Statistische Analyseverfahren
Abschnitt 2: Zufallsvektoren und mehrdimensionale Verteilungen

Dr. Andreas Wünsche

TU Bergakademie Freiberg
Institut für Stochastik

Oktober 2019

2.1 Zufallsvektoren und mehrdimensionale Verteilungen

- Eine endliche Familie $X_1, \ldots, X_n$ von Zufallsgrößen kann als ein Zufallsvektor (ZV)
$$\underline{\mathbf{X}} = (X_1, \ldots, X_n)^{\mathsf{T}}$$
angesehen werden.

- Eine doppelt indizierte Familie $X_{ij}\,, i = 1, \ldots, n\,; j = 1, \ldots, m$, von Zufallsgrößen kann als eine Zufallsmatrix (ZM)
$$\underline{\underline{\mathbf{X}}} = (X_{ij})_{i=1,\ldots,n\,;j=1,\ldots,m} = (X_{ij})$$
angesehen werden.

- Zufallsvariable werden hier im Allgemeinen mit Großbuchstaben bezeichnet, z.B. $\underline{\mathbf{X}}$, deren Realisierungen durch Angabe von ω oder durch Kleinbuchstaben, z.B. $\underline{\mathbf{X}}(\omega) = \underline{\mathbf{x}}$.

- Vektoren werden hier durch einen Unterstrich, Matrizen durch zwei Unterstriche gekennzeichnet.

Bezeichnungen aus der linearen Algebra

- Verwendet werden übliche Bezeichnungen aus der linearen Algebra.
- Vektoren werden als Spaltenvektoren angesehen, $\cdot^{\mathsf{T}}$ bezeichnet die Transponierung eines Vektors oder einer Matrix.
- $\underline{\mathbf{0}}_n\,,\ \underline{\underline{\mathbf{0}}}_{n\times m}$ ... $n-$dim. Nullvektor, Nullmatrix vom Typ $n \times m$.
- $\underline{\mathbf{1}}_n\,,\ \underline{\underline{\mathbf{1}}}_{n\times m}$ ... $n-$dim. Vektor bzw. $n \times m-$Matrix aus Einsen.
- Analog $\underline{\infty}_n\,,\ \underline{-\infty}_n = -\underline{\infty}_n$.
- $\mathbb{M}_{n\times m}\,,\ \mathbb{M}_n$ ... Menge aller $n \times m-$ bzw. $n \times n-$Matrizen.
- $\mathbb{M}_n^{\geq},\ \mathbb{M}_n^{>}$ ... Menge aller positiv semidefiniten bzw. positiv definiten $n \times n-$Matrizen,

  $\forall\ \underline{\mathbf{0}} \neq \underline{\mathbf{x}} \in \mathbb{R}^n : \underline{\mathbf{x}}^{\mathsf{T}} \underline{\underline{\mathbf{a}}}\, \underline{\mathbf{x}} \geq 0\ (\underline{\underline{\mathbf{a}}} \in \mathbb{M}_n^{\geq})$ bzw. $\underline{\mathbf{x}}^{\mathsf{T}} \underline{\underline{\mathbf{a}}}\, \underline{\mathbf{x}} > 0\ (\underline{\underline{\mathbf{a}}} \in \mathbb{M}_n^{>})$.
- $\underline{\underline{\mathbb{I}}}_n$ ... $n \times n-$Einheitsmatrix.
- $\|\cdot\|$ bezeichnet die euklidische Norm.

Verteilung eines Zufallsvektors

- Ein Zufallsvektor $\underline{\mathbf{X}}$ wird im Allgemeinen durch seine Verteilung gegeben.
- Die Verteilung $\mathbb{P}_{\underline{\mathbf{X}}}$ eines n-dimensionalen Zufallsvektors $\underline{\mathbf{X}}$ definiert die Wahrscheinlichkeiten dafür, dass die Realisierungen des Zufallsvektors in geeigneten Teilmengen des $\mathbb{R}^n$ liegen,

$$\mathbb{P}_{\underline{\mathbf{X}}}(B) = \mathbb{P}(\underline{\mathbf{X}} \in B), \qquad B \subseteq \mathbb{R}^n, \text{ geeignet.}$$

- Analog kann die Verteilung einer $n \times m-$Zufallsmatrix für geeignete Teilmengen des Raumes $\mathbb{R}^{n \cdot m}$ genutzt werden. Formal kann dies durch die Vektorisierung der Matrix realisiert werden (Hintereinanderschreiben der Spalten der Matrix zu einem langen Spaltenvektor).

Definition Verteilungsfunktion $n-$dim. Zufallsvektor

- **Bez.**
  $\underline{\mathbf{x}} = (x_1, \dots, x_n)^\mathsf{T} < \underline{\mathbf{y}} = (y_1, \dots, y_n)^\mathsf{T}$, falls $x_i < y_i$ für $i = 1, \dots, n$;
  $\underline{\mathbf{x}} = (x_1, \dots, x_n)^\mathsf{T} \leq \underline{\mathbf{y}} = (y_1, \dots, y_n)^\mathsf{T}$, falls $x_i \leq y_i$ für $i = 1, \dots, n$.

- Für einen $n-$dimensionalen Zufallsvektor $\underline{\mathbf{X}}$ wird die Verteilung eindeutig durch die zugehörige Verteilungsfunktion $F_{\underline{\mathbf{X}}} : \mathbb{R}^n \to [0,1]$ definiert:

$$F_{\underline{\mathbf{X}}}(\underline{\mathbf{x}}) = \mathbb{P}_{\underline{\mathbf{X}}}((-\underline{\infty}_n, \underline{\mathbf{x}})) = \mathbb{P}(\underline{\mathbf{X}} < \underline{\mathbf{x}}), \quad \underline{\mathbf{x}} \in \mathbb{R}^n .$$

- **Bem.**
  Oft wird die Verteilungsfunktion von $\underline{\mathbf{X}}$ auch definiert durch

$$\tilde{F}_{\underline{\mathbf{X}}}(\underline{\mathbf{x}}) = \mathbb{P}_{\underline{\mathbf{X}}}((-\underline{\infty}_n, \underline{\mathbf{x}}]) = \mathbb{P}(\underline{\mathbf{X}} \leq \underline{\mathbf{x}}), \quad \underline{\mathbf{x}} \in \mathbb{R}^n .$$

Eigenschaften Verteilungsfunktion $n-$dim. Zufallsvektor

- **Bez.:** für $i \in \{1, \dots, n\}$, $a < b \in \mathbb{R}$ sei

$$\Delta_{i;a,b} F_{\underline{\mathbf{X}}}(\underline{\mathbf{x}}) :=$$
$$F_{\underline{\mathbf{X}}}(x_1, \dots x_{i-1}, b, x_{i+1}, \dots, x_n) - F_{\underline{\mathbf{X}}}(x_1, \dots x_{i-1}, a, x_{i+1}, \dots, x_n)\,.$$

- Eine Verteilungsfunktion $F_{\underline{\mathbf{X}}}$ besitzt die Eigenschaften
  (i) $\lim\limits_{x_i \to -\infty\,,\, i \in \{1,\dots,n\}} F_{\underline{\mathbf{X}}}(\underline{\mathbf{x}}) = 0\,;\quad \lim\limits_{x_i \to \infty\,,\, i=1,\dots,n} F_{\underline{\mathbf{X}}}(\underline{\mathbf{x}}) = 1\,;$
  (ii) für beliebige $\underline{\mathbf{a}} < \underline{\mathbf{b}}$ gilt

$$\Delta_{1;a_1,b_1} \dots \Delta_{n;a_n,b_n} F_{\underline{\mathbf{X}}}(\underline{\mathbf{x}}) \geq 0\,, \qquad \big(= \mathbb{P}_{\underline{\mathbf{X}}}([\underline{\mathbf{a}}, \underline{\mathbf{b}}))\big)$$

  $\Rightarrow$ $F_{\underline{\mathbf{X}}}$ ist monoton nichtfallend bezüglich jeder Variablen;
  (iii) F_X ist linksseitig stetig bzgl. aller Variablen, d.h. aus $\underline{\mathbf{x}}^{(k)} \uparrow \underline{\mathbf{x}}$ folgt $F_{\underline{\mathbf{X}}}(\underline{\mathbf{x}}^{(k)}) \uparrow F_{\underline{\mathbf{X}}}(\underline{\mathbf{x}})\,,\; k \to \infty\,.$

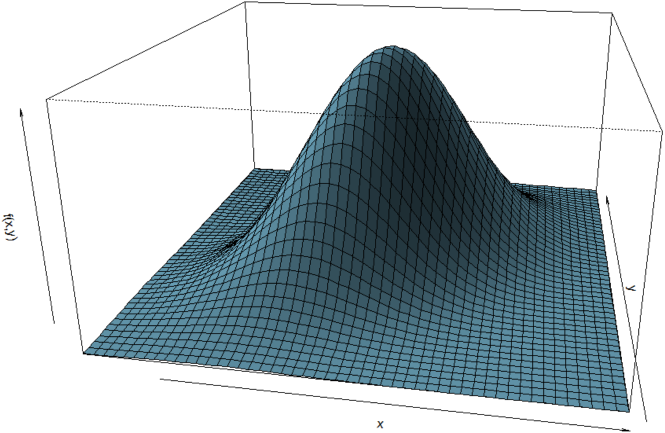
Verteilungsdichte eines $n-$dimensionalen Zufallsvektors

- Für absolut stetige Zufallsvektoren wird die Verteilung auch eindeutig durch die Verteilungsdichte (auch Dichtefunktion) $f_{\underline{\mathbf{X}}} : \mathbb{R}^n \to \mathbb{R}$ gegeben:

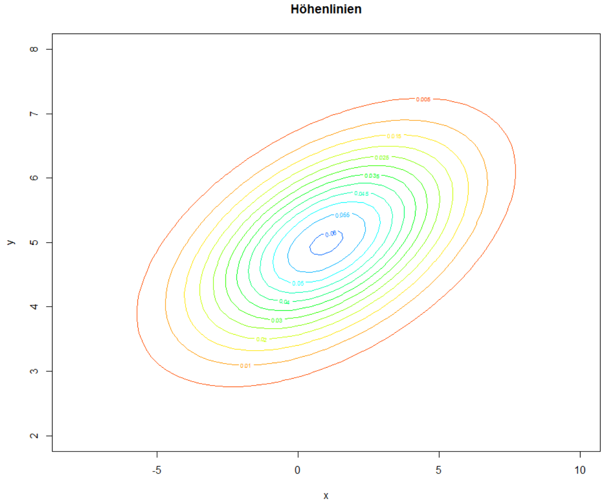
$$F_{\underline{\mathbf{X}}}(\underline{\mathbf{x}}) = \mathbb{P}(\underline{\mathbf{X}} < \underline{\mathbf{x}}) = \int_{-\underline{\infty}_n}^{\underline{\mathbf{x}}} f_{\underline{\mathbf{X}}}(\underline{\mathbf{u}})\,\mathrm{d}\underline{\mathbf{u}}\,, \quad \underline{\mathbf{x}} \in \mathbb{R}^n\,.$$

- Eigenschaften von Dichtefunktionen (etwas vereinfacht):
  - für alle $\underline{\mathbf{x}} \in \mathbb{R}^n : f_{\underline{\mathbf{X}}}(\underline{\mathbf{x}}) \geq 0\,;$
  - $\displaystyle\int_{\mathbb{R}^n} f_{\underline{\mathbf{X}}}(\underline{\mathbf{x}})\,\mathrm{d}\underline{\mathbf{x}} = 1\,.$
- Dann gilt für geeignete $B \subseteq \mathbb{R}^n$ $\mathbb{P}(\underline{\mathbf{X}} \in B) = \displaystyle\int_B f_{\underline{\mathbf{X}}}(\underline{\mathbf{x}})\,\mathrm{d}\underline{\mathbf{x}}\,.$
- Die Dichtefunktion wird im statistischen Kontext, inbesondere bei unbekannten Parametern und einer gegebenen Realisierung (Stichprobe) als Likelihood-Funktion bezeichnet.

Dichte einer zweidimensionalen Normalverteilung



Höhenlinien einer zweidimensionalen Normalverteilungsdichte



Randverteilungen

- Die Verteilungsfunktion eines Teilvektors eines Zufallsvektors (die Randverteilungsfunktion) kann mit Hilfe der Verteilungsfunktion des Zufallsvektors berechnet werden.
- **Bsp.:** $\underline{\mathbf{X}} = (X_1, \dots, X_n)^\mathsf{T} \;\Rightarrow$

$$\forall\, x_1 \in \mathbb{R} \,:\, F_{X_1}(x_1) = \mathbb{P}(X_1 < x_1) = F_{\underline{\mathbf{X}}}(x_1, \infty, \dots, \infty)\,.$$

- Analoges gilt im Fall von absolut stetigen Zufallsvektoren für die Randverteilungsdichten.
- **Bsp.:** $\underline{\mathbf{X}} = (X_1, \dots, X_n)^\mathsf{T} \;\Rightarrow$ (etwas vereinfacht)

$$\forall\, x_1 \in \mathbb{R} \,:\, f_{X_1}(x_1) = \int_{\mathbb{R}^{n-1}} f_{\underline{\mathbf{X}}}(x_1, x_2, \dots, x_n)\,\mathrm{d}x_2 \dots \mathrm{d}x_n\,.$$

- Aus den Randverteilungsfunktionen bzw. -dichten kann man nur in Spezialfällen die Verteilungsfunktion bzw. -dichte des gesamten Zufallsvektors bestimmen.

Satz von Cramér-Wold

Satz von Cramér-Wold

Die Verteilung des $n-$dimensionalen Zufallsvektors $\underline{\mathbf{X}}$ ist vollständig bestimmt durch die Familie der (eindimensionalen) Verteilungen der Zufallsgrößen $\underline{\mathbf{t}}^{\top}\underline{\mathbf{X}}$, wobei $\underline{\mathbf{t}}$ die Menge $\mathbb{R}^n$ durchläuft.

Erwartungswert eines Zufallsvektors

- Erwartungswerte von Zufallsvektoren und Zufallsmatrizen werden komponentenweise definiert und existieren, falls von jeder Komponente der skalare Erwartungswert existiert.

- Erwartungswert des Zufallsvektors $\underline{\mathbf{X}} = (X_1, \ldots, X_n)^{\mathsf{T}}$:

$$\mathbb{E}\underline{\mathbf{X}} := (\mathbb{E}X_1, \ldots, \mathbb{E}X_n)^{\mathsf{T}} .$$

- Erwartungswert der Zufallsmatrix $\underline{\underline{\mathbf{X}}} = (X_{ij})$:

$$\mathbb{E}\underline{\underline{\mathbf{X}}} := (\mathbb{E}X_{ij}) .$$

Kovarianzmatrix eines Zufallsvektors

- Ein Analogon der Varianz für Zufallsvektoren, deren Komponenten endliche zweite Momente besitzen, ist die Kovarianzmatrix (oder auch Varianz-Kovarianz-Matrix) des Zufallsvektors.
- Kovarianzmatrix des Zufallsvektors $\underline{\mathbf{X}} = (X_1, \ldots, X_n)^{\mathsf{T}}$:

$$\mathbb{V}\mathbf{ar}\underline{\mathbf{X}} := \mathbb{E}\left[[\underline{\mathbf{X}} - \mathbb{E}\underline{\mathbf{X}}] \, [\underline{\mathbf{X}} - \mathbb{E}\underline{\mathbf{X}}]^{\mathsf{T}} \right] .$$

- Auf der Hauptdiagonale der Kovarianzmatrix von $\underline{\mathbf{X}}$ stehen die Varianzen $\mathbb{V}\mathbf{ar}X_i$ der Komponenten, an der Stelle (i, j), $i \neq j$, jeweils die Kovarianz der Zufallsgrößen X_i und X_j: $\mathbb{C}\mathbf{ov}[X_i, X_j]$.
- Eigenschaften einer Kovarianzmatrix $\underline{\underline{\mathbf{\Sigma}}} = \mathbb{V}\mathbf{ar}\underline{\mathbf{X}}$
  - $\underline{\underline{\mathbf{\Sigma}}} = \underline{\underline{\mathbf{\Sigma}}}^{\mathsf{T}}$ ($\underline{\underline{\mathbf{\Sigma}}}$ ist symmetrisch);
  - $\forall \, \underline{\mathbf{x}} \in \mathbb{R}^n : \underline{\mathbf{x}}^{\mathsf{T}} \underline{\underline{\mathbf{\Sigma}}} \, \underline{\mathbf{x}} \geq 0$ ($\underline{\underline{\mathbf{\Sigma}}}$ ist positiv semidefinit).

Korrelationsmatrix eines Zufallsvektors

- Gilt für alle Komponenten eines Zufallsvektors $\underline{\mathbf{X}} = (X_1, \dots, X_n)^\mathsf{T}$ $0 < \mathbb{V}\mathbf{ar} X_i < \infty$, kann man die Korrelationsmatrix definieren:

$$\mathbb{C}\mathbf{orr}\underline{\mathbf{X}} := (\mathbb{C}\mathbf{orr}[X_i, X_j])_{i,j=1,\dots,n} \qquad \text{mit}$$

$$\mathbb{C}\mathbf{orr}[X_i, X_j] := \frac{\mathbb{C}\mathbf{ov}[X_i, X_j]}{\sqrt{\mathbb{V}\mathbf{ar} X_i \mathbb{V}\mathbf{ar} X_j}}\,.$$

- Die Elemente auf der Hauptdiagonale sind 1, das Element an der Stelle (i, j) ist der Korrelationskoeffizient von X_i und X_j.
- Die Korrelationsmatrix des Zufallsvektors $\underline{\mathbf{X}} = (X_1, \dots, X_n)^\mathsf{T}$ ist die Kovarianzmatrix der standardisierten Komponenten von $\underline{\mathbf{X}}$.
- Es gilt immer $-1 \le \mathbb{C}\mathbf{orr}[X_i, X_j] \le 1$.
- Im Fall von $|\mathbb{C}\mathbf{orr}[X_i, X_j]| = 1$ besteht eine lineare Beziehung zwischen den Zufallsgrößen X_i und X_j.

Kreuzkovarianz zweier Zufallsvektoren

- Sind $\underline{\mathbf{X}} = (X_1, \ldots, X_n)^\mathsf{T}$ und $\underline{\mathbf{Y}} = (Y_1, \ldots, Y_m)^\mathsf{T}$ zwei Zufallsvektoren, deren Komponenten endliche zweite Momente besitzen, definiert man die Kreuzkovarianzmatrix dieser Zufallsvektoren als $n \times m-$Matrix

$$\mathbb{C}\mathbf{ov}[\underline{\mathbf{X}}, \underline{\mathbf{Y}}] := \mathbb{E}\Big[[\underline{\mathbf{X}} - \mathbb{E}\underline{\mathbf{X}}]\,[\underline{\mathbf{Y}} - \mathbb{E}\underline{\mathbf{Y}}]^\mathsf{T}\Big]\,.$$

- An der Stelle (i, j) der Kreuzkovarianzmatrix steht die Kovarianz $\mathbb{C}\mathbf{ov}[X_i, Y_j]$ der Zufallsgrößen X_i und Y_j.
- Für einen Zufallsvektor $\underline{\mathbf{X}}$ gilt $\mathbb{V}\mathbf{ar}\underline{\mathbf{X}} = \mathbb{C}\mathbf{ov}[\underline{\mathbf{X}}, \underline{\mathbf{X}}] =: \mathbb{C}\mathbf{ov}\underline{\mathbf{X}}$.
- Es gilt $\mathbb{C}\mathbf{ov}[\underline{\mathbf{X}}, \underline{\mathbf{Y}}] = \mathbb{C}\mathbf{ov}[\underline{\mathbf{Y}}, \underline{\mathbf{X}}]^\mathsf{T}$.
- Analog kann man die Kreuzkorrelationsmatrix $\mathbb{C}\mathbf{orr}[\underline{\mathbf{X}}, \underline{\mathbf{Y}}]$ zweier Zufallsvektoren $\underline{\mathbf{X}}$ und $\underline{\mathbf{Y}}$ definieren.
- Gilt $\mathbb{C}\mathbf{ov}[\underline{\mathbf{X}}, \underline{\mathbf{Y}}] = \underline{\underline{\mathbf{0}}}_{n \times m}$, nennt man $\underline{\mathbf{X}}$ und $\underline{\mathbf{Y}}$ unkorreliert.

Eigenschaften bei linearen Operationen I

- **Geg.:** $\underline{\mathbf{X}}\,,\underline{\mathbf{Y}}$ $n-$dim. ZV, $\mathbb{E}\|\underline{\mathbf{X}}\| < \infty\,,$ $\mathbb{E}\|\underline{\mathbf{Y}}\| < \infty\,.$
  - $a, b \in \mathbb{R} \quad \Rightarrow \quad \mathbb{E}[a\underline{\mathbf{X}} + b\underline{\mathbf{Y}}] = a\,\mathbb{E}\underline{\mathbf{X}} + b\,\mathbb{E}\underline{\mathbf{Y}}\,;$
  - $\underline{\underline{\mathbf{d}}} \in \mathbb{M}_{m\times n}\,,\ \underline{\mathbf{c}} \in \mathbb{R}^m \quad \Rightarrow \quad \mathbb{E}\left[\underline{\underline{\mathbf{d}}}\,\underline{\mathbf{X}} + \underline{\mathbf{c}}\right] = \underline{\underline{\mathbf{d}}}\,\mathbb{E}\underline{\mathbf{X}} + \underline{\mathbf{c}}\,.$
- **Geg.:** $\underline{\mathbf{X}}$ $n-$dim. ZV, $\mathbb{E}\|\underline{\mathbf{X}}\|^2 < \infty\,.$
  - $\mathbb{V}\mathbf{ar}\underline{\mathbf{X}} = \mathbb{E}\left[\underline{\mathbf{X}}\,\underline{\mathbf{X}}^{\mathsf{T}}\right] - (\mathbb{E}\underline{\mathbf{X}})\,(\mathbb{E}\underline{\mathbf{X}})^{\mathsf{T}}\,;$
  - $\underline{\mathbf{a}} \in \mathbb{R}^n \quad \Rightarrow \quad \mathbb{V}\mathbf{ar}\left[\underline{\mathbf{a}}^{\mathsf{T}}\underline{\mathbf{X}}\right] = \underline{\mathbf{a}}^{\mathsf{T}}\mathbb{V}\mathbf{ar}[\underline{\mathbf{X}}]\ \underline{\mathbf{a}}\,;$
  - $\underline{\underline{\mathbf{a}}} \in \mathbb{M}_{m\times n}\,,\ \underline{\mathbf{b}} \in \mathbb{R}^m \quad \Rightarrow \quad \mathbb{V}\mathbf{ar}\left[\underline{\underline{\mathbf{a}}}\,\underline{\mathbf{X}} + \underline{\mathbf{b}}\right] = \underline{\underline{\mathbf{a}}}\,\mathbb{V}\mathbf{ar}[\underline{\mathbf{X}}]\ \underline{\underline{\mathbf{a}}}^{\mathsf{T}}\,.$
- **Geg.:** $\underline{\mathbf{X}}^{(1)},\ \underline{\mathbf{X}}^{(2)}$ $n-$dim. ZV, $\underline{\mathbf{Y}}$ $m-$dim. ZV,
  $\mathbb{E}\|\underline{\mathbf{X}}^{(i)}\|^2 < \infty\,,\ i = 1, 2\,,\quad \mathbb{E}\|\underline{\mathbf{Y}}\|^2 < \infty$
  $\Rightarrow \quad \mathbb{C}\mathbf{ov}\left[\underline{\mathbf{X}}^{(1)} + \underline{\mathbf{X}}^{(2)}, \underline{\mathbf{Y}}\right] = \mathbb{C}\mathbf{ov}\left[\underline{\mathbf{X}}^{(1)}, \underline{\mathbf{Y}}\right] + \mathbb{C}\mathbf{ov}\left[\underline{\mathbf{X}}^{(2)}, \underline{\mathbf{Y}}\right]\,.$

Eigenschaften bei linearen Operationen II

- **Geg.:** $\underline{\mathbf{X}}\,,\underline{\mathbf{Y}}$ $n-$dim. ZV, $\quad \mathbb{E}\|\underline{\mathbf{X}}\|^2 < \infty\,, \quad \mathbb{E}\|\underline{\mathbf{Y}}\|^2 < \infty$

  $\Rightarrow \quad \mathbb{V}\mathbf{ar}[\underline{\mathbf{X}} + \underline{\mathbf{Y}}] = \mathbb{V}\mathbf{ar}\underline{\mathbf{X}} + \mathbb{C}\mathbf{ov}[\underline{\mathbf{X}}, \underline{\mathbf{Y}}] + \mathbb{C}\mathbf{ov}[\underline{\mathbf{Y}}, \underline{\mathbf{X}}] + \mathbb{V}\mathbf{ar}\underline{\mathbf{Y}}\,.$

- **Geg.:** $\underline{\mathbf{X}}$ n_1-dim. ZV, $\quad \underline{\mathbf{Y}}$ n_2-dim. ZV, $\quad \mathbb{E}\|\underline{\mathbf{X}}\|^2 < \infty\,,$

  $\mathbb{E}\|\underline{\mathbf{Y}}\|^2 < \infty\,, \quad \underline{\underline{\mathbf{a}}} \in \mathbb{M}_{m_1 \times n_1}\,, \quad \underline{\underline{\mathbf{b}}} \in \mathbb{M}_{m_2 \times n_2}$

  $\Rightarrow \quad \mathbb{C}\mathbf{ov}\left[\underline{\underline{\mathbf{a}}}\,\underline{\mathbf{X}}, \underline{\underline{\mathbf{b}}}\,\underline{\mathbf{Y}}\right] = \underline{\underline{\mathbf{a}}}\,\mathbb{C}\mathbf{ov}[\underline{\mathbf{X}}, \underline{\mathbf{Y}}]\,\underline{\underline{\mathbf{b}}}^{\mathsf{T}}\,.$

Stochastische Unabhängigkeit von Zufallsvektoren

Geg.: $\underline{\mathbf{X}}$ $n-$dim. ZV, $\underline{\mathbf{Y}}$ $m-$dim. ZV, $\underline{\mathbf{Z}} = \left(\underline{\mathbf{X}}^{\mathsf{T}}, \underline{\mathbf{Y}}^{\mathsf{T}}\right)^{\mathsf{T}}$.

Die Zufallsvektoren $\underline{\mathbf{X}}$ und $\underline{\mathbf{Y}}$ sind (stochastisch) unabhängig, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist (Bedingung (iii) nur, falls der $(n+m)-$dimensionale Zufallsvektor $\underline{\mathbf{Z}}$ (absolut) stetig ist).

(i) $\forall\, B_1 \subseteq \mathbb{R}^n,\ \forall\, B_2 \subseteq \mathbb{R}^m,$ geeignet:
$$\mathbb{P}\left(\{\underline{\mathbf{X}} \in B_1\} \cap \{\underline{\mathbf{Y}} \in B_2\}\right) = \mathbb{P}(\underline{\mathbf{X}} \in B_1) \cdot \mathbb{P}(\underline{\mathbf{Y}} \in B_2)\,.$$

(ii) $\forall\, \underline{\mathbf{x}} \in \mathbb{R}^n,\ \forall\, \underline{\mathbf{y}} \in \mathbb{R}^m\ :\ F_{\underline{\mathbf{Z}}}(\underline{\mathbf{x}}, \underline{\mathbf{y}}) = F_{\underline{\mathbf{X}}}(\underline{\mathbf{x}}) \cdot F_{\underline{\mathbf{Y}}}(\underline{\mathbf{y}})\,.$

(iii) $\forall\, \underline{\mathbf{x}} \in \mathbb{R}^n,\ \forall\, \underline{\mathbf{y}} \in \mathbb{R}^m\ :\ f_{\underline{\mathbf{Z}}}(\underline{\mathbf{x}}, \underline{\mathbf{y}}) = f_{\underline{\mathbf{X}}}(\underline{\mathbf{x}}) \cdot f_{\underline{\mathbf{Y}}}(\underline{\mathbf{y}})\,.$

Aus der Unabhängigkeit folgt die Unkorreliertheit, falls die zweiten Momente existieren.

2.2 Mehrdimensionale Normalverteilung (Multinormalverteilung)

- **Def. 2.2.1**
  Ein $m-$dimensionaler Zufallsvektor $\underline{\mathbf{X}} = (X_1, \ldots, X_m)^{\mathsf{T}}$ besitzt eine $m-$dimensionale Standardnormalverteilung, falls $X_i \sim \mathrm{N}(0,1)$, $i = 1, \ldots, m$, i.i.d.

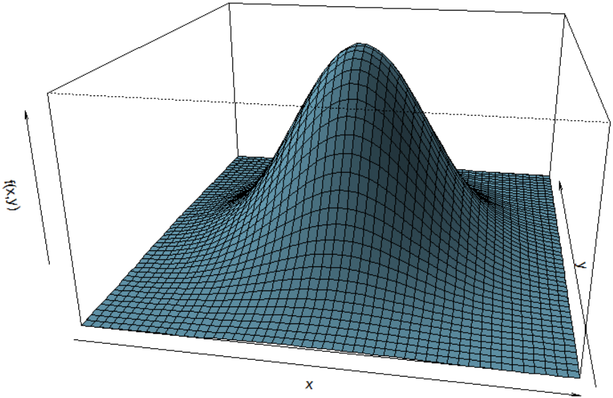
- **Bez.** $\underline{\mathbf{X}} \sim \mathrm{N}_m(\underline{\mathbf{0}}_m, \underline{\underline{\mathbb{I}}}_m)$ oder $\underline{\mathbf{X}} \sim \mathrm{N}(\underline{\mathbf{0}}_m, \underline{\underline{\mathbb{I}}}_m)$.

- **Satz 2.2.2**
  Geg.: Zufallsvektor $\underline{\mathbf{X}} \sim \mathrm{N}_m(\underline{\mathbf{0}}_m, \underline{\underline{\mathbb{I}}}_m)$
  $\Rightarrow$ $\underline{\mathbf{X}}$ ist ein stetiger Zufallsvektor mit Dichtefunktion

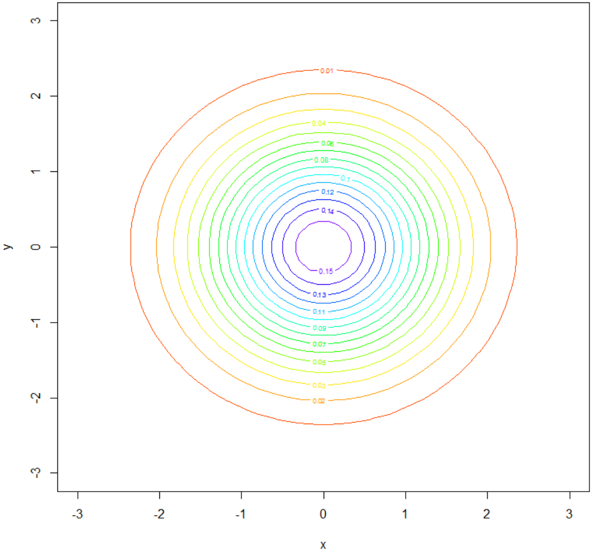
$$f_{\underline{\mathbf{X}}}(\underline{\mathbf{x}}) = (2\pi)^{-m/2} \mathrm{e}^{-\frac{1}{2}\underline{\mathbf{x}}^{\mathsf{T}}\underline{\mathbf{x}}} = (2\pi)^{-m/2} \mathrm{e}^{-\frac{1}{2}\|\underline{\mathbf{x}}\|^2}, \quad \underline{\mathbf{x}} \in \mathbb{R}^m .$$

  Außerdem gelten $\mathbb{E}\underline{\mathbf{X}} = \underline{\mathbf{0}}_m$ und $\mathbb{V}\mathbf{ar}\underline{\mathbf{X}} = \underline{\underline{\mathbb{I}}}_m$.

Dichte der zweidimensionalen Standardnormalverteilung



Höhenlinien der 2-dim. Standardnormalverteilungsdichte



Allgemeine normalverteilte Zufallsvektoren

- **Def. 2.2.3**
  Ein $p-$dimensionaler ZV $\underline{\mathbf{X}} = (X_1, \ldots, X_p)^\mathsf{T}$ besitzt eine *$p-$dimensionale Normalverteilung*, falls für einen $m-$dimensionalen standardnormalverteilten Zufallsvektor $\underline{\mathbf{Z}}$, einen Vektor $\underline{\mu} \in \mathbb{R}^p$ und eine $p \times m-$Matrix $\underline{\underline{\mathbf{a}}}$ gilt
  $$\underline{\mathbf{X}} = \underline{\mu} + \underline{\underline{\mathbf{a}}}\,\underline{\mathbf{Z}}\,.$$
- **Satz 2.2.4**
  Für den Zufallsvektor $\underline{\mathbf{X}}$ aus Def. 2.2.3 gelten
  $$\mathbb{E}\underline{\mathbf{X}} = \underline{\mu} \quad \text{und} \quad \mathbb{V}\mathbf{ar}\underline{\mathbf{X}} = \underline{\underline{\mathbf{a}}}\,\underline{\underline{\mathbf{a}}}^\mathsf{T} =: \underline{\underline{\mathbf{\Sigma}}}\,.$$
- **Bez.** $\underline{\mathbf{X}} \sim \mathrm{N}_p(\underline{\mu}, \underline{\underline{\mathbf{\Sigma}}})$ oder $\underline{\mathbf{X}} \sim \mathrm{N}(\underline{\mu}, \underline{\underline{\mathbf{\Sigma}}})$. Man spricht auch von *Gaußschen Zufallsvektoren*.
- **Bem.** Die $p \times p-$Matrix $\underline{\underline{\mathbf{a}}}\,\underline{\underline{\mathbf{a}}}^\mathsf{T} = \underline{\underline{\mathbf{\Sigma}}}$ ist (z.B. als Kovarianzmatrix) symmetrisch und positiv semidefinit.

Spektraldarstellung reeller symmetrischer Matrizen

Satz 2.2.5

Ist $\underline{\underline{\mathbf{b}}}$ eine reelle symmetrische $p \times p-$Matrix, dann kann sie geschrieben werden als

$$\underline{\underline{\mathbf{b}}} = \underline{\underline{\mathbf{\Gamma}}}\,\underline{\underline{\mathbf{\Lambda}}}\,\underline{\underline{\mathbf{\Gamma}}}^{\mathsf{T}} = \sum_{k=1}^{p} \lambda_k \underline{\gamma}_k \underline{\gamma}_k^{\mathsf{T}},$$

wobei $\underline{\underline{\mathbf{\Lambda}}}$ die Diagonalmatrix der Eigenwerte von $\underline{\underline{\mathbf{b}}}$ ist und $\underline{\underline{\mathbf{\Gamma}}}$ die zugehörige orthogonale Matrix, deren Spalten die standardisierten Eigenvektoren enthalten (es gilt $\underline{\underline{\mathbf{\Gamma}}}\,\underline{\underline{\mathbf{\Gamma}}}^{\mathsf{T}} = \underline{\underline{\mathbf{\Gamma}}}^{\mathsf{T}}\underline{\underline{\mathbf{\Gamma}}} = \underline{\underline{\mathbb{I}}}_p$).

Ist die Matrix $\underline{\underline{\mathbf{b}}}$ positiv semidefinit, dann sind alle Eigenwerte nichtnegativ und es gilt

$$\underline{\underline{\mathbf{b}}} = \underline{\underline{\mathbf{\Gamma}}}\,\underline{\underline{\mathbf{\Lambda}}}^{1/2}\left(\underline{\underline{\mathbf{\Lambda}}}^{1/2}\right)^{\mathsf{T}}\underline{\underline{\mathbf{\Gamma}}}^{\mathsf{T}} = \left(\underline{\underline{\mathbf{\Gamma}}}\,\underline{\underline{\mathbf{\Lambda}}}^{1/2}\right)\left(\underline{\underline{\mathbf{\Gamma}}}\,\underline{\underline{\mathbf{\Lambda}}}^{1/2}\right)^{\mathsf{T}}.$$

Der Rang dieser Matrix ist gleich der Anzahl der Eigenwerte $\neq 0$.

Weitere Eigenschaften

- **Satz 2.2.6**
  Ein $p-$dimensionaler Zufallsvektor $\underline{\mathbf{X}} = (X_1, \ldots, X_p)^{\mathsf{T}}$ ist genau dann normalverteilt, wenn für jeden Vektor $\underline{\mathbf{t}} \in \mathbb{R}^p$ die skalare Zufallsgröße $\underline{\mathbf{t}}^{\mathsf{T}}\underline{\mathbf{X}}$ normalverteilt (oder eine Konstante) ist.

- **Satz 2.2.7**
  Sei $\underline{\mathbf{X}} = (X_1, \ldots, X_p)^{\mathsf{T}} \sim \mathrm{N}_p(\underline{\mu}, \underline{\underline{\mathbf{\Sigma}}})$.
  (i) Für $\underline{\mathbf{b}} \in \mathbb{R}^d$, $\underline{\underline{\mathbf{a}}} \in \mathbb{M}_{d\times p}$ gilt $\underline{\mathbf{Y}} := \underline{\mathbf{b}} + \underline{\underline{\mathbf{a}}}\,\underline{\mathbf{X}} \sim \mathrm{N}_d(\underline{\mathbf{b}} + \underline{\underline{\mathbf{a}}}\,\underline{\mu}, \underline{\underline{\mathbf{a}}}\,\underline{\underline{\mathbf{\Sigma}}}\,\underline{\underline{\mathbf{a}}}^{\mathsf{T}})$.
  (ii) Jeder Teilvektor bzw. jede Komponente von $\underline{\mathbf{X}}$ ist ein normalverteilter Zufallsvektor bzw. eine normalverteilte Zufallsgröße.

- **Bem.**
  Ist jede Komponente eines Zufallsvektors eine normalverteilte (oder konstante) Zufallsgröße, dann muss der Zufallsvektor nicht unbedingt ein normalverteilter Zufallsvektor sein!

Regulär normalverteilte Zufallsvektoren

Def. und Satz 2.2.8
Gilt für einen normalverteilten Zufallsvektor $\underline{\mathbf{X}}$ aus Definition 2.2.3 $p = m$, $\mathbb{V}\mathbf{ar}\underline{\mathbf{X}} = \underline{\underline{\mathbf{a}}}\,\underline{\underline{\mathbf{a}}}^{\mathsf{T}} =: \underline{\underline{\boldsymbol{\Sigma}}}$ mit $\det \underline{\underline{\boldsymbol{\Sigma}}} \neq 0$, dann besitzt $\underline{\mathbf{X}}$ eine reguläre Normalverteilung $\mathrm{N}_p(\underline{\mu}, \underline{\underline{\boldsymbol{\Sigma}}})$ und $\underline{\mathbf{X}}$ ist ein (absolut) stetiger Zufallsvektor mit Dichtefunktion

$$f_{\underline{\mathbf{X}}}(\underline{\mathbf{x}}) = \frac{1}{\sqrt{(2\pi)^p \det \underline{\underline{\boldsymbol{\Sigma}}}}} \mathrm{e}^{-\frac{1}{2}(\underline{\mathbf{x}}-\underline{\mu})^{\mathsf{T}} \underline{\underline{\boldsymbol{\Sigma}}}^{-1} (\underline{\mathbf{x}}-\underline{\mu})}, \qquad \underline{\mathbf{x}} \in \mathbb{R}^p .$$

Der Normierungsfaktor kann auch als $\det(2\pi\underline{\underline{\boldsymbol{\Sigma}}})^{-1/2}$ geschrieben werden, $\underline{\underline{\boldsymbol{\Sigma}}}$ ist eine reelle, symmetrische und positiv definite $p \times p-$Matrix.

Zweidimensionale regulär normalverteilte Zufallsvektoren

Spezialfall $m = p = 2$, $\underline{\mathbf{X}} = (X_1, X_2)^\mathsf{T} \sim \mathrm{N}_2(\underline{\mu}, \underline{\underline{\boldsymbol{\Sigma}}})$ mit

$$\underline{\mu} = (\mu_1, \mu_2)^\mathsf{T} \in \mathbb{R}^2, \quad \underline{\underline{\boldsymbol{\Sigma}}} = \begin{pmatrix} \sigma_1^2 & \varrho\,\sigma_1\sigma_2 \\ \varrho\,\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

$$\sigma_i^2 = \mathbb{V}\mathbf{ar}X_i > 0\,, i = 1, 2, \quad \varrho = \mathbb{C}\mathbf{orr}[X_1, X_2] \in (-1, 1)$$

$$\Rightarrow \qquad \det \underline{\underline{\boldsymbol{\Sigma}}} = \sigma_1^2\sigma_2^2(1 - \varrho^2)\,,$$

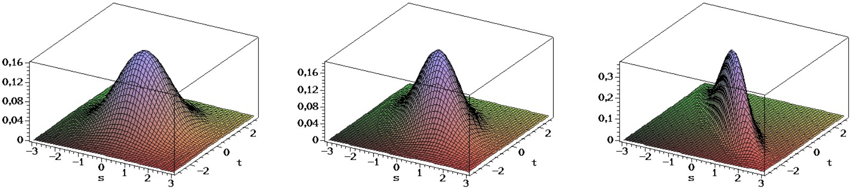
$$\underline{\underline{\boldsymbol{\Sigma}}}^{-1} = \frac{1}{\det \underline{\underline{\boldsymbol{\Sigma}}}} \begin{pmatrix} \sigma_2^2 & -\varrho\,\sigma_1\sigma_2 \\ -\varrho\,\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix},$$

$$f_{(X_1,X_2)}(x_1, x_2) = c \cdot \mathrm{e}^{-\frac{1}{2(1-\varrho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\varrho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]}$$

$$\text{mit} \quad c = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\varrho^2}}\,, \qquad (x_1, x_2)^\mathsf{T} \in \mathbb{R}^2\,.$$

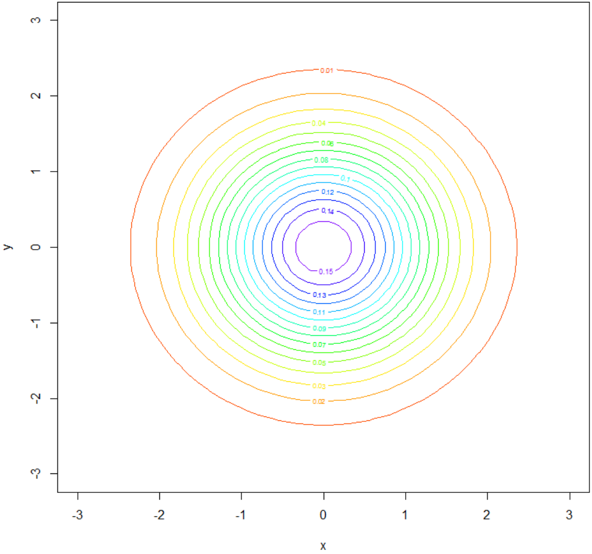
Dichtefunktionsgrafiken zweidimensionaler Normalverteilungen

Dichtefunktionen von normalverteilten Zufallsvektoren $(X_1, X_2)^{\mathsf{T}}$ mit $\mathbb{E}X_1 = \mathbb{E}X_2 = 0\,,\ \mathbb{V}\mathbf{ar}X_1 = \mathbb{V}\mathbf{ar}X_2 = 1$ sowie $\mathbb{C}\mathbf{orr}[X_1, X_2] = 0$ (links), $\mathbb{C}\mathbf{orr}[X_1, X_2] = -0.5$ (mitte) und $\mathbb{C}\mathbf{orr}[X_1, X_2] = -0.9$ (rechts).

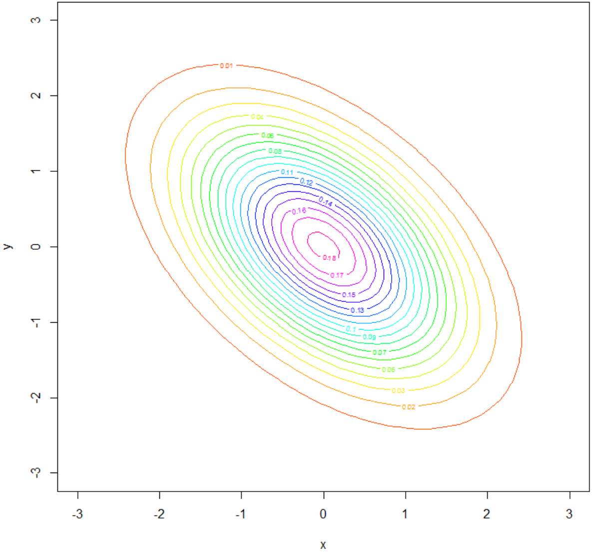


Auf den folgenden Folien folgen die Höhenlinien der jeweiligen Dichtefunktionen.

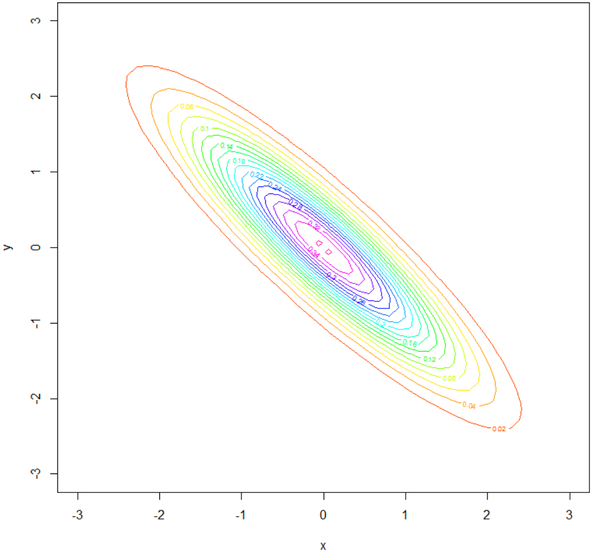
Höhenlinien bei $\mathbb{C}\mathbf{orr}[X_1, X_2] = 0$



Höhenlinien bei $\mathbb{C}\mathbf{orr}[X_1, X_2] = -0.5$



Höhenlinien bei $\mathbb{C}\mathbf{orr}[X_1, X_2] = -0.9$



Unabhängigkeit von Teilvektoren bei Normalverteilung

Satz 2.2.9

Geg. $\underline{\mathbf{X}} = (\underline{\mathbf{X}}_1{}^\mathsf{T}, \underline{\mathbf{X}}_2{}^\mathsf{T})^\mathsf{T} \sim \mathrm{N}_p(\underline{\mu}, \underline{\underline{\mathbf{\Sigma}}})$,

$$\underline{\mathbf{X}}_1 = (X_1, \ldots, X_k)^\mathsf{T}, \underline{\mathbf{X}}_2 = (X_{k+1}, \ldots, X_p)^\mathsf{T}.$$

Dann sind $\underline{\mathbf{X}}_1$ und $\underline{\mathbf{X}}_2$ genau dann stochastisch unabhängig, wenn $\mathbb{C}\mathbf{ov}[\underline{\mathbf{X}}_1, \underline{\mathbf{X}}_2] = \underline{\underline{\mathbf{0}}}_{k\times(p-k)}$ gilt.