Chapter 19
Hypothesis space II

In this chapter we are going to construct the Reproducing kernel Hilbert space (RKHS) using the eigen-pairs of an integral operator defined by the Mercer kernel under consideration. We will show that this operator is a compact and self-adjoint and hence admits a spectral decomposition by the Hilbert-Schmidt theorem 1.1. This approach is more appealing as it follows the same direction of constructing Gaussian measure in Chapter **??**, and hence making the presentation of the book more coherent.

19.1 Kernel-based integral operators

Let us define the following integral operator

$$(L_{\mathsf{K}} f)(\mathbf{x}) := \int_X \mathsf{K}(\mathbf{x}, \mathbf{t}) f(\mathbf{t})\, d\pi(\mathbf{t}),$$

where we assume π is a probability measure on X though all the results, up to a constant, also hold for a general Borel measure. We observe

$$|(L_{\mathsf{K}} f)(\mathbf{x})| \leq \|\mathsf{K}_{\mathbf{x}}\|_{\mathbb{L}^2(X,\pi(\mathbf{x}))} \|f\|_{\mathbb{L}^2(X,\pi(\mathbf{x}))} \leq C_{\mathsf{K}} \|f\|_{\mathbb{L}^2(X,\pi(\mathbf{x}))}, \tag{19.1}$$

which implies L_{K} as a map from $\mathbb{L}^2(X, \pi(\mathbf{x}))$ to $\mathbb{C}(X)$ is (Lipschitz) continuous. Since the inclusion $\mathbb{C}(X) \hookrightarrow \mathbb{L}^2(X, \pi(\mathbf{x}))$ is continuous, we see that L_{K} as a map from $\mathbb{L}^2(X, \pi(\mathbf{x}))$ into $\mathbb{L}^2(X, \pi(\mathbf{x}))$ is continuous and its operator norm is bounded as $\|L_{\mathsf{K}}\| \leq C_{\mathsf{K}}$.

Proposition 19.1. $L_{\mathsf{K}} : \mathbb{L}^2(X, \pi(\mathbf{x})) \to \mathbb{C}(X)$ *is a compact operator. If, in addition,* K *is a Mercer kernel, then* L_{K} *is self-adjoint positive semidefinite compact operator.*

Proof. Let B be a bounded set in $\mathbb{L}^2(X, \pi(\mathbf{x}))$ such that $\|f\|_{\mathbb{L}^2(X,\pi(\mathbf{x}))} \leq M, \forall f \in B$. From (19.1) we see that $L_{\mathsf{K}}(B)$ is uniformly bounded. A similar argument as in (19.1) shows that

$$\left|(L_{\mathsf{K}}f)(\mathbf{x})-(L_{\mathsf{K}}f)\left(\mathbf{x}'\right)\right|\leq\sup_{\mathbf{t}\in X}|\mathsf{K}_{\mathbf{x}}(\mathbf{t})-\mathsf{K}_{\mathbf{x}'}(\mathbf{t})|\,\|f\|_{\mathbb{L}^2(X,\pi(\mathbf{x}))}\leq 2C_{\mathsf{K}}M,$$

which implies that $L_{\mathsf{K}}(B)$ is equicontinuous. By the Arzelá-Ascoli theorem 18.3, the closure of $L_{\mathsf{K}}(B)$ is compact in $\mathbb{C}(X)$. In other words, $L_{\mathsf{K}}(B)$ is relatively compact and by definition 19.2, L_{K} is a compact operator.

The self-adjointness is clear by the Fubini theorem. The positive semidefiniteness of L_{K} is a direct consquence of the positive semidefiniteness of K. Indeed, since X is a compact subset of $\mathbb{R}^k$, without loss of generality, we can subdivide X into n subsets with equal volumes and with "centroids" $\mathbf{x}_1,\ldots,\mathbf{x}_n$. We then have

$$\begin{aligned}(f,L_{\mathsf{K}}f)_{\mathbb{L}^2(X,\pi(\mathbf{x}))}&=\int_{X\times X}\mathsf{K}(\mathbf{x},\mathbf{t})\,f(\mathbf{t})\,f(\mathbf{x})\,d\pi(\mathbf{t})\,d\pi(\mathbf{x})\\&=\lim_{n\to\infty}\frac{\mathscr{V}(X)^2}{n^2}\sum_{i,j=1}^{n}\mathsf{K}(\mathbf{x}_i,\mathbf{t}_j)\,f(\mathbf{t}_j)\,f(\mathbf{x}_i)\geq 0.\end{aligned}$$

□

By the Hilbert-Schmidt Theorem 1.1, L_{K} admits a spectral decomposition with eigenpairs $(\lambda_i,\varphi_i)_{i=1}^{\infty}$, i.e.,

$$L_{\mathsf{K}}f=\sum_{i=1}^{\infty}a_i\lambda_i\varphi_i,$$

for any $f=\sum_{i=1}^{\infty}a_i\varphi_i\in\mathbb{L}^2(X,\pi(\mathbf{x}))$. Moreover the Mercer Theorem 19.3 holds.

Exercise 19.1. Show that

$$\sum_{i=1}^{\infty}\lambda_i=\int_X\mathsf{K}(\mathbf{x},\mathbf{x})\,d\pi(\mathbf{x})\leq\mathscr{V}(X)\,C_{\mathsf{K}}.$$

Deduce that $\lambda_k\leq\dfrac{\mathscr{V}(X)\,C_{\mathsf{K}}}{k}$. •

Without loss of generality, assume that $\lambda_i>0,\forall i$. Let us now define the following space

$$\widehat{\mathscr{H}_{\mathsf{K}}}:=\left\{f\in\mathbb{L}^2(X,\pi(\mathbf{x})):f=\sum_{i=1}^{\infty}a_i\varphi_i\text{ with }\left\{\frac{a_i}{\sqrt{\lambda_i}}\right\}_{i\in\mathbb{N}}\in\ell^2\right\},$$

and equip $\widehat{\mathscr{H}_{\mathsf{K}}}$ with the following inner product

$$(f,g)_{\mathsf{K}}:=\sum_{i=1}^{\infty}\frac{a_ib_i}{\lambda_i}\tag{19.2}$$

for any $f=\sum_{i=1}^{\infty}a_i\varphi_i$ and $g=\sum_{i=1}^{\infty}b_i\varphi_i$ in $\widehat{\mathscr{H}_{\mathsf{K}}}$. Exercise 19.2 shows that $\widehat{\mathscr{H}_{\mathsf{K}}}$ is a Hilbert space.

Exercise 19.2. Show that $\widehat{\mathscr{H}_{\mathsf{K}}}$ with aforementioned inner product is a Hilbert space.

•

Since $L_{\mathsf{K}}^{\frac{1}{2}} : \mathbb{L}^2(X, \pi(\mathbf{x})) \ni f = \sum_{i=1}^{\infty} a_i \varphi_i \mapsto g = L_{\mathsf{K}}^{\frac{1}{2}} f := \sum_{i=1}^{\infty} a_i \sqrt{\lambda_i} \varphi_i \in \widehat{\mathscr{H}_{\mathsf{K}}}$ is an isometry, the space $\widehat{\mathscr{H}_{\mathsf{K}}}$ allows us to define a square root of L_{K}.

Compare $L_{\mathsf{K}}^{\frac{1}{2}}$ and the Cameron-Martin map of a Gaussian measure.

Exercise 19.3. Show that the following hold true:

i) $L_{\mathsf{K}}^{\frac{1}{2}} : \mathbb{L}^2(X, \pi(\mathbf{x})) \ni f = \sum_{i=1}^{\infty} a_i \varphi_i \mapsto g = L_{\mathsf{K}}^{\frac{1}{2}} f := \sum_{i=1}^{\infty} a_i \sqrt{\lambda_i} \varphi_i \in \widehat{\mathscr{H}_{\mathsf{K}}}$ is an isometry.
ii) $g \in \mathbb{L}^2(X, \pi(\mathbf{x}))$.
iii) $L_{\mathsf{K}} = L_{\mathsf{K}}^{\frac{1}{2}} \cdot L_{\mathsf{K}}^{\frac{1}{2}}$.

•

Now comes the main result of this chapter.

Theorem 19.1. *$\widehat{\mathscr{H}_{\mathsf{K}}}$ and $\mathscr{H}_{\mathsf{K}}$ are identical.*

Proof. We just need to show that $\widehat{\mathscr{H}_{\mathsf{K}}}$ has three properties in Theorem 18.1.

- By definition of $\mathsf{K}_{\mathbf{x}}$, Mercer theorem, and the induced norm from (19.2) we have

$$\begin{aligned}\|\mathsf{K}_{\mathbf{x}}(\mathbf{t})\|_{\mathsf{K}} = \|\mathsf{K}(\mathbf{x},\mathbf{t})\|_{\mathsf{K}} = \left\|\sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x}) \varphi_i(\mathbf{t})\right\|_{\mathsf{K}} &= \sqrt{\sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x})^2} \\ &= \sqrt{\mathsf{K}(\mathbf{x},\mathbf{x})} \leq C_{\mathsf{K}} < \infty,\end{aligned}$$

  which shows that $\mathsf{K}_{\mathbf{x}} \in \widehat{\mathscr{H}_{\mathsf{K}}}$ for any $\mathbf{x} \in X$.
- To see that $\widehat{\mathscr{H}_{\mathsf{K}}}$ has the reproducing property we take $f(\mathbf{t}) = \sum_{i=1}^{\infty} a_i \varphi_i(\mathbf{t}) \in \widehat{\mathscr{H}_{\mathsf{K}}}$ and show that $(f, \mathsf{K}_{\mathbf{x}})_{\mathsf{K}} = f(\mathbf{x})$ but this is straightforward by Mercer theorem and the definition of the inner product in $\widehat{\mathscr{H}_{\mathsf{K}}}$. Indeed,

$$(f, \mathsf{K}_{\mathbf{x}})_{\mathsf{K}} = \sum_{i=1}^{\infty} \frac{a_i \lambda_i \varphi_i(\mathbf{x})}{\lambda_i} = f(\mathbf{x}).$$

- To show that $\mathring{\mathscr{H}_{\mathsf{K}}}$ is dense in $\mathscr{H}_{\mathsf{K}}$, it is sufficient to prove that, for any $f \in \widehat{\mathscr{H}_{\mathsf{K}}}$, $(f, \mathsf{K}_{\mathbf{x}}) = 0$ for any $\mathbf{x} \in X$ implies $f = 0$. But this is clear by the reproducing property.

□

We conclude this chapter with the definition and properties of a *feature* map.

Theorem 19.2. *Show the the feature map defined by*

$$\Phi : X \ni \mathbf{x} \mapsto \Phi(\mathbf{x}) := \left\{\sqrt{\lambda_k} \varphi_k(\mathbf{x})\right\}_{k \in \mathbb{N}} \in \ell^2$$

is well-defined, and continuous, and satisfies

$$\mathsf{K}(\mathbf{x},\mathbf{t}) = (\Phi(\mathbf{x}), \Phi(\mathbf{t}))_{\ell^2}.$$

Proof. It is obvious by Mercer theorem:

$$(\Phi(\mathbf{x}), \Phi(\mathbf{t}))_{\ell^2} = \sum_{k=1}^{\infty} \lambda_k \varphi_k(\mathbf{x}) \varphi_k(\mathbf{t}) = \mathsf{K}(\mathbf{x},\mathbf{t}),$$

and

$$\|\Phi(\mathbf{x})\|_{\ell}^2 = \sqrt{\sum_{k=1}^{\infty} \lambda_k \varphi_k(\mathbf{x})^2} = \sqrt{\mathsf{K}(\mathbf{x},\mathbf{x})} \le C_{\mathsf{K}} < \infty.$$

□

19.2 Appendix

Definition 19.1 (Relatively compact). A subset B of a metric space X is relatively compact if its closure is compact in X.

Definition 19.2 (Compact operator). $L : X \to y$ is a compact operator if for any bounded set $B \subset X$, $L(B)$ is relatively compact in y.

Theorem 19.3 (Mercer theorem). *Suppose X is a compact subset of $\mathbb{R}^k$ and K is a Mercer kernel. Let $(\lambda_i, \varphi_i)_{i=1}^{\infty}$ be the eigenpairs of L_{K}. For all $\mathbf{x}, \mathbf{t} \in X$, there holds*

$$\mathsf{K}(\mathbf{x},\mathbf{t}) = \sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x}) \varphi_i(\mathbf{t}),$$

where the convergence is absolute and uniform on pointwise $X \times X$.