Universität Konstanz
Fachbereich Mathematik und Statistik
Prof. Dr. Stefan Volkwein
Tim Seger

28. Mai 2009

Optimierung
3. Übungsblatt

Alle Aufgaben sind zum Vortragen in den Übungen am Dienstag, 09. Juni 2009, vorzubereiten. Die mit ✏ gekenntzeichneten Aufgaben sind auch schriftlich auszuarbeiten und bis Donnerstag, 04. Juni 2009, in den Briefkasten Nr. 16 abzugeben.

Aufgabe 3.1 ✏ Sei $f : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar. Sei $d_k \in \mathbb{R}^n$ eine Abstiegsrichtung am Punkt $x_k \in \mathbb{R}^n$ und vorausgesetzt, dass f auf dem Strahl $\{x_k + td_k \mid t > 0\}$ nach unten beschränkt ist. Zeigen Sie, dass im Fall $0 < \alpha < \varrho < 1$ offene Intervalle für die Schrittweite t existieren, sodass dort die Wolfe-Powell-Bedingungen

$$f(x_k + td_k) \le f(x_k) + \alpha t \nabla f(x_k)^T d_k \tag{1}$$

$$\nabla f(x_k + td_k)^T d_k \ge \varrho \nabla f(x_k)^T d_k \tag{2}$$

bzw. die strengen Wolfe-Powell-Bedingungen

$$\begin{aligned} f(x_k + td_k) &\le f(x_k) + \alpha t \nabla f(x_k)^T d_k \\ \left|\nabla f(x_k + td_k)^T d_k\right| &\le \varrho \left|\nabla f(x_k)^T d_k\right| \end{aligned}$$

gelten[1].

Aufgabe 3.2 Gegeben sei die Funktion $\varphi(t) := f(x_k + td_k)$ mit einer Abstiegsrichtung $d_k \in \mathbb{R}^n$. Ferner sei die Bedingung (1) an der Stelle $t = t_0 > 0$ verletzt.

a) Zeigen Sie, dass das quadratische Interpolationspolynom zu den Punkten $\varphi(0)$, $\varphi'(0)$, $\varphi(t_0)$ gegeben ist durch

$$\Phi_q(t) = \left(\frac{\varphi(t_0) - \varphi(0) - t_0\varphi'(0)}{t_0^2}\right) t^2 + \varphi'(0)t + \varphi(0).$$

b) Zeigen Sie, dass Φ_q eine positive Krümmung hat und dass der Minimierer t_1 von Φ_q die folgende Ungleichung erfüllt[2]:

$$t_1 < \frac{t_0}{2(1-\varrho)}.$$

Aufgabe 3.3 Das Minimum von $f(x) := x_1^2 + x_2^2 - x_1x_2 - 3x_1$ $(x \in \mathbb{R}^2)$ kann leicht durch Differenzieren bestimmt werden. Eine andere Methode ist die der alternierenden Richtungen: Man minimiert eindimensional abwechselnd bzgl. x_1 und x_2. Zeigen Sie, dass die so, aus dem Startwert $x_0 := (0,0)^T$ erhaltene Folge gegen den Minimierer von f konvergiert.

---

[1] Hinweis: Mittelwertsatz

[2] Dies liefert eine grobe Abschätzung der neuen, zulässigen Schrittweite.

Aufgabe 3.4 ✏

a) Sei $\mathcal{H}$ ein reeller Hilbertraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Seien $b \in \mathcal{H}$ und $A : \mathcal{H} \to \mathcal{H}$ eine lineare stetige Abbildung. Betrachten Sie die Abbildung

$$\varphi(x) := \|Ax - b\|_2 \quad (x \in \mathcal{H}).$$

Zeigen Sie möglichst „elementar“[3]: $u \in \mathcal{H}$ ist *genau dann* eine Minimalstelle von φ, wenn die Gauss'schen Normalengleichungen

$$A^*Au = A^*b$$

gelten. Hierbei ist A^* die adjungierte Abbildung - d.h. $\langle Ax, y\rangle = \langle x, A^*y\rangle \ (x, y \in \mathcal{H})$.

b) Lösen Sie mithilfe von a) das Ausgleichsproblem: Gesucht ist eine Gerade $\gamma(t) = x_1 + x_2 t$ so, dass der quadratische Abstand zu den Messpunkten $\sum_{i=1}^{5}(\gamma_i - \gamma(t_i))^2$ minimal wird. Die Messpunkte sind:

t_i	1975	1980	1985	1990	1995
γ_i	30	35	38	42	44

Aufgabe 3.5 Programmieraufgabe (in Zweiergruppen zu bearbeiten)

Implementieren Sie den Schrittweitenalgorithmus (3.11) nach Wolfe-Powell aus der Vorlesung in einer Funktion `wolfepowell.m`[4]. Ändern Sie ferner Ihre Funktion `gradiverf.m` aus Aufgabe 2.4 so ab, dass die Schrittweitenwahl nun auch mit dem Wolfe-Powell-Algorithmus getroffen werden kann. Wählen Sie für die im Algorithmus auftretenden Parameter `gamma = 2, alpha` $= 10^{-4}$ und `rho = 0.9`.

Zur Wahl von $t^{(j)} \in \left[t_1^{(j)} + \tau_1\Delta^{(j)}, t_2^{(j)} - \tau_2\Delta^{(j)}\right]$ beim Punkt (B.1) des Algorithmus verwenden Sie die folgende Strategie:

- Falls der Term $2(\varphi(t) - \varphi(0) - t\varphi'(0))$, der in der quadratischen Interpolierenden aus Aufgabe 3.2 auftritt, größer als ein `epsilon` > 0 ist, nehme fr t das in 3.2. berechnete Minimum von Φ_q.
- Ist $2(\varphi(t) - \varphi(0) - t\varphi'(0)) <$ `epsilon`, so nehme den Intervallmittelpunkt fr t.

Testen Sie das Gradientenverfahren mit der neuen Schrittweitenwahl an der Rosenbrock-Funktion mit den Startpunkten $x_0 = (1.2, 1.2)^T$ und $x_0 = (-1.2, 1)^T$. Verwenden Sie als Anfangsschrittweite $t_0 = 1$.

Während des Verfahrens sollen der Laufindex k, der aktuelle Iterationspunkt x_k, die aktuelle Schrittweite t_k sowie die Werte $f(x_k)$ und $\|\nabla f(x_k)\|_2$ in eine Textdatei `Dokumentation.txt` ausgegeben werden. Hilfreich sind hierbei die Befehle

```
fid=fopen('Dokumentation.txt','w')
```

zum Erstellen der Textdatei und

```
fprintf(fid,'...')
```

zum Schreiben von Text bzw. Eintragen von Werten. Geben Sie ferner den letzten Datensatz auch auf den Bildschirm aus.

Zum Vergleich mit der Armijo-Schrittweitenwahl können Sie das Verfahren auch mit dieser laufen lassen, die jeweiligen Datensätze in zwei Dateien speichern und später vergleichen.

[3] d.h. ohne Verwendung von Differenzierbarkeit und notwendigen bzw. hinreichenden Kriterien

[4] Hinweis: Zur Realisierung des Algorithmus bieten sich zwei separate while-Schleifen und zwei Variablen an, die festlegen, ob diese Schleifen durchlaufen werden an