Prof. Dr. H. Brenner **Osnabrück SS 2014**

Analysis II

Vorlesung 54

Extrema unter Nebenbedingungen

Beispiel 54.1. Ein Nilpferd hat die ganze Nacht an Land gegrast und befindet sich gerade im Punkt $P = (r, s) \in \mathbb{R}^2$. Jetzt kommt plötzlich die heiße Sonne hervor und es muss möglichst schnell zurück in seinen Teich. Es sucht also den Punkt $a = (x, y) \in M$ des Teichufers M, der seiner momentanen Position am nächsten ist, d.h. es soll die Abstandsfunktion

$$h(x,y) = d(P,(x,y)) = \sqrt{(x-r)^2 + (y-s)^2}$$

minimiert werden, wobei allerdings nur die Punkte $(x, y) \in M$ relevant sind. Es geht also um ein Minimierungsproblem, wobei die Punkte die *Nebenbedingung* erfüllen müssen, zum Teichufer zu gehören. Das Teichufer werde mit Hilfe der Funktion

$$f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$$

als

$$M = \left\{(x,y) \in \mathbb{R}^2 \mid f(x,y) = b\right\}$$

zu einem gewissen $b \in \mathbb{R}$ beschrieben, d.h., es liegt als Faser einer Funktion vor. Wenn der Teich beispielsweise eine Ellipse ist, so ist $f(x,y) = \alpha x^2 + \beta y^2$. Wir nehmen weiter an, dass die Funktion f stetig differenzierbar ist und jeder Punkt der Faser regulär ist. Kann man die Punkte des Teichufers, in denen ein lokales Extremum vorliegt, mit Mitteln der Differentialrechnung charakterisieren? Nach dem Satz über implizite Funktionen gibt es lokal eine differenzierbare Parametrisierung des Teichufers, d.h. eine Funktion

$$\gamma\colon I \longrightarrow \mathbb{R}^2$$

auf einem offenen Intervall I, deren Bild gerade ein Ausschnitt aus dem Teichufer ist. Insgesamt erhält man die zusammengesetzte Funktion

$$h \circ \gamma \colon I \longrightarrow \mathbb{R},$$

und genau dann besitzt h in $a \in M$ ein lokales Extremum, wenn $h \circ \gamma$ ein lokales Extremum in $\gamma^{-1}(a) \in I$ besitzt. Auf $h \circ \gamma$ kann man die Kriterien für lokale Extrema (also Satz 48.1 bzw. Satz 19.1) anwenden, da jetzt der Definitionsbereich (man hat die Nebenbedingung sozusagen eliminiert) eine offene Teilmenge von $\mathbb{R}$ ist. Wenn ein lokales Extremum vorliegt, so ist einerseits

$$(h \circ \gamma)'(\gamma^{-1}(a)) = (Dh)_a(\gamma'(\gamma^{-1}(a))) = 0.$$

Andererseits bestimmt $\gamma'(\gamma^{-1}(a))$ den (eindimensionalen) Tangentialraum T_aM, und dieser ist wiederum der Kern des totalen Differentials $(Df)_a$. Daher müssen $(Dh)_a$ und $(Df)_a$ linear abhängig sein. Das Nilpferd muss also nach Punkten $a \in M$ Ausschau halten, für die es ein $\lambda \in \mathbb{R}$ mit

$$(Dh)_a = \lambda (Df)_a$$

gibt.

Satz 54.2. *Es sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und sei*

$$f \colon U \longrightarrow \mathbb{R}^m$$

eine stetig differenzierbare Abbildung. Es sei $M = f^{-1}(b)$ die Faser von f über $b \in \mathbb{R}^m$. Es sei

$$h \colon U \longrightarrow \mathbb{R}$$

eine differenzierbare Funktion und die eingeschränkte Funktion $h|_M$ besitze im Punkt $a \in M$ ein lokales Extremum auf M und a sei ein regulärer Punkt von f. Dann ist

$$T_aM \subseteq \operatorname{kern}(Dh)_a,$$

d.h. die Linearform $(Dh)_a$ verschwindet auf dem Tangentialraum an der Faser von f durch a. Die Linearform $(Dh)_a$ ist eine Linearkombination aus den Linearformen

$$(Df_1)_a, \ldots, (Df_m)_a \,.$$

Beweis. Wir wenden den Satz über implizite Abbildungen auf den Punkt $a \in M$ an. Es gibt also eine offene Menge $a \in W$, $W \subseteq U$, eine offene Menge $V \subseteq \mathbb{R}^{n-m}$ und eine stetig differenzierbare Abbildung

$$\psi \colon V \longrightarrow W$$

derart, dass $\psi(V) \subseteq M \cap W$ ist und ψ eine Bijektion

$$\psi \colon V \longrightarrow M \cap W$$

induziert. Dabei ist ψ in jedem Punkt $Q \in V$ regulär und für das totale Differential von ψ gilt

$$(Df)_{\psi(Q)} \circ (D\psi)_Q = 0.$$

Da $h|_M$ in a ein lokales Extremum besitzt, besitzt auch $h \circ \psi$ in $Q = \psi^{-1}(a)$ (also $a = \psi(Q)$) ein lokales Extremum. Nach Satz 48.1 ist daher

$$(D\,(h \circ \psi))_Q \;=\; (Dh)_{\psi(Q)} \circ (D\psi)_Q \;=\; 0.$$

Somit ist einerseits

$$\mathrm{bild}(D\psi)_Q \;\subseteq\; \mathrm{kern}(Dh)_a$$

und andererseits

$$\mathrm{bild}(D\psi)_Q \;=\; \mathrm{kern}(Df)_a \;=\; T_a M.$$

Der Zusatz folgt, da $\mathrm{kern}(Df)_a$ der Durchschnitt der $\mathrm{kern}(Df_i)_a$, $i = 1, \ldots, m$, ist und somit

$$\bigcap_{i=1}^{m} \mathrm{kern}(Df_i)_a \;\subseteq\; \mathrm{kern}(Dh)_a$$

gilt. Nach Aufgabe 54.1 folgt daraus, dass $(Dh)_a$ zu dem von $(Df_1)_a, \ldots, (Df_m)_a$ erzeugten Untervektorraum gehört. □

Man beachte, dass dieser Satz nur ein notwendiges Kriterium für lokale Extrema angibt, kein hinreichendes. Die auf dem Satz über implizite Abbildungen beruhende Existenz der Bijektion ψ wird zwar im Beweis verwendet, sie muss aber nicht explizit bekannt sein, um die Kandidaten für lokale Extrema zu bestimmen. Eine explizite Bijektion kann aber helfen zu entscheiden, ob in den Kanditaten ein lokales Extremum vorliegt oder nicht. Wenn es nur endliche viele Kandidaten gibt, so kann man die Funktionswerte ausrechnen und auf diesem Weg zumindest die globalen Extrema finden.

KOROLLAR 54.3. *Es sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und seien*

$$f\colon U \longrightarrow \mathbb{R}$$

und

$$h\colon U \longrightarrow \mathbb{R}$$

stetig differenzierbare Funktionen. Es sei $b \in \mathbb{R}$ und $M = f^{-1}(b)$ die Faser von f über b. Die eingeschränkte Funktion $h|_M$ besitze im Punkt $a \in M$ ein lokales Extremum auf M und a sei ein regulärer Punkt von f. Dann ist $(Dh)_a$ ein Vielfaches von $(Df)_a$, d.h. es gibt ein $\lambda \in \mathbb{R}$ mit

$$(Dh)_a \;=\; \lambda (Df)_a.$$

Beweis. Dies folgt unmittelbar aus Satz 54.2. □

Den Faktor λ nennt man *Lagrange-Multiplikator*. Diese Aussage legt folgendes Verfahren nahe, Kandidaten für lokale Extrema (unter Nebenbedingung) zu finden: Man untersucht einfach, für welche (bezüglich f) regulären Punkte $a \in M$ eine lineare Abhängigkeit zwischen $(Df)_a$ und $(Dh)_a$ vorliegt.

Beispiel 54.4. Wir betrachten die Funktion
$$h(x,y) \,=\, x^3 - y^2x$$
auf dem Einheitskreis $K = \{(x,y) \in \mathbb{R}^2 |\, x^2 + y^2 = 1\}$ und interessieren uns für die Punkte $a \in K$, auf denen h ein lokales Extremum annehmen kann. Das Differential von h ist
$$(3x^2 - y^2, -2xy)$$
und das Differential von
$$f(x,y) \,=\, x^2 + y^2$$
ist
$$(2x, 2y)\,.$$
Gemäß Korollar 54.3 müssen wir die Punkte $a \in K$ bestimmen, für die die beiden Differentiale linear abhängig sind. Die Determinante ist
$$\begin{aligned} \det \begin{pmatrix} 2x & 2y \\ 3x^2 - y^2 & -2xy \end{pmatrix} &= -4x^2y - 6x^2y + 2y^3 \\ &= -10x^2y + 2y^3 \\ &= 2y\left(-5x^2 + y^2\right). \end{aligned}$$
Somit liegt bei $y = 0$ und bei $y = \pm\sqrt{5}x$ lineare Abhängigkeit vor. Die Kreisbedingung führt somit auf die Punkte
$$a_1 = (1,\,0)\,,\, a_2 = (-1,\,0)\,,\, a_3 = \left(\sqrt{\frac{1}{6}},\,\sqrt{5}\sqrt{\frac{1}{6}}\right),\, a_4 = \left(\sqrt{\frac{1}{6}},\,-\sqrt{5}\sqrt{\frac{1}{6}}\right),$$
$$a_5 = \left(-\sqrt{\frac{1}{6}},\,\sqrt{5}\sqrt{\frac{1}{6}}\right),\, a_6 = \left(-\sqrt{\frac{1}{6}},\,-\sqrt{5}\sqrt{\frac{1}{6}}\right).$$

Korollar 54.5. *Es sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge, sei*
$$f\colon\, U \longrightarrow \mathbb{R}$$
eine stetig differenzierbare Funktion und sei $M = f^{-1}(b)$ die Faser von f über $b \in \mathbb{R}$. Es sei h eine Linearform auf $\mathbb{R}^n$, deren Einschränkung $h|_M$ auf M im regulären Punkt $a \in M$ ein lokales Extremum besitze. Dann ist h ein Vielfaches von $(Df)_a$, d.h. es gibt ein $\lambda \in \mathbb{R}$ mit
$$h \,=\, \lambda (Df)_a.$$

Beweis. Dies folgt wegen
$$(Dh)_a \,=\, h$$
unmittelbar aus Korollar 54.3. □

Beispiel 54.6. Wir betrachten die Linearform
$$h(x,y,z) \,=\, 3x - 2y + 5z$$
auf der Menge
$$M = \left\{(x,y,z) \in \mathbb{R}^3 |\, f(x,y,z) = x^2 + y^2 + z^4 = 1\right\}.$$

Die Lagrange-Bedingung wird zu

$$\begin{pmatrix} 3 \\ -2 \\ 5 \end{pmatrix} = \lambda \begin{pmatrix} 2x \\ 2y \\ 4z^3 \end{pmatrix} .$$

Dies führt auf

$$\lambda = \frac{3}{2x}$$

(und $x \neq 0$),

$$y = -\frac{1}{\lambda} = -\frac{2x}{3}$$

und

$$z = \sqrt[3]{\frac{5}{4\lambda}} = \sqrt[3]{\frac{5}{6}x}.$$

Dies führt insgesamt zur Bedingung

$$x^2 + \frac{4}{9}x^2 + \frac{5}{6}x\sqrt[3]{\frac{5}{6}x} = 1,$$

die nach dem Zwischenwertsatz mindestens zwei Lösungen hat, die allerdings nicht so einfach explizit anzugeben sind.

Korollar 54.7. *Es sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und sei*

$$f\colon U \longrightarrow \mathbb{R}$$

eine stetig differenzierbare Funktion. Die Faser $M = f^{-1}(b)$ von f zu einem Punkt $b \in \mathbb{R}$ sei kompakt und in jedem Punkt regulär. Dann ist jeder $(n-1)$-dimensionale Unterraum $V \subset \mathbb{R}^n$ für mindestens einen Punkt $a \in M$ gleich dem Tangentialraum T_aM.

Beweis. Der $(n-1)$-dimensionale Untervektorraum $V \subset \mathbb{R}^n$ wird durch eine Linearform beschrieben, sagen wir $V = \operatorname{kern} h$ mit

$$h(x_1, \ldots, x_n) = c_1x_1 + \cdots + c_nx_n,$$

wobei nicht alle c_i gleich 0 sind. Die Funktion h nimmt nach Satz 36.11 auf der kompakten Teilmenge M ihr Maximum an, d.h. es gibt einen Punkt $a \in M$ derart, dass $h|_M$ in a insbesondere ein lokales Extremum besitzt. Da a ein regulärer Punkt ist, folgt nach Korollar 54.5, dass

$$h = (Dh)_a = \lambda(Df)_a$$

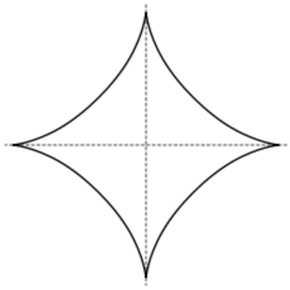
ist ($\lambda \neq 0$) und somit ist

$$V = \operatorname{kern} h = \operatorname{kern}(Df)_a = T_aM.$$

□

Ohne die Kompaktheitsvoraussetzung und ohne die Regularitätsvoraussetzung ist die vorstehende Aussage nicht richtig, wie einfache Beispiele zeigen.

Beispiel 54.8. Es sei
$$f(x,y) = xy$$
und
$$M = \{(x,y) \in \mathbb{R}^2 \mid f(x,y) = 1\}$$
die Standardhyperbel, realisiert als Faser einer Funktion. Jeder Punkt der Hyperbel ist ein regulärer Punkt von f, die Hyperbel ist nicht kompakt. Die beiden Linearformen x bzw. y besitzen kein lokales Extremum auf M und die beiden Koordinatenrichtungen treten nicht als Tangentialräume der Hyperbel auf.



Beispiel 54.9. Wir betrachten
$$f(x,y) = \left(x^2 + y^2 - 1\right)^3 + 27x^2y^2$$
und das zugehörige Nullstellengebilde, also
$$Z = \{(x,y) \in \mathbb{R}^2 \mid f(x,y) = 0\}.$$
Dieses nennt man eine *Astroide*[1]. Dieses Nullstellengebilde liegt innerhalb der abgeschlossenen Kreisscheibe und ist daher kompakt. Die partiellen Ableitungen sind
$$\frac{\partial f}{\partial x} = 3\left(x^2 + y^2 - 1\right)^2 \cdot 2x + 54xy^2 = 6x\left(\left(x^2 + y^2 - 1\right)^2 + 27y^2\right)$$
und
$$\frac{\partial f}{\partial y} = 3\left(x^2 + y^2 - 1\right)^2 \cdot 2y + 54x^2y = 6y\left(\left(x^2 + y^2 - 1\right)^2 + 27x^2\right).$$
Beide partiellen Ableitungen verschwinden genau für die vier Punkte
$$(0,1),\ (0,-1),\ (1,0),\ (-1,0),$$
die alle zu Z gehören. Die x-Achse $\mathbb{R}(1,0)$ tritt nicht als Tangente von Z auf. Die zweite partielle Ableitung verschwindet nämlich nur bei $y = 0$ oder $(x,y) = (0,\pm 1)$, in diesen Fällen verschwinden aber bereits beide partielle Ableitungen.

[1] Solche Gebilde werden im Rahmen der algebraischen Geometrie studiert, siehe den Kurs über algebraische Kurven auf Wikiversity.

Abbildungsverzeichnis

Quelle = AHippoBird2200.jpg , Autor = Benutzer PJ KAPDostie auf Commons, Lizenz = CC-bysa 3.0 1

Quelle = Astroid.svg , Autor = Benutzer Joelholdsworth auf Commons, Lizenz = PD 6