

Übung zur Vorlesung  
Computergestützte Statistik  
Wintersemester 2018/2019  
Übungsblatt Nr. 8

Abgabe ist Montag der 03.12.2018 an CS-abgabe@statistik.tu-dortmund.de oder Briefkasten 138

---

**Aufgabe 1**

**(4 Punkte)**

Betrachten Sie die Quadratic-Interpolation-Search aus dem Skript. **Achtung:** In der ursprünglichen Variante des Skripts ist leider eine falsche Formel angegeben. Korrekt muss es lauten:

$$\theta_{\text{new}} = \frac{1 f(\theta_{\text{upper}})(\theta_{\text{lower}}^2 - \theta_{\text{best}}^2) + f(\theta_{\text{best}})(\theta_{\text{upper}}^2 - \theta_{\text{lower}}^2) + f(\theta_{\text{lower}})(\theta_{\text{best}}^2 - \theta_{\text{upper}}^2)}{2 f(\theta_{\text{upper}})(\theta_{\text{lower}} - \theta_{\text{best}}) + f(\theta_{\text{best}})(\theta_{\text{upper}} - \theta_{\text{lower}}) + f(\theta_{\text{lower}})(\theta_{\text{best}} - \theta_{\text{upper}})}$$

- a) (3 Punkte) Implementieren Sie den Algorithmus.
- b) (1 Punkt) Testen Sie Ihre Implementierung. Denken Sie sich dazu einige (mindestens 4) eindimensionale, unimodale Optimierungsprobleme aus. Bestimmen Sie analytisch die wahren Optimum dieser Probleme und überprüfen Sie, ob Ihre Implementierung diese findet.

**Aufgabe 2**

**(4 Punkte)**

In der Vorlesung haben Sie gelernt, dass zur Bestimmung des Medians eines Vektors  $\mathbf{x}$  dieser Vektor nicht unbedingt sortiert werden muss. Stattdessen kann auch das Minimierungsproblem  $\sum_{i=1}^n |x_i - \beta|$  verwendet werden. Die spannende Frage ist nun: Welche der beiden Varianten ist besser, und was macht R? Der Einfachheit halber wollen wir hier nur Vektoren ungerader Länge betrachten, damit der Median auch eindeutig definiert ist.

- a) (1 Punkt) Betrachten Sie den Beobachtungsvektor  $\mathbf{x} = [15, 51, 33, 20, 66, 35, 72, 3, 34]$ . Bestimmen Sie den Median! Visualisieren Sie dazu das Optimierungsproblem für  $\mathbf{x}$  und bestimmen Sie das Minimum durch *hinschauen*.
- b) (1 Punkt) Führen Sie per Hand die *Golden-Section Suche* zur Bestimmung des Medians durch. Wie viele Funktionsauswertungen haben Sie benötigt? **Hinweis:** Das ursprüngliche Abbruchkriterium der *Golden-Section Suche* ist bei dieser Anwendung nur bedingt geeignet. Brechen Sie stattdessen ab, sobald sich im Suchintervall nur noch eine Beobachtung aus  $\mathbf{x}$  befindet.
- c) (1 Punkt) Bestimmen Sie die Laufzeit dieser Art der Medianbestimmung. Hierzu müssen Sie abschätzen, nach wie vielen Iterationen die GSS terminiert. Verwenden Sie dabei eine geeignete Verteilungsannahme.
- d) (0.5 Punkte) Sie kennen nun 2 Alternativen zur Medianbestimmung - Optimieren und Sortieren. Welche würden Sie bevorzugen? Begründen Sie ausführlich!
- e) (0.5 Punkt) Informieren Sie sich wie in R der Median bestimmt wird. Inwiefern ist dies intelligenter als unsere Ansätze?

### Aufgabe 3

(4 Punkte)

Implementierung und Test des Koordinatenabstiegs:

- a) (3 Punkte) Implementieren Sie das im Skript vorgestellte Koordinatenabstiegsverfahren. Verwenden Sie dazu die Abbruchkriterien 1 (mit  $p = n$ ) und 4 (Folie 399). Denken Sie an eine ordentliche Dokumentation.
- b) (1 Punkt) Testen Sie Ihre Implementierung, indem Sie die negative Dichte einer bivariaten Normalverteilung mit Korrelation  $\rho > 0$  minimieren.