

カーネル法の理論

福水 健次 (統計数理研究所)

2006年7月6~7日

公開講座「カーネル法の最前線

— SVM, 非線形データ解析, 構造化データ —」

4. カーネル法の理論的基礎

- このセクションの目的
 - 正定値性を保つ演算と正定値性の証明
 - 無限次元の問題を有限次元へ還元: Representer定理
 - 正定値カーネルの固有展開表現: Mercerの定理
 - R^m 上の正定値カーネルの特徴づけ: Bochnerの定理

正定値カーネルの導出

■ カーネル法

- 関数空間でのデータ解析
- 関数空間を陽に与えるのではなく、正定値カーネルを与えることが多い

■ 正定値カーネルの定義法

- 正定値性の判定法は？
- ある正定値カーネルから、新たな正定値カーネルを導出する方法は？
- データの構造に即したカーネル → 赤穂、松井

正定値性を保つ演算

■ 復習 (正定値カーネル)

(1) $k(x, y) = k(y, x)$; (2) グラム行列 ($k(x_i, x_j)$) は半正定値

■ 和と積

命題 $k_1(x, y), k_2(x, y) : \Omega$ 上の正定値カーネル. 次も正定値カーネル

(1) 非負実数 a_1 と a_2 に対する線形和 $a_1 k_1(x, y) + a_2 k_2(x, y)$

(2) 積 $k_1(x, y)k_2(x, y)$

証明) 和は自明. 積については, まず k_1 のグラム行列を対角化

$$k_1(x_i, x_j) = \sum_{p=1}^n \lambda_p U_p^i U_p^j \quad (\text{半正定値性より } \lambda_p \geq 0)$$

$$\begin{aligned} \sum_{i,j=1}^n c_i c_j k_1(x_i, x_j) k_2(x_i, x_j) &= \sum_{p=1}^n \sum_{i,j=1}^n c_i c_j \lambda_p U_p^i U_p^j k_2(x_i, x_j) \\ &= \lambda_1 \left(\sum_{i,j=1}^n c_i U_1^i c_j U_1^j k_2(x_i, x_j) \right) + \cdots + \lambda_n \left(\sum_{i,j=1}^n c_i U_n^i c_j U_n^j k_2(x_i, x_j) \right) \geq 0 \end{aligned}$$

■ 正定値カーネルの収束列

命題 $k_1(x, y), k_2(x, y), \dots, k_n(x, y), \dots : \Omega$ 上の正定値カーネルの列.

$$k(x, y) = \lim_{n \rightarrow \infty} k_n(x, y) \quad (\text{任意の } x, y \in \Omega)$$

ならば, $k(x, y)$ も正定値カーネル

■ 正定値カーネルのなす凸錐

ある集合 Ω 上の正定値カーネル全体は, 積について閉じた
(各点位相での)閉凸錐

■ 正規化

□ 関数の積 $k(x, y) = f(x)f(y)$ の形のカーネルは正定値

□ $k(x, y) : \Omega$ 上の正定値カーネル, $f : \Omega \rightarrow \mathbf{R} : 任意の関数$ とすると,

$$\tilde{k}(x, y) = f(x)k(x, y)f(y)$$

は正定値

□ 特に, 正定値カーネルが $k(x, x) > 0$ ($x \in \Omega$ は任意) を満たすとき,

$$\tilde{k}(x, y) = \frac{k(x, y)}{\sqrt{k(x, x)k(y, y)}} = \frac{\langle \Phi(x), \Phi(y) \rangle}{\|\Phi(x)\|_{H_k} \|\Phi(y)\|_{H_k}}$$

は正定値カーネル ... **normalized** カーネル

例)

$$k(x, y) = (x^T y + c)^d \quad (c > 0) \quad \Longrightarrow \quad \tilde{k}(x, y) = \frac{(x^T y + c)^d}{(x^T x + c)^{d/2} (y^T y + c)^{d/2}}$$

正定値カーネルの例：再論（証明）

■ 多項式カーネル

$$x^T y \text{ は正定値} \quad \because) \quad \sum_{i,j=1}^n c_i c_j x_i^T x_j = \left(\sum_{i=1}^n c_i x_i \right)^T \left(\sum_{j=1}^n c_j x_j \right) \geq 0$$

$$\Rightarrow x^T y + c \text{ は正定値} \quad (c \geq 0)$$

$$\Rightarrow (x^T y + c)^d \text{ は正定値} \quad (d \text{ 個の積})$$

■ ガウスカーネル

$$x^T y \text{ は正定値}$$

$$\Rightarrow e^{\frac{1}{\sigma^2} x^T y} = 1 + \frac{1}{\sigma^2} x^T y + \frac{1}{2!} \left(\frac{1}{\sigma^2} x^T y \right)^2 + \frac{1}{3!} \left(\frac{1}{\sigma^2} x^T y \right)^3 + \dots \quad \text{は正定値}$$

$$\Rightarrow e^{-\frac{1}{2\sigma^2} x^T x} e^{\frac{1}{\sigma^2} x^T y} e^{-\frac{1}{2\sigma^2} y^T y} = \exp\left(-\frac{1}{\sigma^2} \|x - y\|^2\right) \quad \text{は正定値}$$

負定値カーネル

Ω : 集合.

$h(x,y) : \Omega \times \Omega \rightarrow \mathbf{R}$ が負定値であるとは, 次の2つを満たすことをいう

1. (対称性) $h(x,y) = h(y,x)$
2. (負定値性) 任意の自然数 n と, 任意の Ω の点 x_1, \dots, x_n と

$$c_1 + \dots + c_n = 0$$

を満たす任意の実数 c_1, \dots, c_n に対し

$$\sum_{i,j=1}^n c_i c_j h(x_i, x_j) \leq 0$$

負定値カーネルと正定値カーネル

■ Schoenberg の定理

$h(x,y) : \Omega \times \Omega \rightarrow \mathbf{R}$ が負定値

⇔

$\exp(-\beta h(x,y))$ が任意の $\beta > 0$ に対して正定値

■ 応用

$\|x - y\|^2$ は負定値

$$\begin{aligned} \therefore) \quad \sum_{i,j} c_i c_j \|x_i - x_j\|^2 &= \sum_{i,j} c_i c_j (\|x_i\|^2 + \|x_j\|^2 - 2x_i^T x_j) \\ &= \sum_i c_i \|x_i\|^2 \sum_j c_j + \sum_i c_i \sum_j c_j \|x_j\|^2 - 2 \sum_{i,j} c_i c_j x_i^T x_j \\ &= 0 + 0 - 2 \left\| \sum_i c_i x_i \right\|^2 \quad (c_1 + \dots + c_n = 0) \end{aligned}$$

⇒ $\exp(-\beta \|x - y\|^2)$ は $\beta > 0$ に対して正定値

Representer Theorem

■ 正則化の問題(復習)

□ 平滑化
$$\min_{f \in H} \sum_{i=1}^N (Y^i - f(X^i))^2 + \lambda \|f\|_H^2$$

□ SVM
$$\min_{f, b} \sum_{i=1}^N (1 - Y^i (f(X^i) + b))_+ + \lambda \|f\|_H^2$$

■ 一般化された問題

k : 正定値カーネル, H : k により定まる再生核ヒルベルト空間

$x_1, \dots, x_N, y_1, \dots, y_N$: データ(固定)

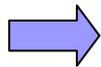
$h_1(x), \dots, h_d(x)$: 固定された関数

$$(*) \quad \min_{\substack{f \in H \\ (b_\ell) \in \mathbf{R}^d}} L\left(\{x_i\}_{i=1}^N, \{y_i\}_{i=1}^N, \left\{f(x_i) + \sum_{\ell=1}^d b_\ell h_\ell(x_i)\right\}_{i=1}^N\right) + \Psi\left(\|f\|_H^2\right)$$

■ Representer Theorem

正則化項の関数 Ψ は, $[0, \infty)$ 上の単調増加関数とする.

$\tilde{H}_N = \text{span}\{k(\cdot, x_1), \dots, k(\cdot, x_N)\}$ $\{k(\cdot, x_j)\}$ の張る N 次元部分空間



(*) の解 f は \tilde{H}_N の中にある. すなわち

$$f(x) = \sum_{i=1}^N \alpha_i k(x, x_i)$$

の形で探してよい.

$$\min_{f \in H} \min_{(b_\ell) \in \mathbf{R}^d} L\left(\{x_i\}_{i=1}^N, \{y_i\}_{i=1}^N, \left\{f(x_i) + \sum_{\ell=1}^d b_\ell h_\ell(x_i)\right\}_{i=1}^N\right) + \Psi\left(\|f\|_H^2\right)$$

$$= \min_{(\alpha_i) \in \mathbf{R}^N} \min_{(b_\ell) \in \mathbf{R}^d} L\left(\{x_i\}_{i=1}^N, \{y_i\}_{i=1}^N, \left\{\sum_{j=1}^N K_{ij} \alpha_j + \sum_{\ell=1}^d b_\ell h_\ell(x_i)\right\}_{i=1}^N\right) + \Psi\left(\alpha^T K \alpha\right)$$

$K = (k(x_i, x_j))$: グラム行列

H (無限次元) 上の最適化が有限次元の最適化に変換できる

■ Representer theorem の証明

$$\min L\left(\{x_i\}_{i=1}^N, \{y_i\}_{i=1}^N, \left\{f(x_i) + \sum_{\ell=1}^d b_\ell h_\ell(x_i)\right\}_{i=1}^N\right) + \Psi\left(\|f\|_H^2\right)$$

$$H = \tilde{H}_N \oplus H_\perp \quad \text{直交分解}$$

$$f = \tilde{f}_N + f_\perp$$

$$\langle f_\perp, k(\cdot, x_i) \rangle = 0 \quad (\forall i)$$

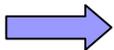


$$\blacksquare f(x_i) = \langle f, k(\cdot, x_i) \rangle = \langle \tilde{f}_N + f_\perp, k(\cdot, x_i) \rangle = \langle \tilde{f}_N, k(\cdot, x_i) \rangle = \tilde{f}_N(x_i)$$

→ L の値は \tilde{f}_N だけで決まる

$$\blacksquare \|f\|_H^2 = \|\tilde{f}_N\|_H^2 + \|f_\perp\|_H^2$$

→ Ψ の値は $f_\perp = 0$ のほうがよい



$f \in \tilde{H}_N$ に最適解がある

(証明終)

正定値カーネルの固有展開～Mercerの定理

■ 正定値カーネルと積分作用素

$(\Omega, \mathcal{B}, \mu)$: 測度空間 $\mu(\Omega) < \infty$.

$k(x,y) : \Omega \times \Omega \rightarrow \mathbf{R}$ は2乗可積分, $\iint k(x,y)^2 d\mu(x)d\mu(y) < \infty$

このとき次の積分作用素が $L^2(\Omega, \mu)$ 上に定まる

$$f \mapsto Kf(x) = \int k(x,y)f(y)d\mu(y)$$

積分核 $k(x,y)$ が次を満たすとき, **正定値**であるという

1. (対称性) $k(x,y) = k(y,x)$

2. (正定値性)

$$\iint k(x,y)f(x)f(y)d\mu(x)d\mu(y) \geq 0$$

* 前の定義と多少異なるが, 緩やかな条件のもと同値になる
(例えば Ω が \mathbf{R}^m のコンパクト集合, かつ $k(x,y)$ が連続)

■ Mercerの定理

積分核 $k(x,y)$ が正定値のとき, 作用素 K は正定値な作用素, すなわち

$$(f, Kf)_{L^2(\Omega)} \geq 0 \quad (\forall f \in L^2(\Omega))$$

→ 作用素 K の固有展開

K の単位固有ベクトル ($L^2(\Omega)$ の正規直交系) $\phi_1(x), \phi_2(x), \dots, \phi_n(x), \dots$
対応する固有値 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq \dots > 0$

$$(*) \quad Kf = \sum_n \lambda_n (f, \phi_n)_{L^2(\Omega)} \phi_n \quad (\text{c.f. 半正定値行列の固有分解})$$

定理 (Mercer)

$$(**) \quad k(x, y) = \sum_{n=1}^{\infty} \lambda_n \phi_n(x) \phi_n(y)$$

がほとんどすべての x, y に対し成り立ち, 一様かつ絶対収束する

注) 積分核 k が $(**)$ の展開を持てば, 作用素 K が正定値で $(*)$ の展開を持つのは明らか.

RKHSの基底展開による表現

■ RKHSの関数と内積

$k(x,y)$: 積分核としての正定値性を満たすカーネル

H_k : k によって定まるRKHS

実は

$$H_k = \left\{ f \in L^2(\Omega) \mid f(x) = \sum_{n=1}^{\infty} a_n \phi_n(x), \quad \sum_{n=1}^{\infty} \frac{a_n^2}{\lambda_n} < \infty \right\}$$

$$\langle f, g \rangle_{H_k} = \sum_{n=1}^{\infty} \frac{a_n b_n}{\lambda_n} \quad f = \sum_{n=1}^{\infty} a_n \phi_n, \quad g = \sum_{n=1}^{\infty} b_n \phi_n$$

注) $k(\cdot, x) = \sum_{n=1}^{\infty} \lambda_n \phi_n(\cdot) \phi_n(x)$ なので

$$\langle k(\cdot, x), f \rangle_{H_k} = \sum_{n=1}^{\infty} \frac{\lambda_n \phi_n(x) a_n}{\lambda_n} = \sum_{n=1}^{\infty} a_n \phi_n(x) = f(x)$$

再生性 !!

■ 有限点上のRKHS

$\Omega = \{1, 2, \dots, n\}$ 上の正定値カーネル $k(x, y)$

$x, y = 1, \dots, n$ の値が決まればよいので, 半正定値 $n \times n$ 行列

$$K_{ij} = k(i, j) \quad (i, j = 1, \dots, n)$$

を与えればよい.

$$K = \sum_{i=1}^n \lambda_i u_i u_i^T \quad \text{固有分解}$$

H_k : Ω 上の関数 $\cong n$ 次元ベクトル空間

H_k の内積: $f = \sum_{i=1}^n \alpha_i u_i, \quad g = \sum_{i=1}^n \beta_i u_i$ に対し

$$\langle f, g \rangle_{H_k} = \sum_{i=1}^n \frac{\alpha_i \beta_i}{\lambda_i} = \underline{\alpha^T K^{-1} \beta}$$

K を分散共分散行列としたときの, Mahalanobis内積

\mathbb{R}^m 上の正定値カーネルの表現

- Fourierカーネル (複素数値の正定値カーネル)

$$\begin{aligned}\exp(\sqrt{-1}\omega^T(x-y)) &= \exp(\sqrt{-1}\omega^T x)\exp(-\sqrt{-1}\omega^T y) \\ &= f(x)\overline{f(y)} \quad \Rightarrow \text{正定値}\end{aligned}$$

- ガウスカーネルの正定値性 (別証明)

- ガウス関数のFourier変換

$$\exp(-\|x\|^2/2) = \frac{1}{(2\pi)^m} \int_{\mathbb{R}^m} \exp(-\|\omega\|^2/2) \exp(\sqrt{-1}\omega^T x) d\omega$$

ガウス関数の Fourier変換は, またガウス関数

- 正定値性の証明

$$\underbrace{\sum_t \exp(-\|\omega_t\|^2/2)}_{\text{正の数}} \underbrace{\exp(\sqrt{-1}\omega_t^T(x-y))}_{\text{正定値カーネル}} \rightarrow \exp(-\|x-y\|^2/2)$$

↓
正定値

Bochner の定理

$k(x,y) = \phi(x-y)$ の形により与えられる \mathbf{R}^m 上のカーネルが正定値であるための必要十分条件は、関数 $\phi(z)$ が、ある非負可積分関数 $f(z)$ によって

$$\phi(z) = \int_{\mathbf{R}^m} \underbrace{f(\omega)}_{\text{非負}} \underbrace{\exp(\sqrt{-1}\omega^T z)}_{\text{Fourierカーネル(正定値)}} d\omega$$

と表されることである。

すなわち、 $\phi(z)$ のFourier変換が非負実数関数となることである。

- 上の積分表示を持つ $\phi(z)$ が正定値カーネルを与えることは、ガウスの場合と同様。Bochnerの定理は、その逆も成り立つことを主張している。
- Fourierカーネル $\exp(\sqrt{-1}\omega^T(x-y))$ が、 $\phi(x-y)$ 型の正定値カーネル全体の成す閉凸錐を張っている。

セクション4のまとめ

■ 正定値性を保つ演算

- 非負の和, 積
 - Normalization
 - 正定値カーネル列の各点収束
- ⇒ あたらしいカーネルの作成 / カーネルの正定値性のチェック

■ Representer theorem

無限次元空間上の正則化問題を有限次元の問題へ

■ Mercerの定理

正定値カーネルの固有展開 → RKHSの基底による陽な表現を与える

■ Bochner の定理

$\phi(x - y)$ 型の \mathbb{R}^m 上のカーネルが正定値 $\Leftrightarrow \phi$ のFourier変換が非負

5. カーネル法とスプライン

- このセクションの目的
 - 正則化・スプラインをRKHS上の最適化として考察する

スプライン

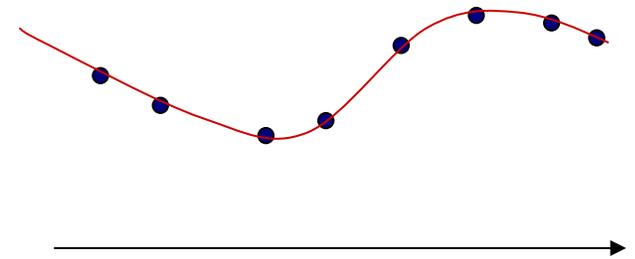
$(X^1, Y^1), \dots, (X^N, Y^N)$: 与えられたデータ $X^i \in \mathbf{R}^n, Y^i \in \mathbf{R}$

P : 微分作用素 e.g. $P = \frac{d^2}{dx^2}$

■ スプライン補間

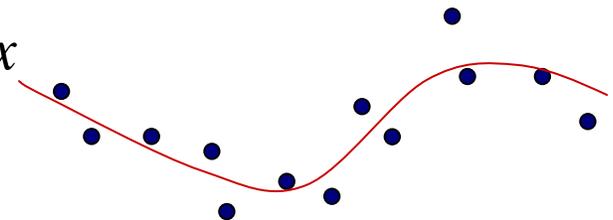
$$\min_f \int |Pf(x)|^2 dx$$

subject to $f(X^i) = Y^i$



■ スプライン平滑化 主にこちらを扱う

$$\min_f \sum_{i=1}^N (Y^i - f(X^i))^2 + \lambda \int |Pf(x)|^2 dx$$



解析学の復習

■ 部分積分

$|f|, |g| \rightarrow 0$ ($\|x\| \rightarrow \infty$) とすると

$$\int f(x) \frac{\partial g(x)}{\partial x} dx = - \int \frac{\partial f(x)}{\partial x} g(x) dx$$

■ 微分のFourier変換

$$\begin{aligned} \frac{\widetilde{\partial f}}{\partial x_i}(u) &= \frac{1}{(2\pi)^{n/2}} \int \frac{\partial f}{\partial x_i}(x) e^{-\sqrt{-1}u^T x} dx \\ &= -\frac{1}{(2\pi)^{n/2}} \int f(x) (-\sqrt{-1}u_i) e^{-\sqrt{-1}u^T x} dx = \sqrt{-1}u_i \tilde{f}(u) \end{aligned}$$

■ Laplacianの定義

$$\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2}$$

■ Green関数 (Laplacianの場合で説明)

□ 微分方程式

$$\text{Solve } f: \quad \Delta f = \varphi$$

φ : 与えられた関数

Δ : Laplacian (自己共役な微分作用素)

$$\text{i.e. } \int \Delta f(x) g(x) dx = \int f(x) \Delta g(x) dx$$

$|f|, |g| \rightarrow 0$ ($\|x\| \rightarrow \infty$) とする

□ Green関数 $G(x, \xi)$

$$\Delta G(x, \xi) = \delta(x - \xi)$$

$G(x, \xi)$: 対称 $G(x, \xi) = G(\xi, x)$

□ 微分方程式の解

$$f(x) = \int G(x, y) \varphi(y) dy$$

$$\begin{aligned} \therefore f(\xi) &= \int f(x) \delta(x - \xi) dx = \int f(x) \Delta G(x, \xi) dx \\ &= \int \Delta f(x) G(x, \xi) dx = \int \varphi(x) G(x, \xi) dx \end{aligned}$$

滑らかさによる正則化

■ 滑らかさを表す正則化項

以降, \mathbf{R}^n 上の関数を考える

m 次微分の L^2 ノルム

$$\begin{aligned} J_m^n(f) &= \sum_{\alpha_1 + \dots + \alpha_n = m} \frac{m!}{\alpha_1! \alpha_2! \dots \alpha_n!} \|D^\alpha f\|_{L^2}^2 \\ &= \sum_{\alpha_1 + \dots + \alpha_n = m} \frac{m!}{\alpha_1! \alpha_2! \dots \alpha_n!} \int \left| \frac{\partial^m f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}} \right|^2 dx \end{aligned}$$

例) $m = n = 2$

$$J_2^2(f) = \int \left\{ \left| \frac{\partial^2 f}{\partial x_1^2} \right|^2 + 2 \left| \frac{\partial^2 f}{\partial x_1 \partial x_2} \right|^2 + \left| \frac{\partial^2 f}{\partial x_2^2} \right|^2 \right\} dx$$

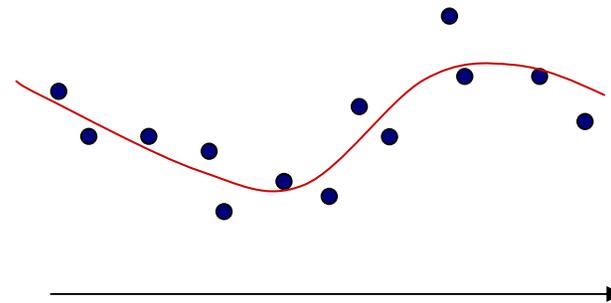
- 本当は有界領域でも考えるべきだが, 境界条件の扱いが面倒なため, ここでは \mathbf{R}^n 全体で考える.

■ 正則化問題

- 以降, 次の形の正則化を考える.

$$\min_f \sum_{i=1}^N (Y^i - f(X^i))^2 + \lambda \sum_{m=0}^{\infty} a_m J_m^n(f)$$

λ : 正則化係数



- Laplacianによる表現

$|D^\alpha f| \rightarrow 0$ ($\|x\| \rightarrow \infty$) とすると, 部分積分により

$$J_m^n(f) = (-1)^m (f, \Delta^m f)_{L^2}$$

したがって

$$\min_f \sum_{i=1}^N (Y^i - f(X^i))^2 + \lambda (f, Af)_{L^2}$$

$$\text{ただし } A = \sum_{m=0}^{\infty} (-1)^m a_m \Delta^m$$

■ 2つのケース

□ $a_0 \neq 0$ の場合

- Green関数が正定値カーネル
- RKHSノルムで正則化していることになる

□ $a_0 = 0$ の場合

- Spline平滑化
- m 次の微分作用素の場合を考える ($a_m = 1$, その他 0)
- Green関数は条件付正定値
- 関数 f を射影したものをRKHSノルムで正則化している

$a_0 \neq 0$ の場合

■ 変分による解法

$$\min_f \sum_{i=1}^N (Y^i - f(X^i))^2 + \lambda (f, Af)_{L^2}$$

変分 \rightarrow
$$\sum_{i=1}^N (Y^i - f(x)) \delta(x - X^i) + \lambda Af = 0$$

$$Af = -\frac{1}{\lambda} \sum_{i=1}^N (Y^i - f(x)) \delta(x - X^i) \quad \text{ただし} \quad A = \sum_{m=0}^{\infty} (-1)^m a_m \Delta^m$$

もし A の Green 関数 $G(x; \xi)$ (i.e. $AG(x; \xi) = \delta(x - \xi)$) が求めれば

$$\begin{aligned} f(\xi) &= -\frac{1}{\lambda} \sum_{i=1}^N \int (Y^i - f(x)) \delta(x - X^i) G(x, \xi) dx \\ &= -\frac{1}{\lambda} \sum_{i=1}^N (Y^i - \underbrace{f(X^i)}_{\text{未知}}) G(\xi, X^i) \end{aligned}$$

解の形は

$$f = \sum_{i=1}^N c_i G(\cdot, X^i)$$

$$AG(x; \xi) = \delta(x - \xi)$$

もとの正則化問題に代入すると

$$\min_{c \in \mathbf{R}^N} \sum_{i=1}^N \left(Y^i - \sum_{j=1}^N c_j G(X^i, X^j) \right)^2 + \lambda \sum_{i,j=1}^N c_i c_j G(X^i, X^j)$$



微分 \rightarrow

$$(G^2 + \lambda G)c = G\mathbf{Y}$$

$$\text{ただし } G_{ij} = G(X^i, X^j) \\ \mathbf{Y} = (Y^1, \dots, Y^N)^T$$

G が非特異とすると

$$c = (G + \lambda I)^{-1} \mathbf{Y}$$

解:

$$f(x) = \mathbf{Y}^T (G + \lambda I)^{-1} g(x)$$

$$\text{ただし } g_i(x) = G(x, X^i)$$

■ Green関数の求め方

Fact: Laplacianが平行移動に対して不変 $\Rightarrow G(x, \xi) = G(x - \xi)$

$$\sum_{m=0}^{\infty} (-1)^m a_m \Delta^m G(z) = \delta(z)$$

Fourier変換



$$\sum_{m=0}^{\infty} a_m \|u\|^{2m} \widetilde{G}_\xi(u) = \frac{1}{(2\pi)^{n/2}}$$

$$\widetilde{G}_\xi(u) = \frac{1}{(2\pi)^{n/2} \left(a_0 + \sum_{m=1}^{\infty} a_m \|u\|^{2m} \right)}$$

$a_0 \neq 0, a_1 \neq 0$ ならば逆Fourier変換可能

$$\Rightarrow G(z) = \int \frac{e^{\sqrt{-1}z^T u}}{(2\pi)^n \left(a_0 + \sum_{m=1}^{\infty} a_m \|u\|^{2m} \right)} du$$

■ 例： ガウスRBFカーネル

$$a_m = \frac{\sigma^{2m}}{m!2^m}, \quad (\sigma > 0)$$

→ $G(z) = \frac{1}{(2\pi)^n} \int \exp\left(-\frac{\sigma^2 \|u\|^2}{2}\right) e^{\sqrt{-1}z^T u} du$

$$= c \exp\left(-\frac{\|z\|^2}{2\sigma^2}\right)$$

ガウスRBFカーネル関数

■ Green関数の正定値性

$a_0 \neq 0$ の場合のGreen関数は正定値

$$G(z) = \frac{1}{(2\pi)^n} \int \frac{1}{\left(a_0 + \sum_{m=1}^{\infty} a_m \|u\|^{2m}\right)} \underbrace{e^{\sqrt{-1}z^T u} du}_{\text{Fourierカーネル}}$$

非負

(Bochnerの定理)

正則化とカーネル法

■ RKHSノルムによる正則化

Green関数 $G(x-y)$ により与えられるRKHS $\cdots H_G$
カーネルによる正則化問題

$$\min_f \sum_{i=1}^N \left(Y^i - f(X^i) \right)^2 + \lambda \| f \|_{H_G}^2$$

representer定理より

$$f = \sum_{i=1}^N c_i G(\cdot, X^i)$$

$$\min_{c \in \mathbf{R}^N} \sum_{i=1}^N \left(Y^i - \sum_{j=1}^N c_j G(X^i, X^j) \right)^2 + \lambda \sum_{i,j=1}^N c_i c_j G(X^i, X^j)$$

$$J(f) = \sum_{m=0}^{\infty} a_m J_m^n(f) \quad (a_0 \neq 0)$$

による正則化項と同一

- $a_0 \neq 0$ の場合 :

滑らかさによる正則化

⇔ Green関数が与えるRKHSノルムによる正則化

$$\min_f \sum_{i=1}^N \left(Y^i - f(X^i) \right)^2 + \lambda \sum_{m=0}^{\infty} a_m J_m^n(f)$$



同値

$$\min_f \sum_{i=1}^N \left(Y^i - f(X^i) \right)^2 + \lambda \| f \|_{H_G}^2$$

スプラインの場合 ($a_0 = 0$)

■ $a_0 = 0$ の特別な場合: Thin-plate spline

□ 正則化問題

$$\min_f \sum_{i=1}^N (Y^i - f(X^i))^2 + \lambda J_m^n(f) \quad m \text{ 次微分}$$

$$J_m^n(f) = \sum_{\alpha_1 + \dots + \alpha_n = m} \frac{m!}{\alpha_1! \alpha_2! \dots \alpha_n!} \|D^\alpha f\|_{L^2}^2$$

□ 関数空間

$$B_m^n : D^\alpha f \in L^2(\mathbf{R}^n) \quad (|\alpha| = m) \quad (\text{本当は超関数で考える})$$

□ 零化空間の存在

$$J_m^n(f) = 0 \quad \Leftrightarrow \quad f \in \mathcal{P}_{m-1}$$

\mathcal{P}_{m-1} : $m-1$ 次以下の多項式全体

■ 変分による解法？ ($a_0 \neq 0$ の場合との違い)

□ Green関数

$$\Delta^m G(z) = \delta(z)$$

解

$$G(z) = \begin{cases} \beta_{m,n} |z|^{2m-n} \log |z| & (2m-n \text{ が偶整数のとき}) \\ \gamma_{m,n} |z|^{2m-n} & (\text{それ以外}) \end{cases}$$

□ Fourier変換では得るのが難しい

$$\widetilde{G}_\xi(u) = \frac{1}{(2\pi)^{n/2} \|u\|^{2m}}$$

Fourier逆変換不可能. 原点で積分できない

- $G(z)$ は正定値ではない (Bochnerの定理は使えない)
実は、条件付正定値と呼ばれるクラス(≡微分すると正定値)

■ RKHSによる定式化

□ 多項式空間の基底

$$\dim \mathcal{P}_{m-1} = \binom{m+n-1}{n} \equiv M$$

次のような点 $a_1, \dots, a_M \in \mathbf{R}^n$ を取る. (一般の位置にとればOK)
「任意の $b_1, \dots, b_M \in \mathbf{R}$ に対してある多項式 $p(x) \in \mathcal{P}_{m-1}$ があって

$$p(a_i) = b_i \quad (1 \leq i \leq M)$$

を満たす」 \dots (unisolvent)

■ \mathcal{P}_{m-1} の基底 ϕ_1, \dots, ϕ_M : $\phi_\ell(a_i) = \begin{cases} 1 & (\ell = i) \\ 0 & (\ell \neq i) \end{cases}$

■ 射影

$$Q: f \mapsto \sum_{\ell=1}^M f(a_\ell) \phi_\ell \quad (\text{Lagrange補間})$$

$$Qp = p \quad (\forall p \in \mathcal{P}_{m-1}) \quad Qf(a_i) = f(a_i)$$

$\Pi = I - Q: B_m^n \rightarrow B_m^n$ \mathcal{P}_{m-1} の補空間 H_* への射影
 (多項式部分を除く写像)
 $B_m^n = \mathcal{P}_{m-1} \oplus H_*$

定理

$m > n/2$ のとき, H_* は内積

$$\langle f, g \rangle_{H_*} = \sum_{|\alpha|=m} \frac{m!}{\alpha_1! \cdots \alpha_n!} (D^\alpha f, D^\alpha g)_{L^2} = \left((-1)^m \Delta^m f, g \right)_{L^2}$$

を持つ再生核ヒルベルト空間である. 特にノルムは

$$\|f\|_{H_*}^2 = J_m^n(f)$$

証明略 (Meinguet 1979を参照)

■ スプライン平滑化とRKHS

□ 上の定理から $m > n/2$ のとき, 次は同値

$$\begin{aligned} & \min_f \sum_{i=1}^N (Y^i - f(X^i))^2 + \lambda J_m^n(f) \\ \iff & \min_f \sum_{i=1}^N (Y^i - f(X^i))^2 + \lambda \|\Pi f\|_{H_*}^2 \\ \iff & \min_{g \in H_*, p \in \mathcal{P}_{m-1}} \sum_{i=1}^N (Y^i - (g(X^i) + p(X^i)))^2 + \lambda \|g\|_{H_*}^2 \end{aligned}$$

□ c.f. $a_0 \neq 0$ の場合

$$\begin{aligned} & \min_f \sum_{i=1}^N (Y^i - f(X^i))^2 + \lambda \sum_{m=0}^{\infty} a_m J_m^n(f) \\ \iff & \min_f \sum_{i=1}^N (Y^i - f(X^i))^2 + \lambda \|f\|_{H_G}^2 \end{aligned}$$

■ 正定値カーネルの導出

H_* の正定値カーネル $K(x,y)$

内積の形から

$$\langle K(\cdot, y), \Pi f \rangle_{H_*} = \left((-1)^m \Delta_x^m K(x, y), f \right)_{L^2}$$

一方

$$\langle K(\cdot, y), \Pi f \rangle_{H_*} = f(y) - \sum_i f(a_i) \phi_i(y) = \left(\delta_y - \sum_i \phi_i(y) \delta_{a_i}, f \right)_{L^2}$$

(再生性)



$$(-1)^m \Delta^m K(\cdot, y) = \delta_y - \sum_i \phi_i(y) \delta_{a_i}$$

Green関数 $G(x-y)$ ($\Delta^m G(z) = \delta(z)$) を使うと

$$K(x, y) = (-1)^m \left\{ G(x-y) - \sum_i \phi_i(y) G(x-a_i) \right\} + \psi(x)$$

$$\psi(x) \in \mathcal{P}_{m-1}$$

$K(x,y) \in H_*$ なので

$$\begin{aligned} K(x, y) &= \Pi_x K(x, y) \\ &= (-1)^m \left\{ G(x-y) - \sum_i \phi_i(y) G(x-a_i) - \sum_i \phi_i(x) G(y-a_i) \right. \\ &\quad \left. + \sum_{i,j} \phi_i(x) \phi_j(y) G(a_i - a_j) \right\} \end{aligned}$$

■ 解の形

Representer定理を使うと、最適解は

$$f(x) = \sum_{i=1}^N c_i K(x - X_i) + \sum_{\ell=1}^M b_{\ell} \phi_{\ell}(x)$$

の形を持つが、さらに K の G による表示により

$$f(x) = \underbrace{\sum_{i=1}^N c_i G(x - X_i)}_{\text{Green関数}} + \underbrace{\sum_{\ell=1}^M b_{\ell} \phi_{\ell}(x)}_{m-1\text{次以下の多項式}}$$

Green関数 $m-1$ 次以下の多項式

条件付正定値カーネル

■ 条件付正定値

$K(x,y) : \Omega \times \Omega \rightarrow \mathbf{R}$ が m 次の条件付正定値カーネルであるとは,

1. $K(x,y) = K(y,x)$

2. Ω の点 x_1, \dots, x_n と実数 c_1, \dots, c_n が, 任意の多項式 $p(x) \in \mathcal{P}_{m-1}$ に対して

$$\sum_{i=1}^n c_i p(x_i) = 0$$

を満たすならば,

$$\sum_{i,j=1}^n c_i c_j K(x_i, x_j) \geq 0$$

- 直感的には、 m 階微分すると正定値になるようなカーネル
- 上の c_1, \dots, c_n は m 階差分(係数)の一般化

$$\text{1階差分} \quad \frac{f(t_{i+1}) - f(t_i)}{t_{i+1} - t_i}$$

$$f(t_i) \text{ の係数} \quad \begin{array}{cccc} \mathbf{C_1} & \mathbf{C_2} & \mathbf{C_3} & \dots \\ \frac{-1}{t_2 - t_1} & \frac{1}{t_2 - t_1} & 0 & \\ 0 & \frac{-1}{t_3 - t_2} & \frac{1}{t_3 - t_2} & \end{array}$$

$$c_1 + c_2 + \dots + c_n = 0 \quad \rightarrow \quad \text{一般化1階差分の係数}$$

$$\text{2階差分} \quad \left\{ \frac{f(t_{i+2}) - f(t_{i+1})}{t_{i+2} - t_{i+1}} - \frac{f(t_{i+1}) - f(t_i)}{t_{i+1} - t_i} \right\} / t_{i+2} - t_{i+1}$$

$$f(t_i) \text{ の係数} \quad \begin{array}{cccc} \mathbf{C_1} & \mathbf{C_2} & \mathbf{C_3} & \mathbf{C_4} \\ \frac{1}{t_2 - t_1} / t_3 - t_1 & \frac{-1}{t_2 - t_1} + \frac{-1}{t_3 - t_2} / t_3 - t_1 & \frac{1}{t_3 - t_2} / t_3 - t_1 & 0 \\ 0 & \frac{1}{t_3 - t_2} / t_4 - t_2 & \frac{-1}{t_3 - t_2} + \frac{-1}{t_4 - t_3} / t_4 - t_2 & \frac{1}{t_4 - t_3} / t_4 - t_2 \end{array} \quad \dots$$

$$\begin{cases} c_1 + c_2 + \dots + c_n = 0 \\ c_1 \times t_1 + c_2 \times t_2 + \dots + c_n \times t_n = 0 \end{cases} \quad \rightarrow \quad \text{一般化2階差分の係数}$$

■ Green関数の条件付正定値性

定理

平行移動に関して不変な定数係数 m 次微分作用素 P に対する Green 関数 $G(z)$, すなわち

$$P^*PG(z) = \delta(z)$$

の解に対し, $G(x-y)$ は条件付正定値カーネルである.

(Gel'fand & Vilenkin 1964, Ch.II.4)

■ 条件付正定値と正定値

定理

$K(x,y)$ を m 次の条件付正定値カーネルとするととき,
unisolvent $\{a_1, \dots, a_M\}$ と \mathcal{P}_{m-1} の基底 ϕ_i ($\phi_i(a_j) = \delta_{ij}$) に対し,

$$R(x, y) = K(x, y) - \sum_i \phi_i(y)K(x, a_i) - \sum_i \phi_i(x)K(y, a_i) \\ + \sum_{i,j} \phi_i(x)\phi_j(y)K(a_i, a_j)$$

は正定値カーネルである.

セクション5のまとめ

- 正則化／スプラインはRKHSで記述できる

$$\min_f \sum_{i=1}^N (Y^i - f(X^i))^2 + \lambda \sum_{m=0}^{\infty} a_m J_m^n(f)$$

$$J_m^n(f) = \sum_{\alpha_1 + \dots + \alpha_n = m} \frac{m!}{\alpha_1! \alpha_2! \dots \alpha_n!} \|D^\alpha f\|_{L^2}^2$$

$$a_0 \neq 0$$

Green関数は正定値

$$\min_f \sum_{i=1}^N (Y^i - f(X^i))^2 + \lambda \|f\|_{H_G}^2$$

解は

$$f = \sum_{i=1}^N c_i G(\cdot, X^i)$$

$$a_0 = 0$$

Green関数は条件付正定値

$$\min_f \sum_{i=1}^N (Y^i - f(X^i))^2 + \lambda \|\Pi f\|_{H^*}^2$$

解は

$$f = \sum_{i=1}^N c_i G(\cdot, X^i) + \sum_{\ell=1}^M b_\ell \phi_\ell$$

5. カーネル法とガウス過程

- このセクションの目的
 - RKHSとガウス過程の関連について述べる
 - カーネル法とガウス過程の推定との関連について述べる

ガウス過程

- 多変量ガウス分布(復習)

- m 変数ガウス分布 $X = (X_1, \dots, X_m)$

確率密度関数

$$g_m(x; \mu, \Sigma) = \frac{1}{(2\pi |\Sigma|)^{1/2}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right)$$

- 平均値 μ と分散共分散 Σ によって定まる

$$\mu_i = E[X_i], \quad \Sigma_{ij} = \text{Cov}[X_i, X_j]$$

■ ガウス過程

□ 集合 Ω を添え字集合とするガウス過程 $\{X_t\}_{t \in \Omega}$

任意の有限個の Ω の点 t_1, \dots, t_n に対し,

同時分布 $(X_{t_1}, \dots, X_{t_n})$ が n 変量ガウス分布となる確率過程

□ ガウス過程は

平均関数 $\mu(t) = E[X_t]$

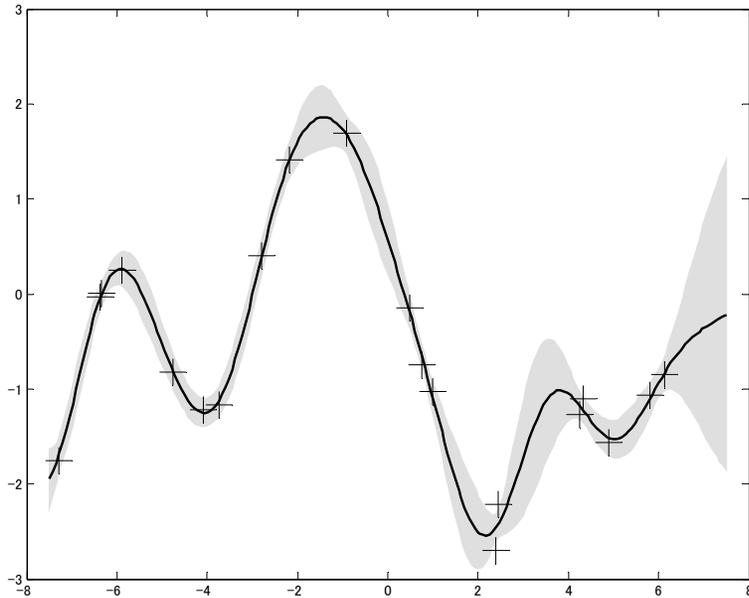
共分散関数 $R(t, s) = \text{Cov}[X_t, X_s]$

によって定まる.

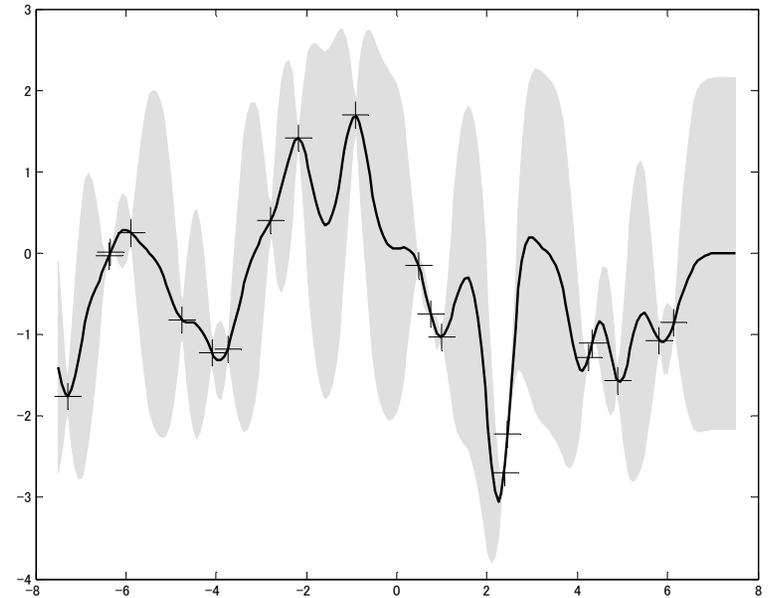
$\mathbf{X} = (X_{t_1}, \dots, X_{t_n})$ が従う n 変量ガウス分布の平均 $\mu_{\mathbf{X}}$ と分散 $\Sigma_{\mathbf{X}}$

$$\mu_{\mathbf{X}} = (\mu(t_1), \dots, \mu(t_n)), \quad \Sigma_{\mathbf{X}} = \begin{pmatrix} R(t_1, t_1) & R(t_1, t_2) & \cdots & R(t_1, t_n) \\ R(t_2, t_1) & R(t_2, t_2) & \cdots & R(t_2, t_n) \\ \vdots & \vdots & \ddots & \vdots \\ R(t_n, t_1) & R(t_n, t_2) & \cdots & R(t_n, t_n) \end{pmatrix}$$

■ ガウス過程の例



$\sigma = 1$



$\sigma = 0.3$

$$\text{共分散関数 } R(s, t) = \exp\left(-\frac{1}{2\sigma^2}(s-t)^2\right)$$

Rasmussen and Williams, matlab gpml toolbox により作成

ガウス過程と正定値カーネル

■ 共分散関数の正定値性

□ 共分散関数 $R: \Omega \times \Omega \rightarrow \mathbf{R}$ は正定値カーネル.

∴) 簡単のため平均0とする.

$$\begin{aligned}\sum_{i,j=1}^n c_i c_j R(t_i, t_j) &= \sum_{i,j=1}^n c_i c_j E[X_{t_i}, X_{t_j}] \\ &= E\left[\sum_{i=1}^n c_i X_{t_i}, \sum_{j=1}^n c_j X_{t_j}\right] = E\left[\left(\sum_{i=1}^n c_i X_{t_i}\right)^2\right] \geq 0\end{aligned}$$

■ ガウス過程 \Rightarrow RKHS

ガウス過程の共分散関数 $R(s, t)$ により, RKHS H_R が定まる.

■ 正定値カーネル \Rightarrow ガウス過程

- 集合 Ω 上の正定値カーネル $k(s, t)$ によって, k を共分散関数に持つ平均0のガウス過程が定まる.

構成法

$k: \Omega \times \Omega \rightarrow \mathbf{R}$: 正定値カーネル

$\mathbf{t} = (t_1, \dots, t_n)$, $t_i \in \Omega$

グラム行列 $\Sigma_{\mathbf{t}} = (k(t_i, t_j))$ は常に半正定値

\Rightarrow

$(X_{t_1}, \dots, X_{t_n})$ の分布が $N(0, \Sigma_{\mathbf{t}})$ となる Ω 上のガウス過程 $\{X_t\}_{t \in \Omega}$ が存在 (Kolmogorovの拡張定理)

■ ガウス過程により生成されるヒルベルト空間

$\{X_t\}_{t \in \Omega}$: Ω 上の平均0のガウス過程

$\Rightarrow X_t$: ある確率空間 (Ξ, \mathcal{B}, P) 上に定義される確率変数 $X_t : \Xi \rightarrow \mathbf{R}$

$\Rightarrow X_t \in L^2(\Xi, \mathcal{B}, P)$

$\overline{\mathcal{L}(X)} = \overline{\{X_t \in L^2(\Xi) \mid t \in \Omega\}}$ $L^2(\Xi)$ の閉部分空間

$\{X_t\}_{t \in \Omega}$ の生成するヒルベルト空間と呼ぶ.

内積

$$(U, V)_{\overline{\mathcal{L}(X)}} = E[UV] \quad U, V \in \overline{\mathcal{L}(X)}$$

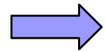
($L^2(\Xi)$ の内積)

■ RKHSとガウス過程

定理 (RKHSとガウス過程の対応)

k : 集合 Ω 上の正定値カーネル

$\{X_t\}_{t \in \Omega}$: k を共分散関数に持つ平均0のガウス過程



$$\overline{\mathcal{L}(X)} \cong H_k \quad \text{内積も込めて同型}$$

対応は $X_t \leftrightarrow k(\cdot, t)$ により定まる.

$$(U, V)_{\overline{\mathcal{L}(\Xi)}} = \langle f, g \rangle \quad U \leftrightarrow f, V \leftrightarrow g$$

注) $(X_t, X_s)_{\overline{\mathcal{L}(\Xi)}} = E[X_t X_s] = k(t, s) = \langle k(\cdot, t), k(\cdot, s) \rangle_{H_k}$
(内積の定義) (共分散) (再生性)

■ 定常過程の場合

\mathbf{R}^m 上の確率過程を考える.

□ 定常な確率過程

$$E[X_{t+h} X_{s+h}] = E[X_t X_s] \quad (\forall t, s, h \in \mathbf{R}^m)$$

共分散関数は

$$R(t, s) \equiv R(t - s)$$

によって決まる.

□ 正定値カーネル

定常ガウス過程に対応する正定値カーネルは次の形

$$K(t, s) = K(t - s)$$

□ Bochnerの定理 \Leftrightarrow Winer-Khinchineの定理

(\mathbf{R}^m 上の定常過程の共分散関数は,
パワースペクトルの逆Fourier変換)

■ n 点上のRKHSと多変量ガウス分布

□ RKHS H_K

正定値カーネル = 半正定値行列 $K_{ij} = k(i, j)$

固有分解 $K = \sum_{p=1}^n \lambda_p u_p u_p^T$ (u_p : 固有ベクトル, λ_p : 固有値)

内積 $\langle f, g \rangle = \alpha^T K^{-1} \beta$ $f = \sum_p \alpha_p u_p, g = \sum_p \beta_p u_p$

□ n 次元ガウス分布

$Z_1, \dots, Z_n \sim N(0,1)$, i.i.d.を用いて, n 次元ガウス確率ベクトルを定義

$$X = \sum_{p=1}^n \sqrt{\lambda_p} Z_p u_p$$

$$E[X_i X_j] = \sum_{p,q} \sqrt{\lambda_p} u_{ip} \sqrt{\lambda_q} u_{jq} E[Z_p Z_q] = \sum_p \lambda_p u_{ip} u_{jq} = K_{ij}$$

→ H_K と X が対応

内積 $\alpha^T K^{-1} \beta$ は、 $\text{Cov}[X]$ によるMahalanobis内積に一致

ガウス過程による推論

■ ガウス過程の推定と回帰問題

□ ガウス過程の推定

X_t : Ω 上のガウス過程

$$Y_t = X_t + \varepsilon_t$$

ε_t : X_t と独立なノイズ $E[\varepsilon_t] = 0$, $\text{Cov}[\varepsilon_t, \varepsilon_s] = \sigma^2 \delta(t - s)$

$R(t, s) = \text{Cov}[X_t, X_s]$: 既知とする. σ^2 も既知

問題

$(X_{t_1}, Y_{t_1}), \dots, (X_{t_n}, Y_{t_n})$ が観測されたとき, t_0 に対する X_{t_0} を推定せよ

□ 回帰問題

$(t_1, Y_{t_1}), \dots, (t_n, Y_{t_n})$ により $t \rightarrow Y_t$ を推定する回帰問題と考えられる

ガウス過程: Bayes推定

- ガウス分布の条件付確率(復習)

$$\begin{pmatrix} Z \\ Z_* \end{pmatrix} \sim N \left(\begin{pmatrix} \mu \\ \mu_* \end{pmatrix}, \begin{pmatrix} \Sigma & \Sigma_* \\ \Sigma_*^T & \Sigma_{**} \end{pmatrix} \right) \quad \text{多変量正規分布}$$



$p(Z_* | Z)$ もガウス分布で,

$$E[Z_* | Z] = \mu_* + \Sigma_*^T \Sigma^{-1} \mu$$

$$\text{Cov}[Z_* | Z] = \Sigma_{**} - \Sigma_*^T \Sigma^{-1} \Sigma_*$$

■ ガウス過程の推定問題の場合

$$\begin{pmatrix} Y_t \\ X_{t_0} \end{pmatrix} \sim N \left(0, \begin{pmatrix} K + \sigma^2 I_n & r \\ r^T & R(t_0, t_0) \end{pmatrix} \right)$$

ただし $K = (R(t_i, t_j)) \in \mathbf{R}^{n \times n}$ $r = (R(t_i, t_0)) \in \mathbf{R}^n$

\therefore) $E[Y_t, Y_s] = R(t, s) + \sigma^2 \delta(t - s), \quad E[Y_t, X_s] = R(t, s)$

■ 推定問題のBayes解

$$E[X_{t_0} | Y_t] = r^T (K + \sigma^2 I_n)^{-1} Y_t$$

ガウス過程：平均2乗解

■ ガウス過程に対する最小2乗推定

- 線形推定量

$$\hat{X}_{t_0} = \sum_{j=1}^n \alpha_j Y_{t_j}$$

を考える

- 平均2乗誤差

$$\min_{\alpha} E | X_{t_0} - \hat{X}_{t_0} |^2 = \min_{\alpha} E | X_{t_0} - \sum_{j=1}^n \alpha_j Y_{t_j} |^2$$

- 最小平均2乗解 = Bayes解

$$\hat{X}_{t_0} = \hat{\alpha}^T Y_t = r^T (K + \sigma^2 I_n)^{-1} Y_t$$

$$\therefore \min_{\alpha} \alpha^T (K + \sigma^2 I_n) \alpha - 2r^T \alpha$$

$$\Rightarrow \hat{\alpha} = (K + \sigma^2 I_n)^{-1} r$$

ガウス過程の推定と正則化問題

- ガウス過程の最小平均2乗解 / Bayes解 = RKHSによる正則化

- ガウス過程の線形2乗誤差推定

$$\min_{\alpha} E | X_{t_0} - \hat{X}_{t_0} |^2 = \min_{\alpha} E | X_{t_0} - \sum_{j=1}^n \alpha_j Y_{t_j} |^2$$

解 $\hat{X}_{t_0} = r^T (K + \sigma^2 I_n)^{-1} Y_t$

- RKHSノルムの正則項付き回帰

$$\min_{f \in H} \sum_{i=1}^N (Y_i - f(t_i))^2 + \lambda \| f \|_H^2$$

解 $f(t) = r(t)^T (K + \lambda I_N)^{-1} Y$

同一
 $\sigma^2 \Leftrightarrow \lambda$

■ RKHSとガウス過程の対応

RKHS

正定値カーネル $K(t,s)$

$$\sum c_i K(\cdot, t_i)$$

$\lim \sum c_i K(\cdot, t_i)$ (完備化)

正則化(平滑化)

$$\min_{f \in H} \sum_{i=1}^N (Y_i - f(t_i))^2 + \lambda \|f\|_H^2$$

$$f(t) = r(t)^T (K + \lambda I_N)^{-1} Y$$

平行移動不変なカーネル

$$K(t,s) = K(t-s)$$

に対するBochnerの定理

ガウス過程

分散共分散関数 $K(t,s) = E[X_t, X_s]$

$$\sum c_i X_{t_i}$$

$\lim \sum c_i X_{t_i}$ (閉包)

線形推定問題

$$\min_{\alpha} E | X_{t_0} - \sum_{j=1}^n \alpha_j Y_{t_j} |^2$$

$$\hat{X}_{t_0} = r^T (K + \sigma^2 I_n)^{-1} Y_t$$

定常過程の共分散関数

$$K(t,s) = K(t-s)$$

に対するWiener-Khinchineの定理

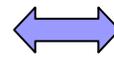
セクション6のまとめ

■ ガウス過程とRKHS

- RKHSとガウス過程が1対1に対応する
正定値カーネル \Leftrightarrow 共分散関数
- ガウス過程の線形最小2乗誤差推定 (Bayes推定) と, RKHSノルムの正則化付き回帰が対応
- $K(x-y)$ の形の正定値カーネル \Leftrightarrow 定常過程の共分散関数

□ 参考)

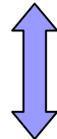
$K(x-y)$ の形の条件付正定値
カーネル



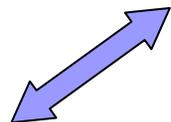
定常過程を積分した過程
(ARIMAなど)

の共分散関数の“定常部分”

(Intrinsic random functions
空間統計のKriging法)



Thin-plate スプライン



参考となる資料

■ ホームページ

- カーネル関連ポータルサイト <http://www.kernel-machines.org>

■ 全般的な文献

- Schölkopf, B. and A. Smola. *Learning with Kernels*. MIT Press. 2002.
- 津田宏治. 「カーネル法の理論と実際」 in 統計科学のフロンティア6: パターン認識と学習の統計学 (岩波書店) 2003.
- Müller K.-R., S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf. (2001) An introduction to kernel-based learning algorithms. *IEEE Trans. Neural Networks*, 12(2), pp.181-201.
- John Shawe-Taylor & Nello Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge Univ. Press. 2004.

■ 個別の話題に関する文献

- Kernel PCA
 - Schölkopf, B., A. Smola, K.-R. Müller. (1998) Nonlinear Component Analysis as a Kernel Eigenvalue Problem. *Neural Computation* 10, 1299–1319.

- Kernel CCA
 - Akaho, S. (2001) A kernel method for canonical correlation analysis. *International Meeting on Psychometric Society (IMPS2001)*.
 - Bach, F.R. and M.I. Jordan. Kernel independent component analysis. *J. Machine Learning Research*, 3, 1-48, 2002.
- スプライン
 - Green, P.J. and Silverman, B.W. *Nonparametric Regression and Generalized Linear Models. A Roughness Penalty Approach*. Chapman & Hall. 1994. (スプラインのわかりやすい解説)
 - Grace Wahba. *Spline Models for Observational Data*. CBMS-NSF Regional Conference Series in Applied Mathematics 59. SIAM. 1990.
 - Meinguet, J. (1979) Multivariate Interpolation at Arbitrary Points Made Simple. *J. Applied Mathematics and Physics (ZAMP)* 30, 292–304.
- Gaussian Process
 - Rasmussen, C.E. and Williams, C.K.I. *Gaussian Processes for Machine Learning*. MIT Press. 2006.
- 数学的理論
 - Berlinet, A. and C. Thomas-Agnan. *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Kluwer Academic Publishers, 2003.
 - Berg, C., J. P. R. Christensen, and P. Ressel. *Harmonic Analysis on Semigroups. Theory of Positive Definite and Related Functions* (Graduate Texts in Mathematics Vol. 100). Springer 1984.
 - Gel'fand, I.M. and Vilenkin, N.Ya. *Generalized Functions Vol.4: Applications of Harmonic Analysis*. Academic Press. 1964.