

ÜBUNG 2

Ausgabedatum: 25. Oktober 2022

Hausaufgabe 2.1 (Eigenschaften von Abstiegsverfahren) 10 Punkte

Abstiegsverfahren sind iterative Optimierungsverfahren mit der Eigenschaft, dass die Folge der Funktionswerte $(f^{(k)})_{k \in \mathbb{N}_0} := (f(x^{(k)}))_{k \in \mathbb{N}_0}$ der Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$ bzgl. der Iterierten $(x^{(k)})_{k \in \mathbb{N}_0}$ **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 $(f^{(k)})_{k \in \mathbb{N}_0}$ eine monoton fallende Folge, (also $f^{(k+1)} \leq f^{(k)}$) mit Häufungspunkt f^* . Zeigen Sie, dass dann bereits die gesamte Folge $(f^{(k)})_{k \in \mathbb{N}_0}$ gegen f^* konvergiert (also $f^{(k)} \xrightarrow{k \rightarrow \infty} f^*$).
- (ii) Es seien $f: \mathbb{R}^n \rightarrow \mathbb{R}$ stetig und $(x^{(k)})_{k \in \mathbb{N}_0}$ eine Folge mit zwei Häufungspunkten x^* und x^{**} sowie $f(x^{(k+1)}) \leq f(x^{(k)})$. Zeigen Sie, dass dann $f(x^*) = f(x^{**})$.
- (iii) Es seien $f: \mathbb{R}^n \rightarrow \mathbb{R}$ stetig und $(x^{(k)})_{k \in \mathbb{N}_0}$ eine Folge mit einem Häufungspunkt x^* sowie $f(x^{(k+1)}) \leq f(x^{(k)})$. 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) Gilt die Aussage in [Punkt \(iv\)](#) auch für das Gradientenverfahren, wenn es nach endlich vielen Schritten in einem Punkt x^* mit $\nabla_M f(x^*) = 0$ abbricht?

Hausaufgabe 2.2 (Exakte Liniensuche für eine quadratische Funktion)

11 Punkte

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

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

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

$$t_{\min} = -\frac{f'(x) d}{d^\top Q d} \quad (*)$$

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

$$t_{\min} = \frac{d^\top M d}{d^\top Q d} \quad (4.16)$$

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 2.3 (Häufungspunkte des Gradientenverfahrens mit spd Hessematrix sind Attraktoren)

14 Punkte

Es seien $f: \mathbb{R}^n \rightarrow \mathbb{R}$ zweimal stetig differenzierbar und $(x^{(k)})_{k \in \mathbb{N}_0}$ eine vom vorkonditionierten Gradientenverfahren ([Algorithmus 4.10](#)) mit Armijo-Backtracking erzeugte Folge.

Zeigen Sie: Ist x^* ein Häufungspunkt der Folge $(x^{(k)})_{k \in \mathbb{N}_0}$ und $f''(x^*)$ positiv definit, dann konvergiert die gesamte Folge gegen x^* , und x^* ist ein strikter lokaler Minimierer.

Hausaufgabe 2.4 (Stetige Abh. der Iterierten im Gradientenverfahren vom Startwert) 16 Punkte

Gegeben sei die quadratische Funktion

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

mit symmetrischer, positiv definiter Matrix $Q \in \mathbb{R}^{n \times n}$ und $c \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$.

- (i) Es seien $(x^{(k)})_{k \in \mathbb{N}_0}$, $(\tilde{x}^{(k)})_{k \in \mathbb{N}_0}$ zwei vom vorkonditionierten Gradientenverfahren ([Algorithmus 4.10](#)) mit *exakter Liniensuche* erzeugte Folgen von Iterierten zu den Startwerten $x^{(0)}$, respektive $\tilde{x}^{(0)}$. Zeigen Sie, dass für jedes ε ein $\delta(\varepsilon, x^{(0)}) > 0$ existiert, so dass

$$\|x^{(k)} - \tilde{x}^{(k)}\| < \varepsilon \quad \text{für alle } k \in \mathbb{N}_0,$$

wenn

$$\|x^{(0)} - \tilde{x}^{(0)}\| < \delta.$$

Beachte: Das bedeutet, dass die Iterierten stetig vom Startwert abhängen.

- (ii) Was bedeutet [Punkt \(i\)](#) für die Anzahl der Iterationen, wenn das Gradientenverfahren mit der Abbruchbedingung $\|\nabla_M f(x^{(k)})\|_M < \text{ATOL}$ terminiert?
- (iii) Kann die Aussage in [Punkt \(i\)](#) auf allgemeine (nicht-quadratische) Funktionen erweitert werden?

Es ist keine Abgabe Ihrer Lösungen zu diesem Übungsblatt vorgesehen.