Extremwertbestimmung

In der Analysis I haben Sie gelernt, wie man Ableitungen benutzen kann, um Extrema einer Funktion zu ermitteln. Mögliche hinreichende Bedingungen sind:

$f'(a)=0 \quad f''(a)>0 \Rightarrow$ in a liegt ein lokales Minimum vor

$f'(a)=0 \quad f''(a)<0 \Rightarrow$ in a liegt ein lokales Maximum vor.

Unser Ziel ist diese Idee auf Funktionen mehrerer Veränderlichen zu verallgemeinern.

__Def__ Sei $U \subset \mathbb{R}^n$ offen, $f: U \to \mathbb{R}$, $a \in U$.
f hat ~~ein~~ in a ein lokales Minimum (bzw. lokales Maximum), falls ein $\varepsilon > 0$ existiert, so dass $f(x) \geq f(a)$ (bzw. $f(x) \leq f(a)$) für alle $x \in U$ mit $\|x-a\| < \varepsilon$.
f hat in a ein globales Minimum (bzw. globales Maximum), falls $f(x) \geq f(a)$ (bzw. $f(x) \leq f(a)$) für alle $x \in U$.
f hat in a ein lokales bzw. globales Extremum, falls f in a ein lokales Minimum oder Maximum bzw. globales besitzt.

__Satz 2.7__ (notwendige Bedingung für ein lokales Extremum) Sei $U \subset \mathbb{R}^n$ offen und $f: U \to \mathbb{R}$ eine partielle differenzierbare Funktion. Hat f in $a \in U$ ein lokales Extremum, so gilt $\operatorname{grad} f(a) = 0$.

__Bew__ Wir betrachten die Funktionen

$$g_j(t) = f(a + t e_j) \qquad j = 1, \dots, n.$$

Die Funktionen g_j ist auf einem Intervall $(-\varepsilon, \varepsilon) \subset \mathbb{R}$, $\varepsilon > 0$, definiert und wegen partieller Differenzierbarkeit von f differenzierbar. Besitzt f ein lokales Extremum in a, so hat auch g_j ein lokales Extremum in 0. Folglich ist $g_j'(0) = 0$ für $j = 1, \dots, n$.

$$0 = g_j'(0) = \frac{d f(a + t e_j)}{dt}\Big|_{t=0} = \lim_{t \to 0} \frac{f(a+te_j) - f(a)}{t} = \frac{\partial f}{\partial x_j}(a)$$

für alle $j = 1, \dots, n$.

$\Rightarrow \operatorname{grad} f(a) = 0$. $\square$

Um zu entscheiden, ob ein lokales Extremum vorliegt, wollen wir wie in einer Dimension die Formel von Taylor heranziehen.

$$f(a+h) = f(a) + \underbrace{\langle \nabla f(a), h \rangle}_{=0} + \frac{1}{2}\langle H_f(a)h, h \rangle + o(\|h\|^2), \ h \to 0$$

Für kleine Werte von h ist $\frac{1}{2}\langle H_f(a)h, h\rangle$ der führende Term. Offensichtlich ist die ganze Information über diesen Term in der Hesse-Matrix enthalten.

Wir stehen vor folgendem Problem:

Sei A eine gegebene $n \times n$ Matrix.

Ist $\langle Ah, h \rangle$ für $h \neq 0$ immer positiv (bzw. negativ)?

Dieses Problem ist nicht neu. Es wird in der linearen Algebra behandelt.

<u>Def</u> Eine symmetrische $(n \times n)$-Matrix heißt

positiv definit, falls $\langle Ah, h \rangle > 0$ für alle $h \in \mathbb{R}^n$, $h \neq 0$.

negativ definit, falls ~~$\langle Ah, h \rangle < 0$ für alle $h \in \mathbb{R}^n$, $h \neq 0$~~ $-A$ positiv definit ist, d.h. falls $\langle Ah, h \rangle < 0$ für alle $h \in \mathbb{R}^n$, $h \neq 0$.

indefinit, falls es Vektoren $h_1, h_2 \in \mathbb{R}^n$ gibt, so dass $\langle Ah_1, h_1 \rangle > 0$ und $\langle Ah_2, h_2 \rangle < 0$.

+ semidefinit

<u>Lemma</u> Sei A eine symmetrische $n\times n$-Matrix. Seien $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$ die Eigenwerte von A. Dann gilt

$$\lambda_1 \|h\|^2 \le \langle Ah, h\rangle \le \lambda_n \|h\|^2$$

für alle $h \in \mathbb{R}^n$.

<u>Beweis</u> Es ist aus der linearen Algebra bekannt, dass jede symmetrische Matrix diagonalisierbar ist. ~~Die Eigenwerte stehen~~ ~~davon~~ Auf der Diagonale stehen dann die Eigenwerte $\lambda_1, \ldots, \lambda_n$ und es gilt eine Orthonormalbasis aus den ~~da~~zugehörigen Eigenvektoren $v^{(1)}, \ldots, v^{(n)}$.
Sei nun $h \in \mathbb{R}^n$. Wir können h als $h = \sum_{i=1}^n \alpha_i v^{(i)}$ darstellen. Dann gilt

$$Ah = A\sum_{i=1}^n \alpha_i v^{(i)} = \sum_{i=1}^n \alpha_i A v^{(i)} = \sum_{i=1}^n \alpha_i \lambda_i v^{(i)},$$

$$\langle Ah, h\rangle = \left\langle \sum_{i=1}^n \alpha_i \lambda_i v^{(i)}, \sum_{j=1}^n \alpha_j v^{(j)} \right\rangle = \sum_{i=1}^n \alpha_i^2 \lambda_i,$$

$$\|h\|^2 = \left\langle \sum_{i=1}^n \alpha_i v^{(i)}, \sum_{j=1}^n \alpha_j v^{(j)} \right\rangle = \sum_{i=1}^n \alpha_i^2 .$$

Die Behauptung folgt dann aus der Ungleichung

$$\lambda_1 \sum_{i=1}^n \alpha_i^2 \le \sum_{i=1}^n \alpha_i^2 \lambda_i \le \lambda_n \sum_{i=1}^n \alpha_i^2 \qquad \square$$

<u>Satz 2.8.</u> Sei A eine symmetrische $(n\times n)$-Matrix. Es gilt:

A ist positiv definit $\Leftrightarrow$ alle Eigenwerte von A sind positiv

A ist negativ definit $\Leftrightarrow$ alle Eigenwerte von A sind negativ.

A ist indefinit $\Leftrightarrow$ Es gibt mindestens einen positiven und mindestens einen negativen Eigenwert von A

<u>Bew</u> Die Behauptungen folgen unmittelbar aus dem Lemma.

Bsp $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$

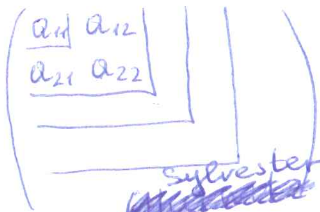
$$\det(A - \lambda E) = \det\begin{pmatrix} 1-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = (1-\lambda)(2-\lambda) - 1 = 2 + \lambda^2 - 3\lambda - 1$$
$$= \lambda^2 - 3\lambda + 1 = 0 \quad \Rightarrow \lambda_{1,2} = \frac{3 \pm \sqrt{9-4}}{2} > 0.$$

$\Rightarrow$ A ist positiv definit.

Für $n \geq 3$ kann es aber schwierig sein alle Eigenwerte zu berechnen. Wir formulieren hier ohne Beweis noch ein Kriterium, das ziemlich praktisch ist.

Sei A eine $n \times n$ - Matrix. Die Hauptminoren von A sind die Zahlen

$$A_k = \det(a_{ij})_{i,j=1,\dots,k} \qquad k = 1, \dots, n$$



(Sylvester-Kriterium) auch Hurwitz-Kriterium

Satz 2.9 Sei A eine symmetrische $n \times n$-Matrix. Dann gilt:

A ist positiv definit $\Longleftrightarrow$ Alle Hauptminoren von A sind positiv.

Bsp $n = 3$ $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

A ist positiv definit $\Longleftrightarrow$ $a_{11} > 0$, $a_{11}a_{22} - a_{21}^2 > 0$, $\det A > 0$.

Bemerkung A ist negativ definit

$\Longleftrightarrow$ $-A$ ist positiv definit

$\Longleftrightarrow$ $(-1)^k A_k > 0$.

$n = 3$ A ist negativ definit $\Longleftrightarrow$

$a_{11} < 0$, $a_{11}a_{22} - a_{12}^2 > 0$, $\det A < 0$.

Wir haben also zwei Verfahren zur Verfügung, um festzustellen, ob eine symmetrische Matrix positiv (bzw. negativ) definit ist:

- durch Eigenwerte
- Hurwitz-Kriterium

Einschub

Sylvester-Kriterium (oder Hurwitz-Kriterium)

Sei A eine symmetrische $n \times n$-Matrix.
Dann gilt:

1) A ist positiv definit $\Longleftrightarrow$ Alle Hauptminoren von A sind positiv ($A_k > 0$ f.a. $k = 1, \dots, n$)

2) A ist negativ definit $\Longleftrightarrow$ Alle Hauptminoren gerader Ordnung sind positiv und alle Hauptminoren ungerader Ordnung sind negativ ($(-1)^k A_k > 0$ für alle $k = 1, \dots, n$)

3) weder $A_k \geq 0$ noch $(-1)^k A_k \geq 0$ für alle $k = 1, \dots, n$
$\Longrightarrow$ A ist indefinit.

Bemerkung $\quad n = 2 \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

A ist positiv definit $\Longleftrightarrow$ $a_{11} > 0$, $a_{11} a_{22} - a_{12}^2 > 0$

A ist negativ definit $\Longleftrightarrow$ $a_{11} < 0$, $\det A > 0$.

A ist indefinit
$\det A < 0 \Longrightarrow A$ ist indefinit

Satz 2.10. (hinreichende Bedingung für ein lokales Extremum)

Sei $\mathcal{U} \subset \mathbb{R}^n$ offen, $f \in C^2(\mathcal{U})$, $a \in \mathcal{U}$.

Dann gilt:

1) Ist $\operatorname{grad} f(a) = 0$ und $H_f(a)$ positiv definit, so besitzt f in a ein lokales Minimum

2) Ist $\operatorname{grad} f(a) = 0$ und $H_f(a)$ negativ definit, so besitzt f in a ein lokales Maximum

3) Ist $\operatorname{grad} f(a) = 0$ und $H_f(a)$ indefinit, so hat f in a kein lokales Extremum. Im dem ~~Fall~~ nennt man a einen Sattelpunkt.

Bew 1) Sei $h \in \mathbb{R}^n$, $h \neq 0$ und $a + h \in \mathcal{U}$.

Nach dem Satz von Taylor ist

$$f(a+h) - f(a) = \underbrace{\langle \operatorname{grad} f(a), h \rangle}_{= 0} + \frac{1}{2}\langle H_f(a)h, h \rangle + o(\|h\|^2), \ h \to 0$$

$$\Rightarrow f(a+h) - f(a) = \frac{1}{2}\langle H_f(a)h, h \rangle + o(\|h\|^2)$$

$$= \frac{\|h\|^2}{2} \cdot \left[\left\langle H_f(a)\frac{h}{\|h\|}, \frac{h}{\|h\|} \right\rangle + o(1) \right]$$

Das Vorzeichen von $f(a+h) - f(a)$ stimmt mit dem Vorzeichen von $\left\langle H_f(a)\frac{h}{\|h\|}, \frac{h}{\|h\|} \right\rangle + o(1)$

überein.

Z.z: $\exists \delta > 0 : \left\langle H_f(a)\frac{h}{\|h\|}, \frac{h}{\|h\|} \right\rangle + o(1) > 0$ für alle h mit $\|h\| < \delta$.

Sei $S := \{x \in \mathbb{R}^n : \|x\| = 1\}$ die Einheitssphäre. S ist beschränkt und abgeschlossen in $\mathbb{R}^n$, deshalb ist S nach dem Satz von Heine-Borel kompakt. Nach dem Satz vom Maximum und Minimum nimmt die stetige Funktion

$$H: \cancel{x} \mapsto \langle H_f \cancel{x}, x \rangle$$

auf S ~~ihr~~ Minimum m und ihr Maximum M an. Da $H_f(a)$ positiv definit ist, gilt

$$\left\langle H_f(a)\frac{h}{\|h\|}, \frac{h}{\|h\|} \right\rangle \geq m > 0.$$

Wegen $\lim_{h\to 0} o(1) = 0$ gibt es ein $\delta > 0$,
sodass $|o(1)| < m$ für alle h mit $\|h\| < \delta$.
Dann gilt für alle h mit $\|h\| < \delta$

$$\left\langle H_f(a)\frac{h}{\|h\|}, \frac{h}{\|h\|}\right\rangle + o(1) \geq m + o(1) > 0$$

2) Ist $H_f(a)$ negativ definit, so ist $H_{-f}(a)$ positiv definit und man kann 1) auf $-f$ anwenden.

3). Seien $e_m, e_M \in S$ die Elemente, für die gilt $\overline{H}(e_m) = m$, $\overline{H}(e_M) = M$. Da $H_f(a)$ indefinit ist, gilt $m < 0 < M$.
Sei nun $h = te_m$ mit t so klein, dass $a + te_m \in U$. Dann ist

$$f(a + te_m) - f(a) = \frac{1}{2}\langle H_f(a)te_m, te_m\rangle + o(1)$$

$$= \frac{1}{2}t^2(m + o(1)), \quad \lim_{t\to 0} o(1) = 0.$$

Für t klein genug ist $m + o(1) < 0$

$\Rightarrow f(a + te_m) - f(a) < 0$

Analog bekommt man mit $h = te_M$, dass $f(a + te_M) - f(a) > 0$ für kleine Werte von t.
Also hat f kein lokales Extremum in a. $\square$

<u>Bsp</u> 1) $f(x,y) = x^2 + y^2$

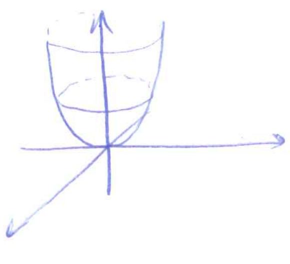
$$\nabla f = \begin{pmatrix} 2x \\ 2y \end{pmatrix} = 0 \Leftrightarrow (x,y) = (0,0)$$

$$H_f(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \Rightarrow H_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Die EW sind $\lambda_1 = 2, \lambda_2 = 2$ und sie sind positiv

$\xRightarrow{\text{Satz 2.8}}$ $H_f(0,0)$ ist positiv definit

$\Rightarrow$ $(0,0)$ ist ein lokales Minimum.



Hier sieht man das sofort $\forall (x,y) \in \mathbb{R}^2$ und $f(x,y) = x^2 + y^2 \geq 0$
$f(x,y) = 0 \Leftrightarrow (x,y) = 0 \Rightarrow$
$(0,0)$ ist ein globales Minimum

2) $f(x,y) = x^3 + 3xy^2 - 39x - 36y + 26$

$\frac{\partial f}{\partial x} = 3x^2 + 3y^2 - 39 = 0$

$\frac{\partial f}{\partial y} = 6xy - 36 = 0$

$\Leftrightarrow \begin{cases} x^2 + y^2 = 13, \\ xy = 6, \end{cases} \quad x = \frac{6}{y} \quad \left(\frac{6}{y}\right)^2 + y^2 = 13.$

$36 + y^4 - 13y^2 = 0 \Rightarrow y^4 - 13y^2 + 36 = 0$

$y^2 = \frac{13 \pm \sqrt{(-13)^2 - 4 \cdot 36}}{2} = \frac{13 \pm \sqrt{169 - 144}}{2} = \frac{13 \pm 5}{2} = 9; 4$

$\Rightarrow y^{\cancel{2}} = \pm 3 ; \pm 2$

Stationäre Punkte:

$(3,2), (-3,-2), (2,3), (-2,-3).$

$\frac{\partial^2 f}{\partial x^2} = 6x \quad \frac{\partial^2 f}{\partial y^2} = 6x \quad \frac{\partial^2 f}{\partial y \partial x} = 6y$

$\Rightarrow \quad H_f(x,y) = \begin{pmatrix} 6x & 6y \\ 6y & 6x \end{pmatrix}$

$\det H_f(x,y) = 36(x^2 - y^2)$

$(3,2)$	$6x > 0$	$\det H_f(3,2) > 0$	$\overset{HK}{\Rightarrow} H_f(3,2)$ pos def.
$(-3,-2)$	$6(-3) < 0$	$\det H_f(-3,-2) > 0$	$\overset{HK}{\Rightarrow} H_f(-3,2)$ neg def.
$(2,3)$	$6 \cdot 2 > 0$	$\det H_f(2,3) < 0$	$\overset{HK}{\Rightarrow} H_f(2,3)$ indef.
$(-2,-3)$	$6 \cdot (-2) < 0$	$\det H_f(-2,-3) < 0$	$\overset{HK}{\Rightarrow} H_f(-2,-3)$ indef.

$(3,2)$ Minimum, $(-3,-2)$ Maximum

$(2,3)$, $(-2,-3)$ kein Extremum.

Bemerkung Sei $\operatorname{grad} f(a) = 0$. Ist die Hessesche Matrix $H_f(a)$ semidefinit, so kann man keine allgemeinen Aussagen machen.

Bsp 1) $f(x,y) = x^2 + y^4$ $\quad \operatorname{grad} f = \begin{pmatrix} 2x \\ 4y^3 \end{pmatrix} \Rightarrow (0,0)$ ist ein st. Punkt

$H_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ semidefinit

f hat ein globales Minimum in $(0,0)$

2) $f(x,y) = x^2 + y^3$ $(0,0)$ ist ein stationärer Punkt.

$H_f(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 3y^2 \end{pmatrix}$ $H_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ semidefinit

Aber f hat in $(0,0)$ weder ein lokales Minimum noch ein lokales Maximum.

$f(0,y) = y^3 < 0$, wenn $y<0$

$f(x,0) = x^2 > 0$, wenn $x \neq 0$.