

Proseminar Mathematik

Johann Seidl

17. Oktober 2019

Inhaltsverzeichnis

Optimalitätskriterien
Johann Seidl

1 Einführung

In der heutigen Welt sind viele mathematische Probleme auf Real-World-Probleme zurückzuführen. Unser unter (??) verstandenes *Minimierungsproblem* spielt unter anderem Rollen in den Ingenieurwissenschaften, der Physik, der Medizin oder der Ökonomie.

1.1 Minimierungsproblem

Unter einem Minimierungsproblem wird folgende Aufgabe verstanden.

Sei $X \subseteq \mathbb{R}^n$ und eine Funktion $f : X \rightarrow \mathbb{R}$, dann wird ein $x^* \in X$ mit folgender Eigenschaft gesucht

$$f(x^*) \leq f(x) \quad \forall x \in X \quad (1)$$

In kurzer Notation lautet diese Aufgabe $\min f(x)$ unter der Nebenbedingung $x \in X$. Ist $X = \mathbb{R}^n$, dann spricht man von unrestringierten Problemen, andernfalls von restringierten Problemen.

1.2 Optimalitätskriterien

In diesem Kapitel gehen wir auf notwendige und hinreichende Bedingungen für lokale Minima ein, welche Ableitungen benutzen. Der erste Satz bezieht ein notwendiges Kriterium erster Ordnung, wobei man von einem Kriterium erster Ordnung spricht, da nur Informationen über die erste Ableitung eingehen.

Seien $X \subseteq^n$ eine offene Menge und $f : X \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion. Ist $x^* \in X$ ein lokales Minimum von f (auf X), so gilt

$$\nabla f(x^*) = 0$$

d.h., x^* ist ein stationärer Punkt von f . Sei $x^* \in X$ ein lokales Minimum von f , jedoch $\nabla f(x^*) \neq 0$. Dann existiert ein Vektor $d \in^n$ mit

$$\nabla f(x^*)d < 0$$

Eine mögliche Wahl für d kann zum Beispiel $d = -\nabla f(x^*)$ sein. Da f nach Voraussetzung stetig differenzierbar ist, gilt für die Richtungsableitung $f'(x^*; d)$ von f in x^* in Richtung d :

$$f'(x^*; d) = \lim_{t \rightarrow 0+} \frac{f(x^* + td) - f(x^*)}{t} = \nabla f(x^*)d < 0.$$

Folglich gibt es ein $t > 0$ mit $x^* + td \in X$ und

$$\frac{f(x^* + td) - f(x^*)}{t} < 0 \quad \forall t \in (0, \bar{t}]$$

Somit ist

$$f(x^* + td) < f(x^*) \quad \forall t \in (0, \bar{t}]$$

Dies steht aber im Widerspruch zur Voraussetzung, dass x^* ein lokales Minimum von f ist.

Offensichtlich ist dieser Satz keine hinreichende Bedingung, da auch ein mögliches Maxima die oben genannten Voraussetzungen erfüllt.

Eine Matrix ist positiv semidefinit, falls für die symmetrische $n \times n$ Matrix M gilt

$$xMx \geq 0 \quad \forall x \in^n \setminus 0 \quad (2)$$

[Satz von Taylor] Seien $f : \mathbb{R}^n \rightarrow \mathbb{R}$ zweimal stetig differenzierbar sowie $x, y \in^n$ gegeben. Dann existiert ein $\theta \in (0, 1)$ mit

$$f(x) = f(y) + \nabla f(y)(x - y) + \frac{1}{2}(x - y)\nabla^2 f(\xi)(x - y) \quad (3)$$

für $\xi = y + \theta(x - y)$.

Betrachten wir nun einen Satz, welcher eine notwendige Bedingung zweiter Ordnung nutzt. Man spricht wieder von Bedingung zweiter Ordnung, da der Satz Informationen über die zweite Ableitung enthält.

Seien $X \subseteq^n$ eine offene Menge und $f : X \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Ist $x^* \in X$ ein lokales Minimum von f (auf X), so ist die Hesse-Matrix $\nabla^2 f(x^*)$ positiv semidefinit. Sei $x^* \in X$ ein lokales Minimum von f , jedoch $\nabla^2 f(x^*)$

nicht positiv semidefinit (??). Dann existiert ein Vektor $d \in^n$ mit

$$d\nabla^2 f(x^*)d < 0 \quad (4)$$

Nutzt man nun (??) ergibt sich mit Satz 2.1 für alle hinreichend kleinen $t > 0$:

$$\begin{aligned} f(x^* + td) &= f(x^*) + \nabla f(x^*)(td) + \frac{1}{2}(td)\nabla^2 f(\xi_t)(td) \\ &= f(x^*) + \frac{1}{2}t^2 d\nabla^2 f(\xi_t)d \end{aligned}$$

Dabei ist $\nabla f(x^*)(td) = 0$, da x^* schon das lokale Minimum ist. Des Weiteren ist $\xi_t = x^* + v_t td$ für ein v_t mit $0 < v_t < 1$. Aus Stetigkeitsgründen folgt hieraus unter Verwendung von (??) die Existenz eines $\bar{t} > 0$ mit

$$f(x^* + td) < f(x^*) \quad \forall t \in (0, \bar{t}]$$

Dies steht aber im Widerspruch dazu, dass x^* ein lokales Minimum ist.

Auch die Bedingungen aus den Sätzen 2.1 und 2.5 zusammen sind *nicht* hinreichend dafür, dass x^* ein lokales Minimum ist.

Betrachte $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x) = x_1^2 - x_2^4$, $x^* = (0, 0)$. Die Ableitung erster Ordnung und die Hesse-Matrix der Funktion sind gegeben durch:

$$\begin{aligned} f'(x) &= (2x_1, -4x_2^3) \\ f''(x) &= \begin{pmatrix} 2 & 0 \\ 0 & -12x_2^2 \end{pmatrix} \end{aligned}$$

Man kann leicht erkennen, dass $f'(x^*) = 0$ und die Eigenwerte der Matrix im Punkt x^* $\lambda_1 = 0$ und $\lambda_2 = -2$ sind, also die Matrix positiv semidefinit ist. Die Voraussetzungen für Satz 2.5 sind also erfüllt. Dennoch ist zum Beispiel für $X = \mathbb{R}^2$ $f(0, 1) < f(0, 0)$.

Seien $X \subseteq^n$ offen und $f: X \rightarrow \mathbb{R}$ zweimal stetig differenzierbar. Gelten

[label=(), nolistsep, topsep=-] $\nabla f(x^*) = 0$ und $\nabla^2 f(x^*)$ ist positiv definit,

so ist x^* ein striktes lokales Minimum von f (auf X)

Aus (b) folgt zunächst die Existenz einer Konstanten ϵ mit

$$d\nabla^2 f(x^*)d \geq \epsilon dd \quad \forall d \in^n$$

Nach (??) gilt für alle hinreichend nahe bei Null gelegenen $d \in^n$:

$$f(x^* + d) = f(x^*) + \nabla f(x^*)(d) + \frac{1}{2}d\nabla^2 f(\xi_d)d \quad (5)$$

mit $\xi_d = x^* + v_d d$ für ein v_d mit $0 < v_d < 1$. Durch Voraussetzung (a) erhalten wir $\nabla f(x^*)(d) = 0$. Wenden wir nun die Cauchy - Schwarz - Ungleichung auf (5) an,

erhalten wir

$$\begin{aligned} f(x^* + d) &= f(x^*) + \frac{1}{2}d\nabla^2 f(x^*)d + \frac{1}{2}d(\nabla^2 f(\xi_d) - \nabla^2 f(x^*))d \\ &\geq f(x^*) + \frac{1}{2}(\epsilon - \|\nabla^2 f(\xi_d) - \nabla^2 f(x^*)\|)\|d\|^2 \end{aligned}$$

Folglich gilt

$$f(x^* + d) > f(x^*) \quad \forall d \neq 0 \text{ und hinreichend nahe } 0$$

Also ist x^* ein striktes lokales Minimum von f .

Dieser Satz liefert ein hinreichendes Kriterium. Man beachte jedoch, dass die Bedingungen (a) und (b) aus Satz 2.7 nicht notwendig für die strikte lokale Minimalität von x^* sind. Genauere Erklärungen liefert nachfolgendes Beispiel.

Betrachte $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x) = x_1^2 + x_2^4$, $x^* = (0, 0)$. Dann ist

$$\begin{aligned} f'(x) &= (2x_1, 4x_2^3) \\ f''(x) &= \begin{pmatrix} 2 & 0 \\ 0 & 12x_2 \end{pmatrix} \end{aligned}$$

so stellt man fest, dass $\nabla^2 f(x^*)$ positiv semidefinit ist und somit Satz 2.7 keine Anwendung finden kann (Voraussetzung (b) verletzt). Man erkennt jedoch leicht, dass $(0, 0)$ das globale Minimum ist.