Analysis II

Apl. Prof. Dr. Axel Grünrock

Heinrich-Heine-Universität Düsseldorf

Sommersemester 2020

Kriterien für lokale Extrema

Im Folgenden sollen - im Wesentlichen aus der Taylorformel - notwendige und hinreichende Bedingungen für lokale Extrema reellwertiger Funktionen *mehrerer* reeller Variablen hergeleitet werden. Die Vorgehensweise ist dabei ganz analog zu den Argumenten, die wir in Analysis I für Funktionen *einer* Veränderlicher vorgebracht haben. Daher zunächst eine

Erinnerung an Analysis I:

Für eine zweimal stetig differenzierbare Funktion $f : \mathbb{R} \supset I \to \mathbb{R}$ kennen wir die folgenden notwendigen und hinreichenden Bedingungen für lokale Extrema:

Erinnerung an Analysis I:

Für eine zweimal stetig differenzierbare Funktion $f : \mathbb{R} \supset I \to \mathbb{R}$ kennen wir die folgenden notwendigen und hinreichenden Bedingungen für lokale Extrema:

(1) f besitzt in $x_0 \in I$ ein lokales Extremum $\Rightarrow f'(x_0) = 0$.

Erinnerung an Analysis I:

Für eine zweimal stetig differenzierbare Funktion $f : \mathbb{R} \supset I \to \mathbb{R}$ kennen wir die folgenden notwendigen und hinreichenden Bedingungen für lokale Extrema:

(1) f besitzt in $x_0 \in I$ ein lokales Extremum $\Rightarrow f'(x_0) = 0$.

(2) f besitzt in $x_0 \in I$ ein lokales Maximum $\Rightarrow f''(x_0) \leq 0$.

Erinnerung an Analysis I:

Für eine zweimal stetig differenzierbare Funktion $f : \mathbb{R} \supset I \to \mathbb{R}$ kennen wir die folgenden notwendigen und hinreichenden Bedingungen für lokale Extrema:

(1) f besitzt in $x_0 \in I$ ein lokales Extremum $\Rightarrow f'(x_0) = 0$.

(2) f besitzt in $x_0 \in I$ ein lokales Maximum $\Rightarrow f''(x_0) \leq 0$.

(3) $(f'(x_0) = 0$ und $f''(x_0) < 0) \Rightarrow f$ hat in x_0 ein isoliertes lokales Maximum.

(2) und (3) gelten entsprechend für lokale Minima, wenn wir die Ungleichheitszeichen umkehren.

Erinnerung an Analysis I:

Für eine zweimal stetig differenzierbare Funktion $f : \mathbb{R} \supset I \to \mathbb{R}$ kennen wir die folgenden notwendigen und hinreichenden Bedingungen für lokale Extrema:

(1) f besitzt in $x_0 \in I$ ein lokales Extremum $\Rightarrow f'(x_0) = 0$.

(2) f besitzt in $x_0 \in I$ ein lokales Maximum $\Rightarrow f''(x_0) \leq 0$.

(3) $(f'(x_0) = 0$ und $f''(x_0) < 0) \Rightarrow f$ hat in x_0 ein isoliertes lokales Maximum.

(2) und (3) gelten entsprechend für lokale Minima, wenn wir die Ungleichheitszeichen umkehren.
Auch wenn es nur geringe Unterschiede zum Eindimensionalen gibt, sollten wir zunächst notieren, was wir im Mehrdimensionalen unter einem lokalen Extremum verstehen wollen:

Definition

Es sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}$ eine Funktion. f besitzt in $x \in \Omega$ ein *lokales Maximum (Minimum)*, falls eine Umgebung $U(x)$ existiert, so dass

$$f(x) \geq f(y) \qquad (f(x) \leq f(y)) \text{ für alle } y \in U(x).$$

Gilt $f(x) > f(y)$ (bzw. $f(x) < f(y)$) für alle $y \in U(x) \setminus \{x\}$, so spricht man von einem *isolierten* lokalen Maximum (bzw. Minimum).

Bemerkungen:

(1) Unter einem *Extremum* verstehen wir ein Maximum oder ein Minimum.

Bemerkungen:

(1) Unter einem *Extremum* verstehen wir ein Maximum oder ein Minimum.

(2) Unter einem *globalen* Maximum ist $\max\{f(x) : x \in \Omega\}$ zu verstehen, wenn dieses existiert. Entsprechend ein globales Minimum.

Bemerkungen:

(1) Unter einem *Extremum* verstehen wir ein Maximum oder ein Minimum.

(2) Unter einem *globalen* Maximum ist $\max\{f(x) : x \in \Omega\}$ zu verstehen, wenn dieses existiert. Entsprechend ein globales Minimum.

(3) Im Gegensatz zum globalen Maximum, welches eindeutig bestimmt ist, kann es eine Vielzahl lokaler Maxima geben.

Das einfache notwendige Kriterium

In Verallgemeinerung der notwendigen Bedingung (1) zeigen wir:

Satz 3: Es sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}$ partiell differenzierbar. Wenn f in $x_0 \in \Omega$ ein lokales Extremum besitzt, so ist $\nabla f(x_0) = 0$.

Beweis: Wir wählen $\varepsilon > 0$ so, dass die offene Kugel $B_\varepsilon(x_0) \subset \Omega$ ist und setzen (für $1 \leq j \leq n$)

$$g_j : (-\varepsilon, \varepsilon) \to \mathbb{R},\ t \mapsto g_j(t) := f(x_0 + te_j).$$

Dann besitzt jedes g_j in $t = 0$ ein lokales Extremum. Nach dem Kriterium (1) ist also für alle $j \in \{1, ..., n\}$

$$0 = g_j'(0) = \lim_{t \to 0} \frac{1}{t}(f(x_0 + te_j) - f(x_0)) = \frac{\partial f}{\partial x_j}(x_0),$$

also $\nabla f(x_0) = 0$. $\square$

Fragestellung zu den Bedingungen an die zweiten Ableitungen

Welche Bedingung an die Hesse-Matrix

$$\mathrm{Hess} f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{1 \le i,j \le n}$$

verallgemeinert in geeigneter Weise die Ungleichungen $f''(x) \geq 0$, $f''(x) > 0$ etc. aus den Kriterien (2) und (3) ?
Da zwischen Matrizen keine offensichtliche $<$ - Relation besteht, müssen wir an dieser Stelle zunächst einige neue Begriffe einführen:

Definition

Eine Matrix $A \in M_n(\mathbb{R})$ heißt

(1) *positiv definit*, wenn für alle $\xi \in \mathbb{R}^n \setminus \{0\}$ gilt $\langle \xi, A\xi \rangle > 0$;

Definition

Eine Matrix $A \in M_n(\mathbb{R})$ heißt

(1) *positiv definit*, wenn für alle $\xi \in \mathbb{R}^n \setminus \{0\}$ gilt $\langle \xi, A\xi \rangle > 0$;

(2) *positiv semidefinit*, wenn für alle $\xi \in \mathbb{R}^n$ gilt $\langle \xi, A\xi \rangle \geq 0$;

Definition

Eine Matrix $A \in M_n(\mathbb{R})$ heißt

(1) *positiv definit*, wenn für alle $\xi \in \mathbb{R}^n \setminus \{0\}$ gilt $\langle \xi, A\xi \rangle > 0$;

(2) *positiv semidefinit*, wenn für alle $\xi \in \mathbb{R}^n$ gilt $\langle \xi, A\xi \rangle \geq 0$;

(3) *negativ (semi-)definit*, wenn $-A$ positiv (semi-)definit ist;

Definition

Eine Matrix $A \in M_n(\mathbb{R})$ heißt

(1) *positiv definit*, wenn für alle $\xi \in \mathbb{R}^n \setminus \{0\}$ gilt $\langle \xi, A\xi \rangle > 0$;

(2) *positiv semidefinit*, wenn für alle $\xi \in \mathbb{R}^n$ gilt $\langle \xi, A\xi \rangle \geq 0$;

(3) *negativ (semi-)definit*, wenn $-A$ positiv (semi-)definit ist;

(4) *indefinit*, falls $\xi, \eta \in \mathbb{R}^n$ existieren, so dass $\langle \xi, A\xi \rangle > 0 > \langle \eta, A\eta \rangle$.

Hierbei bezeichnet $\langle \quad , \quad \rangle$ das Standard-Skalarprodukt auf dem $\mathbb{R}^n$.

Definition

Eine Matrix $A \in M_n(\mathbb{R})$ heißt

(1) *positiv definit*, wenn für alle $\xi \in \mathbb{R}^n \setminus \{0\}$ gilt $\langle \xi, A\xi \rangle > 0$;

(2) *positiv semidefinit*, wenn für alle $\xi \in \mathbb{R}^n$ gilt $\langle \xi, A\xi \rangle \geq 0$;

(3) *negativ (semi-)definit*, wenn $-A$ positiv (semi-)definit ist;

(4) *indefinit*, falls $\xi, \eta \in \mathbb{R}^n$ existieren, so dass $\langle \xi, A\xi \rangle > 0 > \langle \eta, A\eta \rangle$.

Hierbei bezeichnet $\langle \quad , \quad \rangle$ das Standard-Skalarprodukt auf dem $\mathbb{R}^n$.

Für *symmetrische* Matrizen hat man die folgenden Kriterien für Definitheit:

Lemma 1

Eine symmetrische Matrix $A \in M_n(\mathbb{R})$ ist genau dann

(1) positiv definit, wenn alle Eigenwerte $\lambda_1, ...\lambda_n$ von A positiv sind;

Lemma 1

Eine symmetrische Matrix $A \in M_n(\mathbb{R})$ ist genau dann

(1) positiv definit, wenn alle Eigenwerte $\lambda_1, ...\lambda_n$ von A positiv sind;

(2) positiv semidefinit, wenn alle λ_i, $1 \leq i \leq n$, nicht negativ sind;

Lemma 1

Eine symmetrische Matrix $A \in M_n(\mathbb{R})$ ist genau dann

(1) positiv definit, wenn alle Eigenwerte $\lambda_1, ...\lambda_n$ von A positiv sind;

(2) positiv semidefinit, wenn alle λ_i, $1 \leq i \leq n$, nicht negativ sind;

(3) indefinit, wenn es Eigenwerte λ und μ von A gibt, so dass $\lambda > 0 > \mu$.

Beweis des Lemmas

Da A reell und symmetrisch ist, gilt

(1) alle Eigenwerte $\lambda_1, ... \lambda_n$ von A sind reell,

Beweis des Lemmas

Da A reell und symmetrisch ist, gilt

(1) alle Eigenwerte $\lambda_1, ...\lambda_n$ von A sind reell,

(2) es gibt eine Orthonormalbasis (ONB) $v_1, ..., v_n$ des $\mathbb{R}^n$, so dass $Av_j = \lambda_j v_j$ für $1 \leq j \leq n$.

(Das ist der sogenannte "Spektralsatz für symmetrische Matrizen, dieser ist Gegenstand der LA I, wird aber auch im Verlauf dieser Vorlesung noch mit analytischen Mitteln bewiesen.)

Ist nun $\xi = \sum_{i=1}^{n} \alpha_i v_i$, so folgt

$$\langle \xi, A\xi \rangle = \sum_{i,j=1}^{n} \alpha_i \alpha_j \langle v_i, Av_j \rangle = \sum_{i=1}^{n} \alpha_i^2 \lambda_i.$$

Hieraus sind die obigen Behauptungen ablesbar. $\square$

Spezialfall $n = 2$ (I)

Ist $A \in M_2(\mathbb{R})$ symmetrisch, so hat sie die Gestalt

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

mit $a, b, c \in \mathbb{R}$. Die Eigenwerte erhält man als Nullstellen des charakteristischen Polynoms, das ist hier

$$\det(A - \lambda E_2) = \det \begin{pmatrix} a - \lambda & b \\ b & c - \lambda \end{pmatrix}$$

$$= (a - \lambda)(c - \lambda) - b^2 = \lambda^2 - (a + c)\lambda + ac - b^2.$$

Sie sind gerade gegeben durch

Spezialfall $n = 2$ (II)

$$\lambda_{1,2} = \frac{a+c}{2} \pm \sqrt{\frac{(a+c)^2}{4} + b^2 - ac} = \frac{a+c}{2} \pm \sqrt{\frac{(a-c)^2}{4} + b^2}.$$

Diese sind von unterschiedlichem Vorzeichen genau dann, wenn

$$0 > \lambda_1 \lambda_2 = \frac{(a+c)^2}{4} - \frac{(a-c)^2}{4} - b^2 = ac - b^2 = \det A.$$

Also ist A indefinit genau dann, wenn $\det A = ac - b^2 < 0$ ist. Die gleiche Rechnung zeigt: A ist (semi-)definit genau dann, wenn $\det A > 0$ (bzw. ≥ 0) ist, und zwar positiv (semi-)definit, wenn zusätzlich $a > 0$ ($a \geq 0$) oder $c > 0$ ($c \geq 0$) gilt.

Dieser Spezialfall hat eine Verallgemeinerung in dem folgenden ...

Determinantenkriterium von Hurwitz

Lemma 2: Es sei $A = (a_{ij})_{1 \leq i,j \leq n}$ eine reelle symmetrische Matrix und, für $k \in \{1, ..., n\}$, $A_k = (a_{ij})_{1 \leq i,j \leq k}$. Dann gilt: A ist positiv definit genau dann, wenn $\det A_k > 0$ für alle $k \in \{1, ..., n\}$ ist.

Bemerkungen:

Determinantenkriterium von Hurwitz

Lemma 2: Es sei $A = (a_{ij})_{1 \leq i,j \leq n}$ eine reelle symmetrische Matrix und, für $k \in \{1, ..., n\}$, $A_k = (a_{ij})_{1 \leq i,j \leq k}$. Dann gilt: A ist positiv definit genau dann, wenn $\det A_k > 0$ für alle $k \in \{1, ..., n\}$ ist.

Bemerkungen:

(1) Einen Beweis finden Sie im 2. Band des Buches von Kaballo, Satz 20.12.

Determinantenkriterium von Hurwitz

Lemma 2: Es sei $A = (a_{ij})_{1 \leq i,j \leq n}$ eine reelle symmetrische Matrix und, für $k \in \{1, ..., n\}$, $A_k = (a_{ij})_{1 \leq i,j \leq k}$. Dann gilt: A ist positiv definit genau dann, wenn $\det A_k > 0$ für alle $k \in \{1, ..., n\}$ ist.

Bemerkungen:

(1) Einen Beweis finden Sie im 2. Band des Buches von Kaballo, Satz 20.12.

(2) In Lemma 2 kann man *nicht* "positiv definit genau dann, wenn $\det A_k > 0$ für alle $k \in \{1, ..., n\}$" durch "positiv semidefinit genau dann, wenn $\det A_k \geq 0$ für alle $k \in \{1, ..., n\}$" ersetzen. Beispiel: Die 2×2 - Diagonalmatrix mit den Diagonalelementen 0 und -1 ist nicht positiv semidefinit, aber es ist $\det A_1 = \det A_2 = 0$.

Determinantenkriterium von Hurwitz

Lemma 2: Es sei $A = (a_{ij})_{1 \leq i,j \leq n}$ eine reelle symmetrische Matrix und, für $k \in \{1, ..., n\}$, $A_k = (a_{ij})_{1 \leq i,j \leq k}$. Dann gilt: A ist positiv definit genau dann, wenn $\det A_k > 0$ für alle $k \in \{1, ..., n\}$ ist.

Bemerkungen:

(1) Einen Beweis finden Sie im 2. Band des Buches von Kaballo, Satz 20.12.

(2) In Lemma 2 kann man *nicht* "positiv definit genau dann, wenn $\det A_k > 0$ für alle $k \in \{1, ..., n\}$" durch "positiv semidefinit genau dann, wenn $\det A_k \geq 0$ für alle $k \in \{1, ..., n\}$" ersetzen. Beispiel: Die 2×2 - Diagonalmatrix mit den Diagonalelementen 0 und -1 ist nicht positiv semidefinit, aber es ist $\det A_1 = \det A_2 = 0$.

(3) Will man mit diesem Kriterium entscheiden, ob eine Matrix negativ definit ist, so ist Lemma 2 auf $-A$ anzuwenden.

Ergänzendes Beispiel, als Aufgabe formuliert

Die Begriffe "positiv definit", "positiv semidefinit" usw. sind auch für nicht-symmetrische Matrizen formuliert worden, während die beiden Kriterien nur für symmetrische Matrizen gelten. Untersuchen Sie, für welche Winkel $\alpha \in [0, 2\pi)$ die Matrix

$$R_\alpha := \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

positiv definit, negativ definit oder indefinit ist. Beachten Sie, dass R_α eine Drehung in der Ebene um den Winkel α beschreibt. Welche Eigenwerte hat R_α ?

Eine weitere notwendige Bedingung für lokale Extrema

Satz 4: Es sei $\Omega \subset \mathbb{R}^n$ offen. $f \in C^2(\Omega)$ besitze in $x_0 \in \Omega$ ein lokales Maximum. Dann ist $\mathrm{Hess} f(x_0)$ negativ semidefinit.

Beweis: Es reicht, die Ungleichung

$$\langle h, \mathrm{Hess} f(x_0) h \rangle \leq 0$$

für alle $h \in \mathbb{R}^n$ mit $|h| = 1$ zu zeigen. Zu einem solchen h wählen wir $\varepsilon > 0$, so dass $[x_0, x_0 + \varepsilon h] \subset \Omega$. Da f in x_0 ein lokales Maximum besitzt, gilt (ggf. nach Verkleinerung von ε) aufgrund der Taylorschen Formel:

$$0 \geq f(x_0 + \varepsilon h) - f(x_0) = \nabla f(x_0) \varepsilon h + \frac{\varepsilon^2}{2} \langle h, \mathrm{Hess} f(\xi_\varepsilon) h \rangle$$

mit einem $\xi_\varepsilon \in [x_0, x_0 + \varepsilon h]$.

Bitte wenden!

Fortsetzung des Beweises

Da nach Satz 3 $\nabla f(x_0) = 0$ ist, haben wir also

$$\langle h, \mathrm{Hess} f(\xi_\varepsilon) h \rangle \leq 0.$$

Nun ist $f \in C^2(\Omega)$ vorausgesetzt, so dass wir für $\varepsilon \to 0$ erhalten

$$\langle h, \mathrm{Hess} f(x_0) h \rangle \leq 0.$$

□

Folgerung: $f \in C^2(\Omega)$ besitze in $x_0 \in \Omega$ ein lokales Minimum.
Dann ist $\mathrm{Hess} f(x_0)$ positiv semidefinit.

(Um dies einzusehen, muss man nur Satz 4 auf $-f$ anwenden.)

Das hinreichende Kriterium

Satz 5: Es sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^2(\Omega)$, in $x_0 \in \Omega$ gelte $\nabla f(x_0) = 0$ und $\mathrm{Hess} f(x_0)$ sei negativ definit. Dann besitzt f in x_0 ein isoliertes lokales Maximum.

Beweis: Da f zweimal *stetig* differenzierbar ist, existiert ein $\varepsilon > 0$, so dass $\mathrm{Hess} f(\xi)$ negativ definit ist für alle $\xi \in B_\varepsilon(x_0)$. Für $h \in \mathbb{R}^n$ mit $|h| < \varepsilon$ ergibt die Taylorformel

$$f(x_0 + h) - f(x_0) = \nabla f(x_0) h + \frac{1}{2} \langle h, \mathrm{Hess} f(\xi) h \rangle < 0,$$

da nach Voraussetzung $\nabla f(x_0) = 0$ und nach der Vorüberlegung der zweite Summand negativ ist. Also $f(x_0 + h) < f(x_0)$. $\square$

Folgerung: Ist $\nabla f(x_0) = 0$ und $\mathrm{Hess} f(x_0)$ positiv definit, und gelten ansonsten die Voraussetzungen von Satz 5, so besitzt f in x_0 ein isoliertes lokales Minimum.