Nichtlineare Optimierung
2. Rechnerübungsblatt

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Mathematik
Prof. Dr. Stefan Ulbrich
Hannes Meinlschmidt

WS 2011-2012
25.11.2011

Rechnerübung

Aufgabe R1 (Powell-Wolfe-Schrittweitenregel)
Implementieren Sie Algorithmus 4 aus der Vorlesung zur Berechnung der Powell-Wolfe-Schrittweite. Erstellen Sie dazu eine Funktion

```
[sig] = PowellWolfe(xk,sk,stg,fg,fk,gamma,theta).
```

Ändern Sie Ihr Gradientenverfahren aus der ersten Übung so, dass bei der Wahl von `stepsize` = 2 die Powell-Wolfe-Schrittweite gewählt wird. Testen Sie Ihr Verfahren an den Funktionen aus der ersten Rechnerübung.

Aufgabe R2 (Globalisiertes Newton-Verfahren)

(a) Implementieren Sie das globalisierte Newton-Verfahren (Algorithmus 7 der Vorlesung) in Matlab. Verwenden Sie

$$B_k = I, \quad c_1 = 10^{-3}, \quad c_2 = 10^{-1} \quad \text{und} \quad p = 1.$$

Zur Schrittweitenbestimmung bietet sich Ihre Funktion `armijo` aus der ersten Rechnerübung an. Beachten Sie, dass die Schrittweiten-Bestimmung nach Armijo für diesen Algorithmus mit $\gamma \in (0, 1/2)$ statt $\gamma \in (0, 1)$ aufgerufen werden soll (dies garantiert den Übergang zu schneller lokaler Konvergenz!).
Weiterhin müssen die aufgerufenen MATLAB-Funktionen `fg`, die die Funktionswerte und Gradienten der zu minimierenden Funktionen ausgeben, nun auch die Hessematrix liefern, d.h. `[f,g,H] = fg(x)` soll Funktionswert `f`, Gradient `g` und Hessematrix `H` der Funktion `fg` im Punkt x geben.
Verwenden Sie für Ihr Programm wieder einen Eingabeparameter `maxit`, so dass Ihr Verfahren spätestens nach `maxit` Iterationen abbricht.

(b) Testen Sie Ihr Programm an den Funktionen f_1 und f_2 aus der ersten Rechnerübung, sowie

$$f_4(x) = \frac{x^4}{4} - \frac{x^2}{2} \quad \text{mit Startwerten } x_0 \in \left\{ \pm 2, \pm 0.51, \pm\frac{1}{\sqrt{5}}, \pm\frac{1}{\sqrt{3}} \right\}$$

.

Hausübung

Aufgabe H1 (Gauss-Newton-Verfahren) (10 Punkte)
Algorithmus 7 aus der Vorlesung beschreibt die Globalisierung des Newton-Verfahrens für Minimierungsprobleme.
Die Globalisierung des Newton-Verfahrens für *Gleichungssysteme* $F(x) = 0$ mit einer zwei mal stetig differenzierbaren Funktion $F : \mathbb{R}^n \to \mathbb{R}^n$ erfolgt üblicherweise auf Basis der Minimierung einer Energiefunktion für F, nämlich

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} F(x)^T F(x). \tag{1}$$

(a) Bestimmen Sie die Newton-Gleichung von (1).

(b) Im Gauss-Newton-Verfahren bestimmt man die Suchrichtung s_k als Lösung der Gauss-Newton Gleichung

$$F'(x_k)^T F'(x_k) s_k = -F'(x_k)^T F(x_k). \tag{GN}$$

Welcher Term wurde hier im Vergleich mit der Newton-Gleichung für (1) vernachlässigt? Zeigen Sie, dass die Gauss-Newton-Gleichung (GN) zur klassischen Newton-Gleichung für $F(x) = 0$ äquivalent ist, wenn $F'(x_k)$ invertierbar ist.

(c) Sei $\bar{x}$ eine Nullstelle von F und $F'(\bar{x})$ invertierbar. Zeigen Sie, dass (GN) auf ein Newton-artiges Verfahren für das Problem (1) führt, welches für $x_k \to \bar{x}$ die Dennis-Moré-Bedingung erfüllt.

(d) Verwendet man das globalisierte Newton-artige Verfahren (Algorithmus 10 in der Vorlesung) für (1) mit der Matrix $M_k = F'(x_k)^T F'(x_k)$, so nennt man das Verfahren *globalisiertes Gauss-Newton-Verfahren*.

Die Niveaumenge $N_f(x_0)$ zum Startpunkt x_0 sei kompakt. Zeigen Sie mit Sätzen aus der Vorlesung:

i. Jeder Häufungspunkt $\bar{x}$ von (x_k) erfüllt $F'(\bar{x})^T F(\bar{x}) = 0$. Ist $F'(\bar{x})$ invertierbar, so gilt zudem $F(\bar{x}) = 0$.

ii. Hat (x_k) einen Häufungspunkt $\bar{x}$, in dem $F'(\bar{x})$ invertierbar ist, dann konvergiert (x_k) Q-superlinear gegen $\bar{x}$. Ist F' lokal Lipschitz-stetig, so ist die Konvergenz sogar Q-quadratisch.

Aufgabe H2 (CG-Verfahren, oder: Verfahren der konjugierten Gradienten) (10 Punkte)

Gegeben sei die quadratische Funktion

$$q : \mathbb{R}^n \to \mathbb{R} : y \mapsto c^T y + \frac{1}{2} y^T C y$$

mit $c \in \mathbb{R}^n$ und $C \in \mathbb{R}^{n,n}$ symmetrisch so, dass q streng konvex ist.

Zur Bestimmung des eindeutigen globalen Minimums von q betrachten wir folgenden Algorithmus 1:

Algorithmus 1 : Verfahren der konjugierten Gradienten

1 Wähle y_0 und berechne $g_0 := c + C y_0$;
2 **if** $g_0 = 0$ **then**
3 STOP mit Ergebnis y_0.
4 **else**
5 Setze $k \leftarrow 0$ und $d_0 = g_0$;
6 **end**
7 Berechne $\alpha_k := \frac{g_k^T g_k}{d_k^T C d_k}$;
8 Setze $y_{k+1} := y_k - \alpha_k d_k$ sowie $g_{k+1} := g_k - \alpha_k C d_k$;
9 **if** $g_{k+1} = 0$ **then**
10 STOP mit Ergebnis y_{k+1}.
11 **end**
12 Berechne $\beta_k := \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}$;
13 Setze $d_{k+1} := g_{k+1} + \beta_k d_k$;
14 Setze $k \leftarrow k + 1$ und gehe zu 7.

Sei V_{k+1} definiert als $\operatorname{Span}\{g_0, C g_0, \ldots, C^k g_0\}$. Zeigen Sie:

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

i. $d_k \neq 0$,

ii. $V_{k+1} = \operatorname{Span}\{g_0, \ldots, g_k\} = \operatorname{Span}\{d_0, \ldots, d_k\}$,

iii. die Vektoren $d_0, \ldots d_k$ sind paarweise C-konjugiert, d.h.

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

iv. g_{k+1} ist orthogonal zum Unterraum V_{k+1}, also $g_{k+1} \perp V_{k+1}$.

Hinweis: Gehen Sie induktiv vor und behalten Sie den Überblick.

(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.

Hinweis: Betrachten Sie $\{d_0, \ldots, d_k\}$ als Basis von V_{k+1} und formulieren Sie das Minimierungsproblem auf dem Unterraum um in ein Minimierungsproblem über $\mathbb{R}^m$ für ein geeignetes m. Nutzen Sie dann die Konstruktion der y_k aus dem Algorithmus und die in (a) bewiesenen Eigenschaften.