Extrema II

Einleitung

Gegeben sei eine zweimal stetig differenzierbare Funktion $f\colon \Omega \to \mathbb{R}$, wobei $\Omega \subseteq \mathbb{R}^n$ offen ist. Im jetzigen Teil wollen wir nicht die Extrema von f (auf ganz Ω) sondern die Extrema von $f|_S$ studieren, wobei $S := g^{-1}(0)$ für eine zweimal stetig differenzierbare Funktion $g\colon \Omega \to \mathbb{R}^l$ mit $l < n$. Zum Beispiel ist $x_0 \in S$ per Definition ein lokales Minimum, falls eine Umgebung U von x_0 existiert, sodass $f(x_0) \le f(x)$ für alle $x \in S \cap U$. Man beachte, dass die Extrema von $f|_S$ nur von dem Verhalten von f auf S abhängen und nicht von dem Verhalten auf $\Omega \setminus S$. Auch die Funktion g spielt nur insofern eine Rolle, als dass sie die Menge S angibt. Allerdings ist es mit diesen Annahmen — dass f auf ganz Ω definiert ist und dass $S = g^{-1}(0)$ — einfacher konkrete Aussagen zu machen. Man spricht in diesem Zusammenhang von Extrema von f unter der *Nebenbedingung* g.

Die im Folgenden vorgestellte Methode basiert auf der *Lagrange-Multiplikatorenregel*. Bevor wir zu den Details kommen, lohnt es sich zuerst ein paar Begriffe zu besprechen. Im ersten Schritt wird geprüft, ob g am zu untersuchenden Punkt $x_0 \in \mathbb{R}^n$ regulär ist. Das bedeutet, dass $dg(x_0)$ vollen Rang hat, in diesem Fall also Rang l. Falls ja, muss als nächster Schritt die Gleichung

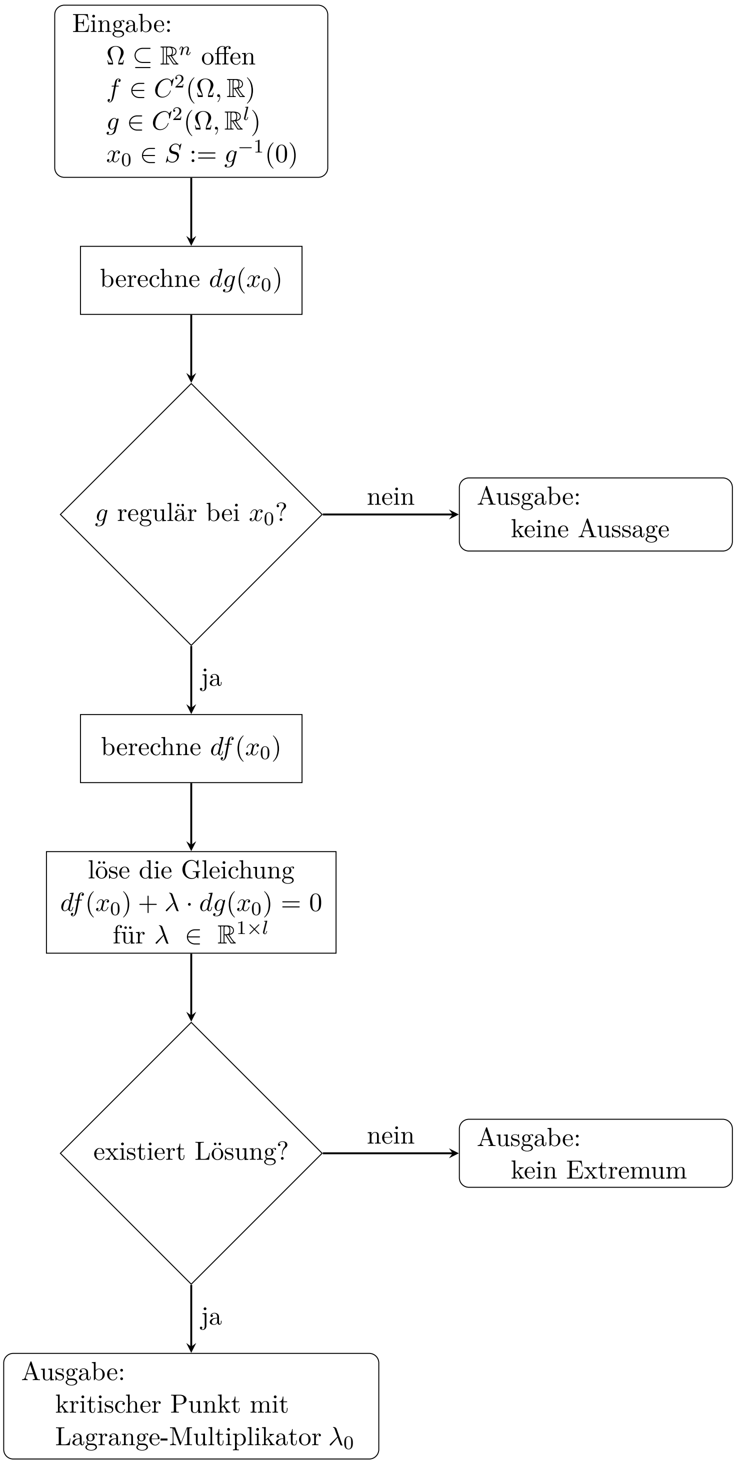
$$df(x_0) + \lambda \cdot dg(x_0) = 0$$

für $\lambda \in \mathbb{R}^{1\times l}$ gelöst werden. Da $dg(x_0)$ Rang l hat, existiert entweder keine oder genau eine Lösung. Gibt es eine Lösung λ_0, genannt Lagrange-Multiplikator, dann ist x_0 ein kritischer Punkt. Zuletzt wird dann ähnlich wie ohne Nebenbedingung die Definitheit von $H := \mathrm{Hess}_{f+\lambda_0\cdot g}(x_0) \in \mathbb{R}^{n\times n}$ untersucht. Allerdings interessieren wir uns jetzt nur für Richtungen im Tangentialraum von S bei x_0. Der Tangentialraum ist in unserer Situation einfach $T_{x_0}S = \ker(dg(x_0))$. Wir unterscheiden folgende Eigenschaften:

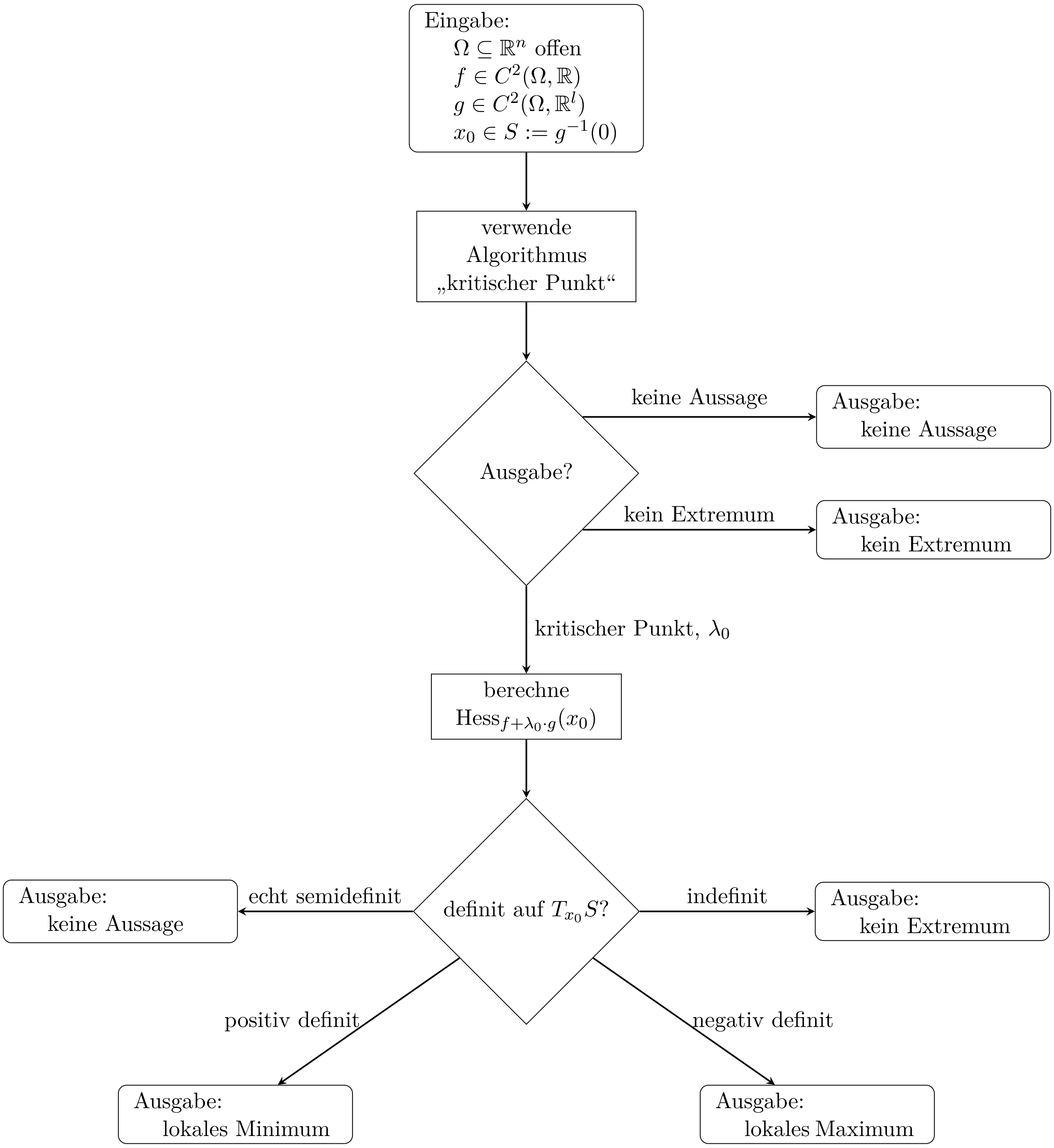
- H positiv definit auf $T_{x_0}S \Leftrightarrow v^tHv > 0$ für alle $v \in T_{x_0}S$
- H negativ definit auf $T_{x_0}S \Leftrightarrow v^tHv < 0$ für alle $v \in T_{x_0}S$
- H positiv semidefinit auf $T_{x_0}S \Leftrightarrow v^tHv \ge 0$ für alle $v \in T_{x_0}S$
- H negativ semidefinit auf $T_{x_0}S \Leftrightarrow v^tHv \le 0$ für alle $v \in T_{x_0}S$
- H echt semidefinit auf $T_{x_0}S \Leftrightarrow H$ (positiv oder negativ) semidefinit aber nicht (positiv oder negativ) definit auf $T_{x_0}S$
- H indefinit auf $T_{x_0}S \Leftrightarrow H$ weder positiv noch negativ semidefinit auf $T_{x_0}S$

Die Matrix H hat immer genau eine der Eigenschaften positiv definit, negativ definit, echt semidefinit und indefinit auf $T_{x_0}S$. Sei B eine Basis von $T_{x_0}S$ in Matrixform. Die Definitheit von H auf $T_{x_0}S$ stimmt mit der gewöhnlichen Definitheit von $B^tHB \in \mathbb{R}^{l\times l}$ überein und kann so leicht ermittelt werden.

Algorithmus „kritischer Punkt von f unter der Nebenbedingung g“



Algorithmus „Extremum von f unter der Nebenbedingung g“



Erläuterungen

Die obigen Algorithmen sind jeweils für einen einzelnen Punkt beschrieben. In der Praxis ist es natürlich meist sinnvoll möglichst viele Punkte gleichzeitig zu behandeln. Sofern zum Beispiel g bei allen $x \in S$ regulär ist, kann man direkt versuchen die Gleichung

$$df(x) + \lambda \cdot dg(x) = 0$$

für $x \in S$ und $\lambda \in \mathbb{R}^{1\times l}$ zu lösen. Dies ist gleichbedeutend damit, das System

$$\begin{aligned} df(x) + \lambda \cdot dg(x) &= 0 \\ g(x) &= 0 \end{aligned}$$

für $x \in \Omega$ und $\lambda \in \mathbb{R}^{1\times l}$ zu lösen.

Beispiel: Seien $f, g\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$f(x, y) = x^3$$

und

$$g(x, y) = x + y - 1$$

für alle $(x, y) \in \mathbb{R}^2$. Wir suchen die Extrema von f unter der Nebenbedingung g. Wir berechnen

$$dg(x, y) = (1, 1)$$

und sehen, dass g für alle Punkte in S regulär ist. Nun betrachten wir die Gleichung

$$0 = df(x, y) + \lambda dg(x, y) = (3x^2, 0) + \lambda(1, 1)$$

für $(x, y) \in S$ und $\lambda \in \mathbb{R}$. Da es keine Lösung (für $(x, y) \in S$) gibt, hat f keine Extrema unter der Nebenbedingung g.

Beispiel: Sei $f\colon \mathbb{R}^3 \to \mathbb{R}$ definiert durch

$$f(x, y, z) = x + z$$

für alle $(x, y, z) \in \mathbb{R}^3$. Sei $g\colon \mathbb{R}^3 \to \mathbb{R}^2$ definiert durch

$$g(x, y, z) = \begin{pmatrix} x^2 + y^2 - 1 \\ xy - z \end{pmatrix}$$

für alle $(x, y, z) \in \mathbb{R}^3$. Wir suchen die Extrema von f unter der Nebenbedingung g. Wir berechnen

$$dg(x, y, z) = \begin{pmatrix} 2x & 2y & 0 \\ y & x & -1 \end{pmatrix}$$

und sehen wieder, dass g für alle Punkte in S regulär ist. Nun betrachten wir die Gleichung

$$0 = df(x,y,z) + (\lambda,\mu)\cdot dg(x,y,z) = (1,0,1) + (\lambda,\mu)\cdot \begin{pmatrix} 2x & 2y & 0 \\ y & x & -1 \end{pmatrix}$$

für $(x,y,z) \in S$ und $(\lambda,\mu) \in \mathbb{R}^{1\times 2}$. Es ergeben sich die Lösungen

$$\begin{aligned} (x,y,z) &= (0,-1,0) \text{ und } (\lambda,\mu) = (0,1), \\ (x,y,z) &= \left(\frac{\sqrt{3}}{2}, \frac{1}{2}, \frac{\sqrt{3}}{4}\right) \text{ und } (\lambda,\mu) = \left(-\frac{\sqrt{3}}{2}, 1\right), \\ (x,y,z) &= \left(-\frac{\sqrt{3}}{2}, \frac{1}{2}, -\frac{\sqrt{3}}{4}\right) \text{ und } (\lambda,\mu) = \left(\frac{\sqrt{3}}{2}, 1\right). \end{aligned}$$

Um zu entscheiden welche dieser kritischen Punkte tatsächlich Extrema sind, berechnen wir

$$\mathrm{Hess}_{f+(\lambda,\mu)\cdot g}(x,y,z) = \begin{pmatrix} 2\lambda & \mu & 0 \\ \mu & 2\lambda & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

und

$$T_{(x,y,z)}S = \ker dg(x,y,z) = \operatorname{span}\begin{pmatrix} y \\ -x \\ y^2 - x^2 \end{pmatrix}.$$

Für die Definitheit der Hesse-Matrix auf dem Tangentialraum, betrachten wir also

$$\begin{pmatrix} y & -x & y^2 - x^2 \end{pmatrix} \cdot \begin{pmatrix} 2\lambda & \mu & 0 \\ \mu & 2\lambda & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} y \\ -x \\ y^2 - x^2 \end{pmatrix} = 2\lambda(x^2+y^2) - 2\mu xy = 2\lambda - 2z.$$

Im letzten Schritt haben wir verwendet, dass wir uns nur für Punkte in S interessieren. Nun können wir leicht ablesen, dass der Punkt $\left(\frac{\sqrt{3}}{2}, \frac{1}{2}, \frac{\sqrt{3}}{4}\right)$ ein lokales Maximum ist und dass der Punkt $\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}, -\frac{\sqrt{3}}{4}\right)$ ein lokales Minimum ist. Für den Punkt $(0,-1,0)$ erhalten wir keine Aussage. Für die weitere Analyse müssen wir also auf ad hoc Argumente zurückgreifen.

Die Punkte $(\pm\sqrt{\epsilon}, -\sqrt{1-\epsilon}, \mp\sqrt{\epsilon-\epsilon^2})$ für $\epsilon > 0$ klein liegen beliebig nahe bei $(0,-1,0)$ und in S. Da

$$f(\pm\sqrt{\epsilon}, -\sqrt{1-\epsilon}, \mp\sqrt{\epsilon-\epsilon^2}) = \pm\sqrt{\epsilon}\,(1-\sqrt{1-\epsilon})$$

ist, kann der Punkt $(0,-1,0)$ kein lokales Extremum sein. Da S kompakt und f stetig ist, muss ein globales Maximum bzw. Minimum existieren. Offensichtlich ist dies dann der Punkt $\left(\frac{\sqrt{3}}{2}, \frac{1}{2}, \frac{\sqrt{3}}{4}\right)$ bzw. $\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}, -\frac{\sqrt{3}}{4}\right)$.

Beispiel: Gegeben $f\colon \mathbb{R}^2 \to \mathbb{R}$ zweimal stetig differenzierbar. Sei $g\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch

$$g(x, y) = (x - y)^2$$

für alle $(x, y) \in \mathbb{R}^2$. Wir berechnen

$$dg(x, y) = 2(x - y) \cdot (1, -1)$$

und sehen, dass dg an allen Punkten mit $x = y$ verschwindet. An diesen Punkten liefert der Algorithmus also keine Aussage.