

Prof. Dr. Christine Müller
M.Sc. Dennis Malcherczyk

Robuste Statistik

Blatt 8

Aufgabe 8.1: (5 Punkte)

Beweisen Sie Satz 4.2.8 für den Fall, dass N ungerade ist (damit keine mengenwertige Zwischenergebnisse entstehen). Zeigen Sie also, dass der Bruchpunkt des MADs $\frac{1}{N} \lfloor \frac{N-1}{2} \rfloor$ beträgt.

Hinweis: Bei der Bestimmung des Bruchpunktes von Streuungsschätzern gehen wir von Beobachtungen aus, die sich in allen Einträgen paarweise unterscheiden.

Aufgabe 8.2: (5 Punkte)

Folgender Datensatz im \mathbb{R}^2 sei gegeben. Ziel dieser Aufgabe ist die Bestimmung der Lokation $\boldsymbol{\mu} \in \mathbb{R}^2$ des Datensatzes mit verschiedenen Methoden im Multivariaten.

$$\mathbf{y}_1 = \begin{pmatrix} 3.43 \\ 0.55 \end{pmatrix}, \mathbf{y}_2 = \begin{pmatrix} 3.77 \\ 2.22 \end{pmatrix}, \mathbf{y}_3 = \begin{pmatrix} 5.56 \\ 1.36 \end{pmatrix}, \mathbf{y}_4 = \begin{pmatrix} 4.07 \\ 1.4 \end{pmatrix},$$

$$\mathbf{y}_5 = \begin{pmatrix} 5.07 \\ 12.95 \end{pmatrix}, \mathbf{y}_6 = \begin{pmatrix} 5.72 \\ 0.44 \end{pmatrix}, \mathbf{y}_7 = \begin{pmatrix} 4.46 \\ 2.79 \end{pmatrix}, \mathbf{y}_8 = \begin{pmatrix} 12.73 \\ 4.5 \end{pmatrix}, \mathbf{y}_9 = \begin{pmatrix} 3.31 \\ 0.97 \end{pmatrix}.$$

- Stellen Sie die Datenpunkte geeignet dar. Welche Daten fallen auf?
- Schätzen Sie $\boldsymbol{\mu}$ durch den komponentenweisen Mittelwert, komponentenweisen Median und l_1 -Median.
- Bestimmen Sie die euklidischen Abstände der drei geschätzten Lokationen in b) zu allen Datenpunkten $\mathbf{y}_1, \dots, \mathbf{y}_9$ und stellen Sie die Abstände grafisch dar. Welche Abstände fallen auf?
- Sei

$$Q(\boldsymbol{\mu}, \hat{\boldsymbol{\mu}}(y)) := \sum_{i=1}^2 (\mu_i - \hat{\mu}_i(y))^2$$

eine Qualitätsfunktion eines Schätzers $\hat{\boldsymbol{\mu}}(y) \in \mathbb{R}^2$ für $\mathbf{y} \in \mathbb{R}^{2 \times N}$, $\boldsymbol{\mu} \in \mathbb{R}^2$. Betrachten Sie $\tilde{\mathbf{y}}$ festgelegt durch \mathbf{y} und einer hinzugefügten Beobachtung $c \cdot \mathbf{1}_2 = \begin{pmatrix} c & c \end{pmatrix}^T \in \mathbb{R}^2$ und plotten Sie $Q(\boldsymbol{\mu}, \hat{\boldsymbol{\mu}}(\tilde{\mathbf{y}}))$ als Funktion in c mit $c \in [-10, 10]$ und Schrittweite 0.01 für die drei Schätzer aus b) in eine Grafik. Nehmen Sie dabei an, dass das wahre $\boldsymbol{\mu}$ den Wert $\begin{pmatrix} 4 & 1 \end{pmatrix}^T$ hat. Interpretieren Sie die Ergebnisse.

Abgabe bis spätestens 27.05.2019, 10 Uhr.