

カーネル法入門

-正定値カーネルによるデータ解析-

8. 平均による確率分布の特徴づけ

8.1 再生核ヒルベルト空間における平均

08T4082A 西野太樹

可測なカーネル

(X, B_X) を可測空間とする.

この X 上のカーネル k が $X \times X$ 上の可測関数であるとき 可測なカーネル という.

命題8.1 可測空間上の可測な正定値カーネル k に対し, 対応する再生核ヒルベルト空間 H に属する関数はすべて可測関数である.

(証明略)

- 以後、可測空間 (X, B_X) 上の可測な正定値カーネル k に対応する再生核ヒルベルト空間を H_k とする.
- ヒルベルト空間はBorel集合族によって可測集合と考える.

確率空間 (Ω, \mathcal{B}, P) があり、 $X: \Omega \rightarrow \mathcal{X}$ は可測写像とする.

このときカーネル法の特徴写像

$$\Phi: \mathcal{X} \rightarrow H_k, \quad x \mapsto k(\cdot, x)$$

は可測である.

したがって、 $\Phi(X)$ は再生核ヒルベルト空間 H_k に値をとる

確率変数である.

8.1.1 ヒルベルト空間に値をとる確率変数(1)

- H : ヒルベルト空間, F : H に値をとる確率変数
- F の平均を考えるために, Rieszの補題(定理A.5, p.202)を用いる.

H に値をとる確率変数 F に対し, 実数値確率変数 $\|F\|$ が

$$E\|F\| < \infty$$

を満たすと仮定する. このとき H 上の線形汎関数 ϕ_F を

$$\phi_F(f) = E[\langle f, F \rangle] \quad (f \in H)$$

により定義する.

8.1.1 ヒルベルト空間に値をとる確率変数(2)

したがって, *Riesz*の補題により, ある $m_F \in H$ があって
任意の $f \in H$ に対し, $\langle f, m_F \rangle = \phi_F(f)$ が成り立つ.

すなわち,

$$E[\langle f, F \rangle] = \langle f, m_F \rangle \quad (\forall f \in H)$$

が成立する. この $m_F \in H$ を F の平均と呼び, $E[F]$ で表す.

このとき

$$E[\langle f, F \rangle] = \langle f, E[F] \rangle \quad (\forall f \in H)$$

であるので, 形式的に 平均操作と内積が交換可能となる.

8.1.2 再生核ヒルベルト空間における平均(1)

- (χ, B) : 可測空間, $X: \chi$ に値をとる確率変数
- χ 上の可測な正定値カーネルを持つ再生核ヒルベルト空間 (H_k, k) を考える
- 上記に対し, $E[\sqrt{k(X, X)}] < \infty$ を仮定.

8.1.2 再生核ヒルベルト空間における平均(2)

仮定から $E\|\Phi(X)\| < \infty$ なので, $\Phi(X)$ の平均 m_X^k が存在.

これを X の H_k における平均と呼ぶ. 前述の

$$E[\langle f, F \rangle] = \langle f, m_F \rangle \quad (\forall f \in H)$$

より,

$$\langle f, m_X^k \rangle = E[\langle f, \Phi(X) \rangle] = E[f(X)] \quad (\forall f \in H_k)$$

であるので, 任意の $f \in H_k$ に対して期待値 $E[f(X)]$ が f と m_X^k との内積で計算される.

8.1.2 再生核ヒルベルト空間における平均(3)

- 平均 m_X^k の関数として陽な表示

m_X^k は H_k の元なので、再生性より任意の $y \in X$ に対し

$$m_X^k(y) = \langle m_X^k, k(\cdot, y) \rangle = E[k(X, y)]$$

である.

- 平均 m_X^k はカーネル関数の期待値として与えられる.
- 平均 m_X^k は X の分布の高次モーメント情報を持つ.

次に、平均に関してよく用いる関係式を示す。(補題 8.2)

8.1.2 再生核ヒルベルト空間における平均(4)

補題8.2 $(\mathcal{X}, \mathcal{B})$ を可測空間, H_k を \mathcal{X} 上の可測な正定値カーネル k による再生核ヒルベルト空間, X と Y を \mathcal{X} に値をとる確率変数とするとき,

$$\langle m_X^k, m_Y^k \rangle = E[k(Y, X)]$$

が成り立つ. ここで期待値は, X と Y が独立とした場合の分布による. 特に

$$\|m_X^k\|^2 = E[k(X, \tilde{X})]$$

ただし, \tilde{X} は X と独立で同一の分布に従う確率変数とする.
(証明略)

8.1.3 平均と標本平均(1)

- 再生核ヒルベルト空間における平均の推定量を考える

(\mathcal{X}, B, P) を確率空間, X, X_1, \dots, X_n を P に従う*i.i.d.*

サンプルとし, k を \mathcal{X} 上の可測な正定値カーネルとする.

このとき, m_X^k の推定量 $\hat{m}_{(n)}^k$ を

$$\hat{m}_{(n)}^k = \frac{1}{n} \sum_{i=1}^n k(\cdot, X_i) = \frac{1}{n} \sum_{i=1}^n \Phi(X_i)$$

により定義する.

8.1.3 平均と標本平均(2)

定理 8.3 先の仮定のもと,

$$E \left\| \hat{m}_{(n)}^k - m_X^k \right\|_{H_k}^2 = \frac{1}{n} \{ E[k(X, X)] - E[k(X, \tilde{X})] \}$$

(\tilde{X} は X と独立で同一の分布 P に従う確率変数)が
成り立つ. 特に,

$$\left\| \hat{m}_{(n)}^k - m_X^k \right\|_{H_k} = O_p(n^{-1/2}) \quad (n \rightarrow \infty)$$

が成り立つ(証明略) .

8.1.3 平均と標本平均(3)

系8.4 定理8.3と同じ仮定のもと

$$\sup_{f \in H_k, \|f\|_{H_k} \leq 1} \left| \frac{1}{n} \sum_{i=1}^n f(X_i) - E[f(X)] \right| = O_p(n^{-1/2}) \quad (n \rightarrow \infty)$$

が成り立つ.

(証明略)

8.1.3 平均と標本平均(4)

- 次に、中心極限定理に関して考える。

$E[k(X, X)] < \infty$ を仮定すると、任意の $f \in H_k$ に対し

$$E[F(X)^2] = E\left|\langle f, k(\cdot, X) \rangle_{H_k}\right|^2 \leq$$

$$\|f\|_{H_k}^2 E\|k(\cdot, X)\|_{H_k}^2 = \|f\|_{H_k}^2 E[k(X, X)] < \infty$$

により、 $f(X)$ は有限の分散 $V(f)$ を持つ。したがって
中心極限定理

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n f(X_i) - E[f(X)]\right) \Rightarrow N(0, V(f)) \quad (n \rightarrow \infty)$$

が成り立つ。

8.1.3 平均と標本平均(5)

これを内積で書き換えると

$$\left\langle f, \sqrt{n}(m_{(n)}^k - m_X^k) \right\rangle_{H_k} \Rightarrow N(0, V(f)) \quad (n \rightarrow \infty)$$

である.

これは, H_k 上の確率変数 $\sqrt{n}(m_{(n)}^k - m_X^k)$ が何らかの
ガウス確率変数に収束することを示唆する.

実際, 次の定理8.5が成り立つ.

8.1.3 平均と標本平均(5)

定理8.5 $E[k(X, X)] < \infty$ を仮定する.

$G_n = \sqrt{n}(m_{(n)}^k - m_X^k)$ は $n \rightarrow \infty$ のとき, H_k に値をとる確率変数として, H_k 上のガウス確率変数 G に法則収束する.

ここで G は平均0, 共分散関数 $R(f, g) = \text{Cov}[f(X), g(X)]$ により定まる.

(証明略)