Numerische Verfahren und Grundlagen der Analysis

Rasa Steuding

Hochschule RheinMain Wiesbaden

Wintersemester 2011/12

Funktionen mehrerer Veränderlicher. II

Taylor-Approximation mit mehreren Veränderlichen. I

Im Folgenden sei wiederum $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$.

Wir wiederholen zunächst Taylor-Approximation von Funktionen einer Veränderlichen ($n = 1$): Gegeben eine $(m+1)$-mal stetig differenzierbare Funktion $f : (a, b) \to \mathbb{R}$, so lässt sich f um $x \in [a, b]$ durch ein Polynom m-ten Grades approximieren:

$$\begin{aligned} f(x+h) &= \sum_{k=0}^{m} \frac{f^{(k)}(x)}{k!} h^k + R_m(h) \\ &= f(x) + f'(x)h + \tfrac{1}{2} f''(x) h^2 + \ldots , \end{aligned}$$

wobei der Fehlerterm von der Form

$$R_m(h) = \frac{f^{(m+1)}(\xi)}{(m+1)!} h^{m+1}$$

mit einem ξ zwischen x und $x + h$ ist. Wie lässt sich dieses Konzept auf Funktionen mehrerer Veränderlicher verallgemeinern?

Taylor-Approximation mit mehreren Veränderlichen. II

Zunächst benötigen wir bei allgemeinem $n \in \mathbb{N}$ einen Ersatz für die erste Ableitung. Hierzu definieren wir den **Gradient von** f **in** $x \in U$ als den Zeilenvektor

$$\mathrm{grad} f(x) = (D_1 f(x), \ldots, D_n f(x)),$$

gebildet aus den partiellen Ableitungen von f.

Sei nun $f : U \to \mathbb{R}$ zweimal partiell differenzierbar. Dann heißt

$$H_f(x) = \big(D_i D_j f(x)\big)_{1 \leq i,j \leq n} = \begin{pmatrix} D_1 D_1 f(x) & D_1 D_2 f(x) & \ldots & D_1 D_n f(x) \\ \vdots & \vdots & \vdots & \vdots \\ D_n D_1 f(x) & D_n D_2 f(x) & \ldots & D_n D_n f(x) \end{pmatrix}$$

die **Hesse-Matrix** von f in $x \in U$. Ist f zweimal stetig partiell differenzierbar, so ist die Hesse-Matrix nach dem Satz von Schwarz symmetrisch.

Beispiel

Die Funktion

$$f(x,y) := \frac{x-y}{x+y}$$

ist auf $\mathbb{R}_+^2$ beliebig oft differenzierbar (als Verkettung beliebig oft differenzierbarer Funktionen), wobei $\mathbb{R}_+$ für die Menge aller positiven reellen Zahlen stehe, und es gilt

$$\mathrm{grad} f(x,y) = \left(\frac{2y}{(x+y)^2}, \frac{-2x}{(x+y)^2}\right)$$

und

$$H_f(x,y) = \begin{pmatrix} \frac{-4y}{(x+y)^3} & \frac{2x-2y}{(x+y)^3} \\ \frac{2x-2y}{(x+y)^3} & \frac{4x}{(x+y)^3} \end{pmatrix}.$$

Satz von Taylor - zweite Ordnung

Satz 13.1

Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ zweimal stetig differenzierbar. Dann gilt für $x \in U$

$$f(x+\zeta) = f(x) + \langle \mathrm{grad}\, f(x), \zeta \rangle + \tfrac{1}{2}\langle \zeta, H_f(x)\zeta \rangle + o(\|\zeta\|^2).$$

Beispiel: Die Funktion $f(x,y)$ aus dem vorangegangenen Beispiel ist um $\binom{x}{y} = \binom{1}{1}$ zu approximieren:

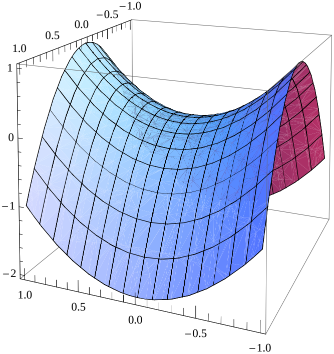
$$\begin{aligned}
f(1+\zeta_1, 1+\zeta_2) &= f(1,1) + \left\langle \mathrm{grad}\, f(1,1), \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} \right\rangle \\
&\quad + \tfrac{1}{2} \left\langle \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix}, H_f(1,1) \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} \right\rangle + o\left(\left\| \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} \right\|^2 \right) \\
&= 0 + (\tfrac{1}{2}, -\tfrac{1}{2}) \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} + \tfrac{1}{2}(\zeta_1, \zeta_2) \begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} \\
&\quad + o(\zeta_1^2 + \zeta_2^2) \\
&= \tfrac{1}{2}(\zeta_1 - \zeta_2) + \tfrac{1}{4}(-\zeta_1^2 + \zeta_2^2) + \mathrm{Fehler}.
\end{aligned}$$

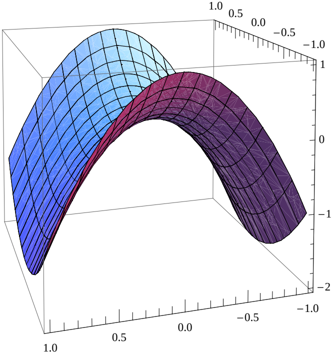
Wir wiederholen kurz die Extremwerttheorie von Funktionen einer Veränderlicher: Die Funktion $f_1(x) = x^2$ besitzt ein Minimum in $x = 0$, denn ihre Ableitung $f_1'(x) = 2x$ verschwindet und die zweite Ableitung $f_2(x) = 2$ ist positiv; ganz analog besitzt $f_2(y) = -2y^2$ in $y = 0$ ein Maximum. Damit besitzt die Funktion

$$f(x, y) = x^2 - 2y^2 \quad (= f_1(x) + f_2(y))$$

weder ein Maximum noch ein Minimum in $(x, y) = (0, 0)$. Bilder dieser Funktion zeigen eine *Sattelfläche* (siehe nächste Seite). Dieses Beispiel zeigt bereits, dass zum eindimensionalen Fall analoge Kriterien für Funktionen mehrerer Veränderlicher zu erwarten sind; tatsächlich werden die erste und zweite Ableitung durch den Gradienten bzw. die Hesse-Matrix ersetzt werden...

Sattelfläche $f(x, y) = x^2 - 2y^2$





Extremwerte. II

Ein Punkt $\xi \in U$ heißt ein **lokales Maximum** von f, falls es eine Umgebung $V \subset U$ von ξ gibt, so dass

$$f(x) \leq f(\xi) \qquad \text{für alle} \quad x \in V$$

gilt; wenn sogar strikte Ungleichung besteht,

$$f(x) < f(\xi) \qquad \text{für alle} \quad x \in V,$$

so ist ξ ein **striktes lokales Maximum** von f. Analog erklärt man ein lokales Minimum: Ein Punkt $\xi \in U$ heißt ein **lokales Minimum** von f, falls es eine Umgebung $V \subset U$ von ξ gibt, so dass

$$f(x) \geq f(\xi) \qquad \text{für alle} \quad x \in V$$

gilt; wenn sogar strikte Ungleichung besteht,

$$f(x) > f(\xi) \qquad \text{für alle} \quad x \in V,$$

so ist ξ ein **striktes lokales Minimum** von f.

Extremwerte. III

Der Oberbegriff für Maximum und Minimum ist **Extremum**.

Wir sprechen von einem **globalen Extremum**, falls die entsprechende strikte Ungleichung nicht nur für alle $x \in V$, sondern für alle x aus dem Definitionsbereich von f besteht.

Notwendiges Kriterium für Extrema

Satz 13.2

Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ einmal stetig differenzierbar. Besitzt f in $\xi \in U$ ein lokales Extremum, so verschwinden alle partiellen Ableitungen in ξ, d.h.

$$\operatorname{grad} f(\xi) = 0.$$

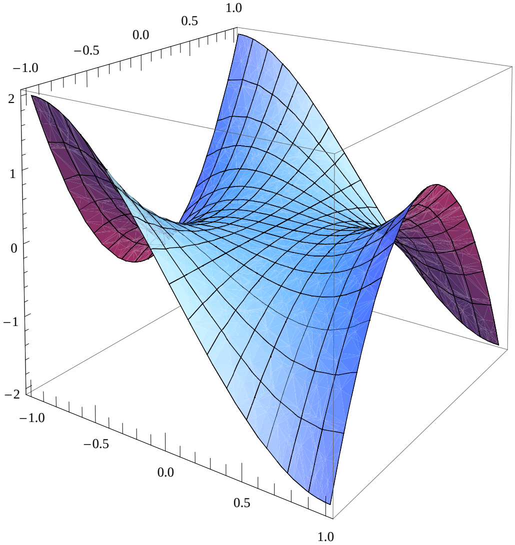
Ist f einmal stetig differenzierbar, so heißt $\xi \in U$ ein **kritischer Punkt** von f, falls $\operatorname{grad} f(\xi) = 0$. Die Extrema finden sich unter den kritischen Punkten!

Beispiele

- Wir greifen mit $f(x, y) = x^2 - 2y^2$ das Eingangsbeispiel noch einmal auf. Wegen $\mathrm{grad}\, f(x, y) = (2x, -4y)$ besitzt f in $(0, 0)$ einen kritischen Punkt, ist aber dort nicht extremal. Solche Punkte nennt man einen **Sattelpunkt**.
- Als weiteres Beispiel sei $g(x, y) = x^3 - 3xy^2$; der Graph von g heißt dann *Affensattel* (siehe nächste Seite) und erinnert wiederum an einen Sattel, bloß besitzt diese Fläche eine weitere Senkung für den Affenschwanz.

Damit ist das Kriterium aus Satz 13.2 nur notwendig für die Existenz eines Extrema, nicht aber hinreichend! Wie im Eindimensionalen können die zweiten Ableitungen (so existent), also die Hesse-Matrix, weitere Informationen liefern.

Affensattel $g(x, y) = x^3 - 3xy^2$



Hinreichendes Kriterium für Extrema

Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ heißt

- **positiv definit**, wenn $\langle x, Ax\rangle > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$,
- **positiv semidefinit**, wenn $\langle x, Ax\rangle \geq 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$,
- **negativ definit**, wenn $\langle x, Ax\rangle < 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$,
- **negativ semidefinit**, wenn $\langle x, Ax\rangle \leq 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$,
- **indefinit**, wenn es $x, y \in \mathbb{R}^n$ gibt mit $\langle x, Ax\rangle < 0 < \langle y, Ay\rangle$.

Satz 13.3

Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ zweimal stetig differenzierbar sowie $\xi \in U$ ein kritischer Punkt. Dann gilt:

1. *Ist $H_f(\xi)$ positiv definit, dann ist ξ ein striktes lokales Minimum;*
2. *ist $H_f(\xi)$ negativ definit, dann ist ξ ein striktes lokales Maximum;*
3. *ist $H_f(\xi)$ indefinit, dann ist ξ kein lokales Extremum.*

Im Falle einer semidefiniten Hesse-Matrix in einem kritischen Punkt kann keine Aussage gemacht werden.

Beispiel

Seien a, b, c reell und

$$f : \mathbb{R}^2 \to \mathbb{R}, \qquad (x, y) \mapsto f(x, y) := ax^2 + 2bxy + cy^2$$

mit $b^2 \neq ac$. Dann ist

$$\mathrm{grad}\, f(x, y) = (2ax + 2by, 2cy + 2bx) = 2(ax + by, bx + cy)$$

und der einzige kritische Punkt ist $(x, y) = (0, 0)$ (denn das homogene lineare Gleichungssystem $ax + by = bx + cy = 0$ besitzt wegen der nach Voraussetzung nicht verschwindenden Determinante $ac - b^2$ nur die triviale Lösung). Ferner ist

$$H_f(x, y) = 2 \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

(unabhängig von (x, y)) und

$$\left\langle \begin{pmatrix} x \\ y \end{pmatrix}, H_f(0, 0) \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = 2(ax^2 + 2bxy + cy^2).$$

Mit Satz 13.3 kann nun für spezielle Werte von a, b, c entschieden werden, ob ein Extremum vorliegt und welchen Typ es gegebenenfalls hat.