

基于Adaboost算法的回声状态网络预报器

韩 敏, 穆大芸

(大连理工大学 电子信息与电气工程学部, 辽宁 大连 116023)

摘要: 把单个回声状态网络(echo state network, ESN)的预测模型作改进, 对整体ESN预测精度的提高是有限的. 针对以上问题, 本文考虑整体ESN. 首先利用Adaboost算法提升单个ESN的泛化性能及预测精度, 并且根据Adaboost算法的结果, 建立一种ESN预报器(Adaboost ESN, ABESN). 这个ESN预报器根据拟合误差不断修正训练样本的权重, 拟合误差越大, 训练样本权重值就越大; 因此, 它在下一次迭代时, 就会侧重在难以学习的样本. 把单个ESN的预测模型经过加权, 然后按照加法组合在一起, 形成最终的ESN预测模型. 将该预测模型应用于太阳黑子、Mackey-Glass时间序列的预测研究, 仿真结果表明所提出的预测模型在实际时间序列预测领域的有效性.

关键词: ESN; Adaboost. RT算法; 非线性时间序列; 预测

中图分类号: TP273 **文献标识码:** A

Improvement of echo state network accuracy with Adaboost

HAN Min, MU Da-yun

(Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian Liaoning 116023, China)

Abstract: Modifying the prediction model of individual echo state network(ESN) improves the total prediction result with limited extent. To solve this problem, we consider an ensemble of ESN. The general performance and prediction accuracy of each individual ESN is boosted by using the Adaboost algorithm. Based on the Adaboost algorithm results, we develop an ESN predictor(ABESN). In this predictor, the weights of training samples are constantly adjusted according to the fitting error, the greater the fitting error, the heavier the weights for the training samples. Therefore, the ESN predictor will focus on the hard-learning samples in the next iteration cycle. The prediction models of individual ESN are weighted and added up to form the final predictor of the ensemble of ESN. The presented model is tested on the benchmark prediction problem of Mackey-Glass time series as well as the time series of sunspots. Simulation results demonstrate its high prediction accuracy and effectiveness.

Key words: ESN; Adaboost. RT algorithm; nonlinear time series; prediction

1 引言(Introduction)

非线性时间序列预测与控制是时间序列分析在工程中应用的一个重要方面. 而最佳控制是在精确预测的基础上进行的, 对非线性时间序列预测精度的提高研究具有重要的意义^[1]. 由于神经网络具有较强的非线性映射能力以及能以任意精度逼近任意非线性函数的特点^[2], 许多研究者利用神经网络对非线性时间序列建模预测展开了研究^[3,4].

2004年, Herbert Jaeger在《科学》中提出一种新的递归神经网络, 即回声状态网络(echo state network, ESN)^[5], 该网络学习算法及稳定性较传统递归神经网络有了较大改善. 因此在非线性时间序列的建模与预测研究中, 预测精度有了较大幅度的提高^[6~8].

尽管单一ESN预测模型的改进和学习算法的完善在一定程度上能够提高预测精度、减少误差, 但是由于非线性时间序列的复杂性以及网络训练算法的局限性等方面限制, 单一ESN预测模型不能在更大的范围内得以推广. 针对以上问题, 本文利用Adaboost算法提升ESN网络泛化性能和学习精度, 即根据拟合误差获得训练样本的权重分布, 但是每一次迭代循环后, 训练样本的权重会根据拟合误差不断地校正, 训练样本的拟合误差越大, 其权重值越大, 由此ESN网络每次迭代时会重点强调学习困难的样本, 由每次迭代产生的ESN网络组合成ABESN预测模型. 最后将该模型应用于太阳黑子和Mackey-Glass时间序列的预测中, 与单一ESN预测模型相比, 得到了较好的预测效果.

2 回声状态网络预测模型(Prediction model of echo state networks)

ESN是由输入层、状态储备池、输出层组成. 其网络结构如图1所示. 基本方程可以写作:

$$\begin{cases} x(t+1) = \tanh(W_x x(t) + W_{in} u(t)), \\ y(t+1) = W^T x(t). \end{cases} \quad (1)$$

其中: $x(t)$, $u(t)$, $y(t)$ 分别代表 t 时刻 ESN 的状态变量、输入变量和输出变量, W_x 和 W_{in} 分别为储备池内部的连接矩阵和外部输入与储备池神经元的连接矩阵, W 为储备池和网络输出的连接权值. \tanh 为 Sigmoid 函数. 给定训练网络的输入输出样本对后, 建立储备池网络的过程如下:

Step 1 选择储备池的规模 N 、内部连接权值 W_x 的稀疏度、内部连接权值 W_x 的谱半径, 并随机初始化储备池网络;

Step 2 选择储备池网络的输入输出样本, 并用输入激发储备池的内部状态. 储备池的初始状态可以任意选择, 在网络输入的作用下, 计算并记录各个时刻储备池的状态变量大小;

Step 3 根据储备池状态变量和期望输出变量之间的回归关系, 计算储备池的输出权值向量 W .

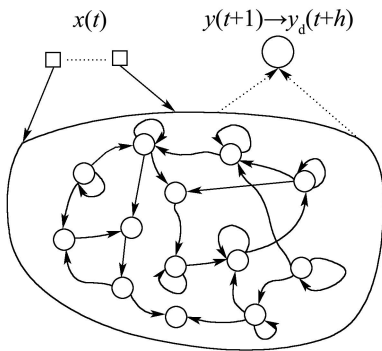


图1 ESN网络结构图
Fig. 1 Structure of ESN

ESN输出权值确定可以通过最小化以下的指标函数来实现:

$$\min_{\hat{W}} \|A \cdot \hat{W} - Y_d\|, \quad (2)$$

其中 $A = [x(\text{Init}), x(\text{Init} + 1), \dots, x(T_m)]^T$, \hat{W} 是对 W 的估计,

$$Y_d = [y_d(\text{Init}), y_d(\text{Init} + 1), \dots, y_d(T_m)],$$

Init 一般不从 1 开始, 以去掉初始暂态的影响. T_m 是训练样本最后一个时间点, 所以有效的训练样本个数为 $T_n = T_m - \text{Init} + 1$.

3 ABESN 预报器(Adaboost-ESN predictor)

AdaBoost 算法是一种自适应提升方法, 应用于分类问题中被证明是一种有效的方法^[9]. Adaboost. RT

(R 代表回归; T 代表阈值)^[9] 是将提升方法应用于回归问题, 并引入了阈值 ϕ 以产生有效系统^[10], 本文利用此算法对 ESN 网络学习能力进行提升, 以提高网络的预测性能.

3.1 Adaboost. RT 基本算法(Adaboost. RT Algorithm)

Adaboost. RT 算法的基本步骤具体如下:

1) 输入. 样本训练集为时间序列 $(x_1, y_1), \dots, (x_m, y_m)$, 这里输出 $y \in \mathbb{R}$; 确定弱学习算法或弱学习器; 确定最大迭代次数; 确定用以判断预报值正确与否的阈值 ϕ ($0 < \phi < 1$).

2) 初始化. $t = 1$ 时, 权重分布 $D_t(i) = \frac{1}{m}$, 初始误差 $\varepsilon_t = 0$.

3) 迭代. 依照权重分布训练弱学习器, 建立回归模型 $f_t(x) \rightarrow y$, 计算训练集的误差:

$$ARE_t(i) = \left| \frac{f_t(x_i) - y_i}{y_i} \right|. \quad (3)$$

计算 $f_t(x)$ 的误差率:

$$\varepsilon_t = \sum_{i: ARE_t(i) > \phi} D_t(i). \quad (4)$$

计算 $\beta_t = \varepsilon_t^n$ ($n=1, 2, 3$). 更新权重 D_t :

$$D_{t+1}(i) = \begin{cases} \frac{D_t(i)}{Z_t} \times \beta_t, & ARE_t(i) \leq \phi, \\ \frac{D_t(i)}{Z_t} \times 1, & \text{其他}. \end{cases} \quad (5)$$

这里 Z_t 为标准化因子, 令 $t = t + 1$.

4) 输出.

$$f_{\text{fin}}(x) = \frac{\sum_t (\lg \frac{1}{\beta}) f_t(x)}{\sum_t (\lg \frac{1}{\beta})}. \quad (6)$$

3.2 基于 Adaboost 算法的 ESN 预测模型(ABESN prediction model)

在时间序列预测研究领域, 对精度的提高具有重要的意义. 新型递归神经网络 ESN 输出权值训练方法简单, 网络稳定性也有保证, 其应用于非线性时间序列的建模与预测研究中也具有较高的预测精度. 尽管单个 ESN 预测模型的改进和学习算法的完善在一定程度上能够提高预测精度、减少误差, 然而由于非线性时间序列的复杂性以及网络训练算法的局限性等方面限制, 单一 ESN 预测模型不能在更大的范围内推广.

如果把单个预测模型按照一定的方法组合起来, 便能进一步改进预测效果, 获得更高的预测精度^[12]. 本文利用 Adaboost 算法对 ESN 网络泛化性能和预测性能进行提升, 首先初始化每个训练样本的权重分布 D_t 为 $\frac{1}{m}$ (m 为样本个数), 利用已知样

本训练ESN网络, 每一次迭代循环后, 计算训练误差 $ARE_t(i)$ 和误差率 ε_t , 根据式(5)更新权重 D_t , 训练样本的权重 D_t 会根据训练误差不断地校正, 训练误差越大, 此样本点权重值越大, 由此ESN网络每次迭代时会重点强调学习困难的样本, 由每次迭代产生的ESN网络加权平均组合成ABESN预测模型, 模型的算法具体流程如下:

开始

输入: $(x_1, y_1), \dots, (x_m, y_m), x, y \in \mathbb{R}$;

初始化训练样本权重: $D_t(i) = \frac{1}{m}$

For $t = 1, \dots, T$

根据权重 $D_t(i)$ 选择训练数据, 训练ESN网络;

得到训练后的ESN网络: $f_t(x)$;

计算训练误差: $ARE_t(i) = \left| \frac{f_t(x_i) - y_i}{y_i} \right|$;

通过 $f_t(x)$ 的误差率 $\varepsilon_t = \sum_{i: ARE_t(i) > \phi} D_t(i)$, 计算

$\beta_t = \varepsilon_t^n$;

根据式(5)更新权重 $D_{t+1}(i)$;

End loop;

由式(6)得到ABESN模型输出 $f_{fin}(x)$;

结束

为分析数值试验结果的优劣, 本文采用均方根误差 E_{rmse} 作为性能指标, 对预测结果进行比较, 定量说明各个模型的预测效果. 设 $y(t)$ 为预测变量的实际观测值, 为该变量的网络预测值, N 对应训练的样本数, 则均方根误差定义为:

$$E_{rmse} = \left[\frac{1}{N-1} \sum_{k=1}^N [\hat{y}(k) - y(k)]^2 \right]^{\frac{1}{2}} \quad (7)$$

均方根误差反映了预测值相对于观测值的偏离程度, 取值大于或等于零, 预测无误差时等于零. 均方根误差值越小, 则预测效果越好, 预测精度越高.

4 仿真实例(Simulations)

采用实际时间序列年太阳黑子对ABESN预测模型进行仿真研究, 为了进一步验证本文所提预测模型的有效性, 又对Mackey-Glass时间序列进行了短期预测, 从而证实ABESN模型具有更强的预测性能, 能更好的应用于时间序列建模预测的研究中.

4.1 太阳黑子数据仿真(Simulation of sunspots data)

将ABESN预测模型应用于太阳黑子时间序列的预测中, 仿真数据为1700年到2003年太阳黑子时间序列. 预测模型中储备池的维数为100, 内部连接权值的稀疏度为0.05, 其谱半径为1, 其中Adaboost中的阈值 ϕ 取0.4, 根据预测误差确定迭代次数为20. 利用前250组数据对模型进行训练, 其余数据进行预测, 预测结果如图2所示.

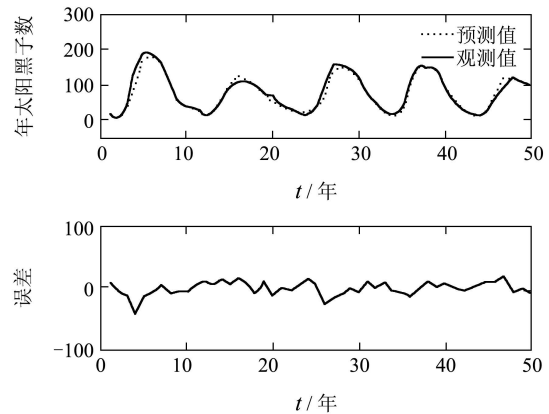


图2 ABESN模型预测太阳黑子预测值、观测值及误差
Fig. 2 Sunspot predictive results using ABESN

分别采用ABESN预测模型及单一ESN预测模型对太阳黑子时间序列进行预测研究, 多次仿真实验得到的均方根误差结果见表1. 由表1结果对比可知, 在用于太阳黑子时间序列预测时, 文[5]中传统单一的ESN预测模型的均方根误差为18.5573, 而ABESN预报器的预测精度提高近三分之一, 与文[8]中改进的单一ESN网络相比, 预测精度也有了一定的提高, 仿真结果证明ABESN预报器具有较好的学习能力.

表1 不同预测模型的结果比较

Table 1 Comparison of different predictive model

模型	ABESN	ESN ^[5]	SVESM ^[8]
E_{rmse} (太阳黑子)	12.0353	18.5573	15.4040

4.2 Mackey-Glass时间序列预测(Prediction of Mackey-Glass time series)

为了验证本文所提出的ABESN预报器对时间序列预测的有效性, 采用Mackey-Glass微分方程产生的时间序列进行仿真研究, 产生长度1000的时间序列, 其中前800组数据用来训练, 剩余数据用于测试. W_x 谱半径和稀疏度分别设为1和0.02, Adaboost中阈值 ϕ 取0.3, 迭代次数为20. 图3是利用ABESN预报器对Mackey-Glass时间序列进行预测的结果图.

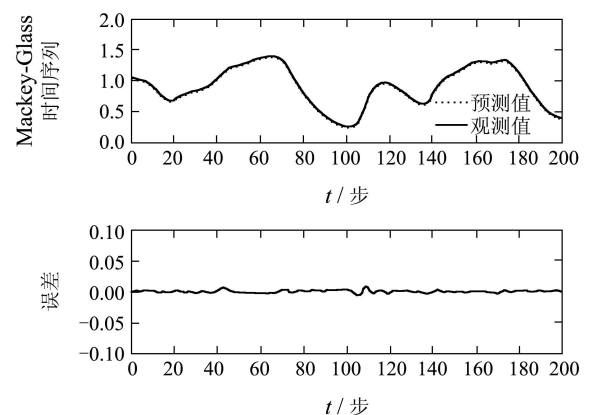


图3 ABESN模型预测Mackey-Glass时序预测值、观测值及误差
Fig. 3 Mackey-Glass predictive results using ABESN

表2中基于ESN迭代的预测精度最高, 本文方法达到的预测精度不及ESN的迭代预测, 而对于其他已有算法, 本文方法具有较大的优势. 无核支持向量机(SVESM)中储备池维数选择1200, 训练样本长度为2200, 所得到的预测误差为0.0085; 基于自组织特征映射(SOM)方法的结果为0.035, 其对应的训练样本长度为1001个; 对于MLPs算法, 其学习样本的长度为500, 所达到的预测误差仅为0.060.

表2 不同模型的预测结果误差对比
Table 2 Errors comparisons of different models

模型	描述	E_{rmse}
ABESN	储备池维数300, 训练长度800	0.00220
ESNs ^[5]	储备池维数1000, 训练长度3000	0.00063
SVESM ^[8]	储备池维数1200, 训练长度2200	0.00850
SOM	训练长度1001	0.03500
MLPs	训练长度500	0.06000

5 结论(Conclusions)

本文提出了一种新型时间序列预测模型ABESN预报器, 即利用Adaboost.RT算法提升新型递归神经网络ESN学习能力, 以提高其预测精度, 通过应用于Mackey-Glass时间序列和太阳黑子实际数据预测, 实例证明, 可得出以下结论:

该模型是研究时间序列预测的一种尝试, 理论上, Adaboost.RT算法根据训练误差将不断修改的权值赋予单个ESN预测模型, 再将单个模型加权按照加法组合在一起, 这样的组合对于提高单一ESN预测模型的学习能力起到了很大的帮助; 通过对Mackey-Glass时间序列、太阳黑子实际数据的短期预测, 与单一ESN网络预测结果相比有了一定提高, 从而证实本文所提出的ABESN网络具有更强的预测性能, 能更好的应用于时间序列建模预测的研究中.

参考文献(References):

- [1] 杨叔子, 吴雅. 时间序列分析的工程应用[M]. 武汉: 华中理工大学出版社, 1991.
(YANG Shuzi, WU Ya. *Time Series Analysis in Engineer Applica-*

- tion*[M]. Wuhan: Huazhong University of Science Technology Press, 1994.)
- [2] HORNIK K, STINCHCOMBE M, WHITE H. Multi-layer feed-forward networks are universal approximators[J]. *Neural Networks*, 1989, 2(3): 359 – 366.
- [3] CAI X D. Time series prediction with recurrent neural networks trained by a hybrid PSO-EA algorithm[J]. *Neurocomputing*, 2007, 70(13/15): 2342 – 2353.
- [4] 张冬青, 宁宣熙, 刘雪妮. 基于RBF神经网络的非线性时间序列在线预测[J]. 控制理论与应用, 2009, 26(2): 151 – 155.
(ZHANG Dongqing, NING Xuanxi, LIU Xueni. On-line prediction of nonlinear time series using RBF neural networks[J]. *Control Theory & Applications*. 2009, 26(2): 151 – 155.)
- [5] JAEGER H, HAAS H. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication[J]. *Science*, 2004, 304(5667): 78 – 80.
- [6] LIN X W, YANG Z H, SONG Y X. Short-term stock price prediction based on echo state networks[J]. *Expert Systems with Applications*, 2009, 36(3): 7313 – 7317.
- [7] GE Q, WEI C J. Multiresolution-based echo state network and its application to the longterm prediction of network traffic[C] //2008 *International Symposium on Computational Intelligence and Design*. Piscataway, N J: IEEE, 2008, 1: 469 – 472.
- [8] SHI Z W, HAN M. Support vector echo state machine for chaotic time series prediction[J]. *IEEE Transactions on Neural Networks*, 2007, 18(2): 359 – 372.
- [9] FREUND Y, SCHAPIRE R E. A decision-theoretic generalization of on-line learning and an application to boosting[J]. *Journal of Computer and System Sciences*, 1997, 55(1): 119 – 139.
- [10] SOLOMATINE D P, SHRESTHA D L. AdaBoost. RT: a boosting algorithm for regression problems[C] //2004 *IEEE International Joint Conference on Neural Networks*. New York: IEEE, 2004, 2: 1163 – 1168.
- [11] SHRESTHA D L, SOLOMATINE D P. Experiments with AdaBoost. RT, an improved boosting scheme for regression[J]. *Neural Computation*, 2006, 18(7): 1678 – 1710.
- [12] GAVIN B, JEREMY W, RACHEL H, et al. Diversity creation methods: A survey and categorization[J]. *Information Fusion Journal, Special Issue on Diversity in Multiple Classifier Systems*, 2005, 6(1): 5 – 20.

作者简介:

韩敏 (1959—), 女, 教授, 博士生导师, 研究方向为神经网络、混沌时间序列分析、专家系统及3S系统, E-mail: minhan@dlut.edu.cn;

穆大芸 (1985—), 女, 硕士研究生, 研究方向为多变量时间序列建模与预测, E-mail: dymu@mail.dlut.edu.cn.