

## Robuste Statistik

### Blatt 10

#### Aufgabe 10.1: (5 Punkte)

Erzeugen Sie den folgenden Datensatz `data10.1`:

```
library(MASS)
set.seed(30619)
data10.1 <- round(mvrnorm(n = 40, c(sqrt(5),sqrt(2)),
  matrix(c(0.2,0.3, 0.3, 0.6), nrow = 2)), digits = 2)
```

- Schätzen Sie die Kovarianzmatrix des Datensatzes `data10.1` mittels empirischer Kovarianz, der MVE-Schätzung, der MCD-Schätzung und der räumlichen Vorzeichen-Kovarianz-Schätzung.
- Plotten Sie eine Verfälschungsfunktion bezüglich der wahren Kovarianzmatrix, indem Sie die vier Schätzer aus a) ausgewertet auf dem Datensatz `data10.1` mit einer hinzugefügten Beobachtung darstellen. Dabei müssen Sie insbesondere ein geeignetes Distanzmaß von der Schätzung zur tatsächlichen Matrix definieren, damit Ihre Verfälschungsfunktion von den reellen Zahlen in die reellen Zahlen abbildet. Die hinzugefügte Beobachtung soll folgende Gestalt aufweisen: (i)  $\begin{pmatrix} c & c \end{pmatrix}^\top$ , (ii)  $\begin{pmatrix} 1.122 & c \end{pmatrix}^\top$ , (iii)  $\begin{pmatrix} c & 5.244 \end{pmatrix}^\top$ , jeweils für geeignete Werte von  $c$ . Beschreiben Sie Ihre Ergebnisse.

#### Aufgabe 10.2: (5 Punkte)

Erzeugen Sie folgende Übungsdaten `data1` bis `data4` für Klassifikationsprobleme:

```
library(mlbench)
set.seed(1)
dat1 <- mlbench.2dnormals(1000, sd = 1.2)
data1 <- data.frame(x = dat1$x, class = dat1$classes)
dat2 <- mlbench.2dnormals(1000, sd = 5)
data2 <- data.frame(x = dat2$x, class = dat2$classes)
dat3 <- mlbench.spirals(1000, 1, 0.1)
data3 <- data.frame(x = dat3$x, class = dat3$classes)
dat4 <- mlbench.circle(1000, 2)
data4 <- data.frame(x = dat4$x, class = dat4$classes)
```

Stellen Sie für die Daten `data1` bis `data4` die DD-Plots (vgl. Kapitel 5.3 in der Vorlesung) für die Halbraum- und Simplex-Tiefe in R dar. Beschreiben und erklären Sie die Resultate.

**Abgabe bis spätestens 11.06.2019 in den Briefkasten 134 oder in der Übung**