Multivariate Normalverteilung und Gauß'sche Prozesse

Tobias Müller

Institut für Stochastik

03.06.13

Inhaltsverzeichnis

Definition

Voraussetzung für multivariate Normalverteilung

Sei $\mu \in \mathbb{R}^n$ und $K = (k_{ij})_{1 \leq i,j \leq n} \in \mathbb{R}^{n \times n}$ eine symmetrisch positiv definite $(n \times n)$ Matrix.

Definition (K regulär d.h $rg(K) = n$ (voll!))

Sei $X = (X_1, \dots, X_n)^T$ ein absolut stetiger Zufallsvektor und

$$f_X(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\sqrt{detK}} \exp\left(-\frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) \quad \forall x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$$

die zugehörige Dichtefunktion.

Der Zufallsvektor $X = (X_1, \dots, X_n)^T$ wird dann als multivariat normalverteilt bezeichnet.
Schreibweise: $X \sim N(\mu, K)$

Definition

Voraussetzung für multivariate Normalverteilung

Sei $\mu \in \mathbb{R}^n$ und $K = (k_{ij})_{1 \leq i,j \leq n} \in \mathbb{R}^{n \times n}$ eine symmetrisch positiv definite $(n \times n)$ Matrix.

Definition (K regulär d.h $rg(K) = n$ (voll!))

Sei $X = (X_1, \ldots, X_n)^T$ ein absolut stetiger Zufallsvektor und

$$f_X(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\sqrt{detK}} \exp\left(-\frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) \quad \forall x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$$

die zugehörige Dichtefunktion.

Der Zufallsvektor $X = (X_1, \ldots, X_n)^T$ wird dann als multivariat normalverteilt bezeichnet.
Schreibweise: $X \sim N(\mu, K)$

Bemerkung

Man kann zeigen, dass

$$\int\limits_{\mathbb{R}} \dots \int\limits_{\mathbb{R}} \exp\left(-\frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) dx_1 \dots dx_n \overset{!}{=} (2\pi)^{n/2}(detK)^{\frac{1}{2}}$$

$\Rightarrow \int\limits_{\mathbb{R}^n} f_X(x)dx = 1$; $f_X(x) \geq 0$ *und* $f_X(x)$ *ist messbar*

$\Rightarrow f_X(x)$ *ist somit die Dichtefunktion von* X

Beweis

- *Da K symmetrisch* $\exists$ *orthogonale Matrix V mit* $V^T KV = diag(\lambda_1, \dots, \lambda_n)$ *und da EW* > 0 *sind und K invertierbar ist*
  $\Rightarrow (V^T KV)^{-1} = V^T K^{-1} V = diag(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n})$
- *Substituiere* $y = \varphi(x) = V^T(x - \mu)$ *wobei* $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ *bijektiv ist*
  $\Rightarrow$ *Jacobi Matrix*

$$J = \begin{pmatrix} \frac{\partial \varphi_1(x)}{\partial x_1} & \dots & \frac{\partial \varphi_1(x)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \varphi_n(x)}{\partial x_1} & \dots & \frac{\partial \varphi_n(x)}{\partial x_n} \end{pmatrix} = V^T$$

Für die Jacobi Determinante von $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ *gilt somit, dass*

$$det\left(\frac{\partial \varphi_i}{\partial x_j}(x_1, \dots, x_n)\right) = det\,(V) = \pm 1$$

, wobei die letzte Gleichung aus $1 = det\left(V^T V\right) = (det V)^2$ *folgt.*

Beweis

- *Da K symmetrisch $\exists$ orthogonale Matrix V mit $V^T K V = diag(\lambda_1, \ldots, \lambda_n)$ und da EW > 0 sind und K invertierbar ist*
  $\Rightarrow (V^T K V)^{-1} = V^T K^{-1} V = diag(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n})$
- *Substituiere $y = \varphi(x) = V^T(x - \mu)$ wobei $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ bijektiv ist*
  $\Rightarrow$ *Jacobi Matrix*

$$J = \begin{pmatrix} \frac{\partial \varphi_1(x)}{\partial x_1} & \cdots & \frac{\partial \varphi_1(x)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \varphi_n(x)}{\partial x_1} & \cdots & \frac{\partial \varphi_n(x)}{\partial x_n} \end{pmatrix} = V^T$$

Für die Jacobi Determinante von $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ gilt somit, dass

$$det \left(\frac{\partial \varphi_i}{\partial x_j}(x_1, \ldots, x_n) \right) = det\,(V) = \pm 1$$

, wobei die letzte Gleichung aus $1 = det\left(V^T V\right) = (det V)^2$ folgt.

Beweis (Fortsetzung)

$$\Rightarrow \int_{\mathbb{R}} \ldots \int_{\mathbb{R}} \exp\left(-\frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) dx_1 \ldots dx_n$$

Beweis (Fortsetzung)

$$\Rightarrow \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) dx_1 \dots dx_n$$

$$\overset{y=V^T(x-\mu)}{=} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2} y^T (V^T K^{-1} V) y\right) dy_1 \dots dy_n$$

Beweis (Fortsetzung)

$$\Rightarrow \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) dx_1 \dots dx_n$$

$$\overset{y=V^T(x-\mu)}{=} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2} y^T (V^T K^{-1} V) y\right) dy_1 \dots dy_n$$

$$= \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} \frac{y_i^2}{\lambda_i}\right) dy_1 \dots dy_n$$

Beweis (Fortsetzung)

$$\Rightarrow \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) dx_1 \dots dx_n$$

$$\overset{y=V^T(x-\mu)}{=} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2} y^T (V^T K^{-1} V) y\right) dy_1 \dots dy_n$$

$$= \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} \frac{y_i^2}{\lambda_i}\right) dy_1 \dots dy_n$$

$$= \prod_{i=1}^{n} \int_{\mathbb{R}} \exp\left(-\frac{1}{2} \frac{y_i^2}{\lambda_i}\right) dy_i \overset{t_i = \frac{y_i}{\sqrt{\lambda_i}}}{=} (2\pi)^{n/2} (detK)^{1/2}$$

Beweis (Fortsetzung)

$$\Rightarrow \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) dx_1 \dots dx_n$$

$$\overset{y=V^T(x-\mu)}{=} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2} y^T (V^T K^{-1} V) y\right) dy_1 \dots dy_n$$

$$= \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \frac{y_i^2}{\lambda_i}\right) dy_1 \dots dy_n$$

$$= \prod_{i=1}^{n} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\frac{y_i^2}{\lambda_i}\right) dy_i \overset{t_i=\frac{y_i}{\sqrt{\lambda_i}}}{=} (2\pi)^{n/2} (detK)^{1/2}$$

$$\left[\int_{\mathbb{R}} \exp\left(-\frac{1}{2}t_i^2\right) dy_i = \sqrt{2\pi} \quad und \quad \prod_{i=1}^{n} \sqrt{\lambda_i} = \sqrt{det(V^T K V)} = \sqrt{detK}\right]$$

Beweis (Fortsetzung)

$$\Rightarrow \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) dx_1 \dots dx_n$$

$$\overset{y=V^T(x-\mu)}{=} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2} y^T (V^T K^{-1} V) y\right) dy_1 \dots dy_n$$

$$= \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \frac{y_i^2}{\lambda_i}\right) dy_1 \dots dy_n$$

$$= \prod_{i=1}^{n} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\frac{y_i^2}{\lambda_i}\right) dy_i \overset{t_i = \frac{y_i}{\sqrt{\lambda_i}}}{=} (2\pi)^{n/2}(detK)^{1/2}$$

$$\left[\int_{\mathbb{R}} \exp\left(-\tfrac{1}{2}t_i^2\right) dy_i = \sqrt{2\pi} \qquad \textit{und} \qquad \prod_{i=1}^{n} \sqrt{\lambda_i} = \sqrt{det(V^T K V)} = \sqrt{detK}\right]$$

charakteristische Funktion

Theorem

Sei $X \sim N(\mu, K)$. Dann ist die charakteristische Funktion gegeben durch

$$\varphi_X(t) = \exp\left(it^T\mu - \frac{1}{2}t^T Kt\right) \quad \forall t \in \mathbb{R}^n$$

Beweis

$$\varphi_X(t) = \mathbb{E}e^{it^T X} = \int_{\mathbb{R}^n} e^{it^T x} f_X(x) dx$$

Beweis

$$\varphi_X(t) = \mathbb{E}e^{it^TX} = \int_{\mathbb{R}^n} e^{it^Tx} f_X(x)dx$$

$$= \frac{1}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(it^Tx - \frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) dx_1 \dots dx_n$$

Beweis

$$\varphi_X(t) = \mathbb{E}e^{it^TX} = \int_{\mathbb{R}^n} e^{it^Tx} f_X(x)dx$$

$$= \frac{1}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(it^Tx - \frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) dx_1 \dots dx_n$$

$$\overset{y=x-\mu}{=} \frac{e^{it^T\mu}}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(it^Ty - \frac{1}{2}y^T K^{-1}y\right) dy_1 \dots dy_n$$

Beweis

$$\begin{aligned}
\varphi_X(t) = \mathbb{E}e^{it^T X} &= \int_{\mathbb{R}^n} e^{it^T x} f_X(x)dx \\
&= \frac{1}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(it^T x - \frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) dx_1 \dots dx_n \\
&\overset{y=x-\mu}{=} \frac{e^{it^T \mu}}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(it^T y - \frac{1}{2}y^T K^{-1} y\right) dy_1 \dots dy_n \\
&\underset{t=Vs}{\overset{y=Vx}{=}} \frac{e^{it^T \mu}}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(is^T x - \frac{1}{2}x^T (V^T K^{-1} V)x\right) dx_1 \dots dx_n
\end{aligned}$$

Beweis

$$\varphi_X(t) = \mathbb{E}e^{it^T X} = \int_{\mathbb{R}^n} e^{it^T x} f_X(x) dx$$

$$= \frac{1}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(it^T x - \frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) dx_1 \dots dx_n$$

$$\overset{y=x-\mu}{=} \frac{e^{it^T\mu}}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(it^T y - \frac{1}{2}y^T K^{-1} y\right) dy_1 \dots dy_n$$

$$\overset{\substack{y=Vx\\t=Vs}}{=} \frac{e^{it^T\mu}}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(is^T x - \frac{1}{2}x^T (V^T K^{-1} V) x\right) dx_1 \dots dx_n$$

$$= \frac{e^{it^T\mu}}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(\sum_{j=1}^{n}\left(is_j x_j - \frac{1}{2}\frac{x_j^2}{\lambda_j}\right)\right) dx_1 \dots dx_n$$

Beweis

$$
\begin{aligned}
\varphi_X(t) &= \mathbb{E}e^{it^TX} = \int_{\mathbb{R}^n} e^{it^Tx} f_X(x)dx \\
&= \frac{1}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \exp\left(it^Tx - \frac{1}{2}(x-\mu)^T K^{-1}(x-\mu)\right) dx_1 \dots dx_n \\
&\overset{y=x-\mu}{=} \frac{e^{it^T\mu}}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \exp\left(it^Ty - \frac{1}{2}y^T K^{-1} y\right) dy_1 \dots dy_n \\
&\overset{\substack{y=Vx \\ t=Vs}}{=} \frac{e^{it^T\mu}}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \exp\left(is^Tx - \frac{1}{2}x^T (V^T K^{-1} V)x\right) dx_1 \dots dx_n \\
&= \frac{e^{it^T\mu}}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \exp\left(\sum_{j=1}^n \left(is_jx_j - \frac{1}{2}\frac{x_j^2}{\lambda_j}\right)\right) dx_1 \dots dx_n \\
&= \frac{e^{it^T\mu}}{(2\pi)^{n/2}(detK)^{1/2}} \prod_{j=1}^n \int_{\mathbb{R}} \exp\left(is_jx_j - \frac{1}{2}\frac{x_j^2}{\lambda_j}\right) dx_j
\end{aligned}
$$

Beweis

$$\varphi_X(t) = \mathbb{E}e^{it^T X} = \int_{\mathbb{R}^n} e^{it^T x} f_X(x)dx$$

$$= \frac{1}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(it^T x - \frac{1}{2}(x-\mu)^T K^{-1}(x-\mu) \right) dx_1 \dots dx_n$$

$$\overset{y=x-\mu}{=} \frac{e^{it^T \mu}}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(it^T y - \frac{1}{2} y^T K^{-1} y \right) dy_1 \dots dy_n$$

$$\overset{\substack{y=Vx \\ t=Vs}}{=} \frac{e^{it^T \mu}}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(is^T x - \frac{1}{2} x^T (V^T K^{-1} V) x \right) dx_1 \dots dx_n$$

$$= \frac{e^{it^T \mu}}{(2\pi)^{n/2}(detK)^{1/2}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \exp\left(\sum_{j=1}^{n} \left(is_j x_j - \frac{1}{2}\frac{x_j^2}{\lambda_j} \right) \right) dx_1 \dots dx_n$$

$$= \frac{e^{it^T \mu}}{(2\pi)^{n/2}(detK)^{1/2}} \prod_{j=1}^{n} \int_{\mathbb{R}} \exp\left(is_j x_j - \frac{1}{2}\frac{x_j^2}{\lambda_j} \right) dx_j$$

Beweis (Fortsetzung)

$$= e^{it^T \mu} \prod_{j=1}^{n} \underbrace{\frac{1}{\sqrt{2\pi\lambda_j}} \int_{\mathbb{R}} \exp\left(i s_j x_j - \frac{1}{2} \frac{x_j^2}{\lambda_j} \right) dx_j}_{(*)}$$

(*) $= \mathbb{E} e^{i s_j X_j} \overset{WR}{=} \exp\left(-\frac{1}{2}\lambda_j s_j^2\right)$, *da (*) der charakteristischen Funktion der* $N(0, \lambda_j)$*-Verteilung genügt*

Beweis (Fortsetzung)

$$= e^{it^T\mu} \prod_{j=1}^{n} \underbrace{\frac{1}{\sqrt{2\pi\lambda_j}} \int_{\mathbb{R}} \exp\left(is_j x_j - \frac{1}{2}\frac{x_j^2}{\lambda_j}\right) dx_j}_{(*)}$$

$$= e^{it^T\mu} \prod_{j=1}^{n} \exp\left(\frac{1}{2}\lambda_j s_j^2\right) = e^{it^T\mu} \exp\left(-\frac{1}{2}\sum_{j=1}^{n} \lambda_j s_j^2\right)$$

(*) $= \mathbb{E}e^{is_j X_j} \overset{WR}{=} \exp\left(-\frac{1}{2}\lambda_j s_j^2\right)$, *da (*) der charakteristischen Funktion der* $N(0, \lambda_j)$*-Verteilung genügt*

Beweis (Fortsetzung)

$$= e^{it^T\mu} \prod_{j=1}^{n} \underbrace{\frac{1}{\sqrt{2\pi\lambda_j}} \int_{\mathbb{R}} \exp\left(is_j x_j - \frac{1}{2}\frac{x_j^2}{\lambda_j} \right) dx_j}_{(*)}$$

$$= e^{it^T\mu} \prod_{j=1}^{n} \exp\left(\frac{1}{2}\lambda_j s_j^2 \right) = e^{it^T\mu} \exp\left(-\frac{1}{2}\sum_{j=1}^{n} \lambda_j s_j^2 \right)$$

$$\overset{s=V^{-1}t}{=} \exp\left(it^T\mu - \frac{1}{2} t^T K t \right)$$

(*) $= \mathbb{E}e^{is_j X_j} \overset{WR}{=} \exp\left(-\frac{1}{2}\lambda_j s_j^2\right)$, *da (*) der charakteristischen Funktion der $N(0, \lambda_j)$-Verteilung genügt*

Beweis (Fortsetzung)

$$= e^{it^T\mu} \prod_{j=1}^{n} \underbrace{\frac{1}{\sqrt{2\pi\lambda_j}} \int_{\mathbb{R}} \exp\left(is_j x_j - \frac{1}{2}\frac{x_j^2}{\lambda_j} \right) dx_j}_{(*)}$$

$$= e^{it^T\mu} \prod_{j=1}^{n} \exp\left(\frac{1}{2}\lambda_j s_j^2\right) = e^{it^T\mu} \exp\left(-\frac{1}{2} \sum_{j=1}^{n} \lambda_j s_j^2 \right)$$

$$\overset{s=V^{-1}t}{=} \exp\left(it^T\mu - \frac{1}{2} t^T K t \right)$$

(*) $= \mathbb{E} e^{is_j X_j} \overset{WR}{=} \exp\left(-\frac{1}{2}\lambda_j s_j^2\right)$, *da (*) der charakteristischen Funktion der* $N(0,\lambda_j)$*-Verteilung genügt*

Eindeutigkeitssatz der charakteristischen Funktion

Eindeutigkeitssatz

Seien $X, Y : \Omega \to \mathbb{R}^n$ beliebige Zufallsvektoren. Dann gilt:

$$X \overset{d}{=} Y \Longleftrightarrow \varphi_X(t) = \varphi_Y(t) \quad \forall t = (t_1, \dots, t_n) \in \mathbb{R}^n$$

Verteilung von beliebigen Teilvektoren

Theorem

Sei $X = (X_1, \ldots, X_n)^T \sim N(\boldsymbol{\mu}, \boldsymbol{K})$ und $\pi : \{1, \ldots, m\} \to \{1, \ldots, m\}$ eine Permutation (bijektiv). Dann gilt:

$$(X_{\pi(1)}, \ldots, X_{\pi(m)})^T \sim N(\boldsymbol{\mu}_{\pi(m)}, \boldsymbol{K}_{\pi(m)}) \quad \forall m = 1, \ldots, n$$

wobei $\boldsymbol{\mu}_{\pi(m)} = (\mu_{\pi(1)}, \ldots, \mu_{\pi(m)})$ und $\boldsymbol{K}_{\pi(m)}$ die $(m \times m)$-Matrix, die die entsprechenden m Zeilen bzw. Spalten von K enthält.

Beweis

o.B.d.A betrachten wir die ersten m Einträge $(\pi(i) = i \quad \forall i = 1, \ldots, m)$

- $\varphi : \mathbb{R}^n \to \mathbb{C}$ *ist die charakteristische Funktion von* $(X_1, \ldots, X_n)^T$
- $\varphi_m : \mathbb{R}^n \to \mathbb{C}$ *ist die charakteristische Funktion von* $(X_1, \ldots, X_m)^T$

Beweis

o.B.d.A betrachten wir die ersten m Einträge $(\pi(i) = i \quad \forall i = 1, \ldots, m)$

- $\varphi : \mathbb{R}^n \to \mathbb{C}$ *ist die charakteristische Funktion von* $(X_1, \ldots, X_n)^T$
- $\varphi_m : \mathbb{R}^n \to \mathbb{C}$ *ist die charakteristische Funktion von* $(X_1, \ldots, X_m)^T$

Dann ist $\varphi_m(\boldsymbol{t}_m) = \varphi((\boldsymbol{t}_m, \underbrace{0, \ldots, 0}_{n-m})) \quad \forall \boldsymbol{t}_m = (t_1, \ldots, t_m)^T \in \mathbb{R}^m$

$\Rightarrow \varphi_m(\boldsymbol{t}_m) = \exp\left(i\boldsymbol{t}_m^T \boldsymbol{\mu}_m - \frac{1}{2}\boldsymbol{t}_m^T \boldsymbol{K}_m \boldsymbol{t}_m\right) \quad \forall \boldsymbol{t}_m \in \mathbb{R}^m$

Beweis

o.B.d.A betrachten wir die ersten m Einträge $(\pi(i) = i \quad \forall i = 1, \dots, m)$

- $\varphi : \mathbb{R}^n \to \mathbb{C}$ *ist die charakteristische Funktion von* $(X_1, \dots, X_n)^T$
- $\varphi_m : \mathbb{R}^n \to \mathbb{C}$ *ist die charakteristische Funktion von* $(X_1, \dots, X_m)^T$

Dann ist $\varphi_m(\boldsymbol{t}_m) = \varphi((\boldsymbol{t}_m, \underbrace{0, \dots, 0}_{n-m})) \quad \forall \boldsymbol{t}_m = (t_1, \dots, t_m)^T \in \mathbb{R}^m$

$\Rightarrow \varphi_m(\boldsymbol{t}_m) = \exp\left(i\boldsymbol{t}_m^T \boldsymbol{\mu}_m - \frac{1}{2}\boldsymbol{t}_m^T \boldsymbol{K}_m \boldsymbol{t}_m\right) \quad \forall \boldsymbol{t}_m \in \mathbb{R}^m$

Außerdem ist $\boldsymbol{K}_m$ *positiv definit und symmetrisch, da* $\boldsymbol{K}$ *positiv definit und symmetrisch ist.*

$\overset{\text{Eindeutigkeitssatz}}{\Longrightarrow} (X_1, \dots, X_m)^T \sim N(\boldsymbol{\mu}_m, \boldsymbol{K}_m) \quad \forall m = 1, \dots, n$

Beweis

o.B.d.A betrachten wir die ersten m Einträge $(\pi(i) = i \quad \forall i = 1, \dots, m)$

- $\varphi : \mathbb{R}^n \to \mathbb{C}$ *ist die charakteristische Funktion von* $(X_1, \dots, X_n)^T$
- $\varphi_m : \mathbb{R}^n \to \mathbb{C}$ *ist die charakteristische Funktion von* $(X_1, \dots, X_m)^T$

Dann ist $\varphi_m(\boldsymbol{t}_m) = \varphi((\boldsymbol{t}_m, \underbrace{0, \dots, 0}_{n-m})) \quad \forall \boldsymbol{t}_m = (t_1, \dots, t_m)^T \in \mathbb{R}^m$

$$\Rightarrow \varphi_m(\boldsymbol{t}_m) = \exp\left(i\boldsymbol{t}_m^T\boldsymbol{\mu}_m - \tfrac{1}{2}\boldsymbol{t}_m^T\boldsymbol{K}_m\boldsymbol{t}_m\right) \quad \forall \boldsymbol{t}_m \in \mathbb{R}^m$$

Außerdem ist $\boldsymbol{K}_m$ *positiv definit und symmetrisch, da* $\boldsymbol{K}$ *positiv definit und symmetrisch ist.*

$$\stackrel{\text{Eindeutigkeitssatz}}{\Longrightarrow} (X_1, \dots, X_m)^T \sim N(\boldsymbol{\mu}_m, \boldsymbol{K}_m) \quad \forall m = 1, \dots, n$$

Lineare Transformation

Theorem

Sei $Y \sim N(\mu, K)$ ein n-dimensionaler normalverteilter Zufallsvektor. Außerdem sei $m \leq n$ und $A \in \mathbb{R}^{m \times n}$ beliebig mit $rg(A) = m$ (voll!) und sei $c \in \mathbb{R}^m$ beliebig. Dann gilt:

$$X = AY + c \sim N(A\mu + c, AKA^T)$$

Beweis

$$\begin{aligned}\varphi_X(t) = \varphi_{AY+c}(t) = \mathbb{E}e^{(it^T(AY+c))} &= e^{it^Tc}\mathbb{E}e^{i(A^Tt)^TY} \\ &= e^{it^Tc}\varphi_Y(A^Tt) = e^{it^Tc}\exp\left(i(A^Tt)^T\mu - \frac{1}{2}(A^Tt)^TK(A^Tt)\right) \\ &= \exp\left(it^T(A\mu + c) - \frac{1}{2}t^T(AKA^T)t\right)\end{aligned}$$

- $rg(A) = m \Rightarrow A^Tx \neq 0 \quad \forall x \in \mathbb{R}^m \setminus \{0\}$ *Dann gilt*
  $$x^TAKA^Tx = \underbrace{x^TA}_{\in\mathbb{R}^{1\times n}\setminus\{0\}} \; K \; \underbrace{A^Tx}_{\in\mathbb{R}^n\setminus\{0\}} \underbrace{>}_{K \text{ pos. def.}} 0$$
- $(AKA^T)^T = AK^TA^T \overset{K\ symm.}{=} AKA^T$

Beweis

$$\begin{aligned}\varphi_X(t) = \varphi_{AY+c}(t) = \mathbb{E}e^{(it^T(AY+c))} &= e^{it^T c}\mathbb{E}e^{i(A^T t)^T Y} \\ &= e^{it^T c}\varphi_Y(A^T t) = e^{it^T c}\exp\left(i(A^T t)^T \mu - \frac{1}{2}(A^T t)^T K (A^T t)\right) \\ &= \exp\left(it^T(A\mu + c) - \frac{1}{2}t^T(AKA^T)t\right)\end{aligned}$$

- $rg(A) = m \Rightarrow A^T x \neq 0 \quad \forall x \in \mathbb{R}^m \setminus \{0\}$ *Dann gilt*

  $$x^T AKA^T x = \underbrace{x^T A}_{\in\mathbb{R}^{1\times n}\setminus\{0\}} \; K \; \underbrace{A^T x}_{\in\mathbb{R}^n\setminus\{0\}} \underbrace{>}_{K \text{ pos. def.}} 0$$

- $(AKA^T)^T = AK^T A^T \overset{K\ symm.}{=} AKA^T$

Mit dem Eindeutigkeitssatz für charakteristische Funktionen folgt die Behauptung, da AKA^T positiv definit und symmetrisch ist

Beweis

$$\begin{aligned}\varphi_X(t) = \varphi_{AY+c}(t) = \mathbb{E}e^{(it^T(AY+c))} &= e^{it^Tc}\mathbb{E}e^{i(A^Tt)^TY} \\ &= e^{it^Tc}\varphi_Y(A^Tt) = e^{it^Tc}\exp\left(i(A^Tt)^T\mu - \frac{1}{2}(A^Tt)^TK(A^Tt)\right) \\ &= \exp\left(it^T(A\mu + c) - \frac{1}{2}t^T(AKA^T)t\right)\end{aligned}$$

- $rg(A) = m \Rightarrow A^Tx \neq 0 \quad \forall x \in \mathbb{R}^m \setminus \{0\}$ *Dann gilt*

  $$x^TAKA^Tx = \underbrace{x^TA}_{\in\mathbb{R}^{1\times n}\setminus\{0\}} \; K \; \underbrace{A^Tx}_{\in\mathbb{R}^n\setminus\{0\}} \; \underbrace{>}_{K \text{ pos. def.}} \; 0$$

- $(AKA^T)^T = AK^TA^T \overset{K \text{ symm.}}{=} AKA^T$

Mit dem Eindeutigkeitssatz für charakteristische Funktionen folgt die Behauptung, da AKA^T *positiv definit und symmetrisch ist*

Singuläre multivariate Normalverteilung

Lemma

Sei K symmetrisch und positiv semidefinit und $rg(K) = r < n$.
Dann $\exists$ Matrix $B \in \mathbb{R}^{n \times r}$ mit $rg(B) = r$, so dass $K = BB^T$

Definition (K singulär d.h $rg(K) = r < n$)

Y heißt singulär normalverteilt, falls

$$Y \stackrel{d}{=} \mu + BZ$$

mit $B \in \mathbb{R}^{n \times r}(rg(B) = r)$ und $Z \sim N(0, I_r)$
Schreibweise: $Y \sim N(\mu, K)$

Singuläre multivariate Normalverteilung

Lemma

Sei K symmetrisch und positiv semidefinit und $rg(K) = r < n$.
Dann $\exists$ Matrix $B \in \mathbb{R}^{n \times r}$ mit $rg(B) = r$, so dass $K = BB^T$

Definition (K singulär d.h $rg(K) = r < n$)

Y heißt singulär normalverteilt, falls

$$Y \overset{d}{=} \mu + BZ$$

mit $B \in \mathbb{R}^{n \times r}(rg(B) = r)$ und $Z \sim N(0, I_r)$
Schreibweise: $Y \sim N(\mu, K)$

Bemerkung

- *Falls* $rg(K) = r < n \Rightarrow Y$ *nicht absolut stetig d.h Verteilung von* Y *hat keine Dichte bzgl. des Lebesguemaßes*
- *Verteilung von* $\mu + BZ$ *hängt nicht von der Wahl von* B *ab*

Grund:
Die charakteristische Funktion von $Y = \mu + BZ$ *ist gegeben durch*
$\varphi_Y(t) \stackrel{!}{=} \exp\left(it^T\mu - \frac{1}{2}t^TKt\right)$

Beweis

$$\varphi_Y(t) = \varphi_{BZ+\mu}(t) = \mathbb{E}e^{it^T(BZ+\mu)} = e^{it^T\mu}\mathbb{E}e^{i(B^Tt)^TZ} = e^{it^T\mu}\varphi_Z(B^Tt)$$

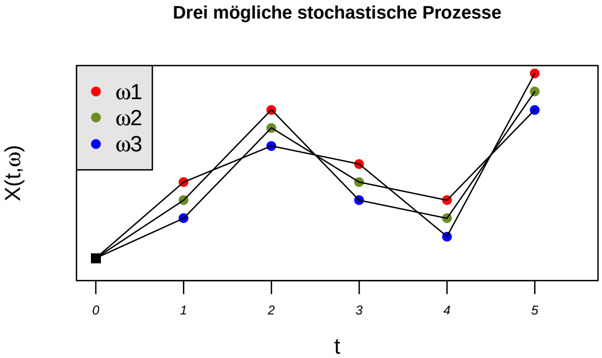
$$\overset{Z\sim N(0,I_r)}{=} \exp\left(it^T\mu\right)\exp\left(-\frac{1}{2}(B^Tt)^T I_r B^T t\right)$$

$$= \exp\left((it^T\mu) - \frac{1}{2}t^T BB^T t\right)$$

$$= \exp\left(it^T\mu - \frac{1}{2}t^T Kt\right)$$

Was ist ein stochastischer Prozess?

Wir betrachten einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ mit $\Omega = \{\omega_1, \omega_2, \omega_3\}$ und eine Folge $(X_t)_{t \in T}$ mit Abbildung $X_t : \Omega \to \mathbb{R}$, $T = \{1, \ldots, 5\}$ und $X_0 = c$



Definition

Ein stochastischer Prozess ist eine Familie $\{X_t, t \in T\}$ von Zufallsvariabeln $X_t : \Omega \to E$, die über einunddemselben Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ definiert sind.

- Indexmenge T kann beliebig sein [d.h diskrete Zeit(T abzählbar), kontinuierliche Zeit($T \subseteq \mathbb{R}$) oder eine Teilmenge von $\mathbb{R}^d$ sein]
- Bildraum kann ein beliebiger Messraum $(E, \mathcal{B}(E))$ sein, wobei $\mathcal{B}(E)$ die Borel σ-Algebra von E ist.

Definition

Ein stochastischer Prozess ist eine Familie $\{X_t, t \in T\}$ von Zufallsvariabeln $X_t : \Omega \to E$, die über einunddemselben Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ definiert sind.

- Indexmenge T kann beliebig sein [d.h diskrete Zeit(T abzählbar), kontinuierliche Zeit($T \subseteq \mathbb{R}$) oder eine Teilmenge von $\mathbb{R}^d$ sein]
- Bildraum kann ein beliebiger Messraum $(E, \mathcal{B}(E))$ sein, wobei $\mathcal{B}(E)$ die Borel σ-Algebra von E ist.

Definition

Seien $n = 1, 2, \ldots$ und $0 \leq t_0 < t_1 < \cdots < t_n$ beliebige Zahlen.

- *stationäre Zuwächse:* Familie $\{X_t, t \geq 0\}$ ist ein Prozess mit stationären Zuwächsen, falls für jedes $h \geq 0$ gilt:
  $(X_{t_1+h} - X_{t_0+h}, \ldots, X_{t_n+h} - X_{t_{n-1}+h}) \stackrel{d}{=} (X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}})$
- *unabhängige Zuwächse:* Familie $\{X_t, t \geq 0\}$ ist ein Prozess mit unabhängigen Zuwächsen, falls gilt:
  Zufallsvariabeln $X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ sind unabhänig

Definition

Seien $n = 1, 2, \ldots$ und $0 \leq t_0 < t_1 < \cdots < t_n$ beliebige Zahlen.

- *stationäre Zuwächse:* Familie $\{X_t, t \geq 0\}$ ist ein Prozess mit stationären Zuwächsen, falls für jedes $h \geq 0$ gilt: $(X_{t_1+h} - X_{t_0+h}, \ldots, X_{t_n+h} - X_{t_{n-1}+h}) \stackrel{d}{=} (X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}})$
- *unabhängige Zuwächse:* Familie $\{X_t, t \geq 0\}$ ist ein Prozess mit unabhängigen Zuwächsen, falls gilt: Zufallsvariabeln $X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ sind unabhänig

Definition

Definition

Ein stochastischer Prozess $\{X_t, t \in T\}$ wird Gauß-Prozess genannt, falls der Zufallsvektor $(X_{t_1}, \ldots, X_{t_n})$ normalverteilt ist $\forall t_1, \ldots, t_n \in T, n \in \mathbb{N}$

Bemerkung

- *T kann diskret, kontinuierlich oder eine Teilmenge von $\mathbb{R}^n$ sein*
- *Der Gauß-Prozess ist vollständig durch die Erwartungswertfunktion $\mu(t) = \mathbb{E}[X_t]$ und die Kovarianzfunktion $K(s,t) = Cov(X_s, X_t)$ bestimmt, wobei $s, t \in T$ ist*

Definition

Definition

Ein stochastischer Prozess $\{X_t, t \in T\}$ wird Gauß-Prozess genannt, falls der Zufallsvektor $(X_{t_1}, \ldots, X_{t_n})$ normalverteilt ist $\forall t_1, \ldots, t_n \in T, n \in \mathbb{N}$

Bemerkung

- *T kann diskret, kontinuierlich oder eine Teilmenge von $\mathbb{R}^n$ sein*
- *Der Gauß-Prozess ist vollständig durch die Erwartungswertfunktion $\mu(t) = \mathbb{E}[X_t]$ und die Kovarianzfunktion $K(s,t) = Cov(X_s, X_t)$ bestimmt, wobei $s, t \in T$ ist*

Wiener Prozess als Beispiel

Definition

Eine Familie $\{W_t, t \in T\}$ heißt Wiener Prozess falls nachfolgende Bedingungen gelten:

- $W_0 = 0$
- $(W_t)_{t \geq 0}$ hat unabhängige + stationäre Zuwächse
- $W_t \sim N(0, t)$
- der Pfad $t \mapsto X_t(\omega)$, $t \in T$ ist für jedes $\omega \in \Omega$ eine stetige Funktion

Bemerke:

- $W_t - W_s \stackrel{d}{=} W_{t+h} - W_{s+h} \stackrel{h=-s}{=} W_{t-s} - W_0 = W_{t-s} \sim N(0, t-s)$
- $Cov(W_s, W_t) = min\{s, t\} \quad \forall s, t \geq 0$

Wiener Prozess als Beispiel

Behauptung: **Wiener Prozess ist ein Gauß-Prozess**

Wiener Prozess als Beispiel

Beweis

Sei $(W_t)_{t\geq 0}$ *ein WP,* $n \in \mathbb{N}$ *und* $t_1, \dots, t_n \in \mathbb{R}^+$ *mit o.B.d.A* $0 = t_0 < t_1 < \dots < t_n$. *Da* $(W_t)_{t\geq 0}$ *stationäre Zuwächse hat und* $X_t \sim N(0, W_t) \overset{!}{\Rightarrow} V = (W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}})^T$ *ist ein Gauß'scher Zufallsvektor.*

Grund:

- $W_{t_i} - W_{t_{i-1}} \overset{d}{=} W_{t_i - t_{i-1}} \sim N(0, t_i - t_{i-1}) \quad \forall i = 1, \dots, n$
- $Cov(W_{t_i - t_{i-1}}, W_{t_j - t_{j-1}}) = \min\{t_i - t_{i-1}, t_j - t_{j-1}\} \, \forall i, j = 1, \dots, n$

Wiener Prozess als Beispiel

Beweis

Sei $(W_t)_{t\geq 0}$ *ein WP,* $n \in \mathbb{N}$ *und* $t_1, \ldots, t_n \in \mathbb{R}^+$ *mit o.B.d.A* $0 = t_0 < t_1 < \cdots < t_n$. *Da* $(W_t)_{t\geq 0}$ *stationäre Zuwächse hat und* $X_t \sim N(0, W_t) \stackrel{!}{\Rightarrow} V = (W_{t_1} - W_{t_0}, \ldots, W_{t_n} - W_{t_{n-1}})^T$ *ist ein Gauß'scher Zufallsvektor.*

Grund:

- $W_{t_i} - W_{t_{i-1}} \stackrel{d}{=} W_{t_i - t_{i-1}} \sim N(0, t_i - t_{i-1}) \quad \forall i = 1, \ldots, n$
- $Cov(W_{t_i - t_{i-1}}, W_{t_j - t_{j-1}}) = \min\{t_i - t_{i-1}, t_j - t_{j-1}\} \forall i, j = 1, \ldots, n$

$\Rightarrow$ *Kovarianzmatrix K des Zufallsvektors ist dann symmetrisch und positiv definit.*

$\Rightarrow V \sim N(0, K)$

Wiener Prozess als Beispiel

Beweis

Sei $(W_t)_{t\geq 0}$ *ein WP,* $n \in \mathbb{N}$ *und* $t_1, \dots, t_n \in \mathbb{R}^+$ *mit o.B.d.A* $0 = t_0 < t_1 < \dots < t_n$. *Da* $(W_t)_{t\geq 0}$ *stationäre Zuwächse hat und* $X_t \sim N(0, W_t) \overset{!}{\Rightarrow} V = (W_{t_1} - W_{t_0}, \dots, W_{t_n} - W_{t_{n-1}})^T$ *ist ein Gauß'scher Zufallsvektor.*
Grund:

- $W_{t_i} - W_{t_{i-1}} \overset{d}{=} W_{t_i - t_{i-1}} \sim N(0, t_i - t_{i-1}) \quad \forall i = 1, \dots, n$
- $Cov(W_{t_i - t_{i-1}}, W_{t_j - t_{j-1}}) = \min\{t_i - t_{i-1}, t_j - t_{j-1}\} \forall i, j = 1, \dots, n$

$\Rightarrow$ *Kovarianzmatrix K des Zufallsvektors ist dann symmetrisch und positiv definit.*
$\Rightarrow V \sim N(0, K)$

Beweis (Fortsetzung)

Wir definieren jetzt eine untere $n \times n$ Dreiecksmatrix A mit lauter Einseinträgen. Durch Multiplikation von A mit dem Zufallsvektor V erhält man wegen der linearen Transformation, dass

$$AV = (W_{t_1}, \dots, W_{t_n})^T \sim N(0, AKA^T)$$

$\Rightarrow (W_t)_{t \geq 0}$ *ist ein Gauß-Prozess*

Quellen:

T.W.Anderson;
An Introduction to Multivariate Statistical Analysis;
Wiley 1984, 2.Ausgabe

M.Lifshits;
Lecture on Gaussian Processes;
Springer 2012, 1.Auflage

V.Schmidt;
Vorlesungsskript Stochastik II;
Universität Ulm WS 2011/2012

Vielen Dank für die Aufmerksamkeit!