2.8 Lokale Extrema im Mehrdimensionalen

Satz 123 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 $h = f \circ g$ auf einem offenen Intervall $I \subseteq \mathbb{R}$ mit $0 \in I$ definiert.) Anwendung der Kettenregel (Satz 116) gibt

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

$$\overset{\text{Satz 116}}{=} \sum_{i=1}^{k} \frac{\partial f}{\partial x_i}(g(0)) \cdot \frac{\partial g_i}{\partial t}(0) = \sum_{i=1}^{k} \frac{\partial f}{\partial x_i}(\underline{a}) \cdot u_i = \langle \operatorname{grad} f(\underline{a}), \underline{u} \rangle$$

Korollar 124 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 $\left| \frac{\partial f}{\partial \underline{u}}(\underline{a}) \right| \le \|\operatorname{grad} f(\underline{a})\|$,

(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 (Korollar 81) erhält man

$$\left| \frac{\partial f}{\partial \underline{u}}(\underline{a}) \right| \overset{\text{Satz 123}}{=} |\langle \operatorname{grad} f(\underline{a}), \underline{u} \rangle| \overset{\text{Kor. 81}}{\le} \|\operatorname{grad} f(\underline{a})\| \cdot \|\underline{u}\| = \|\operatorname{grad} f(\underline{a})\|$$

(ii) Nach dem Zusatz von Satz 79 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 $\|\underline{u}\| = 1$ gilt das genau dann wenn $\underline{u} = \pm \frac{1}{\|\operatorname{grad} f(\underline{a})\|} \operatorname{grad} f(\underline{a})$. Aus

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

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

folgt

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

Bemerkung: Korollar 124 (ii) besagt, dass $\operatorname{grad} f(\underline{a})$ in die Richtung des stärksten Anstiegs von f bei $\underline{a}$ zeigt.

Bemerkung: Wir wollen nun die folgenden bekannten Resultate aus der eindimensionalen reellen Analysis verallgemeinern: Es sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $f: I \to \mathbb{R}$ zweimal differenzierbar und $a \in I$. 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 $\underline{a}$ ein (striktes) lokales Maximum oder ein (striktes) lokales Minimum besitzt.

Satz 125 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. Offenbar besitzt h bei $t = 0$ ein lokales Extremum. Wegen des oben beschriebenen Resultats aus der eindimensionalen reellen Analysis folgt

$$0 = h'(0) = \lim_{t \to 0} \frac{h(t) - h(0)}{t} = \lim_{t \to 0} \frac{f(\underline{a} + t\underline{u}) - f(\underline{a})}{t} = \frac{\partial f}{\partial \underline{u}}(\underline{a}) \overset{\text{Satz 123}}{=} \langle \operatorname{grad} f(\underline{a}), \underline{u} \rangle \quad \text{für } u \in \mathbb{R}^k, \ \|\underline{u}\| = 1.$$

Insbesondere ist $\frac{\partial f}{\partial x_i}(\underline{a}) = \frac{\partial f}{\partial \underline{e}_i}(\underline{a}) = \langle \operatorname{grad} f(\underline{a}), \underline{e}_i \rangle = 0$ 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 (Kor. 110) symmetrisch und definiert daher die quadratische Form

$\underline{x} \mapsto \underline{x}^T \cdot H_f(\underline{a}) \cdot \underline{x} = \langle \underline{x}, H_f(\underline{a}) \cdot \underline{x} \rangle$

Satz 126 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 $\varepsilon > 0$, sodass $B_\varepsilon(\underline{a}) \subseteq U$. Es sei $\underline{v} \in \mathbb{R}^k \setminus \{\underline{0}\}$ und $\|\underline{v}\| < \varepsilon$. Weiters sei wieder $h: I \to \mathbb{R}$, $h(t) = f(\underline{a} + t\underline{v})$ für ein passendes offenes Intervall $I \subseteq \mathbb{R}$ mit $0 \in I$. Zweimaliges anwenden der Kettenregel gibt

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

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 $\xi \in (0,1)$, sodass

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

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

(i) Da $f \in C^2(U)$ gibt es $\delta \in (0, \varepsilon)$, sodass $H_f(\underline{b})$ ebenfalls positiv definit ist, wenn $\|\underline{b} - \underline{a}\| < \delta$.
Ist nun $\|\underline{v}\| < \delta$, 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 $\underline{v} \in B_\delta(\underline{a}) \setminus \{\underline{0}\}$ beliebig war, besitzt f bei $\underline{a}$ ein (striktes) lokales Minimum.

(ii) und (iii) Beweist man analog.

14.1.2025

Korollar 127 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.

Beweis: 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 Satz 118 (i) (bzw. Satz 119) und Satz 126 (i)

(ii) Folgt aus Satz 118 (ii) (bzw. Korollar 120) und Satz 126 (ii).

(iii) Folgt aus Satz 121 und Satz 126 (iii)

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

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

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

und $\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = -\frac{(2x)(2y)}{(x^2+y^2+1)^2} = \frac{-4xy}{(x^2+y^2+1)^2}$. Daher sind $\Delta_1 = \frac{\partial^2 f}{\partial x^2}(0,0) = 2 > 0$

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

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 $x \mapsto f''(a)\cdot x^2$. 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 eintreten.)