Vorlesung Analysis II — SoSe '25 hhu

Teil 1: Differentialrechnung im $\mathbb{R}^n$ — K. Halupczok

an7: Satz von Taylor, Lokale Extrema

Stichworte: Satz von Taylor, Extrema, kritische Stellen, Kriterien, Hessematrix

Literatur: [Hoff] Kapitel 9.6/7, [Forster] Kapitel 7

7.1. Einleitung: Der Satz von Taylor in der mehrdimensionalen Version für Skalarfelder liefert Kriterien zur Erkennung von Extrema anhand Gradienten und Hessematrix.

7.2. Vor.: $f: U \to \mathbb{R}$ mit $U \subseteq \mathbb{R}^n$, $f \in \mathcal{C}^{m+1}(U, \mathbb{R})$, $\overline{ax} \subseteq U$.

Bezeichnung: Für $\alpha \in \mathbb{N}_0^n$, $\alpha = (\alpha_1, \ldots, \alpha_n)$ setze zur einfacheren Notation

$$|\alpha| := \alpha_1 + \ldots + \alpha_n, \quad \alpha! := \alpha_1! \cdots \alpha_n!$$

$$D^\alpha := D_1^{\alpha_1} \circ D_2^{\alpha_2} \circ \ldots \circ D_n^{\alpha_n}, \quad x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \quad X^\alpha := X_1^{\alpha_1} \cdots X_n^{\alpha_n}.$$

Man nennt α auch einen Multi-Index.

Damit kann jedes Polynom $P \in \mathbb{R}[X_1, \ldots, X_n]$, $\deg P = m$, auch in der kurzen Multi-Index-Schreibweise notiert werden als

$$P = \sum_{\substack{\alpha \in \mathbb{N}_0^n \\ |\alpha| \le m}} c_\alpha X^\alpha, \text{ d.h. } P(X_1, \ldots, X_n) = \sum_{\substack{0 \le \alpha_1, \ldots, \alpha_n \le m \\ \alpha_1 + \ldots + \alpha_n \le m}} c_{(\alpha_1, \ldots, \alpha_n)} X_1^{\alpha_1} \cdots X_n^{\alpha_n}.$$

Bsp.: $$\sum_{|\alpha| \le 2} \alpha! X^\alpha = \underbrace{0! X_1^0 \cdots X_n^0}_{\text{Grad } 0} + \underbrace{\sum_{i=1}^n 1! X_i^1}_{\text{Grad } 1} + \underbrace{\sum_{1 \le i < j \le n} 1!1! X_i^1 X_j^1 + \sum_{i=1}^n 2! X_i^2}_{\text{Grad } 2}.$$

Damit kann der Satz von Taylor in einer kurzgefassten Formel notiert werden:

7.3. Satz von Taylor: Unter der Vor. wie in 7.2 gilt:

Beh.: $\exists c \in \overline{ax}$:

$$\boxed{f(x) = \sum_{|\alpha| \le m} \frac{1}{\alpha!} (D^\alpha f)(a) (x-a)^\alpha + \sum_{|\alpha| = m+1} \frac{1}{\alpha!} (D^\alpha f)(c) (x-a)^\alpha}$$

7.4. Bem.: Für $m = 0$ lautet die Beh. $f(x) = f(a) + Df(c)(x-a)$ für ein $c \in \overline{ax}$, dies ist die Aussage des MWS 6.4.

7.5 Kor.: Für $m=1$ lautet die Beh.

$$f(x)=f(a)+\langle \operatorname{grad} f(a), x-a\rangle + \tfrac{1}{2}(x-a)^T H(f;c)(x-a)$$

mit der (laut dem Satz von Schwarz) symmetrischen Matrix $H(f;c):=(D_iD_jf(c))_{n,n}$, die Hessematrix heißt; die zugehörige quadratische Form heißt Hesseform.
(Auch: Schreibweise $\operatorname{Hess} f(c)$ statt $H(f;c)$ üblich.)
(Jede symmetrische Matrix A, wo $A^T=A$, definiert über $\langle x,Ax\rangle = x^TAx$ eine quadratische Form.)

7.6 Bem.: Ist P ein Polynom, $P\in\mathbb{R}[X_1,\dots,X_n]$, $\deg P=m$, dann ist $P\in\mathcal{C}^\infty(\mathbb{R}^n)$ und P ist seine (eigene) Taylorreihe.

7.7. Beweis des Satzes 7.3 von Taylor:
1. Schritt: Für $\varepsilon>0$, $t\in\,]-\varepsilon,1+\varepsilon[\,\subseteq\mathbb{R}$ setze $g(t):=f(a+t(x-a))$ für festes x und a. Es ist also $g\in\mathcal{C}^{m+1}(]-\varepsilon,1+\varepsilon[)$.
Beh.: Für $k\le m+1$ ist

$$\frac{d^kg}{dt^k}(t)\overset{!}{=}\sum_{|\alpha|=k}\frac{k!}{\alpha!}D^\alpha f(a+t(x-a))\,(x-a)^\alpha.$$

Bew.: Setze zunächst $y:=x-a=:(y_1,\dots,y_n)^T$.
Beh.: $\displaystyle\frac{d^kg}{dt^k}(t)=\sum_{i_1,\dots,i_k=1}^{n} D_{i_k}D_{i_{k-1}}\cdots D_{i_1}f(a+ty)\,y_{i_1}\cdots y_{i_k}$.

Bew.: Vollständige Induktion über k:
$k=1$: $\displaystyle\frac{dg}{dt}(t)=\sum_{i=1}^n D_if(a+ty)\,y_i$ nach Kettenregel 5.12
denn $Df(z)=(D_1f(z),\dots,D_nf(z))$,
$D(a+ty)=y=(y_1,\dots,y_n)^T$.

$k\rightsquigarrow k+1$: $\displaystyle\frac{d^{k+1}g}{dt^{k+1}}(t)=\frac{d}{dt}\Big(\sum_{i_1,\dots,i_k=1}^{n} D_{i_k}D_{i_{k-1}}\cdots D_{i_1}f(a+ty)\,y_{i_1}\cdots y_{i_k}\Big)$

$$=\sum_{j=1}^n\sum_{i_1,\dots,i_k=1}^{n} D_jD_{i_k}D_{i_{k-1}}\cdots D_{i_1}f(a+ty)\,y_{i_1}\cdots y_{i_k}\,y_j,$$ setze $i_{k+1}:=j$

$$=\sum_{i_1,\dots,i_{k+1}=1}^{n} D_{i_{k+1}}D_{i_k}\cdots D_{i_1}f(a+ty)\,y_{i_1}\cdots y_{i_k}\,y_{i_{k+1}}$$ nach Kettenregel 5.12 ⌟

Nach dem Satz von Schwarz ist hier beliebiges Umordnen der partiellen Ableitungen möglich. Fassen daher in den $(i_1,\dots,i_k)\in\{1,\dots,n\}^k$ gleiche Indizes zusammen. Dabei komme der Index j darunter α_j-mal vor ($j\in\{1,\dots,n\}$), und wir erhalten zu einem $(i_1,\dots,i_k)$ ein bestimmtes $(\alpha_1,\dots,\alpha_n)$ mit $\alpha_1+\dots+\alpha_n=k$.
Zu diesem $(\alpha_1,\dots,\alpha_n)$ gibt es genau $\frac{k!}{\alpha!}$ viele Möglichkeiten, ein solches passendes $(i_1,\dots,i_k)$ zu finden (beweisbar durch vollständige Induktion, s. 7.8).
Daraus folgt

$$\frac{d^k g}{dt^k}(t)=\sum_{|\alpha|=k}\frac{k!}{\alpha!}D^\alpha f(a+ty)\,y^\alpha \quad \text{für } k\geq 1.$$

($k=0$: Def. von g)

2. Schritt: Wenden nun auf g den eindimensionalen Satz von Taylor An19.3 an:

$$f(x)=g(1)=\sum_{j=0}^{m}\frac{g^{(j)}(0)}{j!}+\frac{1}{(m+1)!}g^{(m+1)}(\vartheta) \quad \text{für } \vartheta\in\,]0,1[\quad \text{(Lagrange-Restglied An19.8)}$$

$$\overset{\text{1. Schritt}}{=}\sum_{j=0}^{m}\frac{1}{j!}\sum_{|\alpha|=j}\frac{j!}{\alpha!}D^\alpha f(a)(x-a)^\alpha+\frac{1}{(m+1)!}\sum_{|\alpha|=m+1}\frac{(m+1)!}{\alpha!}D^\alpha f(c)(x-a)^\alpha$$

mit $c:=a+\vartheta(x-a)$, also $c\in\overline{ax}\setminus\{a,x\}$

$$=\sum_{|\alpha|\leq m}\frac{1}{\alpha!}D^\alpha f(a)(x-a)^\alpha+\sum_{|\alpha|=m+1}\frac{1}{\alpha!}D^\alpha f(c)(x-a)^\alpha.$$

□

7.8. Beh.: Zu $(\alpha_1,\dots,\alpha_n)\in\mathbb{N}_0^n$ gibt es $\frac{k!}{\alpha!}$ viele k-Tupel $(i_1,\dots,i_k)\in\{1,\dots,n\}^k$ so, dass ein (vorgeg.) $j\in\{1,\dots,n\}$ darin α_j-mal vorkommt mit $\alpha_1+\dots+\alpha_n=k$. Was sind die 4-Tupel $(i_1,\dots,i_4)$ zu $\alpha=(\alpha_1,\alpha_2,\alpha_3)=(1,3,0)$? $k=4$, $n=3$ Ü ($j=2$)

Bew. durch vollständige Induktion über k:

$k=1$: Zu $(0,\dots,0,\underset{\text{an Stelle } j}{1},0,\dots,0)$ gibt es $\frac{k!}{\alpha!}=1$ einziges k-Tupel (i_1), so, dass j darin α_j-malig, d.h. 1-malig vorkommt, nämlich (j). ✓

$k\leadsto k+1$: Zu $(\alpha_1,\dots,\alpha_n)$ mit $\alpha_1+\dots+\alpha_n=k+1$ gibt es, falls $i_1=j$ ist, $\underbrace{\alpha_1+\dots+(\alpha_j-1)+\dots+\alpha_n}_{=k}$ nach Ind. vor. genau $\frac{k!}{\alpha_1!\cdots(\alpha_j-1)!\cdots\alpha_n!}$ viele $(k+1)$-Tupel $(i_1,\dots,i_{k+1})$ der geforderten Art, insgesamt sind dies: $\sum_{j=1}^{n}\frac{k!}{\alpha_1!\cdots(\alpha_j-1)!\cdots\alpha_n!}=\frac{k!(\alpha_1+\dots+\alpha_n)}{\alpha_1!\cdots\alpha_n!}=\frac{(k+1)!}{\alpha!}$.

✓ □

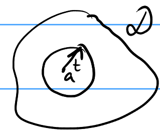
7.9. **Bezeichnung:** Sei $a \in D \subset \mathbb{R}^n$, $f: D \to \mathbb{R}$. Sei $\mathcal{U}_a := \{U \subset \mathbb{R}^n;\ a \in U\}$.

7.10. **Def.:** f hat in a ein relatives $\left\{\begin{matrix}\text{Maximum}\\ \text{Minimum}\end{matrix}\right\} :\Leftrightarrow \exists\, U \in \mathcal{U}_a : f_{\restriction U} \left\{\begin{matrix}\le f(a)\\ \ge f(a)\end{matrix}\right\}$.

Def.: f hat in a ein striktes $\left\{\begin{matrix}\text{Maximum}\\ \text{Minimum}\end{matrix}\right\} :\Leftrightarrow \exists\, U \in \mathcal{U}_a : f_{\restriction U\setminus\{a\}} \left\{\begin{matrix}< f(a)\\ > f(a)\end{matrix}\right\}$.

Ein relatives Extremum heißt auch lokales Extremum,
" " globales Extremum, falls $U = D$ wählbar.

7.11. **Erste notwendige Bedingung:** Sei $f: D \to \mathbb{R}$, $D \subset \mathbb{R}^n$.
Vor.: f habe in $a \in D$ relatives Extremum, $\forall j \in \{1, \dots, n\}$: $D_j f(a)$ ex.
Beh.: $D_j f(a) = 0$ bzw. $\operatorname{grad} f(a) = 0$.
Bew.: Sei $\varphi(t) := f(a + te_j)$, für $|t|$ hinreichend klein:



φ hat in 0 ein relatives Extremum $\Rightarrow \varphi'(t) = 0$
$\Rightarrow D_j f(a) = 0$. □

7.12. **Def.:** a heißt kritischer Punkt von f, falls
$f \in \mathcal{C}^1(D)$, $D \subset \mathbb{R}^n$, und $\operatorname{grad} f(a) = 0$ gilt.

7.13. **Zweite notwendige Bedingung:**
Vor.: f habe in a ein relatives Maximum (/Minimum), $f \in \mathcal{C}^2(D)$, $D \subset \mathbb{R}^n$.
Beh.: $H(f; a)$ ist negativ (/positiv) semidefinit. [Bew. vgl. 7.17]

7.14. **Bezeichnungen** (vgl. Lin. Algebra II): A symmetrische Bilinearform, dann heißt
A positiv (negativ) semidefinit $:\Leftrightarrow \forall h \in \mathbb{R}^n \setminus \{0\}: \langle h, Ah\rangle = h^T A h \left\{\begin{matrix}\ge 0\\ \le 0\end{matrix}\right\}$
A positiv (negativ) definit $:\Leftrightarrow$ " " " $\left\{\begin{matrix}> 0\\ < 0\end{matrix}\right\}$

dort "Kor. 6.5 7", SoSe'25

7.15. **EW-Kriterium für Definitheit** (vgl. Lineare Algebra II): Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch, d.h. $A^T = A$.
Dann gilt: A positiv (negativ) definit $\Leftrightarrow$ alle EWe > 0 (< 0).
A positiv (negativ) semidefinit $\Leftrightarrow$ alle EWe ≥ 0 (≤ 0).

↙ Lineare Algebra II SoSe'25, Satz 6.5.9.

7.15. Hauptminoren-Kriterium (auch: Hurwitz-Kriterium): Sei $A \in \mathbb{R}^{m\times m}$ symmetrisch. Dann:
A positiv definit $\Leftrightarrow$ alle Hauptminoren $\det A_k > 0$, wo A_k die Matrix sei, die aus den ersten k Zeilen und k Spalten von A besteht, $k \in \{1, \dots, m\}$.

7.16. • Bsp.: $g(t) = f(a + t(x-a))$ habe rel. Maximum in $t = 0$.
Dann $0 \geq g''(t) = h^T H(f;a) h$ mit $h = x - a$ laut 7.13.

• Bsp.: $D = \mathbb{R}^2$, $f\binom{x}{y} = y^2 + x^4 + x^3$. Dann ist $D_1 f\binom{x}{y} = 4x^3 + 3x^2$, $D_2 f\binom{x}{y} = 2y$.
Kritische Punkte: $(0,0)$, $(-\frac{3}{4}, 0)$, sind mögliche Extremstellen laut 7.11.

↙ für quadratische Formen: vgl. Lineare Algebra II, SoSe'25, Kor. 6.5.10.

7.17. Hinreichende Bedingung:
Vor.: $f \in \mathcal{C}^2(D)$, $D \subseteq \mathbb{R}^n$, a kritischer Punkt von f, $H(f;a)$ $\left\{\begin{matrix}\text{negativ}\\ \text{positiv}\end{matrix}\right\}$ definit.
Bel.: f hat in a striktes lokales $\left\{\begin{matrix}\text{Max}\\ \text{Min}\end{matrix}\right\}$.
Bew.: Aus 7.3, dem Satz von Taylor für $m = 2$, folgt:
$$f(a+h) = f(a) + \underbrace{0}_{\substack{\operatorname{grad} f(a) = 0 \\ \text{da Krit. Pkt.}}} + \tfrac{1}{2} h^T H(f;c) h \quad \text{für } c \in \overline{a(a+h)} \text{ geeignet.}$$

Da $H(f;c)$ stetig und c nahe bei a liegt (wenn h klein ist), ist auch $H(f;c)$ negativ definit.
Daher ist $f(a+h) < f(a)$, d.h. f hat in a ein striktes lokales Maximum. □
Bem.: Dieser Beweisansatz zeigt auch 7.17.

7.18. Bsp.: $D = \mathbb{R}^2$, $f\binom{x}{y} = y^2 + x^4 + x^3 \Rightarrow D_2 D_1 f\binom{x}{y} = 0$, $D_1^2 f\binom{x}{y} = 6x(2x+1)$, $D_2^2 f\binom{x}{y} = 2$.
$\Rightarrow H(f;0) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$ positiv semidefinit,
$H(f;(-\frac{3}{4}, 0)) = \begin{pmatrix} + & 0 \\ 0 & 2 \end{pmatrix}$ positiv definit $\Rightarrow$ in $(-\frac{3}{4}, 0)$ ex. ein striktes lok. Min.
Ferner ex. in 0 kein Extremum: Denn es ist
$$f\binom{x}{0} = x^3(x+1) \begin{Bmatrix} < 0 = f(0) \\ > 0 = f(0) \end{Bmatrix}, \text{ falls } \begin{Bmatrix} -1 < x < 0 \\ x > 0 \end{Bmatrix}.$$

7.19. Bem.: Ein Ausschluss von Extrema ist wie folgt möglich (Kor. aus 7.13):
Ist f auf $U \subseteq \mathbb{R}^n$ zweimal stetig diff'bar, $\operatorname{grad} f(a) = 0$, $H(f;a)$ indefinit, so hat f in a kein lokales Extremum.

7.20. <u>Praktisches Vorgehen:</u> • Als Extremstellen kommen die Punkte in Frage,
- die kritisch sind (d.h. grad f verschwindet dort),
- die Randpunkte von U sind (falls U nicht offen sein sollte),
- oder die singulär sind (wo f nicht diff'bar ist).

• Die kritischen Punkte ermittelt man durch Lösen des Gleichungssystems (i.a. nichtlin.) $\operatorname{grad} f(x) = 0$, welches n Gleichungen in n Unbekannten hat.
Dieses hat oft nur endlich viele Lösungen.
Der Vergleich der Funktionswerte dort reicht aber nicht aus, deshalb sind Kriterien nötig.

• Dann jeweils Hessematrix berechnen und Kriterien testen.
Hilft das nicht, müssen die Stellen anderweitig untersucht werden.

7.21. <u>Bsp.:</u> Betr. $f_1, f_2, f_3 : \mathbb{R}^2 \to \mathbb{R}$, $f_1(x,y) = x^2 + y^4$, $f_2(x,y) = x^2$, $f_3(x,y) = x^2 + y^3$.
Für alle $i \in \{1,2,3\}$ ist $\operatorname{grad} f_i(0) = 0$ und $H(f_i; 0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ positiv semidefinit.

- Die Fkt. f_1 hat in 0 ein striktes lokales Minimum. $\left(\forall \binom{x}{y} \neq 0 : f_1(x,y) = x^2 + y^2 > 0 = f_1(0)\right)$
- Die Fkt. f_2 hat in 0 ein lokales Minimum $\left(\forall \binom{x}{y} : f_2\binom{x}{y} \geq 0\right)$, das nicht strikt ist, denn in allen Punkten der y-Achse hat f_2 denselben Wert wie in 0. $(\forall y : f_2(0,y) = 0 = f_2(0,0).)$
- Die Fkt. f_3 hat in 0 weder ein lokales Minimum noch lokales Maximum.
$(\forall y < 0 : f_3(0,y) = y^3 < 0 = f_3(0), \quad \forall y > 0 : f_3(0,y) = y^3 > 0 = f_3(0).)$