

文章编号: 1001-0920(2002)03-0306-04

神经模糊系统中模糊规则的优选

贾立, 俞金寿

(华东理工大学 自动化研究所, 上海 200237)

摘要: 提出一种基于两级聚类算法的自组织神经模糊系统, 该系统采用两级聚类算法(改进的最近邻域聚类算法和 Gustafson-Kessel 模糊聚类算法)对输入/输出数据进行模糊聚类, 并由模糊聚类的划分熵确定最优划分, 建立模糊模型, 模型精度可由梯度下降法进一步提高。仿真结果表明, 这种神经模糊系统具有结构简单、规则数少、学习速度快以及建模精度高等特点。

关键词: 神经模糊系统; 两级聚类; 改进的最近邻域聚类算法; GK 模糊聚类算法; 划分熵

中图分类号: TP 18 **文献标识码:** A

Optimal choice of fuzzy rules in neuro-fuzzy systems

JIA Li, YU Jin-shou

(Research Institute of Automation, East China University of Science and Technology, Shanghai 200237, China)

Abstract: A self-organizing neuro-fuzzy system based on two-stage clustering algorithm is proposed. Two-stage clustering algorithm consisting of the nearest-neighborhood clustering algorithm and Gustafson-Kessel fuzzy clustering algorithm with cluster validity criteria is used to partition the input-output space. The optimal number of fuzzy rules can be determined via fuzzy entropy as the criterion of cluster validation. A supervised scheme is utilized for constructing more optimal fuzzy model. Two simulation results show that the proposed method can provide optimal model structure and parameters for fuzzy modeling and possesses high learning efficiency.

Key words: neuro-fuzzy system; two-stage clustering algorithm; proved nearest-neighborhood clustering algorithm; GK fuzzy clustering algorithm; fuzzy partition entropy

1 引言

神经模糊系统是目前较为活跃的研究领域之一, 并已成功应用于复杂系统的建模。它将人工神经网络低层次的学习和计算能力与模糊逻辑系统高层次仿人思维和推理的能力相融合, 使二者各取所长、共生互补^[1-3]。神经模糊系统在设计时常会遇到以下几个问题: 隶属度函数的确定、模糊规则的选取和模糊推理的实现。对于模糊规则的获取, 通常采用

聚类学习算法, 研究表明, 大多数聚类学习算法, 如模糊 c -均值聚类算法、K-Meoid 方法等, 均需用户事先给定聚类个数 c , 且聚类个数 c 直接影响模糊划分的有效性。但如何选择合适的 c 值, 却是模糊建模中的一大难题^[4]。

针对该问题, 本文提出一种基于两级聚类算法的自组织神经模糊系统, 该系统由前向多层神经网络构成, 采用两级聚类学习算法: 改进的最近邻域聚类算法和 GK 模糊聚类算法, 简称为一次聚类和二

收稿日期: 2001-01-11; 修回日期: 2001-04-28

作者简介: 贾立(1975—), 女, 甘肃武威人, 博士生, 从事神经网络与模糊逻辑融合技术等研究; 俞金寿(1939—), 男, 浙江海宁人, 教授, 博士生导师, 从事化工过程控制、工业过程建模与控制等研究。

次聚类。首先, 系统用改进的最近邻域聚类算法对输入输出数据进行预处理, 起到精简样本、滤去样本中重复信息的作用, 从而减少二次聚类算法中的迭代次数; 然后, 系统将一次聚类输出空间的聚类中心作为 GK 模糊聚类的样本, 并由模糊聚类的划分熵确定最优划分, 再根据 Sugeno 和 Yasukawa 提出的映射方法^[5], 将输出空间的模糊聚类映射到输入空间, 从而得到优化的模糊规则。在二次聚类中, 系统直接对输出变量进行聚类, 而不是输入变量。对于多输入系统而言, 这样会减少计算量, 加快学习速度, 更容易得到精简的模糊规则。

需要指出的是, 在二次聚类中系统采用了模糊 c- 均值聚类算法和 Gustafson-Kessel 算法的结合算法, 简称 GK 模糊聚类算法。这是一种提取模糊规则的有效方法, 因为对于每个聚类都有各自不同的几何空间形状。在一次聚类和二次聚类的基础上, 系统采用梯度下降法对模糊模型进一步优化, 以便达到更高的建模精度。

2 模糊规则的优选

本文提出的基于两级模糊聚类的模糊规则优选方法如图 1 所示。首先用改进的最近邻域聚类算法对样本聚类, 并将输入 / 输出空间的聚类中心作为二次聚类的输入 / 输出样本。在二次聚类算法中, 采用 Gustafson-Kessel 模糊聚类算法对输入 / 输出数据中的输出数据进行模糊聚类, 并由模糊聚类的划分熵确定最优划分, 然后根据 Sugeno 和 Yasukawa 提出的映射方法, 将输出空间的模糊聚类映射到输入空间, 从而得到优化的模糊规则。

2.1 一次聚类

采用改进的最近邻域聚类算法对样本聚类, 可描述为: 给定一组输入输出数据 $x_i(k)$ ($i = 1, 2, \dots, M$), $y(k)$ ($k = 1, 2, \dots, P$), 划分输入 / 输出空间, 并将聚类中心点作为二次聚类的输入。算法简述如下:

Step 1: 根据输入数据的分布确定邻域半径 r , 将输入输出数据对 $X(1) = \{x_1(1), x_2(1), \dots, x_M(1), y(1)\}$ 作为输入 / 输出空间的第一个聚类,

并设聚类中心 $c_{1i} = X_i(1)$ ($i = 1, 2, \dots, M + 1$), 此时的聚类个数 $N = 1$ 。

Step 2: 对于第 k 对输入输出数据 $X(k) = \{x_1(k), x_2(k), \dots, x_M(k), y(k)\}$, 计算 $x_i(k)$ ($i = 1, 2, \dots, M + 1$) 到每一个聚类中心 c_{ji} ($j = 1, 2, \dots, N$; $i = 1, 2, \dots, M + 1$) 的距离, 并计算

$$D_j = |X(k) - c_j| = \min_{i=1, \dots, N} |X(k) - c_j|$$

如果 $D_j > r$, 则表明第 k 对输入输出数据不属于已有的聚类, 建立一个新聚类, 令 $c_{N+1} = X(k)$, $N = N + 1$; 如果 $D_j \leq r$, 则表明第 k 对输入输出数据属于第 J 个聚类, 调整聚类中心 $c_j = c_j + \lambda D_j$, 其中 λ 表示学习率, $\lambda = \lambda_0 / (N_j + 1)$ ($\lambda_0 \in [0, 1]$)。

Step 3: 令 $k = k + 1$, 若 $k < P$, 则返回 Step 2。

Step 4: 将输入 / 输出空间的聚类中心作为二次聚类的输入 / 输出数据, 即令

$$\begin{aligned} \bar{x}_i(j) &= c_{ji}, \quad i = 1, 2, \dots, M \\ \bar{y}(j) &= c_{j(M+1)}, \quad j = 1, 2, \dots, N \end{aligned}$$

2.2 二次聚类

2.2.1 预备知识

1) Gustafson-Kessel 模糊聚类算法^[6,7]

GK 模糊聚类算法是模糊 c- 均值聚类算法和 Gustafson-Kessel 算法的结合算法, 其目标函数一般为如下形式

$$J(U, V, Z) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m D_{ij}^2 \quad (1)$$

其中, n 为训练样本数, c 为模糊聚类数, m 为影响隶属度矩阵模糊化程度的指数权重, v_j 为第 i 个聚类的中心向量, u_{ij} 的值在 $[0, 1]$ 之间, 表示样本集中的元素 z_j 属于第 i 个聚类的程度。 v_j 和 u_{ij} 可由 Picard 迭代计算得到

$$v_i = \frac{\sum_{j=1}^n u_{ij}^m z_j}{\sum_{j=1}^n u_{ij}^m}, \quad i = 1, 2, \dots, c \quad (2)$$

$$u_{ij} = \frac{1/D_{ij}^2}{\sum_{h=1}^c 1/D_{hj}^2}, \quad i = 1, 2, \dots, c; j = 1, 2, \dots, n \quad (3)$$

$$D_{ij}^2 = (Z_j - V_i)^T A_i (Z_j - V_i) \quad (4)$$

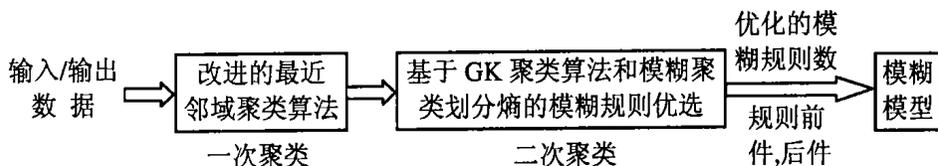


图 1 基于两级模糊聚类的模糊规则优选方法结构

$$A_i = (\rho \det(C_i))^{1/M} C_i^{-1}$$

$$C_i = \frac{\prod_{j=1}^n u_{ij}^m (Z_j - V_i) (Z_j - V_i)^T}{\prod_{j=1}^n u_{ij}^m} \quad (5)$$

2) 模糊聚类的划分熵^[8,9]

样本集 $Z (|Z| = n)$ 的任一模糊 c 划分 U 的划分熵为

$$H_m(U, c) = - \frac{1}{n} \sum_{j=1}^n \sum_{i=1}^c u_{ij} \cdot \log_a(u_{ij}) \quad (6)$$

其中,对数的底 $a (1, +\infty)$,且约定当 $u_{ij} = 0$ 时有 $u_{ij} \cdot \log_a(u_{ij}) = 0$.

由聚类的有效性可知,好的聚类结果应为尽可能明晰的划分,且对应于小的划分熵

$$\min_{\Omega} \{H_m(U, c)\}$$

其中 Ω 为模糊聚类参数,在 GK 模糊聚类算法中有 $\Omega = \{m, U, V, c\}$.因此,选择合适的参数集 Ω 可保证模糊聚类获得有效的划分.

2.2.2 二次聚类算法

在一次聚类的基础上,采用 Gustafson-Kessel 模糊聚类算法对输入/输出数据对 $\{\bar{x}_i(k), \bar{y}(k)\} (i = 1, 2, \dots, M; k = 1, 2, \dots, N)$ 的输出数据进行模糊聚类,并由模糊聚类的划分熵确定最优划分,然后根据 Sugeno 和 Yasukawa 提出的映射方法,将输出空间的模糊聚类映射到输入空间.算法简述如下:

Step 1: 选择最大模糊聚类数 c_{max} , 最大迭代次数 T , 指数权重 m 和停止准则 ϵ

Step 2: 令模糊聚类数 $c = 1, 2, \dots, c_{max}$, 选择隶属度矩阵的初始值 $U(0)$;

Step 3: 令迭代次数 $t = 1, 2, \dots, T$, 由式(3)~(5)分别计算 $v_i(t)$ 和 $u_{ij}(t)$, 如果 $|v_i(t) - v_i(t-1)| < \epsilon$, 则停止迭代, 否则继续 Step 3;

Step 4: 由式(6)计算 $H_m(c)$, 如果 $H_m(c-2) > H_m(c-1), H_m(c) > H_m(c-1)$, 则停止迭代, 得到最优模糊划分数 $c = c-1$, 否则返回 Step 2;

Step 5: 令模糊规则后件参数 $h_i = v_i, i = 1, 2, \dots, c$;

Step 6: 将输出空间的模糊聚类映射到输入空间, 即有如下关系

$$A_1(\bar{x}_1) = A_2(\bar{x}_2) = \dots = A_M(\bar{x}_M) = B(\bar{y})$$

$$i = 1, 2, \dots, c \quad (7)$$

从而得到模糊规则前件隶属度函数的中心点 $c_{ij}, i = 1, 2, \dots, c; j = 1, 2, \dots, M$, 并由下式计算隶属度宽度

$$\sigma_{ji} = \min_{k=1,2,\dots,N, k \neq j} |c_{ji} - c_{ki}| / \rho \quad (8)$$

其中 ρ 表示交迭参数, 通常取 $1/\rho = 2^{[10]}$.

3 参数优化

通过两级聚类算法, 确定了模糊规则数以及模糊规则和隶属度函数的初始值, 但这样得到的模糊模型并不一定满足建模精度的要求, 因此需要采用梯度下降法对参数进一步优化.

各参数调整如下

$$c_{ji}(k+1) = c_{ji}(k) - 2\alpha(y - d)(h_j - y)(x_i - c_{ji})H_j / \sigma_{ji}^2 + \beta(c_{ji}(k) - c_{ji}(k-1)) \quad (9)$$

其中, α 是学习因子, β 是惯性因子, 而

$$H_j = \frac{\exp\left[-\sum_{i=1}^M (x_i - c_{ji})^2 / \sigma_{ji}^2\right]}{\sum_{j=1}^N \exp\left[-\sum_{i=1}^M (x_i - c_{ji})^2 / \sigma_{ji}^2\right]} \quad (10)$$

$$\sigma_{ji}(k+1) = \sigma_{ji}(k) - 2\alpha(y - d)(h_j - y) \times \sum_{i=1}^M (x_i - c_{ji})^2 H_j / \sigma_{ji}^3 + \beta(\sigma_{ji}(k) - \sigma_{ji}(k-1)) \quad (11)$$

$$h_j(k+1) = h_j(k) - \alpha(y - d)H_j + \beta(h_j(k) - h_j(k-1)) \quad (12)$$

4 仿真研究

为了说明本文方法的有效性, 将本文提出的神经模糊系统分别用于非线性系统和化学设备手工操作模型的建模研究.

1) 考虑如下非线性系统

$$y = 0.2 + 0.8 \exp(-x_1) + 0.4 \sin(2\pi x_2)$$

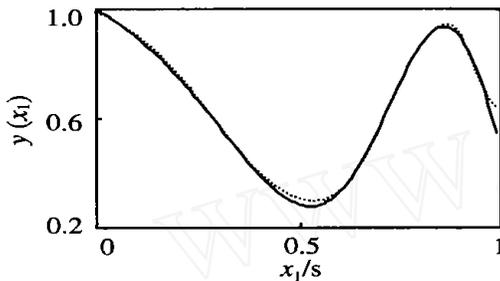
其中 $x_2 = 1 - x_1^2$. 仿真中取 x_1 为 $[0, 1]$ 之间的随机数. 定义模型精度评价指标为

$$MSE = \frac{1}{P} \sum_{k=1}^P [y(k) - d(k)]^2$$

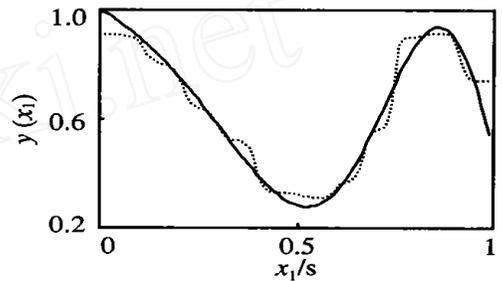
随机产生 $\{x, y\}$ 的 100 个数据对并作为一次聚类的输入, 选择聚类半径 $r = 0.3$. 经改进的最近邻域聚类算法后, 100 对输入/输出数据被划分成 38 个聚类, 将这些聚类的输出空间聚类中心点作为二次聚类的输入, 由模糊划分熵确定最优的聚类数为 5, 即产生 5 条模糊规则, 再将 $[0, 1]$ 区间等分成 100 份, 并以等分后的 100 个数据点作为测试样本. 此时 $MSE = 2.1 \times 10^{-3}$. 经过 100 次迭代优化后, $MSE = 0.233 \times 10^{-3}$, 并且建模精度可通过梯度下降法进一步提高. 优化后的隶属度函数及模糊规则如表

表 1 非线性系统的模糊规则

R		非线性系统的模糊模型			
R_1	If x_1 is exp	$\frac{(x_1 - 1.0052)^2}{0.0742^2}$	and x_2 is exp	$\frac{(x_2 - 0.0200)^2}{0.0742^2}$	then y is 0.5314
R_2	If x_1 is exp	$\frac{(x_1 - 0.9002)^2}{0.1606^2}$	and x_2 is exp	$\frac{(x_2 - 0.3842)^2}{0.1606^2}$	then y is 1.0446
R_3	If x_1 is exp	$\frac{(x_1 - 0.7198)^2}{0.9289^2}$	and x_2 is exp	$\frac{(x_2 - 0.4912)^2}{0.0106^2}$	then y is 0.5990
R_4	If x_1 is exp	$\frac{(x_1 - 0.4117)^2}{0.2346^2}$	and x_2 is exp	$\frac{(x_2 - 0.6426)^2}{0.2346^2}$	then y is 1.1159
R_5	If x_1 is exp	$\frac{(x_1 - 0.0137)^2}{0.1844^2}$	and x_2 is exp	$\frac{(x_2 - 1.2606)^2}{0.1844^2}$	then y is 0.1877



(a) 基于两级聚类算法的神经模糊模型



(b) 基于最近邻域聚类算法的神经模糊模型

图 2 模糊模型的输出和实际输出的比较

1 所示, 模型预测输出如图 2(a) 所示, 图 2(b) 是文献[11]提出的基于最近邻域聚类算法的神经模糊模型($r = 0.125$, 产生了 11 条模糊规则)的预测输出, $MSE = 3.2 \times 10^{-3}$ 。图 2 中实线为实际输出, 虚线为模型输出。从仿真结果可以看出, 本文提出的算法具有结构简单、规则数少以及建模精度高等特点, 优于常规的聚类算法。

2) 将神经模糊系统用于化学设备手工操作模型的建模。这里研究的是产生聚合体的化学设备, 设备的启动是一个非常复杂的过程, 由操作工来控制。通常操作工根据原料浓度、原料浓度的变化量及原料流量确定原料流量的设定点, 然后将原料流量的设定点及实际的流量送入 PD 控制器, 如图 3 所示。

将文献[5]给出的 70 组实测数据作为样本, 经两级聚类后得到 7 条模糊规则, 再由 150 次迭代, 得到设备手工操作的优化模糊模型。模糊模型的输出和实际输出的比较如图 4 所示(其中, 实线为实际输出, 虚线为模型输出), 能够得到较为满意的建模精

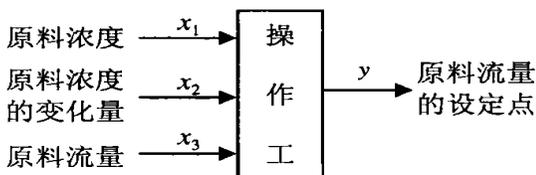


图 3 化学设备的手工操作模型

度。限于篇幅, 这里不再给出各变量的隶属度函数及模糊规则。

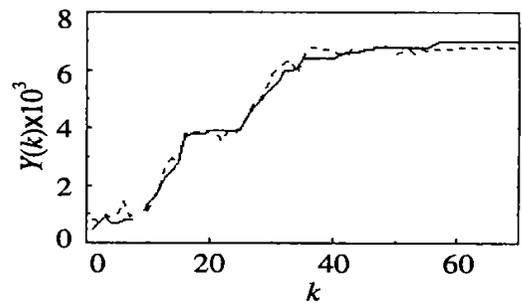


图 4 模糊模型的输出和实际输出的比较

5 结 论

本文针对神经模糊系统设计中的难题, 提出了两级聚类的概念, 通过两级聚类算法: 改进的最近邻域聚类算法和 Gustafson-Kessel 模糊聚类算法对输入/输出数据进行模糊聚类, 并由模糊聚类的划分熵确定最优划分。仿真结果表明, 这种神经模糊系统具有结构简单、规则数少、学习速度快以及建模精度高等特点, 并且建模精度可根据实际要求通过梯度下降法进一步提高。该方法成功地应用于非线性系统的建模, 为神经模糊系统中模糊规则的优选提供了一条新途径。

(下转第 314 页)

$\times 100$ 以下几乎没有影响,随着问题规模的增大影响将逐渐增加。

4) 对于大规模的该模型用普通的分支定界法或割平面法求解,显然是很麻烦的。

5 结 语

本文方法对于目标和约束具有很好的光滑性时,采用经典的算法更好些,可以较好地保证收敛性和速度。算法的复杂性问题与 Step 2 中采用的算法有关,数值例子表明,转化所增加的约束对中等规模问题程序的执行速度影响并不明显。另外我们容易将上述算法与遗传算法结合起来计算,这时的主要问题是收敛速度会变得很慢,有关这方面的内容将另文给出。

本文方法的优点是开辟了一条通用的求解 IP 的途径,可以结合不同的求解无约束问题方法进行计算。根据推论 1,任何变量有界的 IP 或 MIP 转化成等价的 NIP 或 NIP 的方法,对于 0-1 规划 NIP 和 NIP 是一致的,同样对于非 0-1 整数规划问题, NIP 和 NIP 对于求解也是等价的,但究竟哪种形式更有利于求解还有待于进一步研究。在实际中,许多问题的变量都是有界的,但有时变量有界明确给出,因此,我们可以通过约束来求出或估计变量的上下界,一般很容易求出线性整数规划的变量的界。对于无界变量的 IP 和 MIP 问题还有待于进一步研究。

参考文献(References):

- [1] 徐光辉,刘彦佩,程侃. 运筹学基础手册[M]. 北京: 科学出版社, 1999.
- [2] Eiselt H A, Sandblom C L. Integer programming and network models[M]. Berlin: Springer Press, 2000.
- [3] 邢文训,谢金星. 现代优化计算方法[M]. 北京: 清华大学出版社, 1999.
- [4] 高峰,张连生. 多项式 0-1 整数规划的两个连续化途径[J]. 上海大学学报(J of Shanghai Univ), 1999, 5(2): 95-98.
- [5] Dimitris Bertsimas, Georgios Perakis, Sriheer Tayeer. A new algebraic geometry algorithm for integer programming[J]. Management Science, 2000, 46(7): 999-1008.
- [6] 吴志远,邵惠鹤,吴新余. 基于遗传算法的退火精确罚函数非线性约束优化方法[J]. 控制与决策(Control and Decision), 1998, 13(2): 136-140.
- [7] Yang X Q, Mees A I, Campbell K. Simulated annealing and penalty methods for binary multicommodity flow problems[J]. Progress in Optimization, 2000, 6(1): 93-105.
- [8] 赵蔚. 两层多目标规划罚函数法[J]. 自动化学报(Acta Automatica Sinica), 1998, 24(3): 331-337.
- [9] Rubinov A M, Glover B M, Yang X Q. Extended Lagrange and penalty functions in continuous optimization[J]. Optimization, 1999, 46(3): 327-351.
- [10] Rubinov A M, Glover B M, Yang X Q. Decreasing functions with applications to penalization[J]. SIAM J Optim, 1999, 10(1): 289-313.

(上接第 309 页)

参考文献(References):

- [1] Yan Shi, Masaharu M iamoto. A new approach of neuro-fuzzy learning algorithm for tuning fuzzy rules[J]. Fuzzy Sets and System s, 2000, 112: 99-116.
- [2] Mauricio Figueiredo, Fernando Gomide. Design of fuzzy systems using neurofuzzy networks[J]. IEEE Trans on Neural Networks, 1999, 10(4): 815-827.
- [3] Wang L X, Mendel J M. Generating fuzzy rules by learning from examples[J]. IEEE Trans on System s, Man and Cybernet, 1992, 22(6): 1414-1422.
- [4] Milligan G W, Cooper M C. An examination of procedure for detecting the number of clusters in a data set[J]. Psychometrika, 1985, (50): 159-179.
- [5] Michio Sugeno, Takahiro Yasukawa. A fuzzy-logic-based approach to qualitative modeling[J]. IEEE Trans on Fuzzy System s, 1993, 1(1): 7-31.
- [6] D E Gustafson, W C Kessel. Fuzzy clustering with a fuzzy covariance matrix[A]. Proc IEEE Int Conf on Fuzzy System s[C]. San Diego, 1979. 761-766.
- [7] A F Gomez-Skameta, M Delgado, M A Vila. About the use of fuzzy clustering techniques for fuzzy model identification[J]. Fuzzy Sets and System s, 1999, (106): 179-188.
- [8] J C Bezdek. Pattern recognition with fuzzy objective function algorithms[M]. New York: Plenum, 1981.
- [9] D A Linkens, Min-You Chen. Input selection and partition validation for fuzzy modeling using neural network[J]. Fuzzy Sets and System s, 1999, (107): 299-308.
- [10] Lin C T, Lee C S G. Neural-network-based fuzzy logic control and decision system[J]. IEEE Trans on Computer, 1991, 40(12): 1320-1336.
- [11] Yin Wang, Gang Rong. A self-organizing neural network based fuzzy system[J]. Fuzzy Set and System s, 1999, (103): 1-11.