Nichtlineare Optimierung
3. Übungsblatt

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Mathematik
Prof. Dr. Irwin Yousept
Hannes Meinlschmidt

WS 2012/13
8.11.2012

Gruppenübung

Aufgabe G1 (Zum Warmwerden: Exakte Schrittweitensuche und Armijo-Bedingung)
Sei $f(x) = \frac{1}{2}x^TCx + c^Tx$ mit $c \in \mathbb{R}^n$ und $C \in \mathbb{R}^{n\times n}$ symmetrisch positiv definit eine quadratische Funktion von $\mathbb{R}^n$ nach $\mathbb{R}$. Gegeben ist mit $s \in \mathbb{R}^n$ eine Abstiegsrichtung von f im Punkt x und σ^* die durch die exakte Schrittweitenregel $\sigma^* = \arg\min_{\sigma>0} f(x+\sigma s)$ bestimmte Schrittweite. Zeigen Sie: Die exakte Schrittweite σ^* erfüllt die Armijo-Bedingung

$$f(x+\sigma s) - f(x) \leq \sigma\gamma\nabla f(x)^T s$$

für alle $\gamma \in (0,1/2]$, und verletzt die Bedingung für alle $\gamma > 1/2$.

Aufgabe G2 (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 mit den angegebenen Suchrichtungen und Schrittweiten gegen das Minimum von f?

Hinweis: Sie dürfen ohne Beweis benutzen, dass $\lim_{k\to\infty}\prod_{i=1}^{k+1}(1-\frac{1}{2^{i+1}})$ existiert und der Grenzwert $a > 0$ ist.

Aufgabe G3 (Curry-Schrittweitenregel)
Sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar, $x^0 \in \mathbb{R}^n$ und sei ∇f auf der Niveaumenge $N_f(x^0)$ Lipschitz-stetig mit Lipschitz-Konstante $L > 0$. Sei weiter $x^k \in N_f(x^0)$ und $s^k \in \mathbb{R}^n$ eine Abstiegrichtung von f in x^k. Nach der *Curry-Schrittweitenregel* wird die Schrittweite $\sigma_k > 0$ als kleinster stationärer Punkt der Funktion $\Phi(\sigma) = f(x^k + \sigma s^k)$, wobei $\sigma > 0$, berechnet, d.h.

$$\sigma_k = \min\left\{\sigma > 0 : \nabla f(x^k + \sigma s^k)^T s^k = 0\right\}.$$

(a) Zeigen Sie: Ist $\tau > 0$ die kleinste Zahl mit $\nabla f(x^k + \tau s^k)^T s = \frac{1}{2}\nabla f(x^k)^T s^k$, dann gilt

$$f(x^k + \sigma_k s^k) - f(x^k) \leq f(x^k + \tau s^k) - f(x^k) \leq \frac{\tau}{2}\nabla f(x^k)^T s^k.$$

(b) Nutzen Sie nun die Lipschitz-Stetigkeit von ∇f auf $N_f(x^0)$, um

$$\tau \geq \frac{\left|\nabla f(x^k)^T s^k\right|}{2\|s^k\|^2 L}$$

zu zeigen.

(c) Beweisen Sie nun die Zulässigkeit der Curry-Schrittweiten (σ_k).

Hausübung

Aufgabe H1 (Zulässigkeit der Powell-Wolfe Schrittweiten) (6 Punkte)
Beweisen Sie den folgenden Satz über die Zulässigkeit der Powell-Wolfe-Schrittweiten: Sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar. Sei $x^0 \in \mathbb{R}^n$ ein beliebiger Startpunkt und sei ∇f auf der Niveaumenge $N_f(x^0)$ gleichmäßig stetig (z.B. der Fall, wenn $N_f(x^0)$ kompakt ist). Dann ist im Allgemeinen Abstiegsverfahren die mit der Powell-Wolfe-Regel erzeugte Schrittweitenfolge (σ_k) zulässig.
Erinnerung: Die Powell-Wolfe-Schrittweite σ wird so bestimmt, dass die beiden Bedingungen

$$f(x) - f(x + \sigma s) \geq -\gamma\sigma\nabla f(x)^T s \tag{PW1}$$

und

$$\nabla f(x + \sigma s)^T s \geq \theta \nabla f(x)^T s. \tag{PW2}$$

für $\gamma \in \left(0, \frac{1}{2}\right)$ und $\theta \in (\gamma, 1)$ erfüllt sind.
Hinweis: Zeigen Sie die Zulässigkeitsbedingung der Schrittweite durch Kontraposition; nehmen Sie also an, dass es eine Indexmenge $K \subset \mathbb{N}$ mit

$$-\frac{\nabla f(x^k)^T s^k}{\|s^k\|_2} \geq \varepsilon > 0 \quad \text{für ein } \varepsilon > 0 \text{ und alle } k \in K$$

gibt. Folgern Sie daraus mithilfe der zweiten Powell-Wolfe-Bedingung, dass $\|x^{k+1} - x^k\| \geq \delta$ für ein $\delta > 0$ und alle $k \in K$ gilt, und zeigen Sie damit $f(x^k) - f(x^{k+1}) \not\to 0$.

Aufgabe H2 (Unzulässige Suchrichtungen) (6 Punkte)
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

$$s^k = g_\perp^k - \frac{1}{2^{k+3}} g^k.$$

Hierbei sei $g^k = \nabla f(x^k)$ und $g_\perp^k$ so, dass $g_\perp^k \perp g^k$ 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 H3 (Resteverwertung und Rechnerpraktikum)
Bereiten Sie sich auf das Rechnerpraktikum nächste Woche vor. Lesen Sie sich dazu gegebenenfalls wieder in MATLAB ein (zB. mathematische Funktionen modellieren) und gehen Sie auch das Gradientenverfahren und eine mögliche Implementierung der Armijo-Regel durch. Bearbeiten Sie noch Aufgabe G3, falls Ihnen sonst zu langweilig wird.