

结构自适应序贯正则极端学习机时间序列预测及其应用^{*}

何 星, 王宏力, 陆敬辉, 姜 伟

(第二炮兵工程大学 控制工程系, 陕西 西安 710025)

摘 要: 为提高初始小样本情况下时间序列在线预测的精度, 提出了一种结构自适应序贯正则极端学习机 (SA-SRELM)。该方法在在线序贯学习阶段, 针对不同训练样本规模选择不同的递推方式对输出权值进行更新; 同时, 在训练样本达到一定规模后, 为提高预测模型对系统的动态适应性, 在加入新样本的同时对旧样本进行剔除, 完成预测模型的训练。利用3种混沌时间序列预测实例对所提方法的有效性进行了验证。最后, 将所提方法用于航空发动机排气温度预测中, 结果表明该方法相对正则极端学习机 (RELM) 和序贯正则极端学习机 (SRELM) 方法具有更好的泛化性能, 预测精度分别是二者的约6倍和2倍。

关键词: 正则极端学习机; 小样本; 时间序列预测; 动态适应性; 泛化能力

中图分类号: V557 文献标识码: A 文章编号: 1001-4055 (2015) 03-0458-07

DOI: 10.13675/j.cnki.tjjs.2015.03.020

Structure Adaptive Sequential Regularized Extreme Learning Machine Based on Time Series Prediction and Its Application

HE Xing, WANG Hong-li, LU Jing-hui, JIANG Wei

(Department of Control Engineering, The Second Artillery Engineering University, Xi'an 710025, China)

Abstract: To improve the on-line prediction precision of time serials under small initialization sample, a structure adaptive sequential regularized extreme learning machine (SA-SRELM) is proposed. For this new method, at the stage of on-line sequential learning, the output weight of SA-SRELM is renewed by different recursive methods according to training sample size. Meanwhile, when training sample size is added to a destined value, new samples are added and the oldest samples are removed to enhance the dynamic adaptability of prediction model. Based on above process, the prediction model is built. Three typical chaotic time series prediction examples are used to verified the validity of proposed method. In the end, the proposed method is used to predict aeroengine exhaust gas temperature. The results show that SA-SRELM can get better generalization performance compared with regularized extreme learning machine (RELM) and sequential regularized extreme learning machine (SRELM). The prediction accuracy for SA-SRELM is almost six times and twice as high as RELM and SRELM, respectively.

Key words: Regularized extreme learning machine; Small sample; Time series prediction; Dynamic adaptability; Generalization performance

1 引 言

航空发动机是飞行器的核心, 但由于其结构复

杂且长时间在高温、高压及高速等恶劣环境下工作, 已成为飞行器系统中故障率最高、维护成本最高的重要部件^[1,2]。而通过对航空发动机热力性能参数的

^{*} 收稿日期: 2014-03-31; 修订日期: 2014-05-16。

基金项目: 国家自然科学基金 (61004128)。

作者简介: 何 星 (1985—), 男, 博士生, 研究领域为智能故障诊断与预测。E-mail: trees241@163.com

预测从而预知航空发动机性能的变化趋势,可为预防与排除其故障提供决策依据,对于有效实施航空发动机的视情维修具有重要意义。发动机排气温度(Exhaust gas temperature, EGT)是可客观反映航空发动机性能的重要热力性能参数,利用其历史数据构建时间序列模型进行预测是实现航空发动机状态监测的有效途径之一,由于神经网络、支持向量机等人工智能方法具有无需建立准确的数学模型、强大的非线性映射能力等优点,被众多学者成功用于航空发动机排气温度的预测中,如陈果和钟诗胜等学者利用神经网络方法实现了排气温度预测^[3,4],许喆平等对飞机性能参数预测中存在的 uncertainty 处理问题进行了研究^[5],于广滨和范庚等学者则分别利用支持过程向量机和相关向量机对排气温度预测问题进行了深入地研究^[6,7]。但是,传统的人工智能方法存在着参数选取以及在线更新困难等缺点,在用于排气温度的在线预测时受到一定限制。

极端学习机是2006年由Huang等学者^[8]提出的一种新型单隐层前馈神经网络,与传统神经网络方法相比,其训练速度得到极大的提高,泛化性能更好,且有效避免了局部收敛等问题,已成功应用于回归和分类领域^[9~12]。不过由于极端学习机的输出权值通过计算隐层输出矩阵的伪逆得到,在隐层输出矩阵出现奇异或者输入权值过大等情况时会影响ELM的稳定性及泛化性能,为此相关学者提出了RELM^[13],通过在其训练目标中加入输出权值的二范数,从而在训练过程中,不仅最小化训练误差,还对表征了其泛化能力的结构风险进行控制,提高了ELM的稳定性和泛化能力,然而RELM与ELM一样仍然属于离线预测,无法利用系统实时状态参数更新预测模型。为实现RELM的在线预测,张弦等提出了可递推求解出输出权值的序贯正则极端学习机^[14,15],但其只适用于初始训练样本较大的情况,在初始小样本情况下难以取得良好的预测结果。

针对实际中常见的初始训练样本较少的情况,提出了根据样本规模选择输出权值计算方式的结构自适应序贯正则极端学习机(SA-SRELM),并利用典型混沌时间序列和国航某型航空发动机对其预测效果进行验证。

2 正则极端学习机(RELM)

对于时间序列 $\mathbf{S}_N = \{(\mathbf{x}_i, t_i)\}_{i=1}^N$, $\mathbf{x}_i \in \mathbf{R}^n$, $t_i \in \mathbf{R}$, 为

改善ELM方法只考虑经验风险带来的泛化能力不好的缺点,根据结构风险最小化原理,RELM通过引入参数 γ 调节两种风险的比例,可较好的实现经验风险与结构风险间的折中,利用RELM解决回归问题可转化为如下约束优化问题

$$\begin{aligned} & \min \left(\frac{1}{2} \boldsymbol{\beta}_N^T \boldsymbol{\beta}_N + \frac{\gamma}{2} \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} \right) \\ \text{st} \quad & t_j = \sum_{i=1}^L \beta_i g(\alpha_i x_j + b_i) - \varepsilon_j \end{aligned} \quad (1)$$

式中 $j=1, \dots, N$; $g(\cdot)$ 为激活函数; L 为隐层神经元数目; $\boldsymbol{\alpha}_i = [\alpha_{i1} \ \dots \ \alpha_{in}]$ 为连接第 i 个隐层神经元的输入权值; $\boldsymbol{\beta}_N = [\beta_1 \ \dots \ \beta_L]^T$, β_i 为连接第 i 个隐层神经元的输出权值; b_i 是第 i 个隐层神经元的偏差; $\boldsymbol{\varepsilon} = [\varepsilon_1 \ \dots \ \varepsilon_N]^T$, ε_i 为回归误差; 为求解上述优化问题,建立拉格朗日函数

$$\begin{aligned} L(\mathbf{w}, \boldsymbol{\varepsilon}, \boldsymbol{\beta}_N) &= \frac{1}{2} \boldsymbol{\beta}_N^T \boldsymbol{\beta}_N + \frac{\gamma}{2} \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} - \\ & \mathbf{w}(\mathbf{H}_N \boldsymbol{\beta}_N - \mathbf{T}_N - \boldsymbol{\varepsilon}) \end{aligned} \quad (2)$$

式中 $\mathbf{w} = [w_1 \ \dots \ w_N]$ 为拉格朗日权值。

$$\begin{aligned} \mathbf{H}_N &= \begin{bmatrix} G(\boldsymbol{\alpha}_1 \mathbf{x}_1 + b_1) & \dots & G(\boldsymbol{\alpha}_L \mathbf{x}_1 + b_L) \\ \vdots & & \vdots \\ G(\boldsymbol{\alpha}_1 \mathbf{x}_N + b_1) & \dots & G(\boldsymbol{\alpha}_L \mathbf{x}_N + b_L) \end{bmatrix} \\ \mathbf{T}_N &= [t_1 \ \dots \ t_N]^T \end{aligned}$$

对拉格朗日函数各变量求偏导并令偏导数为零可得 $\boldsymbol{\beta}_N$ 的解 $\tilde{\boldsymbol{\beta}}_N$

$$\tilde{\boldsymbol{\beta}}_N = \left(\mathbf{H}_N^T \mathbf{H}_N + \frac{\mathbf{I}}{\gamma} \right)^{-1} \mathbf{H}_N^T \mathbf{T}_N \quad (3)$$

式中 $\mathbf{I} \in \mathbf{R}^{L \times L}$, $\tilde{\boldsymbol{\beta}}_N = [\tilde{\beta}_1 \ \dots \ \tilde{\beta}_L]^T$ 。由此可得经过训练的RELM时间序列预测模型

$$\mathbf{v} = \sum_{i=1}^L \tilde{\beta}_i G(\boldsymbol{\alpha}_i \mathbf{u} + b_i) \quad (4)$$

式中 $\mathbf{u} \in \mathbf{R}^n$ 为模型输入, $\mathbf{v} \in \mathbf{R}$ 为模型输出。

3 结构自适应序贯正则极端学习机

RELM通过引入正则化参数能够有效提高ELM由于病态矩阵等问题带来的泛化能力差的问题,但是其与ELM一样属于离线方式,当训练样本分批加入时,只能通过重复训练来重新建立预测模型,这大大增加了训练时间。借鉴在线序贯极端学习机(Online Sequential Extreme Learning Machine, OS-ELM)的在线递推方法可以实现预测模型的在线更新^[16]。但是OS-ELM方法的隐层节点数目要求不能超过训

训练样本数目,而较多的隐层节点数目则是提高预测模型精度和泛化能力的有效途径,因此对于初始样本数量较少的情况,OS-ELM难以取得满意的精度;同时随着训练样本数目的逐步增加,OS-ELM不断的将新样本加入,而不对旧样本进行任何处理,这使得新旧样本在预测时处于同等地位,从而导致训练得到的预测模型对系统动态变化的跟踪能力不强。针对上述问题,SA-SRELM考虑对不同样本数目采用不同的输出权值计算方法,从而提高预测模型的精度和泛化性能。SA-SRELM具体过程如下:

首先,根据广义逆性质,当隐层输出矩阵 H_0 为行满秩 ($N_0 < L$) 时,其形式如式(5)所示

$$H_0 = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_L, b_L, x_1) \\ \vdots & \cdots & \vdots \\ G(a_1, b_1, x_{N_0}) & \cdots & G(a_L, b_L, x_{N_0}) \end{bmatrix}_{N_0 \times L} \quad (5)$$

RELM初始输出权值可表示^[13]

$$\beta_0 = H_0^T (I/\lambda + H_0 H_0^T)^{-1} T_0 \quad (6)$$

当新样本加入时,由于样本数量较少,因此采取只加入新样本来进行预测模型的更新,而不剔除旧样本。令当前时刻隐层输出矩阵为

$$H_k = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_L, b_L, x_1) \\ \vdots & \cdots & \vdots \\ G(a_1, b_1, x_k) & \cdots & G(a_L, b_L, x_k) \end{bmatrix}_{k \times L} = \begin{bmatrix} h_1^T & \cdots & h_k^T \end{bmatrix}^T \quad (7)$$

输出权值为

$$\beta_k = H_k^T (I/\lambda + H_k H_k^T)^{-1} T_k \quad (8)$$

当新样本 (x_{k+1}, t_{k+1}) 加入时,输出权值变为

$$\beta_{k+1} = H_{k+1}^T (I/\lambda + H_{k+1} H_{k+1}^T)^{-1} T_{k+1} \quad (9)$$

隐层输出矩阵变为 $H_{k+1} = [H_k^T \quad h_{k+1}^T]^T$, 输出向量 $T_{k+1} = [t_1, t_2, \cdots, t_k, t_{k+1}]^T$, 此时

$$\frac{I}{\lambda} + H_{k+1} H_{k+1}^T = \frac{I}{\lambda} + \begin{bmatrix} H_k \\ h_{k+1} \end{bmatrix} \begin{bmatrix} H_k^T & h_{k+1}^T \end{bmatrix} = \begin{bmatrix} H_k H_k^T + I/\lambda & H_k h_{k+1}^T \\ h_{k+1} H_k^T & h_{k+1} h_{k+1}^T + 1/\lambda \end{bmatrix} \quad (10)$$

令 $P_{k+1} = H_{k+1} H_{k+1}^T + I/\lambda$, $P_k = H_k H_k^T + I/\lambda$, $B_{k+1} = H_k h_{k+1}^T$, $D_{k+1} = h_{k+1} h_{k+1}^T + 1/\lambda$

则式(10)表示为

$$P_{k+1} = \begin{bmatrix} P_k & B_{k+1} \\ B_{k+1}^T & D_{k+1} \end{bmatrix} \quad (11)$$

上式根据分块矩阵求逆引理可得^[17]

$$P_{k+1}^{-1} = \begin{bmatrix} P_k & B_{k+1} \\ B_{k+1}^T & D_{k+1} \end{bmatrix}^{-1} = \begin{bmatrix} P_k^{-1} + s_{N+1} P_k^{-1} B_{k+1} B_{k+1}^T P_k^{-1} & -s_{N+1} P_k^{-1} B_{k+1} \\ -s_{N+1} B_{k+1}^T P_k^{-1} & s_{N+1} \end{bmatrix} \quad (12)$$

式中 $s_{N+1} = 1/(D_{k+1} - B_{k+1}^T P_k^{-1} B_{k+1})$ 。

则式(9)可表示为

$$\beta_{k+1} = H_{k+1}^T P_{k+1}^{-1} T_{k+1} \quad (13)$$

由式(12)可以看出,矩阵 P_{k+1}^{-1} 中元素均可在 P_k^{-1} 的基础上递推得到,从而避免了在新样本加入时对隐层输出矩阵反复求逆的过程,可有效节省求解 β_{k+1} 时的计算量。

随着新样本的持续加入,当样本数目超过隐层节点数目后,隐层输出矩阵变为列满秩矩阵,此时,输出权值表示变为

$$\beta_l = (H_l^T H_l + I/\lambda)^{-1} H_l^T T_l \quad (14)$$

当样本增加到一定规模,离当前时刻较远的旧样本对反映系统状态变化贡献逐步减弱,甚至会产生冗余信息影响预测结果的准确性,有必要在加入新样本后对相应的旧样本进行剔除,具体过程如下:

当新样本 (x_{l+1}, t_{l+1}) 加入时,将其加入到样本集中,则隐层输出矩阵变为 $[H_l^T \quad h_{l+1}^T]^T$, 输出向量变为 $[T_l^T \quad t_{l+1}]^T$; 同时剔除与 x_{l+1} 距离最远的旧样本 x_1 , 那么更新之后隐层输出矩阵 H_{l+1} 变为 $[h_2^T \quad h_3^T \quad \cdots \quad h_l^T \quad h_{l+1}^T]^T$, T_{l+1} 变为 $[t_2 \quad t_3 \quad \cdots \quad t_l \quad t_{l+1}]^T$ 。则更新之后输出权值为

$$\beta_{l+1} = (H_{l+1}^T H_{l+1} + I/\lambda)^{-1} H_{l+1}^T T_{l+1} \quad (15)$$

其中

$$H_{l+1}^T H_{l+1} = \begin{bmatrix} h_2^T & h_3^T & \cdots & h_l^T & h_{l+1}^T \end{bmatrix} \cdot \begin{bmatrix} h_2^T & h_3^T & \cdots & h_l^T & h_{l+1}^T \end{bmatrix}^T = h_2^T h_2 + h_3^T h_3 + \cdots + h_l^T h_l + h_{l+1}^T h_{l+1} = H_l^T H_l + h_{l+1}^T h_{l+1} - h_1^T h_1 \quad (16)$$

$$H_{l+1}^T T_{l+1} = \begin{bmatrix} h_2^T & h_3^T & \cdots & h_l^T & h_{l+1}^T \end{bmatrix} \cdot \begin{bmatrix} t_2 & t_3 & \cdots & t_l & t_{l+1} \end{bmatrix}^T = H_l^T T_l + h_{l+1}^T t_{l+1} - h_1^T t_1 \quad (17)$$

借鉴 OS-ELM 做法,令 $P_l = (H_l^T H_l + I/\lambda)^{-1}$, $P_{l+1} = (H_{l+1}^T H_{l+1} + I/\lambda)^{-1}$, 那么式(16)可表示为

$$P_{l+1}^{-1} = P_l^{-1} + h_{l+1}^T h_{l+1} - h_1^T h_1 \quad (18)$$

令 $W_{l+1}^{-1} = P_l^{-1} + h_{l+1}^T h_{l+1}$, 则式(18)变为

$$P_{l+1}^{-1} = W_{l+1}^{-1} - h_1^T h_1 \quad (19)$$

根据 Sherman-Morrison 矩阵求逆引理可得^[17]

$$\begin{aligned} \mathbf{W}_{l+1} &= (\mathbf{P}_l^{-1} + \mathbf{h}_{l+1}^T \mathbf{h}_{l+1})^{-1} = \\ &\mathbf{P}_l - \frac{\mathbf{P}_l \mathbf{h}_{l+1}^T \mathbf{h}_{l+1} \mathbf{P}_l}{1 + \mathbf{h}_{l+1}^T \mathbf{P}_l \mathbf{h}_{l+1}} \end{aligned} \quad (20)$$

$$\begin{aligned} \mathbf{P}_{l+1} &= (\mathbf{W}_{l+1}^{-1} - \mathbf{h}_1^T \mathbf{h}_1)^{-1} = \\ &\mathbf{W}_{l+1} + \frac{\mathbf{W}_{l+1} \mathbf{h}_1^T \mathbf{h}_1 \mathbf{W}_{l+1}}{1 - \mathbf{h}_1^T \mathbf{W}_{l+1} \mathbf{h}_1} \end{aligned} \quad (21)$$

将式(17)、(20)和(21)代入式(15)中,可得训练样本更新后输出权值为

$$\begin{aligned} \beta_{l+1} &= \mathbf{P}_{l+1} \mathbf{H}_{l+1}^T \mathbf{T}_{l+1} = \\ &\mathbf{P}_{l+1} (\mathbf{P}_l^{-1} \mathbf{P}_l \mathbf{H}_l^T \mathbf{T}_l + \mathbf{h}_{l+1}^T \mathbf{t}_{l+1} - \mathbf{h}_1^T \mathbf{t}_1) = \\ &\mathbf{P}_{l+1} (\mathbf{P}_l^{-1} \beta_l + \mathbf{h}_{l+1}^T \mathbf{t}_{l+1} - \mathbf{h}_1^T \mathbf{t}_1) = \\ &\beta_l - \mathbf{P}_{l+1} \mathbf{h}_{l+1}^T (\mathbf{h}_{l+1} \beta_l - \mathbf{t}_{l+1}^T) + \mathbf{P}_{l+1} \mathbf{h}_1^T (\mathbf{h}_1 \beta_l - \mathbf{t}_1^T) \end{aligned} \quad (22)$$

若不对旧样本进行剔除,则只需利用式(20)以及式(22)中前两项来进行输出权值的更新即可,此时,式(22)中 $\mathbf{P}_{l+1}$ 用 $\mathbf{W}_{l+1}$ 替代。

则本文提出的样本规模驱动的结构自适应序贯正则极端学习机的具体步骤如下:

Step 1: 根据初始样本数目 N_0 选取合适隐层节点数 L (一般取 $L > N_0$), 利用式(6)计算初始输出权值 β_0 , 令新到样本数目 k 初始值为 0。

Step 2: 当第 $k+1$ 个新样本到来, 若样本数目 $N < L$, 则将新样本加入样本集, 并根据式(12)、式(13)更新 $\mathbf{P}_{k+1}$ 及输出权值 β_{k+1} , 否则转入 Step 3。

Step 3: 若样本数目 $N > L$, 但 $N \leq \eta$ (η 为预先设置的样本容量阈值), 则只将新样本加入样本集, 不剔除旧样本, 并根据式(20)和(22)更新 $\mathbf{W}_{l+1}$ 及 β_{l+1} ; 若样本数目 $N > L$ 且 $N > \eta$, 则将新样本加入样本集的同时剔除旧样本, 并根据式(21)和(22)更新 $\mathbf{P}_{l+1}$ 及 β_{l+1} 。

Step 4: 令 $k=k+1$, 转至 Step 2。

Step 5: 训练结束, 利用最终的输出权值建立预测模型, 对系统未来状态进行预测。

4 实例验证

分别选取 3 种混沌时间序列: Lorenz, Mackey-Glass 以及太阳黑子时间序列对 SA-SRELM 的预测性能进行验证。其中前 2 种混沌时间序列分别选择 2000 个数据, 嵌入维数分别取 5 和 4, 时间延迟都设为 1, 利用前 1500 个数据来训练模型, 后 500 个数据验证预测模型精度; 太阳黑子选择 1902~2001 年共 100 个数据, 选择嵌入维数为 5, 时间延迟为 1, 利用前 92 个数据滚动生成训练样本来建立预测模型, 对后 8 个数据进行预测。预测效果通过两个指标进行比

较: 平均绝对百分比误差 (MAPE) 和均方根误差 (RMSE)。3 种时间序列预测时初始样本数都设为 5 个。其中 Lorenz 时间序列由式(23)描述

$$\begin{cases} \frac{dx_L(t)}{dt} = 10(y_L(t) - x_L(t)) \\ \frac{dy_L(t)}{dt} = 28x_L(t) - y_L(t) - x_L(t)z_L(t) \\ \frac{dz_L(t)}{dt} = x_L(t)y_L(t) - \frac{8z_L(t)}{3} \end{cases} \quad (23)$$

Mackey-Glass 时间序列由式(24)描述

$$\frac{dx_M(t)}{dt} = \frac{0.2x_M(t-20)}{1 + [x_M(t-20)]^{10}} - 0.1x_M(t) \quad (24)$$

RELM, SRELM 的神经元数目受限于初始样本规模, 因此都取 5 个, 而 SA-SRELM 的神经元数目要受样本规模限制, 本文取 20 个, 同时样本规模阈值设为 30, 即当训练样本在线训练时增加到 30 个后启动剔除功能, 以消除旧样本的影响, 其正则化参数分别取 $\lambda = 2^{10}, 2^{20}, 2^{30}$ 。表 1 为分别利用 RELM, SRELM 以及 SA-SRELM 对混沌时间序列的预测结果。

Table 1 RELM-based, SRELM-based and DR-ELM-based prediction results of chaotic time series

Data	Method	Sigmoid		Radial Basis	
		MAPE/%	RMSE	MAPE/%	RMSE
Lorenz-X	RELM	16.65	0.6766	32.3	5.2034
	SRELM	5.96	0.2795	8.59	0.6408
	SA-SRELM	0.89	0.0512	2.6	0.1639
Mackey-Glass	RELM	15.63	0.1576	10.94	0.1135
	SRELM	2.35	0.0248	2.96	0.0347
	SA-SRELM	0.32	0.004	0.4	0.0051
Sunspot	RELM	24.8	11.6822	28.28	12.0048
	SRELM	16.31	7.4752	21.17	10.9023
	SA-SRELM	7.9	5.8897	8.51	4.7405

可以看出, 三种方法中, SA-SRELM 预测精度最高, RELM 预测精度最低。RELM 预测平均相对误差以及均方根误差最大, 预测精度最低, 这是因为 RELM 由于不具备在线学习能力, 仅能利用 5 个初始训练样本建立离线训练模型, 缺乏对时间序列的动态适应性; SRELM 由于具有在线学习能力, 预测精度相对 RELM 有所提高, 但是由于受限于初始样本规

模,其隐层节点数目最多只能取5个,导致其泛化性能较差;而本文提出的SA-SRELM则不受初始训练样本规模限制,在初始小样本情况下隐层节点数目可以任意选取,本文在选取隐层节点数目为20的情况下就大大提高了对混沌时间序列的预测精度,平均相对误差和均方根误差均远小于RELM和SRELM方法。而且从激活函数选取上可以看出,三种方法Sigmoid型激活函数的预测精度均要高于径向基型激活函数。

5 航空发动机排气温度预测

在航空公司的日常监控中,经发动机生产厂家(Original engine manufacturer,OEM)开发的发动机健康状态监控软件(Engine health monitoring,EHM)提供的巡航状态下的 ΔEGT 是衡量发动机是否超温的主要指标。因此,对 EGT 的预测实际上是对 ΔEGT 的预测^[18]。

本文 ΔEGT 数据来源于国航某发动机巡航数据^[18]。采样间隔约为15个飞行循环,可近似地认为是等间隔采样。从该发动机第1个飞行循环到第12501个飞行循环,共采集101个数据, ΔEGT 变化曲线如图1所示。

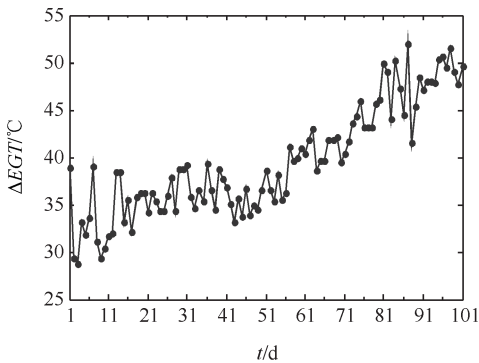


Fig. 1 Sample curve of ΔEGT

选择嵌入维数为6,则可滚动生成95组样本,选择前89组作为训练样本,其余6组作为测试样本。选取初始训练样本数目为5,即RELM仅利用前5组训练样本建立预测模型,SRELM以及SA-SRELM则利用其余84组训练样本对SRELM以及SA-SRELM进行在线更新。三种方法激活函数选择Sigmoid型,并通过留一法交叉验证确定RELM和SRELM隐层节点数目分别为4和5,SA-SRELM隐层节点数目为28,正则化系数 $\lambda=2^{10}$ 。SA-SRELM样本规模阈值设为35。采用一步预测方式,对每种方法分别进行10次仿真,取精度最高结果进行比较。图2和图3为

各方法预测结果及误差曲线,表4~6为预测值及误差结果。

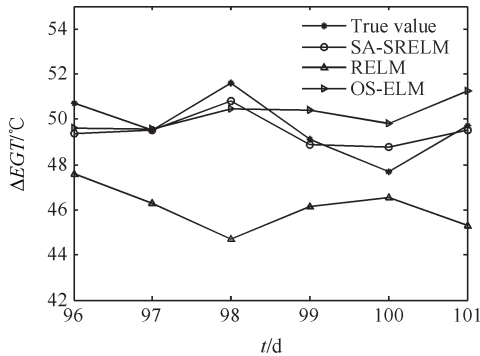


Fig. 2 Prediction result curves of three methods

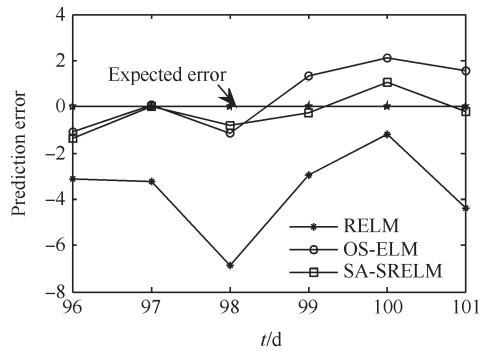


Fig. 3 Prediction error curves of three methods

Table 4 RELM based ΔEGT prediction results

Data point	True value/°C	Prediction value/°C	MAPE/%
96	50.7	47.5867	6.1406
97	49.5	46.2691	6.5271
98	51.6	44.7151	13.3428
99	49.1	46.1219	6.0654
100	47.7	46.5171	2.4799
101	49.7	45.3024	8.8483
MAPE/%	7.23		
RMSE	4.0237		

Table 5 SRELM based ΔEGT prediction results

Data point	True value/°C	Prediction value/°C	MAPE /%
96	50.7	49.6201	2.1300
97	49.5	49.553	0.1071
98	51.6	50.4618	2.2058
99	49.1	50.4133	2.6747
100	47.7	49.809	4.4214
101	49.7	51.2667	3.1523
MAPE/%	2.45		
RMSE	1.3597		

Table 6 SA-SRELM based ΔEGT prediction results

Data point	True value/ $^{\circ}\text{C}$	Prediction value/ $^{\circ}\text{C}$	MAPE /%
96	50.7	49.3563	2.6503
97	49.5	49.5220	0.00444
98	51.6	50.7972	1.5558
99	49.1	48.8504	0.5084
100	47.7	48.7574	2.2168
101	49.7	49.5138	0.3746
MAPE/%		1.23	
RMSE		0.7816	

从预测结果可以看出,RELM 预测精度最差,其相对误差达到了 7.23%,其中第 98 个测试数据相对误差达到 13.3428%;SRELM 通过在线学习,其预测结果相对误差相对 RELM 有所提高,其中第 100 个数据的预测相对误差最高,达到了 4.4214%;本文提出的 SA-SRELM 预测精度最高,为 1.23%,各数据预测相对误差都小于 3%,第 96 个数据预测结果相对误差最大,为 2.6503%,精度明显好于其余两种方法。

表 7 为三种方法所消耗的训练时间,可以看出,RELM 由于只利用 5 个初始样本进行训练,因此训练消耗时间最少;SRELM 和 SA-SRELM 采取了递推更新的方式,训练时间有所增加,而 SA-SRELM 由于选取的隐层节点数目大于初始训练样本数目,使得训练时间比 SRELM 略有增加,但其预测精度得到了有效提高。

Table 7 Training time of three methods

Method	RELM	SRELM	SA-SRELM
Training time/ms	0.19	3.1	6.9

为进一步测试 3 种方法的预测效果,分别选择测试样本数目为 10,20 和 40 三种情况进行仿真验证,参数及仿真条件设置与前述一致。表 8 为不同测试样本数目条件下的预测误差结果。可以看出,三种方法的预测误差随着测试样本数目的增加逐渐变大,其中 RELM 随着测试样本数目的增多,逐渐丧失预测能力,不适合用来建立 ΔEGT 时间序列的预测模型;而本文 SA-SRELM 在三种不同测试样本数目条件下的 MAPE 和 RMSE 均明显小于 RELM 和 SRELM,说明了本文针对不同样本数目选择不同的在线更新方式以及更新过程中采取类似新陈代谢策略能够有效提高预测模型的动态适应性,可用于 ΔEGT 时间序列的在线预测。

Table 8 Prediction results under different testing sample numbers

N_{test}	Error indexes	RELM	SRELM	SA-SRELM
10	MAPE/%	33.86	3.17	2.23
	RMSE	16.9039	1.675	1.2639
20	MAPE/%	58.68	4.57	3.96
	RMSE	27.9812	2.8526	2.3751
40	MAPE/%	94.94	9.8	5.01
	RMSE	44.8965	5.2639	3.0597

6 结 论

(1) 针对 RELM 算法无法满足在线预测要求,以及 SRELM 在线训练过程中隐层节点数目受限于初始训练样本规模限制导致预测模型泛化性能差的缺陷,提出了一种可以根据样本规模自适应选择输出权值计算方法的结构自适应序贯正则极端学习机。

(2) 通过引入正则化参数来克服隐层输出矩阵奇异导致的结果不稳定,在此基础上,根据初始训练以及序贯学习阶段训练样本数目,分别通过不同的方式计算其输出权值,从而实现在初始阶段选择更大范围的隐层节点数目,摆脱训练样本数目的限制,而在训练样本数目增加到一定规模时,通过对旧样本的剔除来提高预测模型追踪系统动态变化的能力。

(3) EGT 是航空发动机重要的健康指标之一,将本文所提方法用于航空发动机排气温度预测,结果表明,SA-SRELM 在初始小样本情况下获得的预测精度相比 RELM 和 SRELM 方法分别提高了约 6 倍和 2 倍,具有一定的应用价值。

参考文献:

[1] Vingerhoeds R A, J anssens P, Netten B D, et al. Enhancing Off- Line and On- Line Condition Monitoring and Fault Diagnosis[J]. *Control Engineering Practice*, 1995, 3 (11): 1515-1528.

[2] Ding G, Zhong S S. Aircraft Engine Lubricating Oil Monitoring by Process Neural Network[J]. *Neural Network World*, 2006, 16(1): 15-24.

[3] 陈 果. 用结构自适应神经网络预测航空发动机性能趋[J]. *航空学报*, 2007, 28(3): 535-539.

[4] 钟诗胜, 雷 达, 丁 刚. 卷积和离散过程神经网络

- 及其在航空发动机排气温度预测中的应用[J]. 航空学报, 2012, 33(3): 438-445.
- [5] 许喆平, 郎荣玲, 邓小乐. 飞机性能参数预测的不确定性处理[J]. 航空学报, 2012, 33(6): 1100-1107.
- [6] 于广滨, 丁 刚, 姚 威, 等. 基于支持过程向量机的航空发动机排气温度预测[J]. 电机与控制学报, 2013, 17(8): 30-36.
- [7] 范 庚, 马登武. 基于组合优化相关向量机的航空发动机性能参数概率预测方法[J]. 航空学报, 2013, 34(9): 2110-2121.
- [8] Huang G B, Zhu Q Y, Siew C K. Extreme Learning Machine: Theory and Applications [J]. *Neurocomputing*, 2006, 70: 489-501.
- [9] Yu Q, Miche Y, Eirola E, et al. Regularized Extreme Learning Machine for Regression with Missing Data [J]. *Neurocomputing*, 2013, 102: 45-51.
- [10] Shrivastava N A, Panigrahi B K. A Hybrid Wavelet-ELM Based Short Term Price Forecasting for Electricity Markets [J]. *Electrical Power and Energy Systems*, 2014, 55: 41-50.
- [11] Huang G B, Ding X J, Zhou H M. Optimization Method Based Extreme Learning Machine for Classification [J]. *Neurocomputing*, 2010, 74: 155-163.
- [12] Huang G B, Zhou H M, Ding X J, et al. Extreme Learning Machine for Regression and Multiclass Classification [J]. *IEEE Trans on Systems, Man, and Cybernetics-part B: Cybernetics*, 2012, 42(2): 513-529.
- [13] 邓万宇, 郑庆华, 陈 琳, 等. 神经网络极速学习方法研究[J]. 计算机学报, 2010, 3(2): 279-287.
- [14] 张 弦, 王宏力. 基于贯序正则极端学习机的时间序列预测及其应用[J]. 航空学报, 2011, 32(7): 1302-1308.
- [15] 张 弦, 王宏力. 利用神经元拓展正则极端学习机预测时间序列[J]. 北京航空航天大学学报, 2011, 37(12): 1510-1514.
- [16] Liang N Y, Huang G B, Saratchandran P, et al. A Fast and Accurate On-Line Sequential Learning Algorithm for Feedforward Networks [J]. *IEEE Trans on Neural Networks*, 2006, 17(6): 1411-1423.
- [17] 张贤达. 矩阵分析与应用[M]. 北京: 清华大学出版社, 2005.
- [18] 栾圣罡. 基于气路参数样本的航空发动机状态监视方法与系统研究[D]. 哈尔滨: 哈尔滨工业大学, 2008.

(编辑:张荣莉)