

# 基于鲁棒模糊聚类的混沌时间序列预测\*

刘福才† 张彦柳 陈 超

(燕山大学电气工程学院自动化系 秦皇岛 066004)

(2007 年 9 月 27 日收到 2007 年 10 月 19 日收到修改稿)

采用一种基于鲁棒模糊聚类算法的模糊辨识方法,通过引入局部划分关联度因子,增强了系统辨识的抗干扰能力,提高了系统辨识的鲁棒性.首先用最近邻模糊聚类法划分初始输入空间,得到模糊规则数及初始聚类中心;然后用鲁棒模糊聚类算法求解并优化模糊隶属度和聚类中心,建立高精度的 T-S 模糊模型;最后利用最小二乘法辨识模型的初始结论参数,进一步利用带遗忘因子的递推最小二乘法优化结论参数.采用该方法对 Mackey-Glass 混沌时间序列进行建模和预测,仿真结果表明利用本方法可以进行准确建模和预测,验证了本方法的鲁棒性、有效性和实用性.

关键词:最近邻模糊聚类,鲁棒模糊聚类,混沌时间序列,最小二乘法

PACC: 0545

## 1. 引 言

非线性混沌时间序列的建模、预测与控制是当今学术界的研究热点.混沌时间序列预测在许多领域都有重要的研究意义,例如通过研究电力负荷的混沌特征,可将混沌时间序列预测算法应用于电力负荷预测<sup>[1]</sup>.

近几年,各种神经网络方法在混沌时间序列预测中得到了广泛的应用,但是神经网络预测方法普遍存在着效率低、预测精度较低等缺陷.张家树等采用少参数二阶 Volterra 滤波器及自适应高阶非线性滤波方法对混沌时间序列进行预测研究<sup>[2]</sup>;王宏伟等采用模糊竞争学习方法,提出一种基于模糊模型的混沌时间序列预测方法<sup>[3]</sup>;Li 等运用自适应方法对几类混沌系统进行了有效的控制和辨识<sup>[4]</sup>;刘涵等用基于模糊 sigmoid 核的支持向量机回归建模方法预测混沌时间序列<sup>[5]</sup>;谭文等运用神经网络对几类混沌系统进行了辨识和预测<sup>[6-10]</sup>.

模糊聚类分析是模糊建模的一种常用的有效的建模方法.模糊聚类算法的优点是可以直接得到输入划分(或乘积的模糊划分).Wang 提出的最近邻模糊聚类法<sup>[11]</sup>,使得聚类算法不再依赖于初始值,扩展了聚类方法的应用前景,本文借鉴其方法,并进行

了简化.Cai 等<sup>[12]</sup>提出了一种快速鲁棒模糊聚类算法,并且应用在图像分割方面,即以图像中某点的像素值为中心,与其他邻近的像素值组成一个  $3 \times 3$  或  $5 \times 5$  的窗口,通过引入局部划分关联度因子,平滑该中心像素值,该方法提高了图像分割的抗干扰能力,具有良好的分割效果.本文所采用的算法,首先用简化的最近邻模糊聚类法<sup>[11]</sup>确定模糊规则数和初始聚类中心;然后改进了 Cai 等提出的快速鲁棒模糊聚类算法,即通过引入一组局部划分关联度因子,对每组输入样本向量进行平滑操作,代替了图像分割中一组局部划分关联度因子只能平滑某一个像素点,得到平滑的输入样本向量后,再用模糊聚类法优化模糊隶属度和聚类中心,从而优化 Takagi-Sugeno(T-S)模糊模型结构;最后用最小二乘法辨识初始结论参数,进一步用带遗忘因子的递推最小二乘法优化结论参数,实现非线性系统的建模和预测.用本文提出的鲁棒模糊聚类法对 Mackey-Glass 混沌时间序列进行建模和预测试验,仿真结果验证本文方法的建模鲁棒性、有效性和实用性.

## 2. T-S 模糊模型

T-S 模型<sup>[13]</sup>是一种基于规则模型,其规则前提参数是模糊变量,而结论部分是输入输出线性函

\* 燕山大学博士基金(批准号: B111)资助的课题.

† E-mail: lfc\_xb@263.net

数,它以局部线性化为基础,通过模糊推理方法实现全局的非线性.T-S 模型一般定义为

$R_i$  若  $x_1$  是  $A_{i1}$ ,  $x_2$  是  $A_{i2}$  ...  $x_r$  是  $A_{ir}$ ,

则

$$y_i = p_{i0} + p_{i1}x_1 + \dots + p_{ir}x_r, \quad 1 \leq i \leq c, \quad (1)$$

式中  $R_i$  为第  $i$  个模糊规则,  $c$  是模糊规则数,  $x_j$  ( $1 \leq j \leq r$ ) 是第  $i$  个规则的第  $j$  个输入变量,  $y_i$  ( $1 \leq i \leq c$ ) 是第  $i$  个规则的输出变量,  $A_{ij}$  为前提模糊子集,  $p_{ij}$  是后件参数(即要辨识的结论参数).

当给定输入  $X = (x_1, x_2, \dots, x_r)^T \in U \subset R^r$  时, T-S 模糊模型推理输出计算如下:

$$y = \frac{\sum_{i=1}^c \omega_i y_i}{\sum_{i=1}^c \omega_i}, \quad (2)$$

$$\omega_i = \mu_{A_{i1}} \wedge \mu_{A_{i2}} \wedge \dots \wedge \mu_{A_{im}}, \quad (3)$$

式中  $\mu_{A_{ij}}$  是由前面的模糊划分得出的,  $\wedge$  符号是取小运算.

### 3. 基于鲁棒模糊聚类的模糊建模方法

本节给出模糊建模方法的详细描述.建模过程如图 1 所示:1)利用最近邻模糊聚类方法初始划分输入数据,得到模糊聚类模糊规则数和初始聚类中心;2)利用鲁棒模糊聚类方法优化模糊聚类中心和模糊隶属度;3)利用最小二乘法辨识模型的初始结论参数,并利用带遗忘因子的递推最小二乘法优化结论参数,实现 T-S 模型的在线学习.

#### 3.1. 输入空间初始划分

最近邻聚类法是一种有效的确定模糊规则数和初始聚类中心的算法.在此算法中,首先把第一个数据作为第一组的聚类中心.接下来,如果一个数据距该聚类中心的距离小于某个预期值,就把这个数据放到此组中,即该组的聚类中心应是和这个数据最接近的;否则,把该数据设为新一组的聚类中心.假设给定  $n$  个输入-输出数据对  $(X_k, y^k) \quad (k = 1, 2, \dots, n)$ ,  $X_k \in R^r$ , 并用  $v_i$  表示第  $i$  个聚类中心向量.算法如下:

步骤 1 令  $k = 1, m = 1$ , 并设定半径  $d$ , 把  $X_1$  设为一个聚类中心  $v_1, v_1 = X_1$ .

步骤 2  $k = k + 1$ .

步骤 3 假定考虑第  $k$  对输入-输出数据  $(X_k$ ;

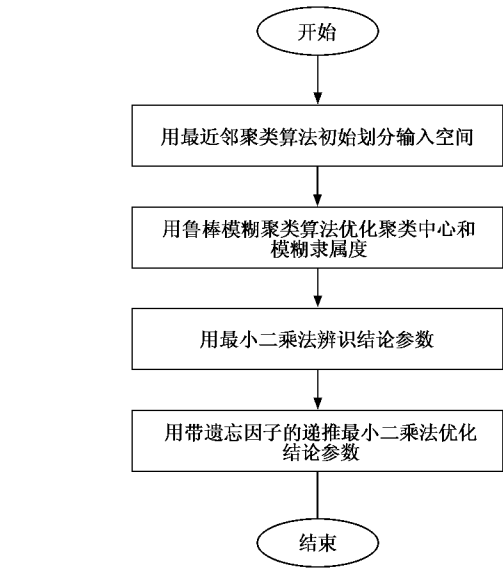


图 1 基于鲁棒模糊聚类算法的模糊建模流程图

$y^k) \quad (k = 2, 3, \dots)$  时,已经存在聚类中心分别为  $v_1, v_2, \dots, v_m$  的  $m$  个聚类.分别计算  $X_k$  到这  $m$  个聚类中心的距离  $\|X_k - v_l\| \quad (l = 1, 2, \dots, m)$ .设这些距离中最小的距离为  $\|X_k - v_{l_k}\|$ , 即  $v_{l_k}$  为  $X_k$  的最近邻聚类中心.

步骤 4 如果  $\|X_k - v_{l_k}\| < d$ , 则转步骤 5; 否则把  $X_k$  作为一个新的聚类中心,  $v_{m+1} = X_k, m = m + 1$ , 转步骤 5.

步骤 5 如果  $k \geq n$ , 则停止; 否则转步骤 2.

由此算法即得模糊规则数  $m$ , 并取  $c = m$ , 和聚类中心向量  $V = (v_1, v_2, \dots, v_c)$ .

#### 3.2. 优化输入空间模糊划分

鲁棒模糊聚类算法的原理是通过计算每组输入样本向量的内部数据间的局部划分关联度,平滑每组输入样本向量,提高了系统建模的抗干扰能力,从而提高了系统建模的鲁棒性和系统辨识精度.在文献 [12] 中,每引入一组局部划分关联度,可平滑某一个数据点,在本文算法中,可同时平滑向量的每一个点,以样本输入向量为单位进行聚类及优化,降低了处理每个数据点的耗时性,提高了系统辨识速度.本方法可同时满足高辨识精度和高辨识速度的要求.

鲁棒模糊聚类包括两部分:1)利用局部划分关联度因子平滑输入样本向量;2)利用模糊聚类算法优化模糊隶属度和聚类中心.

##### 3.2.1. 平滑输入样本向量

在本文所提算法的第二步中,通过引入局部划

分关联度因子平滑输入样本向量. 本文对每组输入样本向量内数据间引入了局部划分关联度  $R_{ij}$ , 以平滑每组输入样本向量. 对于第  $k$  个样本  $X_k$  其表达形式如下:

$$R_{ij}^{(k)} = \begin{cases} R_{s_{-ij}}^{(k)} \times R_{g_{-ij}}^{(k)}, & j \neq i, \\ 0, & i = j, \end{cases} \quad i = 1 \dots, r, j = 1 \dots, r, \quad (4)$$

其中

$$R_{s_{-ij}}^{(k)} = \exp\left(\frac{-|i-j|}{\lambda_s}\right), \quad i = 1 \dots, r, j = 1 \dots, r, \quad (5)$$

$$R_{g_{-ij}}^{(k)} = \exp\left(\frac{-\|x_i - x_j\|^2}{\lambda_s \times \delta_{s_{-i}}^2}\right), \quad i = 1 \dots, r, j = 1 \dots, r, \quad (6)$$

其中

$$\delta_{g_{-i}}^2 = \frac{\sum_{j \in r} \|x_i - x_j\|^2}{\beta} \quad (7)$$

可以提前计算出来.

$\lambda_s$  的参数值, 决定了  $R_{s_{-ij}}^{(k)} = \exp\left(\frac{-|i-j|}{\lambda_s}\right)$  的取值延伸范围, 亦影响了局部划分关联度因子的取

值延伸范围. 一般地,  $\lambda_s$  变化范围根据输入数据的个数进行确定, 可以取为 1, 2, ... 或  $r+1$ . 如图 2 所示, 分别表示  $\lambda_s = 1, 3$  和 6 时, 输入变量个数与  $R_s$  间的关系图.

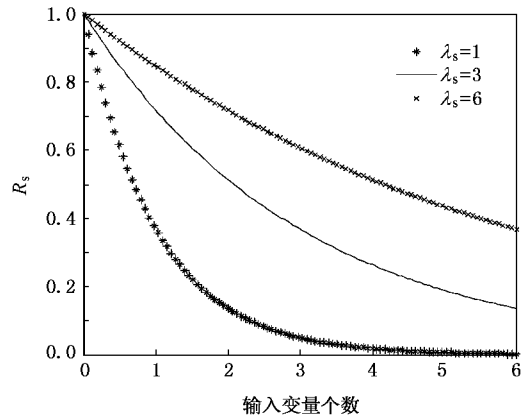


图 2  $R_s$  随输入变量个数变化的趋势

随机抽取混沌时间序列的某一样本向量, 例如随机选取样本向量  $X = (0.5929, 0.6390, 0.6879, 0.7397, 0.7942, 0.8511)$ , 则可根据(4)–(7)式得到对应于样本点  $X$  的  $R_{ij}$  值如表 1 所示.

表 1 样本输入点与其临近输入点间的  $R$  关系

样本	样本输入点与其临近输入点间的 $R$ 关系值					
	$\alpha(1)$	$\alpha(2)$	$\alpha(3)$	$\alpha(4)$	$\alpha(5)$	$\alpha(6)$
$\alpha(1)$	0	0.2937	0.0815	0.0214	0.0053	0.0013
$\alpha(2)$	0.2420	0	0.2258	0.0477	0.0095	0.0018
$\alpha(3)$	0.0279	0.1597	0	0.1465	0.0198	0.0025
$\alpha(4)$	0.0005	0.0056	0.0701	0	0.0622	0.0035
$\alpha(5)$	0.0000	0.0000	0.0000	0.0000	0	0.0048
$\alpha(6)$	0.0000	0.0000	0.0000	0.0000	0.0000	0

$\lambda_g$  的参数值, 决定了  $R_{g_{-ij}}^{(k)}$  的取值延伸范围, 其取值和作用与  $\lambda_s$  相似. 在此不再详述.

通过公式

$$\xi_i^{(k)} = \frac{\sum_{j \in r} R_{ij}^{(k)} \times x_j}{\sum_{j \in r} R_{ij}^{(k)}}, \quad i = 1, 2, \dots, r, k = 1, 2, \dots, m \quad (8)$$

修正原样本点, 并得到修正后的第  $k$  组样本向量为  $\xi^{(k)} = (\xi_1^{(k)}, \xi_2^{(k)}, \dots, \xi_r^{(k)})$ . 引入局部划分关联度因子后, 可以提高抗干扰能力, 从而提高系统建模鲁棒性和精度.

### 3.2.2. 优化聚类中心和模糊隶属度

本文所提算法的第一步中已经由最近邻聚类法, 得到聚类中心  $v_i$  和模糊规则数  $c$ .

模糊聚类方法的目标在于找到  $U \in [\mu_{ik}] \in M_{fc}$  和  $V = (v_1, \dots, v_c)$ , 使得

$$J = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m \| \xi_k - v_i \|^2 \quad (i = 1, 2, \dots, c; k = 1, 2, \dots, m) \quad (9)$$

最小, 其中  $m$  是一个加权指数.

首先利用下式计算  $\mu_{ik}$ :

$$\mu_{ik} = \left[ \sum_{j=1}^c \left( \frac{\xi_k - v_j}{\xi_k - v_i} \right)^{2(m-1)} \right]^{-1}, \quad i = 1, 2, \dots, c; k = 1, 2, \dots, m, \quad (10)$$

并根据下式更新  $v_i$  :

$$v_i = \frac{\sum_{k=1}^n \xi_k(\mu_{ik})^m}{\sum_{k=1}^n (\mu_{ik})^m}, \quad i = 1, 2, \dots, c. \quad (11)$$

由此得到鲁棒模糊聚类算法的完整算法,具体方法如下:

步骤 1 初始化参数. 给定  $c, m$  和  $\beta, c$  为算法第一步中由最近邻聚类算法得到的聚类数,  $m$  一般取 2. 选择  $v_i$  的初始值为第一步中得到的初始聚类中心向量, 并利用(10)式计算初始模糊隶属度  $U$ .

步骤 2 利用(4)-(7)式求解  $R_{ij}^{(k)}$ .

步骤 3 利用(8)式计算修正后的输入样本向量  $\xi^{(k)}$ .

步骤 4 利用(11)式更新  $v_i$ .

步骤 5 利用(10)式更新  $\mu_{ik}$ .

步骤 6 如果  $\|U_{new} - U_{old}\| \leq \epsilon$ ,  $\epsilon$  为阈值, 则  $k = k + 1$ , 并转步骤 2, 否则转步骤 4.

具体分为计算两个矩阵: 隶属函数矩阵  $U$  和聚类中心矩阵  $v$ . 通过仔细研究  $v_i$  的更新方程, 可以看出以下三点. 第一, 分子是样本数据点的加权和; 第二, 分母是隶属函数矩阵第  $i$  行的和; 第三, 有两个地方使用了  $\mu_{ik}$ , 即分子和公共分母的计算. 因为分子、分母都是和的形式, 所以在计算矩阵  $U$  的同时, 可以得到  $v_i$  的分子与分母. 从第三点可以看出, 一旦更新了行与列的和, 矩阵  $U$  的元素就没有用了, 不需要保存矩阵  $U$ . 考虑到这些情况, 在应用 Matlab 软件编程过程中, 可以节省时间.

### 3.3. 初始化及优化结论参数

当给定一组输入-输出数据  $x_{1i}, x_{2i}, \dots, x_{ri} \rightarrow y_i$  ( $i = 1, \dots, m$ ) 时, 定义

$$v_i = \omega_i / \sum_{i=1}^c \omega_i, \quad (12)$$

于是 T-S 模型的输出为

$$\begin{aligned} y &= \sum_{i=1}^c v_i y_i \\ &= \sum_{i=1}^c v_i (p_{i0} + p_{i1} x_1 + \dots + p_{ir} x_r) \\ &= [v_1 \dots v_c v_1 x_1 \dots v_c x_1 \dots v_1 x_r \dots v_c x_r] P. \end{aligned} \quad (13)$$

令  $X$  为  $n \times c(r+1)$  的矩阵,  $Y$  为  $n$  维向量,  $P$  为  $c(n+1)$  维向量:

$$Y = [y^1 \dots y^n],$$

$$P = [p_{10} \dots p_{c0} \quad p_{11} \dots p_{c1} \dots p_{1r} \dots p_{cr}]^T \quad (14)$$

$$X = [v_1 \dots v_c v_1 x_1 \dots v_c x_1 \dots v_1 x_r \dots v_c x_r], \quad (15)$$

则得到

$$Y = XP. \quad (16)$$

本文采用最小二乘估计辨识模型的结论参数  $P$ :

$$P^* = (X^T X)^{-1} X^T Y. \quad (17)$$

对模型规则的结论参数进行在线学习优化. 采用带遗忘因子的递推最小二乘法实现模型参数的在线学习, 即

$$P_{k+1} = P_k + K_{k+1} (y^{k+1} - x_{k+1}^T P_k^*), \quad (18)$$

其中

$$K_{k+1} = \frac{1}{\lambda} \left( K_k - \frac{K_k \cdot x_{k+1} \cdot x_{k+1}^T K_k}{\lambda + x_{k+1}^T K_k \cdot x_{k+1}} \right), \quad (19)$$

其中  $\lambda$  为遗忘因子,  $\lambda$  越小, 对新数据的学习能力越强; 当  $\lambda = 1$  时, 该算法转化为普通的递推最小二乘法.

本文提出的完整的模糊建模算法总结如下:

步骤 1 选定模糊建模型, 用最近邻聚类法初始划分输入数据, 得到模糊聚类的模糊规则数和初始聚类中心值.

步骤 2 平滑输入样本向量.

步骤 3 用模糊聚类算法优化模糊隶属度和聚类中心.

步骤 4 用最小二乘法辨识结论参数.

步骤 5 用带遗忘因子的递推最小二乘法在线优化结论参数.

步骤 6 用得到的结论参数进行在线预测.

在如下仿真试验中, 设置某些参数如  $d = 0.6$ ,  $\lambda_s = 3$ ,  $\lambda_g = 1$ ,  $\epsilon = 0.0001$ .

## 4. Mackey-Glass 混沌时间序列建模与预测

本文将 Mackey-Glass 混沌时间序列(混沌动力系统)建模与预测作为仿真实例, 以证明本文模糊建模方法的有效性. 混沌时间序列是通过下面的时滞微分方程

$$\dot{x} = \frac{\alpha x(t-\tau)}{1 + x^\gamma(t-\tau)} - \beta x(t) \quad (20)$$

得到的<sup>[14]</sup>.

$\tau$  是唯一可调的参数. 首先令  $\alpha = 0.2$ ,  $\beta = 0.1$ ,

$\gamma = 10$ . Mackey-Glass 方程为时滞参数  $\tau$  的函数. 当  $\tau > 17$  时, 系统呈混沌状态. 图 3 为  $\tau = 17$  时, 利用四阶龙格库塔法对该时滞微分方程即(20)式进行求解, 得到的混沌时间序列. 图 4 为  $\tau = 17$  时系统的相图. 利用四阶龙格库塔法对该时滞微分方程进行求解, 可以使输入数据更加逼近非线性系统. 对于 Mackey-Glass 系统预测的目的是根据  $t$  时刻以前的一组数据  $x(\cdot)$ , 去预测  $x(t + \Delta t)$ , 其中  $\Delta t$  为预测时间步幅, 预测方法是令  $\Delta t$  为时滞参数, 取  $D$  个点, 即  $x(t - (D - 1)\Delta t), \dots, x(t - \Delta t), x(t)$  去预测将来时刻  $x(t + \Delta t)$ , 若  $M = D - 1$  为整数, 仿真研究的任务是利用模糊模型构造函数, 即

$$y(t + \Delta t) = f(x(t), x(t - \Delta t), \dots, x(t - M\Delta t)), \quad (21)$$

式中  $f(\cdot)$  为  $M + 1$  维空间上映射,  $y(t + \Delta t)$  为模糊模型的输出,  $x$  为模糊模型的输入量, 对应的模型输出取为  $y(t + \Delta t) = f(x(t), x(t - \Delta t), \dots, x(t - M\Delta t))$  并取  $x(t + \Delta t) = y(t + \Delta t)$ . 选取  $\tau = 17, \Delta t = 1$ , 即选取 1000 对样本数据  $[x(t - 1), x(t - 2), x(t - 3), x(t - 4), x(t - 5), x(t - 6); x(t)]$ , 其中  $t = 19, 20, \dots, 1018$ , 其中前 6 个变量数据作为输入, 最后一个变量数据作为输出.

利用最近邻聚类方法得到模糊规则数  $c$  和模糊聚类中心值  $v$  后, 利用鲁棒模糊聚类方法, 对采样数据的前 500 组当作训练数据, 其余 500 组作为测试数据以验证模糊模型的有效性.

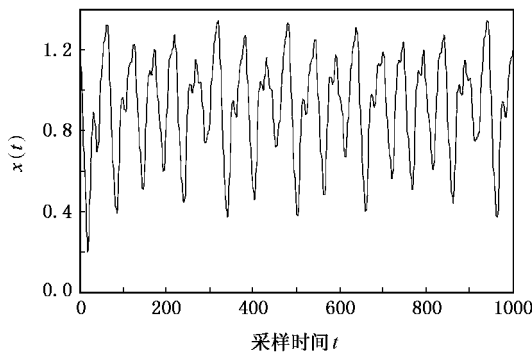


图 3 Mackey-Glass 混沌时间序列图

采用本文方法经过学习后, 模糊规则数为 4, 对图 3 的混沌时间序列进行建模与预测. 先用前 500 组训练数据建立模糊规则并形成混沌动力系统的模糊模型, 使系统性能指标达到期望值; 然后采用后 500 组数据进行预测估计. 图 5 到图 8 给出了仿真结果(仿真环境: MATLAB 6.5, 计算机主频 2.4GHz, 内存 512MB).

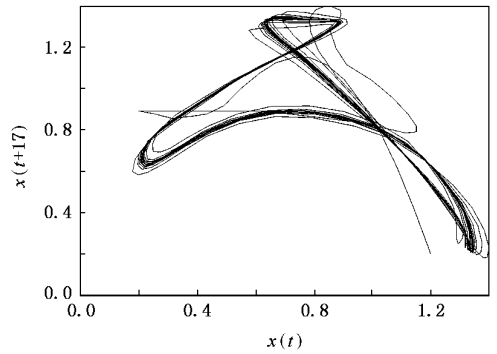


图 4 混沌系统的相图(利用四阶龙格库塔法求解混沌动力系统)

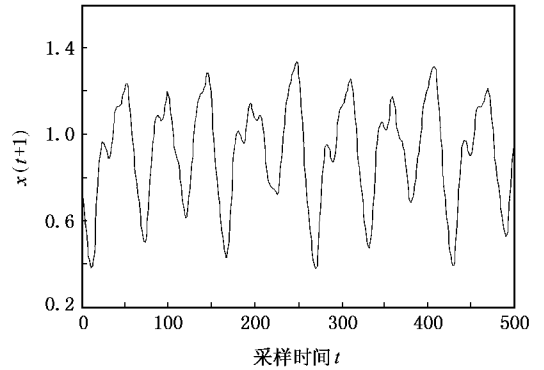


图 5 训练数据模糊模型输出曲线(实线)与预测曲线(虚线)

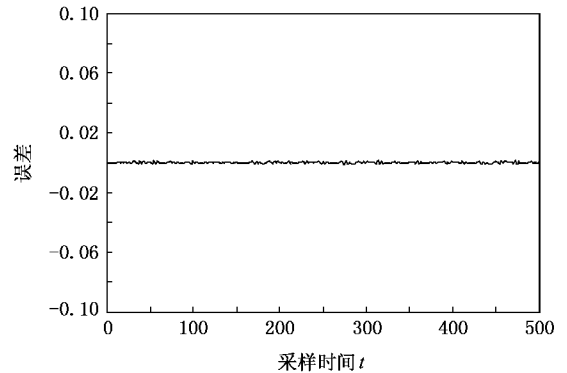


图 6 训练数据预测误差曲线

表 2 给出了鲁棒模糊聚类方法训练数据时模糊聚类前后得到的收敛中心向量.

表 3 给出了鲁棒模糊聚类方法训练数据时模糊聚类得到的结论参数向量.

由仿真结果可以看出, 本文采用的鲁棒模糊聚类建模方法能够对混沌动力系统进行有效的建模与预测.

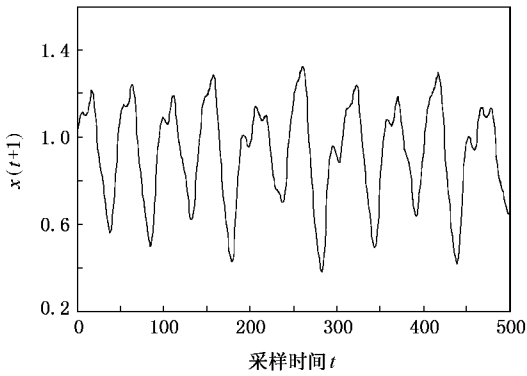


图7 在线检验数据的模糊模型输出曲线(实线)与预测曲线(虚线)

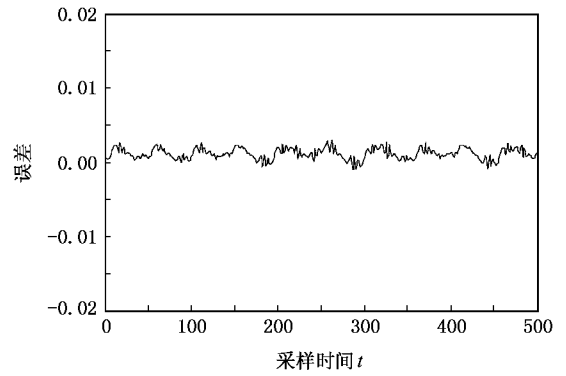


图8 在线检验数据预测误差曲线

表2 鲁棒模糊聚类模糊模型的建模聚类中心值

规则	初始中心向量	优化后中心向量
1	(0.7942 0.8511 0.9103 0.9717 1.0347 1.0984)	(0.8862 0.8862 0.8868 0.8874 0.8875 0.8875)
2	(0.5499 0.5929 0.6390 0.6879 0.7397 0.7942)	(0.7237 0.7232 0.7235 0.7241 0.7243 0.7243)
3	(0.3884 0.3842 0.3941 0.4140 0.4412 0.4735)	(0.5680 0.5681 0.5675 0.5669 0.5667 0.5667)
4	(1.1898 1.1728 1.1575 1.1453 1.1370 1.1326)	(0.8454 0.8451 0.8454 0.8457 0.8457 0.8457)

表3 模糊模型的结论参数向量

规则	结论参数向量
1	[0.0038 4.7841, -10.2226 12.6373, -9.6160 4.2898, -0.8765]
2	[-0.0027 4.8631, -10.4119 12.5705, -8.9962 3.5949, -0.6173]
3	[0.0125 4.3279, -8.6882 10.4289, -7.9457 3.6535, -0.7943]
4	[0.0008 5.0292, -10.9428 13.2965, -9.5937 3.9205, -0.7108]

## 5. 结 论

本文采用基于最近邻模糊聚类和鲁棒模糊聚类

的算法,通过引入局部划分关联度因子,平滑每组输入样本向量,提高了系统建模的鲁棒性、辨识精度和辨识速度。通过一系列的步骤优化 T-S 模糊模型结构,描述动态系统的特性,是一种有效的辨识方法。本文提出的方法不仅克服了聚类中心的任意初始化只能得到一个局部最优的结果,而且通过平滑输入样本向量提高了系统建模鲁棒性,并通过优化算法得到最优的聚类中心,最后用最小二乘法和带遗忘因子的递推最小二乘算法辨识模型结论参数向量,为复杂非线性系统辨识、混沌时间序列预测和控制提供了一条实用有效的途径。

[1] Jiang C W, Yuan Z Q, Hou Z J 2004 *Power System Technology* **28** 25 (in Chinese) [蒋传文、袁智强、侯志俭 2004 电网技术 **28** 25]

[2] Zhang J S, Xiao X C 2001 *Acta Phys. Sin.* **50** 1248 (in Chinese) [张家树、肖先赐 2001 物理学报 **50** 1248]

[3] Wang H W, Ma G F 2004 *Acta Phys. Sin.* **53** 3293 (in Chinese) [王宏伟、马广富 2004 物理学报 **53** 3293]

[4] Li Z H 2001 *Chin. Phys.* **10** 494

[5] Liu H, Liu D 2006 *Control Theory & Applications* **23** 204 (in Chinese) [刘 涵、刘 丁 2006 控制理论与应用 **23** 204]

[6] Tan W, Wang Y N, Liu Z R, Zhou S W 2002 *Acta Phys. Sin.* **51** 2463 (in Chinese) [谭 文、王耀南、刘祖润、周少武 2002 物理学报 **51** 2463]

[7] Zhang S, Liu H X, Gao D T, Du S D 2003 *Chin. Phys.* **12** 594

[8] Guan X P, Tang Y G, Fan Z P, Wang Y Q 2001 *Acta Phys. Sin.* **50** 2112 (in Chinese) [关新平、唐英干、范正平、王益群 2001 物理学报 **50** 2112]

[9] Liu D, Ren H P, Kong Z Q 2003 *Acta Phys. Sin.* **52** 531 (in Chinese) [刘 丁、任海鹏、孔志强 2003 物理学报 **52** 531]

[10] Zhang J S, Xiao X C 2000 *Chin. Phys. Lett.* **17** 88

[11] Wang L X 1994 *Adaptive Fuzzy Systems and Control* (Prentice-Hall, Englewood Cliffs NJ)

[12] Cai W L, Chen S C, Zhang D Q 2007 *The Journal of the Pattern Recognition* **40** 825

[13] Takagi T, Sugeno M 1985 *IEEE Trans. Syst., Man, Cybern.* **15** 116

[14] Tan W, Wang Y N 2003 *Acta Phys. Sin.* **52** 795 (in Chinese) [谭文、王耀南 2003 物理学报 **52** 795]

# Prediction of chaotic time series based on robust fuzzy clustering<sup>\*</sup>

Liu Fu-Cai<sup>†</sup> Zhang Yan-Liu Chen Chao

( Department of Automation , Institute of Electrical Engineering , Yanshan University , Qinhuangdao 066004 , China )

( Received 27 September 2007 ; revised manuscript received 19 October 2007 )

## Abstract

We propose a new method for fuzzy modeling based on a robust fuzzy-clustering algorithm. The induced local spatial similarity improved the system's robustness to noise and outsider and predicted the robustness of the modeling system. Starting from an initial fuzzy partition of input space by a nearest-neighbor clustering method to get the number of rules and the initial clustering center, we can compute and optimize the fuzzy membership and the clustering center with a robust fuzzy-clustering algorithm and get the high precision T-S model. The obtained parameters were identified by the least square method and further optimized by selective recursive least square. The proposed method was applied to simulations of chaotic Mackey-Glass time series modeling and prediction. The results demonstrated the robustness, effectiveness and practicability of the method.

**Keywords** : nearest neighbor clustering , robust fuzzy-clustering , chaotic time series , least square method

**PACC** : 0545

---

<sup>\*</sup> Project supported by the Doctor Function of Yanshan University ( Grant No. B111 ).

<sup>†</sup> E-mail : lfc\_xb@263.net