

一种 T-S 模型的在线辨识算法

钱富才^{1,2}, 伍光宇¹

(1. 西安理工大学 自动化与信息工程学院, 西安 710048; 2. 西安工业大学
新型网络与检测控制国家地方联合工程实验室, 西安 710032)

摘要: 提出一种能通过输入输出数据在线获得 T-S 模型的结构和参数的辨识算法. 首先, 对输入空间进行划分, 并在线优化子空间的形状和个数; 然后, 通过 RLS 更新子模型参数, 使各个子模型逼近当前工况的实际系统; 当子空间生成或形状发生变化时, 调整相应子模型参数和数据矩阵; 最后, 针对非线性动态系统和煤气炉数据进行仿真实验, 验证了所提出算法的有效性.

关键词: T-S 模型; 在线辨识; 子空间; 形状; 数据矩阵

中图分类号: TP273

文献标志码: A

An on-line algorithm for T-S model identification

QIAN Fu-cai^{1,2}, WU Guang-yu¹

(1. School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China; 2. The National-Local Joint Engineering Laboratory for New Network and Detection Control, Xi'an Technological University, Xi'an 710032, China. Correspondent: WU Guang-yu, E-mail: mebest21@163.com)

Abstract: An on-line algorithm for T-S model identification is proposed, which automates the structure and parameter identification simultaneously based on input-target samples. Firstly, several subspaces are produced in the input space, and their shape and distribution are optimized on-line. Then the recursive least square algorithm is employed to update the parameters of each local model, which makes each local model approximate the real system. If a new subspace is produced or the shape of the subspace is changed, an adjustment method is used for each local parameters and data matrix. Finally, a simulation experiment on a dynamic nonlinear system and the gas furnace data is given to verify the effectiveness of the proposed method.

Keywords: T-S model; online identification; subspace; shape; data matrix

0 引言

对于强非线性、多变量和强耦合的复杂工业对象, 建立反映控制对象的全局模型十分困难. 而模糊建模以局部线性化为基础, 通过模糊推理实现系统全局的非线性, 是一种有效的建模方法. 文献[1]证明了 T-S 模糊广义系统是一个通用逼近器, 它能以任意精度逼近在紧集上定义连续函数. T-S 模糊系统具有简单的模型结构和强大的非线性逼近能力, 非常适用于建立非线性系统的近似数学模型. T-S 模糊建模通常分为隶属度函数的个数和形状确定^[2]以及子模型参数的估计^[3]两个部分.

实际的工业过程易出现新的工况, 对象的特性也

易发生变化, 离线辨识很难自适应地反映系统的非线性和不确定性^[4]. 为了实现在线辨识, 文献[5]通过递推 FCM 算法获得模型的前件结构, 然而该算法仍要预先给定初始聚类的中心和个数, 聚类结果对初值很敏感. 文献[6]通过输入数据在线获得聚类的个数和形状, 并在线估计子模型参数. 由于没有考虑聚类过程中聚类分布变化对子模型参数辨识的影响, 导致建模过程误差波动较大, 同时降低了模型的逼近能力. 文献[7]提出了自组织模糊神经网络算法, 然而该算法易产生大量的相似聚类, 对模型逼近效果贡献不大, 反而使模型结构变得复杂. 为了在线获得简洁的模型结构, 本文通过输入数据在线生成和调整聚类形状,

收稿日期: 2013-10-23; 修回日期: 2014-03-14.

基金项目: 国家自然科学基金项目(61273127); 高等学校博士学科点专项科研基金项目(20116118110008).

作者简介: 钱富才(1963—), 男, 教授, 博士生导师, 从事随机系统、最优控制等研究; 伍光宇(1989—), 男, 硕士生, 从事先进控制理论与应用的研究.

删减相似聚类,从而优化输入空间.

文献 [6] 通过局部递推最小二乘法在线辨识子模型参数,然而其子模型参数不能随着聚类形状的改变做出调整. 本文在聚类的形状和个数发生变化时,利用数据信息对子模型参数及数据矩阵进行相应地调整,这样便降低了因子模型参数初值不准确所带来的逼近误差,从而提高在线建模的精度.

1 T-S 模型的描述

T-S 模型的结论部分采用线性函数描述,并通过隶属函数将它们光滑地连成一个全局非线性函数,具体描述如下:

Rule i : If x_1 is A_1^i and \dots x_n is A_n^i ,

$$\text{Then } y_i = p_0^i + p_1^i x_1 + \dots + p_n^i x_n. \quad (1)$$

其中: $i = 1, 2, \dots, r$; x_i 为输入变量; A_j^i 为第 i 条规则的第 j 个模糊集; y_i 为第 i 条规则的输出; p_j^i 为结论部分的权值. A_j^i 的隶属度函数取如下高斯函数:

$$A_j^i(x_j) = \exp\left(\frac{-\|x_j - m_j^i\|^2}{2\sigma_j^i}\right), \quad (2)$$

其中 m_j^i 和 σ_j^i 分别为高斯函数的中心和宽度. 若输入量采用单点模糊集合的模糊化方法,则对于给定的输入 $X = (x_1, x_2, \dots, x_n)$, 其对于规则 i 的激励强度为

$$\omega_i(X) = \prod_{j=1}^n A_{ij}(x_j), \quad (3)$$

这里采用乘积推理运算.

模糊系统的输出量为每条规则输出量的加权平均,即

$$y = \frac{\sum_{i=1}^r \omega_i y_i}{\sum_{i=1}^r \omega_i} = \sum_{i=1}^r \bar{\omega}_i y_i, \quad (4)$$

$$\bar{\omega}_i = \omega_i / \sum_{i=1}^r \omega_i. \quad (5)$$

2 模糊空间的在线划分

通过在线聚类,可将模糊空间划分成若干个子空间,具体算法如下.

Step 1 在出现第 1 个样本 $X^1 = (x_1^1, x_2^1, \dots, x_n^1)$ 时,产生第 1 条规则,其隶属度函数的中心和宽度为

$$m_1^j = x_j^1, \quad (6)$$

$$\sigma_1^j = \sigma_0. \quad (7)$$

其中: $j = 1, 2, \dots, n$; 子空间个数 $r(t) = 1$; 子空间内样本个数 $L_i = 1$.

Step 2 当 $n = 2$ 时,输入空间为二维空间. 定义激励强度阈值 F_{in} , 则 $\omega_i \geq F_{in}$, 即

$$\frac{(x_1 - m_1^i)^2}{\sigma_1^i} + \frac{(x_2 - m_2^i)^2}{\sigma_2^i} \leq -2 \ln(F_{in}) \quad (8)$$

在输入空间中是一个椭圆区域,特别当 $F_{in} = 0.1354$ 时,第 i 个椭圆子空间的长轴为 $2\sigma_1^i$, 短轴为 $2\sigma_2^i$.

对于每次采样得到的输入样本 X^q , 计算其对每条规则激励强度的最大值

$$I = \arg \max_{1 \leq i \leq r(t)} \omega_i(X^q). \quad (9)$$

若 $I \geq F_{in}$, 表明有子空间包含此样本, 则不产生新规则, 数据个数

$$L_i = L_i + 1; \quad (10)$$

若 $I < F_{in}$, 表明此样本在所有子空间以外, 则产生一个新的子空间

$$r(t) = r(t) + 1. \quad (11)$$

Step 3 若子空间内的样本个数 $L_i < \alpha$, α 为子空间数据量阈值, 表明该子空间中连续出现的样本很少, 则删除此规则, 规则数

$$r(t) = r(t) - 1. \quad (12)$$

这样, 当样本中出现孤立的样本时, 不会对其单独聚类, 避免了小聚类的产生.

Step 4 由于产生的子空间往往与其他子空间的重叠区域过大, 在相邻的交叠子空间的模型差异较明显时, 重叠区域模型输出与真实的样本输出误差较大. 因此, 在产生新子空间之前利用输入数据信息来调整当前子空间的形状. 调整时, 计算当前子空间内每个样本的密度值, 形式为^[6]

$$\rho(\chi^i) = \frac{1}{1 + \frac{1}{L_i - 1} \sum_{j=1}^{L_i} (\chi^i - \chi^j)^T (\chi^i - \chi^j)} = \frac{1}{(L_i - 1)(1 + \chi^{iT} \chi^i) - 2\chi^i \sum_{j=1}^{L_i} \chi^{jT} + \sum_{j=1}^{L_i} \chi^{jT} \chi^j}. \quad (13)$$

$$\rho(\chi^\zeta) = \arg \max_{1 \leq i \leq L} \rho(\chi^i). \quad (14)$$

其中: $i = 1, 2, \dots, L$; $\chi(i)$ 为子空间中的样本; χ^ζ 为密度值最大的样本, 并将它作为子空间新的中心, 使各个输入点到其子空间中心的距离尽可能小, 有

$$m_j^{r(t)} = \chi_j^\zeta. \quad (15)$$

新的子空间应能包含原子空间中的所有样本, 新的子空间半径通过以下的迭代方法获得:

$$\sigma_j^{r(t)} = \arg \min_{1 \leq i \leq L} (|\chi_j^\zeta - \chi_j^i|), \quad (16)$$

$$\sigma_j^{r(t)} = (1 + \eta) \sigma_j^{r(t)}, \quad (17)$$

其中 η 为迭代步长, 是介于 0~1 之间的小数.

Step 5 由于输入空间的划分情况发生了变化,对于新样本,返回Step 2重新判断是否产生新规则.若 $I < F_{in}$,则产生新规则的前件参数如下:

$$m_j^{r(t)+1} = x_j^q, \quad (18)$$

$$\sigma_j^{r(t)+1} = \beta_j \cdot \arg \min_{1 \leq i \leq r(t)} (|x_j^q - m_j^i|), \quad (19)$$

$$r(t) = r(t) + 1. \quad (20)$$

其中: $j = 1, 2, \dots, n$; β_j 为产生的新类与其近邻类的重叠系数.

当产生的新类在某个方向上与其最近邻类十分靠近时,产生的新类宽度会很小,若不处理,则输入空间最终会出现许多小聚类,因此有

$$\text{If } \sigma_j^{r(t)+1} < \delta, \quad (21)$$

$$\text{Then } \sigma_j^{r(t)+1} = \sigma_0, \quad (22)$$

其中 δ 为一个很小的值.

Step 6 为了使子空间的个数足够少,对距离较近的进行合并,即

$$\text{If } \sum_{j=1}^n (m_j^i - m_j^{i-t})^2 < \gamma, \quad (23)$$

$$\text{Then } m_j^i = \frac{m_j^i + m_j^{i-t}}{2}, \quad (24)$$

$$\sigma_j^i = \frac{|m_j^i - m_j^{i-t}| + \sigma_j^i + \sigma_j^{i-t}}{2}. \quad (25)$$

其中: $t = 1, 2, \dots, r(t) - 1$, 根据输入空间的范围选择合适的 γ , 通常取输入空间长度的 $0.1 \sim 0.5$ 倍.

Step 7 在聚类的过程中,对于增多的离群点,当其处于近邻类的边缘区域时,即当

$$\arg \min_{1 \leq i \leq r(t)} (X - m^i)(X - m^i)^T < \tau \sum_{j=1}^n \sigma_j^i{}^2 \quad (26)$$

时,通过扩大近邻类的半径来覆盖此离群点.

3 子模型的在线辨识

对于各个子空间建立子模型,使用最小二乘法分别辨识子模型参数.文献[6]以子模型的输出与实际系统的局部输出的误差平方和为代价函数,并将样本的激励强度作为权值,第 i 个子模型的性能指标为

$$J_i = \frac{1}{L} \sum_{q=1}^L [\bar{\omega}_i^q (y^q - h^q \theta_i)]^2, \quad (27)$$

$$\theta_i = (p_0^i, p_1^i, \dots, p_n^i), \quad (28)$$

$$h^q = (1, x_1^q, \dots, x_n^q). \quad (29)$$

其中: θ_i 为参数向量, h^q 为数据向量.

子模型参数可通过递推最小二乘法求得,但当新的子空间产生或其形状发生改变时,全局模型的结构也发生变化,因此各子模型的输出权重 $\bar{\omega}_i^q$ 也发生变化.由于子模型的参数修正量与 $\bar{\omega}_i^q$ 有关,之前辨识得

到的子模型参数不再适合原先的样本,对于同样的样本,全局模型的输出不同.这样,当子空间的数目逐渐增多时,辨识的精度会受到较大影响,一些情况下需要重复学习.

为了避免上述情况,本文将实际系统与子模型间的输出误差平方和作为准则函数,具体形式为

$$J_i = \frac{1}{L} \sum_{q=1}^L (y^q - h^q \theta_i)^2. \quad (30)$$

在辨识过程中,只让样本点对应子空间的子模型逼近样本真实输出;最后,将所有子模型的输出进行模糊加权求和作为全局输出.第 i 个子模型参数的递推最小二乘估计值为

$$\theta_i(q+1) = \theta_i(q) + K(q+1)(y^{q+1} - h(q+1)\theta_i(q)), \quad (31)$$

$$K(q+1) = \frac{P(q)h(q+1)^T}{h(q+1)P(q)h(q+1)^T + 1}, \quad (32)$$

$$P(q+1) = (I - K(q+1)h(q+1))P(q). \quad (33)$$

其中: $\theta_i(q)$ 为 q 时刻的参数估计值, $K(q)$ 为增益矩阵.递推求解过程中的数据向量 $h(q)$ 仅与属于其相应子空间的样本数据有关,这样便减少了计算量.由于最小二乘算法的收敛性,随着时间的增加,模型误差和子模型参数误差均会趋于 0. 子模型辨识的步骤如下.

Step 1: 初始化第 1 条规则的子模型参数 $p_j^1 = 0.01$, 矩阵参数 $P(1) = 10^2 I$, 这里 I 为单位阵.

Step 2: 对于任意一组样本 (X^q, y^q) , 若其属于某个子空间,即 $\bar{\omega}_i^q \geq F_{in}$, 则利用式(31)~(33)对子模型 i 的参数进行递推估计.

Step 3: 当产生新规则时,为了让产生新规则后的全局模型尽可能逼近之前的全局模型,需调用当前类中样本,利用一次完成算法估计新规则的子模型参数和数据矩阵.对于数据矩阵奇异的情况,可采用 SVD 分解算法[8].同样,当聚类发生合并时,以已有类的参数为初值,一次估计合并后类的参数和矩阵.

Step 4: 计算全局模型输出与实际系统的输出误差为

$$e = y - \hat{y}. \quad (34)$$

ε 需根据要求设定.当

$$|e| \leq \varepsilon \quad (35)$$

时可停止参数修正.

4 仿真实验

下面通过两个仿真实验验证本文算法的有效性,

并与其他算法进行对比. 辨识精度指标采用均方误差求出, 即

$$\text{MSE} = \frac{1}{N} \sum_{q=1}^N (y^q - \hat{y}^q)^2. \quad (36)$$

其中: N 为数据长度, \hat{y}^q 为第 q 个样本的全局模型输出.

4.1 非线性动态系统的辨识

考虑如下非线性离散动态系统:

$$y(k+1) = \frac{y(k)y(k-1)[y(k)+2.5]}{1+y^2(k)+y^2(k-1)} + u(k),$$

$$y(0) = 0, y(1) = 0,$$

$$u(k) = \sin(\pi k/10).$$

选取辨识模型

$$y(k+1) = f(u(k), y(k)).$$

其中: k 为 $[0, 20]$ 间均匀分布的随机样本, 输入 $u \in [-1, 1]$. 初始宽度 $b_1^{\text{init}} = b_2^{\text{init}} = 0.6$, 数据量阈值 $\alpha = 6$, 激励强度阈值 $F_{\text{in}} = 0.12$, 精度 $\varepsilon = 0.001$, 重叠系数 $\beta_1 = 0.4, \beta_2 = 0.55$. 相近类距离 $\gamma = 1, \tau = 2$.

图 1 给出了输入空间的划分结果, 图 2 为模型输出与实际输出的效果对比, 辨识精度 $\text{MSE} = 0.0207$. 由图 1 可见, 本文算法能合理地划分输入空间, 避免相似聚类的产生, 模型的结构基本覆盖了所有工况. 由图 2 可见: 模型的输出值与函数的期望值基本重合, 误差范围很小; 随着聚类数目的增多, 发生了相近聚类的合并, 从而改善了模型的结构和逼近效果.

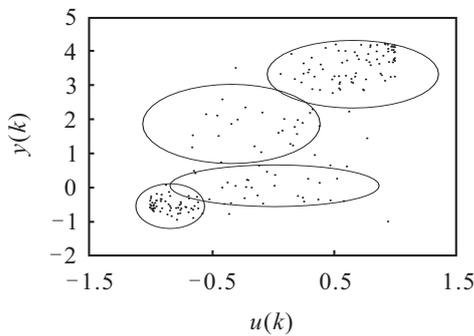


图 1 输入空间划分结果

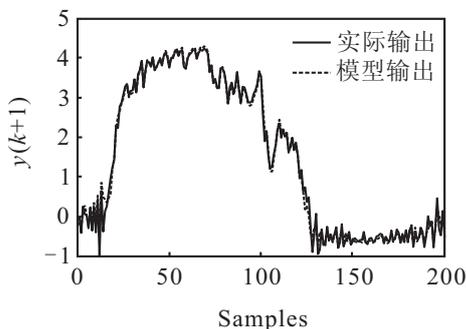


图 2 模型输出与实际输出

选取三输入的辨识模型为

$$y(k+1) = f(u(k), u(k-1), y(k)),$$

$\varepsilon = 0.028, \gamma = 5$. 表 1 给出了本文算法与其他模型的辨识结果对比情况. 由表 1 可以看出, 对于同样的样本数据, 本文算法用较少的规则数目和子模型的参数得到了较高的拟合精度.

表 1 非线性动态系统辨识结果对比情况

模型名称	参数个数	规则数	均方根误差
DFNN	48	6	0.0283
GDFNN	48	6	0.0241
SOFNN	44	5	0.0157
本文模型	30	3	0.0017

4.2 煤气炉数据的辨识

Box-Jenkins 的煤气炉数据常作为检验辨识系统的标准实验数据被许多文献采用, 其数据由 296 对输入输出测量值组成. 输入量 $u(t)$ 为煤气流量, 输出量 $y(t)$ 为二氧化碳浓度. 选取辨识模型 1 为

$$y(k+1) = f(u(k), y(k)), b_1^{\text{init}} = b_2^{\text{init}} = 1.5,$$

$$\alpha = 6, F_{\text{in}} = 0.13, \beta_1 = \beta_2 = 0.7.$$

输入空间的划分结果见图 3, SOFNN 算法^[9]输入空间的划分结果如图 4 所示. 由图 4 可以看出, SOFNN 算法对输入空间的划分很不合理, 子空间之间重叠严重, 且半径过大, 两个已超过输入空间的范围, 使子模型不能很好地描述实际系统, 导致在线辨识精度下降. 图 5 为模型输出与实际输出的效果对比, $\text{MSE} = 0.3367$.

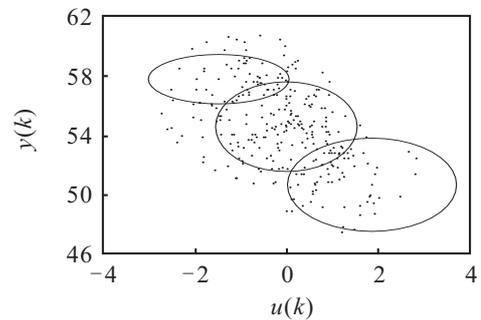


图 3 输入空间划分结果

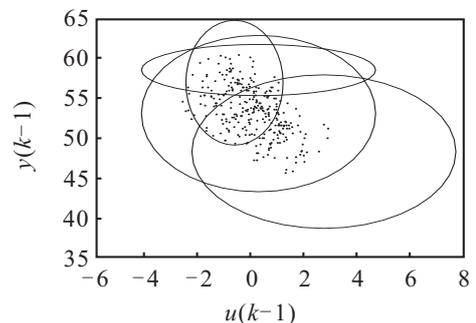


图 4 SOFNN 输入空间划分结果

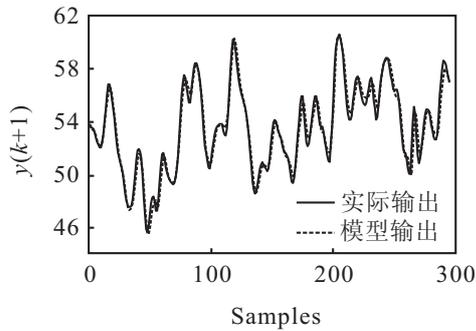


图 5 模型输出与实际输出对比

采用 $y_i = p_1^i u(k-1) + p_2^i u(k) + p_3^i y(k-1) + p_4^i y(k)$ 作为子模型, 用前 190 组样本进行建模得到模型 2, 再用所有样本进行测试, $MSE = 0.1222$. 选取三输入的辨识模型 3 为 $y(k+1) = f(u(k), u(k+1), y(k))$. 表 2 给出了本文算法与早期研究工作的性能对比情况. 从表 2 可以看出, 对于同样的样本数据, 本文算法用较少的规则得到了较高的拟合精度.

表 2 煤气炉数据辨识结果对比

模型名称	输入变量数	规则数	均方误差
Tong	2	18	0.469
Sugeno and Yasukawa	3	6	0.3367
本文模型 1	2	3	0.190
本文模型 2	3	3	0.0801
本文模型 3	2	3	0.1154

5 结 论

本文提出了一种新的 T-S 模型在线辨识算法, 该算法不仅能适应变化的工况, 而且充分考虑了子空间变化对参数辨识的影响, 解决了子模型参数辨识过程中参数初值不准确的问题. 仿真实验表明, 所提出的算法对非线性系统具有良好的逼近能力, 且模型结构简洁, 有较好的可解释性.

参考文献(References)

[1] 马俊峰, 张庆灵. T-S模糊广义系统的逼近性[J]. 控制理论与应用, 2008, 25(5): 837-844.

(Ma J F, Zhang Q L. Approximation property of T-S fuzzy singular systems[J]. Control Theory & Applications, 2008, 25(5): 837-844.)

[2] Angelov P P, Filev D P. An approach to online identification of Takagi-Sugeno fuzzy models[J]. IEEE Trans on Systems, Man, and Cybernetics, 2004, 34(1): 484-498.

[3] Cheng Wei-Yuan, Juang Chia-Feng. An incremental support vector machine-trained TS-type fuzzy system for online classification problems[J]. Fuzzy Sets and Systems, 2011, 163(1): 24-44.

[4] Xue Zhen-Kuang, Li Shao-Yuan. A multi-model identification algorithm based on weighted cost function and application in thermal process[J]. Acta Automatica Sinica, 2005, 31(3): 470-474.

[5] Dejan Dovan, Igor Krjanc. Recursive fuzzy c-means clustering for recursive fuzzy identification of time-varying processes[J]. ISA Trans, 2011, 50(2): 159-169.

[6] Kian Hong Quah, Chai Quek. FISK: Online local learning with genetic fuzzy input Takagi-Sugeno-Kang fuzzy framework for nonlinear system estimation[J]. IEEE Trans on Systems, Man, and Cybernetics, 2006, 36(1): 166-178.

[7] Juang Chia-Feng, Lin Chin-Teng. An on-line self constructing neural fuzzy inference network and its applications[J]. IEEE Trans on Fuzzy Systems, 1998, 6(1): 12-32.

[8] John Yen, Liang Wang, Charles Wayne Gillespie. Improving the interpretability of TSK fuzzy models by combining global learning and local learning[J]. IEEE Trans on Fuzzy Systems, 1998, 6(4): 530-537.

[9] Gang Leng, Thomas Marthin McGinnity, Girigesh Prasad. An approach for on-line extraction of fuzzy rules using a self-organizing fuzzy neural network[J]. Fuzzy Sets and Systems, 2005, 150(2): 211-243.

(责任编辑: 滕 蓉)