Prof. Dr. C. Brandt
Evelyn Herberg Sommersemester 2018

Übungen zur Optimierung - Blatt 6

Besprechung des Übungsblattes am 31. Mai 2018

Aufgabe 1: (Das Gauß - Newton - Verfahren)

Sei $F : \mathbb{R}^n \to \mathbb{R}^m$ zweimal stetig differenzierbar, $m \geq n$ und $J := \mathbb{R}^n \to \mathbb{R}$ mit

$$J(x) = \frac{1}{2} \|F(x)\|^2 \, .$$

Zur Lösung des kleinste Quadrate-Problems

$$\min_{x \in \mathbb{R}^n} J(x) = \min_{x \in \mathbb{R}^n} \frac{1}{2} \|F(x)\|^2 \qquad \text{(KQ)}$$

verwende man das sog. lokale Gauß - Newton - Verfahren, s. Algorithmus 1.

a) Berechnen Sie $J'(x)$ und $J''(x)$.

b) Worin besteht der Unterschied zum lokalen Newton - Verfahren zur Lösung des Optimierungsproblems (KQ)?

c) Geben Sie eine konvexe quadratische Funktion $q_k : \mathbb{R}^n \to \mathbb{R}$ an, für die die Gauß - Newton - Gleichung äquivalent zu $q_k'(d^k) = 0$ ist. Zeigen Sie, dass dann d^k eine Lösung des Problems

$$\min_{d \in \mathbb{R}^n} q_k(d)$$

ist.

- Wähle Startpunkt x^0
- Für $k = 0, 1, \ldots$
  - STOP, falls $F(x^k) = 0$
  - Berechne d^k durch Lösen der Gauß - Newton - Gleichung

$$F'(x^k)^T F'(x^k) d^k + F'(x^k) F(x^k) = 0 \qquad \text{(GN)}$$

  - Setze $x^{k+1} = x^k + d^k$

Algorithm 1: Lokales Gauß - Newton - Verfahren

Aufgabe 2: (Konvergenz des Gauß - Newton - Verfahrens)

a) Möge das Verfahren eine gegen x^* konvergente Folge $\left\{x^k\right\}_k$ erzeugen und sei $F\left(x^*\right) = 0$ und $F'\left(x^*\right)$ habe vollen Spaltenrang. Interpretieren Sie das Gauß - Newton - Verfahren als ein Newton - artiges Verfahren und zeigen Sie, dass die die Konvergenz der Folge q - superlinear ist.

b) Betrachten Sie nun die Funktion $F : \mathbb{R} \to \mathbb{R}^2$,

$$F\left(x\right) = \begin{pmatrix} x \\ \frac{x^2}{4} + 1 \end{pmatrix} .$$

Bestimmen Sie die globale Lösung x^* von (KQ) und zeigen Sie, dass das Gauß - Newton - Verfahren gegen x^* konvergiert für $0 < x^0 < 2$. Bestimmen Sie zudem die Konvergenzrate.

Aufgabe 3 (doppelte Punkte): (Powell - Symmetric - Broyden - Formel)

Aufdatierungsformeln für Quasi-Newton-Verfahren können über folgende Art von Optimierungsproblemen bestimmt werden:

$$\min_{H_+ \in \mathbb{R}^{n\times n}} \left\|H_+ - H_k\right\|_* \qquad \text{u.d.N. } H_+ = H_+^T\,,\; H_+ s^k = y^k.$$

Hierbei seien $H_k \in \mathbb{R}^{n\times n}$ symmetrisch, $s^k \in \mathbb{R}^n \setminus \{0\}$, $y^k \in \mathbb{R}^n$ gegeben. Je nach Wahl der Norm $\left\|\cdot\right\|_*$ ergeben sich unterschiedliche Aufdatierungsformeln.

Im Folgenden betrachten wir die Frobeniusnorm $\|A\|_F^2 = \sum_{i,j=1}^n A_{i,j}^2 = \text{spur}\left(A^T A\right)$ und das Problem

$$\min_{H_+ \in \mathbb{R}^{n\times n}} \left\|H_+ - H_k\right\|_F^2 \qquad \text{u.d.N. } H_+ = H_+^T\,,\; H_+ s^k = y^k. \tag{*}$$

Zeigen Sie, dass die Lösung von (*) eindeutig gegeben ist durch die Powell-symmetric-Broyden-Formel (PSB):

$$H_+^{\mathsf{PSB}} = H_k + \frac{\left(y^k - H_k s^k\right) s^{k^T} + s^k \left(y^k - H_k s^k\right)^T}{\left\|s^k\right\|^2} - \frac{\left(y^k - H_k s^k\right)^T s^k}{\left\|s^k\right\|^4} s^k s^{k^T} .$$

Gehen Sie dazu schrittweise vor und zeigen Sie

a) Das Problem (*) besitzt höchstens eine Lösung.

b) H_+^{PSB} ist zulässig für das Minimierungsproblem.

c) Für jede Matrix A, welche die Nebenbedingungen erfüllt, ist

$$\left\|\left(H_+^{\mathsf{PSB}} - H_k\right) v\right\| \le \left\|\left(A - H_k\right) v\right\|$$

für jedes $v \in \mathbb{R}^n$ mit $v = s^k$ oder $\left\langle v, s^k\right\rangle = 0$.

d) Verwenden Sie die vorherige Aussage um zu beweisen, dass H_+^{PSB} das Minimierungsproblem (*) löst.

Hinweis: Für die Frobeniusnorm einer Matrix $A \in \mathbb{R}^{n \times n}$ gilt mit einer beliebigen Orthonormalbasis $\{v_1, v_2, \ldots, v_n\}$ des $\mathbb{R}^n$

$$\|A\|_F^2 = \sum_{i=1}^{n} \|Av_i\|^2 .$$

Aufgabe 4: (Globales Newton - Verfahren für Optimierungsprobleme)

Implementieren Sie das globale Newtonverfahren und verwenden Sie als Abbruchkriterium

$$\left\| J'\left(x^k\right) \right\| \leq 10^{-6}$$

bzw. brechen Sie nach maximal 200 Iterationen ab. Wählen Sie zunächst $\beta = 0.5$, $\sigma = 10^{-4}$, $t_1 = t_2 = 10^{-6}$ sowie $p = 0.1$.

Testen Sie es anhand von 3 Funktionen:

a) Die Rosenbrock - Funktion mit Startpunkt $x^0 = (-1.2, 1)^T$.

b) Die Gauß - Funktion $J : \mathbb{R}^3 \to \mathbb{R}$ ausgehend von $x^0 = (0.4, 1, 0)$. Dabei ist

$$J(x) = \sum_{i=1}^{15} \left[x_1 \exp\left(-\frac{x_2 \left(t_i - x_3\right)^2}{2} \right) - y_i \right]^2, \qquad t_i = \frac{8-i}{2}$$

und die Parameter y_i, $i = 1, \ldots, 15$ gegeben als

$i \in$	$\{1,15\}$	$\{2,14\}$	$\{3,13\}$	$\{4,12\}$	$\{5,11\}$	$\{6,10\}$	$\{7,9\}$	$\{8\}$
y_i	0.0009	0.0044	0.0175	0.0540	0.1295	0.2420	0.3521	0.3989

c) Die Wood - Funktion * $J : \mathbb{R}^4 \to \mathbb{R}$,

$$J(x) = 100\left(x_2 - x_1^2\right)^2 + (1-x_1)^2 + 90\left(x_4 - x_3^2\right)^2 + (1-x_3)^2 + 10\left(x_2 + x_4 - 2\right)^2 + 0.1\left(x_2 - x_4\right)^2$$

mit Startpunkt $(-3, -1, -3, 1)^T$.

Haben Sie ein Minimum gefunden? (Bemerkung: Das lokale Verfahren findet weder für die Gauß - noch für die Wood - Funktion ein Minimum.) Welchen Einfluss hat die Wahl der Parameter t_1, t_2, p? Ändert sich etwas an Ihrer Beobachtung für die Wood-Funktion, wenn Sie den Newton - Gleichung nicht exakt lösen sondern stattdessen eine modifizierte Cholesky - Zerlegung (mit $\mu = 10^{-6}$) verwenden?

Hinweis: Verwenden Sie die Matlab-Vorlage blatt06.m, welche Sie aus Stine herunterladen können und beachten Sie die Vorgaben. Schicken Sie Ihre Lösung per Email mit dem Betreff "Optimierung - Blatt 6" spätestens am Mittwoch, den 30.5. an Frau Herberg (evelyn.herberg(at)uni-hamburg.de). Nennen Sie ihre Datei blatt06_xy.m, wobei xy für Ihren Vornamen steht.

*Achtung: Auf dem Blatt 5 war die Woodfunktion falsch definiert. In der Matlabvorlage ist die korrigierte Formel implementiert.