Institut für Analysis und Numerik 12.06.2020
Universität Magdeburg
Thomas Richter thomas.richter@ovgu.de
Gozel Judakova gozel.judakova@ovgu.de

Übung Nr. 9 zur Vorlesung Algorithmische Mathematik II
Sommersemester 2020

Abgabe bis zum 19. Juni 2020

Es geht nun um Algorithmen zur Suche von Minima einer Funktion. Dabei werden wir uns auf die Suche nach lokalen Minima beschränken und immer die Optimalitätskriterien berücksichtigen. D.h., im Allgemeinen setzen wir z.B. voraus, dass die Funktion f differenzierbar ist.

Das Gradientenverfahren Man lese Abschnitt *8.2 Abstiegsverfahren und das Gradientenverfahren* bis einschließlich Abschnitt *8.2.1 Das Gradientenverfahren*. Man gebe kurze Antworten:

1. Aus welchen 2 Schritten besteht jedes Abstiegsverfahren?
2. Was zeichnet eine *Abstiegsrichtung* aus? Was ist die Richtung des steilsten Abstiegs?
3. Wir betrachten die Funktion $f(x, y) = x^2 - 3yx^2$ und hier den Punkt $(1, 1)$. Ist die Richtung $(1, 1)$ eine Abstiegsrichtung? In welche Richtung ist im Punkt $(1, 1)$ der steilste Abstieg?

Schrittweitenwahl Die Richtung des steilsten Abstiegs $\mathbf{d}$ ist sehr einfach zu bestimmen. Wir wählen einfach den negativen Gradienten. Weit schwerer ist es, eine gute Schrittweite für die nächste Iteration zu bestimmen. Es gilt nun, die optimale Schrittweite zu finden so dass

$$\min_s f(x + s\mathbf{d}).$$

Hierzu muss selbst wieder ein Minimierungsproblem gelöst werden, zwar ein eindimensionales, aber auch dies ist nicht im allgemeinen Fall zu lösen. Wir untersuchen daher eine algorithmische Umsetzung der Wahl von s. Man lese Abschnitt *8.2.2 Schrittweitenwahl* und gebe kurze Antworten

1. Warum beginnt man bei der Line-Search Methode mit $s = 1$ und reduziert die Schrittweite dann? Genauer gesagt: warum glaubt man, so irgendwann eine Reduktion zu erreichen?
2. Was ist der Vorteil der Armijo-Regel gegenüber des einfachen Line-Search?

Konvergenz des Gradientenverfahrens Es geht nun um den zentralen Beweis zur Konvergenz des Gradientenverfahrens. Dieser Beweis ist (verhältnismäßig) einfach, kann aber auch nur garantieren, dass wir konvergieren. Der Nachweis einer Konvergenzgeschwindigkeit ist viel aufwändiger. Man lese Abschnitt *8.2.3 Konvergenz des Gradientenverfahrens*, d.h. man gehe insbesondere den Beweis genau durch und gebe kurze Antworten

1. Die Aussage von Satz 8.18 ist recht kompliziert. Können wir hieraus in jedem Fall Konvergenz gegen ein Minimum oder Abbruch durch Finden eines stationären Punktes garantieren?

2. Versuchen Sie ein Szenario zu beschreiben, in dem *das Gradientenverfahren nicht mit einem stationären Punkt abbricht, die Funktion nach unten beschränkt ist, das Gradientenverfahren aber nicht gegen einen stationären Punkt konvergiert, diesen nur aus Häufungspunkt hat.*

Aufgabe 9.1:

Wir betrachten das Gradientenverfahren mit fester Schrittweite $s \in \mathbb{R}$ zur Minimierung von

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x) = \|x\|_2^2 = x_1^2 + x_2^2.$$

a) Wie lautet $\nabla f(x)$

b) Es sei $x^0 = (\frac{1}{2}, 0)^T$. Man berechne den negativen Gradienten $-\nabla f(x^0)$ von f in x^0. Wo liegt die nächste Iteration x^1 bei Wahl der Schrittweite $s = 1$?

c) Bestimmen Sie bei Wahl von $s = 1$ die Iterierten x^1, x^2, x^3.

Programmieraufgabe 9.2:

Wir untersuchen das Gradientenverfahren zur Suche des Minimums der Rosenbrock-Funktion

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

a) Man implementiere das Gradientenverfahren mit vorgegebener Schrittweite, d.h., man implementiere eine Funktion `def gradientenverfahren(x,step,tol,maxiter)`, die ausgehend vom Startvektor `x` das Gradientenverfahren durchführt und jeweils die feste Schrittweite `step` nutzt, d.h.

$$x_{k+1} = x_k - \texttt{step}\nabla f(x_k).$$

Die Iteration soll abbrechen, wenn die vorgegebene Toleranz `tol` erreicht wird, d.h.

$$\|\nabla f(x_k)\| < \texttt{tol},$$

oder wenn die maximale Anzahl an Schritten erreicht wird. Man versuche ausgehend vom Startvektor `x=np.array([1,−1])` ein Residuum von `tol=1.e−6` zu erreichen. Man experimentiere mit der geeigneten Wahl der Schrittweite `step` und wähle hinreichend viele Schritte `maxiter`.
Man dokumentiere das gefundene Minimum, die Anzahl der notwendigen Schritte und das hier erreichte Residuum.

b) Man wiederhole Aufgabe **a**) aber mit Line-Search zur Schrittweitenwahl. Hierfür wähle man $\beta = \frac{1}{2}$.

c) Man wiederhole Aufgabe **a**) aber mit der Armijo-Regel zur Schrittweitenwahl. Hierfür teste man z.B. $\beta = \frac{1}{2}$ und $\gamma = 0.01$, versuche aber noch bessere Parameter zu finden.

d) Man stelle den Konvergenzverlauf für alle drei Varianten (jeweils mit Parametern so dass Konvergenz erreicht wird) graphisch dar. Hierfür wähle man eine doppelt-Logarithmische Darstellung.