

文章编号: 1002-0411(2001)06-486-06

几种模糊神经网络系统关系的对比研究

丛 爽

(中国科学技术大学自动化系 合肥 230027)

摘 要: 本文对几种不同结构形式的模糊神经网络系统, 从不同的模糊逻辑算式入手, 对所具有的功能、不同的表达式及其相互之间的关系式, 以及所表现出来的优缺点进行深入的分析与对比研究, 从而揭示模糊逻辑系统的实质内容, 为选择和应用模糊神经网络的设计方法提供一些作者的见解。

关键词: 模糊神经网络; Sugeno 模糊推理法; B 样条基函数; 径向基函数网络

中图分类号: TP13

文献标识码: A

COMPARATIVE RESEARCH ON RELATIONSHIPS BETWEEN SEVERAL FUZZY-NEURAL NETWORK SYSTEMS

CONG Shuang

(Dept. of Automation, University of Science and Technology of China, Hefei 230027)

Abstract: This paper analyses and studies several fuzzy-neural networks with different structures. In order to explore the essential contents of fuzzy logic system, this paper does some deeply comparison on every kinds of fuzzy logic formulations, functions, expressions and relationships of fuzzy-neural network systems. The study and analysis in the paper are useful for selecting and designing the fuzzy-neural networks.

Keywords: fuzzy-neural networks; sugeno fuzzy inference; B-spline basis function; radial basis function network

1 引言(Introduction)

模糊系统已经广泛地应用在模糊控制器、模式识别、模糊辨识和信号处理中。模糊逻辑方法的优点在于它的逻辑性和透明性上, 使人们很容易将先前已知的有关的系统知识, 结合到模糊规则基中。但是, 模糊系统必须包括人为选定的模糊隶属函数的模糊化过程和采用适当方法的模糊规则的推理过程, 每一过程对系统的执行性能都有影响, 而且模糊系统中模糊规则的前提和结论部分通常都是模糊子集, 解模糊也是一项复杂的工作, 模糊系统输入/输出关系式是高度非线性的, 要想得到一个满意的输入/输出关系式, 在众多需要调节的参数面前, 再有经验的专家也难以胜任。而另一方面, 人工神经网络的最大益处就在于它善于对网络参数的自适应学习, 并且具有并行处理及泛化能力。这很自然促使研究者进行将模糊逻辑和神经网络结合成一个系统的研究, 并称其为模糊神经网络或模糊神经系统。通过神经网络实现的模糊逻辑系统结构具有模糊逻辑推

理功能, 同时网络的权值也具有明确的模糊逻辑意义, 从而达到以神经网络及模糊逻辑各自的优点弥补对方不足的目的。

本文从模糊关系式入手, 分别对直接的模糊神经网络的实现、采用 sugeno 模糊推理法实现的模糊系统的特点、B 样条模糊神经网络以及径向基函数神经网络等系统之间的相互联系, 以及各自所表现出来的优缺点进行了性能分析与对比研究, 用此达到使人们对模糊神经网络有更加深入地认识, 更好地设计网络、利用网络系完成不同应用的目的。

2 模糊系统的关系式(Relationship of fuzzy system)

考虑一个一般形式的模糊系统, 其输入是 $x \in R^n$, 输出为 $y(x) \in R$ 。 x_i 和 y 的论域分别是 $x_i \in R$, ($i=1, 2, \dots, n$) 和 $y \in R$ 。 x_i 的语言值是 $A_i^{k_i}$ ($k_i=1, 2, \dots, m_i$ 和 $i=1, 2, \dots, n$), 模糊规则的数量为 $p = m_1 m_2 \dots m_n$ 。第 k 条规则为:

如果 x_1 是 $A_1^{k_1}(x_1)$, 并且 x_2 是 $A_2^{k_2}(x_2)$, 并且 ...

$$x_n \text{ 是 } A_n^{k_n}(x_n), \text{ 那么 } y_k \text{ 为 } y_k(x) \quad (1)$$

其中 $y_k(x) \in Y, k$ 为规则基的数目, 有 $k = 1, 2, \dots, p$, 每个 k 对应一个有序序列 $k_1, \dots, k_i, \dots, k_n$, 其中 $k_i = 1, 2, \dots, m_i$. 若采用乘法法进行与操作, 那么第 k 条模糊规则前提部分的真值 $\mu_k(x)$ 可以表示为:

$$\begin{aligned} \mu_k(x) &= A_1^{k_1}(x_1) \cdot A_2^{k_2}(x_2) \dots A_n^{k_n}(x_n) \\ &= \prod_{i=1}^n A_i^{k_i}(x_i) \end{aligned} \quad (2)$$

从前提部分到结论部分的推理可产生一个结论 C , 它是一个离散的、具有有限点数的模糊子集:

$$C = \{u_k/y_k, | k = 1, 2, \dots, p\}$$

其中 u_k/y_k 表示在点 $y_k \in Y$ 上的 C 的隶属度是 u_k .

采用质心法对模糊集 C 进行解模糊, 系统的实际输出为:

$$y(x) = \frac{\sum_{k=1}^p u_k(x) y_k(x)}{\sum_{k=1}^p u_k(x)} \quad (3)$$

将(2)式中的 $\mu_k(x)$ 的值代入上式中, 其中分母可重写为:

$$\begin{aligned} \sum_{k=1}^p u_k(x) &= \sum_{k_1=1}^{m_1} \sum_{k_2=1}^{m_2} \dots \sum_{k_n=1}^{m_n} A_1^{k_1}(x_1) A_2^{k_2}(x_2) \dots A_n^{k_n}(x_n) \\ &= \sum_{k_1=1}^{m_1} A_1^{k_1}(x_1) \sum_{k_2=1}^{m_2} A_2^{k_2}(x_2) \dots \sum_{k_n=1}^{m_n} A_n^{k_n}(x_n) \end{aligned} \quad (4)$$

所以(3)式变为:

$$y(x) = \frac{\sum_{k=1}^p \prod_{i=1}^n A_i^{k_i}(x_i) y_k(x)}{\sum_{k_1=1}^{m_1} A_1^{k_1}(x_1) \sum_{k_2=1}^{m_2} A_2^{k_2}(x_2) \dots \sum_{k_n=1}^{m_n} A_n^{k_n}(x_n)} \quad (5)$$

由此可见, 模糊系统的输入/输出的求解关系式是相当复杂和费时的. 另一种人们常用的模糊推理方法是所谓的最小推理, 即将关系式(2)中 $\mu_k(x)$ 的操作变为取小运算, 然后再通过(3)式进行解模糊, 并通过制表法来设计模糊逻辑表. 不过这种取小运算将有用信息丢失, 利用信息率较低. 而采用(2)式中的乘法操作考虑了所有信息的影响, 同时便于用神经网络来实现模糊系统, 因而得到广泛地应用.

3 用神经网络直接实现的模糊系统(Fuzzy System realized by Neural Network Directly)

为了能够用计算机对模糊逻辑进行实时操作,

一个很自然的想法是通过神经网络来实现模糊系统. 在采用神经网络实现模糊化、模糊推理和解模糊化过程中, 为了运算简单, 最直接的模糊神经网络系统中采用的模糊推理方式是用乘积-求和代替模糊系统中常用的最小-最大推理法, 且输出变量的隶属函数取单值型^[1].

假定输入为 x_1 和 x_2 , 模糊标记取高(H)、正(P)、零(Z)、负(N)和低(L)五个, 由此可构造一个具有模糊功能的神经网络如图 1 所示.

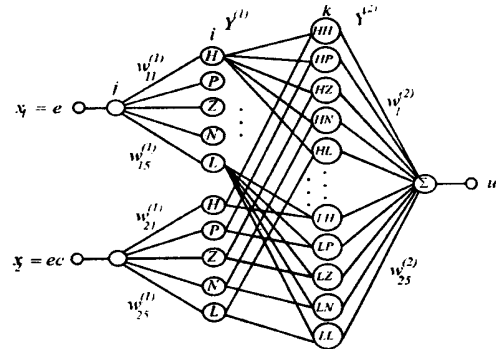


图 1 具有模糊逻辑功能的神经网络的结构

Fig. 1 The structure of neural network with fuzzy logic function

由图 1 可以看出它是一个具有输入层、中间层和输出层的网络结构, 而输入 x_1 和 x_2 各自被送入一个前向人工神经网络, 该网络的激活函数的形状即代表模糊隶属函数的形状, 可由设计者确定采用线性或指数型, 分别可对应模糊隶属函数的三角型和高斯型; 输入层神经元的个数代表模糊标记数, 也是由设计者选定. 通过调节输入层网络的权值和偏差可以达到调节激活函数的宽度及其中心位置的目的.

由输入层输出的 H、P、Z、N 和 L 分别代表模糊化后的结果——隶属度. 中间层执行的是将模糊化得到的隶属度两两相乘的功能. 由 x_1 得到的 H 值与 x_2 得到的 H 值相乘获得 HH 值, 由 x_1 得到的 H 值与 x_2 得到的 P 值相乘获得 HP 值, 以此类推, 可获得由得到的 L 值与 x_2 得到的 L 值相乘得到 LL 值. 所以在图 1 中所看到的 HH、HP、...、HL 等节点是简单的乘法单元, 功能是计算出相应输入的乘积. 最后乘法单元的输出通过 $\omega_1^{(2)}$ 到 $\omega_8^{(2)}$ 的权值送到一个求和节点中, 于是整个网络的输出值是:

$$u = H \cdot H \cdot \omega_1 + H \cdot P \cdot \omega_2 + \dots + L \cdot N \cdot \omega_8 + L \cdot L \cdot \omega_9 \quad (6)$$

这就实现了乘积-求和的模糊推理过程.

此网络的最大优点是用神经网络实现了模糊系

统,从而可以通过训练来求得其中的参数,不必依赖专家经验,人工事先确定,达到了自动确定系统参数的目标.系统的隶属函数的中心、宽度以及规则等全部参数同时得以自寻优.

该模糊神经网络的实现是通过对网络采用监督式训练来确定输入层的权值、偏差以及输出层的权值.由于该网络不是标准的全联接,加上模糊标记一般情况下比较多,所以网络的训练速度通常是相当慢的,对于某些复杂的输入/输出关系,网络的训练常常很难收敛到期望的极小值.另外,由于网络中代表模糊标记的神经元数必须由设计者事先确定,它的数目对网络的精度也是有影响的.

4 Sugeno 模糊推理法(Sugeno fuzzy inference)

迄今为止,人们常用的模糊推理过程被称为玛达尼模糊推理法,玛达尼推理中的结论部分是模糊集合,对模糊集合进行处理后,每个输出变量需要被解模糊.在实践中人们认识到采用单值型隶属函数可以简化计算,提高解模糊的有效性.Sugeno^[2](或 Takagi Sugeno-Kang)推理法对结论就是采用单值型隶属函数的.Sugeno 法的前提部分与玛达尼推理法有相同的结构,而结论部分不同于原来的模糊集合,它用一个前提部分变量的多项式表示,是前提变量的线性函数,并采用加权平均法求解模糊.

零阶 Sugeno 模糊系统的典型模糊规则有如下形式:

如果 x_1 为 A , 并且 x_2 为 B , 那么 $y = k$ (7)

此处 A 和 B 为前提中的模糊集合,而 k 为结论中一个可调参数.一阶 Sugeno 模糊系统具有的模糊规则形式为:

如果 x_1 为 A , 并且 x_2 为 B , 那么 $y = x_1 + qx_2 + r$ (8)

此处 A 和 B 为前提中的模糊集合,而 p 、 q 、和 r 可调参数.想象此一阶系统运作方式的简单方法就是把每一条规则认为是定义一个“移动单值”的位置,即取决于输入值的单值输出可以在输出空间中移动.

作为对比,这里需要指出的是:如果将模糊规则(1)中的 $y_k(x)$ 设置成一个由输入矢量表示的线性组合,即

$$y_k(x) = a_1^k x_1 + a_2^k x_2 + \dots + a_n^k x_n \quad (9)$$

那么由(3)式所示的则为采用 Sugeno 推理法时的系统输出.

当每条规则在系统输入空间中是线性独立时,Sugeno 法可以作为对一个非线性动力学系统进行不同操作条件的复合线性控制器.另外,一个 Sugeno 系统也适用于非线性系统的建模,它是通过插入复合线性模型来实现的.

由于输出的线性组合,使得由 Sugeno 推理法求得的模糊系统比用玛达尼法具有更好的效果.实际上仅从模糊系统角度来看,Sugeno 推理法同样存在如何确定有效的模糊规则的问题.但当把神经网络与其相结合起来后,Sugeno 推理法则显示出它的优越性.

下面给出 Sugeno 推理法与玛达尼推理法性能的对比结论:

	Sugeno 推理法	玛达尼推理法
优点	1) 有效的计算(直接由输入到输出的点到点的计算) 2) 与线性技术(如 PID 等)最优化技术及自适应技术能一起很好的工作(因其本身就是线性复合技术) 3) 保证输出空间的连续性 4) 更适合进行数学分析	1) 较直观 2) 被初学者广泛的应用 3) 较适合于人工输入规则的情况
缺点	1) 参数的确定仍然很麻烦	1) 运算复杂 2) 丢失信息较多

5 B 样条模糊神经网络(B-spline Fuzzy Neural Networks)

通过分析模糊系统的推理过程可知,系统主要有三个可进行选择的部分:隶属函数的形状、模糊规则和解模糊过程.用神经网络实现模糊系统的另一思想是对模糊系统中某一部分的参数实现自动调节.

对于隶属函数的调整,B 样条基函数不失为一种较好的选择,因为它具有期望的数学特性如局部紧支撑、数学计算的简单和单位分割性.将 B 样条基函数作为隶属函数,构造一个 B 样条模糊神经网络系统也是人们一直关注的.

B 样条是基本样条(Basic Spline)的简称.B 样条算法最重要的特性是由其函数的形状而能够产生光滑的输出.另外,B 样条函数是所有样条函数中具有最小局部支撑的样条函数,所以 B 样条基函数可以精确多项式分段插值的方式,对给定的输入/输出数据进行光滑的曲线拟合.现已证明 B 样条网络可以任意精度逼近一个连续的函数^[3].

给定一组单变量 x 的节点序列: $x_1 < x_2 < \dots < x_{N+m}$, 那么在区间 $[x_m, x_{N+1}]$ 之间就可以唯一确定 $N+m$ 个 m 阶线性不相关的 B 样条基函数, 一个连续函数 $y(x)$ 就可以用 $l = N+m$ 个 B 样条基函数的线性组合来近似的表示:

$$y(x) \approx \sum_{i=1}^l \omega_i B_{i,m}(x) \quad x \in [x_m, x_{N+1}] \quad (10)$$

其中, ω_i 为第 i 个 k 阶 B 样条基函数 $B_{i,k}(x)$ 的权值. 当具体计算其函数值时, 这些权值必须被估值.

B 样条基函数满足以下迭代关系式:

$$\left\{ \begin{aligned} B_{i,1}(x) &= 1 \quad x \in (x_i, x_{i+1}] \\ B_{i,k}(x) &= \frac{x - x_i}{x_{i+k-1} - x_i} B_{i,k-1}(x) + \\ &\quad \frac{x_{i+k} - x}{x_{i+k} - x_{i+1}} B_{i+1,k-1}(x), \\ &k = 1, 2, \dots, m \end{aligned} \right. \quad (11)$$

B 样条函数是非负的, 且局部支撑的, 即:

$$B_{i,k}(x) \begin{cases} > 0, & x \in [x_i, x_{i+m}] \\ = 0, & \text{其它} \end{cases} \quad (12)$$

且它们形成一个单份分割, 即

$$\sum_{i=1}^N B_{i,m}(x) \equiv 1, \quad x \in [x_m, x_{N+1}] \quad (13)$$

对于多变量 B 样条, 可以假定 n 维输入矢量 $x = [x_1 \ x_2 \ \dots \ x_n]^T$, 在每个输入轴上定义单变量 B 样条基函数 $B_{i,m_i}^{k_i}(x_i)$, $k_i = 1, 2, \dots, m_i$, $i = 1, 2, \dots, n$,

那么, 第 k 个多变量 B 样条基函数 $M_k(x)$ 由 n 个单变量基函数 $B_{i,m_i}^{k_i}(x_i)$ 乘积 (张积) 组成:

$$M_k(x) = B_{1,m_1}^{k_1}(x_1) B_{2,m_2}^{k_2}(x_2) \dots B_{n,m_n}^{k_n}(x_n) = \prod_{i=1}^n B_{i,m_i}^{k_i}(x_i) \quad (14)$$

$M_k(x)$ 的总数为 $p = m_1 \cdot m_2 \dots m_n$, 所以 $k = 1, 2, \dots, p$.

张积 B 样条基函数仍然保留(11)和(12)式的特性.

那么一个多变量函数 $y(x)$ 可以用多变量 B 样条基函数的线性组合来逼近:

$$y(x) \approx \sum_{k=1}^p M_k(x) \omega_k = \sum_{k=1}^p \prod_{i=1}^n B_{i,m_i}^{k_i}(x_i) \omega_k \quad (15)$$

仔细观察可以发现, B 样条基函数的阶数 m 与不同形状的模糊隶属函数相对应, 如当 $m = 2$ 时, 用 B 样条作的曲线正好代表三角形隶属函数, 当 $m = 3$ 时, 所作曲线与二次型隶属函数相似. 另一方面, 所取的顶点数 N 则对应于模糊标记数. 节点数则在一定范围内决定了个隶属函数的宽度^[4]. 不过, 高于一阶的 B 样条函数并不一定是一个正则(normal)模糊集, 即 B 样条函数的最大值达不到 1, 而通常模糊系统理论要求隶属函数为正则的. 为了满足这一要求, 可对 B 样条基函数乘上一个正则化因子 λ , 以其最大值为 1, 并取隶属函数为:

$$\left. \begin{aligned} \lambda &= 1 / \sup_{x \in x_i} B_{i,m_i}^{k_i}(x_i) \\ A_{i,i}^{k_i}(x_i) &= \lambda B_{i,m_i}^{k_i}(x_i) \quad i = 1, 2, \dots, n \end{aligned} \right| \quad (16)$$

将(16)式代入模糊系统输入/输出关系(5)式中:

$$\begin{aligned} y(x) &= \frac{\sum_{k=1}^p \prod_{i=1}^n \lambda B_{i,m_i}^{k_i}(x_i) y_k(x)}{\sum_{k_1=1}^{m_1} \lambda B_{1,m_1}^{k_1}(x_1) \sum_{k_2=1}^{m_2} \lambda B_{2,m_2}^{k_2}(x_2) \dots \sum_{k_n=1}^{m_n} \lambda B_{n,m_n}^{k_n}(x_n)} \\ &= \frac{\lambda^n \sum_{k=1}^p \prod_{i=1}^n B_{i,m_i}^{k_i}(x_i) y_k(x)}{\lambda^n \sum_{k_1=1}^{m_1} B_{1,m_1}^{k_1}(x_1) \sum_{k_2=1}^{m_2} B_{2,m_2}^{k_2}(x_2) \dots \sum_{k_n=1}^{m_n} B_{n,m_n}^{k_n}(x_n)} \\ &= \sum_{k=1}^p \prod_{i=1}^n B_{i,m_i}^{k_i}(x_i) y_k(x) \end{aligned}$$

与(15)式完全相同. 由(17)式可以看出: 采用 B 样条函数作为模糊系统的隶属函数所得到的输入/输出关系式与正则化因子无关, 与用多变量 B 样条逼近多变量函数 $y(x)$ 的表达式完全相同. 换句话说, 用 B 样条逼近函数的表达式所代表的就是一个模糊系统. 两者研究的出发点虽然不同, 但达到相同的结果, 解决相同的问题.

由(17)式还可看出, 我们可以通过设计求解一个 B 样条网络函数来建立一个模糊系统. 因为两者所表达的实质内容是完全等价的. 设计者只要事先确定了网络的阶数 m (即隶属函数的形状), 拟合点的顶点数 (即模糊标记数), 被拟合的点至少必须有 $N+m$ 个. 网络函数中唯一需要优化的是权值系数 ω_k , 它代表模糊系统中推理规则的结论部分, 它的选

取可以通过不同的优化算法来求解, 以达到优化模糊规则的目的. 由于 B 样条网络采用多项式计算, 使其比单纯的用神经网络的指数函数作为隶属函数的模糊系统的求解节省时间.

由以上分析可得出如下结论:

采用多变量 B 样条网络建立的模糊系统, 相当于采用 B 样条基函数作为隶属函数, 采用优化算法对模糊规则进行自动寻优的模糊系统. 优点是: 算法相对省时, 但缺点是: 1) 隶属函数的模糊标记数事先必须确定; 2) 各隶属函数之间是等距离的, 而不是自动寻优的.

6 径向基函数神经网络(Radial Basis Function Neural Networks)

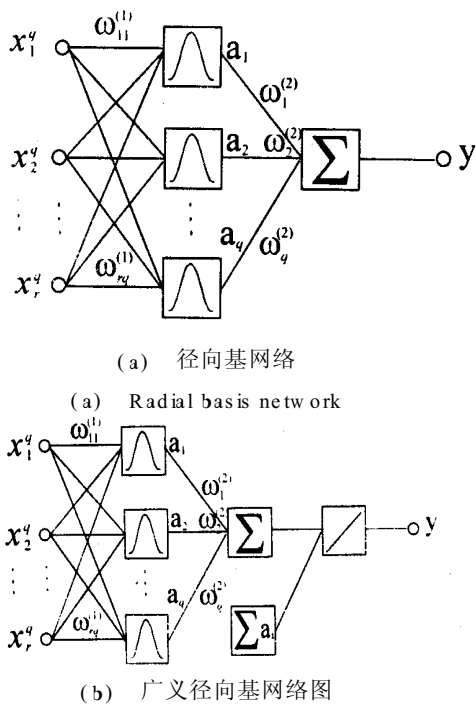
对于直接采用前向网络实现模糊系统所存在的训练时间长, 存在局部极小值问题, 径向基网络(Radial Basis Function Networks, 简称 RBF 网络) 在很大程度上解决了上述问题. RBF 网络不存在局部极小值问题, 它不仅具有全局逼近性能, 而且具有最佳逼近性能, 同时其训练方法快速易行, 为解决模糊神经网络的问题提供了新的思路.

接的前向网络. 隐含层中采用径向基函数——一种高斯型指数函数, 而输出层采用线性激活函数, 由此会使人感觉似乎 RBF 网络就是一个具有径向基函数的 BP 网络. 实际上 RBF 网络是与 BP 网络完全不同的网络, 其原因是: 1) 它不是采用 BP 算法, 即误差的反向传播法来训练网络权值的. 2) 其训练的算法不是梯度下降法. 虽然是两层网络, 径向基网络的权值训练是一层一层进行的. 在对径向基中的权值进行训练时, 采用的是无教师训练. 网络训练的目的是使 $\omega_{ij}^{(1)} = x_j^q$, 其中 j 为输入神经元节点数; q 为输入矢量的数组数. 由于径向基函数在将其输入放置在函数原点时输出为 1. 而对其它不同的输入值的响应均小于 1. 所以权值设计的目的是将每一组输入值 x_j^q 作为一个径向基函数的原点, 而权值 ω 代表径向基函数中心的位置, 通过令 $\omega_{ij}^{(1)} = x_j^q$ 使每一个径向基函数只对一组 x_j^q 响应, 从而迅速识别出 x_j^q 的大小. 然后进行输出层的权值设计. 由于输出层是线性函数, 网络输出是径向基输出的线性组合, 从而很容易的达到了从非线性输入空间向输出空间映射的目的. 理论上已经证明, 径向基网络可以以任意精度逼近任意连续函数.

值得指出的是, 由于径向基网络的权值算法是单层进行, 并且采用每一个 $\omega_{ij}^{(1)} = x_j^q$, 所以虽然从网络结构图中看上去网络是全联接, 实际上工作时网络是局部工作的, 即对每一组输入, 网络只有一个神经元被激活. 其它神经元被激活的程度可以忽略. 所以, 径向基网络是一个局部逼近的神经网络. 正是因为如此, 网络的训练速度要比 BP 网络快上 1~2 各数量级. 这是径向基网络的最大优点. 不过径向基网络隐含层的神经元往往比较多. 从上面的分析来看, 它的数量一般情况下应与被训练的输入数组 Q 相等. 当训练数组很多时, 采用径向基网络可能不容易被人接受.

对此缺点采用的改进算法是: 通过在满足误差目标的前提下, 尽量减少径向基中的神经元数. 具体的做法是在训练时从一个节点开始训练, 通过检查误差目标使网络自动增加径向基中神经元的节点, 每次循环后用使网络产生最大误差所对应的输入矢量产生一个新的径向基的神经元, 然后检查新网络的误差, 重复此过程直到达到误差目标为止. 每增加一个径向基的神经元, 相当于增加了一条模糊规则, 从而可以达到在线调整模糊规则的目的.

从网络结构上看, 每个径向基神经元的输出都是在每组输入值作用下的结果, 与图 1 所表示的



(b) General Radial Basis network

图 2 径向基网络结构图

Fig. 2 The structure of radial basis network

RBF 网络的结构图如图 2 所示. 从网络结构上看, RBF 网络含有一个隐含层, 并且具有标准全连

具有模糊逻辑推理的神经网络结构相比, RBF 网络中的径向基网络就起到了隶属函数的作用, 径向基层就完成了相当于图 1 中的输入层加上中间层所完成的模糊推理功能. 由于 RBF 网络中的径向基层是全连接, 所以通过一层网络, 对所有的输入变量都产生了作用, 即让每一个输入变量在每一个径向基上都有一个输出. 径向基层的输出即为所有输入变量共同作用的综合效果, 其值为 0-1 之间的数, 等价于模糊系统中从前提到结论所获得的所有模糊规则数. 由此可见, RBF 网络用一个隐含层起到了普通模糊神经网络两层的作用. 由于层数的减少, 简化了结构, 而且因其训练是两层分别进行, 必然能够提高训练网络的速度. 另外, 径向基网络输出层的激活函数是线性函数, 所以其网络输出应为 $Y = A \cdot W = \sum a_i \omega_i^{(2)}$, 这与模糊神经网络的表达式(6)完全相同. 这说明, 一个径向基网络的表达式所代表的就是一个模糊神经网络系统.

另一方面, 与 B 样条网络相比较可以发现, 径向基网络与 B 样条网络的表达内容完全相同, 所不同仅在于前者用径向基函数表示隶属函数, 而后的隶属函数是 B 样条基函数.

为了得到更加平滑的解模糊数值, 可在普通的径向基网络输出后加上一个求加权平均值的过程, 如图 2(b) 所示, 它比图 2(a) 中多了一个除法运算符号“/”. 这时的网络输出为

$$y = \frac{\sum a_i(x) \omega_i^{(2)}}{\sum a_i(x)} \quad (18)$$

此网络可称为广义径向基网络(或泛化回归网络).

因为 a_i 的数值为 0~1 的指数型函数输出值, 仔细观察(18)式, 可以发现(18)式的形式与(3)式采用质心法对模糊集合进行解模糊的模糊系统的输出方式完全相同, 只是在广义径向基网络中采用的是径向基函数, 而模糊系统中是某一形状的隶属函数. 正是由于这种径向基形状函数的作用, 使广义径向基网络和模糊系统都具有相同的对局部输入上的激励在一个小的接受域内产生中心加权响应的机理. 两者虽然是基于不同的原理发展而来的, 但是都达到了同一目的, 只是广义径向基网络具有较快的收敛速度, 而模糊系统能够反映出更多数据的物理特性.

完全使广义径向基网络与模糊系统等价, 还必须满足下列条件^[5]:

1) RBF 网络和模糊系统必须采用相同的方法(即或加权平均法或加权求和法)产生其输出.

2) RBF 网络中的接收域的神经元个数等于模糊系统中的如果-那么的模糊规则数.

3) RBF 网络中的每一个径向基函数等于模糊系统中组成某个模糊规则的多维隶属函数的前提部分. 达到这一目的方式之一是在模糊规则中采用与 RBF 网络具有相同的偏差的高斯型隶属函数, 并应用点积来产生规则强度, 从而使这些高斯型隶属函数的乘积变成一个多维高斯函数——RBF 网络中的一个径向基函数.

4) RBF 网络与模糊系统应当对应有相同的响应函数, 即它们应当具有相同的常数项(对于 ω 取常数时, 对应零阶 Sugeno 模糊推理系统, 当取 $\omega = c_0 + c_1 x_1 + c_2 x_2 + \dots + c_n x_n$ 时, 对应于一阶 Sugeno 模糊推理系统).

综上所述, 对于模糊系统的实际应用, 采用广义径向基网络不仅在隶属函数中心位置、宽度上, 而且对模糊规则同时给予优化, 另外具有相当快的收敛速度. 其不足是在一般情况下需要较多的隐含层的神经元数. 不过从模糊推理过程可以得知, 既然模糊标记数和规则数体现出模糊系统的精度, 所以在需要较高精度时, 神经元数目较多是理所当然的, 更何况 RBF 网络对神经元数目还可以自动寻优.

7 总结(Conclusion)

神经网络具有多种结构和学习算法, 模糊逻辑推理也具有多种形式, 本文从多方面研究了在用神经网络描述模糊控制的过程中, 不同模糊神经网络之间的等价特性, 以达到根据模糊推理规则来构造网络结构, 同时利用神经网络的学习能力进行复杂的模糊推理, 提高运算速度, 达到对权值自动寻优的目的. 通过对不同结构及方式的模糊神经网络系统关系的分析与对比, 更加深入地揭示了各种网络的优缺点, 有利于更好地选择和设计网络.

参 考 文 献 (References)

- 1 从 爽. 机电系统中模糊与模糊神经网络控制策略的研究. 中国电机工程学报, 1999, 19(7): 30~32, 37
- 2 M Sugeno. Industrial applications of fuzzy control, Elsevier Science Pub Co., 1985