Einführung in die Numerische Mathematik

Sommersemester 2019

Prof. Dr. Ira Neitzel
Fabian Hoppe

UNIVERSITÄT BONN

Übungsblatt 1

Abgabe am 11. April vor der Vorlesung

Aufgabe 1. (1 + 2 + 1 = 4 Punkte)

Sei $A \in \mathbb{R}^{m\times n}$, $y \in \mathbb{R}^m$ und $\gamma \geq 0$. Wir betrachten

$$f\colon \mathbb{R}^n \longrightarrow \mathbb{R}, \qquad x \longmapsto \frac{1}{2}\|Ax - y\|_2^2 + \frac{\gamma}{2}\|x\|_2^2.$$

a) Berechnen Sie den Gradienten ∇f und die Hessematrix $\nabla^2 f$ von f.

b) Bestimmen Sie den/die kritischen Punkte von f und deren Art in Abhängigkeit von A bzw. γ. Hat f lokale/globale Minima? Sind diese strikt bzw. eindeutig?

c) Sei $\gamma > 0$. Zeigen Sie, dass die Lösung des Minimierungsproblems

$$\min_{x\in\mathbb{R}^n} f(x)$$

differenzierbar von den Daten $(A, y, \gamma) \in \mathbb{R}^{m\times n} \times \mathbb{R}^m \times (0, \infty)$ abhängt.

Aufgabe 2. (1 + 1 + 2 = 4 Punkte)

Sei $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ zweimal stetig differenzierbar. Für einen Punkt $x^* \in \mathbb{R}^n$ sagen wir, dass x^* ein lokales Minimum von f entlang jeder Geraden durch x^* ist, wenn die Funktion

$$g_d\colon \mathbb{R} \longrightarrow \mathbb{R}, \qquad \alpha \longmapsto g_d(\alpha) := f(x^* + \alpha d)$$

für jede Richtung $d \in \mathbb{R}^n \setminus \{0\}$ ein lokales Minimum in $\alpha = 0$ hat.

a) Zeigen Sie: Ist x^* lokales Minimum von f entlang jeder Geraden durch x^*, so gilt $\nabla f(x^*) = 0$.

b) Zeigen Sie: Ist x^* ein lokales Minimum von f, so ist x^* auch lokales Minimum von f entlang jeder Geraden durch x^*.

c) Gilt auch die Umkehrung der Aussage von Teilaufgabe b)?

Hinweis: $f(x_1, x_2) := (x_2 - px_1^2)(x_2 - qx_1^2)$ mit $0 < p < q$ und $x^* = (0, 0)$.

Aufgabe 3. (4 Punkte)

Gegeben sei die Funktion

$$f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}, \qquad (x_1, x_2) \mapsto \frac{1}{2}\left(x_1^2 + x_2^2\right).$$

Für einen Startvektor $x_0 \in \mathbb{R}^2 \setminus \{0\}$ betrachten wir das Abstiegsverfahren mit den Suchrichtungen $s_k := g_k^\perp - \frac{1}{2^{k+3}} g_k$ sowie zulässigen Schrittweiten σ_k, wobei $g_k = \nabla f(x_k)$ gilt und $g_k^\perp \in \mathbb{R}^2$ so gewählt ist, dass

$$\langle g_k^\perp, g_k \rangle = 0 \qquad \text{und} \qquad \|s_k\|_2 = \|g_k\|_2$$

erfüllt ist.

Zeigen Sie, dass dieses Abstiegsverfahren nicht gegen den Minimalpunkt $x^* = (0, 0)$ von f konvergiert und x^* auch nicht Häufungspunkt der Folge der Iterierten ist.

Aufgabe 4. (je 1 Punkt, also insg. 8 Punkte)

Richtig oder Falsch? – Finden Sie für richtige Aussagen einen (kurzen) Beweis und widerlegen Sie die falschen Aussagen mit einem Gegenbeispiel.

a) Für eine konvexe Funktion $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ sind alle Niveaumengen $N(f, c) := \{x \in \mathbb{R}^n : f(x) \le c\}$, $c \in \mathbb{R}$, konvex.

b) Sind für eine Funktion $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ alle Niveaumengen $N(f, c)$, $c \in \mathbb{R}$, konvex, so ist f eine konvexe Funktion.

c) Sei $U \subset \mathbb{R}^n$ eine nichtleere Menge und $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ stetig differenzierbar. Ist $x^* \in U$ globales Minimum von f auf U, dann gilt $\nabla f(x^*) = 0$.

d) Sei $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ zweimal stetig differenzierbar mit $\nabla f(x^*) = 0$ und $\nabla^2 f(x^*) > 0$ (positiv definit) für ein $x^* \in \mathbb{R}^n$. Dann ist x^* striktes lokales Minimum von f mit

$$f(x) \ge f(x^*) + \alpha \|x - x^*\|_2^2 \qquad \text{für alle } x \in \mathbb{R}^n \quad \text{mit} \quad \|x - x^*\|_2 \le \beta \tag{1}$$

mit gewissen $\alpha, \beta > 0$.

e) Sei $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ zweimal stetig differenzierbar mit $\nabla f(x^*) = 0$ und $\nabla^2 f(x^*) > 0$ (positiv definit) für ein $x^* \in \mathbb{R}^n$. Dann kann x^* nicht Häufungspunkt einer Folge von lokalen Minima von f sein.

f) Sei $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ zweimal stetig differenzierbar und $x^* \in \mathbb{R}^n$ ein striktes globales Minimum von f. Dann gibt es $\alpha, \beta > 0$, sodass (1) gilt.

g) Sei $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ differenzierbar. Wenn es $x^* \in \mathbb{R}^n$ gibt und $\alpha, \beta > 0$ derart, dass (1) gilt, so kann x^* nicht Häufungspunkt einer Folge lokaler Minima von f sein.

h) Eine nach unten beschränkte konvexe Funktion $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ hat entweder keinen, einen eindeutig bestimmten globalen Minimierer oder unendlich viele.

Programmieraufgabe 1. (15 Punkte)

Das Mutations-Selektions-Verfahren gehört zu den sogenannten *ableitungsfreien Verfahren*, d.h. es werden nur Funktionsauswertungen, nicht aber Ableitungen der Funktion, benötigt. Der Verzicht auf die Verwendung von Ableitungen wird in der Regel mit vergleichsweise sehr langsamer Konvergenz und dem Fehlen theoretischer Konvergenzaussagen bezahlt.

Mutations-Selektions-Verfahren

Input: Funktion $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, Startwert $x_0 \in \mathbb{R}^n$, Schrittweite $\sigma > 0$, Anzahl der Iterationen $N \in \mathbb{N}$.
Output: Annäherung $x \in \mathbb{R}^n$ für ein lokales Minimum von f

$x \longleftarrow x_0$
for $k = 1, ..., N$ **do**
  Wähle $d \in [-1, 1]^n$ zufällig, z.B. begzüglich der gleichmäßigen Wahrscheinlichkeitsverteilung auf $[-1, 1]^n$;
  $v \longleftarrow x + \sigma d$;
  if $f(v) < f(x)$ **then**
    $x \longleftarrow v$;
  end
end

(Quelle: W. Alt, *Nichtlineare Optimierung*, Vieweg, 2002)

Der oben dargestellte Algorithmus ist "biologisch" motiviert. Man führt "zufällige Mutationen" des aktuellen Iterationspunktes durch und "selektiert" diejenigen "Mutationen", die brauchbar sind, also den Funktionswert verkleinern. Simulated Annealing (siehe unten) stellt eine Verallgemeinerung des Mutations-Selektions-Verfahrens dar. Anstatt "Mutationen" mit höherem Funktionswert immer auszusortieren werden nun auch diese Mutationen zumindest mit einer gewissen Wahrscheinlichkeit als neue Iterierte akzeptiert. Die Formel für die Akzeptanzwahrscheinlichkeit kann physikalisch motiviert werden.

Simulated Annealing

Input: Funktion $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, Monoton fallende "Temperatur-Funktion" $t : \mathbb{N} \longrightarrow (0, \infty)$ mit $\lim_{n\to\infty} t(n) = 0$, Startwert $x_0 \in \mathbb{R}^n$, Anzahl der Iterationen $N \in \mathbb{N}$, Radius $\sigma > 0$.
Output: Annäherung $x_{\text{best}} \in \mathbb{R}^n$ für einen Minimierer von f

$x \longleftarrow x_0$;
$x_{\text{best}} \longleftarrow x_0$;
for n = *1,...,N* **do**
  Wähle $d \in [-1, 1]^n$ zufällig, z.B. bzgl. der gleichmäßigen Wahrscheinlichkeitsverteilung;
  $v \longleftarrow x + \sigma d$;
  if $f(v) \leq f(x)$ **then**
    $x \longleftarrow v$;
    $x_{\text{best}} \longleftarrow x$;
  else
    Setze $x \longleftarrow v$ mit Wahrscheinlichkeit $\exp\left(-\frac{f(v)-f(x)}{t(n)}\right)$;
  end
end

Abgabe der Programmieraufgabe in der Woche 15. - 18. April