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

Analysis II

Vorlesung 47

Zu einer reellwertigen Funktion

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

interessieren wir uns wie schon bei einem eindimensionalen Definitionsbereich für die Extrema, also Maxima und Minima, der Funktion, und inwiefern man dies anhand der Ableitungen (falls diese existieren) erkennen kann. Wenn eine solche Funktion total differenzierbar ist, so ist das totale Differential in einem Punkt eine lineare Abbildung von V nach $\mathbb{K}$. Für solche linearen Abbildungen gibt es einen eigenen Namen.

Definition 47.1. Sei K ein Körper und sei V ein K-Vektorraum. Eine lineare Abbildung

$$V \longrightarrow K$$

heißt auch eine *Linearform* auf V.

Definition 47.2. Sei K ein Körper und V ein K-Vektorraum. Dann heißt der Homomorphismenraum

$$V^* = \operatorname{Hom}_K (V, K)$$

der *Dualraum* zu V.

Wenn $G \subseteq \mathbb{K}^n$ ist, so bilden die partiellen Ableitungen in einem Punkt $P \in G$ eine Matrix mit einer einzigen Zeile, nämlich

$$\left(\frac{\partial f}{\partial x_1}(P), \ldots, \frac{\partial f}{\partial x_1}(P) \right) ,$$

die bei stetigen partiellen Ableitungen das totale Differential repräsentiert. Eine solche Matrix kann man aber ebenso auch als ein n-Tupel in $\mathbb{K}$ und damit als einen Vektor in $\mathbb{K}^n$ auffassen. Dieser Zusammenhang zwischen Vektoren und Linearformen beruht auf dem Standardskalarprodukt des $\mathbb{K}^n$, und lässt sich konzeptioneller mit Hilfe von Bilinearformen erfassen.

Definition 47.3. Sei K ein Körper und sei V ein K-Vektorraum. Eine Abbildung

$$V \times V \longrightarrow K,\, (v,w) \longmapsto \langle v,w \rangle ,$$

heißt *Bilinearform*, wenn für alle $v \in V$ die induzierten Abbildungen

$$V \longrightarrow K,\, w \longmapsto \langle v,w \rangle ,$$

und für alle $w \in V$ die induzierten Abbildungen

$$V \longrightarrow K,\, v \longmapsto \langle v,w \rangle ,$$

K-linear sind.

Eine wichtige Eigenschaft von Bilinearformen, die Skalarprodukte erfüllen, wird in der nächsten Definition formuliert.

Definition 47.4. Sei K ein Körper und sei V ein K-Vektorraum. Eine Bilinearform

$$V \times V \longrightarrow K,\ (v,w) \longmapsto \langle v,w\rangle,$$

heißt *nicht ausgeartet*, wenn für alle $v \in V$, $v \neq 0$, die induzierten Abbildungen

$$V \longrightarrow K,\ w \longmapsto \langle v,w\rangle,$$

und für alle $w \in V$, $w \neq 0$, die induzierten Abbildungen

$$V \longrightarrow K,\ v \longmapsto \langle v,w\rangle,$$

nicht die Nullabbildung sind.

In dieser Vorlesung werden wir für Vektorräume, auf denen eine nicht-ausgeartete Bilinearform gegeben ist, eine bijektive Beziehung zwischen Vektoren und Linearformen beweisen und damit einen Zusammenhang zwischen dem totalen Differential zu einer Funktion in einem Punkt und einem Vektor, dem sogenannten Gradienten der Funktion in diesem Punkt, herstellen.

Der Gradient

Lemma 47.5. *Es sei K ein Körper und V ein K-Vektorraum, der mit einer Bilinearform $\langle -,-\rangle$ versehen sei. Dann gelten folgende Aussagen*

(1) *Für jeden Vektor $u \in V$ sind die Zuordnungen*

$$V \longrightarrow K,\ v \longmapsto \langle u,v\rangle,$$

und

$$V \longrightarrow K,\ v \longmapsto \langle v,u\rangle,$$

K-linear.

(2) *Die Zuordnung*

$$V \longrightarrow V^*,\ u \longmapsto \langle u,-\rangle,$$

ist K-linear.

(3) *Wenn $\langle -,-\rangle$ nicht ausgeartet ist, so ist die Zuordnung in (2) injektiv. Ist V zusätzlich endlichdimensional, so ist diese Zuordnung bijektiv.*

Beweis. (1) folgt unmittelbar aus der Bilinearität. (2). Seien $u_1, u_2 \in V$ und $a_1, a_2 \in K$. Dann ist für jeden Vektor $v \in V$

$$\langle a_1u_1 + a_2u_2, v\rangle \ = \ a_1\langle u_1,v\rangle + a_2\langle u_2,v\rangle,$$

und dies bedeutet gerade die Linearität der Zuordnung. Da die Zuordnung nach (2) linear ist, müssen wir zeigen, dass der Kern davon trivial ist. Sei also $u \in V$ so, dass $\langle u,-\rangle$ die Nullabbildung ist. D.h. $\langle u,v\rangle = 0$ für alle

$v \in V$. Dann muss aber nach der Definition von nicht ausgeartet $u = 0$ sein. Wenn V endliche Dimension hat, so liegt eine injektive lineare Abbildung zwischen Vektorräumen der gleichen Dimension vor, und eine solche ist nach Satz Anhang A.1 bijektiv. □

Wenn es also in einem endlichdimensionalen Vektorraum eine nicht ausgeartete Bilinearform gibt, beispielsweise ein Skalarprodukt, so gibt es zu jeder Linearform einen eindeutig bestimmten Vektor, mit dem diese Linearform beschrieben wird. Wendet man dies auf die Linearform an, die durch das totale Differential zu einer differenzierbaren Funktion $f\colon V \to \mathbb{R}$ gegeben ist, so gelangt man zum Begriff des Gradienten.

Definition 47.6. Sei $(V, \langle -,- \rangle)$ ein euklidischer Vektorraum, $G \subseteq V$ offen und

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

eine in $P \in G$ differenzierbare Funktion. Dann nennt man den eindeutig bestimmten Vektor $w \in V$ mit

$$(Df)_P\,(v) \,=\, \langle w, v \rangle$$

für alle $v \in V$ den *Gradienten* von f in P. Er wird mit

$$\operatorname{Grad} f(P)$$

bezeichnet.

Man beachte, dass wir durchgehend die endlichdimensionalen Vektorräume mit einem Skalarprodukt versehen, um topologische Grundbegriffe wie Konvergenz und Stetigkeit zur Verfügung zu haben, dass diese Begriffe aber nicht von dem gewählten Skalarprodukt abhängen. Dem entgegen hängt aber der Gradient von dem gewählten Skalarprodukt ab.

Bei $V = \mathbb{R}^n$, versehen mit dem Standardskalarprodukt, ist der Gradient einfach gleich

$$\operatorname{Grad} f(P) \,=\, \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}.$$

Bemerkung 47.7. Zu einer differenzierbaren Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ lässt sich der Gradient (bezüglich des Standardskalarproduktes) einfach durch partielles Differenzieren berechnen. Es wäre aber eine künstliche Einschränkung, nur diese Situation zu betrachten. Um dies zu illustrieren sei beispielsweise

$$f\colon \mathbb{R}^3 \longrightarrow \mathbb{R}$$

eine differenzierbare Funktion und $E \subset \mathbb{R}^3$ eine Ebene, die etwa als Lösungsmenge der linearen Gleichung $5x - 4y + 9z = 0$ gegeben sei. Dann induziert das Standardskalarprodukt des $\mathbb{R}^3$ durch Einschränkung ein Skalarprodukt auf E. Diese Ebene ist zwar isomorph zu $\mathbb{R}^2$, es ergibt aber keinen Sinn, das eingeschränkte Skalarprodukt als Standardskalarprodukt anzusprechen. Der

Gradient G zu f in einem Punkt $P \in \mathbb{R}^3$ lässt sich direkt mit den partiellen Ableitungen zu den drei Raumkoordinaten berechnen. Bei $P \in E$ wird im Allgemeinen der Gradient *nicht* auf E liegen. Die eingeschränkte Funktion

$$f|_E \colon E \longrightarrow \mathbb{R}$$

ist aber ebenfalls differenzierbar und besitzt daher einen Gradienten $\tilde{G}$, der auf E liegt, und dieser lässt sich nicht über partielle Ableitungen berechnen, da es auf E keine Standardbasis gibt. Übrigens ist $\tilde{G}$ die orthogonale Projektion von G auf E.

SATZ 47.8. *Sei $(V, \langle -, - \rangle)$ ein euklidischer Vektorraum, sei $G \subseteq V$ offen und sei*

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

eine in $P \in G$ differenzierbare Funktion. Dann gelten folgende Aussagen.

(1) *Für jeden Vektor $v \in V$ ist*

$$|(Df)_P(v)| \leq ||v|| \cdot ||\text{Grad } f(P)|| .$$

(2) *Dabei gilt Gleichheit genau dann, wenn v linear abhängig zum Gradienten ist.*

(3) *Sei* Grad $f(P) \neq 0$. *Unter allen Vektoren $v \in V$ mit $|| v ||= 1$ ist die Richtungsableitung in Richtung des normierten Gradienten maximal, und zwar gleich der Norm des Gradienten.*

Beweis. (1) folgt wegen

$$(Df)_P(v) = \langle v, \text{Grad } f(P) \rangle$$

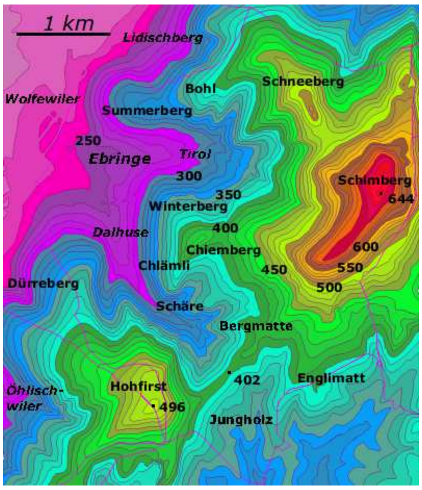
direkt aus der Abschätzung von Cauchy-Schwarz. (2) ergibt sich aus den Zusätzen zur Cauchy-Schwarz-Abschätzung, siehe Aufgabe 32.9. (3). Aus (1) und (2) folgt, dass

$$\begin{aligned} \left| \left\langle \text{Grad } f(P), \pm \frac{\text{Grad } f(P)}{||\text{Grad } f(P)||} \right\rangle \right| &= \left| (Df)_P \left(\pm \frac{\text{Grad } f(P)}{||\text{Grad } f(P)||} \right) \right| \\ &= ||\text{Grad } f(P)|| \end{aligned}$$

gilt, und dass diese beiden Vektoren die einzigen Vektoren der Norm 1 sind, für die diese Gleichung gilt. Wenn man links die Betragstriche weglässt, so gilt die Gleichheit für $\frac{\text{Grad } f(P)}{||\text{Grad } f(P)||}$ nach wie vor, da das Skalarprodukt positiv definit ist. □

Der Gradient gibt demnach die Richtung an, in die die Funktion den stärksten Anstieg hat. In die entgegengesetze Richtung liegt entsprechend der steilste Abstieg vor.

Gradient und Niveaumengen



In einer topographischen Karte wird ein Gebirge durch seine Niveaulinien (Höhenlinien) repräsentiert.

Definition 47.9. Zu einer Funktion

$$f\colon G \longrightarrow \mathbb{K},$$

wobei G ein metrischer Raum sei, nennt man zu $c \in \mathbb{K}$ die Menge

$$N_c = \{x \in G |\, f(x) = c\}$$

die *Niveaumenge* zu f zum Wert c.

Wir werden Niveaumengen (ein anderes Wort ist Faser) später systematischer untersuchen. Die folgende Aussage bedeutet, dass der Gradient stets senkrecht auf den Niveaumengen steht.

Lemma 47.10. *Sei $(V, \langle -, - \rangle)$ ein euklidischer Vektorraum, $G \subseteq V$ offen und*

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

eine in $P \in G$ differenzierbare Funktion. Es sei

$$h\colon I \longrightarrow G$$

eine differenzierbare Kurve mit $h(0) = P$, die ganz innerhalb einer Niveaumenge von f verläuft. Dann steht der Gradient zu f senkrecht auf $h'(0)$.

Beweis. Es sei N die Niveaumenge zu $c \in \mathbb{R}$, in der die Kurve h verlaufe. Dann ist die Hintereinanderschaltung $f \circ h$ konstant gleich c und daher ist unter Verwendung der Kettenregel

$$0 = (D(f \circ h))_0 = (Df)_P \circ (Dh)_0.$$

Daher liegt

$$(Dh)_0(1) = h'(0)$$

im Kern von $(Df)_P$, und das bedeutet, dass $h'(0)$ senkrecht auf dem Gradienten steht. □

Lokale Extrema von Funktionen in mehreren Variablen

Wir wollen mit den Mitteln der Differentialrechnung Kriterien erarbeiten, in welchen Punkten eine Funktion

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

ein lokales Minimum oder ein lokales Maximum annimmt. Wenn man sich den Graph einer solchen Funktion als ein Gebirge über der Grundmenge G vorstellt, so geht es also um die Gipfel und die Senken des Gebirges. Der folgende Satz liefert ein notwendiges Kriterium für die Existenz eines lokalen Extremums, das das entsprechende Kriterium in einer Variablen verallgemeinert.

Satz 47.11. *Sei V ein endlichdimensionaler reeller Vektorraum und $G \subseteq V$ eine offene Teilmenge. Es sei*

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

eine Funktion, die im Punkt $P \in G$ ein lokales Extremum besitzt. Dann gelten folgende Aussagen.

(1) *Wenn f in P in Richtung $v \in V$ differenzierbar ist, so ist*

$$(D_v f)(P) = 0.$$

(2) *Wenn f in P total differenzierbar ist, so verschwindet das totale Differential, also*

$$(Df)_P = 0.$$

Beweis. (1) Zu $v \in V$ betrachten wir die Funktion

$$h\colon I \longrightarrow \mathbb{R},\, t \longmapsto h(t) = f(P + tv),$$

wobei I ein geeignetes reelles Intervall ist. Da die Funktion f in P ein lokales Extremum besitzt, besitzt die Funktion h in $t = 0$ ein lokales Extremum. Nach Voraussetzung ist h differenzierbar und nach Satz 19.1 ist $h'(0) = 0$. Diese Ableitung stimmt aber mit der Richtungsableitung überein, also ist

$$(D_v f)(P) = h'(0) = 0.$$

(2) folgt aus (1) aufgrund von Proposition 46.1. □

Ein lokales Extremum kann also nur in einem sogenannten kritischen Punkt einer Funktion auftreten.

Definition 47.12. Sei V ein endlichdimensionaler reeller Vektorraum, $G \subseteq V$ offen und

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

eine differenzierbare Funktion. Dann heißt $P \in G$ ein *kritischer Punkt* von f (oder ein *stationärer Punkt*), wenn

$$(Df)_P = 0$$

ist. Andernfalls spricht man von einem *regulären Punkt*.

Abbildungsverzeichnis