

文章编号: 1001-0920(2012)11-1681-04

混沌对角递归神经网络的船舶横摇预报方法

李占英^{1,2}, 王科俊¹, 徐亮³, 姚丽君⁴

(1. 哈尔滨工程大学 自动化学院, 哈尔滨 150001; 2. 大连理工大学城市学院
电子与自动化学院, 辽宁 大连 116600; 3. 大连中远船务工程有限公司,
辽宁 大连 116113; 4. 天津工业大学 电气工程与自动化学院, 天津 300387)

摘要: 船舶运动在一定条件下会出现混沌特性, 因此可以利用混沌神经网络对其进行预报. 对传统的混沌对角递归神经网络模型各权值的训练进行优化, 给出了基于Lyapunov函数的各层权所通用的学习速率调整算法的收敛定理并加以证明. 仿真结果表明, 采用优化采样时刻 k 可提高各权值的精确度, 使收敛性得到改善, 能有效提高预报精度和延长预报时间. 与前向神经网络BP预测相对比, 优化后的模型具有很好的预测效果.

关键词: 对角递归神经网络; 混沌; 动量梯度学习算法; 船舶横摇预测; 前向神经网络

中图分类号: TP183

文献标志码: A

Approach of prediction of ship rolling based on chaotic diagonal recurrent neural networks

LI Zhan-ying^{1,2}, WANG Ke-jun¹, XU Liang³, YAO Li-jun⁴

(1. College of Automation, Harbin Engineering University, Harbin 150001, China; 2. School of Electronic Engineering and Automation, City Institute of Dalian University of Technology, Dalian 116600, China; 3. Cosco Dalian Shipyard Company Ltd, Dalian 116113, China; 4. School of Electrical Engineering and Automation, Tianjin Polytechnic University, Tianjin 300387, China. Correspondent: LI Zhan-ying, E-mail: lzy1979@126.com)

Abstract: There is chaotic characteristics in ship motion under certain conditions, so the chaotic diagonal recurrent neural network(CDRNN) is proposed to predict ship swaying motion. A convergence theorem of each weight learn algorithm based on Lyapunov function is given and proved. Simulation results show that, the value of the optimized sampling time k is applied to increase the accuracy of all of the weight, which improves algorithm convergence, and the predicted precision and the forecast time are advanced efficiently. The optimized modeling has better predictive effect using CDRNN than using feed-forward BP neural network to predict.

Key words: diagonal recurrent neural networks(DRNN); chaos; dynamic back-propagation(DBP); ship rolling prediction; feed-forward neural network

1 引言

舰船摇荡运动的极短期预报技术对于改进航态控制系统性能, 提高其武器装备系统的适应力, 动力定位, 以及海上特殊作业等都具有重大意义. 然而, 舰船运动受风、浪等各种非线性因素影响严重, 对于时间序列的分析和预测十分困难. 目前, 大量文献采用时间序列分析法^[1-2], 其基础是将船舶摇荡运动时历作为平稳随机过程来处理, 忽略了船舶运动中的非线性因素. 已有文献证明船舶运动具有混沌特性^[3-4], 使用传统的线性分析方法显然不能满足预测要求, 因此

人们将一些智能方法相结合并应用于船舶预报^[5-8]. 递归网络是一种带有反馈的网络, 这种网络通过存储内部状态使其具备映射动态特征的能力, 能更直接生动地反映系统动态特性, 从而使系统具有适应时变特性的能力, 能够更全面、更完整地对时间序列进行建模.

在各种递归神经网络中, 对角递归神经网络^[9]由于其结构简单、易于实现等特点引起了广大学者的注意, 并对其不断加以改进^[10-15]. 如文献[15]提出了混沌对角递归神经网络. 此网络将神经网络和混沌特性

收稿日期: 2011-07-05; 修回日期: 2011-11-01.

基金项目: 国家自然科学基金项目(60975022); 国家863计划项目(2008AA01Z148).

作者简介: 李占英(1979-), 女, 副教授, 从事智能控制、时间序列预测的研究; 王科俊(1962-), 男, 教授, 博士生导师, 从事人工智能控制、神经网络等研究.

相结合,满足了一定条件下船舶运动中的非线性及混沌特性;但由于在算法推导时并未明确指出所调解权值的采样时刻 k ,使算法推导上存在收敛性差的缺点.本文在原算法的基础上,详细推导了混沌对角递归神经网络的各权值学习算法,并给出了各权值的修正算法,同时,给出了基于Lyapunov函数的各层权所通用的学习速率调整算法的收敛定理并加以证明.仿真结果表明,经过本文的优化,可使系统的收敛性得到改善,能有效提高预报精度和延长预报时间.

2 混沌对角递归神经网络的学习算法收敛性证明

在动态递归神经网络的基础上,文献[15]提出了一种可以学习和逼近动态非线性的混沌神经网络模型,用于非线性映射的函数拟合.网络结构及输入输出描述可参见文献[15].

混沌对角递归神经网络的动态梯度下降法也称为动态BP算法,是在梯度下降算法的基础上引入动量因子 α , $0 < \alpha < 1$,使用前一次的修正结果来影响本次修正量, η 为学习率.

设 $y(k)$ 为系统的实际输出值, $\hat{y}(k)$ 为系统的预测值,取误差代价函数为

$$E(k) = \frac{1}{2}e(k)^2 = \frac{1}{2}[y(k) - \hat{y}(k)]^2. \quad (1)$$

在传统的各权值调节算法的推导中并没有给出其所在的采样时刻 k ,这使得在推导各权值时出现了一些不足.考虑权值采样时刻 k ,对各权值重新进行一次推导.调整算法可见文献[16],本文仅对算法的收敛性进行证明.

DBP算法是对角递归神经网络训练中应用最多的一种学习方法,它是对BP算法的改进,充分考虑了回归单元对算法的影响.网络参数通常采用梯度下降法进行调整.

用动态BP算法进行参数更新时,参数学习率 η 的选择至关重要.如果 η 较小,则算法的收敛性能够得到保证,但训练速度将非常缓慢;如果 η 太大,则算法不稳定,收敛性得不到保证.

定义离散的Lyapunov函数如式(1)所示,在训练过程中,Lyapunov函数的变化为

$$\Delta E(k) = E(k+1) - E(k) = \frac{1}{2}[e^2(k+1) - e^2(k)], \quad (2)$$

则基于Lyapunov函数的各层权所通用的学习速率调整算法有如下的收敛定理.

定理 1 令 η 为学习率,有

$$0 < \eta < \frac{2}{g_{L\max}^2}. \quad (3)$$

其中: $g_{L\max} = \max_k \|g(k)\|$,而

$$g(k) = \frac{\partial O(k)}{\partial W(k)}. \quad (4)$$

$\|\cdot\|$ 定义为欧几里德范数.设 W^I 和 W^O 分别为网络的输入层和输出层权矩阵, W^{D_1} 和 W^{D_2} 分别为网络回归层的权矩阵.有 p 个输入, q 个F神经元和 q 个B神经元,一个输出. $W(k)$ 为一个 n 维向量,是网络所有权值所组成的权值矩阵,则有

$$W = [[W^I]^T [W^T]^T [W^{D_1}]^T [W^{D_2}]^T [W^O]^T]^T, \quad (5)$$

$$W^I = [[W_1^I]^T [W_2^I]^T \cdots [W_p^I]^T]^T, \quad (6)$$

$$W^T = [W_1^T]^T, \quad (7)$$

$$W^{D_1} = [[W_1^{D_1}]^T [W_2^{D_1}]^T \cdots [W_q^{D_1}]^T]^T, \quad (8)$$

$$W^{D_2} = [[W_1^{D_2}]^T [W_2^{D_2}]^T \cdots [W_q^{D_2}]^T]^T, \quad (9)$$

$$W^O = [W_1^O]^T, \quad (10)$$

$$\eta = \begin{bmatrix} \eta^I & & & & \\ & \eta^T & & & \\ & & \eta^{D_1} & & \\ & & & \eta^{D_2} & \\ & & & & \eta^O \end{bmatrix}, \quad (11)$$

$$\Delta W = -\eta \frac{\partial E(k)}{\partial W(k)} = -\eta \frac{\partial e(k)}{\partial W(k)} e(k) = -e(k) \begin{bmatrix} \eta^I & & & & \\ & \eta^T & & & \\ & & \eta^{D_1} & & \\ & & & \eta^{D_2} & \\ & & & & \eta^O \end{bmatrix} \times \left[\begin{bmatrix} \frac{\partial e(k)}{\partial W^I(k)} \end{bmatrix}^T \begin{bmatrix} \frac{\partial e(k)}{\partial W^T(k)} \end{bmatrix}^T \begin{bmatrix} \frac{\partial e(k)}{\partial W^{D_1}(k)} \end{bmatrix}^T \times \begin{bmatrix} \frac{\partial e(k)}{\partial W^{D_2}(k)} \end{bmatrix}^T \begin{bmatrix} \frac{\partial e(k)}{\partial W^O(k)} \end{bmatrix}^T \right]^T, \quad (12)$$

$$\Delta e(k) = \left[\frac{\partial e(k)}{\partial W(k)} \right]^T \Delta W = -e(k) \left[\eta_1 \left\| \frac{\partial e(k)}{\partial W^I(k)} \right\|^2 + \eta_2 \left\| \frac{\partial e(k)}{\partial W^T(k)} \right\|^2 + \eta_3 \left\| \frac{\partial e(k)}{\partial W^{D_1}(k)} \right\|^2 + \eta_4 \left\| \frac{\partial e(k)}{\partial W^{D_2}(k)} \right\|^2 + \eta_5 \left\| \frac{\partial e(k)}{\partial W^O(k)} \right\|^2 \right], \quad (13)$$

$$\Delta E(k) = \Delta e_m(k) \left[e_m(k) + \frac{1}{2} \Delta e_m(k) \right] = \left[\frac{\partial e_m(k)}{\partial W(k)} \right]^T \eta e_m(k) \frac{\partial O(k)}{\partial W(k)} \times \left[e_m(k) + \frac{1}{2} \left[\frac{\partial e_m(k)}{\partial W(k)} \right]^T \eta e_m(k) \frac{\partial O(k)}{\partial W(k)} \right], \quad (14)$$

$$\frac{\partial e_m(k)}{\partial W(k)} = -\frac{\partial O(k)}{\partial W(k)}, \quad (15)$$

$$\Delta E(k) = -\eta e_m^2(k) \left\| \frac{\partial O(k)}{\partial W(k)} \right\|^2 + \frac{1}{2} \eta^2 e_m^2(k) \left\| \frac{\partial O(k)}{\partial W(k)} \right\|^4 = -\lambda e_m^2(k), \quad (16)$$

因此

$$\lambda = \frac{1}{2} \|g(k)\|^2 \eta (2 - \eta \|g(k)\|^2).$$

其中

$$\eta \|g(k)\|^2 = \left[\eta_1 \left\| \frac{\partial e(k)}{\partial W^I(k)} \right\|^2 + \eta_2 \left\| \frac{\partial e(k)}{\partial W^T(k)} \right\|^2 + \eta_3 \left\| \frac{\partial e(k)}{\partial W^{D_1}(k)} \right\|^2 + \eta_4 \left\| \frac{\partial e(k)}{\partial W^{D_2}(k)} \right\|^2 + \eta_5 \left\| \frac{\partial e(k)}{\partial W^O(k)} \right\|^2 \right], \quad (17)$$

$$\|g(k)\|^2 = \left\| \frac{\partial e(k)}{\partial W^I(k)} \right\|^2 + \left\| \frac{\partial e(k)}{\partial W^T(k)} \right\|^2 + \left\| \frac{\partial e(k)}{\partial W^{D_1}(k)} \right\|^2 + \left\| \frac{\partial e(k)}{\partial W^{D_2}(k)} \right\|^2 + \left\| \frac{\partial e(k)}{\partial W^O(k)} \right\|^2. \quad (18)$$

由 Lyapunov 稳定定理可知, 若要保证系统稳定, 即 $\Delta V(k) < 0$, 则需 $\lambda > 0$, 即 $0 < \eta \|g(k)\|^2 < 2$. 令 $g_{L \max} = \max_k \|g(k)\|$, 有 $0 < \eta g_{L \max}^2 < 2$. 再令 $\eta_m = \max_{i=1}^5 \{\eta_i\}$, 则有 $0 < \eta_m < \frac{2}{g_{L \max}^2}$ 成立.

3 仿真实验

以某船舶横摇运动时间序列预测为例, 数据取自我国某船舶在 90° 横浪航行情况的实验数据, 共有 1000 个数据. 通过定性和定量两方面进行混沌特性的验证.

3.1 定性分析

利用功率谱图结合三维相图, 对船舶横摇运动时间序列的混沌特性进行定性分析.

利用傅里叶变换, 可求出船舶横摇运动时间序列的功率谱图, 如图 1 所示.

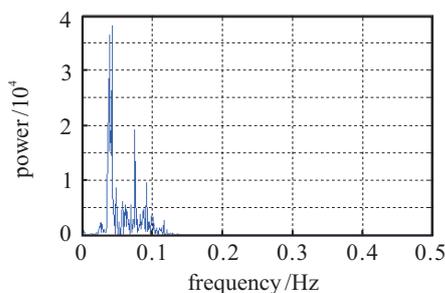


图 1 船舶横摇运动时间序列的功率谱图

根据文献 [17] 所述, 一般地, 如果功率谱图具有单峰 (或几个峰), 则对应的序列为周期 (或拟周期) 序列; 如果没有明显的峰值或峰连成一片, 则可能是混沌序列或湍流. 从图 1 可以看出, 此船舶横摇运动时

间序列样本的功率谱图没有明显的峰值或峰连成一片, 并且出现了宽带的噪声背景, 因此, 表现出一定的混沌性质.

船舶横摇运动时间序列的内部关系如图 2 所示. 由于其嵌入维数可能高于 3, 三维相图无法完整体现吸引子的复杂结构, 但其轨线的反复折叠、相互交叉, 而且被限制在有限的区域内, 此规律性也能清晰地表明奇异吸引子的存在.

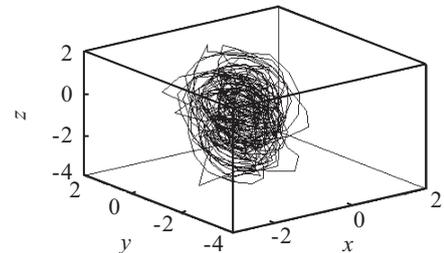


图 2 船舶横摇运动时间序列的三维相图

3.2 定量分析

通过计算混沌信号奇异吸引子的维数和 Lyapunov 特征指数来识别序列是否具有混沌性. 根据 GP (Grassberger-Procaccia) 算法计算出时间序列的关联维 d , 再由 Takens 定理选取嵌入维数 $m = 2d + 1$, 可求得 $m = 8$. 利用互信息法求得相空间重构的时间延迟 $\tau = 2$. 利用小数据量法求得平均周期为 $T = 23.5814$ s, Lyapunov 指数为 $\lambda = 0.07581 > 0$. 由此可知, 船舶横摇运动时间序列具有混沌特性.

3.3 预测仿真

分别用传统的混沌对角递归神经网络和优化后的混沌对角递归神经网络, 对船舶横摇运动时间序列进行预测仿真. 其中: 前 800 个数据用于学习, 其余数据用于预测检验模型和算法的有效性. 有 2 个输入神经元, 当前时刻测量的横摇时间序列值和前一时刻网络输出的横摇预测值; 10 个 F 神经元, 10 个 B 神经元, 10 个 H 神经元, 一个输出神经元, 为网络的横摇预测值.

为清晰显示预测效果, 图 3 给出了其中的 35 个数据的预测结果. 三者的预测误差曲线如图 4 所示.

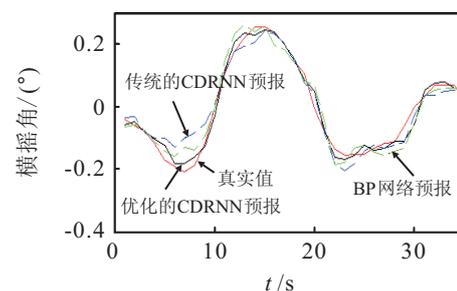


图 3 船舶横摇预测分析

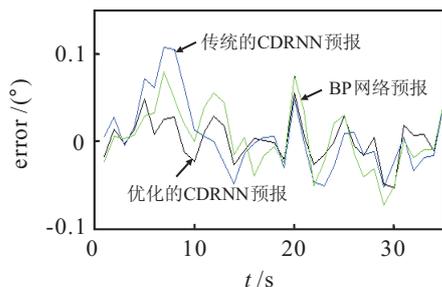


图 4 三模型的预测误差曲线

从图 4 可以看出, 利用反馈神经网络进行船舶横摇预测效果明显好于前向神经网络预测, 且使用优化后的混沌对角递归神经网络进行预测, 系统的收敛性得到了改善, 预测精度也有所提高。

为了对上述 3 种模型预测结果进行定量分析, 本文还分别利用均方根误差 (MSE) 和平均绝对误差 (MAE) 给出了 3 种模型预测的效果对比。其中

$$\text{MSE} = \sqrt{\frac{1}{T} \sum_{i=1}^T (y_{id} - y_i)^2},$$

$$\text{MAE} = \frac{1}{T} \sum_{i=1}^T (y_{id} - y_i).$$

这里: y_{id} 为预测值, y_i 为实测值, T 为预测步数。所用时间是指在同一环境下网络开始训练直至预测结束所用的时间。结果如表 1 所示。

表 1 3 种网络预测结果比较

预测方法	MSE	MAE	所用时间/s
传统 CDRNN 模型	0.029 2	0.010 8	10.457 4
优化 CDRNN 模型	0.021 8	0.002 0	8.281 0
BP 神经网络模型	0.042 2	0.004 2	134.418 7

由图 3, 图 4 及表 1 可以看出, 混沌对角递归网络的预测效果明显好于前向网络 BP 的预测效果, 延长了预测时间。优化后的网络比未优化的网络预测的精度更好, 所用的时间更少, 且预测所用的时间远远小于 BP 网络所用的时间。可见, 本文所做的优化确实改善了网络的预测精度, 加速了收敛。该网络对于船舶摇荡预测具有一定的实用性。

4 结 论

针对船舶横摇运动时间序列预测, 本文提出了一种新的预测方法。将混沌理论与对角递归神经网络相结合, 并将传统的混沌对角递归神经网络进行优化。仿真结果验证了采用优化采样时刻 k 可提高各权值的精确度, 系统的收敛性得到了改善。由于加入采样时刻 k , 修正了网络中含有前一时刻的反馈输入, 充分利用了系统的动态信息, 提高了预测的精确度, 可有效预报 10 s 左右的横摇时历, 优于前向 BP 网络及传

统的混沌对角递归神经网络, 为船舶预测提供了一个新的途径。本文方法也适用于船舶其他自由度的运动。但该网络存在训练过程复杂、记忆渐消等不足, 在未来的研究中需要进行深入研究和探讨。

参考文献(References)

- [1] Ma Jie, Li Teng, Li Guo-bin. Comparison of representative method for time series prediction[C]. Proc of the 2006 IEEE Int Conf on Mechatronics and Automation. Luoyang: IEEE, 2006: 2448-2453.
- [2] Isao Roy Yumori. Real time prediction of ship response to ocean waves using time series analysis[C]. OCEANS81 Conf. Boston: IEEE, 1981: 1082-1089.
- [3] 欧阳茹荃, 朱继懋. 船舶非线性横摇运动与混沌[J]. 中国造船, 1999, 2(1): 21-28.
(Ouyang R Q, Zhu J M. Nonlinear oscillations and chaos of ship rolling motion[J]. Ship Building of China, 1999, 2(1): 21-28.)
- [4] Cai Feng, Shi Ai-guo, Wan Lin. Chaotic dynamics analysis of ship swaying motion[J]. J of Dalian Naval Academy, 2006, 29(5): 36-39.
- [5] Zhou Bo, Shi Ai-guo. LSSVM and hybrid particle swarm optimization for ship motion prediction[C]. Int Conf on Intelligent Control and Information Processing. Dalian: IEEE, 2010: 183-186.
- [6] Fu Hui-xuan, Liu Sheng, Sun Feng. Ship motion prediction based on AGA-LSSVM[C]. Proc of the 2010 IEEE Int Conf on Mechatronics and Automation. Beijing: IEEE, 2010: 202-206.
- [7] Liu Bo, Cai Feng, Shi Ai-guo. Real ship motions' multi-steps prediction based on add-weighted LLE-based model[C]. Int Conf on Intelligent Control and Information Processing. Dalian: IEEE, 2010: 179-182.
- [8] 傅荟璇, 王宇超, 杜春洋. 基于粒子群 BP 神经网络算法舰船航向预报[C]. Proc of the 30th Chinese Control Conf. Yantai: IEEE, 2011: 2748-2752.
(Fu H X, Wang Y C, Du C Y. Ship course prediction based on PSO combined with BP neural network[C]. Proc of the 30th Chinese Control Conf. Yantai: IEEE, 2011: 2748-2752.)
- [9] Ku Chao-Chee, Lee Kwang. Diagonal recurrent neural networks for dynamic systems control[J]. IEEE Trans on Neural Networks, 1995, 6(1): 144-156.
- [10] Barbounis T G, Theocharis J B, Alexiadis M C, et al. Long-term wind speed and power forecasting using local recurrent neural network models[J]. IEEE Trans on Energy Conversion, 2006, 21(1): 273-284.

(下转第 1693 页)