

4 Multivariate Normalverteilungen

4.1 Mathematische Grundlagen

Ein n -dimensionaler Zufallsvektor $\xi \in \mathbb{R}^n$ folgt genau dann einer *multivariaten Normalverteilung* mit *Erwartungswertparameter* $\mu \in \mathbb{R}^n$ und *Kovarianzmatrixparameter* $\Sigma \in \mathbb{R}^{n \times n}$ p.d., d.h. in mathematischer Notation

$$\xi \sim N(\mu, \Sigma) , \quad (1)$$

wenn seine Wahrscheinlichkeitsdichtefunktion (WDF) wie folgt gegeben ist:

$$p(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp \left(-\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right) . \quad (2)$$

Wenn ein Zufallsvektor $\xi \in \mathbb{R}^m$ einer multivariaten Normalverteilung folgt und Realisierungen $x_1, \dots, x_n \in \mathbb{R}^m$ dieses Zufallsvektors vorliegen, können die Verteilungsparameter mithilfe folgender Gleichungen geschätzt werden:

$$\begin{aligned} \hat{\mu} &= \frac{1}{n} \sum_{i=1}^n x_i \\ \hat{\Sigma} &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu}) (x_i - \hat{\mu})^T . \end{aligned} \quad (3)$$

Eine wichtige Eigenschaft der multivariaten Normalverteilung gibt das Theorem zur *linear-affinen Transformation* wieder: Ist $\xi \in \mathbb{R}^n$ multivariat normalverteilt, so ist auch die linear-affine Transformation $v = A\xi + b \in \mathbb{R}^m$ multivariat normalverteilt, wobei sich die Parameter der resultierenden Verteilung aus den Parametern der ursprünglichen Verteilung sowie den Transformationsparametern $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$ ergeben:

$$\xi \sim N(\mu, \Sigma) \quad \Rightarrow \quad v = A\xi + b \sim N(A\mu + b, A\Sigma A^T) , \quad (4)$$

Wir werden in dieser Übung sehen, dass diese Eigenschaft nicht nur für die Verteilungsparameter μ und Σ selbst, sondern auch für deren Schätzer $\hat{\mu}$ und $\hat{\Sigma}$ gilt. Mit anderen Worten: Wird ein Zufallsvektor als multivariat normalverteilt angenommen und werden seine Realisierungen linear-affin transformiert, so ergeben sich die Schätzer für die transformierten Realisierungen durch linear-affine Transformation der ursprünglichen Schätzer:

$$y_i = Ax_i + b, \quad i = 1, \dots, n \quad \Rightarrow \quad \hat{\mu}^{(y)} = A\hat{\mu}^{(x)} + b \quad \text{und} \quad \hat{\Sigma}^{(y)} = A\hat{\Sigma}^{(x)}A^T . \quad (5)$$

4.2 Analyse in R

Die Datei `FADE_SAME.csv` enthält den in der ersten Seminarsitzung vorgestellten Datensatz. Erklären Sie die Funktion des folgenden R-Codes:

```
# Daten einlesen
fname = 'FADE_SAME.csv'           # Dateiname
D      = read.csv(fname)          # Dataframe

# Variablen extrahieren
vars   = c('novelty.FADE', 'novelty.SAME',      # Variablen
            'memory.FADE', 'memory.SAME')
m      = length(vars)                 # Anzahl Variablen
n      = nrow(D)                     # Anzahl Datenpunkte
Y      = matrix(rep(0, m*n), nrow = m)       # Datenmatrix
for (j in 1:m) {
  Y[j,] = D[,vars[j]]                # j-te Variable
}
print(Y[,1:8])                      # erste 8 Datenpunkte
```

4.3 Erste Programmieraufgabe

Berechnen Sie nun die Schätzer für den Erwartungswert- und den Kovarianzmatrixparameter der für die Spalten der Datenmatrix angenommenen multivariaten Normalverteilung. Gehen Sie dazu wie folgt vor:

- Implementieren Sie die obige Formel für $\hat{\mu}$, indem Sie mithilfe der Funktion `rowSums` die Zeilensummen der Matrix Y ermitteln und diese mit $1/n$ multiplizieren. Speichern Sie das Ergebnis als die Variable `mu_est`.
- Implementieren Sie die obige Formel für $\hat{\Sigma}$, indem Sie eine Nullmatrix der Größe $m \times m$ initialisieren, mithilfe einer `for`-Schleife über alle Datenpunkte die Terme $(x_i - \hat{\mu})(x_i - \hat{\mu})^T$ aufsummieren und schließlich mit $1/(n-1)$ multiplizieren. Speichern Sie das Ergebnis als die Variable `Sigma_est`.
- Geben Sie die Resultate ihrer Analyse aus. Sie sollten folgende Ergebnisse erhalten:

```
      [,1]
[1,] -1.8163838
[2,] -0.1993077
[3,] -1.1176979
[4,] -0.5112862
```

```
      [,1]      [,2]      [,3]      [,4]
[1,]  0.41849409 -0.28080692  0.01752792  0.06186873
[2,] -0.28080692  0.43955708 -0.05147251  0.08905219
[3,]  0.01752792 -0.05147251  0.33872503 -0.41351203
[4,]  0.06186873  0.08905219 -0.41351203  0.79214018
```

4.4 Abbildung in R

Im nächsten Schritt soll ein Teil der Ergebnisse visualisiert werden. Erklären Sie den folgenden R-Code und die daraus resultierende Abbildung. Erläutern Sie insbesondere, welchen Einfluss die Indizierung von `mu_est` und `Sigma_est` mit `3:4` auf die dargestellten Variablen hat.

```
# Abbildungsparameter
library(latex2exp)
library(mvtnorm)
par(
  family      = "sans",
  mfcol       = c(1,1),
  pty        = "s",
  bty        = "l",
  lwd        = 1,
  las        = 1,
  mgp        = c(2,1,0),
  xaxs       = "i",
  yaxs       = "i",
  font.main   = 1,
  cex        = 0.9,
  cex.main    = 1.2)

# Wahrscheinlichkeitsdichtefunktion
x_min      = -3
x_max      = +3
x_res      = 1e3
x_1        = seq(x_min, x_max, length.out = x_res)
x_2        = seq(x_min, x_max, length.out = x_res)
X          = expand.grid(x_1, x_2)
p          = matrix(dmvnorm(as.matrix(X), mu_est[3:4], Sigma_est[3:4,3:4]), nrow = x_res)

# Konturplot
contour(x_1, x_2, p,
  xlim     = c(x_min,x_max),
  ylim     = c(x_min,x_max),
  xlab     = TeX("Memory-FADE-Score"),
  ylab     = TeX("Memory-SAME-Score"),
  nlevels  = 5)

# Datenpunkte
points(Y[3,], Y[4,],
  pch     = 21,
  col     = "white",
  bg      = "gray60",
  cex     = 1)

# Speicherung
dev.copy2pdf(
  file    = "Abbildungen/Multivariate_Normalverteilungen_1.pdf",
```

```
width  = 6,  
height = 6)
```

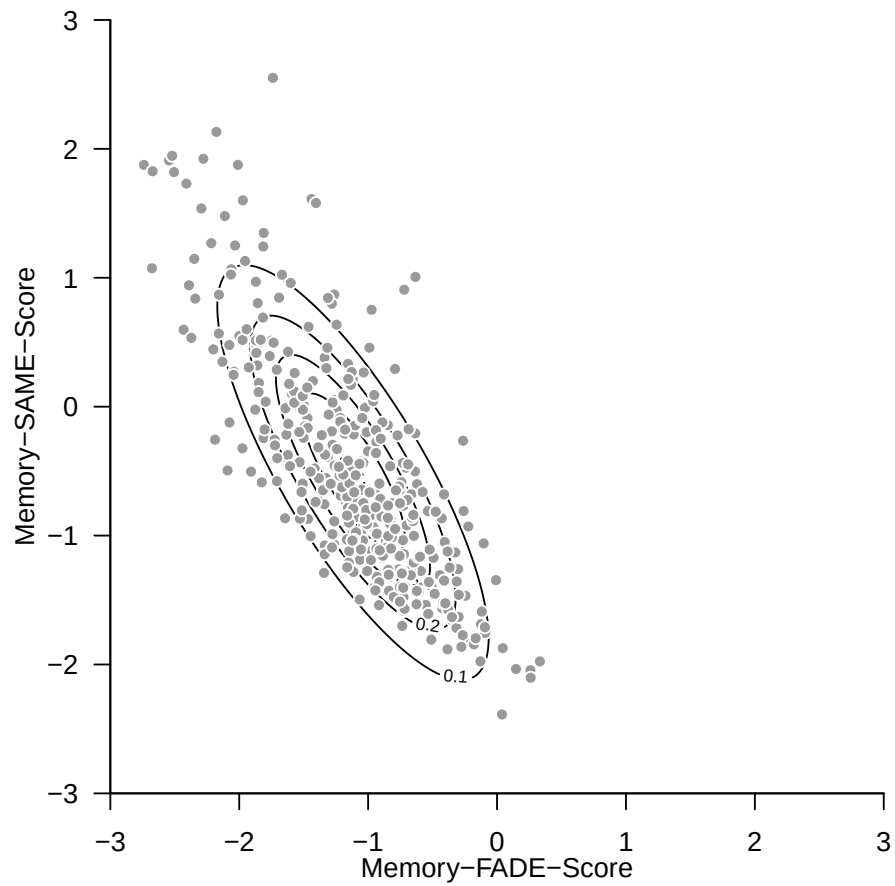


Abbildung 1. Beobachtete Memory-FADE-Scores und Memory-SAME-Scores sowie geschätzte Wahrscheinlichkeitsdichtefunktion der ihnen zugrundeliegenden bivariaten Normalverteilung.

4.5 Zweite Programmieraufgabe

Anhand der geschätzten Parameter der multivariaten Normalverteilung soll nun das in der Vorlesung besprochene Theorem zur linear-affinen Transformation nachvollzogen werden. Gehen Sie dazu wie folgt vor:

- Vergewenwärtigen Sie sich zunächst, wie in Abschnitt 2.5 des Arbeitsblatts (2) *Matrizen* mithilfe einer Matrix-Multiplikation die Mittelwerte der FADE-Scores sowie der SAME-Scores errechnet wurden. In diesem Abschnitt werden wir etwas ähnliches tun, indem mithilfe einer Matrix-Multiplikation die Mittelwerte der Novelty-Scores sowie der Memory-Scores berechnen werden.
- Definieren Sie die Matrix $A = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix} \in \mathbb{R}^{2 \times 4}$ und den Vektor $b = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \mathbb{R}^2$. Erzeugen Sie die Matrix $B \in \mathbb{R}^{2 \times n}$ durch n -fache Replikation des Vektors b .
- Berechnen Sie die linear-affin transformierte Datenmatrix $Z = AY + B$. Vergewenwärtigen Sie sich, was passiert, wenn die Matrix Y von links mit der Matrix A multipliziert wird. Erläutern Sie, woraus die Zeilen der transformierten Datenmatrix Z bestehen.
- Berechnen Sie die linear-affin transformierten Parameterschätzer $\hat{\mu}^{(y)} = A\hat{\mu}^{(x)} + b$ und $\hat{\Sigma}^{(y)} = A\hat{\Sigma}^{(x)}A^T$, wobei $\hat{\mu}^{(x)}$ und $\hat{\Sigma}^{(x)}$ die Schätzer aus Abschnitt 4.3 sind. Speichern Sie die Ergebnisse erneut als die Variablen `mu_est` und `Sigma_est`.
- Erzeugen Sie mit den Ergebnissen die Abbildung wie oben. Ersetzen Sie hierfür `mu_est[3:4]` und `Sigma_est[3:4,3:4]` sowie `Y[3,]` und `Y[4,]` durch die passenden neuen Werte. Sie sollten in etwa folgende Abbildung erhalten:

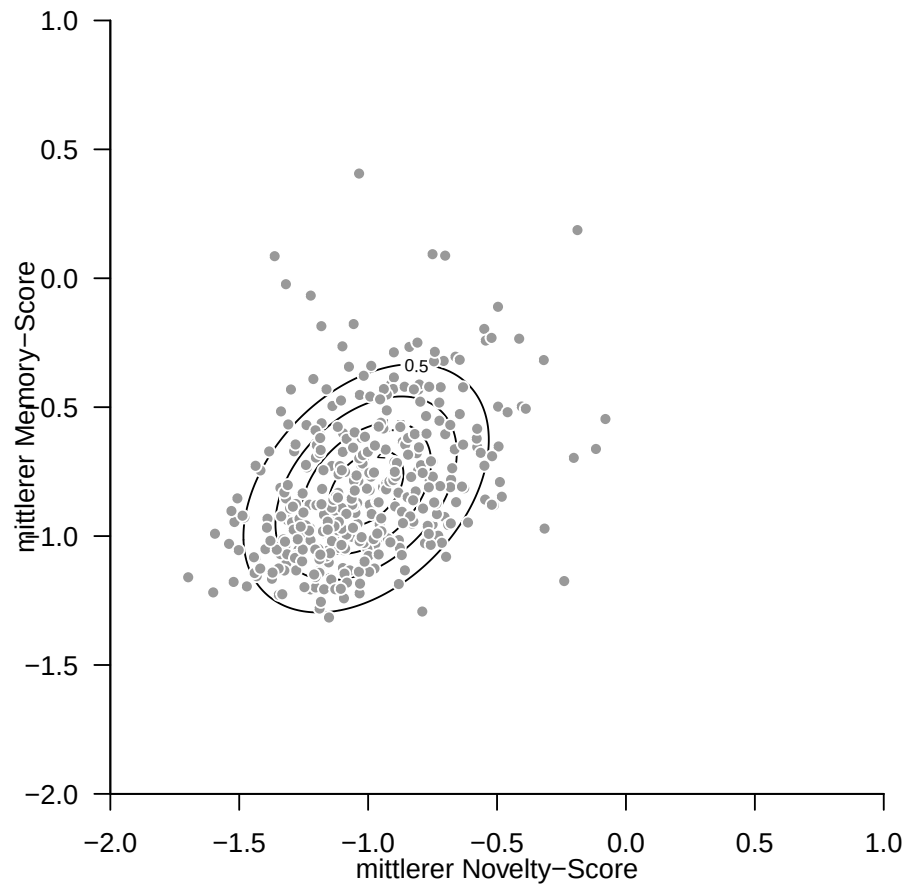


Abbildung 2. Beobachtete mittlere Novelty-Scores und mittlere Memory-Scores sowie geschätzte Wahrscheinlichkeitsdichtefunktion der ihnen zugrundeliegenden bivariaten Normalverteilung.

4.6 Lückentext

Füllen Sie mit den in der Übung gewonnenen Erkenntnissen den folgenden Lückentext aus und präsentieren Sie die Ergebnisse im Seminar:

Lückentext: Die geschätzten Erwartungswerte der vier Score-Variablen sind _____, _____, _____ und _____. Die geschätzten Varianzen der vier Score-Variablen sind _____, _____, _____ und _____. Der Memory-FADE-Score und der Memory-SAME-Score sind im erhobenen Datensatz _____ (positiv/negativ) korreliert. Die mittleren Novelty-Scores und die mittleren Memory-Scores sind im erhobenen Datensatz _____ (positiv/negativ) korreliert. Die geschätzte Normalverteilung in Abbildung _____ kommt einer sphärischen Normalverteilung näher als die geschätzte Normalverteilung in Abbildung _____.

4.7 Mögliche Klausurfrage

Präsentieren Sie im Seminar folgende Klausurfrage und erklären Sie die richtige Antwort:

Frage: Was besagt das Theorem zur linear-affinen Transformation multivariater Normalverteilungen?

- a) Die marginalen Verteilungen einer multivariaten Normalverteilung sind ebenfalls multivariate Normalverteilungen.
- b) Die bedingten Verteilungen einer multivariaten Normalverteilung sind ebenfalls multivariate Normalverteilungen.
- c) Die Parameter einer gemeinsamen Normalverteilungen ergeben sich als linear-affine Transformation der Parameter der sie induzierenden marginalen und bedingten Normalverteilungen.
- d) Jede linear-affine Transformation eines multivariat normalverteilten Zufallsvektors ist ebenfalls multivariat normalverteilt.

4.8 Mathematischer Witz

$$\sqrt{-1} \quad 2^3 \quad \sum \quad \pi$$

isnoilep saw ti dne ...