カーネル法入門

6 章:正定値カーネルの理論

Daiki Tanaka

6.1.1 負定値カーネル

以下のように定める負定値カーネルから正定値カーネルを生成することができる．

定義：負定値

$\mathcal{X}$:set について，$\psi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ が負定値であるとは，ψ がエルミート的で，かつ $\mathcal{X}$ の n 個の任意の点 $x_1, \ldots, x_n$ と $\sum_{i=1}^{n} c_i = 0$ を満たす任意の複素数 $c_1, \ldots, c_n \in \mathbb{C}$ に対して

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

が成り立つことをいう．

[Remark]:負定値カーネルの定義は正定値カーネルの定義:
任意の $n \in \mathbb{N}$，任意の $x_1, \ldots, x_n \in \mathcal{X}$，任意の $c_1, \ldots, c_n \in \mathbb{C}$ に対して

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

の逆符号ではない．(負定値性に関しては制約:$\sum_{i=1}^{n} c_i = 0$ を満たす c_i $(i = 1, \ldots, n)$ に対してだけ考える)

命題 6.1

- (1) k が正定値カーネルならば，$-k$ は負定値カーネル．
- (2) 定数関数は負定値カーネル．
- (3) 任意の関数 f に対して，$\psi(x,y) = f(x) + f(y)$ は負定値カーネル．

[証明]

- (1):k が正定値カーネルである時，任意の $n \in \mathcal{N}$，$x_1, \ldots, x_n \in \mathcal{X}$，$c_1, \ldots, c_n \in \mathbb{C}$ に対し，正値性

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

を満たす．この時，$k' := -k$ は

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

を任意の $c_1, \ldots, c_n \in \mathbb{C}$ について満たすことから，$-k$ は負定値カーネル． □

6.1.1 負定値カーネル

[証明]

- (2):定数関数 $k(\cdot,\cdot) = a$ と $\sum_{i=1}^{n} c_i = 0$ を満たす任意の複素数 $c_1, \ldots, c_n \in \mathbb{C}$ について,

$$
\begin{aligned}
\sum_{i,j=1}^{n} c_i \bar{c_j} k(x_i, x_j) &= \sum_{i,j=1}^{n} c_i \bar{c_j} a \\
&= \sum_{j=1}^{n} \bar{c_j} \left(\sum_{i=1}^{n} c_i a \right) \\
&= 0 \leq 0
\end{aligned}
$$

より, k は負定値カーネル. □

- (3) 任意の関数 f と $\sum_{i=1}^{n} c_i = 0$ を満たす任意の複素数 $c_1, \ldots, c_n \in \mathbb{C}$ に対し,

$$
\begin{aligned}
\sum_{i,j=1}^{n} c_i \bar{c_j} \psi(x_i, x_j) &= \sum_{i,j=1}^{n} c_i \bar{c_j} \left(f(x_i) + f(x_j) \right) \\
&= \sum_{j=1}^{n} \bar{c_j} \left(\sum_{i=1}^{n} c_i f(x_i) \right) + \sum_{i=1}^{n} c_i \left(\sum_{j=1}^{n} \bar{c_j} f(x_j) \right) \\
&= 0 \leq 0
\end{aligned}
$$

より, ψ は負定値カーネル. □

命題 6.2

$\psi_i : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ $(i = 1, \ldots)$ を負定値カーネルとする時，次の 2 つのカーネルも負定値である.

- (1) 非負結合:$\alpha\psi_1 + \beta\psi_2$
- (2) 極限: $\lim_{i\to\infty} \psi_i(x, y)$

ただし，(2) では極限値の存在を仮定した.

$\mathcal{X}$ 上の負定値カーネル全体は閉凸錐である．正定値カーネルとは異なり，負定値カーネルの積は負定値であるとは限らない.

6.1.1 負定値カーネル

負定値カーネルの基本的な例は以下の命題から得られる．

命題 6.3

集合 $\mathcal{X}$ から内積空間 V への写像 $\phi : \mathcal{X} \to V$ について，

$$\psi(x, y) = \|\phi(x) - \phi(y)\|^2$$

は $\mathcal{X}$ 上の負定値カーネルである．

[証明]
$\sum_{i=1}^n c_i = 0$ を満たす任意の複素数 $c_1, \ldots, c_n \in \mathbb{C}$ と $x_1, \ldots, x_n \in \mathcal{X}$ に対して，

$$\sum_{i,j=1}^n c_i \bar{c_j} \|\phi(x) - \phi(y)\|^2 = \sum_{i,j=1}^n c_i \bar{c_j} \langle \phi(x) - \phi(y), \phi(x) - \phi(y) \rangle$$

$$= \sum_{i,j=1}^n c_i \bar{c_j} \left\{ \|\phi(x_i)\|^2 + \|\phi(x_j)\|^2 - \langle \phi(x_i), \phi(x_j) \rangle - \langle \phi(x_j), \phi(x_i) \rangle \right\}$$

$$= 0 + 0 - \left\langle \sum_{i=1}^n c_i \phi(x_i), \sum_{j=1}^n \bar{c_j} \phi(x_j) \right\rangle - \left\langle \sum_{j=1}^n \overline{c_j} \phi(x_j), \sum_{j=1}^n c_i \phi(x_i) \right\rangle$$

$$= - \left\| \sum_{i=1}^n c_i \phi(x_i) \right\|^2 - \left\| \sum_{i=1}^n \overline{c_i} \phi(x_i) \right\|^2 \leq 0$$

6.1.1 負定値カーネル

正定値カーネルと負定値カーネルの間には密接な関連性がある．

補題 6.4

$\psi(x,y)$ を集合 $\mathcal{X} \neq \emptyset$ 上の Hermite 的なカーネル ($\psi(x,y) = \overline{\psi(y,x)}$) とする．
$\forall x_0 \in \mathcal{X}$ に対して，φ を以下のように定義する．

$$\varphi(x,y) := -\psi(x,y) + \psi(x,x_0) + \psi(x_0,y) - \psi(x_0,x_0)$$

この時，ψ が負定値であることと，φ が正定値であることは同値である．

[証明] $x_i \in \mathcal{X}$ を任意の点とする．

- $\Leftarrow$:φ を正定値とする．$\sum_{i=1}^{n} a_i = 0$ を満たす $a_i \in \mathbb{C}$ とすると，正定値性より，

$$\sum_{i,j=1}^{n} a_i \overline{a_j} \varphi(x_i,x_j) \geq 0$$

である．$\sum_{i,j=1}^{n} a_i \overline{a_j} \psi(x_i,x_0) = \sum_{i,j=1}^{n} a_i \overline{a_j} \psi(x_0,x_j) = \sum_{i,j=1}^{n} a_i \overline{a_j} \psi(x_0,x_0) = 0$ であることに注意すると，
$\sum_{i,j=1}^{n} a_i \overline{a_j} \psi(x_i,x_j) \leq 0$ を得られ，ψ は負定値である．

[証明]

- $\Rightarrow$:ψ を負定値とする．$c_i \in \mathbb{C}\ (i = 1, \ldots, n)$ を任意にとり，$c_0 := -\sum_{i=1}^{n} c_i$ とすれば，ψ の負定値性から任意の $x_0, x_1, \ldots, x_n \in \mathcal{X}$ に対して，

$$\sum_{i=0, j=0}^{n} c_i \bar{c}_j \psi\left(x_i, x_j\right) \leq 0$$

が成り立つ．上式の左辺は $i = 0, j = 0$ の場合を外に出せば：

$$\sum_{i=0, j=0}^{n} c_i \bar{c}_j \psi\left(x_i, x_j\right) = \sum_{i=0}^{n} \sum_{j=0}^{n} c_i \bar{c}_j \psi\left(x_i, x_j\right)$$

$$= \sum_{i,j=1}^{n} c_i \bar{c}_j \psi\left(x_i, x_j\right) + \overline{c_0} \sum_{i=1}^{n} c_i \psi\left(x_i, x_0\right) + c_0 \sum_{j=1}^{n} \overline{c_j} \psi\left(x_0, x_j\right) + \left|c_0\right|^2 \psi\left(x_0, x_0\right)$$

$$= \sum_{i,j=1}^{n} c_i \overline{c_j} \psi\left(x_i, x_j\right) - \sum_{i,j=1}^{n} c_i \overline{c_j} \psi\left(x_i, x_0\right) - \sum_{i,j=1}^{n} c_i \overline{c_j} \psi\left(x_0, x_j\right) + \sum_{i,j=1}^{n} c_i \overline{c_j} \psi\left(x_0, x_0\right)$$

$$= -\sum_{i,j=1}^{n} c_i \overline{c_j} \varphi\left(x_i, x_j\right)$$

となって，$\sum_{i,j=1}^{n} c_i \overline{c_j} \varphi\left(x_i, x_j\right) \geq 0$ から φ は正定値である． □

6.1.1 負定値カーネル：カーネル生成に便利な定理

次の定理はカーネルを系統的に生成する際に役立つ.

定理 6.5 Schoenberg の定理

集合 $\mathcal{X} \neq \emptyset$ に対して $\psi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ を $\mathcal{X}$ 上のカーネルとする. この時,

$$\text{「}\psi \text{ が負定値」} \iff \text{「}\exp(-t\psi) \text{ が任意の正数 } t > 0 \text{ に対して正定値」}$$

[証明] 微分の定義から, $\forall x, y \in \mathcal{X}$ に対して,

$$\psi(x, y) = \lim_{t \to +0} \frac{1 - \exp(-t\psi(x, y))}{t}$$

が成り立つ (?). $\forall t > 0$ に対して, $\exp(-t\psi(x, y))$ が正定値ならば,
$\frac{1-\exp(-t\psi(x,y))}{t}$ は負定値であり, その極限値 ψ も負定値である. (命題 6.2)
逆の証明には ψ も負定値である時に $t = 1$ の場合について $\exp(-t\psi(x, y))$ が正定値であることを示せば十分. 任意の $x_0 \in \mathcal{X}$ に対して

$$\varphi(x, y) := -\psi(x, y) + \psi(x, x_0) + \psi(x_0, y) - \psi(x_0, x_0)$$

と定義すると, 補題 6.4 から φ は正定値であり, $\exp(\varphi(x, y))$ も正定値である. 命題 2.5 から

$$\exp(-\psi(x, y)) = \exp(\varphi(x, y)) \exp(-\psi(x, x_0)) \overline{\exp(-\psi(y, x_0))} \exp(\psi(x_0, x_0))$$

は正定値である. □

命題 6.6

集合 $\mathcal{X}$ 上の負定値カーネル $\psi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ が $\psi(x, x) \geq 0$ を満たす時，任意の $0 < p \leq 1$ に対して，

$$\psi(x, y)^p$$

は負定値である．

[証明]$\forall z > 0$ について，ガンマ関数 $\Gamma(z)$ を用いて
$z^p = \frac{p}{\Gamma(1-p)} \int_0^\infty t^{-p-1} \left(1 - e^{-tz}\right) dt$ とできることから，

$$\psi(x, y)^p = \frac{p}{\Gamma(1-p)} \int_0^\infty t^{-p-1} \left(1 - e^{-t\psi(x,y)}\right) dt$$

である．Schoenberg の定理と命題 6.1 により，被積分関数は不定値カーネル．積分がリーマン和の極限であることに注意すれば $\psi(x, y)^p$ も不定値カーネルとなる (命題 6.2 より不定値カーネルの非負結合は不定値カーネ)． □

系 6.7

任意の $0 < p \leq 2$ と $\alpha > 0$ に対して，

$$\exp\left(-\alpha \|x - y\|^p\right)$$

は $\mathbb{R}^n$ 上の正定値カーネルである．

特に，$\alpha = 1, 2$ の時，それぞれラプラスカーネル及びガウス RBF カーネルである．

6.1.2 カーネルを生成する操作

命題 6.8

$\psi : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ を，集合 $\mathcal{X}$ 上の負定値カーネルとする．任意の $x, y \in \mathcal{X}$ について $\psi(x, y) \geq 0$ を満たす時，任意の $\alpha > 0$ について

$$\log(\alpha + \psi(x, y))$$

は負定値カーネルとなる．また，$\psi(x, y) > 0$ である時，

$$\log(\psi(x, y))$$

は負定値カーネルとなる．

[証明]:積分表示

$$\log(1 + \psi(x, y)) = \int_0^\infty \left(1 - e^{-t\psi(x,y)}\right) \frac{e^{-t}}{t} dt$$

により，命題 6.6 と同様に被積分関数が負定値カーネルであることから
$\log(1 + \psi(x, y))$ は負定値カーネルである．したがって，
$\log(\alpha + \psi) = \log\left(1 + \frac{1}{\alpha}\psi\right) + \log\alpha$ も負定値である．

[Remark]:
命題 6.1(3):
「任意の関数 f に対して，$\psi(x, y) = f(x) + f(y)$ は負定値カーネル.」
から，$\psi(x, y) := x + y$ は $\mathbb{R}$ 上の負定値カーネルであるので，
$\psi(x, y) = \log(x + y)$ は $(0, \infty)$ 上の負定値カーネルである．

6.1.2 カーネルを生成する操作

以下の命題を用いれば，負定値カーネルから正定値カーネルを生成できる．

命題 6.9 負定値カーネルから正定値カーネルを生成

負定値カーネル ψ が $\mathrm{Re}\psi\left(x,y\right)\geq 0$ を満たす時，

$$\frac{1}{\psi\left(x,y\right)+a}$$

は正定値カーネルである．ただし，a は正の定数．

[証明]:積分表示

$$\frac{1}{\psi\left(x,y\right)+a}=\int_0^\infty e^{-t(\psi(x,y)+a)}dt$$

より，命題 6.6 と同様にして被積分関数の正定値性から，正定値性をえる． □

6.1.2 カーネルを生成する操作

命題 6.3，命題 6.9 より任意の $0 < p \leq 2$ に対して

$$\frac{1}{1 + |x - y|^p}$$

は $\mathbb{R}$ 上の正定値カーネルである．
特に $p = 2$ の場合は Cauchy 分布の密度関数:

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma\left[1 + \left(\frac{x-x_0}{\gamma}\right)^2\right]} = \frac{1}{\pi} \frac{\gamma}{(x - x_0)^2 + \gamma^2}$$

との類似から Cauchy カーネルと呼ばれる．

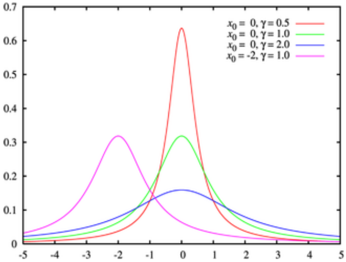

Figure: Cauchy distribution

$\mathbb{R}^n$ 上のカーネル k が平行移動不変である，とは $\mathbb{R}^n$ 上の関数:ϕ があって，
$k(x, y) = \phi(x - y)$ と書けることである．(2 要素の差にのみ依存するカーネル e.g. RBF カーネル) カーネルが平行移動不変であることは
$k(x, y) = k(x + z, y + z)$ $(\forall z \in \mathbb{R}^n)$ と同値である．

定義：正値関数

$\mathbb{R}^n$ 上の関数 ϕ が正値である，とは

$$k(x, y) := \phi(x - y)$$

により定義されるカーネル k が正定値であることをいう．

定理 6.10 Bochner の定理

ϕ を $\mathbb{R}^n$ 上の複素数連続関数とする．この時 ϕ が正値であることの必要十分条件は，$\mathbb{R}^n$ 上の有限な非負 Borel 測度 Λ があって，

$$\phi(x) = \int e^{\sqrt{-1}\omega^\top x} d\Lambda(\omega) \tag{6.1}$$

と表されることである．

- 十分性:

$$\phi(x) = \int e^{\sqrt{-1}\omega^\top x} d\Lambda(\omega)$$

と表されるとする.
$e^{\sqrt{-1}\omega^\top (x-y)} = e^{\sqrt{-1}\omega^\top x} e^{-\sqrt{-1}\omega^\top y} = e^{\sqrt{-1}\omega^\top x}\overline{e^{\sqrt{-1}\omega^\top y}}$ であるから(純虚数 z に対して $-z = \overline{z}$ であることと $\exp(\overline{z}) = \overline{\exp(z)}$ を使った),以下のカーネル:

$$K(x,y) := \phi(x-y) = \int e^{\sqrt{-1}\omega^\top x}\overline{e^{\sqrt{-1}\omega^\top y}} d\Lambda(\omega)$$

の被積分関数は命題 2.5(2) から正定値カーネルである.よって,その積分値として得られる K も正定値カーネルであり,ϕ は正値である. □

- 必要性: 省略.

Bochner の定理は,任意の正値連続関数が $\{e^{\sqrt{-1}\omega^\top x} \mid \omega \in \mathbb{R}^n\}$ の非負結合として表されることを主張している.

平行移動不変な正定値カーネルは周波数領域において陽な形で表現できる．(e.g. RBF カーネル，ラプラスカーネル)
平行移動不変なカーネル K が以下のような形をもつと仮定する．

$$K(x,y) = \int e^{\sqrt{-1}\omega^\top (x-y)} \rho(\omega)\, d\omega$$

ただし，ρ は連続で，$\rho(\omega) > 0$, $\int \rho(\omega)\, d\omega < \infty$.
この時，K を再生核とする RKHS : $\mathcal{H}_K$ は

$$\mathcal{H}_K = \left\{ f \in L^2(\mathbb{R}, dx) \mid \int \frac{\left|\hat{f}(\omega)\right|^2}{\rho(\omega)} d\omega < \infty \right\}$$

$$\langle f, g \rangle = \int \frac{\hat{f}(\omega)\overline{\hat{g}(\omega)}}{\rho(\omega)} d\omega$$

ただし，$\hat{f}$ は f の Fourier 変換 : $\hat{f}(\omega) = \frac{1}{(2\pi)^m} \int f(x)\, e^{-\sqrt{-1}\omega^\top x} d\omega$

6.3 Mercer の定理

連続な正定値カーネルに対する Mercer の定理を導入する．正定値カーネルを積分作用素に関するスペクトル分解によって表現する定理である．

6.3.1 積分核と積分作用素

$(\Omega, \mathcal{B}, \mu)$ を測度空間とし，$K : \Omega \times \Omega \to \mathbb{C}$ を可測なカーネルとする．(K は必ずしも正定値とは限らない．)
6.3 節ではカーネル K に対して常に以下の 2 乗可積分性を仮定する．

$$\int_\Omega \int_\Omega |K(x, y)|^2 \, d\mu(x) \, d\mu(y) < \infty$$

また，$L^2(\Omega, \mu)$(：二乗可積分関数からなる空間) は可分である (稠密な部分集合として可算集合が存在する)：つまり可算正規直交基底をもつ，と仮定する．

L^2 内積

$L^2(\Omega, \mu)$ の内積 (L^2 内積) は

$$\langle f, g \rangle_{L^2} := \int_\Omega f(x) \overline{g(x)} d\mu(x)$$

また，誘導されるノルムは

$$\|f\|_{L^2} := \left(\int_\Omega f(x) \overline{f(x)} d\mu(x) \right)^{\frac{1}{2}} = \left(\int_\Omega |f(x)|^2 \, d\mu(x) \right)^{\frac{1}{2}}$$

である．

カーネル K に対して，$L^2(\Omega, \mu)$ 上の線形作用素
$T_K : L^2(\Omega, \mu) \ni f \mapsto T_K f \in \{g \mid g : \Omega \to \mathbb{C}\}$ を

$$(T_K f)(x) := \int_\Omega K(x, y) f(y)\, d\mu(y), \left(f \in L^2(\Omega, \mu)\right) \tag{6.2}$$

により定義する．ただし，L^p は $\left(\int |f|^p\, d\mu\right)^{\frac{1}{p}} < \infty$ を満たす関数の空間.
この時，Cauchy-Schwarz の不等式により，

$$\begin{aligned}
\int |T_K f(x)|^2\, d\mu(x) &= \int \left\{ \int |K(x, y) f(y)\, d\mu(y)| \right\}^2 d\mu(x) \\
&= \int \langle K(x, \cdot), f \rangle_{L^2}^2\, d\mu(x) \\
&\leq \int \|K(x, \cdot)\|_{L^2}^2 \|f\|_{L^2}^2\, d\mu(x) \\
&= \int \int |K(x, y)|^2\, d\mu(x)\, d\mu(y)\, \|f\|_{L^2}^2 < \infty
\end{aligned}$$

であることから $T_K f \in L^2(\Omega, \mu)$ となり，結局 $T_K : L^2(\Omega, \mu) \to L^2(\Omega, \mu)$ となる．T_K は K を積分核とする積分作用素と呼ばれる.

6.3.1 積分核と積分作用素：Hilbert-Schmidt 作用素

定義：Hilbert-Schmidt 作用素

ヒルベルト空間 $\mathcal{H}_1$ から $\mathcal{H}_2$ への作用素 $A : \mathcal{H}_1 \to \mathcal{H}_2$ が Hilbert-Schmidt 作用素であるとは，$\mathcal{H}_1$ の，ある正規直交基底 $\{\varphi_i\}_{i=1}^{I}, (I \in \mathbb{N} \cup \{\infty\})$ に対し，以下のように定める Hilbert-Schmidt ノルムが

$$\|A\|_{HS}^2 := \sum_{i=1}^{I} \|A\varphi_i\|_{\mathcal{H}_2}^2 < \infty$$

を満たすことである．また，Parseval の等式を用いると，Hilbert-Schmidt ノルムは，$\mathcal{H}_1, \mathcal{H}_2$ の任意の正規直交基底 $\{\varphi_i\}_i, \{\psi_j\}_j$ に対して以下のように表せる：

$$\|A\|_{HS}^2 = \sum_{i=1}^{I} \|A\varphi_i\|_{\mathcal{H}_2}^2 = \sum_{i=1}^{I} \sum_{j=1}^{I'} \langle \psi_j, A\varphi_i \rangle_{\mathcal{H}_2}^2$$

定理 6.11：積分作用素は Hilbert-Schmidt 作用素

2 乗可積分な積分核 K によって，式 6.2 で定まる積分作用素 T_K は Hilbert-Schmidt 作用素であり

$$\|T_K\|_{HS}^2 = \int\int_{\Omega \times \Omega} |K(x, y)|^2 \, d\mu(x) \, d\mu(y)$$

定理 6.11 証明

K の 2 乗可積分性から，ほとんど全ての x に対して，$K(x,\cdot) \in L^2(\Omega,\mu)$ である．$\{\varphi_i\}_{i=1}^{\infty}$ を $L^2(\Omega,\mu)$ の完全正規直交基底とすると，Parseval の等式 (関数解析 p.111) から

$$\begin{aligned}
\int_\Omega |K(x,y)|^2 \, d\mu(y) &= \int_\Omega K(x,y)\overline{K(x,y)} d\mu(y) \quad \left(|z|^2 = z\overline{z}\text{から}\right) \\
&= \|K(x,\cdot)\|_{L^2}^2 \\
&= \sum_{i=1}^{\infty} \left| \langle K(x,\cdot), \varphi_i \rangle_{L^2(\Omega,\mu)} \right|^2 \quad (\text{Parseval の等式から}) \\
&= \sum_{i=1}^{\infty} \left| \int_\Omega K(x,y) \overline{\varphi_i(y)} d\mu(y) \right|^2 \quad \left(L^2\text{内積の定義から}\right) \\
&= \sum_{i=1}^{\infty} \left| T_K \overline{\varphi_i}(x) \right|^2 \quad (\text{積分作用素の定義から})
\end{aligned}$$

が成り立つ．$\{\overline{\varphi_i}\}_{i=1}^{\infty}$ も $L^2(\Omega,\mu)$ の完全正規直交基底であることから

$$\int \int_{\Omega\times\Omega} |K(x,y)|^2 \, d\mu(x) \, d\mu(y) = \sum_{i=1}^{\infty} \left\| T_K \overline{\varphi_i} \right\|^2 = \|T_K\|_{HS}^2$$

をえる． □

6.3.1 積分核と積分作用素

以下の定理は定理 6.11 の逆が成り立つことを主張する．(Hilbert-Schmidt 作用素に対して積分核が一意に存在する)

定理 6.12

$L^2(\Omega, \mu)$ 上の任意の Hilbert-Schmidt 作用素 T に対し，2 乗可積分な積分核 $K(x, y) \in L^2(\Omega \times \Omega, \mu \times \mu)$ が一意に存在し，

$$T\varphi = \int K(x, y)\, \varphi(y)\, d\mu(y) \tag{6.3}$$

となり，$T = T_K$ が成り立つ．

定理 6.11 及び定理 6.12 から，Hilbert-Schmidt 作用素と 2 乗可積分な積分核をもつ積分作用素は一対一に対応する．

[証明]: 存在性を示す．$\{\varphi_i\}_{i=1}^{\infty}$ を $L^2(\Omega, \mu)$ の完全正規直交基底とし，任意の $n \in \mathbb{N}$ に対して

$$K_n(x, y) := \sum_{i=1}^{n} (T\varphi_i)(x)\overline{\varphi_i(y)}$$

と定義する．この時 $\{K_n\}_{n=1}^{\infty}$ は $L^2(\Omega \times \Omega, \mu \times \mu)$ のコーシー列．実際，$m \geq n$ に対して

$$\begin{aligned}
&\int\int_{\Omega\times\Omega} |K_m(x, y) - K_n(x, y)|^2 \, d\mu(x)\, d\mu(y) \\
&= \int\int_{\Omega\times\Omega} \left| \sum_{i=n+1}^{m} (T\varphi_i)(x)\overline{\varphi_i(y)} \right|^2 d\mu(x)\, d\mu(y) \\
&= \sum_{i=n+1}^{m} \sum_{j=n+1}^{m} (T\varphi_i, T\varphi_j)(\varphi_j, \varphi_i) = \sum_{i=n+1}^{m} \|T\varphi_i\|_{L^2(\Omega,\mu)}^2
\end{aligned}$$

であるが，T は Hilbert-Schmidt 作用素なので
$\sum_{i=n+1}^{m} \|T\varphi_i\|_{L^2(\Omega,\mu)}^2 \to 0 \, (n, m \to \infty)$ であり，$\{K_n\}_{n=1}^{\infty}$ は
$L^2(\Omega \times \Omega, \mu \times \mu)$ のコーシー列となる．

$L^2(\Omega \times \Omega, \mu \times \mu)$ は完備であり，2 乗可積分な関数 $K(x, y)$ が $\{K_n\}$ の収束先としてある：$K_n \to K\ (n \to \infty)$．式 (6.2) によって積分作用素 T_K を定義する時，
$T_K = T$ であることを示せば良い．
K_n を積分核とする積分作用素を T_n とすると，定理 6.11 から
$\|T_K - T_n\| \leq \|T_K - T_n\|_{HS} = \|K - K_n\|_{L^2(\Omega\times\Omega,\mu\times\mu)}$ である．すると，任意の $f \in L^2(\Omega, \mu)$ について，一般化フーリエ級数展開：$f = \sum_{i=1}^{\infty} \varphi_i \langle f, \varphi \rangle$ を用いると，

$$\begin{aligned}\|T_K f - Tf\| &= \lim_{n\to\infty} \|T_n f - Tf\| = \lim_{n\to\infty} \left\| \sum_{i=1}^{n} (T\varphi_i) \langle f, \varphi_i \rangle - Tf \right\| \\ &= \lim_{n\to\infty} \left\| T\left(\sum_{i=1}^{n} (f, \varphi_i) \varphi_i \right) - Tf \right\|\end{aligned}$$

となる．$\sum_{i=1}^{n} \langle f, \varphi_i \rangle \varphi_i \to f\ (n \to \infty)$ となるので，上式の最後の極限は 0 となり，
$T_K = T$ をえる．□

[証明]: 一意性を示す．$L^2(\Omega, \mu)$ 上の任意の Hilbert-Schmidt 作用素 T に対して，異なる 2 乗可積分な積分核 $K_1(x, y), K_2(x, y) \in L^2(\Omega \times \Omega, \mu \times \mu)$ が存在し，

$$T\varphi = \int K_1(x, y)\,\varphi(y)\,d\mu(y)$$

$$T\varphi = \int K_2(x, y)\,\varphi(y)\,d\mu(y)$$

が成り立つことを仮定する．つまり，

$$\int (K_1(x, y) - K_2(x, y))\,\varphi(y)\,d\mu(y) = 0$$

であるが，これはほとんど全ての x で $K_1 = K_2$ となることを意味し，仮定に反する． $\square$

6.3.2 Hermite 性から導かれる Hilbert-Schmidt の展開定理

積分核 K が Hermite 的，すなわち $\overline{K(x,y)} = K(y,x)$ を満たすとする．この時，K を積分核とする積分作用素 T_K は自己共役作用素である．実際，

$$\begin{aligned}\langle T_K f, g\rangle_{L^2(\Omega,\mu)} &= \int_\Omega \int_\Omega K(x,y)\, f(y)\, d\mu(y)\, \overline{g(x)} d\mu(x) \\ &= \int_\Omega f(x) \int_\Omega \overline{K(y,x)\, g(x)} d\mu(x)\, d\mu(y) = \langle f, T_K g\rangle\end{aligned}$$

であることが確認できる．自己共役な Hilbert-Schmidt 作用素は固有値分解が可能である．T_K の固有値，固有ベクトルは

$$T_K\phi = \lambda\phi, \left(\lambda \in \mathbb{C}, \phi \in L^2(\Omega,\mu)\right)$$

により定義される．自己共役な Hilbert-Schmidt 作用素の固有値は実数となり，非ゼロ固有値に対する固有ベクトル全体の集合 (固有空間) は有限次元である．固有空間の次元だけ重複を許し $|\lambda_1| \geq |\lambda_2| \geq \cdots > 0$ とし，λ_i に対応する固有ベクトルを ϕ_i とすると，$\{\phi_i\}$ は $L^2(\Omega,\mu)$ の正規直交系 (極大ではない) で，$\forall f \in L^2(\Omega,\mu)$ に対して

$$f = \sum_{i=1}^{\infty} \langle f, \phi_i\rangle_{L^2(\Omega,\mu)} \phi_i + \psi, (\psi \in \mathcal{N}(T_K))$$

$$T_K f = \sum_{i=1}^{\infty} \lambda_i \langle f, \phi_i\rangle_{L^2(\Omega,\mu)} \phi_i$$

と展開できる (Hilbert-Schmidt の展開定理).

6.3.2 積分核の Hilbert-Schmidt 展開

さらに，$L^2(\Omega \times \Omega, \mu \times \mu)$ において $K(x, y)$ は以下のように展開できる．

定理 6.13

Hermite 的な積分核 K に対する積分作用素 T_K の非ゼロ固有値 λ_i と単位固有ベクトル ϕ_i を先述のようにする．この時，$L^2(\Omega \times \Omega, \mu \times \mu)$ において，

$$K(x, y) = \sum_{i=1}^{\infty} \lambda_i \phi_i(x) \overline{\phi_i(y)}$$

の展開が成り立つ．

Fubini の定理

Ω を測度空間とし，$f(x, y)$ が可測かつ可積分であるならば，以下が成立する．

$$\int_\Omega \left(\int_\Omega f(x, y)\, dy \right) dx = \int_\Omega \left(\int_\Omega f(x, y)\, dx \right) dy = \int_{\Omega \times \Omega} f(x, y)\, d(x, y)$$

定理 6.13 証明 1/2

K の 2 乗可積分性から，ほとんど全ての x に対して $K(x,\cdot) \in L^2(\Omega,\mu)$ である．ここで，積分核の定義，及び固有値分解を用いて

$$\left\langle K(x,\cdot), \overline{\phi_i} \right\rangle_{L^2(\Omega,\mu)} = \int K(x,y)\,\phi_i(y)\,d\mu(y) = T_K\phi_i(x) = \lambda_i\phi_i(x)$$

となることに注意する．さらに，$\psi \in \mathcal{N}(T_K)$ ならば

$$\left\langle K(x,\cdot), \overline{\psi} \right\rangle_{L^2(\Omega,\mu)} = T_K\psi(x) = 0$$

となり，測度 μ に関して，ほとんど全ての x に対して一般化フーリエ級数展開を用いると：

$$K(x,\cdot) = \sum_{i=1}^{\infty} \left\langle K(x,\cdot), \overline{\phi_i} \right\rangle_{L^2(\Omega,\mu)} \overline{\phi_i} = \sum_{i=1}^{\infty} \lambda_i\phi_i(x)\,\overline{\phi_i} \tag{6.4}$$

が成立する．Fubini の定理や L^2 ノルムの定義から，

$$\begin{aligned} &\int\int \left| K(x,y) - \sum_{i=1}^{N} \lambda_i\phi_i(x)\,\overline{\phi_i(y)} \right|^2 d\mu(x)\,d\mu(y) \\ &= \int \left\| K(x,\cdot) - \sum_{i=1}^{N} \lambda_i\phi_i(x)\,\overline{\phi_i} \right\|_{L^2(\Omega,\mu)}^2 d\mu(x) \end{aligned} \tag{6.5}$$

となる．

式 6.5 の右辺の被積分関数は式 6.4 からほとんど全ての x に対して $N \to \infty$ で 0 に収束．かつ，

$$
\begin{aligned}
&\left\| K(x,\cdot) - \sum_{i=1}^{N} \lambda_i \phi_i(x) \overline{\phi_i} \right\|_{L^2(\Omega,\mu)}^2 \\
&= \left\langle K(x,\cdot) - \sum_{i=1}^{N} \lambda_i \phi_i(x) \overline{\phi_i}, K(x,\cdot) - \sum_{i=1}^{N} \lambda_i \phi_i(x) \overline{\phi_i} \right\rangle_{L^2(\Omega,\mu)} \\
&= \int |K(x,y)|^2 d\mu(y) - \sum_{i=1}^{N} \lambda_i^2 |\phi_i(x)|^2 \le \int |K(x,y)|^2 d\mu(y)
\end{aligned}
$$

と N によらない 2 乗可積分関数で上から抑えることができ，優収束定理によって式 (6.5) は $N \to \infty$ で 0 に収束する． □

6.3.3 正値積分核と Mercer の定理

定理 6.13 ではカーネル K の Hermite 性のみを仮定して K の $L^2(\Omega \times \Omega, \mu \times \mu)$ における展開を示した．さらに μ がコンパクト Hausdorff 空間 Ω 上の有限 Radon 測度で，K が Ω 上の連続な正定値カーネルの場合には以下に示すようにこの収束は絶対かつ一様である (Mercer の定理).

まず K の正定値性と積分作用素 T_K の正値性:

$$\langle T_K f, f\rangle_{L^2(\Omega,\mu)} = \int_\Omega \int_\Omega K(x, y) f(x) \overline{f(y)} d\mu(x) d\mu(y) \geq 0, \left(\forall f \in L^2(\Omega, \mu)\right) \tag{6.6}$$

の関連を示す.

位相空間 Ω 上の非負 Borel 測度 μ のサポート:$\mathrm{Supp}(\mu)$ を

$$\mathrm{Supp}(\mu) := \{x \in \Omega \mid \mu(U) > 0, U \text{ は } x \text{ を含む任意の開集合 }\}$$

により定義する．サポートは閉集合である．また，$\mathbb{R}^n$ 上の Borel 測度 μ が連続な確率密度 $p(x)$ を持つ時，μ のサポートと $p(x)$ のサポートは一致する.
Hausdorff 空間上の Borel 測度 μ が Radon 測度である，とは任意のコンパクト融合 K に対して $\mu(K) < \infty$ で，任意の可測集合 E に対して
$\mu(E) = \sup\{\mu(K) \mid K \text{ は } K \subset E \text{ なるコンパクト集合 }\}$ が成り立つことである.

Hausdorff 空間

位相空間 S の相異なる任意の 2 点，x_1, x_2 に対して，$U(x_1) \cap U(x_2) = \emptyset$ を満たす開集合 $U(x_1), U(x_2)$ が存在する時，S は Hausdorff 空間，あるいは T_2 空間であるという．

命題 6.14

Ω をコンパクト Hausdorff 空間，μ を Ω 上の有限非負 Radon 測度で，
$\mathrm{Supp}\,(\mu) = \Omega$ とする．
$K\,(\cdot,\cdot) : \Omega \times \Omega \to \mathbb{C}$ を，Ω 上の連続な Hermite 的なカーネルとする時，K が正定値カーネルである必要十分条件は T_K が正値作用素であることである．

証明：

- 必要性:K が正定値カーネルであるとする．Ω 上の任意の連続関数 $g : \Omega \to \mathbb{C}$ と可測集合による Ω の分割 $\{E_i\}_{i=1}^{n}$，(ただし $i \neq j$ ならば $E_i \cap E_j = \emptyset, \cup_{i=1}^{n} E_i = \Omega$) に対して，$K$ の正定値性から，

$$\sum_{i,j=1}^{n} K\,(x_i, x_j)\, g\,(x_i)\, \overline{g\,(x_j)} \mu\,(E_i)\, \mu\,(E_j) \geq 0$$

は成り立つ．式 6.6 : $\langle T_K f, f \rangle$ はこのような和の極限として得られるため，非負である．一般の $f \in L^2\,(\Omega, \mu)$ に対しては任意の $\epsilon > 0$ に対して $\|f - g\|_{L^2(\Omega,\mu)} < \epsilon$ なる連続関数 g をとれば式 6.6 が示される．

- 十分性: T_K が正値作用素であるとする．$\mu = 0$ ならば自明に正定値性は満たされる．そのため，$\mu(\Omega) > 0$ としてよい．背理法で示すために，ある $x_i \in \Omega$ と $c_i \in \mathbb{C}$ と $\delta > 0$ があって，

$$\sum_{i,j=1}^{n} c_i \overline{c_j} K(x_i, x_j) \leq -\delta$$

が成り立つと仮定する (T_K が正値作用素である時に K が正定値カーネルでないことを仮定する)．一般性を失わず，$x_i \neq x_j$ としてよい．K の連続性と Ω の Housdorff 性により，各 x_i の開近傍 U_i があって，$U_i \cap U_j = \emptyset\,(i \neq j)$ かつ，$\forall (x, y) \in U_i \times U_j$ に対して $\left| c_i \overline{c_j} K(x, y) - c_i \overline{c_j} K(x_i, x_j) \right| \leq \frac{\delta}{2n^2}$ とできる (十分小さくすれば取れる)．この時，$\mathrm{Supp}(\mu) = \Omega$ から $\mu(U_i) > 0$ である．

$$f := \sum_{i=1}^{n} \left(\frac{c_i}{\mu(U_i)} \right) I_{U_i} \in L^2(\Omega, \mu)$$

とおく (ただし I_{U_i} は U_i の定義関数)．

このとき T_K の正値性をチェックすると

$$
\begin{aligned}
&\int\int_{\Omega\times\Omega} K(x,y)\, f(x)\, \overline{f(y)} d\mu(x)\, d\mu(y) \\
&\leq \sum_{i,j=1}^{n} \frac{1}{\mu(U_i)\,\mu(U_j)} \int\int_{U_i\times U_j} \left| c_i\overline{c_j} K(x,y) - c_i\overline{c_j} K(x_i,x_j) \right| d\mu(x)\, d\mu(y) \\
&+ \sum_{i,j=1}^{n} \int\int_{U_i\times U_j} \frac{c_i\overline{c_j}}{\mu(U_i)\,\mu(U_j)} K(x_i,x_j)\, d\mu(x)\, d\mu(y) \\
&\leq \frac{\delta}{2} - \delta < 0
\end{aligned}
$$

となり，仮定に矛盾する．□

6.3.3 正値積分核と Mercer の定理

Ω と μ を命題 6.14 のとおりとする．連続な正定値カーネル K によって定まる積分核を T_K とする時，命題 6.14 から T_K は正値作用素で，任意の $f \in L^2(\Omega, \mu)$ に対して，$(T_K f, f) \geq 0$ が成り立つので T_K の固有値は非負実数である．重複度のぶんだけ並べた正の固有値と固有ベクトルをそれぞれ

$$\lambda_1 \geq \lambda_2 \geq \cdots > 0$$

及び ϕ_i で表す．

定理 6.15 Mercer の定理

$$K(x, y) = \sum_{i=1}^{\infty} \lambda_i \phi_i(x) \overline{\phi_i(y)} \tag{6.7}$$

ここで，収束は $\Omega \times \Omega$ 上の絶対かつ一様収束である．

定理 6.13 から，式 6.7 は $L^2(\Omega \times \Omega, \mu \times \mu)$ の収束として成り立つ．また，

$$\lambda_i \phi_i(x) = T_K \phi_i(x) = \int K(x, y) \phi_i(y) d\mu(y)$$

から K の連続性から ϕ_i も連続である．ここで，

$$K_{n,m}(x, y) := \sum_{i=n}^{m} \lambda_i \phi_i(x) \overline{\phi_i(y)}, \quad R_n(x, y) := K(x, y) - K_{1,n-1}(x, y)$$

とおくと，$K_{n,m}(\cdot, \cdot)$ は正定値カーネルであることから．命題 6.14 から正値積分核となる．また，定理 6.13 から $L^2(\Omega \times \Omega, \mu \times \mu)$ において

$$\begin{aligned}\lim_{m\to\infty} K_{n,m}(x, y) &= \sum_{i=n}^{\infty} \lambda_i \phi_i(x) \overline{\phi_i(y)} \\ &= \sum_{i=1}^{\infty} \lambda_i \phi_i(x) \overline{\phi_i(y)} - \sum_{i=1}^{n-1} \lambda_i \phi_i(x) \overline{\phi_i(y)} \\ &= K(x, y) - K_{1,n-1}(x, y) = R_n(x, y)\end{aligned}$$

であり，正値積分核であるという性質は $L^2(\Omega \times \Omega, \mu \times \mu)$ の収束に関して保存されることから，収束先 R_n も正値積分核である．R_n が正値積分核であることから，命題 6.14 を用いれば $R_n(x, y)$ は正定値カーネルである．

この時，正定値カーネルの定義から $\forall x \in \Omega$ と $\forall n \in \mathbb{N}$ に対し，$R_n(x, x) \geq 0$ である．$K(x, x) = R_n(x, x) + K_{1,n-1}(x, x)$ から，
$K_{1,n-1}(x, x) = \sum_{i=1}^{n-1} \lambda_i |\phi_i(x)|^2 \leq K(x, x)$ が成り立ち，

$$\sum_{i=1}^{\infty} \lambda_i |\phi_i(x)|^2 \leq K(x, x) \tag{6.8}$$

をえる．また，Cauchy-Schwarz の不等式から，

$$\begin{aligned} \sum_{i=n}^{m} \left| \lambda_i \phi_i(x) \overline{\phi_i(y)} \right| &\leq \left(\sum_{i=n}^{m} \lambda_i |\phi_i(x)|^2 \right)^{\frac{1}{2}} \left(\sum_{i=n}^{m} \lambda_i |\phi_i(y)|^2 \right)^{\frac{1}{2}} \\ &= K_{n,m}(x, x)^{\frac{1}{2}} K_{n,m}(y, y)^{\frac{1}{2}} \end{aligned} \tag{6.9}$$

である．式 6.8 から，式 6.7 の右辺の級数は各 (x, y) で絶対収束する．この収束先を

$$H(x, y) := \sum_{i=1}^{\infty} \lambda_i \phi_i(x) \overline{\phi_i(y)} \tag{6.10}$$

とおいて，$H = K$ となることを示す．

まず，式 6.9 から

$$\sum_{i=n}^{m} \left| \lambda_i \phi_i(x) \overline{\phi_i(y)} \right| \leq K_{n,m}(x,x)^{\frac{1}{2}} K_{n,m}(y,y)^{\frac{1}{2}} \leq K_{n,m}(x,x)^{\frac{1}{2}} \sup_{z \in \Omega} K(z,z)^{\frac{1}{2}}$$

から，任意に $x \in \Omega$ を固定すれば式 6.10 は $y \in \Omega$ に関して一様に収束する．ゆえに任意の x に対し，

$$\begin{aligned} \int H(x,y)\,\phi_i(y)\,d\mu(y) &= \sum_{n=1}^{\infty} \int \lambda_n \phi_n(x)\,\overline{\phi_n(y)}\phi_i(y)\,d\mu(y) \\ &= \lambda_i \phi_i(x) = \int K(x,y)\,\phi_i(y)\,d\mu(y) \end{aligned}$$

が成り立つ (?)．これは $L^2(\Omega, \mu)$ において $H(x,\cdot) = K(x,\cdot)$ を意味するが，両者は連続関数であることから $H(x,y) = K(x,y)$ を得る．

最後に．Mercer の定理の式が $\Omega \times \Omega$ 上の一様収束であることを示す．各点収束であることは見たので，特に $y = x$ とおくと，

$$\sum_{i=1}^{\infty} \lambda_i \left|\phi_i(x)\right|^2 = K(x, x)$$

が成り立つ．各 $\phi_i(x)$ は連続関数であり，左辺は単調に $K(x, x)$ に収束するので，Dini の定理 (コンパクト集合上の連続関数の単調列がある連続関数に各点収束するならば収束は一様) からこの収束は一様である．したがって，$n, m \to \infty$ の時，$K_{n,m}(x, x)$ は 0 に一様収束し，式 6.9 から $K_{n,m}(x, y)$ は 0 に一様収束する．これは式 6.7 の絶対かつ一様な収束を意味する．(x,y によらず収束することから) □

正定値カーネルに対応する RKHS の陽な表示

Mercer の定理を用いると，正定値カーネルに対応する RKHS と，その内積の陽な表示を与えることができる．(2.2.2.b:有限集合上の RKHS の表示の拡張として与えることができる)

Mercer の定理と同じ条件の元，積分作用素 T_K の非ゼロ固有値に対応する単位固有ベクトルに $\mathcal{N}(T_K)$ の正規直交基底を付け加えて $L^2(\Omega, \mu)$ の完全正規直交基底 $\{\phi_i\}_{i=1}^{\infty}$ を構成する．すると，$\{\phi_i\}_{i=1}^{\infty}$ はシャウダー基底となり，任意の $f \in \mathcal{H}$ はそれらの基底の線型結合として：

$$f = \sum_{i=1}^{\infty} a_i \phi_i \ (\text{ただし } a_i \in \mathbb{R})$$

とできる．これを用いて $L^2(\Omega, \mu)$ の部分ベクトル空間 $\mathcal{H}$ を

$$\mathcal{H} := \left\{ f \in L^2(\Omega, \mu) \mid f = \sum_{i=1}^{\infty} a_i \phi_i, \sum_{i=1}^{\infty} \frac{|a_i|^2}{\lambda_i} < \infty \right\} \tag{6.11}$$

により定義する．また，$f = \sum_{i=1}^{\infty} a_i \phi_i \in \mathcal{H}$ と $g = \sum_{i=1}^{\infty} b_i \phi_i \in \mathcal{H}$ に対して，内積を

$$\langle f, g \rangle_{\mathcal{H}} := \sum_{i=1}^{\infty} \frac{a_i \overline{b_i}}{\lambda_i} \tag{6.12}$$

と定める．このように定めた $\mathcal{H}$ が K を再生核とする RKHS になることを示す．

まず $\mathcal{H}$ がヒルベルト空間である (完備である) ことを示す．
$\{f_n\}_{n=1}^{\infty}$ を $\mathcal{H}$ の Cauchy 列とする．式 6.11 から $f_n = \sum_{i=1}^{\infty} \alpha_{n,i}\phi_i$ とおくことができて，$\sum_{i=1}^{\infty} \frac{|\alpha_{n,i}|^2}{\lambda_i} < \infty$ であるので，

$$t_n := \left\{ \frac{\alpha_{n,i}}{\sqrt{\lambda_i}} \right\}_{i=1}^{\infty}$$

は数列空間 l^2 の Cauchy 列である．l^2 は完備なので，ある $t = \{\beta_i\}_{i=1}^{\infty} \in l^2$ が存在し，$t_n \to t\,(n \to \infty)$ である．この時，$\alpha_i^* := \sqrt{\lambda_i}\beta_i$ とすると，

$$\sum_{i=1}^{\infty} \frac{|\alpha_i^*|^2}{\lambda_i} = \sum_{i=1}^{\infty} |\beta_i|^2 < \infty \ (\{\beta_i\}_{i=1}^{\infty} \in l^2 \text{から})$$

$$\sum_{i=1}^{\infty} \frac{|\alpha_{n,i} - \alpha_i^*|^2}{\lambda_i} \to 0$$

が成り立つ．すると，$f = \sum_{i=1}^{\infty} \alpha_i^*\phi_i \in \mathcal{H}$ に対して，$\|f_n - f\|_{\mathcal{H}} \to 0$ を得る．

$\mathcal{H}$ が正定値カーネルに対応する RKHS であること：証明 2

次に，$\mathcal{H}$ が K を再生核に持つことを示す．
$K(\cdot, x) = \sum_{i=1}^{\infty} \lambda_i \phi_i(\cdot) \overline{\phi_i(x)}$ において，Mercer の定理から
$\sum_{i=1}^{\infty} \lambda_i |\phi_i(x)|^2 < \infty$ であるので $K(\cdot, x) \in \mathcal{H}$ である．$\forall f = \sum_{i=1}^{\infty} a_i \phi_i \in \mathcal{H}$
に対し，

$$\langle f, K(\cdot, x) \rangle_{\mathcal{H}} = \sum_{i=1}^{\infty} \frac{\alpha_i \lambda_i \phi_i(x)}{\lambda_i} = \sum_{i=1}^{\infty} \alpha_i \phi_i(x) = f(x)$$

となり，再生性が確認できた． □

以上から，正定値カーネルに対応する RKHS は式 6.11 で与えられる $\mathcal{H}$ に一致し，その内積は式 6.12 のように，級数として与えられる．

ヒルベルト空間上の確率変数の平均

$\mathcal{H}$：ヒルベルト空間 (関数空間)，$F : \mathcal{H}$ 上の確率変数．ただし，$E[\|F\|] < \infty$．
この時，$f \in \mathcal{H}$ に対して $\mathcal{H}$ 上の線形汎関数 $\phi_F : \mathcal{H} \to \mathbb{R}$ を以下のように定める：

$$\phi_F(f) := E[\langle f, F \rangle]$$

リースの表現定理から，任意の $f \in \mathcal{H}$ に対して，ある $m_F \in \mathcal{H}$ が存在し，
$\langle f, m_F \rangle = \phi_F(f)$ が成り立つ．よって

$$\phi_F(f) = E[\langle f, F \rangle] = \langle f, m_F \rangle \tag{8.1}$$

をえる．この m_F を確率変数 F の平均と呼び，$E[F]$ で表す．この時，

$$E[\langle f, F \rangle] = \langle f, m_F \rangle = \langle f, E[F] \rangle$$

となり，平均と内積の操作は交換可能である．

$(\mathcal{X}, \mathcal{B})$: 可測空間，$X : \mathcal{X}$ に値をとる確率変数，RKHS : $(\mathcal{H}_k, k)$ を考える．ただし，$E[\sqrt{k(X, X)}] < \infty$ を仮定する．
特徴写像 $\Phi(x) = k(\cdot, x)$ に対して，再生性から

$$\|\Phi(X)\|^2 = \langle k(\cdot, X), k(\cdot, X)\rangle = \langle k_X, k_X\rangle = k_X(X) = k(X, X)$$

が成り立つことに注意すれば，$E\|\Phi(X)\| < \infty$ となり，前項の仮定を満たし，確率変数 $\Phi(X)$ の平均 m_X^k が存在する．この時，m_X^k を，X の $\mathcal{H}_k$ における平均，と呼ぶ．式 (8.1) および再生性から，任意の $f \in \mathcal{H}$ に対して

$$\left\langle f, m_X^k \right\rangle = E[\langle f, \Phi(X)\rangle] = E[\langle f, K(X, \cdot)\rangle] = E[f(X)] \tag{8.2}$$

となり，任意の f に対して期待値 $E[f(X)]$ が f と m_X^k の内積で表される．
平均 m_X^k の陽な表示を求める．$m_X^k \in \mathcal{H}$ から，任意の $y \in \mathcal{X}$ について，再生性を用いると

$$\begin{aligned} m_X^k(y) &= \left\langle m_X^k, k(\cdot, y)\right\rangle = \langle E(\Phi(X)), k(\cdot, y)\rangle \\ &= E\langle k(\cdot, X), k(\cdot, y)\rangle = E[k(X, y)] \end{aligned} \tag{8.8}$$

となって，平均 m_X^k はカーネル関数の期待値として与えられる．

RKHS における平均の推定量

$X_1, \ldots, X_n$ を P に従う i.i.d. サンプルとし，k を $\mathcal{X}$ 上の可測なカーネル．m_X^k の推定量 $\widehat{m}_{(n)}^k$ を以下のように定義する：

$$\widehat{m}_{(n)}^k := \frac{1}{n}\sum_{i=1}^{n} k\left(\cdot, X_i\right)$$

これが m_X^k の不偏推定量であることは明らか.

特性的なカーネル

$(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$:可測空間，$\mathcal{P}$:確率測度全体とする．$\mathcal{X}$ 上の有界かつ可測な正定値カーネル k が特性的である，とは，写像：

$$\mathcal{P} \to \mathcal{H}_k, \quad P \mapsto m_P^k$$

が単射であることをいう．ここで，m_P^k は分布 P をもつ $\mathcal{X}$ 上の確率変数の $\mathcal{H}_k$ における平均を表す．

この定義は，任意の $f \in \mathcal{H}_k$ について

$$E_{X \sim P}[f(X)] = E_{X \sim Q}[f(X)] \Rightarrow P = Q$$

と同値である．