Die Karhunen-Loève Zerlegung

Cassandra Uebel und Tim Jaschek

21.01.2016

Setting: Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, D ein abgeschlossenes Intervall und $X : D \times \Omega \to \mathbb{R}$ ein darauf definierter stetiger, zentrierter und quadratintegrierbarer stochastischer Prozess.

Ziel: Dann gibt es für X_t die folgende Darstellung als Reihe:

$$X_t = \sum_{n=1}^{\infty} Z_n \psi_n(t)$$

für gewisse Zufallsvariablen $(Z_n)_{n\in\mathbb{N}}$ und eine Orthonormalbasis $\{\psi_n : n \in \mathbb{N}\}$ von $L^2(D)$.

Zu den Persönlichkeiten

Betrachtung der Thematik erstmals von Damodar Dharmananda Kosambi
$\hookrightarrow$ Veröffentlichung **1943**

Dennoch: Bekannt als Karhunen-Loeve-Theorem

Zu den Persönlichkeiten

Betrachtung der Thematik erstmals von Damodar Dharmananda Kosambi
$\hookrightarrow$ Veröffentlichung **1943**

Dennoch: Bekannt als Karhunen-Loeve-Theorem

$\hookrightarrow$ Kari Karhunen:

- finnischer Herkunft
- 1947 Doktorarbeit auf Deutsch
- **1945** Veröffentlichung des Theorems

$\hookrightarrow$ Michel Loeve:

- israelisch-jüdische Herkunft
- Student unter Levy
- Inhaftierungslager
- Professor in Barkley
- **1948** Veröffentlichung des Theorems

Funktionsanalytischer Hintergrund

Definition

Sei X ein normierter Raum, $A \subseteq X$ heißt **relativ kompakt** *in X, falls $\bar{A}$ kompakt in X ist.*

Funktionsanalytischer Hintergrund

Definition

Sei X ein normierter Raum, $A \subseteq X$ heißt **relativ kompakt** *in X, falls $\bar{A}$ kompakt in X ist.*

Definition

Seien X, Y normierte Räume. Eine Abbildung $A\colon X \to Y$ heißt **linearer Operator**, *falls sie linear ist. Wir schreiben $Ax := A(x)$.*

Funktionsanalytischer Hintergrund

Definition

Sei X ein normierter Raum, $A \subseteq X$ heißt **relativ kompakt** *in X, falls $\bar{A}$ kompakt in X ist.*

Definition

Seien X, Y normierte Räume. Eine Abbildung $A\colon X \to Y$ heißt **linearer Operator***, falls sie linear ist. Wir schreiben $Ax := A(x)$.*

Definition

Gilt zudem für alle beschränkte Mengen $E \subseteq X$, dass $A(E)$ relativ kompakt in Y ist, dann bezeichnen wir diesen Operator als **kompakt***.*

Funktionsanalytischer Hintergrund

Definition

Wir bezeichnen die Funktion $\kappa : D \times D \to \mathbb{R}$ *als einen* **Hilbert-Schmidt-Kern**, *falls:*

$$\int_D \int_D |\kappa(x, y)|^2 dxdy < \infty$$

Funktionsanalytischer Hintergrund

Definition

Wir bezeichnen die Funktion $\kappa : D \times D \to \mathbb{R}$ als einen **Hilbert-Schmidt-Kern**, *falls:*

$$\int_D \int_D |\kappa(x,y)|^2 dxdy < \infty$$

Definiere den Integraloperator T_κ auf $L^2(D)$ durch $T_\kappa : u \mapsto T_\kappa u$ für $u \in L^2(D)$:

$$[T_\kappa u](x) = \int_D \kappa(x,y)u(y)dy$$

.

Funktionsanalytischer Hintergrund

Definiere den Integraloperator T_κ auf $L^2(D)$ durch $T_\kappa : u \mapsto T_\kappa u$ für $u \in L^2(D)$:

$$[T_\kappa u](x) = \int_D \kappa(x, y) u(y) dy$$

.

Lemma

Der oben definierte Integraloperator T_κ mit zugehörigen Hilbert-Schmidt-Kern $\kappa \in L^2(D \times D)$ ist linear, beschränkt und kompakt.
Solch ein Integraloperator wird als **Hilbert-Schmidt-Operator** bezeichnet.

Lemma

T_{κ_X} *ist linear.*

Beweis: Seien $f, g \in L^2(D)$ und $\lambda \in \mathbb{R}$. Dann gilt:

$$
\begin{aligned}
T_\kappa(\lambda f + g) &= \int_D \kappa(s,\cdot)(\lambda f(s) + g(s))ds \\
&= \lambda \int_D \kappa(s,\cdot) f(s) ds + \int_D \kappa(s,\cdot) g(s) ds \\
&= \lambda T_\kappa f + T_\kappa g
\end{aligned}
$$

$\square$

Definiere den Integraloperator K auf $L^2(D)$ durch $K : u \mapsto Ku$ für $u \in L^2(D)$

$$[T_\kappa u](x) = \int_D \kappa(x, y)u(y)dy$$

Beweis:

- T_κ ist linear. $\surd$
- T_κ ist beschränkt.

Definiere den Integraloperator K auf $L^2(D)$ durch $K : u \mapsto Ku$ für $u \in L^2(D)$

$$[T_\kappa u](x) = \int_D \kappa(x, y)u(y)dy$$

Beweis:

- T_κ ist linear. $\surd$
- T_κ ist beschränkt. $\surd$
- T_κ ist kompakt.

Definiere den Integraloperator K auf $L^2(D)$ durch $K : u \mapsto Ku$ für $u \in L^2(D)$

$$[T_\kappa u](x) = \int_D \kappa(x,y)u(y)dy$$

Beweis:

- T_κ ist linear. $\surd$
- T_κ ist beschränkt. $\surd$
- T_κ ist kompakt. $\surd$

$\square$

Spektral Theorie kompakter selbstadjungierter Operatoren

Falls A linearer Operator ist, macht es Sinn nach Eigenvektoren und Eigenwerten im Sinne der Linearen Algebra zu fragen.

Definition

Sei X ein normierte Raum und $A \in \mathcal{L}(X, X)$ ein linearer Operator. $\psi \in X \setminus \{0\}$ heißt **Eigenvektor** *oder auch* **Eigenfunktion** *zum* **Eigenwert** *λ von A, falls gilt:*

$$A\psi = \lambda\psi$$

Beachte, dass dim(X) auch ∞ seien kann. Dies wird in unserem Setting auch der Fall sein.

Spektral Theorie kompakter selbstadjungierter Operatoren

Sei X ein normierter Raum, $A \in \mathcal{L}(X, X)$. Definiere:

Resolvente:

$$\rho(A) = \{\lambda \in \mathbb{C} : ker(\lambda - A) = \{0\}, im(\lambda - A) = X\}$$

Spektrum:

$$\sigma(A) = \mathbb{C} \setminus \rho(A)$$

Spektral Theorie kompakter selbstadjungierter Operatoren

Punktspektrum:

$$\sigma_p(A) = \{\lambda \in \mathbb{C} : ker(\lambda - A) \neq \{0\}\}$$

Stetiges Spektrum:

$$\sigma_c(A) = \left\{\lambda \in \mathbb{C} : \begin{array}{c} ker(\lambda - A) = \{0\} \\ im(\lambda - A) \neq X, \overline{im(\lambda - A)} = X \end{array}\right\}$$

Residuenspektrum:

$$\sigma_r(A) = \{\lambda \in \mathbb{C} : ker(\lambda - A) = \{0\}, \overline{im(\lambda - A)} \neq X\}$$

Definition und Lemma

Sei $(H, < \cdot, \cdot >)$ ein Hilbertraum. $T : H \to H$ ein linearer Operator. T heißt selbstadjungiert, falls für alle $u, v \in H$:

$$< Tu, v > = < u, Tv >$$

Falls T ein Hilbert- Schmidt Operator auf $L^2(D)$ mit zugehörigem Kern κ ist, so gilt: T ist selbstadjungiert $\Leftrightarrow$ κ ist symmetrisch

Beweis:

$$\begin{aligned} < T_\kappa u, v > &= \int_D \int_D \kappa(y, x) u(y) dy v(x) dx \\ < u, T_\kappa v > &= \int_D \int_D \kappa(x, y) v(x) dx u(y) dy \end{aligned}$$

Spektral Theorie

Spektral Theorem

Sei H ein Hilbertraum und $T : H \to H$ ein kompakter, selbstadjungierter Operator. Dann hat H eine Orthonormalbasis $\{\psi_n : n \in \mathbb{N}\}$ aus Eigenvektoren von T zu den zugehörigen Eigenwerten λ_i. Es gilt:

- $\forall i \in \mathbb{N} : \lambda_i \in \mathbb{R}$ mit Null als einzigem Häufungspunkt
- Eigenräume zu verschiedenen Eigenwerten sind orthogonal
- Eigenräume zu von Null verschiedenen Eigenwerten sind endlich dimensional.

Spektral Theorie

Spektral Theorem

Sei H ein Hilbertraum und $T : H \to H$ ein kompakter, selbstadjungierter Operator. Dann hat H eine Orthonormalbasis $\{\psi_n : n \in \mathbb{N}\}$ aus Eigenvektoren von T zu den zugehörigen Eigenwerten λ_i. Es gilt:

- $\forall i \in \mathbb{N} : \lambda_i \in \mathbb{R}$ mit Null als einzigem Häufungspunkt
- Eigenräume zu verschiedenen Eigenwerten sind orthogonal
- Eigenräume zu von Null verschiedenen Eigenwerten sind endlich dimensional.

Beweis:

Functional Analysis lecture, HU Berlin, WS 15/16, Prof. Perkowski, voraussichtlich in den nächsten Wochen

Kovarianzfunktion von X:

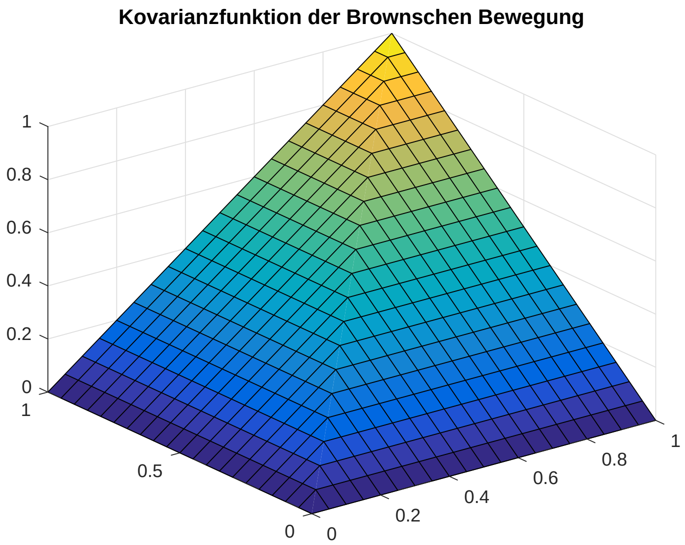
$$\begin{aligned} \kappa_X : \quad D \times D &\to \mathbb{R} \\ (s,t) &\mapsto \kappa_X(s,t) = \mathbb{E}[X_s X_t] \end{aligned}$$

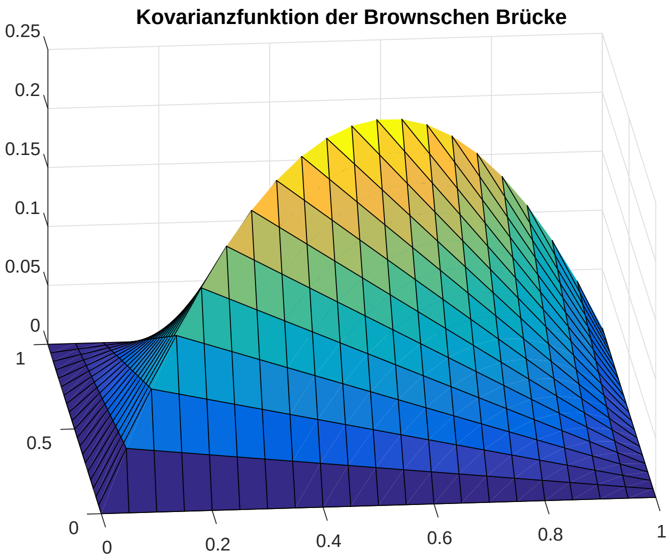
Beispiel: Kovarianzfunktion einer Brownschen Bewegung:

$$\kappa_B(s,t) = s \wedge t$$

Integraloperator induziert durch K_X:

$$\begin{aligned} T_{\kappa_X} : \quad L^2(D) &\to L^2(D) \\ f &\mapsto T_{\kappa_X} f = \int_D \kappa_X(s,\cdot) f(s) ds \end{aligned}$$


Kovarianzfunktion der Brownschen Bewegung
1
0.8
0.6
0.4
0.2
0
1
0.5
0
0
0.2
0.4
0.6
0.8
1


Kovarianzfunktion der Brownschen Brücke
0.25
0.2
0.15
0.1
0.05
0
1
0.5
0
0
0.2
0.4
0.6
0.8
1

Wie sehen die **Eigenwerte und Eigenvektoren** von T_{K_X} aus? Dazu löse:

$$\left(T_{\kappa_X}\psi_n\right)(t) = \int_D \kappa_X(s,t)\psi_n(s)ds = \lambda_n\psi_n(t)$$

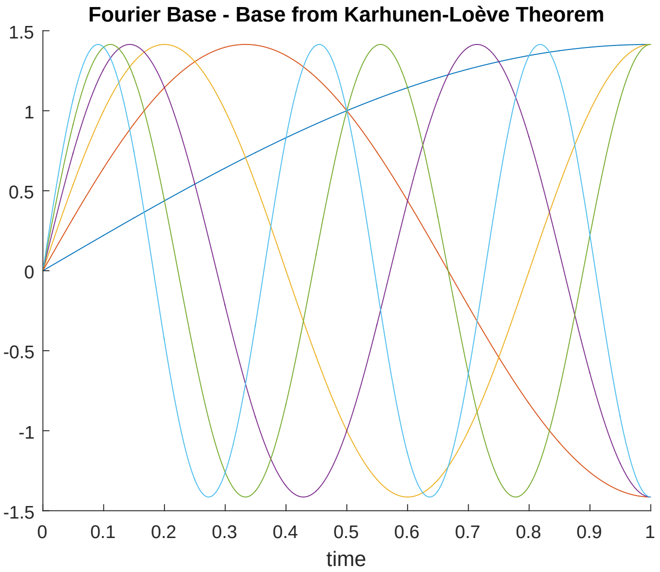
Wie sehen die **Eigenwerte und Eigenvektoren** von T_{K_X} aus?
Dazu löse:

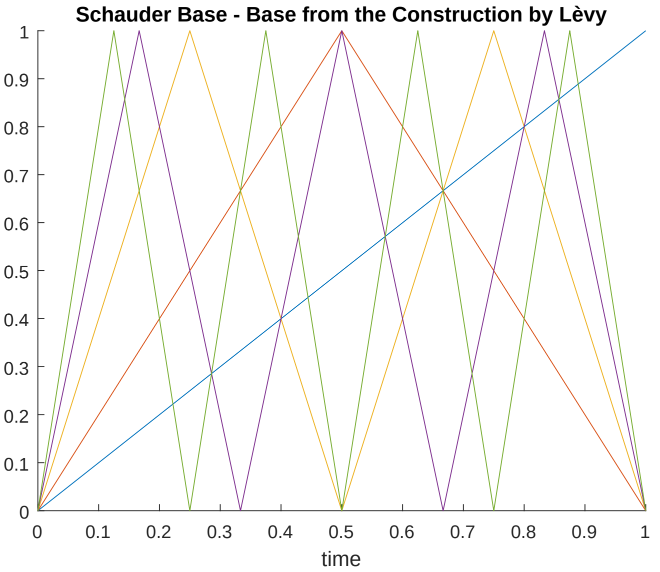
$$\left(T_{\kappa_X}\psi_n\right)(t) = \int_D \kappa_X(s,t)\psi_n(s)ds = \lambda_n\psi_n(t)$$

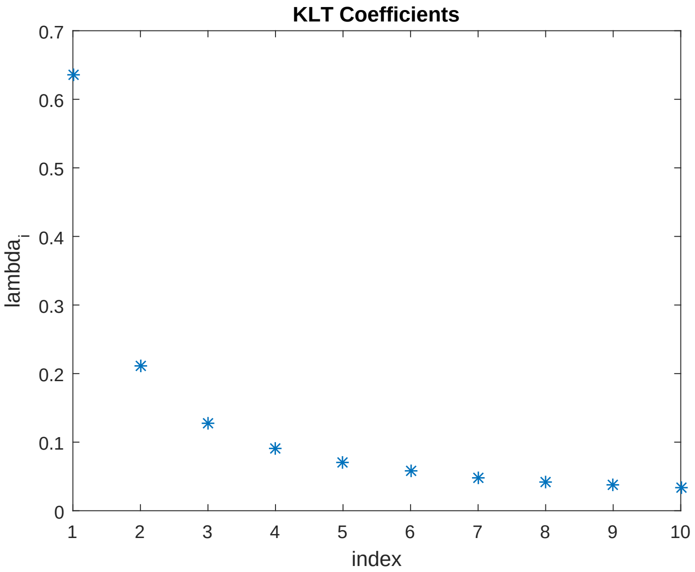
Am Beispiel der Brownschen Bewegung mit $D = [0,1]$ erhält man:

$$\begin{aligned}\psi_n(t) &= \sqrt{2}sin\left(\left(n-\frac{1}{2}\right)\pi t\right)\\ \lambda_n &= \frac{1}{\left(n-\frac{1}{2}\right)^2\pi^2}\end{aligned}$$

$\longrightarrow$ Tafel


Fourier Base - Base from Karhunen-Loève Theorem
1.5
1
0.5
0
-0.5
-1
-1.5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
time


Schauder Base - Base from the Construction by Lèvy
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
time


KLT Coefficients
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
lambda_i
1
2
3
4
5
6
7
8
9
10
index

Definition

Sei $\kappa : D \times D \to \mathbb{R}$ ein Hilbert- Schmidt Kern. Der induzierte Integraloperator T_κ heißt positiv, falls für alle $f \in L^2(D)$:

$$\int_D \int_D \kappa(s,t) f(s) f(t) ds dt \geq 0$$

Mercer's Theorem

Sei $\kappa : D \times D \to \mathbb{R}$ stetig mit positiv definitem induzierten Hilbert-Schmidt Operator T_κ. Seien weiter $\{\lambda_i\}$ und $\{\psi_i\}$ die Eigenwerte und Eigenfunktionen von T_κ. Dann gilt für alle $s, t \in D$:

$$\kappa(s,t) = \sum_{n=1}^{\infty} \lambda_n \psi_n(s) \psi_n(t)$$

wobei die obige Reihe absolut und gleichmäßig konvergiert.

Mercer's Theorem

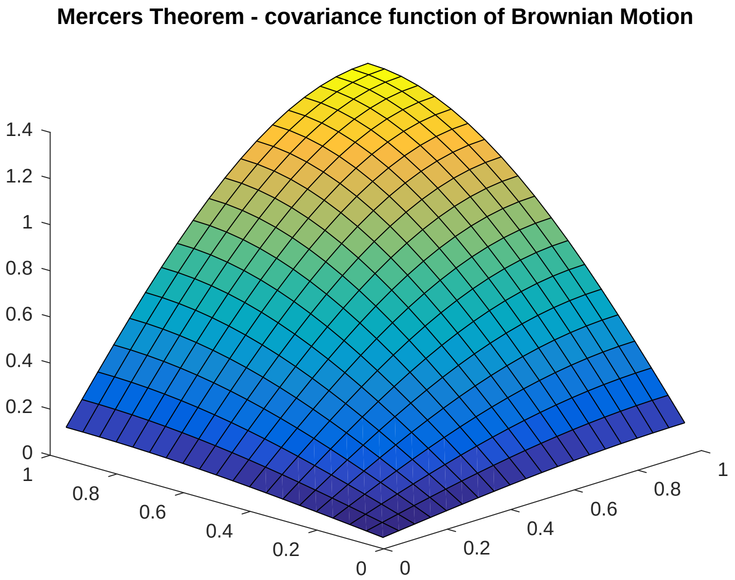
Sei $\kappa : D \times D \to \mathbb{R}$ stetig mit positiv definitem induzierten Hilbert-Schmidt Operator T_κ. Seien weiter $\{\lambda_i\}$ und $\{\psi_i\}$ die Eigenwerte und Eigenfunktionen von T_κ. Dann gilt für alle $s, t \in D$:

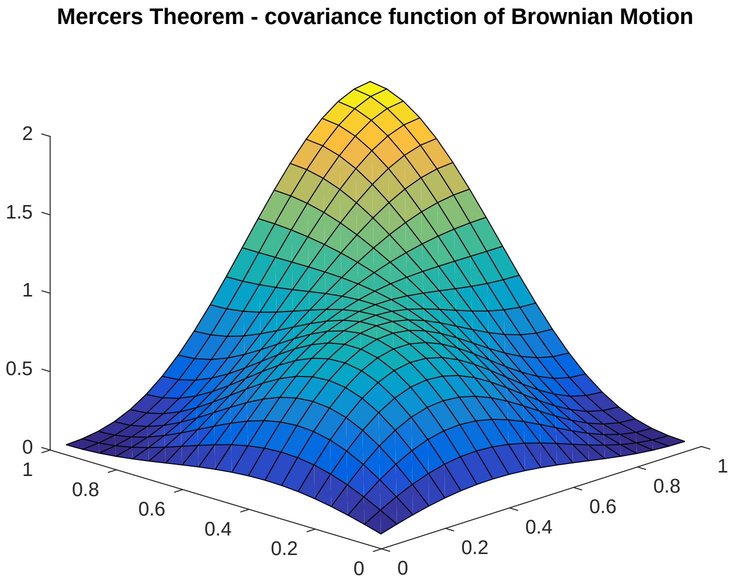
$$\kappa(s,t) = \sum_{n=1}^{\infty} \lambda_n \psi_n(s) \psi_n(t)$$

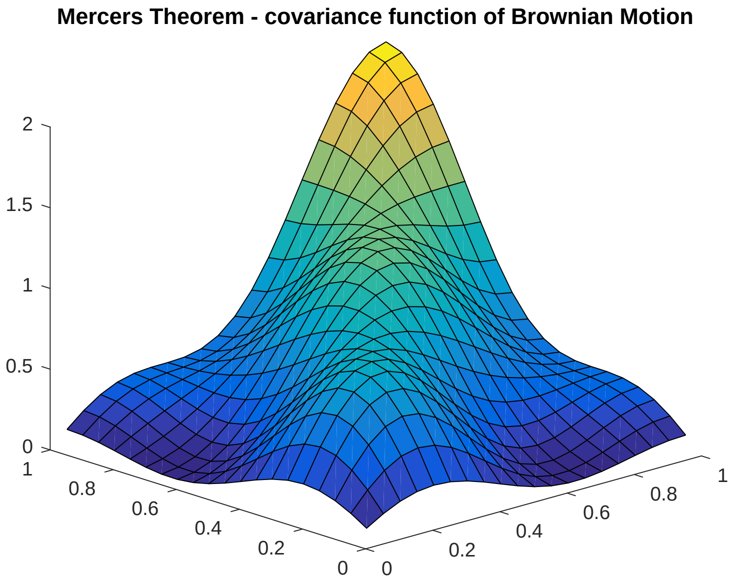
wobei die obige Reihe absolut und gleichmäßig konvergiert.

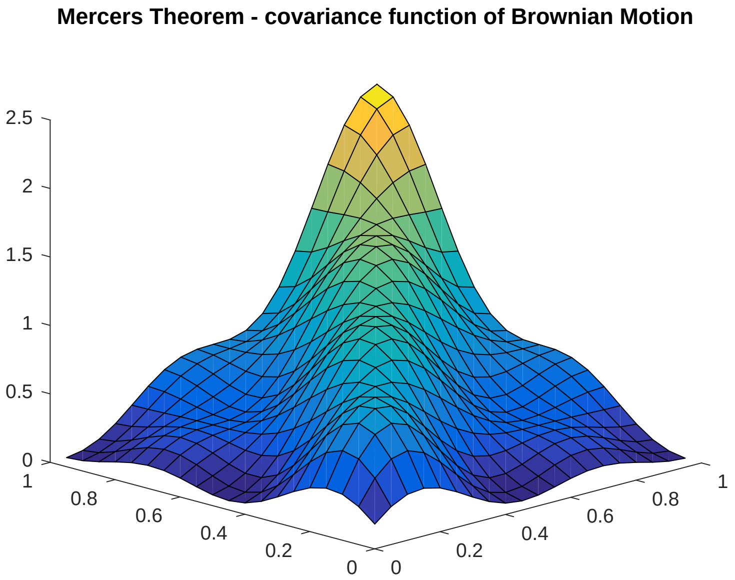
Beweis:

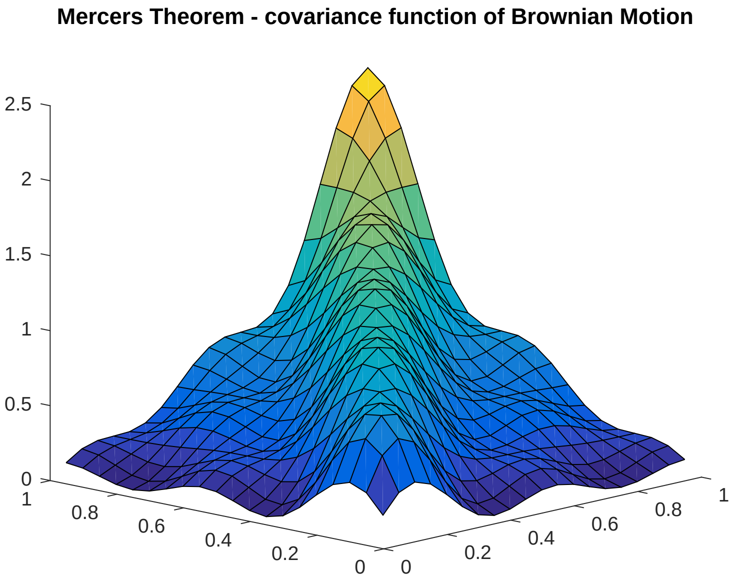
Israel Gohberg, Seymour Goldberg und M.A. Kaashoek. *Basic classes of linear operators.* 2004.


Mercers Theorem - covariance function of Brownian Motion
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1


Mercers Theorem - covariance function of Brownian Motion
2
1.5
1
0.5
0
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1


Mercers Theorem - covariance function of Brownian Motion
2
1.5
1
0.5
0
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1


Mercers Theorem - covariance function of Brownian Motion
2.5
2
1.5
1
0.5
0
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1


Mercers Theorem - covariance function of Brownian Motion
2.5
2
1.5
1
0.5
0
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1

Karhunen-Loève Theorem

Sei $X : D \times \Omega \to \mathbb{R}$ ein stetiger stochastischer Prozess mit $\mathbb{E}[X_t] = 0$ und $X_t \in L^2$ für jedes $t \in D$.
Dann gilt für X_t die folgende Darstellung:

$$X_t = \sum_{n=1}^{\infty} Z_n \psi_n(t)$$

Für Zufallsvariablen Z_n mit $\mathbb{E}[Z_n] = 0$ und $\mathbb{E}[Z_n Z_m] = \delta_{m,n} \lambda_n$ und eine orthonormale Basis $\{\psi_n | n \in \mathbb{N}\}$ von $L^2(D)$. Diese Reihe konvergiert in L^2 und gleichmäßig und es gilt:

$$Z_n = \int_D X_t \psi_n(t) dt$$

Beweisstrategie:

i) κ_X ist ein Hilbert-Schmidt Kern $\surd$

ii) Herleitung der Darstellung bzgl Orthonormalbasis aus i)

Beweisstrategie:

i) κ_X ist ein Hilbert-Schmidt Kern $\surd$

ii) Herleitung der Darstellung bzgl Orthonormalbasis aus i) $\surd$

iii) Nachweis der Eigenschaften der einzelnen Komponenten der Reihe

Beweisstrategie:

i) κ_X ist ein Hilbert-Schmidt Kern $\surd$
ii) Herleitung der Darstellung bzgl Orthonormalbasis aus i) $\surd$
iii) Nachweis der Eigenschaften der einzelnen Komponenten der Reihe $\surd$
iv) Konvergenzaussagen

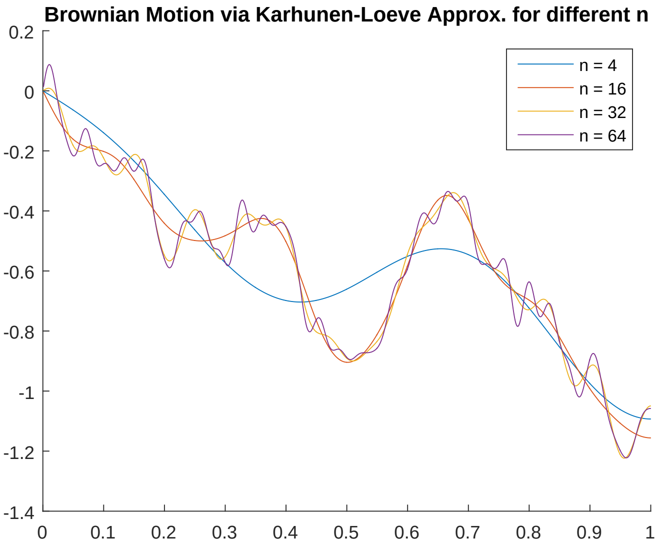
Beweisstrategie:

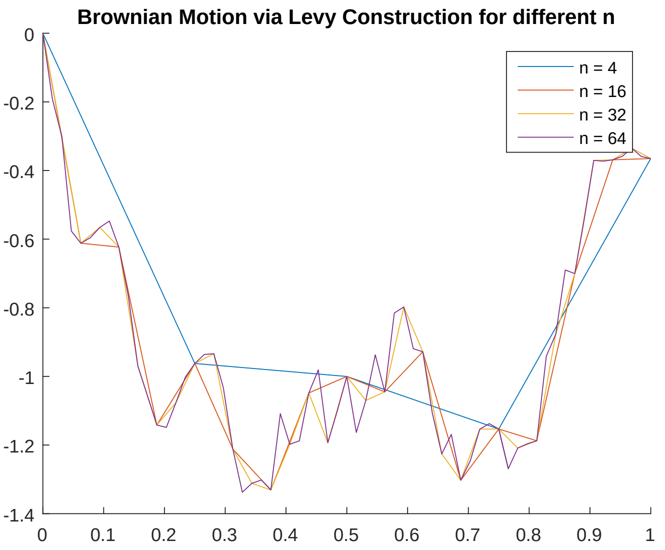
i) κ_X ist ein Hilbert-Schmidt Kern $\surd$
ii) Herleitung der Darstellung bzgl Orthonormalbasis aus i) $\surd$
iii) Nachweis der Eigenschaften der einzelnen Komponenten der Reihe $\surd$
iv) Konvergenzaussagen $\surd$

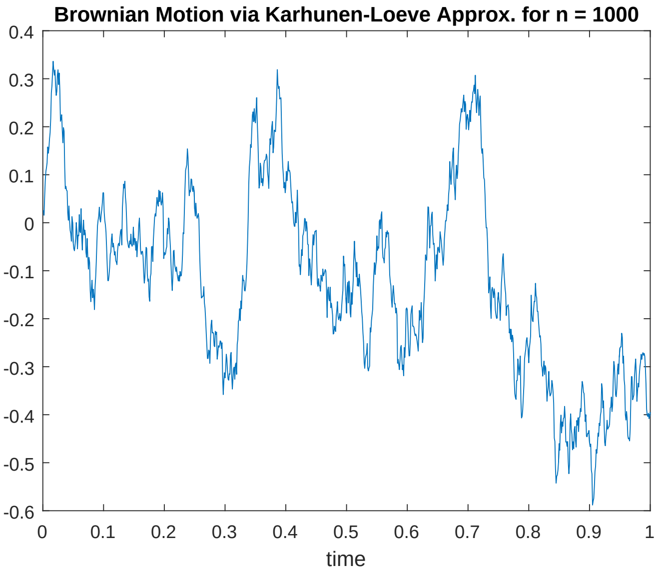
Sei $B = (B_t)_{t\in[0,1]}$ eine **Brownsche Bewegung**. Dann gilt:

$$B_t = \sum_{n=1}^{\infty} Z_n \frac{\sqrt{2}}{\left(n-\frac{1}{2}\right)\pi} sin\left(\left(n-\frac{1}{2}\right)\pi t\right)$$

Für $(Z_t)_{t\in[0,1]}$ u.i.v. $\sim \mathcal{N}(0,1)$.


Brownian Motion via Karhunen-Loeve Approx. for different n
n = 4
n = 16
n = 32
n = 64
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


Brownian Motion via Levy Construction for different n
n = 4
n = 16
n = 32
n = 64
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


Brownian Motion via Karhunen-Loeve Approx. for n = 1000
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
time

Mean squared error

Definition

Sei $(X_t)_{t \in D} \subseteq L^2$ ein stochastischer Prozess und $(\hat{X}_t)_{t \in D} \subseteq L^2$ eine Approximation von X. Dann heißt:

$$MSE(t) := \mathbb{E}[(X - \hat{X})^2]$$

die mittlere quadratische Abweichung von X_t und $\hat{X}_t$. Weiter heißt:

$$TMSE = \int_D MSE(t)dt$$

die totale mittlere quadratische Abweichung.

Satz

Unter allen Reihendarstellungen bezüglich einer Orthonormalbasis minimiert die Karhunen-Loève Zerlegung TMSE.

Satz

Unter allen Reihendarstellungen bezüglich einer Orthonormalbasis minimiert die Karhunen-Loève Zerlegung TMSE.

Definition

Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ *Banachräume über dem selben Körper* $\mathbb{K}$, $U \subseteq X$ *und* $F : U \to Y$. F *heißt Fréchet differenzierbar in* $x \in U$ *falls* $\exists dF(x) \in \mathcal{L}(X, Y)$ *so dass:*

$$\lim_{\substack{\|u\|_X \to 0 \\ u \neq 0}} \frac{\| F(x+u) - F(x) - dF(X)u \|_Y}{\| u \|_X} = 0$$

Falls X *sogar ein Hilbertraum ist und* $Y = \mathbb{R}$, *dann ist* $dF(x)u = < dF(x), u >$.

Varianz des Prozesses

Für alle $t \in D$ gilt:

$$Var[X_t] = \sum_{k=0}^{\infty} \psi_k^2(t) Var[Z_k] = \sum_{k=1}^{\infty} \lambda_k \psi_k^2(t)$$

Gesamte Varianz des Prozesses:

$$Var[X] := \int_D Var[X_t] dt = \sum_{k=1}^{\infty} \lambda_k$$

Varianz der bei $N \in \mathbb{N}$ abgebrochenen Reihe:

$$Var[\hat{X}^{(N)}] = \sum_{k=1}^{N} \lambda_k$$

Betrachte einen diskreten, endlichen Prozess $(X_n)_{n=1,\dots,N}$:

$$\left(T_{K_X}\psi_n\right)(t) = \int_D K_X(s,t)\psi_n(s)ds = \lambda_n\psi_n(t)$$

vereinfacht sich zu:

$$\sum_{i=1}^{N} \Sigma_{ij}\psi_n(i) = \lambda_n\psi_n(j) \quad \Leftrightarrow \quad \Sigma\psi = \lambda\psi$$

Wir müssen also lediglich Eigenwerte und Eigenvektoren von eine $N \times N$ Matrix bestimmen.

...falls noch Zeit ist **Animation.**

Bücher

- MICHEL LOÈVE, Probability Theory
- ISRAEL GOHBERG, Basic Classes of Linear Operators
- DIRK WERNER, Funktionalanalysis
- AMANN/ ESCHER, Analysis II

Papers

- A. ALEXANDERIAN, North Carolina State University, A brief note on the Karhunen-Loève expansion
  http://users.ices.utexas.edu/ alen/articles/KL.pdf
- LIMIN WANG, LSE, Karhunen-Loeve Expansions and their Applications
  http://etheses.lse.ac.uk/2950/1/U615901.pdf