

自适应模糊辨识及其在大系统中的应用*

罗旭光

(西安交通大学系统工程研究所·西安, 710049)

刘 丁

万百五

(西安理工大学自动化学院·西安, 710048) (西安交通大学系统工程研究所·西安, 710049)

摘要: 本文基于 T-S 模糊模型构造了一种新的自适应模糊神经网络, 给出了网络的连接结构和学习算法, 它能自动学习和修正前件参数及模糊规则. 将其用于大系统随机稳态递阶优化的控制建模中, 仿真结果表明, 该方法具有收敛速度快、辨识精度高、泛化能力强等特点, 可当作复杂大系统建模的一种有效手段.

关键词: 模糊逻辑; 神经网络; 动态过程; 稳态模型; 系统辨识; 大系统

1 引 言

稳态大系统递阶控制是解决复杂工业生产过程优化控制的重要方法之一. 在实际应用中, 获得大系统的稳态数学模型是非常困难的, 即使经过多重假设和繁琐推导, 所得到的模型也是相当复杂的, 并且难以保证其精确性, 使得人们不得不采用一些复杂算法(如考虑实际过程的全局反馈等), 影响了控制效果和算法的效率.

神经网络系统作为一般函数估计器已被广泛运用于多种领域, 可是它存在收敛速度慢、知识表达机理不清楚等问题. 而模糊逻辑擅长于定性分析和推理, 具有较强的自然语言处理能力, 更接近于人的思维方式. 模糊逻辑与神经网络从不同的方面近似表现了人脑的机能. Wang 利用 Stone-Weierstrass 定理证明了具有积推理、中心反模糊化、高斯型隶属函数的模糊系统能以任意精度逼近任一闭子集上的实连续函数^[1]. 另一方面, 模糊逻辑与神经网络均可从给定的输入输出数据中建立系统的非线性映射关系以及两者均采用并行处理结构. 正是由于它们之间的共性与差异, 才使得神经网络可以实现模糊逻辑推理, 由此生成的模糊神经网络集中了两者的共同优点. 本文正是依据这样的理由, 利用神经网络的连接主义表达实现 T-S 模糊模型^[2]的模糊逻辑推理, 并由 BP 算法学习模糊规则和精调隶属函数参数, 从而避免了人为主观性. 最后, 将其用于大系统的随机稳态建模中, 获得了满意的结果.

2 模糊推理网原理^[1]

设系统的模糊模型由 T-S 模型描述, 即由以下形式的模糊隐含规则组成

$$R^i: \quad \text{IF } x_1 \text{ is } A_1^i \text{ and } \dots \text{ and } x_n \text{ is } A_n^i \text{ THEN } y_i = f_i(x_1, x_2, \dots, x_n). \quad (1)$$

式中 $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $\bar{x} = (x_1, x_2, \dots, x_n) \subset U$, $A_1, A_2, \dots, A_n \subset F(U)$, 通常用隶属函数 $\mu_{A_j^i}(x_j)$ 表示 A_j^i ($j = 1, 2, \dots, n$), R^i 表示模糊规则基的第 i 条规则, $i = 1, 2, \dots, N$. 若采用代数积推理, 对于给定的一组输入 ($x_1 = x_1^0, x_2 = x_2^0, \dots, x_n = x_n^0$), 则系统的输出为

$$y = f(\bar{x}) = \frac{\sum_{i=1}^N \left(\prod_{j=1}^n \mu_{A_j^i}(x_j^0) \right) \cdot f_i(x_1^0, x_2^0, \dots, x_n^0)}{\sum_{i=1}^N \left(\prod_{j=1}^n \mu_{A_j^i}(x_j^0) \right)}. \quad (2)$$

定义
$$P_i(\bar{x}) = \frac{\prod_{j=1}^n \mu_{A_j^i}(x_j)}{\sum_{i=1}^N \left(\prod_{j=1}^n \mu_{A_j^i}(x_j) \right)}; \quad i = 1, 2, \dots, N; \quad (3)$$

* 国家自然科学基金资助项目(69374007).

本文于 1996 年 11 月 4 日收到, 1997 年 4 月 14 日收到修改稿.

那么

$$y = f(\bar{x}) = \sum_{i=1}^N P_i(\bar{x}) \cdot f_i. \quad (4)$$

则模糊系统(1)式等价于基函数 $P_i(\bar{x})$ 的扩张,文[1]利用 Stone-Weierstrass 定理证明了此种模糊基函数的扩张是一般的函数近似器,故称其为模糊推理网。

3 模糊推理网的神经网络实现

基于模糊推理网原理,采用连接主义表达的神经网络结构如图 1 所示。它由输入层、模糊化层、前件匹配层、强度归一化层和输出层共 5 层以及一个求和节点组成。

1) 输入层:该层的各节点与输入变量 $x_i (i = 1, 2, \dots, n)$ 直接相连,并将它们送至下一层。

2) 模糊化层:该层的每个节点 m_{ij} 对应于一个语言变量,如 NL, NM, NS, ZO, PS, PM 和 PL 等,通过高斯型隶属函数,完成输入变量的模糊化。为简单起见,文中设每个输入变量的论域均划分为 r 个模糊子集,用 c_{ij} 和 b_{ij} 分别表示隶属函数的中心与宽度,即

$$m_{ij} = \mu_{A_j}(x_i) = \exp\{- (x_i - c_{ij})^2 / b_{ij}\}; \quad 1 = 1, 2, \dots, n; \quad j = 1, 2, \dots, r. \quad (5)$$

3) 前件匹配层:该层的每个节点对应于一条模糊规则,它完成模糊规则的前件匹配,通过输入语言变量的组合来形成规则的激励强度,即 $a_k = m_{1j_1} m_{2j_2} \dots m_{nj_n}$,式中 $k = 1, 2, \dots, N; j_1, j_2, \dots, j_n \in \{1, 2, \dots, r\}$ 。

4) 求和节点 S :该节点完成计算 $S = 1 / \sum_{i=1}^N a_i$,为求基函数 $P(\bar{x})$ 做准备。

5) 归一化层:该层节点与前层节点相对应,具有与前层相同的节点个数,它的每一个节点完成相应激励强度的归一化计算,得到模糊推理网的基函数 $P(\bar{x})$,即 $P_k = a_k \cdot S = a_k / \sum_{i=1}^N a_i, k = 1, \dots, N$ 。

6) 输出层:该层节点对基函数 P_k 进行加权线性求和,直接给出清晰化的输出值,避免了去模糊步骤。由(1)式可知,每条规则的后件定义为输入变量的一个函数,所以 $y = \sum_{i=1}^N a_i \cdot f_i / \sum_{i=1}^N a_i = \sum_{i=1}^N P_i \cdot f_i$ 。由图 1 可知, f_i 起着连接权的作用,模糊规则的学习即是对它的修正。

4 学习算法

设期望输出为 y_d ,网络输出为 y ,定义目标函数

$$J = \frac{1}{2} (y_d - y)^2 = \frac{1}{2} \delta^2. \quad (6)$$

采用梯度法,则对第 $i (i = 1, 2, \dots, N)$ 条规则有

$$\frac{\partial J}{\partial f_i} = \frac{\partial J}{\partial y} \cdot \frac{\partial y}{\partial f_i} = - (y_d - y) \cdot P_i = - \delta P_i, \quad (7)$$

$$f_i(t+1) = f_i(t) + \Delta f_i = f_i(t) + \eta \delta P_i. \quad (8)$$

式中 η 为学习效率, $0 < \eta \leq 1$ 。

隶属函数的中心 c_{ij} 和宽度 b_{ij} 亦采用误差反传进行修正,由图 1 及(5)式,得

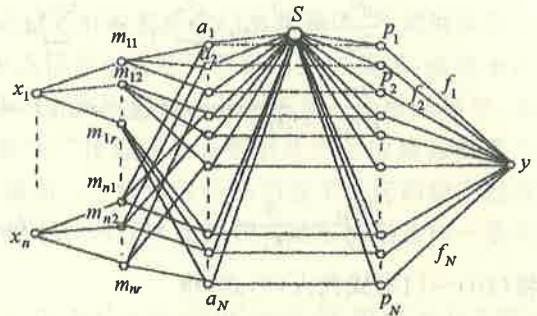


图 1 模糊神经网络结构图

$$\frac{\partial J}{\partial x_{ij}} = \frac{\partial J}{\partial y} \cdot \sum_{k=1}^N \left\{ \frac{\partial y}{\partial P_k} \cdot \frac{\partial P_k}{\partial a_k} \cdot \frac{\partial a_k}{\partial m_{ij}} \cdot \frac{\partial m_{ij}}{\partial x_{ij}} \right\}; \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, r; \quad (9)$$

因为
$$\frac{\partial y}{\partial P_k} = \frac{\partial}{\partial P_k} \left(\sum_{i=1}^N P_i f_i \right) = f_k; \quad (10)$$

$$\frac{\partial P_k}{\partial a_k} = \frac{\partial}{\partial a_k} \left(a_k / \sum_{i=1}^N a_i \right) = \left(\sum_{i=1}^N a_i - a_k \right) / \left(\sum_{i=1}^N a_i \right)^2 = (1 - P_k) / \sum_{i=1}^N a_i; \quad (11)$$

$$\frac{\partial a_k}{\partial m_{ij}} = \frac{\partial}{\partial m_{ij}} (m_{1,j1} m_{2,j2} \dots m_{n,jr}) = m_{1,j1} m_{2,j2} \dots m_{n,jr} / m_{ij} = a_k / m_{ij}; \quad (12)$$

$$j, j1, j2, \dots, jr \in \{1, 2, \dots, r\};$$

$$\frac{\partial m_{ij}}{\partial x_{ij}} = \frac{\partial}{\partial x_{ij}} \exp \{ - (x_i - c_{ij})^2 / b_{ij} \} = 2m_{ij} (x_i - c_{ij}) / b_{ij}; \quad (13)$$

将(10)~(13)式代入(9)式,得

$$\frac{\partial J}{\partial x_{ij}} = -2\delta \left\{ \sum_{k=1}^N f_k P_k (1 - P_k) \right\} (x_i - c_{ij}) / b_{ij}; \quad (14)$$

同理可得

$$\frac{\partial J}{\partial b_{ij}} = -\delta \left\{ \sum_{k=1}^N f_k P_k (1 - P_k) \right\} (x_i - c_{ij})^2 / b_{ij}^2; \quad (15)$$

因为每个输入变量 $x_i (i = 1, 2, \dots, n)$ 的第 $j (j = 1, 2, \dots, r)$ 个语言变量只对全部 $N = r^n$ 条规则中的 $r^{(n-1)}$ 条规则产生作用,隶属函数中的参数 c_{ij} 和 b_{ij} 应该只由受该隶属函数作用的 $r^{(n-1)}$ 条规则进行修正. 所以隶属函数中心 c_{ij} 和宽度 b_{ij} 的修正公式为

$$\Delta c_{ij} = \alpha \delta \left\{ \sum_{k=1}^{r^{(n-1)}} f_k p_k (1 - p_k) \right\} (x_i - c_{ij}) / b_{ij}; \quad (16)$$

$$\Delta b_{ij} = \beta \delta \left\{ \sum_{k=1}^{r^{(n-1)}} f_k p_k (1 - p_k) \right\} (x_i - c_{ij})^2 / b_{ij}^2. \quad (17)$$

式中 α, β 为学习效率, $0 < \alpha \leq 1, 0 < \beta \leq 1$.

5 应用举例

考虑由如下两个子系统组成的大系统,其各子系统的动态过程及目标函数为

$$S1: \begin{cases} y_{11}(k) = 0.11y_{11}(k-1) + 0.43y_{11}(k-2) + 0.47x_{11}(k-1) + 0.6u_{11}(k-1) + \xi_1, \\ u_{11}(k) = 0.4c_{11} - 0.28c_{12} - 0.36y_{11}(k-1), \\ Q_1 = (y_{11}(k) - 1)^2 + c_{11}^2 + c_{12}^2; \end{cases} \quad (18)$$

$$S2: \begin{cases} y_{21}(k) = 0.24y_{21}(k-1) + 0.32y_{21}(k-2) + 0.54x_{21}(k-1) + 0.2u_{21}(k-1) + \xi_2, \\ y_{22}(k) = 0.13y_{22}(k-1) + 0.36y_{22}(k-2) + 0.41x_{21}(k-1) + 0.4u_{22}(k-2) + \xi_3, \\ u_{21}(k) = 0.55c_{21} + 0.5c_{22} - 0.65y_{21}(k-1), \\ u_{22}(k) = 0.27c_{22} + 0.39c_{23} + 0.3y_{22}(k-1), \\ Q_2 = (y_{21}(k) + 1)^2 + (y_{22}(k) - 0.5)^2 + c_{21}^2 + c_{22}^2 + c_{23}^2, \end{cases} \quad (19)$$

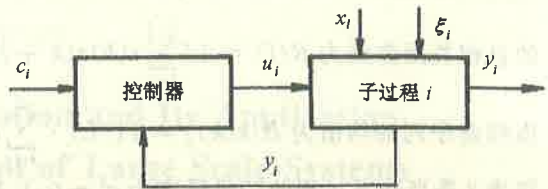
系统的关约束

$$x_{11}(k) = y_{21}(k); \quad x_{21}(k) = y_{11}(k); \quad (20)$$

系统总的目标函数

$$Q = Q_1 + Q_2. \quad (21)$$

式(18)和(19)中的 ξ 是均值为零的高斯型白噪声, c 是由协调器给定的设定点, u 为控制器输出, x 是其它子过程的输出, y 为过程输出, 如图 2 所示. 对于上述动态过程, 当 k 充分大时, 系统达到稳态.

图2 子过程*i*结构图

利用前文给出的基于神经元表达的自适应模糊推理网络, 对上述各子系统的动态过程达到稳态时的模型进行离线辨识. 考虑到实际工业过程中不允许频繁改变控制器设定点以及设定点信号 c 通常为阶跃形式的情况, 这里采用文[3]中的方法获取用于辨识稳态模型的数据. 即先对实际的动态过程施加一个阶跃信号, 利用此时的动态信息辨识过程的动态模型. 文[4]证明了该近似动态模型所产生的稳态模型与实际过程的稳态是一致的, 从而可在动态模型上施加一个阶跃信号 c , 当各子系统的动态过程结束后, 测量各子系统的输出 y , 得到一组辨识数据对 $\{c, y\}$. 以此方式对动态模型施加一系列的阶跃信号, 便得到足够多的稳态模型辨识数据.

依上述方法, 在随机噪声方差水平分别为 $\sigma^2 = 0.0, \sigma^2 = 0.01, \sigma^2 = 0.05$ 和 $\sigma^2 = 0.1$ 时, 用动态过程产生了相应的 200 组数据, 每种噪声水平下均取 140 组数据对模糊神经网络进行训练, 其余 60 组用于检验网络的泛化能力. 以子系统 1 为例, 图 3 是 60 组未参加网络训练的数据在噪声水平为 $\sigma^2 = 0.05$ 时模糊神经网络的泛化情况. 图 4 中实线为模糊神经网络学习时的均方误差曲线, 经 10 次学习后均方误差小于 0.01, 270 次后均方误差已小于 0.001. 用同样数据对含有两个隐层的 BP 神经网络进行训练, 其隐层节点数为 $\{6, 4\}$, 并采用如下变步长算法: $\eta = \max + (\min - \max) \cdot \exp(-\alpha \cdot rms)$, 式中 $\max = 2.0, \min = 0.001, rms$ 为学习时的均方误差, 得到的均方误差曲线如图 4 中虚线所示, 训练 541 次后均方误差小于 0.01, 2000 次后小于 0.002804, 3000 次后小于 0.002798, 此时误差衰减极其缓慢. 另一方面, 模糊神经网络经 1000 次训练后, 对上述 200 组数据进行泛化, 其平均误差为 0.045640, 而 BP 神经网络经 3000 次训练后, 对该 200 组数据进行泛化的平均误差为 0.050814. 由此可见, 本文给出的模糊神经网络具有较快的学习速度和较强的泛化能力. 学习结果还表明, 各语言值的隶属函数均有修正, 避免了人为主观性.

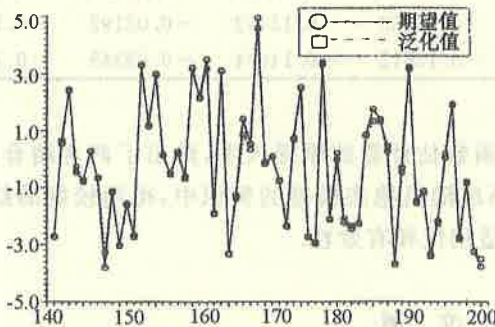


图3 模糊神经网络的泛化情况

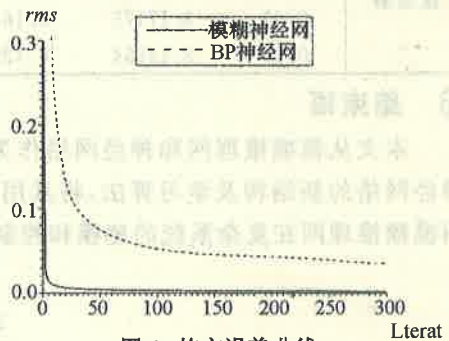


图4 均方误差曲线

为了检验模糊模型的可靠性, 采用自相关函数检验法对过程与模糊模型的输出残差 $\epsilon(k) = y_d(k) - y(k)$ 进行白色性检验^[5]. 若辨识数据长度为 L , 输出残差序列

表1 输出残差均值及检验阈值

$E\{\epsilon_{11}(k)\}$	$E\{\epsilon_{21}(k)\}$	$E\{\epsilon_{22}(k)\}$	$1.98/\sqrt{L}$
-0.000574	0.002637	-0.003324	0.167341

的自相关函数则为 $R(l) = 1/\sum_{k=1}^{L-1} \epsilon(k)\epsilon(k+l)$, $l = 1, 2, \dots, m$ (通常取 $m = 20 \sim 30$). 定义输出残差序列的均值为 $E\{\epsilon(k)\} = (1/L) \cdot \sum_{k=1}^L \epsilon(k) \approx 0$, 自相关系数为 $\rho(l) = R(l)/R(0)$. 表 1 和表 2 是当动态过程的随机噪声为 $\sigma^2 = 0.1$ 、数据长度为 $L = 140$, $m = 20$ 时的检验结果. 由表 1 看到, 三个残差序列的均值都近似为零, 并且 $\{\epsilon_{21}(k)\}$ 的自相关系数 $\rho_{21}(l)$ ($l = 1, 2, \dots, 20$) 全部落在 $|\rho(l)| \leq 1.98/\sqrt{L}$ 的置信区间之内, $\{\epsilon_{11}(k)\}$ 和 $\{\epsilon_{22}(k)\}$ 也各自只有 2 个自相关系数落在置信区间之外, 并且超出的距离很小. 因此, 有 95% 的把握说输出残差序列是白噪声序列. 这进一步验证了模糊神经网络建立的上述动态过程的稳态模型是可靠的.

表 2 输出残差的相关系数

$\rho(l)$	$\{\epsilon_{11}\}$	$\{\epsilon_{21}\}$	$\{\epsilon_{22}\}$	$\rho(l)$	$\{\epsilon_{11}\}$	$\{\epsilon_{21}\}$	$\{\epsilon_{22}\}$	$\rho(l)$	$\{\epsilon_{11}\}$	$\{\epsilon_{21}\}$	$\{\epsilon_{22}\}$
$\rho(1)$	-0.16879	0.05550	-0.02855	$\rho(8)$	0.08875	-0.05046	-0.19395	$\rho(15)$	-0.09429	0.11443	-0.10789
$\rho(2)$	-0.05371	-0.09536	0.06289	$\rho(9)$	-0.05906	-0.09037	0.07455	$\rho(16)$	-0.02896	0.00906	-0.06233
$\rho(3)$	0.02496	-0.02666	0.08334	$\rho(10)$	-0.04950	0.12662	-0.17144	$\rho(17)$	-0.03742	-0.00315	-0.03838
$\rho(4)$	-0.02429	0.06936	-0.03894	$\rho(11)$	0.05361	0.05827	-0.04126	$\rho(18)$	0.02124	-0.02731	0.04453
$\rho(5)$	0.00202	0.02776	0.08096	$\rho(12)$	-0.02202	-0.08443	-0.00551	$\rho(19)$	0.07091	0.06422	-0.08587
$\rho(6)$	0.17736	0.00821	0.02916	$\rho(13)$	0.02775	-0.04661	-0.10169	$\rho(20)$	-0.03526	0.08207	-0.02119
$\rho(7)$	-0.05674	-0.16225	0.07648	$\rho(14)$	0.01485	-0.02981	0.02475				

对于动态过程在噪声水平为 $\sigma^2 = 0.0$, $\sigma^2 = 0.01$, $\sigma^2 = 0.05$ 和 $\sigma^2 = 0.1$ 时的四种情况分别进行建模, 采用关联平衡协调法对四种噪声情况下的过程进行稳态递阶优化, 所得结果见表 3. 从表中可看到协调优化结果是令人满意的, 这也从另一个方面说明了本文给出的基于神经元实现的模糊推理网络是可靠的, 用它进行稳态大系统的建模是一种简单有效的方法.

表 3 大系统的稳态递阶优化结果

	σ^2	Q	c_{11}	c_{12}	c_{21}	c_{22}	c_{23}
解析解		2.10828	0.14677	-0.10291	-0.15620	-0.05547	-0.12490
模型解	0.00	2.11659	0.13101	-0.09241	-0.17576	-0.07899	-0.10715
	0.01	2.15828	0.12367	-0.08202	-0.19314	-0.08382	-0.10926
	0.05	2.17175	0.16327	-0.12883	-0.13492	-0.08192	-0.13057
	0.10	2.19954	0.12019	-0.13612	-0.14694	-0.09349	-0.13580

6 结束语

本文从模糊推理网和神经网络作为一般函数估计器的原理入手, 给出了两者结合的模糊神经网络的新结构及学习算法, 将其用于大系统随机稳态模型的辨识中, 得到较好的结果, 说明模糊推理网在复杂系统的建模和控制中的适用性和有效性.

参 考 文 献

- 1 Wang, L. X. . Fuzzy systems are universal approximators. Proc. IEEE Int. Conf. on Fuzzy Systems, San Diego, 1992, 1163-1170
- 2 Takagi, T. and Sugeno, M. . Fuzzy identification of systems and its application to modeling and control. IEEE Trans. Syst., Man, and Cybern., 1985, 15(1):116-132
- 3 王武义. 工业过程随机稳态优化控制的研究. 西安交通大学博士学位论文, 西安, 1994
- 4 陈庆新, 万百五. 利用工业过程动态信息建立稳态模型及其强一致性分析. ISIO 情形. 控制与决策, 1991, 6(2):90-96

5 方崇智, 萧德云. 过程辨识. 北京: 清华大学出版社, 1988, 474-476

Adaptive Fuzzy Identification and Its Application to the Hierarchical Control of Large Scale Systems

LUO Xuguang

(Systems Engineering Institute, Xi'an Jiaotong University · Xi'an, 710049, PRC)

LIU Ding

(Automation Institute, Xi'an University of Technology · Xi'an, 710048, PRC)

WAN Baiwu

(Systems Engineering Institute, Xi'an Jiaotong University · Xi'an, 710049, PRC)

Abstract: In this paper, a new adaptive fuzzy neural network based on the T-S model is proposed. Also, the network structure and the learning algorithms are given, and it is used in the hierarchical systems for large scale processes operating in random steady state. The simulation results show that the method has the properties of fast convergence, high accuracy and better capability of generalizing. So it is a powerful tool for modeling large scale systems in steady state.

Key words: fuzzy logic; neural network; dynamic processes; steady model; systems identification; large scale systems

本文作者简介

罗旭光 1990年毕业于西北工业大学,获硕士学位,1998年在西安交通大学获博士学位.研究领域:系统建模,智能控制,大系统的稳态优化与控制及多媒体技术等.

刘丁 1957年生.博士,教授.现为西安理工大学自动化学院院长.研究领域:智能控制,大系统的稳态优化与控制,机器人控制等.

万百五 1928年生.教授,博士生导师.1951年毕业于交通大学电信研究所.现为西安交通大学系统工程研究所大系统室主任,自动化学会名誉理事.目前主要研究方向为大系统智能控制.国内外发表论文120余篇.