## Das Lemma von Morse

Die Bedeutung nichtentarteter kritischer Punkte besteht darin, dass Funktionen lokal bereits durch die *Anzahl* der negativen Eigenwerte der Hessischen vollständig charakterisiert sind. Man nennt [1]

$$\operatorname{ind}(c) \coloneqq \operatorname{card}(\operatorname{spec} Hf(c) \cap (-\infty, 0))$$

den *Index* des kritischen Punktes $c$. Alles andere spielt *lokal* keine Rolle, wie der folgende Satz zeigt. Ein Beweis findet man zum Beispiel in Lang, *Real Analysis*, Kapitel VII.

13 **Lemma von Morse** *Die $C^3$-Funktion $f: \mathbb{R}^n \hookrightarrow \mathbb{R}$ besitze einen nichtentarteten kritischen Punkt $c$. Dann existieren Koordinaten $u = (u_1, .., u_n)$ um $c$ so, dass*

$$f(u) = -u_1^2 - .. - u_k^2 + u_{k+1}^2 + .. + u_n^2 + f(c),$$

*wobei $k = \operatorname{ind}(c)$.* ⋊

Das heißt, es gibt eine Koordinatentransformation [2] $\varphi: U_0 \to U_c$ von einer offenen Umgebung $U_0$ von $0$ auf eine offene Umgebung $U_c$ von $c$, so dass

$$(f \circ \varphi)(u) = u_1^2 - .. - u_k^2 + u_{k+1}^2 + .. + u_n^2 + f(c)$$

in den neuen Koordinaten $u = (u_1, .., u_n)$.

In den passenden Koordinaten wird damit $f$ bis auf eine unwichtige additive Konstante zu einer quadratischen Form, die vollständig durch den Index $k$ des kritischen Punktes $c$ bestimmt ist. Da dieser Index nur $n + 1$ verschiedene Werte annehmen kann, erhalten wir folgendes

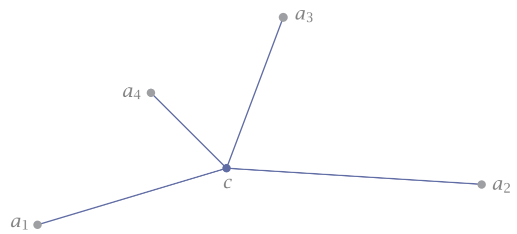
**Korollar** *Im $\mathbb{R}^n$ gibt es genau $n + 1$ verschiedene nichtentartete kritische Punkte, und zwar strikte Minimalstellen, strikte Maximalstellen, und Sattelpunkte mit Index $k = 1, .., n - 1$.* ⋊

[1] $\operatorname{spec} A$ bezeichnet das Spektrum von $A$.

[2] Genauer einen Diffeomorphismus – dazu später mehr.

Abb 6

Punkt mit minimaler Quadratabstandssumme




## Zwei Extremwertaufgaben

Als Anwendungsbeispiele betrachten wir zwei typische Extremwertaufgaben sowie das Maximumprinzip für harmonische Funktionen.

▷ *Extremwertaufgabe* Gegeben sind $m$ Punkte $a_1, .., a_m$ im $\mathbb{R}^n$. Gesucht ist ein Punkt $c$ im $\mathbb{R}^n$, so dass die Summe aller ihrer quadrierten Abstände zum Punkt $c$ minimal wird.

Gesucht ist demnach das Minimum der Funktion

$$f: \mathbb{R}^n \to \mathbb{R}, \quad f(x) = \frac{1}{2m} \sum_{i=1}^{m} \|x - a_i\|^2,$$

wobei die Division durch $m$ nur die folgenden Formeln vereinfacht und sonst keine Bedeutung hat. Das Quadrat der euklidischen Norm ist differenzierbar, mit

$$\nabla(\|x - a_i\|^2) = 2(x - a_i).$$

Also ist

$$\nabla f(x) = \frac{1}{m} \sum_{i=1}^{m} (x - a_i) = x - \frac{1}{m} \sum_{i=1}^{m} a_i.$$

Dieser Gradient besitzt einen einzigen kritischen Punkt in

$$c = \frac{1}{m} \sum_{i=1}^{m} a_i,$$

dem arithmetischen Mittel der Punkte $a_1, .., a_m$, beziehungsweise dem Schwerpunkt des Körpers mit gleichen Massen in den Punkten $a_1, .., a_m$. Aus der Geometrie des Problems ist klar, dass dies ein lokales und sogar globales *Minimum* ist. Die Hessische von $f$ ist übrigens $Hf = E > 0$. ◁

14 ▷ *Extremwertaufgabe mit Nebenbedingung* Zu bestimmen ist derjenige Quader, der bei vorgegebener Kantenlänge das größte Volumen einschließt.

Sind $x, y, z$ die Kantenlängen des Quaders, so ist also die Funktion

$$V = xyz$$

zu maximieren unter der Vorgabe, dass $x+y+z$ konstant ist. Aufgrund der Homogenität des Problems in allen drei Koordinaten können wir die Gesamtkantenlänge auf einen beliebigen positiven Wert fixieren, zum Beispiel auf $x+y+z=3$. Dann ist $z = 3-x-y$ und

$$V = V(x,y) = xy(3-x-y) = 3xy - x^2y - xy^2.$$

Für den Gradienten erhalten wir damit

$$\nabla V(x,y) = \begin{pmatrix} 3y-2xy-y^2 \\ 3x-2xy-x^2 \end{pmatrix}.$$

Dieser Gradient besitzt vier verschiedene Nullstellen, aber nur der kritische Punkt mit $x=y=1$ und damit auch $z=1$ hat positive Koordinaten. Auch hier ist aufgrund geometrischer Überlegungen klar, dass es sich um das globale Maximum der Volumenfunktion auf dem ersten Quadranten in $\mathbb{R}^3$ handeln muss. Das Maximum wird also von einem Quader mit drei gleichen Seiten erreicht, was auch nicht weiter überrascht.

Die Hessische ist übrigens

$$HV(x,y) = \begin{pmatrix} -2y & 3-2x-2y \\ 3-2x-2y & -2x \end{pmatrix},$$

und es gilt

$$HV(x,y)\Big|_{x=y=1} = \begin{pmatrix} -2 & -1 \\ -1 & -2 \end{pmatrix} < 0.$$

Ein allgemeines Verfahren für solche Extremwertaufgaben mit Nebenbedingungen werden wir übrigens später kennenlernen. ◂

15.19

## Das Maximumprinzip für harmonische Funktionen

**Definition** *Eine $C^2$-Funktion $u\colon \mathbb{R}^n \hookrightarrow \mathbb{R}$ heißt harmonisch, falls überall*

$$\Delta u \coloneqq \sum_{i=1}^{n} u_{x_i x_i} = 0.$$

*Bemerkung* Der Differenzialoperator

$$\Delta = \sum_{i=1}^{n} \partial_i^2$$

heißt *Laplaceoperator* und spielt in der Physik und vielen Anwendungen eine fundamentale Rolle. Zum Beispiel beschreiben harmonische Funktionen Gleichgewichtslösungen für viele wichtige partielle Differenzialgleichungen.

A. Auf einem Intervall ist eine Funktion $u$ harmonisch genau dann, wenn

$$u'' = 0,$$

also wenn sie linear ist.

B. Ist $p$ ein beliebiges Polynom mit komplexen Koeffizienten, so ist

$$u\colon\ \mathbb{R}^2 \to \mathbb{R}, \quad u(x,y) = \Re p(x + \mathrm{i}y)$$

harmonisch. Dasselbe gilt für den Imaginärteil $_{\text{A-6}}$.

C. Für $n \geqslant 2$ ist

$$u\colon\ \mathbb{R}^n \setminus \{0\} \to \mathbb{R}, \quad u(x) = \begin{cases} \log|x|, & n = 2 \\ |x|^{2-n}, & n \geqslant 3 \end{cases}$$

harmonisch.

Wir betrachten Funktionen in $C^0(\bar{\Omega}) \cap C^2(\Omega)$, mit $\Omega \subset \mathbb{R}^n$ offen und beschränkt. Diese sind also $C^2$ in $\Omega$ und noch stetig auf dem Abschluss von $\Omega$. Dies schließt aus, dass die Funktionen am Rand von $\Omega$ unbeschränkt werden.

15 **Maximumprinzip** *Sei $\Omega$ beschränkt und offen und $u \in C^0(\bar{\Omega}) \cap C^2(\Omega)$ in $\Omega$ harmonisch. Dann gilt:*

(i) *Die Funktion $u$ nimmt ihr Maximum auf dem Rand an:*

$$\max_{\bar{\Omega}} u = \max_{\partial\Omega} u.$$

(ii) *Ebenso nimmt $|u|$ sein Maximum auf dem Rand an.*

(iii) *Ist $u$ auf dem Rand von $\Omega$ konstant, so ist $u$ überall konstant.*

(i) Betrachte die modifizierte Funktion

$$w = u + \varepsilon v, \qquad \varepsilon > 0,$$

mit $v(x) = x_1^2 + .. + x_n^2 \geqslant 0$. Es ist $\Delta v = 2n > 0$ und damit

$$\Delta w = \Delta u + \varepsilon \Delta v = 2n\varepsilon > 0.$$

Diese Funktion besitzt in $\Omega$ keine Maximalstelle. Denn wäre $c \in \Omega$ eine solche Maximalstelle, so wäre $Hw(c) \leqslant 0$ [11], und somit wären alle Eigenwerte von $Hw(c)$ nichtpositiv. Dann ist aber auch

$$\Delta w(a) = \operatorname{sp} Hw(c) \leqslant 0,$$

im Widerspruch zu $\Delta w > 0$. Somit besitzt $w$ in $\Omega$ kein lokales Maximum.

Andererseits ist $w$ nach Voraussetzung auf $\bar{\Omega}$ stetig. Diese Menge ist abgeschlossen und beschränkt und damit kompakt. Also nimmt $w$ auf $\bar{\Omega}$ sein Maximum an, und es gilt

$$\max_{\bar{\Omega}} w = \max_{\partial\Omega} w = \max_{\partial\Omega} (u + \varepsilon v) \leqslant (\max_{\partial\Omega} u) + \varepsilon r^2$$

für ein hinreichend großes $r$. Also gilt auch

$$\max_{\bar{\Omega}} u \leqslant \max_{\bar{\Omega}} w \leqslant \max_{\partial\Omega} u + \varepsilon r^2.$$

Da dies für jedes $\varepsilon > 0$ gilt, folgt hieraus die erste Behauptung.

(ii) Mit $u$ ist auch $-u$ harmonisch und mit dem gerade Bewiesenen also

$$\max_{\bar{\Omega}} (-u) = \max_{\partial\Omega} (-u).$$

Dies ist äquivalent mit $\min_{\bar{\Omega}} u = \min_{\partial\Omega} u.$, Zusammen mit (i) folgt hieraus die zweite Behauptung.

(iii) Nach Voraussetzung ist $u|_{\partial\Omega} = m$ mit einer reellen Konstanten $m$. Dann ist aber auch $u - m$ harmonisch, also

$$\max_{\bar{\Omega}} |u - m| = \max_{\partial\Omega} |u - m| = 0.$$

Also gilt $u \equiv m$ auf $\bar{\Omega}$. ⟫⟫⟫

*Bemerkung* Man kann auch noch zeigen, dass eine nicht konstante harmonische Funktion $u \in C^0(\bar{\Omega}) \cap C^2(\Omega)$ *nur* auf dem Rand von $\Omega$ ihr Maximum und Minimum annimmt – siehe Hildebrandt 2, Seite 63. ⊸

## 15.4 Konvexe Mengen und Funktionen

**Definition** *Eine Menge $M \subset V$ heißt konvex, wenn sie mit je zwei Punkten auch ihre Verbindungsstrecke enthält:*

$$u, v \in M \quad \Rightarrow \quad [u, v] \subset M,$$

*wobei* $[u, v] \coloneqq \{(1-t)u + tv : 0 \leqslant t \leqslant 1\}$.

A. Auf der reellen Geraden sind die konvexen Mengen genau die Intervalle.

B. Jede offene Kugel $B_r(a)$ eines normierten Vektorraumes ist konvex. Denn seien $u, v \in B_r(a)$, also

$$\|u - a\| < r, \qquad \|v - a\| < r.$$

Für $0 \leqslant t \leqslant 1$ gilt dann aufgrund der Dreiecksungleichung und der positiven Homogenität einer Norm

$$\begin{aligned} \|((1-t)u + tv) - a\| &= \|(1-t)(u-a) + t(v-a)\| \\ &\leqslant (1-t)\,\|u-a\| + t\,\|v-a\| \\ &< (1-t)r + tr = r. \end{aligned}$$

Also ist auch $(1-t)u + tv \in B_r(a)$.

C. Dasselbe gilt für abgeschlossene Kugeln.

D. Der Durchschnitt beliebig vieler konvexer Mengen ist konvex.

E. Ist $M \subset V$ eine beliebige Teilmenge, so ist der Durchschnitt aller $M$ umfassenden konvexen Mengen,

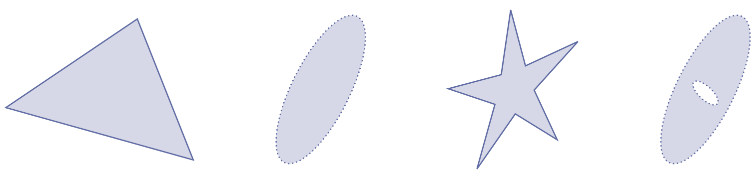
$$\check{M} \coloneqq \bigcap \{K \subset V : M \subset K \text{ und } K \text{ ist konvex}\},$$

eine konvex Menge. Diese heißt die *konvexe Hülle* von $M$ und ist die kleinste konvexe Menge in $V$, die $M$ enthält.

F. Ist $\Omega$ offen und $a \in \Omega$, so ist $\Omega \setminus \{a\}$ *nicht* konvex.

G. Eine Banane ist unkonvex.

Abb 7 Zwei konvexe und zwei nicht-konvexe Mengen



15.22

Man nennt $(1-t)u + tv$ mit $0 \leqslant t \leqslant 1$ eine Konvexkombination der Punkte $u$ und $v$. Allgemeiner heißt eine Linearkombination von $m \geqslant 1$ Punkten,

$$\lambda_1 u_1 + .. + \lambda_m u_m,$$

eine *Konvexkombination* dieser Punkte, wenn $\lambda_1, .., \lambda_m \geqslant 0$ und $\lambda_1 + .. + \lambda_m = 1$. Diese charakterisieren konvexe Mengen ebenfalls.

16 **Satz** *Eine Teilmenge $K$ eines Vektorraumes ist konvex genau dann, wenn jede Konvexkombination aus Punkten in $K$ ebenfalls zu $K$ gehört.* ⋊

⟨⟨⟨⟨⟨ $\Leftarrow$ Wenn jede Konvexkombination aus $K$ wieder zu $K$ gehört, dann gilt das natürlich auch für solche aus zwei Punkten. Also ist $K$ konvex.

$\Rightarrow$ Dies zeigen wir durch Induktion über die Zahl $m$ der konvex kombinierten Punkte. Für $m = 1$ ist nichts zu tun, ebensowenig für $m = 2$, da dies ja der Definition entspricht. Nun gelte die Behauptung für jede Konvexkombination aus $m \geqslant 2$ Punkten aus $K$, und es sei

$$u = \lambda_1 u_1 + .. + \lambda_{m+1} u_{m+1}$$

eine Konvexkombination aus $m+1$ Punkten in $K$. Ist $\lambda_{m+1} = 1$, so ist $u = u_{m+1}$, und wir sind bereits fertig. Andernfalls ist $\lambda \coloneqq \lambda_1 + .. + \lambda_m = 1 - \lambda_{m+1} > 0$. Aufgrund der Induktionsannahme gilt

$$v = \frac{\lambda_1}{\lambda} u_1 + .. + \frac{\lambda_m}{\lambda} u_m \in K,$$

denn dies ist eine Konvexkombination aus $m$ Punkten. Da $K$ auch jede Konvexkombination aus zwei Punkten enthält und $1 - \lambda = \lambda_{m+1}$, ist auch

$$\lambda v + (1-\lambda) u_{m+1} = \lambda_1 u_1 + .. + \lambda_{m+1} u_{m+1} = u \in K. \qquad ⟩⟩⟩⟩⟩$$

Nun zum Begriff der konvexen Funktion.

**Definition** *Eine auf einer konvexen Menge $K$ eines Vektorraumes definierte Funktion $f : K \to \mathbb{R}$ heißt konvex, wenn*

$$f((1-t)u + tv) \leqslant (1-t)f(u) + tf(v)$$

*für alle $u, v \in K$ und alle $0 \leqslant t \leqslant 1$. Sie heißt strikt konvex, wenn außerdem*

$$f((1-t)u + tv) < (1-t)f(u) + tf(v)$$

*für alle $u \neq v$ und alle $0 < t < 1$.* ⋊

Geometrisch bedeutet dies, dass der Graph von $f$ unterhalb jeder Verbindungsstrecke zweier Punkte auf diesem Graphen liegt. Dies lässt sich ebenfalls mit dem Begriff der konvexen Menge ausdrücken.

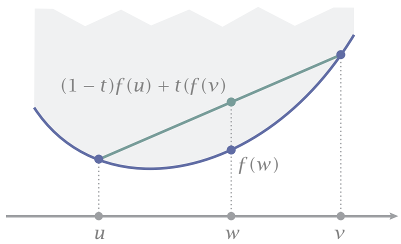



Abb 8

Eine konvexe Funktion und ihr Epigraph

**Notiz** *Eine auf einer konvexen Menge $K$ eines Vektorraumes $V$ definierte Funktion $f: K \to \mathbb{R}$ ist konvex genau dann, wenn ihr Epigraph*

$$\operatorname{Epi}(f) \coloneqq \{(u,z) \in V \times \mathbb{R} : u \in K,\ z \geqslant f(u)\}$$

*konvex in $V \times \mathbb{R}$ ist.* ⋊

▹ Jede Norm $N$ auf einem Vektorraum ist konvex, denn aufgrund der Dreiecksungleichung und der positiven Homogenität gilt, für $0 \leqslant t \leqslant 1$,

$$\begin{aligned} N((1-t)u + tv) &\leqslant N((1-t)u) + N(tv) \\ &= (1-t)N(u) + tN(v). \end{aligned}$$

Sie ist aber nicht strikt konvex, denn für $u = 0$ und $v \neq 0$ gilt

$$N((1-t)u + tv) = N(tv) = tN(v) = (1-t)N(u) + tN(v). \quad ◃$$

Ist eine Funktion konvex, so ist sie es auch bezüglich beliebiger Konvexkombinationen ihrer Argumente. Dies wird genau so bewiesen wie die entsprechende Aussage über konvexe Mengen 16.

17 **Satz** *Eine auf einer konvexen Menge $K$ eines Vektorraumes definierte Funktion $f: K \to \mathbb{R}$ ist konvex genau, wenn*

$$f(\lambda_1 v_1 + .. + \lambda_m v_m) \leqslant \lambda_1 f(v_1) + .. + \lambda_m f(v_m)$$

*für jede Konvexkombination $\lambda_1 v_1 + .. + \lambda_m v_m$ in $K$.* ⋊

Die Konvexität einer Funktion ist punktweise definiert und erfordert keinerlei Stetigkeit. Auf einer abgeschlossenen Menge muss dies auch nicht der Fall sein, wie Abbildung 10 zeigt. Jene Funktion ist offensichtlich konvex auf $[-1, 1]$, aber unstetig. Anders ist dies auf *offenen* Mengen.

**Satz** *Sei $\Omega \subset V$ offen und konvex. Ist $f: \Omega \to \mathbb{R}$ konvex, so ist $f$ stetig und auf jeder kompakten Teilmenge von $\Omega$ sogar lipschitzstetig.* ⋊

Wir benötigen diesen Satz hier nicht und übergehen deshalb den Beweis. Man findet ihn zum Beispiel in Hildebrandt 2, Seite 75–76.