



智能系统学报

CAAI TRANSACTIONS ON INTELLIGENT SYSTEMS

l_1l_1 双范数的最优下边界回归模型辨识

刘小雍, 叶振环

引用本文:

刘小雍, 叶振环. l_1l_1 双范数的最优下边界回归模型辨识[J]. 智能系统学报, 2020, 15(5): 934–942.

LIU Xiaoyong, YE Zhenhuan. Optimal lower boundary regression model based on double norms l_1l_1 optimization[J]. *CAAI Transactions on Intelligent Systems*, 2020, 15(5): 934–942.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.201902006>

您可能感兴趣的其他文章

自适应多阶段线性重构表示分类的人脸识别

Self-adaptive multi-phase linear reconstruction representation based classification for face recognition
智能系统学报. 2020, 15(5): 964–971 <https://dx.doi.org/10.11992/tis.201904002>

基于知识距离的粗糙粒结构的评价模型

Evaluation model of rough granular structure based on knowledge distance
智能系统学报. 2020, 15(1): 166–174 <https://dx.doi.org/10.11992/tis.201904037>

图正则化稀疏判别非负矩阵分解

Graph-regularized, sparse discriminant, non-negative matrix factorization
智能系统学报. 2019, 14(6): 1217–1224 <https://dx.doi.org/10.11992/tis.201811021>

鲁棒的半监督多标签特征选择方法

A robust, semi-supervised, and multi-label feature selection method
智能系统学报. 2019, 14(4): 812–819 <https://dx.doi.org/10.11992/tis.201809017>

快速双非凸回归算法及其电力数据预测应用

Fast double nonconvex regression algorithm for forecast of electric power data
智能系统学报. 2018, 13(4): 665–672 <https://dx.doi.org/10.11992/tis.201708033>

稀疏样本自表达子空间聚类算法

Sparse sample self-representation for subspace clustering
智能系统学报. 2016, 11(5): 696–702 <https://dx.doi.org/10.11992/tis.201601005>

微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.201902006

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20191031.1159.002.html>

ℓ_1 - ℓ_1 双范数的最优下边界回归模型辨识

刘小雍, 叶振环

(遵义师范学院 工学院, 贵州 遵义 563006)

摘要: 考虑到来自传感器测量数据、模型结构以及参数的不确定性等因素, 建模由这些因素导致的下边界模型尤为重要。通过将结构风险最小化理论与逼近误差最小化思想相结合, 提出了 ℓ_1 - ℓ_1 双范数的最优下边界回归模型建模方法。首先, 确定满足下边界回归模型的约束条件。其次, 将结构风险的 ℓ_2 范数转化为简单的 ℓ_1 范数优化问题, 并将回归模型与实际测量数据之间的逼近误差的 ℓ_1 范数融合到结构风险的 ℓ_1 范数优化问题, 再应用较简单的线性规划对双范数的优化问题进行求解获取模型参数。最后, 通过来自测量数据以及模型参数不确定性的实验分析, 论证了提出方法的最优性, 体现在: 下边界模型的建模精度通过逼近误差的 ℓ_1 范数得到保证; 模型结构复杂性在结构风险的 ℓ_1 范数优化条件下得到有效控制, 进而提高其泛化性能。

关键词: ℓ_1 范数的结构风险最小化; 逼近误差的 ℓ_1 范数; 下边界回归模型; 泛化性能; 建模精度; 最优性; 线性规划
中图分类号: TP391.1 **文献标志码:** A **文章编号:** 1673-4785(2020)05-0934-09

中文引用格式: 刘小雍, 叶振环. ℓ_1 - ℓ_1 双范数的最优下边界回归模型辨识 [J]. 智能系统学报, 2020, 15(5): 934-942.

英文引用格式: LIU Xiaoyong, YE Zhenhuan. Optimal lower boundary regression model based on double norms ℓ_1 - ℓ_1 optimization[J]. CAAI transactions on intelligent systems, 2020, 15(5): 934-942.

Optimal lower boundary regression model based on double norms ℓ_1 - ℓ_1 optimization

LIU Xiaoyong, YE Zhenhuan

(College of Engineering, Zunyi Normal University, Zunyi 563006, China)

Abstract: In statistical modeling, regression analysis is a set of statistical processes for estimating the relationships between a dependent variable and one or more independent variables. Considering the uncertainties in the structure and parameters of the model derived from sensor measurement data, a new model called optimal lower boundary model is proposed to remove the uncertainties in parameters and characteristics. The proposed method is a combination of structural risk minimization theory (SRM) and some ideas from approximation error minimization. An optimal lower boundary regression model (LBRM) is presented using ℓ_1 - ℓ_1 double norms optimization. First, constraint conditions subjected to LBRM are defined. Then, ℓ_2 -norm optimization based on structural risk is converted into simple ℓ_1 -norm optimization so that approximation error between the measurements based on ℓ_1 -norm is computed and minimized. Next, LBRM is integrated into ℓ_1 -norm optimization (based on structural risk). Thus, simpler linear programming can be applied to the constructed double-norms optimization problem to solve parameters of LBRM. Finally, the proposed method is demonstrated by experiments regarding uncertain measurements and parameters of nonlinear system. It has the following prominent features: modeling accuracy of LBRM can be guaranteed by introducing the ℓ_1 -norm minimization on approximation error; model's structural complexity is under control by ℓ_1 -norm optimization based on structural risk, thus the performance of the model can be improved further.

Keywords: ℓ_1 -norm-based structural risk minimization; ℓ_1 -norm on approximation error; lower boundary regression model; generalization performance; modeling accuracy; optimality; linear programming

收稿日期: 2019-02-08. 网络出版日期: 2019-10-31.

基金项目: 贵州省科技计划基金项目(黔科合基础[2018]1179); 贵州省教育厅青年基金项目(黔教合 KY 字[2016]254); 遵义师范学院博士项目(遵师 BS[2015]04 号).

通信作者: 刘小雍. E-mail: liuxy204@163.com.

近年来, 在理论或工程上的诸多应用, 包括控制器设计、高级过程仿真、软计算以及故障诊断等, 都离不开对被研究的复杂非线性系统的建模^[1]. 因此, 建立被研究对象的非线性动态数学模型在

实际工程应用中变得尤为重要。由于诸多不确定性的存在,例如模型结构以及参数等,导致非线性系统的机理建模出现了巨大挑战^[2-3]。因此,出现了基于数据的两种经典方法:1)基于经验风险最小化的神经网络(neural network, NN);2)采用结构风险最小化理论的支持向量机(support vector machine, SVM)及其变体最小二乘支持向量机(least squares SVM),都被广泛应用于非线性系统的建模研究。

从理论上讲,神经网络可以以任意的精度逼近任意的非线性系统^[4],在非线性系统的建模领域有着大量的研究^[5]。例如,在文献[6-7]中提出的带随机权值分配的级联神经网络,从某种程度上其建模精度得到了较大改善。为了能够实现非线性系统建模过程中的快速鲁棒收敛,一种自适应二阶算法^[8]被提出用于训练模糊神经网络,获取了满意的建模结果。分层径向基函数神经网络^[9-10]作为NN的另一种变体,通过对污水处理的非线性建模,在实际应用中的预测性能上都达到了较好的效果。然而,上述提到的这些方法仅仅考虑了单隐层结构,在建模精度上仍缺乏显著改进。根据统计学一致逼近理论可知^[11],当NN的隐神经元个数选取较多时,甚至等于训练样本的数量时,单隐层NN就能以足够高的精度去逼近任意的非线性系统;然而,较多的样本数量会引起神经元个数的增加,导致NN的模型结构复杂,泛化性能变差。此外,众多神经网络在参数求解过程中,主要还是采用经验风险最小化理论^[12],即神经网络的参数最终解是以模型预测输出与实际输出之间的平方和达到最小作为标准,进而导致训练获取的神经网络模型复杂,容易产生局部极小与过拟合问题。由Vapnik提出的SVM,通过执行结构风险最小化来代替经验风险最小化,理论上保证了SVM在非线性系统建模上的全局最优,已成为分类和回归应用中的一种重要学习方法。在非线性回归领域,通过大量的实验研究表明,SVM的泛化性能优于神经网络及其变体的非线性建模方法。基于此,在SVM基础上,文献[13-14]提出了基于支持向量学习方法的模糊回归分析,该方法较传统神经网络方法在泛化性能上做了较好的改进。基于此,基于数据的另一种模糊建模方法,也将基于增量平滑SVR的结构风险最小化作为优化问题^[15],进而提高模型泛化性能。近年来,深度学习俨然成为研究的热点,文献[16-17]围绕非线性系统的建模问题,提出了一种基于改进型深度学习的非线性建模方法。

目前,各种数据建模方法主要集中在确定性

数学模型建模的研究,其鲁棒性差,易受外界干扰,很少有针对来自模型结构、参数以及测量数据的不确定性等因素引起的最优下界建模;此外,如何控制所建立模型的结构复杂性,提高泛化性能,也是需要考虑的重点。在本文中,考虑到基于结构风险最小化的支持向量机所具有的良好特性,将其转化为 ℓ_1 范数下的结构风险,并将逼近误差的 ℓ_1 范数思想与之相结合,建立求解最优下界回归模型的优化问题,再应用较简单的线性规划获取下界回归模型的稀疏解。提出的方法可归纳为:1)提出了最小化最大逼近误差的范数定理,作为分别建立下界回归模型的优化问题;2)建立基于结构风险最小化的代价函数,在保证辨识下界回归模型精度的同时,尽可能对模型结构复杂性进行有效控制,进而提高模型的泛化性能;3)下界输出模型包络了由各种不确定性因素引起的输出,进而提高建模的鲁棒性。

1 支持向量回归的 ℓ_1 范数问题转化

1.1 支持向量回归问题

随着Vapnik的不敏感损失函数的引入^[18-19],支持向量机的分类问题被扩展到回归问题,即支持向量回归(SVR),已在最优控制、时间序列预测、区间回归分析等方面得到了广泛应用。SVR方法是对一组带有噪声的测量数据 $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_N, y_N)\}$ 的未知函数进行逼近,其中 $\mathbf{x}_k = (x_k^1, x_k^2, \dots, x_k^d)$ 表示对应测量数据的输入, d 表示每个输入为 d 维的特征模式, y_k 为相应的测量输出, k 表示第 k 个输入和输出, N 为获取的总测量数据个数。因此,函数的逼近问题可转化为寻求如下基函数线性展开的最优参数^[20]:

$$f(\mathbf{x}, \theta) = \sum_{k=1}^m \theta_k g_k(\mathbf{x}) + b \quad (1)$$

式中: $\theta = (\theta_1, \theta_2, \dots, \theta_m)$ 为需要被寻优的参数向量, b 是一个常量,式(1)表示对于 N 个测量数据,可用 m 个无关的基函数线性组合对其建模。进一步,该问题的参数寻优即为寻找满足如下优化问题的非线性函数 f :

$$\min: R(f) = \sum_{k=1}^N L_\varepsilon(y_k - f(\mathbf{x}_k)) + \gamma \|\mathbf{w}\|_2^2 \quad (2)$$

$R(f)$ 为结构风险, γ 表示规则化常量, $\|\mathbf{w}\|_2^2$ 的引入在于控制模型的复杂度, $L_\varepsilon(\cdot)$ 描述 ε -不敏感损失函数,定义为

$$L_\varepsilon(y_k - f(\mathbf{x}_k)) = \begin{cases} 0, & |y_k - f(\mathbf{x}_k)| \leq \varepsilon \\ |y_k - f(\mathbf{x}_k)| - \varepsilon, & \text{其他} \end{cases}$$

从上述 ε 域定义可知,如果 $|y_k - f(\mathbf{x}_k)|$ 的值在该 ε 区域内,损失为0;否则为 $|y_k - f(\mathbf{x}_k)|$ 与 ε 的差值。

通过应用拉格朗日乘子方法,对式(2)的最小化可转化为它的对偶优化问题:

$$\begin{aligned} \min: \quad & W(\alpha^+, \alpha^-) = \varepsilon \sum_{k=1}^N L_\varepsilon(\alpha_k^- + \alpha_k^+) - \sum_{k=1}^N y_k (\alpha_k^+ - \alpha_k^-) + \\ & \frac{1}{2} \sum_{k,i=1}^N \left((\alpha_k^+ - \alpha_k^-) (\alpha_i^+ - \alpha_i^-) \sum_{s=1}^m g_s(\mathbf{x}_k) g_s(\mathbf{x}_i) \right) \\ \text{s.t.} \quad & \sum_{k=1}^N \alpha_k^+ = \sum_{k=1}^N \alpha_k^-, \quad 0 \leq \alpha_k^+, \alpha_k^- \leq \gamma, \\ & \text{for } k = 1, 2, \dots, N \end{aligned} \quad (3)$$

α^+ 、 α^- 、 α_k^+ 、 α_k^- 表示拉格朗日乘子。式(3) $g_s(\mathbf{x})$ 的内积可用如下核函数代替:

$$K(\mathbf{x}_k, \mathbf{x}_i) = \sum_{s=1}^m g_s(\mathbf{x}_k) g_s(\mathbf{x}_i)$$

核函数确定了解的平滑特性,选取时应该更好的反映数据的先验知识。式(3)的优化问题可写为

$$\begin{aligned} \min: \quad & W(\alpha^+, \alpha^-) = \varepsilon \sum_{k=1}^N L_\varepsilon(\alpha_k^- + \alpha_k^+) - \sum_{k=1}^N y_k (\alpha_k^+ - \alpha_k^-) + \\ & \frac{1}{2} \sum_{k,i=1}^N ((\alpha_k^+ - \alpha_k^-) (\alpha_i^+ - \alpha_i^-) K(\mathbf{x}_k, \mathbf{x}_i)) \end{aligned}$$

基于 Vapnik 的研究,SVR 方法的解以核函数的线性展开描述为

$$f(\mathbf{x}, \alpha^+, \alpha^-) = \sum_{k=1}^m (\alpha^+ - \alpha^-) K(\mathbf{x}, \mathbf{x}_k) + b \quad (4)$$

其中常量 b 的计算为

$$b = y_k - \sum_{k=1}^N (\alpha_k^+ - \alpha_k^-) K(\mathbf{x}_k, \mathbf{x}_k) + \varepsilon \cdot \text{sign}(\alpha_k^- - \alpha_k^+)$$

显然,仅当 $(\alpha_k^+ - \alpha_k^-) \neq 0$ 时,对应的样本 \mathbf{x}_k 称为支持向量(SVs)。在 SVR 方法中,其核函数的使用包括高斯核函数(Gaussian radial basis function, GRBF)、多项式核函数、Sigmoid 核函数、可逆多二次核函数(inverse multi-quadric kernel)等。然而,通过大量的实验研究表明,高斯核函数相对于其他核函数在实际应用中易于实现且具有较强的映射能力。因此本文在回归模型辨识中采用高斯核函数,可将式(4)写成:

$$f(\mathbf{x}, \alpha^+, \alpha^-) = \sum_{k=1}^m (\alpha_k^+ - \alpha_k^-) \exp \left\{ \frac{-\|\mathbf{x} - \mathbf{x}_k\|^2}{2\sigma^2} \right\} + b \quad (5)$$

式中 σ 称为高斯核参数。

1.2 SVR 的 ℓ_1 范数优化问题转化

SVR 采用结构风险最小化理论建立求解模型参数的凸二次规划问题,不仅保证了模型建模精度,而且模型结构的稀疏特性也得到了保证,被广泛应用于模式识别以及非线性内动态系统建模。然而,正如 1.1 节的 SVR 回归问题那样,其传

统的二次规划-SVR(quadric programming-support vector regression, QP-SVR)在执行参数的求解过程中,容易产生模型的冗余描述及昂贵的计算成本^[18]。对于 QP-SVR,基于式(2)的优化问题,

$$\begin{aligned} \min: \quad & R(f) = C \sum_{k=1}^N L_\varepsilon(\xi_k + \xi_k^*) + \frac{1}{2} \|\mathbf{w}\|_2^2 \\ \text{s.t.} \quad & \begin{cases} y_k - \langle \mathbf{w}, \varphi(\mathbf{x}_k) \rangle - b \leq \varepsilon + \xi_k, \\ \langle \mathbf{w}, \varphi(\mathbf{x}_k) \rangle + b - y_k \leq \varepsilon + \xi_k^* \\ \xi_k, \xi_k^* \geq 0 \end{cases} \end{aligned}$$

其中 $\varphi(\cdot)$ 表示从输入空间到高维空间的非线性特征映射,即 $\varphi: \mathbf{R}^n \rightarrow \mathbf{R}^m (m > n)$; ξ_k 、 ξ_k^* 为松弛变量,分别对应超出正、负方向偏差值时的大小;常量 C 大于 0,反应非线性 f 与偏差大于 ε 时两者之间的平衡。对于式(5),令 $\beta_k = \alpha_k^+ - \alpha_k^-$, 则有

$$f(\mathbf{x}, \beta) = \sum_{k=1}^N \beta_k \exp \left(\frac{-\|\mathbf{x} - \mathbf{x}_k\|^2}{2\sigma^2} \right) + b \quad (6)$$

$\beta = [\beta_1, \beta_2, \dots, \beta_N]^T$ 。考虑到式(2)的优化问题, $\|\mathbf{w}\|_2^2$ 范数的引入是为了控制模型的复杂度,根据范数的等价性可知,在结构风险中引入其他范数也可以同样对模型复杂性进行控制。接下来,将 QP-SVR 的优化问题(2)变成

$$\min: \quad R(f) = \sum_{k=1}^N L_\varepsilon(y_k - f(\mathbf{x}_k)) + \gamma \|\beta\|_1$$

其中, $f(\mathbf{x})$ 以式(5)形式描述, $\|\beta\|_1$ 表示系数空间的 ℓ_1 范数。因此,新的约束优化问题为

$$\begin{aligned} \min: \quad & R(f) = C \sum_{k=1}^N (\xi_k + \xi_k^*) + \|\beta\|_1 \\ \text{s.t.} \quad & \begin{cases} y_k - \sum_{k=1}^m \beta_k \exp \left\{ \frac{-\|\mathbf{x} - \mathbf{x}_k\|^2}{2\sigma^2} \right\} - b \leq \varepsilon + \xi_k^* \\ \sum_{k=1}^m \beta_k \exp \left\{ \frac{-\|\mathbf{x} - \mathbf{x}_k\|^2}{2\sigma^2} \right\} + b - y_k \leq \varepsilon + \xi_k \\ \xi_k, \xi_k^* \geq 0 \end{cases} \end{aligned} \quad (7)$$

从几何的角度来看, ξ_k 和 ξ_k^* 之间的关系在 SVR 中满足 $\xi_k \xi_k^* = 0$ 。因此,在优化问题(7)中仅引入一个松弛变量 ξ_k 即可^[21], 即

$$\begin{aligned} \min: \quad & R(f) = C \sum_{k=1}^N \xi_k + \|\beta\|_1 \\ \text{s.t.} \quad & \begin{cases} y_k - \sum_{k=1}^m \beta_k \exp \left\{ \frac{-\|\mathbf{x} - \mathbf{x}_k\|^2}{2\sigma^2} \right\} - b \leq \varepsilon + \xi_k \\ \sum_{k=1}^m \beta_k \exp \left\{ \frac{-\|\mathbf{x} - \mathbf{x}_k\|^2}{2\sigma^2} \right\} + b - y_k \leq \varepsilon + \xi_k \\ \xi_k \geq 0 \end{cases} \end{aligned} \quad (8)$$

为了转化上述优化问题为线性规划问题,将 β_k 和 $|\beta_k|$ 进行如下分解:

$$\begin{aligned}\beta_k &= \alpha_k^+ - \alpha_k^- \\ |\beta_k| &= \alpha_k^+ + \alpha_k^-\end{aligned}\quad (9)$$

基于式(9), 优化问题(8)进一步变成:

$$\begin{aligned}\min: \quad & R(f) = C \sum_{k=1}^N \xi_k + \sum_{i=1}^N (\alpha_k^+ + \alpha_k^-) \\ \text{s.t.} \quad & \begin{cases} y_k - \sum_{k=1}^m (\alpha_k^+ - \alpha_k^-) \exp\left\{\frac{-\|\mathbf{x} - \mathbf{x}_k\|^2}{2\sigma^2}\right\} - b \leq \varepsilon + \xi_k \\ \sum_{k=1}^m (\alpha_k^+ - \alpha_k^-) \exp\left\{\frac{-\|\mathbf{x} - \mathbf{x}_k\|^2}{2\sigma^2}\right\} + b - y_k \leq \varepsilon + \xi_k \\ \xi_k \geq 0 \end{cases}\end{aligned}\quad (10)$$

$$\text{现定义向量 } \mathbf{c} = \left(\underbrace{1, 1, \dots, 1}_N, \underbrace{1, 1, \dots, 1}_N, \underbrace{C, C, \dots, C}_N \right)^T,$$

$$\text{向量 } \beta \text{ 的 } L_1 \text{ 范数 } \|\beta\|_1 = \left(\underbrace{1, 1, \dots, 1}_N, \underbrace{1, 1, \dots, 1}_N \right) \begin{pmatrix} \alpha^+ \\ \alpha^- \end{pmatrix},$$

$\alpha^+ = (\alpha_1^+, \alpha_2^+, \dots, \alpha_N^+)^T$, $\alpha^- = (\alpha_1^-, \alpha_2^-, \dots, \alpha_N^-)^T$ 。以向量形式将优化问题(10)构造为线性规划问题如下:

$$\begin{aligned}\min \quad & \mathbf{c}^T \begin{pmatrix} \alpha^+ \\ \alpha^- \\ \xi \end{pmatrix} \\ \text{s.t.} \quad & \begin{cases} \begin{pmatrix} K & -K & -I \\ -K & K & -I \end{pmatrix} \cdot \begin{pmatrix} \alpha^+ \\ \alpha^- \\ \xi \end{pmatrix} \leq \begin{pmatrix} \mathbf{y} + \varepsilon \\ \varepsilon - \mathbf{y} \end{pmatrix} \\ \alpha^+, \alpha^- \geq 0, \xi \geq 0 \end{cases}\end{aligned}\quad (11)$$

其中 $\xi = (\xi_1, \xi_2, \dots, \xi_N)^T$, I 为 $N \times N$ 的单位矩阵, $\mathbf{y} = (y_1, y_2, \dots, y_N)^T$,

$$K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left\{\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right\}$$

线性规划问题(11)可通过单纯型算法或原-对偶内点算法进行求解^[22]。对于二次规划-SVR(QP-SVR), 在 ε 域之外的所有数据点将被选择为 SVs; 而对于线性规划-支持向量回归 (linear programming-support vector regression, LP-SVR), 即便 ε 域选择为 0 时, 由于软约束在优化问题中的使用, LP-SVR 仍然能够获取稀疏解。通常情况下, 稀疏解往往通过设定非零的 ε 域来获取。

2 基于 ℓ_1 范数的回归模型辨识

基于第1节介绍的支持向量回归及优化问题转化的基础上, 该部分将讨论模型参数估计的另一种方法, 即使用 ℓ_1 范数作为建模误差的评判标准。假设通过传感器或数据获取设备一组测量数据 $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_N, y_N)\}$, 其中 $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ 描述输入测量数据, 对应的输出定义为 $\{y_1, y_2, \dots, y_N\}$ 。设测量满足如下非线性系统模型:

$$y_k = g(\mathbf{x}_k), \quad k = 1, 2, \dots, N$$

根据统计学理论可知^[12], 存在以式(6)描述

的非线性回归模型 f 对测量模型 g 的任意逼近, 当逼近精度越小时, 需要的支持向量越少; 反之, 逼近精度越高, 则支持向量越多。因此, 对任意给定的实连续函数 g 及 $\eta > 0$, 存在如下回归模型 f 满足:

$$\sup_{\mathbf{x}_k \in S} |f(\mathbf{x}_k) - g(\mathbf{x}_k)| < \eta \quad \forall k$$

值得指出的是, 较小的 η 值, 对应式(11)较多的支持向量。现讨论回归模型, 式(11)的另一种参数求解方法。在非线形系统模型的逼近情况下, 定义实际输出与由式(6)定义的 SVR 模型输出之间的偏差 e_k :

$$e_k = y_k - f(\mathbf{x}_k) \quad \forall k \quad (12)$$

为了估计 SVR 模型的最优参数, 考虑所有建模误差的最小化:

$$\min_{\mathbf{x}_k \in Z} |y_k - f(\mathbf{x}_k)| \quad \forall k \quad (13)$$

Z 表示整个输入数据集。显然, 这是一个最小(min)优化问题。在式(6)描述的回归模型情况下, 式(12)的最小化可通过两个阶段完成: 1) 核函数中的核宽度 σ 的参数寻优, 通常采用经典的交叉验证或其他方法来实现, 其详细过程在本文中不再讨论; 2) 式(6)的参数确定可通过 min 优化问题求解, 即

$$\beta = \arg \min_{\beta, \mathbf{x}_k \in Z} \left| y_k - \sum_{i=1}^N \beta_i \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_k\|^2}{2\sigma^2}\right) - b \right|$$

3 最优下边界回归模型辨识

假定不确定非线性函数或非线性系统属于函数簇 Γ :

$$\Gamma = \{g: S \rightarrow R^1 | g(\mathbf{x}) = g_{\text{nom}}(\mathbf{x}) + \Delta g(\mathbf{x}), \mathbf{x} \in S\}$$

g_{nom} 为标称函数, 不确定性 $\Delta g(\mathbf{z})$ 满足 $\sup_{\mathbf{x} \in S} |\Delta g(\mathbf{x})| \leq \gamma$, $\gamma \in R$ 。现考虑来自函数簇 Γ 的成员函数 g , $\mathbf{x} \in R^d$, 对应输入 \mathbf{x} 上的测量输出 $\mathbf{Y} = \{y_1, y_1, \dots, y_N\}$, 即 $y_k = g(\mathbf{x}_k)$, $g \in \Gamma$, $\mathbf{x}_k \in S$, $k = 1, 2, \dots, N$ 。LBRM 建模的思想是, 在满足如下约束条件(14)的条件下, 建模下界回归模型 $f(\mathbf{x}_k)$:

$$f(\mathbf{x}_k) \leq g(\mathbf{x}_k) \quad \forall \mathbf{x}_k \in S \quad (14)$$

在式(14)约束的意义下, 来自函数簇的任一成员函数总能在 LBRM 上方中找到。显然, 这样的 LBRM 有无穷多个, 本文的目的就是根据提出的约束(14), 确定尽可能逼近成员函数的下界。为了确定 LBRM 的最优逼近, 提出的方法将逼近误差的 ℓ_1 范数作为模型辨识精度的保证, 而基于 SVM 的结构风险最小化理论用于提高模型的稀疏特性。由式(6)给出下界回归模型的表达式:

$$f(\mathbf{x}, \beta, b) = \sum_{k=1}^N \beta_k \exp\left(\frac{-\|\mathbf{x} - \mathbf{x}_k\|^2}{2\sigma^2}\right) + b$$

下界回归模型 $f(\mathbf{x})$ 可通过线性规划对如下优化问题进行求解:

$$\min_{f, \mathbf{x}_k \in S} \sum_{k=1}^N (y_k - f(\mathbf{x}_k)) \quad \text{s.t.} \quad y_k - f(\mathbf{x}_k) \geq 0 \quad (15)$$

因此,模型 $f(\mathbf{x})$ 的参数 β 、 b 的求解,对应 min 优化问题 (15) 可通过最小化 λ , 且满足如下不等式约束的线性规划求解, 即

$$\begin{aligned} \min: \quad & \lambda = \sum_{k=1}^N \lambda_k \\ & \begin{cases} y_k - \sum_{i=1}^N \beta_i \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_k\|^2}{2\sigma^2}\right) - b \leq \lambda_k, & k = 1, 2, \dots, N \\ y_k - \sum_{i=1}^N \beta_i \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_k\|^2}{2\sigma^2}\right) - \beta_i \geq 0 & k = 1, 2, \dots, N \\ \lambda_k \geq 0 \end{cases} \end{aligned} \quad (16)$$

其中 λ_k 表示逼近误差。

证明 上述定理 2 直接通过定理 1 推出。

从上述回归模型辨识的思想来看, 仅考虑上边模型输出与实际输出之间的逼近误差, 而回归模型本身的结构复杂性却没有被考虑, 这样一来, 通过上述优化问题获取的参数解有可能出现不全为零的情况, 不具有稀疏特性, 对应 N 个样本数据可能都是支持向量, 导致模型结构复杂。为了解决模型稀疏解的问题, 在求解下边回归模型的优化问题中, 有必要将结构风险最小化的思想融合其中, 在保证回归模型逼近精度的同时, 尽可能让模型结构复杂性得到有效控制。基于此, 将下界回归模型优化问题 (16)(式 (16)), 融合到基于结构风险最小化的优化问题 (10)(式 (10))。因此, 对于下界回归模型 $f(\mathbf{x})$ 的优化问题, 有

$$\begin{aligned} \min: \quad & C \sum_{k=1}^N \xi_k + \sum_{i=1}^N (\alpha_k^+ + \alpha_k^-) + \sum_{k=1}^N \lambda_k + b \\ & \begin{cases} \sum_{k=1}^m (\alpha_k^+ - \alpha_k^-) \exp\left\{\frac{-\|\mathbf{x} - \mathbf{x}_k\|^2}{2\sigma^2}\right\} + b - y_k \leq \varepsilon + \xi_k, \\ y_k - \sum_{k=1}^m (\alpha_k^+ - \alpha_k^-) \exp\left\{\frac{-\|\mathbf{x} - \mathbf{x}_k\|^2}{2\sigma^2}\right\} - b \leq \varepsilon + \xi_k, \\ y_k - \sum_{i=1}^N (\alpha_k^+ - \alpha_k^-) \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_k\|^2}{2\sigma^2}\right) - b \leq \lambda_k, \\ \sum_{i=1}^N (\alpha_k^+ - \alpha_k^-) \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_k\|^2}{2\sigma^2}\right) + b - y_k \leq 0, \\ \xi_k \geq 0, \lambda_k \geq 0, k = 1, 2, \dots, N \end{cases} \end{aligned} \quad (17)$$

式中: λ_k 表示最大逼近误差; 参数 α_k^+ 、 α_k^- 、 b 、 ε 、 ξ_k 与第 2 节的定义一样。

从优化问题 (17) 可知, 为典型的线性规划问

题, 可用向量及矩阵形式表述如下:

$$\begin{aligned} \min \quad & \mathbf{c}^T \begin{pmatrix} \alpha^+ \\ \alpha^- \\ \xi \\ \lambda \\ b \end{pmatrix} \\ \text{s.t.} \quad & \begin{cases} \begin{pmatrix} \mathbf{K} & -\mathbf{K} & -\mathbf{I} & \mathbf{Z} & \mathbf{E} \\ -\mathbf{K} & \mathbf{K} & -\mathbf{I} & \mathbf{Z} & \mathbf{E} \\ -\mathbf{K} & \mathbf{K} & \mathbf{Z} & -\mathbf{I} & \mathbf{E} \\ \mathbf{K} & -\mathbf{K} & \mathbf{Z} & \mathbf{Z} & \mathbf{E} \end{pmatrix} \cdot \begin{pmatrix} \alpha^+ \\ \alpha^- \\ \xi \\ \lambda \\ b \end{pmatrix} \leq \begin{pmatrix} \mathbf{y} + \varepsilon \\ \varepsilon - \mathbf{y} \\ -\mathbf{y} \\ \mathbf{y} \end{pmatrix} \\ \alpha^+, \alpha^- \geq 0, \xi \geq 0, 0 \leq \lambda_k \leq 1 \end{cases} \end{aligned} \quad (18)$$

其中, $\mathbf{c} = \left(\underbrace{1, 1, \dots, 1}_N, \underbrace{1, 1, \dots, 1}_N, \underbrace{C, C, \dots, C}_N, 1, 1 \right)^T$, $\mathbf{y} = (y_1, y_2, \dots, y_N)^T$, $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_N)^T$, $\alpha_U^+ = (\alpha_1^+, \alpha_2^+, \dots, \alpha_N^+)^T$, $\alpha_U^- = (\alpha_1^-, \alpha_2^-, \dots, \alpha_N^-)^T$, $\xi = (\xi_1, \xi_2, \dots, \xi_N)^T$, $\mathbf{Z} = \mathbf{0}_{N \times N}$, \mathbf{I} 为 $N \times N$ 单位矩阵, $\mathbf{E} = \mathbf{1}_{N \times 1}$, 核矩阵 \mathbf{K} 的元素定义为

$$K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left\{\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right\}, \sigma \text{ 为可调核参数。}$$

显然, 应用内点法或单纯性方法可以求解优化问题 (18), 进而得到下界回归模型 $f(\mathbf{x})$:

$$f(\mathbf{x}) = \sum_{k=1}^N (\alpha_k^+ - \alpha_k^-) \exp\left(\frac{-\|\mathbf{x} - \mathbf{x}_k\|^2}{2\sigma^2}\right) + b \quad (19)$$

从应用提出方法来建立 $f(\mathbf{x})$ 的整个过程来看, 优化问题既包括了对模型结构复杂性控制的目标函数, 又包括了如何获取较好的模型精度所对应的逼近误差作为目标函数, 而且模型结构复杂性控制和模型精度之间的权衡可以通过规则化参数进行调整。总而言之, 提出方法在保证获取下界模型建模精度的同时, 而且还对模型结构复杂性进行有效控制, 从而提高下界回归模型的泛化性能。

4 实验分析

将通过如下实验分析, 论证所提出方法的最优性与稀疏性; 同时为了更直观地去评判提出的方法, 将考虑如下两个性能指标, 即均方根误差 (root man square error, RMSE) 和支持向量占整个样本数据的百分比SVs%。RMSE定义为

$$\text{RMSE} = \frac{1}{N} \sqrt{\sum_{k=1}^N (y_k - \hat{y}_k)^2}$$

式中: N 表示测试数据的总数; y_k 为实际输出; \hat{y}_k 是模型的被估输出。RMSE反映了用提出方法所建立下界回归模型 (LBRM) 在满足其约束条件下, 即 $y_k - f(\mathbf{x}_k) \geq 0$, 模型输出与实际测量数据之间的逼近程度; RMSE越小, 逼近程度越好, 反之越差。此外, 对应优化问题 (18) 的求解, 若有

$(\alpha_k^+ - \alpha_k^-) \neq 0$ 或 $(\alpha_k^+ - \alpha_k^-) \geq \eta$ (其中在确定样本数据是否为支持向量的过程中, 选择 η 为 1×10^{-11}), 对应第 k 个样本数据为支持向量, 通过条件判断假设共有 N_k 个支持向量, 则 SVs% 的定义如下:

$$\text{SVs\%} = \frac{N_k}{N} \times 100\%$$

显然, 在保证下界模型建模精度的同时, 指标 SVs% 越小越好, 越小则表示求解的下界回归模型有稀疏解, 模型结构简单, 说明具有较好的泛化性能。

接下来将对提出的方法从下边界回归模型的辨识精度以及稀疏特性展开实验分析。当被建模的非线性系统由噪声引起的不确定性输出时, 论证带稀疏特性的最优下边界回归模型辨识。

先考虑如下的非线性动态系统:

$$\begin{aligned} y(t+1) &= \frac{y(t)y(t-1)[y(t)+2.5]}{1+y^2(t)+y^2(t-1)} + u(t) + \text{noise} \\ y(0) &= y(1) = 0, \quad u(t) = \sin(2\pi t/50) \end{aligned} \quad (20)$$

其中, noise 是均值为 0、方差为 0.25 的高斯噪声。基于式 (20) 获取 201 个建模数据。

LBRM 的最优性, 除了应用提出方法在辨识精度与稀疏特性之间取其平衡得以体现之外, 超参数集的选取对 LBRM 的稀疏特性也起着至关重要的作用。在实验分析中, 超参数集的 4 种取值主要是基于 SVR 方法的经验来获取^[23], 其中不敏感域 ε 的取值一般在区间 $[0, 1]$ 之间获取, 正则化参数 γ 一般选取为 $2^n, n = -5, -4, \dots, 15$, 核参数 σ 一般从区间 $[0, 10]$ 获取。当超参数集 $(\varepsilon, \gamma, \sigma)$ 选择为 $(0.001, 1000, 5.0)$ 时, 应用提出方法获取的最优下边界回归模型 (LBRM) 如图 1 所示。

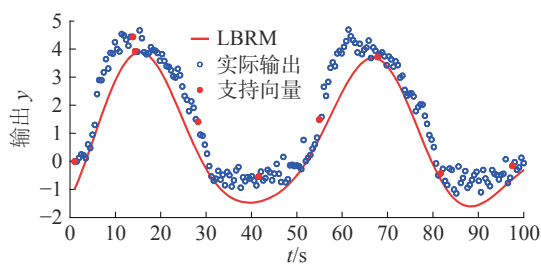


图 1 提出方法所建立的最优下边界回归模型 (核宽度为 5.0)
Fig. 1 Optimal LBRM constructed by our approach, where $\sigma=5.0$

可知, 应用提出方法所建立的最优 LBRM 仅仅需要 9 个支持向量, 即从这 201 个数据中, 建立 LBRM 只用到了其中的 9 个数据, 表明稀疏特性较好, 对应的指标 SVs% 为 4.48%; 辨识精度 RMSE 为 0.814 0。图 2 给出了 LBRM 所对应的逼近误差, 用 $f(x)$ 表示 LBRM, y 表示测量数据, 则有 $f(x) - y \leq 0$, 进一步表明所有的不确定性测量数据都在 LBRM 的上方。当超参数集 $(\varepsilon, \gamma, \sigma)$ 选择为

$(0.0001, 1000, 0.2)$ 时, 如图 3 所示, 提出方法所获取的 LBRM 拟合了所有的不确定性数据, 对应的辨识精度 RMSE 达到了 2.0000×10^{-5} , 对应的 SVs% 为 99.50%, 产生了过拟合, 失去了模型的稀疏特性, 对应的逼近误差如图 4 所示。为了进一步清晰地分析 LBRM 稀疏特性与辨识进度之间的平衡, 表 1 给出了在不同核宽度 σ 下的 SVs% 和 RMSE, 可以发现, LBRM 随核宽度 σ 的增加, SVs% 在逐渐减小, 表明建立 LBRM 所用到的 SVs 个数减小, 模型结构简单, 对应较好的稀疏特性; 相反, 用于反映 LBRM 辨识精度的 RMSE 在增加, 表明模型的辨识精度降低。因此, 反映稀疏特性的 SVs% 和反映模型辨识精度的 RMSE 之间是一对矛盾体, 在核宽度 σ 的选取上, 应从建模的需要从两者之间取其平衡, 不能一味地去追求某个指标, 例如当 $\sigma=0.1$ 时, RMSE 为 4.56×10^{-6} , 表明辨识精度很高, 但 SVs% 达到了 99.00%, 已经完全丧失了模型的稀疏特性, 泛化性能变差, 容易产生过拟合问题。

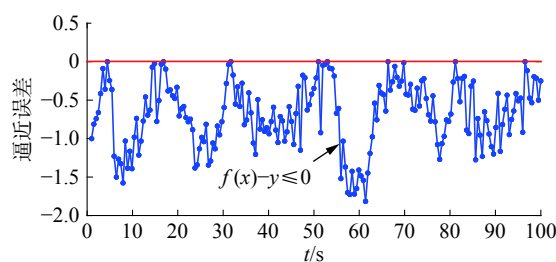


图 2 提出方法所对应的逼近误差
Fig. 2 Approximation error of the proposed method

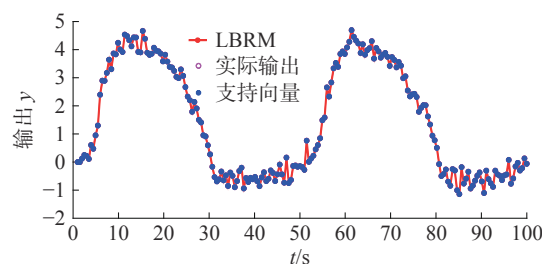


图 3 提出方法所建立的最优下边界回归模型 (核宽度为 0.1)
Fig. 3 Optimal LBRM constructed by our approach, where $\sigma=0.1$

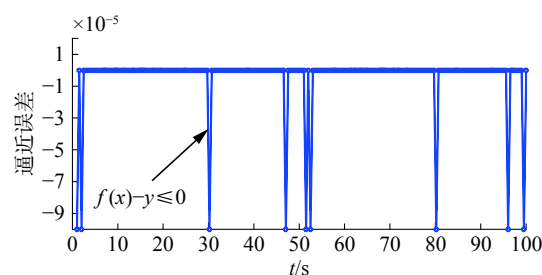


图 4 过拟合所对应的逼近误差
Fig. 4 Approximation error of the proposed method when the over-fitting appeared.

表1 当超参数 $\varepsilon = 0.0001$, $\gamma = 1000$ 时, 在不同核宽度 σ 情况下的SVs%和RMSE

Table 1 Comparison result between SVs% and RMSE when the different kernel width σ is selected for $\varepsilon = 0.0001$ and $\gamma = 1000$

σ	SVs%	RMSE	σ	SVs%	RMSE
0.1	0.9900	4.56×10^{-6}	3.5	0.2139	0.3444
0.2	0.9950	2.00×10^{-4}	4.0	0.1741	0.3869
0.7	0.8607	0.1352	4.5	0.1542	0.3880
0.8	0.7313	0.1615	5.0	0.1493	0.3939
0.9	0.6667	0.1713	5.5	0.1393	0.4141
1.0	0.6418	0.1782	6.0	0.1294	0.4242
1.5	0.4279	0.2750	6.5	0.1144	0.4408
2.0	0.5075	0.2995	7.0	0.1194	0.4549
2.5	0.3035	0.3328	7.5	0.1144	0.4913
3.0	0.2338	0.3356	8.0	0.1542	0.5293

接下来, 考虑由模型结构参数的变化所引起的, 不确定性输出的 LBRM 辨识。描述的不确定非线性系统为

$$\begin{aligned} f_{\text{norm}}(\mathbf{x}) &= \cos \mathbf{x} \sin \mathbf{x} \\ \Delta f(\mathbf{x}) &= \tau \cos(8\mathbf{x}) \\ g(\mathbf{x}) &= f_{\text{norm}}(\mathbf{x}) + \Delta f(\mathbf{x}) \end{aligned} \quad (21)$$

式中: $g(\mathbf{x})$ 由名义函数 $f_{\text{norm}}(\mathbf{x})$ 和不确定性 $\Delta f(\mathbf{x})$ 组成 $0 \leq \tau \leq 1$, 设该函数类的定义域为 $-1 \leq \mathbf{x} \leq 1$ 。提出的方法是基于数据建立的 LBRM, 因此获取式 (21) 的数据是建立 LBRM 的基础。为了更有效地反映提出方法辨识由参数不确定性的 LBRM, 首先获取所需要的样本数据, 不妨取 $\mathbf{x}_k = 0.021k$, $k = -47, -46, \dots, 47$, 图 5 表示由不确定性参数 $\tau (0 \leq \tau \leq 1)$ 所引起的测量输出, 共有 5 组测量数据, 对应 τ 分别取 $\{0, 0.2, 0.4, 0.6, 0.8\}$ 。接下来, 用提出的方法建立这 5 组不确定性测量的最优下边界回归模型。

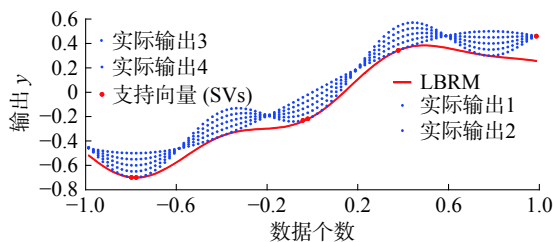


图5 提出方法所建立的最优下边界回归模型 (核宽度为 10.5)

Fig. 5 Optimal LBRM constructed by our approach, where $\sigma=10.5$

当超参数集 $(\varepsilon, \gamma, \sigma)$ 选择为 $(0.1, 0.01, 10.5)$ 时, 求解的 LBRM 如图 5 所示, 从 95 个数据建立最

优 LBRM 仅用了 8 个数据, 即 8 个支持向量, 如图 6 所示, 未标出来的纵坐标 $\alpha_k^+ - \alpha_k^-$ 均不大于 10^{-11} , 表示第 k 个数据对 LBRM 的贡献可忽略不计, 图中只标出对 LBRM 起作用的 $\alpha_k^+ - \alpha_k^-$ 值, 对应的稀疏特性较好, 相应的 SVs% 指标为 8.42%; LBRM 辨识精度对应的 RMSE 为 0.0502。表 2 给出了 $\alpha_k^+ - \alpha_k^- \geq 10^{-11}$ 时所对应的支持向量, 从表 2 可知, 满足此条件的 $\alpha_k^+ - \alpha_k^-$ 的第 k 个数据分别出现在 $k = 9, 10, 28, 46, 47, 68, 69, 95$, 表明从 95 个不确定性数据中, 建立 LBRM 仅仅用了其中的 8 个数据。

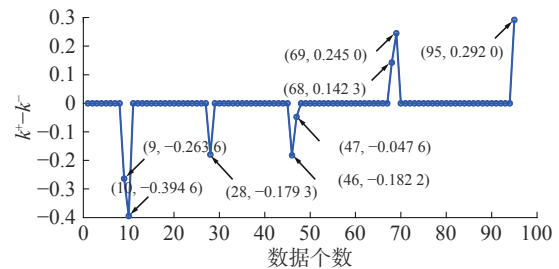


图6 第 k 个支持向量所对应的 $\alpha_k^+ - \alpha_k^-$ 值 ($\alpha_k^+ - \alpha_k^- \geq 10^{-11}$)

Fig. 6 The k -th support vector (SV) corresponding to the values of $\alpha_k^+ - \alpha_k^-$ ($\alpha_k^+ - \alpha_k^- \geq 10^{-11}$)

表2 第 k 个支持向量 (SV) 对应的 $\alpha_k^+ - \alpha_k^-$ 值

Table 2 The k -th support vector (SV) corresponding to the values of $\alpha_k^+ - \alpha_k^-$

第 k 个 SV	$\alpha_k^+ - \alpha_k^-$	第 k 个 SV	$\alpha_k^+ - \alpha_k^-$
9	-0.26357	47	-0.04762
10	-0.39459	68	0.142301
28	-0.17931	69	0.244983
46	-0.18217	95	0.292009

基于式 (19) 可列出 LBRM 的数学模型为

$$\begin{aligned} f(\mathbf{x}) &= -0.26 \cdot \exp\left(\frac{-\|\mathbf{x} + 0.69169\|^2}{2 \times 10.5^2}\right) - \\ &0.40 \cdot \exp\left(\frac{-\|\mathbf{x} + 0.70\|^2}{2 \times 10.5^2}\right) - 0.18 \cdot \exp\left(\frac{-\|\mathbf{x} + 0.37\|^2}{2 \times 10.5^2}\right) - \\ &0.18 \cdot \exp\left(\frac{-\|\mathbf{x} + 0.23\|^2}{2 \times 10.5^2}\right) - 0.05 \cdot \exp\left(\frac{-\|\mathbf{x} + 0.22\|^2}{2 \times 10.5^2}\right) + \\ &0.14 \cdot \exp\left(\frac{-\|\mathbf{x} + 0.37\|^2}{2 \times 10.5^2}\right) + 0.25 \cdot \exp\left(\frac{-\|\mathbf{x} + 0.39\|^2}{2 \times 10.5^2}\right) + \\ &0.29 \cdot \exp\left(\frac{-\|\mathbf{x} + 0.46\|^2}{2 \times 10.5^2}\right) \end{aligned}$$

当超参数集 $(\varepsilon, \gamma, \sigma)$ 选择为 $(0.001, 10, 4.5)$ 时, 其 LBRM 如图 7 所示, 显然模型的辨识精度得到显著提高, 对应 RMSE 为 0.0059, 但所需要的支持向量为 21 个, 描述稀疏特性的 SVs% 指标为 22.11%。为了更清晰地分析 LBRM 之间的稀疏特性和辨识精度, 表 3 列出了不同核宽度 σ 下的

SVs%和RMSE, 从中可以发现, 模型的辨识精度与稀疏特性之间是一对矛盾体, 虽然稀疏特性较好时, 模型结构简单, 泛化性能也较好, 但辨识精度进一步下降。因此, 根据工程应用的需要, 应从两者之间取其平衡。

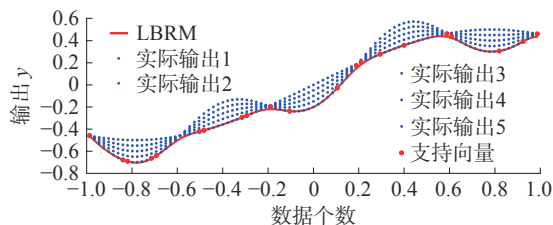


图7 提出方法所建立的最优下边界回归模型(核宽度为4.5)

Fig. 7 Optimal LBRM constructed by our approach, where $\sigma=4.5$

表3 当超参数 $\varepsilon = 0.0001$, $\gamma = 10$ 时, 不同核宽度 σ 下的SVs%和RMSE

Table 3 Comparison result between SVs% and RMSE when the different kernel width σ is selected for $\varepsilon = 0.0001$ and $\gamma = 10$

σ	SVs%	RMSE	σ	SVs%	RMSE
0.5	1.0000	1.03×10^{-5}	5.0	0.2211	0.0078
1.0	0.9895	1.11×10^{-5}	5.5	0.2105	0.0098
1.5	0.7895	0.0029	6.0	0.2	0.0111
2.0	0.6000	0.0037	6.5	0.1895	0.0113
2.5	0.4632	0.0036	7.0	0.1684	0.0111
3.0	0.3895	0.0054	7.5	0.1579	0.0111
3.5	0.3158	0.0052	8.0	0.1579	0.0115
4.0	0.2947	0.0058	8.5	0.1474	0.0113
4.5	0.2211	0.0059	9.0	0.1386	0.0123

5 结束语

基于数据的传统辨识方法主要从模型辨识精度进行研究, 是一种确定性建模方法, 对应点输出, 同时易产生较复杂的模型结构, 致使模型的泛化性能变差。从模型结构参数以及测量数据的不确定性出发, 本文研究了由不确定性引起的下边界回归模型辨识方法, 具有如下显著特点: 1) 建立了不确定性的边界输出, 提高建模的鲁棒性; 2) 边界模型结构可通过引入的结构风险最小化原理进行调整, 提高其泛化性能; 3) 提出方法的最优性体现在模型结构与辨识精度之间的平衡, 即在保证模型稀疏特性的情况下, 尽可能提高模型辨识精度, 该方法可以应用到信息压缩、故障检测等。

参考文献:

- [1] HAN Honggui, GUO Yanan, QIAO Junfei. Nonlinear system modeling using a self-organizing recurrent radial basis function neural network[J]. *Applied soft computing*, 2018, 71: 1105–1116.
- [2] FRAVOLINI M L, NAPOLITANO M R, DEL CORE G, et al. Experimental interval models for the robust fault detection of aircraft air data sensors[J]. *Control engineering practice*, 2018, 78: 196–212.
- [3] FANG Shengen, ZHANG Qiuhe, REN Weixin. An interval model updating strategy using interval response surface models[J]. *Mechanical systems and signal processing*, 2015, 60–61: 909–927.
- [4] LEUNG F H F, LAM H K, LING S H, et al. Tuning of the structure and parameters of a neural network using an improved genetic algorithm[J]. *IEEE transactions on neural networks*, 2003, 14(1): 79–88.
- [5] 赵文清, 严海, 王晓辉. BP神经网络和支持向量机相结合的电容器介损角辨识[J]. *智能系统学报*, 2019, 14(1): 134–140.
ZHAO Wenqing, YAN Hai, WANG Xiaohui. Capacitor dielectric loss angle identification based on a BP neural network and SVM[J]. *CAAI transactions on intelligent systems*, 2019, 14(1): 134–140.
- [6] 刘道华, 张礼涛, 曾召霞, 等. 基于正交最小二乘法的径向基神经网络模型[J]. *信阳师范学院学报(自然科学版)*, 2013, 26(3): 428–431.
LIU Daohua, ZHANG Litao, ZENG Zhaoxia, et al. Radial basis function neural network model based on orthogonal least squares[J]. *Journal of Xinyang Normal University (natural science edition)*, 2013, 26(3): 428–431.
- [7] 刘道华, 张飞, 张言言. 一种改进的RBF神经网络对县级政府编制预测[J]. *信阳师范学院学报(自然科学版)*, 2016, 29(2): 265–269.
LIU Daohua, ZHANG Fei, ZHANG Yanyan. A prediction for the preparation of county government based on improved RBF neural networks[J]. *Journal of Xinyang Normal University (natural science edition)*, 2016, 29(2): 265–269.
- [8] HAN Honggui, GE Luming, QIAO Junfei. An adaptive second order fuzzy neural network for nonlinear system modeling[J]. *Neurocomputing*, 2016, 2014: 837–847.
- [9] LI Fanjun, QIAO Junfei, HAN Honggui, et al. A self-organizing cascade neural network with random weights for nonlinear system modeling[J]. *Applied soft computing*, 2016, 42: 184–193.

- [10] HAN Honggui, QIAO Junfei. Hierarchical neural network modeling approach to predict sludge volume index of wastewater treatment process[J]. *IEEE transactions on control systems technology*, 2013, 21(6): 2423–2431.
- [11] VAPNIK V N. An overview of statistical learning theory[J]. *IEEE transactions on neural networks*, 1999, 10(5): 988–999.
- [12] 唐波, 彭友仙, 陈彬, 等. 基于 BP 神经网络的交流输电线路可听噪声预测模型 [J]. 信阳师范学院学报(自然科学版), 2015, 28(1): 136–140.
TANG Bo, PENG Youxian, CHEN Bin, et al. Audible noise prediction model of ac power lines based on BP neural network[J]. *Journal of Xinyang Normal University (natural science edition)*, 2015, 28(1): 136–140.
- [13] HAO Peiyi. Interval regression analysis using support vector networks[J]. *Fuzzy sets and systems*, 2009, 160(17): 2466–2485.
- [14] HAO Peiyi. Possibilistic regression analysis by support vector machine[C]//Proceedings of 2011 IEEE International Conference on Fuzzy Systems. Taipei, China, 2011: 889–894.
- [15] 石磊, 侯丽萍. 基于改进 PSO 算法参数优化的模糊支持向量分类机 [J]. 信阳师范学院学报(自然科学版), 2013, 26(2): 288–291.
SHI Lei, HOU Liping. Parameter optimization of fuzzy support vector classifiers based on the improved PSO[J]. *Journal of Xinyang Normal University (natural science edition)*, 2013, 26(2): 288–291.
- [16] SOUZA L C, SOUZA R M C R, AMARAL G J A, et al. A parametrized approach for linear regression of interval data[J]. *Knowledge-based systems*, 2017, 131: 149–159.
- [17] DE A LIMA NETO E, DE A T DE CARVALHO F. An exponential-type kernel robust regression model for interval-valued variables[J]. *Information sciences*, 2018, 454–455: 419–442.
- [18] BASAK D, PAL S, PATRANABIS D C. Support vector regression[J]. *Neural information processing-letters and reviews*, 2007, 11(10): 203–224.
- [19] SMOLA A J, SCHÖLKOPF B. A tutorial on support vector regression[J]. *Statistics and computing*, 2004, 14(3): 199–222.
- [20] SOARES Y M G, FAGUNDES R A A. Interval quantile regression models based on swarm intelligence[J]. *Applied soft computing*, 2018, 72: 474–485.
- [21] LU Zhao, SUN Jing, BUTTS K. Linear programming SVM-ARMA2K with application in engine system identification[J]. *IEEE transactions on automation science and engineering*, 2011, 8(4): 846–854.
- [22] RIVAS-PEREA P, COTA-RUIZ J. An algorithm for training a large scale support vector machine for regression based on linear programming and decomposition methods[J]. *Pattern recognition letters*, 2013, 34(4): 439–451.
- [23] HSU C W, CHANG C C, LIN C J. A practical guide to support vector classification. Technical report. Department of Computer Science, National Taiwan University (2003)[EB/OL](2010-4-15). <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>.

作者简介:



刘小雍, 副教授, 博士, 主要研究方向为机器学习与人工智能。发表学术论文 10 余篇。



叶振环, 教授, 博士, 主要研究方向为动态系统故障诊断与容错控制、状态估计。发表学术论文 20 余篇。