2.8 Lokale Extreme in Mehrdimensionalen

Satz 122 Es sei $U \subseteq \mathbb{R}^k$ offen, $f: U \to \mathbb{R}$ und $\underline{a} \in U$. Wenn f bei $\underline{a}$ differenzierbar ist, existieren alle Richtungsableitungen $\frac{\partial f}{\partial \underline{u}}(\underline{a})$ (mit $\underline{u} \in \mathbb{R}^k$, $\|\underline{u}\| = 1$) und es gilt

$$\frac{\partial f}{\partial \underline{u}}(\underline{a}) = \langle \operatorname{grad} f(\underline{a}), \underline{u} \rangle.$$

Beweis: Es sei $g: \mathbb{R} \to \mathbb{R}^k$, $g(t) = \underline{a} + t\underline{u} = (a_1 + tu_1, \dots, a_k + tu_k)$ und $h(t) = f(\underline{a} + t\underline{u})$
(Da U offen ist, ist die Abbildung $h = f \circ g: I \to \mathbb{R}$ auf einem offenen Intervall $I \subseteq \mathbb{R}$ mit $0 \in I$ definiert.) Anwendung von Satz 113 liefert

$$\frac{\partial f}{\partial \underline{u}}(\underline{a}) = \lim_{t \to 0} \frac{f(\underline{a} + t\underline{u}) - f(\underline{a})}{t} = \lim_{t \to 0} \frac{h(t) - h(0)}{t} = h'(0) = (f \circ g)'(0)$$

$$\stackrel{\text{Satz 113}}{=} \sum_{i=1}^{k} \frac{\partial f}{\partial x_i}(g(0)) \cdot \frac{\partial g_i}{\partial t}(0) = \sum_{i=1}^{k} \frac{\partial f}{\partial x_i}(\underline{a}) \cdot u_i = \langle \operatorname{grad} f(\underline{a}), \underline{u} \rangle$$

Korollar 123 Es sei $U \subseteq \mathbb{R}^k$ offen, $f: U \to \mathbb{R}$, $\underline{a} \in U$ und f bei $\underline{a}$ differenzierbar.

(i) $\forall \underline{u} \in \mathbb{R}^k$, $\|\underline{u}\| = 1$ ist $\left| \frac{\partial f}{\partial \underline{u}}(\underline{a}) \right| \leq \|\operatorname{grad} f(\underline{a})\|$,

(ii) Ist $\operatorname{grad} f(\underline{a}) \neq \underline{0}$, so ist $\frac{\partial f}{\partial \underline{u}}(\underline{a}) = \|\operatorname{grad} f(\underline{a})\|$ genau dann wenn $\underline{u} = \frac{1}{\|\operatorname{grad} f(\underline{a})\|} \operatorname{grad} f(\underline{a})$

Beweis: (i) Mit Hilfe der Cauchy-Schwarzschen Ungleichung erhält man

$$\left| \frac{\partial f}{\partial \underline{u}}(\underline{a}) \right| \stackrel{\text{Satz 122}}{=} |\langle \operatorname{grad} f(\underline{a}), \underline{u} \rangle| \stackrel{\text{Satz 78}}{\leq} \|\operatorname{grad} f(\underline{a})\| \cdot \|\underline{u}\| = \|\operatorname{grad} f(\underline{a})\|.$$

(ii) Nach dem Zusatz von Satz 78 gilt $|\langle \operatorname{grad} f(\underline{a}), \underline{u} \rangle| = \|\operatorname{grad} f(\underline{a})\|$ genau dann wenn $\operatorname{grad} f(\underline{a})$ und $\underline{u}$ linear abhängig sind. Da $\operatorname{grad} f(\underline{a}) \neq \underline{0}$ und $\|\underline{u}\| = 1$ ist das genau dann der Fall, wenn $\underline{u} = \pm \frac{1}{\|\operatorname{grad} f(\underline{a})\|} \operatorname{grad} f(\underline{a})$. Da

$$\left\langle \operatorname{grad} f(\underline{a}), \pm \frac{1}{\|\operatorname{grad} f(\underline{a})\|} \operatorname{grad} f(\underline{a}) \right\rangle = \pm \frac{1}{\|\operatorname{grad} f(\underline{a})\|} \langle \operatorname{grad} f(\underline{a}), \operatorname{grad} f(\underline{a}) \rangle$$

$$= \pm \frac{1}{\|\operatorname{grad} f(\underline{a})\|} \|\operatorname{grad} f(\underline{a})\|^2 = \pm \|\operatorname{grad} f(\underline{a})\|.$$

Daraus folgt, dass

$$\frac{\partial f}{\partial \underline{u}}(\underline{a}) = \|\operatorname{grad} f(\underline{a})\| \iff \langle \operatorname{grad} f(\underline{a}), \underline{u} \rangle = \|\operatorname{grad} f(\underline{a})\| \iff \underline{u} = \frac{1}{\|\operatorname{grad} f(\underline{a})\|} \operatorname{grad} f(\underline{a}).$$

Bemerkung: Korollar 123 (ii) besagt, dass $\operatorname{grad} f(\underline{a})$ in die Richtung des stärksten Anstiegs von f bei $\underline{a}$ zeigt.

Bemerkung: Wir wollen im Folgenden einen bekannten Satz aus der eindimensionalen reellen Analysis verallgemeinern: Es sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $a \in I$ und $f: I \to \mathbb{R}$ zweimal differenzierbar. Dann gelten:

(i) Besitzt f bei a ein lokales Extremum, so ist $f'(a) = 0$,

(ii) Ist $f'(a) = 0$ und $f''(a) < 0$ (bzw. $f''(a) > 0$), so besitzt f bei a ein lokales Maximum (bzw. lokales Minimum).

Definition: Es sei $U \subseteq \mathbb{R}^k$ offen, $f: U \to \mathbb{R}$ und $\underline{a} \in U$.

1) f besitzt bei $\underline{a}$ ein (striktes) lokales Maximum, wenn

$\exists \varepsilon > 0: \; f(\underline{x}) < f(\underline{a}) \quad \forall \underline{x} \in B_\varepsilon(\underline{a}) \setminus \{\underline{a}\}$,

2) f besitzt bei $\underline{a}$ ein (striktes) lokales Minimum, wenn

$\exists \varepsilon > 0: \; f(\underline{x}) > f(\underline{a}) \quad \forall \underline{x} \in B_\varepsilon(\underline{a}) \setminus \{\underline{a}\}$,

3) f besitzt bei $\underline{a}$ ein (striktes) lokales Extremum, wenn f bei $\underline{a}$ entweder ein (striktes) lokales Maximum oder ein (striktes) lokales Minimum besitzt.

Satz 124 Es sei $U \subseteq \mathbb{R}^k$ offen, $f: U \to \mathbb{R}$ differenzierbar und $\underline{a} \in U$. Besitzt f bei $\underline{a}$ ein lokales Extremum, so gilt $\frac{\partial f}{\partial x_1}(\underline{a}) = \dots = \frac{\partial f}{\partial x_k}(\underline{a}) = 0$, d.h. $\operatorname{grad} f(\underline{a}) = \underline{0}$.

Beweis: Für jedes $\underline{u} \in \mathbb{R}^k$ mit $\|\underline{u}\| = 1$ ist die Abbildung $h: I \to \mathbb{R}$, $h(t) = f(\underline{a} + t\underline{u})$ auf einem offenen Intervall $I \subseteq \mathbb{R}$ mit $0 \in I$ definiert. Offensichtlich besitzt h bei $t = 0$ ein lokales Extremum. Wegen des oben beschriebenen Resultats aus der eindimensionalen reellen Analysis folgt wie im Beweis von Satz 122

$$0 = h'(0) = \lim_{t \to 0} \frac{1}{t}(h(t) - h(0)) = \lim_{t \to 0} \frac{1}{t}\left(f(\underline{a} + t\underline{u}) - f(\underline{a})\right)$$

$$= \frac{\partial f}{\partial \underline{u}}(\underline{a}) = \langle \operatorname{grad} f(\underline{a}), \underline{u} \rangle \quad \text{für alle } \underline{u} \in \mathbb{R}^k \text{ mit } \|\underline{u}\| = 1.$$

Insbesondere $\frac{\partial f}{\partial x_i}(\underline{a}) = \frac{\partial f}{\partial \underline{e}_i}(\underline{a}) = \langle \operatorname{grad} f(\underline{a}), \underline{e}_i \rangle = 0$ für $1 \le i \le k$.

Definition: Es sei $U \subseteq \mathbb{R}^k$ offen, $f \in C^2(U)$ und $\underline{a} \in U$. Die Matrix

$$H_f(\underline{a}) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{a}) \right)_{1 \le i,j \le k} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\underline{a}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_k}(\underline{a}) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_k}(\underline{a}) & \cdots & \frac{\partial^2 f}{\partial x_k^2}(\underline{a}) \end{pmatrix}$$

wird Hesse-Matrix von f bei $\underline{a}$ genannt.

Bemerkung: Da f eine C^2-Funktion ist, ist die Hesse-Matrix nach dem Satz von Schwarz (Satz 106) symmetrisch und definiert daher eine quadratische Form, nämlich $\underline{x} \longmapsto \underline{x}^T \cdot H_f(\underline{a}) \cdot \underline{x} = \langle \underline{x}, H_f(\underline{a})\, \underline{x} \rangle$

Satz 125 Es sei $U \subseteq \mathbb{R}^k$ offen, $f \in C^2(U)$, $\underline{a} \in U$ und $\operatorname{grad} f(\underline{a}) = \underline{0}$.

(i) Ist $H_f(\underline{a})$ positiv definit, so besitzt f bei $\underline{a}$ ein (striktes) lokales Minimum,

(ii) Ist $H_f(\underline{a})$ negativ definit, so besitzt f bei $\underline{a}$ ein (striktes) lokales Maximum,

(iii) Ist $H_f(\underline{a})$ indefinit, so besitzt f bei $\underline{a}$ kein lokales Extremum.

Beweis: Es gibt ein $\tilde{\varepsilon} > 0$, derart dass $B_{\tilde{\varepsilon}}(\underline{a}) \subseteq U$. Es sei nun $\underline{v} \in \mathbb{R}^k \setminus \{\underline{0}\}$ und $\|\underline{v}\| < \tilde{\varepsilon}$. Es sei wieder $h: I \longrightarrow \mathbb{R}$, $h(t) = f(\underline{a} + t\underline{v})$ für ein passendes offenes Intervall $I \subseteq \mathbb{R}$ mit $0 \in I$. Zweimaliges Anwenden der Kettenregel liefert

$$h'(t) = \sum_{i=1}^{k} \frac{\partial f}{\partial x_i}(\underline{a} + t\underline{v}) \cdot v_i = \langle \operatorname{grad} f(\underline{a} + t\underline{v}), \underline{v} \rangle$$

und

$$h''(t) = \sum_{i=1}^{k} \left(\sum_{j=1}^{k} \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(\underline{a} + t\underline{v}) \cdot v_j \right) v_i = \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{a} + t\underline{v}) \cdot v_i v_j$$

$$= \langle \underline{v}, H_f(\underline{a} + t\underline{v}) \cdot \underline{v} \rangle$$

Nach dem (eindimensionalen) Satz von Taylor gibt es ein $\xi \in (0,1)$, derart dass

$$f(\underline{a} + \underline{v}) = h(1) = h(0) + h'(0) + \frac{1}{2} h''(\xi)$$

$$= f(\underline{a}) + \underbrace{\langle \underbrace{\operatorname{grad} f(\underline{a})}_{=\underline{0}}, \underline{v} \rangle}_{=0} + \frac{1}{2} \langle \underline{v}, H_f(\underline{a} + \xi\underline{v}) \cdot \underline{v} \rangle$$

$$= f(\underline{a}) + \frac{1}{2} \langle \underline{v}, H_f(\underline{a} + \xi\underline{v}) \cdot \underline{v} \rangle$$

(i) Da f eine C^2-Funktion ist, gibt es ein $\varepsilon \in (0, \tilde{\varepsilon}]$, derart dass $H_f(\underline{b})$ ebenfalls positiv definit ist, wenn $\|\underline{b} - \underline{a}\| < \varepsilon$. (Diesen Schritt kann und sollte man genauer begründen. Wir verzichten aus Zeitgründen darauf.)

Ist nun $\|\underline{v}\| < \varepsilon$, so ist $H_f(\underline{a} + \xi\underline{v})$ positiv definit und daher

$$f(\underline{a} + \underline{v}) = f(\underline{a}) + \frac{1}{2} \underbrace{\langle \underline{v}, H_f(\underline{a} + \xi\underline{v}) \cdot \underline{v} \rangle}_{>0} > f(\underline{a}).$$

Da $\underline{v} \in B_\varepsilon(\underline{a})$ beliebig war, besitzt f bei $\underline{a}$ ein (striktes) lokales Minimum.

(ii) und (iii) beweist man analog.

<u>Lemma 126</u> Die quadratische Form $q: \mathbb{R}^2 \to \mathbb{R}$ sei gegeben durch

$$q(x,y) = (x \;\; y)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2bxy + cy^2 \quad (\text{mit } a,b,c \in \mathbb{R}).$$

Ist $\Delta_2 = \begin{vmatrix} a & b \\ b & c \end{vmatrix} = ac - b^2 < 0$, so ist q indefinit.

<u>Beweis:</u> Ist $a \neq 0$, so haben $q(1,0) = a$ und $q(-b,a) = ab^2 - 2ab^2 + ca^2 = a\underbrace{(ac-b^2)}_{<0}$ verschiedenes Vorzeichen.

Ist $c \neq 0$, so haben $q(0,1) = c$ und $q(c,-b) = ac^2 - 2b^2c + b^2c = \underbrace{(ac-b^2)}_{<0}c$ verschiedenes Vorzeichen.

Ist $a = c = 0$, so ist $b \neq 0$, da $-b^2 = \Delta_2 < 0$ und $q(1,1) = 2b$ und $q(1,-1) = -2b$ haben verschiedenes Vorzeichen.

<u>Korollar 127</u> Es sei $U \subseteq \mathbb{R}^2$ offen, $f \in C^2(U)$, $\underline{a} \in U$ und $\operatorname{grad} f(\underline{a}) = \underline{0}$.

(i) Wenn $\frac{\partial^2 f}{\partial x^2}(\underline{a}) > 0$ und $\frac{\partial^2 f}{\partial x^2}(\underline{a}) \cdot \frac{\partial^2 f}{\partial y^2}(\underline{a}) - \left(\frac{\partial^2 f}{\partial x \partial y}(\underline{a})\right)^2 > 0$,

besitzt f bei $\underline{a}$ ein (striktes) lokales Minimum.

(ii) Wenn $\frac{\partial^2 f}{\partial x^2}(\underline{a}) < 0$ und $\frac{\partial^2 f}{\partial x^2}(\underline{a}) \cdot \frac{\partial^2 f}{\partial y^2}(\underline{a}) - \left(\frac{\partial^2 f}{\partial x \partial y}(\underline{a})\right)^2 > 0$,

besitzt f bei $\underline{a}$ ein (striktes) lokales Maximum.

(iii) Wenn $\frac{\partial^2 f}{\partial x^2}(\underline{a}) \cdot \frac{\partial^2 f}{\partial y^2}(\underline{a}) - \left(\frac{\partial^2 f}{\partial x \partial y}(\underline{a})\right)^2 < 0$,

besitzt f bei $\underline{a}$ kein lokales Extremum.

<u>Beweis:</u> Die Hauptminoren der Hesse-Matrix

$$H_f(\underline{a}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(\underline{a}) & \frac{\partial^2 f}{\partial x \partial y}(\underline{a}) \\ \frac{\partial^2 f}{\partial x \partial y}(\underline{a}) & \frac{\partial^2 f}{\partial y^2}(\underline{a}) \end{pmatrix}$$

sind $\Delta_1 = \frac{\partial^2 f}{\partial x^2}(\underline{a})$ und $\Delta_2 = \frac{\partial^2 f}{\partial x^2}(\underline{a}) \cdot \frac{\partial^2 f}{\partial y^2}(\underline{a}) - \left(\frac{\partial^2 f}{\partial x \partial y}(\underline{a})\right)^2$.

(i) Folgt aus Korollar 120 und Satz 125 (i).

(ii) Folgt aus Korollar 121 und Satz 125 (ii).

(iii) Folgt aus Lemma 126 und Satz 125 (iii).

Beispiel: Die Funktion $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \log(x^2+y^2+1)$ besitzt bei $(0,0)$ offensichtlich ein Minimum. Wir wollen das mit Hilfe von Korollar 127 überprüfen: Es ist

$$\frac{\partial f}{\partial x}(x,y) = \frac{2x}{x^2+y^2+1} \quad \text{und} \quad \frac{\partial f}{\partial y}(x,y) = \frac{2y}{x^2+y^2+1}. \quad \text{Aus } \frac{\partial f}{\partial x}(x,y) = \frac{\partial f}{\partial y}(x,y) = 0$$

folgt daher $(x,y) = (0,0)$. Die partiellen Ableitungen zweiter Ordnung sind

$$\frac{\partial^2 f}{\partial x^2}(x,y) = \frac{2(x^2+y^2+1) - 4x^2}{(x^2+y^2+1)^2} = \frac{-2x^2+2y^2+2}{(x^2+y^2+1)^2}$$

und (aus Symmetriegründen)

$$\frac{\partial^2 f}{\partial y^2}(x,y) = \frac{2x^2-2y^2+2}{(x^2+y^2+1)^2}$$

sowie

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = -\frac{(2x)\cdot(2y)}{(x^2+y^2+1)^2} = -\frac{4xy}{(x^2+y^2+1)^2}.$$

Also sind

$$\Delta_1 = \frac{\partial^2 f}{\partial x^2}(0,0) = 2 > 0 \quad \text{und} \quad \Delta_2 = \begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(0,0) & \frac{\partial^2 f}{\partial x \partial y}(0,0) \\ \frac{\partial^2 f}{\partial x \partial y}(0,0) & \frac{\partial^2 f}{\partial y^2}(0,0) \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4 > 0.$$

Aus Korollar 127 (i) folgt nun (wiederum), dass f bei $(0,0)$ ein (striktes) Minimum besitzt.

Bemerkung: Im Fall $k=1$ wird die Hesse-Matrix zur „1×1-Matrix" $f''(a)$ und die quadratische Form $\underline{x} \mapsto \underline{x}^T \cdot H_f(\underline{a}) \cdot \underline{x}$ wird zu $x \mapsto f''(a)\,x^2$. Diese ist genau dann positiv definit (bzw. negativ definit) wenn $f''(a) > 0$ (bzw. $f''(a) < 0$). (Der Fall, dass diese quadratische Form indefinit ist, kann für $k=1$ nicht auftreten.)