Nichtlineare Optimierung
3. Übungsblatt

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Mathematik
Prof. Dr. Stefan Ulbrich
Hannes Meinlschmidt

WS 2011-2012
4.11.2011

Gruppenübung

Aufgabe G1 (Unzulässige Schrittweiten mit der Armijo-Regel)
Wählt man im allgemeinen Abstiegverfahren Suchrichtungen s_k, für die $\|s_k\|$ zu schnell gegen 0 geht, so liefert die Armijo-Bedingung alleine nicht immer zulässige Schrittweiten, selbst wenn die Suchrichtungen zulässig sind. Untersuchen Sie zur Bestätigung dieser Behauptung für eine stetig differenzierbare Funktion $f : \mathbb{R}^n \to \mathbb{R}$ die Suchrichtungen $s_k = -2^{-k}\nabla f(x_k)$.

(a) Zeigen Sie, dass die s_k zulässige Suchrichtungen liefern.

(b) Zeigen Sie am Beispiel $f(x) = x^2/4$, dass mit Startpunkt $x_0 \neq 0$ und mit der Wahl $\gamma \leq 3/4$ in der Armijo-Regel stets $\sigma_k = 1$ gewählt wird und diese Schrittweitenwahl unzulässig ist. Konvergiert der Algorithmus gegen das Minimum?

Hinweis: Sie dürfen benutzen, dass $\lim_{k\to\infty} \prod_{i=1}^{k+1}(1 - \frac{1}{2^{i+1}}) = a$ mit einem $a > 0$.

Aufgabe G2 (Spaß mit Gradienten)
In dieser Aufgabe soll die handelsübliche Definition des Gradienten einer stetig differenzierbaren Funktion $f : \mathbb{R}^n \to \mathbb{R}$, nämlich

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)\right)^T,$$

ein wenig aufgeweicht, bzw. verallgemeinert werden, und zwar zu folgender: Der Gradient von f im Punkt $x \in \mathbb{R}^n$ ist eindeutig gegeben als der Vektor $\nabla f(x) \in \mathbb{R}^n$, für den

$$\langle \nabla f(x), u\rangle = f'(x)u \quad \text{für alle } u \in \mathbb{R}^n$$

gilt, wobei $f'(x)$ die Ableitung bzw. Jacobimatrix von f bezeichne. Klarerweise sind die beiden Definitionen äquivalent (wieso?), allerdings erlaubt die letztere eine interessante Verallgemeinerung:
Sei $M \in \mathbb{R}^{n\times n}$ symmetrisch und positiv definit, und bezeichne $\langle x, y\rangle_M = x^T M y$ das durch M induzierte Skalarprodukt auf $\mathbb{R}^n$. Dann definieren wir den Gradienten von f im Punkt $x \in \mathbb{R}^n$ bzgl. $\langle\cdot,\cdot\rangle_M$ als den eindeutigen Vektor $\nabla_M f(x)$, für den

$$\langle \nabla_M f(x), u\rangle_M = f'(x)u \quad \text{für alle } u \in \mathbb{R}^n$$

gilt. Zeigen Sie:

(a) Der Gradient $\nabla_M f(x)$ ist gegeben durch $M^{-1}\nabla f(x)$.

(b) Die anschauliche Vorstellung des steilsten An- bzw. Abstiegs ist intakt, denn das Problem

$$\min_{\substack{d\in\mathbb{R}^n\\ \|d\|_M=1}} \langle \nabla f(x), d\rangle$$

hat die eindeutige Lösung

$$d^* = -\frac{\nabla_M f(x)}{\|\nabla_M f(x)\|_M} = -\frac{M^{-1}\nabla f(x)}{\|M^{-1}\nabla f(x)\|_M}.$$

Hier ist $\|\cdot\|_M$ die durch

$$\|x\|_M = \sqrt{\langle x, x\rangle_M} = \sqrt{x^T M x}$$

definierte Norm auf $\mathbb{R}^n$.

Hausübung

Aufgabe H1 (Unzulässige Suchrichtungen) (6 Punkte)
Hier wird Beispiel 2.5.1 aus dem Skript nachgeliefert. Wählt man im allgemeinen Abstiegsverfahren Suchrichtungen, die nicht zulässig sind, weil sie fast senkrecht zur Gradientenrichtung, d.h. fast tangential zu den Höhenlinien der Funktion, verlaufen, so kann es auch mit zulässigen Schrittweiten passieren, dass das Abstiegsverfahren nicht gegen einen stationären Punkt konvergiert. Untersuchen Sie dazu die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, gegeben durch $f(x_1, x_2) = \frac{1}{2}(x_1^2 + x_2^2) = \frac{1}{2}\|x\|^2$ für die Suchrichtungen

$$s_k = g_k^{\perp} - \frac{1}{2^{k+3}} g_k.$$

Hierbei sei $g_k = \nabla f(x_k)$ und $g_k^{\perp}$ so, dass $g_k^{\perp} \perp g_k$ und $\|s_k\| = \|g_k\|$.
Zeigen Sie, dass das Abstiegsverfahren mit diesen Suchrichtungen und zulässiger Schrittweitenwahl für keinen Startpunkt $x_0 \neq 0$ gegen den Minimalpunkt $\bar{x} = 0$ von f konvergiert und $\bar{x}$ auch kein Häufungspunkt von (x_k) ist.

Aufgabe H2 (Abstiegsrichtungen und Krümmung) (4 Punkte)
Sei $f : \mathbb{R}^n \to \mathbb{R}$ zweimal stetig differenzierbar, und sei $x \in \mathbb{R}^n$. Beweisen oder widerlegen Sie:
Wenn $\nabla^2 f(x)$ einen negativen Eigenwert besitzt, dann existiert eine Abstiegsrichtung in x, d.h. ein Vektor d mit der Eigenschaft $f(x) > f(x + \alpha d)$ für ein $\alpha > 0$. Die Abstiegsrichtung ist nicht eindeutig in Bezug auf Vorzeichen.

Aufgabe H3 (Matrixnorm und Eigenwerte) (4 Punkte)
Sei $G \in \mathbb{R}^{n,n}$ symmetrisch mit Eigenwerten $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Bezeichne $\|\cdot\|_2$ die euklidische Norm auf $\mathbb{R}^n$ und

$$\|G\|_2 := \sup_{x \neq 0} \frac{\|Gx\|_2}{\|x\|_2}$$

die induzierte Matrixnorm (Spektralnorm). Zudem gelte mit $0 < \mu \leq \eta$

$$\mu\|x\|_2^2 \leq x^T G x \leq \eta\|x\|_2^2 \quad \text{für alle } x \in \mathbb{R}^n.$$

Weisen Sie für alle $x \in \mathbb{R}^n$ die folgenden Ungleichungen bzw. Behauptungen nach:

(a) G ist positiv definit und es gilt $\mu \leq \lambda_1 \leq \cdots \leq \lambda_n \leq \eta$,

(b) die möglichen Werte von $x^T G x$ sind noch genauer eingegrenzt durch

$$\mu\|x\|_2^2 \leq \lambda_1\|x\|_2^2 \leq x^T G x \leq \lambda_n\|x\|_2^2 \leq \eta\|x\|_2^2,$$

(c) die Bezeichnung *Spektralnorm* ist sinnvoll, denn es gilt $\|G\|_2 = \lambda_n \leq \eta$,

(d) G ist invertierbar mit $\|G^{-1}\|_2 = \frac{1}{\lambda_1} \leq \frac{1}{\mu}$, zudem gilt

$$\frac{1}{\eta}\|x\|^2 \leq x^T G^{-1} x \leq \frac{1}{\mu}\|x\|^2 \quad \text{und} \quad \frac{1}{\eta}\|x\|_2^2 \leq \frac{1}{\lambda_n}\|x\|_2^2 \leq x^T G^{-1} x \leq \frac{1}{\lambda_1}\|x\|_2^2 \leq \frac{1}{\mu}\|x\|_2^2.$$

Insbesondere zeigt (d), dass man die obigen Forderungen an G auch äquivalenterweise an G^{-1} stellen kann.