Optimierung und inverse Probleme
Wintersemester 2018/19
Prof. Dr. Bastian von Harrach
M. Sc. Tim Jahn

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

Kombiniertes 6. und 7. Übungsblatt (erschienen am 20.11.2018)

Aufgabe 6.1 (Votieraufgabe)
In der Vorlesung wurde bisher die Armijo-Regel und als Erweiterung die Powell-Wolfe-Regel behandelt. Eine weitere Abänderung der Armijo-Regel ist die **Goldstein-Regel**:

Dabei wird wie bei der Armijo-Regel die tatsächliche Abnahme der Zielfunktion mit der aus der linearen Approximation erwarteten Abnahme verglichen. Es wird dabei eine Schrittweite akzeptiert, wenn die tatsächliche Abnahme einerseits einen Bruchteil der erwarteten Abnahme (beispielsweise 10%) erreicht andererseits aber einen Anteil der erwarteten Abnahme (beispielsweise 90%) nicht übersteigt.

(a) Formulieren Sie diese beiden Bedingungen, wann eine Schrittweite akzeptiert wird, mathematisch.

(b) Schreiben Sie damit einen **Pseudocode**, welcher die Bestimmung einer Schrittweite über die Goldstein-Regel realisiert.

(c) Sei nun d eine Abstiegsrichtung der Zielfunktion f im Punkt x und f sei in Richtung d nach unten beschränkt, d.h.

$$\inf_{t\geq 0} f(x+td) > -\infty.$$

Beweisen Sie, dass Ihr Pseudocode aus Teilaufgabe **(b)** nach endlich vielen Schritten eine Schrittweite liefert, welche die Bedingungen aus Teilaufgabe **(a)** erfüllt. Ist f nach unten beschränkt und $(d_k)_{k\in\mathbb{N}}$ eine Folge von Abstiegsrichtungen, so ist die durch die Goldstein-Regel erzeugte Schrittweitenfolge zulässig.

Aufgabe 6.2 (Schriftliche Aufgabe)[4 Punkte]
Wir betrachten wieder die **Armijo-Schrittweitenregel mit Aufweitung** aus Aufgabe 5.4 und eine Folge von Abstiegsrichtungen $(d_k)_{k\in\mathbb{N}}$. Zeigen Sie, dass eine Schrittweitenfolge, welche die Bedingungen

$$\begin{aligned} f(x_k+s_kd_k) &\leq f(x_k)+\gamma s_k\nabla f(x_k)^\intercal d_k \\ f\left(x_k+\frac{s_k}{\beta}d_k\right) &> f(x_k)+\gamma\frac{s_k}{\beta}\nabla f(x_k)^\intercal d_k \end{aligned}$$

erfüllt, zulässig ist.

Aufgabe 7.1 (Votieraufgabe)
Sei $(x_k)_{k\in\mathbb{N}}\subset\mathbb{R}^n$ eine Folge, die superlinear gegen $\bar{x}$ konvergiert. Zeigen Sie:

$$\lim_{k\to\infty}\frac{\|x_{k+1}-x_k\|}{\|x_k-\bar{x}\|}=1.$$

(In der Nähe der Lösung ist also $\|x_{k+1}-x_k\|\approx\|x_k-\bar{x}\|$.)

Aufgabe 7.2 (Votieraufgabe)
Gegeben sind die C^2-Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ und die Variablentransformation

$$T\colon \mathbb{R}^n \to \mathbb{R}^n : y \mapsto Ay + b =: x$$

mit $A \in \mathbb{R}^{n\times n}$ invertierbar und $b \in \mathbb{R}^n$. Bezeichne $\hat{f}(y) = f(T(y))$ die auf y-Koordinaten transformierte Funktion f. Sei $x \in \mathbb{R}^n$ ein Punkt mit $\nabla f(x) \neq 0$ und invertierbarer Hesse-Matrix $\nabla^2 f(x)$. Weiter bezeichne x_g bzw. x_n das Ergebnis eines Gradienten- bzw. Newton-Schrittes zur Lösung des Problems $\min_{z\in\mathbb{R}^n} f(z)$ ausgehend von $x \in \mathbb{R}^n$:

$$x_g = x - \nabla f(x), \qquad x_n = x - \nabla^2 f(x)^{-1}\nabla f(x).$$

Entsprechend sei y_g bzw. y_n das Ergebnis eines Gradienten- bzw. Newton-Schrittes für $\hat{f}$ ausgehend von $y = T^{-1}(x)$:

$$y_g = y - \nabla \hat{f}(y), \qquad y_n = y - \nabla^2 \hat{f}(y)^{-1}\nabla \hat{f}(y).$$

(a) Drücken Sie $\nabla \hat{f}(y)$ und $\nabla^2 \hat{f}(y)$ mit Hilfe von $\nabla f(x)$, $\nabla^2 f(x)$, A und b aus.

(b) Zeigen Sie, dass das Newton-Verfahren invariant unter der Transformation T ist, d.h. $T(y_n) = x_n$.

(c) Für welche Klasse von Matrizen A ist das Gradientenverfahren invariant unter der Transformation T, d.h. $T(y_g) = x_g$? Für welche nicht?

Aufgabe 7.3 (Votieraufgabe)
Sei $f\colon \mathbb{R}^n \to \mathbb{R}$ zweimal stetig differenzierbar und $\hat{x} \in \mathbb{R}^n$ sei ein lokales Minimum von f mit positiv definiter Hesse-Matrix $\nabla^2 f(\hat{x})$. Wir betrachten das **inexakte Newton-Verfahren** aus Satz 2.55 der Vorlesung. Dies entspricht Algorithmus 6 der Vorlesung, mit dem Unterschied, dass die Richtungen d_k so gewählt werden, dass $\|\nabla f(x_k) + \nabla^2 f(x_k)d_k\| \leq \eta_k \|\nabla f(x_k)\|$ erfüllt ist, mit $(\eta_k)_{k\in\mathbb{N}}$ einer Nullfolge positiver Zahlen, wobei $\max_{k\in\mathbb{N}} \eta_k$ klein genug sein soll. Zeigen Sie, dass dann ein $\delta > 0$ existiert, sodass gilt:

(a) Für jeden Startwert $x_0 \in B_\delta(\hat{x})$ liegen alle vom **inexakten Newton-Verfahren** erzeugten Iterierten in $B_\delta(\hat{x})$. Insbesondere ist für $\delta > 0$ klein genug $\nabla^2 f(x_k)$ an allen Iterierten x_k invertierbar und der Algorithmus ist durchführbar.

(b) Für jeden Startwert $x_0 \in B_\delta(\hat{x})$ terminiert der Algorithmus entweder an einem stationären Punkt $\hat{x}$ oder er erzeugt eine Folge, die superlinear gegen $\hat{x}$ konvergiert.

Aufgabe 7.4 (Multiple Choice)
Sei $f\colon \mathbb{R}^n \to \mathbb{R}$ zweimal stetig differenzierbar und $\hat{x} \in \mathbb{R}^n$ sei ein lokales Minimum von f mit positiv definiter Hesse-Matrix $\nabla^2 f(\hat{x})$.

(a) Eine Folge im $\mathbb{R}^n$, die bezüglich einer Norm superlinear konvergiert, konvergiert auch bezüglich jeder anderen Norm superlinear. wahr ☐ falsch ☐

(b) Die Folge $(x_k) \subset \mathbb{R}^n$ konvergiere superlinear gegen $\bar{x}$. Dann gilt: Für $\gamma \in (0,1)$ konvergiert die Folge (x_k) linear mit Rate γ gegen $\bar{x}$. wahr ☐ falsch ☐

(c) Es existiert ein $\delta > 0$, sodass $\hat{x}$ der einzige stationäre Punkt von f auf $B_\delta(\hat{x})$ ist.
wahr ☐ falsch ☐

(d) Es existieren δ, $\mu > 0$, sodass für alle $x \in B_\delta(\hat{x})$ gilt: $\lambda_{min}(\nabla^2 f(x)) \geq \mu$ und $\nabla^2 f(x)$ ist invertierbar mit $\|\nabla^2 f(x)^{-1}\| \leq \frac{1}{\mu}$. wahr ☐ falsch ☐

Aufgabe 7.5 (Programmieraufgabe)[6 Punkte]

(a) Schreiben Sie eine MATLAB-Funktion

$$\texttt{function [z,fz] = Gradient\_Powell\_Wolfe(f,}\nabla\texttt{f,z0,}\gamma\texttt{,}\eta\texttt{)}$$

welche das allgemeine Abstiegsverfahren (Algorithmus 3) realisiert. Wählen Sie dazu die Suchrichtungen $d_k = -2^{-k}\nabla f(x_k)$ und bestimmen Sie die Schrittweiten mit der Powell-Wolfe Schrittweitenregel (Algorithmus 4).

(b) Schreiben Sie eine MATLAB-Funktion

```
function [z,fz] = Newton(f,grad_f,Hess_f,z0)
```

welche das Newton-Verfahren für Optimierungsprobleme (vgl. Algorithmus 6) aus der Vorlesung realisiert.

(c) Verwenden Sie die Funktion aus (a) um das globale Minimum von $f(x,y) = \sqrt{x^2+y^2+1}$. zu bestimmen. Verwenden Sie dazu den Startwert $z_0 = [0.5, 0.5]$, $\gamma = 0.01$ und $\eta = 0.9$.

(d) Verwenden Sie die Funktion aus (b) um das globale Minimum von $f(x,y) = \sqrt{x^2+y^2+1}$. zu bestimmen. Verwenden Sie dazu den Startwert $z_0 = [0.5, 0.5]$ und $z_0 = [1, 1]$.

(e) Verwenden Sie nun zur Bestimmung der Schrittweite die Armijo-Schrittweitenregel (verwenden Sie weiterhin die Suchrichtungen $d_k = -2^{-k}\nabla f(x_k)$). Bestimmen Sie wieder das globale Minimum von $f(x,y) = \sqrt{x^2+y^2+1}$. Verwenden Sie dazu $z_0 = [0.5, 0.5]$, $\beta = 0.5$ und $\gamma = 0.01$.

(f) Zeichnen Sie die Höhenlinien der Funktion $f(x,y) = \sqrt{x^2+y^2+1}$ (MATLAB-Befehl *contour*) sowie die Folge der Iterierten aus **(c)** bis **(e)**.

Hinweise zur Übungsblattbearbeitung:

- Zu **schriftlichen Aufgaben** soll eine Ausarbeitung/Lösung angefertigt werden, welche am 04.12.2018 in Ihrer Übungsgruppe abzugeben ist.

- Zu **Programmieraufgaben** soll eine kommentierte Ausarbeitung in MATLAB-Code bis zum 04.12.2018 um 10:00 Uhr an **jahn@math.uni-frankfurt.de** geschickt werden. Bitte beginnen Sie die Betreffzeile Ihrer E-Mail mit **"Opt67_1819:"**.

- Beachten Sie, dass der Bearbeitungszeitraum für dieses Blatt **zwei Wochen** beträgt. Der Übungstermin am 27.11.2018 entfällt.