

一种权值直接确定及结构自适应的 Chebyshev 基函数神经网络

张雨浓¹ 陈裕隆¹ 姜孝华¹ 曾庆淡¹ 邹阿金^{1,2}

(中山大学信息科学与技术学院 广州 510275)¹ (广东海洋大学信息学院 湛江 524088)²

摘要 基于函数逼近理论,构造一种 Chebyshev 基函数神经网络模型。推导出该网络模型的权值直接确定方法,可一步计算出权值,克服了传统 BP 神经网络学习率选取困难、学习过程冗长和易陷入局部极小等缺点。在此基础上,设计了基于二分搜索的结构自适应算法,根据精度要求自动确定网络最优结构。理论分析及仿真验证均表明,该网络不仅能够快速地完成网络权值确定和结构自适应,且具有优异的学习与逼近能力,而且对随机加性噪声也具有较好的抑制作用。

关键词 神经网络, Chebyshev 正交基, 权值直接确定, 结构自适应确定

中图分类号 TP183 **文献标识码** A

Weights-directly-determined and Structure-adaptively-tuned Neural Network Based on Chebyshev Basis Functions

ZHANG Yu-nong¹ CHEN Yu-long¹ JIANG Xiao-hua¹ ZENG Qing-dan¹ ZOU A-jin^{1,2}

(School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510275, China)¹

(Information College, Guangdong Ocean University, Zhanjiang 524088, China)²

Abstract A model of neural network based on Chebyshev basis functions was presented to remedy the drawbacks existing in conventional BP neural networks, such as, slow learning rate, local minima, and difficulties in determining optimal structure. The values of the weights in such a neural network can be quickly determined by the proposed weights-direct-determination method. The structure of such a neural network can also be adjusted automatically according to the required precision. Theoretical analysis and simulation results both show that the new neural network model and its algorithm can not only avoid some disadvantages of conventional BP networks, but also have a faster determination speed on the neural-network weights and structure.

Keywords Neural networks, Chebyshev basis, Weights-direct-determination, Structure-adaptive-determination

1 引言

BP(Back Propagation)神经网络是由 Rumelhart 和 McClelland 等人于 1986 年提出的一种前向神经网络模型^[1], 现已成为应用最为广泛的神经网络模型之一^[2-5]。然而,传统 BP 神经网络及其学习算法具有以下一些固有缺陷^[6]: 1) 学习率和网络结构的确定缺乏完善的理论指导; 2) 迭代过程冗长, 收敛速度较慢; 3) 易陷入局部极小而难以达到全局最优。针对上述问题, 国内外研究者已提出不少改进算法^[6,7]。由文献^[8,9]对常用 BP 改进算法的分析可知, 这些改进算法能在一定程度上优化传统 BP 算法, 但并未从根本上克服上述缺陷。

另外, 由于很多实际问题可转化为对非线性目标函数的逼近问题, 因此, 研究神经网络对非线性函数的逼近能力具有重要的实际意义^[10]。另一方面, 函数逼近论揭示了正交基的线性组合对非线性函数有着强大的逼近能力, 由此, 人们研究了大量以正交基作为神经元激励函数的神经网络模型^[11-13]。其中, 文献^[11,12]证明了在最高幂次相同的情况下, 基于

Chebyshev 正交基所构造的网络可比基于其它正交基构造的网络有更高精度的函数逼近能力。

基于上述认识, 本文构造和优化了一种 Chebyshev 基函数神经网络模型, 并推导出其网络权值的直接确定公式, 该公式能够一步计算得到网络最优权值, 从而克服了传统 BP 神经网络学习算法学习率选取困难、迭代学习过程冗长、易陷入局部极小等固有缺陷。此外, 设计了基于二分搜索的网络结构自适应确定算法, 根据学习的目标精度要求自动确定网络的最优结构(最简意义上)。

仿真结果也进一步表明, 本文构造的 Chebyshev 基函数神经网络具有快速、高精度的学习和逼近能力, 且能够根据工程实际要求, 自适应调整网络结构。

2 Chebyshev 基函数神经网络模型

本节首先给出 Chebyshev 基函数定义及其函数逼近相关定理, 然后再给出基于 Chebyshev 基函数的神经网络模型。

2.1 Chebyshev 正交多项式

定义 1^[14] 已知 $g(x)$ 与 $h(x)$ 是在区间 $[a, b]$ 上的函数,

到稿日期: 2008-11-14 返修日期: 2009-02-16 本文受国家自然科学基金(60643004, 60775050), 中山大学科研启动费、后备重点课题资助。
张雨浓 教授, 博士, 博士生导师, 从事神经网络、机器人、高斯过程等研究, E-mail: yangyou@121.com; 陈裕隆 硕士生, 从事神经网络、信息处理等研究; 姜孝华 教授, 博士, 博士生导师, 从事神经网络、信号检测与处理技术等研究。

定义 $g(x)$ 与 $h(x)$ 的带权内积为:

$$\langle g(x), h(x) \rangle = \int_a^b \rho(x) g(x) h(x) dx$$

其中 $\rho(x)$ 为权函数。若 $g(x)$ 与 $h(x)$ 的带权内积为 0, 则称 $g(x)$ 与 $h(x)$ 关于权函数 $\rho(x)$ 正交。其中, 在区间 $[-1, 1]$ 上,

关于权函数 $\rho(x) = \frac{1}{\sqrt{1-x^2}}$ 两两正交的多项式簇称为 Chebyshev 正交基, 它可表示为:

$$T_k(x) = \cos(k \arccos x), k=0, 1, 2, \dots$$

若令 $x = \cos \theta, 0 \leq \theta \leq \pi$, 则 $T_k(x) = \cos k\theta$, 可得到其递推公式如下:

$$\begin{cases} T_0(x) = 1 \\ T_1(x) = x \\ T_{k+2}(x) = 2xT_{k+1}(x) - T_k(x), k=0, 1, 2, \dots \end{cases}$$

定义 2^[15] 记 $C[a, b]$ 为在 $[a, b]$ 上连续的所有函数构成的空间, 设 $f(x)$ 是区间 $[a, b]$ 上的连续目标函数 (记为 $f(x) \in C[a, b]$), ϕ 是 $C[a, b]$ 的一个 n 维子空间, $\{\varphi_k(x)\}_{k=0}^{n-1}$ 为 ϕ 的一组基。取权值 $\{w_k\}_{k=0}^{n-1}$, 使得 $\int_a^b (f(x) - \sum_{k=0}^{n-1} w_k \varphi_k(x))^2 dx$ 最小, 则称 $\sum_{k=0}^{n-1} w_k \varphi_k(x)$ 为 $f(x)$ 在 $[a, b]$ 上的最佳平方逼近。

若定义 2 中的基取 Chebyshev 正交基, 即 $\{\varphi_k(x)\}_{k=0}^{n-1} = \{T_k(x)\}_{k=0}^{n-1}$, 有如下定理。

定理 1^[15] 设 $f(x) \in C[-1, 1]$ 且一阶导数连续, 若 $\sum_{k=0}^{n-1} w_k T_k(x)$ 为 $f(x)$ 在 $[-1, 1]$ 上的最佳平方逼近, 则当 $n \rightarrow \infty$ 时, 级数 $\sum_{k=0}^{n-1} w_k T_k(x)$ 在 $[-1, 1]$ 上一致收敛, 且

$$f(x) = w_0 T_0(x) + w_1 T_1(x) + \dots \quad (1)$$

在实际应用中, 随着 k 的增大, w_k 很快趋于 0, 我们可用式(1)的部分和逼近 $f(x)$ 。定理 1 说明了有限项 Chebyshev 正交基的加权和能以任意精度逼近定义在 $[-1, 1]$ 上的任意有连续导数的函数。

2.2 网络结构

基于上述 Chebyshev 正交基, 本文构造的神经网络模型如图 1 所示。该网络为单隐层神经网络, 其输入、输出层均采用线性函数 $u(x) = x$ 作为激励函数。假设满足精度要求的最小隐神经元数目为 n , 则各隐神经元激励函数 $\{\varphi_k(x)\}_{k=0}^{n-1}$ 可采用 Chebyshev 正交基的前 n 项, 即 $\{\varphi_k(x)\}_{k=0}^{n-1} = \{T_k(x)\}_{k=0}^{n-1}$ 。输入层到隐层的连接权值可恒为 1, 隐层到输出层的权值简记为 $\{w_k\}_{k=0}^{n-1}$, 则网络输出为:

$$y = \sum_{k=0}^{n-1} w_k \varphi_k(x) \quad (2)$$

显然, 只要确定了该网络的隐层神经元数目和隐层到输出层的权值, 就可以唯一确定该网络。

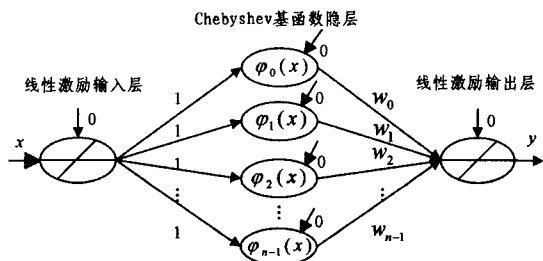


图 1 Chebyshev 基函数神经网络模型

3 权值直接确定法

为了克服传统 BP 算法学习率选取困难、收敛速度慢、易陷入局部极小等缺点, 对于上述 Chebyshev 基函数神经网络, 本文推导了无需迭代过程的权值直接确定公式, 可一步算出其网络最优权值。

下面给出相关定义、定理及权值直接确定公式的基础理论和推导过程。

定义 3^[16] 设长方阵 $A \in R^{m \times n}$ 。如果矩阵 B 同时满足以下 4 个 Moore-Penrose 方程: 1) $ABA = A$; 2) $BAB = B$; 3) $(AB)^T = AB$; 4) $(BA)^T = BA$, 则称 B 为 A 的加号逆 (也称 M-P 伪逆), 记为 $B = A^+$ 。上标 T 表示矩阵或向量的转置。

定理 2^[16] 对任意矩阵 $A \in R^{m \times n}$, 其加号逆存在且惟一。

定理 3^[16] $A \in R^{m \times n}, b \in R^{m \times 1}, x = A^+ b$ 是相容方程组 $Ax = b$ 的惟一极小范数解, 或是不相容方程组 $Ax = b$ 的惟一极小范数最小二乘解。

假设网络的隐神经元数目为 n , 学习样本个数为 m , 输入输出样本对为 $\{(x_i, d_i)\}_{i=0}^{m-1}$, 记隐层到输出层的连接权值向量 w , 样本输出向量 d 、网络的实际输出向量 y 分别为:

$$w = [w_0 \ w_1 \ \dots \ w_{n-1}]^T \in R^{n \times 1}$$

$$d = [d_0 \ d_1 \ \dots \ d_{m-1}]^T \in R^{m \times 1}$$

$$y = [y_0 \ y_1 \ \dots \ y_{m-1}]^T \in R^{m \times 1}$$

上述构造的 Chebyshev 神经网络的输入样本受激励矩阵 A 整合后如下所示:

$$A = \begin{bmatrix} \varphi_0(x_0) & \varphi_1(x_0) & \dots & \varphi_{n-1}(x_0) \\ \varphi_0(x_1) & \varphi_1(x_1) & \dots & \varphi_{n-1}(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_0(x_{m-1}) & \varphi_1(x_{m-1}) & \dots & \varphi_{n-1}(x_{m-1}) \end{bmatrix} \in R^{m \times n}$$

则网络在样本输入作用下的实际输出向量为:

$$y = Aw$$

记误差向量为:

$$e = [e_0 \ e_1 \ \dots \ e_{m-1}]^T = d - y = d - Aw;$$

标量意义下的误差性能指标可相应地设计为 $\epsilon = \sum_{i=0}^{m-1} e_i^2$, 由式(2)和向量二范数的定义, 可推得:

$$\epsilon = \sum_{i=0}^{m-1} e_i^2 = \sum_{i=0}^{m-1} (d_i - \sum_{k=0}^{n-1} w_k \varphi_k(x_i))^2 = \|d - Aw\|_2^2$$

其中, $\|\cdot\|_2$ 表示向量的二范数, 且理想情况下, 上述性能指标应为零 (即 $Aw = d$)。

由定理 3 可知, 不论 $Aw = d$ 是相容方程组还是不相容方程组, 上述神经网络逼近问题的权值解均可表示为:

$$w = A^+ d \quad (3)$$

本文中, 式(3)被称为权值直接确定公式。如此, 我们无须选取学习率, 也不必进行冗长的 BP 迭代计算, 便可由式(3)直接计算出神经网络的最优权值向量 w 使得误差 ϵ 最小。

4 网络结构自适应算法

传统 BP 神经网络隐神经元数目的设定缺乏完善的理论指导, 大多依据设计者经验而定, 且一旦设定, 在应用过程中不再进行调整。若隐神经元数目设定过少, 则无法达到精度要求; 设定过多又可能造成隐神经元冗余, 难以确保网络结构最优。

针对上述缺点, 在权值直接确定的基础上, 本文设计了一

种网络结构自适应算法,能够根据实际问题,快速有效地自动确定网络最优/最简结构,使得在保证隐神经元数目最少的时候又能满足目标精度要求。

算法流程结构图如图 2 所示,其中 floor 函数完成左取整运算。该网络结构自适应算法可说明如下,其中, m 为样本点数, n 为隐神经元数目, N 为输出的最优隐神经元数, $E > 0$ 为均方误差意义下的目标精度要求(如 $E = 10^{-15}$), j 初始化为 1:

- 1) 设定神经元数目 n 的范围 $[p, q] = [2^{j-1}, 2^j]$, 令 $n = q$, 依式(3)一步计算权值并计算对应的性能指标 ϵ ;
- 2) 若 $(\epsilon/m) > E$, 则 $j = j + 1$, 跳转至第 1 步, 否则跳入第 3 步;
- 3) 利用二分搜索法在范围 $[p, q] = [2^{j-1}, 2^j]$ 中确定输出最优神经元数 N 及其对应的权值和误差。

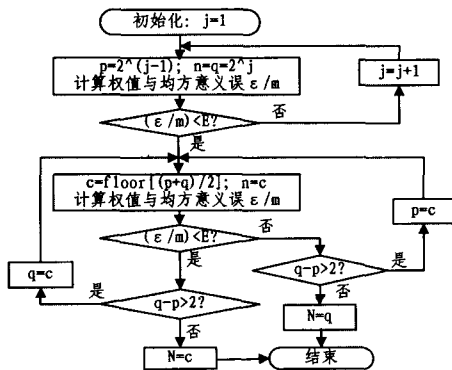


图 2 网络结构自适应算法流程图

5 仿真实验

为验证上述 Chebyshev 基函数神经网络及其算法的可行性,在区间 $[-1, 1]$ 上以间距 0.01 等距离采样 201 个数据作为样本点,即 $\{(x_i, f(x_i))\}_{i=0}^{200}$ 为输入输出样本对。对目标函数 $f(x)$ 进行了以下仿真实验(仿真硬件环境为 P4 1.6G CPU 和 512M 内存)。

实例一:对目标函数 $f(x) = \exp(x^2) \cos(6\pi x)/3$ 进行逼近仿真,仿真结果如图 3 所示,其中实线表示目标函数,虚线表示网络输出。图 3 中的(a)、(b)、(c)分别是当均方误差意义下的目标精度要求为 $E = 10^{-2}, 10^{-3}$ 和 10^{-15} 时的仿真结果,相应的由本文算法确定的最优隐神经元数目分别为 17,

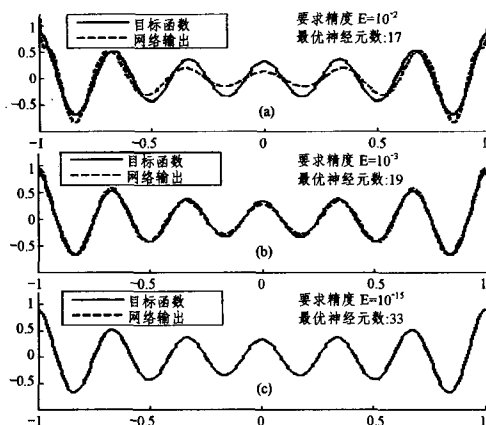


图 3 函数仿真实例 1

19 和 33。值得指出的是,即使当目标精度要求为 $E = 10^{-15}$ 时,仿真计算时间也仅为 0.422000 秒。图 3 仿真结果表明,本文构造的 Chebyshev 基函数神经网络及其算法能够快速高精度地逼近非线性连续函数,且能够根据精度要求自适应调整隐神经元数目,具有结构自适应特性。

事实上,对于某一特定隐神经元数目(如 n 时)的 Chebyshev 基函数神经网络,其也可以通过标准 BP 迭代算法训练权值(权值迭代公式如下):

$$w_j(k+1) = w_j(k) - \frac{\eta}{2} \frac{\partial \epsilon}{\partial w_j} \\ = w_j(k) - \eta \sum_{i=1}^m (\varphi_j(x_i) (\sum_{l=0}^{n-1} w_l(k) \varphi_l(x_i) - d_i))$$

其中, $j = 0, 1, 2, \dots, n-1$, 迭代次数 $k = 0, 1, 2, \dots$, $\eta > 0$ 表示学习率。

表 1 对比了 BP 迭代算法得到的结果和本文权值结构确定算法得到的结果。其中,在迭代算法中我们事先设定隐神经元数(或称隐节点数)分别为最优数目(如 19 与 33)并分别迭代 40000 次和 50000 次。由表 1 可知,当达到相近的目标精度时,本文算法所需的计算时间远少于迭代算法所需的计算时间。

表 1 BP 迭代算法与本文权值结构确定算法的对比

	隐节点数	迭代次数	训练时间(秒)	实际误差 ϵ/m
迭代算法	19	4×10^4	125.38	9.88×10^{-4}
	33	5×10^4	238.84	2.97×10^{-17}
本文算法	19	一步	0.34	9.88×10^{-4}
	33	算出	0.42	2.97×10^{-17}

实例二:对目标函数 $f(x) = 2x/(1+8x^2)$ 叠加一个强度为 -40dBW 的高斯白噪声后进行逼近仿真,以此考查神经网络在有噪声情况下逼近真实目标函数的性能(也即是去噪功能)。仿真结果如图 4 所示,其中实线表示目标函数,虚线表示网络实际输出,点线对应加噪声后的目标函数输出。当均方误差意义下的目标精度要求为 $E = 10^{-3}, 10^{-5}$ 时,对应的最优隐神经元数目分别为 6 和 16(实际误差分别为 6.894994×10^{-4} 和 8.710986×10^{-6})。该结果表明了 Chebyshev 基函数神经网络能够有效地去除随机加性噪声,在有噪声环境下仍能达到较高精度的学习逼近和去噪能力,并完成网络结构的自适应调整。

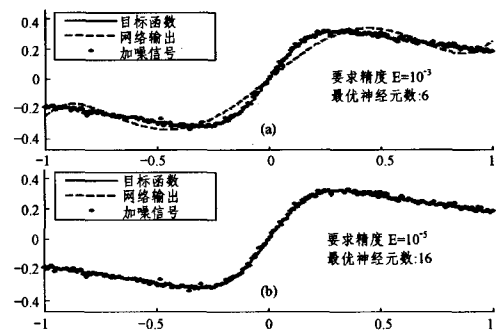


图 4 函数仿真实例 2

实例三:对非连续目标函数 $f(x) = \begin{cases} 1 & |x| < 0.2 \\ 0 & |x| \geq 0.2 \end{cases}$ (即矩

形窗函数)进行逼近仿真。仿真结果如图 5 所示。当均方误差意义下的目标精度要求为 $E = 10^{-3}, 10^{-4}$ 和 10^{-15} 时,相应的最优隐神经元数目分别为 123, 249 和 271。其中,目标精

度要求为 $E=10^{-15}$ 时,计算时间也仅为 4.562624 秒。以上仿真结果表明,本文构造的 Chebyshev 基函数神经网络对非连续函数同样具有较好的学习和逼近能力。这也是对上文中 Chebyshev 正交基逼近理论的延展。

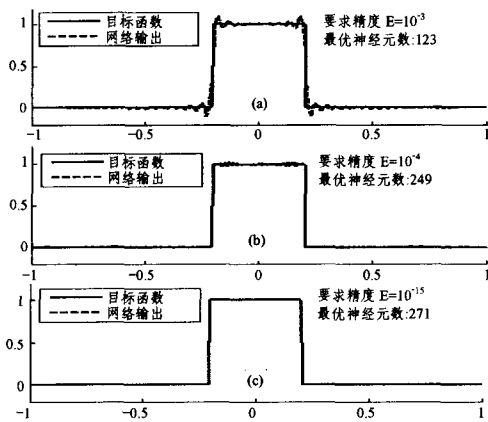


图 5 函数仿真实例 3

结束语 本文构造了一种 Chebyshev 基函数神经网络,并推导出权值直接确定公式,克服了传统 BP 网络学习算法学习率难以选取、迭代学习过程冗长、易陷入局部极小等固有缺陷。在此基础上,也设计了自动确定网络最优隐神经元数目的自适应算法。

仿真结果表明,该网络能够快速有效地完成网络权值的确定和结构的自适应调整,对任意非线性连续函数均具有优良的逼近性能;另外该网络还能够有效地去除随机加性噪声,对非连续函数(如窗函数)也同样具有较好的学习和逼近能力。

参 考 文 献

[1] Rumelhart D, McClelland E. Parallel Distributed Processing: Explorations in the Microstructure of Cognition[M]. Cambridge: MIT Press, 1986

[2] 杨治明,王晓蓉,彭军,等. BP 神经网络在图像分割中的应用[J]. 计算机科学, 2007, 34(3): 234-236

[3] 张雨浓,李巍,刘巍,等. 幂级数前向神经网络的权值直接确定法[C]//2007 年全国模式识别学术会议论文集. 北京: 科学出版社, 2007: 72-77

[4] Zhang Y, Wang J. Obstacle avoidance of kinematically redundant manipulators using a dual neural network[J]. IEEE Transactions on Systems, Man and Cybernetics, 2004, 34(1): 752-759

[5] Zhang Y, Ge S. Design and analysis of a general recurrent neural network model for time-varying matrix inversion [J]. IEEE Transactions on Neural Network, 2005, 16(6): 1477-1490

[6] Wen J, Zhao J, Luo S, et al. The improvements of BP neural network learning algorithm[C]//The 5th International Conference on Signal Processing Proceedings. Beijing, China, 2000

[7] Sun B, Wang X, Wang X, et al. A fast compositive training algorithm of forward neural network[C]//IEEE International Conference on Networking, Sensing and Control. Ft. Lauderdale, Florida, USA, 2006

[8] 高雪鹏,从爽. BP 网络改进算法的性能对比研究[J]. 控制与决策, 2001, 16(2): 167-171

[9] 王科俊,李国斌. 几种变学习率的快速 BP 算法比较研究[J]. 哈尔滨工程大学学报, 1997, 18(3): 31-35

[10] 曹飞龙,张永全,潘星. 构造前向神经网络逼近多项式函数[J]. 模式识别与人工智能, 2007, 20(3): 331-335

[11] 江善和,张杰. 基于 Chebyshev 基函数模糊神经网络的快速辨识方法[J]. 系统仿真学报, 2006, 18(3): 590-593

[12] 吴小俊,王士同,杨静宇,等. 基于正交多项式函数的神经网络及其性质研究[J]. 计算机工程与应用, 2002, 38(9): 25-26

[13] 肖少拥. 正交神经网络的动态动态建模方法研究[J]. 计算机科学, 2000, 27(1): 61-64

[14] 蒿小林,蒋耀林. 现代数值分析[M]. 北京: 国防工业出版社, 2004

[15] 李岳生,黄友谦. 数值逼近[M]. 北京: 人民教育出版社, 1978

[16] 方保镔,周继东,李医民. 矩阵论[M]. 北京: 清华大学出版社, 2004

(上接第 202 页)

$$\text{其中, } n = \frac{(2a-1)}{(1-a)^2}.$$

$$GC(x, y, 0.75) = ite(xy/e | x+y < 2e; (x+y-xy-e)/(1-e) | x+y > 2e; e) \quad (13)$$

a) 当两个团体都反对时,如权威专家团体给出的支持度 $x=0.2$,一般成员团体给出的支持度 $y=0.4$,根据式(12)可得考虑权值后的支持度分别为: $x'=0.0714, y'=0.4286$,再分别带入式(10)和式(13),则合成的结果为 0.0612;

b) 当两个团体都赞成时,如权威专家团体给出的支持度 $x=0.6$,一般成员团体给出的支持度 $y=0.8$,根据式(12)可得考虑权值后的支持度分别为: $x'=0.75, y'=0.75$,再分别带入式(10)和式(13),则合成的结果为 0.875;

c) 当两个团体意见发生冲突时,如权威专家团体给出的支持度 $x=0.2$,一般成员团体给出的支持度 $y=0.9$,根据式(12)可得考虑权值后的支持度分别为: $x'=0.0714, y'=0.8636$,再分别带入式(10)和式(13),则合成的结果为 0.1233。

从上面的合成结果可以看出,在考虑权值后,合成的结果更合理。

结束语 现有的泛组合运算模型描述的是一种理想的等

权情况,而实际复杂系统中的各因素一般都具有不同的权重。因此,本文对不等权泛组合运算模型进行了研究,在深入分析现有加权算子不足的基础上,给出了两种广义加权算子,并据此提出了一种不等权泛组合运算模型,从而可以更准确地处理复杂系统中各因素间关系的不确定性问题。

参 考 文 献

[1] 付利华. 复杂系统的柔性逻辑控制理论及应用研究[D]. 西安: 西北工业大学, 2005

[2] 何华灿,王华,刘永怀,等. 泛逻辑学原理[M]. 北京: 科学出版社, 2001

[3] Yager R R. Full reinforcement operators in aggregation techniques[J]. IEEE Transaction on Systems, Man, and Cybernetics, 1998, 28(6): 757-769

[4] 陈志成. 复杂系统中分形混沌与逻辑的相关性推理研究[D]. 西安: 西北工业大学, 2004

[5] 李士勇. 模糊控制、神经控制与智能控制论[M]. 哈尔滨: 哈尔滨工业大学出版社, 1996

[6] 付利华,何华灿. 模糊推理中相异因子的研究[J]. 计算机科学, 2004, 31(2): 98-100

[7] 付利华,何华灿. 模糊推理中零级泛蕴涵的信息度约束研究[J]. 计算机科学, 2005, 32(1): 162-164