Orthogonalreihendarstellung eines zentrierten Gauß-Prozesses

Thomas Steinle

Seminar Zufällige Felder
Universität Ulm

18. November, 2008

Inhalt

Zur Erinnerung

Ein stochastischer Prozess $\{f(t,\omega) : t \in T, \omega \in \Omega\}$, im folgenden kurz $f(t) := f(t,\cdot)$, heißt **Gauß-Prozess**, falls $\forall\ n \in \mathbb{N}, t_1, ..., t_n \in T$:
$(f(t_1), ..., f(t_n))$ multivariat Normalverteilt ist.

Im Folgenden sei T ein kompakter metrischer Raum.

Zur Erinnerung

Wir nennen den Gauß-Prozess f **zentriert**, falls $\mathbb{E}\{f(t)\} = 0$ für jedes $t \in T$.

Die Funktion

$$Cov(f(s), f(t)) = C(s,t)$$
$$:= \mathbb{E}\{\ (f(s) - \mathbb{E}\{f(s)\})\ (f(t) - \mathbb{E}\{f(t)\})\ \} = \mathbb{E}\{\ f(s)\ f(t)\ \}$$

ist die Kovarianzfunktion des zentrierten Gauß-Prozesses.

Zur Erinnerung

Durch C ist der *zentrierte* Gauß-Prozess eindeutig festgelegt.

Begründung:
Eigenschaft, dass multivariate Normalverteilung durch Erwartungswertvektor und Kovarianzmatrix vollständig bestimmt ist, überträgt sich auf Gauß-Prozess.

Im Folgenden: Kovarianzfunktion immer positiv definit.

Was soll gemacht werden?

Darstellung eines *zentrierten* Gauß-Prozesses $\{f(t) : t \in T\}$ als eine orthogonale Reihe, d.h.

$$f(t) = \sum_{n=0}^{\infty} \xi_n \varphi_n(t).$$

wobei

$\xi_n : \Omega \to \mathbb{R} \sim N(0,1)$ i.i.d,
$\{\varphi_n : T \to \mathbb{R}\}$ Orthonormalbasis eines speziellen Hilbertraumes, der von f abhängt.

Um diese Darstellung herzuleiten ist etwas Funktionalanalysis nötig.

Skalarprodukte

Eine auf einem Vektorraum X über $\mathbb{R}$ gegebene Abbildung $(\cdot,\cdot)_X: \; X \times X \to \mathbb{R}, \quad (x_1, x_2) \mapsto \; (x_1, x_2)_H \; \in \; \mathbb{K}$ heißt **Skalarprodukt** auf X, wenn folgende Regeln gelten:

(S1) $\forall x \in X : (x, x)_X \geq 0; \quad (x, x)_X = 0 \Leftrightarrow x = 0$ (pos. definit)

(S2) $\forall x_1, x_2 \in X : (x_1, x_2)_X = (x_2, x_1)_X$ (symmetrisch)

(S3) $\forall x, x_1, x_2 \in X : (x_1 + x_2, x)_X = (x_1, x)_X + (x_2, x)_X$ (bilinear)

(S4) $\forall x_1, x_2 \in X, \lambda \in \mathbb{K} : (\lambda x_1, x_2)_X = \lambda (x_1, x_2)_X$ (bilinear)

Prä-Hilberträume

Falls ein Skalarprodukt $(\cdot, \cdot)_X$ auf einem Vektorraum X gegeben ist, nennen wir X einen **Prä-Hilbertraum.**

Hilberträume

Ein Prä-Hilbertraum X mit Skalarprodukt $(\cdot, \cdot)_X$ der bezüglich der Norm $\|x\|_X = \sqrt{(x, x)_X}$ vollständig ist, heißt **Hilbertraum.**

(X, d) heißt vollständig, wenn jede Cauchyfolge $(x_n)_{n\in\mathbb{N}}$ in X konvergiert, d.h.

$(\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n, m \geq N : \quad d(x_n, x_m) < \varepsilon)$

$\Rightarrow (\exists \, x \in X \; \forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n \geq N : \quad d(x_n, x) < \varepsilon)$,

wobei x der Grenzwert der Folge (x_n) ist.

Isometrie

Seien (X_1, d_1) und (X_2, d_2) metrische Räume. Dann heißt die Abbildung $\Theta : X_1 \to X_2$ **Isometrie** oder **abstandserhaltend**, falls

$$\forall\, x, y \in X_1 : d_1(x, y) = d_2(\Theta(x), \Theta(y)).$$

Isometrischer Isomorphismus

Jede Isometrie Θ ist injektiv. Ist sie sogar surjektiv, dann nennen wir Θ einen **isometrischen Isomorphismus**. X_1 und X_2 werden als isometrisch isomorph zueinander bezeichnet.

Orthogonalität

Sei X ein Prä-Hilbertraum.

Zwei Vektoren x_1, x_2 heißen **orthogonal**, falls $(x_1, x_2)_X = 0$ ist.

Orthonormalsysteme

Ein System $(x_i)_{i \in I}$ von Vektoren heißt **Orthonormalsystem** (ONS) in X falls keiner der Vektoren der Nullvektor ist, und wenn

$$\forall i, j \in I : \quad (x_i, x_j)_X = \delta_{ij}$$

Bemerkung:
$\|x_i\|_X = \sqrt{(x_i, x_i)_X} = 1, \text{ da } i = j.$

Orthonormalbasen

Ein ONS $(x_i)_{i\in I}$ eines Prähilbertraumes X dessen lineare Hülle $span\{x_i, i \in I\}$ dicht in X liegt, nennt man **vollständiges ONS** oder **Orthonormalbasis** (ONB) von X.

Dichtliegende Teilmengen und lineare Hüllen

Eine Teilmenge A liegt dicht in X, wenn $\overline{A} = X$.

$$span\{x_i, i \in I\} := \{\lambda_{i_1} x_{i_1} + ... + \lambda_{i_n} x_{i_n} : \lambda_{i_k} \in \mathbb{R}, i_k \in I, n \in \mathbb{N}\}$$

Separabilität und Orthonomalbasen

Ein metrischer Raum (X, d) heißt **separabel**, wenn er eine abzählbare dichte Teilmenge besitzt.

Sei X ein unendlich dimensionaler Hilbertraum. Dann gibt es eine abzählbare Orthonormalbasis von X, genau dann wenn X separabel ist.

Beispiel: Separabilität

$\mathbb{Q}$ ist eine abzählbare Teilmenge von $\mathbb{R}$ mit $\overline{\mathbb{Q}} = \mathbb{R}$

Orthogonalreihe

Sei X ein unendlich dimensionaler Prä-Hilbertraum und $(x_k)_{k=0}^{\infty}$ ein abzählbar unendliches Orthonormalsystem in X.

Für beliebige $a_k \in \mathbb{R}$ heißt die Reihe der Form $\sum_{k=0}^{\infty} a_k x_k$ **Orthogonalreihe** mit Koeffizienten a_k.

Fourierreihe

Sei X ein Hilbertraum und $x \in X$, sowie $(x_k)_{k\in\mathbb{N}}$ ONB von X. Dann gilt:

$$x = \sum_{k=0}^{\infty} (x, x_k)_X x_k$$

Dies ist die sog. (verallgemeinerte) **Fourierreihe**.

Was soll gemacht werden?

Darstellung eines *zentrierten* Gauß-Prozesses $\{f(t) : t \in T\}$ als eine orthogonale Reihe, d.h.

$$f(t) = \sum_{n=0}^{\infty} \xi_n \varphi_n(t).$$

wobei

$\xi_n : \Omega \to \mathbb{R} \sim N(0,1)$ i.i.d,
$\{\varphi_n : T \to \mathbb{R}\}$ ONB eines speziellen Hilbertraumes der von f abhängt.

Zu diesem speziellen Hilbertraum:

Konstruieren den sogenannten **Hilbert Raum mit reproduzierendem Kern** des zentrierten Gauß-Prozesses, kurz RKHS von f (engl. reproducing kernel Hilbert space).

RKHS

Sei H ein Hilbertraum von Funktionen auf einer nichtleeren Menge T, dann sagen wir, dass H einen **reproduzierenden Kern** besitzt, falls es eine Funktion $C : T \times T \to \mathbb{R}, (s,t) \mapsto C(s,t)$ mit folgenden Eigenschaften gibt:

(i) $\forall\, t \in T : C(t,\cdot) \in H$
(ii) $\forall\, u \in H\, \forall t \in T : u(t) = (u, C(t,\cdot))_H$

(ii) ist die **Reproduktionseigenschaft** des Kerns, d.h. jede Funktion u wird unter dem Skalarprodukt mit dem Kern reproduziert.

Inhalt

Hilfsmenge S

Wir betrachten die Kovarianzfunktion C für jeweils beliebige aber feste $s_i \in T$ als Funktion in der zweiten Variablen, d.h. $C(s_i, \cdot)$.

Die Menge aller möglichen Linearkombinationen der Kovarianzfunktionen an den Stellen s_i sei

$$S = \left\{ u : T \to \mathbb{R} : u(\cdot) = \sum_{i=1}^{n} a_i C(s_i, \cdot), a_i \in \mathbb{R}, s_i \in T, n \geq 1 \right\}$$

Beispiel für diese Hilfsmenge

Sei $T := \{1, ...N\}$ und f zentrierter Gauß-Prozess auf T mit einer positiv definiten Kovarianzfunktion, welche als positiv definite Kovarianzmatrix $C = (c_{ij})$ angegeben sei.

Der Gauß-Prozess ist also eine multivariate Normalverteilung.

Wie sehen die $u \in S$ hier aus?

O.B.d.A. sei $n \geq N$.
Dann gilt für beliebige $a_i \in \mathbb{R}, s_i \in \{1, ..., n\}, i \in \{1, ..., n\}, n \geq 1$:

$$u(t) \left(\underset{\text{u(t)}}{\overset{\text{Def.}}{=}} \sum_{i=1}^{n} a_i \; C(s_i, t) \right) = \sum_{j=1}^{N} \tilde{a}_j \; c_{jt} = \tilde{a} \; c_t \quad ,$$

wobei $\tilde{a} = (\tilde{a}_1, ..., \tilde{a}_N)$ mit $\tilde{a}_j := \sum_{i: s_i = j} a_i$ und
c_t die t-te Spalte der Matrix C ist.

Wie sieht dann S hier aus?

Es findet sich immer ein $u \in S$, dass eine beliebig vorgegebene reelle Zahl als Wert besitzt, d.h.

$$\forall x \in \mathbb{R} : \exists\, u \in S, t \in T : u(t) = a\, c_t = x.$$

Daher ist:

$$S = \{u : \{1, ..., N\} \to \mathbb{R}\}$$

Ein $u \in S$ nimmt maximal N verschiedene reelle Werte an:

Definiere $u := (u_1, ..., u_N)$ mit $u(1) = u_1,\ ...\ , u(N) = u_N$.

Dann kann man sich S auch als den $\mathbb{R}^N$ vorstellen.

Zurück zur Theorie: Skalarprodukt auf S

Für eine positiv definite Kovarianzfunktion, ist

$$(u, v)_H = \Big(\sum_{i=1}^{n} a_i C(s_i, \cdot), \sum_{j=1}^{m} b_j C(t_j, \cdot) \Big)_H := \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j C(s_i, t_j).$$

ein Skalarprodukt auf S.

Kernreproduktionseigenschaft dieses Skalarproduktes

Für alle $u \in S$:

$$(u, C(t, \cdot))_H = u(t).$$

Jede Funktion u wird unter dem Skalarprodukt mit der Kovarianzfunktion (= Kern) reproduziert.

Beweis:

$$(u, C(t, \cdot))_H \overset{\text{Def.}}{\underset{\text{u}}{=}} \Big(\sum_{i=1}^{n} a_i C(s_i, \cdot), C(t, \cdot) \Big)_H \overset{\text{Def.}}{\underset{\text{SP}}{=}} \sum_{i=1}^{n} a_i C(s_i, t) \overset{\text{Def.}}{\underset{\text{u}}{=}} u(t).$$

Zurück zum Beispiel auf $T = \{1, ..., N\}$

Kernreproduktion:

Da die Kovarianzmatrix C positiv definit ist, existiert deren Inverse C^{-1}.

$$(u, v)_H := u\ C^{-1}\ v.$$

ist Skalarprodukt mit Kernreproduktionseigenschaft auf $S = \mathbb{R}^N$.

Beweis der Kernreproduktionseigenschaft:

Sei c_k die k-te Spalte der Matrix C und $u = (u_1, ..., u_N)$. Dann gilt:

$$(u, c_k)_H \underset{\text{SP}}{\overset{\text{Def.}}{=}} u\ C^{-1}\ c_k = u\ e_k = u_k$$

Anschauung:

Kernreproduktion bedeutet hier, dass die k-te Spalte der Kovarianzmatix unter dem Skalarprodukt mit u die k-te Komponente von u reproduziert.

Aufstellen des RKHS

$H(C) := \overline{S}$ mit dem Skalarprodukt $(\cdot,\cdot)_H$ ist ein Prä-Hilbertraum.

Es kann gezeigt werden, dass $H(C)$ bezüglich der Norm $\|u\|_H := \sqrt{(u,u)_H}$ vollständig ist.

Dann folgt, dass es sich bei $H(C)$ um einen Hilbertraum handelt.

Da das Skalarprodukt die Kernreproduktionseigenschaft besitzt, nennen wir $H(C)$ den von einem zentrierten Gauß-Prozesses f abhängigen **Hilbertraum mit reproduzierendem Kern**, kurz RKHS von f**.**

Bemerkung:

Man spricht vom RKHS **von** f, da dieser vom Gauß-Prozess wie folgt abhängt:

- Zentrierter Gauß-Prozess eindeutig durch C bestimmt.
- C bestimmt RKHS eindeutig.

Im folgenden werden wir eine Orthogonalreihendarstellung eines zentrierten Gauß-Prozesses herleiten.

Genauer gesagt, werden wir folgendes Theorem beweisen:

Orthogonalreihendarstellung des Gauß-Prozesses

Theorem:
f sei ein zentrierter Gauß-Prozess,
$\{\varphi_k : T \to \mathbb{R}\}$ sei eine ONB des $H(C)$ von f und
$\xi_k : \Omega \to \mathbb{R}$ seien unabh. standardnormalverteilte ZV.
Dann gilt:

$$f(t) = \sum_{k=1}^{\infty} \xi_k \, \varphi_k(t) \quad ,$$

Beweis:
an Tafel.

Inhalt

Orthogonalreihe von f auf $T = [0,1]^N$

Unser Ziel ist es, eine Orthogonalreihendarstellung für einen zentrierten Gauß-Prozess f auf einem kompakten $T \subset \mathbb{R}^N$ zu finden.

Im Folgenden beschränken wir uns auf $[0,1]^N$.

Das bedeutet:
Finde eine Orthonormalbasis $(\varphi_n)_{n\in\mathbb{N}}$ des $H(C)$ von f.

Dies führt dann zur sog. Karhunen-Loève Reihendarstellung.

ONB des $H(C)$ von f auf $[0,1]^N$

Man kann zeigen, dass folgendes System eine ONB des $H(C)$ von f ist:

$$(\sqrt{\lambda_n}\psi_n)_{n\in\mathbb{N}} \ ,$$

wobei die λ_n Eigenwerte und die ψ_n normierte Eigenfunktionen des Operators C

$$(\mathrm{C}\ \psi)(t) = \int_T C(s,t)\psi(s)ds$$

sind.

Karhunen-Loève Reihe

Insgesamt führt dies zur sogenannten Karhunen-Loève Reihe:

$$f(t) = \sum_{n=1}^{\infty} \xi_n \, \varphi_n(t) = \sum_{n=1}^{\infty} \xi_n \, \sqrt{\lambda_n} \, \psi_n(t) \, .$$

Im Allgemeinen ist es sehr schwierig diese analytische zu bestimmen.

Jedoch kann die Karhunen-Loève Reihe der Standard Brownschen Bewegung analytisch bestimmt werden.

Standard Brownsche Bewegung

Zur Erinnerung:
Die Standard Brownsche Bewegung $f = W$ auf $T = [0, 1]$ ist ein zentrierter Gauß-Prozess mit Kovarianzfunktion

$$Cov(W_s, W_t) = C(s, t) = min(s, t)$$

für alle $s, t \in T = [0, 1]$.

Karhunen-Loève Reihe der Standard Brownschen Bewegung

Der Operator C sieht hier folgendermaßen aus:

$$(\mathrm{C}\psi)(t) = \int_0^1 min(s,t)\psi(s)ds$$

Wir integrieren also $min(s,t)$ für ein festes $t \in [0,1]$. Daher gilt:

$$(\mathrm{C}\psi)(t) = \int_0^t s\psi(s)ds + t\int_t^1 \psi(s)ds.$$

Es muss also folgendes Eigenwertproblem gelöst werden:

$$\lambda\psi(t) = \int_0^t s\psi(s)ds + t\int_t^1 \psi(s)ds.$$

Diffenzieren nach t auf beiden Seiten liefert uns die Diffentialgleichungen

$\lambda\dot{\psi}(t) = \int_t^1 \psi(s)ds$,

$\lambda\ddot{\psi}(t) = -\psi(t)$

mit den Randwerten $\psi(0) = 0$ und $\dot{\psi}(1) = 0$.

Die Lösungen dieses Randwertproblems sind:

$$\psi_n(t) = \sqrt{2}\sin\left(\frac{1}{2}(2n+1)\pi t\right), \quad \lambda_n = \left(\frac{2}{(2n+1)\pi}\right)^2.$$

Somit sieht die Karhunen-Loève Reihe der Standard Brownschen Bewegung $f(t) = W_t$ folgendermaßen aus:

$$W_t = \frac{\sqrt{2}}{\pi} \sum_{n=0}^{\infty} \xi_n \left(\frac{2}{2n+1} \right) \sin \left(\frac{1}{2}(2n+1)\pi t \right).$$

Zusammenfassung

Jeder zentrierte Gauß-Prozess lässt sich als Orthogonalreihe darstellen mit

Koeffizienten $\xi_n \sim N(0,1)$ *iid* und

einer ONB $\{\varphi_n\}$ des von f abhängigen Hilbertraumes.

Für zentrierte Gauß-Prosse auf $[0, 1]^N$ können wir eine Karhunen-Loève Reihenentwicklung durchführen, welche das Lösen eines Eigenwertproblems beinhaltet.

Für die Standard Brownschen Bewegung kann dies analytisch gelöst werden. Für die meisten anderen Beispiele ist dies jedoch nicht möglich.

Literatur

R. Adler und J. Taylor
Random Fields And Geometry
Springer, Berlin, 2007

W. Balser
Vorlesungsmanuskript zur Funktionalanalysis
Wintersemester 2006/2007

D. Werner
Funktionalanalysis
Springer, 2005