Analysis II (Marciniak-Czochra)

Robin Heinemann

23. Mai 2017

Inhaltsverzeichnis

1 Metrische und normierte Räume **1**
1.1 Metrische Räume 1
1.2 Normierte Räume 3
1.3 Hilberträume 4

2 Stetigkeit und Differenzierbarkeit im $\mathbb{R}^n$ 7
2.1 Satz über implizite Funktionen und der Umkehrsatz 20
2.2 Extremalaufgaben mit Nebenbedingungen 23

3 Gewöhnliche Differentialgleichungen **25**

1 Metrische und normierte Räume

1.1 Metrische Räume

Definition 1.1 Sei M eine Menge, $d : M \times M \to [0, \infty)$ heißt **Metrik** auf M genau dann wenn $\forall x, y, z \in M$

- (D1) $d(x, y) = 0 \iff x = y$ (Definitheit)
- (D2) $d(x, y) = d(y, x)$ (Symmetrie)
- (D3) $d(x, z) \leq d(x, y) + d(z, y)$ (Dreiecksungleichung)

Beispiel 1.2 1. Charakterische (diskrete) Metrik

$$d(x, y) = \begin{cases} 0 & x = y \\ 1 & \text{sonst} \end{cases}$$

2. Sei $X = \mathbb{K}^n (\mathbb{K} = \mathbb{R}$ oder $\mathbb{C})$ mit Metrik

$$d(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^2 \right)^{\frac{n}{2}}$$

(euklidische Metrik)

3. Sei $X = \mathbb{R}^n$. Für $1 \leq \phi \leq \infty$. Sei

$$d_\phi(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^\phi\right)^{\frac{n}{\phi}}$$

Ist $\phi = \infty$, so definieren wir

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

4. $X = \mathbb{R}$ mit Metrik

$$d(x,y) = \frac{|x-y|}{1+|x-y|}$$

5. Der Raum der Folgen $a : \mathbb{N} \to \mathbb{R}$ (beziehungsweise $\mathbb{R}^\mathbb{N}$) kann mit der Metrik

$$d(x,y) = \sum_{k=0}^{\infty} 2^{-k} \frac{|x_k - y_k|}{1+|x_k - y_k|}$$

Definition 1.3 Sei M eine Menge mit Metrik d. Wir definieren für $x \in M, \varepsilon > 0$, die offene ε -Kugel um x durch

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

und eine abgeschlossene Kugel durch

$$K_\varepsilon(x) := \{y \in M \mid d(x,y) \leq \varepsilon\}$$

$A \subset M$ heißt **Umgebung** von $x \in M \iff \exists \varepsilon : K_\varepsilon(x) \subset A$

Konvergenz und Stetigkeit in metrischen Räumen

Definition 1.4 Eine Folge $(x_n)_{n\in\mathbb{N}}$ in einem metrischen Raum (X,d) ist konvergent gegen einem $x \in X$ genau dann wenn $\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} : \forall n \geq n_0 d(x_n, x) < \varepsilon$

Satz 1.5

1. Sei (X,d) ein metrischer Raum. Dann ist $A \subseteq X$ abgeschlossen genau dann wenn $(X_n)_{n\in\mathbb{N}}$ Folge in A mit $x_n \to x \implies x \in A$

2. Seien $(X,d_1), (Y,d_2)$ zwei metrische Räume. Dann ist die Funktion stetig in $x \in X$ genau dann wenn $(x_n)_{n\in\mathbb{N}}$ Folge in X mit $x_n \to x \implies f(x_n) \to f(x)$.

Definition 1.6 (Cauchy Folgen und Vollständigkeit) Sei (X,d) ein metrischer Raum. Eine Folge $(x_n)_{n\in\mathbb{N}}$ heißt Cauchy-Folge falls $d(x_n,x_m) \to 0$ für $n,m \to \infty$. Der metrische Raum heißt **vollständig**, falls jede Cauchy-Folge konvergent ist.

1.2 Normierte Räume

Definition 1.7 Ein normierter Raum $(X, \|\cdot\|)$ ist ein Paar bestehend aus einem $\mathbb{K}$ -Vektorraum X und einer Abbildung $\|\cdot\| : X \to [0, \infty)$ mit

1. $\|x\| = 0 \iff x = 0$
2. $\|\lambda x\| = |\lambda|\|x\| \forall \lambda \in \mathbb{K}, x \in X$
3. $\|x + y\| \leq \|x\| + \|y\| \forall x, y \in X$

Bemerkung

1. Die Norm $\|\cdot\|$ induziert auf X eine Metrik $d(x, y) = \|x - y\|$
2. Eine Metrik d auf einem Vektorraum definiert die Norm $\|d(x, 0)\|$ nur dann, wenn

$$\forall \lambda \in \mathbb{K} \forall x, y, z \in X : d(\lambda x, \lambda y) = |\lambda| d(x, y) \qquad \text{(Homogenität)}$$
$$d(x + z, y + z) = d(x, y) \qquad \text{(Translationsinvarianz)}$$

Definition 1.8 (Banachraum) Ein normierter Raum $(X, \|\cdot\|)$ heißt vollständig, falls X als metrischer Raum mit der Metrik $d(x, y) = \|x - y\|$ vollständig ist. Ein solcher vollständiger normierter Raum heißt **Banachraum**

Beispiel 1.9

1. $(\mathbb{R}^n, \|\cdot\|_2)$, wobei

$$\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{\frac{n}{2}}$$

2. Sei K eine kompakte Menge:

$$C_{\mathbb{K}} := \{f : K \to \mathbb{K} \mid f \text{ stetig}\}$$

$$\|\cdot\|_\infty = \max_{\lambda \in K} |f(x)|$$

$\left(C_{\mathbb{K}(K)}, \|\cdot\|_\infty\right)$ ist ein Banachraum.

Bemerkung

1. Jede Cauchy-Folge in $\mathbb{K}^n$ konvergiert, das heißt $(\mathbb{K}^n, \|\cdot\|)$ ist vollständig
2. Jede beschränkte Folge in $\mathbb{K}^n$ besitzt eine konvergente Teilfolge. (Der Satz von Bolzano-Weierstraß gilt in $\mathbb{R}^n$) (Beweis für $\mathbb{R}^n$ zum Beispiel in RR Ana2 Satz 1.1)

Satz 1.10 (Äquivalenz von Normen) Auf dem endlich dimensionalen Vektorraum $\mathbb{K}^n$ sind alle Normen **äquivalent** zur Maximumnorm, das heißt zu jeder Norm $\|\cdot\|$ gibt es positive Konstanten w, M mit denen gilt

$$m\|x\|_\infty \leq \|x\| \leq M\|x\|_\infty, x \in \mathbb{K}^n$$

Beweis Sei $\|\cdot\|$ irgendeine Norm $\forall x \in \mathbb{K}^n$ gilt

$$\|x\| \leq \sum_{k=1}^{n} |x_k| \left\| e^{(k)} \right\| \leq M \|x\|_\infty$$

mit

$$M := \sum_{k=1}^{n} \left\| e^{(k)} \right\|$$

Wir setzen

$$S_1 := \{x \in \mathbb{K}^m \mid \|x\|_\infty = 1\}, m := \inf\{\|x\|, x \in S_1\} \geq 0$$

Zu zeigen $m > 0$ (dann ergibt sich für $x \neq 0$ wegen $\|x\|_\infty^{-1} x \in S_1$ auch $m \leq \|x\|_\infty^{-1} \|x\| \implies 0 < m\|x\|_\infty \leq \|x\| \quad x \in \mathbb{K}^n$) Sei also angenommen, dass $m = 0$

Dann gibt eine eine Folge $\left(x^{(k)}\right)_{k \in \mathbb{N}} \in S_1$ mit $\left\| x^{(k)} \right\| \xrightarrow{k \to \infty} 0$. Da die Folge bezüglich $\|\cdot\|_\infty$ beschränkt ist, gibt es nach dem B.-W. Satz eine Teilfolge auch von $\left(x^{(k)}\right)$, die bezüglich $\|\cdot\|_\infty$ gegen ein $x \in \mathbb{K}^n$ konvergiert.

$$|1 - \|x\|_\infty| = \left| \left\| x^{(k)} \right\|_\infty - \|x\|_\infty \right| \leq \left\| x^{(k)} - x \right\|_\infty \to 0 \implies \|x\|_\infty = 1 \implies x \in S_1$$

Anderseits gilt

$$\forall k \in \mathbb{N} : \|x\| \leq \left\| x - x^{(k)} \right\| + \left\| x^{(k)} \right\| \leq M \left\| x - x^{(k)} \right\|_\infty + \left\| x^{(k)} \right\| \xrightarrow{k \to \infty} \implies x = 0$$

↯zu $x \in S_1$ □

Definition 1.11 Eine Menge $M \subset K^n$ heißt kompakt (folgenkompakt), wenn jede beliebige Folge in M eine konvergente Teilfolge besitzt, deren Grenzwert ebenfalls in M enthalten ist.

Beispiel 1.12 Mit Hilfe von dem Satz von B.W. folgt, dass alle abgeschlossene Kugeln im $\mathbb{R}^n$ ($K_r(a), a \in K^n$) kompakt sind. Ferner ist für beschränkte Mengen M der Rand ∂M kompakt. Jede endliche Menge ist auch kompakt.

1.3 Hilberträume

Definition 1.13 Sei $H\mathbb{K}$ Vektorraum. Ein **Skalarprodukt** auf eine Abbildung

$$(\cdot, \cdot) : H \times H \to \mathbb{K}$$

mit

1. $\forall x, y, z \in H, \lambda \in \mathbb{K} : (z, x + \lambda y) = (z, x) + \lambda(z, y)$
2. $\forall x, y \in H : (x, y) = \overline{(y, x)}$
3. $\forall x \in H : (x, x) \geq 0 \wedge (x, x) = 0 \iff x = 0$

$(H, (\cdot, \cdot))$ nennt man einen Prähilbertraum.

Bemerkung Für $\mathbb{K} = \mathbb{C}$ ist das Skalarprodukt linear in der zweiten Komponente aber antilinear in der ersten ($(\lambda x, y) = \bar{\lambda}(x, y)$).

Lemma 1.14 (Cauchy-Schwarz Ungleichung) Sei $(H, (\cdot, \cdot))$ Prähilbertraum, dann gilt

$$\forall x, y \in H : |(x, y)|^2 \leq (x, x)(y, y)$$

Beweis Da die Ungleichung für $y = 0$ bereits erfüllt ist, können wir ohne Beschränkung der Allgemeinheit annehmen $y \neq 0$. Für ein beliebiges $\alpha \in \mathbb{K}$ gilt

$$0 \leq (x + \alpha y, x + \alpha y) = (x, x) + \bar{\alpha}(y, x) + \alpha(x, y) + \alpha\bar{\alpha}(y, y)$$

Setze nun $\alpha := -(x, y)(y, y)^{-1}$

$$\begin{aligned}
&= (x, x) - \overline{(x, y)}(y, y)^{-1} - (x, y)(y, y)^{-1}(x, y) - \left|(x, y)^2\right|(y, y)^{-1} \\
&= (x, x) - \underbrace{((y, x)(y, x) + (x, y)(x, y))(y, y)^{-1}}_{>0} - |(x, y)|^2(y, y)^{-1} \\
&\leq (x, x) - |(x, y)|^2(y, y)^{-1} \\
&\iff |(x, y)|^2 \leq (x, x)(y, y)
\end{aligned}$$

□

Korollar 1.15 Sei $(H, (\cdot, \cdot))$ ein Prähilbertraum, dann ist $\|x\| := \sqrt{(x, x)}$ eine Norm auf H.

Beweis Es ist nur die Dreiecksungleichung zu beweisen, weil der Rest klar ist. Für $x, y \in H$ gilt

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

□

Definition 1.16 Ein Prähilbertraum $(H, (\cdot, \cdot))$ heißt Hilbertraum, falls $(H, \|\cdot\|)$ mit $\|x\| := \sqrt{(x, x)}$ ein Banachraum ist.

Beispiel 1.17

1. $H = \mathbb{R}^n$ versehen mit $\underbrace{(x, y) := \sum_{i=1}^{n} x_i y_i}_{\text{euklidisches Skalarprodukt}}$ ist ein Hilbertraum

2. $H = \mathbb{C}^n$ mit $\underbrace{(x, y) := \sum_{i=1}^{n} \bar{x}_i y_i}_{\text{euklidisches Skalarprodukt}}$ ist ein Hilbertraum

3. Sei $l^2\mathbb{K} := \{(x_k)_{k \in \mathbb{N}} \mid x_k \in \mathbb{K}, \forall k \in \mathbb{N} \wedge \sum_{i=1}^{\infty} |x_k|^2 < \infty\}$ versehen mit $(x, y) := \sum_{i=1}^{\infty} \bar{x}_i y_i$ ist ein Hilbertraum.

$$\sum_{i=1}^{n} |x_i||y_i| \leq \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} |y_i|^2\right)^{\frac{1}{2}} \leq \|x\|_{l^2}\|y\|_{l^2} < \infty$$

Lemma 1.18 (Hölder-Ungleichung) Für das euklidische Skalarprodukt $(\cdot,\cdot)_2$ gilt für beliebige p, q mit $1 < p, q < \infty$ und $\frac{1}{p} + \frac{1}{q} = 1$ die Ungleichung

$$\forall x, y \in \mathbb{K}^n : |(x,y)_2| \leq \|x\|_p \|y\|_q, \|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$$

Darüber hinaus gilt die Ungleichung auch für $p = 1, q = \infty$

Lemma 1.19 (Young'sche Ungleichung) Tür $p, q \in \mathbb{R}, 1 < p, q < \infty, \frac{1}{p} + \frac{1}{q} = 1$ gilt

$$\forall x, y \in \mathbb{K} : |(x,y)| \leq \frac{|x|^p}{p} + \frac{|y|^q}{q}$$

Lemma 1.20 (Minkowski-Ungleichung) Für ein beliebiges $p \in [1, \infty]$ gilt

$$\forall x, y \in \mathbb{K}^n : \|x+y\|_p \leq \|x\|_p + \|y\|_p$$

Satz 1.21 (Banachscher Fixpunktsatz) Sei (M, d) ein vollständiger, metrischer Raum und $f : M \to M$ ist eine strenge Kontraktion, das heißt

$$\exists 0 < \alpha < 1 \forall x, y \in M : d(f(x), f(y)) < \alpha d(x,y)$$

Dann existiert ein eindeutiger Fixpunkt von f, das heißt es existiert ein eindeutiges $x^* \in M : f(x^*) = x^*$

Beweis Existenz:
Wähle ein $x_0 \in M$ beliebig, aber fest und definiere dann $x_1 := f(x_0), x_2 := f(x_1), \ldots$. Dann gilt für $n \leq m$

$$\begin{aligned} d(x_n, x_m) &= d(f(x_{n-1}), f(x_{m-1})) < \alpha d(x_{n-1}, x_{m-1}) \\ &= \alpha d(f(x_{n-2}), f(x_{m-2})) < \cdots < \alpha^n d(x_0, x_{m-n}) \end{aligned}$$

Nun gilt aber

$$\begin{aligned} d(x_0, x_{m-n}) &\leq d(x_0, x_1) + d(x_1, x_2) + \cdots + d(x_{m-n-1}, x_{m-n}) \\ &\leq d(x_0, x_1) + \alpha d(x_0, x_1) + \cdots + a^{m-n-1} d(x_0, x_1) \\ &= d(x_0, x_1) \sum_{i=0}^{m-n-1} \alpha^i \leq d(x_0, x_1) \sum_{i=0}^{\infty} \alpha^i \\ &= \frac{d(x_0, x_1)}{1-\alpha} < \infty \\ \Longrightarrow d(x_n, x_m) &\leq \frac{\alpha^n}{1-\alpha} d(x_0, x_1) \end{aligned}$$

Also ist $(x_k)_{k\in\mathbb{N}}$ Cauchy-Folge. Da (M,d) vollständig ist existiert $x^* \in M$, sodass $x_k \xrightarrow{k\to\infty} x^*$. Zeige, dass x^* Fixpunkt von f ist:

$$\begin{aligned} 0 \le d(x^*, f(x^*)) &\le d(x^*, x_k) + d(x_k, f(x^*)) \\ &\le d(x^*, x_k) + \alpha d(x_{k-1}, x^*) \xrightarrow{k\to\infty} 0 \end{aligned}$$

$\implies f(x^*) = x^*$
Eindeutigkeit: Angenommen $\exists x' \in M, x' \neq x^* : f(x') = x'$:

$$0 < d\big(x^*, x'\big) = d\big(f(x^*), f\big(x'\big)\big) < \alpha d\big(x^*, x'\big) \implies \alpha > 1`$$

□

2 Stetigkeit und Differenzierbarkeit im $\mathbb{R}^n$

Definition 2.1 Eine Funktion $f : D \subset \mathbb{K}^n \to \mathbb{K}^m, m, n \in \mathbb{N} \setminus \{0\}, D \neq \emptyset$, ist stetig in einem $a \in D$, wenn

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in D : \|x - a\| < \delta \implies \|f(x) - f(a)\| < \varepsilon$$

Bemerkung Es gelten auch im Mehrdimensionalen die Permanenzeigenschaften, das heißt f, g stetig $\implies f + g, f \circ g$ sind stetig.

Satz 2.2 Eine stetige Funktion $f : D \subset \mathbb{K}^n \to \mathbb{K}^m$ ist auf einer kompakten Menge $K \subset D$ beschränkt, das heißt für jede kompakte Menge K existiert eine Konstante M_k, sodass

$$\forall x \in K \|f(x)\| < M_k$$

Beweis Angenommen f wäre auf K unbeschränkt, dann gäbe es zu jedem $k \in \mathbb{N}$ ein $x_k \in K$ mit $\|f(x_k)\| > K$. Da K kompakt hat die Folge $(x_k)_{k\in\mathbb{N}}$ eine konvergente Teilfolge $\big(x_{k_j}\big)_{j\in\mathbb{N}}$ für die gilt $x_{k_j} \xrightarrow{j\to\infty} x \in K$. Da f stetig $f\big(x_{k_j}\big) \to f(x)$ und $\|f(x)\| < \infty$, was im Widerspruch steht zu $\|f(x_k)\| \xrightarrow{k\to\infty} \infty$. □

Satz 2.3 Eine stetige Funktion $f : D \subset \mathbb{K}^n \to \mathbb{R}$ nimmt auf jeder (nicht leeren) kompakten Menge $K \subset D$ ihr Minimum und Maximum an.

Beweis Nach Satz 2.2 besitzt f eine obere Schranke auf K

$$\mathcal{K} := \sup_{x\in K} f(x)$$

Dazu $(x_k)_{k\in\mathbb{N}} \subseteq K$, sodass $f(x_k) \xrightarrow{k\to\infty} \mathcal{K}$. Da K kompakt existiert eine konvergente Teilfolge $\big(x_{k_j}\big)_{j\in\mathbb{N}}$ und ein x_{max}, sodass $x_{k_j} \xrightarrow{j\to\infty} x_{max}$. Da f stetig, gilt $f\big(x_{k_j}\big) \to f(x_{max})$. □

Bemerkung Auf diese Weise lassen sich die Ergebnisse der Stetigkeit aus dem Eindimensionalen ins Mehrdimensionale verallgemeinern.

Im folgenden Teil sei $D \subseteq \mathbb{R}^n$ offen, $\mathbb{K} = \mathbb{R}$

Definition 2.4 Eine Funktion $f : D \to \mathbb{R}$ heißt in einem Punkt $x \in D$ partiell differenzierbar bezüglich der i-ten Koordinatenrichtung, falls der Limes

$$\lim_{h\to 0} \frac{f(x+he_i) - f(x)}{h} =: \frac{\partial f}{\partial x_i}(x) =: \partial_i f(x)$$

existiert. Existieren in allen Punkten $x \in D$ **alle** partiellen Ableitungen, so heißt f partiell differenzierbar. Sind alle partiellen Ableitungen stetig auf D, so heißt f stetig partiell differenzierbar. Eine Funktion $f : D \to \mathbb{R}^m$ heißt (stetig) partiell differenzierbar, wenn $f_i, i = 1, \ldots, m$ (stetig) partiell differenzierbar.

Bemerkung Die Ableitungsregeln aus dem Eindimensionalen übertragen sich auf partielle Ableitungen.

Beispiel

1. Polynome sind stetig partiell differenzierbar. Sei $p : D \subset \mathbb{R}^2 \to \mathbb{R}, (x_1, x_2) \mapsto a_{01}x_2 + a_{11}x_1x_2 + a_{02}x_2^2 + a_{21}x_1^2x_2$. Dann ist

$$\frac{\partial p}{\partial x_1}(x_1, x_2) = a_{11}x_2 + 2a_{21}x_1x_2 \quad \frac{\partial p}{\partial x_2} = a_{01} + a_{11}x_1 + 2a_{02}x_2 + a_{21}x_1^2$$

2. $\|\cdot\|_2 : \mathbb{R}^k \setminus \{0\} \to \mathbb{R}$ ist stetig partiell differenzierbar, da

$$\frac{\partial\|\cdot\|}{x_i} = \frac{1}{2}\frac{2x_i}{\left(x_1^2 + \cdots + x_k^2\right)^{\frac{1}{2}}} = \frac{x_i}{\|x\|_2}$$

3. $f : \mathbb{R}^2 \to \mathbb{R}, (x_1, x_2) \mapsto \frac{x_1x_2}{\left(x_1^2+x_2^2\right)^2}$ für $x \neq 0, f(0) = 0$

$$\frac{\partial f}{\partial x_1}(x) = \frac{x_2}{\left(x_1^2 + x_2^2\right)^2} - 4\frac{x_1^2x_2}{\left(x_1^2 + x_2^2\right)^3}, x \neq 0$$

Für $x = 0$ ist $f(0) = 0$

$$\Longrightarrow \lim_{h\to 0} \frac{f(xe_i) - f(0)}{h} = 0$$

Sei $x_\varepsilon(\varepsilon, \varepsilon)$ und damit gilt $\|x_\varepsilon\|_2 \xrightarrow{\varepsilon\to 0} 0$

$$f(x_\varepsilon) = \frac{\varepsilon^2}{4\varepsilon^4} = \frac{1}{4\varepsilon^2} \xrightarrow{\varepsilon\to 0} \infty$$

Satz 2.5 Die Funktion $f : D \to \mathbb{R}$ habe in einer Kugelumgebung $K_r(x) \subset D$ eines Punktes $x \in D$ beschränkte partielle Ableitungen, das heißt

$$\sup_{y\in K_r(x)} \left|\frac{\partial f}{\partial x_i}\right| \leq M, i = 1, \ldots, n$$

dann ist f stetig in x.

Beweis Es genügt $n = 2$. Für $(y_1, y_2) \in K_r(x)$

$$f(y_1, y_2) - f(x_1, x_2) = f(y_1, y_2) - f(x_1, y_2) + f(x_1, y_2) - f(x_1, x_2)$$

Nach dem 1-D Mittelwertsatz existieren $\xi, \eta \in K_r(x)$, sodass

$$\begin{aligned} |f(y_1, y_2) - f(x_1, x_2)| &= \frac{\partial f}{\partial x_1}(\xi, y_2)(y_1 - x_1) + \frac{\partial f}{\partial x_2}(x_1, \eta)(y_2 - x_2) \\ &\leq M(|y_1 - x_1| + |y_2 - x_2|) \end{aligned}$$

□

Höhere partielle Ableitungen definieren sich durch sukzessives Ableiten, das heißt

$$\frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial x_k} f(x) = \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}$$

Beispiel

$$\frac{x_1}{x_2} := \frac{x_1^3 x_2 - x_1 x_2^3}{x_1^2 + x_2^2}$$

für $(x_1, x_2) \neq (0, 0)$, $f(0, 0) = 0$. f zweimal partiell differenzierbar, aber

$$\frac{\partial^2}{\partial x_1 \partial x_2} f(0, 0) \neq \frac{\partial^2}{\partial x_2 \partial x_1} f(0, 0)$$

Satz 2.6 Eine Funktion $f : D \to \mathbb{R}$ sei in einer Umgebung $K_r(x) \subset D$ eines Punktes $x \in D$ zweimal stetig partiell differenzierbar, dann gilt

$$\frac{\partial^2}{\partial x_i \partial x_j} f(x) = \frac{\partial^2}{\partial x_j \partial x_i} f(x), i, j = 1, \ldots, n$$

Beweis $n = 2$. Sei $A := f(x_1 - h_1, x_2 + h_2) - f(x_1 + h_1, x_2) - f(x_1, x_2 + h_2) + f(x_1, x_2)$.

$$\varphi(x_1) := f(x_1, x_2 + h_2) - f(x_1, x_2) \implies A = \varphi(x_1 + h_1) - \varphi x_1$$

Mit dem Mittelwertsatz erhalten wir $A = h_1 \varphi'(x_1 + \theta_1 h_1), \theta_1 \in (0, 1)$.

$$\varphi'(x_1) = \frac{\partial}{\partial x_1} f(x_1, x_2 + h_2) - \frac{\partial}{\partial x_1} f(x_1, x_2) = h_2 \frac{\partial^2}{\partial x_2 x_1} f\big(x_1, x_2 + \theta_1' h_2\big), \theta_1' \in (0, 2)$$

Analog verfahre man mit x_2 und erhalte für $\psi(x_2) := f(x_1 + h_1, x_2) - f(x_1, x_2)$

$$A = \psi(x_2 - h_2) - \psi(x_2) = h_2 \psi'(x_2 + \theta_2 h_2) = h_1 h_2 \frac{\partial^2}{\partial x_1 \partial x_2} f\big(x_1 + \theta_2 h_1, x_2 \theta_2' h_2\big)$$

$$\begin{aligned} \implies \frac{\partial^2}{\partial x_2 \partial x_1} f\big(x_1 + \theta_1 h_1, x_2 + \theta_1' h_2\big) &= \frac{\partial^2}{\partial x_1 \partial x_2} f\big(x_1 + \theta_2 h_1, x_2 + \theta_2' h_2\big) \\ \xRightarrow{h_1, h_2 \to 0} \frac{\partial^2}{\partial x_2 \partial x_1} f(x_1, x_2) &= \frac{\partial^2}{\partial x_1 \partial x_2} f(x_1, x_2) \end{aligned}$$

□

Definition 2.7 $f : D \to \mathbb{R}$ partiell differenzierbar.

$$\operatorname{grad} f(x) := \left(\frac{\partial}{\partial x_1} f, \ldots, \frac{\partial}{\partial x_n} f\right)^T \in \mathbb{R}^n$$

heißt **Gradient** von f in $x \in D$. Man schreibt $\nabla f(x) := \operatorname{grad} f$. $f : D \to \mathbb{R}^n$ partiell differenzierbar.

$$\operatorname{div} f(x) := \frac{\partial}{\partial x_1} f_1(x) + \cdots + \frac{\partial}{\partial x_n} f_n(x)$$

Es gilt:

$$\operatorname{div} \operatorname{grad} f(x) := \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} f_i =: \Delta f(x)$$

Definition 2.8 $f : D \to \mathbb{R}^m$ partiell differenzierbar. Die Matrix der ersten partiellen Ableitungen

$$J_f := \left(\frac{\partial f_i}{\partial x_j}\right)_{\substack{i=1,\ldots,w \\ j=1,\ldots,n}} \in \mathbb{R}^{n \times w}$$

heißt die **Jacobi-Matrix** (manchmal auch **Fundametalmatrix**) von f in x. Im Fall $n = m$ bezeichnet man $\det(J_f)$ als **Jacobideterminante**.

Definition 2.9 $f : D \to \mathbb{R}$ zweimal partiell differenzierbar. Die Matrix der zweiten Ableitungen

$$H_f(x) := \left(\frac{\partial^2}{\partial x_i \partial x_j} f\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,w}} \in \mathbb{R}^{n \times m}$$

heißt **Hesse-Matrix**.

Definition 2.10 Sei $f : D \to \mathbb{R}^m$, dann nennen wir f in einem Punkt $x \in D$ (total differenzierbar), wenn die Funktion f in x sich linear approximieren lässt, das heißt es gibt eine lineare Abbildung $Df(x) : \mathbb{R}^n \to \mathbb{R}^m$ (Differential) sodass in einer kleinen Umgebung von x gilt:

$$f(x+h) = f(x) + Df(x)h + w(h), h \in \mathbb{R}^n, x+h \in D$$

mit einer Funktion $w : D \to \mathbb{R}^m$, die die Eigenschaft hat

$$\lim_{\substack{x+h \in D \\ \|h\|_2 \to 0}} \frac{\|wh\|_2}{\|h\|_2} = 0$$

alternativ: $w(h) = \wr(\|h\|_2)$

Satz 2.11 Für Funktionen $f : D \to \mathbb{R}^m$ gilt:

1. Ist f in $x \in D$ differenzierbar, so ist f auch in x partiell differenzierbar und das Differential von f ist gegeben durch die Jacobi-Matrix.

2. Ist f partiell differenzierbar in einer Umgebung von x und sind zusätzlich die partiellen Ableitungen stetig in x, so ist f in x differenzierbar.

Beweis 1. Für differenzierbares f gilt für $i = 1, 2$:

$$\lim_{h\to 0} \frac{f(x+he_i) - f(x)}{h} = \lim_{h\to 0}\left(Df(x)e_i + \frac{w(h)}{h}\right) = Df(x)e_i$$

2. Für ein stetig partiell differenzierbares f gilt mit $h = (h_1, h_2)$:

$$f(x+h) - f(x) = f(x_1+h_1, x_2+h_2) - f(x_1+h_1, x_2) + f(x_1+h_1, x_2) - f(x_1, x_2)$$

Mittelwertsatz

$$\begin{aligned}
&= h_2 \frac{\partial f}{\partial x_2}(x_1+h_1, x_2+\theta_2 h_2) + h_1 \frac{\partial f}{\partial x_1}(x_1+\theta_1 h_1, x_2) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \theta_1, \theta_2 \in (0,1) \\
&= h_2\left(\frac{\partial f}{\partial x_2}(x_1, x_2) + \omega_2(h_1, h_2)\right) + h_1\left(\frac{\partial f}{\partial x_1}(x_1, x_2) + \omega_1(h_1, h_2)\right)
\end{aligned}$$

$$\omega_1(h_1, h_2) := \frac{\partial f}{\partial x_1}(x_1+\theta_1 h_1, x_2) - \frac{\partial f}{\partial x_1}(x_1, x_2) \xrightarrow{h_1, h_2 \to 0} 0$$

$$\omega_2(h_1, h_2) := \frac{\partial f}{\partial x_2}(x_1+h_1, x_2+\theta_2 h_2) - \frac{\partial f}{\partial x_2}(x_1, x_2) \xrightarrow{h_1, h_2 \to 0} 0$$

Also ist f differenzierbar mit Ableitungen $Df(x) = \nabla f(x)$. □

Bemerkung Es gelten folgende Implikationen: stetig partiell differenzierbar $\Longrightarrow$ (total) differenzierbar $\Longrightarrow$ partiell differenzierbar.

Satz 2.12 Seien $D_f \subset \mathbb{R}^n, Dg \subseteq \mathbb{R}^m$ offen und $g : D_g \to \mathbb{R}^n, f : D_f \to \mathbb{R}^r$. Ist g im Punkt $x \in D_g$ differenzierbar und f in $y = g(x) \in D_f$ differenzierbar, so ist die Komposition $h = f \circ g$ im Punkt x differenzierbar. Es gilt $D_x h(x) = D_y f(g(x)) \cdot D_x g(x)$. Hierbei ist $\cdot$ die Matrixmultiplikation.

Beweis Nach Voraussetzung $x \in D_g$ sodass $g(x) = y \in D_f$. Da sowohl f als auch g differenzierbar

$$\begin{aligned}
g(x+h_1) &= g(x) + D_x g(x)h_1 + \omega_g(h_1)\\
f(y+h_2) &= f(y) + D_y f(y)h_2 + \omega_f(h_2)\\
\lim_{\substack{x+h_1\in D_y\\ \|h_1\|\to 0}} \frac{\|\omega_g(h_1)\|}{\|h_1\|} &= 0\\
\lim_{\substack{y+h_2\in D_y\\ \|h_2\|\to 0}} \frac{\|\omega_f(h_2)\|}{\|h_2\|} &= 0\\
(f\circ g)(x+h_1) &= f(g(x+h_1)) = f(y+\eta), \quad \eta := D_x g(x)h_1 + \omega_g(h_1)\\
&= f(y) + D_y f(y)\eta + \omega_f(\eta)\\
&= f(y) + D_y f(y) D_x g(x) h_1 + D_y f(y)\omega_g(h_1) + \omega_f(D_x g(x)h_1 + \omega_g(h_1))\\
&= (f\circ g)(x) + D_y f(y) D_x g(x) h_1 + \omega_{f\circ g}(h_1)\\
\omega_{f\circ g}(h_1) &:= D_y f(y)\omega_g(h_1) + \omega_f(D_x g(x)h_1 + \omega_g(h_1))
\end{aligned}$$

Es bleibt zu zeigen $\omega_{f\circ g} = \wr(h_1)$. Nach Voraussetzung gilt $\omega_{f\circ g} \xrightarrow{h_1\to 0} 0$

□

Lemma 2.13 Sei $A : [a,b] \to \mathbb{R}^{n\times m}$ stetig, dann gilt

$$\left\|\int 0^1 A(s)\mathrm{d}s\right\|_M \leq \int_0^1 \|A(s)_M \mathrm{d}s\|, \|A\|_M := \max\{|\lambda| \mid \lambda \in \sigma(A)\}$$

$\int A = \left(\int a_{ij}\right)_{ij}, \sigma(A) :=$ Menge der Eigenwerte von A

Satz 2.14 Sei $f : D \to \mathbb{R}^m$ stetig differenzierbar mit J_f als Jacobi-Matrix, so gilt

$$f(x+h) - f(x) = \left(\int_0^1 J_f(x+sh)\mathrm{d}s\right)h$$

Beweis Definiere $g_j(s) := f_j(x+sh)$, dann ist $g_{j_1} : [0,1] \to \mathbb{R}$, also gilt

$$f_j(x+sh) - f_j(x) = g_j(1) - g_j(0) = \int_0^1 g_j'(s)\mathrm{d}s = \int_0^1 \sum_{i=1}^n \frac{\partial f_j}{\partial x_i}(x+sh)h_i\mathrm{d}s$$

□

Bemerkung Im Fall $m = 1$ kann man aus dem Mittelwertsatz für Integrale schließen, dass

$$f(x+h) - f(x) = \int_0^1 J_f(x+sh)h\mathrm{d}s = J_f(x+\tau h)h$$

$x_1 + h = x_2 \implies h = x_2 - x_1$

Korollar 2.15 Sei $f : D \to \mathbb{R}^m$ stetig differenzierbar. Ferner sei $x \in D$ mit $K_r(x) \subset D, r > 0$, dann gilt

$$\|f(x) - f(y)\|_2 \leq M\|x - y\|_2, y \in K_r(x), M := \sup_{z \in K_r(x)} \|J_f(z)\|_M$$

das heißt die Abbildung ist in D lokal Lipschitz-stetig.

Beweis Nach Satz 2.14 gilt mit $h = y - x$

$$\begin{aligned}\|f(y) - f(x)\|_2 = \|f(x+h) - f(x)_2\| &= \left\| \int_0^1 J_f(x+sh)h\mathrm{d}s \right\|_2 \\ &\leq \int_0^1 \|J_f(x+sh)h\|_2 \mathrm{d}s \leq \int_0^1 \|J_f(x+sh)\|_m \|h\|_2 \mathrm{d}s \\ &\leq \underbrace{\sup_{0<s<1} \|J_f(x+sh)\|_2}_{M} \underbrace{\|h\|_2}_{\|y-x\|_2}\end{aligned}$$

□

Bemerkung Korollar 2.16 gilt mit beliebigen von Vektor-Matrix-norm induzierter Norm, siehe Übung 2.1.

Taylor-Entwicklung und Extremwerte in $\mathbb{R}^n$

Definition 2.16 (Multiindex Notation) Ein n-dimensionaler **Multiindex** ist ein Tupel $\alpha = (\alpha_1, \dots, \alpha_n)$ mit $\alpha_i \in \mathbb{N}$. Für Multiindizes sind die **Ordnung** $|\alpha|$ und die Fakultät $\alpha!$ definiert durch

$$\begin{aligned}|\alpha| &:= \alpha_1 + \alpha_2 + \dots + \alpha_n \\ \alpha! &:= \alpha_1! \cdot \ldots \cdot \alpha_n!\end{aligned}$$

Für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ wird gesetzt

$$x^\alpha := x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}$$

Für eine $|\alpha|$ -mal stetig differenzierbare Funktion wird gesetzt

$$\partial^\alpha f := \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n} f := \frac{\partial^{|\alpha|} f}{\partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n}}$$

Bemerkung Wegen der Stetigkeit der Ableitung ist dieser Ausdruck unabhängig von der Reihenfolge der partiellen Ableitungen. Wir definieren

$$\sum_{|\alpha|=0}^{r} a_\alpha := \sum_{k=0}^{r} \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|=k}} a_\alpha$$

Beispiel 2.17 Für $n = 3$ sind die Multiindizes $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ der Ordnung $|\alpha| = 2$ gegeben durch

$$(2,0,0), (0,2,0), (0,0,2), (1,1,0), (1,0,1), (0,0,1)$$

Die zugehörigen partiellen Ableitungen sind

$$\begin{aligned}\partial^\alpha f &= \left(\partial^2_{x_1} f, \partial^2_{x_2} f, \partial^2_{x_3} f, \partial_{x_1}\partial_{x_2} f, \partial_{x_2}\partial_{x_3} f, \partial_{x_2}\partial_{x_3} f\right)\\ \alpha! &= (2,2,2,1,1,1)\end{aligned}$$

Schließlich ist

$$\sum_{|\alpha|=2} \partial^\alpha f = \partial^2_{x_1} f + \partial^2_{x_2} f + \partial^2_{x_3} f + \partial_{x_1}\partial_{x_2} f + \partial_{x_2}\partial_{x_3} f + \partial_{x_2}\partial_{x_3} f$$

Satz 2.18 (Taylor-Formel) Sei $D \subset \mathbb{R}^n$ eine offene Menge und $f : D \to \mathbb{R}$ eine $(r+1)$ -mal stetig differenziebare Funktion. Dann gilt für jeden Vektor $h \in \mathbb{R}^n$ mit $x + sh \in D, s \in [0,1]$ die Taylor-Formel

$$f(x+h) = \sum_{|\alpha|<r} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha + R^f_{r+1}(x,h)$$

in differentieller Form

$$R^f_{r+1}(x,h) = \sum_{|\alpha|=r+1} \frac{\partial^\alpha f(x+\theta h)}{\alpha!} h^\alpha, \theta \in (0,1)$$

oder in integraler Form

$$R^f_{r+1}(x,h) = (r+1) \int_0^1 \sum_{|\alpha|=r+1} \frac{\partial^\alpha f(x+th)}{\alpha!} h^\alpha (1-t)^r \mathrm{d}t$$

Beweis Wir nehmen $g : [0,1] \to \mathbb{R}$ mit $g(t) := f(x+th)$. g ist $(r+1)$ mal stetig differenzierbar mit der k-ten Ableitung

$$g^{(k)}(t) = \sum_{i_1,\ldots,i_k=1}^{n} \partial_{i_k} \ldots \partial_{i_1} f(x+th) h_{i_1} \ldots h_{i_k}$$

Wir zeigen des durch Induktion nach k (mit Hilfe von Kettenregel). Für $k = 1$ gilt

$$g'(t) = \frac{\mathrm{d}}{\mathrm{d}t} f(x_1 + th_1, \ldots, x_n + th_n) = \sum_{i=1}^{n} \partial_i f h_i$$

Sei die Behauptung als richtig angenommen für $k-1 \geq 1$. Dann gilt

$$
\begin{aligned}
g^{(k)}(t) = \frac{\mathrm{d}}{\mathrm{d}t} g^{(k-1)}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i_1,\ldots,i_{k-1}=1}^{n} \partial_{i_{k-1}} \ldots \partial_{i_1} f(x+th) h_1 \ldots h_{i_{k-1}} \right) \\
&= \sum_{i=1}^{n} \partial_i \left(\sum_{i_1 \ldots i_{k-1}=1}^{n} \partial_{i_{k-1}} \ldots \partial_{i_1} f(x+th) h_{i_1} \ldots h_{i_{k-1}} \right) h_1 \\
&= \sum_{i_1 \ldots i_k=1}^{n} \partial_{i_k} \ldots \partial_{i_1} f(x+th) h_{i_1} \ldots h_{i_k}
\end{aligned}
$$

Es gilt

$$\partial_{i_k} \ldots \partial_{i_1} f(x+th) h_{i_1} \ldots h_{i_k} = \partial_1^{\alpha_1} \ldots \partial_n^{\alpha_n} f(x+th) h_1^{\alpha_1} \ldots h_n^{\alpha_n}$$

(der Index $i \in \{1, \ldots, n\}$ kommt genau α_i mal vor und wegen Vertauschbarkeit der Ableitungen). Die Anzahl der k -Tupel $(i_1, \ldots, i_k)$ von Zahlen $i_j \in \{1, \ldots, n\}$, bei denen die Zahl $i \in \{1, \ldots, n\}$ genau α_i -mal vorkommt mit $\alpha_1 + \cdots + \alpha_n = k$ ist

$$\frac{k!}{\alpha_1! \ldots \alpha_n!}$$

(Lemma unten) Wir bekommen

$$
\begin{aligned}
g^{(k)}(t) &= \sum_{|\alpha|=k} \frac{k!}{\alpha_1! \ldots \alpha_n!} \partial_1^{\alpha_1} \ldots \partial_n^{\alpha_n} f(x+th) h_1^{\alpha_1} \ldots h_n^{\alpha_n} \\
&= \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(x+th) h^\alpha
\end{aligned}
$$

Wir wenden die 1-dimensionale Taylor-Formel auf $g(t)$ an. $\exists \theta \in [0,1]$ sodass

$$g(1) = \sum_{k=0}^{r} \frac{g^k(0)}{k!} + \frac{g^{(r+1)}(\theta)}{(r+1)!} = \sum_{k=0}^{n} \frac{g^{(k)}}{k!} + \frac{1}{r!} \int_0^1 g^{(r+1)}(t)(1-t)^r \mathrm{d}t$$

Man erhält

$$
\begin{aligned}
\frac{g^{(k)}(0)}{k!} &= \sum_{|\alpha|=k} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha \\
\frac{g^{(r+1)}(\theta)}{(r+1)!} &= \sum_{|\alpha|=r+1} \frac{\partial^\alpha f(x+\theta h)}{\alpha!} h^\alpha \\
\frac{1}{r!} \int_0^1 g^{(r+1)}(t)(1-t)^r \mathrm{d}t &= (r+1) \int_0^1 \sum_{|\alpha|=r+1} \frac{\partial^\alpha f(x+th)}{\alpha!} h^\alpha (1-t)^r \mathrm{d}t
\end{aligned}
$$

Dies impliziert die Taylor-Formel mit den Restgliedern in differentieller oder integraler Form. $\square$

Lemma 2.19 (2.20) Sei $\alpha = (\alpha_1, \dots, \alpha_n)$ mit $|\alpha| = k \geq 1$. Dann ist die Anzahl $N_\alpha(k)$ der k-Tupel von Zahlen $i_j = \{1, \dots, n\}$, bei denen die Zahl $i \in \{1, \dots, n\}$ genau α_i -mal vorkommt, bestimmt durch

$$N_\alpha(k) = \frac{k!}{\alpha_1! \dots \alpha_n!}$$

Beweis Wir ordnen die Indizes in dem k-Tupel

$$(i_1, \dots, i_k) = \left(\underbrace{1, \dots, 1}_{\alpha_1}, \underbrace{2, \dots, 2}_{\alpha_2}, \dots, \underbrace{n, \dots, n}_{\alpha_n \text{ mal}} \right)$$

$\alpha_1 + \dots + \alpha_n = k$. Die Anzahl der möglichen Permutationen der k Elemente des k-Tupel ist $k!$. Das k-Tupel bleibt unverändert bei Permutationen von gleichen Elementen i. Insgesamt bekommen wir

$$N_\alpha(k) = \frac{k!}{\alpha!}$$

□

Korollar 2.20 (2.21) Sei $D \subset \mathbb{R}^n$ eine offene Menge und $f : D \to \mathbb{R}$ eine $r + 1$ mal stetig differenzierbare Funktion. Dann gilt für $x \in D$ und $h \in \mathbb{R}^n$ mit $x + sh \in D$, $s \in [0, 1]$:

$$f(x + h) = \sum_{|\alpha| \leq r+1} \frac{\partial^a f(x)}{\alpha!} h^\alpha + \omega_{r+1}(x, h)$$

wobei $\omega_{r+1}(x, 0) = 0$ und $\omega_{r+1}(x, h) = \wr\left(\|h\|_2^{r+1} \right)$.
Im Fall $r = 0$ gilt

$$f(x + h) = f(x) + (\nabla f(x), h)_2 + \omega_1(x, h)$$

Im Fall $r = 1$ gilt:

$$f(x + h) = f(x) + (\nabla f(x), h)_2 + \frac{1}{2}(H_f(x)h, h)_2 + \omega_2(x, h)$$

Beweis

$$\begin{aligned} f(x + h) &= \sum_{|\alpha| \leq r} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha + \sum_{|\alpha| = r+1} \frac{\partial^\alpha f(x + \theta h)}{\alpha!} h^\alpha \\ &= \sum_{|\alpha| \leq r+1} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha + \sum_{|\alpha| = r+1} r_\alpha(x, h) h^\alpha \end{aligned}$$

wobei

$$r_\alpha(x, h) := \frac{\partial^\alpha f(x + \theta h) - \partial^\alpha f(x)}{\alpha!}$$

$\lim_{h \to 0} r_\alpha(x, h) = 0$, wegen der Stetigkeit von $\partial^\alpha f$ für $|\alpha| = r + 1$. Wir setzen $\omega_{r+1}(x, h) := \sum_{|\alpha| = r+1} r_\alpha(x, h) h^\alpha$. Es gilt

$$\lim_{h \to 0} \frac{\omega(h)}{\|h\|_2^{r+1}} = 0$$

weil

$$\frac{|h^\alpha|}{\|h\|_2^\alpha} = \frac{|h_1^{\alpha_1}| \cdot \ldots \cdot h_n^{\alpha_n}}{\|h\|_2^{\alpha_1} \cdot \ldots \cdot \|h\|_2^{\alpha_n}} \leq 1 \qquad |\alpha| = r + 1$$

Für $r = 0$ gilt

$$\begin{aligned} f(x+h) &= \sum_{|\alpha| \leq 1} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha + \omega_1(x,h) \\ &= f(x) + \sum_{|\alpha|=1} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha + \omega_1(x,h) \\ &= f(x) + \sum_{i=1}^n \partial_i f(x) h_i + \omega_1(x,h) \\ &= f(x) + (\nabla f(x), h)_2 + \omega_1(x,h) \end{aligned}$$

Für $r = 1$

$$\begin{aligned} f(x+h) &= \sum_{|\alpha| \leq 2} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha + \omega_2(x,h) \\ &= f(x)(\nabla f(x), h)_2 + \sum_{|\alpha|=2} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha + \omega_2(x,h) \\ &= f(x) + (\nabla f(x), h)_2 + \frac{1}{2} \sum_{i,j=1}^n \partial_i \partial_j f(x) h_i h_j + \omega_2(x,h) \\ &= f_1(x) + (\nabla f(x), h)_2 + \frac{1}{2} (H_f(x) h, h)_2 + \omega_2(x,h) \end{aligned}$$

□

Definition 2.21 Sei $D \subset \mathbb{R}^n$ eine offene Menge, $x \in D$ und $f : D \to \mathbb{R}$ beliebig oft differenzierbar.

$$F_\infty^f(x+h) = \sum_{|\alpha|=0}^\infty \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha$$

heißt die Taylor-Reihe von f in x

Korollar 2.22 Sei $D \subset \mathbb{R}^n$ eine offene Menge, $f : D \to \mathbb{R}$ beliebig oft differenzierbar. Dann konvergiert die Taylor-Reihe von f und stellt f dar, wenn

$$R_{r+1}^f(x,h) \xrightarrow{r \to \infty} 0 \quad x \in D$$

Hinreichend dafür ist, dass die partielle Ableitung gleichmäßig beschränkt sind:

$$\sup_{|\alpha| \geq 0} \sup_{x \in D} |\partial^\alpha f(x)| < \infty$$

Beweis

$$\left\|R_{r+1}^f(x,h)\right\|_\infty \leq \sum_{|\alpha|=r+1} \frac{|\partial^\alpha f(x+\theta h)|}{\alpha!}\|h\|_\infty^{|\alpha|} \leq M(f) \sum_{|\alpha|=r+1} \frac{1}{\alpha!}\|h\|_\infty^{|\alpha|} \to 0$$

□

Definition 2.23 Eine Funktion $f : D \to \mathbb{R}$ hat in einem Punkt $x \in D \subset \mathbb{R}^n$ ein lokales Extremum, wenn auf einer $K_\sigma(x) \subset \mathbb{R}^n$ (Kugelumgebung) gilt

$$f(x) = \sup_{y \in K_\sigma(x) \cap D} f(y) \quad \text{oder} \quad f(x) = \inf_{y \in K_\sigma(x) \cap D} f(x)$$

Das Extremum heißt strikt, wenn es in $K_\sigma(x) \cap D$ nur in dem Punkt angenommen wird. Das Extremum heißt global, wenn $f(x) = \sup_{y \in D} f(y)$ (oder $\inf_{y \in D}$)

Satz 2.24 (Notwendige Extremalbedingung) Sei $f : D \to \mathbb{R}$ stetig differenzierbar, D offen. Hat f in einem Punkt $\hat{x} \in D$ ein lokales Extremum, so gilt $\nabla f(\vec{x}) = 0$

Beweis Angenommen $f : D \to \mathbb{R}$ hat in $x \in D$ ein lokales Extremum. Wir nehmen $g_i(t) := f\left(\vec{x} + te^{(1)}\right), i = 1, \ldots, n, e^{(i)}$ Einheitsvektor in $\mathbb{R}^n$. g_i ist auf einem nichtleeren $(-\delta_i, \delta_i) \subset \mathbb{R}$ definiert und hat lokales Extremum in $t = 0 \implies g_i'(0) = 0$

$$0 = g_i'(0) = \sum_{j=1}^n \partial_j f(\vec{x})\delta_{ij} = \partial_i f(\vec{x}) \quad i = 1, \ldots, n \implies \nabla f(\vec{x}) = 0$$

□

Satz 2.25 (Hinreichende Extremalbedingung) Sei $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ zweimal stetig differenzierbar und $\nabla f(\vec{x}) = 0$ in einem $\vec{x} \in D$. Ist die Hesse Matrix $H_f(x)$ in $\vec{x}$ **positiv definit** (das heißt alle Eigenwerte positiv), so liegt in $\vec{x}$ ein striktes lokales Minimum. Ist sie negativ definit (das heißt alle Eigenwerte negativ), so liegt in $\vec{x}$ ein striktes lokals Maximum. Ist sie indefinit (hat sowohl positive als auch negative Eigenwerte), so kann in $\vec{x}$ kein lokales Extremum liegen.

Beweis Nach Korollar 2.21 gilt

$$f(x+h) = f(x) + (\nabla f(x), h)_2 + \frac{1}{2}(H_f(x)h, h)_2 + \omega_2(x,h)$$

wobei

$$\lim_{\substack{h \to 0 \\ h \neq 0}} \frac{\omega_2(x,h)}{\|h\|_2^2} = 0$$

$$\nabla f(\vec{x}) = 0 \implies f(\vec{x}+h) - f(\vec{x}) = \frac{1}{2}(H_f(\vec{x})h, h)_2 + \omega_2(\vec{x}, h)$$

Ist $H_f(\vec{x})$ positiv definit, so gilt

$$(H_f(\vec{x})h, h)_2 \geq \lambda \|h\|_2^2, h \in \mathbb{R}^n$$

wobei λ der kleinste Eigenwert ist.

$$\implies f(\vec{x}+h) - f(\vec{x}) \geq \frac{1}{2}\lambda\|h\|_2^2 + \omega(\vec{x}, h)$$

Für kleines $\|h\|_2 < \sigma, h \neq 0$ ist

$$|\omega_2(\vec{x}, h)| < \frac{1}{2}\lambda\|h\|_2^2$$

und somit

$$f(\vec{x}+h) - f(\vec{x}) > \frac{1}{2}\lambda\|h\|_2^2 - \frac{1}{2}\lambda\|h\|_2^2 = 0$$

$\implies$ $\vec{x}$ ist ein lokales Maximum. Ist $H_f(\vec{x})$ negativ definit $\implies$ $\vec{x}$ ist ein lokales Maximum (analog).

Ist $H_f(\vec{x})$ indefinit $\implies$ $\exists\lambda_+ > 0$ (mit Eigenvektor z_+) und $\exists\lambda_- < 0$ (mit EV z_-)

$$(H_f(\vec{x})z_+, z_+)_2 = \lambda_+\|z_+\|_2^2 > 0$$
$$(H_f(\vec{x})z_-, z_-)_2 = \lambda_-\|z_-\|_2^2 < 0$$

Für genübend kleines $t > 0$ gilt dann

$$f(\vec{x}+tz_+) - f(\vec{x}) > 0 \qquad f(\vec{x}+tz_-) - f(\vec{x}) < 0$$

$\implies$ kein Extremum in $\vec{x}$ □

Beispiel 2.26 1. $f_1(x) = a + x_1^2 + x_2^2$

$$\nabla f_2(x) = (2x_1, 2x_2) = 0 \iff \vec{x}_1 = 0 \wedge \vec{x}_2 = 0$$

$$H_{f_1}(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

positiv definit $\implies$ $\vec{x} = 0$ ist Minimum.

2. $f_2(x) = a - x_1^2 - x_2^2$

$$\nabla f_2(x) = (-2x_1, -2x_2) \implies \vec{x} = 0, H_{f_2}(x) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

negativ definit $\implies$ $\vec{x} = 0$ ist Maximum.

Bemerkung Ist die Hesse Matrix in einer Nullstelle das Gradienten semidefinit (des heißt $\exists\lambda_i = 0$), so lassen sich keine allgemeinen Aussagen über lokale Extrema machen.

2.1 Satz über implizite Funktionen und der Umkehrsatz

Problemstellung: $F(x,y) = x^2 + y^2 - 1$. Betrachte $F(x,y) = 0$

$$\implies y(x) = \pm\sqrt{1-x^2}$$

Satz 2.27 (Satz über implizite Funktionen) Sei $U_1 \subseteq \mathbb{R}^n, U_2 \subseteq \mathbb{R}^m$ offene Menge und $F : U_1 \times U_2 \to \mathbb{R}^m, (x,y) \mapsto F(x,y)$ sei eine stetig differenzierbare Funktion. Sei $(a,b) \in U_1 \times U_2$ mit $F(a,b) = 0$. Die $(m \times n)$ Matrix

$$\frac{\partial F}{\partial y} := \left(\frac{\partial F_i}{\partial l_j}\right)_{i,j=1,\ldots,m}$$

in (a,b) invertierbar. Dann gibt es offene Mengen $V_1 \subseteq U_1, V_2 \subseteq U_2$, V_1 Umgebung von a, V_2 Umgebung von b sowie eine eindeutige stetig differenzierbare Funktion $\varphi : V_1 \to V_2$ mit $\varphi(a) = b$ und $F(x,\varphi(x)) = 0 \forall x \in V_1$. (Eindeutigkeit: Ist $(x,y) \in V_1 \times V_2$ mit $F(x,y) = 0 \implies y = \varphi(x)$.)

Beweis Ohne Beschränkung der Allgemeinheit $(a,b) = (0,0)$. Wir setzen

$$B := \frac{\partial F}{\partial y}(0,0) \in \mathrm{GL}(m,\mathbb{R})$$

und betrachten $G : U_1 \times U_2 \to \mathbb{R}^m$ durch $G(x,y) := y - B^{-1}F(x,y)$ definiert. G ist stetig differenzierbar, weil F es ist. Dann gilt

$$\frac{\partial G}{\partial y} = \not{\Vdash} - B^{-1}\frac{\partial F}{\partial y}(x,y)$$

mit

$$\frac{\partial G}{\partial y}(0,0) = \not{\Vdash} - B^{-1}B = 0$$

Es gilt: $F(x,y) = 0 \iff G(x,y)y$.
Aufgrund der Stetigkeit von $\frac{\partial G}{\partial y}$ gibt es $W_1 \subseteq U_1, W_2 \subseteq U_2$ (jeweils um 0), sodass

$$\left\|\frac{\partial G}{\partial y}\right\|_2 \leq \frac{1}{2} \forall (x,y) \in W_1 \times W_2$$

Wähle $r > 0$, sodass $V_2 := \{y \in \mathbb{R}^n \mid \|y\|_2 \leq r\} \subseteq W_2$ und da $G(0,0) = 0$ gibt es offene Umgebung $V_1 \subset W_1$, sodass

$$\sup_{x \in V_1} \|G(x,0)\|_2 =: \varepsilon \leq \frac{r}{2}$$

Es gilt für alle $x \in V_1$ und $y, \eta \in V_2$:

$$\|G(x,y) - G(x,y)\| \leq \frac{1}{2}\|y - \eta\|$$

Ferner gilt

$$\begin{aligned} \|G(x,y)\| &\leq \|G(x,y) - G(x,0)\| + \|G(x,0)\| \\ &\leq \frac{1}{2}\|y\| + \frac{r}{2} \leq \frac{r}{2} + \frac{r}{2} = r \end{aligned}$$

Die Abbildung $y \mapsto G(x,y)$ bildet V_2 in sich selbst ab und ist eine Kontraktion. Also existiert ein eindeutiger Fixpunkt y nach Banachschem Fixpunktsatz sodass $G(x,y) = y$ beziehungsweise $y = \varphi(x), F(x,\varphi(x)) = 0$. Wir setzen

$$A := \{\varphi \in C_b(V_1, \mathbb{R}^m) \mid \|\varphi\|_\infty \leq r\} = \{\varphi \in C_b(V_1, \mathbb{R}^m) \mid \varphi(V_1) \subset V_2\}$$

Definiere $\Phi : A \to A, \varphi \mapsto G(x, \varphi(x))$.

$$\begin{aligned} \|\Phi(\varphi_1) - \Phi(\varphi_2)\|_\infty &= \sup_{x \in V_1} \|G(x,\varphi_1(x)) - G(x,\varphi_2(x))\| \leq \frac{1}{2} \sup_{x \in V_1} \|\varphi_1(x) - \varphi_2(x)\| \\ &= \frac{1}{2}\|\varphi_1 - \varphi_2\|_\infty \end{aligned}$$

$\implies$ es existiert ein eindeutiges $\varphi \in C_b(V_1, \mathbb{R}^m)$ mit $\Phi(\varphi) = \varphi \iff G(x,\varphi(x)) = \varphi(x)$. Nach eventueller Verkleinerung von V_1 könne wir annehmen, dass $\frac{\partial F}{\partial y}$ in jedem Punkt $(x, (\varphi(x))), x \in V_1$ invertierbar ist. Wir zeigen de Differenziebarkeit von φ nur in 0.

$$A := \frac{\partial F}{\partial x}(0,0) \in M(m \times n, \mathbb{R}), \quad B := \frac{\partial F}{\partial y}(0,0) \in \mathrm{GL}(m, \mathbb{R})$$

Aus der Differenzierbarkeit von F in $(0,0)$ folgt: $F(x,y) = Ax + By + \omega(x,y)$. Nun gilt $F(x,\varphi(x)) = 0 \forall x \in V_1$, das heißt

$$\varphi(x) = -B^{-1}Ax - B^{-1}\omega(x,\varphi(x))$$

Es muss also gezeigt werden, dass $\omega(x,\varphi(x)) = \wr(\|(x,\varphi(x))\|)$. Zeige dazu, dass es eine Umgebung $V_1 \subset V_1$ von 0 gibt und eine Konstate $K > 0$, sodass

$$\|\varphi(x)\| \leq K\|x\| \forall x \in V_1' \quad p_1 := \left\|B^{-1}A\right\| \quad c_2 := \left\|B^{-1}\right\|$$

und wegen $\omega(x,y) = \prime(\|x,y\|)$ gibt es zu $\varepsilon := 1/(2c_2)$ eine Umgebung $V' \subset V_1 \times V_2$ von $0, 0$, sodass

$$\|\omega(x,y)\| = \varepsilon\|(x,y)\| \leq \frac{1}{2c_2}(\|x\| + \|y\|) \forall (x,y) \in V'$$

Wegen der Stetigkeit von φ gibt es eine Nullumgebung $V_1' \subset V_1$, sodass der Graph $\varphi \mid_{V_1'}$ ganz in V' enthalten ist. Dahit gilt

$$\|\omega(x,\varphi(x))\| \leq \frac{1}{2c_2}(\|x\| + \|\varphi(x)\|)$$

Außerdem gilt

$$\begin{aligned} \|\varphi(x)\| &\leq c_1\|x\| + c_2\|\omega(x,\varphi(x))\| \\ &leq\left(c_1 + \frac{1}{2}\right)\|x\| + \frac{1}{2}\|\varphi(x)\| \end{aligned}$$

$$\implies \|\varphi(x)\| \leq \underbrace{2\left(c_1 + \frac{1}{2}\right)}_{=:K}\|x\| \qquad \square$$

Beispiel 2.28 $F(x,y) = x^2 + y^2 - 1 = 0 \implies D_y F = 2y$. Wir können demnach in einer Umgebung von $\left(\hat{x}^2, \hat{y}^2\right)$, $\hat{x}^2 + \hat{y}^2 - 1 = 0$ mit $\hat{y} \neq 0$ eindeutig nach y auflösen und erhalten

$$y = \pm\sqrt{1 - x^2}$$

Definition 2.29 (2.27) Sei $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}^n$ heißt **regulär** in einem Punkt $\hat{x} \in D$, wenn f in einer Umgebung $K_\delta(\hat{x}) \subset D$ von $\hat{x}$ stetig differenzierbar und die Jacobi-Matrix J_f regulär ist. (invertierbar). f heißt regulär in D, wenn f in jedem Punkt regulär ist.

Satz 2.30 (Satz von der Umkehrabbildung) Sei $D \subset \mathbb{R}^n$ offen, $f : D \to \mathbb{R}^n$ regulär in einem Punkt $\hat{x} \in D$. Dann gibt es eine offene Umgebung $V(\hat{x}) \subset D$, die von f bijektiv auf eine offene Umgebung $U(\hat{y}) \subset \mathbb{R}^n$ ($\hat{y} = f(\hat{x})$) abgebildet wird. Die Umkehrabbildung ist ebenfalls regulär in $\hat{y}$. $f^{-1} : U(\hat{y}) \to V(\hat{x})$. Für die Funktionalmatrix und -determinante gilt:

$$J_{f^{-1}}(\hat{y}) = (J_f(\hat{x}))^{-1}, \quad \det J_{f^{-1}}(\hat{y}) = \frac{1}{\det J_f(\hat{x})}$$

Beweis Sei $\hat{x} \in D$ und definiere $\hat{y} := f(\hat{x})$. Betrachte $F : \mathbb{R}^n \times D \to \mathbb{R}^n$, $F(x,y) = y - f(x)$ und offenbar gilt $F(\hat{y}, \hat{x}) = 0$ und $D_x F(y,x) = -J_f(x)$ und damit regulär in $\hat{x}$. Nach dem Satz über implizite Funktionen existieren Umgebungen $U(\hat{y})$ und $U(\hat{x})$, sowie eine eindeutige, stetige differenzierbare Funktion $\varphi : U(\hat{y}) \to U(\hat{x})$ sodass $0 = F(y, \varphi(y)) = y - f(\varphi(y)), y \in U(\hat{y})$. Das bedeutet zu jedem $y \in U(\hat{y})$ kann man genau ein $x = \varphi(y) \in U(\hat{x})$ finden mit $y = f(x)$. Wir setzen

$$V(\hat{x}) := U(\hat{x}) \cap f^{-1}(U(\hat{y})) = \{x \in U(\hat{x}) \mid f(x) \in U(\hat{y})\}$$

$V(\hat{x})$ offen. Fermer wird $V(\hat{x})$ bijktiv von f abgebildet mit zugehörgen Umkehrabbildung $f^{-1} = \varphi$. Wegen $J_{f \circ f^{-1}} = J_{\text{id}} = I$ und der Ketteregel gilt

$$J_f(x) \cdot J_{f^{-1}}(f(x)) = I \implies J_{f^{-1}}(f(x)) = (J_f(x))^{-1} \qquad \square$$

Beispiel 2.31 Transformation der Polarkoordinaten auf kartesische Koordinaten. Polarkoodinaten: $(r, \theta) \to$ kartesische Koordinaten (x_1, x_2).

$$(x_1, x_2) = f(r, \theta) := (r\cos\theta, r\sin\theta) \qquad f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}^2$$

$$J_f(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \quad \det J_f(r,\theta) = r > 0$$

f ist also auf $D = \mathbb{R}_+ \times \mathbb{R}$ regulär. Nach dem Satz über Umkherabbildung ist f also überall in D lokal umkehrbar

$$J_{f^{-1}}(x_1, x_2) = J_f(r,\theta)^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -r^{-1}\sin\theta & r^{-1}\cos\theta \end{pmatrix}$$

Umrechnung in die Variablen $(x_1, x_2) = (r\cos\theta, r\sin\theta)$ liefert

$$r = \sqrt{x_1^2 + x_2^2}, \cos\theta = \frac{x_1}{r}, \sin\theta = \frac{x_2}{r}$$

$$J_{f^{-1}}(x_1, x_2) = \frac{1}{\sqrt{x_1^2 + x_2^2}} \begin{pmatrix} x_1\sqrt{x_1^2 + x_2^2} & x_1\sqrt{x_1^2 + x_2^2} \\ -x_2 & x_1 \end{pmatrix}$$

Wir haben bekommen die Jacobi-Matrix von f^{-1} ohne f^{-1} explizit zu berechnen. Wir berechnen jetzt die f^{-1} $f : U \to V$ wit $U := \mathbb{R}_+ \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), V := \mathbb{R}_+ \times \mathbb{R}$ ist bijktiv

$$f^{-1}(x_1, x_2)\left(\sqrt{x_1^2 + x_2^2}, \arctan\left(\frac{x_2}{x_1}\right)\right)$$

2.2 Extremalaufgaben mit Nebenbedingungen

Sei $f : D \to \mathbb{R}$ und $g : D \to \mathbb{R}$ differenziebare Funktionen auf einer offenen Meng $D \subset \mathbb{R}^n$. Wir suchen $\hat{x} \in D$, sodass

$$f(\hat{x}) = \inf\{f(x) \mid x \in U(\hat{x}), g(\hat{x}) = 0\}$$

für eine Umgebung $U(\hat{x})$ von $\hat{x}$, oder

$$f(\hat{x}) = \sup\{f(x) \mid x \in U(\hat{x}), g(\hat{x}) = 0\}$$

Satz 2.32 (Lagrange Multiplikatoren) Sei $D \subset \mathbb{R}^n$ offen und $f, g : D \to \mathbb{R}$ stetig differenzierbar. Ferner sei $\hat{x} \in D$ ein Punkt, in dem f ein lokales Extremum unter der Nebenbegingung $g(\hat{x}) = 0$ hat. Das heißt

$$f(\hat{x}) = \inf_{x \in U \cap Ng} f(x) \qquad \sup_{x \in U \cap Ng} f(x)$$

wobei $Ng := \{x \in D \mid g(x) = 0\}$. Ist dass $\nabla g(\hat{x}) \neq 0$, so gilt es ein $\hat{\lambda} \in \mathbb{R}$

$$\nabla f(\hat{x}) = \hat{\lambda} \nabla g(\hat{x})$$

Der Parameter $\hat{\lambda}$ ist der sogenannte **Lagrange-Multiplikator**.

Beweis Wegen $\nabla g(\hat{x}) \neq 0$ können wir (nach evetueller Umnummerierung der Koordinaten) annehmen, dass $\partial_n g(\hat{x}) \neq 0$

$$\hat{x} := \left(\hat{x}', \hat{x}_n\right) \in \mathbb{R}^n, \hat{x}' \in \mathbb{R}^{n-1}$$

Nach dem Impliziten Funktionen Satz existieren für die Gleichung $F(x', x_n) := g(x) = 0$ die Umgebungen $U(\hat{x}') \subset \mathbb{R}^{n-1}$ und $U(\hat{x}_n) \subset \mathbb{R}$ mit $U(\hat{x}') \times U(\hat{x}_n) \subset D$ und eine eindeutige Funktion $\varphi : U(\hat{x}') \to U(\hat{x}_n)$ stetig differenzierbar und sodass

$$F\big(x', \varphi(x')\big) = 0 \quad x' \in U(\hat{x})$$

$$Ng \cap (U(\hat{x}_n) \times U(\hat{x})) = \{x \in U(\hat{x}_n) \times U\big(\hat{x}'\big) : x_n = \varphi\big(x'\big)\}$$

Mit Hilfe der Kettenregel bekommen wir

$$\partial_i g(\hat{x}) + \partial_n g(\hat{x}) \partial_i \varphi\big(\hat{x}'\big) = 0 \quad i = 1, \ldots, n-1$$

Da f auf Ng im Punkt $\hat{x}$ ein lokales Extremum hat, hat die Funktion $f(x', \varphi(x'))$ auf $U(\hat{x}')$ ein lokales Extremum.

$$\begin{aligned}
&\implies 0 = \partial_i f(\hat{x}) + \partial_n f(\hat{x}) \partial_i \varphi(\hat{x}) \quad i = 1, \dots, n-1 \\
&\implies \partial_n f(\hat{x}) = \hat{\lambda} \partial_n g(\hat{x}) \qquad \hat{\lambda}_n := \frac{\partial_n f(\hat{x})}{\partial_n g(\hat{x})} \\
&\implies \partial_i f(\hat{x}) = \hat{\lambda} \partial_i g(\hat{x}) \quad i = 1, \dots, n \\
&\implies \nabla f(\hat{x}) = \hat{\lambda} \nabla g(\hat{x})
\end{aligned}$$

□

Bemerkung Jedes lokale Minimum $\vec{x}$ der Funktion f unter der Nebenbedingung $g(\hat{x}) = 0$ korrespondiert zu einem sogenanntem „stationären Punkt der Lagrange Funktion“

$$\mathcal{L}(x, \lambda) := f(x) - \lambda g(x) \quad (x, \lambda) \in D \times \mathbb{R}$$

$$\nabla_{x,\lambda} \mathcal{L}(\hat{x}, \hat{y}) = \begin{pmatrix} \nabla_x f(\hat{x}) - \hat{\lambda} \nabla_x g(\hat{x}) \\ g(\hat{x}) \end{pmatrix} = 0$$

Beispiel 2.33 $f(x) := (x_1 \cdot \ldots \cdot x_n)^2, f : \mathbb{R}^n \to \mathbb{R}$. Wir suchen das Maximum von f auf der Sphäre $S_1 = \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$ das heißt

$$g(x) := \|x\|_2 - 1 = \sum_{i=1}^{n} x_1^2 - 1$$

Nebenbedingung: $g(x) = 0$. $s \subset \mathbb{R}^n$ kompakt $\implies$ f nimmt auf S_1 sein Maximum und Minimum an.

$$|_{x \in S_1} f(x) = 0 \qquad \max_{x \in S_1} f(x) > 0$$

Ferner $\nabla g(x) = 2x \neq 0$ auf S_1. Nach dem Satz 2.30 sind die Extremalpunkte die Lösungen $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}$ vom Gleichungssystem

$$\partial_i f(x) = \lambda \partial_i g(x) \quad i = 1, \dots, n$$

$$\implies 2(x_1 \cdot \ldots \cdot x_n)^2 = 2\lambda x_i$$

$$\implies (x_1 \cdot \ldots \cdot x_n)^2 = \lambda x_i^2 \quad i = 1, \dots, n$$

Weil $x_i \neq 0$ im Maximum $\implies$ $\lambda \neq 0$

$$\implies \sum_{i=1}^{n} (x_1 \cdot \ldots \cdot x_n)^2 = \lambda \sum_{i=1}^{n} x_i^2 = \lambda$$

$$\implies n(x_1 \cdot \ldots \cdot x_n)^2 = \lambda$$

$$\implies x_i^2 = \frac{1}{n} \quad i = 1, \dots, n$$

3 Gewöhnliche Differentialgleichungen

Grundbegriffe
Zu einer gegebenen Funktion $f : \mathbb{R} \to \mathbb{R}$ suchen wir eine differenzierbare Funktion $x : \mathbb{R} \to \mathbb{R}$, deren Ableitung durch $f(\cdot)$ beschrieben wird. Wir suchen also eine Funktion sodass

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(t)\forall t \in \mathbb{R}$$

Bewerkung zur Notation

$$x' = f$$
$$\dot{x} = f$$

Beispiel 3.1 Für gegebene Geschwindigkeit (in Ableitung von Zeit) suchen wir die Position des Körpers auf einer festen eindimensionalen Achse.

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(t)\forall t \in \mathbb{R}$$

Wir müssen noch die Position zu irgendwelchen Zeitpunkt nennen. Das heißt die Lösung ist nicht eindeutig solag wir keinen Wert $x(t_0) \in \mathbb{R}$ festlegen. Das Problem

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(t)$$
$$x(t_0) = x_0$$

lässt sich lösen wenn $f : \mathbb{R} \to \mathbb{R}$ stetig ist. Dann besagt nämlich des Hauptsatz der Integralrechnung, dass

$$x(\cdot) : \mathbb{R} \to \mathbb{R}, t \to x_0 + \int_{t_0}^{t} f(s)\mathrm{d}s$$

differenzierbar ist und die Ableitung $f(t)$ begrenzt ist.

Ziel:

- Existenz von Lösung
- Eindeutigkeit von Lösung
- Verhalten

Beispiel 3.2

$$\frac{\mathrm{d}x}{\mathrm{d}t} = rx$$

r: Konstante. In $t_0 = 0 : x(0) = x_0$

$$x(\cdot) = c \cdot e^{rt}$$
$$x_0 = x(0) = c$$
$$\implies x(t) = x_0 e^{rt}$$

Definition 3.3 Gegeben seien nicht leere Teilmenge $D \subset \mathbb{R} \times \mathbb{R}^m$ und eine Funktion $f : D \to \mathbb{R}^m$. Dann nennt man

$$x' = f(\cdot, x)$$

eine explizite Gewöhnliche Differenzialgleichung (GDGL)(ODE - ordinary differential equation) 1. Ordnung. Im Fall $m = 0$ wird die Gleichung als **Skalar** bezeichnet. Eine solche Differentialgleichung heißt **autonom** falls f nicht explizit von t abhängt (sonst: **nichtautonom**). Für $m > 1$ bekommen wir ein System von Gewöhnlichen Differentialgleichungen. Eine Funktion $x : I \to \mathbb{R}^m, I \subset \mathbb{R}$, heißt eine Lösung der Differentialgleichung, wenn

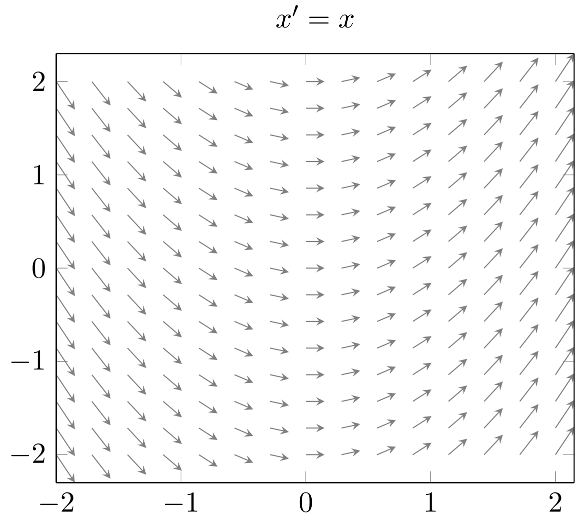
1. $\forall t \in R \subset \mathbb{R}$ liegt $(t, x(t)) \in D$
2. $x(\cdot)$ ist differenzierbar, das heißt

$$\forall t \in I \exists x'(t) = \lim_{\substack{h \to 0 \\ t+h \in I}} \frac{x(t+h) - x(t)}{h} \in \mathbb{R}^m$$

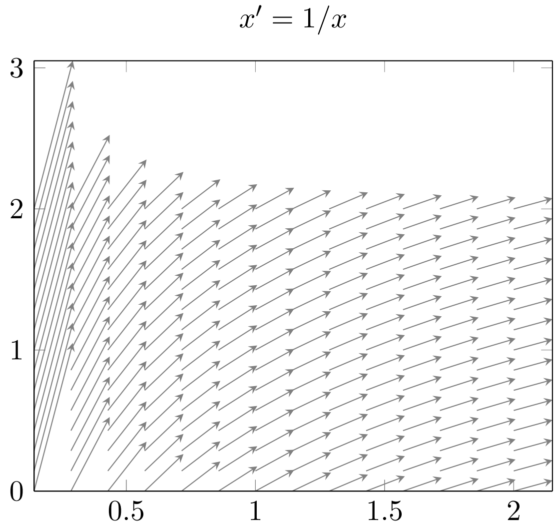
3. $\forall t \in I$ gilt $x'(t) = f(t, x(t))$

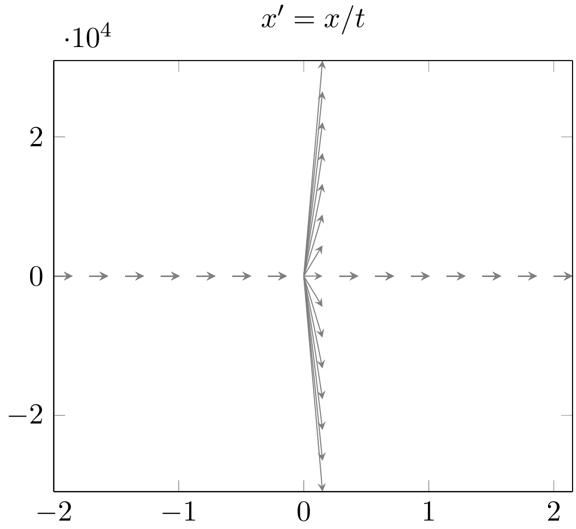
Bei **Anfangswertproblemen** zu dieser Gewöhnlichen Differentialgleichung ist noch ein Tupel $(t_0, x_0) \in D$ gegeben und gesucht ist eine Funktion die Bedingungen 1. bis 3. und $x(t_0) = x_0$ erfüllt.

Konstruktion von Lösungen
Geometrische Interpretation: Eine skalare Gleichung $x' = f(t, x)$ bestimmt ein **Richtungsfeld**, das heißt $\forall (t, x) \in \mathbb{R}^2$ wird durch $x' = f(t, x)$ eine **Steigung** gegeben. Gesucht sind $x(t)$ deren Graph $G(x) = \{(t, x)\}$ in jedem Punkt die vorgegebene Steigung hat. In einfachen Fällen kann mit



aus ihrem Richtungsfeld die mögliche Lösung ergeben.





Methode der Trennung der Variablen
Wir betrachten die separable Differentialgleichung

$$x' = f(x,t) = a(t)g(x)$$

Sei x einen Lösung. Falls $g(t) \neq 0$ bekommen

$$\int_{t_0}^{t} \frac{x'(s)}{g(x(s))} \mathrm{d}s = \int_{t_0}^{t} a(s) \mathrm{d}s$$

Mit Hilge die Substitution $z := x(s)$ ergibt sich (mit $\frac{\mathrm{d}z}{\mathrm{d}x} = x'(s)$)

$$\int_{x_0}^{x(t)} \frac{1}{g(z)} \mathrm{d}z = \int_{t_0}^{t} a(s) \mathrm{d}s$$

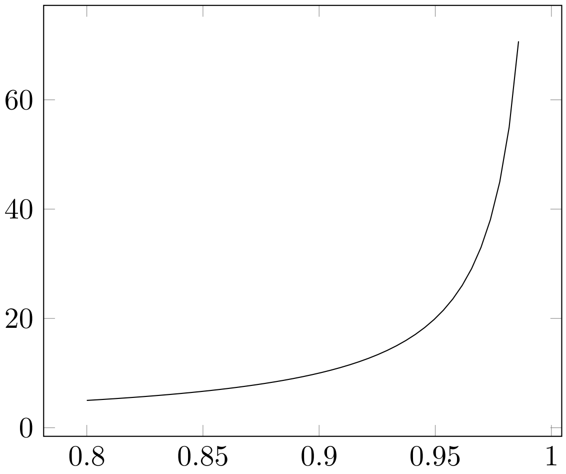
Beispiel 3.4 (3.4)

$$\begin{cases} x' = x^2 \\ x(t_0) = x_0 \end{cases}$$

$$\begin{aligned} \int_{x_0}^{x(t)} \frac{\mathrm{d}z}{z^2} &= \int_{t_0}^{t} 1 \mathrm{d}s \\ -\frac{1}{2}\Big|_{x_0}^{x(t)} &= t - t_0 \\ t - t_0 &= \frac{1}{x_0} - \frac{1}{x(t)} \\ x(t) &= \frac{x_0}{1 - x_0(t - t_0)} \end{aligned}$$

Falls $t_0 = 0, x(0) = 1$:

$$x(t) = \frac{1}{1-t}$$



Dies ist keine **globale** ($\forall t \in \mathbb{R}_+$) Lösung, da man $x(t)$ mich nach $t = t^*$ vorsetzen kann.

Methode der Variation die Konstanten
Wir betrachten Differentialgleichung $x' = a(t)x(t) + b(t), t \in I = [t_0, t_0 + \tau] \subset \mathbb{R}$ mit stetigen Funktionen $a, b : I \to \mathbb{R}$ Die zugehörige homogene Differentialgleichung $y' = ay$ hat eine Lösung die Form

$$y(t) = c \exp \int_{t_0}^{t} a(s) \mathrm{d}s, \quad c \in \mathbb{R}$$

(Seperation der Variablen). Sei $y(t)$ eine Lösung mit $c = 1$. Zur Bestimmung einer Lösung der **inhomogenen Differentialgleichung** wird c als Funktion von t angesetzt. Ansatz: $x(t) = c(t)y(t)$

$$\begin{aligned} \implies x'(t) &= c'(t)y(t) + c(t)y'(t) \\ &= c'(t) \exp \int_{t_0}^{t} a(s)\mathrm{d}s + a(t)x(t) \\ &\overset{?}{=} a(t)x(t) + b(t) \iff c'(t) \exp\left(\int_{t_0}^{t} a(x)\mathrm{d}s\right) = b(t) \end{aligned}$$

Wir bekommen

$$c(t) = \int_{t_0}^{t} \exp\left(-\int_{t_0}^{\tau} a(s)\mathrm{d}s\right) b(\tau)\mathrm{d}\tau + r$$

mit einer freien Konstante $r \in \mathbb{R}$. Damit wird

$$x(t) = \exp\left(\int_{t_0}^{t} a(s)\mathrm{d}s\right) \int_{t_0}^{t} \exp\left(-\int_{t_0}^{t} a(s)\mathrm{d}s\right) b(\tau)\mathrm{d}\tau + r \exp\left(\int_{t_0}^{t} a(s)\mathrm{d}s\right)$$

Durch die Wahl der Konstanten $r = x_0$ ergibt sich $x(t_0) = x_0$

$$\implies x(t) = \exp\left(\int_{t_0}^{t} a(s)\mathrm{d}s\right)[x_0 + \int_{t_0}^{t} \exp\left(-\int_{t_0}^{\tau} a(s)\mathrm{d}s\right) b(\tau)\mathrm{d}\tau]$$

Beispiel 3.5

$$x' = ax(t) + b(t), \quad x(0) = x_0$$

a: Konstante

$$\implies x(t) = x_0 e^{at} + \int_{t_0}^{t} e^{a(t-\tau)} b(\tau) \mathrm{d}\tau$$

$$\big(c(t)e^{at}\big)' = c'e^{at} + ce^{at}a = ae^{at} + b$$

$$\implies c' = b(t)e^{-at}$$
$$c(t) = \int_{t_0}^{t} b(\tau)e^{-a\tau} \mathrm{d}\tau$$
$$x(t) = x_0 e^{at} + c(t)e^{at}$$

Anfangswertproblem

$$x' = f(t,x)$$
$$x(0) = x_0$$

Existenzsatz von Peano

Satz 3.6 (Peano) Die Funktion $f(t,x)$ sei **stetig** auf einem Zylinder

$$D = \{(t,x) \in \mathbb{R}^1 \times \mathbb{R}^m \mid |t - t_0| \leq \alpha, \|x - x_0\| \leq \beta\}$$

Dann existiert eine Lösung $x(t)$ auf dem Intervall $I := [t_0 - T, t_0 + T]$ wobei

$$T := \min\left(\alpha, \frac{\beta}{M}\right), \quad M := \max_{(t,x)\in D} \|f(t,x)\|$$

Beweis Mit Hilfe der Differenzenmethode konstruieren wir eine Folge von stückweise linearer Funktionen, welche eine Teilfolge besitzt, die (gleichmäßig) gegen eine Lösung des Anfangswertproblems konvergiert. Ohne Beschränkung der Allgemeinheit genügt es das Halbintervall $I = [t_0, t_0 + T]$ zu betrachten. Zu einem Schrittweitenparameter $h > 0$ wird eine äquidistante Unterteilung des I gewählt.

$$t_0 < \cdots < t_N = t_0 + T \qquad h = t_n - t_{n-1}$$

Ausgehend von $x_0^h := x_0$ erzeugt dann das sogenannte Eulersche Polygonzugverfahren Werte für x_n^h durch

$$x_n^h = x_{n-1}^h + hf\Big(t_{n-1}, x_{n-1}^h\Big), n \geq 0$$

Diese diskreten Funktionswirte wirden linear interpoliert zu einer stetigen Funktion:

$$x_n^h(t) := x_{n-1}^h + (t - t_{n-1})f\Big(t_{n-1}, x_{n-1}^h\Big)$$

Schritt 1: Wir zeigen Graph $(x^n) \subset D$.
Sei $\left(t, x^h(t)\right) \in D$ für $t_0 \leq t \leq t_{k-1}$. Es gilt

$$\left(x(t)^h\right)' = f\left(t_{k-1}, x_{k-1}^k\right), t \in [t_{k-1}, t_k]$$

Nach Konstruktion gilt dann für $t \in [t_{k-1}, t_k]$

$$x^h(t) - x_0 = x^h(t) - x_{k-1}^h + \sum_{i=1}^{k-1}\left(x_i^h - x_{i-1}^h\right)$$

$$= (t - t_{k-1}) f\left(t_{k-1}, x_{k-1}^h\right) + h \sum_{i=1}^{k-1} f\left(t_{i-1}, x_{i-1}^h\right)$$

$$\implies \left\|x^h(t) - x_0\right\| \leq (t - t_{k-1})M + (t_{k-1} - t_0)M = (t - t_0)M$$

Also $\left(t, x^h(t)\right) \in D$ für $0 \leq t \leq t_k$
Schritt 2: Wir zeigen gleichgradige Stetigkeit
Seien dazu $t, \tilde{t} \in I, \tilde{t} \leq t$ mit $t \in [t_{k-1}, t_k], \tilde{t} \in [t_{j-1}, t_j]$ für gewisse $t_j \leq t_k$. Im Fall $t, \tilde{t} \in [t_{k-1}, t_k]$ gilt

$$x^h(t) - x^h(\tilde{t}) = \left(t - \tilde{t}\right) f\left(t_{k-1}, x^h(t_{k-1})\right)$$

$$\implies \left\|x^h(t) - x^h(\tilde{t})\right\| \leq M\left(t - \tilde{T}\right)$$

Für $t_j < t_k$

$$x^h(t) - x^h(\tilde{t}) = (t - t_{k-1}) f\left(t_{k-1}, x_{k-1}^h\right) + h \sum_{i=j}^{k-1} f\left(t_{i-1}, x_{i-1}^h\right) + \left(t_{j-1} - \tilde{t}\right) f\left(t_{j-1}, x_{j-1}^h\right)$$

$$= (t - t_{k-1}) f\left(t_{k-1}, x_{k-1}^h\right) + h \sum_{i=j+1}^{k-1} f\left(t_{i-1}, x_{i-1}^h\right) + \left(h + t_{j-1} - \tilde{t}\right) f\left(t_{j-1}, x_{j-1}^h\right)$$

$$\implies \left\|x^h(t) - x^h(\tilde{t})\right\| \leq M\left((t - t_{k-1}) + (t_{k-1} - t_j) + \left(t_j - \tilde{t}\right)\right) \leq M\left|t - \tilde{t}\right|$$

Also $x_{h>0}^h$ gleichgradig stetig. Die Funktionen sind auch gleichmäßig beschränkt:

$$\left\|x^h(t)\right\| \leq \left\|x^h(t) - x_0\right\| + \|x_0\| \leq MT + \|x_0\|, t \in (t_0, t_0 + T)$$

Arzela-Ascoli Satz: $\exists$ eine Nullfolge $(h_i)_{i \in \mathbb{N}}$ und stetiges $x(t)$ sodass

$$\left\|x^{h_i}(t) - x(t)\right\| \xrightarrow{i \to \infty} 0$$

und Graph $(x) \subset D$ $\square$