

基于 T-S 逻辑的新型模糊神经网络模型

姜长元¹, 朱庆保²

(1. 湖州师范学院 理学院, 浙江 湖州 313000)

2 南京师范大学 数学与计算机科学学院, 江苏 南京 210097)

[摘要] 模糊神经网络是神经网络与模糊逻辑系统的有机结合, 具有强大的自学习和自整定功能. 针对目前模糊逻辑与神经网络技术在融合中存在的问题, 根据模糊逻辑与神经网络的本质和内在联系, 提出了一种新型的结构优化的模糊神经网络. 用模糊系统理论中的 Stone-Weierstrass 定理证明了该网络能以任意精度逼近任意一个定义在致密集上的实连续函数. 仿真实验表明该网络模型算法是可行且有效的.

[关键词] 模糊逻辑, 模糊神经网络, T-S 逻辑

[中图分类号] TP183 **[文献标识码]** A **[文章编号]** 1672-1292(2006)03-0051-05

A New Fuzzy Neural Network Model Based on T-S Logic

JIANG Changyuan¹, ZHU Qingbao²

(1. Faculty of Science, Huzhou Teachers College, Huzhou 313000, China)

2. School of Mathematics and Computer Science, Nanjing Normal University, Nanjing 210097, China)

Abstract Fuzzy neural network, composed of the neural network and fuzzy logic system, is good at self-learning and self-tuning. According to the inner relationship between the neural network and fuzzy logic inference, a new fuzzy neural network can be obtained with effective algorithm and structure. It can be proved to approximate at any degree of accuracy to any real continuous function in a compact domain by using the famous Stone-Weierstrass law. The simulation shows that the model is effective and feasible.

Key words fuzzy logic; fuzzy neural network; T-S logic

0 引言

近年来, 模糊神经网络在智能控制、模糊决策、专家系统、模式识别等领域发挥了重要作用, 其理论、模型、算法和应用技术一直是这些领域的重要研究课题, 并以它的特有优势受到高度的重视. 模糊神经网络的研究大多数都是围绕 Takagi-Sugeno 模糊逻辑系统进行的, 王立新^[1]证明了基于乘积型、高斯型隶属函数、中心反模糊化输出的模糊逻辑系统具有全局收敛性. KOSKO^[2]提出加性模糊神经网络, 并且证明了该模糊系统能以任意精度逼近一个致密集上的任意实连续函数. J Zhang^[3]提出了一种仅由乘积构成的模糊神经网络, 结构过于简单. CHEN Y C^[4]也提出过类似的网络结构, 但它们都有一些缺点, 如网络结构复杂、收敛速度慢、计算量大, 难以在实际中应用. 针对已有模糊神经网络的模型收敛速度慢、全局逼近能力差等缺点并在研究了许多模糊神经网络结构的基础上^[5, 6], 综合这些网络结构的优点, 提出了一种结构简单、算法清晰的新型模糊神经网络模型, 它具有收敛速度快、全局逼近能力强的优点, 理论证明和仿真结果表明该模型具有一定的实用价值.

1 模型原理

本文提出的模糊神经网络是一个 5 层的 BP 神经网络, 如图 1 所示, 分别为输入层、隶属度函数生成层、

收稿日期: 2005-11-20

作者简介: 姜长元(1975-), 讲师, 主要从事智能优化等方面的教学与研究. E-mail: jyc@hutc.zj.cn

通讯联系人: 朱庆保(1955-), 教授, 主要从事智能控制等方面的教学与研究. E-mail: zhuqingbao@njnu.edu.cn

推理层、求和层和反模糊化层. 其中推理层的节点个数 m 一般按照 K-means 方法对样本聚类后求得, 并可根据实际需要进行调整. 因为高斯型函数具有良好的平滑性, 故隶属函数采用高斯型函数, 如式 (1) 所示:

$$\mu_{ij} = \exp\left[-\frac{(x_i - m_{ij})^2}{\delta_{ij}^2}\right] \quad 1 \leq i \leq n, \quad 1 \leq j \leq m \tag{1}$$

式中, μ_{ij} 与隶属度生成层的各个节点相对应; m_{ij} 和 δ_{ij} 分别对应节点的隶属度函数的均值和方差, 也是在模型学习过程中要优化的两个参数. 隶属函数生成层中, 各节点对应的输出 μ_{ij} 的标注依次为: $\mu_{11}, \mu_{12}, \dots, \mu_{1m}; \mu_{21}, \mu_{22}, \dots, \mu_{2m}; \dots; \mu_{n1}, \mu_{n2}, \dots, \mu_{nm}$. 对于 m_{ij} 和 δ_{ij} 因为它们是 μ_{ij} 中的两个参数, 故它们的下标与 μ_{ij} 一样.

在此模糊神经网络中, 第一层至第三层网络的连接权值均为 1 0 为每一层的输出函数, 即输入到下一层的变量. 该模糊神经网络各层的输入输出分别如下:

第一层: 第一层的节点只是把输入值直接传送给下一层
 $O_i^1 = x_i \quad 1 \leq i \leq n$ (2)

第二层: 用单一节点计算简单的隶属函数, 因而有:

$$O_j^2 = \exp\left[-\frac{(O_i^1 - m_{ij})^2}{\delta_{ij}^2}\right] \quad 1 \leq i \leq n, \quad 1 \leq j \leq m \tag{3}$$

式中, m_{ij} 和 δ_{ij} 分别为第 i 个输入语言变量 x_i 的第 j 个术语的高斯型隶属函数的中心和均方差.

第三层分两个部分:

(1) 上半部分, 是用 Π 标记的所有节点 $O_{\Pi j}^3 = \mu_{1j} \cdot \mu_{2j} \cdot \dots \cdot \mu_{nj} = \prod_{i=1}^n \mu_{ij} \quad 1 \leq j \leq m$ (4)

(2) 下半部分, 是用 Σ 标记的所有节点: $O_{\Sigma j}^3 = \mu_{1j} + \mu_{2j} + \dots + \mu_{nj} = \sum_{i=1}^n \mu_{ij} \quad 1 \leq j \leq m$ (5)

第四层: 是将第三层的上半部分乘以系数 α ($0 < \alpha < 1$) 与下半部分乘以系数 $1 - \alpha$ 对应相加:

$$O_j^4 = \alpha O_{\Pi j}^3 + (1 - \alpha) O_{\Sigma j}^3 = \alpha \prod_{i=1}^n \mu_{ij} + (1 - \alpha) \sum_{i=1}^n \mu_{ij} \quad 1 \leq j \leq m \tag{6}$$

第五层: 输出层:

$$y = O_j^5 = \sum_{j=1}^m w_j O_j^4 = \sum_{j=1}^m w_j \left[\alpha \prod_{i=1}^n \mu_{ij} + (1 - \alpha) \sum_{i=1}^n \mu_{ij} \right] \tag{7}$$

该层实现了模糊神经网络最后清晰量的输出.

根据此模型, 可以看出:

(1) 常规的基于模糊系统的神经网络输出为 $y = \frac{\sum_{j=1}^m \left(w_j \cdot \prod_{i=1}^n \mu_{ij} \right)}{\sum_{j=1}^m \prod_{i=1}^n \mu_{ij}}$, 而该网络模型的输出为:

$$y = \sum_{j=1}^m w_j f_j = \sum_{j=1}^m w_j \left[\alpha \prod_{i=1}^n \mu_{ij} + (1 - \alpha) \sum_{i=1}^n \mu_{ij} \right], \text{ 因而该模型是一个新型的模糊神经网络.}$$

(2) 此网络的每层推导均有明显的物理意义, 较好地利用了模糊隶属度的信息, 克服加法推理收敛过慢和乘法推理收敛过速的缺点.

2 全局逼近性证明

虽然新型模糊神经网络模型具有上述特点, 但要用于实际建模, 还必须证明其具有全局逼近性质.

$$y = \sum_{j=1}^m w_j \left[\alpha \prod_{i=1}^n \mu_{ij} + (1 - \alpha) \sum_{i=1}^n \mu_{ij} \right] \quad 0 < \alpha < 1 \tag{8}$$

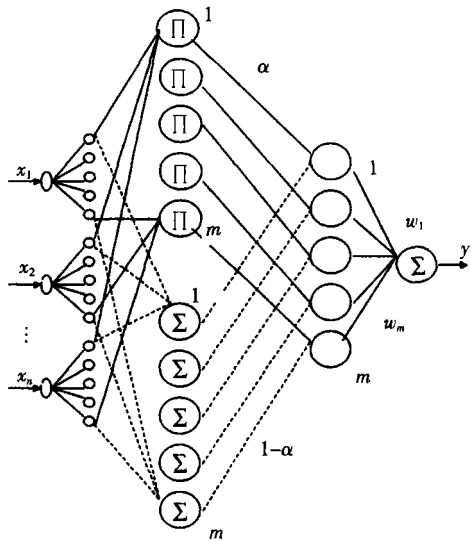


图 1 新型模糊神经网络结构

设 Y 为一个形如式 (8) 所示的模糊系统组成的集合. 下面用 Stone-Weierstrass 定理^[1] 证明该模型具有全局逼近性质. 首先证明 Y 符合该定理中给出的 3 个条件.

引理 1 (Y, d_∞) 是一个代数.

证明 任设 $f_1, f_2 \in Y$, 由此 f_1, f_2 详细表述如下:

$$f_1(x) = \sum_{j=1}^{m_1} w_j^1 \left[\alpha \prod_{i=1}^n \mu_{ij}^1(x_i) + (1-\alpha) \sum_{i=1}^n \mu_{ij}^1(x_i) \right] \quad (9)$$

$$f_2(x) = \sum_{j=1}^{m_2} w_j^2 \left[\alpha \prod_{i=1}^n \mu_{ij}^2(x_i) + (1-\alpha) \sum_{i=1}^n \mu_{ij}^2(x_i) \right] \quad (10)$$

于是有:

$$\begin{aligned} f_1(x) \cdot f_2(x) &= \left[\sum_{j=1}^{m_1} w_j^1 \left[\alpha \prod_{i=1}^n \mu_{ij}^1(x_i) + (1-\alpha) \sum_{i=1}^n \mu_{ij}^1(x_i) \right] \right] \cdot \\ &\left[\sum_{j=1}^{m_2} w_j^2 \left[\alpha \prod_{i=1}^n \mu_{ij}^2(x_i) + (1-\alpha) \sum_{i=1}^n \mu_{ij}^2(x_i) \right] \right] = \\ &\sum_{j_1=1}^{m_1} \sum_{j_2=1}^{m_2} w_{j_1}^1 w_{j_2}^2 \left[\left[\alpha \prod_{i=1}^n \mu_{ij_1}^1(x_i) + (1-\alpha) \sum_{i=1}^n \mu_{ij_1}^1(x_i) \right] \cdot \right. \\ &\left. \left[\alpha \prod_{i=1}^n \mu_{ij_2}^2(x_i) + (1-\alpha) \sum_{i=1}^n \mu_{ij_2}^2(x_i) \right] \right] \end{aligned} \quad (11)$$

由于 $\mu_{ij_1}^1(x_i)$ 和 $\mu_{ij_2}^2(x_i)$ 均为高斯型, 因而两者的乘积仍为高斯型. 所以式 (11) 的形式完全等同于式 (8). 由此可知, $f_1(x) \cdot f_2(x) \in Y$.

对于任意的 $c \in R$, 有:

$$cf_1(x) = \sum_{j=1}^{m_1} cw_j^1 \left[\alpha \prod_{i=1}^n \mu_{ij}^1(x_i) + (1-\alpha) \sum_{i=1}^n \mu_{ij}^1(x_i) \right] \quad (12)$$

显而易见, 式 (12) 的形式完全等同于式 (8). 因此, $cf_1(x) \in Y$. 同样,

$$\begin{aligned} f_1(x) + f_2(x) &= \sum_{j=1}^{m_1} w_j^1 \left[\alpha \prod_{i=1}^n \mu_{ij}^1(x_i) + (1-\alpha) \sum_{i=1}^n \mu_{ij}^1(x_i) \right] + \\ &\sum_{j=1}^{m_2} w_j^2 \left[\alpha \prod_{i=1}^n \mu_{ij}^2(x_i) + (1-\alpha) \sum_{i=1}^n \mu_{ij}^2(x_i) \right] \end{aligned} \quad (13)$$

式 (13) 的形式显然也等价于式 (8), 亦即: $f_1(x) + f_2(x) \in Y$. 至此, 引理 1 得证.

引理 2 (Y, d_∞) 能离析 U 上的点.

证明 构造一个所需的 f 来证明上述条件, 即给定定义在 U 和 R 上的模糊集合的数目、高斯型隶属函数的参数、模糊规则的数目以及模糊规则的表达式, 使所得的 f (形如式 (8) 的形式) 具有如下特征:

对任意给定的 $x^0, y^0 \in U$, 当 $x^0 \neq y^0$ 时, 有 $f(x^0) \neq f(y^0)$ 成立. 设 $x^0 = (x_1^0, x_2^0, \dots, x_n^0)$, $y^0 = (y_1^0, y_2^0, \dots, y_n^0)$. 若 $x_i^0 \neq y_i^0$ 时, 在 U 上的第 i 个子空间上定义两个模糊集合 $(A_i^1, \mu_{A_i^1}^1)$ 及 $(A_i^2, \mu_{A_i^2}^2)$, 其相应的隶属函数分别为:

$$\mu_{A_i^1}^1(x_i) = \exp \left[-\frac{(x_i - x_i^0)^2}{2} \right] \quad (14)$$

$$\mu_{A_i^2}^2(x_i) = \exp \left[-\frac{(x_i - y_i^0)^2}{2} \right] \quad (15)$$

这里, 如果 $x_i^0 = y_i^0$, 则有 $A_i^1 = A_i^2$, 且 $\mu_{A_i^1}^1 = \mu_{A_i^2}^2$, 也就是说此时的第 i 个子空间上只定义了一个模糊集合. 在输出论域 R 上定义了两个模糊集合 $(B^1, \mu_{B^1}^1)$ 及 $(B^2, \mu_{B^2}^2)$, 其相应的隶属函数如下:

$$\mu_{B^i}^i(w) = \exp \left[-\frac{(w - w_j)^2}{2} \right] \quad (16)$$

式中, $j = 1, 2, \dots, w_j$ 待以后确定.

此外, 设模糊规则库由两条模糊规则构成 (即 $m = 2$), 到此除 w_j 以外, 其余的设计参数均已确定, 亦即已得到了形如式 (8) 的 f , 其中 $m = 2$ 其表达形式如下:

$$f(x^0) = (\alpha + n(1 - \alpha))w_1 + w_2 \left[\alpha \prod_{i=1}^n \exp \left[-\frac{1}{2}(x_i^0 - y_i^0)^2 \right] + (1 - \alpha) \sum_{i=1}^n \exp \left[-\frac{1}{2}(x_i^0 - y_i^0)^2 \right] \right]$$
$$f(y^0) = (\alpha + n(1 - \alpha))w_2 + w_1 \left[\alpha \prod_{i=1}^n \exp \left[-\frac{1}{2}(x_i^0 - y_i^0)^2 \right] + (1 - \alpha) \sum_{i=1}^n \exp \left[-\frac{1}{2}(x_i^0 - y_i^0)^2 \right] \right]$$

假设 $f(x^0) = f(y^0)$, 并设 $\alpha \prod_{i=1}^n \exp \left[-\frac{1}{2}(x_i^0 - y_i^0)^2 \right] + (1 - \alpha) \sum_{i=1}^n \exp \left[-\frac{1}{2}(x_i^0 - y_i^0)^2 \right] = A$, 由于 $f(x^0) = f(y^0)$, 也即 $(\alpha + n(1 - \alpha))w_1 + w_2A = (\alpha + n(1 - \alpha))w_2 + w_1A$, 上式变为 $(w_1 - w_2) \cdot (\alpha + n(1 - \alpha) - A) = 0$. 由于 $x^0 \neq y^0$, 由此推知必然存在某些 i 使得 $x_i^0 \neq y_i^0$; 这样就有 $\prod_{i=1}^n \exp \left[-\frac{1}{2}(x_i^0 - y_i^0)^2 \right] < 1$, $\sum_{i=1}^n \exp \left[-\frac{1}{2}(x_i^0 - y_i^0)^2 \right] < n$. 故 $\alpha \prod_{i=1}^n \exp \left[-\frac{1}{2}(x_i^0 - y_i^0)^2 \right] + (1 - \alpha) \sum_{i=1}^n \exp \left[-\frac{1}{2}(x_i^0 - y_i^0)^2 \right] = A < \alpha + n(1 - \alpha)$.

所以, 不妨取 $w_1 = 2, w_2 = 1$ 时, 必然有 $(w_1 - w_2) \cdot (\alpha + n(1 - \alpha) - A) \neq 0$ 与假设矛盾, 也即 $f(x^0) \neq f(y^0)$ 得证. 本引理证毕.

引理 3 (Y, d_∞) 上所有点均不为零.
证明 仔细观察式 (8), 只要简单选取 $w_j > 0, i = 1, 2, \dots, m$ 即可. 即 $w_j > 0$ 所对应的任何 $f \in Y$ 均可作为所需的 f . 本引理证毕.

定理 1 图 1 所示的新型模糊神经网络是一个全局逼近器.
证明 从式 (8) 可知, Y 是定义在 U 上的一组连续实函数, 根据 Stone-Weierstrass 定理以及引理 1、2、3 可以直接推出其具有全局逼近之性质.
至此, 就证明了图 1 所示的新型模糊神经网络模型的全局逼近性.

3 仿真实验与结果分析

在此模糊神经网络中采用自适应调整学习效率的 BP 算法, 该算法主要过程为: 当连续两次迭代的梯度方向相同时, 表明下降太慢, 这时可使步长加倍; 当连续两次迭代的梯度方向相反时, 表明下降过头, 这时可使步长减半. 在网络学习过程中, 对 m_{ij}, δ_{ij} 以及 w_j 进行调整.

采用文献 [6] 的数据对本文提出的模糊神经网络模型进行了仿真实验, 在收敛速度、推理的精确度、误差大小等方面与其他文献的结果进行比较. 将数据分成两部分: 一部分为训练样本集, 训练样本集用于确定模糊规则; 另一部分为测试样本, 测试样本集用于比较推理的精度和网络的收敛速度, 即用来检验该模型是否具有预测能力和泛化能力.

用 K-means 法将样本进行聚类后, 得到的规则数为 5. 由 $m = 5$ 可得到该网络的结构, 它是一个三输入单输出的、规则数为 25 的模糊神经网络. 用训练样本的 20 组数据, 对网络进行训练, 可得结果如表 1 所示 (左边为训练样本, 右边为评价样本). 其中 y 为期望输出, y^* 为实际输出, x_1, x_2, x_3 为输入变量. 由评价样本数据可知, 该网络有较好的泛化能力. 用 40 组样本数据对网络进行训练的结果, 限于篇幅略.

表 1 训练 20 组数据后网络运行结果

x_1	x_2	x_3	y	y^*	x_1	x_2	x_3	y	y^*
1	3	1	11.110	12.136	1	1	5	9.545	9.862
1	5	2	6.521	6.848	1	3	4	6.043	5.941
1	1	3	10.190	10.250	1	5	3	5.724	5.624
1	3	4	6.043	5.941	1	1	2	11.250	11.601
1	5	5	5.242	5.236	1	3	1	11.110	12.136
5	1	4	19.020	18.538	5	5	2	14.360	14.510
5	3	3	14.150	14.223	5	1	3	19.610	19.081
5	5	2	14.360	14.510	5	3	4	13.650	13.736
5	1	1	27.426	26.046	5	5	5	12.430	12.191
5	3	2	15.390	15.736	5	1	4	19.020	18.538
1	5	3	5.724	5.624	1	3	3	6.380	6.142
1	1	4	9.766	10.048	1	5	2	6.521	6.848

x_1	x_2	x_3	y	y^*	x_1	x_2	x_3	y	y^*
1	3	5	5.870	5.755	1	1	1	16.000	16.924
1	5	4	5.406	5.423	1	3	2	7.219	7.368
1	1	3	10.190	10.250	1	5	3	5.724	5.624
5	3	2	15.390	15.736	5	1	4	19.020	18.538
5	5	1	19.680	19.513	5	3	5	13.390	13.270
5	1	2	21.060	20.703	5	5	4	12.680	12.609
5	3	3	14.150	14.223	5	1	3	19.610	19.081
5	5	4	12.680	12.609	5	3	2	15.390	15.736

模型的精度定义为: $J = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - y_i^*|}{y_i} \cdot 100(\%)$. 其中, y_i 为网络的期望输出, y_i^* 为对应网络的实际输出, n 为样本数. 可以得到各种模型的精度如表 2 所示.

表 2 各种模型精度

取 20 组数据为训练样本		取 40 组数据为训练样本	
文献 [6] 方法	4.362 788%	文献 [6] 方法	3.656 668%
新型网络模型方法	2.414 682%	文献 [5] 方法	3.894 576%
		新型网络模型方法	1.000 278%

由表 2 中的数据可知, 本文提出的模糊神经网络模型在推理精度方面比文献 [5]、[6] 所提供的方法有较大提高. 当选择 20 组样本作为网络训练数据时, 本文的方法所得到的误差平方和是文献 [6] 的 11%; 当选择 40 组样本作为网络训练数据时, 本文的方法所得到的误差平方和是文献 [5] 的 6%, 是文献 [6] 的 4%, 整体性能比文献 [6] 优. 根据模型精度的定义, 该网络模型精度较高.

4 结语

由全局逼近性的证明可见, 本文提出的新型模糊神经网络模型能以任意精度逼近任意一个定义在致密集上的实连续函数. 此网络结构简单, 层数和节点数均比常规模糊神经网络少, 网络采用隶属度加法推理和乘法推理, 克服加法推理收敛过慢和乘法推理收敛过速的缺点. 仿真实验表明: 学习和工作的速度较快, 收敛速度明显提高, 这一点在实际应用中尤为重要. 网络不需要对输入变量进行模糊化分割, 利用初始数据自动提取模糊推理规则, 而不用考虑常规模糊神经网络必须由相关领域的专家提供模糊推理规则, 这一点大大简化了网络的复杂性. 仿真结果表明, 本文提出的新型模糊神经网络模型具有推理精度高、收敛速度快的特点, 具有一定的应用价值.

[参考文献] (References)

[1] 王立新, 王迎军. 模糊系统与模糊控制教程 [M]. 北京: 清华大学出版社, 2003: 54-95
WANG Lixin, WANG Yingjun. A Course in Fuzzy Systems and Control [M]. Beijing: Tsinghua University Press, 2003: 54-95. (in Chinese)

[2] KOSKO B. Neural Network and Fuzzy Systems [M]. New York: Prentice-Hall Press, 1992: 23-47.

[3] ZHANG J, MORRIS A J. Fuzzy neural networks for nonlinear system modeling [J]. IEEE Proceedings: Control Theory and Applications, 1995, 142(6): 551-561.

[4] CHEN Y C, TENG C C. A model reference control structure using a fuzzy neural networks [J]. Fuzzy Sets and Systems, 1995, 73: 291-392.

[5] 於东军, 王士同. 新模糊神经网络模型及其建模应用研究 [J]. 华东船舶工业学院学报, 1998, 12(2): 22-28
YU Dongjun, WANG Shitong. A new fuzzy neural network model and its modelling application research [J]. Journal of East China Shipbuilding Institute, 1998, 12(2): 22-28. (in Chinese)

[6] 周志坚, 毛宗源. 一种基于遗传算法的模糊神经网络结构和参数优化 [J]. 华南理工大学学报, 1999, 27(1): 25-31.
ZHOU Zhijian, MAO Zongyuan. Fuzzy neural network structure and its parameters optimization by genetic algorithm [J]. Journal of South China University of Technology, 1999, 27(1): 25-31. (in Chinese)

[责任编辑: 刘 健]