カーネル法の基礎 4.3節

参考文献

- 金森, 統計的学習理論, 講談社, 2015.

Table of contents

カーネル関数から生成される内積空間

4.1節の議論を思い出すと，推定量$\hat{f}(x)$はカーネル関数の線形結合で与えられる．
$\rightsquigarrow$ カーネル関数の線形結合で生成される線形空間$\mathcal{H}_0$を定義．

$$\mathcal{H}_0 = \left\{ f(x) = \sum_{i=1}^{m} \alpha_i k(z_i, x) \;\middle|\; \alpha_i \in \mathbb{R},\ z_i \in \mathcal{X},\ m \in \mathbb{N} \right\} \quad (4.4)$$

ここで，$\mathcal{H}_0$は，$\mathcal{X}$上で定義される実数値関数

$$x \mapsto \sum_{i=1}^{m} \alpha_i k(z_i, x)$$

の集合である．

- mが可変
- z_iは入力空間の任意の点

- 回帰関数や判別関数を学習するアルゴリズムの記述には，$\mathcal{H}_0$を定めておけば十分
- 一般に$\mathcal{H}_0$は完備距離空間とならない．
  - 学習結果の統計的性質を調べるときの収束性の議論が困難になる．

$\mathcal{H}_0$に内積を定義する．$\mathcal{H}_0$上の双線形関数を

$$\langle k(x_1,\cdot), k(x_2,\cdot)\rangle = k(x_1,x_2),\ x_1,x_2 \in \mathcal{X} \quad (4.5)$$

で定める．

- $k(x',\cdot) \in \mathcal{H}_0$ は $x \mapsto k(x',x)$ で定義される$\mathcal{X}$上の関数
  - $x' \in \mathcal{X}$: given で，$x \in \mathcal{X}$を与えたときに，$k(x',x)$を返す関数

$\rightsquigarrow$ $\mathcal{X}$上の双線形関数としてwell-definedであることを示す．

(4.5)式より，$f = \sum_i \alpha_i k(x_i, \cdot)$，$g = \sum_j \beta_j k(x_j, \cdot) \in \mathcal{H}_0$に対して，

$$\langle f, g \rangle = \sum_i \sum_j \alpha_i \beta_j k(x_i, x_j) = \sum_j f(x_j) \beta_j = \sum_i g(x_i) \alpha_i.$$

- $\langle f, g \rangle$ の値は関数f, gの関数値のみに依存
  - 線形結合の表し方に依存しない

$\rightsquigarrow$ (4.5)に基づいて$\mathcal{H}_0$の双線形関数が定義される．$g = k(x, \cdot)$とおくと，

$$\langle f, k(x, \cdot) \rangle = f(x),\ f \in \mathcal{H}_0 \quad (4.6)$$

- 再生性 (reproducing property)
  - 再生性を用いて，データ点における関数の評価値から内積を計算することができる

以下，$\langle f, g \rangle$が$\mathcal{H}_0$上の (実数体における) 内積の公理を満たすことを示す．

線形性と対称性は明らかなので，非負値性と非退化性を示す．

非負値性 ($\langle f, f \rangle \geq 0$)
カーネル関数の定義より，$\alpha := (\alpha_1, \ldots, \alpha_n)^{\mathrm{T}}$とおくと，

$$\langle f, f \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j k(x_i, x_j) = \alpha^{\mathrm{T}} K \alpha \geq 0$$

が示されるので，非負値性が示される．

非退化性 ($\langle f, f \rangle = 0 \Rightarrow f = 0$)
非負値性より，$\forall f, g \in \mathcal{H}_0$ および $\forall t \in \mathbb{R}$に対して，

$$\langle f + tg, f + tg \rangle = \langle f, f \rangle + 2t \langle f, g \rangle + t^2 \langle g, g \rangle \geq 0$$

が成り立つ．

tの2次式として見ると，判別式からSchwarzの不等式

$$|\langle f, g\rangle|^2 \leq \langle f, f\rangle\langle g, g\rangle$$

を得る．よって，$\langle f, f\rangle = 0$のとき，再生性 (4.6) とSchwarzの不等式より，$\forall x \in \mathcal{X}$に対して

$$|f(x)|^2 = |\langle f, k(x, \cdot)\rangle|^2 \leq \langle f, f\rangle k(x, x) = 0$$

が成り立つ．よって，$f = 0$である．

- 1つ目の等号: 再生性
- 2つ目の不等号: Schwarz

以上より，$\langle f, g\rangle$が$\mathcal{H}_0$の内積であることが示される．

内積空間の完備化

$(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$: ノルム空間
$\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{H}$がCauchy列
$\overset{\text{def}}{\Longleftrightarrow} \lim_{n,m\to\infty} \|f_n - f_m\|_{\mathcal{H}} = 0$

完備性
$\overset{\text{def}}{\Longleftrightarrow} \forall\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{H}$: Cauchy列, $\exists f \in \mathcal{H}, \lim_{n\to\infty} \|f_n - f\|_{\mathcal{H}} = 0$

完備性をもつノルム空間をBanach空間という.

(4.4)式の$\mathcal{H}_0$は，内積から誘導されるノルムに関して一般に完備ではない.
$\rightsquigarrow$ 完備化という操作により，$\mathcal{H}_0$を稠密に含む完備な内積空間$\mathcal{H}$を構成できる.

Def 4.3 (Hilbert空間)

内積空間$(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$が，内積から誘導されるノルムに関して完備性をもつとき，$\mathcal{H}$をHilbert空間という．

カーネル関数kから生成される内積空間$\mathcal{H}_0$は一般に完備性を満たさないが，完備化して以下の条件を満たすHilbert空間 $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$ を構成できる．

1. 線形写像$j : \mathcal{H}_0 \to \mathcal{H}$が存在し，$\forall f, g \in \mathcal{H}_0$に対して，$\langle f, g\rangle = \langle j(f), j(g)\rangle_{\mathcal{H}}$
   i.e., jは等長写像．
2. $j(\mathcal{H}_0)$は$\mathcal{H}$のなかで稠密, i.e.,
   $\forall f \in \mathcal{H},\ \exists\{f_n\} \subset j(\mathcal{H}_0),\ \lim_{n\to\infty} \|f_n - f\|_{\mathcal{H}} = 0$
   - $\|f\|_{\mathcal{H}} := \sqrt{\langle f, f\rangle_{\mathcal{H}}}, \forall f \in \mathcal{H}$

等長性からjは1対1写像となり，$\mathcal{H}_0$と$j(\mathcal{H}_0)$は内積空間として同一視できる．

完備化

- $\|f\| := \sqrt{\langle f, f \rangle},\ \forall f \in \mathcal{H}_0$
- $\tilde{\mathcal{H}}_0 := \{\{f_n\} \subset \mathcal{H}_0 \mid \lim_{n,m\to\infty} \|f_n - f_m\| = 0\}$
    - $\mathcal{H}_0$のCauchy列全体
    - $a, b \in \mathbb{R}$に対して，$a\{f_n\} + b\{g_n\} := \{af_n + bg_n\}$ とすると，線形空間

$\tilde{\mathcal{H}}_0$に同値関係

$$\{f_n\} \sim \{g_n\} \iff \lim_{n\to\infty} \|f_n - g_n\| = 0$$

を定義し，商空間$\tilde{\mathcal{H}} := \tilde{\mathcal{H}} \backslash \sim$ を定義する．代表元を，

$$[\{f_n\}] := \{\{g_n\} \in \tilde{\mathcal{H}}_0 \mid \{f_n\} \sim \{g_n\}\} \in \tilde{\mathcal{H}}$$

と表す．

線形写像$j : \mathcal{H}_0 \to \tilde{\mathcal{H}}$を，

$$j(f) := [\{f, f, f, \ldots\}]$$

と定め，$\tilde{\mathcal{H}}$における内積を

$$\langle [\{f_n\}], [\{g_n\}] \rangle_{\tilde{\mathcal{H}}} := \lim_{n\to\infty} \langle f_n, g_n \rangle$$

と定める．

$\rightsquigarrow$ $\tilde{\mathcal{H}}$ が$\mathcal{H}_0$を完備化した線形空間

Remark

- $\tilde{\mathcal{H}}$の元は$\mathcal{H}_0$のコーシー列の同値類．
  - このままでは関数ではないので，統計モデルとして考えることができない．
- $\mathcal{H}_0$の内積に関する再生性を用いて，$\tilde{\mathcal{H}}$の元を$\mathcal{X}$上の関数と対応付けできる．

$[\{f_n\}] \in \tilde{\mathcal{H}}$に対して，極限値$\lim_{n\to\infty} f_n(x)$が存在し，それは代表元のとり方に依存しないことを示す．

- 各$x \in \mathcal{X}$に対して，$\{f_n(x)\}_{n\in\mathbb{N}}$は実数上のCauchy列．極限値$\bar{f}(x)$が存在する．
- $\|f_n - g_n\| \to 0$のとき，$g_n(x)$も収束して，極限値は$\bar{f}(x)$となる．
  $\because \mathcal{H}_0$の再生性 (4.6) とSchwarzの不等式より，
  $$|f_n(x) - g_n(x)| = |\langle f_n - g_n, k(x,\cdot)\rangle| \le \|f_n - g_n\|\sqrt{k(x,x)} \to 0$$
  となる．

$\rightsquigarrow [\{f_n\}] \in \tilde{\mathcal{H}}$ を $\mathcal{X}$上の関数

$$x \mapsto \bar{f}(x) = \lim_{n\to\infty} f_n(x)$$

と対応付けることができる．

上記の対応関係は線形かつ1対1.
$\rightsquigarrow$ $\tilde{\mathcal{H}}$を$\mathcal{X}$上の関数からなる線形空間$\mathcal{H}$と同一視できる.

$\mathcal{H}$の内積$\langle \cdot, \cdot \rangle_{\mathcal{H}}$を

$$\langle \bar{f}, \bar{g} \rangle_{\mathcal{H}} := \langle [\{f_n\}], [\{g_n\}] \rangle_{\tilde{\mathcal{H}}} \ \left(= \lim_{n \to \infty} \langle f_n, g_n \rangle \right)$$

とする. このとき,

$$\langle \bar{f}, k(x, \cdot) \rangle_{\mathcal{H}} = \lim_{n \to \infty} \langle f_n, k(x, \cdot) \rangle = \lim_{n \to \infty} f_n(x) = \bar{f}(x)$$

より, $\mathcal{H}$上の再生性が示される.

1. $\mathcal{X}$上のカーネル関数kから内積空間$\mathcal{H}_0$を生成
2. 内積空間$\mathcal{H}_0$を完備化して$\tilde{\mathcal{H}}$を構成
3. $\tilde{\mathcal{H}}$の要素を$\mathcal{X}$上の関数と同一視して, $\mathcal{X}$上の関数からなるHilbert空間$(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$を構成

上記の流れで，カーネル関数からHilbert空間 $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$ を導出できる．

Hilbert空間の例

$\mathcal{L}_2(Q)$

- 2乗可積分関数から定義
- $f, g \in L^2(Q)$に対し，

$$\mathbb{E}_{X\sim Q}\left[\mathbf{1}[f(X) \neq g(X)]\right] = 0$$

のとき，$f \sim g$とする同値関係を導入し，$\mathcal{L}^2(Q) := L^2(Q) \backslash \sim$とする．

Remark

- 同値類では関数の値$f(x)$は意味をもたない．
- カーネル関数から構成される (再生核) Hilbert空間の元は関数と同一視できるので，通常の統計モデルとして用いられる．

再生核Hilbert空間とカーネル関数

- 再生核Hilbert空間の定義
- カーネル関数との対応関係

Def 4.4 (集合$\mathcal{X}$上の再生核Hilbert空間)

集合$\mathcal{X}$上の関数からなるHilbert空間を$(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$とする．
関数$k : \mathcal{X}^2 \to \mathbb{R}$が存在して，任意の$x \in \mathcal{X}$と$f \in \mathcal{H}$に対して，

$$k(x,\cdot) \in \mathcal{H},\ \langle f, k(x,\cdot)\rangle_{\mathcal{H}} = f(x)$$

が成り立つとき，$\mathcal{H}$を再生核Hilbert空間 (RKHS) という．また，関数kを再生核という．

Lemma 4.5

再生核Hilbert空間 $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$ の再生核はカーネル関数．
また，再生核Hilbert空間$\mathcal{H}$の再生核は一意的である．

証明

再生核を$k(x, x')$とする．内積の対称性より

$$k(x, x') = \langle k(x, \cdot), k(x', \cdot) \rangle_{\mathcal{H}} = \langle k(x', \cdot), k(x, \cdot) \rangle_{\mathcal{H}} = k(x', x)$$

となる．また，$x_1, \ldots, x_n \in \mathcal{X}$と$c_1, \ldots, c_n \in \mathbb{R}$に対して，
$\sum_i \sum_j c_i c_j k(x_i, x_j) = \sum_i \sum_j c_i c_j \langle k(x_j, \cdot), k(x_i, \cdot) \rangle_{\mathcal{H}}$
$= \langle \sum_j c_j k(x_j, \cdot), \sum_i c_i k(x_i, \cdot) \rangle_{\mathcal{H}} = \| \sum_j c_j k(x_j, \cdot) \|_{\mathcal{H}}^2 \geq 0$
より，$k(x_i, x_j)$をi, j成分とする$n \times n$行列は対称であり，非負定値．

一意性を示す．

k_1, k_2: $\mathcal{H}$の再生核．対称性より，

$$k_1(x,x') = \langle k_1(x,\cdot), k_2(x',\cdot)\rangle_{\mathcal{H}} = \langle k_2(x',\cdot), k_1(x,\cdot)\rangle_{\mathcal{H}} = k_2(x',x) = k_2(x,x')$$

より，$k_1 = k_2$から一意性が示される．

- 再生核Hilbert空間に対して，再生核となるようなカーネル関数が一意に存在する
- カーネル関数$k : \mathcal{X}^2 \to \mathbb{R}$が与えられたとき，$\mathcal{X}$上の関数からなるHilbert空間$\mathcal{H}$を定義することができる
  - $\mathcal{H}$はkを再生核にもつHilbert空間である
    - $\mathcal{H}$とkの間には再生性が成り立つ

$\rightsquigarrow$ カーネル関数と再生核Hilbert空間の間に1対1対応が存在する．

Th 4.6

$\mathcal{X}$上のカーネル関数$k_1(x, x'), k_2(x, x')$に対して再生核Hilbert空間$\mathcal{H}_1, \mathcal{H}_2$が対応するとする．
このとき，$k_1(x, x') + k_2(x, x')$に対応する再生核Hilbert空間は，

$$\mathcal{H}_1 + \mathcal{H}_2 = \{f : \mathcal{X} \to \mathbb{R} \mid \exists f_1 \in \mathcal{H}_1, \exists f_2 \in \mathcal{H}_2, f = f_1 + f_2\}$$

となる．ここで，

$$\|f\|^2 = \min\left\{ \|f_1\|^2_{\mathcal{H}_1} + \|f_2\|^2_{\mathcal{H}_2} \;\middle|\; f = f_1 + f_2, f_1 \in \mathcal{H}_1, f_2 \in \mathcal{H}_2 \right\}$$

である．

Th 4.6 (続)

また，$k_1(x,x')k_2(x,x')$に対応する再生核Hilbert空間は
$\mathcal{H}_1 \otimes \mathcal{H}_2 =$ ベクトル空間としてのテンソル積
$\{f = \sum_{i=1}^n f_1^{(i)} f_2^{(i)} \mid f_1^{(i)} \in \mathcal{H}_1, f_2^{(i)} \in \mathcal{H}_2\} \in \mathcal{H}_1 \otimes \mathcal{H}_2$
$\langle \sum_{i=1}^n f_1^{(i)} f_2^{(i)}, \sum_{j=1}^m g_1^{(j)} g_2^{(j)} \rangle_{\mathcal{H}_1 \otimes \mathcal{H}_2} = \sum_{i=1}^n \sum_{j=1}^m \langle f_1^{(i)}, g_1^{(i)} \rangle_{\mathcal{H}_1} \langle f_2^{(i)} g_2^{(j)} \rangle_{\mathcal{H}_2}$
となる (参考: Introduction to Kernel Methods, 3. カーネル法の数理的基礎).

カーネル法の性質から，再生核Hilbert空間の元の性質が定まる．
特に，以下の定理が成り立つ．

Th 4.7

有界で連続なカーネル関数$k : \mathcal{X}^2 \to \mathbb{R}$に対応する再生核Hilbert空間$\mathcal{H}$の元は連続関数である．

Th 4.7で，入力空間$\mathcal{X}$は一般の位相空間とする．

Th 4.7の証明

$\mathcal{H}_0$: カーネル関数$k(x, \cdot),\ x \in \mathcal{X}$で張られる線形空間
$\rightsquigarrow$ 完備化．$(\mathcal{H}, \langle \cdot \rangle_{\mathcal{H}})$: RKHS.

- $\mathcal{H}_0$の元は連続関数
- $\mathcal{H}_0$は$(\mathcal{H}, \langle \cdot \rangle_{\mathcal{H}})$で稠密

$\rightsquigarrow \forall f \in \mathcal{H},\ \exists \{f_n\}_{n \in \mathbb{N}} \subset \mathcal{H}_0,\ \|f_n - f\|_{\mathcal{H}} \to 0.$

よって，
$\sup_{x \in \mathcal{X}} |f_n(x) - f(x)| = \sup_{x \in \mathcal{X}} |\langle f_n - f, k(x, \cdot) \rangle_{\mathcal{H}}|\ \ (\because \text{再生性})$
$\leq \|f_n - f\| \sqrt{k(x, \cdot)}\ \ (\because \text{Schwarz})$
$\to 0$
連続関数列$\{f_n\}$がfに一様収束するので，fも連続関数．

Hilbert空間の分類とRKHS

ここでは，任意のHilbert空間はRKHSとして表現できることを証明する．
(数学的にはRKHSとHilbert空間は等価な概念)
以下，Hilbert空間は$\mathbb{R}$上の線形空間とする．

証明の際，次元によるHilbert空間の分類を考える．
$\rightsquigarrow$ Hilbert空間の次元を定義する．

Def 4.8

$(\mathcal{H}, \langle\cdot\rangle_{\mathcal{H}})$: Hilbert空間
$S = \{e_i \mid i \in I\} \in \mathcal{H}$ (Iは適当な添字集合)
次の条件が成り立つとき，Sを$\mathcal{H}$の正規直交基底という．

- 任意の (可分とは限らない) Hilbert空間は正規直交基底をもつ

Def 4.8 (続き)

1. $\forall i \in I,\ \langle e_i, e_i \rangle = 1$
2. $\forall i, j \in I,\ \langle e_i, e_j \rangle = 0\ (i \neq j)$
3. Sで張られる線形空間

$$\left\{ \sum_{k=1}^{n} \alpha_k e_{i_k} \,\middle|\, n \in \mathbb{N}, \alpha_1, \ldots, \alpha_n \in \mathbb{R}, i_i, \ldots, i_n \in I \right\}$$

は$\mathcal{H}$の中で稠密

   - $f \in \mathcal{H}$が任意の$h \in S$に対して$\langle f, h \rangle = 0 \Rightarrow f = 0$ と等価.

このとき，Iの濃度を$\mathcal{H}$の次元という.

Hilbert空間$\mathcal{H}$の正規直交基底の添字集合Iに対して$\ell^2(I)$を

$$\ell^2(I) = \left\{ f : I \to \mathbb{R} \;\middle|\; \sum_{i \in I} |f(i)|^2 < \infty \right\}$$

を定義する.

- $f \in \ell^2(I)$に対して$f(i) \neq 0$になる$i \in I$は高々有限個.
- $\ell^2(I)$は$\langle f, g \rangle = \sum_{i \in I} f(i) g(i)$ を内積とすることでHilbert空間となる.

Th 4.9

1. $\mathcal{H}$と$\ell^2(I)$は等長同型.
2. $\ell^2(I)$はRKHS.

Th 4.9の証明

(1の証明)

(Uが等長かつ単射であること)

- $\{e_i\}_{i \in I} \subset \mathcal{H}$: $\mathcal{H}$の正規直交基底
- $x \in \mathcal{H}$に対して，$U : \mathcal{H} \to \ell(I)$を$Ux : I \to \mathbb{R}, Ux(i) = \langle x, e_i \rangle,\ i \in I$とする
    - $Ux \in \ell^2(I)$となる．

ここで，$x = \sum_{i \in I} \langle x, e_i \rangle e_i$より，

$\|x\|^2 = \sum_{i \in I} |\langle x, e_i \rangle|^2 = \sum_{i \in I} |Ux(i)|^2 \quad (\because \text{definition})$

$= \|Ux\|^2_{\ell^2(I)} \quad (\because \ell^2(I)$における内積の定義)

となる．写像Uは等長的で，Uは単射である．

(Uが全射であること)
定義より，$f \in \ell^2(I)$に対して，$\sum_{i \in I} |f(i)|^2 < \infty$である．
$\rightsquigarrow \sum_{i \in I} f(i) e_i \in \mathcal{H}$ ($\because$ リース・フィッシャーの定理)

この元を$x_0 \in \mathcal{H}$とおくと，

$$Ux_0(i) = \left\langle \sum_{j \in I} f(j) e_j, e_i \right\rangle = f(i) \quad (\because \{e_i\}_{i \in I}\text{が正規直交基底})$$

より，$Ux_0 = f$となる．

(2の証明)
関数$k : I^2 \to \mathbb{R}$を

$$k(i, i') = \begin{cases} 1, & i = i' \\ 0, & i \neq i' \end{cases}$$

と定義すると，明らかに，$k(i, \cdot) \in \ell^2(I)$となる．
また，$f \in \ell^2(I)$に対して，$\langle f, k(i, \cdot) \rangle = f(i)$が成り立つ．
$\rightsquigarrow \ell^2(I)$はカーネル関数$k(i, i')$から生成されるRKHS．

以上より，任意のHilbert空間に対して等長同型なRKHSが存在することが分かる．

Remark

上の対応関係では，入力空間$\mathcal{X}$として正規直交基底の添字集合を用いている．
$\rightsquigarrow$ 必ずしも応用上有用なRKHSの構成法を与えるわけではない．