Prof. Dr. H. Brenner **Osnabrück SS 2010**

Mathematik II

Vorlesung 44

Der Gradient

Lemma 46.1. *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 $v \in V$
$$\langle a_1 u_1 + a_2 u_2, v \rangle \;=\; a_1 \langle u_1, v \rangle + a_2 \langle u_2, v \rangle,$$
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 Korollar 12.10 bijektiv. □

Wenn es also in einem endlichdimensionalen Vektorraum eine nicht ausgeartete Bilinearform gibt, bspw. 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 $f : V \to \mathbb{R}$ gegeben ist, so gelangt man zum Begriff des Gradienten.

DEFINITION 46.2. Sei $(V, \langle -, - \rangle)$ ein euklidischer Vektorraum, $G \subseteq V$ offen und

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

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

$$(D\varphi)_P(v) \;=\; \langle u, v \rangle$$

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

$$\text{grad } f(P)$$

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

$$\text{grad } f(P) \;=\; \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}.$$

SATZ 46.3. *Sei $(V, \langle -, - \rangle)$ ein euklidischer Vektorraum, sei $G \subseteq V$ offen und sei*

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

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* $\text{grad } f(P) \neq 0$. *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, siehe Aufgabe 46.14. (3). Aus (1) und (2) folgt, dass

$$|\,\langle \text{grad } f(P), \pm\frac{\text{grad } f(P)}{||\,\text{grad } f(P)\,||} \rangle\,| \;\;=\;\; |\,(Df)_P(\pm\frac{\text{grad } f(P)}{||\,\text{grad } f(P)\,||})\,|$$

$$= \; ||\operatorname{grad} f(P)||$$

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 $\frac{\operatorname{grad} f(P)}{||\operatorname{grad} f(P)||}$ 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.

Lokale Extrema von Funktionen in mehreren Variablen

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

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

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

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

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

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

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 28.1 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 45.1. □

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

Definition 46.5. Sei V ein endlichdimensionaler reeller Vektorraum, $G \subseteq V$ offen und

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

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

$$(Df)_P = 0$$

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

Wir sind natürlich auch an hinreichenden Kriterien für das Vorliegen von lokalen Extrema interessiert. Wie schon im eindimensionalen Fall muss man sich die zweiten Ableitungen anschauen, wobei die Situation natürlich dadurch wesentlich verkompliziert wird, dass es zu je zwei Richtungsvektoren v und w eine zweite Richtungsableitung $D_{vw} = D_v D_w$ gibt. Die zweite Richtungsableitung wird dadurch handhabbar, dass man sie in die sogenannte Hesse-Form bzw. Hesse-Matrix zusammenfasst. Als solche ist sie eine symmetrische Bilinearform, die mit Methoden der linearen Algebra analysiert werden kann. Diese Methoden werden wir im Folgenden entwickeln und insbesondere auf die Hesse-Form anwenden, um schließlich hinreichende Kriterien für die Existenz von lokalen Extrema zu erhalten.

Definition 46.6. Es sei V ein endlichdimensionaler reeller Vektorraum, $G \subseteq V$ eine offene Menge und

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

eine zweimal stetig differenzierbare Funktion. Zu $P \in G$ heißt die Abbildung

$$\operatorname{Hess}_P f : V \times V \longrightarrow \mathbb{R},\ (u, v) \longmapsto D_u D_v f(P),$$

die *Hesse-Form* im Punkt $P \in G$.

Definition 46.7. Es sei V ein endlichdimensionaler reeller Vektorraum, $G \subseteq V$ eine offene Menge und

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

eine zweimal stetig differenzierbare Funktion. Es sei eine Basis v_i, $i = 1, \ldots, n$, von V gegeben mit den zugehörigen Richtungsableitungen $D_i = D_{v_i}$, $i = 1, \ldots, n$. Zu $P \in G$ heißt dann die Matrix

$$\begin{pmatrix} D_1 D_1 f(P) & \cdots & D_1 D_n f(P) \\ \vdots & \ddots & \vdots \\ D_n D_1 f(P) & \cdots & D_n D_n f(P) \end{pmatrix}$$

die *Hesse-Matrix* zu f im Punkt P bzgl. der gegebenen Basis.

Die Hesse-Matrix ist beispielsweise die Gramsche Matrix der Hesse-Form bzgl. der Standardbasis im $\mathbb{R}^n$.

Eigenschaften von Bilinearformen

Definition 46.8. Es sei K ein Körper, V ein endlichdimensionaler K-Vektorraum und $\langle -,-\rangle$ eine Bilinearform. Es sei $v_1, \ldots, v_n$ eine Basis von V. Dann heißt die $n \times n$-Matrix

$$\langle v_i, v_j \rangle_{1 \leq i,j \leq n}$$

die *Gramsche Matrix* von $\langle -,-\rangle$ bzgl. der Basis.

Lemma 46.9. *Es sei K ein Körper, V ein endlichdimensionaler K-Vektorraum und $\langle -,-\rangle$ eine Bilinearform. Es seien $\mathfrak{v} = v_1, \ldots, v_n$ und $\mathfrak{w} = w_1, \ldots, w_n$ zwei Basen von V und es seien G bzw. H die Gramschen Matrizen von $\langle -,-\rangle$ bzgl. diesen Basen. Zwischen den Basiselementen gelten die Beziehungen*

$$w_i = \sum_{j=1}^{n} a_{ij} v_j ,$$

die wir durch die Übergangsmatrix $A = (a_{ij})_{i,j}$ ausdrücken. Dann besteht zwischen den Gramschen Matrizen die Beziehung

$$H = AGA^t.$$

Beweis. Es ist

$$\begin{aligned} \langle w_r, w_s \rangle &= \langle \sum_{j=1}^{n} a_{rj} v_j, \sum_{k=1}^{n} a_{sk} v_k \rangle \\ &= \sum_{1 \leq j,k \leq n} a_{rj} a_{sk} \langle v_j, v_k \rangle \\ &= \sum_{1 \leq j \leq n} a_{rj} \big(\sum_{1 \leq k \leq n} a_{sk} \langle v_j, v_k \rangle \big) \\ &= (A \circ G \circ A^t)_{rs}. \end{aligned}$$

□

Definition 46.10. Es sei K ein Körper, V ein K-Vektorraum und $\langle -,-\rangle$ eine Bilinearform. Die Bilinearform heißt *symmetrisch*, wenn

$$\langle v, w \rangle = \langle w, v \rangle$$

für alle $v, w \in V$ gilt.

Definition 46.11. Es sei V ein reeller Vektorraum mit einer symmetrischen Bilinearform $\langle -,-\rangle$. Diese Bilinearfrom heißt

(1) *positiv definit*, wenn $\langle v, v \rangle > 0$ ist für alle $v \in V$, $v \neq 0$.
(2) *negativ definit*, wenn $\langle v, v \rangle < 0$ ist für alle $v \in V$, $v \neq 0$.
(3) *positiv semidefinit*, wenn $\langle v, v \rangle \geq 0$ ist für alle $v \in V$.
(4) *negativ semidefinit*, wenn $\langle v, v \rangle \leq 0$ ist für alle $v \in V$.
(5) *indefinit*, wenn $\langle -,-\rangle$ weder positiv semidefinit noch negativ semidefinit ist.

Positiv definite symmetrische Bilinearformen nennt man auch Skalarprodukte. Eine Bilinearform auf V kann man auf einen Untervektorraum $U \subseteq V$ einschränken, wodurch sich eine Bilinearform auf U ergibt. Wenn die ursprüngliche Form positiv definit ist, so überträgt sich dies auf die Einschränkung. Allerdings kann eine indefinite Form eingeschränkt auf gewisse Unterräume positiv definit werden und auf andere negativ definit. Dies führt zu folgender Definition.

Definition 46.12. Es sei V ein endlichdimensionaler reeller Vektorraum mit einer symmetrischen Bilinearform $\langle -,- \rangle$. Man sagt, dass eine solche Bilinearform den *Typ*

$$(p,q)$$

besitzt, wobei

$$p = \max\left(\dim_{\mathbb{R}}(U), U \subseteq V,\ \langle -,- \rangle|_U \text{ positiv definit}\right)$$

und

$$q = \max\left(\dim_{\mathbb{R}}(W), W \subseteq V,\ \langle -,- \rangle|_W \text{ negativ definit}\right)$$

ist.

Wie für Skalarprodukte nennt man zwei Vektoren $v, w \in V$ *orthogonal* bzgl. einer Bilinearform, wenn $\langle v, w \rangle = 0$ ist, und wie im Fall eines Skalarproduktes kann man zeigen, dass es Orthogonalbasen gibt. Die folgende Aussage nennt man den *Trägheitssatz von Sylvester*.

Satz 46.13. *Es sei V ein endlichdimensionaler reeller Vektorraum mit einer symmetrischen Bilinearform $\langle -,- \rangle$ vom Typ (p,q). Dann ist die Gramsche Matrix von $\langle -,- \rangle$ bzgl. einer jeden Orthogonalbasis eine Diagonalmatrix mit p positiven und q negativen Einträgen.*

Beweis. Bzgl. einer Orthogonalbasis $u_1, \ldots, u_n$ hat die Gramsche Matrix natürlich Diagonalgestalt. Es sei p' die Anzahl der positiven Diagonaleinträge und q' die Anzahl der negativen Diagonaleinträge. Die Basis sei so geordnet, dass die ersten p' Diagonaleinträge positiv, die folgenden q' Diagonaleinträge negativ und die übrigen 0 seien. Auf dem p'-dimensionalen Unterraum $U = \langle u_1, \ldots, u_{p'} \rangle$ ist die eingeschränkte Bilinearform positiv definit, so dass $p' \leq p$ gilt. Sei $W = \langle u_{p'+1}, \ldots, u_n \rangle$, auf diesem Unterraum ist die Bilinearform negativ semidefinit. Dabei ist $V = U \oplus W$, und diese beiden Räume sind orthogonal zueinander.

Angenommen, es gebe einen Unterraum U', auf dem die Bilinearform positiv definit ist, und dessen Dimension größer als p' ist. Dann ist $U' \not\subseteq U$. Wir können annehmen, dass $U \subseteq U'$ ist, da man eine Orthogonalbasis von U zu einer Orthogonalbasis von $U + U'$ ergänzen kann, und dabei die positive Definitheit erhalten bleibt. Es gibt dann ein $u \in U'$, $u \notin U$, das orthogonal zu U ist. Daher ist $u \in W$. Dann ist aber zugleich $\langle u, u \rangle$ positiv und nicht positiv, ein Widerspruch. $\square$