Nichtlineare Optimierung
9. Übungsblatt

TECHNISCHE UNIVERSITÄT DARMSTADT

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

WS 2010/2011
14. Januar 2011

Rechnerübung

Aufgabe R1 (Trust-Region-Verfahren)
Programmieren Sie das Trust-Region-Verfahren (Algorithmus 14 der Vorlesung) in `Matlab`. Schreiben Sie dazu eine Funktion

```
function [xn] = tr(x0,fgH,tol,trsolv).
```

Hierbei sei `x0` der Startpunkt, `fgH` eine Funktion, welche Funktionswert, Gradient und Hessematrix der zu minimierenden Funktion zurückgibt, `tol` eine Toleranz für die (relative) Abbruchbedingung und `trsolv` ein Parameter, der steuert, welches Verfahren zur Lösung der Trust-Region-Teilprobleme benutzt werden soll. Verwenden Sie folgende Konstanten

$$\Delta_{\min} = 0.1, \quad \eta_1 = 10^{-3}, \quad \eta_2 = 0.8, \quad \beta_1 = 0.5 \quad \text{und} \quad \beta_2 = 2.$$

(a) Implementieren Sie das sogenannte *Steihaug-CG-Verfahren* zur näherungsweisen Lösung der Trust-Region-Teilprobleme (Schritt 3 von Algorithmus 14). Schreiben Sie also eine Funktion

```
function [sn] = steihaugcg(C,c,Del),
```

wobei `c` der Gradient der Zielfunktion im aktuellen Iterationspunkt, `C` die Hessematrix im aktuellen Iterationspunkt und `Del` der aktuelle Trust-Region-Radius sei. Verwenden Sie dazu Ihr CG-Newton-Verfahren zur Bestimmung der Suchrichtung aus der dritten Rechnerübung. Alternativ steht auf der Veranstaltungsseite die Datei `cg_newton.m` zum Download bereit. Ändern Sie das Verfahren ab, indem Sie folgende Abbruchbedingungen verwenden:

Steihaug-CG-Verfahren:
Wende das CG-Verfahren zur Minimierung von $q = q_k$ an. Terminiere mit $s_k = s^S$ nach folgenden Regeln: Sei $\epsilon \in]0,1[, \nu > 0$, $g_0 = \nabla q(0)$ und sonstige Bezeichnungen wie in Aufgabe H2 von Übungsblatt 5.

1. Falls $\|g_j\| \leq \min(\epsilon\|g_0\|, \nu\|g_0\|^2)$: STOP mit $s^S = y_j$.
2. Falls $d_j^T H d_j \leq 0$: STOP mit
$$s^S = y_j - \alpha^* \operatorname{sgn}(g_j^T d_j) d_j,$$
wobei $\alpha^* \geq 0$, so dass $\|s^S\| = \Delta_k$.
3. Falls $\|y_{j+1}\| > \Delta_k$: STOP mit
$$s^S = y_j - \alpha^* d_j,$$
mit $\alpha^* \geq 0$, so dass $\|s^S\| = \Delta_k$.

Verwenden Sie die Konstanten

$$\epsilon = 0.01, \quad \nu = 1.$$

Testen Sie dieses Programm am quadratischen Modell der Funktion

$$f(x_1, x_2) = 10(x_2 - x_1^2)^2 + (1 - x_1)^2$$

in den Punkten $x = (0, -1)$ und $x = (0, 0.5)$ und für verschiedene Trust-Region-Radien zwischen 0 und 2. Veranschaulichen Sie Ihr Ergebnis, indem Sie sich die Höhenlinien des Quadratischen Modells, sowie den Rand der Trust Region (= Höhenlinie zum Niveau Δ^2 der Funktion $s \mapsto s^T s$) und den erhaltenen Punkt s^S plotten lassen.

(b) Testen Sie nun Ihr Trust-Region-Verfahren `tr` unter Verwendung des Steihaug-CG-Verfahrens an den Funktionen
$f_1(x_1,x_2) = 10(x_2 - x_1^2)^2 + (1-x_1)^2$, mit Startpunkten $(0,-1)$ und $(0,0.5)$.
$f_2(x_1,x_2) = 100(x_2 - x_1^2)^2 + (1-x_1)^2$ (globales Minimum bei $(1,1)$),
mit Startwerten $(0,2)$ und $(20,10)$,
$f_3(x_1,x_2) = x_1^2 + 20x_2^2$ mit Startwerten $(20,10)$, $(1,2)$ und $(-1,2)$.

(c) Implementieren Sie nun das folgende, sogenannte *Dogleg-Verfahren* zur näherungsweisen Lösung der TR-Teilprobleme (Schritt 3 von Algorithmus 14):

Dogleg-Verfahren:
Berechne den Cauchy-Punkt s_k^c und zudem den Newton-Schritt $s_k^N = -H_k^{-1}\nabla f(x_k)$, falls H_k positiv definit ist. Setze

$$s_k = \begin{cases} s_k^N, & \text{falls } H_k \text{ pos.def. und } \|s_k^N\| \leq \Delta_k, \\ s_k^c, & \text{falls } H_k \text{ nicht pos.def.}, \\ s_k^{DL}(t^*), & t^* = \underset{t\in[0,1],\|s_k^{DL}(t)\|\leq\Delta_k}{\operatorname{argmin}} \quad q_k(s_k^{DL}(t)) \text{ sonst}, \end{cases}$$

wobei

$$s_k^{DL}(t) := s_k^c + t(s_k^N - s_k^c).$$

Beachten Sie, dass die Abbildung $[0,1] \ni t \mapsto q_k(s_k^{DL}(t))$ für positiv definite H_k streng konvex ist. Somit besitzt sie in der Trust-Region ein eindeutiges Minimum.
Testen Sie die Hessematrix H_k auf positive Definitheit, indem Sie eine Cholesky-Zerlegung versuchen. Verwenden Sie hierzu die `Matlab`-Funktion `chol.m` mit dem Aufruf

```
[R,p] = chol(H),
```

mit $H = H_k$. Hierbei ist `p` ein Flag, das angibt, ob die Zerlegung erfolgreich war: Ist `H` positiv definit, so ist `p=0`. Ansonsten ist `p` eine positive natürliche Zahl. Benutzen Sie ausserdem die Ergebnisse aus Aufgabe H2 aus der 7. Übung, um den Cauchy-Punkt zu berechnen. Beachten Sie auch Lemma 2.12.8 aus der Vorlesung zur Berechnung des Dogleg-Punktes. Testen Sie nun Ihr Trust-Region-Verfahren `tr` unter Verwendung des Dogleg-Verfahrens an den Funktionen aus Aufgabe (b).

Hausübung

Aufgabe H1 (Linearisierungskegel) (2 Punkte)
Es sei die Menge Z gegeben durch $Z = \{(x_1,x_2)^T \in \mathbb{R}^2 : x_2 = 0\}$. Dann hat man mit $h(x) = x_2$ und $\tilde{h}(x) = x_2^2$ die Darstellungen

$$Z = \{x \in \mathbb{R}^2 : h(x) = 0\} = \{x \in \mathbb{R}^2 : \tilde{h}(x) = 0\}.$$

Berechnen Sie $T_L(h;x)$, $T_L(\tilde{h};x)$ und $T(Z;x)$. Welche der Mengen stimmen überein, welche nicht und welche Inklusionen gelten dann?

Aufgabe H2 (Optimalpunkt ohne Constraint Qualification) (8 Punkte)
Betrachten Sie das Minimierungsproblem

$$\min_{x\in\mathbb{R}^2} -x_1 \quad \text{u.d.N. } x_2 - (1-x_1)^3 \leq 0, \quad -x_1 \leq 0, \quad -x_2 \leq 0. \tag{1}$$

(a) Ermitteln Sie graphisch die Lösung $\bar{x}$.

(b) Ermitteln Sie den Tangentialkegel $T(Z,\bar{x})$ sowie den Linearisierungskegel $T_L(c,h;\bar{x})$ für $\bar{x} = (1,0)$ und den Zulässigkeitsbereich Z. Weisen Sie nach, dass keine Constraint Qualification in $\bar{x} = (1,0)$ gelten kann.

(c) Geben Sie zwei (unabhängige) lineare Nebenbedingungen an, sodass das Hinzufügen jeweils einer dieser Bedingungen zu (1) bewirkt, dass im Punkt $\bar{x} = (1,0)$ die Abadie Constraint Qualification (ACQ) erfüllt ist, die zulässige Menge jedoch unverändert bleibt.

(d) Ersetzen Sie die erste Nebenbedingung in (1) durch

$$\tilde{c}_1(x) := x_2 - \max((1-x_1)^3, 0) \ \leq \ 0.$$

Überprüfen Sie nun $\bar{x} = (1,0)$ erneut auf (ACQ). Ist $\bar{x}$ für das veränderte Problem optimal?

Aufgabe H3 (KKT-Bedingungen) (4 Punkte)

Gegeben seien folgende Optimierungsprobleme:

$$
(\mathrm{P1})\quad \begin{array}{lrcl} \max & 2x_1 + 3x_2 & & \\ \text{s.t.} & x_1 + x_2 & \leq & 8 \\ & -x_1 + 2x_2 & \leq & 4 \\ & x_1, x_2 & \geq & 0 \end{array}
\qquad
(\mathrm{P2})\quad \begin{array}{lrcl} \min & (x_1 - \frac{9}{4})^2 & + & (x_2 - 2)^2 \\ \text{s.t.} & x_2 - x_1^2 & \geq & 0 \\ & x_1 + x_2 & \leq & 6 \\ & x_1, x_2 & \geq & 0 \end{array}
$$

(a) Formulieren Sie die KKT-Bedingungen für (P1). Verifizieren Sie für jeden Eckpunkt (algebraisch und geometrisch), ob die KKT-Bedingungen gelten. Was ist die globale Lösung?

(b) Formulieren Sie die KKT-Bedingungen für (P2) und verifizieren Sie, dass diese im Punkt $x_* = (\frac{3}{2}, \frac{9}{4})$ erfüllt sind. Interpretieren Sie die KKT-Bedingungen in x_* geometrisch. Gilt diese Interpretation allgemein für Probleme mit Ungleichungrestriktionen?