TU München
Lehrstuhl Mathematische Optimierung
Prof. Dr. M. Ulbrich
Dipl.-Math. B. von Loesch

Sommersemester 2007

Blatt 9

Übungen zur Optimierung 3

H29. Inverses BFGS-Verfahren (ca. 6 Punkte)

a) Zeigen Sie, dass die inverse BFGS-Aufdatierung 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 $\rho_k = 1/({d^k}^T y^k)$.

b) Zur Berechnung der Suchrichtung $s^k = -B_k \nabla f(x^k)$ betrachten wir das folgende rekursive Verfahren:

Algorithmus $v = \mathsf{bfgsrec}(k, w)$

1. Falls $k = 0$, kehre zurück mit Ergebnis $v = B_0 w$.
2. Berechne $\rho = 1/({d^{k-1}}^T y^{k-1})$, $\alpha = \rho {d^{k-1}}^T w$ und setze $w^1 = w - \alpha y^{k-1}$.
3. Berechne $w^2 = \mathsf{bfgsrec}(k-1, w^1)$.
4. Kehre zurück mit Ergebnis $v = w^2 + (\alpha - \rho {y^{k-1}}^T w^2) d^{k-1}$.

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

H30. Powell-Symmetric-Broyden-Formel (ca. 8 Punkte)

In der Vorlesung wurde begründet, dass es bei der Entwicklung von Quasi-Newton-Verfahren sinnvoll ist, nach Aufdatierungsformeln zu suchen, die das folgende Optimierungsproblem lösen:

$$\min_{H \in \mathbb{R}^{n \times n}} \|H - H_k\|_* \quad \text{u.d.N.} \quad H = H^T, \;\; H d^k = y^k.$$

Hierbei sind die symmetrische Matrix $H_k \in \mathbb{R}^{n \times n}$, die Vektoren $d^k, y^k \in \mathbb{R}^n$ und die Matrix-Norm $\|\cdot\|_*$ gegeben. Je nach Wahl von $\|\cdot\|_*$ ergeben sich unterschiedliche Aufdatierungsformeln. Wir betrachten nun speziell die Frobenius-Norm

$$\|A\|_F = \sqrt{\langle A, A \rangle} \quad \text{mit} \quad \langle A, B \rangle = \sum_{i,j=1}^{n} A_{ij} B_{ij}.$$

Das Problem lautet dann

$$\min_{H \in \mathbb{R}^{n \times n}} \frac{1}{2} \|H - H_k\|_F^2 \quad \text{u.d.N.} \quad H = H^T, \;\; H d^k = y^k, \qquad (*)$$

wobei wir o.E. die Zielfunktion quadriert und mit $1/2$ multipliziert haben.

Ziel der Aufgabe ist es, zu zeigen, dass die Lösung von $(*)$ gegeben ist durch die

Powell-Symmetric-Broyden- (PSB-) Formel:

$$H_{k+1}^{PSB} = H_k + \frac{(y^k - H_k d^k){d^k}^T + d^k (y^k - H_k d^k)^T}{\|d^k\|^2} - \frac{(y^k - H_k d^k)^T d^k}{\|d^k\|^4} d^k {d^k}^T.$$

– bitte wenden –

a) Zeigen Sie, dass der zulässige Bereich $\mathcal{H} = \{H \in \mathbb{R}^{n\times n} : H = H^T,\ Hd^k = y^k\}$ von $(*)$ abgeschlossen und konvex ist.

b) Berechnen Sie den Gradienten der Zielfunktion $f : H \in \mathbb{R}^{n\times n} \mapsto \frac{1}{2}\|H - H_k\|_F^2$ und begründen Sie, dass f quadratisch und streng konvex ist.

Bemerkung: Da H eine Matrix ist, lässt sich $\nabla f(H)$ am einfachsten ebenfalls als Matrix darstellen mit $(\nabla f(H))_{ij} = \frac{d}{dH_{ij}} f(H)$.

c) Verwenden Sie Sätze über konvexe Funktionen aus der Vorlesung, um zu begründen, dass $(*)$ genau eine Lösung H_* besitzt und dass diese eindeutig charakterisiert ist durch

$$H_* \in \mathcal{H}, \quad \langle \nabla f(H_*), H - H_* \rangle \geq 0 \quad \forall\, H \in \mathcal{H}. \tag{$\bullet$}$$

Begründen Sie weiter, dass gilt:

$$\{H - H_* : H \in \mathcal{H}\} = \{M \in \mathbb{R}^{n\times n} : M = M^T,\ Md^k = 0\}.$$

d) Zeigen Sie, dass $H_* = H_{k+1}^{PSB}$ der Bedingung ($\bullet$) genügt. Warum ist also H_{k+1}^{PSB} die eindeutige Lösung von $(*)$? **Tip:** Benutzen Sie $\langle A, uv^T \rangle = u^T Av$.

H31. Der Cauchy-Punkt (ca. 5 Punkte)

Bei Trust-Region-Verfahren wird in jeder Iteration ein Problem der Form

$$\min \ \ q(s) \quad \text{u.d.N} \quad \|s\| \leq \Delta$$

(approximativ) gelöst. Hierbei ist $\Delta > 0$ und $q : \mathbb{R}^n \to \mathbb{R}$ ist eine quadratische Funktion:

$$q(s) = g^T s + \frac{1}{2} s^T H s, \quad g \in \mathbb{R}^n \setminus \{0\},\ \ H \in \mathbb{R}^{n\times n} \ \text{symmetrisch.}$$

Der Cauchy-Punkt $s_c \in \mathbb{R}^n$ ist definiert als Lösung des Problems

$$\min \ \ q(s) \quad \text{u.d.N} \quad s = -tg,\ \ t \geq 0,\ \ \|s\| \leq \Delta. \tag{CP}$$

a) Wir betrachten zunächst die Funktion

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

Zeigen Sie, dass das Problem

$$\min \ \ \varphi(t) \quad \text{u.d.N} \quad 0 \leq t \leq \tau \tag{P$_\varphi$}$$

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

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

Hierbei interpretieren wir im Fall $\beta = 0$ das erste Argument von min als $+\infty$.

b) Wenden Sie nun a) mit $\varphi(t) = q(-tg)$ und geeignetem $\tau > 0$ an, um den Cauchy-Punkt s_c zu berechnen, und zeigen Sie:

$$q(s_c) \leq -\frac{\|g\|}{2} \min\left\{ \frac{\|g\|}{\|H\|}, \Delta \right\}.$$

Abgabe: Mittwoch, 20.06.2007 bis 14:15 Uhr im Briefkasten Optimierung 3 im Untergeschoss
Bitte schreiben Sie neben **Namen** und **Matrikelnummer** auch die **Nummer Ihrer Übungsgruppe** auf die Abgabe.