Reellwertige Funktoinen mehrerer Variabler

Teilnehmer:

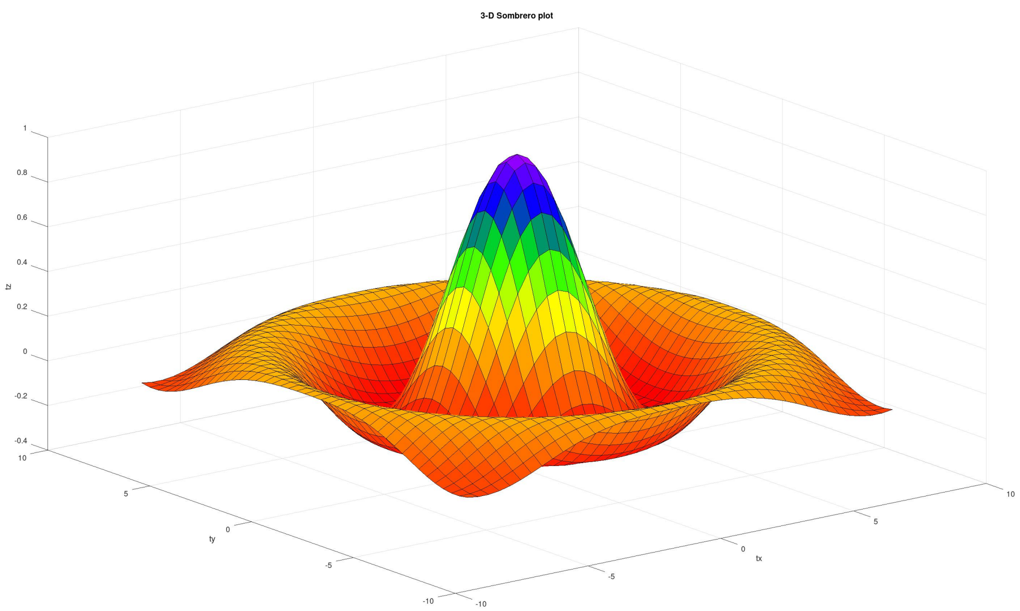
1 Teilnehmer des	Andreas-Gymnasiums
2 Teilnehmende des	Heinrich-Hertz-Gymnasiums
3 Teilnehmnende des	Herder-Gymnasiums
2 Teilnehmende des	Immanuel-Kant-Gymnasiums
2 Teilnehmende des	Käthe-Kollwitz-Gymnasiums

mit tatkräftiger Unterstützung durch:

Antje Maeß und Falk Ebert	Herder-Gymnasium

Gruppenleiter:

Barbara Grabowski	HTW des Saarlandes Saarbrücken

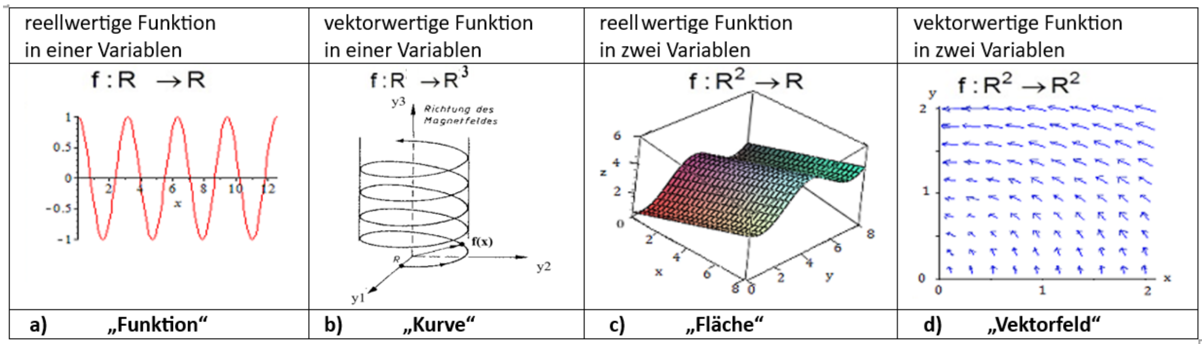


$$f(x, y) = \frac{sin(\sqrt{x^2 + y^2})}{\sqrt{x^2 + y^2}}$$

1. Einleitung

In der Schule haben wir uns ausführlich mit Funktienen $y = f(x)$ in ***einer*** Variablen x befasst. Viele Probleme aus Wirtschaft, Technik und Wissenschaft lassen sich aber durch Funktionen in einer Variablen nicht lösen. Bei den meisten praktischen Problemen der realen Welt treten Funktionen ***mehrerer Variabler*** auf und es geht um die Beschreibung des Zusammenhangs zwischen mehreren Eingangsvariablen $x_1, \ldots, x_n$ (den sogenannten unabhängigen Variablen) und m Ausgangsvariablen $y_1, \ldots, y_m$ (den sogenannten abhängigen Variablen).

Die Abbildung unten zeigt einige Beispiele:



Wir haben uns in unserer Arbeitsgruppe diese Woche speziell mit reellwertigen Funktionen in mehreren Veränderlichen beschäftigt. Wir haben uns Grundkenntnisse der Differentialrechnung für solche Funktionen erarbeitet und diese auf einige Beispiele angewandt.

Einen besonderer Fokus lag dabei auf der Veranschaulichung der genannten Begriffe für Funktionen $y = f(x, y)$ in 2 reellen Variablen x und y .

2. Grundlagen

2.1. Vektorrechnung

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

Definition 1. (Verbindungsvektor)
Seien $A = (a_1, \ldots, a_n)$ und $B = (b_1, \ldots, b_n)$ zwei Punkte aus $\mathbb{R}^n$. Das Gebilde $\vec{v}$ für welches gilt, dass $\vec{v}$ die gerichtete Strecke zwischen A und B ist, heißt Vektor von A nach B.

Schreibweise: $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \overrightarrow{AB}$, mit $v_i = b_i - a_i$ für $i = 1, \ldots, n$.

Definition 2. (Rechenoperationen mit Vektoren)
Die Addition und Subtraktion zweier Vektoren, sowie die Multiplikation eines Vektors mit einer reellen Zahl (Skalar genannt) erfolgt komponentenweise.

Definition 3. (Betrag eines Vektors)
Für einen Vektor $\vec{a} \in \mathbb{R}^n$ heißt

$$|\vec{a}| = \sqrt{a_1^2 + a_2^2 + ... + a_n^2}$$

Betrag (bzw. Norm) des Vektors $\vec{a}$. Ein Vektor mit Betrag 1 heißt normiert. Der Betrag eines Vektors beschreibt die Länge des Vektors vom Fußpunkt bis zur Spitze.

Definition 4. (Skalarprodukt)
Das Skalarprodukt zweier Vektoren $\vec{a}, \vec{b} \in \mathbb{R}^n$ ist definiert als:

$$\vec{a} \cdot \vec{b} = a_1 \cdot b_1 + a_2 \cdot b_2 + ... + a_n \cdot b_n$$

Das Skalarprodukt zweier Vektoren beschreibt den Winkel $\varphi = \measuredangle(\vec{a}, \vec{b}) \in [0, \pi]$ zwischen den Vektoren $\vec{a}$ und $\vec{b}$. Es gilt folgender Satz:

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 $\gamma = 90^o$)*
  $\vec{a} \perp \vec{b} \Leftrightarrow \vec{a} \cdot \vec{b} = 0$.

2.2. Matrizen

Definition 5. (Matrizen)
Eine $m \times n$ -Matrix A ist ein rechteckiges Zahlenschema mit m Zeilen und n Spalten. Unter der Transponierten A^T von A versteht man diejenige Matrix, die man erhält, wenn man die Zeilen und Spalten von A vertauscht.

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

A heißt symmetrisch, wenn $A = A^T$.

Definition 6. (Addition und Subtraktion zweier Matrizen)
Die Addition und Subtraktion zweier Matrizen erfolgt wieder komponentenweise. Dazu müssen ihre Zeilen- und Spaltenzahl übereinstimmen.
Ebenfalls erfolgt die Multiplikation einer Matrix mit einer reellen Zahl komponentenweise.

Definition 7. (Multiplikation zweier Matrizen)
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, d.h. jeweils das Skalarprodukt jeder Zeile von A mit jeder Spalte von B bildet. Dazu muss die Spaltenzahl von A mit der Zeilenzahl von B übereinstimmen und es ensteht eine $m \times k$ Matrix.

Für 2x2-Matrizen definieren wir noch die sogenannte ***Determinante*** der Matrix.

Definition 8. (Determinante einer 2x2 Matrix)
Die Determinante einer 2×2 Matrix A ist definiert durch:

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

2.3. Kurven

Definition 9. (Kurven)
Eine stetige vektorwertige Abbildung $\vec{r} : t \in T = [t_1, t_2] \subset \mathbb{R} \to \vec{r}(t) \subseteq B \in \mathbb{R}^n$ heißt **Weg**. Als **Kurve** C bezeichnet man die Bildmenge eines Weges und schreibt: $C : \vec{r}(t),\ t \in [t_1, t_2]$. t ist der Parameter der Kurve C und T ihr Parameterbereich.

Definition 10. (Ableitung einer parametrisierten Kurve)
Sei $\vec{r}(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T : t \in T \subset \mathbb{R} \to \mathbb{R}^n$ eine parametrisierte Kurve. Dann heißt der Vektor seiner komponentenweisen Ableitungen nach t

$$\vec{r}'(t) = \frac{d\vec{r}(t)}{dt} := \left(x_1'(t), x_2'(t), \ldots, x_n'(t)\right)^T = \lim_{h \to 0} \frac{\vec{r}(t+h) - \vec{r}(t)}{h}$$

Tangentialvektor von $\vec{r}(t)$ im Parameter $t \in T$.
Für die Ableitung einer Kurve nach einem Parameter wird i.A. auch die Bezeichnung

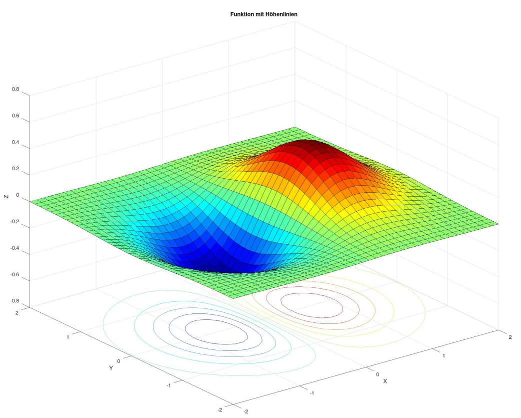
$$\dot{\vec{r}}(t), \text{ statt } \vec{r}'(t)$$

verwendet.

3. Funktionen mehrere Variablen

Definition 11. (Funktionen mehrerer Variablen)
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.

Definition 12. (Höhenlinien)
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.
Eine c-Höhenlinie ist die Menge aller Punkte $P = (x, y)$, die denselben Funktonswert c haben, d.h. für die gilt $f(P) = c$.



Da Höhenlinien darstellen, an welchen Stellen die Funktion bestimmte Funktionswerte c erreicht, helfen sie auch Extrempunkte zu finden. Durch sie kann man Intervalle um die Extrempunkte ablesen. In der Abbildung sieht man die Funktion $f(x, y) = x \cdot e^{-x^2-y^2}$ und ihre zugehörigen Höhenlinien.

4. Richtungsableitungen

Die Richtungsableitung ist der Anstieg einer Tangenten an $f(P_0)$ über einer Richtungsgeraden $P_0 + \lambda\vec{a}, \lambda \in \mathbb{R}$.
Der Anstieg einer Sekanten zwischen $f(P_0 + \lambda\vec{a})$ und $f(P_0)$ lässt sich in der Form:

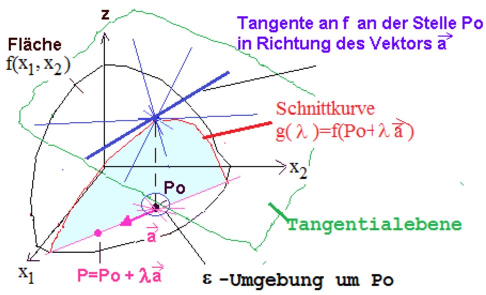
$$\frac{f(P_0 + \lambda\vec{a}) - f(P_0)}{\lambda}$$

darstellen. Der Anstieg der Tangente an f im Punkt P_0 ist dann der Grenzwert des Sekantenanstiegs für λ gegen Null.

Definition 13. (Richtungsableitung)
Sei $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ eine reellwertige Funktion. Sei $P_0 \in D$ 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 1. Sei $f(x, y) = 4 - x^2 - y^2$. Wir bestimmen den Anstieg der Tangenten an $f(x, y)$ im Punkt $P_0 = (0; 1)$ in Richtung des Vektors $\vec{a} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$.
Wir erhalten wegen

$$P_0 + \lambda \cdot \vec{a} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\lambda \\ 1 \end{pmatrix}$$

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

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

5. Partielle Ableitungen und Gradient

5.1. Partielle Ableitung

Die partielle Ableitung ist ein Spezialfall der Richtungsableitung mit einem Richtungsvektor $\vec{a}$, der parallel zu den Koordinatenachsen ist. Es wird nur nach einer Variablen x_i abgeleitet, die anderen Variablen werden dabei als Konstanten betrachtet.

Definition 14. (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.

Beispiel 2. Sei $f(x_1, x_2) = x_1 \cdot x_2 + x_1^2$. Wir leiten nach den Variablen x_1, x_2 ab.

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = x_2 + 2x_1$$
$$\frac{\partial f}{\partial x_2}(x_1, x_2) = x_1$$

5.2. Gradient

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

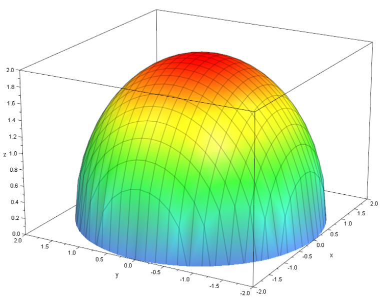
Definition 15. (Gradient)
Als Gradient von f and der Stelle P_0 bezeichnet man den Vektor der partiellen Ableitungen 1. Ordnung von f an der Stelle P_0:

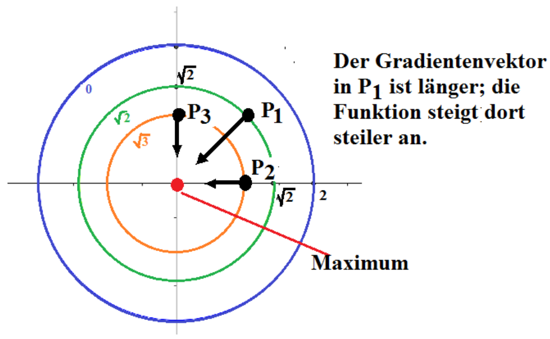
$$\text{grad } f(P_0) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(P_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(P_0) \end{pmatrix}.$$

Beispiel 3. Sei $f(x, y) = 4 - x^2 - y^2$. Wir bestimmen den Gradienten von f.

$$\frac{\partial f}{\partial x}(x, y) = -2x \qquad \frac{\partial f}{\partial y}(x, y) = -2y$$

$$\text{grad } f(x, y) = \begin{pmatrix} -2x \\ \\ -2y \end{pmatrix} \qquad \text{grad } f(P_0) = \begin{pmatrix} 0 \\ \\ -2 \end{pmatrix}$$



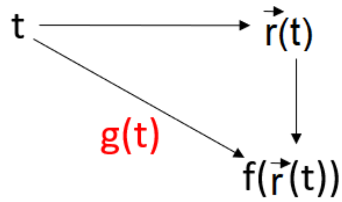


Sichtbare Eigenschaften des Gradienten:
1. Der Gradientenvektor in einem Punkt P_0 zeigt in die Richtung des steilsten Anstiegs von f vom Punkt P_0 aus betrachtet.
2. Der Gradientenvektor in einem Punkt P_0 steht senkrecht zum Tangentenvektor an die Höhenlinie im Punkt P_0.

5.3. Die verallgemeinerte Kettenregel

Die Kettenregel besagt, wie wir eine Funktion f in mehreren Variablen über einer Kurve $\vec{r}(t)$ ableiten können.

Satz 2. *(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:



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

Mit Hilfe der Kettenregel haben wir die folgenden Eigenschaften des Gradienten bewiesen.

5.4. Eigenschaften des Gradienten

Satz 3. *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 $f(\vec{x})$ 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:*

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

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) mit $g(t) = f(\vec{r}(t))$ und $t_0 = 0$:

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

□

Satz 4. *Seien die Voraussetzungen von Satz 3 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.*

Satz 5. *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 $\vec{r_c}(t)$, $t_1 \leq t \leq t_2$ beschreibbar, diese in einer Umgebung des Punktes P_0 nach t differenzierbar und es gelte $\vec{r_c}(t_0) = P_0$. Dann gilt:

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

5.5. Tangentialebene

Wir benutzen die Tangentialebene, um $f(P)$ für Punkte $P \in D$ in der Nähe eines Punktes $P_0 \in D$ zu approximieren.

Ansatz. Die allgemeine Ebenengleichung in $\mathbb{R}^2$ lautet, auch für die Tangentialebene,

$$T(x, y) = ax + by + c$$

Wir können Eigenschaften von Tangenten auf die Tangentialebene übertragen:

1. $T(P_0) = f(P_0)$ (Funktion f und Ebene T berühren sich im Punkt P_0)
2. $\frac{\partial T}{\partial x}(P_0) = \frac{\partial f}{\partial x}(P_0)$ (Im Punkt P_0 haben f und T den gleichen Anstieg in $x-$Achsenrichtung)
3. $\frac{\partial T}{\partial y}(P_0) = \frac{\partial f}{\partial y}(P_0)$ (Im Punkt P_0 haben f und T den gleichen Anstieg in $y-$Achsenrichtung)

Aus der zweiten und dritten Bedingung folgt direkt $a = \frac{\partial f}{\partial x}(P_0)$ und $b = \frac{\partial f}{\partial y}(P_0)$. Durch Einsetzen in die Ebenengleichung ergibt sich dann $c = f(P_0) - \frac{\partial f}{\partial x}(P_0) \cdot x_0 - \frac{\partial f}{\partial y}(P_0) \cdot y_0$. Also folgt für die Tangentialebene an f an einem Punkt P_0:

$$\begin{aligned} T_{P_0}(x,y) &= f_x(P_0) \cdot (x - x_0) + f_y(P_0) \cdot (y - y_0) + f(P_0) \\ T_{P_0}(P) &= f(P_0) + \text{grad } f(P_0) \cdot \overrightarrow{P_0P} \end{aligned}$$

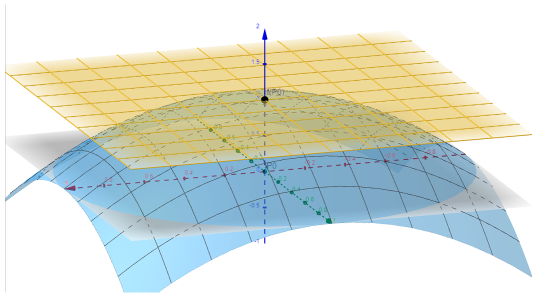
Definition 16. Für eine in $P_0 \in D \subseteq \mathbb{R}^n$ differenzierbare Funktion $f : D \to \mathbb{R}$ heißt die Ebene

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

Tangentialebene von f an der Stelle P_0.

Wir können $f(P)$ in einer Umgebung des Punktes P_0 durch $T_{P_0}(P)$ annähern. Je weiter aber P von P_0 entfernt ist, desto größer ist der Fehler der Approximation. Wir schreiben für die Approximation:

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



Deswegen benötigen wir eine genauere Formel zur Bestimmung von $f(P)$ ausgehend von $f(P_0)$.

6. Hessematrix

Definition 17. (Hesse-Matrix)
Die Matrix der 2. partiellen Ableitungen von f an der Stelle P_0

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

heißt Hesse-Matrix von f in P_0.

Durch den folgenden Satz von Schwarz ist die Hesse-Matrix symmetrisch.

Satz 6. *(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.

7. Taylor-Polynome

Satz 7. *Sei $I = (x_0 - \varepsilon, x_0 + \varepsilon)$ und $f : I \to \mathbb{R}$ eine $(k+1)$-mal stetig differenzierbare Funktion auf I. Dann existiert für alle $x \in I$ eine Zahl ξ zwischen x und x_0, so dass sich $f(x)$ wie folgt darstellen lässt:*

$$f(x) = f(x_0) + \frac{f^{'}(x_0)}{1!}(x-x_0)^1 + \frac{f^{''}(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + o(|x-x_0|^k)$$

$o(|x-x_0|^k)$ ist ein Ausdruck für den gilt:

$$\lim_{|x-x_0|\to 0} \frac{o(|x-x_0|^k)}{|x-x_0|^k} = 0,$$

d.h. das $o(|x-x_0|^k)$ strebt schneller gegen 0 als $|x-x_0|^k$; folglich wird die Approximation $f(x) \approx T_{k,x_0}$ mit wachsendem k immer besser.

Der Satz von Taylor gilt auch allgemein für reellwertige Funktionen in $n \geq 1$ Variablen:

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

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

8. Extremwerte

Definition 18. (Extremwerte)
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.

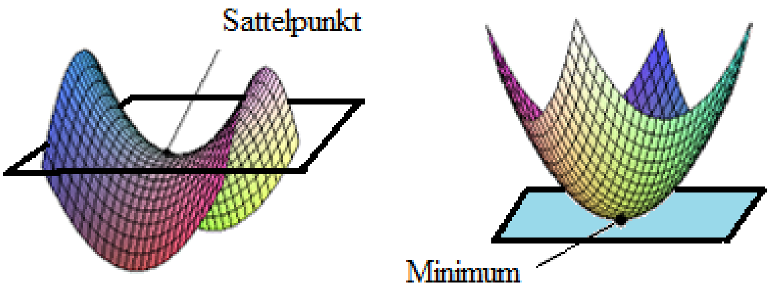
Um solche Punkte zu bestimmen, wird bei Funktionen mit einer Variablen eine notwendige und eine hinreichende Bedingung benötigt. Bei Funktionen mit mehreren Variablen ist das genauso. Um für diese Funktionen eine notwendige Bedingung zu formulieren, benötigt man zuerst den Begriff der stationären Punkte.

Definition 19. (Stationäre Punkte)
Sei $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$. Dann heißen Punkte $P_0 \in D$ mit grad $f(P_0) = \vec{0}$ stationäre Punkte von f.

Satz 9. *(Notwendiges Kriterium)*
Ein Extrempunkt ist immer ein stationärer Punkt. An jedem Extrempunkt (P_0) gilt also:

$$\mathit{grad}\ f(P_0) = \vec{0}$$

Da dieses Kriterium aber nur notwendig und nicht hinreichend ist, ist nicht jeder Punkt der dieser Bedingung erfüllt auch ein Extrempunkt. Sattelpunkte erfüllen ebenfalls das notwendige Kriterium.



Um eine hinreichende Bedingung aufzustellen, benötigt man den Begriff der Definitheit von Matrizen.

Definition 20. (Definitheit von Matrizen)
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:

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

Satz 10. *(Hinreichende Bedingung)*
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. Dieser Satz wird mithilfe vom Satz von Taylor bewiesen. Nach Taylor gilt nämlich

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

Da $o\left(|\overrightarrow{P_0P}|^2\right)$ schneller als quadratisch gegen 0 geht und $(\overrightarrow{P_0P})^T \cdot H(P_0) \cdot \overrightarrow{P_0P}$ nur quadratisch 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.

9. Determinantenkriterium für $\mathbb{R}^2$

Die hinreichende Bedingung für die Überprüfung eines Extremwertes benötigt die positive/negative Definitheit der Hesse-Matrix. Die positive Definitheit einer Hesse-Matrix zu bestimmen ist nicht immer einfach, vor allem im $\mathbb{R}^n$, $n > 2$. Für den R^2 lässt sich hierfür aber das Determinantenkriterium verwenden:

Die Hesse-Matrix der 2. partiellen Ableitungen in $P = (x, y)$ hat folgende Gestalt:

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

und ihre Determinante ist $\det(H(P)) = \frac{\partial^2 f}{\partial x^2}(P) \cdot \frac{\partial^2 f}{\partial y^2}(P) - \left(\frac{\partial^2 f}{\partial x \partial y}(P)\right)^2$. Diese benötigt man für das Determinantenkriterium:

Satz 11. *Determinantenkriterium*

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

Hieraus ergibt sich eine neue hinreichende Bedingung

Satz 12. *(Hinreichendes Extremwertkriterium für $f(x, y)$)*

(a) Wenn $\operatorname{grad} f(P_0) = \vec{0}$, $\det(H(P_0)) > 0$*, so ist P_0 ein lokales Maximum von f für* $\frac{\partial^2 f}{\partial x^2}(P_0) < 0$.
(b) Wenn $\operatorname{grad} f(P_0) = \vec{0}$, $\det(H(P_0)) > 0$*, so ist P_0 ein lokales Minimum von f für* $\frac{\partial^2 f}{\partial x^2}(P_0) > 0$.

Andere Fälle: *Ist* $\operatorname{grad} f(P_0) = \vec{0}$ *und* $\det(H(P_0)) < 0$*, so liegt in P_0 ein Sattelpunkt vor.*
Wenn $\det(H(P_0)) = 0$*, so lässt sich mit diesem Kriterium keine Aussage über den Punkt P_0 treffen und erfordert eine komplexere Betrachtung.*

10. Beispiele

Beispiel 1. Es sollen die Extrempunkte der Funktion $f(x, y) = xe^{-(x^2+y^2)}$ bestimmt werden.
Für die notwendige Bedingung wird folgendes Gleichungssystem gelöst:

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

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

Da die e-Funktion immer positiv ist, ergibt sich hierraus $x = \pm\frac{1}{\sqrt{2}}$ und $y = 0$. Die zu überprüfenden Punkte sind also $P_1(\frac{1}{\sqrt{2}}, 0)$ und $P_2(-\frac{1}{\sqrt{2}}, 0)$.

Aufgrund des Satzes von Schwarz müssen wir nur die unterschiedlichen partiellen Ableitungen 2. Grades berechnen.

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

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

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

Setzt man die Punkte P_1 und P_2 ein, so erhält man für die Hesse-Matrizen

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

Die Determinanten beider Hessematrizen sind > 0. In In der ersten Matrix $H(P_1)$ ist $\frac{\partial^2 f}{\partial x^2}(x,y)$ negativ, in $H(P_2)$ positiv.
Daraus folgt: $P_1 = P(\frac{1}{\sqrt{2}}; 0)$ ist ein lokales Maximum und $P_2 = P(-\frac{1}{\sqrt{2}}; 0)$ ist ein lokales Minimum.

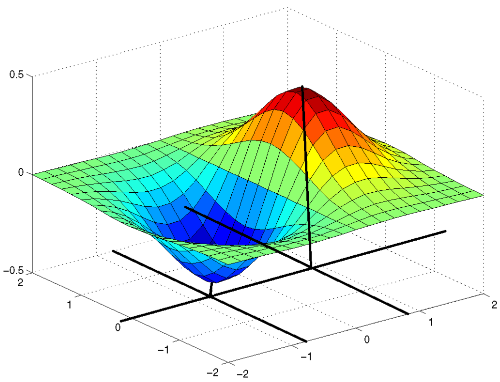


Abbildung 1: Die betrachtete Funktion

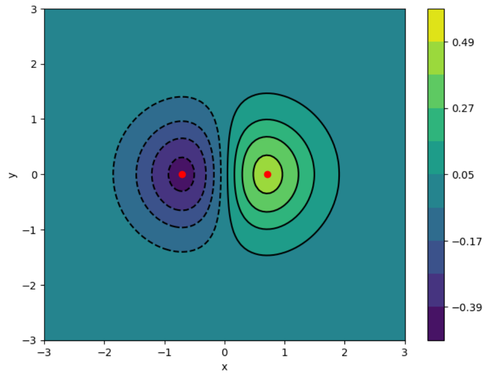


Abbildung 2: Höhenliniendiagramm der Funktion

Beispiel 2. Es soll der Quader mit der minimalen Oberfläche gefunden werden, welcher ein bestimmtes Volumen V_0 hat. Es soll also das Minimum der Funktion

$$f(a,b,c) = 2(ab + ac + bc)$$

unter der Nebenbedingung $abc = V_0$ gefunden werden. Diese Funktion wird mithilfe der Nebenbedingung auf eine Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ reduziert:

$$f(b,c) = 2\left(\frac{V_0}{b} + bc + \frac{V_0}{c}\right)$$

Um Punkte zu finden, an denen der Gradient verschwindet, muss folgendes Gleichungssystem gelöst werden:

$$\frac{\partial f}{\partial b}(b,c) = 2\left(-\frac{V_0}{b^2} + c\right) = 0 \quad \text{und} \quad \frac{\partial f}{\partial c}(b,c) = 2\left(-\frac{V_0}{c^2} + b\right) = 0$$

Durch das Lösen dieses Gleichungssystems ergibt sich $b = \sqrt[3]{V_0}$, $c = \sqrt[3]{V_0}$.

Für die hinreichende Bedingung wird die Hesse-Matrix mit den 2. partiellen Ableitungen aufgestellt:

Partielle Ableitungen:

$$\frac{\partial^2 f}{\partial b^2} = \frac{4V_0}{b^3}$$
$$\frac{\partial^2 f}{\partial c^2} = \frac{4V_0}{c^3}$$
$$\frac{\partial^2 f}{\partial b \partial c} = \frac{\partial^2 f}{\partial c \partial b} = 2$$

Hesse-Matrix:

$$H_f(b,c) = \begin{pmatrix} \frac{4V_0}{b^3} & 2 \\ 2 & \frac{4V_0}{c^3} \end{pmatrix}$$

Setzt man $b = c = \sqrt[3]{V_0}$ ein, so ergibt sich:

$$H_f = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix} \quad \Rightarrow \quad \det(H_f) = 16 - 4 = 12 > 0$$

Da $\frac{\partial^2 f}{\partial b^2} = 4 > 0$ ist, folgt, dass es sich beim Punkt $P(\sqrt[3]{V_0}; \sqrt[3]{V_0})$ um ein lokales Minimum der reduzierten Funktion $f(b,c)$ handelt.
Aus der Nebenbedingung folgt dann auch $a = \sqrt[3]{V_0}$. Der optimale Quader, der mit minimaler Oberfläche das Volumen V_o besitzt, ist also ein Würfel mit der Kantenlänge $\sqrt[3]{V_0}$.