

確率数理要論 8

● Gauss 過程

Def

\mathbb{R}^m 値確率変数 (X_1, \dots, X_m) が多変量正規分布に従うとは、 $\forall t_1, \dots, t_m \in \mathbb{R}$ に対して

$$t_1 X_1 + t_2 X_2 + \dots + t_m X_m$$

が \mathbb{R} 上の正規分布に従うことと定義する。

ただし、分散 0 でも正規分布であると言う。

平均: $\mu = (E[X_1], \dots, E[X_m])^T \in \mathbb{R}^m$

共分散: $\Sigma = (E[X_i X_j] - E[X_i]E[X_j])_{i,j} \in \mathbb{R}^{m \times m}$

$\text{Cov}(X_i, X_j)$

を用いて、 $N(\mu, \Sigma)$ と書く。

//

※ この定義から Σ がランク落ちしてもよい。

(通常の密度関数を用いた定義では扱えない。)

実際、 $\Sigma \succ 0$ の際には、密度関数を用いた定義と同値であることをチェックする。

$\mu = 0$ としなくても一般性を失わない。

$t_1 X_1 + \dots + t_m X_m$ の平均は 0、分散は $t^T \Sigma t$ である。

2次元正規分布に従うので、その特性関数は

$$\phi_t(\beta) = E[e^{i\beta(t_1 X_1 + \dots + t_m X_m)}] = \exp\left(-\frac{1}{2} \beta^2 (t^T \Sigma t)\right) \quad (*)$$

である。 $\beta = 1$ とすると、 $\phi_t(1)$ は $(X_1, \dots, X_m)^T$ の特性関数に他ならない。

実際、 $\phi(t) = E[e^{i t^T X}] = E[e^{i(t_1 X_1 + \dots + t_m X_m)}]$

$$= \exp\left(-\frac{1}{2} t^T \Sigma t\right) \quad (:\text{ } *)$$

これは平均 0、分散共分散 Σ の多変量正規分布の特性関数。

特性関数の一意性より、両者の定義が同値であることを示す。

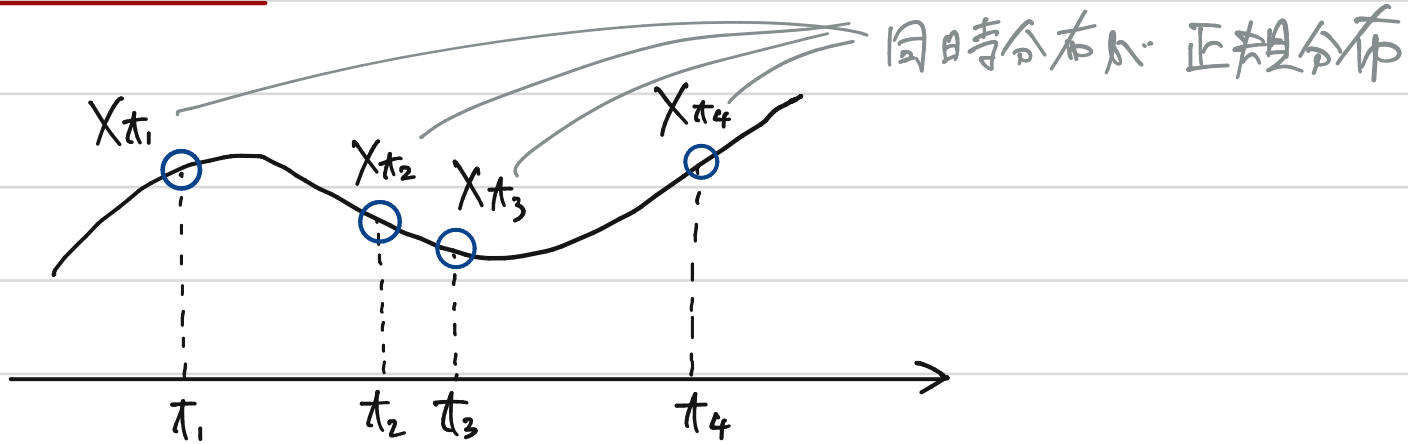
Def (Gauss過程, Gaussian Process)

T : インデックスの集合 ($T = \mathbb{R}$ とは限らな u . 一般の集合)

確率過程 $X = (X_t)_{t \in T}$ が任意の $t_1, \dots, t_n \in T$ に対し、

$(X_{t_1}, \dots, X_{t_n})$ が多変量正規分布に従う時、

Gauss過程 と言う。



$m(t) = E[X_t]$: 平均関数

$k(t, t') = E[(X_t - m(t))(X_{t'} - m(t'))]$: 共分散関数

$GP(m, k)$ と書く。

- 再生核ヒルベルト空間

共分散関数 k は以下の性質を持つ:

- ・ 対称: $k(t, s) = k(s, t)$
- ・ 正定値: $\forall \alpha_1, \dots, \alpha_m \in \mathbb{R}, \forall t_1, \dots, t_m \in T,$
$$\sum_{1 \leq i, j \leq m} \alpha_i \alpha_j k(t_i, t_j) \geq 0$$

この2つの性質を持つ関数 $k: T \times T \rightarrow \mathbb{R}$ を 正定値カーネル とする。

Def (再生核ヒルベルト空間, Reproducing kernel Hilbert space, RKHS)

集合 T 上の RKHS \mathcal{H} は、 T 上の (実数値) 関数からなるヒルベルト空間で、 $\forall t \in T$ に対し $k_t \in \mathcal{H}$ が存在して

$$(\text{再生性}) \quad \langle f, k_t \rangle_{\mathcal{H}} = f(t) \quad (\forall f \in \mathcal{H})$$

が成り立つものとする。

特に、 $k(x, y) = \langle k_x, k_y \rangle_{\mathcal{H}}$ を \mathcal{H} に付随した再生核 (Reproducing kernel) と呼ぶ。 //

※ $k(x, y)$ は正定値カーネルになっている。

つまり、RKHS が 1) と 2) なら、1) 正定値カーネルが定まる。

実は、以下のよりの逆も成り立つ。

Thm (Moore - Aronszajn)

$k: T \times T \rightarrow \mathbb{R}$ を正定値カーネルとする。

\Rightarrow k を再生核とする RKHS が一意に存在する。 //

(略証) $H_0 = \{ f(\cdot) = \sum_{j=1}^m d_j k(\cdot, t_j) \mid d_j \in \mathbb{R}, t_j \in T, m \in \mathbb{N} \}$

とし、

H_0 の元 $f = \sum_{j=1}^m d_j k(t_j, \cdot)$, $g = \sum_{j=1}^{m'} \beta_j k(t'_j, \cdot)$ に対し、

$$\langle f, g \rangle_{H_0} = \sum_{i=1}^m \sum_{j=1}^{m'} d_i \beta_j k(t_i, t'_j)$$

と“内積”を定める。

この内積から定まるノルムで H_0 を完備化したものを H とする。

つまり、 H は H_0 内のコーシー列の集合で、

H_0 内のコーシー列 $f = (f_n)_n$, $g = (g_n)_n$ に対し、

$$\langle f, g \rangle_H := \lim_{n \rightarrow \infty} \langle f_n, g_n \rangle_{H_0}$$

とし、 $\langle f-g, f-g \rangle_H = 0$ である2つのコーシー列は同一視する。

$f \in H$ に対し、 f に対応するコーシー列 $(f_n)_n$ を用いて、

$$f(t) = \lim_{n \rightarrow \infty} \langle f_n, k(t, \cdot) \rangle_{H_0}$$

とすれば、 $f \in H$ は T 上の実数値関数とみなせる。

実は、この H は $\langle \cdot, \cdot \rangle_H$ を内積として持つヒルベルト空間になっていることを示せる。

さらに、 H_0 内のコーシー列 $(k_{t_n})_{n=1}^{\infty}$ も $k_t \in H$ と書くと、

H は k_t を再生核とする RKHS になっている。 //

$$\begin{array}{ccc} k & \overset{\text{一対一対応}}{\longleftrightarrow} & H \\ \text{(正定値カーネル)} & & \text{(RKHS)} \end{array}$$

Thm

ある実数値関数のヒルベルト空間 H を考える。

$\forall t \in T$ における $f \mapsto f(t)$ が H のノルムによる連続

$\Leftrightarrow H$ は RKHS. //

(証明は略, Rieszの表現定理より従う)

Thm (Mercerの定理 (一般形))

T : ハウスドルフ位相空間

ν : T 上の非負 Borel 測度で $\text{supp}(\nu) = T$ とする.

"Borel 測度" の
定義は全実数に
対して
成り立つ

ただし、全てのコンパクトな $A \in \mathcal{B}(T)$ には $\nu(A) < \infty$ とする.

$K: T \times T \rightarrow \mathbb{R}$ は 連続な正定値カーネル である.

$$\int K(x, x) d\nu(x) < \infty$$

とする.

このとき、 K によって決まる RKHS \mathcal{H} は可分である.

$$\begin{cases} \cdot \mu_1 \geq \mu_2 \geq \dots \geq 0 \\ \cdot (e_n)_n: L^2(\nu) \text{ 内の ONB (正規直交系)} \end{cases}$$

が存在し、 $\forall x \in T$ には

$$K(x, x') = \sum_{n=1}^{\infty} \mu_n e_n(x) e_n(x') \quad (\text{Mercer 展開})$$

が $\forall x, x' \in T$ で絶対収束する.

さらに、任意のコンパクト集合 $A \subset T$ に対し、 $A \times A$ 上で x, x' に関して一様収束する. //

元々は、Mercer (1909) による。 $T = [a, b]$ で示された。

上記の一般化に関する。

Steinwart & Scovel: Mercer's theorem on general domains:

On the interaction between measures, kernels and RKHSs.

Constructive Approximation, 35:363-417, 2012.

Mercer 展開を用いる。

$$f \in \mathcal{H} \text{ は } \sum_{n=1}^{\infty} d_n^2 < \infty \text{ とある } (d_n)_n \text{ を用いる}$$

$$f = \sum_{n=1}^{\infty} d_n \sqrt{\mu_n} e_n$$

と与えられる。

$$f = \sum_{n=1}^{\infty} d_n \sqrt{\mu_n} e_n, g = \sum_{n=1}^{\infty} \beta_n \sqrt{\mu_n} e_n \quad k \neq l$$

$$\langle f, g \rangle_{\mathcal{H}} = \sum_{n=1}^{\infty} d_n \beta_n$$

とある。

Thm (Karhunen - Loève 展開)

- 正定値カーネル $k: T \times T \rightarrow \mathbb{R}$ が Mercer の定理の条件を満たして与えられる。この時、確率過程

$$\underline{X_t = \sum_{\lambda=1}^{\infty} \xi_{\lambda} \sqrt{\mu_{\lambda}} e_{\lambda}(t)}$$

← RKHS の π の展開と比較せよ。

$$\xi_{\lambda} \sim N(0, 1) \quad (\text{i.i.d.})$$

は $GP(0, k)$ になる。

- 逆に、 $(X_t)_t$ が $GP(0, k)$ であるとす。

$$\xi_{\lambda}(\omega) = \frac{1}{\sqrt{\mu_{\lambda}}} \int X_t(\omega) e_{\lambda}(t) dV(t)$$

とすると、 $(\xi_{\lambda})_{\lambda}$ は独立で $\xi_{\lambda} \sim N(0, 1)$ かつ $E[\xi_{\lambda} \xi_j] = \delta_{\lambda j}$ である。
任意の有限な I と任意の $t \in T$ に対し、

$$E \left[\left(X_t - \sum_{\lambda \in I} \xi_{\lambda} \sqrt{\mu_{\lambda}} e_{\lambda}(t) \right)^2 \right] = k(t, t) - \sum_{\lambda \in I} \mu_{\lambda} e_{\lambda}^2(t)$$

であり、特に

$$\int E \left[\left(X_t - \sum_{\lambda \in I} \xi_{\lambda} \sqrt{\mu_{\lambda}} e_{\lambda}(t) \right)^2 \right] dV(t) = \sum_{\lambda \notin I} \mu_{\lambda}$$

- KL-展開は Gauss 過程の具体的な構成法を与えてくれる。
- 2つめの主張より $X_t = \sum_{\lambda} \xi_{\lambda} \sqrt{\mu_{\lambda}} e_{\lambda}(t)$ なる展開は $L^2(P \times V)$ 内の収束という意味で成り立つ。

確認

$$\begin{aligned} E[X_t X_{t'}] &= E \left[\sum_{\lambda=1}^{\infty} \sum_{j=1}^{\infty} \xi_{\lambda} \xi_j \sqrt{\mu_{\lambda}} \sqrt{\mu_j} e_{\lambda}(t) e_j(t') \right] \\ &= \sum_{\lambda=1}^{\infty} \sum_{j=1}^{\infty} \underbrace{E[\xi_{\lambda} \xi_j]}_{\delta_{\lambda j}} \sqrt{\mu_{\lambda}} \sqrt{\mu_j} e_{\lambda}(t) e_j(t') \\ &= \sum_{\lambda=1}^{\infty} \mu_{\lambda} e_{\lambda}(t) e_{\lambda}(t') = k(t, t') \end{aligned}$$

Thm (Driscollの定理)

H が無限次元 \Rightarrow 確率1で $X \notin H$

H が有限次元 \Rightarrow (X のあるバージョン \tilde{X} が存在して)

確率1で $\tilde{X} \in H$

(確認)

KL-展開を用いて

$$\|X\|_H^2 = \sum_{i=1}^{\infty} \xi_i^2 \mathbb{1}(\mu_i \neq 0)$$

と書ける.

無限次元 : $\forall n \exists \mu_n \neq 0$ となる.

無限

$$\sum_{i=1}^{\infty} \xi_i^2 = \infty \quad (\text{a.s.})$$

有限次元 : $\exists N \forall n > N \exists \mu_n = 0$ となる.

有限

$$\sum_{i=1}^N \xi_i^2 < \infty \quad (\text{a.s.})$$

以下、 $T = \mathbb{R}_+$ の状況を考へる。伊藤積分に話を向けておく。

Def (Brown 運動, Brownian motion, Wiener process)

確率過程 $B = (B_t)_{t \geq 0}$ が次をみたるとき、

(標準) Brown 運動 (standard Brownian motion) とする。

(1) B は Gauss 過程

(2) $\forall t, s \in [0, \infty)$, $E[B_t] = 0$, $E[B_t B_s] = t \wedge s$

(3) a.s. $\omega \in \Omega$ に対し、 $t \mapsto B_t(\omega)$ は連続

(連続なパスを持つ) とする。 //

Cor

B を Brown 運動とする。

(1)' $B_0 = 0$ a.s.

(2)' B_t は加法過程

(3)' $\forall t \geq 0, \forall s > 0$, $B_{t+s} - B_t \sim N(0, s)$

特に、Brown 運動は Lévy 過程になる。

(1)' ~ (3)' と (3) を Brown 運動の定義としても良い。

Proof

(1)' $E[B_0^2] = 0$ より、O.K.

(2)' $0 \leq a < b \leq c < d$ に対し、

$$E[(B_d - B_c)(B_b - B_a)]$$

$$= d \wedge b - c \wedge b - d \wedge a + c \wedge a$$

$$= b - b - a + a = 0$$

より、 $B_d - B_c$ と $B_b - B_a$ は独立。

(共分散 0 の正規分布は独立)

一般の n についても同様。

(3)' $B_{t+s} - B_t$ が平均 0 の正規分布に従うことは明らか。

$$E[(B_{t+s} - B_t)^2] = (t+s) \wedge (t+s) - 2(t+s) \wedge t + t \wedge t$$

$$= t+s - 2t + t = s$$
 //

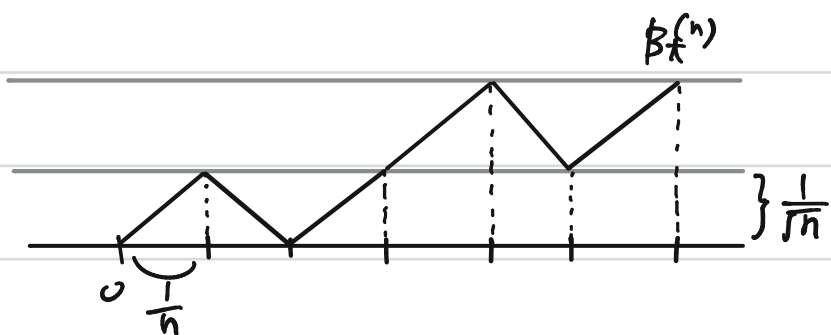
B_t の直感的な構成

X_1, X_2, \dots は i.i.d. で

$$P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$$

とする。

$$B_t^{(n)} := \frac{1}{\sqrt{n}} \left(\sum_{k=1}^{\lfloor nt \rfloor} X_k + \left(t - \frac{\lfloor nt \rfloor}{n} \right) \cdot X_{\lfloor nt \rfloor + 1} \right) \quad (0 \leq t \leq 1)$$



(2)' は $n \rightarrow \infty$ で $X_{\lfloor nt \rfloor + 1}$ 項は無視できると成り立つことを示す。

$$(3)' \text{ は } B_{t+s}^{(n)} - B_t^{(n)} = \frac{1}{\sqrt{n}} \left(\sum_{k=\lfloor nt \rfloor + 1}^{\lfloor n(t+s) \rfloor} X_k + O(1) \right)$$

← 合計 $n \cdot s$ 個

$$\rightsquigarrow N(0, s) \quad (\because \text{中心極限定理})$$

から示す。

(厳密に扱うには $B_t^{(n)}$ が確率過程として B_t に収束するを示す (連続関数空間上の分布収束))

• $(B_t)_{t \in [0,1]}$ の KL-展開

$$k(t,s) = t \wedge s \quad (t, s \in [0,1])$$

より、

$$\int_0^1 k(t,s) e(s) ds = \mu e(t)$$

を解けば良い。

$$\text{左辺} = \int_0^1 (t \wedge s) e(s) ds \text{ に対し、両辺 } t \text{ について微分すると}$$

$$\int_0^1 \mathbb{1}[s \geq t] e(s) ds = \mu e'(t)$$

$$\Rightarrow -e(t) = \mu e''(t)$$

← エルミット多項式

これを解くと、

$$\begin{cases} \mu_n = \frac{4}{(2n-1)^2 \pi^2} \\ e_n(t) = \sqrt{2} \sin\left(\frac{(2n-1)\pi}{2} t\right) \end{cases} \quad (n=1, 2, \dots)$$

が固有値固有関数になる。

KL-展開より、

$$\begin{aligned} B_t &= \sum_{n=1}^{\infty} \sqrt{\mu_n} e_n(t) \xi_n \\ &= \sum_{n=1}^{\infty} \frac{2\sqrt{2}}{\pi(2n-1)} \sin\left(\frac{(2n-1)\pi}{2} t\right) \xi_n \end{aligned}$$

← 各周波数成分が
独立(合計±2σ²)
である。

(ただし、 $\xi_n \sim N(0, 1)$ (i.i.d.))

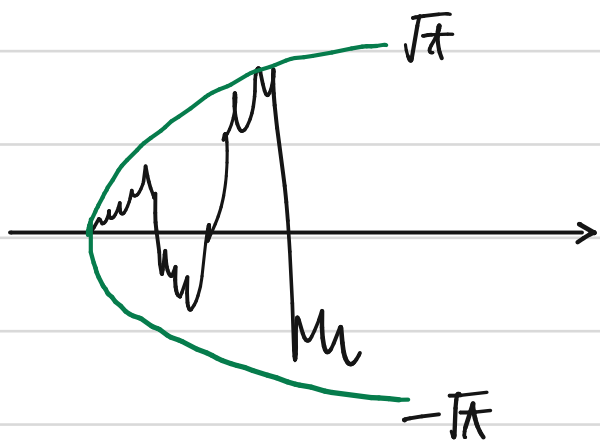
なる表現を持つ。

Brown運動の性質

- a.s. $\omega \in \Omega$ に対し、ほとんど全ての t で微分不可能
- $\lim_{\delta \rightarrow 0} \sup_{\substack{|t-s| < \delta \\ t, s \in [0, 1]}} \frac{|B_t - B_s|}{|t-s|^\varepsilon} = \begin{cases} 0 & (0 < \varepsilon < \frac{1}{2}) \\ \infty & (\varepsilon \geq \frac{1}{2}) \end{cases} \quad \text{a.s.}$

→ B_t は大体 \sqrt{t} の大さきで振動する。

(Karatzas and Shreve Thm 2.2.8, Thm 2.9.25)



Cor

$(B_t)_{t \geq 0}$ は Brown 運動である。

すると、以下は全 2 Brown 運動に付き:

(1) $X_t = -B_t$

(2) $X_t = B_{t+c} - B_c \quad (c > 0)$

(3) $X_t = \sqrt{c} B_{t/c}$

(証明は略) ← 42... せよ.

Thm (Lévy - Itô の分解)

Lévy 過程 $(X_t)_{t \geq 0}$ に対し,

- 連続関数 $m(t)$ s.t. $m(0) = 0$

- 連続・単調非減少関数 $v(t)$ s.t. $v(0) = 0$

- Poisson random measure N s.t.

平均測度 μ が:

$$\mu(\{t\} \times (\mathbb{R} - \{0\})) = 0,$$

$$\int_{s \leq t} \int_{|u| > 0} (u^2 \wedge 1) \mu(ds du) < \infty$$

をみたす。

- N とは独立な Brown 運動 $(B_t)_{t \geq 0}$

が存在し.

$$X_t = m(t) + B_{v(t)}$$

← かつ 2 過程の部分

$$+ \lim_{n \rightarrow \infty} \iint_{\substack{0 \leq s \leq t \\ |u| \geq \frac{1}{n}} (u N(ds du) - \frac{u}{1+u^2} \mu(ds du))$$

と書ける.

Poisson 過程の部分

Lévy 過程は大體、Brown 運動と Poisson random process と書ける (と)

こと.

(伊藤清, 確率論, 定理 5.29)