

研究论文

DOI: 10.11949/j.issn.0438-1157.20170771

基于 ROLS 算法的递归 RBF 神经网络结构设计

乔俊飞^{1,2}, 马士杰^{1,2}, 杨翠丽^{1,2}

(1北京工业大学信息学部, 北京 100124; 2计算智能与智能系统北京市重点实验室, 北京 100124)

摘要: 针对递归 RBF 神经网络结构难以自适应问题, 提出一种基于递归正交最小二乘(recursive orthogonal least squares, ROLS)算法的结构设计方法。首先, 利用 ROLS 算法来计算隐含层神经元的独立贡献度和损失函数, 以此判断增加或归为不活跃组的神经元, 同时调整神经网络的拓扑结构, 并且利用奇异值分解(singular value decomposition, SVD)决定最佳的隐含层神经元个数, 以此来删除不活跃组中相对不活跃的神经元, 有效地解决了递归 RBF 神经网络结构冗余和难以自适应问题。其次, 利用梯度下降算法更新递归 RBF 神经网络的参数来保证神经网络的精度。最后, 通过对 Mackey-Glass 时间序列预测、非线性系统辨识和污水处理过程中关键水质参数动态建模, 证明了该结构设计方法的可行性和有效性。

关键词: 神经网络; 结构设计; 算法; 奇异值分解; 动态建模

中图分类号: TP 183

文献标志码: A

文章编号: 0438—1157 (2018) 03—1191—09

Structure design for recurrent RBF neural network based on recursive orthogonal least squares

QIAO Junfei^{1,2}, MA Shijie^{1,2}, YANG Cuili^{1,2}

(¹Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China; ²Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124, China)

Abstract: Aiming at the problem of recurrent radial basis function (RRBF) neural network structure which is difficult to be self-adaptive, this paper proposes a structure design method based on recursive orthogonal least square (ROLS) algorithm. Firstly, ROLS algorithm is used to calculate the contribution and the loss function of hidden layer neurons, which determines to increase or be grouped into inactive neurons, and the topology structure of neural network is adjusted accordingly. At the same time, singular value decomposition (SVD) is applied to determine the best number of hidden layer neurons in order to delete the neurons of the inactive group, which effectively solves the problems of RRBF neural network structure which is redundant and hardly self-adaptive. Secondly, the gradient descent algorithm is utilized to update the parameters of RRBF neural network in order to ensure the accuracy of neural network. Finally, several experiments including the Mackey-Glass time series prediction, nonlinear system identification and key water quality parameters dynamic modeling in wastewater treatment process are conducted, and the simulation results prove the feasibility and effectiveness of the structure design method.

Key words: neural network; structure design; algorithm; singular value decomposition; dynamic modeling

2017-06-14 收到初稿, 2017-08-14 收到修改稿。

联系人及第一作者: 乔俊飞 (1968—), 男, 教授。

基金项目: 国家自然科学基金项目 (61533002, 61603012); 北京市教育委员会科研计划项目 (KM201710005025)。

Received date: 2017-06-14.

Corresponding author: Prof. QIAO Junfei, mashijie@emails.bjut.edu.cn

Foundation item: supported by the National Natural Science Foundation of China(61533002, 61603012) and the Beijing Municipal Education Commission Foundation(KM201710005025).

引 言

径向基函数(radial basis function, RBF)神经网络由于结构简单、全局逼近能力强,在信号处理、模式识别、非线性系统建模和控制中得到了广泛的应用^[1-4]。但是 RBF 神经网络对非线性系统建模的能力有限,并且无法自适应较复杂的动态环境^[5-6]。另外,由于递归 RBF 神经网络是具有反馈连接的,具有较好的动态特性,能够自适应较复杂的动态环境,所以递归 RBF 神经网络在实际应用和理论研究中得到了更广泛的应用^[7-9]。

递归 RBF 神经网络应用的核心问题是其结构设计及学习算法设计。学习算法的研究^[10-14]已经日趋成熟,而结构设计仍是一个普遍关注而尚未很好解决的问题。为了实现递归 RBF 神经网络的结构自组织设计,Karayiahnnis 等^[15]提出了一种结合有监督、无监督学习和网络增长的方法来增加隐含层神经元的数量。但是他们忽略了结构调整后参数的调整,导致神经网络学习算法的收敛速度较慢。因此,Subrahmanya 等^[16]提出了一种基于粒子群算法和协方差矩阵的自适应(particle swarm optimization and covariance matrix adaptation, PSOCMA-RNN)策略来同时调整递归 RBF 神经网络隐含层神经元的结构和参数。Hsu^[17]提出一种基于尔曼的自组织结构设计(Elman-based self-organizing RBF neural network, ESRNN)方法来同时调整隐含层神经元的结构和参数。但是 PSOCMA-RNN 和 ESRNN 的网络尺寸随着冗余的隐含层神经元未被移除而逐渐增加,所以在实际的一些应用中会由于网络尺寸过大而影响到网络的预测精度。为了删除这些冗余的隐含层神经元,Zheng 等^[18]介绍了一种基于动作电位和稳态机制的策略来设计递归神经网络隐含层神经元的结构;Leung 等^[19]提出一种局部卡尔曼滤波的方法来删除递归神经网络的隐含层神经元。尽管这些递归神经网络通过增加和删减机制取得了一些效果,比如,较少冗余的隐含层神经元和权值连接,但是由于单独的进行增长和删减使大部分的时间消耗在网络的训练上。因此,递归 RBF 神经网络结构优化设计方法仍是一个开放的问题^[20]。

鉴于以上存在的问题,本文提出了一种基于 ROLS 算法^[21]的递归 RBF 神经网络(recurrent RBF

neural network based on recursive orthogonal least squares, ROLS-RRBF)。ROLS-RRBF 利用 ROLS 算法来计算隐含层神经元的独立贡献度和损失函数,以此判断增加或归为不活跃组的神经元。然后通过 SVD 的方法决定最佳的隐含层神经元个数来删除不活跃组的一些神经元以实现减少隐含层神经元的冗余。同时利用梯度下降算法更新递归 RBF 神经网络的参数来保证神经网络的精度。最后,利用其对典型非线性函数逼近和污水处理过程中关键水质参数建模。

1 递归 RBF 神经网络

递归 RBF 神经网络主要由输入层、隐含层、输出层组成。此外,递归 RBF 神经网络的输出层通过时延 Z^{-1} 反馈到隐含层,将神经网络迁移时刻的输出值反馈到隐含层,作为新的输入加入到隐含层。递归 RBF 神经网络的结构如图 1 所示。递归 RBF 神经网络的结构介绍如下。

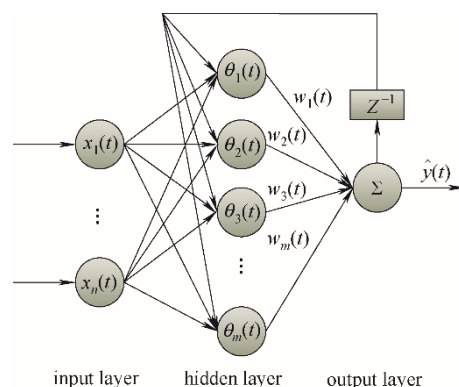


图 1 递归 RBF 神经网络结构

Fig.1 Structure of RRBF neural network

输入层: 输入层有 n 个神经元分别代表 n 个输入变量。输入矢量 $\mathbf{x}(t)$ 的表达式如下

$$\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_i(t), \dots, x_n(t)] \quad (1)$$

式中, $x_i(t)$ 是在时间 t 时刻的第 i 个输入, $i=1, 2, \dots, n$ 。

隐含层: 隐含层有 m 个神经元, 每个神经元分别连接着输入层神经元和输出层神经元。由于高斯函数的连续性和微分特性, 将其作为隐含层的激活函数。第 j 个隐含层神经元的输出值 $v_j(t)$ 如式(2)所示。

$$v_j(t) = \exp\left(-\sum_{i=1}^n (x_i(t) + u_j(t) \times \hat{y}(t-1) - c_{ij})^2 / 2\sigma_j^2\right) \quad (2)$$

式中, $u_j(t)$ 是第 j 个隐含层神经元的反馈权值, c_{ij} 和 σ_j 是第 j 个隐含层神经元的中心和宽度, $\hat{y}(t-1)$

是反馈量, $j=1, 2, \dots, m$ 。输出层的表达式如式(3)所示。

$$\hat{y}(t) = \sum_{j=1}^m v_j(t) w_j(t) \quad (3)$$

式中, $\hat{y}(t)$ 是在时间 t 时刻的输出, $w_j(t)$ 是第 j 个隐含层神经元的输出权值。

输出层: 输出层仅有一个神经元, 输出层接收来自隐含层神经元的信息。此外, 递归 RBF 神经网络的输出又反馈到隐含层。

2 ROLS-RRBF 结构设计

本文提出的 ROLS-RRBF 利用 ROLS 算法来计算隐含层神经元的独立贡献度并且结合 k 个输入样本前后的网络误差变化, 以此判断是否需要增加隐含神经元。同时利用 ROLS 算法来计算隐含层神经元的损失函数, 以此来判断归为不活跃组的神经元。另外, 利用奇异值分解(singular value decomposition, SVD)决定最佳的隐含层神经元个数, 以此来删除不活跃组中相对不活跃的神经元, 有效地解决了递归 RBF 神经网络结构难以自适应问题。这部分详细介绍了 ROLS-RRBF 的结构设计。

2.1 递归正交最小二乘算法

由于本文提出的 ROLS-RRBF 是基于递归正交最小二乘算法(ROLS)^[22]。因此, 递归正交最小二乘算法(ROLS)介绍如下, 随意选取 N 个输入、输出样本, 则有

$$Y(t) = \hat{Y}(t) + E(t) = V(t)W(t) + E(t) \quad (4)$$

式中, $Y(t)$ 是网络在时刻 t 的输出矩阵, $W(t)$ 是隐含层神经元在时刻 t 的输出权值矩阵, $V(t)$ 是隐含层神经元在时刻 t 的输出矩阵, $E(t)$ 是在时刻 t 的误差矩阵。

由式(4)有

$$J(t) = \|E(t)\|_F = \|Y(t) - V(t)W(t)\|_F \quad (5)$$

式中, $\|\cdot\|_F$ 定义为矩阵的 F 范数。

根据 QR 正交变换^[23-24], 式(5)可化为

$$J(t) = \left\| \begin{bmatrix} \hat{Y}_1(t) \\ \hat{Y}_2(t) \end{bmatrix} - \begin{bmatrix} R(t) \\ 0 \end{bmatrix} W(t) \right\|_F \quad (6)$$

$$= \left\| \begin{bmatrix} \hat{Y}_1(t) - R(t)W(t) \\ \hat{Y}_2(t) \end{bmatrix} \right\|_F$$

式中, $\hat{Y}_1(t)$ 是 m 个输出失量(m 接近于 N), $\hat{Y}_2(t)$ 为 $N-m$ 个输出失量, $R(t)$ 为维数为 m 的上三角矩阵。

若删除第 j 个隐含层神经元, 则有 $R'(t)=[r_1(t), \dots, r_{j-1}(t), \dots, r_{j+1}(t), \dots, r_m(t)]$, 使用 QR 正交分解法得到最终的价值函数为

$$J_j(t) = \left\| \begin{bmatrix} \hat{Y}_{1j}(t) - R_j(t)W_j(t) \\ \hat{Y}_{2j}(t) \end{bmatrix} \right\|_F \quad (7)$$

式中, $R_j(t)W_j(t)=\hat{Y}_{1j}(t)$, 余项为

$$\|\hat{Y}_{2j}(t)\|_F^2 = \|\hat{Y}_{2j}^T(t)\|_F^2 + \|\hat{Y}_2(t)\|_F^2 \quad (8)$$

比较式(6)和式(7)可以得出, $\|\hat{Y}_{2j}^T(t)\|_F^2$ 是由移除第 j 个隐含层神经元引起的。

在递归正交最小二乘算法中, $\hat{Y}_1(t)$ 包括了网络的主要信息^[22]。因此, 每个隐含层神经元的独立贡献度定义如下

$$I_j(t) = \frac{\|\hat{Y}_{1j}(t)\hat{Y}_{1j}^T(t)\|_F}{\|\hat{Y}_1(t)\|_F} \quad (9)$$

式中, $\hat{Y}_{1j}(t)$ 是输出 $\hat{Y}_1(t)$ 在第 j 行的输出值, $I_j(t)$ 是在 t 时刻第 j 个隐含层神经元的独立贡献度, $j=1, 2, \dots, m$ 。

2.2 递归 RBF 神经网络的自组织机制

2.2.1 神经元增加机制 结合独立贡献度 $I_j(t)$ 和 k 个输入样本前后的网络误差对比作为神经元增加机制的判定条件。当 k 个输入样本后的网络误差变大时, 说明网络对于输入动态信息处理能力不足, 此时应分裂当前时刻具有最大独立贡献度的隐含层神经元来增加一个新的神经元。神经元增加机制的具体判定条件如式(10)所示。

$$\begin{cases} e(t) - e(t-k) > 0 \\ j = \arg \max (I_j(t)) \end{cases} \quad (10)$$

式中, j 是具有最大独立贡献度的隐含层神经元, $e(t-k)$ 和 $e(t)$ 分别是 k 个输入样本前后的训练误差, k 是间隔的样本个数, 下角标 $j=1, 2, \dots, m$ 。根据当前时刻具有最大独立贡献度的隐含层神经元 j , 对新增的神经元参数设置如下

$$\begin{cases} c_{\text{new}}(t) = \varepsilon c_j(t) + (1-\varepsilon)x(t) \\ \sigma_{\text{new}}(t) = \varepsilon \sigma_j(t) \\ u_{\text{new}}(t) = \varepsilon u_j(t) \\ w_{\text{new}}(t) = (w_j(t)v_j(t) - e(t))(M_{\text{new}}^{-1}y_{\text{new}}(t))^{-1} \end{cases} \quad (11)$$

式中, $c_j(t)$ 、 $\sigma_j(t)$ 、 $u_j(t)$ 、 $w_j(t)$ 分别是第 j 个神经元分裂前的中心、宽度、反馈权值和输出权值, $\varepsilon \in [0, 1]$, $c_{\text{new}}(t)$ 、 $\sigma_{\text{new}}(t)$ 、 $u_{\text{new}}(t)$ 、 $w_{\text{new}}(t)$ 分别是新增神经元在 t 时刻的中心、宽度、反馈权值和输出权值, M_{new}

为新增的神经元个数, 式(11)中 $M_{\text{new}}=1$ 。

2.2.2 神经元分组机制 分组机制^[21]中所用到的最终预测误差(final prediction error, FPE)^[25]和损失函数 $V(t)$ 定义如下

$$\text{FPE} = \frac{1+\beta(n_p/N)}{1-\beta(n_p/N)} V(t) \quad (12)$$

$$V(t) = \|\hat{\mathbf{y}}_{2j}(N)\|_{\text{F}}^2 / (N) \quad (13)$$

式中, n_p 是权值的数量, β 是学习率, $V(t)$ 是损失函数, 由式(8)、式(13)中求得。

通过 FPE 和损失函数找出对隐含层神经元影响最小的一个神经元, 然后把该神经元归为不活跃组。具体步骤如下。

(1) 初始化 FPE。

(2) 用式(8)和式(13)计算新的 $V_j(t)$ 。

(3) 设 $i = \arg \min(V_j(t))$, 计算 FPE_i 。

(4) 若 $\text{FPE}_i < \text{FPE}$, 把中心 i 归为不活跃组, 转向步骤(2)。

(5) 当 $\text{FPE}_i > \text{FPE}$ 时, 停止分组。

2.2.3 神经元删减机制 删减机制中用 SVD^[26]来计算隐含层最佳神经元个数, 然后删除分组机制中归为不活跃组中的相对不活跃的神经元。步骤如下。

(1) 令 $\mathbf{\Omega} \in \mathbf{R}^{N \times m} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_m)$, 在这里, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0 = \sigma_{n+1} = \dots = \sigma_N$ 。

(2) 令 $\min \{(\lambda \sigma_k + (1-\lambda) \varepsilon_k) \leq r\}$, 在这里, $\varepsilon_k = \sigma_k - \sigma_{k+1}$, $\lambda \in [0, 1]$, 文献[27]提供了可用的数据手段来构建合适的阈值 r 。

(3) 由步骤(1)和步骤(2)求出最佳隐含层神经元个数 k 。

根据神经元分组机制删减不活跃组中对隐含层神经元影响最小的一个神经元 j , 删除后的神经元及参数变化为

$$\begin{cases} \mathbf{c}'_{j'}(t) = \mathbf{c}_j(t) \\ \sigma'_{j'}(t) = \sigma_j(t) \\ u'_{j'}(t) = u_j(t) \\ w'_{j'}(t) = w_j(t) + w_j(t) v_j(t) (v_j(t))^{-1} \end{cases} \quad (14)$$

式中, $\mathbf{c}'_{j'}(t)$ 、 $\mathbf{c}_j(t)$ 分别是第 j' 个神经元在第 j 个神经元删除后和删除前的中心; $\sigma'_{j'}(t)$ 、 $\sigma_j(t)$ 分别是第 j' 个神经元在第 j 个神经元删除后和删除前的宽度; $u'_{j'}(t)$ 、 $u_j(t)$ 、 $w'_{j'}(t)$ 、 $w_j(t)$ 分别是第 j' 个神经元在第 j 个神经元删除后和删除前的反馈权值; $w_j(t)$ 是删除前第 j 个神经元的宽度; $v_j(t)$ 是删除前第 j 个隐含层神经元输出。

基于以上分析, ROLS-RRBF 神经网络学习算法的具体步骤如下。

(1) 给定一个递归 RBF 神经网络, 初始化隐含层神经元个数为较小自然数, 利用梯度下降算法调整神经网络的中心 $\mathbf{c}$ 、宽度 σ 、反馈权值 u 和权值 w , 初始化 $\mathbf{R}(0) = \alpha \mathbf{I}$, $\alpha > 0$, $\hat{\mathbf{y}}_1(0) = \|\hat{\mathbf{y}}_2(0)\|_{\text{F}}^2 = 0$ 。

(2) 在迭代 t 时, 根据式(9)计算神经元独立贡献度 $I_j = \|\hat{\mathbf{y}}_{1j}(t) \hat{\mathbf{p}}_{1j}^T(t)\|_{\text{F}} / \|\hat{\mathbf{y}}_1(t)\|_{\text{F}}$, $j=1, 2, \dots, m$, 令具有最大贡献度 I 的中心为 $\mathbf{c}_j(t)$ 。当满足式(10)时增加新的神经元, 根据式(11)设定新增神经元的初始参数。

(3) 根据式(12)、式(13)计算最终预测误差 FPE 和损失函数 $V(t)$, 然后结合分组机制步骤来判定归为不活跃组的神经元。

(4) 根据神经元删减机制中的奇异值分解 SVD 找出最佳的隐含层神经元个数 k , 进而删除不活跃组中的相对不活跃的神经元, 根据式(14)设定删减神经元后的参数。

(5) 达到期望误差或最大迭代步数时停止计算。

3 收敛性分析

ROLS-RRBF 神经网络的收敛性影响最终网络的性能。因此, 为了保证 ROLS-RRBF 神经网络的实际应用, 以下给出神经网络的收敛性证明。ROLS-RRBF 神经网络的收敛性主要分为 3 个部分: 隐节点增加、删减和固定阶段。

假设 ROLS-RRBF 神经网络在 t 时刻有 m 个隐含层神经元, 误差为 $e_m(t)$ 。

(1) 隐节点增加阶段。若神经网络隐含层神经元 j 需要分裂, 分裂后的新神经元数为 M_{new} 个, 则分裂神经元后递归 RBF 神经网络隐含层神经元个数为变为 $m+M_{\text{new}}-1$ 个, 输出误差将变为

$$\begin{aligned} e_{m+M_{\text{new}}-1}(t) &= \sum_{k=1}^{m+M_{\text{new}}-1} w_k(t) v_k(t) - y_d(t) = \\ &= \sum_{k=1}^m w_k(t) v_k(t) + \sum_{i=1}^{M_{\text{new}}} w_i(t) v_i(t) - y_d(t) \end{aligned} \quad (15)$$

式中, $y_d(t)$ 是 t 时刻的期望输出, 由式(11)给定的参数调整规则可得

$$\begin{aligned} \sum_{i=1}^{M_{\text{new}}} w_i(t) v_i(t) - w_j(t) v_j(t) &= \\ \sum_{i=1}^{M_{\text{new}}} (w_j(t) v_j(t) - e_m(t)) (M_{\text{new}} v_i(t))^{-1} v_i(t) - & \\ w_j(t) v_j(t) &= -e_m(t) \end{aligned} \quad (16)$$

则神经元分裂后的神经网络的误差即为

$$\begin{aligned}
e_{m+M_{\text{new}}-1}(t) &= \sum_{k=1}^m w_k(t)v_k(t) - y_d(t) + \\
&\sum_{i=1}^{M_{\text{new}}} w_i(t)v_i(t) - w_j(t)v_j(t) = \\
&\sum_{k=1}^m w_k(t)v_k(t) - y_d(t) - e_m(t) = \\
&e_m(t) - e_m(t) = 0
\end{aligned} \quad (17)$$

由以上分析可以看出, 增加神经元后, t 时刻的网络输出误差为 0, 一定程度上加快了网络的收敛性, 从而提高网络的性能。

(2) 隐节点删减阶段。在隐节点删减阶段, 若满足删减条件时删除第 j 个隐节点。此时神经元的数量变为 $m-1$, 而神经网络的输出误差变为

$$e_{m-1}(t) = \sum_{k=1}^m w_k(t)v_k(t) - w_j(t)v_j(t) - y_d(t) \quad (18)$$

根据式(14)给定的参数调整规则, 输出误差可变为

$$\begin{aligned}
e_{m-1}(t) &= \sum_{k=1, k \neq j'}^m w_k(t)v_k(t) - w_j(t)v_j(t) - y_d(t) + \\
&\left(w_{j'}(t) + w_j(t)v_j(t)(v_{j'}(t))^{-1} \right) v_{j'}(t) = \\
&\sum_{k=1, k \neq j'}^m w_k(t)v_k(t) - w_j(t)v_j(t) - y_d(t) + \\
&w_{j'}(t)v_{j'}(t) + w_j(t)v_j(t) = \\
&\sum_{k=1, k \neq j'}^m w_k(t)v_k(t) + w_{j'}(t)v_{j'}(t) - y_d(t) = \\
&\sum_{k=1}^m w_k(t)v_k(t) - y_d(t) = e_m(t)
\end{aligned} \quad (19)$$

以上分析表明, 第 j 个神经元删除前后神经网络的输出误差相等, 因此可以得出, 结构删减对神经网络的误差没有产生影响, 进而表明神经元删减过程对神经网络的收敛性并没有产生影响, 反而提高了神经网络结构的紧凑性。

(3) 隐节点固定阶段。在递归 RBF 神经网络结构不变时, 利用梯度下降算法可以保证最终神经网络的收敛。ROLS-RRBF 神经网络在结构固定阶段属于文献[7]所提到的网络结构, 并且学习算法也与其相同, 因此, ROLS-RRBF 能够保证最终神经网络的收敛性。

综上所述, 通过对隐节点增加、删减和固定阶段进行收敛性证明, 可以保证 ROLS-RRBF 的收敛性。因此证明了该结构设计方法的有效性。

4 仿真实验

本文提出的 ROLS-RRBF 能够根据研究对象的复杂度在线调整神经网络隐含层神经元个数, 减少

神经网络的冗余并且提高神经网络的性能。利用 ROLS-RRBF 对 Mackey-Glass 时间序列预测和非线性系统辨识, 以及污水处理过程中关键水质参数进行建模, 证明该结构设计方法的有效性。

进一步, 为了表达预测精度, 均方根误差被定义为

$$\text{RMSE}(p) = \sqrt{\frac{1}{2t} \sum_{p=1}^t (y_d(p) - \hat{y}(p))^2} \quad (20)$$

式中, $\hat{y}(p)$ 和 $y_d(p)$ 分别是时间 p ($p=1, 2, \dots, t$) 时刻的神经网络输出值和期望输出值。

4.1 Mackey-Glass 时间序列预测

本实验中, 通过对 Mackey-Glass 时间序列预测来验证 ROLS-RRBF 的有效性, 其表达式如下

$$x(t+1) = (1-a)x(t) + \frac{bx(t-\tau)}{1+x^{10}(t-\tau)} \quad (21)$$

式中, $a=0.1$, $b=0.2$, $\tau=17$ 。选取样本 1000 组, 其中 500 组用来训练, 剩余 500 组用来预测。神经网络的初始隐含层神经元是 2 个, 初始网络结构为 4-2-1。递归 RBF 神经网络的学习算法为梯度下降法。

首先, 与固定结构的递归 RBF 神经网络(隐含层神经元为 12 个)作对比来验证 ROLS-RRBF 自组织结构设计的合理性, 其两种方法的测试效果如图 2 所示。从图 2 可以清晰地看出, 与具有固定结构的递归 RBF 神经网络对比, ROLS-RRBF 能更准确地预测 Mackey-Glass 时间序列。图 3 为两种方法的测试误差曲线。从图 3 可以看出, 与具有固定结构的递归 RBF 神经网络对比, ROLS-RRBF 的测试误差仅在一个很小的范围内波动。

为了进一步证明 ROLS-RRBF 的有效性和结构的紧凑性, 与 SORBF^[28]、D-RBF^[20]以及 F-RBF^[29]

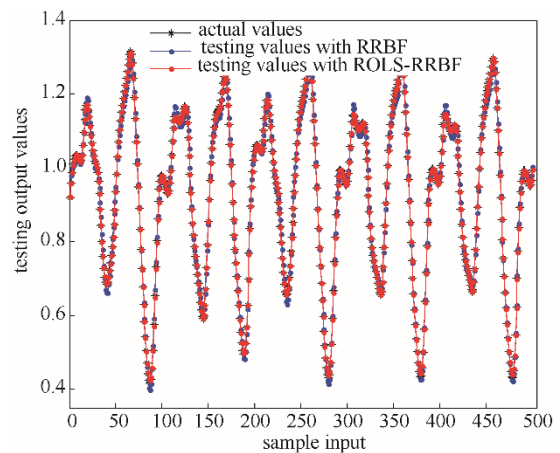


图2 Mackey-Glass 时间序列预测测试结果

Fig.2 Result of testing for Mackey-Glass time series prediction

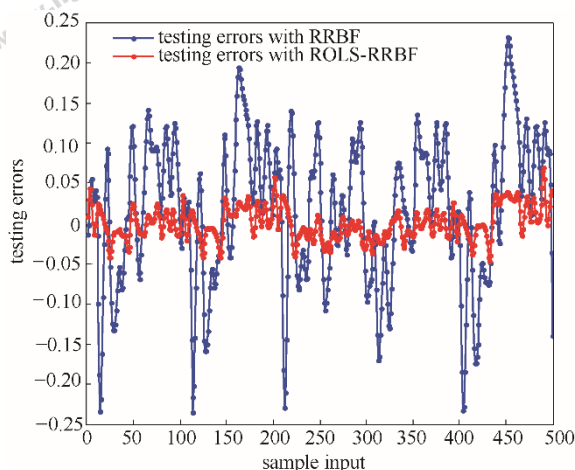


图 3 Mackey-Glass 时间序列预测测试误差

Fig.3 Error of testing for Mackey-Glass time series prediction

进行比较。以上几种结构设计方法的性能参数(独立运行 50 次求平均结果)如表 1 所示。从表 1 可以看出 ROLS-RRBF 以最少的隐含层神经元个数 8 和最少的训练时间(8.92 s)获得最小的训练 RMSE(0.0088)和测试 RMSE(0.0092)。图 4 为 ROLS-RRBF 和 RRBf 隐含层神经元个数变化曲线, 最终 ROLS-RRBF 隐

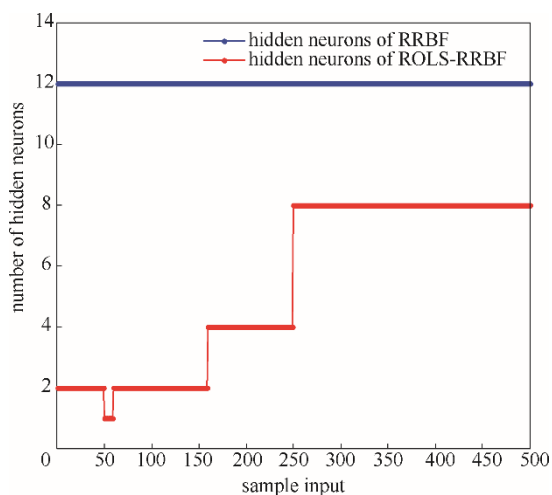


图 4 Mackey-Glass 时间序列预测隐含层神经元个数

Fig.4 Number of hidden neurons for Mackey-Glass time series prediction

表 1 不同方法的性能对比 (Mackey-Glass 时间序列预测)

Table 1 Performance comparison of different method(Mackey-Glass time series prediction)

Different method	Number of hidden neurons	Training RMSE	Testing RMSE	Training time/s
ROLS-RRBF	8	0.0088	0.0092	8.92
RRBF ^[7]	12	0.0204	0.0260	11.26
SORBF ^[28]	12	0.0100	0.0107	9.09
D-RBF ^[20]	10	0.0107	0.0125	9.12
F-RBF ^[29]	10	0.0116	0.0129	10.33

含层神经元个数稳定在 8 个。因此, 与以上几种方法对比, ROLS-RRBF 更适合于 Mackey-Glass 时间序列的非线性系统建模。

4.2 非线性系统辨识

式(22)给出了一个典型的非线性系统, 用该系统来验证 ROLS-RRBF 结构设计方法的有效性。

$$y(t+1) = \frac{y(t)y(t-1)[y(t)+2.5]}{1+y^2(t)+y^2(t-1)} + u(t) \quad (22)$$

式中, $u(t)=\sin(2\pi t/25)$, $t \in [1, 600]$; $y(0)=0$, $y(1)=0$ 。由式(22)可知, 该模型由 3 个输入 $[y(t), y(t-1), u(t)]$ 、一个输出 $(y(t+1))$ 组成。由式(22)在 $t \in [1, 600]$ 上得到 600 组样本数据, 其中前 500 组用于网络训练, 其余 100 组用于测试。隐含层神经元的初始个数为 2。递归 RBF 神经网络的学习算法为梯度下降法。

图 5 给出了固定结构的递归 RBF 神经网络(隐含层神经元为 12 个)和 ROLS-RRBF 自组织结构的对比辨识效果图。从图 5 中的拟合效果对比可以看出, 与具有固定结构的递归 RBF 神经网络对比, ROLS-RRBF 能更准确地辨识非线性系统。图 6 为两种方法的辨识误差曲线。从图 6 中可以清晰地看出, 与具有固定结构的递归 RBF 神经网络对比, ROLS-RRBF 的辨识误差仅在一个很小的范围内波动。从而验证了 ROLS-RRBF 自组织结构设计方法的有效性。图 7 为 ROLS-RRBF 和 RRBf 隐含层神经元个数变化曲线, 最终 ROLS-RRBF 隐含层神经元个数稳定在 7 个。

为了进一步证明 ROLS-RRBF 辨识非线性系统的有效性和结构的紧凑性, 与 SORBF^[28]、D-RBF^[20]以及 F-RBF^[29]进行比较。以上几种结构设计方法的性能参数(独立运行 50 次求平均结果)如表 2 所示。从

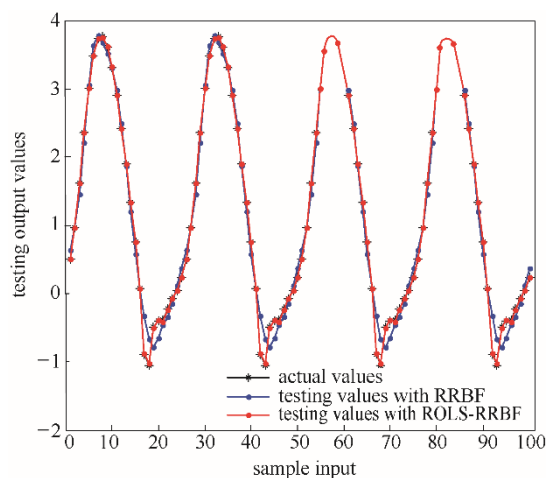


图 5 非线性系统辨识测试结果

Fig.5 Result of testing for nonlinear system identification

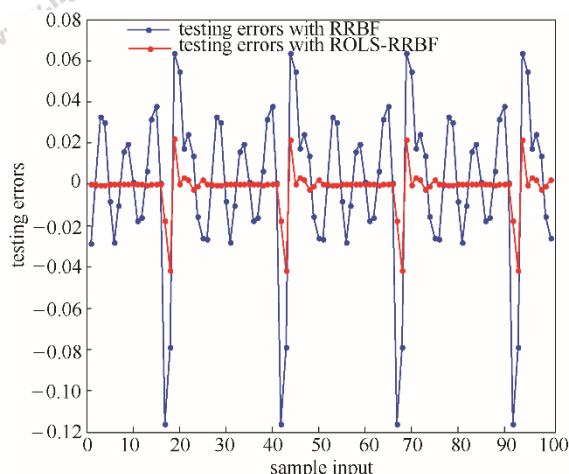


图6 非线性系统辨识测试误差

Fig.6 Error of testing for nonlinear system identification

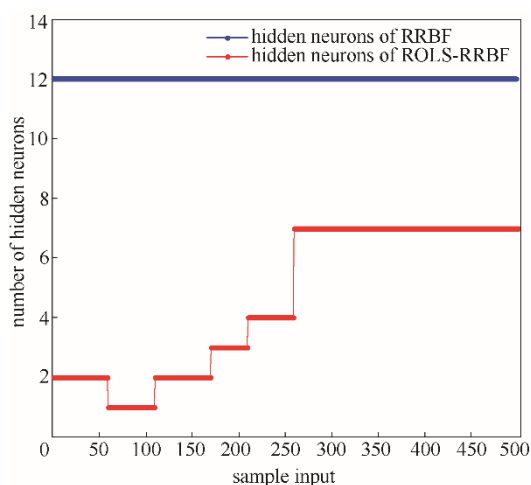


图7 非线性系统辨识隐含层神经元个数

Fig.7 Number of hidden neurons for nonlinear system identification

表2可以看出 ROLS-RRBF 以最少的隐含层神经元个数 7 和最少的训练时间(8.56 s)获得最小的训练 RMSE(0.0091)和测试 RMSE(0.0105)。与以上几种方法对比,进一步说明了 ROLS-RRBF 更适合于非线性系统辨识。

4.3 非线性系统建模

由于在线测量技术的限制,导致很难及时准确

表2 不同方法的性能对比(非线性系统辨识)

Table2 Performance comparison of different method (nonlinear system identification)

Different method	Number of hidden neurons	Training RMSE	Testing RMSE	Training time/s
ROLS-RRBF	7	0.0091	0.0105	8.56
RRBF ^[7]	12	0.0284	0.0320	11.08
SORBF ^[28]	12	0.0135	0.0186	8.98
D-RBF ^[20]	9	0.0118	0.0152	9.15
F-RBF ^[29]	8	0.0098	0.0108	9.24

的预测污水处理过程中出水氨氮的浓度而造成超标排放。因此,为了及时和准确地预测污水处理过程中出水氨氮的浓度,本文利用 ROLS-RRBF 对污水处理过程中出水氨氮进行建模,选取温度 (T)、出水 ORP (oxidation reduction potential, ORP)、出水硝态氮 ($\text{NO}_3\text{-N}$)、出水酸碱度 (pH)、好氧末端 DO (dissolved oxygen, DO)作为 ROLS-RRBF 的输入变量,ROLS-RRBF 的输出变量为出水氨氮。选取北京市某污水处理厂 2015 年全年数据进行仿真。在剔除不正常数据后,还剩余 530 组数据,其中,330 组数据用于训练,200 组数据用于测试。神经网络的初始隐含层神经元数是 2 个,初始连接权值为任意值,初始中心给定为 0 到 1 的任意值,初始函数宽度给定为 $\sigma=1$ 。在此条件下进行验证,对系统建模过程中的神经网络隐含层神经元个数变化如图 9 所示,固定结构的递归 RBF 神经网络(隐含层神经元为 12 个)和 ROLS-RRBF 输出值与系统实际值的比较如图 8 所示,固定结构的递归 RBF 神经网络和 ROLS-RRBF 输出值与系统实际值的误差如图 10 所示。

仿真结果表明:与固定结构的递归 RBF 神经网络相比,ROLS-RRBF 能够很好地建立该非线性系统的模型;图 8 显示实际系统输出值与基于 ROLS-RRBF 模型的输出值基本吻合,证明基于 ROLS-RRBF 方法的污水处理过程关键水质参数出水氨氮模型的有效性;图 10 可以清晰地看出 ROLS-RRBF 的测试误差仅在一个很小的范围内波动,从而说明了 ROLS-RRBF 神经网络具有较好的泛化能力。

为了进一步证明 ROLS-RRBF 建立非线性系统模型的有效性和预测的准确性,除了与固定结构的

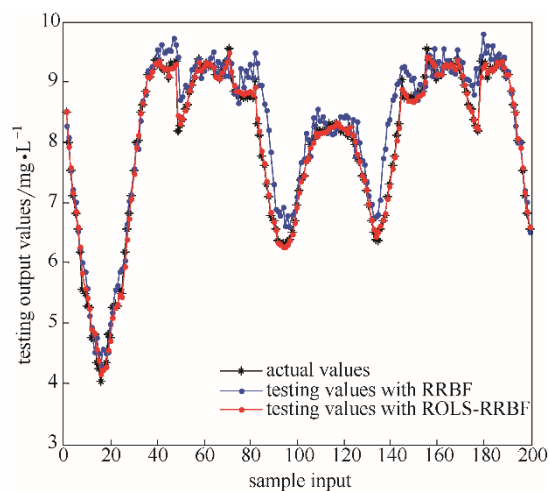


图8 非线性系统建模测试结果

Fig.8 Result of testing for nonlinear system modeling

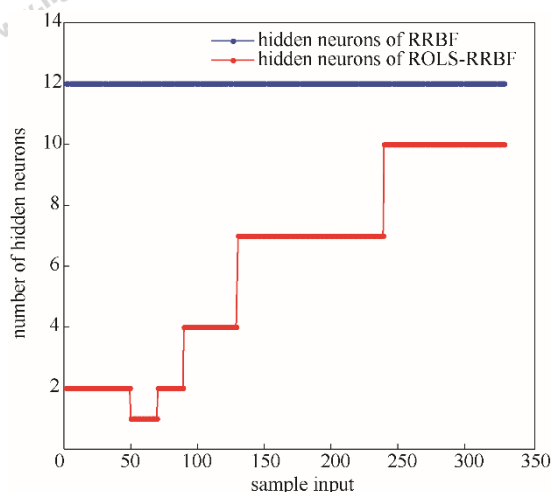


图 9 非线性系统建模隐含层神经元个数

Fig.9 Number of hidden neurons for nonlinear system modeling

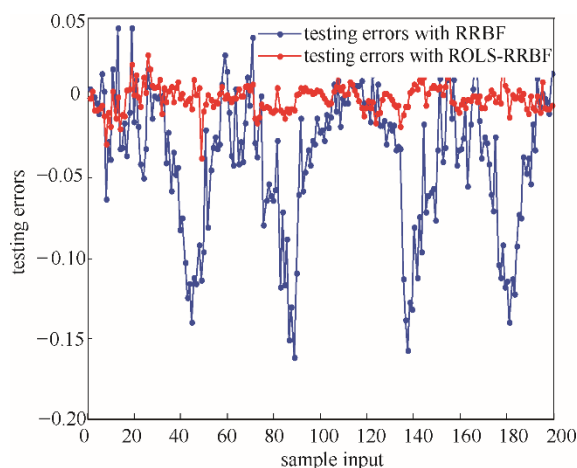


图 10 非线性系统建模测试误差

Fig.10 Error of testing for nonlinear system modeling

递归 RBF 神经网络对比外, 另外与 RBF^[30]以及 BP^[31]进行比较。以上方法的性能参数(独立运行 50 次求平均结果)如表 3 所示。从表 3 可以看出 ROLS-RRBF 以最小的训练 RMSE(0.0215)、测试 RMSE(0.0316)和最少的训练时间(10.35 s)获得最高的预测准确度(97.82%)。因此, 与以上几种方法对比, ROLS-RRBF 更适合于污水处理过程中的非线性系统建模。

表 3 不同方法的性能对比 (非线性系统建模)

Table 3 Performance comparison of different method(nonlinear system modeling)

Different method	Training RMSE	Testing RMSE	Prediction accuracy/%	Training time/s
ROLS-RRBF	0.0215	0.0316	97.82	10.35
RRBF ^[7]	0.0584	0.1020	92.58	12.38
RBF ^[30]	0.1425	0.2857	90.63	15.68
BP ^[31]	0.3512	0.3517	88.42	18.24

5 结 论

针对递归 RBF 神经网络的结构设计问题, 提出基于 ROLS 的递归 RBF 网络结构调整方式, 有效地解决了递归 RBF 神经网络结构动态设计问题。提出的 ROLS-RRBF 自组织方法有以下特点。

(1) 通过与其他结构自组织方法进行比较得出, 文中所提 ROLS-RRBF 神经网络具有较好的预测精度。

(2) 给出了神经网络结构动态调整阶段收敛性分析, 证明 ROLS-RRBF 神经网络具有较好的收敛性, 并且利用梯度下降的参数算法保证了最终的收敛, 保证 ROLS-RRBF 的成功应用。

(3) 通过若干个实验结果表明, 文中所提的 ROLS-RRBF 有效地解决了递归 RBF 神经网络结构设计问题, 实现了 Mackey-Glass 时间序列预测、非线性系统辨识和污水处理过程关键水质参数建模。

References

- [1] YE J, GE L D, WU Y X. An application of improved RBF neural network in modulation recognition[J]. *Acta Automatica Sinica*, 2007, **33**(6): 652-654.
- [2] CHEN S, WANG X X, BROWN D J. Sparse incremental regression modeling using correlation criterion with boosting search[J]. *IEEE Signal Processing Letters*, 2005, **12**(3): 198-201.
- [3] CHEN S, WOLFGANG A, HARRIS C J, *et al.* Symmetric RBF classifier for nonlinear detection in multiple-antenna-aided systems[J]. *IEEE Transactions on Neural Networks*, 2008, **19**(5): 737-745.
- [4] CARA A, POMARES H, ROJAS I. A new methodology for the online adaptation of fuzzy self-structuring controllers[J]. *IEEE Transactions on Fuzzy Systems*, 2011, **19**(3): 449-464.
- [5] FAN H J, SONG Q. A linear recurrent kernel online learning algorithm with sparse updates[J]. *Neural Networks*, 2014, **50**(3): 142-153.
- [6] NIU B, LI L. Adaptive neural network tracking control for a class of switched strict feedback nonlinear systems with input delay[J]. *Neurocomputing*, 2016, **173**(3): 2121-2128.
- [7] HAN H G, LI Y, GUO Y N, *et al.* A soft computing method to predict sludge volume index based on a recurrent self-organizing neural network[J]. *Applied Soft Computing*, 2016, **38**(1): 477-486.
- [8] ZHANG H, WANG Z, LIU D. A comprehensive review of stability analysis of continuous-time recurrent neural networks[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, **25**(7): 1229-1262.
- [9] LEE W, LEE S, PARK P. Improved stability criteria for recurrent neural networks with interval time-varying delays via new Lyapunov functionals[J]. *Neurocomputing*, 2015, **155**(1): 128-134.
- [10] XIE T, YU H, HEWLETT J, *et al.* Fast and efficient second-order method for training radial basis function networks[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2012, **23**(4): 609-619.
- [11] WANG X, YU J, LI C, *et al.* Robust stability of stochastic fuzzy delayed neural networks with impulsive time window[J]. *Neural*

- Networks, 2015, **67**(1): 84-91.
- [12] HAN H G, WU X L, QIAO J F. Nonlinear systems modeling based on self-organizing fuzzy-neural-network with adaptive computation algorithm[J]. IEEE Transactions on Cybernetics, 2014, **44**(4): 554-564.
- [13] MENG D, PEI Z. Dynamic adaptive learning algorithm based on two-fuzzy neural-networks[J]. Neurocomputing, 2014, **125**(1): 88-94.
- [14] CRUZ D, MAIA R, SILVA L, *et al.* A bee-inspired data clustering approach to design RBF neural network classifiers[J]. Neurocomputing, 2016, **172**(3): 427-437.
- [15] KARAYIANNIS N B, MIG W. Growing radial basis neural networks: merging supervised and unsupervised learning with network growth techniques[J]. IEEE Transactions on Neural Network, 1997, **8**(6): 1492-1506.
- [16] SUBRAHMANYA N, SHIN Y C. Constructive training of recurrent neural networks using hybrid optimization[J]. Neurocomputing, 2010, **73**(13): 2624-2631.
- [17] HSU C F. Adaptive backstepping Elman-based neural control for unknown nonlinear systems[J]. Neurocomputing, 2014, **136**(1): 170-179.
- [18] ZHENG P S, DIMITRAKAKIS C, TRIESCH J. Network self-organization explains the statistics and dynamics of synaptic connection strengths in cortex[J]. Computational Biology, 2013, **9**(1): 1-8.
- [19] LEUNG C S, LAM P M. A local training-pruning approach for recurrent neural networks[J]. International Journal of Neural Systems, 2003, **13**(1): 25-38.
- [20] 乔俊飞, 韩红桂. RBF神经网络的结构动态优化设计[J]. 自动化学报, 2010, **36**(6): 865-872.
- QIAO J F, HAN H G. Optimal structure design for RBFNN structure[J]. Acta Automatica Sinica, 2010, **36**(6): 865-872.
- [21] TOK D K, YU D L, MATHEWS C. Adaptive structure radial basis function network model for processes with operating region migration[J]. Neurocomputing, 2015, **155**: 186-193.
- [22] GOMM J B, YU D L. Selecting radial basis function network centers with recursive orthogonal least squares training[J]. IEEE Transactions on Neural networks, 2000, **11**(2): 306-314.
- [23] BROOMHEAD D S, LOWE D. Multivariable functional interpolation and adaptive networks[J]. Complex Systems, 1988, **2**: 321-355.
- [24] KAMINSKI W, STRUMILLO P. Kernel orthonormalization in radial basis function neural networks[J]. IEEE Transactions on Neural Networks, 1997, **8**(5): 1177-1183.
- [25] AKAIKE H. Fitting autoregressive models for prediction[J]. Annals of the Institute of Statistical Mathematics, 1969, **21**: 425-439.
- [26] TEOH E J, TAN K C, XIANG C. Estimating the number of hidden neurons in a feedforward network using the singular value decomposition[J]. IEEE Transactions on Neural Networks, 2006, **17**(6): 1623-1629.
- [27] KONSTANTINIDES K, YAO K. Statistical analysis of effective singular values in matrix rank determination[J]. IEEE Transactions on Acoustics, Speech Signal Process, 1988, **36**(5): 757-763.
- [28] HAN H G, CHEN Q L, QIAO J F. An efficient self-organizing RBF neural network for water quality prediction[J]. Neural Networks, 2011, **24**(7): 717-725.
- [29] 韩红桂, 乔俊飞, 薄迎春. 基于信息强度的 RBF 神经网络结构设计研究[J]. 自动化学报, 2012, **38**(7): 1083-1090.
- HAN H G, QIAO J F, BO Y C. On structure design for RBF neural network based on information strength[J]. Acta Automatica Sinica, 2012, **38**(7): 1083-1090.
- [30] 乔俊飞, 安茹, 韩红桂. 基于 RBF 神经网络的出水氨氮预测研究[C]//杨辉. 第 26 届中国过程控制会议. 南昌: 中国自动化学会, 2015: 120-127.
- QIAO J F, AN R, HAN H G. Based on the RBF neural network water ammonia nitrogen prediction research[C]//YANG H. The 26th Process Control Meeting in China. Nanchang: China Society of Automation, 2015: 120-127.
- [31] 杨琴, 谢淑云. BP 神经网络在洞庭湖氨氮浓度预测中的应用[J]. 水资源与水工程学报, 2006, **17**(1): 65-67.
- YANG Q, XIE S Y. Application of BP neural network into predicting $\text{NH}_3\text{-N}$ concentration of Dongting Lake[J]. Journal of Water Resources & Water Engineering, 2006, **17**(1): 65-70.