Analysis III
WS 2017/18

M. Fuchs

Inhalt

20

Höhere Ableitungen, Satz von Taylor, Extremwerte, Konvexität

Höhere Ableitungen werden rekursiv definiert.

Situation : $\begin{cases} X, Y \text{ endlich dim. } \mathbb{R}\text{- V.R. }, \; U \text{ offen } \subset X \\ f : \; U \to Y \text{ differenzierbar} \end{cases}$

Dann definiert f' eine Abbildung $U \to L(X, Y)$, die jedem $x \in U$ eine **lineare** Abbildung $X \to Y$ zuordnet.

Definition 20.1 : *(***2^{te} Ableitung***)*

a) Ist f' diff'bar an einer Stelle $a \in U$, so heißt $(f')'(a)$ die **2^{te} Ableitung** *von f in a.*

b) f heißt auf U **2^{mal} diff'bar***, wenn $(f')'(a)$ an jeder Stelle $a \in U$ existiert.*

Schreibweise: $f''(a)$ *oder* $f^{(2)}(a)$ *statt* $(f')'(a)$ *bzw.* $D^2 f(a)$

Interpretation von $\mathbf{f''(a)}$:

Sei $Z = L(X, Y) \overset{\text{per Def}}{\Longrightarrow} f''(a) : \; X \longrightarrow Z$ linear, also $f''(a) \in L(X, Z) = L\Big(X, L(X, Y)\Big)$.

D.h.: für jeden Vektor $v \in X$ ist $f''(a)(v)$ eine lineare Abbildung $X \to Y$,

kann also nochmal auf einen Vektor $w \in X$ angewendet werden.

$\Big(f''(a)(v)\Big)(w)$ ist dann ein Element von Y.

Man schreibt: $\boxed{f''(a)(v, w) \text{ und interpretiert } f''(a) \text{ als Bilinearform } X \times X \longrightarrow Y}$

Bemerkungen :

1) $\Big(f''(a)(v)\Big)(w)$ ist offenbar linear in v und w.
2) Wir geben gleich Formeln für diesen Ausdruck an.

Rekursive Definition von k-maliger Differenzierbarkeit:

$U \subset X$ offen, Y $\mathbb{R}$-V.R., $\dim X, \dim Y < \infty$ wobei $Y_\ell := L(X, Y_{\ell-1})$, $Y_0 := Y$.
(beachte: Y_ℓ ist für alle ℓ endlich dim. $\mathbb{R}$-Vektorraum)
setze nun: $f^{(\ell)} := (f^{(\ell-1)})' : U \to Y_\ell$

Definition 20.2 :

$$\begin{aligned}
C^0(U,Y) &:= \{f: U \to Y \mid \ \ f \ \textit{stetig}\},\\
C^1(U,Y) &:= \{f: U \to Y \mid \ \ f' \ \textit{existiert und ist stetig}\},\\
C^k(U,Y) &:= \{f: U \to Y \mid \ \ f^{(k)} \ \textit{existiert und ist stetig}\},\\
C^\infty(U,Y) &:= \bigcap_{k=0}^{\infty} C^k(U,Y).
\end{aligned}$$

Man kann sich (mit Induktion!) überlegen:
| $f^{(k)}(a)$ kann aufgefaßt werden als k-lineare Abbildung $X \times \ldots \times X \longrightarrow Y$.

<u>Kehren wir zurück zu</u> $f''(a): X \times X \longrightarrow Y$.

Wir hatten definiert $\Big(f''(a)$ ist <u>lin. Abb.</u> $X \to L(X,Y)\Big)$
$f''(a)(v,w) := \Big(\underbrace{f''(a)(v)}_{\in L(X,Y)}\Big)(w),$

und es sieht so aus, als ob man die Reihenfolge von v, w beachten müßte. Dem ist nicht so!

Satz 20.1 :

Sei $f: X \supset U \to Y$ 2-mal (total !) diff'bar in $a \in U$. Dann gilt für alle $v, w \in X$:

$$\boxed{f''(a)(v,w) = \partial_w(\partial_v f)(a) = \partial_v(\partial_w f)(a) = f''(a)(w,v)}$$

(Das ist einerseits eine Formel zur Berechnung von $f''(a)(v,w)$ mit Richtungsableitungen, andererseits bekommt man Vertauschbarkeit der Reihenfolge)

$f''(a)$ ist also eine **symmetrische Bilinearform**.
[zur Erinnerung : $\partial_u f(x) = \frac{d}{dt_{|0}} \; f(x+tu) = f'(x)u$]

Korollar :

Sei f **k-mal diff'bar** auf U. Dann ist $f^{(k)}(a) : \; X \times \ldots \times X \longrightarrow Y$ k-linear und **symmetrisch**. Es gilt:

$$f^{(k)}(a)\;(v_1, \ldots, v_k) = \underbrace{\partial_{v_k}\,(\partial_{v_{k-1}}(\ldots(\partial_{v_1} f)\ldots))}_{k\text{-fach iterierte Richtungsabl.}}\;(a).$$

Bemerkungen :

1) **Vertauschbarkeit der partiellen Ableitungen im Euklidischen Fall**

   Seien $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$. Dann wissen wir:
   Existenz nur von $\partial_i(\partial_j f)$ auf U für alle $i, j \not\Longrightarrow \partial_i(\partial_j f) = \partial_j(\partial_i f)$ (vgl. Analysis II)

   Satz 20.1 zeigt dagegen:

   $$f^{(2)}(x) \text{ existiert } \Longrightarrow \partial_i(\partial_j f)(x) = \partial_j(\partial_i f)(x) \quad \forall i, j$$

   Wenn also in x die (totale) 2^{te} Ableitung existiert, so kann man bei allen 2^{ten} partiellen Ableitungen die Reihenfolge beliebig ändern.

   | Dasselbe gilt sinngemäßfür k-mal diff'bare Funktionen und k-fache partielle Ableitungen.

2) **Wie berechnet man $\mathbf{f''(x)}$ im Euklidischen Fall ?**

   Sei $X = \mathbb{R}^n$, $Y = \mathbb{R}$ (f $\mathbb{R}^m$-wertig : analog!).
   Nach unserer Interpretation ist $f''(a)$ eine **Bilinarform** auf $\mathbb{R}^n \times \mathbb{R}^n$,
   $\{e_1, \ldots, e_n\}$ = kanonische Basis;
   $v, w \in \mathbb{R}^n \implies v = \sum_{i=1}^{n} v_i\, e_i, \; w = \sum_{k=1}^{n} w_k\, e_k$

   Damit: $\boxed{f''(a)(v,w) = \sum_{i=1}^{n} v_i\, w_k\; \partial_i(\partial_k\, f)\,(a)}$

   Also: $f''(a)$ **entspricht der symmetrischem Matrix**

   $$\Big(\partial_i(\partial_k f)\,(a)\Big)_{1\le i,k\le n} = \Big(\frac{\partial^2 f}{\partial x_i\,\partial x_k}\,(a)\Big)_{1\le i,k\le n}$$, der sogenannten **Hesse-Matrix**.

   Für $f^{(k)}(a)$, $k > 2$, bildet man eine entsprechende "Multimatrix".

3) Kriterien für die Existenz von $f^{(k)}(a)$: s. Satz 20.2
   (in Termen k^{ter} partieller Ableitungen)

$\square$

Beweis von 20.1 :

$f: U \to Y$ sei 2^{mal} diff'bar in a, d.h. (Def. von $f''(a)$)

(1) $\lim\limits_{v\to 0} \frac{1}{|v|} |\underbrace{f'(a+v)}_{\in L(X,Y)!} - f'(a) - f''(a)(v)| = 0$, wobei $f''(a): X \to L(X,Y)$ linear.

Setze R_a : Nullumgebung in $X \to L(X,Y)$,

$$R_a(v) := \begin{cases} 0, & v = 0 \\ \frac{1}{|v|}\left(f'(a+v) - f'(a) - f''(a)v\right), & v \neq 0 \end{cases}$$

Gemäß (1) ist R_a stetig in 0, und wir können schreiben

(2) $f'(a+v) - f'(a) - f''(a)v = |v| R_a(v)$

Seien jetzt $u, v \in X$ beliebig. Für $t \neq 0$ ist

$$\begin{aligned} & \tfrac{1}{t}\Big\{(\partial_u f)(a+tv) - (\partial_u f)(a)\Big\} \\ \overset{\partial_u f(x) = f'(x)u}{=} \; & \tfrac{1}{t}\Big\{f'(a+tv)(u) - f'(a)(u)\Big\} \\ \overset{\text{(2) mit } tv \text{ statt } v}{=} \; & \tfrac{1}{t}\Big\{|tv| R_a(tv)(u) + f''(a)(tv)(u)\Big\} \\ \overset{\text{Linearität von } f''(a)(\ldots)(u)}{=} \; & |v| \cdot R_a(tv)(u) + f''(a)(v)(u). \end{aligned}$$

Also existiert $\lim\limits_{t\to 0} \frac{1}{t}\Big\{(\partial_u f)(a+tv) - (\partial_u f)(a)\Big\}$, mit Wert $f''(a)(v)(u)$.

Es folgt nach Def. der Richtungsableitung

(3) $\partial_v(\partial_u f)(a) = f''(a)(v)(u)$

und analog

(4) $\partial_u(\partial_v f)(a) = f''(a)(u)(v)$.

Zum Beweis der Symmetrie argumentieren wir heuristisch:

$$\begin{aligned} \partial_v(\partial_u f)(a) \quad & \approx \quad \tfrac{1}{t}\Big\{(\partial_u f)(a+tv) - (\partial_u f)(a)\Big\} \\ & \approx \quad \tfrac{1}{t}\Big\{\tfrac{1}{s}\Big[f(a+tv+su) - f(a+tv)\Big] - \tfrac{1}{s}\Big[f(a+su) - f(a)\Big]\Big\} \\ & \overset{\text{setze } s=t}{=} \quad \tfrac{1}{s^2}\Big\{f(a+su+sv) - f(a+su) - f(a+sv) + f(a)\Big\} \\ & =: \quad \tfrac{1}{s^2} \cdot B_a(su, sv) \end{aligned}$$

$$\boxed{B_a(\tilde{u}, \tilde{v}) := f(a+\tilde{u}+\tilde{v}) - f(a+\tilde{u}) - f(a+\tilde{v}) + f(a)}$$

Der Ausdruck $B_a(\tilde{u},\tilde{v})$ ist symmetrisch in $\tilde{u}$ und $\tilde{v}$, so dass $f''(a)(u)(v) = f''(a)(v)(u)$ und damit alles streng bewiesen ist wenn wir

(5) $\lim\limits_{s\to 0} \frac{1}{s^2}\, B_a(su,sv) = f''(a)(v)(u)$ zeigen können.

zu (5) :

O.E. f $\mathbb{R}$-wertig (sonst wähle Basis in Y und betrachte die Komponenten von f).
Seien $u, v \in X,\ \delta > 0$ sei so klein, dass für $0 < s < \delta :\ a+su,\ a+sv,\ a+s(u+v) \in U$.
Sei ein $s \in (0,\delta)$ fixiert, $\tilde{u} := su,\ \tilde{v} := sv$.

Setze $F(t) := f(a + t\tilde{u} + \tilde{v}) - f(a + t\tilde{u})$ für $0 \le t \le 1$.

$\Longrightarrow$ $F'(t) = f'(a + t\tilde{u} + \tilde{v})(\tilde{u}) - f'(a + t\tilde{u})(\tilde{u})$

$\overset{MWS}{\Longrightarrow}$ $F(1) - F(0) = B_a(su,sv) = F'(\vartheta)$ für ein $0 < \vartheta < 1$

Also:

$$
\begin{aligned}
B_a(su,sv) &= f'(a+\vartheta\tilde{u}+\tilde{v})(\tilde{u}) - f'(a+\vartheta\tilde{u})(\tilde{u})\\
&= f'(a+\vartheta\tilde{u}+\tilde{v})(\tilde{u}) - f'(a)(\tilde{u}) - \Big(f'(a+\vartheta\tilde{u})(\tilde{u}) - f'(a)(\tilde{u})\Big)\\
&\overset{(2)}{=} |\vartheta\tilde{u}+\tilde{v}| \cdot R_a(\vartheta\tilde{u}+\tilde{v})(\tilde{u}) + f''(a)(\vartheta\tilde{u}+\tilde{v})(\tilde{u})\\
&\qquad -\Big(|\vartheta\tilde{u}| \cdot R_a(\vartheta\tilde{u})(\tilde{u}) + f''(a)(\vartheta\tilde{u})(\tilde{u})\Big)\\
&\underset{\substack{\uparrow\\ \text{Bilinearität von} f''(a)\\ \downarrow}}{=} f''(a)(\tilde{v})(\tilde{u}) + |\vartheta\tilde{u}+\tilde{v}|\ R_a(\vartheta\tilde{u}+\tilde{v})(\tilde{u}) - |\vartheta\tilde{u}|\ R_a(\vartheta\tilde{u})(\tilde{u})\\
&= s^2\Big\{f''(a)(v)(u) + |\vartheta u+v|\ R_a(\vartheta\tilde{u}+\tilde{v})(u) - |\vartheta u|\ R_a(\vartheta\tilde{u})(u)\Big\}
\end{aligned}
$$

$\Longrightarrow$ $\frac{1}{s^2}\, B_a(su,sv) = f''(a)(v)(u) + |\vartheta u+v|\ R_a(\vartheta\tilde{u}+\tilde{v})(u) - |\vartheta u|\ R_a(\vartheta\tilde{u})(u)$

und man erhält $\lim\limits_{s\downarrow 0} s^{-2}B_a(su,sv) = f''(a)(v)(u)$

gemäß $\lim\limits_{w\to 0} R_a(w) = 0.$

$\square$

Satz 20.2 : *(***Kriterium für k-fache stetige Diff'barkeit***)*

Sei $U \subset \mathbb{R}^n$ *offen und* $f : U \longrightarrow \mathbb{R}^m$. *Dann gilt:*
$f \in C^k(U, \mathbb{R}^m)$ *(d.h.* $f^{(1)}, \ldots, f^{(k)}$ *existieren auf* U *und sind stetig)*

$$\Updownarrow$$

für jeden *Multiindex* $\alpha \in \mathbb{N}_0^n$, $|\alpha| \leq k$, *existieren die partiellen Ableitungen*

$$\partial^\alpha f := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \left(\frac{\partial^{\alpha^2}}{\partial x_2^{\alpha_2}} \left(\ldots \left(\frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} \, f \right) \ldots \right) \right)$$

und sind **stetig** *auf* U $\quad \Big(\alpha = (\alpha_1, \ldots, \alpha_n), \ |\alpha| = \alpha_1 + \ldots + \alpha_n\Big)$

Beweis : (Induktion über k)

Satz 18.4 ergibt den Induktionsanfang für $k = 1$.

□

Beispiel :

Polynome und **rationale** Funktionen $\mathbb{R}^n \longrightarrow \mathbb{R}$ sind auf ihrem Definitionsbereich unendlich oft differenzierbar.

Als Anwendung des Konzepts der höheren Ableitung diskutieren wir nun

Taylor Formel für reellwertige Funktionen.

Satz 20.3 : *(wird in der Extremwertdiskussion eingesetzt, s.u.)*

Sei U *offen in* $\mathbb{R}^n$ *und* $f \in C^{\ell+1}(U, \mathbb{R})$ *für ein* $\ell \geq 1$. *Sei* $v \in \mathbb{R}^n$ *mit* $a + tv \in U$ *für alle* $0 \leq t \leq 1$, $a \in U$ *fixierter Punkt.*

Dann gilt:

$$\boxed{f\,(a + v) = \sum_{k=0}^{\ell} \tfrac{1}{k!} f^{(k)}(a) \underbrace{(v, \ldots, v)}_{k\ fach} + R_{a,\ell}(v)}$$

mit dem **Restglied**

$$\boxed{R_{a,\ell}(v) = \tfrac{1}{(\ell+1)!} f^{(\ell+1)}(a + \vartheta v) \underbrace{(v, \ldots, v)}_{\ell+1\ fach}}$$

an einer geeigneten Zwischenstelle $\vartheta \in (0, 1)$.

Bemerkungen :

1) $f^{(k)}(b),\ b \in U$ ist eine k-lineare Abbildung $\mathbb{R}^n \times \ldots \times \mathbb{R}^n \to \mathbb{R}$; $f^{(k)}(b)(w, \ldots, w)$ heißt, k-mal den selben Vektor w in diese Form einzutragen.
(Erinnerung : $f^{(k)}(b) = \partial_w(\partial_w \ldots (\partial_w f) \ldots)(b) = \frac{d^k}{dt^k|_0}\ f(b + tw)$)

2) Es gilt folgende **Integraldarstellung des Restglieds**:

$$R_{a,\ell}(v) = \frac{1}{\ell!} \int_0^1 (1-t)^\ell\ f^{(\ell+1)}(a+tv)(v, \ldots, v)\ dt$$

3) $f \in C^\infty(U) \implies$ die **formale Taylorreihe** $\sum\limits_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(a)(x-a, \ldots, x-a), x \in U$
mit Entwicklungsmitte a ist bildbar

4) f heißt **reell analytisch auf** U $:\Longleftrightarrow$ jeder Punkt $a \in U$ hat eine Umgebung, auf der die formale Taylorreihe gegen f konvergiert

Beweis von Satz 20.3 :

Benutze die Taylorformel für $F(t) := f(a+tv),\ 0 \le t \le 1 \implies F(1) = \ldots$. Die Integraldarstellung 2) folgt aus der Darstellung des Restglieds von F.

$\square$

Lokale Extrema :

Definition 20.3 :

Sei $U \subset \mathbb{R}^n$ offen und $f : U \longrightarrow \mathbb{R}$

a) Ist f auf U differenzierbar, so heißt $a \in U$ **kritischer Punkt von** *f,*
wenn $f'(a) = 0$ ist, also $\nabla f(a) = 0$ (da $f'(a)v = \langle \nabla f(a), v \rangle$)

b) $a \in U$ heißt **lokales Maximum** *(Minimum) von f*

$$\Longleftrightarrow$$

$\exists\, r > 0 :\ f(x) \le (\ge) f(a)$ *für alle* $x \in B_r(a)$
("strikt" , wenn $<$ $(>)$ für $x \in B_r(a) - \{a\}$)

Satz 20.4 :

Sei $f :\ U \to \mathbb{R}$ differenzierbar und $a \in U$ lokales Max. (Min.).
Dann ist a kritischer Punkt.

allgemeiner:

a lokaler Extremwert
$\implies\ \partial_v f(a) = 0$ für jede Richtung v, für die $\partial_v f(a)$ existiert.

Bemerkung :

$$\nabla f(a) = 0 \quad \not\Longrightarrow \quad a \text{ lokaler Extremwert (ist ja schon falsch für } n = 1,\ x \mapsto x^3)$$

Beweis von Satz 20.4 :

Sei a lokales <u>Max</u> $\Longrightarrow$

$$f(a + tv) \leq f(a) \quad \forall\ |t| \text{ klein genug}$$

Also: $\frac{1}{t}\ (f(a + tv) - f(a)) \begin{cases} \geq 0, & t < 0 \\ \leq 0, & t > 0 \end{cases} \quad \Longrightarrow \quad \partial_v f(a) = 0.$

Ist f diff'bar in a, so folgt $\partial_v f(a) = 0$ für alle v, d.h. $f'(a) = 0$.

$\square$

Beispiel :

$f(x, y) := x^3 + y^3 - 3xy$ auf $U = \mathbb{R}^2$ $\quad$ f ist beliebig oft diff'bar:

Kritische Punkte :
$$\frac{\partial f}{\partial x}(x, y) = 3x^2 - 3y \overset{!}{=} 0$$
$$\frac{\partial f}{\partial y}(x, y) = 3y^2 - 3x \overset{!}{=} 0$$

$\Longrightarrow \quad (0, 0), (1, 1)$ als Lösungen.

<u>in $(0, 0)$ liegt kein lokaler Extremwert vor</u>:

es ist $\quad f(x, 0) = x^3 \quad \begin{array}{l} > 0 \text{ für } x > 0 \\ < 0 \text{ für } x < 0 \end{array}$

d.h. f nimmt in jeder Umgebung von $(0, 0)$ sowohl positive als auch negative Werte an.

f hat <u>keine absoluten Extremwerte</u>:

$$\left.\begin{array}{lcll} f(n, 0) & = & n^3 & \longrightarrow \infty, \\ f(-n, 0) & = & -n^3 & \longrightarrow -\infty \end{array}\right\} \text{ bei } n \to \infty$$

<u>Was passiert in $(1, 1)$</u>?

Um dies zu entscheiden, formulieren wir <u>Kriterien mit der 2^{ten} Ableitung</u>.

Einschub aus der Algebra :

$X = \mathbb{R}$ - V.R. endlicher Dimension,

$B : X \times X \to \mathbb{R}$ symmetrische Bilinearform $\Big(B(u,v) = B(v,u)\Big)$

$Q(v) := B(v,v)$ zugehörige quadratische Form.

Es gilt:

$$\left|\begin{array}{lll} Q \textit{ positiv (negativ) definit} & :\Longleftrightarrow & Q(v,v) > 0 (< 0) \quad \forall\, v \neq 0 \\ \ldots\ldots(\ldots) \textit{ semidefinit} & :\Longleftrightarrow & Q(v,v) \geq 0, Q(w,w) < 0 \quad \forall\, v \\ Q \textit{ indefinit} & :\Longleftrightarrow & \exists\, v,w \neq 0 \text{ mit } Q(v,v) > 0, Q(w,w) < 0 \end{array}\right.$$

(man sagt auch: B ist ... definit und meint das zugehörige Q)

Sei $X = \mathbb{R}^n$, A eine symmetrische $(n \times n)$ - Matrix, d.h. $\langle v, Aw\rangle = \langle Av, w\rangle \forall\, v, w \in \mathbb{R}^n$

bilde $B(v,w) := \langle v, Aw\rangle \overset{Bilinearform (symm.!)}{\Longrightarrow} Q(v) = \langle v, Av\rangle$

Wie entscheidet man, ob dieses Q definit ist?

Betrachte dazu:

A symmetrisch $\Longrightarrow$ $\exists$ Orthonormalbasis $v_1, \ldots, v_n$ aus *Eigenvektoren*
(also: $Av_i = \lambda_i v_i$ mit $\lambda_i \in \mathbb{R}$, $i = 1, \ldots, n$)

schreibe nun $w \in \mathbb{R}^n$ als $w = \sum\limits_{i=1}^{n} \alpha_i v_i \Longrightarrow Q(w) = \langle w, Aw\rangle = \sum\limits_{i,j=1}^{n} \alpha_i \alpha_j \langle v_i, Av_j\rangle = \sum\limits_{i=1}^{n} |\alpha_i|^2 \lambda_i$,

und damit kann man sich sehr leicht überlegen, dass gilt:

$$\left|\begin{array}{lll} Q \textit{ positiv (negativ) definit} & \Longleftrightarrow & \textbf{alle} \text{ EW von } A \text{ positiv (negativ)} \\ \quad\text{''} \qquad \textit{semidefinit} & \Longleftrightarrow & \ldots\ldots \geq 0 \ (\leq 0) \\ Q \textit{ indefinit} & \Longleftrightarrow & \exists \text{ EW } > 0 \textbf{ und } < 0 \end{array}\right.$$

$\to$ Wie schlägt man die Brücke zu lokalen Extremwerten?

Sei $f : U \to \mathbb{R}$, U offen $\subset \mathbb{R}^n$, von der Klasse C^2, a ein Punkt aus U. Dann gilt:

$$f''(a)(v,w) \;=\; \sum_{i,j=1}^{n} v_i w_j \, \frac{\partial^2 f}{\partial x_i \partial x_j}\,(a) \;=\; \Big\langle v,\; \underbrace{\Big(\frac{\partial^2 f}{\partial x_i \partial x_j}\,(a)\Big)_{1\le i,j\le n}}_{\uparrow\, A} \, w \Big\rangle$$

mit *symmetrischer Matrix* A (Hesse-Matrix).

Nach Taylor ist:

$$\begin{aligned} f\,(a+v) &= f\,(a) \;+\; \langle \nabla f\,(a), v\rangle \;+\; \tfrac{1}{2}\, f^{(2)}\,(a+\vartheta v)\,(v,v) \\ &= f\,(a) \;+\; \langle \nabla f\,(a), v\rangle \;+\; \tfrac{1}{2}\, f^{(2)}\,(a)\,(v,v) \\ &\quad + \tfrac{1}{2}\,\Big(f^{(2)}\,(a+\vartheta v)\,(v,v) \;-\; f^{(2)}\,(a)\,(v,v)\Big) \qquad \text{mit } \vartheta \in (0,1). \end{aligned}$$

Gemäß Vorraussetzung ist $f^{(2)} : U \to L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}))$ **stetig**, woraus unschwer folgt ($\to$ Übung!)

$$(1) \quad \lim_{x\to a} \;\sup_{|w|=1} \Big| f^{(2)}(x)(w,w) - f^{(2)}\,(a)(w,w) \Big| \;=\; 0$$

Sei nun a **kritischer Punkt** ($\nabla f(a) = 0$) und $f^{(2)}(a)$ **positiv** definit. Dann folgt

$$(2) \quad \min_{|w|=1} \; f^{(2)}(a)(w,w) =: c_\circ > 0.$$

(wegen der Stetigkeit von $w \mapsto f^{(2)}(a)(w,w)$ und der Kompaktheit von $\{w \in \mathbb{R}^n : |w| = 1\}$)

Für $v \neq 0$ folgt nach (2)

$$\begin{aligned} f(a+v) \;-\; f(a) &= \tfrac{1}{2}\, f^{(2)}(a)(v,v) \;+\; \tfrac{1}{2}\, \Big(f^{(2)}(a+\vartheta v)(v,v) \;-\; f^{(2)}(a)(v,v)\Big) \\ &\ge \tfrac{1}{2}\,|v|^2\, c_\circ \;-\; \tfrac{1}{2}\,|v|^2 \sup_{|w|=1} \Big| f^{(2)}(a+\vartheta v)(w,w) \;-\; f^{(2)}(a)(w,w)\Big|. \end{aligned}$$

Wählt man $|v| < \delta$ mit passendem δ, so kann man gemäß (1)

$$\sup_{|w|=1} |\ldots| \;\le\; \frac{1}{2}\, c_\circ$$

erreichen, also

$$f(a+v) - f(a) \;\le\; \frac{1}{4}\, c_\circ\, |v|^2 \quad \forall\, |v| < \delta.$$

f hat also in a ein strenges lokales Minimum. $\Longrightarrow$

Satz 20.5 :

Sei $f \in C^2(U, \mathbb{R}),\ U \subset \mathbb{R}^n$, *und* <u>*a kritischer Punkt von* f</u>. *Dann gilt:*

i)	*f" (a)* **positiv** *definit*	$\Longrightarrow$	*a ist striktes lokales* **Min.**
ii)	*f" (a)* **negativ** *definit*	$\Longrightarrow$	" **Max.**
iii)	*f" (a)* **indefinit**	$\Longrightarrow$	*a ist* **keine lokale** *Extremstelle*

Beweis :

(i), (ii) $\surd$

(iii) $f''(a)$ indefinit $\Longrightarrow\ \exists$ Vektoren $u_\circ, v_\circ$ mit Länge 1 und

$$f^{(2)}(a)(u_\circ, u_\circ)\ <\ 0\ <\ f^{(2)}(a)(v_\circ, v_\circ)$$

Daraus folgt für $|t| \neq 0$ klein (diskutiere $t \mapsto f\,(a+tu_\circ),\ f(a+tv_\circ)$ als Fkt. $\mathbb{R} \to \mathbb{R}$)

$$f(a+tu_\circ) - f(a)\ < 0 <\ f(a+tv_\circ) - f(a),$$

also kein Extremwert in a.

$\square$

Bemerkung :

<u>Die Kriterien sind nur hinreichend</u>! $f : t \mapsto t^4$ hat absolutes Min. in 0, $f''(0) = 0$.

Korollar :

Sei $f \in C^2\ (U, \mathbb{R})$. Dann gilt:
a lokales Max. (Min.) $\Longrightarrow f''(a)$ negativ (positiv) **semidefinit**

Beweis :

Sei a Max. $f''(a)$ nicht negativ semidefinit
$\Longrightarrow\ \ \exists\ v_\circ, |v_\circ|\ =\ 1$ mit $f''(a)(v_\circ, v_\circ)\ <\ 0$
$\Longrightarrow\ \ t \mapsto f(a+tv_\circ)\ =\ \varphi(t)$ erfüllt $\varphi'(0)\ =\ 0$ und $\varphi''(0)\ =\ f''(a)(v_\circ, v_\circ)\ >\ 0$
$\Longrightarrow\ \ f(a)\ <\ f(a+tv_\circ)$ für $|t|\ \neq\ 0$ klein genug, Wspr.!

$\square$

Beispiel :

$$f(x,y) \;=\; x^3 + y^3 - 3xy \text{ in } (1,1)$$

$$\frac{\partial^2 f}{\partial x^2} = 6x, \quad \frac{\partial^2 f}{\partial y^2} = 6y, \quad \frac{\partial^2 f}{\partial x \partial y} = -3 \quad \Longrightarrow \quad f^{(2)}(1,1) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$$

Zeige (→ Übung):

Eine <u>symmetrische</u> 2×2 Matrix $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ ist

positiv definit $\iff$ $a > 0$ und $\det > 0$

negativ definit $\iff$ $a < 0$ und $\det > 0$

indefinit $\iff$ $\det < 0$

hier: $a = 6 > 0,\ \det = 27 > 0 \Longrightarrow$ <u>$(1,1)$ lok. Min.</u>

Bemerkung :

Die Matrix $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\ (a)\right)_{1<i,j\leq n}$ heißt <u>Hesse-Matrix</u> von f bei a.

Kriterien für Konvexität:

Erinnerung :

$U \subset \mathbb{R}^n$ <u>konvex</u>; $f : U \to \mathbb{R}$ <u>konvexe Funktion</u> $\iff$

$f\Big(tx + (1-t)\ y\Big) \;\leq\; t\ f(x) + (1-t) f(y) \quad \forall\ x, y \in U,\ 0 \leq t \leq 1$

(streng konvex, konkav ...)

Satz 20.6 :

Sei $U \subset \mathbb{R}^n$ konvex und offen, $f : U \to \mathbb{R}$ diff'bar. Dann gilt:

(i) *f konvex* $\iff f(x) \;\geq\; \overbrace{f(x_\circ) \;+\; \langle \nabla f(x_\circ),\ x - x_\circ \rangle}^{\text{Graph davon ist Ebene in } \mathbb{R}^{n+1} \text{ durch } (x_\circ,\ f(x_\circ))\ \text{"Stützebene"}}$ *für alle $x,\ x_\circ \in U$*

(ii) *f streng konvex* $\iff >$ *analog* *für alle $x \neq x_\circ$*

(iii) *konkav (streng)* *analog*

Beweis : nur (i)

"$\Longrightarrow$": Sei f konvex.

Mit $v := x - x_\circ,\ t \in (0,1)$ ist $f(x_\circ + tv) \ \leq\ t\, f(x_\circ + v) \ +\ (1-t)\ f(x_\circ)$

$\Longrightarrow\ \frac{1}{t}\,\Big(f(x_\circ + tv)\ -\ f(x_\circ)\Big)\ \leq\ f(x_\circ + v)\ -\ f(x_\circ)$

und mit $t \downarrow 0 :\quad \langle \nabla f(x_\circ), v\rangle\ \leq\ f(x)\ -\ f(x_\circ)$

"$\Longleftarrow$": Wir wissen jetzt: $f(y+w)\ \geq\ f(y) + \langle \nabla f(y), w\rangle \quad \forall\, y \in U,\ y+w \in U.$

Seien $x_1 \neq x_2$ aus $U,\ t \in (0,1)$

setze $x_\circ := t \cdot x_1\ +\ (1-t)\ x_2,\ v := x_1 - x_\circ$

dann: $f(x_\circ + v)\ \geq\ f(x_\circ)\ +\ \Big\langle \nabla f(x_\circ),\ v\Big\rangle \quad \Longleftrightarrow$

(1) $\quad f(x_1)\ \geq\ f\Big(tx_1 + (1-t)\ x_2\Big)\ +\ \Big\langle \nabla f(x_\circ), v\Big\rangle$

analog: $f(x_2)\ =\ f\Big(x_\circ\ +\ [-\frac{t}{1-t} \cdot v]\Big)\ \geq\ f(x_\circ) - \frac{t}{1-t}\,\Big\langle \nabla f(x_\circ),\ v\Big\rangle \quad \Longrightarrow$

(2) $\quad f(x_2)\ \geq\ f\Big(tx_1\ +\ (1-t)\ x_2\Big)\ -\ \frac{t}{1-t}\,\Big\langle \nabla f(x_\circ),\ v\Big\rangle$

Nun multipliziere (1) mit t, (2) mit $(1-t)$ und addiere die Resultate.

$\square$

Für C^2-Funktionen läßt sich Konvexität, etc., einfach prüfen.

Satz 20.7 :

Sei $U \subset \mathbb{R}^n$ offen und konvex, $\boxed{f \in C^2\ (U)}$

(i) f konvex $\Longleftrightarrow$ $f^{(2)}(x)$ positiv semidefinit für alle x

(ii) $f^{(2)}$ positiv definit auf U $\Longrightarrow$ f streng konvex

(iii) "konkav"

(iv) "konkav"

Bemerkung :

<u>Umkehrung in (ii)</u> <u>falsch</u> bereits für $n = 1 :\ t \mapsto t^4,\ t \in \mathbb{R}$, ist streng konvex.

Beweis : nur (i)

"$\Longleftarrow$":

Wähle $x_\circ \neq x$ aus $U \overset{\text{Taylor}}{\Longrightarrow}$ es existiert $\vartheta \in (0,1)$, so dass gilt:

$f(x) = f(x_\circ) + \Big\langle \nabla f(x_\circ),\ x - x_\circ \Big\rangle + \frac{1}{2}\, f''\Big(x_\circ + \vartheta(x - x_\circ)\Big)(x - x_\circ, x - x_\circ)$

nach Vor. ist $f''(\ldots)(x - x_\circ,\ x - x_\circ)\ \geq\ 0$

also $f(x)\ \geq\ f(x_\circ)\ +\ \Big\langle \nabla f(x_\circ),\ x - x_\circ \Big\rangle$; nun benutze 20.6 .

"$\Longrightarrow$": (indirekt)

$f^{(2)}$ **nicht** positiv semidefinit auf U

$\Longrightarrow\ \exists\ x_\circ \in U$ und $v \in \mathbb{R}^n,\ |v| = 1$, mit

$$f^{(2)}(x_\circ)(v, v) < 0.$$

Da $f^{(2)}$ auf U stetig ist, gibt es $B_\delta(x_\circ) \subset U$ mit

$$f^{(2)}(y)(v, v)\ < 0 \quad \forall\ y \in B_\delta(x_\circ). \qquad (*)$$

Wähle $t \in \mathbb{R} \setminus \{0\}$ mit $|t \cdot v| < \delta$
und setze $w := tv,\ x := x_\circ + w \in B_\delta(x_\circ) \overset{\text{Taylor}}{\Longrightarrow}$

$$\begin{aligned} f(x) &= f(x_\circ) + \langle \nabla f(x_\circ),\ w\rangle + \int_0^1 (1-s) f^{(2)} \overbrace{(x_\circ + sw)}^{\in B_\delta(x_\circ)}(w, w)\, ds \\ &= f(x_\circ) + \langle \nabla f(x_\circ),\ w\rangle + t^2 \int_0^1 (1-s) f^{(2)}(x_\circ + sw)(v, v)\, ds \\ &\overset{(*)}{<} f(x_\circ) + \langle \nabla f(x_\circ),\ w\rangle. \end{aligned}$$

<u>Da f konvex ist, widerspricht dies 20.6 i)</u>.

$\square$

$\rightarrow$ Übungen: weitere Beispiele!