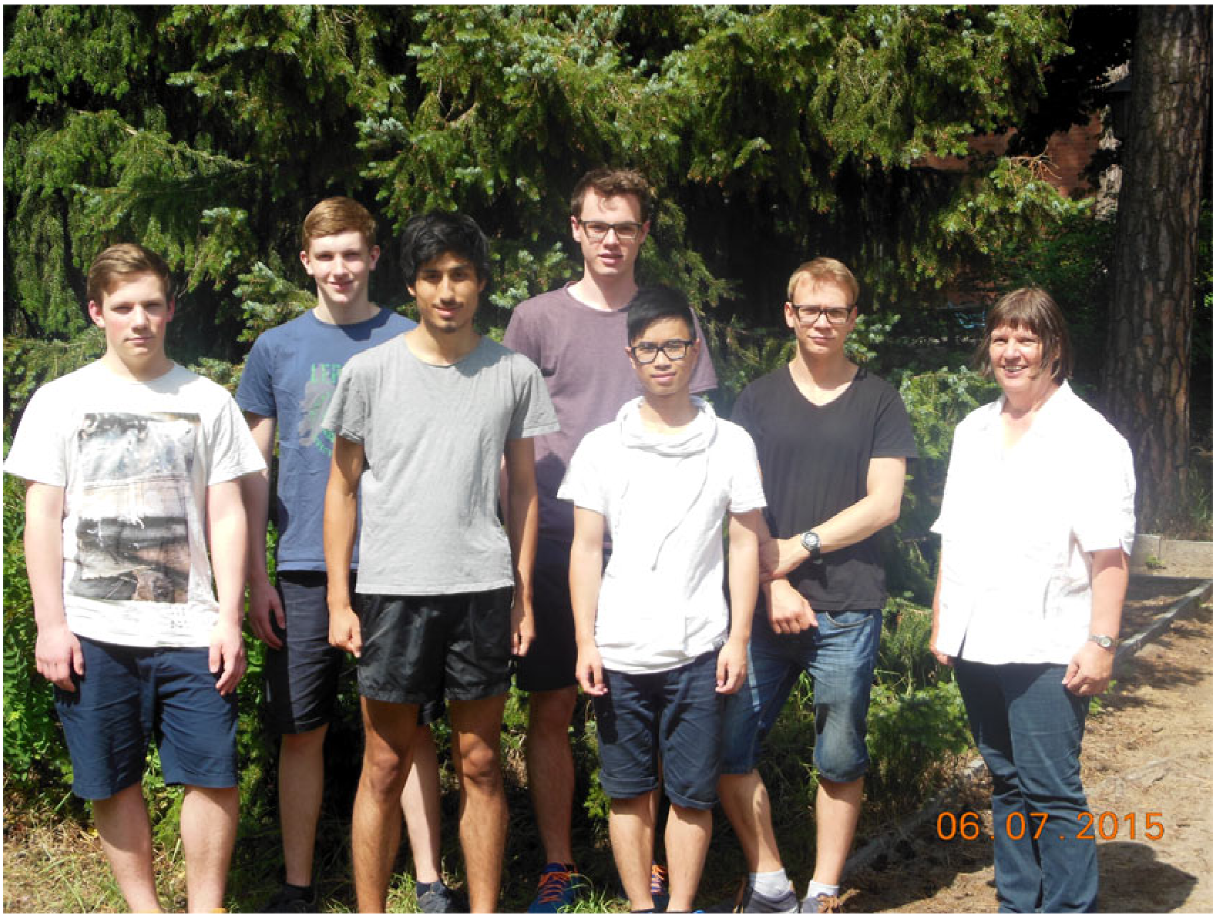
Reellwertige Funktionen mehrerer Variabler

Teilnehmer:

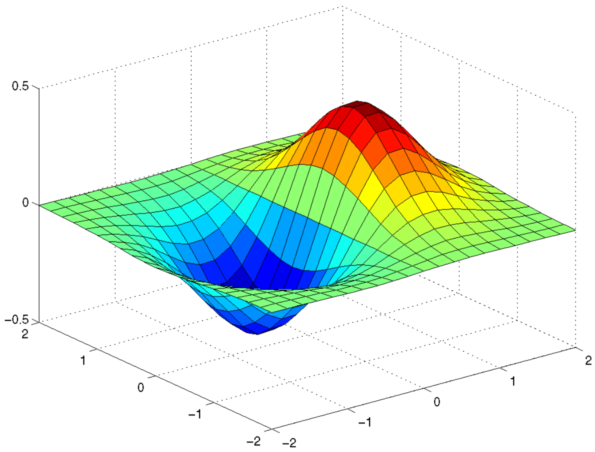
Maximilian Ringleb	Herder-Oberschule, Berlin
Jakob Napiontek	Heinrich-Hertz-Oberschule, Berlin
Kay Makowsky	Heinrich-Hertz-Oberschule, Berlin
Mallku Schlagowski	Heinrich-Hertz-Oberschule, Berlin
Trung Duc Nguyen	Andreas-Gymnasium, Berlin
Alexander Reinecke	Schüler-Praktikant, Humboldt-Universität zu Berlin

Gruppenleiter:

Prof. Dr. Barbara Grabowski HTW Saar



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



Grundlagen

Definition 1. Seien $P = (p_1, ..., p_n) \in \mathbb{R}^n$ und $Q = (q_1, ..., q_n) \in \mathbb{R}^n$ zwei Punkte im $\mathbb{R}^n$. Dann heißt $\vec{a} = (a_1, \cdots, a_n)^T$, mit $a_i = q_i - p_i$ für $i = 1, ..., n$, Vektor von P nach Q.

1. $|\vec{a}| = \sqrt{\sum_{i=1}^{n} a_i^2}$ heißt Betrag des Vektors $\vec{a}$ und ist gleich seiner Länge.
2. Addition, Subtraktion zweier Vektoren $\vec{a}, \vec{b} \in \mathbb{R}^n$ und Multiplikation eines Vektor mit einer reellen Zahl $\lambda \in \mathbb{R}$ erfolgen immer komponentenweise.
3. Das Produkt $\vec{a} \cdot \vec{b} = \sum_{i=1}^{n} a_i \cdot b_i$ heißt Skalarprodukt.

Satz 1. *1. Der Vektor $\frac{\vec{a}}{|\vec{a}|}$ hat stets die Länge 1, d.h. es ist $\left|\frac{\vec{a}}{|\vec{a}|}\right| = 1$.*

2. Es gilt: $\vec{a} \cdot \vec{b} = |\vec{a}| \cdot |\vec{b}| \cdot cos(\gamma)$, wobei γ der kleinere der beiden von $\vec{a}$ und $\vec{b}$ eingeschlossenen Winkel ist.

3. $\vec{a}$ und $\vec{b}$ stehen senkrecht aufeinander $\Leftrightarrow \vec{a} \cdot \vec{b} = 0$.

Definition 2. Eine (stetige) vektorwertige Abbildung
$\vec{x} : t \in [t_1, t_2] \subseteq \mathbb{R} \to \vec{x}(t) \in \mathbb{R}^n$
von einem Intervall $t \in [t_1, t_2]$ in den $\mathbb{R}^n$ heißt Weg.

Als Kurve C bezeichnet man das Bild eines Weges, $C : \vec{x}(t), t \in [t_1, t_2]$.

Beispiele.

1. Die Kurve $C : \vec{x}(t) = P_0 + t \cdot \vec{a}, t \in R$, mit $P_0, \vec{a} \in \mathbb{R}^3$, ist eine Gerade im $\mathbb{R}^3$ mit dem Aufpunkt P_0 und dem Richtungsvektor $\vec{a}$.

2. Die Kurve $C : \vec{x}(t) = \begin{pmatrix} R \cdot \cos(t) \\ R \cdot \sin(t) \end{pmatrix}, t \in [0, 2 \cdot \pi]$, beschreibt einen Kreis im $\mathbb{R}^2$ mit dem Mittelpunkt (0,0) und dem Radius R.

Reellwertige Funktion in mehreren Variablen

Definition 3. Eine Abbildung

$$f : P = (x_1, x_2, ..., x_n) \in D \subseteq \mathbb{R}^n \longmapsto y = f(P) \in \mathbb{R}$$

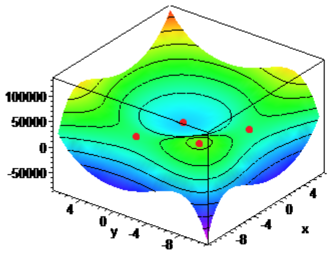
heißt reellwertige Funktion in n Variablen bzw. n Veränderlichen.

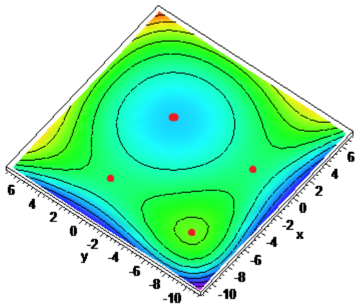
Beispiele. 1. In der Physik werden diverse skalare Größen, z.B. Temperatur, Druck usw. in Abhängigkeit vom Ort und der Zeit angegeben. Somit ergeben sich mit den drei Raumrichtungen und der Zeit Funktionen in vier Variablen.
2. In der Wirtschaft lässt sich der Gewinn eines Unternehmens aus seinen Einnahmen und Ausgaben ermitteln, wodurch eine Funktion mit sehr vielen Variablen entstehen kann.

Höhen- und Schnittlinien

Setzen wir eine Variable x_i konstant, so ergibt sich eine Funktionsschar von **Schnittkurven** $f_c(P) = f(x_1, .., c, .., x_n)$.

Eine **Höhenlinie** ist die Menge aller Punkte P, die denselben Funktionswert haben, d.h. für die gilt: $f(P) = c$.





Beispiele.
1. Bei Rotationsflächen $f : P \in \mathbb{R}^2 \longmapsto \mathbb{R}$ mit $f(x_1, x_2) = g(\sqrt{x_1^2 + x_2^2})$ sind die Höhenlinien Kreise, bei Ebenen $f(x_1, x_2) = a \cdot x_1 + b \cdot x_2 + c$ sind die Höhenlinien Geraden.
2. In der Meteorologie werden Luftdruck- und Temperaturbereiche durch sogenannte Isobaren bzw. Isothermen, also Zonen gleichen Drucks respektive gleicher Temperatur, veranschaulicht. Diese sind nichts anderes als die Höhenlinien der Funktion, die dem Ort den Luftdruck bzw. die Temperatur zuordnet.

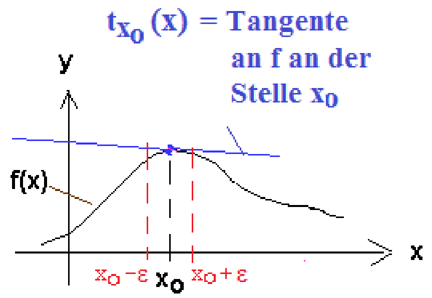
Richtungsableitung und partielle Ableitung

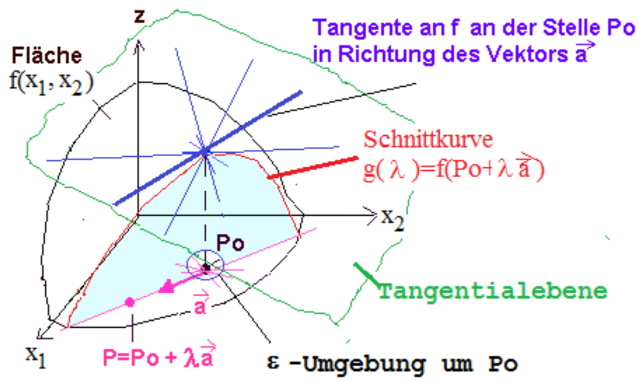
Wir betrachten eine Funktion $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$. Möchte man an der Funktionsfläche über einem Punkt P_0 die Tangente anlegen, so fällt auf, dass es in diesem Punkt unendlich viele Tangenten gibt. Sie liegen alle in einer Ebene. Legt man allerdings eine Richtung $\vec{a}$ auf der $x_1 - x_2$-Ebene fest, so gibt es genau eine Tangente in P_0, und zwar die Tangente an die Schnittkurve $g(\lambda) = f(P_0 + \lambda\vec{a})$ von f über der Geraden $P = P_0 + \lambda \cdot \vec{a}$. Der Anstieg dieser Tangenten wird als Richtungsableitung bezeichnet. Dieses Prinzip ist auch auf Funktionen mit mehr als zwei Veränderlichen anwendbar.

Definition 4. Seien $y = f(x_1, \ldots, x_n)$, $P_0 = (x_1^0, \ldots, x_n^0)$, $\vec{a} = (a_1, \ldots, a_n)^T$ ein Vektor der Länge 1 und $g(\lambda) = f(P_0 + \lambda \cdot \vec{a})$. Der Grenzwert

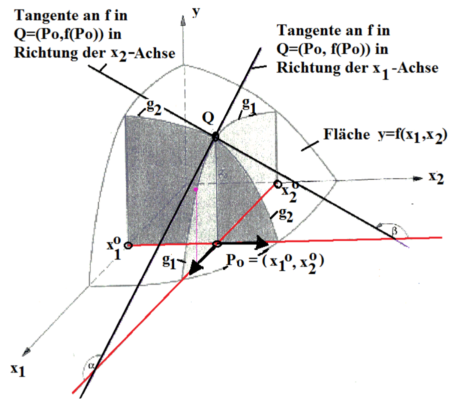
$$\frac{\partial f}{\partial \vec{a}}(P_o) = \lim_{\lambda \to 0} \frac{f(P_0 + \lambda \vec{a}) - f(P_0)}{\lambda} = \lim_{\lambda \to 0} \frac{g(\lambda) - g(0)}{\lambda} = g'(0)$$

heißt **Richtungsableitung von f im Punkt P_0 in Richtung $\vec{a}$.**





Leitet man eine reellwertige Funktion f mehrerer Veränderlicher in Richtung einer Koordinatenachse ab, so spricht man von einer partiellen Ableitung von f.



Definition 5. Seien $y = f(x_1, \ldots, x_n)$ und $P_0 = (x_1^0, \ldots, x_n^0)$ gegeben, so heißt der Grenzwert

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

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

Der Gradient und seine Eigenschaften

Für eine Funktion $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ sind damit genau n partielle Ableitungen definiert. Diese können in einem Vektor zusammengefasst werden.

Definition 6. Der Vektor

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

heißt **Gradient von f im Punkt P_0.**

Aus der Schule sind Ableitungsregeln für Funktionen in einer Veränderlichen bekannt, wie z.B. die Summenregel oder die Kettenregel. Analoge Regeln gelten für die Richtungsableitung einer Funktion in mehreren Veränderlichen. Im Folgenden benötigen wir vor allen Dingen die **verallgemeinerte Kettenregel**.

Satz 2. *Seien $f : \vec{x} \in D_f \subseteq \mathbb{R}^n \longrightarrow f(\vec{x}) \in \mathbb{R}$ und $\vec{x} : t \in D_{\vec{x}} \subseteq \mathbb{R}^n$. Sei t_0 ein Punkt in $\mathbb{R}$.*
Innerhalb einer Umgebung von $\vec{x}(t_o)$ seien die Funktion $f(\vec{x})$ und ihre partiellen Ableitungen $\dfrac{\partial f}{\partial x_1}(\vec{x}), \ldots, \dfrac{\partial f}{\partial x_n}(\vec{x})$ definiert und stetig und innerhalb einer Umgebung von t_0 sei $\vec{x}(t)$ definiert und dort differenzierbar, d.h. die einzelnen Komponenten $x_1(t), \ldots, x_n(t)$ der Funktion $\vec{x}(t)$ seien differenzierbar.
Dann ist die zusammengesetzte Funktion $g(t) = f(\vec{x}(t))$ bei t_0 differenzierbar und es gilt an dieser Stelle:

$$\frac{dg}{dt}(t_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{x}(t_0))\frac{dx_i}{dt}(t_0) = gradf(\vec{x}(t_0)) \cdot \frac{d\vec{x}}{dt}(t_0). \tag{1}$$

Aus Formel (1) ergibt sich mit $t = \lambda$, $t_0 = \lambda_0 = 0$, $\vec{x}(\lambda) = P_0 + \lambda \cdot \vec{a}$ und $g(\lambda) = f(\vec{x}(\lambda))$ sofort folgende hilfreiche Eigenschaft 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 seinen innerhalb einer Umgebung von P_0 die Funktion $f(\vec{x})$ und ihre partiellen Ableitungen $\dfrac{\partial f(\vec{x})}{\partial x_1}, \ldots, \dfrac{\partial f(\vec{x})}{\partial x_n}$ definiert und stetig. Sei $\vec{a}$ ein Vektor der Länge 1. Dann gilt:*

$$\frac{\partial f}{\partial \vec{a}}(P_0) = \vec{a} \cdot gradf(P_0). \tag{2}$$

Satz 4. *Sind die Voraussetzungen von Satz 3 erfüllt und $gradf(P_0) \neq \vec{0}$, dann zeigt der Vektor $gradf(P_0)$ in Richtung des steilsten Anstieges von f, vom Punkt P_0 aus betrachtet.*

Beweis: Sei $\vec{b} \in \mathbb{R}^n$ ein Vektor der Länge 1, der in die gleiche Richtung wie der Gradient $\text{grad}f(P_0)$ zeigt. Dann ist der Winkel γ zwischen den Gradienten und $\vec{b}$ gleich 0. Sei $\vec{a} \in \mathbb{R}^n$ nun ein beliebiger Vektor der Länge 1. Dann erhalten wir unter Verwendung von Satz 1

$$
\begin{aligned}
\frac{\partial f}{\partial \vec{a}}(P_0) &= \vec{a} \cdot \text{grad}f(P_0) = |\vec{a}| \cdot |\text{grad}f(P_0)| \cdot cos(\gamma) \\
&\leq |\vec{a}| \cdot |\text{grad}f(P_0)| \cdot cos(0) = 1 \cdot |\text{grad}f(P_0)| \cdot cos(0) \\
&= |\vec{b}| \cdot |\text{grad}f(P_0)| \cdot cos(0) = \vec{b} \cdot \text{grad}f(P_0) = \frac{\partial f}{\partial \vec{b}}(P_0).
\end{aligned}
$$

$\square$

Satz 5. *Sei $f : D_f \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ 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 $P_0 \in C$. Sei M_c durch die Kurve $\vec{x}_c(t), t_1 \leq t \leq t_2$ beschreibbar, welche in einer Umgebung des Punktes P_0 differenzierbar ist und es gelte $\vec{x}_c(t_0) = P_0$. Dann gilt:*

$$
gradf(P_0) \perp \frac{d\vec{x}}{dt}(t_0). \tag{3}
$$

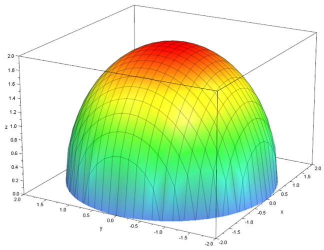
Formel (3) besagt, dass der Gradient von f in P_0 senkrecht auf dem Tangentenvektor in P_0 an die c-Höhenlinie, zu der P_0 gehört, steht.

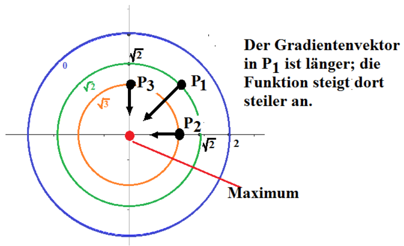
Beweis: Nach Kettenregel (1) ist $\frac{d(f \circ \vec{x_c})}{dt}(t_0) = \text{grad}f(P_0) \cdot \frac{d\vec{x_c}}{dt_0}(t_0)$. Aus $f(\vec{x_c}(t)) = c$ folgt $\frac{d(f \circ \vec{x_c})}{dt}(t_0) = 0$ und damit folgt die Behauptung sofort aus Satz 1. $\square$

Beispiel 1. Die folgende Abbildung zeigt die Funktion f definiert durch

$$f(x_1, x_2) = \sqrt{4 - x_1^2 - x_2^2},$$

ihre Höhenlinien $f(x_1, x_2) = c$ für $c = 0, c = \sqrt{2}, c = \sqrt{3}$ und die Gradienten von f in den Punkten $P_1 = (1, 1)$, $P_2 = (1, 0), P_3 = (0, 1)$.





Extremwertberechnung

Im folgenden Abschnitt betrachten wir nur Funktionen in zwei Variablen $f : (x_1, x_2) \in \mathbb{R}^2 \to \mathbb{R}$.

Definition 7. Ein Punkt $P_0 \in \mathbb{R}^n$ heißt (strenges) lokales Extremum von f, wenn alle Punkte $P \neq P_0$ in der Umgebung $U_\varepsilon(P_0)$ von P_0 einen größeren bzw. kleineren Funktionswert besitzen als P_0. P_0 heißt
lokales Maximum von f, falls gilt: $f(P_0) > f(P)$ für alle $P \in U_\varepsilon(P_0)$,
lokales Minimum von f, falls gilt: $f(P_0) < f(P)$ für alle $P \in U_\varepsilon(P_0)$.

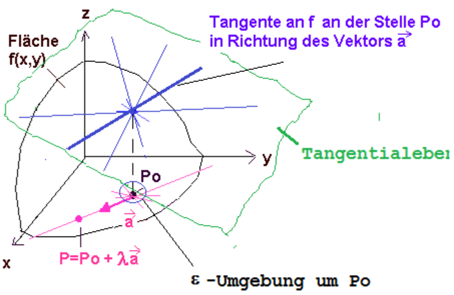
Wie bereits aus der Schule bekannt, lautet die Gleichung der Tangente an eine Funktion f an der Stelle x_0:

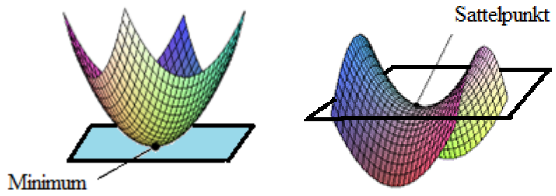
$$t_{x_0}(x) = f(x_0) + f'(x_0) \cdot (x - x_0) \tag{4}$$

Auffällig ist hier, dass sie vom Anstieg $f'(x_0)$ abhängt. Schwieriger wird es im $\mathbb{R}^2$, weil es hier in einem Punkt P_0 unendlich viele Tangenten gibt. Glücklicherweise liegen diese auf einer Ebene, welche als Tangentialebene bezeichnet wird. Eine Ebene kann man mit einem Punkt und zwei Richtungen beschreiben. Wir haben die folgende **Ebenengleichung der Tangentialebene** hergeleitet:

$$T_{P_0}(P) \;=\; f(P_0) + (\mathrm{grad} f(P_0)) \cdot \overrightarrow{P_0 P} \tag{5}$$

Man sieht die Analogie zu (4): Anstelle der Ableitung $f'(x_0)$ steht nun der Gradient $\mathrm{grad} f(P_0)$.





Wir sehen

1. Wenn f in P_0 ein lokales Minimum (oder lokales Maximum) besitzt, so ist $\mathrm{grad} f(P_0) = 0$.
2. Wenn aber $\mathrm{grad} f(P_0) = 0$ ist, so muss P_0 nicht unbedingt ein lokaler Extremwert sein. P_0 kann auch ein Sattelpunkt sein.

Um zu erkennen, ob f in P_0 ein lokales Extremum besitzt, benötigen wir zusätzlich die partiellen Ableitungen zweiter Ordnung von f in P_0. f besitzt in einem Punkt P zwei partielle Ableitungen 1. Ordnung und folglich vier partielle Ableitungen 2. Ordnung. Diese fassen wir zur sogenannten Hessematrix zusammen.

Definition 8. Die Matrix der 2. partiellen Ableitungen von f in einem Punkt $P = (x_1, x_2)$

$$H(P) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(P) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(P) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(P) & \frac{\partial^2 f}{\partial x_2^2}(P) \end{pmatrix}$$

heißt **Hessematrix** von f in P.

$$\mathrm{Det}(H(P)) = \frac{\partial^2 f}{\partial x_1^2}(P) \cdot \frac{\partial^2 f}{\partial x_2^2}(P) - \frac{\partial^2 f}{\partial x_1 \partial x_2}(P) \cdot \frac{\partial^2 f}{\partial x_2 \partial x_1}(P)$$

heißt **Determinante** der Matrix $H(P)$.

Die Grundlage für ein hinreichendes Extremwertkriterium bildet der mehrdimensionale Satz von Taylor.

Satz 6. *Mehrdimensionaler Satz von Taylor*

Sei $f : (x_1, x_2) \in U_\varepsilon(P_0) \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ auf $U_\varepsilon(P_0)$ in beiden Variablen x_1, x_2, (k+1)-mal stetig partiell differenzierbar. Dann gilt für alle $P \in U_\varepsilon(P_0)$:

$$f(P)=f(P_0)+gradf(P_0)\cdot\overrightarrow{P_0P}+\frac{1}{2}(\overrightarrow{P_0P})^T H(P_0)(\overrightarrow{P_0P})+o(|\overrightarrow{P_0P}|^2) \tag{6}$$

wobei $o(|\overrightarrow{P_0P}|^2)$ *eine Größe ist mit der Eigenschaft*

$$\lim_{|\overrightarrow{P_0P}|\to 0} \frac{o(|\overrightarrow{P_0P}|^2)}{|\overrightarrow{P_0P}|^2} = 0\ . \tag{7}$$

Mit Hilfe dieses Satzes 6 haben wir uns Kriterien für das Vorliegen von lokalen Extremwerten wie folgt verdeutlicht.

Für ein lokales Extremmum P_0 gilt $\mathrm{grad} f(P_0) = \vec{0}$. Weiterhin folgt aus Formel (7), dass der Fehlerterm $o(|\overrightarrow{P_0P}|^2)$ für alle Punkte P in einer kleinen Umgebung $U_\epsilon(P_0)$ von P_0 vernachlässigbar ist. Damit wird die Formel (6) für alle $P \in U_\epsilon(P_0)$ zu:

$$f(P) = f(P_0) + \tfrac{1}{2} \cdot (\overrightarrow{P_0P})^T \cdot H(P_0) \cdot (\overrightarrow{P_0P}). \tag{8}$$

Daraus folgt, wenn $(\overrightarrow{P_0P})^T \cdot H(P_0) \cdot (\overrightarrow{P_0P}) > 0$ für alle $P \in U_\epsilon(P_0)$ ist, so gilt für diese P $f(P_0) < f(P)$ und folglich ist P_0 ein lokales Minimum. Wenn für alle $P \in U_\epsilon(P_0)$ gilt $(\overrightarrow{P_0P})^T \cdot H(P_0) \cdot (\overrightarrow{P_0P}) < 0$, so folgt für diese P $f(P_0) > f(P)$ und folglich ist P_0 ein lokales Maximum.

Um daraus ein praktikables Extremwertkriterium zu erhalten, haben wir nachfolgenden Satz verwendet.

Satz 7. *Wenn* $Det(H(P_0)) > 0$ *und* $\frac{\partial^2 f}{\partial x_1^2}(P_0) > 0$ *so gilt*
$(\overrightarrow{P_0P})^T \cdot H(P_0) \cdot (\overrightarrow{P_0P}) > 0$ *für alle* $P \in U_\epsilon(P_0)$.
Wenn $Det(H(P_0)) > 0$ *und* $\frac{\partial^2 f}{\partial x_1^2}(P_0) < 0$ *so gilt*
$(\overrightarrow{P_0P})^T \cdot H(P_0) \cdot (\overrightarrow{P_0P}) > 0$ *für alle* $P \in U_\epsilon(P_0)$.

Daraus ergibt sich das folgende hinreichende Extremwertkriterium für Funktionen in zwei Veränderlichen.

Satz 8. *Hinreichendes Kriterium für lokale Extremwerte*

Wenn $\text{grad}f(P_0) = \vec{0}$ *und* $Det(H(P_0)) > 0$, *so ist* P_0 *ein lokales Extremum von* f. *Ist* $\dfrac{\partial^2 f}{\partial x_1^2}(P_0) < 0$, *so ist* P_0 *lokales Maximum; ist* $\dfrac{\partial^2 f}{\partial x_1^2}(P_0) > 0$, *so ist* P_0 *lokales Minimum.*

Beispiel 2. Sei $f(P) = f(x,y) = x \cdot e^{-(x^2+y^2)}$, wobei $-2 \le x; y \le 2$.

1. Bildung der 1. partiellen Ableitungen:
   $\frac{\partial f}{\partial x}(P) = e^{-(x^2+y^2)} + x \cdot (-2x) \cdot e^{-(x^2+y^2)} = e^{-(x^2+y^2)} \cdot (1 - 2x^2)$,
   $\frac{\partial f}{\partial y}(P) = -2xy \cdot e^{-(x^2+y^2)}$.
2. Die 1.partielle Ableitungen auf Nullstellen überprüfen und so die extremwertverdächtigen Punkte finden:
   $\frac{\partial f}{\partial x}(P) = 0 \Leftrightarrow x = \pm\sqrt{\frac{1}{2}}$ und $\frac{\partial f}{\partial y}(P) = 0 \Leftrightarrow y = 0$ oder $x = 0$.
   Die extremwertverdächtigen Punkte sind also
   $P_1 = (\sqrt{\frac{1}{2}}; 0)$ und $P_2 = (-\sqrt{\frac{1}{2}}; 0)$.
3. Hesssematrix in P_1 und P_2 untersuchen, um zu erkennen, ob es sich wirklich um lokale Minima oder lokale Maxima handelt. Wir erhalten
   $$H(P_1) = \begin{pmatrix} -\frac{4}{\sqrt{2}} \cdot e^{-1/2} & 0 \\ 0 & -\frac{2}{\sqrt{2}} \cdot e^{-1/2} \end{pmatrix}$$
   und folglich ist P_1 wegen $\text{Det}(H(P_1)) = \frac{4}{e} > 0$ und $\frac{\partial^2 f}{\partial x_1^2}(P_1) < 0$ ein lokales Maximum. Weiterhin gilt
   $H(P_2) = \begin{pmatrix} \frac{4}{\sqrt{2}} \cdot e^{-1/2} & 0 \\ 0 & \frac{2}{\sqrt{2}} \cdot e^{-1/2} \end{pmatrix}$ und folglich ist P_2
   wegen $\text{Det}(H(P_2)) = \frac{4}{e} > 0$ und $\frac{\partial^2 f}{\partial x_1^2}(P_2) > 0$ ein lokales Minimum.