Extremwerte von Funktionen mehrerer reeller Variabler

Bei der Bestimmung der lokalen 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, d.h. eine notwendige Bedingung für das Vorliegen eines (lokalen) Extremums anzugeben.

Es sei daran erinnert, dass im eindimensionalen Fall, also bei $f : \mathbb{R} \to \mathbb{R}$, die entsprechende Bedingung lautete, 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 $\mathrm{grad} f(x^0) = \vec{0}$.

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 $\mathrm{grad} f(x^0) \cdot \vec{a}$ verschwindet. Weil $\vec{a}$ beliebig ist, muss damit $\mathrm{grad} f(x^0) = \vec{0}$ sein. $\square$

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

Zur Gewinnung eines hinreichenden Kriteriums für das Vorliegen eines Extremums nehmen wir an, dass f zweimal stetig differenzierbar ist. Dann kann die (symmetrische) Matrix $H(x)$ der zweiten partiellen Ableitungen gebildet werden, die sog. **Hesse-Matrix** oder **zweite Ableitung** von f .

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

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

Indem f (lokal) durch das Taylor-Polynom zweiten Grades approximiert wird, stellt sich heraus, dass für das Vorliegen eines Extremums die durch die Hesse-Matrix definierte quadratische Form von wesentlicher Bedeutung ist.

Definition. Sei A eine $n \times n$ Matrix. Dann heißt die Abbildung $Q : \mathbb{R}^n \to \mathbb{R}$ mit $Q(x) = x^T A x$ eine **quadratische Form**.

Beispiel. Sei $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ -4 & 0 & 3 \end{pmatrix}$. Die dadurch definierte quadratische Form ist

$$Q(x_1, x_2, x_3) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ -4 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$$

$$= \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} x_1 + x_3 \\ 2x_2 \\ -4x_1 + 3x_3 \end{pmatrix} = x_1^2 + x_1 x_3 + 2x_2^2 - 4x_1 x_3 + 3x_3^2 =$$

$$= x_1^2 - 3x_1 x_3 + 2x_2^2 + 3x_3^2 \ .$$

Eine quadratische Form $Q(x) = x^T A x$ 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. Sind alle Eigenwerte von A reell, wie es etwa im Fall einer reellen symmetrischen Matrix der Fall ist, dann können die obigen Bedingungen 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. (**Sattelpunkt**)

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.

Häufig sucht man die Extremwerte einer Funktion $f : \mathbb{R}^n \to \mathbb{R}$ unter eingeschränkten Bedingungen für die Variablen $x_1, x_2, ..., x_n$.

Beispiel. Man bestimme die Extrema von $f(x,y) = \cos^2 x - 2\sin^2 y$, wobei $y - x = \frac{\pi}{2}$.

Die entscheidende Aussage in dieser Fragestellung ist der folgende

Satz. (**Lagrange'sche Multiplikatorregel**)

Sei $f : \mathbb{R}^n \to \mathbb{R}$ differenzierbar auf einer offenen Menge $X \subseteq \mathbb{R}^n$. Weiters seien $g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar auf X , und linear unabhängig.

Hat unter diesen Voraussetzungen f an x^0 ein Extremum, wobei nur jene $x \in X$ betrachtet werden, für die $g_1(x) = 0, \ldots, g_m(x) = 0$ ist, dann gibt es Zahlen $\lambda_1, \lambda_2, ..., \lambda_m$ sodass

$$\frac{\partial f}{\partial x_\nu}(x^0) + \sum_{l=1}^{m} \lambda_l \frac{\partial g_l}{\partial x_\nu}(x^0) = 0 \quad \text{für} \quad \nu = 1, 2, ..., n \ .$$

$\lambda_1, \lambda_2, ..., \lambda_m$ heißen **Lagrange'sche Parameter** bzw. auch **Lagrange'sche**

Multiplikatoren .

Die Funktionen $g_1, g_2, \ldots, g_m$ beschreiben dabei die m **Nebenbedingungen**.

Zur praktischen Berechnung.

Gegeben seien also $f : \mathbb{R}^n \to \mathbb{R}$ und des weiteren die m Nebenbedingungen $g_1 : \mathbb{R}^n \to \mathbb{R}, ..., g_m : \mathbb{R}^n \to \mathbb{R}$.

Man bilde die Funktion

$F(x_1, ..., x_n, \lambda_1, ..., \lambda_m) = f(x_1, ..., x_n) + \lambda_1 g_1(x_1, ..., x_n) + \lambda_2 g_2(x_1, ..., x_n) + \ldots + \lambda_m g_m(x_1, ..., x_n)$

mit den $n+m$ Unbekannten $x_1, ..., x_n, \lambda_1, ..., \lambda_m$.

Die mögliche Extremalstelle x^0 ergibt sich dann aus dem System der $n+m$ Gleichungen

$$g_l(x^0) = 0 \quad \text{für } l = 1, 2, ..., m$$

$$\frac{\partial f}{\partial x_\nu}(x^0) + \sum_{l=1}^{m} \lambda_l \frac{\partial g_l}{\partial x_\nu}(x^0) = 0 \quad \text{für } \nu = 1, 2, ..., n \ ,$$

also aus den Gleichungen

$$F_{\lambda_1} = 0 \ , \ ... \ , \ F_{\lambda_m} = 0 \quad \text{sowie} \quad F_{x_1} = 0 \ , \ ... \ , \ F_{x_m} = 0 \ .$$

Beispiel. Ein Kegelvolumen $V(r, H) = \frac{r^2 \pi H}{3}$ soll unter der Nebenbedingung $g(r, H) = r^2 + H^2 - 2RH = 0$ (R ... fest) maximal werden.

Wir betrachten $F(r, H, \lambda) = V + \lambda g = \frac{r^2 \pi H}{3} + \lambda(r^2 + H^2 - 2RH)$.

Damit erhalten wir die Gleichungen

$$F_\lambda = r^2 + H^2 - 2RH = 0$$

$$F_r = \frac{2\pi}{3} rH + \lambda 2r = 0$$

$$F_H = \frac{\pi}{3} r^2 + \lambda(2H - 2R) = 0$$

Aus der zweiten Gleichung folgt $\lambda = -\frac{\pi}{3}H$, eingesetzt in die dritte Gleichung erhalten wir $r^2 = 2H(H - R)$.

Unter Berücksichtigung der ersten Gleichung erhalten wir nun $H = \frac{4}{3}R$ und $r = \frac{2\sqrt{2}}{3}R$.