Extrema multivariater Funktionen

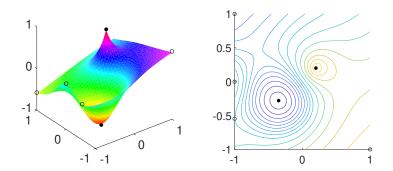
Lokale Extremstellen x_* (Minima oder Maxima) einer Funktion $f: \mathbb{R}^n \ni D \to \mathbb{R}$ können an folgenden Punkten auftreten:

- Unstetigkeitsstellen der partiellen Ableitungen,
- kritischen Punkten, d.h. Punkten mit grad $f(x_*) = (0, ..., 0)^t$,
- Randpunkten des Definitionsbereichs D.

Eine hinreichende Bedingung für ein lokales Minimum (Maximum) in einem kritischen Punkt x_* im Innern des Definitionsbereichs D ist, dass ebenfalls die zweiten partiellen Ableitungen stetig sind und alle Eigenwerte der Hesse-Matrix $Hf(x_*)$ positiv (negativ) sind. Gibt es Eigenwerte mit verschiedenen Vorzeichen, so handelt es sich um einen Sattelpunkt, also kein lokales Extremum. Ist mindestens ein Eigenwert Null bei gleichen Vorzeichen der von Null verschiedenen Eigenwerte, so kann der Typ des kritischen Punktes x_* anhand der zweiten Ableitungen nicht klassifiziert werden.

Für lokale Minima (Maxima) an Randpunkten des Definitionsbereichs D muss die Richtungsableitung $\partial_{\nu} f(x_*)$ für jede ins Innere von D zeigende Richtung $\geq 0 \ (\leq 0)$ sein.

Eine globale Extremstelle von f kann durch Vergleich der Funktionswerte an allen Frage kommenden Punkten ermittelt werden.



Die Abbildung illustriert die verschiedenen Möglichkeiten für eine bivariate Funktion. Dabei sind lokale Extrema durch Kreise und globale Extrema

durch Punkte gekennzeichnet. Extrema im Innern werden von Niveaulinien umschlossen. In glatten Randbereichen berührt die durch eine Extremstelle verlaufene Niveaulinie den Rand.

Die Randbetrachtung ist für multivariate Funktionen komplizierter als für Funktionen einer Variablen. Bei der Bestimmung globaler Extrema wird man deshalb versuchen, Randpunkte als Kandidaten auszuschließen. Existiert beispielsweise für einen beschränkten Definitionsbereich D ein Punkt $y \in D$ mit einem kleineren Funktionswert als das Minimum der Funktionswerte auf dem Rand von D, dann kommen nur Punkte im Innern von D als Minimalstellen in Betracht. Gibt es nur eine Unstetigkeisstelle einer partiellen Ableitung bzw. einen kritischen Punkt im Innern von D, so muss es sich um die Minimalstelle von f auf D handeln.

Für einen unbeschränkten Definitionsbereich muss man zusätzlich zeigen, dass f(x) > f(y) für |x| hinreichend groß, um die Existenz einer Minimalstelle im Innern von D zu folgern.

Bei der Bestimmung globaler Maxima verfährt man analog.

Beweis

betrachte die n-variate Funktion f entlang von Geraden, d.h. für einen beliebigen Vektor v mit |v|=1 die univariate Funktion

$$g(t) = f(x_* + tv), \quad t \in \mathbb{R}$$

Kettenregel ⇒

$$g'(t) = \sum_{k=1}^{n} \partial_k f(x_* + tv) v_k = \operatorname{grad} f(x_* + tv)^{\mathsf{t}} v$$

$$g''(t) = \sum_{i=1}^{n} \sum_{k=1}^{n} \partial_j \partial_k f(x_* + tv) v_j v_k = v^{\mathsf{t}} \operatorname{H} f(x_* + tv) v_j v_k$$

(i) Notwendige Bedingung für Extrema:

 x_* lokale Extremstelle von $f \implies t = 0$ lokale Extremstelle von g und folglich

$$0 = g'(0) = \operatorname{grad} f(x_*))^{\mathsf{t}} v = \partial_v f(x_*)$$

Richtung v beliebig \Longrightarrow grad $f(x_*) = (0, \dots, 0)^t$ Für ein lokales Minimum x_* am Rand folgt aus $g'(0) \ge 0$ für jede ins Innere von D zeigende Richtung v, dass $\partial_v f(x_*) \ge 0$ analoges Argument für ein lokales Maximum am Rand (ii) Hinreichende Bedingung: Eigenwerte von H $f(x_*)$ positiv \iff H $f(x_*)$ positiv definit \iff

$$v^{\mathrm{t}} H f(x_*) v \geq c > 0$$

Stetigkeit \Longrightarrow

Die Ungleichung bleibt für Hf(x) und $x = x_* + tv$ in einer Umgebung $U: t < \delta$ erhalten mit der kleineren Konstanten c/2 an Stelle von c.

Taylor-Approximation für $x \in U$ \Longrightarrow

$$\begin{split} f(x) &= g(t) = g(0) + \underbrace{g'(0)}_{0} t + \frac{1}{2} g''(\underbrace{s}_{\in [0,\delta)}) t^{2} \\ &= f(x_{*}) + \frac{1}{2} v^{t} H f(\underbrace{x + sv}_{\in U}) v t^{2} \ge f(x_{*}) + \frac{1}{2} t^{2} \frac{c}{2} > f(x_{*}) \end{split}$$

 \implies x_* lokale Minimalstelle von f in U analoges Argument für ein lokales Maximum im Innern

Beispiel

Extrema einer quadratischen Funktion

$$f(x) = \frac{1}{2}x^t Ax - x^t b + c, \quad A = A^t$$

Gradient und Hesse-Matrix

$$\operatorname{grad} f = Ax - b$$
, $\operatorname{H} f = A$

kritische Punkte x_* : Lösungen von Ax = bEigenwerte von A ausschließlich positiv oder negativ \implies $\det A \neq 0$ und Ax = b eindeutig lösbar

ightarrow genau ein lokales und damit ebenfalls globales Extremum von f

z.B.

$$f(x,y) = \frac{3}{2}x^2 + 2xy + \frac{3}{2}y^2 + x + 4y - 3$$
, $A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}$, $b = \begin{pmatrix} -1 \\ -4 \end{pmatrix}$

$$\operatorname{grad} f = (0,0)^{\operatorname{t}} \quad \sim$$

$$\left(\begin{array}{cc} 3 & 2 \\ 2 & 3 \end{array}\right) \left(\begin{array}{c} x \\ y \end{array}\right) = \left(\begin{array}{c} -1 \\ -4 \end{array}\right)$$

mit der Lösung
$$(x_*, y_*) = (1, -2)$$

det $A = 5$ und Spur $A = 6$ positiv

$$\implies$$
 Eigenwerte von A positiv

$$\implies$$
 $(1,-2)$ ist das globale Minimum von f

Beispiel

Bestimmung der globalen Extrema der Funktion

$$f(x,y) = \cos x + \cos y + \cos(x+y)$$

f ist 2π -periodisch bezüglich x und y und f(y,x)=f(x,y)=f(-x,-y) \implies Es genügt, den Bereich

$$D = (-\pi, \pi] \times [0, \pi]$$

zu untersuchen.

keine Randpunkte und Unstetigkeitstellen von partiellen Ableitungen

→ nur kritische Punkte relevant

grad
$$f = \begin{pmatrix} -\sin x - \sin(x+y) \\ -\sin y - \sin(x+y) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\iff$$

$$\sin x = -\sin(x+y) = \sin y$$

 $\sin x = \sin y \text{ mit } 0 \le y \le \pi \text{ und } -\pi < x \le \pi$ $\Rightarrow \text{ zwei m\"{o}gliche F\"{a}lle}$

$$x = y \quad \lor \quad y = \pi - x$$

betrachte für beide Fälle die zweite Gleichung $\sin x = -\sin(x+y)$ (i) x = y:

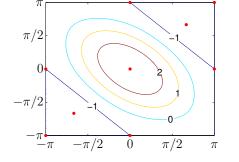
$$\sin x = -\sin(2x) = -2\sin x \cos x \quad \Longleftrightarrow \quad (1/2 + \cos x)\sin x = 0$$

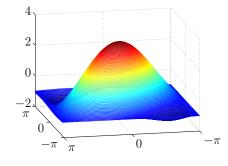
 \leadsto kritische Punkte (0,0), (π , π) und ($2\pi/3,2\pi/3$) im betrachteten Bereich

(ii)
$$y = \pi - x$$
:

$$\sin(x) = -\sin(\pi) = 0$$

 \rightsquigarrow kritische Punkte $(0,\pi)$ und $(\pi,0)$ im betrachteten Bereich





Vergleich der Funktionswerte \implies globales Maximum mit Wert f(0,0)=3 bei (0,0) globales Minimum mit Wert $f(2\pi/3,2\pi/3)=-3/2$ bei $(2\pi/3,2\pi/3)$

Typbestimmung der anderen kritischen Punkte mit Hilfe der Hesse Matrix

$$Hf = \begin{pmatrix} -\cos x - \cos(x+y) & -\cos(x+y) \\ -\cos(x+y) & -\cos y - \cos(x+y) \end{pmatrix}$$

Sattelpunkt bei (π, π) , da

$$\mathsf{H}\,f(\pi,\pi)=\left(egin{array}{cc} 0 & -1 \ -1 & 0 \end{array}
ight),\quad \mathsf{det}\,\mathsf{H}\,f=-1<0$$

Sattelpunkte bei $(\pi,0)$ und $(0,\pi)$ (symmetrische Lage), da

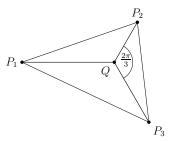
$$\mathsf{H}\,f(\pi,0)=\left(egin{array}{cc} 2 & 1 \ 1 & 0 \end{array}
ight),\quad \mathsf{H}\,f(0,\pi)=\left(egin{array}{cc} 0 & 1 \ 1 & 2 \end{array}
ight),\quad \det\mathsf{H}\,f=-1<0$$

Punktsymmetrie bzgl. (0,0) und Periodizität \sim globale Maxima bei $(2k\pi,2\ell\pi)$, globale Minima bei $(2\pi/3+2k\pi,2\pi/3+2\ell\pi)$, $(-2\pi/3+2k\pi,-2\pi/3+2\ell\pi)$, Sattelpunkte bei $(k\pi,\ell\pi)$, jeweils mit $k,\ell\in\mathbb{Z}$

Beispiel

Steiners Problem: Bestimme den Punkt $Q \in \mathbb{R}^2$, so dass die Summe der Abstände zu vorgegebenen Punkten P_k minimal wird.

Spezialfall dreier Punkte P_1 , P_2 , $P_3 \leadsto Q$ Randpunkt des Dreiecks $[P_1, P_2, P_3]$ oder $\sphericalangle(P_k, Q, P_\ell) = 2\pi/3 \ \forall k, \ell$



zeige die Charakterisierung für einen inneren Punkt: d_k : Abstand zwischen Q = (x, y) und $P_k = (x_k, y_k)$, d.h.

$$d_k^2 = (x - x_k)^2 + (y - y_k)^2$$

Kettenregel ⇒

$$2d_k \frac{\partial d_k}{\partial x} = 2(x - x_k)$$

bzw.

$$\frac{\partial d_k}{\partial x} = \frac{x - x_k}{d_k}$$

und entsprechend $\partial d_k/\partial y=(y-y_k)/d_k$

$$\Longrightarrow$$

$$\operatorname{grad} d_k = \frac{1}{d_k} \left(\begin{array}{c} x - x_k \\ y - y_k \end{array} \right), \quad d_k = \left| \left(\begin{array}{c} x - x_k \\ y - y_k \end{array} \right) \right|$$

Minimum von

$$f = d_1 + d_2 + d_3$$

bei innerem Punkt $Q \implies$

$$\operatorname{grad} f(Q) = \sum_{k} \operatorname{grad} d_{k} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
,

da f in einer Umgebung von Q stetig differenzierbar ist $|\operatorname{grad} d_k| = 1 \implies$

Gradienten bilden gleichseitiges Dreieck (Vektorsumme null)

 \leadsto Winkel $2\pi/3$ zwischen den Gradientenrichtungen