RKHS Notes

Kai Zhang

May 2018

1 Examples of Mercer's Theorem

1.1 R^d

Let $(R^d, < \cdot, \cdot >_E)$ be the Euclidean space. Let $\mathcal{D} = \{1, 2, \ldots, d\}$. Given a symmetric positive definite matrix $K \in R^{d\times d}$, find the RKHS $(\mathcal{H}, < \cdot, \cdot >_{\mathcal{H}}, K)$ such that

- $\mathcal{H} \subset R^d$
- $K(\cdot, j) \in \mathcal{H}, \forall j \in \mathcal{D}$
- $< \theta, K(\cdot, j) >_{\mathcal{H}} = \theta_j, \forall \theta \in \mathcal{H}, j \in \mathcal{D}$

, where $K(\cdot, j)$ denotes the j^{th} column vector of the matrix K.

Note that **Moore-Aronszajn Theorem** guarantees the existence and uniqueness of such RKHS. It turns out that the RKHS associated with K is $\mathcal{H} = R^d, < \theta_1, \theta_2 >_{\mathcal{H}} = \sum_{n=1}^{d} \frac{<\theta_1, e_n>_E \cdot <\theta_2, e_n>_E}{\lambda_n}$, in which $\{e_n\}_{n=1,2,\ldots,n}$ is an orthonormal set of eigen-vectors of K and λ_n is the corresponding eigen-value. Check condition 3,

$$K = \sum_{n=1}^{d} \lambda_n e_n e_n^T \qquad \text{(Spectral Theorem)}$$

$$K(i, j) = \sum_{n=1}^{d} \lambda_n e_n(i) e_n^T(j)$$

$$K(\cdot, j) = \sum_{n=1}^{d} \lambda_n e_n(\cdot) e_n^T(j) = \sum_{n=1}^{d} \lambda_n e_n(j) \cdot e_n$$

$$\theta = \sum_{n=1}^{d} < \theta, e_n >_E \cdot e_n$$

$$< \theta, K(\cdot, j) >_{\mathcal{H}} = \sum_{n=1}^{d} < \theta, e_n >_E \cdot e_n(j) = \theta_j$$

Let's compare the RKHS $(\mathcal{H}, <\cdot,\cdot>_{\mathcal{H}}, K)$ with the Eculidean space $(R^d, <\cdot,\cdot>_E)$ to get some intuition. First, $\mathcal{H} = R^d$ means they contain exactly the same set of vectors; the only difference lies in the inner product. Second, $<\theta,\theta>_{\mathcal{H}} \neq <\theta,\theta>_E$. This indicates that the underlying normed space is changed, which will affect the Tikhonov regularization term $\|\theta\|$ in optimization.

A very simple example is $K = I$.

1.2 $L_2([a,b])$

Let $(L_2([a,b]), <\cdot,\cdot>_{L_2})$ be the Hilbert space consisting of square-integrable functions $f : [a,b] \to R, \int_a^b f^2(x)dx < \infty$. Given a **continuous** symmetric positive definite function $k(x,x')$, find the RKHS $(\mathcal{H}, <\cdot,\cdot>_{\mathcal{H}}, k)$ such that

- $\mathcal{H} \subset L_2([a,b])$
- $k(\cdot,x) \in \mathcal{H}, \forall x \in [a,b]$
- $<f, k(\cdot,x)>_{\mathcal{H}} = f(x), \forall f \in \mathcal{H}, x \in [a,b]$

Since $k(\cdot,x) \in L_2([a,b])$, the integral operator associated with k is $T : f(\cdot) \to g(x) = <f(\cdot), k(\cdot,x)>_{L_2}$. Clearly, $g(x) \in L_2([a,b])$. T is a self-adjoint ($<Tf,g>_{L_2} = <f,Tg>_{L_2}$), positive ($<f,Tf>_{L_2} \geq 0$), linear, continuous, bounded operator, thus by the **Spectral Theorem**, the eigen-values $\{\lambda_n\}_{n=1,2,...}$ forms a decreasing sequence with $\lambda_n > 0, \lim_{n\to\infty} \lambda_n = 0$ and the orthonormal set of **continuous** eigen-functions $\{e_n\}_{n=1,2,...}$ forms a complete basis of $L_2([a,b])$. This means,

- $\forall f \in L_2([a,b]), f = \sum_{n=1}^{\infty} <f, e_n>_{L_2} \cdot e_n$; any functions that only differ at a zero-measure set have the same series expansion.
- $\forall f \in L_2([a,b])$, the function sequence $g_m = \sum_{n=1}^{m} <f, e_n>_{L_2} \cdot e_n, m = 1, 2, \ldots$ converges in norm to f, i.e., $\lim_{m\to\infty} \|g_m - f\|_{L_2} = 0$
- if $f \in C([a,b]) \subset L_2([a,b])$, then $\{g_m\}_{m=1,2,...}$ also converges point-wise to f, i.e., $\forall x \in [a,b], \lim_{m\to\infty} |g_m(x) - f(x)| = 0$

.

$$k(x,x') = \sum_{n=1}^{\infty} \lambda_n e_n(x) e_n(x'), \quad \forall x, x' \in [a,b] \qquad \text{(Mercer's Theorem)}$$

, where the right-hand side series converges absolutely, and the function represented by this series converges uniformly to the left-hand side function, i.e.,

$$\lim_{m\to\infty} \sup_{x,x'\in[a,b]} \left| k(x,x') - \sum_{n=1}^{m} \lambda_n e_n(x) e_n(x') \right| = 0$$

. The RKHS $\mathcal{H}$ associated with $k(x,x')$ is

$$\mathcal{H} = \{f \in C([a,b]) \big| f = \sum_{n=1}^{\infty} < f, e_n >_{L_2} \cdot e_n, \sum_{n=1}^{\infty} (\frac{< f, e_n >_{L_2}}{\sqrt{\lambda_n}})^2 < \infty\}$$

, with the inner product $< f, g >_{\mathcal{H}} = \sum_{n=1}^{\infty} \frac{<f,e_n>_{L_2}}{\sqrt{\lambda_n}} \cdot \frac{<g,e_n>_{L_2}}{\sqrt{\lambda_n}}$. Check condition 2,

$$k(\cdot, x) = \sum_{n=1}^{\infty} \lambda_n e_n(\cdot) e_n(x) = \sum_{n=1}^{\infty} \lambda_n e_n(x) \cdot e_n$$
$$< k(\cdot, x), e_n >_{L_2} = \lambda_n e_n(x)$$
$$\sum_{n=1}^{\infty} \frac{< k(\cdot, x), e_n >_{L_2}^2}{\lambda_n} = \sum_{n=1}^{\infty} \lambda_n e_n^2(x) = k(x,x) < \infty$$
$$k(\cdot, x) = \sum_{n=1}^{\infty} < k(\cdot, x), e_n >_{L_2} \cdot e_n \in \mathcal{H}$$

Check condition 3,

$$f(x) = \sum_{n=1}^{\infty} < f, e_n >_{L_2} \cdot e_n(x) \qquad \text{(point-wise convergence)}$$
$$< f, k(\cdot, x) >_{\mathcal{H}} = \sum_{n=1}^{\infty} < f, e_n >_{L_2} \cdot e_n(x) = f(x)$$

Note that $\mathcal{H} \subset C([a,b])$.

A very simple example $k(x,x') = \mathbb{1}\{x = x'\}$