Technische Universität Chemnitz
Prof. Dr. R. Herzog, J. Blechschmidt, A. Schäfer

Chemnitz, 31. Oktober 2014
Abgabe am 7. November 2014

Grundlagen der Optimierung

Übung 3

Aufgabe 9: Gradientenverfahren mit exakter Liniensuche am Beispiel

Wir betrachten das Problem

$$\text{Minimiere} \quad x^2 + (y-3)^2.$$

Bestimmen Sie die ersten zwei Iterierten des Gradientenverfahrens mit Startpunkt $x_0 = 0$, $y_0 = 0$ und exakter Liniensuche gemäß (4.2).

Aufgabe 10: Implementierung der Armijo-Schrittweitenstrategie

Implementiere die Armijo-Schrittweitenstrategie aus der Vorlesung in Matlab. Erstelle dazu eine Datei `ls_armijo.m` und verwende
`function t = ls_armijo(fhandle, x, d, s, sigma, beta)`
als erste Zeile.
Dabei bezeichnet `fhandle` das Handle auf eine Funktion (`help function_handle` und `doc function_handle`; Beispiel: `@rosenbrock`), die den Funktionswert und ggf. den Gradienten zurückgibt, `x` den Ausgangspunkt, `d` eine Abstiegsrichtung, `s` die Startschrittweite und `sigma` und `beta` die in der Vorlesung eingeführten Parameter des Algorithmus. Zurückgegeben werden soll eine Schrittweite `t`, welche die Armijo-Bedingung erfüllt.

Teste die Funktion an folgenden Beispielen und Eingabewerten.

(a) An der Rosenbrock-Funktion

$$f(x, y) = (1-x)^2 + 100(y-x^2)^2$$

mit `x`$= [1.7; 1.5]$, `d`$= [-1; 0]$, `s`$= 4$, `sigma`$=0.1$ und `beta`$= 0.5$.

(b) An der Rosenbrock-Funktion mit `x`$= [0; 0]$, `d`$= [1; 0]$, `s`$= 1$, `sigma`$=0.1$ und `beta`$= 0.5$.

(c) An

$$\varphi(x) := \omega(c_1)\sqrt{(1-x)^2 + c_2^2} + \omega(c_2)\sqrt{x^2 + c_1^2}$$

mit $\omega(c) = \sqrt{1+c^2} - c$, $c_1 = 0.01$, $c_2 = 0.001$, `x`$= 0$, `d`$= 1$, `s`$= 1$, `beta`$= 0.5$. Variiere `sigma` zwischen 0.01 und 0.5. Was stellt man fest?

Hinweis 1: Es kann die Datei `rosenbrock.m` von der Webseite der Lehrveranstaltung für Teil (a) und (b) verwendet werden. Für Teil (c) empfiehlt es sich, eine ähnliche Datei zu erstellen.
Hinweis 2: Der Quelltext ist angemessen zu kommentieren.

Aufgabe 11: Lineare Regression

Gegeben seien die Messwerte $(x_i, y_i), i = 1, \ldots, m$. Ein zugrunde liegendes physikalisches Modell sagt den linearen Zusammenhang

$$y(x) = ax + b$$

voraus. Anhand der Messwerte sollen die Parameter a und b bestimmt werden, sodass die Summe der Fehlerquadrate minimiert wird, d.h. es ist das unbeschränkte Optimierungsproblem

$$\min_{(a,b)\in\mathbb{R}^2} \quad f(a,b) = \sum_{i=1}^{m} (y(x_i) - y_i)^2 = \sum_{i=1}^{m} (ax_i + b - y_i)^2$$

zu lösen.

(a) Wie lauten die notwendigen Optimalitätsbedingungen erster Ordnung (die sogenannte Normalengleichung) (vgl. Satz 3.1 und Hausaufgabe 11 (a))?

(b) Zeige, dass $\nabla^2 f$ überall positiv semidefinit ist und sogar positiv definit, falls mindestens zwei x_i verschieden sind.

Aufgabe 12: Ein Häufungspunkt des Gradientenverfahrens kann kein Maximum sein

Zeige: Bricht das Gradientenverfahren (Algorithmus 4.4) nicht nach endlich vielen Schritten ab (mit $\nabla f(x_k) = 0$) und ist x^* ein Häufungspunkt der durch dieses Verfahren konstruierten Folge $\{x_k\}$, so ist x^* kein lokales Maximum der stetigen Funktion f. Gilt diese Aussage auch, wenn der Algorithmus nach endlich vielen Schritten in einem Punkt x^* abbricht?

Hausaufgabe 10: Exakte Liniensuche für eine quadratische Funktion

Die exakte Liniensuche (vgl. (4.2)) lautet:

„Bestimme $t_{\min}$ so, dass $f(x + t_{\min} d) = \min_{t\geq 0} f(x + t\,d)$ gilt.“

Berechne $t_{\min}$ für die quadratische Funktion

$$f(x) = \frac{1}{2} x^\top Q\, x + c^\top x + \gamma$$

mit einer symmetrischen, positiv definiten Matrix $Q \in \mathbb{R}^{n\times n}$, $c \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$ und einer Richtung d. Bestätige damit die Richtigkeit von Gleichung (4.12).

Hausaufgabe 11: Gradient, Hessematrix und Darstellung von ausgewählten quadratischen Funktionen und der Rosenbrock-Funktion

Bestimme Gradienten und Hessematrix folgender Funktionen.

(a) $f_1(x) = \frac{1}{2}x^\top Qx + b^\top x + d$ für $Q \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, $d \in \mathbb{R}$

(b) $f_2(x, y) = (1 - x)^2 + 100(y - x^2)^2$ (*Rosenbrock-Funktion*)

Wie vereinfacht sich die Darstellung, falls $Q = Q^\top$ vorausgesetzt wird?

Zeichne $f_1(x)$ für $b = \begin{pmatrix} 3 & 1 \end{pmatrix}^\top$, $d = 2$ und $Q =$

(i) $\begin{pmatrix} 5 & -3 \\ -3 & 4 \end{pmatrix}$, (ii) $\begin{pmatrix} 5 & -3 \\ -3 & 1 \end{pmatrix}$, (iii) $\begin{pmatrix} 5 & -3 \\ -3 & -4 \end{pmatrix}$, (iv) $\begin{pmatrix} -5 & -3 \\ -3 & -4 \end{pmatrix}$

auf $[-10, 10] \times [-10, 10]$ und $f_2(x, y)$ auf $[-2, 2] \times [-2, 2]$. Nutze dazu die Matlab-Funktionen `mesh` und `contour` (es ist günstig, geeignete Niveaus vorzugeben; siehe Beispieldatei `Plot_Himmelblau.m`).

Hausaufgabe 12: Implementierung des Gradientenverfahrens

Implementiere das Gradientenverfahren mit allgemeinen Skalarprodukt und Armijo-Schrittweitenstrategie in Matlab (Algorithmus 4.4 aus der Vorlesung). Verwende das Abbruchkriterium aus Bemerkung 4.7 (a) und (b).

Erstelle dazu eine Datei `steepest_descent.m` und verwende
`function X = steepest_descent(fhandle, M, x0, tol, s, sigma, beta)`
als erste Zeile. Dabei bezeichnet `fhandle` das Handle auf die Zielfunktion, `M` die Skalarprodukt induzierende Matrix, `x0` den Startpunkt, `tol` eine Struktur, welche die vier Parameter `ATOL_f`, `RTOL_f`, `ATOL_x` und `RTOL_x` für das Abbruchkriterium aus Bemerkung 4.7 (a) und (b) enthält, und `s`, `sigma` und `beta` die Parameter für die Armijo-Schrittweitensuche. Zurückgegeben werden soll eine Matrix $X = [x0, x1, x2, \ldots]$, welche den gesamten Iterationsverlauf enthält.

Teste das implementierte Verfahren an folgenden Funktionen und Eingabewerten und visualisiere den Iterationsverlauf im eindimensionalen Fall bzw. zweidimensionalen Fall in einem `plot` bzw. `contour-plot` mit Hilfe der zurückgegebenen Matrix X. Sofern nichts anderes angegeben, verwende dabei `M`$= I$, `sigma=1e-2`, `beta=0.5` sowie `1e-2` für die vier Toleranzen. Interpretiere dabei auch die Anzahl der benötigten Iterationen.

(a) An der Cosinus-Funktion mit `x0`$= 1.1656$

(b) An der Himmelblau-Funktion

$$f(x_1, x_2) = \left(x_1^2 + x_2 - 11\right)^2 + \left(x_1 + x_2^2 - 7\right)^2$$

mit den Startwerten `x0`$= [-0.27; -0.91]$, `x0`$= [-0.271; -0.91]$, `x0`$= [-0.25; -1.1]$ und `x0`$= [-0.25; -1]$

(c) An der Funktion $f_1(x)$ aus Hausaufgabe 11 (a) mit `x0`$= [4; 1]$

(d) An der Rosenbrock-Funktion aus Hausaufgabe 11 (b) mit `x0`$= [1; -0.5]$ und `1e-6` für die vier Toleranzen.

Experimentiere in Teil (d) mit der Matrix M. Finde eine Matrix M, sodass in Teil (d) weniger Iterationen benötigt werden.

Schicke die erzeugten Dateien an jan.blechschmidt@mathematik.tu-chemnitz.de (Betreff: HA Grundlagen der Optimierung Übung 3)!

Hinweis 1: Es können die Dateien von der Webseite der Lehrveranstaltung verwendet werden (insbesondere: `himmelblau.m`, `rosenbrock.m` und `ls_armijo.m`).
Hinweis 2: Der Quelltext ist angemessen zu kommentieren.

Hausaufgabe 13: Travelling Salesperson Problem (TSP)

Ein Pizzabote beliefert von der Pizzeria 1 die Kunden $2, \dots, n$. Die Fahrzeiten von $i \in \{1, \dots, n\}$ nach $j \in \{1, \dots, n\} \setminus \{i\}$ sind durch $c_{ij} \geq 0$ gegeben. Es gilt $c_{ij} = c_{ji}, \forall i \neq j$. Gesucht ist eine Fahrtroute zu allen Kunden und zurück zur Pizzeria, so dass die Gesamtfahrzeit so gering wie möglich ist. Formuliere eine entsprechende Optimierungsaufgabe.