

一种基于模糊神经网络的非线性系统模型辨识方法¹

李 映 白本督* 焦李成

(西安电子科技大学雷达信号处理重点实验室 西安 710071)

*(西安电子科技大学 296 信箱 西安 710071)

摘 要 该文提出一种非线性系统的模型辨识方法。通过结构的辨识(学习)和参数的辨识(学习),构造了一个模糊神经网络,经调整网络的权值,获得一个精确的模糊模型。对两个非线性系统辨识的仿真结果验证了该方法的有效性。

关键词 模糊神经网络, 结构辨识, 参数辨识, 系统辨识

中图分类号 TP391.4, TN-052

1 引 言

到目前为止,线性系统的辨识已是一个成熟的领域,存在着大量定义完善和十分有效的辨识模型及算法。但在实际应用中,几乎所有的过程都呈现出不同程度的非线性。而线性化作为一种数学的理想化,只是对实际系统的一个近似,在某些情况下才可能有效。因此非线性系统的辨识方法在过去一直是人们致力研究的课题。但总的来说,无论在理论上还是实际应用中,非线性系统辨识还远没有达到完善的地步。

近几年,在把神经网络和模糊系统结合起来(简称为模糊神经系统)的基础上,研究者提出了许多新的非线性系统辨识及控制方法^[1-5]。尽管这些方法在某些场合取得了一定的效果,但是大多数方法只涉及参数的辨识或学习,较少考虑结构的辨识或学习。而正如文献[3]中所指出的,除了参数辨识,结构辨识问题涉及到输入输出空间的划分,模糊规则数目以及前提和结论部分初始参数的确定等等。

本文提出一种同时包括结构辨识和参数辨识的非线性系统的模型辨识方法。在结构辨识中,采用一个无监督的聚类算法^[6],无需有关数据结构的任何先验假设,仅从已知的输入输出数据中就能自动地划分输入输出空间,确定模糊规则的数目及每条规则中前提和结论部分的初始参数。因此结构辨识的目的是为了生成一个能描述给定输入输出数据结构的初始的模糊模型。在参数辨识中,构造了一个四层的模糊神经网络与模糊模型的模糊推理机制相匹配。通过调整网络的权值,也即模糊模型的参数使其具有更高的精度。同时此网络具有通用逼近的特性,这是在诸如建模和控制应用中一个非常重要的特性。最后,我们给出两个例子验证了此方法的有效性。

2 结构辨识

当我们手边只具有被辨识系统的输入输出数据,而对系统的其它信息一无所知,这时聚类方法就成为构造模糊系统结构的一个最有效的方法之一。人们通常采用模糊 C 均值(FCM)算法来构造初始的模糊规则库,但是在 FCM 算法中聚类的类别数必须预先指定。因此这里我们采用一种可以自动地对已知的输入输出数据聚类,而无需有关数据结构的任何先验假设的聚类算法,其主要思想是关系度大的矢量具有相似的特性,可以把这些矢量划分为一类。因此首先选择一点做为参考矢量,找到与参考矢量具有大的关系度的那些矢量,然后用那些矢量的均值代替参考矢量作为聚类中心,算法的具体步骤如下:

¹ 2000-01-12 收到, 2000-11-16 定稿

国家教育部博士点基金和国家部级基金资助项目

设 $X = \{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^p\}$ 是包含 p 个输入输出数据的集合, 每个点表示为 $\mathbf{x}^k = \{x_1^k, x_2^k, \dots, x_n^k, x_{n+1}^k\}$, 其中 $\{x_1^k, x_2^k, \dots, x_n^k\}$ 是第 k 个输入输出数据对的输入矢量, x_{n+1}^k 是相应的输出, 则

(1) 定义 p 个矢量 $\mathbf{v}^k (k = 1, 2, \dots, p)$, 令 $\mathbf{v}^k = \mathbf{x}^k$, 即 \mathbf{x}^k 是 \mathbf{v}^k 的初值.

(2) 计算参考矢量 \mathbf{v}^k 与比较矢量 \mathbf{v}^l 之间的关系度:

$$r_{kl} = \exp[-\|\mathbf{v}^k - \mathbf{v}^l\|^2 / (2\sigma^2)], \quad k = 1, 2, \dots, p, \quad l = 1, 2, \dots, p \quad (1)$$

其中 $\|\mathbf{v}^k - \mathbf{v}^l\|$ 表示 \mathbf{v}^k 和 \mathbf{v}^l 之间的欧氏距离, σ 是高斯函数的宽度.

(3) 改变 \mathbf{v}^k 与 \mathbf{v}^l 之间的关系度, 使

$$r_{kl} = \begin{cases} 0, & \text{若 } r_{kl} < \xi \\ r_{kl}, & \text{其它} \end{cases} \quad (2)$$

其中 ξ 是一个小的常数.

(4) 由下式计算 $\mathbf{w}^k = \{w_1^k, w_2^k, \dots, w_{n+1}^k\}$

$$\mathbf{w}^k = \sum_{l=1}^p r_{kl} \mathbf{v}^l / \sum_{l=1}^p r_{kl}, \quad k = 1, 2, \dots, p \quad (3)$$

(5) 若所有矢量 \mathbf{w}^k 和 $\mathbf{v}^k (k = 1, 2, \dots, p)$ 相同, 则转到 (6), 否则令 $\mathbf{v}^k = \mathbf{w}^k$ 返回 (2).

(6) 基于最终结果 $\mathbf{v}^k, k = 1, 2, \dots, p$, 可以确定聚类的数目等于收敛矢量的数目, 具有相同收敛矢量的原始数据被划分为同一类, 并且收敛矢量是聚类的中心.

通过上面的聚类算法就把已知的数据集分为 m 个类别, 并得到相应的聚类中心 $c_j = (c_{1j}, c_{2j}, \dots, c_{nj}, c_{n+1,j})$, 于是我们就可以如下构造一个初始模糊模型的规则库:

$$\text{规则 } R^j: \text{ If } x_1 \text{ is } A_1^j(x_1), \dots, \text{ and } x_n \text{ is } A_n^j(x_n) \text{ then } y = w_j, \quad j = 1, 2, \dots, m \quad (4)$$

其中 m 是模糊规则的数目, 规则中前提部分的隶属函数 $\mu_{A_i^j}(x_i)$ 用高斯函数表示:

$$\mu_{A_i^j}(x_i) = \exp[-(x_i^2 - m_{ij})^2 / (\sigma_{ij})^2], \quad j = 1, 2, \dots, m \quad (5)$$

其中 m_{ij} 和 σ_{ij} 分别是高斯函数的中心和宽度, $\mathbf{m}_j = (m_{1j}, \dots, m_{nj}) = (c_{1j}, \dots, c_{nj}), \sigma_j = \{\sigma_{1j}, \dots, \sigma_{nj}\}$ 的值通过计算第 j 类的输入数据点与其聚类中心的偏差即可得到. 规则中结论部分的实数 w_j 表示为

$$w_j = c_{n+1,j} \quad (6)$$

模糊模型总的输出用所有规则结论的线性组合来表示:

$$y = \sum_{j=1}^m \alpha_j w_j, \quad \alpha_j = \prod_{i=1}^n \mu_{A_i^j}(x_i) \quad (7)$$

3 参数辨识

3.1 模糊神经网络

为了与所构造的模糊模型的模糊推理机制相匹配, 我们设计了一个 4 层的模糊神经网络. 这个网络的结构如图 1 所示, 由输入层、隶属函数层、规则层和输出层组成.

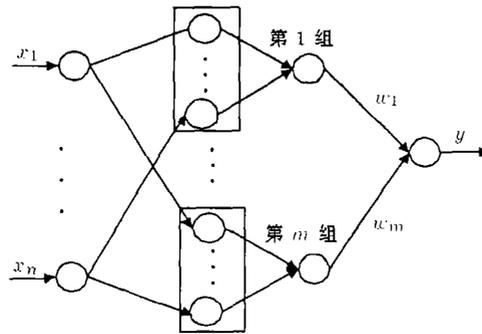


图 1 模糊神经网络的结构框图

第 1 层 (输入层) 每个输入神经元接收一个信号 x_i , 并把它传递到第 2 层。

第 2 层 (隶属函数层) 由 m 组 (表示 m 个规则) 组成, 每组有 n 个神经元 (表示 n 个规则前提)。第 j 组的第 i 个神经元 ($i = 1, 2, \dots, n; j = 1, 2, \dots, m$) 只与第 1 层中的第 i 个神经元相连接, 其生成的隶属函数为 (5) 式。

第 3 层 (规则层) 也即前提集合层, 用来实现每个规则的前提匹配。这一层有 m 个神经元, 其中第 j 个神经元只接收来自第 2 层中第 j 组神经元的输入, 其输出为

$$\alpha_j = \prod_{i=1}^n \mu_{A_i^j}(x_i) \quad (8)$$

第 4 层 (输出层) 也即去模糊化层, 或规则集合层。由于模型总的输出为所有规则结论的线性组合, 因此这一层神经元的输出简单定义为

$$y = \sum_{j=1}^m \alpha_j w_j = \sum_{j=1}^m w_j \prod_{i=1}^n \mu_{A_i^j}(x_i) \quad (9)$$

3.2 通用逼近性

显然, 图 1 的模糊神经网络在它的结构中对一个模糊 if-then 规则集进行编码, 并以与所采用的模糊推理机制相匹配的方式来处理信息, 这两点保证了此网络具有对偶的特性, 也就是既可以把它看作是一个模糊系统, 也可以把它看作是一个神经网络。一些类型的模糊系统已经被证明是通用逼近器^[7-9], 因此下面的定理 1 表明所提出的模糊神经网络具有通用逼近的特性。

定理 1 设 $f: U \subset R^n \rightarrow R$ 是定义在致密集 U 上的一连续函数, $x \in U$ 使 $x = (x_1, x_2, \dots, x_n)$ 。考虑用上面定义的模糊规则类型和推理机制来构造模糊神经网络, 则存在一模糊规则集或一网络使得对于任何 $x \in U$ 和 $\varepsilon > 0$ 满足下式:

$$\sup\{|f(x) - y| | x \in U\} \leq \varepsilon \quad (10)$$

其中 y 是输入为 x 时网络的输出。

这个定理的证明类似于文献 [7]。

3.3 有监督学习

我们采用有监督的学习算法来调整模糊神经网络的权值 $(m_{ij}, \sigma_{ij}, w_j)$, 也即规则的前提和结论参数, 以期得到一个精确的模糊模型。该算法的基本原理是简单的梯度下降法。定义误差函数为

$$E = \frac{1}{2}(y - y^*)^2 \quad (11)$$

其中 y 为网络输出, y^* 为期望输出。通过递归应用链式规则, 最后得到参数的调整公式为

$$\Delta w_j = -\eta \cdot (y - y^*) \cdot \alpha_j \quad (12)$$

$$\Delta m_{ij} = -\eta \cdot (y - y^*) \cdot w_j \cdot \alpha_j \cdot 2(x_i - m_{ij}) / \sigma_{ij}^2 \quad (13)$$

$$\Delta \sigma_{ij} = -\eta \cdot (y - y^*) \cdot w_j \cdot \alpha_j \cdot 2(x_i - m_{ij})^2 / \sigma_{ij}^3 \quad (14)$$

式中 η 是学习率, 在迭代过程中可自适应地调整。

4 实验结果

本节通过两个例子来验证上面所提出的辨识方法的有效性。

例 1 气相沉积过程的辨识

气相沉积过程是碳 / 碳复合材料生成的一种方法, 整个过程属于复杂的非线性过程。系统输入包括 C_3H_6 和 N_2 两种气体参加反应的体积比 V 及反应温度 W , 系统输出为热解碳真密度 ρ 。我们取 100 个样本用做训练样本, 取另外 49 个样本用做测试样本。

第 1 步进行结构辨识, 首先对样本进行归一化处理, 然后利用聚类算法对所有训练样本进行聚类, 聚类结果表明, 样本被划分为 5 个子类, 有 5 个聚类中心。因此模糊规则的数目确定为 5, 用得到的聚类中心构造一个初始模糊模型的规则库, 规则前提和结论部分的参数如表 1 所示。第 2 步进行参数辨识, 即设计一个模糊神经网络来调整规则前提部分和结论部分的参数, 从而得到一个精确的模糊模型。根据输入变量和规则的数目, 确定该网络的输入层有 2 个神经元, 隶属函数层有 5×2 个 (分为 5 组, 每组 2 个) 神经元, 规则层有 5 个神经元, 输出层有 1 个神经元。利用有监督的梯度下降学习算法对网络进行训练, 得到了模糊模型最终的参数, 如表 2 所示。最后用得到的模糊模型对测试样本进行辨识, 图 2(a) 和图 2(b) 分别画出了实际系统与辨识系统的输出 p 和 p^* 。仿真结果表明, 用本文提出的辨识方法对碳 / 碳气相沉积过程热解碳真密度进行辨识的结果是令人满意的。

表 1 结构辨识中得到的模糊模型的初始参数

j	前提部分	结论部分
	$(m_{1j}, m_{2j}, \sigma_{1j}, \sigma_{2j})$	w_j
1	(0.3538, 0.1216, 0.2048, 0.1050)	0.2863
2	(0.1562, 0.5794, 0.0998, 0.2041)	0.7139
3	(0.5318, 0.9234, 0.2006, 0.0939)	0.1892
4	(0.5554, 0.4924, 0.2051, 0.1013)	0.3525
5	(0.8707, 0.0978, 0.1056, 0.0993)	0.2236

表 2 参数辨识中得到的模糊模型的最终参数

j	前提部分	结论部分
	$(m_{1j}, m_{2j}, \sigma_{1j}, \sigma_{2j})$	w_j
1	(0.3212, 0.1038, 0.1894, 0.0946)	0.2312
2	(0.1329, 0.5988, 0.0753, 0.1643)	0.6743
3	(0.5286, 0.8713, 0.1740, 0.0381)	0.2005
4	(0.6015, 0.5576, 0.4174, 0.2361)	0.3186
5	(1.0109, 0.1224, 0.1443, 0.1345)	0.1862

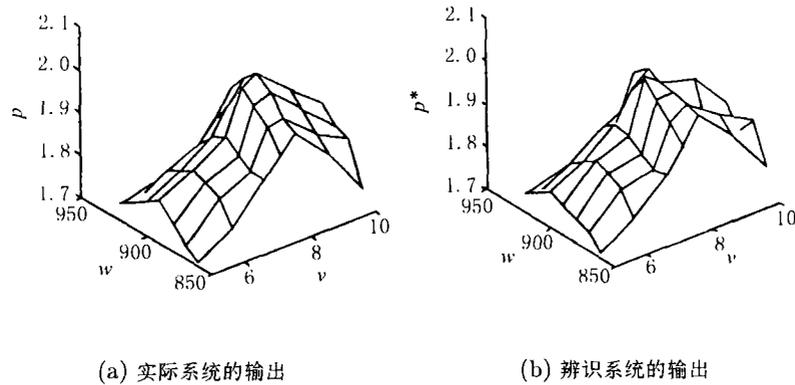


图 2

例 2 辨识对象由如下非线性系统^[10]产生

$$y = 0.5(1 + \sin(2\pi x_1) \cos(2\pi x_2)) + \xi \quad (15)$$

式中两个输入变量 x_1 和 x_2 定义在 $[0, 1] \times [0, 1]$ 上, 噪声分量 ξ 的方差为 0.05。每个输入变量都按照一个均匀概率密度函数随机产生, 共产生 400 个输入输出数据对。在结构辨识中, 生成 12 个聚类中心, 得到一个初始的模糊模型。在参数辨识中, 根据聚类结果设计一个模糊神经网络对初始参数进行调整, 得到模型的最终参数。图 3(a) 和图 3(b) 分别显示了实际非线性系统与辨识系统的输出 y 和 y^* 。显然, 辨识系统具有很好的逼近能力。

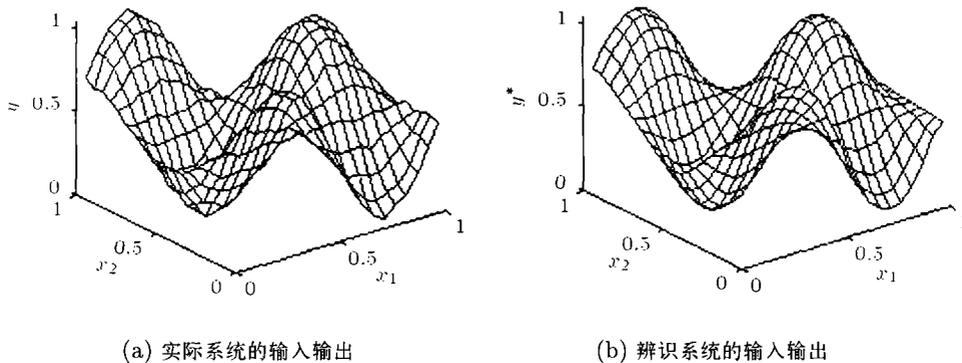


图 3

5 结 束 语

本文提出一种新的非线性系统模型辨识方法。该方法不但能够进行参数辨识, 而且能够进行结构辨识。在结构辨识中, 首先通过聚类算法自动地从已知的输入输出数据中生成一个初始的模糊模型。接着在参数辨识中, 设计了一个模糊神经网络, 通过有监督学习逐步调整网络的权值, 也即模糊模型的参数以使模型具有更高的精度。该网络的结构使它以 if-then 模糊规则的形式编码所学的知识, 并且按照模糊推理的原则来处理数据。同时网络具有通用逼近的能力。在我们的方法中, 由于采用了无监督学习和有监督学习相结合的混合算法, 能根据环境的不同自动地对辨识系统的结构和参数进行学习, 因此具有一定的自适应能力。

参 考 文 献

- [1] J. Nie, Constructing fuzzy model by self-organizing counterpropagation network, IEEE Trans. on system, Man and Cybernetics 1995, 25(6), 963-970.
- [2] R. P. Li, M. Mukaidono, Fuzzy modeling and clustering neural network. Control and Cybernetics, 1996, 25(2), 225-242.
- [3] J. Jang, C. Sun, Neuro-fuzzy modeling and control. Proceeding of IEEE, 1995, 83(3), 378-406.
- [4] J. Keller, H. Tahani, Neural network implementation of fuzzy logic. Fuzzy Sets and Systems, 1992, 45(1): 1-12.
- [5] Y. C. Chen, C. C. Teng, A model reference control structure using a fuzzy neural network, Fuzzy Sets and Systems, 1995, 73(1), 291-312.
- [6] C. C. Wong, C. C. Chen, A hybrid clustering and gradient descent approach for fuzzy modeling, IEEE Trans. on system, Man and Cybernetics-Part B, 1999, 29(6), 686-693.
- [7] L. X. Wang, M Mendel, Fuzzy basis function, universal approximation, and orthogonal least-squares learning, IEEE Trans. on Neural Network, 1992, 3(5), 807-814.
- [8] J. Castro, M. Delgado, Fuzzy systems with defuzzification are universal approximators, IEEE Trans. on system, Man and Cybernetics, 1996, 26(1): 149-152.
- [9] B. Kosko, Fuzzy systems as universal approximators, Proceeding of IEEE International Conference on Fuzzy System, San Diego, CA, 1992, 1153-1162.
- [10] F. Mauricio, G. Fernando, Design of fuzzy system using neurofuzzy networks, IEEE Trans. on Neural Network, 1999, 10(4), 815-827.

A MODEL IDENTIFICATION APPROACH OF
NONLINEAR-SYSTEMS BASED ON FUZZY NEURAL NETWORKS

Li Ying Bai Bendu* Jiao Licheng

*(Key Lab. for Radar Signal Processing, Xidian University, Xi'an 710071, China)***(P. O. Box 296, Xidian University, Xi'an 710071, China)*

Abstract This paper presents a model identification approach of nonlinear-systems. Through structure identification (learning) and parameter identification (learning), a fuzzy neural network is constructed. Having adjusted the weights of the fuzzy neural network, a precise fuzzy model is obtained. Finally, the effectiveness of the proposed technique is confirmed by the simulation results of two nonlinear systems.

Key words Fuzzy neural networks, Structure identification, Parameter identification, System identification

李 映: 女, 1969 年生, 博士生, 主要研究方向为模式识别和信号处理, 计算智能与混合系统等.

白本督: 男, 1972 年生, 硕士生, 主要兴趣为模式识别, 神经网络等.

焦李成: 男, 1959 年生, 教授, 博士生导师, 主要研究方向包括: 计算智能和集成系统, 非线性智能信号和图象处理.