

武汉理工大学

硕士学位论文

非线性系统的模糊辨识方法与应用研究

姓名：杨建林

申请学位级别：硕士

专业：机械制造及其自动化

指导教师：黄之初

20051001

摘 要

模糊模型是当前系统辨识方法中经常用到的模型,它具有能够有效地综合利用专家知识和系统的输入输出数据的能力,并且得到的模型相对其它黑箱模型来说具有可解释性。目前为止,已经有很多模糊建模方法被提出来,然而,仍然存在着许多待解决的问题。

本文中,将研究的问题集中在下面三个方面:模糊模型的结构辨识,模型的精确性和可解释性间的平衡,模型的参数辨识:

一、模型的的结构辨识和参数辨识。在这里采用模糊C均值聚类的方法从数据中获取模糊规则前提部分的模糊集。模糊聚类依据输入数据空间上的相似度测试逐渐地将输入输出数据分成组,然后从每组中选择一条零阶或一阶TS形式的模糊IF-THEN规则,以组成一个初步的模糊规则库,此规则库可以直接用作模糊推理或用参数辨识的方法进行优化。与其它方法相比,这种获取规则的方法具有能以更小的代价获得更简化的模型的优点。

二、模型的精确性和可解释性间的平衡。模糊建模相对于其他建模方法最大的优点就是其结果具有可解释性。就目前的研究情况来看,不论何种方法,获取的模型都较复杂,模糊集合重叠,具有很大的冗余。从得到的结果中我们很难解释每个模糊集的物理意义。这里使用基于相似度合并方法提供了一个较为合理且快速的模糊集合并机制,藉此减少模型的复杂性,并可避免一般聚类方法容易产生多余模糊集的问题。

三、模型的参数辨识。得到初步的模糊规则库后,将这个TS模糊模型融合到神经网络结构中,利用神经网络的学习能力调节模型参数。主要讨论误差反向传播学习方法和一种混合学习方法。

这三个方面侧重点不一样,同时这三个方面是相辅相成的,在研究具体的问题时它们之间是一个反复完善的过程,可以依据对模型要求的不同,具体安排。

仿真及实验结果表明模糊建模方法能有效地建立非线性系统模型,模型较好的精确性。相对于其它建模方法而言,得到的模型具有可解释性,可能用专家知识等优点。

关键词:非线性系统辨识,模糊系统,混合式学习算法,模糊C聚类,可解释性

ABSTRACT

Fuzzy model has been recognized as a powerful tool which can facilitate the effective development of models by combining information from different sources, such as empirical models, the expert knowledge or the system input-output data. And the result model is more interpretable than other black-box methods. By now there are many way to build the fuzzy model, but there also are some unsolved problems.

In this thesis, we will discuss how to extract the model from system input-output data, how to insure the model precision and interpretability. We will major discuss two types questions:

A. The estimate of Fuzzy model's structures and parameters. Firstly, construction methods based on fuzzy C-means clustering originate from data analysis and pattern recognition, where the concept of fuzzy membership is used to represent the degree to which a given data object is similar to some prototypical object. Then, extract the antecedent membership functions by projecting the cluster onto the individual variables. The number of clusters is equal to the number of rules in the rule base.

B. The between model precision and interpretability. Contrast with others method, the most advantage of fuzzy model is the interpretability of the result. However, many methods don't handle this issue. Moreover, in general no model reduction is applied, while this may simplify the model by removing redundant information. The model is very complex and there are heavily redundancies from which we can't know the behavior of the system. In this thesis, we merge similar fuzzy sets based on similarity measure between fuzzy sets in order to reduce the rule base. The simulation results show it's effective.

C. At last, use optimization algorithm to estimate the rule base's parameters. In case, it can be formed as a fuzzy network model, the cluster can be defined as linear subspaces of the system. And use the error back propagation or a hybrid learning algorithm to training the parameters. Compare whit other methods, it is more advantage which can get a more compact model with less cost. These three items place different emphasis on the fuzzy model identification issue, they are interrelated. When deal with a problem, repeatedly iterates between them, however, they can be rearranged corresponding to specific needs.

The results of simulations and experiments show that proposed methods can build a fuzzy model from the data and expert knowledge. The model has good performance

in precision and interpretability.

Keywords: nonlinear system identification, fuzzy system, hybrid learning algorithm, fuzzy-c means clustering, interpretability.

第一章 绪论

1.1 研究的背景及意义

人类总是想把握自然。

人们总是想预先知道在一定条件下事物会出现什么样的结果，但是时间是人类无法改变的。发生了的是历史，不再重演；没有发生的是将来，不可预知。幸运的是事物的发生有其必然性，人们从已发生了的实践中总结规律，对实体特征和它的变化规律进行抽象，构建它的简化模型，试图以这种简化了的模型来预测事物在一定条件下会发生什么。这就是系统建模的主要任务。^[1]

建立系统数学模型的方法可分为机理分析法和测试法两大类，本文讨论的是一种测试分析法。系统的输入输出信号一般总是可以测量的，系统的动态特性必然表现在这些输入输出数据中。因此，可以利用系统的输入输出数据提供的信息，经过加工处理来建立系统的数学模型。这种建模方法又叫系统辨识。

系统辨识的常用方法有：阶跃响应法、脉冲响应法、相关函数法、最小二乘类参数辨识方法、极度大似然法等，这些方法都是建立对象的所谓精确的数学模型。然而，随着社会、经济的迅速发展，对科学管理、预测、监测和控制深度和广度方面的要求不断提高，以及科学研究的不断深入，所研究的对象越来越复杂，影响因素也越来越多，甚至存在一些难以精确描述的人为因素，且存在着大量延时、非线性、时变等非常复杂的现象，导致很难建立精确的数学模型。因此，在复杂的大系统中，复杂性与精确性形成了尖锐的矛盾。要想确切地描述复杂现象的任何现实的物理状态，事实上是不可能的。为了对整个问题描述有意义，我们必须在模型的准确性与简明性之间取得平衡。

基于模糊集合理论，用系统的输入、输出数据来辨识系统的模糊模型，是对复杂系统建模系统建模与辨识的一种有效途径。这种模糊模型在复杂系统的建模、预测、控制与决策中得到了广泛应用，并且越来越显示出其重要作用。另外，在实际的一些复杂过程控制中，也经常会出现这种情况，一些控制器很难控制的问题，操作人员和工程师凭其经验和知识却能得到满意的控制效果，但操作人员却不能够用语言确切地描述出在什么情况下采用什么动作的完整过程。因此，通过数据对人的控制行为进行模糊建模也是很有意义的问题。

总之，无论是要模拟操作人员的控制行为，还是要建立基于模型的控制；无论是要对复杂系统的行为进行预测，还是要为达到某个目标而进行决策，对系统进行模糊建模都是非常有意义和必要的。

模糊辨识与传统的辨识方法在辨识步骤和任务上大体相同，模糊系统相对于

其他方法有其独特的优点具体体现在以下几个方面:

1) 相对于经典辨识方法, 它能有效地辨识复杂和病态结构 (ill-defined) 的系统, 能够有效地辨识具有大时延、时变的非线性系统^[2, 3, 4]。现代工业过程中的系统越来越复杂, 模糊辨识方法的这些优点显得尤为突出。

2) 相对于神经网络等黑箱建模方法, 它可以得到性能优越的人类控制器^[5, 6]。一个好的系统模型应该具有明确的物理意义, 无论是科学家还是工程操作人员都能够理解它的结构和参数。黑箱模型所有的信息是隐含的, 赋予它的权值是它工作性能的关键, 然而却无法知道权值和神经网络在做什么, 如果要理解它几乎是不可能的。模糊系统是基于规则库的, 用模糊系统来辨识对象, 可以从观测数据中提取系统的模糊规则, 得到关于系统的IF-THEN规则形式的描述。另外, 从这些规则的物理含义, 很直观的就可以判断某组参数是否正确, 模糊系统的参数是否合适。模糊控制和人类的控制策略, 无论是专家还是一线工作人员都很容易理解它的操作原理。

3) 可得到被控对象的定性与定量相结合的模型。对绝大多数的工程系统来说, 其重要的信息来源有两种: 提供测量数据的传感器和提供系统性能描述的专家。我们称来自传感器的信息为数据信息, 来自专家的信息为语言信息^[7]。现代系统辨识理论指出: 并非所有系统都可以从它的输入、输出完全确定其结构和参数, 专家知识在系统建模中也有相当重要的作用。

从以上三个方面来讲, 模糊辨识方法相对于其它方法具有本质的优点, 但模糊模型的建立也存在着实现上的困难, 下面将讨论国内外在这种方法研究上的现状。

1.2 国内外研究现状调查

非线性系统辨识与控制是模糊理论的一个重要的应用领域。模糊系统理论随着模糊数学和人工智能技术的发展而不断完善, 但其中最重要的问题之一就是模糊规则的获取。模糊控制理论的发展从某种意义上说也是围绕着这一问题而展开的^[8]。近来, 这一研究领域取得了许多理论成果和实际应用, 但还是存在许多难题, 在理论研究与实际应用间还存在很大差距。模糊系统设计的核心问题就是如何从数据中获取合适的IF-THEN模糊规则。

模糊辨识是通过输入输出测量数据, 对模糊模型中的结构和参数进行的辨识。事实证明模糊模型在非线性动力系统建模, 基于规则的学习控制, 模式识别等领域起到了很大的作用。最近几年, 基于模糊推理的建模方法大概分为以下几类:

1. 基于模糊关系模型的建模方法^[9, 10, 11];
2. 基于模糊神经网络的模糊建模方法^[12, 13, 14, 15];
3. 基于Takagi-Sugeno (TS) 的模糊线性函数模型的建模方法。^[16]

下面主要针对第三种建模方法,对目前的理论与实际应用加以简单综述。文献^[16]提出的模糊模型,也称为一阶TS模型。国内外许多学者在此模型的基础上进行了卓有成效的研究。TS模型具有很多优点:由于其规则前件是模糊变量,而结论部分是关于输入输出变量的线性函数,它以局部线性化为基础,通过模糊推理方法实现了全局的非线性。这种模型使用了局部线性化函数,能克服以往模糊模型的高维问题,所以已成为人们广泛使用的模糊模型。总的来说,这方面的研究主要集中于TS模型的结构辨识和参数辨识两个方面:

一、TS模型的结构辨识

查表法:这种方法^[17]的主要思想是按照某一确定的过程,例如等分,来划分输入模糊空间,即确定输出语言项与模糊区域的映射关系。把模糊规则库描述成一个两位输入情况下的可查询表格,每个子格代表一条规则,它确定了模糊规则的结构。这种设计方法简单易用,但对于较复杂的系统应用中存在问题。划分准则是确定的,不具有学习功能。模糊空间划分愈细,辨识结果愈好,但这样会使模糊规则的数目快速增长。而且由查表法得到的模糊规则库可能是不完备的。

模糊聚类法:B. Moshiri等人^[18, 19, 20, 21]利用模糊聚类方法确定系统模糊规则。模糊聚类法把输入输出乘积空间划分成不相交的子集或组,然后每组数据仅用一条规则来描述,聚类中心的数目就是模糊模型中模糊规则的数目,对每个区域可以用简单的函数,如线性子模型来逼近系统曲面。将聚类数组投影到前提变量上,这样就可以获得规则前提部分的模糊子集,用这些模糊子集来构造合适的模糊系统。模糊聚类方法能够有效地减少模糊规则的数目,但是这种方法产生的结果可能存在模糊集相互重叠的情况,在本文的后面将作解释。本文用一种基于模糊集合相似度测量的方法对聚类的结果进行测量,合并相似的模糊规则,减少模型的冗余度。

二、TS模型的参数辨识

梯度下降法根据先验知识或一般模糊网格法初步确定模糊模型,然后利用梯度下降法优化模糊模型的参数。它是一种具有学习功能的算法。其缺点是系统结构的确定和初始参数的设置,这需要大量的启发性知识。当输入量增加时,学习复杂性成指数增长。

遗传算法遗传算法是模拟达尔文的生物进化机制(选择、交叉、变异)的一种自适应全局优化概率搜索算法。对规则前、后件参数进行编码,借助于遗传算法优化模糊系统^[22],这种方法能得到系统的全局最优解。但二进制编码或经典的

遗传算法在解决多维、高精度、连续问题时染色体的长度往往过长,使计算量随之猛增,效率很低,而实值编码方法^[23]可较好地解决这一问题。

模糊系统与神经网络相结合将神经网络优秀的学习能力和模糊推理整合在一起,形成Neuro-Fuzzy模型^[24]。然而,许多Neuro-Fuzzy设计方法仅涉及到参数的辨识,回避了结构的辨识,一般假设结构已预先确定好了。

总的来说,从模型的精度来看各种算法均能得到较满意的结果,但是很少涉及到结构优化的参数辨识。即大多数研究者将结构辨识和参数辨识这两个问题是分开考虑的。

1.3 研究目的、所解决的关键问题

1、研究的目的

将模糊系统应用到非线性系统辨识问题中的主要难点就在于模糊规则的获取及如何保证模型的精确性和可解释性。本文将围绕这个问题展开讨论,结合其他算法解决模糊规则的获取及模糊系统初始值的确定方法,并将这些方法应用到非线性系统的辨识问题中。

本文的目的就是:

1) 选择一种合适的自适应模糊系统,讨论此方法在工程系统中综合地利用数据和语言两种信息;

2) 将自适应模糊系统应用于非线性系统的辨识问题,构造模糊辨识方法的一般步骤;

3) 解决模糊模型的精度与模型的可解释性间的平衡问题。

2、所解决的关键问题

对于非线性系统的模糊辨识,本研究主要解决以下几个方面的问题:

1) 减少模糊聚类方法得到模型的冗余度,提高模型的可解释性;

2) 误差反向传播算法与全局最小二乘法,最小二乘法相结合,辨识模糊模型的参数;

3) 模糊聚类算法和模糊神经网络方法的应用及软件实现;

4) 模糊系统辨识方法在双容水箱系统液位控制中的应用。

1.4 论文结构

本文将分为六个部分:

第一章为绪论,主要介绍本文的目的和意义,简单介绍国内外在模糊系统辨识方面的研究现状。

第二章介绍模型系统理论方面的知识。主要介绍模糊集、模糊规则和模糊推理的基础理论，它们是构建本文内容的数学基础。在这一章中还介绍了两种常用的模糊模型，Mamdani模型和Takagi-Sugeno (TS) 模型，重点讨论了TS模型的结构，并给出了TS模型的数学表达式。本文中讨论的辨识方法均在这种模型的基础上展开讨论的。

第三章主要讨论模糊系统辨识的结构辨识问题。通过对数据进行模糊聚类，将数据划分为若干类，然后将每类数据投影到输入空间的每个变量上，作为输入变量的模糊集。在本章中还讨论了另一种模糊聚类方法，Gath-Geva模糊聚类。该方法中采用模糊最大似然估计作为数据点到聚类中心的距离。最后讨论了模型的可解释性问题。通过两种方法来简化模糊规则库：第一种方式为合并相似隶属度函数的方法减少模型中的冗余模糊集和规则，另一种方式是在模型参数优化过程中添加约束参数，以这种方式约束模糊集间的重合度，提高模型的可解释性。

第四章中主要讨论TS模型的参数辨识问题。主要阐述了误差反向传播算法和几种混合学习算法。混合学习方法中采用最小二乘法和误差反向传播学习相结合的方式，两者分别训练模型的结论参数和前提参数。最后提出将全局最小二乘法和局部最小二乘法相结合的方式，以达到在模型的精确性和局部模型的可解释性间的平衡。

第五章主要讨论模糊模型辨识方法在双容水箱系统中的应用。首先利用本文阐述的方法根据系统的输入输出数据结合局部模型的约束参数，得到系统的模糊模型。然后结合预测控制和内模控制对实际系统进行控制，结果表明了模糊模型的优越性。

最后一部分为结论与展望。总结了本文中提出的各种方法的优点与不足，并对今后进一步研究方向作了初步探讨。

第二章 模糊系统模型

模糊系统早期最有成就的当属于模糊控制在消费品和工业过程控制中的应用。同时因为大多实践中只需要我们自己制定几条模糊规则就可以实现，这就使得有这样一种说法，称模糊系统和模糊控制是经验学科，缺少严格的理论支持和指导。我们应当注意到模糊系统经历了近40年的发展，它不仅仅是一种简单的控制方法，已成为一种重要的计算方法^[25]。一个重要的原因就是模糊系统可以用一些简单的线性函数逼近复杂的非线性曲面。模糊系统的精华在于它将一个问题划分为多个局部区域，用线性函数描绘局部区域的性能，各区域的线性函数赋有相应的连接权值，使得这些线性函数能够平滑地连接达到逼近平滑非线性函数的目的。

2.1 模糊系统基础理论

模糊系统是一种基于知识或规则的系统。它的核心就是有所谓的IF-THEN规则所组成的知识库^[26]。下面将简单介绍模糊系统的有关概念^[17]：

2.1.1 模糊集

模糊集可以认为是经典数学理论中集合概念的扩展，它为模糊理论的发展应用奠定了基础。经典集合论中集合是指具有某种性质或用途的元素的全体。集合 A 可以定义为集合中元素的穷举（列举法），或描述为集合中元素所具有的性质（描述法），也可用隶属度法定义。隶属度法引入了集合 A 的 $0,1$ 隶属度函数，用 μ_A 表示，它满足：

$$\mu_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases} \quad (2-1)$$

可以看出经典集合论中要求元素具有一个定义的很准确的性质，论域中的元素非此即彼。但是在现实生活中我们发现这种方式往往不能表达我们要陈述的意思。例如，我们说“那个人的是个年青人。”具体多大年纪范围内位年青人，我们并没有一个非常严格的定义。

为了克服经典集合理论中的这种局限性，需要一种新理论——模糊集合理论，来弥补它的局限性。

定义 2.1.1. 论域 U 上的模糊集合是用隶属度函数 $\mu_A(x)$ 来表征的， $\mu_A(x)$ 的取值范围是 $[0, 1]$ 。

经典集合论中元素的隶属度只允许取两个值——0或1，而模糊集合中元素的隶属度则是在区间 $[0, 1]$ 上的一个连续函数。所以说模糊集合一点都不模糊，它只是一个带有连续隶属函数的集合。

2.1.1.1 模糊集合的一些基本概念

这里简单介绍一下本文中将要用到的几种基本概念。

支撑集(support)和模糊单值(singleton) 论域 U 上模糊集 A 的支撑集 $supp(A)$ 是一个清晰集，它包括了 U 中所有在 A 上具有非零隶属度值的元素，即

$$supp(A) = \{x \in U | \mu_A(x) > 0\} \quad (2-2)$$

如果模糊集合的支撑集仅包含 U 中的一个点，则称该模糊集为模糊单值。

投影(projections) 令 A 是 R^n 上的一个模糊集，其隶属度函数为 $\mu_A(x) = \mu_A(x_1, \dots, x_n)$ ，令 H 为 R^n 中的一个超平面(hyperplane)，定义 H 为 $H = \{x \in R^n | x_1 = 0\}$ 。定义 A 在 H 的投影为在 R^{n-1} 上的模糊集合 A_H ，其隶属度函数为

$$\mu_{A_H}(x_2, \dots, x_n) = \sup_{x_1 \in R} \mu_A(x_1, \dots, x_n) \quad (2-3)$$

模糊算子(operator) 设 A 和 B 是 U 上的两个模糊集合。下表为两种常用的模糊算子：

例如，对于下面的这条规则：

表 2-1: 常用模糊逻辑算子

	A and B	A or B	not A
Zadeh	$\min(\mu_A, \mu_B)$	$\max(\mu_A, \mu_B)$	$1 - \mu_A$
probabilistic	$\mu_A \cdot \mu_B$	$\mu_A + \mu_B - \mu_A \cdot \mu_B$	$1 - \mu_A$

$$R_i: \text{if } x_1 \text{ is } A_{i1} \text{ and } x_2 \text{ is } A_{i2}, \dots, \text{ and } x_p \text{ is } A_{ip} \text{ then } y_i = a_i^T x + b_i \quad (2-4)$$

用上面两种算子得到这条规则的隶属度分别为：

$$\beta_i(x) = \min(\mu_{A_{i1}}(x_1), \mu_{A_{i2}}(x_2), \dots, \mu_{A_{ip}}(x_p)) \quad (2-5)$$

和

$$\beta_i(x) = \mu_{A_{i1}}(x_1) \cdot \mu_{A_{i2}}(x_2) \cdot \dots \cdot \mu_{A_{ip}}(x_p) \quad (2-6)$$

通常，并和交的运算分别用算子 \vee 和 \wedge 来表示。

2.1.2 语言变量和模糊IF-THEN规则

2.1.2.1 语言变量

如果一个变量取自自然语言中的词语为值，则称其为语言变量。这里，词语由定义在论域上的模糊集合来描述的，变量也是在论域上定义的。语言变量在某种意义上实数值变量的一种扩展，即允许语言变量取模糊集为值。如图2-1，“年龄”是一个语言变量，它的术语集合 $T(\text{年纪})$ 可以是 $T(\text{年纪}) = \{Young, Middle, Old\}$ ，

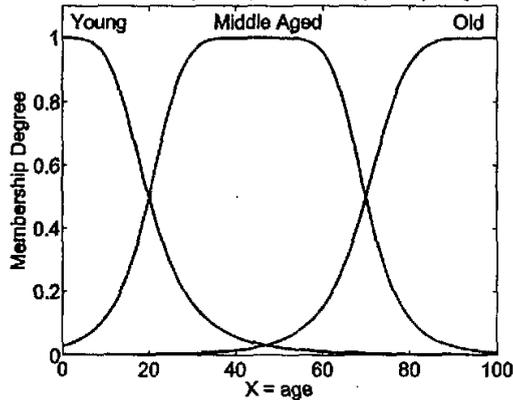


图 2-1: TS模型对非线性函数的分段逼近能力

语言变量是人类知识表达中最基本的元素，传感器会给出数值，而专家会给出语言变量。例如，当使用流量计来测量水的流速时，流量计会给出23ml/s, 45ml/s等数字；而当让某人告诉我们水的流速时，通常会说“水流的慢”、“水流的很快”等话语。在一些情况下语言变量比精确的数学描述更方便，易理解。例如，我们说“今天特别闷热！”，我们想表达的意思就是感觉到天气闷热，不舒服。如果说“今天是华氏 98.60°，气压960百帕。”我们可能反而不知所云。引入语言变量的概念是人类知识系统有效地嵌入工程系统的第一步。

2.1.3 模糊推理

在模糊逻辑中，命题都是由模糊集表述的模糊命题。模糊逻辑的最终目的是将模糊集理论作为一种主要工具，为不精确命题的近似推理提供理论基础。为实现这一目的，引入：

广义取式推理 (Generalized Modus Ponens)：

前提1: x 为 A'

前提2: 如果 x 为 A ，则 y 为 B

结论: y 为 B'

即, $\mu_{B'}(y) = \sup_{x \in U} [\mu_{A'}(x), \mu_{A \rightarrow B}(x, y)]$

广义拒式推理 (Generalized Modus Tollens):

前提1: y 为 B'

前提2: 如果 x 为 A , 则 y 为 B

结论: x 为 A'

即, $\mu_{A'}(x) = \sup_{y \in V} [\mu_{B'}(y), \mu_{A \rightarrow B}(x, y)]$

广义拒式推理 (Generalized Modus Tollens):

前提1: x 为 A , 则 y 为 B

前提2: 如果 y 为 B' , 则 z 为 C

结论: 如果 x 为 A , 则 z 为 C'

即, $\mu_{A \rightarrow C'}(x, z) = \sup_{y \in V} [\mu_{A \rightarrow B}(x, y), \mu_{B' \rightarrow C}(y, z)]$

2.2 模糊模型

在系统辨识中,经常使用到基于模糊规则库的模糊模型。模糊模型由模糊产生器、模糊规则库、模糊推理机和反模糊化器四部分组成,其基本结构如图2-13所示。设 $x \in U = U_1 \times \dots \times U_n \subseteq X_1 \times \dots \times X_n$ 为模糊系统的输入, $y \in V \subset R$ 为模



图 2-2: 模糊模型的结构图

糊系统的输出,则模糊系统构成了由子空间 U 到子空间 V 上的一个映射。

模糊产生器将论域 U 上的点一一映射为 U 上的模糊集合,反模糊化器将论域 V 上的模糊集合一一映射为 V 上确定的点,模糊推理机根据模糊规则库中的模糊推理知识以及由模糊产生器产生的模糊集合,推理出模糊结论,即论域 V 上的模糊集合,并将其输入到反模糊化器。

在模糊模型中,由于采用模糊产生器、模糊规则库、模糊推理机和反模糊化器的方法很多,每一种组合都会产生不同类型的模糊模型。一般常用的模糊模型为 Mamdani 模型和 Takagi-Sugeno (TS) 模型:

2.2.0.1 Mamdani 模型

在此模型中,前提(规则的IF部分)和结论(规则的THEN部分)均为模糊命

题:

$$R^{(l)}: \text{ if } x_1 \text{ is } A_1^l \text{ and } \cdots x_n \text{ is } A_n^l \text{ then } y \text{ is } B^l, \quad l = 1, 2, \dots, M \quad (2-7)$$

这里 $A_i^l (i = 1, 2, \dots, n)$ 和 B_i 分别是前提和结论语言变量(如,“小”,“大”,等),它们用模糊集表示。 M 是模型中规则的数目。

2.2.0.2 TS模型

Mamdani模型经常用在专家系统中,而TS模型在数据辨识系统中应用的越来越多。没有特别说明的情况下,本文中采用TS模型作为辨识的对象。在这种模型中模型的前提部分也是模糊命题,与Mamdani模型不同的是它的结论部分是输入变量的线性函数:

$$R^{(l)}: \text{ if } x_1 \text{ is } A_1^l \text{ and } \cdots x_n \text{ is } A_n^l \text{ then } y^l = \mathbf{a}_l^T \mathbf{x} + b_l, \quad l = 1, 2, \dots, M \quad (2-8)$$

式中,向量 \mathbf{a}_l 和参数 b_l 为结论参数。

Mamdani模型和TS模型的不同之处在于他们的结论部分,前者的结论部分为模糊集,后者的结论为线性函数。

TS模型将语言变量和标准回归函数结合在一起:它描述了输入空间各模糊划分区域对应的线性函数。带有中心平均解模糊器的模型输出可由下式来计算:

$$y = \frac{\sum_{l=1}^M \omega_l(\mathbf{x}) y^l}{\sum_{l=1}^M \omega_l(\mathbf{x})} = \frac{\sum_{l=1}^M \omega_l(\mathbf{x}) (\mathbf{a}_l^T \mathbf{x} + b_l)}{\sum_{l=1}^M \omega_l(\mathbf{x})} \quad (2-9)$$

式中, $\omega_l(\mathbf{x})$ 是第 l 条规则的高度,若采用乘积推理:

$$\omega_l(\mathbf{x}) = \prod_{i=1}^n \mu_{A_i^l}(x_i) \quad (2-10)$$

前提部分的模糊集描述输入空间的不同区域,结论部分的参数 \mathbf{a}_l, b_l 逼近前提部分描述的局部区域的线性模型。从这种意义上讲,可以把TS模型看作多个局部线性模型的加权求和,如图2.2.0.2所示。由于TS模型具有函数逼近能力和参数学习能力,本文中主要以这种模型来进行系统辨识。

2.3 模糊模型和系统辨识

假设一个多输入单输出系统有 n 维输入变量和一个输出变量 y 。经过测试得到系统的 N 对输入输出数据样本:

$$X = [\mathbf{x}_1, \dots, \mathbf{x}_N]^T, \quad \mathbf{y} = [y_1, \dots, y_N]^T \quad (2-11)$$

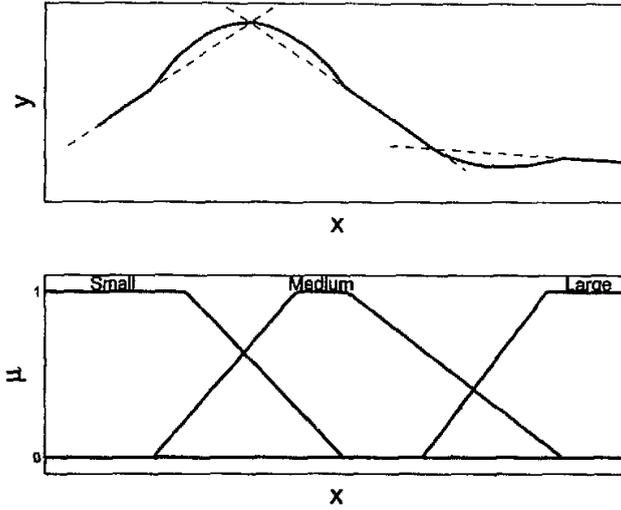


图 2-3: TS模型对非线性函数的分段逼近能力

式中, $x = [x_1, \dots, x_n]^T$ 为系统的输入变量。

在传统辨识方法中, 通过建立预测输出与过去时刻输入输出间的数学关系得到系统的模型。通常表示为带外生变量的非线性自回归 (NARX) 模型:

$$y(t) = f(y(t-1), \dots, y(t-n_y), u(t-d), \dots, u(t-d-n_u)) \quad (2-12)$$

式中, n_y, n_u 为输出、输入变量的最大延迟, d 为系统延迟时间, f 代表 NARX 模型的映射关系。

本文的目的就是从数据样本中获取一个模糊模型, 此模型由下面的一组规则表示:

$$R^{(l)}: \text{if } x_1 \text{ is } A_1^l \text{ and } \dots x_n \text{ is } A_n^l \text{ then } y^l = a_l^T x + b_l, \quad l = 1, 2, \dots, M \quad (2-13)$$

式中, A_i^l 表示输入变量的语言变量, 由隶属数 $\mu_{A_i^l}$ 定义。上面这个结构称为模糊 TS 模型。在这种意义上讲, 模糊辨识是一种特殊的系统辨识方法, 因为它得到的模型是一个模糊系统。因此, 模糊模型设计的主要任务是: 模型的结构辨识和参数估计。结构辨识主要是从输入输出样本中选择合适的模型前提变量、结论变量, 及模糊规则的数目 M ; 参数估计主要是确定每条规则的前提变量的模糊集 A_i^l 和结论参数 a_l, b_l 。

2.4 模糊模型的结构和参数辨识

从TS模糊模型可以看出,这种模糊系统设计的主要问题就是确定规则的数目 M ,及每条规则对应的隶属函数 $A_1^i, \dots, A_n^i, a_i, b_i$,也就是模型的结构和参数确定的问题。

模型结构的选择非常重要,因为它决定了模型逼近未知系统的适应性。一个模型的结构越详细对系统的逼近越精确,但模型的泛化能力越差。模型的泛化能力是指对于同一过程,根据有限样本得到的模型,对其他样本集也有良好的预测能力。模糊模型的构造中可以使用专家知识和输入输出数据,专家知识可以是一种近似的表述(定性的,启发式的heuristics)。有两种方式可以综合这两种信息:

- 1、首先利用专家知识构造规则库,这样就可以构造初始模型。然后利用输入输出数据对调节模型的参数(前提参数,结论参数)。

- 2、利用输入输出数据对直接构造模糊规则及相关的参数。这种情况下,与其它黑箱模型(如神经网络模型)相比,这种网络结构模型有一个突出优点,这种模型得到的结果具有可解释性。我们可以理解得到规则的实际意义,可以修改规则,或者添加规则以增加模型的合理性。

实际问题中,甚至可以将两种方式结合起来使用。大致确定模型的结构后,可以用多种优化方法确定模型的参数。

本文将主要就这两个问题展开讨论,首先将讨论利用模糊聚类法来构造合适的模糊规则库,然后将讨论模型参数辨识的几种方法。

第三章 模糊聚类方法构造模糊系统规则

3.1 引言

对于模糊模型来说模型结构的确定主要包括两个方面：

1、输入变量的选择。这包括物理输入量 x 和状态变量的选择。可以依据专家知识、对过程特性的理解和模型的用途来考虑如何选择合适的输入变量。然后可以参考某种准则，使用自动数据驱动选择程序来比较不同结构的性能，选择较好的模型结构。

2、隶属函数的数量，形式及规则数量的选择。这两种结构参数是相互关联的，因为有更多的隶属函数相应地必须定义更多的规则。它们决定了模型的复杂程度，及对输入空间划分的间隔。在模糊系统的设计过程中，规则数目的确定很重要。规则太多会令模糊系统变的复杂，同时可能会存在不必需的规则；规则太少有会令模糊系统的作用削弱，也许难以完成该问题的目标。模型的用途，可获取的信息（知识和数据）可以用来作为选择的参考。还可以用其它的算法来选择隶属函数和规则的数目。

模糊聚类方法由于无须监督控制，可以按着某一规则寻优，快速且有效，因此采用模糊聚类进行模糊辨识是一种有效的方法。本章采用模糊聚类的方法来确定模糊系统IF-THEN规则。基本思想是采用模糊聚类法把输入输出数据样本分类，再将每类数据视为一个输入输出数据对，即一类数据采用一条模糊规则。所以模糊规则的数目等于聚类的数目。

3.2 聚类划分

聚类分析是数理统计中研究“物以类聚”的一种多元分析方法。在数学上，把按一定要求对事物进行分类的方法叫做聚类分析，所要进行分类的对象叫做样本。因此，聚类分析的任务在于通过数学分析定量地确定样本的亲疏关系，从而客观地分型划类。

联合系统的输入输出数据样本(2-11)构成聚类数据集 Z ：

$$Z = [X; y]^T \quad (3-1)$$

模糊聚类方法将样本的笛卡尔空间 $\mathbb{R}^{N \times n} \times \mathbb{R}^{N \times 1}$ 划分为一些特定的局部区域，利用这些局部区域的线性模型逼近系统特性。

聚类可以看作是数据集的子集，依据得到的子集是清晰集还是模糊集可以将聚类方法分为硬聚类和模糊聚类。硬聚类法基于经典集合论，数据点只有属于或

不属于某一聚类两种情况。它将数据集 Z 中的数据划分到 M 个不相交的子集(聚类)中。模糊聚类法中数据点以不同的隶属度同时属于几个聚类,数据集 Z 被划分到 M 个模糊子集。由于事物本身在很多情况下都带有模糊(Fuzzy)性,因此把模糊数学方法引入聚类分析,就会使分类更切合实际。在模糊聚类法中,数据点不是被强制地划分到某个聚类中,而是赋给它一个 $[0,1]$ 间的隶属度,表示它属于某个聚类的程度。另外,硬聚类划分的离散本质性导致它在分析和算法上难以实现,因为它的解析函数是不可微的^[27]。

3.2.1 硬划分

聚类的目的是将数据集 Z 划分到 M 个聚类中。假设,由先前知识或试凑的方法得到 M 。

由经典集合论,一个硬划分可以定义为一组子集 $\{A_i | 1 \leq i \leq M \subset P(Z)$,它满足:

$$\begin{cases} \bigcup_{i=1}^M A_i = Z, \\ A_i \cap A_j = \emptyset, \quad 1 \leq i \neq j \leq M \\ \emptyset \subset A_i \subset Z, \quad 1 \leq i \leq M \end{cases} \quad (3-2)$$

这些条件是指数据 Z 中的所有数据均被唯一地划分到子集 A_i 中,且子集 A_i 不相交,不为空。写成隶属函数的形式为:

$$\begin{cases} \bigvee_{i=1}^M \mu_{A_i} = 1, \\ \mu_{A_i} \wedge \mu_{A_j} = 0, \quad 1 \leq i \neq j \leq M \\ 0 < \mu_{A_i} < 1, \quad 1 \leq i \leq M \end{cases} \quad (3-3)$$

式中, μ_{A_i} 是子集 A_i 的特征函数,它的值为1或0。

简单起见,将 μ_{A_i} 记为 μ_i ,将 $\mu_i(z_k)$ 记为 μ_{ik} ,划分可以写成矩阵 $U \in \mathbb{R}^{M \times N}$ 的形式,且矩阵 $U_{M \times N} = [\mu_{ik}]$ 满足下式:

$$\begin{cases} \mu_{ij} \in \{0, 1\}, & \forall i, \forall k \\ \sum_{i=1}^M \mu_{ik} = 1, & \forall k \\ 0 < \sum_{k=1}^N \mu_{ik} < N, & \forall i \end{cases} \quad (3-4)$$

定义 3.2.1. 硬划分空间 令 $Z = [z_1, z_2, \dots, z_N]$ 为有限集合, $2 \leq M \leq N$ 为正整

数, Z 的硬划分空间为集合:

$$M_{hc}\{U \in R^{N \times M} \mid \mu_{ik} \in \{0, 1\}, \forall i, k; \sum_{i=1}^M \mu_{ik} = 1, \forall k; 0 < \sum_{k=1}^N \mu_{ik} < N, \forall i\} \quad (3-5)$$

3.2.2 模糊划分

模糊划分可以看作是硬划分的推广, u_{ik} 可以取 $[0, 1]$ 间的值。这时表示模糊划分的矩阵 $U \in R^{M \times N} = [\mu_{ik}]$ 的约束条件为:

$$\begin{cases} \mu_{ik} \in [0, 1], & \forall i, \forall k \\ \sum_{i=1}^M \mu_{ik} = 1, & \forall k \\ 0 < \sum_{k=1}^N \mu_{ik} < N, & \forall i \end{cases} \quad (3-6)$$

定义 3.2.2. 模糊划分空间 令 $Z = [z_1, z_2, \dots, z_N]$ 为有限集合, $2 \leq M \leq N$ 为正整数, Z 的模糊划分空间为集合:

$$M_{fc}\{U \in R^{N \times M} \mid \mu_{ik} \in [0, 1], \forall i, k; \sum_{i=1}^M \mu_{ik} = 1, \forall k; 0 < \sum_{k=1}^N \mu_{ik} < N, \forall i\} \quad (3-7)$$

矩阵 U 的每个元素 u_{ik} 表示样本 z_k 属于聚类 i 的程度, 也称之为隶属度。 U 的每列元素之和为1, 也就是说每个样本到所有聚类的隶属度之和为1。但每个隶属度的取值没有限制。在硬划分中隶属度的取值只能为1或0, U 中的每列元素中只能有一个为1, 其它为0。这是模糊聚类和硬聚类本质区别在形式上的体现。

另外, 需要注意的是 U 矩阵是一个二维矩阵, 但因 $z_k = [x_1, x_2, \dots, x_n, y]$ 及 v_i 是 $n+1$ 维空间上的点, 所以元素 u_{ik} 表征的是多维空间中的关系, 它的取值为正实数。

3.3 模糊C均值聚类

3.3.1 K-均值算法

硬聚类较直观, 下面先来讨论它的实现方法。我们希望在众多可能的分类中寻求合理的分类结果, 为此, 就要确立合理的聚类准则。在硬分类时, 常使用的聚类准则是最小平方误差和。K-均值和K-中心点聚类法中使数据点 z_k 到 M 个聚类中心 v_i 的距离平方和最小, 即:

$$\begin{aligned} \min J_1(Z; U, V) &= \sum_{i=1}^M \sum_{k \in A_i} \|z_k - v_i\|_2 \\ \text{s.t. } U &\in M_{hc} \end{aligned} \quad (3-8)$$

式中, A_i 是第*i*个聚类中的数据集, v_i 是这个聚类中数据的均值。用 d_{ik} 表示第*i*类中的样本 z_k 与第*i*类的原型样本 v_i 之间的距离, 即 $d_{ik}^2 = \|z_k - v_i\|_2$ 。利用 μ_{ik} , $J_1(Z; U, V)$ 也可以表示为:

$$\min J_1(Z; U, V) = \sum_{k=1}^N \sum_{i=1}^M \mu_{ik} (d_{ik})^2 \quad (3-9)$$

聚类准则为寻求最佳组对 (U, V) , 以使得在满足约束条件 $\mu_{ik} \in M_{hc}$ 下 $J_1(Z; U, V)$ 为最小。

K-均值聚类中称 v_i 为原型样本, 如聚类的中心点:

$$v_i = \frac{\sum_{k=1}^{N_i} z_k}{N_i}, z_k \in A_i \quad (3-10)$$

式中, N_i 是 A_i 中数据点的数目。

3.3.2 模糊C均值聚类

Dunn按照Ruspini定义的模糊划分的概念, 把硬聚类的目标函数推广到模糊聚类的情况。为了避免产生平凡解, 保证这一推广有意义, Dunn对每个样本与每类原型间的距离用其隶属度平方加权, 从而把类内误差平方和目标函数扩展为类内加权误差平方和目标函数:

$$\begin{aligned} \min J_2(Z; U, V) &= \sum_{i=1}^M \sum_{k=1}^N (\mu_{ik})^m D_{ikA}^2 \\ \text{s.t. } U &\in M_{fc} \end{aligned} \quad (3-11)$$

在上述目标函数中, 样本 z_k 与第*i*类的聚类原型 v_i 之间的距离度量一般定义为

$$D_{ikA}^2 = \|z_k - v_i\|_A = (z_k - v_i)^T A (z_k - v_i) \quad (3-12)$$

式中, A 为 $(n+1) \times (n+1)$ 阶的对称正定矩阵, 当 A 取单位矩阵 I 时对应于欧几里德(欧氏)距离。图3-1说明了模糊C聚类的直观意义:

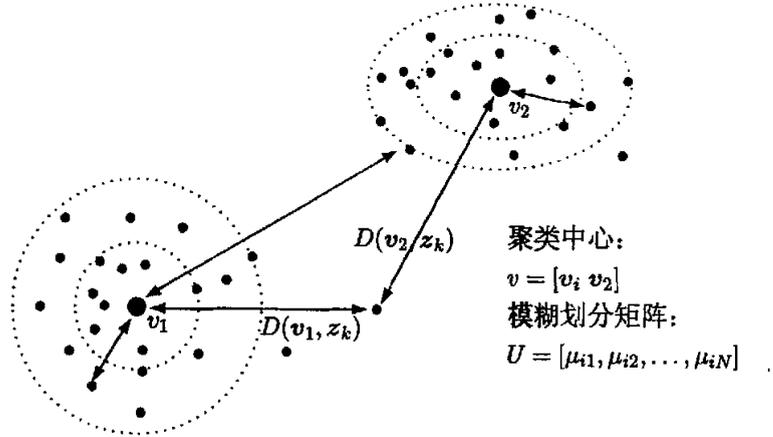


图 3-1: 模糊聚类

从统计学意义上讲, 式(3-11)可以看作求 z_k 到 v_i 总误差的最小值。聚类准则(3-11)的求解是一个非线性优化问题, 解决这类优化问题有很多方法, 如群协同最小化、过拟退火、遗传算法等。最常用的方法就是简单的Picard迭代法, 即所谓的模糊C-均值聚类 (FCM) 法。

聚类推则(3-11)可以用拉格朗日乘子来求解:

$$\bar{J}(Z; U, V, \lambda) = \sum_{i=1}^M \sum_{k=1}^N (\mu_{ik})^m D_{ikA}^2 + \sum_{k=1}^N \lambda_k (\sum_{i=1}^M \mu_{ik} - 1) \quad (3-13)$$

若, $D_{ikA}^2 > 0, \forall i, k$ 和 $m > 1$, 则 $(U, V) \in M_{fc} \times R^{n \times M}$ 。式3-11取最小值的必要条件是, \bar{J} 对 U, V, λ 的一阶微分为0, 即:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^M (D_{ikA}/D_{jkA})^{2/(m-1)}}, \quad 1 \leq i \leq M, 1 \leq k \leq N \quad (3-14)$$

和

$$v_i = \frac{\sum_{k=1}^N \mu_{ik}^m x_k}{\sum_{k=1}^N \mu_{ik}^m}, \quad 1 \leq i \leq M \quad (3-15)$$

式(3-15)说明 v_i 是聚类中数据的加权平均值, 权值就是它们的隶属度。因此这种算法叫做“C-平均”。另外, FCM算法就是式(3-15)和式(3-14)间的迭代。

FCM的详细算法如下[28]:

初始化参数：对于给定的数据集 Z ，选择聚类数 $1 < M < N$ ，权值数 $m > 1$ ，中止阈值 $\varepsilon > 0$ ，内模积矩阵 A 。随机初始化划分矩阵，满足 $U^{(0)} \in M_{fc}$ 。

进行下面的迭代过程：

步骤1 计算聚类中心：

$$v_i^{(q)} = \frac{\sum_{k=1}^N (u_{ik}^{(q-1)})^m x_k}{\sum_{k=1}^N (u_{ik}^{(q-1)})^m}, \quad 1 \leq i \leq M \quad (3-16)$$

步骤2 计算距离：

$$D_{ikA}^2 = (z_k - v_i)^T A (z_k - v_i), \quad 1 \leq i \leq M, 1 \leq k \leq N \quad (3-17)$$

步骤3 更新划分矩阵：

$$\mu_{ik}^{(q)} = \frac{1}{\sum_{j=1}^M (D_{ikA} / D_{jkA})^{2/(m-1)}} \quad (3-18)$$

步骤4 如果

$$\|U^{(q)} - U^{(q-1)}\| < \varepsilon \quad (3-19)$$

则停止，并输出划分矩阵 U 和聚类原型 V ，否则令 $q = q + 1$ 返回步骤1。

□

由以上算法不难看出，整个计算过程就是反复修改聚类中心和分类矩阵的过程，因此常称这种方法为动态聚类或者逐步聚类法。

3.4 Gath-Kessel(GK)聚类

与模糊C均值聚类不同的是，在GK聚类法中使用一种自适应的距离度量，使其更适合辨识出数据的特征划分区域。GK聚类法中距离 D_{ik}^2 定义为：

$$D_{ik}^2 = (z_k - v_i)^T \{[\det(F_i)]^{1/(n+1)} F_i^{-1}\} (z_k - v_i), \quad (3-20)$$

$$i = 1, \dots, M, \quad k = 1, \dots, N$$

式中， F_i 为聚类的模糊协方差：

$$F_i = \frac{\sum_{k=1}^N (\mu_{ik})^m (z_k - v_i)(z_k - v_i)^T}{\sum_{k=1}^N (\mu_{ik})^m}, \quad i = 1, \dots, M \quad (3-21)$$

GK聚类方法如下：

初始化参数，与FCM方法相同。

步骤1计算聚类中心：

$$v_i^{(q)} = \frac{\sum_{k=1}^N (u_{ik}^{(q-1)})^m z_k}{\sum_{k=1}^N (u_{ik}^{(q-1)})^m}, \quad i = 1, \dots, M \quad (3-22)$$

步骤2计算协方差矩阵：

$$F_i = \frac{\sum_{k=1}^N (\mu_{ik}^{(q-1)})^m [z_k - v_i^{(q)}][z_k - v_i^{(q)}]^T}{\sum_{k=1}^N (\mu_{ik}^{(q-1)})^m}, \quad i = 1, \dots, M \quad (3-23)$$

步骤3计算距离：

$$D_{ik}^2 = (z_k - v_i^{(q)})^T (\det(F_i)^{1/(n+1)} F_i^{-1}) (z_k - v_i^{(q)}), \quad (3-24)$$

$$i = 1, \dots, M, \quad k = 1, \dots, N$$

步骤4更新划分矩阵：

$$\mu_{ik}^{(q)} = \begin{cases} \frac{1}{\sum_{j=1}^M (D_{ik}/D_{jk})^{2/(m-1)}}, & \text{若 } D_{ik} > 0 \\ 0, & \text{其它} \end{cases} \quad (3-25)$$

且，当 $D_{ik} > 0$ 时 $\mu_{ik}^{(q)}$ 须满足约束体条件

直至

$$\|U^{(q)} - U^{(q-1)}\| < \varepsilon \quad (3-26)$$

则停止，否则令 $q = q + 1$ 返回步骤1。

□

3.5 模糊系统的设计

每个聚类代表系统的一个特征域，聚类数目 M 等于规则数目。由划分矩阵 U 可以得到规则前提部分的模糊集 A_i^j 。前面已说明，元素 $\mu_{ik} \in [0, 1]$ 为数据点 z_k 属于聚类 i 的隶属度。 U 的第 i 行包含一个定义在多维模糊子集上的一个点。将多维模糊子集投影到输入变量 z_k 的空间上就得到规则(2-8)中的单变量模糊集 A_i^j ：

$$\mu_{A_i^j}(x_{ik}) = \text{proj}_i(\mu_{ik}) \quad (3-27)$$

式中， proj 为点投影操作。如图3-2。

然后可利用第4章的方法辨识结论参数。

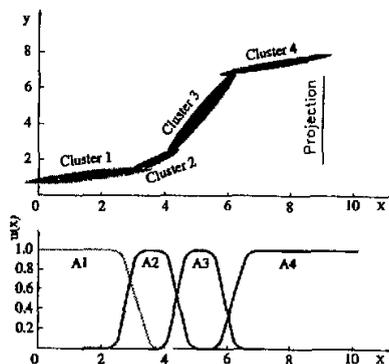


图 3-2: 模糊聚类的投影得到规则前提中的模糊集

3.6 合并相似隶属函数

模糊聚类能有效减少模糊规则数目，

模糊的系统的最大优点体现在它们的语言可解释性。然而这个问题经常被忽略，人们的大多将注意力放在模型的精确性上。正如Nauck和Kruse所说^[29]，如果模型的可解释性不是我们要关注的主要问题，还不如放弃模糊建模方法，使用其它方法得到的结果可能更准确。

为了保证模型的可解释性，第一步就是要确定、合并隶属函数的相似规则。使用聚类法得到的结构中存在许多含相似的隶属函数。这样模型的可解释性很差，很对参数的调整显得是多余的，需要消耗大量的计算资源。将非常相似的规则合并就很有意义了。

这里对规则库的简化主要是建立在隶属函数的相似分析上的，有很多文献提出了不同的相似性测量方法，这里介绍下面这种相似测度方法：

假设A和B是2个模糊集，模糊集A 和B 的相似度定义为：

$$S(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (3-28)$$

式中， $|\cdot|$ 为集合的基数 (cardinality)，。很明显，计算两个模糊集的相似度就是计算两个模糊集的交集和并集的比。当定义在离散域 $D = \{x_j | j = 1, 2, \dots, m\}$ 上时，可将式(3-28)写成隶属函数的形式为：

$$S(A, B) = \frac{\sum_{j=1}^m [\mu_A(x_j) \wedge \mu_B(x_j)]}{\sum_{j=1}^m [\mu_A(x_j) \vee \mu_B(x_j)]} \quad (3-29)$$

\wedge 和 \vee 是最小和最大算子。如果 $S(A, B) = 1$ 则两个模糊集合A, B是相等的。当两个

集合没有重合部分的时候 $S(A, B) = 0$ 。 S 越大则认为两个模糊集合 A, B 的相似度越大。

这里设定两个阈值 $\theta_u \in (0, 1)$ (集合与全集的相似度阈值), $\theta_m \in (0, 1)$ (两个集合间相似度阈值)。

反复合并两两相似的模糊集, 测量每个输入变量的所有模糊集间的相似度。模糊集间的相似度 $S > \theta_m$ 则将它们合并。两个模糊集合并后作为一个新的模糊集, 减少了规则库中的模糊集, 更新规则库, 计算更新规则库中模糊集间的相似度, 直到没有模糊集间的相似度 $S > \theta_m$ 为止。这样可以减少模型中模糊集合的数目, 同时可以增加模型的透明性。最后, 检查更新规则库中的模糊集中是否存在与全集相似的集合, 若存在则将它们从规则前提中删除, 如图3-3所示。

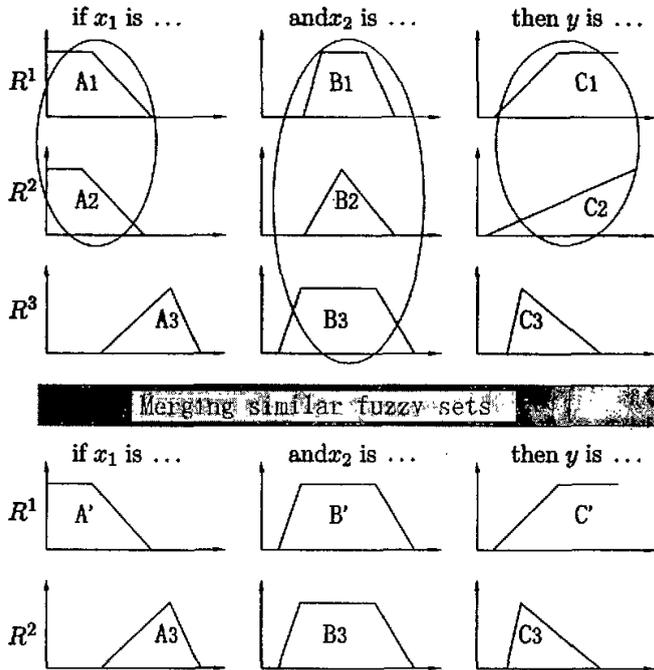


图 3-3: 相似集合的合并

此算法总结为:

已知初始模糊规则库 $R = \{R_i | i = 1, \dots, N\}$:

$$R^i: \text{if } x_1 \text{ is } A_1^i \text{ and } \dots x_n \text{ is } A_n^i \text{ then } y^i \quad (3-30)$$

式中, $A_i^l (l = 1, \dots, M)$ 为模糊集, 它的隶属度为 $\mu_{A_i^l} : x_i \rightarrow [0, 1]$ 。选择两个阈值 $\theta_u, \theta_m \in (0, 1)$ 。

Repeat:

1. 选择规则库中两个最相似的模糊集。对每个输入变量 $x_i, i = 1, \dots, n$ 的所有模糊集 A_i^1, \dots, A_i^M , 计算 $S_i^{j,k} = S(A_i^j, A_i^k), j, k = 1, \dots, M$ 。选择其中相似度最大的两个集合 A_i^p, A_i^q , 即 $S_i^{p,q} = \max_{j \neq k} \{S_i^{j,k}\}$ 。
2. 合并两个最相似的规则, 更新规则库。如果 $S_i^{p,q} \geq \theta_m$, 则合并集合 A_i^p, A_i^q 得到新的模糊集 A , 在相应的模糊规则前提中用替换它们 $A_{ql} = A, A_{qm} = A$ 。

Until: $S_i^{p,q} < \theta_m$ 。

删除与全集相似的模糊集。对每个模糊集 A_i^l 计算 $S(A_i^l, U), \mu_U(x) = 1, \forall x$, 如果 $S(A_i^l, U) \geq \theta_u$, 则 A_i^l 将从 R^l 中删除。

□

这种方法中每次只合并一对模糊集, 对不同变量的模糊集的操作是相互独立的。下面将讨论一个细节问题, 即两个集合如何合并为一个新的集合。在^[30]中, 使用的合并方法是取两个模糊集合的并集作为新的集合, 即 $A = A_i^p \cup A_i^q$ 。这种方法操作简单, 但是它会使得模糊集的支撑集越来越大, 最后可能使很多模糊集都与全集相似, 与初始模糊集相去甚远, 影响模型的精度。

这里采取一种“中庸”之法, 使新的模糊集在两个模糊集之间, 如图3-4, 虚线表示合并前的两个模糊集 A, B , 实线表示合并后得到的模糊集 C 。

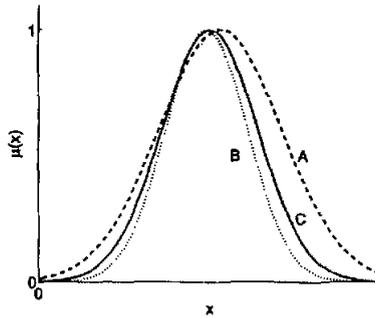


图 3-4: 相似集合的合并

具体来说, 赋给集合 A, B 一个合并权值, 然后将它们合并:

$$C_p = \frac{n_A A_p + n_B B_p}{n_A + n_B} \quad (3-31)$$

新模糊集 C 的参数是由初始函数的加权平均得到的。公式(3-31)中, n_A, n_B 分别表

示集合 A, B 经过了合并的次数。如 A 是初始集合则 $n_A = 1$ ； A 是由两个初始集合合并得到的，则 $n_A = 2$ ，如此类推。 C_p, A_p, B_p 是定义模糊集 C, A, B 的参数向量。

例如，取模糊集 A, B 为式(4-4)形式的高斯函数， $n_A = 2, n_B = 3$ ，这两个模糊集合并后得到模糊集 C 的参数为：

$$\begin{aligned} C_c &= \frac{n_A A_c + n_B B_c}{n_A + n_B} = \frac{2A_c + 3B_c}{5} \\ C_\sigma &= \frac{n_A A_\sigma + n_B B_\sigma}{n_A + n_B} = \frac{2A_\sigma + 3B_\sigma}{5} \end{aligned} \quad (3-32)$$

第四章 TS模糊模型网络系统的参数辨识

引入TS模型数学表达式的目的就是为了利用各种学习方法调节模型的参数。下面将讨论几种TS模糊模型参数辨识算法。

神经网络是受人脑组织的生理学知识启发而创立的。它们是由一系列互相联系的、相同的、单元(神经元)组成的。相互间的联系可以在不同的神经元之间传递增强或抑制信号。增强或抑制是通过调整相互间联系的权重系数(weight)实现的。神经网络的性质中最有意义的是它的学习能力。它通过调整突触的权值和阈值的迭代过程从周围环境进行学习以不断地改进自身的性能。在神经网络中,可将“学习”定义为:神经网络参数对来自周围环境的激励不断适应的过程。学习的形式由参数变化的方法所确定。本节主要论述模糊推理网络系统的结构,然后阐述禁忌搜索算法和最小二乘法的混合算法在模糊神经网络参数辨识中的应用。

4.1 模糊推理网络系统的结构

将TS模型的数学模型转化为网络结构,将公式2-9中的算子和函数看作网络中的节点。为了说明方便首先在这里考虑一个多输入单输出的一阶TS模型,假设这个模型有两条模糊规则:

$$R^i: \text{ if } x_1 \text{ is } A_1^i \text{ and } \dots \text{ and } x_n \text{ is } A_n^i \text{ then } y^i = a_i^T x + b_i \quad (4-1)$$

图4-1解释了这种TS模型的推理机制;该模型相应等效的网络结构如图4-2。

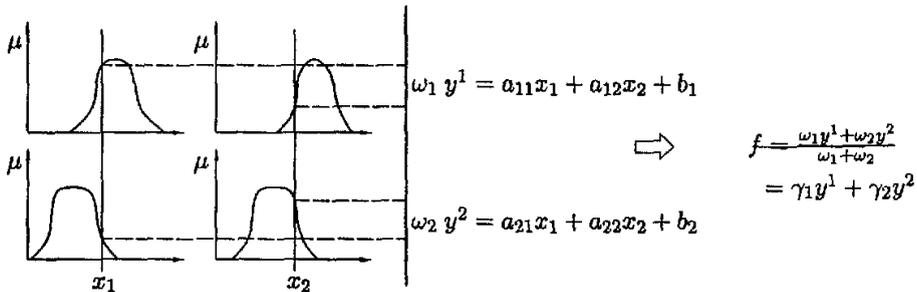


图 4-1: TS模型的推理过程

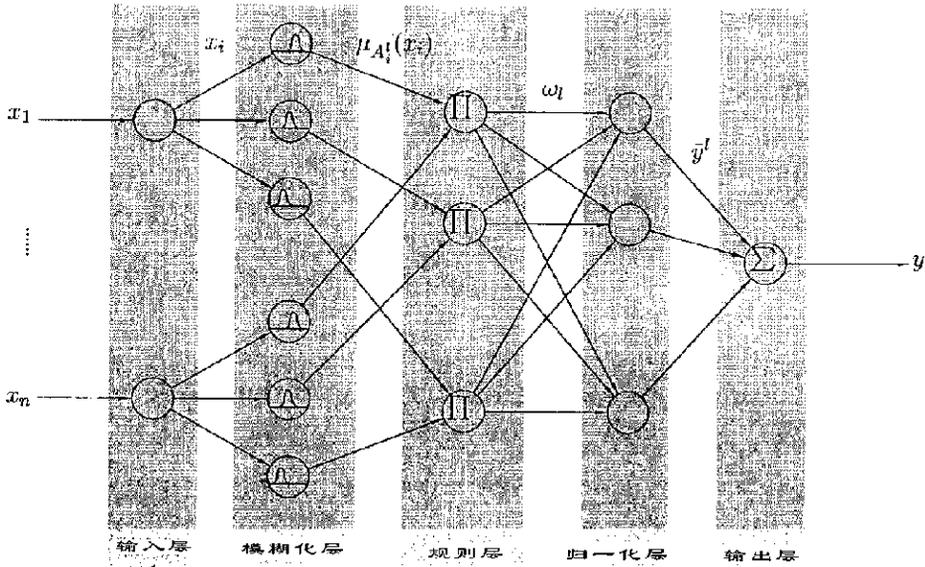


图 4-2: TS模型的网络结构

下面将要说明该网络结构。

Layer 1: 输入层, 从外界接收数据并将它们传递到下一层。 $X^T = [x_1, x_2]$ 为网络的输入信息, 网络的节点输出为 $O_i^{(1)}$:

$$O_i^{(1)}: x_i \quad (4-2)$$

式中, $i = 1, 2$ 代表输入变量 x_1, x_2 。

Layer 2: 模糊化层, 每个节点表示一个模糊语言变量, 如“小”, “中”, “大”等, 节点的激发函数可以用任意合适的参数化隶属函数表示。输出为各输入 x_i 分别满足语言变量 $A_i^l (l = 1, \dots, M)$ 的隶属度 $\mu_{A_i^l}(x_i)$:

$$O_{i,l}^{(2)}: \mu_{A_i^l}(x_i) \quad (4-3)$$

这里的隶属函数可以是任意合适的参数化隶属函数, 如一般的高斯函数:

$$\mu_{A_i^l}(x_i; c_i^l, \sigma_i^l) = \exp\left(-\frac{(x_i - c_i^l)^2}{2(\sigma_i^l)^2}\right) \quad (4-4)$$

式中, $\{c_i^l, \sigma_i^l\}$ 是参数集, 分别表示第 l 个输入变量的第 l 个隶属函数的中心和标准差, 当这些参数的值改变时, Gaussian 函数也随之改变, 这样实现了模糊集隶属函数形式的改变。本层参数称为前提参数。

Layer 3: 规则层, 每个结点表示一条规则, 是一个标以 Π 的固定结点, 它用t-范数将规则前提的各部分连接起来, 每个结点的输出表示一条规则的激励强度。通常取代数乘积算子, 这样得到的模型曲面光滑, 而且可以直接求导, 它的输出是所有输入信号的积:

$$O_i^{(3)}: \omega_i = \prod_{i=1}^n \mu_{A_i^l}(x_i) \quad (4-5)$$

一般来说, 本层的结点函数可以用任意其它执行模糊“与”的T范式算子。

Layer 4: 归一化层, 每个节点表示输出变量的一种线性回归方程, 输出为该方程的激励强度:

$$O_i^{(4)}: \bar{y}^i = \frac{\omega_i y^i}{\sum_{l=1}^M \omega_l} = \frac{\omega_i (a_i^T x + b_i)}{\sum_{l=1}^M \omega_l} \quad (4-6)$$

式中, $\{a_{il}, b_i\}, (i = 1, 2, \dots, n)$ 是该节点的参数集。本层的参数称为结论参数。

Layer 5: 输出层, 每个结点是一个标以 Σ 的固定结点, 它计算所有传来信号之和作为总输出:

$$O^{(5)}: y = \sum \bar{y}^i = \frac{\sum_i \omega_i y^i}{\sum_i \omega_i} \quad (4-7)$$

这样, 我们建立了一个功能上与TS模糊模型等价的自适应网络:

$$y = \sum_{l=1}^M \gamma_l(x) (a_l^T x + b_l) \quad \text{其中, } \gamma_l(x) = \frac{\prod_{i=1}^n \exp\left(-\frac{(x_i - c_i^l)^2}{2(\sigma_i^l)^2}\right)}{\sum_{l=1}^M \prod_{i=1}^n \exp\left(-\frac{(x_i - c_i^l)^2}{2(\sigma_i^l)^2}\right)} \quad (4-8)$$

4.2 误差反向传播算法

模糊系统的网络模型有5层, 这里可以把它看作一个多层感知器 (MLP), 下面将利用MLP的学习方法来训练这个网络模型。即与误差修正学习算法所发展的误差反向传播 (BP) 算法使多层感知器成为有监督学习方法中一个非常有效的工具。

反向传播算法的思想是由Werbos于1974年在他的博士论文中提出。误差反向传播的过程包括正向通过网络的各个层和反向通过网络的各个层。正向通过时, 输入向量作用于网络的输入节点, 通过网络逐层传播其影响, 此时, 网络的突触权值是固定不变的。反向通过时, 所有的权值都按规定误差修正规则加以调整, 使网络的实际响应最接近要求的响应。

4.2.1 误差修正学习

考虑一个多输入单输出的神经网络，有 N 对输入输出数据样本 $\{(\mathbf{x}_k, \mathbf{y}_k^*) | k = 1, 2, \dots, N\}$ 。将这个样本集写成矩阵形式，输入矩阵 $\mathbf{X} \in \mathbb{R}_{N \times n}$ ，它的每一行为输入向量 \mathbf{x}_k 。输出列向量 $\mathbf{y}^* \in \mathbb{R}_N$ ， \mathbf{y}_k^* 为系统的实际输出值：

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]^T, \quad \mathbf{y}^* = [y_1^*, y_2^*, \dots, y_N^*]^T \quad (4-9)$$

误差为实际输出 \mathbf{y}^* 与网络输出 \mathbf{y} 的差，网络输出是权值 W 的非线性函数（为了表述方便将所有的权值写成一个向量 W ）。目标函数定义为：

$$J(W) = \frac{1}{2} \sum_{k=1}^N e_k^2 \quad \text{其中 } e_k = y_k^* - y_k \quad (4-10)$$

网络的训练问题就是一个非线性优化问题，可以有多种方法：

- 误差反向传播（一阶梯度下降）
- 牛顿法及牛顿修正法（二阶梯度下降）
- 其他非倒数优化方法，免疫算法、模拟退火、随机搜索方法等。

一阶梯度下降法 一阶梯度下降法使用以下方式学习权值 W ：

$$W(q+1) = W(q) - \alpha(q) \nabla J(W(q)) \quad (4-11)$$

式中， $q = 1, 2, \dots$ ， $W(q)$ 是第 q 次迭代时的权值向量， $\alpha(q)$ 是一个可变的学習速率常数， $\nabla J(W)$ 是网络的Jacobian阵：

$$\nabla J(W) = \left[\frac{\partial J(W)}{\partial W_1}, \frac{\partial J(W)}{\partial W_2}, \dots, \frac{\partial J(W)}{\partial W_M} \right]$$

非线性优化问题的求解是借助目标函数的一阶泰勒展开式，所以称为一阶梯度下降法。

二阶梯度下降法 同样地也可以利用目标函数的二阶泰勒展开式求解，即二阶梯度下降法：

$$J(W) \approx J(W_0) + \nabla J(W_0)^T (W - W_0) + \frac{1}{2} (W - W_0)^T \mathbf{H}(W_0) (W - W_0)$$

式中， $\mathbf{H}(W_0)$ 为海森矩阵（Hessian），由 $J(W)$ 的二次偏导数构成。经过几步迭代之后，参数 W 的迭代公式为：

$$W(q+1) = W(q) - \alpha(q) \mathbf{H}^{-1}(W(q)) \nabla J(W(q)) \quad (4-12)$$

式(4-11)与式(4-12)的本质区别是，梯度下降的步长和方向不同，如图4-3所示。

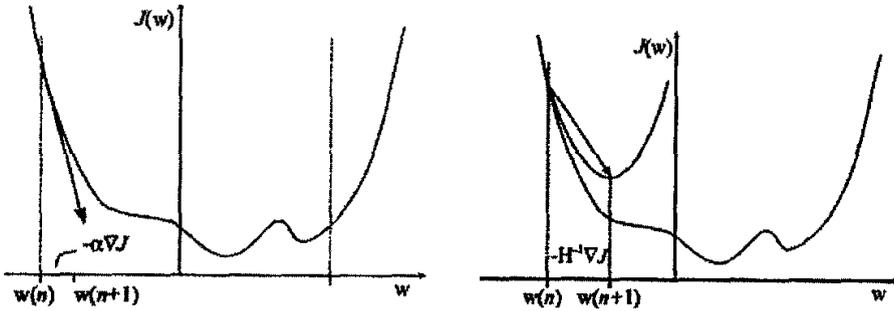


图 4-3: 一阶和二阶梯度优化

模糊推理网络系统的误差反向传播学习 通常二阶方法比一阶方法更有效一些, 但一阶方法更简便一些。这里首先来看看模糊推理网络模型的误差反向传播方法(一阶梯度下降)。

结论参数的学习 从输出层的权值开始, 它的Jacobian为:

$$\frac{\partial J}{\partial b_l} = \frac{\partial J}{\partial e} \cdot \frac{\partial e}{\partial y} \cdot \frac{\partial y}{\partial b_l} = -\gamma_l e, \quad l = 1, 2, \dots, M \quad (4-13)$$

代入式(4-11), 得到输出层权值的迭代式:

$$b_l(q+1) = b_l(q) + \alpha(q)\gamma_l e \quad (4-14)$$

同样地有:

$$a_{il}(q+1) = a_{il}(q) + \alpha(q)\gamma_l x_i e \quad (4-15)$$

前提参数的学习 前提参数为Gaussian函数的中心和宽度。由图4-1可知, y (于是 e 也如此) 仅通过 ω_l 依赖于 c_i^l 。所以根据复合函数求导规则, 有:

$$\begin{aligned} \frac{\partial J}{\partial c_i^l} &= \frac{\partial J}{\partial e} \cdot \frac{\partial e}{\partial y} \cdot \frac{\partial y}{\partial \omega_l} \cdot \frac{\partial \omega_l}{\partial c_i^l} \\ &= e \frac{y^l - y}{b} \omega_l \frac{2(x_i - c_i^l)}{(\sigma_i^l)^2} \\ &= 2e\gamma_l(y^l - y) \frac{x_i - c_i^l}{(\sigma_i^l)^2} \end{aligned} \quad (4-16)$$

将上式代入式4-11, 得到前提参数的迭代式:

$$c_i^l(q+1) = c_i^l(q) + 2\alpha(q)e\gamma_l(y^l(q) - y) \frac{x_i - c_i^l(q)}{(\sigma_i^l(q))^2} \quad (4-17)$$

同样地有:

$$\sigma_i^l(q+1) = \sigma_i^l(q) + 2\alpha(q)e\gamma_l(y^l(q) - y) \frac{(x_i - c_i^l(q))^2}{(\sigma_i^l(q))^3} \quad (4-18)$$

学习算法(4-14)、(4-15)、(4-17)、(4-18)完成的是一个误差反向传播程序。为了训练 b_l, a_{il} , 误差 e 被反向传播到 b_l, a_{il} 所在层; 则 b_l, a_{il} 可分别用式(4-14)、(4-15)来调整。为了训练 c_i^l, σ_i^l , 标推误差 e 与 γ_l 及 $y^l - y$ 的乘积被反向传播到Layer 2的处理单元; 则 c_i^l, σ_i^l 可分别用式(4-17)、(4-18)来调整。这样就可以实现网络参数的不断修正, 使模型的参数达到最优。因此, 也称这一算法为误差反向传播BP学习算法(error back-propagation training algorithm)。

现在, 将这一算法总结如下:

1. 结构的确定和初始参数的设置。选择形如式4-1的模糊规则并确定 M 。 M 越大, 产生的参数越多, 运算也就越复杂, 但结出的逼近精度越高。设定初始参数 $c_i^l(0), \sigma_i^l(0), a(0)_l, b(0)_l$ 。这些初始参数可能是根据专家的语言规则确定的, 也可能是由均匀地覆盖输入-输出空间的相应的隶属度函数确定的。对于数据对较复杂的问题可以先用第3章的模糊聚类法来设计初始模型系统。
2. 给出输入数据并计算模糊系统的输出。对于给定的输入-输出数据对 (x_k, y_k^*) , $k = 1, 2, \dots$, 在学习的第 q ($q = 0, 1, 2, \dots$)阶段, 把输入 x_k 作为图4-1中的模糊系统的输入层, 然后根据式(4-2)-(4-7)计算第1-5层的输出。
3. 调整参数。采用学习算法(4-14)、(4-15)、(4-17)和(4-18)计算要调整的参数 $b_l(q+1), a_{il}(q+1), c_i^l(q+1), \sigma_i^l(q+1)$ 。
4. 令 $q = q+1$, 返回步骤2重新计算。直至误差 e 计小于一个很小的数 ϵ 或直至 q 等于一个预先指定的。
5. 今 $p = p+1$, 返回步骤2重新计算。即, 用下一个输入-输出数据对 (x_{k+1}, y_{k+1}^*) 来调整参数。
6. 如果有必要的活, 令 $p = 1$, 并重新计算步骤2-步骤5直至所设计的模糊系统令人满意。对于在线控制和动态系统辨识问题, 这一步是不可行的, 因为该问题给出的输入-输出数据对是以实时方式一一对应的; 而对于模式识别问题, 因为其输入-输出数据对是离线的, 所以通常这一步是可行的。

□

4.2.2 仿真实例及结果分析

用上述方法来逼近一个静态函数 $y = \sin(x)$, 结论部分使用一阶线性函数。根据 $y = \sin(x)$ 函数曲线近似选择以下五条规则:

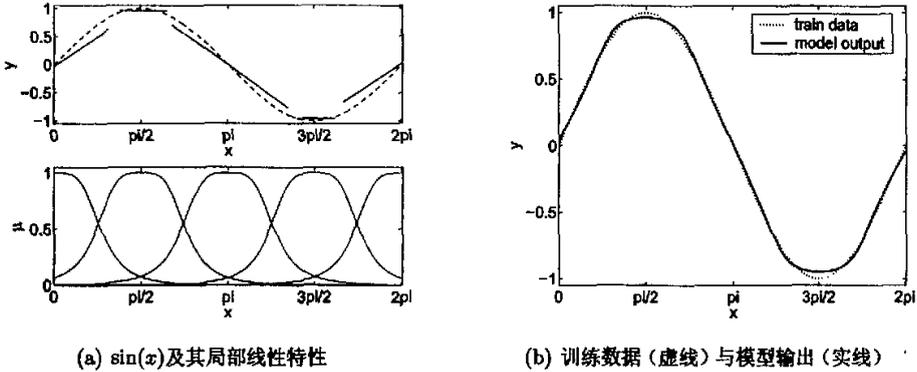


图 4-4: 初始TS模型的逼近特性

$$\begin{aligned} \text{if } x \text{ is } A_1 \text{ then } y &= 0.67x - 0.03 \\ \text{if } x \text{ is } A_2 \text{ then } y &= 0.001x + 0.96 \\ \text{if } x \text{ is } A_3 \text{ then } y &= -0.7x + 2.2 \\ \text{if } x \text{ is } A_4 \text{ then } y &= -0.001x - 0.95 \\ \text{if } x \text{ is } A_5 \text{ then } y &= 0.67x - 4.18 \end{aligned}$$

如图4.4(a)所示, 为5个局部线性函数和隶属函数, 从左到右依次为 $A_1 \sim A_5$ 。图4.4(b)为这个模糊模型的输出曲线。

可以看出这个模型具有很好的可解释性, 它由一个非线性函数的5个局部近似线性函数组成。它却具有较好的逼近精度(均方根误差 $RMSE^1=0.5230$)。使用上面的混合学习方法调整模型的参数, 将使逼近效果更好。经过100次迭代后得到的模糊规则如下:

$$\begin{aligned} \text{if } x \text{ is } A'_1 \text{ then } y &= 0.94x - 0.085 \\ \text{if } x \text{ is } A'_2 \text{ then } y &= 0.01x + 0.95 \\ \text{if } x \text{ is } A'_3 \text{ then } y &= -0.91x + 2.83 \\ \text{if } x \text{ is } A'_4 \text{ then } y &= 0.01x - 1.02 \\ \text{if } x \text{ is } A'_5 \text{ then } y &= 0.92x - 5.67 \end{aligned}$$

模型输出如图4.5(a), 从图上已经看不出明显的误差。均方根误差 $RMSE=0.0020$, 比初始模型的精度提高了260倍。可以看到对隶属函数只是做了较小的改动, 如图4.5(b), 但是结果的精度得到了很大改善。

¹本文中用均方根误差 (RMSE) 来衡量模型的精度: $RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^N (y_k^* - y_k)^2}$

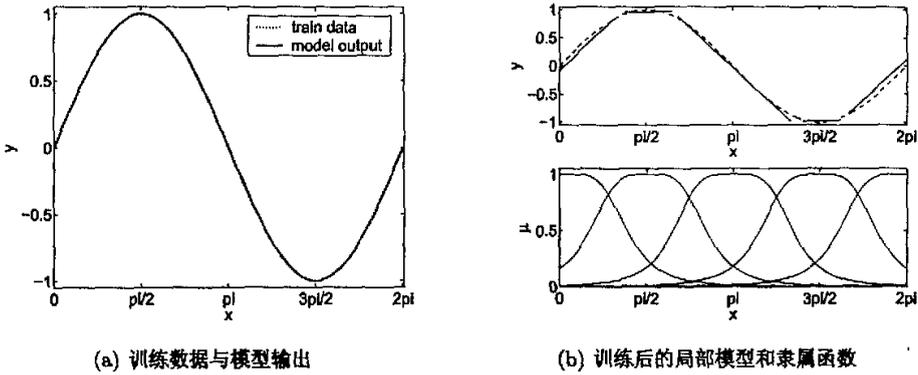


图 4-5: 训练后的TS模型的逼近特性

4.3 混合学习算法

观察图4-2所示的模糊推理网络系统结构, 我们得到, 当固定前提隶属函数的参数值时, 系统总的输出可以表示为结论参数的线性组合, 即将式4-8写为:

$$\begin{aligned}
 y &= \gamma_1 y^1 + \gamma_2 y^2 + \dots + \gamma_M y^M \\
 &= \left(\sum_{l=1}^M \gamma_l \right) b_l + \left(\sum_{l=1}^M \gamma_l x_1 \right) a_{l1} + \dots + \left(\sum_{l=1}^M \gamma_l x_n \right) a_{ln}
 \end{aligned} \tag{4-19}$$

它是结论参数 $a_{li}, b_l (l = 1, 2, \dots, M)$ 的线性函数。

这样处理的好处是将参数优化问题简化为一个线性优化问题, 可以使用效率更高的线性优化算法。所以把 γ_l 看作已知函数, y^l 看作待估计的未知参数, 并用最小二乘法辨识结论参数。在反向通道中, 误差信号反传, 并用梯度法更新前提参数^[31]。下面首先讨论最小二乘估计的数学表达方法。

4.3.1 最小二乘估计

在一般的最小二乘问题中, 线性模型 y 的输出是参数化的线性表达式^[32]:

$$y = f(x) = \sum_{i=1}^m w_i \phi_i(x) \tag{4-20}$$

其中 $x = [x_1 \ x_2 \ \dots \ x_p]^T$ 是模型的输入向量, ϕ_1, \dots, ϕ_n 是 x 的已知函数, w_1, \dots, w_n 是待估计的未知参数。

为了辨识未知参数 w_i , 将数据对 $\{(x_i; y_i^*), i = 1, \dots, N\}$ 代入方程(4-20)产生一

组 m 个线性方程,使用矩阵表示为:

$$Vw = y \quad (4-21)$$

为了唯一地识别出未知向量,要求 $m \geq n$ 。如果 A 是非奇异的方阵,则从式(4-21)可求得 w :

$$w = V^{-1}y \quad (4-22)$$

但是,通常 $m > n$,表明数据对的数目多于拟合参数的数目。在这种情况下,满足所有 m 个方程的精确解是不可能的,因为数据可能会受到噪声的污染,或者模型有可能不适合描述目标系统。因此,必须修改方程(4-21),加入误差向量 e ,以便对随机噪声或建模误差加以考虑:

$$Vw + e = y \quad (4-23)$$

现在,需搜索使如下定义的误差平方和最小的 $w = \hat{w}$,而不是求取式(4-21)的精确解:

$$E(w) = \sum_{i=1}^m (y_i^* - \phi_i w)^2 = e^T e = (y^* - Vw)^T (y^* - Vw) \quad (4-24)$$

其中 $e = y^* - Vw$ 是特定取值所产生误差向量。 $E(w)$ 是二次型,且有唯一的最小值 $w = \hat{w}$ 。如果 $V^T A$ 是非奇异,最小值 \hat{w} 由下式唯一确定:

$$\hat{w} = (V^T V)^{-1} V^T y^* \quad (4-25)$$

在误差最小化方面任意最小二乘估计法没有明显的差别。但如果要考虑结论参数对局部模型可解释性的影响,则要慎重考虑方法的选择。ANFIS模型,特别是一阶模型,对大多逼近问题出现拟合参数过多的问题。这带来过拟合和无意义参数的问题。下面这个例子将说明这个问题。

例4.1 假设我们用一个一阶ANFIS模型逼近一个二阶多项式 $y = f(u) = 3u^2 - 5u + 6$ 。首先选择两点 t_1, t_2 ,并定义初始三角隶属函数 $t_1 \leq u \leq t_2$:

$$\mu_{A_1}(u) = \frac{u - t_1}{t_2 - t_1}, \quad \mu_{A_2}(u) = 1 - \mu_{A_1}(u) \quad (4-26)$$

模型由两条规则组成:

$$R^i: \text{ if } u \text{ is } A^i \text{ then } y^i = a_i u + b_i, \quad i = 1, 2$$

将隶属函数式4-26代入2-9, 得到TS模型的输出为:

$$y = \frac{a_1 - a_2}{t_1 - t_2} u^2 + \frac{t_2 a_2 - t_1 a_1 + b_1 - b_2}{t_1 - t_2} u + \frac{t_2 b_2 - t_1 b_1}{t_1 - t_2}$$

这是关于 u 的一个二阶多项式, 这个模型可以很好地表达给定的非线性系统。但是它有4个自由参数 (a_1, a_2, b_1, b_2) , 事实上3个参数已足够表达这个多项式, 出现了过多的参数。这是一个非常简单的例子, 过多参数的本质在未知的复杂系统的逼近问题中同样会出现。

□

可以增加局部拟合的标准, 或约束参数值解决过参数的问题。下面讨论几种方法:

4.3.2 全局最小二乘估计

对于TS模型(4-8)全局最小二乘估计法计算参数使下面的误差标准最小化^[33]:

$$\begin{aligned} J_G &= \arg \min \sum_{k=1}^N (y_k^* - y_k)^2 \\ &= \arg \min \sum_{k=1}^N \left(y_k^* - \sum_{l=1}^M \gamma_l(x_k) [\mathbf{x}_k^T \mathbf{1}] \theta_l \right)^2 \end{aligned}$$

式中 $\theta_l = [a_l \ b_l]^T$ 是所有结论参数的向量, N 为训练数据的数目。对于输入-输出数据对(4-9), 将这个标准写为矩阵的形式:

$$\begin{aligned} J_G &= \arg \min (\mathbf{y}^* - \Lambda \theta)^T (\mathbf{y}^* - \Lambda \theta) \\ \Lambda &= [\Gamma_1 \varphi \ \dots \ \Gamma_M \varphi] \end{aligned} \quad (4-27)$$

式中, $\mathbf{y}^* = [y_1^* \ y_2^* \ \dots \ y_N^*]^T \in \mathbb{R}^N$, $\varphi = [X \ 1]$, $\Gamma_l = \text{diag}(\gamma_l(x_1) \ \dots \ \gamma_l(x_N))$, 即一个以 $\gamma_l(x_k)$ 为对角矩阵的第 k 个对角元素。

因为前提参数暂时是固定的, J_G 中的未知参数只有参数向量 θ , 它的元素是TS模型线性回归方程的参数。所以式(4-27)是一个线性优化问题, 它的最优解可通过下式求解:

$$\theta = (\Lambda^T \Lambda)^{-1} \Lambda^T \mathbf{y}^* \quad (4-28)$$

4.3.3 局部最小二乘估计

全局最优解得到的是最小预测误差, 但可能会出现局部模型的参数估计偏差较大。要得到局部模型的最优参数, 则应使用最小二乘法估计每条规则的参数。可以利用 K 个局部LS标准:

$$J_l = \arg \min (\mathbf{y}^* - \varphi \theta_l)^T \Gamma_l (\mathbf{y}^* - \varphi \theta_l), \quad l = 1, 2, \dots, M \quad (4-29)$$

它的最优解为：

$$\theta_l = (\varphi^T \Gamma_l \varphi)^{-1} \varphi^T \Gamma_l y^*, \quad l = 1, 2, \dots, M \quad (4-30)$$

在这种情况下单独估计每条规则的结论参数，所以结论不受规则相互间的影响。但模型的最终预测误差要比全局最小二乘法的大。

例4.2考虑例4.1的TS模型，分别使用局部和全局最小二乘估计，最后得到的模型如图4-6。注意到局部最小二乘法能很好描述函数的局部特性，但拟合精度较差。全局最小二乘估计则相反，较好的精度，但结果与系统的局部特性不相干。

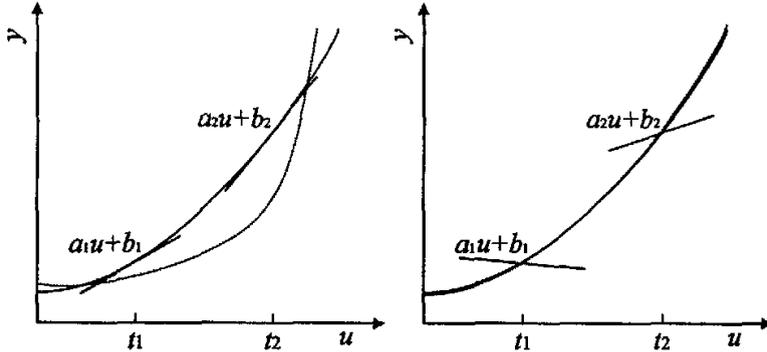


图 4-6: 局部(a)和全局(b)最小二乘法估计结论参数得到的结果。虚线为模型的输出。

□

4.3.4 多目标优化

全局LSE将所有规则的结论参数作为一个最小二乘问题，局部LSE并不同步顾及所有的参数，而是将参数优化问题分解为M（模糊规则的数目）个相互独立的参数优化问题，使结果与局部模型的误差最小。

为了在TS模型的全局逼近特性和局部可解释性间取得平衡这里采用多目标优化方法解决这一问题：它通过求解下面的联合准则求解模型参数：

$$J_C = \alpha J_G + (1 - \alpha) J_L \quad (4-31)$$

式中， α 为正常数，可根据对模型的要求相应地调整 α 的取值。

4.4 仿真实例：PH值中和过程辨识及结果分析

4.4.1 PH值中和过程描述

连续搅拌反应釜(CSTR)有三个输入溶液和一个输出溶液，保持输入溶液的

浓度不变, 输入输出溶液的流量如下:

- 输入酸溶液 Q_1 (HNO_3)
- 输入缓冲溶液 Q_2 (NaHCO_3)
- 输入碱基溶液 Q_3 (NaOH 和少量 NaHCO_3)
- 输出溶液 Q_4

这个问题的难点在于pH值的变化有很强的非线性, 根据经验 $Q_3(k)$ 与pH间的关系即滴定曲线是导致系统非线性的主要原因, 所以这里只考虑 $Q_3(k)$ 对pH值的影响。即保持 $Q_1(k), Q_2(k)$ 恒定, 流量 $Q_3(k)$ 可调作为控制量。控制的目标是使输出溶液的pH值达到设定值7。

4.4.2 过程辨识

辨识数据和检验数据来源于^[34], 随机变化输入碱基溶液流量 Q_3 , 保持输入缓冲剂和酸液的流量不变。输出值为罐内溶液的PH值。辨识数据包含 $N = 499$ 个样本, 采样时间为15s, 如图4-7。

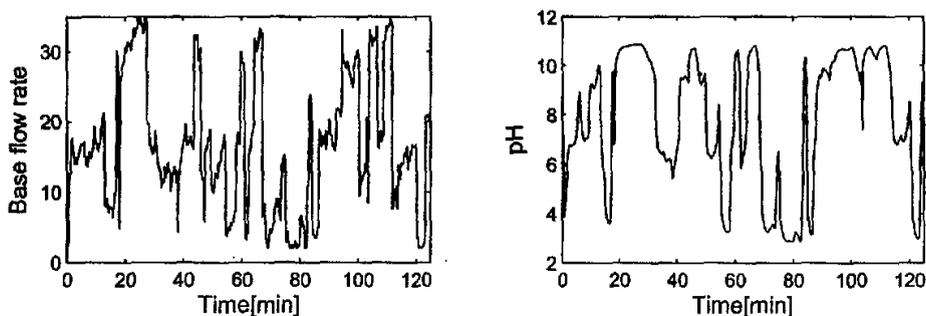


图 4-7: 辨识数据

通过对此中和过程的分析知道, 它可看作一个一阶离散NARX模型:

$$\text{pH}(k+1) = f(\text{pH}(k), Q_3(k)) \quad (4-32)$$

式中, k 表示采样时刻, f 为需辨识的未知关系。利用模糊模型来辨识此未知关系, 模型的输入变量为 $Q_3(k), \text{pH}(k)$, 输出变量为 $\text{pH}(k+1)$ 。采用如下形式规则:

if $Q_3(k)$ is A_i and $\text{pH}(k)$ is B_l then $\text{pH}(k+1) = a_{11}Q_3(k) + a_{12}\text{pH}(k) + b_l$

下面将从以下几个方面来讨论模糊模型的辨识方法:

一、模糊聚类得到数据集的模糊划分, 混合学习算法训练结论参数 根据模糊聚类方法, 设聚类中心为5, 将聚类数据分别投影到两输入变量上, 得到模型

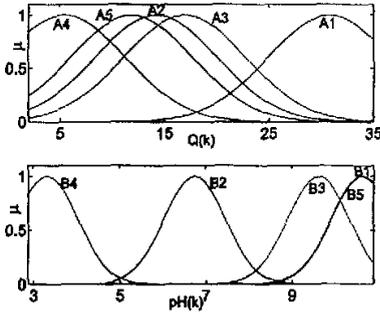


图 4-8: 模糊聚类得到输入变量的隶属函数

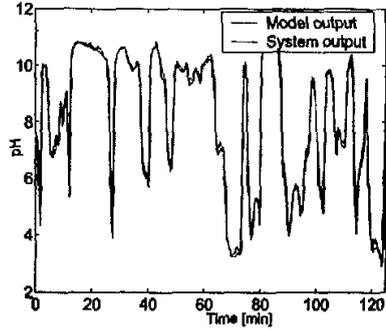


图 4-9: 模型输出与系统输出曲线

的初始模糊集。因为流量 $Q_3(k) \geq 0$, $Q_3(k)$ 表示碱性溶液的流量, 从物理意义上讲 $Q_3(k)$ 的存在将使 $pH(k)$ 增大, 即 $a_{11} \geq 0$ 。在参数训练过程中加入此约束条件。然后使用误差反向传播算法训练前提参数, 最小二乘法算法训练结论参数。最后得到得到隶属函数如图4-8, 模糊规则库的结论参数如表4-1: 利用检验数据得到模

表 4-1: 训练得到的结论参数

l	a_{l1}	a_{l2}	b_l
1	0.03939	0.8382	0.4741
2	0.1405	0.8057	-0.566
3	0.1373	0.551	1.833
4	0.1071	0.6611	0.56
5	0.06921	1.794	-9.026

型输出与系统输出如图4-9, $RMSE=0.0146$ 。

二、合并相似模糊集, 简化模糊规则库 从图4-8可看出一些模糊集合比较相似, 根据式(3-28),(3-31), 并取 $\theta_m = 0.7$ 。通过手动计算即可将图4-8中的 A_2, A_3, A_5 合并, B_1, B_3, B_5 合并, 得到新的模糊集。更新模糊规则库得到3条模糊规则, 训练模型参数, 得到简化了的模糊模型, 经仿真, 此模型对检验数据的 $RMSE=0.0167$ 。

三、ARX模型 将此过程用ARX模型来近似。ARX模型中, 输入序列 $u(t)$ 和输出序列是用线性差分方程:

$$y(t) + a_1y(t-1) + \dots + a_nay(t-na) = b_1u(t-nk) + \dots + bnbu(t-nk-nb+1) \quad (4-33)$$

取 $na = 4, nb = 4, nk = 1$ 通过参数辨识得到模型为:

$$\begin{aligned} y(t) - 1.092y(t-1) - 0.1021y(t-2) + 0.2933y(t-3) - 0.04403y(t-4) \\ = 0.05931u(t-1) - 0.0004227u(t-2) - 0.04597u(t-3) + 0.01207u(t-4) \end{aligned} \quad (4-34)$$

模型对检验数据的一步向前预测曲线如图4-10, RMSE=0.0169。

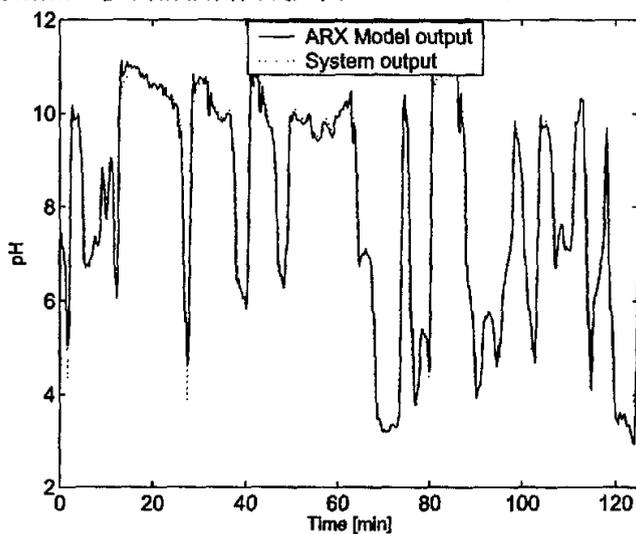


图 4-10: ARX模型输出曲线

4.4.3 结果分析

方法一中, 首先将数据聚类, 得到适当数量的模糊集, 然后利用参数辨识方法训练模型的参数, 并通过对此过程的专家经验, 在训练过程中加入了适当的约束条件, 以保证模型在物理意义上的合理性。仿真结果表明模型能很好地逼近检验数据曲线, 此模型可用于实际过程的预测。方法二中, 按照一定的标准将相似的模糊集和合并, 将模糊集的数目减少为3个。从仿真结果来看, 模型精度影响不大。从模糊规则中更易看出输入输出变量的关系, 提高了模型的可解释性。方法三中, 建立系统的数据对看作时间序列, 辨识得到ARX模型, 用此模型来预测系统的输出值。这种方法的结果也具有较好的精度, 但是它不具有模糊模型辨识方法得到的结果具有可解释性的特点。从输入输出的线性差分方程中很难看出 Q_3 与pH之间的关系。但是根据此过程的模糊规则库, 或由图4-8和表table:ph-para可以看出 Q_3 与pH之间是成正向关系的, 及在各局部区域他们之间的大致对应关系。

第五章 模糊系统辨识及控制仿真实验

在本章中将结合上述方法, 对一个双容单容水箱系统辨识和控制, 以验证本文的模糊建模方法的有效性。这个实验中, 根据系统的输入输出数据和系统的一些先验知识用模糊辨识方法建立水箱系统的模型, 并将得到的模型用于预测控制。

5.1 双容水箱系统结构

双容水箱系统如图5-1所示, 本装置由被控对象和控制两部分组成。

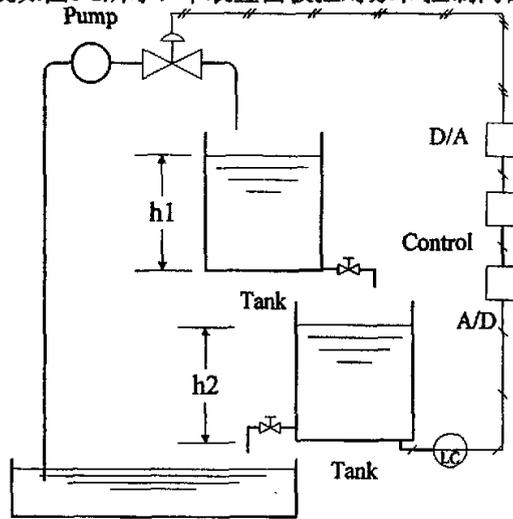


图 5-1: 双容水箱系统

下水箱里的液位 h 作为被控量, 控制算法在计算机里实现的, 过程与计算机间通过数据采集卡连接。辨识和控制程序在MATLAB环境中编写的, 程序中利用Real-Time工具箱和硬件通信, 采样时间为2.5s。

系统动力支路由三相(380V交流)磁力驱动泵、电动调节阀、直流电磁阀、涡轮流量计及手动调节阀组成。详细资料如表5-1。

5.2 模糊模型辨识

设上水箱和下水箱的液位分别为 h_1 和 h_2 , 控制量进水阀的开启角度为 φ 。 q_0 为水箱中水流出的流量, 它由水箱的液位 h 和出水阀的开度 α_0 (不变的)决定。主要

表 5-1: 系统硬件组成

	设备	详细资料	型号
被控对象	水箱	数目3个, 容积均为450cm ³ , 最大液位高度20cm	--
执行机构	磁力驱动泵	380V恒压驱动, 流量32L/min, 扬程8m, 功率180W	16CQ-8P
	电动调节阀	电源220VAC 50HZ, 控制信号为4-20mA DC或1-5V DC, 输出为4-20mA DC, 行程10mm	QSVP-16K
检测装置	压力传感器、变送器	测量范围0.5-100m, 精度0.5级, 供电电压12-36VDC, , 输出信号4-20mA	---
	流量传感器、变送器	供电电源: 电压: 12V±10%, 电流: ≤10mA, 流量范围0.2-1.2m ³ /h, 精确度0.5, 输出信号4-20mA	LWGY-10

参数及取值如表5-2:

表 5-2: 参数和初始值

参数	代表的物理量	初始值
Q_{in}	泵的最大供水量	32L/min
φ	阀的开度	
q_{1in}, q_{2in}	上水箱、下水箱的进水量	
α_1, α_2	出水阀的开度	
h_1, h_2	上水箱、下水箱液位高度	

流入水箱中水的流量 q_i 由进水阀的角度 φ 决定, 同时阀的开度受液位的高度及最大进水流量 Q_{in} 的限制。液位的高度由输入/输出流量和水箱的截面积 A 决定。由水箱的质量平衡有, 水箱中液体的变化量等于输入输出流量之差。这个过程可以用下面的微分方程表示:

$$A_1 \frac{dh_1}{dt} = q_{1in} - q_{1out} \quad (5-1a)$$

$$A_2 \frac{dh_2}{dt} = q_{2in} - q_{2out} \quad (5-1b)$$

上水箱的输入流量 q_{1in} 为:

$$q_{1in} = Q_{in} \sin(\varphi(t)), \quad \varphi \in [0, \pi/2] \quad (5-2)$$

下水箱的输入流量等于上水箱的输出流量, 即:

$$q_{2in} = q_{1out} \quad (5-3)$$

根据Torricelli定律有两水箱的输出流量为:

$$q_{1out} = \alpha_1 \sqrt{2gh_1}, \quad g = 9.81ms^{-2} \quad (5-4a)$$

$$q_{2out} = \alpha_2 \sqrt{2gh_2}, \quad g = 9.81ms^{-2} \quad (5-4b)$$

因此, 得到系统的非线性方程为:

$$\frac{A_1}{\alpha_1} \frac{dh_1}{dt} = -\sqrt{2gh_1} + \frac{Q_{in} \sin(\varphi(t))}{\alpha_1} \quad (5-5a)$$

$$\frac{A_2}{\alpha_2} \frac{dh_2}{dt} = \sqrt{2gh_1} - \sqrt{2gh_2} \quad (5-5b)$$

由以上的分析可以知道, 这个水箱系统是一个二阶滞后环节, 选择模糊模型的前提变量为 $h_2(k)$ 因此设系统的TS模糊模型结构为:

$$R^i: \text{ if } h_2(k) \text{ is } A_1^i \text{ then } h_2(k+1)^i = a_1^i h_2(k) + a_2^i h_2(k-1) + a_3^i u(k-1) + b^i, \quad i = 1, \dots, 6 \quad (5-6)$$

前提变量被划分6个模糊子集。根据液位变化的范围和手动优化之后选择模糊集合的核为{1.8, 4.5, 9, 11.3, 16, 18}。

首先产生辨识和检验数据, 辨识数据集包含 $N = 1400$ 个样本, 输入信号如图5-2¹。这些信号包含了动态过程的一些重要工作频率, 在后面将说明这些信号的频率特性并不是随意选取的, 而是符合实验的实际情况的。

文献^[36]指出在辨识模型中加入合适的先验知识约束条件可以改善模糊控制器的性能, 在这里加入约束条件: 增加时间常数约束, K 的取值在[0, 2.5]间; 每个局部模型的滞后时间约束, 如表5-3。

水箱系统辨识结构如图5-3, 使用第4章的学习算法训练模型参数得到水箱系统的模糊模型, 模型的仿真输出与系统的实际输出如图5-4, 可以看出两条曲线, 几乎完全重合。

¹在这里, 过程的输入输出信号显示为相对于最大值的百分比

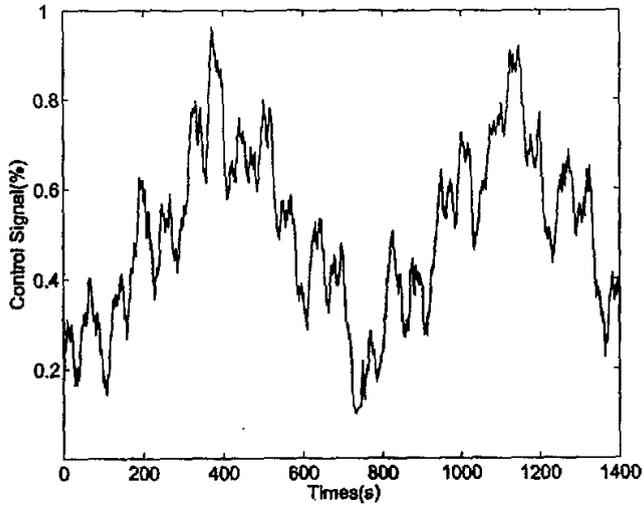


图 5-2: 控制序列的训练数据集

表 5-3: 局部模型的滞后时间约束

	1	2	3	4	5	6
τ_{min}	15	40	60	70	80	80
τ_{max}	30	60	100	120	140	140

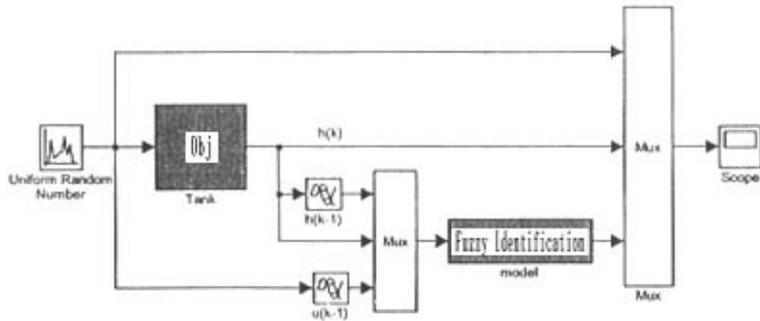


图 5-3: 辨识结构

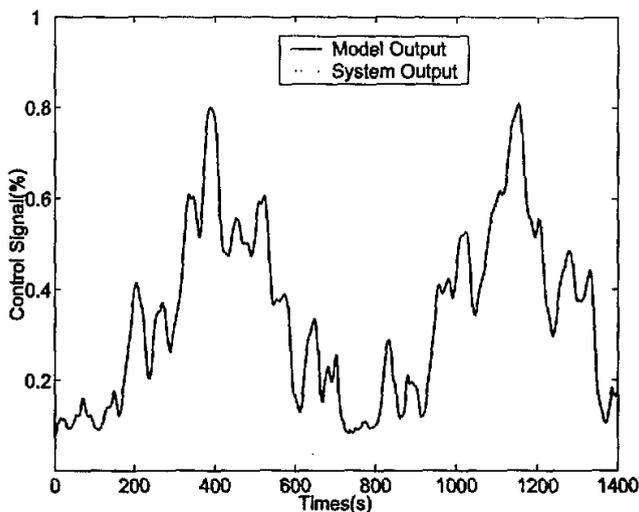


图 5-4: 模型仿真输出与系统实际输出

5.3 基于模糊模型的预测控制

将辨识得到的模糊模型用于对水箱系统的广义预测控制 (GPC)。GPC算法是使以下目标函数值最小^[36]:

$$J = \sum_{i=H_{p1}}^{H_{p2}} (w(k+j) - \hat{y}(k+j))^2 + \lambda \sum_{j=1}^{H_c} \Delta u^2(k+j-1) \quad (5-7)$$

式中, $\hat{y}(k+j)$ 表示预测输出, H_{p2} 表示预测域, H_c 表示控制域, λ 为权系数。使用下面的近似阶跃响应计算预测输出:

$$\hat{y}(k+j) = \sum_{i=1}^j g_i \Delta u(k+i-1) + p_j \quad (5-8)$$

式中, g_i 表示阶跃响应系数, 由模糊模型在当前时刻的响应求得, p_j 表示模糊模型在 $k+j$ 时刻的响应值, 并假设未来时刻的控制信号为常数。把预测输出记成向量的形式 $\hat{\mathbf{y}} = [\hat{y}(k+H_{p1}), \dots, \hat{y}(k+H_{p2})]$, 则GPC算法可写为^[37]:

$$\hat{\mathbf{y}} = \mathbf{G}\Delta\mathbf{u} + \mathbf{p} \quad (5-9)$$

式中, $\Delta\mathbf{u} = [\Delta u(k), \dots, \Delta u(k+H_c)]$, $\mathbf{p} = [p_1, \dots, p_{H_{p2}}]$, \mathbf{p} 的元素值由模糊模型计

算得出:

$$\begin{aligned} p_1 &= f(\hat{y}(k), \hat{y}(k-1), u(k-2)) \\ p_2 &= f(p_1, \hat{y}(k), u(k-1)) \\ p_j &= f(p_{j-1}, p_{j-2}, u(k-1)), j > 2 \end{aligned} \quad (5-10)$$

G 是 $(H_{p2} - H_{p1} + 1) \times H_c$ 阶矩阵, 其中元素 $g_{ij}(j - i > H_{p1})$ 为0:

$$G = \begin{bmatrix} g_{H_{p1}} & g_{H_{p1}-1} & \dots & 0 \\ g_{H_{p1}-1} & g_{H_{p1}} & g_{H_{p1}-1} & \\ \vdots & \vdots & & \vdots \\ g_{H_{p2}} & g_{H_{p2}-1} & \dots & g_{H_{p2}-H_c} \end{bmatrix} \quad (5-11)$$

这个矩阵的元素 G_j 是根据模糊模型在当前操作点计算得出的阶跃响应系数得出的:

$$g_j = \begin{cases} 0, & j \leq n_d \\ -\sum_{i=1}^j a_i g_{j-i} + \sum_{i=1}^j b_i, & j > n_d \end{cases} \quad (5-12)$$

考虑到约束条件时, 式(5-7)的最小值求解可以写成:

$$\min\{(G\Delta u + p - w)^T(G\Delta u + p - w) + \lambda u^T \lambda u\} = \min\{\frac{1}{2} \lambda u^T H \lambda u + d u\} \quad (5-13)$$

其中, $H = 2(G^T G + \lambda I)$, $d = -2G^T(w - p)$, I 是 $H_c \times H_c$ 单位矩阵。

$u, \Delta u$ 的约束条件可以写成不等式的形式:

$$\begin{pmatrix} I_{\Delta u} \\ -I_{\Delta u} \\ I_{H_c} \\ -I_{H_c} \end{pmatrix} \Delta u \leq \begin{pmatrix} u_{max} - I_u u(k-1) \\ -u_{min} + I_u u(k-1) \\ \Delta u_{max} \\ -\Delta u_{min} \end{pmatrix} \quad (5-14)$$

式中, I_{H_c}, I_u 是 $H_c \times H_c$ 阶单位阵, $I_{\Delta u}$ 是 $H_c \times H_c$ 阶下三角矩阵, 其中非零元素为1, $\Delta u_{min}, \Delta u_{max}, u_{min}, u_{max}$ 是 H_c 维向量, 分别表示 $\Delta u_{min}, \Delta u_{max}, u_{min}, u_{max}$ 。

5.4 控制结果

为比较不同模糊模型的控制效果定义下面的性能指标:

$$SS = \sum_{k=1}^T \{(w(k) - y(k))^2 + \lambda \Delta u^2(k)\} \quad (5-15)$$

T 为实验过程的

选择预测域和控制域为 $H_{p1} = 3, H_{p2} = 10, H_c = 2$, 滑动抑制系数 $\lambda = 0.15$ 。约束条件为 $\Delta u_{min} = -\Delta u_{max} = 0.25, u_{max} = 1, u_{min} = 0$

为抑制模型误差和不可测量干扰, 将GPC与内模控制 (IMC) [38]结合起来, 如图5-5

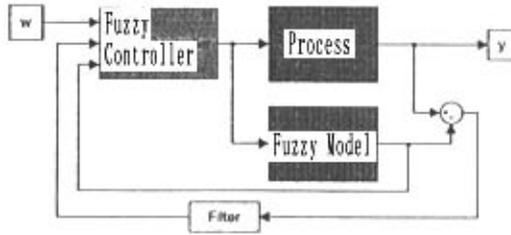


图 5-5: GPC与内模控制的结合控制结构

控制器的输入为参考信号 w , 预测液位 \hat{y} , 和模型误差滤波信号 e_{mf} 。使用如下的一阶Butterworth低通滤波:

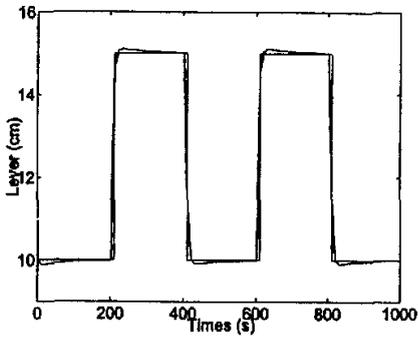
$$e_{mf}(k) = b_{f1}e_m(k) + b_{f2}e_m(k-1) - a_{f2}e_m(k-1) \quad (5-16)$$

式中, $b_{f1} = 0.2929, b_{f2} = 0.2929, a_{f2} = -0.4142$ 。滤波器参数是由期望的截止频率得到的, 以滤去可检测误差并得到快速响应。

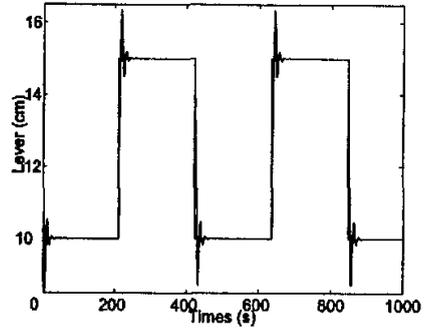
上述方法的预测控制效果如图5.6(a), 对应的 $SS = 0.73$, 使用PID控制得到的结果如图5.6(b)。可以看出使用本文的方法得到的响应曲线效果更好, 使得超调量减小。另外在工程实际中很难确定PID参数, 大多靠多次试调的方法得到较好的参数组合, 是否能够得到好的控制效果只得依靠工程人员的实际经验。而本文的方法不存在这个问题。若不使用先验知识约束条件得到的控制结果的 $SS = 1.2$, 说明在辨识过程中使用合适的约束条件, 可使得模糊模型的局部ARX模型更符合实际, 因此控制效果更好。

5.4.1 结果分析

利用模糊辨识方法建立双容水箱的系统模型, 并在参数辨识过程中加入了适当的专家经验, 实验结果表明得到的模型具有精确性高, 具有一定的可解释性的特点。预测控制结果优于一般控制方法, 且不存在PID控制方法难以确定参数的特点。



(a) 模糊模型的GPC仿真控制



(b) PID仿真控制

图 5-6: 控制效果比较

第六章 结论与展望

结论 本文主要讨论了如何从系统的输入输出数据中获取系统的模型,如何保证模型的精确性和可解释性两个方面的主要问题。

文中讨论了使用模糊C均值聚类的方法构建模糊规则库,这种方法有效地减少了规则数目,为模型参数辨识提供了更精简的模型,减少了运算量,同时可提高模型的可解释性。

文中还讨论了提出了一种简化模型,提高模型可解释性的方法。这种方法是基于合并相似的前提模糊集,结果证明它能简化模型规则库,提高模型的可解释性。可以手动选择调节的程度,在模型的可解释性和模型精度间取得较好的平衡。

在模型参数辨识方面,使用模糊网络结构的方式来辨识模型参数。将网络模型的参数分为前提参数和结论参数。可以使用误差反向传播算法来优化这两类参数,也可将模型的前提参数固定,模型输出看作结论参数的线性函数,使用最小二乘法或其他方法估计模型的结论参数,然后再利用误差方向传播法辨识模型的前提参数。仿真结果证明这种方法比单纯使用误差反向传播方法更有效。

最后将模糊辨识方法应用于水箱系统建模,结合预测控制、内模控制对双容水箱进行控制,得到较好的控制效果。

展望 模糊系统经历了近40年的发展,已从一种简单的控制方法发展为一种新的数值计算方法。本文讨论的模糊模型辨识问题是模糊系统一个较成功的应用方面,这里讨论了如何构建模糊模型的一般框架,但也有不少需要完善的地方:

- 1、将模糊系统辨识方法应用到实际系统控制中。本文讨论的方法大多是离线状态进行的,如何开发更精简的算法,以推进模糊系统辨识实现过程控制的在线辨识,将有很大的研究前景。

- 2、深入分析模糊系统的结构特性及逼近精度,建立一套完整的理论体系,使人们在应用模糊系统时做到心中有数。我们对其的了解就越清楚,也就越能知道它适用于干什么及不适用于干什么。对模糊系统来说这方面研究的具体内容包括,不同的结构参数是怎样具体地影响逼近精度;逼近误差随着参数个数的增多是以什么样的速率在减小;局部性和全局性能否得到平衡;什么样的非线性结构特别适于表示等等。

人们力图把握自然,人类的需求总是超前于目前所掌握的技术,更好的发展模糊系统辨识方法,从数据中提取易于人们理解的模糊模型,任重而道远。

参考文献

- [1] L. X.Wang. A Course in Fuzzy Systems and Control[M]. Englewood Cliffs,N.J.:Prentice Hall, 1997
- [2] Alexandre Evsukoff, Antonio C. S. Branco, Sylvie Galichet. Structure identification and parameter optimization for non-linear fuzzy modeling[J]. Fuzzy Sets Syst. 2002, 132(2):173-188
- [3] T.and Sugeno Tomohiro. Fuzzy Identification of Systems and Its Applications to Modeling and Control[J]. IEEE Transon SMC. 1985, 15:116 132
- [4] Takeshi Horikawa, S.I. Fuzzy Modeling Using Fuzzy Networks with the Back-Propagation Algorithm[J]. IEEE Transon NN. 1992, 3:801 806
- [5] N. Sanornoi, P. Sooraksa. Artificial intelligence based on fuzzy behavior for game programming[C]. In Proceedings of the 2004 ACM SIGCHI international Conference on Advances in Computer Entertainment Technology. Singapore: ACM Press, New York, 2005, vol. 74, 277-279
- [6] Adam Hess, Jason Holt, Jared Jacobson, Kent E. Seamons. Content-triggered trust negotiation[J]. Aug. 2004, 7(3):428-456
- [7] 王立新. 自适应模糊系统与控制——设计与稳定性分析[M]. 北京: 国防工业出版社, 1995
- [8] 张平安李人厚. 关于模糊辨识的理论与应用实际问题[J]. 控制理论与应用. 1995, 12(2):129-137
- [9] Rajshekhar Sunderraman, Haibin Wang. Paraconsistent Intuitionistic Fuzzy Relational Data Model[Z], Oct,2004
- [10] Klir GJ. Higashi M. Identification of Fuzzy Relation System[J]. IEEE Trans on SMC. 1984, 14:349 - 355
- [11] Shyue-Liang Wang, Jenn-Shing Tsai, Tzung-Pei Hong. Mining functional dependencies from fuzzy relational databases[C]. SAC '00: Proceedings of the 2000 ACM symposium on Applied computing. New York, NY, USA: ACM Press, 2000, 490-493
- [12] Meng Joo Er, Yang Gao. Online adaptive fuzzy neural identification and control of non-linear dynamic systems[J]. 2003:373-402
- [13] Sung-Kwun Oh, Witold Pedrycz, Ho-Sung Park. Hybrid identification in fuzzy-neural networks[J]. Fuzzy Sets Syst. 2003, 138(2):399-426
- [14] Teo Lian Seng, Marzuki Khalid, Rubiyah Yusof, Sigeru Omatu. Adaptive Neuro-fuzzy Control System by RBF and GRNN Neural Networks[J]. J Intell Robotics Syst. 1998, 23(2-4):267-289
- [15] Di Wang, Chai Quek, Geok See Ng. Novel Self-Organizing Takagi Sugeno Kang Fuzzy Neural Networks Based on ART-like Clustering[J]. Neural Process Lett. 2004, 20(1):39-51

- [16] T. Takagi, M. Sugeno. Fuzzy identification of systems and its application to modeling and control[J]. *IEEE Trans Systems, Man and Cybernetics*. 1985, 15(1):116-132
- [17] 王立新. 模糊系统与模糊控制教程[M]. 北京: 清华大学出版社, 2003
- [18] B. Moshiri, S. Chaychi. Maleki. Identification of a Nonlinear Industrial Process Via Fuzzy Clustering[C]. *IEA/AIE '98: Proceedings of the 11th international conference on Industrial and engineering applications of artificial intelligence and expert systems*. London, UK: Springer-Verlag, 1998, 768-775
- [19] R. Babuska, M. Oosterom. Fuzzy clustering for multiple-model approaches in system identification and control[J]. 2001:306-323
- [20] Mohamed Marzouk, Osama Moselhi. On the use of fuzzy clustering in construction simulation[C]. *WSC '01: Proceedings of the 33rd conference on Winter simulation*. Washington, DC, USA: IEEE Computer Society, 2001, 1547-1555
- [21] George Tsekouras, Haralambos Sarimveis, George Bafas. A method for fuzzy system identification based on clustering analysis[J]. *Syst Anal Model Simul*. 2002, 42(6):797-823
- [22] Magne Setnes, Hans Roubos. GA-Fuzzy Modeling and Classification: Complexity and Performance[J]. *IEEE Trans on Fuzzy Systems*. 2000, (8):509-522
- [23] Verdegay J L Herrera F, Lozano M. Tackling real-coded genetic algorithms: Operators and tools for behavioral analysis[J]. *Artificial Intell Rev*. 1998, 12:265-319
- [24] Jang J-SR. Neuro-fuzzy modeling and control[J]. *Proceedings of IEEE*. 1995, 83(3):378-406
- [25] James J. Buckley William Siler. *Fuzzy Expert Systems and Fuzzy Reasoning*[M]. Wiley, 2004
- [26] Timothy J. Ross. *Fuzzy Logic with Engineering Applications, 2nd Edition*[M], 2nd edition . Wiley, 2004
- [27] Kessel W. Gustafson, D.E. Fuzzy clustering with a fuzzy covariance matrix[J]. *Proc IEEE-CDC*. 1979, 2:761-766
- [28] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithm*[J]. Plenum Press. 1981
- [29] Kruse Nauck. Neuro-fuzzy systems for function approximation[J]. *Fuzzy Sets and Systems*. 1999, 101:261-271
- [30] Uzay Kaymak et al Magne Setnes, Robert Babuska. Similarity measures in fuzzy rule base simplification[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*. 1998, 28(3):376-386
- [31] The Mathworks. *Fuzzy logic toolbox for use with MATLAB*[C]. The Mathworks, Inc., Natick, MA, USA, June 2001
- [32] 符曦. *系统最优化及控制*[M]. 北京: 机械工业出版社, 2000.7
- [33] Steven M.Kay著 罗鹏飞张文明等译. *统计信号处理基础—估计与检测理论*[M]. 电子工

业出版社, 2003

- [34] T. Johansen. Operating regime based process modeling and identification[Z]. Technical Report 94-109 W. Dr. Ing. thesis, 1994
- [35] T.A. Johansen. Identification of non-linear systems using empirical data and a priori knowledge - an optimisation approach[J]. Automatica. 1996, (32):337 - 356
- [36] P.S.C.Tuffs D.W.Clarke, C.Mothadi. Generalized predictive control-part I. the basic algorithm.[J]. Automatica. 1989, (23):137-148
- [37] Heidar A. Malki Rayanallur S., Ranganathan. Fuzzy predictive PI control for processes with large time delays[J]. Expert Systems. 2002, 19(1):21-26
- [38] C.E. Garcia, M. Morari. Internal model control:1.A unifying review and some new results[J]. Ind Eng Chem Process ResDev. 1982, (21):308 - 323

发表文章目录

- [1] 黄之初, 杨建林, 蒋冬青, 二级可变论域模糊控制器的设计与仿真试验, 武汉理工大学学报, 2005,27(8),85-90
- [2] 杨建林, 黄之初, 一种可变论域模糊控制器的设计及仿真, 计算机仿真, 已录用

致 谢

我非常感谢我在武汉理工大学读研期间所有帮助过我的人。没有他们的帮助，这篇论文不可能完成。在这里不可能提到每一个人的名字，但是我要向每位表达我诚恳的谢意。

首先，我非常感谢我的导师黄之初教授给予我无私的帮助，不断的激励和周全的指导。他渊博的知识，广泛的研究兴趣和充沛的精力给我很大的影响。他总是鼓励我独立思考，是他的信任激发了我的自信。“夫子循循然善诱人，博我以文，约我以礼，欲罢不能。”

我感谢我的父母和朋友对我的默默地奉献和关怀，我的每一次进步都离不开他们的支持。感谢实验室的师兄弟生活和学习上给予的关怀和帮助。

在这即将离开母校之际，感谢所有帮助过我的老师和同学。

作者: [杨建林](#)
学位授予单位: [武汉理工大学](#)

参考文献(38条)

1. [L X Wang](#) [A Course in Fuzzy Systems and Control](#) 1997
2. [Alexandre Evsukoff](#), [Antonio C. S Branco](#), [Sylvie Galichet](#) [Structure identification and parameter optimization for non-linear fuzzy modeling](#)[外文期刊] 2002(02)
3. [Tand Sugeno Tomohiro](#) [Fuzzy Identification of Systems and Its Applications to Modeling and Control](#) 1985
4. [Takeshi Horikawa](#), [S I](#) [Fuzzy Modeling Using Fuzzy Networks with the Back-Propagation Algorithm](#) 1992
5. [N Sanornoi](#), [P Sooraksa](#) [Artificial intelligence based on fuzzy behavior for game program ming](#) 2005
6. [Adam Hess](#), [Jason Holt](#), [Jared Jacobson](#), [Kent E. Seamons](#) [Content-triggered trust nego tiation](#) 2004(03)
7. [王立新](#) [自适应模糊系统与控制—设计与稳定性分析](#) 1995
8. [张平安](#), [李人厚](#) [关于模糊辨识的理论与应用实际问题](#) 1995(02)
9. [Rajshekhar Sunderraman](#), [Haibin Wang](#) [Paraconsistent Intuitionistic Fuzzy Relational Data Model](#) 2004
10. [Klir GJ](#), [Higashi M](#) [Identification of Fuzzy Relation System](#) 1984
11. [Shyue-Liang Wang](#), [Jenn-Shing Tsai](#), [Tzung-Pei Hong](#) [Mining functional dependencies from fuzzy relational databases](#) 2000
12. [Meng Joo Er](#), [Yang Gao](#) [Online adaptive fuzzy neural identification and control of non- linear dynamic systems](#) 2003
13. [Sung-Kwon Oh](#), [Witold Pedrycz](#), [Ho-Sung Park](#) [Hybrid identification in fuzzy-neural networks](#)[外文期刊] 2003(02)
14. [Teo Lian Seng](#), [Marzuki Khalid](#), [Rubiyah Yusof](#), [Sigeru Omatu](#) [Adaptive Neuro-fuzzy Control System by RBF and GRNN Neural Networks](#) 1998(2-4)
15. [Di Wang](#), [Chai Quek](#), [Geok See Ng](#) [Novel Self-Organizing Takagi Sugeno Kang Fuzzy Neural Networks Based on ART-like Clustering](#)[外文期刊] 2004(01)
16. [T Takagi](#), [M Sugeno](#) [Fuzzy identification of systems and its application to modeling and control](#) 1985(01)
17. [王立新](#), [王迎新](#) [模糊系统与模糊控制教程](#) 2003
18. [B Moshiri](#), [S Chaychi Maleki](#) [Identification of a Nonlinear Industrial Process Via Fuzzy Clustering](#) 1998
19. [R Babuaka](#), [M Oosterom](#) [Fuzzy clustering for multiple-model approaches in system identification and control](#) 2001
20. [Mohamed Marzouk](#), [Osama Moselhi](#) [On the use of fuzzy clustering in construction simula tion](#) 2001
21. [George Tsekouras](#), [Haralambos Sarimveis](#), [George Bafas](#) [A method for fuzzy system identification based on clustering analysis](#)[外文期刊] 2002(06)
22. [Magne Setnes](#), [Hans Roubos](#) [GA-Fuzzy Modeling and Classification:Complexity and Performance](#) 2000(08)
23. [Verdegay J L](#), [Herrera F](#), [Lozano M](#) [Tackling real-coded genetic algorithms:Operators and tools for](#)

[behavioral analysis](#) 1998

24. [Jang J-SR Neuro-fuzzy modeling and control](#)[外文期刊] 1995(03)
25. [James J. Buckley William Siler Fuzzy Expert Systems and Fuzzy Reasoning](#) 2004
26. [Timothy J Ross Fuzzy Logic with Engineering Applications](#) 2004
27. [Kessel W. Gustafson D E Fuzzy clustering with a fuzzy covariance matrix](#) 1979
28. [J C Bezdek Pattern Recognition with Fuzzy Objective Function Algorithm](#) 1981
29. [Kruse Nauck Neuro-fuzzy systems for function approximation](#)[外文期刊] 1999
30. [Uzay Kaymak, Magne Setnes, Robert Babuska Similarity measures in fuzzy rule base simplification](#) 1998(03)
31. [The Mathworks Fuzzy logic toolbox for use with MATLAB](#) 2001
32. [符曦 系统最优化及控制](#) 2000
33. [Steven M Kay, 罗鹏飞, 张文明 统计信号处理基础-估计与检测理论](#) 2003
34. [T Johansen Operating regime based process modeling and identification](#) 1994
35. [T A Johansen Identification of non-linear systems using empirical data and a priori knowledge - an optimisation approach](#) 1996(32)
36. [P S C Tuffs, D W Clarke, C Mothadi Generalized predictive control-part I. the basic algorithm](#) 1989(23)
37. [Heidar A Malki, Rayanallur S Ranganathan Fuzzy predictive PI control for processes with large time delays](#)[外文期刊] 2002(01)
38. [C E Garcia, M Morari Internal model control:1.A unifying review and some new results](#) 1982(21)

本文读者也读过(10条)

1. [姚泽宏 模糊辨识方法的研究](#)[学位论文]2004
2. [邵青, 冯汝鹏, SHAO Qing, FENG Ru-peng 非线性系统模糊辨识的新方法](#)[期刊论文]-[控制与决策](#)2001, 16(1)
3. [刘福才, 关新平, 裴润 一种基于模糊规则的非线性系统快速模糊辨识方法](#)[期刊论文]-[系统仿真学报](#)2002, 14(5)
4. [程锋章 非线性系统模糊辨识方法的研究](#)[学位论文]2006
5. [刘福才, 关新平, 裴润 基于一种新模糊模型的非线性系统模糊辨识](#)[期刊论文]-[控制理论与应用](#)2003, 20(1)
6. [孙立萍 基于T-S模型的模糊辨识方法及其应用研究](#)[学位论文]2006
7. [李秀英 非线性系统的模糊辨识与控制](#)[学位论文]2005
8. [王宏伟, 马广富, WANG Hongwei, MA Guangfu 基于模糊规则的非线性系统建模方法](#)[期刊论文]-[控制理论与应用](#) 2000, 17(3)
9. [刘福才, 路平立, 裴润, LIU Fu-cai, LU Ping-li, PEI Run 基于正交最小二乘的非线性系统模糊辨识](#)[期刊论文]-[系统工程学报](#)2005, 20(3)
10. [翟彦景 具有BP算法的模糊、神经网络在非线形动态系统辨识中的应用](#)[学位论文]2007

本文链接: http://d.g.wanfangdata.com.cn/Thesis_Y813083.aspx