

文章编号: 1002-0411(2003)02-181-04

## 递归复合型模糊神经网络结构研究

王京慧 李宏光

(北京化工大学信息科学与技术学院 北京 100029)

**摘要:** 针对一类能够有效引入过程先验知识的复合型模糊神经网络, 研究了其动态结构. 通过对复合型模糊神经网络的函数网络的第二层引入动态递归环节, 使其具有动态映射能力, 实现了对动态系统的良好响应. 本文采用了动态非线性模型对其进行仿真研究, 结果表明, 对于处理动态非线性系统, 此动态复合模糊神经网络较之静态网络在收敛速度、预测精度和网络规模等方面都有较大的改善.\*

**关键词:** 复合型模糊神经网络; 递归; 动态系统

中图分类号: TP13

文献标识码: A

### STUDY ON THE STRUCTURE OF RECURRENT COMPOUND FUZZY NEURAL NETWORK

WANG Jing-hui LI Hong-guang

(Department of Information Science, Beijing University of Chemical Technology, Beijing 100029)

**Abstract:** This paper studies the dynamic structure of a type of compound fuzzy neural network, which can effectively use the process knowledge. The network has the ability of dynamic mapping by adding recurrent nodes in the second layer of the function network of the compound fuzzy neural network, so it would realize the good response to the dynamic system. Simulation has been made with a dynamic nonlinear model, and the result shows that the compound dynamic fuzzy neural network is better than the static network in convergence rate and prediction precision, and has a smaller network size.

**Keywords:** compound fuzzy neural network, recurrent, dynamic system

### 1 引言(Introduction)

模糊逻辑与神经网络相结合是一种新的思想和技术, 它可以从神经网络和模糊系统两个方面获得好处, 具有学习能力, 并可以提供更加有效的智能行为. 模糊神经网络系统目前已有广泛的应用, 但由于大都不具备反馈结构, 而使其应用范围还仅限于处理静态问题. 实际上, 过程对象往往具有动态特性, 过程变量是时间的函数, 因此, 研究具有动态特性的模糊神经网络具有更为实际的意义.

动态神经网络的研究已取得了一定的进展, 其中有: 反馈型神经网络, 如 Hopfield 网络、双向联想存储器 BAM 和 Boltzmann 机; 动态递归神经网络, 包括广义递归神经网络、Elman 网络及局部递归网络等. 它们或是引入反馈, 使网络本身成为一个动态系统, 或是通过延时单元, 把以前的状态存入延时单元中. 这些方法都是设法使网络具有记忆功能, 因而

能够处理与时间有关的对象.

目前, 对于动态模糊神经网络的研究相对较少, 本文针对一类复合型模糊神经网络<sup>[1]</sup>, 提出了递归复合型模糊神经网络(RCFNN)解决方案. 递归复合型模糊神经网络由函数网络和规则网络组成, 并在函数逼近网络的第二层内部引入了递归环节, 这使得它不但能更好地利用对象的先验知识, 而且对动态系统具有很好的响应. 作者还采用了一类动态非线性数学模型对其进行了仿真研究, 结果表明, 此网络结构能有效地提高对复杂动态非线性系统的逼近能力, 而且具有较小的网络结构.

### 2 递归复合型模糊神经网络(Recurrent compound fuzzy neural networks)

多层递归动态复合型模糊神经网络的拓扑结构如图 1 所示, 它由两个模糊神经网络构成: 函数逼近

\* 收稿日期: 2002-04-10

网络和规则推理网络. 函数网络通过一个改进型模糊神经网络进行相应过程操作区域内的非线性函数逼近, 而规则网络则应用了基于规则推理和学习的模糊神经网络结构, 通过对对象先验知识的利用, 实现网络操作区域的划分. 这种复合型结构与普通的模糊神经网络相比, 能更好地从实际意义上引入对象的先验知识, 如操作和控制经验、过程领域知识等, 从而较大地发挥“模糊”的优势.

函数逼近网络由 5 部分组成: 输入层  $L_1$ 、模糊化层  $L_2$ 、规则层  $L_3$ 、解模糊化层  $L_4$  和输出层  $L_5$ . 在第二层  $L_2$  引入了递归神经元, 由于递归神经元有内部反馈连接, 其作为内部记忆的形式自动涉及动态因素, 是一种动态映射, 可以捕获系统的动态响应, 因而网络具有更强的处理动态系统的能力并能简化网络模型. 规则推理网络由 4 部分组成: 输入层  $L_1$ 、模糊化层  $L_2$ 、规则层  $L_3$  和输出层  $L_5$ .

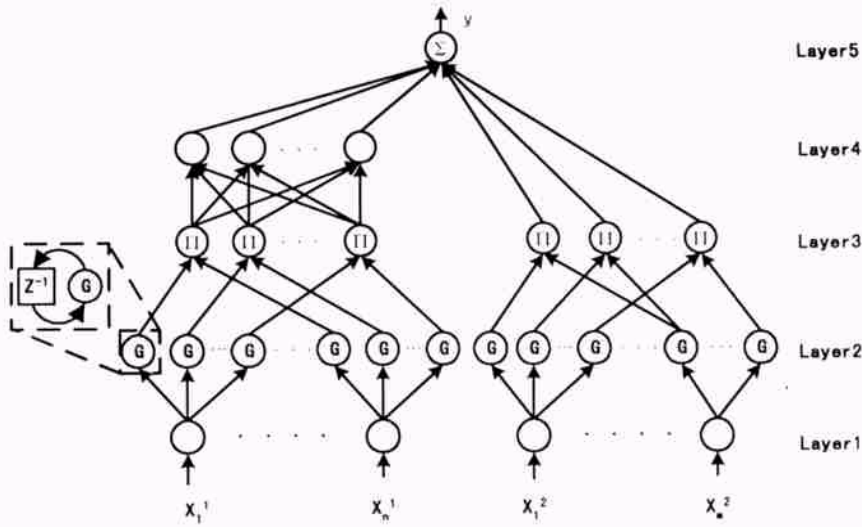


图1 RCFNN 的网络结构图

Fig.1 The structure of the RCFNN

第一层为输入层  $L_1$ : 此层的各个节点直接与输入向量各分量  $x_i$  相连, 并将各输入值  $X = [x_1 \ x_2 \dots \ x_n]^T$  传到下一层.  $n_m$  为输入变量的个数,  $m=1$  为函数网络,  $m=2$  为规则网络.

$$O_i^{(1)m} = u_i^{(1)m} = x_i^{(1)m} \quad (i = 1, 2, \dots, n_m; m = 1, 2) \quad (1)$$

第二层为模糊化加递归层  $L_2$ : 每个节点代表一个语言变量值, 如  $NB, ZE, PS$  等. 它计算各输入分量属于各语言变量值模糊集合的隶属度函数, 这里采用高斯函数.

$$O_{ij}^{(2)m} = \exp\left\{-\frac{(u_{ij}^{(2)m} - c_{ij}^m)^2}{(\sigma_{ij}^m)^2}\right\} \quad (i = 1, 2, \dots, n_m; j = 1, 2, \dots, s_m; m = 1, 2) \quad (2)$$

其中,  $c_{ij}$  和  $\sigma_{ij}$  分别表示隶属函数的中心和宽度,  $s_m$  为  $x_{im}$  的模糊分割数, 对于规则网络, 推理的前提条件的模糊分割数根据先验知识来决定, 从而可以提高网络的收敛速度. 另外, 对于函数网络, 此层的输入为:  $u_{ij}^{(2)}(k) = O_i^{(1)}(k) + O_{ij}^f(k) = O_i^{(1)}(k) + O_{ij}^{(2)}(k-1) \cdot \theta_{ij}$ ,  $\theta_{ij}$  表示递归单元的连接权. 此层中  $O_{ij}^{(2)}(k-1)$  记录了网络前一时刻的信息, 因而可实现动态映射.

第三层为规则层  $L_3$ : 每个节点代表一条模糊规则, 用来匹配模糊规则的前件, 完成模糊与操作, 计算出每条规则的适用度. 函数网络的规则数可依实际情况调整, 而规则推理网络的模糊规则和模糊规则数须根据事先确定的系统推理规则决定. 下面给出函数网络的计算公式, 规则网络的计算由具体的规则连接而定.

$$O_i^{(3)} = \prod_j u_{ji}^{(3)} = \prod_j O_{ji}^{(2)} \quad (i = 1, 2, \dots, s_1; j = 1, 2, \dots, n_1) \quad (3)$$

第四层为解模糊化层  $L_4$ : 实现函数逼近网络的模糊结论的解模糊化运算. 此层节点与第三层中规则网络的节点数相同,  $w_{ij}$  为规则层与解模糊化层之间的连接权.

$$O_j^{(4)1} = \sum_i w_{ji} u_i^{(4)1} = \sum_i w_{ji} O_i^{(3)1} \quad (j = 1, 2, \dots, s_2) \quad (4)$$

第五层为输出层  $L_5$ : 采用重心法计算函数逼近网络和规则推理网络的联合输出.

$$y = O^{(5)} = \frac{\sum_i O_i^{(4)1} O_i^{(3)2}}{\sum_i O_i^{(3)2}} \quad (5)$$

### 3 学习算法(Learning algorithm)

若模糊分割数是确定的,则需要学习的参数为函数网络的权值  $w_{ij}$ 、第二层隶属函数的中心  $c_{ij}$ 和宽度  $\sigma_{ij}$ 以及递归连接权值  $\theta_{ij}$ 。这里采用误差反传的方法来调整参数,误差代价函数为:

$$E = \frac{1}{2}(y_{di} - y_i)^2 \quad (6)$$

$y_{di}$ 和  $y_i$  分别为期望输出和实际输出。

误差反传计算如下:

$$\delta^s = - \frac{\partial E}{\partial O^{(s)}} = y_{di} - y_i \quad (7)$$

$$\delta_j^{41} = - \frac{\partial E}{\partial O_j^{(41)}} = - \frac{\partial E}{\partial O^{(5)}} \frac{\partial O^{(5)}}{\partial O_j^{(41)}} = \frac{\delta^s O_j^{(3)2}}{\sum_i O_i^{(3)2}} \quad (8)$$

因此权值  $w_{ij}$ 有:

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial O_i^{(41)}} \frac{\partial O_i^{(41)}}{\partial w_{ij}} = - \delta_i^{41} u_j^{(4)1} \quad (9)$$

再计算:

$$\delta_i^{31} = - \frac{\partial E}{\partial O_i^{(3)1}} = - \sum_j \frac{\partial E}{\partial O_j^{(41)}} \frac{\partial O_j^{(41)}}{\partial O_i^{(3)1}} \quad (10)$$

$$= \sum_j \delta_j^{41} w_{ij}$$

$$\delta_j^{32} = - \frac{\partial E}{\partial O_j^{(3)2}} = - \frac{\partial E}{\partial O^{(5)}} \frac{\partial O^{(5)}}{\partial O_j^{(3)2}} = \frac{\delta^s}{[\sum_{i=1}^{s_2} O_i^{(3)2}]^2} \cdot [O_j^{(4)1} \sum_{i=1}^{s_2} O_i^{(3)2} - \sum_{k=1}^{s_2} O_k^{(4)1} O_k^{(3)2}] \quad (11)$$

$$\delta_{ij}^{21} = - \frac{\partial E}{\partial O_{ij}^{(2)1}} = - \frac{\partial E}{\partial O_i^{(3)1}} \frac{\partial O_i^{(3)1}}{\partial O_{ij}^{(2)1}} \quad (12)$$

$$= \delta_i^{31} O_j^{(3)1}$$

因此,对于函数网络的隶属函数的中心  $c_{ij}$ 、宽度  $\sigma_{ij}$ 和  $\theta_{ij}$ 有:

$$\frac{\partial E}{\partial c_{ij}} = \frac{\partial E}{\partial O_{ij}^{(2)1}} \frac{\partial O_{ij}^{(2)1}}{\partial c_{ij}} = - \delta_{ij}^{21} \frac{2(x_i + O_{ij}^{(2)1}(k-1)\theta_{ij} - c_{ij})}{\sigma_{ij}^2} \quad (13)$$

$$\frac{\partial E}{\partial \sigma_{ij}} = \frac{\partial E}{\partial O_{ij}^{(2)1}} \frac{\partial O_{ij}^{(2)1}}{\partial \sigma_{ij}} = - \delta_{ij}^{21} \frac{2(x_i + O_{ij}^{(2)1}(k-1)\theta_{ij} - c_{ij})^2}{\sigma_{ij}^3} \quad (14)$$

$$\frac{\partial E}{\partial \theta_{ij}} = \frac{\partial E}{\partial O_{ij}^{(2)1}} \frac{\partial O_{ij}^{(2)1}}{\partial \theta_{ij}} = - \delta_{ij}^{21} \cdot$$

$$\cdot - \frac{2(x_i + O_{ij}^{(2)1}(k-1)\theta_{ij} - c_{ij})O_{ij}^{(2)1}(k-1)}{\sigma_{ij}^2} \quad (15)$$

对于规则网络要根据节点的连接情况决定。最后可得出参数调整的学习算法为:

$$w_{ij}(k+1) = w_{ij}(k) - \beta \frac{\partial E}{\partial w_{ij}},$$

$$i = 1, 2, \dots, r; j = 1, 2, \dots, s_1(s_2) \quad (16)$$

$$c_{ij}(k+1) = c_{ij}(k) - \beta \frac{\partial E}{\partial c_{ij}},$$

$$i = 1, 2, \dots, n; j = 1, 2, \dots, s_1(s_2) \quad (17)$$

$$\sigma_{ij}(k+1) = \sigma_{ij}(k) - \beta \frac{\partial E}{\partial \sigma_{ij}},$$

$$i = 1, 2, \dots, n; j = 1, 2, \dots, s_1(s_2) \quad (18)$$

$$\theta_{ij}(k+1) = \theta_{ij}(k) - \beta \frac{\partial E}{\partial \theta_{ij}},$$

$$i = 1, 2, \dots, n; j = 1, 2, \dots, s_1(s_2) \quad (19)$$

其中  $\beta > 0$  为学习率。

### 4 仿真研究(Simulation)

作者通过辨识一类具有多步时延的非线性动态系统<sup>31</sup>,仿真验证 RCFNN 处理动态非线性系统的能力,并将结果与静态模糊神经网络(CFNN)进行了比较。

非线性动态系统对象:

$$y_p(k+1) = f(y_p(k), y_p(k-1), y_p(k-2), u(k), u(k-1))$$

其中:

$$f(x_1, x_2, x_3, x_4, x_5) = \frac{x_1 x_2 x_3 x_5 (x_3 - 1) + x_4}{1 + x_2^2 + x_3^2}$$

辨识实验采用以下输入<sup>[21]</sup>信号:

$$u(k) = \begin{cases} \sin(\pi k/25), & (k < 250) \\ 1.0, & (250 \leq k < 500) \\ -1.0 & (500 \leq k < 750) \\ 0.3 \sin(\pi k/25) + 0.1 \sin(\pi k/32) \\ \quad + 0.6 \sin(\pi k/10), & (750 \leq k < 1000) \end{cases}$$

假设对于系统输出  $y_p$  存在观测变量  $x_1^{(2)} \sim x_3^{(2)}$ ,且可以获得  $y_p$  与观测变量之间的以下先验知识:

IF  $x_1^{(2)}$  is low and  $x_2^{(2)}$  is low and  $x_3^{(2)}$  is high,

THEN  $y$  is low

IF  $x_1^{(2)}$  is medium and  $x_3^{(2)}$  is medium,

THEN  $y$  is medium

IF  $x_1^{(2)}$  is high and  $x_2^{(2)}$  is high and  $x_3^{(2)}$  is low,

THEN  $y$  is high

选择网络规则数为 8,学习速率为 0.01,  $\beta$  初始

值为 0.3. 图 2 给出了 RCFNN 的辨识结果, 实线为期望输出, 点线为 RCFNN 输出.

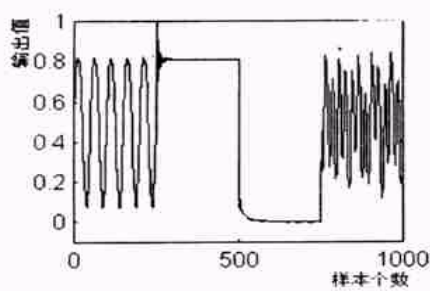


图 2 仿真曲线

Fig. 2 Simulation curve

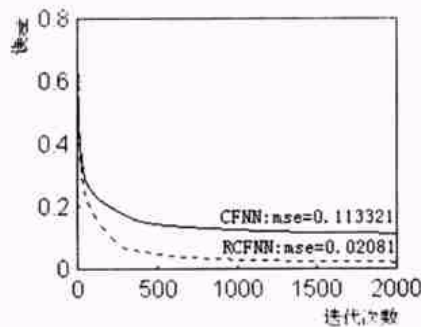


图 3 RCFNN 和 CFNN 的误差曲线

Fig. 3 MSE of RCFNN and CFNN

在相同条件下采用了 CFNN 对系统进行了实验, 图 3 给出了 RCFNN、CFNN 的误差曲线. 同样训练 2000 次, RCFNN 的训练时间为 49s, 误差为 0.02081; 而 CFNN 的训练时间为 44s, 误差为 0.113321. RCFNN 的训练时间虽然稍长, 但误差确是

CFNN 的 5 分之一, 可见, RCFNN 具有更高的精度. 而若使 CFNN 达到与 RCFNN 相同的精度, 则 CFNN 要增加网络节点, 这说明动态模糊神经网络比普通的模糊神经网络具有更小的网络结构.

## 5 结束语 (Conclusion)

递归型模糊神经网络是一个全局逼近器, 适用于动态非线性系统, 它的网络参数具有明确的物理意义, 可根据经验选择参数的初始值, 其规则网络能够充分利用过程的先验知识, 加快了网络的响应速度, 由于函数网络含有递归环节, 因而具有动态映射能力, 可处理暂态问题. 仿真结果证明此类 RCFNN 具有较高的收敛精度及较小的网络结构.

目前, 作者正在对此递归模糊神经网络进行噪声影响的研究, 并准备将其应用到实际系统中.

## 参 考 文 献 (References)

- 1 何 谦, 李宏光. 非线性系统建模的复合型模糊神经网络研究 [J]. 北京化工大学学报. 2000, 27(4): 67~ 71
- 2 Juang Chia-feng, Lin Chin-teng. A recurrent self-organizing neural fuzzy inference network [J]. IEEE Transactions on Neural Networks, 1999, 10(4): 828~ 845
- 3 Ching-hung Lee, Ching-cheng Teng. Identification and control of dynamic systems using recurrent fuzzy neural networks [J]. IEEE Trans. Fuzzy System, 2000, 8(4): 349~ 366

## 作者简介

王京慧(1978- ), 女, 硕士研究生. 研究领域为智能控制.

李宏光(1963- ), 男, 副教授. 研究领域为智能控制.

(上接第 180 页)

- for commercial application [A]. Proceedings of the South African Symposium on Communication and signal Processing [C]. 1992, 99
- 7 Moayer B, Fu K. A tree system approach for fingerprint pattern recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8: 376
  - 8 Verma M R, Majumdar A K, Chatterjee B. Edge detection in fingerprints [J]. Pattern Recognition, 1987, 20: 513
  - 9 Stefanelli R, Rosenfeld A. Some parallel thinning algorithms for digital pictures [J]. ACM, 1971, 18: 255
  - 10 Arcelli C, Bajaj S D. A width independent fast thinning algorithm [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1984, 7: 463

- 11 冯星奎, 李林艳, 颜祖泉. 一种新的指纹图像细化算法 [J]. 中国图形图像学报, 1999, 4(10): 835
- 12 黄煦涛. 二维数字信号处理 [M]. 北京: 科学出版社, 1985. 238 ~ 239

## 作者简介

沈 江(1957- ), 男, 博士生. 研究领域为信息安全管理、模式识别、工业工程管理等.

齐二石(1954- ), 男, 教授, 博士生导师. 研究领域为工业工程管理、CIMS、系统安全管理等.

师文轩(1977- ), 男, 硕士生. 研究领域为系统安全认证、计算机信息系统等.