

基于无迹卡尔曼滤波的 LSSVR 在线多步时间序列预测

刘小雍^{1*}, 方华京², 熊中刚¹, 许宁¹

(1. 遵义师范学院 工学院, 贵州 遵义 563006; 2. 华中科技大学 自动化学院, 湖北 武汉 430074)

摘要: 准确宽范围多步预测在时间序列预测应用中带来了巨大挑战. 提出了一种基于最小二乘支持向量回归(LSSVR)和无迹卡尔曼滤波(UKF)的在线多步预测方法, 利用时间滑动窗口减小算法的计算负荷, UKF方法实现 LSSVR 模型参数更新以提高预测精度. 当预测范围达到预定步长 p 时, 由核宽度 σ 、支持值参数 $\{\alpha_k\}_{k=1}^L$ 以及偏移项 b 所构成的模型参数通过新的测量值和 UKF 进行在线更新. 提出的方法不仅以较少的训练数据建立在线预测模型(所需训练数据集大小为相空间维数与滑动窗口长度之和), 且多步预测值的精度相比于传统方法得到进一步提高. 最后, 通过几个实验研究论证了提出方法的有效性和优越性.

关键词: 在线多步预测; 最小二乘支持向量回归(LSSVR); 无迹卡尔曼滤波(UKF); 时间滑动窗口
中图分类号: TP391.9 **文献标志码:** A **开放科学(资源服务)标识码(OSID):** 

Multi-step-ahead Time Series Prediction based on LSSVR using UKF with Sliding-Windows

LIU Xiaoyong^{1*}, FANG Huajing², XIONG Zhonggang¹, XU Ning¹

(1. College of Engineering, Zunyi Normal University, Zunyi 563006, China;

2. School of Automation, Huazhong University of Science & Technology, Wuhan 430074, China)

Abstract: Accurate multi-step-ahead prediction over long future horizons posts great challenges for the application of time series prediction. A novel online multi-step-ahead prediction method based on least squares support vector regression (LSSVR) is proposed. Taking the superiorities of using sliding-windows to reduce largely computation burden and implementing LSSVR model updating by Unscented Kalman Filter (UKF), the proposed method not only can construct online predicted model in much fewer training data (such as the size of original training data set required is only the sum of embedding dimension corresponding to phase-space-reconstruction and the length of sliding-windows), but also has the better accuracy over multi-step-ahead prediction. When the prediction horizon reached the predefined step p in the process of predicting, model parameters consisted of kernel width σ , support values $\{\alpha_k\}_{k=1}^L$ and bias term b are updated by new arrived measurements and UKF. Finally, several simulations are provided to show the validity and applicability of the proposed method.

Key words: online multi-step-ahead prediction; least squares support vector regression (LSSVR); Unscented Kalman Filter (UKF); sliding-windows

0 引言

时间序列预测被广泛应用于工程领域, 尤其是时间序列预测分析及多步预测一直是许多研究领域的焦点^[1,2]. 然而, 跟单步预测不一样, 多步预测

正面临着来自外部不同干扰源的不确定性, 导致多步预测输出结果很不理想, 尤其是预测过程中出现的预测累积误差和先验信息缺失使得准确的多步预测变得更困难^[3]. 关于多步预测已提出多种预测方法, 有自回归滑动平均预测模型 (ARIMA)^[4]、

收稿日期: 2018-07-30; 修订日期: 2018-12-01; * 通信联系人, E-mail: liuxy204@163.com

基金项目: 国家自然科学基金项目(61473127); 贵州省教育厅青年项目(黔科合 KY 字[2016]254); 贵州省科学技术基金项目(黔科合基础[2018]1179, 黔科合 LH 字[2015]7012 号, 黔科合 LH 字[2016]7003 号); 贵州省千层次创新人才项目(遵师科合人才[2017]1179 号); 遵义师范学院博士项目(遵师 BS[2015]04 号)

作者简介: 刘小雍(1982—), 男, 贵州遵义人, 副教授, 博士, 主要从事数据模型融合建模及故障检测研究.

神经网络^[5]、支持向量机(SVR)及其变体最小二乘支持向量回归(LSSVR)等.目前在多步预测的策略上,通常使用两种方法,迭代预测和直接多步预测.对于 p 步迭代预测方法,是通过单步预测 p 次迭代来处理的,且每次单步预测值作为下一次预测的输入^[6].直接多步预测迭代预测的不同点是直接建立 p 步预测模型.多数情况下,基于SVR的时间序列预测采用直接多步预测方法实现多步预测^[7].简言之,直接预测模型实际上就是SVR回归估计的模型.正是考虑到SVR具有优越的泛化性能,基于SVR的直接预测模型被广泛使用.然而,在执行 p 步预测过程中,训练 p 个直接预测模型,显然需要消耗大量的运算时间,而迭代多步预测仅需要一个预测模型,缺陷是当使用迭代方法解决多步问题时,会出现预测误差的累积,导致预测精度严重下降.此外,迭代方法还存在预测过程中出现预测结果的稳定性问题.为了解决该问题,文献^[8]引入了一种多输入-多输出策略的多步预测方法,在预测值中的随机相关性以时间序列为特征,为建模时间序列的潜在动态特性变得更容易.此时预测值并不是一个标量,而是由预测值构成的向量,其大小等于预测步长.有关单步以及多步预测中的不同预测方法优缺点,可参考早期综述研究^[7].

尽管多步时间序列预测是一项具有挑战性的问题,但近年来不乏出现了一些有意义的研究.尤其是基于数据驱动方法,包括人工神经网络、自回归滑动平均模型、SVR等.将卡尔曼滤波模型(KFM)与神经网络(NN)相结合,提出了一种新的多步时间序列多步预测^[9],其中NN用于实现多步预测,KFM对预测值进行平滑,进而提高预测值的准确性.

然而,作为SVR的变体,最小二乘支持向量回归(LSSVR)^[10]也具备优良的泛化能力,通过将SVR的不等式约束变成等式约束,使用二范数损失的凸线性问题取代凸二次规划,极大地加速了模型的训练问题.通过大量的实验研究表明,LSSVR的泛化性能并不亚于SVR^[11].因此,LSSVR在时间序列预测、子空间辨识、信号处理等方面得到了广泛应用.因此,本文将从减少预测误差累积和提高预测精度两方面考虑,提出基于带滑动窗口的LSSVR及UKF方法的在线多步预测方法.该方法不仅可以以较小的训练数据建立在线预测模型,而且多步预测精度也进一步得到提高.

1 最小二乘支持向量回归

通过大量的实验研究表明,LSSVR的泛化性能并不亚于SVR.下面将对LSSVR做简要介绍,对于输入-输出数据 $\mathbf{x}_k = (x_{1k}, x_{2k}, \dots, x_{nk})^T$ 和 $\mathbf{y}_k (k = 1, 2, \dots, N)$,LSSVR^[10]可建模为

$$f(\mathbf{x}) = \sum_{k=1}^N \alpha_k K(\mathbf{x}, \mathbf{x}_k) + b, \quad (1)$$

式中 \mathbf{x} 表示新的测试数据.由于本文采用高斯核函数,式(1)可进一步描述为,

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

式中 σ 为高斯核函数, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_N)$ 为支持值向量, b 为偏差项,以上参数可通过如下优化求解,

$$\min_{\mathbf{w}, b, \xi_k} J_1(\mathbf{w}, b, \xi_k) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \gamma \frac{1}{2} \sum_{k=1}^N \xi_k^2, \quad (3)$$

$$\text{s.t. } y_k = \mathbf{w}^T \boldsymbol{\Phi}(\mathbf{x}_k) + b + \xi_k, \quad (4)$$

$$k = 1, 2, \dots, N,$$

式中 $\boldsymbol{\Phi}$ 表示从非线性空间到线性空间的特征映射,参数 $\gamma \in \mathbf{R}^+$ 表示规则化常量,用于对模型拟合度与解平滑性之间相对重要性的控制.对式(3)和式(4)使用拉格朗日方法得到如下无约束优化问题:

$$L(\mathbf{w}, b, \xi, \boldsymbol{\alpha}) = J_1(\mathbf{w}, b, \xi) - \sum_{k=1}^N \alpha_k (\mathbf{w}^T \boldsymbol{\Phi}(\mathbf{x}_k) + b + \xi_k - y_k). \quad (5)$$

根据KKT条件有,

$$\frac{\partial L}{\partial \mathbf{w}} = 0 \Rightarrow \mathbf{w} = \sum_{k=1}^N \alpha_k \boldsymbol{\Phi}(\mathbf{x}_k), \quad (6)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{k=1}^N \alpha_k = 0, \quad (7)$$

$$\frac{\partial L}{\partial \xi_k} = 0 \Rightarrow \alpha_k = \gamma \xi_k. \quad (8)$$

因此,LSSVR的学习过程可由式(6)~式(8)建立如下方程组求解,

$$\begin{pmatrix} \mathbf{0} & \mathbf{1}_N^T \\ \mathbf{1}_N & \boldsymbol{\Omega} + \gamma^{-1} I_N \end{pmatrix} \begin{pmatrix} b \\ \boldsymbol{\alpha} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{y} \end{pmatrix}, \quad (9)$$

其中

$$\mathbf{y} = (y_1, y_2, \dots, y_N)^T,$$

$$\mathbf{1}_N = (1, 1, \dots, 1)^T,$$

$$\boldsymbol{\Omega}_{ij} = K(\mathbf{x}_i, \mathbf{x}_j) = \boldsymbol{\Phi}^T(\mathbf{x}_i) \boldsymbol{\Phi}(\mathbf{x}_j),$$

对任意的 $i, j = 1, 2, \dots, N, K(\cdot, \cdot)$ 为满足Mercer条件的正定核函数.

2 基于 UKF 和 LSSVR 的在线多步时间序列预测

根据统计学理论以及本文中采用的高斯核函数,基于 LSSVR 的回归预测模型可表示为:

$$f(x) = \sum_{k=1}^N \alpha_k \exp\left(-\frac{\|x - x_k\|^2}{2\sigma^2}\right) + b, \quad (10)$$

其中 σ 为高斯核参数(gkp).在训练过程结束之后,LSSVR 预测模型的拓扑结构由支持向量 $\{x_k\}_{k=1}^L$ 、支持值参数 $\{\alpha_k\}_{k=1}^L$ 、偏移项 b 和 σ 组成.对于一个时变系统,考虑到获取到的实时运行数据是系统特征的较好反映,因此通过连续更新相关的模型参数来建立回归预测模型可以很好地捕捉系统新的特征.本文提出了一种基于 LSSVR 和 UKF 方法来实现预测模型的在线更新,进而提高多步预测的精度,如图 1 所示.

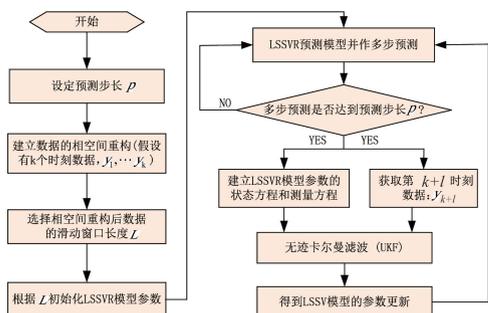


图 1 基于提出方法的在线多步预测

Fig. 1 Online multi-step prediction of the proposed method

首先确定预测步长 p ,根据相空间重构维数和滑动窗口长度初始化 LSSVR 预测模型,再借助获取到的第 $k+1$ 时刻的测量数据 y_{k+1} 和建立 LSSVR 预测模型参数更新的状态方程和测量方程,应用 UKF 方法对模型参数进行更新.

2.1 建立参数更新的 UKF 系统模型

在时间序列预测的过程中,建立相应的相空间重构是必不可少的.对于获取到的一维时间序列,使用嵌入理论建立对应时间序列的相空间重构^[12].假设来自非线性动态系统产生的单变量时间序列,即 $\{z_1, z_2, \dots, z_N\}$,其维数为 d 的相空间重构过程如下:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N-d} \end{pmatrix} = \begin{pmatrix} z_1 & \cdots & z_d \\ z_2 & \cdots & z_{d+1} \\ \vdots & \ddots & \vdots \\ z_{N-d} & \cdots & z_{N-1} \end{pmatrix}, \quad (11)$$

其中: $x_t = (z_t, z_{t+1}, \dots, z_{t+d-1})^T$ 为输入向量, $(z_{d+1}, z_{d+2}, \dots, z_{d+N-d})^T$ 为对应输出, $t = 1, 2, \dots, N-d$. 值得

注意的是,在许多应用中,对时间窗口大小的确定变得尤为重要,通常选取距离当前时刻最近的数据为重要数据,其余数据将自动舍弃,从而极大地减小计算负荷.假设时间窗口长度为 L ,即从相空间重构选取 L 个数据,且这些数据被认为是新的输入向量 $\{x_t, x_{t+1}, \dots, x_{t+L}\}$,对应的输出为 $z_{t+L+d-1}$.因此,基于滑动窗口大小的 LSSVR 预测模型描述的系统模型测量方程为,

$$g(x_k, s_k) = \sum_{l=1}^L \alpha_l \exp\left(-\frac{\|x_k - x_l\|^2}{2\sigma^2}\right) + b + v_k, \quad (12)$$

其中: $s_k \in \mathbf{R}^{L+2}$ 表示 LSSVR 模型在 k 时刻的参数向量, $x_k \in \mathbf{R}^d$ 预测模型在 k 时刻的输入向量.对于建立 UKF 系统模型的状态方程,考虑将要求解的后件参数向量作为状态变量 s_k :

$$s_{k+1} = \mathbf{A}s_k + w_k, s_k = (\alpha_k \quad b_k \quad \sigma_k)^T, \quad (13)$$

其中: \mathbf{A} 表示 $(L+2) \times (L+2)$ 的状态转移矩阵, w_k 为过程噪声(或状态噪声),且测量噪声 v_k 和过程噪声 w_k 都被假定为统计独立、零均值的高斯白噪声过程,其 v_k 和 w_k 所对应的协方差矩阵分别为:

$$E[w_i^T w_j] = \begin{cases} Q, & i=j, \\ 0, & i \neq j, \end{cases} \quad (14)$$

$$E[v_i^T v_j] = \begin{cases} r, & i=j, \\ 0, & i \neq j, \end{cases} \quad (15)$$

$$E[w_i v_j] = 0, \quad \forall i, j. \quad (16)$$

对于系统模型的状态方程,考虑到该状态由参数向量构成,通常情况下认为系统参数的变化相对于系统状态本身是缓慢变化的^[13],或者它的变化是一个平稳过程^[14],因此状态方程(13)的状态转移矩阵 \mathbf{A} 可用单位矩阵 \mathbf{I} 表示.在式(12)的预测过程中,输入向量 x_k 是已知的,且在下一时刻新的测量数据到来时被更新,例如在下一时刻 $k+1$, $x_{k-d+1} = (z_{k-d+1}, \dots, z_k)^T$.

2.2 使用 UKF 方法的在线参数更新

无迹卡尔曼滤波(UKF)^[15]解决了扩展卡尔曼滤波(EKF)的逼近问题,其思想是选择一组确定的 sigma 点,准确捕捉原始状态分布的均值和协方差;紧接着,假设系统模型噪声服从标准高斯分布,采用一组与 sigma 点相关的加权样本点来描述.为了在线估计回归预测模型 LSSVR 的参数,基于式(12)和式(13)所建立 UKF 系统模型的状态方程和测量方程如下:

$$g(x_k, s_k) = \sum_{l=1}^L \alpha_l \exp\left(-\frac{\|x_k - x_l\|^2}{2\sigma^2}\right) +$$

$$b + v_k, \quad (17)$$

$$s_{k+1} = A s_k + w_k,$$

$$s_k = (\alpha_k \quad b_k \quad \sigma_k)^T. \quad (18)$$

本文方法考虑当有新的测量数据到来时,动态更新 LSSVR 模型参数.为了执行在 k 时刻的多步预测,在 k 时刻首先通过 UKF 和新的测量数据估计回归预测模型 LSSVR 参数.其估计主要包括两个步骤:预测和更新.在预测阶段,从初始时刻到 $k-1$ 时刻的测量数据通过嵌入理论被重构,并从相空间产生最近的 L 个输入向量并堆栈到滑动窗口,基于这 L 个输入向量实现 LSSVR 参数的预测.当在 k 时刻获取到新的可靠测量数据时, LSSVR 参数通过 UKF 更新.在预测过程中,达到预定的步长 p 时, k 时刻的 LSSVR 参数将作为 UKF 下一时刻模型更新的初始参数.

基于式(17)和(18),通过 UKF 实现 LSSVR 参数更新的主要过程如下:

1) 初始化 UKF

设定初始时刻从 0 开始,定义先验初始状态分布为 $s_0 = (\alpha_0, b_0, \sigma_0)^T$, $P_{0|0} = P_0$ 以及 λ , 可通过如下式子计算点 τ_i :

$$\tau_i = \left[\sqrt{(L+2+\lambda) \cdot P_{k-1|k-1}} \right]_i, \quad (19)$$

其中 $[\cdot]_i$ 表示取其内容的第 i 列.

2) 预测阶段

设定中心点 $s_{k-1}^0 = \hat{s}_{k-1|k-1}$ 及相应权值:

$$w_{k-1}^{0(c)} = w_{k-1}^{0(m)} = \frac{\lambda}{L+2+\lambda}.$$

for $i = 1, \dots, L+2$

do

计算 sigma 点 $\chi_{k|k-1}^i = \hat{s}_{k-1|k-1} \pm \tau_i$;

使用状态方程(18)计算状态预测输出 $\nu_{k|k-1}^i$ 并

分配权值 $w_{k-1}^{i(c)} = \frac{1}{2(L+2+\lambda)}$;

通过参数向量 $\nu_{k|k-1}^i$ 及测量方程(17)计算测量

值的预测输出 $\zeta_{k|k-1}^i$ 并分配权值 $w_{k-1}^{i(m)} = \frac{1}{2(L+2+\lambda)}$;

end for

计算状态预测均值 $\hat{s}_{k|k-1}$ 和测量预测均值 $\hat{z}_{k|k-1}$

如下:

$$\hat{s}_{k|k-1} = E[f(s_k) | z_{k-1}] \approx \sum_{i=0}^{2(L+2)} w_i^m \nu_{k|k-1}^i, \quad (20)$$

$$\hat{z}_{k|k-1} = E[g(x_{k-1}, s_k) | z_{k-1}] \approx \sum_{i=0}^{2(L+2)} w_i^c \zeta_{k|k-1}^i, \quad (21)$$

$$\nu_{k|k-1}^i = f(\chi_i),$$

$$\zeta_{k|k-1}^i = g(\nu_{k|k-1}^i), i = 0, 1, \dots, 2(L+2). \quad (22)$$

计算状态和测量预测协方差矩阵 $P_{s(k|k-1)}$ 和 $P_{z(k|k-1)}$ 如下:

$$P_{s(k|k-1)} = E[(f(s_k) - \hat{s}_{k|k-1})(f(s_k) - \hat{s}_{k|k-1})^T] + Q_k, \quad (23)$$

$$P_{z(k|k-1)} = E[(g(x_{k-1}, s_k) - \hat{z}_{k|k-1})(g(x_{k-1}, s_k) - \hat{z}_{k|k-1})^T] + R_k. \quad (24)$$

3) 更新阶段

计算测量预测和状态预测的交叉协方差矩阵 $P_{sz(k|k-1)}$ 和卡尔曼增益矩阵 K_k :

$$P_{sz(k|k-1)} = \sum_{i=0}^{2(L+2)} w_i^c [\nu_{k|k-1}^i - \hat{s}_{k|k-1}][g(\nu_{k|k-1}^i) - \hat{z}_{k|k-1}]^T, \quad (25)$$

$$K_k = P_{sz(k|k-1)} P^{-1}_{z(k|k-1)}. \quad (26)$$

当前时刻 k 的状态更新 $\hat{s}_{k|k}$ 以及状态协方差矩阵更新 $P_{s(k|k)}$:

$$\hat{s}_{k|k} = \hat{s}_{k|k-1} + K_k(z_k - \hat{z}_{k|k-1}), \quad (27)$$

$$P_{s(k|k)} = P_{s(k|k-1)} - K_k P_{z(k|k-1)} K_k^T. \quad (28)$$

2.3 基于更新后回归预测模型的多步预测

通过 2.2 节的内容,便可完成 LSSVR 模型参数的估计,即完成回归预测模型的更新.在进行多步预测过程中,包括直接多步预测和迭代递归预测两种方法.但在本文中,考虑到建立直接预测模型需要更多的建模时间,迭代递归预测方法将被采用^[3].尽管迭代预测在执行长期预测的任务过程中,面临多步预测性能的下降,但提出的方法会对多步预测达到预定步长时,通过新到的测量数据以及应用 UKF 方法对 LSSVR 预报模型进行不断更新,避免了传统多步预测方法采用建立起的一个确定模型在预测过程中,始终保持不变而导致预测精度严重下降的问题.

对于递归预测方法,假设单变量时间序列 $\{z\}_{i=1}^N$, 时间延迟常量 $\tau=1$, 滑动窗口长度为 L , N 表示时间序列的长度,对该序列通过相空间重构之后且嵌入维数为 d , 可确定的一步预测模型为:

$$\begin{aligned} \hat{z}_{t+1} &= g_k(x_t, x_{t-1}, \dots, x_{t-L+1}, s_k), \\ x_t &= (z_t, z_{t-1}, \dots, z_{t-d+1}), \end{aligned} \quad (29)$$

其中: k 表示预报模型在第 k 次被更新, s_k 表示更新后模型所对应的第 k 个 LSSVR 模型参数, x_t 为根据式(29)进行构造的数据.在 t 时刻的 p 步预测需要先前的 p 个一步预测来实现,即 $\hat{z}_{t+1}, \hat{z}_{t+2}, \dots, \hat{z}_{t+p-1}$, p 为预定义的步长.也就是说, p 步预测的数学模型可表述为如下:

$$\begin{aligned} \hat{z}_{t+1} &= g_k(\mathbf{x}_t, \mathbf{x}_{t-1}, \dots, \mathbf{x}_{t-L+1}, \mathbf{s}_k), \\ \mathbf{x}_t &= (z_t, z_{t-1}, \dots, z_{t-d+1}), \end{aligned} \quad (30)$$

$$\begin{aligned} \hat{z}_{t+2} &= g_k(\hat{\mathbf{x}}_{t+1}, \mathbf{x}_t, \dots, \mathbf{x}_{t-L+2}, \mathbf{s}_k), \\ \hat{\mathbf{x}}_{t+1} &= (\hat{z}_{t+1}, z_t, \dots, z_{t-d+2}), \end{aligned} \quad (31)$$

$$\begin{aligned} \dots, \\ \hat{z}_{t+p} &= g_k(\hat{\mathbf{x}}_{t+p-1}, \hat{\mathbf{x}}_{t+p-2}, \dots, \mathbf{x}_{t-L+p}, \mathbf{s}_k), \\ \hat{\mathbf{x}}_{t+p-1} &= (\hat{z}_{t+p-1}, \dots, \hat{z}_{t+1}, z_t, \dots, z_{t+d-p}). \end{aligned} \quad (32)$$

显然,当预测范围未达到预测步长时,多步预测可通过式(30)~(32)来进行;此外,当预测步长等于 p ,且在 k 时刻获取到新的测量数据时,LSSVR 模型参数通过 UKF 被更新,更新后的模型为 $g_k(M, s_k)$,其中 M 表示由 L 个向量 $\mathbf{x}_{i=1}^L$ 构成.

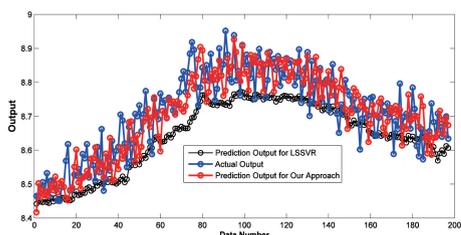


图 2 提出方法的四步预测输出

Fig. 2 Ahead four-step prediction output of the proposed method

表 1 提出方法在训练数据集 L_{tr} 较小时,不同预测步长的预测输出

Tab. 1 Results of multi-step-ahead prediction on RMSE for the proposed method when the size of L_{tr} is the smaller

预测步长	RMSE	预测步长	RMSE	预测步长	RMSE
1	0.2692	6	0.4627	12	0.7335
2	0.3130	7	0.5062	15	0.8559
3	0.3301	8	0.5547	18	0.9799
4	0.3879	9	0.5998	20	1.0609
5	0.4322	10	0.6430		

表 2 提出方法在训练数据集 L_{tr} 较大时,两种不同方法不同预测步长的预测输出

Tab. 2 Results of multi-step-ahead prediction on RMSE for the proposed method when the size of L_{tr} is relative large

预测步长	RMSE		预测步长	RMSE	
	LSSVR	本文方法		LSSVR	本文方法
1	1.1144	0.2840	7	2.1882	0.5285
2	1.4588	0.3273	8	2.3021	0.5847
3	1.7981	0.3318	9	2.3640	0.6362
4	1.8382	0.3983	10	2.5229	0.6734
5	1.9939	0.4470	12	2.5280	0.6744
6	2.0077	0.4793	15	2.9360	0.9015

首先考虑如下非线性动态系统:

$$\begin{aligned} y(k+1) &= \frac{y(k)y(k-1)[y(k)+2.5]}{1+y^2(k)+y^2(k-1)} + \\ &\sin \frac{2\pi k}{50} + v(k), \end{aligned} \quad (34)$$

3 实验研究

这部分将提出的方法与传统 LSSVR 方法进行多步预测的性能比较,其中 LSSVR 最优超参数是通过在训练数据集上的十阶交叉验证来获取,但当原始训练数据集较小时,十阶交叉验证不能用于超参数的寻优.对于不同方法的性能评估,在测试数据上将采用均方根误差(Root Mean Squared Error, RMSE),

$$RMSE = \sqrt{\sum_{k=1}^{N_{test}} (y_k - \hat{y}_k)^2 / N_{test}}, \quad (33)$$

其中 y_k 是对应测试数据集中第 k 个样本 \mathbf{x}_k 的目标输出, \hat{y}_k 不同方法的预测输出, N_{test} 是测试数据的总样本数.接下来考虑两种情况下的仿真研究:

情况 1) 当训练数据集较小时,不能采用十阶交叉验证寻优参数,即不能建立 LSSVR 预测模型或建立 LSSVR 模型的意义不大,此时不考虑 LSSVR 与提出方法的比较,因此超参数可设定为任意大于 0 作为 UKF 的初始值;

情况 2) 训练数据集具备采用十阶交叉验证寻优的条件,将寻优后的超参数作为 UKF 的初始值,再将提出方法与 LSSVR 进行比较.

其中 $y(0) = 0, y(1) = 0, 1 \leq k \leq 100, v(k)$ 表示均值是 0,方差为 0.25 的高斯噪声.对于情况 1,定义训练数据集大小为 L_{tr} ,其大小被选择为 $L_{win} + IN_{dim}$, L_{win}, IN_{dim} 分别表示滑动窗口长度和相空间中每个输入数据的输入维数大小.在该实验中,若选择 L_{win}

为 4, IN_{dim} 为 2, 则训练数据集 L_{tr} 为 6, 显然 L_{tr} 太小, 不足以建立 LSSVR 回归预测模型, 只能用本文提出的方法进行多步预测. 表 1 和图 2 分别给出了本文方法的 RMSE 和四步预测结果, 清晰地表明, 本文方法在训练数据集 L_{tr} 较小时, 也能获取到较好的四步预测结果. 对于情况 2, 首先从式 (34) 获取 501 个数据, 其中前 100 个数据作为训练模型, 其余数据用作测试. 为了将提出方法与传统 LSSVR 预测方法有一个公平的比较, 采用十阶交叉验证方

法获取 LSSVR 预测模型的最优超参数, 并以此参数作为本文方法中 UKF 的初始参数, 其中 L_{win} 、 IN_{dim} 分别选取为 4 和 10. 两种不同方法的不同预测步长所对应的 RMSE 如表 2 所示, 图 3 和图 4 分别给出了两种方法在测试数据上的三步和九步预测结果. 显然, 尽管本文方法在这些测试数据上的预测结果不是最好的, 但相比于传统的预测方法, 确实减小了累积误差, 实现更好的多步预测.

表 3 两种不同方法不同预测步长的预测输出比较

Tab. 3 Comparisons of multi-step-ahead prediction on RMSE between the proposed method and LSSVR

预测步长	RMSE		预测步长	RMSE	
	LSSVR	本文方法		LSSVR	本文方法
1	0.0794	0.0625	7	0.1117	0.0692
2	0.0835	0.0639	8	0.1236	0.0704
3	0.0843	0.0646	10	0.1368	0.0725
4	0.0914	0.0657	15	0.1783	0.0831
5	0.0984	0.0675	20	0.2158	0.0932
6	0.1050	0.0685			

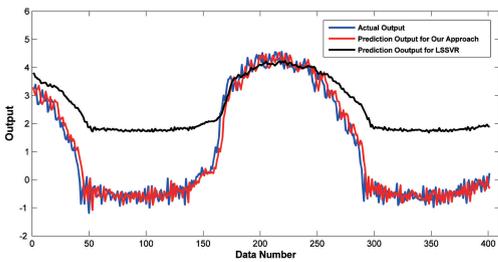


图 3 提出方法与 LSSVR 方法的三步预测输出
Fig. 3 Ahead three-step prediction output of the proposed method

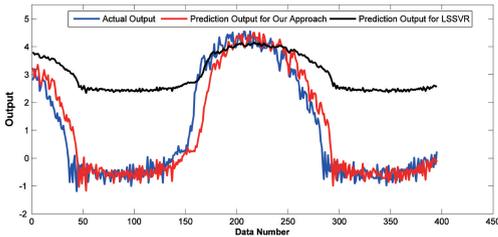


图 4 提出方法与 LSSVR 方法的九步预测输出
Fig. 4 Ahead nine-step prediction output of the proposed method

为了进一步验证本文方法的优越性, 下面将考虑来自田纳西过程 (TEP) [16] 的一组故障 4 的真实数据, 对应的仿真平台可 <http://brahms.scs.uiu-c.edu/> 下载, 由反应器冷却水入口温度的阶跃变化引起, 包含 960 个测量值. 使用前 250 个测量值作为训练数据, 剩下的用作测试数据. 同理, 为了获取传统预测方法的最佳预测, 采用十阶交叉验证建立其

最优超参数. 在本文方法中, L_{win} 和 IN_{dim} 分别选取为 2 和 10. 表 3 给出了本文方法与 LSSVR 方法在不同预测步长下的 RMSE 比较, 预测步长为 8 和 12 的预测输出如图 5 和图 6 所示, 进一步表明, 无论是从 RMSE, 还是不同预测步长的预测输出结果来看, 本文方法有着更好的预测性能.

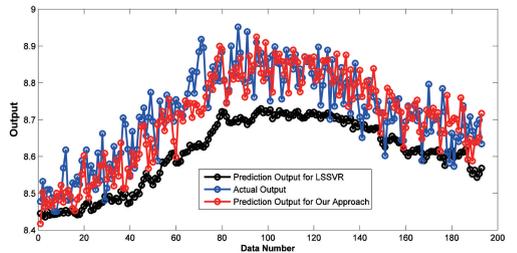


图 5 提出方法与 LSSVR 方法的八步预测输出
Fig. 5 Ahead eight-step prediction output of the proposed method

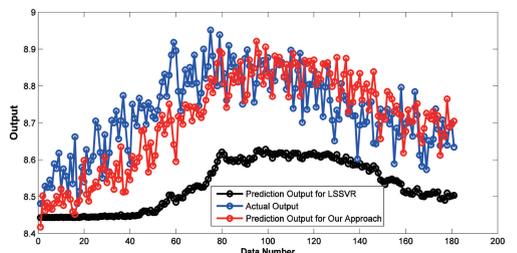


图 6 提出方法与 LSSVR 方法的十二步预测输出
Fig. 6 Ahead twelve-step prediction output of the proposed method

4 结论

本文提出了一种基于 UKF 及滑动窗口的在线 LSSVR 多步预测方法,该方法不仅可以以较小的训练数据集(如训练数据集的大小仅仅是滑动窗口长度和对应相空间中输入数据维数之和)建立在线预

测模型,而且在训练数据集较大时,尽管本文方法在测试数据上的预测输出结果并不一定是最好的,但它却是能够很大程度上减少预测输出的累计误差,相比于传统的 LSSVR 预测方法有更好的预测性能。

参考文献:

- [1] 殷礼胜,何怡刚,董学平,等.交通流量 VNNMF 神经网络模型多步预测研究[J].自动化学报,2014,40(9):2067-2073.
YAN Lisheng, HE Yigang, DONG Xueping, et al. Research on the multi-step prediction of volterra neural network for traffic flow[J]. Acta Automatica Sinica, 2014, 40(9): 2067-2073.
- [2] 南新元,曾庆凯,孟宪强.基于改进蛙跳算法优化的支持向量回归机 ORP 预测[J].新疆大学学报(自然科学版),2016,33(4):393-398.
NAN Xinyuan, ZENG Qingkai, MENG Xianqiang. Research on prediction of oxidation reduction potential based on improved SFLA-SVR algorithm[J]. Journal of Xinjiang University(Natural Science Edition), 2016, 33(4): 393-398.
- [3] LI Zhang, ZHOU Weida, PEI-CHANN Chang, et al. Iterated time series prediction with multiple support vector regression models[J]. Neurocomputing, 2013(99): 411-422.
- [4] QIN Mengjiao, LI Zhihang, DU Zhenhong. Red tide time series forecasting by combining ARIMA and deep belief network [J]. Knowledge-Based Systems, 2017(125): 39-52.
- [5] 肖迁,李文华,李志刚,等.基于改进的小波-BP 神经网络的风速和风电功率预测[J].电力系统保护与控制,2014,42(15):80-86.
XIAO Qian, LI Wenhua, LI Zhigang, et al. Wind speed and power prediction based on improved wavelet-BP neural network [J]. Power System Protection and Control, 2014, 42(15): 80-86.
- [6] TAIEB S B, BONTEMPI G, ATIYA A F, et al. A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition[J]. Expert Systems with Applications, 2012(39): 7067-7083.
- [7] TAIEB S B, ANTTI S, BONTEMPI G. Multiple-output modeling for multi-step-ahead time series forecasting[J]. Neurocomputing, 2010(73): 1950-1957.
- [8] BAO Yukun, XIONG Tao, HU Zhongyi. Multi-step-ahead time series prediction using multiple-output support vector regression[J]. Neurocomputing, 2014(129): 482-493.
- [9] XIAO Qinkun, QIN ChuChao, LI Zhao. Time series prediction using dynamic Bayesian network[J]. Optik-International Journal for Light and Electron Optics, 2017(35): 98-103.
- [10] SUYKENS J A, VANDEWALLE J. Least squares support vector machine classifiers[J]. Neural Processing Letters, 1999, 9(3):293-300.
- [11] 阎威武,邵惠鹤.支持向量机和最小二乘支持向量机的比较及应用研究[J].控制与决策,2003,18(3):358-360.
YAN Weiwu, SHAO Huihe. Application of support vector machines and least squares support vector machines to heart disease diagnoses[J]. Control and Decision, 2003, 18(3): 358-360.
- [12] 马红光,李夕海,王国华,等.相空间重构中嵌入维和时间延迟的选择[J].西安交通大学学报,2004,38(4):335-338.
MA Hongguang, LI Xihai, WANG Guohua, et al. Selection of embedding dimension and delay time in phase space reconstruction[J]. Journal of Xi'an Jiaotong University, 2004, 38(4): 335-338.
- [13] HAYKIN S S, HAYKIN S S, HAYKIN S S. Kalman filtering and neural networks[M]. 2ed Edition. Singapore: Wiley-Interscience, 2001: 65.
- [14] AGUIRRE L A, TEIXEIRA B O S, TORRES L A B. Using data-driven discrete-time models and the unscented Kalman filter to estimate unobserved variables of nonlinear systems[J]. Physical Review E, 2005, 72(2): 026226.
- [15] JULIER S J, UHLMANN J K. Unscented filtering and nonlinear estimation[J].Proceedings of the IEEE, 2004, 92(3): 401-422.
- [16] CHIANG L H, PELL R J, SEASHOLTZ M B. Exploring process data with the use of robust outlier detection algorithms[J]. Journal of Process Control, 2003, 13(5): 437-449.

责任编辑:郭红建