2.2 Lokale Extrema ohne Nebenbedingungen

Das Maximum einer Funktion $f: \Omega \to \mathbb{R}$ ist der größte Funktionswert von f, sofern dieser existiert. Entsprechend ist das Minimum der kleinste Funktionswert. Die Punkte $x \in \Omega$, an denen diese Werte angenommen werden, heißen Maximal- bzw. Minimalstellen. Extrema sind Maxima oder Minima, Extrem(al)stellen sind Maximal- oder Minimalstellen.

Mitunter benutzt man für den größten (kleinsten) Funktionswert auch die Bezeichnung "globales Maximum (Minimum)" bzw. "absolutes Maximum (Minimum)", um von der folgenden Situation zu unterscheiden:

Def.: Es sei $\Omega \subset \mathbb{R}^n$ und $f: \Omega \to \mathbb{R}$ eine Funktion.

(1) $x^* \in \Omega$ heißt eine lokale (relative) Maximalstelle von f, wenn ein $\delta > 0$ existiert, so dass für alle $x \in \Omega$ mit $|x - x^*| < \delta$ gilt, dass $f(x^*) \geq f(x)$. $f(x^*)$ heißt dann ein lokales (relatives) Maximum von f.

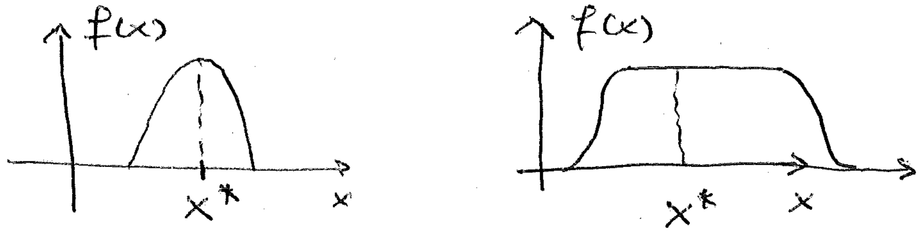
(2) $x_* \in \Omega$ heißt eine lokale (relative) Minimalstelle von f, wenn x_* eine lokale (relative) Maximalstelle von $-f$ ist. In diesem Fall heißt $f(x_*)$ ein lokales (relatives) Minimum von f.

Bem.: (1) lokales (relatives) Extremum = lok. (rel.) Maximum oder Minimum.

(2) Isoliertes lokales Extremum: Anstelle von $f(x^*) \geq f(x)$ bzw. $f(x_*) \leq f(x)$ hat man $f(x^*) > f(x)$

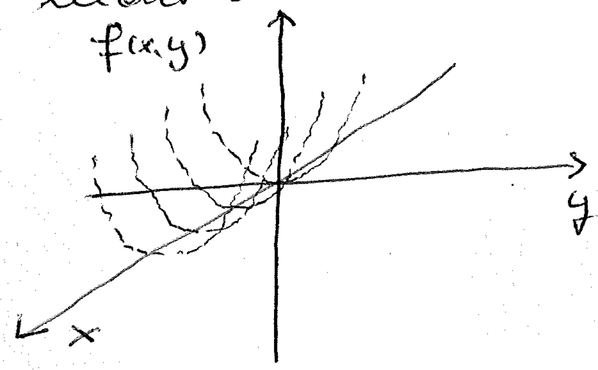
bzw. $f(x_*) < f(x)$ für alle $x \in \Omega \setminus \{x_*^*\}$ mit $|x-x_*^*|<\delta$. Bsp.:

$n=1$:

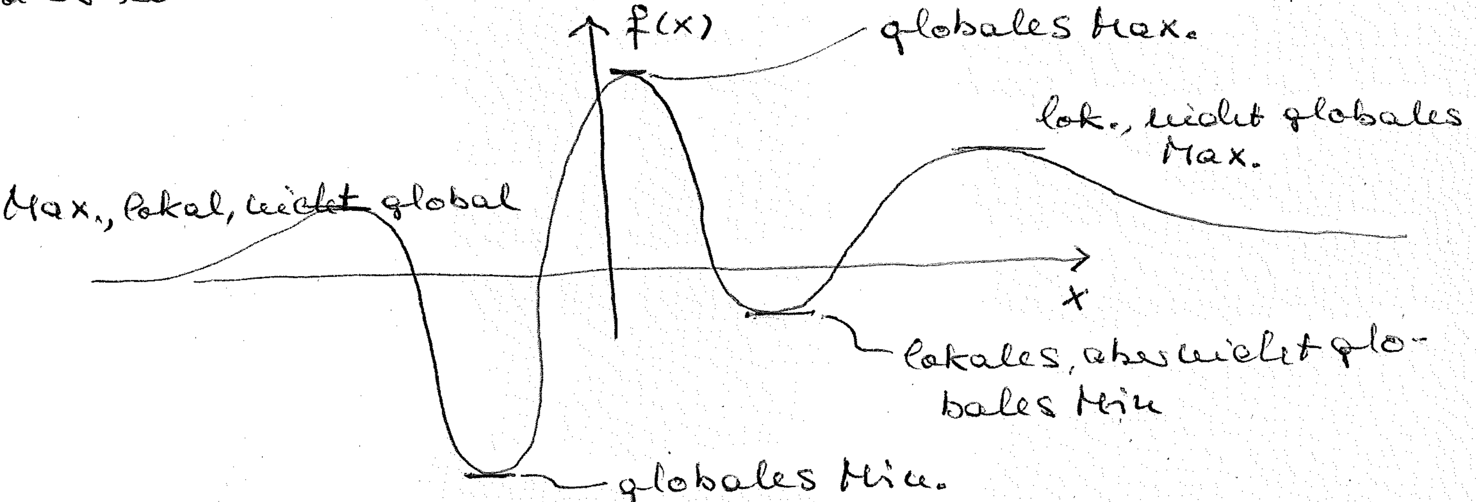


isoliert — nicht isoliert

$n=2$, $f:\mathbb{R}^2\to\mathbb{R}$, $\binom{x}{y}\mapsto y^2$



(3) Jedes globale Extremum ist auch ein lokales, aber nicht umgekehrt. Bsp für $n=1$:



Nachdem wir also die Begriffe geklärt haben, nehmen wir im Folgenden an, dass $f:\Omega\to\mathbb{R}$ zweimal stetig differenzierbar ist. Weiter setzen wir die Offenheit von Ω voraus.

<u>Satz</u>: Wenn $f:\Omega\to\mathbb{R}$ in $x_0\in\Omega$ ein lokales Extremum besitzt, ist $\nabla f(x_0)=0$.

Begründung: Für Funktionen lediglich einer Variablen ist diese notwendige Bedingung hinlänglich bekannt, und wir müssen nur den Fall $n \geq 2$ auf den Fall $n=1$ zurückführen. Dazu definieren wir die Hilfsfunktion

$$\varphi_k(t) = f(x_0 + te_k).$$

Da f in x_0 ein lokales Extremum besitzt, hat φ_k ein ebensolches lokales Extremum in $t_0 = 0$. Also gilt

$$0 = \varphi_k'(0) = \frac{d}{dt} f(x_0+te_k)\Big|_{t=0} = \frac{\partial f}{\partial x_k}(x_0).$$

Das gilt für alle $k \in \{1, \ldots, n\}$, also ist $\nabla f(x_0) = 0$.

Def.: $x \in \Omega$ mit $\nabla f(x) = 0$ heißt eine <u>kritische Stelle</u> von f. ("Kandidaten" für lokale Extremstellen)

Was tritt an die Stelle der hinreichenden Bedingung $f''(x_0) > 0$ $(f''(x_0) < 0)$ für ein Minimum (Maximum)? Für Matrizen wie die Hesse-Matrix gibt es ja keine Ordnungsrelation "$\leq$"! Zur Beantwortung dieser Frage greifen wir auf die Taylor-Approximation zweiter Ordnung zurück:

$$f(x+h) = f(x) + \nabla f(x) \cdot h + \frac{1}{2} h^T \cdot \mathrm{Hess} f(\xi) h \quad \text{Zwischenstelle } \xi \in [x, x+h]$$

Nehmen wir an, dass in $x \in \Omega$ ein lokales Maximum vorliegt. Dann sind

$$\left.\begin{array}{l} \nabla f(x) = 0 \quad \text{und} \\ f(x+h) - f(x) \leq 0 \end{array}\right\} \Rightarrow h^T \cdot \mathrm{Hess} f(\xi) h \leq 0 \quad \forall h \in B_\delta(0)$$

$$\underset{\substack{f \text{ 2x stetig} \\ \text{diff'bar}, \ \xi \to 0}}{\Longrightarrow} h^T \cdot \mathrm{Hess} f(x) h \leq 0 \quad \text{(Notwendige Bedingung)}$$

Nehmen wir hingegen an, dass

$$\nabla f(x) = 0 \quad \text{und} \quad h^T \cdot \mathrm{Hess} f(x) h < 0 \quad \forall h \in \mathbb{R}^n \setminus \{0\}$$

Dann folgt für alle hinreichend nahe bei x liegenden $\xi \in [x, x+h]$ dass $h^T \cdot \mathrm{Hess} f(\xi) h < 0$ (C^2-Vor. an f!) und damit

$$f(x+h) - f(x) = \frac{1}{2} h^T \cdot \mathrm{Hess} f(\xi) \cdot h < 0 \quad \text{für } h \in B_\delta(0)$$

D.h. $f(x) > f(x+h)$ $\forall h \in B_\delta(0)$; f besitzt in x ein <u>isoliertes</u> lokales Maximum. (Für ein Minimum sind die $\leq$ bzw $<$ - Zeichen umzukehren.)

Diese Überlegungen führen zu folgender Begriffsbildung:

<u>Def.</u> $A \in \mathbb{R}^{n \times n}$ heißt

(a) positiv semidefinit, wenn $h^T \cdot A h \geq 0$ ist $\forall h \in \mathbb{R}^n$;

(b) positiv definit, wenn $h^T \cdot A h > 0$ ist $\forall h \in \mathbb{R}^n \setminus \{0\}$;

(c) negativ (semi)definit, wenn $-A$ positiv (semi) definit ist, und

(d) indefinit, wenn $h_1 \in \mathbb{R}^n$ und $h_2 \in \mathbb{R}^n$ existieren, so dass $h_1^T \cdot A h_1 < 0 < h_2^T A h_2$.

<u>Bem.</u>: Für $h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathbb{R}^n$ und $A = (a_{ij})_{1 \leq i,j \leq n} \in \mathbb{R}^{n \times n}$ ist

$$h^T \cdot A h = \langle h, A h \rangle = \sum_{i,j=1}^{n} a_{ij} h_i h_j$$

Mit Hilfe des Definitheitsbegriffs können wir die oben gezogenen Schlussfolgerungen aus der Taylorapproximation in zwei Sätzchen zusammenfassen:

Satz 1 (Notwendige Bedingung für lokale Extrema): Es seien $\Omega \subset \mathbb{R}^n$ offen und $f: \Omega \to \mathbb{R}$ zweimal stetig differenzierbar. Besitzt f in $x_0 \in \Omega$ ein lokales Maximum, so gelten

(i) $\nabla f(x_0) = 0$ und

(ii) $\operatorname{Hess} f(x_0)$ ist negativ semidefinit.

Folgerungen:

(1) Besitzt f in $x_0 \in \Omega$ ein lokales Minimum, so gelten $\nabla f(x_0) = 0$ und $\operatorname{Hess} f(x_0)$ ist positiv semidefinit.

(2) Ist $\operatorname{Hess} f(x_0)$ indefinit, so besitzt f in x_0 kein Extremum.

Satz 2 (Hinreichende Bedingung für lokale Extrema): Es seien $\Omega \subset \mathbb{R}^n$ offen und $f: \Omega \to \mathbb{R}$ zweimal stetig differenzierbar. In $x_0 \in \Omega$ gelte

(i) $\nabla f(x_0) = 0$ und

(ii) $\operatorname{Hess} f(x_0)$ ist negativ definit.

Dann nimmt f in x_0 ein isoliertes lokales Maximum an.

Folgerung: Wenn in $x_0 \in \Omega$ $\nabla f(x_0) = 0$ und $\operatorname{Hess} f(x_0)$ positiv definit ist, besitzt f in x_0 ein isoliertes lokales Minimum.

Bem.: Ist $\nabla f(x_0) = 0$ und $\operatorname{Hess} f(x_0)$ semidefinit (positiv oder negativ), so kann verschiedenes passieren:

(i) f besitzt ein isoliertes lokales Extremum;

Bsp.: $f(x,y) = x^4 + y^4$, $x_0 = (0,0)$, $\text{Hess} f(x_0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

(ii) f besitzt ein nicht isoliertes lokales Extremum;

Bsp.: $f(x,y) = y^4$, $x_0 = (0,0)$, $\text{Hess} f(x_0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

(iii) f besitzt kein Extremum;

Bsp.: $f(x,y) = x^4 - y^4$, $x_0 = (0,0)$, $\text{Hess} f(x_0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Insofern besteht eine Lücke zwischen der notwendigen und der hinreichenden Bedingung. In dieser befinden sich u.a. die nicht isolierten Extrema.

Zur Lösung von Extremwertaufgaben gehört also die Untersuchung der Hesse-Matrix auf Definitheit. Die lineare Algebra stellt dazu zwei Kriterien zur Verfügung:

(1) Eigenwertkriterium: Eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ mit Eigenwerten $\lambda_1, \dots, \lambda_n$ ist

(a) positiv semidefinit $\Leftrightarrow \lambda_i \geq 0 \;\; \forall i \in \{1, \dots, n\}$,

(b) positiv definit $\Leftrightarrow \lambda_i > 0 \;\; \forall i \in \{1, \dots, n\}$;

(c) negativ semidefinit $\Leftrightarrow \lambda_i \leq 0 \;\; \forall i \in \{1, \dots, n\}$;

(d) negativ definit $\Leftrightarrow \lambda_i < 0 \;\; \forall i \in \{1, \dots, n\}$;

(e) indefinit $\Leftrightarrow \exists\, i, j \in \{1, \dots, n\}$, sodass $\lambda_i < 0 < \lambda_j$.

Der Fall $n=2$, d.h. $A=\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ mit $a,b,c\in\mathbb{R}$.

Die Eigenwerte sind allgemein die Nullstellen des charakteristischen Polynoms

$$P(\lambda)=\det(\lambda I-A),$$

im Fall einer symmetrischen 2×2-Matrix also von

$$P(\lambda)=\det\begin{pmatrix} a-\lambda & b \\ b & c-\lambda \end{pmatrix}=(a-\lambda)(c-\lambda)-b^2=\lambda^2-(a+c)\lambda+ac-b^2.$$

$$\Rightarrow P(\lambda)=0 \iff \lambda_{1,2}=\frac{a+c}{2}\pm\left(\left(\frac{a+c}{2}\right)^2-ac+b^2\right)^{\frac{1}{2}}.$$

In diesem Fall gilt:

$$A \text{ ist (semi-)definit} \iff \lambda_1\lambda_2>0 \;(\lambda_1\lambda_2\geq 0) \iff ac-b^2 \underset{(=)}{>} 0,$$

(d.h. auch A ist indefinit $\iff ac-b^2<0$!)

und zwar positiv (semi-)definit, wenn $\overbrace{a+c}^{=\lambda_1+\lambda_2} \underset{(=)}{>} 0$.

Wir können also die Definitheit von $A=\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ an Determinante und Spur ablesen, ohne die Eigenwerte exakt zu berechnen. Das funktioniert ähnlich auch für größere Matrizen, für die diese Beobachtung ungleich wichtiger ist. (Es gibt keine allgemeine Lösungsformel für Gleichungen 5. oder höheren Grades!)

2. Determinantenkriterium (Hurwitz)

Für eine symmetrische Matrix $A=(a_{ij})_{1\leq i,j\leq n}$ bildet man die Untermatrizen $A_k=(a_{ij})_{1\leq i,j\leq k}$, wobei $k\leq n$ ist, also

$$A_1=(a_{11}),\quad A_2=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},\quad A_3=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},\ \text{usw.}.$$

Dann gilt: A ist positiv definit $\iff \det A_k>0\ \forall k\in\{1,\dots,n\}$.

Vorsicht: Gilt nicht entsprechend für semidefinit oder negativ definit! Wende stattdessen das Kriterium auf $-A$ an!

Bsp.: Für $(x,y) \in \mathbb{R}^2$ sei – mit reellem Parameter $\lambda > 0$ –

$$f(x,y) = x^2 + y^2 - \lambda xy + 2(x-y).$$

1. Berechnung des Gradienten und Bestimmung der krit. Stellen:

$$\nabla f(x,y) = (2x - \lambda y + 2,\ 2y - \lambda x - 2) \overset{!}{=} (0,0)$$

$$\begin{array}{lll} 2x - \lambda y = -2 \quad (1) & \quad (1)+(2): (2-\lambda)(x+y) = 0 & \quad (1') \\ -\lambda x + 2y = 2 \quad (2) & \quad (1)-(2): (2+\lambda)(x-y) = -4 & \quad (2') \end{array}$$

Fallunterscheidung:

(a) $\lambda = 2$: $(1')$ stets erfüllt, $(2')$ ergibt $x - y = -1 \Leftrightarrow y = x+1$

Alle Punkte auf der Geraden $\{(x,y) \mid y = x+1\}$ sind kritisch.

(b) $\lambda \neq 2$: $(1')$ ergibt $y = -x$. Einsetzen in $(2')$ liefert

$$(2+\lambda)\, 2x = -4 \Rightarrow x = \frac{-2}{2+\lambda} \Rightarrow (x_c, y_c) = \frac{2}{2+\lambda}(-1,1).$$

2. Berechnung der 2. Ableitungen und Anordnung zur Hesse-Matrix

$$\operatorname{Hess} f(x,y) = \begin{pmatrix} 2 & -\lambda \\ -\lambda & 2 \end{pmatrix}$$

unabhängig von (x,y), da ein quadratisches Polynom untersucht wird! Ausnahme!

Wert der Determinante

$$\det \operatorname{Hess} f(x,y) = 4 - \lambda^2 \begin{cases} > 0, & \text{falls } 0 < \lambda < 2 \quad (i) \\ = 0, & \text{falls } \lambda = 2 \quad (ii) \\ < 0, & \text{falls } \lambda > 2 \quad (iii). \end{cases}$$

3. Auswertung

(i) $\operatorname{Hess} f(x,y)$ ist definit, und zwar positiv, da $a + c = 2 + 2 > 0$.

In $(x_c, y_c) = \frac{2}{2+\lambda}(-1,1)$ liegt ein isoliertes Minimum vor, und zwar ein globales, wie die Taylorentwicklung zeigt:

$$f(x_c+h, y_c+k) = f(x_c, y_c) + 0 + \frac{1}{2}(h,k)\begin{pmatrix} 2 & -\lambda \\ -\lambda & 2 \end{pmatrix}\begin{pmatrix} h \\ k \end{pmatrix} > f(x_c, y_c)$$
$$\forall \begin{pmatrix} h \\ k \end{pmatrix} \in \mathbb{R}^2 \setminus \{(0,0)\}$$

(ii) $\lambda = 2$. Hess $f(x,y)$ ist semidefinit. Unsere Kriterien erlauben keine Entscheidung, aber

$$f(x,y) = x^2 + y^2 - 2xy + 2(x-y) = (x-y)^2 + 2(x-y) + 1 - 1$$
$$= (x-y+1)^2 - 1,$$

und auf der kritischen Geraden $\{(x,y) : y = x+1\}$ wird der erste Summand $= 0$. Hier liegt also ein nicht isoliertes, globales Minimum vor.

(iii) $\lambda > 2$. Hess $f(x,y)$ ist indefinit, f besitzt keine Extrema.