ist $L(V,W)$ ein Vektorraum. Durch die Einschränkung können wir beschränkte lineare Abbildungen als W wertige beschränkte Abbildungen auf dem abgeschlossenen Einheitsball in W auffassen. Damit ist auch $\|.\|_{V\to W}$ eine Norm. Sei A_n ein Cauchyfolge. Dann ist für jedes x Ax eine Cauchyfolge in Y. Ist Y vollständig so existiert ein Grenzwert und wir definieren $Ax = \lim_{n\to\infty} A_n x$. Man überzeugt sich leicht davon, dass A eine lineare Abbildung ist. Da die Einschränkung auf den abgeschlossenen Einheitsball konvergiert ist A eine beschränkte lineare Abbildung und $A_n \to A$ in $L(V,W)$. □

Wir nennen $\|.\|_{L(V,W)}$ die Operatornorm. Sind $(V_j, \|.\|_{V_j})$ für $j = 1,2,3$ normierte Vektorräume, und ist $S \in L(V_1,V_2)$ and $T \in L(V_2,V_3)$ so ist $T\mathring{S} \in L(V_1,V_3)$ und es gilt

$$\|TS\|_{V_1\to V_3} \le \|S\|_{V_1\to V_2}\|T\|_{V_2\to V_3}. \tag{2.1}$$

2.4 Matrizen und lineare Abbildungen

Seien $N,M \in \mathbb{N}$, $N,M \ge 1$ und $A = (a_{nm})_{1\le n\le N, 1\le m\le M} \in M_{n\times M}(\mathbb{R})$ eine reelle $N \times M$ Matrix. Wir versehen die Matrizen mit der Hilbert-Schmidtnorm

$$\|A\|_{HS} = \sqrt{\sum_{n=1}^{N}\sum_{m=1}^{M} |a_{nm}|^2}$$

Mit dieser Norm und dem offensichtlichen inneren Produkt sind die $N \times M$ Matrizen ein euklidscher Vektorraum.

Wir identifizieren $N \times M$ Matrizen mit linearen Abbildungen von $\mathbb{R}^M$ nach $\mathbb{R}^N$. Es ist leicht zu sehen, dass diese linearen Abbildungen beschränkt sind. Die Operatornorm definiert eine zweite Norm auf den Matrizen.

Satz 2.4. *Es gilt*

$$\|A\|_{\mathbb{R}^M\to\mathbb{R}^N} \le \|A\|_{HS} \le \sqrt{M}\|A\|_{\mathbb{R}^M\to\mathbb{R}^N}$$

Beweis. Die erste Ungleichung folgt mit der Cauchy-Schwarz-Ungleichung

$$\begin{aligned}
|Ax|^2_{\mathbb{R}^N} =& \sum_{n=1}^N \left(\sum_{m=1}^M a_{nm} x_m \right)^2 \\
\leq & \sum_{n=1}^N \left(\sum_{m=1}^N |a_{nm}|^2 \right) \left(\sum_{m'=1}^M |x_m|^2 \right) \\
=& \|A\|^2_{HS} |x|^2.
\end{aligned}$$

Aus der triviallen Abschätzung der Summe der Quadrate folgt

$$\|A\|^2_{HS} \leq \sqrt{M} \max\{\sum_{n=1}^N |a_{nm}|^2 : 1 \leq m \leq M\}.$$

Da

$$\sum_{n=1}^N |a_{nm}|^2 \leq \|Ae_m\|^2 \leq \|A\|^2_{\mathbb{R}^M \to \mathbb{R}^N}$$

erhalten wir die zweite Ungleichung. □

In diesen Überlegungen können wir $\mathbb{R}$ durch $\mathbb{C}$ ersetzen. Wir betrachten nun zwei Beispiele.

[19.04.2018]

[23.04.2018]

Satz 2.5. *Es sei A eine invertierbare $N \times N$ Matrix und B eine $N \times N$ Matrix für die gilt:*

$$\|B\|_{\mathbb{R}^N \to \mathbb{R}^N} < \|A^{-1}\|^{-1}_{\mathbb{R}^N \to \mathbb{R}^N}.$$

Dann ist $A - B$ invertierbar und es gilt

$$(A - B)^{-1} = \sum_{j=0}^\infty A^{-1}(BA^{-1})^j.$$

Insbesondere konvergiert die Reihe (d.h. die Folge der Partialsummen konvergiert). Es gilt

$$\|A - B\|_{\mathbb{R}^N \to \mathbb{R}^N} \leq \frac{\|A^{-1}\|_{\mathbb{R}^N \to \mathbb{R}^N}}{1 - \|A^{-1}\|_{\mathbb{R}^N \to \mathbb{R}^N} \|B\|_{\mathbb{R}^N \to \mathbb{R}^N}}.$$

Man überzeugt sich wie bei den reellen Zahlen, dass die Folge der Partialsummen $\sum_{j=0}^\infty x_n$ in einem Banachraum konvergiert falls die Reihe

$$\sum_{j=0}^\infty \|x_n\| < \infty$$

konvergiert.

Beweis. Wir verifizieren zunächst die Konvergenz und wenden (2.1) mehrfach an:

$$\|A^{-1}(BA^{-1})^j\|_{\mathbb{R}^N\to\mathbb{R}^N} \le \|A^{-1}\|_{\mathbb{R}^N\to\mathbb{R}^N}\|(BA^{-1})^j\|_{\mathbb{R}^N\to\mathbb{R}^N}$$

und rekursiv

$$\|(BA^{-1})^J\|_{\mathbb{R}^N\to\mathbb{R}^N} \le \|BA^{-1}\|_{\mathbb{R}^N\to\mathbb{R}^N} \le \|B\|_{\mathbb{R}^N\to\mathbb{R}^N}\|A^{-1}\|_{\mathbb{R}^N\to\mathbb{R}^N} =: \theta < 1$$

Daraus erhalten wir

$$\|A^{-1}(BA^{-1})^j\|_{\mathbb{R}^N\to\mathbb{R}^N} \le \theta^j\|A^{-1}\|_{\mathbb{R}^N\to\mathbb{R}^N}$$

und die Konvergenz folgt mit der geometrischen Reihe, genauso wie die Abschätzung.

Da

$$(A-B)\sum_{j=0}^{n} A^{-1}(BA^{-1})^j = 1_{\mathbb{R}^N} - (BA^{-1})^{n+1} \to 1_{\mathbb{R}^N}$$

ist die Reihe das Inverse zu $A - B$. □

Wir erhalten zwei interessante Schlussfolgerungen: Die Menge der invertierbaren Matrizen ist offen (folgt direkt aus dem Satz) und die Abbildung auf die inverse Matrix stetig. Zum Nachweis der Stetigkeit sei A eine invertierbar Matrix und $r = \frac{1}{2\|A^{-1}\|}$. Ist $\|B\| \le r$ so folgt mit dem Satz $\|(A-B)^{-1}\| \le 2\|A^{-1}\|$. Da

$$(A-B)^{-1} - A^{-1} = (A-B)^{-1}BA^{-1}$$

folgt

$$\|(A-B)^{-1} - A^{-1}\| \le 2\|A^{-1}\|^2\|B\|.$$

Sei $\varepsilon > 0$ und

$$\delta = \min\{\delta\|A^{-1}\|^{-2}/2, \frac{1}{2\|A^{-1}\|}\}.$$

Ist $\|C - A\| < \delta$ so folgt

$$\|C^{-1} - A^{-1}\| < \varepsilon$$

und damit die Stetigkeit. Alle Normen hier sind die Operatornormen.

Als zweites Beispiel betrachten wir die Matrixexponentialfunktion der $N\times N$ Matrix A:

$$\exp(A) = \sum_{j=0}^{\infty} \frac{A^j}{j!}$$

Da wieder

$$\|\frac{A^j}{j!}\| = \frac{1}{j!}\|A^j\| \le \frac{1}{j!}\|A\|^j$$

erhalten wir

$$\|\exp A\| \le \exp(\|A\|) \tag{2.2}$$

und die Exponentialreihe konvergiert. Wie bei der reellen und der komplexen Exponentialfunktion folgt

$$\exp(A+B) = \exp A \exp B$$

falls $AB = BA$, da wir bei der reellen Exponentialfunktion die Kommutativität die Multiplikation ausgenutzt haben. Im Allgemeinen ist die Identität falsch. Wir halten die wichtigsten Aussagen fest.

Satz 2.6. *Die Reihe der Matrixexponentialfunktion konvergiert und es gilt* (2.2)*. Für $s, t \in \mathbb{K}$ gilt*

$$\exp((s+t)A) = \exp(sA)\exp(tA)$$

$$\exp(0_{\mathbb{K}^N}) = 1_{\mathbb{K}^N}$$

$$\lim_{t\to 0} \frac{\exp(tA) - 1_{\mathbb{R}^N}}{t} = A$$

Für den Beweis verwenden wir, dass $(tA)(sA) = (sA)(tA)$. Für den Limes argumentieren wir genauso wie bei der reellen Exponentialfunktion.

Beispiele:

1. $N = 1$. Wir erhalten die gewöhnliche Exponentialfunktion.
2. $N = 2$. Wir betrachten zuerst spezielle Matrizen der Form $A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ und berechnen

   $$A^2 = (a^2+bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

   und ersetzen die Quadrate in der Reihe durch die Multiplikation:

   $$\begin{aligned} \exp A &= \sum_{j=1}^{\infty} \frac{1}{(2j)!} A^{2j} + \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} A^{2n+1} \\ &= \sum_{j=1}^{\infty} \frac{(a^2+bc)^j}{(2j)!}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sum_{n=0}^{\infty} \frac{(a^2+bc)^n}{(2n+1)!} A \\ &= \frac{e^{\mu}+e^{-\mu}}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{e^{\mu}-e^{-\mu}}{2\mu} \sum_{n=0}^{\infty} \frac{(a^2+bc)^{2n}}{(2n+1)!} A \end{aligned}$$

   wobei μ eine komplexe Zahl ist mit dem Quadrat $\mu^2 = a^2 + bc$. Die letzte Identität folgt mit einem Koeffizientenvergleich.

3. Der Fall der Diagonalmatrix.

$$\exp\begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_N \end{pmatrix} = \begin{pmatrix} e^{a_1} & 0 & \dots & 0 \\ 0 & e^{a_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{a_N} \end{pmatrix}$$

4. Der Fall eines Jordanblocks

$$\exp\begin{pmatrix} 0 & t & 0 & \dots & 0 \\ 0 & 0 & t & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & t \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} = \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{N-1}}{(N-1)!} \\ 0 & 1 & t & \dots & \frac{t^{N-2}}{(N-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & t \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

3 Kompakte Mengen

Definition 3.1 (Kompaktheit)**.** *Wir nennen einen metrischen Raum* (X, d) *kompakt, wenn jede offene Überdeckung eine endliche Teilüberdeckung besitzt. Eine offene Überdeckung ist eine Menge* $\mathcal{A}$ *von offenen Teilmengen von* X*, so dass*

$$X \subset \bigcup_{U \in A} U$$

von diesen Mengen überdeckt, d.h. falls jedes $x \in X$ *in mindestens einer dieser Teilmengen liegt. Eine endliche Teilüberdeckung ist eine endliche Menge dieser offenen Mengen, die* X *immer noch überdeckt.*

Wir nennen X *folgenkompakt, wenn jede Folge eine konvergente Teilfolge besitzt.*

Wir nennen X *präkompakt, wenn für jedes* $r > 0$ *endlich viele Punkte* $(x_j)_{1 \leq j \leq n}$ *existieren, so dass*

$$X \subset \bigcup_{j=1} B_r(x_j).$$

Satz 3.1. *Für einen metrischen Raum* $(X.d)$ *sind die folgenden Aussagen äquivalent:*

1. (X, d) *ist kompakt.*

2. (X, d) *ist folgenkompakt.*

3. (X, d) *ist vollständig und präkompakt.*

Beweis. 1. Schritt. Sei (X,d) kompakt und $(x_n)_{n\in\mathbb{N}}$ eine Folge in X. Wir nehmen an, die Folge habe keinen Häufungspunkt und führen das zu einem Widerspruch. Ist x kein Häufungspunkt, so gibt es ein $r(x) > 0$ so dass in $B_{r(x)}(x)$ nur endlich viele Folgenglieder liegen. Da wir annehmen, es gäbe keinen HP, so gibt es einen deratigen Ball für jedes x und diese Bälle überdecken X - sogar die Mittelpunkte überdecken X. Da (X,d) kompakt ist gibt es ein $(x_j)_{1\le j\le n}$ so dass

$$X \subset \bigcup_{j=1}^{n} B_{r(x_j)}(x_j).$$

Das ist absurd, da dann die Folge nur endlich viele Folgenglieder haben kann.

2. Schritt. Sei (X,d) folgenkompakt. Wir führen einen Widerspruchsbeweis und nehmen an, es gäbe $r > 0$ so dass es keine Überdeckung mit endlich vielen Bällen mit Radius r gibt. Wir konstruieren rekursiv eine Folge $(x_n)_{n\in\mathbb{N}}$ mit der Eigenschaft

$$d(x_n, x_m) \ge r$$

falls $n \ne m$. Wir nehmen an, wir haben die Punkte (x_n) für $1 \le n \le N$ gefunden. Da es keine endliche Überdeckung mit $N+2$ Bällen mit Radius r gibt muss es einen Punkt x_{N+1} geben, der zu allen früheren Mittelpunkten mindestens den Abstand r hat. Diese Folge (x_n) hat keine konvergente Teilfolge.

[23.04.2018]

[26.04.2018]

3. Schritt. Sei (X,d) folgenkompakt. Dann konvergent insbesondere jede Cauchyfolge und X is vollständig.

4. Schritt. Sei (X,d) vollständig und präkompakt. Sei $\mathcal{A}$ eine Menge offener Teilmengen, die X überdecken. Wir nehmen an, es gäbe keine endliche Teilüberdeckung und führen das zu einem Widerspruch. Sei $\mathcal{B}$ die Menge aller Teilmengen, die nicht von endlich vielen Mengen in $\mathcal{A}$ übedeckt werden. Dann ist insbesondere $X \in \mathcal{B}$. Wir kontruieren eine Folge $(x_n)_{n\in\mathbb{N}}$ so dass

$$D_n = \bigcap_{j=1}^{n} B_{1/j}(x_j) \in \mathcal{B}.$$

Wir machen dies rekursiv und nehmen an, wir haben $(x_j)_{1\le j\le n}$ gefunden. Sei $r = \frac{1}{n+1}$. Da X präkompakt ist existieren $(y_k)_{1\le k\le M}$ so dass

$$X \subset \bigcup_{k=1}^{M} B_{\frac{1}{n+1}}(y_k).$$

Für mindestens ein k_0 ist

$$D_{n+1} = D_n \cap B_{\frac{1}{n+1}}(y_{k_0}) \in \mathcal{B}$$

da sonst D_n von endlich vielen Mengen in $\mathcal{A}$ überdeckt würde. Wir wählen $x_{n+1} = y_{k_0}$. Die Mengen D_n sind nicht leer. Daher muss

$$d(x_n, x_m) \leq \frac{1}{n} + \frac{1}{m}$$

gelten und (x_n) ist eine Cauchyfolge. X ist vollständig und die Folge konvergiert gegen ein $x \in X$. Da die Mengen in $\mathcal{A}$ X überdecken existiert ein $U \in \mathcal{A}$ mit $x \in U$. Da U offen ist, existiert $r > 0$ so dass $B_r(x) \in U$. Das ist aber ein Widerspruch zu $D_m \in \mathcal{B}$ für $m > \frac{2}{r}$ da dann $D_m \subset B_r \subset U$ folgt. $\square$

Wir sagen, eine Teilmenge $A \subset X$ ist kompakt, wenn der metrische Raum $(A, d|_{A\times A})$ kompakt ist. Die folgenden Beobachtungen sind nützlich.

1. $U \subset A$ ist genau dann offen in $(A, d|_{A\times A})$ wenn eine offene Menge $V \subset X$ existiert, so dass $U = V \cap A$.
2. Eine Teilmenge $A \subset X$ eines vollständigen metrischen Raumes ist genau dann vollständig, wenn sie abgeschlossen ist. Dies folgt aus der Charakterisierung abgeschlossener Mengen in Satz 1.3.
3. Der Durchmesser kompakter Mengen X ist beschränkt: Sei $x_0 \in X$. Dann überdecken die Bälle $B_n(x_0)$ den Raum X. Da X kompakt ist, gibt es eine endliche Teilüberdeckung, und daher folgt $X \subset B_n(x_0)$ für ein n. Insbesondere folgt

$$\sup_{x,y\in X} d(x, y) \leq 2n.$$

Satz 3.2. *Eine Teilmenge eines kompakten metrischen Raumes ist genau dann kompakt, wenn sie abgeschlossen ist. Seien (X, d_X) und (Y, d_Y) metrische Räume, (X, d) kompakt und $f : X \to Y$ stetig. Dann ist das Bild*

$$f(X) = \{y : \text{ es existiert } x \in X \text{ mit } y = f(x)\}$$

kompakt.

Satz 3.3. *(Satz vom Maximum) Eine stetige Funktion auf einer kompakten Menge nimmt das Maximum an.*

Die Dimension eines Vektorraums V ist die Kardinalität einer Basis. Die Dimension ist also genau dann endlich, wenn es eine bijektive lineare Abbildung $\phi \in L(\mathbb{K}^N, V)$ gibt. Ist V ein normierter Vektorraum, so definiert

$$\|x\| = \|\phi x\|_V$$

eine Norm auf $\mathbb{K}^N$.

Satz 3.4. *Sei V ein endlichdimensionaler reeller Vektorraum mit den Normen $\|.\|_1$ und $\|.\|_2$. Dann existiert $C > 1$ so dass*

$$C^{-1}\|x\|_1 \le \|x\|_2 \le C\|x\|_1$$

für alle $x \in V$ gilt.

Dieser Satz hat die interessante Konsequenz, dass beide Normen dieselben offenen, abgeschlossen und kompakten Mengen, Cauchyfolgen und konvergente Folgen definieren. Wir sagen, die Normen sind äquivalent.

Beweis. Mit der obigen Beobachtung genügt es, $V = \mathbb{R}^N$ zu betrachten. Es genügt weiterhin, den Fall $\|.\|_1 = |.|$ zu betrachten (wir wenden die Aussage in diesem Spezialfall zweimal an, um die allgemeine Aussage zu erhalten). Sei also $\|.\|$ eine zweite Norm auf $\mathbb{R}^N$. Sei (e_j) die Standardbasis. Es gilt dann

$$\|\sum x_j e_j\| \le \sum_{j=1}^N |x_j| \max \|e_j\| \le \sqrt{N} \max \|e_j\| |x|$$

und wir erhalten die erste Ungleichung. Mit der Dreiecksungleichung

$$\|x\| - \|y\| \le \|x - y\| \le \sqrt{N} \max \|e_j\| |x - y|$$

und die Abbildung $\mathbb{R}^N \ni x \to \|x\|$ ist stetig. Nach dem Satz vom Maximum gibt es ein $x \in \partial B_1(0) \subset \mathbb{R}^N$ so dass

$$-\|x_0\| \ge -\|y\|$$

für jedes $y \in \partial B_1(0)$. Also folgt

$$|x|\|x_0\| \le |x|\|x/|x|\| = \|x\|$$

und wir erhalten die zweite Ungleichung. □

30.04.2018

4 Wege und Kurven

Vorbemerkungen: Wir hatten die Ableitung über den Differenzenquotienten defininiert. In dieser Definition können wir Abbildungen nach $\mathbb{R}$ durch Abbildungen nach $\mathbb{R}^d$ oder allgemeiner in Banachräume ersetzen ohne wesentliche Änderungen in den folgenden Aussagen und Beweisen.

Genauso können wir das Riemannintegral für stetige Funktionen in den $\mathbb{R}^d$ ohne weitere Änderungen definieren. Dann gilt für $f : [a, b] \to \mathbb{R}^d$ stetig:

$$\left(\int_a^b f(t)dt\right)_j = \int_a^b f_j(t)dt$$

wobei links die j-te Komponente des Integral steht (das Integral ist ein Vektor in $\mathbb{R}^d$).

Definition 4.1 (Weg)**.** *Sei (X, d) ein metrischer Raum. Ein Weg ist eine stetige Abbildung eines Intervalls*

$$\gamma : I \to X.$$

Wir nennen das Bild

$$(\gamma) = \{x \in X : \exists t \in I \text{ mit } \gamma(t) = x$$

Bahn von γ. Ein injektiver Weg wird einfacher Weg oder Jordanweg genannt. Ein Weg $\gamma : [a, b] \to X$ heißt geschlossener Jordanweg, wenn $\gamma|_{[a,b)}$ injektiv ist und $\gamma(a) = \gamma(b)$. Wir sagen, γ parametrisiert die Kurve (γ).

Wir nennen (X, d) wegzusammenhängend, wenn es zu je zwei Punkten $x, y \in X$ einen Weg $\gamma : [0, 1] \to X$ gibt mit $\gamma(0) = x$ und $\gamma(1) = y$.

Beispiel 4.1. $I = [0, 2\pi)$

$$\gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

Beispiel 4.2. $I = [-1, 1]$,

$$\gamma(t) = \begin{pmatrix} t^3 \\ t^2 \end{pmatrix}$$

Satz 4.1. *Seien $\gamma_1 : I \to X$ und $\gamma_2 : J \to X$ Wege mit $(\gamma_1) = (\gamma_2)$. Dann existiert eine bijektive Abbildung $\alpha : I \to J$, die $\mathring{I}$ bijektiv auf $\mathring{J}$ abbildet und dort streng monoton ist.*

Die Kurve bestimmt also die Parametrisierung bis auf eine streng monotone Abbildug α.

Definition 4.2 (Länge eines Weges)**.** *Sei $\gamma : I \to X$ ein Weg. Wir definieren die Länge des Weges*

$$L(\gamma) = \sup\{\sum_{j=1}^{N} d(\gamma(t_j), \gamma(t_{j-1})) : t_j \in I, t_0 < t_1 < \ldots t_N\} \in [0, \infty].$$

Wir nennen $\tau = (t_0, t_1, \ldots t_N)$ wie oben Zerlegung. In der Analysis 1 hatten wir Zerlegungen und Verfeinerungen im Rahmen des Riemannintegrals betrachtet. Zu je zwei Zerlegungen gibt es eine gemeinsame Verfeinerung. Bei einer Verfeinerung wird die Summe im Supremum nicht kleiner. Es genügt also, feine Zerlegungen zu betrachten.

Für einfache Wege hängt $L(\gamma)$ offensichtlich nur von der Bahn (γ) und nicht der Parametrisierung ab.

Satz 4.2. *Sei $\gamma : [a,b] \to \mathbb{R}^d$ differenzierbar mit stetiger Ableitung. Dann gilt*

$$L(\gamma) = \int_a^b |\gamma'(t)| dt$$

Beispiel 4.3. Wir kommen zurück auf das erste Beispiel.

$$\gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

$$\gamma'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}$$

und daher $|\gamma'(t)| = 1$. Mit $I = [0, T]$ erhalten wir

$$L(\gamma) = T$$

Definition 4.3. *Ein Jordanweg γ heißt regulär, wenn γ differenzierbar ist mit stetiger Ableitung, und wenn immer $|\gamma'(t)| \neq 0$ gilt.*

Satz 4.3. *Sei $\gamma : [a,b] \to X$ ein regulärer Jordanweg. Dann existiert*

$$\alpha : [0, L(\gamma)] \to [a,b]$$

bijektiv, differenzierbar mit stetiger Ableitung und $\alpha'(t) > 0$ für alle t so dass

$$|(\gamma \circ \alpha)'(t)| = 1 \qquad \text{in } [0, L(\gamma)].$$

Beweis. α ist die Umkehrabbildung von

$$[a,b] \ni s \to \int_a^s |\gamma'(\tau)| d\tau.$$

□

03.05.2018

Definition 4.4 (Rektifizierbarkeit)**.** *Wir nennen einen Weg γ bzw die Bahn eines Jordanweges rekitifizierbar, wenn $L(\gamma) < \infty$.*

Beispiel 4.4. Sei $A \in M_{d\times d}(\mathbb{K})$ eine reelle oder komplexe $d \times d$ Matrix. Wir betrachten

$$f(t) = \exp(tA)$$

Dann gilt

$$\begin{aligned} f'(t) =& \frac{d}{dt} \exp(tA)(t) \\ =& \lim_{h\to 0} \frac{1}{h}(\exp((t+h)A) - \exp(tA)) \\ =& \lim_{h\to 0} \frac{1}{h}(\exp(hA)\exp(tA) - \exp(tA)) \\ =& \lim_{h\to 0} \frac{1}{h}(\exp(hA) - \exp(0))\exp(tA) \\ =& A\exp(tA) \end{aligned}$$

nach Satz 2.6.

Definition 4.5. *Sei $A \subset \mathbb{R}^d$. Ein Vektorfeld auf A ist eine Abbildung $A \to \mathbb{R}^d$, die an jeden Punkt einen Vektor anheftet.*

Im folgenden sei $\gamma : [a,b] \to \mathbb{R}^d$ ein differenzierbarer Weg mit stetiger Ableitung.

Definition 4.6. *Sei $f : (\gamma) \to \mathbb{R}$ stetig. Wir definieren das Wegintegral über γ bezüglich der Bogenlänge durch*

$$\int_\gamma f ds = \int_a^b f(\gamma(s))|\gamma'(s)|ds$$

Sei $F : (\gamma) \to \mathbb{R}^d$ stetig. Wir definieren das Wegintegral

$$\int_\gamma F d\vec{x} = \int_a^b \sum_{j=1}^d F_j(\gamma(t))\gamma_j'(t)dt$$

Beispiel 4.5. Sei $\gamma(t)$ die Position eines Fahrzeuges zur Zeit t. Wir nehmen an, γ sei regulärer Jordanweg. Für $x \in (\gamma)$ sei $f(x)$ der Verbrauch des Fahrzeuges in Liter je Kilometer. Dann ist der gesamte Verbrauch zwischen dem Zeitpunkt t_0 und t_1 (mit γ definiert auf $[t_0, t_1]$

$$\int_\gamma f ds$$

in geeigneten Einheiten.

Beispiel 4.6. Ist γ ein regulärer Weg mit $|\gamma'| = 1$, so gilt

$$\int_\gamma F d\vec{x} = \int_\gamma \sum_{j=1}^d F_j \gamma_j' ds$$

Lemma. *Sei $F : [a,b] \to \mathbb{R}^d$ stetig. Dann gilt:*

$$\left| \int_a^b F(t)dt \right| \leq \int_a^b |F(t)|dt$$

Beweis. Sei $w \in \mathbb{R}^d$. Dann ist

$$\sum_{j=1}^d \left(\int_a^b F(t)dt \right)_j w_j = \int_a^b \sum_{j=1}^d F_j w_j dt$$

Wir nehmen nun Beträge:

$$\left| \sum_{j=1}^d \left(\int_a^b F(t)dt \right)_j w_j \right| \leq \int_a^b |F(t)||w|dt \leq |w| \int_a^b |F(t)|dt$$

Wenn wir $w = \int_a^b F dt$ nehmen, so erhalten wir

$$\left| \sum_{j=1}^{d} \left(\int_a^b F(t)dt \right)_j w_j \right|^2 \leq |w| \int_a^b |F(t)|dt \left| \sum_{j=1}^{d} \left(\int_a^b F(t)dt \right)_j w_j \right|$$

woraus die Aussage folgt. □

Wir erinnern an den folgenden Satz:

Satz 4.4 (Fixpunktsatz von Banach)**.** *Sei (X, d) ein vollständiger metrischer Raum, $f : X \to X$ eine Kontraktion, d.h. es existierte $\theta < 1$ so dass*

$$d(f(x), f(y)) \leq \theta d(x, y)$$

für alle $x, y \in X$ gilt. Dann gibt es genau ein $x \in X$ so dass

$$x = f(x)$$

gilt.

Ein typisches Vektorfeld auf γ ist durch die Ableitung γ' eines regulären Weges gegeben. Wir wenden uns nun der Frage zu: Gibt es zu jedem Vektorfeld und Anfangspunkt einen Weg, dessen Ableitung mit dem Vektorfeld übereinstimmt?

Konkret: Sei $F : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$. Wir bezeichnen Punkte in $\mathbb{R} \times \mathbb{R}^d$ mit (t, x). Sei $x_0 \in X$. Gesucht ist nun $\gamma : \mathbb{R} \to \mathbb{R}^d$ differenzierbar, so dass

1. $\gamma(0) = x_0$
2. $\gamma'(t) = F(t, \gamma(t))$

Zunächst formulieren wir das Problem um: Falls γ existiert, so gilt auch nach dem Hauptsatz

$$\gamma(t) = x_0 + \int_0^t F(s, \gamma(s))ds \tag{4.1}$$

Umgekehrt: Genügt γ dieser Integralgleichung, so ist nach dem Hauptsatz γ differenzierbar mit stetiger Ableitung, und es gilt $\gamma(0) = x_0$ und $\gamma'(t) = F(t, \gamma(t))$.

Wir treffen nun die Annahme: F ist stetig und es existiert $C > 0$ so dass

$$|F(t, 0)| \leq C \tag{4.2}$$

$$|F(t, x) - F(t, y)| \leq C|x - y|. \tag{4.3}$$

und betrachten die Abbildung

$$C_b(\mathbb{R},\mathbb{R}^d) \ni f \to Tf = e^{-2C|t|}x_0 + e^{-2C|t|}\int_0^t F(s, e^{2C|s|}f(s))ds \in C_b(\mathbb{R},\mathbb{R}^d).$$

Der Raum $C_b(\mathbb{R},\mathbb{R}^d)$ ist ein Banachraum nach Satz **??**.
Es gilt nun

$$\|Tf\|_{C_b(\mathbb{R},\mathbb{R}^d)} \le |x_0| + C\sup_t \int_0^t e^{-2C|t-s|}ds + C\left|\sup_t \int_0^t e^{-2C(|t|-|s|)}ds\right|$$

und

$$\|Tf - Tg\|_{C_b(\mathbb{R},\mathbb{R}^d)} \le \frac{1}{2}\|f-g\|_{C_b(\mathbb{R},\mathbb{R}^d)}.$$

Nach dem Fixpunktsatz von Banach, Satz 4.4, hat diese Abbildung genau einen Fixpunkt, d.h., es gibt genau eine Funktion f für die

$$f(t) = e^{-2C|t|}x_0 + e^{-2C|t|}\int_0^t F(s, e^{2C|s|}f(s))ds$$

Diese Gleichung ist äquivalent dazu, dass $\gamma(t) = e^{2Ct}f(t)$ (4.1) erfüllt. Nach dem Hauptsatz ist γ dann stetig differenzierbar, und genügt dem Anfangswertproblem.

Satz 4.5 (Picard)**.** *F sei stetig und genüge* (4.2) *und* (4.3)*. Sei $x_0 \in \mathbb{R}^d$. Dann existiert genau ein $\gamma : \mathbb{R} \to \mathbb{R}^d$ stetig mit stetiger Ableitung. das das Anfangswertproblem*

1. $\gamma(0) = x_0$
2. $\gamma'(t) = F(t, \gamma(t))$

löst.

Beweis. Da wir nicht annehmen, dass $e^{-2C|t|}\gamma \in C_b$ gilt, müssen wir noch nachprüfen, dass es keine weiteren Lösungen gibt. Es genügt, die obigen Überlegungen auf einem konpakten Intervall durchzuführen, das 0 enthält. □

Beispiel:

$$\frac{d}{dt}U(t) = A(t)U(t) \qquad U(0) = 1$$

Die eindeutige Lösung ist $U(t) = \exp(tA)$. Wir können aber auch die hinter dem Satz von Banach stehende Iteration ausführen: $U_0 = 1$,

$$U_1 = 1 + \int_0^t A$$

$$U_k = \sum_{j=0}^{k} \frac{1}{k!} t^k A^k.$$

[26.04.2018]

[06.05.2018]

5 Die totale Ableitung

Idee und Ziel: Verallgemeinerung des aus der Analysis 1 bekannnten Differentierbarkeitsbegriffs von reellen Funktion auf Abbildungen $f : U \to \mathbb{R}^M$, $U \subset \mathbb{R}^d$ offen.

Erinnerung: Sei $I \subset \mathbb{R}$ ein offenes Intervall. $f : I \to \mathbb{R}$ heißt differenzierbar in $a \in \mathbb{R}$, falls der Grenzwert der Differenzenquotienten

$$\lim_{x \to a, x \neq a} \frac{f(x) - f(a)}{x - a}$$

existiert. Wir nennen den Grenzwert Ableitung in x_0. Ist f in jedem Punkt differenzierbar, so nennen wir f in I differenzierbar.
In der Analysis I haben wir eine lineare Approximationseigenschaft als alternative Definition der Differenzierbarkeit kennengelernt.

Satz 5.1. *Sei $U \subset \mathbb{R}^d$ offen, $f : U \to \mathbb{R}^m$ eine Abbildung und $a \in U$. Dann sind folgende Aussagen äquivalent:*

1. *Es gibt eine lineare Abbildung $L : \mathbb{R}^d \to \mathbb{R}^m$ mit*

$$\lim_{x \to x_0, x \neq x_0} \frac{f(x) - f(a) - L(x-a)}{|x - a|} = 0.$$

2. *Es gibt eine lineare Abbildung $L : \mathbb{R}^d \to \mathbb{R}^m$ und eine Abbildung $\phi : U \to \mathbb{R}^m$ so dass*

   (a) $f(x) = f(a) + L(x-a) + \phi(x)$ in U

   (b) $\lim_{\substack{x \to x_0 \\ x \neq a}} \frac{\phi(x)}{|x-a|} = 0$.

3. *Es gibt eine in a stetige lineare Abbildung $\tilde{\delta} : U \to L(\mathbb{R}^d; \mathbb{R}^m)$ so dass*

$$f(x) = f(a) + \tilde{\delta}(x)(x - a).$$

Beweis. (1) $\Longrightarrow$ (2): Setze $\phi(x) := f(x) - f(a) - L(x-a)$.

(2) $\Longrightarrow$ (3): Definiere für $x \in U$

$$\tilde{\delta}(x)(v) := \begin{cases} L(v) + \frac{\langle x-a, v\rangle}{|x-a|^2}\phi(x) & \text{falls } x \neq a \\ L(v) & \text{falls } x = a \end{cases}$$

Damit erhalten wir

$$f(x) = f(x) + \tilde{\delta}(x)(x-a)$$

und

$$(\delta(x) - L)(x-a) = \phi(x) \to 0 \qquad \text{für } x \to a.$$

Damit erhalten wir die Stetigkeit.
(3) $\Longrightarrow$ (1). Wir definieren $L = \tilde{\delta}(a)$. Dann ist

$$f(x) - f(a) - L(x-a) = \tilde{\delta}(x)(x-a) - \tilde{\delta}(a)(x-a) = (\tilde{\delta}(x) - \tilde{\delta}(a))(x-a)$$

und damit

$$\left|\frac{f(x) - f(a) - L(x-a)}{|x-a|}\right| = |\tilde{\delta}(x) - \tilde{\delta}(a)| \to 0 \qquad \text{für } x \to a.$$

□

Bemerkung: Die lineare Abbildung L ist eindeutig bestimmt, wenn eine der Bedingungen erfüllt ist. Das folgt aus dem Beweis.

Definition 5.1 (Totale Differenzierbarkeit, Jacobimatrix)**.** *Sei $U \subset \mathbb{R}^d$ offen und $f : U \to \mathbb{R}^m$ eine Abbildung. Wir nennen f in $a \in U$ total differenzierbar, falls eine der äquivalenten Bedingungen des Satzes erfüllt ist. Ist f in jedem Punkt in U total differenzierbar, so nennen wir f in U total differenzierbar. Die lineare Abbildung L nennen wir totale Ableitung in a und bezeichnen sie mit $Df(a)$. Die zugehörige Matrix heißt Jacobimatrix oder Funktionalmatrix.*

Satz 5.2. *Sei $U \subset \mathbb{R}^d$ offen, $a \in U$, $f : U \to \mathbb{R}^m$ sei in a total differenzierbar. Dann ist f in a stetig.*

Beweis. Das folgt direkt aus der Eigenschaft (3). □

Beispiel 5.1. Sei $f(x) = c, c \in \mathbb{R}^m$ Setze $\tilde{\delta} = 0$ und damit ist $Df(x) = 0$ für alle $x \in \mathbb{R}^d$.

Beispiel 5.2. Sei $f : \mathbb{R}^d \to \mathbb{R}^m$ linear. Dann ist $f(x) - f(a) = f(x-a)$. Setze $\tilde{\delta}(x) = f$. Dann ist $Df(a) = f$.

Beispiel 5.3. Sei $A \in M_{d \times d}$ symmetrisch und definiere

$$f(a) = x^T Ax$$

Dann ist

$$\begin{aligned} f(x) - f(a) &= x^T Ax - a^T Aa \\ &= (x-a)^T Ax * a^T A(x-a) \\ &= x^T A(x-a) + a^T A(x-a) \\ &= (x+a)^T A(x-a). \end{aligned}$$

Damit ist $\tilde{\delta}(x) = (x+a)^T A$ und $Df(a) = 2a^T A$ wobei wir Matrizen und lineare Abbildungen identifizieren.

Beispiel 5.4. Sei A eine reelle invertierbare $d \times d$ Matrix und B eine $d \times d$ Matrix mit

$$\|B\|_{\mathbb{R}^d \to \mathbb{R}^d} \|A^{-1}\|_{\mathbb{R}^d \to \mathbb{R}^d} < 1$$

Dann ist $A - B$ invertierbar und es gilt

$$(A-B)^{-1} = \sum_{n=0}^{\infty} A^{-1}(BA^{-1})^j$$

wobei die Reihe konvergiert. Insbesondere ist die Menge der invertierbaren matrizen offen. Die Abbildung auf die inverse Abbildung ist stetig. Wir rechnen für invertierbare Matrizen A, C

$$C^{-1} - A^{-1} = C^{-1}(A-C)A^{-1} = \tilde{\delta}(C)[C-A]$$

wobei $\tilde{\delta}(C)$ die durch

$$\tilde{\delta}(C)[H] = -C^{-1}HA^{-1}$$

definierte lineare Abbildung ist. Wir bezeichnen die Abbildung auf die Inverse mit f. Dann gilt

$$Df(A)[H] = -A^{-1}HA^{-1}.$$

Satz 5.3. *Sei $U \subset \mathbb{R}^d$ offen, $f, g : U \to \mathbb{R}^n$ in a total differenzierbar und $r, s \in \mathbb{R}$. Dann ist auch $rf + sg$ in a differenzierbar und es gilt*

$$D(rf + sg)(a) = rDf(a) + rDg(a).$$

Der Beweis folgt direkt aus der Definition.

Satz 5.4 (Kettenregel)**.** *Sei $U \subset \mathbb{R}^d$, $V \subset \mathbb{R}^k$ offen $f : U \to V$ sei in $a \in U$ total differenzierbar, $g : V \to \mathbb{R}^m$ sei in $f(a)$ total differenzierbar. Dann ist $g \circ f$ in a total differenzierbar und es gilt*

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

Beweis. Sei $b = f(a)$,

$$f(x) = f(a) + \tilde{\delta}_f(x)(x-a)$$
$$g(y) = g(b) + \tilde{\delta}_g(b)(y-b).$$

Es folgt

$$(g \circ f)(x) - g \circ f(a) = \tilde{\delta}_g(f(x))(f(x) - f(a)) = (\tilde{\delta}_g(f(x)) \circ \delta_f(x))(x-a)$$
$$Dg \circ f(a) = \delta_g(b)\delta_f(a).$$

Der Beweis ist vollständig. □

[06.05.2018]

[14.05.2018]

6 Die partielle Ableitung

Sei $U \subset \mathbb{R}^d$, $f : U \to \mathbb{R}^m$.

Definition 6.1 (Partielle Differenzierbarkeit)**.** *Sei $a \in U$, $1 \leq j \leq d$. Wir sagen, f ist in a in Richtung e_j partiell differenzierbar, falls der Limes*

$$\lim_{h \to 0, h \neq 0} \frac{f(a + he_j) - f(a)}{h} =: \frac{\partial f}{\partial x_j}(a) = \left(\frac{\partial f_k}{\partial x_j}(a)\right)_{1 \leq k \leq m}$$

existiert. f ist in a partiell differenzhierbar, wenn die Limiten in alle Richtungen e_j existieren. f ist in U partiell differenzierbar, wenn f in jedem Punkt partiell differenzierbar ist. Wir nennen f stetig partiell differenzierbar, wenn alle partiellen Ableitungen stetig sind.

Satz 6.1. *Ist f in a total differenzierbar, so ist f in a auch partiell differenzierbar. Die Jacobimatrix ist*

$$\left(\frac{\partial f_k}{\partial x_j}(a)\right)_{1 \leq k \leq m, 1 \leq j \leq d} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \frac{\partial f_1}{\partial x_d}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \dots & \frac{\partial f_m}{\partial x_d}(a) \end{pmatrix}.$$

Beweis. Sei f in a total differenzierbar mit Jacobimatrix A. Sei $1 \leq j \leq d$. Dann gilt

$$\lim_{h \to 0, h \neq 0} \left| \frac{f(a + he_j) - (f(a) + A(He_j)}{|h|} \right| = 0$$

und damit

$$\frac{\partial f}{\partial x_j}(a) = Ae_j,$$

Wir erhalten die angegebene Jacobimatrix. □

Definition 6.2 (Richtungsableitung)**.** *Sei $a \in U$, $v \in \mathbb{R}^d$. Die Richtungsableitung in Richtung v ist*

$$\lim_{h \to 0, h \neq 0} \frac{f(a+hv) - f(a)}{h} = \frac{\partial f}{\partial v}(a)$$

falls der Grenzwert existiert.

Bemerkung: Ist f in a total differenzierbar, so ist es auch in a in jede Richtung differenzierbar. Der Beweis ist analog zur Aussage mit den partiellen Aussagen in Satz 6.1

Beispiel 6.1. Wir definieren $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x) = \begin{cases} 0 & \text{falls } x = 0 \\ \frac{x_1 x_2}{|x|} & \text{falls } x \neq 0 \end{cases}$$

Dann ist

$$\frac{\partial f}{\partial x_1}(x) \begin{cases} 0 & \text{falls } x = 0 \\ \frac{x_2}{|x|} - \frac{x_1^2 x_2}{|x|^3} & \text{falls } x \neq 0 \end{cases}$$

$$\frac{\partial f}{\partial x_2}(x) \begin{cases} 0 & \text{falls } x = 0 \\ \frac{x_1}{|x|} - \frac{x_1 x_2^2}{|x|^3} & \text{falls } x \neq 0 \end{cases}$$

Damit ist f partiell differenzierbar. Da

$$f((t,t)^T) = \frac{1}{\sqrt{2}}|t|$$

ist f in Null nicht in Richtung $(1,1)^T$ differenzierbar und damit auch nicht in $x = 0$ total differenzierbar.

Beispiel 6.2. Wir definieren $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x) = \begin{cases} 0 & \text{falls } x = 0 \\ \frac{x_1 x_2^2}{|x|^2} & \text{falls } x \neq 0 \end{cases}$$

Sei $v \in \mathbb{R}^2 \backslash \{0\}$. Dann ist

$$f_v(t); = f(tv) = t \frac{v_1 v_2^2}{v_1^2 + v_2^2}$$

differenzierbar und alle Richtungsableitungen existieren in jedem Punkt. Liegt v auf einer Koordinatenachse, so ist $f_v(t) = 0$ für $t \in \mathbb{R}$. Damit verschwinden alle partiellen Ableitungen. Wäre f in 0 differenzierbar, so wäre die totale Ableitung die Nullabbildung, da alle Einträge der Jacobimatrix verschwinden. Dann wären auch alle Richtungsableitungen in $x = 0$ Null, und das ist ein Widerspruch. Daher ist f in $x = 0$ nicht total differenzierbar, obwohl alle Richtungsableitungen existieren.

Satz 6.2. *Ist f stetig partiell differenzierbar, so ist f auch total differenzierbar*

Beweis. Es genügt, den Fall $m = 1$ zu betrachten, da f offensichtlich in einem Punkt genau dann total differenzierbar ist, wenn dies für jede Komponente gilt. Sei also $m = 1$ und $x \neq a$. Wir definieren $\hat{x}_j \in \mathbb{R}^d$ für $0 \leq j \leq d$ durch

$$\hat{x}_j = \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ a_{j+1} \\ \vdots \\ a_d \end{pmatrix}.$$

Dann gilt

$$f(x) - f(a) = \sum_{j=1}^{d} f(\hat{x}_j) - f(\hat{x}_{j-1}).$$

Mit dem Mittelwertsatz der Differentialrechnung (angewandt in der j Koordinate) erhalten wir

$$f(\hat{x}_j) - f(\hat{x}_{j-1}) = \partial_j f(\xi_j)(x_j - a_j)$$

für ein ξ_j zwischen x_{j-1} und x_j. Es folgt

$$f(x) - f(a) - \sum_{j=1}^{d} \frac{\partial f}{\partial x_j}(a)(x - a) = \sum_{j=1}^{d} (\frac{\partial f}{\partial x_j}(\xi_j) - \frac{\partial f}{\partial x_j}(a))(x_j - a_j).$$

Sei $\varepsilon > 0$. Dann existiert $r > 0$ so dass wegen der Stetigkeit der Ableitungen der Betrag der rechten Seite nicht größer als $\varepsilon|x - a|$ ist. Damit folgt die totale Differenzierbarkeit. $\square$

6.1 Der Satz von Schwarz

Definition 6.3. *Sei $k \in \mathbb{N}$ und $U \subset \mathbb{R}^d$ offen. Eine Funktion $F : U \to \mathbb{R}^d$ heißt $k + 1$ mal stetig partiell differenzierbar, wenn F k-mal differenzierbar ist, und wenn die partiellen Ableitungen der Ordnung k partiell differenzierbar sind. Wir sagen, F ist k mal stetig differenzierbar, wenn die k ten partiellen Ableitungen stetig sind.*

Insbesondere ist $F : U \to \mathbb{R}^m$ zweimal partiell differenzierbar, wenn die partiellen Ableitungen $\frac{\partial}{\partial x_j} f$ partiell differenzierbar sind, was wir als

$$\frac{\partial}{\partial x_k} \frac{\partial}{\partial x_j} f$$

schreiben. Es gilt der wichtige Satz

Satz 6.3 (Satz von Schwarz)**.** *Sei $U \subset \mathbb{R}^d$ offen und $f : U \to \mathbb{R}^m$ zweimal stetig partiell differenzierbar. Dann gilt*

$$\frac{\partial}{\partial x_k}\frac{\partial}{\partial x_j} f(x) = \frac{\partial}{\partial x_j}\frac{\partial}{\partial x_k} f(x)$$

Beweis. Es genügt, die Ableitungen im Punkt $x = 0$ to betrachten, unter der Annahme $0 \in U$. Außerdem dürfen wir uns auf den Fall $m = 1$ beschränken, da es genügt, die Aussage für Komponenten zu beweisen. Da nur zwei Ableitungen vorkommen, reicht es aus, den Fall $d = 2$ zu betrachten. Sei also $f : U \to \mathbb{R}$ zweimal stetig differenzierbar, $U \subset \mathbb{R}^2$ offen und $0 \in U$. Der Einfachheit halber betrachten wir $U = \mathbb{R}^2$, um die Notation einfacher zu halten. Wir definieren

$$F_y(x_1) = f\begin{pmatrix} x_1 \\ y \end{pmatrix} - f\begin{pmatrix} x_1 \\ 0 \end{pmatrix}$$

Nach dem Mittelwertsatz der Differentialrechnung existiert ξ zwischen 0 und x_1 so dass

$$F_y'(\xi) = F_y(x_1) - F_y(0).$$

Nun ist

$$F_y'(x_1) = \frac{\partial f}{\partial x_1}\begin{pmatrix} \xi \\ y \end{pmatrix} - \frac{\partial f}{\partial x_1}\begin{pmatrix} \xi \\ 0 \end{pmatrix}.$$

Wir wenden den Mittelwertsatz der Differentialrechnung auf die Abboldung

$$y \to \frac{\partial f}{\partial x_1}\begin{pmatrix} \xi \\ y \end{pmatrix}$$

an:

$$\frac{\partial}{\partial x_2}\frac{\partial}{\partial x_1} f\begin{pmatrix} \xi \\ \eta \end{pmatrix} = \frac{\partial f}{\partial x_1}\begin{pmatrix} \xi \\ y \end{pmatrix} - \frac{\partial f}{\partial x_1}\begin{pmatrix} \xi \\ 0 \end{pmatrix}.$$

Durch Vertauschen der Rollen von x_1 und x_2 erhalten wir

$$\frac{\partial}{\partial x_2}\frac{\partial}{\partial x_1} f\begin{pmatrix} \xi \\ \eta \end{pmatrix} x_1x_2 = f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - f\begin{pmatrix} x_1 \\ 0 \end{pmatrix} - f\begin{pmatrix} 0 \\ x_2 \end{pmatrix} + f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2} f\begin{pmatrix} \tilde{\xi} \\ \tilde{\eta} \end{pmatrix} x_1x_2.$$

Ist $x_1x_2 \neq 0$ erhalten wir

$$\frac{\partial}{\partial x_2}\frac{\partial}{\partial x_1} f\begin{pmatrix} \xi \\ \eta \end{pmatrix} x_1x_2 = \frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2} f\begin{pmatrix} \tilde{\xi} \\ \tilde{\eta} \end{pmatrix} x_1x_2.$$

Lassen wir $x_1, x_2 \to 0$ streben, so erhalten wir die erwünschte Gleichheit. Der Beweis ist vollständig.

□

[14.05.2018]

[25.05.2018]

7 Lokale Extrema und die Taylorformel

Sei $U \subset \mathbb{R}^d$ offen, $f : U \to \mathbb{R}$ in a partiell differenzierbar. Wir definieren den Gradienten

$$\nabla f(a) = \begin{pmatrix} \partial_{x_1} f(a) \\ \dots \\ \partial_{x_x} f(a) \end{pmatrix} = (Df(x))^T.$$

Ist f zweimal stetig differenzierbar, so definieren wir die Hessematrix

$$Hf(x) = \left(\partial_{x_i}\partial_{x_j} f(x)\right)_{1\le i,j\le d}.$$

Nach dem Satz von Schwarz ist die Hessematrix symmetrisch.

Definition 7.1. *Sei (X,d) ein metrischer Raum. Wir sagen, f besitzt in a ein lokales Maximum (Minimum), falls eine offene Umgebung $V \subset U$ von a existiert (d.h. eine offene Menge V, die a enthält) mit*

$$f(x) \le f(x) \quad (f(x) \ge f(a)) \qquad \text{für alle } x \in V.$$

a heißt lokale Maximal(Minimal)stelle und $f(a)$ lokales Maximum (Minimum). Dieses heißt isoliert falls für $a \neq x \in V$ $f(x) \neq f(a)$ gilt. In beiden Fällen nennen wir a (isolierte) Extremalstelle.

Satz 7.1. *$a \in U$ sei lokale Extremalstelle. Ist f in a partiell differenzierbar, so ist $\nabla f(a) = 0$.*

Beweis. Für $1 \le j \le d$ betrachten wir $\phi(t) = f(a + te_j)$. $t = 0$ is lokale Extremalstelle, und damit $\phi'(0) = 0$, woraus $\partial_{x_j} f(a) = 0$ folgt. □

Wir nennen a einen kritischen Punkt von f, falls $\nabla f(a) = 0$.
Sei jetzt f zweimal stetig differenzierbar. Für $v \in \mathbb{R}^d$ definieren wir

$$\phi(t) = f(a + tv).$$

Damit gilt

$$\phi'(t) = \partial_v f(a + th) = Df(a + th)(v)$$

und

$$\phi''(t) = \sum_{i,j=1}^{d} \partial_{x_i}\partial_{x_j} f(a + th) v_i v_j = v^T Hf(a + th)v. \tag{7.1}$$

Nach der Taylorformel gilt

$$\phi(1) = \phi(0) + \phi'(0) + \int_0^1 \phi''(t)(1-t)dt$$

woraus

$$f(a+v) = f(a) + Df(a)v + \int_0^1 (1-t)v^T Hf(a+tv)vdt.$$

Ist a ein kritischer Punkt, so ist der zweite Term auf der rechten Seite 0.

Definition 7.2. *Sei A eine symmetrische $d \times d$ Matrix. a heißt*

1. *postiv definit, falls $v^T Av > 0$ für alle $v \neq 0$*
2. *positiv semidefinit, falls $v^T Av \geq 0$ für alle $v \in \mathbb{R}^d$*
3. *negativ (semi) definit, falls $-A$ positiv (semi) definit.*
4. *indefinit, wenn A weder positiv noch negativ semidefinit ist.*

Wir bezeichnen den normierten Vektorraum der symmetrischen $d \times d$ Matrizen mit $S(d)$. Wir schreiben $A > B$ für $A, B \in S(d)$ und $A - B$ positiv definit, und $A \geq B$ für $A - B$ positiv semidefinit.

Lemma.
- *Für $v \in \mathbb{R}^d$ ist die Menge $\{A \in S(d) : v^T Av > 0$ offen.*
- *Ist $A \in S(d)$ positiv definit, so existiert $\alpha > 0$ so dass für $C \in S(d)$ mit $\|C - A\|_{\mathbb{R}^d \to \mathbb{R}^d} < \alpha$ gilt*

$$v^T Av \geq \alpha |v|^2$$

Beweis. Die erste Aussage folgt aus der Stetigkeit von $A \to v^T Av$. Sei jetzt $A > 0$. Dann ist

$$\alpha = \inf\{v^T Av : |v| = 1\}/2 > 0$$

da $\{v : |v| = 1\}$ kompakt ist, $v^T Av > 0$ für alle v vom Betrag 1, und nach dem Satz vom Maximum das Minimum auf $\partial B_1(0)$ angenommen wird. Ist $\|C - A\|_{\mathbb{R}^d \to \mathbb{R}^d} < \alpha$ so folgt

$$v^T CV = v^T Av + v^T (C - A)v \geq 2\alpha - \alpha = \alpha$$

für $|v| = 1$, woraus

$$v^T Cv \geq \alpha |v|^2$$

folgt. $\square$

Satz 7.2. *Sei $U \subset \mathbb{R}^d$ offen, $f : U \to \mathbb{R}$ zweimal stetig differenzierbar, $a \in U$ sei ein kritischer Punkt und $A = Hf(a)$ die Hessematrix in a. Es gilt*

1. *$A > 0 \quad \Longrightarrow f$ hat ein isoliertes Minimum in a.*
2. *f hat ein lokales Minimum in $a \Longrightarrow A \geq 0$*

3. $A < 0 \quad \Longrightarrow f$ *hat ein isoliertes Maximum in* a.

4. f *hat ein lokales Maximum in* $a \Longrightarrow 0 \geq A$

5. A *indefinit* $\Longrightarrow a$ *ist keine Extremalstelle.*

Beweis. a) Wegen $Hf(a) > 0$ gibt es nach dem Lemma ein α so dass

$$v^t C v \geq \delta |v|^2$$

für alle $C \in S(d)$ mit $\|C - Hf(a)\|_{\mathbb{R}^d \to \mathbb{R}^d} \leq \alpha$ und alle $v \in \mathbb{R}^d$. Da $x \to Hf(x)$ stetig ist, existiert $\delta > 0$ so dass $v^T Hf(x) v \geq \alpha |v|^2$ for $x \in B_\delta(a)$. Damit folgt nach (7.1) dass

$$f(x) \geq f(a) + \frac{\alpha}{2}|x-a|^2$$

für $x \in B_\delta(a)$.

b) Hat f ein lokales Minimum, so hat $\phi(t)$ für jedes v ein lokales Minimum in 0. Es folgt $\phi''(0) \geq 0$ und damit $Hf(a) \geq 0$.

c) und d) folgen aus a) und b) angewamdt auf $-f$. e) folgt aus b) und *d*). □

[25.05.2018]

[28.05.2018]

Satz 7.3 (Hurwitz). *Für* $A \in S(d)$ *und* $1 \leq k \leq d$ *sei* $A_k = (a_{ij})_{1 \leq i,j \leq k} \in S(k)$ *die linke obere* $k \times k$ *Matrix von* A. *DAnn gilt*

1. $\det A_k > 0$ *für alle* $k \Longleftrightarrow A > 0$

2. $(-1)^k \det A_k > 0$ *für alle* $k \Longleftrightarrow A < 0$

3. Falls $1 \leq k \leq d/2$ *existiert mit* $\det A_{2k} < 0$, *so ist* A *indefinit.*

Beweis. Später in Kapitel 10. □

Wie betrachten zur konkreten Anwendung dieses Satzes von Hurwitz ein

Beispiel 7.1. Betrachte $f\colon \mathbb{R}^3 \to \mathbb{R}$ gegeben durch

$$f\colon \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto 6xy - 3y^2 - 2x^3 - yz^2.$$

Wir wollen die lokalen Extrema dieser Funktion herausfinden. Wir berechnen dann den Gradienten sowie die Hessematrix als

$$Df(x,y,z) = \begin{pmatrix} 6y - 6x^2 \\ 6x - 6y - z^2 \\ -2yz \end{pmatrix}, \quad Hf(x,y,z) = \begin{pmatrix} -13x & 6 & 0 \\ 6 & -6 & -2z \\ 0 & -2z & -2y \end{pmatrix}.$$

Nun gilt $Df(x,y,z)=0$ genau dann, wenn $y=x^2$, $yz=0$ sowie $z^2=6x-6y$ allesamt erfüllt sind. Gilt $y=0$, so muss wegen der ersten Gleichung auch $x=0$ und dann wegen der dritten Gleichung auch $z=0$ gelten. Fall $y\neq 0$, so muss wegen der zweiten Gleichung sodann notwendigerweise $z=0$ gelten. Damit folgt $x=y$ nach der dritten Gleichung und daher mit der ersten $x^2=x$, $x=y$, $z=0$. Es gibt nun zwei Möglichkeiten: $x\in\{0,1\}$. Ist $x=0$, so führt und das gerade auf $x=y=z=0$, was wir oben schon erhalten hatten. Ist $x=1$, so folgt $y=1$ und $z=0$. Also sind die einzigen kritischen Punkte gegeben durch

$$\xi_1 := \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{sowie} \quad \xi_2 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Um nun herauszufinden, ob es sich in diesen Punkten um lokale Minimal- bzw. Maximalstellen handelt, verwenden wir das Kriterium nach Hurwitz. Die Hessematrix in ξ_1 lautet

$$Hf(\xi_1) = \begin{pmatrix} 0 & 6 & 0 \\ 6 & -6 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Die Determinanten der $1., 2., 3.$ Abschnittsmatrizen lauten in dieser Reihenfolge: $0, -36, 0$. Damit ist $Hf(\xi_1)$ indefinit und es liegt in ξ_1 *kein* lokales Extremum vor. In ξ_2 erhalten wir hingegen

$$Hf(\xi_2) = \begin{pmatrix} -12 & 6 & 0 \\ 6 & -6 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Hier rechnen wir wieder für die Determinanten in dieser Reihenfolge $1., 2., 3.$ Abschnittsmatrizen: $-12, 36, -72$. Nach dem Satz von Hurwitz ist Hf in ξ_2 also negativ definit, also liegt in ξ_2 ein isoliertes lokales Maximum der Funktion f vor.

7.1 Höhere Ableitungen und der Satz von Taylor

Unser Ziel ist nun, die Taylorformel aus der Analysis 1 auf Funktionen $f\colon \mathbb{R}^d \to \mathbb{R}^m$ zu verallgemeinern. Damals war wie in dieser Vorlesung die Idee hinter der Taylorformel, hinreichend oft differenzierbare Funktionen durch Polynome zu approximieren. Hierzu wollen wir die mehrdimensionale Situation auf die Eindimensionale zurückführen.

Bemerkung: Für das Folgende können wir ohne Einschränkung $m=1$ annehmen. Betrachten wir dann Abbildungen $f=(f_1,...,f_m)^\mathsf{T}\colon \mathbb{R}^d \to \mathbb{R}^m$ mit $m\in\mathbb{N}$, so wenden wir die nachfolgenden Theorie auf jede einzelne Komponente f_j, $j=1,...,m$, an.

Sei im Folgenden $U \subset \mathbb{R}^d$ offen, $k \in \mathbb{N}$ und $f \in C^{k+1}(U)$, wobei wir an die Schreibweise

$$C^j(U) := \{g\colon U \to \mathbb{R}\colon\ g \text{ ist } j\text{-mal stetig differenzierbar}\}$$

für $j \in \mathbb{N}$ erinnern. Seien weiters $x_0 \in U$ und $x \in U$ so, dass die Verbindungsstrecke $[x_0, x] := \{\lambda x + (1-\lambda)x_0\colon 0 \leq \lambda \leq 1\}$ in U enthalten ist. Wir setzen $h := x - x_0$ und ferner

$$\varphi\colon [0,1] \to \mathbb{R}, \quad \varphi(t) := f(x_0 + th).$$

Dann folgt mittels Kettenregel, dass

$$\varphi'(t) = \sum_{i_1=1}^{n} (\partial_{i_1} f)(x_0 + th) h_{i_1},$$

und induktiv folgt

$$\begin{aligned}\varphi^{(k)}(t) &= \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} (\partial_{i_1}\partial_{i_2}...\partial_{i_k} f)(x_0 + th) h_{i_1}...h_{i_k} \\ &= \sum_{i_1,\dots,i_k=1}^{n} (\partial_{i_1}\partial_{i_2}...\partial_{i_k} f)(x_0 + th) h_{i_1}...h_{i_k}.\end{aligned}$$

Aus der Analysis 1 ist die folgende Integraldarstellung des Restglieds bekannt: Ist $I \subset \mathbb{R}$ ein echtes Intervall und $g \in C^{k+1}(I)$ mit k-tem Taylorpolynom

$$P_k(t) = \sum_{j=0}^{k} \frac{g^{(j)}(t_0)}{j!}(t - t_0)^j$$

in t_0, so gilt

$$g(t) = P_k(t) + R_k(t) =: P_k(t) + \frac{1}{k!}\int_{t_0}^{t} (t-s)^k g^{(k+1)}(s) ds.$$

Wir wenden das nun auf φ von oben an. Es ergibt sich sodann mit $I = [0,1]$, $t = 1$ und $t_0 = 0$

$$\begin{aligned}f(x) = \varphi(1) &= P_k(1) + \frac{1}{k!}\int_0^1 (1-s)^k \varphi^{(k+1)}(s) ds \\ &= \sum_{j=0}^{k} \frac{1}{j!} \sum_{i_1,\dots,i_j=1}^{d} (\partial_{i_1}...\partial_{i_j} f)(x_0) h_{i_1}...h_{i_j} \\ &\quad + \frac{1}{k!}\int_0^1 (1-s)^k \sum_{i_1,\dots,i_{k+1}=1}^{d} (\partial_{i_1}...\partial_{i_{k+1}} f)(x_0 + sh) h_{i_1}...h_{i_{k+1}} ds =: (*).\end{aligned} \tag{7.2}$$

Diese Darstellung ist recht umständlich, und um eine bequeme abkürzende Schreibweise zu haben, führen wir *Multi-Indices* ein.

Multiindices. Sei $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n$. Wir setzen

$$\begin{aligned}
&|\alpha| := \alpha_1 + ... + \alpha_d, \\
&\alpha! := \alpha_1!...\alpha_d!, \\
&x^\alpha := x_1^{\alpha_1}...x_d^{\alpha_d} \qquad \text{für } x = (x_1, ..., x_d)^\mathsf{T} \in \mathbb{R}^d, \\
&\partial^\alpha := \partial_1^{\alpha_1}...\partial_d^{\alpha_d}.
\end{aligned}$$

Wir setzen auch noch $\partial^{(0,...,0)} := \mathrm{id}$. Fassen wir die in (7.2) auftretenden Summen über Multiindices zusammen, so treten nach dem Satz von Schwarz in (7.2) Ableitungen, die zu einem fixierten Multiindex α gehören, mehrfach auf.

Es sei daher $j \in \{0, ..., k+1\}$. Dann beträgt die Anzahl der Tupel $(i_1, ..., i_j)$ in der Summe

$$\sum_{i_1,...,i_j=1}^{d} (\partial_{i_1}...\partial_{i_j} f)(x_0) h_{i_1}...h_{i_j},$$

in denen nach der Koordinaten x_ν genau α_ν-mal differenziert wird,

$$\binom{j}{\alpha_1}\binom{j-\alpha_1}{\alpha_2}...\binom{j-\alpha_1-...-\alpha_{d-1}}{\alpha_d} = \frac{j!}{\alpha!}.$$

Damit folgt für $(*)$ (cf. (7.2)):

$$\begin{aligned}
f(x) = (*) = &\sum_{|\alpha|=0}^{k} \frac{1}{\alpha!} (\partial^\alpha f)(x_0) h^\alpha \\
&+ \frac{1}{k!} \int_0^1 (1-s)^k \sum_{|\alpha|=k+1} \frac{(k+1)!}{\alpha!} (\partial^\alpha f)(x_0 + sh) h^\alpha ds.
\end{aligned}$$

Damit haben wir den folgenden Satz über die Taylorapproximation bewiesen:

Satz 7.4. *Sei $U \subset \mathbb{R}^d$ offen, $x_0, x_0 + h \in U$ mit $[x_0, x_0 + h] \subset U$. Sei $f \in C^k(U; \mathbb{R}^m)$. Dann gilt*

$$\begin{aligned}
f(x_0 + h) = &\sum_{|\alpha|=0}^{k} \frac{1}{\alpha!} (\partial^\alpha f)(x_0) h^\alpha \\
&+ (k+1) \sum_{|\alpha|=k+1} \int_0^1 (1-s)^k \frac{1}{\alpha!} (\partial^\alpha f)(x_0 + sh) h^\alpha ds.
\end{aligned}$$

Im vorherigen Satz haben wir die integrale Form des Restglieds (der eindimensionalen Taylorapproximation) verwendet. Verwendet man stattdessen das Lagrangesche Restglied, so erhalten wir den nachfolgenden

Satz 7.5. *In der Situation des vorherigen Satzes existiert ein $\xi \in [x_0, x_0+h]$ mit*

$$f(x_0 + h) = \sum_{|\alpha|\leq k} \frac{\partial^\alpha f(x_0)}{\alpha!} h^\alpha + \sum_{|\alpha|=k+1} \frac{\partial^\alpha f(\xi)}{\alpha!} h^\alpha.$$

Wie in Analysis führen wir nun eine kanonische Sprechweise für die vorausgegangenen Sätze ein:

Definition 7.3. *Für $U \subset \mathbb{R}^d$ offen, $x_0 \in U$, $f \in C^k(U; \mathbb{R}^m)$ heißt*

$$T_k^{x_0} f(h) := \sum_{|\alpha|\leq k} \frac{\partial^\alpha f(x_0)}{\alpha!} h^\alpha$$

das Taylorpolynom *in h zu f vom Grad k in x_0. Ist $f \in C^\infty(U; \mathbb{R}^m) := \bigcap_k C^k(U; \mathbb{R}^m)$, so heißt*

$$T^{x_0} f(h) := \sum_{\alpha\in\mathbb{N}^n} \frac{\partial^\alpha f(x_0)}{\alpha!} h^\alpha$$

die Taylorreihe *in h zu f in x_0. Wir nennen f* reell-analytisch*, falls zu jedem $x_0 \in U$ ein $\delta > 0$ existiert sodass $T^a f(h)$ für alle $|h| < \delta$ summierbar ist und ihre Summe $f(x_0 + h)$ ist.*

Wir besprechen nun die Approximationsgüte der Taylorapproximation. Aus dem Satz über das Lagrangesche Restglied folgt sofort

Lemma. *In der Situation von Satz 7.4 gilt*

$$\lim_{h\to 0} \frac{|f(x_0 + h) - T_k^{x_0} f(h)|}{|h|^k} = 0.$$

Weiters ist in der Situation von Satz 7.4 $T_k^{x_0} f$ das eindeutig bestimmte Polynom P vom Grad kleiner gleich k mit $\partial^\alpha P(x_0) = \partial^\alpha f(x_0)$ für alle Multiindices α mit $|\alpha| \leq k$.

Für praktische Berechnungen lässt sich das Taylorpolynom der Ordnung zwei kompakt schreiben als

$$T_2^{x_0} f(h) = f(x_0) + \langle \nabla f(x_0), h\rangle + \frac{1}{2}\langle Hf(x_0)h, h\rangle,$$

wobei $\langle\cdot,\cdot\rangle$ das Euklidische Skalarprodukt in $\mathbb{R}^d$ bezeichne. Wir betrachten dahingehend nun ein

Beispiel 7.2. Sei $F\colon \mathbb{R}^2 \ni (x,y)^\mathsf{T} \mapsto x^y \in \mathbb{R}$. Dann ist F in $(0,\infty)^2$ differenzierbar. Es gilt sodann

$$\begin{aligned} T_2^{(1,1)^\mathsf{T}} F(h) &= F(1,1) + \langle \nabla F(1,1), h\rangle + \frac{1}{2}\langle HF(x_0)h, h, \rangle \\ &= 1 + h_1 + h_1 h_2, \end{aligned}$$

wobei $h = (h_1, h_2)^\mathsf{T}$, da

$$\nabla F(1,1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad Hf(1,1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

8 Der Satz über implizite Funktionen

8.1 Diffeomorphismen und Umkehrfunktionen

Sei $A \in M_{d\times d}(\mathbb{R})$. Die lineare Algebra behandelt die Lösbarkeit von Gleichungen $Ax = b$, wobei $b \in \mathbb{R}^d$ gegeben ist. Wir wissen, dass diese Gleichung für jedes $b \in \mathbb{R}^d$ genau dann lösbar ist, wenn $A \in M_{d\times d}(\mathbb{R})$ invertierbar ist. Sei nun $F\colon \mathbb{R}^d \ni x \mapsto F(x) := Ax \in \mathbb{R}^d$. Dann übersetzt sich dieses Problem in $F(x) = b$. Es ist dann natürlich zu fragen, wie die Lösbarkeit solcher Gleichungen behandelt werden kann, wenn F nichtlinear ist.

Ein Schlüsselkonzept ist dahingehend die *Linearisierung*; wir führen nichtlineare Problem auf lineare zurück.

Bevor wir dies genauer behandeln, formulieren wir zuerst die Fragen von Interesse genauer. Sei hierzu $D \subset \mathbb{R}^d$ offen, $F \in C^1(D; \mathbb{R}^d)$ und $y \in \mathbb{R}^d$.

1. Für welche y ist $F(x) = y$ in lösbar in U?
2. Wann ist die Lösung eindeutig?
3. Hängen die Lösungen stetig von den Daten y ab?

Für die Beantwortung dieser Fragen führen wir zuerst eine neue Sprechweise ein.

Definition 8.1 (C^k-Diffeomorphismen)**.** *Seien $U, D \subset \mathbb{R}^d$ offen und $k \in \mathbb{N} \cup \{\infty\}$. Eine bijektive C^k-Abbildung $\Psi\colon D \to U$ heißt C^k-Diffeomorphismus, falls auch $\Psi^{-1}\colon U \to D$ eine C^k-Abbildung ist.*

Weiters heißt Ψ ein lokaler C^k-Diffeomorphismus, falls für alle $x \in D$ eine Umgebung $V \subset D$ von x und eine Umgebung $\widetilde{V}$ von $f(x)$ existieren, sodass $f|_V\colon V \to \widetilde{V}$ ein C^k-Diffeomorphismus ist.

Der zentrale Satz dieses Abschnitts lautet nun wie folgt:

Satz 8.1 (Lokaler Umkehrsatz)**.** *Es seien $D \subset \mathbb{R}^d$ offen, $1 \leq k \leq \infty$. Eine Abbildung $f \in C^k(U; \mathbb{R}^d)$ ist genau dann ein lokaler C^k-Diffeomorphismus, wenn für alle $x \in D$ die totale Ableitung $Df(x)$ invertierbar ist.*

In der Praxis prüfen wir also für die Anwendbarkeit des Satzes nach, ob $\det(Df(x)) \neq 0$ gilt.

Man beachte, dass der Satz nichts über die globale Umkehrbarkeit einer Abbildung aussagt:

Beispiel 8.1. Wir betrachten die komplexe Exponentialfunktion in reellen Koordinaten, also

$$f \colon (x, y)^\mathsf{T} \mapsto \begin{pmatrix} e^x \cos(y) \\ e^x \sin(y) \end{pmatrix}.$$

Dann gilt für alle $x, y \in \mathbb{R}$

$$\det\left(Df(x, y)\right) = e^{-2x} \neq 0,$$

aber man sieht leicht, dass f nicht global injektiv ist.

[28.05.2018]

[04.06.2018]

8.2 Der Satz über implizite Funktionenen

Beispiele: $F(x) = |x|^2 - 1$, Matrizen, $f \in C^\infty(\mathbb{R}^2)$ definiert durch

$$f(x) = x_2^5 + (1 + x_1^2)x_2 + x^4 - x^3$$

Satz 8.2. *Sei $k \in \mathbb{N} \cup \infty$, $d, m \in \mathbb{N}$, $1 \leq d, k, m$, $U \subset \mathbb{R}^d \times \mathbb{R}^m = \mathbb{R}^{d+m}$ offen. Wir schreiben $\begin{pmatrix} x \\ y \end{pmatrix}$ für Punkte in $\mathbb{R}^d \times \mathbb{R}^m$. Sei $F \in C^k(U; \mathbb{R}^m)$. Wir schreiben*

$$DF \begin{pmatrix} x \\ y \end{pmatrix} = (D_x F \begin{pmatrix} x \\ y \end{pmatrix}, D_y F \begin{pmatrix} x \\ y \end{pmatrix}).$$

Sei $DF \begin{pmatrix} a \\ b \end{pmatrix} \in U$ so dass

1. *$F \begin{pmatrix} a \\ b \end{pmatrix} = 0$*

2. *$D_y F \begin{pmatrix} a \\ b \end{pmatrix} = 0$ ist invertierbar.*

Dann existieren $\rho > 0$, $r > 0$ und $f \in C^k(B_\rho(a);\mathbb{R}^m$ so dass

$$B_\rho(a) \times B_r(b) \subset U,$$

$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x \in B_\rho(a), y \in B_r(b), F\begin{pmatrix} x \\ y \end{pmatrix} = 0 \right\} = \left\{ \begin{pmatrix} x \\ f(x) \end{pmatrix} : x \in B_\rho(x) \right\}$$

und

$$Df(x) = -\left(D_y F \begin{pmatrix} x \\ f(x) \end{pmatrix} \right)^{-1} D_x F \begin{pmatrix} x \\ f(x) \end{pmatrix}. \tag{8.1}$$

Beweis. Wir beginnen mit Vorbemerkungen und Vereinfachungen. Sei $U \in \mathbb{R}^d$, $G \in C^1(U)$, und $x, y \in U$. Wir nehmen an, dass auch die Verbindungsstrecke in U liegt. Wir definieren für $t \in [0,1]$

$$g(t) = G(x + t(y-x)).$$

Nach dem Hauptsatz gilt

$$\begin{aligned} G(y) - G(x) =& g(1) - g(0) \\ =& \int_0^1 g'(t)dt \\ =& \int_0^1 DG(x+t(y-x))(y-x)dt \\ =& \int_0^1 DG(x+t(y-x))dt(y-x) \end{aligned}$$

wobei die dritte Identität mit der Kettenregel folgt, und die letzt aufgrund der Linearität des Integrals, wobei das letzte Integral ein Integral über eine Matrix ist, und wie immer als die Matrix mit den Integralen der Komponenten als Einträge ist.

Dieses Argument funktioniert genauso mit Integralen mit Integranden, die Werte in $\mathbb{R}^d$ annehmen. Sei also $G \in C^1(U;\mathbb{R}^m)$. Dann gilt

$$\begin{aligned} |G(y) - G(x)| =& \left| \int_0^1 DG(x+t(y-x))(y-x)dt \right| \\ \leq & \int_0^1 |DG(x+t(y-x))(y-x)|dt \\ \leq & \int_0^1 \|DG(x+t(y-x))\|_{\mathbb{R}^d \to \mathbb{R}^m} dgt|y-x|. \end{aligned}$$

Diese Art von Abschätzung werden wir im Folgenden wiederholt verwenden.

Wir definieren

$$\tilde{F}\begin{pmatrix} x \\ y \end{pmatrix} = D_y \begin{pmatrix} a \\ b \end{pmatrix}^{-1} F \begin{pmatrix} a+x \\ b+y \end{pmatrix}.$$

Damit ist $\tilde{F}\begin{pmatrix}0\\0\end{pmatrix} = 0$ und $D_y\begin{pmatrix}0\\0\end{pmatrix} = 1_{\mathbb{R}^m}$. $\tilde{F}$ ist auf einer Menge $\tilde{U}$ definiert, die aus U durch Verschieben um $\begin{pmatrix}a\\b\end{pmatrix}$ erhalten wird.

Schritt 1: Wir arbeiten mit $\tilde{F}$ und lassen in diesem und dem nächsten Schritt die Tilde weg, betrachten also F für $a = 0$, $b = 0$ und $D_yF\begin{pmatrix}0\\0\end{pmatrix} = 1_{\mathbb{R}^m}$.
Da die Ableitung stetig ist, existiert $r > 0$, so dass für $|x| < r$ and $|y| \le r$

$$\|D_yF\begin{pmatrix}x\\y\end{pmatrix} - 1_{\mathbb{R}^m}\|_{\mathbb{R}^m\to\mathbb{R}^m} \le \frac{1}{2}.$$

Außerdem existiert $r \ge \rho > 0$ so dass für $|x| < \rho$

$$|F(x,0)| < r/2.$$

Sei jetzt $|x| < \rho$. Wir definieren $\phi : \overline{B_r^{\mathbb{R}^m}(0)} \to \mathbb{R}^m$ durch

$$\phi(y) = y - F\begin{pmatrix}x\\y\end{pmatrix}.$$

Es gilt

$$\begin{aligned}\phi(y_2) - \phi(y_1) =& y_2 - y_1 - \left(F\begin{pmatrix}x\\y_2\end{pmatrix} - F\begin{pmatrix}x\\y_2\end{pmatrix}\right)\\ =& \int_0^1 \left(1_{\mathbb{R}^m} - D_yF\begin{pmatrix}x\\y_1 + t(y_2 - y_1)\end{pmatrix}\right)(y_2 - y_1)dt\end{aligned}$$

und mit den Abschätzungen aus dem nullten Schritt

$$\begin{aligned}|\phi(y_2) - \phi(y_1)| \le& \sup\left\|1_{\mathbb{R}^m} - DF\begin{pmatrix}x\\y_1 + t(y_2 - y_1)\end{pmatrix}\right\|_{\mathbb{R}^m\to\mathbb{R}^m}|y_2 - y_1|\\ \le& \frac{1}{2}|y_2 - y_1|.\end{aligned}$$

Da

$$\phi(y) = \phi(y) - \phi(0) + \phi(0) = \phi(y) - \phi(0) - F\begin{pmatrix}x\\0\end{pmatrix}$$

folgt

$$|\phi(y)| < \frac{1}{2}|y| + \frac{r}{2} \le r$$

und ϕ bildet $\overline{B_r^{\mathbb{R}^m}(0)}$ auf sich selbst ab. Außerdem ist ϕ eine Kontraktion mit Kontraktionskonstanten $\frac{1}{2}$. Nach dem Fixpunktsatz von Banach existiert genau ein y in dem Ball mit $y = \phi(y)$. Das ist äquivalent zu $F\begin{pmatrix}x\\y\end{pmatrix} = 0$.
WIr definieren die Abbildung $f(x) := y$.

Schritt 2: f ist Lipschitz stetig. Wenn wir r und δ notfalls kleiner wählen, so dürfen wir annehmen, dass

$$\left\| D_x F \begin{pmatrix} x \\ y \end{pmatrix} - D_x F \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\|_{\mathbb{R}^d \to \mathbb{R}^m} \leq 1.$$

Sei $x_1, x_2 \in B_\rho^{\mathbb{R}^d}(0)$ und $y_1 = f(x_1)$, $y_2 = f(x_2)$. Aufgrund der Definition gilt

$$\begin{aligned} 0 = & F \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} - F \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \\ = & F \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} - F \begin{pmatrix} x_1 \\ y_2 \end{pmatrix} + F \begin{pmatrix} x_1 \\ y_2 \end{pmatrix} - F \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \end{aligned}$$

Wie oben erhalten wir mit $\|D_x F \begin{pmatrix} 0 \\ 0 \end{pmatrix} \|_{\mathbb{R}^d \to \mathbb{R}^m} =: C$

$$\left| F \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} - F \begin{pmatrix} x_1 \\ y_2 \end{pmatrix} \right| \leq (C+1)|x_2 - x_1|$$

und

$$F \begin{pmatrix} x_1 \\ y_2 \end{pmatrix} - F \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = y_2 - y_1 - \int_0^1 (1_{\mathbb{R}^m} - D_y F \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} (y_2 - y_1) dt$$

und daher

$$|y_2 - y_1| \leq \left| F \begin{pmatrix} x_1 \\ y_2 \end{pmatrix} - F \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \right| + \frac{1}{2}|y_2 - y_1|$$

und zusammen erhalten wir die Lipschitzstetigkeit

$$|y_2 - y_1| \leq 2(C+1)|x_2 - x_1|.$$

Schritt 3: f ist in $x = 0$ total differenzierbar. Sei $\varepsilon > 0$. Wir wählen $r > 0$ so klein dass

$$\left\| DF \begin{pmatrix} x \\ y \end{pmatrix} - DF \begin{pmatrix} x \\ y \end{pmatrix} \right\|_{\mathbb{R}^{d+m} \to \mathbb{R}^m} \leq \frac{\varepsilon}{2(1+C)}.$$

Es folgt

$$y - 0 + D_x F \begin{pmatrix} 0 \\ 0 \end{pmatrix} x = \int_0^1 (1_{\mathbb{R}^m} - D_y F \begin{pmatrix} tx \\ ty \end{pmatrix} y + \left[D_x F \begin{pmatrix} 0 \\ 0 \end{pmatrix} - D_x F \begin{pmatrix} x \\ y \end{pmatrix} \right] x \, dt$$

und daher

$$\left| y + D_x F \begin{pmatrix} 0 \\ 0 \end{pmatrix} x \right| < \frac{\varepsilon}{C+2}(|x| + |y|) \leq \varepsilon |x|.$$

Damit ist

$$Df(0) = -D_x F \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

[04.06.2018]

[07.06.2018]

Schritt 4: Rückkehr zu F wie im Satz. Wir versehen die Abbildungen in den ersten drei Schritten wieder mit einer Tilde. Dann ist

$$f(x) = \tilde{f}(x-a) + b$$

und

$$Df(a) = -\left(D_y F \begin{pmatrix} a \\ b \end{pmatrix} \right)^{-1} D_x F \begin{pmatrix} a \\ b \end{pmatrix}.$$

Wir können dieses Argument in für jedes $x \in B_\rho(a)$ anwenden und erhalten die Formel für die Ableitung. Die rechte Seite der Formel wird aus stetigen Funktionen gebildet, und damit folgt $f \in C^1(B_\rho(a); \mathbb{R}^m =$.

Sei nun $k \geq 2$. Wie bei der Stetigkeit sehen wir, dass die partiellen Ableitungen von f total differenzierbar sind. Wir berechnen die zweiten Ableitung (und verwenden, dass die totale Ableitung der Inversen an der invertieren Matrix A die lineare Abbildung $H \to -A^{-1} H A^{-1}$ ist:

$$\frac{\partial}{\partial x_j} \frac{\partial}{\partial x_k} f(x) = (D_y F)^{-1} \partial_{x_j} (D_y F)(D_y F)^{-1} \partial_{x_k} F - (D_y F)^{-1} \partial_{x_j} \partial_{x_k} F.$$

Rekursiv erhalten wir $f \in C^k(B_\rho(a); \mathbb{R}^m)$, da sich die Struktur der Formel für die Ableitungen nicht ändert.

□

Der Satz von der Umkehrfunktion ist eine direkte Konsequenz aus dem Satz über implizite Funktionen.

Beweis des Satzes von der Umkehrfunktion: Es seien $U \subset \mathbb{R}^d$ offen, $g \in C^k(U; \mathbb{R}^d)$, $a \in U$ und $Df(a)$ invertierbar. Wir definieren

$$F \in C^k(\mathbb{R}^d \times U; \mathbb{R}^d)$$

mit $m = d$ durch

$$F(x, y) = x - f(y).$$

Dann ist $D_y F = -Df$, $D_x F = 1_{\mathbb{R}^d}$ und wir können den Satz über implizite Funktionen anwenden. Wir erhalten eine Abbildung $g \in C^k(B_\rho(f(a)); \mathbb{R}^d)$ mit

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

und

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

□

9 Mannigfaltigkeiten des $\mathbb{R}^d$

Definition 9.1. *Seien $k \in \mathbb{N} \cap \infty$, $d, p \in \mathbb{N}$, $d, k, p \geq 1$ und $p \leq d$. Eine Menge $S \subset \mathbb{R}^d$ heißt p dimensionale C^k Mannigfaltigkeit, falls zu jedem $a \in S$ eine offene Umgebung $V \subset U$, eine offene Menge $W \subset \mathbb{R}^d$ und ein C^k Diffeomorphismus $\Phi : V \to W$ existieren, so dass*

$$\Phi(S \cap V) = (\mathbb{R}^p \times \{0_{\mathbb{R}^{d-p}}) \cap W.$$

Satz 9.1. *Für eine Menge $S \subset \mathbb{R}^d$ sind äquivalent:*

1. *S ist eine p dimensionale C^k Mannigfaltigkeit des $\mathbb{R}^d$.*

2. *Zu jedem $a \in S$ existiert eine offene Umgebung $V \subset \mathbb{R}^d$ und ein $f \in C^k(V; \mathbb{R}^{d-p})$ so dass der Rang von $Df(x)$ immer $d - p$ ist und*

$$S \cap V = \{x \in V : f(x) = 0\}$$

3. *Zu jedem $a \in S$ gibt es nach einer Permutation der Koordinaten offene Mengen $V_1 \subset \mathbb{R}^p$, $V_2 \subset \mathbb{R}^{d-p}$ und $g \in C^k(V_1, \mathbb{R}^{d-p}$ so dass $a \in V_1 \times V_2 \subset U$, $g : V_1 \to V_2$ und*

$$S \cap (V_1 \times V_2) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_2 = g(x_1), x_1 \in V_1 \right\}.$$

4. *Zu jedem $a \in S$ gibt es eine offene Umgebung $V \subset U$, eine offene Menge $W \subset \mathbb{R}^p$ und eine C^k Abbildung $\psi : W \to \mathbb{R}^d$ so dass der Rang von $D\psi$ immer p ist, $\psi(W) = S \cap V$ und $\psi : W \to S \cap V$ ist ein Homöomorphism, d.h. eine stetige bijektive Abbildung mit stetiger Umkehrabbildung.*

Es ist instruktiv, das vermutlich einfachste nichttriviale Beispiel zu betrachten mit $d = 2$, $p = 1$ und $k = \infty$:

$$S = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}.$$

Wir definieren die Funktion

$$f(x) = |x|^2 - 1.$$

Dann ist $Df(x) = 2x^T$ was nur dann Null ist, wenn $x = 0$, sonst ist der Rang 1. Also genügt S der Bedingung in (b).
Ist $x_2 > 0$ so gilt $x_2 = \sqrt{1 - x_1^2}$ und S ist der Graph dieser Funktion in einer Umgebung von $x \in S$. Die Modifikationen für $x_2 < 0$, $x_1 > 0$ und $x_1 < 0$ sind offensichtlich. Wir können also S als Graph schreiben, möglicherweise nach einer Permutation der Koordinaten.

Eine Parametrisierung wie in Teil d erhalten wir durch

$$\psi(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}.$$

[07.06.2018]

[11.06.2018]

Beweis. "(a) $\Longrightarrow$ (b)". Sei Φ wie in der Definition. Wir definieren

$$f(x) = (\phi_j(x))_{p<j\leq d}.$$

f hat offensichtlich die gewünschen Eigenschaften.

"(b) $\Longrightarrow$ (c)" Sei $a \in S$ und f wie in (b). Dann ist der Rang von Df $d-p$, und nach Permutation der Koordinaten ist $(\partial_j f_k)_{p+1\leq j\leq d, 1\leq k\leq d-p=:m}$ invertierbar. Nach dem Satz über implizite Funktionen können wir nach den letzten Koordinaten auflösen.

"(c) $\Longrightarrow$ (d)" Nach einer Permutation der Koordinaten definieren wir

$$\psi(u) = \begin{pmatrix} u \\ g(u) \end{pmatrix}.$$

"(d) $\Longrightarrow$ (a)" Sei $a = \Psi(b)$. Die Ableitung $D\Psi$ hat den Rang p. Nach einer Permutation der Koordinaten dürfen wir annehmen, dass die ersten p Zeilen linear unabhängig sind. Wir definieren

$$\Phi : V \times \mathbb{R}^{d-p} \to \mathbb{R}^d$$

$$\Phi \begin{pmatrix} u \\ s \end{pmatrix} = \Psi(u) + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ s_1 \\ \vdots \\ s_{d-p} \end{pmatrix}.$$

Dann ist

$$D\Psi \begin{pmatrix} b \\ y \end{pmatrix} = \begin{pmatrix} D\Psi_1(b) & 0 \\ D\Psi_2(b) & 1 \end{pmatrix}$$

wobei die Zerlegung von $D\Psi$ der des $\mathbb{R}^d = \mathbb{R}^p \times \mathbb{R}^{d-p}$ entspricht. Diese lineare Abbildung ist invertierbar, das $D_1\Psi(b)$ nach Konstruktion und Annahme invertierbar ist. Nach dem Satz über die Umkehrfunktion ist Ψ in einer Umgebung ein Diffeomorphismus, der die gewünschten Eigenschaften hat.

$\square$

Beispiel 9.1. Wir betrachten die Sphäre

$$S^{d-1} := \partial B_1(0) \subset \mathbb{R}^d.$$

Sei

$$f(x) = |x|^2 - 1.$$

Wir oben sieht man, dass S die Nullstellenmenge von f ist, und dass der Rang der Ableitung in S 1 ist. Damit ist S eine Mannigfaltigkeit.

Mit Polarkoordinaten bekommen wir Abbildung wie in der Definition der Mannigfaltigkeit. Im Fall $d = 2$ definieren wir

$$\Phi \begin{pmatrix} r \\ \phi \end{pmatrix} = \begin{pmatrix} r\cos\phi \\ r\sin\phi \end{pmatrix}$$

für $r > 0$ und $\phi \in \mathbb{R}$. Wir berechnen die Jacobimatrix

$$D\Phi \begin{pmatrix} r \\ \phi \end{pmatrix} = \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix}$$

und $\det D\Phi = r \neq 0$. Daher ist Φ ein lokaler Diffeomorphismus, der die Menge $\{r = 1\}$ auf S^1 abbildet, und wir bekommen aus der Umkehrabbildung durch Subtraktion der Konstanten eine Abbildung wie in der Definition. Wenn wir ϕ auf ein Intervall der Länge $< 2\pi$ einschränken, so erhalten wir einen Diffeomorphismus auf das Bild.

Im Fall $d = 3$ definieren wir

$$\Psi \begin{pmatrix} \theta \\ \phi \end{pmatrix} = \begin{pmatrix} r\cos\phi\sin\theta \\ r\sin\phi\sin\theta \\ r\cos\theta \end{pmatrix}$$

für $0 < \phi < 2\pi$, $0 < \theta < \pi$. Wir berechnen

$$D\Psi \begin{pmatrix} r \\ \theta \\ \phi \end{pmatrix} = \left(\frac{\partial \Psi}{\partial(r,\theta,\phi)} \right) = \begin{pmatrix} \cos\phi\sin\theta & r\cos\phi\cos\theta & -r\sin\phi\sin\theta \\ \sin\phi\sin\theta & r\sin\phi\cos\theta & r\cos\phi\sin\theta \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}.$$

Dann ist

$$\det D\Psi \begin{pmatrix} r \\ \theta \\ \phi \end{pmatrix} = r^2 \sin\theta \neq 0$$

$$\Phi : (0,\infty) \times (0,\pi) \times (0,\pi) \to \mathbb{R}^3 \backslash \{x : x_2 = 0, x_1 \geq 0\}$$

ist ein Diffeomorphismus, und die Umkehrabbildung minus e_1 ist eine Abbildung wie in der Definition der Mannigfaltigkeit für S^2.

Analog können wir S^d durch

$$\psi(\theta_1, \theta_2, \dots, \theta_{d-2}, \phi) = \begin{pmatrix} \cos\phi \sin\theta_1 \dots \sin\theta_{n-2} \\ \sin\phi \sin\theta_1 \dots \sin\theta_{n-2} \\ \cos\theta_1 \dots \sin\theta_{n-2} \\ \vdots \\ \cos\theta_{n-2} \end{pmatrix}$$

parametrisieren.

Definition 9.2. *Sei $f \in C^k(U; \mathbb{R}^m)$. $y \in \mathbb{R}^m$ heißt regulärer Wert, wenn $Df(x)$ den Rang m hat für alle x mit $f(x) = y$. Die Menge*

$$S = \{x : f(x) = y\}$$

heißt dann reguläre Niveaumenge.

Ist y ein regulärer Wert, so genügt in jedem Punkt von S $f - y$ den Anforderungen von Teil b) von Satz 9.1 und daher ist eine reguläre Niveaumenge eine Mannigfaltigkeit.

Beispiel 9.2. Wir betrachten Quadriken. Sei A eine symmetrische $d \times d$ Matrix und

$$f(x) = x^T A x - 1.$$

Die Menge

$$S = \{x : f(x) = 0\}$$

ist eine Quadrik. Da

$$Df(x) = 2x^T A$$

impliziert $Df(x) = 0$ auch $f(x) = 0$ und daher ist die Quadrik eine C^∞ Mannigfaltigkeit. Konkrete Beispiele

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Hyperbel

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Ellipse

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Zweischaliges Hyperboloid

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Einschaliges Hyperboloid.

Definition 9.3. *$S \subset \mathbb{R}^d$ sei eine p dimensionale C^k Mannigfaltigkeit und $a \in S$. Der Tangentialraum in a besteht aus allen Vektoren v des $\mathbb{R}^d$, für die ein differenzierbarer Weg in S existiert mit $\gamma(0) = a$ und $\gamma'(a) = 0$.*

Satz 9.2. *Sei S eine Mannigfaltigkeit, $a \in S$ und f wie oben. Dann ist v genau dann im Tangentialraum, wenn $Df(a)(v) = 0$, d.h. wenn v im Kern von $Df(a)$ liegt.*

Beweis. Sei γ ein Weg in S mit $\gamma(0) = a$. Da $f \circ \gamma(t) = 0$ ist für alle t folgt $Df(a)\gamma'(0) = 0$.
Umgekehrt: Sei $a \in S$, und v im Kern von $Df(a)$. Sei Φ wie in der Definition der Mannigfaltigkeit, $v' = D\Phi(a)(v)$. Aufgrund der Definition gilt

$$v' \in \mathbb{R}^p \times \{0\}.$$

Wir definieren γ' als den affinen Weg $\gamma(t) = \Phi(a) + tv'$ und

$$\gamma(t) = \Phi^{-1}(\Phi(a) + tv').$$

Dann ist γ ein Weg in S, $\gamma(0) = a$ und

$$\gamma'(0) = (D\Phi(a))^{-1}v' = v.$$

Der Beweis ist vollständig. □

[11.06.2018]

[14.06.2018]

Satz 9.3. *Sei S eine Mannigfaltigkeit, $a \in S$, T_aS der Tangentialraum von S, und $N_a(S)$ der zu $T_a(S)$ orthogonale Teilraum. Dann existieren offene Mengen V_T und V_N in $T_a(S)$ und $N_a(S)$ und $g \in \mathbb{C}^k(V_T; V_N)$ sodass $g(V_T) \subset V_N)$ und*

$$V \cap (V_1 \times V_2) = \{u + g(u) : u \in V_T\}$$

Definition 9.4. *$N_a(S)$ heißt Normalenraum von S in a. Ist f eine Abbildung wie in Satz 9.1, Teil b, so sind die Transponierten der Zeilen von Df eine Basis.*

Beweis. Im Fall $T_aS = \mathbb{R}^p \times \{0\}$ folgt dies aus Satz 9.1 bzw dessen Beweis. Im allgemeinen Fallen wählen wir geeignete Koordinaten, um uns auf den ersten Fall zurückzuziehen: Wir wählen eine ONB in T_aS und in N_aS, die zusammen eine ONB von $\mathbb{R}^d$ bilden. Sei O die aus diesen Vektoren bestehende orthogonale Matrix. Wir betrachten dann $x \to f(Ox)$ und

$$O^T S = \{O^T x : x \in S.\}$$

□

Beispiel 9.3. Die orthogonale Gruppe $O(\mathbb{R}^d)$. Wir definieren

$$O(\mathbb{R}^d) = \{A : A^T A = 1_{\mathbb{R}^d} = AA^T\}.$$

Diese Bedingung ist äquivalent zu

$$\langle x, y\rangle = \langle Ax, Ay\rangle$$

für alle $x, y \in \mathbb{R}^d$ oder sogar zu

$$|Ax| = |x|$$

für alle $x \in \mathbb{R}^d$: Für die einfache Richtung nehmen wir $A \in O(d)$ an. Dann folgt

$$|Ax|^2 = \langle Ax, Ax\rangle = \langle x, A^T Ax\rangle = \langle x, x\rangle = |x|^2.$$

Es gelte nun $|Ax| = |x|$ für alle $x \in \mathbb{R}^d$. Es folgt

$$\begin{aligned}\langle Ax, Ay\rangle =& \frac{1}{4}\left(\langle A(x+y), A(x+y)\rangle - \langle A(x-y), A(x-y)\rangle\right)\\ =& \frac{1}{4}\left(|A(x+y)|^2 - |A(x-y)|^2\right)\\ =& \frac{1}{4}\left(|x+y|^2 - |x-y|^2\right)\\ =& \langle x, y\rangle\end{aligned}$$

für $x, y \in \mathbb{R}^d$. Es folgt

$$\langle A^T Ax, y\rangle = \langle x, y\rangle.$$

Einsetzen der Vektoren der Standardbasis zeigt dass $A^T A = 1_{\mathbb{R}^d}$.
Wir erinnern an die Definition $S(\mathbb{R}^d)$ als die symmetrischen $d \times d$ Matrizen. Die Abbildung $F : M_{d\times d} \to S(\mathbb{R}^d)$

$$f(A) = A^T A - 1_{\mathbb{R}^d}$$

hat die Ableitung

$$Df(A)(H) = H^T A + (H^T A)^T$$

Sei jetzt $B \in S(\mathbb{R}^d)$ und $A \in O(\mathbb{R}^d)$. Wir definieren

$$H = \frac{1}{2}(AB)$$

und überprüfen

$$H^T A = \frac{1}{2}B^T = \frac{1}{2}B$$

Damit ist Df surjektiv für $A \in O(\mathbb{R}^d)$ und hat damit den maximalen Rang. Damit ist $O(\mathbb{R}^d)$ eine Mannigfaltigkeit. Der Tangentialraum in 1 besteht aus dem Kern von $df(1)$, also aus den Matrizen, die

$$A^T + A = 0$$

genügen. Diese Matrizen nennen wir schiefsymmetrisch und bezeichnen sie mit $A(\mathbb{R}^d)$. Die Dimension von $S(\mathbb{R}^d)$ ist die Dimension der oberen Dreiecksmatrizen, also $\frac{d}{d+1}2$. Die Dimension von $A(\mathbb{R}^d)$ ist die Dimension des Raumes der strikten oberen Dreiecksmatrizen, also $\frac{(d-1)d}{2}$.
Wir erarbeiten einen Diffeomorphismus in der Nähe der 1 wie in der Definition der Mannigfaltigkeit.

Lemma. Ist A eine schiefsymmetrische $d \times d$ Matrix, so ist

$$\exp(A) \in O(\mathbb{R}^d).$$

Beweis. Da $\exp(A^T) = (\exp(A))^T$ ist

$$\begin{aligned} \exp(A)(\exp(A))^T &= \exp(A)\exp(A^T) \\ &= \exp(A)\exp(-A) \\ &= \exp(A - A) = \exp(0) = 1_{\mathbb{R}^d}. \end{aligned}$$

□

Lemma. Die Abbildung $M_{d\times d} \ni A \to \exp(A) \in M_{d\times d}$ ist C^∞. Die Ableitung in 1 ist die Identität.

Beweis. Wir zeigen nur, dass $A \to \exp(A)$ stetig differenzierbar ist. Mittels vollständiger Induktion folgt $\exp \in C^\infty(M_{d\times d}, M_{d\times d})$. Da

$$(A+H)^j - A^j = HA^{j-1} + (A+H)HA^{j-2}\dot{+}(A+H)^jH$$

ist $A \to f_j(A) = A^j$ differenzierbar und

$$Df_j(A)(H) = HA^j + \cdots + A^jH.$$

Außerdem gilt

$$\begin{aligned} (A+H)^j - (A^j + Df_j(A)H) =& (A+H-A)HA^{j-2} \\ &+ ((A+H)^2 - A^2)HA^{j-2} \\ &+ \dots ((A+H)^{j-2} - A^{j-2})H \end{aligned}$$

und, da die Norm des Produktes von Matrizen nicht größer als das Produkt der Normen ist,

$$\|Df_j(A)(H)\|_{\mathbb{R}^d\to\mathbb{R}^d} \le (j+1)\|A\|^j_{\mathbb{R}^d\to\mathbb{R}^d}\|H\|_{\mathbb{R}^d\to\mathbb{R}^d}$$

sowie

$$\|(A+H)^{j+2} - A^{j+2} - Df_j(A)(H)\|_{\mathbb{R}^d\to\mathbb{R}^d} \le (j+2)(j+1)\|A\|^j_{\mathbb{R}^d\to\mathbb{R}^d}\|H\|^2_{\mathbb{R}^d\to\mathbb{R}^d}$$

Also ist

$$\left\| \sum_{j=1}^{\infty} \frac{1}{j!} Df_j(A)(H) \right\|_{\mathbb{R}^d \to \mathbb{R}^d} \leq \sum_{j=0}^{\infty} \frac{\|A\|^j_{\mathbb{R}^d \to \mathbb{R}^d}}{j!} \|H\|_{\mathbb{R}^d \to \mathbb{R}^d}$$
$$= \exp(\|A\|_{\mathbb{R}^d \to \mathbb{R}^d}) \|H\|_{\mathbb{R}^d \to \mathbb{R}^d}$$

und

$$\left\| \exp(A+H) - \exp(A) - \sum_{j=1}^{\infty} \frac{1}{j!} Df_j(A)(H) \right\|_{\mathbb{R}^d \to \mathbb{R}^d} \leq$$
$$\exp(\|A\|_{\mathbb{R}^d \to \mathbb{R}^d} + \|H\|_{\mathbb{R}^d \to \mathbb{R}^d}) \|H\|^2_{\mathbb{R}^d \to \mathbb{R}^d}.$$

Damit folgt

$$D\exp(A)(H) = \sum_{j=1}^{\infty} \frac{1}{j!} Df_j(A)(H).$$

Wir wählen die Standardbasis E_{ij} der Matrizen mit einem Eintrag 1 und 0 sonst. Dann ist die partielle Ableitung

$$\frac{\partial \exp(A)}{\partial E_{ik}} = \sum_{j=1}^{\infty} \frac{1}{j!} Df_j(A) E_{ik}.$$

Die Argumentation für höhere Ableitungen erfolgt genauso. □

Da die totale Ableitung von exp in 0 die Identität (und damit invertierbar) ist, definiert exp nach dem Satz von der Umkehrfunktion einen C^∞ Diffeomorphismus einer Umgebung der Null in eine Umgebung der 1.

$A(R^d)$ ist der zu $S(\mathbb{R}^d)$ orthonale Teilraum (im euklidschen Vektorraum mit Hilbert-Schmidt Norm) und umgekehrt. $S(\mathbb{R}^d) = T_1 O(\mathbb{R}^d)$. Also ist $O(\mathbb{R}^d)$ der Graph einer Abbildung g in einer Umgebung der Identität. Sei P die Projektion auf $A(d)$ entlang $S(d)$. Die Ableitung von

$$A(d) \ni A \to P \exp(A) \in A(d)$$

in $A = 0$ ist die Identität. Also ist die Abbildung nach dem Satz von der Umkehrabbildung ein Diffeomorphismus von geeigneten Umgebung. Damit erfüllt die Umkehrabbildung der Matrixexponentialabbildung die in der Definition geforderten Bediungungen in einer Umgebung der Identität, nach Wahl geigneter Koordinaten.

[14.06.2018]

[18.06.2018]

9.1 Extremwertaufgaben mit Nebenbedingungen

Satz 9.4. *Sei $U \subset \mathbb{R}^d$ offen, $S \subset U$ eine p dimensionale C^k Mannigfaltigkeit, $h \in C^1(U;\mathbb{R})$. Besitzt $f|_S$ ein lokales Extremum in $a \in S$, so ist $\nabla h(a) \subset N_a(S)$.*

Beweis. Zu zeigen ist, dass $\nabla h(a)$ senkrecht auf jedem Tangentialvektor steht. Sei $\gamma : (-\varepsilon, \varepsilon) \to S$ ein C^1 Weg mit $\gamma(0) = a$. Da $t \to h \circ \gamma$ ein lokales Extremum in $t = 0$ hat folgt

$$0 = \frac{d}{dt} h \circ \gamma(0) = Df(a)\gamma'(0).$$

Der Beweis ist komplett. □

Insbesondere existieren sogenannte Lagrangesche Multiplikatoren λ_j, so dass

$$Dh(a) = (\lambda_1, \dots, \lambda_{d-p}) D\tilde{f}(a)$$

oder, in einer äquivalenten Formulierung

$$\nabla h(a) = \sum_{j=1}^{d-p} \lambda_j \nabla f_{p+j}(a)$$

wobei wir die Funktion im Satz 9.1, Teil b hier mit f bezeichnen.

Beispiel 9.4. Wir das Maximum von

$$h(x) = 4x_1^2 - 3x_1x_2$$

auf $\overline{B_1(0)}$. Wir berechnen

$$Dh(x) = (8x_1 - 3x_2, -3x_1).$$

Wir suchen zunächst lokale Extrema im Inneren und dafür Nullstellen der totalen Ableitung. Aus der zweiten komponenten lesen wir die Bedingung $x_1 = 0$ ab, und dann aus der ersten $x_1 = 0$. Dort nimmt h den Wert Null an.

Auf S^1 suchen wir einen Lagrangeschen Multiplikator, d.h. Lösungen von

$$(8x_1 - 3x_2, -3x_1) = \lambda(x_1, x_2)$$

mit $x_1^2 + x_2^2 = 1$, da $S^1 = \{x : f(x) = \frac{1}{2}(|x|^2 - 1) = 0\}$ und $\nabla f(x) = x$, also

$$x_1 = -\frac{\lambda x_2}{3}$$

und damit

$$(\frac{8\lambda}{3} - 3)x_2 = -\lambda^2 x_2/3.$$

Ist $x_2 = 0$, so müsste auch $x_1 = 0$ sein, was $x_1^2 + x_2^2 = 1$ widerspricht. Sonst erhalten wir

$$0 = \lambda^2 + 8\lambda - 9 = (\lambda + 9)(\lambda - 1)$$

Wir erhalten zwei Punkte, die den Gleichungen

$$3x_1 = x_2 \qquad \text{und } x_1^2 + x_2^2 = 1$$

also

$$\begin{pmatrix} \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix}, \begin{pmatrix} -\frac{3}{\sqrt{10}} \\ -\frac{1}{\sqrt{10}} \end{pmatrix}$$

bwz

$$x_1 = -3x_2 \qquad \text{und } x_1^2 + x_2^2 = 1$$

und damit

$$\begin{pmatrix} \frac{1}{\sqrt{10}} \\ -\frac{3}{\sqrt{10}} \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{pmatrix}.$$

Nach dem Satz vom Maximum wird das Maximum angenommen. Da wir die notwendigen Bedingungen getestet haben, muss das Maximum an einem dieser fünf Punkte angenommen werden. Wir testen

$$h(0) = 0, h(\begin{pmatrix} \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix} = h\begin{pmatrix} -\frac{3}{\sqrt{10}} \\ -\frac{1}{\sqrt{10}} = \frac{27}{10} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{10}} \\ -\frac{3}{\sqrt{10}} \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{pmatrix} = \frac{13}{10}.$$

Also ist 2.7 das Maximum.

9.2 Die Schurzerlegung für symmetrische und selbstadjungierte Matrizen

Satz 9.5 (Schurzerlegung, Hauptachsentransformation)**.** *Sei V ein d dimensionaler Euklidscher reeller (komplexer) Vektorraum über $\mathbb{C}$ und $A : V \to V$ symmetrisch (selbstadjungiert). Dann existiert eine orthogonale (unitäre) Abbildung $O : \mathbb{R}^d \to V$ $(U : \mathbb{C}^d \to V)$ und eine Diagonalmatrix D mit reellen Einträgen so dass*

$$O^T AU = D \qquad (U^* AU = D).$$

Die Spalten von O (U) sind Eigenvektoren und die Einträge von D Eigenwerte.

Bemerkungen. Wir können die Gleichung auch in der Form

$$AO = OD$$

schreiben. Rechts steht eine Matrix, deren Spalten Vielfache der Spalten von U sind.

Wir nennen A symmetrisch (unitär) , falls

$$\langle v, Aw\rangle = \langle Av, w\rangle$$

und $O : \mathbb{R}^d \to V$ orthogonal (unitär), falls

$$\langle Ov, Ow\rangle_V = \langle v, w\rangle_{\mathbb{R}^d}$$

Dann definiert

$$\langle v, Ow\rangle_V = \langle O^T v, w\rangle$$

für $v \in V, w \in \mathbb{R}^d$ die transponierte Abbildung $O^T : V \to \mathbb{R}^d$. Analog definieren wir die adjungierte Abbildung.

Seien z^j Eigenvektoren zu $d^j \in \mathbb{R}$, $j = 1, 2$. Ist $d_1 \neq d_2$ so folgt

$$d_1\langle z^2, z^1\rangle = \langle z^2, Az^1\rangle = \langle Az^2, z^1\rangle = d_2\langle z^2, z^1\rangle$$

Daraus lesen wir ab, dass Eigenvektoren zu verschiedenen Eigenwerten orthogonal sind.

Wir können das Ergebnis auch anders interpretieren: Ist A symmetrisch (unitär), so existiert eine Basis aus Eigenvektoren, in der die lineare Abbildung durch eine Diagonalmatrix dargestellt wird.

Beweis. Nach einem Basiswechsel genügt es, die Aussage für $V = \mathbb{R}^d$ bzw $V = \mathbb{C}^d$ zu beweisen. Wir betrachten zunächst den reellen Fall und definieren $h \in C^\infty(\mathbb{R}^d)$ durch

$$h(x) = x^T Ax \in \mathbb{R}.$$

Nach dem Satz vom Maximum nimmt f das Maximum auf der kompakten Menge S^{d-1} in einem Punkt a mit $|a| = 1$ an. Die totale Ableitung von h und $f(x) = |x|^2 - 1$ läßt sich leicht bestimmen: $\nabla h(x) = 2Ax$ und $\nabla f(x) = 2x$. Nach dem Satz 9.4 existiert $\lambda \in \mathbb{R}$ mit

$$Aa = \lambda a$$

also ist a ein Eigenvektor und λ ein Eigenwert.

Wir beweisen die Aussage nun mittels vollständiger Induktion nach d. Der Fall $d = 1$ is trivial. Sei also $d > 1$ und die Aussage des Satzes wahr für $d - 1$. Sei v mit $|v| = 1$ ein Eigenvektor zum Eigenwert λ.

Sei $\langle < v, w\rangle = 0$. Es folgt

$$\langle v, Aw\rangle = \langle Av, w\rangle = \lambda\lambda v, w\rangle = 0$$

d.h., A definiert eine Abbildung des orthogonalen TRs. Sei

$$V' = \{w : \langle v, w\rangle = 0\}$$

und $A' : V' \to V'$ die durch die Einschränkung von A definierte Abbildung. Dann ist A' symmetrisch. Wir wenden die Induktionsannahme auf A' an und erhalten

$$(O')^T A' O' = D'$$

und die Spalten von O' sind Eigenvektoren von A. Wir ergänzen die orthogonale Abbildung $O' : \mathbb{R}^{d-1} \to V' \subset \mathbb{R}^d$ durch Hinzufügen der Spalte v zu einer orthogonalen $d \times d$ Matrix, und D' durch Hinzufügen des Diagonalelements λ zu einer $d \times d$ Diagonalmatrix. Wir erhalten damit die gewünschte Zerlegung.

Im komplexen Fall betrachten wir V als reellen Vektorraum der doppelten Dimension. Da jeweils mit z auch iz Eigenvektor ist, haben die Eigenräume jeweils die reelle Dimension 2. Der komplexe Fall folgt damit aus dem reellen Fall.

$\square$

[18.06.2018]

[21.06.2018]