フォック空間とガウス型核に対する
普遍近似定理

防衛大学校・数学教育室　瀬戸　道生
Michio Seto
Department of Mathematics,
National Defense Academy

1　はじめに

ガウス過程回帰は機械学習において非常に汎用性のある強力な手法である（例えば，ビショップ [1], 持橋–大羽 [5], S–伊吹–畑中 [9] を参照せよ）．そのガウス過程回帰の数学的保証を与える次の定理は普遍近似定理としてよく知られている．

定理 1.1. $\gamma > 0$ を定数，K を $\mathbb{R}^n$ 内の任意のコンパクト集合とする．K 上で定義された任意の実数値連続関数 f と任意の $\varepsilon > 0$ に対し，

$$\sup_{\boldsymbol{x}\in K}\left| f(\boldsymbol{x}) - \sum_{j=1}^{N} c_j e^{-\gamma\|\boldsymbol{x}-\boldsymbol{a}_j\|^2_{\mathbb{R}^n}} \right| < \varepsilon$$

をみたす $N \geq 1$, $c_1, \ldots, c_N \in \mathbb{R}$, $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N \in K$ が存在する．

ガウス過程回帰では，予測はガウス核 $e^{-\gamma\|\boldsymbol{x}-\boldsymbol{a}\|^2_{\mathbb{R}^n}}$ $(\boldsymbol{a} \in \mathbb{R}^n)$ の線形結合で与えられる．従って，普遍近似定理はガウス過程回帰により応用上十分広い範囲の関数を扱えることを述べている．

さて，普遍近似定理には複数の証明法が知られている (Steinwart [10], Micchelli–Xu–Zhang [4], Guella [2])．この小論では，Kuwahara–S [3] と S [8] で与えた手法から一種の非有界関数環を考えると，普遍近似定理の比較的簡単な証明が得られることを紹介する．

2　準備

この節では，再生核ヒルベルト空間に対する専門家の間ではよく知られた結果をまとめる（実際，Nikolski [6] の p. 320 では演習問題になっている）．以下，係

数体 $\mathbb{K}$ は $\mathbb{R}$ でも $\mathbb{C}$ でもよい．まず，次の定理は多くの証明の中に埋め込まれている．

定理 2.1. $\mathcal{H}$ をヒルベルト空間，$\mathcal{V}$ をベクトル空間とする．$\mathcal{H}$ から $\mathcal{V}$ への線形写像 T の核 $\ker T$ が閉ならば，T の値域 $\operatorname{ran} T$ は

$$\langle Tx, Ty\rangle_T = \langle P_{(\ker T)^\perp}x, P_{(\ker T)^\perp}y\rangle_{\mathcal{H}}$$

を内積とするヒルベルト空間になる．

Proof. 線形写像に対する準同型定理とヒルベルト空間論における射影定理により，

$$\operatorname{ran} T \cong \mathcal{H}/\ker T \cong (\ker T)^\perp$$

が成り立つ． □

この定理 2.1 から，引き戻し構成法とよばれるヒルベルト空間の構成法が得られる．引き戻し構成法の一般論については Paulsen–Raghupathi [7] の Section 5.5 と Chapter 7 にまとめられている．ここでは，対角線写像と指数写像について考える．

2.1 対角線写像

まず，X を集合とし，$\mathcal{H}_k$ を X 上の再生核ヒルベルト空間とする．$\mathcal{H}_k$ の n 重テンソル積 $\mathcal{H}_k^{\otimes n}$ と n 次元対角線写像

$$\Delta_n : X \to X^n, \quad x \to (x, \dots, x)$$

に対し，$\mathcal{H}_k^{\otimes n}$ から X 上の関数全体からなるベクトル空間への線形写像 Δ_n を

$$(\Delta_n F)(x) = F \circ \Delta_n(x) = F(x, \dots, x) \quad (F \in \mathcal{H}_k^{\otimes n},\ x \in X)$$

と定める．$\Delta_n F$ に対し，本来は $\Delta_n^* F$ のような記号を使うべきかもしれないが，このまま話を進める．今，

$$(\Delta_n F)(x) = F(x, \dots, x) = \langle F, k_x^{\otimes n}\rangle_{\mathcal{H}_k^{\otimes n}} \tag{2.1}$$

から $\ker \Delta_n := \{F \in \mathcal{H}_k^{\otimes n} : \Delta_n F = 0\}$ は閉であることがわかる．従って，$\mathcal{H}_k^{\otimes n}$ と Δ_n に定理 2.1 を適用できる．

定義 2.2. Δ_n を用いた引き戻し構成法により定まるヒルベルト空間を $\mathcal{H}_k^n$ と表す．すなわち，内積

$$\langle \Delta_n F, \Delta_n G\rangle_{\mathcal{H}_k^n} = \langle P_{(\ker \Delta_n)^\perp}F, P_{(\ker \Delta_n)^\perp}G\rangle_{\mathcal{H}_k^{\otimes n}}$$

により，$\Delta_n \mathcal{H}_k^{\otimes n}$ はヒルベルト空間になる．

$\mathcal{H}_k^n$ の性質を以下にまとめる．

命題 2.3. $\mathcal{H}_k^n$ は $\Delta_n k_x^{\otimes n} = k_x^n$ を再生核とする再生核ヒルベルト空間である．また，$\|k_x^n\|_{\mathcal{H}_k^n}^2 = \|k_x\|_{\mathcal{H}_k}^{2n}$ が成り立つ．

Proof. まず，(2.1) から，$k_x^{\otimes n} \in (\ker \Delta_n)^\perp$ がわかる．よって，$\mathcal{H}_k^n$ の内積の定め方から，

$$\begin{aligned}
\langle \Delta_n F, k_x^n \rangle_{\mathcal{H}_k^n} &= \langle P_{(\ker \Delta_n)^\perp} F, P_{(\ker \Delta_n)^\perp} k_x^{\otimes n} \rangle_{\mathcal{H}_k^{\otimes n}} \\
&= \langle F, k_x^{\otimes n} \rangle_{\mathcal{H}_k^{\otimes n}} \\
&= F(x, \ldots, x) \\
&= (\Delta_n F)(x)
\end{aligned}$$

が成り立つ．同様に，

$$\begin{aligned}
\|k_x^n\|_{\mathcal{H}_k^n}^2 &= \langle k_x^n, k_x^n \rangle_{\mathcal{H}_k^n} \\
&= \langle P_{(\ker \Delta_n)^\perp} k_x^{\otimes n}, P_{(\ker \Delta_n)^\perp} k_x^{\otimes n} \rangle_{\mathcal{H}_k^{\otimes n}} \\
&= \langle k_x^{\otimes n}, k_x^{\otimes n} \rangle_{\mathcal{H}_k^{\otimes n}} \\
&= \|k_x\|_{\mathcal{H}_k}^{2n}
\end{aligned}$$

が成り立つ． □

2.2 指数写像

次に，重みを付けた内積

$$\langle (f_0, f_1, \ldots), (g_0, g_1, \ldots) \rangle_{\mathcal{F}} = \sum_{n=0}^{\infty} \frac{1}{n!} \langle f_n, g_n \rangle_{\mathcal{H}_k^n} \quad (f_n, g_n \in \mathcal{H}_k^n)$$

により定まるヒルベルト空間を $\mathcal{F}$ とし，写像

$$\Gamma(f_0, f_1, \ldots)^\top = \sum_{n=0}^{\infty} \frac{1}{n!} f_n \quad ((f_0, f_1, \ldots)^\top \in \mathcal{F})$$

を考える．ただし，$\mathcal{H}_k^0 = \mathbb{K}$ とおいた．この Γ の性質を補題としてまとめておく．

補題 2.4. Γ は $\mathcal{F}$ から X 上の関数全体からなるベクトル空間への線形写像であり，$\ker \Gamma$ は $\mathcal{F}$ の閉部分空間である．

Proof. まず，任意の $F = (f_0, f_1, \ldots)^\top \in \mathcal{F}$ に対し，命題 2.3 から，

$$\begin{aligned}
\left|\sum_{j=n}^{m} \frac{1}{j!} f_j(x)\right| &\leq \sum_{j=n}^{m} \frac{1}{j!} |f_j(x)| \\
&\leq \sum_{j=n}^{m} \frac{1}{j!} \|f_j\|_{\mathcal{H}_k^j} \|k_x^j\|_{\mathcal{H}_k^j} \\
&\leq \left(\sum_{j=n}^{m} \frac{1}{j!} \|f_j\|_{\mathcal{H}_k^j}^2\right)^{1/2} \left(\sum_{j=n}^{m} \frac{1}{j!} \|k_x\|_{\mathcal{H}_k}^{2j}\right)^{1/2} \qquad (2.2)
\end{aligned}$$

が成り立つ．従って，Γ は $\mathcal{F}$ から X 上の関数全体からなるベクトル空間への線形写像である．さらに，(2.2) から

$$|(\Gamma F)(x)| \leq \|F\|_{\mathcal{F}} \exp \frac{\|k_x\|_{\mathcal{H}_k}^2}{2}$$

が導かれる．よって，$F_n \in \ker \Gamma$ かつ $F_n \to F$ のとき，

$$\begin{aligned}
|(\Gamma F)(x)| &= |(\Gamma F)(x) - (\Gamma F_n)(x)| \\
&= |(\Gamma(F - F_n))(x)| \\
&\leq \|F - F_n\|_{\mathcal{F}} \exp \frac{\|k_x\|_{\mathcal{H}_k}^2}{2} \to 0
\end{aligned}$$

を得る．従って，$F \in \ker \Gamma$ となり，$\ker \Gamma$ は $\mathcal{F}$ の閉部分空間であることがわかった． □

この補題 2.4 により $\mathcal{F}$ と Γ に定理 2.1 を適用できる．

定義 2.5. Γ を用いた引き戻し構成法により定まるヒルベルト空間を $\exp \mathcal{H}_k$ と表す．

命題 2.3 と同様に次が成り立つ．

命題 2.6. $\exp \mathcal{H}_k$ は $\exp k_x$ を再生核とする再生核ヒルベルト空間である．また，$\|\exp k_x\|_{\exp \mathcal{H}_k}^2 = \exp \|k_x\|_{\mathcal{H}_k}^2$ が成り立つ．

ここで，$\Delta = \operatorname{diag} \Delta_n$ と定めると，図式

$$\bigoplus_{n=0}^{\infty} \mathcal{H}_k^{\otimes n} \xrightarrow{\Delta} \mathcal{F} \xrightarrow{\Gamma} \exp \mathcal{H}_k$$

が得られる．このように，$\exp \mathcal{H}_k$ の構成法の背後には（重み付き）フォック空間 $\bigoplus_{n=0}^{\infty} \mathcal{H}_k^{\otimes n}$ が隠れている．次の節では，普遍近似定理に対して，フォック空間のテンソル代数構造を利用した証明を与える．

3 普遍近似定理の証明

この節では，X を局所コンパクトなハウスドルフ空間とし，$k=k(x,y)$ を $X\times X$ 上の連続な実数値カーネル関数とする．また，K を X 内のコンパクト集合とし，$C(K)$ を K 上で定義された実数値連続関数からなるバナッハ環とする．ここで，$C(K)$ のノルムは $\|f\|_{\infty,K}=\sup_{x\in K}|f(x)|$ と定める．このとき，$\exp k$ も連続であり，$\exp k$ から構成される再生核ヒルベルト空間 $\exp\mathcal{H}_k$ の中の関数もすべて連続であることに注意する．

補題 3.1. 台が有限な元からなる $\mathcal{F}$ の部分空間を $\mathcal{F}_0$ により表す．このとき，$\Gamma\mathcal{F}_0$ は環である．

Proof. 任意の $\Psi\in\mathcal{H}_k^{\otimes m}$ に対し，$\psi=\Psi\circ\Delta_m$,

$$L_\Psi:\mathcal{H}_k^{\otimes n}\to\mathcal{H}_k^{\otimes m+n},\quad F\mapsto\Psi\otimes F$$

と定める．このとき，M_ψ により掛け算作用素を表せば，図式

$$\begin{array}{ccc}\mathcal{H}_k^{\otimes n} & \xrightarrow{L_\Psi} & \mathcal{H}_k^{\otimes m+n}\\ {\scriptstyle\Delta_n}\big\downarrow & & \big\downarrow{\scriptstyle\Delta_{m+n}}\\ \mathcal{H}_k^{n} & \xrightarrow[M_\psi|_{\mathcal{H}_k^n}]{} & \mathcal{H}_k^{m+n}\end{array}$$

は可換である．よって，$\Gamma\mathcal{F}_0$ は環である． □

次の定理は S [8] で述べたことを少々改良したものである．

定理 3.2 (S [8]). X を局所コンパクトなハウスドルフ空間，$k=k(x,y)$ を $X\times X$ 上の連続な実数値カーネル関数とする．この k から構成される再生核ヒルベルト空間 $\mathcal{H}_k$ が X の任意の 2 点を分離するならば，任意のコンパクト集合 $K\subset X$, 任意の $\varepsilon>0$, 任意の $f\in C(K)$ に対し，

$$\left\|f-\sum_{j=1}^{N}c_j\exp k(\cdot,a_j)\right\|_{\infty,K}<\varepsilon$$

をみたす $N\geq 1$, $c_1,\ldots,c_N\in\mathbb{R}$, $a_1,\ldots,a_N\in K$ が存在する．

Proof. まず，$\Gamma\mathcal{F}_0\subset\exp\mathcal{H}_k\subset C(X)$ に注意し，ノルム $\|\cdot\|_{\infty,K}$ に関する $\{f|_K : f\in\Gamma\mathcal{F}_0\}$ の閉包を A と表す．補題 3.1 により，A は $C(K)$ の部分バナッハ環である．さらに，ストーン–ワイエルシュトラスの定理により，$A=C(K)$ が成り立つ．よって， 任意の $\varepsilon>0$ と任意の $f\in C(K)$ に対し，$\|f-g\|_{\infty,K}<\varepsilon$ をみたす $g\in\Gamma\mathcal{F}_0$ が存在する．ここで，$\mathcal{K}$ を $\{\exp k_x : x\in K\}$ で生成される $\exp\mathcal{H}_k$ の閉

部分空間とする．このとき，$P_{\mathcal{K}}$ を $\mathcal{K}$ の上への直交射影とすれば，任意の $x \in K$ に対し，

$$(P_{\mathcal{K}}g)(x) = \langle P_{\mathcal{K}}g, \exp k_x \rangle_{\exp \mathcal{H}_k} = \langle g, \exp k_x \rangle_{\exp \mathcal{H}_k} = g(x)$$

が成り立つ．定理の主張は関数の K 上での値に関するものであるから，初めから $g \in \mathcal{K}$ と仮定してよい．このとき，$\|g - h\|_{\exp \mathcal{H}_k} < \varepsilon$ をみたす $\exp k_x$ $(x \in K)$ の線形結合 h が存在する．従って，

$$M_K = \sup_{x \in K} \exp \frac{k(x,x)}{2}$$

とおけば，M_K は有限であり，任意の $x \in K$ に対し，命題 2.6 から

$$\begin{aligned} |f(x) - h(x)| &\leq |f(x) - g(x)| + |g(x) - h(x)| \\ &\leq \|f - g\|_{\infty, K} + \|g - h\|_{\exp \mathcal{H}_k} \|\exp k_x\|_{\exp \mathcal{H}_k} \\ &< (1 + M_K)\varepsilon \end{aligned}$$

が成り立つ． □

以上の準備の下，普遍近似定理は簡単に導かれる．まず，$k(\boldsymbol{x}, \boldsymbol{y}) = 2\gamma \langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbb{R}^n}$ と定めると，k は $\mathbb{R}^{2n}$ 上の連続な実数値カーネル関数である．このとき，任意の $f \in C(K)$ に対して，定理 3.2 により，

$$\sup_{\boldsymbol{x} \in K} \left| e^{\gamma \|\boldsymbol{x}\|_{\mathbb{R}^n}^2} f(\boldsymbol{x}) - \sum_{j=1}^N d_j e^{2\gamma \langle \boldsymbol{x}, \boldsymbol{a}_j \rangle_{\mathbb{R}^n}} \right| < \varepsilon$$

をみたす $N \geq 1$, $d_1, \ldots, d_N \in \mathbb{R}$, $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N \in K$ が存在する．よって，$c_j = d_j e^{\gamma \|\boldsymbol{a}_j\|_{\mathbb{R}^n}^2}$ とおけば，任意の $\boldsymbol{x} \in K$ に対し，

$$\begin{aligned} \left| f(\boldsymbol{x}) - \sum_{j=1}^N c_j e^{-\gamma \|\boldsymbol{x} - \boldsymbol{a}_j\|_{\mathbb{R}^n}^2} \right| &= \left| f(\boldsymbol{x}) - e^{-\gamma \|\boldsymbol{x}\|_{\mathbb{R}^n}^2} \sum_{j=1}^N c_j e^{-\gamma \|\boldsymbol{a}_j\|_{\mathbb{R}^n}^2} e^{2\gamma \langle \boldsymbol{x}, \boldsymbol{a}_j \rangle_{\mathbb{R}^n}} \right| \\ &= e^{-\gamma \|\boldsymbol{x}\|_{\mathbb{R}^n}^2} \left| e^{\gamma \|\boldsymbol{x}\|_{\mathbb{R}^n}^2} f(\boldsymbol{x}) - \sum_{j=1}^N d_j e^{2\gamma \langle \boldsymbol{x}, \boldsymbol{a}_j \rangle_{\mathbb{R}^n}} \right| \\ &\leq \sup_{\boldsymbol{x} \in K} \left| e^{\gamma \|\boldsymbol{x}\|_{\mathbb{R}^n}^2} f(\boldsymbol{x}) - \sum_{j=1}^N d_j e^{2\gamma \langle \boldsymbol{x}, \boldsymbol{a}_j \rangle_{\mathbb{R}^n}} \right| \\ &< \varepsilon \end{aligned}$$

が成り立つ．

参考文献

[1] C. M. ビショップ（元田 浩，栗田 多喜夫，樋口 知之，松本 裕治，村田 昇（監訳）），パターン認識と機械学習 上・下，丸善出版，2012.

[2] J. C. GUELLA, *On Gaussian kernels on Hilbert spaces and kernels on hyperbolic spaces.* J. Approx. Theory **279** (2022), Paper No. 105765, 36 pp.

[3] S. KUWAHARA AND M. SETO, *Exponentials of de Branges-Rovnyak kernels.* Canad. Math. Bull. **65** (2022), no. 2, 447–455.

[4] C. A. MICCHELLI, Y. XU AND H. ZHANG, *Universal kernels.* J. Mach. Learn. Res. **7** (2006), 2651–2667.

[5] 持橋 大地，大羽 成征，ガウス過程と機械学習，講談社，2019.

[6] N. K. NIKOLSKI, *Operators, functions, and systems: an easy reading. Vol. 1. Hardy, Hankel, and Toeplitz.* Translated from the French by Andreas Hartmann. Mathematical Surveys and Monographs, **92**. American Mathematical Society, Providence, RI, 2002.

[7] V. I. PAULSEN AND M. RAGHUPATHI, *An introduction to the theory of reproducing kernel Hilbert spaces.* Cambridge Studies in Advanced Mathematics, **152**. Cambridge University Press, Cambridge, 2016.

[8] M. SETO, *An unbounded approach to the theory of kernel functions*, preprint.

[9] 瀬戸 道生，伊吹 竜也，畑中 健志，機械学習のための関数解析入門，内田老鶴圃，2021.

[10] I. STEINWART, *On the influence of the kernel on the consistency of support vector machines.* J. Mach. Learn. Res. **2** (2002), no. 1, 67–93.

Michio Seto
National Defense Academy
Yokosuka 239-8686 JAPAN
E-mail: mseto@nda.ac.jp