Available online at www.sciencedirect.com

ScienceDirect

Journal of Complexity 24 (2008) 241–258

Journal of COMPLEXITY

www.elsevier.com/locate/jco

The covering number for some Mercer kernel Hilbert spaces☆

Sheng Baohuai*, Wang Jianli, Li Ping

Department of Mathematics, Shaoxing College of Arts and Sciences, Shaoxing, Zhejiang 312000, PR China

Received 13 October 2006; accepted 4 April 2007

Available online 4 May 2007

Abstract

In the present paper, we investigate the estimates for the covering number of a ball in a Mercer kernel Hilbert space on [0, 1]. Let $P_l(x)$ be the Legendre orthogonal polynomial of order l, $a_l > 0$ be real numbers satisfying $\sum_{l=0}^{+\infty} l a_l < +\infty$. Then, for the Mercer kernel function

$$K(x,t) = \sum_{l=0}^{+\infty} a_l P_l(x) P_l(t), \quad x, t \in [0, 1],$$

we provide the upper estimates of the covering number for the Mercer kernel Hilbert space reproducing from $K(x,t)$. For some particular a_l we give the lower estimates. Meanwhile, a kind of l^2-norm estimate for the inverse Mercer matrix associated with the Mercer kernel $K(x,t)$ is given.

Keywords: Mercer kernel Hilbert spaces; Covering number; Legendre polynomials

1. Introduction

Let X be a compact set of the Euclidean space $\mathcal{R}^n$, $L^2(X)$ be the space of real square integrable functions with respect to a Borel measure v on X.

A function $K : X \times X \to \mathcal{R}$ which is continuous, symmetric and positive definite, i.e., for any finite set $\{x_1, \ldots, x_m\} \subset X$, the matrix $(K(x_i, x_j))_{i,j=1}^m$ is positive definite is called a Mercer kernel. The reproducing kernel Hilbert space $\mathcal{H}_K$ associated with the kernel K is defined to be

☆ This work is supported by the national NSF of China.

* Corresponding author.

E-mail address: shengbaohuai@hotmail.com (S. Baohuai).

0885-064X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.

doi:10.1016/j.jco.2007.04.001

the closure of the linear span of the set of functions $\{K(t, x) : t \in X\}$ with the inner product satisfying

$$\langle K(\cdot, x), f(\cdot)\rangle_{\mathcal{H}_K} = f(x), \quad x \in X, \; f \in \mathcal{H}_K.$$

Define a Hilbert–Schmidt integral operator by means of this kernel as

$$L_K(f, x) = \int_X K(x, t) f(t)\, dv(t), \quad x \in X, \; f \in L^2_v(X).$$

Then, $L_K(f, x)$ is a positive, compact operator and its range lies in $C(X)$.

Let $(\lambda_j)_{j=1}^{+\infty}$ denote the nonincreasing sequence of eigenvalues of L_K and $(\phi_j)_{j=1}^{+\infty}$ be the corresponding eigenfunctions. Then,

$$K(x, t) = \sum_{j=1}^{+\infty} \lambda_j \phi_j(x)\phi_j(t), \quad x, t \in X,$$

where the series converges uniformly and absolutely.

$\mathcal{H}_K$ can be imbedded into $C(X)$, and we denote the inclusion as $I_K : \mathcal{H}_K \to C(X)$. For this facts, see [2].

Let $R > 0$ and B_R be the ball of $\mathcal{H}_K$ with radius R:

$$B_R := \{f \in \mathcal{H}_K : \|f\|_K \leqslant R\}.$$

Then $I_K(B_R) \subset C(X)$. Denote its closure in $C(X)$ as $\overline{I_K(B_R)}$ which is a compact subset in $C(X)$.

Let $\mathcal{N}$ be the set of natural numbers, S be a compact set in a metric space and $\eta > 0$. The covering number $\mathcal{N}(S, \eta)$ of S is defined to be the minimal integer m such that there exist m disks with radius η covering S.

The covering number is often used to bounding the error between the empirical function and the target function (see, [2,15]). Thus, the estimates for the covering number of $\overline{I_K(B_R)}$ are needed in kernel machine learning. Both the upper bounds and the lower bounds for $\mathcal{N}(\overline{I_K(B_R)}, \eta)$ have been investigated in the literature [3,10,14,16,17].

A theorem of Zhou (see, [16]) has concluded the estimates for the covering numbers to the estimates of the l^2-norm of the inverse of the Mercer matrix A_N. It is known $\|A_N^{-1}\|_{l^2}$ equals the inverse of the smallest eigenvalue of the matrix A_N. This problem has a closely connection with the radial interpolation approximation (see, e.g., [13]) but are difficult to deal with. Some investigations of this field can be found from [5–7,11,13]. In the present paper, we shall provide, with the help of the Legendre orthogonal polynomials, a kind of estimate for the smallest eigenvalue of the matrix A_N for a kind of Mercer kernel function on $[0, 1] \times [0, 1]$ and thus give some lower and upper estimates for the bounds of the covering number.

Let $X = [0, 1]$, $P_k(x)$ be the Legendre orthogonal polynomial of order k on $[0, 1]$. Then, $P_k(x)$ satisfy $|P_k(x)| \leqslant 1$, $x \in [0, 1]$, and

$$\int_0^1 P_n(x)P_m(x)\, dx = \frac{\delta_{n,m}}{2n + 1},$$

where $\delta_{n,m}$ is the δ function whose value is 1 if $n = m$ and whose value is 0 if $n \neq m$.

Let $\phi \in L^1[0,1]$ and its Fourier–Legendre coefficients $a_k(\phi) = \int_0^1 \phi(u) P_k(u)\, du$ satisfy $a_k(\phi) > 0$ and $\sum_{l=0}^{+\infty} l a_l(\phi) < +\infty$. Then, the Mercer kernel $K(x, y)$ defined on $[0,1] \times [0,1]$ with $a_k(\phi)$ being its eigenvalues and $P_k(x)$ being its eigenfunctions has the representation

$$K(x, y) := \sum_{l=0}^{+\infty} (2l+1) a_l(\phi) P_l(x) P_l(y), \quad x, y \in [0,1]. \tag{1}$$

In fact, $K(x, y)$ is a kind of translation of ϕ. By [1,9] we know

$$K(x, y) = \frac{1}{\pi} \int_0^1 \phi\left[(2x-1)(2y-1) + 4(2u-1)\sqrt{x(1-x)y(1-y)}\right] \times (u(1-u))^{-1/2}\, du, \quad x, y \in [0,1]. \tag{2}$$

The generating function of the Legendre polynomials yields

$$\frac{1}{\sqrt{(1+q)^2 - 4xq}} = \sum_{l=0}^{+\infty} q^l P_l(x), \quad x \in [0,1], \ \ 0 < q < 1.$$

Take $\phi_q(x) = \frac{1}{\sqrt{(1+q)^2 - 4xq}}$, then $a_l(\phi_q) = \frac{q^l}{2l+1}$, $l = 0, 1, 2, \ldots$. We know by (1) and (2) that the Mercer kernel

$$K(x, y) = \sum_{l=0}^{+\infty} q^l P_l(x) P_l(y), \quad x, y \in [0,1] \tag{3}$$

is

$$K(x, y) = \frac{1}{\pi} \int_0^1 \Big((1+q)^2 - 4q(2x-1)(2y-1) - 16q(2u-1) \times \sqrt{x(1-x)y(1-y)}\Big)^{-1/2} \big(u(1-u)\big)^{-1/2}\, du, \quad x, y \in [0,1].$$

Let $K_t(x) = K(x, t)$ and $\mathcal{H}_K$ be the closure of the linear span of $\{K_t(x) : t \in [0,1]\}$. It is easy to check that for any distinct points $t_1, t_2, \ldots, t_m \in [0,1]$ the matrix $\{K_{t_i}(t_k)\}_{m \times m}$ is symmetric and positive definite. $K_t(x)$ is therefore a Mercer kernel. Defining a binary operation in $\mathcal{H}_K$ by

$$(f, g)_{\mathcal{H}_K} = \sum_{k,i} d_i c_k K_{t_i}(t_k),$$

for $f(x) = \sum_k c_k K_{t_k}(x) \in \mathcal{H}_K$ and $g(x) = \sum_i d_i K_{t_i}(x) \in \mathcal{H}_K$, we have by [2] that $\mathcal{H}_K$ will become a reproducing kernel Hilbert space.

Let $I_K : \mathcal{H}_K \to C[0,1]$ be the embedding operator from $\mathcal{H}_K$ to $C[0,1]$, $I_K(B_R)$ be the embedding of the ball with radius R.

We shall estimate the lower and the upper bounds of $\mathcal{N}(I_K(B_R), \eta)$. The paper is organized as follows. The lower bounds for $\mathcal{N}(I_K(B_R), \eta)$ is estimated in the second section. Choosing the knot set $X_N = \{x^{(N)}\}_{k=0}^{N-1}$ as the zeroes of the Legendre polynomial of order N and taking $A_N = \{K_t(x)\}_{x,t\in X_N}$, we shall provide an upper estimate of $\|A_N^{-1}\|_{l^2(X_N)}$ with the help of the Gauss integral formula and the Lagrange interpolation operators for the algebraic polynomials. In the third section, we shall use the method given by Zhou in [16] to give some estimates for the upper bounds of the covering number under the condition that the series $\sum_{l=0}^{+\infty} l^s a_l$ is convergent for some real numbers $s\geqslant 4$. In the fourth section, we shall construct a kind of local algebraic polynomials reproducing basis functions, with which we give an upper estimate of the covering number $\mathcal{N}(I_K(B_R),\eta)$ for the Mercer kernel

$$K(x,y)=\sum_{l=0}^{+\infty}\frac{1}{(1+l)^\alpha}\times P_l(x)P_l(y),\quad \alpha>2,\ \ x,y\in[0,1].$$

Throughout this paper, we shall denote by $\mathcal{N}$ the set of natural number. By $\mathcal{R}^N$ we denote the N-dimensional Euclidean space, and by $\mathcal{P}_N$ we denote the set of all algebraic polynomials of order $\leqslant N$. The biggest integer which $\leqslant a$ is denoted by $[a]$.

2. The lower bound estimates

In this section, we shall investigate the lower bound estimates of the covering numbers for some particular Mercer kernel Hilbert spaces.

Theorem 2.1. *Let a_l be a decreasing sequence such that $0<a_{l-1}\leqslant \frac{C'}{l^\alpha}$ for two given constants $C'>0,\alpha>2$ and all l, $K(x,y)$ be a Mercer kernel defined as* (1). *Then, there exists a constant $c_0>0$ such that for $0<\eta\leqslant\frac{R}{2c_0}$ there holds*

$$\ln\mathcal{N}\left(\overline{I_K(B_R)},\eta\right)\geqslant\frac{\ln 2}{2^{1+(2/\alpha)}c_0^{2/\alpha}}\left(\frac{R}{\eta}\right)^{2/\alpha}-\ln 2. \tag{4}$$

Corollary 2.1. *Let $K(x,y)$ be the Mercer kernel defined by* (3). *Then, there is a constant $c_0>0$ such that for $0<\eta\leqslant\frac{R}{2c_0}$ there holds*

$$\ln\mathcal{N}\left(\overline{I_K(B_R)},\eta\right)\geqslant\frac{\ln 2}{2^{5/3}c_0^{2/3}}\left(\frac{R}{\eta}\right)^{2/3}-\ln 2. \tag{5}$$

To prove Theorem 2.1, we first give a lemma.

Lemma 2.1. *Let $\phi\in L^1[0,1]$ satisfy $a_l(\phi)>0$ and $\sum_{l=0}^{+\infty} l a_l(\phi)<+\infty$. $N\geqslant 2$ is a given integer, $X_N=\{x_k^{(N)}\}_{k=0}^{N-1}$ is the knot set of the zeroes of the Legendre algebraic polynomial $P_N(x)$ and the zeroes are arranged in the increasing order, i.e.,*

$$0<x_0^{(N)}<x_1^{(N)}<\cdots<x_{N-1}^{(N)}<1.$$

$K(x,y)$ is defined as (1) *and the matrix $A_N:=(K(x,y))_{x,y\in X_N}$. Then, there exists a constant $C>0$ such that*

$$\|A_N^{-1}\|_{l^2(X_N)}\leqslant\frac{C}{N\min_{0\leqslant l\leqslant N-1}a_l(\phi)}. \tag{6}$$

Proof. For any $V = (v_0, \dots, v_{N-1})^\top \in \mathcal{R}^N$ there holds

$$\begin{aligned}
V^\top A_N V &= \sum_{k,j=0}^{N-1} v_k K_{x_k^{(N)}}(x_j^{(N)}) v_j \\
&= \sum_{k,j=0}^{N-1} v_k \sum_{l=0}^{+\infty} (2l+1) a_l(\phi) P_l(x_k^{(N)}) P_l(x_j^{(N)}) v_j \\
&= \sum_{l=0}^{+\infty} a_l(\phi) \left| \sum_{k=0}^{N-1} v_k \sqrt{2l+1} P_l(x_k^{(N)}) \right|^2 \\
&\geqslant \min_{0 \leqslant l \leqslant N-1} a_l(\phi) \sum_{l=0}^{N-1} \left| \sum_{k=0}^{N-1} v_k \sqrt{2l+1} P_l(x_k^{(N)}) \right|^2 \\
&= \min_{0 \leqslant l \leqslant N-1} a_l(\phi) \int_0^1 \left| \sum_{l=0}^{N-1} \left(\sum_{k=0}^{N-1} v_k \sqrt{2l+1} P_l(x_k^{(N)}) \right) \sqrt{2l+1} P_l(x) \right|^2 dx \\
&= \min_{0 \leqslant l \leqslant N-1} a_l(\phi) \int_0^1 \left| \sum_{k=0}^{N-1} v_k \left(\sum_{l=0}^{N-1} (2l+1) P_l(x_k^{(N)}) P_l(x) \right) \right|^2 dx.
\end{aligned}$$

Take $K_N(x, y) = \sum_{l=0}^{N-1} (2l+1) P_l(y) P_l(x)$, then

$$\begin{aligned}
\sum_{k=0}^{N-1} v_k \left(\sum_{l=0}^{N-1} (2l+1) P_l(x_k^{(N)}) P_l(x) \right) &= \sum_{k=0}^{N-1} v_k K_N(x, x_k^{(N)}) \\
&= \sum_{k=0}^{N-1} \frac{v_k}{\lambda_k^{(N)}} \times \lambda_k^{(N)} K_N(x, x_k^{(N)}).
\end{aligned}$$

Let

$$l_{k,N}(x) = \frac{P_N(x)}{P'_N(x_k^{(N)})(x - x_k^{(N)})}, \quad x \in [0, 1], \quad k = 0, 1, \dots, N-1$$

be the Lagrange basic interpolating functions based on X_N, and

$$L_N(x) = \sum_{k=0}^{N-1} \frac{v_k}{\lambda_k^{(N)}} l_{k,N}(x), \quad x \in [0, 1],$$

with the Cotes numbers on X_N being defined by

$$\lambda_k^{(N)} = \int_0^1 l_{k,N}^2(t)\, dt, \quad k = 0, 1, \dots, N-1.$$

Then, the interpolating property of Lagrange basis functions makes

$$L_N(x_k^{(N)}) = \frac{v_k}{\lambda_k^{(N)}}, \quad k = 0, 1, 2, \dots, N-1.$$

Therefore,

$$\sum_{k=0}^{N-1} \frac{v_k}{\lambda_k^{(N)}} \times \lambda_k^{(N)} K_N(x, x_k^{(N)}) = \sum_{k=0}^{N-1} \lambda_k^{(N)} L_N(x_k^{(N)}) K_N(x, x_k^{(N)}).$$

Since $L_N(\cdot)K_N(x,\cdot) \in \mathcal{P}_{2N-1}$, the Gauss integral formula (see, [12])

$$\int_0^1 p(x)\,dx = \sum_{k=0}^{N-1} \lambda_k^{(N)} p(x_k^{(N)}), \quad p \in \mathcal{P}_{2N-1}$$

yields

$$\sum_{k=0}^{N-1} \lambda_k^{(N)} L_N(x_k^{(N)}) K_N(x, x_k^{(N)}) = \int_0^1 L_N(u) K_N(x,u)\,du = L_N(x).$$

Hence,

$$\begin{aligned} V^\top A_N V &\geqslant \min_{0\leqslant l\leqslant N-1} a_l(\phi) \int_0^1 \left| L_N(x) \right|^2 dx \\ &= \min_{0\leqslant l\leqslant N-1} a_l(\phi) \int_0^1 \left| \sum_{k=0}^{N-1} \frac{v_k}{\lambda_k^{(N)}} l_{k,N}(x) \right|^2 dx \\ &= \min_{0\leqslant l\leqslant N-1} a_l(\phi) \sum_{k=0}^{N-1} \frac{v_k^2}{\lambda_k^{(N)}} \\ &\geqslant \frac{\min_{0\leqslant l\leqslant N-1} a_l(\phi)}{\max_{0\leqslant l\leqslant N-1} \lambda_l^{(N)}} \sum_{k=0}^{N-1} \left| v_k \right|^2 \\ &= \frac{\min_{0\leqslant l\leqslant N-1} a_l(\phi)}{\max_{0\leqslant l\leqslant N-1} \lambda_l^{(N)}} V^\top V, \end{aligned}$$

where we have used the fact that (see, [12])

$$\int_0^1 l_{k,N}(x) l_{j,N}(x)\,dx = \lambda_k^{(N)} \delta_{k,j}.$$

It follows that $\frac{\min_{0\leqslant l\leqslant N-1} a_l(\phi)}{\max_{0\leqslant l\leqslant N-1} \lambda_l^{(N)}}$ is smaller than any eigenvalues of the matrix A_N. Hence,

$$\|A_N^{-1}\|_{l^2(X_N)} \leqslant \frac{\max_{0\leqslant l\leqslant N-1} \lambda_l^{(N)}}{\min_{0\leqslant l\leqslant N-1} a_l(\phi)}. \tag{7}$$

On the other hand, by [8] we know

$$\lambda_k^{(N)} = \left[\sum_{l=0}^{N-1} \left| P_l(x_k^{(N)}) \right|^2 \right]^{-1} \sim \frac{\sqrt{x_k^{(N)}(1 - x_k^{(N)})}}{N},$$

i.e., there is a constant number $C > 0$ such that

$$\frac{1}{C}\frac{\sqrt{x_k^{(N)}(1-x_k^{(N)})}}{N} \leqslant \lambda_k^{(N)} \leqslant \frac{C\sqrt{x_k^{(N)}(1-x_k^{(N)})}}{N}.$$

Then, (7) makes (6). □

Lemma 2.1 provides an upper estimator for the l^2-norm of the inverse of the Mercer kernel matrix, itself is independence.

Proof of Theorem 2.1. We first recall the general lower bound estimates given by Zhou in [17]:

Let $X \subset \mathcal{R}^n$ be a compact set, K be a Mercer kernel on X, $N \in \mathcal{N}$, and $X_N := \{x_1, x_2, \ldots, x_N\} \in X$ yield an invertible Gramian matrix $A_N := (K(x_i, x_j))_{i,j=1}^N$. Then,

$$\mathcal{N}\left(\overline{I_K(B_R)}, \frac{\eta}{2}\right) \geqslant 2^N - 1$$

provided that $\|A_N^{-1}\|_{l^2} \leqslant \frac{1}{N}(\frac{R}{\eta})^2$.

Let X_N and A_N be defined as in Lemma 2.1. Since a_l are decreasing on l and $a_l \leqslant \frac{C'}{(1+l)^\alpha}$, we have by Lemma 2.1 that

$$\|A_N^{-1}\|_{l^2(X_N)} \leqslant \frac{c_0^2}{N} \times N^\alpha, \quad c_0 = \sqrt{\frac{C}{C'}}.$$

Hence, $c_0^2 N^\alpha \leqslant \left(\frac{R}{\eta}\right)^2$. For $0 < \eta \leqslant \frac{R}{c_0}$ we can choose $N \in \mathcal{N}$ such that

$$c_0^2 N^\alpha \leqslant \left(\frac{R}{\eta}\right)^2 < c_0^2 (N+1)^\alpha.$$

It follows $N + 1 > \left[\frac{1}{c_0^2}\left(\frac{R}{\eta}\right)^2\right]^{1/\alpha}$. Therefore,

$$N \geqslant \frac{N+1}{2} > \frac{1}{2c_0^{2/\alpha}}\left(\frac{R}{\eta}\right)^{2/\alpha},$$

and

$$\mathcal{N}\left(\overline{I_K(B_R)}, \frac{\eta}{2}\right) \geqslant 2^N - 1 \geqslant 2^{N-1} \geqslant 2^{\frac{1}{2c_0^{2/\alpha}}\left(\frac{R}{\eta}\right)^{2/\alpha} - 1}.$$

Consequently,

$$\ln \mathcal{N}\left(\overline{I_K(B_R)}, \frac{\eta}{2}\right) \geqslant \left(\frac{1}{2c_0^{2/\alpha}}\left(\frac{R}{\eta}\right)^{2/\alpha} - 1\right)\ln 2.$$

(4) thus holds. □

Proof of Corollary 2.1. Let $h = \frac{1}{q} - 1$. Then for $N \geqslant 2$ we have $(1+h)^{N-1} > C_{N-1}^2 h^2 = \frac{(N-1)(N-2)}{2} h^2$. It follows

$$\frac{q^{N-1}}{2N-1} \leqslant \frac{2}{(2N-1)(N-1)(N-2)h^2} \leqslant \frac{8}{N^3 h^2} = 8\left(\frac{q}{1-q}\right)^2 \frac{1}{N^3}.$$

Taking $\alpha = 3$, $C' = 8(\frac{q}{1-q})^2$ and $c_0 = \sqrt{\frac{C}{C'}}$ in Theorem 2.1, we have (5). □

3. Mercer kernels with smoothness

If $a_l > 0$ and there exists $s \geqslant 4$ such that $\sum_{l=0}^{+\infty} l^s a_l < +\infty$, then the Mercer kernel $K(x, y) = \sum_{l=0}^{+\infty} a_l P_l(x) P_l(y)$ has certain smoothness. In this case, we can estimate the upper bounds of $\mathcal{N}\left(\overline{I_K(B_R)}, \eta\right)$ by the way given in [16]. We give the following results.

Theorem 3.1. *Let $s \geqslant 1$ be an integer, $a_l > 0$ and $\sum_{l=0}^{+\infty} l^{4s} a_l < +\infty$. $K(x, y) = \sum_{l=0}^{+\infty} a_l P_l(x) P_l(y)$ is thus a Mercer kernel on $[0, 1] \times [0, 1]$, $C_1 \geqslant \frac{(2s)^{2s}}{(s-1)!} \left(\sum_{l=s+1}^{+\infty} l^{4s} a_l\right)^{1/2}$ is a constant number. Then for $0 < \eta \leqslant 2RC_1 s^{-s}$ there holds*

$$\ln \mathcal{N}\left(\overline{I_K(B_R)}, \eta\right) \leqslant \left[1 + 2\left(\frac{2RC_1}{\eta}\right)^{1/s}\right] \ln\left[16\left(\sum_{l=0}^{+\infty} a_l\right)\left(1 + 2\left(\frac{2RC_1}{\eta}\right)^{1/s}\right)^{1/2} \times (2C_1)^{1/s} 4^{((2RC_1)/\eta)^{1/s}} \left(\frac{R}{\eta}\right)^{1+(1/s)}\right]. \quad (8)$$

Some special cases of Theorem 3.1 are following Corollaries.

Corollary 3.1. *Let $K(x, y) = \sum_{l=0}^{+\infty} a_l P_l(x) P_l(y)$ be a Mercer kernel on $[0, 1] \times [0, 1]$ and there exists a constant $C > 0$ such that $0 < a_l \leqslant \frac{C}{(1+l)^\alpha}$, $\alpha > 5$. Then, for $C_1 \geqslant \frac{4(2^\alpha - \alpha + 5)}{2^\alpha(\alpha - 5)}$ and $0 < \eta \leqslant 2RC_1$ there holds*

$$\ln \mathcal{N}\left(\overline{I_K(B_R)}, \eta\right) \leqslant \left(1 + \frac{4C_1 R}{\eta}\right) \ln\left[\frac{16C}{\alpha - 1}\left(1 + \frac{4C_1 R}{\eta}\right)^{1/2} \times 2C_1 4^{(2C_1 R)/\eta} \left(\frac{R}{\eta}\right)^2\right]. \quad (9)$$

Corollary 3.2. *Let $K(x, y)$ be defined as (3). Then, for $C_1 \geqslant 4\left(\frac{24q^4}{(1-q)^5} + \frac{36q^3}{(1-q)^4} - \frac{8q+8q^2}{(1-q)^3} + \frac{4q}{(1-q)^2} - q\right)$ and $0 < \eta \leqslant 2RC_1$ there holds*

$$\ln \mathcal{N}\left(\overline{I_K(B_R)}, \eta\right) \leqslant \left(1 + \frac{4C_1 R}{\eta}\right) \ln\left[\frac{16}{1-q}\left(1 + \frac{4C_1 R}{\eta}\right)^{1/2} \times 2C_1 4^{(2C_1 R)/\eta} \left(\frac{R}{\eta}\right)^2\right]. \quad (10)$$

To prove Theorem 3.1, we first give some lemmas.

Assume that $\{X_N : N \in \mathcal{N}\}$ is a family of finite subsets of X such that

$$d_N := \max_{x\in X} \min_{y\in X_N} d(x, y) \to 0, \quad (N \to +\infty).$$

This means that the discrete knot X_N becomes dense in X as N tends to the infinity. Let the function measuring the regularity of K be defined by

$$\varepsilon_K(N) := \sup_{x\in X}\left[\inf\left(K(x,x) - 2\sum_{y\in X_N} w_y K(x,y) + \sum_{y,t\in X_N} w_y K(y,t) w_t : w_y \in \mathcal{R}^1\right)^{1/2}\right],$$

the cardinality of the set X_N be ΞX_N, and A_N be the positive definite matrix $A_N := [K(y,t)]_{y,t\in X_N}$. Then, Zhou gave in [16] the following general upper estimate for the covering number $\mathcal{N}(\overline{I_K(B_R)}, \eta)$.

Lemma 3.1 (*see, Zhou [16]*). *Let $K(x, y)$ be a Mercer kernel, I_K be given as in Section* 1. *Then for $0 < \eta \leqslant \frac{R}{2}$ there holds*

$$\ln \mathcal{N}(\overline{I_K(B_R)}, \eta) \leqslant (\Xi X_N) \ln\left[8\|K\|_\infty^{3/2}(\Xi X_N)\|A_N^{-1}\|_{l^2(X_N)}\frac{R}{\eta}\right], \tag{11}$$

where N is any integer satisfying $\varepsilon_K(N) \leqslant \frac{\eta}{2R}$.

Lemma 3.2. *Let $s \geqslant 1$ be a given integer, $a_l > 0$ satisfy $\sum_{l=0}^{+\infty} l^{4s} a_l < +\infty$. $K(x, y) = \sum_{l=0}^{+\infty} a_l P_l(x) P_l(y)$, $x, y \in [0, 1]$. Choose the knot set as $X_N = \{\frac{j}{N}\}_{j=0}^{N-1}$ and take*

$$\varepsilon_K(N) = \sup_{x\in[0,1]}\left[\inf\left(K(x,x) - 2\sum_{y\in X_N} w_y K(x,y) + \sum_{y,t\in X_N} w_y K(y,t) w_t : w_y, w_t \in \mathcal{R}^1\right)^{1/2}\right],$$

then, for $N \geqslant s$ there holds

$$\varepsilon_K(N) \leqslant \frac{(2s)^{2s}}{N^s(s-1)!}\left(\sum_{l=s+1}^{+\infty} l^{4s} a_l\right)^{1/2}. \tag{12}$$

Proof. Since $X_N = \{\frac{j}{N}\}_{j=0}^{N-1}$, we have $d_N = \frac{1}{N} \to 0, (N \to +\infty)$. As in [16], we define Lagrange interpolation functions

$$w_{l,s}(t) = \prod_{j\in\{0,1,2,\ldots,s\}\backslash\{l\}} \frac{t - j/s}{l/s - j/s} = \prod_{j\in\{0,1,2,\ldots,s\}\backslash\{l\}} \frac{st - j}{l - j}.$$

Then, $w_{l,s}(\frac{m}{s}) = \delta_{l,m}, l, m = 0, 1, 2, \ldots, s$. For $x \in [0, 1]$ we can find an $m \in \{0, 1, \ldots, N-s\}$ such that $x \in [\frac{m}{N}, \frac{m+s}{N}]$. Choose

$$w_{j/N} = \begin{cases} w_{i,s}\left(\dfrac{Nx-m}{s}\right), & j = m+i, \quad i \in \{0, \ldots, s\}, \\ 0 & \text{otherwise,} \end{cases}$$

then,

$$\begin{aligned}
&K(x,x) - 2\sum_{y \in X_N} w_y K(x,y) + \sum_{y,t \in X_N} w_y K(y,t) w_t \\
&= K(x,x) - 2\sum_{i=0}^{s} w_{i,s}\left(\frac{Nx-m}{s}\right) K\left(x, \frac{m+i}{N}\right) \\
&\quad + \sum_{i,j=0}^{s} w_{i,s}\left(\frac{Nx-m}{s}\right) K\left(\frac{m+i}{N}, \frac{m+j}{N}\right) w_{j,s}\left(\frac{Nx-m}{s}\right) \\
&= \sum_{l=0}^{+\infty} a_l P_l^2(x) - 2\sum_{i=0}^{s} w_{i,s}\left(\frac{Nx-m}{s}\right) \sum_{l=0}^{+\infty} a_l P_l(x) P_l\left(\frac{m+i}{N}\right) \\
&\quad + \sum_{i,j=0}^{s} w_{i,s}\left(\frac{Nx-m}{s}\right)\left(\sum_{l=0}^{+\infty} a_l P_l\left(\frac{m+i}{N}\right) P_l\left(\frac{m+j}{N}\right)\right) w_{j,s}\left(\frac{Nx-m}{s}\right) \\
&= \sum_{l=0}^{+\infty} a_l \left(P_l^2(x) - 2\sum_{i=0}^{s} w_{i,s}\left(\frac{Nx-m}{s}\right) P_l(x) P_l\left(\frac{m+i}{N}\right)\right. \\
&\quad \left. + \sum_{i,j=0}^{s} w_{i,s}\left(\frac{Nx-m}{s}\right) P_l\left(\frac{m+i}{N}\right) P_l\left(\frac{m+j}{N}\right) w_{j,s}\left(\frac{Nx-m}{s}\right)\right) \\
&= \sum_{l=0}^{+\infty} a_l \left| P_l(x) - \sum_{i=0}^{s} P_l\left(\frac{m+i}{N}\right) w_{i,s}\left(\frac{Nx-m}{s}\right)\right|^2 \\
&= \sum_{l=s+1}^{+\infty} a_l \left| \sum_{i=0}^{s} \left(P_l(x) - P_l\left(\frac{m+i}{N}\right)\right) w_{i,s}\left(\frac{Nx-m}{s}\right)\right|^2 \\
&= \sum_{l=s+1}^{+\infty} a_l \left| \sum_{i=0}^{s} \left(\sum_{k=1}^{s-1} \frac{1}{k!} \frac{\partial^k}{\partial y^k} P_l(x) \left(\frac{m+i}{N} - x\right)^k \right.\right. \\
&\quad \left.\left. + \frac{1}{s!} \frac{\partial^s}{\partial y^s} P_l(\xi_i) \left(\frac{m+i}{N} - x\right)^s\right) w_{i,s}\left(\frac{Nx-m}{s}\right)\right|^2,
\end{aligned}$$

where ξ_i is a real number between x and $\frac{m+i}{N}$. Since $\{w_{i,s}(x)\}_{i=0}^s$ are Lagrange basic interpolating functions based on $\{\frac{i}{s}\}_{i=0}^s$, we have

$$\begin{aligned}
&\sum_{i=0}^{s} w_{i,s}\left(\frac{Nx-m}{s}\right)\left(\frac{m+i}{N}-x\right)^k \\
&\quad = \sum_{i=0}^{s} w_{i,s}\left(\frac{Nx-m}{s}\right)\left(\frac{m+s\times(i/s)}{N}-x\right)^k \\
&\quad = \left(\frac{m+s\times(Nx-m)/s}{N}-x\right)^k = 0, \quad \forall 1\leqslant k\leqslant s-1.
\end{aligned}$$

Then, the original equation

$$\begin{aligned}
&= \sum_{l=s+1}^{+\infty} a_l \left|\sum_{i=0}^{s}\frac{1}{s!}\frac{\partial^s}{\partial y^s}P_l(\xi_i)\left(\frac{m+i}{N}-x\right)^s w_{i,s}\left(\frac{Nx-m}{s}\right)\right|^2 \\
&\leqslant \left(\frac{1}{s!}\right)^2 \sum_{l=s+1}^{+\infty} a_l \left\|\frac{\partial^s}{\partial y^s}P_l\right\|_\infty^2 \left|\sum_{i=0}^{s}\left|\frac{m+i}{N}-x\right|^s \left|w_{i,s}\left(\frac{Nx-m}{s}\right)\right|\right|^2 \\
&\leqslant \left(\frac{1}{s!}\right)^2 \left(\frac{2s^2}{N}\right)^{2s} \sum_{l=s+1}^{+\infty} a_l \left\|\frac{\partial^s}{\partial y^s}P_l\right\|_\infty^2 \left|\sum_{i=0}^{s}\left|w_{i,s}\left(\frac{Nx-m}{s}\right)\right|\right|^2 \\
&\leqslant \left[\frac{1}{s!}\left(\frac{2s^2}{N}\right)^s\right]^2 \sum_{l=s+1}^{+\infty} a_l \left\|\frac{\partial^s}{\partial y^s}P_l\right\|_\infty^2 \left(s2^s\right)^2 \\
&\leqslant \left[\frac{(2s)^{2s}}{(s-1)!N^s}\right]^2 \sum_{l=s+1}^{+\infty} l^{4s} a_l,
\end{aligned}$$

where we have used the facts that $\|P_l\|_\infty \leqslant 1$, $\|\frac{\partial^s}{\partial y^s}P_l\|_\infty \leqslant l^{2s}$ and (see, [16])

$$\sum_{\alpha\in X_N} |w_{\alpha,N}(x)| \leqslant N2^N, \quad x\in[0,1]. \quad \square \tag{13}$$

Proof of Theorem 3.1. Since $X_N = \{\frac{j}{N}\}_{j=0}^N$, we have by [16, Theorem 1] that for any $x\in[0,1]$ there holds

$$\left|f(x) - \sum_{\alpha\in X_N} f(\alpha)w_\alpha(x)\right| \leqslant \|f\|_{\mathcal{H}_K}\varepsilon_K(N) \leqslant R\times\frac{(2s)^{2s}}{N^s(s-1)!}\left(\sum_{l=s+1}^{+\infty} l^{4s}a_l\right)^{1/2}.$$

Then for $N > \left[\frac{2RC_1}{\eta}\right]^{1/s}$ one has

$$\left\|f(x) - \sum_{\alpha\in X_N} f(\alpha)w_\alpha(x)\right\|_{C[0,1]} \leqslant \frac{\eta}{2}.$$

On the other hand, since

$$|f(x)| = |(K(x,\cdot), f(\cdot))_{\mathcal{H}_K}| \leqslant R\sqrt{K(x,x)} \leqslant R\left(\sum_{l=0}^{+\infty} a_l\right)^{1/2},$$

we have

$$\|\{f(\alpha)\}\|_{l^2(X_N)} \leqslant R\left(\sum_{l=0}^{+\infty} a_l\right)^{1/2} (N+1)^{1/2}.$$

By [2] we know that if E is a finite dimension space with $\dim E = m$, then,

$$\ln \mathcal{N}(B_r, \varepsilon) \leqslant m \ln\left(\frac{4r}{\varepsilon}\right).$$

The dimension of $l^2(X_N)$ is $N+1$. Let

$$r := R\left(\sum_{l=0}^{+\infty} a_l\right)^{1/2} (N+1)^{1/2}$$

and $\varepsilon := \frac{\eta}{2(N2^N)}$. Then, there exists $\{c^l : l = 1, 2, \ldots, [\frac{4r}{\varepsilon}]^{N+1}\} \subset l^2(X_N)$ such that for any $d \in l^2(X_N)$ with $\|d\|_{l^2(X_N)} \leqslant r$, we can find some l satisfying

$$\|d - c^l\|_{l^2(X_N)} \leqslant \varepsilon.$$

By (13) we have

$$\begin{aligned}\left\|\sum_{\alpha\in X_N} c_\alpha^l w_\alpha(x) - \sum_{\alpha\in X_N} d_\alpha w_\alpha(x)\right\|_C &\leqslant \left\|\sum_{\alpha\in X_N}\left(c_\alpha^l - d_\alpha\right) w_\alpha(x)\right\|_C \\ &\leqslant \|d - c^l\|_{l^2(X_N)} \| \|w_\alpha(x)\|_{l^2(X_N)}\|_{C[-,1]} \\ &\leqslant N2^N\varepsilon \leqslant \frac{\eta}{2}.\end{aligned}$$

Since $\|\{f(\alpha)\}_{\alpha\in X_N}\|_{l^2} \leqslant r$, we have

$$\begin{aligned}\left\|f(x) - \sum_{\alpha\in X_N} c_\alpha^l w_\alpha(x)\right\|_C &\leqslant \left\|f(x) - \sum_{\alpha\in X_N} f(\alpha) w_\alpha(x)\right\|_C \\ &\quad + \left\|\sum_{\alpha\in X_N} c_\alpha^l w_\alpha(x) - \sum_{\alpha\in X_N} f(\alpha) w_\alpha(x)\right\|_C \\ &\leqslant \frac{\eta}{2} + \frac{\eta}{2} = \eta.\end{aligned}$$

We then have covered $I_K(B_R)$ by balls with radii η and centers $\sum_{\alpha\in X_N} c_\alpha^l w_\alpha(x)$. Therefore,

$$\mathcal{N}\left(\overline{I_K(B_R)}, \eta\right) \leqslant \left(\frac{4r}{\varepsilon}\right)^{N+1},$$

i.e.,

$$\ln \mathcal{N}\left(\overline{I_K(B_R)}, \eta\right) \leqslant (N+1)\ln\left(\frac{4r}{\varepsilon}\right) \leqslant (N+1)\ln\left[8\left(\sum_{l=0}^{+\infty} a_l\right)^{1/2}(N+1)^{1/2}(N2^N)\frac{R}{\eta}\right].$$

Since $0 < \eta \leqslant 2RC_1 s^{-s}$, for $N \geqslant (\frac{2RC_1}{\eta})^{1/s}$ we have $N \geqslant s$ and $\frac{2RC_1}{\eta} \geqslant 1$. Therefore, we can find $N \in \mathcal{N}$ such that

$$N \leqslant 2\left(\frac{2RC_1}{\eta}\right)^{1/s}.$$

Consequently,

$$\begin{aligned}\ln \mathcal{N}\left(\overline{I_K(B_R)}, \eta\right) &\leqslant \left(1+2\left(\frac{2RC_1}{\eta}\right)^{1/s}\right)\ln\left[8\left(\sum_{l=0}^{+\infty} a_l\right)\left[1+2\left(\frac{2RC_1}{\eta}\right)^{1/s}\right]^{1/2}\right. \\ &\quad \left.\times\left[2\left(\frac{2RC_1}{\eta}\right)^{1/s} 2^{2(2RC_1/\eta)^{1/s}}\right]\frac{R}{\eta}\right] \\ &= \left[1+2\left(\frac{2RC_1}{\eta}\right)^{1/s}\right]\ln\left[16\left(\sum_{l=0}^{+\infty} a_l\right)\left[1+2\left(\frac{2RC_1}{\eta}\right)^{1/s}\right]^{1/2}\right. \\ &\quad \left.\times(2C_1)^{1/s}4^{(2RC_1/\eta)^{1/s}}\left(\frac{R}{\eta}\right)^{1+(1/s)}\right]. \qquad \square\end{aligned}$$

Proof of Corollary 3.1. Let $s = 1$ in Theorem 4.1. Then, we have by

$$\sum_{l=0}^{+\infty} \frac{l^4}{(1+l)^\alpha} \leqslant \int_0^{+\infty} \frac{dl}{(1+l)^{\alpha-4}} - \frac{1}{2^\alpha} = \frac{2^\alpha - \alpha + 5}{2^\alpha(\alpha-5)}.$$

that (9) holds. $\square$

Proof of Corollary 3.2. Taking $s = 1$, $\sum_0^{+\infty} q^l = \frac{1}{1-q}$, and

$$\sum_{l=0}^{+\infty} l^4 q^l = \frac{24q^4}{(1-q)^5} + \frac{36q^3}{(1-q)^4} - \frac{8q+8q^2}{(1-q)^3} + \frac{4q}{(1-q)^2} - q$$

in Theorem 4.1,we have (10). $\square$

4. The general mercer kernels

Theorem 3.1 requires that $K(x, y)$ has certain smoothness which even does not suit to the usual sequence $a_l = \frac{1}{(1+l)^3}$. For such case we should ask for another way. The interpolating property and the uniformly boundedness of the local polynomial reproducing basis functions (see, [4,13]) remind us to construct a local polynomial reproducing basis functions associating with the knot set X_N to take the place of Lagrange basic functions used in the proof of Theorem 3.1. We show this way by a special Mercer kernel.

Theorem 4.1. *Let* $K(x,y)=\sum_{l=0}^{+\infty}\frac{1}{(1+l)^{\alpha}}P_l(x)P_l(y)$, $\alpha>2$. *Then, for* $0<\eta\leqslant\min\{\frac{3R}{\sqrt{(\alpha-1)2^{\alpha-7}}},\frac{R}{2}\}$ *there holds*

$$\ln\mathcal{N}\left(\overline{I_K(B_R)},\eta\right)\leqslant\frac{3}{\sqrt{(\alpha-1)2^{\alpha-11}}}\frac{R}{\eta}\ln\left[\frac{3}{(\alpha-1)^2\sqrt{2^{\alpha-17}}}\left(\frac{R}{\eta}\right)^2\right]. \tag{14}$$

To prove Theorem 4.1, we first give some lemmas.

Let V be a finite-dimensional vector space with norm $\|\cdot\|_V$ and let $Z\subset V^*$ be a finite set consisting of N functionals. Here, V^* denotes the dual space of V consisting of all linear and continuous functionals defined on V. If the mapping $T:V\to T(V)\subset\mathcal{R}^N$ defined by $T(v)=\{z(v)\}_{z\in Z}$ is injective, we call T a sampling operator and Z a norming set for V.

Lemma 4.1 (*see, Wendland [13, Theorem 3.4]*). *Suppose V is a finite-dimensional normed linear space and* $Z=\{z_1,z_2,\dots,z_N\}$ *is a norming set for V, T being the corresponding sampling operator. For every* $\psi\in V^*$ *there exists a vector* $u\in\mathcal{R}^N$ *depending only on* ψ *such that, for every* $v\in V$, $\psi(v)=\sum_{j=1}^{N}u_jz_j(v)$ *and*

$$\|u\|_{(\mathcal{R}^N)^*}\leqslant\|\psi\|_{V^*}\|T^{-1}\|, \tag{15}$$

where

$$\|T^{-1}\|=\sup_{v\in V\setminus\{0\}}\frac{\|v\|_V}{\|T(v)\|_{\mathcal{R}^N}}.$$

Lemma 4.2. *Let* X_N *be the knot set in Lemma* 2.1. *Then, there holds*

$$\tfrac{1}{6}\|p\|_{L^\infty[0,1]}\leqslant\max_{x\in X_N}|p(x)|\leqslant\|p\|_{L^\infty[0,1]},\qquad p\in\mathcal{P}_{[N/2]}. \tag{16}$$

Proof. Let $J_N(x)$ be the Legendre orthogonal polynomial of order N on $[-1,1]$. The zeroes of $J_N(x)$ are $\{x_{k,(N)}\}_{k=0}^{N-1}$ in the increasing order. Taking $x_{k,(N)}=\cos\theta_k$, $0\leqslant k\leqslant N-1$, we have by [12, Theorem 6.3.2] that $|\theta_{k+1}-\theta_k|\leqslant\frac{2\pi}{2N+1}$. Set $u=\frac{1+x}{2}$, then, $P_N(u)=J_N(x)$. We thus have $x_k^{(N)}=x_{k,(N)}$, $k=0,1,\dots,N-1$. Let $p\in\mathcal{P}_{N/2}$ and $p^*(x)=p(\frac{1+x}{2})$, $x\in[-1,1]$. Then,

$$\|p\|_{L^\infty[0,1]}=|p(u_0)|=|p^*(x_0)|=\|p^*\|_{L^\infty[-1,1]},\quad u_0=\frac{1+x_0}{2}.$$

Let $x_{k_0,(N)}$ be the nearest knot to x_0, $x_0=\cos\theta_0$, $x_{k_0,(N)}=\cos\theta_{k_0}$. Taking $T(\theta)=p^*(\cos\theta)$, we have by the Bernstein inequality for trigonometrical polynomial and the mean theorem that

$$\begin{aligned}\|p\|_{L^\infty[0,1]}=\|p^*\|_{L^\infty[-1,1]}=|p^*(x_0)|=\|T(\theta_0)\|&\leqslant|T(\theta_0)-T(\theta_{k_0})|+|T(\theta_{k_0})|\\&\leqslant\frac{\pi}{2N+1}\|T'\|_{L^\infty[0,2\pi]}+|T(\theta_{k_0})|\\&\leqslant\frac{N\pi}{2(2N+1)}\|T\|_{L^\infty[0,2\pi]}+\max_{\theta_k}|T(\theta_k)|\\&\leqslant\frac{N\pi}{4N+2}\|p^*\|_{L^\infty[-1,1]}+\max_{x\in X_N}|p(x)|\\&\leqslant\tfrac{5}{6}\|p\|_{L^\infty[0,1]}+\max_{x\in X_N}|p(x)|.\end{aligned}$$

It follows that $\max_{x\in X_N}|p(x)|\geqslant\frac{1}{6}\|p\|_{L^\infty[0,1]}$. (16) thus holds. □

Lemma 4.3. *Let $X_N = \{x_\alpha^{(N)}\}_{0\leqslant\alpha\leqslant N-1}$ be the knot set in Lemma 2.1. Then, for every $x \in [0, 1]$ there exist real numbers $u_j(x)$ such that $\sum_{j=0}^{N-1} \left|u_j(x)\right| \leqslant 6$ and*

$$\sum_{j=0}^{N-1} u_j(x) p(x_j^{(N)}) = p(x), \quad p \in \mathcal{P}_{[N/2]}. \tag{17}$$

Proof. Let $(V, \|\cdot\|_V) = (\mathcal{P}_N, \|\cdot\|_{L^\infty})$. Defining a sampling operator by $T(p) = \{p(x_\alpha^{(N)})\}_{0\leqslant\alpha\leqslant N-1} \in \mathcal{R}^N$ and equipping with $\mathcal{R}^N$ the l^∞-norm, we know $(\mathcal{R}^N, \|\cdot\|_{l^\infty})^* = (\mathcal{R}^N, \|\cdot\|_{l^1})$ and by (16) $\|T^{-1}\| \leqslant 6$. Noticing that $|\delta_x(p)| = |p(x)| \leqslant \|p\|_{L^\infty}$,we have by Lemma 4.1 that there are functions $u_\alpha(x)$, $\alpha = 1, 2, \ldots, N-1$, such that (17) holds and

$$\sum_{0\leqslant\alpha\leqslant N-1} |u_\alpha(x)| \leqslant \|\delta_x\| \|T^{-1}\| \leqslant 6.$$

Proof of Theorem 4.1. By Lemma 3.1 what we need to do is to estimate $\varepsilon_K(N)$. In fact, let $X_N = \{x_\alpha^{(N)}\}_{0\leqslant\alpha\leqslant N-1}$ be defined as in Lemma 2.1 and $w_{x_\alpha^{(N)}} = u_\alpha(x)$, $\alpha = 0, 1, \ldots, N-1$. Then, by Lemma 4.3 and the fact that $|P_l(x)| \leqslant 1$ we have

$$\begin{aligned}
&K(x, x) - 2\sum_{y\in X_N} w_y K(x, y) + \sum_{y,t\in X_N} w_y K(y, t) w_t \\
&= \sum_{l=0}^{+\infty} \frac{1}{(1+l)^\alpha} \left| P_l(x) - \sum_{\alpha=0}^{N-1} P_l(x_\alpha^{(N)}) u_\alpha(x) \right|^2 \\
&\leqslant \sum_{l=[\frac{N}{2}]+1}^{+\infty} \frac{1}{(1+l)^\alpha} \left| \sum_{\alpha=0}^{N-1} (P_l(x) - P_l(x_\alpha^{(N)})) u_\alpha(x) \right|^2 \\
&\leqslant 4 \sum_{l=[\frac{N}{2}]+1}^{+\infty} \frac{1}{(1+l)^\alpha} \left(\sum_{\alpha=0}^{N-1} |u_\alpha(x)| \right)^2 \\
&\leqslant 144 \sum_{l=[\frac{N}{2}]+1}^{+\infty} \frac{1}{(1+l)^\alpha} \\
&\leqslant \frac{144}{(\alpha-1)([\frac{N}{2}]+2)^{\alpha-1}} \\
&\leqslant \frac{3^2 \times 2^{3+\alpha}}{(\alpha-1)(N+2)^{\alpha-1}}.
\end{aligned}$$

Hence, $\varepsilon_K(N) \leqslant \frac{3\times 2^{\frac{3+\alpha}{2}}}{\sqrt{(\alpha-1)(N+2)^{\alpha-1}}}$. If $\frac{3\times 2^{\frac{3+\alpha}{2}}}{\sqrt{(\alpha-1)(N+2)^{\alpha-1}}} < \frac{\eta}{2R}$, then $N + 2 \geqslant \left(\frac{3^2\times 2^{5+\alpha}}{\alpha-1}\right)^{1/(\alpha-1)} \left(\frac{R}{\eta}\right)^{2/(\alpha-1)}$. From $\left(\frac{3^2\times 2^{5+\alpha}}{\alpha-1}\right)^{1/(\alpha-1)} \left(\frac{R}{\eta}\right)^{2/(\alpha-1)} \geqslant 4$ we have $0 < \eta < 3\sqrt{\frac{2^{7-\alpha}}{\alpha-1}} R$. We can thus find $N \in \mathcal{N}$ such that $N \leqslant 12\sqrt{\frac{2^{7-\alpha}}{\alpha-1}}\frac{R}{\eta}$. Lemma 3.1 makes

$$\ln \mathcal{N}\left(\overline{I_K(B_R)}, \eta\right) \leqslant N \ln \left(8\left(\sum_{l=0}^{+\infty} \frac{1}{(1+l)^\alpha}\right)^{3/2} \frac{NR}{\eta}\right)$$
$$\leqslant 12\sqrt{\frac{2^{7-\alpha}}{\alpha-1}}\frac{R}{\eta} \ln\left[\frac{3\times 2^{(17-\alpha)/2}}{(\alpha-1)^2}\left(\frac{R}{\eta}\right)^2\right]$$
$$\leqslant \frac{3}{\sqrt{(\alpha-1)2^{\alpha-11}}}\frac{R}{\eta} \ln\left[\frac{3}{(\alpha-1)^2\sqrt{2^{\alpha-17}}}\left(\frac{R}{\eta}\right)^2\right]. \qquad \square$$

We now give a corollary to compare Theorem 2.1 with Theorem 4.1.

Corollary 4.1. *Let the Mercer kernel $K(x, y)$ be defined as in Theorem 4.1. c_0 is the constant in Theorem 2.1. If $\frac{\sqrt{(\alpha-1)2^{\alpha-7}}}{3} \geqslant 2\max\{c_0, 1\}$, then for $\frac{R}{\sqrt{(\alpha-1)2^{\alpha-7}}} < \eta \leqslant \frac{3R}{\sqrt{(\alpha-1)2^{\alpha-7}}}$ there are constants $C_1 > 0$, $C_2 > 0$, which depend only on c_0 and α, such that*

$$\frac{1}{C_1}\left(\frac{R}{\eta}\right)^{2/(\alpha+1)} - \frac{1}{C_2} \leqslant \ln \mathcal{N}(\overline{I_K(B_K)}, \eta) \leqslant C_1\left(\frac{R}{\eta}\right)^{2/(\alpha+1)} \pm C_2. \tag{18}$$

If $\frac{\sqrt{(\alpha-1)2^{\alpha-7}}}{3} < 2\max\{c_0, 1\}$, then for $\frac{R}{6\max\{c_0,1\}} < \eta \leqslant \frac{R}{2\max\{c_0,1\}}$ there are $C_1' > 0$, $C_2' > 0$, which depend only on c_0 and α, such that

$$\frac{1}{C_1'}\left(\frac{R}{\eta}\right)^{2/(\alpha+1)} - \frac{1}{C_2'} \leqslant \ln \mathcal{N}(\overline{I_K(B_K)}, \eta) \leqslant C_1'\left(\frac{R}{\eta}\right)^{2/(\alpha+1)} \pm C_2', \tag{19}$$

where the $\pm$ are determined by the $\mp$ of $\ln \frac{3}{(\alpha-1)^2\sqrt{2^{\alpha-17}}}$.

Proof. The Mercer kernel $K(x, y)$ can be rewritten as

$$K(x, y) = \sum_{l=0}^{+\infty}(2l+1) \times \frac{P_l(x)P_l(y)}{(2l+1)(1+l)^\alpha}, \quad x, y \in [0, 1],$$

and $a_{l-1} = \frac{1}{(2l-1)l^\alpha} \leqslant \frac{1}{l^{\alpha+1}}$, $l > 1$. We have by Theorem 2.1 that for $c_0 = \sqrt{C}$ and $0 < \eta \leqslant \frac{R}{2c_0}$ there holds

$$\ln \mathcal{N}\left(\overline{I_K(B_R)}, \eta\right) \geqslant \frac{\ln 2}{2^{1+\frac{2}{\alpha+1}} c_0^{2/(\alpha+1)}}\left(\frac{R}{\eta}\right)^{2/(\alpha+1)} - \ln 2. \tag{20}$$

If

$$\min\left\{\frac{3R}{\sqrt{(\alpha-1)2^{\alpha-7}}}, \frac{R}{2\max\{c_0, 1\}}\right\} = \frac{3R}{\sqrt{(\alpha-1)2^{\alpha-7}}},$$

then, the η and R in Theorem 4.1 satisfy $0 < \eta \leqslant \frac{3R}{\sqrt{(\alpha-1)2^{\alpha-7}}}$. In this case, for

$$\frac{R}{\sqrt{(\alpha-1)2^{\alpha-7}}} < \eta \leqslant \frac{3R}{\sqrt{(\alpha-1)2^{\alpha-7}}},$$

i.e.,

$$\frac{\sqrt{(\alpha-1)2^{\alpha-7}}}{3} \leqslant \frac{R}{\eta} < \sqrt{(\alpha-1)2^{\alpha-7}},$$

we have

$$\begin{aligned}\ln \mathcal{N}\left(\overline{I_K(B_R)}, \eta\right) &\leqslant \frac{3}{\sqrt{(\alpha-1)2^{\alpha-11}}}\frac{R}{\eta}\ln\left(\frac{R}{\eta}\right)^2 \\ &\quad - \frac{3}{\sqrt{(\alpha-1)2^{\alpha-11}}}\frac{R}{\eta}\ln\frac{3}{(\alpha-1)^2\sqrt{2^{\alpha-17}}} \\ &\leqslant \frac{3}{\sqrt{(\alpha-1)2^{\alpha-11}}}\left(\frac{R}{\eta}\right)^{2/(\alpha+1)}\left(\frac{\sqrt{(\alpha-1)2^{\alpha-7}}}{3}\right)^{-(\alpha-1)/(\alpha+1)} \\ &\quad \times \ln\left[(\alpha-1)2^{\alpha-7}\right] \\ &\quad - \frac{3A}{\sqrt{(\alpha-1)2^{\alpha-11}}}\ln\frac{3}{(\alpha-1)^2\sqrt{2^{\alpha-17}}}, \end{aligned} \tag{21}$$

where for $\frac{3}{(\alpha-1)^2\sqrt{2^{\alpha-17}}} \geqslant 1$ we have $A = \frac{\sqrt{(\alpha-1)2^{\alpha-7}}}{3}$, and for $0 < \frac{3}{(\alpha-1)^2\sqrt{2^{\alpha-17}}} < 1$ we have $A = \sqrt{(\alpha-1)2^{\alpha-7}}$. (20) and (21) make (18).

If

$$\min\left\{\frac{3R}{\sqrt{(\alpha-1)2^{\alpha-7}}}, \frac{R}{2\max\{c_0,1\}}\right\} = \frac{R}{2\max\{c_0,1\}},$$

then the η and R in Theorem 4.1 satisfy $0 < \eta \leqslant \frac{R}{2\max\{c_0,1\}}$. In this case, for

$$\frac{R}{6\max\{c_0,1\}} < \eta \leqslant \frac{R}{2\max\{c_0,1\}},$$

i.e.,

$$2\max\{c_0,1\} \leqslant \frac{R}{\eta} < 6\max\{c_0,1\},$$

we have

$$\begin{aligned}\ln \mathcal{N}\left(\overline{I_K(B_R)}, \eta\right) &\leqslant \frac{3}{\sqrt{(\alpha-1)2^{\alpha-11}}}\left(\frac{R}{\eta}\right)^{2/(\alpha+1)}(2\max\{c_0,1\})^{-(\alpha-1)/(\alpha+1)} \\ &\quad \times \ln\left[36\max\{c_0^2,1\}\right] \\ &\quad - \frac{3B}{\sqrt{(\alpha-1)2^{\alpha-11}}}\ln\frac{3}{(\alpha-1)^2\sqrt{2^{\alpha-17}}}, \end{aligned} \tag{22}$$

where for $\frac{3}{(\alpha-1)^2\sqrt{2^{\alpha-17}}} \geqslant 1$ we have $B = 2\max\{c_0,1\}$, and for $0 < \frac{3}{(\alpha-1)^2\sqrt{2^{\alpha-17}}} < 1$ we have $B = 6\max\{c_0,1\}$. (20) and (22) make (19).

Acknowledgments

The authors thank the referees for their valuable comments on this paper which made them rewrite this paper in a better form.

References

[1] R. Askey, S. Wainger, A convolution structure for Jacobi series, Am. J. Math. 91 (1969) 463–485.
[2] F. Cucker, S. Smale, On the mathematical foundations of learning, Bull. Am. Math. Soc. 39 (1) (2001) 1–49.
[3] Y. Guo, P.L. Bartlett, J. Shawe-Taylor, R.C. Williamson, Covering numbers for support vector machines, IEEE Trans. Inf. Theory 48 (2002) 239–250.
[4] K. Jetter, J. Stŏckler, J.D. Ward, Error estimates for scattered data interpolation on spheres, Math. Comput. 68 (226) (1999) 733–747.
[5] F.J. Narcowich, N. Sivakumar, J.D. Ward, On condition numbers associated with radial function interpolation, J. Math. Anal. Appl. 186 (1994) 457–485.
[6] F.J. Narcowich, J.D. Ward, Norms of inverses and condition numbers for matrices associated with scattered data, J. Approx. Theory 64 (1991) 69–94.
[7] F.J. Narcowich, J.D. Ward, Norms estimates for the inverses of a general class of scattered data radial function interpolation matrices, J. Approx. Theory 69 (1992) 84–109.
[8] P. Nevai, P. Vertesi, Mean convergence of Hermite–Fejér interpolation, J. Math. Anal. Appl. 105 (1) (1985) 26–58.
[9] V.S. Pawelke, Ein satz vom Jacksonschen type für algebraische polynome, Acta Sci. Math (Szeged) 33 (1972) 323–336.
[10] M. Pontil, A note on different covering numbers in learning theory, J. Complexity 19 (2003) 665–671.
[11] R. Schaback, Lower bounds for norms of inverses interpolation matrices for radial basis functions, J. Approx. Theory 79 (1994) 287–306.
[12] G. Szegŏ, Orthogonal Polynomials, American Mathematical Society, New York, 1967.
[13] H. Wendland, Scattered Data Approximation, Cambridge University Press, Cambridge, 2005.
[14] R.C. Williamson, A.J. Smola, B. Schŏlkopf, Generalization performance of regularization networks and support vector machine via entropy numbers of compact operators, IEEE Trans. Inf. Theory 47 (6) (2001) 2516–2532.
[15] Q. Wu, Y.M. Ying, D.X. Zhou, Learning theory: from regression to classification, in: Topics in Multivariate Approximation and Interpolation, Elsevier B.V., Amsterdam, 2004.
[16] D.X. Zhou, The covering number in learning theory, J. Complexity 18 (2002) 739–767.
[17] D.X. Zhou, Capacity of reproducing kernel spaces in learning theory, IEEE Trans. Inf. Theory 49 (7) (2003) 1743–1752.