文章编号: 1000-0887(2007) 04-0428-05

一类递归小波神经网络的稳定性研究*

邓 韧, 李著信, 樊友洪

(后勤工程学院, 重庆 400016)

(陈正汉推荐)

摘要: 在小波神经网络(WNNs) 和递归神经网络(RNNs) 的基础上, 提出了一类递归小波神经网络(RWNNs) 模型, 它具有两种网络模型的优点. 根据 Liapunov 渐近稳定理论, 对该模型的渐近稳定性进行了研究, 并给出了相关的定理和公式. 仿真结果表明该模型对非线性动态系统有良好的辨识效果.

关 键 词: 递归小波神经网络; 渐近稳定性; 非线性系统; Liapunov 函数
中图分类号: TP183 文献标识码: A

引 言

小波神经网络(wavelet neural networks(WNNs)) 是近年来在小波分析研究获得突破的基础上提出的一种前馈型网络. 其特点是: 1) 小波基元(wavelet basic unit(WBU)) 及整个网络结构的确定, 有可靠的理论依据, 可避免 BP 网络等结构设计上的盲目性; 2) 网络权系数线性分布和学习目标函数的凸性, 使网络训练过程从根本上避免了局部最优等非线性优化问题; 3) 有较强的函数学习能力和推广能力[1]. 因此, 小波神经网络在函数逼近, 非线性动态系统辨识, 时序数据预测等方面得到了广泛应用. 然而利用单纯的小波前馈网络对动态系统进行辨识, 实际上是将动态时间建模问题变为静态空间建模问题, 这必然会出现许多问题: 如需要先验假定系统的 NARMA 模型类, 需要对模型结构进行定阶, 特别是随着系统阶次的增加或阶次未知时, 迅速膨胀的网络结构将使学习收敛速度变缓. 此外, 多的输入节点也将使相应的辨识系统对外部噪声特别敏感. 因此, 能够更直接更生动地反映系统动态特性的网络应该是动态神经网络, 即递归神经网络(recurrent neural networks(RNNs))[2]. RNNs 能够逼近任何动力学系统; 对于某些问题, 一个很小的反馈系统可能等价于一个很大的甚至无限的前馈系统; 而且反馈系统具有递归计算和反应状态信息的能力; 所以 RNNs 特别适合于非线性动力学系统的辨识, 但是 RNNs 学习算法较为复杂, 而且存在稳定性问题[3]. 本文提出一类递归小波神经网络(recurrent wavelet neural networks(RWNNs)), 它集成了小波神经网络和递归神经网络的优点. 对该模型稳定性的问题, 本文根据 Liapunov 渐近稳定理论进行了讨论, 提出了模型渐近稳定性的充分条件, 最后, 仿真实验也表明了模型的可行性与稳定性.

---

* 收稿日期: 2005-01-20; 修订日期: 2006-10-22
作者简介: 邓韧(1976—), 男, 四川泸州人, 博士;
李著信(1940—), 男, 山东人, 教授, 博士生导师(联系人. Tel: + 86-23-68756374; E-mail: dengdror@tom.com).

1　网络描述

1.1　一维情形

一维递归小波神经网络的结构如图1所示．它由一个输入层，一个隐含层和一个输出层构成．输入层和输出层与一般小波神经网络类似．信号或其采样通过输入层进入网络，而输出层是隐含层输出的线性组合．隐含层是由一系列二维小波神经元构成，它的输入来自于输入层和输出层的反馈．一维递归小波神经网络的动态方程描述如下：

$$\begin{cases} \dot{x}_i(t) = -a_i x_i(t) + \varphi_{j,k}(I(t)) \cdot \varphi_{j_o,k_o}(O(t)) = q(x_i, t), \\ O(t) = \sum_{i=1}^{n} w_i \cdot x_i(t), \end{cases} \tag{1}$$

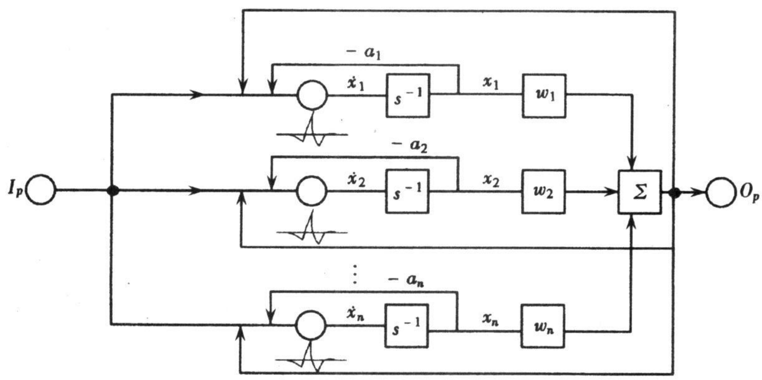


图1　一维递归小波神经网络结构图

$i = 1, 2, \cdots, n;\ j, k \in Z^2, j_o, k_o \in Z^2$．$x_i(t)$ 为一维递归小波神经网络的内部状态，a_i 为一正数向量，$I(t)$ 为网络的输入，$O(t)$ 为网络的输出，w_i 为第 i 个隐含层神经元到输出的权值，$\varphi_{j,k}(I(t)) \cdot \varphi_{j_o,k_o}(O(t)) = \psi_i(I(t), O(t))$ 为二维小波函数，由小波理论知，它由一维小波通过形式为 $\psi(x) = \prod_{i=1}^{2} \varphi(x_i)$ 的直积得到，可以证明，它满足 $L^2(R^n)$ 中的框架条件[4]，即 $A\|f\|^2 \leqslant |\langle \psi_{j,k,j_o,k_o}; f\rangle| \leqslant B\|f\|^2$，其中 $f \in L^2(R)$，A、B 为框架界．从(1)式最终可以得到

$$O_{\text{fin}}(t) = \sum_{i=1}^{n} C_i \cdot \psi_i(I(t), O(t)), \tag{2}$$

其中 C_i 将通过网络学习或在线辨识得到．它的形式正是函数 f 的小波框架表示，所以该模型可以任意精度逼近函数 f．同时可以看出，递归小波神经网络的回归变量，能将网络隐单元状态的信息以紧凑形式保留系统所有以前的输入信息．为书写方便，以下公式省略 t．

假设RWNNs存在一个稳定点 $X^e = (x_1^e, x_1^e, \cdots, x_n^e)$(本文在此不讨论其稳定点的存在性)，将 X^e 代入(1)式，得到

$$-a_i x_i^e + \varphi_{j,k}(I) \cdot \varphi_{j_{oe},k_{oe}}(O_e) = 0,\quad O_e = \sum_{i=1}^{n} w_i \cdot x_i^e, \tag{3}$$

其中 $j_{oe}, k_{oe} \in Z^2$，其余定义如上述．现定义 $z_i = x_i - x_i^e$，(1)式重写为

$$\begin{cases} \dot{z}_i = (x_i - x_i^{\mathrm{e}})' = \dot{x}_i = -a_i(z_i + x_i^{\mathrm{e}}) + \varphi_{j,k}(I) \cdot \varphi_{j_o,k_o}(O), \\ O = \sum_{i=1}^{n} w_i \cdot x_i . \end{cases} \tag{4}$$

定义 $\phi(z_i, x_i^{\mathrm{e}}) = \varphi_{j,k}(I) \cdot \varphi_{j_o,k_o}(O) - \varphi_{j,k}(I) \cdot \varphi_{j_{oe},k_{oe}}(O_{\mathrm{e}})$.

由(1)、(3)、(4)式可以得到

$$\dot{z}_i = -a_i z_i + \phi(z_i, x_i^{\mathrm{e}}),\quad O - O_{\mathrm{e}} = \sum_{i=1}^{n} w_i \cdot z_i . \tag{5}$$

将上述等式以矩阵的形式表达为

$$\dot{\boldsymbol{Z}} = -\boldsymbol{TZ} + \phi,\quad \boldsymbol{O} - \boldsymbol{O}^{\mathrm{e}} = \boldsymbol{WZ}, \tag{6}$$

其中 $\boldsymbol{Z} = [z_1\ \ z_2\ \ \cdots\ \ z_n]^{\mathrm{T}}$, $\boldsymbol{T} = \mathrm{diag}[a_1\ \ a_2\ \ \cdots\ \ a_n]$, $\phi = [\phi(z_1, x_1^{\mathrm{e}})\ \ \phi(z_2, x_2^{\mathrm{e}})\ \ \cdots\ \ \phi(z_n, x_n^{\mathrm{e}})]$, $\boldsymbol{W} = [w_1\ \ w_2\ \ \cdots\ \ w_n]$. 通过以上转换, (6)式中平衡点 $\boldsymbol{Z}^{\mathrm{e}} = \boldsymbol{0}$.

1.2 m 维情形

对于输入 $I \in R^m$, $m > 1$ 的情形, 输入层将有 m 个输入节点, 隐含层的小波神经元节点传递函数为 $\prod_{l=1}^{m} \varphi_{j_l,k_l}(I_l) \cdot \varphi_{j_o,k_o}(O), j_l, k_l \in Z^2$, 其余和一维情形类似.

2 渐近稳定性条件

定理1(全局稳定性) 如果以下两个条件同时满足

1) $j_o = j_{oe}$;

2) $k_o = k_{oe} - 2^{j_o} \cdot (O^{\mathrm{e}} - O)/b$, 其中 $0 < b < 3.5$[4].

则RWNNs稳定点 $X = X^{\mathrm{e}}$ 是全局渐近稳定的.

证明 对于RWNNs(6)式, 定义正的Liapunov函数 $\boldsymbol{V}(\boldsymbol{Z}) = \boldsymbol{Z}^{\mathrm{T}} \cdot \boldsymbol{Z}/2$. 根据(6)式, 得到 $\dot{\boldsymbol{V}}(\boldsymbol{Z}) = \boldsymbol{Z}^{\mathrm{T}}\dot{\boldsymbol{Z}} = -\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{TZ} + \boldsymbol{Z}^{\mathrm{T}}\phi$, 对于一维情形, 上式右端

$$\boldsymbol{Z}^{\mathrm{T}}\phi = \sum_{i=1}^{n} z_i[\varphi_{j,k}(I) \cdot \varphi_{j_o,k_o}(O) - \varphi_{j,k}(I) \cdot \varphi_{j_{oe},k_{oe}}(O_{\mathrm{e}})] =$$

$$\sum_{i=1}^{n} z_i \varphi_{j,k}(I)[\varphi(2^{j_o}O - k_o \cdot b) - \varphi(2^{j_{oe}}O_{\mathrm{e}} - k_{oe} \cdot b)].$$

将条件1)、2)代入, 得到 $\boldsymbol{Z}^{\mathrm{T}}\phi = \boldsymbol{0}$, $\dot{\boldsymbol{V}}(\boldsymbol{Z}) = -\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{TZ} \leqslant \boldsymbol{0}$, 当且仅当 $\boldsymbol{Z} = \boldsymbol{0}$, $\dot{\boldsymbol{V}}(\boldsymbol{Z}) = \boldsymbol{0}$. 根据Liapunov渐近稳定理论, RWNNs(1)式的所有初始状态 $X = X^0 \in (R^n)$ 将全局稳定于平衡点 X^{e}; 对于 m 维情形, 也将得到同样的结论. 证毕.

注1 实际上, 当网络状态离平衡点较远的时候, 系统的能量是非常大的, 从条件1)、2)和等式(2)可以看出输出反馈是很强的, 因此系统逐渐向平衡点靠拢; 当系统接近了平衡点时, 系统的能量降低, 输出反馈几近于0.

注2 条件1)、2)在网络结构设计时是比较容易满足的, 可以先根据输入、输出信号的时频域, 用"分解–综合"方法[1]给出输入、输出信号的尺度和平移的范围, 即确定 j、k、j_{oe}、k_{oe} 所在区间. 由于 O、O_{e} 的频谱和取值范围都是一致的, 所以 $j_o = j_{oe}$ 自然满足, 而 k_o 则取 k_{oe} 区间中满足条件2)的值. 当网络用于时序数据预测时, 条件更容易满足, 因为 j、k 和 j_{oe}、k_{oe} 所在的区间是一致的, 可以只分析输入或输出信号的尺度和平移范围.

定理2(局部稳定性) 如果以下两个不等式同时满足:

1) $|D_i^{\mathrm{e}} w_i| \geqslant \left|\sum_{\substack{j=1 \\ j \neq i}}^{n} D_j^{\mathrm{e}} w_i\right|$;

2) $D_i^e w_i \leqslant 0$;

其中, 一维情形, $D_i^e = \varphi_{j,k}(I) \cdot \varphi'_{j_o,k_o}(O^e)$; m 维情形, $D_i^e = \prod_{l=1}^{m} \varphi_{j_l,k_l}(I_l) \cdot \varphi'_{j_o,k_o}(O^e)$, $i = 1, 2, \cdots, n$.

则对于RWNNs(1)式, 初始状态 $X = X^0$ 在 $S(X^e, \varepsilon) \in R^n$ 邻域的平衡点 $X = X^e$ 是局部渐近稳定的.

证明 当小波函数 $\varphi(x)$ 取为Morlet, Gauss, Mexico hat 等显式可微小波时, (1)式中 $q(x_i, t)$ 是连续可微的, 并在平衡点 $S(X^e, \varepsilon)$ 邻域内有任意阶导数, 则对任意 $x_i (i = 1, 2, \cdots, n)$ 有

$$\dot{x}_i = q(X^e) + \sum_{j=1}^{n} \frac{\partial q_i}{\partial x_j}\Big|_{x_j = x_j^e}(x_j - x_j^e) + Q(n).$$

显然, 当系统状态趋向于平衡点时, 高阶项 $Q(n)$ 将趋向于0, 所以

$$q(X^e) = 0, \quad \dot{x}_i = \sum_{j=1}^{n} \frac{\partial q_i}{\partial x_j}\Big|_{x_j = x_j^e}(x_j - x_j^e),$$

其中 $\sum_{j=1}^{n} \frac{\partial q_i}{\partial x_j}\Big|_{x_j = x_j^e}(x_j - x_j^e)$ 是Jacobi矩阵 $\boldsymbol{J}(X)$ 中元素,

$$\boldsymbol{J}(X) = \begin{bmatrix} -a_1 + D_1^e w_1 & D_1^e w_2 & \cdots & D_1^e w_n \\ D_2^e w_1 & -a_2 + D_2^e w_2 & \cdots & D_2^e w_n \\ \vdots & \vdots & \vdots & \vdots \\ D_n^e w_1 & D_n^e w_2 & \cdots & -a_n + D_n^e w_n \end{bmatrix}.$$

应用条件1)可得

$$a_i + |-D_i^e w_i| \geqslant \left|\sum_{\substack{j=1 \\ j \neq i}}^{n} -D_j^e w_i\right|, \qquad i = 1, 2, \cdots, n.$$

应用条件2)可得

$$|a_i - D_i^e w_i| \geqslant \left|\sum_{\substack{j=1 \\ j \neq i}}^{n} -D_j^e w_i\right|, \qquad a_i - D_i^e w_i > 0.$$

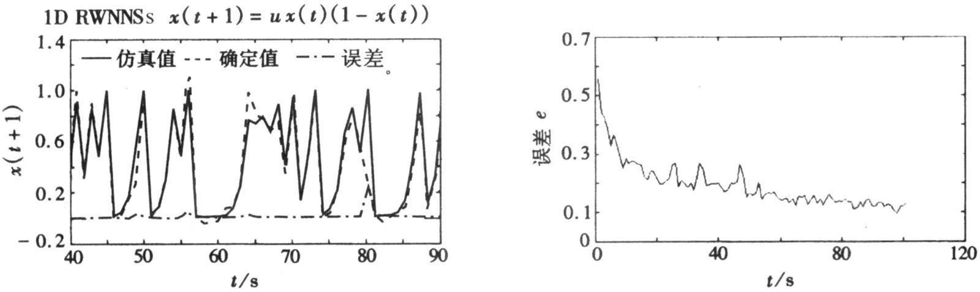
因此, $-\boldsymbol{J}(X^e)$ 是严格的对角占优矩阵, 且对角元均大于0, 所以 $-\boldsymbol{J}(X^e)$ 正定, 因而 $\boldsymbol{J}(X^e)$ 负定, 因此, RWNNs(1) 式在平衡点 $S(X^e, \varepsilon) \in R^n$ 邻域渐近稳定. 证毕.

3 实验仿真

考虑下列系统

$$\dot{x}(t) = \mu x(t)(1 - x(t)). \tag{7}$$

初始条件: $\mu = 4$, $x(0) = 0.02$. 由(7)式描述的系统是一个典型的混沌系统, 对其的辨识与模拟, 将很好地表明递归小波神经网络的性能和稳定性. 设计一维递归小波神经网络, 输入为 $x(t)$, 输出为 $x(t+1)$, 小波神经元母波函数为Gauss小波函数 $f(t) = t\mathrm{e}^{-t^2/2}$, 400个样本点, 根据文献[1]中"分解-综合"方法, 确定 j、k、j_{oe}、k_{oe} 和 j_o、k_o 所在区间, j、j_o、$j_{oe} \in \{5, 6, 7\}$, 取 $j_o = j_{oe}$; k, $k_{oe} \in [1, 128]$, 根据定理1中条件2), 取 $b = 1$, 则 $k_{oe} \in [1, 67]$. 对系统(7)辨识结果如图2(a)、(b)所示, 其均方差为0.135 3. 仿真表明递归小波神经网络在前文所述的充分条件下是稳定的, 对非线性动态系统辨识有较好的结果.



(a) 辨识结果 (b) 误差曲线

图 2 一维递归小波神经网络作动态系统辨识结果

4 结 论

本文提出了一类递归小波神经网络，它继承了小波神经网络和递归神经网络的优点，能够对非线性动态系统进行较好的辨识．通过论证分析和实验仿真表明在上文所述的充分条件下，系统是渐近稳定的．但是同时应该指出，文中所述的仅仅是充分条件，它对网络结构设计有指导意义的同时也加强了约束．递归小波神经网络是一个复杂的非线性动力系统，其稳定性、收敛性都是一个基本的和困难的问题，值得进一步地研究．

[参 考 文 献]

[1] 李银国，张邦礼，曹长修．小波神经网络及其结构设计方法[J]．模式识别与人工智能，1997，**10**(3)：197-205.

[2] 谢庆国，沈轶，万淑芸．Elman 人工神经网络的收敛性分析[J]．计算机工程与应用，2002，**38**(6)：65-81.

[3] 林毅．一类新的离散时间递归 RBF 神经网络[J]．计算技术与自动化，1999，**18**(3)：18-21.

[4] Daubechies I. Orthonormal basis of compactly supported wavelets[J]. Comm Pure Appl Math, 1988, **41**(7): 909-996.

[5] LIU Mei-qin, CHEN Ji-da, LIAO Xiao-xin. Discussion of stability in a class of models on recurrent radial basis function neural networks[J]. Control Theory & Applications, 2000, **17**(6): 919—928.

Discussion of Stability in a Class of Models on Recurrent Wavelet Neural Networks

DENG Ren, LI Zhu-xin, FAN You-hong
(Logistic Engineering University, Chongqing 400016, R. P. China)

Abstract: Based on wavelet neural networks (WNNs) and recurrent neural networks (RNNs), a class of models on recurrent wavelet neural networks (RWNNs) was proposed. The new networks possess the advantages of WNNs and RNNs. Asymptotic stability of RWNNs was researched according to the Liapunov theorem. Some theorems and formulae were given. The simulation results show the excellent performance of the networks in nonlinear dynamic system recognition.

Key words: recurrent wavelet neural networks; asymptotic stability; nonlinear dynamic system; Liapunov function