Nichtlineare Optimierung
8. Übungsblatt

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Mathematik
Prof. Dr. Stefan Ulbrich
M.Sc. Franziska Kartzow
Dipl.-Math. Sebastian Pfaff

WS 2010/2011
17. Dezember 2010

Gruppenübung

Aufgabe G1 (Exakte Lösung eines TR-Problems)
Für die quadratische Funktion

$$q(s) = c^T s + \frac{1}{2} s^T H s, \qquad \text{mit } c \in \mathbb{R}^n,\ H \in \mathbb{R}^{n\times n} \text{ symmetrisch},$$

betrachten wir für $\Delta > 0$ das Trust-Region-Problem

$$\min q(s) \quad \text{s.t.} \quad \|s\|_2 \le \Delta. \tag{TP}$$

Zeige:

(a) Das Trust-Region-Problem (TP) besitzt eine Lösung.

(b) Wenn es ein $\lambda \ge 0$ gibt, so dass die Bedingungen
   (i) $(H + \lambda I)\bar{s} = -c$,
   (ii) $(H + \lambda I)$ positiv semidefinit,
   (iii) entweder $\|\bar{s}\|_2 = \Delta$ oder $\|\bar{s}\|_2 < \Delta$ und $\lambda = 0$

   erfüllt sind, so ist $\bar{s}$ eine Lösung des Trust-Region-Problems (TP).
   Hinweis: Betrachte zunächst die Funktion $\tilde{q}(s) := c^T s + \frac{1}{2} s^T (H + \lambda I)s = q(s) + \frac{\lambda}{2}\|s\|_2^2$.

Bemerkung: Die Bedingungen unter (b) sind nicht nur hinreichend, sondern auch *notwendige* Optimalitätsbedingungen, wie wir in den Hausübungen noch sehen werden.

Aufgabe G2 (Beispiel für ein TR-Problem)
Gegeben sei

$$\min c^T s + \frac{1}{2} s^T H s \quad \text{s.t.} \quad \|s\|_2 \le \sqrt{2}, \text{ mit } c = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

(a) Geben Sie die notwendigen und hinreichenden Optimalitätsbedingungen an (siehe letzte Aufgabe).

(b) Berechnen Sie $s(\lambda) := -(H + \lambda I)^{-1} c$, für $\lambda > \max\{0, -\lambda_{\min}(H)\}$ und $p(\lambda) := \|s(\lambda)\|_2$.

(c) Bestimmen Sie die Optimallösung des Trust-Region-Problems. (Hinweis: Der optimale Lagrange-Multiplikator λ ist ganzzahlig.) Zeichen Sie die Funktionen $p(\lambda) - \Delta$ und $\frac{1}{p(\lambda)} - \frac{1}{\Delta}$. Was beobachten Sie?

Aufgabe G3 (Optimallösungen des Trust-Region-Problems)
Wir betrachten das Trust-Region-Problem aus Aufgabe G1. Man kann die Bedingungen (i)-(iii) verwenden, um eine Lösung des Trust-Region-Problems zu finden (siehe auch Aufgabe G2). Erfüllt $\lambda = 0$ die Bedingungen mit einem $\|s\|_2 \le \Delta$, so sind wir fertig. Für $\lambda > \max\{0, -\lambda_{\min}(H)\}$ definieren wir

$$s(\lambda) := -(H + \lambda I)^{-1} c.$$

Um neben (i) und (ii) auch Bedingung (iii) zu erfüllen, versucht man nun, eine Nullstelle $\lambda > 0$ der Funktion

$$\phi(\lambda) := \|s(\lambda)\|_2 - \Delta$$

zu finden.

(a) Sei $Q = (q_1, q_2, \dots, q_n)$ eine orthogonale Matrix mit $Q^T H Q = \operatorname{diag}(\lambda_1, \dots \lambda_n)$, wobei $\lambda_1 \le \lambda_2 \le \dots \le \lambda_n$ die Eigenwerte von H sind. Zeigen Sie, dass für $\lambda \ne -\lambda_i$ $(i = 1, \dots n)$ gilt:

$$\|s(\lambda)\|_2^2 = \sum_{i=1}^n \frac{(q_i^T c)^2}{(\lambda_i + \lambda)^2}.$$

(b) Diskutieren Sie die Existenz der Nullstellen $\lambda > \max\{0, -\lambda_{\min}(H)\}$ von ϕ, falls $q_1^T c \ne 0$ und $\lambda_1 < 0$.
Hinweis: Untersuchen Sie $\lim_{\lambda\to\infty} \|s(\lambda)\|$ und $\lim_{\lambda\searrow -\lambda_1} \|s(\lambda)\|$.

(c) Untersuchen Sie den Fall, wenn H positiv definit ist und $\|s(0)\|_2 > \Delta$.

(d) Angenommen, man wendet unter den Voraussetzungen von b) das Newton-Verfahren zur Nullstellensuche auf ϕ an. Was passiert für λ nahe $-\lambda_1$? Warum ist es besser, die Gleichung $\Phi(\lambda) := \frac{1}{\|s(\lambda)\|} - \frac{1}{\Delta} = 0$ zu lösen?

Hausübung

Aufgabe H1 (CG-Verfahren - Ein Beispiel) (5 Punkte)
Verwenden Sie das CG-Verfahren aus der Aufgabe G1 von Übungsblatt 5 um das globale Minimum der Funktion

$$q : y \mapsto \frac{1}{2} y^T C y + c^T y, \qquad \text{mit} \qquad C = \operatorname{diag}(1, 2), \quad c = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

zu bestimmen. Starten sie im Punkt $y_0 = (0, 0)^T$. Rechnen Sie mit Brüchen ohne zu runden. Geben Sie alle Vektoren g_k, d_k, y_k, sowie die Werte α_k, β_k an.
Fertigen Sie danach eine Zeichnung an. Diese soll folgendes enthalten:

- die Höhenlinien für $q(y) = -0.2r$ für $r = 0, 1, 2, 3$
- die Iterierten y_k
- die Höhenlinien durch y_k
- Pfeile an jedem y_k, die die Richtung des jeweiligen Gradienten angeben
- die Verbindungsstrecke zwischen y_k und $y_k - d_k$

Wie kann man aus der Zeichnung die Werte für α_k ablesen?

Aufgabe H2 (Notwendigkeit der Bedingungen aus G1) (10 Punkte)
Wir betrachten erneut Aufgabe G1 (auf diesem Übungsblatt). $\bar{s}$ sei optimal für (TP). Zeigen Sie, dass die Bedingungen unter G1(b) notwendigerweise gelten. Gehen Sie dabei wie folgt vor:

(a) Zeigen Sie, dass für $\|\bar{s}\|_2 < \Delta$ die Bedingungen aus G1(b) mit einem $\lambda \ge 0$ erfüllt sind.

Im weiteren sei $\|\bar{s}\|_2 = \Delta$.

(b) Zeigen Sie nun, dass für alle $v \in \mathbb{R}^n \setminus \{0\}$ mit $\|\bar{s} + v\|_2 \le \Delta$ die Richtungsableitung $q'(\bar{s}, v) = \lim_{t\searrow 0} \frac{q(\bar{s}+tv)-q(\bar{s})}{t}$ in Richtung v nichtnegativ ist.

(c) Zeigen Sie

$$\nabla q(\bar{s})^T \hat{v} \ge 0 \qquad \text{für alle } \hat{v} \in \mathbb{R}^n \text{ mit } \hat{v}^T \bar{s} \le 0.$$

Hinweis: Zeigen Sie zuerst, dass es im Falle $\hat{v}^T \bar{s} < 0$ ein $\bar{t} > 0$ gibt, sodass $\|\bar{s} + t\hat{v}\|_2 \le \Delta$ für alle $t \in (0, \bar{t}]$ gilt und verwenden Sie $q'(\bar{s}, \hat{v}) = \frac{1}{\|\hat{v}\|_2} \nabla q(\bar{s})^T \hat{v}$.

(d) Wir erinnern uns an das Lemma von Farkas aus der "Einführung in die Optimierung":
Für $A \in \mathbb{R}^{m\times n}$ und $b \in \mathbb{R}^m$ hat genau eines der folgenden Systeme eine Lösung:

$$\begin{array}{rcl} Ax & = & b \\ x & \ge & 0 \end{array} \quad \dot{\bigvee} \quad \begin{array}{rcl} y^T A & \ge & 0 \\ y^T b & < & 0. \end{array}$$

Zeigen Sie mit Hilfe des Farkas-Lemmas und Aufgabenteil (c), dass (i) in G1(b) erfüllt wird, d.h. es gibt ein $\lambda \ge 0$, sodass $(H + \lambda I)\bar{s} = -c$.

(e) Zeigen Sie, dass mit λ aus (d) die Matrix $(H + \lambda I)$ positiv semidefinit ist. Betrachten Sie zuerst für $y \in \mathbb{R}^n$ mit $y^T \bar{s} \neq 0$ die Differenz

$$\tilde{q}(\bar{s}) - \tilde{q}(\tilde{s})$$

wobei

$$\tilde{s} := \bar{s} + \alpha y, \qquad \alpha := -2\frac{y^T \bar{s}}{y^T y}.$$

Zeigen Sie auf diesem Weg $y^T(H + \lambda I)y \geq 0$ für den Fall, dass $y^T \bar{s} \neq 0$ gilt. Argumentieren Sie, warum daraus schon $y^T(H + \lambda I)y \geq 0$ für alle $y \in \mathbb{R}^n$ folgt.

Abschließend wünschen wir Ihnen

Frohe Weihnachten und ein erfolgreiches Jahr 2011!

Die nächsten Übungen finden am 14. bzw. 17. Januar 2011 in den entsprechenden Rechnerpools statt.