

一种推广的模糊神经网络及学习算法

章云 毛宗源 周其节 徐建闽

杨宜民

(华南理工大学自动控制工程系·广州, 510641) (广东工业大学自动化所·广州, 510090)

摘要: 本文采用广义模糊神经网络实现分段建模的思想, 给出了一种广义 k -均值聚类算法, 该算法能同时确定模糊规则的个数和相应的参数. 仿真结果表明该算法是可行和有效的.

关键词: 模糊系统; 神经网络; 系统辨识; 局部模型

1 引言

对复杂非线性系统采用一个全局数学模型来描述它的全部特性常常是困难的, 其结果势必造成网络模型结构复杂, 网络学习速度慢. 如果把对象的工作空间划分为功能上相互独立的不同区域, 然后对各区域分别采用局部模型去逼近对象的局部特性, 往往是十分有效的. 这种分段建模的思想事实上可用模糊神经网络模型来实现. 在模糊神经网络模型中, 模糊规则的前件相当于对象工作空间的一个划分, 模糊规则的后件类似于对象的局部模型. 通常的模糊神经网络的规则后件常常是常数型的, 这也是模糊神经网络在逼近复杂非线性系统需要较多规则的原因所在. 本文从分段建模的思想来考察模糊神经网络, 扩展了规则后件的类型, 提出了一种新的学习算法, 该算法的核心是聚类算法与最小二乘法的综合, 因而具有较快的收敛速度.

2 模糊神经网络与分段建模

图 1 为模糊神经网络模型^[1], 其中第一层节点是输入节点, 第二层节点是输入语言值节点, 第三层节点是规则节点, 第四层节点是输出语言值节点, 第五层是输出节点. 第三层规则节点的输入是模糊规则的前提, 它的输出是模糊规则的结论. 通常采用的模糊规则为

$$\text{if } x_1 \text{ is } F_1^i, x_2 \text{ is } F_2^i, \dots, x_d \text{ is } F_d^i, \text{ then } y \text{ is } G^i. \quad (1)$$

如果采用高斯型隶属函数、乘积推理规则、中心去模糊, 则网络模型的输出与输入有如下关系^[2]:

$$p_j(x) = \frac{\exp(-(x - m_j)^T Q_j^{-1}(x - m_j))}{\sum_{j=1}^q \exp(-(x - m_j)^T Q_j^{-1}(x - m_j))}, \quad (2a)$$

$$y = \sum_{j=1}^q c_j p_j(x). \quad (2b)$$

式中 $x \in \mathbb{R}^d, m_j \in \mathbb{R}^d, Q_j = \text{diag}\{\sigma_{ij}\} \in \mathbb{R}^{d \times d}, (j = 1, 2, \dots, q)$.

如果从分段建模的思想考察式(2), 可以得知 c_j 起局部模型的作用, $p_j(x)$ 反映局部模型

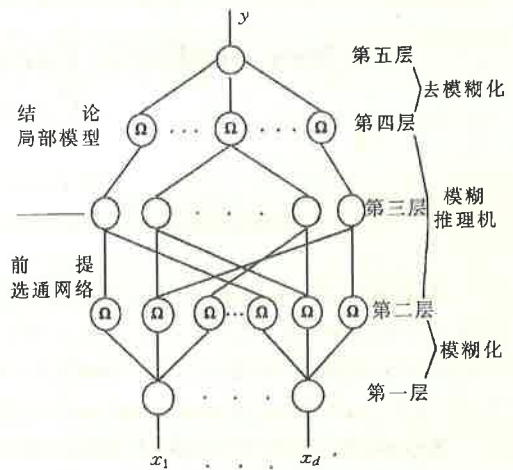


图 1 模糊神经网络

* 广东省自然科学基金(960101)和国家攀登计划认知科学(神经网络)重大关键项目(B1-131-343)资助课题.
本文于 1996 年 7 月 29 日收到, 1997 年 1 月 17 日收到修改稿.

适用的程度. 由于 c_j 是一种常数型的局部模型, 所以模糊神经网络在逼近复杂非线性系统需要较多的规则. 为了减小网络规模和改善网络性能, 采用一种推广的模糊规则是有意义的. 即将式(1)的模糊规则推广为

$$\text{if } x_1 \text{ is } F_1^j, x_2 \text{ is } F_2^j, \dots, x_d \text{ is } F_d^j, \text{ then } y = f_j(x). \quad (3)$$

这样的模糊神经网络称为广义模糊神经网络, 它的输入与输出映射关系为

$$f(x) = \sum_{j=1}^q f_j(x) p_j(x). \quad (4)$$

式中 $p_j(x)$ 按式(2a) 计算.

以什么样的函数作为局部模型现没有定论. 由于多项式是一类有着良好性质的函数, 常被用作局部模型. 不失一般性, 本文以一次多项式作为局部模型来讨论, 即 $f_j(x)$ 为

$$f_j(x) = \theta_{0j} + x_1 \theta_{1j} + \dots + x_d \theta_{dj}, \quad (j = 1, 2, \dots, q). \quad (5)$$

3 学习算法

假定待重构的非线性函数为 $f_d(x)$, 我们希望

$$\|f_d(x) - f(x)\| = \left\| f_d(x) - \sum_{j=1}^q f_j(x) p_j(x) \right\| = \min. \quad (6)$$

一种做法是先确定模糊规则数 q , 再以梯度信息对网络中所有参数进行寻优. 由于这是一种非线性寻优过程, 所以常常存在局部极值、收敛速度慢的缺陷^[3]. 另一种学习方法是以 k -均值聚类算法得到隶属函数中的非线性参数 m_j 和 Q_j , 再以梯度信息对网络中剩下的参数进行寻优^[1,2], 这样做可以加快学习的速度. 但是 k -均值聚类算法完全没有考虑输出样本信息的作用, 所以得到的参数 m_j 和 Q_j 完全取决于输入样本分布的情况, 而这与输入样本产生的方式有较大的关系. 为了得到参数 m_j 和 Q_j 的合理值, 应该同时考虑输入与输出样本信息的影响. 基于此, 本文给出如下的广义 k -均值聚类算法.

给定一组学习样本 $\{x^i \in \mathbb{R}^d, f_d(x^i) \in \mathbb{R}, i = 1, 2, \dots, N\}$, 记

$$E(x^i, \theta_j) = f_d(x^i) - f_j(x^i). \quad (7)$$

广义 k -均值聚类算法由下述迭代公式组成:

$$h_j(x^i) = \|x^i - m_j^{(t)}\|^2 + \lambda E^2(x^i, \theta_j^{(t)}), \quad (8a)$$

$$w_{ij}^{(t)} = \begin{cases} 1, & h_j(x^i) < h_k(x^i), j \neq k, k \in (1, 2, \dots, q), \\ 0, & \text{otherwise,} \end{cases} \quad (8b)$$

$$\{m_j^{(t+1)}, \theta_j^{(t+1)}\} = \arg \min_{m, \theta} \left\{ \sum_{i=1}^N w_{ij}^{(t)} [\|x^i - m\|^2 + \lambda E^2(x^i, \theta)] \right\}. \quad (8c)$$

从式(5)和式(8c)知, 参与优化的局部模型参数 θ_j 是一类线性参数, 用最小二乘法可容易求得, 不存在局部极值问题. 式(8)中权重 λ 的大小对模型分类有影响, λ 太小将以数据的密度来分类, λ 太大又会太强调输出样本信息的作用, 使得最后得到的模型趋于雷同, 本文算法对 λ 进行自适应变化, 以求得到合理的 λ 值, 当 w_{ij} 稳定时迭代学习完成. 记

$$S_j = \{x^i | w_{ij} = 1, i = 1, 2, \dots, N\}, \quad (9)$$

集合 S_j 给出了输入空间的一个划分.

下面讨论用广义 k -均值聚类算法确定广义模糊神经网络参数的具体步骤. 本算法分两个部分, 首先确定需要局部模型的个数以及输入空间的一个划分 S_j ; 然后对 S_j 进行必要的细分, 得到隶属函数中的参数, 同时也得到了所需的模糊规则数. 在进行第一部分运算时, 先将 λ 的值取小一步, 使得样本数据按其本身的密度来分类; 然后, 逐步加大 λ 的值使得局部模型产生

的误差作用加强,从而可以得到较好的局部模型参数.

Step 1 取初值 $k \geq 2$, 一个较小的数 $\lambda \in [0, 1]$, 随机选取 $m_j^{(0)}$ 和 $\theta_j^{(0)}$ ($j = 1, 2, \dots, k$);

Step 2 按式(8)迭代, 直到 w_{ij} 稳定;

Step 3 λ 增大一倍, 重复 Step 2, 若本次的 w_{ij} 与上次一样转 Step 4;

Step 4 计算并判断 $\sum_{x^i \in S_j} E^2(x^i, \theta_j) < \epsilon$? 若不满足, 则需分裂第 j 个局部模型, 这时只需对

S_j 中的数据重复上述步骤即可, k 变为 $k + 1$;

Step 5 假定经过上述步骤得到了 k 个局部模型和输入空间的一个划分 S_j , 由于 S_j 中的数据可能是多段稠密, 因而需要对 S_j 中的数据分布进行判断和必要的细分. 计算 S_j 中的数据方差 σ_{lj} ($l = 1, 2, \dots, d$), 取

$$\tilde{S}_j = \{x^i \mid |x_l^i - m_{lj}| \leq \sigma_{lj}, l = 1, 2, \dots, d\}; \tag{10}$$

Step 6 判断 $\tilde{S}_j = \tilde{S}_j \cap S_j$? 若关系成立, 则 S_j 不需细分; 否则转 Step 7;

Step 7 将 S_j 的中心增加一个, k 变为 $k + 1$, 按 k -均值聚类算法细分 S_j , 也就是在式(8)中取 $\lambda = 0$ 便可;

Step 8 将得到新划分, 重复 Step 5~Step 7, 直到勿需细分为止, 此时 $q = k$.

4 仿真结果

给定二个期望曲面式(11)和式(12), 其期望图形如图 2(a)和图 3(a)所示. 局部模型选用式(5)的一次多项式, 每个曲面仿真的样本点均为 300 个. 采用本文算法分别对其进行仿真研究. 图 2(b)和图 3(b)是仿真得到的结果. 可看出, 由本文算法得到的模糊神经网络对非线性曲面有较好的重构性.

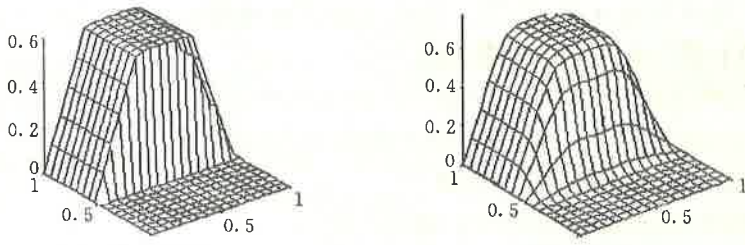


图 2 式(11)的期望图形和重构图

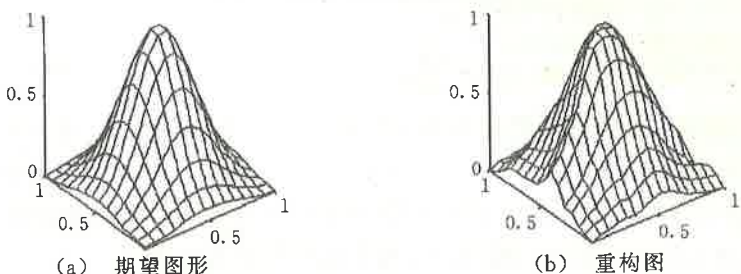


图 3 式(12)的期望图形和重构图

$$f_d(x) = \begin{cases} 2x_2, & 1 \geq x_1 > 0.5, & 0.3 \geq x_2 \geq 0, \\ 0.6, & 1 \geq x_1 > 0.5, & 0.7 > x_2 > 0.3, \\ -2x_2 + 2, & 1 \geq x_1 > 0.5, & 1 \geq x_2 \geq 0.7, \\ 0, & 0.5 \geq x_1 \geq 0, & 1 \geq x_2 \geq 0, \end{cases} \tag{11}$$

$$f_d(x) = \exp(-10(x_1 - 0.5)^2 - 10(x_2 - 0.5)^2), \quad x_1, x_2 \in [0, 1]. \tag{12}$$

5 结束语

本文将分段建模思想与模糊神经网络相结合, 拓广了模糊神经网络研究的思路. 给出的广义 k - 均值聚类算法能同时考虑输入与输出样本信息的影响, 从而得到更加合理的隶属函数参数值, 避免了对非线性参数的寻优. 另外, 该算法同时对局部模型中的参数进行了优化, 也得到了模糊规则数. 仿真结果表明本文提出的算法是可行和有效的.

参 考 文 献

- 1 Lin, C. T. . Neural Fuzzy Control Systems with Structure and Parameter Learning. London: World Scientific, 1994
- 2 王立新. 自适应模糊系统与控制. 北京: 国防工业出版社, 1995
- 3 张铃, 张钊. 神经网络中 BP 算法分析. 模式识别与人工智能, 1994, 7(3): 191—195
- 4 Jacobs, R. A. and Jordan, M. I. . Learning piecewise control strategies in a modular neural network architecture. IEEE Trans. Syst. Man, and Cybern. , 1993, 23(2): 337—345

A Generalized Fuzzy Neural Network and Its Learning Algorithm

ZHANG Yun, MAO Zongyuan, ZHOU Qijie and XU Jianmin

(Department of Automatic Control Engineering, South China University of Technology • Guangzhou, 510641, PRC)

YANG Yimin

(Research Institute of Automation, Guangdong University of Technology • Guangzhou, 510090, PRC)

Abstract: This paper presents a generalized fuzzy neural network that can realize the strategy about approximating a function using piecewise models. A generalized k - means algorithm is given. The number and parameters of fuzzy rules can be simultaneously obtained by the algorithm. Computer simulations show that the method is feasible and efficient.

Key words: fuzzy system; neural network; identification; local model

本文作者简介

章 云 1963 年生. 副教授. 1982 年、1986 年先后在湖南大学获学士、硕士学位, 随后留校任教, 1989 年调入广东工业大学, 现在华南理工大学攻读在职博士. 研究领域: 智能控制, 机器人, 离散事件动态系统.

毛宗源 1936 年生. 1962 年 7 月毕业于大连工学院. 现为华南理工大学自动控制工程系教授, 博士生导师. 研究兴趣包括工业自动化, 智能控制. 已出版著作 6 本, 发表论文 70 篇.

周其节 见本刊 1998 年第 1 期第 38 页.

徐建闽 1960 年生. 1982 年 1 月在江西工学院电机工程系获工学学士学位, 1988 年和 1994 年在华南理工大学自动化系获工学硕士和博士学位. 现任华南理工大学交通学院副院长, 教授. 主要研究领域为自适应控制, 鲁棒控制, 智能控制, 现代交通工程, CIMS, 机器人控制等.

杨宜民 1945 年生. 教授, 博士生导师. 1969 年毕业于广东工学院, 毕业后留校任教, 先后在东北大学、北方工业大学、中山大学研究学习四年. 在日本筑波大学、东京大学从事“微驱动器、微机器人”研究两年半. 现为广东工业大学自动化研究所所长. 研究领域: 自动控制, 机器人.