

基于迭代误差补偿的混沌时间序列最小二乘支持向量机预测算法*

唐舟进[†] 任峰 彭涛 王文博

(北京邮电大学信息与通信工程学院, 北京 100876)

(2013年10月24日收到; 2013年11月26日收到修改稿)

本文分析了传统支持向量机预测算法产生的误差特性, 发现产生的预测误差不同于噪声, 具有较强的规律性, 单一的预测模型遗漏了许多混沌序列中的确定性分量. 经过误差补偿后, 残差的冗余信息减少, 随机性增强. 在此基础上, 本文提出一种基于迭代误差补偿的最小二乘支持向量机预测算法, 能够通过多模型联合预测更加有效地逼近混沌系统的映射函数, 在预测精度上取得了大幅度的提升. 此外, 算法通过留一交叉验证法的方法能够在预测前自动优化模型参数组合, 克服了现有算法无法仅利用先验信息优化预测模型参数的缺陷. 对 MackeyGlass 和 Lorenz 混沌时间序列进行了仿真实验, 实验结果优于相关文献记载方法的预测性能, 在性能指标上好于现有算法一个数量级.

关键词: 混沌时间序列预测, 最小二乘支持向量机, 迭代误差补偿, 参数组合优化

PACS: 05.45.Tp, 05.45.Gg

DOI: 10.7498/aps.63.050505

1 引言

从1927年 Yule 通过自回归技术预测了太阳黑子数开始, 研究者对混沌时间序列的预测已经进行了多年的研究. 传统的预测都是基于线性的预测模型, 而自然界中广泛存在的时间序列通常都是混沌时间序列, 同时具有线性特征和非线性特征, 非线性特征很难通过线性预测模型来描述. 提高混沌时间序列的预测精度对于多个领域的发展都具有重要意义.

目前关于混沌时间序列的研究主要集中于人工神经网络模型^[1-10]和支持向量机模型^[11-14]. 人工神经网络具有以任意精度逼近非线性函数的特征, 但传统神经网络模型收敛速度慢, 容易过拟合. 近年已有学者提出了回声状态网络, 简单快速, 其改进算法也引入了结构风险最小化原理, 在混沌时间序列的预测中取得了较好的效果^[2,9,10]. 文献^[2]利用正则化回声状态网络预测模型进行预测误

差补偿取得较好的预测精度. 文献^[9]在回声状态网络预测模型的基础上引入了支持向量机的结构风险最小化原理, 在含噪混沌时间序列预测上取得了较好的效果. 文献^[10]比较了通过回声状态网络进行建模的预测性能优势. 但神经网络始终存在网络结构过分依赖经验, 缺乏普适性的缺点, 其网络结构没有一种通用有效的设计方法.

支持向量机是统计学习理论的研究成果, 建立在 VC 维理论和结构风险最小化原理的基础之上, 同样能够以任意精度逼近非线性函数, 是一种有效的非线性建模预测的方法^[15]. 经过大量的研究, 支持向量机有多种改进算法, 其中最小二乘支持向量机在保持较高训练效率的基础上简化了算法复杂度, 具有模型简单, 收敛速度快, 可实现性强的优点^[16]. 基于上述优点, 支持向量机及其改进算法已经广泛应用于混沌时间序列预测中, 取得了较好的预测效果.

但在目前有关支持向量机的混沌时间序列预

* 中国国防科技预研项目 (批准号: 208010201) 资助的课题.

[†] 通讯作者. E-mail: tangzhoujin@gmail.com

测算法中, 还存在一些问题:

1) 相比神经网络类的预测算法, 基于支持向量机的预测算法在预测精度上没有明显优势^[9].

2) SVM 及其改进算法的参数选择目前还主要是依靠人工选择. 在文献 [11] 中有学者提出了一种基于差分进化算法的参数组合自动优化算法, 预测性能较好, 但选择最优参数组合时需要输入待预测的测试样本, 无法应用到实际的预测系统中.

本文提出了一种基于迭代误差补偿的最小二乘支持向量机混沌时间序列预测算法, 通过迭代误差补偿建立多个模型进行联合预测提高预测精度, 通过基于模式搜索的参数自动优化在先验信息仅有训练数据的情况下能够自动选择优选参数组合. Lorenz 混沌时间序列和 MackeyGlass 序列的预测仿真实验结果表明了算法的有效性, 本文所提算法相比于现有算法在预测精度上有大幅度的提高, 达到一个数量级以上.

2 最小二乘支持向量机

LSSVM 是基于 SVM 改进后的一种算法, 通过引入等式化约束和最小二乘损失函数的方法, 使最优化问题的求解变为求解线性方程, 避免了解二次规划问题, 使得算法的复杂度降低, 相比 SVM, LSSVM 的运算速度较快. 基于 LSSVM 的回归预测可以描述为:

首先对训练数据 X 重构后, 获得相空间矩阵 $S = \{\mathbf{x}_i, y_i\}_{i=1}^k$, \mathbf{x}_i 为 $m-1$ 维输入向量, y_i 为一维输出向量, k 为重构后的训练样本个数. \mathbf{x}_i 与 y_i 之间通常为非线性关系, 因此将 \mathbf{x}_i 映射到高维特征空间中, LSSVM 主要是在高维空间中对训练样本进行回归:

$$y = \boldsymbol{\omega}^T \varphi(\mathbf{x}) + b, \quad (1)$$

其中, $\varphi(\mathbf{x})$ 为非线性映射函数, $\boldsymbol{\omega}$ 为法向量, b 为偏置量, 对以上问题的求解可描述如下:

$$\min_{\boldsymbol{\omega}, b, e} J(\boldsymbol{\omega}, b, e) = \frac{1}{2} \|\boldsymbol{\omega}\|^2 + \frac{C}{2} \sum_{i=1}^k \xi_i^2,$$

$$\text{s.t. } y_i - \xi_i = \boldsymbol{\omega}^T \varphi(\mathbf{x}_i) + b, \quad i = 1, \dots, k, \quad (2)$$

C 为惩罚因子, ξ_i 为训练误差. 为求解此优化问题, 可引入 Lagrange 函数

$$L(\boldsymbol{\omega}, b, e, a)$$

$$= J(\boldsymbol{\omega}, b, e) + \sum_{i=1}^k \alpha_i [y_i - \xi_i - \boldsymbol{\omega}^T \varphi(\mathbf{x}_i) - b], \quad (3)$$

其中 $\alpha_i, i = 1, \dots, k$ 为 Lagrange 乘子, 由 KKT 条件得到如下关系式:

$$\begin{aligned} \frac{\partial L}{\partial \boldsymbol{\omega}} = 0 &\rightarrow \boldsymbol{\omega} = \sum_{i=1}^k \alpha_i \varphi(\mathbf{x}_i), \\ \frac{\partial L}{\partial \alpha_i} = 0 &\rightarrow y_i - \xi_i - \boldsymbol{\omega}^T \varphi(\mathbf{x}_i) - b = 0, \\ \frac{\partial L}{\partial b} = 0 &\rightarrow \sum_{i=1}^k \alpha_i = 0, \\ \frac{\partial L}{\partial \xi_i} = 0 &\rightarrow \alpha_i = C \xi_i. \end{aligned} \quad (4)$$

关系式的求解可化为

$$\begin{bmatrix} 0 & \mathbf{e}^T \\ \mathbf{e} & Q + C^{-1}I \end{bmatrix} \times \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix}, \quad (5)$$

其中 Q 是元素为 K_{ij} 的 $k \times k$ 阶核矩阵, $K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j) = \langle \varphi(\mathbf{x}_i), \varphi(\mathbf{x}_j) \rangle$, I 为单位矩阵, 向量 $\mathbf{e} = [1, \dots, 1]^T$, 向量 $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_k]^T$, 向量 $\mathbf{y} = [y_1, \dots, y_k]^T$.

定义 $Q_n = Q + \frac{I}{C}$, 这样我们就可以得到 $\boldsymbol{\alpha}$ 和 b 的表达式:

$$\begin{aligned} b &= \frac{\mathbf{e}^T Q_n^{-1} \mathbf{y}}{\mathbf{e}^T Q_n^{-1} \mathbf{e}}, \\ \boldsymbol{\alpha} &= Q_n^{-1} (\mathbf{y} - \mathbf{e} \times b). \end{aligned} \quad (6)$$

通过 (5) 式和 (6) 式可得到 LSSVM 的混沌时间序列回归模型为

$$\begin{aligned} f(\mathbf{x}) &= \sum_{i=1}^k \alpha_i K(\mathbf{x}_i, \mathbf{x}) + b, \\ y &= f(\mathbf{x}). \end{aligned} \quad (7)$$

对应的预测输入样本为 \mathbf{x}_p , 则得到的预测值为

$$y_p = f(\mathbf{x}_p) = \sum_{i=1}^k \alpha_i K(\mathbf{x}_i, \mathbf{x}_p) + b. \quad (8)$$

本文采用径向基核函数 (RBF) 作为核函数

$$K(\mathbf{x}_i, \mathbf{x}_p) = \exp\left(-\|\mathbf{x}_i - \mathbf{x}_p\|^2 / 2\delta^2\right), \quad (9)$$

式中 δ^2 是核函数的方差.

3 传统 LSSVM 预测误差特性分析

在文献 [11] 中, 张文专等研究了最优参数组合条件下传统 LSSVM 预测模型的预测性能, 但是最

优的参数组合需要得到待预测数据之后才能通过寻优得到, 是一种后验的优化方法, 难以在实际情况下应用. 本文通过分析这种最优 LSSVM 预测模型在预测 Lorenz 混沌序列时的误差特性来引出迭代误差补偿的研究思路. 利用获得的最优参数组合建立 LSSVM 模型按照文献 [11] 的方法对 Lorenz 预测样本进行预测, 绝对误差如图 1 所示.

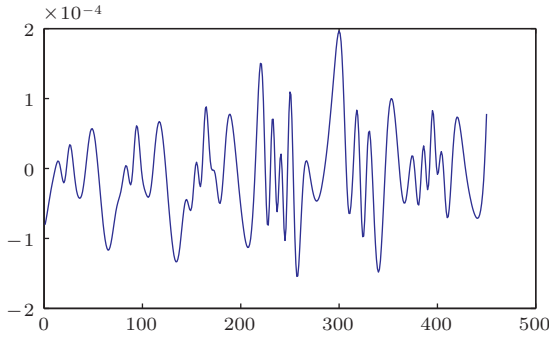


图1 Lorenz 混沌序列最优 LSSVM 模型预测绝对误差

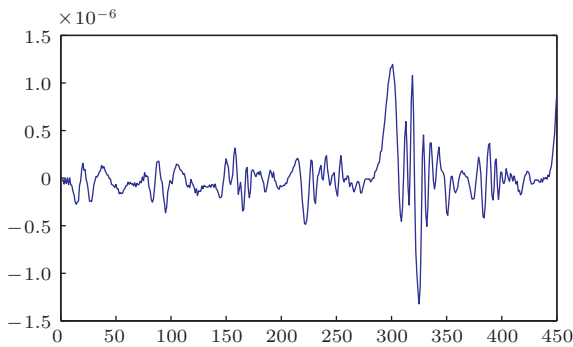


图2 经过最优 LSSVM 误差补偿后的 Lorenz 序列预测绝对误差

如果预测模型能够逼近混沌序列的映射函数, 那么产生的误差应当主要由混沌序列中的随机成分引起的, 从图 1 中可以看出传统 LSSVM 预测模型在最优条件下产生的误差仍然十分平滑, 说明预测模型遗漏了部分混沌序列中的确定性分量, 否则误差应当接近于噪声. 在文献 [2] 中, 韩敏等提出一种基于误差补偿的预测模型, 通过自回归移动平均模型预测混沌序列中的线性部分, 使残差仅有非线性成分, 再通过正则化回声状态网络模型预测序列中的非线性残差部分. 本文也引入这样的误差

补偿技术, 但不是将序列简单地划分为线性和非线性部分, 而是继续使用最优 LSSVM 预测模型预测图 1 中的绝对误差, 训练数据为上一个预测模型的拟合误差. 同样通过文献 [11] 的方法建立最优预测模型, 误差序列的预测效果如图 2 所示.

从图 2 中可以看出, 经过一次误差补偿后, 预测绝对误差小了一个数量级以上, 而且序列随机性增强, 不再平滑. 但是误差序列与随机噪声的区别仍然很大, 还是具有一定的规律性, 可以继续执行误差补偿. 那么将误差补偿一直进行下去, 直到最终的预测误差收敛, 这样的思想构成了本文所提算法的核心.

4 基于误差补偿和参数优化的预测算法具体步骤

通过上文的分析, 在理想的条件下, 可以通过最优 LSSVM 预测模型迭代进行误差补偿, 提高预测精度. 但是要使算法具有实用价值, 还需要解决传统的预测算法难以利用先验信息确定预测模型参数组合的问题.

本文提出了一种基于迭代误差补偿的混沌时间序列最小二乘支持向量机预测算法, 具体步骤如下:

首先, 将训练数据 X 按照下式归一化,

$$X_1(i) = \frac{X(i) - \min(X)}{\max(X) - \min(X)}. \quad (10)$$

得到归一化训练数据 X_1 , 其中 $\max()$ 是取最大值操作, $\min()$ 是取最小值操作. 将 X_1 重构相空间, 可得相空间矩阵

$$S_1 = \{x_1^i, X_1(i + m - 1)\}_{i=1}^{n-m+1}.$$

通过留一交叉验证法 [17] (leave-one-out cross validation) 选择验证样本

$$y_v^1 = [X_1(n - s), \dots, X_1(n - 1), X_1(n)],$$

其余部分作为训练数据, 通过前文中的 LSSVM 回归预测算法, 利用剩余的训练数据建立预测模型. 其中验证样本预测模型最优参数组合 $[m, \tau, C, \delta^2]$ 通过求解下式:

$$\min_{m, \tau, C, \delta} \sqrt{\sum_{i=1}^{s+1} (f_1'(m, \tau, C, \delta, x_1^{n-m-s+i}) - y_v^1(i))^2 / (s+1)}, \quad (11)$$

$$f'_1(m, \tau, C, \delta, \mathbf{x}_1^{n-m-s+i}) = \sum_{j=1}^{n-m-s} \alpha'_1(j) K(\mathbf{x}_1^j, \mathbf{x}_1^{n-m-s+i}) + b'_1. \quad (12)$$

由于LSSVM的黑盒特性, 难以直接求解最优参数组合, 本文选择模式搜索算法对LSSVM预测模型参数组合 $[m, \tau, C, \delta^2]$ 进行寻优, 具有准确度高, 收敛速度快的特性^[18]. 将寻找的最优预测模型预测值与验证样本的均方根误差记为 R_1 . 由于时间上的相关性, 验证样本的最优模型参数对于预测样本也是较优的选择, 由验证样本的优选参数组合 $[m_1, \tau_1, C_1, \delta_1^2]$ 和全部训练数据建立测试集的LSSVM预测模型, 由下式表达:

$$f_1(m_1, \tau_1, C_1, \delta_1, \mathbf{x})$$

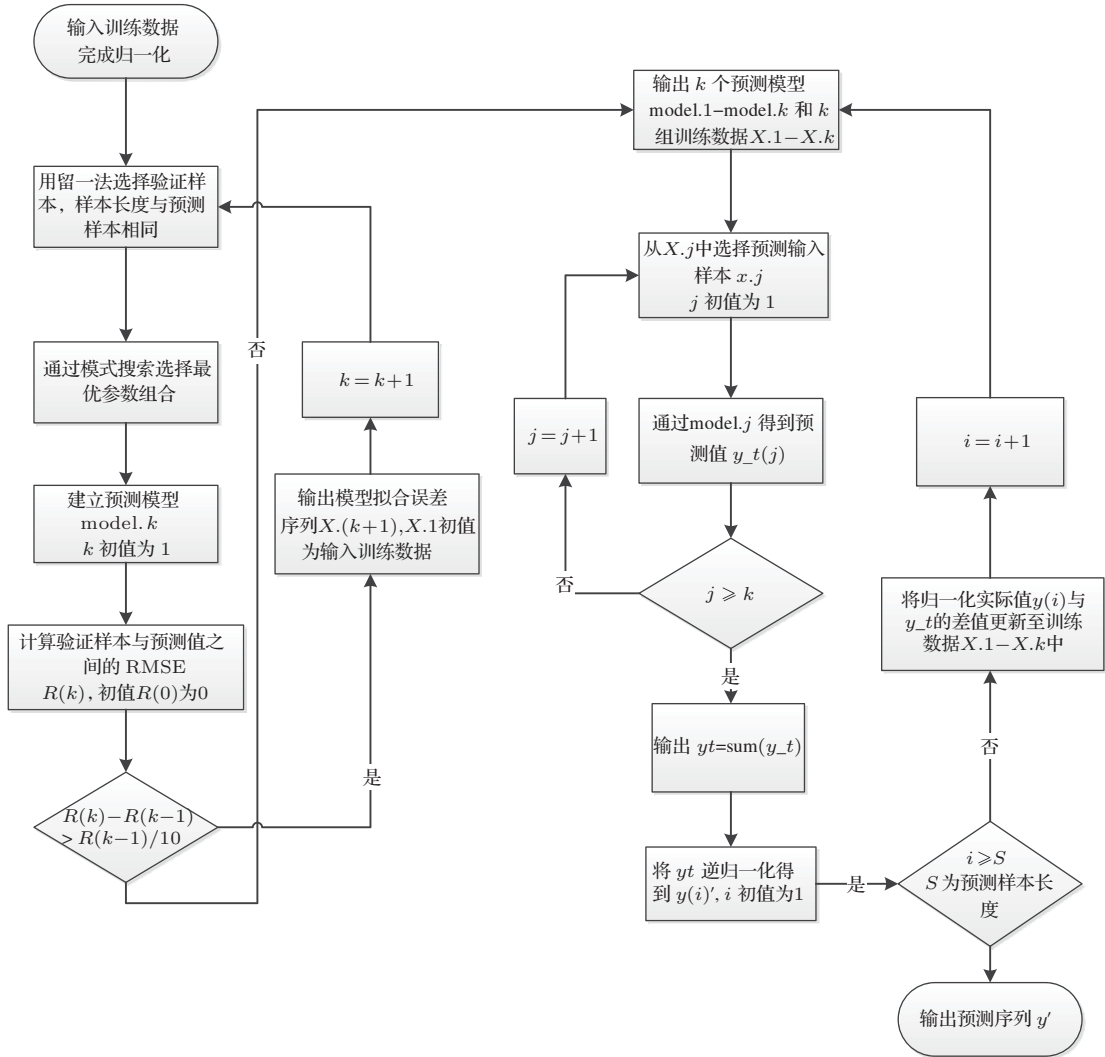


图 3 基于误差补偿和参数优化的预测算法流程图

$$= \sum_{j=1}^{n-m+1} \alpha_1(j) K(\mathbf{x}_1^j, \mathbf{x}) + b_1. \quad (13)$$

由上式可得预测模型的拟合误差表达式

$$X_2(i) = 0; \quad i \leq (m_1 - 1) \times \tau_1, \\ X_2(i) = f_1(m, \tau, C, \delta, \mathbf{x}_1^{i-(m_1-1) \times \tau_1}) - X_1(i); \\ n \geq i > (m_1 - 1) \times \tau_1, \quad (14)$$

式中, 当 $i \leq (m_1 - 1) \times \tau_1$, 为数据补零是为了使各预测模型的训练数据长度相等.

同样通过上述的方法, 将拟合误差序列作为训练数据建立验证样本的预测模型

$$f'_2(m_2, \tau_2, C_2, \delta_2, \mathbf{x}_2^{n-m-s+i}) = \sum_{j=1}^{n-m-s} \alpha'_2(j) K(\mathbf{x}_2^j, \mathbf{x}_2^{n-m-s+i}) + b'_2. \quad (15)$$

通过模式搜索选择优选参数组合 $[m_2, \tau_2, C_2, \delta_2^2]$, 并得到验证样本的预测均方根误差 R_2 , 此时可得拟合误差的预测模型

$$f_2(m_2, \tau_2, C_2, \delta_2, \mathbf{x}) = \sum_{j=1}^{n-m+1} \alpha_2(j)K(\mathbf{x}_2^j, \mathbf{x}) + b_2. \quad (16)$$

若 $R_2 < R_1 \times T$, 则继续对 f_2 的拟合误差通过上述方法建立预测模型 f_3 , 反复迭代直到 $R_{k+1} > R_k \times T$. 式中的 T 表示迭代中断阈值, 通过作者大量的测试, 在预测混沌序列时, 设置为 0.9—0.95 较为合适.

输出全部的预测模型 $\{f_1, f_2, \dots, f_k\}$ 和模型的训练数据 $\{X_1, X_2, \dots, X_k\}$ 可得预测值的表达式

$$\hat{y}_p = \sum_{i=1}^k \text{Invnorm}(f_i(m_i, \tau_i, C_i, \delta_i, \mathbf{x}_p^i)) = \sum_{i=1}^k \sum_{j=1}^{n-m+1} \text{Invnorm}(\alpha_i(j)K(\mathbf{x}_i^j, \mathbf{x}_p^i) + b_i), \quad (17)$$

式中的 $\text{Invnorm}(\cdot)$ 表示逆归一化操作, \mathbf{x}_p^i 表示第 i 个预测模型的预测输入向量, \hat{y}_p 表示预测值.

最后通过离线单步预测的方式完成预测测试集的预测.

算法流程如图 3 所示.

5 实验仿真与分析

为了验证基于误差补偿和参数优化的预测算法的有效性, 我们用理论混沌序列 Lorenz 时间序列和 MackeyGlass 时间序列对算法性能进行验证. 采用五种指标说明预测性能, 如表 1 所示.

5.1 Lorenz 混沌时间序列

Lorenz 方程组表达式如下:

$$\begin{aligned} \frac{dx(t)}{dt} &= a(-x(t) + y(t)), \\ \frac{dy(t)}{dt} &= -bx(t) - y(t) - x(t)z(t), \\ \frac{dz(t)}{dt} &= x(t)y(t) - cz(t), \end{aligned} \quad (18)$$

其中 a, b, c 都是无量纲参数, 其典型取值为 $a = 10, b = 28, c = 8/3$, 初始状态取值为 $x(0) = 12,$

$y(0) = 2, z(0) = 9$ 时, 此时的 Lorenz 系统具有混沌特性^[2]. 取步长 0.02, 选用四阶龙格库塔法^[19] 产生长度为 1250 的 x 分量混沌时间序列, 将前 800 个数据用于训练, 后 450 个数据用于测试. 首先选择训练数据中的后 450 个数据作为验证样本, 通过前文所述的方法依次建立数据预测模型和迭代残差预测模型. 模式搜索算法参数配置如表 2 所示.

表 1 仿真性能指标

指标	计算公式
均方根误差	$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^S [y(t) - \hat{y}(t)]^2}{(S-1)}}$
平均绝对百分率误差	$\text{SMAPE} = \frac{1}{S} \sum_{t=1}^S \left \frac{y(t) - \hat{y}(t)}{y(t) + \hat{y}(t)} \right $
正则化均方误差	$\text{NMSE} = \frac{\sum_{t=1}^S [y(t) - \hat{y}(t)]^2}{\sum_{t=1}^S [y(t) - \bar{y}]^2}$
预测误差绝对值均值	$\frac{1}{S} \sum_{t=1}^S y(t) - \hat{y}(t) $
预测误差绝对值最大值	$\max(y(t) - \hat{y}(t))$

表 2 混沌序列预测模式搜索参数配置表

搜索参数	m, C, δ^2, τ
起始点	2, 1, 1, 1
上界	30, $10^8, 10^8, 10$
下界	2, 0.0001, 0.0001, 1
初始步长	1
步长变化倍率	2, 0.5
退出条件	步长小于 2^{-20} 或者搜索次数大于 400 次

每次迭代产生的预测模型优化参数和预测效果如表 3 所示.

图 4 是本文提出的预测方法对于 Lorenz 序列的预测效果图, 从图 4(a)—(d) 中可以看出, 前 3 次迭代的误差都是平滑的, 能够通过建立不同预测模型很好地进行拟合, 最后的误差序列已经接近于噪声, 如图 4(f), 误差数量级也达到了 10^{-7} .

验证样本的预测精度要好于预测样本的精度, 结果与前文中的分析一致, 算法通过留一交叉验证法对参数寻优, 最后选择的是预测验证样本的最优参数组合, 而验证样本仅与预测样本存在相关性, 所以寻优所得参数不是预测样本的最优参数. 通过模型对误差序列的准确拟合说明了在先验信息仅

有训练数据的条件下,很难通过一个预测模型准确地逼近混沌系统的映射关系,余下的误差序列中还保留着确定性的成分,因为以任意精度逼近映射函数的LSSVM模型的最优参数组合在得到预测样本之前是无法确定的.而先验信息仅有训练数据的条

件下联合多个模型去逼近混沌系统的映射关系在次优的条件下能够保证较高的预测精度,最后剩下的残差接近噪声也说明多个模型的联合预测能够比较完整地提取出系统中的确定成分,剩下的几乎都是无法预测的随机部分.

表3 迭代误差补偿预测模型参数及其Lorenz序列预测效果

迭代次数	模型参数				描述	验证样本 RMSE	预测样本 RMSE
	m	C	δ^2	τ			
1	8.0	9.9841×10^8	747.125	1.0	模型1是原始数据序列预测模型	6.1542×10^{-4}	1.4757×10^{-4}
2	13.0	9.9964×10^8	234779.0079	1.0	模型2是模型1的预测误差预测模型	1.1638×10^{-6}	5.1513×10^{-7}
3	6.0	452.6256	348.0805	1.0	模型3是模型2的预测误差预测模型	1.4613×10^{-7}	1.1936×10^{-7}
4	15.0	53.6495	27.0216	1.0	模型4是模型3的预测误差预测模型	5.2364×10^{-8}	1.1846×10^{-7}
5	2.0	2234765.0	24052.0	3.0	模型5是模型4的预测误差预测模型	5.2250×10^{-8}	1.1893×10^{-7}
6	2.0	1.011×10^{-4}	9.999×10^8	1.0	模型6是模型5的预测误差预测模型	5.0733×10^{-8}	1.1893×10^{-7}
7	2.0	1.011×10^{-4}	9.999×10^8	1.0	模型7是模型6的预测误差预测模型	5.0733×10^{-8}	1.1893×10^{-7}
8	2.0	1.011×10^{-4}	9.999×10^8	1.0	模型8是模型7的预测误差预测模型	5.0733×10^{-8}	1.1893×10^{-7}

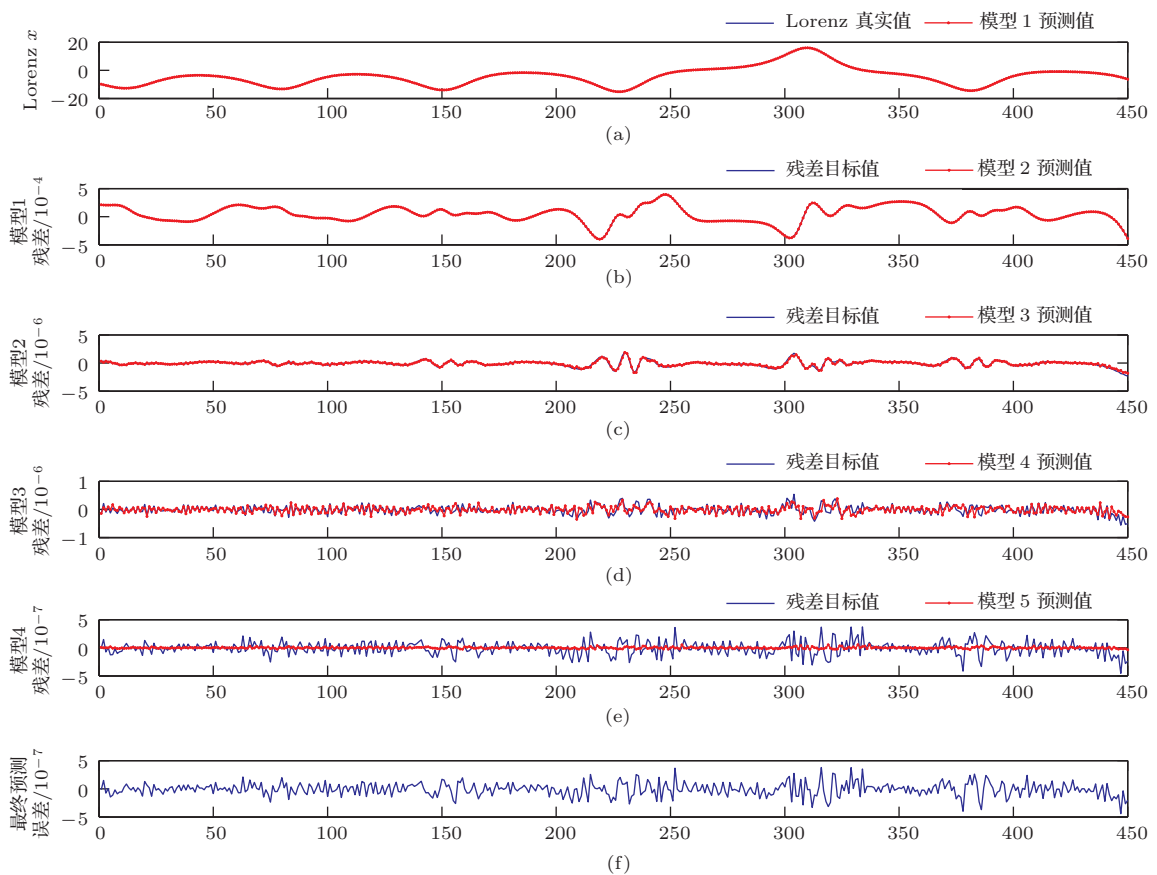


图4 Lorenz混沌序列单步预测效果图 (a) LSSVM预测Lorenz曲线; (b)—(e)各模型残差序列预测曲线; (f)本文方法预测Lorenz的误差曲线

从表3中也可以看出预测误差在迭代预测中逐次减少,每一个预测模型都对应着混沌系统映射关系中的一部分,相比不进行迭代仅依靠单个

LSSVM模型预测,基于迭代误差补偿的预测算法在预测精度上提高了3个数量级.此外,随着迭代次数的增加,误差补偿带来的性能增益越来越小,

这是因为随着确定性分量的提取, 随机分量所占比例越来越大, 直到最后的误差接近于噪声, 已经无法对预测样本进行有效预测.

同时从表 3 中可以看出, 迭代次数为 4 时, 预测性能就很难再提升, 预测样本的误差就已经收敛, 也表明本文设置的迭代中断阈值设置为 0.95 是合适的, 能够确保算法在接近收敛时及时退出迭代.

表 4 中比较了多篇参考文献中的 Lorenz 时间序列预测算法与本文所提算法的预测误差绝对值均值, 预测误差绝对值最大值, RMSE, NMSE, SMAPE. 从表中可以看出, 在各项误差评价指标中, 相比于现有算法本文所提算法都取得了至少一个数量级以上的大幅度性能提升.

表 4 Lorenz 序列预测性能对比

算法	性能指标				
	RMSE	NMSE	误差绝对值均值	误差绝对值最大值	SMAPE
PJE-NN ^[1]			1.32×10^{-5}	1.9×10^{-4}	
ESGP ^[2]	9.1258×10^{-5}				1.7845×10^{-5}
SVESM ^[2]	1.2676×10^{-5}				7.6531×10^{-7}
ARMA-RESN ^[2]	3.0264×10^{-6}				3.3689×10^{-7}
WESN ^[3]	9.58×10^{-4}				
RBF-NN ^[4]				5×10^{-3}	
SL-CCRNN ^[5]	6.36×10^{-3}	7.72×10^{-4}			
NL-CCRNN ^[5]	8.20×10^{-3}	1.28×10^{-3}			
FBMNN ^[7]	2.046×10^{-5}			7×10^{-5}	
IEOP ^[8]	1.0537×10^{-5}			4.8×10^{-5}	
DE-LSSVM ^[11]				2.8×10^{-6}	
MKL-SVR ^[12]	0.2595				
LNF ^[13]		6.4×10^{-5}			
本算法	1.1846×10^{-7}	3.3298×10^{-16}	8.9627×10^{-8}	4.4072×10^{-7}	2.4115×10^{-8}

5.2 MackeyGlass 混沌时间序列预测

为了检验所提算法的适应性, 本文选用 MackeyGlass 混沌序列测试算法的预测精度. Mackey-Glass 混沌时间序列具有典型的混沌特性, 通过下面的时滞微分方程产生:

$$\frac{dx}{dt} = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t), \quad (19)$$

此处的 τ 是时滞参数, 当 $\tau > 16.8$ 时, 时间序列具有混沌特性^[5], 对初值非常敏感. 我们选择参数 $\tau = 17$, 初值 $x(0) = 1.2$, 并用龙格库塔法产生长度为 1200 的 MackeyGlass 混沌时间序列. 其中前 1000 个作为训练数据, 后 200 个作为测试样本. 通过本文所提算法进行预测, 模式搜索参数配置参见表 2, 预测效果如图 5 所示.

MackeyGlass 迭代误差补偿预测每次迭代产生的预测模型优化参数和预测效果如表 5 所示.

表 5 迭代误差补偿预测 MackeyGlass 序列模型参数及其预测效果

迭代次数	模型参数				验证样本 RMSE	预测样本 RMSE
	m	C	δ^2	τ		
1	19.0	9.9738^6	21.99	1.0	8.7656×10^{-6}	1.4757×10^{-4}
2	30.0	9.9355×10^6	2.3288×10^6	1.0	2.9385×10^{-6}	3.1489×10^{-6}
3	20.0	9.9998×10^6	4.31	1.0	2.7159×10^{-6}	2.8934×10^{-6}
4	12.0	3.5544	11.8499	1.0	2.5473×10^{-6}	2.8381×10^{-6}
5	2.0	7.1413×10^6	5.9842×10^6	1.0	1.6686×10^{-6}	2.8176×10^{-6}
6	26.0	5.2275×10^5	9.7359×10^8	4.0	1.6574×10^{-6}	

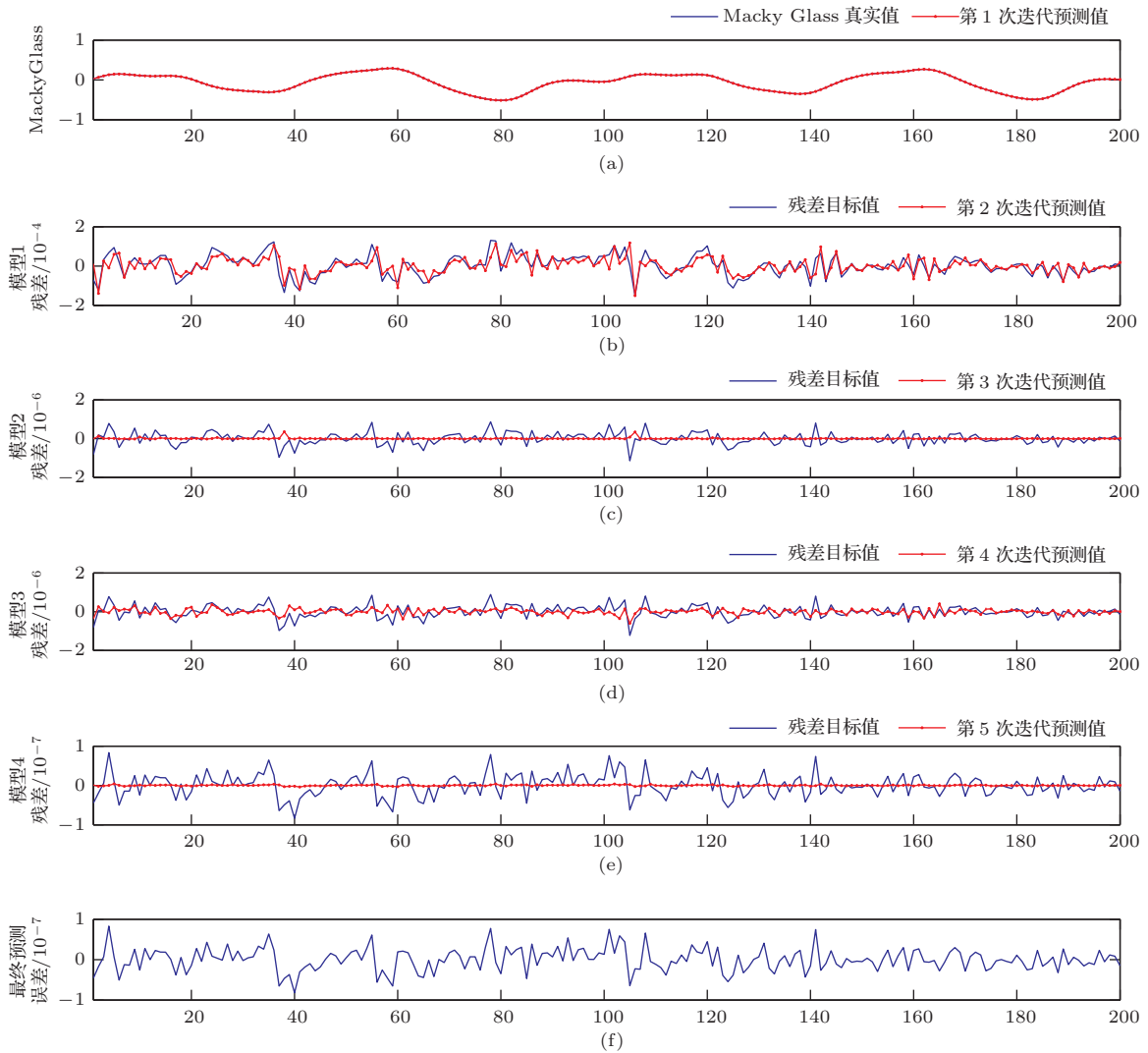


图5 MackeyGlass 混沌序列单步预测效果图 (a) LSSVM 预测 MackeyGlass 曲线; (b)–(e) 各模型残差序列预测曲线; (f) 本文方法预测 MackeyGlass 的误差

表6 MackeyGlass 序列预测性能对比

算法	性能指标		
	RMSE	NMSE	误差绝对值最大值
SL-CCRNN ^[5]	9.4×10^{-3}	6.31×10^{-4}	
NL-CCRNN ^[5]	1.2×10^{-2}	1.11×10^{-3}	
FBMNN ^[7]	3.02×10^{-4}		
IEOP ^[8]	1.941×10^{-5}		
LNF ^[13]	7.9×10^{-4}		
LLNF ^[13]	2.0×10^{-3}		
ODCF-SVM ^[14]	1.6×10^{-2}		6.4×10^{-5}
R-SSA ^[20]			7.9×10^{-5}
本算法	2.82×10^{-6}	2.8×10^{-11}	8.32×10^{-6}

从图5中可以看出, 本文所提算法同样适用于 MackeyGlass 混沌时间序列的预测, 经过多次迭代

后, 残差中的确定性成分能够被充分的预测. 从表5中可以看出, 在前2次迭代中, 验证样本和预测样本的预测误差呈现出快速下降的趋势; 算法在进行第6次迭代时, 验证样本预测误差满足迭代中断条件, 而预测样本的预测误差也接近于收敛. 表6中比较了已有的 MackeyGlass 混沌时间序列预测算法与本文所提算法的预测误差绝对值均值, 预测误差绝对值最大值, RMSE, NMSE, SMAPE. 从表6中可以看出, 在各项评价指标中, 相比于现有算法本文所提算法都取得了较大幅度的性能提升.

6 结论

本文提出了一种基于迭代误差补偿的最小二乘支持向量机预测方法, 通过多个模型的联合预测更加有效地逼近了混沌系统的映射函数. 相比于传

统的单预测模型, 本文所提方法在 Lorenz 时间序列预测和 MackeyGlass 预测的仿真实验中体现出大幅度的预测精度提升. 其原因在于传统预测方式中, 单个的预测模型难以描述混沌系统, 会遗漏部分确定性分量, 表现为预测误差变化平滑规律性较强. 而迭代误差补偿建立的多模型联合预测能够通过多个模型预测残差中的确定性分量, 直到预测误差主要由随机分量组成. 通过实验验证了本文所提算法在预测精度上相对已有算法有大幅度的提升, 达到一个数量级以上; 此外实验表明引入迭代中断阈值后, 算法在确保收敛的同时能够有效抑制迭代引起的复杂度增长. 本文还给出了一种先验信息仅有训练数据的条件下, 通过留一交叉验证法和模式搜索算法自动选择预测模型优化参数组合的方法, 简单高效, 具有较强的实用性.

参考文献

- [1] Chen D Y, Liu Y, Ma X Y 2012 *Acta Phys. Sin.* **61** 100501 (in Chinese)[陈帝伊, 柳焯, 马孝义 2012 物理学报 **61** 100501]
- [2] Han M, Xu M L 2013 *Acta Phys. Sin.* **62** 120510 (in Chinese)[韩敏, 许美玲 2013 物理学报 **62** 120510]
- [3] Song T, Li H 2012 *Acta Phys. Sin.* **61** 080506 (in Chinese)[宋彤, 李菡 2012 物理学报 **61** 080506]
- [4] Lei Z, Fengchun T, Shouqiong L, Lijun D, Xiongwei P, Xin Y 2013 *Sensors and Actuators B: Chemical* **182** 71
- [5] Rohitash C, Mengjie Z 2012 *Neurocomputing* **86** 116
- [6] Cheng-Jian L, Cheng-Hung C, Chin-Teng L 2009 *IEEE Trans. Sys. Tem. Man Cyber.* **39** 55
- [7] Ma Q L, Zheng Q L, Peng H, Tan J W 2009 *Acta Phys. Sin.* **58** 1410 (in Chinese)[马千里, 郑启伦, 彭宏, 覃姜维 2009 物理学报 **58** 1410]
- [8] Zhang C T, Ma Q L, Peng H 2010 *Acta Phys. Sin.* **59** 7623 (in Chinese)[张春涛, 马千里, 彭宏 2010 物理学报 **59** 7623]
- [9] Shi Z W, Han M 2007 *IEEE Trans. Neural Netw.* **18** 359
- [10] Chatzis S P, Demiris Y 2011 *IEEE Trans. Neural Netw.* **22** 1435
- [11] Zhang W Z, Long W, Jiao J J 2012 *Acta Phys. Sin.* **61** 220506 (in Chinese)[张文专, 龙文, 焦建军 2012 物理学报 **61** 220506]
- [12] Zhang J F, Hu S S 2008 *Acta Phys. Sin.* **57** 2708 (in Chinese)[张军峰, 胡寿松 2008 物理学报 **57** 2708]
- [13] Arash M, Majid A 2013 *IEEE Trans. Neural Netw.* **24** 207
- [14] Yu Y H, Song J D 2012 *Acta Phys. Sin.* **61** 170516 (in Chinese)[于艳华, 宋俊德 2012 物理学报 **61** 170516]
- [15] Vapnik V N, 1999 *The Nature of Statistical Learning Theory* (2nd Ed.) (New York, Springer) pp183–190
- [16] Sapankevych N I, Sankar R 2009 *IEEE Comput. Intell. Mag.* **4** 24
- [17] Cai C Z, Fei J F, Wen Y F, Zhu X J, Xiao T T 2009 *Acta Phys. Sin.* **58** S008
- [18] Ligang Z, Kin K L, Lean Y 2009 *Soft Comput.* **13** 149
- [19] Chen M H, Ceckbum B, Reitich F 2005 *J. Sci. Comput.* **22** 205
- [20] Mirmomeni M, Lucas C, Araabi B N, Moshiri B, Bidar M R 2011 *IET Signal Process.* **5** 515

A least square support vector machine prediction algorithm for chaotic time series based on the iterative error correction*

Tang Zhou-Jin[†] Ren Feng Peng Tao Wang Wen-Bo

(School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China)

(Received 24 October 2013; revised manuscript received 26 November 2013)

Abstract

This paper analyzes the error characteristic of traditional support vector machine prediction algorithm, where the error series are smooth and regular. This is because a single prediction model is incapable of fitting chaotic system mapping function and omitting some deterministic component. On this basis, a prediction algorithm that consists of an iterative error correction and a least square support vector machine (LSSVM) is proposed. The algorithm creates multiple predictive models via the method of iterative error correction to approximate the chaotic system mapping function and obtain significant improvements of predictive performance. In addition, the optimal parameters of the prediction model are automatically obtained from the pattern search algorithm which is simple and effective. Experiment conducted on Lorenz time series and MackeyGlass time series indicates that the proposed algorithm has a much better performance than that recorded in the literature.

Keywords: chaos time series prediction, least square support vector machine, iterative error correction, parameter composite optimization

PACS: 05.45.Tp, 05.45.Gg

DOI: [10.7498/aps.63.050505](https://doi.org/10.7498/aps.63.050505)

* Project supported by the Chinese Defence Advanced Research Program of Science and Technology, China(Grant No. 208010201).

[†] Corresponding author. E-mail: tangzhoujin@gmail.com