第 4 期
2001 年 4 月

电　子　学　报
ACTA ELECTRONICA SINICA

Vol.29　No.4
April　2001

基于模糊聚类的改进模糊辨识方法

王宏伟 ,詹荣开 ,贺汉根
(国防科学技术大学机电工程与自动化学院 ,长沙 410073)

摘　要: 针对以往模糊建模方法不能很好优化模糊模型输入空间的问题 ,本文提出了一种基于新的目标函数的模糊聚类方法 ,从而使模型的输入输出空间映射空间具有逼近实际输出的能力 ,从而达到优化模型结构的目的. 仿真实例表明 ,该方法能够辨识非线性系统 ,能显著提高建模的精度.

关键词: 目标函数; 模糊辨识; 模糊聚类; 非线性系统
中图分类号: TP273　**文献标识码**: A　**文章编号**: 0372-2112 (2001) 04-0436-003

Improving Fuzzy Identifying Method Based on Fuzzy Clustering

WANG Hong-wei ,ZHAN Rong-kai ,HE Han-gen
(College of Mechatronics Engineering and Automation , National University of Defense Technology , Changsha 410073 , China)

Abstract: In accordance with the problem that the past fuzzy identifying method can not optimize the input-output space of fuzzy model ,this paper proposes a fuzzy clustering method based on a new objective function. This method enables the input-output mapping space to approach the real output ,thus optimizing the model structure. Simulation example demonstrates that this method can identify non-linear systems and significantly improve modeling accuracy.

Key words: objective function ;fuzzy clustering ;fuzzy identification ;non-linear systems

1　引言

对于非线性动态系统的辨识 ,目前还存在着很多困难 ,常用的方法有两种: 一是用多线性模型在平衡点附近近似描述非线性系统 ,这对于严重的非线性系统如何做到平稳切换 ,减小系统误差仍然缺乏有效的手段; 二是根据被控对象已知的信息 ,选择与之相近的非线性模型. 基于 T-S 模型的模糊辨识方法 ,其本质就是非线性模型 ,易于表达复杂非线性系统的动态模型.

设辨识对象为 $P(U, Y)$,它是一个多输入单输出系统 ,T-S 模型经过变形后其形式如下:

$$R^i: \text{if } Z \text{ is } \bar{Z}_i, \mu_i \text{ then } y^i = \theta_i^T \vec{Z} (i = 1, 2, \cdots, c) \qquad (1)$$

其中 , Z 是输入向量 , $Z = (x_1, x_2, \cdots, x_M)^T \subset R^M$; $\bar{Z}_i$ 是第 i 类的中心向量 , $\bar{Z}_i = (\bar{x}_{i1}, \bar{x}_{i2}, \cdots, \bar{x}_{iM})^T$; μ_i 是输入向量在第 i 类里的隶属度; y^i 是第 i 类规则的后件输出; θ_i 为第 i 类规则的后件参数向量 , $\theta_i = (p_0^i, p_1^i, \cdots, p_M^i)^T$; $\vec{Z}$ 是后件部分的输入向量 , $\vec{Z} = (1, x_1, x_2, \cdots, x_M)^T$; c 为规则数.

模型(1)输出为

$$\bar{y} = \sum_{i=1}^{c} \mu_i y^i \qquad (2)$$

下面基于上述描述的模糊模型来研究模糊辨识方法.

2　问题的描述

在文献[2]中 ,讨论了对于模型(1)的基于模糊聚类目标函数 J_1 和辨识目标函数 J_2 的模糊辨识方法 ,见式(3)和(4). 在研究中 ,发现该方法不能提高系统辨识精度. 主要是没有考虑系统输出对输入空间的影响 ,从而输入空间划分并不是最优的 ,造成一些数据落到输出空间之外 ,造成误差 ,模型的泛化能力下降.

$$J_1 = \sum_{k=1}^{n} \sum_{i=1}^{c} (\mu_i^k)^q \| Z(k) - \bar{Z}_i \|^2$$

$$(q > 1; i = 1, 2, \cdots, c; k = 1, 2, \cdots, n) \qquad (3)$$

$$J_2 = \frac{1}{2} \sum_{k=1}^{n} (y_k - \bar{y}_k)^2 \qquad (4)$$

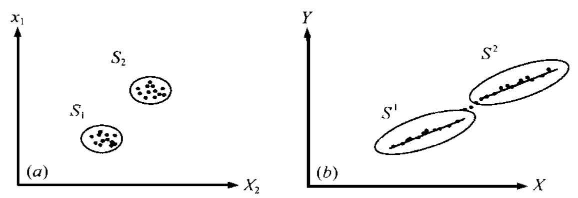


图 1　(a) 对输入量的聚类 x_2; (b) 模型的输出逼近实际输出 X 的情况

上式中 , $Z(k)$ 表示第 k 次输入向量; y_k 为第 k 次系统输出的实际值; $\bar{y}_k$ 为模糊模型(1)的第 k 次输出值 , n 为样本总数; 其余符号与式(1)中的符号相同. 为了详细说明这个问题 ,以两个输入 ,一个输出为例进行阐述. 对输入量 x_1 , x_2 聚为两

收稿日期:1999-11-15 ;修回日期:2000-05-30
基金项目:国防预研基金

类,见图 1(a),图 1(b)给出了辨识的情况.经过辨识后,S_1 类映射成区域 S^1,S_2 类映射成区域 S^2.图 1(b)中小点表示实际值,实线表示后件线性函数,由于没有考虑实际输出对输入空间的影响,而只对输入量 x_1, x_2 聚类,使得一些输出样本点漏到了图 1(b)的椭圆外边,这样使得辨识误差增大.

3　改进的模糊辨识方法

首先,引入下面的目标函数.

定义 1　式(5)称为模糊聚类目标函数 J

$$J(\mu_1, \bar{Z}_i, f_i, \theta_i) = \sum_{k=1}^{n}\sum_{i=1}^{c}(\mu_i^k)^q \| Z(k) - \bar{Z}_i \|^2 + \sum_{k=1}^{n}\sum_{i=1}^{c}(f_i^k)^q (y_k - \theta_i^T Z(k))^2 \tag{5}$$

其中,μ_i^k 是第 k 次输入向量在第 i 类里的隶属度;$Z(k)$ 表示第 k 次输入向量;$\bar{Z}_i$ 是第 i 类的中心向量,$\bar{Z}_i = (\bar{x}_{i1}, \bar{x}_{i2}, \cdots, \bar{x}_{iM})^T$;$f_i^k$ 表示模型的后件线性函数的输出逼近实际输出的加权值,其满足 $\sum_{i=1}^{c} f_i^k = 1$;$f_i^k \in [0,1]$;y_k 为系统的第 k 次实际输出值,$k = 1,2,\cdots,n$;θ_i 为第 i 条规则的后件参数向量,$\theta_i = (p_0^i, p_1^i, \cdots, p_M^i)^T (i = 1,2,\cdots,c)$;$\vec{Z}(k)$ 是后件部分的第 k 次输入向量,$\vec{Z}(k) = [1, x_1(k), x_2(k), \cdots, x_M(k)]^T$;$c$ 为规则数.q 为加权因子,n 为样本总数.

式(5)中的第一项定义了输入向量到中心向量的距离尽可能小,反映了模糊输入空间的划分;第二项表达了在这个划分输入空间下,模糊模型的后件输出逼近实际输出的能力,这说明模糊模型的映射空间的输出与实际输出误差尽可能小,使聚类更加有效.

根据这个模糊聚类目标函数,可以定义一个新的辨识目标函数

定义 2　式(6)称为基于模糊聚类目标函数(5)的辨识目标函数 J'

$$J' = \sum_{k=1}^{n}(y_k - \sum_{i=1}^{c}\mu_i^k \cdot f_i^k \cdot y_k^i)^2 \tag{6}$$

其中,y_k^i 表示模糊模型(1)第 k 次采样第 i 条规则后件线性函数的输出.式(6)中其余符号的含义与定义 1 中的符号含义相同.

由于模糊模型中每条规则都是由输入子空间和相应的输入输出映射子空间组成的.模糊目标函数 J 和目标函数 J 考虑了这两个问题.首先,利用 μ_i^k 考虑了输出空间的划分.其次,利用 f_i^k 考虑了模型输出值逼近实际输出值的程度,这实际考虑了输入输出映射子空间的划分.第三,当 $Z(k)$ 属于第 i 个输入子空间 S_i 时,前件部分的隶属度 μ_i^k 趋于 1,那么 f_i^k 也应当趋近于 1;当 $Z(k)$ 不属于第 i 个输入子空间 S_i 时,前件部分的隶属度 μ_i^k 趋于 0,那么 f_i^k 也应当趋近于 0.即

$$\begin{aligned} &\text{if } Z(k) \in S_i, \text{then } \mu_i^k \text{ 和 } f_i^k \text{ 全都接近于 } 1 \\ &\text{if } Z(k) \notin S_i, \text{then } \mu_i^k \text{ 和 } f_i^k \text{ 全都接近于 } 0 \end{aligned} \tag{7}$$

根据式(7),辨识目标函数 J 可以进行如下变化:

$$\begin{aligned} J' = &\sum_{k\in S_1}(y_k - \mu_1^k f_1^k y_k^1)^2 + \sum_{k\in S_2}(y_k - \mu_2^k f_2^k y_k^2)^2 + \cdots \\ &+ \sum_{k\in S_c}(y_k - \mu_c^k f_c^k y_k^c)^2 \approx \sum_{k\in S_2}(\mu_1^k f_1^k)^2 (y_k - y_k^1)^2 \\ &+ \sum_{k\in S_2}(\mu_2^k f_2^k)^2 (y_k - y_k^2)^2 + \cdots + \sum_{k\in S_c}(\mu_c^k f_c^k)^2 (y_k - y_k^c)^2 \end{aligned} \tag{8}$$

这样根据式(7)和式(8)的结果,将目标函数 J 转换为下面目标函数

$$J' = \sum_{k=1}^{n}(\mu_i^k f_i^k)^2 (y_k - y_k^i)^2 \tag{9}$$

定理 1　假定 θ_i 已知,给定加权因子 q,规则数 c,那么满足下面条件能使 J 最小.

$$\mu_i^k = \frac{1}{\left[\sum_{j=1}^{c}\frac{\| Z(k) - \bar{Z}_i \|^2}{\| Z(k) - \bar{Z}_j \|^2}\right]^{\frac{1}{(q-1)}}} \tag{10-a}$$

$$f_i^k = \frac{1}{\sum_{j=1}^{c}\left[\frac{(| y_k - \theta_i^T \vec{Z}(k)|)}{(| y_k - \theta_j^T \vec{Z}(k)|)}\right]^{\frac{1}{(q-1)}}} \tag{10-b}$$

$$\bar{Z}_i = \sum_{k=1}^{n}(\mu_i^k)^q Z(k) \Big/ \sum_{k=1}^{n}(\mu_i^k)^q \tag{10-c}$$

式(10)中的符号含义与定义 1 中的符号含义相同.

证明　考虑拉格朗日乘子集.因为 $\sum_{i=1}^{c}\mu_i^k = 1$;$\sum_{i=1}^{c} f_i^k = 1$,有

$$\begin{aligned} \bar{J}(\mu_i, \bar{Z}_i, f_i, \lambda, \beta) = &\sum_{k=1}^{n}\sum_{i=1}^{c}(\mu_i^k)^q \| Z(k) - \bar{Z}_i \|^2 \\ &+ \sum_{k=1}^{n}\sum_{i=1}^{c}(f_i^k)^q (y_k - \theta_i^T Z)^2 + \sum_{k=1}^{n}\lambda q \\ &\cdot [1 - \sum_{i=1}^{c}\mu_i^k] + \sum_{i=1}^{n}\beta q [1 - \sum_{i=1}^{c} f_i^k] \end{aligned} \tag{11}$$

由于 μ_i^k 之间相互独立,故 $\bar{J}(\mu_i, \bar{Z}_i, f_i, \lambda, \beta)$ 关于 μ_i^k 的最小化等价于下面目标函数关于 μ_i^k 的最小化.

$$J(\mu_i^k) = (\mu_i^k)^q \| Z(k) - \bar{Z}_i \|^2 + q\lambda [1 - \sum_{i=1}^{c}\mu_i^k] \tag{12}$$

通过式(12)对 μ_i^k 求导,并令其等于零,就能得出式(10-a).

$\bar{J}(\mu_i, \bar{Z}_i, f_i, \lambda, \beta)$ 关于 $\bar{Z}_i$ 的最小化等价于下面目标函数关于 $\bar{Z}_i$ 的最小化.

$$J(\bar{Z}_i) = \sum_{k=1}^{n}\sum_{i=1}^{c}(\mu_i^k)^q \| Z(k) - \bar{Z}_i \|^2 \tag{13}$$

通过式(13)对 $\bar{Z}_i$ 求导,并令其等于零,就能导出式(10-c).

同样 f_i^k 之间也相互独立,所以 $\bar{J}(\mu_i, \bar{Z}_i, f_i, \lambda, \beta)$ 关于 f_i^k 的最小化等价于下面目标函数关于 f_i^k 的最小化.

$$J(f_i^k) = (f_i^k)^q (y_k - \theta_i^T \vec{Z})^2 + q\beta [1 - \sum_{i=1}^{c} f_i^k] \tag{14}$$

通过式(14)对 f_i^k 求导,并令其等于零,就能得出式(10-b).证毕.

定理 2　假定 μ_i^k 和 f_i^k 已知,那么后件参数向量 θ_i 满足下列条件时,能够使 J 最小化.

$$\theta_i = (\Phi^T U_i^2 F_i^2 \Phi)^{-1} (\Phi^T U_i^2 F_i^2 Y) \tag{15}$$

其中,$U_i = \text{diag}[\mu_i^k]_{n\cdot n} (i = 1,2,\cdots,c; k = 1,2,\cdots,n)$;

$F_i = \text{diag}[f_i^k]_{n\cdot n} (i = 1,2,\cdots,c; k = 1,2,\cdots,n)$;

$\Phi = [\vec{Z}(1), \vec{Z}(2), \cdots, \vec{Z}(n)]^T$;

证明:由于 μ_i^k 和 f_i^k 已知,因此关于 θ_i 使目标函数(10)最

小化的问题,实际上是熟知的最小二乘估计问题.因此可以得出式(15)的结果.

通过上述介绍,本文将模糊辨识算法总结如下:

(1) 初始设置 c、q、θ_i,给定 $\overline{Z}_i$, $i=1,\cdots,c$; $k=1,\cdots,n$.

(2) 根据式(10)的模糊聚类算法对辨识样本进行聚类,确定 μ_i^k,同时计算 f_i^k.

(3) 通过样本数据和它的隶属度 μ_i^k 和 f_i^k 来计算 $\overline{Z}_i$ 和 θ_i.

(4) 计算 $J\left(J=\frac{1}{n}\sum_{k=1}^{n}(y_k-\overline{y}_k)^2\right)$.如果 J 小于阈值或相邻两次 J 不变,则转(5),否则转(2).

(5) 如果 J 满足辨识精度,则辨识算法结束,否则 $c=c+1$,转(1).

在辨识时,由于规则数 c 不能根据经验确定,所以只能按照递增的方式不断寻找,通过不断递增寻找到一个使目标函数最小的 c 值.当然 c 值也不能太大,因为这涉及一个聚类有效性的问题.聚类有效性有这样的特性,当 c 值达到一定时,其聚类指标函数值将增大.

4　仿真实例

设非线性系统如下形式[2]:

$$y(k+1)=\frac{y(k)y(k-1)y(k-2)u(k-1)u(k-1)(y(k-2)-1)+u(k)}{1+y(k-1)^2+y(k-2)^2} \tag{16}$$

其中 $u(k)=$

$$\begin{cases}\sin(\pi k/25), & k<250\\ 1.0, & 250\le k<500\\ -1.0, & 500\le k<750\\ 0.3\sin(\pi k/25)+0.1\sin(\pi k/32)+0.6\sin(\pi k/10), & 750\le k<1000\end{cases} \tag{17}$$

利用本文模糊辨识方法与采用文献[2,3]中方法分别对上述非线性系统进行辨识.在辨识中将 $y(k)$、$y(k-1)$、$u(k)$ 作为输入量,$c=3$,$q=2$.辨识结果为 $J=2.5271e-4$,其模糊模型的描述为

R_1:if　Z is $\overline{Z}_1$ then $y(k+1)=$

$-0.0385+1.0755y(k)-0.0999y(k-1)-0.0113u(k)$

R_2:if　Z is $\overline{Z}_2$ then $y(k+1)=$

$-0.0966-0.1944y(k)-0.1207y(k-1)+0.9324u(k)$

R_3:if　Z is $\overline{Z}_3$ then $y(k+1)=$

$$0.2498+0.3797y(k)-0.6432y(k-1)+0.4648u(k) \tag{18}$$

其中,中心向量为

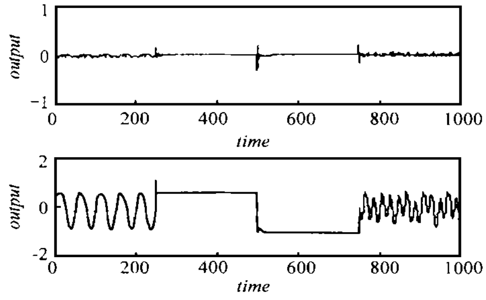


图 2　误差变化和辨识结果

$$[\overline{Z}_1\quad \overline{Z}_2\quad \overline{Z}_3]=\begin{bmatrix}-0.9323 & -0.2926 & 0.5113\\ -0.9295 & -0.2804 & 0.5055\\ -0.9751 & -0.3330 & 0.8550\end{bmatrix} \tag{19}$$

图 2 给出了误差变化和辨识结果,其上图为均方误差的变化曲线,下图为辨识结果.可以看出模型的输出结果基本上与样本数据基本吻合,另外从表 1 来看,本文模糊辨识方法又很高的辨识精度.

表 1　本文方法与其他辨识方法的比较

模型名称	规则数	均方误差
文献[2]	3	0.01
文献[3]	3	5.0137e-3
本方法	3	2.5271e-4

5　结论

通过新的目标函数的确定,能够把输出情况反映到模型的输入空间,使模型的输入输出映射空间具有逼近输出的能力,从而使模型的前件输入空间得到优化,使模型结构得到优化.从仿真的实例来看,这种方法能够使建模精度得到显著提高,能够辨识非线性系统.

参考文献:

[1]　王宏伟.基于模糊模型的模糊辨识方法及其应用研究[D].博士学位论文.哈尔滨工业大学,1999,7:1-5.

[2]　Sugeno.M.and Takahiro.Y.A fuzzy-logic-based approach to qualitative modeling[J].IEEE Trans on Fuzzy Systems.1993,(1):7-31.

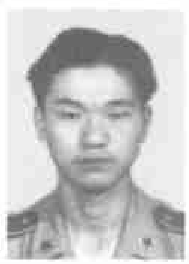
[3]　Q.Chen,Y.G.Xi,Z.J.Zhang.A clustering algorithm for fuzzy models identification[J].Fuzzy Sets and Systems.1998,(98):319-329.

[4]　王宏伟,马广富,王子才.一种基于模糊规则的模糊辨识方法[J].系统仿真学报.1998,10(4):61-64.

[5]　王宏伟,王清,王子才.基于模糊聚类多变量系统的模糊辨识方法[J].哈尔滨工业大学学报,1998,30(4):25-27.

作者简介:

王宏伟　1969 年生,1999 年获哈尔滨工业大学工学博士学位.现在国防科技大学机电工程与自动化学院流动站工作.



詹荣开　1977 年生,现为国防科技大学机电工程与自动化学院硕士研究生.

贺汉根　1942 年生,教授,博士生导师.现于国防科技大学机电工程与自动化学院工作.