Universität Hamburg
Fachbereich Mathematik
Dr. Matthias Voigt, Dr. Heiko Kröner

Sommersemester 2019

Optimierung – 5. Übungsblatt.

Bitte bearbeiten Sie die Übungsaufgaben in festen Zweiergruppen.
Abgabetermin: *Donnerstag, 16.05.2019 vor der Übung. Bitte senden Sie die Lösung der Programmieraufgabe bis zum 15.05.2019, 18:00 Uhr an* `heiko.kroener@uni-hamburg.de`*. Alle zur Ausführung nötigen Dateien sollten in einem Archiv mit dem Namen* `opt_blattnr_aufgabennr_name1_name2.zip` *enthalten sein.*

Aufgabe 1 (6 Punkte): Es sei $\{A_k\}$ eine Folge von Matrizen in $\mathbb{R}^{n\times n}$, die in einem Quasi-Newton-Verfahren mit BFGS-Update berechnet wurde. Zudem seien die Vektoren s^k, $y^k \in \mathbb{R}^n$ für $k = 0, 1, 2, \ldots$ wie in der Vorlesung definiert. Beweisen Sie die folgenden Aussagen:

a) Ist A_k symmetrisch und positiv definit, so ist die Matrix

$$\widetilde{A}_k := A_k - \frac{\left(A_k s^k\right)\left(A_k s^k\right)^{\mathsf{T}}}{\left\langle s^k, A_k s^k\right\rangle}$$

symmetrisch und postiv *semi*definit.

b) Die Matrix

$$A_{k+1} := A_k + \frac{y^k\left(y^k\right)^{\mathsf{T}}}{\left\langle y^k, s^k\right\rangle}$$

ist symmetrisch und positiv definit genau dann, wenn $\left\langle y^k, s^k\right\rangle > 0$ gilt.

Aufgabe 2 (6 Punkte): Oftmals ist es besser anstelle einer Approximation A_k der Hesse-Matrix $f''\left(x^k\right)$ eine Approximation B_k der inversen Hesse-Matrix $f''\left(x^k\right)^{-1}$ aufzubauen. Es seien s^k, $y^k \in \mathbb{R}^n$ für $k = 0, 1, 2, \ldots$ wie in der Vorlesung definiert.

a) Es sei $B_k = A_k^{-1}$. Zeigen Sie, daß dann die Matrix

$$B_{k+1} := \left(I_n - \frac{s^k\left(y^k\right)^{\mathsf{T}}}{\left\langle y^k, s^k\right\rangle}\right) B_k \left(I_n - \frac{s^k\left(y^k\right)^{\mathsf{T}}}{\left\langle y^k, s^k\right\rangle}\right) + \frac{s^k\left(s^k\right)^{\mathsf{T}}}{\left\langle y^k, s^k\right\rangle}$$

die Inverse von A_{k+1} ist.

b) Es sei nun die Folge $\{B_k\}$ von Matrizen mit $B_0 = \left\|f\left(x^0\right)\right\|^{-1} I_n$ definiert wie oben und n sei groß. Beschreiben Sie, wie man die Matrizen B_1, B_2, $\ldots$ effizient speichern und Matrix-Vektor-Produkte $B_k z$ für $z \in \mathbb{R}^n$ effizient auswerten kann.

c) Motivieren aus Ihren Überlegungen in b) die Verwendung des „limited memory BFGS“. Dabei wird für ein festes $m > 1$ folgende Konstruktion ausgeführt:

- Für $k \leq m$ werden die Matrizen B_k wie in a) konstruiert mit $B_0 = \left\|f\left(x^0\right)\right\|^{-1} I_n$.
- Für $k > m$ werden die Matrizen B_k wie in a) konstruiert, aber mit $B_{k-m} = \left\|f\left(x^{k-m}\right)\right\|^{-1} I_n$.

D. h. zur Konstruktion von B_k müssen sich nur jeweils die letzten m Vektoren $s^{k-m}, \ldots, s^{k-1}$ und $y^{k-m}, \ldots, y^{k-1}$ „gemerkt“ werden.

Aufgabe 3 (8 Punkte): Implementieren Sie ein Quasi-Newton-Verfahren mit BFGS-Update und Powell-Schrittweitenregel.

a) Implementieren Sie zunächst die Schrittweitenregel in einer `MATLAB`-Funktion mit dem Interface

```
sigmaP = powell( x, fun, grad, d, sigma0, beta, delta )
```

und den folgenden Argumenten:

- `x`: Startpunkt x der Iteration,
- `fun, grad`: zwei Function Handles zur Auswertung der Funktionen f und ∇f,
- `d`: die Suchrichtung,
- `sigma0`: die Startschrittweite,
- `beta, delta`: die Parameter β und δ, die das Powell-Verfahren beschreiben (siehe Vorlesung),
- `sigmaP`: eine Schrittweite, die die Powell-Bedingungen erfüllt.

b) Implementieren Sie nun eine `MATLAB`-Funktion für das Quasi-Newton-Verfahren. Diese soll das Interface

```
[ xmin, fmin ] = bfgs( x0, fun, grad )
```

mit den folgenden Argumenten haben:

- `x0`: Startpunkt x^0 der Iteration,
- `fun, grad`: zwei Function Handles zur Auswertung der Funktionen f und ∇f,
- `xmin, fmin`: der Minimierer und der minimale Funktionswert.

Testen Sie Ihre Implementierung an der Rosenbrock-Funktion von Übungsblatt 1 mit dem Startpunkt $x^0 = \left[\begin{smallmatrix} -1.9 \\ 2 \end{smallmatrix}\right]$. Wählen Sie $A_0 = \left\|f\left(x^0\right)\right\| I_2$ (oder $B_0 = \left\|f\left(x^0\right)\right\|^{-1} I_2$). Wählen Sie für das Powell-Verfahren die Parameter $\beta = 0.9$ und $\delta = 0.1$ und als Startschrittweite $\sigma_{k,0} = 1$. Illustrieren Sie die Konvergenz des Verfahrens, indem Sie die Iterierten und die Niveaulinien der Rosenbrock-Funktion in einem gemeinsamen Plot darstellen.