Reellwertige Funktionen mehrerer Variabler

Teilnehmende:

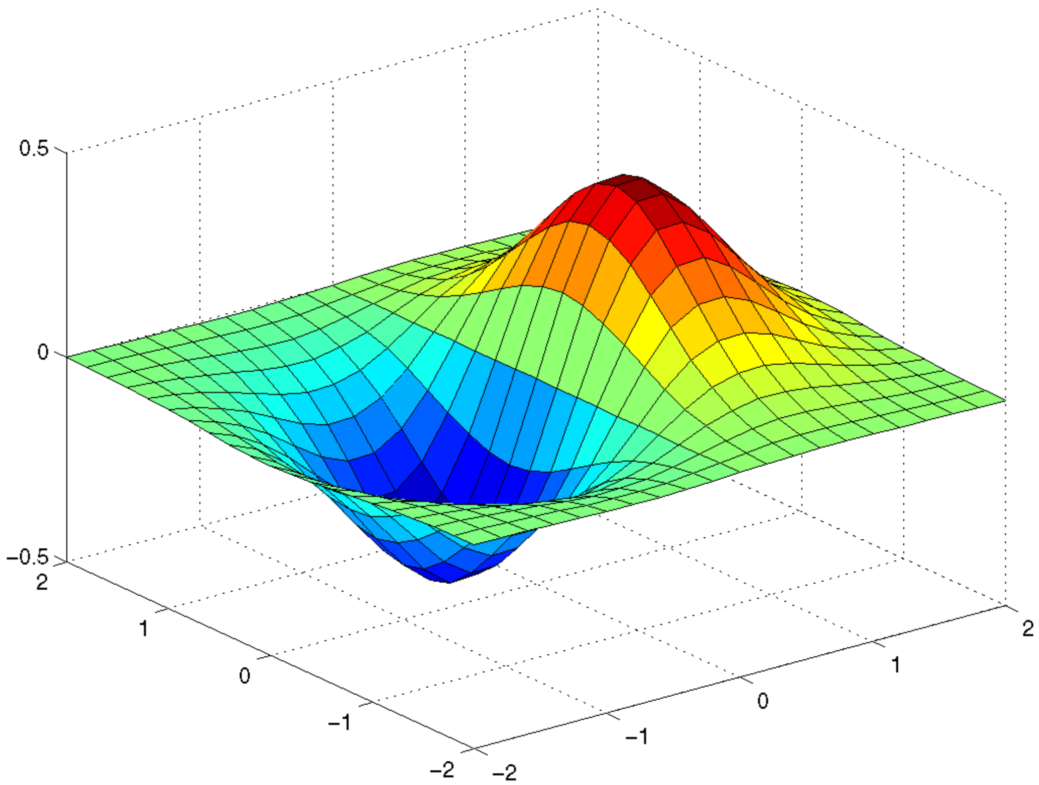
Name	Schule
2 Teilnehmende des Andreas-Gymnasium	
2 Teilnehmdende des Heinrich-Hertz-Oberschule	
2 Teilnehmende des Herder-Gymnasium	
1 Teilnehmer des Immanuel-Kant-Gymnasium	
2 Teilnehmende des Käthe-Kollwitz-Gymnasium	

mit tatkräftiger Unterstützung durch:

Karl Krum	Humboldt-Universität zu Berlin

Gruppenleiterin:

Barbara Grabowski	HTW des Saarlandes Saarbrücken



Unsere Arbeitsgruppe hat sich mit reellwertigen Funktionen in mehreren Veränderlichen beschäftigt. Diese sind von großer Bedeutung, da sich viele Probleme aus Wirtschaft, Technik und Wissenschaft durch Funktionen in einer Variablen nicht lösen lassen. Untersucht werden die Fragestellungen, wo lokale und globale Extrema auftreten, wie sehr geringfügige Änderungen in den Eingangswerten der Funktion eine Abweichung des Funktionswertes verursachen und welche Eingangswerte gleiche Funktionswerte hervorrufen.

1. Grundlagen

1.1. Vektorrechung

Da die Vektorrechnung grundlegend für die Betrachtung reellwertiger Funktionen mehrerer Veränderlicher ist, haben wir uns zunächst mit den Grundlagen der Vektorrechnung befasst.

Definition 1. (Verbindungsvektor)

Seien $P = (p_1, \dots, p_n)$ und $Q = (q_1, \dots, q_n)$ zwei Punkte aus $\mathbb{R}^n$. Das Gebilde $\vec{v}$ für welches gilt, dass $\vec{v}$ die gerichtete Strecke zwischen P und Q ist, heißt Vektor von P nach Q.

Schreibweise: $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \overrightarrow{PQ}$, mit $v_i = q_i - p_i$ für $i = 1, \dots, n$.

Wir bezeichnen mit $P + \vec{v}$ das Abtragen des Vektors $\vec{v}$ am Punkt P. Das Ergebnis ist der Punkt Q.

Definition 2. (Rechenoperationen mit Vektoren)

1. Die Addition und Subtraktion zweier Vektoren, sowie die Multiplikation eines Vektors mit einer reellen Zahl (Skalar genannt) erfolgt komponentenweise.
2. $|\vec{v}| = \sqrt{v_1^2 + v_2^2 + ... + v_n^2}$ heißt Betrag des Vektors $\vec{v} \in \mathbb{R}^n$ und ist gleich seiner Länge. Ein Vektor der Länge 1 heißt normiert.
3. Das Skalarprodukt zweier Vektoren $\vec{a}, \vec{b} \in \mathbb{R}^n$ ist definiert als: $\vec{a} \cdot \vec{b} = a_1 \cdot b_1 + a_2 \cdot b_2 + ... + a_n \cdot b_n$.

Satz 1. *Seien $\vec{a}, \vec{b} \in \mathbb{R}^n$ zwei Vektoren. Dann gilt:*

1 (Winkeldarstellung des Skalarproduktes)
$\vec{a} \cdot \vec{b} = |\vec{a}| \cdot |\vec{b}| \cdot \cos\gamma$, *wobei γ der Einschlusswinkel von $\vec{a}$ und $\vec{b}$ mit $0^o \leq \gamma \leq 180^o$ ist.*

2 (Folgerung aus 1. für $\gamma = 90^o$)
$\vec{a}, \vec{b}$ stehen senkrecht aufeinander $\Leftrightarrow \vec{a} \cdot \vec{b} = 0$.

1.2. Matrizen

Eine $m \times n$-Matrix A ist ein rechteckiges Zahlenschema mit m Zeilen und n Spalten. Man kann sie sich als Aneinanderreihung von n (Spalten-)Vektoren der Länge m vorstellen.
Unter der Transponierten A^T von A versteht man diejenige Matrix, die man erhält, wenn man die Zeilen und Spalten von A vertauscht.

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}, \qquad A^T = \begin{pmatrix} a_{1,1} & \cdots & a_{m,1} \\ \vdots & \ddots & \vdots \\ a_{1,n} & \cdots & a_{m,n} \end{pmatrix}$$

Eine $(m \times n)$ Matrix A und eine $(n \times k)$ Matrix B werden miteinander multipliziert, indem man die Zeilen von A mit den Spalten von B komponentenweise multipliziert. Dazu muss die Spaltenzahl von A mit der Zeilenzahl von B übereinstimmen und es ensteht eine $m \times k$ Matrix.

Definition 3. Eine $m \times n$ Matrix A wird mit einer $n \times k$ Matrix B wie folgt multipliziert:

$$A \cdot B = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \cdot \begin{pmatrix} b_{1,1} & \cdots & b_{1,k} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \cdots & b_{n,k} \end{pmatrix} = \begin{pmatrix} c_{1,1} & \cdots & c_{1,k} \\ \vdots & \ddots & \vdots \\ c_{m,1} & \cdots & c_{m,k} \end{pmatrix}$$

mit

$$c_{i,j} = \sum_{s=1}^{k} a_{is} \cdot b_{sj}, i = 1, \cdots, m; j = 1 \cdots, n.$$

Definition 4. Die Determinante einer 2×2 Matrix A ist definiert durch:

$$\det(A) = \det(\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}) = a_{1,1} \cdot a_{2,2} - a_{1,2} \cdot a_{2,1}$$

1.3. Parametrische Darstellung von Kurven

Definition 5. Eine stetige vektorwertige Abbildung $\vec{r} : t \in [t_1, t_2] \subset \mathbb{R} \to \vec{r}(t) \in \mathbb{R}^n$ heißt **Weg**. Als **Kurve** C bezeichnet man das Bild eines Weges: $C : \vec{r}(t),\ t \in [t_1, t_2]$.

Beispiele für $n = 2$ (ebene Kurven).

1. Ein Kreis mit dem Mittelpunkt $(0; 0)$ und dem Radius r kann durch die Gleichung $x^2 + y^2 = r^2$ im kartesischen Koordinatensystem beschrieben werden. Andererseits kann dieser Kreis auch durch eine Kurve in Abhängigkeit des Winkels definiert werden:

$$C : \vec{r}(t) = \begin{pmatrix} r \cdot cos(t) \\ r \cdot sin(t) \end{pmatrix}, t \in [0°, 360°].$$

2. Die Gerade, welche zwei Punkte P und Q aus $\mathbb{R}^n$ verbindet, lässt sich durch die folgende Kurve beschreiben:

$$\vec{r}(t) = P + t \cdot \overrightarrow{PQ},\ t \in \mathbb{R}$$

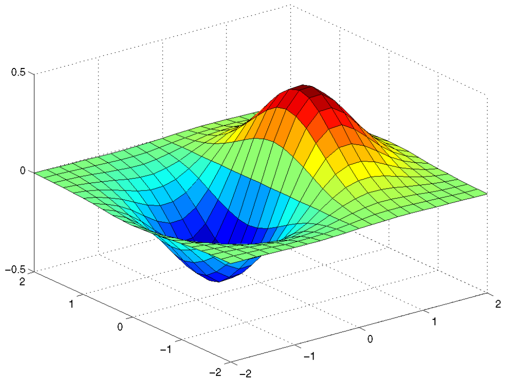
2. Reellwertige Funktion in mehreren Veränderlichen

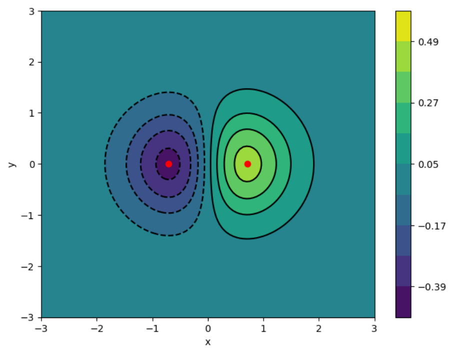
Definition 6. Eine Abbildung $f : \vec{x} = (x_1, ..., x_n) \in D \subseteq \mathbb{R}^n \to y = f(\vec{x}) \in \mathbb{R}$ heißt reellwertige Funktion in n Veränderlichen (bzw. Variablen) $(x_1, ..., x_n)$.
$(x_1, ..., x_n)$ sind die unabhängigen Variablen, y ist die abhängige Variable. D ist der Definitionsbereich. Die Menge $B = \{f(\vec{x}) | \vec{x} \in D\}$ der Bildwerte von f heißt Bildbereich.

Beispiele. 1. In der Physik werden Größen, z.B. Temperatur, Druck usw. in Abhängigkeit von Ort und Zeit angegeben.
2. In der Wirtschaft lässt sich der Gewinn eines Unternehmens aus seinen Einnahmen und Ausgaben ermitteln.

Definition 7. Sei $f : (x, y) \in D \subseteq \mathbb{R}^2 \to z = f(x, y) \in \mathbb{R}$ eine reellwertige Funktion in 2 Variablen x, y.
1. Wird eine unabhängige Variable x bzw. y konstant gesetzt, so ergibt sich eine Funktionsschar von **Schnittkurven** $z_C = f(C, y)$ bzw. $z_C = f(x, C)$, wobei $C \in \mathbb{R}$.

2. Eine **C-Höhenlinie** ist die Menge aller Punkte $P = (x, y)$, die denselben Funktonswert C haben, d.h. für die gilt $f(P) = C$.





Beispiele. 1. Eine reellwertige Funktion $f : (x,y) \in D \subseteq \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) = g(\sqrt{x^2+y^2})$, die also für alle (x,y) auf einem Kreisbogen den selben Funktionswert hat, heißt Rotationsfläche.
Die C-Höhenlinien einer Rotationsfläche sind **Kreise** mit dem Radius $\sqrt{C}$.

2. In der Metereologie werden Luftdruck- und Temperaturbereiche gleichen Drucks- bzw. gleicher Temperatur durch sogenannte Isobare bzw. Isotherme dargestellt. Diese sind nicht anderes als die Höhenlinien der Funktion, die dem Ort den Luftdruck bzw. die Temperatur zuordnet.

3. Richtungsableitung und partielle Ableitung

Bei Funktionen mehrerer Variablen ist - wie bei Funktionen in einer Variablen - das Ableiten eine wichtige Grundlage zum Untersuchen dieser Funktionen bezüglich Monotonie, Krümmung und Extrema.

3.1. Richtungsableitung

Im eindimensionalen Fall ($n = 1$) wird die Ableitung über den Differentialquotienten berechnet. Im mehrdimensionalen Fall ist es ähnlich, jedoch kann man hier in unendlich viele Richtungen ableiten. Die Ableitungen an einem Punkt P_0 verlaufen in Richtung von Geraden der Form $\vec{r}(\lambda) = P_0 + \lambda\vec{a}$ mit normiertem Richtungsvektor $\vec{a}$. Diese Ableitungen nennt man Richtungsableitungen. Sie sind gleich dem Anstieg der Tangenten an $f(P_o)$, die in Richtung der Geraden $\vec{r}(t)$ gebildet werden. Der Differenzenquotient, der den Sekantenanstieg zwischen $P_0 + \lambda\vec{a}$ und P_0 angibt, lässt sich somit in der Form:

$$\frac{f(P_0 + \lambda\vec{a}) - f(P_0)}{\lambda}$$

darstellen. Um den Anstieg der Tangente in P_0 zu finden, lässt man nun λ gegen Null laufen.

Definition 8. (Richtungsableitung) Sei $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ eine reellwertige Funktion. Sei $P_0 \in \mathbb{D}$ und f in einer Umgebung $U \subseteq D$ von P_0 stetig. Sei $\vec{a} \in \mathbb{R}$ ein Vektor der Länge 1. Dann heisst der Grenzwert (falls er existiert)

$$\frac{\partial f}{\partial \vec{a}}(P_0) = \lim_{\lambda \to 0} \frac{f(P_0 + \lambda\vec{a}) - f(P_0)}{\lambda}$$

Richtungsableitung von f im Punkt P_0 in Richtung des Vektors $\vec{a}$.

Beispiel. Sei $f(x,y) = 4 - x^2 - y^2$. Wir bestimmen den Anstieg der Tangenten an $f(x,y)$ im Punkt

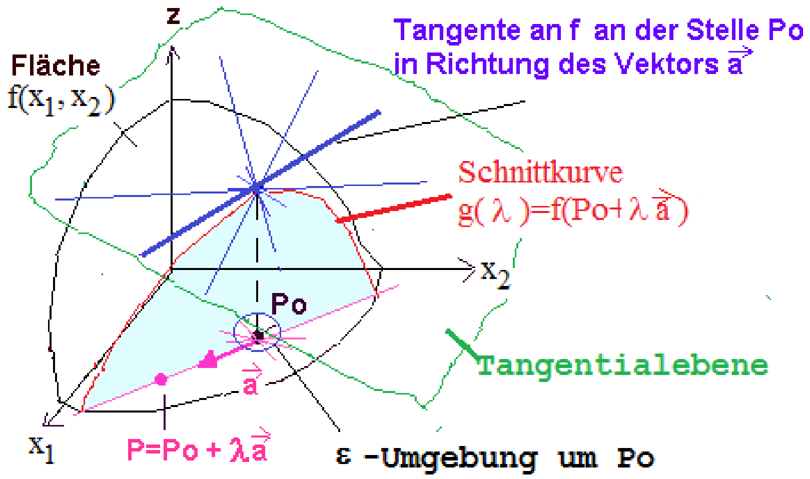


Abbildung 1: Tangente an $f(P_o)$ in Richtung $\vec{a}$ im $\mathbb{R}^2$

$P_0 = (0;1)$ in Richtung des Vektors $\vec{a} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$. Wir erhalten wegen $P_0 + \lambda \cdot \vec{a} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 0 \end{pmatrix}) = \begin{pmatrix} -\lambda \\ 1 \end{pmatrix}$:

$$\frac{\partial f}{\partial \vec{a}}(P_0) = \lim_{\lambda \to 0} \frac{f(P_0 + \lambda\vec{a}) - f(P_0)}{\lambda} = \lim_{\lambda \to 0} \frac{f(-\lambda, 1) - f(0,1)}{\lambda}$$
$$= \lim_{\lambda \to 0} \frac{4 - (-\lambda)^2 - 1^2 - (4 - 1^2)}{\lambda} = \lim_{\lambda \to 0} \lambda = 0.$$

Die Tangente an $f(P)$ im Punkt $P_0 = (0;1)$ in Richtung des Vektors $\vec{a}$ verläuft also waagerecht.

Satz 2. *(Summenregel für die Richtungsableitung)*

Seien a, b zwei reelle Zahlen sowie $f : D_f \subseteq \mathbb{R}^n \to \mathbb{R}$ und $g : D_g \subseteq \mathbb{R}^n \to \mathbb{R}$ reellwertige Funktionen von $\mathbb{R}^n$ nach $\mathbb{R}$. Sei $P_0 \in D_f \cap D_g$ und f und g in P_0 differenzierbar in Richtung $\vec{a} \in \mathbb{R}^n$. Dann gilt für die Richtungsableitung der Funktion $F : P \in D_f \cap D_g \to a \cdot f(P) + b \cdot g(P) \in \mathbb{R}$:

$$\frac{\partial F}{\partial \vec{a}}(P_0) = \frac{\partial(a \cdot f + b \cdot g)}{\partial \vec{a}}(P_0) = a \cdot \frac{\partial f}{\partial \vec{a}}(P_0) + b \cdot \frac{\partial g}{\partial \vec{a}}(P_0)$$

3.2. Partielle Ableitung

Die partielle Ableitung ist ein Spezialfall der Richtungsableitung mit einem Richtungsvektor $\vec{a}$, der parallel zu den Koordinatenachsen ist. Der Richtungsvektor, der parallel zur x_i-Achse verläuft, ist damit der i-te Einheitsvektor im $\mathbb{R}^n$ (mit einer 1 an der i.ten Stelle) und sieht (als Zeilenvektor geschrieben) wie folgt aus:

$$\vec{a}^T = (0, \dots, 0, 1, 0, \dots, 0).$$

Definition 9. (Partielle Ableitung)
Seien $y = f(x_1, ..., x_n) und P_0 = (x_1^0, ..., x_n^0)$. Der Grenzwert

$$\frac{\partial f}{\partial x_i}(P_0) = f_{x_i}(P_0) = \lim_{\lambda \to 0} \frac{f(x_1^0, ..., x_{i-1}^0, x_i^0 + \lambda, x_{i+1}^0, ..., x_n^0) - f(P_0)}{\lambda}$$

heißt partielle Ableitung 1. Ordnung von f nach x_i an der Stelle P_0.

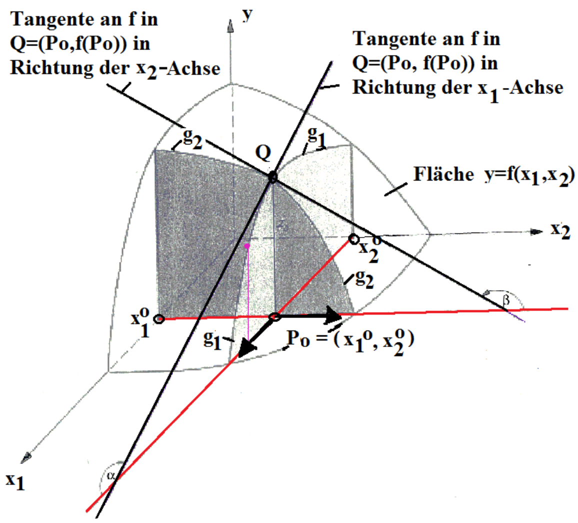


Abbildung 2: Richtungsableitung mit Einheitsvektoren

Wir sehen, dass mit dem i.ten Einheitsvektor als Richtungsvektor $f(x_1, \ldots, x_i, \ldots, x_n)$ nur noch in der einen Variable x_i variiert werden kann, die anderen bleiben konstant. Somit kann man wie bisher bei Funktionen in einer Variablen nach dieser einen Variablen x_i ableiten. Die anderen werden dabei als Konstanten betrachtet.

Die partiellen Ableitungen erster Ordnung können wir wiederum nach allen Variablen $x_i, i = 1, ..., n$ ableiten und erhalten so die partiellen Ableitungen 2. bzw. höherer Ordnung.

Bezeichnung. (Bezeichnungen partieller Ableitungen)

1 Ordnung: $f_{x_i}(P_0) = \frac{\partial f}{\partial x_i}(P_0)$, 2. Ordnung: $f_{x_i x_i}(P_0) = \frac{\partial^2 f}{\partial x_i^2}(P_0)$, $f_{x_i x_j}(P_0) = \frac{\partial^2 f}{\partial x_i \partial x_j}(P_0)$,

n. Ordnung: $f_{x_i x_i \cdots x_i}(P_0) = \frac{\partial^n f}{\partial x_i^n}(P_0)$.

4. Der Gradient und seine Eigenschaften

4.1. Der Gradient

Alle partiellen Ableitungen erster Ordnung an der Stelle P_0 fasst man zu einem Vektor zusammen, dieser wird Gradient genannt.

Definition 10. (Gradient)
Als Gradient von f and der Stelle P_0 bezeichnet man den Vektor der partiellen Ableitungen 1. Ordnung von f an der Stelle P_0:

$$\operatorname{grad} f(P_0) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(P_0) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(P_0) \end{pmatrix}.$$

4.2. Die verallgemeinerte Kettenregel

Nützliche Eigenschaften des Gradienten basieren auf der Kettenregel von Richtungsableitungen, die analog zur Kettenregel für Funktionen in einer Variablen gilt. Diese besagt, wie wir eine Funktion f in mehreren Variablen über einer Kurve $\vec{r}(t)$ ableiten können.

$$\begin{tikzcd} t \arrow[r] \arrow[dr, "g(t)"'] & \vec{r}(t) \arrow[d] \\ & f(\vec{r}(t)) \end{tikzcd}$$

Satz 3. *(Kettenregel für die Verkettung von Funktionen in mehreren Veränderlichen)*

Seien $f : \vec{x} \in D_f \subseteq \mathbb{R}^n \to f(\vec{x}) \in \mathbb{R}$ eine Funktion aus dem $\mathbb{R}^n$ in $\mathbb{R}$ und $\vec{r} : t \in D_{\vec{r}} \subseteq \mathbb{R} \to \vec{r}(t) \in D_f \subseteq \mathbb{R}^n$ eine Kurve im Definitionsbereich von f. Seien weiterhin innerhalb einer Umgebung des Kurvenpunktes $\vec{x}_0 = \vec{r}(t_0)$ die Funktionen $f(\vec{x})$ und ihre partiellen Ableitungen $\frac{\partial f}{\partial x_1}(\vec{x})$, ..., $\frac{\partial f}{\partial x_n}(\vec{x})$ definiert und stetig. Außerdem sei $\vec{r}(t)$ in t_0 differenzierbar.

Dann ist die zusammengesetzte Funktion $g(t) = f(\vec{r}(t))$ in t_0 differenzierbar und es gilt:

$$\frac{dg}{dt}(t_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{r}(t_0)) \cdot \frac{dr_i}{dt}(t_0) = \mathit{grad} f(\vec{r}(t_0)) \cdot \frac{d\vec{r}}{dt}(t_0) \tag{1}$$

4.3. Eigenschaften des Gradienten

Aus der Kettenregel ergibt sich sofort folgender nützlicher Satz.

Satz 4. *Sei $f : D_f \subseteq \mathbb{R}^n \to f(\vec{x}) \in \mathbb{R}$ eine Funktion aus $\mathbb{R}^n$ in $\mathbb{R}$ und seien innerhalb einer Umgebung von P_0 die Funktion $f(\vec{r})$ und ihre partiellen Ableitungen $\frac{\partial f(\vec{x})}{\partial x_1}$, ..., $\frac{\partial f(\vec{x})}{\partial x_n}$ definiert und (jede als Funktionen von n Veränderlichen) stetig. Sei $\vec{a}$ ein Vektor der Länge 1. Dann gilt:*

$$\frac{\partial f}{\partial \vec{a}}(P_0) = \mathit{grad}\ f(P_0) \cdot \vec{a}.$$

Beweis:

Sei $\vec{r}(t) = P_0 + t \cdot \vec{a}$. Dann gilt $\frac{d\vec{r}}{dt}(t) = \vec{a}$, da man eine Kurve $\vec{r}(t)$ komponentenweise nach t ableitet. Weiterhin gilt $\vec{r}(0) = P_0$.

Nun folgt sofort aus der Kettenregel (1):

$$\frac{\partial f}{\partial \vec{a}}(P_0) = \frac{d}{dt}(f(\vec{r}(0))) = \text{grad } f(\vec{r}(0)) \cdot \frac{d\vec{r}}{dt}(0) = \text{grad } f(P_0) \cdot \vec{a}$$

□

Satz 5. *Seien die Voraussetzungen von Satz 4 erfüllt und grad $f(P_0) \neq \vec{0}$. Dann zeigt der Vektor grad $f(P_0)$ in Richtung des steilsten Anstieges von f vom Punkt P_0 aus betrachtet.*

Beweis: Es gilt

$$\frac{\partial f}{\partial \vec{a}}(P_0) = \vec{a} \cdot \text{grad} f(P_0) = |\vec{a}| \cdot |\text{grad} f(P_0)| \cdot \cos\varphi = |\text{grad} f(P_0)| \cdot \cos\varphi$$

für alle Vektoren $\vec{a}$ mit Länge 1, wobei φ der Einschlusswinkel der beiden Vektoren ist.

Da der Betrag des Gradient unabhängig von $\vec{a}$ und damit konstant ist, hängt die Größe des Anstiegs nur von $\cos\varphi$ ab. Dieser Term ist maximal, falls $\varphi = 0$. Da φ der Einschlusswinkel der beiden Vektoren ist, folgt $\vec{a} \uparrow\uparrow \text{grad} f(P_0)$, falls $\vec{a}$ in die Richtung des stärksten Anstiegs gerichtet ist. Also zeigt auch $\text{grad} f(P_0)$ in die Richtung des stärksten Anstiegs. □

Satz 6. *Sei $f : D_f \subseteq \mathbb{R}^2 \to \mathbb{R}$ in einer Umgebung des Punktes P_0 differenzierbar.*
Sei $M_c = \{(x, y) \in D_f | f(x, y) = c\}$ eine c-Höhenlinie von f mit $P_0 \in M_c$. Sei M_c durch die Kurve $\vec{r_c}(t),\ t_1 \le t \le t_2$ beschreibbar, diese in einer Umgebung des Punktes P_0 nach t differenzierbar und es gelte $\vec{r_c}(t_0) = P_0$. Dann gilt:

$$grad\ f(P_0) \perp \frac{d\vec{r}_c}{dt}(t_0)$$

(Der Gradient von f in P_0 steht senkrecht zum Tangentenvektor an die Höhenlinie, auf der P_0 liegt.)

Beweis: Es gilt

$$\text{grad } f(P_0) \cdot \frac{d\vec{r}_c}{dt}(t_0) = \text{grad } f(\vec{r}_c(t_0)) \cdot \frac{d\vec{r}_c}{dt}(t_0) = \frac{df(\vec{r}_c(t_0))}{dt} = \frac{d(c)}{dt} = 0$$

wobei in der zweiten Gleichung die Kettenregel verwendet wurde. Da das Skalarprodukt aus Gradient und Tangentenvektor der Höhenlinie gleich 0 ist, folgt nach Satz 1 die Behauptung

$$\text{grad } f(P_0) \perp \frac{d\vec{r}_c}{dt}(t_0)$$

□

5. Tangentialebene

Mehrdimensionale Funktionen $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ kann man mithilfe der sogenannten Tangentialebene für Punkte $P \in D$ in der Nähe eines Punktes $P_0 \in D$ approximieren.
Die Gleichung dieser Tangentialebene $T : \mathbb{R}^n \to \mathbb{R}$ soll im Folgenden für den Fall $n = 2$ hergeleitet werden.

Ansatz. Die allgemeine Ebenengleichung lautet $z = ax+by+c$, weshalb man auch für die Tangentialebene - hier als $T(x, y)$ bezeichnet - schreiben kann

$$T(x, y) = ax + by + c \tag{2}$$

Wir bestimmen nun die Koeffizienten $a, b, c \in \mathbb{R}$ so, dass die Ebene $T(x, y)$ analoge Eigenschaften besitzt, wie sie die Tangenten an eine Funktion in einer Veränderlichen besitzen. Wir fordern:

1. $T(P_0) = f(P_0)$ (Funktion f und Ebene T berühren sich im Punkt P_0)
2. $T_x(P_0) = f_x(P_0)$ (Im Punkt P_0 haben f und T den gleichen Anstieg in $x-$Achsenrichtung)
3. $T_y(P_0) = f_y(P_0)$ (Im Punkt P_0 haben f und T den gleichen Anstieg in $y-$Achsenrichtung)

Aus der zweiten und dritten Bedingung folgt direkt $a = f_x(P_0)$ und $b = f_y(P_0)$. Durch Einsetzen in die Gleichung (2) ergibt sich dann $c = f(P_0) - f_x(P_0) \cdot x_0 - f_y(P_0) \cdot y_0 = f(P_0) - \text{grad } f(P_0) \cdot P_0$. Also folgt für die Tangentialebene an f an einem Punkt P_0:

$$T_{P_0}(x, y) = f_x(P_0) \cdot x + f_y(P_0) \cdot y + f(P_0) - \text{grad } f(P_0) \cdot P_0$$

$$T_{P_0}(P) = f(P_0) + \text{grad } f(P_0) \cdot \overrightarrow{P_0P}$$

Diese Form kann auch auf $n > 2$ verallgemeinert werden.

Definition 11. Für eine in $P_0 \in D \subseteq \mathbb{R}^n$ differenzierbare Funktion $f : D \to \mathbb{R}$ heisst die Ebene

$$T_{P_0}(P) = f(P_0) + \text{grad} f(P_0) \cdot \overrightarrow{P_0P}$$

Tangentialebene von f an der Stelle P_0.

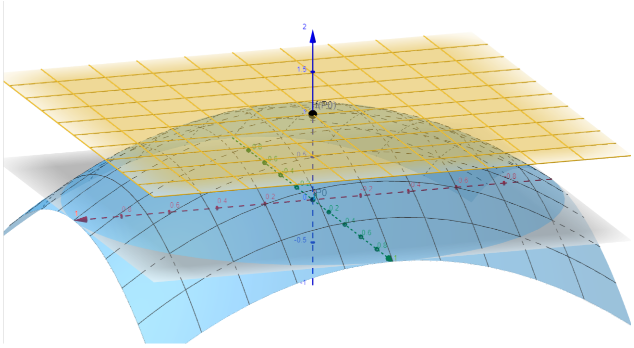


Abbildung 3: Tangentialebene an die Funktion $f(x, y) = 1 - x^2 - y^2$ an der Stelle P_0

In Abbildung 3 kann man in gelb die Tangentialebene an die Funktion $f(x) = 1 - x^2 - y^2$ an der Stelle eines Punktes P_0 erkennen. Man kann ebenfalls erkennen, dass diese die Funktion f in einer Umgebung von P_0 approximieren kann:

$$f(P) \approx f(P_0) + \mathrm{grad} f(P_0) \cdot \overrightarrow{P_0P}$$

6. Mehrdimensionaler Satz von Taylor

Definition 12. (Hesse-Matrix)
Die Matrix der 2. partiellen Ableitungen von f an der Stelle P_0

$$H(P_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(P_0) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(P_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(P_0) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(P_0) \end{pmatrix}$$

heißt Hesse-Matrix von f in P_0.

Durch den Satz von Schwarz ist die Hesse-Matrix symmetrisch, solange die 2. partiellen Ableitungen alle stetig auf ihren Definitionsbereich sind.

Satz 7. *(Satz von Schwarz)*
Sei D eine offene Menge im $\mathbb{R}^n$ und $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ eine Funktion. Falls alle partiellen Ableitungen der Ordnung $\leq k$ existieren und stetig sind, so sind alle partiellen Ableitungen bis zur Ordnung k unabhängig von der Reihenfolge des Differenzierens.

Satz 8 (Satz von Taylor)**.**
Sei $f : U \subset \mathbb{R}^n \to \mathbb{R}$ eine auf der offenen Menge U $k+1$ mal stetig differenzierbare Funktion (das heißt, alle partiellen Ableitungen von f existieren bis zur Ordnung $k+1$ und sind stetig). Seien P_0 und P zwei Punkte aus U, für die die gesamte Strecke von P_0 bis P in U liegt. Dann gilt:

(a) $f(P) = f(P_0) + \text{grad } f(P_0) \cdot \overrightarrow{P_0P} + o(|\overrightarrow{P_0P}|)$

(b) $f(P) = f(P_0) + \text{grad } f(P_0) \cdot \overrightarrow{P_0P} + \frac{1}{2}(\overrightarrow{P_0P})^T H(P_0)(\overrightarrow{P_0P}) + o(|\overrightarrow{P_0P}|^2)$

wobei $o(|\overrightarrow{P_0P}|^k)$ eine Größe ist, für die gilt: $\lim\limits_{|P-P_0|\to 0} \frac{o(|\overrightarrow{P_0P}|^k)}{|\overrightarrow{P_0P}|^k} = 0$ *für* $k = 1, 2$.

7. Extremwertbestimmung

Definition 13. Sei $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ eine Funktion in n Veränderlichen. Der Punkt $P_0 \in D$ heißt striktes lokales Minimum (Maximum) von f, falls gilt:

$$f(P_0) \underset{(>)}{<} f(P)$$

für alle Punkte $P \in D$ mit $P_0 \neq P$, die in einer ε-Umgebung $U_\epsilon(P_0)$ um den Punkt P_0 liegen. Wir sprechen von einem globalen Minimum (Maximum), wenn die jeweilige Ungleichung auf dem gesamten Definitionsbereich von f gilt.

Definition 14. Sei $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$. Dann heißen Punkte $P_0 \in D$ mit $\operatorname{grad} f(P_0) = \vec{0}$ **stationäre Punkte** von f.

7.1. Hinreichende Bedingungen für einen lokalen Extremwert

Definition 15. Eine $n \times n$-Matrix A heißt positiv (bzw. negativ) definit, falls für alle Vektoren $\vec{h} \in \mathbb{R}^n$ ungleich dem Nullvektor gilt:

$$\vec{h}^T A \vec{h} \underset{(<)}{>} 0.$$

Andernfalls heißt sie indefinit.

Satz 9. *Sei D eine offene Menge und $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$. Weiterhin nehmen wir an, dass auf D alle partiellen Ableitungen von f bis zur 2. Ordnung existieren und stetig sind. Sei $P_0 \in D$. Dann gilt: Wenn*

1. *grad $f(P_0) = \vec{0}$ (P_0 ist ein stationärer Punkt) und*
2. *Hesse-Matrix $H(P_0)$ ist negativ (bzw. positiv) definit,*

so nimmt f in P_0 ein striktes lokales Maximum (bzw. striktes lokales Minimum) an.

Beweisidee. Vorausgesetzt wird nach 1., dass P_0 ein stationärer Punkt ist und 2. die Hesse-Matrix in P_0 negativ (bzw. positiv) definit ist. Nach dem Satz von Taylor (Satz 8) folgt dann:

$$f(P) = f(P_0) + \underbrace{(\text{grad } f(P_0))^T \cdot \overrightarrow{P_0P}}_{=\,0} + \underbrace{\frac{1}{2}\overrightarrow{P_0P}^T \cdot H(P_0)\overrightarrow{P_0P}}_{\underset{(>)}{<}\; 0} + \; \mathrm{o}\left(|\overrightarrow{P_0P}|^2\right)$$

Da $o\left(|\overrightarrow{P_0P}|^2\right)$ schneller als quadratisch, d.h. schneller als der quadrierte Abstand $|\overrightarrow{P_0P}|^2$ gegen 0 geht und $(\overrightarrow{P_0P})^T \cdot H(P_0) \cdot \overrightarrow{P_0P}$ nur quadratisch, d.h. genauso schnell wie $|\overrightarrow{P_0P}|^2$ gegen 0 geht, folgt für alle Punkte P in einer gewissen ε-Umgebung von P_0:

$$f(P) \underset{(>)}{<} f(P_0)$$

und damit die Behauptung des Satzes.

7.2. Determinantenkriterium für 2 Variablen

Für Funktionen $f(x,y)$ in 2 Variablen haben wir ein praktikableres Kriterium anstelle der positiven bzw. negativen Definitheit der Hesse-Matrix verwendet. Die Hesse-Matrix der 2. partiellen Ableitungen in $P=(x,y)$ hat folgende Gestalt

$$H(P) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(P) & \frac{\partial^2 f}{\partial x \partial y}(P) \\ \frac{\partial^2 f}{\partial x \partial y}(P) & \frac{\partial^2 f}{\partial y^2}(P) \end{pmatrix}$$

und ihre Determinante ist $\det(H(P)) = \frac{\partial^2 f}{\partial x^2}(P) \cdot \frac{\partial^2 f}{\partial y^2}(P) - \frac{\partial^2 f}{\partial x \partial y}(P) \cdot \frac{\partial^2 f}{\partial x \partial y}(P)$. Wir haben nun folgenden Satz verwendet.

Satz 10.
Wenn $\det(H(P_0)) > 0$ und $\frac{\partial^2 f}{\partial x_1^2}(P_0) > 0$, so ist $H(P_0)$ positiv definit.
Wenn $\det(H(P_0)) > 0$ und $\frac{\partial^2 f}{\partial x_1^2}(P_0) < 0$, so ist $H(P_0)$ negativ definit.

Daraus ergibt sich mit Satz 9 das folgende hinreichende Extremwertkriterium:

Satz 11. *(Hinreichendes Extremwertkriterium im $\mathbb{R}^2$)*

1. *Wenn grad $f(P_0) = \vec{0}$, $\det(H(P_0)) > 0$ und $\frac{\partial^2 f}{\partial x_1^2}(P_0) < 0$, so ist P_0 ein lokales Maximum von f.*

2. *Wenn grad $f(P_0) = \vec{0}$, $\det(H(P_0)) > 0$ und $\frac{\partial^2 f}{\partial x_1^2}(P_0) > 0$, so ist P_0 ein lokales Minimum von f.*

Bemerkung Es gilt auch: Ist grad $f(P_0) = \vec{0}$ und $\det(H(P_0)) < 0$, so liegt in P_0 ein Sattelpunkt vor.

8. Anwendungsbeispiel

Wir ermitteln nun die Extremwerte der Funktion $f(x,y) = x \cdot e^{-(x^2+y^2)}$ mittels Satz 11.

1. Partielle Ableitungen von f ermitteln

$$\frac{\partial f}{\partial x}(x,y) = (1-2x^2) \cdot e^{-(x^2+y^2)}$$
$$\frac{\partial f}{\partial y}(x,y) = -2xy \cdot e^{-(x^2+y^2)}$$

2. Stationäre Punkte bestimmen

$$\frac{\partial f}{\partial x}(x,y) = (1-2x^2) \cdot e^{-(x^2+y^2)} = 0$$
$$\frac{\partial f}{\partial y}(x,y) = -2xy \cdot e^{-(x^2+y^2)} = 0$$

Da $e^{-(x^2+y^2)}$ eine e-Funktion ist, kann diese nicht Null werden. Damit folgt aus der ersten Gleichung: $1-2x^2 = 0 \iff x = \pm\frac{1}{\sqrt{2}}$.

Für beide gefundenen x-Werte ergibt sich aus der 2. Gleichung $y=0$, da $-2xy=0$ sein muss. Wir erhalten damit genau zwei stationäre Punkte $P_1(\frac{1}{\sqrt{2}};0)$ und $P_2(-\frac{1}{\sqrt{2}};0)$.

3. Hesse-Matrix in Punkten P_1 und P_2 bestimmen.

   Aufgrund des Satzes von Schwarz müssen wir nur drei unterschiedliche partielle Ableitungen 2. Grades berechnen.

$$\frac{\partial^2 f}{\partial x^2}(x,y) = -4x \cdot e^{-(x^2+y^2)} + (1-2x^2)(-2x) \cdot e^{-(x^2+y^2)}$$
$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = 8xy \cdot e^{-(x^2+y^2)} - 2ye^{-(x^2+y^2)} + 4x^2ye^{-(x^2+y^2)}$$
$$\frac{\partial^2 f}{\partial y^2}(x,y) = -2xe^{-(x^2+y^2)} - 2xy(-2y)e^{-(x^2+y^2)}$$

   Setzen wir P_1 und P_2 in diese 3 Gleichungen ein, so erhalten wir folgende Hessematrizen:

$$H(P_1) = \begin{pmatrix} -\frac{4}{\sqrt{2}}(e^{-\frac{1}{\sqrt{2}}}) & 0 \\ 0 & -\frac{2}{\sqrt{2}}(e^{-\frac{1}{2}}) \end{pmatrix} \text{ und } \; H(P_2) = \begin{pmatrix} \frac{4}{\sqrt{2}}(e^{-\frac{1}{\sqrt{2}}}) & 0 \\ 0 & \frac{2}{\sqrt{2}}(e^{-\frac{1}{2}}) \end{pmatrix}$$

4. Entscheidung mittels Determinantenkriterium (Satz 11).

   Für beide Punkte P_1 und P_2 ist die Determinate der Hessematrix > 0. Für den Punkt P_1 ist $\frac{\partial^2 f}{\partial x^2}(P_1) < 0$, folglich ist $P_1(\frac{1}{\sqrt{2}}; 0)$ ein lokales Maximum. Für den Punkt P_2 ist $\frac{\partial^2 f}{\partial x^2}(P_2) > 0$, folglich ist $P_2(-\frac{1}{\sqrt{2}}; 0)$ ein lokales Minimum.