

Error Analysis of Coefficient Regularized Regression and Non-Iid Sampling

Xinxin Chang¹, Shuang Chen²

¹Hebei University of Technology, Tianjin

²Beijing Institute of Petrochemical Technology, Beijing
Email: 958014075@qq.com

Received: Oct. 4th, 2017; accepted: Oct. 20th, 2017; published: Oct. 30th, 2017

Abstract

In this paper, we study the error analysis of indefinite kernel network with coefficient regularization for non-iid sampling. The framework under investigation is different from classical kernel learning. The kernel function satisfies only the continuity and uniform boundedness; the standard bound assumption for output data is abandoned and we carry out the error analysis with output sample values satisfying a generalized moment hypothesis. Satisfactory error bounds and learning rates independent of capacity are derived by the techniques of integral operator for this learning algorithm.

Keywords

Coefficient Regularization Regression, Indefinite Kernel, Strong Mixing Condition, Integral Operator

基于非独立同分布样本的系数正则化回归算法的误差分析

常欣欣¹, 陈爽²

¹河北工业大学, 天津

²北京石油化工学院, 北京
Email: 958014075@qq.com

收稿日期: 2017年10月4日; 录用日期: 2017年10月20日; 发布日期: 2017年10月30日

摘要

本文分析了基于非独立同分布样本的最小二乘系数正则化算法的误差。全文的框架不同于以往的经典核

学习方法。核函数仅仅满足连续性和一致有界性; 我们进行基于输出样本满足广义矩理论的误差分析, 而不再考虑标准的输出有界假设。最后通过利用积分算子技术得到了满意的与容量无关的误差界和学习率。

关键词

系数正则化回归, 不定核, 强混合条件, 积分算子

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在机器学习理论中, 最小二乘回归问题可以陈述如下:

设 X 为一个紧子集, $Y = R$. ρ 为 $Z: X \times Y$ 上未知的概率测度, $z = \{(x_i, y_i)\}_{i=1}^m \in Z^m$ 为服从 ρ 的一组独立同分布的取样。在回归学习中定义关于函数 $f: X \rightarrow Y$ 的推广误差,

$$\varepsilon(f) = \int_Z (f(x) - y)^2 d\rho$$

回归函数定义为

$$f_\rho(x) = E(y|x) = \int_Y y d\rho(y|x), x \in X$$

记 $\rho(y|x)$ 为 ρ 在 $X = x$ 时的条件分布。实际上, 由于 ρ 是未知的, f_ρ 是不能直接求出来的, 于是我们的目标是利用样本 z 产生关于 f_ρ 的一个最佳逼近。

传统的学习算法所选取的假设空间通常为与 Mercer 核相关的再生核 Hilbert 空间。基于核的学习算法具有良好的运算性质, 这得益于再生核 Hilbert 空间所特有的再生性。核函数最初用于模式识别, 现在已经广泛成功应用于机器学习的很多领域。在基于核函数的学习中, 核函数 $K: X \times X \rightarrow R$ 扮演着基函数的角色。

下面给出 Mercer 核和再生核 Hilbert 空间的定义:

定义 1.1 (Mercer 核) 设函数 $K: X \times X \rightarrow R$ 为连续的, 对称的, 半正定的, 则称 K 为 Mercer 核。

定义 1.2 (再生核 Hilbert 空间) 由 $K_x = K(x, \cdot), x \in X$ 的有限线性组合成的线性函数空间,

$$\left\{ \sum_{i=1}^m a_i K_{x_i} : a_i \in R, x_i \in X, i = 1, 2, \dots, m; m \in N \right\}$$

在该函数空间中定义内积

$$\left\langle \sum_{i=1}^n a_i K_{x_i}, \sum_{j=1}^m b_j K_{x_j} \right\rangle_K = \sum_{i=1}^n \sum_{j=1}^m a_i b_j K(x_i, x_j)$$

则该函数空间成为内积空间。其完备化的 Hilbert 空间称由核函数 K 所生成的再生核 Hilbert 空间 (RKHS), 记为 H_K 。

H_K 的再生性指

$$f(x) = \langle f, K_x \rangle_K \text{ 对任意 } f \in H_K, x \in X$$

大量的文献都有对 Mercer 核的研究学习, 可以见参考文献[1] [2] [3] [4]。

为了防止过拟合, 一般采用正则化方法[5]。常用的回归学习算法有正则化最小二乘算法与系数正则化算法。

正则化最小二乘算法形式:

$$f_{Z,K} = \arg \min_{f \in H_K} \frac{1}{m} \sum_{i=1}^m (y_i - f(x_i))^2 + \lambda \|f\|_K^2 \quad (1.1)$$

关于这个算法的学习可以看文献[6] [7] [8]。

接下来给出系数正则化算法(CRKN) [9] [10] [11]:

$$f_{Z,\lambda} = f_{\alpha_Z} \text{ 其中 } \alpha_Z = \arg \min_{\alpha \in R^m} \left\{ \frac{1}{m} \sum_{i=1}^m (f_{\alpha}(x_i) - y_i)^2 + \lambda \Omega_Z(f) \right\}, \lambda > 0$$

这里 $f_{\alpha} = \sum_{i=1}^m \alpha_i K(\cdot, x_i)$ 是与 Mercer 核有关的。在系数正则化算法中的假设空间取为:

$$H_{K,x} = f_{\alpha}(x) = \sum_{i=1}^m \alpha_i K(x, x_i) : \alpha = (\alpha_1, \alpha_2, \dots, \alpha_m) \in R^m, m \in N$$

注: 假设空间依赖输入数据 $x = \{x_i : i = 1, 2, \dots, m\}$ 。系数正则化被 Vapnik 在文献[5]中为了设计线性规划支持向量机而首次引用。系数正则化是有一些优点的, 我们可以见参考文献[10]。

本文中, 我们考虑不定核而不再考虑 Mercer 核, 所谓一般核 K 是指 K 只满足有界性和连续性, 不再满足对称性或者半正定性。对于非正定核通常通过以下定义来处理。

$$\text{定义 1.3 } \tilde{K}(u, v) = E_u K(x, u) K(t, v), \quad (1.2)$$

即 $\tilde{K}$ 是一个 Mercer 核。

更多的关于不定核的学习, 有: 文献[12] Qu 和 Sun 学习了对于不定核核 l^q 范数正则化学习算法的渐进性能, 并且得到了满意的学习率; 通过奇异值分解的方法, 文献[13] Q 考虑不定核的情况下, 得出了与容量无关的误差界。

在这篇文章里, 记 $k = \sup_{x,t \in X} |K(x, t)| < \infty$, 我们学习 l^2 系数正则化算法:

$$f_{Z,\lambda} = f_{\alpha_Z} \text{ 其中 } \alpha_Z = \arg \min_{\alpha \in R^m} \left\{ \frac{1}{m} \sum_{i=1}^m (f_{\alpha}(x_i) - y_i)^2 + \lambda m \sum_{i=1}^m \alpha_i^2 \right\}, \lambda > 0 \quad (1.3)$$

这里 $f_{\alpha} = \sum_{i=1}^m \alpha_i \tilde{K}(\cdot, x_i)$ 。由(1.2), 有 $k^2 = \sup_{u,v} |K(u, v)|$ 。

本文我们研究 non-iid 抽样下的系数正则化学习算法的性能。考虑基于 Z 上的概率分布序列 $\{\rho^i\}$, 因此对于任意的 $i \geq 1$, ρ^i 在 x 处的条件分布为 $\rho(\cdot|x)$ 。第 i 个样本 $z_i = (x_i, y_i)$ 的选取依赖于 $\rho^{(i)}$ 。我们假设样本序列 $z_i = (x_i, y_i), i \geq 1$ 来自于强平稳过程并且样本的相关性满足强混合条件。

关于强混合条件的定义:

定义 1.4 对于两个 σ 域 Γ 和 D , 定义 α -数为 $\alpha(\Gamma, D) = \sup_{A \in \Gamma, B \in D} |P(A \cap B) - P(A)P(B)|$ 。给定一系列样

本 $\{z_i\}_{i=1}^m$, 定义 M_a^b 为由随机变量 $z_a, z_{a+1}, \dots, z_b$ 生成的 σ 域。当 $i \rightarrow \infty$ 时, $\alpha_i = \sup_{k \geq 1} \alpha(M_1^k, M_{k+i}^{\infty}) \rightarrow 0$, 称随机过程 $z_i, i \geq 1$ 满足强混合条件(或 α 混合条件)。

对于正则化学习算法的误差分析, 已经有很多研究工作, 其中绝大多数假设输出数据是一致有界的, 即存在某一常数 $M > 0$, 使得几乎处处有 $|y| \leq M$, 也即条件分布 $\rho(\cdot|x)$ 的支撑集几乎处处包含在 $[-M, M]$ 中, 见文献[3] [4] [6] [10]。文献[14] [15]在输出数据满足以下矩假设条件下得出算法(1.3)的误差

界:

定义 1.5 (矩有界假设) 存在常数 $\hat{M} > 0$ 及 $C_l > 0$ 使得

$$\int_Y |y|^l d\rho(y/x) \leq C_l l! \hat{M}_l, \forall l \in N, x \in X \quad (1.4)$$

文献[1]和[2], 在以下弱化的矩有界假设下分别研究了回归学习算法(1.1)和(1.3)。

本文给出的有关输出数据的假设:

定义 1.6 (弱化的矩有界假设) 存在两个常数 $M > 0$ 及 $p \geq 2$ 使得

$$\int_Z |y|^p d\rho \leq M \quad (1.5)$$

注: 矩假设(1.4)要求输出数据 y 的所有矩皆有界, 而本文对矩假设进一步弱化, 只要求 y 的某个 $p(p \geq 2)$ 阶矩有界。

本文研究对输出变量满足弱化的矩有界条件(1.5)下基于非正定核的惩罚项为 l^2 -范数的系数正则化算法的误差分析。下一节, 我们将给出一些假设条件和本文主要结论。

2. 假设与结论

设 X 是紧的距离空间, $\tilde{K}$ 是 Mercer 核, ρ_X 为定义在 X 上非退化的 Borel 测度, 对任意的 $f \in L^2_\nu(t)$, 定义积分算子为:

$$L_{\tilde{K}, \nu} f(x) = \int_X \tilde{K}(x, t) f(t) d\rho_X(t), x \in X \quad (2.1)$$

假设 1 (Non-iid Sampling Condition): 存在一些正整数 N_0 , 在 X 上有两个正的有界的 Borel 测度, 使得

$$\tilde{\mu} \leq \mu \leq \hat{\mu}, \forall m \geq N_0. \quad (2.2)$$

其中 $\mu \equiv \frac{1}{m} \sum_{i=1}^m \rho_X^{(i)}$ 。

假设 2 (Approximation Condition): 存在 $a_0 > 0, 0 < \beta \leq 1$, 使得

$$D(\lambda) = \min_{f \in H_{\tilde{K}}} \left\{ \|f - f_\rho\|_{L^2_{\tilde{\mu}}}^2 + \lambda \|f\|_{\tilde{K}}^2 \right\} \leq a_0^2 \lambda^\beta, \forall \lambda > 0 \quad (2.3)$$

通过求导运算, 有

$$D(\lambda) = \lambda \left\| \left(\lambda I + L_{\tilde{K}, \hat{\mu}} \right)^{-\frac{1}{2}} f_\rho \right\|_{L^2_{\tilde{K}, \hat{\mu}}}^2 \quad (2.4)$$

即, $f_\rho \in \text{Range} \left(L_{\tilde{K}, \hat{\mu}}^{\frac{\beta}{2}} \right)$ [16]。

定理 2.1 假设 Non-iid Sampling Condition (2.2)和 Approximation Condition (2.3)成立, 样本序列 $z_i = (x_i, y_i), 1 \leq i \leq m$ 满足强混合条件, 并且 $0 < \lambda \leq 1$ 。则对任何 $0 < \eta < 1$, $1 - \eta$ 的概率下有

$$\begin{aligned} \|f_{Z, \lambda} - f_\rho\|_{\rho_X} &\leq \frac{2a_0 \lambda^{\beta/4}}{\eta} + \frac{C}{\eta} m^{-1/2} \lambda^{-1/4} \left[1 + m^{-1/4} \lambda^{-1/4} \left(1 + \sum_{i=1}^m \alpha_i \right)^{1/4} \right] \\ &\quad \times \left[1 + \lambda^{(\beta-1)(p-2)/4p} \left(\sum_{i=1}^m \alpha_i^{(p-2)/2} \right)^{1/2} + \lambda^{(\beta-1)/4} \left(1 + \sum_{i=1}^m \alpha_i \right)^{1/2} \right] \end{aligned}$$

成立。

其中 C 是与 k, M, p, a_0 有关的常数。

基于定理 2.1, 通过选择合适的 λ (与 m 相关) 可以得到算法(1.3)的收敛速率。即:

定理 2.2 在定理 2.1 的假设下, 且对于一些 $a > 0$ 和 $t > 0$, 满足 $\alpha_t \leq al^{-t}$, 则有

$$\|f_{Z,\lambda} - f_\rho\|_{L_\mu^2} \leq \tilde{C} \left(m^{-\theta(\beta,t)\min\left\{\frac{p-2}{p}, 1\right\}} (\log_m)^{3/4} \right)$$

$$\text{这里 } \theta(\beta,t) \text{ 定义为: } \theta(\beta,t) = \begin{cases} \frac{p\beta}{4(\beta+p-1)}, & 0 < \beta < 1, 0 < t < \frac{p}{p-2} \\ \frac{\beta}{4}, & 0 < \beta < 1, t \geq \frac{p}{p-2} \\ \frac{1}{4}, & \beta = 1 \end{cases}$$

定理 2.1, 定理 2.2 的证明会在第 4 部分-误差分析给出。

下一节我们会给出一些基本引理。□

3. 基本引理

证明定理 2.1 还需要一些结论。首先我们用以下引理[3]来处理 α -混合条件。定义向量值随机变量 ξ 的 u 阶矩为: $\|\xi\|_\mu = \left(E\|\xi\|_H^u\right)^{1/u}$ ($1 < u < \infty$), $\|\xi\|_\infty = \sup\|\xi\|_H$ 。

引理 3.1 假设随机样本 $z_i = (x_i, y_i), i \geq 1$ 满足 α -混合条件。设 ξ, η 为取值在可分 Hilbert 空间 H 中, 分别关于 σ 代数 J 与 D 可测的随机变量。对任意的 $1 < p, q, t < \infty$, 满足 $p^{-1} + q^{-1} + t^{-1} = 1$ 或者 $p = q = \infty, t = 1$, 则有

$$|E(\xi, \eta) - (E\xi, E\eta)| < 15\alpha^{\frac{1}{t}}(J, D)\|\xi\|_p\|\eta\|_q \quad (3.1)$$

讨论假设空间的逼近能力, 需要以下函数

$$g_{\mu,\lambda} = \arg \min_{f \in H_{\tilde{K}}} \left\{ \|f - f_\rho\|_{L_\mu^2}^2 + \lambda \|f\|_{\tilde{K}}^2 \right\} \quad (3.2)$$

通过对(3.2)的求导运算, 可以得到

$$g_{\mu,\lambda} = L_{\tilde{K},\mu} \left(L_{\tilde{K},\mu} + \lambda I \right)^{-1} f_\rho \quad (3.3)$$

同样, 也需要以下引理。

引理 3.2 对 $\forall f \in H_{\tilde{K}}$, $L_{\tilde{K},\mu}^{-\frac{1}{2}}$ 为从 L_μ^2 到 $H_{\tilde{K}}$ 的等距同构映射, 则有

$$\|f\|_{L_\mu^2} = \left\| L_{\tilde{K},\mu}^{-\frac{1}{2}} f \right\|_{\tilde{K}} \quad (3.4)$$

引理 3.3 在假设 1 (Non-iid Sampling Condition) (2.2)和假设 2 (Approximation Condition) (2.3)成立的条件下, 则有

$$\left\| \left(\lambda I + L_{K,\mu} \right)^{-\frac{1}{2}} f_\rho \right\|_{L_\mu^2} \leq a_0 \lambda^{(\beta-1)/2} \quad (3.5)$$

$$\begin{aligned}
D(\lambda) &\geq \min_{f \in H_{\tilde{K}}} \left\{ \|f - f_\rho\|_{L^2_\mu}^2 + \lambda \|f\|_{\tilde{K}}^2 \right\} \\
&= \|g_{\mu, \lambda} - f_\rho\|_{L^2_\mu}^2 + \lambda \|g_{\mu, \lambda}\|_{\tilde{K}}^2 \\
\text{证: } &= \left\| \lambda (L_{\tilde{K}, \mu} + \lambda I)^{-1} f_\rho \right\|_{L^2_\mu}^2 + \lambda \left\| L_{\tilde{K}, \mu}^{-\frac{1}{2}} (L_{\tilde{K}, \mu} + \lambda I)^{-1} f_\rho \right\|_{L^2_\mu}^2 \\
&= \lambda \left\| (\lambda I + L_{\tilde{K}, \mu})^{-\frac{1}{2}} f_\rho \right\|_{L^2_\mu}^2
\end{aligned}$$

在给出下一个引理之前, 这里先给出一些概念: 定义样本算子 $S_x: H_K \rightarrow R^m$ 为

$$S_x f = \{f(x_i)\}_{i=1}^m, \text{ 对任意的 } f \in H_K,$$

其中 $x = \{x_i\}_{i=1}^m \subseteq X$ 。

那么, S_x 的共轭算子 $S_x^*: R^m \rightarrow H_{\tilde{K}}$ 为

$$S_x^* c = \sum_{i=1}^m c_i \tilde{K}(\cdot, x_i)$$

这里, $c = (c_1, \dots, c_m) \in R^m$ 。

在文献[10][13]里已经证明有

$$\begin{aligned}
f_{Z, \lambda} &= \left(\lambda I + \left(\frac{1}{m} S_x^* S_x \right)^2 \right)^{-1} \frac{1}{m^2} S_x^* S_x S_x^* y \\
&= \left(\lambda I + \left(\frac{1}{m} S_x^* S_x \right)^2 \right)^{-1} \left(\frac{1}{m} S_x^* S_x \right) \frac{1}{m} S_x^* y
\end{aligned} \tag{3.6}$$

接下来给出 Hilbert Schmidt [7] 的定义和性质:

令 $HS(H_{\tilde{K}})$ 是 $H_{\tilde{K}}$ 上所有的 Hilbert Schmidt 类, 它构成了一个 Hilbert 空间, 其内积为

$$\langle T, S \rangle_{HS} := \sum_{i=1}^{\infty} \langle T \varphi_i, S \varphi_i \rangle_{\tilde{K}}$$

其中 $\{\varphi_i\}_{i=1}^{+\infty}$ 是 H_K 的一组正交基并且这个定义不依赖于基的选择。

我们需要 Hilbert Schmidt 算子的下列性质。

1) 对 $\forall h \in H_{\tilde{K}}$, 定义一秩算子 $h \otimes h(f) = \langle h, f \rangle_{\tilde{K}}$, 并且它是一个 Hilbert Schmidt 算子满足 $\|h \otimes h\|_{HS} = \|h\|_{\tilde{K}}^2$ 。记

$$E \frac{1}{m} S_x^* S_x = E \frac{1}{m} \sum_{i=1}^m \tilde{K}_{x_i} \otimes \tilde{K}_{x_i} = \frac{1}{m} \sum_{i=1}^m L_{\tilde{K}, \rho_X^{(i)}} = L_{\tilde{K}, \mu} \tag{3.7}$$

2) 对任意 Hilbert Schmidt 算子 T , 有 $\|T\| \leq \|T\|_{HS}$ 。

接下来的引理给出了经验积分算子的收敛性:

引理 3.4 假设随机变量序列 $z_i = (x_i, y_i), 1 \leq i \leq m$ 满足强混合条件, 则下列不等式成立

$$\left\| \frac{1}{m} S_x^* S_x - L_{\tilde{K}, \mu} \right\| \leq \frac{k^2}{\sqrt{m}} \left(1 + 30 \sum_{l=1}^{m-1} \alpha_l \right)^{\frac{1}{2}} \tag{3.8}$$

证: 由(3.7)式和 $L_{\tilde{K}, \mu}$ 是一 Hilbert Schmidt 算子, 有

$$\begin{aligned}
E \left\| \frac{1}{m} S_x^* S_x - L_{\tilde{K}, \mu} \right\|^2 &\leq E \left\| \frac{1}{m} \sum_{i=1}^m \tilde{K}(\cdot, x_i) \otimes \tilde{K}(\cdot, x_i) - L_{\tilde{K}, \mu} \right\|_{HS}^2 \\
&= E \left\| \frac{1}{m} \sum_{i=1}^m \tilde{K}_{x_i} \otimes \tilde{K}_{x_i} \right\|_{HS}^2 - \|L_{\tilde{K}, \mu}\|_{HS}^2 \\
&= E \frac{1}{m^2} \sum_{i=1}^m \|\tilde{K}_{x_i} \otimes \tilde{K}_{x_i}\|_{HS}^2 + \frac{1}{m^2} \sum_{i \neq j} E \langle \tilde{K}_{x_i} \otimes \tilde{K}_{x_i}, \tilde{K}_{x_j} \otimes \tilde{K}_{x_j} \rangle_{HS} - \|L_{\tilde{K}, \mu}\|_{HS}^2
\end{aligned}$$

记 $\|\tilde{K}_{x_i} \otimes \tilde{K}_{x_i}\|_{HS}^2 \leq k^4$ [11], 由引理 3.1, 当 $p = q = \infty$ 且 $t = 1$ 时有, 当 $i \neq j$,

$$\begin{aligned}
E \langle \tilde{K}_{x_i} \otimes \tilde{K}_{x_i}, \tilde{K}_{x_j} \otimes \tilde{K}_{x_j} \rangle_{HS} &\leq \langle E \tilde{K}_{x_i} \otimes \tilde{K}_{x_i}, E \tilde{K}_{x_j} \otimes \tilde{K}_{x_j} \rangle_{HS} + 15 \alpha_{|i-j|} \|\tilde{K}_{x_i} \otimes \tilde{K}_{x_i}\|_{\infty} \|\tilde{K}_{x_j} \otimes \tilde{K}_{x_j}\|_{\infty} \\
&\leq \|L_{\tilde{K}, \mu}\|_{HS}^2 + 15 k^4 \alpha_{|i-j|}
\end{aligned}$$

然后有

$$\begin{aligned}
E \left\| \frac{1}{m} S_x^* S_x - L_{\tilde{K}, \mu} \right\|^2 &\leq \frac{k^4}{m} + \frac{m^2 - m}{m^2} \|L_{\tilde{K}, \mu}\|_{HS}^2 + \frac{30 k^4}{m} \sum_{l=1}^{m-1} \alpha_l \|L_{\tilde{K}, \mu}\|_{HS}^2 \\
&\leq \frac{k^4}{m} \left(1 + 30 \sum_{l=1}^{m-1} \alpha_l \right)
\end{aligned}$$

4. 误差分析

为了分析算法误差, 像文献[2]中一样, 这里需要引进一个正则化函数 $f_{\mu, \lambda}$:

$$f_{\mu, \lambda} = L_{\tilde{K}, \mu}^2 (\lambda I + L_{\tilde{K}, \mu}^2)^{-1} f_{\rho}, 0 < \lambda < 1 \quad (4.1)$$

所以我们可以把误差 $f_{Z, \lambda} - f_{\rho}$ 分解成两部分:

$$f_{Z, \lambda} - f_{\rho} = \{f_{Z, \lambda} - f_{\mu, \lambda}\} + \{f_{\mu, \lambda} - f_{\rho}\} \quad (4.2)$$

这里右边第一部分叫做样本误差, 第二部分叫做正则误差。

4.1. 正则误差分析

在这一小节中, 我们分析正则化误差 $\|f_{\mu, \lambda} - f_{\rho}\|_{L_{\mu}^2}$ 的界。

命题 4.1 在假设 1 (Non-iid Sampling Condition) (2.2) 和假设 2 (Approximation Condition) (2.3) 成立的条件下, 当 $0 < \lambda \leq 1$ 时, 下列正则化误差的界成立:

$$\|f_{\mu, \lambda} - f_{\rho}\|_{L_{\mu}^2} \leq 2 a_0 \lambda^{\beta/4} \quad (4.3)$$

证: 由引理 3.2, 引理 3.3 和下列积分算子不等式

$$\left\| \left(\lambda + L_{\tilde{K}, \mu}^2 \right)^{-\frac{1}{4}} \left(\sqrt{\lambda} + L_{\tilde{K}, \mu} \right)^{\frac{1}{2}} \right\| \leq 2$$

则有

$$\|f_{\mu, \lambda} - f_{\rho}\|_{L_{\mu}^2} = \left\| \lambda \left(\lambda I + L_{\tilde{K}, \mu}^2 \right)^{-1} \left(\sqrt{\lambda} + L_{\tilde{K}, \mu} \right)^{\frac{1}{2}} \left(\sqrt{\lambda} + L_{\tilde{K}, \mu} \right)^{\frac{1}{2}} f_{\rho} \right\|_{L_{\mu}^2} \leq 2 a_0 \lambda^{\beta/4}$$

相似地, 有

$$\|f_{\mu,\lambda}\|_{\tilde{K}} = \left\| L_{\tilde{K},\mu}^{\frac{3}{2}} (\lambda I + L_{\tilde{K},\mu}^2)^{-1} (\sqrt{\lambda} + L_{\tilde{K},\mu})^{\frac{1}{2}} (\sqrt{\lambda} + L_{\tilde{K},\mu})^{-\frac{1}{2}} f_{\rho} \right\|_{L_{\mu}^2} \leq 2a_0 \lambda^{(\beta-1)/4} \quad (4.4)$$

4.2. 样本误差分析

在这一个小分节中, 我们分析样本误差 $\|f_{Z,\lambda} - f_{\mu,\lambda}\|_{L_{\mu}^2}$ 的界。

与文献[10]同样的方法, 记 $\lambda f_{\mu,\lambda} = L_{\tilde{K},\mu}^2 (f_{\rho} - f_{\mu,\lambda})$, 有

$$\begin{aligned} f_{Z,\lambda} - f_{\mu,\lambda} &= \left[\lambda I + \left(\frac{1}{m} S_x^* S_x \right)^2 \right]^{-1} \left\{ \left(\frac{1}{m} S_x^* S_x \right) \frac{1}{m} S_x^* y - \left[\left(\frac{1}{m} S_x^* S_x \right)^2 + \lambda I \right] f_{\mu,\lambda} \right\} \\ &= \left[\lambda I + \left(\frac{1}{m} S_x^* S_x \right)^2 \right]^{-1} \left\{ \left(\frac{1}{m} S_x^* S_x \right) \frac{1}{m} \sum_{i=1}^m (y_i - f_{\mu,\lambda}(x_i)) \tilde{K}_{x_i} - L_{\tilde{K},\mu}^2 (f_{\rho} - f_{\mu,\lambda}) \right\} \\ &= \left[\lambda I + \left(\frac{1}{m} S_x^* S_x \right)^2 \right]^{-1} \left(\frac{1}{m} S_x^* S_x \right) U + \left[\lambda I + \left(\frac{1}{m} S_x^* S_x \right)^2 \right]^{-1} VW \end{aligned} \quad (4.5)$$

其中

$$\begin{aligned} U &= \frac{1}{m} \sum_{i=1}^m (y_i - f_{\mu,\lambda}(x_i)) \tilde{K}_{x_i} - L_{\tilde{K},\mu}^2 (f_{\rho} - f_{\mu,\lambda}) \\ V &= \frac{1}{m} S_x^* S_x - L_{\tilde{K},\mu} \\ W &= L_{\tilde{K},\mu} (f_{\rho} - f_{\mu,\lambda}) \end{aligned} \quad (4.6)$$

记 $f_{Z,\lambda}$ 和 $f_{\mu,\lambda}$ 都在 $H_{\tilde{K}}$ 中。由引理 3.2, 有

$$\begin{aligned} \|f_{Z,\lambda} - f_{\mu,\lambda}\|_{L_{\mu}^2} &= \left\| L_{\tilde{K},\mu}^{\frac{1}{2}} (f_{Z,\lambda} - f_{\mu,\lambda}) \right\|_{\tilde{K}} \\ &\leq \left\| L_{\tilde{K},\mu}^{\frac{1}{2}} \left[\lambda I + \left(\frac{1}{m} S_x^* S_x \right)^2 \right]^{-1} \left(\frac{1}{m} S_x^* S_x \right) U \right\|_{\tilde{K}} + \left\| L_{\tilde{K},\mu}^{\frac{1}{2}} \left[\lambda I + \left(\frac{1}{m} S_x^* S_x \right)^2 \right]^{-1} VW \right\|_{\tilde{K}} \\ &= I_1 + I_2 \end{aligned} \quad (4.7)$$

接下来分别分析 I_1 和 I_2 。

由算子单调不等式([17], 定理 2.1), 有

$$\left\| L_{\tilde{K},\mu}^{\frac{1}{2}} - \left(\frac{1}{m} S_x^* S_x \right)^{\frac{1}{2}} \right\| \leq \left\| L_{\tilde{K},\mu} - \left(\frac{1}{m} S_x^* S_x \right) \right\|^{\frac{1}{2}} \quad (4.8)$$

于是, 可以得到

$$\begin{aligned} I_1 &\leq \left\| L_{\tilde{K},\mu} - \left(\frac{1}{m} S_x^* S_x \right) \right\|^{\frac{1}{2}} \left\| \left[\lambda I + \left(\frac{1}{m} S_x^* S_x \right)^2 \right]^{-1} \left(\frac{1}{m} S_x^* S_x \right) \right\| \|U\|_{\tilde{K}} + \left\| \left[\lambda I + \left(\frac{1}{m} S_x^* S_x \right)^2 \right]^{-1} \right\| \|U\|_{\tilde{K}} \\ &\leq \left(\lambda^{-\frac{1}{2}} \|V\|^{\frac{1}{2}} + \lambda^{-\frac{1}{4}} \right) \|U\|_{\tilde{K}} \end{aligned} \quad (4.9)$$

类似地得到 I_2 的界

$$\begin{aligned}
I_2 &\leq \left\| L_{\bar{K}, \mu} - \left(\frac{1}{m} S_x^* S_x \right) \right\|_{\bar{K}}^{\frac{1}{2}} \left\| \left[\lambda I + \left(\frac{1}{m} S_x^* S_x \right)^2 \right]^{-1} \right\|_{\bar{K}} \left\| V \right\|_{\bar{K}} \left\| W \right\|_{\bar{K}} + \left\| \left[\lambda I + \left(\frac{1}{m} S_x^* S_x \right)^2 \right]^{-1} \left(\frac{1}{m} S_x^* S_x \right) \right\|_{\bar{K}}^{\frac{1}{2}} \left\| V \right\|_{\bar{K}} \left\| W \right\|_{\bar{K}} \\
&\leq \left(\lambda^{-1} \left\| V \right\|_{\bar{K}}^2 + \lambda^{-\frac{3}{4}} \right) \left\| V \right\|_{\bar{K}} \left\| W \right\|_{\bar{K}}
\end{aligned} \quad (4.10)$$

下面我们会依次给出 $\|V\|, \|W\|_{\bar{K}}, \|U\|_{\bar{K}}$ 的估计。引理 3.4 已经给出了 $\|V\|$ 的估计。接下来考虑 $\|W\|_{\bar{K}}$ 类似命题 4.1, 得到以下引理:

引理 4.1 在假设 1 (Non-iid Sampling Condition) 和假设 2 (Approximation Condition) 成立的条件下, 当 $0 < \lambda \leq 1$ 时, 有以下不等式成立

$$\|W\|_{\bar{K}} \leq 2a_0 \lambda^{(\beta+1)/4} \quad (4.11)$$

最后分析 $\|U\|_{\bar{K}}$, 需要有下列引理:

引理 4.2 在假设 1 (Non-iid Sampling Condition) (2.2) 和假设 2 (Approximation Condition) (2.3) 成立的条件下, 有

$$\left(E \|U\|_{\bar{K}}^2 \right)^{\frac{1}{2}} \leq \left(\frac{C_3}{m} \right)^{\frac{1}{2}} \left\{ 1 + \lambda^{(\beta-1)(p-1)/4p} \left(\sum_{i=1}^{m-1} \alpha_i^{(p-2)/p} \right)^{\frac{1}{2}} \right\} \quad (4.12)$$

证: 定义在 $H_{\bar{K}}$ 上的向量值随机变量 $\xi(z) = (y - f_{\mu, \lambda}(x)) \tilde{K}(\cdot, x)$ 。因此,

$$U = \frac{1}{m} \sum_{i=1}^m \xi(z_i) - L_{\bar{K}, \mu}(f_\rho - f_{\mu, \lambda}), E \frac{1}{m} \sum_{i=1}^m \xi(z_i) = L_{\bar{K}, \mu}(f_\rho - f_{\mu, \lambda})$$

所以,

$$\begin{aligned}
E \|U\|_{\bar{K}}^2 &= E \left\| \frac{1}{m} \sum_{i=1}^m \xi(z_i) - E \frac{1}{m} \sum_{i=1}^m \xi(z_i) \right\|_{\bar{K}}^2 \\
&= E \left\| \frac{1}{m} \sum_{i=1}^m \xi(z_i) \right\|_{\bar{K}}^2 - \left\| L_{\bar{K}, \mu}(f_\rho - f_{\mu, \lambda}) \right\|_{\bar{K}}^2
\end{aligned} \quad (4.13)$$

由引理 3.1, 当 $p > 2$ 时, 令 $u = v = p$, $t = \frac{p}{p-2}$, 当 $j < i$, 有

$$\begin{aligned}
E \langle \xi(z_i), \xi(z_j) \rangle_{\bar{K}} &\leq \langle E \xi(z_i), E \xi(z_j) \rangle_{\bar{K}} + 15 \alpha^{(p-2)/p} (M_1^j, M_1^i) \|\xi(z_i)\|_p \\
&= \left\| L_{\bar{K}, \mu}(f_\rho - f_{\mu, \lambda}) \right\|_{\bar{K}}^2 + 15 (\alpha_{i-j})^{(p-2)/p} \|\xi\|_p^2
\end{aligned} \quad (4.14)$$

所以有

$$\begin{aligned}
E \left\| \frac{1}{m} \sum_{i=1}^m \xi(z_i) \right\|_{\bar{K}}^2 &= \frac{1}{m^2} \sum_{i=1}^m E (y_i - f_{\mu, \lambda}(x_i))^2 \tilde{K}(x_i, x_i) + \frac{1}{m^2} \sum_{i \neq j} E \langle \xi(z_i), \xi(z_j) \rangle_{\bar{K}} \\
&\leq \frac{k^2}{m^2} \sum_{i=1}^m E (y_i - f_{\mu, \lambda}(x_i))^2 + \frac{m^2 - m}{m^2} \left\| L_{\bar{K}, \mu}(f_\rho - f_{\mu, \lambda}) \right\|_{\bar{K}}^2 + \frac{30}{m} \sum_{i=1}^{m-1} \alpha_i^{(p-2)/p} \|\xi\|_p^2
\end{aligned} \quad (4.15)$$

$$\text{故有 } E \left\| \frac{1}{m} \sum_{i=1}^m \xi(z_i) - E \xi \right\|_{\bar{K}}^2 \leq \frac{k^2}{m^2} \sum_{i=1}^m E (y_i - f_{\mu, \lambda}(x_i))^2 + \frac{30}{m} \sum_{i=1}^{m-1} \alpha_i^{(p-2)/p} \|\xi\|_p^2 \quad (4.16)$$

接下来分别估计 $\frac{k^2}{m^2} \sum_{i=1}^m E (y_i - f_{\mu, \lambda}(x_i))^2$ 和 $\|\xi\|_p$ 。对于 $\frac{k^2}{m^2} \sum_{i=1}^m E (y_i - f_{\mu, \lambda}(x_i))^2$, 有

$$\begin{aligned} \frac{k^2}{m^2} \sum_{i=1}^m E(y_i - f_{\mu,\lambda}(x_i))^2 &\leq \frac{k^2}{m^2} \sum_{i=1}^m E(y_i - f_\rho)^2 + \frac{k^2}{m^2} \sum_{i=1}^m E(f_\rho - f_{\mu,\lambda}(x_i))^2 \\ &= \frac{k^2}{m} (E|y|^2 - E|f_\rho|^2) + \frac{k^2}{m} \|f_\rho - f_{\mu,\lambda}\|_{L_\mu^2}^2 \end{aligned} \quad (4.17)$$

由 Hölder 不等式, 得到

$$\int_Z y^2 d\rho \leq \left(\int_Z |y|^p d\rho \right)^{2/p} \leq M^{2/p} \quad (4.18)$$

则

$$\|f_\rho\|_{\rho_X} \leq \int_X \left(\int_Y y d(y/x) \right)^2 d\rho_X \leq \int_Z y^2 d\rho \leq M^{2/p} \quad (4.19)$$

因此

$$Ey^2 - \|f_\rho\|_{\rho_X}^2 < M^{2/p} \quad (4.20)$$

将(4.3)和(4.20)代入(4.17), 得到

$$\frac{k^2}{m^2} \sum_{i=1}^m E(y_i - f_{\mu,\lambda}(x_i))^2 \leq \frac{k^2}{m} (M^{2/p} + 4a_0^2 \lambda^{\beta/2}) \quad (4.21)$$

接下来评估 $\|\xi\|_p^2$,

$$\begin{aligned} \|\xi\|_p^2 &= \left[E \left((y - f_{\mu,\lambda})^2 \tilde{K}(x, x) \right)^{p/2} \right]^{2/p} \\ &\leq k^2 \left(E |y - f_{\mu,\lambda}(x)|^p \right)^{2/p} \\ &\leq k^2 \left[\left(E |y|^p \right)^{1/p} + \left(E |f_{\mu,\lambda}(x)|^p \right)^{1/p} \right]^2 \\ &\leq 2k^2 \left[\left(E |y|^p \right)^{2/p} + \left(E |f_{\mu,\lambda}(x)|^p \right)^{2/p} \right] \end{aligned} \quad (4.22)$$

现在, 我们需要给出 $\left(E |f_{\mu,\lambda}(x)|^p \right)^{2/p}$ 的上界。

由(4.4), 有

$$|f_{\mu,\lambda}(x)| \leq k \|f_{\mu,\lambda}\|_{\tilde{K}} \leq 2ka_0 \lambda^{(\beta-1)/4} \quad (4.23)$$

由 $f_{\mu,\lambda}$ 的定义, 有

$$\|f_{\mu,\lambda}\|_{\rho_X}^2 \leq \|f_\rho\|_{\rho_X}^2 \leq Ey^2 \leq M^{2/p} \quad (4.24)$$

所以

$$\begin{aligned} \left(E |f_{\mu,\lambda}(x)|^p \right)^{2/p} &= E \left(|f_{\mu,\lambda}(x)|^2 |f_{\mu,\lambda}|^{p-2} \right)^{2/p} \\ &\leq \left(\|f_{\mu,\lambda}\|_{\rho_X}^2 \right)^{2/p} \left(|f_{\mu,\lambda}(x)|^{p-2} \right)^{2/p} \\ &\leq \left(M^{2/p} \right)^{2/p} \left(\left(2ka_0 \lambda^{(\beta-1)/4} \right)^{p-2} \right)^{2/p} \\ &= 4^{(p-2)/p} M^{4/p^2} k^{2(p-2)/p} a_0^{2(p-2)/p} \lambda^{(\beta-1)(p-2)/2p} \end{aligned} \quad (4.25)$$

将(4.18)和(4.25)代入(4.22), 有

$$\|\xi\|_p^2 \leq 2k^2 \left(M^{2/p} + 4^{(p-2)/p} M^{4/p} k^{2(p-2)/p} a_0^{2(p-2)/p} \lambda^{(\beta-1)(p-2)/2p} \right) \quad (4.26)$$

将 $\frac{k^2}{m^2} \sum_{i=1}^m E(y_i - f_{\mu,\lambda}(x_i))^2$ 和 $\|\xi\|_p^2$ 的上界与(4.16)联系, 有

$$\left(E \|U\|_{\bar{K}}^2 \right) \leq \frac{k^2}{m} \left(M^{2/p} + 4a_0 \lambda^{\beta/2} \right) + \frac{30C_1}{m} \sum_{l=1}^{m-1} \alpha_l^{(p-2)/p} \times \lambda^{(\beta-1)(p-2)/2p} \quad (4.27)$$

故有

$$\left(E \|U\|_{\bar{K}}^2 \right)^{\frac{1}{2}} \leq \left(\frac{C_2}{m} \right)^{\frac{1}{2}} \left\{ 1 + \lambda^{(\beta-1)(p-2)/4p} \left(\sum_{l=1}^{m-1} \alpha_l^{(p-2)/p} \right)^{\frac{1}{2}} \right\} \quad (4.28)$$

其中 C_1, C_2 是与 k, M, p, a_0 相关的常数。随即我们可以有以下命题:

命题 4.2 在假设 1 (Non-iid Sampling Condition) (2.2) 和假设 2 (Approximation Condition) (2.3) 成立的条件下, 当 $0 < \lambda \leq 1$ 时, 有以下不等式成立:

$$\begin{aligned} \|f_{Z,\lambda} - f_{\mu,\lambda}\|_{L_\mu^2} &\leq C m^{-1/2} \lambda^{-1/4} \left[1 + m^{-1/4} \lambda^{-1/4} \left(1 + \sum_{l=1}^m \alpha_l \right)^{1/4} \right] \\ &\quad \times \left[1 + \lambda^{(\beta-1)(p-2)/4p} \left(\sum_{l=1}^m \alpha_l^{(p-2)/p} \right)^{1/2} + \lambda^{(\beta-1)/4} \left(1 + \sum_{l=1}^m \alpha_l \right)^{1/2} \right] \end{aligned}$$

其中 C 是与 k, M, p, a_0 相关的常数。

证:

$$\begin{aligned} \|f_{Z,\lambda} - f_{\mu,\lambda}\|_{L_\mu^2}^2 &\leq \left[\lambda^{-1/4} \|U\|_{\bar{K}} + 2a_0 k^2 \lambda^{(\beta-2)/4} m^{-1/2} \left(1 + \sum_{l=1}^{m-1} \alpha_l \right)^{1/2} \right] \\ &\quad \times \left[k \lambda^{-1/4} m^{-1/4} \left(1 + \sum_{l=1}^{m-1} \alpha_l \right)^{1/4} + 1 \right] \\ &= C_3 m^{-1/2} \lambda^{-1/4} \left[1 + m^{-1/4} \lambda^{-1/4} \left(1 + \sum_{l=1}^{m-1} \alpha_l \right)^{1/4} \right] \\ &\quad \times \left[1 + \lambda^{(\beta-1)(p-2)/4p} \times \left(\sum_{l=1}^{m-1} \alpha_l^{(p-2)/p} \right)^{1/2} + 2a_0 \lambda^{(\beta-1)/4} \left(1 + \sum_{l=1}^{m-1} \alpha_l \right)^{1/2} \right] \\ &= C m^{-1/2} \lambda^{-1/4} \left[1 + m^{-1/4} \lambda^{-1/4} \left(1 + \sum_{l=1}^m \alpha_l \right)^{1/4} \right] \\ &\quad \times \left[1 + \lambda^{(\beta-1)(p-2)/4p} \left(\sum_{l=1}^m \alpha_l^{(p-2)/2} \right)^{1/2} + \lambda^{(\beta-1)/4} \left(1 + \sum_{l=1}^m \alpha_l \right)^{1/2} \right] \end{aligned} \quad (4.29)$$

由马尔科夫不等式和三角不等式, 有

$$P\{\|f_{Z,\lambda} - f_\rho\| > \eta\} \leq \frac{1}{\eta} \left\{ E \|f_{Z,\lambda} - f_{\mu,\lambda}\|_{L_\mu^2} + E \|f_{\mu,\lambda} - f_\rho\|_{L_\mu^2} \right\} \quad (4.30)$$

将(4.3)和(4.29)代入(4.30)中, 可以得到想要的界, 即完成定理 2.1 的证明。

最后证明定理 2.2。为了得到学习率, 这里需要对 $\alpha_i \rightarrow 0$ 收敛速度做一些规定。

定义 4.1 随机样本序列 $z_i = (x_i, y_i)$, 满足 $i \geq 1$ 足多项式强混合条件, 如果对一 $a > 0$, $t > 0$ 。 α -混合系数 α_i 满足,

$$\alpha_i \leq al^{-t}, l \geq 1 \quad (4.31)$$

通过以下式子可以很容易地得到学习率

$$\sum_{l=1}^{m-1} l^{-t} = \begin{cases} o(m^{1-t}), & t < 1 \\ o(\log m), & t = 1 \\ o(1), & t > 1 \end{cases}$$

定理 2.2 的证明。当 $0 < t < 1$, 由(4.31), 有

$$1 + \sum_{l=1}^{m-1} \alpha_l \leq 1 + a \sum_{l=1}^{m-1} l^{-t} \leq (1+a)m^{1-t}$$

$$\sum_{l=1}^{m-1} \alpha_l^{(p-2)/p} \leq a^{(p-2)/p} m^{1-(p-2)t/p}$$

当 $0 < \beta < 1$, 由定理 2.1, 有

$$\|f_{Z,\lambda} - f_\rho\|_{L_\mu^2} \leq C \left\{ \lambda^{\beta/4} + m^{-1/2} \lambda^{-1/4} \left(1 + m^{-1/4} \lambda^{-1/4} m^{(1-t)/4} \right) \right. \\ \left. \times \left(1 + \lambda^{(\beta-1)(p-2)/4p} m^{1/2-(p-2)t/2p} + \lambda^{(\beta-1)/4} m^{(1-t)/2} \right) \right\}$$

为了平衡正则误差和样本误差, 选取 $\lambda = m^{\frac{(p-2)t}{p+\beta-1}}$ 。然后有

$$\|f_{Z,\lambda} - f_\rho\|_{L_\mu^2} \leq C \left(m^{\frac{(p-2)\beta t}{4(p+\beta-1)}} \right)$$

类似地, 当 $\beta = 1$ 时, 取 $\lambda = m^{\frac{3pt-4t}{3p}}$ 。另一种情况也可用相同的方法考虑。

致 谢

本论文是在陈爽导师的悉心指导下和同学的无私帮助下完成的。论文写作期间遇到了很多困难和障碍, 但都在同学, 导师和自己的努力下一一克服了。尤其要强烈感谢我的导师-陈爽老师, 没有她对我进行的不厌其烦的帮助和指导, 无私的为我进行论文的修改和改进, 就不会有我这篇论文的最终完成。在此, 我向我也向指导和帮助我的所有老师和同学表示衷心的感谢。同时, 我也要感谢本论文所引用的各学者的专著, 如果没有这些学者的研究成果的启发和帮助, 我将无法本篇论文的最终写作。

参考文献 (References)

- [1] Chun, X.R. and Sun, H.W. (2013) Regularized Least Square Regression with Unbounded and Dependent Sampling. *Abstract and Applied Analysis*, **2013**, 900-914.
- [2] Gao, Q. and Ye, P.X. (2016) Coefficient-Based Regularized Regression with Dependent and Unbounded Sampling. *International Journal of Wavelets Multiresolution and Information Processing*, **14**, Article ID: 1650039.
- [3] Sun, H. and Guo, Q. (2011) Coefficient Regularized Regression with Non-Iid Sampling. *International Journal of Computer Mathematics*, **88**, 3113-3124. <https://doi.org/10.1080/00207160.2011.587511>
- [4] Zhang, M.J. and Sun, H.W. (2017) Regression Learning with Non-Identically and Non-Independently Sampling. *International Journal of Wavelets Multiresolution and Information Processing*, **15**, Article ID: 1750007.

<https://doi.org/10.1142/S0219691317500072>

- [5] Vapnik, V. (1998) Statistical Learning Theory. Wiley, New York, 3185.
- [6] Pan, Z.W. and Xiao, Q.W. (2009) Least-Square Regularized Regression with Non-Iid Sampling. *Journal of Statistical Planning and Inference*, **139**, 3579-3587. <https://doi.org/10.1016/j.jspi.2009.04.007>
- [7] Sun, H.W. and Wu, Q. (2012) Regularized Least Square Regression with Dependent Sampling. *Advances in Computational Mathematics*, **32**, 175-189.
- [8] Qiang, W., Ying, Y. and Ding, X.Z. (2006) Learning Rates of Least-Square Regularized Regression. *Foundations of Computational Mathematics*, **6**, 171-192. <https://doi.org/10.1007/s10208-004-0155-9>
- [9] Sheng, B.H., Ye, P.X. and Wang, J.L. (2012) Learning Rates for Least Square Regression with Coefficient Regularization. *Acta Mathematica Sinica, English Series*, **8**, 2205-2212. <https://doi.org/10.1007/s10114-012-0607-0>
- [10] Sun, H. and Wu, Q. (2011) Least Square Regression with Indefinite Kernels and Coefficient Regularization. *Applied and Computational Harmonic Analysis*, **30**, 96-109. <https://doi.org/10.1016/j.acha.2010.04.001>
- [11] Zuo, L., Li, L.Q. and Chen, C. (2015) The Graph Based Semi-Supervised Algorithm with l^1 -Regularizer. *Neurocomputing*, **149**, 966-974.
- [12] Qu, Z.F. and Sun, H.W. (2016) Indefinite Kernel Network with l^q -Norm Regularization. *Discrete Dynamics in Nature and Society* Volume, Article ID: 6516258.
- [13] Wu, Q. (2013) Regularization Networks with Indefinite Kernels. *Journal of Approximation Theory*, **166**, 1-18.
- [14] Cai, J. (2013) Coefficient-Based Regression with Non-Identical Unbounded Sampling. *Hindawi Publishing Corporation Abstract and Applied Analysis*, 415-425. <https://doi.org/10.1155/2013/134727>
- [15] Cai, J. and Wang, C. (2013) Coefficient-Based Regularized Regression with Indefinite Kernels by Unbounded Sampling. *Scientia Sinica Mathematica*, **43**, 613-624.
- [16] Cucker, F. and Zhou, D.X. (2007) Learning Theory: An Approximation Theory Viewpoint. Cambridge University Press. <https://doi.org/10.1017/CBO9780511618796>
- [17] Sun, H. and Guo, Q. (2009) A Note On Application of Integral Operator in Learning Theory. *Applied and Computational Harmonic Analysis*, **26**, 416-421.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-2251, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: sa@hanspub.org