Extrema I

Einleitung

Oft interessieren wir uns für *Extrema*[1] einer Funktion $f\colon \Omega \to \mathbb{R}$. Unter einem *lokalen Extremum* verstehen wir einen Punkt $x_0 \in \Omega$, der entweder ein *lokales Minimum* oder *lokales Maximum* ist. Wir sprechen von einem lokalen Minimum (bzw. Maximum), falls eine Umgebung[2] $U \subseteq \Omega$ von x_0 existiert, sodass $f(x_0) \leq f(x)$ (bzw. $f(x_0) \geq f(x)$) für alle $x \in U$ gilt. Unter einem *globalen Extremum* verstehen wir einen Punkt $x_0 \in \Omega$, der entweder ein *globales Minimum* oder *globales Maximum* ist. Wir sprechen von einem globalen Minimum (bzw. Maximum), falls $f(x_0) \leq f(x)$ (bzw. $f(x_0) \geq f(x)$) für alle $x \in \Omega$ gilt. Extrema (lokale oder globale) müssen weder existieren noch eindeutig sein.

Hier wollen wir uns mit einer ganz bestimmten Situation beschäftigen. Sei $\Omega \subseteq \mathbb{R}^n$ offen und $f \in C^2(\Omega, \mathbb{R})$. Unter diesen Voraussetzungen haben wir bestimmte Werkzeuge zur Verfügung, die uns beim Finden von Extrema helfen können. Hierbei ist es aber wichtig sich vor Augen zu halten, dass diese Werkzeuge nicht immer erfolgreich sind. In manchen Fällen liefern sie uns keine Möglichkeit zu entscheiden, ob ein Punkt ein Extremum ist oder nicht. Das bedeutet aber noch lange nicht, dass es überhaupt nicht entschieden werden kann. In solchen Fällen müssen wir dann auf ad hoc Methoden zurückgreifen, die an unser Problem angepasst sind.

Sind wir nicht in obiger Situation, sind also die Voraussetzung nicht erfüllt, so können wir die beschriebenen Algorithmen auch nicht anwenden (zumindest nicht direkt). Wir müssen dann auf andere generelle oder ad hoc Methoden zurückgreifen.

Bevor wir zu den eigentlichen Algorithmen kommen, wollen wir noch den Begriff der *Definitheit* einer (symmetrischen) Matrix $A \in \mathbb{R}^{n\times n}$ besprechen. Wir sagen A ist *positiv definit* (bzw. *negativ definit*), falls $v^T Av > 0$ (bzw. $v^T Av < 0$) für alle $v \in \mathbb{R}^n \setminus \{0\}$ gilt. Wir sagen A ist *positiv semidefinit* (bzw. *negativ semidefinit*), falls $v^T Av \geq 0$ (bzw. $v^T Av \leq 0$) für alle $v \in \mathbb{R}^n$ gilt. Wir sagen A ist *echt semidefinit*[3], falls A entweder positiv semidefinit aber nicht positiv definit ist oder negativ semidefinit aber nicht negativ definit ist. Wir sagen A ist indefinit, falls es sowohl ein $v \in \mathbb{R}^n$ gibt, sodass $v^T Av > 0$, als auch ein $v \in \mathbb{R}^n$ gibt, sodass $v^T Av < 0$. Jede Matrix hat also genau eine der Eigenschaften: positiv definit, negativ definit, echt semidefinit oder indefinit. Wie die Definitheit einer Matrix ermittelt werden kann, wird in der Linearen Algebra besprochen.

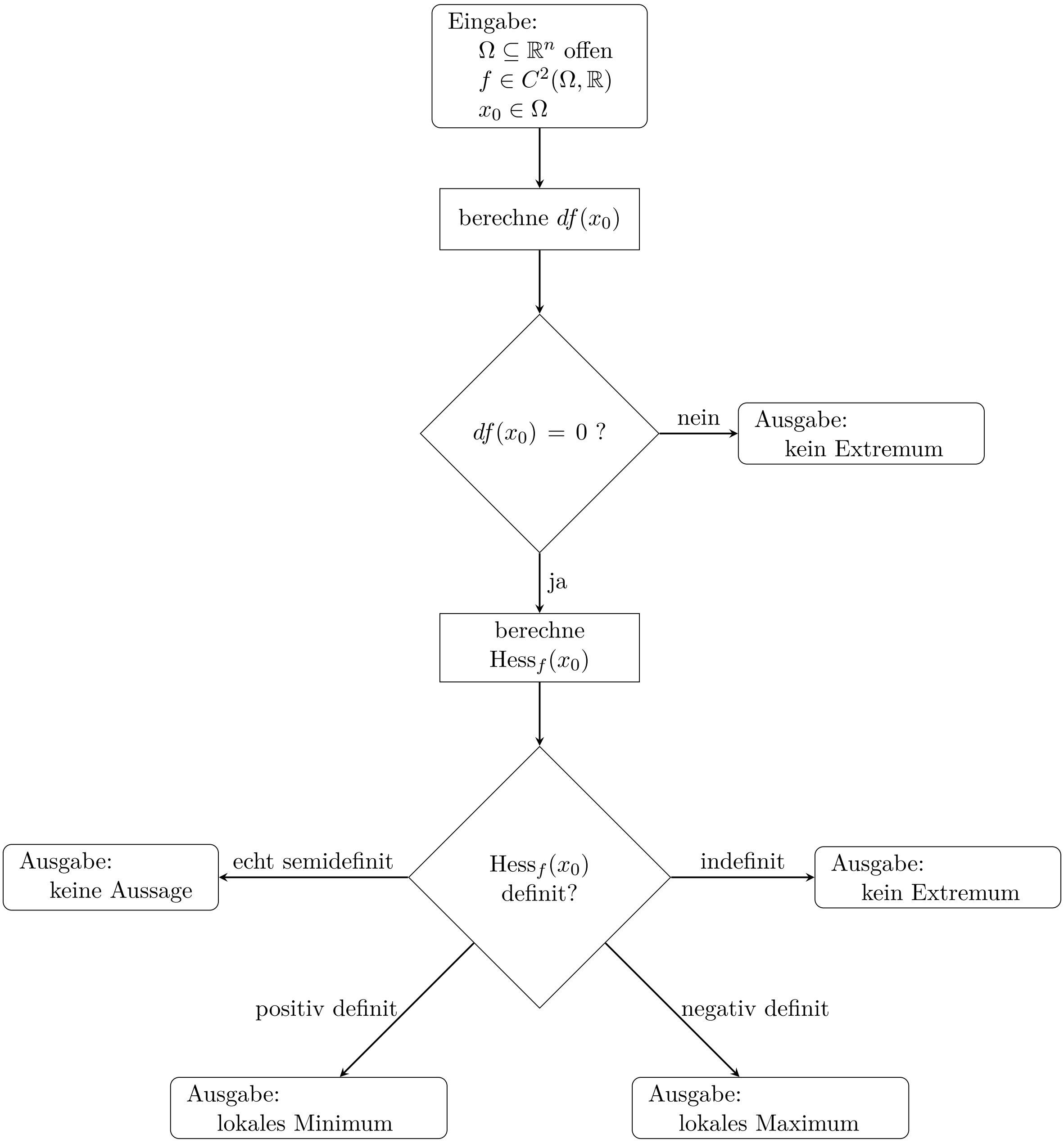
Es folgen nun zwei verwandte Algorithmen die uns als Werkzeuge zum Finden von Extrema dienen. Danach erläutern wir diese Algorithmen an Beispielen.

[1] Manchmal wird sprachlich zwischen „Extremumsstelle“ und „Extremum“ (im Sinne von dem Wert an dieser Stelle) unterschieden. Wir verzichten darauf.

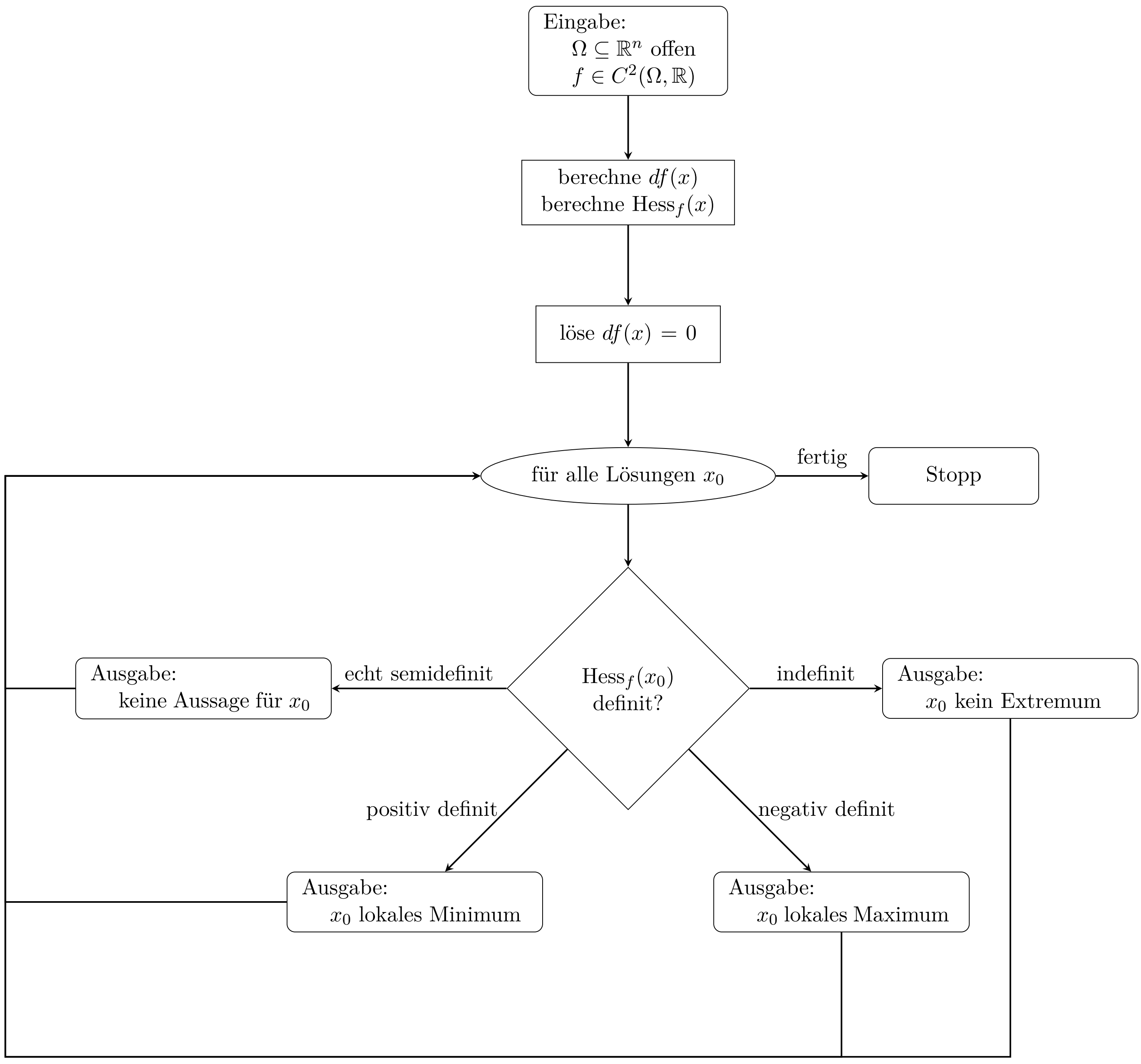
[2] Eine Menge $U \subseteq \Omega$ ist eine Umgebung von x_0, falls x_0 ein innerer Punkt von U ist.

[3] Diese Begriffsbildung ist wenig gebräuchlich, aber hier praktisch.

Algorithmus *„untersuche einen Punkt“*



Algorithmus *„untersuche alle Punkte“*



Erläuterungen

Beide Algorithmen beruhen auf dem selben Prinzip. Der zweite entsteht im Wesentlichen durch Anwendung des ersten auf alle Punkte gleichzeitig. Wir sprechen daher im Folgenden nur von einem Algorithmus.

Beispiel: Gegeben $f\colon [-1,1]^2 \to \mathbb{R}$. Wir können unseren Algorithmus nicht direkt anwenden, da $[-1,1]^2$ nicht offen ist. Sofern $f|_{(-1,1)^2}$ zweimal stetig differenzierbar ist, können wir so zumindest nach Extrema im Inneren suchen. Über mögliche Extrema am Rand erfahren wir nichts.

Beispiel: Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x,y) = (x+y)^4.$$

Offensichtlich ist $(0,0)$ ein globales Minimum. Wie wir aber gleich sehen werden, scheitert unser Algorithmus. Wir berechnen

$$df(0,0) = 0$$

und

$$\mathrm{Hess}_f(0,0) = 0.$$

Wir sehen, dass $\mathrm{Hess}_f(0,0)$ echt semidefinit ist. Der Algorithmus liefert also keine Aussage für $(0,0)$.

Beispiel: Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x,y) = (1+x)y.$$

Wir wollen nur den Punkt $(0,0)$ untersuchen und berechnen

$$df(0,0) = (0,1).$$

Es handelt sich also um kein Extremum.

Beispiel: Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch

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

Es ist leicht zu sehen, dass f zweimal stetig differenzierbar ist. Wir wollen nur den Punkt $(0,0)$ untersuchen und berechnen

$$df(0,0) = 0.$$

Es handelt sich also um einen Kandidaten für ein Extremum. Wir berechnen weiter

$$\mathrm{Hess}_f(0,0) = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}.$$

Da diese Matrix indefinit ist, ist $(0,0)$ kein Extremum.

Beispiel: Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x,y) = e^{(-x^2)} - xy - y^2.$$

Wir berechnen

$$df(x,y) = (-2xe^{(-x^2)} - y, -x-2y)$$

und

$$\mathrm{Hess}_f(x,y) = \begin{pmatrix} (-2+4x^2)e^{(-x^2)} & -1 \\ -1 & -2 \end{pmatrix}.$$

Wir erhalten

$$df(x,y) = 0 \Leftrightarrow (x,y) = (0,0) \text{ oder } (x,y) = \sqrt{\ln(4)} \cdot \left(1, -\frac{1}{2}\right).$$

Da

$$\mathrm{Hess}_f(0,0) = \begin{pmatrix} -2 & -1 \\ -1 & -2 \end{pmatrix}$$

negativ definit ist, ist $(0,0)$ ein lokales Maximum. Da

$$\mathrm{Hess}_f\left(\sqrt{\ln(4)} \cdot \left(1, -\frac{1}{2}\right)\right) = \begin{pmatrix} \left(-\frac{1}{2} + \ln(4)\right) & -1 \\ -1 & -2 \end{pmatrix}$$

indefinit ist, ist $\sqrt{\ln(4)} \cdot \left(1, -\frac{1}{2}\right)$ kein Extremum.

Insbesondere existieren keine Minima (lokal oder global). Um zu sehen, dass $(0,0)$ kein globales Maximum ist, müssen wir noch eine ad hoc Methode verwenden. Da

$$\lim_{x\to\infty} f(x,-1) = \infty$$

ist, existiert kein globales Maximum.

Beispiel: Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x,y) = \begin{cases} \sin\left(\frac{1}{x+y}\right)(x+y) + x^2 & \text{für } y \neq -x, \\ x^2 & \text{für } y = -x. \end{cases}$$

Die Funktion f ist stetig, aber nur C^2 auf $\{(x,y) \in \mathbb{R}^2 : y \neq -x\}$. Deswegen suchen wir dort zuerst mittels unseres Algorithmus nach Extrema. Sei also $(x,y) \in \mathbb{R}^2$ mit $y \neq -x$. Wir berechnen

$$df(x,y) = \left[\sin\left(\frac{1}{x+y}\right) - \cos\left(\frac{1}{x+y}\right)\frac{1}{x+y}\right] \cdot (1,1) + (2x,0)$$

und

$$\mathrm{Hess}_f(x,y) = -\sin\left(\frac{1}{x+y}\right)\frac{1}{(x+y)^3} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Wir erhalten

$$df(x,y) = 0 \Leftrightarrow x = 0 \text{ und } \tan\left(\frac{1}{y}\right) = \frac{1}{y}.$$

Sei y_0 eine Lösung von $\tan\left(\frac{1}{y}\right) = \frac{1}{y}$. Für

$$y_0 \in \pm\left(\left(0, \frac{\pi}{2}\right) + 2\pi\mathbb{N}_0\right)^{-1}$$

ist $\mathrm{Hess}_f(0, y_0)$ indefinit und $(0, y_0)$ daher kein Extremum. Für

$$y_0 \in \pm\left(\left(\pi, \frac{3\pi}{2}\right) + 2\pi\mathbb{N}_0\right)^{-1}$$

ist $\mathrm{Hess}_f(0, y_0)$ positiv definit und $(0, y_0)$ daher ein lokales Minimum. Wir wollen noch mittels einer ad hoc Methode zeigen, dass es sogar ein globales Minimum gibt. Unter all diesen Minima wird der kleinste Wert für

$$y_0^* \in \pm\left(\pi, \frac{3\pi}{2}\right)^{-1}$$

angenommen. Da $f(0, \pm y_0^*) < 0$,

$$\lim_{(x,y)\to 0} f(x,y) = 0 \quad \text{und} \quad \liminf_{(x,y)\to\infty} = 1,$$

sind die beiden Punkte $(0, \pm y_0^*)$ globale Minima.

Es bleiben noch die Punkte $(x,y) \in \mathbb{R}^2$ mit $y = -x$ zu untersuchen. Hier können wir unseren Algorithmus nicht anwenden. Betrachten wir f entlang der Geraden $y = -x$, so ist offensichtlich, dass nur der Punkt $(0,0)$ als weiteres Extremum in Frage kommt. Allerdings existieren in jeder Umgebung von $(0,0)$ sowohl Punkte, an denen f positive Werte annimmt, als auch Punkte, an denen f negative Werte annimmt. Der Punkt $(0,0)$ ist daher kein Extremum.

Abschliessend merken wir noch an, dass es keine lokalen Maxima gibt und damit auch kein globales.