

文章编号:1000-582X(2007)02-0065-05

模糊反向传播算法及其收敛性*

魏延^{1,2},李世宏^{1,2},曹长修¹,曾绍华¹

(1. 重庆大学自动化学院,重庆 400030; 2. 重庆师范大学数学与计算机科学学院,重庆 400047)

摘要:针对 S. Stoeva 提出的基于相同样本及网络输出的模糊神经网络模型,通过对基于极大-极小模糊算子的模糊神经网络模型的研究,证明了其与 S. Stoeva 提出的网络模型的等价性.在此基础上提出了依赖于模糊逻辑神经元输出的调整模糊权值的模糊反向传播学习算法,并进一步研究了其收敛性.最后以汽轮发电机组的状态监测为例进行仿真分析.结果表明:在网络输入神经元满足样本输出介于样本输入的极大与极小之间时,所提出的模糊反向传播学习算法是收敛的.

关键词:模糊神经网络;极大-极小模糊算子;学习算法;收敛性

中图分类号:TP18

文献标识码:A

模糊神经网络(Fuzzy Neural Network)是模糊和神经网络技术的融合,克服了神经网络和模糊逻辑在知识处理方面的缺点,是一个集语言计算、逻辑推理、分布式处理和非线性动力学过程为一体的系统^[1].作为新的智能信息处理方法,对人工智能和模式识别技术发展有重要作用,具有良好的发展前景^[2].近年来对模糊神经网络的研究主要集中在以下几个方面:将神经网络引入模糊系统的建模中,加强模糊系统的自学习能力;将模糊的概念引入神经网络的研究中,改进神经网络的性能;模糊神经网络的应用研究,主要集中在与机器人、工业控制、模式的分类和识别中.

一段时间以来,人们已对常见的模糊反向传播神经网络做了许多理论与应用工作,代表性的有文献[2-8]等.此类网络的特点用模糊极大-极小算子代替传统的神经网络积-和算子而构成的,它与传统的神经网络相比有很大区别,即信息在网络的传递过程中有较大的损失,与此同时,网络的鲁棒性得到了增强.最近 S. Stoeva^[3]提出了一个非常有创新意义的模糊反向传播神经网络 FBP(Fuzzy Backpropagation),并且对其学习规则的收敛性进行了研究.文献[4]提出了基于极大-极小模糊算子的模糊神经网络及其学习规则.但文献[3-4]的模糊神经网络是基于相同的样本输出,且权值的修改条件是基于输出值与所有权值的比较的.文中首先研究了基于极大-极小模糊算子的模

糊神经网络结构,并证明了其与文献[3]模糊神经网络模型的等价性,在此基础上,改进了模糊反向传播学习算法,对其收敛性进行了研究,最后以汽轮发电机组的状态监测为例进行了仿真.

1 模糊神经网络模型

1.1 模糊神经网络结构模型

文中所研究的模糊神经网络是由输入层、隐含层和输出层 3 层组成的多输入单输出模糊神经网络,具体结构描述如图 1 所示.

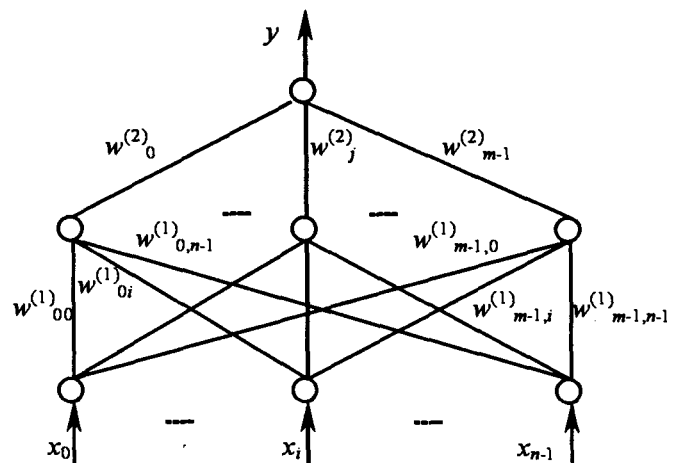


图 1 模糊神经网络结构模型

* 收稿日期:2006-08-28

基金项目:重庆市教委科学技术研究项目(KJ050809);重庆市高等学校优秀中青年骨干教师资助计划

作者简介:魏延(1970-),男,重庆师范大学副教授,重庆大学博士研究生,主要从事模糊推理、软计算方法及数据挖掘研究,

E-mail: weiyan@cqu.edu.cn

假设网络输入层有 n 个神经元,对输入训练样本 $P = (X, d)$, 输入 $X = (x_0, x_1, \dots, x_{n-1})$, 其中 $x_i \in [0, 1]$, $i = 0, 1, \dots, n - 1$. 隐含层有 m 个神经元, 输入层到隐含层的连接权值为 $w_{ji}^{(1)} \in [0, 1]$, 其中 $i = 0, 1, \dots, n - 1$, $j = 0, 1, \dots, m - 1$. 隐含层第 j 个模糊神经元的净输入为:

$$net_j^{(1)} = \bigvee_{i=0}^{n-1} (x_i \wedge w_{ji}^{(1)}), j = 0, 1, \dots, m - 1. \quad (1)$$

隐含层第 j 个模糊神经元的净输出为:

$$o_j^{(2)} = f(net_j^{(1)}), j = 0, 1, \dots, m - 1$$

输出层只有 1 个神经元, 隐含层与输出层的连接权值为 $w_j^{(2)} \in [0, 1]$, 其中 $j = 0, 1, \dots, m - 1$. 输出层的净输入为:

$$net^{(2)} = \bigvee_{j=0}^{m-1} (o_j^{(2)} \wedge w_j^{(2)}), \quad (2)$$

输出层的输出为:

$$y = f(net^{(2)}).$$

以上 $f(x)$ 为传递函数, 也称变换函数, 通常采用硬限幅函数、Sigmoid 函数、高斯函数、线性函数等. 文中为保证传递函数的值分布在 $[0, 1]$ 区间内, 取传递函数 $f(x)$ 为线性函数为:

$$f(x) = \frac{1}{1 + e^{-6(x-0.5)}}.$$

1.2 与文献[3]模糊神经网络计算过程的等价性

模糊神经网络的动力学过程有 2 类: 一类是计算过程, 该过程将实现神经网络的活跃状态的模式变换, 即由神经网络的样本输入得到网络的输出; 另一类是学习过程, 该过程中神经元之间的连接权值得到调整, 使之与环境信息相符合, 通常说的学习算法就是连接权的调整方法.

在文献[3]讨论的反向传播算法中, 引入了一个集合 $P(I)$, $P(I)$ 是网络输入指针集 I 的幂集, I_{ji} 是网络隐含层第 j 个神经元的第 i 个输入, 接下来定义一个函数 $g: P(I) \rightarrow [0, 1]$ 如下:

$$\begin{aligned} g(\emptyset) &= 0, \\ &\vdots \\ g(\{I_{ji}\}) &= w_{ji}^{(1)}, i = 0, 1, \dots, n - 1 \\ &\vdots \\ g(\{I_{j1}, I_{j2}, \dots, I_{jr}\}) &= w_{j1}^{(1)} \vee w_{j2}^{(1)} \vee \dots \vee w_{jr}^{(1)}, \\ &\{i_1, i_2, \dots, i_r\} \subset \{0, 1, \dots, n - 1\} \\ &\vdots \\ g(\{I_{j0}, I_{j1}, \dots, I_{j,n-1}\}) &= 1, \end{aligned} \quad (3)$$

函数 g 体现了网络权值与输入之间的幂函数关系. 同样定义一个函数 $h: P(J) \rightarrow [0, 1]$, $P(J)$ 是网络隐含层指针集 J 的幂集, 定义方法同上. 这样对隐含层神经元有:

$$net_j^{(1)} = \bigvee_{G \in P(I)} [(\bigwedge_{p \in G} x_p) \wedge g(G)], \quad (4)$$

$$o_j^{(2)} = f(net_j^{(1)}), j = 0, 1, \dots, m - 1.$$

对输出层神经元则有:

$$\begin{aligned} net^{(2)} &= \bigvee_{G \in P(J)} [\bigwedge_{p \in G} o_p^{(2)} \wedge h(G)], \\ y &= f(net^{(2)}). \end{aligned} \quad (5)$$

事实上, 文献[3]中的结论将(3)式改为 $g(\{I_{j0}, I_{j1}, \dots, I_{j,n-1}\}) = w_{j0}^{(1)} \vee w_{j1}^{(1)} \vee \dots \vee w_{j,n-1}^{(1)}$ 也是成立的, 读者可自行证明. 下面证明文献[3]的计算过程与文中的网络计算过程是等价的.

定理 1 文献[3]的模糊神经网络计算过程与文中的模糊神经网络计算过程是等价的.

证明 要证明网络计算过程的等价性, 只需证明神经元净输入计算(1)和(4)式以及(2)和(5)式的等价性.

先考虑(1)式和(4)式的等价性, 由函数 g 的定义和幂集的特点, 将(4)式的右边展开得:

$$\begin{aligned} net_j^{(1)} &= \bigvee_{G \in P(I)} [(\bigwedge_{p \in G} x_p) \wedge g(G)] = \bigvee_{i=0}^{n-1} (x_i \wedge w_{ji}^{(1)}) \vee \bigvee_{i,k=0, i < k}^{n-1} [(x_i \wedge x_k) \wedge (w_{ji}^{(1)} \vee w_{jk}^{(1)})] \\ &\vee \bigvee_{i,k,l=0, i < k < l}^{n-1} [(x_i \wedge x_k \wedge x_l) \wedge (w_{ji}^{(1)} \vee w_{jk}^{(1)} \vee w_{jl}^{(1)})] \vee \dots \vee [(\bigwedge_{i=0}^{n-1} x_i) \wedge (\bigvee_{i=0}^{n-1} w_{ji}^{(1)})] \end{aligned}$$

从而有:

$$\begin{aligned} net_j^{(1)} &= \bigvee_{i=0}^{n-1} (x_i \wedge w_{ji}^{(1)}) \vee \bigvee_{i,k=0, i < k}^{n-1} (x_i \wedge x_k \wedge w_{ji}^{(1)}) \\ &\vee \bigvee_{i,k=0, i < k}^{n-1} (x_i \wedge x_k \wedge w_{jk}^{(1)}) \vee \dots \vee [(\bigwedge_{i=0}^{n-1} x_i) \wedge w_{j0}^{(1)}] \vee [(\bigwedge_{i=0}^{n-1} x_i) \wedge w_{j1}^{(1)}] \vee \dots \vee [(\bigwedge_{i=0}^{n-1} x_i) \wedge w_{j,n-1}^{(1)}] = \\ &\{(x_0 \wedge w_{j0}^{(1)}) \vee \bigvee_{i=0}^{n-1} (x_0 \wedge x_i \wedge w_{j0}^{(1)}) \vee \bigvee_{i,k=0, i < k}^{n-1} (x_0 \wedge x_i \wedge x_k \wedge w_{j0}^{(1)}) \vee \dots \vee [(\bigwedge_{i=0}^{n-1} x_i) \wedge w_{j0}^{(1)}]\} \\ &\vee \{(x_1 \wedge w_{j1}^{(1)}) \vee \bigvee_{i=0, i \neq 1}^{n-1} (x_1 \wedge x_i \wedge w_{j1}^{(1)}) \vee \bigvee_{i,k=0, i, k \neq 1, i < k}^{n-1} (x_1 \wedge x_i \wedge x_k \wedge w_{j1}^{(1)}) \vee \dots \vee [(\bigwedge_{i=0}^{n-1} x_i) \wedge w_{j1}^{(1)}]\} \vee \dots \vee \{(x_{n-1} \wedge w_{j,n-1}^{(1)}) \vee \bigvee_{i=0, i \neq n-1}^{n-1} (x_{n-1} \wedge x_i \wedge w_{j,n-1}^{(1)}) \vee \bigvee_{i,k=0, i, k \neq n-1, i < k}^{n-1} (x_{n-1} \wedge x_i \wedge x_k \wedge w_{j,n-1}^{(1)}) \vee \dots \vee [(\bigwedge_{i=0}^{n-1} x_i) \wedge w_{j,n-1}^{(1)}]\} \end{aligned}$$

由极大-极小算子运算的性质, 有:

$$\begin{aligned} net_j^{(1)} &= (x_0 \wedge w_{j0}^{(1)}) \vee (x_1 \wedge w_{j1}^{(1)}) \vee \dots \vee (x_{n-1} \wedge w_{j,n-1}^{(1)}) = \bigvee_{i=0}^{n-1} (x_i \wedge w_{ji}^{(1)}) \\ \text{即 } net_j^{(1)} &= \bigvee_{i=0}^{n-1} (x_i \wedge w_{ji}^{(1)}), \text{ 从而(1)式与(4)式} \end{aligned}$$

等价。

同理可证(2)式和(5)式是等价的. 证毕

由定理1可见,文中极大-极小算子的网络计算比文献[3]就简单、直观得多,而且具有很强的可操作性。

2 模糊反向传播学习算法

众所周知,对模糊神经网络的研究,网络的学习算法及其是否收敛是至关重要的. 文献[3-4,7-8]提出相应的基于极大-极小模糊算子的模糊神经网络学习规则。

在此基础上,考虑基于模糊神经网络训练输出与实际输出的关系,依据各神经元的相应输出与权值的比较来修改其权值,提出了改进的模糊神经网络学习规则。

设 d 为训练样本 P 的实际输出,并令

$$\delta_j^{(1)} = |d - o_j^{(2)}|, j = 0, 1, \dots, m-1,$$

$$\delta^{(2)} = |d - y|.$$

则改进的模糊神经网络学习算法为:

1) 若 $y = d$, 则权值不再进行调整;

2) 若 $y < d$, 则

① 对权值 $w_j^{(2)}$ ($j=0, 1, \dots, m-1$), 当 $w_j^{(2)} < o_j^{(2)}$ 时, $w_j^{(2)'} = (w_j^{(2)} + \alpha\delta^{(2)}) \wedge 1$; 当 $w_j^{(2)} \geq o_j^{(2)}$ 时, $w_j^{(2)'} = w_j^{(2)}$.

② 对权值 $w_{ji}^{(1)}$ ($i=0, 1, \dots, n-1; j=0, 1, \dots, m-1$), 当 $w_{ji}^{(1)} < x_i$ 时, $w_{ji}^{(1)'} = (w_{ji}^{(1)} + \alpha\delta_j^{(1)}) \wedge 1$; 当 $w_{ji}^{(1)} \geq x_i$ 时, $w_{ji}^{(1)'} = w_{ji}^{(1)}$.

3) 若 $y > d$, 则

① 对权值 $w_j^{(2)}$ ($j=0, 1, \dots, m-1$), 当 $w_j^{(2)} > o_j^{(2)}$ 时, $w_j^{(2)'} = (w_j^{(2)} - \alpha\delta^{(2)}) \vee 0$; 当 $w_j^{(2)} \leq o_j^{(2)}$ 时, $w_j^{(2)'} = w_j^{(2)}$.

② 对权值 $w_{ji}^{(1)}$ ($i=0, 1, \dots, n-1; j=0, 1, \dots, m-1$), 当 $w_{ji}^{(1)} > x_i$ 时, $w_{ji}^{(1)'} = (w_{ji}^{(1)} - \alpha\delta_j^{(1)}) \vee 0$; 当 $w_{ji}^{(1)} \leq x_i$ 时, $w_{ji}^{(1)'} = w_{ji}^{(1)}$.

其中 $\alpha \in [0, 1]$ 为调整步幅系数, 或称步长. α 的大小将影响学习算法的收敛速度. 该学习算法与文献[3-4]的学习算法的区别是: 网络权值的修改条件不是依赖网络输出, 而是依赖相应神经元的输出. 这样, 权值的修改将更加精确, 网络学习更加平稳。

相对于反向传播的 BP 算法, 文中学习算法是相当简洁的, 而且有很强的直观性和可操作性。

3 模糊反向传播学习算法的收敛性

文献[3-4]研究了所提出的学习算法的收敛性, 是基于单训练模式的, 对多训练模式的证明是在假定了样本有相同输出的前提下进行的. 显然这有背于神经

网络的自学习、自适应特性。

定义 对模糊神经网络训练样本 $(X_k; d_k), k=1, 2, \dots, K$, 存在训练学习次数 s , 使第 s 次以后的任意样本 $(X_k; d_k), k=1, 2, \dots, K$, 均满足 $y_k = d_k$, 则称学习算法是收敛的。

定理2 对模糊神经网络训练样本 $(X_k; d_k), k=1, 2, \dots, K$, 则学习算法收敛的充要条件是任意训练样本 $(X_k; d_k), k=1, 2, \dots, K, X_k = (x_{0k}, x_{1k}, \dots, x_{n-1,k})$, 满足

$$\bigwedge_{i=0}^{n-1} x_{ik} \leq d_k \leq \bigvee_{i=0}^{n-1} x_{ik}, k = 1, 2, \dots, K. \quad (6)$$

证明 1) 充分性

用反证法来证明. 假设样本集不满足(6)式, 显然存在样本 $(X_k; d_k), k \in \{1, 2, \dots, K\}, X_k = (x_{0k}, x_{1k}, \dots, x_{n-1,k})$, 对任意的 $i, 0 \leq i \leq n-1$, 都有 $x_{ik} < d_k$ 或都有 $x_{ik} > d_k$.

若 $x_{ik} < d_k, i=0, 1, \dots, n-1$, 有:

$$o_j^{(2)} = \bigvee_{i=0}^{n-1} (x_{ik} \wedge w_{ji}^{(1)}) < \bigvee_{i=0}^{n-1} (d_k \wedge w_{ji}^{(1)}) =$$

$$d_k \wedge \left(\bigvee_{i=0}^{n-1} w_{ji}^{(1)} \right) \leq d_k, j=0, 1, \dots, m-1$$

从而有:

$$y_k = \bigvee_{j=0}^{m-1} (o_j^{(2)} \wedge w_j^{(2)}) < d_k \wedge \left(\bigvee_{j=0}^{m-1} w_j^{(2)} \right) \leq d_k,$$

显然这样模糊神经网络不具有收敛性。

同理, 对 $x_{ik} > d_k, i=0, 1, \dots, n-1$, 同样地证明不具有收敛性。

2) 必要性

假设模糊神经网络不收敛, 即不存在训练次数 s , 使 $y_k = d_k, k=1, 2, \dots, K$, 则对某次训练或者 $y_k < d_k$ 成立, 或者 $y_k > d_k$ 成立。

对 $k \in \{1, 2, \dots, K\}$, 若 $y_k < d_k$ 成立, 则对任意的训练次数, 均有 $y_k < d_k$ 成立. 事实上, 设在 s' 次训练时

$$y_k(s') = \bigvee_{j=0}^{m-1} (o_j^{(2)} \wedge w_j^{(2)}) < d_k. \quad (7)$$

对 $y_k(s'+1) = \bigvee_{j=0}^{m-1} (o_j^{(2)'} \wedge w_j^{(2)'})$, 必存在 $j \in \{0, 1, \dots, m-1\}$, 使

$$y_k(s'+1) = o_j^{(2)'} \wedge w_j^{(2)'}. \quad (8)$$

对这个 j , 由(7)式, 可知 $o_j^{(2)} \wedge w_j^{(2)} \leq y_k < d_k$, 此时, 若 $o_j^{(2)} < w_j^{(2)}$, 即 $o_j^{(2)} \leq y_k < d_k$, 由文献[3]可知 $o_j^{(2)'} < d_k$, 由(8)式知 $y_k(s'+1) < d_k$; 若 $o_j^{(2)} > w_j^{(2)}$, 即 $w_j^{(2)} \leq y_k \leq d_k$, 由模糊神经网络学习规则有:

$$w_j^{(2)'} = (w_j^{(2)} + \alpha\delta^{(2)}) \wedge 1 =$$

$$(w_j^{(2)} + \alpha |d_k - y_k|) \wedge 1 \leq$$

$$\begin{aligned} (w_j^{(2)} + |d_k - y_k|) \wedge 1 &= \\ (w_j^{(2)} + d_k - y_k) \wedge 1 &\leq \\ (y_k + d_k - y_k) \wedge 1 &= d_k \wedge 1 = d_k \end{aligned}$$

显然由(8)式知 $y_k(s'+1) < d$.

从而对任意训练次数,由模糊神经网络输出公式及学习算法权值修改规则,对 $j=0,1,\dots,m-1$,有 $o_j^{(2)} < d_k$,显然进而易得对 $i=0,1,\dots,n-1$,有 $x_i < d_k$,这与(6)式矛盾.

同样地,若 $y_k > d_k$ 成立,易得对 $i=0,1,\dots,n-1$,有,这也与(6)式矛盾.

综上所述,当且仅当训练样本 $(X_k; d_k), k=1,2,\dots,K, X_k = (x_{0k}, x_{1k}, \dots, x_{n-1,k})$, 满足条件 $\bigwedge_{i=0}^{n-1} x_{ik} \leq d_k \leq \bigvee_{i=0}^{n-1} x_{ik}, k=1,2,\dots,K$,模糊神经网络学习算法是收敛的. 证毕

4 在汽轮发电机组状态监测中的应用

文献[9]中对汽轮发电机组的状态测试了6组样本数据,每个样本用10个指标表示时域、频域参数.根据不同指标的特性建立相应隶属函数,进而获得每个样本10个数据(如表1所示).并建立了基于乘积-求和的模糊BP神经网络,从而判断发电机组运行的良好、正常或不正常状态.

文中用改进的基于极大-极小模糊算子的模糊神经网络来实现汽轮发电机组的状态监测.改进的模糊神经网络结构由3层组成,其中输入层采用10个模糊神经元,对应每个样本的10个指标参数,隐含层模糊神经元数目仍按文献[9]中的选取原则,用5个模糊神经元表示,输出层由于是监测发电机组运行的状态,所以采用1个模糊神经元作为网络输出.

表1 气轮发电机组各状态的隶属函数值^[9]

模糊子集	样本1	样本2	样本3	样本4	样本5	样本6	隶属函数
波形偏差小	x_0	0.90	0.34	0.14	0.25	0.88	$\exp(-0.0002u^2)$
波动(方差)小	x_1	0.68	0.35	0.45	0.81	0.71	$(400-u)/300$
相关性好	x_2	0.29	0.93	0.96	0.49	0.84	$10u-9$
一倍频幅值小	x_3	0.62	0.29	0.46	0.82	0.69	$(500-u)/400$
二倍频幅值小	x_4	0.89	0.49	0.52	0.46	0.48	$(200-u)/200$
三倍频幅值小	x_5	0.66	0.57	0.50	0.42	0.52	$(60-u)/60$
四倍频幅值小	x_6	0.88	0.18	0.43	0.47	0.69	$(50-u)/50$
五倍频幅值小	x_7	0.87	0.34	0.76	0.76	0.62	$(45-u)/45$
六倍频幅值小	x_8	0.90	0.19	0.57	0.41	0.65	$(40-u)/40$
分频幅值小	x_9	0.39	0.93	0.24	0.21	0.65	$(30-u)/30$
输出 贴近度 d		0.639	0.402	0.451	0.468	0.654	0.343

表2 气轮发电机组模糊神经网络知识库结构

j	0	1	2	3	4	5	6	7	8	9	$w_j^{(2)}$
0	0.0011	0.3071	0.4020	0.3172	0.2529	0.2518	0.0084	0.4927	0.0907	0.1590	0.4793
1	0	0.2939	0.4587	0.3284	0.2162	0.1875	0.0156	0.4783	0.0364	0.1462	0.6244
2	0	0.3286	0.3591	0.4494	0.2281	0.2244	0.0405	0.5044	0.0685	0.1380	0.3901
3	0.0012	0.3080	0.4073	0.3230	0.2313	0.2609	0.0083	0.4907	0.0912	0.1641	0.4798
4	0	0.3805	0.3841	0.3820	0.2134	0.2431	0.0462	0.4963	0.0354	0.1308	0.3801

对改进的模糊神经网络进行训练,样本为表1中的数据,训练所得的模糊权值数据的集合即为模糊知识库.通过 Visual C++ 语言程序仿真实现,得到一种可能的模糊知识库表示,如表2所示.

在程序仿真过程中,考虑了极大-极小模糊算子对训练样本的偶然性,对样本的使用采用了均值的方法.对6个样本训练后,通过每个样本 $\delta^{(2)} = |d - y|$ 的计算平均值,然后利用这个均值来对每个样本训练时的模糊权值进行调整.同时考虑了训练学习精度,若

$|y - d| \leq \eta (0 < \eta < 1)$, 则权值不再进行调整.

针对模糊神经网络训练获得的模糊知识库,根据网络的输出 y 值利用反模糊化原则来得到发电机组的运行状态:

$$g(y) = \begin{cases} \text{良好} & y > 0.60 \\ \text{正常} & 0.40 \leq y \leq 0.60 \\ \text{不正常} & y < 0.40 \end{cases}$$

例如,对一组新的监测样本^[9]{波形偏差,波动方差,相关性,一倍频幅值,二倍频幅值,三倍频幅值,四

倍频幅值,五倍频幅值,六倍频幅值,分频幅值} = {48.50, 230.64, 0.958, 260.75, 59.67, 40.64, 30.12, 18.16, 20.18, 12.78}. 依照表1相应隶属函数对这些实际数据进行模糊化,得到模糊输入为{0.62, 0.56, 0.58, 0.60, 0.70, 0.32, 0.40, 0.60, 0.50, 0.57},采用表2模糊知识库进行模糊神经网络计算,得到量化模糊输出为0.469 8,最后得到汽轮发电机组运行状况正常.

文献[9]中采用基于乘积-求和的模糊BP神经网络对上例得到的量化输出为0.545,得到汽轮发电机组运行状况正常,与文中的监测结果是一致的.其输出结果的差异主要是由于网络中使用乘积-求和算子与极大-极小算子的差异带来的.通过仿真比较,文献[9]的学习训练次数为2 310次,而文中采用基于极大-极小模糊算子的神经网络学习次数为1 486次,学习速度更快,效果更好.

5 结束语

在基于极大-极小模糊算子的模糊神经网络模型的基础上,证明了文中的模糊神经网络模型与文献[3]提出的网络模型的等价性.同时提出了基于模糊神经网络训练输出与实际输出的关系,依据各神经元的相应输出与权值的比较来修改其权值的改进的模糊反向传播学习算法,并对其收敛性进行了研究,同时在

汽轮发电机组的状态检测中进行的仿真.结果表明,在网络输入满足一定条件时模糊神经网络是收敛的.

参考文献:

- [1] 高隽.智能信息处理方法导论[M].北京:机械工业出版社,2004.
- [2] 丛爽.几种模糊神经网络系统关系的对比研究[J].信息与控制,2001,30(6):486-491.
- [3] STOIEVA S, NIKOV A. A fuzzy backpropagation algorithm[J]. Fuzzy Sets and Systems, 2000, 112(1): 27-39.
- [4] 朱晓铭,王士同.单体模糊神经网络的学习规则及其收敛性研究[J].计算机研究与发展,2001,38(9):1057-1060.
- [5] STOIEVA S. A weight-learning algorithm for fuzzy production systems with weighting coefficients[J]. Fuzzy Sets and Systems, 1992, 48(1): 87-97.
- [6] WANG SHITONG. Fuzzy basis functions, fuzzy neural networks, and the new hybrid learning algorithm[J]. Journal of Software Research, 1999, 30(3): 325-334.
- [7] KELLER M, TAHANI H. Backpropagation neural network for fuzzy logic [J]. Information Science, 1992, 62(3): 205-221.
- [8] 梁久祯,何新贵,黄德双.关于单体模糊神经网络感知机收敛定理的讨论[J].电子学报,2002,30(3):407-409.
- [9] 吴今培,肖健华.智能故障诊断与专家系统[M].北京:科学出版社,1997.

Fuzzy Backpropagation Algorithms and Their Convergence

WEI Yan^{1,2}, LI Shi-hong^{1,2}, CAO Chang-xiu¹, ZENG Shao-hua¹

(1. College of Automation, Chongqing University, Chongqing 400030, China;

2. College of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047, China)

Abstract: The equivalence between fuzzy neural networks model for max-min fuzzy operator and S. Stoeva's is proved by studying the fuzzy neural networks model for max-min fuzzy based on S. Stoeva's. Then the paper proposes the fuzzy backpropagation learning algorithms for changing fuzzy power and probes their convergence properties. Finally, it simulates experiment such as state monitoring of turbo-generator set. The results show that the fuzzy backpropagation learning algorithms presented are convergent on condition that the output of training sample is between maximum and minimum of its input.

Key words: fuzzy neural networks; max-min fuzzy operator; learning algorithm; convergence