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

TECHNISCHE
UNIVERSITÄT
DARMSTADT

WS 2010/2011
10. Dezember 2010

7. Übungsblatt zur „Nichtlinearen Optimierung“

Gruppenübung

Aufgabe G1 (Inverser BFGS-Update)

(a) Zeigen Sie, dass die inverse BFGS-Aufdatierung (siehe Skript, Gleichung (2.63)) auch in der Form

$$B_{k+1}^{BFGS} = V_k^T B_k V_k + \rho_k d_k d_k^T$$

geschrieben werden kann, mit $V_k = I - \rho_k y_k d_k^T$ und $\rho_k = \frac{1}{d_k^T y_k}$.

(b) Zur Berechnung der Suchrichtung $s_k = -B_k \nabla f(x_k)$ betrachten wir das folgende rekursive Verfahren:

Algorithmus $v = \texttt{bfgsrek}(k, w)$

1. Falls $k = 0$: STOP mit Ergebnis $v = B_0 w$.
2. Berechne $\rho = \frac{1}{d_{k-1}^T y_{k-1}}$, $\alpha = \rho d_{k-1}^T w$. Setze $w_1 = w - \alpha y_{k-1}$.
3. Berechne $w_2 = \texttt{bfgsrek}(k-1, w_1)$.
4. STOP mit Ergebnis $v = w_2 + (\alpha - \rho y_{k-1}^T w_2) d_{k-1}$.

Zeigen Sie, dass der Aufruf $v = \texttt{bfgsrek}(k, w)$ das Ergebnis $v = B_k w$ liefert, wobei B_k die k-te inverse BFGS-Matrix ist.

Aufgabe G2 (Invertierbarkeit von DFP- und BFGS-Updates)

Sei H_k symmetrisch und invertierbar. Zeigen Sie:
Gilt $y_k^T d_k \neq 0$, $d_k^T H_k d_k \neq 0$ und $y_k^T H_k^{-1} y_k \neq 0$, so sind H_{k+1}^{DFP} sowie H_{k+1}^{BFGS} invertierbar und es gilt

$$(H_{k+1}^{DFP})^{-1} = \Phi^{BFGS}(H_k^{-1}, y_k, d_k)$$

und

$$(H_{k+1}^{BFGS})^{-1} = \Phi^{DFP}(H_k^{-1}, y_k, d_k).$$

Hinweis:
Wegen $y_k^T d_k \neq 0$ lässt sich jeder Vektor schreiben als $v = u + \lambda d_k$, mit $u \perp y_k$. Berechnen Sie nun zunächst $H_{k+1}^{DFP} v$, um die erste Gleichung zu zeigen. Nutzen Sie ausserdem, dass H_{k+1}^{DFP} die Quasi-Newton-Gleichung erfüllt.

Hausübung

Aufgabe H1 (BFGS-Verfahren mit exakter Schrittweitensuche) (6 Punkte)

Sei

$$f(x_1, x_2) := \frac{1}{2}(x_1^2 - 4x_1x_2 + 8x_2^2) - 4x_2.$$

Wenden Sie das BFGS-Verfahren mit exakter-Schrittweitensuche und den Startwerten

$$x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad H_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

zur Minimierung von f an. Bestätigen Sie, dass $H_2 = \nabla^2 f$.

Zeigen Sie weiter, dass H_{k+1} nicht nur die Quasi–Newton–Gleichung sondern zudem

$$(*) \qquad H_{k+1}d_j = y_j \ , \quad j = 0, ..., k, \ \ k = 0, ..., n-1,$$

erfüllt, wobei $y_j = \nabla f(x_{j+1}) - \nabla f(x_j)$, $d_j = x_{j+1} - x_j$ und $n = 2$.

Bemerkung: Allgemein gilt $(*)$ für streng konvexe, quadratische Funktionen, sowie $H_n = \nabla^2 f$, und nach spätestens n Schritten hat man das Optimum erreicht.

Aufgabe H2 (Der Cauchy-Punkt) (6 Punkte)

In Trust-Region-Verfahren muss das Trust-Region-Problem nur "hinreichend gut" gelöst werden. Eine der einfachsten Näherungslösungen für das Problem erhält man, in dem man sich nur auf den Strahl in Richtung des steilsten Abstiegs beschränkt. Die Lösung des folgenden (eindimensionalen) Minimierungsproblems

$$\min q_k(s) := \nabla f(x_k)^T s + \frac{1}{2} s^T H_k s \quad \text{u.d.N.} \quad s = -t\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}, \quad t \in [0, \Delta_k],$$

mit $\Delta_k > 0$ nennt man *Cauchy-Punkt.* Wir wollen im folgenden die Bestimmung des Cauchy-Punkts näher untersuchen.

(a) Hierzu betrachten wir zunächst die Funktion

$$\phi : \mathbb{R} \mapsto \mathbb{R}, \quad \phi(t) := \alpha t + \beta t^2, \quad \alpha, \beta \in \mathbb{R}, \alpha < 0.$$

Zeigen Sie, dass das Problem

$$\min \phi(t) \quad \text{u.d.N.} \quad 0 \leq t \leq \tau$$

für jedes $\tau > 0$ genau eine Lösung t^* besitzt und dass folgende Abschätzung gilt:

$$\phi(t^*) \leq \frac{\alpha}{2} \min\{\frac{|\alpha|}{2|\beta|}, \tau\}.$$

Interpretieren Sie hierbei für $\beta = 0$ das erste Argument als $+\infty$.

(b) Wenden Sie a) nun mit $\phi(t) = q_k(-t\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|})$ und geeignetem $\tau > 0$ an, um den Cauchy-Punkt zu berechnen und zeigen Sie:

$$q_k(s_k^c) \leq -\frac{\|\nabla f(x_k)\|}{2} \min\{\frac{\|\nabla f(x_k)\|}{\|H_k\|}, \Delta_k\}.$$