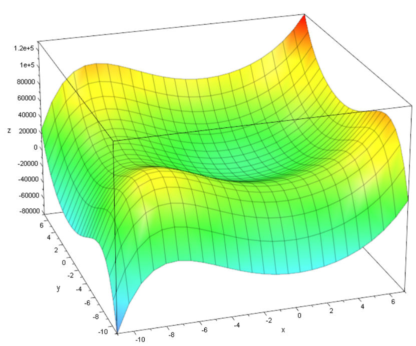
Reellwertige Funktionen mehrerer Veränderlicher

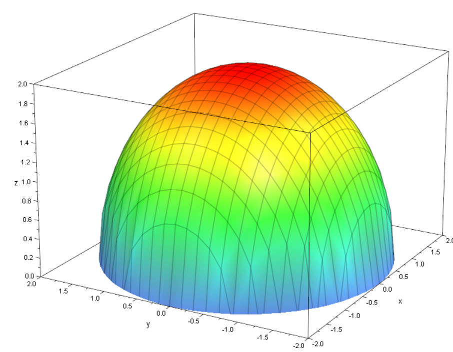
Teilnehmer:

Philipp Besel	Heinrich-Hertz-Oberschule, Berlin
Joschka Braun	Ludwigs-Georgs-Gymnasium, Darmstadt
Robert Courant	Heinrich-Hertz-Oberschule, Berlin
Florens Greßner	Immanuel-Kant-Oberschule, Berlin
Tim Jaschek	Immanuel-Kant-Oberschule, Berlin
Leroy Odunlami	Herder-Oberschule, Berlin
Gloria Xiao	Heinrich-Hertz-Oberschule, Berlin

Gruppenleiter:

Prof. Dr. Barbara Grabowski HTW Saar





Unsere Arbeitsgruppe hat sich mit reellwertigen Funktionen mehrerer Veränderlicher beschäftigt. Diese sind in der Mathematik von großer Bedeutung, da sich viele praktische 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 stark sich eine geringfügige Änderung der Argumente der Funktion auf den Funktionswert auswirkt und bei welchen Argumentwerten die Funktionswerte gleich sind. Zunächst soll geklärt werden, worum es sich bei reellwertigen Funktionen mehrerer Veränderlicher handelt.

1 Grundlagen

Definition 1.1. 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 Veränderlichen.

Beispiel. In der Physik werden diverse Größen, wie z.B. die Temperatur, in Abhängigkeit vom Ort und der Zeit angegeben. Somit ist mit den drei Raumkoordinaten und der Zeit die Temperatur eine Funktion in vier Veränderlichen.
In der Wirtschaft lässt sich der Gewinn eines Unternehmens als Funktion seiner Einnahmen und Ausgaben darstellen, wodurch eine Funktion mit sehr vielen Veränderlichen entstehen kann.

1.1 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: $f(P) = c$.

Beispiel. 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.

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 Stetigkeit

Sei $P = (x_1, \ldots, x_n) \in \mathbb{R}^n$ und $\|P\| = \sqrt{\sum_{i=1}^n x_i^2}$ die **Euklidische Norm** im $\mathbb{R}^n$. Dann ist durch $U_\varepsilon(P_0) = \{P \in \mathbb{R}^n | \|P - P_0\| < \varepsilon\}$ eine **ε-Umgebung** von $P_0 \in \mathbb{R}^n$ definiert.

Definition 2.1. Eine Funktion $f : D \subseteq \mathbb{R}^n \longmapsto \mathbb{R}$ heißt **stetig in einem Punkt** $P_0 \in D$, wenn gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall P \in D : \|P - P_0\| < \delta \Rightarrow |f(P) - f(P_0)| < \varepsilon.$$

Anschaulich bedeutet diese Definition, dass für jede beliebig kleine ε-Umgebung um den Wert $f(P_0)$ eine δ-Umgebung um P_0 existiert, so dass das Bild der δ-Umgebung von P_0 in der ε-Umgebung von $f(P_0)$ liegt.

3 Ableitungen von Funktionen in n Veränderlichen

Für Funktionen in einer Veränderlichen sind lokale Änderungsraten von großer Bedeutung. Diese werden durch Ableitungen beschrieben. Kennt man diese, so kann man beispielsweise auf Extrem- und Wendestellen schließen. Nun stellt sich die Frage, ob analog zur Ableitung einer Funktion in einer Veränderlichen auch eine Ableitung von Funktionen mehrerer Veränderlicher definiert werden kann.

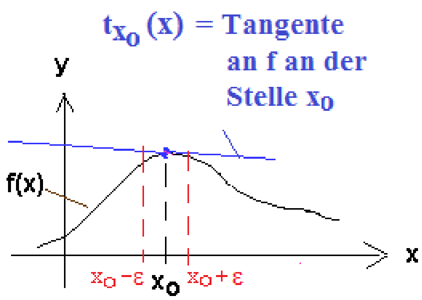
3.1 Richtungsableitung

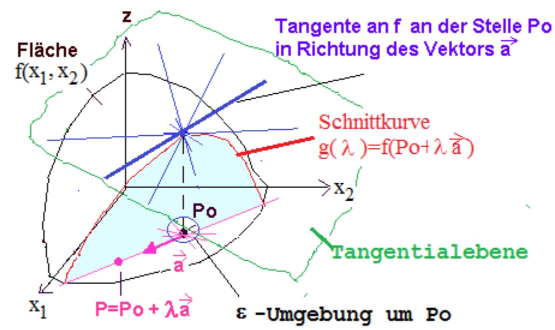
Wir betrachten eine Funktion $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$. Möchte man an der Funktionsfläche über einem Punkt P_0 die Tangente anlegen, so fällt auf, dass es in diesem Punkt unendlich viele Tangenten gibt. Sie liegen alle in einer Ebene. Legt man allerdings eine Richtung $\vec{a}$ auf 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 3.1. Seien $y = f(x_1, \ldots, x_n), P_0 = (x_1^0, \ldots, x_n^0), \vec{a} = (a_1, \ldots, a_n)$ 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 $\vec{a}$.**





Beispiel. Die Funktion $f(x_1,x_2) = \sqrt{R^2 - x_1^2 - x_2^2}$ soll im Punkt $P_0 = (x_1^0, x_2^0)$ in Richtung $\vec{a} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$ abgeleitet werden. Dazu leiten wir unter Verwendung der bekannten Ableitungsregeln die Funktion

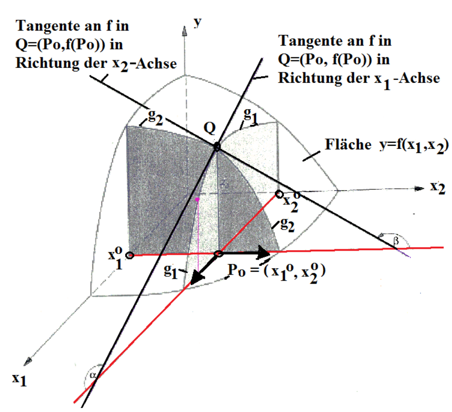
$$g(\lambda) = f(P_0 + \lambda\vec{a}) = \sqrt{R^2 - (x_1^0 + \lambda/\sqrt{2})^2 - (x_2^0 + \lambda/\sqrt{2})^2}$$

nach λ ab und erhalten für die gesuchte Richtungsableitung:

$$\frac{\partial f}{\partial \vec{a}}(P_0) = g'(0) = \frac{-x_1^0 - x_2^0}{\sqrt{2}\sqrt{R^2 - (x_1^0)^2 - (x_2^0)^2}}.$$

3.2 Partielle Ableitungen und Gradient

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



Definition 3.2. 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.

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 3.3. 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.

3.3 Ableitungsregeln

3.3.1 Verknüpfung von Funktionen

Satz 3.1. *Seien a, b zwei reelle Zahlen und seien $f : D_f \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ und $h : D_h \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ reellwertige Funktionen. Sei $P_0 \in D_f \cap D_h$ und f und h in P_0 differenzierbar in Richtung $\vec{a} \in \mathbb{R}^n$. Dann gilt für die Richtungsableitung der Funktion $F : P \in D_f \cap D_h \longrightarrow a \cdot f(P) + b \cdot h(P)$:*

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

3.3.2 Die verallgemeinerte Kettenregel

Satz 3.2. *Seien $P : t \in T \subseteq \mathbb{R} \longrightarrow \mathbb{R}^n$ eine durch t parametrisierte Kurve in $\mathbb{R}^{\ltimes}$ und $f : P \in D_f \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ eine Funktion auf dieser Kurve. Seien weiterhin $t_0 \in T$ und $P_0 = P(t_0) \in D_f$.*
Innerhalb einer Umgebung von $P_0 = P(t_0)$ seien die Funktion $f(P)$ und ihre partiellen Ableitungen $\dfrac{\partial f}{\partial x_1}(P), \ldots, \dfrac{\partial f}{\partial x_n}(P)$ definiert und stetig und innerhalb einer Umgebung von t_0 sei $P(t)$ definiert und dort differenzierbar, d.h.,

die einzelnen Komponenten $x_1(t), \ldots, x_n(t)$ der Funktion $P(t)$ seien dort differenzierbar.
Dann ist die zusammengesetzte Funktion $g(t) = f(P(t))$ bei t_0 differenzierbar und es gilt an dieser Stelle:

$$\frac{dg}{dt}(t_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(P_0)\frac{dx_i}{dt}(t_0) = gradf(P_0) \cdot \frac{dP}{dt}(t_0) \tag{3.1}$$

3.4 Eigenschaften des Gradienten

Aus (3.1) ergibt sich mit $\lambda = t$ und $P(\lambda) = P_0 + \lambda \cdot \vec{a}$ sofort folgender Satz:

Satz 3.3. *Seien $\vec{a} \in \mathbb{R}^n$ ein Vektor der Länge 1, $f : D_f \subseteq \mathbb{R}^n \to \mathbb{R}$ und seien innerhalb einer Umgebung eines Punktes $P_0 \in D_f$ die Funktion $f(P)$ und ihre partiellen Ableitungen $\frac{\partial f}{\partial x_1}(P), \ldots, \frac{\partial f}{\partial x_n}(P)$ definiert und stetig. Dann gilt:*

$$\frac{\partial f}{\partial \vec{a}}(P_0) = \vec{a} \cdot grad\ f(P_0) \tag{3.2}$$

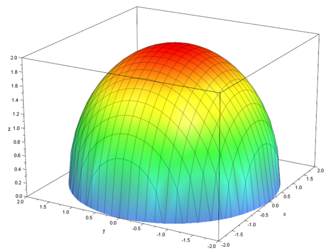
Mit Hilfe dieses Satzes und der Kettenregel haben wir zwei weitere sehr anschauliche und hilfreiche Eigenschaften des Gradienten bewiesen.

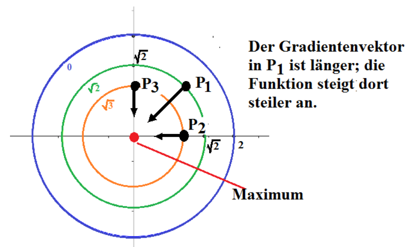
Satz 3.4. *Sind die Voraussetzungen von Satz 3.3 erfüllt und $gradf(P_0) \neq \vec{0}$, dann zeigt der Vektor $gradf(P_0)$ in Richtung des steilsten Anstieges von f, vom Punkt P_0 aus betrachtet.*

Satz 3.5. *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 $P_0 \in M_c$. Sei M_c durch die Kurve $P_c(t), t_1 \leq t \leq t_2$ beschreibbar, welche in einer Umgebung des Punktes P_0 differenzierbar ist, und es gelte $P_c(t_0) = P_0$. Dann gilt:*

$$gradf(P_0) \perp \frac{dP}{dt}(t_0)$$

Beispiel. Die folgende Abbildung zeigt die Funktion $f(x_1, x_2) = \sqrt{4 - x_1^2 - x_2^2}$, 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 $P_1 = (1, 1)$, $P_2 = (1, 0), P_3 = (0, 1)$.





4 Tangentialebene

Wie bereits aus der Schule bekannt, kann man Funktionswerte $f(x)$ für $x \in (x_0-\epsilon, x_0+\epsilon)$ mit Hilfe einer Tangente an f an der Stelle x_0 linear annähern. Die bekannte Gleichung dieser Tangente lautet:

$$t_{x_0}(x) = f(x_0) + f'(x_0) \cdot (x - x_0) \tag{4.1}$$

Auffällig ist hier, dass sie von dem 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:

$$\begin{aligned} T_{P_0}(P) &= f(P_0) + \frac{\partial f}{\partial x_1} \cdot (x_1 - x_1^0) + \cdots + \frac{\partial f}{\partial x_n}(P_0) \cdot (x_n - x_n^0) \\ &= f(P_0) + (\mathrm{grad} f(P_0)) \cdot \overrightarrow{P_0 P} \end{aligned}$$

Man sieht die Analogie zu (4.1): Anstelle der Ableitung $f'(x_0)$ steht nun der Gradient *grad* $f(P_0)$.

Die beste lineare Approximation von $f(P)$ für alle Punkte P in einer kleinen Umgebung des Punktes P_o erhalten wir dann durch diese Tangentialebene:

$$f(P) \approx T_{P_0}(P) = f(P_0) + (\mathrm{grad} f(P_0)) \cdot \overrightarrow{P_0 P} \tag{4.2}$$

Beispiel. Mit (4.2) haben wir für verschiedene Funktionen untersucht, wie sich der Funktionswert einer Funktion in einer Umgebung eines Punktes P_0 verhält. Aus Platzgründen müssen wir hier auf weitere Erläuterungen verzichten.

5 Der Satz von Taylor

5.1 Taylor-Polynome für Funktionen einer Veränderlichen

Um das Verhalten von Funktionswerten $f(x)$ in einer Umgebung $(x_0-\varepsilon, x_0+\varepsilon)$ einer Stelle x_0 abschätzen zu können, kann man die Funktion f durch eine andere Funktion annähern. Wie könnte eine solche Funktion aussehen?

Der einfachste Fall ist die Näherung $f(x) \approx f(x_0)$. Je weiter man sich mit x von der Stelle x_0 entfernt, desto größer wird allerdings der Fehler.

Die Approximation wird genauer, wenn man als Näherung die Tangente $t_{x_0}(x)$ in (4.1), d.h. $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$ wählt. Noch genauer kann die Funktion angenähert werden, indem ein quadratischer Term in $(x - x_0)$ und die 2. Ableitungen von $f(x_0)$ hinzugefügt werden. Die Verallgemeinerung dieser Erkenntnis ist:

Satz 5.1. *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:*

$$f(x) = T_{k,x_0}(x) + R_{k,x_0}(x)$$

*mit dem sogenannten **k-ten Taylor-Polynom an der Entwicklungsstelle** x_0:*

$$T_{k,x_0} = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0)^1 + \cdots + \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k = \sum_{i=0}^{k} \frac{f^{(i)}(x_0)}{i!}(x-x_0)^i$$

*und dem **Restglied** $R_{k,x_0}(x)$:*

$$R_{k,x_0}(x) = \frac{f^{(k+1)}(\xi)}{(k+1)!}(x - x_0)^{k+1}.$$

Lemma 5.2. *Unter der Voraussetzung von Satz 5.1 gilt für das Restglied:*

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

D.h., das Restglied 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.

Beweis. Da $f^{(k+1)}(x)$ für alle x $\in$ I stetig ist, ist $f^{(k+1)}(x)$ auch in x_0 stetig. Mit $x \to x_0$ strebt nun auch die Folge der zwischen x und x_0 liegenden ξ-Werte gegen x_0. Aus der Stetigkeit folgt deshalb:

$$\lim_{|x-x_0|\to 0} f^{(k+1)}(\xi) = f^{(k+1)}(x_0)$$

und wir erhalten:

$$\begin{aligned} |R_{k,x_0}(x)| &= \frac{|f^{(k+1)}(\xi)|}{|k+1|!}|x-x_0|^k|x-x_0| \\ \Rightarrow \lim_{|x-x_0|\to 0} \frac{|R_{k,x_0}(x)|}{|x-x_0|^k} &= \frac{1}{|k+1|!} \lim_{|x-x_0|^k} |f^{(k+1)}(\xi)||x-x_0| \\ &= \frac{1}{|k+1|!} \cdot |f^{(k+1)}(x_0)| \lim_{|x-x_0|\to 0} |x-x_0| = 0. \end{aligned}$$

$\square$

5.2 Taylor-Polynome für Funktionen mehrerer Veränderlicher

Auch im $\mathbb{R}^n$ kann eine Funktion mittels des Verfahrens von Taylor abgeschätzt werden.

Satz 5.3. *Sei $f : U_\varepsilon(P_0) \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ auf $U_\varepsilon(P_0)$ in allen Variablen $x_i, i = 1, \ldots, n$, $(k+1)$-mal stetig partiell differenzierbar. Dann gilt für alle $P \in U_\varepsilon(P_0)$:*

(i) Für $k = 1$: $f(P) = f(P_0) + grad\ f(P_0) \cdot \overrightarrow{P_0P} + o(|| \overrightarrow{P_0P} ||)$

(ii) Für $k = 2$:

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

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

$$\lim_{||P-P_0||\to 0} \frac{o(|| \overrightarrow{P_0P} ||^k)}{|| \overrightarrow{P_0P} ||^k} = 0\ .$$

Bemerkung: $H(P_0)$ ist die Matrix der 2. partiellen Ableitungen von f im Punkt P_0. Diese wird im folgenden Abschnitt für $n = 2$ näher erläutert.

6 Extremwertbestimmung

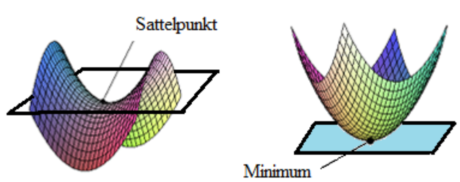
Definition 6.1. Ein Punkt $P_0 \in \mathbb{R}^n$ heißt (strenges) lokales Extremum von f, wenn alle Punkte $P \neq P_0$ in der Umgebung von P_0 einen größeren bzw. kleineren Funktionswert besitzen als P_0.
P_0 heißt lokales Maximum von f, falls gilt: $f(P_0) > f(P) \quad \forall P \in U_\varepsilon(P_0)$,
P_0 heißt lokales Minimum von f, falls gilt: $f(P_0) < f(P) \quad \forall P \in U_\varepsilon(P_0)$.

Entsprechend zu $\mathbb{R}^1$ existieren in $\mathbb{R}^n$ ein notwendiges und ein hinreichendes Kriterium zur Bestimmung von Extremalstellen.

Notwendiges Kriterium:

Wenn P_0 ein lokales Extremum ist, dann ist grad $f(P_0) = \vec{0}$.

Das hinreichende Kriterium haben wir nur für den Fall $n = 2$ untersucht.



Wie man an der Grafik sieht, erfüllen auch Sattelpunkte das notwendige Kriterium. Um zu erkennen, ob es sich um einen Extremwert handelt, benötigen wir die 2. partiellen Ableitungen von f in P_0.

Die Hesse-Matrix:
f besitzt in einem Punkt P zwei partielle Ableitungen 1. Ordnung und folglich vier partielle Ableitungen 2. Ordnung. Diese lassen sich vorteilhaft in der Matrixschreibweise darstellen.

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

Hinreichendes Kriterium:

Wenn grad$f(P_0) = \vec{0}$ und $\mathrm{Det}(H(P_0)) > 0$, 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.

Mit Hilfe des mehrdimensionalen Satzes 5.3 von Taylor haben wir uns dieses Kriterium im $\mathbb{R}^2$ verdeutlicht.

7 Extremwertaufgaben

Beispiel. Wir suchen die Extremstellen und Sattelpunkte der Funktion $f(x, y) = x^5 + 1000x^2 + 1000y^2 + y^5$ mit $x, y \in \mathbb{R}$.

Wir berechnen den Gradienten von f und setzen ihn gleich 0:

$$\frac{\partial f}{\partial x}(x, y) = 5x^4 + 2000x \qquad\qquad \frac{\partial f}{\partial y}(x, y) = 5y^4 + 2000y$$

$$\begin{aligned} \tfrac{\partial f}{\partial x} = 0 \quad &\Leftrightarrow \quad 5x^4 + 2000x = 0 \quad \Leftrightarrow \quad x = 0 \ \vee \ x = -\sqrt[3]{400} \\ \tfrac{\partial f}{\partial y} = 0 \quad &\Leftrightarrow \quad 5y^4 + 2000y = 0 \quad \Leftrightarrow \quad y = 0 \ \vee \ y = -\sqrt[3]{400} \end{aligned}$$

Damit erhalten wir 4 extremwertverdächtige Punkte:

$P_1 = (0; 0),\ P_2 = (0; -\sqrt[3]{400}),\ P_3 = (-\sqrt[3]{400}; 0),\ P_4 = (-\sqrt[3]{400}, -\sqrt[3]{400}).$

Wir untersuchen nun die zweiten partiellen Ableitungen von $f(P)$ in diesen 4 Punkten. Die Hesse-Matrix $H(P)$ ist

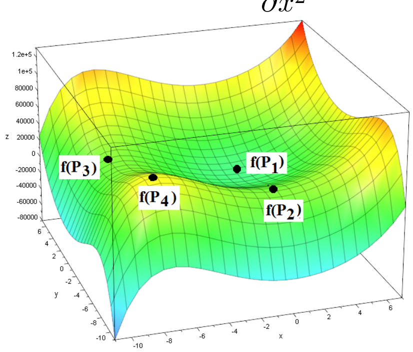
$$H(x, y) = \begin{pmatrix} 20x^3 + 2000 & 0 \\ 0 & 20y^3 + 2000 \end{pmatrix}$$

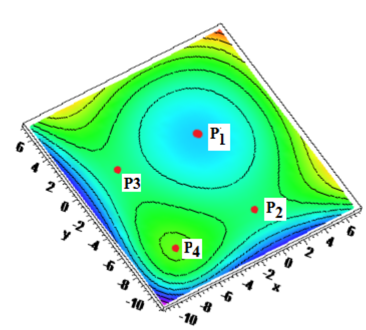
und wir erhalten mit $\mathrm{Det}(H(P)) = (20x^3 + 2000) \cdot (20y^3 + 2000)$ das Ergebnis:

$\mathrm{Det}(H(P_1)) > 0$ und $\dfrac{\partial^2 f}{\partial x^2}(P_1) > 0 \longrightarrow P_1$ ist ein lokales Minimum,

$\mathrm{Det}(H(P_2)) < 0$ und $\mathrm{Det}(H(P_3)) < 0 \longrightarrow P_2$ und P_3 sind Sattelpunkte,

$\mathrm{Det}(H(P_4)) > 0$ und $\dfrac{\partial^2 f}{\partial x^2}(P_4) < 0 \longrightarrow P_4$ ist ein lokales Maximum.





Beispiel. Die Funktion $f(x,y) = xy + 1$ ist auf $D = \{(x,y)|x^2 + y^2 \leq 1\}$ auf Extrema zu untersuchen. Denkbar ist, dass f die Wärmeverteilung auf einer runden Metallplatte mit Radius $r = 1$ beschreibt.
Wir berechnen die partiellen Ableitungen von f und setzen sie gleich 0, um die extremwertverdächtigen Punkte zu bestimmen.

$$\frac{\partial f}{\partial x}(x,y) = y = 0 \qquad \frac{\partial f}{\partial y}(x,y) = x = 0.$$

Die Hesse-Matrix des einzigen verdächtigen Punktes $P_0 = (x_0, y_0) = (0,0)$ ist:

$$H(P_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x_0,y_0) & \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0) \\ \frac{\partial^2 f}{\partial y \partial x}(x_0,y_0) & \frac{\partial^2 f}{\partial y^2}(x_0,y_0) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ mit } \mathrm{Det}(H(P_0)) = -1.$$

Daraus folgt, dass f im Inneren der Kreisscheibe nur einen Sattelpunkt in $P_0 = (0,0)$ besitzt. Da die Kreisscheibe D eine beschränkte und abgeschlossene Menge ist und f auf dieser Menge stetig ist, muss f nach dem Satz von Weierstraß ihr globales Maximum und ihr globales Minimum auf D annehmen. Diese Stellen müssen folglich auf dem Rand $x^2 + y^2 = 1$ von D liegen.

Durch Parametrisierung des Kreisrandes in Polarkoordinaten lassen sich die Funktionswerte auf dem Kreisbogen untersuchen:

$$x = cos(\varphi), y = sin(\varphi) \longrightarrow f(x,y) = cos(\varphi) \cdot sin(\varphi) + 1 =: k(\varphi).$$

Die Lösung der Gleichung $k'(\varphi) = \cos^2(\varphi) - \sin^2(\varphi) = 0$ ergibt die kritischen Winkel $\varphi_1 = \frac{1}{4}\pi,\ \varphi_2 = \frac{3}{4}\pi,\ \varphi_3 = \frac{5}{4}\pi,\ \varphi_4 = \frac{7}{4}\pi$. Durch Einsetzen dieser in x und y erhalten wir die kritischen Punkte von f:
$P_1 = (1/\sqrt{2}, 1/\sqrt{2})$, $\quad P_2 = (-1/\sqrt{2}, 1/\sqrt{2})$, $\quad P_3 = (-1/\sqrt{2}, -1/\sqrt{2})$,
$P_4 = (1/\sqrt{2}, -1/\sqrt{2})$.

Wegen $f(P_1) = f(P_3) = \frac{3}{2}$ sind P_1 und P_3 globale Maxima von f und wegen $f(P_2) = f(P_4) = \frac{1}{2}$ sind P_2 und P_4 globale Minimalwerte von f.