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

TECHNISCHE UNIVERSITÄT DARMSTADT

WS 2010/2011
26. November 2010

5. Übungsblatt zur „Nichtlinearen Optimierung“

Gruppenübung

Aufgabe G1 (Verfahren der konjugierten Gradienten (CG-Verfahren))
Gegeben sei die streng konvexe quadratische Funktion $q : \mathbb{R}^n \longrightarrow \mathbb{R}$, $q(y) = c^T y + \frac{1}{2} y^T C y$. Zur Bestimmung des eindeutigen globalen Minimums von q betrachten wir das

---

Verfahren der konjugierten Gradienten:

1. Wähle y_0 und berechne $g_0 := c + Cy_0$. Falls $g_0 = 0$: STOP mit Ergebnis y_0.
   Sonst setze $d_0 := g_0$ und $k := 0$.
2. Berechne $\alpha_k := \frac{g_k^T g_k}{d_k^T C d_k}$ und setze $y_{k+1} := y_k - \alpha_k d_k$ sowie $g_{k+1} := g_k - \alpha_k C d_k$.
   Falls $g_{k+1} = 0$: STOP mit Ergebnis y_{k+1}.
3. Berechne $\beta_k := \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}$ und setze $d_{k+1} := g_{k+1} + \beta_k d_k$.
4. Ersetze k durch $k+1$ und gehe zu 2.

Zeigen Sie:

a) Solange $g_k \neq 0$ ist, gilt:

   i) $d_k \neq 0$,

   ii) $V_{k+1} := \text{span}\{g_0, Cg_0, \dots, C^k g_0\} = \text{span}\{g_0, \dots, g_k\} = \text{span}\{d_0, \dots, d_k\}$,

   iii) Die Vektoren $d_0, \dots d_k$ sind paarweise C-konjugiert, d.h.

   $$d_i^T C d_j = 0 \quad \text{für alle } i \neq j,\ 0 \leq i, j \leq k,$$

   iv) $g_{k+1} \perp V_{k+1}$.

b) Es gilt $q(y_{k+1}) = \min_{y \in V_{k+1}} q(y_0 + y)$. Und das Verfahren berechnet in höchstens n Schritten das globale Minimum von q.

Aufgabe G2 (Konvergenz des Newtonverfahrens in $\mathbb{R}$)
Sei x^* eine m-fache Nullstelle der Funktion $F \in C^{m+1}(\mathbb{R})$.

(a) Bestimmen Sie die Konvergenzordung des Newtonverfahrens

$$x_{k+1} = x_k - \frac{F(x_k)}{F'(x_k)}.$$

in Anhängigkeit von m und im Falle linearer Konvergenz die Konvergenzrate.

(b) Betrachten Sie nun das folgende veränderte Newtonverfahren :

$$x_{k+1} = x_k - n\frac{F(x_k)}{F'(x_k)} \qquad (n > 1).$$

Für welche m konvergiert dieses Verfahren? Bestimmen Sie für die m, für die das Verfahren konvergiert, die Konvergenzordnung in Abhängigkeit von n.

Hinweis: Benutzen Sie die Taylorentwicklung von $F(x)$ und $F'(x)$ mit Entwicklungspunkt x^* und

$$\frac{1+\mathcal{O}(y)}{1+\mathcal{O}(y)} = 1 + \mathcal{O}(y) .$$

Hausübung

Aufgabe H1 (Modifiziertes Newton-Verfahren) (10 Punkte)

Sei die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x_1, x_2) := x_1^4 - 3x_1^2 + 2 + 2x_2^2$. Wir betrachten das Newton-Verfahren zur Minimierung der Funktion $f(x)$ mit der folgenden Modifikation: Falls die Hessematrix $\nabla^2 f(x_k)$ nicht positiv definit ist, so soll die modifizierte Newton-Richtung

$$s_k := -(\nabla^2 f(x_k) + \mu_k I)^{-1} \nabla f(x_k)$$

verwendet werden, wobei I die Einheitsmatrix im $\mathbb{R}^2$ bezeichnet. Hierbei soll μ_k so gewählt werden, dass die Matrix $\nabla^2 f(x_k) + \mu_k I$ positiv definit ist. Hierzu wählen wir μ_k so, dass gilt:

$$\mu_k \geq \mu + \max\{0, -\lambda_{\min}(\nabla^2 f(x_k))\},$$

mit einer Konstante $\mu > 0$.

(a) Berechnen Sie die ersten beiden Schritte dieses Verfahrens mit dem Startpunkt $x_0 = (\frac{1}{2}, 1)^T$. Verwenden Sie zur Wahl von μ_k die Konstante $\mu = 1$ und bestimmen Sie die Schrittweiten nach der Armijo-Regel mit den Parametern $\gamma = \frac{1}{4}$ und $\beta = \frac{1}{2}$.

(b) Skizzieren Sie die Höhenlinien von f und zeichnen Sie im Startpunkt x_0 die klassische Newton-Richtung, den negativen Gradienten und die Richtungen aus Aufgabenteil (a) ein.

Aufgabe H2 (Inexaktes CG-Newton-Verfahren) (10 Punkte)

Wir betrachten inexakte Newton-Verfahren zur Minimierung einer zweimal stetig differenzierbaren Funktion $f : \mathbb{R}^n \to \mathbb{R}$. Um eine inexakte Lösung der Newton-Gleichung zu berechnen, kann man das CG-Verfahren (siehe Gruppenübung) verwenden, welches für symmetrische, positiv definite Matrizen $C \in \mathbb{R}^{n\times n}$, die Lösung eines Gleichungssystems $Cy = -c$ liefert. Hierbei wird in höchstens n Schritten die streng konvexe Funktion $p(y) = c^T y + \frac{1}{2} y^T C y$ minimiert.
Die Idee ist nun, das CG-Verfahren auf die Funktion $q_k(s) := \nabla f(x_k)^T s + \frac{1}{2} s^T \nabla^2 f(x_k) s$ anzuwenden und abzubrechen, wenn das Residuum $\|\nabla f(x_k) + \nabla^2 f(x_k)s\|_2$ klein genug ist. Damit erhalten wir eine näherungsweise Lösung der Newton-Gleichung $\nabla^2 f(x_k)s = -\nabla f(x_k)$.
Im allgemeinen Abstiegsverfahren Algorithmus 3 verwenden wir also zur Berechnung der Suchrichtung s_k das folgende modifizierte CG-Verfahren:

Inexaktes CG-Newton-Verfahren zur Bestimmung der Suchrichtung:
Seien $\alpha, \nu \in (0, 1)$ beliebig fest.

1. Wähle $y_0 = 0$, setze $g_0 = \nabla f(x_k)$ und $d_0 := \nabla f(x_k)$, $j := 0$

2. Falls $\|g_j\|_2 \le \min\{\nu, \|\nabla f(x_k)\|_2\}\|\nabla f(x_k)\|_2$: (relatives Residuum klein genug)
   STOP mit $s_k = y_j$.

   Falls $d_j^T \nabla^2 f(x_k) d_j \le 0$: (Richtung nichtpositiver Krümmung)
   STOP mit Ergebnis $s_k = y_j - sgn(\nabla f(x_k)^T d_j)\|\nabla f(x_k)\|_2 \frac{d_j}{\|d_j\|_2}$.

   Sonst: Berechne $\alpha_j = \frac{g_j^T g_j}{d_j^T \nabla^2 f(x_k) d_j}$ setze $y_{j+1} = y_j - \alpha_j d_j$ und $g_{j+1} := g_j - \alpha_j \nabla^2 f(x_k) d_j$.

   Falls $-\nabla f(x_k)^T y_{j+1} < \min\{\alpha, \|\nabla f(x_k)\|_2\}\|\nabla f(x_k)\|_2 \|y_{j+1}\|_2$: (Abstiegsrichtung wird unzureichend)
   STOP mit Ergebnis $s_k = y_j$.

3. Berechne $\beta_j := \frac{g_{j+1}^T g_{j+1}}{g_j^T g_j}$ und setze $d_{j+1} := g_{j+1} + \beta_j d_j$.

4. Setze $j := j + 1$ und gehe nach 2.

Zur **Bestimmung der Schrittweite** werde die Armijo-Regel mit Parametern $\gamma \in (0, \frac{1}{2})$ und $\beta \in (0, 1)$ verwendet. Sei nun $x_0 \in \mathbb{R}^n$ und die Niveaumenge $N_f(x_0)$ kompakt. Ausserdem dürfen Sie folgende Ungleichungen ohne Beweis verwenden:

$$-\nabla f(x_k)^T s_k \ge \frac{\|\nabla f(x_k)\|_2^2}{1 + 2\|\nabla^2 f(x_k)\|_2} \tag{1}$$

und

$$-\nabla f(x_k)^T y_j \ge \frac{\|\nabla f(x_k)\|_2^2}{1 + 2\|\nabla^2 f(x_k)\|_2}. \tag{2}$$

Zeigen Sie:

(a) Es gilt:$\|s_k\|_2 \ge \delta \|\nabla f(x_k)\|_2$ und $\|y_j\|_2 \ge \delta \|\nabla f(x_k)\|_2$, mit einem $\delta > 0$.
Hinweis: Nutzen Sie die Kompaktheit der Niveaumenge $N_f(x_0)$ und die Ungleichungen (1) und (2).

(b) Die erzeugten Suchrichtungen sind zulässig. (Zeigen Sie, dass die verallgemeinerte Winkelbedingung erfüllt ist.)
Hinweis: Nutzen Sie die Ergebnisse aus a).

(c) Die mit der Armijo-Regel erzeugten Schrittweiten σ_k sind zulässig.

(d) Ist $\nabla^2 f(\bar{x})$ positiv definit und gilt $x_k \to \bar{x}$, so konvergiert $x_k \to \bar{x}$ Q-superlinear oder sogar Q-quadratisch, falls $\nabla^2 f(x)$ in einer Umgebung von $\bar{x}$ lokal Lipschitz-stetig ist .
Hinweis: Zeigen Sie, dass es ein $K > 0$ gibt, so dass das inexakte Newton-CG-Verfahren in allen Iterationen $k \ge K$ des Abstiegsverfahrens nur abbricht, wenn das Residuum klein genug ist – mit Hilfe von Ungleichung (2). Zeigen Sie nun, dass die Voraussetzungen des Satzes 2.9.1 ii) bzw. iii) für $F(x) = \nabla f(x)$ erfüllt sind.

(e) Führen Sie jeweils zwei Schritte des Abstiegsverfahrens mit der obigen Suchrichtungsbestimmung für die Funktionen $f_1(x) = x_1^2 + 10x_2^2$ mit Startpunkt $x_0 = (10, 20)$ und $f_2(x) = x_1^4 - 2x_2^2 + x_2^4$, mit Startpunkt $x_0 = (\frac{1}{2}, \frac{1}{2})$ durch.