Nichtlineare Optimierung - 3. Hausaufgabe

Claudia Wohlgemuth	*366323*
Thorsten Lucke	*363089*
Felix Thoma	*358638*

Tutor: Mathieu Rosière

4. Juni 2017

3.1	3.2	3.3	3.4	Σ

Anmerkungen:

Aufgabe 3.1

Obwohl die in der Aufgabenstellung formulierte vermeintliche Charakterisierung der Abstiegsrichtung außerordentlich intuitiv erscheint, ist sie im Allgemeinen nicht mit dem in der Vorlesung eingeführten Begriff der Abstiegsrichtung vereinbar. Das folgende Gegenbeispiel dient als Untermauerung dieser kühnen Behauptung.

Für $n = 1$ bemühen wir die Sinusfunktion $\sin\colon \mathbb{R} \to \mathbb{R}, x \mapsto \sin(x)$ und betrachten den Punkt $x_0 = \pi/2$. Dann ist mit $d = 1$ wegen

$$\nabla \sin(\pi/2) \cdot d = \cos(\pi/2) \cdot 1 = 0 \not< 0$$

keine Abstiegsrichtung im Sinne der Definition (4.1.1) gegeben. Setzen wir allerdings $\widetilde{\sigma} := \pi/2$, so ist für alle $\sigma \in (0, \widetilde{\sigma}]$ die geforderte Abschätzung

$$0 \leq \sin(x + \sigma d) < 1 = \sin(x)$$

dennoch erfüllt. Es kann also keine Äquivalenz gelten.

Aufgabe 3.2

Es sei $f\colon \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar, streng konvex und koerziv. Damit hat f ein eindeutiges Minimum $\bar{x} \in \mathbb{R}^n$ und $\nabla f(x) = 0$ genau dann, wenn $x = \bar{x}$ ist. (†) Weiter genüge die im Abstiegsverfahren gewählte Richtung $x \mapsto d(x)$ folgender Bedingung: Es gibt Konstanten $c_1, c_2 > 0$ und $p > -1$ mit

$$\frac{|\langle \nabla f(x), d(x) \rangle|}{\|\nabla f(x)\| \|d(x)\|} \geq \min\{c_1, c_2 \|\nabla f(x)\|^p\} \tag{$\star$}$$

für alle $x \in \mathbb{R}^n$. Weiter seien die Schrittweiten $(x, d(x)) \mapsto \sigma(x)$ effizient gewählt, d.h. es gibt ein $c_3 > 0$ mit

$$f(x + d(x)\sigma(x)) \leq f(x) - c_3 \left(\frac{\langle \nabla f(x), d(x) \rangle}{\|d(x)\|} \right)^2 \tag{$\star\star$}$$

für alle $x \in \mathbb{R}^n$.

Sei $x^0 \in \mathbb{R}^n$ beliebig. Wir zeigen zunächst, dass die Folge der Gradienten $(\nabla f(x^k))_k$ des Verfahrens gegen Null konvergiert.

Wegen

$$c_3 \left(\frac{\langle \nabla f(x^k), d(x^k) \rangle}{\|d(x^k)\|} \right)^2 \geq 0$$

und mit $(\star\star)$ ist die Folge $(f(x^k))_k$ monoton fallend und aufgrund der Beschränktheit von f auch konvergent. Von $(\star\star)$ ausgehend, erhalten wir zudem für $k \in \mathbb{N}$

$$c_3 \left(\frac{\langle \nabla f(x^k), d(x^k) \rangle}{\|d(x^k)\|} \right)^2 \leq f(x^k) - f(x^{k+1}) \overset{k\to\infty}{\to} 0.$$

Setzen wir nun $(\star)$ ein, so ergibt sich

$$c_3 \|\nabla f(x^k)\|^2 (\min\{c_1, c_2 \|\nabla f(x^k)\|^p\})^2 \leq c_3 \left(\frac{\langle \nabla f(x^k), d(x^k) \rangle}{\|d(x^k)\|} \right)^2 \leq f(x^k) - f(x^{k+1}) \overset{k\to\infty}{\to} 0.$$

Damit gilt $c_3 c_1^2 \|\nabla f(x^k)\|^2 \to 0$ oder $c_3 c_2^2 \|\nabla f(x^k)\|^{2+2p} \to 0$ für $k \to \infty$, was genau dann erfüllt ist, falls $(\nabla f(x^k))_k$ gegen Null konvergiert.

Aus der Koerzivität und der Stetigkeit von f folgt, dass die Niveaumenge $N(f, f(x_0))$ kompakt ist. Damit hat die jede Teilfolge der Folge $(x^k)_k$ eine konvergente Teilfolge $(x^{k_{l_m}})_m$, deren Grenzwert wir mit $\tilde{x}_l \in \mathbb{R}^n$ bezeichnen wollen. Mit der Stetigkeit der Ableitung von f folgt

$$\nabla f(\tilde{x}) = \lim_{m\to\infty} \nabla f(x^{k_{l_m}}) = 0.$$

Mit (†) folgt, dass $\tilde{x}_l$ der (eindeutige) Minimierer $\bar{x}$ von f ist. Insgesamt konvergiert also eine Teilfolge jeder Teilfolge gegen $\bar{x}$ und somit konvergiert die ganze Folge.

Aufgabe 3.3

Aufgabe 3.4