Multivariate Verfahren

MSc Psychologie & MSc Klinische Psychologie und Psychotherapie
Wintersemester 2024/2025

Joram Soch

(4) Multivariate Normalverteilungen

Überblick

Die multivariate Normalverteilung ist zentraler Bestandteil vieler Modelle.

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 Quantifikation 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* und Einheit (4) in *Allgemeines Lineares Modell*).

Kontinuierliche Zufallsvektoren

Erwartungswerte und Kovarianzen

Multivariate Normalverteilungen

Marginale und bedingte Verteilungen

Selbstkontrollfragen

Kontinuierliche Zufallsvektoren

Erwartungswerte und Kovarianzen

Multivariate Normalverteilungen

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

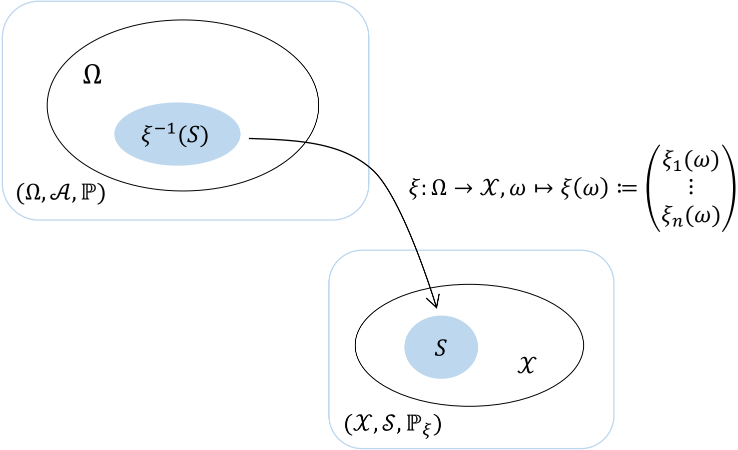
$$\xi : \Omega \to \mathcal{X}, \omega \mapsto \xi(\omega) := \begin{pmatrix} \xi_1(\omega) \\ \vdots \\ \xi_n(\omega) \end{pmatrix} \tag{1}$$

mit der *Messbarkeitseigenschaft*

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

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.



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

Definition (Diskreter Zufallsvektor, Multivariate WMF)

ξ sei ein Zufallsvektor mit Ergebnisraum $\mathcal{X}$. ξ heißt *diskreter Zufallsvektor* wenn der Ergebnisraum $\mathcal{X}$ endlich oder abzählbar ist und eine Funktion

$$p_\xi : \mathcal{X} \to [0,1], x \mapsto p_\xi(x) \tag{3}$$

existiert, für die gilt

(1) $\sum_{x \in \mathcal{X}} p_\xi(x) = 1$ und

(2) $\mathbb{P}_\xi(\xi = x) = p_\xi(x)$ für alle $x \in \mathcal{X}$.

Ein entsprechende Funktion p heißt *multivariate Wahrscheinlichkeitsmassefunktion (WMF)* von ξ.

Bemerkungen

- Der Begriff der multivariaten WMF ist analog zum Begriff der WMF.
- Man spricht häufig auch einfach von der *WMF eines Zufallsvektors*.
- Wie univariate WMFen sind multivariate WMFen nicht-negativ und normiert.

Definition (Kontinuierlicher Zufallsvektor, Multivariate WDF)

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

$$p_\xi : \mathbb{R}^n \to \mathbb{R}_{\geq 0}, x \mapsto p_\xi(x), \tag{4}$$

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:

$$\mathbb{P}_\xi(\xi = x) = \mathbb{P}_\xi(x \leq \xi \leq x) = \int_{x_1}^{x_1} \cdots \int_{x_n}^{x_n} p_\xi(s_1, ..., s_n)\, ds_1 \cdots ds_n = 0\ . \tag{5}$$

Kontinuierliche Zufallsvektoren

Erwartungswerte und Kovarianzen

Multivariate Normalverteilungen

Marginale und bedingte Verteilungen

Selbstkontrollfragen

Definition (Erwartungswert eines Zufallsvektors)

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

$$\mathbb{E}(\xi) := \begin{pmatrix} \mathbb{E}(\xi_1) \\ \vdots \\ \mathbb{E}(\xi_m) \end{pmatrix}. \tag{6}$$

Bemerkung

- Der Erwartungswert eines Zufallsvektors ist der Vektor der Erwartungswerte seiner Komponenten.

Definition (Erwartungswert einer Zufallsmatrix)

Ξ sei eine durch die spaltenweise Konkatenation der m-dimensionalen Zufallsvektoren $\xi_1, ..., \xi_n$ gebildete Zufallsmatrix. Dann ist der *Erwartungwert von* Ξ ist definiert als die $m \times n$ Matrix

$$\mathbb{E}(\Xi) := \begin{pmatrix} \mathbb{E}\left(\xi_{1_1}\right) & \cdots & \mathbb{E}\left(\xi_{n_1}\right) \\ \vdots & \ddots & \vdots \\ \mathbb{E}\left(\xi_{1_m}\right) & \cdots & \mathbb{E}\left(\xi_{n_m}\right) \end{pmatrix}. \tag{7}$$

Bemerkung

- Der Erwartungswert einer Zufallsmatrix ist die Matrix der spaltenweise konkatenierten Erwartungswerte $\mathbb{E}\left(\xi_1\right), ..., \mathbb{E}\left(\xi_n\right)$.

Theorem (Eigenschaften von Erwartungswerten)

(1) (Linear-affine Transformation einer Zufallsmatrix) Ξ sei ein $m \times n$-dimensionale Zufallsmatrix und es seien $A \in \mathbb{R}^{l \times m}, B \in \mathbb{R}^{n \times p}$ und $C \in \mathbb{R}^{l \times p}$. Dann gilt

$$\mathbb{E}(A\Xi B + C) = A\mathbb{E}(\Xi)B + C. \tag{8}$$

(2) (Linear-affine Transformation eines Zufallsvektors) ξ sei ein m-dimensionaler Zufallvektor und es seien $A \in \mathbb{R}^{n \times m}$ und $b \in \mathbb{R}^n$. Dann gilt

$$\mathbb{E}(A\xi + b) = A\mathbb{E}(\xi) + b. \tag{9}$$

(3) (Lineare Kombination zweier Zufallsvektoren) ξ und υ seien m-dimensionale Zufallvektoren und es seien $A, B \in \mathbb{R}^{n \times m}$. Dann gilt

$$\mathbb{E}\left(A\xi + B\upsilon\right) = A\mathbb{E}(\xi) + B\mathbb{E}(\upsilon). \tag{10}$$

Bemerkungen

- Die Aussagen sind im Wesentlichen analog zu den Eigenschaften des Erwartungswerts bei Zufallsvariablen.
- Eigenschaft (2) ist ein Spezialfall von Eigenschaft (1).

Beweis

(1) Wir halten zunächst fest, dass für $i = 1, ..., l$ und $j = 1, ...p$ mit den Regeln von Matrixmultiplikation und Matrixaddition und der Linearität des Erwartungswertes von Zufallsvariablen gilt, dass

$$\mathbb{E}(A\Xi B + C)_{ij} = \mathbb{E}\left(\sum_{r=1}^{m}\sum_{s=1}^{n} a_{ir}\xi_{s_r}b_{sj} + c_{ij}\right) = \sum_{r=1}^{m}\sum_{s=1}^{n} a_{ir}\mathbb{E}(\xi_{s_r})b_{sj} + c_{ij}. \tag{11}$$

Dann aber gilt mit der Definition des Erwartungswerts einer Zufallsmatrix, dass

$$\begin{aligned}
\mathbb{E}(A\Xi B + C) &= (\mathbb{E}(A\Xi B + C)_{ij})_{1\le i\le l, 1\le j\le p} \\
&= \left(\sum_{r=1}^{m}\sum_{s=1}^{n} a_{ir}\mathbb{E}(\xi_{s_r})b_{sj} + c_{ij}\right)_{1\le i\le l, 1\le j\le p} \\
&= \left(\sum_{r=1}^{m}\sum_{s=1}^{n} a_{ir}\mathbb{E}(\xi_{s_r})b_{sj}\right)_{1\le i\le l, 1\le j\le p} + \left(c_{ij}\right)_{1\le i\le l, 1\le j\le p} \\
&= A\mathbb{E}(\Xi)B + C.
\end{aligned} \tag{12}$$

Beweis (fortgeführt)

(2) Für $i = 1, ..., n$ gilt mit der Definition des Erwartungswert eines Zufallsvektors und der Linearität des Erwartungswert einer Zufallsvariable, dass

$$\mathbb{E}(A\xi + b)_i = \mathbb{E}\left(\sum_{j=1}^m a_{ij}\xi_j + b_i\right) = \sum_{j=1}^m a_{ij}\mathbb{E}(\xi_j) + b_i. \tag{13}$$

Also gilt

$$\mathbb{E}(A\xi + b) = \begin{pmatrix} \sum_{j=1}^m a_{1j}\mathbb{E}(\xi_j) + b_1 \\ \vdots \\ \sum_{j=1}^m a_{nj}\mathbb{E}(\xi_j) + b_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m a_{1j}\mathbb{E}(\xi_j) \\ \vdots \\ \sum_{j=1}^m a_{nj}\mathbb{E}(\xi_j) \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = A\mathbb{E}(\xi) + b. \tag{14}$$

(3) Für $i = 1, ..., n$ gilt mit der Definition des Erwartungswert eines Zufallsvektors und der Linearität des Erwartungswert einer Zufallsvariable, dass

$$\mathbb{E}\left(A\xi + B\upsilon\right)_i = \mathbb{E}\left(\sum_{j=1}^m a_{ij}\xi_j + \sum_{j=1}^m b_{ij}\upsilon_j\right) = \sum_{j=1}^m a_{ij}\mathbb{E}\left(\xi_j\right) + \sum_{j=1}^m b_{ij}\mathbb{E}\left(\upsilon_j\right). \tag{15}$$

Also gilt

$$\mathbb{E}\left(A\xi + B\upsilon\right) = \begin{pmatrix} \sum_{j=1}^m a_{1j}\mathbb{E}\left(\xi_j\right) + \sum_{j=1}^m b_{1j}\mathbb{E}\left(\upsilon_j\right) \\ \vdots \\ \sum_{j=1}^m a_{nj}\mathbb{E}\left(\xi_j\right) + \sum_{j=1}^m b_{nj}\mathbb{E}\left(\upsilon_j\right) \end{pmatrix} = A\mathbb{E}(\xi) + B\mathbb{E}(\upsilon). \tag{16}$$

$\square$

Definition (Kovarianzmatrix eines Zufallsvektors)

ξ sei ein m-dimensionaler Zufallvektor. Dann ist die *Kovarianzmatrix von* ξ definiert als die $m \times m$ Matrix

$$\mathbb{C}(\xi) := \mathbb{E}\left((\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^{\mathrm{T}}\right). \tag{17}$$

Bemerkungen

- Die Kovarianzmatrix ist formal analog zur Kovarianz zweier Zufallsvariablen definiert.
- Der äußere Erwartungswert ist der Erwartungswert einer Matrix, die inneren diejenigen von Vektoren.

Theorem (Eigenschaften der Kovarianzmatrix)

ξ sei ein m-dimensionaler Zufallsvektor und $\mathbb{C}(\xi)$ sei seine Kovarianzmatrix. Dann gelten

(1) (Elemente) Die Elemente von $\mathbb{C}(\xi)$ sind die Kovarianzen der Komponenten von ξ,

$$\mathbb{C}(\xi) = \left(\mathbb{C}(\xi_i, \xi_j)\right)_{1 \le i,j \le m}. \tag{18}$$

(2) (Kovarianzmatrixverschiebungssatz) Es gilt

$$\mathbb{C}(\xi) = \mathbb{E}\left(\xi\xi^{\mathrm{T}}\right) - \mathbb{E}(\xi)\mathbb{E}(\xi)^{\mathrm{T}}. \tag{19}$$

(3) (Linear-affine Transformation) Für $A \in \mathbb{R}^{n \times m}$ und $b \in \mathbb{R}^n$ gilt

$$\mathbb{C}(A\xi + b) = A\mathbb{C}(\xi)A^{\mathrm{T}}. \tag{20}$$

(4) (Matrixeigenschaften) $\mathbb{C}(\xi)$ ist symmetrisch und positiv-semidefinit.

Bemerkungen

- Die Diagonalelemente von $\mathbb{C}(\xi)$ sind die Varianzen der Komponenten von ξ, da

$$\mathbb{V}(\xi_i) = \mathbb{C}(\xi_i, \xi_i) \text{ für } i = 1, ..., m. \tag{21}$$

- Eigenschaften (2) und (3) sind im Wesentlichen analog zu den Eigenschaften der Varianz.

Beweis

(1) Es gilt

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

Erwartungswerte und Kovarianzen

Beweis (fortgeführt)

(2) Mit den Eigenschaften von Erwartungswerten gilt

$$
\begin{aligned}
\mathbb{C}(\xi) &= \mathbb{E}\left((\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^{\mathrm{T}}\right) \\
&= \mathbb{E}\left(\xi\xi^{\mathrm{T}} - \xi\mathbb{E}(\xi)^{\mathrm{T}} - \mathbb{E}(\xi)\xi^{\mathrm{T}} + \mathbb{E}(\xi)\mathbb{E}(\xi)^{\mathrm{T}}\right) \\
&= \mathbb{E}\left(\xi\xi^{\mathrm{T}}\right) - \mathbb{E}(\xi)\mathbb{E}(\xi)^{\mathrm{T}} - \mathbb{E}(\xi)\mathbb{E}(\xi)^{\mathrm{T}} + \mathbb{E}(\xi)\mathbb{E}(\xi)^{\mathrm{T}} \\
&= \mathbb{E}\left(\xi\xi^{\mathrm{T}}\right) - \mathbb{E}(\xi)\mathbb{E}(\xi)^{\mathrm{T}}.
\end{aligned} \tag{22}
$$

(3) Mit den Eigenschaften von Erwartungswerten gilt

$$
\begin{aligned}
\mathbb{C}(A\xi + b) &= \mathbb{E}\left((A\xi + b - \mathbb{E}(A\xi + b))(A\xi + b - \mathbb{E}(A\xi + b))^{\mathrm{T}}\right) \\
&= \mathbb{E}\left((A\xi + b - A\mathbb{E}(\xi) - b)(A\xi + b - A\mathbb{E}(\xi) - b)^{\mathrm{T}}\right) \\
&= \mathbb{E}\left((A(\xi - \mathbb{E}(\xi)))(A(\xi - \mathbb{E}(\xi)))^{\mathrm{T}}\right) \\
&= \mathbb{E}\left(A(\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^{\mathrm{T}}A^{\mathrm{T}}\right) \\
&= A\mathbb{E}\left((\xi - \mathbb{E}(\xi))(\xi - \mathbb{E}(\xi))^{\mathrm{T}}\right)A^{\mathrm{T}} \\
&= A\mathbb{C}(\xi)A^{\mathrm{T}}.
\end{aligned} \tag{23}
$$

Beweis (fortgeführt)

(4) Die Symmetrie von $\mathbb{C}(\xi)$ folgt aus der Symmetrie der Kovarianz einer Zufallsvariable mit

$$\mathbb{C}(\xi_i, \xi_j) = \mathbb{C}(\xi_j, \xi_i) \text{ für alle } i, j = 1, ..., m. \tag{24}$$

Um die positive Semidefinitheit von $\mathbb{C}(\xi)$ nachzuweisen, ist zu zeigen, dass $a^{\mathrm{T}}\mathbb{C}(\xi)a \geq 0$ für alle $a \in \mathbb{R}^m$ mit $a \neq 0_m$. Sei also $a \in \mathbb{R}^m$ mit $a \neq 0$. Dann gilt mit Aussage (3) für $A := a^{\mathrm{T}} \in \mathbb{R}^{1\times m}$, dass

$$a^{\mathrm{T}}\mathbb{C}(\xi)a = \mathbb{C}(a^{\mathrm{T}}\xi). \tag{25}$$

Weiterhin gilt mit der Definition der Kovarianzmatrix aber, dass

$$\mathbb{C}(a^{\mathrm{T}}\xi) = \mathbb{E}\left(\left(a^{\mathrm{T}}\xi - \mathbb{E}(a^{\mathrm{T}}\xi)\right)^2\right) = \mathbb{V}\left(a^{\mathrm{T}}\xi\right). \tag{26}$$

Da die Varianz der Zufallsvariable $a^{\mathrm{T}}\xi$ aber nicht-negativ ist, folgt

$$a^{\mathrm{T}}\mathbb{C}(\xi)a = \mathbb{V}(a^{\mathrm{T}}\xi) \geq 0 \tag{27}$$

und damit die Positiv-Semidefinitheit von $\mathbb{C}(\xi)$. $\square$

Definition (Korrelationsmatrix eines Zufallsvektors)

ξ sei ein m-dimensionaler Zufallsvektor. Dann ist die *Korrelationsmatrix* von ξ definiert als die $m \times m$ Matrix

$$\mathbb{R}(\xi) := \left(\rho_{ij}\right)_{1 \le i,j \le m} = \left(\frac{\mathbb{C}(\xi_i, \xi_j)}{\sqrt{\mathbb{V}(\xi_i)}\sqrt{\mathbb{V}(\xi_j)}}\right)_{1 \le i,j \le m} \tag{28}$$

Bemerkungen

- Mit $\mathbb{V}(\xi_i) = \mathbb{C}(\xi_i, \xi_i), i = 1, ..., m$ ergeben sich die Diagonaleinträge der Korrelationsmatrix zu 1.
- Es gelten $\rho_{ij} \in [-1, 1]$ für $1 \le i, j \le m$ und $\rho_{ii} = 1$ für $1 \le i \le m$.

Kontinuierliche Zufallsvektoren

Erwartungswerte und Kovarianzen

Multivariate Normalverteilungen

Marginale und bedingte Verteilungen

Selbstkontrollfragen

Definition (Normalverteilte Zufallsvariable)

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

$$p : \mathbb{R} \to \mathbb{R}_{>0}, x \mapsto p(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right). \tag{29}$$

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

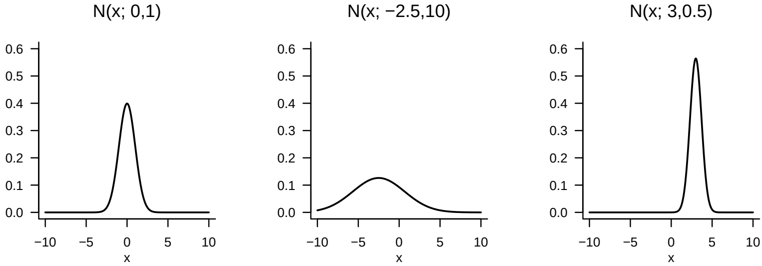
$$N\left(x; \mu, \sigma^2\right) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right). \tag{30}$$

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*.

Multivariate Normalverteilungen

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

$$\begin{aligned} \xi_1 &:= \sigma_1\zeta_1 + \mu_1 \\ \xi_2 &:= \sigma_2\left(\rho\zeta_1 + \sqrt{1-\rho^2}\zeta_2\right) + \mu_2. \end{aligned} \tag{31}$$

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

$$p : \mathbb{R}^2 \to \mathbb{R}_{>0},\, x \mapsto p(x) := \frac{1}{\sqrt{(2\pi)^n|\Sigma|}}\exp\left(-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)\right), \tag{32}$$

wobei

$$n := 2, \mu := \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \text{ und } \Sigma := \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix} \tag{33}$$

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*.

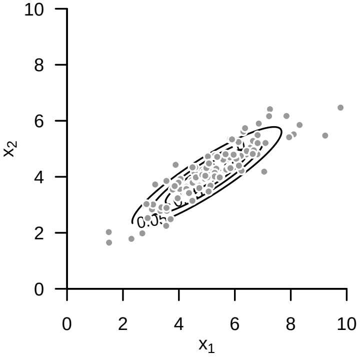
Multivariate Normalverteilungen

Konstruktion bivariater Normalverteilungen

$$\mu_1 = 5,\ \mu_2 = 4,\ \sigma_1 = 1.5,\ \sigma_2 = 1.0,\ \rho = 0.9.$$

• $n = 100$ Realisierungen von $\xi = (\xi_1, \xi_2)^{\mathrm{T}}$

— Wahrscheinlichkeitsdichtefunktion von $\xi \sim N(\mu, \Sigma)$



Definition (Multivariate Normalverteilung)

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

$$p : \mathbb{R}^n \to \mathbb{R}_{>0},\, x \mapsto p(x) := \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)\right). \tag{34}$$

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

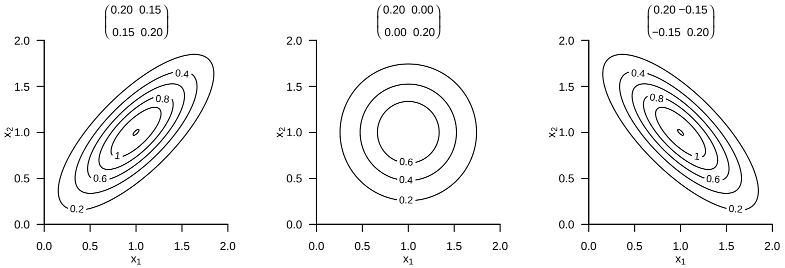
$$N(x; \mu, \Sigma) := \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)\right). \tag{35}$$

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.

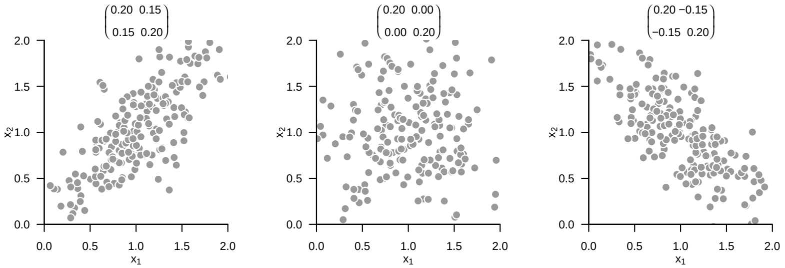
Visualisierung bivariater Normalverteilungsdichtefunktionen

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



Realisierung bivariater normalverteilter Zufallsvektoren

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



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

$$\mathbb{E}(\xi) = \mu \quad \text{und} \quad \mathbb{C}(\xi) = \Sigma\,. \tag{36}$$

Bemerkungen

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

Theorem (Linear-affine Transformation)

$\xi \sim N(\mu, \Sigma)$ sei ein multivariat 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

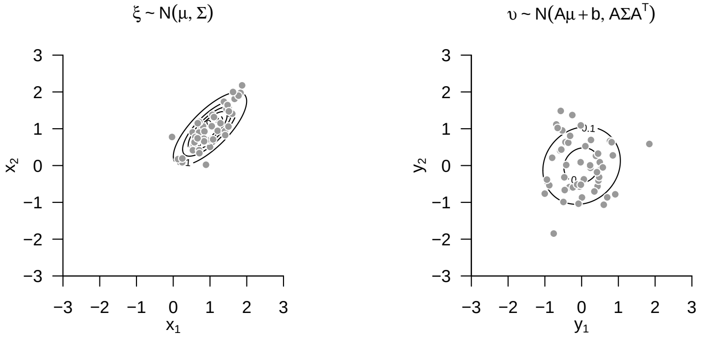
$$\upsilon \sim N\left(A\mu + b, A\Sigma A^{\mathrm{T}}\right) \tag{37}$$

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

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

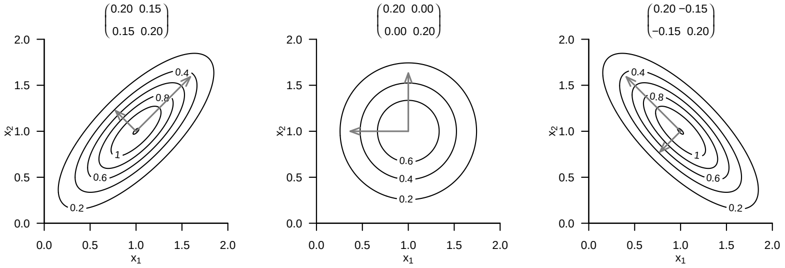


Multivariate Normalverteilungen und Eigenanalyse

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

— Wahrscheinlichkeitsdichtefunktion von $\xi \sim N(\mu, \Sigma)$

$\rightarrow \sqrt{\lambda_i}v_i$, Eigenwert-skalierte Eigenvektoren von Σ



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. Sie spielen aber auch im Rahmen der Schätzung kanonischer Korrelationen eine Rolle (siehe Einheit (7) Kanonische Korrelationsanalyse).

Theorem (Marginale Normalverteilungen)

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

$$\mu = \begin{pmatrix} \mu_\upsilon \\ \mu_\zeta \end{pmatrix} \in \mathbb{R}^m, \tag{38}$$

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

$$\Sigma = \begin{pmatrix} \Sigma_{\upsilon\upsilon} & \Sigma_{\upsilon\zeta} \\ \Sigma_{\zeta\upsilon} & \Sigma_{\zeta\zeta} \end{pmatrix} \in \mathbb{R}^{m \times m}, \tag{39}$$

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:

$$\upsilon \sim N(\mu_\upsilon, \Sigma_{\upsilon\upsilon}) \text{ and } \zeta \sim N(\mu_\zeta, \Sigma_{\zeta\zeta}) \,. \tag{40}$$

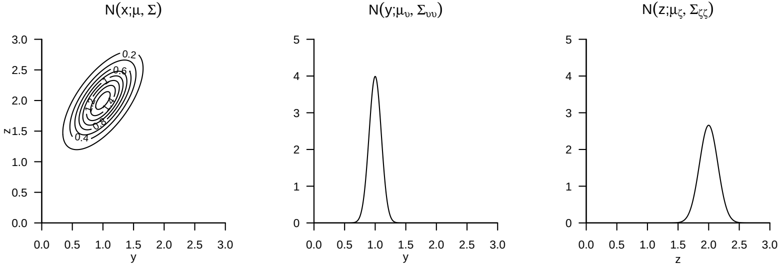
Bemerkungen

- 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 und bedingte Verteilungen

Marginale Normalverteilungen

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



Marginale und bedingte Verteilungen

Theorem (Bedingte Normalverteilungen)

$(\xi_1, ..., \xi_m, \upsilon_1, ..., \upsilon_n)^{\mathrm{T}}$ sei ein $m+n$-dimensionaler normalverteilter Zufallsvektor mit WDF

$$p_{\xi,\upsilon} : \mathbb{R}^{m+n} \to \mathbb{R}_{>0}, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto p_{\xi,\upsilon}\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) := N\left(\begin{pmatrix} x \\ y \end{pmatrix}; \mu_{\xi,\upsilon}, \Sigma_{\xi,\upsilon}\right) \tag{41}$$

mit

$$\mu_{\xi,\upsilon} = \begin{pmatrix} \mu_\xi \\ \mu_\upsilon \end{pmatrix} \text{ und } \Sigma_{\xi,\upsilon} = \begin{pmatrix} \Sigma_{\xi\xi} & \Sigma_{\xi\upsilon} \\ \Sigma_{\upsilon\xi} & \Sigma_{\upsilon\upsilon} \end{pmatrix}, \tag{42}$$

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

$$p_{\xi|\upsilon}(\cdot|y) : \mathbb{R}^m \to \mathbb{R}_{>0}, x \mapsto p_{\xi|\upsilon}(x|y) := N(x; \mu_{\xi|\upsilon}, \Sigma_{\xi|\upsilon}) \tag{43}$$

mit

$$\mu_{\xi|\upsilon} = \mu_\xi + \Sigma_{\xi\upsilon}\Sigma_{\upsilon\upsilon}^{-1}(y - \mu_\upsilon) \in \mathbb{R}^m \tag{44}$$

und

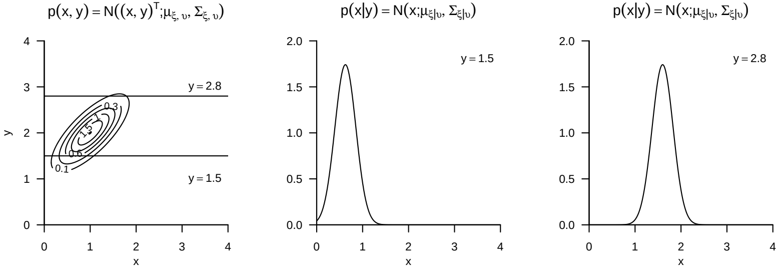
$$\Sigma_{\xi|\upsilon} = \Sigma_{\xi\xi} - \Sigma_{\xi\upsilon}\Sigma_{\upsilon\upsilon}^{-1}\Sigma_{\upsilon\xi} \in \mathbb{R}^{m\times m}. \tag{45}$$

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

$$p_\xi : \mathbb{R}^m \to \mathbb{R}_{>0},\, x \mapsto p_\xi(x) := N(x; \mu_\xi, \Sigma_{\xi\xi}) \text{ mit } \mu_\xi \in \mathbb{R}^m, \Sigma_{\xi\xi} \in \mathbb{R}^{m\times m}, \tag{46}$$

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

$$p_{\upsilon|\xi}(\cdot|x) : \mathbb{R}^n \to \mathbb{R}_{>0},\, y \mapsto p_{\upsilon|\xi}(y|x) := N(y; A\xi + b, \Sigma_{\upsilon\upsilon}) \text{ mit } \Sigma_{\upsilon\upsilon} \in \mathbb{R}^{n\times n}. \tag{47}$$

Dann ist der $m+n$-dimensionale Zufallsvektor $(\xi, \upsilon)^{\mathrm{T}}$ normalverteilt mit (gemeinsamer) WDF

$$p_{\xi,\upsilon} : \mathbb{R}^{m+n} \to \mathbb{R}_{>0},\, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto p_{\xi,\upsilon}\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = N\left(\begin{pmatrix} x \\ y \end{pmatrix}; \mu_{\xi,\upsilon}, \Sigma_{\xi,\upsilon}\right), \tag{48}$$

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

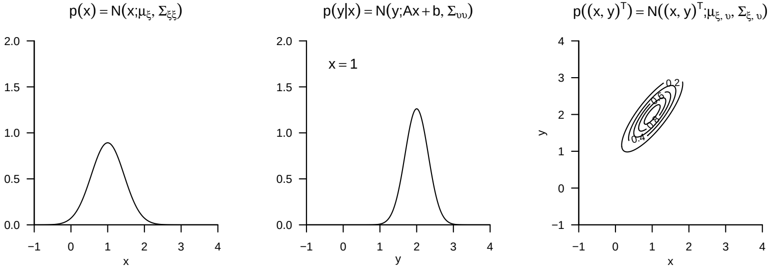
$$\mu_{\xi,\upsilon} = \begin{pmatrix} \mu_\xi \\ A\mu_\xi + b \end{pmatrix} \text{ und } \Sigma_{\xi,\upsilon} = \begin{pmatrix} \Sigma_{\xi\xi} & \Sigma_{\xi\xi}A^{\mathrm{T}} \\ A\Sigma_{\xi\xi} & \Sigma_{\upsilon\upsilon} + A\Sigma_{\xi\xi}A^{\mathrm{T}} \end{pmatrix}. \tag{49}$$

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.

Gemeinsame Normalverteilungen

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



Kontinuierliche Zufallsvektoren

Erwartungswerte und Kovarianzen

Multivariate Normalverteilungen

Marginale und bedingte Verteilungen

Selbstkontrollfragen

Selbstkontrollfragen

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

$$\mu := \begin{pmatrix} 10 \\ 15 \end{pmatrix} \text{ und } \Sigma := \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}. \tag{50}$$

12. Generieren Sie 100 Realisierungen aus dieser Verteilung und visualisieren Sie diese.
13. Geben Sie das Theorem zur linear-affinen Transformation multivariater Normalverteilungen wieder.

Anhang

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

- Linearität des Erwartungswerts einer Zufallsvariable
- Zusammenhang zwischen Kovarianzmatrix und Erwartungswerten
- Symmetrie und Positiv-Semidefinitheit der Kovarianzmatrix
- Erwartungswert eines multivariat normalverteilten Zufallsvektors
- Kovarianzmatrix eines multivariat normalverteilten Zufallsvektors
- Linear-affine Transformation eines multivariat normalverteilten Zufallsvektors
- 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.