

## Mathematik für Anwender II

### Vorlesung 52

#### Hinreichende Kriterien für lokale Extrema

Wir kommen jetzt zu hinreichenden Kriterien für die Existenz von lokalen Extrema einer Funktion

$$f: G \longrightarrow \mathbb{R},$$

die auf Eigenschaften der zweiten Richtungsableitungen, genauer der Hesse-Form, beruhen und die entsprechenden Kriterien in einer Variablen verallgemeinern. Zunächst brauchen wir ein Lemma, das beschreibt, wie die Definitheit (oder der „Definitheitstyp“) der Hesse-Form vom Punkt abhängt.

LEMMA 52.1. *Es sei  $V$  ein endlichdimensionaler reeller Vektorraum,  $G \subseteq V$  eine offene Teilmenge und*

$$f: G \longrightarrow \mathbb{R}$$

*eine zweimal stetig differenzierbare Funktion. Es sei  $P \in G$  ein Punkt, in dem die Hesse-Form  $\text{Hess}_P f$  positiv (negativ) definit sei. Dann gibt es eine offene Umgebung  $U$ ,  $P \in U \subseteq G$ , derart, dass die Hesse-Form  $\text{Hess}_Q f$  in jedem Punkt  $Q \in U$  positiv (negativ) definit ist.*

*Beweis.* Sei  $v_1, \dots, v_n$  eine Basis von  $V$ , und sei  $H(Q)$  die Gramsche Matrix zur Hesse-Form  $\text{Hess}_Q f$  im Punkt  $Q \in G$  bezüglich dieser Basis. Aufgrund der Differenzierbarkeitsvoraussetzungen hängt  $H(Q)$  stetig von  $Q$  ab. Daher hängen auch die Determinanten der quadratischen Untermatrizen von  $H(Q)$  stetig von  $Q$  ab. Die Determinanten

$$D_k(P) = \det((H(P)_{i,j})_{1 \leq i,j \leq k})$$

sind nach Korollar 44.13 alle von 0 verschieden. Daher gibt es eine offene Umgebung  $U$ ,  $P \in U \subseteq G$ , derart, dass für alle  $Q \in U$  die Determinanten

$$D_k(Q) = \det((H(Q)_{i,j})_{1 \leq i,j \leq k})$$

das gleiche Vorzeichen haben wie  $D_k(P)$ . Da diese Vorzeichen nach Korollar 44.13 über die Definitheit entscheiden, folgt die Behauptung.  $\square$

SATZ 52.2. *Es sei  $V$  ein endlichdimensionaler reeller Vektorraum,  $G \subseteq V$  eine offene Teilmenge und*

$$f: G \longrightarrow \mathbb{R}$$

*eine zweimal stetig differenzierbare Funktion. Es sei  $P \in G$  mit  $(Df)_P = 0$ . Dann gelten folgende Aussagen.*

- (1) *Wenn  $\text{Hess}_P f$  negativ definit ist, so besitzt  $f$  ein isoliertes lokales Maximum in  $P$ .*

- (2) Wenn  $\text{Hess}_P f$  positiv definit ist, so besitzt  $f$  ein isoliertes lokales Minimum in  $P$ .
- (3) Wenn  $\text{Hess}_P f$  indefinit ist, so besitzt  $f$  in  $P$  weder ein lokales Minimum noch ein lokales Maximum.

*Beweis.* (1). Aufgrund von Lemma 52.1 gibt es ein  $\delta > 0$  derart, dass die Hesse-Form  $\text{Hess}_Q f$  für alle  $Q \in U(P, \delta)$  negativ definit ist. Für alle Vektoren  $v \in V$ ,  $v \in U(0, \delta)$ , gibt es nach Satz 50.5 ein  $c = c(v) \in [0, 1]$  mit

$$f(P + v) = f(P) + \sum_{|r|=2} \frac{1}{r!} D^r f(P + cv) \cdot v^r = f(P) + \frac{1}{2} \text{Hess}_{P+cv} f(v, v),$$

wobei die erste Formulierung sich auf eine fixierte Basis bezieht und wobei die zweite Identität auf Aufgabe 52.15 beruht. Da die Hesse-Form negativ definit ist, steht rechts für  $v \neq 0$  eine Zahl, die echt kleiner als  $f(P)$  ist. Daher liegt ein isoliertes lokales Maximum vor. (2) wird wie (1) bewiesen oder durch betrachten von  $-f$  darauf zurückgeführt. (3). Sei  $\text{Hess}_P f$  indefinit. Dann gibt es Vektoren  $v$  und  $w$  mit

$$\text{Hess}_P f(v, v) > 0 \text{ und } \text{Hess}_P f(w, w) < 0.$$

Wegen der stetigen Abhängigkeit der Hesse-Form gelten diese Abschätzungen auch für  $\text{Hess}_Q f$  für  $Q$  aus einer offenen Umgebung von  $P$  (mit den gleichen Vektoren  $v$  und  $w$ ). Wir können durch Skalierung von  $v$  und  $w$  annehmen, dass  $P + v$  und  $P + w$  zu dieser Umgebung gehören. Wie im Beweis zu Teil (1) gilt daher ( $v$  und  $w$  sind nicht 0)

$$f(P + v) = f(P) + \frac{1}{2} \text{Hess}_{P+cv} f(v, v) > f(P)$$

und

$$f(P + w) = f(P) + \frac{1}{2} \text{Hess}_{P+dw} f(w, w) < f(P)$$

mit  $c, d \in [0, 1]$ . Also kann in  $P$  kein lokales Extremum vorliegen.  $\square$

BEISPIEL 52.3. Wir betrachten die Funktion

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}, (x, y) \longmapsto x + 3x^2 - 2xy - y^2 + y^3.$$

Die partiellen Ableitungen sind

$$\frac{\partial f}{\partial x} = 1 + 6x - 2y \text{ und } \frac{\partial f}{\partial y} = -2x - 2y + 3y^2.$$

Zur Berechnung der kritischen Punkte dieser Funktion eliminieren wir  $x$  und erhalten die Bedingung

$$9y^2 - 8y + 1 = 0,$$

die zu

$$y = \frac{\pm\sqrt{7}}{9} + \frac{4}{9}$$

führt. Die kritischen Punkte sind also

$$P_1 = \left( \frac{2\sqrt{7}-1}{54}, \frac{\sqrt{7}}{9} + \frac{4}{9} \right) \text{ und } P_2 = \left( \frac{-2\sqrt{7}-1}{54}, \frac{-\sqrt{7}}{9} + \frac{4}{9} \right).$$

Die Hesse-Form ist in einem Punkt  $Q = (x, y)$  gleich

$$\text{Hess}_Q f = \begin{pmatrix} 6 & -2 \\ -2 & -2 + 6y \end{pmatrix}.$$

Zur Bestimmung des Definitheitstyps ziehen wir Korollar 44.13 heran, wobei der erste Minor, also 6, natürlich positiv ist. Die Determinante der Hesse-Matrix ist

$$-16 + 36y,$$

was genau bei  $y > \frac{4}{9}$  positiv ist. Dies ist im Punkt  $P_1$  der Fall, aber nicht im Punkt  $P_2$ . Daher ist die Hesse-Matrix im Punkt  $P_1$  nach Korollar 44.13 positiv definit und somit besitzt die Funktion  $f$  im Punkt  $P_1$  nach Satz 52.2 ein isoliertes lokales Minimum, das zugleich ein globales Minimum ist. In  $P_2$  ist die Determinante negativ, so dass dort die Hesse-Form indefinit ist und somit, wiederum nach Satz 52.2, kein Extremum vorliegen kann.

BEISPIEL 52.4. Wir betrachten die Abbildung

$$\varphi: \mathbb{R}_+ \times \mathbb{R} \longrightarrow \mathbb{R}, (x, y) \longmapsto x^y.$$

Es ist

$$x^y = e^{(\ln x) \cdot y}.$$

Die partiellen Ableitungen sind

$$\frac{\partial \varphi}{\partial x} = \frac{y}{x} \cdot e^{(\ln x) \cdot y} = \frac{y}{x} \cdot x^y \text{ und } \frac{\partial \varphi}{\partial y} = (\ln x) \cdot e^{(\ln x) \cdot y} = (\ln x) \cdot x^y.$$

Da die Exponentialfunktion stets positiv ist, ist  $P = (1, 0)$  der einzige kritische Punkt. Die Hesse-Matrix in einem Punkt  $(x, y)$  ist

$$\begin{pmatrix} \frac{-y+y^2}{x^2} \cdot e^{(\ln x) \cdot y} & \frac{1+y \ln x}{x} \cdot e^{(\ln x) \cdot y} \\ \frac{1+y \ln x}{x} \cdot e^{(\ln x) \cdot y} & (\ln x)^2 \cdot e^{(\ln x) \cdot y} \end{pmatrix} = \begin{pmatrix} \frac{-y+y^2}{x^2} \cdot x^y & \frac{1+y \ln x}{x} \cdot x^y \\ \frac{1+y \ln x}{x} \cdot x^y & (\ln x)^2 \cdot x^y \end{pmatrix}.$$

In  $P$  ist dies

$$\text{Hess}_P \varphi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Nach Korollar 44.13 ist daher die Hesse-Form im kritischen Punkt weder positiv definit noch negativ definit. Man kann direkt zeigen, dass diese Matrix indefinit ist (vom Typ  $(1, 1)$ ), da diese Bilinearform auf  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  positiv und auf

$\begin{pmatrix} 1 \\ -1 \end{pmatrix}$  negativ definit ist. Nach Satz 52.2 liegt in diesem Punkt also kein Extremum vor.

Dies kann man auch ohne Differentialrechnung erkennen. Für  $x = 1$  oder  $y = 0$  ist  $x^y = 1$ . Ansonsten gelten die folgenden Beziehungen.

- (1) Für  $0 < x < 1$  und  $y > 0$  ist  $x^y < 1$ .
- (2) Für  $x > 1$  und  $y > 0$  ist  $x^y > 1$ .
- (3) Für  $0 < x < 1$  und  $y < 0$  ist  $x^y > 1$ .
- (4) Für  $x > 1$  und  $y < 0$  ist  $x^y < 1$ .

Daher gibt es in jeder Umgebung von  $(1, 0)$  Punkte, an denen die Funktionswerte größer bzw. kleiner als 1 sind.

BEMERKUNG 52.5. Es sei

$$g: [a, b] \longrightarrow \mathbb{R}$$

eine stetige Funktion und

$$a = x_0 < x_1 < x_2 < \dots < x_n < b = x_{n+1}$$

eine Unterteilung des Intervalls durch  $n$  Zwischenpunkte (in  $n + 1$  Teilintervalle). Dazu gehört die Treppenfunktion, die auf  $[x_i, x_{i+1}[$  den konstanten Wert  $g(x_i)$  annimmt. Wenn  $g$  monoton wachsend ist, so ist dies eine untere Treppenfunktion, und das zugehörige Treppenintegral ist eine untere Schranke für das bestimmte Integral  $\int_a^b g(t)dt$ . Das Treppenintegral ist durch

$$f(x_1, \dots, x_n) = \sum_{i=0}^n g(x_i)(x_{i+1} - x_i)$$

gegeben. Wir fragen uns, für welche Intervallunterteilung mit  $n$  Teilpunkten das Treppenintegral maximal oder minimal wird. Dazu kann man die differentiellen Methoden zur Bestimmung von Extrema für Funktionen in mehreren Variablen verwenden (nämlich den variablen Unterteilungspunkten  $x_1, \dots, x_n$ ), vorausgesetzt, dass  $g$  (hinreichend oft) differenzierbar (in einer Variablen) ist. In diesem Fall sind die partiellen Ableitungen von  $f$  gleich

$$\frac{\partial f}{\partial x_i} = g'(x_i)(x_{i+1} - x_i) - g(x_i) + g(x_{i-1})$$

für  $i = 1, \dots, n$  (wobei  $x_0 = a$  und  $x_{n+1} = b$  zu lesen ist). Als Definitionsbereich von  $f$  kann man die offene Menge

$$\{(x_1, \dots, x_n) \mid a < x_1 < x_2 < \dots < x_n < b\} \subseteq \mathbb{R}^n$$

oder aber  $[a, b]^n$  wählen. Es ist im Allgemeinen schwierig, die kritischen Punkte dieser Abbildung zu bestimmen.

BEISPIEL 52.6. Wir wollen für die Funktion

$$g: \mathbb{R} \longrightarrow \mathbb{R}, t \longmapsto g(t) = 1 - t^3,$$

und das Einheitsintervall  $[0, 1]$  bestimmen, für welche zwei Unterteilungspunkte  $0 < x < y < 1$  das Treppenintegral der zugehörigen (dreistufigen) unteren Treppenfunktion maximal wird. Das Treppenintegral wird durch die Funktion

$$\begin{aligned} f(x, y) &= x(1 - x^3) + (y - x)(1 - y^3) \\ &= x - x^4 + y - y^4 - x + xy^3 \end{aligned}$$

$$= -x^4 + y - y^4 + xy^3$$

beschrieben. Die partiellen Ableitungen dieser Funktion sind

$$\frac{\partial f}{\partial x} = -4x^3 + y^3$$

und

$$\frac{\partial f}{\partial y} = 1 - 4y^3 + 3xy^2.$$

Wir bestimmen die kritischen Punkte. Aus der ersten partiellen Ableitung ergibt sich die Bedingung

$$y = \sqrt[3]{4x}$$

und daraus ergibt sich mit der zweiten partiellen Ableitung die Bedingung

$$1 - 16x^3 + 3 \cdot 4^{2/3}x^3 = 0,$$

also

$$(16 - 3 \cdot 4^{2/3})x^3 = 1$$

bzw.

$$x = \frac{1}{\sqrt[3]{16 - 3 \cdot 4^{2/3}}}.$$

Somit ist

$$P = \left( \frac{1}{\sqrt[3]{16 - 3 \cdot 4^{2/3}}}, \frac{\sqrt[3]{4}}{\sqrt[3]{16 - 3 \cdot 4^{2/3}}} \right) \cong (0,4911, 0,7796)$$

der einzige kritische Punkt. Wir bestimmen die Hesse-Matrix in diesem Punkt, sie ist

$$\text{Hess}_P f = \begin{pmatrix} -12x^2 & 3y^2 \\ 3y^2 & -12y^2 + 6xy \end{pmatrix}$$

und in  $P$  gleich

$$\begin{pmatrix} -2,8942 & 1,8233 \\ 1,8233 & -4,9961 \end{pmatrix},$$

also negativ definit nach Korollar 44.13. Daher liegt in  $P$  ein Maximum nach Satz 52.2 vor.

BEISPIEL 52.7. Wir wollen für die Funktion

$$g: \mathbb{R} \longrightarrow \mathbb{R}, t \longmapsto t,$$

und das Einheitsintervall  $[0, 1]$  bestimmen, für welche  $n$  Unterteilungspunkte  $0 < x_1 < \dots < x_n < 1$  das Treppenintegral der zugehörigen  $((n + 1)$ -stufigen) unteren Treppenfunktion maximal wird. Das Treppenintegral wird durch die Funktion

$$\begin{aligned} f(x_1, \dots, x_n) &= x_1(x_2 - x_1) + x_2(x_3 - x_2) + \dots \\ &\quad + x_{n-1}(x_n - x_{n-1}) + x_n(1 - x_n) \\ &= \sum_{i=1}^{n-1} x_i x_{i+1} + x_n - \sum_{i=1}^n x_i^2 \end{aligned}$$

beschrieben. Die partiellen Ableitungen dieser Funktion sind

$$\frac{\partial f}{\partial x_1} = x_2 - 2x_1,$$

$$\frac{\partial f}{\partial x_i} = x_{i-1} + x_{i+1} - 2x_i$$

für  $i = 2, \dots, n-1$  und

$$\frac{\partial f}{\partial x_n} = x_{n-1} + 1 - 2x_n.$$

Wir bestimmen die kritischen Punkte, indem wir die partiellen Ableitungen gleich 0 setzen. Die ersten  $n-1$  Gleichungen ergeben sukzessive die Bedingungen

$$x_i = ix_1$$

für alle  $i$ . Dies zeigt man durch Induktion, der Induktionsanfang ( $i = 1$ ) ist trivial,  $i = 2$  folgt direkt aus der ersten Gleichung und der Induktionsschritt ergibt sich aus

$$x_{i+1} = -x_{i-1} + 2x_i = -(i-1)x_1 + 2ix_1 = (i+1)x_1.$$

Aus der letzten Gleichung folgt schließlich

$$0 = x_{n-1} + 1 - 2x_n = 1 + (n-1-2n)x_1 = 1 - (n+1)x_1$$

und somit  $x_1 = \frac{1}{n+1}$ . Der einzige kritische Punkt liegt also in der äquidistanten Unterteilung vor. Die Hesse-Matrix ist (unabhängig vom Punkt) gleich

$$\begin{pmatrix} -2 & 1 & 0 & \dots & \dots & 0 \\ 1 & -2 & 1 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -2 & 1 \\ 0 & \dots & \dots & 0 & 1 & -2 \end{pmatrix}.$$

Diese Matrix ist negativ definit nach Korollar 44.13. Daher liegt in der äquidistanten Unterteilung nach Satz 52.2 das Maximum vor.

## Abbildungsverzeichnis

- Erläuterung: Die in diesem Text verwendeten Bilder stammen aus Commons (also von <http://commons.wikimedia.org>) und haben eine Lizenz, die die Verwendung hier erlaubt. Die Bilder werden mit ihren Dateinamen auf Commons angeführt zusammen mit ihrem Autor bzw. Hochlader und der Lizenz. 7
- Lizenzklärung: Diese Seite wurde von Holger Brenner alias Bocardodarapti auf der deutschsprachigen Wikiversity erstellt und unter die Lizenz CC-by-sa 3.0 gestellt. 7