

动态 T-S 递归神经网络及其应用

彭晓波¹, 桂卫华²

(1. 湖南工业大学 电气与信息工程学院, 湖南 株洲 412007;

2. 中南大学 信息科学与工程学院, 湖南 长沙 410083)

摘要: 基于递归神经网络和模糊系统, 给出了一种动态 T-S 递归模糊神经网络 (DTRFNN)。该神经网络用 BP 算法进行网络权值的学习, 并在权值学习的基础上采用改进的 BP 算法克服局部极小。以动态系统的辨识为例进行仿真实验研究, 并与一般的模糊神经网络进行了比较。结果表明, DTRFNN 的辨识误差较小, 取得了很好的辨识效果。该神经网络应用于某金属温度软测量时, 能很好地实现温度的在线检测。

关键词: 动态 T-S 递归模糊神经网络; BP 学习算法; 软测量

中图分类号: TP273+.4

文献标志码: A

文章编号: 1673-9833(2011)03-0047-04

Dynamic T-S Recurrent Networks and Their Application

Peng Xiaobo¹, Gui Weihua²

(1. School of Electric and Information Engineering, Hunan University of Technology, Zhuzhou Hunan 412007, China;

2. School of Information Science and Engineering, Central South University, Changsha 410083, China)

Abstract: Based on the recurrent neural network and fuzzy system, dynamic T-S recurrent fuzzy neural network (DTRFNN) is proposed. The DTRFNN adopts BP algorithm for net weight learning and uses improved BP algorithm to overcome the local minima. With dynamic system identification as an example, makes a simulation research and compares it with general fuzzy neural network. The result shows that identification error of DTRFNN is smaller than general fuzzy neural network and achieves better identification effect. When the DTRFNN applying to Soft metal temperature measurement, it well realizes the on-line detection.

Keywords: dynamic T-S recurrent fuzzy neural networks (DTRFNN); BP learning algorithm; soft measurement

0 引言

神经网络在系统辨识、模式识别、智能控制等领域有着广泛的应用前景。模糊系统适合表达模糊或定性的知识, 但单纯的模糊系统缺乏自学习和自适应能力, 而神经网络不适宜表达基于规则的知识。T-S 模糊神经网络^[1-2]系统集合了模糊系统与人工神经网络理论的优点, 可用较少的模糊规则表达高度

非线性的复杂系统, 且具有较强的自学习和自适应能力。此外, 系统的参数调整较方便。结合递归神经网络和 T-S 模糊模型, 本文采用 T-S 动态递归型模糊神经网络 (dynamic T-S recurrent fuzzy neural networks, DTRFNN)^[3-8]。利用递归网络实现模糊推理, 能很好地反映动态映射关系, 并具定性知识表达能力, 使网络的内部知识具有明确的物理意义, 且可较易地确定网络的结构和神经元的参数。

收稿日期: 2011-03-24

基金项目: 国家自然科学基金资助项目 (61074067)

作者简介: 彭晓波 (1974-), 男, 湖南株洲人, 湖南工业大学讲师, 博士, 主要从事工业过程优化方面的研究,

E-mail: pengxiaobo88@yahoo.com.cn

本文给出一种动态递归神经网络，其采用BP算法进行网络权值的学习，并通过选择合适的学习率，保证了DTRFNN能克服局部极小。并将其应用于动态系统的辨识和某重金属温度的软测量中，以验证所给出的神经网络的有效性。

1 DTRFNN 的结构

本文给出的DTRFNN结构如图1所示。

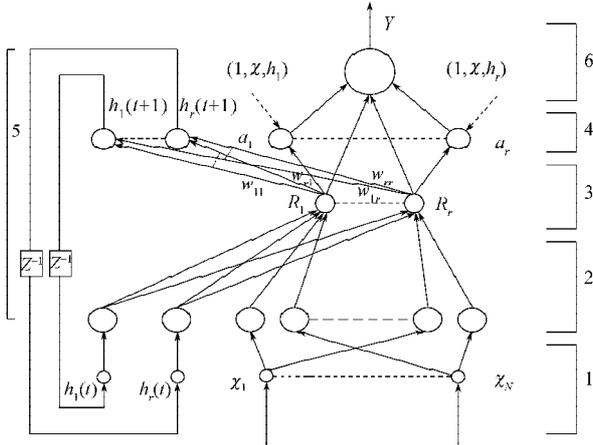


图1 DTRFNN 的结构

Fig. 1 The structure of DTRFNN

从图1可看出，所设计的DTRFNN被分为6层，其输入形式如下：

Rule r : IF $\chi_1(t)$ is A_{r1} and $\chi_2(t)$ is A_{r2} and $h_r(t)$ is G ,
Then $y(t+1) = \alpha_{r0} + \alpha_{r1}\chi_1(t) + \dots + \alpha_{rn}\chi_n(t) + \alpha_{rN+1}h_1(t)$
and $h_1(t+1)$ is w_{1r} and \dots $h_r(t+1)$ is w_{rr} ;
其中： A 和 G 是模糊集， w 是 y 的后置参数， α 是 h 的后置参数。

本文在每一层中定义输入描述为 $u_i^{(k)}$ ，输出为 $O_i^{(k)}$ 。

DTRFNN系统结构描述为：

第一层 $O_i^{(1)} = u_i^{(1)}$;

第二层 $O_i^{(2)} = \exp\left\{-\frac{(u_j^{(2)} - m_{ij})^2}{\sigma_{ij}^2}\right\}$,

其中 $u_j^{(2)} = O_j^{(1)}$;

第三层 $O_i^{(3)} = \prod_{i=1}^r O_{(i)}^{(5)} \prod_{j=1}^{n+1} O_j^{(2)}$;

第四层 $O_i^{(4)} = \sum_{j=0}^{n+1} a_{ij} u_j^{(4)} = a_{i0} + \sum_{j=1}^n a_{ij} \chi_j + a_{in+1} h_i$,

其中 $h_i = \sum_{j=1}^r O_j^{(3)} w_{ij}$;

第五层 $O_i^{(5)} = \frac{1}{1 + e^{-h_i}}$;

第六层 $y = O_i^6 = \frac{\sum_{j=1}^r O_j^{(3)} O_j^{(4)}}{\sum_{j=1}^r O_j^{(3)}}$ 。

2 DTRFNN 的 BP 学习算法

网络结构和规则确定后，需对参数 m_{ij}, α_{ij} 进行学习，常采用BP方法调整参数。设误差代价函数为：

$$\begin{cases} E(t+1) = \frac{1}{2} (y(t+1) - y^d(t+1))^2, \\ e_i(k) = y(t+1) - y^d(t+1). \end{cases}$$

式中： $y^d(t+1)$ 为期望输出； $y(t+1)$ 为实际输出。

参数 a_{ij} 的学习公式为： $a_{ij}(t+1) = a_{ij}(t) - \eta \frac{\partial E(t+1)}{\partial a_{ij}(t)}$,

($i=1,2,\dots,r(t), j=0,1,\dots,n+1$),

式中： $\frac{\partial E(t+1)}{\partial a_{ij}(t)} = (y(t+1) - y^d(t+1)) \frac{O_i^{(3)} u_j^{(4)}}{\sum_{k=1}^r O_k^{(3)}}$ 。

参数 m_{ij} 根据式(1)学习：

$$m_{ij}(t+1) = m_{ij}(t) - \eta \frac{\partial E(t+1)}{\partial m_{ij}(t)}, \tag{1}$$

式中： $\frac{\partial E(t+1)}{\partial m_{ij}(t)} = (y(t+1) - y^d(t+1)) \times$

$$\left(\frac{O_i^{(4)} - y(t+1)}{\sum_{k=1}^r O_k^{(3)}(t)} \times \frac{\partial O_i^{(3)}(t)}{\partial m_{ij}} + \frac{\sum_{i=1}^r O_i^{(3)} a_{in+1} \frac{\partial h_i(t)}{\partial m_{ij}}}{\sum_{i=1}^r O_i^{(3)}} \right),$$

其中，

$$\frac{\partial h_i(t)}{\partial m_{ij}} = \sum_{i=1}^r \left(w_{ii} \times \prod_{i=1, i \neq i}^r O_i^{(5)}(t-1) \times \frac{\partial O_i^{(5)}(t-1)}{\partial m_{ij}} F^i(t-1) \right) +$$

$$\sum_{i=1}^r \left(w_{ii} \times \left[F^i(t-1) \times \prod_{i=1}^r O_i^{(5)}(t-1) \times 2 \frac{\chi_i(t-1) - m_{ij}}{\alpha_{ij}^2} \right] \right),$$

$$\frac{\partial O_i^{(3)}(t)}{\partial m_{ij}} = \prod_{i=1, i \neq p}^r O_i^{(5)} \frac{\partial O_i^{(5)}}{\partial m_{ij}} \times F^p + \prod_i^r O_i^{(5)} \times F^i \times 2 \times$$

$$\frac{\chi_q(t) - m_{ij}}{\alpha_{ij}^2} + \prod_i^r O_i^{(5)} \times F^i \times 2 \times \frac{\chi_q(t) - m_{ij}}{\alpha_{ij}^2},$$

$$\frac{\partial O_i^{(5)}(t)}{\partial m_{ij}} = O_i^{(5)} (1 - O_i^{(5)}(t)) \times \frac{\partial h_i(t)}{\partial m_{ij}},$$

$$F^i(\chi) = \prod_{k=1}^n O_k^{(2)} = \exp\left\{-\left(\sum_{j=1}^n \frac{(\chi_j(t) - m_{ij})^2}{\alpha_{ij}^2}\right)\right\}.$$

3 DTRFNN 局部极小的解决方案

BP 算法中, 学习率是由初始值设定的, BP 算法使用梯度搜索技术, 在整个网络学习过程中, 学习率的选择与网络学习性能密切相关, 学习率代表了权值修正的幅度, 容易陷于局部极小。为了解决这个问题, 笔者引入一种改进的 BP 算法, 无需限制任何前提条件, 通过引入因子即能克服局部极小, 达到全局最小。

step1 如果满足以下 2 式:

$$|y(k+1) - y(k)| < \alpha, \quad 0 < \alpha \ll 1;$$

$$|y(k) - y^d(k)| > \beta, \quad 0 < \beta < 0.2。$$

说明神经网络达局部极小, 必须采用更大的学习率。

step2 用更加大的学习率替换 $\Delta W_i = -\eta \frac{\partial E_i(K)}{\partial W_i}$

中的 η 。判断重新回到旧的局部极小的条件, 是这次输出的 y 和上次输出的 y' 的空间距离小于一个极小的值 r , r 一般取收敛误差的 0.1 倍, 即

$$\|y - y'\| < r。$$

step3 如果没有跳出旧的局部极小, 则将进入

该局部极小的次数 T 加 1, $\Delta W_i = -\xi \frac{\partial E_i(K)}{\partial W_i}$, $\xi > \eta, \xi$ 的选择用下式表示: $\xi = \eta + \varepsilon \lfloor (T-1)/2 \rfloor$, 式中 ε 为学习率增量; 方括号为取整函数。

经过以上 3 步, 可得出要求的学习率, 直到跳出局部极小, 达到最优结果。结合对参数调整的 BP 学习算法和对局部极小的处理, DTRFNN 能达到很好的效果, 下面采用 DTRFNN 进行系统辨识和软测量某金属温度。

4 DTRFNN 的应用实例

4.1 动态系统辨识中的应用

对于非线性系统辨识, 假设系统模型如下:

$$\begin{cases} y_p(k+1) = f(y_p(k), y_p(k-1), y_p(k-2), u(k), u(k-1)), \\ f(x_1, x_2, x_3, x_4) = \frac{x_1 x_2 x_3 (x_3 - 5) + x_4}{1 + x_3^2 + x_2^2}。 \end{cases}$$

含 DTRFNN 的动态系统结构如图 2 所示。

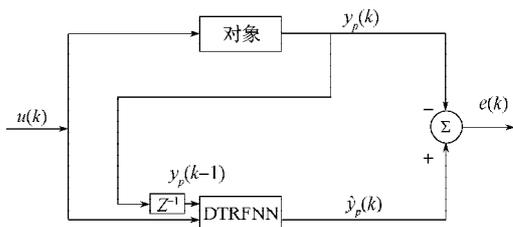


图2 动态系统结构图

Fig. 2 The structure of dynamic system

由图 2 可看出, 含 DTRFNN 的动态系统中, 输入有 $u(k)$ 和 $y_p(k-1)$ 2 个神经元, 隐层有 4 个神经元, 输出仅 $y_p(k)$ 1 个神经元。初始权值为 $[-1, 1]$ 之间的随机数, 学习率的初始值为: $\eta_i^m = \eta_i^\alpha = \eta_i^a = \eta_i^o = 0.011$, $\eta_i^o = 0.055$ 。训练完成后采用如下输入:

$$u(k) = \begin{cases} \sin(\pi k / 25), & k < 250; \\ 1.0, & 250 \leq k < 500; \\ 0.4 \sin(\pi k / 25) + 0.2 \sin(\pi k / 32) + \\ \quad 0.5 \sin(\pi k / 10), & 750 \leq k < 1000。 \end{cases} \quad (2)$$

用式 (2) 来验证 DTRFNN 的有效性。验证后的结果和一般的模糊神经网络 (fuzzy neural networks, FNN) 进行对比, 在相同的输出数为 1, 规则数为 5, 训练样本为 1 000, 训练周期为 200, 初始学习率相同的情况下, 其比较结果见表 1。

表1 DTRFNN 和 FNN 的比较

Table 1 The comparison of DTRFNN and FNN

神经网络	比较项目		
	输入数	节点数	最小均方误差
DTRFNN	2	13	0.000 11
FNN	5	32	0.002 50

从表 1 中可知, 在相同的条件下, DTRFNN 的节点数少, 最后得到的均方误差也小得多, 因此辨识的误差也更加小。可见, DTRFNN 比 FNN 取得的辨识效果好。

4.2 某金属冶炼过程中温度软测量的应用

4.2.1 金属软测量与熔炼技术简介

软测量是指根据比较容易测量的过程辅助变量来估计不可测量过程的主要输出, 以完成实际检测中不能完成的测量。软测量的技术实施主要包括辅助变量的选取、数据处理、软测量模型的建立等几个方面。

某重金属冶炼过程是将经选矿处理后的精矿从反应塔顶部喷入塔内, 并使塔内形成高温, 以使矿石在炉内完成氧化反应, 形成重金属。在金属温度的测量中, 现场的大部分情况依靠人工取样, 然后进行肉眼判断, 因此, 温度的实际值与测量值常有很大的差距。表 2 为某冶炼厂反应塔输入输出变量。

金属温度这一质量指标是对闪速炉的控制目标量, 即 DTRFNN 输出参数; 氧、风、溶剂、精矿量、灰量、精矿成分等 9 个量为操作参数, 即 DTRFNN 的 9 个输入参数。

首先根据现场采集到的大量数据对 DTRFNN 进行学习。即 t 时刻的数据 (可以是化验分析值) 是与 $t-S$ 时刻的输入数据对应的, S 一般为 60 min。经过

大量的数据分析,最后确定 DTRFNN 的结构为 9-53-5-5-1,即输入变量为表 2 中所示 9 个变量,第二层的节点数是 53 个,第三层和第四层的节点数都是 5 个,输出变量为 1 个,即金属温度。

表 2 反应塔输入与输出变量

Table 2 The input and output variables of reactor

类别	变 量	变化范围
控制量 (输入)	氧流量 / ($\text{NM}^3 \cdot \text{h}^{-1}$)	[10 000, 31 000]
	风流量 / ($\text{NM}^3 \cdot \text{h}^{-1}$)	[10 000, 60 000]
干扰量	矿石添加速度 / ($\text{t} \cdot \text{h}^{-1}$)	[80, 176]
	灰添加速度 / ($\text{t} \cdot \text{h}^{-1}$)	[3, 16]
	熔剂添加速度 / ($\text{t} \cdot \text{h}^{-1}$)	[5, 12]
	S/Fe/MgO/SiO ₂	
控制量 (输入)	金属温度 / $^{\circ}\text{C}$	[1 100, 1 400]

4.2.2 熔炼过程中的仿真结果

对给出的 500 组金属温度软测量数据进行训练,并取 58 组数据进行验证。所得结果如图 3 所示,图中 DTRFNN 的结果用星号来表示,实际值用圆圈来表示。

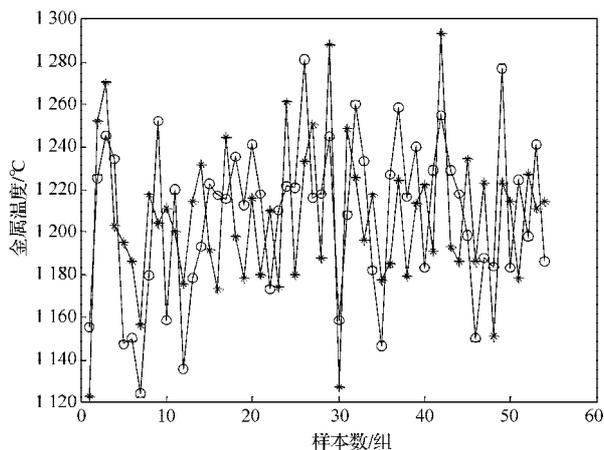


图 3 金属温度的 DTRFNN 软测量

Fig. 3 DTRFNN soft measurement of metal temperature

从图 3 中可看出,软测量温度平均准确率达到到了满意值,说明 DTRFNN 能很好地对系统的参数进行软测量。

5 结语

本文提出了 DTRFNN 的动态模型,并对它的结

构、参数的学习、局部极小问题的解决进行了描述。应用实践结果表明, DTRFNN 动态模型能很好地对动态系统进行辨识,且能很好地对金属温度系统参数进行软测量。

参考文献:

- [1] Takagi T, Sugeno M. Fuzzy Identification of Systems and Its Applications to Modeling and Control[J]. IEEE Trans on SMC, 1985, 15(1): 116-132.
- [2] 干思权,刘贺平,申祝江.一种改进型 T-S 模糊神经网络[J].控制工程,2005,12(5): 442-445.
Gan Siqun, Liu Heping, Shen Zhujang. An Improved T-S Fuzzy Neural Network[J]. Control Engineering of China, 2005, 12(5): 442-445.
- [3] Ku C C, Lee K Y. Diagonal Recurrent Neural Networks for Dynamic Systems Control[J]. IEEE Trans on Neural Networks, 1995, 6(1): 144-156.
- [4] Ma S, Ji C. Fast Training of Recurrent Neural Networks Based on the EM Algorithm[J]. IEEE Trans on Neural Networks, 1998, 9(1): 11-26.
- [5] Sundareshan M K, Condarcuru T A. Recurrent Neural Network Training by a Learning Automaton Approach for Trajectory Learning and Control System Design[J]. IEEE Trans on Neural Networks, 1998, 9(3): 354-368.
- [6] Liang X B, Wang J. A Recurrent Neural Network for Nonlinear Optimization with a Continuously Differentiable Objective Function and Bound Constraints[J]. IEEE Trans on Neural Networks, 2000, 11(6): 1251-1262.
- [7] 从爽,戴谊.递归神经网络的结构研究[J].计算机应用,2004,24(8): 18-20.
Cong Shuang, Dai Yi. Structure of Recurrent Neural Networks[J]. Journal of Computer Applications, 2004, 24(8): 18-20.
- [8] 从爽,高雪鹏.几种递归神经网络及其在系统辨识中的应用[J].系统工程与电子技术,2003,25(2): 195-197.
Cong Shuang, Gao Xuepeng. Recurrent Neural Networks and Their Application in System Identification[J]. Systems Engineering and Electronics, 2003, 25(2): 195-197.

(责任编辑:廖友媛)