E. Herberg, M. Hashemi, V. Stein
Universität Heidelberg

Übung 2

Ausgabedatum: 23. Oktober 2023
Abgabedatum: 29. Oktober 2023

Hausaufgabe 2.1 (Untersuchung von Optimalstellen)

Bestimmen Sie für die unten stehenden Minimierungsaufgaben sämtliche stationären Punkte und entscheiden Sie, ob es sich um Extrempunkte handelt und ob diese lokal, global und strikt bzw. nicht strikt sind. Begründen Sie Ihre Antworten.

(i)

$$\text{Minimiere} \quad f(x) := 3x^3 + 26x^2 - 12x + 5 \quad \text{über } x \in \mathbb{R}$$

(ii)

$$\text{Minimiere} \quad f(x) := x_1^4 + x_2^4 - 4x_1x_2 \quad \text{über } x = (x_1, x_2)^\top \in \mathbb{R}^2$$

Hausaufgabe 2.2 (Notwendige Bedingungen zweiter Ordnung sind nicht hinreichend)

Wir betrachten die Optimierungsaufgabe

$$\text{Minimiere} \quad f(x) = \left(x_1 - x_2^2\right)\left(2x_1 - x_2^2\right) = 2x_1^2 - 3x_1x_2^2 + x_2^4 \quad \text{über } x \in \mathbb{R}^2.$$

(i) Zeigen Sie, dass die notwendigen Optimalitätsbedingungen erster und zweiter Ordnung im Punkt $(0,0)^\top$ erfüllt sind.

(ii) Zeigen Sie, dass der Punkt $(0,0)^\top$ ein lokaler Minimierer von f entlang jeder geraden Linie durch den Ursprung ist.

(iii) Zeigen Sie, dass der Punkt $(0,0)^\top$ *kein* lokaler Minimierer von f ist.

Hausaufgabe 2.3 (Eigenschaften von Abstiegsverfahren)

Abstiegsverfahren sind iterative Optimierungsverfahren mit der Eigenschaft, dass die Folge der Funktionswerte $\big(f^{(k)}\big)_{k\in\mathbb{N}_0} := \big(f(x^{(k)})\big)_{k\in\mathbb{N}_0}$ der Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ bzgl. der Iterierten $\big(x^{(k)}\big)_{k\in\mathbb{N}_0}$ strikt fällt, also dass $f(x^{(k+1)}) < f(x^{(k)})$ für alle $k \in \mathbb{N}_0$. Ein Vertreter dieser Algorithmen aus dem Skript ist das vorkonditionierte Gradientenverfahren (Algorithmus 4.10).

Wir wollen zeigen, dass die Iterierten von Abstiegsverfahren nicht gegen lokale Maximierer konvergieren können, und dass für einige interessante Eigenschaften schon "nicht-Aufstieg" der Funktionswertfolge ausreicht.

(i) Es sei $\big(f^{(k)}\big)_{k\in\mathbb{N}_0}$ eine monoton fallende Folge, (also $f^{(k+1)} \le f^{(k)}$) mit Häufungspunkt f^*. Zeigen Sie, dass dann bereits die gesamte Folge $\big(f^{(k)}\big)_{k\in\mathbb{N}_0}$ gegen f^* konvergiert (also $f^{(k)} \xrightarrow{k\to\infty} f^*$).

(ii) Es seien $f\colon \mathbb{R}^n \to \mathbb{R}$ stetig und $\big(x^{(k)}\big)_{k\in\mathbb{N}_0}$ eine Folge mit zwei Häufungspunkten x^* und x^{**} sowie $f(x^{(k+1)}) \le f(x^{(k)})$. Zeigen Sie, dass dann $f(x^*) = f(x^{**})$.

(iii) Es seien $f\colon \mathbb{R}^n \to \mathbb{R}$ stetig und $\big(x^{(k)}\big)_{k\in\mathbb{N}_0}$ eine Folge mit einem Häufungspunkt x^* sowie $f(x^{(k+1)}) \le f(x^{(k)})$. Zeigen Sie, dass x^* nur dann ein strikter lokaler Maximierer von f sein kann, wenn x^* ein isolierter Punkt der Folge ist.

(iv) Zeigen Sie, dass der Punkt x^* in Punkt (iii) für eine von einem Abstiegsverfahren generierte Folge von Iterierten gar kein lokaler Maximierer – also auch kein nicht strikter Maximierer – sein kann. Können Sie die Bedingung hierfür abschwächen?

(v) Geben Sie ein Beispiel an, in dem das Gradientenverfahren nach endlich vielen Schritten in einem Punkt x^* mit $\nabla_M f(x^*) = 0$ abbricht, obwohl kein Minimierer vorliegt.

Hausaufgabe 2.4 (Exakte Liniensuche für eine quadratische Funktion)

Die Vorschrift der exakten Liniensuche für eine Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ entlang einer Abstiegsrichtung $d \in \mathbb{R}^n$ lautet:

$$\text{Bestimme } t_{\min} \text{ so, dass } f(x + t_{\min} d) = \min_{t\ge 0} f(x + t\, d),$$

siehe Gleichung (4.2) im Skript.

Gegeben sei die quadratische Funktion

$$f(x) = \frac{1}{2} x^\top Q\, x + c^\top x + \gamma$$

mit einer symmetrischen, positiv definiten Matrix $Q \in \mathbb{R}^{n\times n}$, $c \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$, einem Punkt (Iterierten) $x \in \mathbb{R}^n$ und einer Richtung $d \in \mathbb{R}^n$, $d \neq 0$.

(i) Beweisen Sie die Darstellung

$$t_{\min} = -\frac{f'(x)\,d}{d^\intercal Q\,d} \tag{$*$}$$

und zeigen Sie, dass im Gradientenverfahren für $d = -M^{-1}f'(x) = -\nabla_M f(x)$ die Darstellung

$$t_{\min} = \frac{d^\intercal M\,d}{d^\intercal Q\,d} \tag{4.16}$$

aus dem Skript gilt.

(ii) Wie ändert sich die Schrittweite $t_{\min}$ und die *Korrektur* $\Delta x := t_{\min} d$, wenn Sie einen Vorkonditionierer M durch einen Vorkonditionierer $\tilde{M} = \alpha\, M$ für $\alpha \in \mathbb{R}, \alpha \neq 0$ ersetzen? Was bedeutet das für die Wahl des Vorkonditionierers bei Verwendung der exakten Schrittweite?

(iii) In der Implementierung von Abstiegsverfahren wird die Liniensuchfunktion für gewöhnlich als separater, austauschbarer Parameter an das Abstiegsverfahren übergeben. Beim Aufruf der Liniensuche im Abstiegsverfahren reicht das Verfahren dann für gewöhnlich nur die Informationen runter, die eine allgemeine Liniensuche benötigt, also eine Routine zur Auswertung des Schnitts $\varphi(t) := f(x + td)$ durch die Funktion entlang der Suchrichtung und ihrer Ableitungen. Der Term Gleichung ($*$) kann dann (weil Q nicht konkret auswertbar vorliegt) also so nicht ausgewertet werden. Drücken Sie Gleichung ($*$) nur durch Auswertungen der Funktion φ und ihrer ersten Ableitung φ' aus. Verwenden Sie insbesondere die Auswertungen von $\varphi(0)$ und $\varphi'(0)$ (diese werden bspw. vom Gradientenverfahren ohnehin berechnet und nach unten weitergereicht).

Hausaufgabe 2.5 (Programmieraufgabe - Gradientenverfahren)

Bearbeiten Sie `programmierung_gradientenverfahren.ipynb`.

https://scoop.iwr.uni-heidelberg.de/teaching/2023ws/lecture-grundlagen-der-optimierung