Extremwerte von Funktionen mehrerer reeller Variabler

Bei der Bestimmung der Extrema von (differenzierbaren) Funktionen $f : \mathbb{R}^n \to \mathbb{R}$ ist es sinnvoll, zuerst jene Stellen zu bestimmen, an denen überhaupt ein Extremum auftreten kann.

Im eindimensionalen Fall, also bei $f : \mathbb{R} \to \mathbb{R}$, war ja die entsprechende Bedingung die, dass $f'(x_0) = 0$.

Satz. Sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar auf einer offenen Umgebung $U(x^0)$ von $x^0 \in \mathbb{R}^n$.

Hat f an x^0 ein relatives (= lokales) Extremum, dann gilt notwendigerweise $\operatorname{grad} f(x^0) = \vec{0}$.

Beweis. f muß an x^0 in jeder Richtung $\vec{a}$ als Funktion des Geradenparameters ein relatives Extremum besitzen. Wegen des entsprechenden Kriteriums für Funktionen einer reellen Variablen bedeutet dies, dass an x^0 die entsprechende Richtungsableitung $\operatorname{grad} f(x^0) \cdot \vec{a}$ verschwindet. Weil $\vec{a}$ beliebig ist, muss damit $\operatorname{grad} f(x^0) = \vec{0}$ sein. $\Box$

Bemerkung. Mit der Bedingung $\operatorname{grad} f(x) = \vec{0}$ erhält man also jene Stellen, die für ein lokales Extremum in Frage kommen.

Zur Gewinnung eines hinreichenden Kriteriums für das Vorliegen eines Extremums approximieren wir f lokal durch das Taylor-Polynom zweiten Grades bzw. stellen f mittels des Satzes von Taylor durch das Taylor-Polynom $T_1(x, x^0)$ und dem entsprechenden Restglied dar.

Das Taylor-Polynom zweiten Grades hat mit $\vec{h} = \vec{x} - \vec{x}^0$ die Gestalt

$$T_2(x, x^0) = f(x^0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x^0) h_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} h_i \frac{\partial^2 f}{\partial x_i \partial x_j}(x^0) h_j$$

Definition. Die für eine Funktion $f \in C^2(U(x^0))$ (i.e. f sei zweimal stetig differenzierbar auf einer Umgebung $U(x^0)$ von x^0) definierte (symmetrische) Matrix der zweiten partiellen Ableitungen

$$\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix} = \frac{d^2 f}{dx^2} = f''(x)$$

heißt **Hesse-Matrix** $H(x)$ oder **zweite Ableitung** von f .

Beispiel. Sei $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = e^{xy}$.

Dann ist $f_x = ye^{xy}$, $f_y = xe^{xy}$ und weiters

$f_{xx} = y^2 e^{xy}$, $f_{xy} = f_{yx} = e^{xy} + xye^{xy}$ und $f_{yy} = x^2 e^{xy}$.

Somit ist $H(x, y) = \begin{pmatrix} y^2 e^{xy} & e^{xy} + xye^{xy} \\ e^{xy} + xye^{xy} & x^2 e^{xy} \end{pmatrix}$

Bemerkung. Die quadratische Approximation von f durch das Taylor-Polynom zweiten Grades

$$\widehat{f}(x) = f(x^0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x^0) h_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} h_i \frac{\partial^2 f}{\partial x_i \partial x_j}(x^0) h_j$$

heißt **Schmiegquadrik** . Sie hat an x^0 dieselbe Tangentialhyperebene wie f sowie dieselbe Krümmung (die im wesentlichen durch die Hesse-Matrix bestimmt ist).

Im Fall $n = 2$ stellt die Schmiegquadrik eine Fläche im $\mathbb{R}^3$ und ist entweder ein elliptisches Paraboloid, ein hyperbolisches Paraboloid oder ein parabolischer Zylinder (ersichtlich durch affine Transformation auf eine von drei sog. Normalformen) .

Das für das Vorliegen eines Extremums entscheidende Glied im Taylor-Polynom $T_2(x, x^0)$ ist $\sum_{i=1}^{n} \sum_{j=1}^{n} h_i \frac{\partial^2 f}{\partial x_i \partial x_j}(x^0) h_j$ und damit eine sogenannte **quadratische Form**, d.h. eine Abbildung $Q : \mathbb{R}^n \to \mathbb{R}$ der Form $Q(x) = x^T Ax$, wobei A eine $n \times n$ Matrix ist.

Eine quadratische Form $Q(x) = x^T Ax$ heißt nun

- **positiv definit**, wenn $Q(x) > 0 \quad \forall\, x \neq 0$
- **negativ definit**, wenn $Q(x) < 0 \quad \forall\, x \neq 0$
- **positiv semidefinit**, wenn $Q(x) \geq 0 \quad \forall\, x$
- **negativ semidefinit**, wenn $Q(x) \leq 0 \quad \forall\, x$
- **indefinit**, wenn sowohl positive als auch negative Werte angenommen werden.

Bemerkung. Die obigen Bedingungen können auch durch die Eigenwerte von A ausgedrückt werden.

- $Q > 0 \;\Leftrightarrow\;$ alle Eigenwerte sind > 0
- $Q < 0 \;\Leftrightarrow\;$ alle Eigenwerte sind < 0
- $Q \geq 0 \;\Leftrightarrow\;$ alle Eigenwerte sind ≥ 0
- $Q \leq 0 \;\Leftrightarrow\;$ alle Eigenwerte sind ≤ 0
- Q ist indefinit $\Leftrightarrow$ es gibt positive als auch negative Eigenwerte.

Bemerkung. Der Spezialfall $n = 2$ läßt sich einfach beschreiben.

Eine 2×2 Matrix A ist

- positiv definit $\Leftrightarrow$ $\det A > 0$ und $a_{11} > 0$
- negativ definit $\Leftrightarrow$ $\det A > 0$ und $a_{11} < 0$
- semidefinit $\Leftrightarrow$ $\det A \geq 0$
- indefinit $\Leftrightarrow$ $\det A < 0$.

Satz. Sei $f : \mathbb{R}^n \to \mathbb{R}$, $x^0 \in \mathbb{R}^n$ und $f \in C^2(U(x^0))$ auf einer offenen Umgebung von x^0 . Weiters sei $\operatorname{grad} f(x^0) = \vec{0}$. Dann gilt :

- $H(x^0)$ ist positiv definit $\Leftrightarrow$ f hat an x^0 ein isoliertes relatives Minimum

- $H(x^0)$ ist negativ definit $\Leftrightarrow$ f hat an x^0 ein isoliertes relatives Maximum

- $H(x^0)$ ist indefinit $\Leftrightarrow$ f hat an x^0 kein relatives Extremum.

Beweis. Sei zunächst $H(x^0)$ positiv definit. Wegen $\mathrm{grad} f(x^0) = \vec{0}$ und einer früheren Aussage gilt

$$f(x) = f(x^0) + \tfrac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} h_i \tfrac{\partial^2 f}{\partial x_i \partial x_j}(x^0) h_j + g(x) \|\vec{h}\|^2$$

wobei $g(x) \to 0$ für $x \to x^0$ bzw. $\vec{h} \to \vec{0}$.

Die quadratische Form $h^T H(x^0) h$ nimmt als stetige Funktion auf der (kompakten !) Einheitssphäre $\{x \in \mathbb{R}^n \ : \ \|x\| = 1\}$ ein Minimum $m > 0$ an.

Wegen $h^T H(x^0) h = \|h\|^2 \left(\left(\frac{h}{\|h\|} \right)^T H(x^0) \frac{h}{\|h\|} \right) \geq m \, \|h\|^2$ folgt dann für eine hinreichend kleine Umgebung $U_\varepsilon(x^0)$, in welcher $|g(x)| < \frac{m}{4}$ ist, dass

$$f(x) \geq f(x^0) + \tfrac{m}{4} \|h\|^2 \ .$$

Dann nimmt aber f in dieser Umgebung nur in x^0 den Wert $f(x^0)$ an und ist sonst überall größer, i.e. an der Stelle x^0 liegt ein isoliertes Minimum vor.

Der Fall, dass $H(x^0)$ negativ definit ist, wird durch $f \to -f$ auf den eben behandelten Fall zurückgeführt.

Der letzte Fall wird so gezeigt, dass der Vorzeichenwechsel bei $h^T H(x^0) h$ "stärker wiegt" als der Term $g(x) \|\vec{h}\|^2$. Ein solcher Punkt heißt auch **Sattelpunkt** . $\square$

Beispiel.
Man bestimme die relativen Extrema von $f(x, y) = x^2 + 4y + \frac{1}{y}$.

Die Kandidaten für ein mögliches Extremum ergeben sich aus der Bedingung $\operatorname{grad} f(x) = \vec{0}$, hier also aus $f_x = 2x = 0$ und $f_y = 4 - \frac{1}{y^2} = 0$ bzw. $x = 0$ und $y = \pm\frac{1}{2}$.

Somit sind die Punkte $P_1(0, \frac{1}{2})$ und $P_2(0, -\frac{1}{2})$ mögliche Extrema.

$f_{xx} = 2$, $f_{xy} = 0$, $f_{yy} = \frac{2}{y^3}$,

also ist $H(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & \frac{2}{y^3} \end{pmatrix}$ und $\det H(x,y) = \frac{4}{y^3}$.

Mit $\det H(x,y)|_{P_1} = 32 > 0$ und $f_{xx}|_{P_1} = 2 > 0$ folgt, dass in P_1 ein relatives Minimum vorliegt.

Wegen $\det H(x,y)|_{P_2} = -32 < 0$ liegt in P_2 ein Sattelpunkt vor.