

6 Tag 6

Topologische Grundbegriffe im $\mathbf{R}^n$

Sei $n \in \mathbf{N}$ und $D \subseteq \mathbf{R}^n$.

- D heißt *beschränkt*, wenn es ein $S \in \mathbf{R}$ gibt mit $|a| \leq S$ für alle $a \in D$.
- D heißt *konvex*, wenn für zwei Punkte $a, b \in D$ die gesamte Verbindungsstrecke $\overline{ab} = \{a + t(b - a) \in \mathbf{R}^n \mid t \in [0, 1]\}$ in D enthalten ist.
- $\overline{D} = \{x \in \mathbf{R}^n \mid \text{Es gibt eine Folge } (a_k)_k \text{ in } D \text{ mit } \lim_{k \rightarrow \infty} a_k = x\}$ heißt der *Abschluss* von D und D heißt *abgeschlossen*, wenn $\overline{D} = D$ gilt.
- $D^\circ = \mathbf{R}^n \setminus \overline{\mathbf{R}^n \setminus D}$ heißt das *Innere* von D . Ein Punkt $a \in D^\circ$ heißt *innerer Punkt* von D . Für $a \in \mathbf{R}^n$ gilt $a \in D^\circ$ genau dann, wenn es ein $\varepsilon \in \mathbf{R}^+$ mit $U_\varepsilon(a) \subseteq D$ gibt, wobei $U_\varepsilon(a) = \{x \in \mathbf{R}^n \mid |x - a| < \varepsilon\}$ die ε -*Umgebung* von a ist. D heißt *offen*, wenn $D = D^\circ$ gilt. $\partial D = \overline{D} \setminus D^\circ$ ist der *Rand* von D .
- D heißt *kompakt*, wenn D beschränkt und abgeschlossen ist.
Beispiele kompakter Mengen sind Kreise, Ellipsen, Kugeln und Ellipsoide.

6.1 Stetigkeit

- *Weierstraßscher Satz vom Minimum und Maximum*: Sei $\emptyset \neq D \subseteq \mathbf{R}^n$ kompakt und $f: D \rightarrow \mathbf{R}$ stetig. Dann nimmt f auf D ein (globales/absolutes) Maximum und ein (globales/absolutes) Minimum an, d.h. es gibt $a, b \in D$ mit $f(a) \leq f(x) \leq f(b)$ für alle $x \in D$.
- *Nullstellensatz von Bolzano*: Seien $a, b \in \mathbf{R}$ mit $a < b$, $f: [a, b] \rightarrow \mathbf{R}$ stetig und es gelte $f(a) \cdot f(b) < 0$. Dann gibt es ein $\xi \in (a, b)$ mit $f(\xi) = 0$.
Mit der *Intervallhalbierungsmethode* können wir ein solches ξ approximieren: Man erhält eine gegen eine Nullstelle konvergente Folge $(a_k)_k$, indem man $a_1 = a$, $b_1 = b$ setzt und für $k \in \mathbf{N}$ rekursiv a_{k+1}, b_{k+1} wie folgt definiert:
 - * Falls $f(a_k) \cdot f\left(\frac{a_k+b_k}{2}\right) \leq 0$ gilt: Setze $a_{k+1} = a_k$ und $b_{k+1} = \frac{a_k+b_k}{2}$.
 - * Falls $f(a_k) \cdot f\left(\frac{a_k+b_k}{2}\right) > 0$ gilt: Setze $a_{k+1} = \frac{a_k+b_k}{2}$ und $b_{k+1} = b_k$.

6.2 Differentiation in einer und in mehreren Veränderlichen

Ableitungen

Sei $n \in \mathbf{N}$, $D \subseteq \mathbf{R}^n$, $f: D \rightarrow \mathbf{R}$, $a \in D^\circ$ und $v \in \mathbf{R}^n$.

Wenn der Differentialquotient genannte Grenzwert

$$\partial_v f(a) = \lim_{h \rightarrow 0} \frac{f(a + hv) - f(a)}{h}$$

existiert, so nennt man ihn die *Ableitung längs v* von f in a . Falls $|v| = 1$ gilt, so heißt $\partial_v f(a)$ eine *Richtungsableitung*.

Sei $j \in \{1, \dots, n\}$. Für den j -ten Standardbasisvektor $e_j \in \mathbf{R}^n$ heißt $\partial_{e_j} f(a)$ die *partielle Ableitung* von f in a nach x_j . Alternative Notationen:

$$\frac{\partial f}{\partial x_j}(a), \quad \frac{\partial}{\partial x_j} f(a), \quad f_{x_j}(a), \quad \frac{\partial}{\partial x_j} f(x) \Big|_{x=a} \quad \text{und} \quad \partial_j f(a).$$

Entsprechend heißt f in a partiell nach x_j differenzierbar, falls der Grenzwert $\lim_{h \rightarrow 0} \frac{f(a + he_j) - f(a)}{h}$ existiert. Ist f in a partiell nach allen x_j differenzierbar, so heißt

$$\nabla f(a) = \text{grad } f(a) = \begin{pmatrix} \partial_1 f(a) \\ \partial_2 f(a) \\ \vdots \\ \partial_n f(a) \end{pmatrix} \in \mathbf{R}^n$$

der *Gradient* von f in a .

Für $n = 1$ heißt f in a differenzierbar, falls der Grenzwert $\lim_{h \rightarrow 0} \frac{f(a + h) - f(a)}{h}$ existiert und diesenfalls heißt $f'(a) = \frac{d}{dx} f(a) = \partial_{e_1} f(a)$ die *Ableitung* von f in a . Ferner ist f dann stetig in a .

Sei nun D offen. Falls f in allen Punkten $a \in D$ partiell nach x_j differenzierbar ist, so erhalten wir eine Funktion

$$f_{x_j}: D \rightarrow \mathbf{R}: x \mapsto f_{x_j}(a),$$

welche die *partielle Ableitung* von f nach x_j genannt wird.

Ist f in allen $x \in D$ partiell nach allen x_j differenzierbar und sind f sowie alle f_{x_j} stetig, so heißt f stetig (partiell) differenzierbar. Die Menge aller stetig partiell differenzierbaren Funktionen auf D wird mit $C^1(D)$ bezeichnet.

Gilt $f \in C^1(D)$ und $a \in D$, so folgt $\partial_v f(a) = \nabla f(a) \cdot v$ für alle $v \in \mathbf{R}^n$.

Sei nun $f: D \rightarrow \mathbf{R}^m$ eine (vektorwertige) Funktion. Sind alle Komponentenfunktionen $f_k: D \rightarrow \mathbf{R}: x \mapsto f(x)_k$ mit $1 \leq k \leq m$ in a partiell nach allen x_j mit $1 \leq j \leq n$ differenzierbar, so heißt die $(m \times n)$ -Matrix

$$Jf(a) = \begin{pmatrix} \nabla f_1(a)^\top \\ \nabla f_2(a)^\top \\ \vdots \\ \nabla f_m(a)^\top \end{pmatrix} = \begin{pmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) & \cdots & \partial_n f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) & \cdots & \partial_n f_2(a) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 f_m(a) & \partial_2 f_m(a) & \cdots & \partial_n f_m(a) \end{pmatrix}$$

die *Jacobimatrix* von f in a .

Ist D offen und gilt $f_k \in C^1(D)$ für alle $k \in \{1, \dots, m\}$, so ist f in $a \in D$ *total differenzierbar*, d.h. es gilt

$$\lim_{x \rightarrow a} \frac{|f(x) - f(a) - Jf(a)(x - a)|}{|x - a|} = 0.$$

Ableitungsregeln

Seien $n, m, \ell \in \mathbf{N}$.

- Sei $D \subseteq \mathbf{R}^n$ offen, $f, g: D \rightarrow \mathbf{R}^m$ und $\alpha, \beta \in \mathbf{R}$. Sind f und g in $a \in D$ total differenzierbar, so ist auch $\alpha f + \beta g$ in a total differenzierbar mit

$$J(\alpha f + \beta g)(a) = \alpha Jf(a) + \beta Jg(a).$$

- Sei $D \subseteq \mathbf{R}^n$ offen, $f: D \rightarrow E$ mit $E \subseteq \mathbf{R}^m$ offen und $g: E \rightarrow \mathbf{R}^\ell$. Ist f in $a \in D$ total differenzierbar und ist g in $b = f(a) \in E$ total differenzierbar, so ist $g \circ f$ in a total differenzierbar mit $J(g \circ f)(a) = Jg(b) \cdot Jf(a)$. (*Kettenregel*)
- Sei $D \subseteq \mathbf{R}^n$ offen und $f, g \in C^1(D)$. Dann gilt auch $f + g \in C^1(D)$ und $f \cdot g \in C^1(D)$. Hat g keine Nullstelle in D , so ist auch $\frac{f}{g} \in C^1(D)$. Ferner gilt für $a \in D$:

* *Summenregel:* $\nabla(f + g)(a) = \nabla f(a) + \nabla g(a)$

* *Produktregel:* $\nabla(f \cdot g)(a) = g(a) \cdot \nabla f(a) + f(a) \cdot \nabla g(a)$.

* *Quotientenregel:* $\nabla \left(\frac{f}{g} \right) (a) = \frac{g(a) \cdot \nabla f(a) - f(a) \cdot \nabla g(a)}{g(a)^2}$

- Seien I und J reelle Intervalle, sowie $f: I \rightarrow J$ stetig und bijektiv mit Umkehrfunktion $f^{-1}: J \rightarrow I$. Ist f in $x_0 \in I$ differenzierbar mit $f'(x_0) \neq 0$, so ist auch f^{-1} in $y_0 = f(x_0)$ differenzierbar mit $(f^{-1})'(y_0) = \frac{1}{f'(x_0)}$.

Ausgewählte Ableitungen

Seien $z, s, n \in \mathbf{N}$.

- Sei $A \in \mathbf{R}^{z \times s}$. Die lineare Abbildung $f: \mathbf{R}^s \rightarrow \mathbf{R}^z: x \mapsto Ax$ ist in allen $a \in \mathbf{R}^s$ total differenzierbar mit $Jf(a) = A$.
- Sei $A \in \mathbf{R}^{n \times n}, t \in \mathbf{R}^n$. Die affine Abbildung $f: \mathbf{R}^n \rightarrow \mathbf{R}^n: x \mapsto Ax + t$ ist in allen $a \in \mathbf{R}^n$ total differenzierbar mit $Jf(a) = A$.
- Reelle Potenzreihen sind innerhalb ihres Konvergenzintervalls differenzierbar mit

$$\frac{d}{dx} \sum_{k=0}^{\infty} a_k (x - x_0)^k = \sum_{k=0}^{\infty} \frac{d}{dx} a_k (x - x_0)^k = \sum_{k=1}^{\infty} a_k k (x - x_0)^{k-1}.$$

- Tabellarische Übersicht von Ableitungen wichtiger reeller Funktionen auf jeweils geeignet gewählten Definitionsbereichen: (mit $a \in \mathbf{R}$ und $b \in \mathbf{R}^+$)

$f(x)$	$f'(x)$	$f(x)$	$f'(x)$
x^a	ax^{a-1}	$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$
e^x	e^x	$\arccos x$	$\frac{-1}{\sqrt{1-x^2}}$
b^x	$\ln b \cdot b^x$	$\arctan x$	$\frac{1}{1+x^2}$
$\ln x $	$\frac{1}{x}$	$\sinh x$	$\cosh x$
$\sin x$	$\cos x$	$\cosh x$	$\sinh x$
$\cos x$	$-\sin x$	$\operatorname{arsinh} x$	$\frac{1}{\sqrt{x^2+1}}$
$\tan x$	$\frac{1}{\cos^2 x} = 1 + \tan^2 x$	$\operatorname{arcosh} x$	$\frac{1}{\sqrt{x^2-1}}$

Höhere Ableitungen

Seien $n \in \mathbf{N}$, $D \subseteq \mathbf{R}^n$ offen und $f \in C^1(D)$. Ist f_{x_j} in allen $x \in D$ partiell nach x_ℓ differenzierbar, so schreiben wir $f_{x_j x_\ell} = (f_{x_j})_{x_\ell}$. Allgemein heißt $f_{x_{j_1} x_{j_2} \dots x_{j_k}}$ für $k \in \mathbf{N}$ eine k -te *partielle Ableitung* von f . f heißt k -mal stetig (partiell) differenzierbar, wenn alle ℓ -ten partiellen Ableitungen für $\ell \in \{0, \dots, k\}$ existieren und stetig sind. (Hierbei ist f als nullte partielle Ableitung zu interpretieren.) Die Menge aller k -mal stetig partiell differenzierbaren Funktion auf D wird mit $C^k(D)$ bezeichnet.

Sei $f \in C^2(D)$ und $a \in D$. Die Matrix

$$Hf(a) = \begin{pmatrix} f_{x_1 x_1}(a) & f_{x_1 x_2}(a) & \cdots & f_{x_1 x_n}(a) \\ f_{x_2 x_1}(a) & f_{x_2 x_2}(a) & \cdots & f_{x_2 x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_n x_1}(a) & f_{x_n x_2}(a) & \cdots & f_{x_n x_n}(a) \end{pmatrix} \in \mathbf{R}^{n \times n}$$

nennt man die *Hessematrix* von f in a . Der Satz von Schwarz besagt, dass diese Matrix stets symmetrisch ist.

Der Satz von Taylor

Sei $n \in \mathbf{N}$, $D \subseteq \mathbf{R}^n$ offen und konvex, $m \in \mathbf{N}_0$, $f \in C^{m+1}(D)$ und $a \in D$. Dann gibt es für jedes $x \in D$ ein $\vartheta \in \overline{ax}$ mit

$$f(x) = T_m(f, x, a) + R_m(f, x, a),$$

wobei

$$T_m(f, x, a) = \sum_{k=0}^m \frac{\partial_{x-a}^k f(a)}{k!} = f(a) + \partial_{x-a} f(a) + \cdots + \frac{\partial_{x-a}^m f(a)}{m!}$$

das *Taylorpolynom* der Stufe m von f um den Entwicklungspunkt a und

$$R_m(f, x, a) = \frac{\partial_{x-a}^{m+1} f(\vartheta)}{(m+1)!}$$

das zugehörige *Restglied nach Lagrange* ist.

- Es gilt $T_1(f, x, a) = f(a) + \partial_{x-a} f(a) = f(a) + \nabla f(a) \cdot (x - a)$ und $T_2(f, x, a) = f(a) + \nabla f(a) \cdot (x - a) + \frac{1}{2}(x - a)^\top \cdot Hf(a) \cdot (x - a)$.
- Für $n = 2$ ist der Graph von $T_1(f, x, a)$ die *Tangentialebene* an der Stelle a an den Graphen von f . Diese erfüllt die Gleichung $x_3 = T_1(f, (x_1, x_2), a)$ in $\mathbf{R}^3$.

- Für $n = 1$ gilt für die Richtungsableitungen $\partial_{x-a}^k f(a) = f^{(k)}(a) \cdot (x - a)^k$. Ist zudem f beliebig oft in a differenzierbar, so heißt

$$T(f, x, a) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x - a)^k = f(a) + f'(a) \cdot (x - a) + \frac{f''(a)}{2} (x - a)^2 + \dots$$

die *Taylorreihe* von f um den Entwicklungspunkt a .

Extrema

Sei $n \in \mathbf{N}$, $D \subseteq \mathbf{R}^n$ und $f: D \rightarrow \mathbf{R}$.

- $a \in D$ heißt lokale Extremalstelle von f , wenn f in a ein lokales (oder relatives) Extremum besitzt, d.h. wenn es ein $\varepsilon \in \mathbf{R}^+$ gibt mit $f(a) \leq f(x)$ für alle $x \in U_\varepsilon(a) \cap D$ (lokales Minimum) oder mit $f(a) \geq f(x)$ für alle $x \in U_\varepsilon(a) \cap D$ (lokales Maximum).
- $a \in D$ heißt *kritische Stelle* von f , wenn $\nabla f(a) = 0$ gilt. In diesem Fall heißt $(a, f(a))$ ein *kritischer Punkt* von f .
- Sei D offen und $f \in C^1(D)$. Ist $a \in D$ eine lokale Extremalstelle von f , so ist a eine kritische Stelle von f .
- Definitheitskriterium: Sei $n \in \mathbf{N}$, $D \subseteq \mathbf{R}^n$ offen und $f \in C^2(D)$. Ferner sei a eine kritische Stelle von f . Dann gilt:
 - * Ist $Hf(a)$ positiv definit, d.h. alle Eigenwerte von $Hf(a)$ sind positiv, so besitzt f in a ein lokales Minimum.
 - * Ist $Hf(a)$ negativ definit, d.h. alle Eigenwerte von $Hf(a)$ sind negativ, so besitzt f in a ein lokales Maximum.
 - * Ist $Hf(a)$ indefinit, d.h. $Hf(a)$ besitzt positive und negative Eigenwerte, so besitzt f in a kein lokales Extremum. Man spricht von einem *Sattelpunkt*.
- Ist D kompakt, so liefert der Weierstraßsche Satz vom Minimum und Maximum die Existenz von (globalen) Extrema. Gegebenenfalls kann der Satz auch auf eine kompakte Teilmenge von D angewendet werden.
- Für $t \in \mathbf{R}$ heißt $N_t = \{x \in D \mid f(x) = t\}$ die Niveaumenge von f zum Niveau t . N_0 ist die Nullstellenmenge von f . Diese kann für $n = 2$ skizziert und die dabei entstehenden Bereiche, in denen $f(x)$ positiv bzw. negativ ist, können

entsprechend markiert werden. (Vorzeichenverteilung)

Treffen sich an einer kritischen Stelle Bereiche mit unterschiedlichen Vorzeichen, so liegt ein Sattelpunkt vor.

Extrema unter Nebenbedingungen

Sei $n \in \mathbf{N}$, $D \subseteq \mathbf{R}^n$ offen und $f, g \in C^1(D)$. Wir suchen Extrema der Funktion f unter der Nebenbedingung $g(x) = 0$, d.h. Extrema der Funktion

$$K \rightarrow \mathbf{R}: x \mapsto f(x)$$

mit $K = \{x \in D \mid g(x) = 0\}$.

- Falls K mit einer geeigneten Funktion $p: I \subseteq \mathbf{R} \rightarrow K$ parametrisiert werden kann, so sind die Extrema von $f \circ p: I \rightarrow \mathbf{R}$ die gesuchten.
- Die *Multiplikatormethode nach Lagrange* liefert ein notwendiges Kriterium: $a \in D$ erfülle $g(a) = 0$ und $\nabla g(a) \neq 0$. Ist $(a, f(a))$ ein relatives Extremum von f unter der Nebenbedingung $g(x) = 0$, dann gibt es ein $\lambda \in \mathbf{R}$, *Lagrange-Multiplikator* genannt, mit

$$\nabla f(a) + \lambda \nabla g(a) = 0.$$

Durch Lösen dieses Gleichungssystems (zusammen mit der Gleichung $g(a)=0$) erhält man Kandidaten für Extremstellen (auch hier kritische Stellen genannt). Zusätzlich muss man noch argumentieren, dass man tatsächlich Extrema gefunden hat, z.B. mit dem Weierstraßschen Satz vom Minimum und Maximum, angewendet auf kompaktes K oder eine kompakte Teilmenge von K . (Hinweis: Die Nullstellenmenge $\{x \in D \mid g(x) = 0\}$ ist abgeschlossen, da g stetig ist. Für die Kompaktheit muss zusätzlich nur noch die Beschränktheit nachgewiesen werden.)

Geometrische Eigenschaften von Gradienten

Sei $n \in \mathbf{N}$, $D \subseteq \mathbf{R}^n$ offen, $f \in C^1(D)$ und $a \in D$.

- Die Richtung von $\nabla f(a)$ ist die Richtung des steilsten Anstiegs von f in a . Die umgekehrte Richtung (also die von $-\nabla f(a)$) ist die der maximalen Abnahme von f .

- Für jede Kurve durch a in der Niveaumenge $\{x \in D | f(x) = f(a)\}$ ist $\nabla f(a)$ senkrecht zur Richtung der Tangente an die Kurve in a .
- Sei $n = 2$ und $\nabla f(a) \neq 0$. Dann wird die Tangente an die Niveaumenge $\{x \in D | f(x) = f(a)\}$ in a durch die Gleichung $\nabla f(a) \cdot (x - a) = 0$ beschrieben.
- Sei $n = 3$ und $\nabla f(a) \neq 0$. Dann wird die Tangentialebene an die Niveaumenge $\{x \in D | f(x) = f(a)\}$ in a durch die Gleichung $\nabla f(a) \cdot (x - a) = 0$ beschrieben.