Allgemeines Lineares Modell

BSc Psychologie, SoSe 2024

Joram Soch

(4) Normalverteilungen

Überblick

Die multivariate Normalverteilung ist zentraler Bestandteil des ALMs.

Sie ist die multivariate Generalisierung der univariaten Normalverteilung.

Die Motivation von Normalverteilungsannahmen liegt immer im Zentralen Grenzwertsatz:

- Probabilistische Terme repräsentieren die Summation sehr vieler Prozesse, die durch die deterministischen Bestandteile eines Modell, also eine mechanistische wissenschaftliche Theorie, nicht erklärt werden.
- Nach dem Zentralen Grenzwertsatz ist die Summe dieser Prozesse normalverteilt.

Darüberhinaus hat die Normalverteilung günstige mathematische Eigenschaften, die auch in der Quantifizierung subjektiver Unsicherheit genutzt werden können.

Zur Erarbeitung der Inhalte dieser Einheit bietet sich eine Wiederholdung von Zufallsvektoren an (vgl. Einheit (4) in *Wahrscheinlichkeitstheorie und Frequentistische Inferenz*).

Kontinuierliche Zufallsvektoren

Konstruktion und Definition

Transformationen

Sphärische Verteilungen

Marginale und bedingte Verteilungen

Selbstkontrollfragen

Kontinuierliche Zufallsvektoren

Konstruktion und Definition

Transformationen

Sphärische Verteilungen

Marginale und bedingte Verteilungen

Selbstkontrollfragen

Kontinuierliche Zufallsvektoren

Definition (Zufallsvektor)

$(\Omega, \mathcal{A}, \mathbb{P})$ sei ein Wahrscheinlichkeitsraum und $(\mathcal{X}, \mathcal{S})$ sei ein n-dimensionaler Messraum. Ein n-dimensionaler *Zufallsvektor* ist definiert als eine Abbildung

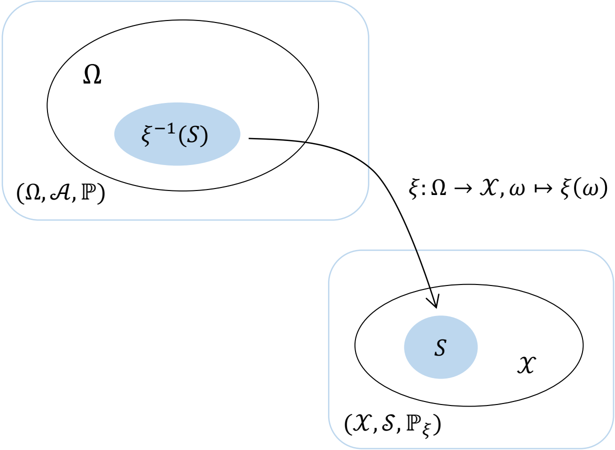
$$\xi : \Omega \to \mathcal{X}, \omega \mapsto \xi(\omega) := \begin{pmatrix} \xi_1(\omega) \\ \vdots \\ \xi_n(\omega) \end{pmatrix} \tag{1}$$

mit der *Messbarkeitseigenschaft*

$$\{\omega \in \Omega | \xi(\omega) \in S\} \in \mathcal{A} \quad \text{für alle} \quad S \in \mathcal{S}. \tag{2}$$

Bemerkungen

- ξ ist hier eine univariate, vektorwertige Abbildung.
- Das Standardbeispiel für $(\mathcal{X}, \mathcal{S})$ ist $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.
- Wir verzichten auf eine explizite Einführung n-dimensionaler σ-Algebren wie $\mathcal{B}(\mathbb{R}^n)$.
- Ohne Beweis halten wir fest, dass ξ messbar ist, wenn die Funktionen $\xi_1, ..., \xi_n$ messbar sind.
- Die Komponentenfunktionen eines Zufallsvektors sind Zufallsvariablen.
- Ein n-dimensionaler Zufallsvektor ist die Konkatenation von n Zufallsvariablen.
- Ein ein-dimensionaler Zufallsvektor ($n := 1$) ist eine Zufallsvariable.



$$\xi : \Omega \to \mathcal{X}, \omega \mapsto \xi(\omega) := \begin{pmatrix} \xi_1(\omega) \\ \vdots \\ \xi_n(\omega) \end{pmatrix}$$

$$\mathbb{P}\big(\xi^{-1}(S)\big) = \mathbb{P}(\{\omega \in \Omega | \xi(\omega) \in S\}) =: \mathbb{P}_\xi(S)$$

Definition (Kontinuierlicher Zufallsvektor, Multivariate WDF)

Ein Zufallsvektor ξ heißt *kontinuierlich*, wenn $\mathbb{R}^n$ der Ergebnisraum von ξ ist und eine Funktion

$$p_\xi : \mathbb{R}^n \to \mathbb{R}_{\geq 0}, x \mapsto p_\xi(x), \tag{3}$$

existiert, für die gilt

(1) $\int_{\mathbb{R}^n} p_\xi(x)\, dx = 1$ und

(2) $\mathbb{P}_\xi(x_1 \leq \xi \leq x_2) = \int_{x_{1_1}}^{x_{2_1}} \cdots \int_{x_{1_n}}^{x_{2_n}} p_\xi(s_1, ..., s_n)\, ds_1 \cdots ds_n$.

Eine entsprechende Funktion p heißt *multivariate Wahrscheinlichkeitsdichtefunktion* (WDF) von ξ.

Bemerkungen

- Der Begriff der multivariaten WDF ist analog zum Begriff der WDF.
- Man spricht häufig auch einfach von der *WDF eines Zufallsvektors*.
- Wie univariate WDFen sind multivariate WDFen nicht-negativ und normiert.
- Wie für kontinuierliche Zufallsvariablen gilt für kontinuierliche Zufallsvektoren:

$$\mathbb{P}_\xi(\xi = x) = \mathbb{P}_\xi(x \leq \xi \leq x) = \int_{x_1}^{x_1} \cdots \int_{x_n}^{x_n} p_\xi(s_1, ..., s_n)\, ds_1 \cdots ds_n = 0\,. \tag{4}$$

Definition (Erwartungswert und Kovarianzmatrix von Zufallsvektoren)

ξ sei ein n-dimensionaler Zufallvektor. Dann ist der *Erwartungwert* von ξ definiert als der n-dimensionale Vektor

$$\mathbb{E}(\xi) := \begin{pmatrix} \mathbb{E}(\xi_1) \\ \vdots \\ \mathbb{E}(\xi_n) \end{pmatrix} \tag{5}$$

und die *Kovarianzmatrix* von ξ ist definiert als die $n \times n$ Matrix

$$\mathbb{C}(\xi) := \mathbb{E}\left((\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^{\mathrm{T}}\right) \tag{6}$$

Bemerkungen

- Der Erwartungswert von ξ ist der Vektor der Erwartungswerte $\mathbb{E}(\xi_1), ..., \mathbb{E}(\xi_n)$.
- Der Erwartungswert eines Zufallsvektors ist also *elementweise* definiert.
- Die Kovarianzmatrix ist formal analog zur Kovarianz zweier Zufallsvariablen definiert.

Theorem (Kovarianzmatrix eines Zufallsvektors)

ξ sei ein n-dimensionaler Zufallvektor und $\mathbb{C}(\xi)$ sei seine Kovarianzmatrix. Dann gilt:

$$\mathbb{C}(\xi) = \big(\mathbb{C}(\xi_i, \xi_j)\big)_{1 \le i,j \le n} = \begin{pmatrix} \mathbb{C}(\xi_1, \xi_1) & \mathbb{C}(\xi_1, \xi_2) & \cdots & \mathbb{C}(\xi_1, \xi_n) \\ \mathbb{C}(\xi_2, \xi_1) & \mathbb{C}(\xi_2, \xi_2) & \cdots & \mathbb{C}(\xi_2, \xi_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{C}(\xi_n, \xi_1) & \mathbb{C}(\xi_n, \xi_2) & \cdots & \mathbb{C}(\xi_n, \xi_n) \end{pmatrix}. \tag{7}$$

Bemerkung

- Die Kovarianzmatrix $\mathbb{C}(\xi)$ ist also die Matrix der Kovarianzen der Komponenten von ξ.

Kontinuierliche Zufallsvektoren

<u>Beweis</u>

Es gilt

$$
\begin{aligned}
\mathbb{C}(\xi) &:= \mathbb{E}\left((\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^{\mathrm{T}}\right) \\
&= \mathbb{E}\left(\left(\begin{pmatrix}\xi_1 \\ \vdots \\ \xi_n\end{pmatrix} - \begin{pmatrix}\mathbb{E}(\xi_1) \\ \vdots \\ \mathbb{E}(\xi_n)\end{pmatrix}\right)\left(\begin{pmatrix}\xi_1 \\ \vdots \\ \xi_n\end{pmatrix} - \begin{pmatrix}\mathbb{E}(\xi_1) \\ \vdots \\ \mathbb{E}(\xi_n)\end{pmatrix}\right)^{\mathrm{T}}\right) \\
&= \mathbb{E}\left(\begin{pmatrix}\xi_1 - \mathbb{E}(\xi_1) \\ \vdots \\ \xi_n - \mathbb{E}(\xi_n)\end{pmatrix}\begin{pmatrix}\xi_1 - \mathbb{E}(\xi_1) \\ \vdots \\ \xi_n - \mathbb{E}(\xi_n)\end{pmatrix}^{\mathrm{T}}\right) \\
&= \mathbb{E}\left(\begin{pmatrix}\xi_1 - \mathbb{E}(\xi_1) \\ \vdots \\ \xi_n - \mathbb{E}(\xi_n)\end{pmatrix}\begin{pmatrix}\xi_1 - \mathbb{E}(\xi_1) & \dots & \xi_n - \mathbb{E}(\xi_n)\end{pmatrix}\right) \\
&= \mathbb{E}\begin{pmatrix}(\xi_1 - \mathbb{E}(\xi_1))(\xi_1 - \mathbb{E}(\xi_1)) & \dots & (\xi_1 - \mathbb{E}(\xi_1))(\xi_n - \mathbb{E}(\xi_n) \\ \vdots & \ddots & \vdots \\ (\xi_n - \mathbb{E}(\xi_n))(\xi_1 - \mathbb{E}(\xi_1)) & \dots & (\xi_n - \mathbb{E}(\xi_n))(\xi_n - \mathbb{E}(\xi_n))\end{pmatrix} \\
&= \left(\mathbb{E}\left((\xi_i - \mathbb{E}(\xi_i))(\xi_j - \mathbb{E}(\xi_j))\right)\right)_{1 \le i,j \le n} \\
&= \left(\mathbb{C}(\xi_i, \xi_j)\right)_{1 \le i,j \le n}
\end{aligned}
$$

$\square$

Definition (Normalverteilte Zufallsvariable)

ξ sei eine Zufallsvariable mit Ergebnisraum $\mathbb{R}$ und WDF

$$p : \mathbb{R} \to \mathbb{R}_{>0}, x \mapsto p(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right). \tag{8}$$

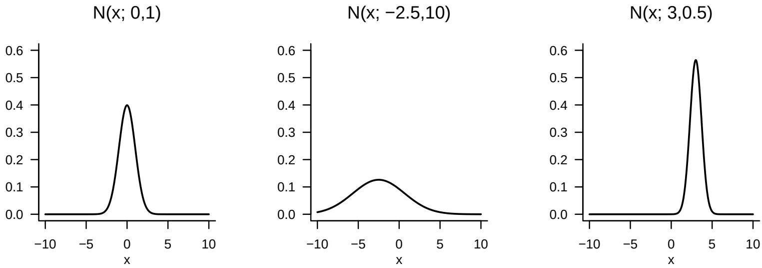
Dann sagen wir, dass ξ einer *Normalverteilung (oder Gauß-Verteilung) mit Erwartungswertparameter* $\mu \in \mathbb{R}$ *und Varianzparameter* $\sigma^2 > 0$ unterliegt und nennen ξ eine *normalverteilte Zufallsvariable*. Wir kürzen dies mit $\xi \sim N(\mu, \sigma^2)$ ab. Die WDF einer normalverteilten Zufallsvariable bezeichnen wir mit

$$N\left(x; \mu, \sigma^2\right) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right). \tag{9}$$

Bemerkungen

- Es gelten $\mathbb{E}(\xi) = \mu$ und $\mathbb{V}(\xi) = \sigma^2$.
- Der Parameter μ entspricht dem Wert höchster Wahrscheinlichkeitsdichte.
- Der Parameter σ^2 spezifiziert die Breite der WDF.
- $\xi \sim N(0, 1)$ heißt auch *standardnormalverteilt*.

Visualisierung univariater Normalverteilungsdichtefunktionen



Theorem (Konstruktion bivariater Normalverteilungen)

$\zeta_1 \sim N(0,1)$ und $\zeta_2 \sim N(0,1)$ seien zwei unabhängige standardnormalverteilte Zufallsvariablen. Weiterhin seien $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$ und $\rho \in]-1,1[$. Schließlich seien

$$
\begin{aligned}
\xi_1 &:= \sigma_1 \zeta_1 + \mu_1 \\
\xi_2 &:= \sigma_2 \left(\rho\zeta_1 + \sqrt{1-\rho^2}\zeta_2 \right) + \mu_2 .
\end{aligned} \tag{10}
$$

Dann hat die WDF des Zufallsvektors $\xi := (\xi_1, \xi_2)^{\mathrm{T}}$, also der gemeinsamen Verteilung von ξ_1 und ξ_2, die Form

$$
p : \mathbb{R}^2 \to \mathbb{R}_{>0},\, x \mapsto p(x) := \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2} (x-\mu)^{\mathrm{T}} \Sigma^{-1} (x-\mu) \right), \tag{11}
$$

wobei

$$
n := 2, \quad \mu := \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{und} \quad \Sigma := \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix} \tag{12}
$$

Bemerkungen

- Für einen Beweis, siehe DeGroot and Schervish (2012), S. 338-339.
- Man nennt die gemeinsame Verteilung von ξ_1 und ξ_2 *bivariate Normalverteilung*.

Konstruktion bivariater Normalverteilungen

```r
# Parameterdefinitionen
mu_1   = 5.0                                             # \mu_1
mu_2   = 4.0                                             # \mu_2
sig_1  = 1.5                                             # \sigma_1
sig_2  = 1.0                                             # \sigma_2
rho    = 0.9                                             # \rho

# Realisierungen der standardnormalverteilten ZVen
nr     = 100                                             # Anzahl Realisierungen
zeta_1 = rnorm(nr)                                       # \zeta_1 \sim N(0,1)
zeta_2 = rnorm(nr)                                       # \zeta_1 \sim N(0,1)

# Evaluation von Realisierungen von \xi_1 und \xi_2
xi_1   = sig_1*zeta_1 + mu_1                             # Realsierungen von zeta_1
xi_2   = sig_2*(rho*zeta_1 + sqrt(1-rho^2)*zeta_2) + mu_2 # Realsierungen von zeta_2

# Parameter der gemeinsamen Verteilung von \xi_1 und \xi_2
mu     = matrix(c(mu_1,                                  # \mu \in \mathbb{R}^2
                  mu_2),
                nrow = 2, byrow = TRUE)
Sigma  = matrix(c(sig_1^2,          rho*sig_1*sig_2,     # \Sigma \in \mathbb{R}^{2 x 2}
                  rho*sig_1*sig_2, sig_2^2),
                nrow = 2, byrow = TRUE)
print(mu)
```

```
>      [,1]
> [1,]    5
> [2,]    4
```

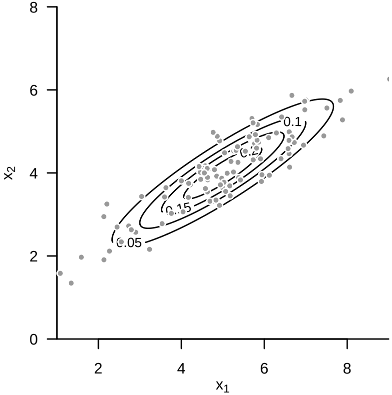
```r
print(Sigma)
```

```
>      [,1] [,2]
> [1,] 2.25 1.35
> [2,] 1.35 1.00
```

Konstruktion bivariater Normalverteilungen

- Realisierungen von $\xi = (\xi_1, \xi_2)^{\mathrm{T}}$

— Isokonturen (Linien gleicher Wahrscheinlichkeitsdichte) von p



Animation zur Konstruktion

Definition (Multivariate Normalverteilung)

ξ sei ein n-dimensionaler Zufallsvektor mit Ergebnisraum $\mathbb{R}^n$ und WDF

$$p : \mathbb{R}^n \to \mathbb{R}_{>0},\, x \mapsto p(x) := \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)\right). \tag{13}$$

Dann sagen wird, dass ξ einer *multivariaten (oder n-dimensionalen) Normalverteilung* mit *Erwartungswertparameter* $\mu \in \mathbb{R}^n$ und positiv-definitem *Kovarianzmatrixparameter* $\Sigma \in \mathbb{R}^{n \times n}$ unterliegt und nennen ξ einen *(multivariat) normalverteilten Zufallsvektor*. Wir kürzen dies mit $\xi \sim N(\mu, \Sigma)$ ab. Die WDF eines multivariat normalverteilten Zufallsvektors bezeichnen wir mit

$$N(x; \mu, \Sigma) := \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)\right). \tag{14}$$

Bemerkungen

- Der Parameter $\mu \in \mathbb{R}^n$ entspricht dem Wert höchster Wahrscheinlichkeitsdichte.
- Die Diagonalelemente von Σ spezifizieren die Breite der WDF bezüglich $\xi_1, ..., \xi_n$.
- Das (i, j)te Element von Σ spezifiziert die Kovarianz on ξ_i und ξ_j.
- Der Term $1/\sqrt{(2\pi)^n|\Sigma|}$ ist die Normalisierungskonstante für den Exponentialfunktionsterm.

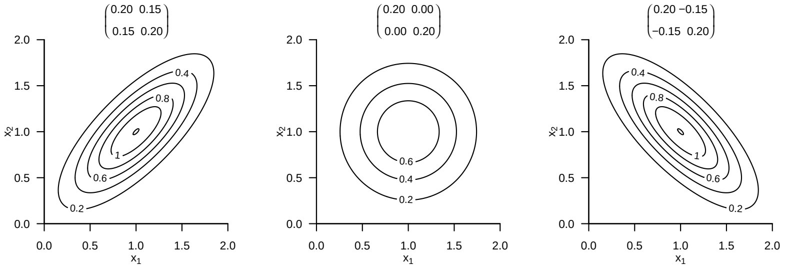
Visualisierung bivariater Normalverteilungsdichtefunktionen

```
library(mvtnorm)                                      # Tools für die multivariate Normalverteilung

# Ergebnisraumdefintion
x_min  = 0                                            # x_i Minimum
x_max  = 2                                            # x_i Maximum
x_res  = 1e3                                          # x_i Auflösung
x_1    = seq(x_min, x_max, length.out = x_res)        # x_1 Raum
x_2    = seq(x_min, x_max, length.out = x_res)        # x_2 Raum
X      = expand.grid(x_1,x_2)                         # X = (x_1,x_2)^T Raum

# Parameterdefinition
mu     = c(1,1)                                       # \mu \in \mathbb{R}^2
S      = list(matrix(c(0.2,  0.15,  0.15, 0.2), 2),   # \Sigma in \mathbb{R}^{2 \times 2}
              matrix(c(0.2,  0.00,  0.00, 0.2), 2),   # \Sigma in \mathbb{R}^{2 \times 2}
              matrix(c(0.2, -0.15, -0.15, 0.2), 2))   # \Sigma in \mathbb{R}^{2 \times 2}

# Schleife über Kovarianzmatrixparametervarianten
for (Sigma in S){

    # Berechnung der Wahrscheinlichkeitsdichtefunktion
    p  = matrix(dmvnorm(as.matrix(X), mu, Sigma),     # Matrixkonversion des von
                nrow = x_res)                         # dmvnorm() ausgegebenen Vektors

    # Visualisierung der Wahrscheinlichkeitsdichtefunktion
    contour(x_1, x_2, p,
        xlim      = c(x_min,x_max),
        ylim      = c(x_min,x_max),
        nlevels   = 5)

}
```

Konstruktion und Definition

Visualisierung bivariater Normalverteilungsdichtefunktionen

$$\mu = (1,1)^{\mathrm{T}},\ \Sigma \in \mathbb{R}^{2\times 2} \text{ wie im Abbildungstitel vermerkt.}$$



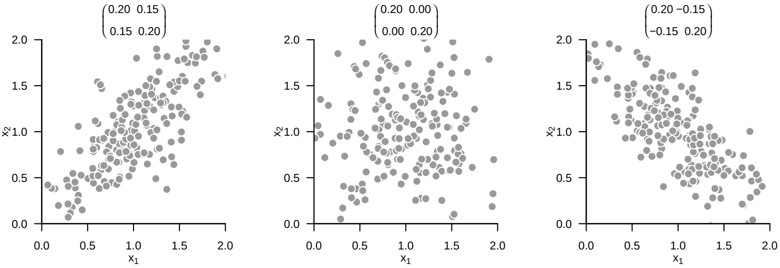
Realisierung bivariater normalverteilter Zufallsvektoren

```
library(mvtnorm)                                   # Tools für die multivariate Normalverteilung

# Parameterdefinition
mu     = c(1,1)                                    # \mu \in \mathbb{R}^2
Sigma  = matrix(c(0.2, 0.15, 0.15, 0.2), 2)        # \Sigma in \mathbb{R}^{2 \times 2}

# Realisierungen des Zufallsvektors
t(rmvnorm(n = 10, mu, Sigma))                      # multivariate Zufallsvektoren
```

```
>      [,1]  [,2]   [,3]  [,4]  [,5]  [,6]  [,7] [,8] [,9] [,10]
> [1,] 1.08 0.921 0.0148 0.436 0.646 0.467 1.087 1.23 1.30 0.581
> [2,] 1.48 0.876 0.5765 0.540 1.700 0.649 0.627 1.13 1.27 1.116
```

Realisierung bivariater normalverteilter Zufallsvektoren

$\mu = (1, 1)^{\mathrm{T}}, \ \Sigma \in \mathbb{R}^{2\times 2}$ wie im Abbildungstitel vermerkt ($n = 200$).



Theorem (Erwartungswert und Kovarianzmatrix)

$\xi \sim N(\mu, \Sigma)$ sei ein multivariate normalverteilter Zufallsvektor mit Erwartungswertparameter $\mu \in \mathbb{R}^n$ und Kovarianzmatrixparameter $\Sigma \in \mathbb{R}^{n \times n}$ p.d. . Dann gelten für den Erwartungswert und die Kovarianzmatrix von ξ, dass

$$\mathbb{E}(\xi) = \mu \quad \text{und} \quad \mathbb{C}(\xi) = \Sigma\,. \tag{15}$$

Bemerkungen

- Wir verzichten auf einen Beweis. Verweise befinden sich im Anhang.
- Das Theorem ist die direkte Generalisierung der Eigenschaften univariater normalverteilter Zufallsvariablen

Kontinuierliche Zufallsvektoren

Konstruktion und Definition

Transformationen

Sphärische Verteilungen

Marginale und bedingte Verteilungen

Selbstkontrollfragen

Theorem (Linear-affine Transformation)

$\xi \sim N(\mu, \Sigma)$ sei ein normalverteilter n-dimensionaler Zufallsvektor und es sei $\upsilon := A\xi + b$ mit $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$. Dann gilt

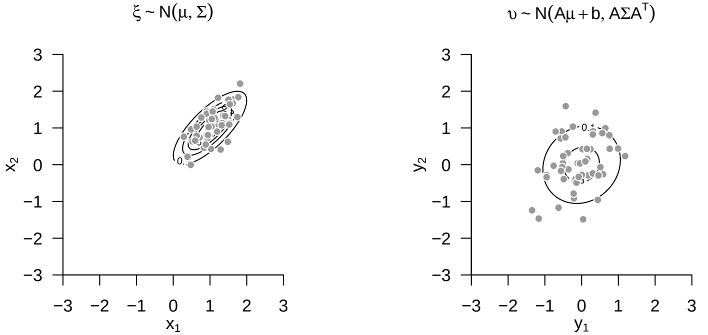
$$\upsilon \sim N\left(A\mu + b, A\Sigma A^{\mathrm{T}}\right) \tag{16}$$

Bemerkung

- Für einen Beweis, siehe Anderson (2003), Abschnitt 2.4 oder die Verweise im Anhang.
- Die linear-affine Transformation eines multivariat normalverteilten Zuallsvektors ergibt wieder einen multivariat normalverteilten Zufallsvektor. Die Parameter des resultierenden normalverteilten Zufallsvektors ergeben sich dabei aus den Parametern des ursprünglichen Zufallsvektors und den Transformationsparametern.

Linear-affine Transformation

$$n := 2,\ \mu := \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ \Sigma := \begin{pmatrix} 0.20 & 0.15 \\ 0.15 & 0.20 \end{pmatrix},\ A := \begin{pmatrix} -2 & 1 \\ -1 & 2 \end{pmatrix},\ b := \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$



Theorem (Invertierbare lineare Transformation)

$\xi \sim N(\mu, \Sigma)$ sei ein normalverteilter n-dimensionaler Zufallsvektor und es sei $\upsilon := A\xi$ mit einer invertierbaren Matrix $A \in \mathbb{R}^{n \times n}$. Dann gilt

$$\upsilon \sim N\left(A\mu, A\Sigma A^{\mathrm{T}}\right) \tag{17}$$

Bemerkung

- Es handelt sich hierbei um einen Spezialfall der linear-affinen Transformation mit $A \in \mathbb{R}^{n \times n}$ und $b = 0_n$.
- Die invertierbare lineare Transformation eines multivariat normalverteilten Zuallsvektors ergibt wieder einen multivariat normalverteilten Zufallsvektor. Die Parameter des resultierenden normalverteilten Zufallsvektors ergeben sich dabei aus den Parametern des ursprünglichen Zufallsvektors und der Transformationsmatrix.

Theorem (Sphärische multivariate Normalverteilung)

Für $i = 1, ..., n$ seien $N(x_i; \mu_i, \sigma^2)$ die WDFen von n unabhängigen univariaten normalverteilten Zufallsvariablen $\xi_1, ..., \xi_n$ mit $\mu_1, ..., \mu_n \in \mathbb{R}$ und $\sigma^2 > 0$. Weiterhin sei $N(x; \mu, \sigma^2 I_n)$ die WDF eines n-dimensionalen Zufallsvektors ξ mit Erwartungswertparameter $\mu := (\mu_1, ..., \mu_n)^{\mathrm{T}} \in \mathbb{R}^n$. Dann gilt

$$p_\xi(x) = p_{\xi_1,...,\xi_n}(x_1, ..., x_n) = \prod_{i=1}^n p_{\xi_i}(x_i) \tag{18}$$

und damit im vorliegenden Fall normalverteilter Zufallsvariablen

$$N\left(x; \mu, \sigma^2 I_n\right) = \prod_{i=1}^n N\left(x_i; \mu_i, \sigma^2\right). \tag{19}$$

Bemerkungen

- Das Theorem ist für die Theories des Allgemeinen Linearen Modells zentral.
- Die Zufallsvariablen ξ_i sind unabhängig, aber nicht notwendigerweise identisch verteilt.
- Einen Kovarianzmatrixparameter der Form $\Sigma = \sigma^2 I_n$ nennt man auch *sphärisch*; eine multivariate Normalverteilung mit spährischem Kovarianzmatrixparameter nennt man auch *sphärische Normalverteilung*.
- Sphärische n-variate Normalverteilungen entsprechen n unabhängigen univariaten Normalverteilungen und umgekehrt; eine Realisierung einer n-variaten sphärischen Normalverteilung entspricht n Realisierungen von unabhängigen univariaten Normalverteilungen.

Beweis

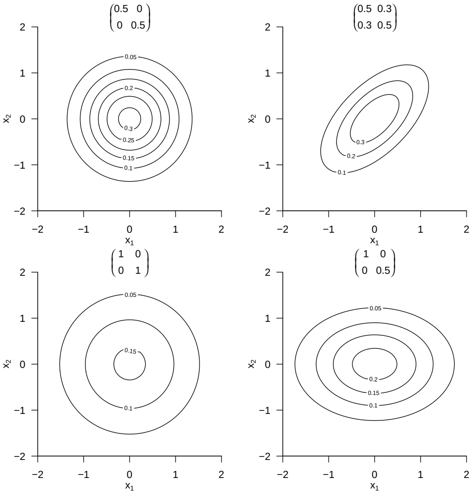
Wir zeigen die Identität der multivariaten WDF $N(x;\mu,\sigma^2 I_n)$ mit dem Produkt von n univariaten WDFen $N(x_i;\mu_i,\sigma^2 I_n)$, wobei μ_i der $i$te Eintrag von $\mu \in \mathbb{R}^n$ ist. Es ergibt sich

$$
\begin{aligned}
N\left(x;\mu,\sigma^2 I_n\right) &= \frac{1}{\sqrt{(2\pi)^n\left|\sigma^2 I_n\right|}} \exp\left(-\frac{1}{2}(x-\mu)^{\mathrm{T}}(\sigma^2 I_n)^{-1}(x-\mu)\right) \\
&= \left(\prod_{i=1}^n \frac{1}{\sqrt{2\pi}}\right) \frac{1}{\sqrt{(\sigma^2)^n}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^{\mathrm{T}}(x-\mu)\right) \\
&= \left(\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}}\right) \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu_i)^2\right) \\
&= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \prod_{i=1}^n \exp\left(-\frac{1}{2\sigma^2}(x_i-\mu_i)^2\right) \\
&= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x_i-\mu_i)^2\right) \\
&= \prod_{i=1}^n N\left(x_i;\mu_i,\sigma^2\right).
\end{aligned}
\tag{20}
$$

$\square$

Beispiele für sphärische/nicht-sphärische Verteilungen

$\mu = (0,0)^{\mathrm{T}},\ \Sigma \in \mathbb{R}^{2\times 2}$ wie im Abbildungstitel vermerkt.



Kontinuierliche Zufallsvektoren

Konstruktion und Definition

Transformationen

Sphärische Verteilungen

Marginale und bedingte Verteilungen

Selbstkontrollfragen

Marginale und bedingte Verteilungen

Multivariate Normalverteilungen haben die Eigenschaft, dass auch alle anderen assoziierten Verteilungen Normalverteilungen sind, deren Erwartungswert- und Kovarianzmatrixparameter aus den Parametern der jeweils komplementären Verteilung errechnet werden können.

Insbesondere gelten:

(1) Die uni- und multivariaten Marginalverteiltungen multivariater Normalverteilungen sind Normalverteilungen.
(2) Die uni- und multivariaten bedingten Verteiltungen multivariater Normalverteilungen sind Normalverteilungen.
(3) Wie alle gemeinsamen Verteilungen lassen sich multivariate Normalverteilungen multiplikativ in eine marginale und eine bedingte Verteilung zerlegen. Diese sind bei multivariaten Normalverteilungen auch Normalverteilungen, deren Parameter aus den Parametern der gemeinsamen Verteilung errechnet werden können und umgekehrt.

Die Resultate dieses Abschnitts sind insbesondere für generalisierte ALMs zentral, bspw. Linear Mixed Models bzw. Hierarchische Normalverteilungsmodelle, die Varianzkomponentenschätzung und die Bayesianische Inferenz bei Normalverteilungsannahmen, inklusive z.B. Kalman-Bucy-Filtering. Sie spielen aber auch im Rahmen der partiellen Korrelation bei ALMs eine Rolle (siehe Einheit (12) in *Allgemeines Lineares Modell*).

Theorem (Marginale Normalverteilungen)

Es sei $m := k + l$ und $\xi = (\xi_1, ..., \xi_m)^{\mathrm{T}}$ sei ein m-dimensionaler normalverteilter Zufallsvektor mit Erwartungswertparameter

$$\mu = \begin{pmatrix} \mu_\upsilon \\ \mu_\zeta \end{pmatrix} \in \mathbb{R}^m, \tag{21}$$

mit $\mu_\upsilon \in \mathbb{R}^k$ und $\mu_\zeta \in \mathbb{R}^l$ und Kovarianzmatrixparameter

$$\Sigma = \begin{pmatrix} \Sigma_{\upsilon\upsilon} & \Sigma_{\upsilon\zeta} \\ \Sigma_{\zeta\upsilon} & \Sigma_{\zeta\zeta} \end{pmatrix} \in \mathbb{R}^{m \times m}, \tag{22}$$

mit $\Sigma_{\upsilon\upsilon} \in \mathbb{R}^{k \times k}$, $\Sigma_{\upsilon\zeta} \in \mathbb{R}^{k \times l}$, $\Sigma_{\zeta\upsilon} \in \mathbb{R}^{l \times k}$ und $\Sigma_{\zeta\zeta} \in \mathbb{R}^{l \times l}$. Dann sind $\upsilon := (\xi_1, ..., \xi_k)^{\mathrm{T}}$ und $\zeta := (\xi_{k+1}, ..., \xi_m)^{\mathrm{T}}$ k- bzw. l-dimensionale normalverteilte Zufallsvektoren und es gilt:

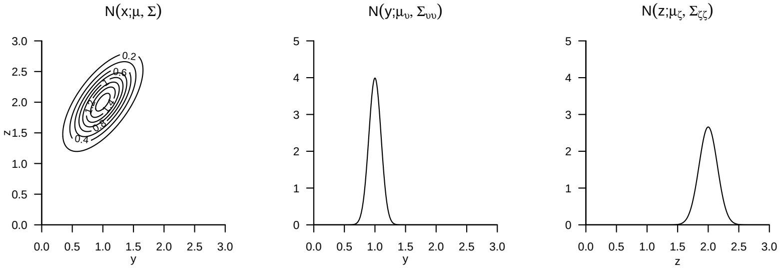
$$\upsilon \sim N(\mu_\upsilon, \Sigma_{\upsilon\upsilon}) \quad \text{und} \quad \zeta \sim N(\mu_\zeta, \Sigma_{\zeta\zeta})\,. \tag{23}$$

Bemerkungen

- Die Marginalverteilungen einer multivariaten Normalverteilung sind auch Normalverteilungen.
- Die Parameter der Marginalverteilungen ergeben sich aus den Parametern der gemeinsamen Verteilung.
- Für Beweise, siehe z.B. Mardia, Kent, and Bibby (1979), Kapitel 3 oder Anderson (2003), Kapitel 2.

Marginale Normalverteilungen

$$m := 2,\ k := 1,\ l := 1,\ \mu := \begin{pmatrix} 1 \\ 2 \end{pmatrix},\ \Sigma := \begin{pmatrix} 0.10 & 0.08 \\ 0.08 & 0.15 \end{pmatrix}$$



Theorem (Bedingte Normalverteilungen)

$\begin{pmatrix} \xi \\ \upsilon \end{pmatrix}$ sei ein $m + n$-dimensionaler normalverteilter Zufallsvektor mit WDF

$$p_{\xi,\upsilon} : \mathbb{R}^{m+n} \to \mathbb{R}_{>0}, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto p_{\xi,\upsilon}\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) := N\left(\begin{pmatrix} x \\ y \end{pmatrix}; \mu_{\xi,\upsilon}, \Sigma_{\xi,\upsilon}\right) \tag{24}$$

mit

$$\mu_{\xi,\upsilon} = \begin{pmatrix} \mu_\xi \\ \mu_\upsilon \end{pmatrix} \quad \text{und} \quad \Sigma_{\xi,\upsilon} = \begin{pmatrix} \Sigma_{\xi\xi} & \Sigma_{\xi\upsilon} \\ \Sigma_{\upsilon\xi} & \Sigma_{\upsilon\upsilon} \end{pmatrix}, \tag{25}$$

wobei $x, \mu_\xi \in \mathbb{R}^m, y, \mu_\upsilon \in \mathbb{R}^n$ and $\Sigma_{\xi\xi} \in \mathbb{R}^{m\times m}, \Sigma_{\xi\upsilon} \in \mathbb{R}^{m\times n}, \Sigma_{\upsilon\upsilon} \in \mathbb{R}^{n\times n}$. Dann ist die bedingte Verteilung von ξ gegeben υ eine m-dimensionale Normalverteilung mit bedingter WDF

$$p_{\xi|\upsilon}(\cdot|y) : \mathbb{R}^m \to \mathbb{R}_{>0}, x \mapsto p_{\xi|\upsilon}(x|y) := N(x; \mu_{\xi|\upsilon}, \Sigma_{\xi|\upsilon}) \tag{26}$$

mit

$$\mu_{\xi|\upsilon} = \mu_\xi + \Sigma_{\xi\upsilon}\Sigma_{\upsilon\upsilon}^{-1}(y - \mu_\upsilon) \in \mathbb{R}^m \tag{27}$$

und

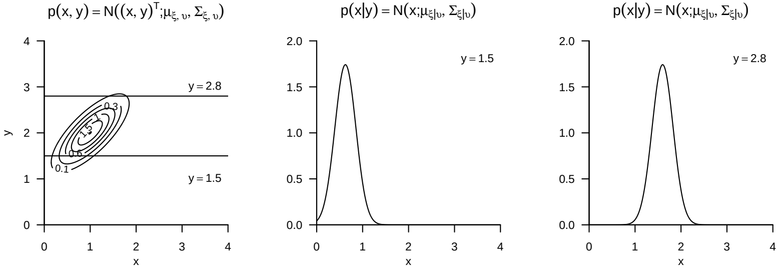
$$\Sigma_{\xi|\upsilon} = \Sigma_{\xi\xi} - \Sigma_{\xi\upsilon}\Sigma_{\upsilon\upsilon}^{-1}\Sigma_{\upsilon\xi} \in \mathbb{R}^{m\times m}. \tag{28}$$

Bemerkungen

- Die Parameter einer bedingten (multivariaten) Normalverteilung ergeben sich aus den Parametern einer gemeinsamen multivariaten Normalverteilung. Im Zusammenspiel mit den vorherigen Theoremen können die Parameter bedingter und marginale Normalverteilungen aus den Parametern der komplementären bedingten und marginalen Normalverteilungen bestimmt werden.

Bedingte Normalverteilungen

$$m := 2,\ n := 1,\ \mu := \begin{pmatrix} 1 \\ 2 \end{pmatrix},\ \Sigma := \begin{pmatrix} 0.12 & 0.09 \\ 0.09 & 0.12 \end{pmatrix}$$



Theorem (Gemeinsame Normalverteilungen)

ξ sei ein m-dimensionaler normalverteilter Zufallsvektor mit WDF

$$p_\xi : \mathbb{R}^m \to \mathbb{R}_{>0},\, x \mapsto p_\xi(x) := N(x; \mu_\xi, \Sigma_{\xi\xi}) \quad \text{mit} \quad \mu_\xi \in \mathbb{R}^m,\ \Sigma_{\xi\xi} \in \mathbb{R}^{m\times m}, \tag{29}$$

$A \in \mathbb{R}^{n\times m}$ sei eine Matrix, $b \in \mathbb{R}^n$ sei ein Vektor und υ sei ein n-dimensionaler bedingt normalverteilter Zufallsvektor mit bedingter WDF

$$p_{\upsilon|\xi}(\cdot|x) : \mathbb{R}^n \to \mathbb{R}_{>0},\, y \mapsto p_{\upsilon|\xi}(y|x) := N(y; A\xi + b, \Sigma_{\upsilon\upsilon}) \quad \text{mit} \quad \Sigma_{\upsilon\upsilon} \in \mathbb{R}^{n\times n}. \tag{30}$$

Dann ist der $m+n$-dimensionale Zufallsvektor $\begin{pmatrix}\xi \\ \upsilon\end{pmatrix}$ normalverteilt mit (gemeinsamer) WDF

$$p_{\xi,\upsilon} : \mathbb{R}^{m+n} \to \mathbb{R}_{>0},\, \begin{pmatrix}x \\ y\end{pmatrix} \mapsto p_{\xi,\upsilon}\left(\begin{pmatrix}x \\ y\end{pmatrix}\right) = N\left(\begin{pmatrix}x \\ y\end{pmatrix}; \mu_{\xi,\upsilon}, \Sigma_{\xi,\upsilon}\right), \tag{31}$$

mit $\mu_{\xi,\upsilon} \in \mathbb{R}^{m+n}$ and $\Sigma_{\xi,\upsilon} \in \mathbb{R}^{(m+n)\times(m+n)}$ und insbesondere

$$\mu_{\xi,\upsilon} = \begin{pmatrix}\mu_\xi \\ A\mu_\xi + b\end{pmatrix} \quad \text{und} \quad \Sigma_{\xi,\upsilon} = \begin{pmatrix}\Sigma_{\xi\xi} & \Sigma_{\xi\xi}A^{\mathrm{T}} \\ A\Sigma_{\xi\xi} & \Sigma_{\upsilon\upsilon} + A\Sigma_{\xi\xi}A^{\mathrm{T}}\end{pmatrix}. \tag{32}$$

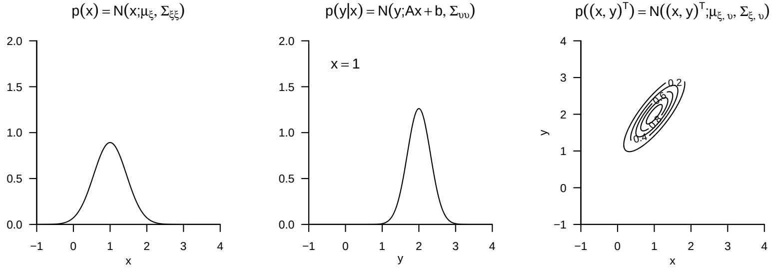
Bemerkungen

- Eine marginale und eine bedingte multivariate Normalverteilung induzieren eine gemeinsame Normalverteilung.
- Die Parameter der gemeinsamen Verteilungen ergeben sich als linear-affine Transformation der Parameter der induzierenden Verteilungen.

Marginale und bedingte Verteilungen

Gemeinsame Normalverteilungen

$$m := 1,\ n := 1,\ \mu_\xi := 1,\ \Sigma_{\xi\xi} := 0.2,\ A := 1,\ b := 1,\ \Sigma_{\upsilon\upsilon} := 0.1$$

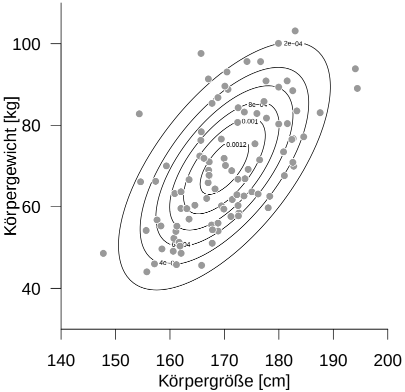


Marginale und bedingte Verteilungen

Beispiel: Gemeinsame Normalverteilung von Körpergröße und Körpergewicht ($n = 100$)

$$x_i \sim N(170, 10^2) \quad \text{und} \quad y_i \sim N(x_i - 100, 12^2) \quad \text{für} \quad i = 1, ..., n \qquad (33)$$

$$\Rightarrow \quad \begin{pmatrix} x_i \\ y_i \end{pmatrix} \sim N\left(\begin{pmatrix} 170 \\ 70 \end{pmatrix}, \begin{pmatrix} 100 & 100 \\ 100 & 244 \end{pmatrix}\right) \quad \text{für} \quad i = 1, ..., n \qquad (34)$$



Selbstkontrollfragen

1. Geben Sie die Definitionen des Erwartungswerts und der Kovarianzmatrix eines Zufallsvektors wieder.
2. Was repräsentieren die Diagonalelemente der Kovarianzmatrix eines Zufallsvektors?
3. Was repräsentieren die Nichtdiagonalelemente der Kovarianzmatrix eines Zufallsvektors?
4. Geben Sie die Definition eines multivariat normalverteilten Zufallsvektors wieder.
5. Erläutern Sie die Definition eines multivariat normalverteilten Zufallsvektors.
6. Welche Werte haben der Erwartungswert und die Kovarianzmatrix eines normalverteilten Zufallsvektors?
7. Visualisieren Sie die WDF eines 2-dimensionalen normalverteilten Zufallsvektors mit den Parameterwerten

$$\mu := \begin{pmatrix} 10 \\ 15 \end{pmatrix} \text{ und } \Sigma := \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}. \tag{35}$$

8. Generieren Sie 100 Realisierungen aus dieser Verteilung und visualisieren Sie diese.
9. Geben Sie das Theorem zur linear-affinen Transformation multivariater Normalverteilungen wieder.
10. Geben Sie das Theorem zur invertierbaren linearen Transformation multivariater Normalverteilungen wieder.
11. Geben Sie das Theorem zu sphärischen Normalverteilungen wieder.
12. Erläutern Sie den Begriff des sphärischen Kovarianzmatrixparameters.
13. Skizzieren Sie den Beweis des Theorems zu sphärischen Normalverteilungen.

Anhang

Beweise einiger in dieser Vorlesung nicht bewiesener Theoreme (englisch):

- Erwartungswert eines multivariat normalverteilten Zufallsvektors
- Kovarianzmatrix eines multivariat normalverteilten Zufallsvektors
- Linear-affine Transformation eines multivariat normalverteilten Zufallsvektors
- Invertierbare lin. Transformation eines multivariat normalverteilten Zufallsvektors
- Äquivalenz unabhängiger univariater mit sphärischer multivariater Normalverteilung
- Marginale Verteilung einer multivariaten Normalverteilung
- Bedingte Verteilung einer multivariaten Normalverteilung

Referenzen

Anderson, T. W. 2003. *An Introduction to Multivariate Statistical Analysis*. 3rd ed. Wiley Series in Probability and Statistics. Hoboken, N.J: Wiley-Interscience.

DeGroot, Morris H., and Mark J. Schervish. 2012. *Probability and Statistics*. 4th ed. Boston: Addison-Wesley.

Mardia, K. V., J. T. Kent, and J. M. Bibby. 1979. *Multivariate Analysis*. Probability and Mathematical Statistics. London ; New York: Academic Press.