

文章编号:1671-6833(2022)02-0028-07

# 光滑孪生参数化不敏感支持向量回归机

黄华娟<sup>1</sup>, 韦修喜<sup>1</sup>, 周永权<sup>1,2</sup>

(1.广西民族大学 人工智能学院,广西 南宁 530006; 2.广西民族大学 广西混杂计算与集成电路设计分析重点实验室,广西 南宁 530006)

**摘要:**作为机器学习方法之一的孪生参数化不敏感支持向量回归机(TPISVR)有着简洁的数学模型,良好的学习性能,特别适合于求解带有结构异方差噪声的数据回归问题,然而TPISVR的训练速度较低,训练效率有待提高。TPISVR的传统算法可以归结为通过转化对偶问题的方法求解2个带有不等式约束的二次规划问题,然而这种求解二次规划问题的方法对于样本数目较大的问题将受到时间和内存的制约,这是导致TPISVR训练效率低的关键所在。针对此问题,首先,引入正号函数,将TPISVR的2个二次规划问题转化为2个不可微的无约束优化问题;其次,引入CHKS光滑函数和正则项,对TPISVR模型进行正则化,并对不可微的无约束优化问题进行光滑逼近,从而将不可微的模型转化为可微的无约束优化问题,并用收敛速度快的Newton-Armijo方法求解新模型,提出光滑孪生参数化不敏感支持向量回归机(STPISVR);最后,从理论上证明了STPISVR模型是收敛的,并具有任意阶光滑性。为了验证所提算法的有效性和可行性,对机器学习常用的人工数据集和UCI数据集进行仿真实验。实验结果表明:和其他机器学习方法相比,STPISVR在保证精度不下降的前提下,获得了更高的训练效率。

**关键词:**孪生参数化不敏感支持向量回归机;光滑技术;异方差噪声;Newton法;训练效率

**中图分类号:** TP18 **文献标志码:** A **doi:**10.13705/j.issn.1671-6833.2022.02.005

## 0 引言

支持向量机(support vector machine, SVM)是一种优秀的机器学习方法,具有扎实的理论基础和核映射能力,一经提出就得到了学者们的极大关注,目前在很多工程领域都得到了应用<sup>[1-3]</sup>。随着研究的深入,研究者们发现SVM虽然分类能力强,但是在训练阶段,花费的时间代价比较高。为了解决这个问题,目前有很多关于提高SVM训练速度的改进算法。例如,2001年,Fung等<sup>[4]</sup>提出了近似支持向量机(proximal support vector machine, PSVM),其基本思想是在每类样本集中设置与样本点邻近的平行超平面,并且使2个超平面之间的距离达到最大。2006年,Mangasarian等<sup>[5]</sup>对PSVM进行改进,提出了广义特征近似支持向量机(proximal support vector machine via generalized eigenvalues, GEPSVM)。GEPSVM通过求解2个广义特征值问题的最小特征值来获得全

局极值。2007年,Jayadeva等<sup>[6]</sup>在GEPSVMs基础上提出了孪生支持向量机(twin support vector machines, TWSVM)。和SVM的不同之处在于TWSVM的数学模型是2个带有不等式约束条件的二次规划问题,寻找的是一对不平行的分类超平面,隐含的并行性从理论上可以保证TWSVM的训练速度有很大的提高。由于TWSVM具有良好的分类性能,目前已被应用于说话人识别<sup>[7]</sup>、医学检测<sup>[8]</sup>等领域。

对于回归问题,2010年,Peng等<sup>[9]</sup>提出了孪生支持向量回归机(twin support vector regression, TSVR)。TSVR的数学模型是一对二次规划问题,每个二次规划问题所含约束条件的数目仅为传统支持向量回归机(support vector regression, SVR)的一半,并且TSVR的对偶问题中没有等式约束,这使得TSVR的训练速度大为提高<sup>[10-12]</sup>。然而,和很多机器学习方法一样,TSVR算法不适用于求解带有异方差噪声的样本,因为这类噪声

收稿日期:2021-07-01;修订日期:2021-09-10

基金项目:国家自然科学基金资助项目(61662005);广西自然科学基金资助项目(2018GXNSFAA294068)。

通信作者:韦修喜(1980—),男,广西百色人,广西民族大学副教授,主要从事智能计算、机器学习方面的研究,

E-mail:weixiuxi@163.com。

的分布区域是不一致的。为了解决这类问题,2010年,Hao<sup>[13]</sup>引入一种参数化不敏感损失函数,提出了参数化不敏感支持向量回归机(par-v-SVR),其实验结果表明,par-v-SVR在求解该类噪声的样本具有优势。2012年,Peng<sup>[14]</sup>借鉴这个思想,提出孪生参数化不敏感支持向量回归机(twin parametric insensitive support vector regression, TPISVR)。和TSVR相比,TPISVR更擅长处理异方差噪声的问题,并且TPISVR的训练速度有了显著的提高。2017年,丁世飞等<sup>[15]</sup>引入最小二乘方法,并采用混沌布谷鸟优化算法优化其参数,提出了最小二乘孪生参数化不敏感支持向量回归机(least squares twin parametric insensitive support vector regression, LSTPISVR),效率得到了一定的提高。然而由于算法缺乏稀疏性,其性能有待进一步提升。

为了进一步提高TPISVR的训练效率,本文引入光滑技术,提出了光滑孪生参数化不敏感支持向量回归机(smooth twin parametric insensitive support vector regression, STPISVR)。在STPISVR中,引入光滑技术,直接在原始空间将TPISVR带约束条件的二次规划问题转化为无约束优化问题,从而可以通过各种快速算法如Newton法进行求解。理论上证明了算法的光滑性和收敛性,在标准数据集上的实验结果表明,和其他机器学习方法相比,STPISVR的学习性能有明显提升。

### 1 TPISVR 基本理论

假设给定训练集  $\{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\} \in \mathbf{R}^n \times \mathbf{R}, i = 1, 2, \dots, l$ 。令  $A_{l \times n}$  为训练样本输入数据集,即  $\{x_k\}_{k=1}^l$ ; 令  $Y_{l \times n}$  为训练样本输出数据集,即  $A_{l \times n}$  对应的回归为  $Y_{l \times n} = [y_1, y_2, \dots, y_l]^T$ 。

首先简要地回顾一下TPISVR算法<sup>[14]</sup>。TPISVR需要寻找下面的一对函数:

$$\begin{cases} f_1(x) = w_1^T x + b_1; \\ f_2(x) = w_2^T x + b_2. \end{cases} \quad (1)$$

通过式(1)确定最终的回归函数。求解如下二次规划问题可以得到这对函数的参数。

$$\begin{aligned} \min & \left[ \frac{1}{2} \|w_1\|^2 - \frac{v_1}{l} e_2^T (A_2 w_1 + b_1 e_2) + \frac{c_1}{l} e_2^T \xi \right], \\ \text{s.t.} & \quad Y \geq A w_1 + b_1 e_1 - \xi, \xi \geq 0; \quad (2) \\ \min & \left[ \frac{1}{2} \|w_2\|^2 + \frac{v_2}{l} e^T (A_1 w_2 + b_2 e_1) + \frac{c_2}{l} e_1^T \eta \right], \end{aligned}$$

$$\text{s.t.} \quad Y \leq A w_2 + b_2 e_2 + \eta, \eta \geq 0. \quad (3)$$

式中:  $v_1, c_1, v_2, c_2$  为惩罚参数;  $\xi, \eta$  为松弛变量;  $e_1 \in \mathbf{R}^l$  和  $e_2 \in \mathbf{R}^l$  为分量全为1的列向量。

引入拉格朗日乘子  $\alpha$ , 式(2)变为

$$\begin{aligned} L = & \frac{1}{2} \|w_1\|^2 - \frac{v_1}{l} e^T (A w_1 + e b_1) + \frac{c_1}{l} e^T \xi - \\ & \alpha^T (Y - (A w_1 + b_1 e) + \xi) - \gamma^T \xi. \end{aligned} \quad (4)$$

根据最优化理论的KKT(Karush-Kuhn-Tucker)条件,可以得到式(4)的对偶形式:

$$\begin{aligned} \max & \left( -\frac{1}{2} \alpha^T A A^T \alpha - Y^T \alpha + \frac{v_1}{l} e^T A A^T \alpha \right), \\ \text{s.t.} & \quad 0 \leq \alpha \leq \frac{c_1}{l} e, e^T \alpha = v_1. \end{aligned} \quad (5)$$

优化后可以计算出第1个超平面的解:

$$w_1 = A^T \left( \frac{v_1}{l} e - \alpha \right), b_1 = \frac{1}{|\mathbf{SV}_1|} \sum_{i \in \mathbf{SV}_1} (y_i - A w_1). \quad (6)$$

式中:  $\mathbf{SV}_1$  为满足  $\alpha_i \in \left(0, \frac{c_1}{l}\right), i = 1, 2, \dots, l$  的样本指标的集合;  $|\cdot|$  为集合的基数。

采用同样的方法,引入拉格朗日乘子  $\beta$ , 可以得到式(3)的解:

$$w_2 = A^T \left( \beta - \frac{v_2}{l} e \right), b_2 = \frac{1}{|\mathbf{SV}_2|} \sum_{i \in \mathbf{SV}_2} (y_i - A w_2). \quad (7)$$

式中:  $\mathbf{SV}_2$  为满足  $\beta_i \in \left(0, \frac{c_2}{l}\right), i = 1, 2, \dots, l$  的样本指标的集合;  $|\cdot|$  为集合的基数。

### 2 光滑孪生参数化不敏感支持向量回归机

为了提高TPISVR的学习性能,在本节中,引入光滑技术,将TPISVR的数学模型直接转化为无约束优化问题,进而可以采用具有快速寻优能力的Newton法求解的算法,即光滑孪生参数化不敏感支持向量回归机(STPISVR)。

首先,对TPISVR的二次规划问题式(2)和式(3)进行稍微改动:①在目标函数中加入所求超平面偏移的平方,如此可以保证模型解的唯一性<sup>[16]</sup>;②将松弛变量  $\xi$  和  $\eta$  由原来的1范式修改为2范式,如此可以在一定程度上避免解的不稳定性<sup>[17-18]</sup>。改进后的2个二次规划问题如下:

$$\begin{aligned} \min & \left[ \frac{1}{2} (\|w_1\|^2 + b_1^2) - \frac{v_1}{l} e_2^T (A_2 w_1 + b_1 e_2) + \frac{c_1}{2l} e_2^T \xi \right], \\ \text{s.t.} & \quad Y \geq A_1 w_1 + b_1 e_1 - \xi, \xi \geq 0; \quad (8) \end{aligned}$$

$$\min \left[ \frac{1}{2} (\|w_2\|^2 + b_2^2) + \frac{v_2}{l} e_1^T (A_1 w_2 + b_2 e_1) + \frac{c_2}{2l} e_1^T \eta \right],$$

s.t.  $Y \leq A_2 w_2 + b_2 e_2 + \eta, \eta \geq 0$ . (9)

为方便计算,令  $W_1 = (w_1^T, b_1)^T, W_2 = (w_2^T, b_2)^T, E = (A_1, e_1)^T, F = (A_2, e_2)^T$ , 则引出如下的定理。

**定理 1** 式(8)等价于如下的无约束优化问题:

$$\min \left[ \frac{1}{2} W_1^T W_1 - \frac{v_1}{l} e_2^T F W_1 + \frac{c_1}{2l} \|(-E W_1)_+\|^2 \right].$$

(10)

式中:  $(\cdot)_+$  称为正号函数,表示把其中向量为负的分量标记为零。

**证明** 由式(8)的约束条件,得到  $\xi \geq (-A_1 w_1 - b_1 e_1)_+ = (-E W_1)_+$ , 只要证明  $(W_1^*, \xi^*)$  是优化问题式(8)最优解的必要条件是  $\xi = (-E W_1)_+$  即可。采用反证法证明,设  $(W_1^*, \xi^*)$  是优化问题(8)的最优解,且有  $\xi^* > (-E W_1)_+$  成立。令  $\xi' > (-E W_1)_+$ , 同时设优化问题式(8)的目标函数为  $F(W, \xi)$ , 则有  $F(W_1^*, \xi') < F(W_1^*, \xi^*)$  成立。由于  $\xi' \neq \xi^*$  存在,所以  $(W_1^*, \xi^*)$  不是优化问题式(8)的最优解,与假设产生矛盾。证毕。

**定理 2** 式(10)是连续且不光滑的无约束的优化问题。

**证明** 式(10)的可导性和光滑性完全取决于正号函数  $(-E W_1)_+$ , 记  $u_+ = (-E W_1)_+$ , 由  $u_+$  的定义  $u_+ = \begin{cases} u, u \geq 0 \\ 0, u < 0 \end{cases}$  可知,  $\lim_{u \rightarrow 0^+} u_+ = 1, \lim_{u \rightarrow 0^-} u_+ = 0$ , 因此  $u_+$  在  $u = 0$  处不可微。同时,  $u_+$  也可以表示为  $u_+ = \frac{|u| + u}{2}$ , 显然  $u_+$  是连续函数。因此,  $u_+$  连续但不光滑。证毕。

对于无约束优化问题(式(10)), 可以使用具有快速寻优能力的 Newton 法进行求解, 由此可以快速提高 TPISVR 的训练速度。然而, Newton 法要求目标函数具有二阶可微性, 而由定理 2 可知, 式(10)是不可微的。一条可行的途径是利用光滑技术对不可微优化问题作光滑处理, 从而易于求解。引入 CHKS(Chen-Harker-Kanzow-smale) 光滑函数后, 可得到如下的优化问题:

$$\min \left[ \frac{1}{2} W_1^T W_1 - \frac{v_1}{l} e_2^T F W_1 + \frac{c_1}{2l} \|\rho(-E W_1, \omega)\|^2 \right].$$

(11)

式中:  $\rho(x, \omega) = \frac{x + \sqrt{x^2 + 4\omega^2}}{2}$ , 这里光滑函数  $\rho(x, \omega)$  对  $(\cdot)_+$  进行了光滑处理。

**定理 3** CHKS 光滑函数  $\rho(x, \omega)$  关于  $x$  是任意阶光滑的。

实际上, CHKS 光滑函数  $\rho(x, \omega)$  已经被成功应用到支持向量机和孪生支持向量机中, 关于它的任意光滑性已经得到证明<sup>[19]</sup>。

同样地, 我们也可以得到式(9)的光滑逼近形式:

$$\min \left[ \frac{1}{2} W_2^T W_2 + \frac{v_2}{l} e_1^T F W_2 + \frac{c_2}{2l} \|\rho(-F W_2, \omega)\|^2 \right].$$

(12)

由定理 3 可知, STPISVR 的优化问题式(11)和式(12)可以用 Newton 法快速求解。

求解非线性问题时, 在 STPISVR 数学模型中引入核函数即可。

基于核空间的式(8)和式(9)的优化问题为

$$\min \left\{ \frac{1}{2} (\|w_1\|^2 + b_1^2) - \frac{v_1}{l} e_2^T [K(A_2, A) w_1 + b_1 e_2] + \frac{c_1}{2l} e_2^T \xi \right\},$$

$$\text{s.t. } Y \geq K(A_2, A) w_1 + b_1 e_1 - \xi, \xi \geq 0; \quad (13)$$

$$\min \left\{ \frac{1}{2} (\|w_2\|^2 + b_2^2) + \frac{v_2}{l} e_1^T [K(A_1, A) w_2 + b_2 e_1] + \frac{c_2}{2l} e_1^T \eta \right\},$$

$$\text{s.t. } Y \leq K(A_1, A) w_2 + b_2 e_2 + \eta, \eta \geq 0. \quad (14)$$

式中:  $A = [A_1, A_2]$ ;  $K(\cdot, \cdot)$  是选定的核函数。

引入 CHKS 光滑函数, 则可得到非线性 STPISVR 的数学模型:

$$\min \left[ \frac{1}{2} W_1^T W_1 - \frac{v_1}{l} e_2^T K_2 W_1 + \frac{c_1}{2l} \|\rho(-K_1 W_1, \omega)\|^2 \right];$$

(15)

$$\min \left[ \frac{1}{2} W_2^T W_2 + \frac{v_2}{l} e_1^T K_1 W_2 + \frac{c_2}{2l} \|\rho(-K_2 W_2, \omega)\|^2 \right].$$

(16)

式中:  $K_1 = [K(A_1, A), e_1]$ ;  $K_2 = [K(A_2, A), e_2]$ 。

前面定理也适用于非线性 STPISVR, 即优化问题式(15)和式(16)是任意阶光滑的, 可使用 Newton 法快速求解。

### 3 实验与分析

为了测试 STPISVR 算法的学习能力, 本节进行 2 组实验: 第 1 组实验是在 1 个常用的人工数据集上; 第 2 组实验是在 10 个标准数据集上。实验结果与 TSVR、TPISVR 和 LSTPISVR<sup>[15]</sup> 这 3 个算法的结果进行对比。运行环境设置为 Intel(R) Core(TM)2 Duo CPU E4500, 2 G 内存和 MATLAB。

TSVR、TPISVR、LSTPISVR 和 STPISVR 的参数都是采用网格划分方法进行确定,搜索范围为  $[2^{-7}, 2^{12}]$ 。本文采用的核函数为高斯核函数,其表达式为  $K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right)$ , Newton 法结束时下降方向的模  $\varepsilon_1 = 1.0 \times 10^{-3}$ , CHKS 函数的参数  $\omega = 1.0 \times 10^{-5}$ 。

### 3.1 人工数据集上的实验

sinc 函数经常被用来测试各种机器学习方法的回归性能<sup>[19]</sup>,其表达式为

$$y = \text{sinc}(x_i) = \frac{\sin x_i}{x_i} + e_i, x_i \sim U[-3\pi, 3\pi]。$$

为了更好地测试算法的学习性能,在训练样

本中,分别加入如下的异方差结构噪声:

$$\text{Type A: } e_i = \left(-\frac{|x_i|}{8\pi} + 0.5\right) \varepsilon_i, \varepsilon_i \sim U[-0.5, 0.5];$$

$$\text{Type B: } e_i = \left(-\frac{|x_i|}{8\pi} + 0.5\right) \varepsilon_i, \varepsilon_i \sim N(0, 0.25^2)。$$

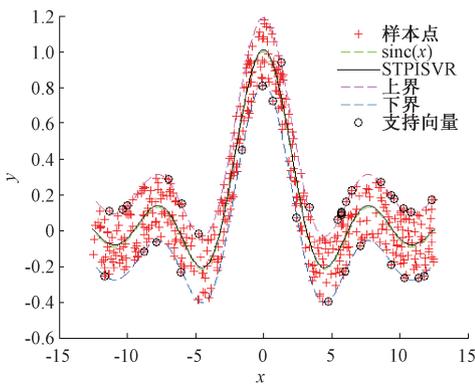
式中:  $U[a, b]$  表示的是在  $[a, b]$  上的均匀随机变量,  $N(c, d^2)$  表示的是均值为  $c$ 、方差为  $d^2$  的高斯随机变量。

该数据集的训练样本和测试样本都设置为 5 000。表 1 为 LSTPISVR、TSVR、TPISVR 和 STPISVR 4 种算法独立运行 30 次的平均结果。图 1 和图 2 分别为 LSTPISVR、TSVR、TPISVR 和 STPISVR 对带不同噪声的 sinc 函数的拟合效果。

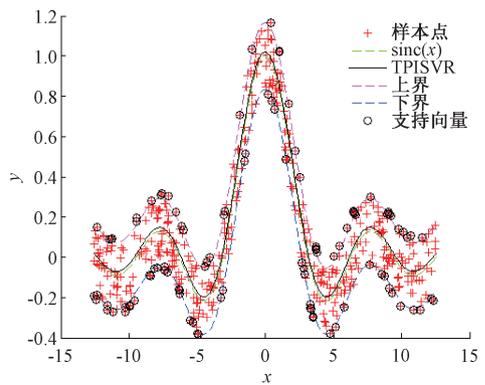
表 1 4 种算法在带有 2 种不同类型噪声 sinc 函数下的比较结果

Table 1 Comparison results of four algorithms with two different types of noise sinc functions

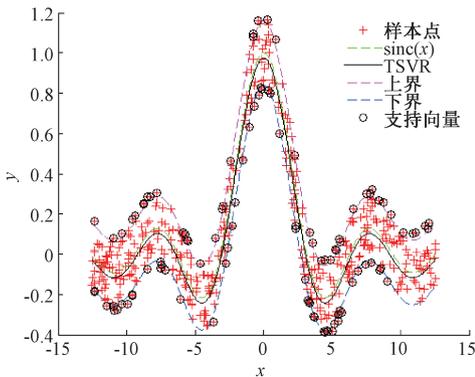
算法	均方根误差 RMSE		训练时间/s	
	Type A	Type B	Type A	Type B
STPISVR	<b>0.083 2±0.001 5</b>	<b>0.086 3±0.001 1</b>	<b>0.056 4</b>	<b>0.065 3</b>
TPISVR	0.092 1±0.002 4	0.091 2±0.002 5	0.092 4	0.589 6
TSVR	0.101 2±0.002 5	0.136 6±0.001 5	0.095 2	2.082 1
LSTPISVR	0.096 2±0.002 1	0.101 2±0.002 3	0.331 2	0.321 4



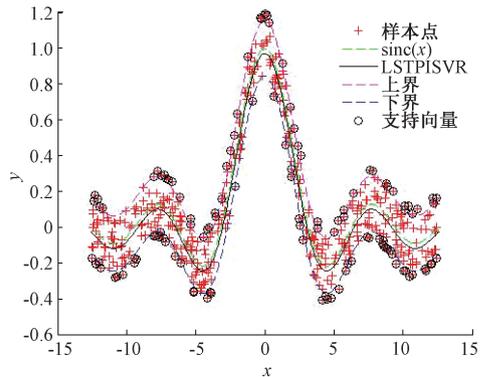
(a) STPISVR对带有Type A噪声的sinc(x)的拟合结果



(b) TPISVR对带有Type A噪声的sinc(x)的拟合结果



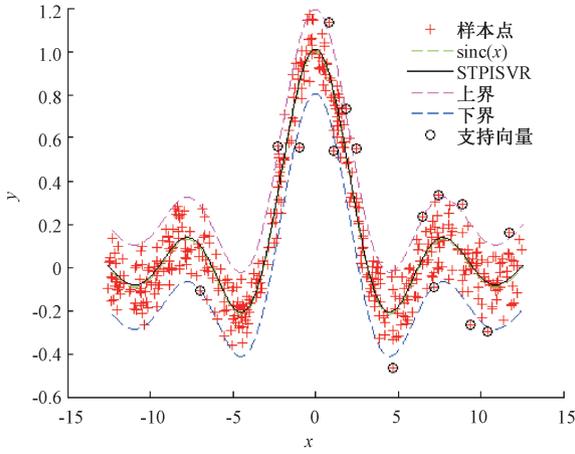
(c) TSVR对带有Type A噪声的sinc(x)的拟合结果



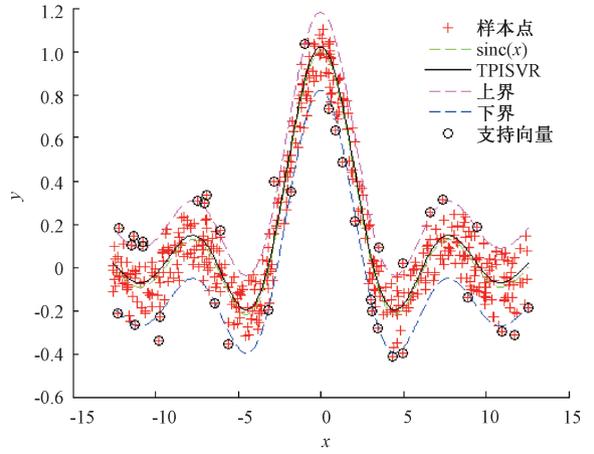
(d) LSTPISVR对带有Type A噪声的sinc(x)的拟合结果

图 1 4 种算法对带有 Type A 噪声的 sinc(x) 的拟合结果

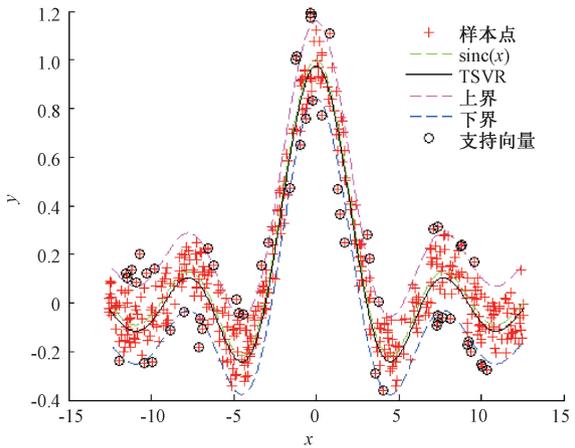
Figure 1 Fitting results of four algorithms with sinc(x) of Type A noise



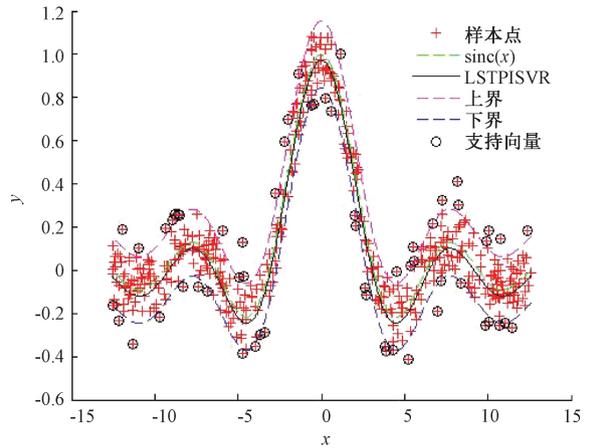
(a) STPISVR对带有Type B噪声的sinc(x)的拟合结果



(b) TPISVR对带有Type B噪声的sinc(x)的拟合结果



(c) TSVR对带有Type B噪声的sinc(x)的拟合结果



(d) LSTPISVR对带有Type B噪声的sinc(x)的拟合结果

图2 4种算法对带有Type B噪声的sinc(x)的拟合结果

Figure 2 Fitting results of four algorithms with sinc(x) of Type B noise

从表1可以看出,对于带有异方差噪声的sinc函数,STPISVR、TPISVR和LSTPISVR的拟合效果都比TSVR的效果好,说明前3种算法在处理异方差噪声方面有独特的优势。而在STPISVR、TPISVR和LSTPISVR这3种算法中,STPISVR在回归精度不下降甚至更好的情况下,其使用的训练时间是最小的。表1的实验结果表明,引入光滑技术,采用Newton法求解算法模型,可以快速提高训练的效率。图1是4种算法对带有Type A噪声的函数sinc(x)的拟合结果。从图1可以直观地看出,STPISVR算法(图1(a))的支持向量最少,所以使用的训练时间最少。图2是4种算法对带有Type B噪声的sinc(x)函数的拟合结果。从图2也可以看出,STPISVR算法(图2(a))的回归函数和sinc函数基本重合,说明其拟合精度是比较令人满意的。同时和其他算法相比,STPISVR的支持向量是最少的,所以其训练

的效率也得到了提升。

### 3.2 UCI数据集上的实验

为了进一步验证算法的性能,本文对10个常用的UCI数据集进行测试,这些数据集常被用来测试机器学习算法的性能。实验环境、参数设置和3.1节一样。表2显示的是4种算法的运行结果。

由表2可以看出,和其他算法相比,STPISVR在拟合效果上不一定是最好的,比如在Boston、Mach CPU和Servo这3个数据集上,STPISVR的RMSE不是最好的,但是精度还在可接受的范围内。最重要的,STPISVR所花费的训练时间是最少的。对于机器学习算法来说,能在满意解内取得最高的训练效率是很重要的。Triazines是一个60维的高维数据集,Wine quality white是一个具有4898个样本的较大数据集,从表2可以看出,STPISVR处理这2个数据集时同样具有不错的性能。

表 2 4 种算法在 UCI 数据集的测试结果  
Table 2 Test results of four algorithms on UCI Data sets

数据集	均方根误差 <i>RMSE</i>				训练时间/s			
	STPISVR	TPISVR	TSVR	LSTPISVR	STPISVR	TPISVR	TSVR	LSTPISVR
Motorcycle (133×2)	0.204±0.005	0.231±0.001	0.265±0.006	0.218±0.002	0.021 3	0.0512	0.043 2	0.042 5
Boston (506×14)	0.143±0.009	0.124±0.013	0.165±0.015	0.152±0.042	0.083 3	0.176 5	0.160 6	0.106 8
Auto-Mpg (398×8)	0.109±0.032	0.231±0.011	0.198±0.022	0.134±0.045	0.053 3	0.232 1	0.123 2	0.073 6
Math.CPU (209×7)	0.067±0.003	0.055±0.012	0.012±0.008	0.092±0.006	0.013 3	0.024 5	0.034 5	0.018 5
Servo (167×5)	0.177±0.034	0.199±0.023	0.109±0.018	0.181±0.092	0.015 5	0.031 3	0.043 2	0.023 5
Auto price (159×16)	0.104±0.023	0.332±0.023	0.154±0.021	0.160±0.043	0.013 4	0.039 9	0.226 5	0.024 3
Triazines (186×60)	0.109±0.011	0.228±0.256	0.465±0.012	0.466±0.123	0.029 8	0.127 6	0.488 7	0.297 6
Concrete Cs (1 030×9)	0.104±0.098	0.158±0.231	0.589±0.124	0.376±0.322	0.076 6	0.154 3	0.456 5	0.198 8
Wine quality white (4 898×12)	0.053±0.067	0.199±0.145	0.676±0.028	0.643±0.122	0.265 4	0.567 7	1.254 3	0.897 4
Lorenz <sub>0.05</sub> (3 000×6)	0.112±0.178	0.232±0.078	0.797±0.265	0.358±0.028	0.064 4	0.174 3	0.982 6	0.267 4

### 4 结论

孪生参数化不敏感支持向量回归机 (TPISVR) 的模型是一对带有约束条件的二次规划问题。如果直接在对偶空间求解 TPISVR, 所花费的训练时间比较多。为了提高 TPISVR 的训练效率, 本文引入光滑技术, 将 TPISVR 的带有不等式约束的二次规划问题等价地转化为 2 个无约束的优化问题, 并从理论上证明改进后的目标函数具有二阶可微性, 由此可以直接采用具有快速寻优能力的 Newton 法进行求解, 提出了光滑孪生参数化不敏感支持向量回归机 (STPISVR)。在 UCI 数据集和加入异方差噪声的人工数据集上的实验结果都表明了和其他算法相比, 在保证拟合效果不下降的情况下, STPISVR 可以得到较高的训练效率。

### 参考文献:

[1] CORTES C, VAPNIK V. Support-vector networks [J]. Machine learning, 1995, 20(3): 273-297.  
 [2] MELLO A R, STEMMER M R, KOERICH A L. Incremental and decremental fuzzy bounded twin support vector machine [J]. Information sciences, 2020, 526: 20-38.

[3] 孙国栋, 江亚杰, 徐亮, 等. BP 网络预测阈值的仪表重影字符识别方法研究 [J]. 郑州大学学报 (工学版), 2020, 41(4): 28-33.  
 [4] FUNG G, MANGASARIAN O L. Proximal support vector machine classifiers [C]//Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining-KDD'01. New York: ACM, 2001: 77-86.  
 [5] MANGASARIAN O L, WILD E W. Multisurface proximal support vector machine classification via generalized eigenvalues [J]. IEEE transactions on pattern analysis and machine intelligence, 2006, 28(1): 69-74.  
 [6] JAYADEVA, KHEMCHANDANI R, CHANDRA S. Twin support vector machines for pattern classification [J]. IEEE transactions on pattern analysis and machine intelligence, 2007, 29(5): 905-910.  
 [7] CHEN W J, SHAO Y H, LI C N, et al. N-projection twin support vector machine for pattern classification [J]. Neurocomputing, 2020, 376: 10-24.  
 [8] ZHANG X S, GAO X B, WANG Y. Twin support tensor machines for MCs detection [J]. Journal of electronics (China), 2009, 26(3): 318-325.  
 [9] PENG X J. TSVR: an efficient Twin Support Vector Machine for regression [J]. Neural networks, 2010, 23(3): 365-372.

- [10] LI Y M, SUN H J, YAN W Z, et al. Multi-output parameter-insensitive kernel twin SVR model[J]. *Neural networks*, 2020, 121: 276–293.
- [11] 黄文锋, 徐珊珊, 孙焱, 等. 基于多分辨率卷积神经网络的火焰检测[J]. *郑州大学学报(工学版)*, 2019, 40(5): 79–83.
- [12] LÓPEZ J, MALDONADO S. Robust twin support vector regression via second-order cone programming [J]. *Knowledge-based systems*, 2018, 152: 83–93.
- [13] HAO P Y. New support vector algorithms with parametric insensitive/margin model [J]. *Neural networks*, 2010, 23(1): 60–73.
- [14] PENG X J. Efficient twin parametric insensitive support vector regression model [J]. *Neurocomputing*, 2012, 79: 26–38.
- [15] 丁世飞, 黄华娟. 最小二乘孪生参数化不敏感支持向量回归机 [J]. *软件学报*, 2017, 28(12): 3146–3155.
- [16] 丁世飞, 黄华娟, 史忠植. 加权光滑 CHKS 孪生支持向量机 [J]. *软件学报*, 2013, 24(11): 2548–2557.
- [17] 黄华娟, 丁世飞, 史忠植. 光滑 CHKS 孪生支持向量回归机 [J]. *计算机研究与发展*, 2015, 52(3): 561–568.
- [18] 王震. 基于非平行超平面支持向量机的分类问题研究 [D]. 长春: 吉林大学, 2014.
- [19] 黄华娟. 孪生支持向量机关键问题的研究 [D]. 徐州: 中国矿业大学, 2014.

## Smooth Twin Parametric Insensitive Support Vector Regression

HUANG Huajuan<sup>1</sup>, WEI Xiuxi<sup>1</sup>, ZHOU Yongquan<sup>1,2</sup>

(1. College of Artificial Intelligence, Guangxi University for Nationalities, Nanning 530006, China; 2. Guangxi Key Laboratory of Hybrid Computation and IC Design Analysis, Guangxi University for Nationalities, Nanning 530006, China)

**Abstract:** As one of the machine learning methods, twin parametric insensitive support vector regression (TPISVR) had a simple mathematical model and good learning performance. It was especially suitable for solving data regression problems with structural heteroscedasticity noise. However, the training speed of TPISVR was low, and the training efficiency needs to be improved. The traditional algorithm of TPISVR could be reduced to solve two quadratic programming problems with inequality constraints by transforming dual problems. However, this method of solving quadratic programming problems with large number of samples would be restricted by time and memory, which was the key to the low training efficiency of TPISVR. In this study, the positive sign function was introduced to transform the two quadratic programming problems of TPISVR into two non-differentiable unconstrained optimization problems. Secondly, CHKS smooth function and regular term were introduced to regularize TPISVR model, and smooth approximation was made to the non-differentiable unconstrained optimization problem, so as to transform the non-differentiable model into a differentiable unconstrained optimization problem. The new model was solved by Newton Armijo method with fast convergence speed, and a smooth twin parameterized insensitive support vector regression machine (STPISVR) was proposed. Finally, it was proved theoretically that STPISVR model was convergent and had arbitrary order smoothness; In order to verify the effectiveness and feasibility of the algorithm, simulation experiments were carried out on the artificial data set and UCI data set commonly used in machine learning. The experimental results showed that compared with other machine learning methods, STPISVR achieved higher training efficiency without reducing the accuracy.

**Keywords:** twin parametric insensitive support vector regression; smooth technology; heteroscedastic noise; Newton method; training efficiency