第23卷 第4期　　自 动 化 学 报　　Vol. 23, No. 4
1997年7月　　ACTA AUTOMATICA SINICA　　Jul, 1997

短 文

模糊规则的学习及其在非线性系统建模中的应用[1)]

陈建勤　席裕庚　张钟俊

（上海交通大学自动化研究所　上海　200030）

摘　要　探讨用神经网络的学习算法及模糊推理方法为非线性系统建模的问题．给出了学习模糊规则的新算法．这个算法首先用竞争学习为训练样本的输入空间进行聚类，然后为其确定区域划分边界，并按样本输入区域学习模糊规则．文中对于模糊规则提出了相应的模糊推理算法．并用算例验证了本文算法的有效性．

关键词　模糊系统，竞争学习，非线性系统建模，神经网络．

1　引　言

用神经网络和模糊系统研究非线性函数逼近、非线性系统建模、信号处理、模式识别等在国内外已成为一个重要的研究领域．目前用神经网络学习模糊系统从而建立非线性模型主要存在如下问题：用神经网络学习模糊规则普遍存在过拟合(overfitting)和泛化性(generalization)之间的矛盾[1]；重复率较差，即对于同一模型，其学习结果与样本的随机性有关；为了达到要求精度，需要对输入空间划分相当数目的区域，当输入变量较多时，划分区域将以指数增长[2]．

针对上述问题，本文采用神经网络的竞争学习算法、模糊边界划分、以及递推最小二乘法，实现训练样本输入空间的区域划分及规则学习，并用自适应模糊推理实现输入输出映射．

2　模糊规则及其推理方法

对于非线性系统建模问题，模糊系统的输入和输出处理的都是精确数据，所以，一般来说，模糊系统可以简化为模糊规则表达及推理两部分[1,3]，即

模糊规则：if x is v_i, Δ_i, then $y_i=u_i+\theta_i^{T}(x-v_i)$, $i=1,2,\cdots,q$.

---

1)　国家自然科学基金和中国博士后基金资助项目．

收稿日期　1996-01-08

模糊推理：$y=\frac{\sum_i s_i y_i}{\sum_i s_i}, s_i=\begin{cases}1-\frac{\|x-v_i\|}{\Delta_i} & \text{if } \|x-v_i\| \le \Delta_i, \\ 0 & \text{if } \|x-v_i\| > \Delta_i,\end{cases} i=1,2,\cdots,q.$ (1)

其中，v_i 是第 i 条规则对应的输入区域中心，Δ_i 是规则输入区域的宽度，y_i 是规则对应的输出重心，q 表示规则的个数，结论部分可以选择成线性函数 $y_i=u_i+\theta_i^{\mathrm{T}}(x-v_i)$.

3 模糊规则学习及自适应模糊推理

3.1 利用竞争学习对输入空间进行聚类

常规模糊系统中，输入空间都是人为划分的. 为了达到要求精度，常需对每个分量划分较多的区间. 这样，当输入变量维数较多时，输入划分子空间将以指数形式增长. 为此，本文采用聚类方法，根据输入空间数据的密集程度进行输入空间划分. 设 $x(t)$ 是输入空间的任一向量，v_i 是输入空间的聚类中心，采用带有明显统计意义的竞争学习算法如下[4]：

1）选择聚类个数为 q，初始化中心点 $v_i(i=1,2,\cdots q)$，

2）对于任一输入 $x(t)$，依输入样本密度确定最近中心点 v_c，

$$p_c\|x(t)-v_c\| = \min_j p_j\|x(t)-v_j\|, \ p_j=n_j/\sum_{i=1}^{q} n_i. \tag{2}$$

其中，p_j 表示中心点 v_j 附近样本的密集程度，n_j 为 v_j 被选为最近点的次数.

3）修正 v_c，

$$\begin{aligned} &n_c=n_c+1,\ a_c=\frac{1}{n_c}, \\ &v_c(k+1)=v_c(k)+a_c(x(t)-v_c(k)), \\ &v_j(k+1)=v_j(k),\ (j\ne c). \end{aligned} \tag{3}$$

使用 $p_j\|x-v_j\|$ 进行比较可使 p_j 参与竞争，从而避免一般竞争学习中的死点问题. 这是因为作为死点，它被选为中心 v_j 的次数几乎为零，引入 p_j 后，意味着 n_j 参与比较，使死点有机会竞争.

3.2 确定输入空间的区域划分边界

对于非线性系统模糊建模，输入空间的边 界划分是很重要的环节. 然而，大多数文献都忽略了这个方面. 事实上，区域边界的选择对解决过拟合和泛化性之间的矛盾以及减少划分区域的数目都是很有意义的. 本文在输入空间聚类的基础上，确定简便易行的区域划分边界，其方法如下：

1）固定聚类中心 $v_i(i=1,2,\cdots,q)$，并初始化半径 $r_i=0(i=1,2,\cdots,q)$，

2）对任意输入样本 $x(t)$，确定 $\|x(t)-v_c\|=\min\limits_j\|x(t)-v_j\|$， (3)

3）$\begin{aligned}&\text{if } \|x(t)-v_c\|>r_c(k),\ \text{then } r_c(k+1)=\|x(t)-v_c\|, \\ &\text{if } \|x(t)-v_c\|\le r_c(k),\ \text{then } r_c(k+1)=r_c(k).\end{aligned}$ (4)

显然，区分边界是一个球面，其中心是聚类中心，其半径由属于此区域的最远的样本决定，每个球面的半径不一定相同. 每个区域包含了所有属于自己的样本，并且一定程度上含有邻近区域的样本. 这样的区分边界是相互重叠的，具有模糊性. 由此得到的区域

其并集一定包含了所有训练样本,所以训练结果会有很好的记忆特性. 同时,由于划分区域是相互重叠的,所以如果训练样本足够充分,新的数据不属于划分区域的并集的可能性很小.

3.3 规则结论部分的学习

由式(1)可知,每条规则的前提部分对应一个输入划分区域,而结论部分则对应一个超平面 $y_i=u_i+\theta_i^{\mathrm{T}}(x-v_i)$, 其中 u_i 是第 i 条规则输出部分超平面的中心,θ_i 是相应系数. 给定输入输出训练样本,可通过下列算法学习这些参数

3.3.1　学习 $u_i(i=1,2,\cdots,q)$

1) 固定输入区域划分参数 $v_i,r_i(i=1,2,\cdots,q)$,

2) 随机选择训练样本 $x(t),y(t)(t=1,2,\cdots N)$,

3) 修正 $u_i(i=1,2,\cdots,q)$.

if $\|x(t)-v_i\|\leq r_i$

then $n_i(k+1)=n_i(k)+1,a_i=\dfrac{1}{n_i(k+1)},u_i(k+1)=u_i(k)+a_i(y(t)-u_i(k))$.　(5)

if $\|x(t)-v_i\|>r_i$,

then $u_i(k+1)=u_i(k)$.　(6)

3.3.2　学习参数 $\theta_i,i=1,2,\cdots,q$(通过 RLS)

1) 固定 $v_i,r_i,u_i(i=1,2,\cdots q)$, 并初始化 $f_i=\alpha I$, α 是一个较大的数,I 是单位矩阵,

2) 随机选择训练样本 $x(t),y(t),t=1,2,\cdots,N$,

3) 修正 $\theta_i(i=1,2,\cdots q)$.

if $\|x(t)-v_i\|\leq r_i$,

$$\text{then } f_i(k+1)=f_i(k)-\frac{f_i(k)(x(t)-v_i)(x(t)-v_i)^{\mathrm{T}}f_i(k)}{1+(x(t)-v_i)^{T}f_i(k)(x(t)-v_i)},\tag{7}$$

$$\theta_i(k+1)=\theta_i(k)+f_i(k+1)(y(t)-u_i-(x(t)-v_i)^{\mathrm{T}}\theta_i)(x(t)-v_i).$$

if $\|x(t)-v_i\|>r_i$,

then $f_i(k+1)=f_i(k)$,　(8)

$\theta_i(k+1)=\theta_i(k)$.

在 u_i,θ_i 的学习过程中,每次 u_i,θ_i是否修正根据训练样本是否属于某一或某些输入区域而定. 如果训练样本$\{x(t),y(t)\}$属于多个输入划分区域,那么它将同时修正多个 u_i,θ_i. 这一点不同于其它文献所用方法[1]. 显然,本文的方法具有模糊特性.

3.4 自适应模糊推理

利用训练样本学到模糊规则后,对于任意新的输入可通过模糊推理得到相应的输出. 由于区分边界的重叠性,所以当新的输入在某一划分区域内或某些划分区域内时,可用式(1)进行模糊推理. 然而,由于训练过程中样本可能不会充满整个输入空间,或者由于输入空间聚类个数不一定很多,可能会出现划分好的输入区域的并集小于输入空间的情况,在这种情况下,可以采用如下自适应模糊推理方法进行推理.

对于任一输入 $x(t)$:

1) 设 $\mathrm{adap}_i(0)=r_i(i=1,2,\cdots q)$以及 $k=1$,

2) $\mathrm{adap}_i(k)=\mathrm{adap}_i(k-1)+\beta r_i(i=1,2,\cdots q)$,

$$\text{if } \| x(t)-v_i \| < \text{adap}_i(k), \text{then } s_i = 1 - \frac{\| x(t)-v_i \|}{\text{adap}_i(k)}, \text{ else } s_i = 0 \tag{9}$$

3) if $\sum_{i=1}^{q} s_i = 0$, then $k = k+1$, goto (2),

$$\text{if } \sum_{i=1}^{q} s_i \neq 0, \text{then 进行模糊推理 } y = \frac{\sum_{i=1}^{q} s_i (u_i + (x(t)-v_i)^{\mathrm{T}} \theta_i)}{\sum_{i=1}^{q} s_i}. \tag{10}$$

其推理过程是,以 βr_i 为步长，逐渐增加每个输入区域的半径,直到某一区域或某些区域包含新的输入 $x(t)$. β 的选择与聚类个数也有关系,当聚类个数较多时,输入区域并集接近或等于输入空间,新的输入在输入划分区域内的可能性较大. 这时,系数 β 可以取较小的值而不影响推理结果.

4 实验结果

本文采用文献[5]所用的常用算例来说明本文提出的算法具有很高的可靠性,很好的泛化性,可以减少输入空间的划分区域，而且每个参数的选择都不敏感. 设被辨识非线性模型为

$$y(t+1) = \frac{y(t) + y(t-1)}{1 + y^2(t) + y^2(t-1)} + u^3(t), \tag{11}$$

这个模型等价于映射 $y(t+1) = f(y(t), y(t-1), u(t))$. 训练数据通过 $u(t)$ 在[-1,1]随机产生. 泛化性检验数据由输入 $u(t) = \sin(2\pi t/25)\sin(2\pi t/10)$ 产生 . $\beta = 0.6$，聚类数 q 分别选为 10，20，…，70. 每种情况下,学习十次，其中给出误差最小、最大及平均值的

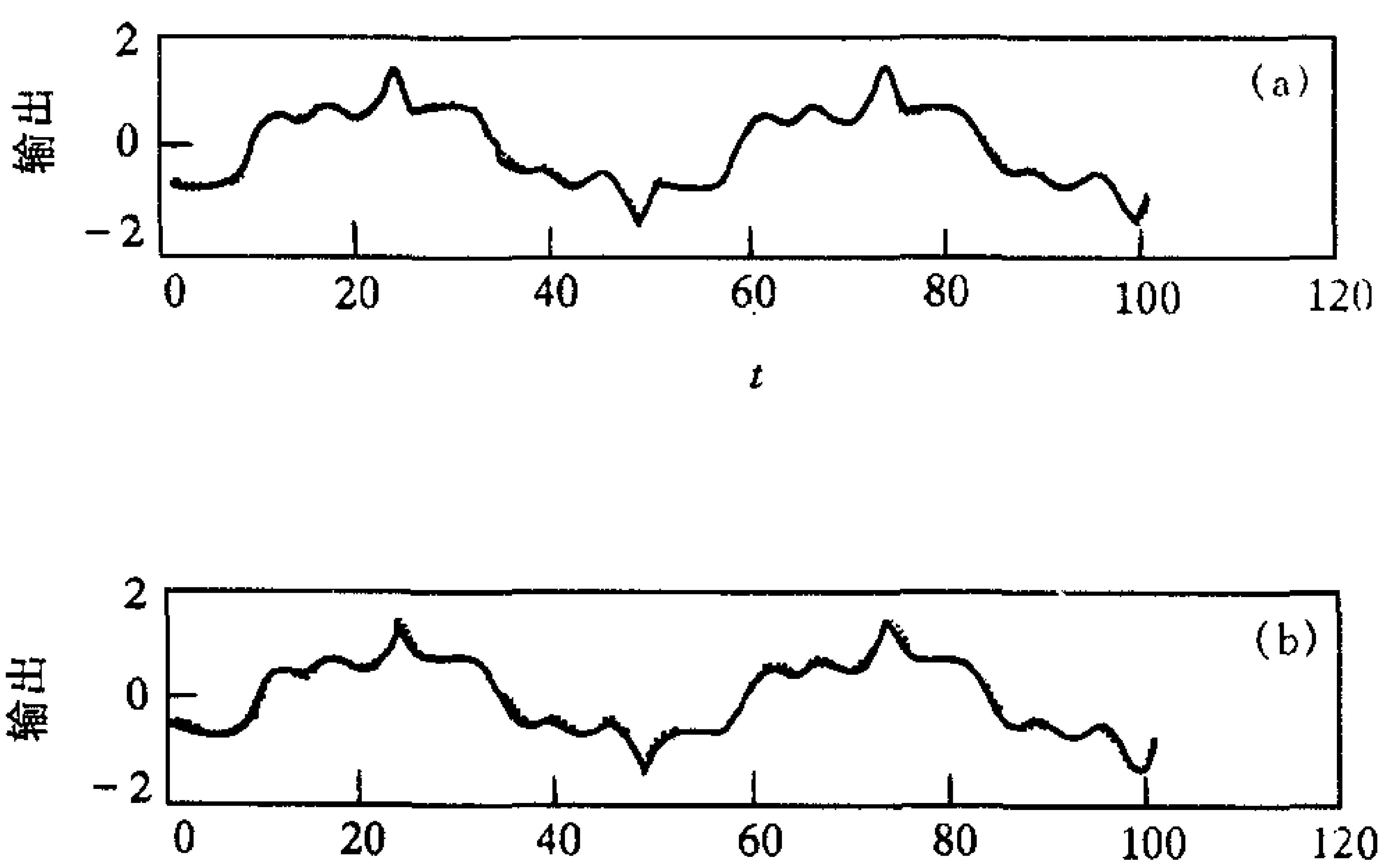

图 1 $\beta=0.6$，$q=60$ 时,实际输出(实线)与估计输出(虚线)比较结果

图(a) 对应误差最小情况. 图(b) 对应误差最大情况

情况. 图 1 给出了 $\beta=0.6$ 和聚类数 $q=60$ 时,误差最小及误差最大的情况. 可以看出,图(a)几乎在所有点处,估计输出都非常逼近实际输出. 图(b)中,在 $t=16$，$t=52$，$t=72$ 附近,估计输出与实际输出有微小的差别. 为了达到满意的结果，文献[5]中需要 1000 个隐

层节点，而本文仅需几十条规则.

5 结论

本文考虑用竞争学习算法、模糊规则及模糊推理对非线性系统进行建模的问题. 采用聚类算法为输入空间聚类，用重叠边界为输入空间划分区域. 提出了自适应模糊推理概念及算法. 得到以下好的结果:解决了现有文献中存在的过拟合和泛化矛盾问题;算法可靠,学习结果重复率高;容易为算法选择参数,因为参数对于学习结果不敏感;使用较少的输入划分区域. 大量实验结果表明，本文给出的学习算法明显优于其它文献给出的学习算法.

参 考 文 献

[1] Nie J. Constructing fuzzy model by self-organizing counterpropagation network. *IEEE Trans. SMC.*, 1995, **25** (6): 963—970.

[2] Wang L X, Mendel J M. Generating fuzzy rules by learning from examples. *IEEE Trans. SMC.*, 1992, **22** (6):1414—1427.

[3] 张化光. 热工过程的模糊建模与控制. 东南大学博士论文, 1992. 12

[4] Xu L, Krzyzak A, Oja E. Rival penalized competitive learning for clustering analysis, RBF net, and curve detection. *IEEE Trans. NN.*, 1993, **4**(4):636—649.

[5] Choi C H, Choi J Y. Constructive neural networks with piecewise interpolation capabilities for function approximations. *IEEE Trans. NN.*, 1994, **5**(6):936—944.

LEARNING OF FUZZY RULES AND ITS APPLICATION TO NONLINEAR SYSTEMS MODELING

CHEN JIANQIN　XI YUGENG　ZHANG ZHONGJUN

(*Institute of Automation, Shanghai Jiao Tong University, Shanghai* 200030)

Abstract　This paper discusses fuzzy inference systems and their application to nonlinear system modeling. The paper presents a new algorithm for learning fuzzy rules. The algorithm first partitions input data into some clusters by competitive learning, then determines the decision margins for each input cluster, and finally, learns the fuzzy rules for each input local region. The paper proposes and adaptive fuzzy inference method for the fuzzy rules. Examples are provided to demonstrate the presented learning algorithm and the computing results show that the method of the paper is superior to those in the reference.

Key words　Fuzzy systems, competitive learning, nonlinear system modeling, neural networks.