Chapter 18
Hypothesis space I

Recall from Chapter 17 we have assumed that the hypothesis space $\mathscr{H}$ is compact subset of $\mathbb{C}(X)$. In this chapter we show that this assumption is valid when the hypothesis space is taken as a bounded subset of the reproducing kernel Hilbert space of Mercer kernel.

18.1 Reproducing kernel Hilbert spaces (RKHS)

In this book we are interested in hypothesis space associated with RKHS. Assume X is a metric space and let $\mathsf{K} : X \times X \to \mathbb{R}$ be *symmetric positive semidefinite* in the sense $\mathsf{K}(\mathbf{x}, \mathbf{x}') = \mathsf{K}(\mathbf{x}', \mathbf{x})$ and any $n \times n$ *Gramian* matrix whose ij-element is $\mathsf{K}(\mathbf{x}_i, \mathbf{x}_j)$ is semipositive definite for all n and $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$. We say K is a *Mercer* kernel if it is continuous, symmetric, and positive semidefinite. Clearly the positive semidefiniteness implies $\mathsf{K}(\mathbf{x}, \mathbf{x}) \geq 0$ for any $\mathbf{x} \in X$, and this allows us to define

$$C_{\mathsf{K}} := \sup_{\mathbf{x} \in X} \sqrt{\mathsf{K}(\mathbf{x}, \mathbf{x})}.$$

Exercise 18.1. Show that

$$C_{\mathsf{K}} = \sup_{\mathbf{x}, \mathbf{x}' \in X} \sqrt{|\mathsf{K}(\mathbf{x}, \mathbf{x}')|}$$

•

For any $\mathbf{x} \in X$, by $\mathsf{K}_{\mathbf{x}}$ we denote the function $\mathsf{K}_{\mathbf{x}} : X \ni \mathbf{x}' \mapsto \mathsf{K}_{\mathbf{x}}(\mathbf{x}') := \mathsf{K}(\mathbf{x}, \mathbf{x}') \in \mathbb{R}$. We form the space $\mathring{\mathscr{H}}_{\mathsf{K}}$ as

$$\mathring{\mathscr{H}}_{\mathsf{K}} := \operatorname{span}\{\mathsf{K}_{\mathbf{x}} : \mathbf{x} \in X\},$$

and equip it with the following inner product: for $f = \sum_{i=1}^{I} f_i \mathsf{K}_{\mathbf{x}^i} \in \mathring{\mathscr{H}}_{\mathsf{K}}$ and $g = \sum_{j=1}^{J} g_j \mathsf{K}_{\mathbf{t}^j} \in \mathring{\mathscr{H}}_{\mathsf{K}}$,

$$(f,g)_{\mathscr{H}_{\mathsf{K}}} = \sum_{i,j}^{I,J} f_i g_j \mathsf{K}_{\mathbf{x}^i} \mathsf{K}_{\mathbf{t}^j}. \tag{18.1}$$

Theorem 18.1 (Reproducing Kernel Hilbert Space). *Let $\mathscr{H}_{\mathsf{K}}$ be a Hilbert space with with the following properties:*

1. $\mathsf{K}_{\mathbf{x}} \in \mathscr{H}_{\mathsf{K}}$ *for any* $\mathbf{x} \in X$.
2. $\mathring{\mathscr{H}}_{\mathsf{K}}$ *is dense in* $\mathscr{H}_{\mathsf{K}}$.
3. **Reproducing property:** *for any* $f \in \mathscr{H}_{\mathsf{K}}$ *and* $\mathbf{x} \in X$, *we have* $f(\mathbf{x}) = (\mathsf{K}_{\mathbf{x}}, f)_{\mathscr{H}_{\mathsf{K}}}$.

Then $\mathscr{H}_{\mathsf{K}}$ is unique. Moreover, $\mathscr{H}_{\mathsf{K}} \subset \mathbb{C}(X)$ and the inclusion $\mathrm{i}_{\mathsf{K}} : \mathscr{H}_{\mathsf{K}} \hookrightarrow \mathbb{C}(X)$ is bounded.

Proof. Note sure we need this result, so let's not type the proof for now. □

18.2 Hypothesis space associated with RKHS

Proposition 18.1. *Suppose K be a Mercer kernel on a compact metric space X, and $\mathscr{H}_{\mathsf{K}}$ is the associated RKHS. For any $R > 0$, the ball $B(R) := \left\{ f \in \mathscr{H}_{\mathsf{K}} : \|f\|_{\mathscr{H}_{\mathsf{K}}} \le R \right\}$ is a closed subset of $\mathbb{C}(X)$.*

Proof. From Theorem 18.1 it is sufficient to show that $B(R)$ is closed in $\mathbb{C}(X)$. To that end, suppose $\{f_n\}_{n\in\mathbb{N}} \in B(R)$ converges to $f^* \in \mathbb{C}(X)$, i.e.,

$$\lim_n f_n(\mathbf{x}) = f^*(\mathbf{x}),$$

and we need to show that $f^* \in B(R)$. Since, by Lemma 18.1, $B(R)$ is weakly compact, there exists a subsequence $\left\{f_{n_k}\right\}_{k\in\mathbb{N}}$ converging to $\hat{f} \in B(R)$, i.e.,

$$\lim_k \left(f_{n_k}, g\right)_{\mathscr{H}_{\mathsf{K}}} = \left(\hat{f}, g\right)_{\mathscr{H}_{\mathsf{K}}}, \quad \forall g \in \mathscr{H}_{\mathsf{K}}.$$

Now taking $g = \mathsf{K}_{\mathbf{x}}$ and using the reproducing property we have

$$f^*(\mathbf{x}) = \lim_k f_{n_k}(\mathbf{x}) = \lim_k \left(f_{n_k}, \mathsf{K}_{\mathbf{x}}\right)_{\mathscr{H}_{\mathsf{K}}} = \left(\hat{f}, \mathsf{K}_{\mathbf{x}}\right)_{\mathscr{H}_{\mathsf{K}}} = \hat{f}(\mathbf{x}), \quad \forall \mathbf{x} \in X.$$

Since both f^* and $\hat{f}$ are continuous, they must be identical and this ends the proof. □

Theorem 18.2. *Suppose K be a Mercer kernel on a compact metric space X, and $\mathscr{H}_{\mathsf{K}}$ is the associated RKHS. The inclusion $\mathrm{i}_{\mathsf{K}} : \mathscr{H}_{\mathsf{K}} \hookrightarrow \mathbb{C}(X)$ is compact. In other words, the set $\mathrm{i}_{\mathsf{K}}(B(R))$ is compact for any $R > 0$.*

Proof. Theorem 18.1 and Proposition 18.1 show that $\mathrm{i}_{\mathsf{K}}(B(R))$ is closed and bounded. By the Arzelá-Ascoli theorem, what remains is to show the equicontinuity. We have

$$\begin{aligned}|f(\mathbf{x})-f(\mathbf{t})| &= \left|(f,\mathsf{K}_{\mathbf{x}}-\mathsf{K}_{\mathbf{t}})_{\mathscr{H}_{\mathsf{K}}}\right| \le \|f\|_{\mathscr{H}_{\mathsf{K}}}\|\mathsf{K}_{\mathbf{x}}-\mathsf{K}_{\mathbf{t}}\|_{\mathscr{H}_{\mathsf{K}}} \\ &= R\sqrt{(\mathsf{K}_{\mathbf{x}}-\mathsf{K}_{\mathbf{t}},\mathsf{K}_{\mathbf{x}}-\mathsf{K}_{\mathbf{t}})_{\mathscr{H}_{\mathsf{K}}}} = R\sqrt{(\mathsf{K}_{\mathbf{x}}(\mathbf{x})-\mathsf{K}_{\mathbf{x}}(\mathbf{t})+\mathsf{K}_{\mathbf{t}}(\mathbf{t})-\mathsf{K}_{\mathbf{t}}(\mathbf{x}))_{\mathscr{H}_{\mathsf{K}}}}\end{aligned}$$

Now since K is continuous on the compact set $X\times X$, it is uniformly continuous on $X\times X$, i.e., for all $\mathbf{x},\mathbf{t},\mathbf{t}'\in X$ such that $\left\|\mathbf{t}-\mathbf{t}'\right\|_X \le \delta$ (δ does not depend on $\mathbf{x},\mathbf{t},\mathbf{t}'$ implies

$$\left|\mathsf{K}_{\mathbf{x}}(\mathbf{t})-\mathsf{K}_{\mathbf{x}}\left(\mathbf{t}'\right)\right| \le \varepsilon.$$

We thus have

$$|f(\mathbf{x})-f(\mathbf{t})| \le R\sqrt{2\varepsilon}, \quad \forall \mathbf{t},\mathbf{x} : \|\mathbf{x}-\mathbf{t}\| \le \delta, \quad \forall f\in \mathrm{i}_{\mathsf{K}}(B(R)),$$

and this concludes the proof. □

In this book we generally take the hypothesis space $\mathscr{H}$ as a compact subset of $\mathbb{C}(X)$. As shown in Theorem 18.2, this can be justified by choosing $\mathscr{H}$ as a closed ball in the RKHS associated with the kernel under consideration.

18.3 Appendix

Lemma 18.1 (Weak compactness of closed balls in Hilbert space). *If B be a closed ball in a Hilbert space $\mathscr{H}$, it is weakly compact. In other words, every sequence $\{f_n\}_{n\in\mathbb{N}} \subset B$ has a weakly convergence subsequence $\left\{f_{n_k}\right\}_{k\in\mathbb{N}}$. That is, there exists some $f^*\in B$ such that*

$$\lim_{k\to\infty}\left(f_{n_k},g\right)_{\mathscr{H}} = (f^*,g)_{\mathscr{H}}, \quad \forall g\in\mathscr{H}.$$

Definition 18.1 (Equicontinuity). A subset K of $\mathbb{C}(X)$ is equicontinuous at $\mathbf{x}\in X$ if for any $\varepsilon>0$ there exists a neighborhood B of x such that $\forall \mathbf{t}\in X$ and $\forall f\in K$ we have $\|f(\mathbf{x})-f(\mathbf{t})\|_\infty < \varepsilon$. K is equicontinuous if it is equicontinuous at every $\mathbf{x}\in X$.

Theorem 18.3 (Arzelá-Ascoli theorem). *Let X be compact. $K\subset\mathbb{C}(X)$ is compact if and only if K is closed, bounded, and equicontinuous.*