

非线性系统建模的复合型模糊神经网络研究

何 谦 李宏光

(北京化工大学信息科学与技术学院, 北京 100029)

摘 要: 针对非线性系统建模问题, 提出了一类由函数逼近和规则推理网络构成的复合型模糊神经网络, 其规则网络基于过程先验知识用于对操作区间的划分, 而函数网络采用改进型模糊神经网络结构完成非线性函数逼近。采用一类非线性函数模型进行了仿真研究, 结果表明, 复合型模糊神经网络较之普通模糊神经网络在建模收敛速度和预测精度等方面都有较大的改善。

关键词: 模糊神经网络; 过程知识; 非线性系统; 建模

中图分类号: TP 183; TP 273.4

引 言

神经网络以其强大的函数映射能力, 已经广泛用于工业过程非线性系统建模领域, 它能够通过输入输出数据对过程进行有效的学习。但是, 一般的神经网络只能建立起过程的“黑箱”模型, 往往难于理解并导致模型缺乏鲁棒性。

改进神经网络建模技术的一个有效途径是在其学习过程中有效地利用过程的先验知识。模糊逻辑系统在此方面具有优势, 它较适合表达模糊或定性的知识, 其缺点是缺乏自学习和自适应能力。而模糊神经网络技术则是将模糊逻辑和神经网络结合起来, 取长补短, 因而在解决基于知识的非线性系统建模领域具有良好的研究前景^[1,2]。

目前, 模糊神经网络建模多是基于其能以任意精度逼近紧致集上的实连续函数的性质^[3,4], 与普通神经网络相比, 模糊神经网络在学习速度和精度方面都具有优势。但是, 这类模糊神经网络并未从实际意义上引入对象的先验知识, 如操作和控制经验, 以及过程的领域知识等, 这使得模糊神经网络模型仍是一个“黑箱”模型。

针对非线性系统建模问题, 本文提出了一种结合过程的规则推理和函数逼近功能的复合型模糊神经网络结构。其中, 规则网络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, 实现操作区域的划分。而函数网络则利用

一类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。此类复合型模糊神经网络由于强调了对过程先验知识的利用和学习, 因此能够有效地提高对复杂非线性系统在大区段范围的逼近和泛化能力。作者采用了一类非线性数学模型对此复合型模糊神经网络的建模性能进行了仿真研究, 结果表明了它解决问题的有效性。此研究为模糊神经网络的非线性系统建模提供了思路。

1 网络结构

复合型模糊神经网络由两个模糊神经网络——函数网络和规则网络组合而成, 该网络的拓扑结构如图 1 所示。

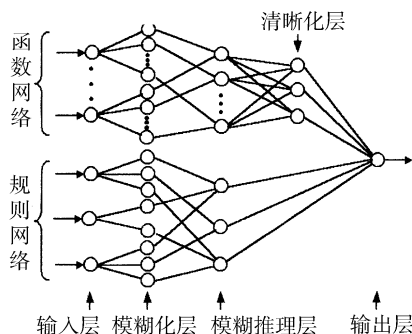


图 1 复合型模糊神经网络结构示意图

Fig. 1 Schematic of compound fuzzy neural network

函数网络采用一类改进型模糊神经网络结构^[3], 用于实现对象建模的函数运算功能, 网络输入是影响系统输出的主导变量。此网络结构的特点是各输入变量的模糊集合分割数 (即模糊化层的节

点数)等于模糊推理层的规则数,且模糊推理层各节点的输入与模糊化层各节点的输出一一对应。网络采用了较一般模糊神经网络计算更为简单的清晰化输出 $u = \sum_{i=1}^m \mu_i$, 其中, μ_i 为模糊推理层和清晰化层之间的连接权值, μ_i 为模糊推理层各单元节点的输出值, m 为规则数。模糊神经网络及对应的模糊系统可以作为一个全局逼近器, 仿真研究已经证明了它具有收敛快、精度高等良好性能, 可以作为普通型函数逼近器^[3]。

规则网络采用一般模糊神经网络结构, 用于实现对象操作区域的划分, 网络的输入是操作区域划分的规则条件变量。规则网络的输入层各单元的模糊分割数和模糊推理层各单元的连接情况取决于事先获知的区域推理规则, 按不同的建模对象分别确定。

复合型模糊神经网络具有将函数网络和规则网络有机结合的特点, 使得整个网络具有能够充分利用对象先验知识进行非线性系统建模的能力, 从而可以提高网络的收敛速度, 增强网络的建模精度。

2 网络功能

复合型模糊神经网络分为五层结构, 分别为输入层、模糊化层、模糊推理层、清晰化层和输出层。

第一层为输入层, 函数网络的输入节点直接与系统输入向量的各个分量相连, 而规则网络的输入节点则对应于区域推理规则的前提变量。

第二层为模糊化层, 它对输入向量的每个分量进行模糊集合分类处理。由于函数网络采用了改进型的模糊神经网络结构, 因此它的模糊集合分类数等于其第三层的规则数, 而对应于规则网络, 区域推理规则的前提变量的模糊集合分类处理要根据先验规则确定。本层采用高斯函数作为模糊等级的隶属

度值, 即 $\mu_i = e^{-\frac{(x_i - c_i)^2}{2\sigma_i^2}}$, 其中, x_i 为输入量, c_i 为高斯函数的中心点, σ_i 为高斯函数的半径。

第三层为模糊推理层, 每个节点代表一个推理结论。函数网络的推理结论数(即规则数)等于各输入分量的模糊集合分类数, 数量可根据实际情况调整。规则网络的规则及规则数须根据系统推理规则确定。本层采用积运算推理方法。

第四层为清晰化层, 其作用是实现函数网络的模糊结论的清晰化计算。这里采用的改进型算法清

晰化输出为: $u_k = \sum_{i=1}^r \mu_{ki}$, $k = 1, 2, \dots, m$, 其中, μ_i 为推理层 r 个节点的输出, μ_{ki} 为清晰化层与推理层之间的连接权, 而清晰化输出的个数应等于规则网络的推理结论数, 一般对应于操作区域划分数。

第五层为输出层, 采用重心法计算函数网络和规则网络的综合输出值, 对于单输出网络, 其输出为

$y = \frac{\sum_{k=1}^m \mu_k}{\sum_{k=1}^m \mu_k}$, 其中 μ_k 为输出层与清晰化层之间的连接权, μ_k 为规则网络的模糊推理层节点输出, $k = 1, 2, \dots, m$ 。

3 学习算法

由于网络的各个输入分量的模糊分割数已预先确定, 因此要学习的参数主要是函数网络的模糊推理层与清晰化层之间的连接权, 以及整个网络的第二层各单元节点的隶属度函数的中心点和半径。由于实际上是一种多层前馈网络, 因此易于采用误差反传(BP)算法^[4], 其中, 向量 x, μ, u, y 分别为网络第一层到第五层单元节点的输出, d 为期望输出, c, σ 分别为高斯函数的中心点和半径, w 为连接权值, s 表示网络号, 即 $s = 1$ 为函数网络, $s = 2$ 为规则网络, p 为样本号, n 为输入层节点数, r 为规则数(模糊推理层的节点数)。

⑧ 前向计算

第一层 $x_{pi}^{(s)}$ = 各输入向量 ($i = 1, 2, \dots, n_s$; $s = 1, 2$)

第二层 $\mu_{pij}^{(s)} = \exp \left[- \left(\frac{x_{pi}^{(s)} - c_{ij}^{(s)}}{\sigma_{ij}^{(s)}} \right)^2 \right]$ ($i = 1, 2, \dots, n; j = 1, 2, \dots, r_s$)

第三层

(a) 函数网络 $\mu_{pi}^{(1)} = \sum_{j=1}^{n_1} \mu_{pki}^{(1)}$ ($k = 1, 2, \dots, n_1; i = 1, 2, \dots, r_1$)

(b) 规则网络 $\mu_{p1}^{(2)} = \mu_{p11}^{(2)} \cdot \mu_{p21}^{(2)} \cdot \mu_{p31}^{(2)}$
 $\mu_{p2}^{(2)} = \mu_{p12}^{(2)} \cdot \mu_{p22}^{(2)} \cdot \mu_{p32}^{(2)}$
 $\mu_{p3}^{(2)} = \mu_{p13}^{(2)} \cdot \mu_{p23}^{(2)} \cdot \mu_{p33}^{(2)}$

第四层(仅函数网络) $u_{pi} = \sum_{k=1}^{r_1} \mu_{pki}^{(1)}$ ($i = 1, 2, \dots, m; k = 1, 2, \dots, r_1$)

第五层 $y_p = \frac{\sum_{i=1}^m u_{pi} \cdot \mu_{pi}^{(2)}}{\sum_{i=1}^m \mu_{pi}^{(2)}} \quad (i = 1, 2, \dots, m)$

® 误差代价函数 $E_p = \frac{1}{2} (d_p - y_p)^2$

® 误差反传计算

对于权值 $_{ij} : \frac{\partial E_p}{\partial _{ij}} = \frac{\partial E_p}{\partial y_p} \cdot \frac{\partial y_p}{\partial u_{pj}} \cdot \frac{\partial u_{pj}}{\partial _{ij}} = - (d_p - y_p) \cdot \frac{(2)}{m} \cdot \frac{(1)}{pi} \cdot \frac{(2)}{pk}$

对于半径 $_{ij}^{(s)}$ 和中心点 $c_{ij}^{(s)}$:

(a)函数网络 $\frac{\partial E_p}{\partial _{ij}^{(1)}} = \frac{\partial E_p}{\partial y_p} \cdot \frac{\partial y_p}{\partial u_{pk}} \cdot \frac{\partial u_{pk}}{\partial _{pj}^{(1)}} \cdot \frac{\partial _{pj}^{(1)}}{\partial \mu_{pij}^{(1)}}$
 $\frac{\partial \mu_{pij}^{(1)}}{\partial _{ij}^{(1)}} = - (d_p - y_p) \cdot \sum_{k=1}^m \left[\frac{(2)}{3} \cdot \frac{(2)}{pk} \cdot \frac{(1)}{jk} \right] \cdot \sum_{i=1}^{n_1} \mu_{pij}^{(1)}$
 $2\exp \left[- \left(\frac{x_{pi}^{(1)} - c_{ij}^{(1)}}{(1)} \right)^2 \right] \cdot \frac{(x_{pi}^{(1)} - c_{ij}^{(c)})^2}{((1))^3}$

(b)规则网络 视各节点的连接情况定,以 $_{11}^{(2)}$ 为例,有

$$\frac{\partial E_p}{\partial _{11}^{(2)}} = \frac{\partial E_p}{\partial y_p} \cdot \frac{\partial y_p}{\partial u_{pk}^{(2)}} \cdot \frac{\partial u_{pk}^{(2)}}{\partial \mu_{p11}^{(2)}} \cdot \frac{\partial \mu_{p11}^{(2)}}{\partial _{11}^{(2)}} = - (d_p - y_p) \cdot \frac{u_{p1}^{(2)} - (u_{pk}^{(2)} \cdot \frac{(2)}{pk})}{\left(\frac{(2)}{pk} \right)^2} \cdot \left(\frac{(2)}{p21} \cdot \mu_{p31}^{(2)} \right) \cdot 2\exp \left[- \left(\frac{x_{pi}^{(2)} - c_{ij}^{(2)}}{(2)} \right)^2 \right] \cdot \frac{(x_{pi}^{(2)} - c_{ij}^{(2)})^2}{((2))^3}$$

4 仿真实例

4.1 对象模型

选用 4 输入 1 输出的非线性系统进行仿真：

$$z = 1 + \frac{2(x_1^{(1)}) + (x_2^{(1)})^2}{1 + (x_3^{(1)})^{0.5} + 0.8(x_4^{(1)})^{1.5}} \tag{1}$$

其中, $x_1^{(1)} \sim x_4^{(1)}$ 是系统输入变量,且为[1,5]内的整数,z 为系统输出。

为了模拟实际系统在不同操作区域内的不同非线性特性,模型(1)设置了 变量,并设 $x_1^{(1)} = 1$ 时, $= 0.5$; $x_1^{(1)} \in [2,4]$ 时, $= 1$; $x_1^{(1)} = 5$ 时, $= 2$ 。

对于许多实际工业系统,往往可以找到一些与系统输出量相关的变量,如工业尿素生产 IDR 汽提塔液位系统,汽提塔气相出口温度等受液位变化的影响,当液位处于偏高、正常或偏低区域时,汽提塔气相出口温度反映出明显的不同。此类现象在许多复杂工业过程中往往都是存在的,为此在仿真研究时进行此类假设具有实际意义。

对于模型(1),假设存在系统输出 z 的观测变量 $x_1^{(2)} \sim x_3^{(2)}$, z 与输出观测变量有如下关系：

$$x_1^{(2)} = z^2, x_2^{(2)} = z^{0.5}, x_3^{(2)} = 1 + \cos(z) \tag{2}$$

其中, 为圆周率。

在自变量定义域内随机产生 100 组数据,输入模型发生仿真数据,将前 50 组数据作为训练样本,后 50 组数据作为测试样本,仿真数据如表 1 所示。

表 1 模型仿真数据

Table 1 Data of model simulation

样本号	$x_1^{(1)}$	$x_2^{(1)}$	$x_3^{(1)}$	$x_4^{(1)}$	$x_1^{(2)}$	$x_2^{(2)}$	$x_3^{(2)}$	z
1	4	2	3	4	0.003 4	0.241 8	1.983 1	2.314 0
2	3	3	3	1	0.049 2	0.471 0	1.766 7	5.246 8
3	3	1	4	3	0.001 5	0.199 4	1.992 2	1.978 0
4	2	3	1	4	0.005 1	0.267 4	1.974 8	2.547 6
5	2	2	3	5	0.000 5	0.153 1	1.997 2	1.685 1
6	4	4	3	2	0.063 9	0.502 9	1.700 4	5.805 0
7	1	3	5	5	0.001 2	0.188 6	1.993 7	1.903 0
8	4	1	3	5	0.000 7	0.168 0	1.996 0	1.770 7
9	5	5	3	4	0.196 0	0.665 4	1.178 6	9.212 8
10	4	4	3	5	0.009 9	0.315 9	1.951 2	3.055 4
.....

通过观测量 $x_1^{(2)} \sim x_3^{(2)}$ 与输出量 z 的数据统计分析,可以得到以下关于系统的区域规则：

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

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

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

采用如图 1 所示的复合型模糊神经网络结构,对于具体问题,网络选用 7 输入 1 输出结构,其中函数网络为 4 输入 1 输出形式,实现(1)式的函数模型计算。规则网络采用 3 输入 1 输出形式,实现按(2)式获得的规则推理。

4.2 网络结构参数确定

复合型模糊神经网络的结构具有一定的灵活性并且较容易确定。对于实际问题,规则网络往往是根据事先所获得的经验知识而确定的,因此其结构确定应根据具体问题而定。对于函数网络,由于它实际上是普通的改进型模糊神经网络,一般其规则数应等于模糊分割数,具体数量在事先不了解对象情况的条件下可随意选取。实际上,它也是复合型模糊神经网络结构中唯一需要进一步研究确定的结构参数。

目前确定模型结构的方法很多,如 Akaike 信息准则、预报误差准则、高阶相关检验等^[5~7],笔者根据神经网络 AIC 准则函数思想^[8],提出对复合型模糊神经网络结构确定的 AIC 准则 $A(r) = N \cdot \lg \left[\frac{2E_S(t)}{N} \right] + 4r$, 其中, N 为训练样本总数, $E_S(t)$ 为第 t 次训练的误差平方和, r 为规则数。

分别对规则数 $r=2, 3, 4, 5$ 的情况进行了仿真实验,得到 AIC 参数曲线,如图 2 所示。

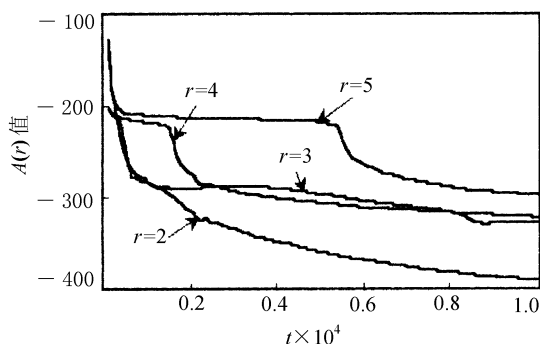


图 2 AIC 曲线

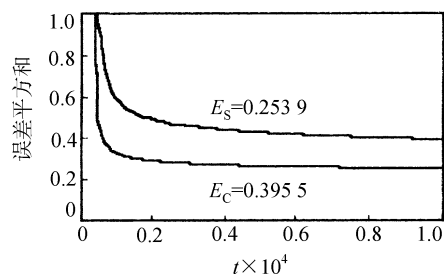
Fig. 2 AIC curves

根据 AIC 原理,由仿真结果可见,在规则数 r 取 2, 3 时效果较好, r 取 4 和 5 时较差。此观点与笔者进行的具体模型仿真结果相符。可见,采用

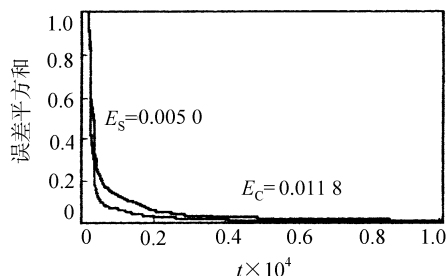
AIC 准则方法对于确定复合型模糊神经网络的结构是有效的。

4.3 网络学习及数据拟合

取函数网络规则数 $r=2$, 训练次数 $t=10\ 000$, 采用复合型模糊神经网络对研究对象进行仿真。并在相同条件下,对普通型模糊神经网络(即单纯函数神经网络结构)同样进行仿真实验,两类网络的学习误差和数据检验拟合曲线见图 3 和图 4,其中 E_S , E_C 分别为训练和测试误差。



(a) 普通型模糊神经网络



(b) 复合型模糊神经网络

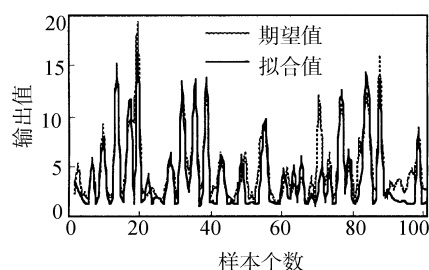
图 3 学习误差曲线

Fig. 3 Learning error curves

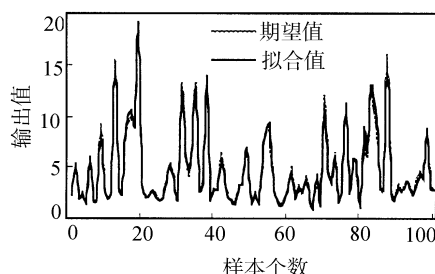
研究表明,虽然复合型模糊神经网络较普通型模糊神经网络在结构上有所增大,但学习收敛速度和精度方面却有较大的提高,并且在系统输出的跟踪能力方面也有改进,就本次仿真实例而言,训练精度和测试精度分别提高了约 50 和 33 倍。

5 结束语

针对多变量非线性系统建模问题,提出了一种复合型模糊神经网络,研究了网络结构,给出了网络学习算法,并采用具体仿真对象对其进行了深入的研究。研究表明,此类复合型模糊神经网络由于兼备函数建模和规则推理的结构特点,因此较普通型模糊神经网络在建模能力上有较大的改善。本文的工作是为解决实际工业过程建模及预测问题而展开的,并且已经在具体问题处理中取得满意的效果。



(a) 普通型模糊神经网络



(b) 复合型模糊神经网络

图 4 拟合曲线

Fig. 4 Fitted curves

目前,作者关于模糊神经网络建模的研究还在进行,如引入内部反馈形成递归网络,以处理动态特性较

强的非线性系统等。

参 考 文 献

- [1] Zhang J, Morris A J. Recurrent neuro-fuzzy networks for nonlinear process modeling. IEEE Trans On Neural Networks, 1999, 10(2): 313 ~ 325
- [2] 喻宗泉. 模糊神经系统的典型结构. 自动化与仪表, 1999, 14(1): 6 ~ 10
- [3] 王士同. 神经模糊系统及其应用. 北京: 北京航空航天大学出版社, 1998
- [4] 孙增圻, 张再兴, 邓志东. 智能控制理论与技术. 北京: 清华大学出版社, 1997
- [5] Akaike H. A new look at the statistical model identification. IEEE Trans On Automatic Control, 1974, 19(6): 716 ~ 723
- [6] 方崇智, 萧德云. 过程辨识. 北京: 清华大学出版社, 1988
- [7] Billings S A, Voon W S F. Correlation based model validity tests for non-linear models. Int J Control, 1986, 44(1): 235 ~ 244
- [8] Billings S A, Zhang G L. Radial basis function network configuration using genetic algorithms. Neural Networks, 1995, 8(6): 877 ~ 890

Compound fuzzy neural net work proposed for modeling of nonlinear systems

HE Qian LI Hong-guang

(College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China)

Abstract: A type of compound fuzzy neural network that has a rule network and a function network is proposed in this paper for building nonlinear system models. Based on process knowledge, the process operation is partitioned into a number of local operating regions by the rule network, and the local behavior of the process is approximated by the function network with improved fuzzy neural network structure. Simulation has been made with a nonlinear model, and the results of which show that the compound fuzzy neural network is better than the ordinary fuzzy neural network in convergence rate and predication precision.

Key words: fuzzy neural network; process knowledge; nonlinear system; modeling