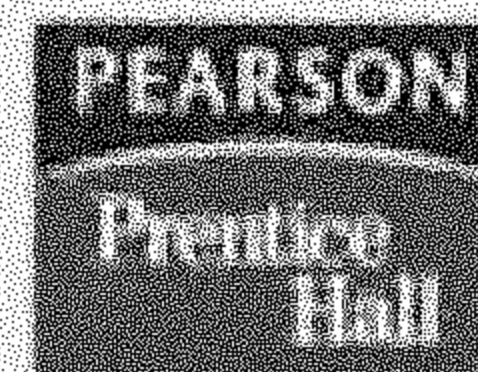
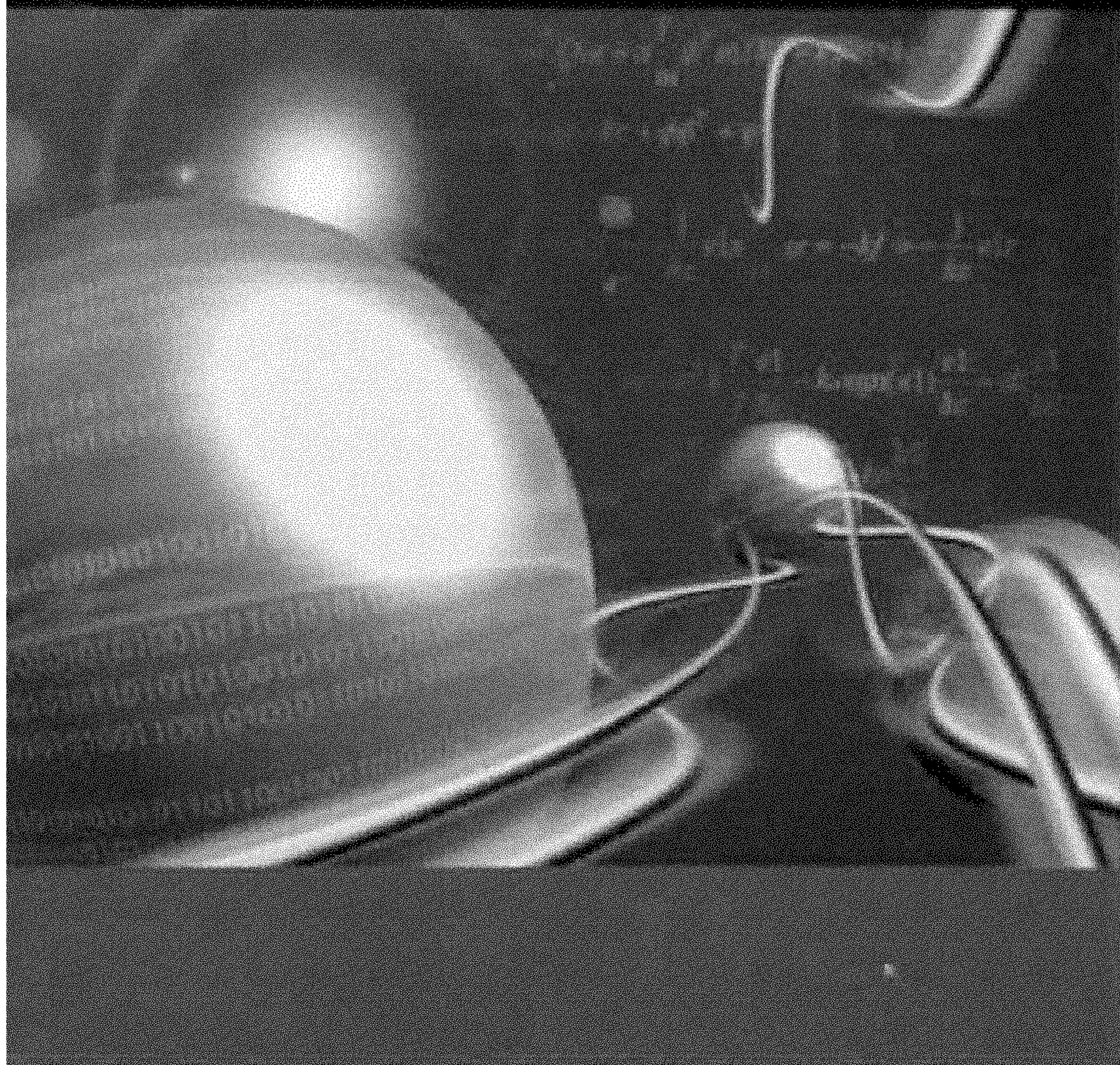


王立新 著 王迎军 译



A COURSE IN FUZZY SYSTEMS & CONTROL

模糊系统与模糊控制 教程



清华大学出版社



清华大学出版社

清华商务出版在线（最新图书资料与服务信息）

<http://www.tupub.com/>

请填写此表，传真或寄回出版社发行部

书名		ISBN	
作者		单价	
订货单位		数量（册）	
地址		邮编	
到站	慢件	订货码洋	
	快件	订货日期	
开户行		经办人	
账号		电话	
税号		传真	

发行部：

地区省份 电 话 分机（总机：010-62770177）

华东区 010-62782535 3522 /3517

华南区 010-62782543 3514 /3512

西北区 010-62776966 3532/3531

西南区 010-62781146 3519 /3518

华中区 010-62781827 3515/3516

东北区 010-62771153 3520/3521

华北区 010-62771377 3525/3513/3523

邮购电话：62786544 分机：3507 联系人：王晓龙

传真：010-62776996（华北） 010-62771137 010-62776466

开户行：北京市建设银行清华园分理处 户名：清华大学出版社发行部

帐号：6510005092610001135 税号：110108H52628661

地址：清华大学东门外科技园区学研大厦 A 座 5 层

文泉经管研发部：

选题策划：徐学军 010-62781829 市场推广：白 杰 010-62794748

编辑助理：刘冰利 010-62794748

编辑部传真：010-62794748

模糊系统与模糊控制教程

王立新 著
王迎军 译

清华大学出版社
北京

内 容 简 介

《模糊系统与模糊控制教程》共 31 章,内容包括模糊数学与模糊逻辑中一些对于模糊系统理论有用的概念和原理,模糊系统的各组成部分,设计模糊系统的四种方法,模糊控制与非自适应模糊控制方法,以及若干不属于本书主体结构但很重要的专题,如模糊 c-均值算法、模糊关系方程、模糊算术、模糊线性规划和可能性理论等。

本书有效地结合了理论的前沿性和方法的实用性,具有极强的指导意义,可以使硕士、博士研究生在选择论文题目或考虑到解决问题的方案以及遇到新问题采取对策或措施时形成更多的思路。从国内有关教材来看,绝大部分书籍以模糊数学为主要内容,缺少模糊系统中的建模方法及优化方法。而该书则为读者提供了模糊系统与模糊控制这一领域的概要及基本原理,是一本真正优秀的模糊系统与模糊控制教材。

该书是一本针对研究生和高年级本科生的教科书,适用于 60 学时以下的模糊系统或模糊控制以及相关专业的教学,如适合控制科学与工程、管理科学与工程、应用数学等专业教学使用。同时,该书也可供工程人员研究参考。

A Course in Fuzzy Systems and Control

Li-Xin Wang

ISBN: 0-13-540882-2

Copyright ©1997 by Prentice-Hall, Inc.

Original English language edition published by Prentice-Hall, Inc

Simplified Chinese edition copyright © 1997 by Pearson Education Limited and Tsinghua University Press.

Published by arrangement with Prentice-Hall, a Pearson Education Company.

This edition is authorized for sale only in People's Republic of China (excluding the Special Administrative Region of Hong Kong and Macau, and Taiwan).

All Rights Reserved.

本书中文简体版由培生教育集团授权清华大学出版社在中国境内(香港、澳门特别行政区和台湾地区除外)独家出版、发行。

版权所有,翻印必究。

本书封面贴有 **Pearson Education**(培生教育出版集团)激光防伪标签,无标签者不得销售。

北京市版权局著作权合同登记号: 01-2002-2998

图书在版编目(CIP)数据

模糊系统与模糊控制教程 = A Course in Fuzzy Systems and Control/王立新著;王迎军译.

—北京:清华大学出版社,2003

ISBN 7-302-06561-6

I. 模… II. ①王…②王… III. ①模糊系统—高等学校—教材—英文
②模糊控制—高等学校—教材—英文 IV. ①N94 ②TP13

中国版本图书馆 CIP 数据核字(2003)第 027705 号

出 版 者: 清华大学出版社(北京清华大学学研大厦,邮编 100084)

<http://www.tup.tsinghua.edu.cn>

<http://www.tup.com.cn>

责任编辑:徐学军

印 刷 者:

发 行 者: 新华书店总店北京发行所

开 本: 787×960 1/16 印张: 23.25 字数: 516 千字

版 次: 2003 年 6 月第 1 版 2003 年 6 月第 1 次印刷

书 号: ISBN 7-302-06561-6/F·523

印 数: 0001 ~ 5000

定 价: 48.00 元

商务出版在线 www.tupub.com

译者序

这是一本关于模糊系统与模糊控制的著作。我最初接触这个领域是在上海宝钢研究院自动化所做博士后期间,当时我全程参加了本书作者王立新教授关于本书的报告和讨论活动,然后我通过独立编程和建立多个模糊系统模型,运用模糊系统建模方法完成了博士后报告中的核心内容,并将所得成果应用于宝钢最新的转炉炼钢模型。这些切身的体会使我从此对该领域产生浓厚兴趣,一直关注该领域的发展,且积极地将模糊系统理论应用于管理科学领域。

模糊系统理论是由 Zadeh 于 1965 年创立的。随着模糊系统理论在工业界的成功应用,越来越多的学者开始研究模糊系统理论。20 世纪 90 年代,这一学科在理论上解决了两个重要难题:如何从数据中获得模糊规则和如何保证模糊系统的稳定性,对此阐述最为详尽、研究最为深入的就是现在我们想向读者推荐的《模糊系统与模糊控制教程》一书。该书是国际上从事模糊系统与模糊控制研究的学者人手一册的指导性专著,是国际流行的最优秀的教材,在国际学术界影响颇深。该书作者是 Automatica 与 IEEE Trans. On Fuzzy Systems 的副主编、香港科技大学的王立新教授。作为 Zadeh 学生的王立新教授在模糊系统与模糊控制领域的工作成果被同行们广为引用。

《模糊系统与模糊控制教程》是一本针对研究生和高年级本科生的教科书(如适用于 60 学时以下的模糊系统或模糊控制以及相关专业的教学),如控制科学与工程、管理科学与工程、应用数学等专业,也适于工程人员使用。本书有效地结合了理论的前沿性和方法的实用性,具有极强的指导意义,可以使硕士、博士研究生在选择论文题目或考虑到解决问题的方案以及遇到新问题采取对策或措施时形成更多的思路。从国内有关教材来看,绝大部分书籍以模糊数学为主要内容,缺少模糊系统中的建模方法及优化方法。而《模糊系统与模糊控制教程》则为读者提供了模糊系统与模糊控制这一领域的概要及基本原理,是一本真正的模糊系统与模糊控制教材。

全书字数约为 40 万字左右,其要点如下:

- ◆ 第一部分(第二章至第六章)介绍模糊数学与模糊逻辑中一些对于模糊系统理论有用的概念和原理,包括模糊集合及模糊集合的基本运算和其他运算、模糊逻辑与近似推理、语言变量与模糊 IF-THEN 规则等。
- ◆ 第二部分(第七章至第十一章)详细研究模糊系统的各组成部分:模糊器、解模糊器、模糊规则库与模糊推理机,推导出不同类型模糊系统的严密的数学公式,并研究这些模糊系统的逼近特性,证明了模糊系统是万能逼近器以及如何设计模糊系统才会使其达到所需精度。

- ◆ 第三部分(第十二章至第十五章)介绍了根据输入输出数据设计模糊系统的四种方法:查表法、梯度下降法、递推最小二乘法和聚类法,并给出大量的具体应用例子来检验这些方法。
- ◆ 第四部分(第十六章至第二十二章)和第五部分(第二十三章至第二十六章)集中研究了模糊控制。其中,第四部分研究了非自适应模糊控制,给出了设计模糊控制器的试错法和线性系统与非线性系统模糊控制器的设计方法,并对模糊控制系统的稳健性进行了分析;第五部分研究了自适应模糊控制,给出了基本自适应模糊控制器和高级自适应模糊控制器的设计方法。
- ◆ 第六部分(第二十七章至第三十一章)综述了若干不属于本书主体结构但很重要的专题,如模糊 c -均值算法、模糊关系方程、模糊算术、模糊线性规划和可能性理论等。

本书的译者包括东北大学工商管理学院的王迎军博士和陈桂云博士、黑龙江科技学院陈培友副教授及东北大学工商管理学院的博士生高峻峻。译者们在繁重的工作和学习之余,前后付出了近一年半的时间完成了这项颇为艰巨的任务。全书共三十一章,各章初步翻译的分工如下:前言、目录、索引和第一章至第十五章由王迎军与高峻峻共同翻译;第十六章至第二十二章由陈桂云翻译;第二十三章至第三十一章由陈培友翻译。王迎军、陈桂云和高峻峻对全书的初稿进行了统一修改、校对,最后由王迎军定稿。此外,徐慧、胡乐江、周麟祥、段琰、谭凌、李秋正、吕芹等研究生对书稿的打印、整理工作提供了很多帮助,在此表示感谢。本书在翻译和出版过程中得到清华大学出版社文泉公司经管研发部的大力支持,在此一并表示感谢。

由于现今中文本的模糊系统理论专著很少,因此书中有些专业术语尚无统一的中文译法。在翻译过程中,虽然几经推敲,但由于时间紧,译者水平有限,仍难免有疏漏和错误,敬请读者不吝指正。对于本书译稿,若有指正或需与译者商讨,请通过电子邮件 yjwang@rocket.mail 或 gao_jun_jun@sohu.com 与译者联系。

王迎军

2003年2月于东北大学

中文版序言

首先,非常感谢王迎军教授以及他所领导的小组对本书的中文翻译所付出的大量时间和心血。同时,也非常感谢清华大学出版社文泉经管研发部出版此书。为了增强广大读者对本书的了解以及增加学习模糊理论的兴趣,下面谈谈我是怎样进入模糊系统这个研究领域的,以及多年来在这个领域从事研究的观察与体会。

一、我是怎样进入这个模糊领域的

我的本科和硕士是在国内完成的,学的都是传统的自动控制。当时只听说过模糊数学,没有接触过模糊系统和模糊控制。1989年到美国后,第一年从事神经网络在信号处理中的应用方面的研究,用 Hopfield 网处理探油地震数据,估计地层反射系数;还用 BP 网估计高阶矩模型的参数。这些研究结果后来在地质学和信号处理的刊物上发表,至今在 SCI 上还被引用。1990 年下半年时,由于模糊控制在日本的广泛应用而成为热门课题,我的导师 Mendel 教授建议我看看这方面的论文,谁知从此走上了不归路。

当时找了很多模糊理论的书和论文来看,最深的体会是越看越模糊。体现在两个方面:一是论文引言中所声称要做的和后来论文中实际所做的互相矛盾,如引言中说传统理论太精确而不利于解决复杂问题,需要引入模糊概念,而后来论文中给出的却都是精确的定义和严格的数学推理;二是不少论文逻辑混乱,谦虚点说是以我的水平看不懂。与传统自动控制领域的论文相比,模糊领域论文的整体质量实在不敢恭维。那么是什么原因让我继续下去呢?这要从我在国内读硕士时说起。

读硕士第一年我的导师戴冠中教授去了美国做访问学者。我因此而没人管,放了羊,在书的草原上漫游。控制论与人机系统方面的一些书和论文很吸引我的兴趣。我当时有一种隐隐的感觉,就是控制论要得以开放性的发展,必须将人的因素作为一个有机的部分考虑进去。控制论的控制对象不应局限于物理和化学系统,而应扩展到社会经济等人与人交往而形成的系统。社会经济系统与物理化学系统的一个本质区别,在于人是社会经济系统的核心组成部分。物理化学变量之间的关系受物理化学定律的约束,自然科学研究的目的是发现这些定律。社会经济系统的复杂性就在于它的运作主体——人,是通过自然语言来交流的,而至今还没有一套可以和自然科学定律相比的数学体系,来描述自然语言交流中的各种现象。当时的理想是建立一套广义控制论,不仅适用于物理系统,而且适用于社会经济系统。出生牛犊,志大才疏,具体怎么做无从下手。

回到 1990 年下半年的美国。虽然大部分模糊论文让我失望,但剥开模糊的浓雾,我看到了模糊理论一个独特且杰出的贡献,那就是模糊理论给出了一套系统而有效的方法,将以

自然语言表达的知识转换成数学形式,从而使这些知识得以有效的应用。模糊系统与通常的专家系统的不同之处在于,模糊系统是把专家知识先转换成数学形式,然后加以应用,而通常的专家系统是把专家知识用计算机语言来表述。即模糊系统把符号(自然语言表达的知识)转换成数学函数,而专家系统把一种符号转换成另一种符号。一旦语言转换成了数学函数,我们就可以借助于丰富的数学工具建立一套理论体系,把人和物理系统统一起来加以描述和研究。难道这就是我设想的广义控制论的突破口?直觉告诉我应该走下去。就这样,一走就是十二年。

二、模糊系统与模糊控制理论的三个突破

有一种说法,称模糊系统与模糊控制是经验学科,缺少严格的理论支持和指导。倒退至1990年,这种说法是完全正确的,也是当时我所面临的现状。我当时觉得造成理论研究困难的核心因素,是人们把模糊控制器的设计分成几个相对独立的部分来进行,如隶属函数的确定,规则的获取,控制器的合成等。这样做的好处是把问题简化,便于初学者上手。但这样做的问题是设计好的系统很难进行理论分析及优化设计。于是,我尝试把模糊控制器作为一个整体来看,从输入输出的角度研究其性能,统一其设计。由于我当时正在进行神经网络的研究,而神经网络中最常用的BP网之所以被广泛应用,其核心原因是有人证明BP网是万能逼近器,即任何非线性函数都可用BP网任意逼近。当时模糊控制器已经在很多不同的领域得到成功应用,于是我想模糊控制器是不是(是否)也是万能逼近器?这就有了下面第一个突破。

不管其内部是如何进行运算的,从输入输出的角度讲模糊系统是一个非线性函数(当模糊系统被当作控制器使用时称作模糊控制器)。这个非线性函数的具体形式取决于多种因素,如隶属函数,推理规则,解模糊方法等。现在的问题是对于一个任意的非线性连续函数,我们是否一定能找出一类隶属函数,一种推理规则,一个解模糊方法,使得设计出的模糊系统能够任意逼近这个非线性函数。如果答案是肯定的,那么我们就从理论上解释了为什么模糊控制器会在众多不同领域取得成功,也为模糊系统在更广泛领域的推广应用提供了强有力的理论支持。要从数学上严格证明这个结果不是件容易的事,更何况当时没人知道答案到底是肯定的还是否定的。当时是我到美国的第三学期,在数学系修一门研究生的高级泛函分析课,要写一篇课程设计,于是我用课上所学的Stone-Weierstrass定理,现买现卖,证明了一类模糊系统是万能逼近器。这个结果最终在IEEE神经网络汇刊上发表^①,开辟了模糊逼近这个领域。10年来,模糊逼近一直是模糊理论的热门课题,到2003年2月止这篇论文在SCI上已被引用过261次。

第二个突破是从数据中通过学习获取模糊规则。当时模糊控制器应用的一个瓶颈是模

^① L. X. Wang and J. M. Mendel, "Fuzzy basis functions, universal approximation, and orthogonal least squares learning," IEEE Trans. on Neural Networks, Vol. 3, No. 5, pp. 807-814, 1992.

糊规则的获取。对于简单系统可以用专家问卷及反复尝试法,但当变量增多这些方法就行不通了。由于没有数学模型可用,信息的最大来源是量测到的数据,因此如何从数据中获取模糊规则就成为模糊系统理论和应用能否得以长足发展的核心问题。当时我正在修 Kosko 的神经网络与模糊系统课程,要做课程设计,于是就发明了现在被别人称做的 Wang-Mendel 方法^①。此方法的特点是简单有效,现已成为该领域的经典方法,新提出的方法一般都要和此方法进行比较。此方法开辟了后来十分热门的模糊神经网络领域。当时我不喜欢模糊神经网络这个提法,因为学习算法是多种多样的,不应局限于神经网络中的方法。到 2003 年 2 月止这一方法在 SCI 上已被引用过 279 次。

第三个突破是如何设计出能够确保稳定性的模糊控制器。当时虽然有一些模糊控制器的稳定性分析,但都局限于线性被控对象或被控对象模型已知,这与模糊控制器的基本理念相违背,因为模糊控制器的被控对象应该是非线性系统且模型未知。由于有了第一个突破中将模糊控制器作为一个整体来看的观点,于是我就提出了一种能确保稳定的自适应模糊控制器的设计方法^②,把模糊控制器的设计从一门艺术向一门科学推进。多年来,能确保稳定的自适应模糊控制理论得到了长足的发展。现在如果还有人说模糊控制是经验学科,缺少严格的理论支持和指导,那么他一定不了解模糊控制的最新发展。我们有一套完全可以和其它控制方法相比的理论体系,设计出保证稳定及其它性能要求的模糊控制器。自适应模糊控制器的设计方法开辟了这个领域,到 2003 年 2 月止这一设计方法在 SCI 上已被引用过 155 次。

俗话说千里马常有,而伯乐不常有。以上这些工作在当时我并没有觉得有什么特殊,顺理做下来而已。谁知一个电话,使我和 Zadeh 结下了不解之缘,也把我拴在了模糊的十字架上。

三、我所相识的 Zadeh 教授

Wang-Mendel 方法首先以研究报告的形式于 1991 年初发表, Mendel 寄给了 Zadeh 一份(当时我并不知道)。一天下午,办公室电话铃响,师兄去接的,说是找我。电话里的对方说他是 Zadeh,刚看完我的研究报告,觉得很好。作为模糊之父, Zadeh 在我心中是高高在上的。作为二年级的博士生,突然接到鼻祖的电话,我受宠若惊,兴奋不已。Zadeh 问我什么时候毕业,我说学分还没修够,论文进展顺利。Zadeh 又问我毕业后有没有兴趣去他那里工作,我说当然愿意。一星期后,我又接到 Zadeh 的电话,他说已经专门为我申请好了经费,由 Rockwell 出资,我可以随时去做博士后。就这样,我于 1992 年初通过博士论文答辩,那时距我到美国只有两年半。

^① L. X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," IEEE Trans. on Systems, Man, and Cybern., Vol. 22, No. 6, pp. 1414-1427, 1992.

^② L. X. Wang, "Stable adaptive fuzzy control of nonlinear systems," IEEE Trans. on Fuzzy Systems, Vol. 1, No. 2, pp. 146-155, 1993.

1992年的夏天是令人兴奋的。退掉了洛杉矶的房子,将所有家当装进我的那辆破车,开上了前往旧金山的高速公路。沙丘在车旁奔跑,伴随着风的欢笑,在阳光里划出一道弧线。要做 Zadeh 的学生了,而据 Mendel 讲 Zadeh 有意培养我做他的接班人。做大师的接班人,发扬光大模糊理论,任重道远,义不容辞。当时兴奋的心情至今记忆犹新。到 Berkeley 后先找地方住下,然后马上去见 Zadeh。敲门,门开了,出现在我面前的是一位犀利的眼光中透着和蔼的老人。

Zadeh 于 1921 年出生在前苏联的阿塞拜疆(Azerbaijan)共和国,母亲是俄罗斯人,父亲是伊朗人,从事进出口生意,家境富裕。Zadeh 在前苏联上小学,接受十月革命后的共产主义教育。后来由于政局动荡,Zadeh 的父母把他送到伊朗的美国学校,并在伊朗 University of Teheran 完成了大学教育。1944 年 Zadeh 到了美国,就读 MIT,1946 年获得硕士学位后转到 Columbia 大学继续深造,于 1949 年获博士学位,并留校任教。Zadeh 是现代控制理论的开创者之一,他首先提出了状态空间的概念,并是系统理论(system theory)这个名词的发明者。1959 年 Zadeh 离开 Columbia 转去 Berkeley,一直工作到现在。

我到 Berkeley 时 Zadeh 已是退休教授,不用上课,其它研究条件不变。当时 Zadeh 有一个博士生(J. R. Jang,著名的模糊神经网络 ANFIS 的发明者)和我一个博士后。在和 Zadeh 相处的一年多时间里,我们经常一起吃午饭,讨论各种问题。其实具体的学术问题我们讨论的并不多,我想主要原因是我所研究的具体方向(模糊系统与模糊控制)与 Zadeh 所研究的具体方向(computing with words)有所不同。在开创模糊理论之后 Zadeh 就转到计算机系,因为他当时觉得模糊理论的主要应用领域应该是计算机科学。后来 Zadeh 自己也说没想到模糊理论的突破发生在他离开了的系统与控制领域。Zadeh 对我的学术指导只有一句话,那就是我第一天见他时问他我应该做什么,他说:“不要变,继续沿你现有的方向做下去。”于是我在 Berkeley 期间主要完成了我的第一本专著^①,其中译版于 1995 年 10 月由国防工业出版社出版。该书当时被 Astrom 称作模糊控制最好的书,到 2003 年 2 月止该书在 SCI 上已被引用过 423 次。

与 1992 年的夏天相比,1993 年的夏天是沉重的。Zadeh 把我作为第一候选人,全力推举我留 Berkeley 任教。Zadeh 退休后,Berkeley 再没有教授专门从事模糊理论的研究(到现在依然如此)。虽然我已经到了最后一轮(final list),但最终没能如愿。Zadeh 建议我再多作一年博士后,第二年再争取。当时的我由于一路非常顺利,年青气盛,竟然拒绝了 Zadeh 的建议,毅然奔向了灿烂的东方明珠。

四、本书是怎样完成的

到香港后想作的第一件事是写一本模糊系统与模糊控制的标准教材。一个领域要得以

^① L. X. Wang, Adaptive Fuzzy Systems and Control: Design and Stability Analysis, Prentice-Hall Inc., Englewood Cliffs, NJ, 1994.

长期稳定的发展,一本标准教材是必不可少的。虽然当时模糊理论的书不少,但没有一本可作为模糊系统与模糊控制的标准教材使用。我自己进入这个领域时就因为没有一本好书而浪费了不少时间。从领域发展的角度讲,当时急需一本标准教材,使希望从事模糊系统与模糊控制研究的学者和学生有路可寻。而一本标准教材也是一个领域走向成熟的表现。万事开头难,既然是 Zadeh 的接班人,完成这第一本教材便成为我义不容辞的责任。经过两年多的教学实践,终于完成了此书,于 1997 年由美国的 Prentice-Hall 出版,这就是本书的英文原著。

本书英文版出版 5 年来,已被众多国家的多所大学采用。不少模糊领域的新秀是读着此书进入本领域的。到 2003 年 2 月止本书英文版已在 SCI 上被引用过 151 次,而且引用的增长速度在加快。由于当时时间紧,现在回过头来看看有很多地方需要改进。一直有出第二版的想法,但由于太懒,一拖再拖,不知何时才能完成。

五、天苍苍,雾茫茫,拨开迷雾见太阳

模糊理论是一个充满争议的领域,主要原因之一是人们对模糊这两个字的误解。模糊理论并非“模糊的理论”,而是研究模糊现象,利用模糊信息的精确理论。与其它实用性理论一样,模糊系统理论所追求的目标是为解决各种实际问题提供更加有效的思路与方法。模糊系统理论有以下特点:

- 强调充分利用各种所能获得的信息,包括数据信息,模型信息,语言信息等。
- 将各种信息融为一体,在一个统一的数学框架下进行研究。
- 强调实用性与理论完备性相结合,以实用性为先导。
- 应用对象不局限于某个特定领域,而广泛适用于工程、经济、金融、管理、社会等各个领域。

天苍苍,广义控制论的梦想不知现在何方?

雾茫茫,模糊理论是不是正确的前进方向?

拨开迷雾,等待我们的是不是那灿烂的阳光?

亲爱的读者,希望你喜欢此书,就像我喜欢模糊理论一样。

王立新

2003 年 3 月于香港科技大学

目 录

前言	XIX
第 1 章 引言	1
1.1 为什么研究模糊系统	1
1.2 什么是模糊系统	2
1.3 模糊系统应用的领域及方式	5
1.3.1 模糊洗衣机	5
1.3.2 数字图像稳定器	6
1.3.3 汽车中的模糊系统	6
1.3.4 水泥窑的模糊控制	7
1.3.5 地铁的模糊控制	7
1.4 模糊理论的主要研究领域	8
1.5 模糊理论及其应用的简史	9
1.5.1 20 世纪 60 年代: 模糊理论的萌芽	9
1.5.2 20 世纪 70 年代: 模糊理论继续发展并出现了实际的应用	10
1.5.3 20 世纪 80 年代: 模糊理论的大规模应用使其产生巨大飞跃	10
1.5.4 20 世纪 90 年代: 模糊理论仍有更多的挑战	11
1.6 本章小结及导读	11
1.7 习题	12

第一部分 模糊系统与模糊控制的模糊数学原理

第 2 章 模糊集合及其基本运算	14
2.1 由经典集合到模糊集合	14
2.2 模糊集合的一些基本概念	18
2.3 模糊集合的运算	20
2.4 本章小结及导读	21
2.5 习题	22
第 3 章 模糊集合的其他运算	23
3.1 模糊补	23
3.2 模糊并—— s -范数	25
3.3 模糊交—— t -范数	29

3.4	平均算子·····	32
3.5	本章小结及导读·····	33
3.6	习题·····	34
第4章	模糊关系与扩展原理·····	35
4.1	由经典关系到模糊关系·····	35
4.1.1	关系·····	35
4.1.2	投影和柱状扩展·····	37
4.2	模糊关系的合成·····	39
4.3	扩展原理·····	42
4.4	本章小结及导读·····	42
4.5	习题·····	43
第5章	语言变量与模糊 IF-THEN 规则·····	44
5.1	由数值变量到语言变量·····	44
5.2	语言限定词·····	45
5.3	模糊 IF-THEN 规则·····	46
5.3.1	模糊命题·····	46
5.3.2	模糊 IF-THEN 规则的解释·····	48
5.4	本章小结及导读·····	53
5.5	习题·····	54
第6章	模糊逻辑与近似推理·····	55
6.1	由经典逻辑到模糊逻辑·····	55
6.1.1	经典逻辑概要·····	55
6.1.2	模糊逻辑的基本原理·····	57
6.2	推理合成规则·····	59
6.3	规则的特性·····	61
6.3.1	广义取式推理·····	61
6.3.2	广义拒式推理·····	64
6.3.3	广义假言推理·····	65
6.4	本章小结及导读·····	66
6.5	习题·····	66

第二部分 模糊系统及其特性

第7章	模糊规则库与模糊推理机·····	70
7.1	模糊规则库·····	70

7.1.1	模糊规则库的结构	70
7.1.2	规则集合的特性	71
7.2	模糊推理机	73
7.2.1	组合推理	73
7.2.2	独立推理	74
7.2.3	各类推理机	75
7.3	本章小结及导读	80
7.4	习题	80
第 8 章	模糊器和解模糊器	81
8.1	模糊器	81
8.2	解模糊器	84
8.2.1	重心解模糊器	84
8.2.2	中心平均解模糊器	85
8.2.3	最大值解模糊器	86
8.2.4	解模糊器的比较	87
8.3	本章小结及导读	89
8.4	习题	90
第 9 章	模糊系统作为非线性映射	91
9.1	几类模糊系统的公式	91
9.1.1	带有中心平均解模糊器的模糊系统	91
9.1.2	带有最大值解模糊器的模糊系统	94
9.2	模糊系统作为万能逼近器	96
9.3	本章小结及导读	99
9.4	习题	99
第 10 章	模糊系统的逼近特性 I	100
10.1	基本概念	100
10.2	模糊系统的设计	102
10.3	模糊系统的逼近精度	104
10.4	本章小结及导读	108
10.5	习题	108
第 11 章	模糊系统的逼近特性 II	110
11.1	具有二阶逼近精度的模糊系统	110
11.2	带有最大值解模糊器的模糊系统的逼近精度	114
11.3	本章小结及导读	118

11.4	习题	118
第三部分 根据输入—输出数据设计模糊系统		
第 12 章	查表法设计模糊系统	121
12.1	设计模糊系统的查表法	121
12.2	查表法在倒车控制中的应用	124
12.3	查表法在时间序列预测中的应用	128
12.4	本章小结及导读	131
12.5	习题与设计	131
第 13 章	梯度下降法设计模糊系统	133
13.1	模糊系统结构的选择	133
13.2	模糊系统参数的设计	134
13.3	在非线性动态系统辨识中的应用	137
13.3.1	辨识器的设计	137
13.3.2	初始参数的选择	138
13.3.3	仿真	139
13.4	本章小结及导读	142
13.5	习题与设计	142
第 14 章	递推最小二乘法设计模糊系统	144
14.1	模糊系统的设计	144
14.2	递推最小二乘法的推导	145
14.3	在非线性通信信道均衡中的应用	146
14.3.1	均衡问题及其几何公式	146
14.3.2	模糊系统在均衡问题中的应用	149
14.4	本章小结及导读	152
14.5	习题与设计	152
第 15 章	聚类法设计模糊系统	154
15.1	一种最优的模糊系统	154
15.2	聚类法设计模糊系统	156
15.3	在非线性自适应控制系统中的应用	160
15.4	本章小结及导读	163
15.5	习题与设计	163

第四部分 非自适应模糊控制

第 16 章	模糊控制器设计的试错法	166
16.1	模糊控制与传统控制	166
16.2	模糊控制器设计的试错法	167
16.3	实例研究 I：水泥窑的模糊控制	167
16.3.1	水泥窑的工艺流程	167
16.3.2	水泥窑流程的模糊控制器的设计	168
16.3.3	模糊控制器的实施	170
16.4	实例研究 II：废水处理过程的模糊控制	171
16.4.1	污物废水处理过程	171
16.4.2	模糊控制器的设计	172
16.5	本章小结及导读	174
16.6	习题	174
第 17 章	线性系统模糊控制 I：稳定控制器	176
17.1	单输入单输出系统的稳定模糊控制	176
17.1.1	模糊控制系统的指数稳定性	177
17.1.2	模糊控制系统的输入—输出稳定性	179
17.2	多输入—多输出系统的稳定模糊控制	180
17.2.1	指数稳定性	180
17.2.2	输入—输出稳定性	183
17.3	本章小结及导读	183
17.4	习题	183
第 18 章	线性系统模糊控制 II：最优控制器与稳健控制器	185
18.1	最优模糊控制	185
18.1.1	Pontryagin 最小值原理	185
18.1.2	最优模糊控制器的设计	186
18.1.3	在球—横木系统中的应用	188
18.2	稳健模糊控制	189
18.3	本章小结及导读	190
18.4	习题	191
第 19 章	非线性系统模糊控制 I：滑动控制	192
19.1	作为滑动控制的模糊控制：分析	192
19.1.1	滑动控制的基本原理	192

19.1.2	根据滑动控制原理分析模糊控制器	194
19.2	作为滑动控制的模糊控制：设计	194
19.2.1	滑动控制律的连续逼近	194
19.2.2	根据平滑滑动控制律设计模糊控制器	196
19.3	本章小结及导读	198
19.4	习题	199
第 20 章	非线性系统模糊控制 II：监督控制	200
20.1	模糊系统中的多层控制	200
20.2	用非模糊监控实现稳定的模糊控制	201
20.2.1	监督控制器的设计	201
20.2.2	监督控制器在倒立摆平衡中的应用	203
20.3	用模糊系统调节 PID 控制器的增益	206
20.3.1	PID 控制器	206
20.3.2	调节 PID 增益的模糊系统	206
20.4	本章小结及导读	209
20.5	习题	210
第 21 章	模糊系统模型的模糊控制	211
21.1	TSK 模糊系统	211
21.2	带有模糊控制器的模糊模型的闭环动态特性	212
21.3	动态 TSK 模糊系统的稳定性分析	214
21.4	模糊模型的稳定模糊控制器的设计	218
21.5	本章小结及导读	220
21.6	习题	220
第 22 章	模糊控制的定性分析和多级模糊系统	222
22.1	模糊控制系统的相平面分析	222
22.2	稳定性的稳健性指标	224
22.2.1	一维情况	224
22.2.2	n 维情况	225
22.3	多级模糊控制	227
22.3.1	维度灾难	227
22.3.2	多级模糊系统的构造	227
22.3.3	多级模糊系统的性质	228
22.4	本章小结及导读	229
22.5	习题	229

第五部分 自适应模糊控制

第 23 章	基本自适应模糊控制器 I	233
23.1	自适应模糊控制器的分类	233
23.2	间接型自适应模糊控制器的设计	234
23.2.1	问题描述	234
23.2.2	模糊控制器的设计	234
23.2.3	自适应律的设计	237
23.3	在倒立摆跟踪控制中的应用	238
23.4	本章小结及导读	242
23.5	习题	242
第 24 章	基本自适应模糊控制器 II	244
24.1	直接型自适应模糊控制器的设计	244
24.1.1	问题描述	244
24.1.2	模糊控制器的设计	244
24.1.3	自适应律的设计	245
24.1.4	仿真	246
24.2	组合型自适应模糊控制器的设计	249
24.2.1	问题描述	249
24.2.2	模糊控制器的设计	250
24.2.3	自适应律的设计	250
24.2.4	收敛性分析	252
24.3	本章小结及导读	253
24.4	习题	253
第 25 章	高级自适应模糊控制器 I	254
25.1	保证状态有界的监督控制法	254
25.1.1	在间接型自适应模糊控制系统中的应用	254
25.1.2	在直接型自适应模糊控制系统中的应用	256
25.2	保证参数有界的投影法	257
25.2.1	在间接型自适应模糊控制系统中的应用	257
25.2.2	在直接型自适应模糊控制系统中的应用	258
25.3	稳定的直接型自适应模糊控制系统	259
25.3.1	稳定性和收敛性分析	259
25.3.2	仿真	260
25.4	本章小结及导读	261

25.5	习题	262
第 26 章	高级自适应模糊控制器 II	263
26.1	稳定的间接型自适应模糊控制系统	263
26.1.1	稳定性和收敛性分析	263
26.1.2	非线性参数化	264
26.2	一般非线性系统的自适应模糊控制	266
26.2.1	输入—输出线性化的直观概念	266
26.2.2	基于输入—输出线性化的自适应模糊控制器的设计	268
26.2.3	在球—横木系统中的应用	269
26.3	本章小结及导读	271
26.4	习题	271
第六部分 其他专题		
第 27 章	模糊 c -均值算法	274
27.1	为什么模糊模型可用于模式识别	274
27.2	硬 c -划分与模糊 c -划分	275
27.3	硬 c -均值算法和模糊 c -均值算法	276
27.3.1	目标函数聚类法和硬 c -均值算法	276
27.3.2	模糊 c -均值算法	278
27.4	模糊 c -均值算法的收敛性	280
27.5	本章小结及导读	282
27.6	习题	282
第 28 章	模糊关系方程	284
28.1	引言	284
28.2	求解模糊关系方程	284
28.3	模糊关系方程的可解度	287
28.3.1	两个模糊集的贴近度	287
28.3.2	模糊关系方程的可解度	288
28.4	近似解——神经网络算法	289
28.5	本章小结及导读	293
28.6	习题	294
第 29 章	模糊算术	295
29.1	模糊数和分解定理	295
29.2	模糊数的加和减	296
29.2.1	α -截集法	296

29.2.2	扩展原理法	297
29.3	模糊数的乘和除	298
29.3.1	α -截集法	298
29.3.2	扩展原理法	299
29.4	模糊方程	300
29.5	模糊排序	301
29.6	本章小结及导读	303
29.7	习题	303
第 30 章	模糊线性规划	305
30.1	模糊线性规划问题分类	305
30.2	模糊资源型线性规划问题	307
30.3	模糊目标系数型线性规划问题	308
30.4	模糊约束系数型线性规划问题	310
30.5	随机线性规划和模糊线性规划的比较	311
30.6	本章小结及导读	312
30.7	习题	313
第 31 章	可能性理论	315
31.1	引言	315
31.2	可能性的直观方法	315
31.2.1	可能性分布与可能性测度	315
31.2.2	边际可能性分布与非关联性	316
31.2.3	条件可能性分布	317
31.3	可能性的公理化方法	318
31.3.1	似真性测度和信任性测度	318
31.3.2	可能性测度和必要性测度	319
31.4	可能性与概率	320
31.4.1	可能性与概率的无休止的论战	320
31.4.2	两种理论的主要差别	321
31.4.3	从工程师的观点怎样看待这些争议	322
31.5	本章小结及导读	323
31.6	习题	323
参考文献		325
重要词汇英汉对照		335

前 言

近年来,模糊系统和模糊控制领域取得了迅速的进展。模糊控制在消费品和工业过程控制中的成功实践促使越来越多的学者对模糊系统和模糊控制理论进行严密的研究。这些学者试图通过理论研究解释实践成功的原因,并使模糊系统的现有研究方法系统化,从而逐步形成更加强大的方法。所有的这些努力都使整个模糊系统和模糊控制理论的前景变得更加清晰。虽然目前已有许多模糊理论方面的著作,但大部分著作要么是只研究某一专题,要么是论文集或是模糊数学方面的著作。因此,我们迫切需要一本真正的模糊系统和模糊控制方面的教材,它应该能够为我们提供这一领域的基本纲要,概括这一领域的基本原理。

本书是在香港科技大学的一门课程的基础上发展形成的,它旨在用作研究生和高年级本科生的教材,也可用作职业工程师的自学指导书籍。本书的特色如下。

1. 结构完好。本书并不是简单收集了至今为止模糊系统与模糊控制的成果,而是在编写具体内容前,先建立了模糊系统和模糊控制理论应该遵循的合理结构,然后再将具体内容嵌入这个结构。例如,研究模糊控制系统时,应该考虑系统的稳定性、最优性和稳健性;对方法进行分类时,应该考虑研究对象是线性的、非线性的,还是用模糊系统建模的。幸运的是,现今的主要成果与这一结构很匹配,因此,本书覆盖了现今成果的具体内容。与其他主流领域相比,这一领域尚不成熟,理论结构上还存在许多未解难题。对于这类专题,本书会给出它的初步方法,或指出问题尚未解决。

2. 清晰准确。对任何一本书来讲,清晰且有逻辑的表达都是非常关键的,尤其是对于一本与“模糊”相关的书。模糊理论本身是精确的;“模糊”存在于现象中,这也正是模糊理论试图研究的现象。一旦模糊描述(如,“热天”)通过模糊理论得以确切表达,一切就不再是模糊的了。书中采用准确的语言来介绍概念,推导方法,并证明结论。

3. 实践性强。实践应用是模糊系统和模糊控制理论发展的驱动力。本书中的大部分方法对于实际问题的解决都有很重要的意义。事实上,本书的一个主要目标就是告诉学生和职业工程师怎样使用模糊系统的方法来解决控制、信号处理和通信中的工程问题。

4. 丰富严密。本书中的大部分方法对于学生来讲是智力上的挑战。它除了强调实践性外,还给出了许多理论成果(当然,这些理论成果具有实际相关性和重要性)。所有的定理和引理都经过了严格的数学证明,其中某些成果是可以被一般学生理解的。

5. 易作教材。本书是按教科书方式编写的,所设计的每个章节可花费1学时~1.5学时。有时,两个学时可以包含三个章节的内容,反之亦然,这主要取决于教师的讲课重点和学生所具有的背景知识。每章都包含相应的习题和一些小设计,它们是教材完整不可缺少的部分。

本书分为六个部分:

第一部分(第2章至第6章)介绍模糊理论领域中的基本概念和基本理论,它们在模糊系统和模糊控制领域中非常有用。

第二部分(第7章至第11章)研究模糊系统,仔细分析了模糊系统各组成部分的运算,研究了模糊系统的一些特性(如逼近性和精度)。

第三部分(第12章至第15章)介绍了根据传感器测量的数据设计模糊系统的四种方法,并大量地在控制、信号处理或通信问题中检验了这些方法。

第四部分(第16章至第22章)和第五部分(第23章至第26章)集中研究了模糊控制,其中,第四部分研究了非自适应模糊控制,第五部分研究了自适应模糊控制。

第六部分(第27章至第31章)综述了大量不属于本书主体结构的专题,这些重要的专题与模糊系统和模糊控制高度相关。

本书的学习方法,取决于教师和读者的特定兴趣。第1章至第15章包含了可以应用于各种工程问题的一般资料。第16章至第26章专门研究控制问题。如果课程不是控制课程,则第16章至第26章中的一些资料可以忽略,将节省下来的时间用于更详细地学习第1章至第15章和第27章至第31章。相反,如果是控制课程,则应该仔细地学习第16章至第26章。如果是神经网络与模糊系统方面的课程,则本书可以与神经网络方面的教材一起使用。这时,教材中的第1章至第15章和第16章至第31章中的部分内容可作为模糊系统那一半课程的教材。如果职业工程师想快速地学会模糊系统和模糊控制理论,那么定理和推论的证明是可以跳过的。

本书的顺利完成受益于许多同事、同学和朋友的关注与支持。首先,感谢我的导师 Lotfi Zadeh 和 Jerry Mendel,感谢他们一直以来的鼓励。感谢 Karl Åström 在1995年夏天,推荐他的学生 Mikael Johansson 帮助我准备手稿。与模糊理论领域内的学者 Kevin Passino, Frank Lewis, Jyh-Shing Jang, Hua Wang, Hideyuki Takagi 以及其他学者的讨论对材料的组织也有很大的帮助。此外,本书还得到了香港科技大学参加课程的学生们的帮助。

非常感谢香港研究评审委员会(Hong Kong Research Grants Council)对作者的支持。

最后,感谢香港科技大学及本系,他们为我提供了极好的研究环境和教学环境。特别地,我要感谢我的同事 Xiren Cao, Zexiang Li, Li Qiu, Erwei Bai, Justin Chuang, Philip Chan 和 Kwan-Fai Cheung,感谢他们的协作和他们对模糊理论各种专题的批评指正。

王立新

香港科技大学

1.1 为什么研究模糊系统

根据牛津英语词典的解释,“模糊”(fuzzy)一词的含义为“朦胧的,模糊的;不精确的;不合乎逻辑的,不明白的”。我们希望读者忽略这种定义,将“模糊”一词看作一个技术形容词。实际上,模糊系统是一个被精确定义的系统,模糊控制也是一种被精确定义的特殊非线性控制。线性系统与控制中“线性”(linear)一词用于界定“系统与控制”,类似地“模糊系统与模糊控制”中的“模糊”一词也是用于界定一类特殊的“系统与控制”。从本质上看,我们想要强调的是,尽管模糊系统理论描述的现象可能是模糊的,但理论本身却是精确的。

在文献中,认为研究模糊系统理论的原因有两类:

- 现实世界太复杂以至于无法做到精确描述,所以为了得到一个合理的且可跟踪的模型就必须引入近似性(即模糊性)概念。
- 随着向信息时代的迈进,人类知识变得日益重要。我们需要一种理论,能系统地描述人类知识并将其同其他信息(如数学模型和感官测量)一起嵌入到工程系统中。

第一类原因虽然正确,但它并未概括出模糊系统理论独一无二的性质。实际上,几乎工程中的所有理论都是以一种近似的方法来描述现实世界的。举例来说,大多数实际系统都是非线性的,但我们却对线性系统的研究做了大量的努力。一个好的工程理论应该是精确的,应既能概括出现实世界的关键性质,又便于用数学分析的方法跟踪研究。在此方面,模糊系统理论同其他工程理论是没有区别的。

第二类原因,描述了模糊系统理论的独有特征并证明了它是作为工程学中的一个独立分支而存在的。根据一般性原则,一个好的工程理论应该能够有效利用所有可得的信息。对于大多数实际系统来说,有两个重要的信息来源:一个是用自然语言描述系统性能的专家;另一个是传感器提供的测量数据和根据自然法则推导出来的数学模型。因此,一项重要的任务就是怎样将这两类信息整合到系统设计中去。实现这种整合的关键在于怎样将人类知识整合到同传感器测量结果及数学模型类似的“框架”中。换句话说,关键在于怎样把一个人类知识库转换成一个数学公式。为了理解模糊系统如何实现这种转换,首先必须了解什么是模糊系统。

1.2 什么是模糊系统

模糊系统是一种基于知识或基于规则的系统。它的核心就是由所谓的 IF-THEN 规则所组成的知识库。一个模糊的 IF-THEN 规则就是一个用连续隶属度函数对所描述的某些句子所做的 IF-THEN 形式的陈述。例如,下面就是一个模糊 IF-THEN 规则:

如果一辆汽车的速度快,则施加给油门的力较小 (1.1)

这里,词语“快”和“较小”分别用图 1.1 和图 1.2 中的隶属度函数进行描述。模糊系统就是通过组合模糊 IF-THEN 规则构造而成的。下面看两个例子:

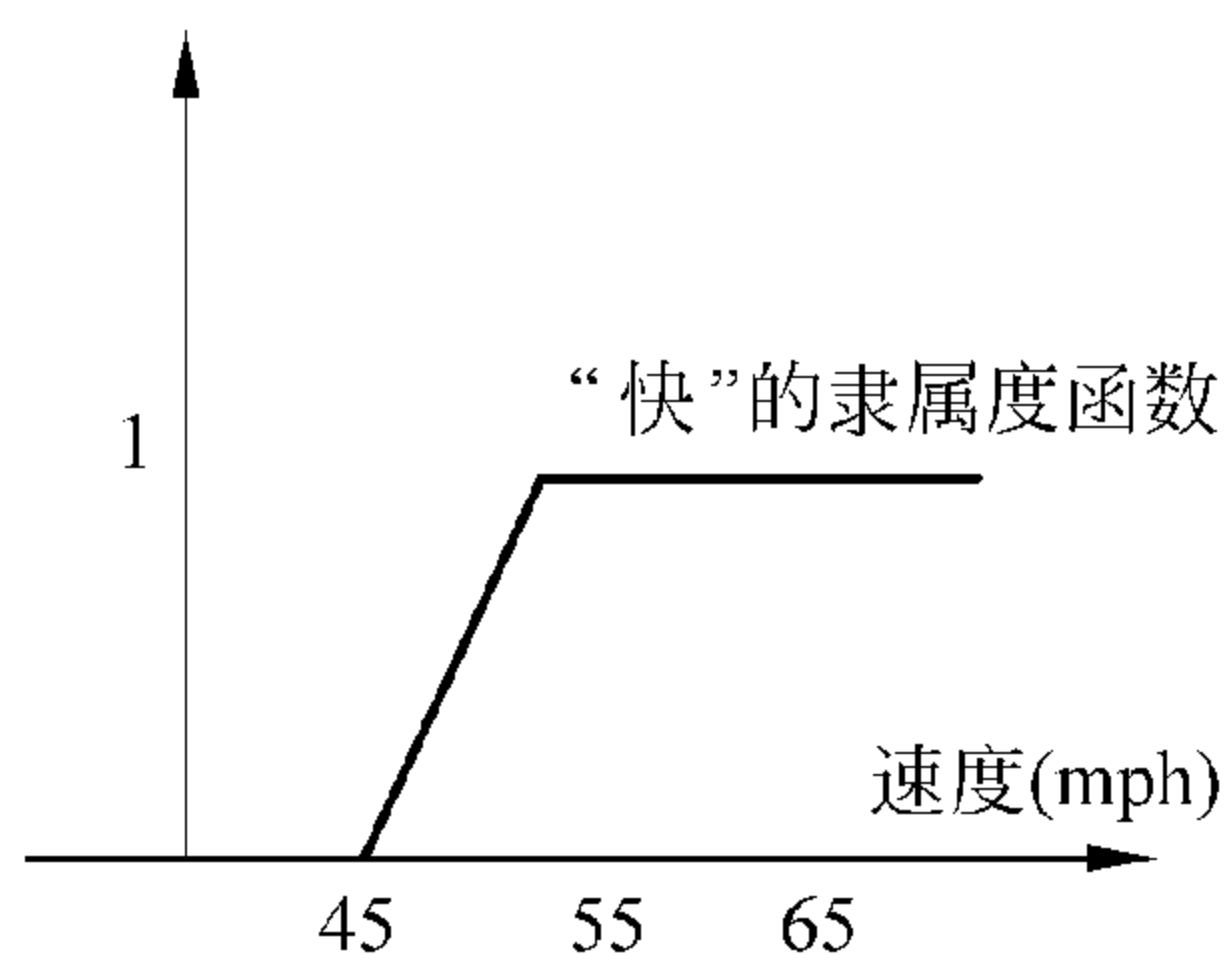
例 1.1 设想设计一个可以自动控制汽车速度的控制器。从概念上讲,有两种设计控制器的方法:第一种方法是采用传统的控制理论,比如设计一个 PID 控制器;第二种方法是模仿人类司机,也就是把人类司机所采用的规则转换到自动化控制器中来。现在,让我们看一下第二种方法。粗略地讲,司机在一般环境下采用如下三类规则来驾驶汽车:

如果速度慢,则施加给油门较大的力 (1.2)

如果速度适中,则施加给油门正常大小的力 (1.3)

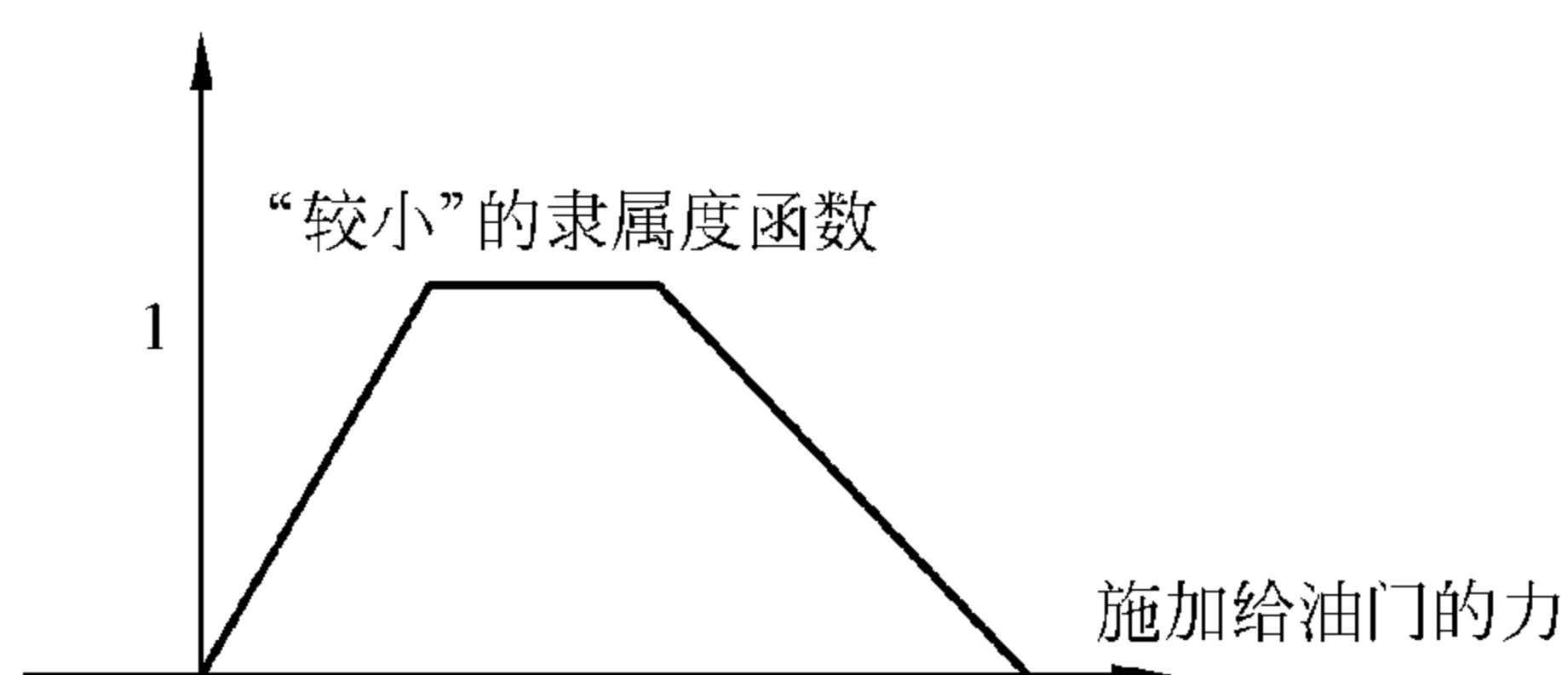
如果速度快,则施加给油门较小的力 (1.4)

这里“慢”,“较大”,“适中”,“正常大小”,“快”,“较小”都是用与图 1.1 和图 1.2 类似的隶属度函数来描述的。^①当然,实际情况可能需要更多的规则。我们可以根据这些规则来构造模糊系统。当把模糊系统作为控制器来使用时,也可以把这种控制器叫做模糊控制器。



注:图中横轴表示汽车的速度,
纵轴表示“快”的隶属度值。

图 1.1 “快”的隶属度函数



注:图中横轴表示施加给油门的力,
纵轴表示“较小”的隶属度值。

图 1.2 “较小”的隶属度函数

^① 隶属度函数详细的定义和分析将在第 2 章中给出。此处,根据图 1.1 和图 1.2 足以直观地理解隶属度函数。

例 1.2 在例 1.1 中,规则是一种控制指令。即,规则代表的是在一般环境下司机的行为。人类知识的另一类型就是对系统进行描述。设想一个人在给气球打气时,总希望知道在气球爆炸之前最多能打多少气,那么一些关键变量之间的关系就会非常有用。气球有三个关键变量:气球内的空气、空气上升量和表面张力。我们可以将这些变量之间的关系用以下模糊 IF-THEN 规则来描述:

如果空气数量少且空气数量稍微上升,则表面张力会稍微上升 (1.5)

如果空气数量少且空气数量显著上升,则表面张力会显著上升 (1.6)

如果空气数量多且空气数量稍微上升,则表面张力会适度上升 (1.7)

如果空气数量多且空气数量显著上升,则表面张力会非常显著上升 (1.8)

这里,“少”,“稍微”,“显著”等都是用与图 1.1 和图 1.2 中类似的隶属度函数描述出来的。把这些规则组合到一个模糊系统里,就会得到气球的模型。

总的说来,构造一个模糊系统的出发点就是要得到一组来自于专家或基于该领域知识的模糊 IF-THEN 规则,然后将这些规则组合到单一系统中。不同的模糊系统可采用不同的组合原则。接下来的问题就是:什么是最常用的模糊系统?

文献中最常见的模糊系统有三类:(1)纯模糊系统;(2)TSK(Takagi-Sugeno-Kang)模糊系统;(3)具有模糊器和解模糊器的模糊系统。下面简要介绍一下这三种模糊系统。

纯模糊系统的基本框图如图 1.3 所示。模糊规则库可表示为若干条模糊 IF-THEN 规则的集合。例如,对于例 1.1 中的汽车控制器来说,模糊规则库由式(1.2)~式(1.4)三条规则组成,而对于例 1.2 中的气球模型来说,模糊规则库由式(1.5)~式(1.8)四条规则组成。模糊推理机根据模糊逻辑原理,通

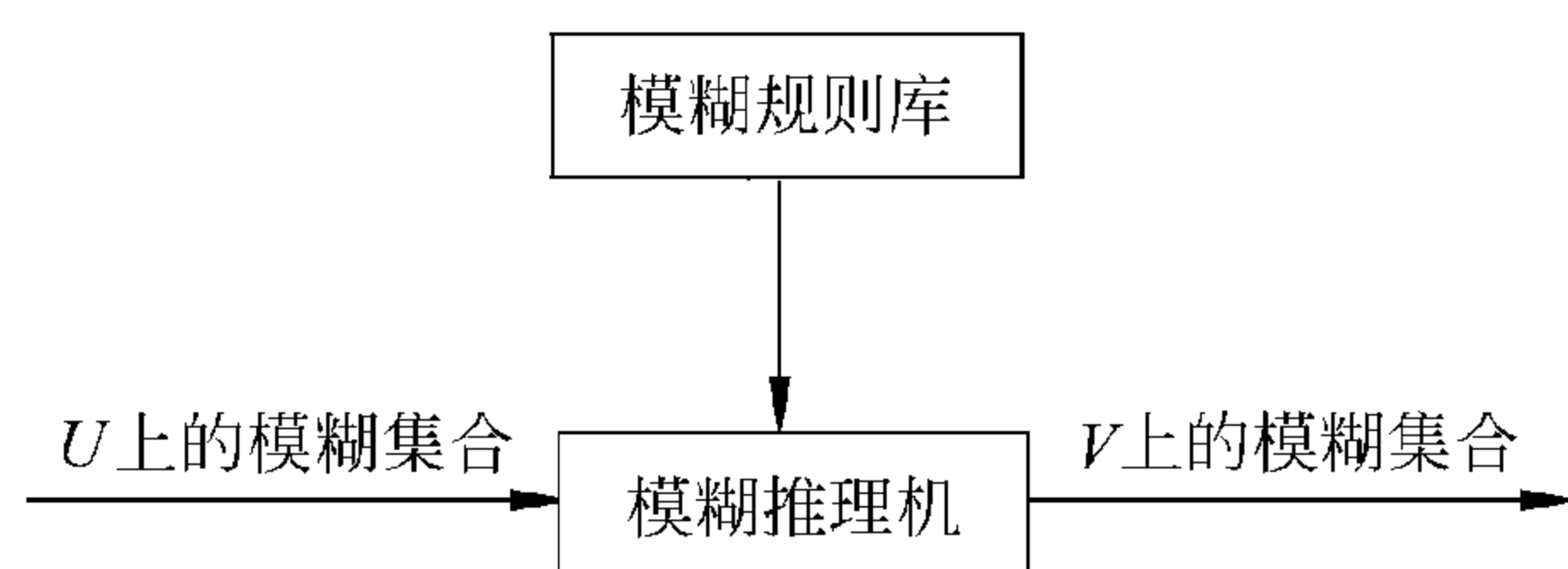


图 1.3 纯模糊系统的基本框图

过组合这些模糊 IF-THEN 规则来决定如何将输入论域 $U \subset R^n$ 上的模糊集合映射到输出论域 $V \subset R$ 上的模糊集合。^① 如果图 1.3 中画有少量反馈线,那么系统就变成了所谓的模糊动态系统。

纯模糊系统的主要问题在于其输入变量和输出变量均为模糊集合(即为采用自然语言描述的词语),而在工程系统中输入变量与输出变量均为真值(real-valued)变量。为了解决这一问题,Takagi,Sugeno 和 Kang(Takagi 和 Sugeno[1985],Sugeno 和 Kang[1988])提出了另一种输入变量和输出变量均为真值变量的模糊系统。

TSK 模糊系统未采用式(1.1)所示的 IF-THEN 规则,而是采用如下形式:

^① 模糊集合的精确定义将在第 2 章中给出,此处,把模糊集合看做“快”这类词语即可,“快”的隶属度函数见图 1.1。

如果一辆汽车的速度 x 为快, 则施加给油门的力为 $y = cx$ (1.9)

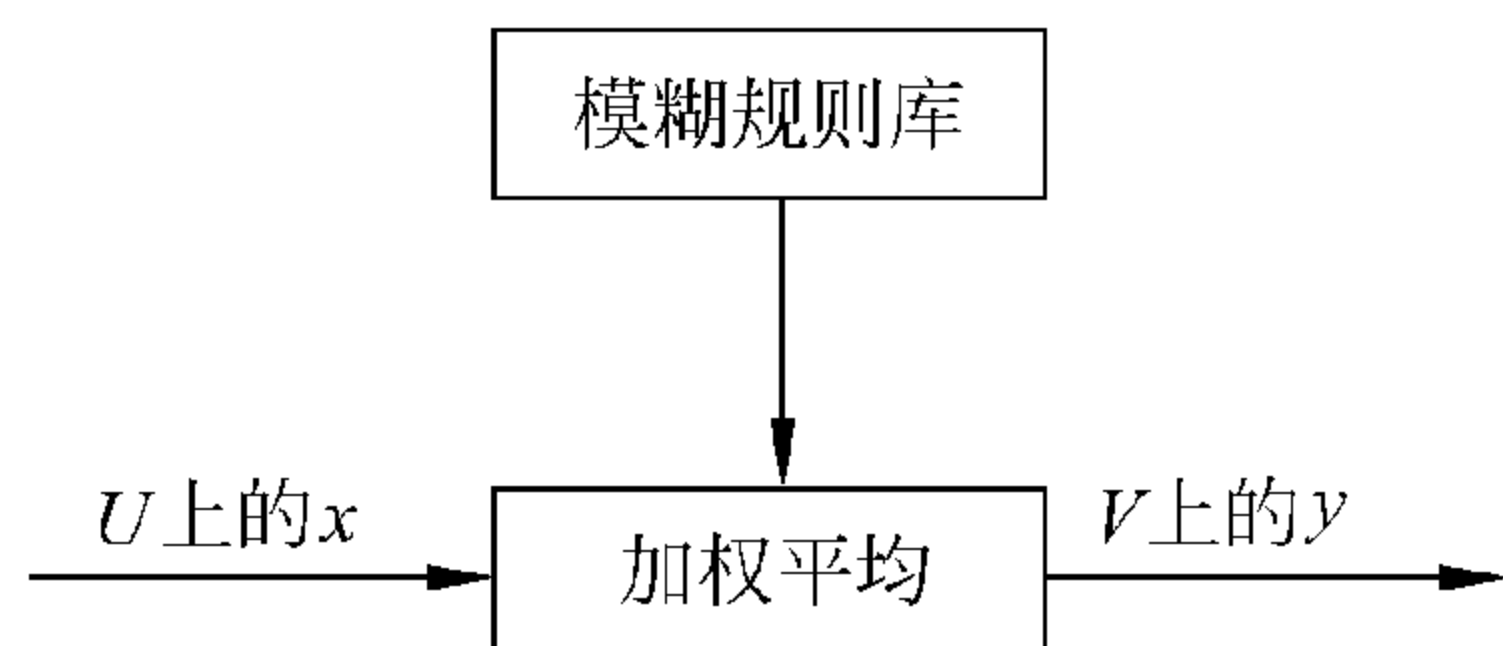


图 1.4 TSK 模糊系统的基本框图

这里的“快”和式(1.1)中的含义是一样的, c 是一个常数。通过比较式(1.9)和式(1.1)可以看出, 规则的 THEN 部分由自然语言描述的词语变成了一个简单的数学公式。这一变化使规则的组合更为容易。实际上, TSK 模糊系统就是对规则的 THEN 部分的值进行了加权平均。TSK 模糊系统的基本框图见图 1.4。

TSK 模糊系统的主要问题在于: (i) 其规则的 THEN 部分是一个数学公式, 所以可能无法提供一个自然的体系来表达人类知识; (ii) 模糊逻辑的各种原理得到应用的自由度也很有限, 以至于在该体系中没有很好地体现模糊系统的广泛用途。为了解决这些问题, 我们将采用第三类模糊系统——具有模糊器和解模糊器的模糊系统。

要想在工程系统中应用纯模糊系统, 一个简单的方法就是在纯模糊系统的输入端加上一个模糊器, 将真值变量转变成模糊集合, 在输出端加上一个解模糊器, 将模糊集合转变成真值变量。具有模糊器和解模糊器的模糊系统如图 1.5 所示, 这种模糊系统克服了纯模糊系统和 TSK 模糊系统的缺陷。从现在起, 除非特殊指明, 书中提到的模糊系统都是指这种具有模糊器和解模糊器的模糊系统。

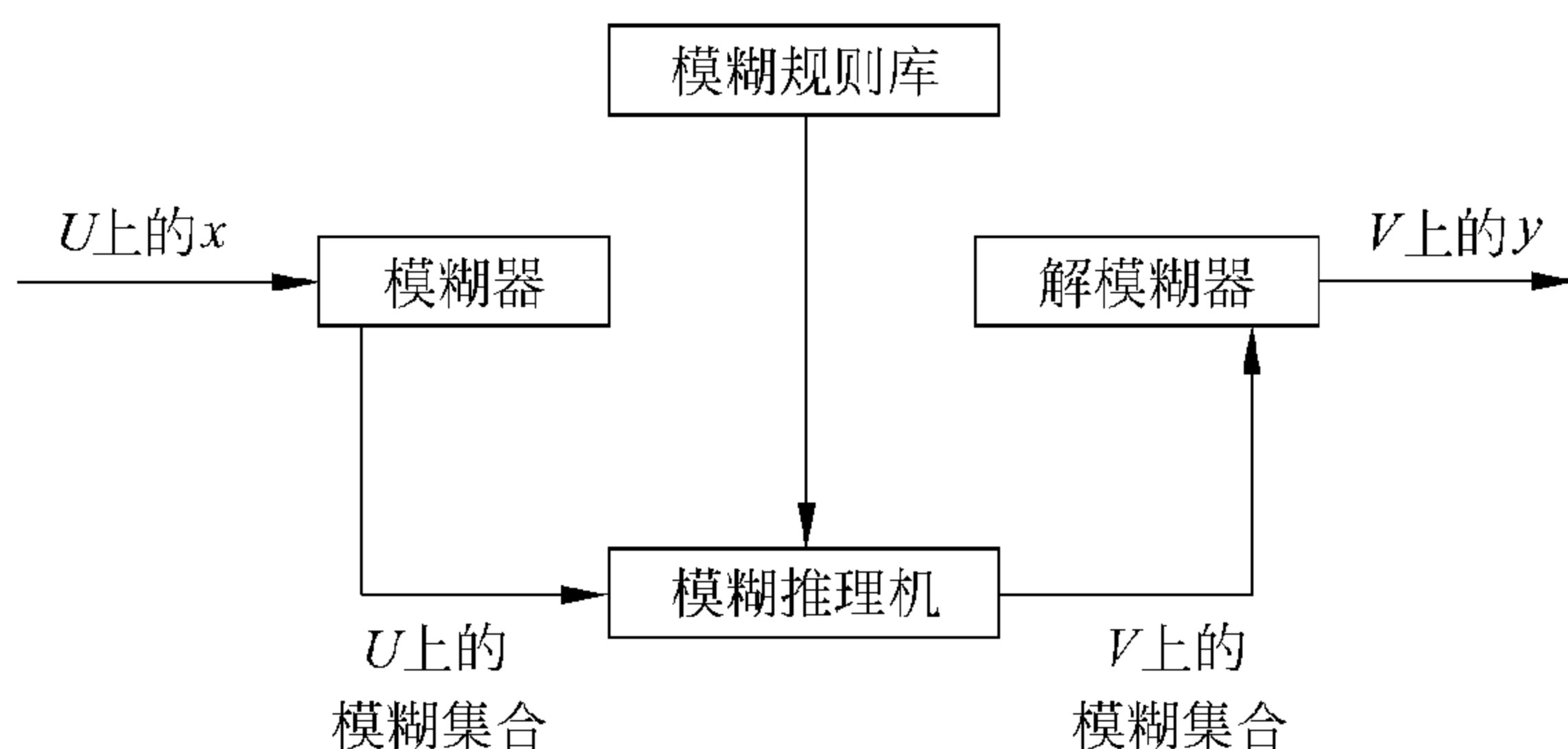


图 1.5 具有模糊器和解模糊器的模糊系统的基本框图

为总结本节内容, 我们要再次强调一下模糊系统的卓越特性: 一方面, 模糊系统是由一个实值向量向一个实值标量所作的多输入单输出映射(多输出映射可以被分解成一个单输出映射的集合), 并且能够得到这些映射的精确的数学公式(详见第 9 章); 另一方面, 模糊系统是由来自于以模糊 IF-THEN 规则为形式的人类知识所组成的基于知识的系统。模糊系统理论的一个重大贡献就是它为从知识库向非线性映射的转换提供了一套系统的程序。正是由于这一转换, 我们才能将基于知识的系统(模糊系统)采用同数学模型及传感器测量一样的方式, 应用到工程应用中(控制、信号处理及通信系统等)。这样, 最终组合而成的系统的分析和设计就会以数学这种严密方式来进行。本书的目的就是要说明这一转换的实现

过程以及如何分析与设计这种系统。

1.3 模糊系统应用的领域及方式

模糊系统一直广泛应用于各个领域,从控制、信号处理、通信、集成电路制造到商业专家系统、医药、行为科学等。不过,其最重大的应用一直集中在控制问题上,所以,这里不准备罗列模糊系统在不同领域的应用,而是集中在一些模糊系统占重要地位的控制问题上。

图 1.5 所示的模糊系统既可用作开环控制器(见图 1.6),也可用作闭环控制器(见图 1.7)。

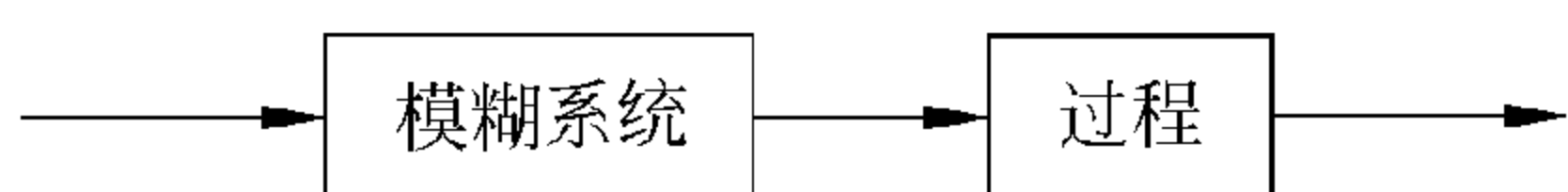


图 1.6 作为开环控制器的模糊系统

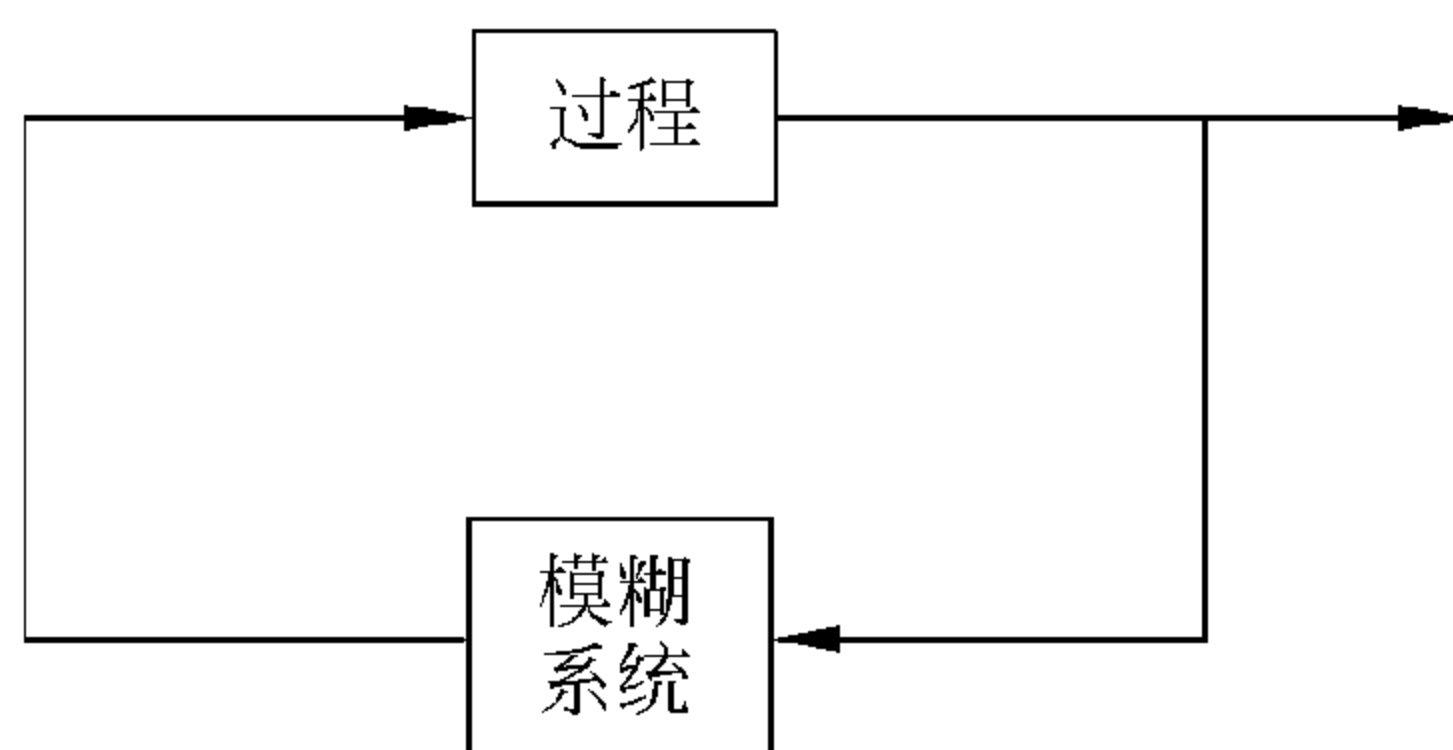


图 1.7 作为闭环控制器的模糊系统

当模糊系统作为开环控制器时,它通常会设置一些控制参数,然后系统根据这些控制参数来运行。许多应用于电子消费品的模糊系统就属于开环控制器。当模糊系统作为闭环控制器时,它会测量过程的输出,并连续地对过程进行控制。应用于工业过程的模糊系统就属于这类控制器。接下来,我们将简要描述模糊系统是如何应用于一些商业产品和工业系统的。

1.3.1 模糊洗衣机

模糊洗衣机是第一个应用模糊系统的消费产品,它是由日本松下(Matsushita)电子工业公司于 1990 年前后生产的。该洗衣机根据污物的种类、数量及机器负载量,运用模糊系统来自动设定正确的周期。具体地讲,它所用的模糊系统是一个三维输入一维输出的模糊系统,系统的三个输入变量是脏度、脏物的类型和负载量,输出变量是正确的周期。该洗衣机是通过传感器将三个输入变量输入到模糊系统中的。首先,光学传感器会射出一道穿过水的光线并计算有多少光线到达了另一端。水越脏,到达的光线越少。然后,光学传感器要辨别脏物是泥污还是油脂。泥污是很快能洗干净的。如果光的读数快速到达最小值的话,则脏物是泥污;如果下降较慢的话,则脏物是油脂;如果曲线斜率介于上述两斜率之间,则脏物是泥污油脂混合物。同时机器还有一个负载传感器,它能感知衣物的重量。很明显,衣物

量越大,所需的洗衣时间也就越长。将以上的启发式规则用几条模糊 IF-THEN 规则进行概括,然后再根据这些模糊 IF-THEN 规则构造模糊系统。

1.3.2 数字图像稳定器

任何使用过便携式摄像机的人都知道,人很难在拿着摄像机时手不发生轻微晃动。消除这种图像晃动,将会诞生极具流行商业价值的新一代摄像机。松下开发了一种叫做数字图像稳定器(digital image stabilizer)的产品,它以模糊系统为基础,当手晃动时会让画面保持稳定。数字图像稳定器是根据下面的启发式方法构造而成的模糊系统:

如果画面中所有的点都向同一个方向移动,则手晃动了 (1.10)

如果画面中仅有某些点移动,则手没有晃动 (1.11)

而且,数字图像稳定器还能将当前画面同存储器中以前的画面进行比较。如果整体特征变化明显,则根据规则(1.10)可知手晃动了,这时模糊系统会调整画面,否则将不做任何调整。由此可知,当一辆汽车穿过田野时,如果仅有一部分画面发生变化,则该摄像机是不会尝试修复画面的。从这个角度看,尽管手晃动了,画面依然会保持稳定。

1.3.3 汽车中的模糊系统

一辆汽车是许多系统的集合体——发动机、传动杆、刹车、离合器、方向盘等等——而模糊系统则几乎可以被应用到所有这些系统中去。例如,尼桑(Nissan)曾经根据如下的观察发明了一个模糊自动化传动杆,它可以节约 12% ~ 17% 的燃料:一般的传动杆只要汽车超过了某一速度就要换档,所以它换档频繁且每次换档都要耗费汽油。而司机不仅要考虑不能频繁换档,而且还要考虑一些非速度因素。例如在加速爬坡时,他们可能会延迟换档。尼桑的模糊自动化传动杆装置把这些启发式规则总结到模糊 IF-THEN 规则库中,并用它构造了一个能够指导换档的模糊系统。

尼桑还开发了一个模糊反锁刹车系统,它的挑战是,在将最大压力施加给刹车时不会引起熄火。尼桑系统考虑了以下启发式规则:

如果汽车非常快地减速,则系统会假设刹车上锁,并释放压力 (1.12)

1992年3月,三菱(Mitsubishi)公司公布了一种模糊多用途系统,它能控制一辆汽车的传动杆、离合器、牵引、四轮方向盘、四轮驱动器及空调。模糊传动杆会按曲线调到低速档并会防止汽车不恰当地曲线地调高车速档或者在司机踩下油门时调高车速档。模糊离合器包含了一个位于汽车前面的传感器,它可以感知路面的起伏和高度变化,然后用一种更平稳的方式来调整离合器。模糊牵引可用来防止汽车在转弯时超速,还可根据光滑的路面是平路还是坡路来改善对汽车的控制。最后,模糊方向盘会根据路面和汽车速度来调整后轮的反

应角。而模糊空调则会通过监控阳光、温度和湿度来改善车内环境。

1.3.4 水泥窑的模糊控制

水泥是通过磨细水泥渣块制成的。水泥渣块则是石灰石、粘土和砂子的混合物经水泥窑加热制成的。由于水泥窑表现出来的是时变的非线性行为,所以几乎没有可用的测量结果,很难用传统的控制理论控制水泥窑。

在20世纪70年代后期,丹麦的 Holmblad 和 Østergaard 开发了一个控制水泥窑的模糊系统。这个模糊系统(模糊控制器)有四个输入变量和两个输出变量(它可以看做是形如图1.5的两个模糊系统,这两个模糊系统的输入相同)。四个输入变量分别为:(1)废气中氧的比例;(2)排出气体的温度;(3)水泥窑的驱动扭矩;(4)炼砖的公升重量(可表明燃烧区的温度和炼砖质量)。两个输出变量分别为:(1)煤的供给率;(2)气流。通过所建立的模糊 IF-THEN 规则集合能将输入变量和输出变量对应起来。例如,可用到下面的两条规则:

如果氧的比例高且温度低,则增加气流 (1.13)

如果氧的比例高且温度高,则稍微降低煤的供给率 (1.14)

将这些规则组合到模糊系统中就形成了模糊控制器。1978年6月,模糊控制器在丹麦的 F. L. Smidth 公司的水泥窑运行了六天——这是模糊控制在整个工业过程实验中的首次成功,它证明了模糊控制器既可稍微改善人工操作者的工作效果,也可降低燃料的耗费。对此我们将在第16章作更详细的阐述。

1.3.5 地铁的模糊控制

迄今为止,模糊系统最显著的应用可能就是在日本仙台(Sendai)地铁中的应用。该地铁单行南北线全长13.6公里,途经16个站点,但其运行却非常地平稳。该模糊控制系统同时考虑了四种性能指标:安全性、乘坐舒适性、目标速度的可跟踪性和停车距离的准确性。它由两部分组成:常量速度控制器(用于发动地铁并使地铁速度保持在安全界限以下)和自动刹车控制器(用于调整地铁速度以使地铁能停在目标位置)。常量速度控制器是根据下面的规则设计的:

安全性:如果地铁速度接近极限速度,则选择最大刹车级 (1.15)

乘坐舒适性:如果地铁速度在允许范围内,则不改变控制级 (1.16)

至于实际系统的目标速度的可跟踪性及其他因素用到的规则就更多了。

自动停车控制器是由下面的规则构造而成的:

乘坐舒适性:如果火车将在允许区域内停下来,则不改变控制级 (1.17)

乘坐舒服和安全性：如果火车在允许区域内，则控制级将由“加速”调到“轻微刹车”。 (1.18)

至于实际系统的停车距离的准确性及其他因素用到的规则就更多了。到 1991 年为止，仙台的地铁已经运行了四年，而且仍然是最先进的地铁系统之一。

1.4 模糊理论的主要研究领域

我们所说的模糊理论是指用到了模糊集合的基本概念或连续隶属度函数的理论。根据图 1.8 可将模糊理论进行大致的分类。主要有五个分支：

- (i) 模糊数学，它用模糊集合取代经典集合从而扩展了经典数学中的概念；
- (ii) 模糊逻辑与人工智能，它引入了经典逻辑学中的近似推理，且在模糊信息和近似推理的基础上开发了专家系统；
- (iii) 模糊系统，它包含了信号处理和通信中的模糊控制和模糊方法；

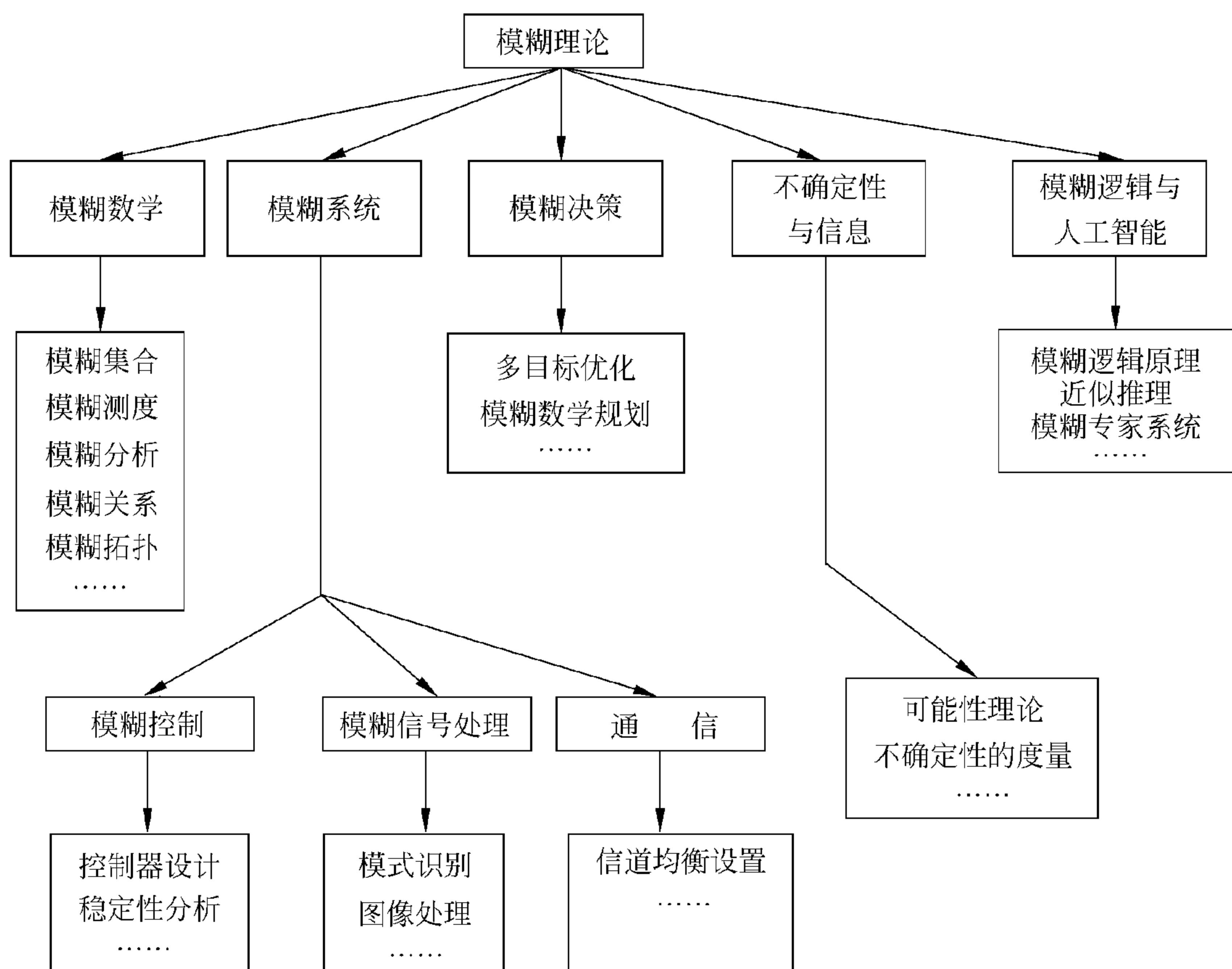


图 1.8 模糊理论的分类

(iv) 不确定性和信息,它用于分析各种不确定性;

(v) 模糊决策,它用软约束来考虑优化问题。

当然,这五个分支并不是完全独立的,他们之间有紧密的联系。例如,模糊控制就会用到模糊数学和模糊逻辑中的概念。

从实际应用的观点来看,模糊理论的应用大部分集中在模糊系统上,尤其集中在模糊控制上,这从1.3节中的例子就可以看出来。也有一些模糊专家系统应用于医疗诊断和决策支持(Terano, Asai 和 Sugeno [1994])。由于模糊理论从理论和实践的角度看仍然是新生事物,所以我们期望,随着模糊领域的成熟,将会出现更多可靠的实际应用。

由图1.8可知,模糊理论是一个包含了多种研究课题的广阔领域。在本书中,我们集中于模糊系统与模糊控制。首先,我们将学习模糊系统和模糊控制所用到的模糊数学和模糊逻辑中的基本概念(第2章至第6章)。然后,详细研究模糊系统和模糊控制(第7章至第26章)。最后,简要回顾一下模糊理论中其他领域的一些问题(第27章至第31章)。

1.5 模糊理论及其应用的简史

1.5.1 20世纪60年代:模糊理论的萌芽

模糊理论是由 Lotfi A. Zadeh 于1965年在名为“模糊集合”(Zadeh[1965])的开创性文章中创立的。在研究模糊理论以前,Zadeh 是一位很有威望的控制论学者。他提出的“状态”的概念形成了现代控制论的基础。在20世纪60年代初期,他认为经典控制论过于强调精确性反而无法处理复杂的系统,正如他在1962年的文章中提到的,“在处理生物系统时,需要一种彻底不同的数学——关于模糊量的数学,该数学不能用概率分布来描述”(Zadeh [1962])。后来,他将这些思想正式形成文章“模糊集合”(Fuzzy Sets)。

自模糊理论诞生之日起,它就一直处于各派的激烈争论之中。一些学者,如 Richard Bellman 认可了这一理论并开始着手在这一新领域中进行研究。而其他一些学者则反对这一理论,认为“模糊化”与基本的科学原则相违背。不过最大的挑战还是来自于统计和概率论领域的数学家们,他们认为概率论已足以描述不确定性,而且任何模糊理论可以解决的问题,概率论也都可以解决得一样好或更好(见第31章)。由于模糊理论在初期没有实际应用,所以它很难击败这种纯哲学观点的质疑。当时几乎世界上所有的大型研究机构都未将模糊理论作为一个重要的研究领域。

尽管模糊理论没有成为主流,但世界各地仍有许多学者毕生致力于这一新领域的研究。在20世纪60年代后期,这些学者提出了许多新的模糊方法,如模糊算法、模糊决策等。

1.5.2 20 世纪 70 年代：模糊理论继续发展并出现了实际的应用

公平地说,模糊理论成为一个独立的领域,很大程度上归功于 Zadeh 的贡献及其杰出的研究工作。模糊理论的大多数基本概念都是由 Zadeh 在 20 世纪 60 年代末 70 年代初提出来的。他在 1965 年提出模糊集合后,又在 1968 年提出模糊算法的概念(Zadeh[1968]),在 1970 年提出模糊决策(Bellman 和 Zadeh[1970]),在 1971 年提出了模糊排序(Zadeh[1971b])。1973 年他发表了另一篇开创性文章《分析复杂系统和决策过程的新方法纲要》,该文建立了研究模糊控制的基础理论,在引入语言变量这一概念的基础上,提出了用模糊 IF-THEN 规则来量化人类知识。

20 世纪 70 年代的一个重大事件就是诞生了处理实际系统的模糊控制器。在 1975 年, Mamdani 和 Assilian 创立了模糊控制器的基本框架(实质上就是图 1.5 中的模糊系统),并将模糊控制器用于控制蒸汽机。他们的研究成果发表在文章《带有模糊逻辑控制器的语言合成实验》(Mamdani 和 Assilian [1975])中,这是关于模糊理论的另一篇具有开创性的文章。他们发现模糊控制器非常易于构造且运作效果较好。后来,在 1978 年, Holmblad 和 Østergaard 为整个工业过程开发出了第一个模糊控制器——模糊水泥窑控制器(见 1.3 节)。

总的说来,公认的模糊理论的基础创建于 20 世纪 70 年代。随着许多新概念的引进,模糊理论作为一门新领域的前景已经日益清晰了。像模糊蒸汽机控制器和模糊水泥窑控制器这类最初的应用也已经表明了这一领域的潜力。通常来说,一个领域的开拓应该是通过大型研究机构将主要资源放在该领域来实现的。然而不幸的是,实际情况并非如此。20 世纪 70 年代末 80 年代初,尤其是在美国,模糊理论的许多学者都是由于无法找到继续研究的支持而转向其他领域的。

1.5.3 20 世纪 80 年代：模糊理论的大规模应用使其产生巨大飞跃

从理论角度讲,20 世纪 80 年代初这一领域的进展缓慢。这期间没有提出什么新的概念和方法,这是因为几乎没有人继续从事该领域的研究,只有模糊控制方面的应用保存下来。

日本工程师们,以其对新技术的敏感,迅速地发现模糊控制器对许多问题来讲都是易于设计的,而且操作效果也非常好。因为模糊控制不需要过程的数学模型,所以它可以应用到很多因数学模型未知而无法使用传统控制论的系统中去。1980 年, Sugeno 开创了日本的首次模糊应用——控制一家富士(Fuji)电子水净化工厂。1983 年,他又开始研究模糊机器人,这种机器人能够根据呼唤命令来自动控制汽车的停放(Sugeno 和 Nishida [1985])。20 世纪 80 年代初,来自于日立公司的 Yasunobu 和 Miyamoto 开始给仙台地铁开发模糊系统。他们于 1987 年结束了该项目,并创造了世界上最先进的地铁系统。模糊控制的这一应用非常

振奋人心并引起了模糊领域的一场巨变。

1987年7月,第二届国际模糊系统协会年会在东京召开。会议是在仙台地铁开始运行后三天召开的,与会者们梦幻般地历经了一次愉快的旅行。并且,Hirota还在会议上演示了一种模糊机器人手臂,他能实时地做二元空间内的乒乓动作(Hirota, Arai 和 Hachisu [1989]),Yamakawa也证明了模糊系统可以保持倒立摆的平衡(Yamakawa[1989])。此前,模糊理论在日本一直没有盛行起来。但此后,支持模糊理论的浪潮迅速蔓延到工程、政府以及商业团体中。到了20世纪90年代初,市场上已经出现了大量的模糊消费产品(见1.3节中的例子)。

1.5.4 20世纪90年代:模糊理论仍有更多的挑战

日本模糊系统的成功震惊了美国和欧洲主流学者们。一些学者仍对模糊理论持批评态度,但更多的学者不仅已经转变观念,而且还给予了模糊理论发展壮大机会。1992年2月,首届IEEE模糊系统国际会议在圣地亚哥召开了,这次大会标志着模糊理论已被世界上最大的工程师协会——IEEE所接受,而且IEEE还于1993年创办了IEEE模糊系统会刊。

从理论角度来看,模糊系统与模糊控制在20世纪80年代末90年代初的发展是迅猛的。尽管很难说有什么突破,但对于模糊系统与模糊控制中的一些基本问题上的研究已经取得了可喜的进步,例如利用神经网络技术系统地确定隶属度函数及严格分析模糊系统的稳定性。尽管模糊系统应用于控制理论的整体图景已经越来越清晰,但仍有大量的工作要做,大多数的方法和分析仍停留在初级阶段。我们确信只有当顶尖研究机构将重要人力放在模糊理论研究上时,该理论才能产生巨大的进步。

1.6 本章小结及导读

本章主要阐述了以下内容:

- 使用模糊系统的目的是为了将人类知识以一种系统的、有效的、可分析的方式嵌入工程系统中;
- 最常用的模糊系统的基本框架;
- 用于某些工业过程和消费品中的模糊IF-THEN规则;
- 模糊理论和应用的分类及简史。

McNeill 和 Freiburger[1993]对模糊理论及其应用做了非常好的非技术性的介绍,文中包括了对许多重要事件的访问和描述。Kruse, Gebhardt 和 Klawonn[1994]做了一些历史性的评论。Klir 和 Yuan[1995]出版了可能是一本关于模糊集合和模糊逻辑的最具综合性的书。模糊控制的早期应用都被收集在 Sugeno[1985]的书中。而近年来的应用(主要是日

本)则在 Terano, Asai 和 Sugeno[1994]的书中有总结。

1.7 习题

习题 1.1 模糊洗衣机是一个开环控制系统还是一个闭环控制系统? 模糊水泥窑控制系统呢? 请解释你的结论。

习题 1.2 列举 4 条 ~6 条模糊理论在实际问题中的应用(不包括 1.3 节中给出的那些应用), 并指出你所找到的这些应用所在的参考文献。

习题 1.3 假设我们想设计一个模糊系统来平衡如图 1.9 所示的倒立摆, 令摆角 θ 和其导数 $\dot{\theta}$ 作为模糊系统的输入, 施加给小车的力 u 作为模糊系统的输出。

(a) 根据常识中倒立摆的平衡问题确定 3 条 ~5 条模糊 IF-THEN 规则。

(b) 假设(a)中的规则可以成功地控制一个特定的倒立摆系统。现在, 如果想用这些规则来控制另一个倒立摆, 其 m_c, m 和 l 值与原来不同, 那么规则的什么部分应该改变, 规则的什么部分可保持不变?

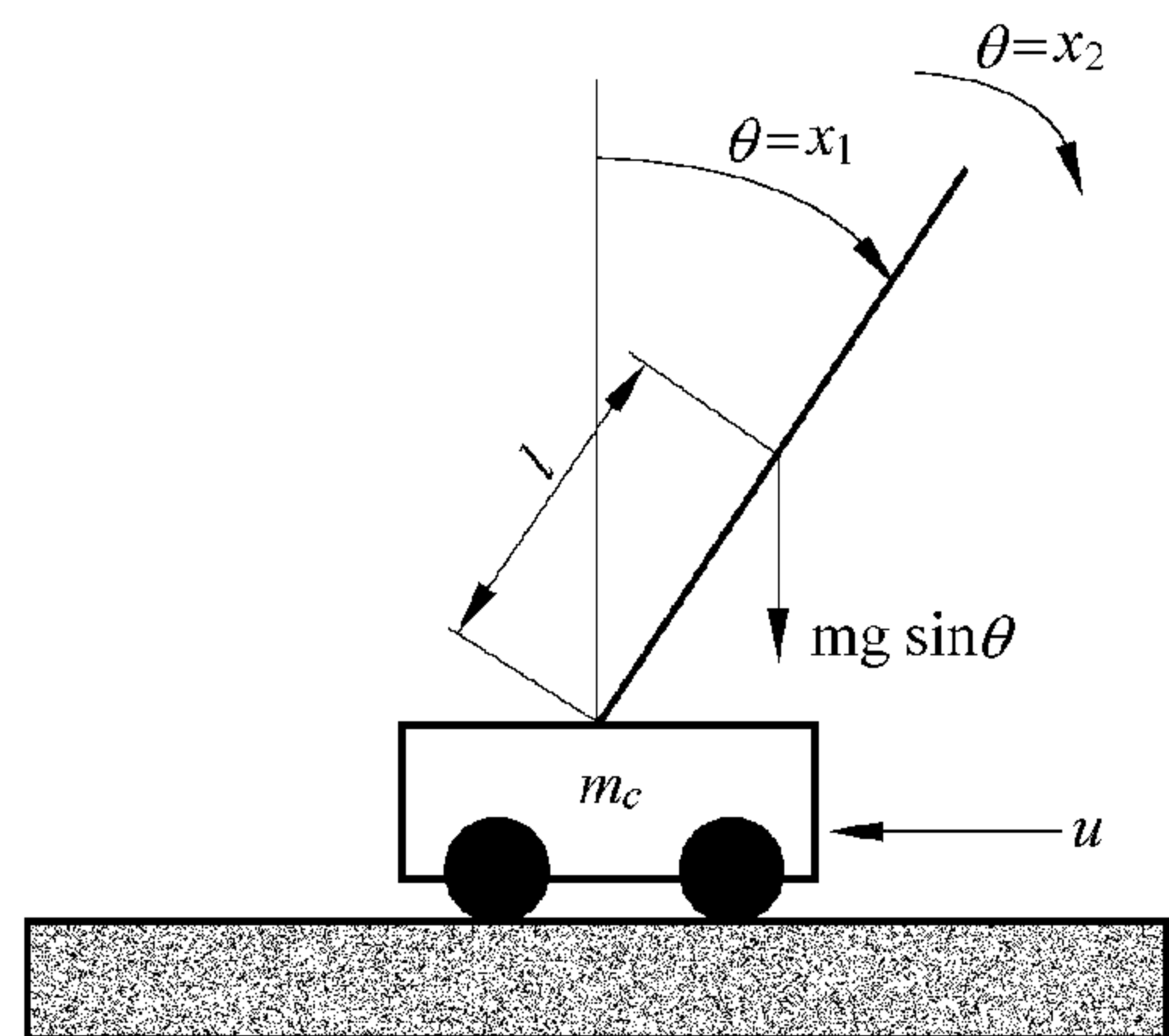


图 1.9 倒立摆系统

第一部分

模糊系统与模糊控制的 模糊数学原理

模糊数学为模糊系统与模糊控制的发展提供了起点和基本语言。模糊数学本身就是一个巨大的领域,其原理是由用模糊集合的概念取代经典数学理论中的集合概念而发展来的。按照这种方式,所有的经典数学分支都可以被“模糊化”,于是诞生了模糊测度理论、模糊拓扑、模糊算术和模糊分析等等分支。显然,模糊数学中仅有一部分可以应用到工程中去。因此,在接下来的5章中,我们将学习模糊数学中对模糊系统与模糊控制有用的那些概念和原理。

在第2章中,我们将学习模糊理论中最基本的概念——模糊集合。在第3章中,将详细学习模糊集合的集合运算,如模糊补、模糊并和模糊交。在第4章中,将学习模糊关系,并介绍模糊理论中的一条重要原理——扩展原理。语言变量和模糊 IF-THEN 规则将在第5章中作精确定义和研究,它也是模糊系统与模糊控制的本质所在。最后,第6章将集中研究模糊逻辑的三条基本原理,它们对于模糊系统中的模糊推理机是很有用的。

2.1 由经典集合到模糊集合

设 U 为论域或全集,它是具有某种特定性质或用途的元素的全体。回顾一下论域 U 中经典(清晰)集合 A 的概念,或简单地说,集合 A 的概念。集合 A 可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)。列举法仅用于有限集,所以其使用范围有限;描述法则比较常用。在描述法中,集合 A 可以表示为

$$A = \{x \in U \mid x \text{ 满足某些条件}\} \quad (2.1)$$

还有第三种定义集合 A 的方法——隶属度法,该方法引入了集合 A 的 0-1 隶属度函数(也可叫做特征函数、差别函数或指示函数),用 $\mu_A(x)$ 表示,它满足

$$\mu_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases} \quad (2.2)$$

集合 A 等价于其隶属度函数 $\mu_A(x)$,从这个意义上讲,知道 $\mu_A(x)$ 与知道 A 是一样的。

例 2.1 设伯克利(Berkeley)的所有汽车的集合作为论域 U ,我们可以根据汽车的特征来定义 U 上的不同集合。图 2.1 给出了可用于定义 U 上的集合的两类特征:(a)美国汽车或非美国汽车;(b)汽缸数量。例如,定义 U 上所有具有四个汽缸的汽车为集合 A ,即

$$A = \{x \in U \mid x \text{ 具有 4 个汽缸}\} \quad (2.3)$$

或

$$\mu_A(x) = \begin{cases} 1 & \text{如果 } x \in U \text{ 且 } x \text{ 有 4 个汽缸} \\ 0 & \text{如果 } x \in U \text{ 且 } x \text{ 没有 4 个汽缸} \end{cases} \quad (2.4)$$

美国 汽车	4 个汽缸
	6 个汽缸
非美国 汽车	8 个汽缸
	其 他

注:把伯克利的所有汽车这一集合分割成子集:

(a) 美国汽车和非美国汽车;(b) 汽缸数量。

图 2.1 伯克利汽车集合的子集分割图

如果想根据汽车是美国汽车还是非美国汽车来定义一个 U 上的集合的话,将存在一定困难。一种办法是,如果汽车具有美国汽车制造商的商标,则认为该汽车是美国汽车;否则就认为该汽车是非美国汽车。不过,很多人感觉美国汽车与非美国汽车之间的差异并不是那么分明,因为美国汽车(如福特、通用和克莱斯勒)的许多零部件都不是在美国生产的。此外,有一些“非美国”汽车却是在美国制造的。那么,怎样处理这类问题呢?

从本质上看,例 2.1 中的困难说明了某些集合并不具有清晰的边界。经典集合理论中的集合要求具有一个定义得很准确的性质,因此,经典集合无法定义诸如“伯克利的所有美国汽车”这样的集合。为了克服经典集合理论的这种局限性,模糊集合的概念应运而生。它也说明了经典集合的这种局限性是本质上的,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 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 $[0,1]$ 。

因此,模糊集合是经典集合的一种推广,它允许隶属度函数在区间 $[0,1]$ 内任意取值。换句话说,经典集合的隶属度函数只允许取两个值——0 或 1,而模糊集合的隶属度函数则是区间 $[0,1]$ 上的一个连续函数。由定义可以看出,模糊集合一点都不模糊,它只是一个带有连续隶属度函数的集合。

U 上的模糊集合 A 可以表示为一组元素与其隶属度值的有序对的集合,即

$$A = \{(x, \mu_A(x)) \mid x \in U\} \quad (2.5)$$

当 U 连续时(如 $U = R$), A 一般可以表示为

$$A = \int_U \mu_A(x) / x \quad (2.6)$$

这里的积分符号并不表示积分,而是表示 U 上隶属度函数为 $\mu_A(x)$ 的所有点 x 的集合。

当 U 取离散值时, A 一般可以表示为:

$$A = \sum_U \mu_A(x) / x \quad (2.7)$$

这里的求和符号并不表示求和,而是表示 U 上隶属度函数为 $\mu_A(x)$ 的所有点 x 的集合。

现在我们回到例 2.1,考虑怎样用模糊集合的概念来定义美国汽车和非美国汽车。

例 2.1(续) 可以根据汽车的零部件在美国制造的百分比,将集合“伯克利的美国汽车”(用 D 表示)定义为一个模糊集合。具体来说,可用如下的隶属度函数来定义 D :

$$\mu_D(x) = p(x) \quad (2.8)$$

其中, $p(x)$ 是汽车的零部件在美国制造的百分比,它在 0% 至 100% 之间取值。例如,如果某汽车 x_0 有 60% 的零件在美国制造,则可以说汽车 x_0 属于模糊集合 D 的程度为 0.6。

类似地,可以用下面的隶属度函数来定义集合“伯克利的非美国汽车”(用 F 表示):

$$\mu_F(x) = 1 - (px) \quad (2.9)$$

其中, $p(x)$ 的含义与式(2.8)中的 $p(x)$ 相同, 那么, 如果某汽车 x_0 有 60% 的零件在美国制造, 则可以说汽车 x_0 属于模糊集合 F 的程度为 $1 - 0.6 = 0.4$ 。式(2.8)和式(2.9)的定义可参见图 2.2 中的例子。显然, 一种元素可以以相同或不同的程度属于不同的模糊集合。

现在, 让我们考虑另一个模糊集合的例子, 从中可以得出一些结论。

例 2.2 令 Z 表示模糊集合“接近于 0 的数”, 则 Z 的隶属度函数可能为

$$\mu_Z(x) = e^{-x^2} \quad (2.10)$$

式中, $x \in R$ 。这是一个均值为 0, 标准差为 1 的高斯函数。根据这一隶属度函数可知, 0 和 2 属于模糊集合 Z 的程度分别为 $e^0 = 1$ 和 e^{-4} 。

也可以将 Z 的隶属度函数定义为

$$\mu_Z(x) = \begin{cases} 0 & x < -1 \\ x+1 & -1 \leq x \leq 0 \\ 1-x & 0 \leq x < 1 \\ 0 & x \geq 1 \end{cases} \quad (2.11)$$

根据此隶属度函数可知, 0 和 2 属于模糊集合 Z 的程度分别为 1 和 0。式(2.10)和式(2.11)的图形分别见图 2.3 和图 2.4。我们还可以选择许多其他的隶属度函数来描述“接近于 0 的数”。

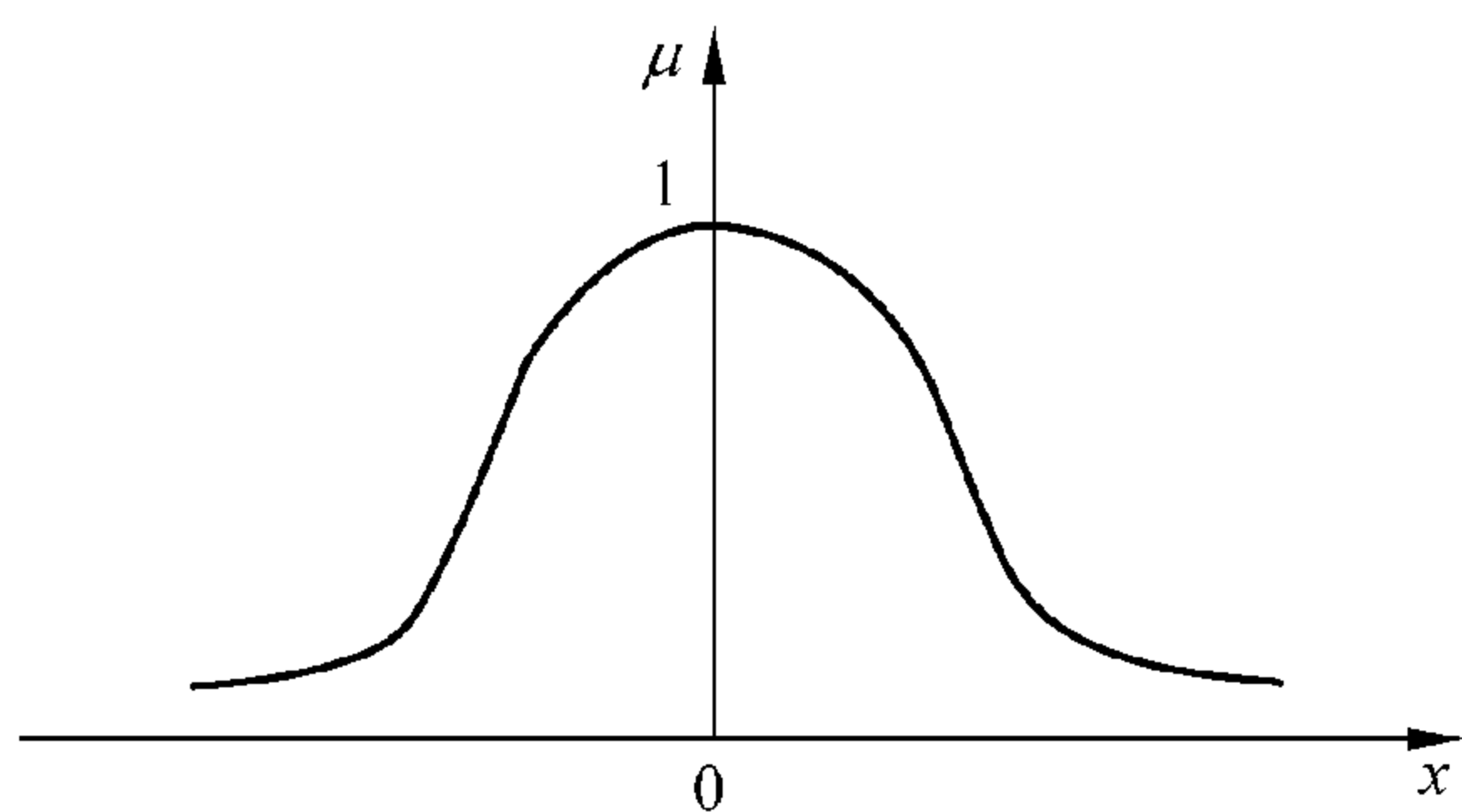


图 2.3 “接近于 0 的数”的隶属度函数的一种可能形式

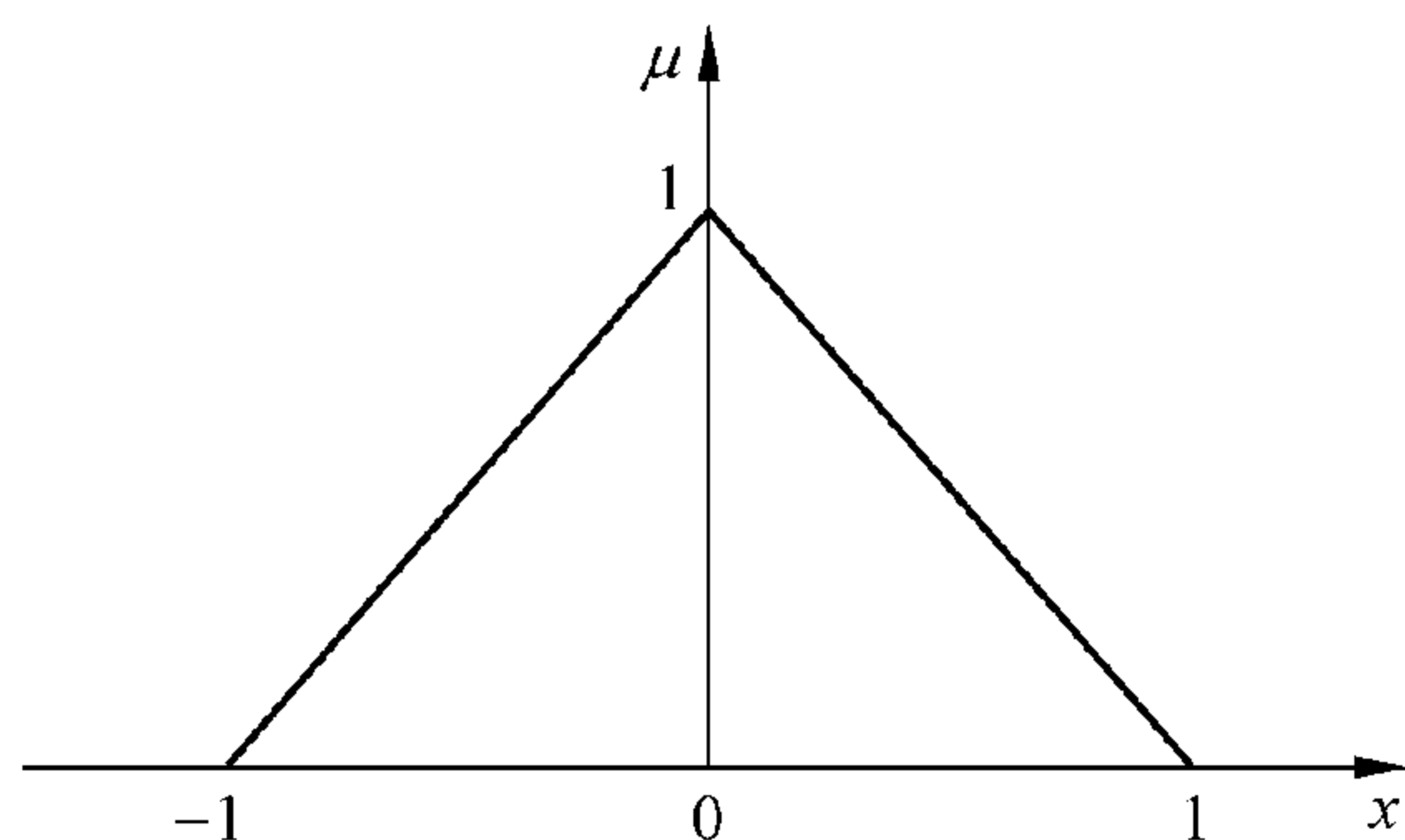
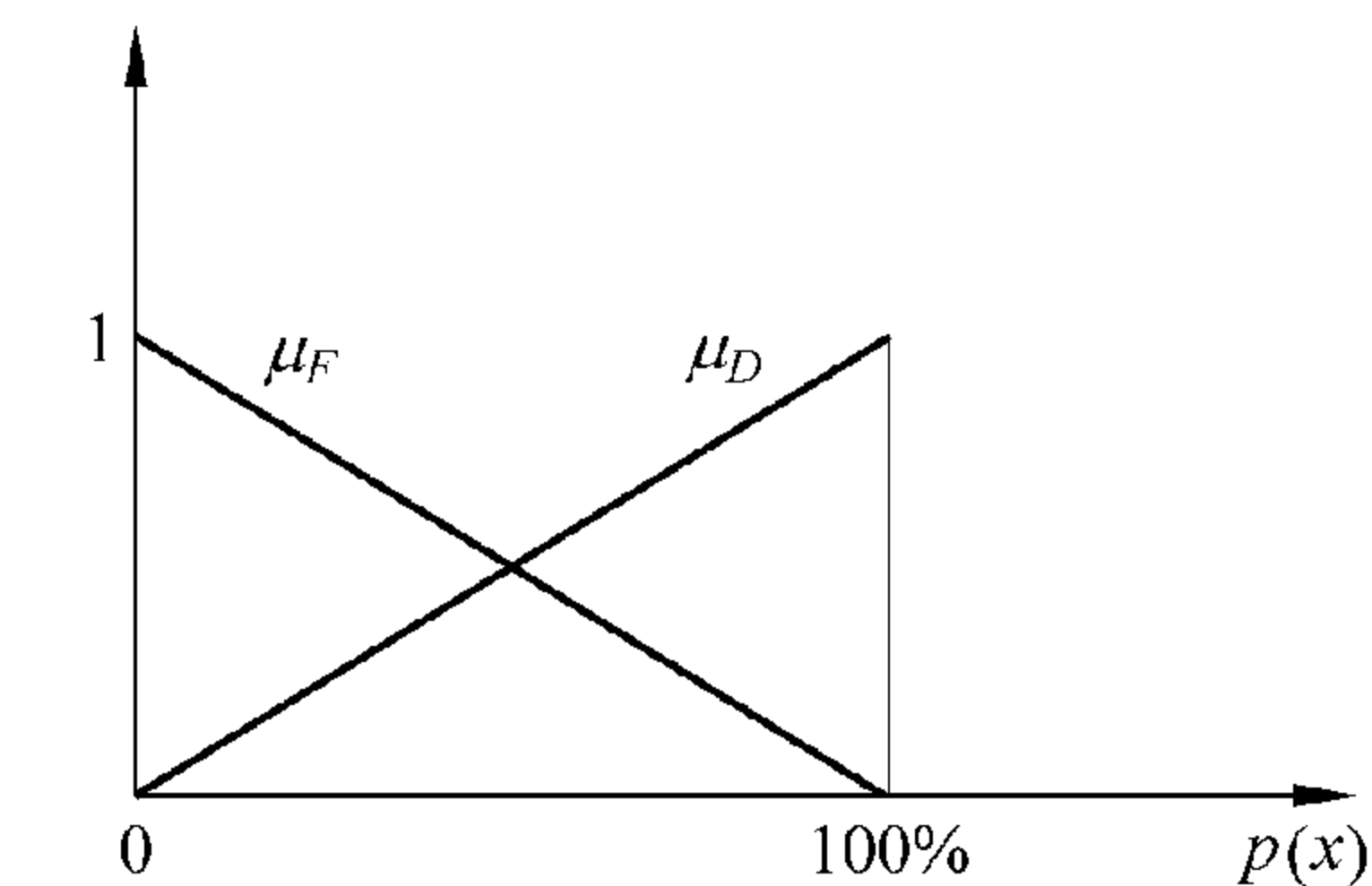


图 2.4 “接近于 0 的数”的隶属度函数另一种可能的形式



注: μ_D 和 μ_F 是根据零部件在美国制造的百分比 $[p(x)]$ 来确定的。

图 2.2 美国汽车的隶属度函数(μ_D)和非美国汽车的隶属度函数(μ_F)

函数式(2.10)或式(2.11)表征了“接近于0的数”,则一切都将不再模糊了。因此,用隶属度函数表征一个模糊描述后,实质上就将模糊描述的模糊消除了。一种常见的对模糊集合理论误解就是,认为模糊集合理论试图使世界模糊化。实际上恰恰相反,我们看到的是,模糊集合消除了世界的模糊。

- 紧跟前一结论的另一个重要问题是:怎样确定隶属度函数?隶属度函数有各种各样的选择,怎样从中进行选择呢?从概念上讲,有两种确定隶属度函数的方法。第一种方法是利用人类专家的知识,也就是说请该领域的专家来指定隶属度函数。由于模糊集合通常用于描述人类知识,所以,隶属度函数也就代表了部分人类知识。通常,这种方法仅能够给出隶属度函数的一个粗略的公式,还必须对其进行“微调”;第二种方法是从各种传感器中收集数据来确定隶属度函数。具体地讲,首先要指定隶属度函数的结构,然后根据数据对隶属度函数的参数进行“微调”。这两种方法,尤其是第二种,将会在后面的章节中详细讨论。
- 最后,应该强调的是,尽管式(2.10)和式(2.11)都用来描述“接近于0的数”,但它们是不同的模糊集合。因此,严格地讲,应采用不同的说明性短语表达模糊集合式(2.10)和式(2.11)。例如,用 $\mu_{z_1}(x)$ 表示式(2.10),用 $\mu_{z_2}(x)$ 表示式(2.11)。一个模糊集合与其隶属度函数应具有一一对应的关系。也就是说,当给出一个模糊集合时,必须有一个与之对应的独一无二的隶属度函数;反过来,当给出一个隶属度函数时,它也仅能表达一个模糊集合。在此意义上,模糊集合与其隶属度函数是等价的。

下面让我们再看两个模糊集合的例子,一个是连续域上的模糊集合,而另一个是离散域上的模糊集合。这两个集合都是从Zadeh的开创性文章(Zadeh[1965])中摘录的经典例子。

例 2.3 令 U 表示普通人的年龄区间 $[0, 100]$,那么就可以将模糊集合“年轻”和“年老”用下面隶属度函数(见图2.5)来定义(采用式(2.6)中的积分符号)

$$\text{年轻} = \int_0^{25} 1/x + \int_{25}^{100} \left(1 + \left(\frac{x-25}{5}\right)^2\right)^{-1} / x \quad (2.12)$$

$$\text{年老} = \int_{50}^{100} \left(1 + \left(\frac{x-50}{5}\right)^{-2}\right)^{-1} / x \quad (2.13)$$

例 2.4 令 U 表示从1到10的整数,即 $U = \{1, 2, \dots, 10\}$,则模糊集合“几个”可以定义为[用式(2.7)中的求和符号]

$$\text{几个} = 0.5/3 + 0.8/4 + 1/5 + 1/6 + 0.8/7 + 0.5/8 \quad (2.14)$$

即,5和6隶属于模糊集合“几个”的程度为1,4和7隶属于模糊集合“几个”的程度为0.8,3和8隶属于模糊集合“几个”的程度为0.5,1、2、9和10隶属于模糊集合“几个”的程度为0(见图2.6)。

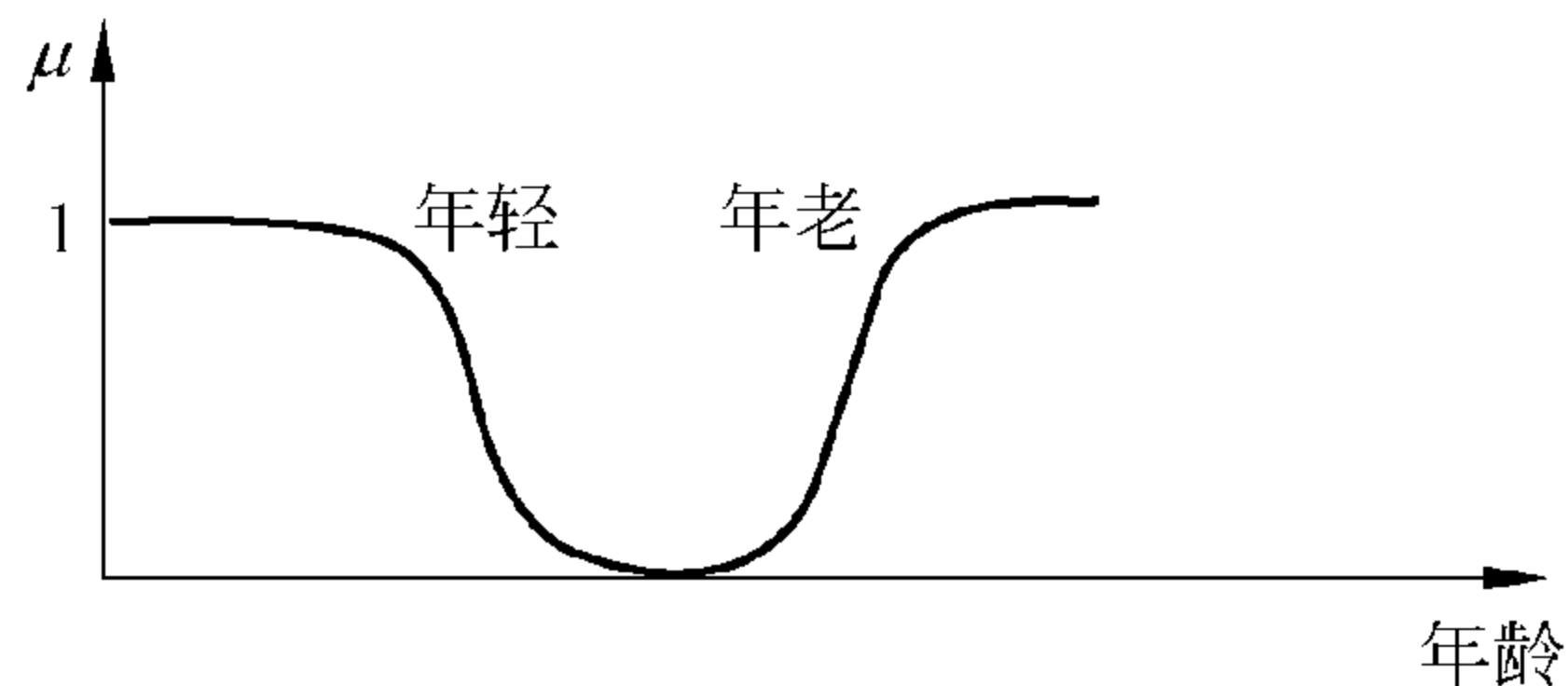


图 2.5 “年轻”和“年老”的图示

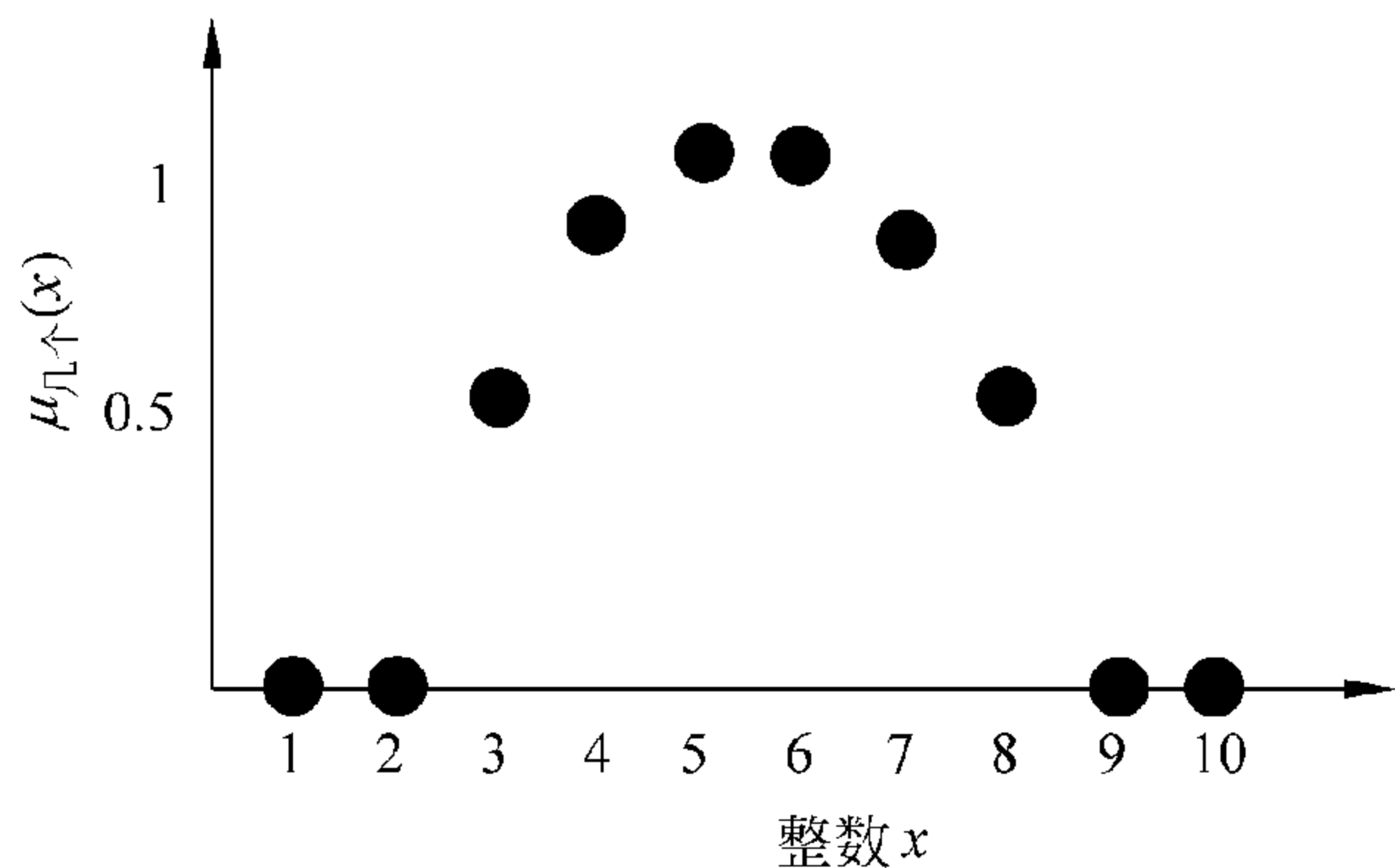


图 2.6 模糊集合“几个”的隶属度函数

2.2 模糊集合的一些基本概念

现在我们学习一些与模糊集合相关的基本概念和术语。模糊集合的许多概念和术语都是由经典(清晰)集合的基本概念推广来的,但有一些概念是模糊集合体系所独有的。

定义 2.2 支撑集(support)、模糊单值(fuzzy singleton)、中心(center)、交叉点(crossover point)、高度(height)、标准模糊集(normal fuzzy set)、 α -截集(α -cut)、凸模糊集(convex fuzzy set)及投影(projections)的概念定义如下:

论域 U 上模糊集 A 的支撑集是一个清晰集,它包含了 U 中所有在 A 上具有非零隶属度值的元素,即

$$\text{supp}(A) = \{x \in U \mid \mu_A(x) > 0\} \quad (2.15)$$

式中, $\text{supp}(A)$ 表示模糊集 A 的支撑集。例如,图 2.6 中的模糊集合“几个”的支撑集就是 $\{3, 4, 5, 6, 7, 8\}$ 这个整数集合。如果一个模糊集的支撑集是空的,则称该模糊集为空模糊集。

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。

模糊集的中心可定义如下,如果模糊集的隶属度函数达到其最大值的所有点的均值是有限值,则将该均值定义为模糊集的中心;如果该均值为正(负)无穷大,则将该模糊集的中心定义为所有达到最大隶属值的点中的最小(最大)点的值。图 2.7 中给出了一些典型模糊集的中心。一个模糊集的交叉点就是 U 中

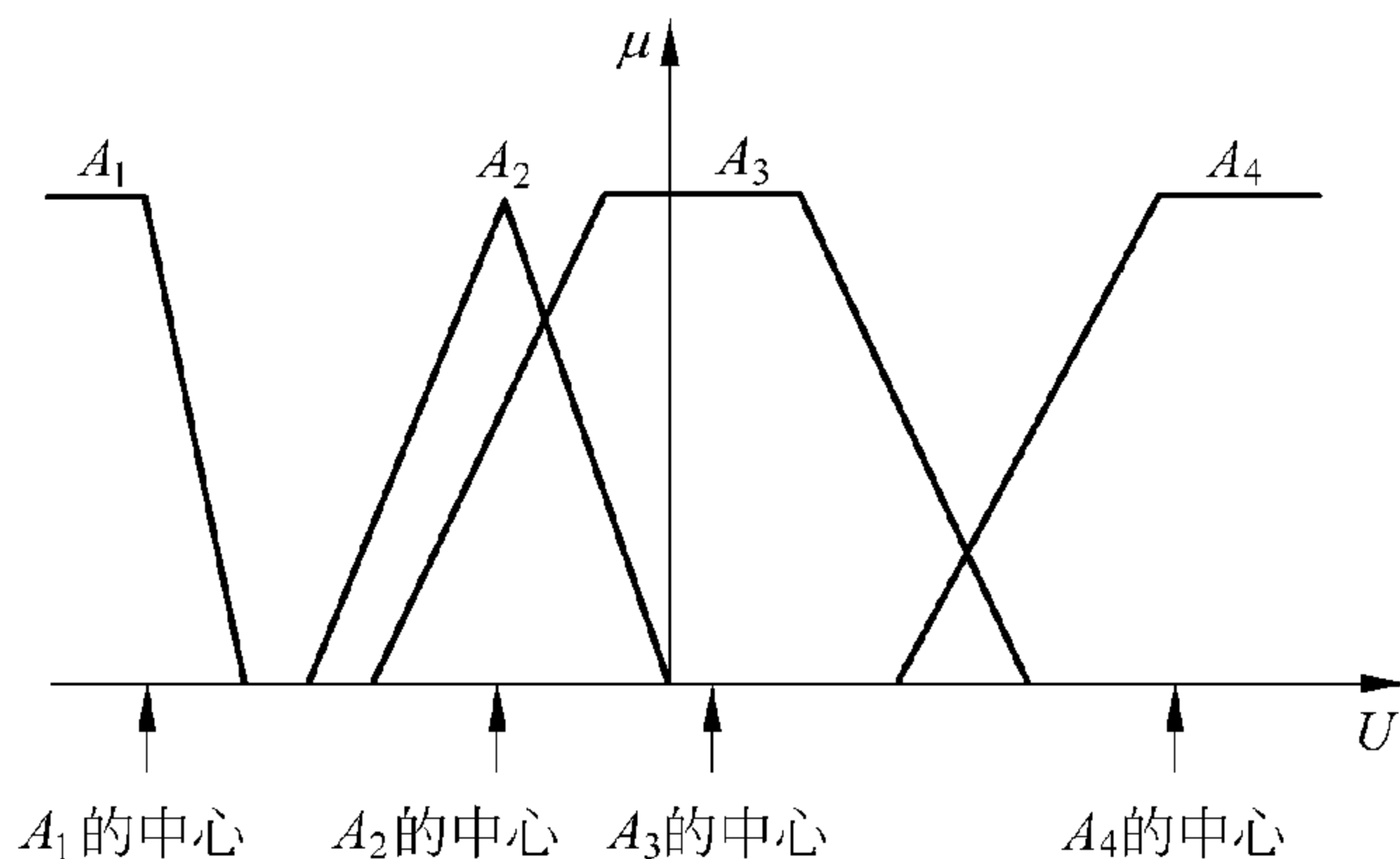


图 2.7 一些典型模糊集的中心

隶属于 A 的隶属度值等于 0.5 的点。

模糊集的高度,是指任意点所达到的最大隶属度值。例如,图 2.2 ~ 图 2.4 中所有模糊集的高度都等于 1。如果一个模糊集的高度等于 1,则称之为标准模糊集。因此图 2.2 ~ 图 2.4 中的所有模糊集都是标准模糊集。

一个模糊集 A 的 α -截集是一个清晰集 A_α ,它包含了 U 中所有隶属于 A 的隶属度值大于等于 α 的元素,即

$$A_\alpha = \{x \in U \mid \mu_A(x) \geq \alpha\} \quad (2.16)$$

例如,当 $\alpha = 0.3$ 时,模糊集(2.11)(见图 2.4)的 α -截集就是清晰集 $[-0.7, 0.7]$;而当 $\alpha = 0.9$ 时,模糊集(2.11)的 α -截集就是 $[-0.1, 0.1]$ 。

当论域 U 为 n 维欧氏空间 R^n 时,凸集的概念可以推广至模糊集合。对于任意 α ,当且仅当模糊集 A 在区间 $(0, 1]$ 上的 α -截集 A_α 为凸集时,模糊集 A 是凸模糊集。

引理 2.1 对任意 $x_1, x_2 \in R^n$ 和任意 $\lambda \in [0, 1]$,称 R^n 上的模糊集合 A 是凸模糊集,当且仅当下式成立

$$\mu_A[\lambda x_1 + (1 - \lambda)x_2] \geq \min[\mu_A(x_1), \mu_A(x_2)] \quad (2.17)$$

证明 首先,假设 A 是凸模糊集,证明式(2.17)是成立的。

令 x_1, x_2 为 R^n 上的任意点,为不失一般性,假设 $\mu_A(x_1) \leq \mu_A(x_2)$,因为 $\mu_A(x_1) = 0$ 时式(2.17)必定成立,所以令 $\mu_A(x_1) = \alpha > 0$ 。由 A 的 α -截集 A_α 是凸集的性质和 $x_1, x_2 \in A_\alpha$ (因为 $\mu_A(x_2) \geq \mu_A(x_1) = \alpha$),可得,对所有 $\lambda \in [0, 1]$ 有 $\lambda x_1 + (1 - \lambda)x_2 \in A_\alpha$ 。因此, $\mu_A[\lambda x_1 + (1 - \lambda)x_2] \geq \alpha = \mu_A(x_1) = \min[\mu_A(x_1), \mu_A(x_2)]$ 。

反过来,证明在式(2.17)成立的条件下, A 为凸模糊集。

令 α 为 $(0, 1]$ 上的任意点,如果 A_α 是空集,则 A 为凸模糊集(因为空集是凸集)。如果 A_α 是非空的,则存在 $x_1 \in R^n$,使得 $\mu_A(x_1) = \alpha$ (根据 A_α 的定义)。令 x_2 为 A_α 中的任一元素,则有 $\mu_A(x_2) \geq \alpha = \mu_A(x_1)$ 。因为根据假设,式(2.17)是成立的,所以对任意 $\lambda \in [0, 1]$ 有 $\mu_A[\lambda x_1 + (1 - \lambda)x_2] \geq \min[\mu_A(x_1), \mu_A(x_2)] = \mu_A(x_1) = \alpha$,这表明 $\lambda x_1 + (1 - \lambda)x_2 \in A_\alpha$,所以, A_α 是凸集。因为 α 是 $(0, 1]$ 上的任意点,所以由 A_α 是凸集可知 A 为凸模糊集。

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 $H = \{x \in R^n \mid x_1 = 0\}$ (为简化起见,这里只考虑了这个特殊的超平面,由它可直接推广到一般的超平面)。定义 A 在 H 上的投影为在 R^{n-1} 上的模糊集合 A_H ,其隶属度函数为

$$\mu_{A_H}(x_2, \dots, x_n) = \sup_{x_1 \in R} \mu_A(x_1, \dots, x_n) \quad (2.18)$$

式中, $\sup_{x_1 \in R} \mu_A(x_1, \dots, x_n)$ 表示,当 x_1 在 R 中取值时,函数 $\mu_A(x_1, \dots, x_n)$ 的最大值。

2.3 模糊集合的运算

在 2.1 节和 2.2 节中学习的基本概念仅与单一的模糊集合有关。本节中,我们将学习多个模糊集合的基本运算。以下均假设 A 和 B 是定义在同一论域 U 上的模糊集合。

定义 2.3 两个模糊集合 A 和 B 的等价 (equality)、包含 (containment)、补集 (complement)、并集 (union) 和交集 (intersection) 的定义如下:

对任意 $x \in U$, 当且仅当 $\mu_A(x) = \mu_B(x)$ 时, 称 A 和 B 是等价的。对任意 $x \in U$, 当且仅当 $\mu_A(x) \leq \mu_B(x)$ 时, 称 B 包含 A , 记为 $A \subset B$ 。定义集合 A 的补集为 U 上的模糊集合, 记为 \bar{A} , 其隶属度函数为

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \quad (2.19)$$

U 上模糊集 A 和 B 的并集也是模糊集, 记为 $A \cup B$, 其隶属度函数为

$$\mu_{A \cup B}(x) = \max[\mu_A(x), \mu_B(x)] \quad (2.20)$$

U 上模糊集 A 和 B 的交集也是模糊集, 记为 $A \cap B$, 其隶属度函数为

$$\mu_{A \cap B}(x) = \min[\mu_A(x), \mu_B(x)] \quad (2.21)$$

读者也许会思考, 为什么用“max”表示并集, 用“min”表示交集。现在我们给出一种直观的解释。定义并集的一种直观的方式是, A 和 B 的并集是包含 A 和 B 的“最小”的模糊集合。更准确地说, 如果 C 是任意一个包含 A 和 B 的模糊集, 则它也包含 A 和 B 的并集。现证明这一直观的定义与式(2.20)等价。首先由 $\max[\mu_A, \mu_B] \geq \mu_A$ 和 $\max[\mu_A, \mu_B] \geq \mu_B$ 可知, 式(2.20)中定义的 $A \cup B$ 包含了 A 和 B 。而且, 若 C 是任意一个包含 A 和 B 的模糊集, 则有 $\mu_C \geq \mu_A, \mu_C \geq \mu_B$ 。从而有 $\mu_C \geq \max[\mu_A, \mu_B] = \mu_{A \cup B}$ 。这样就证明了式(2.20)中定义的 $A \cup B$ 是包含 A 和 B 的“最小”模糊集。同理可证, 式(2.21)中定义的交集是 A 和 B 所包含的“最大”模糊集。

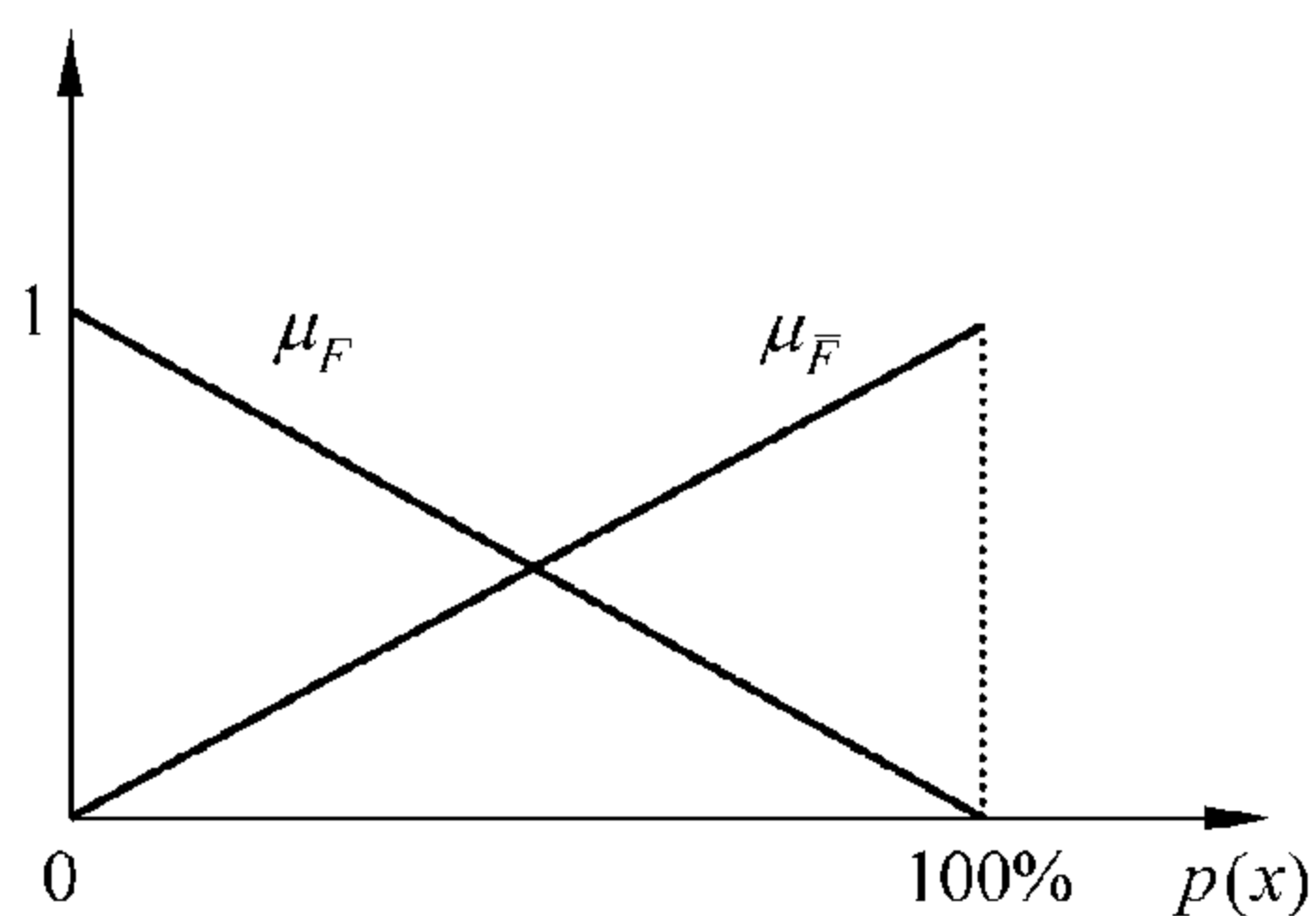


图 2.8 \bar{F} 和 F 的隶属度函数

例 2.5 考虑式(2.8)和式(2.9)所定义的两个模糊集合 D 和 F (见图 2.2), 定义 F 的补集 \bar{F} 的隶属度函数(见图 2.8)为

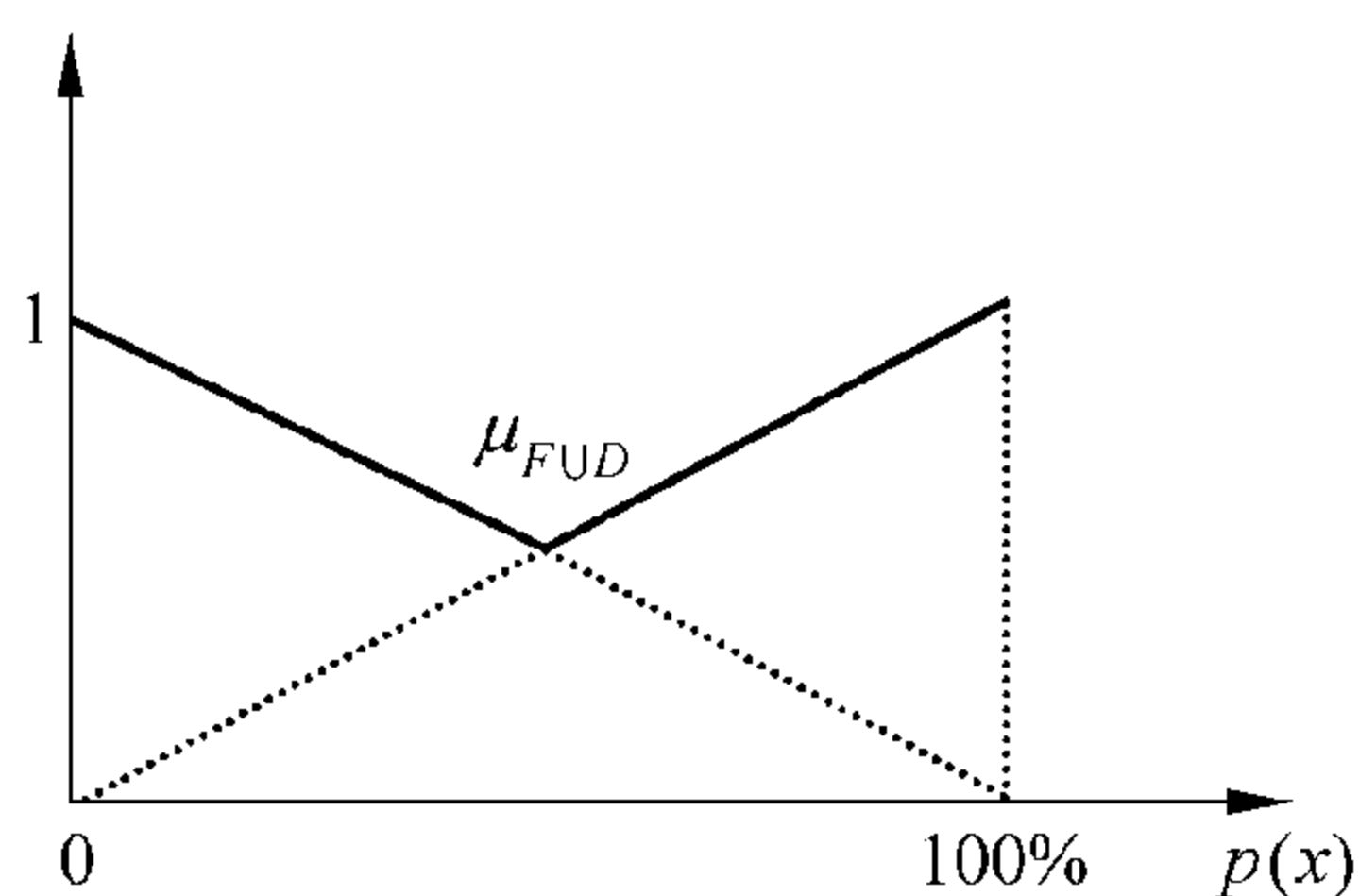
$$\mu_{\bar{F}}(x) = 1 - \mu_F(x) = 1 - p(x) \quad (2.22)$$

比较式(2.22)与式(2.9), 可以看出, $\bar{F} = D$ 。这说明, 如果一辆汽车不是非美国汽车 (F 的补集) 就是美国汽车。或者更准确地说, 一辆汽车越不是非美国汽车, 就越是美国汽车。 F 和 D 的并集 $F \cup D$ (见图 2.9) 可定义为

$$\mu_{F \cup D}(x) = \max[\mu_F, \mu_D] = \begin{cases} \mu_F(x) & 0 \leq p(x) \leq 0.5 \\ \mu_D(x) & 0.5 \leq p(x) \leq 1 \end{cases} \quad (2.23)$$

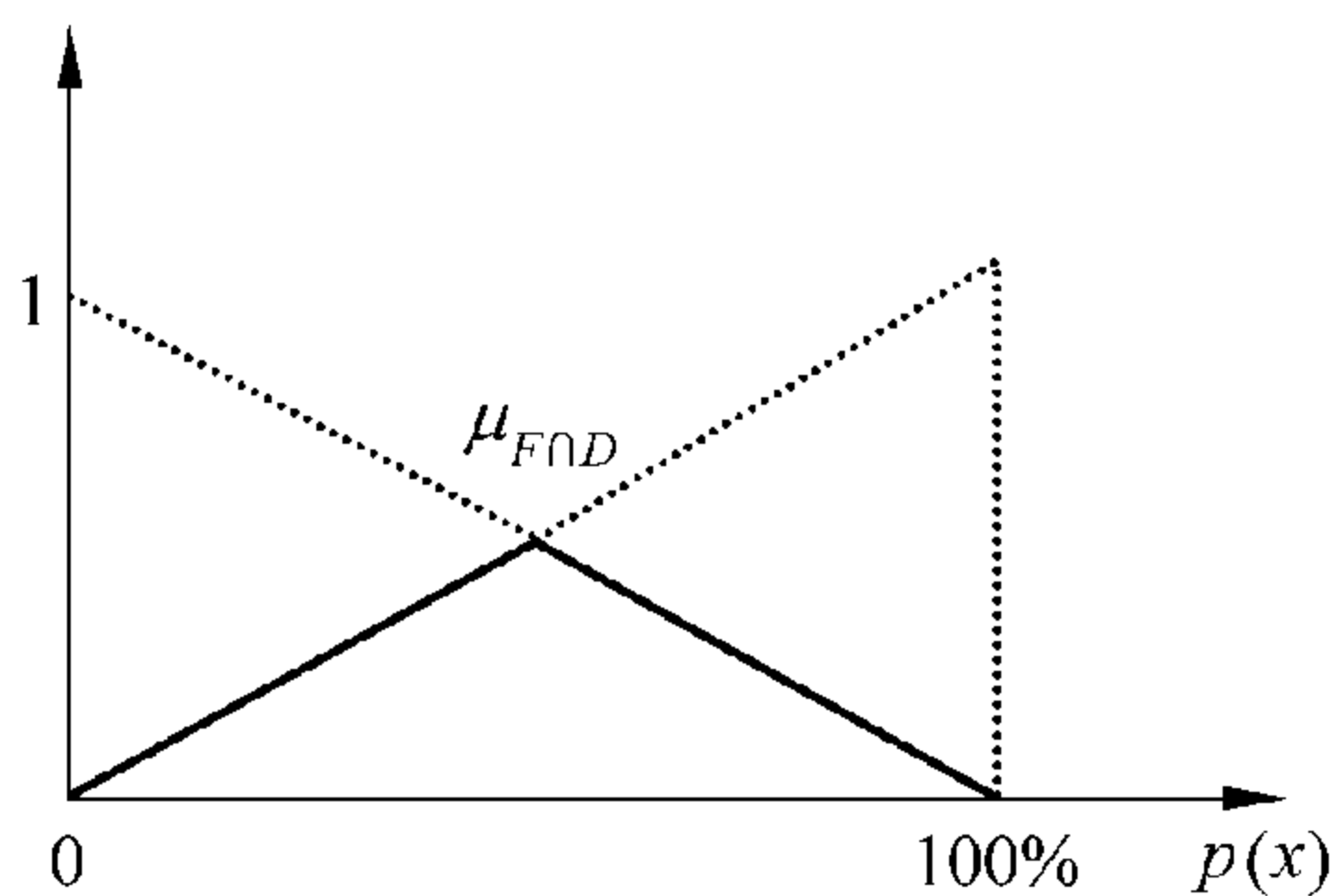
F 和 D 的交集 $F \cap D$ (见图 2.10) 可定义为

$$\mu_{F \cap D}(x) = \min[\mu_F, \mu_D] = \begin{cases} \mu_D(x) & 0 \leq p(x) \leq 0.5 \\ \mu_F(x) & 0.5 \leq p(x) \leq 1 \end{cases} \quad (2.24)$$



注：这里 F 和 D 的定义参考图 2.2。

图 2.9 $F \cup D$ 的隶属度函数



注：这里 F 和 D 的定义参考图 2.2。

图 2.10 $F \cap D$ 的隶属度函数

对于式(2.19)、式(2.20)和式(2.21)中所定义的补、并、交运算来说,许多在经典集合中成立的基本性质(并不是全部)是可以扩展到模糊集合中来的,以下面的引理为例。

引理 2.2 德·摩根定律(The De Morgan's Laws)对于模糊集合也是成立的。即,假设 A 和 B 是模糊集合,则有

$$\overline{A \cup B} = \overline{A} \cap \overline{B} \quad (2.25)$$

$$\overline{A \cap B} = \overline{A} \cup \overline{B} \quad (2.26)$$

证明 仅证明式(2.25)成立。式(2.26)可采用相同的方法证明,其证明留作习题。

首先,证明下面的等式是成立的:

$$1 - \max[\mu_A, \mu_B] = \min[1 - \mu_A, 1 - \mu_B] \quad (2.27)$$

要证明上式成立,须考虑两种可能情况: $\mu_A \geq \mu_B$ 和 $\mu_A < \mu_B$ 。若 $\mu_A \geq \mu_B$, 则有 $1 - \mu_A \leq 1 - \mu_B$, 进而有 $1 - \max[\mu_A, \mu_B] = 1 - \mu_A = \min[1 - \mu_A, 1 - \mu_B]$, 即式(2.27)成立; 若 $\mu_A < \mu_B$, 则有 $1 - \mu_A > 1 - \mu_B$, 进而有 $1 - \max[\mu_A, \mu_B] = 1 - \mu_B = \min[1 - \mu_A, 1 - \mu_B]$, 即式(2.27)成立。因此,式(2.27)是成立的。由定义式(2.19)~式(2.21)及两个模糊集等价的定义可知,式(2.27)成立和式(2.25)成立是等价的。

2.4 本章小结及导读

本章主要阐述了以下内容:

- 模糊集的定义,与模糊集有关的一些基本概念(支撑集、 α -截集、凸集等)及模糊集的

基本运算(补、并、交运算等);

- 隶属度函数的直观含义以及如何直观地确定某个特定模糊描述的隶属度函数;
- 模糊集的几个特殊例子,模糊集及其运算的基本性质的证明。

Zadeh 原始的文献 Zadeh[1965] 仍然是学习模糊集合及其相关概念的最佳材料,这是一篇值得阅读的好文章。与模糊集合相关的基本运算和概念在 Zadeh[1965] 中也有介绍。

2.5 习题

习题 2.1 确定“矮人”、“高人”、“胖人”的合理的隶属度函数。

习题 2.2 将下面的表达式用模糊集合表达:

(a) 努力学习的学生; (b) 顶尖的学生; (c) 聪明的学生。

习题 2.3 考虑一下定义在区间 $U = [0, 10]$ 上的模糊集合 F, G 和 H , 其隶属度函数为

$$\mu_F(x) = \frac{x}{x+2}, \mu_G(x) = 2^{-x}, \mu_H(x) = \frac{1}{1+10(x-2)^2} \quad (2.28)$$

确定下面每个模糊集合的隶属度函数的数学公式和隶属度函数的图形:

(a) $\overline{F}, \overline{G}, \overline{H}$

(b) $F \cup G, F \cup H, G \cup H$

(c) $F \cap G, F \cap H, G \cap H$

(d) $F \cup G \cup H, F \cap G \cap H$

(e) $F \cap \overline{H}, \overline{G \cap H}, \overline{F \cup H}$

习题 2.4 确定习题 2.3 中的模糊集合 F, G 和 H 在 (a) $\alpha = 0.2$, (b) $\alpha = 0.5$, (c) $\alpha = 0.9$, (d) $\alpha = 1$ 时的 α -截集。

习题 2.5 令模糊集合 A 为定义在封闭平面 $U = [-1, 1] \times [-3, 3]$ 上的模糊集, 其隶属度函数为

$$\mu_A(x_1, x_2) = e^{-(x_1^2 + x_2^2)} \quad (2.29)$$

分别确定 A 在超平面 $H_1 = \{x \in U \mid x_1 = 0\}$ 和 $H_2 = \{x \in U \mid x_2 = 0\}$ 上的投影。

习题 2.6 证明排中律(the excluded middle)对于模糊集合是不成立的, 即当 F 是模糊集时, $F \cup \overline{F} = U$ 不成立。

习题 2.7 证明引理 2.2 中的等式(2.26)是成立的。

习题 2.8 证明两个凸模糊集的交集是凸模糊集。请回答, 两个凸模糊集的并集是不是凸模糊集?

模糊集合的其他运算

在第 2 章中,我们学习了以下基本算子(operators),模糊集的补集、并集和交集:

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \quad (3.1)$$

$$\mu_{A \cup B}(x) = \max[\mu_A(x), \mu_B(x)] \quad (3.2)$$

$$\mu_{A \cap B}(x) = \min[\mu_A(x), \mu_B(x)] \quad (3.3)$$

前面曾提到过,式(3.2)中所定义的模糊集合 $A \cup B$ 是包含 A 和 B 的最小模糊集合,式(3.3)中所定义的模糊集合 $A \cap B$ 是 A 和 B 所包含的最大模糊集合。所以,式(3.1)~式(3.3)所定义的只是模糊集合的一种算子,还有可能存在其他的算子。例如,可以将 $A \cup B$ 定义为任意一个包含 A 和 B 的模糊集(并不一定是最小的模糊集)。本章将研究关于模糊集的补集、并集和交集的其他类型的算子。

为什么需要其他类型的算子呢?主要原因在于,在某些条件下,算子式(3.1)~式(3.3)也许并不令人满意。例如,当取两个模糊集的交集时,可能希望较大的模糊集对结果产生影响,但如果模糊交集选用式(3.3)中的最小(min)算子,则可能较大的模糊集是无法产生影响的。另一个原因在于,从理论上研究何种类型的算子对模糊集合可行是很有意义的。大家知道,对于非模糊集来说,只有一种并集、补集和交集算子是可行的,而对于模糊集来说,可能还有其他类型的算子可行,那么这些算子是什么类型呢?这些新算子有什么性质呢?这些都是本章将要回答的问题。

新算子是基于公理提出来的。为使运算合理,本章将从几个补集、并集和交集应满足的公理出发,列举一些满足这些公理的特定公式。

3.1 模糊补

令映射 $c: [0,1] \rightarrow [0,1]$ 表示由模糊集 A 的隶属度函数向其补集的隶属度函数转换的映射,即

$$c[\mu_A(x)] = \mu_{\bar{A}}(x) \quad (3.4)$$

由式(3.1),有 $c[\mu_A(x)] = 1 - \mu_A(x)$ 。为使函数 c 适合于计算模糊补的隶属度函数(等价于模糊补),它必须至少满足以下两个必要条件:

公理 **c1** $c(0) = 1, c(1) = 0$ (有界性)。

公理 **c2** 当 $a, b \in [0, 1]$ 时, 如果 $a < b$, 则有 $c(a) \geq c(b)$ (非增性)。

这里(在本章中) a, b 表示某些个模糊集合的隶属度函数, 即, $a = \mu_A(x), b = \mu_B(x)$ 。

公理 **c1** 说明, 如果一个元素隶属于一个模糊集合的隶属度为 0 (或 1), 则它隶属于这个模糊集的补集的隶属度就为 0 (或 1)。公理 **c2** 说明, 一个模糊集合的隶属度值的上升必将导致其补集的隶属度值的下降或不变。显然, 违背其中任何一个必要条件的函数都不能用作模糊补算子。

定义 **3.1** 任意满足公理 **c1** 和 **c2** 的函数 $c: [0, 1] \rightarrow [0, 1]$, 都叫做模糊补集。

Sugeno 模糊补(Sugeno[1977])可定义为

$$c_\lambda(a) = \frac{1-a}{1+\lambda a} \quad (3.5)$$

它是一类模糊补。其中, $\lambda \in (-1, \infty)$, 对于参数 λ 的每个取值, 都可以得到一个特定的模糊补, 很容易证明式(3.5)是满足公理 **c1** 和 **c2** 的。图 3.1 给出了不同 λ 值下的 Sugeno 模糊补。请注意, 当 $\lambda = 0$ 时, Sugeno 模糊补就变成了基本模糊补(3.1)。

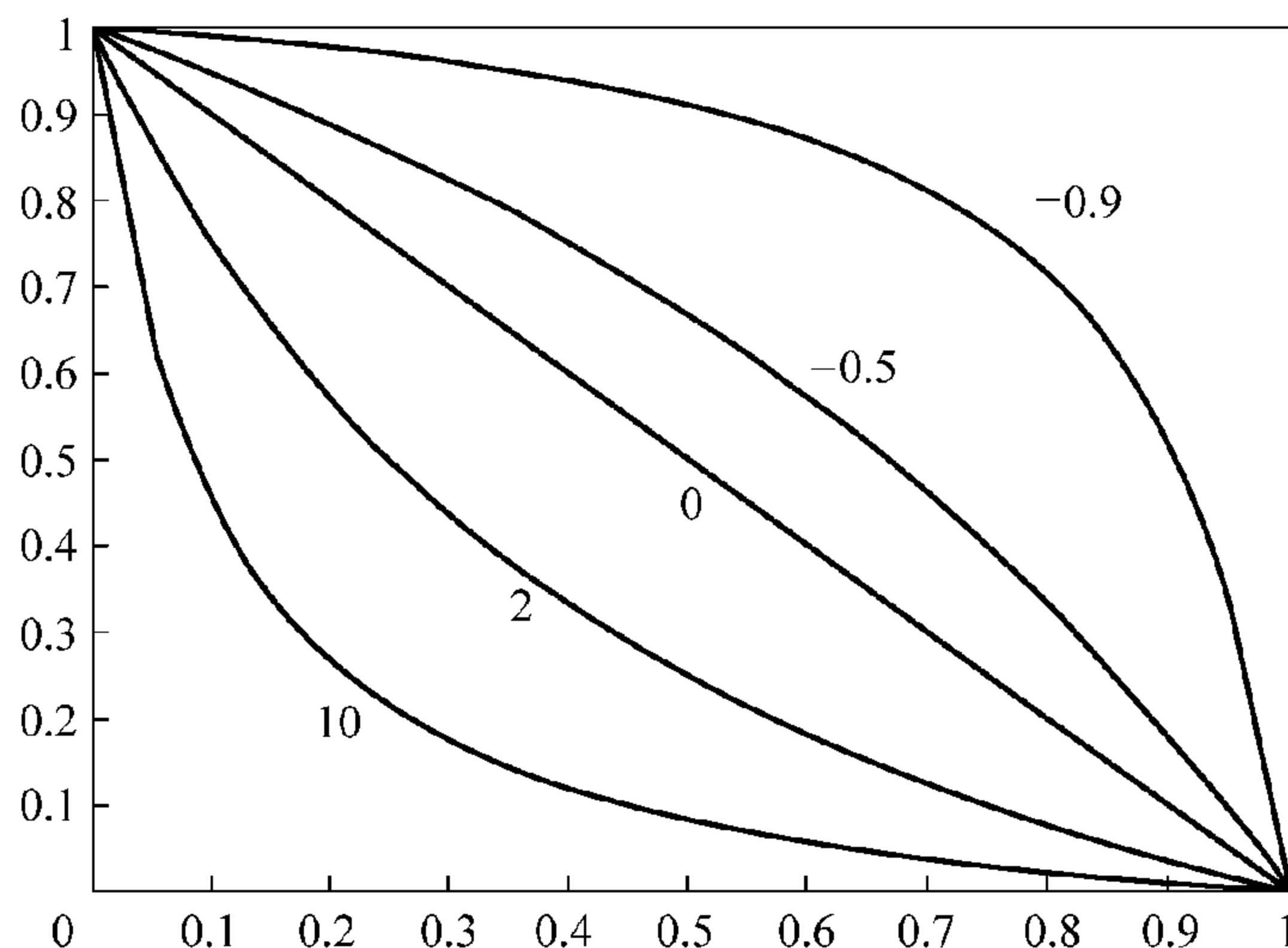
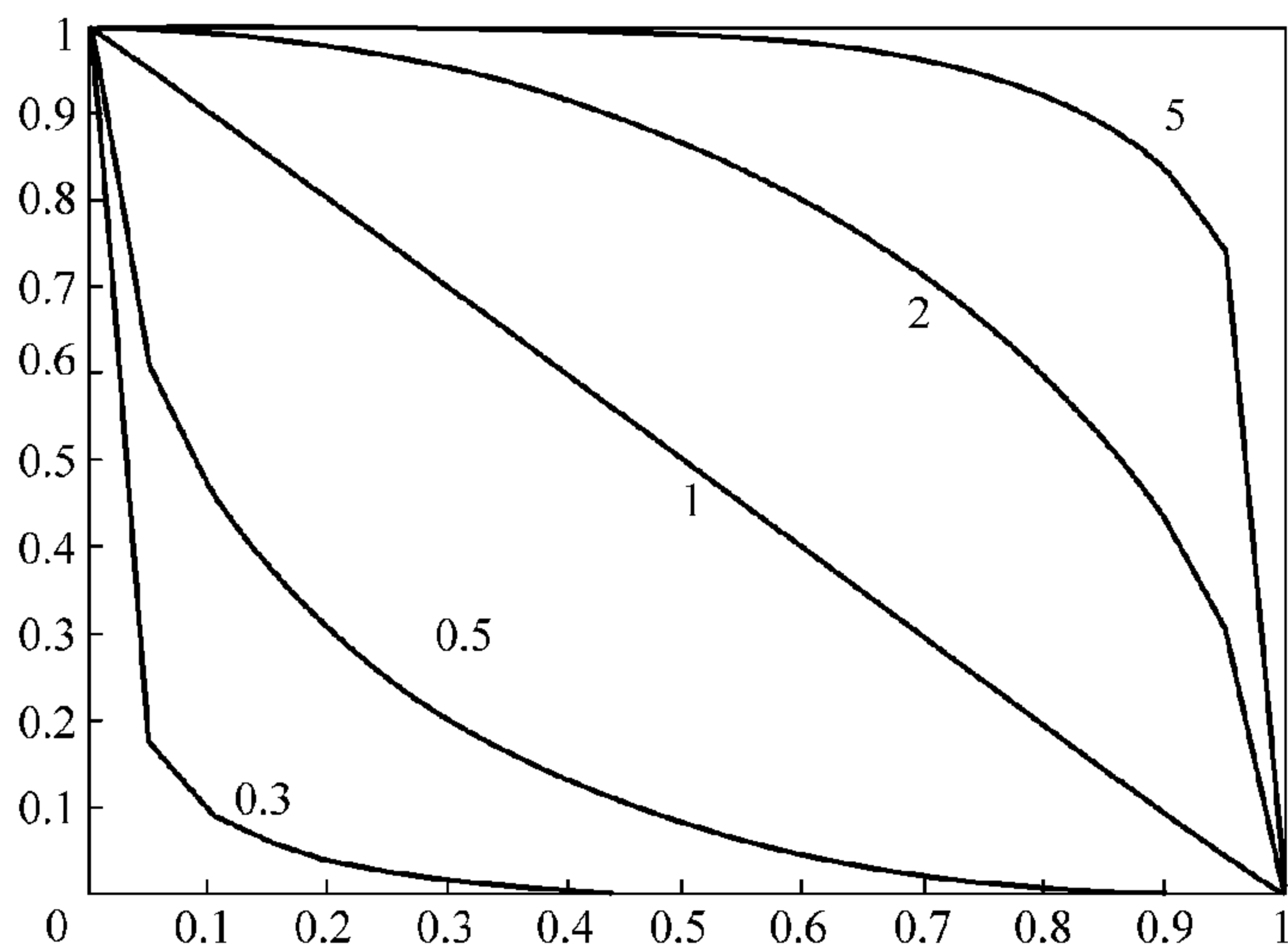


图 3.1 不同 λ 值下的 Sugeno 模糊补 $c_\lambda(a)$

Yager 模糊补(Yager[1980])可定义为

$$c_\omega(a) = (1 - a^\omega)^{1/\omega} \quad (3.6)$$

它是另一类模糊补, 其中, $\omega \in (0, \infty)$, 对于每个 ω 值, 都可以得到一个特定的模糊补。很容易证明式(3.6)是满足公理 **c1** 和 **c2** 的。图 3.2 给出了不同的 ω 值下的 Yager 模糊补。当 $\omega = 1$ 时, 式(3.6)就变成了式(3.1)。

图 3.2 不同 ω 值下的 Yager 模糊补 $c_\omega(a)$

3.2 模糊并—— s -范数

令 $s: [0,1] \times [0,1] \rightarrow [0,1]$, 表示由模糊集 A 和 B 的隶属度函数到 A 和 B 的并集的隶属度函数的映射, 即

$$s[\mu_A(x), \mu_B(x)] = \mu_{A \cup B}(x) \quad (3.7)$$

根据式(3.2)可知, $s[\mu_A(x), \mu_B(x)] = \max[\mu_A(x), \mu_B(x)]$ 。为使函数 s 适合于计算模糊并的隶属度函数, 它必须至少满足以下四个必要条件:

公理 **s1** $s(1,1) = 1, s(0,a) = s(a,0) = a$ (有界性)

公理 **s2** $s(a,b) = s(b,a)$ (交换性)

公理 **s3** 如果 $a \leq a'$ 且 $b \leq b'$, 则 $s(a,b) \leq s(a',b')$ (非减性)

公理 **s4** $s(s(a,b),c) = s(a,s(b,c))$ (结合性)

公理 **s1** 是模糊并集函数在边界处的特性; 公理 **s2** 保证运算结果与模糊集的顺序无关; 公理 **s3** 给出了模糊并的通用必要条件: 两个模糊集合的隶属度值的上升会导致这两个模糊集的并集的隶属度值的上升; 公理 **s4** 则把模糊并运算扩展至两个模糊集合以上。

定义 3.2 任意一个满足公理 **s1** ~ 公理 **s4** 的函数 $s: [0,1] \times [0,1] \rightarrow [0,1]$ 都叫做 s -范数 (s -Norm)。

很容易证明式(3.2)中的基本模糊并 \max 是一个 s -范数, 现给出其他三种 s -范数:

- Dombi 的 s -范数 (Dombi [1982]):

$$s_\lambda(a, b) = \frac{1}{1 + \left[\left(\frac{1}{a} - 1 \right)^{-\lambda} + \left(\frac{1}{b} - 1 \right)^{-\lambda} \right]^{-1/\lambda}} \quad (3.8)$$

其中, 参数 $\lambda \in (0, \infty)$ 。

- Dubois-Prade 的 s -范数(Dubois and Prade[1980]):

$$s_\alpha(a, b) = \frac{a + b - ab - \min(a, b, 1 - \alpha)}{\max(1 - a, 1 - b, \alpha)} \quad (3.9)$$

其中, 参数 $\alpha \in [0, 1]$ 。

- Yager 的 s -范数(Yager[1980]):

$$s_w(a, b) = \min[1, (a^w + b^w)^{1/w}] \quad (3.10)$$

其中, 参数 $w \in (0, \infty)$ 。

只要参数选定, 式(3.8) ~ 式(3.10)都定义了一个特定的 s -范数。可以直接证明式(3.8) ~ 式(3.10)是满足公理 s_1 ~ 公理 s_4 的。通过从不同角度推广经典集合的并集运算, 就可以得到这些 s -范数。

下面的文献中还提出了许多其他的 s -范数, 列于此处仅供参考:

$$\bullet \text{ 直和: } s_{ds}(a, b) = \begin{cases} a & b = 0 \\ b & a = 0 \\ 1 & \text{其他} \end{cases} \quad (3.11)$$

$$\bullet \text{ 爱因斯坦和: } s_{es}(a, b) = \frac{a + b}{1 + ab} \quad (3.12)$$

$$\bullet \text{ 代数和: } s_{as}(a, b) = a + b - ab \quad (3.13)$$

- 最大: 式(3.2)

为什么在文献中会提出这么多的 s -范数呢? 理论上的原因是因为, 当隶属度值严格为 0 或 1 时, 这些 s -范数是完全一样的。也就是说, 它们全部都是非模糊并集的扩展。实践上的原因是因为, 一些 s -范数在某些应用中可能比其他 s -范数更有效。

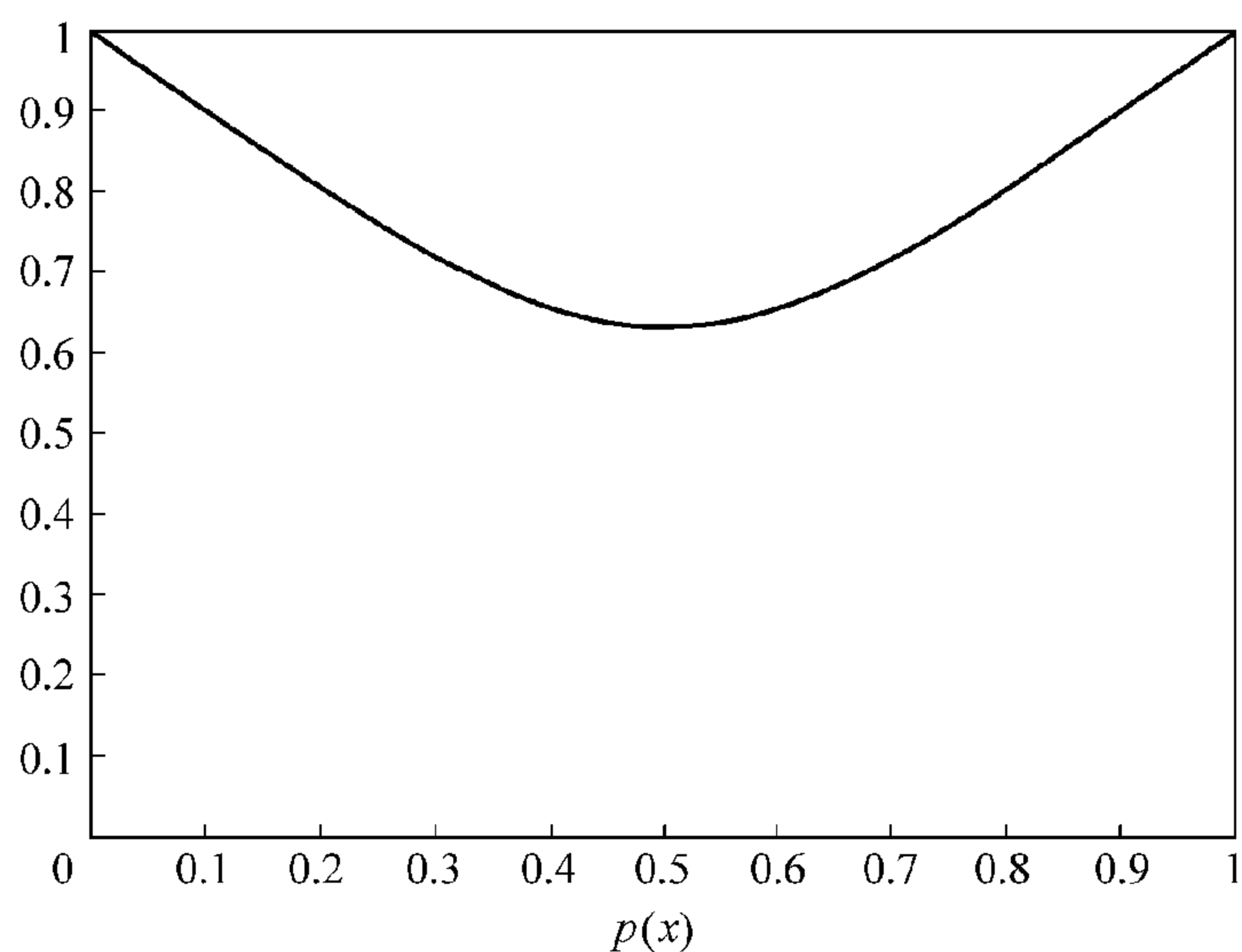
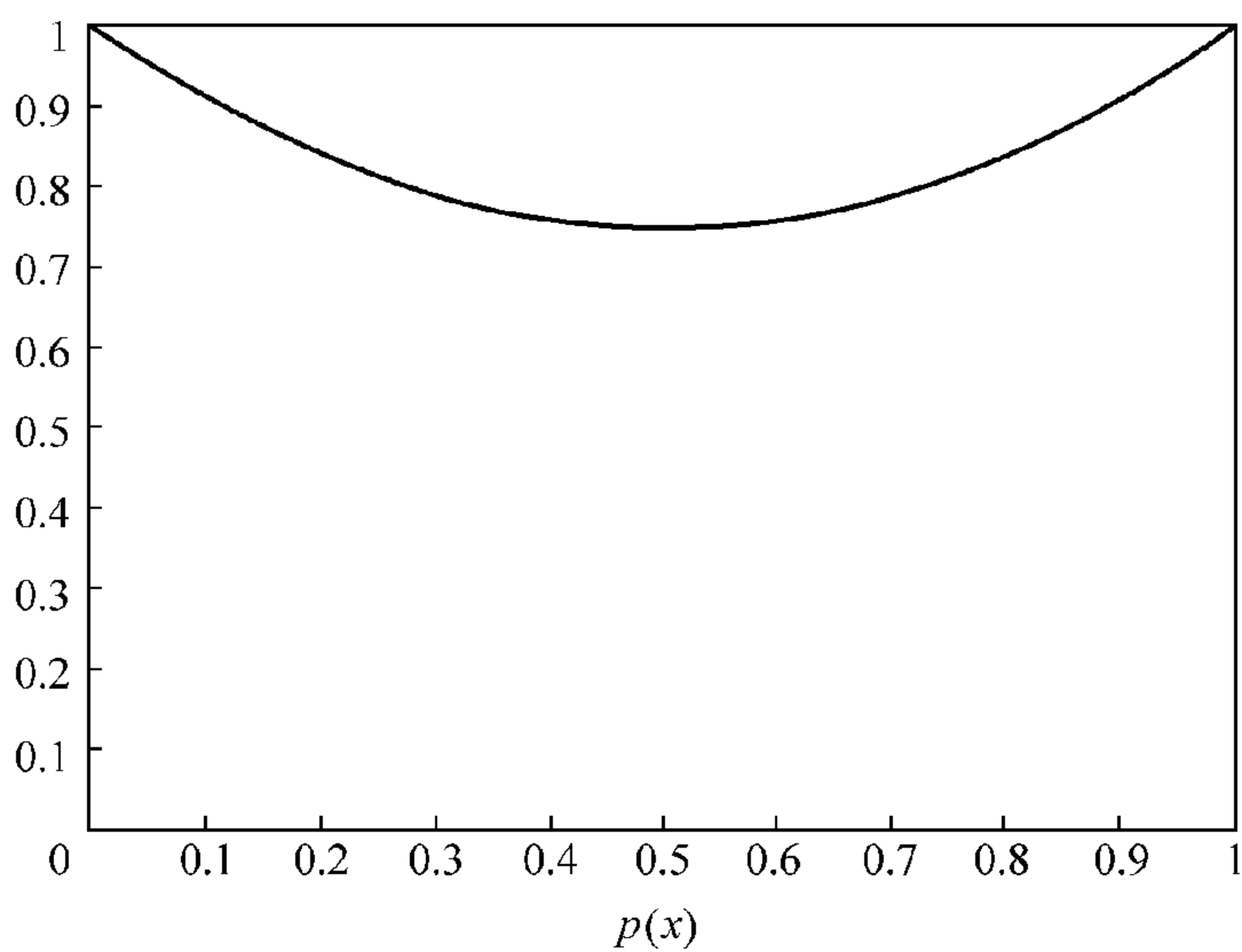
例 3.1 考虑第 2 章中的例 2.1 所定义的模糊集合 D 和 F (分别见式(2.8)和式(2.9))。如果用 Yager 的 s -范数(3.10)计算模糊并的隶属度函数, 则模糊集 $D \cup F$ 的隶属度函数为

$$\mu_{D \cup F}(x) = s_w[\mu_D(x), \mu_F(x)] = \min[1, ((p(x))^w + (1 - p(x))^w)^{1/w}] \quad (3.14)$$

图 3.3 给出了 $w = 3$ 时 $\mu_{D \cup F}(x)$ 的图形。如果用代数和(3.13)计算模糊并的隶属度函数, 则模糊集 $D \cup F$ 的隶属度函数为

$$\begin{aligned} \mu_{D \cup F}(x) &= s_{as}[\mu_D(x), \mu_F(x)] = p(x) + (1 - p(x)) - p(x)(1 - p(x)) \\ &= 1 - p(x) + (p(x))^2 \end{aligned} \quad (3.15)$$

将图 3.3 和图 3.4 与图 2.9 比较, 可以看出, Yager 的 s -范数和代数和 s -范数都大于最大算子。总之, 可以证明, 最大值算子(3.2)是最小的 s -范数, 直和(3.11)是最大的 s -范数。

图 3.3 用 Yager s -范数(3.10)($w=3$)计算 $D \cup F$ 的隶属度函数图 3.4 用代数和(3.13)计算 $D \cup F$ 的隶属度函数

定理 3.1 对任意一个 s -范数,即对任意一个满足公理 $s_1 \sim$ 公理 s_4 的函数 $s: [0,1] \times [0,1] \rightarrow [0,1]$ 来说,当 $a, b \in [0,1]$ 时,下面的不等式是成立的:

$$\max(a, b) \leq s(a, b) \leq s_{ds}(a, b) \quad (3.16)$$

证明 先证明 $\max(a, b) \leq s(a, b)$ 。由公理 s_3 的非减性条件和公理 s_1 的有界性条件,可得

$$s(a, b) \geq s(a, 0) = a \quad (3.17)$$

再由公理 s_2 的交换性条件,可得

$$s(a, b) = s(b, a) \geq s(b, 0) = b \quad (3.18)$$

合并式(3.17)和式(3.18),可得 $s(a, b) \geq \max(a, b)$ 。

下一步证明 $s(a, b) \leq s_{ds}(a, b)$ 。

如果 $b = 0$, 则由公理 s1, 有 $s(a, b) = s(a, 0) = a$, 进而有 $s(a, b) = s_{ds}(a, b)$ 。由公理 s2 的交换性条件, 如果 $a = 0$, 有 $s(a, b) = s_{ds}(a, b)$; 如果 $a \neq 0$ 且 $b \neq 0$, 则有

$$s_{ds}(a, b) = 1 \geq s(a, b) \quad (3.19)$$

因此, 对于所有 $a, b \in [0, 1]$, 都有 $s(a, b) \leq s_{ds}(a, b)$ 。

最后, 证明 Dombi 的 s -范数 $s_\lambda(a, b)$ [见式(3.8)] 的一条有趣的性质: 在 λ 趋于无穷小时, $s_\lambda(a, b)$ 收敛于基本的模糊并集 $\max(a, b)$, 而在 λ 趋于零时, $s_\lambda(a, b)$ 收敛于直和 $s_{ds}(a, b)$ 。因此, Dombi 的 s -范数覆盖了 s -范数的整个空间。

引理 3.1 令 $s_\lambda(a, b)$ 为式(3.8)所定义的 s -范数, $s_{ds}(a, b)$ 为式(3.11)所定义的 s -范数, 则有

$$\lim_{\lambda \rightarrow \infty} s_\lambda(a, b) = \max(a, b) \quad (3.20)$$

$$\lim_{\lambda \rightarrow 0} s_\lambda(a, b) = s_{ds}(a, b) \quad (3.21)$$

证明 首先证明式(3.20)是成立的。如果 $a = b \neq 0$, 则由式(3.8)可得

$$\lim_{\lambda \rightarrow \infty} s_\lambda(a, b) = \lim_{\lambda \rightarrow \infty} \left[1 / \left(1 + 2^{-1/\lambda} \left(\frac{1}{a} - 1 \right) \right) \right] = a = \max(a, b)$$

如果 $a = b = 0$, 则有

$$\lim_{\lambda \rightarrow \infty} s_\lambda(a, b) = \lim_{\lambda \rightarrow \infty} 1 / (1 + 0^{-1/\lambda}) = 0 = \max(a, b)$$

如果 $a \neq b$, 并不失一般性地假设 $a < b$ (由公理 s2), 令

$$z = \left[\left(\frac{1}{a} - 1 \right)^{-\lambda} + \left(\frac{1}{b} - 1 \right)^{-\lambda} \right]^{-1/\lambda}$$

则根据罗必塔法则, 可得

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \ln(z) &= \lim_{\lambda \rightarrow \infty} - \frac{\ln \left[\left(\frac{1}{a} - 1 \right)^{-\lambda} + \left(\frac{1}{b} - 1 \right)^{-\lambda} \right]}{\lambda} \\ &= \lim_{\lambda \rightarrow \infty} \frac{\left(\frac{1}{a} - 1 \right)^{-\lambda} \ln \left(\frac{1}{a} - 1 \right) + \left(\frac{1}{b} - 1 \right)^{-\lambda} \ln \left(\frac{1}{b} - 1 \right)}{\left(\frac{1}{a} - 1 \right)^{-\lambda} + \left(\frac{1}{b} - 1 \right)^{-\lambda}} \\ &= \lim_{\lambda \rightarrow \infty} \frac{\left[\left(\frac{1}{a} - 1 \right) / \left(\frac{1}{b} - 1 \right) \right]^{-\lambda} \ln \left(\frac{1}{a} - 1 \right) + \ln \left(\frac{1}{b} - 1 \right)}{\left[\left(\frac{1}{a} - 1 \right) / \left(\frac{1}{b} - 1 \right) \right]^{-\lambda} + 1} \\ &= \ln \left(\frac{1}{b} - 1 \right) \end{aligned} \quad (3.22)$$

由此可得

$$\lim_{\lambda \rightarrow \infty} z = \frac{1}{b} - 1$$

$$\lim_{\lambda \rightarrow \infty} s_{\lambda}(a, b) = \lim_{\lambda \rightarrow \infty} \frac{1}{1+z} = b = \max(a, b) \quad (3.23)$$

然后,证明式(3.21)是成立的。若 $a=0, b \neq 0$, 则有

$$s_{\lambda}(a, b) = 1 / \left[1 + \left(\frac{1}{b} - 1 \right)^{-\lambda \frac{-1}{\lambda}} \right] = b = s_{ds}(a, b)$$

若 $b=0, a \neq 0$, 则根据交换性, 有

$$s_{\lambda}(a, b) = a = s_{ds}(a, b)$$

若 $a \neq 0, b \neq 0$, 则有

$$\lim_{\lambda \rightarrow 0} s_{\lambda}(a, b) = \lim_{\lambda \rightarrow 0} 1 / [1 + 2^{-1/\lambda}] = 1 = s_{ds}(a, b)$$

最后,若 $a=b=0$, 则有

$$\lim_{\lambda \rightarrow 0} s_{\lambda}(a, b) = \lim_{\lambda \rightarrow 0} 1 / [1 + 0^{-1/\lambda}] = 0 = s_{ds}(a, b)$$

同理可证,当 w 趋于无穷大时, Yager 的 s -范数(3.10)收敛于基本模糊并集 $\max(a, b)$; 当 w 趋于 0 时, Yager 的 s -范数(3.10)收敛于直和 $s_{ds}(a, b)$, 其证明留作习题。

3.3 模糊交—— t -范数

令映射 $t: [0, 1] \times [0, 1] \rightarrow [0, 1]$, 表示由模糊集 A 和 B 的隶属度函数向 A 和 B 的交集的隶属度函数转换的一个函数, 即

$$t[\mu_A(x), \mu_B(x)] = \mu_{A \cap B}(x) \quad (3.24)$$

根据式(3.3)可知

$$t[\mu_A(x), \mu_B(x)] = \min[\mu_A(x), \mu_B(x)]$$

为使函数 t 适合于计算模糊交的隶属度函数, 它至少应满足以下的四个必要条件:

公理 **t1** $t(0, 0) = 0, t(a, 1) = t(1, a) = a$ (有界性)

公理 **t2** $t(a, b) = t(b, a)$ (交换性)

公理 **t3** 如果 $a \leq a'$ 且 $b \leq b'$, 则 $t(a, b) \leq t(a', b')$ (非减性)

公理 **t4** $t[t(a, b), c] = t[a, t(b, c)]$ (结合性)

上述公理的证明和公理 $s1 \sim$ 公理 $s4$ 的证明完全相同。

定义 3.3 任意一个满足公理 $t1 \sim$ 公理 $t4$ 的函数 $t: [0, 1] \times [0, 1] \rightarrow [0, 1]$ 都叫做 t -范数。

可以证明, 式(3.3)中的基本模糊交算子 \min 是 t -范数。对于任意一个 t -范数, 一定有一个与之相关的 s -范数, 反之亦然。因此, 与 Dombi 的 s -范数, Dubois-Prade 的 s -范数和 Yager 的 s -范数(见式(3.8)~式(3.10))相对应的 t -范数有 Dombi 的 t -范数, Dubois-Prade

的 t -范数和 Yager 的 t -范数, 它们的定义如下:

- Dombi 的 t -范数 (Dombi [1982])

$$t_\lambda(a, b) = \frac{1}{1 + \left[\left(\frac{1}{a} - 1 \right)^\lambda + \left(\frac{1}{b} - 1 \right)^\lambda \right]^{1/\lambda}} \quad (3.25)$$

其中, $\lambda \in (0, \infty)$ 。

- Dubois-Prade 的 t -范数 (Dubois and Prade [1980])

$$t_\alpha(a, b) = \frac{ab}{\max(a, b, \alpha)} \quad (3.26)$$

其中, $\alpha \in [0, 1]$ 。

- Yager 的 t -范数 (Yager [1980]):

$$t_w(a, b) = 1 - \min[1, ((1-a)^w + (1-b)^w)^{1/w}] \quad (3.27)$$

其中, $w \in (0, \infty)$ 。

只要参数选定, 式(3.25) ~ 式(3.27)都可以定义一个特定的 t -范数。可以证明, 式(3.25) ~ 式(3.27)是满足公理 t1 ~ 公理 t4 的。同式(3.11) ~ 式(3.13)及式(3.2)中的 s -范数相对应的 t -范数有以下几种:

- 直积: $t_{dp}(a, b) = \begin{cases} a & b = 1 \\ b & a = 1 \\ 0 & \text{其他} \end{cases} \quad (3.28)$

- 爱因斯坦积: $t_{ep}(a, b) = \frac{ab}{2 - (a + b - ab)} \quad (3.29)$

- 代数积: $t_{ap}(a, b) = ab \quad (3.30)$

- 最小: 式(3.3)

例 3.2 考虑第 2 章中例 2.1 所定义的模糊集合 D 和 F , 如果用 Yager 的 t -范数(3.27)来计算模糊交集的隶属度函数, 则 $D \cap F$ 的隶属度函数为

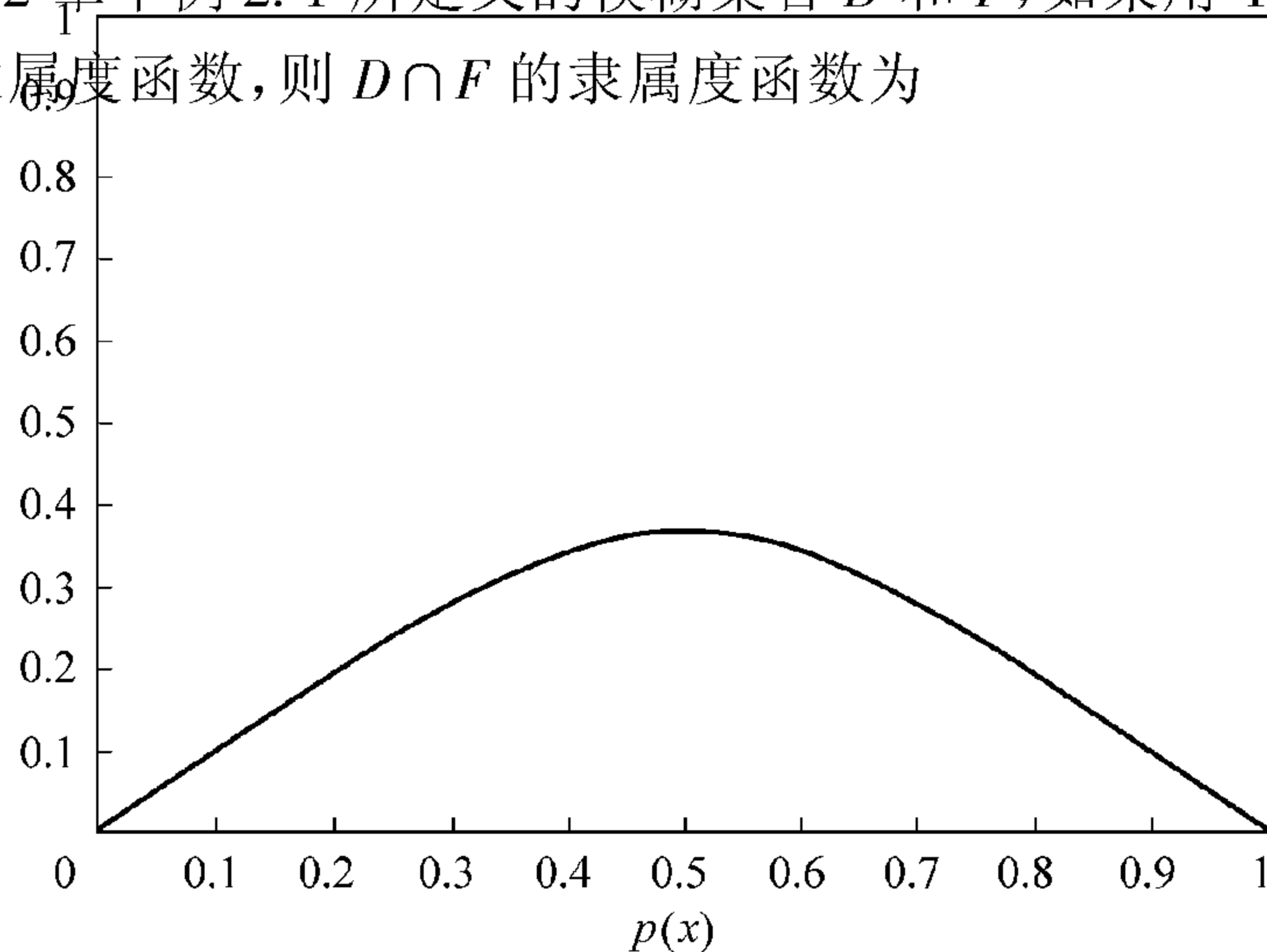


图 3.5 用 Yager t -范数(3.27)($w=3$)计算 $D \cap F$ 的隶属度函数

$$\mu_{D \cap F}(x) = t_w[\mu_D(x), \mu_F(x)] = 1 - \min[1, ((1 - p(x))^w + (p(x))^w)^{1/w}] \quad (3.31)$$

图 3.5 给出了 $w = 3$ 时 $\mu_{D \cap F}(x)$ 的图形。如果用代数积(3.30)来计算模糊交集,则模糊集 $D \cap F$ 的隶属度函数(见图 3.6)为

$$\mu_{D \cap F}(x) = t_{ap}[\mu_D(x), \mu_F(x)] = p(x)(1 - p(x)) \quad (3.32)$$

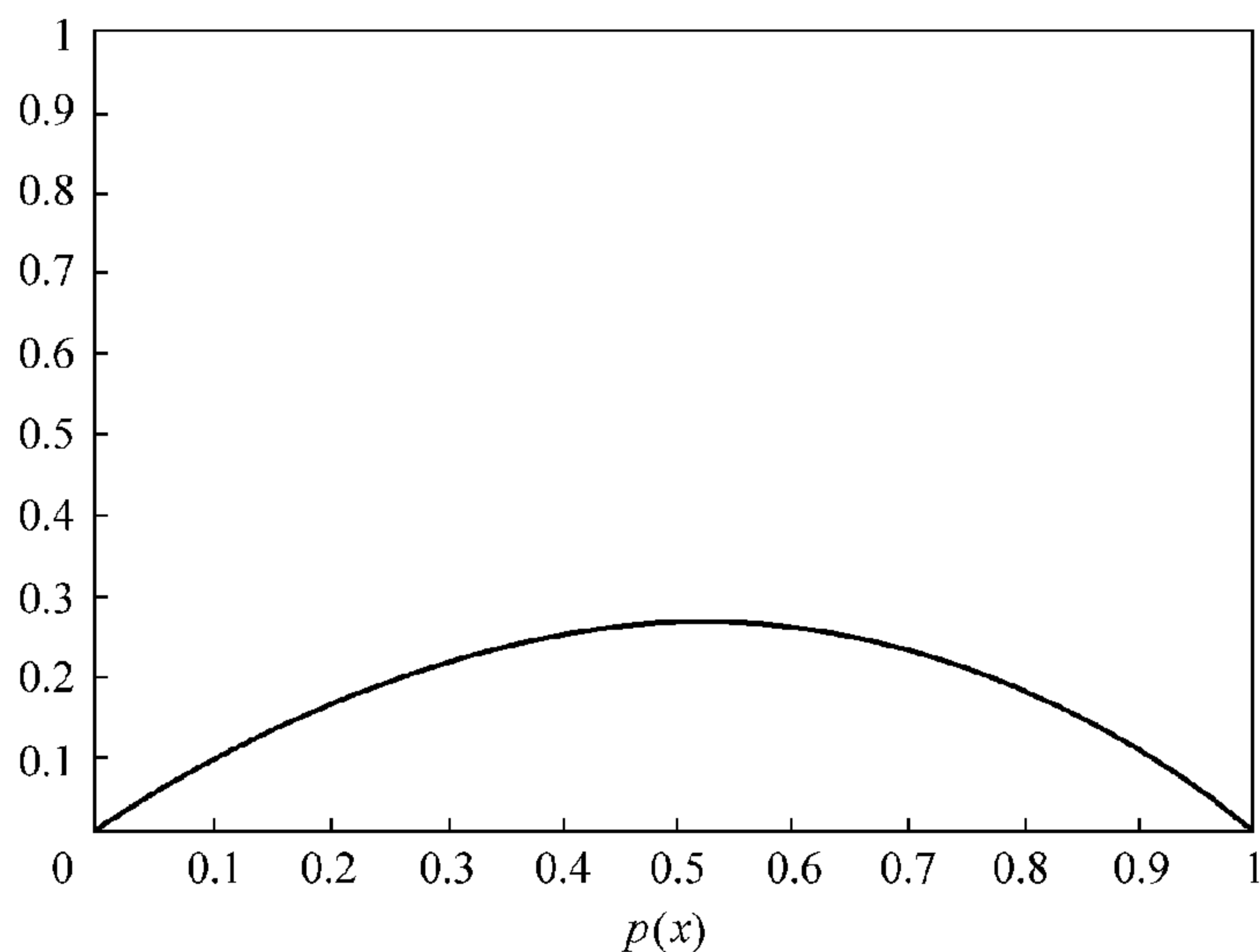


图 3.6 用代数积(3.30)计算 $D \cap F$ 的隶属度函数

将图 3.5 和图 3.6 与图 2.10 比较,可以看出, Yager 的 t -范数和代数积比最小(\min)算子小。总之,可以证明最小算子是最大的 t -范数,而直积是最小的 t -范数。

定理 3.2 对任意一个 t -范数,即对于任意一个满足公理 t1 ~ 公理 t4 的函数 $t: [0, 1] \times [0, 1] \rightarrow [0, 1]$ 来说,当 $a, b \in [0, 1]$ 时,下面的不等式是成立的:

$$t_{dp}(a, b) \leq t(a, b) \leq \min(a, b) \quad (3.33)$$

该定理的证明和定理 3.1 很相似,留作习题。与引理 3.1 类似,也可证明,当 λ 趋于无穷大时,式(3.25)中的 Dombi 的 t -范数 $t_\lambda(a, b)$ 收敛于基本模糊交 $\min(a, b)$; 当 λ 趋于 0 时,式(3.25)中的 Dombi 的 t -范数 $t_\lambda(a, b)$ 收敛于直积 $t_{dp}(a, b)$ 。所以, Dombi 的 t -范数覆盖了 t -范数的整个空间。

引理 3.2 令 $t_\lambda(a, b)$ 为式(3.25)所定义的 $t_\lambda(a, b)$, 则有

$$\lim_{\lambda \rightarrow \infty} t_\lambda(a, b) = \min(a, b) \quad (3.34)$$

$$\lim_{\lambda \rightarrow 0} t_\lambda(a, b) = t_{dp}(a, b) \quad (3.35)$$

该引理的证明和引理 3.1 的证明类似。

将式(3.8) ~ 式(3.13)分别和式(3.25) ~ 式(3.30)比较,可以看出,每个 s -范数都有一个与之相对应的 t -范数,不过“相对应”的含义是什么呢? 它是指,存在一个模糊补,使得模糊补、并和交三个模糊集都满足德·摩根定律。具体地讲,就是认为 s -范数 $s(a, b)$ 和 t -范数 $t(a, b)$ 以及模糊补集 $c(a)$, 在下式成立的条件下形成了关联组(associated class):

$$c[s(a, b)] = t[c(a), c(b)] \quad (3.36)$$

例 3.3 式(3.10)中 Yager 的 s -范数 $s_w(a, b)$, 式(3.27)中 Yager 的 t -范数 $t_w(a, b)$ 和基本模糊补(3.1)可以形成一组关联组。下面, 证明该判断是成立的。由式(3.1)和式(3.10), 可得

$$c[s_w(a, b)] = 1 - \min[1, (a^w + b^w)^{1/w}] \quad (3.37)$$

其中, $c(a)$ 表示基本模糊补(3.1)。再由式(3.1)和式(3.27), 可得

$$t_w[c(a), c(b)] = 1 - \min[1, ((1 - 1 + a)^w + (1 - 1 + b)^w)^{1/w}] \quad (3.38)$$

最后, 由式(3.37)和式(3.38)即可得到式(3.36)。

例 3.4 代数和(3.13)和代数积(3.30)及基本模糊补(3.1)可以组成一个关联组。下面证明该判断是成立的。由式(3.1)和式(3.12), 可得

$$c[s_{as}(a, b)] = 1 - a - b + ab \quad (3.39)$$

再由式(3.1)和式(3.30), 可得

$$t_{ap}[c(a), c(b)] = (1 - a)(1 - b) = 1 - a - b + ab \quad (3.40)$$

所以, 该关联组满足德·摩根定律(3.36)。

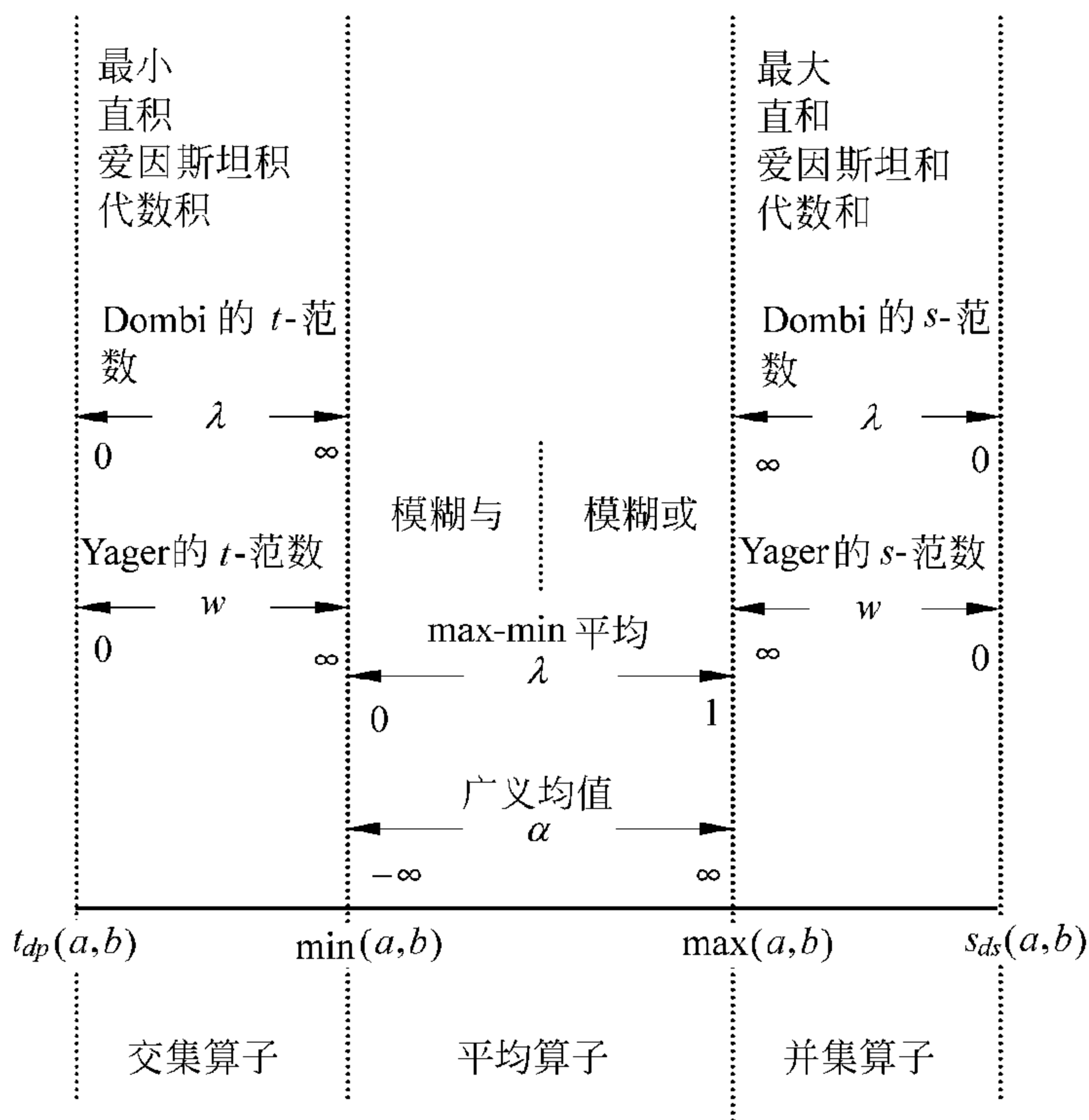


图 3.7 模糊集合算子的全集空间

3.4 平均算子

对于任意模糊集 A 和 B 的隶属度值 $a = \mu_A(x)$ 和 $b = \mu_B(x)$, 由定理 3.1 可知, $A \cup B$ (由任一 s -范数定义) 的隶属度值在区间 $[\max(a, b), s_{ds}(a, b)]$ 上。同样, 由定理 3.2 可知, 交集 $A \cap B$ (由任一 t -范数定义) 的隶属度值在区间 $[t_{dp}(a, b), \min(a, b)]$ 上 (见图 3.7)。因此, 并集算子和交集算子并不能覆盖 $\min(a, b)$ 和 $\max(a, b)$ 之间的区间, 这里将覆盖区间 $[\min(a, b), \max(a, b)]$ 的算子叫做平均算子。与 s -范数和 t -范数类似, 平均算子也是一个由 $[0, 1] \times [0, 1]$ 到 $[0, 1]$ 的函数, 记作 v 。

文献中提出过许多平均算子, 这里给出四个定义:

- 最大—最小平均(max-min averages)

$$v_\lambda(a, b) = \lambda \max(a, b) + (1 - \lambda) \min(a, b) \quad (3.41)$$

其中, $\lambda \in [0, 1]$ 。

- 广义均值(Generalized means)

$$v_\alpha(a, b) = \left(\frac{a^\alpha + b^\alpha}{2} \right)^{1/\alpha} \quad (3.42)$$

其中, $\alpha \in R (\alpha \neq 0)$ 。

- “模糊与”(Fuzzy and)

$$v_p(a, b) = p \min(a, b) + \frac{(1-p)(a+b)}{2} \quad (3.43)$$

其中, $p \in [0, 1]$ 。

- “模糊或”(Fuzzy or)

$$v_\gamma(a, b) = \gamma \max(a, b) + \frac{(1-\gamma)(a+b)}{2} \quad (3.44)$$

其中, $\gamma \in [0, 1]$ 。

显然, 当参数 λ 由 0 变至 1 时, max-min 均值覆盖了整个 $[\min(a, b), \max(a, b)]$ 区间; “模糊与”覆盖了由 $\min(a, b)$ 至 $(a+b)/2$ 的整个闭区间; “模糊或”覆盖了从 $(a+b)/2$ 至 $\max(a, b)$ 的整个闭区间。也可以证明, 当 α 从 $-\infty$ 变化至 ∞ 时, 广义均值覆盖了由 $\min(a, b)$ 至 $\max(a, b)$ 的整个闭区间。

3.5 本章小结及导读

本章主要阐述了以下内容:

- 模糊补、 s -范数(模糊并)、 t -范数(模糊交)的公理化定义;
- 几类特别的模糊补、 s -范数、 t -范数和平均算子及其性质;
- 怎样证明这些特定的模糊补、 s -范数、 t -范数及平均算子的各种性质。

本章中的内容摘自 Klir 和 Yuan[1995], 从中可以找到更多关于这些算子的详细讲解。Dubois 和 Prade[1985] 则对模糊并、模糊交和平均算子做了很好的综述。

3.6 习题

习题 3.1 定义模糊补 c 的均衡点(equilibrium)为 $c(a) = a (a \in [0, 1])$ 。

- 确定 Yager 模糊补(3.6)的均衡点;
- 证明每个模糊补至多有一个均衡点;
- 证明一个连续模糊补只有一个均衡点。

习题 3.2 证明当 w 趋于无穷时, Yager 的 s -范数(3.10)收敛于基本模糊并(3.2); 当 w 趋于 0 时, Yager 的 s -范数(3.10)收敛于直和(3.11)。

习题 3.3 令模糊集 F 和 G 为习题 2.3 中所定义的 F 和 G 。

- 用 Yager 的 s -范数(3.10)和 t -范数(3.27)(其中 $w = 2$)计算 $F \cup G$ 和 $F \cap G$ 的隶属度函数;
- 式(3.1)用作模糊补, 代数和(3.13)用作模糊并, 代数积(3.30)用作模糊交, 计算 $\overline{F \cap G}, \overline{F \cap G}, \overline{F \cup G}$ 的隶属度函数。

习题 3.4 证明定理 3.2。

习题 3.5 如果对于任意的 $\alpha \in [0, 1]$, 模糊补 c 满足 $c[c(a)] = a$, 则称模糊补 c 是自反的(involutive)。

- 证明 Sugeno 模糊补(3.5)和 Yager 模糊补(3.6)是自反的;
- 令 c 为一自反的模糊补, t 为任一 t -范数, 证明由下式定义的算子 $u: [0, 1] \times [0, 1] \rightarrow [0, 1]$ 是一个 s -范数

$$u(a, b) = c[t(c(a), c(b))] \quad (3.45)$$

- 证明(b)中的 c, t, u 组成了一个关联组。

习题 3.6 确定 s -范数 $s_x(a, b)$, 使得 $s_x(a, b)$ 、最小值 t -范数(3.3)和 Yager 补(3.6)(其中 $w = 2$)可以形成一个关联组。

习题 3.7 证明下面的两组 t -范数和 s -范数与任意一个模糊补 c 都会形成一个关联组

- (\min, \max, c) ; (b) (t_{dp}, s_{ds}, c) 。

习题 3.8 证明在 $\alpha \rightarrow -\infty$ 和 $\alpha \rightarrow \infty$ 时, 广义均值(3.42)分别变成了 \min 和 \max 算子。

4.1 由经典关系到模糊关系

4.1.1 关系

令 U 和 V 为两个任意经典(非模糊、清晰)集合。 U 和 V 的笛卡儿积,记为 $U \times V$,是所有有序对 (u, v) 的非模糊集合,其中 $u \in U, v \in V$,即

$$U \times V = \{(u, v) \mid u \in U \text{ 且 } v \in V\} \quad (4.1)$$

应注意, U 和 V 出现的顺序是很重要的。也就是说,如果 $U \neq V$,则 $U \times V \neq V \times U$ 。一般地,任意 n 个非模糊集 U_1, U_2, \dots, U_n 的笛卡儿积 $U_1 \times U_2 \times \dots \times U_n$,是所有 n 对有序对 (u_1, u_2, \dots, u_n) 的非模糊集合,其中 $u_i \in U_i (i \in \{1, 2, \dots, n\})$; 即

$$U_1 \times U_2 \times \dots \times U_n = \{(u_1, u_2, \dots, u_n) \mid u_1 \in U_1, u_2 \in U_2, \dots, u_n \in U_n\} \quad (4.2)$$

(非模糊)集合 U_1, U_2, \dots, U_n 中的一个(非模糊)关系是笛卡儿积 $U_1 \times U_2 \times \dots \times U_n$ 上的一个模糊子集。即,若用 $Q(U_1, U_2, \dots, U_n)$ 表示 U_1, U_2, \dots, U_n 中的一个关系,则

$$Q(U_1, U_2, \dots, U_n) \subset U_1 \times U_2 \times \dots \times U_n \quad (4.3)$$

作为特例,(非模糊)集合 U 和 V 之间的二元关系是笛卡儿积 $U \times V$ 的一个子集。

例 4.1 令 $U = \{1, 2, 3\}, V = \{2, 3, 4\}$,则 U 和 V 的笛卡儿积为集合 $U \times V = \{(1, 2), (1, 3), (1, 4), (2, 2), (2, 3), (2, 4), (3, 2), (3, 3), (3, 4)\}$ 。 U 和 V 之间的关系是 $U \times V$ 的一个子集。如,若令 $Q(U, V)$ 表示关系“第一个元素不小于第二个元素”,则

$$Q(U, V) = \{(2, 2), (3, 2), (3, 3)\} \quad (4.4)$$

由于关系本身就是一个集合,所以,集合的所有基本运算都可以应用于关系且无须做任何修正。同时,也可以用下面的隶属度函数来表达一个关系:

$$\mu_Q(u_1, u_2, \dots, u_n) = \begin{cases} 1 & (u_1, u_2, \dots, u_n) \in Q(U_1, U_2, \dots, U_n) \\ 0 & \text{其他} \end{cases} \quad (4.5)$$

一般地,当定义在 $U \times V$ 上的二元关系 $Q(U, V)$ 包含的元素个数有限时,可将隶属度函数 μ_Q 的值放到一个关系矩阵中,如下面的例子。

例 4.1(续) 式(4.4)中的关系 $Q(U, V)$ 可用下面的关系矩阵来表述:

$$\begin{array}{cccc}
 & & & V \\
 & & 2 & 3 & 4 \\
 & 1 & 0 & 0 & 0 \\
 U & 2 & 1 & 0 & 0 \\
 & 3 & 1 & 1 & 0
 \end{array} \tag{4.6}$$

经典关系表达的是集合间的清晰关系,即,集合间要么有关系,要么没有关系。但对某些关系来说,很难给出一个 0-1 评价,如下面的例子。

例 4.2 令 $U = \{\text{旧金山, 香港, 北京}\}$, $V = \{\text{波士顿, 香港}\}$, 现打算确定这两个城市集合间“非常远”这一关系。显然,经典关系无法用于该问题,因为“非常远”的概念在经典集合关系的体系中没有完美的定义。而“非常远”又确实有其含义,所以应该可以找到一种数学体系来描述它。如果用区间 $[0, 1]$ 上的一个数来表达“非常远”的程度,则概念“非常远”可用下面的(模糊)关系阵来表述:

$$\begin{array}{ccc}
 & & V \\
 & & \text{波士顿} & \text{香港} \\
 & \text{旧金山} & 0.3 & 0.9 \\
 U & \text{香港} & 1 & 0 \\
 & \text{东京} & 0.95 & 0.1
 \end{array} \tag{4.7}$$

例 4.2 表明,要进一步描述更多的实际关系,需要推广一下经典关系的概念。于是,在此引入模糊关系的概念。

定义 4.1 模糊关系是一个定义在清晰集 U_1, U_2, \dots, U_n 的笛卡儿积上的模糊集。用模糊集的表示方法(2.5),可将 $U_1 \times U_2 \times \dots \times U_n$ 上的模糊关系 Q 定义为如下模糊集合:

$$Q = \{((u_1, u_2, \dots, u_n), \mu_Q(u_1, u_2, \dots, u_n)) \mid (u_1, u_2, \dots, u_n) \in U_1 \times U_2 \times \dots \times U_n\} \tag{4.8}$$

其中, $\mu_Q: U_1 \times U_2 \times \dots \times U_n \rightarrow [0, 1]$ 。

作为特例,二元模糊关系是一个定义在两个清晰集的笛卡儿积上的模糊集。通常用模糊关系阵来表示有限笛卡儿积上的二元关系,即,这个矩阵中的元素是,相应有序对属于该模糊关系的隶属度值。例如,式(4.7)就是一个模糊关系矩阵,它表述的是两个城市集合之间“非常远”的模糊关系。

例 4.3 令 U 和 V 为实数集,即 $U = V = R$, 模糊关系“ x 约等于 y ”,记作 AE ,可用下面的隶属度函数定义:

$$\mu_{AE}(x, y) = e^{-(x-y)^2} \tag{4.9}$$

类似地,模糊关系“ x 远大于 y ”,记作 ML ,可用下面的隶属度函数定义:

$$\mu_{ML}(x, y) = \frac{1}{1 + e^{-(x-y)}} \tag{4.10}$$

当然,也可以用其他隶属度函数来表述这些模糊关系。

4.1.2 投影和柱状扩展

由于清晰关系是定义在两个或多个集合的乘积空间上的,所以,这里引入了投影和柱状扩展的概念。如,集合 $A = \{(x, y) \in R^2 \mid (x-1)^2 + (y-1)^2 \leq 1\}$ 为 $U \times V = R^2$ 中的一个关系,则 A 在 U 上的投影为 $A_1 = [0, 1] \subset U$, A 在 V 上的投影为 $A_2 = [0, 1] \subset V$ (见图 4.1)。 A_1 扩展至 $U \times V = R^2$ 的柱状扩展是 $A_{1E} = [0, 1] \times (-\infty, \infty) \subset R^2$ 。这些概念都可推广至模糊关系。

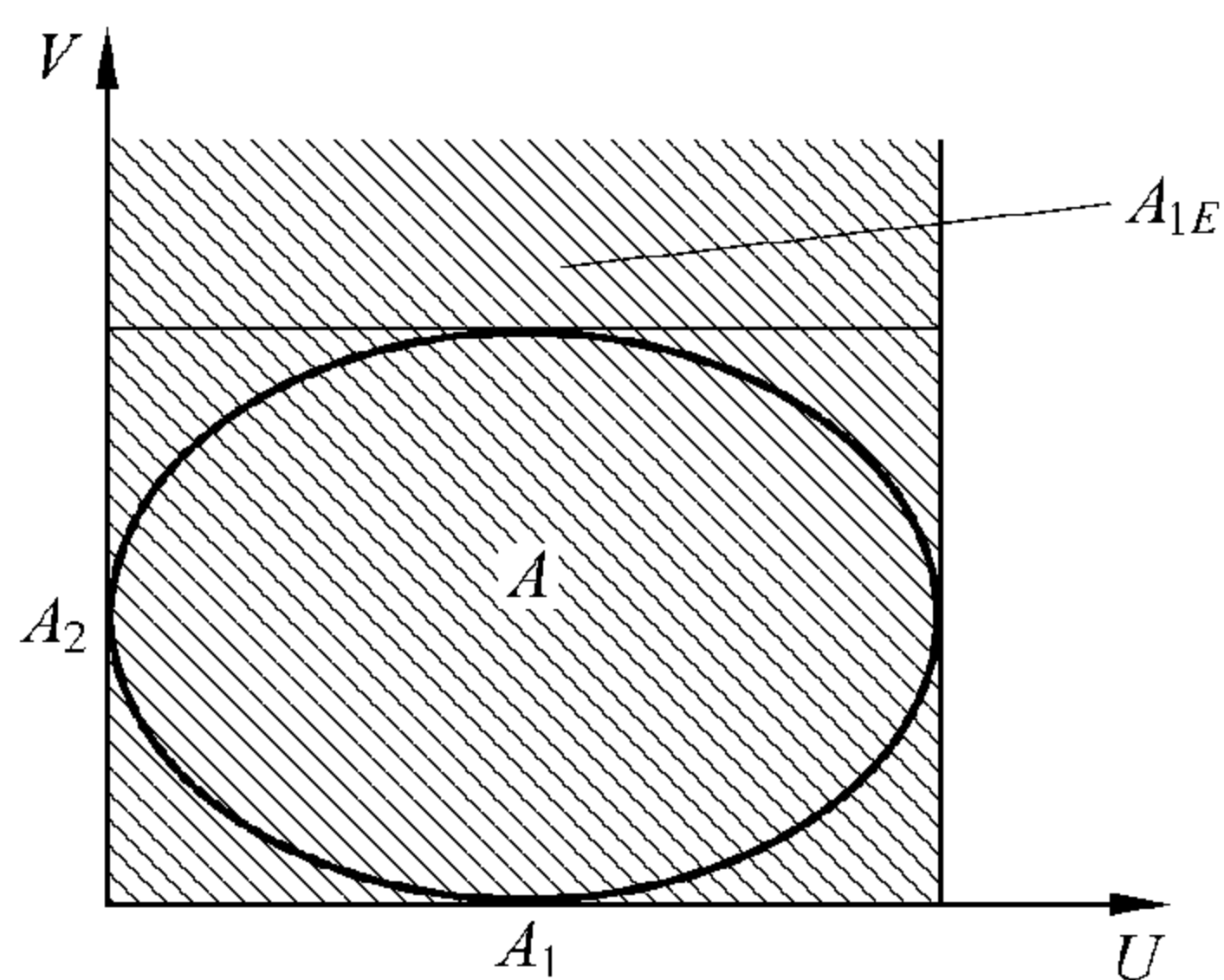


图 4.1 一个关系的投影和柱状扩展

定义 4.2 令 Q 为 $U_1 \times U_2 \times \cdots \times U_n$ 中的一个模糊关系, $\{i_1, \dots, i_k\}$ 为 $\{1, 2, \dots, n\}$ 的一个子集,则 Q 在 $U_{i_1} \times \cdots \times U_{i_k}$ 上的投影是 $U_{i_1} \times \cdots \times U_{i_k}$ 中的一个模糊关系 Q_P ,它可用下面的隶属度函数来定义:

$$\mu_{Q_P}(u_{i_1}, \dots, u_{i_k}) = \max_{u_{j_1} \in U_{j_1}, \dots, u_{j_{n-k}} \in U_{j_{n-k}}} \mu_Q(u_1, \dots, u_n) \quad (4.11)$$

其中, $\{u_{j_1}, \dots, u_{j_{n-k}}\}$ 是 $\{u_{i_1}, \dots, u_{i_k}\}$ 关于 $\{u_1, \dots, u_n\}$ 的补集。

作为特例,如果 Q 是 $U \times V$ 中的一个二元模糊关系,则 Q 在 U 上的投影,记作 Q_1 ,是 U 上的一个模糊集,它可用下面的隶属度函数来定义:

$$\mu_{Q_1}(x) = \max_{y \in V} \mu_Q(x, y) \quad (4.12)$$

应注意, Q 为清晰关系时,式(4.12)仍然成立。例如,如果 Q 是图 4.1 中的清晰关系 A ,则由式(4.12)定义的 A 的投影 Q_1 就是图 4.1 中的 A_1 。因此,式(4.11)所定义的模糊关系的投影是清晰关系的投影的一种自然推广。

例 4.4 根据式(4.12),可得模糊关系(4.7)在 U 和 V 上的投影分别为模糊集

$$Q_1 = 0.9/\text{旧金山} + 1/\text{香港} + 0.95/\text{东京} \quad (4.13)$$

和
$$Q_2 = 1/\text{波士顿} + 0.9/\text{香港} \quad (4.14)$$

类似地,式(4.9)所定义的 AE 在 U 和 V 上的投影分别为模糊集

$$AE_1 = \int_U \max_{y \in V} e^{-(x-y)^2} / x = \int_U 1/x \quad (4.15)$$

和
$$AE_2 = \int_V \max_{x \in U} e^{-(x-y)^2} / y = \int_V 1/y \quad (4.16)$$

应注意, AE_1 等于清晰集 U , AE_2 等于清晰集 V 。

投影将模糊关系约束于一个子空间;相反,柱状扩展则把模糊关系从一个子空间扩展到了整个空间。于是有下面的正式定义:

定义 4.3 令 Q_P 表示 $U_{i_1} \times \cdots \times U_{i_k}$ 中的一个模糊关系, $\{i_1, \dots, i_k\}$ 是 $\{1, 2, \dots, n\}$ 的子

集, 则 Q_P 扩展至 $U_1 \times \cdots \times U_n$ 的柱状扩展为 $U_1 \times \cdots \times U_n$ 中的一个模糊关系 Q_{PE} , 其定义为

$$\mu_{Q_{PE}}(u_1, \cdots, u_n) = \mu_{Q_P}(u_{i_1}, \cdots, u_{i_k}) \quad (4.17)$$

作为特例, 如果 Q_1 是 U 上的一个模糊集合, 则 Q_1 扩展至 $U \times V$ 的柱状扩展就是 $U \times V$ 中的一个模糊关系 Q_{1E} , 其定义为

$$\mu_{Q_{1E}}(x, y) = \mu_{Q_1}(x) \quad (4.18)$$

定义(4.17)对于清晰关系也是成立的(见图 4.1 中的示例)。

例 4.5 考虑例 4.4 中的投影 Q_1 和 Q_2 [见式(4.13)和式(4.14)], 根据式(4.18)可知, 它们扩展至 $U \times V$ 的柱状扩展为

$$Q_{1E} = 0.9/(旧金山, 波士顿) + 0.9/(旧金山, 香港) + 1/(香港, 波士顿) \\ + 1/(香港, 香港) + 0.95/(东京, 波士顿) + 0.95/(东京, 香港) \quad (4.19)$$

$$Q_{2E} = 1/(旧金山, 波士顿) + 1/(香港, 波士顿) + 1/(东京, 波士顿) \\ + 0.9/(旧金山, 香港) + 0.9/(香港, 香港) + 0.9/(东京, 香港) \quad (4.20)$$

类似地, 式(4.15)和式(4.16)中的 AE_1 和 AE_2 扩展至 $U \times V$ 的柱状扩展为

$$AE_{1E} = \int_{U \times V} 1/(x, y) = U \times V \quad (4.21)$$

$$AE_{2E} = \int_{U \times V} 1/(x, y) = U \times V \quad (4.22)$$

从例 4.4 和例 4.5 可以看出, 当得到一个模糊关系的投影并柱状扩展它时, 就会得到一个大于最初关系的模糊关系。为了正式描述这一性质, 首先介绍一下模糊集合的笛卡儿积的概念。令 A_1, \cdots, A_n 分别表示 U_1, \cdots, U_n 上的模糊集合, 则 A_1, \cdots, A_n 的笛卡儿积为 $U_1 \times U_2 \times \cdots \times U_n$ 中的一个模糊关系 $A_1 \times \cdots \times A_n$, 其隶属度函数为

$$\mu_{A_1 \times \cdots \times A_n}(u_1, \cdots, u_n) = \mu_{A_1}(u_1) \star \cdots \star \mu_{A_n}(u_n) \quad (4.23)$$

其中, \star 表示任一 t -范数算子。

引理 4.1 如果 Q 为 $U_1 \times U_2 \times \cdots \times U_n$ 中的一个模糊关系, Q_1, \cdots, Q_n 是 Q 分别在 U_1, \cdots, U_n 上的投影, 则(见图 4.2)

$$Q \subset Q_1 \times \cdots \times Q_n \quad (4.24)$$

其中, $Q_1 \times \cdots \times Q_n$ 的定义(4.23)中的 t -范数采用最小(min)算子。

证明 把式(4.11)代入式(4.17), 可得

$$\mu_{Q_{PE}}(u_1, \cdots, u_n) = \max_{u_{j_1} \in U_{j_1}, \cdots, u_{j_{(n-k)}} \in U_{j_{(n-k)}}} \mu_Q(u_1, \cdots, u_n) \quad (4.25)$$

因此, 对于所有的 $i = 1, 2, \cdots, n$, 都有

$$Q \subset Q_{iE} \quad (4.26)$$

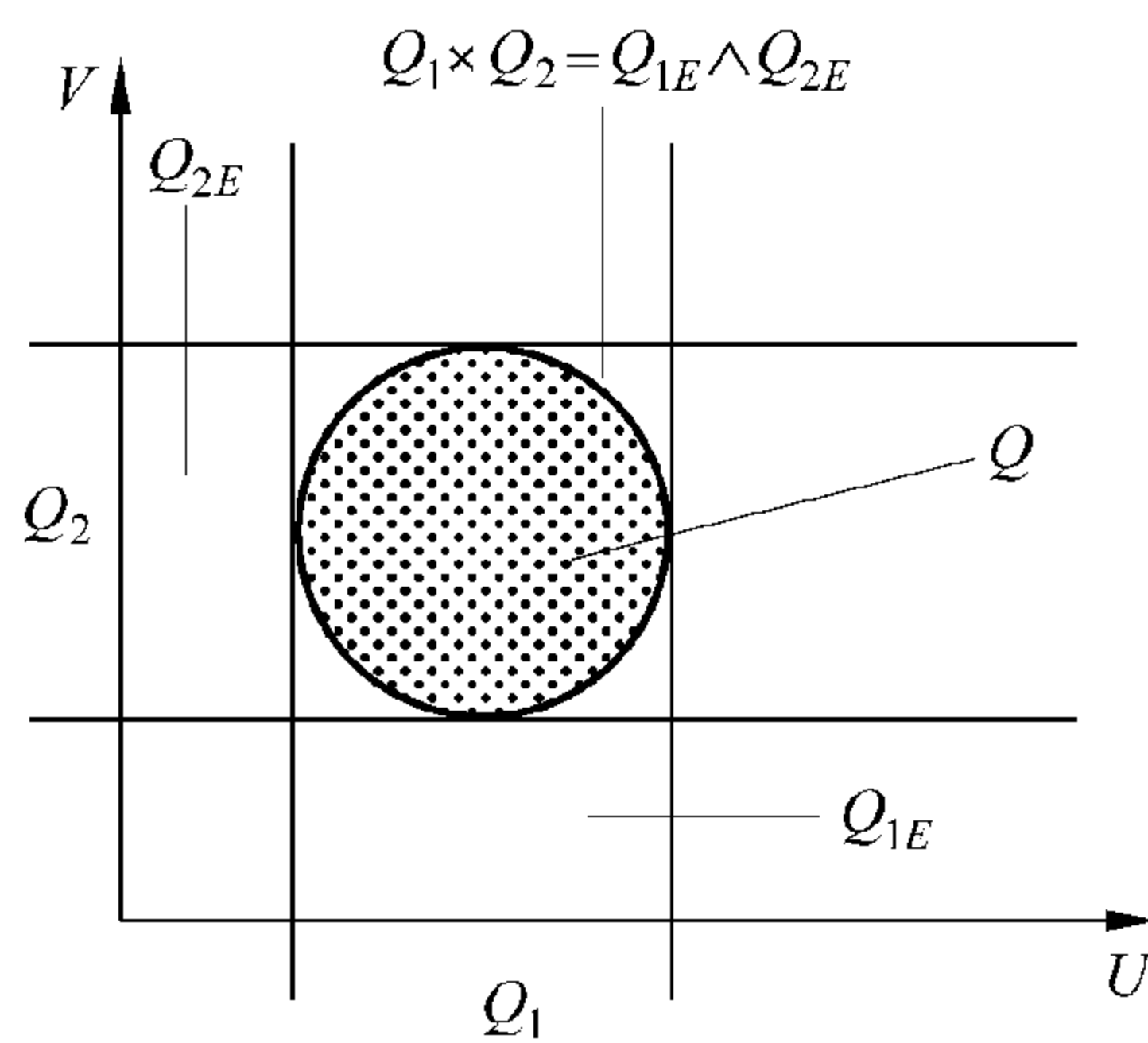


图 4.2 笛卡儿积和柱状集合的交集之间的关系

其中, Q_{iE} 是 Q_i 扩展至 $U_1 \times \cdots \times U_n$ 的柱状扩展。那么, 如果交集采用 \min , 则有

$$Q \subset Q_{1E} \cap \cdots \cap Q_{nE} = Q_1 \times \cdots \times Q_n \quad (4.27)$$

4.2 模糊关系的合成

令 $P(U, V)$ 和 $Q(V, W)$ 表示两个共用一个公共集 V 的清晰二元关系。定义 P 和 Q 的合成为 $U \times W$ 中的一个关系, 记作 $P \circ Q$, 它满足 $(x, z) \in P \circ Q$ 的充要条件是至少存在一个 $y \in V$ 使 $(x, y) \in P$ 且 $(y, z) \in Q$ 。通过使用关系的隶属度函数表达式(4.5), 可以得到合成的一个等价定义, 现以引理形式列到下面:

引理 4.2 对任意 $(x, z) \in U \times W$, 当且仅当

$$\mu_{P \circ Q}(x, z) = \max_{y \in V} t[\mu_P(x, y), \mu_Q(y, z)] \quad (4.28)$$

时, $P \circ Q$ 是 $P(U, V)$ 和 $Q(V, W)$ 的合成, 其中, t 表示任一 t -范数。

证明 根据合成的定义, 首先证明, 如果 $P \circ Q$ 是关系的合成, 则式(4.28)成立。如果 $P \circ Q$ 是关系的合成, 则 $(x, z) \in P \circ Q$ 表明, 存在 $y \in V$ 使 $\mu_P(x, y) = 1$ 且 $\mu_Q(y, z) = 1$ 。由此可得 $\mu_{P \circ Q}(x, z) = 1 = \max_{y \in V} t[\mu_P(x, y), \mu_Q(y, z)]$, 即, 式(4.28)成立; 如果 $(x, z) \notin P \circ Q$, 则对任意 $y \in V$ 有 $\mu_P(x, y) = 0$ 或 $\mu_Q(y, z) = 0$ 。由此可得 $\mu_{P \circ Q}(x, z) = 0 = \max_{y \in V} t[\mu_P(x, y), \mu_Q(y, z)]$ 。因此, 对任意 $(x, z) \in U \times W$, 式(4.28)都是成立的。

反过来, 证明如果式(4.28)成立, 则 $(x, z) \in P \circ Q$, 即 $\max_{y \in V} t[\mu_P(x, y), \mu_Q(y, z)] = 1$, 这说明至少存在一个 $y \in V$ 使 $\mu_P(x, y) = \mu_Q(y, z) = 1$ (见 3.3 节的公理 t_1), 这是合成的定义。对于 $(x, z) \notin P \circ Q$, 由式(4.28)可知, $\max_{y \in V} t[\mu_P(x, y), \mu_Q(y, z)] = 0$, 这说明不存在 $y \in V$ 使 $\mu_P(x, y) = \mu_Q(y, z) = 1$ 。所以, 根据合成的定义和式(4.28)可知, $P \circ Q$ 是关系的合成。

现将合成的概念推广到模糊关系。由引理 4.2 可知, 如果用式(4.28)定义模糊关系合成(假设 P 和 Q 是模糊关系), 则对于 P 和 Q 是清晰关系这一特例, 该定义仍然成立。由此, 给出下面的定义。

定义 4.4 模糊关系 $P(U, V)$ 和 $Q(V, W)$ 的合成, 记作 $P \circ Q$, 是 $U \times W$ 中的一个模糊关系, 其隶属度函数由式(4.28)给出。

因为式(4.28)中的 t -范数可以有多种, 所以每种 t -范数都能得到一个特定的关系合成。文献中最常用的两种关系合成就是最大—最小(max-min)合成和最大—代数积(max-product)合成, 它们的定义如下:

- 模糊关系 $P(U, V)$ 和 $Q(V, W)$ 的最大—最小合成是指由下面的隶属度函数定义的 $U \times W$ 中的模糊关系 $P \circ Q$:

$$\mu_{P \circ Q}(x, z) = \max_{y \in V} \min[\mu_P(x, y), \mu_Q(y, z)] \quad (4.29)$$

其中, $(x, z) \in U \times W$ 。

- 模糊关系 $P(U, V)$ 和 $Q(V, W)$ 的最大一代数积合成是指由下面的隶属度函数定义的 $U \times W$ 中的模糊关系 $P \circ Q$:

$$\mu_{P \circ Q}(x, z) = \max_{y \in V} [\mu_P(x, y) \mu_Q(y, z)] \quad (4.30)$$

其中, $(x, z) \in U \times W$ 。

可以看出, 最大一最小合成和最大一代数积合成是指定义(4.28)中的 t -范数分别采用最小 t -范数和代数积 t -范数时的合成。下面给出两个例子说明模糊合成运算的过程。

例 4.6 令 U 和 V 的定义同例 4.2, $W = \{\text{纽约, 北京}\}$ 。令 $P(U, V)$ 表示式(4.7)所定义的模糊关系“非常远”, 则可将 $V \times W$ 中的模糊关系“非常近”, 记作 $Q(V, W)$, 定义为如下关系矩阵:

		W		
		纽约	北京	
V	波士顿	0.95	0.1	(4.31)
	香港	0.1	0.9	

用式(2.7)中的求和符号, 可将 P 和 Q 表示为

$$P = 0.3/(\text{旧金山, 波士顿}) + 0.9/(\text{旧金山, 香港}) + 1/(\text{香港, 波士顿}) \\ + 0/(\text{香港, 香港}) + 0.95/(\text{东京, 波士顿}) + 0.1/(\text{东京, 香港}) \quad (4.32)$$

$$Q = 0.95/(\text{波士顿, 纽约}) + 0.1/(\text{波士顿, 北京}) + 0.1/(\text{香港, 纽约}) \\ + 0.9/(\text{香港, 北京}) \quad (4.33)$$

下面计算 P 和 Q 的最大一最小合成和最大一代数积合成。首先, 已知 $U \times W$ 中包含六个元素: (旧金山, 纽约), (旧金山, 北京), (香港, 纽约), (香港, 北京), (东京, 纽约), (东京, 北京)。因此, 任务是确定六个元素的隶属度值 $\mu_{P \circ Q}$ 。用最大一最小合成(4.29), 可得

$$\mu_{P \circ Q}(\text{旧金山, 纽约}) = \max \{ \min [\mu_P(\text{旧金山, 波士顿}), \mu_Q(\text{波士顿, 纽约})], \\ \min [\mu_P(\text{旧金山, 香港}), \mu_Q(\text{香港, 纽约})] \} \\ = \max \{ \min(0.3, 0.95), \min(0.9, 0.1) \} = 0.3 \quad (4.34)$$

类似地, 可得

$$\mu_{P \circ Q}(\text{旧金山, 北京}) = \max \{ \min [\mu_P(\text{旧金山, 波士顿}), \mu_Q(\text{波士顿, 北京})], \\ \min [\mu_P(\text{旧金山, 香港}), \mu_Q(\text{香港, 北京})] \} \\ = \max \{ \min(0.3, 0.1), \min(0.9, 0.9) \} = 0.9 \quad (4.35)$$

最终的 $P \circ Q$ 为

$$P \circ Q = 0.3/(\text{旧金山, 纽约}) + 0.9/(\text{旧金山, 北京}) + 0.95/(\text{香港, 纽约}) \\ + 0.1/(\text{香港, 北京}) + 0.95/(\text{东京, 纽约}) + 0.1/(\text{东京, 北京}) \quad (4.36)$$

如果用最大一代数积合成(4.30), 则通过与上面相同的运算程序(用代数积 t -范数代

替最小 t -范数), 即可得到

$$P \circ Q = 0.285 / (\text{旧金山, 纽约}) + 0.81 / (\text{旧金山, 北京}) + 0.95 / (\text{香港, 纽约}) \\ + 0.1 / (\text{香港, 北京}) + 0.9025 / (\text{东京, 纽约}) + 0.095 / (\text{东京, 北京}) \quad (4.37)$$

由式(4.36)、式(4.37)和关系矩阵式(4.7)、式(4.31)可以看出, 一种较简便地计算 $P \circ Q$ 的方法是, 用关系矩阵和矩阵相乘。具体地讲, 用 P 和 Q 分别表示模糊关系 $P(U, V)$ 和 $Q(V, W)$ 的关系矩阵, 则模糊合成 $P \circ Q$ 的关系矩阵可由以下计算方法得到:

- 最大—最小合成法: 写出矩阵乘积 PQ 中的每个元素, 只不过将每个乘积运算看做一个 \min 运算, 每个求和运算看做一个 \max 运算。
- 最大—代数积合成法: 写出矩阵乘积 PQ 中的每个元素, 只不过将每个求和运算看做一个 \max 运算。

下面检验一下, 看是否能用此方法得到式(4.36)和式(4.37), 具体如下:

最大—最小合成法

$$\begin{pmatrix} 0.3 & 0.9 \\ 1 & 0 \\ 0.95 & 0.1 \end{pmatrix} \circ \begin{pmatrix} 0.95 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} = \begin{pmatrix} 0.3 & 0.9 \\ 0.95 & 0.1 \\ 0.95 & 0.1 \end{pmatrix} \quad (4.38)$$

最大—代数积合成法

$$\begin{pmatrix} 0.3 & 0.9 \\ 1 & 0 \\ 0.95 & 0.1 \end{pmatrix} \circ \begin{pmatrix} 0.95 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} = \begin{pmatrix} 0.285 & 0.81 \\ 0.95 & 0.1 \\ 0.9025 & 0.095 \end{pmatrix} \quad (4.39)$$

在例 4.6 中, 论域 U, V 和 W 中包含的元素个数是有限的。而在大多数工程应用中, U, V 和 W 是包含无限个元素的实值空间。为此, 下面给出在连续区域内计算模糊关系合成的例子。

例 4.7 令 $U = V = W = R$, 考虑例 4.3 中由式(4.9)和式(4.10)所定义的模糊关系 AE (约等于) 和 ML (远大于), 应用最大—代数积合成法计算合成 $AE \circ ML$, 可得

$$\mu_{AE \circ ML}(x, z) = \max_{y \in R} \left[\frac{e^{-(x-y)^2}}{1 + e^{-(y-z)}} \right] \quad (4.40)$$

若想计算式(4.40)的右端, 须确定 $y \in R$ 使 $\frac{e^{-(x-y)^2}}{1 + e^{-(y-z)}}$ 达到最大值, 其中, x 和 z 在 R 中是固定不变的。则求得此 y 的必要条件是

$$\frac{\partial}{\partial y} \left[\frac{e^{-(x-y)^2}}{1 + e^{-(y-z)}} \right] = 0 \quad (4.41)$$

因为不可能得到方程式(4.41)的封闭解, 所以无法进一步简化式(4.40)。在实际中, 对于给定的 x 和 z 值, 可先确定方程式(4.41)的数值解, 然后将其代入式(4.40)。将这个例子与例 4.6 比较, 可以看出, 在连续区域上计算模糊关系的合成比在离散区域上计算要困难得多。

4.3 扩展原理

扩展原理(The Extension Principle)是一个基本恒等式,它允许一个函数的区域从 U 上的清晰点扩展至 V 上的模糊集合,更具体地讲,令 $f: U \rightarrow V$ 表示一个从清晰集 U 至清晰集 V 的函数。假定已知 U 上的模糊集 A ,确定一个 V 上由 f 诱导出的模糊集 $B = f(A)$ 。如果 f 是一个一对一的映射,则定义 B 的隶属度函数为

$$\mu_B(y) = \mu_A[f^{-1}(y)], y \in V \quad (4.42)$$

其中, $f^{-1}(y)$ 是 f 的逆,即 $f[f^{-1}(y)] = y$ 。如果 f 不是一对一的映射,则当 U 中的两个点或更多不同的点(具有不同的隶属度值)映射到 V 中的同一点时,模糊就发生了。例如,可能存在 $f(x_1) = f(x_2) = y$,但 $x_1 \neq x_2$ 且 $\mu_A(x_1) \neq \mu_A(x_2)$,这样等式(4.42)的右端就可能取两个不同的值 $\mu_A(x_1 = f^{-1}(y))$ 或 $\mu_A(x_2 = f^{-1}(y))$ 。为了解决这种模糊性,我们设两个隶属度值中较大的那个为 $\mu_B(y)$,即定义 B 的隶属度函数为

$$\mu_B(y) = \max_{x \in f^{-1}(y)} \mu_A(x), y \in V \quad (4.43)$$

其中, $f^{-1}(y)$ 表示所有 $x \in U$ 满足 $f(x) = y$ 的点的集合。恒等式(4.43)就叫做扩展原理。

例 4.8 令 $U = \{1, 2, \dots, 10\}$, $f(x) = x^2$, 令“小”为 U 上的模糊集合,其定义为

$$\text{小} = 1/1 + 1/2 + 0.8/3 + 0.6/4 + 0.4/5 \quad (4.44)$$

则由式(4.43),可得

$$\text{小}^2 = 1/1 + 1/4 + 0.8/9 + 0.6/16 + 0.4/25 \quad (4.45)$$

4.4 本章小结及导读

本章主要阐述了以下内容:

- 模糊关系、投影和柱状扩展的概念;
- 模糊关系的最大—最小合成和最大—代数积合成;
- 扩展原理及其应用。

Zadeh[1971b]和 Zadeh[1975]提出了模糊关系、投影、柱状扩展和模糊关系的合成以及扩展原理的基本思想。Zadeh 的原始文献写得很清晰易懂,仍然是学习这些基础概念的最佳资料。

4.5 习题

习题 4.1 给定 n -ary 关系, 则可得到此关系的多少个不同投影?

习题 4.2 考虑定义在 $U_1 \times \cdots \times U_4$ 中的模糊关系 Q :

$$Q = 0.4/(b, t, y, i) + 0.6/(a, s, x, i) + 0.9/(b, s, y, i) + 1/(b, s, y, j) \\ + 0.6/(a, t, y, j) + 0.2/(c, s, y, i)$$

其中, $U_1 = \{a, b, c\}$, $U_2 = \{s, t\}$, $U_3 = \{x, y\}$, $U_4 = \{i, j\}$ 。

(a) 计算 Q 在 $U_1 \times U_2 \times U_4$, $U_1 \times U_3$ 和 U_4 上的投影;

(b) 计算(a)中的投影扩展至 $U_1 \times U_2 \times U_3 \times U_4$ 的柱状扩展。

习题 4.3 考虑三个二元模糊关系, 它们由下面的关系矩阵定义:

$$Q_1 = \begin{pmatrix} 1 & 0 & 0.7 \\ 0.3 & 0.2 & 0 \\ 0 & 0.5 & 1 \end{pmatrix}, Q_2 = \begin{pmatrix} 0.6 & 0.6 & 0 \\ 0 & 0.6 & 0.1 \\ 0 & 0.1 & 0 \end{pmatrix}, Q_3 = \begin{pmatrix} 1 & 0 & 0.7 \\ 0 & 1 & 0 \\ 0.7 & 0 & 1 \end{pmatrix} \quad (4.46)$$

用最大—最小合成法和最大—代数积合成法计算 $Q_1 \circ Q_2$, $Q_1 \circ Q_3$ 和 $Q_1 \circ Q_2 \circ Q_3$ 。

习题 4.4 已知模糊集 $A = 0.5/-1 + 0.8/0 + 1/1 + 0.4/2$, 函数 $f(x) = x^2$, 根据扩展原理确定模糊集 $f(A)$ 。

习题 4.5 当 $(x, z) = (0, 0), (0, 1), (1, 0), (1, 1)$ 时, 计算例 4.7 中的 $\mu_{AE \circ ML}(x, z)$ 。

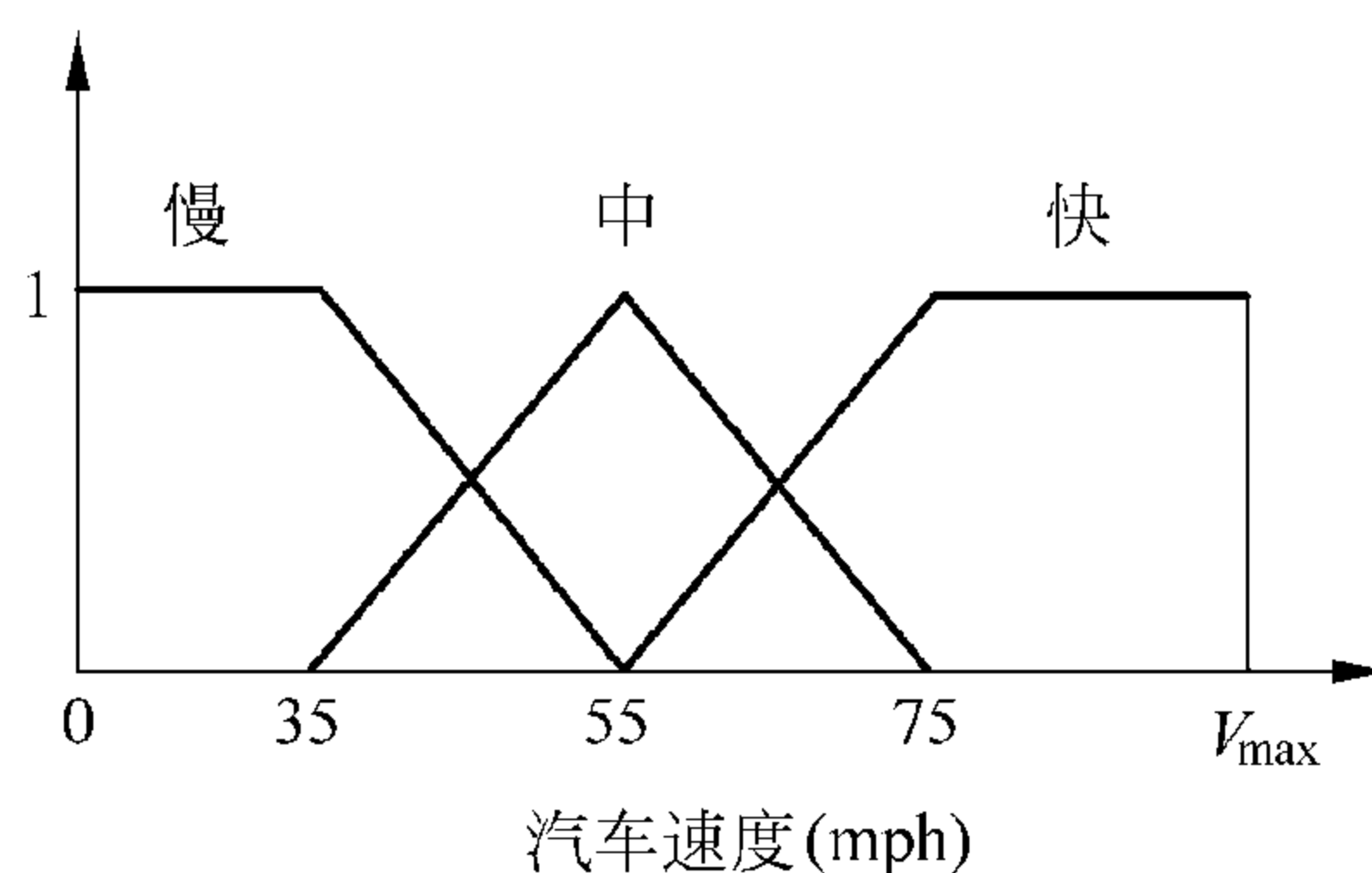
语言变量与模糊 IF-THEN 规则

5.1 由数值变量到语言变量

在日常生活中,变量常用词语来描述,如,当说“今天热”,或者说“今天的气温高”时,就是用词语“高”来描述变量“今天的气温”。也就是说,变量“今天的气温”取值为“高”。当然,变量“今天的气温”也可以取值为 25°C 、 19°C 等。当一个变量取数值时,已经有一个完善的数学体系对其进行描述。而当一个变量取语言值时,在经典数学理论中并没有一个正式的体系可对其进行描述。为了提供这样一种正式体系,引入了语言变量的概念。粗略地讲,如果一个变量取自然语言中的词语为值,则称其为语言变量。现在,问题是怎样用数学术语来描述这些词语。这里用模糊集来描述这些词语。于是,有下面的定义:

定义 5.1 如果一个变量能够取普通语言中的词语为值,则称该变量为语言变量。这里,词语由定义在论域上的模糊集合来描述,变量也是在论域上定义的。

例 5.1 汽车速度是一个变量 x ,其取值范围为区间 $[0, V_{\max}]$,这里, V_{\max} 是汽车的最快速度。在 $[0, V_{\max}]$ 内定义如图 5.1 所示的三个模糊集“慢速”、“中速”、“快速”。如果把 x 看作一个语言变量,则它可取“慢速”、“中速”和“快速”为值,即“ x 为慢速”、“ x 为中速”、“ x 为快速”。当然, x 也可取区间 $[0, V_{\max}]$ 上的数值为值,如 $x = 50\text{mph}$, 35mph 等。



注:汽车速度可以取模糊集“慢速”、“中速”和“快速”为值。

图 5.1 汽车的速度作为一个语言变量

定义 5.1 给出了语言变量的一个简单而直观的定义。在模糊理论文献中,语言变量的更正式的定义见 Zadeh[1973]和[1975]。该定义如下:

定义 5.2 语言变量可表征为四元组 (X, T, U, M) ,其中

- X 为语言变量名称;在例 5.1 中, X 指汽车的速度。
- T 为语言变量 X 取值的术语集合;在例 5.1 中, $T = \{\text{慢速}, \text{中速}, \text{快速}\}$ 。
- U 是语言变量 X 取值的论域;在例 5.1 中, $U = [0, V_{\max}]$ 。
- M 是研究 X 取值的语义规则,即将 T 中的每个语言值和 U 中的模糊集连接起来的

语义规则；在例 5.1 中， M 将“慢速”、“中速”、“快速”同图 5.1 中的隶属度函数连接起来。

比较定义 5.1 和定义 5.2，可以看出，这两种定义实际上是等价的。定义 5.1 更直观些，而定义 5.2 则更正式些。由定义可以看出，语言变量某种意义上是数值变量的一种扩展，即允许语言变量取模糊集为值（见图 5.2）。

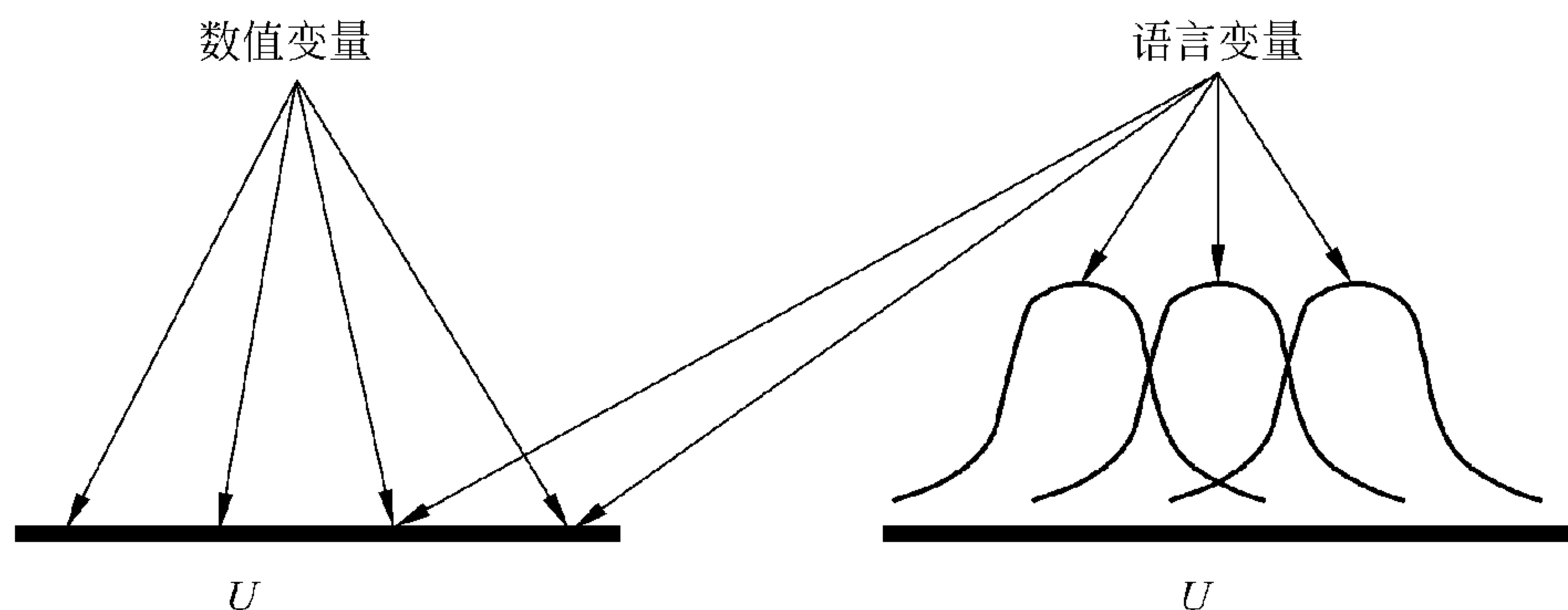


图 5.2 由数值变量到语言变量

为什么语言变量的概念重要呢？这是因为，语言变量是人类知识表达中最基本的元素，当用传感器测量一个变量时，传感器会给出一个数值；而当征求专家对一个变量的评价时，专家会给出语言。例如，当使用雷达枪来测量汽车速度时，雷达枪会给出诸如 39mph, 42mph 等数字；而当让某个人告诉我们汽车速度时，他/她通常会说如“它开得慢”、“它开得快”等话语。因此，引入语言变量的概念就会使自然语言的模糊描述形成精确的数学描述，这是人类知识系统有效地嵌入工程系统的第一步。

5.2 语言限定词

根据语言变量的概念，我们可以将词语赋给语言变量。在日常生活中，我们常用一个单词以上的词语来描述一个变量。例如，如果将汽车速度看做一个语言变量，则它的值可能为“不慢”，“非常慢”，“稍快”，“差不多中速”等等。一般说来，一个语言变量的取值是一个合成术语 $x = x_1 x_2 \cdots x_n$ ，即为元辞 x_1, x_2, \cdots, x_n 的串接。这些元辞可以分成三类：

- 基本术语，它是模糊集的说明性短语；如例 5.1 中，基本术语就是“慢速”、“中速”和“快速”。
- 连接词，“非”、“且”和“或”。
- 限定词，如“非常”、“稍微”、“差不多”等等。

连接词“非”、“且”、“或”在第 2 章和第 3 章中已经研究过了，本节主要研究限定词。

尽管在日常生活中，限定词“非常”并没有很完善的定义，但本质上它起的作用就是一

种增强器。根据这种思想,可得到两种最常用的限定词——非常和差不多的定义。

定义 5.3 令 A 为 U 上的一个模糊集合,则非常 A 也是一个 U 上的模糊集合,它可用如下隶属度函数来定义:

$$\mu_{\text{非常}A}(x) = [\mu_A(x)]^2 \quad (5.1)$$

差不多 A 也是 U 上的一个模糊集合,它可用如下隶属度函数来定义:

$$\mu_{\text{差不多}A}(x) = [\mu_A(x)]^{\frac{1}{2}} \quad (5.2)$$

例 5.2 令 $U = \{1, 2, \dots, 5\}$, 则模糊集合“小”可定义为

$$\text{小} = 1/1 + 0.8/2 + 0.6/3 + 0.4/4 + 0.2/5 \quad (5.3)$$

然后,由式(5.1)和式(5.2),可得

$$\text{非常小} = 1/1 + 0.64/2 + 0.36/3 + 0.16/4 + 0.04/5 \quad (5.4)$$

$$\text{非常非常小} = \text{非常}(\text{非常小}) = 1/1 + 0.4096/2 + 0.1296/3 + 0.0256/4 + 0.0016/5 \quad (5.5)$$

$$\text{差不多小} = 1/1 + 0.8944/2 + 0.7746/3 + 0.6325/4 + 0.4472/5 \quad (5.6)$$

5.3 模糊 IF-THEN 规则

第 1 章中曾提到,在模糊系统与模糊控制中,人类知识可以用模糊 IF-THEN 规则来表述。一条模糊 IF-THEN 规则就是一个条件陈述句,它可以表述为

$$\text{IF} \langle \text{模糊命题} \rangle, \text{THEN} \langle \text{模糊命题} \rangle \quad (5.7)$$

因此,要想理解模糊 IF-THEN 规则,就必须先知道什么是模糊命题。

5.3.1 模糊命题

有两种类型的模糊命题(Fuzzy Propositions): 子模糊命题和复合模糊命题。子模糊命题是一个单独的陈述句

$$x \text{ 为 } A \quad (5.8)$$

这里, x 是语言变量, A 是语言变量 x 的值(即 A 是一个定义在 x 的论域上的模糊集合)。子模糊命题通过连接词“且”、“或”、“非”连接起来而构成的命题叫做复合模糊命题,这里“且”、“或”、“非”分别表示模糊交、模糊并和模糊补。如例 5.1 中,如果用 x 表示汽车的速度,则有以下模糊命题(前三个为子模糊命题,后三个为复合模糊命题):

$$x \text{ 为 } S \quad (5.9)$$

$$x \text{ 为 } M \quad (5.10)$$

$$x \text{ 为 } F \quad (5.11)$$

$$x \text{ 为 } S \text{ 或 } x \text{ 非 } M \quad (5.12)$$

$$x \text{ 非 } S \text{ 或 } x \text{ 非 } F \quad (5.13)$$

$$(x \text{ 为 } S \text{ 且 } x \text{ 非 } F) \text{ 或 } x \text{ 为 } M \quad (5.14)$$

这里, S 、 M 和 F 分别表示模糊集“慢速”、“中速”和“快速”。

应注意, 在一个复合模糊命题中, 子模糊命题是独立的。即, 对于命题(5.12) ~ 命题(5.14)中的同一个变量 x 而言, x 的取值可能是不同的。实际上, 在复合模糊命题中的语言变量一般也是不同的。例如, 令 x 表示汽车速度, $y = \dot{x}$ 表示汽车的加速度, 如果将加速度取值为模糊集合大(L), 则有如下复合模糊命题:

$$x \text{ 为 } F \text{ 且 } y \text{ 为 } L$$

所以, 复合模糊命题应该被理解为一种模糊关系。那么, 怎样确定这些模糊关系的隶属度函数呢?

- 用模糊交表示连接词“且”。具体地讲, 令 x 和 y 分别为定义域 U 和 V 上的语言变量, A 和 B 分别为 U 和 V 上的模糊集合, 则下面的复合模糊命题

$$x \text{ 为 } A \text{ 且 } y \text{ 为 } B \quad (5.15)$$

可以解释为 $U \times V$ 中的模糊关系 $A \cap B$ ^①, 其隶属度函数为

$$\mu_{A \cap B}(x, y) = t[\mu_A(x), \mu_B(y)] \quad (5.16)$$

其中, $t: [0, 1] \times [0, 1] \rightarrow [0, 1]$ 是任意 t -范数。

- 用模糊并表示连接词“或”。具体地讲, 下面的复合模糊命题

$$x \text{ 为 } A \text{ 或 } y \text{ 为 } B \quad (5.17)$$

可以解释为 $U \times V$ 中的模糊关系 $A \cup B$, 其隶属度函数为

$$\mu_{A \cup B}(x, y) = s[\mu_A(x), \mu_B(y)] \quad (5.18)$$

其中, $s: [0, 1] \times [0, 1] \rightarrow [0, 1]$ 是任意 s -范数。

- 用模糊补表示连接词“非”。即, 把非 A 用 \bar{A} 来替代, 它可由第3章中的模糊补算子来定义。

例 5.3 模糊命题(5.14), 即

$$FP = (x \text{ 为 } S \text{ 且 } x \text{ 非 } F) \text{ 或 } x \text{ 为 } M \quad (5.19)$$

是乘积空间 $[0, V_{\max}]^3$ 中的一个模糊关系, 其隶属度函数为

$$\mu_{FP}(x_1, x_2, x_3) = s\{t[\mu_S(x_1), c(\mu_F(x_2))], \mu_M(x_3)\} \quad (5.20)$$

其中, s 、 t 和 c 分别表示 s -范数、 t -范数和模糊补算子, 模糊集(其定义见图 5.1) $S =$ 慢速, $M =$ 中速, $F =$ 快速, $x_1 = x_2 = x_3 = x$ 。

解释形如式(5.7)的模糊 IF-THEN 规则的准备工作已经就绪。

① 注意在第2章和第3章中, A 和 B 都是定义在同一论域 U 上的模糊集合, 且 $A \cup B$ 和 $A \cap B$ 都是 U 上的模糊集合; 这里, $A \cup B$ 和 $A \cap B$ 是 $U \times V$ 中的模糊关系, U 可能等于也可能不等于 V 。

5.3.2 模糊 IF-THEN 规则的解释

由于模糊命题是用模糊关系来解释的,所以剩下的关键问题是怎样解释 IF-THEN 的运算。在经典命题运算中,表达式 *IF p THEN q* 可以写成 $p \rightarrow q$, \rightarrow 可以看做是由表 5.1 所定义的一种连接,这里 p 和 q 都是命题变量,其值为真(T)或为假(F)。

表 5.1 $p \rightarrow q$ 的真值表

p	q	$p \rightarrow q$	p	q	$p \rightarrow q$
T	T	T	F	T	T
T	F	F	F	F	T

由表 5.1 可以看出,如果 p 和 q 的值都为真或都为假,则 $p \rightarrow q$ 为真;如果 p 的值为真, q 的值为假,则 $p \rightarrow q$ 为假;如果 p 的值为假, q 的值为真,则 $p \rightarrow q$ 为真。因此,认为 $p \rightarrow q$ 等价于

$$\bar{p} \vee q \quad (5.21)$$

和

$$(p \wedge q) \vee \bar{p} \quad (5.22)$$

即式(5.21)和式(5.22)与 $p \rightarrow q$ 都符合真值表(表 5.1)的规律,这里 $\bar{}$ 、 \vee 和 \wedge 分别代表(经典)逻辑运算符“非”、“或”和“与”。

由于模糊 IF-THEN 规则可以解释为用模糊命题取代了 p 和 q ,所以,模糊 IF-THEN 规则也可以解释为用模糊补、模糊并和模糊交来分别取代式(5.21)和式(5.22)中的 $\bar{}$ 、 \vee 和 \wedge 算子。因为模糊补、模糊并和模糊交算子有很多种,所以文献中提出的模糊 IF-THEN 规则也就有很多种不同的解释。下面列举了其中的一部分。

将式(5.7)重写为 *IF* $\langle FP_1 \rangle$ *THEN* $\langle FP_2 \rangle$,用 FP_1 和 FP_2 来分别取代式(5.21)和式(5.22)中的 p 和 q ,这里 FP_1 和 FP_2 都是模糊命题。假设 FP_1 是一个定义在 $U = U_1 \times \cdots \times U_n$ 上的模糊关系, FP_2 是一个定义在 $V = V_1 \times \cdots \times V_m$ 上的模糊关系, x 和 y 分别是 U 和 V 上的语言变量(向量)。

- **Dienes-Rescher** 含义:把式(5.21)中的逻辑运算符 $\bar{}$ 和 \vee 分别用式(3.1)和式(3.2)中的基本模糊补和基本模糊并来取代,就可得到 Dienes-Rescher 含义。具体地讲,模糊 IF-THEN 规则 *IF* $\langle FP_1 \rangle$ *THEN* $\langle FP_2 \rangle$,可以解释为 $U \times V$ 中的一个模糊关系 Q_D ,其隶属度函数为

$$\mu_{Q_D}(x, y) = \max[1 - \mu_{FP_1}(x), \mu_{FP_2}(y)] \quad (5.23)$$

- **Lukasiewicz** 含义:用 Yager 的 s -范数(3.10)($w = 1$)来取代式(5.21)中的 \vee ,用基本模糊补(3.1)来取代式(5.21)中的 $\bar{}$,就可得到 Lukasiewicz 含义。具体地讲,模糊 IF-THEN 规则 *IF* $\langle FP_1 \rangle$ *THEN* $\langle FP_2 \rangle$,可以解释为 $U \times V$ 中的一个模糊关系

Q_L , 其隶属度函数为

$$\mu_{Q_L}(x, y) = \min[1, 1 - \mu_{FP_1}(x) + \mu_{FP_2}(y)] \quad (5.24)$$

- **Zadeh** 含义: 这里的模糊 IF-THEN 规则 $IF \langle FP_1 \rangle THEN \langle FP_2 \rangle$ 可以解释为 $U \times V$ 中的一个模糊关系 Q_Z , 其隶属度函数为

$$\mu_{Q_Z}(x, y) = \max[\min(\mu_{FP_1}(x), \mu_{FP_2}(y)), 1 - \mu_{FP_1}(x)] \quad (5.25)$$

显然, 式(5.25)是通过用基本模糊补(3.1)和基本模糊并(3.2)以及基本模糊交(3.3)来分别取代式(5.22)中 $\bar{\cdot}$ 、 \vee 和 \wedge 而得到的。

- **Gödel** 含义: Gödel 含义是经典逻辑中一个众所周知的含义公式。通过将其推广至模糊命题, 得到的模糊 IF-THEN 规则 $IF \langle FP_1 \rangle THEN \langle FP_2 \rangle$ 可以解释为 $U \times V$ 中的一个模糊关系 Q_G , 其隶属度函数为

$$\mu_{Q_G}(x, y) = \begin{cases} 1 & \mu_{FP_1}(x) \leq \mu_{FP_2}(y) \\ \mu_{FP_2}(y) & \text{其他} \end{cases} \quad (5.26)$$

研究这些含义之间的关系是很有趣的。下面的引理就说明了 Zadeh 含义是小于 Dienes-Rescher 含义的, 而 Dienes-Rescher 含义又是小于 Lukasiewicz 含义的。

引理 5.1 对于所有的 $(x, y) \in U \times V$, 都有下式成立

$$\mu_{Q_Z}(x, y) \leq \mu_{Q_D}(x, y) \leq \mu_{Q_L}(x, y) \quad (5.27)$$

证明 由 $0 \leq 1 - \mu_{FP_1}(x) \leq 1$ 且 $0 \leq \mu_{FP_2}(y) \leq 1$, 可得 $\max[1 - \mu_{FP_1}(x), \mu_{FP_2}(y)] \leq 1 - \mu_{FP_1}(x) + \mu_{FP_2}(y)$ 和 $\max[1 - \mu_{FP_1}(x), \mu_{FP_2}(y)] \leq 1$ 。于是可得, $\mu_{Q_D}(x, y) = \max[1 - \mu_{FP_1}(x), \mu_{FP_2}(y)] \leq \min[1, 1 - \mu_{FP_1}(x) + \mu_{FP_2}(y)] = \mu_{Q_L}(x, y)$ 。再由 $\min[\mu_{FP_1}(x), \mu_{FP_2}(y)] \leq \mu_{FP_2}(y)$, 可得 $\max[\min(\mu_{FP_1}(x), \mu_{FP_2}(y)), 1 - \mu_{FP_1}(x)] \leq \max[\mu_{FP_2}(y), 1 - \mu_{FP_1}(x)]$, 即 $\mu_{Q_Z}(x, y) \leq \mu_{Q_D}(x, y)$ 。

理论上讲, 用模糊补, s -范数和 t -范数来分别代替式(5.21)和式(5.22)中的 $\bar{\cdot}$ 、 \vee 和 \wedge , 会得到模糊 IF-THEN 规则的一种特定含义。但这样又引发了一个问题: 应该依据什么标准选择模糊补、 s -范数和 t -范数的组合呢? 这是一个很重要的问题, 我们将在第7章至第10章中对其进行讨论。另一个问题是: 当 p 和 q 是模糊命题时, 命题(5.21)和命题(5.22)仍然与 $p \rightarrow q$ 等价吗? 等价的含义是什么? 现在, 试着回答这个问题。当 p 和 q 为清晰命题时 (即 p 和 q 的值为真或者为假), $p \rightarrow q$ 为全局含义, 如表 5.1 所示, 它涵盖了所有可能的情况。但当 p 和 q 为模糊命题时, $p \rightarrow q$ 可能仅是一个局部含义, 即仅当 p 和 q 具有较大真值时, $p \rightarrow q$ 才有较大的真值。例如, 当说“如果速度为快速, 则阻力大”时, 我们在一定程度上关心的仅是一个局部条件, 这个规则说明, 一切都与“速度为慢速”和“速度为中速”等环境无关。因此, 模糊 IF-THEN 规则

$$IF \langle FP_1 \rangle THEN \langle FP_2 \rangle \quad (5.28)$$

应解释为

$$IF \langle FP_1 \rangle THEN \langle FP_2 \rangle ELSE \langle NOTHING \rangle \quad (5.29)$$

这里, *NOTHING* 指这条规则就不存在了。在逻辑术语中, 它就变成了

$$p \rightarrow q = p \wedge q \quad (5.30)$$

用最小算子或代数积来代替式(5.30)中的 \wedge , 就得到了 Mamdani 含义。

- **Mamdani 含义:** 模糊 IF-THEN 规则(5.28)可以解释为 $U \times V$ 中的一个模糊关系 Q_{MM} 或 Q_{MP} , 其隶属度函数为

$$\mu_{Q_{MM}}(x, y) = \min[\mu_{FP_1}(x), \mu_{FP_2}(y)] \quad (5.31)$$

$$\mu_{Q_{MP}}(x, y) = \mu_{FP_1}(x) \mu_{FP_2}(y) \quad (5.32)$$

Mamdani 含义是在模糊系统与模糊控制中使用最广泛的含义。其成立的论据是, 模糊 IF-THEN 规则为局部含义。尽管, 也许有一些人不同意这一观点。如, 有人可能认为, 当说“速度快则阻力很大”时, 其内在含义是“速度慢则阻力小”。从这个意义上讲, 模糊 IF-THEN 规则又是非局部的。这种争论表明, 当用模糊 IF-THEN 规则来表述人类的知识时, 不同的人会有不同的解释。因此, 要求用不同的含义来处理不同的问题。例如, 如果专家认为其规则是局部的, 则会使用 Mamdani 含义; 否则就会考虑全局含义式(5.23)~式(5.26)。

下面给出 Q_D, Q_L, Q_Z, Q_{MM} 和 Q_{MP} 运算的一些例子。

例 5.4 令 x_1 表示汽车的速度, x_2 表示加速度, y 是施加给油门的力。考虑下面的模糊 IF-THEN 规则

$$\text{如果 } x_1 \text{ 为慢速, } x_2 \text{ 为小, 则 } y \text{ 为大} \quad (5.33)$$

这里“慢速”是定义在图 5.1 中的模糊集, 即

$$\mu_{\text{慢速}}(x_1) = \begin{cases} 1 & x_1 \leq 35 \\ \frac{55 - x_1}{20} & 35 < x_1 \leq 55 \\ 0 & x_1 > 55 \end{cases} \quad (5.34)$$

“小”是定义在加速度的值域上的一个模糊集合, 其隶属度函数为

$$\mu_{\text{小}}(x_2) = \begin{cases} \frac{10 - x_2}{10} & 0 \leq x_2 \leq 10 \\ 0 & x_2 > 10 \end{cases} \quad (5.35)$$

“大”是定义在施加给油门的力的值域上的一个模糊集合, 其隶属度函数为

$$\mu_{\text{大}}(y) = \begin{cases} 0 & y \leq 1 \\ y - 1 & 1 \leq y \leq 2 \\ 1 & y > 2 \end{cases} \quad (5.36)$$

令 x_1, x_2 和 y 的值域分别为 $U_1 = [0, 100], U_2 = [0, 30], V = [0, 3]$ 。如果用代数积来替换式(5.16)中的 t -范数, 则下面的模糊命题就是一个定义在 $U_1 \times U_2$ 中的模糊关系

$$FP_1 = x_1 \text{ 为慢且 } x_2 \text{ 为小} \quad (5.37)$$

其隶属度函数为

$$\mu_{FP_1}(x_1, x_2) = \mu_{慢速}(x_1)\mu_{小}(x_2) = \begin{cases} 0 & x_1 \geq 55 \text{ 或 } x_2 > 10 \\ \frac{10 - x_2}{10} & x_1 \leq 35 \text{ 且 } x_2 \leq 10 \\ \frac{(55 - x_1)(10 - x_2)}{200} & 35 < x_1 \leq 55 \text{ 且 } x_2 \leq 10 \end{cases} \quad (5.38)$$

$\mu_{FP_1}(x_1, x_2)$ 的计算可参见图 5.3。

如果使用 Dienes-Rescher 含义(5.23), 则模糊 IF-THEN 规则就可以解释为 $U_1 \times U_2 \times V$ 中的一个模糊关系 $Q_D(x_1, x_2, y)$, 其隶属度函数为

$$\mu_{Q_D}(x_1, x_2, y) = \max[1 - \mu_{FP_1}(x_1, x_2), \mu_{大}(y)] \quad (5.39)$$

由式(5.38), 可得

$$1 - \mu_{FP_1}(x_1, x_2) = \begin{cases} 1 & x_1 \geq 55 \text{ 或 } x_2 > 10 \\ x_2/10 & x_1 \leq 35 \text{ 且 } x_2 \leq 10 \\ 1 - \frac{(55 - x_1)(10 - x_2)}{200} & 35 < x_1 \leq 55 \text{ 且 } x_2 \leq 10 \end{cases} \quad (5.40)$$

为有助于用最大(max)算子组合(5.40)中的 $1 - \mu_{FP_1}(x_1, x_2)$ 和式(5.36)中的 $\mu_{大}(y)$, 图 5.4 对 $1 - \mu_{FP_1}(x_1, x_2)$ 和 $\mu_{大}(y)$ 的区域的划分及其组合做了解释。由图 5.4 可知

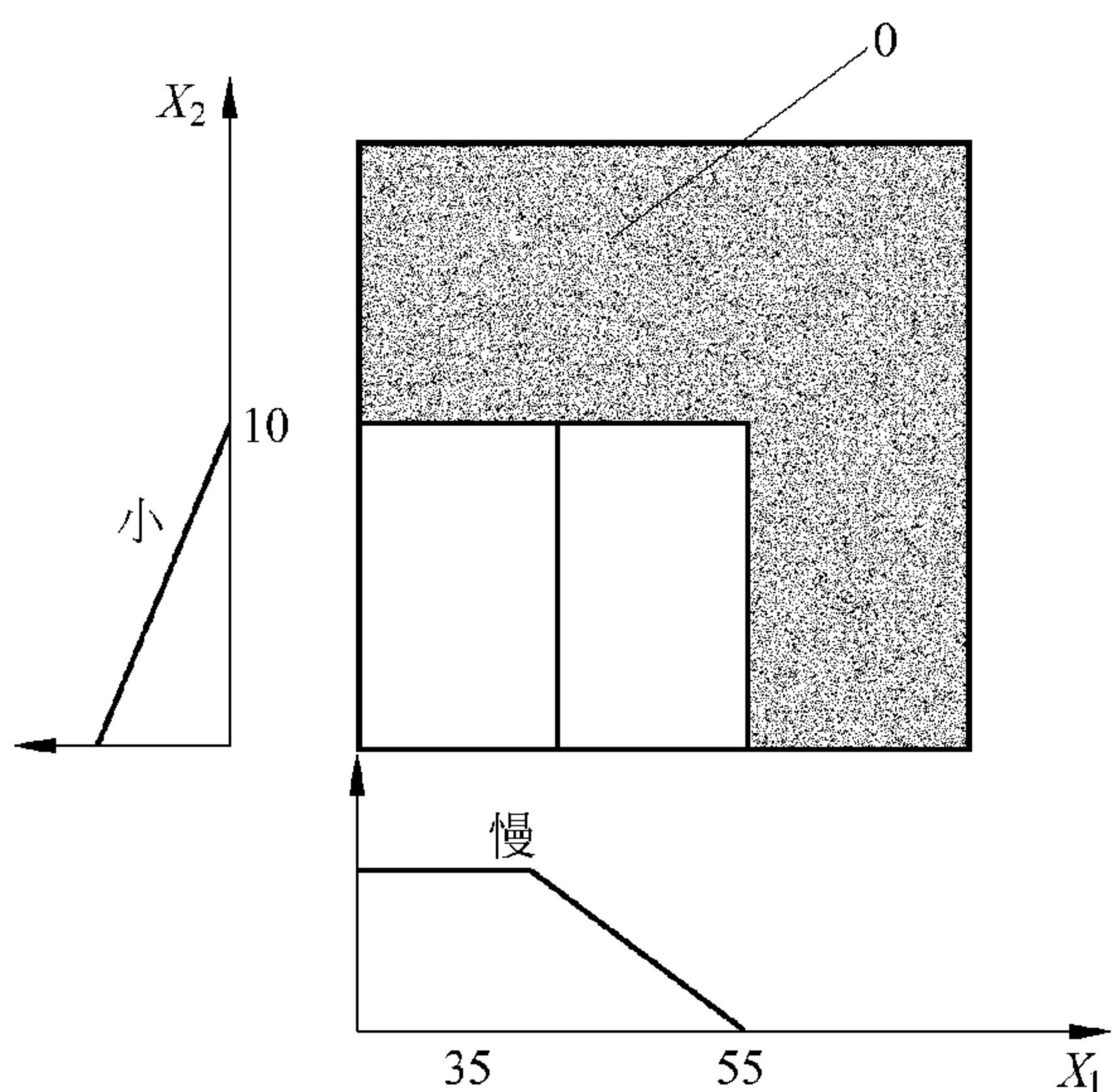


图 5.3 计算例 5.4 中的 $\mu_{慢}(x_1)$ 和 $\mu_{小}(x_2)$ 的图示

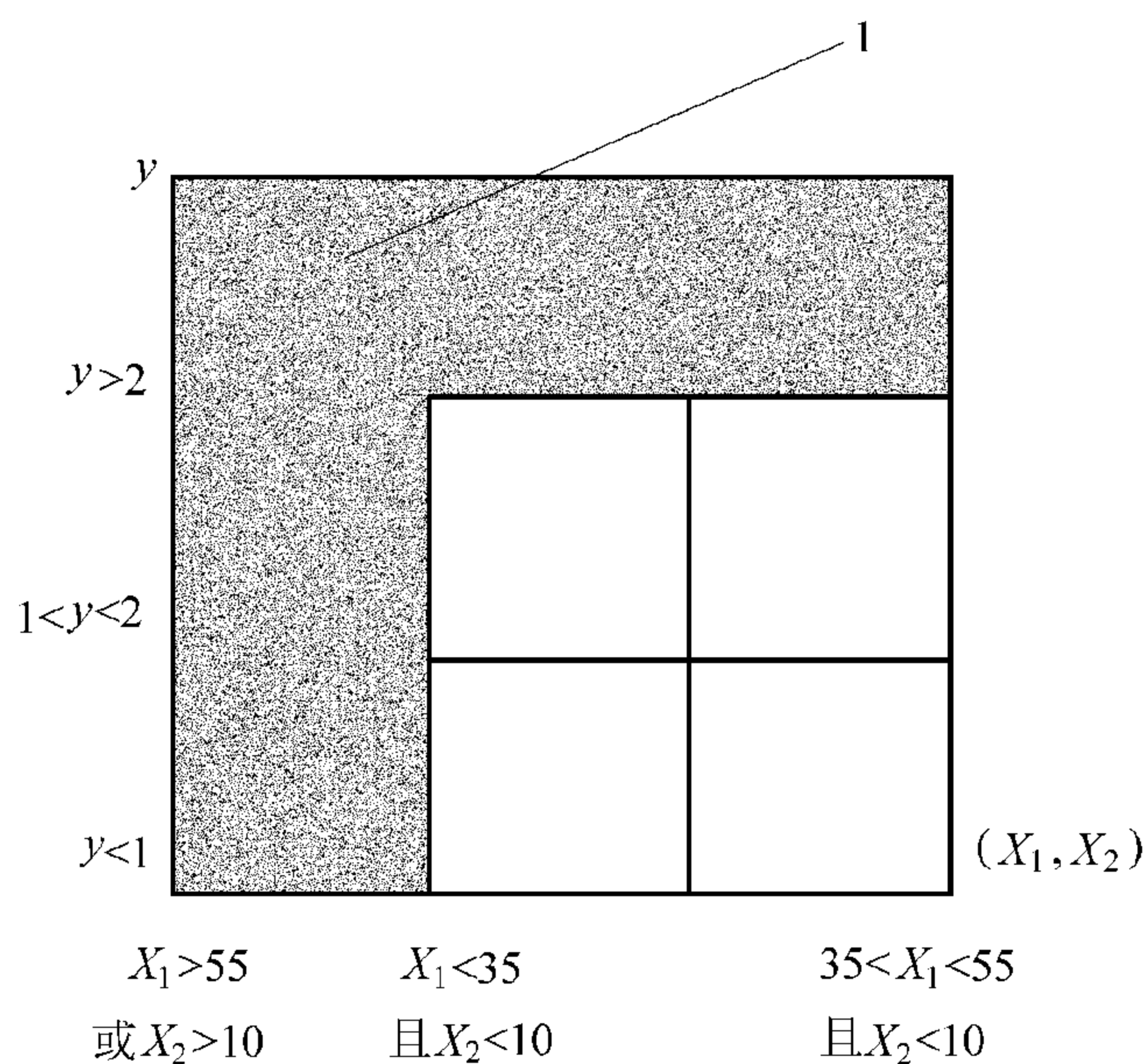


图 5.4 例 5.4 中 $1 - \mu_{FP_1}(x_1, x_2)$ 和 $\mu_{大}(y)$ 及其组合的区域分配

$$\mu_{Q_D}(x_1, x_2, y) = \begin{cases} 1 & x_1 \geq 55 \text{ 或 } x_2 > 10 \text{ 或 } y > 2 \\ x_2/10 & x_1 \leq 35 \text{ 且 } x_2 \leq 10 \text{ 且 } y \leq 1 \\ 1 - \frac{(55-x_1)(10-x_2)}{200} & 35 < x_1 \leq 55 \text{ 且 } x_2 \leq 10 \text{ 且 } y \leq 1 \\ \max[y_1 - 1, x_2/10] & x_1 \leq 35 \text{ 且 } x_2 \leq 10 \text{ 且 } 1 < y \leq 2 \\ \max\left[y - 1, 1 - \frac{(55-x_1)(10-x_2)}{200}\right] & 35 < x_1 \leq 55 \text{ 且 } x_2 \leq 10 \text{ 且 } 1 < y \leq 2 \end{cases} \quad (5.41)$$

可以用相同的方法确定 Lukasiewicz 含义、Zadeh 含义和 Mamdani 含义的隶属度函数。

由例 5.4 可以看出, 如果子模糊命题的隶属度函数不是平滑函数[如式(5.34)~式(5.36)], 则最终的隶属度函数 μ_{Q_D}, μ_{Q_Z} 等的计算虽然直接但很难处理。解决这一复杂性的一种方法就是用一个平滑函数来逼近这个非平滑函数。请看下例:

例 5.4(续) 假设用下式来逼近式(5.34)中的 $\mu_{\text{慢速}}(x_1)$

$$\mu_{\text{慢速}}(x_1) = \frac{1}{1 + e^{\frac{x_1-45}{5}}} \quad (5.42)$$

用下式来逼近式(5.35)中的 $\mu_{\text{小}}(x_2)$

$$\mu_{\text{小}}(x_2) = \frac{1}{1 + e^{\frac{x_2-5}{2}}} \quad (5.43)$$

用下式来逼近式(5.36)中的 $\mu_{\text{大}}(y)$

$$\mu_{\text{大}}(y) = \frac{1}{1 + e^{2(-y+1.25)}} \quad (5.44)$$

现在, 如果使用 Mamdani 乘积含义(5.32), 并用代数积来代替式(5.16)中的 t -范数, 则隶属度函数 $\mu_{Q_{MP}}(x_1, x_2, y)$ 可以很容易地计算如下:

$$\begin{aligned} \mu_{Q_{MP}}(x_1, x_2, y) &= \mu_{\text{慢}}(x_1) \mu_{\text{小}}(x_2) \mu_{\text{大}}(y) \\ &= \frac{1}{(1 + e^{\frac{x_1-45}{5}})(1 + e^{\frac{x_2-5}{2}})(1 + e^{2(-y+1.5)})} \end{aligned} \quad (5.45)$$

例 5.5 令 $U = \{1, 2, 3, 4\}, V = \{1, 2, 3\}$ 。假设已知 $x \in U$ 是 $y \in V$ 的逆命题。为描述这一知识, 可以使用下面的模糊 IF-THEN 规则:

$$\text{如果 } x \text{ 为大, 则 } y \text{ 为小} \quad (5.46)$$

这里, 模糊集“大”和“小”的定义如下:

$$\text{大} = 0/1 + 0.1/2 + 0.5/3 + 1/4 \quad (5.47)$$

$$\text{小} = 1/1 + 0.5/2 + 0.1/3 \quad (5.48)$$

如果使用 Dienes-Rescher 含义(5.23), 则模糊 IF-THEN 规则(5.46)就可以表述为如下 $U \times V$ 中的模糊关系 Q_D

$$Q_D = 1/(1,1) + 1/(1,2) + 1/(1,3) + 1/(2,1) + 0.9/(2,2) + 0.9/(2,3) + 1/(3,1) + 0.5/(3,2) + 0.5/(3,3) + 1/(4,1) + 0.5/(4,2) + 0.1/(4,3) \quad (5.49)$$

如果使用 Lukasiewicz 含义(5.24),则规则(5.46)就变成了

$$Q_L = 1/(1,1) + 1/(1,2) + 1/(1,3) + 1/(2,1) + 1/(2,2) + 1/(2,3) + 1/(3,1) + 1/(3,2) + 0.6/(3,3) + 1/(4,1) + 0.5/(4,2) + 0.1/(4,3) \quad (5.50)$$

如果使用 Zadeh 含义(5.25)和 Gödel 含义(5.26),则有

$$Q_Z = 1/(1,1) + 1/(1,2) + 1/(1,3) + 0.9/(2,1) + 0.9/(2,2) + 0.9/(2,3) + 0.5/(3,1) + 0.5/(3,2) + 0.5/(3,3) + 1/(4,1) + 0.5/(4,2) + 0.1/(4,3) \quad (5.51)$$

$$Q_G = 1/(1,1) + 1/(1,2) + 1/(1,3) + 1/(2,1) + 1/(2,2) + 1/(2,3) + 1/(3,1) + 1/(3,2) + 0.1/(3,3) + 1/(4,1) + 0.5/(4,2) + 0.1/(4,3) \quad (5.52)$$

最后,如果使用 Mamdani 含义(5.31)和(5.32),则模糊 IF-THEN 规则(5.46)就变成了

$$Q_{MM} = 0/(1,1) + 0/(1,2) + 0/(1,3) + 0.1/(2,1) + 0.1/(2,2) + 0.1/(2,3) + 0.5/(3,1) + 0.5/(3,2) + 0.1/(3,3) + 1/(4,1) + 0.5/(4,2) + 0.1/(4,3) \quad (5.53)$$

$$Q_{MP} = 0/(1,1) + 0/(1,2) + 0/(1,3) + 0.1/(2,1) + 0.05/(2,2) + 0.01/(2,3) + 0.5/(3,1) + 0.25/(3,2) + 0.05/(3,3) + 1/(4,1) + 0.5/(4,2) + 0.1/(4,3) \quad (5.54)$$

由式(5.49)~式(5.52)可以看出,对于不包含在规则式(5.46)之内的数据对,即(1,1),(1,2)和(1,3)(因为 $\mu_x(1)=0$), Q_D , Q_L , Q_Z 和 Q_G 给出的隶属度值是1,而 Q_{MM} 和 Q_{MP} 给出的隶属度值是0。这同前面的结论,Dienes-Rescher 含义、Lukasiewicz 含义、Zadeh 含义以及 Gödel 含义为全局含义,而 Mamdani 含义为局部含义的结论是一致的。

5.4 本章小结及导读

本章主要阐述了以下内容:

- 语言变量的概念和限定词的描述;
- 模糊命题和模糊 IF-THEN 规则的概念;
- 模糊 IF-THEN 规则的不同含义,包括 Dienes-Rescher 含义,Lukasiewicz 含义,Zadeh 含义和 Gödel 含义以及 Mamdani 含义;
- 这些含义的性质及其计算。

语言变量是在 Zadeh 的开创性文章 Zadeh[1973]中引入的。本文是另一篇学术文章,建议读者认真研读。Zadeh[1975]是由三部分组成的,文中概括了许多与语言变量有关的概念和原理。

5.5 习题

习题 5.1 给出三个语言变量的例子, 并把这些语言变量组合成一个复合模糊命题, 然后确定其隶属度函数。

习题 5.2 给出一些 5.2 节中以外的语言限定词, 并给出它们的合理运算公式。

习题 5.3 令 Q_L, Q_G, Q_{MM} 和 Q_{MP} 分别为式(5.24), 式(5.26), 式(5.31)和式(5.32)中所定义的模糊关系, 证明

$$Q_{MP} \subset Q_{MM} \subset Q_G \subset Q_L \quad (5.55)$$

习题 5.4 用基本模糊算子式(3.1) ~ 式(3.3)分别表示“非”、“或”和“且”, 确定模糊命题式(5.12)和式(5.13)的隶属度函数并画出其图形。

习题 5.5 考虑模糊 IF-THEN 规则(5.33), 规则中的模糊集“慢速”、“小”和“大”分别由式(5.42)、式(5.43)和式(5.44)定义。用 \min 来代替式(5.16)中的 t -范数, 并计算模糊关系 $Q_D, Q_L, Q_Z, Q_G, Q_{MM}$ 和 Q_{MP} 。

习题 5.6 令 Q 表示 $U \times U$ 中的一个模糊关系。当 $\mu_Q(u, U) = 1 (u \in U)$ 时, 称 Q 是自反的(reflexive)。证明: 如果 Q 是自反的, 则(a) $Q \circ Q$ 也是自反的; (b) $Q \subseteq Q \circ Q$ 。这里, \circ 是最大—最小合成。

6.1 由经典逻辑到模糊逻辑

逻辑学是研究推理的方法和原理的一门学科,这里推理表示由现命题得到新命题。在经典逻辑中,可能要求命题为真或为假,也就是说,一个命题的真值为 0 或 1。模糊逻辑允许一个命题的真值为区间 $[0,1]$ 上的任意值,从而推广了仅取两个值的经典逻辑。这种推广要用到近似推理,即由一组不精确的前提集合(模糊命题)推出不精确的结论(模糊命题)。本章首先回顾经典逻辑中的一些基本概念和原理,然后研究由经典逻辑向模糊逻辑推广的问题。

6.1.1 经典逻辑概要

在经典逻辑中,命题间的关系通常用一个真值表来描述。析取“ \vee ”,合取“ \wedge ”,蕴含“ \rightarrow ”,等价“ \leftrightarrow ”和否定“ \neg ”的基本真值表都放在表 6.1 中,表中的符号 T 和 F 分别表示真和假。

表 6.1 命题中常用的五种运算的真值表

p	q	$p \wedge q$	$p \vee q$	$p \rightarrow q$	$p \leftrightarrow q$	\bar{p}
T	T	T	T	T	T	F
T	F	F	T	F	F	F
F	T	F	T	T	F	T
F	F	F	F	T	T	T

给定 n 个基本命题 p_1, \dots, p_n , 则通过一个把某一特定的真值(给定命题的真值的一个组合)赋给新命题的函数就可以定义一个新命题,通常称新命题为逻辑函数。因为 n 个命题可能产生 2^n 个真值组合,所以就能得到由 2^n 个可能的逻辑函数定义的 n 个命题。由于 n 较大时, 2^{2^n} 是一个巨大的数,因此,经典逻辑中的一个关键问题就是把所有的逻辑函数仅用几个基本逻辑算子来表达;称基本的逻辑算子为初始全集。最常用的初始全集是否定 \neg 、析

取 \vee 和合取 \wedge 。将 \neg 、 \vee 和 \wedge 组合到适当的代数表达式中,就形成了逻辑公式,这样就能形成任意其他的逻辑函数了。于是,将逻辑函数做如下递归定义

- (a) 真值0和1都是逻辑公式;
- (b) 如果 p 是一个命题,则 p 和 \bar{p} 都是逻辑公式;
- (c) 如果 p 和 q 是逻辑公式,则 $p \vee q$ 和 $p \wedge q$ 也是逻辑公式;
- (d) 只有(a)~(c)中所定义的公式才是逻辑公式。

当由一个逻辑公式所表达的命题总为真,而不论公式中的基本命题的真值为何值时,称该逻辑公式为永真公式(tautology);而当其所表达的命题总为假时,称该逻辑公式为永假公式(contradiction)。

例 6.1 下面的逻辑公式就是永真公式

$$(p \rightarrow q) \leftrightarrow (\bar{p} \vee q) \quad (6.1)$$

$$(p \rightarrow q) \leftrightarrow ((p \wedge q) \vee \bar{p}) \quad (6.2)$$

可以用真值表法来证明式(6.1)和式(6.2),即,列出式(6.1)和式(6.2)的所有可能值,然后看它们是否都为真。表 6.2 给出的结果表明式(6.1)和式(6.2)都是永真公式。

表 6.2 $(p \rightarrow q) \leftrightarrow (\bar{p} \vee q)$ 和 $(p \rightarrow q) \leftrightarrow ((p \wedge q) \vee \bar{p})$ 的证明

p	q	$p \rightarrow q$	$\bar{p} \vee q$	$(p \wedge q) \vee \bar{p}$	$(p \rightarrow q) \leftrightarrow (\bar{p} \vee q)$	$(p \rightarrow q) \leftrightarrow ((p \wedge q) \vee \bar{p})$
T	T	T	T	T	T	T
T	F	F	F	F	T	T
F	T	T	T	T	T	T
F	F	T	T	T	T	T

永真公式的各类形式都可用于作推理,这就是所谓的推理规则。三个最常用的推理规则如下:

- 取式推理(**Modus Ponens**)。这一推理规则表述的是,给定两个命题 p 和 $p \rightarrow q$ (称之为前提),可以推出命题 q 为真命题(称之为结论),记为

$$(p \wedge (p \rightarrow q)) \rightarrow q \quad (6.3)$$

取式推理可更直观地表述为

前提 1: x 为 A

前提 2: 如果 x 为 A , 则 y 为 B

结论: y 为 B

- 拒式推理(**Modus Tollens**)。这一推理规则陈述的是,给定两个命题 \bar{q} 和 $p \rightarrow q$,可以推出命题 \bar{p} 为真命题,记为

$$(\bar{q} \wedge (p \rightarrow q)) \rightarrow \bar{p} \quad (6.4)$$

拒式推理可更直观地表述为

前提 1: y 非 B

前提 2: 如果 x 为 A , 则 y 为 B

结论: x 非 A

- 假言推理(**Hypothetical Syllogism**)。这一推理规则陈述的是, 给定两个命题 $p \rightarrow q$ 和 $q \rightarrow r$, 可以推出命题 $p \rightarrow r$ 为真命题, 记为

$$((p \rightarrow q) \wedge (q \rightarrow r)) \rightarrow (p \rightarrow r) \quad (6.5)$$

假言推理可更直观地表述为

前提 1: 如果 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

6.1.2 模糊逻辑的基本原理

在模糊逻辑中, 命题都是由模糊集表述的模糊命题(其定义见第 5 章)。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, 引入了所谓的广义取式推理、广义拒式推理和广义假言推理, 它们都是模糊逻辑中的基本原理。

- 广义取式推理(**Generalized Modus Ponens**)。这一推理规则陈述的是, 给定两个模糊命题“ x 为 A' ”和“如果 x 为 A , 则 y 为 B ”, 可推出一个新模糊命题“ y 为 B' ”。这里, A, A', B 和 B' 都是模糊集且 A' 和 A 很近似, B' 和 B 很近似, 即

前提 1: x 为 A'

前提 2: 如果 x 为 A , 则 y 为 B

结论: y 为 B'

表 6.3 给出了在广义取式推理中, 与前提 1 和结论有关的一些直观准则。注意到, 如果在前提 2 中, “ x 为 A ”和“ y 为 B ”之间的推理关系不太强时, 准则 $p3$ 和 $p5$ 可以成立。准则 $p7$ 可以理解为“如果 x 为 A , 则 y 为 B , 否则 y 非 B ”。尽管这一关系在经典逻辑中并不成立, 但在日常推理中却经常做出这类推理。

表 6.3 在广义取式推理中, 前提 2 给定, 与前提 1 和结论有关的直观准则

	x 为 A' (前提 1)	y 为 B' (结论)
准则 $p1$	x 为 A	y 为 B
准则 $p2$	x 为非常 A	y 为非常 B
准则 $p3$	x 为非常 A	y 为 B
准则 $p4$	x 为差不多 A	y 为差不多 B

续表

	x 为 A' (前提 1)	y 为 B' (结论)
准则 $p5$	x 为差不多 A	y 为 B
准则 $p6$	x 非 A	y 未知
准则 $p7$	x 非 A	y 非 B

- 广义拒式推理(**Generalized Modus Tollens**)。这一推理规则陈述的是,给定两个模糊命题“ y 为 B' ”和“如果 x 为 A ,则 y 为 B ”,可推出一个新命题“ x 为 A' ”。这里, A,A',B 和 B' 都是模糊集, B' 和 B 之间的差异越大, A' 和 A 之间的差异就越大,即

前提 1: y 为 B'

前提 2: 如果 x 为 A ,则 y 为 B

结论: x 为 A'

表 6.4 给出了在广义拒式推理中,与前提 1 和结论有关的一些直观的准则。和表 6.3 中的准则类似,表 6.4 中的某些准则在经典逻辑中也是不成立的,但在现实生活可以近似地使用它们。

表 6.4 在广义拒式推理中,前提 2 给定,与前提 1 和结论有关的直观准则

	y 为 B' (前提 1)	x 为 A' (结论)
准则 $t1$	y 非 B	x 非 A
准则 $t2$	y 为非非常 B	x 为非非常 A
准则 $t3$	y 为非差不多 B	x 为非差不多 A
准则 $t4$	y 为 B	x 未知
准则 $t5$	y 为 B	x 为 A

- 广义假言推理(**Generalized Hypothetical Syllogism**)。这一推理规则陈述的是,给定两个模糊命题,“如果 x 为 A ,则 y 为 B ”和“如果 y 为 B' ,则 z 为 C ”,可推出一个新的命题“如果 x 为 A ,则 z 为 C' ”。这里, A,B,B',C 和 C' 是模糊集合, B 和 B' 很近似, C 和 C' 很近似,即

前提 1: 如果 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'

表 6.5 给出了在广义假言推理中,与“ y 为 B' ”和“ z 为 C' ”有关的一些直观准则。准则 $s2$ 是根据以下直觉知识得到的:为了使前提 1 中的 B 和前提 2 中的相匹配,即

“ $B' = \text{非常 } B$ ”，可以将前提1变为“如果 x 为非常 A ，则 y 为非常 B ”，这样就会得出“如果 x 为非常 A ，则 z 为 C ”的结论。然后用限定词“差不多”来代替“非常”，则有“如果 x 为 A ，则 z 差不多为 C ”，即为准则 s_2 。其他准则都可用类似方法来说明。

表 6.5 在广义假言推理中，与前提 2 中的 y 为 B' 和结论中的 z 为 C' 有关的直观准则

	y 为 B' (前提 2)	z 为 C' (结论)
准则 s_1	y 为 B	z 为 C
准则 s_2	y 为非常 B	z 为差不多 C
准则 s_3	y 为非常 B	z 为 C
准则 s_4	y 为差不多 B	z 为非常 C
准则 s_5	y 为差不多 B	z 为 C
准则 s_6	y 非 B	z 未知
准则 s_7	y 非 B	z 非 C

称表 6.3 ~ 表 6.5 中的准则为直观准则，是因为它们在选择特定模糊集合时并不一定成立，这也正是近似推理的内涵所在。尽管这些准则并不绝对成立，但它们是很有意义的。在设计某一推理时，可以将其作为指导原则(或软约束)。

现已给出了模糊逻辑的三条基本原理(广义取式推理，广义拒式推理和广义假言推理)的基本思想，下一个问题就是，在给定前提条件下，怎样确定结论中模糊命题的隶属度函数。为了解答这一问题，提出了推理合成规则。

6.2 推理合成规则

推理合成规则是下列程序(见图 6.1)的一种推广：假设从 $x \in U$ 到 $y \in V$ 上有一条曲线 $y = f(x)$ ，并给定 $x = a$ ，则由 $x = a$ 和 $y = f(x)$ 可以推出 $y = b = f(a)$ 。

将上述程序进行推广，即假设 a 为一个区间， $f(x)$ 为一个区间值函数(见图 6.2)。为找到由 a 和 $f(x)$ 所推出的区间 b ，首先要构建一个由 a 扩展而来的柱状集合 a_E ，然后找到它和区间值曲线的交集 I 。那么， I 在 V 上的投影就是所求的区间 b 。

再进行下一步推广，假设 A' 是 U 上的一个模糊集合， Q 是 $U \times V$ 中的一个模糊关系。则由 A' 的柱状扩

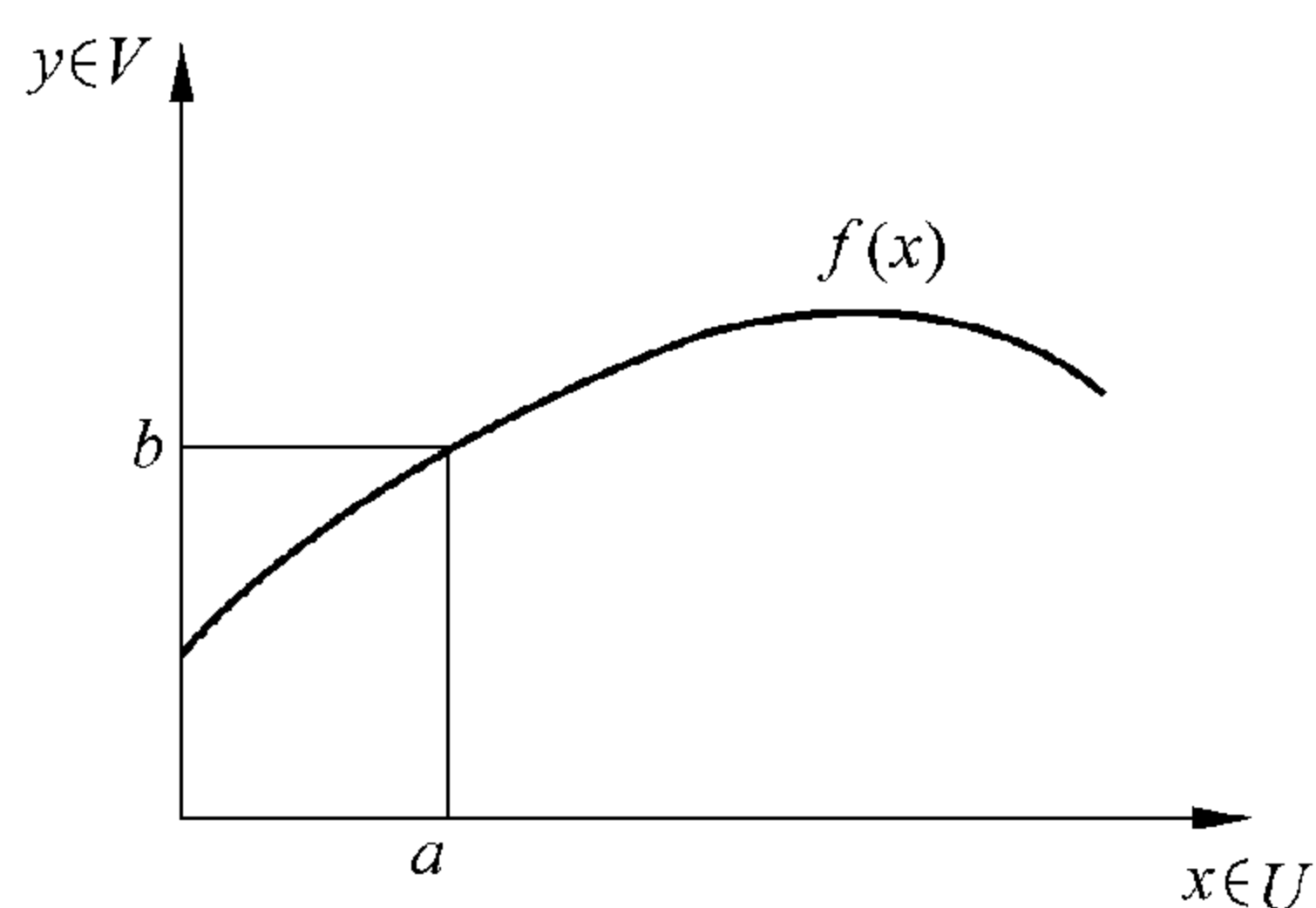


图 6.1 由 $x = a, y = f(x)$ 推出 $y = b$

展 A'_E 和其模糊关系 Q 的交集(见图 6.3),即可得到模糊集 $A'_E \cap Q$,它和图 6.2 中的交集 I 类似。那么, $A'_E \cap Q$ 在 y 轴上的投影即为模糊集合 B' 。

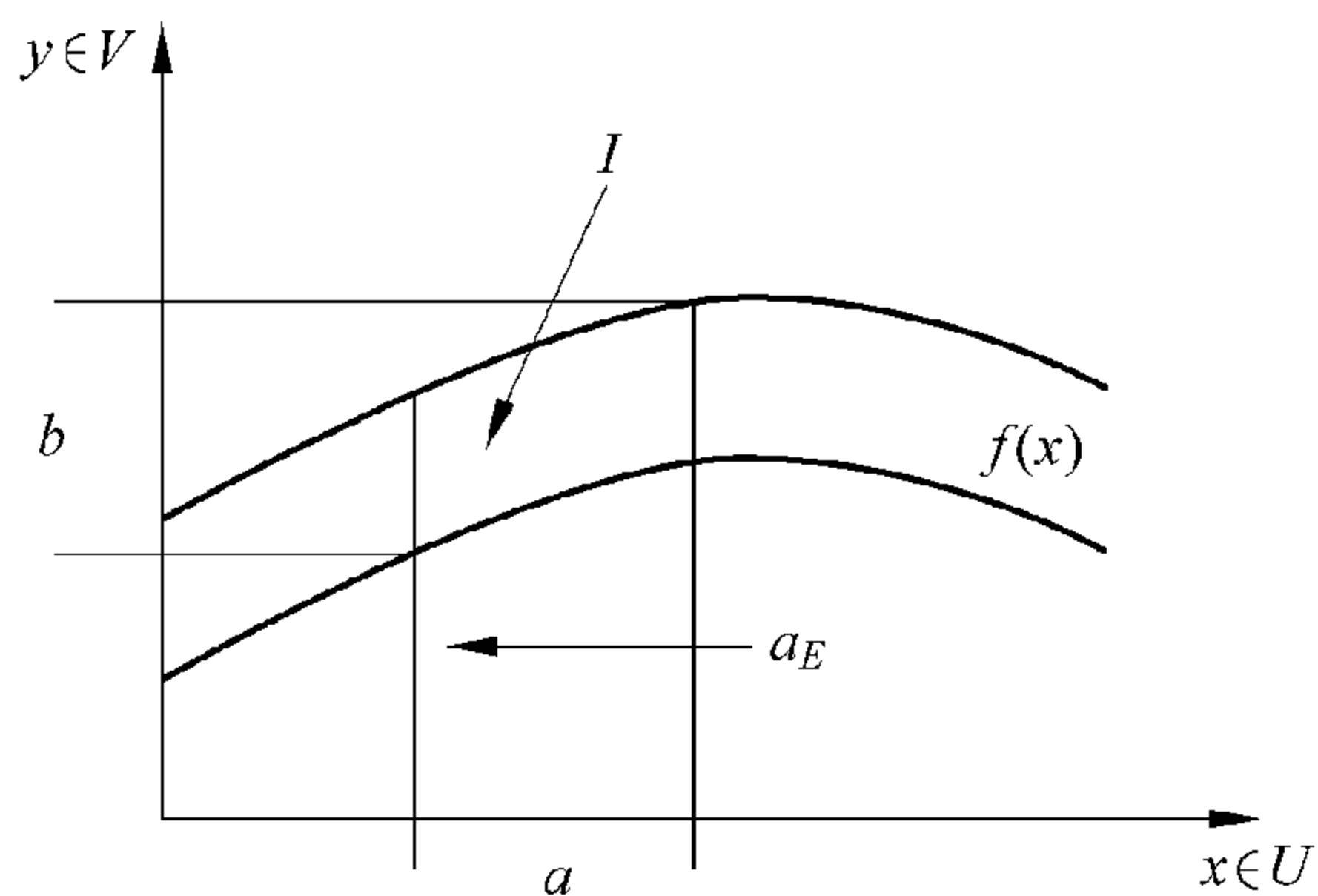


图 6.2 由区间 a 和区间值函数 $f(x)$ 推出区间 b

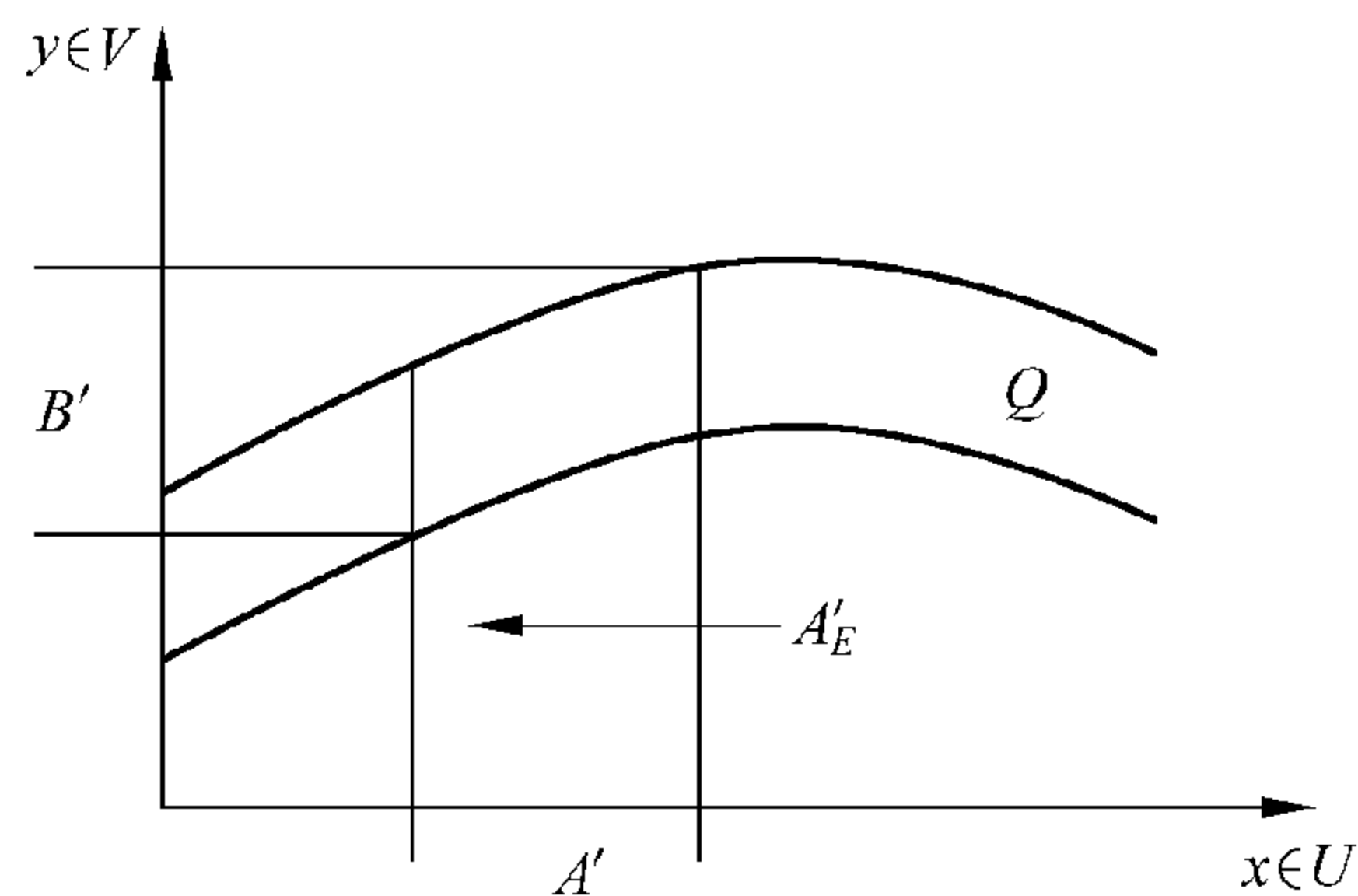


图 6.3 由模糊集 A' 和模糊关系 Q 推出模糊集 B'

更具体地讲,给定 $\mu_{A'}(x)$ 和 $\mu_Q(x, y)$, 可得[见式(4.18)]

$$\mu_{A'_E}(x, y) = \mu_{A'}(x) \quad (6.6)$$

进而可得

$$\mu_{A'_E \cap Q}(x, y) = t[\mu_{A'_E}(x, y), \mu_Q(x, y)] = t[\mu_{A'}(x), \mu_Q(x, y)] \quad (6.7)$$

最后,由式(4.12),可得 B' , 它为 $A'_E \cap Q$ 在 V 上的投影,即

$$\mu_{B'}(y) = \sup_{x \in U} t[\mu_{A'}(x), \mu_Q(x, y)] \quad (6.8)$$

称式(6.8)为推理合成规则(compositional rule of inference)。文献中常用符号“ \star ”表示 t -范数算子,所以,式(6.8)也可写成

$$\mu_{B'}(y) = \sup_{x \in U} [\mu_{A'}(x) \star \mu_Q(x, y)] \quad (6.9)$$

推理合成规则也可叫做 $sup-\star$ 合成。

第 5 章曾讲过,模糊 IF-THEN 规则,如“如果 x 为 A , 则 y 为 B ”可以解释为 x, y 论域的笛卡儿积中的一个模糊关系。含义不同,模糊关系也就不同,可参见式(5.23)~式(5.26)、式(5.31)和式(5.32)。因此,在广义取式推理和广义拒式推理中,前提 2 也可以看做是式(6.9)中的模糊关系 Q 。至于广义假言推理,可以简单地将其看成是两个模糊关系的合成,所以,可以根据合成(4.28)即可得到结论。综上,可以得到广义取式推理、广义拒式推理和广义假言推理中计算结论的详细公式,具体如下:

- 广义取式推理: 给定模糊集 A' (表明前提为“ x 为 A' ”)和 $U \times V$ 中的模糊关系 $A \rightarrow B$ (表明前提为“如果 x 为 A , 则 y 为 B ”),可推出 U 上的一个模糊集合 B' (表明结论为“ y 为 B' ”),其隶属度函数为

$$\mu_{B'}(y) = \sup_{x \in U} t[\mu_{A'}(x), \mu_{A \rightarrow B}(x, y)] \quad (6.10)$$

- 广义拒式推理: 给定模糊集合 B' (表明前提为“ y 为 B' ”)和 $U \times V$ 中的模糊关系

$A \rightarrow B$ (表明前提为“如果 x 为 A , 则 y 为 B ”), 可推出 U 上的模糊集合 A' (表明结论为“ x 为 A' ”), 其隶属度函数为

$$\mu_{A'}(x) = \sup_{y \in V} t[\mu_{B'}(y), \mu_{A \rightarrow B}(x, y)] \quad (6.11)$$

- 广义假言推理: 给定 $U \times V$ 中的模糊关系 $A \rightarrow B$ (表明前提为“ x 为 A , 则 y 为 B ”) 和 $V \times W$ 中的模糊关系 $B' \rightarrow C$ (表明前提为“如果 y 为 B' , 则 z 为 C' ”), 可推出 $U \times W$ 中的一个模糊关系 (表明结论为“如果 x 为 A , 则 z 为 C' ”), 其隶属度函数为

$$\mu_{A \rightarrow C'}(x, z) = \sup_{y \in V} t[\mu_{A \rightarrow B}(x, y), \mu_{B' \rightarrow C}(y, z)] \quad (6.12)$$

若在式(6.10) ~ 式(6.12)中采用不同的 t -范数和不同的含义规则式(5.23) ~ 式(5.26), 式(5.31)和式(5.32), 则可得到多个结论。结论说明了这些含义规则的特性, 下面将对这些特性中的部分特性进行研究。

6.3 规则的特性

本节将特定含义的规则和 t -范数应用于式(6.10) ~ 式(6.12), 然后, 在 A' 和 B' 的某些典型情况下, 观察 $\mu_{B'}(y)$, $\mu_{A'}(x)$ 和 $\mu_{A \rightarrow C'}(x, z)$ 的特性。下面将依次在不同含义的规则下研究广义取式推理、广义拒式推理和广义假言推理。

6.3.1 广义取式推理

例 6.2 假设广义取式推理(6.10)中的 t -范数用最小(\min)算子取代, $\mu_{A \rightarrow B}(x, y)$ 用 Mamdani 含义(5.32)取代。分别在 A' 的四种情况: (a) $A' = A$, (b) $A' =$ 非常 A , (c) $A' =$ 差不多 A , (d) $A' = \bar{A}$ 下确定相应的 B' 的隶属度函数。假设 $\sup_{x \in U} [\mu_A(x)] = 1$ (模糊集 A 为标准模糊集)。当 $A' = A$ 时, 有

$$\begin{aligned} \mu_{B'}(y) &= \sup_{x \in U} \{ \min[\mu_A(x), \mu_A(x)\mu_B(y)] \} \\ &= \sup_{x \in U} [\mu_A(x)\mu_B(y)] = \mu_B(y) \end{aligned} \quad (6.13)$$

当 $A' =$ 非常 A 时, 有

$$\mu_{B'}(y) = \sup_{x \in U} \{ \min[\mu_A^2(x), \mu_A(x)\mu_B(y)] \} \quad (6.14)$$

由于 $\sup_{x \in U} [\mu_A(x)] = 1$, 且 x 可以在 U 中取任意值, 所以, 对任意 $y \in V$ 存在 $x \in U$ 使得 $\mu_A(x) \geq \mu_B(y)$, 于是, 式(6.14)可以简化为

$$\mu_{B'}(y) = \sup_{x \in U} [\mu_A(x)\mu_B(y)] = \mu_B(y) \quad (6.15)$$

当 $A' =$ 差不多 A 时, 由 $\mu_A^{1/2}(x) \geq \mu_A(x) \geq \mu_A(x)\mu_B(y)$, 可得

$$\mu_{B'}(y) = \sup_{x \in U} \{ \min[\mu_A^{1/2}(x), \mu_A(x)\mu_B(y)] \} = \mu_B(y) \quad (6.16)$$

最后, 当 $A' = \bar{A}$ 时, 有

$$\mu_{B'}(y) = \sup_{x \in U} \{ \min[1 - \mu_A(x), \mu_A(x)\mu_B(y)] \} \quad (6.17)$$

因为, 对于固定不变的 $y \in V$ 来说, $\mu_A(x)\mu_B(y)$ 是一个关于 $\mu_A(x)$ 的增函数, 而 $1 - \mu_A(x)$ 是一个关于 $\mu_A(x)$ 的减函数, 所以, 当 $1 - \mu_A(x) = \mu_A(x)\mu_B(y)$ 时, 即当 $\mu_A(x) = \frac{1}{1 + \mu_B(y)}$ 时,

可达到式(6.17)中的 $\sup \min$ 值, 由此可得

$$\mu_{B'}(y) = \frac{\mu_B(y)}{1 + \mu_B(y)} \quad (6.18)$$

由式(6.13)、式(6.15)、式(6.16)和式(6.18)及表 6.3, 可以看出, 该例中考虑的这个特殊的广义取式推理满足准则 $P1, P3, P5$, 但不满足准则 $P2, P4, P6$ 和 $P7$ 。

例 6.3 本例仍假设广义取式推理(6.10)中的 t -范数用最小 (\min) 算子取代, $\mu_{A \rightarrow B}(x, y)$ 用 Zadeh 含义(5.25)取代。考虑例 6.2 中的四种情况并假设 $\sup_{x \in U} [\mu_A(x)] = 1$ 。

(a) 当 $A' = A$ 时, 有

$$\mu_{B'}(y) = \sup_{x \in U} \min \{ \mu_A(x), \max[\min(\mu_A(x), \mu_B(y)), 1 - \mu_A(x)] \} \quad (6.19)$$

因为 $\sup_{x \in U} \mu_A(x) = 1$, 所以当 x_0 满足

$$\mu_A(x_0) = \max[\min(\mu_A(x_0), \mu_B(y)), 1 - \mu_A(x_0)] \quad (6.20)$$

时, 在点 $x_0 \in U$ 处可达到式(6.19)中的 $\sup \min$ 值。如果 $\mu_A(x_0) < \mu_B(y)$, 则式(6.20)变为

$$\mu_A(x_0) = \max[\mu_A(x_0), 1 - \mu_A(x_0)] \quad (6.21)$$

上式在 $\mu_A(x_0) \geq 0.5$ 时成立, 所以, 由式(6.19)和式(6.20), 可得, $\mu_{B'}(y) = \mu_A(x_0)$ 。因为 $\sup_{x \in U} [\mu_A(x)] = 1$, 所以 $\mu_A(x_0) = 1$ 是一定成立的, 但这又使得 $\mu_{B'}(y) > \mu_A(x_0) = 1$, 这是不可能的。因此, 不存在 $\mu_A(x_0) < \mu_B(y)$ 的情况。现在考虑惟一可能的情况, 即, $\mu_A(x_0) \geq \mu_B(y)$ 。在这种情况下, 式(6.20)变成了

$$\mu_A(x_0) = \max[\mu_B(y), 1 - \mu_A(x_0)] \quad (6.22)$$

如果 $\mu_B(y) < 1 - \mu_A(x_0)$, 则有 $\mu_A(x_0) = 1 - \mu_A(x_0)$, 这在 $\mu_A(x_0) = 0.5$ 时是成立的。如果 $\mu_B(y) \geq 1 - \mu_A(x_0)$, 则由式(6.22), 可得, $\mu_A(x_0) = \mu_B(y) \geq 0.5$ 。于是可得

$$\mu_{B'}(y) = \mu_A(x_0) = \max[0.5, \mu_B(y)] \quad (6.23)$$

(b) 当 $A' =$ 非常 A 时, 有

$$\mu_{B'}(y) = \sup_{x \in U} \min \{ \mu_A^2(x), \max[\min(\mu_A(x), \mu_B(y)), 1 - \mu_A(x)] \} \quad (6.24)$$

和 $A' = A$ 的情况类似, 当 x_0 满足

$$\mu_A^2(x_0) = \max[\min(\mu_A(x_0), \mu_B(y)), 1 - \mu_A(x_0)] \quad (6.25)$$

时,在点 $x_0 \in U$ 处可以达到 $\sup_{x \in U} \min$ 值。如果 $\mu_A(x_0) < \mu_B(y)$, 则有

$$\mu_A^2(x_0) = \max[\mu_A(x_0), 1 - \mu_A(x_0)] \quad (6.26)$$

上式只有在 $\mu_A(x_0) = 1$ 时才成立,但这又产生了矛盾 $\mu_B(y) > 1$ 。所以, $\mu_A(x_0) \geq \mu_B(y)$ 才是惟一可能的情况。如果 $\mu_A(x_0) \geq \mu_B(y)$, 则式(6.25)变为

$$\mu_A^2(x_0) = \max[\mu_B(y), 1 - \mu_A(x_0)] \quad (6.27)$$

如果 $\mu_B(y) < 1 - \mu_A(x_0)$, 则有 $\mu_A^2(x_0) = 1 - \mu_A(x_0)$, 该式在 $\mu_A(x_0) = \frac{\sqrt{5}-1}{2}$ 时成立。所以, 当

$\mu_B(y) < 1 - \mu_A(x_0) = \frac{3-\sqrt{5}}{2}$ 时, 可得 $\mu_{B'}(y) = \mu_A^2(x_0) = \frac{3-\sqrt{5}}{2}$ 。当 $\mu_B(y) \geq 1 - \mu_A(x_0)$ 时, 可

得 $\mu_{B'}(y) = \mu_A^2(x_0) = \mu_B(y) \geq \frac{3-\sqrt{5}}{2}$ 。综上所述可得

$$\mu_{B'}(y) = \mu_A^2(x_0) = \max\left[\frac{3-\sqrt{5}}{2}, \mu_B(y)\right] \quad (6.28)$$

(c) 当 $A' =$ 差不多 A 时, 有

$$\mu_{B'}(y) = \sup_{x \in U} \min\{\mu_A^{1/2}(x), \max[\min(\mu_A(x), \mu_B(y)), 1 - \mu_A(x)]\} \quad (6.29)$$

当 x_0 满足

$$\mu_A^{1/2}(x_0) = \max[\min(\mu_A(x_0), \mu_B(y)), 1 - \mu_A(x_0)] \quad (6.30)$$

时,在点 $x_0 \in U$ 处可以达到 $\sup_{x \in U} \min$ 值。与 $A' =$ 非常 A 的情况类似, 可以证明 $\mu_A(x_0) < \mu_B(y)$ 是不成立的。当 $\mu_A(x_0) \geq \mu_B(y)$ 时, 有

$$\mu_A^{1/2}(x_0) = \max[\mu_B(y), 1 - \mu_A(x_0)] \quad (6.31)$$

如果 $\mu_B(y) < 1 - \mu_A(x_0)$, 则有 $\mu_A^{1/2}(x_0) = 1 - \mu_A(x_0)$, 它在 $\mu_A(x_0) = \frac{3-\sqrt{5}}{2}$ 时成立。所以, 当

$\mu_B(y) < 1 - \mu_A(x_0) = \frac{\sqrt{5}-1}{2}$ 时, 可得 $\mu_{B'}(y) = \mu_A^{1/2}(x_0) = \frac{\sqrt{5}-1}{2}$ 。当 $\mu_B(y) \geq 1 - \mu_A(x_0)$ 时, 可

得 $\mu_{B'}(y) = \mu_A^{1/2}(x_0) = \mu_B(y) \geq \frac{\sqrt{5}-1}{2}$ 。综上所述可得

$$\mu_{B'}(y) = \mu_A^{1/2}(x_0) = \max\left[\frac{\sqrt{5}-1}{2}, \mu_B(y)\right] \quad (6.32)$$

(d) 最后, 当 $A' = \bar{A}$ 时, 有

$$\mu_{B'}(y) = \sup_{x \in U} \min\{1 - \mu_A(x), \max[\min(\mu_A(x), \mu_B(y)), 1 - \mu_A(x)]\} \quad (6.33)$$

由式(6.33)可以看出,若选择 $x_0 \in U$ 使 $\mu_A(x_0) = 0$, 则 $1 - \mu_A(x_0) = 1$, $\max[\min(\mu_A(x), \mu_B(y)), 1 - \mu_A(x)] = 1$, 这样在 $x = x_0$ 处就可以达到 $\sup_{x \in U} \min$ 值。因此,在这种情况下有

$$\mu_{B'}(y) = 1 \quad (6.34)$$

由式(6.23),式(6.28),式(6.32)和式(6.34)可以看出,表 6.3 中的所有准则中,只有准则 P6 是满足的(近似推理确实近似)。

6.3.2 广义拒式推理

例 6.4 和例 6.2 类似,广义拒式推理(6.11)中的 t -范数用最小(min)算子取代, $\mu_{A \rightarrow B}(x, y)$ 用 Mamdani 含义(5.32)取代。考虑 B' 的四种情况:(a) $B' = \bar{B}$; (b) $B' =$ 不非常 B ; (c) $B' =$ 非常不多 B ; (d) $B' = B$ 。假设 $\sup_{y \in V} [\mu_B(y)] = 1$ 。当 $B' = \bar{B}$ 时,由式(6.11),可得

$$\mu_{A'}(x) = \sup_{y \in V} \min[1 - \mu_B(y), \mu_A(x) \mu_B(y)] \quad (6.35)$$

当 $y_0 \in V$ 满足 $1 - \mu_B(y_0) = \mu_A(x) \mu_B(y_0)$ 时,可在 y_0 处达到 $\sup_{y \in V} \min$ 值,由此可得

$$\mu_B(y_0) = \frac{1}{1 + \mu_A(x)}$$

因此,有

$$\mu_{A'}(x) = 1 - \mu_B(y_0) = \frac{\mu_A(x)}{1 + \mu_A(x)} \quad (6.36)$$

当 $B' =$ 不非常 B 时,有

$$\mu_{A'}(x) = \sup_{y \in V} \min[1 - \mu_B^2(y), \mu_A(x) \mu_B(y)] \quad (6.37)$$

其中,当 $y_0 \in V$ 满足 $1 - \mu_B^2(y_0) = \mu_A(x) \mu_B(y_0)$ 时,可在 y_0 处达到 $\sup_{y \in V} \min$ 值,由此可得

$$\mu_B(y_0) = \frac{\sqrt{\mu_A^2(x) + 4} - \mu_A(x)}{2}。因此,有$$

$$\mu_{A'}(x) = \mu_A(x) \mu_B(y_0) = \frac{\mu_A(x) \sqrt{\mu_A^2(x) + 4} - \mu_A^2(x)}{2} \quad (6.38)$$

当 $B' =$ 非差不多 B 时,有

$$\mu_{A'}(x) = \sup_{y \in V} \min[1 - \mu_B^{1/2}(y), \mu_A(x) \mu_B(y)] \quad (6.39)$$

当 $y_0 \in V$ 满足 $1 - \mu_B^{1/2}(y_0) = \mu_A(x) \mu_B(y_0)$ 时,可在 y_0 处达到 $\sup_{y \in V} \min$ 值,由此可得

$$\mu_B(y_0) = \frac{1 + 2\mu_A(x) - \sqrt{\mu_A^2(x) + 1}}{2\mu_A^2(x)}$$

因此,有

$$\mu_{A'}(x) = \mu_A(x)\mu_B(y_0) = \frac{1 + 2\mu_A(x) - \sqrt{\mu_A^2(x) + 1}}{2\mu_A(x)} \quad (6.40)$$

最后,当 $B' = B$ 时,有

$$\mu_{A'}(x) = \sup_{y \in V} \min[\mu_B(y), \mu_A(x)\mu_B(y)] = \sup_{y \in V} \mu_A(x)\mu_B(y) = \mu_A(x) \quad (6.41)$$

由式(6.36)、式(6.38)、式(6.40)和式(6.41)可以看出,表6.4中的7个直观准则中只有准则 t5 是满足的。

6.3.3 广义假言推理

例 6.5 与例 6.2 和例 6.4 类似,广义假言推理(6.12)中的 t -范数用最小(\min)算子取代, $\mu_{A \rightarrow B}(x, y)$ 和 $\mu_{B' \rightarrow C}(y, z)$ 用 Mamdani 含义取代。假设 $\sup_{y \in V} [\mu_B(y)] = 1$, 并考虑 B' 的四种情况: (a) $B' = B$; (b) $B' =$ 非常 B ; (c) $B' =$ 差不多 B ; (d) $B' = \bar{B}$ 。当 $B' = B$ 时,由式(6.12), 可得

$$\begin{aligned} \mu_{A \rightarrow C'}(x, z) &= \sup_{y \in V} \min[\mu_A(x)\mu_B(y), \mu_B(y)\mu_C(z)] \\ &= (\sup_{y \in V} \mu_B(y)) \min[\mu_A(x), \mu_C(z)] = \min[\mu_A(x), \mu_C(z)] \end{aligned} \quad (6.42)$$

当 $B' =$ 非常 B 时,有

$$\mu_{A \rightarrow C'}(x, z) = \sup_{y \in V} \min[\mu_A(x)\mu_B(y), \mu_B^2(y)\mu_C(z)] \quad (6.43)$$

如果 $\mu_A(x) > \mu_C(z)$, 则当 $\mu_A(x)\mu_B(y) > \mu_B^2(y)\mu_C(z)$ 时它总是成立的, 因此, 有 $\mu_{A \rightarrow C'}(x, z) = \sup_{y \in V} \mu_B^2(y)\mu_C(z) = \mu_C(z)$ 。

如果 $\mu_A(x) \leq \mu_C(z)$, 则当 $y_0 \in V$ 满足 $\mu_A(x)\mu_B(y_0) = \mu_B^2(y_0)\mu_C(z)$ 时, 可在 y_0 处达到 $\sup_{y \in V}$ \min 值, 由此可得 $\mu_B(y_0) = \frac{\mu_A(x)}{\mu_C(z)}$ 。因此, 有 $\mu_{A \rightarrow C'}(x, z) = \mu_A(x)\mu_B(y_0) = \frac{\mu_A^2(x)}{\mu_C(z)}$ 。综上可得

$$\mu_{A \rightarrow C'}(x, z) = \begin{cases} \mu_C(z) & \mu_C(z) < \mu_A(x) \\ \frac{\mu_A^2(x)}{\mu_C(z)} & \mu_C(z) \geq \mu_A(x) \end{cases} \quad (6.44)$$

当 $B' =$ 差不多 B 时, 采用与 $B' =$ 非常 B 同样的方法, 可得

$$\mu_{A \rightarrow C'}(x, z) = \begin{cases} \mu_A(x) & \mu_A(x) < \mu_C(z) \\ \frac{\mu_C^2(z)}{\mu_A(x)} & \mu_A(x) \geq \mu_C(z) \end{cases} \quad (6.45)$$

最后,当 $B' = \bar{B}$ 时,有

$$\mu_{A \rightarrow C'}(x, z) = \sup_{y \in V} \min[\mu_A(x)\mu_B(y), (1 - \mu_B(y))\mu_C(z)] \quad (6.46)$$

其中,当 $y_0 \in V$ 满足 $\mu_A(x)\mu_B(y_0) = (1 - \mu_B(y_0))\mu_C(z)$ 时,即当 $\mu_B(y_0) = \frac{\mu_C(z)}{\mu_A(x) + \mu_C(z)}$ 时可在 y_0 处达到 $\sup_{y \in V} \min$ 值。因此,有

$$\mu_{A \rightarrow C'}(x, z) = \frac{\mu_A(x)\mu_C(z)}{\mu_A(x) + \mu_C(z)} \quad (6.47)$$

6.4 本章小结及导读

本章主要阐述了以下内容:

- 用真值表证明命题的等价;
- 基本推理规则(取式推理,拒式推理,假言推理)及其向模糊命题的推广(广义取式推理,广义拒式推理,广义假言推理);
- 推理合成规则的思想及其应用;
- 在不同含义的规则和前提的典型情况下确定结论的隶属度函数。

Rescher[1969]研究了各种多值逻辑的处理问题。经典逻辑原理向模糊逻辑原理的推广是20世纪70年代在Zadeh[1973],Zadeh[1975]和Zadeh的其他文章中提出来的。推理合成规则也包含在Zadeh的 these 文章中。

6.5 习题

习题 6.1 用真值表法来证明以下各式均为永真公式:

(a)取式推理(6.3);(b)拒式推理(6.4);(c)假言推理(6.5)。

习题 6.2 令 $U = \{x_1, x_2, x_3\}, V = \{y_1, y_2\}$ 。假设已知一条模糊 IF-THEN 规则:“如果 x 为 A ,则 y 为 B ”,其中 $A = 0.5/x_1 + 1/x_2 + 0.6/x_3, B = 1/y_1 + 0.4/y_2$,请在给定“ x 为 A' ”($A' = 0.6/x_1 + 0.9/x_2 + 0.7/x_3$)的条件下,运用广义取式推理(6.10)推出结论“ y 为 B ”。这里模糊关系 $A \rightarrow B$ 可解释为如下含义:

- (a) Dienes-Rescher 含义(5.23);
- (b) Lukasiewicz 含义(5.24);
- (c) Zadeh 含义(5.25);
- (d) Mamdani 乘积含义(5.32)。

习题 6.3 在 $A = 0.6/x_1 + 1/x_2 + 0.9/x_3, B = 0.6/y_1 + 1/y_2, A' = 0.5/x_1 + 0.9/x_2 + 1/x_3$ 的条件下重做习题 6.2。

习题 6.4 令 U, V, A 和 B 与习题 6.2 中的 U, V, A 和 B 相同。请在已知“ y 为 B' ”($B' = 0.9/y_1 + 0.7/y_2$)的条件下,运用广义拒式推理(6.11)推出结论“ x 为 A' ”,这里模糊关系 $A \rightarrow B$ 可解释为如下含义:

(a) Lukasiewicz 含义(5.24);

(b) Mamdani 含义(5.32)。

习题 6.5 用最小(min)算子和 Lukasiewicz 含义(5.24)分别取代广义取式推理(6.10)中的 t -范数和 $\mu_{A \rightarrow B}(x, y)$ 。分别在(a) $A' = A$, (b) $A' =$ 非常 A , (c) $A' =$ 差不多 A , (d) $A' = \bar{A}$ 的情况下确定隶属度函数 $\mu_{B'}(y)$ (以 $\mu_B(y)$ 的形式表示)。

习题 6.6 用最小(min)算子和 Dienes-Rescher 含义(5.23)分别取代广义取式推理(6.10)中的 t -范数和 $\mu_{A \rightarrow B}(x, y)$ 。分别在(a) $A' = A$, (b) $A' =$ 非常 A , (c) $A' =$ 差不多 A , (d) $A' = \bar{A}$ 的情况下确定隶属度函数 $\mu_{B'}(y)$ (以 $\mu_B(y)$ 的形式表示)。

习题 6.7 用最小(min)算子和 Mamdani 含义(5.31)分别取代广义拒式推理(6.11)中的 t -范数和 $\mu_{A \rightarrow B}(x, y)$ 。分别在(a) $B' = \bar{B}$, (b) $B' =$ 不非常 B , (c) $B' =$ 非差不多 B ; (d) $B' = B$ 的情况下确定隶属度函数 $\mu_{A'}(x)$ (以 $\mu_A(x)$ 的形式表示)。

习题 6.8 考虑一个基于标准算子($\min, \max, 1 - a$)的模糊逻辑。对于任意两个命题, A 和 B , 从逻辑上讲, 下面的恒等式是成立的:

$$\overline{A \wedge B} = B \vee (\bar{A} \wedge \bar{B}) \quad (6.48)$$

这一条件意味着 A 和 B 的真值严格受 $[0, 1]^2$ 的一个子集约束。请准确说明 A 和 B 是怎样受约束的。

第二部分

模糊系统及其特性

在第 1 章中已经提到了模糊系统的四个组成部分：模糊规则库、模糊推理机、模糊器和解模糊器(见图 1.5)。在本部分(第 7 章至第 11 章)将详细研究模糊系统的这四个组成部分,将学习如何将第一部分中的模糊数学原理和逻辑原理应用于模糊系统,推导出不同类型模糊系统的严密的数学公式,并研究这些模糊系统的逼近性质。

第 7 章,将分析模糊规则库的结构,并给出一些特定的模糊推理机。第 8 章,将给出一些模糊器和解模糊器,并对其做详细分析。第 9 章,将把第 7 章和第 8 章中给出的模糊推理机、模糊器和解模糊器组合到一起以得到某些特定的模糊系统,并将证明这些模糊系统具有万能逼近性质。第 10 章和第 11 章,将详细研究模糊系统的逼近精度问题,并证明如何设计模糊系统才会使其达到所需精度。

模糊规则库与模糊推理机

考虑图 1.5 所示的模糊系统, 图中 $U = U_1 \times U_2 \times \cdots \times U_n \subset R^n, V \subset R$ 。这里仅考虑多维输入一维输出的情况, 因为一个多维输出的系统总是可以分解成多个一维输出的系统。例如, 若要求设计一个四维输入三维输出的模糊系统, 则可以先分别设计三个四维输入一维输出的模糊系统, 然后将它们像图 7.1 那样组合到一起。

本章将详细研究模糊规则库和模糊推理机, 模糊器和解模糊器将在下一章中研究。

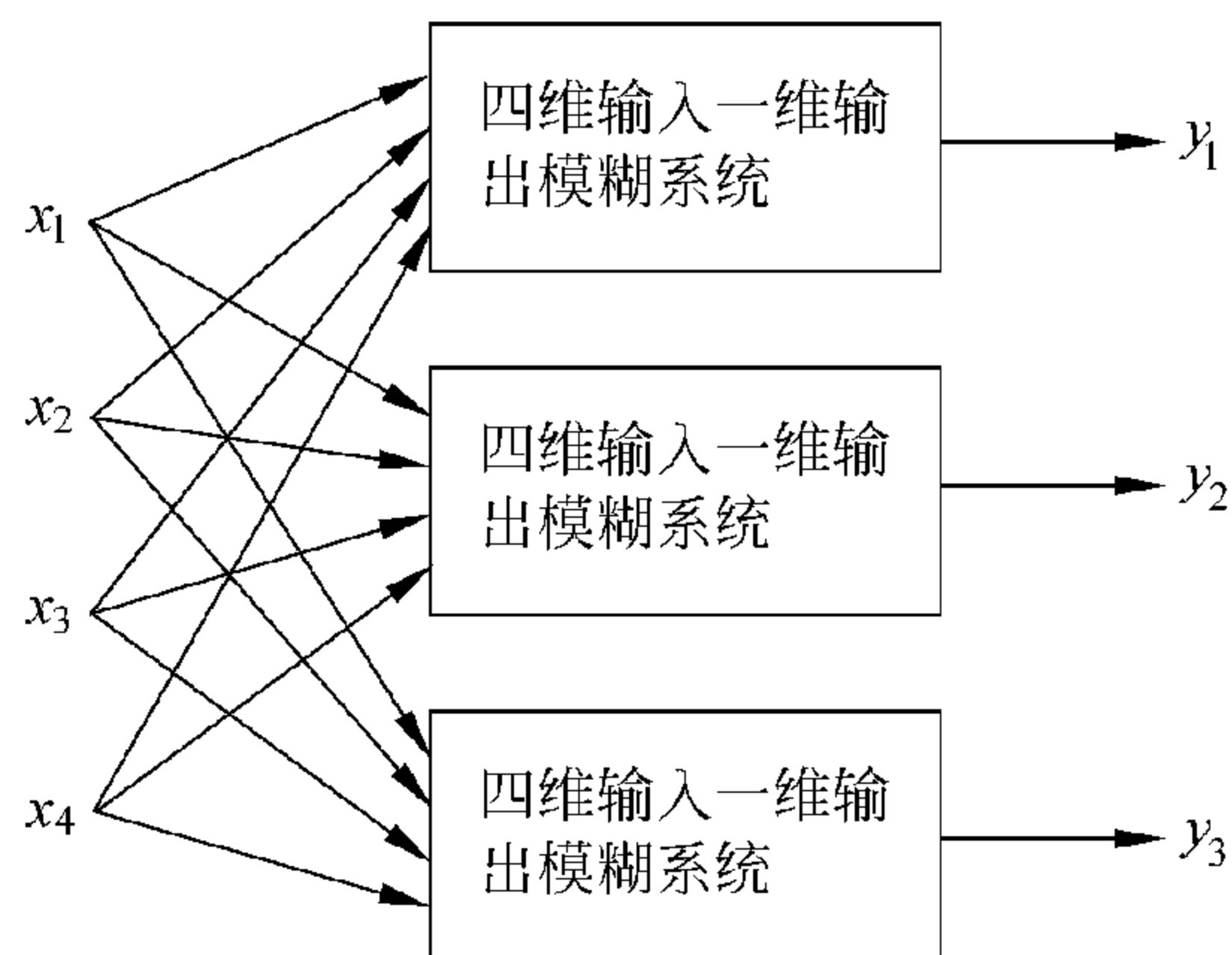


图 7.1 一个多输入多输出模糊系统可以被分解成多个多维输入一维输出模糊系统

7.1 模糊规则库

7.1.1 模糊规则库的结构

模糊规则库是由模糊 IF-THEN 规则集合组成的。它是模糊系统的核心, 从这个意义上讲, 模糊系统的其他组成部分都是以一种合理而有效的方式来执行这些规则的。具体地讲, 模糊规则库是由以下模糊 IF-THEN 规则组成的:

$$R_u^{(l)}: \text{如果 } x_1 \text{ 为 } A_1^l \text{ 且 } \cdots \text{ 且 } x_n \text{ 为 } A_n^l, \text{ 则 } y \text{ 为 } B^l \quad (7.1)$$

其中, A_i^l 和 B^l 分别是 $U_i \subset R$ 和 $V \subset R$ 上的模糊集合, $x = (x_1, x_2, \cdots, x_n)^T \in U$ 和 $y \in V$ 分别是模糊系统的输入和输出(语言)变量。令 M 为模糊规则库式(7.1)中的规则数目, 即 $l = 1, 2, \cdots, M$ 。将形如式(7.1)的规则叫做标准模糊 IF-THEN 规则, 因为正如下面的引理所表明的那样, 它包含了许多其他类型的模糊规则及特殊的模糊命题。

引理 7.1 形如式(7.1)的标准模糊 IF-THEN 规则包含了下列特例:

(a) “不完整规则”

$$\text{如果 } x_1 \text{ 为 } A_1^l \text{ 且 } \cdots \text{ 且 } x_m \text{ 为 } A_m^l, \text{ 则 } y \text{ 为 } B^l \quad (7.2)$$

其中, $m < n$ 。

(b) “或规则”

$$\text{如果 } x_1 \text{ 为 } A_1^l \text{ 且 } \cdots \text{ 且 } x_m \text{ 为 } A_m^l \text{ 或 } x_{m+1} \text{ 为 } A_{m+1}^l \text{ 且 } \cdots \text{ 且 } x_n \text{ 为 } A_n^l, \text{ 则 } y \text{ 为 } B^l \quad (7.3)$$

(c) 单一模糊陈述

$$y \text{ 为 } B^l \quad (7.4)$$

(d) “逐级变化规则”

$$x \text{ 越小, 则 } y \text{ 越大} \quad (7.5)$$

(e) 非模糊规则(即传统的扩展规则)

证明 不完整规则(7.2)等价于

$$\text{如果 } x_1 \text{ 为 } A_1^l \text{ 且 } \cdots \text{ 且 } x_m \text{ 为 } A_m^l \text{ 且 } x_{m+1} \text{ 为 } I \text{ 且 } \cdots \text{ 且 } x_n \text{ 为 } I, \text{ 则 } y \text{ 为 } B^l \quad (7.6)$$

其中, I 是 R 上的模糊集合, 对于所有 $x \in R$ 都有 $\mu_I(x) \equiv 1$, 因为该模糊规则如规则(7.1), 所以(a)得证。根据逻辑算子“或”的直观含义可知, “或规则”(7.3)等价于下面的两条规则:

$$\text{如果 } x_1 \text{ 为 } A_1^l \text{ 且 } \cdots \text{ 且 } x_m \text{ 为 } A_m^l, \text{ 则 } y \text{ 为 } B^l \quad (7.7)$$

$$\text{如果 } x_{m+1} \text{ 为 } A_{m+1}^l \text{ 且 } \cdots \text{ 且 } x_n \text{ 为 } A_n^l, \text{ 则 } y \text{ 为 } B^l \quad (7.8)$$

由(a)可知, 这两条规则(7.7)和(7.8)都是规则(7.1)的特例, 于是(b)得证。模糊陈述(7.4)等价于

$$\text{如果 } x_1 \text{ 为 } I \text{ 且 } \cdots \text{ 且 } x_n \text{ 为 } I, \text{ 则 } y \text{ 为 } B^l \quad (7.9)$$

因为该规则也形如规则(7.1), 所以(c)得证。至于(d), 令 S 表示模糊集“较小”, 如 $\mu_S(x) = 1/(1 + \exp(5(x+2)))$, 令 B 表示模糊集“较大”, 如 $\mu_B(y) = 1/(1 + \exp(-5(y-2)))$, 则“逐级变化规则”(7.5)等价于

$$\text{如果 } x \text{ 为 } S, \text{ 则 } y \text{ 为 } B \quad (7.10)$$

该规则也形如规则(7.1), 于是(d)得证。最后, 当 A_i^l 和 B^l 的隶属度函数仅取值为 1 或 0 时, 规则(7.1)就变成了非模糊规则。

在模糊系统体系中, 人类知识是不得不以模糊 IF-THEN 规则(7.1)的形式来表达的, 也就是说, 只有用模糊 IF-THEN 规则的形式来描述人类知识, 才能利用人类知识。幸运的是, 引理 7.1 确保了这些规则能够提供一个一般性的知识表达式。

7.1.2 规则集合的特性

因为模糊规则库是由规则集合组成的, 所以这些规则之间的关系及规则整体都蕴涵着有趣的问题。如, 这些规则涵盖了模糊系统所有可能面对的情况了吗? 这些规则之间有冲突吗? 为回答这些问题, 引入下列概念。

定义 7.1 如果对任意 $x \in U$, 在模糊规则库中都至少存在一条规则, 称之为规则 $R_u^{(l)}$ [形如规则(7.1)], 对于所有 $i = 1, 2, \dots, n$ 都满足

$$\mu_{A_i^l}(x_i) \neq 0 \quad (7.11)$$

则称这个模糊 IF-THEN 规则集合是完备的。

直观地讲,一个规则集合的完备性表明,输入空间中的任意点上至少存在一条可用规则。也就是说,规则的 IF 部分的隶属度值在这一点上是非零的。

例 7.1 考虑 $U = U_1 \times U_2 = [0,1] \times [0,1]$ 和 $V = [0,1]$ 上的一个二维输入一维输出模糊系统。在 U_1 上定义三个模糊集合 S_1, M_1, L_1 , 在 U_2 上定义两个模糊集合 S_2, L_2 (见图 7.2)。为了保证模糊规则库的完备性,该模糊规则库必须包含以下六条规则,且六条规则的 IF 部分应由 S_1, M_1, L_1 和 S_2, L_2 的所有可能的组合组成:

$$\begin{aligned}
 & \text{如果 } x_1 \text{ 为 } S_1 \text{ 且 } x_2 \text{ 为 } S_2, \text{ 则 } y \text{ 为 } B^1 \\
 & \text{如果 } x_1 \text{ 为 } S_1 \text{ 且 } x_2 \text{ 为 } L_2, \text{ 则 } y \text{ 为 } B^2 \\
 & \text{如果 } x_1 \text{ 为 } M_1 \text{ 且 } x_2 \text{ 为 } S_2, \text{ 则 } y \text{ 为 } B^3 \\
 & \text{如果 } x_1 \text{ 为 } M_1 \text{ 且 } x_2 \text{ 为 } L_2, \text{ 则 } y \text{ 为 } B^4 \\
 & \text{如果 } x_1 \text{ 为 } L_1 \text{ 且 } x_2 \text{ 为 } S_2, \text{ 则 } y \text{ 为 } B^5 \\
 & \text{如果 } x_1 \text{ 为 } L_1 \text{ 且 } x_2 \text{ 为 } L_2, \text{ 则 } y \text{ 为 } B^6
 \end{aligned} \tag{7.12}$$

其中, $B^l (l=1,2,\dots,6)$ 为 V 上的模糊集合。如果这六条规则中缺少任意一条规则,则都会找到一个点 $x^* \in U$, 在这点上,所有余下规则的 IF 部分的命题的隶属度值为零。例如,如果式(7.12)中缺少第二条规则,则 $x^* = (0,1)$ 。(为什么?)

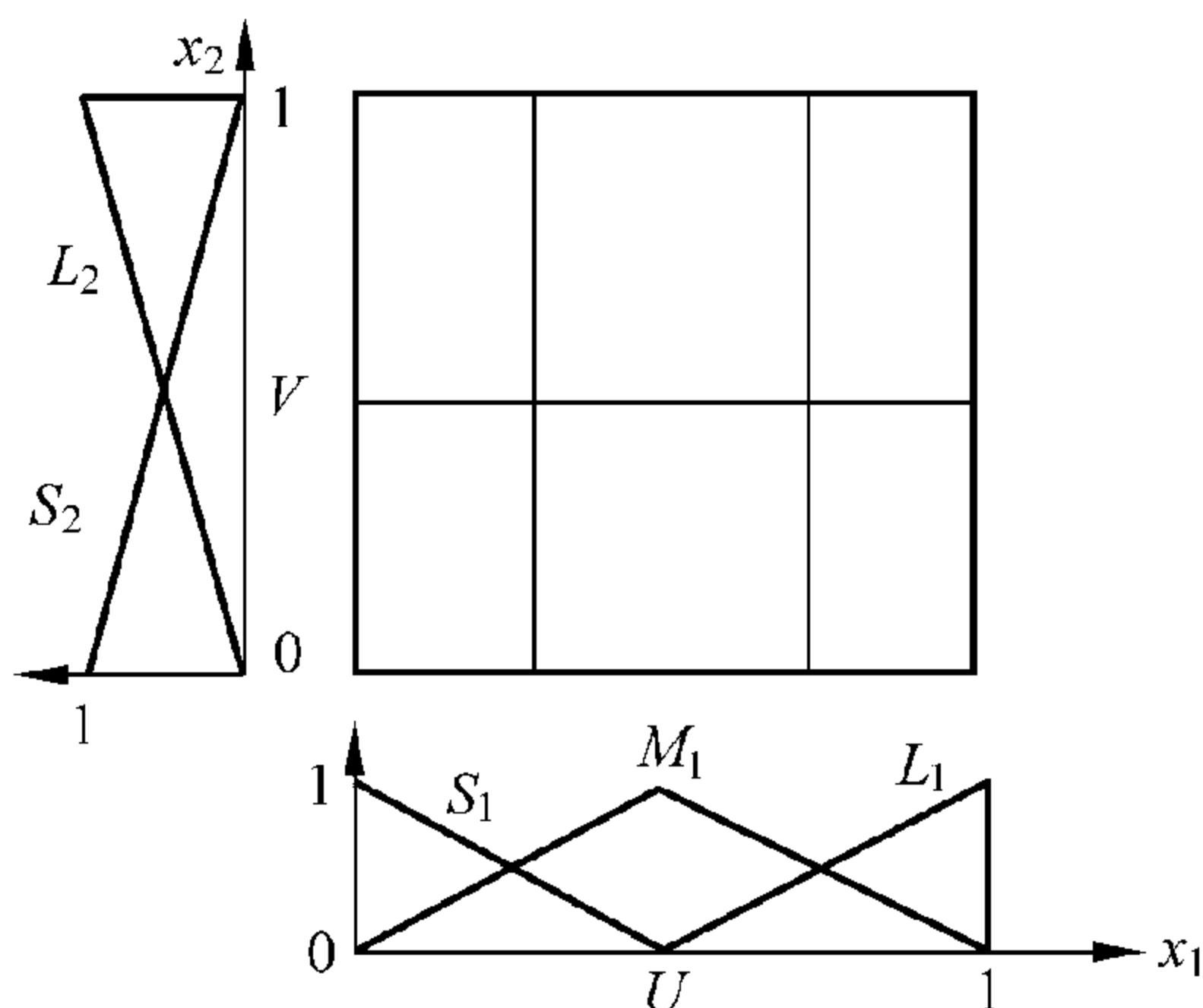


图 7.2 一个二维输入模糊系统的隶属度函数的例子

由例 7.1 可以看出,如果用图 7.2 中所示的三角形隶属度函数,则完备模糊规则库中规则的数目就会随输入空间 U 的维度的上升而呈指数上升,称该问题为“维度灾难”,其进一步讨论见第 22 章。

定义 7.2 如果模糊 IF-THEN 规则集合不存在“IF 部分相同,THEN 部分不同”的规则,则认为这个模糊 IF-THEN 规则集合是一致的。

对于非模糊规则来说,一致性是一条重要的要求,因为如果存在冲突规则,则很难继续搜索。不过后面将会讲到,对于模糊规则来说,一致性并不是关键所在,因为如果存在冲突规则,模糊推理机和解模糊器将会把它们自动平均并输出一个折衷结果。当然,首先最好有一个一致的模糊规则库。

定义 7.3 当邻近规则的 THEN 部分的模糊集的交集不为空集时,称该模糊 IF-THEN 规则集合是连续的。

直观上讲,连续性表明,模糊系统的输入—输出行为应该是平滑的。对此概念很难再做更详细的解释,因为还一直未能推导出模糊系统的完整公式,不过到第9章时,该问题会得到明确地解决。

7.2 模糊推理机

在模糊推理机中,模糊逻辑原理常用于将模糊规则库中的模糊 IF-THEN 规则组合成一个从 U 上的模糊集 A' 到 V 上的模糊集 B' 的映射。第6章曾讲过,模糊 IF-THEN 规则可以解释为输入—输出空间 $U \times V$ 中的一个模糊关系,而且还提出了一些含义来详细说明这些模糊关系。如果模糊规则库仅由一条规则组成,则广义取式推理(6.10)就可指定一个由 U 上的模糊集 A' 向 V 上的模糊集 B' 的映射。因为任意一个现实的模糊规则库都是由一条以上的规则组成的,所以这里的关键问题是,怎样用一系列规则来推理。有两种推理方法:组合推理和独立推理,下面分别进行讨论。

7.2.1 组合推理

在组合推理中,模糊规则库中的所有规则都被组合到 $U \times V$ 中的单一模糊关系中,并将这一模糊关系看做单独的模糊 IF-THEN 规则。因此,关键问题是怎样进行这种组合。首先,应当直观地理解规则的含义是什么,然后再用适当的逻辑算子对其进行组合运算。

关于规则集合的含义有两种对立的观点。第一种观点认为,规则是独立的条件陈述。如果接受了这一观点,则组合规则的一个合理的算子就是“并”。第二种观点认为,规则是具有极强耦合性的陈述,为使整个规则集合具有影响力,须使所有规则的条件都得到满足。如果接受这种观点,则应该用算子“交”来组合规则。第二种观点可能看起来有点奇怪,不过正如将在本节中很快看到的那样,对于某些含义来说,如 Gödel 含义(5.26),它是很有意义的。下面给出了这两种观点的详细内容。

令 $R_u^{(l)}$ 为 $U \times V$ 中的一个模糊关系,用其表示模糊 IF-THEN 规则(7.1)。即, $R_u^{(l)} = A_1^l \times \cdots \times A_n^l \rightarrow B^l$ 。由第5章可知, $A_1^l \times \cdots \times A_n^l$ 是 $U = U_1 \times \cdots \times U_n$ 中的一个模糊关系,其定义为

$$\mu_{A_1^l \times \cdots \times A_n^l}(x_1, \cdots, x_n) = \mu_{A_1^l}(x_1) \star \cdots \star \mu_{A_n^l}(x_n) \quad (7.13)$$

其中, \star 表示任一 t -范数算子。在 $R_u^{(l)}$ 中,含义 \rightarrow 是根据不同的含义(5.23) ~ (5.26), (5.31)和(5.32)定义的。如果接受了规则集合的第一种观点,则式(7.1)的 M 条规则就可

以解释为 $U \times V$ 中的一个独立模糊关系 Q_M , 其定义如下

$$Q_M = \bigcup_{l=1}^M R_{u^{(l)}} \quad (7.14)$$

这一组合称为 Mamdani 组合。如果用符号 $\dot{+}$ 来表示 s -范数, 则式(7.14)变为

$$\mu_{Q_M}(x, y) = \mu_{R_{u^{(1)}}}(x, y) \dot{+} \cdots \dot{+} \mu_{R_{u^{(M)}}}(x, y) \quad (7.15)$$

对于规则集合的第二种观点, 式(7.1)的 M 条模糊 IF-THEN 规则可以解释为 $U \times V$ 中的一个模糊关系 Q_G , 其定义如下

$$Q_G = \bigcap_{l=1}^M R_{u^{(l)}} \quad (7.16)$$

或定义为

$$\mu_{Q_G}(x, y) = \mu_{R_{u^{(1)}}}(x, y) \star \cdots \star \mu_{R_{u^{(M)}}}(x, y) \quad (7.17)$$

其中, \star 表示 t -范数。这一组合称为 Gödel 组合。

令 A' 为 U 上的一个任意模糊集合并将其输入到模糊推理机中, 将 Q_M 或 Q_G 看做一条独立的模糊 IF-THEN 规则并使用广义取式推理(6.10), 可得模糊推理机的输出(使用 Mamdani 组合)为

$$\mu_{B'}(y) = \sup_{x \in U} t[\mu_{A'}(x), \mu_{Q_M}(x, y)] \quad (7.18)$$

或(使用 Gödel 组合)为

$$\mu_{B'}(y) = \sup_{x \in U} t[\mu_{A'}(x), \mu_{Q_G}(x, y)] \quad (7.19)$$

组合推理的计算程序可归纳如下:

步骤 1 对于式(7.1)的 M 条模糊 IF-THEN 规则, 根据式(7.13)确定隶属度函数, $\mu_{A_1^l \times \cdots \times A_n^l}(x_1, \cdots, x_n), l = 1, 2, \cdots, M$;

步骤 2 把 $A_1^l \times \cdots \times A_n^l$ 看做 FP_1 , 把 B^l 看做含义(5.23) ~ (5.26), (5.31) 和 (5.32) 中的 FP_2 , 根据这些含义中的某个含义确定 $\mu_{R_{u^{(l)}}}(x_1, \cdots, x_n, y) = \mu_{A_1^l \times \cdots \times A_n^l \rightarrow B^l}(x_1, \cdots, x_n, y), l = 1, 2, \cdots, M$;

步骤 3 根据式(7.15)或式(7.17)确定 $\mu_{Q_M}(x, y)$ 或 $\mu_{Q_G}(x, y)$;

步骤 4 对于给定的输入 A' , 模糊推理机会根据式(7.18)或式(7.19)输出 B' 。

7.2.2 独立推理

在独立推理中, 模糊规则库中的每条规则都确定一个输出模糊集合, 整个模糊推理机的输出是这 M 个独立模糊集合的组合。这一组合既可用并集得到, 也可用交集得到。

独立推理的运算程序归纳如下:

步骤 1 和 **步骤 2** 同组合推理的步骤 1 和步骤 2;

步骤3 对于 U 上给定的输入模糊集合 A' , 根据广义取式推理(6.10)确定每条规则 $Ru^{(l)}$ 的输出模糊集 B_l , 即当 $l=1, 2, \dots, M$ 时, 有

$$\mu_{B_l}(y) = \sup_{x \in U} t[\mu_{A'}(x), \mu_{Ru^{(l)}}(x, y)] \quad (7.20)$$

步骤4 模糊推理机的输出是 M 个模糊集 $\{B'_1, \dots, B'_M\}$ 的并组合或交组合, 当取并组合时, 有

$$\mu_{B'}(y) = \mu_{B'_1}(y) \dot{+} \dots \dot{+} \mu_{B'_M}(y) \quad (7.21)$$

当取交组合时, 有

$$\mu_{B'}(y) = \mu_{B'_1}(y) \star \dots \star \mu_{B'_M}(y) \quad (7.22)$$

其中, $\dot{+}$ 和 \star 分别表示 s -范数算子和 t -范数算子。

7.2.3 各类推理机

由前两节可知, 对模糊推理机可有多种选择。具体讲, 有以下几种选择:

(i) 组合推理或独立推理, 其中组合推理又包括 Mamdani 组合推理或 Gödel 组合推理;

(ii) Dienes-Rescher 含义(5.23), Lukasiewicz 含义(5.24), Zadeh 含义(5.25), Gödel 含义(5.26)或 Mamdani 含义(5.31)~(5.32);

(iii) 在不同的推理机公式中, 选用不同的 t -范数算子和 s -范数算子。

接下来, 读者很自然就会问: 怎样选择推理机呢?

一般认为, 有如下三条准则:

- 直观力 直观的选择才有意义。例如, 如果由专家给定一个规则集合, 且专家相信这些规则是相互独立的, 则这些规则应该用“并组合”。
- 计算效率高 应选择把 B' 和 A' 关联起来的易于计算的公式。
- 特性 某些选择可能会令推理机具备某些特性, 如果这些特性是有价值的, 则应选择该推理机。

下面详细地给出了在模糊系统与模糊控制中, 最常用的一些模糊推理机的公式。

- 乘积推理机(**Product Inference Engine**)。乘积推理机用到以下推理:

(i) 使用模糊并组合的独立推理(7.21);

(ii) Mamdani 积含义(5.32);

(iii) 所有 t -范数算子都选用代数积算子, 所有 s -范数算子都选用最大(max)算子。

具体地讲, 由式(7.20), 式(7.21), 式(5.32), 式(7.13), 可得如下乘积推理机:

$$\mu_{B'}(y) = \max_{l=1}^M \left[\sup_{X \in U} \left(\mu_{A'}(x) \prod_{i=1}^n \mu_{A'_i}(x_i) \mu_{B'_l}(y) \right) \right] \quad (7.23)$$

即, 给定 U 上一个模糊集合 A' , 乘积推理机可由式(7.23)给出 V 上的模糊集合 B' 。

- 最小推理机(**Minimum Inference Engine**)。最小推理机用到以下推理:
 - (i) 使用模糊并组合的独立推理(7.21);
 - (ii) Mamdani 最小含义(5.31);
 - (iii) 所有 t -范数算子都选用最小(\min)算子,所有 s -范数算子都选用最大(\max)算子。

具体地讲,由式(7.20),式(7.21),式(5.31)和式(7.13),可得

$$\mu_{B'}(y) = \max_{l=1}^M \left[\sup_{x \in U} \min(\mu_{A'}(x), \mu_{A_1'}(x_1), \dots, \mu_{A_n'}(x_n), \mu_{B^l}(y)) \right] \quad (7.24)$$

即,给定 U 上的模糊集 A' ,最小推理机可由式(7.24)给出 V 上的模糊集 B' 。

乘积推理机和最小推理机在模糊系统与模糊控制中都是最常用的推理机。它们的主要优势在于运算的简便性,尤其是乘积推理机。而且,它们对许多现实问题来说更直观,特别是对模糊控制。下面请看乘积推理机和最小推理机的一些特性。

引理 7.2 如果用“带有 Mamdani 组合的组合推理(7.15)”替换“使用并组合的独立推理(7.21)”,则乘积推理机不发生变化。

证明 由式(7.15)和式(7.18),可得

$$\mu_{B'}(y) = \sup_{x \in U} \left[\mu_{A'}(x) \max_{l=1}^M (\mu_{R_{u^l}(D)}(x, y)) \right] \quad (7.25)$$

将式(5.32)和式(7.13)带入式(7.25),则式(7.25)变为

$$\mu_{B'}(y) = \sup_{x \in U} \max_{l=1}^M \left[\mu_{A'}(x) \prod_{i=1}^n \mu_{A_i'}(x_i) \mu_{B^l}(y) \right] \quad (7.26)$$

因为 $\max_{l=1}^M$ 和 $\sup_{x \in U}$ 是可以互换的,所以式(7.26)和式(7.23)是等价的。

引理 7.3 如果模糊集 A' 是一个模糊单值,即如果

$$\mu_{A'}(x) = \begin{cases} 1, & x = x^* \\ 0, & \text{其他} \end{cases} \quad (7.27)$$

其中, x^* 是 U 上的某一点,则乘积推理机可被简化为

$$\mu_{B'}(y) = \max_{l=1}^M \left[\prod_{i=1}^n \mu_{A_i'}(x_i^*) \mu_{B^l}(y) \right] \quad (7.28)$$

最小推理机可被简化为

$$\mu_{B'}(y) = \max_{l=1}^M \left[\min(\mu_{A_1'}(x_1^*), \dots, \mu_{A_n'}(x_n^*), \mu_{B^l}(y)) \right] \quad (7.29)$$

证明 把式(7.27)代入式(7.23)和式(7.24)中,可以看出,在 $x = x^*$ 处可达到 $\sup_{x \in U}$ 。所以,由式(7.23)可推出式(7.28),由式(7.24)可推出式(7.29)。

引理 7.2 说明,尽管独立推理和组合推理在概念上有差别,但它们在某些重要例子中产生的模糊推理机是一样的。引理 7.3 表明,如果输入的是一个模糊单值,则模糊推理机中的计算可以大大简化[在式(7.23)和式(7.24)中最难计算的就是在式(7.28)和式(7.29)中

没有出现的 $\sup_{X \in U}$]

乘积推理机和最小推理机的缺陷在于,如果 $\mu_{A_i}(x_i)$ ($x \in U$)非常小,则由式(7.23)或式(7.24)得到 $\mu_{B'}(y)$ 也将会非常小,这在应用中可能会引发问题。下面是克服这一缺陷的三种模糊推理机:

- **Lukasiewicz** 推理机。Lukasiewicz 推理机用到以下推理:

- (i) 使用模糊交组合的独立推理(7.22);
- (ii) Lukasiewicz 含义(5.24);
- (iii) 所有 t -范数算子都选用最小(min)算子。

具体来讲,由式(7.22),式(7.20),式(5.24)和式(7.13),可得

$$\begin{aligned} \mu_{B'}(y) &= \min_{l=1}^M \left[\sup_{X \in U} \min(\mu_{A'}(x), \mu_{R_{u(l)}}(x, y)) \right] \\ &= \min_{l=1}^M \left[\sup_{X \in U} \min \left[\mu_{A'}(x), \min(1, 1 - \min_{i=1}^n(\mu_{A_i}(x_i)) + \mu_{B'}(y)) \right] \right] \\ &= \min_{l=1}^M \left[\sup_{X \in U} \min \left[\mu_{A'}(x), 1 - \min_{i=1}^n(\mu_{A_i}(x_i)) + \mu_{B'}(y) \right] \right] \end{aligned} \quad (7.30)$$

即,给定 U 上的模糊集 A' , Lukasiewicz 推理机可由式(7.30)给出 V 上的模糊集 B' 。

- **Zadeh** 推理机。Zadeh 推理机用到以下推理:

- (i) 使用基于模糊交组合的独立推理(7.22);
- (ii) Zadeh 含义(5.25);
- (iii) 所有 t -范数算子都选用最小(min)算子。

具体来讲,由式(7.22),式(7.20),式(5.25)和式(7.13),可得

$$\begin{aligned} \mu_{B'}(y) &= \min_{l=1}^M \left\{ \sup_{X \in U} \min \left[\mu_{A'}(x), \max(\min(\mu_{A_1}(x_1), \dots, \mu_{A_n}(x_n)), \mu_{B'}(y)), \right. \right. \\ &\quad \left. \left. 1 - \min_{i=1}^n(\mu_{A_i}(x_i)) \right) \right] \right\} \end{aligned} \quad (7.31)$$

- **Dienes-Rescher** 推理机。Dienes-Rescher 推理机与 Zadeh 推理机的运算方法相同,只不过它把 Zadeh 含义(5.25)替换为 Dienes-Rescher 含义(5.23)。具体来讲,由式(7.22),式(7.20),式(5.23)和式(7.13),可得

$$\mu_{B'}(y) = \min_{l=1}^M \left\{ \sup_{X \in U} \min \left[\mu_{A'}(x), \max \left(1 - \min_{i=1}^n(\mu_{A_i}(x_i)), \mu_{B'}(y) \right) \right] \right\} \quad (7.32)$$

和引理 7.3 类似,可得到以下与 Lukasiewicz 推理机、Zadeh 推理机和 Dienes-Rescher 推理机有关的结论。

引理 7.4 如果 A' 是式(7.27)所定义的一个模糊单值,则 Lukasiewicz 推理机、Zadeh 推理机和 Dienes-Rescher 推理机可以分别简化为

$$\mu_{B'}(y) = \min_{l=1}^M \left[1, 1 - \min_{i=1}^n(\mu_{A_i}(x_i^*)) + \mu_{B'}(y) \right] \quad (7.33)$$

$$\mu_{B'}(y) = \min_{l=1}^M \left\{ \max \left[\min(\mu_{A_1}(x_1^*), \dots, \mu_{A_n}(x_n^*), \mu_{B_l}(y)), 1 - \min_{i=1}^n (\mu_{A_i}(x_i^*)) \right] \right\} \quad (7.34)$$

$$\mu_{B'}(y) = \min_{l=1}^M \left\{ \max \left[1 - \min_{i=1}^n (\mu_{A_i}(x_i^*)), \mu_{B_l}(y) \right] \right\} \quad (7.35)$$

证明 采用与引理 7.3 相同的方法即可证明本引理。

现已给出了五个模糊推理机。下一步将通过两个例子对这些推理机进行比较。

例 7.2 假设一个模糊规则库仅由一条规则组成

$$\text{如果 } x_1 \text{ 为 } A_1 \text{ 且 } \dots \text{ 且 } x_n \text{ 为 } A_n, \text{ 则 } y \text{ 为 } B \quad (7.36)$$

其中

$$\mu_B(y) = \begin{cases} 1 - |y| & -1 \leq y \leq 1 \\ 0 & \text{其他} \end{cases} \quad (7.37)$$

假设 A' 是式(7.27)所定义的一个模糊单值。下面打算将画出由这五个模糊推理机得到的 $\mu_{B'}(y)$ 的图形。令 B'_P, B'_M, B'_L, B'_Z 和 B'_D 分别为由乘积推理机、最小推理机、Lukasiewicz 推理机、Zadeh 推理机和 Dienes-Rescher 推理机得到的模糊集 B' , 并令 $\min[\mu_{A_1}(x_1^*), \dots,$

$\mu_{A_n}(x_n^*)] = \mu_{A_p}(x_p^*)$, $\prod_{i=1}^n \mu_{A_i}(x_i^*) = \mu_A(x^*)$ 。则由式(7.28), 式(7.29)和式(7.33) ~ 式(7.35), 可得

$$\mu_{B'_P}(y) = \mu_A(x^*) \mu_B(y) \quad (7.38)$$

$$\mu_{B'_M}(y) = \min[\mu_{A_p}(x_p^*), \mu_B(y)] \quad (7.39)$$

$$\mu_{B'_L}(y) = \min[1, 1, -\mu_{A_p}(x_p^*) + \mu_B(y)] \quad (7.40)$$

$$\mu_{B'_Z}(y) = \max[\min(\mu_{A_p}(x_p^*), \mu_B(y)), 1 - \mu_{A_p}(x_p^*)] \quad (7.41)$$

$$\mu_{B'_D}(y) = \max[1 - \mu_{A_p}(x_p^*), \mu_B(y)] \quad (7.42)$$

当 $\mu_{A_p}(x_p^*) \geq 0.5$ 时, $\mu_{B'_P}(y)$ 和 $\mu_{B'_M}(y)$ 的图形见图 7.3,

$\mu_{B'_L}(y)$, $\mu_{B'_Z}(y)$ 和 $\mu_{B'_D}(y)$ 的图形见图 7.4。当 $\mu_{A_p}(x_p^*) < 0.5$ 时, $\mu_{B'_P}(y)$ 和 $\mu_{B'_M}(y)$ 的图形见图 7.5, $\mu_{B'_L}(y)$, $\mu_{B'_Z}(y)$ 和 $\mu_{B'_D}(y)$ 的图形见图 7.6。

由图 7.3 ~ 图 7.6, 可有以下发现:

(i) 当点 x^* 处的 IF 部分的隶属度值很小 [如 $\mu_{A_p}(x_p^*) < 0.5$] 时, 乘积推理机和最小推理机会给出非常小的隶属度值, 而 Lukasiewicz 推理机、Zadeh 推理机和 Dienes-Rescher 推理机却会给出非常大的隶属度值;

(ii) 乘积推理机和最小推理机是很相近的,

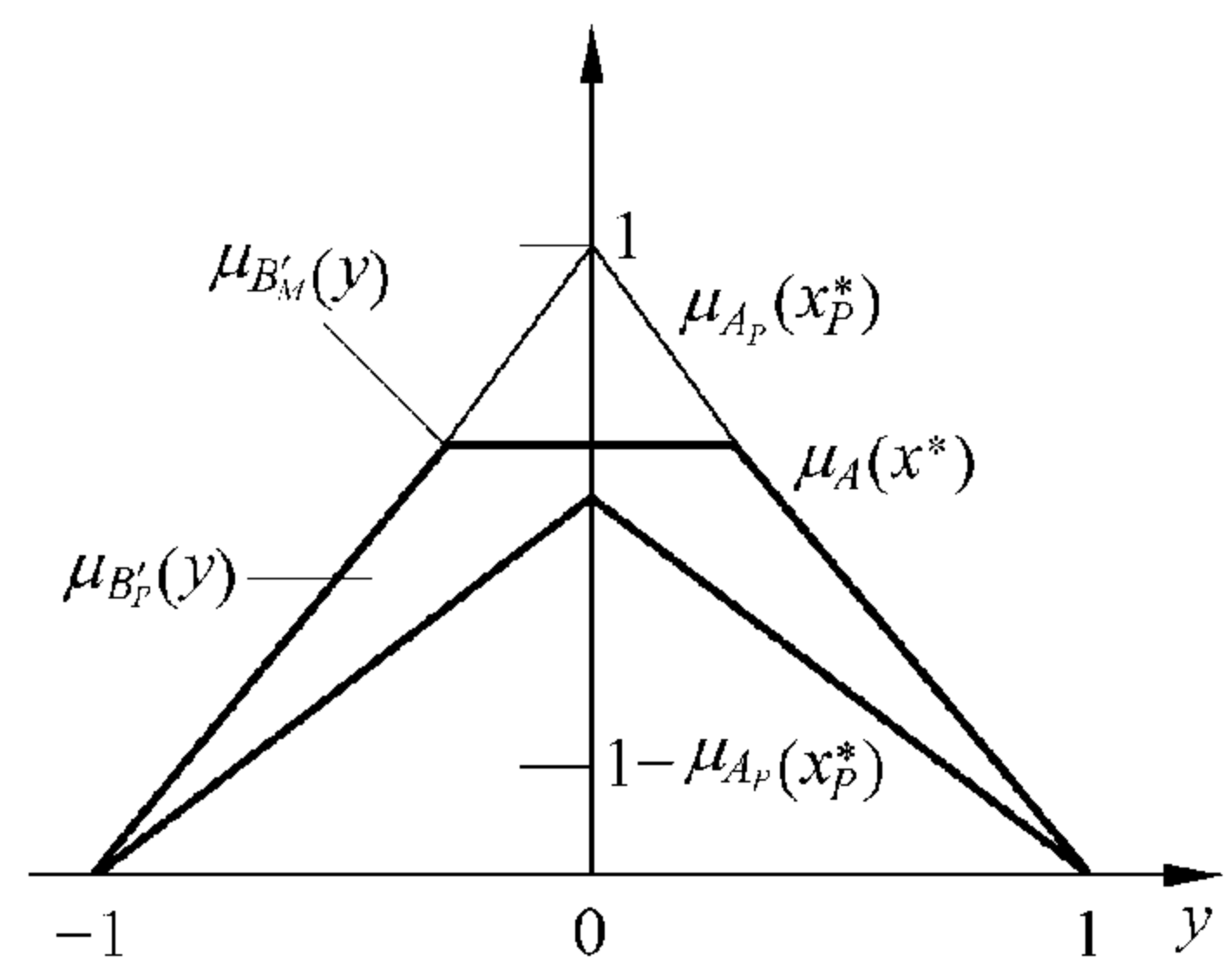


图 7.3 当 $\mu_{A_p}(x_p^*) \geq 0.5$ 时, 由

Lukasiewicz 推理机、Zadeh 推理机和

Dienes-Rescher 推理机输出的隶属度函数

Lukasiewicz 推理机、Zadeh 推理机和 Dienes-Rescher 推理机也是很相近的,但这两个推理机群之间却存在着很大的差异;

(iii) 在所有推理机中, Lukasiewicz 推理机输出的隶属度值最大, 乘积推理机输出的隶属度值最小, 其他三种推理机则介于这两者之间。

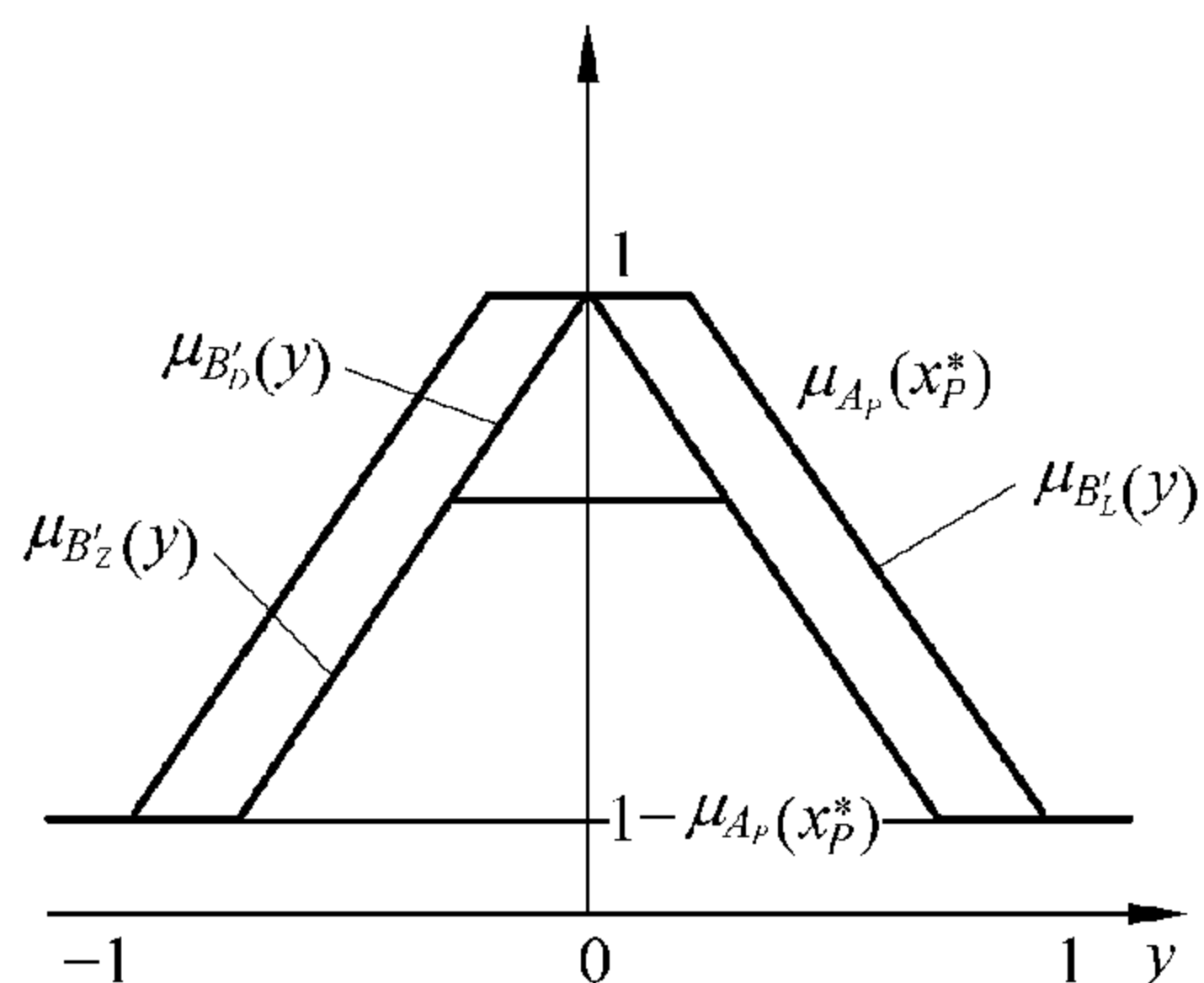


图 7.4 当 $\mu_{A_p}(x_p^*) \geq 0.5$ 时, 由乘积推理机和最小推理机输出的隶属度函数

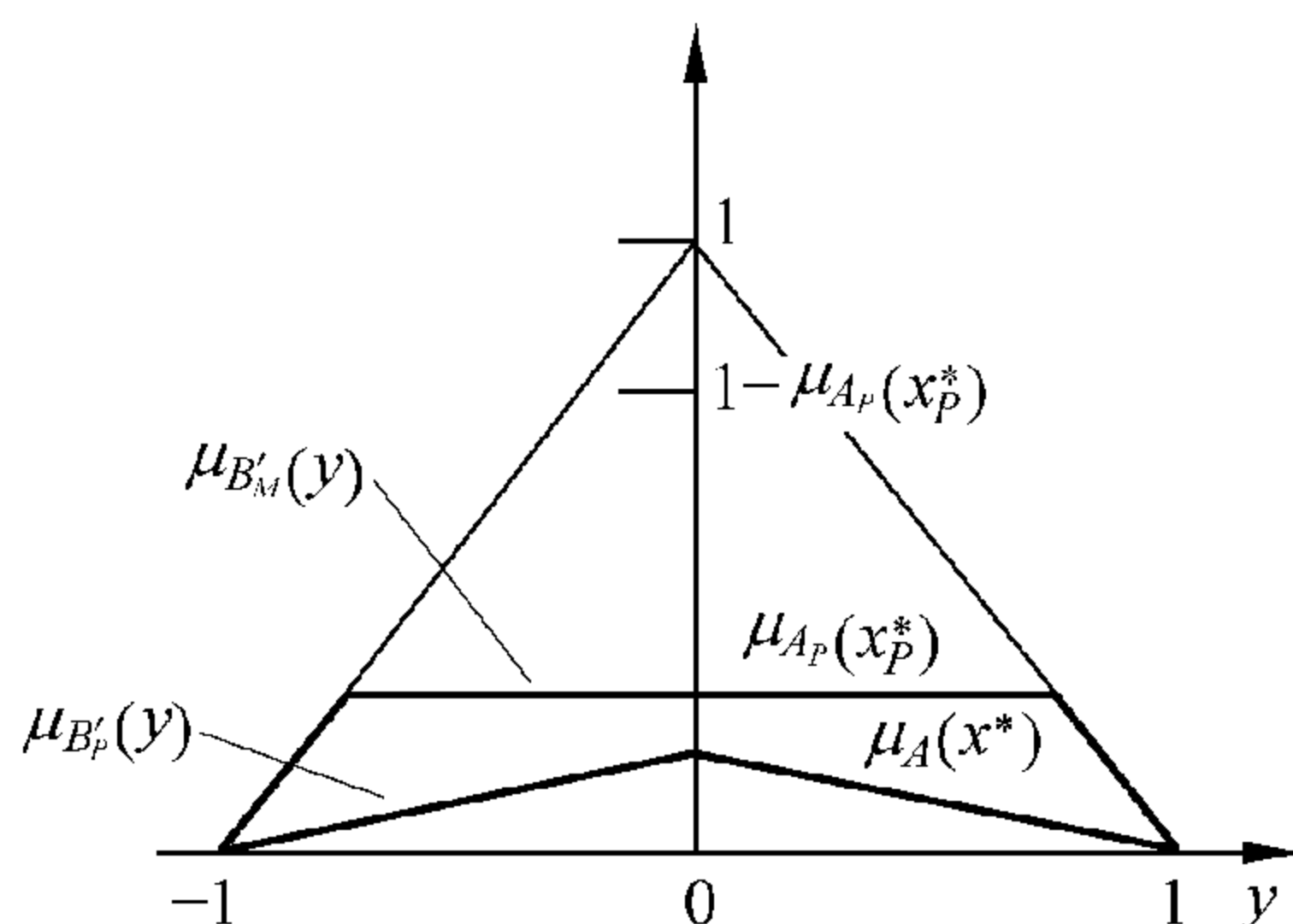


图 7.5 当 $\mu_{A_p}(x_p^*) < 0.5$ 时, 由乘积推理机和最小推理机输出的隶属度函数

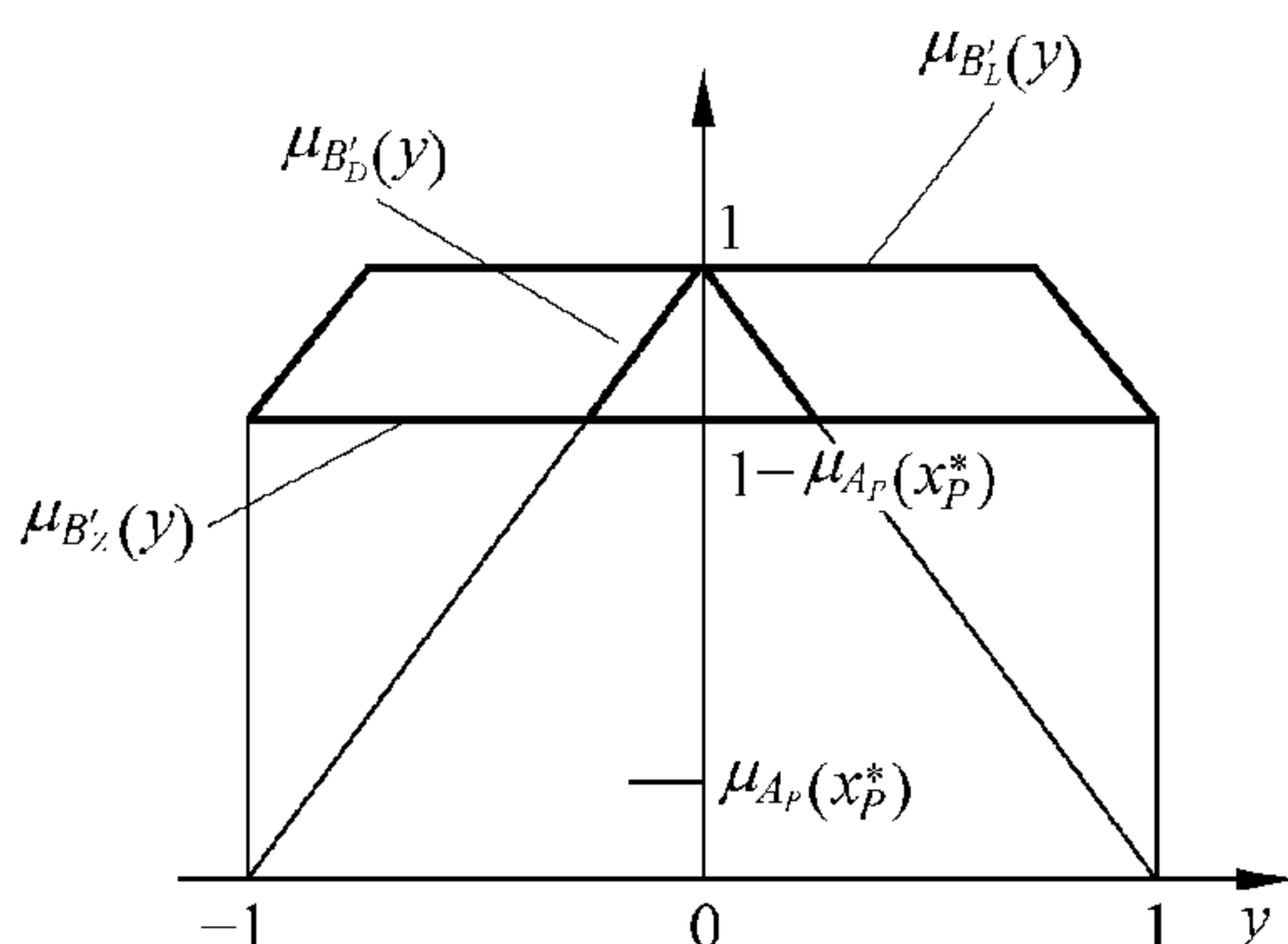


图 7.6 当 $\mu_{A_p}(x_p^*) < 0.5$ 时, 由 Lukasiewicz 推理机、Zadeh 推理机和 Dienes-Rescher 推理机输出的隶属度函数

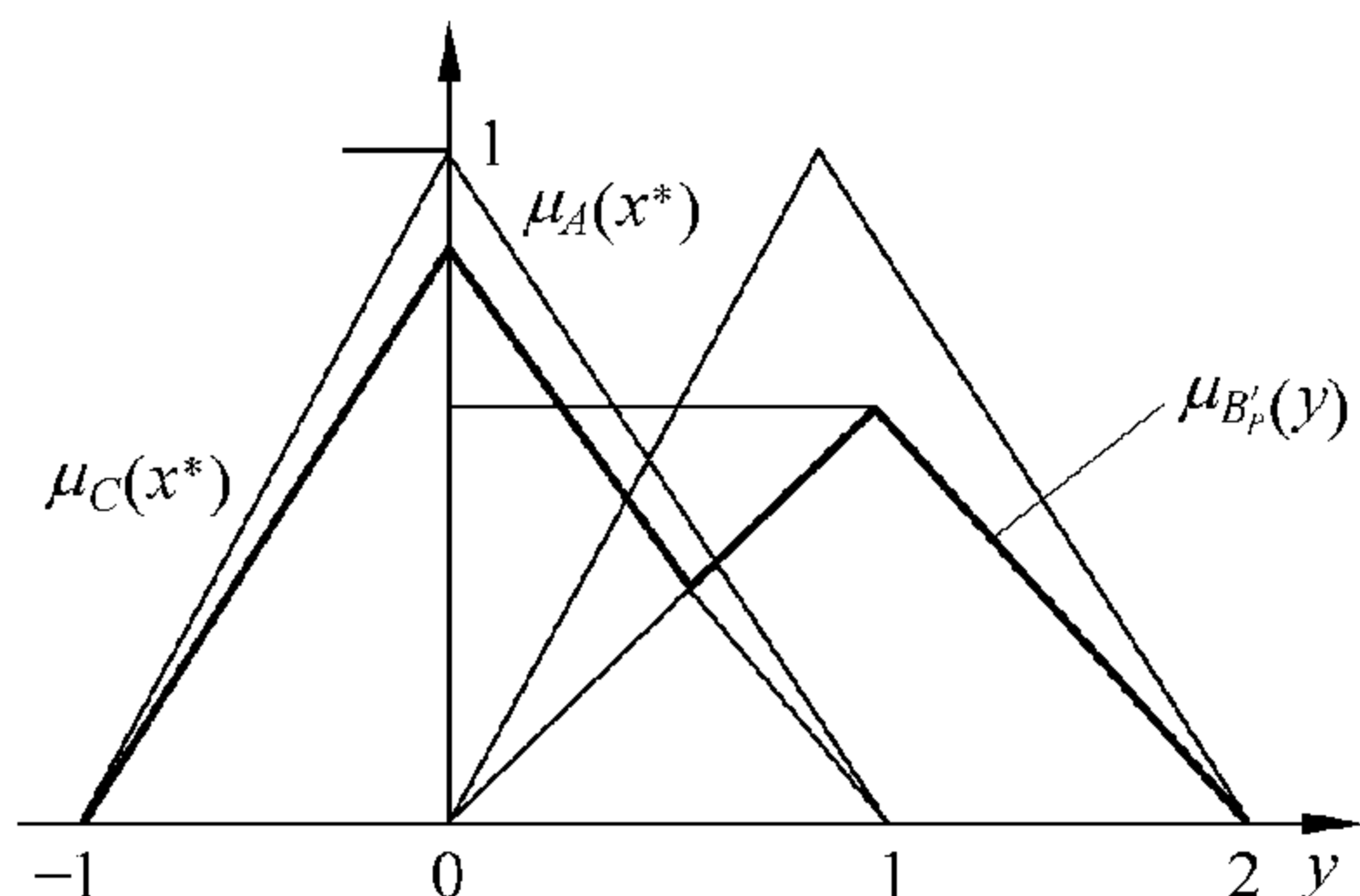


图 7.7 两条规则下, 由乘积推理机输出的隶属度函数

例 7.3 本例中, 考虑包含两条规则的模糊系统, 一条与规则(7.36)一致, 另一条为
如果 x_1 为 C_1 且...且 x_n 为 C_n , 则 y 为 D (7.43)

其中

$$\mu_D(y) = \begin{cases} 1 - |y - 1| & 0 \leq y \leq 2 \\ 0 & \text{其他} \end{cases} \quad (7.44)$$

再次假设 A' 是由式(7.27)定义的模糊单值。可画出由乘积推理机输出的 $\mu_{B_p}(y)$ [即 $\mu_{B_p}(y)$] 的图形, 也就是图 7.7 给出的 $\mu_{B_p}(y)$ 的图形, 图中 $\mu_A(x^*) = \prod_{i=1}^n \mu_{A_i}(x_i^*), \mu_C(x^*)$

$$= \prod_{i=1}^n \mu_{C_i}(x_i^*)。$$

7.3 本章小结及导读

本章主要阐述了以下内容：

- 标准模糊 IF-THEN 规则的结构和评价一个规则集合的准则；
- 组合推理和独立推理的计算程序；
- 五种特定模糊推理机的详细公式：乘积推理机，最小推理机，Lukasiewicz 推理机，Zadeh 推理机和 Dienes-Rescher 推理机，以及它们的特殊算例。

Lee[1990]对模糊规则库和模糊推理机做了非常好的综述，文中给出了各种与模糊规则库和模糊推理机相关的专题的直观分析。Driankov, Hellendoorn 和 Reinfrank[1993]采用和本章类似的方法，对模糊推理机做了数学分析。

7.4 习题

习题 7.1 如果不考虑式(7.12)中的第三条和第六条规则(见例 7.1)，那么所有余下规则的 IF 部分的命题会在哪些点上出现隶属度值？

习题 7.2 给出一个模糊集合 B^1, \dots, B^6 的例子，使其六条规则的集合(7.12)是连续的。

习题 7.3 假设一个模糊规则库仅由一条规则(7.36)组成，其隶属度函数为

$$\mu_B(y) = \exp(-y^2) \quad (7.45)$$

令模糊推理机的输入 A' 为模糊单值(7.27)，分别用(a)乘积推理机，(b)最小推理机，(c)Lukasiewicz 推理机，(d)Zadeh 推理机，(e)Dienes-Rescher 推理机输出 $\mu_{B'}(y)$ ，并画出输出的隶属度函数 $\mu_{B'}(y)$ 的图形。

习题 7.4 考虑例 7.3，分别用：(a)Lukasiewicz 推理机，(b)Zadeh 推理机输出 $\mu_{B'}(y)$ ，并画出输出的隶属度函数 $\mu_{B'}(y)$ 的图形。

习题 7.5 试用 Gödel 含义给出一个所谓的“Gödel 推理机”。

由第 7 章可知,模糊推理机可将模糊规则库中的规则组合成一个由 U 上的模糊集 A' 向 V 上的模糊集 B' 的映射。因为在大多数应用中,模糊系统的输入和输出是实值数,所以,必须在模糊推理机和外部环境之间构建某些界面,此即如图 1.5 所示的模糊器和解模糊器。

8.1 模糊器

模糊器(Fuzzifiers)可定义为由一实值点 $x^* \in U \subset R^n$ 向 U 上的模糊集 A' 的映射。那么设计模糊器的准则是什么呢? 第一条准则,模糊器应该考虑到输入是在清晰点 x^* 处输入的这一事实,即在点 x^* 处模糊集 A' 应该有一个更大的隶属度值; 第二条准则,如果模糊系统的输入受到噪声干扰,则要求模糊器有助于克服噪声的影响; 第三条准则,模糊器应有助于简化模糊推理机的计算。由式(7.23),式(7.24)和式(7.30)~式(7.32)可以看出,在模糊推理机中最复杂的计算是 $\sup_{x \in U}$, 所以,设计模糊器的目标应是简化包含 $\sup_{x \in U}$ 的计算。

现提出三种模糊器:

- 单值模糊器。单值模糊器将一个实值点 $x^* \in U$ 映射成 U 上的一个模糊单值 A' , A' 在 x^* 点上的隶属度值为 1, 在 U 中其他所有点上的隶属度值为 0, 即

$$\mu_{A'}(x) = \begin{cases} 1 & x = x^* \\ 0 & \text{其他} \end{cases} \quad (8.1)$$

- 高斯模糊器。高斯模糊器将 $x^* \in U$ 映射成 U 上的模糊集 A' , 它具有如下高斯隶属度函数:

$$\mu_{A'}(x) = e^{-\left(\frac{x_1 - x_1^*}{a_1}\right)^2} \star \cdots \star e^{-\left(\frac{x_n - x_n^*}{a_n}\right)^2} \quad (8.2)$$

其中,参数 a_i 是正数, t -范数 \star 通常选用代数积算子或最小(min)算子。

- 三角形模糊器。三角形模糊器将 $x^* \in U$ 映射成 U 上的模糊集 A' , 它具有如下三角形隶属度函数:

$$\mu_{A'}(x) = \begin{cases} \left(1 - \frac{|x_1 - x_1^*|}{b_1}\right) \star \cdots \star \left(1 - \frac{|x_n - x_n^*|}{b_n}\right) & |x_i - x_i^*| \leq b_i (i = 1, 2, \cdots, n) \\ 0 & \text{其他} \end{cases} \quad (8.3)$$

其中,参数 $b_i (i=1,2,\dots,n)$ 是正数, t -范数 \star 通常选用代数积算子或最小 (\min) 算子。

由式(8.1)~式(8.3)可以看出,这三种模糊器都满足 $\mu_{A^l}(x^*)=1$, 即,它们满足前面提到的第一条准则。由引理 7.3 和引理 7.4 可以看出,单值模糊器会大大地简化模糊推理机设计中的计算量。下面将证明,如果规则(7.1)中的模糊集 A_i^l 具有高斯隶属度函数或三角形隶属度函数,则高斯模糊器或三角形模糊器也会简化某些模糊推理机设计中的计算量。

引理 8.1 假设模糊规则库由形如式(7.1)的 M 条规则组成,且

$$\mu_{A_i^l}(x_i) = e^{-\left(\frac{x_i - \bar{x}_i^l}{\sigma_i^l}\right)^2} \quad (8.4)$$

式中,参数 \bar{x}_i^l 和 σ_i^l 为常数 ($i=1,2,\dots,n; l=1,2,\dots,M$)。如果选用高斯模糊器(8.2),则有

(a) 若式(8.2)中的 t -范数 \star 选用代数积算子,则乘积推理机(7.23)可以简化为

$$\mu_{B^l}(y) = \max_{l=1}^M \left[\prod_{i=1}^n e^{-\left(\frac{x_i^l - \bar{x}_i^l}{\sigma_i^l}\right)^2} e^{-\left(\frac{x_i^l - x_i^*}{a_i}\right)^2} \mu_{B^l}(y) \right] \quad (8.5)$$

式中

$$x_{iP}^l = \frac{a_i^2 \bar{x}_i^l + (\sigma_i^l)^2 x_i^*}{a_i^2 + (\sigma_i^l)^2} \quad (8.6)$$

(b) 若式(8.2)中的 t -范数 \star 选用最小 (\min) 算子,则最小推理机(7.24)可以简化为

$$\mu_{B^l}(y) = \max_{l=1}^M \left[\min \left(e^{-\left(\frac{x_i^l - \bar{x}_i^l}{\sigma_i^l}\right)^2}, \dots, e^{-\left(\frac{x_i^l - \bar{x}_i^l}{\sigma_i^l}\right)^2}, \mu_{B^l}(y) \right) \right] \quad (8.7)$$

式中

$$x_{iM}^l = \frac{a_i \bar{x}_i^l + \sigma_i^l x_i^*}{a_i + \sigma_i^l} \quad (8.8)$$

证明 (a) 将式(8.2)和式(8.4)代入式(7.23), \star 取为代数积算子,可得

$$\begin{aligned} \mu_{B^l}(y) &= \max_{l=1}^M \left[\sup_{X \in U} \prod_{i=1}^n e^{-\left(\frac{x_i - x_i^*}{a_i}\right)^2} e^{-\left(\frac{x_i - \bar{x}_i^l}{\sigma_i^l}\right)^2} \mu_{B^l}(y) \right] \\ &= \max_{l=1}^M \left[\prod_{i=1}^n \sup_{X \in U} e^{-\left(\frac{x_i - x_i^*}{a_i}\right)^2 - \left(\frac{x_i - \bar{x}_i^l}{\sigma_i^l}\right)^2} \mu_{B^l}(y) \right] \end{aligned} \quad (8.9)$$

因为

$$-\left(\frac{x_i - x_i^*}{a_i}\right)^2 - \left(\frac{x_i - \bar{x}_i^l}{\sigma_i^l}\right)^2 = -k_1 \left(x_i - \frac{a_i^2 \bar{x}_i^l + (\sigma_i^l)^2 x_i^*}{a_i^2 + (\sigma_i^l)^2} \right)^2 + k_2 \quad (8.10)$$

式中, k_1, k_2 不是 x_i 的函数, 所以式(8.9)中的 $\sup_{X \in U}$ 会在 $x_{iP} \in U (i = 1, 2, \dots, n)$ 处得到, 于是有式(8.6)。

(b) 将式(8.2)和式(8.4)代入式(7.24), $\star S$ 取为最小(min)算子, 可得

$$\mu_{B'}(y) = \max_{l=1}^M \left\{ \min \left[\sup_{X \in U} \min \left(e^{-\left(\frac{x_1-x_1^*}{a_1}\right)^2}, e^{-\left(\frac{x_1-\bar{x}_1^l}{\sigma_1^l}\right)^2} \right), \dots, \sup_{X \in U} \min \left(e^{-\left(\frac{x_n-x_n^*}{a_n}\right)^2}, e^{-\left(\frac{x_n-\bar{x}_n^l}{\sigma_n^l}\right)^2} \right), \mu_{B^l}(y) \right] \right\} \quad (8.11)$$

显然, 当下式得到满足时

$$e^{-\left(\frac{x_i-x_i^*}{a_i}\right)^2} = e^{-\left(\frac{x_i-\bar{x}_i^l}{\sigma_i^l}\right)^2} \quad (8.12)$$

可在 $x_i = x_{iM}^l$ 处得到 $\sup_{X \in U} \min$ 值, 于是可得式(8.8)。把 $x_i = x_{iM}^l$ 代入式(8.11)即可得式(8.7)。

当采用三角形模糊器时, 可以得到与引理 8.1 类似的结果。如果 $a_i = 0$, 则由式(8.6)和式(8.8), 有 $x_{iP}^l = x_{iM}^l = x_i^*$ 。即, 在这种情况下, 高斯模糊器变成了单值模糊器。如果 a_i 远大于 σ_i^l , 则由式(8.6)和式(8.8)可以看出, x_{iP}^l 和 x_{iM}^l 将会非常接近 \bar{x}_i^l , 即 x_{iP}^l 和 x_{iM}^l 对于输入 x_i^* 的变化是不敏感的。所以, 只要选择较大的 a_i , 高斯模糊器是可以克服输入 x_i^* 中的噪声的。更具体地讲, 假定输入 x_i^* 受到了噪声干扰, 即

$$x_i^* = x_{i0}^* + n_i^* \quad (8.13)$$

其中, x_{i0}^* 是有用的信号, n_i^* 是噪声。将式(8.13)代入式(8.6), 可得

$$x_{iP}^l = \frac{a_i^2 \bar{x}_i^l + (\sigma_i^l)^2 x_{i0}^*}{a_i^2 + (\sigma_i^l)^2} + \frac{(\sigma_i^l)^2}{a_i^2 + (\sigma_i^l)^2} n_i^* \quad (8.14)$$

由式(8.14)可以看出, 噪声穿过高斯模糊器之后, 会被因子 $\frac{(\sigma_i^l)^2}{a_i^2 + (\sigma_i^l)^2}$ 克服掉。当 a_i 远大于 σ_i^l 时, 噪声将会被大大地克服掉。类似地, 可以证明三角形模糊器也具有这种克服噪声的能力。

总的说来, 关于这三种模糊器有以下几点结论:

- 对于任意可能采用的模糊 IF-THEN 规则的隶属度函数类型, 单值模糊器都可以大大简化模糊推理机的计算。
- 如果模糊 IF-THEN 规则中的隶属度函数分别为高斯隶属度函数或三角形隶属度函数, 则高斯模糊器或三角形模糊器也能简化模糊推理机的计算。
- 高斯模糊器或三角形模糊器能克服输入变量中包含的噪声, 而单值模糊器却不能。

8.2 解模糊器

解模糊器(Defuzzifier)可定义为由 $V \subset R$ 上模糊集 B' (模糊推理机的输出)向清晰点 $y^* \in V$ 的一种映射。从概念上讲,解模糊器的任务是确定一个最能代表模糊集 B' 的 V 上的点。这和一个随机变量的均值是很类似的。不过,由于 B' 是以某种特殊方式(见第7章)构造的,所以在确定这一代表点上有多种选择。下面是在选择解模糊方法中应考虑三条准则:

- 言之有据。点 y^* 可直观地代表 B' ; 例如,它可能位于 B' 的支撑集的中心附近或在 B' 中有很大的隶属度值。
- 计算简便。因为模糊控制器是实时运作的,所以这一准则对模糊控制尤为重要。
- 连续性。 B' 的微小变化不会造成 y^* 的大幅度变动。

下面,提出三种解模糊器。这些解模糊器都假设,模糊集 B' 是由第7章中的五个模糊推理机之一得到的,即 B' 是由式(7.23),式(7.24),式(7.30),式(7.31)或式(7.32)给出的。由这些方程可以看出, B' 是 M 个独立模糊集的模糊并合成或模糊交合成。

8.2.1 重心解模糊器

重心解模糊器(center of gravity Defuzzifier)所确定的 y^* 是 B' 的隶属度函数所涵盖区域的中心。即

$$y^* = \frac{\int_V y \mu_{B'}(y) dy}{\int_V \mu_{B'}(y) dy} \quad (8.15)$$

式中, \int_V 是常规积分。图8.1表明了这一计算过程。

如果将 $\mu_{B'}(y)$ 看做一个随机变量的概率密度函数,则重心解模糊器给出的就是这个随机变量的均值。有时应消去那些在 B' 中的隶属度值太小的 $y \in V$,这使得重心解模糊器变为

$$y^* = \frac{\int_{V_\alpha} y \mu_{B'}(y) dy}{\int_{V_\alpha} \mu_{B'}(y) dy} \quad (8.16)$$

其中, V_α (α 为常数)定义为

$$V_\alpha = \{y \in V \mid \mu_{B'}(y) \geq \alpha\} \quad (8.17)$$

重心解模糊器的优点在于其直观合理,言

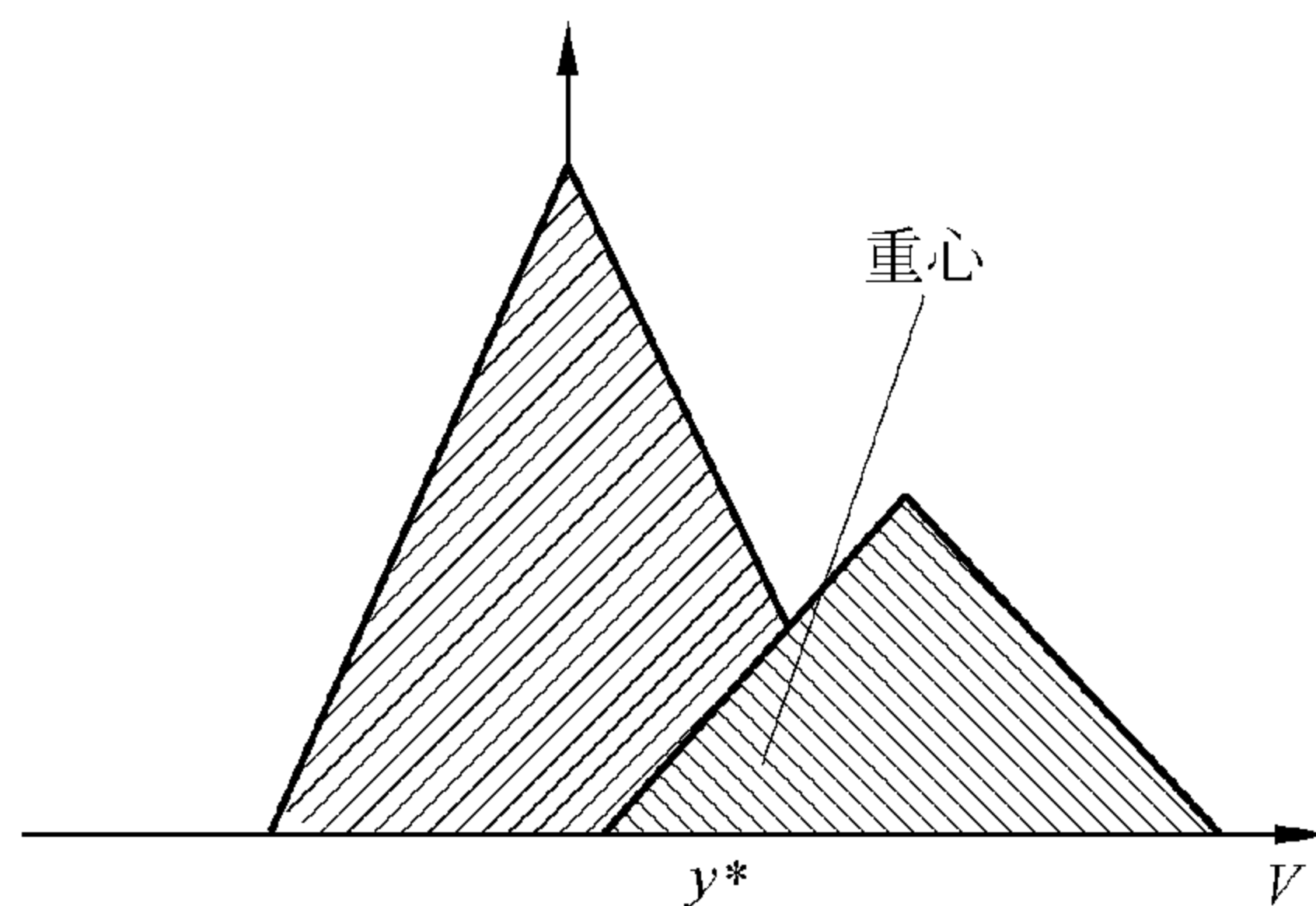


图8.1 重心解模糊器的示意图

之有据。缺点在于其计算要求高。实际上,隶属度函数 $\mu_{B'}(y)$ 通常是不规则的,因而式(8.15)和式(8.16)中的积分很难计算。下一个解模糊器将通过一个简单的公式来逼近式(8.15)从而克服重心解模糊器的这一缺陷。

8.2.2 中心平均解模糊器

由于模糊集 B' 是 M 个模糊集的模糊并合成或模糊交合成,所以式(8.15)的一个好的逼近就是 M 个模糊集中心的加权平均,其权重等于相应模糊集的高度。具体地讲,令 \bar{y}^l 为第 l 个模糊集的中心, w_l 为其高度,则中心平均解模糊器(Center Average Defuzzifier)可由下式确定 y^*

$$y^* = \frac{\sum_{l=1}^M \bar{y}^l w_l}{\sum_{l=1}^M w_l} \quad (8.18)$$

图 8.2 给出了当 $M=2$ 时的一个简单例子,用图形说明了其运算过程。

中心平均解模糊器是在模糊系统与模糊控制中最常用的解模糊器。它计算简便,直观合理。下面用一个简单的例子来比较重心解模糊器和中心平均解模糊器。

例 8.1 假设模糊集 B' 是图 8.2 中所示的两个模糊集的并,图中 $\bar{y}^1 = 0, \bar{y}^2 = 1$, 则由中心平均解模糊器可得

$$y^* = \frac{w_2}{w_1 + w_2} \quad (8.19)$$

下面,计算由重心解模糊器得到的结果。首先注意到,这两个模糊集在 $y = \frac{w_1}{w_1 + w_2}$ 处相交,因此,可得

$$\begin{aligned} \int_V \mu_{B'}(y) dy &= \text{第一个模糊集区域} + \text{第二个模糊集区域} - \text{它们的交集区域} \\ &= w_1 + w_2 - \frac{1}{2} \frac{w_1 w_2}{w_1 + w_2} \end{aligned} \quad (8.20)$$

由图 8.2,可得

$$\int_V y \mu_{B'}(y) dy = \int_{-1}^0 y w_1 (1 + y) dy + \int_0^{\frac{w_1}{w_1 + w_2}} y w_1 (1 - y) dy$$

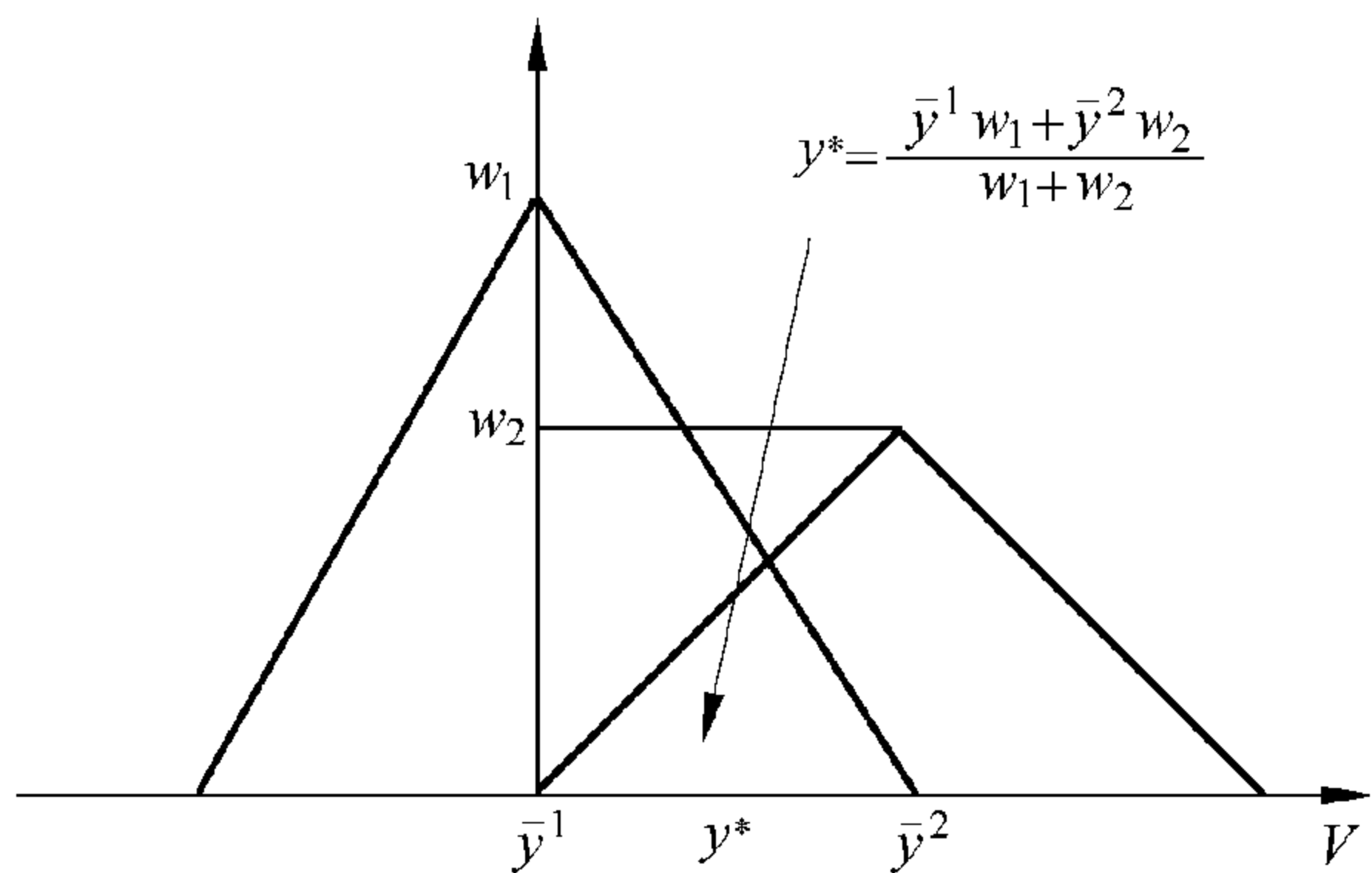


图 8.2 中心平均解模糊器的示意图

$$\begin{aligned}
& + \int_{\frac{w_1}{w_1+w_2}}^1 y w_2 y dy + \int_1^2 y w_2 (2-y) dy \\
& = w_1 \left(\frac{1}{2} y^2 + \frac{1}{3} y^3 \right) \Big|_{-\frac{w_1}{w_1+w_2}}^0 + w_1 \left(\frac{1}{2} y^2 - \frac{1}{3} y^3 \right) \Big|_{\frac{w_1}{w_1+w_2}}^0 + w_2 \frac{1}{3} y^3 \\
& \quad \Big|_{\frac{w_1}{w_1+w_2}}^1 + w_2 \left(y^2 - \frac{1}{3} y^3 \right) \Big|_1^2 \\
& = -\frac{1}{6} w_1 + w_2 + \frac{1}{6} \frac{w_1^3}{(w_1 + w_2)^2} \tag{8.21}
\end{aligned}$$

用式(8.21)除以式(8.20),就得到了重心解模糊器的输出 y^* 。表 8.1 给出了特定的 w_1 和 w_2 值下,采用这两种解模糊器所得到的 y^* 的值。可以看出,重心解模糊器的运算比中心平均解模糊器的运算要复杂得多。

表 8.1 在例 8.1 中重心解模糊器和中心平均解模糊器的比较

w_1	w_2	y^* (重心)	y^* (中心平均)	相对误差
0.9	0.7	0.4258	0.4375	0.0275
0.9	0.5	0.5457	0.5385	0.0133
0.9	0.2	0.7313	0.7000	0.0428
0.6	0.7	0.3324	0.3571	0.0743
0.6	0.5	0.4460	0.4545	0.0192
0.6	0.2	0.6471	0.6250	0.0342
0.3	0.7	0.1477	0.1818	0.2308
0.3	0.5	0.2155	0.2500	0.1600
0.3	0.2	0.3818	0.4000	0.0476

8.2.3 最大值解模糊器

从概念上讲,最大值解模糊器(Maximum Defuzzifier)把 y^* 确定为 V 上 $\mu_{B'}(y)$ 取得其最大值的点。定义集合

$$\text{hgt}(B') = \left\{ y \in V \mid \mu_{B'}(y) = \sup_{y \in V} \mu_{B'}(y) \right\} \tag{8.22}$$

即, $\text{hgt}(B')$ 是 V 上所有 $\mu_{B'}(y)$ 取得其最大值的点的集合。最大值解模糊器定义 y^* 为 $\text{hgt}(B')$ 中的任意元素,即

$$y^* = \text{hgt}(B') \text{ 中的任意一点} \tag{8.23}$$

如果 $\text{hgt}(B')$ 仅包含一个点,则 y^* 惟一确定。如果 $\text{hgt}(B')$ 包含一个以上的点,则仍可采用

式(8.23),或者大中取小,或者大中取大,或者大中取均值,选择其一构造解模糊器。具体来说,大中取小的解模糊器为

$$y^* = \inf \{y \in \text{hgt}(B')\} \quad (8.24)$$

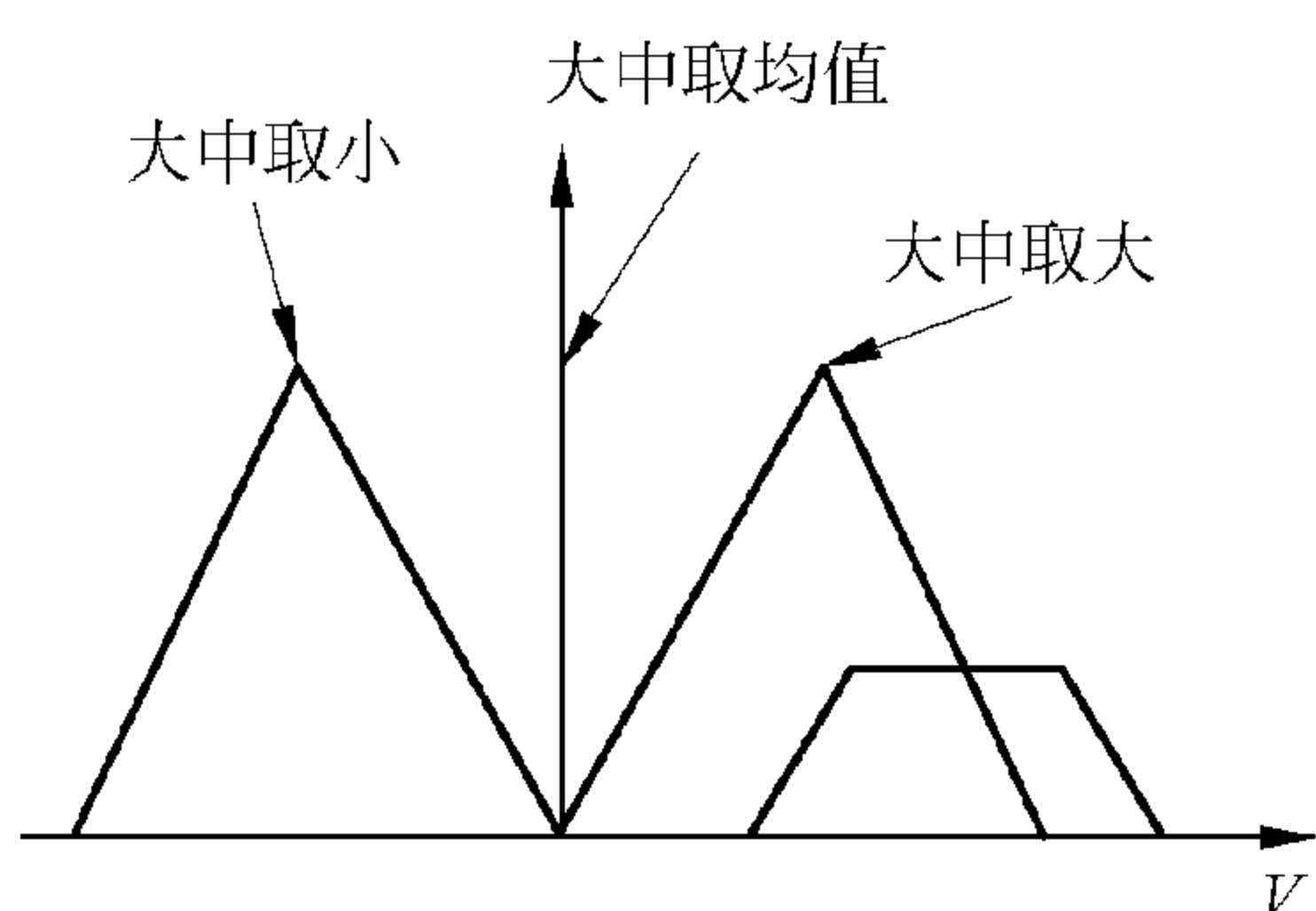
大中取大的解模糊器为

$$y^* = \sup \{y \in \text{hgt}(B')\} \quad (8.25)$$

大中取均值的解模糊器为

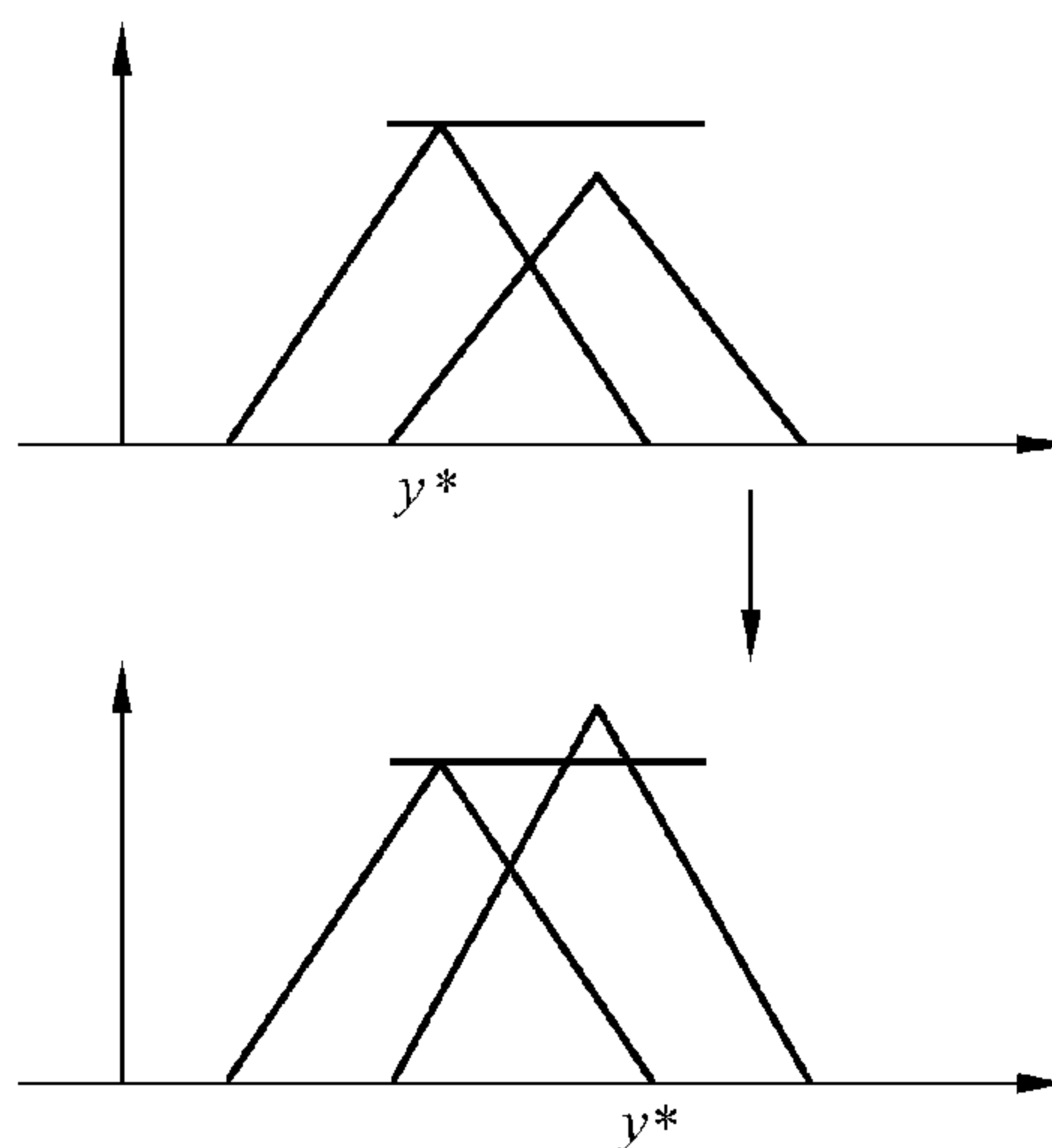
$$y^* = \frac{\int_{\text{hgt}(B')} y dy}{\int_{\text{hgt}(B')} dy} \quad (8.26)$$

其中, $\int_{\text{hgt}(B')}$ 可以是 $\text{hgt}(B')$ 连续部分的常规积分,也可以是 $\text{hgt}(B')$ 离散部分的求和。不过可能会发现,大中取均值的解模糊器给出的结果直观上与最大隶属度有矛盾。如,由大中取均值的解模糊器产生的 y^* 可能在 B' 上有非常小的隶属度值(见图 8.3)。这一问题是由隶属度函数 $\mu_{B'}(y)$ 的非凸性引起的。



注:在此例中,大中取均值的解模糊器给出了一个直观上和最大隶属度相矛盾的结果。

图 8.3 最大值解模糊器的示意图



注:该例中 B' 的微小变化都会造成 y^* 的很大变化。

图 8.4 最大解模糊器的一个例子

最大值解模糊器直观合理且运算简便,但 B' 的微小变化可能会造成 y^* 的很大变化(见图 8.4)。如果图 8.4 中的情况不发生,则最大值解模糊器也是一个不错的选择。

8.2.4 解模糊器的比较

表 8.2 依据言之有据、计算简便和连续性这三条准则来比较这三种解模糊器。由表

8.2 可以看出,中心平均解模糊器是最好的解模糊器。

表 8.2 重心解模糊器、中心平均解模糊器、最大值解模糊器
关于言之有据、计算简便和连续性的比较

	重 心	中 心 平 均	最 大
言之有据	是	是	是
计算简便	否	是	是
连续性	是	是	否

最后请看一个具有某些特定隶属度函数的解模糊器的算例。

例 8.2 设想有一个二维输入一维输出的模糊系统,它是由以下两条规则构造而成的:

$$\text{如果 } x_1 \text{ 为 } A_1 \text{ 且 } x_2 \text{ 为 } A_2, \text{ 则 } y \text{ 为 } A_1 \quad (8.27)$$

$$\text{如果 } x_1 \text{ 为 } A_2 \text{ 且 } x_2 \text{ 为 } A_1, \text{ 则 } y \text{ 为 } A_2 \quad (8.28)$$

其中, A_1 和 A_2 是 R 上的模糊集,其隶属度函数分别为

$$\mu_{A_1}(u) = \begin{cases} 1 - |u| & -1 \leq u \leq 1 \\ 0 & \text{其他} \end{cases} \quad (8.29)$$

$$\mu_{A_2}(u) = \begin{cases} 1 - |u - 1| & 0 \leq u \leq 2 \\ 0 & \text{其他} \end{cases} \quad (8.30)$$

假定模糊系统的输入为 $(x_1^*, x_2^*) = (0.3, 0.6)$, 模糊系统选用单值模糊器。在下述情况下确定模糊系统的输出 y^* : (a) 乘积推理机(7.23)和中心平均解模糊器(8.18); (b) 乘积推理机(7.23)和重心解模糊器(8.15); (c) Lukasiewicz 推理机(7.30)和重心解模糊器(8.26); (d) Lukasiewicz 推理机(7.30)和中心平均解模糊器(8.18)。

(a) 因为选用单值模糊器,所以由引理 7.3 [式(7.28)和式(8.29)~式(8.30)], 可得

$$\begin{aligned} \mu_{B'}(y) &= \max[\mu_{A_1}(0.3)\mu_{A_2}(0.6)\mu_{A_1}(y), \mu_{A_2}(0.3)\mu_{A_1}(0.6)\mu_{A_2}(y)] \\ &= \max[0.7 \times 0.6 \times \mu_{A_1}(y), 0.3 \times 0.4 \times \mu_{A_2}(y)] \\ &= \max[0.42\mu_{A_1}(y), 0.12\mu_{A_2}(y)] \end{aligned} \quad (8.31)$$

其图形见图 8.2, 图中 $\bar{y}^1 = 0, \bar{y}^2 = 1, w_1 = 0.42, w_2 = 0.12$ 。因此,由式(8.19), 可得中心平均解模糊器的输出结果为

$$y^* = \frac{0.12}{0.42 + 0.12} = 0.2222 \quad (8.32)$$

(b) 由 (a) 及式(8.20)~式(8.21), 可得

$$\int_V \mu_{B'}(y) dy = 0.42 + 0.12 - \frac{1}{2} \frac{0.42 \times 0.12}{0.42 + 0.12} = 0.4933 \quad (8.33)$$

$$\int_V y \mu_{B'}(y) dy = -\frac{1}{6} 0.42 + 0.12 + \frac{1}{6} \frac{(0.42)^3}{(0.42 + 0.12)^2} = 0.0923 \quad (8.34)$$

因此,此情况下的 y^* 为

$$y^* = 0.0923 / 0.4933 = 0.1871 \quad (8.35)$$

(c) 如果用 Lukasiewicz 推理机, 则由引理 7.4 [式(7.33)], 可得

$$\begin{aligned} \mu_{B'}(y) &= \min \{1, 1 - \min[\mu_{A_1}(0.3), \mu_{A_2}(0.6)] + \mu_{A_1}(y), 1 - \min[\mu_{A_2}(0.3), \\ &\quad \mu_{A_1}(0.6)] + \mu_{A_2}(y)\} \\ &= \min[1, 0.4 + \mu_{A_1}(y), 0.7 + \mu_{A_2}(y)] \end{aligned} \quad (8.36)$$

其图形见图 8.5。由图 8.5 可以看出, $\sup_{y \in V} \mu_{B'}(y)$ 可在区间 $[0.3, 0.4]$ 中得到, 由大中取均值解模糊器可得

$$y^* = 0.35 \quad (8.37)$$

(d) 由图 8.5 可以看出, 在 $\bar{y}^1 = 0, \bar{y}^2 = 1, w_1 = 1, w_2 = 1$ 的情况下, 中心平均解模糊器 (8.18) 的输出为

$$y^* = \frac{0 \times 1 + 1 \times 1}{1 + 1} = 0.5 \quad (8.38)$$

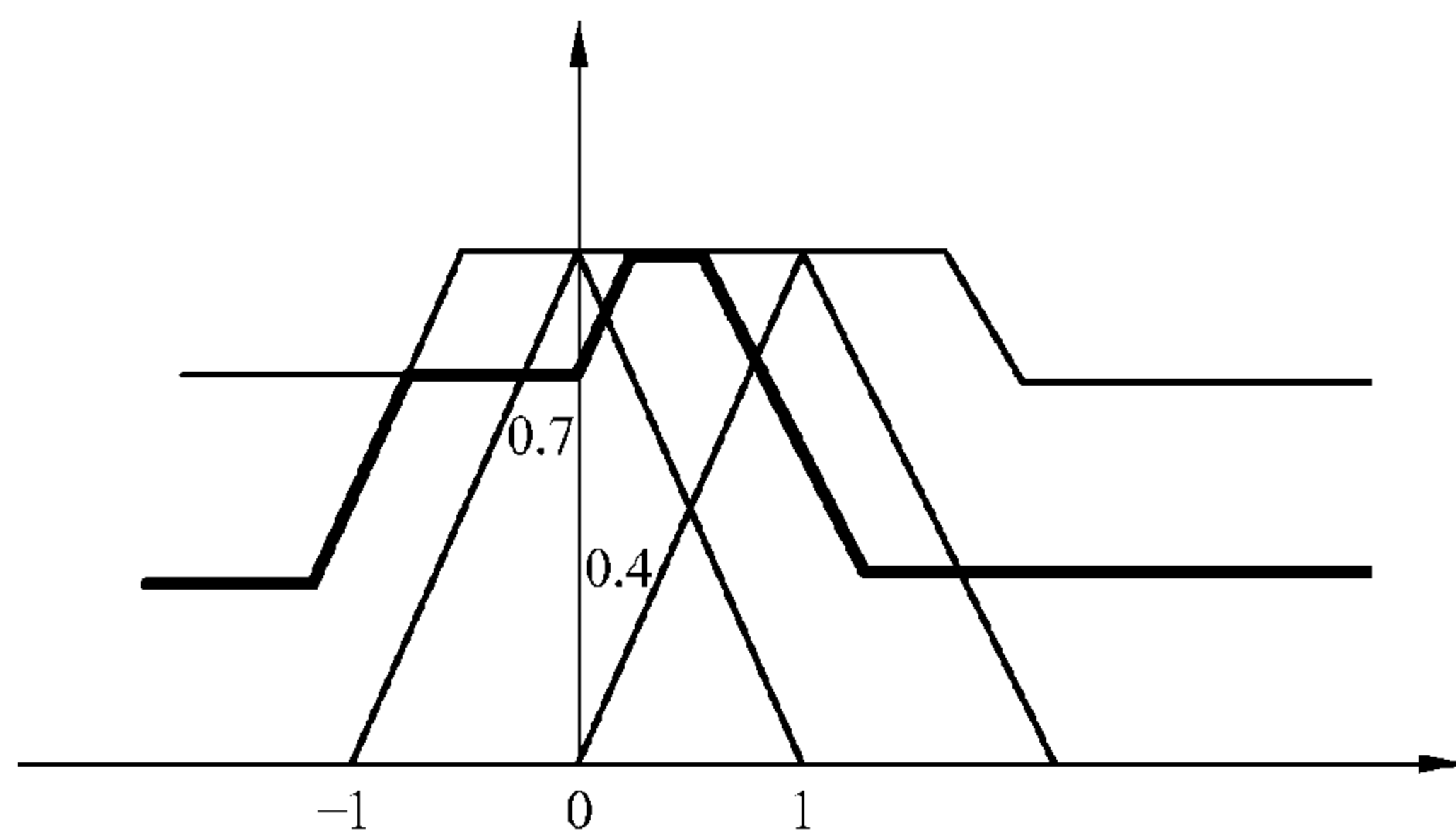


图 8.5 式(8.36)中隶属度函数 $\mu_{B'}(y)$ 的示意图

8.3 本章小结及导读

本章主要阐述了以下内容:

- 单值模糊器、高斯模糊器和三角形模糊器的定义和直观含义, 以及重心解模糊器、中心平均解模糊器和最大值解模糊器的定义和直观含义;
- 在某些特定的模糊器、解模糊器和模糊推理机的不同组合下, 计算模糊系统的输出。

Driankov, Hellendoorn 和 Reifrank [1993] 以及 Yager 和 Filev [1994] 都对不同的解模糊方法进行了详细的研究。

8.4 习题

习题 8.1 假设模糊规则库由形如式(7.1)的 M 条规则组成,其隶属度函数为

$$\mu_{A_i^l}(x_i) = \begin{cases} 1 - \frac{|x_i - \bar{x}_i^l|}{\sigma_i^l} & |x_i - \bar{x}_i^l| \leq \sigma_i^l \\ 0 & \text{其他} \end{cases} \quad (8.39)$$

选用三角形模糊器(8.3),试确定下列情况下模糊推理机的输出 $\mu_{B'}(y)$:

(a) 乘积推理机(7.23),其中所有的 \star = 代数积;

(b) 最小推理机(7.24),其中所有的 \star = 最小。

习题 8.2 考虑例 8.1,用重心解模糊器($\alpha = 0.1$)确定 y^* ,同时计算 y^* 的值(取表 8.1 中的 w_1 和 w_2)。

习题 8.3 考虑例 8.2,模糊系统的输入为 $(x_1^*, x_2^*) = (0.3, 0.6)$,在下列情况下确定模糊系统的输出 y^* :

(a) Zadeh 推理机(7.31)和最大值(大中取均值)解模糊器。

(b) Dienes-Rescher 推理机(7.32)和最大值(大中取均值)解模糊器。

习题 8.4 用一个特例,如移动机器人轨道规划问题,来证明当模糊集 B' 为非凸模糊集时,重心解模糊器和中心平均解模糊器都可能出现问题。

习题 8.5 当模糊集 B' 非凸模糊集时,给出一种叫做最大区域解模糊器中心(center of largest area defuzzifier)的方法。该方法可确定具有最大区域的 B' 的凸集部分,并定义清晰输出 y^* 为这一凸集部分的重心。试建立一个特殊的非凸模糊集 B' ,然后用最大区域解模糊器中心来确定解模糊后的输出 y^* 。

模糊系统作为非线性映射

9.1 几类模糊系统的公式

由第7章和第8章可知,模糊推理机、模糊器和解模糊器都有多种选择。具体地讲,我们提出了五种模糊推理机(乘积推理机、最小推理机、Lukasiewicz推理机、Zadeh推理机和Dienes-Rescher推理机),三种模糊器(单值模糊器、高斯模糊器和三角形模糊器),三种解模糊器(重心解模糊器、中心平均解模糊器、最大值解模糊器)。那么,组合这些不同类型的模糊推理机、模糊器和解模糊器至少可以得到 $5 \times 3 \times 3 = 45$ 种模糊系统。本章将详细推导各类模糊系统的公式,且由这些公式可以看出,某些模糊系统非常有用,而另一些则无太大意义。也就是说,并不是每一个组合都会产生有用的模糊系统。因为已在第8章中证明了,重心解模糊器计算“成本高”,中心平均解模糊器是其很好的近似。所以,本章将不考虑重心解模糊器,而将模糊系统分成两类:带有中心平均解模糊器的模糊系统和带有最大值解模糊器的模糊系统。

9.1.1 带有中心平均解模糊器的模糊系统

引理 9.1 假设规则(7.1)中的模糊集是 B^l 标准模糊集,其中心为 \bar{y}^l ,则带有模糊规则库(7.1)、乘积推理机(7.23)和单值模糊器(8.1)及中心平均解模糊器(8.18)的模糊系统形式为:

$$f(x) = \frac{\sum_{l=1}^M \bar{y}^l \left(\prod_{i=1}^n \mu_{A_i^l}(x_i) \right)}{\sum_{l=1}^M \left(\prod_{i=1}^n \mu_{A_i^l}(x_i) \right)} \quad (9.1)$$

式中, $x \in U \subset R^n$ 是模糊系统的输入, $f(x) \in V \subset R$ 是模糊系统的输出。

证明 将式(8.1)代入式(7.23),可得

$$\mu_{B^l}(y) = \max_{l=1}^M \left[\prod_{i=1}^n \mu_{A_i^l}(x_i^*) \mu_{B^l}(y) \right] \quad (9.2)$$

因为对于给定的输入 x_i^* ,式(9.2)中第 l 个模糊集[即隶属度函数为 $\mu_{A_i^l}(x_i^*) \mu_{B^l}(y)$ 的模糊集]的中心是 B^l 的中心,所以,式(8.18)中的 \bar{y}^l 和本引理中的 \bar{y}^l 是相同的。而且,式

(9.2)中第 l 个模糊集的高度[即式(8.18)中的 w_l]为 $\prod_{i=1}^n \mu_{A_i^l}(x_i^*) \mu_{B^l}(\bar{y}^l) = \prod_{i=1}^n \mu_{A_i^l}(x_i^*)$ (因为 B^l 是标准模糊集)。因此,将式(9.2)代入中心平均解模糊器(8.18),可得

$$y^* = \frac{\sum_{l=1}^M \bar{y}^l \left(\prod_{i=1}^n \mu_{A_i^l}(x_i^*) \right)}{\sum_{l=1}^M \left(\prod_{i=1}^n \mu_{A_i^l}(x_i^*) \right)} \quad (9.3)$$

采用本引理中的符号,令 $x^* = x, y^* = f(x)$,于是式(9.3)变为式(9.1)。

由引理9.1可知,模糊系统是一个由 $x \in U \subset R^n$ 到 $f(x) \in V \subset R$ 的非线性映射,式(9.1)给出了这一映射的详细公式。形如式(9.1)的模糊系统是文献中最为常用的模糊系统,这种模糊系统计算简便且富有直观吸引力。由式(9.1)可以看出,模糊系统的输出就是规则 THEN 部分的模糊集的中心加权平均,其权重等于规则 IF 部分的模糊集在输入点上的隶属度值。总之,输入点与一条规则的 IF 部分的一致性程度越高,赋给这条规则的权重就越大;这也是有其直观意义的。

引理9.1也揭示了模糊系统理论的一个重大贡献,现归纳如下:

- 模糊系统的双重角色:一方面,模糊系统是基于规则库的系统,它是由一系列语言规则构造而成的;另一方面,模糊系统又是非线性映射,在许多情况下,它可以用形如式(9.1)那样的准确而严密的公式来表达。模糊系统理论的重要贡献在于,它提供了一个把语言规则集合转变成非线性映射的系统化程序。由于非线性映射易于实现,所以模糊系统也就找到其转化成各种工程应用的方式。

通过选择隶属度函数 $\mu_{A_i^l}$ 和 μ_{B^l} 的不同形式,可以得到不同子类型的带有中心平均解模糊器的模糊系统。 $\mu_{A_i^l}$ 和 μ_{B^l} 的一种选择是高斯隶属度函数。具体地讲,若选择 $\mu_{A_i^l}$ 和 μ_{B^l} 为如下高斯隶属度函数

$$\mu_{A_i^l}(x_i) = a_i^l \exp\left[-\left(\frac{x_i - \bar{x}_i^l}{\sigma_i^l}\right)^2\right] \quad (9.4)$$

$$\mu_{B^l}(y) = \exp\left[-(y - \bar{y}^l)^2\right] \quad (9.5)$$

式中, $a_i^l \in (0, 1]$, $\sigma_i^l \in (0, \infty)$, $\bar{x}_i^l, \bar{y}^l \in R$ 均为实值参数,则引理9.1中的模糊系统变为

$$f(x) = \frac{\sum_{l=1}^M \bar{y}^l \left[\prod_{i=1}^n a_i^l \exp\left(-\left(\frac{x_i - \bar{x}_i^l}{\sigma_i^l}\right)^2\right) \right]}{\sum_{l=1}^M \left[\prod_{i=1}^n a_i^l \exp\left(-\left(\frac{x_i - \bar{x}_i^l}{\sigma_i^l}\right)^2\right) \right]} \quad (9.6)$$

形如式(9.6)的模糊系统叫做带有乘积推理机、单值模糊器、中心平均解模糊器和高斯隶属度函数的模糊系统。也常选择 $\mu_{A_i^l}$ 和 μ_{B^l} 为三角形和梯形隶属度函数,对此将在第10章和第11章中进行详细研究。

另一类常用的模糊系统是,将引理9.1中的乘积推理机替换为最小推理机而得到的模

糊系统。采用与引理 9.1 中相同的证明过程,即可得到带有模糊规则库(7.1),最小推理机(7.24),单值模糊器(8.1)和中心平均解模糊器(8.18)的模糊系统,其形式为

$$f(x) = \frac{\sum_{l=1}^M \bar{y}^l \left(\min_{i=1}^n \mu_{A_i^l}(x_i) \right)}{\sum_{l=1}^M \left(\min_{i=1}^n \mu_{A_i^l}(x_i) \right)} \quad (9.7)$$

式中,变量的含义与式(9.1)中相同。

在第8章(引理 8.1)中已经证明了,如果 A_i^l 的隶属度函数是高斯隶属度函数,则高斯模糊器会明显地简化模糊推理机的计算。那么,在此情况下模糊系统是什么形式呢?请看下面的引理。

引理 9.2 带有模糊规则库(7.1)、乘积推理机(7.23)、高斯模糊器(8.2)(其中 $\star =$ 代数积)和中心平均解模糊器(8.18)以及高斯隶属度函数(9.4)和(9.5)(其中 $a_i^l = 1$)的模糊系统形式为

$$f(x) = \frac{\sum_{l=1}^M \bar{y}^l \left[\prod_{i=1}^n \exp\left(-\frac{(x_i - \bar{x}_i^l)^2}{a_i^2 + (\sigma_i^l)^2}\right) \right]}{\sum_{l=1}^M \left[\prod_{i=1}^n \exp\left(-\frac{(x_i - \bar{x}_i^l)^2}{a_i^2 + (\sigma_i^l)^2}\right) \right]} \quad (9.8)$$

如果用最小推理机(7.24)来代替乘积推理机(7.23),且式(8.2)中的 $\star = \min$,则模糊系统变为

$$f(x) = \frac{\sum_{l=1}^M \bar{y}^l \left[\min_{i=1}^n \exp\left(-\frac{(x_i - \bar{x}_i^l)^2}{a_i + \sigma_i^l}\right) \right]}{\sum_{l=1}^M \left[\min_{i=1}^n \exp\left(-\frac{(x_i - \bar{x}_i^l)^2}{a_i + \sigma_i^l}\right) \right]} \quad (9.9)$$

证明 把式(8.6)代入式(8.5),并用 x 代替 x^* ,可得

$$\begin{aligned} \mu_{B^l}(y) &= \max_{l=1}^M \left\{ \prod_{i=1}^n \exp\left[-\frac{\left(\frac{a_i^2 \bar{x}_i^l + (\sigma_i^l)^2 x_i}{a_i^2 + (\sigma_i^l)^2} - \bar{x}_i^l\right)^2}{\sigma_i^l} - \frac{\left(\frac{a_i^2 \bar{x}_i^l + (\sigma_i^l)^2 x_i}{a_i^2 + (\sigma_i^l)^2} - x_i\right)^2}{a_i}\right] \mu_{B_i^l}(y) \right\} \\ &= \max_{l=1}^M \left\{ \prod_{i=1}^n \exp\left[-\frac{(x_i - \bar{x}_i^l)^2}{a_i + \sigma_i^l}\right] \mu_{B_i^l}(y) \right\} \end{aligned} \quad (9.10)$$

采用与引理 9.1 相同的证明过程,并将中心平均解模糊器(8.18)代入式(9.10),即可得到式(9.8)。同理,将式(8.8)代入式(8.7),可得

$$\mu_{B^l}(y) = \max_{l=1}^M \left\{ \min \left[\exp\left(-\frac{\left(\frac{a_1 \bar{x}_1^l + \sigma_1^l x_1}{a_1 + \sigma_1^l} - \bar{x}_1^l\right)^2}{\sigma_1^l}\right), \dots, \exp\left(-\frac{\left(\frac{a_n \bar{x}_n^l + \sigma_n^l x_n}{a_n + \sigma_n^l} - \bar{x}_n^l\right)^2}{\sigma_n^l}\right) \right], \mu_{B_i^l}(y) \right\}$$

$$= \max_{l=1}^M \left\{ \min \left[\exp \left(- \left(\frac{x_1 - \bar{x}_1^l}{a_1 + \sigma_1^l} \right)^2 \right), \dots, \exp \left(- \left(\frac{x_n - \bar{x}_n^l}{a_n + \sigma_n^l} \right)^2 \right), \mu_{B^l}(y) \right] \right\} \quad (9.11)$$

再将中心平均解模糊器(8.18)代入式(9.11)中,即可得式(9.9)。

由第7章可知,乘积推理机和最小推理机,与Lukasiewicz推理机、Zadeh推理机及Dienes-Rescher推理机有很大的不同。那么带有Lukasiewicz推理机、Zadeh推理机以及Dienes-Rescher推理机的模糊系统是什么形式呢?请看下面的引理。

引理 9.3 如果规则(7.1)中的模糊集 B^l 是标准模糊集,其中心为 \bar{y}^l ,则带有模糊规则库(7.1)、Lukasiewicz推理机(7.30)或Dienes-Rescher推理机(7.32)、单值模糊器(8.1)或高斯模糊器(8.2)或三角形模糊器(8.3)以及中心平均解模糊器(8.18)的模糊系统形式为

$$f(x) = \frac{1}{M} \sum_{l=1}^M \bar{y}^l \quad (9.12)$$

证明 由 $\mu_{B^l}(\bar{y}^l) = 1$,可得 $1 - \min_{i=1}^n (\mu_{A_i^l}(x_i)) + \mu_{B^l}(\bar{y}^l) \geq 1$ 。所以,式(7.30)中的第 l 个模糊集的高度为

$$\begin{aligned} w_l &= \sup_{x \in U} \min [\mu_{A'}(x), 1 - \min_{i=1}^n (\mu_{A_i^l}(x_i)) + \mu_{B^l}(\bar{y}^l)] \\ &= \sup_{x \in U} \mu_{A'}(x) \\ &= 1 \end{aligned} \quad (9.13)$$

上式用到了所有这三种模糊器式(8.1)~式(8.3)的 $\sup_{x \in U} \mu_{A'}(x) = 1$ 这个事实。类似地,式(7.32)中的第 l 个模糊集的高度也等于1。再应用中心平均解模糊器(8.18)即可得式(9.12)。

因为不管模糊系统输入的是什么,形如式(9.12)的模糊系统给出的都是一个常量输出,所以,研究该模糊系统没有太大的意义。因此,引理9.3中的模糊推理机、模糊器、解模糊器的组合并未产生一个有用的模糊系统。

9.1.2 带有最大值解模糊器的模糊系统

引理 9.4 假设规则(7.1)中的模糊集 B^l 是标准模糊集,其中心为 \bar{y}^l ,则带有模糊规则库(7.1)、乘积推理机(7.23)、单值模糊器(8.1)和最大值解模糊器(8.23)的模糊系统形式为

$$f(x) = \bar{y}^{l^*} \quad (9.14)$$

式中, $l^* \in \{1, 2, \dots, M\}$,它对所有的 $l = 1, 2, \dots, M$ 都满足

$$\prod_{i=1}^n \mu_{A_i^{l^*}}(x_i) \geq \prod_{i=1}^n \mu_{A_i^l}(x_i) \quad (9.15)$$

证明 由式(7.28)(引理7.3)及 $x^* = x$,可得

$$\sup_{y \in V} \mu_{B^l}(y) = \sup_{y \in V} \max_{l=1}^M \left[\prod_{i=1}^n \mu_{A_i^l}(x_i) \mu_{B^l}(y) \right] \quad (9.16)$$

由于 $\sup_{y \in V}$ 与 $\max_{l=1}^M$ 可以互换, 且 B^l 为标准模糊集, 所以有

$$\begin{aligned} \sup_{y \in V} \mu_{B^l}(y) &= \max_{l=1}^M \left[\sup_{y \in V} \prod_{i=1}^n \mu_{A_i^l}(x_i) \mu_{B^l}(y) \right] \\ &= \max_{l=1}^M \left[\prod_{i=1}^n \mu_{A_i^l}(x_i) \right] \\ &= \prod_{i=1}^n \mu_{A_i^{l^*}}(x_i) \end{aligned} \quad (9.17)$$

式中, l^* 是根据式(9.15)定义的。当 $l \neq l^*$ 且 $\mu_{B^l}(\bar{y}^{l^*}) = 1$ 时, 可得 $\mu_{B^l}(\bar{y}^{l^*}) \leq 1$, 由此可得

$$\mu_{B^l}(\bar{y}^{l^*}) = \max_{l=1}^M \left[\prod_{i=1}^n \mu_{A_i^l}(x_i) \mu_{B^l}(\bar{y}^{l^*}) \right] = \prod_{i=1}^n \mu_{A_i^{l^*}}(x_i) \quad (9.18)$$

因此, 可在 \bar{y}^{l^*} 处得到式(9.16)中的 $\sup_{y \in V}$ 值。再应用最大值解模糊器(8.23), 即可得到式(9.14)。

由引理 9.4 可以看出, 在这种情况下的模糊系统是一个简单的函数——分段常值函数, 并且这些常值都是规则 THEN 部分模糊集的隶属度函数的中心。由式(9.15)可知, 只要某条规则 IF 部分模糊集的隶属度值的乘积大于等于其他规则, 就可认为该模糊系统的输出是不变的。因此, 这类模糊系统对于输入和隶属度函数 $\mu_{A_i^l}(x_i)$ 的微小波动是稳健的。不过, 这类模糊系统是不连续的, 即当 l^* 从一个数到另一个数发生变化时, $f(x)$ 会以离散方式发生变化。如果模糊系统用于决策或其他开环应用, 则这种突变是可以容忍的, 但在闭环控制中通常是不可接受的。

下面的引理说明, 用最小推理机也可得到一个类似的结果。

引理 9.5 如果将引理 9.4 中的乘积推理机替换为最小推理机(7.24), 其他部分保持不变, 则该模糊系统与式(9.14)相同, 其中 l^* 由下式确定:

$$\min_{i=1}^n \left[\mu_{A_i^{l^*}}(x_i) \right] \geq \min_{i=1}^n \left[\mu_{A_i^l}(x_i) \right] \quad (9.19)$$

式中, $l = 1, 2, \dots, M$ 。

证明 根据式(7.29)(引理 7.3)和 $\sup_{y \in V}$ 与 $\max_{l=1}^M$ 的互换性以及 B^l 为正交模糊集, 可得

$$\begin{aligned} \sup_{y \in V} \mu_{B^l}(y) &= \max_{l=1}^M \left[\sup_{y \in V} \min(\mu_{A_1^l}(x_1), \dots, \mu_{A_n^l}(x_n), \mu_{B^l}(y)) \right] \\ &= \max_{l=1}^M \left[\min_{i=1}^n (\mu_{A_i^l}(x_i)) \right] \\ &= \min_{i=1}^n (\mu_{A_i^{l^*}}(x_i)) \end{aligned} \quad (9.20)$$

由式(7.29)可知, $\mu_{B^l}(\bar{y}^{l^*}) = \min_{i=1}^n (\mu_{A_i^{l^*}}(x_i))$, 那么式(9.20)中的 $\sup_{y \in V}$ 就可在 \bar{y}^{l^*} 处得到。然

后应用最大值解模糊器(8.23)即可得到式(9.14)。

此外,还可得到一类为简单函数的模糊系统。

对于带有最大值解模糊器和 Lukasiewicz 推理机或 Zadeh 推理机或 Dienes-Rescher 推理机的模糊系统来说,很难得到模糊系统的闭型公式。其原因在于,一般情况下 $\sup_{y \in V}$ 和 \min 算子是无法互换的,所以,由式(7.30)~式(7.32)可以看出,最大值解模糊器问题变成了一个非平滑函数的优化问题。在此情况下,给定一个输入 x ,模糊系统的输出将不得不按一步步(step by step)的方式进行计算,即依次计算模糊器、模糊推理机和解模糊器的输出。应注意,模糊推理机的输出是一个函数,而不是一个单值,所以,其计算非常复杂。因此,本书的余下部分将不再采用这类带有最大值解模糊器、Lukasiewicz 推理机、Zadeh 推理机或 Dienes-Rescher 推理机的模糊系统。

9.2 模糊系统作为万能逼近器

上一节给出了几类模糊系统的严格非线性公式。一方面,这些严格的公式简化了模糊系统的运算;另一方面,它们给了我们一个更详细分析模糊系统的机会。可以看出,模糊系统是非线性函数的特殊类型,所以不管模糊系统是否用于控制器、决策、信号处理或任何其他功能,从函数逼近的角度了解模糊系统的能力都是有意义的。如,模糊系统能代表或逼近什么类型的非线性函数?精度如何?如果模糊系统只能以有限精度逼近某些类型的非线性函数,那么在一般应用中模糊系统不会十分有用。但是,如果模糊系统可以以任意精度逼近任意非线性函数,那么它将可以广泛应用到各类问题中去。本节将证明,在上一节中研究过的几类模糊系统都是具有这种万能逼近能力的。请看下面的定理:

定理 9.1(万能逼近定理) 假定输入论域 U 是 R^n 上的一个紧集,则对于任意定义在 U 上的实连续函数 $g(x)$ 和任意的 $\varepsilon > 0$,一定存在如式(9.6)的模糊系统 $f(x)$ 使下式成立:

$$\sup_{x \in U} |f(x) - g(x)| < \varepsilon \quad (9.21)$$

即,带有乘积推理机、单值模糊器、中心平均解模糊器和高斯隶属度函数的模糊系统是万能逼近器。

该定理的一种证明方法是根据下面的众所周知的 Stone-Weierstrass 定理进行证明。

Stone-Weierstrass 定理(Rudin[1976]) 令 Z 为紧集 U 上的一个实连续函数的集合,如果(i) Z 是代数,即集合 Z 在加、乘积和标量积下是闭集合;(ii) Z 分离了 U 上的点,即对任意 $x, y \in U, x \neq y$,存在 $f \in Z$ 使 $f(x) \neq f(y)$ 成立;(iii) Z 使得 U 中的点不为零,即对于任意 $x \in U$,存在 $f \in Z$ 使 $f(x) \neq 0$,则对 U 上的任意实连续函数 $g(x)$ 和任意 $\varepsilon > 0$,都存在 $f \in Z$ 使 $\sup_{x \in U} |f(x) - g(x)| < \varepsilon$ 成立。

定理 9.1 的证明 令 Y 为所有形如式(9.6)的模糊系统的集合。现在要证明, Y 是代

数, Y 分离了 U 上的点, Y 使得 U 中的点不为零。

令 $f_1, f_2 \in Y$, 于是可将 f_1, f_2 写成

$$f_1(x) = \frac{\sum_{l=1}^{M1} \bar{y}1^l \left[\prod_{i=1}^n a1_i^l \exp\left(-\left(\frac{x_i - \bar{x}1_i^l}{\sigma1_i^l}\right)^2\right)\right]}{\sum_{l=1}^{M1} \left[\prod_{i=1}^n a1_i^l \exp\left(-\left(\frac{x_i - \bar{x}1_i^l}{\sigma1_i^l}\right)^2\right)\right]} \quad (9.22)$$

$$f_2(x) = \frac{\sum_{l=1}^{M2} \bar{y}2^l \left[\prod_{i=1}^n a2_i^l \exp\left(-\left(\frac{x_i - \bar{x}2_i^l}{\sigma2_i^l}\right)^2\right)\right]}{\sum_{l=1}^{M2} \left[\prod_{i=1}^n a2_i^l \exp\left(-\left(\frac{x_i - \bar{x}2_i^l}{\sigma2_i^l}\right)^2\right)\right]} \quad (9.23)$$

由此可得

$$f_1(x)f_2(x) = \frac{\sum_{l1=1}^{M1} \sum_{l2=1}^{M2} (\bar{y}1^{l1} + \bar{y}2^{l2}) \left[\prod_{i=1}^n a1_i^{l1} + a2_i^{l2} \exp\left(-\left(\frac{x_i - \bar{x}1_i^{l1}}{\sigma1_i^{l1}}\right)^2 - \left(\frac{x_i - \bar{x}2_i^{l2}}{\sigma2_i^{l2}}\right)^2\right)\right]}{\sum_{l1=1}^{M1} \sum_{l2=1}^{M2} \left[\prod_{i=1}^n a1_i^{l1} a2_i^{l2} \exp\left(-\left(\frac{x_i - \bar{x}1_i^{l1}}{\sigma1_i^{l1}}\right)^2 - \left(\frac{x_i - \bar{x}2_i^{l2}}{\sigma2_i^{l2}}\right)^2\right)\right]} \quad (9.24)$$

因为 $a1_i^{l1} a2_i^{l2} \exp\left(-\left(\frac{x_i - \bar{x}1_i^{l1}}{\sigma1_i^{l1}}\right)^2 - \left(\frac{x_i - \bar{x}2_i^{l2}}{\sigma2_i^{l2}}\right)^2\right)$ 可以用式(9.4)的形式表示, $\bar{y}_1 + \bar{y}_2$ 又可以看做是形如式(9.5)的模糊集的中心, 所以, $f_1(x) + f_2(x)$ 可以表示成式(9.6)的形式, 即 $f_1 + f_2 \in Y$ 。同理

$$f_1(x) + f_2(x) = \frac{\sum_{l1=1}^{M1} \sum_{l2=1}^{M2} (\bar{y}1^{l1} + \bar{y}2^{l2}) \left[\prod_{i=1}^n a1_i^{l1} a2_i^{l2} \exp\left(-\left(\frac{x_i - \bar{x}1_i^{l1}}{\sigma1_i^{l1}}\right)^2 - \left(\frac{x_i - \bar{x}2_i^{l2}}{\sigma2_i^{l2}}\right)^2\right)\right]}{\sum_{l1=1}^{M1} \sum_{l2=1}^{M2} \left[\prod_{i=1}^n a1_i^{l1} a2_i^{l2} \exp\left(-\left(\frac{x_i - \bar{x}1_i^{l1}}{\sigma1_i^{l1}}\right)^2 - \left(\frac{x_i - \bar{x}2_i^{l2}}{\sigma2_i^{l2}}\right)^2\right)\right]} \quad (9.25)$$

上式也可以表示成式(9.6)的形式, 所以, $f_1 f_2 \in Y$ 。最后, 对任意 $c \in R$, 有

$$cf_1(x) = \frac{\sum_{l=1}^{M1} c \bar{y}1^l \left[\prod_{i=1}^n a1_i^l \exp\left(-\left(\frac{x_i - \bar{x}1_i^l}{\sigma1_i^l}\right)^2\right)\right]}{\sum_{l=1}^{M1} \left[\prod_{i=1}^n a1_i^l \exp\left(-\left(\frac{x_i - \bar{x}1_i^l}{\sigma1_i^l}\right)^2\right)\right]} \quad (9.26)$$

上式也可以表示成式(9.6)的形式, 所以, $cf_1 \in Y$ 。因此, Y 是代数。

下面通过构造一个所需的模糊系统 $f(x)$ 来证明, Y 分离了 U 上的点。令 $x^0, z^0 \in U$ 为两个任意点且 $x^0 \neq z^0$ 。选择如式(9.6)的 $f(x)$ 的参数为: $M = 2, \bar{y}^1 = 0, \bar{y}^2 = 1, a_i^l = 1, \sigma_i^l = 1, \bar{x}_i^1 = x_i^0, \bar{x}_i^2 = z_i^0 (i = 1, 2, \dots, n, l = 1, 2)$ 。这一特定的模糊系统为

$$f(x) = \frac{\exp(-\|x - z^0\|_2^2)}{\exp(-\|x - x^0\|_2^2) + \exp(-\|x - z^0\|_2^2)} \quad (9.27)$$

由此可得

$$f(x^0) = \frac{\exp(-\|x^0 - z^0\|_2^2)}{1 + \exp(-\|x^0 - z^0\|_2^2)} \quad (9.28)$$

和

$$f(z^0) = \frac{1}{1 + \exp(-\|x^0 - z^0\|_2^2)} \quad (9.29)$$

因为 $x^0 \neq z^0$, 所以, 根据式(9.28)和式(9.29), 可得 $\exp(-\|x^0 - z^0\|_2^2) \neq 1$, 进而可得 $f(x^0) \neq f(z^0)$ 。因此, Y 分离了 U 上的点。

下面证明, Y 使得 U 中的点不为零。显然, 任意形如式(9.6)的模糊系统 $f(x)$ ($\bar{y}^l > 0$) 都具有 $f(x) > 0$ ($\forall x \in U$) 这一性质。故 Y 使得 U 中的点不为零。

归纳上面的证明, 并由 Stone-Weierstrass 定理, 即可得到本定理的结论。

定理 9.1 说明了, 模糊系统可以以任意精度逼近连续函数。下面的推论又将这一结论扩展至离散函数。

推论 9.1 对于任意一个紧集 $U \subset R^n$ 上的平方可积函数, 即对任意 $g \in L_2(U) = \{g: U \rightarrow R \mid \int_U |g(x)|^2 dx < \infty\}$, 一定存在形如式(9.6)的模糊系统 $f(x)$ 使得下式成立:

$$\left(\int_U |f(x) - g(x)|^2 dx \right)^{1/2} < \varepsilon \quad (9.30)$$

证明 因为 U 为紧集, 所以 $\int_U dx = E < \infty$ 。又因为 U 上的连续函数形成了 $L_2(U)$ 的一个稠密子集(Rudin[1976]), 所以, 对于任意 $g \in L_2(U)$, 一定存在一个 U 上的连续函数 \bar{g} 使得 $\left(\int_U |g(x) - \bar{g}(x)|^2 dx \right)^{1/2} < \varepsilon/2$ 成立。由定理 9.1 可知, 存在 $f \in Y$ 使得 $\sup_{x \in U} |f(x) - \bar{g}(x)| < \varepsilon/(2E^{1/2})$ 。由此可得

$$\begin{aligned} \left(\int_U |f(x) - g(x)|^2 dx \right)^{1/2} &\leq \left(\int_U |f(x) - \bar{g}(x)|^2 dx \right)^{1/2} + \left(\int_U |\bar{g}(x) - g(x)|^2 dx \right)^{1/2} \\ &< \left(\int_U \left(\sup_{x \in U} |f(x) - \bar{g}(x)| \right)^2 dx \right)^{1/2} + \varepsilon/2 \\ &< \left(\frac{\varepsilon^2}{2^2 E} E \right)^{1/2} + \varepsilon/2 = \varepsilon \end{aligned} \quad (9.31)$$

定理 9.1 和推论 9.1 为模糊系统的广泛应用提供了证明。即, 它们证明了对于任意非线性运算总能设计出一个模糊系统, 并使其以任意精度完成所需运算。同时, 定理 9.1 和推论 9.1 也对模糊系统在实际中的成功应用提供了一种理论上的解释。

不过, 定理 9.1 和推论 9.1 仅给出了存在性结论。即, 它们仅证明了存在形如式(9.6)的能以任意精度逼近任意函数的模糊系统, 但并不知道怎样找到这样一个模糊系统。对于工程应用而言, 知道一种理想的模糊系统的存在性是不够的, 必须开发一种能为特定应用找

到合适模糊系统的方法。因为,根据所提供的信息可能找到也可能找不到理想的模糊系统,所以接下来的几章将研究各种设计模糊系统的方法。

9.3 本章小结及导读

本章主要阐述了以下内容:

- 几类有用的模糊系统的严格公式;
- 如果这类严格公式存在,则如何推导出任意一种模糊系统的严格公式;
- 怎样应用 Stone-Weierstrass 定理。

模糊系统的数学公式是最近推导出的,其相关参考文献见 Wang[1994]。万能逼近定理及其证明都摘自 Wang[1992]。该问题的其他研究方法见 Buckley[1992b]与 Zeng 和 Singh[1994]。

9.4 习题

习题 9.1 推导带有模糊规则库(7.1)、Zadeh 推理机(7.31)、单值模糊器(8.1)和中心平均解模糊器(8.18)的模糊系统的严格公式。

习题 9.2 用 Lukasiewicz 推理机替代习题 9.1 中 Zadeh 推理机,然后重做习题 9.1。

习题 9.3 证明定理 9.1 中形如式(9.1)的模糊系统具有万能逼近特性。

习题 9.4 你能用 Stone-Weierstrass 定理证明形如式(9.7)或式(9.6)的模糊系统($a_i^l = 1$)是万能逼近器吗?并做出解释。

习题 9.5 用 Stone-Weierstrass 定理证明多项式是万能逼近器。

习题 9.6 画出 $x \in U = [-1, 2] \times [-1, 2]$ 时的模糊系统 $f_1(x)$ 和 $f_2(x)$ 的图形,这里 $f_1(x)$ 是带有两条模糊规则(8.27)和(8.28)、乘积推理机(7.23)、单值模糊器(8.1)和最大值解模糊器(8.23)的模糊系统, $f_2(x)$ 除了采用中心平均解模糊器(8.18)而不是最大值解模糊器外,其余组成部分都与 $f_1(x)$ 相同。

模糊系统的逼近特性 I

在第 9 章中已经证明了模糊系统是万能逼近器。即,模糊系统能以任意精度逼近紧集上的任意函数。但是,这一结论仅证明了最优模糊系统的存在性,并未给出找到最优模糊系统的方法。事实上,找到最优模糊系统要比证明它的存在性难得多。根据所提供的信息,可能找到也可能找不到最优的模糊系统。

为了回答怎样找到最优模糊系统这一问题,首先必须知道对于要去逼近的非线性函数 $g(x): U \subset R^n \rightarrow R$ 来说,什么信息是可以获得的。一般来说,可能遇到以下三种情况:

- $g(x)$ 的解析公式已知;
- $g(x)$ 的解析公式未知,但对于任意 $x \in U$ 可以确定相应的 $g(x)$,即 $g(x)$ 是一个黑箱——仅知道 $g(x)$ 的输入—输出行为,而不知道黑箱内部的详细情况;
- $g(x)$ 的解析公式未知,仅知道有限数量的输入—输出数据对 $(x^j, g(x^j))$,其中 $x^j \in U$ 是不能任意选择的。

第一种情况不是很有意义,因为若 $g(x)$ 的解析公式已知,就可以用 $g(x)$ 来达到模糊系统想要达到的任何目标,几乎不会用模糊系统去取代 $g(x)$ 。所以,暂不单独考虑第一种情况,而是把第一种情况作为第二种情况的特例,寻求处理第二种情况的方法。

第二种情况更具现实性。本章及余下章节中将会对其做详细研究。

第三种情况在实践中则更为常见,尤其在模糊控制中更是如此,因为控制系统的稳定性可能要求不可以任意选择输入值。第 12 章至第 15 章将对这种情况进行详细研究。

因此,本章假定 $g(x)$ 的解析公式未知,但对任意 $x \in U$,可以确定输入—输出数据对 $(x; g(x))$ 。本章的主要任务是设计一个能以某种最优方式逼近 $g(x)$ 的模糊系统。

10.1 基本概念

首先介绍一些基本概念。

定义 10.1 四边形隶属度函数。令 $[a, d] \subset R$,模糊集 A 的四边形隶属度函数是 R 上的一个连续函数,其定义为

$$\mu_A(x; a, b, c, d, H) = \begin{cases} I(x), x \in [a, b) \\ H, x \in [b, c] \\ D(x), x \in (c, d] \\ 0, x \in R - (a, d) \end{cases} \quad (10.1)$$

式中, $a \leq b \leq c \leq d, 0 \leq H \leq 1, 0 \leq I(x) \leq 1$ 是 $[a, b)$ 上的一个非减函数, $0 \leq D(x) \leq 1$ 是 $(c, d]$ 上的一个非增函数。当模糊集 A 为标准模糊集(即 $H = 1$)时, 其隶属度函数可以简单地记为 $\mu_A(x; a, b, c, d)$ 。

图 10.1 给出了四边形隶属度函数的一些例子。如果论域是有边界的, 则 a, b, c, d 是有限数。四边形隶属度函数包含了许多常用的隶属度函数, 如, 若选定

$$I(x) = \frac{x-a}{b-a}, D(x) = \frac{x-d}{c-d} \quad (10.2)$$

则四边形隶属度函数就变成了梯形隶属度函数。如果 $b = c, I(x)$ 和 $D(x)$ 同式(10.2)中的 $I(x)$ 和 $D(x)$, 就会得到三角形隶属度函数, 记为 $\mu_A(x; a, b, d)$ 。如果选定 $a = -\infty, b = c = \bar{x}, d = \infty$, 且

$$I(x) = D(x) = \exp\left(-\left(\frac{x-\bar{x}}{\sigma}\right)^2\right) \quad (10.3)$$

则四边形隶属度函数就变成了高斯隶属度函数。所以说, 四边形隶属度函数是一个隶属度函数族。

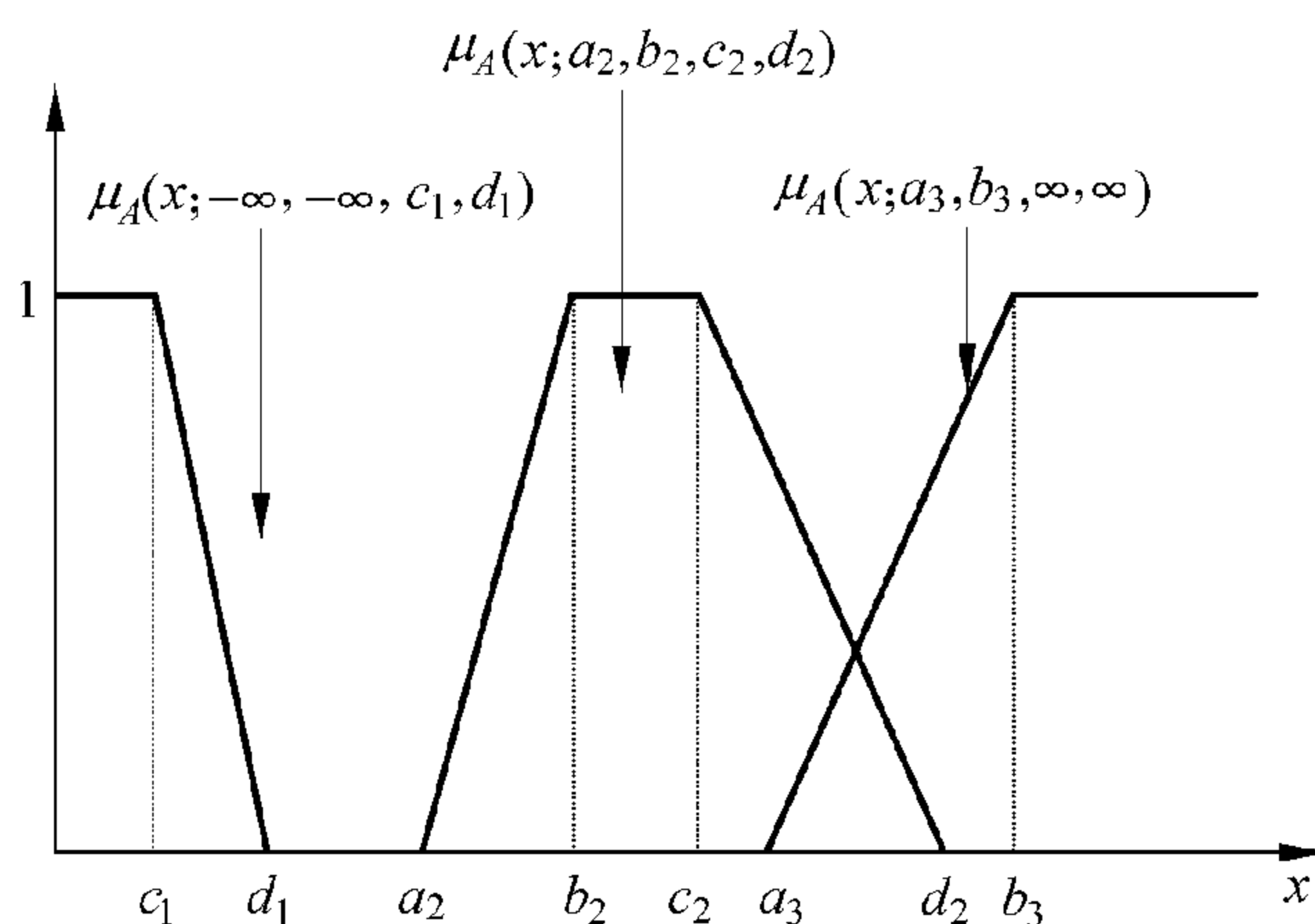


图 10.1 四边形隶属度函数的一个例子

定义 10.2 模糊集的完备性。如果对任意 $x \in W$ 都存在 A^j 使 $\mu_{A^j}(x) > 0$, 则认为 $W \subset R$ 上的模糊集 A^1, A^2, \dots, A^N 在 W 上是完备的。

定义 10.3 模糊集的一致性。如果对某个 $x \in W$ 有 $\mu_{A^i}(x) = 1$ 成立且对所有 $i \neq j$ 都有 $\mu_{A^j}(x) = 0$ 成立, 则认为 $W \subset R$ 上的模糊集 A^1, A^2, \dots, A^N 在 W 上是一致的。

定义 10.4 模糊集的高峰集。 $W \subset R$ 上的一个模糊集 A 的高峰集是 W 上的一个子集, 其定义为

$$hgh(A) = \left\{ x \in W \mid \mu_A(x) = \sup_{x' \in W} \mu_A(x') \right\} \quad (10.4)$$

如果 A 是一个具有四边形隶属度函数 $\mu_A(x; a, b, c, d)$ 的标准模糊集, 则 $hgh(A) = [b, c]$ 。

定义 10.5 模糊集的排序。对于 $W \subset R$ 上的两个模糊集 A 和 B , 如果 $hgh(A) > hgh(B)$, 则称 $A > B$ [即如果 $x \in hgh(A), x' \in hgh(B)$, 则 $x > x'$]。

下面证明一些具有四边形隶属度函数的模糊集的性质。

引理 10.1 如果 A^1, A^2, \dots, A^N 是 $W \subset R$ 上一致的、标准的模糊集, 其具有四边形隶属度函数 $\mu_{A^i}(x; a_i, b_i, c_i, d_i), i = 1, 2, \dots, N$, 则存在一个 $\{1, 2, \dots, N\}$ 的重排列 $\{i_1, i_2, \dots, i_N\}$ 使得

$$A^{i_1} < A^{i_2} < \dots < A^{i_N} \quad (10.5)$$

证明 对于任意 $i, j \in \{1, 2, \dots, N\}, [b_i, c_i] \cap [b_j, c_j] = \emptyset$ 一定是成立的, 否则模糊集 A^1, A^2, \dots, A^N 将是不一致的。这样, 就存在一个 $\{1, 2, \dots, N\}$ 的重排列 $\{i_1, i_2, \dots, i_N\}$ 使得下式成立

$$[b_{i_1}, c_{i_1}] < [b_{i_2}, c_{i_2}] < [b_{i_N}, c_{i_N}] \quad (10.6)$$

上式说明了式(10.5)是成立的。

引理 10.1 表明, 为不失一般性, 总是假设 $A^1 < A^2 < \dots < A^N$ 。(证明)

引理 10.2 令 $W \subset R$ 上的模糊集 A^1, A^2, \dots, A^N 为标准的、一致的、完备的模糊集, 其具有四边形隶属度函数 $\mu_{A^i}(x; a_i, b_i, c_i, d_i)$ 。如果 $A^1 < A^2 < \dots < A^N$, 则当 $i = 1, 2, \dots, N-1$ 时, 有

$$c_i \leq a_{i+1} < d_i \leq b_{i+1} \quad (10.7)$$

引理 10.2 的示例见图 10.2, 其证明很简单, 留作习题。

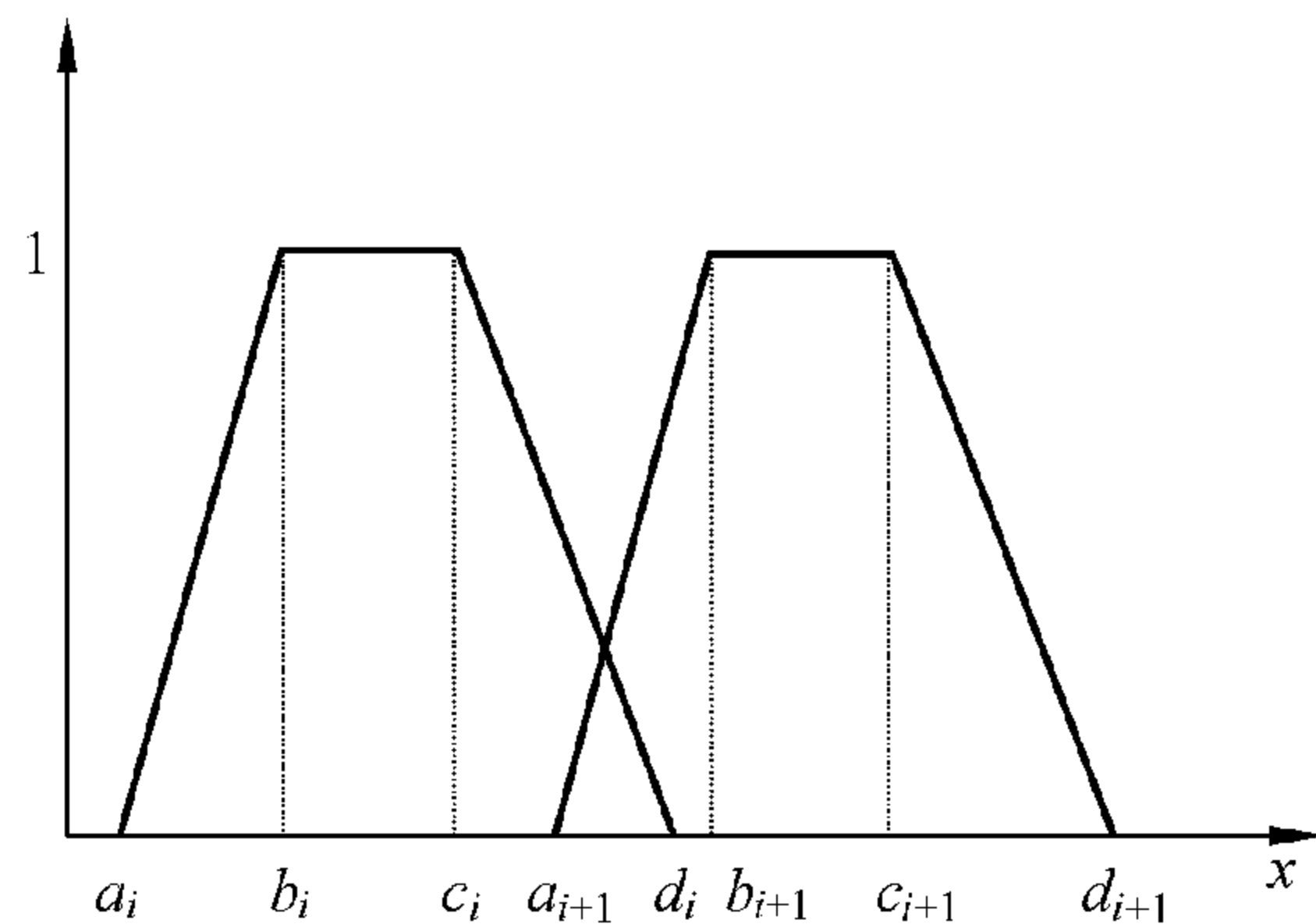


图 10.2 引理 10.2 的一个例子:
 $c_i \leq a_{i+1} < d_i \leq b_{i+1}$

10.2 模糊系统的设计

现在准备设计一种具有某些良好特性的特定类型的模糊系统。为使符号简便和图示便捷, 这里只考虑二维输入的模糊系统。其实, 其方法和结论对于一般的 n 维输入的模糊系统也是有效的, 即, 可用适当的相同程序来设计 n 维输入的模糊系统。首先, 对这类问题做以下描述。

问题 令 $g(x)$ 为紧集 $U = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2] \subset R^2$ 上的一个函数, $g(x)$ 的解析形式未知。假设对任意一个 $x \in U$, 都能得到 $g(x)$, 任务是设计一个逼近 $g(x)$ 的模糊系统。

下面,逐步地设计一个此类模糊系统。

模糊系统的设计

步骤 1 在 $[\alpha_i, \beta_i]$ 上定义 $N_i (i=1, 2)$ 个标准的、一致的和完备的模糊集 $A_i^1, A_i^2, \dots, A_i^{N_i}$ 。它们具有四边形隶属度函数 $\mu_{A_i^1}(x_i; a_i^1, b_i^1, c_i^1, d_i^1), \dots, \mu_{A_i^{N_i}}(x_i; a_i^{N_i}, b_i^{N_i}, c_i^{N_i}, d_i^{N_i})$ 且 $A_i^1 < A_i^2 < \dots < A_i^{N_i}$, 其中 $a_i^1 = b_i^1 = \alpha_i, c_i^{N_i} = d_i^{N_i} = \beta_i$ 。定义 $e_1^1 = \alpha_1, e_1^{N_1} = \beta_1, e_1^j = \frac{1}{2}(b_1^j + c_1^j), j=2, 3, \dots, N_1 - 1$ 。类似地, 定义 $e_2^1 = \alpha_2, e_2^{N_2} = \beta_2, e_2^j = \frac{1}{2}(b_2^j + c_2^j), j=2, 3, \dots, N_2 - 1$ 。图 10.3 给出了在 $N_1 = 3, N_2 = 4, \alpha_1 = \alpha_2 = 0, \beta_1 = \beta_2 = 1$ 时的一个例子。

步骤 2 以如下形式组建 $M = N_1 \times N_2$ 条模糊集 IF-THEN 规则:

$R_u^{i_1 i_2}$: 如果 x_1 为 $A_1^{i_1}$ 且 x_2 为 $A_2^{i_2}$,

则 y 为 $B^{i_1 i_2}$ (10.8)

其中, $i_1 = 1, 2, \dots, N_1, i_2 = 1, 2, \dots, N_2$, 将模糊集 $B^{i_1 i_2}$ 的中心(用 $\bar{y}^{i_1 i_2}$ 表示)选择为

$\bar{y}^{i_1 i_2} = g(e_1^{i_1}, e_2^{i_2})$ (10.9)

对于图 10.3 中的例子来说, 有 $3 \times 4 = 12$ 条规则, $B^{i_1 i_2}$ 的中心等于图 10.3 中所示的 12 个黑点上所估计 $g(x)$ 的值。

步骤 3 用乘积推理机(7.23)、单值模糊器(8.1)和中心平均解模糊器(8.18)(见引理 9.1), 根据式(10.8)中的 $N_1 \times N_2$ 条规则来构造模糊系统 $f(x)$

$$f(x) = \frac{\sum_{i_1=1}^{N_1} \sum_{i_2=1}^{N_2} \bar{y}^{i_1 i_2} (\mu_{A_1^{i_1}}(x_1) \mu_{A_2^{i_2}}(x_2))}{\sum_{i_1=1}^{N_1} \sum_{i_2=1}^{N_2} (\mu_{A_1^{i_1}}(x_1) \mu_{A_2^{i_2}}(x_2))} \quad (10.10)$$

由于模糊集 $A_i^1, \dots, A_i^{N_i}$ 是完备的, 在每个 $x \in U$ 处都存在 i_1, i_2 使得 $\mu_{A_1^{i_1}}(x_1) \mu_{A_2^{i_2}}(x_2) \neq 0$, 所以模糊系统(10.10)可以被定义的很好, 即其分母总是非零的。

由步骤 2 可以看出, 规则(10.8)的 IF 部分包含了为每个输入变量所定义的模糊集的所有可能组合。所以, 若将设计程序推广到 n 维输入的模糊系统, 并根据每个输入变量定义 N 个模糊集, 则得到总的规则数量为 N^n 。即, 通过这种设计方法, 规则的数量会随输入空间维数的上升而呈指数上升, 这种情况被称作维度灾难, 它是在所有高维逼近问题中都会遇到的一个一般性的问题。第 22 章将会对此问题再次阐述。

最后, 由设计程序可发现, 必须知道 $g(x)$ 在 $x = (e_1^{i_1}, e_2^{i_2}), (i_1 = 1, 2, \dots, N_1, i_2 = 1, 2, \dots, N_2)$

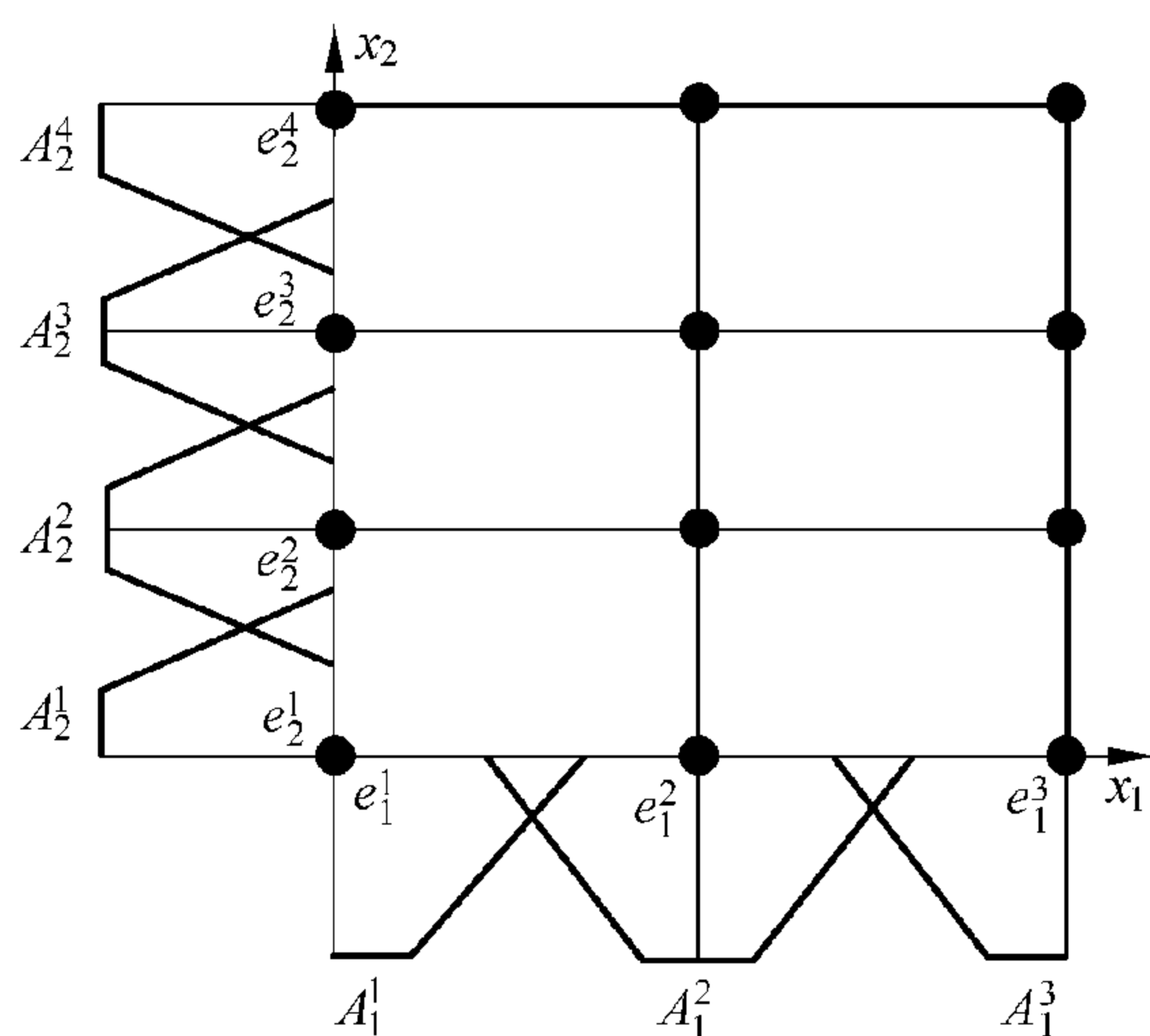


图 10.3 设计程序步骤 1 中定义的模糊集的例子

处的值,因为 $(e_1^{i_1}, e_2^{i_2})$ 是 U 上的任意点,所以,这与知道 $g(x)$ 在任意 $x \in U$ 处的值是一样的。

下面将研究所设计的 $f(x)$ 逼近上述未知函数 $g(x)$ 的逼近精度问题。

10.3 模糊系统的逼近精度

定理 10.1 令 $f(x)$ 为式(10.10)中的模糊系统, $g(x)$ 为式(10.9)中的未知函数。如果 $g(x)$ 在 $U = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2]$ 上是连续可微的,则

$$\|g - f\|_{\infty} \leq \left\| \frac{\partial g}{\partial x_1} \right\|_{\infty} h_1 + \left\| \frac{\partial g}{\partial x_2} \right\|_{\infty} h_2 \quad (10.11)$$

式中,无穷维范数 $\|*\|_{\infty}$ 定义为 $\|d(x)\|_{\infty} = \sup_{x \in U} |d(x)|$, $h_i = \max_{1 \leq j \leq N_i - 1} |e_i^{j+1} - e_i^j|$ ($i = 1, 2$)。

证明 令 $U^{i_1 i_2} = [e_1^{i_1}, e_1^{i_1+1}] \times [e_2^{i_2}, e_2^{i_2+1}]$, 其中, $i_1 = 1, 2, \dots, N_1 - 1$; $i_2 = 1, 2, \dots, N_2 - 1$ 。因为 $[\alpha_i, \beta_i] = [e_i^1, e_i^2] \cup [e_i^2, e_i^3] \cup \dots \cup [e_i^{N_i-1}, e_i^{N_i}]$ ($i = 1, 2$), 所以有

$$U = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2] = \bigcup_{i_1=1}^{N_1-1} \bigcup_{i_2=1}^{N_2-1} U^{i_1 i_2} \quad (10.12)$$

这表明对任意 $x \in U$, 都存在 $U^{i_1 i_2}$ 使得 $x \in U^{i_1 i_2}$ 。现假定 $x \in U^{i_1 i_2}$, 即 $x_1 \in [e_1^{i_1}, e_1^{i_1+1}]$, $x_2 \in [e_2^{i_2}, e_2^{i_2+1}]$ (因为 x 是确定的, 所以 i_1 和 i_2 在次序上也是确定的)。由于模糊集 $A_1^1, \dots, A_1^{N_1}$ 为标准的、一致的、完备的模糊集, 所以, 当 $j_1 = 1, 2, \dots, N_1$ 时, 至少有一个至多有两个 $\mu_{A_1^{j_1}}(x_1)$ 非零。由 $e_1^{j_1}$ ($j_1 = 1, 2, \dots, N_1 - 1$)的定义可知, 两个可能非零的 $\mu_{A_1^{j_1}}(x_1)$ 是 $\mu_{A_1^{j_1}}(x_1)$ 和 $\mu_{A_1^{j_1+1}}(x_1)$ 。类似地, 两个可能非零的 $\mu_{A_2^{j_2}}(x_2)$ ($j_2 = 1, 2, \dots, N_2$)是 $\mu_{A_2^{j_2}}(x_2)$ 和 $\mu_{A_2^{j_2+1}}(x_2)$ 。因此, 将式(10.9)代入式(10.10), 可将模糊系统 $f(x)$ (10.10)简化为

$$\begin{aligned} f(x) &= \frac{\sum_{j_1=i_1+1}^{i_1+1} \sum_{j_2=i_2+1}^{i_2+1} \bar{y}^{j_1 j_2} (\mu_{A_1^{j_1}}(x_1) \mu_{A_2^{j_2}}(x_2))}{\sum_{j_1=i_1+1}^{i_1+1} \sum_{j_2=i_2+1}^{i_2+1} (\mu_{A_1^{j_1}}(x_1) \mu_{A_2^{j_2}}(x_2))} \\ &= \sum_{j_1=i_1+1}^{i_1+1} \sum_{j_2=i_2+1}^{i_2+1} \left[\frac{\mu_{A_1^{j_1}}(x_1) \mu_{A_2^{j_2}}(x_2)}{\sum_{j_1=i_1+1}^{i_1+1} \sum_{j_2=i_2+1}^{i_2+1} (\mu_{A_1^{j_1}}(x_1) \mu_{A_2^{j_2}}(x_2))} \right] g(e_1^{j_1}, e_2^{j_2}) \end{aligned} \quad (10.13)$$

因为

$$\sum_{j_1=i_1+1}^{i_1+1} \sum_{j_2=i_2+1}^{i_2+1} \left[\frac{\mu_{A_1^{j_1}}(x_1) \mu_{A_2^{j_2}}(x_2)}{\sum_{j_1=i_1+1}^{i_1+1} \sum_{j_2=i_2+1}^{i_2+1} (\mu_{A_1^{j_1}}(x_1) \mu_{A_2^{j_2}}(x_2))} \right] = 1 \quad (10.14)$$

所以

$$\begin{aligned}
|g(x) - f(x)| &\leq \sum_{j_1=i_1}^{i_1+1} \sum_{j_2=i_2}^{i_2+1} \left[\frac{\mu_{A_1^{j_1}}(x_1) \mu_{A_2^{j_2}}(x_2)}{\sum_{j_1=i_1}^{i_1+1} \sum_{j_2=i_2}^{i_2+1} (\mu_{A_1^{j_1}}(x_1) \mu_{A_2^{j_2}}(x_2))} \right] |g(x) - g(e_1^{j_1}, e_2^{j_2})| \\
&\leq \max_{j_1=i_1, i_1+1; j_2=i_2, i_2+1} |g(x) - g(e_1^{j_1}, e_2^{j_2})| \quad (10.15)
\end{aligned}$$

由均值定理,可进一步把式(10.15)写为

$$|g(x) - f(x)| \leq \max_{j_1=i_1, i_1+1; j_2=i_2, i_2+1} \left(\left\| \frac{\partial g}{\partial x_1} \right\|_{\infty} |x_1 - e_1^{j_1}| + \left\| \frac{\partial g}{\partial x_2} \right\|_{\infty} |x_2 - e_2^{j_2}| \right) \quad (10.16)$$

因为 $x \in U^{i_1 i_2}$, 即 $x_1 \in [e_1^{i_1}, e_1^{i_1+1}]$, $x_2 \in [e_2^{i_2}, e_2^{i_2+1}]$, 所以有 $|x_1 - e_1^{j_1}| \leq |e_1^{i_1+1} - e_1^{i_1}|$ 和 $|x_2 - e_2^{j_2}| \leq |e_2^{i_2+1} - e_2^{i_2}|$ ($j_1 = i_1, i_1 + 1; j_2 = i_2, i_2 + 1$)。于是,式(10.16)变为

$$|g(x) - f(x)| \leq \left\| \frac{\partial g}{\partial x_1} \right\|_{\infty} |e_1^{i_1+1} - e_1^{i_1}| + \left\| \frac{\partial g}{\partial x_2} \right\|_{\infty} |e_2^{i_2+1} - e_2^{i_2}| \quad (10.17)$$

由此可得

$$\begin{aligned}
\|g - f\|_{\infty} &= \sup_{x \in U} |g(x) - f(x)| \\
&\leq \left\| \frac{\partial g}{\partial x_1} \right\|_{\infty} \max_{1 \leq i_1 \leq N_1 - 1} |e_1^{i_1+1} - e_1^{i_1}| + \left\| \frac{\partial g}{\partial x_2} \right\|_{\infty} \max_{1 \leq i_2 \leq N_2 - 1} |e_2^{i_2+1} - e_2^{i_2}| \\
&= \left\| \frac{\partial g}{\partial x_1} \right\|_{\infty} h_1 + \left\| \frac{\partial g}{\partial x_2} \right\|_{\infty} h_2 \quad (10.18)
\end{aligned}$$

定理 10.1 是一个重要的定理,从中可以得出一些结论,归纳如下:

- 由式(10.11)可以得出结论,形如式(10.10)的模糊系统是万能逼近器。具体来讲,因为 $\left\| \frac{\partial g}{\partial x_1} \right\|_{\infty}$ 和 $\left\| \frac{\partial g}{\partial x_2} \right\|_{\infty}$ 是有限数(是紧集上的一个连续有界函数),所以,对任意给定的 $\varepsilon > 0$,都可将 h_1 和 h_2 选的足够小,使 $\left\| \frac{\partial g}{\partial x_1} \right\|_{\infty} h_1 + \left\| \frac{\partial g}{\partial x_2} \right\|_{\infty} h_2 < \varepsilon$ 成立。因此,由式(10.11),可得 $\sup_{x \in U} |g(x) - f(x)| = \|g - f\|_{\infty} < \varepsilon$ 。
- 由式(10.11)及 h_1 和 h_2 的定义可以看出,通过对每个 x_i 定义更多的模糊集可以得到更为准确的逼近。这一结论直观地说明了:规则越多,所产生的模糊系统越有效。
- 由式(10.11)可以看出,为了设计一个具有预定精度的模糊系统,必须知道 $g(x)$ 关于 x_1 和 x_2 的导数边界,即 $\left\| \frac{\partial g}{\partial x_1} \right\|_{\infty}$ 和 $\left\| \frac{\partial g}{\partial x_2} \right\|_{\infty}$ 。同时,在设计过程中,还必须知道 $g(x)$ 在 $x = (e_1^{i_1}, e_2^{i_2})$ ($i_1 = 1, 2, \dots, N_1; i_2 = 1, 2, \dots, N_2$) 处的值。也就是说,为设计可达到任意预定精度的模糊系统,这种方法是需要知道以上两类信息的。

- 由定理 10.1 的证明可以看出,把 $\mu_{A_1^{i_1}}(x_1)\mu_{A_2^{i_2}}(x_2)$ 变为 $\min[\mu_{A_1^{i_1}}(x_1),\mu_{A_2^{i_2}}(x_2)]$ 后,其证明仍然是成立的。所以,如果在设计程序中使用最小推理机,其他部分保持不变,则所设计的模糊系统仍具有定理 10.1 中逼近性质。因此,带有最小推理机、单值模糊器、中心平均解模糊器和四边形隶属度函数的模糊系统也是万能逼近器。

定理 10.1 给出了作为逼近器的 $f(x)$ 逼近 $g(x)$ 的精度。下面的引理证明了 $g(x)$ 和 $f(x)$ 在什么点处恰好相等。

引理 10.3 令 $f(x)$ 为模糊系统(10.10), $e_1^{i_1}$ 和 $e_2^{i_2}$ 为设计程序中定义的 $f(x)$ 的点,则当 $i_1 = 1, 2, \dots, N_1; i_2 = 1, 2, \dots, N_2$ 时,有

$$f(e_1^{i_1}, e_2^{i_2}) = g(e_1^{i_1}, e_2^{i_2}) \quad (10.19)$$

证明 由 $e_1^{i_1}$ 和 $e_2^{i_2}$ 的定义以及 $A_i^{i_i}$ 为标准模糊集可知, $\mu_{A_1^{i_1}}(e_1^{i_1}) = \mu_{A_2^{i_2}}(e_2^{i_2}) = 1$ 。因为模糊集 $A_i^1, A_i^2, \dots, A_i^{N_i} (i=1, 2)$ 是一致的,所以有 $\mu_{A_1^{j_1}}(e_1^{i_1}) = \mu_{A_2^{j_2}}(e_2^{i_2}) = 0 (j_1 \neq i_1, j_2 \neq i_2)$ 。因此,由式(10.10)和式(10.9),可得, $f(e_1^{i_1}, e_2^{i_2}) = \bar{y}^{i_1 i_2} = g(e_1^{i_1}, e_2^{i_2})$ 。

引理 10.3 说明了,模糊系统(10.10)可以被看做是函数 $g(x)$ 在论域 U 中的某些固定点 $(e_1^{i_1}, e_2^{i_2}) (i_1 = 1, 2, \dots, N_1; i_2 = 1, 2, \dots, N_2)$ 上的插值。这是很直观的。

最后,给出两个例子来说明怎样使用定理 10.1 设计所需的模糊系统。

例 10.1 设计一个模糊系统 $f(x)$, 使之一致地逼近定义在 $U = [-3, 3]$ 上的连续函数 $g(x) = \sin(x)$, 所需精度为 $\varepsilon = 0.2$, 即 $\sup_{x \in U} |g(x) - f(x)| < \varepsilon$ 。

因为 $\left\| \frac{\partial g}{\partial x} \right\|_{\infty} = \|\cos(x)\|_{\infty} = 1$, 所以,由式(10.11)可知, $h = 0.2$ 的模糊系统是满足要求的。

因此,在 $U = [-3, 3]$ 上定义 31 个具有如下三角形隶属度函数的模糊集 A^j :

$$\mu_{A_1}(x) = \mu_{A_1}(x; -3, -3, -2.8) \quad (10.20)$$

$$\mu_{A_{31}}(x) = \mu_{A_{31}}(x; 2.8, 3, 3) \quad (10.21)$$

和

$$\mu_{A_j}(x) = \mu_{A_j}(x; e^{j-1}, e^j, e^{j+1}), j = 2, 3, \dots, 30 \quad (10.22)$$

其中, $e^j = -3 + 0.2(j-1)$ 。这些隶属度函数的图形见图 10.4。根据式(10.10)可知,所设计的模糊系统为

$$f(x) = \frac{\sum_{j=1}^{31} \sin(e^j) \mu_{A_j}(x)}{\sum_{j=1}^{31} \mu_{A_j}(x)} \quad (10.23)$$

上式的图形见图 10.5, 图中 $g(x) = \sin(x)$ 。由图 10.5 可知, $f(x)$ 和 $g(x)$ 是基本一致的。

例 10.2 设计一个模糊系统来一致地逼近定义在 $U = [-1, 1] \times [-1, 1]$ 的函数 $g(x) = 0.52 + 0.1x_1 + 0.28x_2 - 0.06x_1x_2$, 所需精度为 $\varepsilon = 0.1$ 。

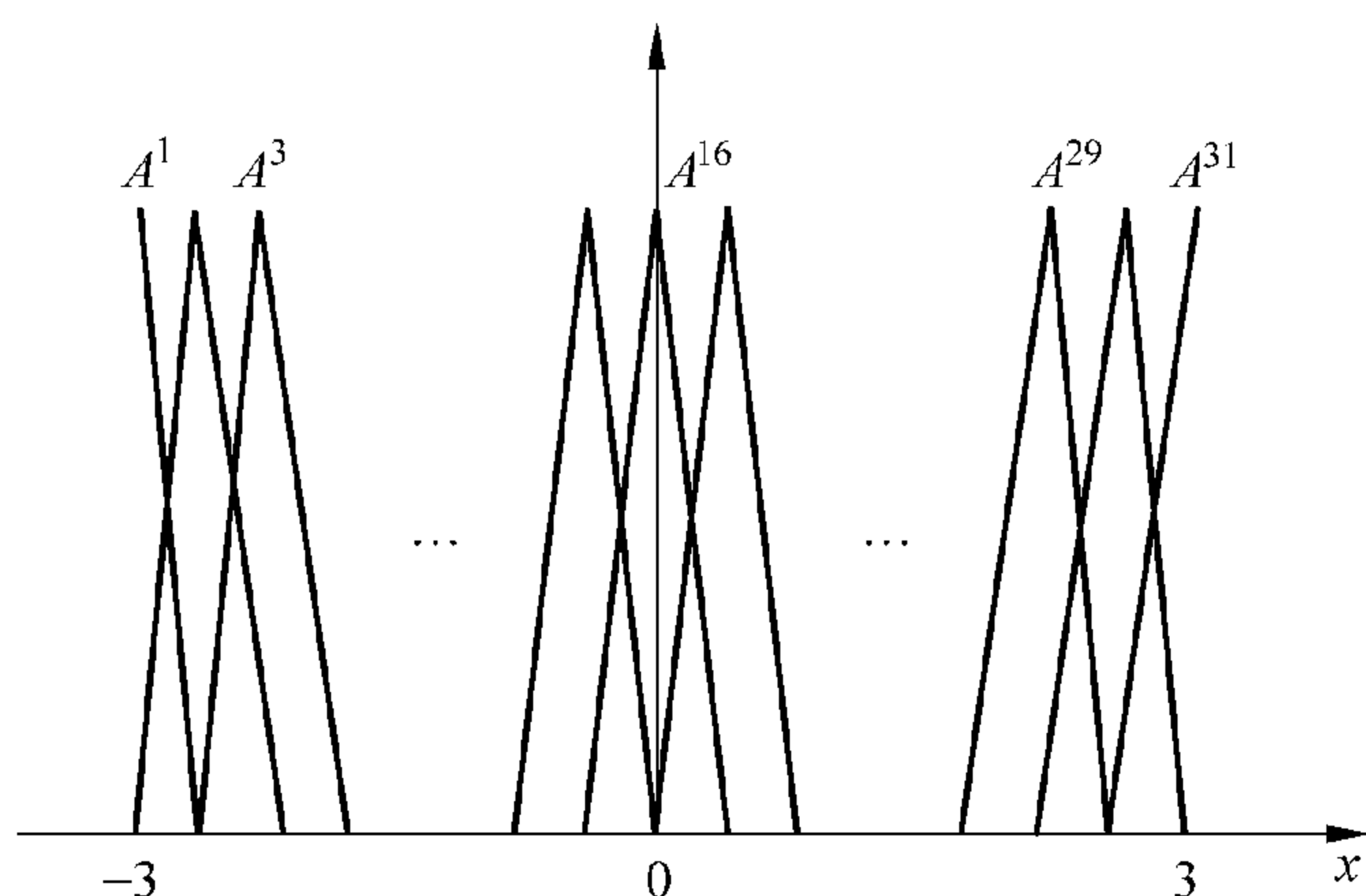
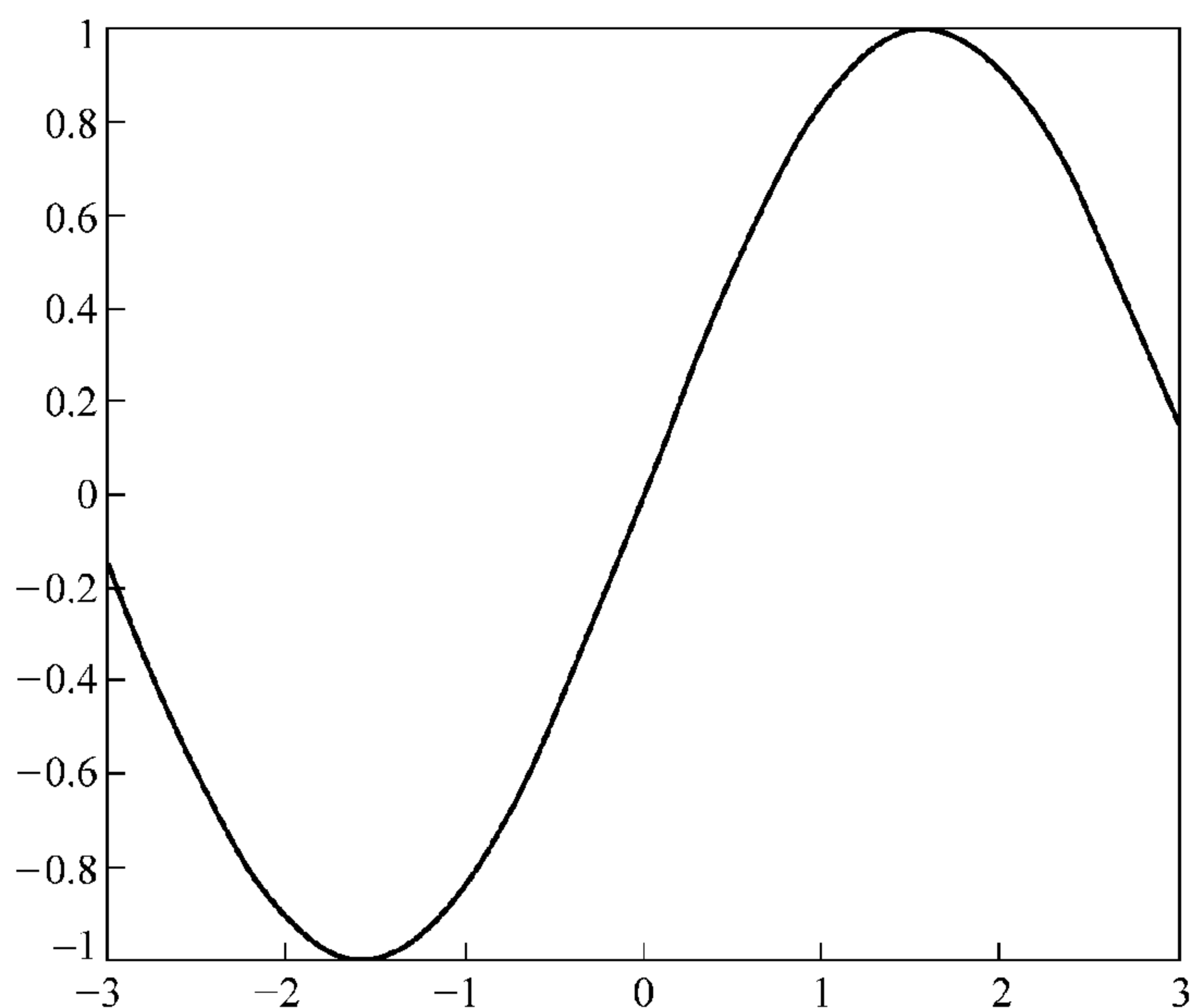


图 10.4 例 10.1 中的隶属度函数



注: $f(x)$ 和 $g(x) = \sin x$ 几乎是完全相同的。

图 10.5 所设计的模糊系统 $f(x)$ 和函数 $g(x) = \sin x$

因为 $\left\| \frac{\partial g}{\partial x_1} \right\|_{\infty} = \sup_{x \in U} |0.1 - 0.06x_2| = 0.16$, $\left\| \frac{\partial g}{\partial x_2} \right\|_{\infty} = \sup_{x \in U} |0.28 - 0.06x_1| = 0.34$, 所以, 由式(10.11)可知, 当 $h_1 = h_2 = 0.2$ 时有 $\|g - f\|_{\infty} \leq 0.16 \times 0.2 + 0.34 \times 0.2 = 0.1$ 。由此可得, 定义在 $[-1, 1]$ 的 11 个模糊集 A^j 的三角形隶属度函数为

$$\mu_{A^1}(x) = \mu_{A^1}(x; -1, -1, -0.8) \quad (10.24)$$

$$\mu_{A^{11}}(x) = \mu_{A^{11}}(x; 0.8, 1, 1) \quad (10.25)$$

$$\mu_{A^j}(x) = \mu_{A^j}(x; e^{j-1}, e^j, e^{j+1}), j = 2, 3, \dots, 10 \quad (10.26)$$

其中, $e^j = -1 + 0.2(j-1)$ 。模糊系统是由以下 $11 \times 11 = 121$ 条规则构造的:

如果 x_1 为 A^{i_1} 且 x_2 为 A^{i_2} , 则 y 为 $B^{i_1 i_2}$ (10.27)

其中, $i_1, i_2 = 1, 2, \dots, 11$, $B^{i_1 i_2}$ 的中心为 $\bar{y}^{i_1 i_2} = g(e^{i_1}, e^{i_2})$, 则最终的模糊系统为

$$f(x) = \frac{\sum_{i_1=1}^{11} \sum_{i_2=1}^{11} g(e^{i_1}, e^{i_2}) (\mu_{A^{i_1}}(x_1) \mu_{A^{i_2}}(x_2))}{\sum_{i_1=1}^{11} \sum_{i_2=1}^{11} (\mu_{A^{i_1}}(x_1) \mu_{A^{i_2}}(x_2))} \quad (10.28)$$

由例 10.2 可知, 需用 121 条规则来逼近函数 $g(x)$ 。然而, 真的需要这么多条规则吗? 或者说, 能否修改式(10.11)中的边界, 从而仅用较少的规则, 以相同的精度逼近同一个函数? 答案是肯定的, 这将在下一章中进行详细研究。

10.4 本章小结及导读

本章主要阐述了以下内容:

- 三类逼近问题(根据可获得的信息进行分类);
- 完备性概念、一致性概念和模糊集的排序以及它们在具有四边形隶属函数的模糊集中的应用;
- 按照给定的逼近精度, 如何设计一个模糊系统, 使其逼近某一给定函数;
- 提供逼近边界(10.11)的思想。

Ying[1994]和 Zeng 和 Singh[1995]对模糊系统的逼近精度做过分析。不过, 逼近精度问题仍是一个相对较新的专题, 参考文献很少。

10.5 习题

习题 10.1 令 $U = [a, b]$ 上的模糊集 $A^j (j = 1, 2, \dots, N)$ 为标准的、一致的和完备的模糊集, 其四边形隶属函数为 $\mu_{A^j}(x) = \mu_{A^j}(x; a_j, b_j, c_j, d_j)$ 。假定 $A^1 < A^2 < \dots < A^N$, 定义模糊集 B^j 的隶属函数为

$$\mu_{B^j}(x) = \frac{\mu_{A^j}(x)}{\sum_{i=1}^N \mu_{A^i}(x)} \quad (10.29)$$

证明

- $\mu_{B^j}(x) (j = 1, 2, \dots, N)$ 也是四边形隶属度函数;
- 模糊集 $B^j (j = 1, 2, \dots, N)$ 也是标准的、一致的和完备的模糊集;
- $B^1 < B^2 < \dots < B^N$;

(d) 如果 $\mu_{A_j}(x) = \mu_{A_j}(x; a_j, b_j, c_j, d_j)$ 是梯形隶属度函数, 其中 $c_i = a_{i+1}, d_i = b_{i+1}$ ($i = 1, 2, \dots, N-1$), 则 $\mu_{B_j}(x) = \mu_{A_j}(x)$ ($j = 1, 2, \dots, N-1$)。

习题 10.2 设计一个在 $U = [-1, 1]$ 上的模糊系统, 使其以精度 $\varepsilon = 0.1$ 一致地逼近 $g(x) = \sin(x\pi) + \cos(x\pi) + \sin(x\pi)\cos(x\pi)$ 。

习题 10.3 设计一个在 $U = [-1, 1] \times [-1, 1]$ 上的模糊系统, 使其以精度 $\varepsilon = 0.1$ 一致地逼近函数 $g(x) = \sin(x_1\pi) + \cos(x_2\pi) + \sin(x_1\pi)\cos(x_2\pi)$ 。

习题 10.4 将 10.2 节中的设计方法推广到 n 维输入的模糊系统。

习题 10.5 令 $U = [0, 1]^3$ 上的函数 $g(x)$ 由下式给出

$$g(x_1, x_2, x_3) = 1 + \sum_{k_1, k_2, k_3 \in K} x_1^{k_1} x_2^{k_2} x_3^{k_3} \quad (10.30)$$

式中, $K = \{k_1 k_2 k_3 \mid k_i = 0, 1; i = 1, 2, 3 \text{ 且 } k_1 + k_2 + k_3 > 0\}$ 。设计一个模糊系统以 $\varepsilon = 0.05$ 为逼近精度一致地逼近 $g(x)$ 。

习题 10.6 证明, 如果 10.2 节的设计程序中的模糊集 $A_1^1, A_1^2, \dots, A_i^{N_i}$ 是不完整的, 则模糊系统式(10.10)定义得不好。如果这些模糊集是不标准的或不一致的模糊集, 那么模糊系统式(10.10)定义得好吗?

习题 10.7 画出 $U = [-1, 1] \times [-1, 1]$ 上的模糊系统(10.28), 并将其与 $g(x) = 0.52 + 0.1x_1 + 0.28x_2 - 0.06x_1x_2$ 进行比较。

模糊系统的逼近特性 II

由第 10 章可知,采用定理 10.1 所定义的边界后,通常需要大量的规则来逼近某个简单的函数。如,例 10.2 就用 121 条规则去逼近一个二元平方函数。由式(10.11)可知,边界是 h_i 的线性函数。由于 h_i 通常很小,所以如果边界是 h_i^2 的线性函数,则这一边界将会比式(10.11)中的边界小的多。即,如果能得到比式(10.11)中更小的边界,就能用更少的规则以相同的精度去逼近同一个函数了。

正如定理 10.1 证明的那样,在逼近理论中(Powell[1981]),如果 $g(x)$ 是一个 U 上的给定函数, $U^{i_1 i_2} (i_1 = 1, 2, \dots, N_1, i_2 = 1, 2, \dots, N_2)$ 是 U 的一个划分,则当 $\|g - f\|_\infty \leq M_g h^k$ 时,可认为 $f(x)$ 是 $g(x)$ 的第 k 阶精确逼近器,这里, M_g 是一个与 g 有关的常数, h 是这一区域划分的范数,本例中 h 为 $\max(h_1, h_2)$ 。本章将首先设计一个作为二阶精确逼近器的模糊系统。

11.1 具有二阶逼近精度的模糊系统

首先逐步地设计一个模糊系统,然后研究其逼近精度。与第 10 章一样,为使符号简便,这里只考虑二维输入的模糊系统,其方法和结论对 n 维输入的模糊系统仍然有效。要设计的模糊系统与 10.2 节中相同。

设计具有二阶逼近精度的模糊系统

步骤 1 在 $[\alpha_i, \beta_i]$ 上定义 $N_i (i = 1, 2)$ 个模糊集 $A_i^1, A_i^2, \dots, A_i^{N_i}$, 这些模糊集均为标准的、一致的和完备的模糊集,其三角形隶属度函数为

$$\mu_{A_i^1}(x_i) = \mu_{A_i^1}(x_i; e_i^1, e_i^1, e_i^2) \quad (11.1)$$

$$\mu_{A_i^j}(x_i) = \mu_{A_i^j}(x_i; e_i^{j-1}, e_i^j, e_i^{j+1}), j = 2, 3, \dots, N_i - 1 \quad (11.2)$$

$$\mu_{A_i^{N_i}}(x_i) = \mu_{A_i^{N_i}}(x_i; e_i^{N_i-1}, e_i^{N_i}, e_i^{N_i}) \quad (11.3)$$

其中, $i = 1, 2, \alpha_i = e_i^1 < e_i^2 < \dots < e_i^{N_i} = \beta_i$ 。图 11.1 给出了 $N_1 = 4, N_2 = 5, \alpha_1 = \alpha_2 = 0, \beta_1 = \beta_2 = 1$ 时的一个例子。

步骤 2 和步骤 3 同 10.2 节中设计程序的步骤 2 和步骤 3。即,所构造的模糊系统是由式(10.10)给定的,其中 $\bar{y}^{i_1 i_2}$ 由式(10.9)给定, $A_1^{i_1}, A_2^{i_2}$ 由式(11.1) ~ 式(11.3)给定。

因为根据以上步骤设计的模糊系统只是 10.2 节中设计的模糊系统的一个特例,所以定

理 10.1 对此模糊系统仍然有效。下面的定理给出了更好的结论。

定理 11.1 令 $f(x)$ 为通过以上三步设计的模糊系统。如果 $g(x)$ 在 U 上是二次连续可微的, 则下式成立

$$\|g - f\|_{\infty} \leq \frac{1}{8} \left[\left\| \frac{\partial^2 g}{\partial x_1^2} \right\|_{\infty} h_1^2 + \left\| \frac{\partial^2 g}{\partial x_2^2} \right\|_{\infty} h_2^2 \right] \quad (11.4)$$

式中, $h_i = \max_{1 \leq j \leq N_i-1} |e_i^{j+1} - e_i^j|$ ($i = 1, 2$)

证明 与定理 10.1 的证明一样, 将 U 分割成 U

$$= \bigcup_{i_1=1}^{N_1-1} \bigcup_{i_2=1}^{N_2-1} U^{i_1 i_2}, \text{ 其中, } U^{i_1 i_2} = [e_1^{i_1}, e_1^{i_1+1}] \times [e_2^{i_2}, e_2^{i_2+1}].$$

所以, 对任意的 $x \in U$, 都存在 $U^{i_1 i_2}$ 使得 $x \in U^{i_1 i_2}$ 。现假

设 $x \in U^{i_1 i_2}$, 则由模糊集 $A_1^1, A_1^2, \dots, A_1^{N_1}$ ($i = 1, 2$) 的一致性和完备性, 可把模糊系统简化为(与式(10.13)相同)

$$f(x) = \sum_{j_1=i_1}^{i_1+1} \sum_{j_2=i_2}^{i_2+1} \left[\frac{\mu_{A_1^{j_1}}(x_1) \mu_{A_2^{j_2}}(x_2)}{\sum_{j_1=i_1}^{i_1+1} \sum_{j_2=i_2}^{i_2+1} \mu_{A_1^{j_1}}(x_1) \mu_{A_2^{j_2}}(x_2)} \right] g(e_1^{i_1}, e_2^{i_2}) \quad (11.5)$$

由于 $\mu_{A_i^j}(x_i)$ 是由式(11.1) ~ 式(11.3)给定的特定三角形隶属度函数, 所以, 当 $i = 1, 2$ 时, 有

$$\mu_{A_i^{j_1}}(x_i) + \mu_{A_i^{j_1+1}}(x_i) = 1 \quad (11.6)$$

由此可得

$$\sum_{j_1=i_1}^{i_1+1} \sum_{j_2=i_2}^{i_2+1} \mu_{A_1^{j_1}}(x_1) \mu_{A_2^{j_2}}(x_2) = \sum_{j_1=i_1}^{i_1+1} \mu_{A_1^{j_1}}(x_1) \left[\sum_{j_2=i_2}^{i_2+1} \mu_{A_2^{j_2}}(x_2) \right] = 1 \quad (11.7)$$

于是模糊系统(11.5)可以简化为

$$f(x) = \sum_{j_1=i_1}^{i_1+1} \sum_{j_2=i_2}^{i_2+1} [\mu_{A_1^{j_1}}(x_1) \mu_{A_2^{j_2}}(x_2)] g(e_1^{i_1}, e_2^{i_2}) \quad (11.8)$$

令 $C^2(U^{i_1 i_2})$ 为 $U^{i_1 i_2}$ 上的所有二次连续可微函数的集合, 并将 $C^2(U^{i_1 i_2})$ 上的线性算子 L_1 和 L_2 定义为

$$(L_1 g)(x) = \sum_{j_1=i_1}^{i_1+1} (\mu_{A_1^{j_1}}(x_1)) g(e_1^{j_1}, x_2) \quad (11.9)$$

$$(L_2 g)(x) = \sum_{j_2=i_2}^{i_2+1} (\mu_{A_2^{j_2}}(x_2)) g(x_1, e_2^{j_2}) \quad (11.10)$$

由于 $\mu_{A_1^{j_1}}(x_1)$ 和 $\mu_{A_2^{j_2}}(x_2)$ 是 $U^{i_1 i_2}$ 上的线性函数, 所以它们是二次连续可微的。因此, 由

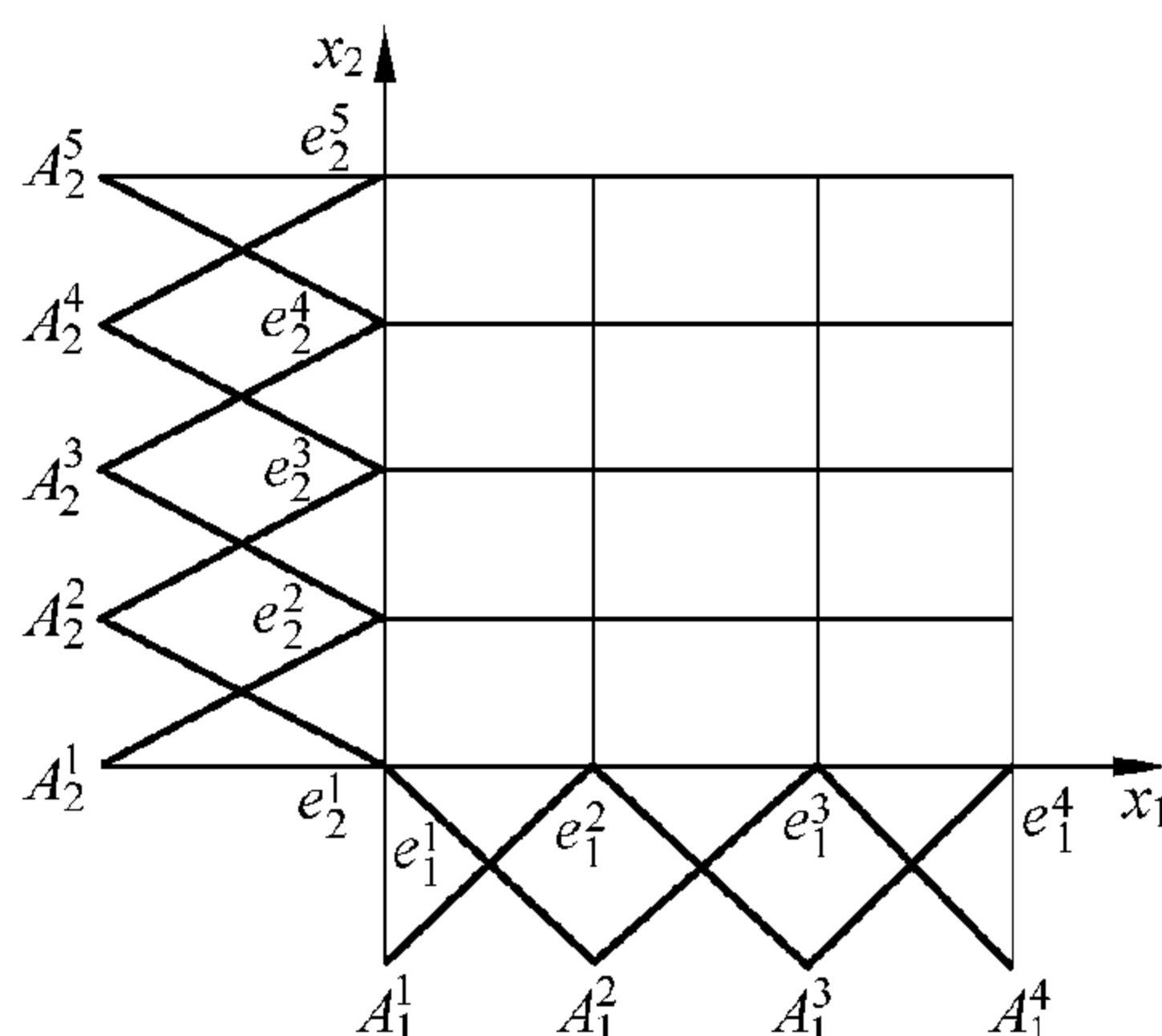


图 11.1 模糊集合的一个例子

$g \in C^2(U^{i_1 i_2})$ 可知, $L_1 g \in C^2(U^{i_1 i_2}), L_2 g \in C^2(U^{i_1 i_2})$ 。由式(11.9)和式(11.6), 可得

$$\|L_1 g\|_\infty \leq \|g\|_\infty \left| \sum_{j_1=i_1}^{i_1+1} \mu_{A_1^{j_1}}(x_1) \right| = \|g\|_\infty \quad (11.11)$$

合并式(11.9)和式(11.10)并观察式(11.8), 可得

$$(L_1 L_2 g)(x) = \sum_{j_1=i_1}^{i_1+1} \mu_{A_1^{j_1}}(x_1) \left[\sum_{j_2=i_2}^{i_2+1} \mu_{A_2^{j_2}}(x_2) g(e_1^{j_1}, e_2^{j_2}) \right] = f(x) \quad (11.12)$$

所以, 由式(11.11)和式(11.12), 可得

$$\begin{aligned} \|g - f\|_\infty &= \|g - L_1 L_2 g\|_\infty \\ &\leq \|g - L_1 g\|_\infty + \|L_1(g - L_2 g)\|_\infty \\ &\leq \|g - L_1 g\|_\infty + \|g - L_2 g\|_\infty \end{aligned} \quad (11.13)$$

因为 $x \in U^{i_1 i_2} = [e_1^{i_1}, e_1^{i_1+1}] \times [e_2^{i_2}, e_2^{i_2+1}]$, 所以由单变量线性插值(Powell[1981])中的结论, 可得

$$\begin{aligned} \|g - L_1 g\|_\infty &= \left\| \sum_{j_1=i_1}^{i_1+1} \mu_{A_1^{j_1}}(x_1) \left[g(x_1, x_2) - g(e_1^{j_1}, x_2) \right] \right\|_\infty \\ &\leq \left\| g(x_1, x_2) - g(e_1^{j_1}, x_2) \right\|_\infty \\ &\leq \frac{1}{8} (e_1^{i_1+1} - e_1^{i_1})^2 \left\| \frac{\partial^2 g}{\partial x_1^2} \right\|_\infty \end{aligned} \quad (11.14)$$

同理可得

$$\|g - L_2 g\|_\infty \leq \frac{1}{8} (e_2^{i_2+1} - e_2^{i_2})^2 \left\| \frac{\partial^2 g}{\partial x_2^2} \right\|_\infty \quad (11.15)$$

将式(11.14)和式(11.15)代入式(11.13)并由 h_i 的定义, 即可得到式(11.14)。

由定理 11.1 可知, 如果选择特殊的三角形隶属度函数, 就能得到一个二阶精确逼近器。现设计一个模糊系统来逼近例 10.1 和例 10.2 中的同一个函数 $g(x)$ [用新边界(11.4)]。将会发现, 用较少的规则可达到相同的精度。

例 11.1 现在采用边界(11.4), 其余与例 10.1 都相同。由于 $\left\| \frac{\partial^2 g}{\partial x_2^2} \right\|_\infty = 1$, 所以, 由式

(11.4)可知, 如果选择 $h = 1$, 则有 $\|g - f\|_\infty \leq \frac{1}{8} < \varepsilon$ 。因此, 定义 7 个形如式(11.1) ~ 式(11.3)的模糊集 A^j , 其中 $e^j = -3 + (j-1)$ ($j=1, 2, \dots, 7$)。于是可得, 所设计的模糊系统为

$$f(x) = \frac{\sum_{j=1}^7 \sin(e^j) \mu_{A^j}(x)}{\sum_{j=1}^7 \mu_{A^j}(x)} \quad (11.16)$$

比较式(11.16)和式(10.23),可以看出,规则的数量由 31 降为 7,但精度仍保持不变。式(11.16)中 $f(x)$ 的图形见图 11.2,图中 $g(x) = \sin(x)$ 。

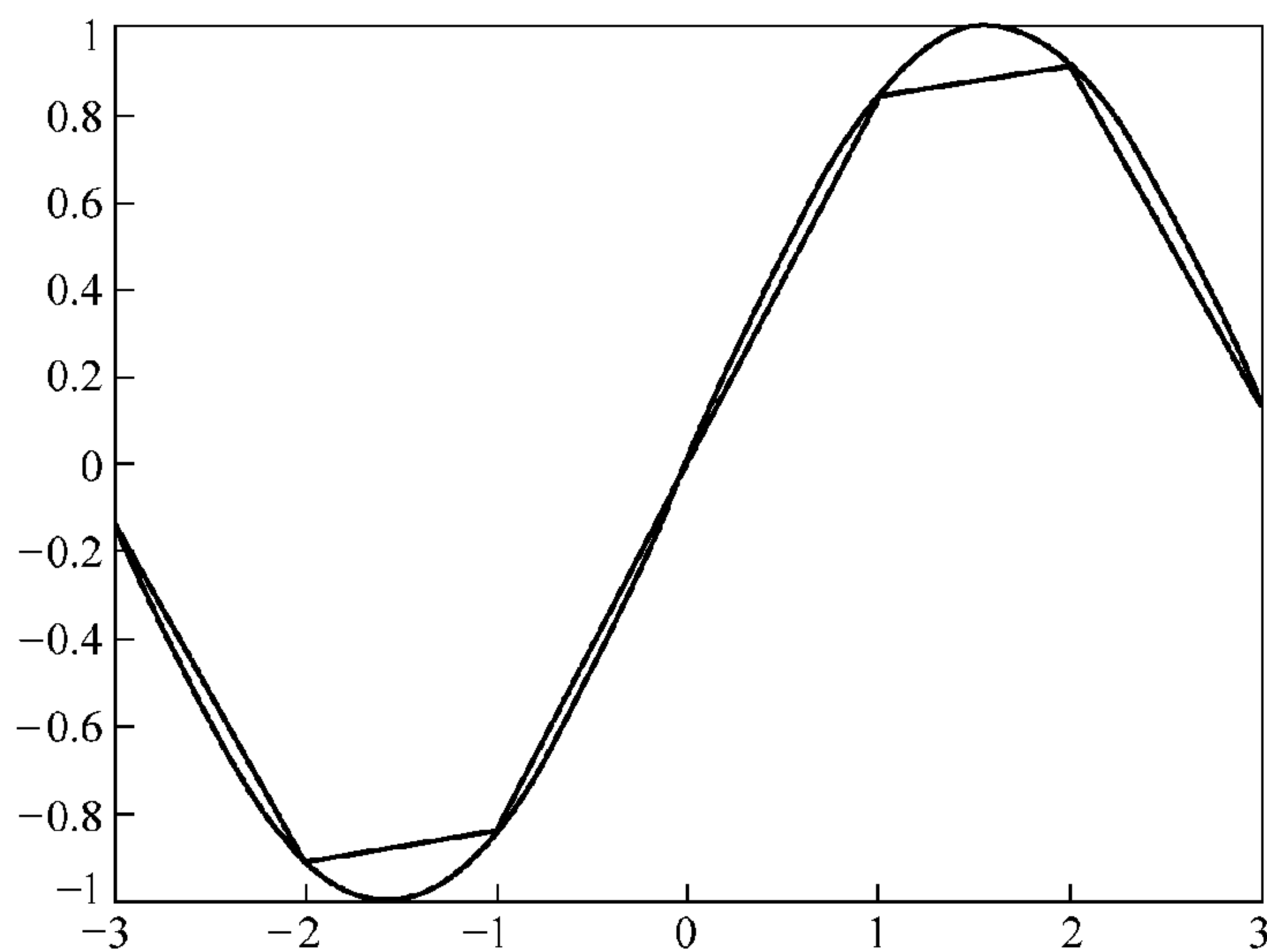


图 11.2 由式(11.16)设计的模糊系统和要逼近的函数 $g(x) = \sin(x)$

例 11.2 采用边界(11.4),其余与例 10.2 都相同。因为 $\frac{\partial^2 g}{\partial x_i^2} = 0 (i = 1, 2)$,所以,由式(11.4)可知, $f(x) = g(x) (x \in U)$ 。事实上,只要选定 $h_i = 2, e_i^1 = -1, e_i^2 = 1 (i = 1, 2)$,即 $N_1 = N_2 = 2$,就会得到如下模糊系统:

$$f(x) = \frac{\sum_{i_1=1}^2 \sum_{i_2=1}^2 g(e_1^{i_1}, e_2^{i_2}) (\mu_{A_{1i_1}}(x_1) \mu_{A_{2i_2}}(x_2))}{\sum_{i_1=1}^2 \sum_{i_2=1}^2 \mu_{A_{1i_1}}(x_1) \mu_{A_{2i_2}}(x_2)} \quad (11.17)$$

其中,隶属度函数由式(11.1)~式(11.3)给定。当 $i = 1, 2$ 和 $x \in U$ 时,对于该特例有

$$\mu_{A_1^1}(x_i) = \mu_{A_1^1}(x_i; -1, -1, 1) = \frac{1}{2}(1 - x_i) \quad (11.18)$$

$$\mu_{A_2^1}(x_i) = \mu_{A_2^1}(x_i; -1, 1, 1) = \frac{1}{2}(1 + x_i) \quad (11.19)$$

$g(e_1^1, e_2^1) = g(-1, -1) = 0.08, g(e_1^1, e_2^2) = g(-1, 1) = 0.76, g(e_1^2, e_2^1) = g(1, -1) = 0.4,$
 $g(e_1^2, e_2^2) = g(1, 1) = 0.84$ 。将上述结果代入式(11.17),可得

$$f(x) = \left[\frac{0.08}{4}(1 - x_1)(1 - x_2) + \frac{0.76}{4}(1 - x_1)(1 + x_2) + \frac{0.4}{4}(1 + x_1)(1 - x_2) + \frac{0.84}{4}(1 + x_1)(1 + x_2) \right] / \left[\frac{1}{4}(1 - x_1)(1 - x_2) + \frac{1}{4}(1 - x_1)(1 + x_2) + \frac{1}{4}(1 + x_1)(1 - x_2) + \frac{1}{4}(1 + x_1)(1 + x_2) \right]$$

$$= 0.52 + 0.1x_1 + 0.28x_2 - 0.06x_1x_2 \quad (11.20)$$

上式中的 $f(x)$ 恰好等于 $g(x)$, 这也证明了 $f(x)$ 可以准确拟合 $g(x)$ 这一结论。在例 10.2 中, 用 121 条规则来构造模糊系统, 而在本例中仅用 4 条规则就达到了零拟合误差。

为了推广例 11.2, 由式(11.4)可发现, 对任意满足 $\left\| \frac{\partial^2 g}{\partial x_i^2} \right\|_{\infty} = 0$ 的函数来说, 所设计的模糊系统 $f(x)$ 都可以通过这三个步骤来准确拟合 $g(x)$ 。即, 对所有的 $x \in U$, 有 $f(x) = g(x)$ 。请看下面的定理 11.1 的推论。

推论 11.1 令 $f(x)$ 为按本节中的三个步骤所设计的模糊系统, 如果函数 $g(x)$ 为下述形式, 则对于所有 $x \in U$ 都有 $f(x) = g(x)$ 。

$$g(x) = \sum_{k_1=0}^1 \sum_{k_2=0}^1 a_{k_1 k_2} x_1^{k_1} x_2^{k_2} \quad (11.21)$$

式中, $a_{k_1 k_2}$ 为常数。

证明 因为 $g(x)$ 满足 $\frac{\partial^2 g}{\partial x_1^2} = \frac{\partial^2 g}{\partial x_2^2} = 0$, 所以由式(11.4)可以立刻得出结论。

11.2 带有最大值解模糊器的模糊系统的逼近精度

在第 9 章中学到了, 带有最大值解模糊器的模糊系统同带有中心平均解模糊器的模糊系统有很大的不同。本节将研究带有最大值解模糊器的模糊系统的逼近性质。

与 11.1 节及 10.2 节中的方法类似, 首先设计一个特定的模糊系统, 然后研究其性质。

设计有最大值解模糊器的模糊系统

步骤 1 与 11.1 节中设计程序的步骤 1 相同。

步骤 2 与 11.1 节中设计程序的步骤 2 相同。

步骤 3 采用乘积推理机(7.23), 单值模糊器(8.1)和最大值解模糊器(8.23), 根据形如规则(10.8)的 $N_1 \times N_2$ 条规则构造模糊系统 $f(x)$ 。根据引理 9.4 可知, 该模糊系统为

$$f(x) = \bar{y}^{i_1^* i_2^*} = g(e_1^{i_1^*}, e_2^{i_2^*}) \quad (11.22)$$

其中, 当 $i_1 = 1, 2, \dots, N_1$; $i_2 = 1, 2, \dots, N_2$ 时, i_1^*, i_2^* 使得

$$\mu_{A_1^{i_1^*}}(x_1) \mu_{A_2^{i_2^*}}(x_2) \geq \mu_{A_1^{i_1}}(x_1) \mu_{A_2^{i_2}}(x_2) \quad (11.23)$$

下面的定理证明了上面所设计的模糊系统逼近 $g(x)$ 的精度是一阶的。

定理 11.2 令 $f(x)$ 为根据上述三步所设计的模糊系统(11.22), 如果 $g(x)$ 在 $U = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2]$ 上是连续可微的, 则有

$$\|g - f\|_{\infty} \leq \left\| \frac{\partial g}{\partial x_1} \right\|_{\infty} h_1 + \left\| \frac{\partial g}{\partial x_2} \right\|_{\infty} h_2 \quad (11.24)$$

其中, $h_i = \max_{1 \leq j \leq N_i-1} |e_i^{j+1} - e_i^j|, i = 1, 2$ 。

证明 与定理 11.1 的证明相同, 把区域 U 分为

$$U = \bigcup_{i_1=1}^{N_1-1} \bigcup_{i_2=1}^{N_2-1} U^{i_1 i_2}$$

这里

$$U^{i_1 i_2} = [e_1^{i_1}, e_1^{i_1+1}] \times [e_2^{i_2}, e_2^{i_2+1}]$$

然后, 再把区域 $U^{i_1 i_2}$ 分为

$$U^{i_1 i_2} = U_{00}^{i_1 i_2} \cup U_{01}^{i_1 i_2} \cup U_{10}^{i_1 i_2} \cup U_{11}^{i_1 i_2}$$

其中

$$U_{00}^{i_1 i_2} = \left[e_1^{i_1}, \frac{1}{2} (e_1^{i_1} + e_1^{i_1+1}) \right] \times \left[e_2^{i_2}, \frac{1}{2} (e_2^{i_2} + e_2^{i_2+1}) \right]$$

$$U_{01}^{i_1 i_2} = \left[e_1^{i_1}, \frac{1}{2} (e_1^{i_1} + e_1^{i_1+1}) \right] \times \left[\frac{1}{2} (e_2^{i_2} + e_2^{i_2+1}), e_2^{i_2+1} \right]$$

$$U_{10}^{i_1 i_2} = \left[\frac{1}{2} (e_1^{i_1} + e_1^{i_1+1}), e_1^{i_1+1} \right] \times \left[e_2^{i_2}, \frac{1}{2} (e_2^{i_2} + e_2^{i_2+1}) \right]$$

$$U_{11}^{i_1 i_2} = \left[\frac{1}{2} (e_1^{i_1} + e_1^{i_1+1}), e_1^{i_1+1} \right] \times \left[\frac{1}{2} (e_2^{i_2} + e_2^{i_2+1}), e_2^{i_2+1} \right]$$

如图 11.3 所示。所以, 对任意 $x \in U$, 都存在 $U_{pq}^{i_1 i_2}$ ($p, q = 0$ 或 1) 使得 $x \in U_{pq}^{i_1 i_2}$ 。如果 x 在区域 $U_{pq}^{i_1 i_2}$ 中, 则借助图 11.3 可知, $\mu_{A_1^{i_1+p}}(x_1) > 0.5, \mu_{A_2^{i_2+q}}(x_2) > 0.5$, 其他所有的隶属度函数值都小于 0.5 。所以, 由式 (11.22) 和式 (11.23), 可得

$$f(x) = g(e_1^{i_1+p}, e_2^{i_2+q}) \quad (11.25)$$

应用均值定理并由 $x \in U^{i_1 i_2}$, 可得

$$\begin{aligned} & |g(x) - f(x)| \\ &= \left| g(x_1, x_2) - g(e_1^{i_1+p}, e_2^{i_2+q}) \right| \\ &\leq \left\| \frac{\partial g}{\partial x_1} \right\|_{\infty} |x_1 - e_1^{i_1+p}| + \left\| \frac{\partial g}{\partial x_2} \right\|_{\infty} |x_2 - e_2^{i_2+q}| \end{aligned}$$

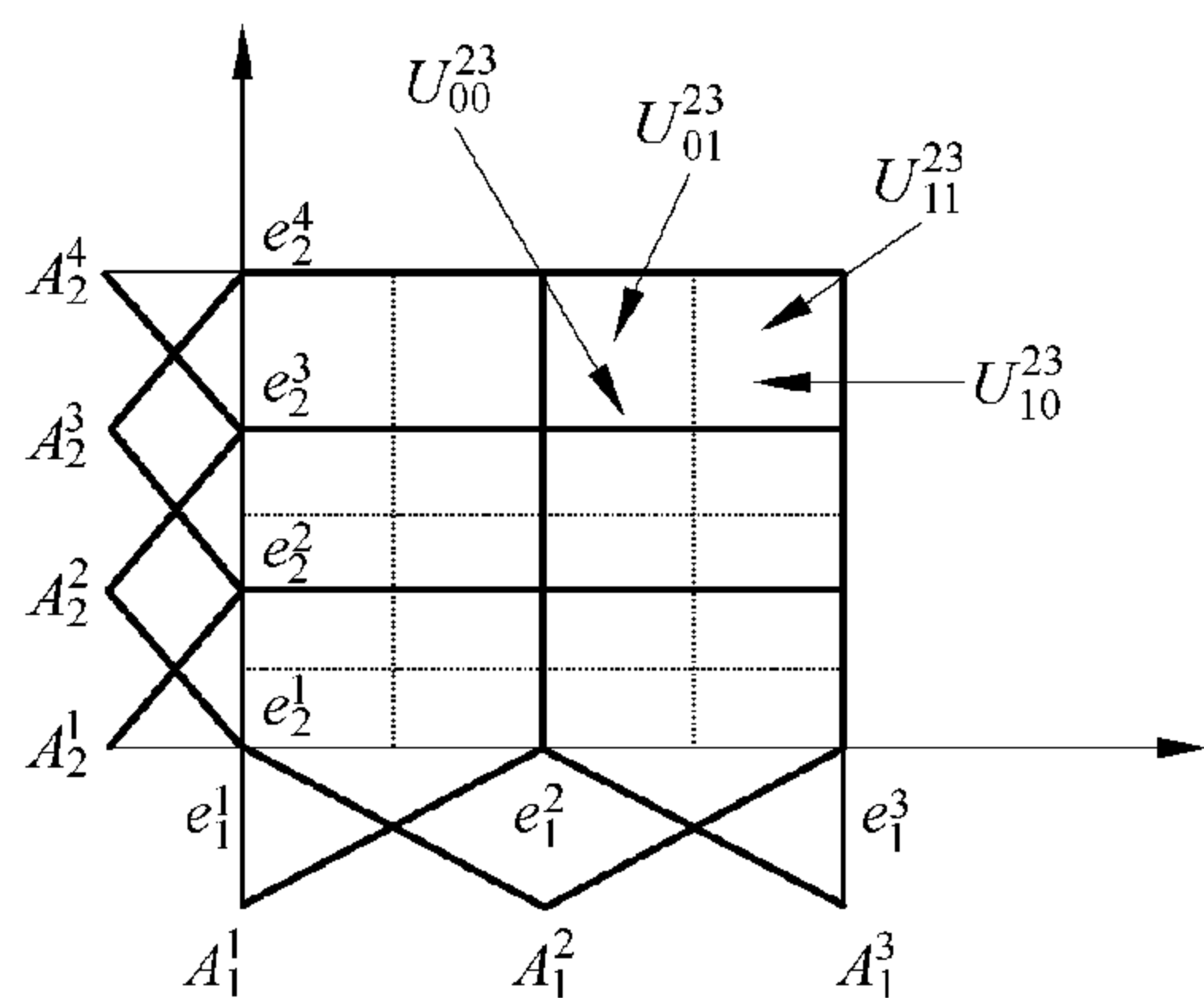


图 11.3 $U^{i_1 i_2}$ 分割成子集的一个例子

$$\leq \left\| \frac{\partial g}{\partial x_1} \right\|_{\infty} \left| e_1^{i_1+1} - e_1^{i_1} \right| + \left\| \frac{\partial g}{\partial x_2} \right\|_{\infty} \left| e_2^{i_2+1} - e_2^{i_2} \right| \quad (11.26)$$

如果 x 在 $U_{pq}^{i_1 i_2}$ 的边界上, 则借助图 11.3 可知, $f(x)$ 可以从至多有四个元素的集合 $\left\{ g\left(e_1^{i_1}, e_2^{i_2}\right), g\left(e_1^{i_1}, e_2^{i_2+1}\right), g\left(e_1^{i_1+1}, e_2^{i_2}\right), g\left(e_1^{i_1+1}, e_2^{i_2+1}\right) \right\}$ 中任意取值, 所以, 式(11.26)在此情况下仍然成立。最后, 由式(11.26)可知式(11.24)成立。

由定理 11.2 可以看出, 带有乘积推理机、单值模糊器和最大值解模糊器的模糊系统是万能逼近器。实际上, 根据式(11.24)可知, 当 h_1 和 h_2 足够小时, 对任意 $\varepsilon > 0$, 都有 $\|g - f\|_{\infty} < \varepsilon$ 成立。

现在, 用模糊系统(11.22)来逼近例 11.1 和例 11.2 中的函数 $g(x)$ 。

例 11.3 除了应用模糊系统(11.22)外, 其余与例 10.1 都相同。因为 $\left\| \frac{\partial g}{\partial x} \right\|_{\infty} = 1$, 所以选定 $h = 0.2$, 定义 31 个形如式(10.20) ~ 式(10.22)的模糊集 A^j (见图 10.4)。令 $e^j = -3 + 0.2(j-1)$, 则模糊系统(11.22)变为

$$f(x) = \begin{cases} \sin(-3) & x \in [-3, -2.8] \\ \sin(e^j) & x \in (e^j - 0.1, e^j + 0.1), j = 2, 3, \dots, 30 \\ \sin(3) & x \in (2.8, 3] \end{cases} \quad (11.27)$$

其图形见图 11.4。

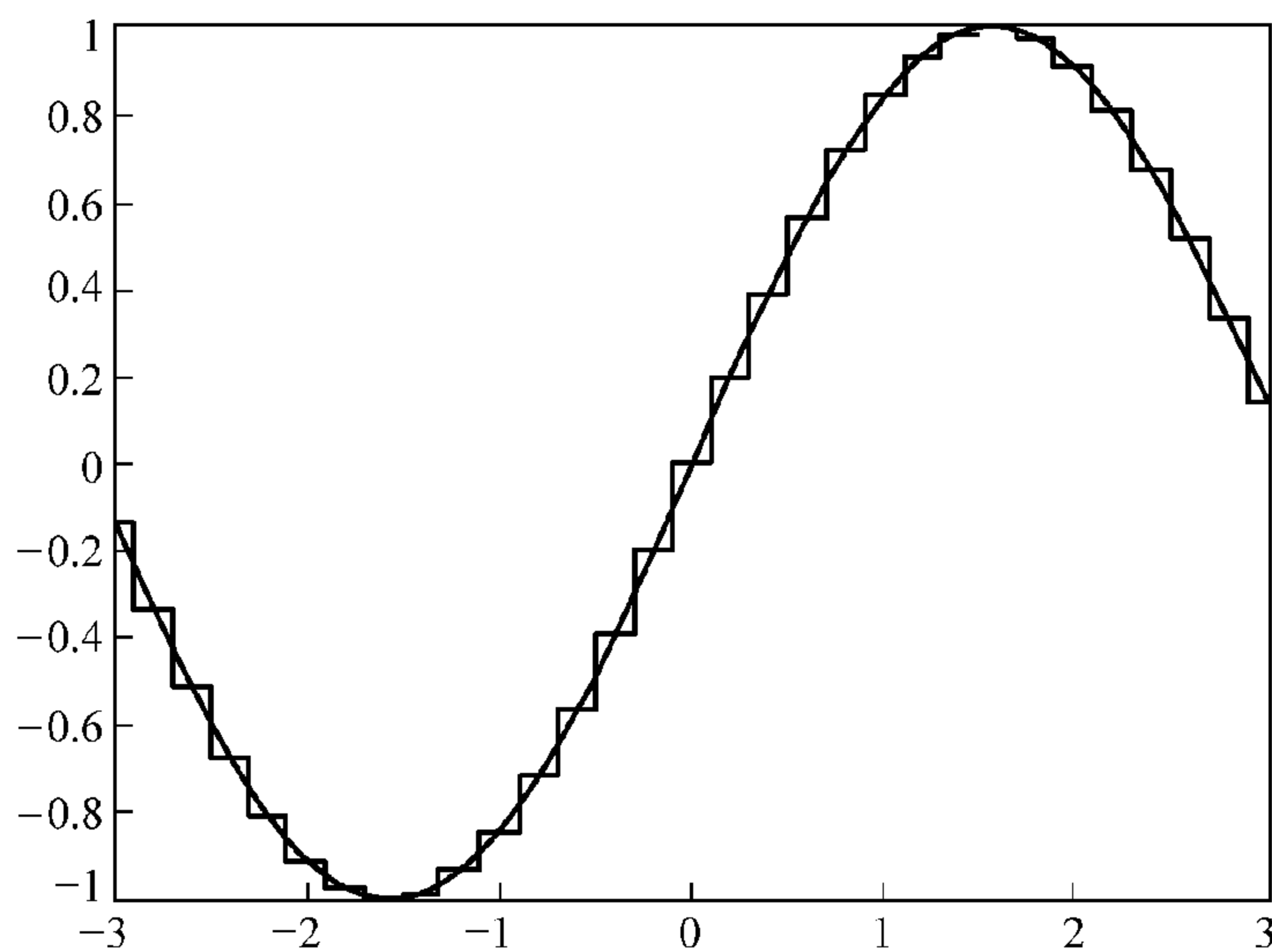


图 11.4 在例 11.3 中设计的模糊系统(11.22)及其所逼近的函数 $g(x) = \sin(x)$

例 11.4 模糊系统采用式(11.22), 其余与例 10.2 都相同。和例 10.2 一样, 选择 $h_1 = h_2 = 0.2$, 并在 $[-1, 1]$ 上定义 11 个模糊集 A^j [它们由式(10.24) ~ 式(10.26)给定]。用如

规则(10.27)的 121 条规则来构造模糊系统。在该例中, $e^j = -1 + 0.2(j-1), j = 1, 2, \dots, 11$; $U^{i_1 i_2} = [e^{i_1}, e^{i_1+1}] \times [e^{i_2}, e^{i_2+1}], i_1, i_2 = 1, 2, \dots, 10$ 。正如图 11.3 中所示的那样, 进一步

分解 $U^{i_1 i_2}$ 为 $U^{i_1 i_2} = \bigcup_{p=0}^1 \bigcup_{q=0}^1 U_{pq}^{i_1 i_2}$, 其中

$$U_{00}^{i_1 i_2} = \left[e^{i_1}, \frac{1}{2} (e^{i_1} + e^{i_1+1}) \right] \times \left[e^{i_2}, \frac{1}{2} (e^{i_2} + e^{i_2+1}) \right]$$

$$U_{01}^{i_1 i_2} = \left[e^{i_1}, \frac{1}{2} (e^{i_1} + e^{i_1+1}) \right] \times \left[\frac{1}{2} (e^{i_2} + e^{i_2+1}), e^{i_2+1} \right]$$

$$U_{10}^{i_1 i_2} = \left[\frac{1}{2} (e^{i_1} + e^{i_1+1}), e^{i_1+1} \right] \times \left[e^{i_2}, \frac{1}{2} (e^{i_2} + e^{i_2+1}) \right]$$

$$U_{11}^{i_1 i_2} = \left[\frac{1}{2} (e^{i_1} + e^{i_1+1}), e^{i_1+1} \right] \times \left[\frac{1}{2} (e^{i_2} + e^{i_2+1}), e^{i_2+1} \right]$$

则模糊系统(11.22)变为

$$f(x) = g(e^{i_1+p}, e^{i_2+q}), x \in U_{pq}^{i_1 i_2} \quad (11.28)$$

上式是通过以下两个步骤进行计算的: (i) 对于给定的 $x \in U$, 确定 i_1, i_2, p, q 使其满足 $x \in U_{pq}^{i_1 i_2}$; (ii) $f(x)$ 等于 $g(e^{i_1+p}, e^{i_2+q})$ 。

在 11.1 节中已经证明了带有中心平均解模糊器的模糊系统是二阶精度逼近器。定理 11.2 说明了, 带有最大值解模糊器的模糊系统是一阶精度逼近器。因此, 我们很自然地想知道带有最大值解模糊器的模糊系统是不是二阶逼近器? 很不幸, 下面的例子说明了, 它并不是二阶精度逼近器。

例 11.5 令 $U = [0, 1]$ 上 $g(x) = x, \mu_{A_i}(x) = \mu_{A_i}(x; e^{i-1}, e^i, e^{i+1})$, 其中 $e^0 = 0, e^{N+1} = 1, e^i = \frac{i-1}{N-1} (i = 1, 2, \dots, N)$, N 为任意正整数。所以, 在此情况下, 有 N 条规则, 且 $h = \frac{1}{N-1}$ 。令 $U^i = [e^i, e^{i+1}] (i = 1, 2, \dots, N-1)$, $f(x)$ 为模糊系统(11.22)。如果 $x \in U^i$, 则有

$$\max_{x \in U^i} |g(x) - f(x)| = \max_{x \in [e^i, e^{i+1}]} |x - g(e^i) \text{ 或 } g(e^{i+1})| = \frac{1}{2}(e^{i+1} - e^i) = \frac{1}{2}h \quad (11.29)$$

由于对任意正整数 N 都有 $h \left(= \frac{1}{N-1} \right) \geq h^2$, 所以, 模糊系统(11.22)不能以二阶精度逼近简单函数 $g(x) = x$ 。由此可归纳出, 带有最大值解模糊器的模糊系统不是二阶精度逼近器。因此, 带有中心平均解模糊器的模糊系统优于带有最大值解模糊器的模糊系统。

11.3 本章小结及导读

本章主要阐述了以下内容:

- 用二阶边界(11.4)按所需精度设计模糊系统;
- 设计带有最大值解模糊器的模糊系统并使其满足逼近精度的要求;
- 证明了带有中心平均解模糊器的模糊系统具有二阶边界(定理 11.1),带有最大值解模糊器的模糊系统具有一阶边界(定理 11.2)。

再次重申,本章中讨论的专题可以参考的文献极少。最相关的文章是 Ying[1994], Zeng 和 Singh[1995]。

11.4 习题

习题 11.1 分别用一阶边界(10.11)和二阶边界(11.4)设计两个具有中心平均解模糊器的模糊系统,使其一致地逼近 $U = [-1, 1] \times [-1, 1]$ 上的函数 $g(x_1, x_2) = \frac{1}{3 + x_1 + x_2}$, 所达到的逼近精度为 $\varepsilon = 0.1$ 。画出所设计的两个模糊系统,并对二者进行比较。

习题 11.2 当 $g(x_1, x_2) = \frac{1}{1 + x_1^2 + x_2^2}$ 时,重做习题 11.1。

习题 11.3 设计一个具有最大值解模糊器的模糊系统,使其以 $\varepsilon = 0.1$ 的精度一致地逼近习题 11.1 中 $U = [-1, 1] \times [-1, 1]$ 上的 $g(x_1, x_2)$ 。画出所设计的模糊系统。

习题 11.4 当 $g(x_1, x_2)$ 为习题 11.2 中的 $g(x_1, x_2)$ 时,重做习题 11.3。

习题 11.5 将 11.1 节中的设计程序推广至 n 维输入的模糊系统并证明所设计的模糊系统 $f(x)$ 满足

$$\|g - f\|_{\infty} \leq \frac{1}{8} \left(\sum_{i=1}^n \left\| \frac{\partial^2 g}{\partial x_i^2} \right\|_{\infty} h_i^2 \right) \quad (11.30)$$

习题 11.6 用图来证明式(11.29)是成立的。

第三部分

根据输入—输出数据设计模糊系统

模糊系统是用于描述人类知识的。所以,重要的是,人类知识通常采用什么形式来表述?大体上讲,关于某个特定工程问题的知识可分为两类:显性知识和隐性知识。显性知识是指那种能用清晰的语言表达出来的知识,而隐性知识是指那种专家知道去做什么,但无法恰当地用语言表达出怎样去做的知识。例如,有经验的卡车司机知道怎样在非常困难的情况下驾驶汽车(他们具有隐性知识),但是他们很难用精确的语言对其进行描述,即使他们对其进行了语言描述,该描述通常对于完成这一任务仍然是不完整的和不充分的。

对于显性知识,可以简单地让专家将该知识描述为模糊 IF-THEN 规则,从而将其嵌入模糊系统。对于隐性知识,可以让人类专家去阐述这种知识,然后把其作为一个黑箱,测量其输入和输出。也就是说,可以将该知识放到一个输入—输出数据对的集合中。从这个角度看,隐性知识是可以转化为一个输入—输出数据对集合的(见图 12.1)。因此,根据输入—输出数据对构造模糊系统的过程,也就是将隐性知识嵌入模糊系统的过程。

本部分(第 12 章至第 15 章)将开发一些方法,用这些方法来实现根据输入—输出数据

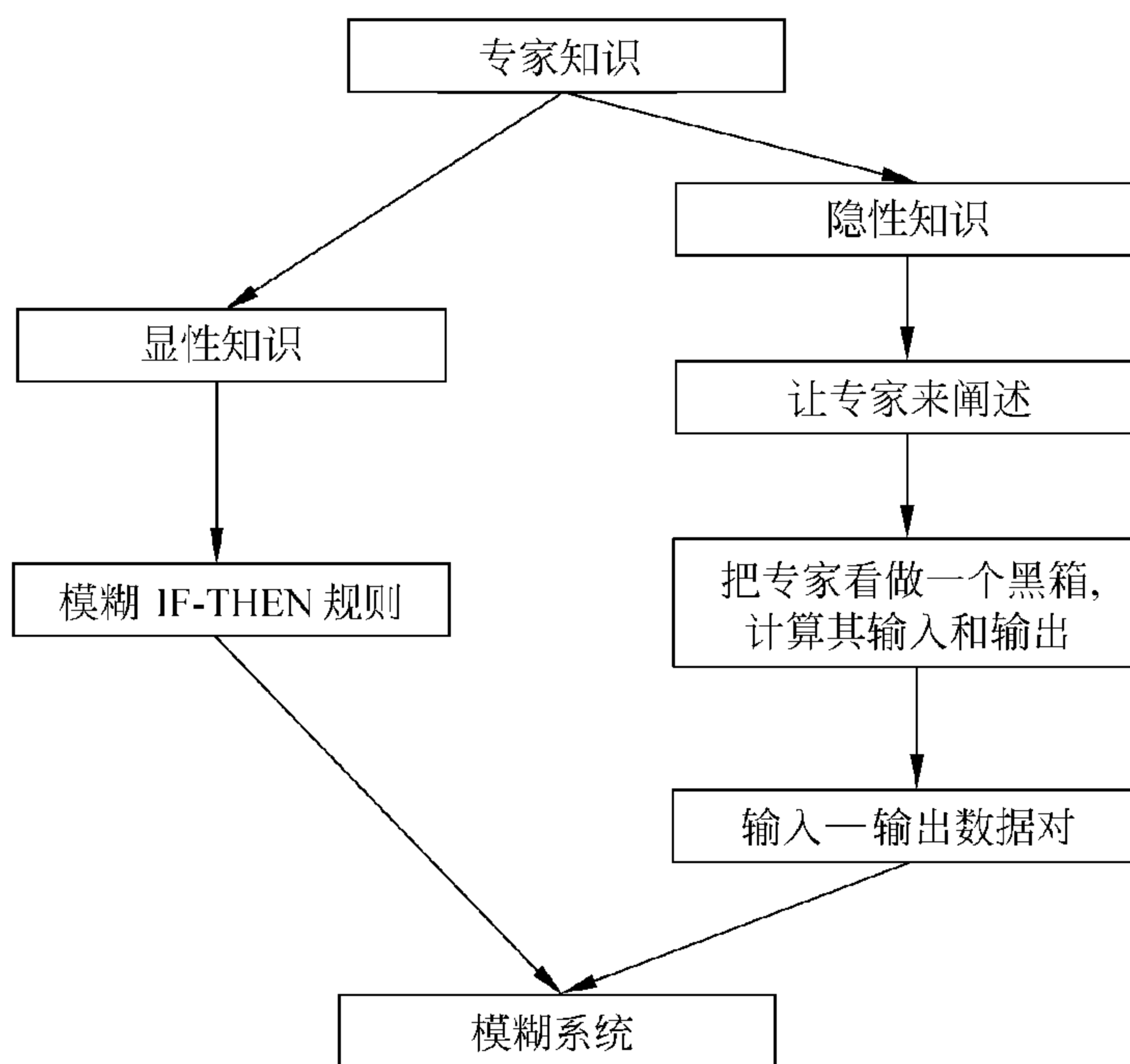


图 12.1 把专家知识转化成模糊系统

对设计模糊系统。由于在许多特定场合下,我们仅能给出有限的输入—输出数据对,而无法得到任意输入下的输出,所以,第 10 章和第 11 章中设计模糊系统的方法在这里是无法应用的(因为第 10 章和第 11 章中的方法都要求能对任意输入 $x \in U$ 确定输出 $g(x)$)。因此,这里考虑在第 10 章开始部分所讨论的第三种情况,即设计一个模糊系统,用其来描述由输入—输出数据对所表示的输入—输出行为。

在第 12 章中,将开发一种根据输入—输出数据对设计模糊系统的简单启发式方法,并将该方法应用于倒车控制和时间序列预测问题。在第 13 章中,首先将详细描述模糊系统的结构,然后用梯度下降法来调整结构参数,并用所设计的模糊系统辨识非线性动态系统。在第 14 章中,将采用递推最小二乘法设计模糊系统,并将所设计的模糊系统作为一种非线性通信信道的均衡器来使用。最后,第 15 章将给出怎样应用聚类的思想设计模糊系统。

12.1 设计模糊系统的查表法

假设给出如下输入—输出数据对：

$$(x_0^p; y_0^p), p = 1, 2, \dots, N \quad (12.1)$$

式中, $x_0^p \in U = [\alpha_1, \beta_1] \times \dots \times [\alpha_n, \beta_n] \subset R^n, y_0^p \in V = [\alpha_y, \beta_y] \subset R$ 。目的是根据以上 N 对输入—输出数据设计一个模糊系统 $f(x)$ 。下面给出利用查表法设计模糊系统的五个步骤：

步骤 1 把输入和输出空间划分为模糊空间。

具体做法为, 在每个区间 $[\alpha_i, \beta_i]$ ($i = 1, 2, \dots, n$) 上定义 N_i 个模糊集 A_i^j ($j = 1, 2, \dots, N_i$), 且 A_i^j 在 $[\alpha_i, \beta_i]$ 上是完备模糊集, 即对任意 $x_i \in [\alpha_i, \beta_i]$, 都存在 A_i^j 使得 $\mu_{A_i^j}(x_i) \neq 0$ 。比如, 可选择 $\mu_{A_i^j}(x_i)$ 为四边形隶属度函数：

$$\mu_{A_i^j}(x_i) = \mu_{A_i^j}(x_i; a_i^j, b_i^j, c_i^j, d_i^j)$$

其中 $a_i^1 = b_i^1 = \alpha_i, c_i^j = a_i^{j+1} < b_i^{j+1} = d_i^j$ ($j = 1, 2, \dots, N_i - 1$), $c_i^{N_i} = d_i^{N_i} = \beta_i$ 。

类似地, 定义 N_y 个模糊集 $B^j, j = 1, 2, \dots, N_y$, 它们在 $[\alpha_y, \beta_y]$ 上也是完备模糊集。也选择 $\mu_{B^j}(y)$ 为四边形隶属度函数

$$\mu_{B^j}(y) = \mu_{B^j}(y; a^j, b^j, c^j, d^j)$$

其中 $a^1 = b^1 = \alpha_y, c^j = a^{j+1} < b^{j+1} = d^j$ ($j = 1, 2, \dots, N_y - 1$), $c^{N_y} = d^{N_y} = \beta_y$ 。图 12.2 给出了 $n = 2$ 中的一个例子, 图中 $N_1 = 5, N_2 = 7, N_y = 5$, 所有的隶属度函数都是三角形隶属度函数。

步骤 2 由一个输入—输出数据对产生一条模糊规则。

首先, 根据每个输入—输出数据对 $(x_{01}^p, \dots, x_{0n}^p; y_0^p)$, 确定 x_{0i}^p ($i = 1, 2, \dots, n$) 隶属于模糊集 A_i^j ($j = 1, 2, \dots, N_i$) 的隶属度值和 y_0^p 隶属于模糊集 B^l 的隶属度值 ($l = 1, 2, \dots, N_y$), 即计算 $\mu_{A_i^j}(x_{0i}^p)$ ($j = 1, 2, \dots, N_i, i = 1, 2, \dots, n$) 和 $\mu_{B^l}(y_0^p)$ ($l = 1, 2, \dots, N_y$)。对于图 12.2 的例子, 可以近似地认为: x_{01}^1 隶属于 $B1$ 的隶属度值为 0.8, 隶属于 $B2$ 的隶属度值为 0.2, 隶属于其他模糊集的隶属度值为 0; x_{02}^1 隶属于 $S1$ 的隶属度值为 0.6, 隶属于 $S2$ 的隶属度值为 0.4, 隶属于其他模糊集的隶属度值为 0; y_0^1 隶属于 CE 的隶属度值为 0.8, 隶属于 $B1$ 的隶属度值为 0.2, 隶属于其他模糊集的隶属度值为 0。

然后, 对每个输入变量 x_i ($i = 1, 2, \dots, n$), 确定使 x_{0i}^p 有最大隶属度值的模糊集, 即, 确定 $A_i^{j^*}$

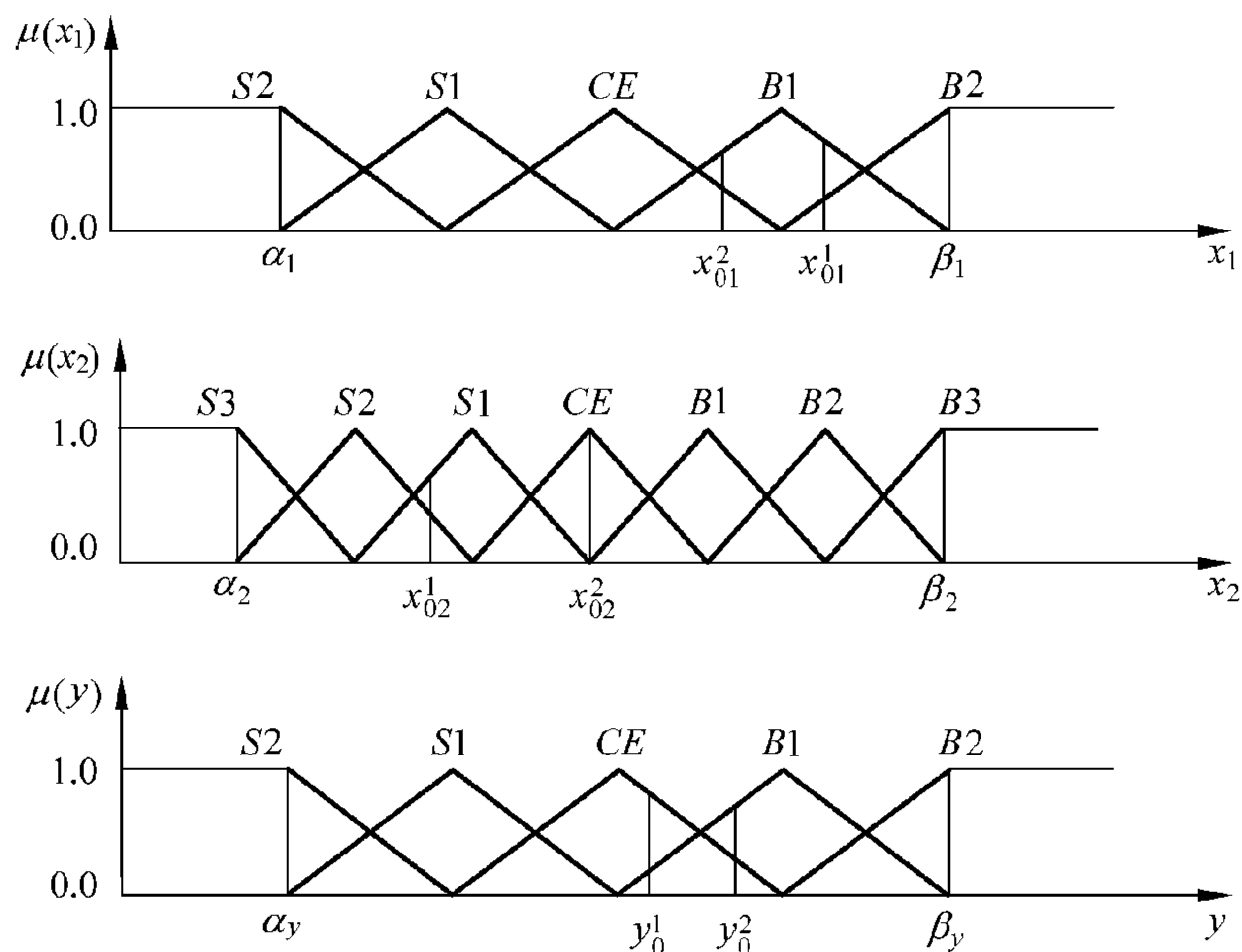


图 12.2 隶属度函数和二维输入情况下输入—输出数据对的函数的一个

使得 $\mu_{A_i^*}(x_{0i}^p) \geq \mu_{A_i}(x_{0i}^p)$ ($j=1, 2, \dots, N_i$)。类似地, 确定 B^{l*} 使得 $\mu_{B^{l*}}(y_0^p) \geq \mu_{B^l}(y_0^p)$ ($l=1, 2, \dots, N_y$)。对于图 12.2 中的例子来说, 由输入—输出数据对 $(x_{01}^1, x_{02}^1; y_0^1)$ 可得, $A_1^{j*} = B1, A_2^{j*} = S1, B^{l*} = CE$, 由输入—输出数据对 $(x_{01}^2, x_{02}^2; y_0^2)$ 可得, $A_1^{j*} = B1, A_2^{j*} = CE, B^{l*} = B1$ 。

最后, 可以得到下面的模糊 IF-THEN 规则

$$\text{如果 } x_1 \text{ 为 } A_1^{j*} \text{ 且 } \dots \text{ 且 } x_n \text{ 为 } A_n^{j*}, \text{ 则 } y \text{ 为 } B^{l*} \quad (12.2)$$

对于图 12.2 中的例子, 由数据对 $(x_{01}^1, x_{02}^1; y_0^1)$ 可得出规则: 如果 x_1 为 $B1$ 且 x_2 为 $S1$, 则 y 为 CE ; 由数据对 $(x_{01}^2, x_{02}^2; y_0^2)$ 可得出规则: 如果 x_1 为 $B1$ 且 x_2 为 CE , 则 y 为 $B1$ 。

步骤 3 对步骤 2 中的每条规则赋予一个强度。

由于输入—输出数据对的数量通常都较大, 且每对数据都会产生一条规则, 所以很可能会有冲突的规则。即, 规则的 IF 部分相同, 而 THEN 部分不同。为了解决这一冲突, 可赋予步骤 2 中的每条规则一个强度, 从而使得一个冲突群中仅有一条规则具有最大强度。这样, 不仅冲突问题解决了, 而且规则数量也大大减少了。

规则的强度定义如下, 假定规则(12.2)是由输入—输出数据对 $(x_0^p; y_0^p)$ 产生的, 则其强度可定义为

$$D(\text{规则}) = \prod_{i=1}^n \mu_{A_i^*}(x_{0i}^p) \mu_{B^{l*}}(y_0^p) \quad (12.3)$$

对于图 12.2 中的例子, 由数据对 $(x_{01}^1, x_{02}^1; y_0^1)$ 产生的规则的强度为

$$D(\text{规则 1}) = \mu_{B1}(x_{01}^1) \mu_{S1}(x_{02}^1) \mu_{CE}(y_0^1)$$

$$= 0.8 \times 0.6 \times 0.8 = 0.384 \quad (12.4)$$

由数据对 $(x_{01}^2, x_{02}^2; y_0^2)$ 产生的规则强度为

$$\begin{aligned} D(\text{规则 } 2) &= \mu_{B_1}(x_{01}^2) \mu_{CE}(x_{02}^2) \mu_{B_1}(y_0^2) \\ &= 0.6 \times 1 \times 0.7 = 0.42 \end{aligned} \quad (12.5)$$

如果输入—输出数据对具有不同的可靠性且能用一个数来评价它的话,则可以把这一信息也合并到规则强度中。具体来说,假定输入—输出数据对 $(x_0^p; y_0^p)$ 的可靠程度为 $\mu^p (\in [0, 1])$, 则由 $(x_0^p; y_0^p)$ 产生的规则强度可定义为

$$D(\text{规则}) = \prod_{i=1}^n \mu_{A_i^*}(x_{0i}^p) \mu_{B_i^*}(y_0^p) \mu^p \quad (12.6)$$

在实践中,如果输入—输出数据对数目小的话,可以让专家去检测数据,并评价其可靠程度 μ^p 。或者,如果知道在数据对中噪声的特性,也可以选择 μ^p 来反映噪声的强度。如果不能知道在输入—输出数据对间的差异,则可简单地令所有 $\mu^p = 1$, 这样式(12.6)就变了式(12.3)。

步骤 4 创建模糊规则库。

模糊规则库由以下三个规则集合组成:

- 步骤 2 中产生的与其他规则不发生冲突的规则;
- 一个冲突规则群体中具有最大强度的规则,其中冲突规则群体是指那些具有相同的 IF 部分的规则;
- 来自于专家的语言规则(主要指专家的显性知识)。

由于前两个规则集合是由隐性知识得到的,所以最终的规则库是由显性知识和隐性知识组成的。

直观上,可以把一个模糊规则库描述成一个两维输入情况下的可查询的表格。例如,图 12.3 就给出了与图 12.2 中的模糊集相对应的模糊规则库的表格示意图。每个格子代表 $[\alpha_1, \beta_1]$ 中的模糊集和 $[\alpha_2, \beta_2]$ 中的模糊集的一个组合,由此能得到一条可能的规则。一个冲突规则群体是由同一个格子中的规则组成的。该方法也可以看做是用恰当的规则来填充这个表格,这就是称其为查表法的原因。

步骤 5 基于模糊规则库构造模糊系统。

根据步骤 4 中产生的模糊规则库,可用第 9 章中的任一方案来构造模糊系统。例如,可以选择带有乘积推理机、单值模糊器、中心平均解模糊器(引理 9.1)的模糊系统。

现对由输入—输出数据对设计模糊系统的五个步骤做几点评论:

- 第 10 章与第 11 章中的方法和查表法之间的根本差异在于,第 10 章和第 11 章

S3					
S2					
S1					
x_2 CE					
B1					
B2					
B3					
	S2	S1	CE	B1	B2
	x_1				

图 12.3 由查表法得到的模糊规则库示意图

中的方法要求能对任意的输入 $x \in U$ 确定正确的输出 $g(x)$, 而查表法却能在给定的输入—输出数据对中自由选择输入点。而且, 第 10 章和第 11 章中的方法为了以所需精度设计模糊系统, 需要知道所逼近函数的一阶或二阶导数的边界, 查表法则不需要这一信息。

- 如果输入—输出数据对恰为式(10.9)中的 $(e_1^{i_1}, e_2^{i_2}; \bar{y}^{i_1 i_2})$, 则很容易证明, 通过这五步所设计的模糊系统与 10.2 节中设计的模糊系统是一样的。所以, 可以把查表法看做是 10.2 节中设计方法的一种推广, 而 10.2 节中的输入—输出数据对是不能任意选择的。
- 最终的模糊规则库中规则的数量受两个数约束: N (输入—输出数据对的数量) 和 $\prod_{i=1}^n N_i$ (由输入变量定义的模糊集的所有可能组合的数量)。如果输入空间 n 的维数很大, 则 $\prod_{i=1}^n N_i$ 将是一个非常大的数, 可能大于 N 。此外, 某些输入—输出数据对可能对应于图 12.3 中的同一个格子, 但仅有一条规则对模糊规则库是有用的。所以, 模糊规则库中的规则的数量都可能小于 $\prod_{i=1}^n N_i$ 和 N 。这样, 由查表法得到的模糊规则库可能是不完备的。为了使模糊规则库完备, 可能需要通过插补现存的规则填充模糊规则库中的空格, 细节留给读者自己去思考。

下面, 将查表法应用于一个控制问题和一个时间序列预测问题。

12.2 查表法在倒车控制中的应用

将卡车后倒至一装卸车位是一个非线性控制问题。采用传统的控制方法, 首先会建立一个该系统的数学模型, 然后根据非线性控制理论设计一个控制器 (Walsh, Tilbury, Sastry, Murray 和 Laumond [1994])。还有一种方法就是, 设计一个控制器来模仿人类司机。本文采用后一种方法。假设有一个有经验的司机, 且已知卡车的状态以及司机将卡车后倒至车位的控制行为, 即, 能够获得输入—输出 (状态—控制) 数据对集合。那么就可以采用上一节中的查表法, 根据这些输入—输出数据对设计一个模糊系统, 并用所设计的模糊系统来

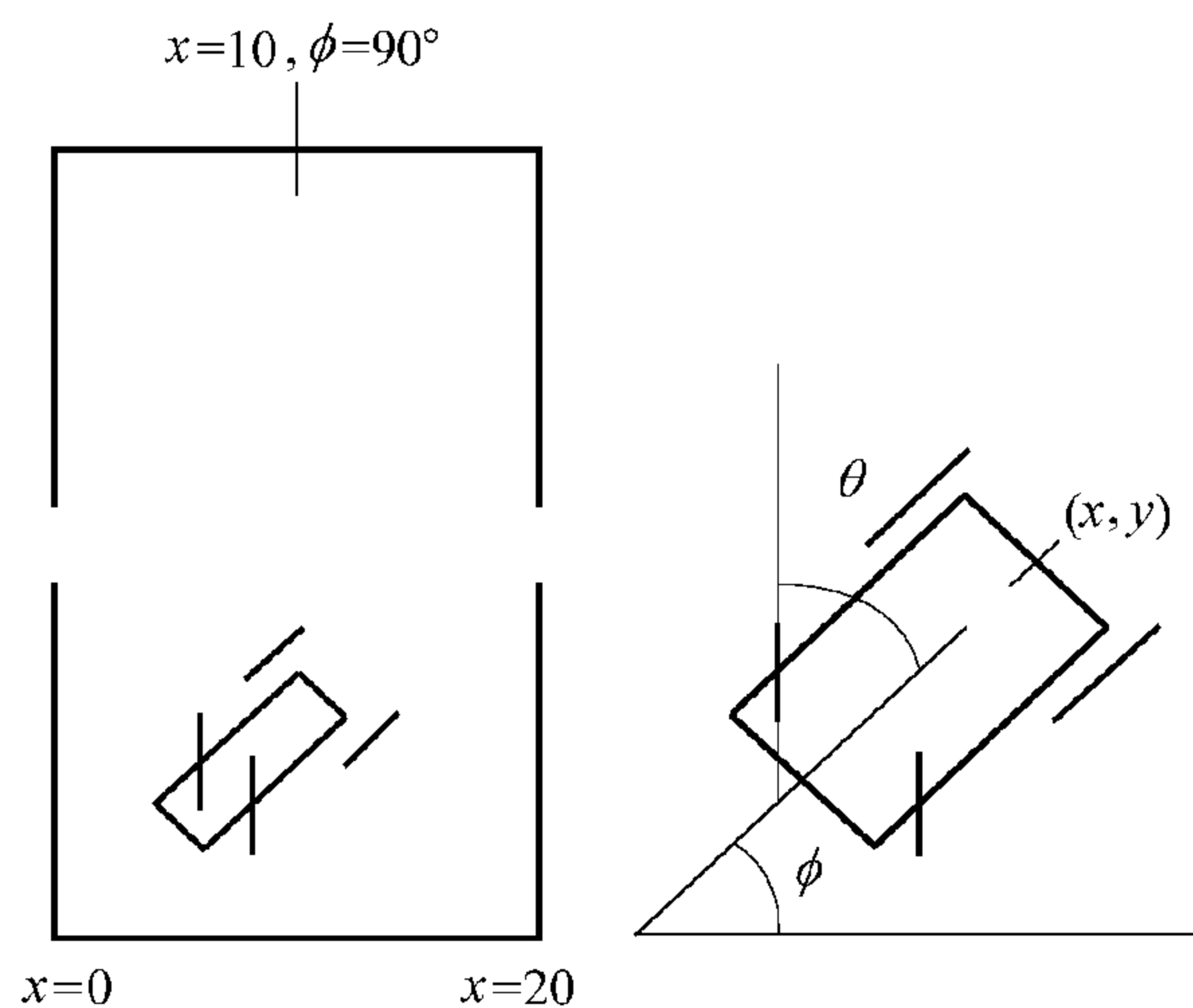


图 12.4 仿真的卡车和装卸区

取代人类司机。

仿真的卡车和装卸区位如图 12.4 所示。卡车的位置是由三个状态变量 ϕ, x 和 y 确定的, 其中 ϕ 是卡车与水平线的夹角, θ 为卡车的控制量即方向盘转动角(见图 12.4)。

卡车只允许后倒, 卡车每隔一段时间都会向后移动, 移动的距离是固定的。为简化该问题, 假设在卡车和装卸车位之间有足够间隙, 这样 y 就是一个状态变量了。设计一个输入为 (x, ϕ) , 输出为 θ 的控制器, 要求卡车的最终状态为 $(x_f, \phi_f) = (10, 90^\circ)$ 。假设

$$x \in [0, 20], \phi \in [-90^\circ, 270^\circ], \theta \in [-40^\circ, 40^\circ]$$

即

$$U = [0, 20] \times [-90^\circ, 270^\circ], V = [-40^\circ, 40^\circ]$$

首先, 产生一组输入—输出数据对 $(x^p, \phi^p; \theta^p)$ 。利用试错法产生输入—输出数据对: 在每一时刻(此时 x 和 ϕ 是给定的)当卡车从一个初始状态开始后倒时, 根据常识确定控制量 θ (实际上这种常识就是在某种状态下如何控制方向盘角度的经验)。经过几次实验后, 就可以找到与最平滑最成功的轨迹相对应的输入—输出数据对。以下 14 个初始状态将分别用来产生所需的初始输入—输出数据对:

$$\begin{aligned} (x_0, \phi_0^\circ) = & (1, 0), (1, 90), (1, 270); (7, 0), (7, 90), (7, 180), (7, 270); \\ & (13, 0), (13, 90), (13, 180), (13, 270); (19, 90), (19, 180), (19, 270) \end{aligned}$$

表 12.1 给出了初始状态为 $(x_0, \phi_0) = (1, 0^\circ)$ 时的输入—输出数据对。其他 13 种初始状态下的输入—输出数据对也可用类似的方法得到。总的说来, 可以得到 250 对输入—输出数据对。现在, 根据这些输入—输出数据对, 利用上一节中的查表法来设计一个模糊系统。

表 12.1 理想的轨迹 (x_t, ϕ_t°) 和从 $(x_0, \phi_0) = (1, 0^\circ)$ 出发的对应控制量 θ_t°

t	x_t	ϕ_t°	θ_t°	t	x_t	ϕ_t°	θ_t°
-----	-------	----------------	------------------	-----	-------	----------------	------------------

0	1.00	0.00	-19.00	9	8.72	65.99	-9.55
1	1.95	9.37	-17.95	10	9.01	70.75	-8.50
2	2.88	18.23	-16.90	11	9.28	74.98	-7.45
3	3.79	26.57	-15.85	12	9.46	78.70	-6.40
4	4.65	34.44	-14.80	13	9.59	81.90	-5.34
5	5.45	41.78	-13.75	14	9.72	84.57	-4.30
6	6.18	48.60	-12.70	15	9.81	86.72	-3.25
7	7.48	54.91	-11.65	16	9.88	88.34	-2.20
8	7.99	60.71	-10.60	17	9.91	89.44	0.00

步骤 1 在 $[-90^\circ, 270^\circ]$ 上定义 7 个模糊集, 在 $[0, 20]$ 上定义 5 个模糊集, 在 $[-40^\circ, 40^\circ]$ 上定义 7 个模糊集, 它们的隶属度函数的图形见图 12.5。

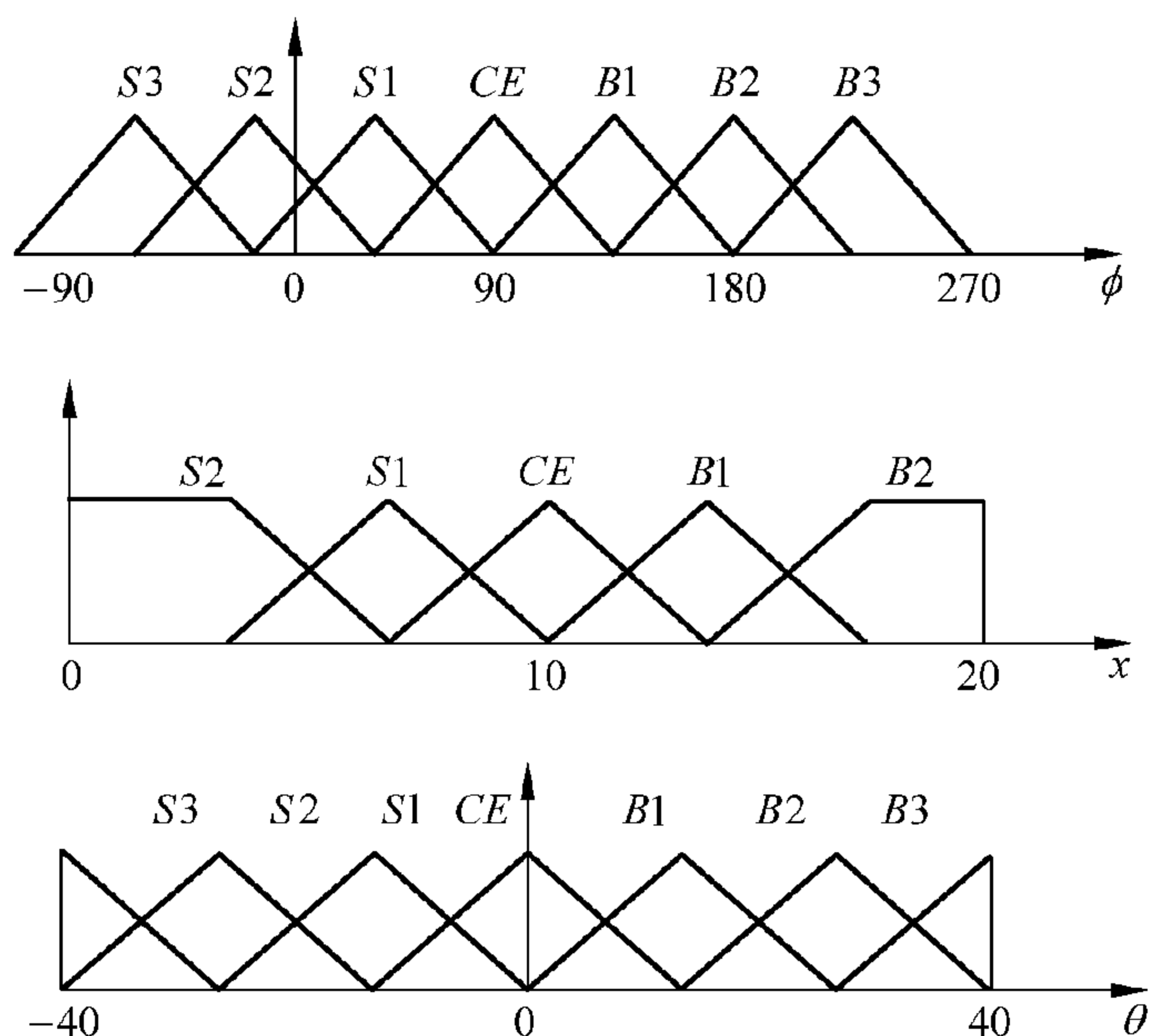


图 12.5 倒车控制问题的隶属度函数

步骤 2 和步骤 3 由一个输入—输出数据对产生一条规则并计算该规则的强度。表 12.2 说明了, 规则及其强度都是由表 12.1 中对应的输入—输出数据对产生的。

表 12.2 根据表 12.1 中的输入—输出数据对产生的模糊 IF—THEN 规则及规则强度

x 为	ϕ 为	θ 为	强度
-------	----------	------------	----

S2	S2	S2	1.00
S2	S2	S2	0.92
S2	S2	S2	0.35
S2	S2	S2	0.12
S2	S2	S2	0.07
S1	S2	S1	0.08
S1	S1	S1	0.18
S1	S1	S1	0.53
S1	S1	S1	0.56
S1	S1	S1	0.60
CE	S1	S1	0.35

续表

x 为	ϕ 为	θ 为	强度
<i>CE</i>	<i>S1</i>	<i>S1</i>	0.21
<i>CE</i>	<i>S1</i>	<i>CE</i>	0.16
<i>CE</i>	<i>CE</i>	<i>CE</i>	0.32
<i>CE</i>	<i>CE</i>	<i>CE</i>	0.45
<i>CE</i>	<i>CE</i>	<i>CE</i>	0.54
<i>CE</i>	<i>CE</i>	<i>CE</i>	0.88
<i>CE</i>	<i>CE</i>	<i>CE</i>	0.92

步骤4 产生的最终模糊规则库(如图 12.6 所示,可以看出,某些格子是空的,所以,输入—输出数据对并没有覆盖所有的状态空间;不过,仍可看出,图 12.6 中的规则对于控制倒车,使其从一个初始的位置开始后倒至指定的位置,已经是足够的了)。

ϕ <i>S3</i>	<i>S2</i>	<i>S3</i>			
<i>S2</i>	<i>S2</i>	<i>S3</i>	<i>S3</i>	<i>S3</i>	
<i>S1</i>	<i>B1</i>	<i>S1</i>	<i>S2</i>	<i>S3</i>	<i>S2</i>
<i>CE</i>	<i>B2</i>	<i>B2</i>	<i>CE</i>	<i>S2</i>	<i>S2</i>
<i>B1</i>	<i>B2</i>	<i>B3</i>	<i>B2</i>	<i>B1</i>	<i>S1</i>
<i>B2</i>		<i>B3</i>	<i>B3</i>	<i>B3</i>	<i>B2</i>
<i>B3</i>				<i>B3</i>	<i>B2</i>
	<i>S2</i>	<i>S1</i>	<i>CE</i>	<i>B1</i>	<i>B2</i>
	x				

图 12.6 倒车控制问题的最终模糊规则库

步骤5 采用带有乘积推理机、单值模糊器、中心平均解模糊器的模糊系统。即,设计形如式(9.1)的模糊系统,其规则如图 12.6 所示。

现在,将上面设计的模糊系统作为一个倒车控制器应用在卡车上。为了仿真该控制系统,需要一个控制倒车的数学模型。选用以下逼近模型(Wang 和 Mendel[1992b]):

$$x(t+1) = x(t) + \cos[\phi(t) + \theta(t)] + \sin[\theta(t)]\sin[\phi(t)] \quad (12.7)$$

$$y(t+1) = y(t) + \sin[\phi(t) + \theta(t)] - \sin[\theta(t)]\cos[\phi(t)] \quad (12.8)$$

$$\phi(t+1) = \phi(t) - \sin^{-1}\left[\frac{2\sin(\theta(t))}{b}\right] \quad (12.9)$$

式中, b 为卡车的长度并在仿真中假定 $b=4$ 。图 12.7 给出了采用所设计的模糊系统作为控

制器时,从两个初始条件 $(x_0, \phi_0) = (3, -30^\circ)$ 和 $(x_0, \phi_0) = (13, 30^\circ)$ 出发的倒车轨迹。可以看出,模糊控制器可以成功地用于控制倒车至指定的位置上。

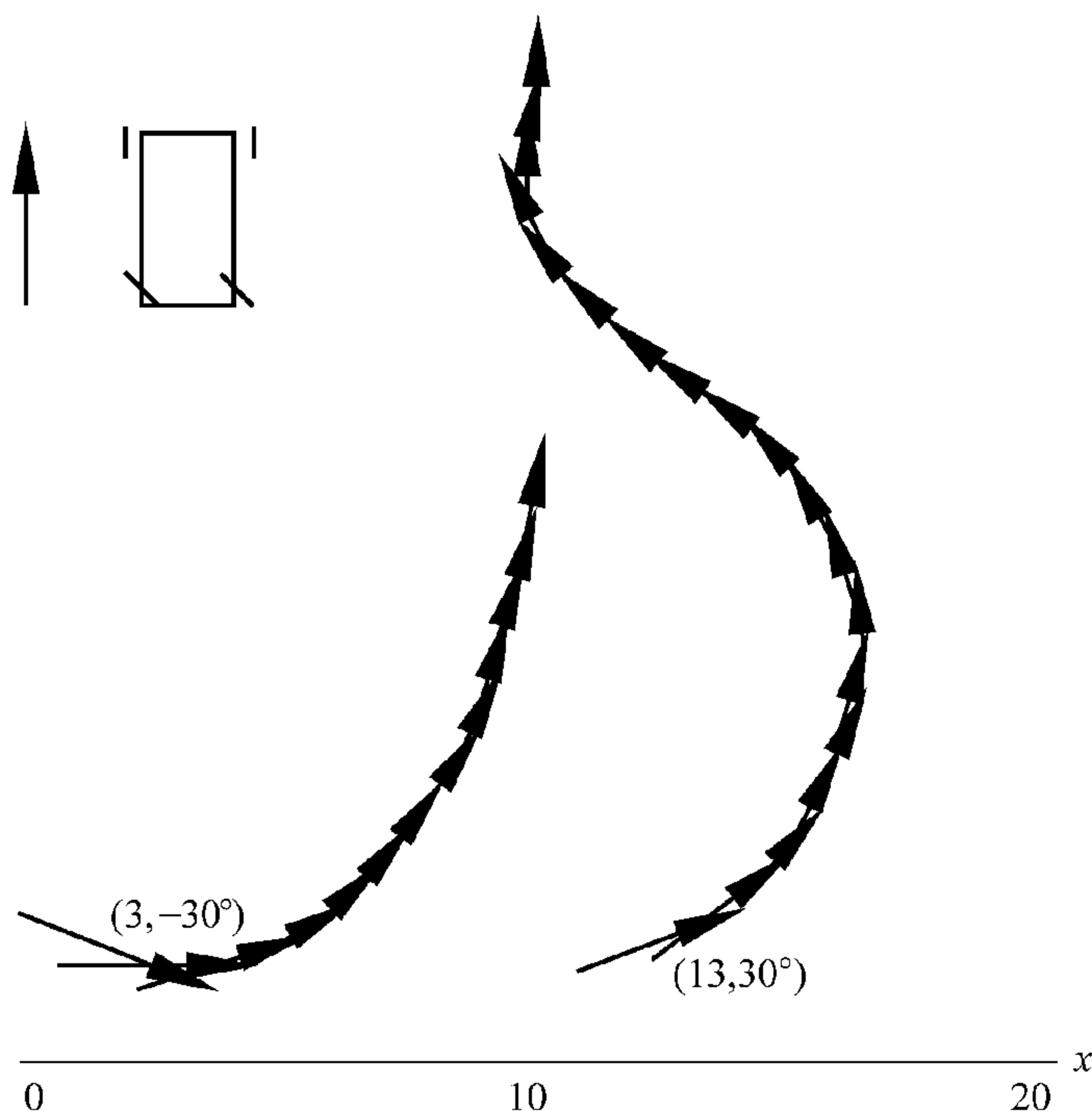


图 12.7 采用模糊控制器的倒车轨迹

12.3 查表法在时间序列预测中的应用

时间序列预测是一个重要的实际问题。时间序列预测可应用于经济学、企业计划、库存和生产控制、天气预报、信号处理、控制及许多其他领域。本节采用查表法所设计的模糊系统来预测 Mackey-Glass 混沌时间序列,它由下面的含有时滞的微分方程产生:

$$\frac{dx(t)}{dt} = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t) \quad (12.10)$$

当 $\tau > 17$ 时,式(12.10)显示了混沌行为。这里选定 $\tau = 30$ 。

令 $x(k)$ ($k = 1, 2, 3, \dots$) 为由式(12.10)所产生的时间序列数据[在式(12.10)每隔 1 秒所产生的连续曲线 $x(t)$ 中抽样]。图 12.8 给出了这段混沌时间序列 $x(k)$ 的 600 个点。时间序列预测问题可做如下描述:给定 $x(k-n+1), x(k-n+2), \dots, x(k)$, 预测 $x(k+1)$, 其中 n 是正整数。即,任务是确定一个从 $[x(k-n+1), x(k-n+2), \dots, x(k)] \in R^n$ 到 $[x(k+1)] \in R$ 的映射。本例中这一映射就是以输入—输出数据对为基础设计的模糊系统。设 $x(1), x(2), \dots, x(k)$ ($k > n$), 可以产生如下 $k-n$ 对输入—输出数据:

$$\begin{aligned}
 & [x(k-n), \dots, x(k-1); x(k)] \\
 & [x(k-n-1), \dots, x(k-2); x(k-1)] \\
 & \dots \\
 & [x(1), \dots, x(n); x(n+1)]
 \end{aligned} \tag{12.11}$$

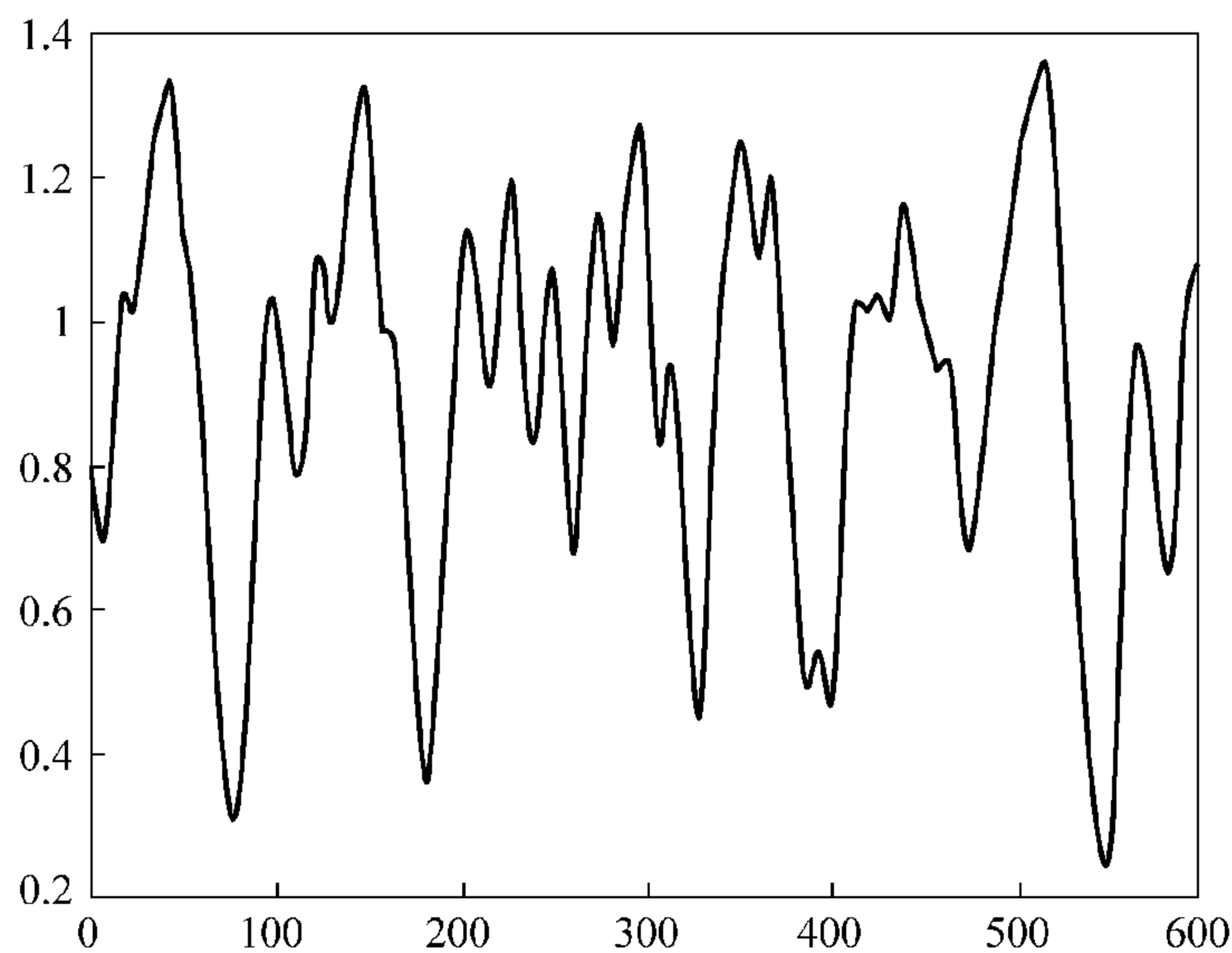


图 12.8 Mackey-Glass 混沌时间序列的一部分

根据 12.2 节中的查表法, 利用这些输入—输出数据对设计模糊系统 $f(x)$, 然后用 $f(x)$ 来预测 $x(k+l)$ ($l=1, 2, \dots$), 当预测 $x(k+l)$ 时, $f(x)$ 的输入是 $[x(k-n+l), \dots, x(k-1+l)]$ 。

现在用图 12.8 中的前 300 个点来组建输入—输出数据对, 然后用所设计的模糊系统来预测余下的 300 个点。考虑两种情况: (i) $n=4$, 对每个输入变量根据图 12.9 定义 7 个模糊集; (ii) $n=4$, 对每个输入变量根据图 12.10 定义 15 个模糊集。这两种情况的预测结果分别见图 12.11 和图 12.12。比较图 12.11 和图 12.12 可以看出, 对每个输入变量定义较多的模糊集可以提高预测精度。

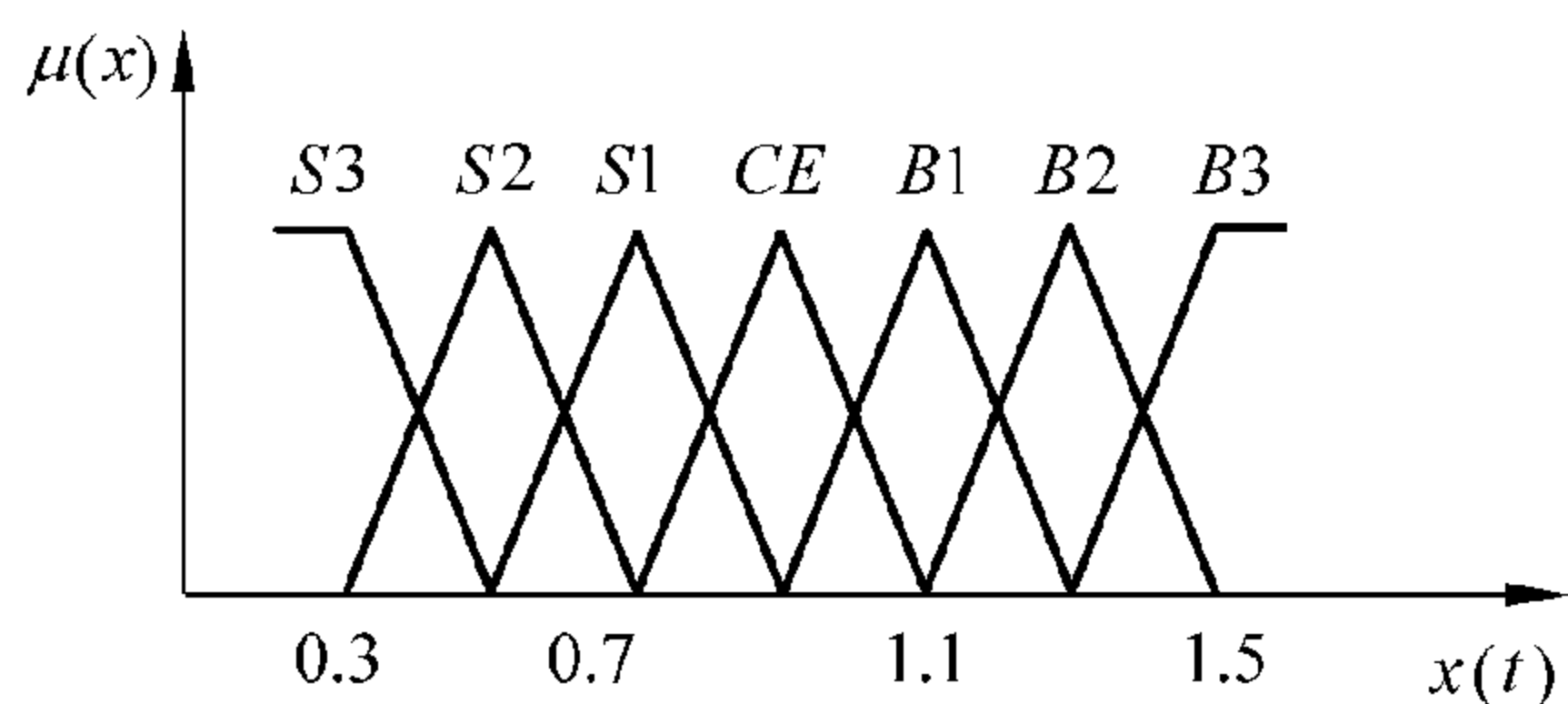


图 12.9 每个输入变量的隶属度函数的第一个选择

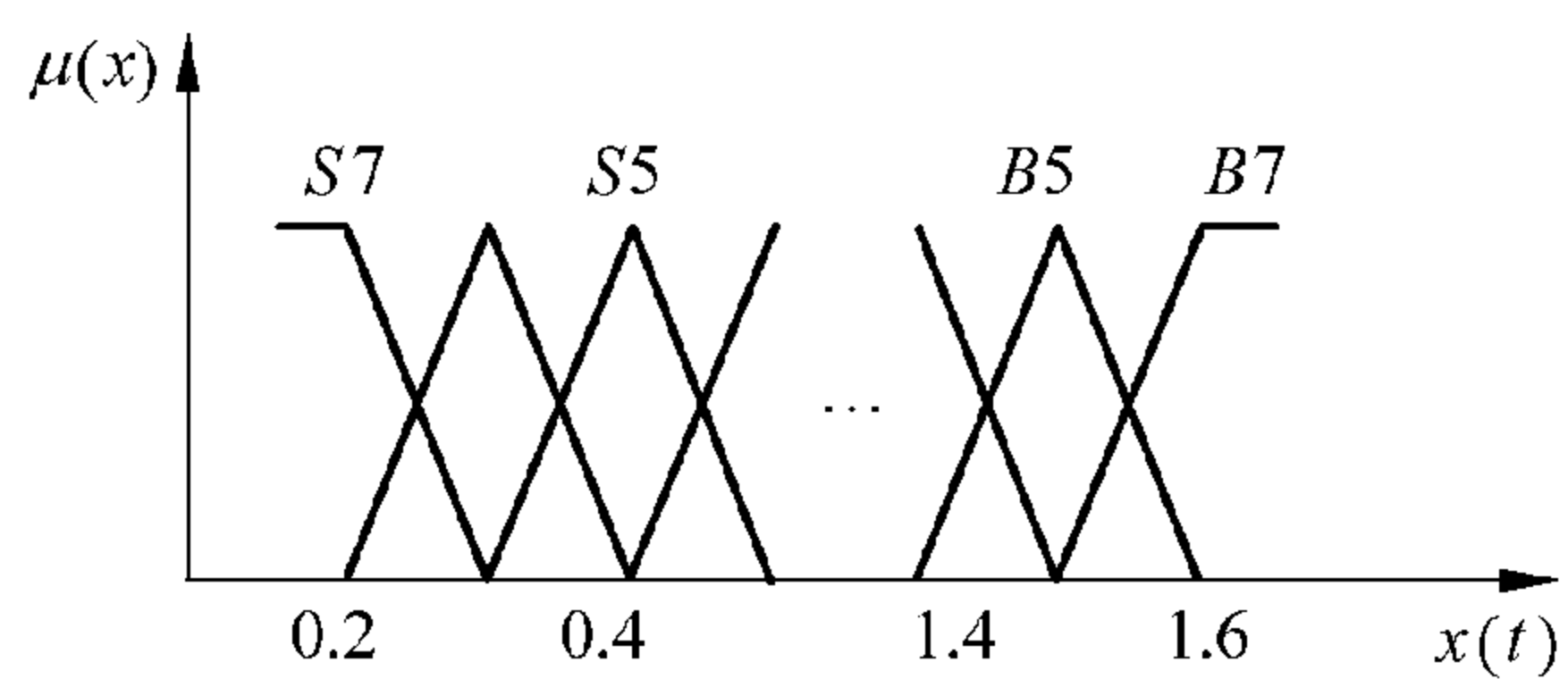


图 12.10 每个输入变量的隶属度函数的第二个选择

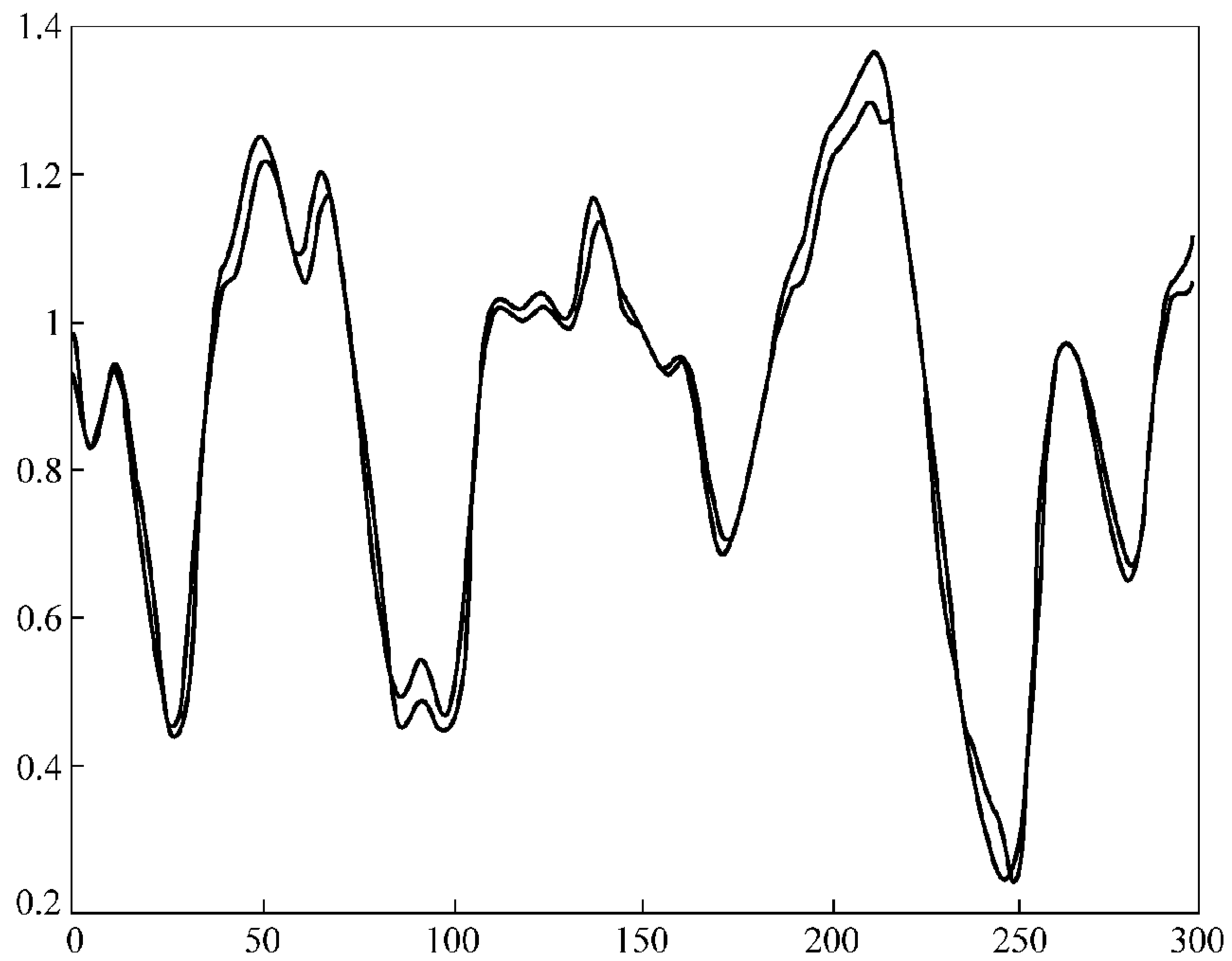


图 12.11 采用图 12.9 中所定义的隶属度函数所得到的时间序列的预测值和实值

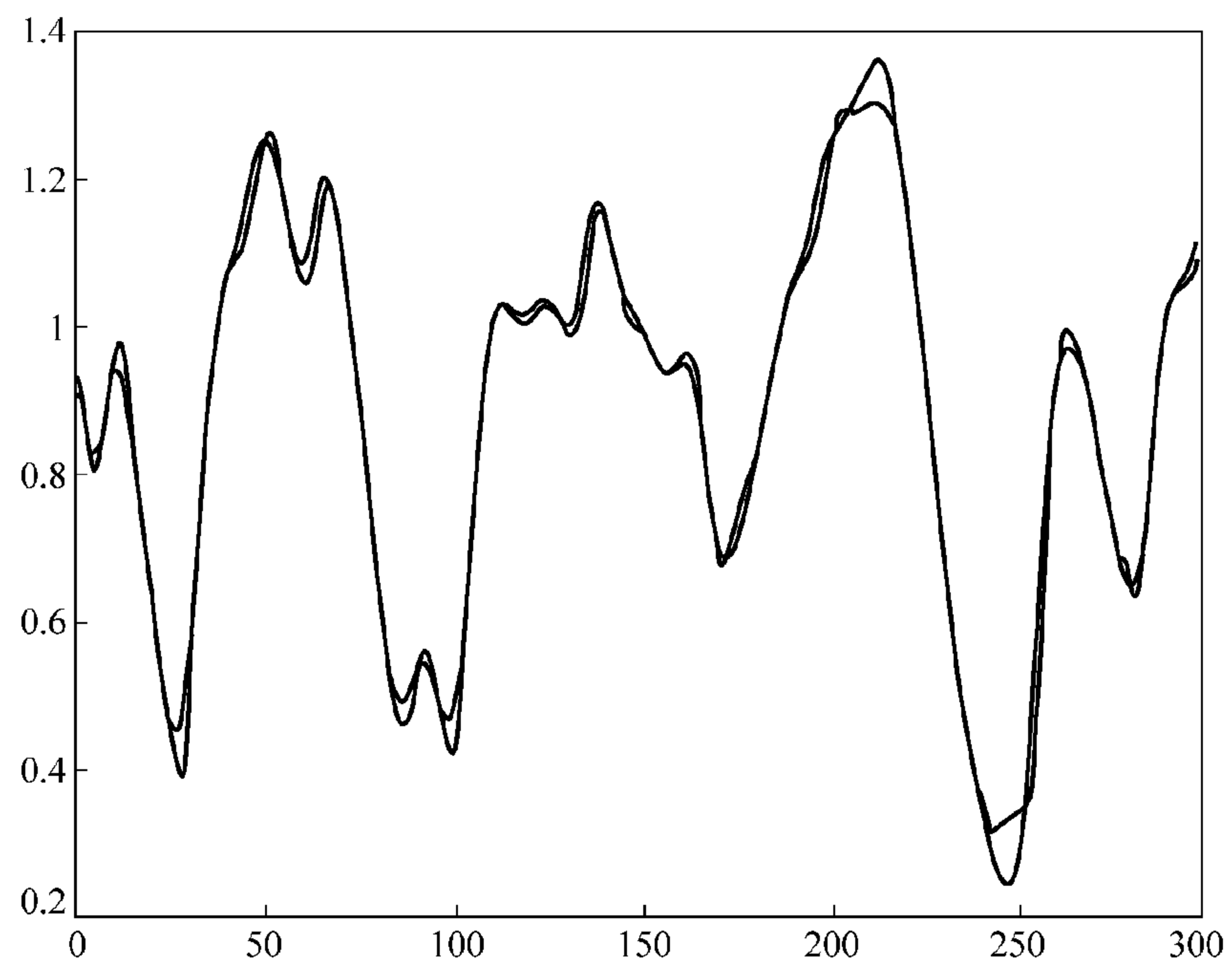


图 12.12 采用图 12.10 中所定义的隶属度函数所得到的时间序列的预测值和实值

12.4 本章小结及导读

本章主要阐述了以下内容：

- 根据输入—输出数据对,用查表法设计模糊系统的详细步骤;
- 如何用查表法控制倒车过程和时间序列预测问题;
- 怎样应用查表法把显性知识和隐性知识组合到模糊系统中去。

查表法摘自于 Wang 和 Mendel[1992b], Wang[1994], 文中还给出了关于这一方法的更多详细介绍并给出了更多的例子。查表法在金融数据预测方面的应用见 Cox[1994]。

12.5 习题与设计

习题 12.1 用查表法设计一个二维输入—一维输出的模糊系统。

假设步骤 1 中定义的模糊集如图 12.2 所示, 图中 $\alpha_1 = \alpha_2 = \alpha_y = 0, \beta_1 = \beta_2 = \beta_y = 1$ 。模糊集的隶属度函数都为三角形隶属度函数且空间相等。

(a) 使图 12.2 中的每个模糊集在所产生的规则中至少出现一次的输入—输出数据对的最小数目是多少? 给出一个这一输入—输出数据对最小集合的例子。

(b) 使得所产生的模糊规则库是完备的规则库的输入—输出数据对的最小数目是多少? 给出一个这一输入—输出数据对最小集合的例子。

习题 12.2 考虑 12.3 节中的倒车控制问题。

(a) 用常识产生一个从初始状态 $(x_0, \phi_0) = (1, 90^\circ)$ 至最终状态 $(x_f, \phi_f) = (10, 90^\circ)$ 的驾驶卡车的输入—输出数据对集合。

(b) 根据(a)中所产生的输入—输出数据对,用查表法建立一个模糊规则库,步骤 1 中的隶属度函数由图 12.5 给定。

(c) 构造一个基于(b)中的模糊规则库的模糊系统,用该模糊系统来控制初始状态分别为 $(x_0, \phi_0) = (0, 90^\circ)$ 和 $(x_0, \phi_0) = (-3, 90^\circ)$ 的卡车,并评价仿真结果。

习题 12.3 令 $f(x)$ 为用查表法所设计的模糊系统。你能确定 $|f(x_0^p) - y_0^p|$ 的误差边界吗? 请解释你的答案。

习题 12.4 提出一种填充由查表法所产生的模糊规则库中的空格的方法。证明你的方法并用例子来验证它。

习题 12.5 给定一个时间序列的 10 个点 $x(1), x(2), \dots, x(10)$, 预测 $x(12)$ 。

(a) 如果用模糊系统 $f[x(10), x(8)]$ 来预测 $x(12)$, 那么请列出所有构造这个模糊系

统的输入—输出数据对。

(b) 如果用模糊系统 $f[x(10), x(9), x(8)]$ 来预测 $x(12)$, 则列出所有组建这个模糊系统的输入—输出数据对。

12.6(设计) 编写一个计算机程序来实现查表法, 并将你的程序应用于 12.4 节中的时间序列预测问题。为使编码可行, 程序中可能还需要有一种填充表格中空格的方法。

梯度下降法设计模糊系统

13.1 模糊系统结构的选择

在第 12 章的查表法中, 第一步中的隶属度函数是固定不变的且不依赖于输入—输出数据对, 即, 隶属度函数并不根据输入—输出数据对进行优化。本章提出了设计模糊系统的另一种方法, 该方法中的隶属度函数是在进行了某种程度的优化后选择出来的。

从概念的角度来看, 根据输入—输出数据对设计模糊系统有两种方法。第一种方法: 首先由输入—输出数据对产生模糊 IF-THEN 规则, 然后根据这些规则和选定的模糊推理机、模糊器、解模糊器来构造模糊系统。第 12 章中的查表法就属于这种方法。第二种方法: 首先描述模糊系统的结构, 然后允许模糊系统结构中的一些参数自由变化, 最后根据输入—输出数据对确定这些自由参数。本章采用第二种方法。

首先, 设计模糊系统的结构。这里把模糊系统选定为式(9.6)给定的带有乘积推理机、单值模糊器、中心平均解模糊器和高斯隶属度函数的模糊系统。即, 假定将要设计的模糊系统形式为

$$f(x) = \frac{\sum_{l=1}^M \bar{y}^l \left[\prod_{i=1}^n \exp\left(-\left(\frac{x_i - \bar{x}_i^l}{\sigma_i^l}\right)^2\right)\right]}{\sum_{l=1}^M \left[\prod_{i=1}^n \exp\left(-\left(\frac{x_i - \bar{x}_i^l}{\sigma_i^l}\right)^2\right)\right]} \quad (13.1)$$

式中, M 是固定不变的, \bar{y}^l, \bar{x}_i^l 和 σ_i^l 是自由变化的参数(令 $a_i^l = 1$)。尽管模糊系统的结构已经选定为式(13.1), 但由于参数 \bar{y}^l, \bar{x}_i^l 和 σ_i^l 还未确定, 所以模糊系统并未设计好。一旦参数 $\bar{y}^l, \bar{x}_i^l, \sigma_i^l$ 确定了, 模糊系统也就设计好了。即, 设计模糊系统和确定参数 $\bar{y}^l, \bar{x}_i^l, \sigma_i^l$ 是等价的。

把式(13.1)中的模糊系统表述为一个前馈网络有助于以某种最优的方式确定这些参数。具体来讲, 从输入 $x \in U \subset R^n$ 到输出 $f(x) \in V \subset R$ 的映射可以根据下面的运算得到: 首先, 输入 x 通过一个乘积高斯算子运算而变成了 $z^l = \prod_{i=1}^n \exp\left(-\left(\frac{x_i - \bar{x}_i^l}{\sigma_i^l}\right)^2\right)$; 然后, z^l 再通

过一个求和运算和一个加权求和运算得到 $b = \sum_{l=1}^M z^l$ 和 $a = \sum_{l=1}^M \bar{y}^l z^l$; 最后, 计算模糊系统的输出 $f(x) = a/b$ 。这三个阶段的运算见图 13.1 中的三层前馈网络。

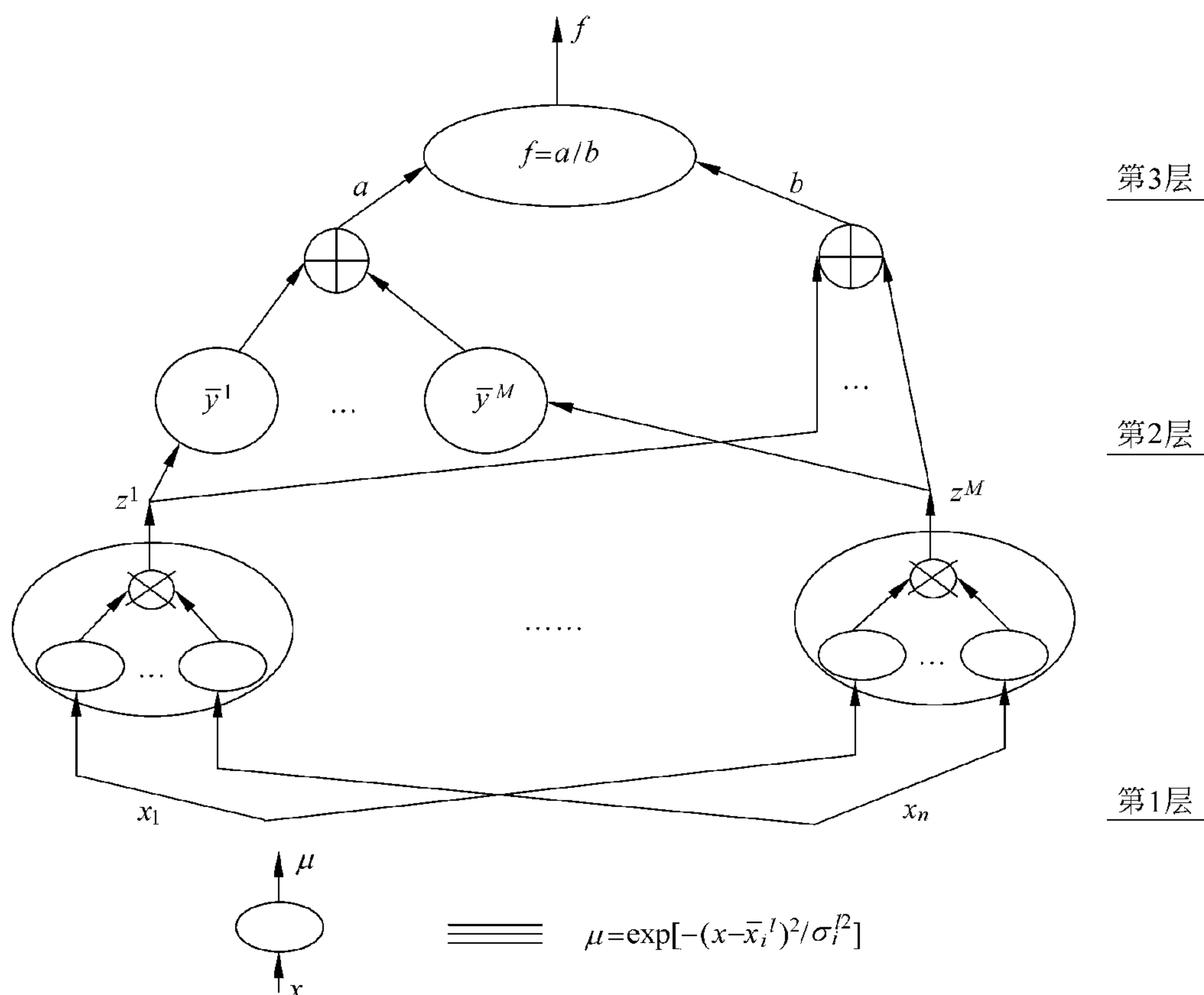


图 13.1 模糊系统的网络示意图

13.2 模糊系统参数的设计

与第12章一样,数据是由式(12.1)给定的输入—输出数据对,任务是设计一个形如式(13.1)的模糊系统 $f(x)$,使得下面的拟合误差最小

$$e^p = \frac{1}{2} [f(x_0^p) - y_0^p]^2 \quad (13.2)$$

即,任务是确定参数 $\bar{y}^l, \bar{x}_i^l, \sigma_i^l$ 使式(13.2)中的 e^p 最小。接下来,分别用 e, f 和 y 来表示 $e^p, f(x_0^p)$ 和 y_0^p 。

下面用梯度下降法来确定参数。具体地讲,就是用下面的算法来确定 \bar{y}^l

$$\bar{y}^l(q+1) = \bar{y}^l(q) - \alpha \left. \frac{\partial e}{\partial \bar{y}^l} \right|_q \quad (13.3)$$

式中, $l = 1, 2, \dots, M, q = 0, 1, 2, \dots, \alpha$ 为定步长。如果 q 趋于无穷时, $\bar{y}^l(q)$ 收敛,则由式(13.3)可知,在收敛的 \bar{y}^l 处有 $\frac{\partial e}{\partial \bar{y}^l} = 0$,这表明收敛点 \bar{y}^l 是 e 的一个局部极小点。由图 13.1

可知, f (于是 e 亦如此) 仅通过 a 依赖于 \bar{y}^l , 其中, $f = a/b$, $a = \sum_{l=1}^M (\bar{y}^l z^l)$, $b = \sum_{l=1}^M z^l$, $z^l = \prod_{i=1}^n \exp\left(-\left(\frac{x_i^l - \bar{x}_i^l}{\sigma_i^l}\right)^2\right)$; 因此, 根据复合函数求导规则, 有

$$\frac{\partial e}{\partial \bar{y}^l} = (f - y) \frac{\partial f}{\partial a} \frac{\partial a}{\partial \bar{y}^l} = (f - y) \frac{1}{b} z^l \quad (13.4)$$

把式(13.4)代入式(13.3), 即可得 \bar{y}^l 的学习算法为

$$\bar{y}^l(q+1) = \bar{y}^l(q) - \alpha \frac{f - y}{b} z^l \quad (13.5)$$

式中, $l = 1, 2, \dots, M, q = 0, 1, 2, \dots$ 。

用下式确定 \bar{x}_i^l

$$\bar{x}_i^l(q+1) = \bar{x}_i^l(q) - \alpha \left. \frac{\partial e}{\partial \bar{x}_i^l} \right|_q \quad (13.6)$$

式中, $i = 1, 2, \dots, n, l = 1, 2, \dots, M, q = 0, 1, 2, \dots$ 。由图 13.1 可以看出, f (于是 e 亦如此) 仅通过 z^l 依赖于 \bar{x}_i^l 。所以, 根据复合函数求导规则, 有

$$\frac{\partial e}{\partial \bar{x}_i^l} = (f - y) \frac{\partial f}{\partial z^l} \frac{\partial z^l}{\partial \bar{x}_i^l} = (f - y) \frac{\bar{y}^l - f}{b} z^l \frac{2(x_{0i}^p - \bar{x}_i^l)}{\sigma_i^{l2}} \quad (13.7)$$

把式(13.7)代入式(13.6), 可得 \bar{x}_i^l 的学习算法为

$$\bar{x}_i^l(q+1) = \bar{x}_i^l(q) - \alpha \frac{f - y}{b} (\bar{y}^l(q) - f) z^l \frac{2(x_{0i}^p - \bar{x}_i^l(q))}{\sigma_i^{l2}(q)} \quad (13.8)$$

式中, $i = 1, 2, \dots, n, l = 1, 2, \dots, M, q = 0, 1, 2, \dots$ 。

用同样的步骤, 可得 σ_i^l 的学习算法为

$$\begin{aligned} \sigma_i^l(q+1) &= \sigma_i^l(q) - \alpha \left. \frac{\partial e}{\partial \sigma_i^l} \right|_q \\ &= \sigma_i^l(q) - \alpha \frac{f - y}{b} (\bar{y}^l(q) - f) z^l \frac{2(x_{0i}^p - \bar{x}_i^l(q))^2}{\sigma_i^{l3}(q)} \end{aligned} \quad (13.9)$$

式中, $i = 1, 2, \dots, n, l = 1, 2, \dots, M, q = 0, 1, 2, \dots$ 。

学习算法(13.5)、(13.8)和(13.9)完成的是一个误差反向传播程序。为了训练 \bar{y}^l , 标准误差 $(f - y)/b$ 被反向传播到 \bar{y}^l 所在层; 则 \bar{y}^l 可用式(13.5)来调整, 这里, z^l 是 \bar{y}^l 的输入 (参见图 13.1)。为了训练 \bar{x}_i^l 和 σ_i^l , 标准误差 $(f - y)/b$ 与 $\bar{y}^l - f$ 及 z^l 的乘积被反向传播到第一层的处理单元 (其输出为 z^l); 则 \bar{x}_i^l 和 σ_i^l 可分别用式(13.8)和式(13.9)来调整, 余下变量 \bar{x}_i^l, x_{0i}^p 和 σ_i^l (即式(13.8)和式(13.9)右边的除 $\frac{(f - y)}{b} (\bar{y}^l - f) z^l$ 以外的变量) 也可被局部地

得到。因此, 也称这一算法为误差反向传播学习算法 (error back-propagation training algorithm)。

现在,将这一算法总结如下:

用梯度下降法设计模糊系统。

步骤 1 结构的确定和初始参数的设置。选择形如式(13.1)的模糊系统并确定 M 。 M 越大,产生的参数越多,运算也就越复杂,但给出的逼近精度越高。设定初始参数 $\bar{y}^l(0)$, $\bar{x}_i^l(0)$, $\sigma_i^l(0)$ 。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。对于特殊的应用可用特殊的方法(见 13.3 节中的例子)。

步骤 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 $(x_0^p; y_0^p)$, $p = 1, 2, \dots$, 在学习的第 q ($q = 0, 1, 2, \dots$) 阶段,把输入 x_0^p 作为图 13.1 中的模糊系统的输入层,然后计算第 1 层 ~ 第 3 层的输出。即,计算

$$z^l = \prod_{i=1}^n \exp\left(-\left(\frac{x_{0i}^p - \bar{x}_i^l(q)}{\sigma_i^l(q)}\right)^2\right) \quad (13.10)$$

$$b = \sum_{l=1}^M z^l \quad (13.11)$$

$$a = \sum_{l=1}^M \bar{y}^l(q) z^l \quad (13.12)$$

$$f = a/b \quad (13.13)$$

步骤 3 调整参数。采用学习算法(13.5)、(13.8)和(13.9)计算要调整的参数 $\bar{y}^l(q+1)$, $\bar{x}_i^l(q+1)$, $\sigma_i^l(q+1)$, 其中 $y = y_0^p$, z^l, b, a, f 都等于步骤 2 中算出的 z^l, b, a, f 。

步骤 4 令 $q = q + 1$ 返回步骤 2 重新计算,直至误差 $|f - y_0^p|$ 小于一个很小的数 ε , 或直至 q 等于一个预先指定的。

步骤 5 令 $p = p + 1$, 返回步骤 2 重新计算。即,用下一个输入—输出数据对 $(x_0^{p+1}; y_0^{p+1})$ 来调整参数。

步骤 6 如果有必要的话,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。

因为学习算法(13.5)、(13.8)和(13.9)属于梯度下降算法,所以初始参数的选择对于算法的成功是很关键的。如果所选择的初始参数接近于最优参数,则算法就很可能收敛于最优解; 否则,算法可能收敛于一个非最优解甚至不收敛。使用模糊系统的优点在于,有办法来选择较好的初始参数。要记住,参数 \bar{y}^l 是规则的 THEN 部分的模糊集的中心,而参数 \bar{x}_i^l 和 σ_i^l 是规则中 IF 部分的高斯模糊集的中心和宽度。因此,如果给定一个形如式(13.1)的模糊系统,则参数 \bar{y}^l, \bar{x}_i^l 和 σ_i^l 的物理意义不仅可以用来修复构造模糊系统的模糊 IF-THEN 规则,而且还有助于以一种用户友好的方式解释所设计的模糊系统。

下面将此方法应用于非线性动态系统辨识问题。

13.3 在非线性动态系统辨识中的应用

13.3.1 辨识器的设计

系统辨识是根据传感器测量结果确定一个合适的系统模型的过程。这一过程之所以重要,是因为工程中的许多方法都要依赖于系统模型。由于模糊系统是一种强有力的万能逼近器,所以把它作为非线性系统的辨识模型是合适的。本节将采用带有学习算法(13.5)、(13.8)和(13.9)的模糊系统(13.1)来逼近动态系统中未知的非线性部分动态的特性。

考虑如下离散时间非线性动态系统:

$$y(k+1) = f(y(k), \dots, y(k-n+1); u(k), \dots, u(k-m+1)) \quad (13.14)$$

式中, f 是一个要辨识的未知函数, u 和 y 分别是系统的输入和输出, n 和 m 是正整数。任务是根据模糊系统辨识未知函数 f 。

令 $\hat{f}(x)$ 为形如式(13.1)的模糊系统。用 $\hat{f}(x)$ 取代式(13.14)中的 $f(x)$, 于是可得下面的辨识模型:

$$\hat{y}(k+1) = \hat{f}(y(k), \dots, y(k-n+1); u(k), \dots, u(k-m+1)) \quad (13.15)$$

任务是调整 $\hat{f}(x)$ 中的参数, 使得当 k 趋于无穷时, 辨识模型 $\hat{y}(k+1)$ 的输出收敛于真实系统 $y(k+1)$ 的输出。图 13.2 给出了这一辨识方案。

该问题的输入—输出数据对是 $(x_0^{k+1}; y_0^{k+1})$, 其中

$$\begin{aligned} x_0^{k+1} &= (y(k), \dots, y(k-n+1); \\ &\quad u(k), \dots, u(k-m+1)), \\ y_0^{k+1} &= y(k+1), k=0, 1, 2, \dots \end{aligned}$$

因为系统是动态的, 所以这些输入—输出数据对可以在某一个时刻聚到一起。辨识过程的运算同 13.2 节中步骤 1 ~ 步骤 5 是一样的, 只不过 13.2 节中的 p 在这里是式(13.14)和式(13.15)中的 k , 式(13.1)中的 n 在这里等于 $n+m$ 。

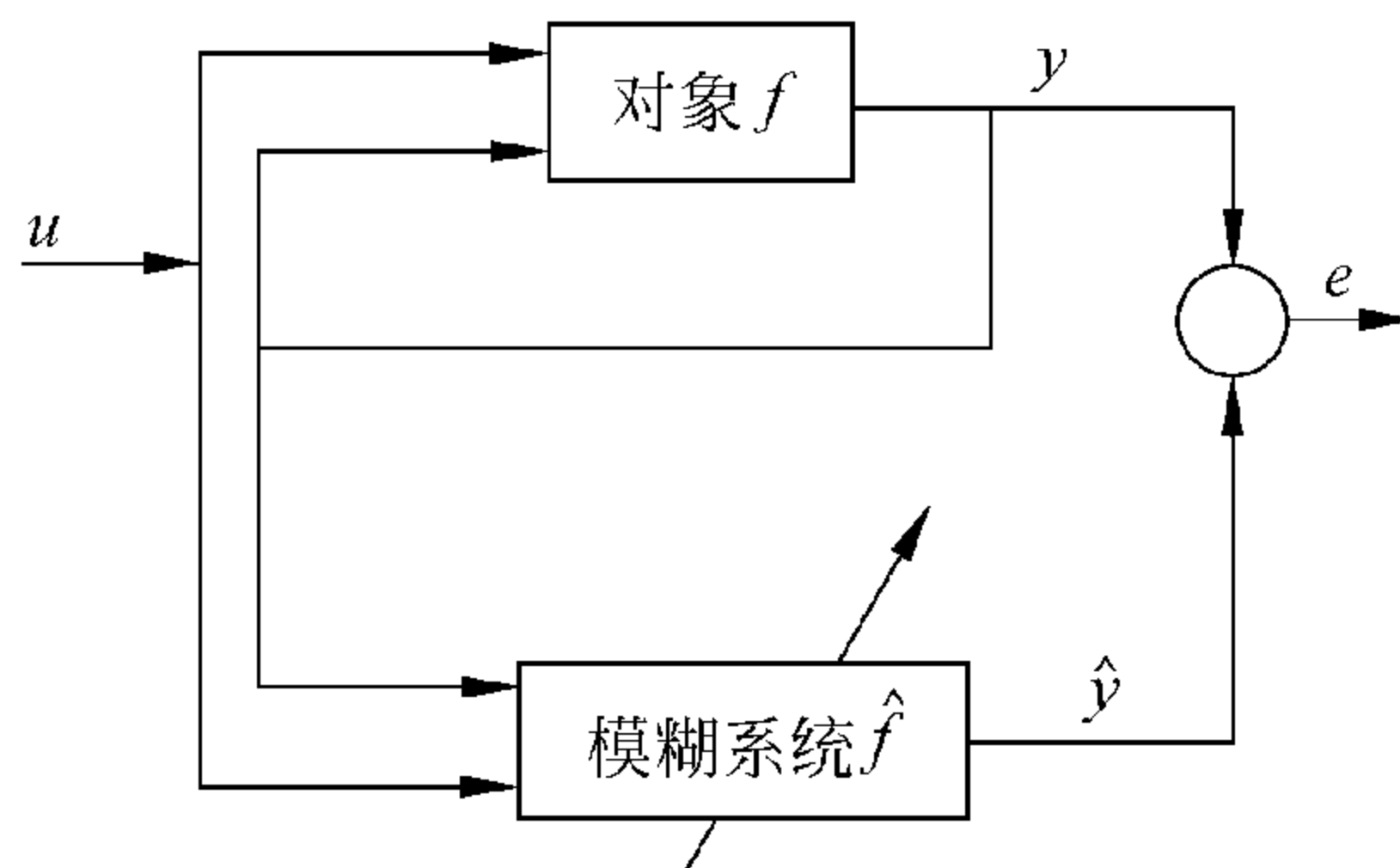


图 13.2 采用模糊系统建立非线性动态系统的辨识模型的基本方案

13.3.2 初始参数的选择

正如在13.2节中讨论的那样,一个好的初始 \hat{f} 对于梯度下降法的成功是很关键的。对于这类特殊的辨识问题,提出以下方法来在线选择初始参数,并提供一个理论上的证明(引理13.1)来解释为什么这是一个好方法。

在线初始参数选择方法:针对最初 M 个时刻点 $k=0,1,\dots,M-1$,收集输入—输出数据对 $(x_0^{k+1}; y_0^{k+1}) = (y(k), \dots, y(k-n+1), u(k), \dots, u(k-m+1); y(k+1))$,当 $k=M-1$ 时(即令式(13.5),式(13.8)和式(13.9)中的 $q=k-M$)开始此学习算法。设初始参数为 $\bar{y}^l(0) = y_0^l, \bar{x}_i^l(0) = x_{0i}^l, \sigma_i^l(0)$ 是很小的数(见引理13.1)或设 $\sigma_i^l(0) = [\max(x_{0i}^l: l=1,2,\dots,M) - \min(x_{0i}^l: l=1,2,\dots,M)]/M$,其中 $l=1,2,\dots,M, i=1,2,\dots,n+m$ 。设定 $\sigma_i^l(0)$ 的第二种方法是令隶属度函数均匀地覆盖 x_{0i}^l 由 $l=1$ 至 $l=M$ 的所有区域。

现要证明的是,只要 $\sigma_i^l(0)$ 的取值足够小,具有上述初始参数的模糊系统就能将所有 M 对输入—输出数据 $(x_0^{k+1}; y_0^{k+1}) (k=0,1,\dots,M-1)$ 拟合到任意精度。

引理13.1 对于任意 $\varepsilon > 0$,都存在 $\sigma^* > 0$,使得形如式(13.1)的模糊系统 $\hat{f}(x)$ (具有上述初始参数 \bar{y}_i^l, \bar{x}_i^l 和 $\sigma_i^l = \sigma^*$),在 $k=0,1,\dots,M-1$ 时,具有如下性质

$$|\hat{f}(x_0^{k+1}) - y_0^{k+1}| < \varepsilon \quad (13.16)$$

证明 把初始参数 $\bar{y}^l(0), \bar{x}_i^l(0)$ 代入式(13.1)并令 $\sigma_i^l = \sigma^*$,可得

$$\hat{f}(x) = \frac{\sum_{l=1}^M y_0^l \left[\prod_{i=1}^{n+m} \exp\left(-\left(\frac{x_i - x_{0i}^l}{\sigma^*}\right)^2\right)\right]}{\sum_{l=1}^M \left[\prod_{i=1}^{n+m} \exp\left(-\left(\frac{x_i - x_{0i}^l}{\sigma^*}\right)^2\right)\right]} \quad (13.17)$$

令式(13.17)中的 $x = x_0^{k+1}$,并由 $1 \leq k+1 \leq M$,可得

$$\begin{aligned} \hat{f}(x_0^{k+1}) &= \frac{\sum_{l=1}^M y_0^l \left[\prod_{i=1}^{n+m} \exp\left(-\left(\frac{x_{0i}^{k+1} - x_{0i}^l}{\sigma^*}\right)^2\right)\right]}{\sum_{l=1}^M \left[\prod_{i=1}^{n+m} \exp\left(-\left(\frac{x_{0i}^{k+1} - x_{0i}^l}{\sigma^*}\right)^2\right)\right]} \\ &= \frac{y_0^{k+1} + \sum_{l=1, l \neq k+1}^M y_0^l \left[\prod_{i=1}^{n+m} \exp\left(-\left(\frac{x_{0i}^{k+1} - x_{0i}^l}{\sigma^*}\right)^2\right)\right]}{1 + \sum_{l=1, l \neq k+1}^M \left[\prod_{i=1}^{n+m} \exp\left(-\left(\frac{x_{0i}^{k+1} - x_{0i}^l}{\sigma^*}\right)^2\right)\right]} \end{aligned} \quad (13.18)$$

由此可得

$$|\hat{f}(x_0^{k+1}) - y_0^{k+1}| = \left| \frac{\sum_{l=1, l \neq k+1}^M (y_0^l - y_0^{k+1}) \left[\prod_{i=1}^{n+m} \exp\left(-\left(\frac{x_{0i}^{k+1} - x_{0i}^l}{\sigma^*}\right)^2\right)\right]}{1 + \sum_{l=1, l \neq k+1}^M \left[\prod_{i=1}^{n+m} \exp\left(-\left(\frac{x_{0i}^{k+1} - x_{0i}^l}{\sigma^*}\right)^2\right)\right]} \right| \quad (13.19)$$

如果 $x_0^{l_1} \neq x_0^{l_2}$ ($l_1 \neq l_2$), 则当 $\sigma^* \rightarrow 0$ 时, 有 $\prod_{i=1}^{n+m} \exp\left(-\left(\frac{x_{0i}^{k+1} - x_{0i}^l}{\sigma^*}\right)^2\right) \rightarrow 0$ ($l \neq k+1$)。所以, 只

要令 σ^* 足够小, 就可使 $|\hat{f}(x_0^{k+1}) - y_0^{k+1}| < \varepsilon$ 成立。类似地, 可以证明在 $x_0^l = x_0^{k+1}$ ($l \neq k+1$) 的条件下, 式(13.16)仍然成立。这留作习题。

根据引理 13.1 可以认为, 上面的初始参数的选择方法是一个好方法, 因为具有这些初始参数的模糊系统至少使最初的 M 对输入—输出数据达到了拟合精度。如果最初的 M 对输入—输出数据包含了未知非线性函数 $f(x)$ 的重要特性, 那么就可以期望在由时刻点 M 开始训练之后, 模糊辨识器会非常迅速地收敛于未知的非线性系统。事实上, 在下一小节的仿真中, 仿真结果也证实的确如此。不过, 不能把 σ_i^l 选的太小, 因为尽管 σ_i^l 很小时, 模糊系统可以让最初的 M 对数据拟合得很好, 但它对于其他的输入—输出数据对可能会产生较大的拟合误差。所以, 在接下来的仿真中, 将使用在线初始参数选择法所描述的第二种 σ_i^l 的选择方法。

13.3.3 仿真

例 13.1 辨识对象满足如下差分方程:

$$y(k+1) = 0.3y(k) + 0.6y(k-1) + g[u(k)] \quad (13.20)$$

式中, 未知函数为 $g(u) = 0.6\sin(\pi u) + 0.3\sin(3\pi u) + 0.1\sin(5\pi u)$ 。由式(13.15)可知, 辨识模型满足如下差分方程:

$$\hat{y}(k+1) = 0.3y(k) + 0.6y(k-1) + \hat{f}[u(k)] \quad (13.21)$$

式中, $\hat{f}[*]$ 形如式(13.1), $M=10$ 。在学习算法(13.5), (13.8)和(13.9)中, 取 $\alpha=0.5$, 并选用在线参数选择方法。辨识器从时刻 $k=10$ 开始进行学习, 然后调整一个周期内每个时刻点的参数 \bar{y}^l , \bar{x}_i^l 和 σ_i^l 。即, 在每个时刻用式(13.5)、式(13.8)和式(13.9)进行一次运算。图 13.3 给出了学习过程终止于 $k=200$ 时系统的输出和辨识模型的输出, 此时输入为 $u(k) = \sin(2\pi k/200)$ 。由图 13.3 可以看出, 辨识模型的输出几乎立刻就跟踪上了系统的输出并且一直保持这种跟踪状态到终点 $k=200$ 。

例 13.2 本例将说明模糊辨识器对于一个多维输入多维输出的系统是怎样进行工作的, 模糊辨识器可由如下方程描述:

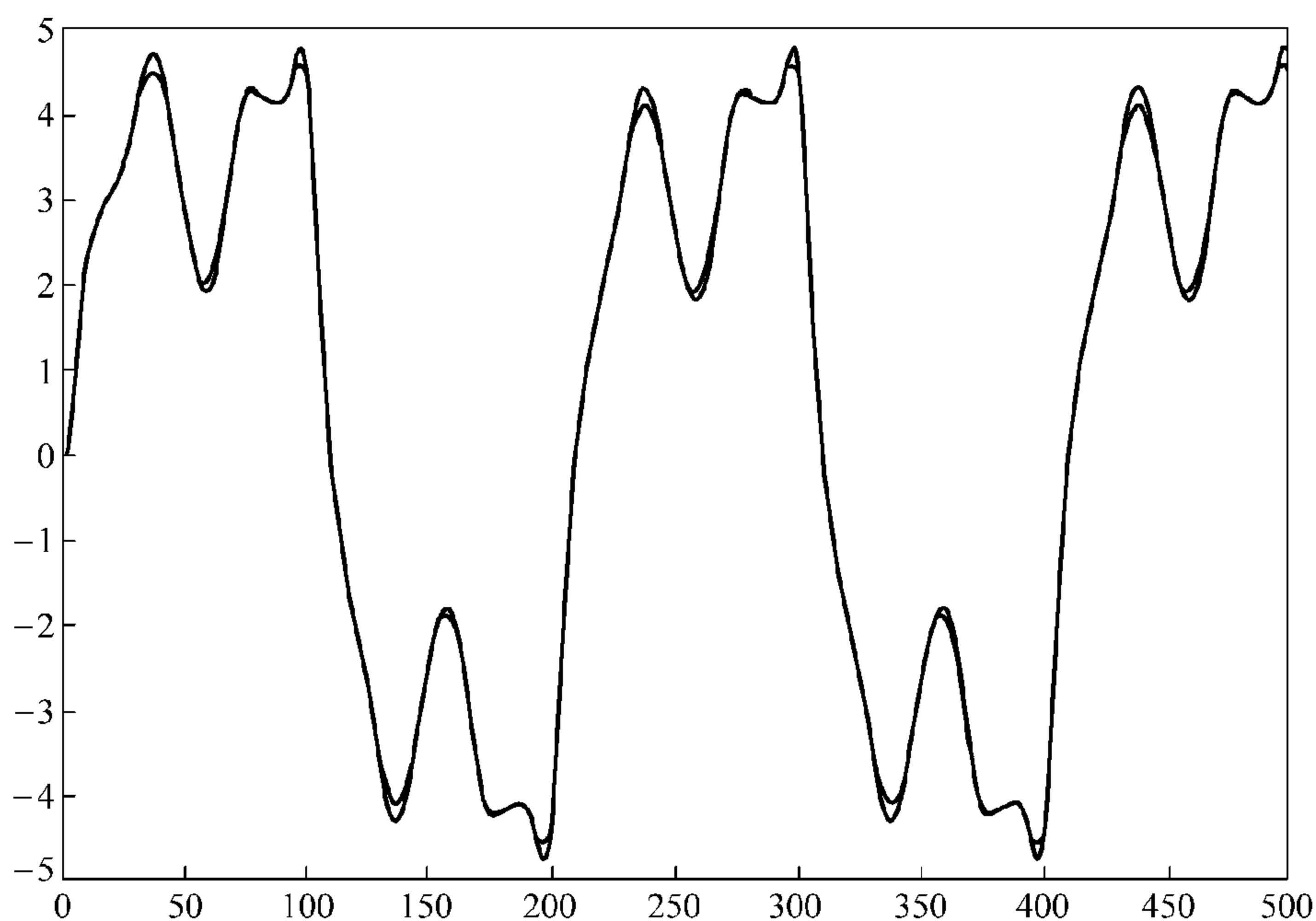


图 13.3 例 13.1 中的学习过程终止于 $k=200$ 时的系统的输出和辨识模型的输出

$$\begin{bmatrix} y_1(k+1) \\ y_2(k+1) \end{bmatrix} = \begin{bmatrix} \frac{y_1(k)}{1+y_2^2(k)} \\ \frac{y_1(k)y_2(k)}{1+y_2^2(k)} \end{bmatrix} + \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix} \quad (13.22)$$

辨识模型由 \hat{f}^1 和 \hat{f}^2 两个模糊系统组成, 它们分别由如下方程描述:

$$\begin{bmatrix} \hat{y}_1(k+1) \\ \hat{y}_2(k+1) \end{bmatrix} = \begin{bmatrix} \hat{f}^1(y_1(k), y_2(k)) \\ \hat{f}^2(y_1(k), y_2(k)) \end{bmatrix} + \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix} \quad (13.23)$$

\hat{f}^1 和 \hat{f}^2 都形如式(13.1), $M=121$ 。下面, 利用随机输入对模糊辨识器进行 5 000 步的学习, 其中随机输入 $u_1(k)$ 和 $u_2(k)$ 选用均匀分布在 $[-1, 1]$ 上的随机数, 并取 $\alpha=0.5$ 。初始参数由在线初始参数选择方法确定, 学习过程每循环一步参数就调整一次。图 13.4 和图 13.5 分别给出了辨识对象和辨识学习模型的输出 $y_1(k)$, $\hat{y}_1(k)$ 和 $y_2(k)$, $\hat{y}_2(k)$, 其中辨识模型的输入向量为 $[u_1(k), u_2(k)] = [\sin(2\pi k/25), \cos(2\pi k/25)]$ 。可以看出, 模糊辨识器几乎是完美地跟踪着真实系统。

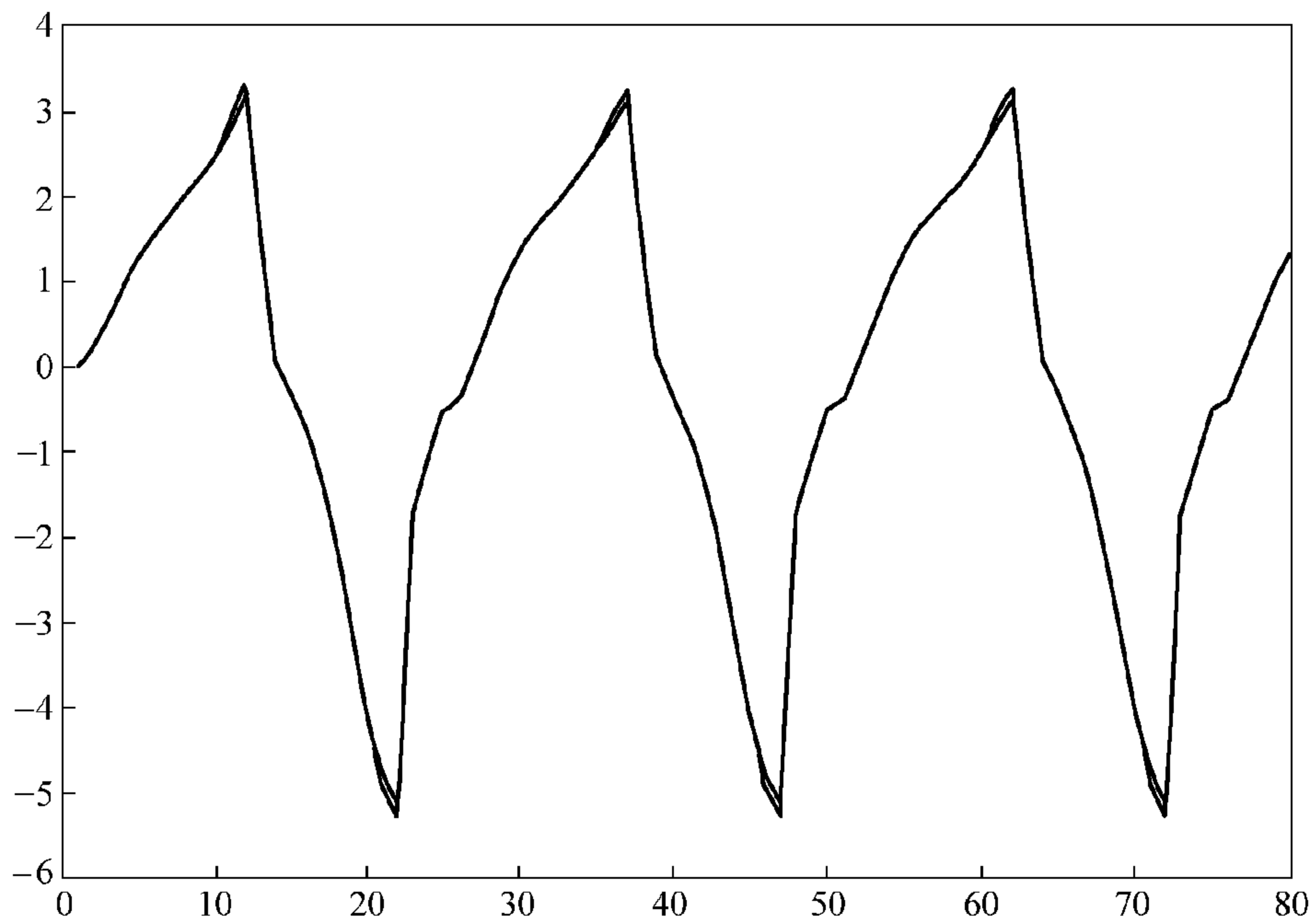


图 13.4 例 13.2 中经过 5 000 步训练后辨识对象的输出 $y_1(k)$ 和辨识学习模型的输出 $\hat{y}_1(k)$

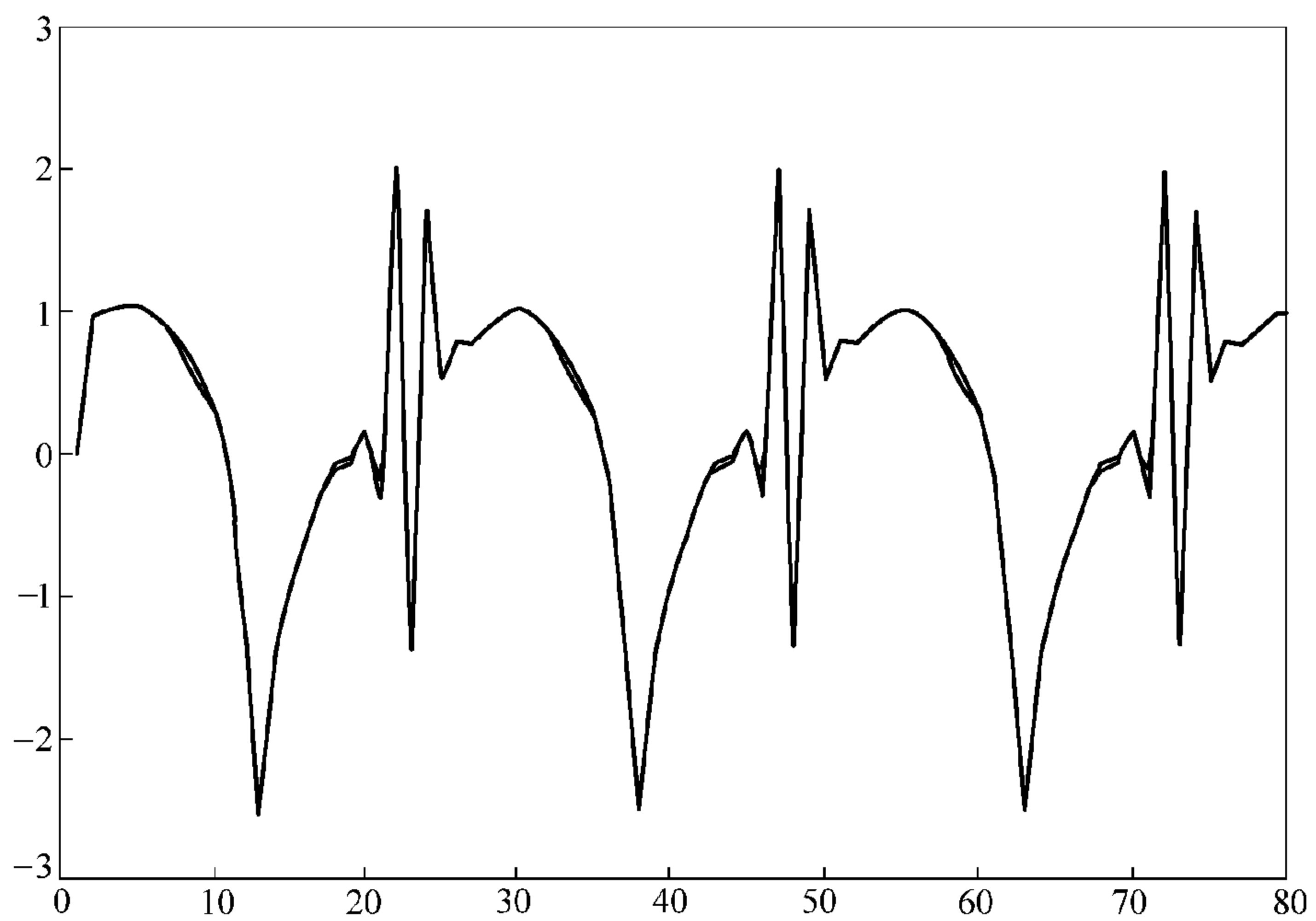


图 13.5 例 13.2 中经过 5 000 步训练后辨识对象的输出 $y_2(k)$ 和辨识学习模型的输出 $\hat{y}_2(k)$

13.4 本章小结及导读

本章主要阐述了以下内容:

- 梯度下降学习算法的推导;
- 选择模糊辨识模型的初始参数的方法及证明(见引理 13.1);
- 用梯度下降学习算法和初始参数选择方法设计模糊系统;
- 将梯度下降学习算法和初始参数选择方法应用于非线性动态系统辨识问题和其他问题。

本章中的方法是受神经网络中的误差反向传播算法(见 Werbos[1974])的启发而得来的。在文献中还可找到许多类似的方法(参见 Jang[1993]和 Lin[1994])。更多的仿真结果可以在 Wang[1994]中找到。Narendra 和 Parthasarathy[1990]给出了神经网络辨识器辨识非线性系统(如例 13.1 和例 13.2)的更广泛的仿真结果。

13.5 习题与设计

习题 13.1 假设式(13.1)中的参数 \bar{x}_i^l 和 σ_i^l 是常数, \bar{y}^l 是可以自由变化的参数。

(a) 证明如果学习算法(13.5)收敛,则它收敛于式(13.2)中的 e^p 的全局最小值。

(b) 令 $e^p(\bar{y}^l)$ 为式(13.2)中的 e^p ,通过优化 $e^p\left[\bar{y}^l(q) - \alpha \frac{f - y^l}{b} z^l\right]$ 找到最优的步长 α 。

习题 13.2 为什么当 $x_0^l = x_0^{k+1}$ ($l \neq k+1$) 时,引理 13.1 的证明是不成立的? 证明此条件下的引理 13.1。

习题 13.3 假设给定了 k 对输入—输出数据 $(x_0^p; y_0^p)$ ($p = 1, 2, \dots, k$), 现想设计一个形如式(13.1)的模糊系统 $f(x)$, 使得下面的误差平方和最小:

$$J = \sum_{p=1}^k [f(x_0^p) - y_0^p]^2 \quad (13.24)$$

令参数 \bar{x}_i^l 和 σ_i^l 为常数, 仅 \bar{y}^l 可以自由变化。确定使 J 最小的最优 \bar{y}^l ($l = 1, 2, \dots, M$) (用闭型表达式来表达最优的 \bar{y}^l)。

习题 13.4 令 $[x(k)]$ 为一个由梯度下降法产生的实值向量序列

$$x(k+1) = x(k) - \alpha \nabla e(x(k)) \quad (13.25)$$

式中, $e: R^n \rightarrow R$ 是一个成本函数, $e \in C^2$ (如 e 有连续的二阶导数)。假设对于某个紧致集 D 的所有 $x(k) \in D \subset R^n$, 都存在 $\varepsilon > 0$ 和 $L > 0$ 使得如果

$$0 < \varepsilon \leq \alpha \leq \frac{2 - \varepsilon}{L} \quad (13.26)$$

那么

- (a) 如果 $\nabla e(x(k)) \neq 0$, 则 $e(x(k+1)) < e(x(k))$;
- (b) 如果 $e(x)$ 是有下边界的, 则当 $k \rightarrow \infty$ 时, 有 $x(k) \rightarrow x^* \in D$;
- (c) x^* 是 $e(x)$ 的局部极小值。

习题 13.5 众所周知, 13.2 节中的梯度下降法是收敛于局部极小点的。创建一个学习样本集合, 使之能从数值上说明这一现象。即, 证明在不同的初始条件下, 最终可能产生不同的收敛解。

习题 13.6 解释显性知识和隐性知识是怎样用 13.2 节中的设计方法组合到模糊系统中去的。显性知识可能在最终所设计的模糊系统中不出现吗? 为什么? 如果想在最终的模糊系统中出现显性知识, 怎样修正 13.2 节中的设计程序?

13.7(设计) 编写一个计算机程序来实现学习算法(13.5)、(13.8)和(13.9), 并将你的编码应用于第 12 章中的时间序列预测问题。

递推最小二乘法设计模糊系统

14.1 模糊系统的设计

第 13 章中的梯度下降算法试图使式(13.2)中的 e^p 达到最小,但它仅考虑了某一输入—输出数据对 $(x_0^p; y_0^p)$ 的拟合误差。即,这种学习算法是在某一时刻通过调整参数以拟合输入—输出数据对的。本章将提出另一种学习算法,它能令所有由 1 至 p 的输入—输出数据对的拟合误差之和达到最小。即,现在的目标是设计一个模糊系统 $f(x)$,使得下式最小

$$J_p = \sum_{j=1}^p [f(x_0^j) - y_0^j]^2 \quad (14.1)$$

此外,还要递推地设计该模糊系统,即如果 f_p 就是使 J_p 最小的模糊系统,则 f_p 应该可以被表述为 f_{p-1} 的函数。下面,用递推最小二乘法来设计模糊系统。

用递推最小二乘法设计模糊系统:

步骤 1 假设 $U = [\alpha_1, \beta_1] \times \cdots \times [\alpha_n, \beta_n] \subset R^n$ 。在每个区间 $[\alpha_i, \beta_i]$ ($i = 1, 2, \dots, n$) 上定义 N_i 个模糊集 $A_i^{l_i}$ ($l_i = 1, 2, \dots, N_i$), 它们在 $[\alpha_i, \beta_i]$ 上是完备模糊集。如可选 $A_i^{l_i}$ 为四边形模糊集: $\mu_{A_i^{l_i}}(x_i) = \mu_{A_i^{l_i}}(x_i; a_i^{l_i}, b_i^{l_i}, c_i^{l_i}, d_i^{l_i})$, 其中, $a_i^1 = b_i^1 = \alpha_i, c_i^j \leq a_i^{j+1} < d_i^j \leq b_i^{j+1}$ ($j = 1, 2, \dots, N_i - 1$), $c_i^{N_i} = d_i^{N_i} = \beta_i$ 。

步骤 2 根据如下形式的 $\prod_{i=1}^n N_i$ 条模糊 IF-THEN 规则来构造模糊系统

$$\text{如果 } x_1 \text{ 为 } A_1^{l_1} \text{ 且 } \cdots \text{ 且 } x_n \text{ 为 } A_n^{l_n}, \text{ 则 } y \text{ 为 } B^{l_1 \cdots l_n} \quad (14.2)$$

其中, $l_i = 1, 2, \dots, N_i$ ($i = 1, 2, \dots, n$), $B^{l_1 \cdots l_n}$ 是中心为 $\bar{y}^{l_1 \cdots l_n}$ (可自由变化) 的任意模糊集。具体地讲,就是选择带有乘积推理机、单值模糊器、中心平均解模糊器的模糊系统。即,所设计的模糊系统为

$$f(x) = \frac{\sum_{l_1=1}^{N_1} \cdots \sum_{l_n=1}^{N_n} \bar{y}^{l_1 \cdots l_n} \left[\prod_{i=1}^n \mu_{A_i^{l_i}}(x_i) \right]}{\sum_{l_1=1}^{N_1} \cdots \sum_{l_n=1}^{N_n} \left[\prod_{i=1}^n \mu_{A_i^{l_i}}(x_i) \right]} \quad (14.3)$$

其中, $\bar{y}^{l_1 \cdots l_n}$ 是要设计的自由参数, $A_i^{l_i}$ 在步骤 1 中给定。然后将自由参数 $\bar{y}^{l_1 \cdots l_n}$ 放到 $\prod_{i=1}^n N_i$ 维

向量中

$$\theta = (\bar{y}^{1\dots 1}, \dots, \bar{y}^{N_1 1\dots 1}, \bar{y}^{12 1\dots 1}, \dots, \bar{y}^{N_1 2 1\dots 1}, \dots, \bar{y}^{1 N_2 \dots N_n}, \dots, \bar{y}^{N_1 N_2 \dots N_n})^T \quad (14.4)$$

则式(14.3)可变为

$$f(x) = b^T(x)\theta \quad (14.5)$$

其中

$$b(x) = (b^{1\dots 1}(x), \dots, b^{N_1 1\dots 1}(x), b^{12 1\dots 1}(x), \dots, b^{N_1 2 1\dots 1}(x), \dots, b^{1 N_2 \dots N_n}(x), \dots, b^{N_1 N_2 \dots N_n}(x))^T \quad (14.6)$$

$$b^{l_1 \dots l_n}(x) = \frac{\prod_{i=1}^n \mu_{A_i^{l_i}}(x_i)}{\sum_{l_1=1}^{N_1} \dots \sum_{l_n=1}^{N_n} \left[\prod_{i=1}^n \mu_{A_i^{l_i}}(x_i) \right]} \quad (14.7)$$

步骤3 根据以下过程选择初始参数 $\theta(0)$: 如果专家(显性知识)能提供与式(14.2)的 IF 部分相同的语言规则, 则选择 $\bar{y}^{l_1 \dots l_n}(0)$ 为这些语言规则中的 THEN 部分的模糊集中心; 否则, 在输出空间 $V \subset R$ 上任意选择 $\theta(0)$ (如, 选定 $\theta(0) = 0$ 或 $\theta(0)$ 中的元素在 V 上均匀分布)。由此可知, 最初的模糊系统是由显性知识组建而成的。

步骤4 当 $p = 1, 2, \dots$, 时, 用以下递推最小二乘法计算参数 θ :

$$\theta(p) = \theta(p-1) + K(p)[y_0^p - b^T(x_0^p)\theta(p-1)] \quad (14.8)$$

$$K(p) = P(p-1)b(x_0^p)[b^T(x_0^p)P(p-1)b(x_0^p) + 1]^{-1} \quad (14.9)$$

$$P(p) = P(p-1) - P(p-1)b(x_0^p)[b^T(x_0^p)P(p-1)b(x_0^p) + 1]^{-1}b^T(x_0^p)P(p-1) \quad (14.10)$$

式中, $\theta(0)$ 是在步骤3中选定的, $P(0) = \sigma I$ (σ 是一个很大的常数)。在所设计的形如式(14.3)的模糊系统中的参数 $\bar{y}^{l_1 \dots l_n}$ 等于 $\theta(p)$ 中的对应元素。

递归最小二乘算法(14.8)~算法(14.10)是通过令式(14.1)中的 J_p 最小而得到的, 其中 $(f(x_0^j))$ 形如式(14.3)。下面将会给出其推导过程。

14.2 递推最小二乘法的推导

令 $Y_0^{p-1} = (y_0^1, \dots, y_0^{p-1})^T$, $B_{p-1} = (b(x_0^1), \dots, b(x_0^{p-1}))^T$, 则由式(14.5), 可把 J_{p-1} 改写为

$$\begin{aligned} J_{p-1} &= \sum_{j=1}^{p-1} [f(x_0^j) - y_0^j]^2 \\ &= (B_{p-1}\theta - Y_0^{p-1})^T (B_{p-1}\theta - Y_0^{p-1}) \end{aligned} \quad (14.11)$$

因为 J_{p-1} 是 θ 的二次函数, 所以使 J_{p-1} 达到最小的最优 θ (用 $\theta(p-1)$ 表示) 为

$$\theta(p-1) = (B_{p-1}^T B_{p-1})^{-1} B_{p-1}^T Y_0^{p-1} \quad (14.12)$$

当已知输入—输出数据对 $(x_0^p; y_0^p)$ 时, 式(14.1)中 J_p 可改写为

$$J_p = \left[\begin{pmatrix} B_{p-1} \\ b^T(x_0^p) \end{pmatrix} \theta - \begin{pmatrix} Y_0^{p-1} \\ y_0^p \end{pmatrix} \right]^T \left[\begin{pmatrix} B_{p-1} \\ b^T(x_0^p) \end{pmatrix} \theta - \begin{pmatrix} Y_0^{p-1} \\ y_0^p \end{pmatrix} \right] \quad (14.13)$$

和式(14.12)类似, 使 J_p 达到最小的最优的 θ (用 $\theta(p)$ 表示) 可由下式得到

$$\begin{aligned} \theta(p) &= \left[(B_{p-1}^T b(x_0^p)) \begin{pmatrix} B_{p-1} \\ b^T(x_0^p) \end{pmatrix} \right]^{-1} (B_{p-1}^T b(x_0^p)) \begin{pmatrix} Y_0^{p-1} \\ y_0^p \end{pmatrix} \\ &= [B_{p-1}^T B_{p-1} + b(x_0^p) b^T(x_0^p)]^{-1} [B_{p-1}^T Y_0^{p-1} + b(x_0^p) y_0^p] \end{aligned} \quad (14.14)$$

为进一步简化式(14.14), 需使用矩阵恒等式

$$(P^{-1} + bb^T)^{-1} = P - Pb(b^T P b + 1)^{-1} b^T P \quad (14.15)$$

定义 $P(p-1) = (B_{p-1}^T B_{p-1})^{-1}$, 然后, 由式(14.15), 可将式(14.14)改写为

$$\begin{aligned} \theta(p) &= \{P(p-1) - P(p-1)b(x_0^p)[b^T(x_0^p)P(p-1)b(x_0^p) + 1]^{-1} b^T(x_0^p)P(p-1)\} \\ &\quad [B_{p-1}^T Y_0^{p-1} + b(x_0^p) y_0^p] \end{aligned} \quad (14.16)$$

因为 $P(p-1)B_{p-1}^T Y_0^{p-1} = (B_{p-1}^T B_{p-1})^{-1} B_{p-1}^T Y_0^{p-1} = \theta(p-1)$ (参见式(14.12)), 所以, 可将式(14.16)简化为

$$\begin{aligned} \theta(p) &= \theta(p-1) - P(p-1)b(x_0^p)[b^T(x_0^p)P(p-1)b(x_0^p) + 1]^{-1} b^T(x_0^p)\theta(p-1) \\ &\quad + P(p-1)b(x_0^p)[1 - (b^T(x_0^p)P(p-1)b(x_0^p) + 1)^{-1} b^T(x_0^p)P(p-1)b(x_0^p)] y_0^p \\ &= \theta(p-1) + P(p-1)b(x_0^p)[b^T(x_0^p)P(p-1)b(x_0^p) + 1]^{-1} \\ &\quad (y_0^p - b^T(x_0^p)\theta(p-1)) \end{aligned} \quad (14.17)$$

定义 $K(p) = P(p-1)b(x_0^p)[b^T(x_0^p)P(p-1)b(x_0^p) + 1]^{-1}$, 即可得到式(14.8)和式(14.9)。

最后, 推导式(14.10)。由定义, 可得

$$\begin{aligned} P(p) &= \left[(B_{p-1}^T b(x_0^p)) \begin{pmatrix} B_{p-1} \\ b^T(x_0^p) \end{pmatrix} \right]^{-1} \\ &= [B_{p-1}^T B_{p-1} + b(x_0^p) b^T(x_0^p)]^{-1} \end{aligned} \quad (14.18)$$

根据矩阵恒等式(14.15)和 $B_{p-1}^T B_{p-1} = P^{-1}(p-1)$, 即可由式(14.18)得到式(14.10)。

14.3 在非线性通信信道均衡中的应用

14.3.1 均衡问题及其几何公式

通信信道中的非线性畸变对于进一步提高快速传输数据的准确率来说是一个重要因素 (Biglieri, Gersho, Gitlin 和 Lim[1984])。由于通过非线性信道接收到的信号是原信号传输值的非线性函数, 而且非线性畸变随时间和位置的不同发生变化。所以, 有效的非线性信道

的均衡器应该非线性和自适应的。在本节中,我们利用递推最小二乘法设计一个作为非线性信道均衡器的模糊系统。

这里用到的数字通信系统原理如图 14.1 所示。信道包括发送端滤波器、传输介质、接收端匹配滤波器及其他元件。假设发送信号序列为一个以相同概率在 $\{-1, 1\}$ 取值的独立序列 $s(k)$ 。信道输出加入噪声后得到的 $x(k), x(k-1), \dots, x(k-n+1)$ 作为均衡器的输入。均衡器在采样时刻 k 的任务根据 $x(k), x(k-1), \dots, x(k-n+1)$ 中所包含的信息,估计发送信号 $s(k-d)$ 。这里,整数 n 和 d 分别为均衡器的阶数和延迟。

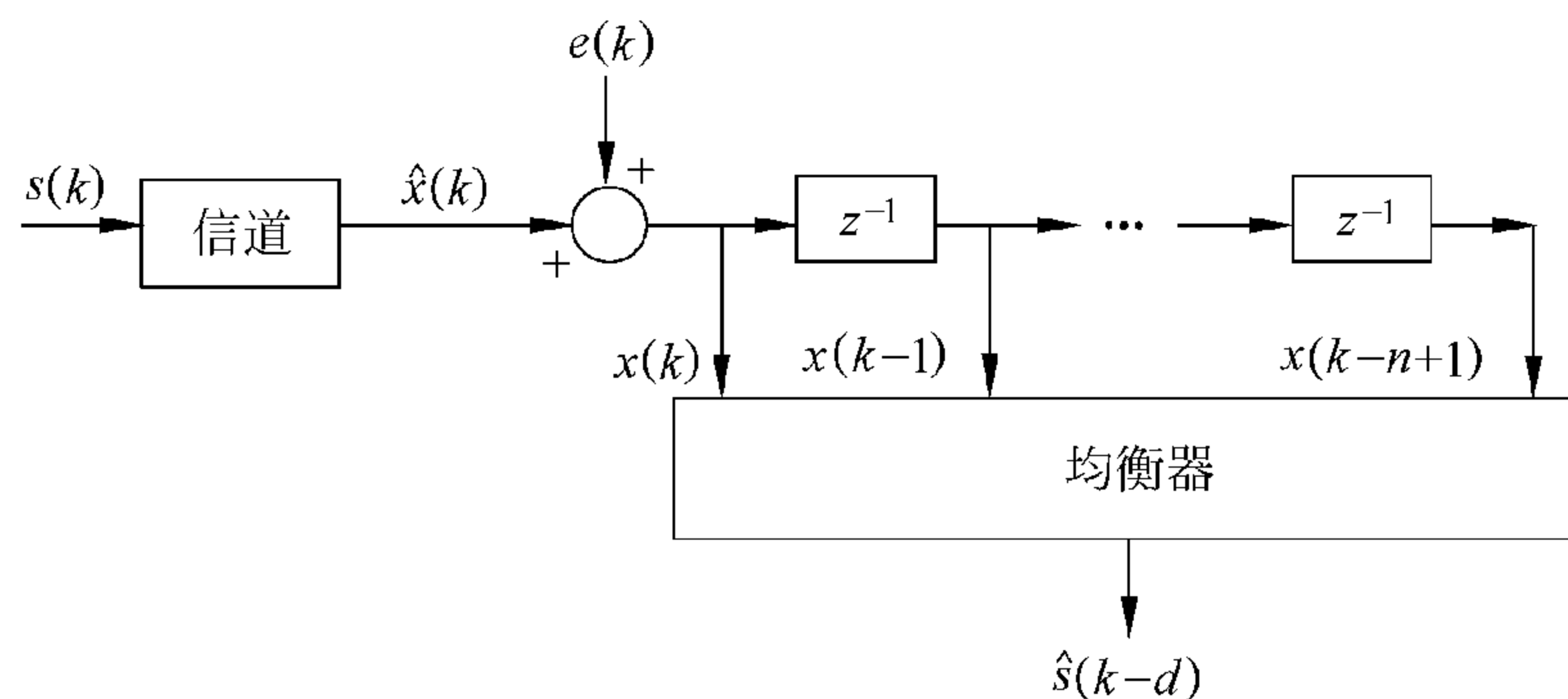


图 14.1 数据传输系统框图

下面仍沿用 Chen, Gibson, Cowan 和 Grand[1990]研究均衡问题的几何公式,其定义为

$$P_{n,d}(1) = \{\hat{x}(k) \in R^n \mid s(k-d) = 1\} \quad (14.19)$$

$$P_{n,d}(-1) = \{\hat{x}(k) \in R^n \mid s(k-d) = -1\} \quad (14.20)$$

式中

$$\hat{x}(k) = [\hat{x}(k), \hat{x}(k-1), \dots, \hat{x}(k-n+1)]^T \quad (14.21)$$

$\hat{x}(k)$ 为信道的无噪声输出(如图 14.1 所示), $P_{n,d}(1)$ 和 $P_{n,d}(-1)$ 分别表示信道的无噪声输出向量 $\hat{x}(k)$ 的两个集合,这两组向量分别从包含 $s(k-d) = 1$ 和 $s(k-d) = -1$ 的信道输入序列中产生。

均衡器可由下列函数来描述:

$$g_k: R^n \rightarrow \{-1, 1\} \quad (14.22)$$

$$\hat{s}(k-d) = g_k(x(k)) \quad (14.23)$$

式中

$$x(k) = [x(k), x(k-1), \dots, x(k-n+1)]^T \quad (14.24)$$

是信道的可观测输出向量。令 $p_1[x(k) \mid \hat{x}(k) \in P_{n,d}(1)]$ 和 $p_{-1}[x(k) \mid \hat{x}(k) \in P_{n,d}(-1)]$ 分别为 $x(k)$ 的条件概率密度函数,此时 $\hat{x}(k) \in P_{n,d}(1)$ 和 $\hat{x}(k) \in P_{n,d}(-1)$ 是给定的。Chen, Gibson, Cowan 和 Grand[1990]已经证明了,下式定义的均衡器在其阶数 n 和延迟 d

给定后能够使位误差率达到最小:

$$f_{opt}(x(k)) = \text{sgn}[p_1(x(k)|\hat{x}(k) \in P_{n,d}(1)) - p_{-1}(x(k)|\hat{x}(k) \in P_{n,d}(-1))] \quad (14.25)$$

式中, 如果 $y \geq 0$ ($y < 0$), 则符号函数 $\text{sgn}(y) = 1$ (-1)。如果噪声 $e(k)$ 为零均值高斯噪声, 其协方差矩阵为

$$Q = Ee(k), \dots, e(k-n+1)^T \quad (14.26)$$

由 $x(k) = \hat{x}(k) + e(k)$, 可得

$$\begin{aligned} & p_1[x(k)|\hat{x}(k) \in P_{n,d}(1)] - p_{-1}[x(k)|\hat{x}(k) \in P_{n,d}(-1)] \\ &= \sum \exp\left[-\frac{1}{2}(x(k) - \hat{x}_+)^T Q^{-1}(x(k) - \hat{x}_+)\right] \\ & \quad - \sum \exp\left[-\frac{1}{2}(x(k) - \hat{x}_-)^T Q^{-1}(x(k) - \hat{x}_-)\right] \end{aligned} \quad (14.27)$$

式中第一项(第二项)求和是对所有的 $\hat{x}_+ \in P_{n,d}(1)$ ($\hat{x}_- \in P_{n,d}(-1)$) 点的求和。

现在考虑如下形式的非线性信道:

$$\hat{x}(k) = s(k) + 0.5s(k-1) - 0.9[s(k) + 0.5s(k-1)]^3 \quad (14.28)$$

高斯噪声 $e(k)$ 的方差为 $E[e^2(k)] = 0.2$ 。当 $n=2, d=0$ 时, 最优的决策区域为

$$[x(k) \in R^2 | p_1[x(k)|\hat{x}(k) \in P_{2,0}(1)] - p_{-1}[x(k)|\hat{x}(k) \in P_{2,0}(-1)] \geq 0] \quad (14.29)$$

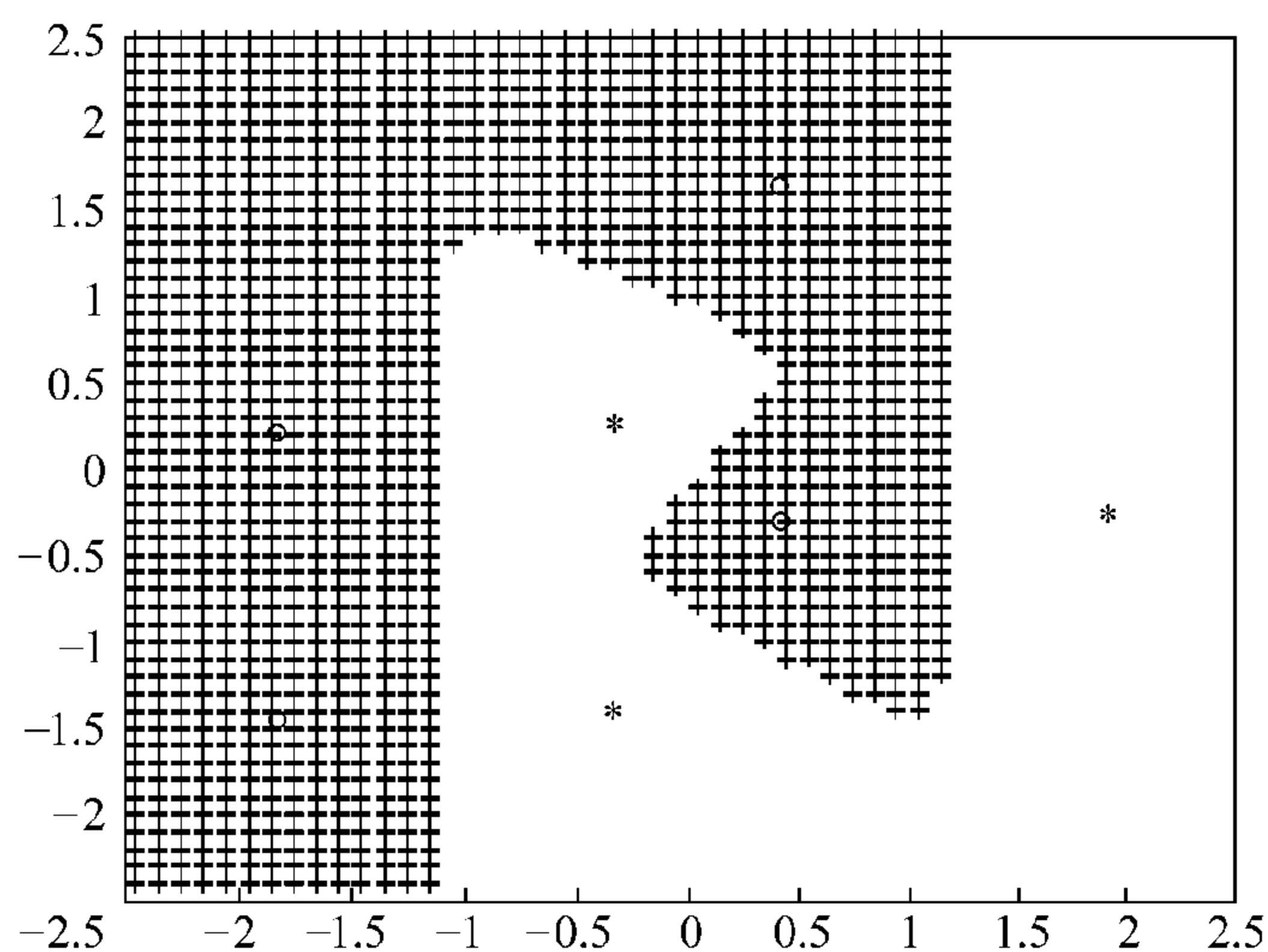
即为图 14.2 中的阴影部分, 集合 $P_{2,0}(1)$ 和 $P_{2,0}(-1)$ 中的元素分别用“o”和“*”表示。由图 14.2 可以看出, 本例中的最优决策域边界呈严格非线性。

14.3.2 模糊系统在均衡问题中的应用

使用将模糊系统(14.3)作为图 14.1 中的均衡器。运算过程包括以下两个阶段:

- 学习阶段。在此阶段, 发送信号 $s(k)$ 已知, 任务是设计均衡器(即模糊系统(14.3))。采用 14.1 节中的设计方法来确定均衡器的结构和参数。此问题的输入—输出数据对为: $x_0^k = (x(k), \dots, x(k-n+1))^T$ 和 $y_0^k = s(k-d)$ 。14.1 节中的指标 p 在这里变成了时间指标 k 。
- 应用阶段。在此阶段, 发送信号 $s(k)$ 未知, 用所设计的均衡器(模糊系统(14.3))来估计 $s(k-d)$ 。具体地讲, 如果模糊系统的输出大于等于零, 则估计 $\hat{s}(k-d) = 1$; 否则 $\hat{s}(k-d) = -1$ 。

例 14.1 考虑非线性信道(14.28)。假定 $n=2, d=0$, 最优决策域如图 14.2 所示。任务是设计一个模糊系统, 其输入—输出行为和图 14.2 中的输入—输出行为相接近, 这里模



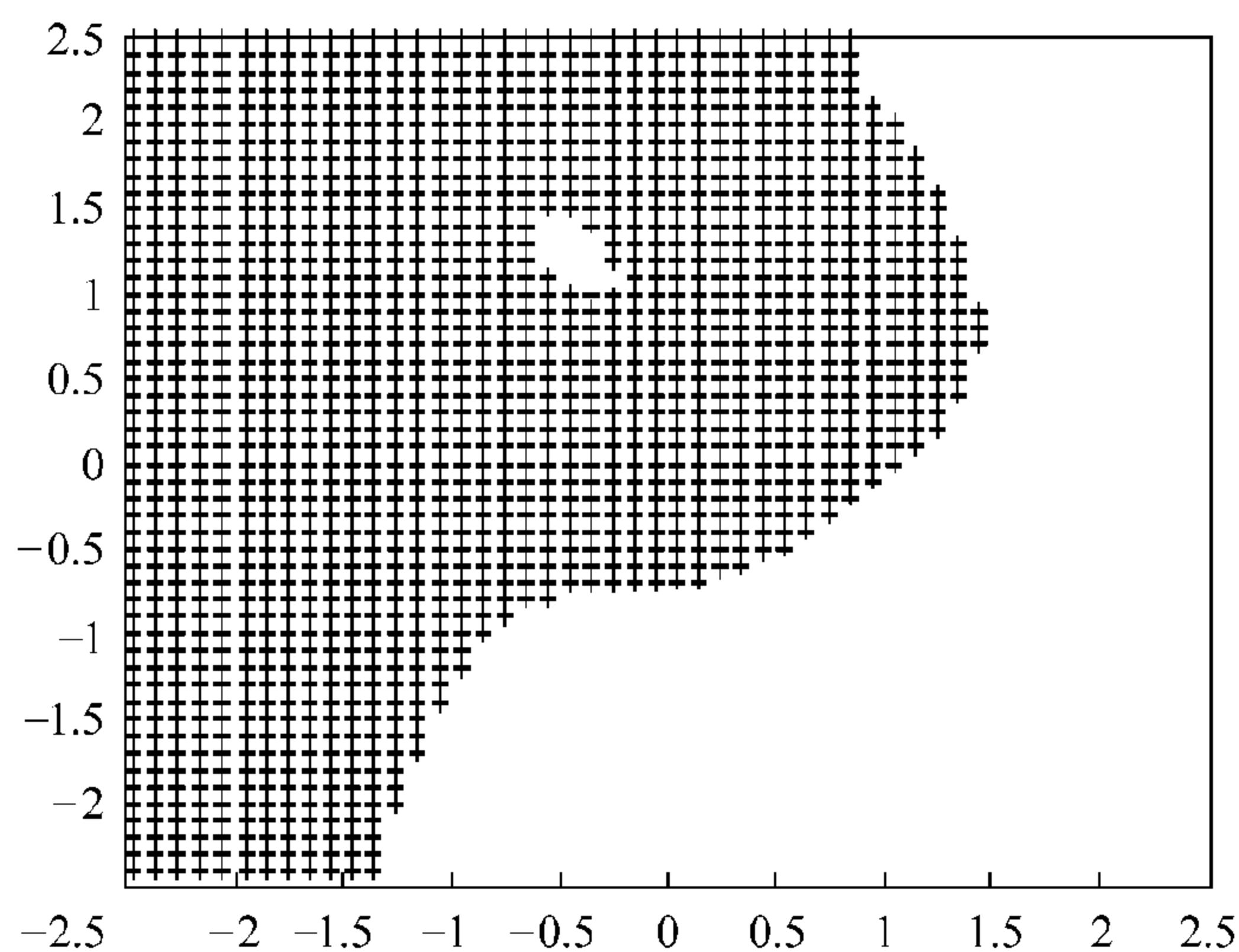
注：图中横轴表示 $x(k)$ ，纵轴表示 $x(k-1)$ 。

图 14.2 当高斯噪声的方差为 $\sigma_e^2 = 0.2$ ，均衡器的阶数 $n = 2$ ，延迟 $d = 0$ 时信道(14.28)的最优决策域

糊系统的输出由“应用阶段”中给出的方法确定。下面采用 14.1 节中的设计程序。在步骤 1 中，选定

$$N_1 = N_2 = 9, \mu_{A_i^l}(x_i) = \exp\left(-\left(\frac{x_i - \bar{x}_i^l}{0.3}\right)^2\right) \quad (i = 1, 2), \text{ 其中, } \bar{x}_i^l = -2 + 0.5(l-1)$$

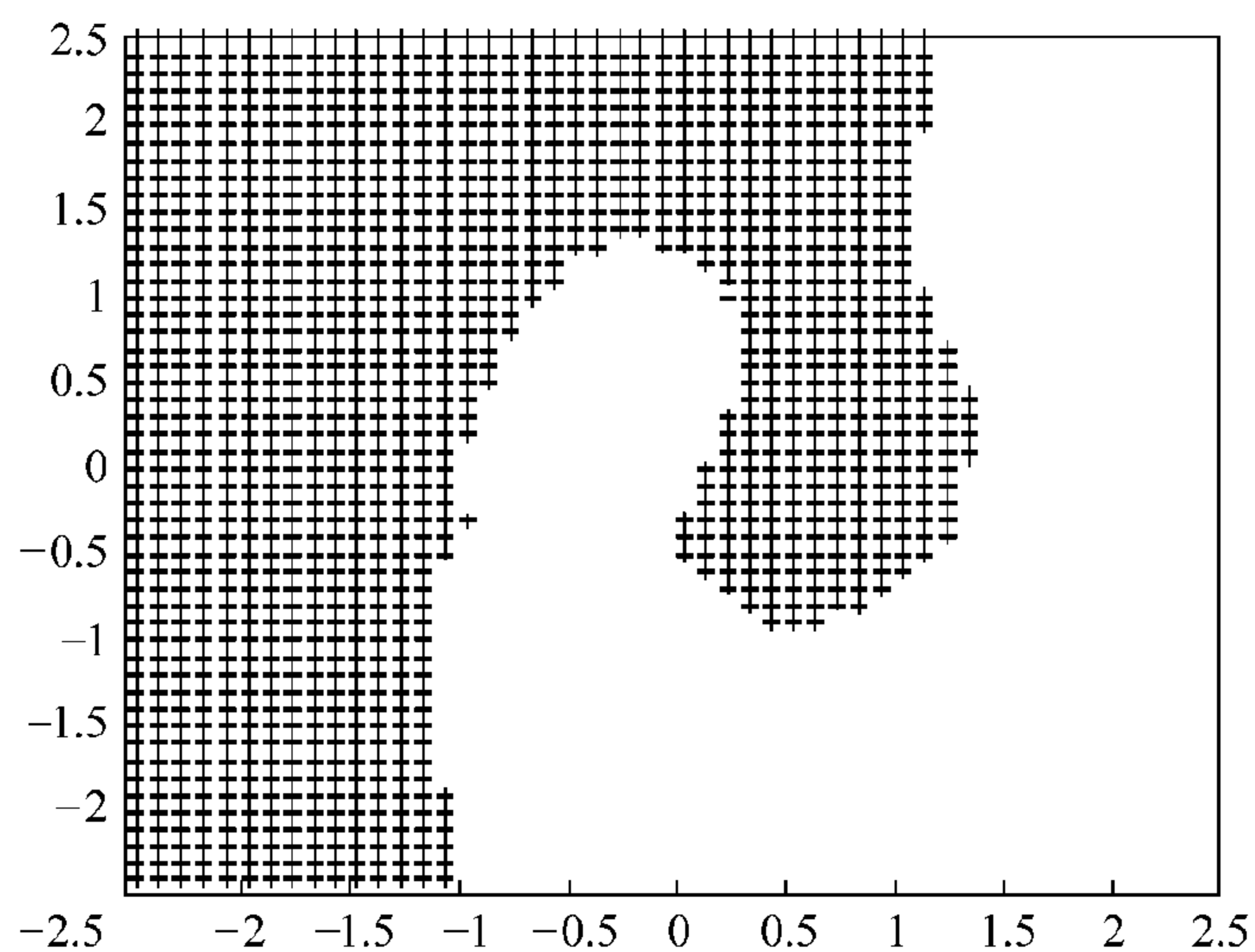
($l = 1, 2, \dots, 9$)。在步骤 3 中，选初始参数 $\theta(0)$ 为区间 $[-0.3, 0.3]$ 上的随机值。在步骤 4 中，选定 $\sigma = 0.1$ 。图 14.3 ~ 图 14.5 分别给出了在学习步数(步骤 4 中)终止于 $k = 30, 50, 100$ (即式(14.8) ~ 式(14.10)中的 $p = 30, 50, 100$)的情况下，所设计的模糊系统的决策域。



注：图中横轴表示 $x(k)$ ，纵轴表示 $x(k-1)$ 。

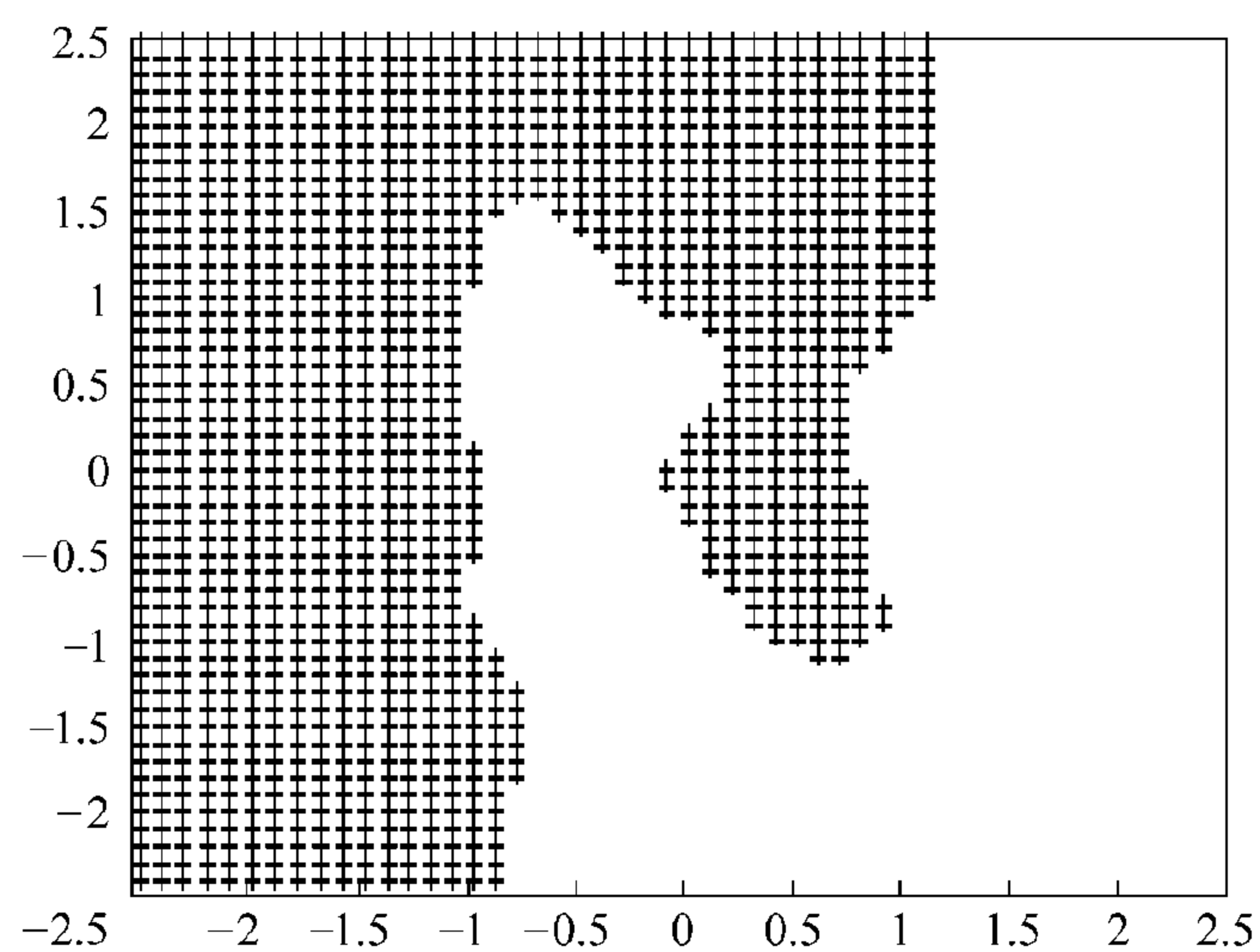
图 14.3 学习步数终止于 $k = 30$ 时，模糊系统均衡器的决策域

由图 14.3 ~ 图 14.5 可知,决策域随着学习步数的增多而趋向收敛于最优决策域。



注: 图中横轴表示 $x(k)$, 纵轴表示 $x(k-1)$ 。

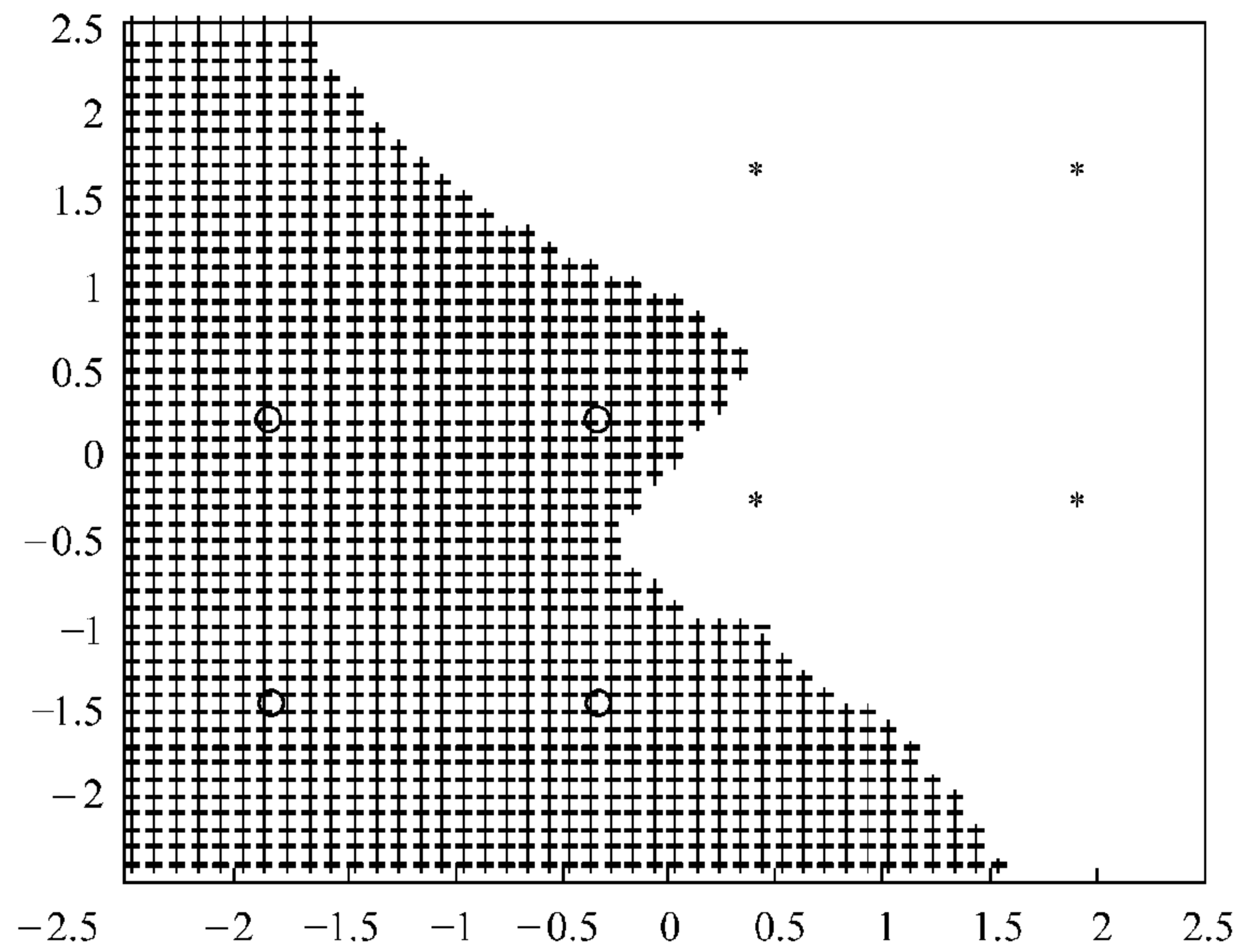
图 14.4 学习步数终止于 $k=50$ 时,
模糊系统均衡器的决策域



注: 图中横轴表示 $x(k)$, 纵轴表示 $x(k-1)$ 。

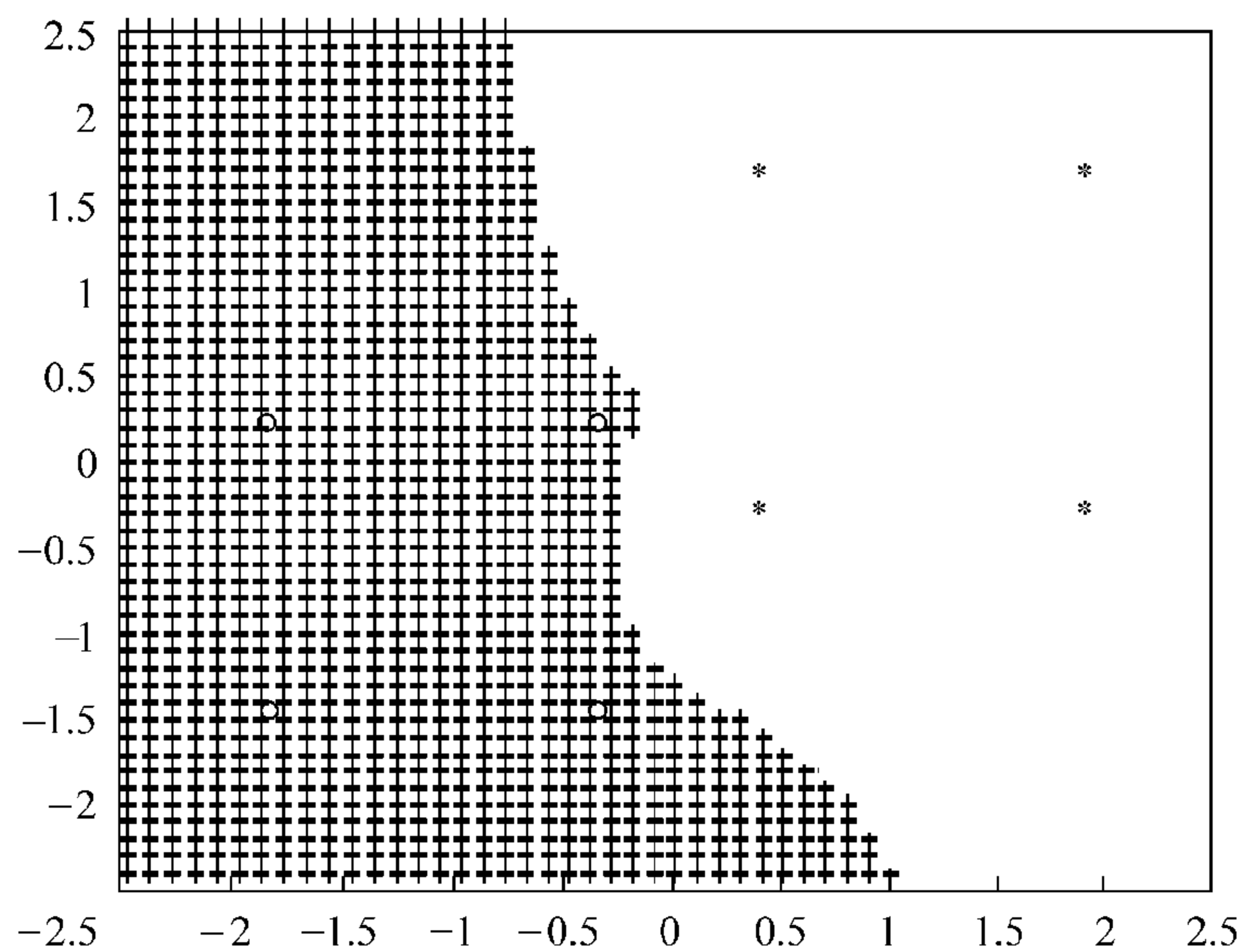
图 14.5 学习步数终止于 $k=100$ 时,
模糊系统均衡器的决策域

例 14.2 在本例中,考虑与例 14.1 相同的外部环境,只不过选定 $d=1$,而非 $d=0$ 。该例的最优决策域如图 14.6 所示。图 14.7 和图 14.8 分别给出了在步骤 4 中的学习步数终止于 $k=20$ 和 $k=50$ 情况下,模糊系统的决策域。再次看到,学习步数越多,决策域越收敛于最优决策域。



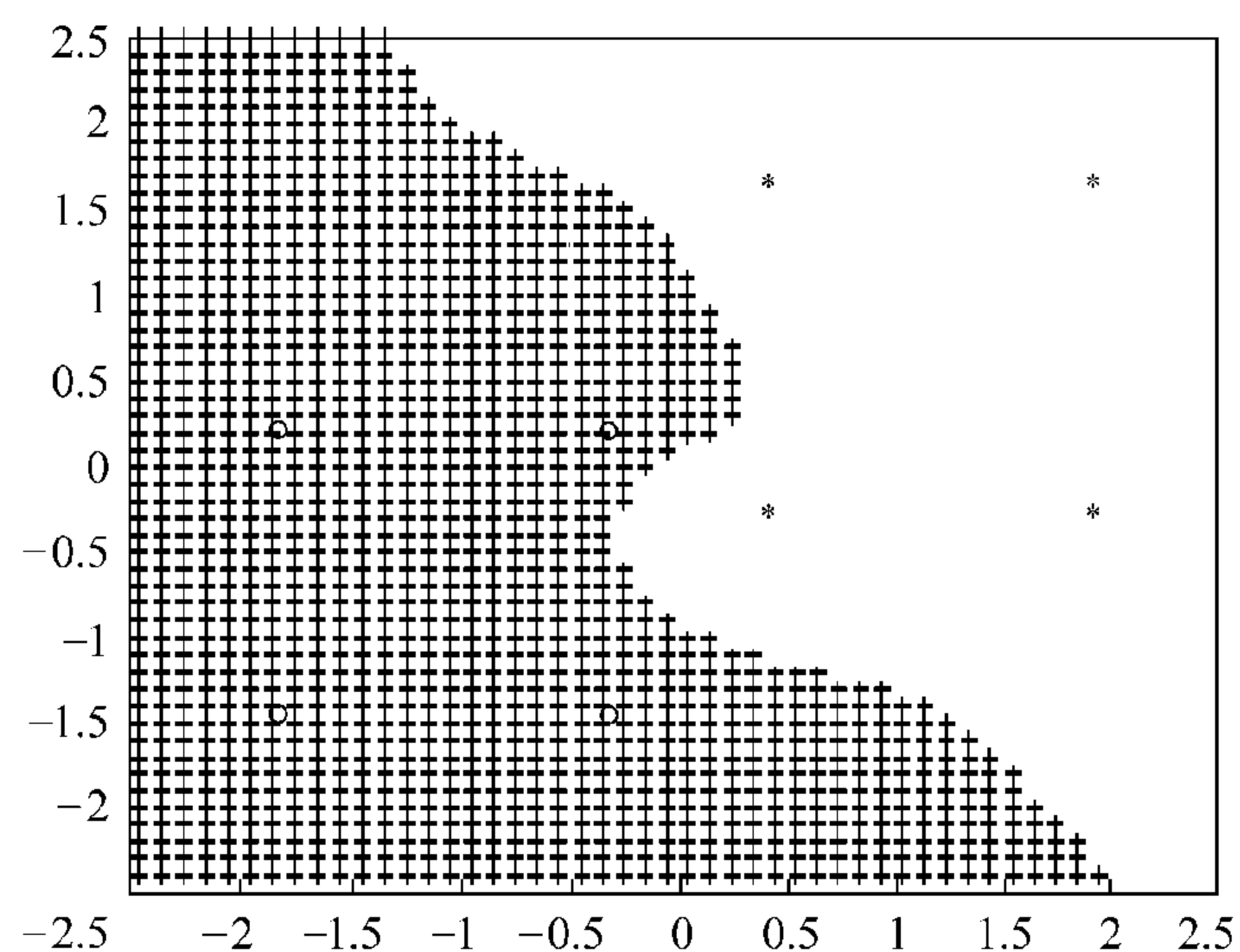
注：图中横轴表示 $x(k)$ ，纵轴表示 $x(k-1)$ 。

图 14.6 当高斯噪声的方差为 $\sigma_e^2 = 0.2$ ，
均衡器的阶数 $n = 2$ ，延迟 $d = 1$ 时
信道(14.28)的最优决策域



注：图中横轴表示 $x(k)$ ，纵轴表示 $x(k-1)$ 。

图 14.7 例 14.2 中的学习步数终止于 $k = 20$ 时，
模糊系统均衡器的决策域



注：图中横轴表示 $x(k)$ ，纵轴表示 $x(k-1)$ 。

图 14.8 例 14.2 中的学习步数终止于 $k=50$ 时，模糊系统均衡器的决策域

14.4 本章小结及导读

本章主要阐述了以下内容：

- 用递推最小二乘法设计模糊系统的参数；
- 推导递推最小二乘法的技术；
- 作为模式识别问题的信道均衡问题的公式，以及用递推最小二乘法设计的模糊系统在均衡问题和类似模式识别问题中的应用。

递推最小二乘算法在许多有关预测理论和自适应滤波器的标准教材中都被详细研究过，例如，Mendel[1994]，Cowan 和 Grant[1985]。本章的方法摘自 Wang[1994a]，Wang 和 Mendel[1993]，从中可以找到更多的仿真结果。至于采用的类似神经网络的方法，见 Chen, S., G. J. Gibson, C. F. N. Cowan 和 P. M. Grand[1990]。

14.5 习题与设计

习题 14.1 定义 XoR 函数为

x_1	x_2	$x_1 Xor x_2$	x_1	x_2	$x_1 Xor x_2$
-1	-1	1	1	-1	-1
-1	1	-1	1	1	1

(a) 设计一个形如式(14.3)的模糊系统 $f(x_1, x_2)$ 使得 $\text{sgn}[f(x_1, x_2)]$ 就是 xor 函数, 其中如果 $f \geq 0$ 则符号函数 $\text{sgn}(f) = 1$, 如果 $f < 0$ 则 $\text{sgn}(f) = -1$ 。

(b) 画出决策域 $\{x \in U \mid \text{sgn}[f(x)] \geq 0\}$, 其中 $U = [-2, 2] \times [-2, 2]$, $f(x)$ 是在(a)中所设计的模糊系统。

习题 14.2 讨论式(14.8) ~ 式(14.10)的物理含义, 解释为什么初始 $p(0) = \sigma I$ 应该较大。

习题 14.3 证明矩阵恒等式(14.15)。

习题 14.4 假定将式(14.1)中的标准 J_k 变为

$$J'_k = \sum_{j=1}^k \lambda^{k-j} [f(x_0^j) - y_0^j]^2 \quad (14.30)$$

式中, $\lambda \in (0, 1]$ 是遗忘因子, 并且仍采用形如式(14.5)的模糊系统。推导新标准 J'_k 下类似于式(14.8) ~ 式(14.10)的递推最小二乘法。

习题 14.5 用式(14.30)中的 J'_k 的目的是给历史数据赋予较小的权重, 从而低估历史数据。而另一种方法是仅考虑最近的 N 对数据, 即其标准为

$$J_k = \sum_{j=k-N+1}^k [f(x_0^j) - y_0^j]^2 \quad (14.31)$$

令 $f(x_0^j)$ 为形如式(14.5)的模糊系统。推导使得 J_k 最小的类似于式(14.8) ~ 式(14.10)的递推最小二乘法。

习题 14.6 确定以下点集的正确位置:

(a) 图 14.2 中的 $P_{2,0}(1)$ 和 $P_{2,0}(-1)$ 。

(b) 图 14.6 中的 $P_{2,1}(1)$ 和 $P_{2,1}(-1)$ 。

14.7(设计) 编写一个计算机程序来实现 14.1 节中的设计方法并将你的程序应用于第 13 章中的非线性系统辨识问题。

在第 12 章至第 14 章中,我们提出了三种设计模糊系统的方法。在这些方法中,我们并未提出一种系统地确定模糊系统规则数目的程序。更具体地讲,第 13 章中的梯度下降法是在训练前调整规则数目,而第 12 章中的查表法和第 13 章中的递推最小二乘法是通过调整 IF 部分的模糊集来依次设定规则数目的边界。在设计模糊系统的过程中,选择一个适当的规则数目是很重要的。因为规则太多会令模糊系统变得复杂,这种复杂的模糊系统也许对于该问题不是必需的;而规则太少又会令模糊系统的作用削弱,这种无效的模糊系统也许难以完成该问题的目标。

在本章中,我们把模糊系统中规则的数目作为设计参数并根据输入—输出数据对来确定规则的数目。基本思想是把输入—输出数据对分成组,一组采用一条模糊规则,即模糊规则的数目等于组的数量。首先,我们构造一个在某种意义上是最优的模糊系统,使其能以任意精度拟合所有输入—输出数据;即使输入—输出数据对的数量少,该最优模糊系统仍然是有效的。然后,采用最近邻原则聚类算法将输入—输出数据对分组,再将每组数据视为一个输入—输出数据对,用最优模糊系统进行拟合。

15.1 一种最优的模糊系统

设给定 N (N 较小,如 $N = 20$) 对输入—输出数据对 $(x_0^l; y_0^l)$ ($l = 1, 2, \dots, N$), 想要构造一个模糊系统 $f(x)$, 它能将 N 对数据拟合到任意精度。即,对于任意 $\varepsilon > 0$, 都有 $|f(x_0^l) - y_0^l| < \varepsilon$ ($l = 1, 2, \dots, N$) 成立。

最优的模糊系统可构造如下:

$$f(x) = \frac{\sum_{l=1}^N y_0^l \exp\left(-\frac{|x - x_0^l|^2}{\sigma^2}\right)}{\sum_{l=1}^N \exp\left(-\frac{|x - x_0^l|^2}{\sigma^2}\right)} \quad (15.1)$$

显然,模糊系统 (15.1) 是由如式 (7.1) 的 N 条规则构造而成的,其中 $\mu_{A_i}(x_i) = \exp\left(-\frac{|x_i - x_{0i}^l|}{\sigma^2}\right)$, B^l 的中心等于 y_0^l , 且模糊系统采用了乘积推理机、单值模糊器和中心平

均解模糊器。下面的定理说明了,通过适当选择参数 σ ,模糊系统(15.1)能够以任意精度拟合所有的 N 对输入—输出数据。

定理 15.1 对任意 $\varepsilon > 0$,存在 $\sigma^* > 0$,使得模糊系统(15.1)在 $\sigma = \sigma^*$ 时,对 $l = 1, 2, \dots, N$,有

$$|f(x_0^l) - y_0^l| < \varepsilon \quad (15.2)$$

证明 把 x_0^l 和 y_0^l 看做是引理 13.1 中的 x_0^{k+1} 和 y_0^{k+1} ,并使用与引理 13.1 相同的证明方法,即可证明该定理成立。

σ 是一个平滑参数: σ 越小,拟合误差 $|f(x_0^l) - y_0^l|$ 越小,但 $f(x)$ 越不光滑。我们知道,如果 $f(x)$ 不光滑,则它可能对不在学习集中的数据点很难进行较好的泛化。为此,应该适当选择 σ 以保持拟合和泛化之间的平衡。因为 σ 是个一维参数,所以通常对某一特定问题确定一个适当的 σ 是不难的。有时通过几次试错,就可能会确定一个很好的 σ 。一般性的规律是,较大的 σ 能去除带噪声的数据,而较小的 σ 能使 $f(x)$ 达到准确逼近样本数据所需的非线性程度。

函数 $f(x)$ 是一个广义的非线性回归模型,能在可观测数据点 $(x_0^l; y_0^l)$ 之间进行平滑插值,当 σ 非常小时,这一插值过程会更好。

例 15.1 本例将考虑参数 σ 对平滑性和最优模糊系统的拟合精度的影响。设想一个简单的一维输入的例子。假定有五对输入—输出数据: $(-2, 1), (-1, 0), (0, 2), (1, 2)$ 和 $(2, 1)$,最优的模糊系统就是如式(15.1)的模糊系统,当 $l = 1, 2, \dots, 5$ 时, $(x_0^l; y_0^l)$ 分别等于 $(-2, 1), (-1, 0), (0, 2), (1, 2)$ 和 $(2, 1)$ 。

图 15.1 ~ 图 15.3 分别给出了当 $\sigma = 0.1, 0.3$ 和 0.5 时 $f(x)$ 的图形。这些图也印证了较小的 σ 能给出较小的拟合误差,但是函数不很光滑的结论。

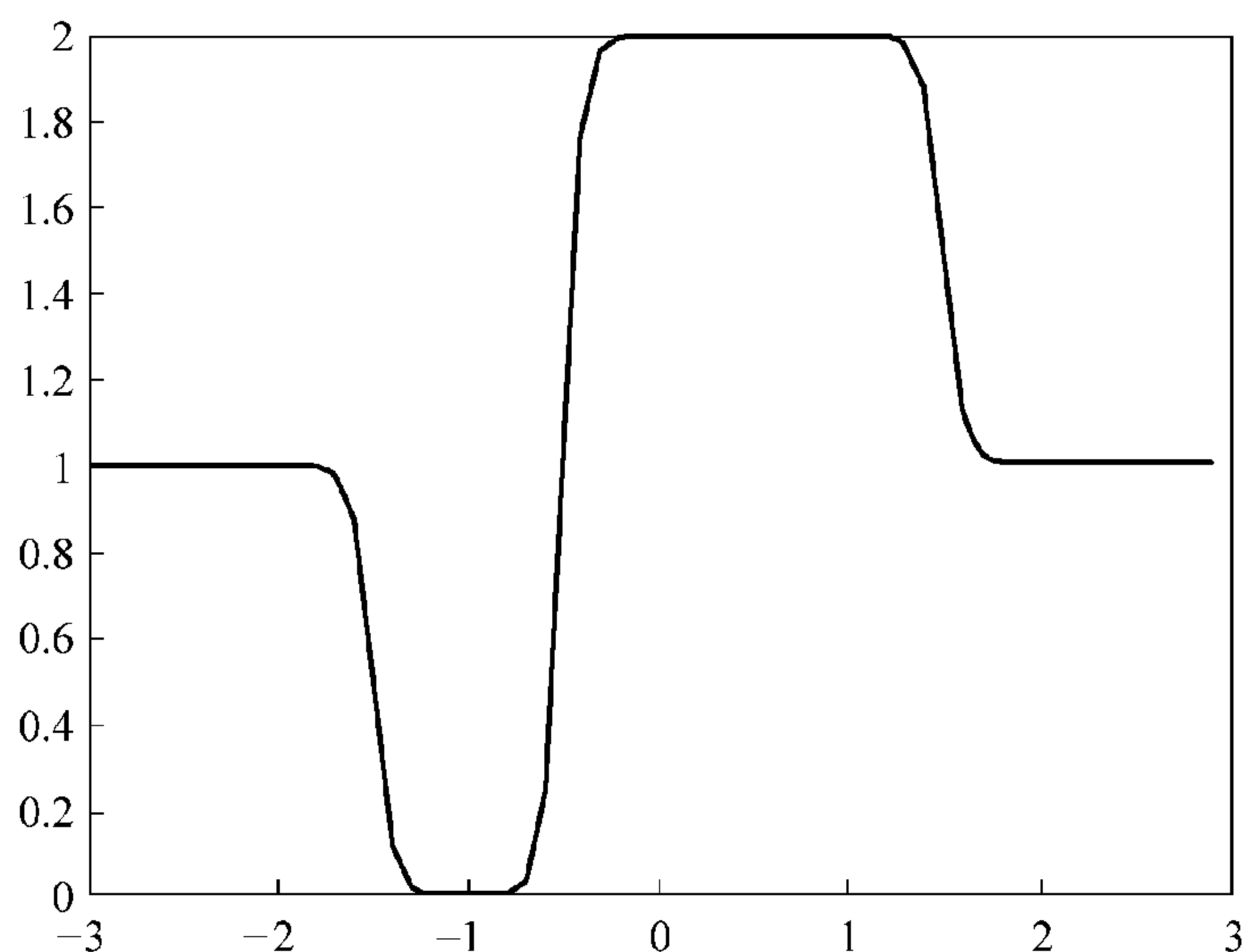
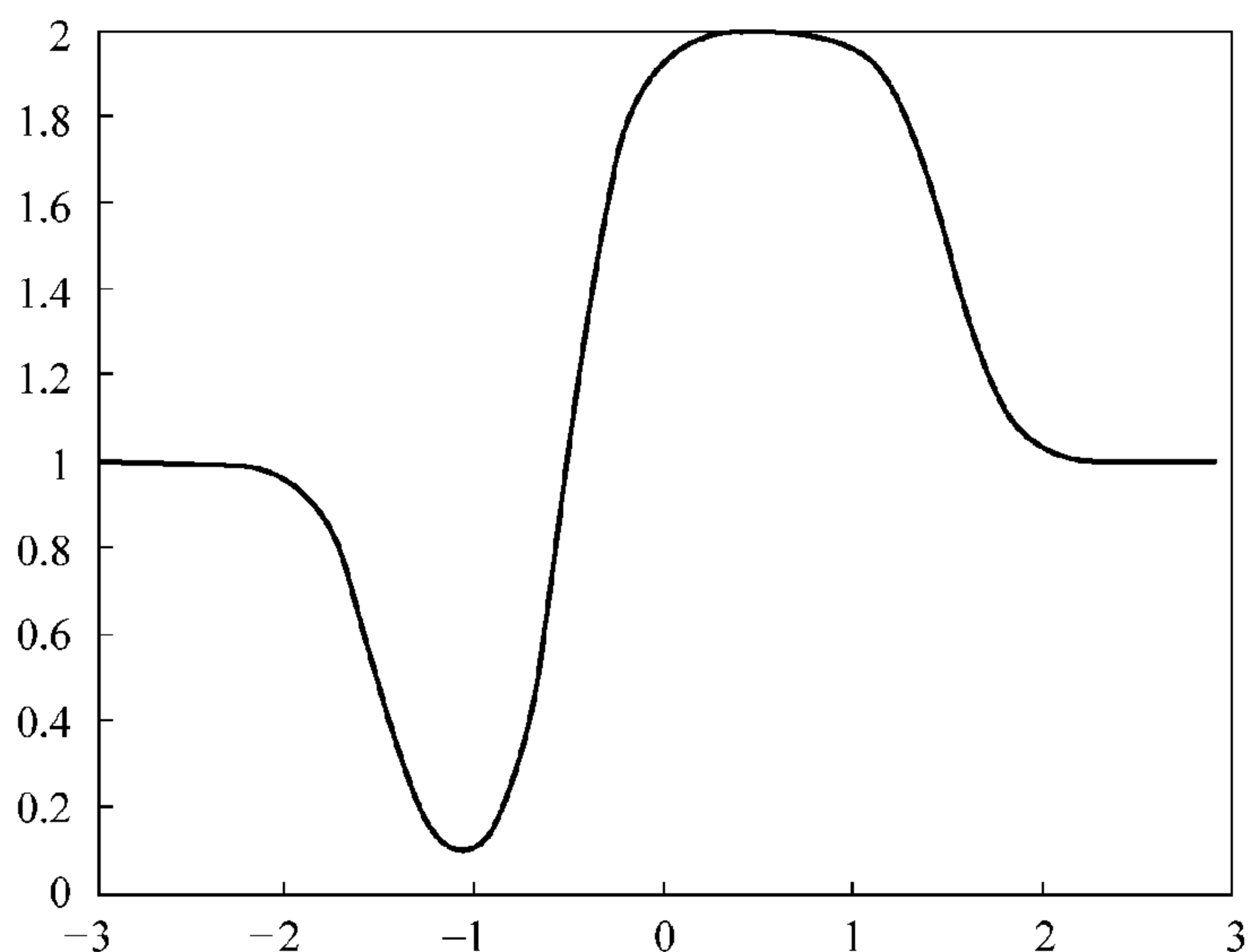
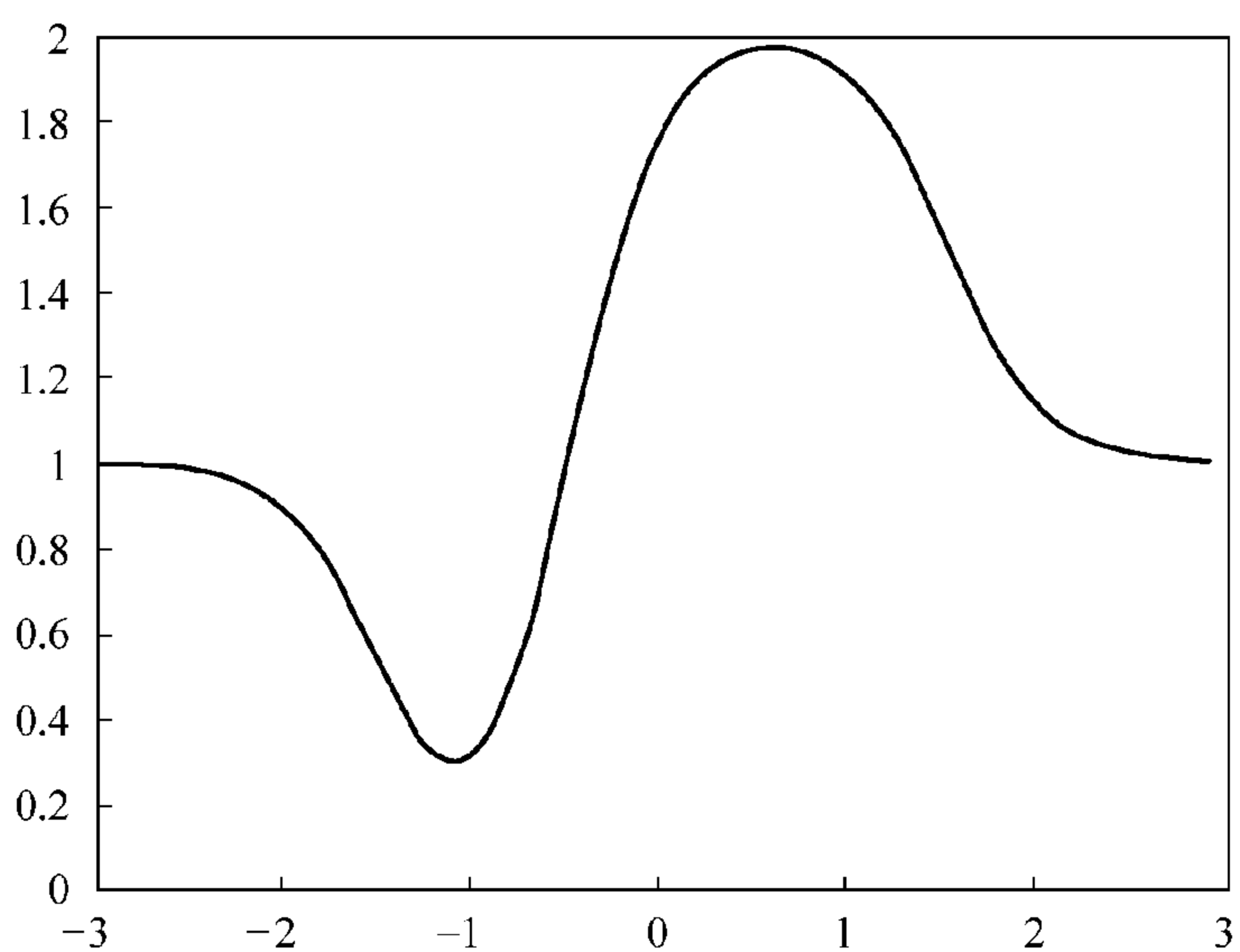


图 15.1 例 15.1 中 $\sigma = 0.1$ 时的最优模糊系统

图 15.2 例 15.1 中 $\sigma = 0.3$ 时的最优模糊系统图 15.3 例 15.1 中 $\sigma = 0.5$ 时的最优模糊系统

15.2 聚类法设计模糊系统

最优的模糊系统(15.1)是将一条规则应用于一个输入—输出数据对,但如果输入—输出数据对数目很大的话,这种系统不再是一个实际系统了。对于大样本问题,需要用不同的聚类技术将输入—输出数据对分组,使一组数据仅对应一条规则。

从一般意义来讲,聚类意味着把一个数据集合分割成不相交的子集或组,一组中的数据应具有某些能将它们同其他组数据区分开来的性质。对于这里的问题,应首先把输入—输

出数据对根据输入点的分布进行分组,然后每组仅用一条规则来描述。图 15.4 就阐述了这样一个例子,六对输入—输出数据被分成两组,然后产生两条用于构造模糊系统的规则。下面将会给出详细的算法。

最近邻聚类法是一种最简单的聚类算法。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。详细的算法如下:

用最近邻聚类法设计模糊系统。

步骤 1 从第一个输入—输出数据对 $(x_0^1; y_0^1)$ 开始,把 x_0^1 设为一个聚类中心 x_c^1 ,并令 $A^1(1) = y_0^1, B^1(1) = 1$,设定半径 r 。

步骤 2 假定考虑第 k 对输入—输出数据 $(x_0^k; y_0^k)$ ($k = 2, 3, \dots$) 时,已经存在聚类中心分别为 $x_c^1, x_c^2, \dots, x_c^M$ 的 M 个聚类。分别计算 x_0^k 到这 M 个聚类中心的距离 $|x_0^k - x_c^l|$ ($l = 1, 2, \dots, M$)。设这些距离中最小的距离为 $|x_0^k - x_c^{l_k}|$,即 $x_c^{l_k}$ 为 x_0^k 的最近邻原则聚类,则:

(a) 如果 $|x_0^k - x_c^{l_k}| > r$,则把 x_0^k 作为一个新的聚类中心 $x_c^{M+1} = x_0^k$,令 $A^{M+1}(k) = y_0^k, B^{M+1}(k) = 1$,并令 $A^l(k) = A^l(k-1), B^l(k) = B^l(k-1)$ ($l = 1, 2, \dots, M$)。

(b) 如果 $|x_0^k - x_c^{l_k}| \leq r$,则做如下计算:

$$A^{l_k}(k) = A^{l_k}(k-1) + y_0^k \quad (15.3)$$

$$B^{l_k}(k) = B^{l_k}(k-1) + 1 \quad (15.4)$$

当 $l \neq l_k, l = 1, 2, \dots, M$ 时,令

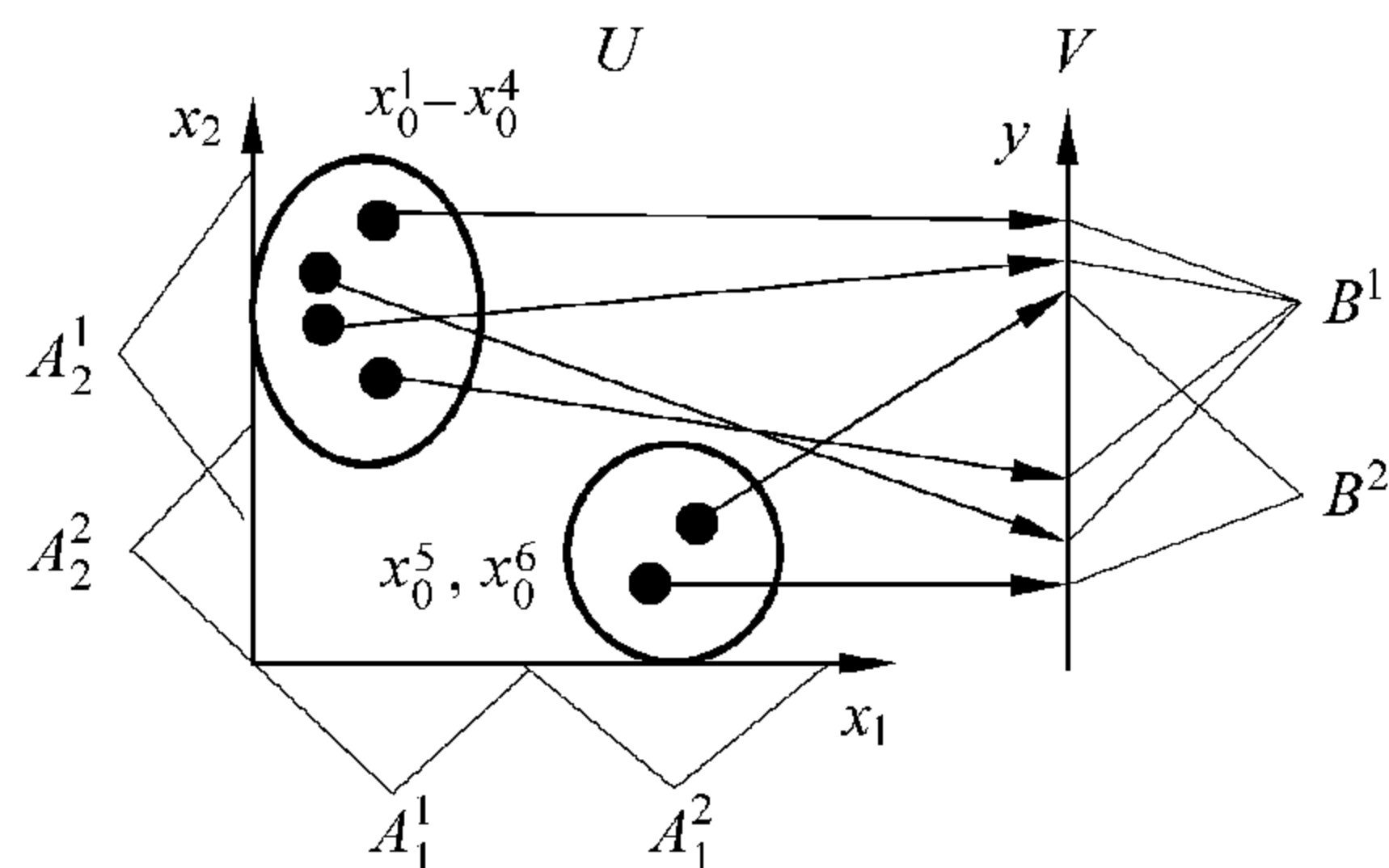
$$A^l(k) = A^l(k-1) \quad (15.5)$$

$$B^l(k) = B^l(k-1) \quad (15.6)$$

步骤 3 如果 x_0^k 并未建立一个新的聚类,则根据 k 对输入—输出数据 $(x_0^j; y_0^j)$ ($j = 1, 2, \dots, k$) 设计如下模糊系统:

$$f_k(x) = \frac{\sum_{l=1}^M A^l(k) \exp\left(-\frac{|x - x_c^l|^2}{\sigma}\right)}{\sum_{l=1}^M B^l(k) \exp\left(-\frac{|x - x_c^l|^2}{\sigma}\right)} \quad (15.7)$$

如果 x_0^k 建立了一个新的聚类,则所设计的模糊系统为



如果 x_1 为 A_1^1 且 x_2 为 A_2^1 , 则 y 为 B^1

如果 x_1 为 A_1^2 且 x_2 为 A_2^2 , 则 y 为 B^2

注: 图中的六对输入—输出数据 $(x_0^1; y_0^1), \dots, (x_0^6; y_0^6)$ 被分为两组, 并由此产生上面的两条规则。

图 15.4 根据输入—输出数据对构造模糊 IF-THEN 规则的一个例子

$$f_k(x) = \frac{\sum_{l=1}^{M+1} A^l(k) \exp\left(-\frac{|x - x_c^l|^2}{\sigma}\right)}{\sum_{l=1}^{M+1} B^l(k) \exp\left(-\frac{|x - x_c^l|^2}{\sigma}\right)} \quad (15.8)$$

步骤4 令 $k = k + 1$, 返回步骤2。

从式(15.3) ~ 式(15.6)可以看出, 变量 $B^l(k)$ 等于第 l 组中已使用了 k 对输入—输出数据后的输入—输出数据对的数目, $A^l(k)$ 等于第 l 组中输入—输出数据对的输出值的总和。所以, 如果每个输入—输出数据对都建立了一个聚类中心, 那么所设计的模糊系统(15.8)就变成了最优的模糊系统(15.1)。因为最优的模糊系统(15.1)可以看做是用一条规则来对应一个输入—输出数据对, 所以模糊系统(15.7)或(15.8)就可以看做是用一条规则来对应一组输入—输出数据对。由于每个输入—输出数据对都有可能产生一个新的聚类, 因此, 所设计的模糊系统中规则的数目在设计过程中也是不断变化的。组(或规则)的数目取决于输入—输出数据对中输入点的分布及半径 r 。

半径 r 决定了模糊系统的复杂性。 r 越小, 所得到的组的数目就越多, 从而使得模糊系统越复杂。当 r 较大时, 所设计的模糊系统会比较简单但缺乏力度。实际中, 可以通过试错法找到一个适当的半径 r 。

例 15.2 现在针对例 15.1 中的五个输入—输出数据对, 采用本节中的设计程序来设计一个模糊系统。如果 $r < 1$, 则五个输入—输出数据对中的每一对输入—输出数据都能成为一个聚类中心, 这样所设计的模糊系统 $f_5(x)$ 就和例 15.1 中的模糊系统一样了。现在, 设计一个 $r = 1.5$ 的模糊系统。

首先, 建立第一组的聚类中心 $x_c^1 = -2$, 并设 $A^1(1) = y_0^1 = 1, B^1(1) = 1$ 。

然后, 当 $k = 2$ 时, $|x_0^2 - x_c^1| = |x_0^2 - x_c^1| = |-1 - (-2)| = 1 < r = 1.5$, 则 $A^1(2) = A^1(1) + y_0^2 = 1 + 0 = 1, B^1(2) = B^1(1) + 1 = 2$ 。当 $k = 3$ 时, 有 $|x_0^3 - x_c^1| = |x_0^3 - x_c^1| = |0 - (-2)| = 2 > r$, 则应建立一个新的聚类中心 $x_c^2 = x^3 = 0$, 此时, $A^2(3) = y_0^3 = 2, B^2(3) = 1$ 。 A^1, B^1 仍保持不变, 即 $A^1(3) = A^1(2) = 1, B^1(3) = B^1(2) = 2$; 当 $k = 4$ 时, 有 $|x_0^4 - x_c^1| = |x_0^4 - x_c^2| = |1 - 0| = 1 < r$, 则 $A^2(4) = A^2(3) + y_0^4 = 2 + 2 = 4, B^2(4) = B^2(3) + 1 = 2, A^1(4) = A^1(3) = 1, B^1(4) = B^1(3) = 2$ 。最后, 当 $k = 5$ 时, 有 $|x_0^5 - x_c^1| = |x_0^5 - x_c^2| = |2 - 0| = 2 > r$, 则应建立一个新的聚类中心 $x_c^3 = x_0^5 = 2$, 此时, $A^3(5) = y_0^5 = 1, B^3(5) = 1$, 其他变量仍保持不变, 即 $A^1(5) = A^1(4) = 1, B^1(5) = B^1(4) = 2, A^2(5) = A^2(4) = 4, B^2(5) = B^2(4) = 2$ 。最终的模糊系统为

$$f_5(x) = \frac{\sum_{l=1}^3 A^l(5) \exp\left(-\frac{(x - x_c^l)^2}{\sigma}\right)}{\sum_{l=1}^3 B^l(5) \exp\left(-\frac{(x - x_c^l)^2}{\sigma}\right)}$$

$$= \frac{\exp\left(-\frac{(x+2)^2}{\sigma}\right) + 4\exp\left(-\frac{x^2}{\sigma}\right) + \exp\left(-\frac{(x-2)^2}{\sigma}\right)}{2\exp\left(-\frac{(x+2)^2}{\sigma}\right) + 2\exp\left(-\frac{x^2}{\sigma}\right) + \exp\left(-\frac{(x-2)^2}{\sigma}\right)} \quad (15.9)$$

当 $\sigma = 0.3$ 时, 上面的模糊系统的图形见图 15.5。

比较图 15.5 和图 15.2 可以看出, 正如预期的那样, 模糊系统(15.9)对于五个输入—输出数据对的拟合误差是大于最优模糊系统的。

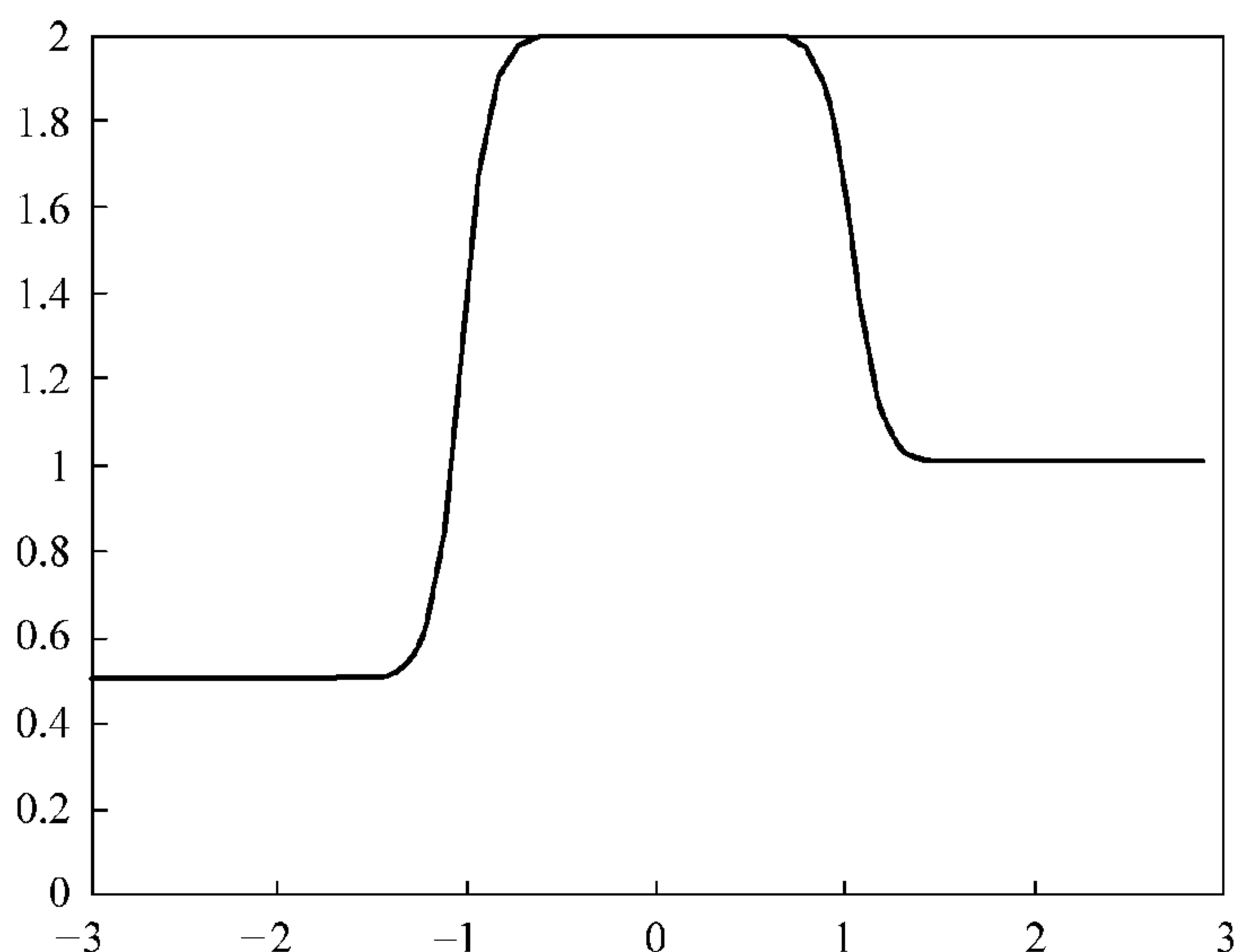


图 15.5 例 15.2 中设计的模糊系统 $f_5(x)$ (15.9)($\sigma = 0.3$)

因为式(15.7)和式(15.8)中的系数 $A^l(k)$ 和 $B^l(k)$ 是用递推方程(15.3)~(15.6)确定的, 所以可以很容易地在式(15.3)~式(15.6)中加入一个遗忘因子, 以使模糊系统在为具有时变特性的系统建模时, 达到满意的效果。在此情况下, 可把式(15.3)和式(15.4)用下式替换:

$$A^{lk}(k) = \frac{\tau - 1}{\tau} A^{lk}(k-1) + \frac{1}{\tau} y_0^k \quad (15.10)$$

$$B^{lk}(k) = \frac{\tau - 1}{\tau} B^{lk}(k-1) + \frac{1}{\tau} \quad (15.11)$$

把式(15.5)和式(15.6)用下式替换:

$$A^l(k) = \frac{\tau - 1}{\tau} A^l(k-1) \quad (15.12)$$

$$B^l(k) = \frac{\tau - 1}{\tau} B^l(k-1) \quad (15.13)$$

式中, τ 可视为一个指数衰变函数的时间常数。通常在实践中, $B^l(k)$ 应该选择较低的阈值。这样如果某个聚类在较长的一段时间内一直没有得到更新(原因是比阈值更小), 那么这个

聚类将被取消。

15.3 在非线性自适应控制系统中的应用

在本节中,将式(15.7)或式(15.8)所设计的模糊系统作为离散时间非线性动态系统的自适应模糊控制器的基本组成单元。这里考虑两个例子,其方法可以推广到其他情况。

例 15.3 考虑由差分方程所描述的离散时间非线性系统:

$$y(k+1) = g[y(k), y(k-1)] + u(k) \quad (15.14)$$

其中,非线性函数

$$g[y(k), y(k-1)] = \frac{y(k)y(k-1)[y(k) + 2.5]}{1 + y^2(k) + y^2(k-1)} \quad (15.15)$$

假设该非线性函数未知,控制的目标是要根据模糊系统(15.7)或(15.8)设计一个控制器 $u(k)$,使得闭环系统的输出 $y(k)$ 能跟踪下面的参考模型的输出 $y_m(k)$:

$$y_m(k+1) = 0.6y_m(k) + 0.2y_m(k-1) + r(k) \quad (15.16)$$

式中, $r(k) = \sin(2\pi k/25)$ 。即,需保证 $e(k) = y(k) - y_m(k)$ 在 k 趋于无穷大时收敛于 0。

如果函数 $g[y(k), y(k-1)]$ 已知,则可用下式来构造控制器:

$$u(k) = -g[y(k), y(k-1)] + 0.6y(k) + 0.2y(k-1) + r(k) \quad (15.17)$$

将上式带入式(15.14),可得

$$y(k+1) = 0.6y(k) + 0.2y(k-1) + r(k) \quad (15.18)$$

合并式(15.16)和式(15.18),可得

$$e(k+1) = 0.6e(k) + 0.2e(k-1) \quad (15.19)$$

由此可得 $\lim_{k \rightarrow \infty} e(k) = 0$ 。然而,因为 $g[y(k), y(k-1)]$ 未知,所以控制器(15.17)还不能实现。为了解决这一问题,把式(15.17)中的 $g[y(k), y(k-1)]$ 用模糊系统(15.7)或(15.8)来取代,即,用下面的控制器:

$$u(k) = -f_k[y(k), y(k-1)] + 0.6y(k) + 0.2y(k-1) + r(k) \quad (15.20)$$

其中, $f_k[y(k), y(k-1)]$ 是形如式(15.7)或式(15.8)的模糊系统, $x = (y(k), y(k-1))^T$ 。这样,就必须用下面的非线性差分方程来描述闭环系统的动态特性:

$$y(k+1) = g[y(k), y(k-1)] - f_k[y(k), y(k-1)] + 0.6y(k) + 0.2y(k-1) + r(k) \quad (15.21)$$

整个控制系统的框图如图 15.6 所示。由图 15.6 可以看出,控制器由辨识器和控制器两部分组成。辨识器利用模糊系统 f_k 来逼近未知非线性函数 g ,然后再把 f_k 复制到控制器中。

对以下两类情况进行仿真:

- 第一种情况:先不连接图 15.6 中的控制器,只用辨识器来辨识未知的系统。在辨识

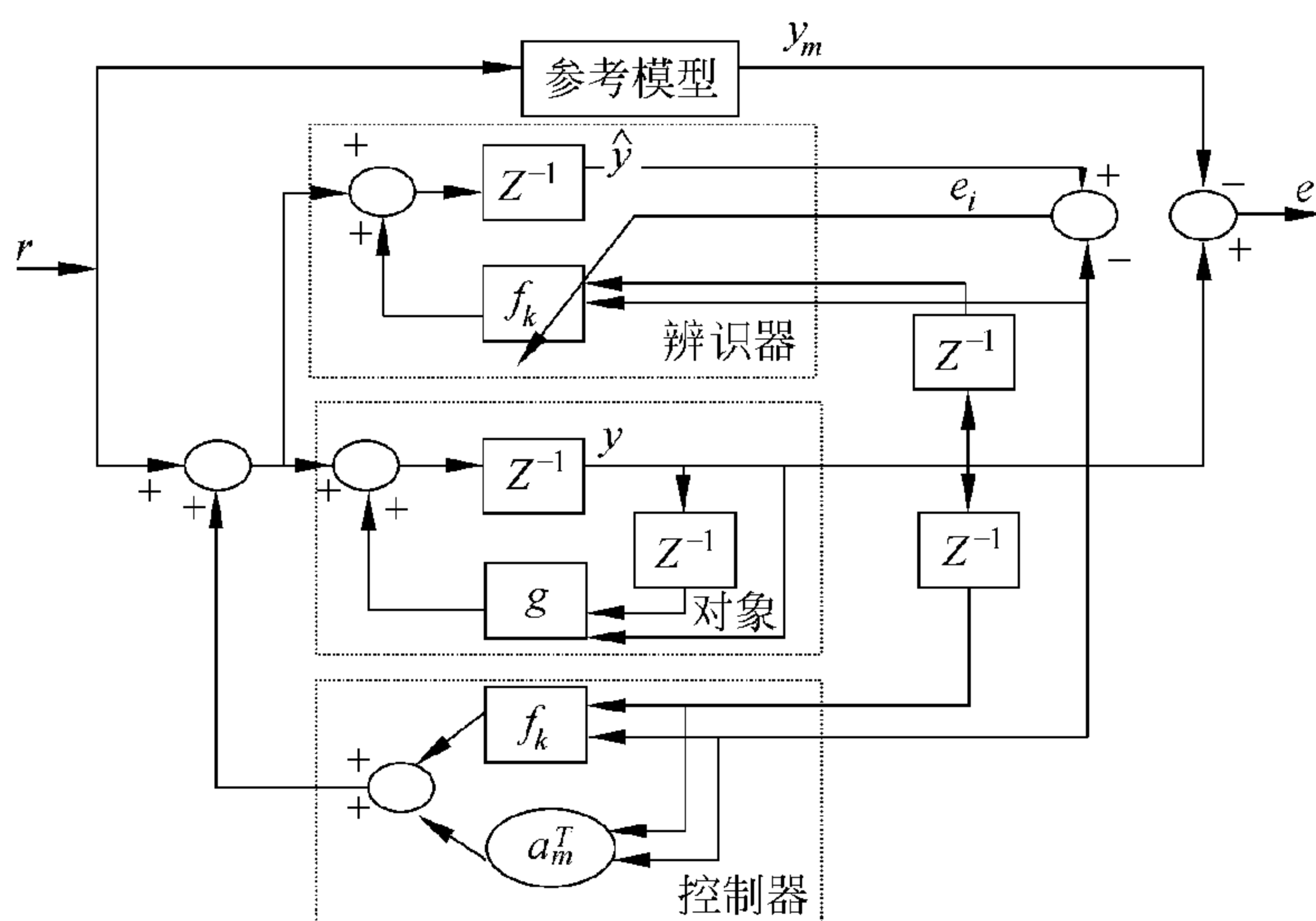


图 15.6 例 15.3 中的自适应控制系统的整体框图

阶段, 选定输入 $u(k)$ 为独立同分布(i. i. d)的随机信号, 随机信号是均匀分布在区间 $[-3, 3]$ 上的。当辨识过程结束之后, 用式 (15.20) 来产生控制器的输入。即, 图 15.6 中的控制器中的 f_k 是由应用辨识器得到的 f_k 而复制来的。图 15.7 和图 15.8 分别给出了当辨识过程分别终止于 $k = 100$ 和 $k = 500$ 时, 采用这种控制器的闭环系统的输出 $y(k)$ 和参考模型的输出 $y_m(k)$ 。在这些仿真中都选定 $\sigma = 0.3, r = 0.3$ 。由仿真结果可以看出: (i) 只有经过 100 步的训练, 辨识器才能产生出精确的模型, 才能得到良好的跟踪效果; (ii) 学习步数越多, 控制性能越好。

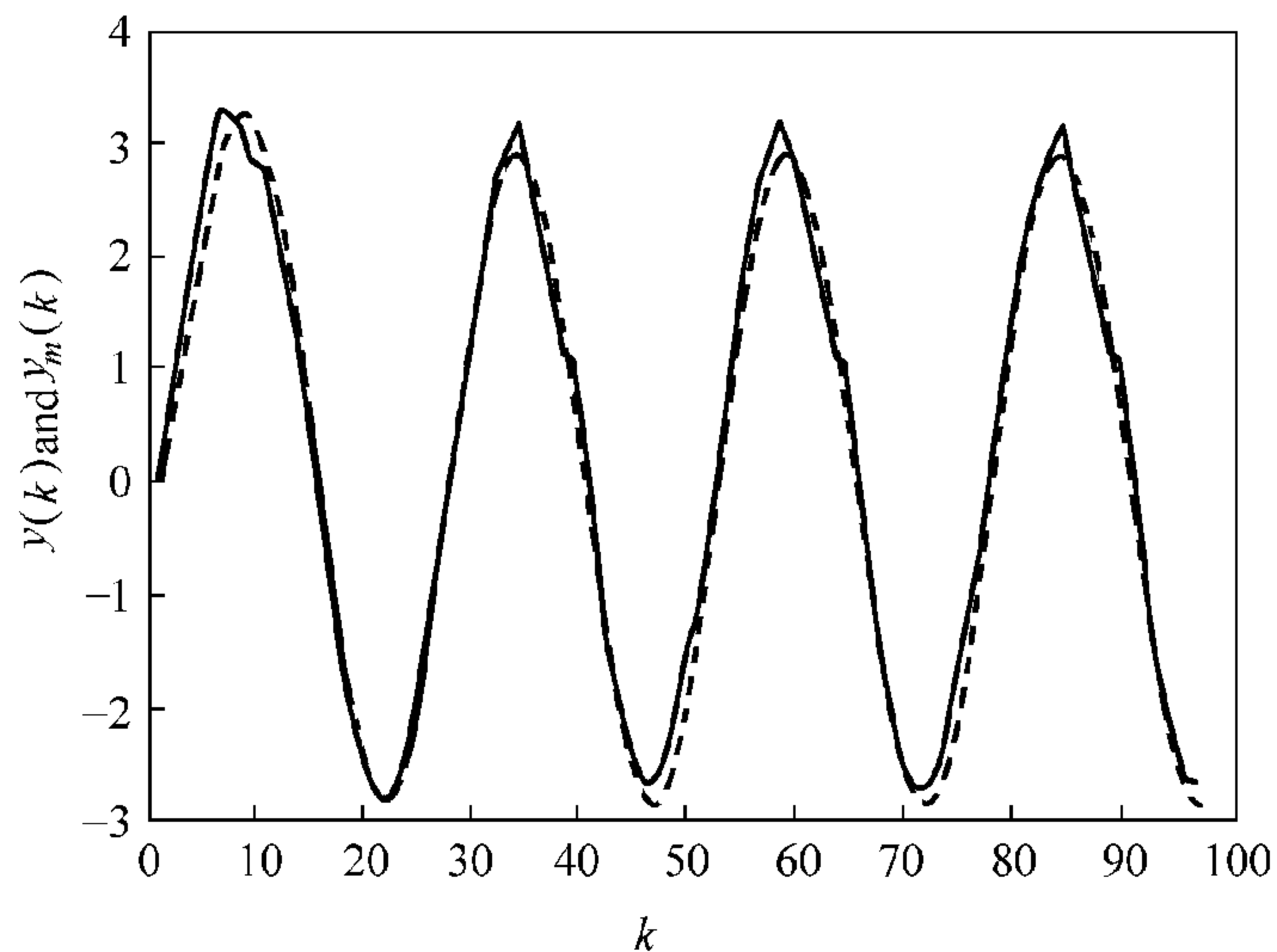


图 15.7 例 15.3 中的第一种情况, 当辨识程序终止于 $k = 100$ 时闭环系统的输出 $y(k)$ (实线) 和参考轨迹 $y_m(k)$ (虚线)

- 第二种情况: 控制器和辨识器同时从 $k = 0$ 开始运作(如图 15.6 所示)。仍选定 $\sigma =$

0.3, $r=0.3$ 。图 15.9 给出了这一仿真的输出结果 $y(k)$ 和 $y_m(k)$ 。

例 15.4 在本例中,考虑下面的系统

$$y(k+1) = \frac{5y(k)y(k-1)}{1+y^2(k)+y^2(k-1)+y^2(k-2)} + u(k) + 0.8u(k-1) \quad (15.22)$$

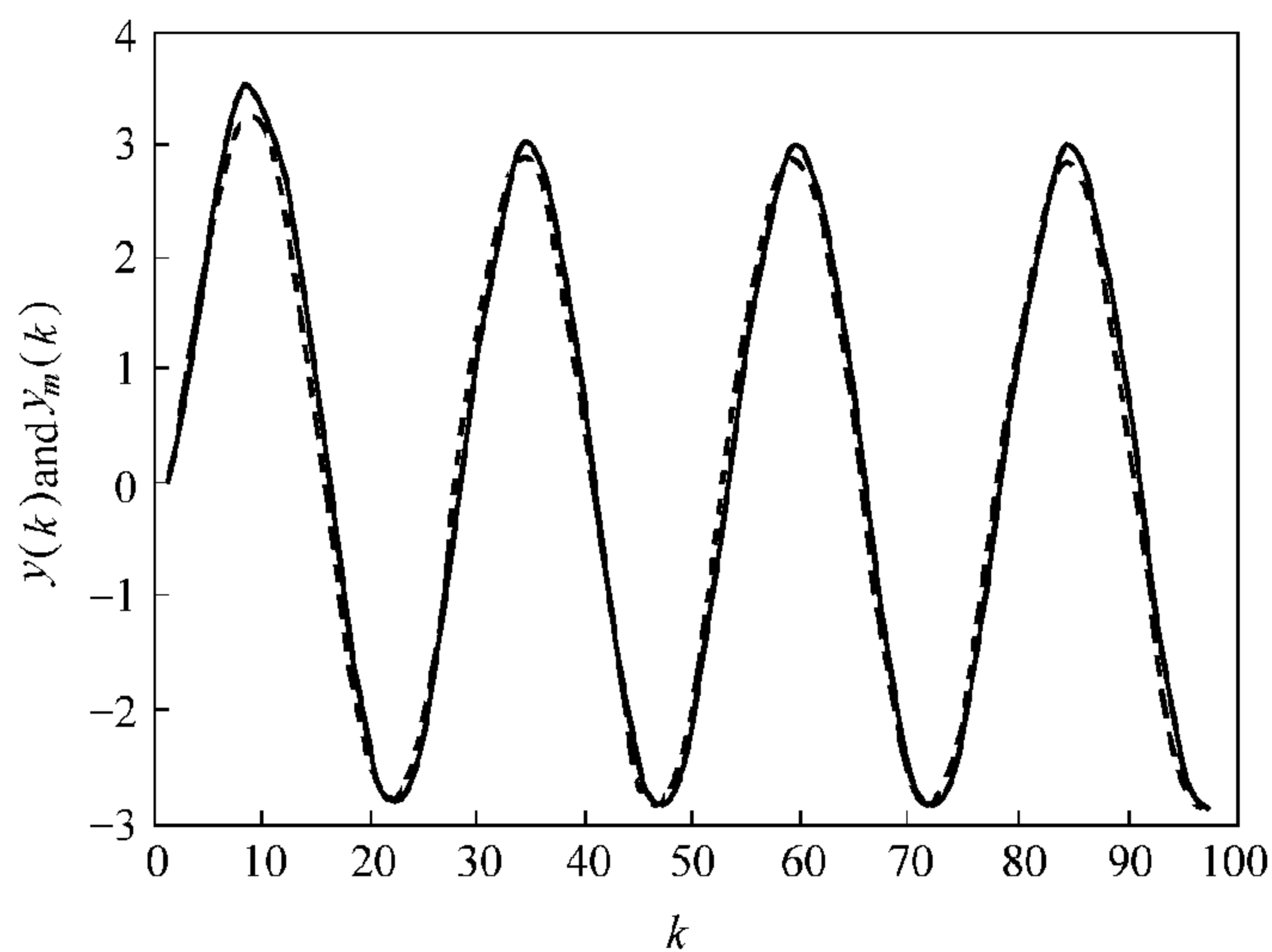


图 15.8 例 15.3 中的第一种情况,当辨识程序终止于 $k=500$ 时闭环系统的输出 $y(k)$ (实线)和参考轨迹 $y_m(k)$ (虚线)

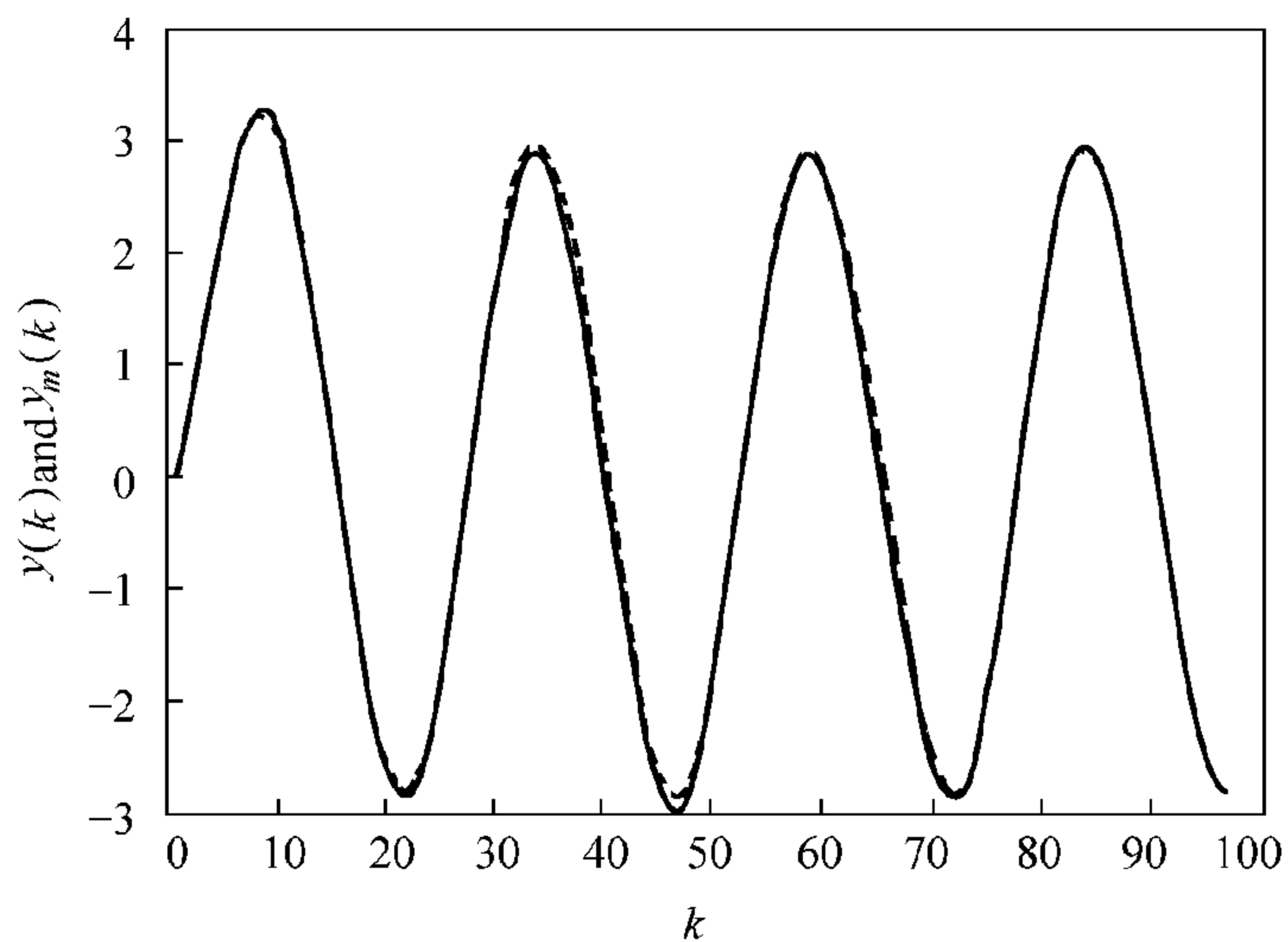


图 15.9 例 15.3 中的第二种情况下闭环系统的输出 $y(k)$ (实线)和参考轨迹 $y_m(k)$ (虚线)

这里,假设非线性函数是未知的。目标是设计一个控制器 $u(k)$,使得 $y(k)$ 能跟踪下面的参考模型:

$$y_m(k+1) = 0.32y_m(k) + 0.64y_m(k-1) - 0.5y_m(k-2) + \sin(2\pi k/25) \quad (15.23)$$

采用与例 15.3 相同的思想,选定:

$$u(k) = -f_k[y(k), y(k-1), y(k-2)] - 0.8u(k-1) + 0.32y(k) + 0.64y(k-1) - 0.5y(k-2) + \sin(2\pi k/25) \quad (15.24)$$

其中, $f_k[y(k), y(k-1), y(k-2)]$ 是形如式(15.7)或式(15.8)的模糊系统, 控制方案的整体框图与图 15.6 相同。图 15.10 给出了当辨识器和控制器都从 $k=0$ 开始工作时的仿真结果 $y(k)$ 和 $y_m(k)$ 。在此仿真中, 同样仍选择 $\sigma=0.3, r=0.3$ 。

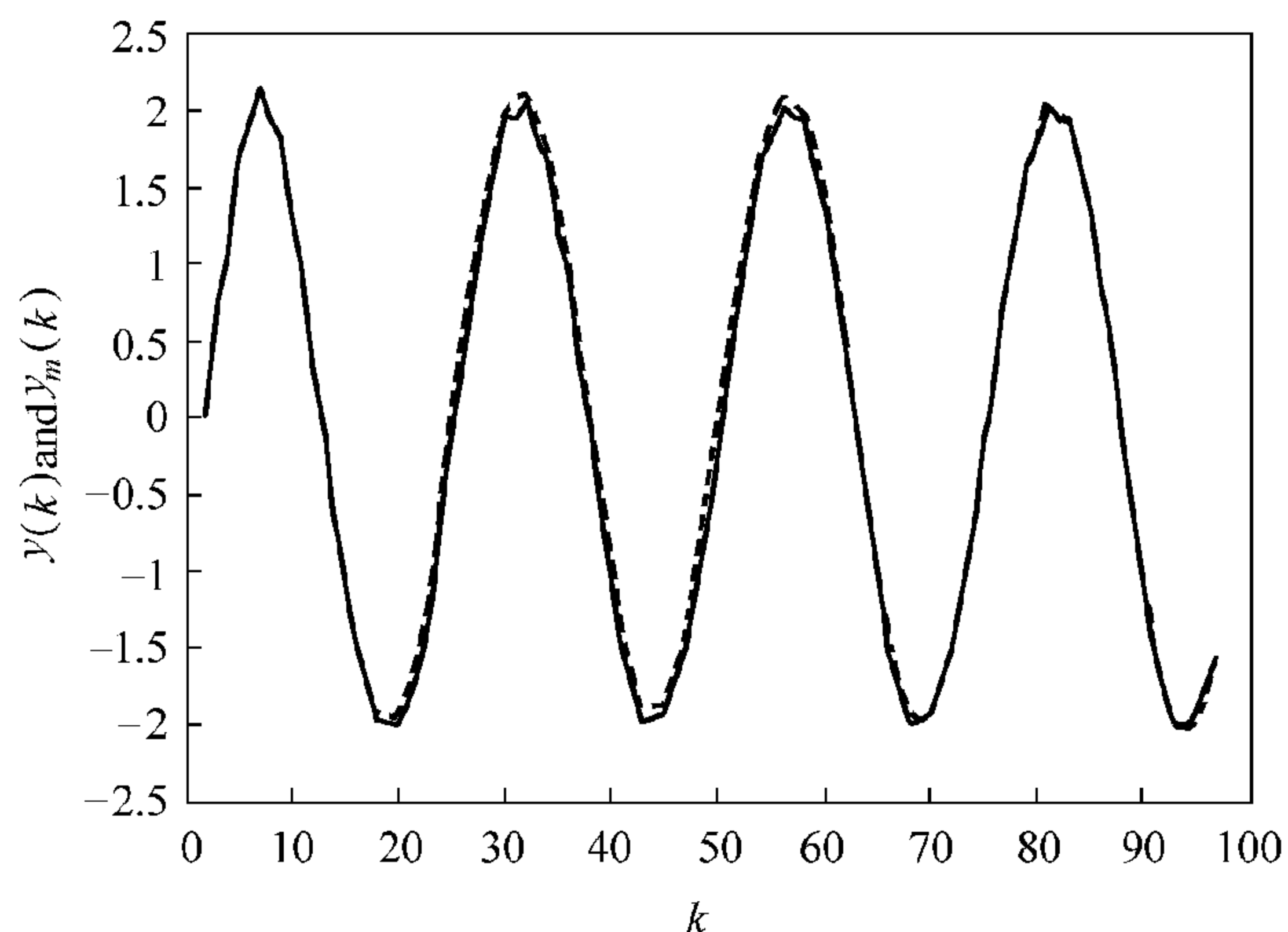


图 15.10 例 15.4 中的闭环系统的输出 $y(k)$ (实线) 和参考轨迹 $y_m(k)$ (虚线)

15.4 本章小结及导读

本章主要阐述了以下内容:

- 最优模糊系统的思想和构造;
- 根据输入—输出数据对, 利用最近邻聚类法设计模糊系统的详细步骤;
- 所设计的模糊系统在离散时间动态系统的自适应控制和其他问题中的应用。

在模式识别的教材中可以找到不同的聚类算法, 其中 Duda 和 Hart[1973] 仍是最好的教材之一。本章中的方法摘自 Wang[1994a], 从中可以找到更多的例子。

15.5 习题与设计

习题 15.1 在 $r=2.2$ 的情况下重新计算例 15.2。

习题 15.2 修正 15.2 节中的设计方法, 令聚类中心为每组中各输入点的均值, 参数 $A^l(k)$ 是该组中输入点的平均值, 删掉参数 $B^l(k)$ 。

习题 15.3 试用一个案例来说明,即使输入—输出数据对的集合相同,只要输入—输出数据对所用的阶数不同,15.2 节中的聚类法就可以得到不同的模糊系统。

习题 15.4 分级聚类的基本思想见图 15.11,用分级聚类思想提出一种设计模糊系统的方法。一步步地详细证明你的方法并用一个简单的例子来说明它。

习题 15.5 考虑下面的二维输入二维输出系统:

$$\begin{bmatrix} y_1(k+1) \\ y_2(k+2) \end{bmatrix} = \begin{bmatrix} \frac{y_1(k)}{1+y_2^2(k)} \\ \frac{y_1(k)y_2(k)}{1+y_2^2(k)} \end{bmatrix} + \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}$$

(15.25)

这里,假设非线性函数未知。

(a) 用模糊系统(15.7)或(15.8)设计一个该黑箱系统的辨识器,并解释辨识器的工作程序。

(b) 为该系统设计一个辨识器,使得闭环系统的输出能跟踪下面的参考模型:

$$\begin{bmatrix} y_{m1}(k+1) \\ y_{m2}(k+1) \end{bmatrix} = \begin{bmatrix} y_{m1}(k) + y_{m2}(k) \\ y_{m2}(k) \end{bmatrix} + \begin{bmatrix} r_1(k) \\ r_2(k) \end{bmatrix} \quad (15.26)$$

这里, $r_1(k)$ 和 $r_2(k)$ 是已知的参考信号。在什么条件下,学习误差会收敛于0?

15.6(设计) 编写一个计算机程序来实现 15.2 节中的设计方法,并将你的程序应用于第 12 章和第 13 章中的时间序列预测问题和非线性系统辨识问题。

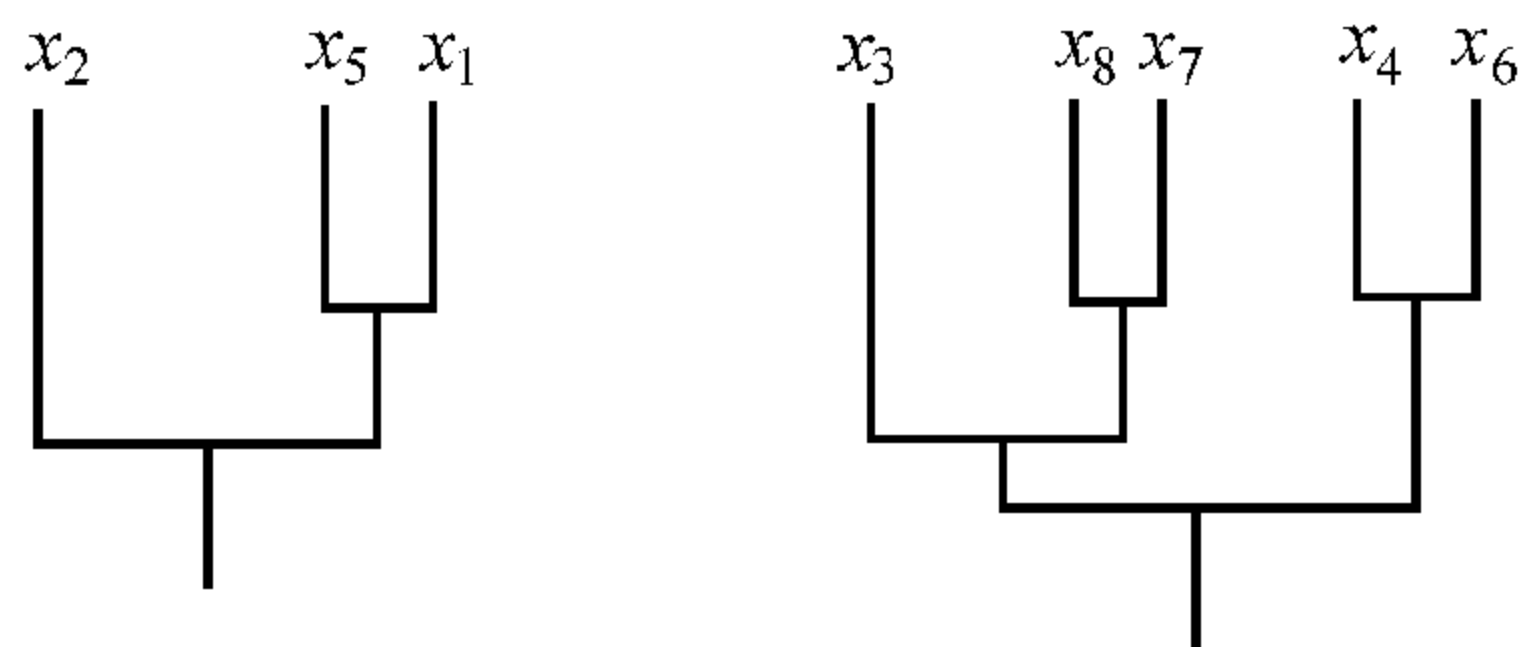


图 15.11 分层聚类的基本思想

第四部分

非自适应模糊控制

模糊系统用作控制器时称为模糊控制器。如果用模糊系统为一个过程建模,并基于此模型设计控制器,则生成的控制器也称为模糊控制器。因此,模糊控制器是具有特殊结构的非线性控制器。模糊控制代表了将模糊理论用于解决实际问题最成功的应用。

模糊控制可分为非自适应模糊控制和自适应模糊控制两类。在非自适应模糊控制系统中,模糊控制器的结构和参数都是固定的,在实时运行中不改变。在自适应模糊控制系统中,模糊控制器的结构或/和参数在实时运行中改变。非自适应模糊控制器比自适应模糊控制器简单,但需要较多过程模型的知识或启发式规则。而自适应模糊控制的应用代价虽然高昂,但它只需较少量的信息,且能达到更好的效果。本部分(第 16 至第 22 章)将研究非自适应模糊控制问题。

在第 16 章中,将通过两个实例检验模糊控制器设计的试错法,两个例子是:水泥窑的模糊控制和废水处理过程的模糊控制。第 17 章和第 18 章,将分别给出线性对象的稳定控制器和最优模糊控制器的设计方法。第 19 章和第 20 章,将进一步探讨非线性对象的模糊控制器的设计,其稳定性用滑动控制和监督控制的思想予以保证。第 20 章研究了 PID 控制器的模糊增益调度问题。第 21 章中的受控对象和控制器都用 TSK 模糊系统建模,并说明了如何选择参数以保证闭环系统的稳定。最后,第 22 章介绍了模糊控制系统的一些稳健性指标,并给出了多级模糊系统的一些基本知识。

模糊控制器设计的试错法

16.1 模糊控制与传统控制

模糊控制与传统控制有相似性,也有相异性。相似方面如下:

- 解决的问题类型相似,即都解决控制问题。所以二者均须阐述所有控制问题中都要遇到的问题,如稳定性和性能指标。
- 用于分析所设计的控制系统的数学工具相似,这是因为它们研究同一类系统的相同问题(稳定性、收敛性等)。

虽然如此,模糊控制与传统控制之间也存在根本的差别:

- 传统控制是从过程的数学建模开始的,控制器是为数学模型设计的;而模糊控制是从收集整理启发式规则和人类经验(即模糊 IF-THEN 规则)开始的,通过分析这些规则来设计控制器。亦即,用于构造两种控制器的信息基础是不同的,如图 16.1 所示,高级模糊控制器可以兼用启发式规则和数学模型(见第 24 章)。

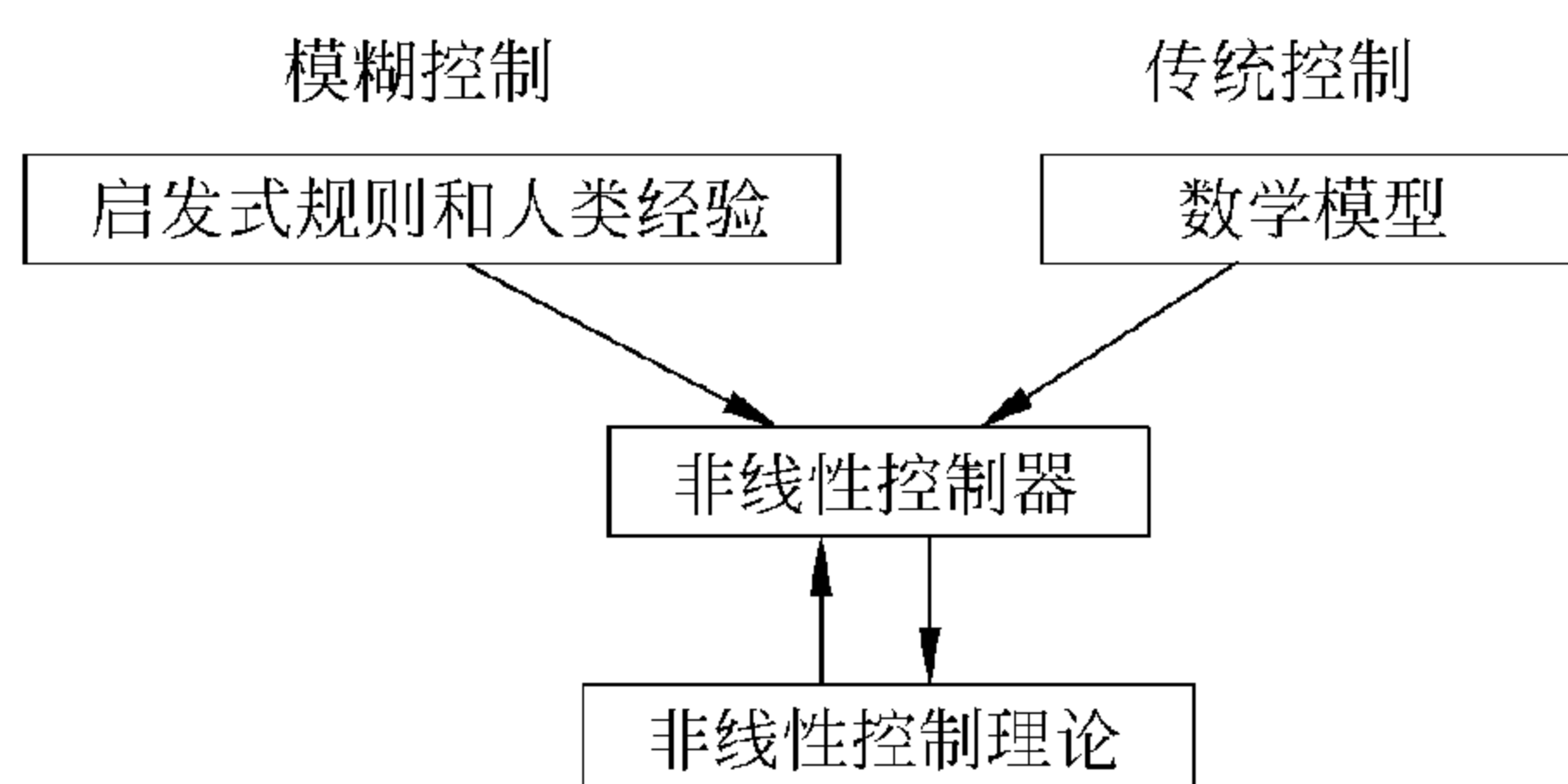


图 16.1 模糊控制与传统控制的比较

对于许多实际控制问题(如,工业过程控制),很难得到精确而又简单的数学模型,而在这种情况下,人类却能提供对过程控制有意义的启发式规则和经验原则。对于这类问题,模糊控制显得非常有用。正如将在本章和以后几章中所谈到的,在过程的数学模型未知或部分未知的情况下,模糊控制器的设计可以采用系统化的设计方法以保证其满足某些关键性能指标。

模糊控制器的设计方法可分成两类:试错法(trial-and-error approach)和理论法(theoretical approach)。在试错法中,首要任务是建立一个模糊 IF-THEN 规则集合,规则可以通过总结经验知识(如,操作手册)而得到,也可以通过精心组织的问卷向领域内专家请教而得到;然后,基于这些模糊 IF-THEN 规则建立模糊控制器;最后,在实际系统中检验模糊控制器,如果性能指标不满意,则对上述规则进行细微调整或通过反复用试错法进行再设计,直至效果满意为止。在理论法中,模糊控制器的结构和参数是以保证某一性能标准

(如,稳定性)为设计原则的。当然,在实际系统的模糊控制器设计中,要尽可能地综合运用这两种方法以得到最好的控制器。本章将通过两个实例——水泥窑模糊控制系统和废水处理模糊控制系统,来说明试错法。理论法将在第 17 章至第 21 章中学习。

16.2 模糊控制器设计的试错法

模糊控制器设计的试错法大体上可概括成以下三步:

步骤 1 分析实际系统并选择状态变量和控制变量。状态变量应能描述系统的关键特征,控制变量应该能够影响系统的状态。状态变量是模糊控制器的输入,控制变量是模糊控制器的输出。本质上说,这一步骤界定了模糊控制器的应用范围。

步骤 2 推演连接状态变量与控制变量的模糊 IF-THEN 规则。有两种启发式方法可以形成这些规则。最常见的是,自我总结归纳经验并以语言表达出来,这种方法的典型案例就是下一节要介绍的水泥窑手动操作。另一种方法包括,以精心设计的问卷向有经验的专家或操作者进行咨询。通过上述两种途径,就能得到模糊控制规则的原型了。

步骤 3 将推演出的模糊 IF-THEN 规则组合成模糊系统,并检验以此模糊系统作为控制器的闭环系统。换言之,试运行带有模糊控制器的闭环系统,如果不满意其性能指标,则用试错法对模糊控制器进行微调或再设计,重复此过程,直至结果满意为止。

下面就以水泥窑系统和废水处理过程两个实际系统为例,说明如何用试错法设计模糊控制器。

16.3 实例研究 I: 水泥窑的模糊控制

正如我们在第 1 章中提到的,水泥窑的模糊控制是模糊控制最早成功应用于工业系统全过程的实例之一。在本节中,将概括介绍 20 世纪 70 年代末由 Holmblad 和 Østergaard [1982]提出的水泥窑模糊控制系统。

16.3.1 水泥窑的工艺流程

水泥是通过磨细水泥渣块制成的。水泥渣块则是由石灰石、粘土和沙子的混合物经水泥窑加热制成的。湿法水泥窑流程中,原料混合物呈泥浆状,如图 16.2 所示。接下来的过程包括如下四个阶段:在第一阶段,水分被除去,称为预热;第二阶段,对混合原料加热并发生焙烧分解(CO_2 被分离出去);第三阶段,原料在大约 $1\ 430^\circ\text{C}$ 左右的温度下燃烧,熟石灰(CaO)与其他物质结合在一起形成水泥渣;第四阶段,对水泥渣进行风冷。

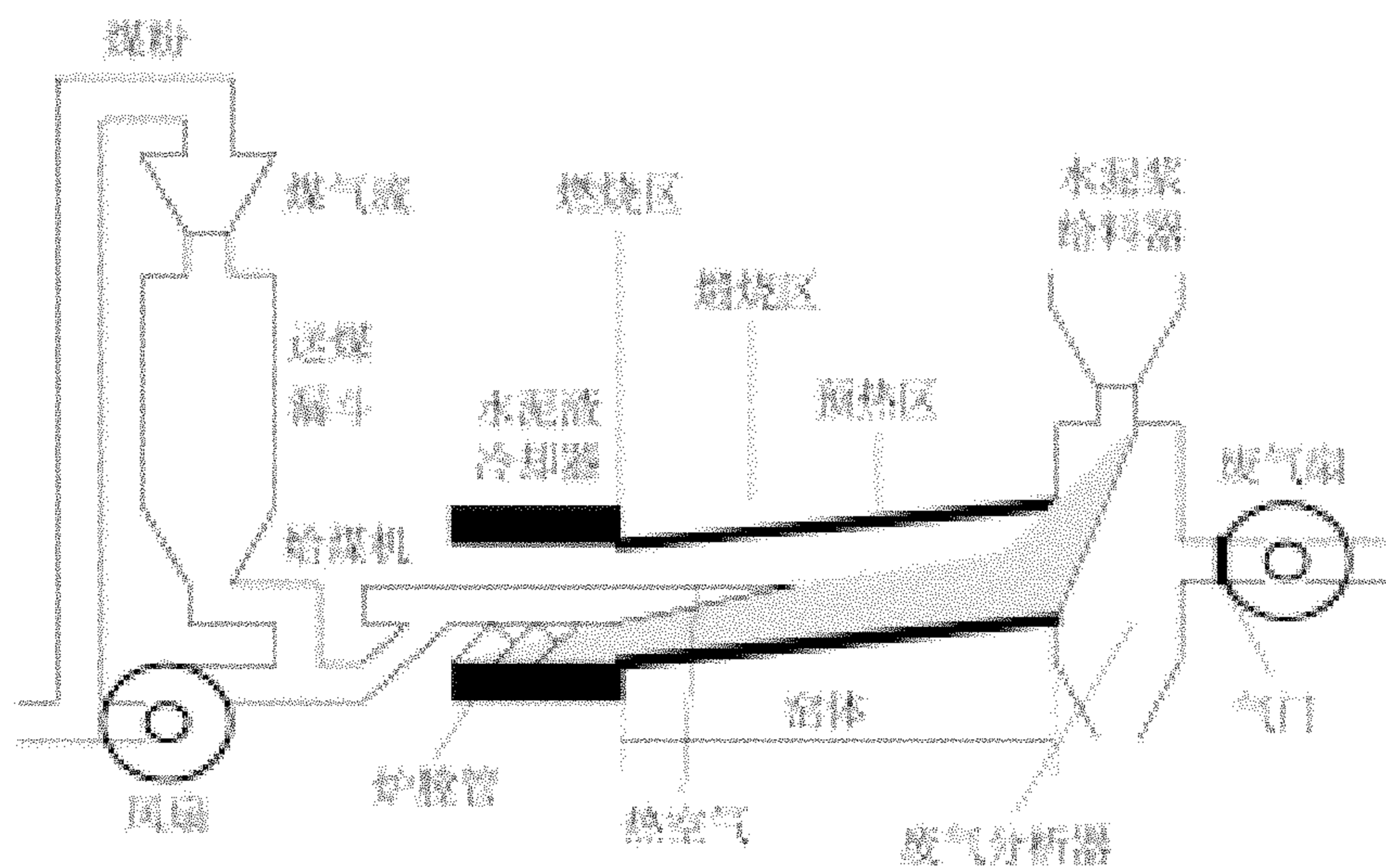


图 16.2 水泥窑的工艺流程

水泥窑是长约 165 米,直径为 5 米的钢管。水泥窑管在安置时相对于水平面会有小的倾斜角度,且每分钟旋转 1 次~2 次。水泥渣的生产是一个连续过程。水泥浆由水泥窑的高端(后部)供给,热量则由水泥窑的低(前)端供给(见图 16.2)。由于窑管的倾斜和旋转,原料在窑内进行 3 小时~4 小时的缓慢传输,并由热气加热。燃烧过程中产生的气体由废气扇排出水泥窑,并由窑上端的风扇控制(见图 16.2)。

水泥窑系统具有时变非线性的特征。实践表明,描述该流程的数学模型或因过于简单而没有实用价值,或因对特定过程考虑因素过多、针对性太强而没有通用性。然而,工作人员却可以在相对较短的时间内通过培训成为熟练操作者。所以,水泥窑尤其适合于应用模糊控制。

16.3.2 水泥窑流程的模糊控制器的设计

首先,确定本系统的状态变量和控制变量。状态变量必须能描述水泥窑流程的主要特征,且它们的量值能够通过传感器测得。控制变量必须能影响状态变量的量值。通过对系统的分析,选定如下三个状态变量:

- 燃烧区的温度,以 BZ 表示。
- 废气中氧的比例,以 OX 表示。
- 窑尾端的温度,以 BE 表示。

BZ 的量值可以从每升水泥渣的重量中得到,它是可测量的。OX 和 BE 的值可以从窑的后端的废气分析仪中得到(见图 16.2)。选定如下两个控制变量:

- 煤的给料率,以 CR 表示。
- 废气风扇的定位,以 DP 表示。

废气风扇的定位会影响窑内通过的空气量。每个控制变量对预热、焙烧、形成渣和冷却

等各阶段都会有影响,且具有不同的时滞,时滞时间短则几分钟,长则数小时。从通常的观点来看,在水泥窑的控制中,窑的速度和水泥浆的原料的供给速率也可以用作控制变量,但因为水泥窑正常情况下都是恒速生产的,所以原料供给速率的调整和窑的传输速度很少用于调节控制。

第二步,导出联系状态变量 BZ、OX、BE 和控制变量 CR、DP 的模糊 IF-THEN 规则。这里,根据水泥窑控制的操作手册导出这些规则。图 16.3 摘自于一本面向水泥窑操作者的美国教材。这部分描述了,为什么操作者必须根据不同情况——以燃烧区的温度(BZ)、废气含氧百分比(OX)、水泥窑尾端的温度(BE)——来调节燃料供给率、窑的传输速率、窑中的空气通过量。可以看到,环境因素和控制行为以定性术语描述为“高”、“适中”、“低”、“轻微增加”等。

实例	条 件	采取的措施	原 因
10	BZ 适中 OX 低 BE 低	a. 提高风扇速度	提高尾部温度,为行动“b”增加氧量
11	BZ 适中 OX 低 BE 适中	b. 提高燃料供给速度 a. 轻微减少燃料供给速度	保持燃烧区温度 增加氧量
12	BZ 适中 OX 低 BE 高	a. 减少燃料供给速度 b. 降低风扇速度	为行动“b”增加氧量 降低尾部温度且保持燃烧区温度
13	BZ 高 OX 适中 BE 适中	a. 提高风扇速度 b. 提高燃料供给速度	提高尾部温度 保持燃烧区温度
14	BZ 适中 OX 适中 BE 适中	无须行动,但要避免盲目自信,密切观察状态变化	
15	BZ 适中 OX 适中 BE 高	当氧量处于高位值时 a. 降低风扇速度 当氧量处于低位值时 b. 减少燃料供给速度 c. 降低风扇速度	降低尾部温度 为行动“c”增加氧量 降低尾部温度且保持燃烧区温度
16	BZ 适中 OX 高 BE 低	a. 提高风扇速度 b. 提高燃料供给速度	提高尾部温度 保持燃烧区温度且减少氧量
17	BZ 适中 OX 高 BE 适中	a. 轻微减少风扇速度	减少氧量

图 16.3 美国水泥窑操作手册选摘

为了把图 16.3 的指令变成模糊 IF-THEN 规则,首先要定义不同变量取值为“低”、“适中”或“高”的隶属度函数。图 16.4 给出了变量 OX 取值为上述三个术语的隶属度函数。其他的与此类似,这里略去细节。从图 16.3 中可以得到模糊 IF-THEN 规则集合。例如,从图 16.3 中的第 10 项,可以得到如下的模糊 IF-THEN 规则:

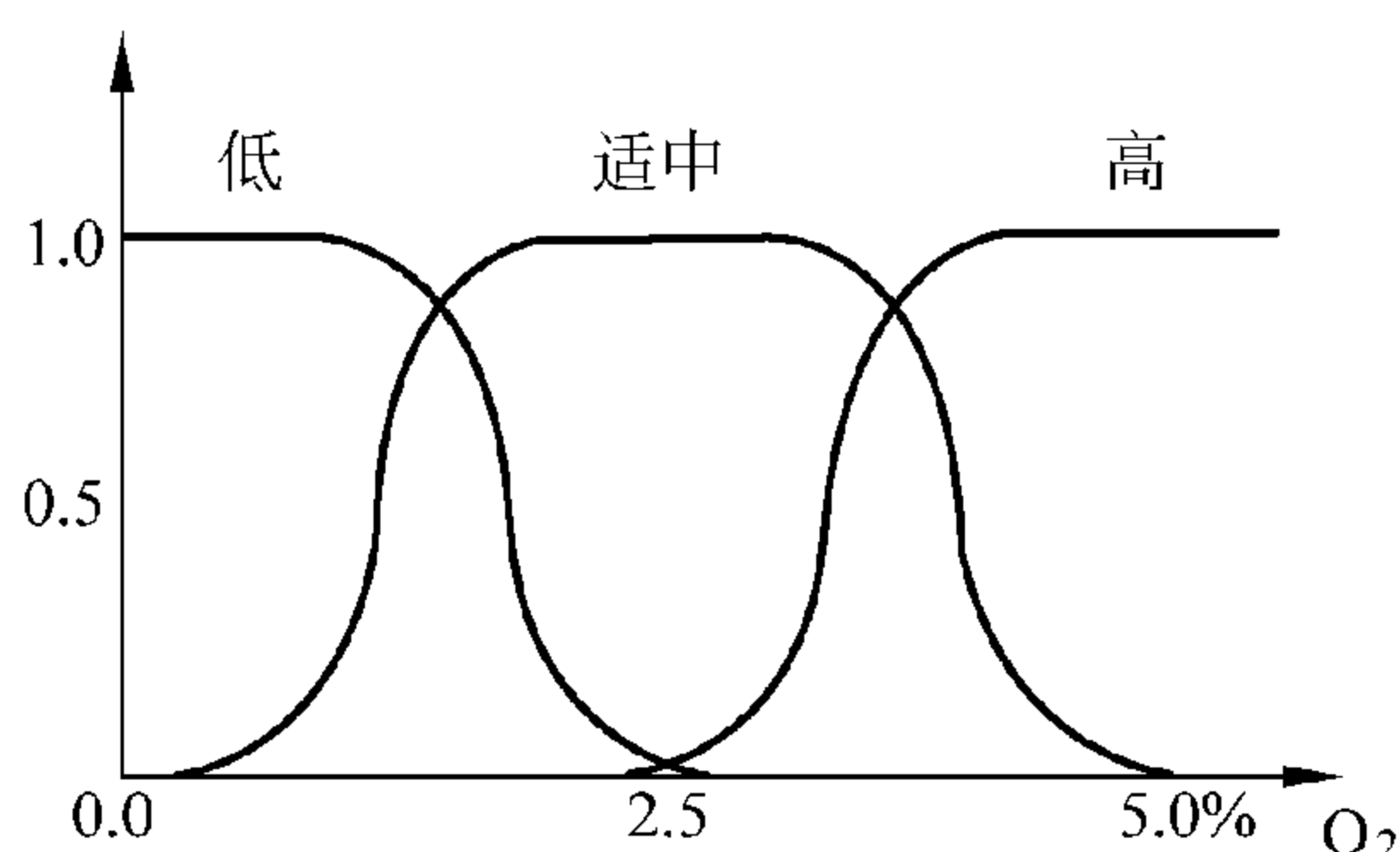


图 16.4 变量 OX 取值为“低”、“适中”和“高”的隶属度函数

如果 BZ 是适中的,OX 是低的,且 BE 是低的,则 CR 是大量的,DP 是大量的
(16.1)

类似地,第 11 项给出了规则:

如果 BZ 是适中的,OX 是低的,且 BE 是适中的,则 CR 是小量的
(16.2)

在实际应用中,全套的模糊 IF-THEN 规则共有 27 条。

16.3.3 模糊控制器的实施

为了开发水泥窑过程的模糊控制系统,仅有一个模糊控制器是不够的。应设计一系列的模糊控制器来满足不同模式下系统运行的需要。具体地讲,考虑下面两种系统运作模式:

- 系统运行非常稳定,通过对水泥窑的驱动扭力的测定,显示扭力只有细微变化。
- 系统处于不稳定状态,通过对水泥窑的驱动扭力的测定,显示扭力出现大幅度摆动式变化。

在 Holmblad 和 Østerguard[1982]开发的水泥窑模糊控制系统中,共有八个运行子程序,其中每个子程序代表一个模糊控制器或某些支持操作。图 16.5 列举了这些子程序。子程序 A 监控水泥窑是否处于稳定运行状态,它能在 8 小时内调整扭力的变化。如果运行是不稳定的,过程控制将由子程序 B 来完成,该子程序仅通过调整煤量就可使运行重归稳定。

在稳定运行中,状态变量的目标值是:每升水泥渣的重量约为 1 350 克,尾气中的氧含量约为 1.8%,窑的尾端的温度约为 197℃。为了达到和保持目标状态,子程序 C 调整煤的供给速率和尾气风扇的位置,该子程序即是上节提到的模糊控制器,它完成主要控制操作。

由于水泥渣密度是由人工通过渣样确定的,并由人工通过操作控制板加入到系统中,所以当将水泥渣的密度值添加到系统中时,与窑的实际状态相比,它代表的是 1 小时~2 小时以前某个时刻的状态值。这种时滞在于子程序 D 和 E 中均有考虑。子程序 D 负责监测尾气中 NO 成分的变化,如果 NO 成分上升,是表明燃烧区温度升高,进而导致密度值增大。NO 含量的升降信息在子程序 E 中被用来调节添加的密度值。例如,如果添加的密度值为高,

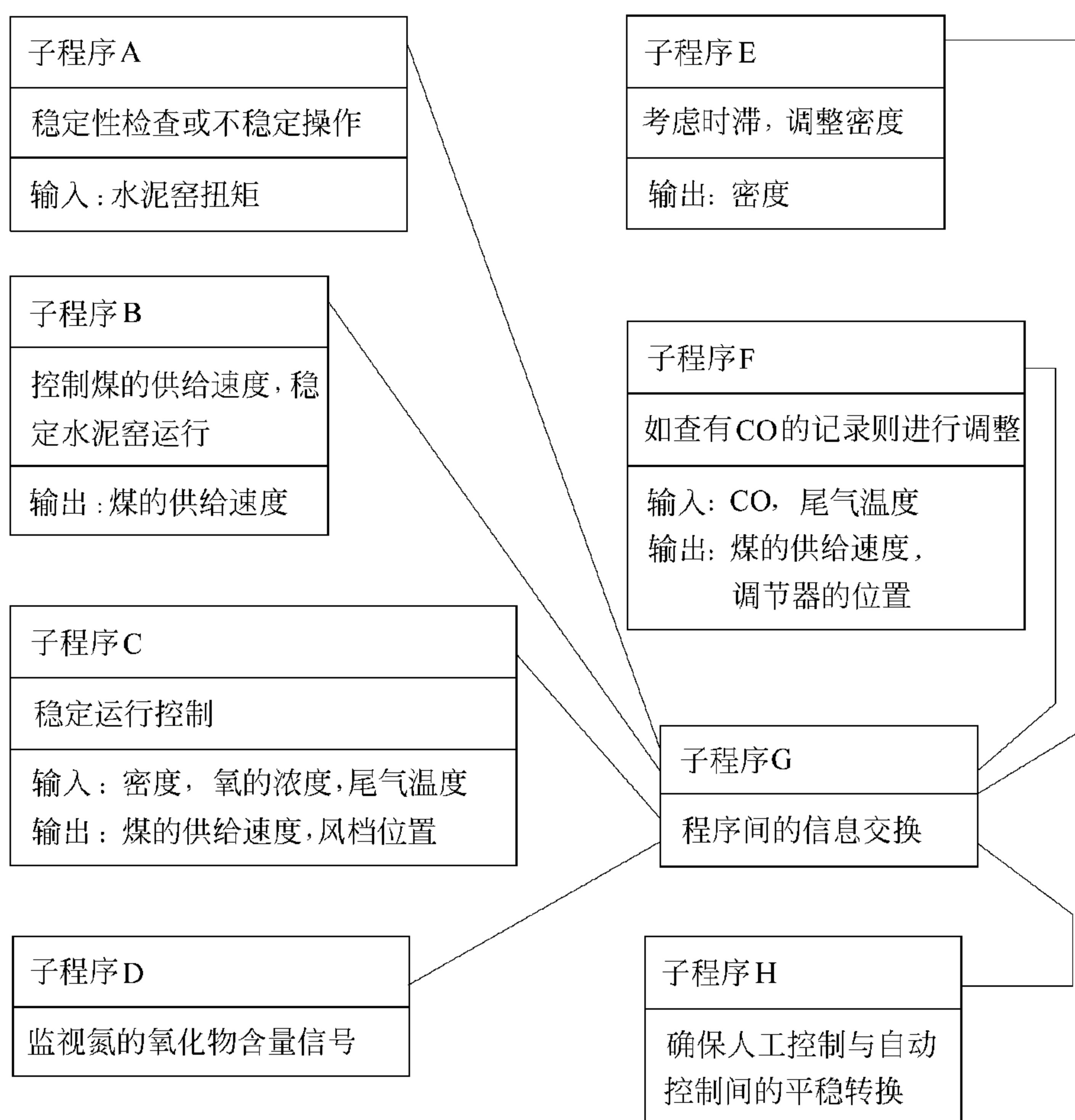


图 16.5 水泥窑控制流程图

而 NO 的含量应该被调低时,则密度值将根据 NO 含量所需的降幅调低。子程序 E 还考虑了工业过程在进行控制调整后达到预期状态所需的时间。

无论是在稳定还是在非稳定操作中,子程序 F 均担任尾气中 CO 含量的测量工作,并对之实施适当的控制。子程序 G 负责各子程序间信息的传递。最后,子程序 H 确保操作人员能在水泥窑的手工控制和自动控制间平稳转换。

16.4 实例研究 II: 废水处理过程的模糊控制

16.4.1 污物废水处理过程

污物激活流程是常用的污水处理过程。图 16.6 是该系统的示意图。此过程(虚线框内的部分)包括一个充气水池和一个沉淀池。首先,进入此处理流程的废水和回收的污水是混在一起的。然后,空气通过沿充气水池安置的扩散装置吹入混合液中。在充气池中会

发生复杂的生物、化学反应,这样水和废物才能分开。最后,处理过的混合液进入沉淀池,废物在此处沉淀下来,净水则被排放。

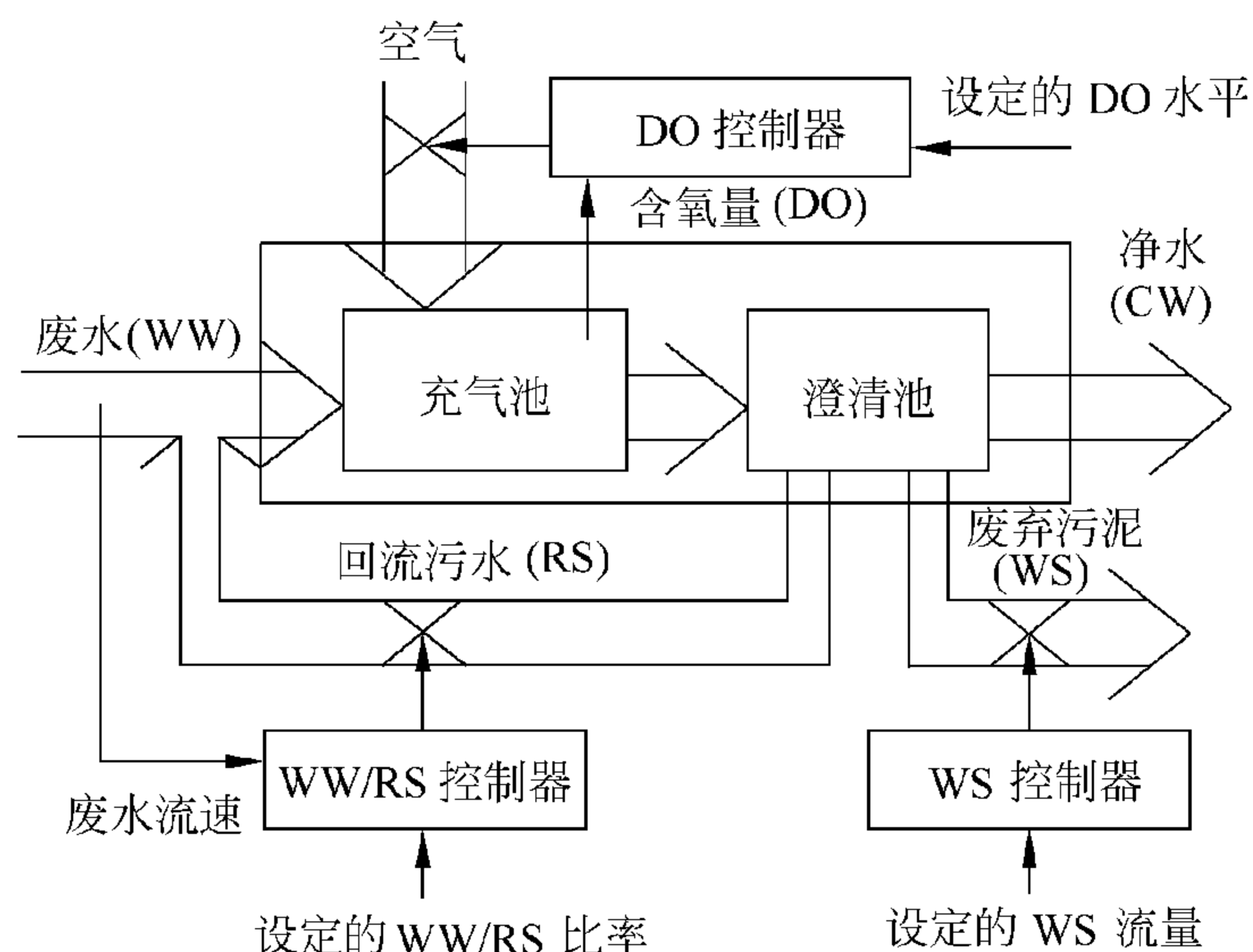


图 16.6 污物激活法污水处理过程的示意图

如图 16.6 所示,系统共有三个下级控制器。其中,WW/RS 控制器将废水流速与回收污水流速之比控制在要求的量值范围内。这个控制器的目标是维持理想的基质/有机物浓度比率,控制污水在充气池和沉淀池的分布。DO 控制器控制空气流速,使充气池保持要求的氧气溶解度。因为要使废水里的含氮废物氧化,所以必须具有较高的氧气溶解度(这被称作氮的硝化作用)。最后,WS 控制器通过控制污水废水流速,控制充气池和沉淀池中污水的总量和平均滞留时间。为了完成氮的硝化过程,污水通常需要有较高的滞留时间。

因为对流程中基本的生物机理知之甚少,很难得到可用的数学模型。而实际中 WW/RS 比率值、DO 水平值和 WS 流速的目标值又是人工设定和调整的。所以我们的目标是,将人工操作的经验总结成模糊系统,从而使操作人员从在线操作中解放出来。亦即,这里我们将设计一个模糊控制器,它将能给出 WW/RS 比率、DO 水平和 WS 流速的目标值。此模糊控制器即成为一个上级决策机,下级的直接控制则分别由 WW/RS、DO 和 WS 控制器实现。

16.4.2 模糊控制器的设计

首先,选出状态变量和控制变量,显然,可得如下三个控制变量:

- WW/RS: 将 WW/RS 比率调节至要求值。
- DO: 将 DO 调节至要求值。
- WS: 将 WS 流调节至要求值。

状态变量应能表征系统的基本特征。既然废水处理过程的最终目标是将输出净水中生化氧(biochemical oxygen)和固体悬浮物的含量控制在某一标准之下,那么这两个变量就应

作为状态变量。另外,混合液离开充气池时,固体悬浮物含量和回流污水中固体悬浮物含量也很重要,也应作为状态变量。最后,输出净水中的氨氮含量($\text{NH}_3\text{-N}$)和污水流速也应视为状态变量。归纳可得如下六个状态变量:

- TBO: 输出净水中生化氧的总量。
- TSS: 输出净水中固体悬浮物的总量。
- MSS: 离开充气池的混合液中的固体悬浮物的量。
- RSS: 回流污水中固体悬浮物的量。
- $\text{NH}_3\text{-N}$: 输出净水中氨—氮的含量。
- WSR: 污水流速。

下一步的任务是推演联系状态变量与控制变量的模糊 IF-THEN 规则。根据人工操作的经验,由 Tong, Beck 和 Latten[1980]提出了的 15 条规则(见表 16.1),其中, S、M、L、SN、LN、SP、LP、VS 和 NL 分别对应模糊集“少量”、“中等”、“大量”、“负值少量”、“负值大量”、“正小量”、“正大量”、“极少量”和“不大”等。规则 1~规则 2 是调整规则,指示如果过程处理在理想状态,而 WSR 的值异常,则需要对 WSR 做相应的调整。规则 3~规则 6 处理的是输出净水中 $\text{NH}_3\text{-N}$ 的含量高的情况。规则 7~规则 8 负责处理输出净水中固体物质偏多的情况。规则 9~规则 13 描述了 MSS 超标时所需的控制操作。最后,规则 14~规则 15 处理的是输出净水中生化氧含量高的情况。

表 16.1 污水处理模糊控制器的模糊 IF-THEN 规则

规则号	TBO	TSS	MSS	RSS	$\text{NH}_3\text{-N}$	WSR	$\Delta\text{WW}/\text{RS}$	ΔDO	ΔWS
1	S	S	M	M	S	S			SP
2	S	S	M	M	S	L			SN
3		S			M		SP		
4		S			M				SN
5		S			L		LP		
6		S			L				LN
7	NL	M						SP	
8	NL	L						LP	
9			L						LP
10			S						SN
11			VS						LN
12			VS			S		SP	
13			L			L		SN	
14	M	S			S				SN
15	L	S			S				LN

此系统的详细描述见 Tong, Beck 和 Latten[1980]。

16.5 本章小结及导读

本章主要阐述了以下内容:

- 试错法设计模糊控制器的一般步骤;
- 水泥窑系统及其模糊控制器的设计;
- 实现水泥窑模糊控制系统所要考虑的因素;
- 污物废水激活处理过程的模糊控制器的设计。

模糊控制的实例研究很多, Sugeno 和 Nishida[1985]是早期的应用成果, Terano, Asai 和 Sugeno[1994]是近期的应用成果。本章研究的两个实例节选自 Holmblad 和 Østergaard [1982]及 Tong, Beck 和 Latten[1980],在原著中有更多详细阐述。

16.6 习题

习题 16.1 考虑图 1.9 中的倒立摆系统,以 $x_1 = \theta, x_2 = \dot{\theta}$, 作为状态变量。倒立摆系统的状态方程是

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \frac{g \sin x_1 - \frac{m l x_2^2 \cos x_1 \sin x_1}{m_c + m}}{l \left(\frac{4}{3} - \frac{m \cos^2 x_1}{m_c + m} \right)} + \frac{\frac{\cos x_1}{m_c + m}}{l \left(\frac{4}{3} - \frac{m \cos^2 x_1}{m_c + m} \right)} u \end{aligned} \quad (16.3)$$

其中, $g = 9.8 \text{ m/s}^2$ 是重力加速度, m_c 是小车的质量, m 是倒立摆的质量, l 是倒立摆的长度的一半, u 是施加的外力(控制)。用试错法设计平衡此倒立摆的模糊控制器。令 $m_c = 1 \text{ kg}$, $m = 0.1 \text{ kg}$, $l = 0.5 \text{ m}$, 通过计算机仿真检验你的模糊控制器。

习题 16.2 考虑 16.7 所示的球—横木系统,令 $x = (r, \dot{r}, \theta, \dot{\theta})^T$ 为系统的状态向量, $y = r$ 为系统的输出变量。这样系统的状态空间模型可表示为

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} x_2 \\ \alpha(x_1 x_4^2 - \beta \sin x_3) \\ x_4 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} u \quad (16.4)$$

$$y = x_1 \quad (16.5)$$

其中,控制变量 u 是 θ 的加速度, α, β 是参数,用试错法设计一个模糊控制器,把球从某些初

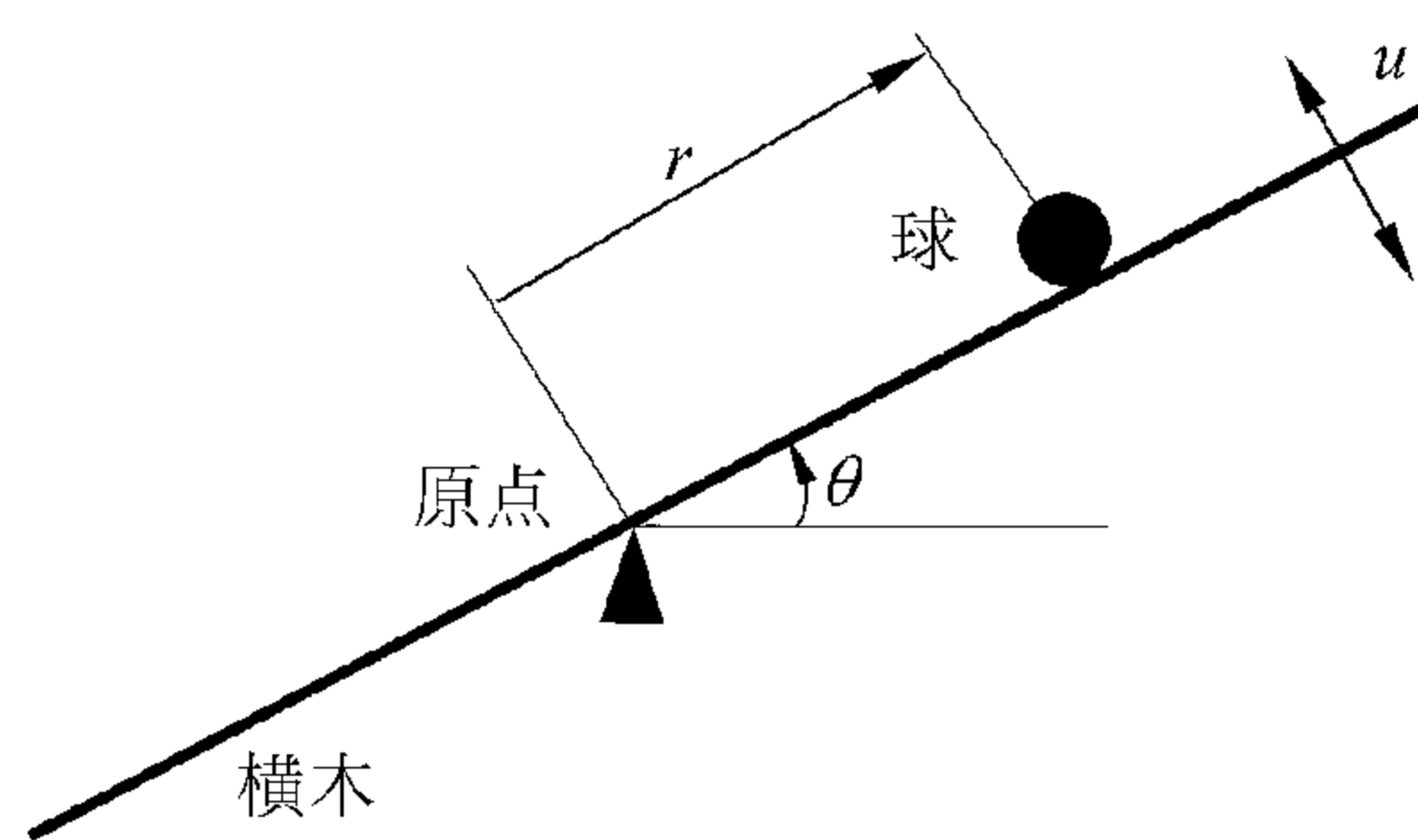


图 16.7 球—横木系统

始位置控制到原点。用计算机仿真检验 $\alpha = 0.7143, \beta = 9.81$ 时的模糊控制器。

习题 16.3 在 16.1 节中讨论的模糊控制和传统控制之间异同点的基础上,进一步探讨二者的异同。

习题 16.4 在 Sugeno, Nishida[1994]或 Terano, Asai, Sugeno[1994]中找一个模糊控制的应用实例,并描述模糊控制系统工作原理的一些细节。

线性系统模糊控制 I: 稳定控制器

由第 16 章可知,设计模糊控制器的起点是,根据专家的操作手册收集模糊 IF-THEN 规则。这意味着,在设计模糊控制器时,不必知道受控过程的数学模型。但是,为了从理论上分析所设计的闭环模糊控制系统,至少要有一个粗略的过程模型。此外,为保证设计的模糊控制器能达到某一性能指标要求,如全局指数稳定性指标,必须为过程假定一个数学模型,以便通过进行数学分析给出要设计系统的特性。本章将研究过程为线性系统而控制器为模糊系统的一个实例。本章的目的是,给出保证闭环模糊控制系统稳定的模糊控制器应满足的条件。

17.1 单输入单输出系统的稳定模糊控制

对任何控制系统(包括模糊控制系统)而言,稳定性都是重要指标,因为不稳定控制系统是典型的无用之物,而且还存在潜在的危险。从概念上讲,稳定性有两类: Lyapunov(李亚普诺夫)稳定性和输入—输出稳定性。在每一类稳定性中都有一系列特殊的稳定性概念,这里先简要回顾一下这些概念。

对于 Lyapunov 稳定性,先考虑自治系统的情形:

$$\dot{x}(t) = g[x(t)] \quad (17.1)$$

式中, $x \in R^n$, $g(x)$ 是一个 $n \times 1$ 的向量函数。假定 $g(0) = 0$, 则 $x = 0$ 是此系统的一个平衡点。

定义 17.1 (Lyapunov 稳定性): 对于平衡点 $x = 0$, 如果对任意一个 $\varepsilon > 0$, 存在 $\delta > 0$, 使得 $\|x(0)\| < \delta$ 时, 对所有 $t \geq 0$ 都有 $\|x(t)\| < \varepsilon$, 则称平衡点 $x = 0$ 是稳定的。如果平衡点 $x = 0$ 是稳定的, 且存在 $\delta' > 0$, 使得由 $\|x(0)\| < \delta'$, 有当 $t \rightarrow \infty$ 时, $x(t) \rightarrow 0$, 则可知该平衡点是渐近稳定的。最后, 如果存在正数 α, λ 和 γ , 对于所有 $t \geq 0$, $\|x(0)\| < \gamma$, 有

$$\|x(t)\| \leq \alpha \|x(0)\| e^{-\lambda t} \quad (17.2)$$

则称平衡点 $x = 0$ 是指数稳定的。如果对任意初始点 $x(0)$, 平衡点 $x = 0$ 都是渐近稳定或指数稳定的, 则相应地称该平衡点是全局渐近稳定的或全局指数稳定的。

下面考虑输入为 $u(t) \in R^r$, 输出为 $y(t) \in R^m$ 的任意系统的输入—输出稳定性。

定义 17.2 (输入—输出稳定性): 令 L_p^n 为所有向量函数集 $g(t) = (g_1(t), \dots, g_n(t))^T$:

$[0, \infty) \rightarrow R^n$, 它使 $\|g\|_p = \left(\sum_{i=1}^n \|g_i\|_p^2 \right)^{1/2} < \infty$ 成立, 其中 $\|g_i\|_p = \left(\int_0^\infty |g_i(t)|^p dt \right)^{1/p}$, $p \in [1, \infty)$, 且 $\|g_i\|_\infty = \sup_{t \in [0, \infty)} |g_i(t)|$ 。一个输入为 $u(t) \in R^r$, 输出 $y(t) \in R^m$ 的系统:

$$u(t) \in L_p^r, \text{能推出 } y(t) \in L_p^m \quad (17.3)$$

则称此系统为 L_p -稳定的, 其中 $p \in [1, \infty)$ 。特别地, 如果 $u(t) \in L_\infty^r$, 可推出 $y(t) \in L_\infty^m$, 则称该系统为 L_∞ 稳定(即有界输入—有界输出稳定)的。

现假定所要控制的过程是单输入—单输出(SISO)的时不变线性系统, 由其状态变量表示的模型如下:

$$\dot{x}(t) = Ax(t) + bu(t) \quad (17.4)$$

$$y(t) = cx(t) \quad (17.5)$$

式中, $u \in R$ 是控制变量, $y \in R$ 是输出变量, $x \in R^n$ 是状态变量。下面介绍有关此线性系统的几个重要概念。

定义 17.3 [可控性(controllability), 可观性(observability), 正实域性(Positive Real)] 如果由式(17.4)和式(17.5)组成的系统满足

$$\text{rank}[b \ AB \ \cdots \ A^{n-1}b] = n \quad (17.6)$$

则称系统是控的。若满足

$$\text{rank} \begin{bmatrix} c \\ cA \\ \cdots \\ cA^{n-1} \end{bmatrix} = n \quad (17.7)$$

则称系统是可观的。如果系统的传递函数 $h(s) = c(sI - A)^{-1}b$ 满足

$$\inf_{\omega \in R} \text{Re}[h(j\omega)] > 0 \quad (17.8)$$

则称系统是严格正实域的。

至此将控制理论的某些相关结果应用于模糊控制系统的准备工作已经就绪。

17.1.1 模糊控制系统的指数稳定性

假定由式(17.4)和式(17.5)组成的系统的控制变量 $u(t)$ 是一个模糊系统, 其输入是 $y(t)$, 即

$$u(t) = -f[y(t)] \quad (17.9)$$

其中, f 是一个模糊系统。将式(17.9)代入式(17.4)和式(17.5)中, 就可得到如图 17.1 所示的闭环模糊控制系统。下面引述控制理论中的一个著名的结果(其证明可在 Vidyasagar [1993]中找到)。

命题 17.1 考虑由式(17.4)、式(17.5)和式(17.9)组成的闭环系统, 假定(a) A 的所

有特征值都位于复平面的左半开平面;(b)系统(17.4)和(17.5)是能控的和能观的;(c)系统(17.4)和(17.5)的传递函数属严格正实域。如果非线性函数 f 满足 $f(0) = 0$, 且

$$yf(y) \geq 0, \forall y \in R \quad (17.10)$$

则闭环系统(17.4)、(17.5)和(17.9)的平衡点 $x = 0$ 是全局指数平衡点。

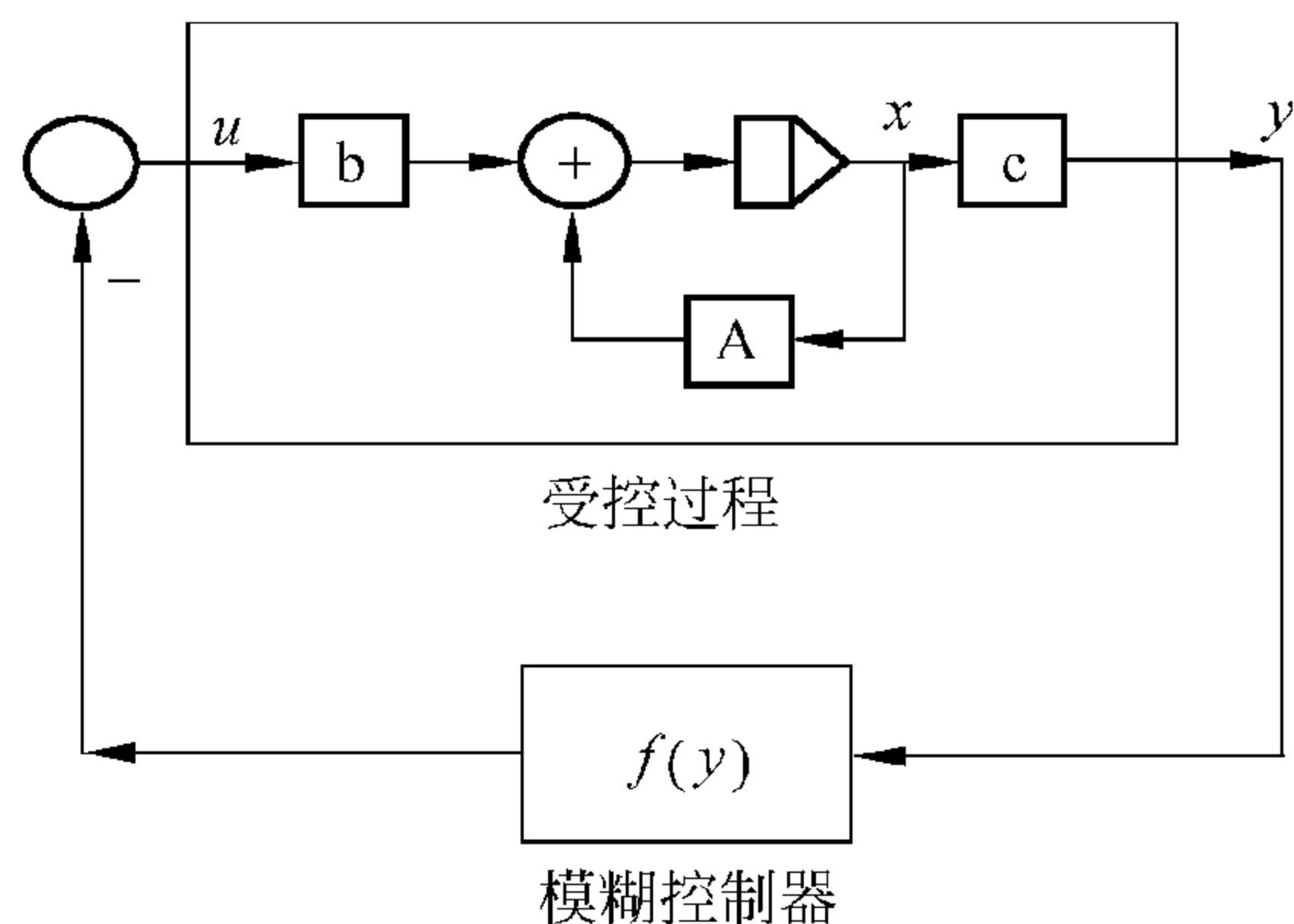


图 17.1 闭环模糊控制系统

命题 17.1 中的条件(a)~(c)是对所控制过程的限定,而不是对控制器 $f(y)$ 的限制。这是很强的限制条件,实质上是要求开环系统稳定且运行良好。从概念上讲,这些系统并不难以控制,所以对模糊控制器施加的限制条件 $f(0) = 0$ 和式(17.10)等也不是很强。当所要控制的过程是线性的且满足命题 17.1 中的条件(a)~(c)时,如果所设计的模糊控制器 $f(y)$ 满足 $f(0) = 0$ 和式(17.10),则命题 17.1 能够保证闭环系统是全局指数稳定的。下面将设计这样一个模糊控制器。

稳定模糊控制器的设计

步骤 1 假定输出 $y(t)$ 的值域是闭区间 $U = [\alpha, \beta] \subset R$ 。在 U 上定义 $2N+1$ 个模糊集 A^l , 如图 17.3 所示,它们是标准的、一致的且完备的,并具有三角形隶属度函数。即,这里用 N 个模糊集 A^1, \dots, A^N 覆盖了负值区域 $[\alpha, 0)$, 另外 N 个模糊集 A^{N+2}, \dots, A^{2N+1} 涵盖了正值区域 $(0, \beta]$, 并设模糊集 A^{N+1} 的中心 \bar{x}^{N+1} 为原点。

步骤 2 考虑下面的 $2N+1$ 条模糊 IF-THEN 规则:

$$\text{如果 } y \text{ 为 } A^l, \text{ 则 } u \text{ 为 } B^l \quad (17.11)$$

其中 $l = 1, 2, \dots, 2N+1$, 且模糊集 B^l 的中心 \bar{y}^l 根据下面方法选定:

$$\bar{y}^l \begin{cases} \leq 0 & l = 1, \dots, N \\ = 0 & l = N+1 \\ \geq 0 & l = N+2, \dots, 2N+1 \end{cases} \quad (17.12)$$

步骤 3 用乘积推理机、单值模糊器和中心平均解模糊器,根据式(17.11)中的 $2N+1$ 条模糊规则设计模糊控制器,即设计出如下的模糊控制器:

$$u = -f(y) = - \frac{\sum_{l=1}^{2N+1} \bar{y}^l \mu_{A^l}(y)}{\sum_{l=1}^{2N+1} \mu_{A^l}(y)} \quad (17.13)$$

其中, $\mu_{A^l}(y)$ 如图 17.2 所示, \bar{y}^l 满足式(17.12)。

下面证明,根据以上三步设计的模糊控制器会形成一个稳定的闭环系统。

定理 17.1 考虑图 17.1 中的闭环模糊控制系统。如果模糊控制器 $u = -f(y)$ 是通过以上三步设计的[亦即, u 由式(17.13)给出], 且控制过程满足命题(17.1)中的条件(a)~

(c), 则闭环模糊控制系统的平衡点 $x=0$ 一定是全局指数稳定的。

证明: 根据命题 17.1, 这里只须证明 $f(0)=0$ 和 $yf(y) \geq 0$ 对于所有 $y \in R$ 均成立即可。由式(17.13)和图 17.2 可知, $f(0) = \bar{y}^{N+1} = 0$ 。如果 $y < 0$, 则由式(17.13)和图 17.2, 可得

$$f(y) = \frac{\bar{y}^{l_1} \mu_{A^{l_1}}(y) + \bar{y}^{l_1+1} \mu_{A^{l_1+1}}(y)}{\mu_{A^{l_1}}(y) + \mu_{A^{l_1+1}}(y)} \quad (17.14)$$

或者 $f(y) = \bar{y}^{l_1}, l_1 \in \{1, 2, \dots, N\}$ 。由于 $\bar{y}^{l_1} \leq 0, \bar{y}^{l_1+1} \leq 0$, 而由隶属度函数的非负性可知, $f(y) \leq 0$; 从而, $yf(y) \geq 0$ 。类似地, 可以证明, 当 $y > 0$ 时, $yf(y) \geq 0$ 。

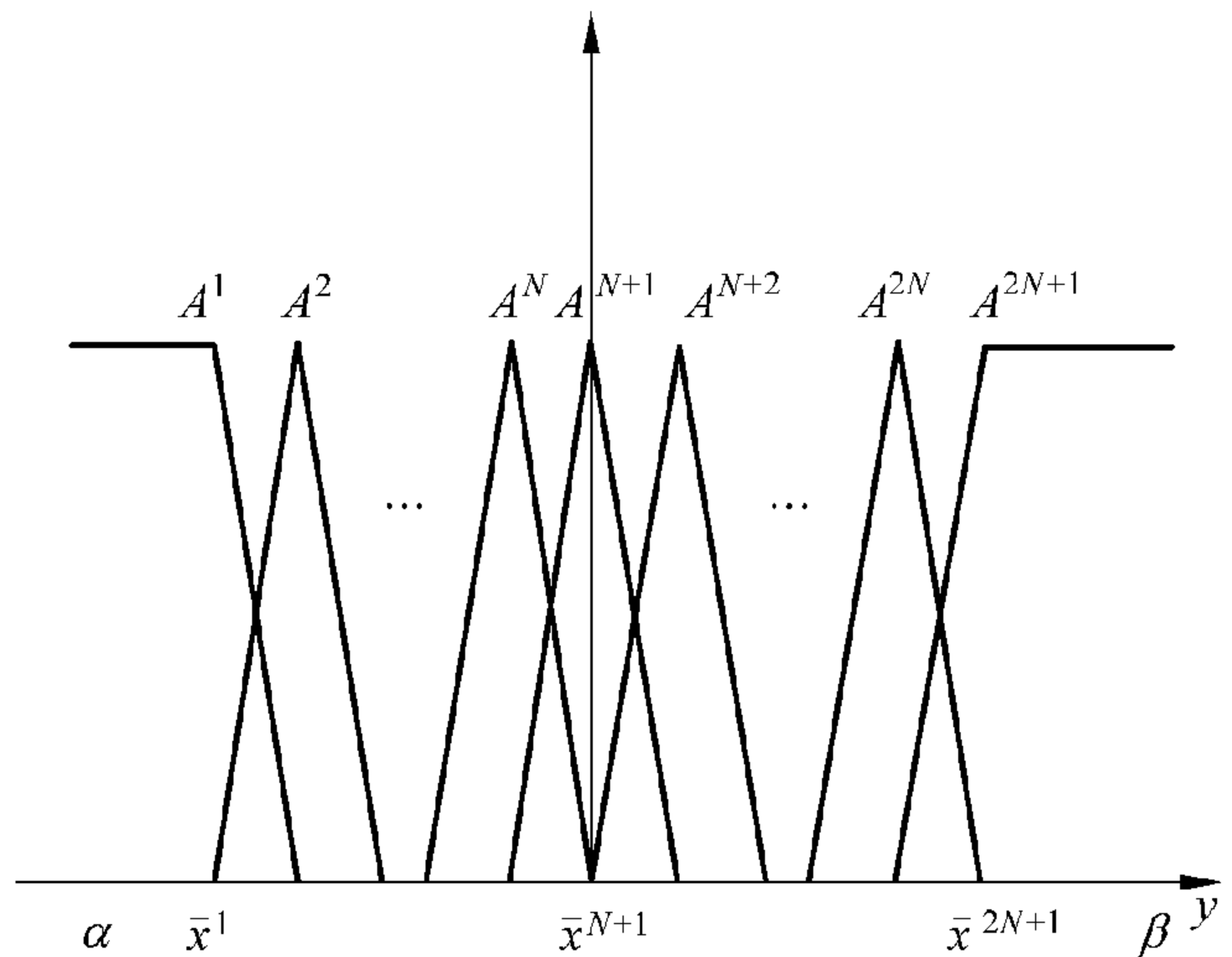


图 17.2 模糊控制器的隶属度函数

从设计过程的三个步骤可知, 在设计模糊控制器 $f(y)$ 的过程中, 不需要知道所要控制过程的参数 A, b 和 c 的确切值。与此同时, 在选择模糊控制器参数时可以有很大自由度。事实上, 只要求使所有的 \bar{y}^l 都满足式(17.12)和隶属度函数 A^l 具有如图 17.2 所示的形式即可。在第 18 章中将给出一种选择模糊控制器参数的优化方法。

17.1.2 模糊控制系统的输入—输出稳定性

图 17.1 所示的闭环模糊控制系统没有明确的输入。为了研究输入—输出稳定性, 这里

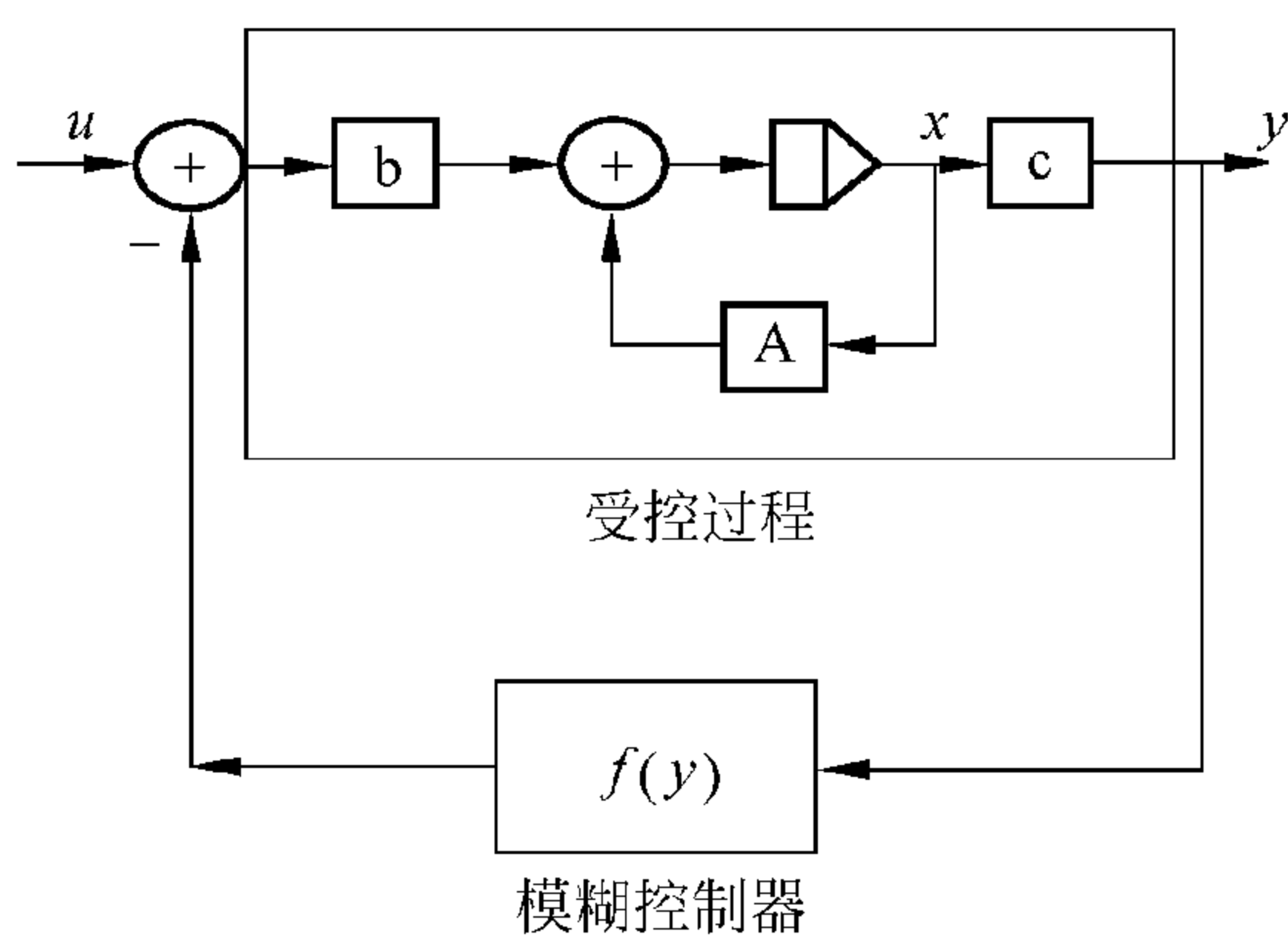


图 17.3 带有外部输入的闭环模糊控制系统

引入一个特殊的输入, 系统如图 17.3 所示。下面引述控制理论中的一个著名的结果(证明见 Vidyasagar[1993])。

命题 17.2 考虑如图 17.3 所示的系统, 假设非线性控制器 $f(y)$ 是全局 Lipschitz 连续的, 即

$$|f(y_1) - f(y_2)| \leq \alpha |y_1 - y_2|, \forall y_1, y_2 \in R \quad (17.15)$$

式中 α 是常数。如果开环非强制系统 $\dot{x} = Ax$ 是全局指数稳定的(或等价地, A 的特征值都在复平面的左半开平面), 则如图 17.3 所示的强制闭环系统对于所有 $p \in [1, \infty)$ 都是 L_p -稳定的。

据命题 17.2, 当 A 的特征值在复平面的左半开平面上时, 只要能证明设计的模糊控制器 $f(y)$ 满足 Lipschitz 连续条件(17.15), 就可认为图 17.3 所示的闭环模糊控制系统是 L_p -

稳定的。有趣的是,无论模糊控制器(17.13)是否是根据17.1.1节中的三个步骤设计的,它都满足式(17.15)的Lipschitz连续条件。这里先说明式(17.13)中的模糊控制器 $f(y)$ 是连续有界的分段线性函数,然后由此说明它满足Lipschitz条件。

引理 17.1 式(17.13)中的模糊控制器 $f(y)$ 是连续的、有界的和分段线性的。

证明 令 \bar{x}^l 为图17.2所示的模糊集 A^l 的中心($l=1,2,\dots,2N+1$)。因为图17.2所示的隶属度函数是连续的,所以式(17.13)中的 $f(y)$ 也是连续的。由于 $y \leq \bar{x}^1$ 时 $f(y) = \bar{y}^1$, $y \geq \bar{x}^{2N+1}$ 时 $f(y) = \bar{y}^{2N+1}$,当 $y \in [\bar{x}^l, \bar{x}^{l+1}]$ ($l=1,2,\dots,2N$)时, $f(y)$ 是 \bar{y}^l 和 \bar{y}^{l+1} 的加权平均值,所以可推知, $f(y)$ 是有界的。为了证明 $f(y)$ 是分段线性函数,首先将实轴 R 分段,使

$$R = (-\infty, \bar{x}^1] \cup [\bar{x}^1, \bar{x}^2] \cup \dots \cup [\bar{x}^{2N}, \bar{x}^{2N+1}] \cup [\bar{x}^{2N+1}, \infty)$$

当 $y \in (-\infty, \bar{x}^1]$ 和 $y \in [\bar{x}^{2N+1}, \infty)$ 时,分别有 $f(y) = \bar{y}^1$ 和 $f(y) = \bar{y}^{2N+1}$,均是线性函数。对于 $y \in [\bar{x}^l, \bar{x}^{l+1}]$, $l \in \{1,2,\dots,2N\}$,根据图17.2有

$$\begin{aligned} f(y) &= \frac{\bar{y}^l \mu_{A^l}(y) + \bar{y}^{l+1} \mu_{A^{l+1}}(y)}{\mu_{A^l}(y) + \mu_{A^{l+1}}(y)} \\ &= \frac{\bar{y}^l \frac{\bar{x}^{l+1} - y}{\bar{x}^{l+1} - \bar{x}^l} + \bar{y}^{l+1} \frac{y - \bar{x}^l}{\bar{x}^{l+1} - \bar{x}^l}}{\frac{\bar{x}^{l+1} - y}{\bar{x}^{l+1} - \bar{x}^l} + \frac{y - \bar{x}^l}{\bar{x}^{l+1} - \bar{x}^l}} \\ &= \frac{(\bar{y}^{l+1} - \bar{y}^l)y + \bar{y}^l \bar{x}^{l+1} - \bar{y}^{l+1} \bar{x}^l}{\bar{x}^{l+1} - \bar{x}^l} \end{aligned} \quad (17.16)$$

上式是关于 y 的线性函数。由于 $f(y)$ 连续,所以 $f(y)$ 是分段线性函数。

综合引理17.1和命题17.2,可得如下定理。

定理 17.2 考虑图17.3中的闭环模糊控制系统。假设所设计的模糊控制器 $f(y)$ 形如式(17.13),且 A 的所有特征值都在复平面的左半开平面上。则图17.3中的闭环模糊控制系统对于所有 $p \in [0, \infty)$ 是 L_p -稳定的。

证明: 由于一个连续、有界、分段线性函数满足Lipschitz条件(17.15),所以此定理可由命题17.2和引理17.1得证。

17.2 多输入—多输出系统的稳定模糊控制

17.2.1 指数稳定性

考虑如下多输入—多输出(MIMO)线性系统:

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (17.17)$$

$$y(t) = Cx(t) \quad (17.18)$$

式中, 输入变量 $u(t) \in R^m$, 输出变量 $y(t) \in R^m$, 状态变量 $x(t) \in R^n$ 。假定输入变量数目和输出变量数目相等, 称为方正系统(“squared” systems)。本例中, 控制变量 $u(t) = (u_1(t), \dots, u_m(t))^T$ 由 m 个模糊系统构成, 即

$$u_j(t) = -f_j[y(t)] \quad (17.19)$$

式中, $j=1, 2, \dots, m$, 且 $f_j[y(t)]$ 是 m 维输入 1 维输出系统。此闭环模糊控制系统的结构仍如图 17.1 所示, 不同之处在于, 以矩阵 B 替换了向量 b , 以矩阵 C 替换了向量 c , 以向量函数 $f = (f_1, \dots, f_m)^T$ 替换了标量函数 f 。对于由式(17.18)和式(17.19)组成的 MIMO 系统, 其能控性和能观性的定义是式(17.6)和式(17.7), 只是将其中的 b 换为 B , c 换为 C 而已。定义严格正实域性为: 令 $H(s) = C(sI - A)^{-1}B$ 为传递矩阵, 如果

$$\inf_{\omega \in R} \lambda_{\min}[H(j\omega) + H^*(j\omega)] > 0 \quad (17.20)$$

则称 $H(s)$ 是严格正实域。其中, $*$ 表示共轭转置, $\lambda_{\min}[H(j\omega) + H^*(j\omega)]$ 为矩阵 $H(j\omega) + H^*(j\omega)$ 的最小特征值。

与前面一样, 这里先引述控制理论中的一个结果, 然后设计一个满足假设条件的模糊控制器。下述命题引自 Vidyasagar[1993]。

命题 17.3 考虑式(17.17)~式(17.19)组成的闭环系统, 假定: (a) A 的所有特征值都在复平面的左半开平面上; (b) 系统(17.17)~(17.18)是能控和能观的; (c) 传递矩阵 $H(s) = C(sI - A)^{-1}B$ 是严格正实域。如果控制向量 $f(y)$ 满足 $f(0) = 0$, 且

$$y^T f(y) \geq 0, \forall y \in R^m \quad (17.21)$$

则闭环系统(17.17)~(17.19)的平衡点 $x=0$ 是全局指数稳定的。

需要注意的是, 为了满足式(17.21), m 维模糊系统 $\{f_1(y), \dots, f_m(y)\}$ 不能以各自独立的方式设计。下面将给出使模糊控制器 $f(y) = (f_1(y), \dots, f_m(y))^T$ 满足式(17.21)的 m 维模糊系统的设计方法。

稳定模糊控制器的设计

步骤 1 假定输出变量 $y_i(t)$ 的取值区间为 $U_i = [\alpha_i, \beta_i] \subset R$, 其中 $i=1, 2, \dots, m$ 。在 U_i 上定义 $2N_i + 1$ 个模糊集 $A_i^{l_i}$, 他们均是标准的、一致的和完备的, 并具有如图 17.2 所示的三角形隶属度函数(所有变量均有下角标 i)。

步骤 2 考虑 m 组模糊 IF-THEN 规则, 其中, 第 j 组($j=1, 2, \dots, m$)由下面的 $\prod_{i=1}^m (2N_i + 1)$ 条规则构成。

$$\text{如果 } y_1 \text{ 是 } A_1^{l_1}, \text{ 且 } \dots \text{ 且 } y_m \text{ 是 } A_m^{l_m}, \text{ 则 } u \text{ 是 } B_j^{l_1 \dots l_m} \quad (17.22)$$

其中, $l_i = 1, 2, \dots, 2N_i + 1, i=1, 2, \dots, m$, 模糊集 $B_j^{l_1 \dots l_m}$ 的中心 $\bar{y}^{l_1 \dots l_m}$ 选定为:

$$\bar{y}_j^{l_1 \dots l_m} \begin{cases} \leq 0 & l_j = 1, 2, \dots, N_j \\ = 0 & l_j = N_j + 1 \\ \geq 0 & l_j = N_j + 2, \dots, 2N_j + 1 \end{cases} \quad (17.23)$$

其中,当 $i=1,2,\dots,m, i \neq j$ 时, l_i 可以取 $\{1,2,\dots,2N_i+1\}$ 中的任意值。

步骤3 使用乘积推理机、单值模糊器、中心平均解模糊器,根据式(17.22)中的 $\prod_{i=1}^m (2N_i+1)$ 条规则设计 m 维模糊系统的每个一个分量 $f_j(y)$ 。亦即,设计如下模糊控制器 $f_j(y)$:

$$u_j = -f_j(y) = - \frac{\sum_{l_1=1}^{2N_1+1} \cdots \sum_{l_m=1}^{2N_m+1} \bar{y}_j^{l_1 \cdots l_m} \left(\prod_{i=1}^m \mu_{A_i^{l_i}}(y_i) \right)}{\sum_{l_1=1}^{2N_1+1} \cdots \sum_{l_m=1}^{2N_m+1} \left(\prod_{i=1}^m \mu_{A_i^{l_i}}(y_i) \right)} \quad (17.24)$$

其中, $j=1,2,\dots,m$ 。

和前面的 SISO 系统一样,稳定模糊控制器设计实例的三个步骤也无须知道过程参数 A 、 B 和 C ,只要隶属度函数形如图 17.2 所示且参数 $\bar{y}_j^{l_1 \cdots l_m}$ 满足式(17.23)即可。这些参数的选择有很大的自由度。下面将给出形如式(17.24)的模糊控制器 $u = (u_1, \dots, u_m)^T$ 中的分量 u_j ,它能保证闭环系统的稳定。

定理 17.3 考虑由式(17.17)~式(17.19)构成的闭环模糊控制系统。如果模糊控制器 $u = (u_1, \dots, u_m)^T = (-f_1(y), \dots, -f_m(y))^T$ 是根据以上三个步骤设计的[即 $u_j = -f_j(y)$ 由式(17.24)给出]且所要控制的过程满足命题 17.3 中的条件(a)~(c),则闭环系统的平衡点 $x=0$ 是全局指数稳定的。

证明:如果能说明对于所有 $y \in R^m$ 都有 $f(0)=0$ 和 $y^T f(y) \geq 0$ 成立,其中 $f(y) = (f_1(y), \dots, f_m(y))^T$,则此定理可由命题 17.3 得证。由图 17.2、式(17.23)和式(17.24)可得

$$f_j(0) = \bar{y}_j^{(N_1+1) \cdots (N_m+1)} = 0, j=1,2,\dots,m$$

由于

$$y^T f(y) = y_1 f_1(y) + \cdots + y_m f_m(y)$$

如果能说明对于所有 $y_j \in R(j=1,2,\dots,m)$,都有 $y_j f_j(y) \geq 0$,则可得,对于所有的 $y \in R^m$,都有 $y^T f(y) \geq 0$ 。下面设 $y_j < 0$,则根据图 17.2 有 $\mu_{A_j^{l_j}}(y_j) = 0, l_j = N_j+2, \dots, 2N_j+1$ 。然后,根据式(17.24),可得

$$f_j(y) = \frac{\sum_{l_1=1}^{2N_1+1} \cdots \sum_{l_{j-1}=1}^{2N_{j-1}+1} \sum_{l_j=1}^{N_j+1} \sum_{l_{j+1}=1}^{2N_{j+1}+1} \cdots \sum_{l_m=1}^{2N_m+1} \bar{y}_j^{l_1 \cdots l_m} \left(\prod_{i=1}^m \mu_{A_i^{l_i}}(y_i) \right)}{\sum_{l_1=1}^{2N_1+1} \cdots \sum_{l_{i-1}=1}^{2N_{i-1}+1} \sum_{l_j=1}^{N_j+1} \sum_{l_{j+1}=1}^{2N_{j+1}+1} \cdots \sum_{l_m=1}^{2N_m+1} \left(\prod_{i=1}^m \mu_{A_i^{l_i}}(y_i) \right)} \quad (17.25)$$

根据式(17.23),有 $\bar{y}_j^{l_1 \cdots l_m} \leq 0, l_j = 1,2,\dots,N_j+1$,从而有 $f_j(y) \leq 0$ 和 $y_j f_j(y) \geq 0$ 。类似地,可证明当 $y_j > 0$ 时, $y_j f_j(y) \geq 0$ 。

17.2.2 输入—输出稳定性

再次考虑闭环模糊控制系统(17.17)~(17.19)。与命题 17.2 相似,对于本控制系统有如下与 L_p -稳定性有关的结果。

命题 17.4(Vidgasagar[1993])。考虑闭环控制系统(17.17)~(17.19),并假定开环非强制系统 $\dot{x} = Ax$ 是全局指数稳定的。如果非线性控制器 $f(y) = (f_1(y), \dots, f_m(y))^T$ 是全局 Lipschitz 连续的,即

$$\|f(y_1) - f(y_2)\| \leq \alpha \|y_1 - y_2\|, \forall y_1, y_2 \in R^m \quad (17.26)$$

式中, α 是常数,则称闭环系统(17.17)~(17.19)对于所有的 $p \in [1, \infty)$ 都是 L_p -稳定的。

采用与引理 17.1 相同的推理过程,可以证明模糊系统(17.24)是连续有界的分段线性函数。由于连续有界的分段线性函数向量满足 Lipschitz 条件(17.26),所以根据命题 17.4 可以得到如下结果。

定理 17.4 模糊控制系统(17.17)~(17.19)的模糊控制分量 $f_j(y)$ 由式(17.24)给出,只要 A 的特征值都在复平面的左半开平面上,就称该系统就是 L_p -稳定的,其中 $p \in [1, \infty)$ 。

17.3 本章小结及导读

本章主要阐述了以下内容:

- 带有线性过程—非线性控制器的经典控制系统(命题 17.1~命题 17.4);
- SISO 和 MIMO 线性系统的稳定(指数稳定或输入—输出稳定)模糊控制器的设计,及所设计的模糊控制器的性质证明。

Vidyassagar[1993]、Slotine 和 Li[1991]是两本非线性控制方面的优秀教材。前者给出了有关非线性控制的严谨论述,后者则侧重于可读性,方便应用人员阅读。将非线性控制理论的结论用于模糊稳定控制器的先驱者是 Langari 和 Tomizuka[1990], Chiu, Chand, Moor 和 Chaudhary[1991]也给出了关于线性被控对象的稳定模糊控制的充分条件。本章中的方法是非常初步的,但在其他文献中尚未见到。

17.4 习题

习题 17.1 假定 $\Phi: R^m \rightarrow R^m$, 常数 $a, b \in R$, 且 $a < b$ 。若(a): $\Phi(0) = 0$, 且(b):

$$[\Phi(y) - ay]^T [by - \Phi(y)] \geq 0, \forall y \in R^m \quad (17.27)$$

则称 Φ 属于区间 $[a, b]$ 。试设计一个 SISO 模糊系统,使其中心平均解模糊器属于区间

$[a, b]$ 。

习题 17.2 设计一个 SISO 模糊系统,使其最大值解模糊器属于区间 $[a, b]$ 。

习题 17.3 设计一个二维输入二维输出的模糊系统(即两个二维输入一维输出的模糊系统),使其中心平均解模糊器属于区间 $[a, b]$ 。

习题 17.4 设计一个二维输入二维输出的模糊系统,使其最大值解模糊器属于区间 $[a, b]$ 。

习题 17.5 证明系统

$$\dot{x} = -x^2 \quad (17.28)$$

中的平衡点 $x=0$ 是渐近稳定的,但不是指数稳定的。并探讨系统

$$\dot{x} = -(1 + \sin^2 x)x \quad (17.29)$$

在平衡点 $x=0$ 的稳定性。

习题 17.6 假定有常数 $a, b, c, r > 0, p \geq 1$, 连续函数 $V: R^n \rightarrow R$, 使得

$$a \|x\|^p \leq V(x) \leq b \|x\|^p \quad (17.30)$$

$$\dot{V}(x) \leq c \|x\|^p \quad (17.31)$$

证明平衡点 $x=0$ 是指数稳定的。

习题 17.7 模仿 17.1 节中模糊控制器的设计方法,为具有传递函数

$$h(s) = \frac{1}{(s+1)(s+2)(s+3)} \quad (17.32)$$

的线性系统设计模糊控制器。自行选择模糊控制器的参数,要求满足设计步骤中的条件。在不同初始条件下画出闭环系统的输出的图形。

线性系统模糊控制 II： 最优控制器与稳健控制器

第 17 章给出了闭环模糊控制系统稳定的条件。通常情况下,需要确定模糊控制器的参数的取值范围,从而当参数在规定范围内取值时,系统的稳定性能得到保证。但我们还没有说明如何在取值范围内选取参数。所以在本章中,首要说明的是如何选取特殊的模糊控制器参数以使某个性能指标得到优化;也就是先考虑如何设计线性系统的最优模糊控制器。这比设计单纯稳定的模糊控制器要困难得多。本章介绍的方法是很初步的。在设计最优模糊控制器的过程中,必须给出被控对象参数 A 、 B 和 C 的取值,而这在 17 章里的纯稳定模糊控制器的设计中是无须给出的。

本章的第二个主题是线性系统的稳健模糊控制。此领域中有很多尚未解决的问题,这里只说明用小增益定理设计稳健模糊控制器的一些最基本思想。

18.1 最优模糊控制

这里首先简要回顾一下求解最优控制问题的 Pontryagin 最小值原理。然后,考虑控制器为模糊系统时,怎样确定实现最佳性能的控制器的参数。

18.1.1 Pontryagin 最小值原理

考虑系统

$$\dot{x}(t) = g[x(t), u(t)] \quad (18.1)$$

其初始条件为 $x(0) = x_0$ 。其中, $x \in R^n$ 是状态变量, $u \in R^m$ 是控制变量, g 是线性或非线性函数。系统(18.1)的最优控制问题为确定控制 $u(t)$ 使性能指标

$$J = S[x(T)] + \int_0^T L[x(t), u(t)] dt \quad (18.2)$$

达到最小,其中, S 和 L 是给定函数,终止时间 T 可能也是给定的。

此最优控制问题可利用 Pontryagin 最小值原理求解,方法如下:

首先,定义哈密顿函数

$$H(x, u, p) = L[x(t), u(t)] + p^T(t)g[x(t), u(t)] \quad (18.3)$$

寻找 $u = h(x, p)$, 使 $H(x, u, p)$ 达到最小。将 $u = h(x, p)$ 代入式(18.3)并定义

$$H^*(x, p) = H[x, h(x, p), p] \quad (18.4)$$

然后, 求解 $2n$ 个偏微分方程组成的方程组(已给出两点的边值条件)

$$\dot{x}(t) = \frac{\partial H^*}{\partial p}, x(0) = x_0 \quad (18.5)$$

$$\dot{p}(t) = -\frac{\partial H^*}{\partial x}, p(T) = \frac{\partial S}{\partial x} \Big|_{x(T)} \quad (18.6)$$

再令 $x^*(t)$ 和 $p^*(t)$ 分别为式(18.5)和式(18.6)的解[称 $x^*(t)$ 和 $p^*(t)$ 为最优轨迹], 最后可得最优控制为

$$u^*(t) = h[x^*(t), p^*(t)] \quad (18.7)$$

18.1.2 最优模糊控制器的设计

假定所要控制的系统为时不变线性系统

$$\dot{x}(t) = Ax(t) + Bu(t), x(0) = x_0 \quad (18.8)$$

式中, $x \in R^n, u \in R^m$, 二次型性能指标为

$$J = x^T(T)Mx(T) + \int_0^T [x^T(t)Qx(t) + u^T(t)Ru(t)]dt \quad (18.9)$$

其中矩阵 $M \in R^{n \times n}, Q \in R^{n \times n}, R \in R^{m \times m}$, 都是对称正定方阵。

现假定控制器 $u(t)$ 是形如式(17.24)的模糊系统, 只是将式(17.24)中的输出变量 y 改为状态变量 x , 即 $u(t) = (u_1, \dots, u_m)^T$, 其中

$$u_j = -f_j(x) = -\frac{\sum_{l_1=1}^{2N_1+1} \cdots \sum_{l_n=1}^{2N_n+1} \bar{y}_j^{l_1 \cdots l_n} \left(\prod_{i=1}^n \mu_{A_i}^{l_i}(x_i) \right)}{\sum_{l_1=1}^{2N_1+1} \cdots \sum_{l_n=1}^{2N_n+1} \left(\prod_{i=1}^n \mu_{A_i}^{l_i}(x_i) \right)} \quad (18.10)$$

假设隶属度函数是固定的, 形如图 17.2 所示或其他形式。下面的任务是确定参数 $\bar{y}_j^{l_1 \cdots l_n}$ 使式(18.9)中的 J 达到极小。

定义模糊基函数 $b(x) = (b_1(x), \dots, b_N(x))^T$ 为

$$b_l(x) = \frac{\prod_{i=1}^n \mu_{A_i}^{l_i}(x_i)}{\sum_{l_1=1}^{2N_1+1} \cdots \sum_{l_n=1}^{2N_n+1} \left(\prod_{i=1}^n \mu_{A_i}^{l_i}(x_i) \right)} \quad (18.11)$$

式中, $l_i = 1, 2, \dots, 2N_i + 1, l = 1, 2, \dots, N, N = \prod_{i=1}^n (2N_i + 1)$ 。定义参数矩阵 $\Theta \in R^{m \times N}$ 为

$$\Theta = \begin{bmatrix} -\theta_1^T \\ \cdots \\ -\theta_m^T \end{bmatrix} \quad (18.12)$$

式中, $\Theta_j^T \in R^{1 \times N}$, 它包含 N 个参数 $\bar{y}_j^{l_1 \cdots l_n}$, $l_i = 1, 2, \cdots, 2N_i + 1$, 顺序与 $b_l(x)$ 相同, $l = 1, 2, \cdots, N$ 。利用这些符号将模糊控制器 $u = (u_1, \cdots, u_m)^T = (-f_1(x), \cdots, -f_n(x))^T$ 改写为

$$u = \Theta b(x) \quad (18.13)$$

为了实现最优化, 假定参数矩阵 Θ 是时变的, 即 $\Theta = \Theta(t)$ 。

将式(18.13)代入式(18.8)和式(18.9), 得到闭环系统

$$\dot{x}(t) = Ax(t) + B\Theta(t)b[x(t)] \quad (18.14)$$

其相应的性能指标为

$$J = x^T(T)Mx(T) + \int_0^T [x^T(t)Qx(t) + b^T(x(t))\Theta^T(t)R\Theta(t)b(x(t))] dt \quad (18.15)$$

于是, 设计最优模糊控制器问题就变成了确定 $\Theta(t)$ 的最优值, 使式(18.15)中的 J 达到最小值的问题。将 $\Theta(t)$ 视为 Pontryagin 最小值原理中的控制 $u(t)$, 则可根据式(18.3) ~ 式(18.7)确定最优的 $\Theta(t)$ 值。具体地, 定义哈密顿函数

$$H(x, \Theta, p) = x^T Qx + b^T(x)\Theta^T R\Theta b(x) + p^T [Ax + B\Theta b(x)] \quad (18.16)$$

由 $\frac{\partial H}{\partial \Theta} = 0$, 即

$$\frac{\partial H}{\partial \Theta} = 2R\Theta b(x)b^T(x) + B^T p b^T(x) = 0 \quad (18.17)$$

可得

$$\Theta = -\frac{1}{2}R^{-1}B^T p b^T(x)[b(x)b^T(x)]^{-1} \quad (18.18)$$

将式(18.18)代入式(18.16), 可得

$$\begin{aligned} H^*(x, p) &= x^T Qx + p^T Ax + \frac{1}{4}b^T(x)[b(x)b^T(x)]^{-1}b(x)p^T B R^{-1} B^T p b^T(x) \\ &\quad [b(x)b^T(x)]^{-1}b(x) - \frac{1}{2}p^T B R^{-1} B^T p b^T(x)[b(x)b^T(x)]^{-1}b(x) \\ &= x^T Qx + p^T Ax + [\alpha^2(x) - \alpha(x)]p^T B R^{-1} B^T p \end{aligned} \quad (18.19)$$

式中, $\alpha(x)$ 的定义为

$$\alpha(x) = \frac{1}{2}b^T(x)[b(x)b^T(x)]^{-1}b(x) \quad (18.20)$$

将此 H^* 代入式(18.5)和式(18.6), 可得

$$\dot{x} = \frac{\partial H^*}{\partial p} = Ax + 2[\alpha^2(x) - \alpha(x)]B R^{-1} B^T p \quad (18.21)$$

$$p = -\frac{\partial H^*}{\partial x} = -2Qx - A^T p - [2\alpha(x) - 1] \frac{\partial \alpha(x)}{\partial x} p^T B R^{-1} B^T p \quad (18.22)$$

且满足初始条件 $x(0) = x_0$ 和边界条件 $p(T) = 2Mx(T)$, 令 $x^*(t), p^*(t) (t \in [0, T])$ 分别为式(18.21)和式(18.22)的解, 则最优模糊控制器的参数为

$$\Theta^*(t) = -\frac{1}{2} R^{-1} B^T p^*(t) b^T(x^*(t)) [b(x^*(t)) b^T(x^*(t))]^{-1} \quad (18.23)$$

最优模糊控制器为

$$u^* = \Theta^*(t) b(x) \quad (18.24)$$

值得注意的是, 式(18.24)中的最优模糊控制器是具有时变系数的状态反馈控制器。此最优模糊控制器的设计过程归纳如下:

最优模糊控制器的设计。

步骤1 在整个状态空间上给出隶属度函数 $\mu_{A_i^{l_i}}(x_i)$, 其中, $l_i = 1, 2, \dots, 2N_i + 1, i = 1, 2, \dots, n$ 。因为图 17.2 所示的隶属度函数将使函数 $\alpha(x)$ 不可微[式(18.22)要求 $\frac{\partial \alpha(x)}{\partial x}$ 存在], 所以这里不能把隶属度函数选择为图 17.2 所示的形式。这里将 $\mu_{A_i^{l_i}}(x_i)$ 选为高斯函数。

步骤2 根据式(18.11)计算模糊基函数 $b_l(x)$, 根据式(18.20)计算 $\alpha(x)$, 并计算相应的偏导数 $\frac{\partial \alpha(x)}{\partial x}$ 。

步骤3 求解偏微分方程组(18.21)和(18.22)的两点的边值, 并令解为 $x^*(t)$ 和 $p^*(t), t \in [0, T]$ 。根据式(18.23)计算 $\Theta^*(t), t \in [0, T]$ 。

步骤4 得到最优模糊控制器(18.24)。

需要说明的是, 这里步骤 1 至步骤 3 是离线操作, 即先按照步骤 1 至步骤 3 计算 $\Theta^*(t)$, 并把 $t \in [0, T]$ 时的 $\Theta^*(t)$ 存储在计算机里, 然后, 当在线操作时, 将储存的 $\Theta^*(t)$ 代入式(18.24)即可得到最优模糊控制器。

在设计最优模糊控制器时, 最困难的是求解偏微分方程组(18.21)和(18.22)的两点的边值。由于这些偏微分方程是非线性的, 所以通常需要用数值积分才能求解。

18.1.3 在球—横木系统中的应用

考虑图 16.7 所示的球—横木系统, 状态空间模型由式(16.4)~式(16.5)给出。由于球—横木系统是非线性的, 可以为了应用最优模糊控制器, 在平衡点 $x = 0$ 附近将系统线性化。则线性化后的系统形如式(18.8), 且

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -\alpha\beta & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (18.25)$$

下面为线性化的球—横木系统设计一个最优模糊控制器,并把设计的控制器应用于初始的非线性系统(16.4)~(16.5)。选定 $M=0, Q=I, R=I, T=30, N_i=2, i=1,2,3,4$ 。隶属度函数 $\mu_{A_i^{l_i}}(x_i)$ 选为

$$\mu_{A_i^{l_i}}(x_i) = \exp[-2(x_i - \bar{x}_i^{l_i})^2] \quad (18.26)$$

式中, $i=1,2,3,4, l_i=1,2,\dots,5$, 且 $\bar{x}_i^{l_i} = a_i + b_i(l_i - 1)$, $a_1 = a_2 = -2, a_3 = a_4 = -1, b_1 = b_2 = 1, b_3 = b_4 = 0.5$ 。图 18.1 给出了三种初始条件: $x(0) = (1,0,0,0)^T, (2,0,0,0)^T$ 和 $(3,0,0,0)^T$ 下的闭环系统的输出 $y(t)$ 。从图 18.1 可以看出,虽然这里的最优模糊控制器是基于线性化模型设计的,但它还是能把球从一系列初始位置平滑地控制到原点。

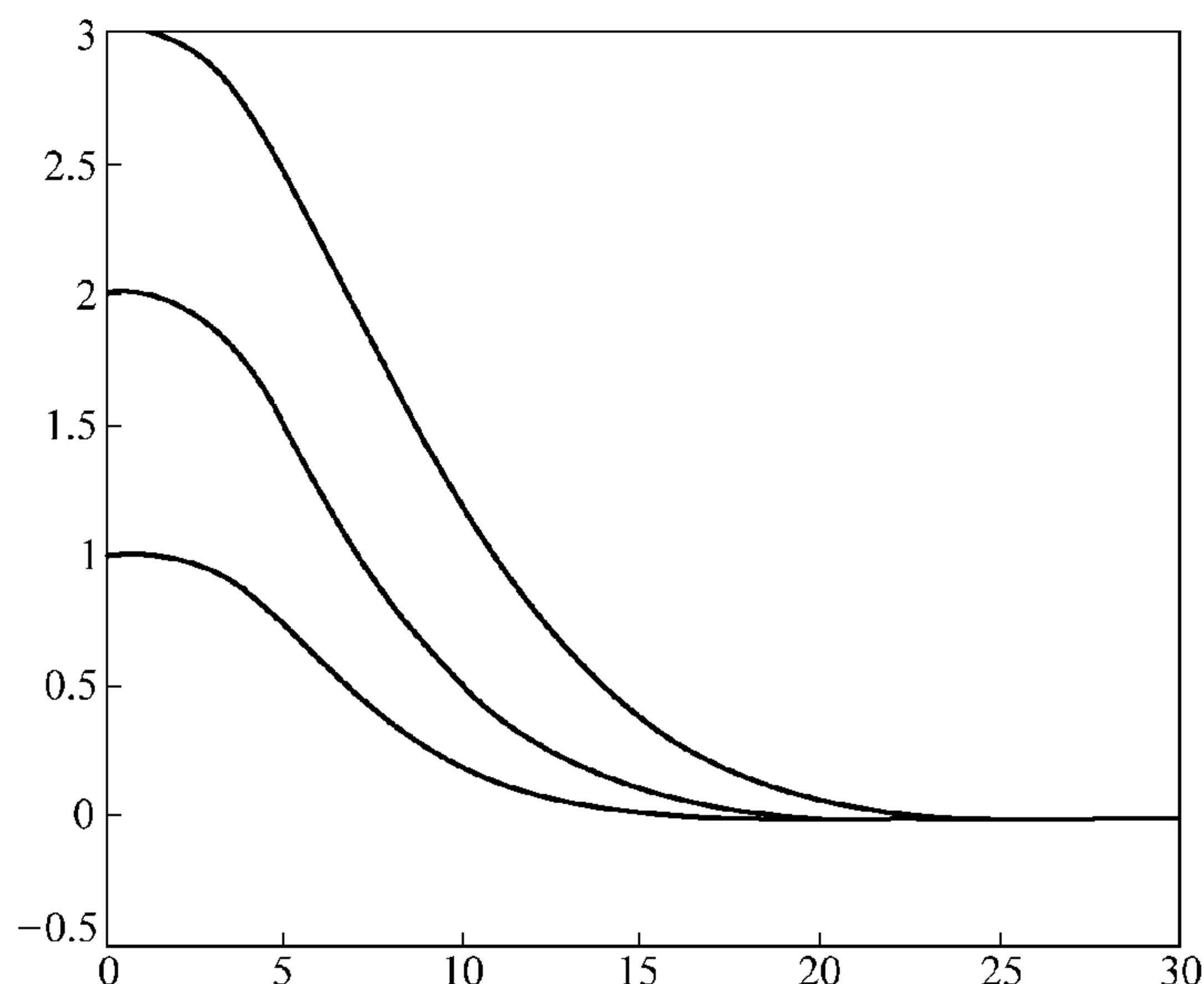


图 18.1 三种初始条件下带有最优模糊控制器的闭环系统的输出 $y(t)$ ($=r(t)$, 球的位置)

18.2 稳健模糊控制

几乎在所有的实际系统中,尤其是在工业系统中,从来不能准确地获得被控对象的模型。由于控制器是为模型设计的,理论上和仿真中(采用了模型)的效果好并不意味着实际应用中效果好(此处,控制器被直接用于现实系统,而不是模型)。稳健控制的目标是为相当广泛的被控对象群设计出一个能够保证闭环系统稳定性的控制器。被控对象的范围越

广,则控制器的稳健性越强。下面只考虑线性被控对象的稳健模糊控制器的设计。

考虑如图 17.1 所示的模糊控制系统,其中所要控制过程以传递函数 $G(s)$ 表征。为简便起见,这里假定 $G(s)$ 是 SISO 系统。定义

$$\gamma(G) = \sup_{\omega \in R} |G(j\omega)| \quad (18.27)$$

且

$$\Omega_\alpha = \{G(s): \gamma(G) \leq \alpha\} \quad (18.28)$$

显然, α 越大, Ω_α 中包含的被控对象越多。稳健模糊控制的目标是,在允许 Ω_α 中包含的被控对象尽可能多的情况下,设计一个使 Ω_α 内的所有被控对象都稳定受控的模糊控制器。该含义是很直观的,因为 Ω_α 越大,就有越多的被控对象可由模糊控制器控制其稳定运行,也就意味着模糊控制器的稳健性越好。

实现此目标的主要的数学工具就是有名的小增益定理(Vidyasagar[1993])。为了理解此定理,这里首先引进非线性映射增益的概念。

定义 18.1 令 $f: R \rightarrow R$ 对于某些 γ 和全部 $x \in R$ 满足 $|f(x)| \leq \gamma|x|$ 。然后定义增益 f 如下:

$$\gamma(f) = \inf \{\gamma: |f(x)| \leq \gamma|x|, \forall x \in R\} \quad (18.29)$$

小增益定理表述如下:

小增益定理。假定线性对象 $G(s)$ 是稳定的,非线性映射 $f(y)$ 在输入有界时是有界的。则图 17.1 所示的闭环系统[所要控制过程以 $G(s)$ 表征]在满足如下条件时是稳定的:

$$\gamma(G)\gamma(f) < 1 \quad (18.30)$$

由式(18.30)可知, $\gamma(f)$ 的值越小, $\gamma(G)$ 的允许值越大。根据式(18.27)~式(18.28), $\gamma(G)$ 的值越大意味着控制器 $f(y)$ 能稳定化的被控对象集 Ω_α 越大。这样,为了实现稳健控制的目的,应选择尽可能小的 $\gamma(f)$ 。极端的情况是 $f=0$,这表明根本不需要控制。但这是不切实际的,因为需要有反馈控制以改善性能指标。由于最常用的模糊控制器是 \bar{y}^l (THEN 部分隶属度函数的中心值)的加权平均,所以这里建议按如下方法设计稳健模糊控制器:用第 16 章至第 26 章中的任何方法设计模糊控制器,但在设计中尽量采用小一点的 \bar{y}^l 。

稳健模糊控制是一个广阔的研究领域,仍有许多问题有待研究。

18.3 本章小结及导读

本章主要阐述了以下内容:

- 求解最优控制问题的 Pontryagin 最小值原理;
- 如何用 Pontryagin 最小值原理设计线性被控对象的最优模糊控制器;
- 小增益定理和设计稳健模糊控制器的基本思想。

在最优控制方面的好教科书很多,如 Anderson 和 Moore[1990], Bryson 和 Ho[1975]。

Green 和 Limebeer[1995]也是一本稳健控制方面的好书。

18.4 习题

习题 18.1 汽车线性运动的一个简化模型是 $\dot{x} = u$, 其中 $x(t)$ 是汽车的速度, $u(t)$ 是加速度或减速度。汽车的初速度为 x_0 米/分。利用 Pontryagin 最小值原理设计一个最优的 $u(t)$, 使速度 $x(t_f)$ 在最少时间 t_f 内达到零值。设时刻 t 加速或减速的约束为 $|u(t)| \leq M$, 其中 M 是一个常数。

习题 18.2 利用 Pontryagin 最小值原理解最优化问题:

$$(a) \dot{x}_1 = x_2, \dot{x}_2 = u, x_1(0) = 1, x_2(0) = 1, x_1(2) = 0, x_2(2) = 0,$$

$$\text{使 } J = \frac{1}{2} \int_0^2 u^2(t) dt \text{ 最小。}$$

$$(b) \dot{x} = u, x(0) = 0, x(1) = 1, \text{使 } J = \int_0^1 (x^2 + u^2) dt \text{ 最小。}$$

习题 18.3 写出 18.1.3 小节的例子中微分方程(18.21)~(18.22)的详细表达式 ($\frac{\partial \alpha(x)}{\partial x}$ 的详细表达式可写可不写)。

习题 18.4 在 $U = [-1, 1]$ 上利用中心平均解模糊器设计一个至少有五条规则的模糊系统 $f(x)$, 使得: (a) $f(0) = 0$; (b) $\gamma(f) = 1$ 。其中式(18.29)中的 R 变为 U 。

习题 18.5 在 $U = [-1, 1]$ 上利用最大值解模糊器设计一个至少带有五条规则的模糊系统 $f(x)$, 使得: (a) $f(0) = 0$; (b) $\gamma(f) = 1$ 。其中式(18.29)中的 R 变为 U 。

非线性系统模糊控制 I：滑动控制

滑动控制对非线性和不确定系统来说是一种有效的控制方法。此方法是一种稳健性好的控制方法,适用于模型具有不确定性、参数处于扰动状态的情况,其条件是这些不确定性和扰动的边界值是已知的。通过认真对比滑动控制和模糊控制可知,二者在很多地方具有相似性。本章将揭示模糊控制和滑动控制间的关系,并根据滑动控制的基本原理设计模糊控制器。

19.1 作为滑动控制的模糊控制：分析

19.1.1 滑动控制的基本原理

考虑单输入—单输出(SISO)非线性系统

$$x^{(n)} = f(\mathbf{x}) + u \quad (19.1)$$

式中, $u \in R$ 是输入控制, $x \in R$ 是输出控制, $\mathbf{x} = (x, \dot{x}, \dots, x^{(n-1)})^T \in R^n$ 是状态变量。在式(19.1)中,函数 $f(\mathbf{x})$ 是隐函数,但函数 $f(\mathbf{x})$ 的不确定性可由 \mathbf{x} 的一个已知函数来界定,即

$$f(\mathbf{x}) = \hat{f}(\mathbf{x}) + \Delta f(\mathbf{x}) \quad (19.2)$$

和

$$|\Delta f(\mathbf{x})| \leq F(\mathbf{x}) \quad (19.3)$$

其中, $\Delta f(\mathbf{x})$ 未知,而 $\hat{f}(\mathbf{x})$ 和 $f(\mathbf{x})$ 均已知。控制目标是确定状态反馈控制 $u = u(\mathbf{x})$,使闭环系统的状态变量 \mathbf{x} 达到要求的状态 $\mathbf{x}_d = (x_d, \dot{x}_d, \dots, x_d^{(n-1)})^T$ 。即轨迹误差

$$\mathbf{e} = \mathbf{x} - \mathbf{x}_d = (e, \dot{e}, \dots, e^{(n-1)})^T \quad (19.4)$$

应收敛于零,其中 $e = x - x_d$ 。

滑动控制的基本思想如下。定义标量函数

$$s(\mathbf{x}, t) = \left(\frac{d}{dt} + \lambda \right)^{n-1} e = e^{n-1} + C_{n-1}^1 \lambda e^{(n-2)} + C_{n-1}^2 \lambda^2 e^{(n-3)} + \dots + \lambda^{n-1} e \quad (19.5)$$

式中, λ 是正常数。则

$$s(\mathbf{x}, t) = 0 \quad (19.6)$$

定义状态空间 R^n 中的时变表面函数 $S(t)$ 。如, $n=2$ 则表面函数 $S(t)$ 为

$$s(\mathbf{x}, t) = \dot{e} + \lambda e = \dot{x} + \lambda x - \dot{x}_d - \lambda x_d = 0 \quad (19.7)$$

如图 19.1 所示,它在 $x - \dot{x}$ 的相平面图中是一条直线。由于 \dot{x}_d 和 x_d 通常是时变函数,所以 $S(t)$ 也是时变的。如果初始状态 $\mathbf{x}(0)$ 等于初始理想状态 $\mathbf{x}_d(0)$,即设 $\mathbf{e}(0) = 0$,则由式(19.5)和式(19.6)可知,如果对于所有 $t \geq 0$ 状态向量 \mathbf{x} 都处于表面 $S(t)$ 上,则对于所有 $t \geq 0$ 都有 $\mathbf{e}(t) = 0$ 。事实上, $s(\mathbf{x}, t) = 0$ 表示,初始条件为 $\mathbf{e}(0) = 0$ 时,线性微分方程有惟一解 $\mathbf{e}(t) = 0$ 。因此,跟踪控制问题与将标量函数 $s(\mathbf{x}, t)$ 保持在零点是等价的。为达此目标,先选择控制 u 使得当状态处于 $S(t)$ 以外时,有

$$\frac{1}{2} \frac{d}{dt} s^2 \leq -\eta |s| \quad (19.8)$$

式中, η 是正常数。式(19.8)称为滑动条件,它保证了 $|s(\mathbf{x}, t)|$ 在 \mathbf{x} 不在表面 $S(t)$ 上时递减。亦即,状态轨迹将如图 19.1 所示向表面 $S(t)$ 移动。表面 $S(t)$ 称为滑动表面,在此表面上的系统称为滑动模式,且确保式(19.8)成立的控制称为滑动模式控制或滑动控制。为总结以上讨论,这里给出如下引理:

引理 19.1 考虑非线性系统(19.1),令 $s(\mathbf{x}, t)$ 定义如式(19.5)。如果能设计控制器 u 使滑动条件(19.8)得到满足,则

- (a) 状态将在有限时间内达到滑动表面 $S(t)$;
- (b) 一旦状态在滑动表面上,则它将停留在那里;
- (c) 如果状态停留在滑动表面上,则跟踪误差 $\mathbf{e}(t)$ 收敛于零。

因此,目标变为设计控制器 u 使闭环系统满足滑动条件(19.8)。

下面给出二阶系统滑动控制的详细过程,即 $n = 2$ 时的系统(19.1)。在此例中,式(19.8)变为

$$s[f(x) + u - \ddot{x}_d + \lambda \dot{e}] \leq -\eta |s| \quad (19.9)$$

这里用到了式(19.5)和式(19.1)。如果选择

$$u = -\hat{f}(\mathbf{x}) + \ddot{x}_d - \lambda \dot{e} - K(x, \dot{x}) \operatorname{sgn}(s) \quad (19.10)$$

则式(19.9)变为

$$\operatorname{sgn}(s)[f(\mathbf{x}) - \hat{f}(\mathbf{x}) - K(x, \dot{x}) \operatorname{sgn}(s)] \leq -\eta \quad (19.11)$$

式中, $s > 0$ 时, $\operatorname{sgn}(s) = 1$, $s < 0$ 时, $\operatorname{sgn}(s) = -1$,且同式(19.2)中一样, $\hat{f}(\mathbf{x})$ 是对 $f(\mathbf{x})$ 的一个估计。而且式(19.11)等价于

$$K(x, \dot{x}) \geq \eta + \operatorname{sgn}(s)[\Delta f(\mathbf{x})] \quad (19.12)$$

因此是,如果选择

$$K(x, \dot{x}) = \eta + F(\mathbf{x}) \quad (19.13)$$

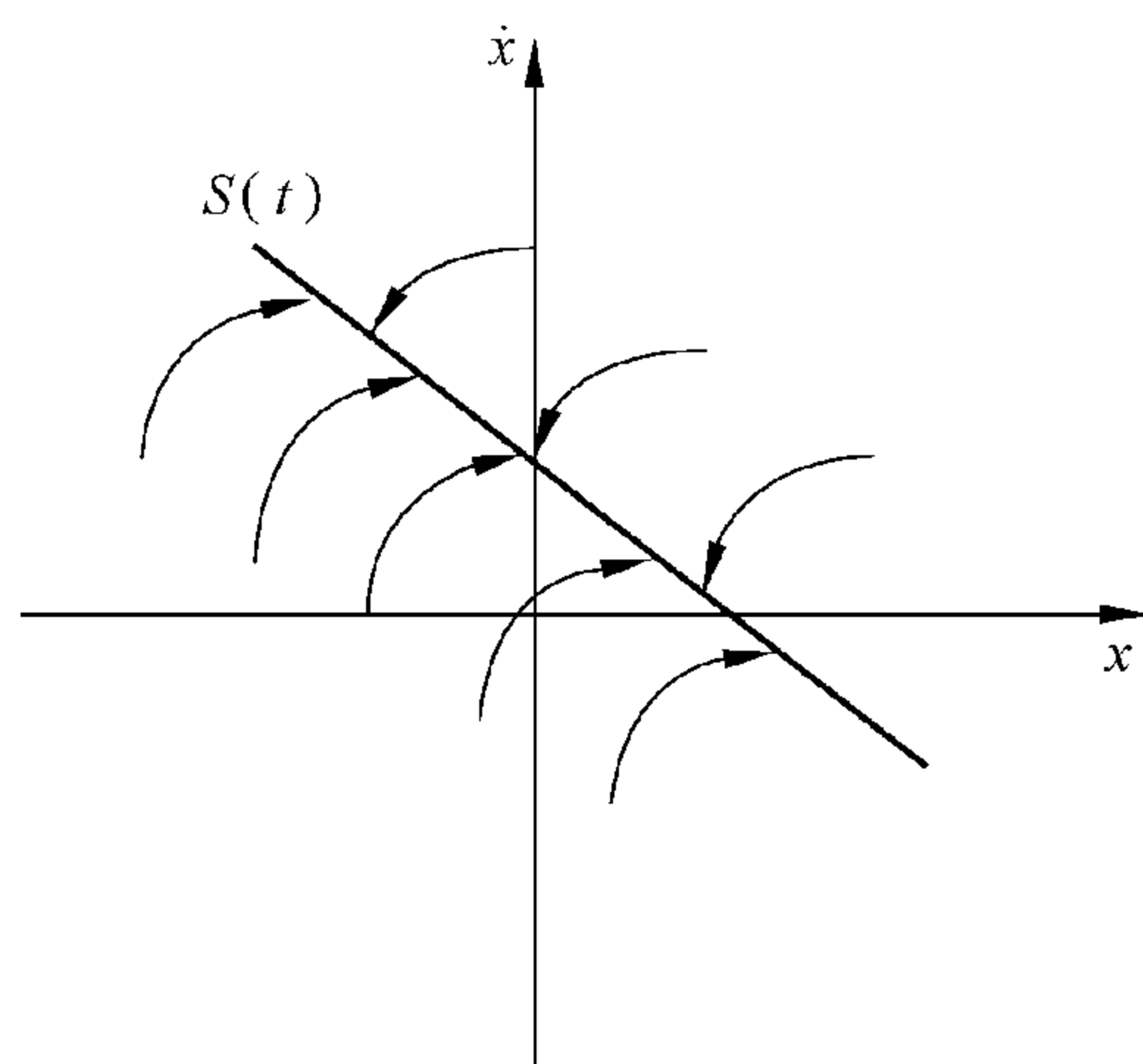


图 19.1 二维相平面的滑动表面

则由式(19.3)可知,式(19.12)是能得到满足,这反过来又说明滑动条件(19.8)也是满足的。总之,滑动控制由式(19.10)给出,其中 $K(x, \dot{x})$ 由式(19.13)给出。

19.1.2 根据滑动控制原理分析模糊控制器

现在使用滑动控制原理来分析模糊控制器。为简单起见,这里考虑 $n = 2$ 时的系统(19.1)。但此方法也可用于高阶系统。假设将控制 u 选为模糊控制 $u_{fuzz}(\mathbf{x})$, 即

$$u = u_{fuzz}(\mathbf{x}) \quad (19.14)$$

下面的定理给出了使跟踪误差 \mathbf{e} 收敛于零的模糊控制 $u_{fuzz}(\mathbf{x})$ 的条件。

定理 19.1 考虑 $n = 2$ 时的非线性系统(19.1), 设控制 u 由式(19.14)给出, 如果模糊控制器 $u_{fuzz}(\mathbf{x})$ 满足下面条件:

$$u_{fuzz}(\mathbf{x}) \leq -\eta - [f(\mathbf{x}) + \lambda \dot{e} - \ddot{x}_d], \text{sgn}(s) > 0 \quad (19.15)$$

$$u_{fuzz}(\mathbf{x}) \geq \eta - [f(\mathbf{x}) + \lambda \dot{e} - \ddot{x}_d], \text{sgn}(s) < 0 \quad (19.16)$$

式中, η 和 λ 是正常数, 且 $s = \dot{e} + \lambda e$, 则此控制保证了跟踪误差 $\mathbf{e} = (x - x_d, \dot{x} - \dot{x}_d)^T$ 收敛于零。

证明 将式(19.14)代入滑动条件(19.9), 有

$$\text{sgn}(s)[f(\mathbf{x}) + u_{fuzz}(\mathbf{x}) - \ddot{x}_d + \lambda \dot{e}] \leq -\eta \quad (19.17)$$

显然, 如果 $u_{fuzz}(\mathbf{x})$ 满足式(19.15)和式(19.16), 则式(19.17)成立, 这表明滑动条件(19.8)是满足的, 从而根据引理 19.1 可知跟踪误差收敛于零。

由定理 19.1 可知, 如果设计的模糊控制器满足式(19.15)和式(19.16), 则能保证跟踪误差收敛于零。然而, 因为模糊控制器 $u_{fuzz}(\mathbf{x})$ 必须不连续地跨过滑动面 $s = 0$, 所以使用式(19.15)和式(19.16)作为约束条件设计模糊控制器是困难的。因此, 定理 19.1 与其说是具有设计价值, 还不如说是具有分析价值。此外, 不连续控制(19.10)或(19.15)~(19.16)将产生跨越滑动面的振荡(chattering)(见下节), 这是不希望看到的。下节将基于滑动控制原理在另一条件下设计模糊控制器。

19.2 作为滑动控制的模糊控制: 设计

19.2.1 滑动控制律的连续逼近

由上节可知, 滑动控制[如式(19.10)]不得不非连续地横穿滑动面 $S(t)$ 。由于采用开关控制不是理想的手段, 所以需要在数字控制系统中给出信号采样, 这样就导致了振荡的产生, 如图 19.2 所示。振荡是不希望看到的, 因为这将涉及高强度的控制活动, 而且可能激发高频动态性能。去除振荡的办法是在滑动面邻近引入一个窄的边界层

$$B(t) = \{ \mathbf{x} : |s(\mathbf{x}, t)| \leq \Phi \} \quad (19.18)$$

这样控制就可以在此边界层内连续变化了(见图 19.2)。 Φ 称为边界层的厚度, $\varepsilon = \Phi/\lambda^{n-1}$ 称为边界层的宽度。下面说明, 如果此控制律保证在边界层 $B(t)$ 外部时滑动条件(19.8)是满足的, 则跟踪误差一定在精确度 ε 的范围内。

引理 19.2 如果外滑动条件(19.8)在(19.18)的边界层 $B(t)$ 外是满足的, 则在有限时间内一定有

$$|e(t)| \leq \varepsilon \quad (19.19)$$

证明 由于在 $B(t)$ 之外, 滑动条件(19.8)是满足的, 所以状态向量 \mathbf{x} 将在有限时间内进入边界层, 且进入之后便停留在里面。亦即, 无论初始状态 $\mathbf{x}(0)$ 的位置如何, 在有限时间之后, 可有 $|s(\mathbf{x}, t)| \leq \Phi$ 。令

$$y_i(t) = \left(\frac{d}{dt} + \lambda \right)^{n-i-1} e \quad (19.20)$$

则有

$$y_0(t) = s(\mathbf{x}, t) \quad (19.21)$$

$$y_{n-1}(t) = e(t) \quad (19.22)$$

和

$$y_i(t) = \left(\frac{d}{dt} + \lambda \right) y_{i+1}(t) \quad (19.23)$$

式中, $i = 0, 1, \dots, n-2$ 。于是有

$$y_{i+1}(t) = \int_0^t e^{-\lambda(t-\tau)} y_i(\tau) d\tau \quad (19.24)$$

当 $i=0$ 时, 由式(19.24)、式(19.21)和 $|s(\mathbf{x}, t)| \leq \Phi$ 有

$$|y_1(t)| \leq \int_0^t e^{-\lambda(t-\tau)} |s(\mathbf{x}, \tau)| d\tau \leq \Phi \frac{1 - e^{-\lambda t}}{\lambda} \leq \Phi/\lambda \quad (19.25)$$

当 $i=1$ 时, 有

$$|y_2(t)| \leq \int_0^t e^{-\lambda(t-\tau)} |y_1(\tau)| d\tau \leq (\Phi/\lambda) \frac{1 - e^{-\lambda t}}{\lambda} \leq \Phi/\lambda^2 \quad (19.26)$$

重复上述步骤直到 $i = n-2$, 有

$$|y_{n-1}(t)| = |e(t)| \leq \Phi/\lambda^{n-1} = \varepsilon \quad (19.27)$$

上式即为式(19.19)。

引理 19.2 说明了, 如果愿意牺牲精度, 即从精确跟踪 $e(t) = 0$ 过渡到跟踪精度为 $|e(t)| \leq \varepsilon$, 则控制律被简化了, 从原来要求在所有时间内满足滑动条件(19.8)变为只有在 $\mathbf{x}(t)$ 处于边界层 $B(t)$ 之外时满足滑动条件。因此, 可以设计一个平滑控制器, 不再需要不

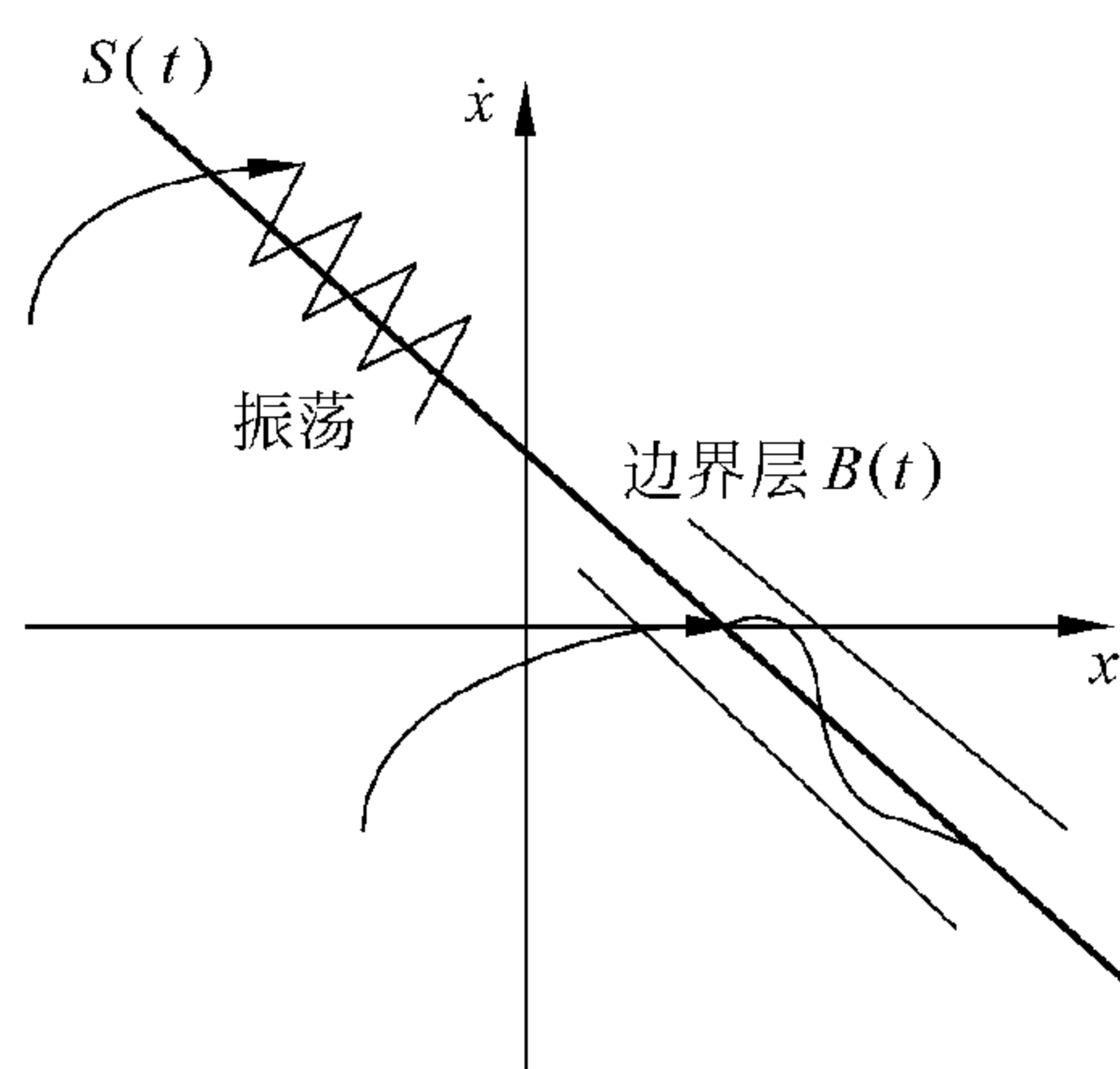


图 19.2 振荡和边界层

连续地进行穿越滑动面的开关转换。特别地,对于二阶系统,控制律(19.10)变为

$$u = -\hat{f}(\mathbf{x}) + \ddot{x}_d - \lambda \dot{e} - K(x, \dot{x}) \text{sat}(s/\Phi) \quad (19.28)$$

其中,饱和函数 $\text{sat}(s/\Phi)$ 的定义为

$$\text{sat}(s/\Phi) = \begin{cases} -1, & s/\Phi \leq -1 \\ s/\Phi, & -1 < s/\Phi \leq 1 \\ 1, & s/\Phi > 1 \end{cases} \quad (19.29)$$

显然,如果状态位于边界层之外,即,如果 $|s/\Phi| > 1$, 则 $\text{sat}(s/\Phi) = \text{sgn}(s)$, 于是控制律(19.28)是等价于控制律(19.10)的。因此,当 $K(x, \dot{x})$ 由式(19.13)给出时,控制律(19.28)能保证在边界层 $B(t)$ 的外部,滑动条件(19.28)是满足的。控制律(19.28)是平滑控制律,不需要不连续跨越滑动面的开关控制。

19.2.2 根据平滑滑动控制律设计模糊控制器

由上一小节可知,如果根据式(19.28)设计模糊控制器,则能保证跟踪误差在有限时间内满足式(19.19)。对于一给定的精度 ε , 可以选择 Φ 和 λ , 使 $\varepsilon = \Phi/\lambda^{n-1}$ 。即,可以通过给定设计参数 Φ 和 λ 使跟踪误差收敛在任意精确区间 $|e(t)| \leq \varepsilon$ 内。因为(19.28)中的控制 u 是 x 和 \dot{x} 的平滑函数,所以可以设计一个模糊控制器来逼近(19.28)中的 u 。由上一小节,可有如下定理:

定理 19.2 考虑 $n=2$ 时的非线性系统(19.1), 设控制 u 是模糊控制器 u_{fuzz} 。如果模糊控制器设计为

$$u_{fuzz}(\mathbf{x}) = -\hat{f}(\mathbf{x}) + \ddot{x}_d - \lambda \dot{e} - [\eta + F(\mathbf{x})] \text{sat}(s/\Phi) \quad (19.30)$$

则有限时间后,跟踪误差 $e(t) = x(t) - x_d(t)$ 将满足式(19.19)。

证明 由于状态在边界层 $B(t)$ 的外部时,模糊控制器(19.30)满足滑动条件(19.8), 则本定理由引理 19.2 得证。

下面的任务是设计模糊控制器用来逼近式(19.30)的右端项。由于式(19.30)右端项中的所有函数都已知,所以,可以计算相平面内式(19.30)右端在某些规则点处的值,并可使用第 10 章和第 11 章中的方法设计模糊控制器。具体地,给出下面的设计方法。

模糊控制器的设计:

步骤 1 为 e 和 \dot{e} 确定有意义的值域。即,确定区间 $[\alpha_1, \beta_1]$ 和 $[\alpha_2, \beta_2]$, 使得 $\mathbf{e} = (e, \dot{e})^T = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2]$ 。

步骤 2 令 $g(e, \dot{e}) = -\hat{f}(\mathbf{x}) + \ddot{x}_d - \lambda \dot{e} - [\eta + F(\mathbf{x})] \text{sat}(s/\Phi)$, 并将 $g(e, \dot{e})$ 看做式(10.9)中的 g 。根据 10.2 节或 11.1 节中的步骤设计模糊控制器。即,把模糊控制器设计为模糊系统(10.10)。

此方法设计的模糊控制器逼近式(19.30)中的理想模糊控制器 $u_{fuzz}(\mathbf{x})$ 的精度可由定

理 10.1 或定理 11.1 给出。由这些定理可知,只要在取值区间内抽取足够的样本,逼近误差就能像要求的那样小。

例 19.1 考虑一阶非线性系统

$$\dot{x}(t) = \frac{1 - e^{-x(t)}}{1 + e^{-x(t)}} + u(t) \quad (19.31)$$

其中,假定非线性函数 $f(x) = \frac{1 - e^{-x(t)}}{1 + e^{-x(t)}}$ 未知。下面的任务是基于平滑滑动控制律设计模糊控制器,使 $x(t)$ 收敛于零。

由于

$$f(x) = 1 - \frac{-2e^{-x(t)}}{1 + e^{-x(t)}} \quad (19.32)$$

且选定 $\hat{f}(x) = 1, \Delta f(x) = \frac{-2e^{-x(t)}}{1 + e^{-x(t)}}, F(x) = 2$, 则有 $|\Delta f(x)| \leq F(x)$ 。对于此例, $x_d = 0, e = x - 0 = x$, 滑动条件(19.8)变成

$$x(f+u) \leq -\eta|x| \quad (19.33)$$

因此,滑动控制律为

$$u(x) = -\hat{f}(x) - K(x)\text{sgn}(x) \quad (19.34)$$

其中, $K(x) = \eta + F(x) = \eta + 2$ 。图 19.3 给出了四个初始条件下应用此滑动控制器的闭环系统的状态 $x(t)$, 其中 $\eta = 0.1$, 采样速率为 0.02 秒。可以看出,振荡发生了。

平滑控制器为

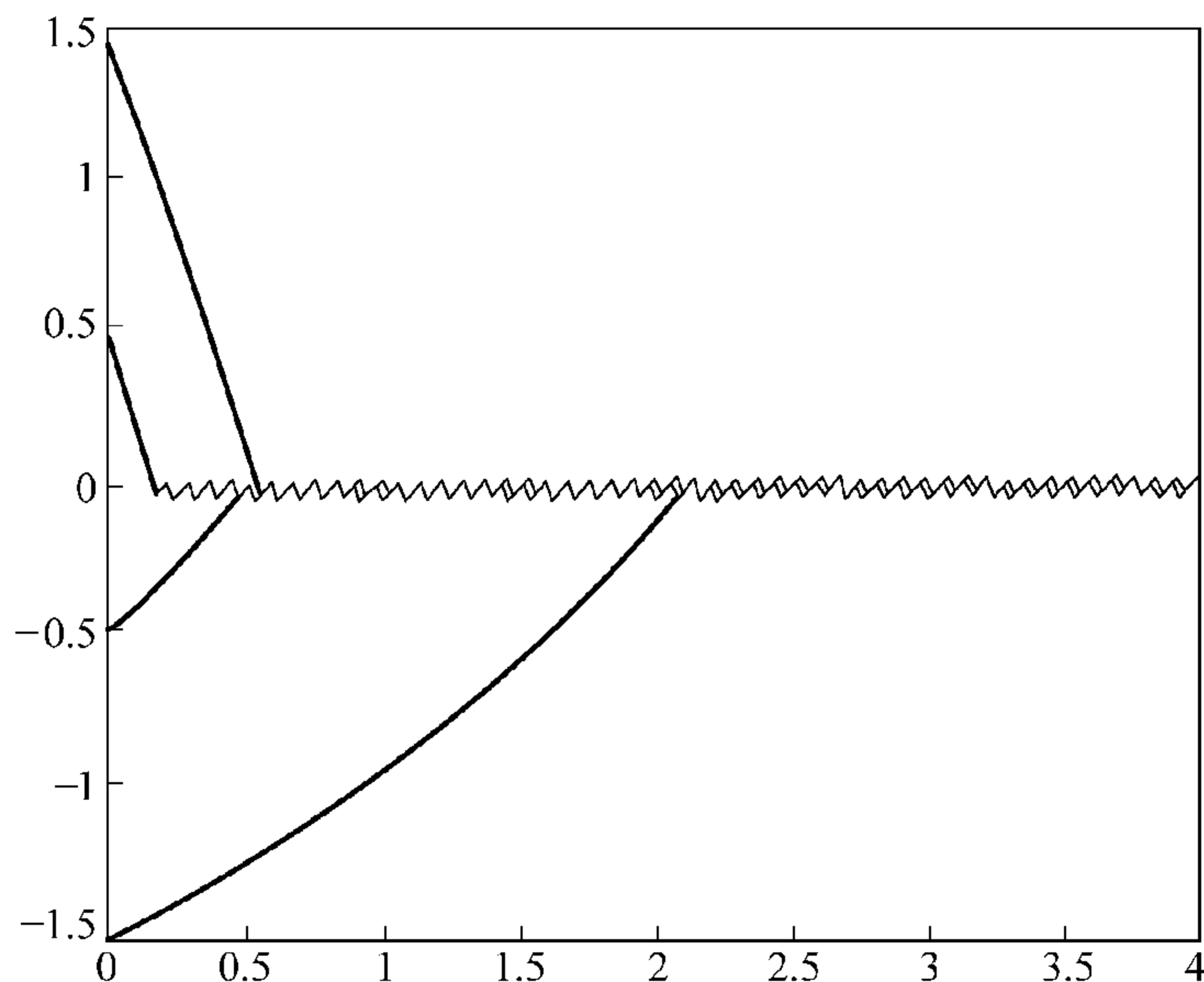


图 19.3 四种不同初始条件下带有滑动控制器(19.34)的非线性系统(19.31)的闭环状态 $x(t)$

$$u_{sm}(x) = -\hat{f}(x) - K(x) \text{sat}(x/\Phi) \quad (19.35)$$

将此 $u_{sm}(x)$ 视为式(10.9)中的 $g(x)$, 根据 11.1 节中的三个步骤设计模糊控制器(这是一个一维模糊系统, 是 11.1 节中系统的一个特例), 选定 $\Phi = 0.2$, $U = [-2, 2]$, $N = 9$, 各个 e^j 的所有值都均匀地分布在区间 $[-2, 2]$ 上。图 19.4 给出了与图 19.3 相同的初始条件下应用此模糊控制器时的闭环系统状态 $x(t)$ 。对比图 19.4 和图 19.3, 可以看到, 振荡消失了, 但出现了稳定的跟踪误差。正如预料的那样, 振荡被平滑掉了, 但却是以牺牲精度为代价的。

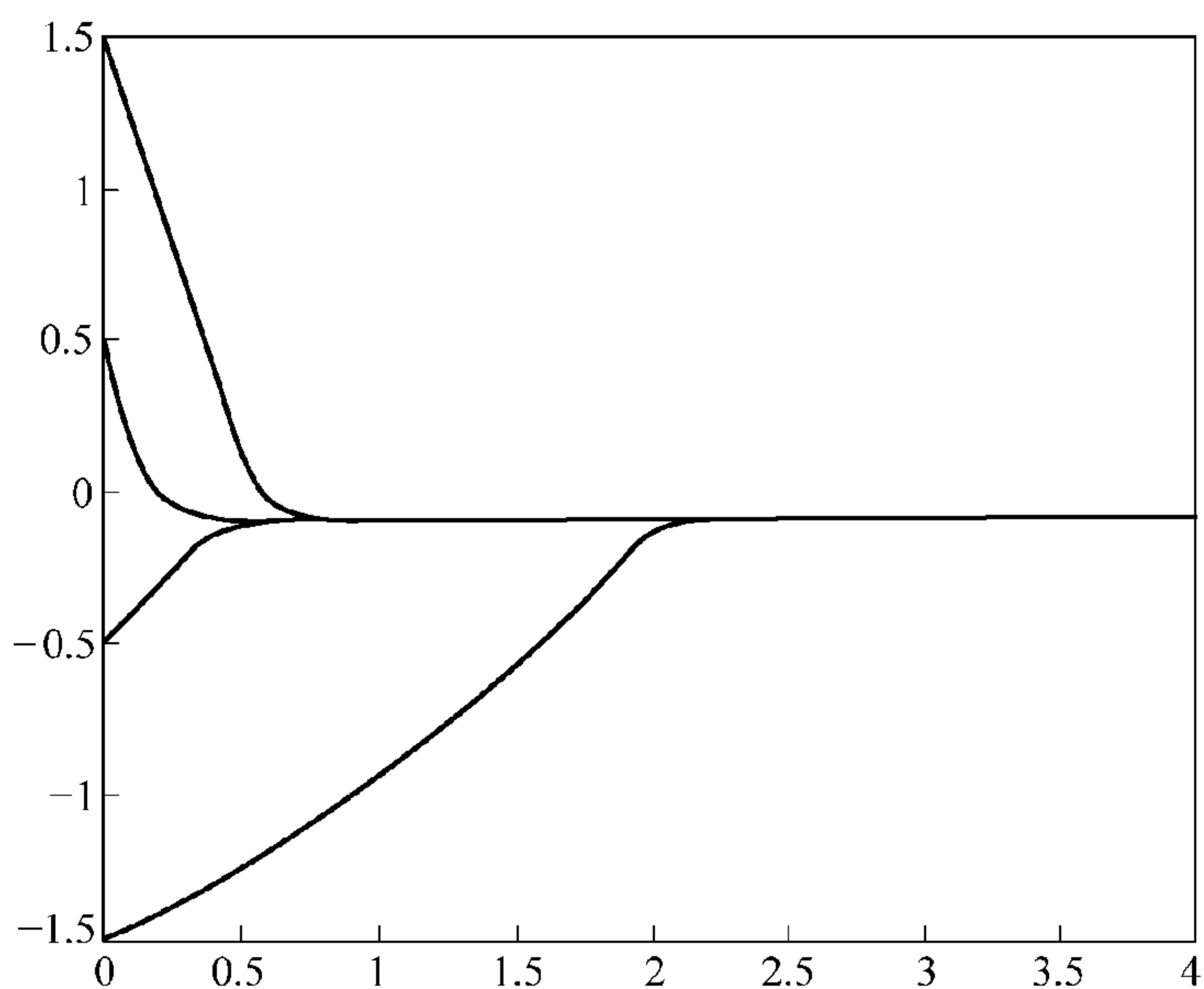


图 19.4 四种不同初始条件下带有模糊控制器的非线性系统(19.31)的闭环状态 $x(t)$

19.3 本章小结及导读

本章主要阐述了以下内容:

- 如何设计非线性系统的滑动控制器及其基本假设;
- 什么是振荡和如何平滑振荡(在跟踪精度与平滑控制之间取个平衡点);
- 如何根据平滑滑动控制律设计模糊控制器。

滑动控制的研究成果可见 Utkin[1978], Slotine 和 Li[1991]。而将滑动模式概念应用于模糊控制则主要归功于 Palm[1992]。在 Driankov, Hellendoorn 和 Reinfrank[1993] 与 Yager 和 Filev[1994] 的书中也对滑模模糊控制做了介绍。

19.4 习题

习题 19.1 考虑二阶系统

$$\ddot{x} + a(t)\dot{x}^2 \cos 3x = u \quad (19.36)$$

式中, $a(t)$ 未知, 但可证明

$$1 \leq a(t) \leq 2 \quad (19.37)$$

设计一滑动控制器 u 使 x 收敛于预期轨迹 x_d 。

习题 19.2 考虑非线性系统

$$\ddot{x} = f(x, \dot{x}) + g(x, \dot{x})u \quad (19.38)$$

式中, f 和 g 未知, 但 $f(\mathbf{x}) = \hat{f}(\mathbf{x}) + \Delta f(\mathbf{x})$, $|\Delta f(\mathbf{x})| \leq F(\mathbf{x})$, $\hat{f}(\mathbf{x})$ 和 $F(\mathbf{x})$ 已知, $0 < g_{\min}(\mathbf{x}) \leq g(\mathbf{x}) \leq g_{\max}(\mathbf{x})$, $g_{\min}(\mathbf{x})$ 和 $g_{\max}(\mathbf{x})$ 已知。令 $\hat{g}(\mathbf{x}) = [g_{\min}(\mathbf{x})g_{\max}(\mathbf{x})]^{1/2}$ 为 $g(\mathbf{x})$ 的估计, 且 $\beta = [g_{\max}(\mathbf{x})/g_{\min}(\mathbf{x})]^{1/2}$ 则可证明下面的控制律满足滑动条件(19.8):

$$u = \hat{g}^{-1}(\mathbf{x})[-\hat{f}(\mathbf{x}) + \ddot{x}_d - \lambda \dot{e} - K(\mathbf{x})\text{sgn}(s)] \quad (19.39)$$

其中

$$K(\mathbf{x}) \geq \beta(F + \eta) + (\beta - 1)|-\hat{f}(\mathbf{x}) + \ddot{x}_d - \lambda \dot{e}| \quad (19.40)$$

习题 19.3 对例 19.1 的滑动控制器(19.34)采用不同的采样速度进行仿真, 可观察到一种现象: 采样速度愈大, 振荡愈强(如术语“冲击巨大”)。

习题 19.4 考虑非线性系统

$$\dot{x}_1 = \sin x_2 + \sqrt{t+1}x_2 \quad (19.41)$$

$$\dot{x}_2 = \alpha_1(t)x_1^4 \cos x_2 + \alpha_2(t)u \quad (19.42)$$

其中, $\alpha_1(t)$ 和 $\alpha_2(t)$ 是未知时变函数, 已知其取值范围为

$$|\alpha_1(t)| \leq 10, 1 \leq \alpha_2(t) \leq 2, \forall t \geq 0 \quad (19.43)$$

根据平滑滑动控制器设计模糊控制器, 使闭环系统状态 $x_1(t)$ 能够跟踪某一给定轨迹 $x_d(t)$ 。

习题 19.5 应用 19.1.1 小节中的方法为倒立摆系统设计一个滑动控制器, 应用第 16 章中的试错法设计一个模糊控制器。通过画图对两个控制器加以比较: (a) 两个控制器的控制面; (b) 应用上述两个控制器时闭环系统的响应。给出你比较后的结论。

20.2 用非模糊监控实现稳定的模糊控制

20.2.1 监督控制器的设计

从概念上看,至少有两种不同的方法可以确保模糊控制系统的稳定性。第一种方法是为模糊控制器选定特殊的结构和参数,使带有模糊控制器的闭环系统稳定(如第 17 章和第 19 章中的模糊控制器)。在第二种方法中,设计模糊控制器时先不考虑稳定性,然后将另一个控制器添加到模糊控制器上以满足其稳定性需要。因为第二种方法中模糊控制器设计有很大的灵活性,所以用此方法设计的模糊控制系统可望获得更好的性能。

第二种方法的关键是设计添加的第二层非模糊控制器,使稳定性得到保证。因为模糊控制器是执行主要控制操作的,所以第二层控制器最好是一个安全警卫,而不是主控制器。因此,这里选择第二层控制器执行以下监督职能:如果模糊控制器运行良好,则第二层控制器停止工作;如果纯模糊控制系统趋于不稳定,则第二层控制器开始工作以确保稳定性。所以,第二层控制器被称为监督控制器。

考虑由下面微分方程确定的非线性系统

$$\dot{x}^{(n)} = f(x, \dot{x}, \dots, x^{(n-1)}) + g(x, \dot{x}, \dots, x^{(n-1)})u \quad (20.1)$$

其中, $x \in R$ 是系统的输出, $u \in R$ 是控制, $\mathbf{x} = (x, \dot{x}, \dots, x^{(n-1)})^T$ 是状态向量并假定其可测或可计算, f 和 g 是未知非线性函数。这里设 $g > 0$ 。由非线性控制理论(Isidori[1989])可知,此系统是一个标准型,许多常见的非线性系统都可转换成此种形式。其最主要的限制是,控制 u 必须在方程中以线性关系出现。

现假定模糊控制器已经为系统设计了一个

$$u = u_{fuzz}(\mathbf{x}) \quad (20.2)$$

这是可以实现的,如用第 16 章中的试错法。下面的任务是确保闭环系统的稳定性,同时,不改变模糊控制器的现存设计 $u_{fuzz}(\mathbf{x})$ 。更具体地讲,需要设计一个控制器,其主要控制行为是模糊控制 $u_{fuzz}(\mathbf{x})$,且要求带有此控制器的闭环系统是全局稳定的。即,状态变量 \mathbf{x} 是均匀有界的,对任意 $t > 0$,有 $|\mathbf{x}(t)| \leq M_x, \forall t > 0$,其中 M_x 是设计者给定的常数。

为完成此任务,在模糊控制器 $u_{fuzz}(\mathbf{x})$ 上添加一个监督控制器 $u_s(\mathbf{x})$,它只有在状态变量达到约束集 $\{\mathbf{x}: |\mathbf{x}| \leq M_x\}$ 的边界时才不为零。即,控制变为

$$u = u_{fuzz}(\mathbf{x}) + I^* u_s(\mathbf{x}) \quad (20.3)$$

式中, I^* 为指示函数(indicator function),当 $|\mathbf{x}| \geq M_x$ 时 $I^* = 1$; 当 $|\mathbf{x}| < M_x$ 时 $I^* = 0$ 。因此,主要控制行为仍然是模糊控制器 $u_{fuzz}(\mathbf{x})$ 。现在的任务是设计 u_s ,使对所有 $t > 0$ 有 $|\mathbf{x}(t)| \leq M_x$ 。

首先计算一下,能否不借助附加假设来设计一个这类监督控制器。将式(20.3)代入式

(20.1), 可得闭环系统满足

$$\dot{\mathbf{x}}^{(n)} = f(\mathbf{x}) + g(\mathbf{x})u_{fuzz}(\mathbf{x}) + g(\mathbf{x})I^*u_s(\mathbf{x}) \quad (20.4)$$

现假设 $\|\mathbf{x}\| = M_x$, 则 $I^* = 1$ 。因为这里假设 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 是完全未知的且是任意非线性函数, 所以对任意 $u_s(\mathbf{x})$ 都可找到 $f(\mathbf{x})$ 和 $g(\mathbf{x})$, 使式(20.4)的右端为正, 从而使 $\|\mathbf{x}\| < M_x$ 。因此, 为了使 u_s 的设计成为可能, 必须对 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 做进一步的假设。假设如下:

假设 20.1: 假设可确定函数 $f^U(\mathbf{x})$ 和 $g_L(\mathbf{x})$, 使得 $|f(\mathbf{x})| \leq f^U(\mathbf{x}), 0 < g_L(\mathbf{x}) \leq g(\mathbf{x})$ 。即, 假设 $|f(\mathbf{x})|$ 的上界和 $g(\mathbf{x})$ 的下界是已知的。

在实际系统中, 通常边界 $f^U(\mathbf{x})$ 和 $g_L(\mathbf{x})$ 不难找到, 因为往往只需要知道一个大致的边界, 即 $f^U(\mathbf{x})$ 可以很大, $g_L(\mathbf{x})$ 可以很小。同时, 还需要知道与状态变量相关的边界, 这比要求有固定边界的条件还弱。

在设计监督控制器 u_s 之前, 需要将闭环系统的方程写成向量形式。首先, 定义

$$u^* = \frac{1}{g(\mathbf{x})}[-f(\mathbf{x}) - \mathbf{k}^T \mathbf{x}] \quad (20.5)$$

其中, $\mathbf{k} = (k_n, \dots, k_1)^T \in R^n$ 使得多项式 $s^n + k_1 s^{n-1} + \dots + k_n$ 的所有根都在复平面的左半开平面上。利用此 u^* , 可将式(20.4)改写为

$$\dot{\mathbf{x}}^{(n)} = -\mathbf{k}^T \mathbf{x} + g[u_{fuzz} - u^* + I^* u_s] \quad (20.6)$$

定义

$$\mathbf{\Lambda} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ -k_n & -k_{n-1} & \cdots & \cdots & \cdots & \cdots & -k_1 \end{bmatrix} \quad (20.7)$$

$$\mathbf{b} = \begin{bmatrix} 0 \\ \cdots \\ 0 \\ g \end{bmatrix} \quad (20.8)$$

则式(20.6)可以改写成向量形式

$$\dot{\mathbf{x}} = \mathbf{\Lambda} \mathbf{x} + \mathbf{b}[u_{fuzz} - u^* + I^* u_s] \quad (20.9)$$

下面设计监督控制器 u_s 以保证 $\|\mathbf{x}\| \leq M_x$ 。定义候选 Lyapunov 函数为

$$V = \frac{1}{2} \mathbf{x}^T P \mathbf{x} \quad (20.10)$$

其中, P 是满足 Lyapunov 方程的正定对称阵

$$\mathbf{\Lambda}^T P + P \mathbf{\Lambda} = -Q \quad (20.11)$$

其中, $Q > 0$ 由设计者给定。因为 $\mathbf{\Lambda}$ 是稳定的, 所以这样的 P 总是存在的。利用式(20.9)和式(20.11)并 $\|\mathbf{x}\| \geq M_x$, 可得

$$\dot{V} = -\frac{1}{2}\mathbf{x}^T Q \mathbf{x} + \mathbf{x}^T P \mathbf{b} [u_{fuzz} - u^* + u_s] \leq |\mathbf{x}^T P \mathbf{b}| (|u_{fuzz}| + |u^*|) + \mathbf{x}^T P \mathbf{b} u_s \quad (20.12)$$

现在的目标是,设计 u_s 使 $\dot{V} \leq 0$ 。即,使式(20.12)的右端非正。根据式(20.12)和式(20.5),可选定 u_s 为如下形式:

$$u_s = -\text{sign}(\mathbf{x}^T P \mathbf{b}) \left[\frac{1}{g_L} (f^U + |\mathbf{k}^T \mathbf{x}|) + |u_{fuzz}| \right] \quad (20.13)$$

将式(20.13)代入式(20.12),有 $\dot{V} \leq 0$ 。因此,式(20.13)中的监督控制器 u_s 能保证 $|\mathbf{x}|$ 在 $|\mathbf{x}(t)| \leq M_x$ 时递减。因此,如果选择初始条件为 $|\mathbf{x}(0)| \leq M_x$,则对于所有 $t \geq 0$ 都有 $|\mathbf{x}(t)| \leq M_x$ 。因为 $g > 0$, \mathbf{x} 和 P 都可得,所以式(20.13)中的 $\text{sign}(\mathbf{x}^T P \mathbf{b})$ 是可以计算出来的。同时,式(20.13)中其他所有项也都可得,所以式(20.13)中的 u_s 可以被在线添加到模糊控制器中。

因为式(20.3)中的 I^* 是一个阶跃函数,每当 \mathbf{x} 碰到边界 $|\mathbf{x}| = M_x$ 时,监督控制器就开始工作,每当 \mathbf{x} 回到约束条件 $|\mathbf{x}| \leq M_x$ 的内部时,监督控制器就停止工作,因此,系统在跨越边界时将受到冲击。克服这种“振荡”的一个办法是,令 I^* 在 0 到 1 之间连续变化。具体地,可以选择如下 I^* :

$$I^* = \begin{cases} 0 & |\mathbf{x}| < a \\ \frac{|\mathbf{x}| - a}{M_x - a} & a \leq |\mathbf{x}| < M_x \\ 1 & |\mathbf{x}| \geq M_x \end{cases} \quad (20.14)$$

式中, $a \in (0, M_x)$, 它是由设计者给定的一个参数。 I^* 形如式(20.3), 则当 \mathbf{x} 从 a 变到 M_x 时, 监督控制器 u_s 将从停止状态连续变化到“最大值”1。显然, 此 I^* 也能保证 $|\mathbf{x}| \leq M_x$ (习题 20.1)。

20.2.2 监督控制器在倒立摆平衡中的应用

本小节将一个模糊控制器与一个监督控制器同时应用于倒立摆平衡问题。控制目标是在平衡倒立摆的同时保证状态处于边界范围内。倒立摆系统如图 1.9 所示,其动态方程为

$$\dot{x}_1 = x_2 \quad (20.15)$$

$$\dot{x}_2 = \frac{g \sin x_1 - \frac{m l x_2^2 \cos x_1 \sin x_1}{m_c + m}}{l \left(\frac{4}{3} - \frac{m \cos^2 x_1}{m_c + m} \right)} + \frac{\frac{\cos x_1}{m_c + m}}{l \left(\frac{4}{3} - \frac{m \cos^2 x_1}{m_c + m} \right)} u \quad (20.16)$$

其中, $g = 9.8 \text{ m/s}^2$ 是重力加速度, m_c 是小车的质量, m 是摆的质量, l 是摆长度的一半, u 是施加的外力(控制)。在下面仿真中选定 $m_c = 1 \text{ kg}$, $m = 0.1 \text{ kg}$, $l = 0.5 \text{ m}$ 。显然,式(20.15) ~

式(20.16)是形如式(20.1)的,这样就可以将前面得到的方法应用于此系统了。

假设模糊控制 u_{fuzz} 是根据下面的四条 IF-THEN 规则构造而成的:

$$\text{如果 } x_1 \text{ 是正且 } x_2 \text{ 是正, 则 } u \text{ 是负最大值} \quad (20.17)$$

$$\text{如果 } x_1 \text{ 是正且 } x_2 \text{ 是负, 则 } u \text{ 是零} \quad (20.18)$$

$$\text{如果 } x_1 \text{ 是负且 } x_2 \text{ 是正, 则 } u \text{ 是零} \quad (20.19)$$

$$\text{如果 } x_1 \text{ 是负且 } x_2 \text{ 是负, 则 } u \text{ 是正最大值} \quad (20.20)$$

其中,模糊集“正”、“负”、“负最大值”、“零”和“正最大值”的隶属度函数分别为

$$\mu_{\text{正}}(x) = \frac{1}{1 + e^{-30x}} \quad (20.21)$$

$$\mu_{\text{负}}(x) = \frac{1}{1 + e^{30x}} \quad (20.22)$$

$$\mu_{\text{负最大值}}(u) = e^{-(u+5)^2} \quad (20.23)$$

$$\mu_{\text{零}}(u) = e^{-u^2} \quad (20.24)$$

$$\mu_{\text{正最大值}}(u) = e^{-(u-5)^2} \quad (20.25)$$

应用中心平均解模糊器和乘积推理机,可以得到如下模糊控制器 u_{fuzz} :

$$u_{fuzz}(\mathbf{x}) = \left(5 \frac{1}{1 + e^{30x_1}} \frac{1}{1 + e^{30x_2}} - 5 \frac{1}{1 + e^{-30x_1}} \frac{1}{1 + e^{-30x_2}} \right) / \left(\frac{1}{1 + e^{30x_1}} \frac{1}{1 + e^{30x_2}} + \frac{1}{1 + e^{-30x_1}} \frac{1}{1 + e^{30x_2}} + \frac{1}{1 + e^{30x_1}} \frac{1}{1 + e^{-30x_2}} + \frac{1}{1 + e^{-30x_1}} \frac{1}{1 + e^{-30x_2}} \right) \quad (20.26)$$

要想设计监督控制器,首先须确定边界 f^U 和 g_L 。对本系统,有

$$|f(x_1, x_2)| = \left| \frac{g \sin x_1 - \frac{m l x_2^2 \cos x_1 \sin x_1}{m_c + m}}{l \left(\frac{4}{3} - \frac{m \cos^2 x_1}{m_c + m} \right)} \right| \leq \frac{9.8 + \frac{0.025}{1.1} x_2^2}{\frac{2}{3} - \frac{0.05}{1.1}} = 15.78 + 0.0366 x_2^2 = f^U(x_1, x_2) \quad (20.27)$$

如果要求 $|x_1| \leq \pi/9$ (这里将给定设计参数以使此要求得到满足),则

$$|g(x_1, x_2)| \geq \frac{\cos \pi/9}{1.1 \left(\frac{2}{3} + \frac{0.05}{1.1} \cos^2 \pi \right)} \doteq 1.1 = g_L(x_1, x_2) \quad (20.28)$$

控制的目标是能将任意初始角度 $x_1 \in [-\pi/9, \pi/9]$ 的倒立摆控制到平衡点,并同时保证 $\|(x_1, x_2)\|_2 \leq \pi/9 \equiv M_x$ 。

给定如下设计参数: $a = \pi/18, k_1 = 2, k_2 = 1$ (此时 $s^2 + k_1 s + k_2$ 是稳定的), $Q = \text{diag}(10, 10)$ 。则求解 Lyapunov 方程(20.11),可得

$$P = \begin{bmatrix} 15 & 5 \\ 5 & 5 \end{bmatrix} \quad (20.29)$$

这里给出三种情况下的仿真结果：(i) 没有监督控制器，即，只有模糊控制器(20.26)；(ii) 共同使用监督控制器与模糊控制器；(iii) 除了将表示轻微干扰的高斯噪声(方差为 3) 添加到 u 上外，其余都同(ii)。这里给出了闭环系统在每种情况下，从五个初始条件： $(x_1(0), x_2(0)) = (4^\circ, 0), (8^\circ, 0), (12^\circ, 0), (16^\circ, 0), (20^\circ, 0)$ 出发所得到的仿真结果。(i)、(ii) 和 (iii) 三种情况下的仿真结果分别见图 20.3、图 20.4 和图 20.5。图中将五个初始条件的角度 $x_1(t)$ 表示成 t 的函数。从结果可以看出：(a) 纯模糊控制器能将倒立摆从较小的初始角度 $4^\circ, 8^\circ$ 和 12° 控制到平衡点，但当初始角度较大，为 16° 和 20° 时，系统开始不稳定；(b) 将监督控制器添加到模糊控制器后，就可以成功地将倒立摆从五个初始值控制到平衡点，且能保证角度值在 $[-20^\circ, 20^\circ]$ 之间；(c) 模糊控制器对于随机干扰是稳健的。

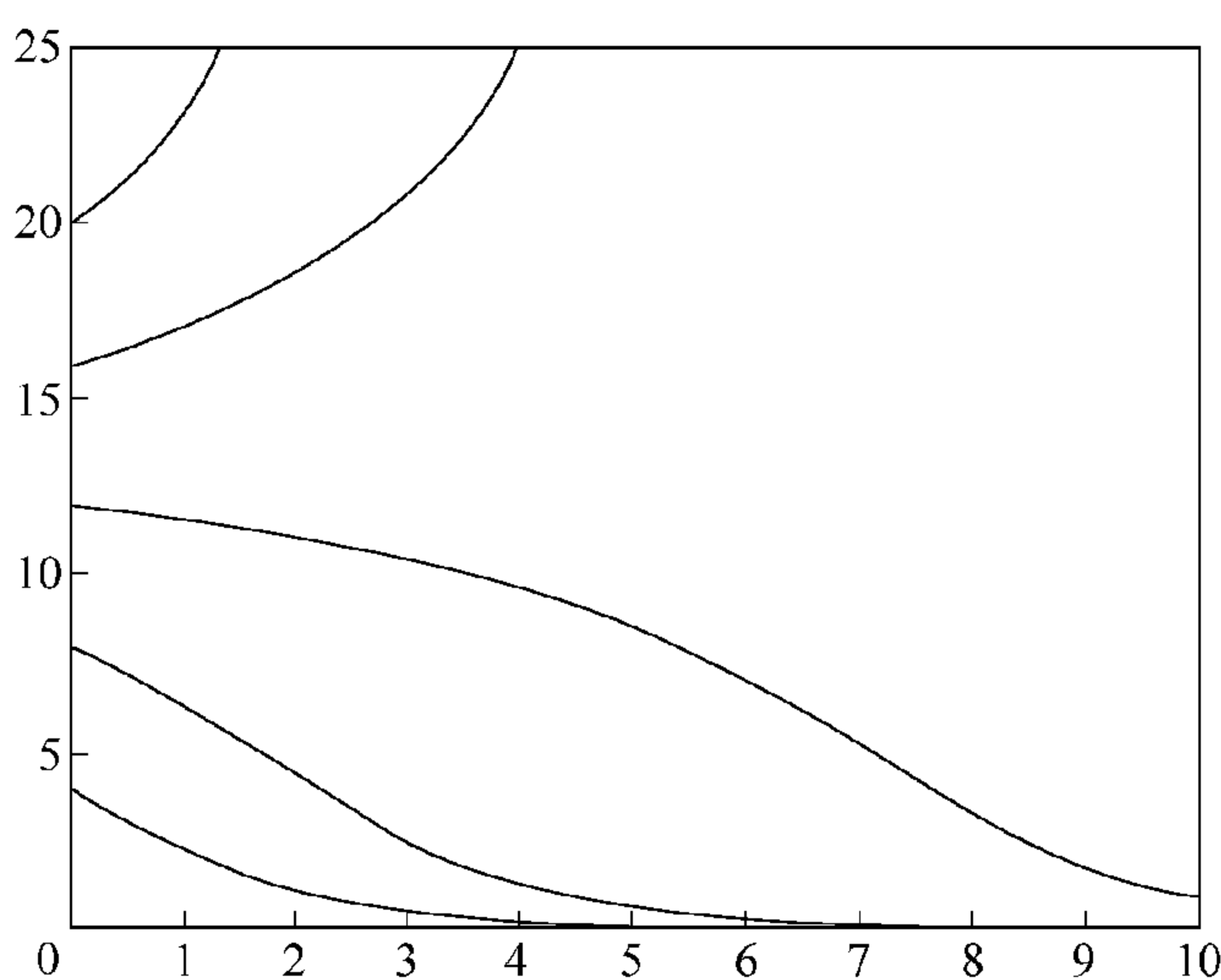


图 20.3 纯模糊控制器在五种初始条件下的闭环系统状态 $x_1(t)$

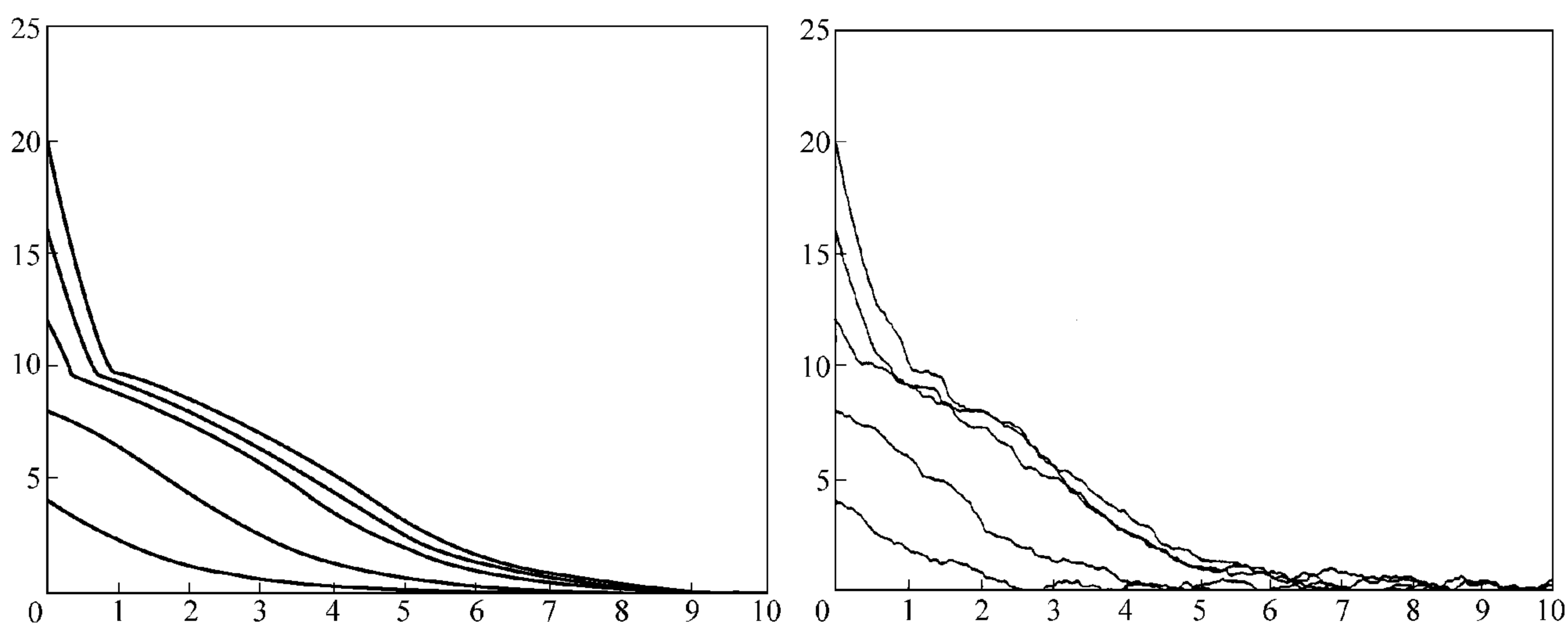


图 20.4 带有监督控制器的模糊控制器在五种初始条件下的闭环系统状态 $x_1(t)$

图 20.5 在图 20.4 所示系统的控制 u 中添加了方差为 3 的高斯噪声

20.3 用模糊系统调节 PID 控制器的增益

20.3.1 PID 控制器

由于比例—积分—微分(PID)控制器因其结构简单、稳健性好而成为工业过程控制中最常用的控制器。PID 控制器的传递函数形式如下:

$$G(s) = K_p + K_i/s + K_d s \quad (20.30)$$

其中, K_p 、 K_i 和 K_d 分别称为比例增益、积分增益和微分增益。PID 控制器的另一等价形式为

$$u(t) = K_p \left[e(t) + \frac{1}{T_i} \int_0^t e(\tau) d\tau + T_d \dot{e}(t) \right] \quad (20.31)$$

其中, $T_i = K_p/K_i$, $T_d = K_d/K_p$ 分别称为积分时间常数和微分时间常数。

PID 控制器的成功应用有赖于 PID 增益的正确选择。将 PID 增益调至最佳状态亦非无关紧要。在实践中, PID 增益通常由有经验的专家人员根据“简单规则”(“rule of thumb”)来调整。在下一小节中, 先通过分析系统的典型反应给出一组确定 PID 增益的调节规则(模糊 IF-THEN 规则), 然后将这些规则整合到模糊系统中, 以实现 PID 增益的在线调节。这里采用的是 Zhao, Tomizuka 和 Isaba[1993]中给出的方法。

20.3.2 调节 PID 增益的模糊系统

考虑图 20.2 所示的两层控制系统, 其中传统控制器采用的是形如式(20.30)[等价于式(20.31)]的 PID 控制器, 并用模糊系统对 PID 增益进行实时调整。模糊系统是由模糊规则集合构造而成的, 该规则集合描述了在各种情况下如何选择 PID 增益。所以, 我们首先对问题进行主题表述, 然后导出模糊 IF-THEN 规则。

假设可以确定区间 $[K_{pmin}, K_{pmax}] \subset R$ 和 $[K_{dmin}, K_{dmax}] \subset R$, 使得有比例增益 $K_p \in [K_{pmin}, K_{pmax}]$ 和微分增益 $K_d \in [K_{dmin}, K_{dmax}]$ 。为方便起见, 用下面线性变换将 K_p 和 K_d 归一化成 0—1 之间的数:

$$K'_p = \frac{K_p - K_{pmin}}{K_{pmax} - K_{pmin}} \quad (20.32)$$

$$K'_d = \frac{K_d - K_{dmin}}{K_{dmax} - K_{dmin}} \quad (20.33)$$

假定积分时间常数可参考微分时间常数给出

$$T_i = \alpha T_d \tag{20.34}$$

由此可得

$$K_i = K_p / (\alpha T_d) = K_p^2 / (\alpha K_d) \tag{20.35}$$

于是,需要用模糊系统调节参数 K'_p 、 K'_d 和 α 。如果能确定这些参数,则 PID 增益能由式(20.32)、式(20.33)和式(20.35)得到。假定模糊系统的输入为 $e(t)$ 和 $\dot{e}(t)$,则模糊系统调节器由三个两维输入一维输出的模糊系统组成(见图 20.6)。下面给出构造这些模糊系统的模糊 IF-THEN 规则。

令模糊 IF-THEN 规则具有如下形式:

$$\text{如果 } e(t) \text{ 是 } A^l, \text{ 且 } \dot{e}(t) \text{ 是 } B^l, \text{ 则 } K'_p \text{ 是 } C^l, K'_d \text{ 是 } D^l, \alpha \text{ 是 } E^l \tag{20.36}$$

其中, A^l 、 B^l 、 C^l 、 D^l 和 E^l 是模糊集, $l = 1, 2, \dots, M$ 。假设 $e(t)$ 和 $\dot{e}(t)$ 有意义的区域分别为 $[e_M^-, e_M^+]$ 和 $[e_{Md}^-, e_{Md}^+]$, 并在上面定义 7 个如图 20.7 所示的模糊集。于是,完整的模糊规则库由 49 条规则构成。为简单起见,假设 C^l 和 D^l 是模糊集“大”,或是模糊集“小”,它们的隶属度函数如图 20.8 所示。最后,假定 E^l 可为图 20.9 所示的四个模糊集。下面,就可以推导模糊规则了。

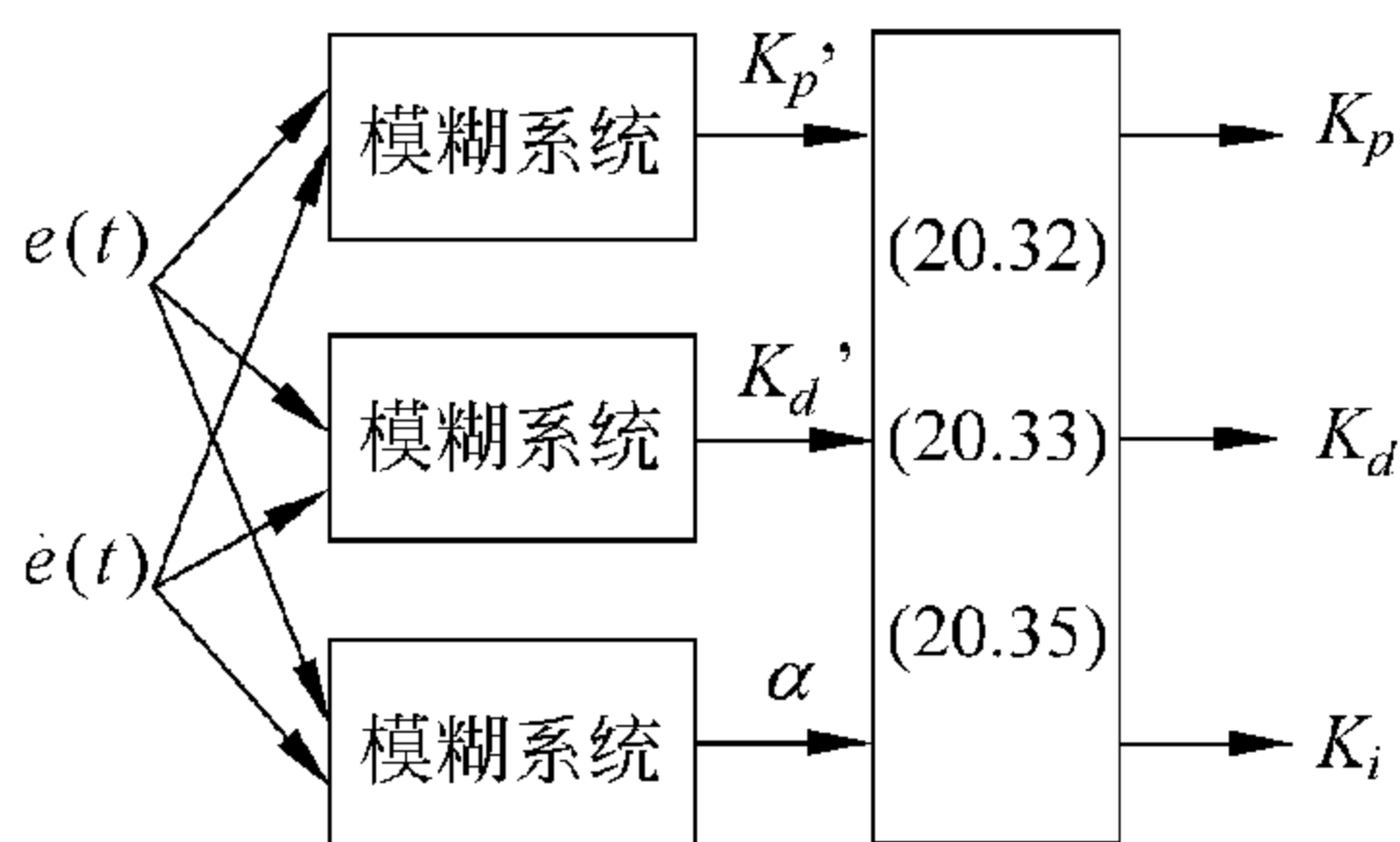


图 20.6 PID 增益的模糊系统调节器

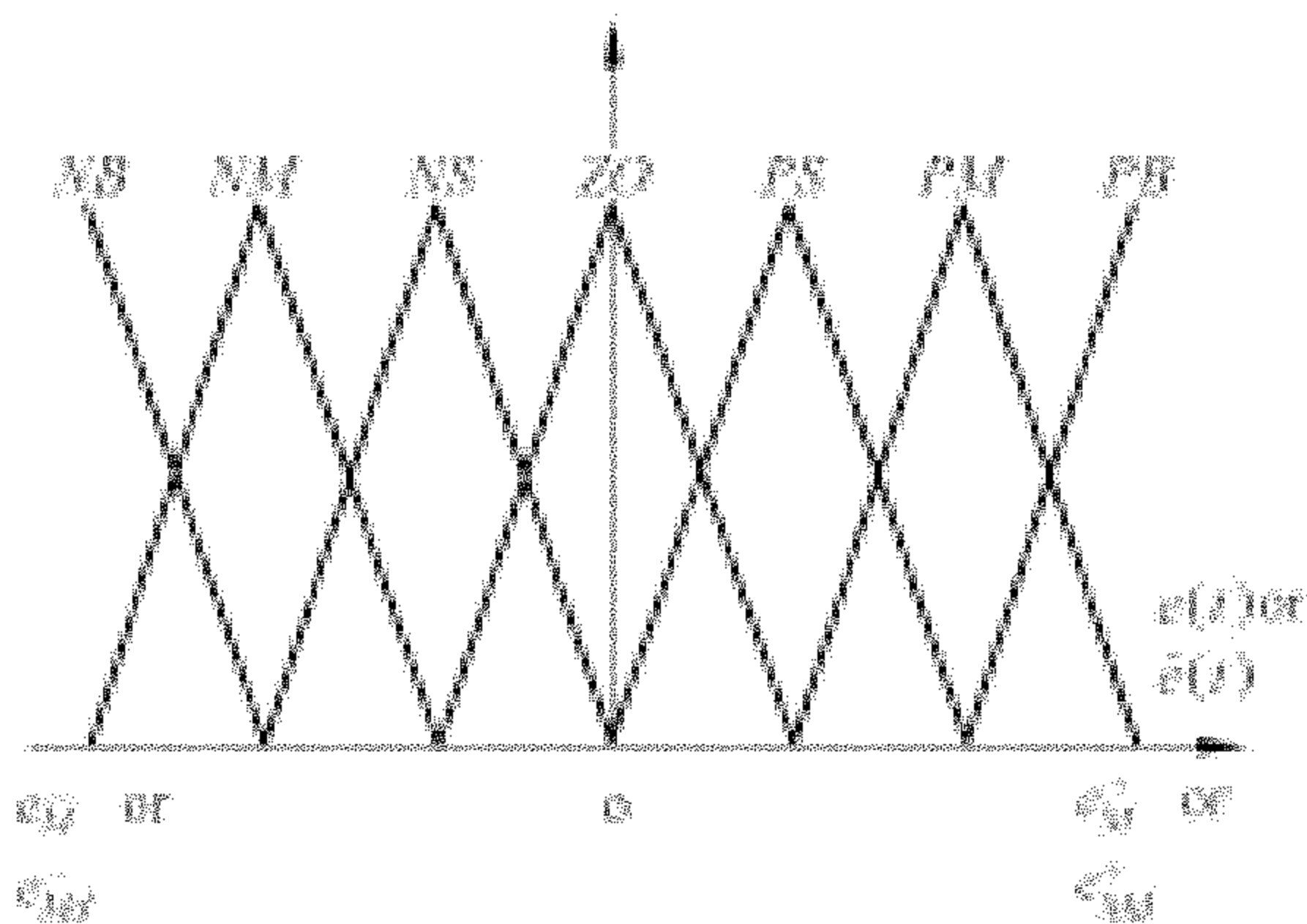


图 20.7 $e(t)$ 和 $\dot{e}(t)$ 的隶属度函数

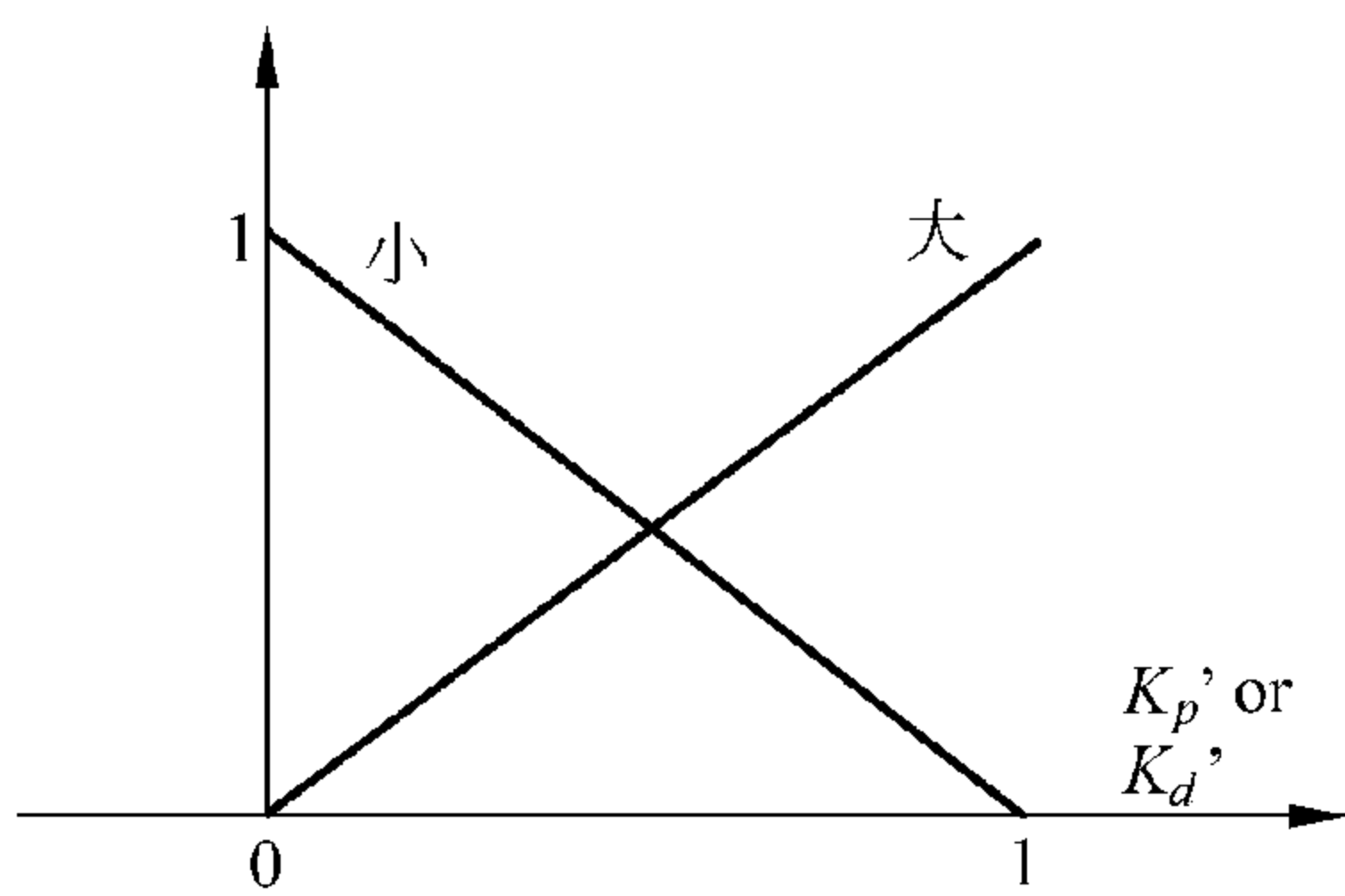


图 20.8 K'_p 、 K'_d 的隶属度函数

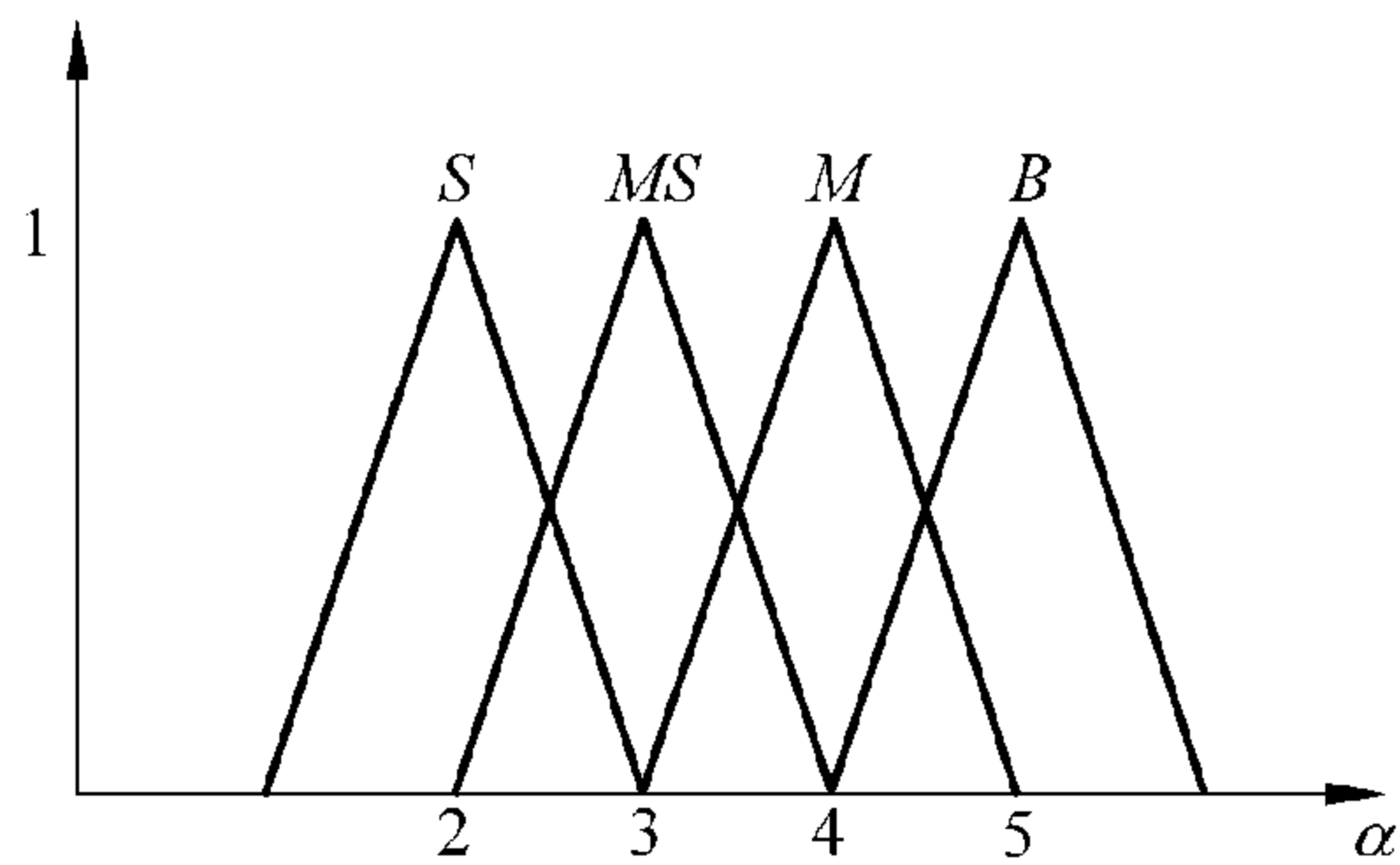


图 20.9 α 的隶属度函数

这里根据过程的典型阶跃反应来推导经验性规则。图 20.10 给出了典型时间反应的例子。刚开始时,即在 a_1 附近,需要一个大的控制信号以快速启动时间。为了产生一个大的控制信号,需要一个大的比例增益 K'_p 、一个小的微分增益 K'_d 和一个大的积分增益。由式

(20.35)可以看出,当 K_p 和 K_d 固定不变时,积分增益和 α 成反比。所以,积分增益大说明 α 小。由此可知,在 a_1 附近的规则为

$$\text{如果 } e(t) \text{ 是 } PB \text{ 且 } \dot{e}(t) \text{ 是 } ZO, \text{ 则 } K'_p \text{ 是大, } K'_d \text{ 是小, } \alpha \text{ 是 } S \quad (20.37)$$

其中,模糊集 PB 、 ZO 、大、小和 S 的隶属度函数分别见图20.7 ~ 图 20.9。

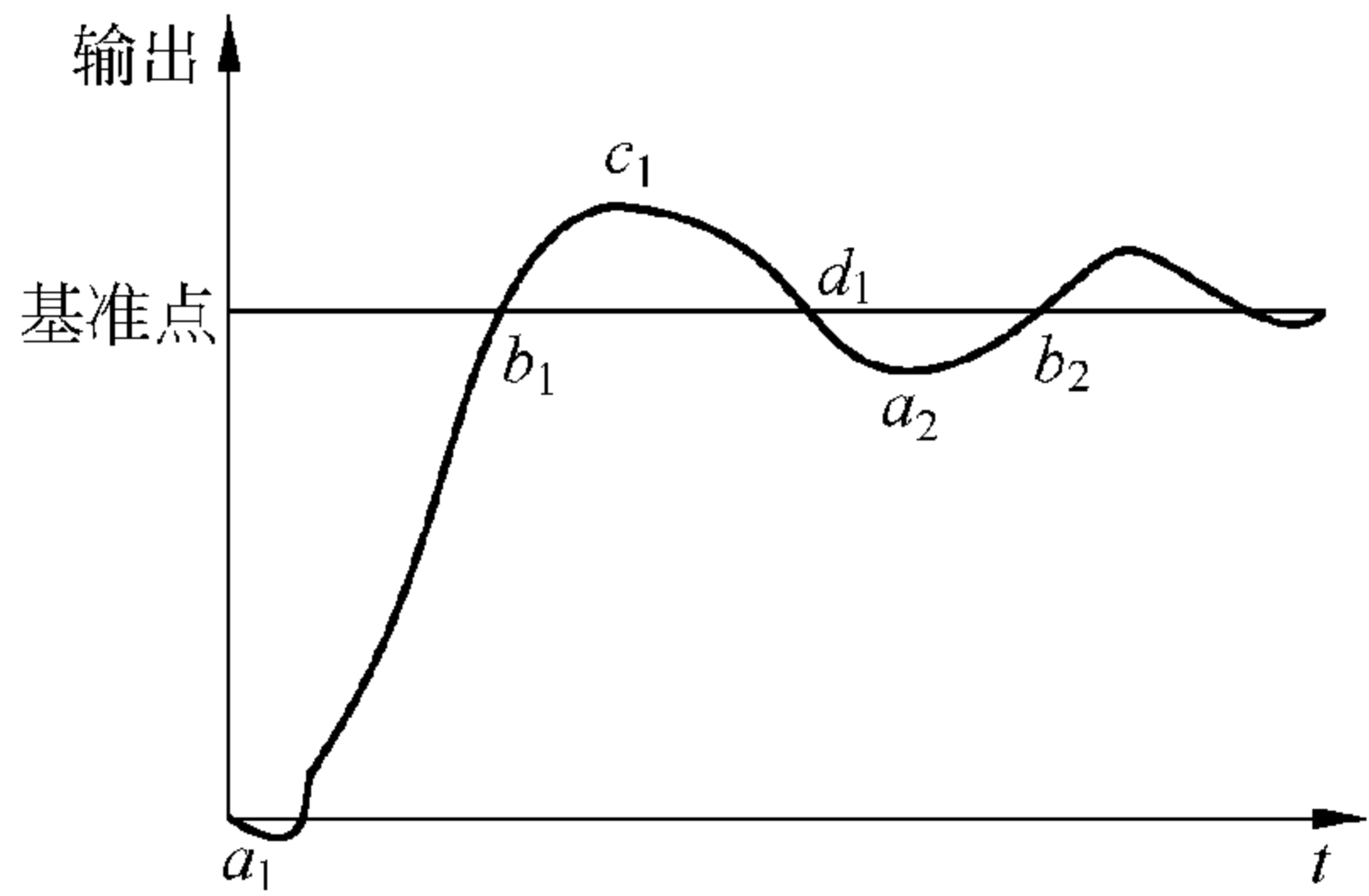


图 20.10 过程的典型阶跃响应

在图 20.10 中的 b_1 附近,需要一个小的控制信号以避免错过目标。所以这里需要一个小的比例增益、一个大的微分增益和一个小的积分增益。因此,需要采用下面的规则:

$$\text{如果 } e(t) \text{ 是 } ZO \text{ 且 } \dot{e}(t) \text{ 是 } NB, \text{ 则 } K'_p \text{ 是小, } K'_d \text{ 是大, } \alpha \text{ 是 } B \quad (20.38)$$

在点 c_1 和 d_1 附近的控制行为分别与点 a_1 和 b_1 附近的相似。利用这个思路可以确定关于 K'_p 、 K'_d 和 α 的三个规则集合,每个规则集合包含 49 条规则,这些规则集合分别见图20.11 ~ 图 20.13。

		$\dot{e}(t)$						
		NB	NM	NS	ZO	PS	PM	OB
$e(t)$	NB	B	B	B	B	B	B	B
	BM	S	B	B	B	B	B	S
	NS	S	S	B	B	B	S	S
	ZO	S	S	S	B	S	S	S
	PS	S	S	B	B	B	S	S
	PM	S	B	B	B	B	B	S
	PB	B	B	B	B	B	B	S

图 20.11 K'_p 的模糊调节规则

		$\dot{e}(t)$						
		NB	NM	NS	ZO	PS	PM	OB

$e(t)$	NB	S	S	S	S	S	S	S
	BM	B	B	S	S	S	B	B
	NS	B	B	B	S	B	B	B
	ZO	B	B	B	B	B	B	B
	PS	B	B	B	S	B	B	B
	PM	B	B	S	S	S	B	B
	PB	S	S	S	S	S	S	S

图 20.12 K'_d 的模糊调节规则

		$\dot{e}(t)$						
		NB	NM	NS	ZO	PS	PM	OB
$e(t)$	NB	2	2	2	2	2	2	2
	BM	3	3	2	2	2	3	3
	NS	4	3	3	2	3	3	4
	ZO	5	4	3	3	3	4	5
	PS	4	3	3	2	3	3	4
	PM	3	3	2	2	2	3	3
	PB	2	2	2	2	2	2	2

图 20.13 α 的模糊调节规则

将每个规则集合中的 49 条规则用乘积推理机、单值模糊器和中心平均解模糊器相组合,即,参数 K'_p 、 K'_d 和 α 可根据以下各式在线调节:

$$K'_p(t) = \frac{\sum_{l=1}^{49} \bar{y}_p^l \mu_{A^l}(e(t)) \mu_{B^l}(\dot{e}(t))}{\sum_{l=1}^{49} \mu_{A^l}(e(t)) \mu_{B^l}(\dot{e}(t))} \quad (20.39)$$

$$K'_d(t) = \frac{\sum_{l=1}^{49} \bar{y}_d^l \mu_{A^l}(e(t)) \mu_{B^l}(\dot{e}(t))}{\sum_{l=1}^{49} \mu_{A^l}(e(t)) \mu_{B^l}(\dot{e}(t))} \quad (20.40)$$

$$\alpha(t) = \frac{\sum_{l=1}^{49} \bar{y}_\alpha^l \mu_{A^l}(e(t)) \mu_{B^l}(\dot{e}(t))}{\sum_{l=1}^{49} \mu_{A^l}(e(t)) \mu_{B^l}(\dot{e}(t))} \quad (20.41)$$

式中, A^l 和 B^l 如图 20.7 所示, \bar{y}_p^l 、 \bar{y}_d^l 和 \bar{y}_α^l 是图 20.8 和图 20.9 所示的相应模糊集的中心(根据图 20.11 ~ 图 20.13 中的规则推得)。

可在 Zhao, Tomizuka 和 Isaka[1993] 中找到仿真结果和与古典方法的对比性研究成果。

20.4 本章小结及导读

本章主要阐述了以下内容:

- 两层模糊控制的结构和工作原理;
- 如何设计非模糊监督控制器以保证模糊控制系统的全局稳定性;
- 如何由典型阶跃反应的启发式分析推导出调节 PID 增益的模糊 IF-THEN 规则。

多层控制在智能机器人领域有研究,见 Valavanis 和 Saridis[1992]和 Tzafestas[1991]。本章中的监督控制思想由 Wang[1994b]给出。有几位研究者都研究过用模糊系统调节 PID 参数问题,本章中的方法要归功于 Zhao, Tomizuka 和 Isaka[1993]。

20.5 习题

习题 20.1 说明如果控制器(20.3)中的指示函数 I^* 采用其连续形式(20.14),且 u_s 由式(20.13)给出,则当 $|\mathbf{x}(0)| \leq M_x$ 时,对于所有 $t \geq 0$,仍有 $|\mathbf{x}(t)| \leq M_x$ 。并证明此种情况下存在 $M_x^* \leq M_x$,使得当 $|\mathbf{x}(0)| \leq M_x^*$ 时,对于所有 $t \geq 0$,都有 $|\mathbf{x}(t)| \leq M_x^*$ 。

习题 20.2 讨论本章中的监督控制器与第 19 章中的滑动控制器的异同点。

习题 20.3 在模糊控制规则有轻微改变和具有不同的初始条件的情况下,重复 20.22 节中的仿真计算。

习题 20.4 在 PID 控制器(20.30)中, K_p 、 K_i/s 和 $K_d s$ 分别称为比例模式、积分模式和微分模式。举例说明:

(a) 比例模式能快速调整所调变量,它虽然能减少误差但不能提供零稳定状态;它能加速动态反应,但如果调节不正确会导致不稳定。

(b) 积分模式能达到零点稳态补偿,但调节所调变量的速度比比例模式慢,且如果调节不正确也会导致不稳定。

(c) 微分模式不影响误差的最终稳态值,它能根据被控变量的变化率进行快速校正,但可能引起操作变量的高频变动,这是我们所不希望的。

习题 20.5 考虑图 20.2 中的两层模糊控制系统,其过程如下:

$$G(s) = \frac{27}{(s+1)(s+3)^3} \quad (20.42)$$

传统的控制器是 PID 控制器,且模糊系统由式(20.39)~式(20.41)给出。对此系统进行仿真并绘制过程输出的图形。可任意选取 K_{pmin} 、 K_{pmax} 、 K_{dmin} 、 K_{dmax} 、 e_M^- 、 e_M^+ 、 e_{Md}^- 、 e_{Md}^+ 的合理值。

模糊系统模型的模糊控制

在第 17 章至第 20 章中研究的模糊控制系统中,受控过程都是用公共的线性或非线性动态系统模型来表征的。而在大多数实际问题中,可能是由专家人员提供受控过程的语言描述(以模糊 IF-THEN 规则的形式),然后将这些描述整合到过程模型中;我们称此模型为模糊系统模型。因此,研究过程是模糊系统模型而反馈控制器是模糊控制器的模糊控制系统是很有意义的,这也是本章的研究主题。

本章将首先引入 Takagi-Sugeno-Kang(TSK)模糊系统,并推导闭环系统的详细公式,在闭环系统中的过程 and 控制器都是由 TSK 模糊系统表示的。然后,分析此闭环系统的稳定性。最后,介绍稳定模糊控制器的设计程序。

21.1 TSK 模糊系统

TSK(Takagi-Sugeno-Kang)模糊系统是作为本书大部分章节中的模糊系统的替代形式而建立的。TSK 模糊系统是由下面的规则构造而成的:

$$\text{如果 } x_1 \text{ 是 } C_1^l \text{ 且 } \cdots \text{ 且 } x_n \text{ 是 } C_n^l, \text{ 则 } y^l = c_0^l + c_1^l x_1 + \cdots + c_n^l x_n \quad (21.1)$$

其中, C_i^l 是模糊集, c_i^l 是常数, $l = 1, 2, \dots, M$ 。即,规则的 IF 部分与普通模糊 IF-THEN 规则的 IF 部分相同,但规则的 THEN 部分却是输入变量的线性组合。当给定 TSK 模糊系统的输入 $x = (x_1, \dots, x_n)^T \in U \subset R^n$ 时,其输出 $f(x) \in V \subset R$ 可计算为式(21.1)中 y^l 的加权平均,即

$$f(x) = \frac{\sum_{l=1}^M y^l w^l}{\sum_{l=1}^M w^l} \quad (21.2)$$

式中,权重 w^l 计算如下

$$w^l = \prod_{i=1}^n \mu_{C_i^l}(x_i) \quad (21.3)$$

可以看出,TSK 模糊系统仍然是从 $U \subset R^n$ 到 $V \subset R$ 的映射。规则(21.1)的物理意义是,当 x 被约束在以规则的 IF 部分表征的模糊区间内时,其输出是输入变量的线性函数。因

此, TSK 模糊系统可以看做是某种分段线性函数, 它从一段到另一段的变化是光滑的, 而不是陡峭的。如果 $c_i^l = 0 (i = 1, 2, \dots, n)$ 且 c_0^l 等于公共模糊 IF-THEN 规则(7.1)中模糊集 B^l 的中心 \bar{y}^l , 则 TSK 模糊系统与采用乘积推理机、单值模糊器和中心平均解模糊器的模糊系统是一致的[通过比较式(21.2) ~ 式(21.3)与式(9.1)]。

如果 TSK 模糊系统的输出作为它自身的一个输入, 则能够得到所谓的动态 TSK 模糊系统。具体地讲, 一个动态 TSK 模糊系统可由以下规则构造而成:

$$\begin{aligned} & \text{如果 } x(k) \text{ 是 } A_1^p \text{ 且 } \dots \text{ 且 } x(k-n+1) \text{ 是 } A_n^p \text{ 且 } u(k) \text{ 是 } B^p, \\ & \text{则 } x^p(k+1) = a_1^p x(k) + \dots + a_n^p x(k-n+1) + b^p u(k) \end{aligned} \quad (21.4)$$

其中, A_i^p 和 B^p 是模糊集, a_i^p 和 $b^p (p = 1, 2, \dots, N)$ 是常数, $u(k)$ 是系统的输入, $x(k) = (x(k), x(k-1), \dots, x(k-n+1))^T \in R^n$ 是系统的状态向量。动态 TSK 模糊系统的输出由下式计算为

$$x(k+1) = \frac{\sum_{p=1}^N x^p(k+1)v^p}{\sum_{p=1}^N v^p} \quad (21.5)$$

其中, $x^p(k+1)$ 由式(21.4)给出, 且

$$v^p = \prod_{i=1}^n \mu_{A_i^p}[x(k-i+1)] \mu_{B^p}[u(k)] \quad (21.6)$$

下面, 将应用此动态 TSK 模糊系统为所要控制的过程建立模型。

21.2 带有模糊控制器的模糊模型的闭环动态特性

考虑如图 21.1 所示的反馈控制系统, 其中, 受控过程模型是动态 TSK 模糊模型(21.5), 控制器是 TSK 模糊系统(21.2)(其中 $c_0^l = 0, x_i = x(k-i+1), i = 1, 2, \dots, n$)。下面的定理给出了图 21.1 中的控制系统的闭环动态特性(closed-loop dynamics)。

定理 21.1 图 21.1 中的闭环模糊控制系统等价于由下述规则构造的动态 TSK 模糊系统:

$$\begin{aligned} & \text{如果 } x(k) \text{ 是 } (C_1^l \text{ 和 } A_1^p) \text{ 且 } \dots \text{ 且 } x(k-n+1) \text{ 是 } (C_n^l \text{ 和 } A_n^p) \text{ 且 } u(k) \text{ 是 } B^p, \\ & \text{则 } x^l(k+1) = \sum_{i=1}^n (a_i^p + b^p c_i^l) x(k-i+1) \end{aligned} \quad (21.7)$$

其中, $u(k)$ 是控制器的输出, $l = 1, 2, \dots, M, p = 1, 2, \dots, N$, 模糊集 (C_i^l 和 A_i^p) 的隶属度函数为 $\mu_{C_i^l}(x(k-i+1))$ 和 $\mu_{A_i^p}(x(k-i+1))$ 。此动态

TSK 模糊系统的输出由下式计算

$$x(k+1) = \frac{\sum_{l=1}^M \sum_{p=1}^N x^l(k+1) w^l v^p}{\sum_{l=1}^M \sum_{p=1}^N w^l v^p} \quad (21.8)$$

其中

$$w^l = \prod_{i=1}^n \mu_{C_i^l}(x(k-i+1)) \quad (21.9)$$

$$v^p = \prod_{i=1}^n \mu_{A_i^p}(x(k-i+1)) \mu_{B^p}(u(k)) \quad (21.10)$$

证明 由图 21.1 可以看出, 式(21.4)中的 $u(k)$ 等于式(21.2)中的 $f(x)$ 。于是, 式(21.4)中的 $x^p(k+1)$ 变为

$$\begin{aligned} x^p(k+1) &= a_1^p x(k) + \dots + a_n^p x(k-n+1) + b^p \frac{\sum_{l=1}^M [c_1^l x(k) + \dots + c_n^l x(k-n+1)] w^l}{\sum_{l=1}^M w^l} \\ &= \frac{\sum_{l=1}^M [\sum_{i=1}^n (a_i^p + b^p c_i^l) x(k-i+1)] w^l}{\sum_{l=1}^M w^l} \end{aligned} \quad (21.11)$$

将式(21.11)代入式(21.5), 可得到闭环系统的输出为

$$x(k+1) = \frac{\sum_{l=1}^M \sum_{p=1}^N [\sum_{i=1}^n (a_i^p + b^p c_i^l) x(k-i+1)] w^l v^p}{\sum_{l=1}^M \sum_{p=1}^N w^l v^p} \quad (21.12)$$

上式即为式(21.8), 故式(21.8)成立。

例 21.1 假设图 21.1 中的过程模型是二阶动态 TSK 模糊系统, 其建模遵循以下两条规则:

L^1 : 如果 $x(k)$ 是 A_1^1 且 $x(k-1)$ 是 A_2^1 且 $u(k)$ 是 B^1 ,

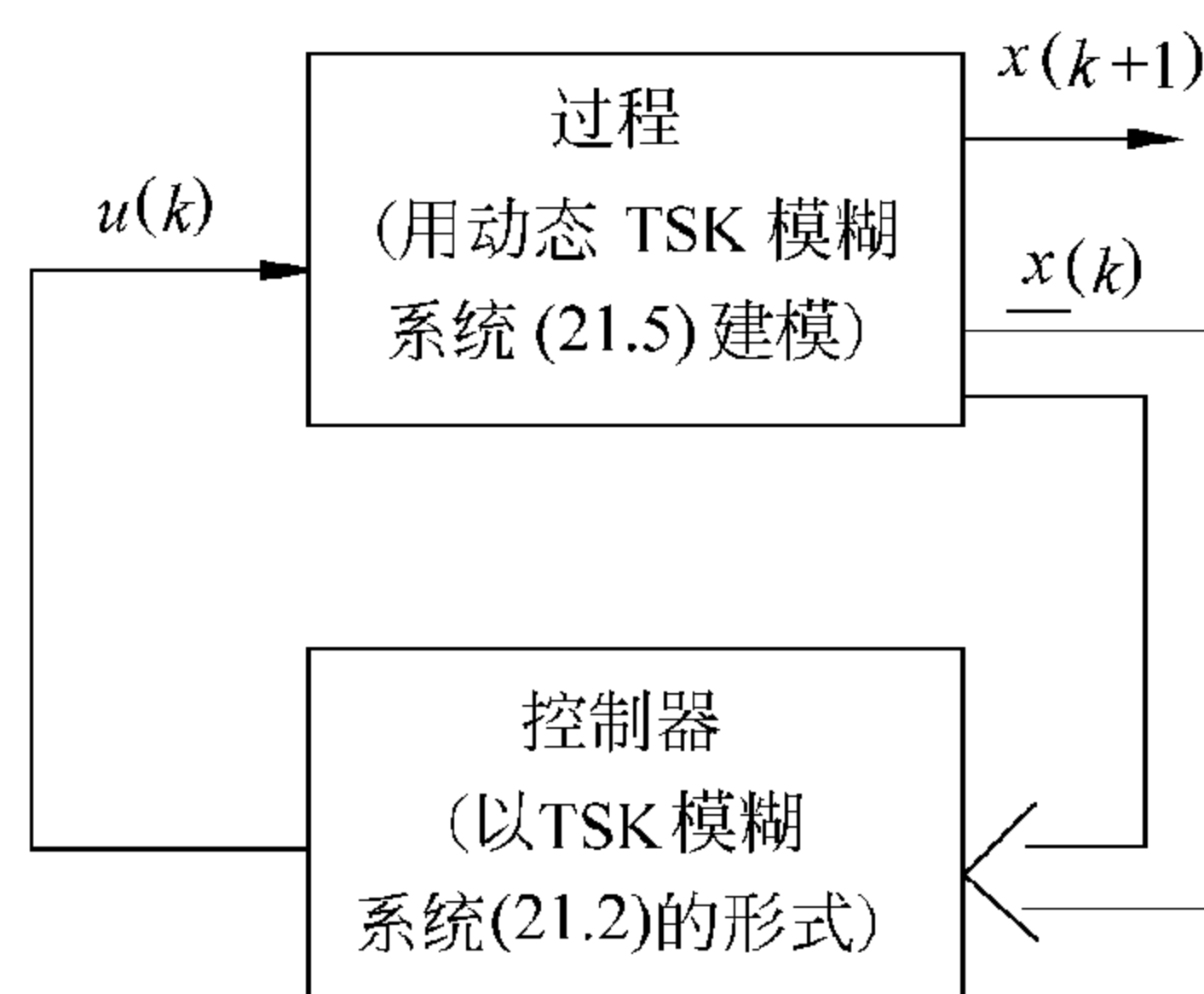


图 21.1 模糊系统模型的模糊控制

$$\text{则 } x^1(k+1) = 1.5x(k) + 2.1x(k-1) - u(k) \quad (21.13)$$

L^2 : 如果 $x(k)$ 是 A_1^2 且 $x(k-1)$ 是 A_2^2 且 $u(k)$ 是 B^2 ,

$$\text{则 } x^2(k+1) = 0.3x(k) - 3.4x(k-1) + 0.5u(k) \quad (21.14)$$

图 21.1 中控制器是由以下两条规则构造的 TSK 模糊系统:

$$R^1: \text{如果 } x(k) \text{ 是 } C_1^1 \text{ 且 } x(k-1) \text{ 是 } C_2^1, \text{ 则 } u^1(k) = k_1^1 x(k) + k_2^1 x(k-1) \quad (21.15)$$

$$R^2: \text{如果 } x(k) \text{ 是 } C_1^2 \text{ 且 } x(k-1) \text{ 是 } C_2^2, \text{ 则 } u^2(k) = k_1^2 x(k) + k_2^2 x(k-1) \quad (21.16)$$

然后根据定理 21.1 可知, 闭环系统是一个由以下四条规则构造而成的动态 TSK 模糊系统:

$$S^{11}: \text{如果 } x(k) \text{ 是 } (A_1^1 \text{ 和 } C_1^1) \text{ 且 } x(k-1) \text{ 是 } (A_2^1 \text{ 和 } C_2^1) \text{ 且 } u(k) \text{ 是 } B^1,$$

$$\text{则 } x^{11}(k+1) = (1.5 - k_1^1)x(k) + (2.1 - k_2^1)x(k-1) \quad (21.17)$$

$$S^{12}: \text{如果 } x(k) \text{ 是 } (A_1^1 \text{ 和 } C_1^2) \text{ 且 } x(k-1) \text{ 是 } (A_2^1 \text{ 和 } C_2^2) \text{ 且 } u(k) \text{ 是 } B^1,$$

$$\text{则 } x^{12}(k+1) = (1.5 - k_1^2)x(k) + (2.1 - k_2^2)x(k-1) \quad (21.18)$$

$$S^{21}: \text{如果 } x(k) \text{ 是 } (A_1^2 \text{ 和 } C_1^1) \text{ 且 } x(k-1) \text{ 是 } (A_2^2 \text{ 和 } C_2^1) \text{ 且 } u(k) \text{ 是 } B^2,$$

$$\text{则 } x^{21}(k+1) = (0.3 + 0.5k_1^1)x(k) + (-3.4 + 0.5k_2^1)x(k-1) \quad (21.19)$$

$$S^{22}: \text{如果 } x(k) \text{ 是 } (A_1^2 \text{ 和 } C_1^2) \text{ 且 } x(k-1) \text{ 是 } (A_2^2 \text{ 和 } C_2^2) \text{ 且 } u(k) \text{ 是 } B^2,$$

$$\text{则 } x^{22}(k+1) = (0.3 + 0.5k_1^2)x(k) + (-3.4 + 0.5k_2^2)x(k-1) \quad (21.20)$$

动态方程可由式(21.8) ~ 式(21.10)得到。

因为图 21.1 中的闭环模糊控制系统等价于一个动态 TSK 模糊系统, 所以研究动态 TSK 模糊系统的稳定性很重要, 这也是下一节的研究主题。

21.3 动态 TSK 模糊系统的稳定性分析

考虑式(21.4)中 b^p 为零时的动态 TSK 模糊系统(21.5)。在这里假设 $b^p = 0$, 因为比较式(21.4)与式(21.7)可知, 式(21.7)中不包含 $b^p u(k)$ 这一项。定义状态变量为 $x(k) = (x(k), x(k-1), \dots, x(k-n+1))^T$ 和

$$A_p = \begin{bmatrix} a_1^p & a_2^p & \cdots & a_{n-1}^p & a_n^p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \quad (21.21)$$

则动态 TSK 模糊系统(21.5)可以改写成

$$\mathbf{x}(k+1) = \frac{\sum_{p=1}^N A_p \mathbf{x}(k) v^p}{\sum_{p=1}^N v^p} \quad (21.22)$$

式中, v^p 的定义由式(21.6)给出。因为当 $\mathbf{x}(k) = \mathbf{0}$ 时, 式(21.22)的右端等于零, 所以 R^n 中的原点是动态系统(21.22)的一个平衡点。下面应用众所周知的 Lyapunov 稳定性定理研究动态系统(21.22)的稳定性。

Lyapunov 稳定性定理: 考虑由下式描述的离散系统

$$\mathbf{x}(k+1) = f[\mathbf{x}(k)] \quad (21.23)$$

其中, $\mathbf{x}(k) \in R^n, f(\mathbf{0}) = \mathbf{0}$ 。假设存在一个标量函数 $V[\mathbf{x}(k)]$, 使得(a) $V(\mathbf{0}) = 0$, (b) $\mathbf{x}(k) \neq \mathbf{0}$ 时, $V[\mathbf{x}(k)] > 0$, (c) $\|\mathbf{x}(k)\| \rightarrow \infty$ 时, $V[\mathbf{x}(k)] \rightarrow \infty$, (d) $\mathbf{x}(k) \neq \mathbf{0}$ 时, $\Delta V[\mathbf{x}(k)] = V[\mathbf{x}(k+1)] - V[\mathbf{x}(k)] < 0$ 则系统(21.23)的平衡点 $\mathbf{0}$ 是全局渐进稳定的。

为了将此定理应用于系统(21.22), 需引入下面的引理。

矩阵不等式引理: 如果 P 是一个正定矩阵, 且满足

$$A^T P A - P < 0 \text{ 且 } B^T P B - P < 0 \quad (21.24)$$

其中, $A, B, P \in R^{n \times n}$, 则有

$$A^T P B + B^T P A - 2P < 0 \quad (21.25)$$

此引理的证明留作习题。应用 Lyapunov 稳定性定理和矩阵不等式引理, 可得到下面的动态 TSK 模糊系统(21.22)的稳定性定理。

定理 21.2 如果对于所有 $p = 1, 2, \dots, N$, 存在一个公共的正定矩阵 P 使得

$$A_p^T P A_p - P < 0 \quad (21.26)$$

则动态 TSK 模糊系统(21.22)的平衡点 $\mathbf{0}$ 是稳定的。

证明 考虑 Lyapunov 函数

$$V[\mathbf{x}(k)] = \mathbf{x}^T(k) P \mathbf{x}(k) \quad (21.27)$$

式中, P 为正定矩阵, $V[\mathbf{x}(k)]$ 满足 Lyapunov 稳定性定理中的条件(a) ~ (c)。下面将说明它也满足条件(d)。由式(21.22), 可得

$$\Delta V[\mathbf{x}(k)] = \mathbf{x}^T(k+1) P \mathbf{x}(k+1) - \mathbf{x}^T(k) P \mathbf{x}(k)$$

$$\begin{aligned}
&= \mathbf{x}^T(k) \left[\left(\frac{\sum_{p=1}^N A_p^T v^p}{\sum_{p=1}^N v^p} \right) P \left(\frac{\sum_{p=1}^N A_p v^p}{\sum_{p=1}^N v^p} \right) - P \right] \mathbf{x}(k) \\
&= \frac{\sum_{p=1}^N \sum_{q=1}^N v^p v^q \mathbf{x}^T(k) (A_p^T P A_q - P) \mathbf{x}(k)}{\sum_{p=1}^N \sum_{q=1}^N v^p v^q} \quad (21.28)
\end{aligned}$$

改变求和顺序,式(21.28)变为

$$\begin{aligned}
\Delta V[\mathbf{x}(k)] &= \left[\sum_{p=1}^N (v^p)^2 \mathbf{x}^T(k) (A_p^T P A_p - P) \mathbf{x}(k) + \sum_{p < q}^N v^p v^q \mathbf{x}^T(k) (A_p^T P A_p \right. \\
&\quad \left. + A_q^T P A_p - 2P) \mathbf{x}(k) \right] / \left[\sum_{p=1}^N \sum_{q=1}^N v^p v^q \right] \quad (21.29)
\end{aligned}$$

由式(21.26)、矩阵不等式引理和 $v^p \geq 0$, 可得 $\Delta V[\mathbf{x}(k)] < 0$ 。因此,此定理可由 Lyapunov 稳定性定理得证。

定理 21.2 给出了保证动态 TSK 模糊系统(21.22)稳定的充分条件。直观上,可以猜想,如果所有小区间内的近似线性化系统 A_p ($p = 1, 2, \dots, N$) 是稳定的(当 A_p 的所有特征值都在单位圆内时认为线性系统 $\mathbf{x}(k+1) = A_p \mathbf{x}(k)$ 是稳定的),则非线性系统(21.22)是稳定的。然而,这并不普遍成立。事实上,即使假设所有的 A_p 's(小)都稳定,也有可能不存在一个公共的正定矩阵使式(21.26)成立。下面的例子即说明了两个小区域内稳定的线性系统产生了一个不稳定的非线性系统的情况。

例 21.2 考虑一个用下述两条规则构造的动态 TSK 模糊系统:

$$\text{如果 } x(k-1) \text{ 是 } F_1, \text{ 则 } x^1(k+1) = x(k) - 0.5x(k-1) \quad (21.30)$$

$$\text{如果 } x(k-1) \text{ 是 } F_2, \text{ 则 } x^2(k+1) = -x(k) - 0.5x(k-1) \quad (21.31)$$

其中,模糊集 F_1 和 F_2 的隶属度函数如图 21.2 所示。对于此系统,有

$$A_1 = \begin{bmatrix} 1 & -0.5 \\ 1 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} -1 & -0.5 \\ 1 & 0 \end{bmatrix} \quad (21.32)$$

A_1 和 A_2 的特征值分别为 $\frac{1 \pm j}{2}$ 和 $\frac{-1 \pm j}{2}$, 均在单位圆内,所以两个局部逼近的线性系统

$\mathbf{x}(k+1) = A_1 \mathbf{x}(k)$ 和 $\mathbf{x}(k+1) = A_2 \mathbf{x}(k)$ 是稳定的。而由规则(21.30)和(21.31)构造的动态 TSK 模糊系统

$$\mathbf{x}(k+1) = \frac{A_1 \mathbf{x}(k) \mu_{F_1}(x(k-1)) + A_2 \mathbf{x}(k) \mu_{F_2}(x(k-1))}{\mu_{F_1}(x(k-1)) + \mu_{F_2}(x(k-1))} \quad (21.33)$$

则是不稳定的,其图形见图 21.3。图中的 $x(k)$ 是式(21.33)的结果,其初始条件为 $\mathbf{x}(1) = (x(1), x(0))^T = (-1.7, 1.9)^T$ 。

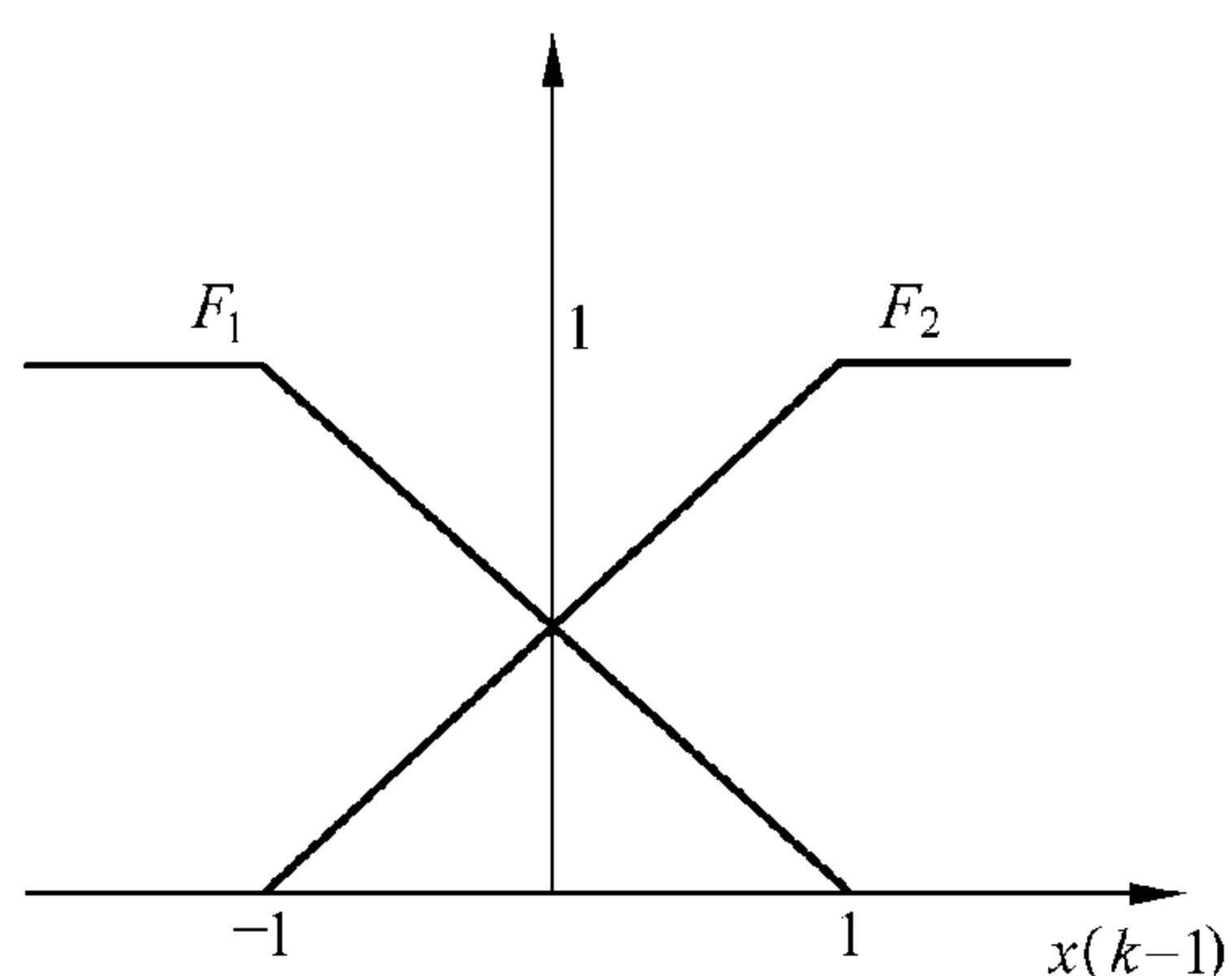
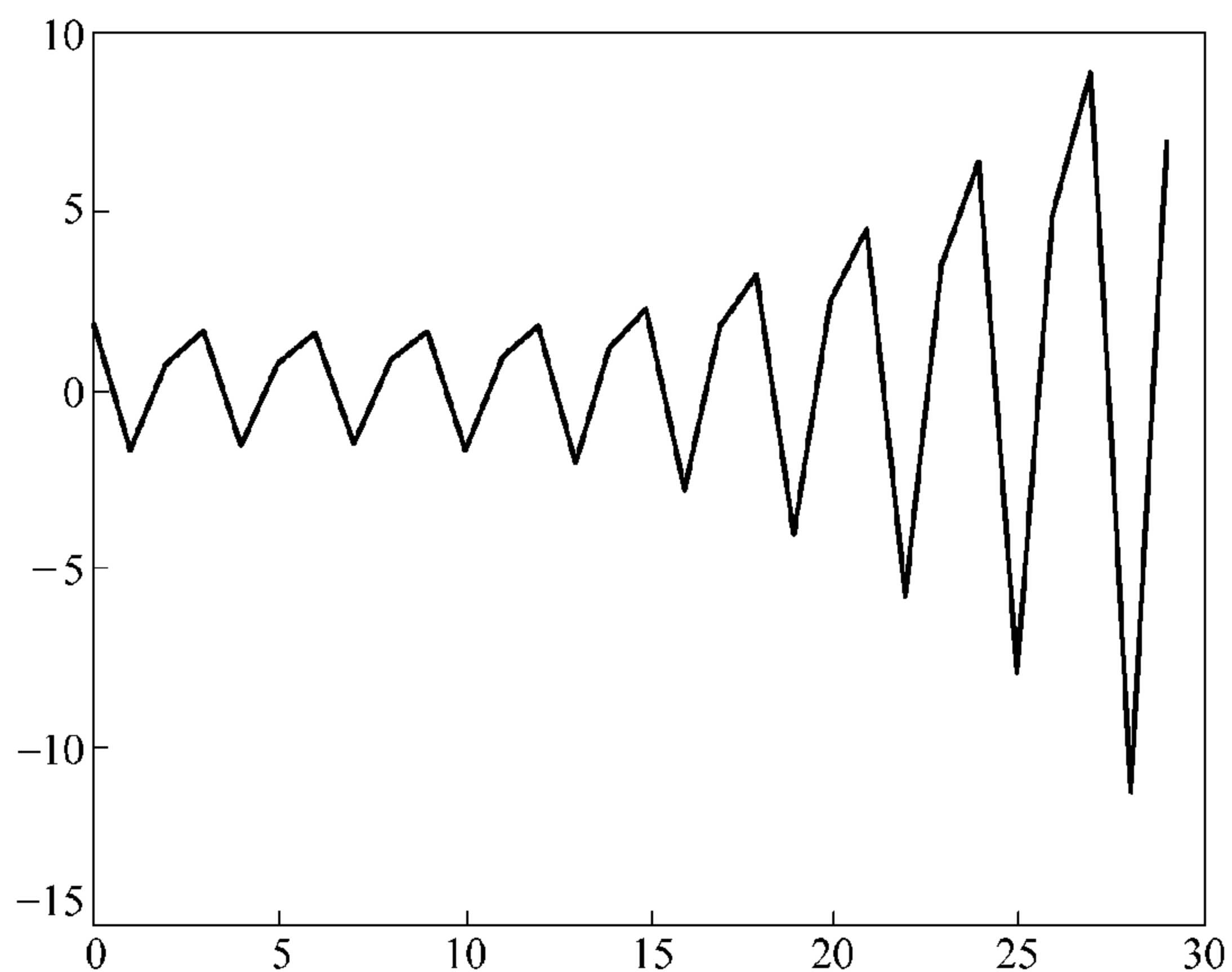


图 21.2 规则(21.30)和(21.31)中的模糊集的隶属度函数



注: 初始条件为 $\mathbf{x}(1) = (x(1), x(0))^T = (-1.7, 1.9)^T$ 。

图 21.3 动态 TSK 模糊系统(21.33)的响应

显然,因为最终的动态 TSK 模糊系统是不稳定的,所以例 21.2 中不存在使得式(21.26)成立的公共的 P 。下面给出保证存在公共的 P 的必要条件。

引理 21.1 假设 $A_p (p=1, 2, \dots, N)$ 是稳定的非奇异矩阵。如果存在一个公共正定矩阵 P 使得 $A_p^T P A_p - P < 0 (p=1, 2, \dots, N)$ 成立,则对于所有的 $p, q=1, 2, \dots, N$, 有 $A_p A_q$ 为稳定矩阵。

此引理的证明留作习题。引理 21.1 指出,如果任意一个 $A_p A_q$ 是不稳定矩阵,则公共的 P 不存在,进而可知动态 TSK 模糊系统可能是不稳定的。对于例 21.2, 有

$$A_1 A_2 = \begin{bmatrix} -1.5 & -0.5 \\ -1 & -0.5 \end{bmatrix} \quad (21.34)$$

其特征值是 $\frac{-2 \pm \sqrt{3}}{2}$, 其中一个在单位圆外。

没有一个通用的程序能保证找到这样一个公共的 P , 在下一节中将采用常用的试错法

来寻找公共的 P 。

21.4 模糊模型的稳定模糊控制器的设计

定理 21.2 给出了保证动态 TSK 模糊系统稳定的充分条件。定理 21.1 证明了图 21.1 中的闭环模糊控制系统可以表述成一个动态 TSK 模糊系统。因此,可以应用定理 21.2 来设计模糊系统模型的稳定模糊控制器,此即本节主题。

因为一种找到定理 21.2 中公共 P 的系统方法是不存在的,所以下面给出的设计程序本质上是试错法。

为模糊系统模型设计稳定模糊控制器

步骤 1 应用定理 20.1 可将闭环模糊控制系统表示成动态 TSK 模糊系统。系统中,参数 a_i^p 、 b^p 和隶属度函数 $\mu_{A_i^p}$ 已知,控制器参数(即 c_i^l 和 $\mu_{C_i^l}$)需要设计。通常令 $\mu_{C_i^l}$ 固定不变,根据定理 21.2 设计 c_i^l 。

步骤 2 根据式(21.7)选择参数 c_i^l 使所有的局部逼近线性系统是稳定的,则其局部逼近线性系统为 $\mathbf{x}(k+1) = A_{lp} \mathbf{x}(k)$, 且

$$A_{lp} = \begin{bmatrix} a_1^p + b^p c_1^l & a_2^p + b^p c_2^l & \cdots & a_{n-1}^p + b^p c_{n-1}^l & a_n^p + b^p c_n^l \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \quad (21.35)$$

其中, $l = 1, 2, \dots, M, p = 1, 2, \dots, N$ 。

步骤 3 对于 $l = 1, 2, \dots, M, p = 1, 2, \dots, N$, 找到一个正定矩阵 P_{lp} 使

$$A_{lp}^T P_{lp} A_{lp} - P_{lp} < 0 \quad (21.36)$$

如果对于固定的 $l^* \in \{1, 2, \dots, M\}$ 和 $p^* \in \{1, 2, \dots, N\}$, 存在一个 $P_{l^* p^*}$ ($l = 1, 2, \dots, M, p = 1, 2, \dots, N$) 使得

$$A_{lp}^T P_{l^* p^*} A_{lp} - P_{l^* p^*} < 0 \quad (21.37)$$

则可以选择此 $P_{l^* p^*}$ 作为公共的 P , 否则转到步骤 2, 重新设计参数 c_i^l 直至找到 $P = P_{l^* p^*}$ 。

由于在步骤 2 中寻找参数 c_i^l 时的自由度很大, 所以可能在迭代几次后才会找到公共的

P 。当然,我们不能保证一定会找到这样一个公共的 P 。下面通过一个例子来检验设计程序。

例 21.3 假设过程模型是由如下规则构造的动态 TSK 模糊系统模型:

$$\text{如果 } x(k) \text{ 是 } G_1, \text{ 则 } x^1(k+1) = 2.18x(k) - 0.59x(k-1) - 0.603u(k) \quad (21.38)$$

$$\text{如果 } x(k) \text{ 是 } G_2, \text{ 则 } x^2(k+1) = 2.26x(k) - 0.36x(k-1) - 1.120u(k) \quad (21.39)$$

且其控制器是由下面两条规则构造的 TSK 模糊系统:

$$\text{如果 } x(k) \text{ 是 } G_1, \text{ 则 } u^1(k) = c_1^1x(k) + c_2^1x(k-1) \quad (21.40)$$

$$\text{如果 } x(k) \text{ 是 } G_2, \text{ 则 } u^2(k) = c_1^2x(k) + c_2^2x(k-1) \quad (21.41)$$

其中, G_1 和 G_2 的隶属度函数见图 21.4, 需要设计的控制器参数 c_1^1 、 c_2^1 、 c_1^2 和 c_2^2 应能保证闭环系统的

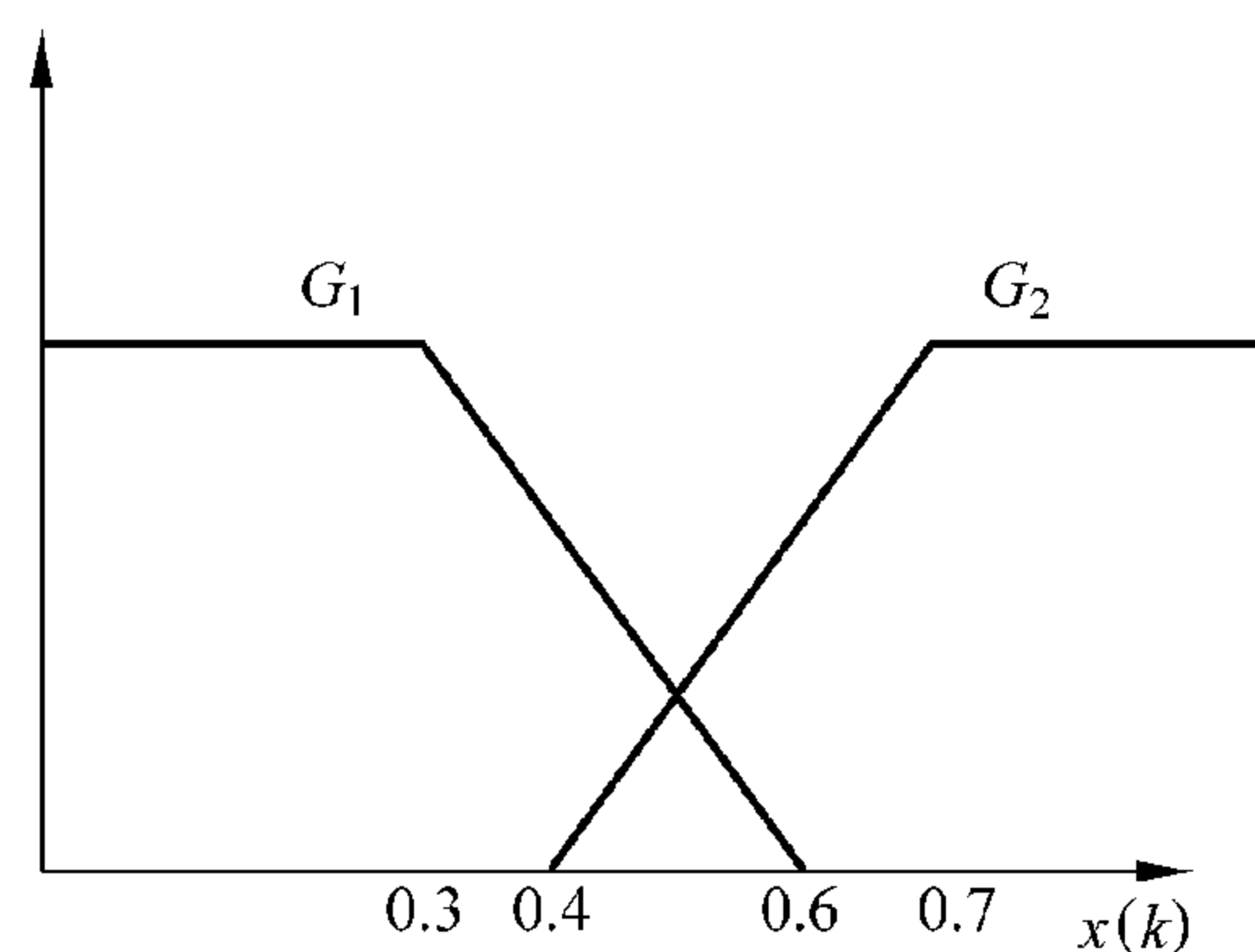


图 21.4 例 21.3 中模糊集的隶属度函数

稳定。

步骤 1 由定理 21.1 可知, 闭环系统是由下面四条规则构造的动态 TSK 模糊系统:

$$\begin{aligned} \text{如果 } x(k) \text{ 是 } (G_1 \text{ 和 } G_2), \text{ 则 } x^{11}(k+1) &= (2.18 - 0.603c_1^1)x(k) \\ &+ (-0.59 - 0.603c_2^1)x(k-1) \quad (21.42) \end{aligned}$$

$$\begin{aligned} \text{如果 } x(k) \text{ 是 } (G_1 \text{ 和 } G_2), \text{ 则 } x^{12}(k+1) &= (2.18 - 0.603c_1^2)x(k) \\ &+ (-0.59 - 0.603c_2^2)x(k-1) \quad (21.43) \end{aligned}$$

$$\begin{aligned} \text{如果 } x(k) \text{ 是 } (G_1 \text{ 和 } G_2), \text{ 则 } x^{21}(k+1) &= (2.26 - 1.120c_1^1)x(k) \\ &+ (-0.36 - 1.120c_2^1)x(k-1) \quad (21.44) \end{aligned}$$

$$\begin{aligned} \text{如果 } x(k) \text{ 是 } (G_1 \text{ 和 } G_2), \text{ 则 } x^{22}(k+1) &= (2.26 - 1.120c_1^2)x(k) \\ &+ (-0.36 - 1.120c_2^2)x(k-1) \quad (21.45) \end{aligned}$$

步骤 2 四个线性子系统(21.42) ~ (21.45)的四个矩阵 A_p 分别为

$$A_{11} = \begin{bmatrix} 2.18 - 0.603c_1^1 & -0.59 - 0.603c_2^1 \\ 1 & 0 \end{bmatrix} \quad (21.46)$$

$$A_{12} = \begin{bmatrix} 2.18 - 0.603c_1^2 & -0.59 - 0.603c_2^2 \\ 1 & 0 \end{bmatrix} \quad (21.47)$$

$$A_{21} = \begin{bmatrix} 2.26 - 1.120c_1^1 & -0.36 - 1.120c_2^1 \\ 1 & 0 \end{bmatrix} \quad (21.48)$$

$$A_{22} = \begin{bmatrix} 2.26 - 1.120c_1^2 & -0.36 - 1.120c_2^2 \\ 1 & 0 \end{bmatrix} \quad (21.49)$$

通过反复试错,可发现,如果选择

$$c_1^1 = 1.564, c_2^1 = -0.223, c_1^2 = 0.912, c_2^2 = 0.079 \quad (21.50)$$

则以上四个线性子系统都是稳定的。

步骤3: 结果表明,可为式(21.50)中的参数找到一个公共的 P 。因此,这里的稳定模糊控制器是由规则(21.40)~(21.41)两条规则构造的 TSK 模糊系统,其参数由式(21.50)给出。

21.5 本章小结及导读

本章主要阐述了以下内容:

- 静态 TSK 模糊系统与动态 TSK 模糊系统;
- 闭环系统的动态方程,其过程模型为一个动态 TSK 模糊系统模型,其控制器为一个静态 TSK 模糊系统;
- 确保上述闭环系统稳定的条件;
- 如何设计使得上述闭环系统稳定的模糊控制器。

Takagi 和 Sugeno[1985]与 Sugeno 和 Kang[1988]研究了利用 TSK 模糊系统建立非线性系统模型的问题。本章中对稳定性的分析可参见 Tanaka 和 Sugeno[1992],从中可以找到大量例子。

21.6 习题

习题 21.1 证明矩阵不等式引理。

习题 21.2 如果将 Lyapunov 稳定性定理中的条件(d)改为:在 0 的邻域内,对于 $x(k)$ 有 $\Delta V[x(k)] = V[x(k+1)] - V[x(k)] < 0$,则结论变为,该系统是渐近稳定的。请证明

系统

$$x_1(k+1) = x_1(k)[x_1^2(k) + x_2^2(k)] \quad (21.51)$$

$$x_2(k+1) = x_2(k)[x_1^2(k) + x_2^2(k)] \quad (21.52)$$

是渐近稳定的。

习题 21.3 证明引理 21.1。

习题 21.4 考虑 $N=2$ 时的动态 TSK 模糊系统(21.22)。假设 A_1 和 A_2 是稳定的且 $A_1A_2 = A_2A_1$ 。给定 $P_0 > 0$, 根据

$$A_1^T P_1 A_1 - P_1 = -P_0 \quad (21.53)$$

$$A_2^T P_2 A_2 - P_2 = -P_1 \quad (21.54)$$

可确定矩阵 P_1 和 P_2 。证明 P_2 是定理 21.2 中的公共矩阵 P 。

习题 21.5 将习题 21.4 中找到公共矩阵 P 的程序推广到 N 为任意值的情况。

习题 21.6 假设所要控制的过程模型是由规则(21.38)和规则(21.39)构造而成的一个动态 TSK 模糊系统模型。设计一个线性控制器 $u(k) = Kx(k)$ (即, 确定常数 K)使得闭环系统稳定。

习题 21.7 将图 21.2 中的模糊集 F_1 和 F_2 分别替换成习题 21.6 中的模糊集 G_1 和 G_2 , 重做习题 21.6。

模糊控制的定性分析和多级模糊系统

22.1 模糊控制系统的相平面分析

相平面分析是研究二阶系统的图形方法。其基本思想是为状态空间中的系统生成对应于不同初始条件的运动轨道,然后检查轨迹的性能。利用这种方法,可以得到系统的稳定性信息和稳健性信息及其他性能信息。二阶系统的状态空间称为相平面。

相平面分析的主要优点是它的图形特征,它允许在不求方程解析解的情况下看到不同初始条件下非线性系统的运动走向。其主要的缺点是严格限于二阶系统,很难推广到高阶系统中去。

下面应用相平面法分析二阶模糊控制系统,令 $x = (x_1, x_2)^T$ 为状态变量,考虑模糊控制系统

$$\dot{x} = f(x) + bu \quad (22.1)$$

$$u = \Phi(x) \quad (22.2)$$

其中, $f(x)$ 是非线性向量函数(向量域),代表研究对象的动态, b 是二维向量, u 是标量型控制变量, $\Phi(x)$ 是一个两维输入一维输出的模糊系统。因此,研究此闭环模糊控制系统,并识别某一轨迹上有哪些规则起作用是有意义的。假定定义 N_1 个和 N_2 个模糊集分别覆盖 x_1 和 x_2 所在的区域,且 $\Phi(x)$ 的模糊规则库由 $N_1 \times N_2$ 条规则组成。令规则库中的第 l 条规则为

$$\text{如果 } x_1 \text{ 是 } A_1^l, \text{ 且 } x_2 \text{ 是 } A_2^l, \text{ 则 } y \text{ 是 } B^l \quad (22.3)$$

如果对于所有的 $l \neq l_*$ 有下式成立,则认为相平面中的点 (x_1, x_2) “属于规则 l_* ”

$$\begin{aligned} & \mu_{A_1^*}(x_1) \star \mu_{A_2^*}(x_2) \\ & \geq \mu_{A_1^l}(x_1) \star \mu_{A_2^l}(x_2) \end{aligned} \quad (22.4)$$

式中, \star 表示 t -范数。

考虑图 22.1 中的模糊规则库,其中 $N_1 = N_2 = 5$,该规则库包含 25 条规则。如图 22.1 所示,有意义的状态空间可被分成 25 个区域,每个区域归属于一条规则。对于一个给定的初始条件,闭环系

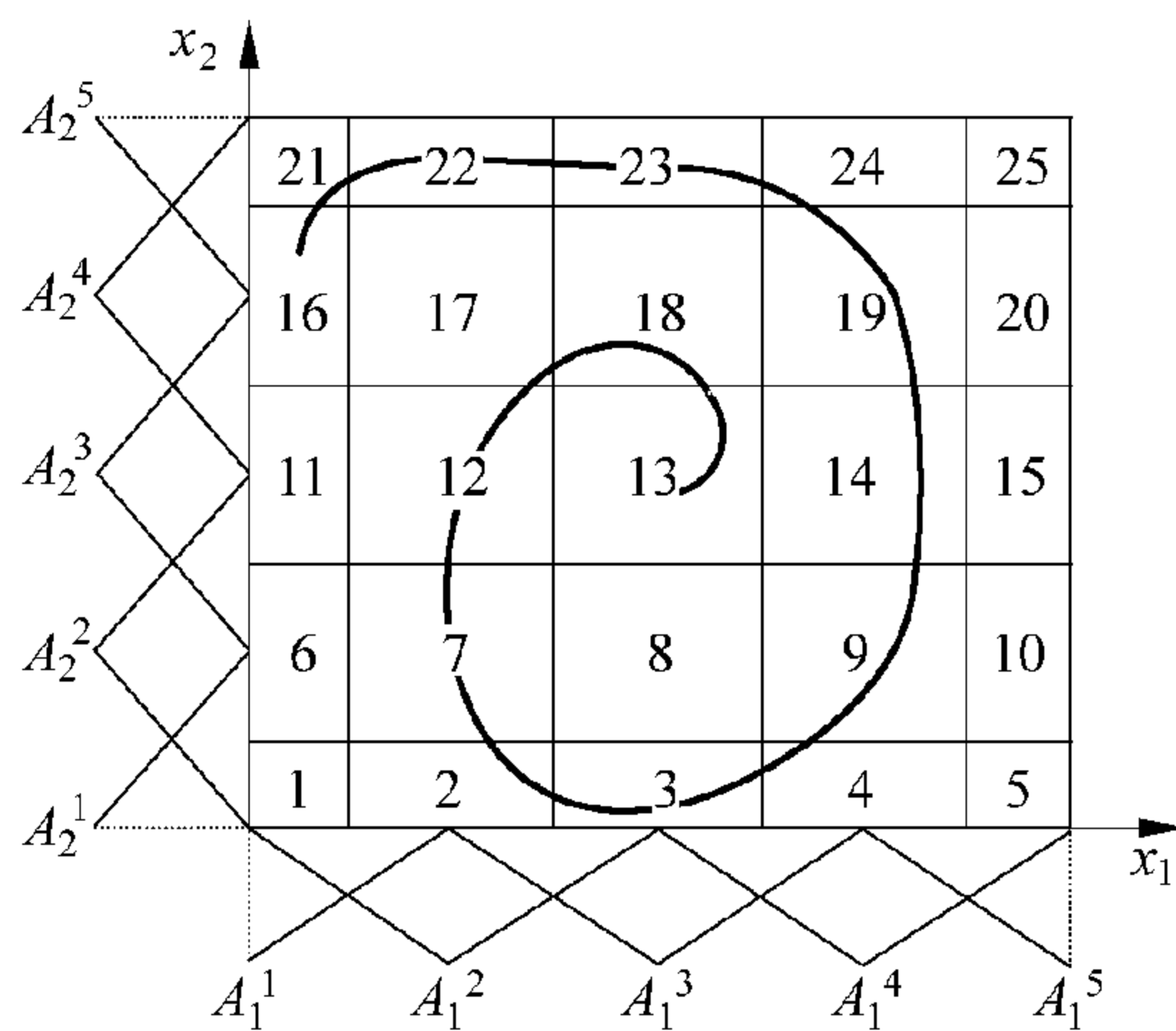


图 22.1 语言轨迹示例

统的轨迹式(22.1)~式(22.2)可映射到图 22.1 中的被分割的状态空间中。一条轨迹对应一系列沿轨迹方向起作用的规则,这一系列规则称为语言轨迹,它与状态轨迹相对应。对应于图 22.1 中轨迹的语言轨迹为

$$\text{语言轨迹} = (\text{规则 16, 规则 21, 规则 22, 规则 23, 规则 24, 规则 19, 规则 14, 规则 9, 规则 4, 规则 3, 规则 2, 规则 7, 规则 12, 规则 17, 规则 18, 规则 13}) \quad (22.5)$$

通过修正语言轨迹中的规则,可以改变相应的状态轨迹。例如,如果认为图 22.1 中状态轨迹在初始阶段收敛太慢,则可以修正规则 22 和规则 23,使收敛速度快些。应注意,状态轨迹的切线(tangent vector)等于向量域 $f(x)$ 和向量域 $b\Phi(x)$ 之和(见图 22.2),然后可以再修正上述规则。如果图 22.2 中的点 x 属于规则 23,则增加规则 23 中的 THEN 部分模糊集的中心值,从而增加此点的 $b\Phi(x)$ 值,也因此提高了收敛速度。相平面分析的优点是,能帮助识别影响运动轨迹的规则,从而只需通过修正这些规则就可得到期望轨迹。

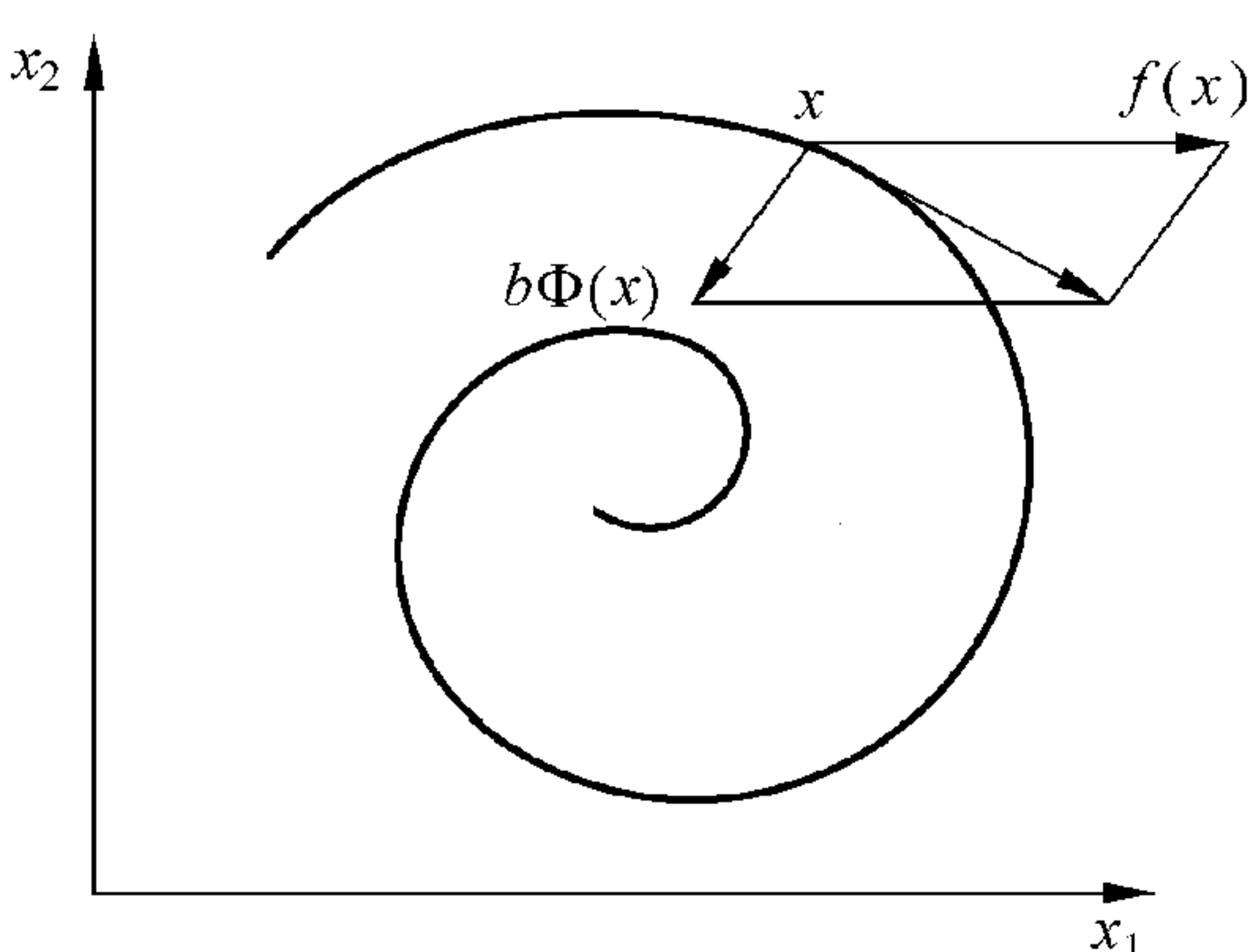


图 22.2 沿向量 $f(x) + b\Phi(x)$ 向移动的状态轨迹

对于闭环模糊控制系统(22.1)和(22.2),应该特殊关注其子空间 $\Phi(x) = 0$ 的情况,它在相平面中是一条线。此线将相平面分隔成正、负控制域两部分。此线也叫做开关线(switching line),因为每当状态轨迹跨越此线时,控制会由正变负,反之亦然。所以,当状态向量远离开关线时,控制向量 $b\Phi(x)$ 通常比对象的 $f(x)$ 部分对闭环系统的影响大。在状态向量趋近于开关线时, $b\Phi(x)$ 值变小,于是 $f(x)$ 对闭环系统的影响更大, $b\Phi(x)$ 和 $f(x)$ 之间的关系决定了闭环系统的行为。最有意义的两种情况是:

- 稳定的闭环系统。通常来说,开环系统 $\dot{x} = f(x)$ 稳定,控制 $u = \Phi(x)$ 尽量引领系统轨迹靠近开关线 $\Phi(x) = 0$ 。一旦轨迹接近开关线,对象的 $f(x)$ 将有较大影响力,它促使轨迹收敛于平衡点,如图 22.3 中的示例。
- 有界循环(limit cycles)。这可能是开环系统 $\dot{x} = f(x)$ 不稳定,而控制 $u = \Phi(x)$ 努力使系统稳定的情况。当状态远离开关线 $\Phi(x) = 0$ 时,控制 $\Phi(x)$ 的影响力大,使状态轨迹趋向于开关线。当状态轨迹接近开关线时,不稳定的对象因素 $f(x)$ 的影响力变得更大些,使状态轨迹远离平衡点,控制与对象因素的交互作用使状态在平衡点周围振荡,这样,有界循环就形成了,如图 22.4 中的示例。

虽然相平面分析是定性的,但它对表征一个模糊控制系统的动态行为是有帮助的。它还能被用作适当选择或修改规则的依据。

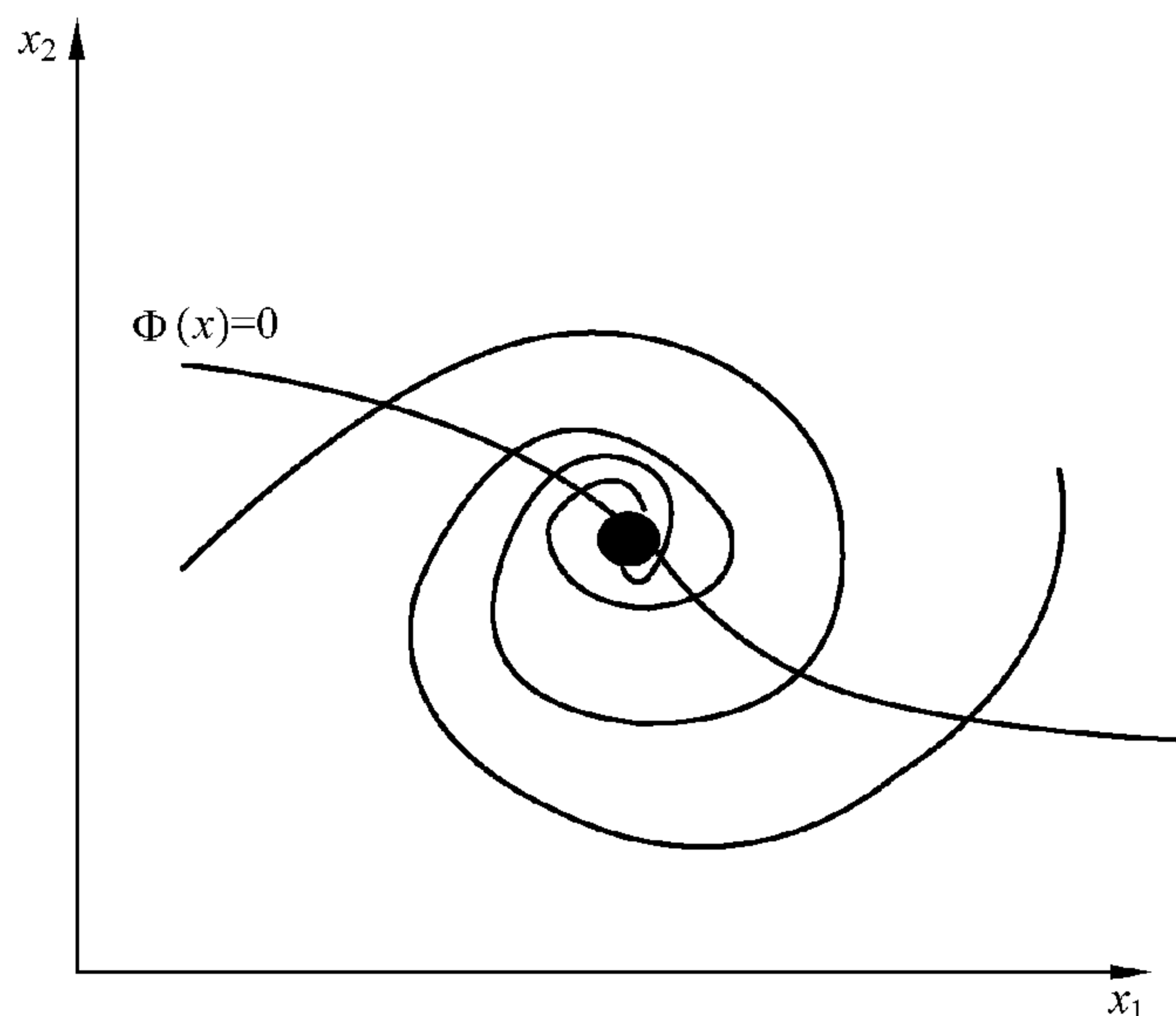


图 22.3 稳定闭环系统示例

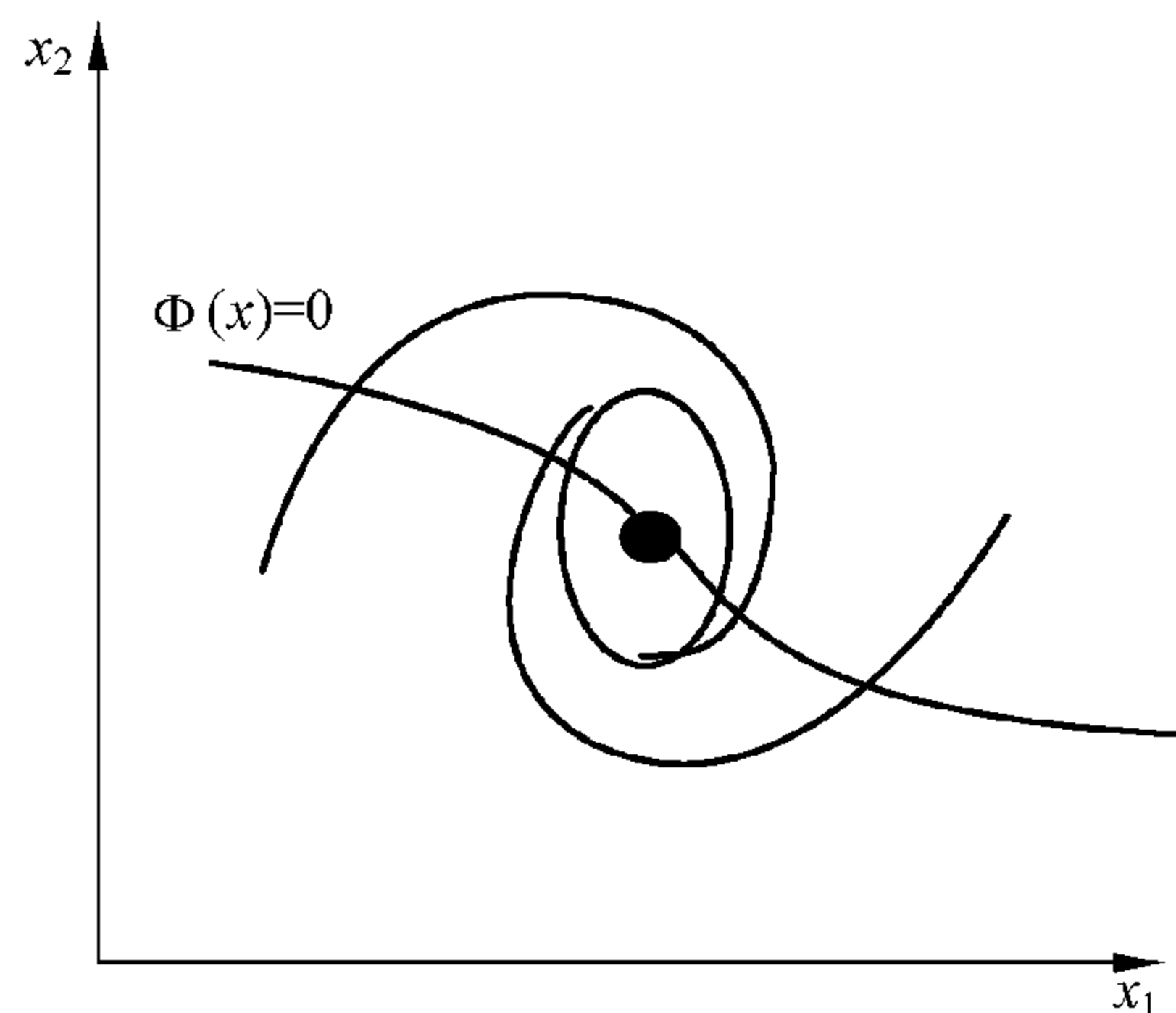


图 22.4 有界循环示例

22.2 稳定性的稳健性指标

众所周知,如果一个闭环系统的特征值都在复平面的左半开平面上,则此线性系统是稳定的。如果特征值靠近虚轴,则参数的微小变动就能使一些特征值移到右半开平面,从而引起不稳定。因此,特征值离虚轴越远,稳定性的稳健性越好。最小特征值距虚轴的距离可作为系统稳定性的稳健性度量指标。本节将对模糊控制系统进行此类分析并建立一些稳定性的稳健性指标。这里先借助一维情况来介绍稳健性的基本概念,然后将其扩展到 n 维情况。

22.2.1 一维情况

考虑 $x \in R$ 为标量且 $b = 1$ 时的模糊控制系统(22.1) ~ (22.2)。假设 $f(x)$ 是单调递增函数, $f(0) = 0$, 模糊控制器 $\Phi(x)$ 满足 $\Phi(0) = 0$, 则闭环系统的平衡点可由下式确定

$$\dot{x} = f(x) + \Phi(x) = 0 \quad (22.6)$$

因为 $f(0) = \Phi(0) = 0$, 所以原点是一个平衡点。此平衡点稳定的一个充分条件是

$$\frac{d}{dx}[f(x) + \Phi(x)]|_{x=0} = f'(0) + \Phi'(0) < 0 \quad (22.7)$$

上式可由 Lyapunov 线性化定理得证(Vidyasagar[1993])。因此,如果下述两个条件得到满足,则闭环系统(22.1) ~ (22.2)是全局稳定的:

$$\text{条件 1: } f'(0) + \Phi'(0) < 0 \quad (22.8)$$

$$\text{条件 2: } |\Phi(x)| < |f(x)|, \forall x \neq 0 \quad (22.9)$$

条件 1 保证原点是稳定平衡点,而条件 2 保证了曲线 $\Phi(x)$ 和 $-f(x)$ 不相交,从而阻止了其他平衡点的出现。

一旦稳定性失去,则要么原点这一平衡点变得不稳定,要么根据 $\Phi(x)$ 和 $-f(x)$ 的相交产生新的平衡点。由条件(22.8)可看出,数

$$I_1 = -(f'(0) + \Phi'(0)) \quad (22.10)$$

可用来度量平衡点原点的稳定性的稳健性。 I_1 的值越大,原点的稳定性的稳健性越好。式(22.10)中的 I_1 是给出的模糊控制系统的稳定性的第一个稳健性指标。

类似地,可以给出与条件 2 相关联的稳健性度量指标。应用 $\Phi(x)$ 和 $-f(x)$ 间的最小距离,即 $\min |\Phi(x) + f(x)|$ 来度量稳健性。但因为 $x=0$ 时, $\min |\Phi(x) + f(x)|$ 为零,所以应在原点的一个邻域内找 \min ,这就是稳定性的第二条稳健性指标:

$$I_2 = \min_{x \in R - (-\alpha, \alpha)} |\Phi(x) + f(x)| \quad (22.11)$$

其中, α 是一个正常数。 I_2 的值越大,模糊控制系统的稳定性的稳健性越好。

22.2.2 n 维情况

这里由 $n=2$ 开始,推导 n 为任意值的情况。假设 $f(0)=0$,且 $\Phi(0)=0$,则原点是闭环系统(22.1)~(22.2)的一个平衡点。由 Lyapunov 线性化定理(Vidyasagar[1993])可知,如果在原点附近的线性化系统的两个特征值实部为负,则原点稳定。通常,在以下两种情况下系统会变得不稳定:

- 一个实特征值穿越虚轴变为正,这称作静态分岔点(Static bifurcation)。
- 一对复特征值穿越虚轴,二者有正实部,这称作 Hopf 分岔点(Hopf bifurcation)。

下面检验何种情况下会出现这两种分岔点。

令

$$J = \frac{\partial(f(x) + b\Phi(x))}{\partial x} \Big|_{x=0} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad (22.12)$$

为原点处 $f(x) + \Phi(x)$ 的雅可比矩阵(Jacobian matrix),则在原点附近的线性化系统的特征值为特征方程:

$$|sI - J| = s^2 - (a_{11} + a_{22})s + a_{11}a_{22} - a_{12}a_{21} = 0 \quad (22.13)$$

的解。静态分岔点在一个实特征值穿越虚轴时产生。即,在方程(22.13)的一个根为零时产生。由式(22.13)可以看出,只要 $a_{11}a_{22} - a_{12}a_{21} = 0$,就会产生静态分岔点。 $|a_{11}a_{22} - a_{12}a_{21}|$ 的值越大,系统越远离静态分岔点。因此,定义

$$I_1 = |a_{11}a_{22} - a_{12}a_{21}| = |\det(J)| \quad (22.14)$$

为二阶模糊控制系统稳定性的稳健性指标。类似地,当两个复特征值的实部为零时,会产生 Hopf 分岔点。由方程(22.13)可以看出,Hopf 分岔点只在 $a_{11} + a_{22} = 0$ 时产生。 $|a_{11} +$

a_{22} 的值越大, Hopf 分歧点越不可能产生。因此, 定义

$$I'_1 = |a_{11} + a_{22}| = |\text{tr}(J)| \quad (22.15)$$

为另一稳健性指标。总之, I_1 和 I'_1 的值越大, 模糊控制系统的稳健性越好。

与一维情况相似, 当模糊控制器的向量域 $b\Phi(x)$ 正好补偿了对象向量域 $f(x)$ 时, 可能失去稳定性。令 $b = (b_1, b_2)$, $f(x) = (f_1(x), f_2(x))$, 则仅当状态空间区域中对象的方向为 (b_1, b_2) 时, 对象与控制器向量域的补偿才能发生。即, 定义此区域为

$$C = \left\{ x \in R^2 \mid \frac{f_1(x)}{b_1} = \frac{f_2(x)}{b_2} \right\} \quad (22.16)$$

与一维情况(22.11)中的 I_2 相仿, 定义稳健性指标为

$$I_2 = \min_{x \in C-B} |f(x) + b\Phi(x)| \quad (22.17)$$

其中, $B = \{x \in R^2 \mid |x|^2 \leq \alpha\}$ 是围绕原点的一个球。

对于 $n > 2$ 的一般情况, 令

$$J = \frac{\partial}{\partial x} [f(x) + b\Phi(x)] \Big|_{x=0} \quad (22.18)$$

为原点附近闭环系统的雅可比矩阵, 且

$$P(s) = |sI - J| = s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n \quad (22.19)$$

为其特征多项式。与 $n = 2$ 的情况相似, $a_n = 0$ 时可产生静态分歧点。因此, 定义

$$I_1 = a_n = |\det(J)| \quad (22.20)$$

为一个稳健性指标。与式(22.15)中的 I'_1 相似, 可根据 Hopf 分歧点定义稳健性的另一个指标。最后, 将式(22.17)中的 I_2 推广到 $n > 2$ 的情况即可得到该稳定性指标, 只不过此时式(22.16)中的子空间 C 变成了

$$C = \left\{ x \in R^n \mid \frac{f_1(x)}{b_1} = \frac{f_2(x)}{b_2} = \cdots = \frac{f_n(x)}{b_n} \right\} \quad (22.21)$$

例 22.1 考虑模糊控制系统

$$\dot{x} = x_2 \quad (22.22)$$

$$\dot{x}_2 = -12.74x_1 - 2.22x_2 + 2.22x_1^2x_2 + 12.74\Phi(x_1, x_2) \quad (22.23)$$

此系统在原点的雅可比矩阵为

$$J = \begin{pmatrix} -12.74 + 12.74 \frac{\partial^0 \Phi(x_1, x_2)}{\partial x_1} \Big|_{x_1=x_2=0} & -2.22 + 12.74 \frac{\partial^1 \Phi(x_1, x_2)}{\partial x_2} \Big|_{x_1=x_2=0} \end{pmatrix} \quad (22.24)$$

如果没有控制, 即如果 $\Phi(x_1, x_2) = 0$, 则稳健性指标 I_1 和 I'_1 为

$$I_1 = 12.74, I'_1 = 2.22 \quad (22.25)$$

如果用模糊控制器 $\Phi(x_1, x_2)$, 则上述稳健性指标变为

$$I_1 = \left| -12.74 + 12.74 \frac{\partial \Phi(x_1, x_2)}{\partial x_1} \Big|_{x_1=x_2=0} \right|,$$

$$I'_1 = \left| -2.22 + 12.74 \frac{\partial \Phi(x_1, x_2)}{\partial x_2} \right|_{x_1=x_2=0} \quad (22.26)$$

因此,如果设计模糊控制器使 $\frac{\partial \Phi(x_1, x_2)}{\partial x_1} \Big|_{x_1=x_2=0} < 0$ 和 $\frac{\partial \Phi(x_1, x_2)}{\partial x_2} \Big|_{x_1=x_2=0} < 0$ 同时成立,则闭环模糊控制系统将比不加控制的开环系统更加稳健。

22.3 多级模糊控制

22.3.1 维度灾难

在第 10 章和第 11 章中可以看到,要使设计的模糊系统达到所需精度,规则数量将不得不随模糊系统输入变量数目的增加而指数性地上升。具体地讲,假设有 n 个输入变量,每个输入变量定义为 m 个模糊集,则模糊系统的规则数为 m^n 。当 n 值较大时, m^n 将是一个巨大的数。在实际中,有五个输入变量并不鲜见。当 $n=5, m=3$ (通常,每个变量上至少要定义三个模糊集)时, $m^n=243$; 如果 $m=5$, 这比 $m=3$ 更接近实际,则 $m^n=3120$ 。在一个模糊系统中要采用数千条规则是不现实的,所以模糊系统应用所面临的一个严重的考验就是,如何处理这一规则爆炸问题。

事实上,问题的复杂度随涉及变量数目的增加而指数性上升是一个常见现象,这不为模糊系统所独有。Bellman 称这种现象为“维度灾难”。已经提出了几种解决此难题的方法,其中包括使用多级模糊系统。下面将会看到,多级模糊控制系统是具有良好性能的,其构造模糊系统所需的规则数仅与变量数线性相关。

22.3.2 多级模糊系统的构造

多级模糊系统的思想是将输入变量变成一个低维模糊系统集,而不是通常情况下的单级高维模糊系统。每一低维模糊系统构成多级模糊系统的一级。假设有 n 个输入变量 x_1, \dots, x_n , 则可构造如下多级模糊系统:

- 第一级模糊系统有 n_1 个输入变量 x_1, \dots, x_{n_1} , 它是由下面的规则构造而成的:

$$\text{如果 } x_1 \text{ 是 } A_1^l, \text{ 且 } \dots \text{ 且 } x_{n_1} \text{ 是 } A_{n_1}^l, \text{ 则 } y_1 \text{ 是 } B_1^l \quad (22.27)$$

其中, $2 \leq n_1 < n, l=1, 2, \dots, M_1$ 。

- 第 i 级 ($i > 1$) 是一个有 $n_i + 1$ ($n_i \geq 1$) 个输入变量的模糊系统, 它是由下面的规则构造而成的:

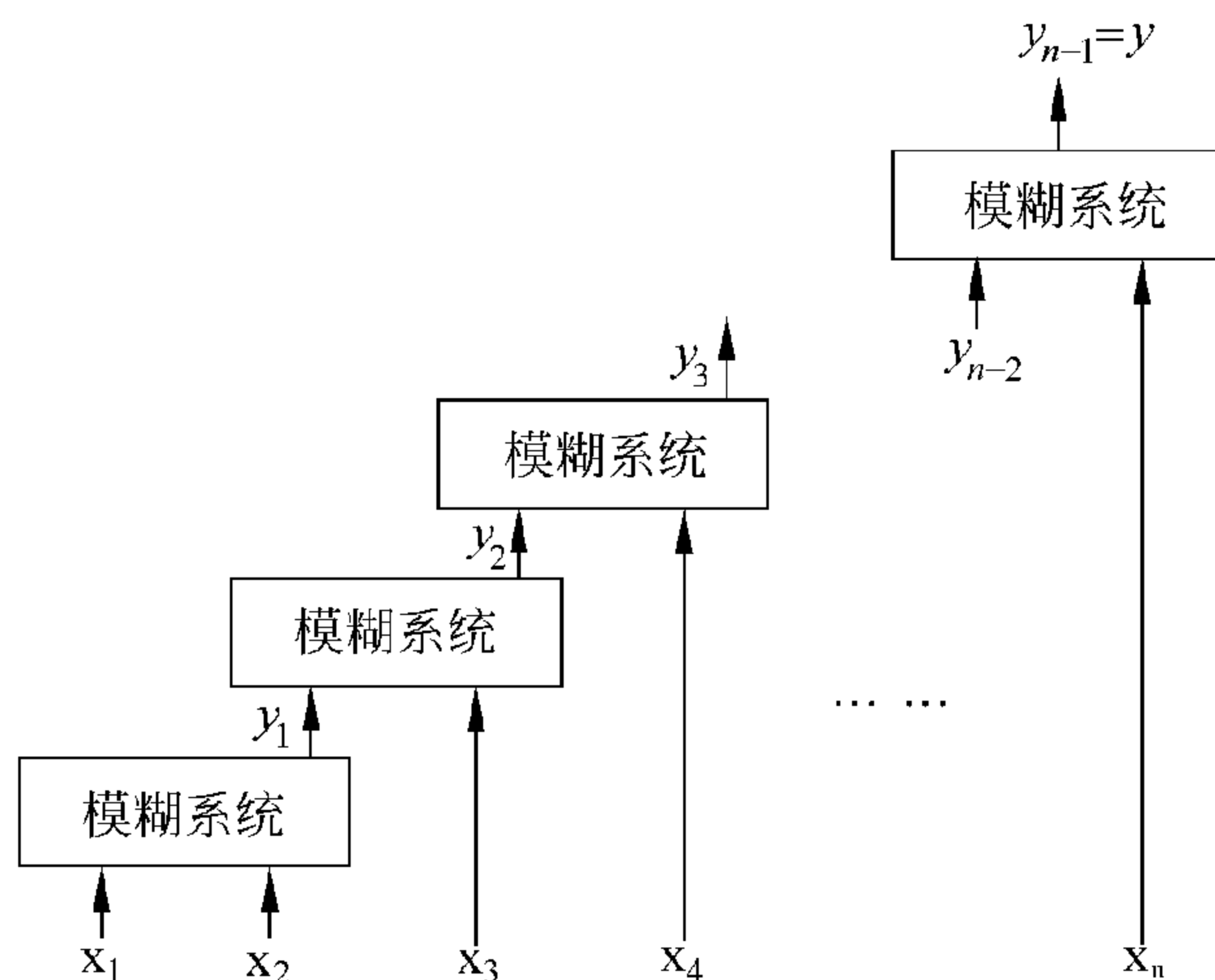
$$\text{如果 } x_{N_i+1} \text{ 是 } A_{N_i+1}^l, \text{ 且 } \dots \text{ 且 } x_{N_i+n_i} \text{ 是 } A_{N_i+n_i}^l, \text{ 且 } y_{i-1} \text{ 是 } C_{i-1}^l, \text{ 则 } y_i \text{ 是 } B_i^l \quad (22.28)$$

其中, $N_i = \sum_{j=1}^{i-1} n_j, l=1, 2, \dots, M_i$ 。

- 直至 $i = L$ 满足 $\sum_{j=1}^L n_j = n$ 时, 多级模糊系统的构造约束, 即直到所有变量都在某一级中得到使用为止。

可以看到第一级中的 n_1 个变量 x_1, \dots, x_{n_1} 变成一个变量 y_1 送入第二级中。在第二级中另外 n_2 个变量 $x_{n_1+1}, \dots, x_{n_1+n_2}$ 和变量 y_1 合并成另一个变量 y_2 被送到第三级中。此过程延续至所有变量 x_1, \dots, x_n 都被使用为止。

多级模糊系统的一个特例是选择 $n_1 = 2, n_i = 1, i = 2, 3, \dots, L$ 。在这种情况下, 多级结构中的所有模糊系统都有两个输入, 共有 $L = n - 1$ 级。这种特殊的多级模糊系统的示例见图 22.5。



注: $n_1 = 2, n_i = 1, i = 2, 3, \dots, n - 1$ 本多级模糊系统由 $n - 1$ 个二输入模糊系统构成。

图 22.5 多级模糊系统示例

下面将讨论多级模糊系统的性质。

22.3.3 多级模糊系统的性质

下面证明, 多级模糊系统的规则数是输入变量数的线性函数(定理 22.1)和图 22.5 的示例中的规则数会达到最小(定理 22.2)。

定理 22.1 令 n, n_i 和 L 在设计多级模糊系统时指代同样的变量。假定为包括输入变量 x_1, \dots, x_n 和中间变量 y_1, \dots, y_n 在内的每一个变量, 都定义了 m 个模糊集, 第 i 级 ($i = 2, 3, \dots, L$) 模糊系统由 m^{n_i+1} 条规则构成(完备的模糊规则库), 而第一级有 m^{n_1} 条规则。如果对于 $i = 2, 3, \dots, L$ 有 $n_1 = n_i + 1 = C$ (常数), 则整个多级模糊系统的总体规则数为

$$M = \frac{m^c}{c-1} (n-1) \quad (22.29)$$

证明 显然可以得到

$$M = m^{n_1} + \sum_{i=2}^L m^{n_i+1} = Lm^c \quad (22.30)$$

由 $n = \sum_{i=1}^L n_i = c + \sum_{i=2}^L (c-1) = Lc - L + 1$, 可得

$$L = \frac{n-1}{c-1} \quad (22.31)$$

将式(22.31)代入式(22.30), 即可得到式(22.29)。

由于 $\frac{m^c}{c-1}$ 是一个常数, 所以由式(22.29)可知, 多级模糊系统的规则数随输入变量数的增加而线性增大。例如, 当 $m=3, c=2, n=5$ 时, 规则数为 $M=3^2 \times 4=36$; 如果利用传统的模糊系统, 则规则数为 $m^n=3^5=243$ 。当 m 与 n 量值较大时, 这种减少更为显著。

最后, 需要指出的是, 当 $c=2$ 时, 规则数目 M 会达到最小值。

定理 22.2 令定理 22.1 中的假设均成立。如果 $m \geq 2$, 则在 $c=2$ 时, 即, 当模糊系统的各级都有两个输入时(如图 22.5 所示), 式(22.29)中的规则数 M 会达到最小。

此定理的证明留作习题。

22.4 本章小结及导读

本章主要阐述了以下内容:

- 二阶模糊控制系统的相平面分析, 包括语言轨迹的概念和如何利用它修改规则;
- 各种稳健性指标的定义和涵义;
- 多级模糊系统的构造及其基本性质。

对模糊控制系统进行语言分析的概念是由 Braae 和 Rutherford[1979]提出来的, 模糊控制系统的稳定性指标在 Aracil, Ollero 和 Garcia-Cerezo[1989]中有所研究。Driankov, Hellendoorn 和 Beinfrank[1993]的书中也包含有对模糊控制系统的定性分析和对稳健性指标的讨论。多级模糊系统由 Raju, Zhou 和 Kisner[1991]提出。Wang[1996]证明了多级模糊系统也是万能逼近器。

22.5 习题

习题 22.1 说明非线性系统

$$\dot{x} = x_2 - x_1(x_1^2 + x_2^2 - 1) \quad (22.32)$$

$$\dot{x}_2 = -x_1 - x_2(x_1^2 + x_2^2 - 1) \quad (22.33)$$

无论其初始状态如何都收敛于一个有界循环。

习题 22.2 为图 1.9 中的倒立摆系统设计一个由 25 条规则构造而成的模糊控制器, 并确定初始状态为 $(\theta(0), \dot{\theta}(0)) = (10^\circ, 0)$, $(5^\circ, 0)$ 和 $(-8^\circ, 0)$ 时的语言轨迹。

习题 22.3 为图 16.7 中的球—横木系统设计一个 25 条规则的模糊控制器, 并确定初始状态为 $(r(0), \theta(0)) = (1, 0^\circ)$, $(1, 10^\circ)$ 和 $(-1, 5^\circ)$ 时的语言轨迹。

习题 22.4 为例 22.1 设计一个模糊系统 $\Phi(x_1, x_2)$, 使闭环系统的稳健性优于开环系统的稳健性。

习题 22.5 考虑 $n=3$ 时, 图 22.5 所示的多级模糊系统。设 $g(x_1, x_2, x_3)$ 是一个未知函数, 但它在 $(x_1, x_2, x_3) = (e_i, e_j, e_k)$, $i, j, k = 1, 2, \dots, 6$ 时的值是已知的, 其中 $e_i = 0.2(i-1)$ 。令 $f_1(x_1, x_2)$ 和 $f_2(y_1, x_3)$ 为两个模糊系统, $f(x_1, x_2, x_3) = f_2[f_1(x_1, x_2), x_3]$ 为多级模糊系统。设计 f_1 和 f_2 使对于 $i, j, k = 1, 2, \dots, 6$, 有 $f(e_i, e_j, e_k) = g(e_i, e_j, e_k)$ 。

习题 22.6 利用习题 22.5 和第 10 章中的理论证明多级模糊系统是一个万能逼近器。

习题 22.7 证明定理 22.2。

第五部分

自适应模糊控制

模糊控制器通常是在被控对象的参数和结构存在诸多不确定性因素或者未知变动时才被采用。一般来讲,自适应控制的目的是在系统出现这些不确定因素时,仍保持一致性。因此,高级模糊控制应该具有自适应性。

图 23.1 给出了一个自适应模糊控制系统的基本框架。这个参考模型通常用于刻画一个模糊控制系统应该遵循的理想响应。被控对象中含有若干个未知的部件,模糊控制器则根据带有可调参数 θ 的模糊系统来设计。自适应律实时调节这些参数 θ ,使得被控对象的输出 $y(t)$ 能跟踪参考模型的输出 $y_m(t)$ 。

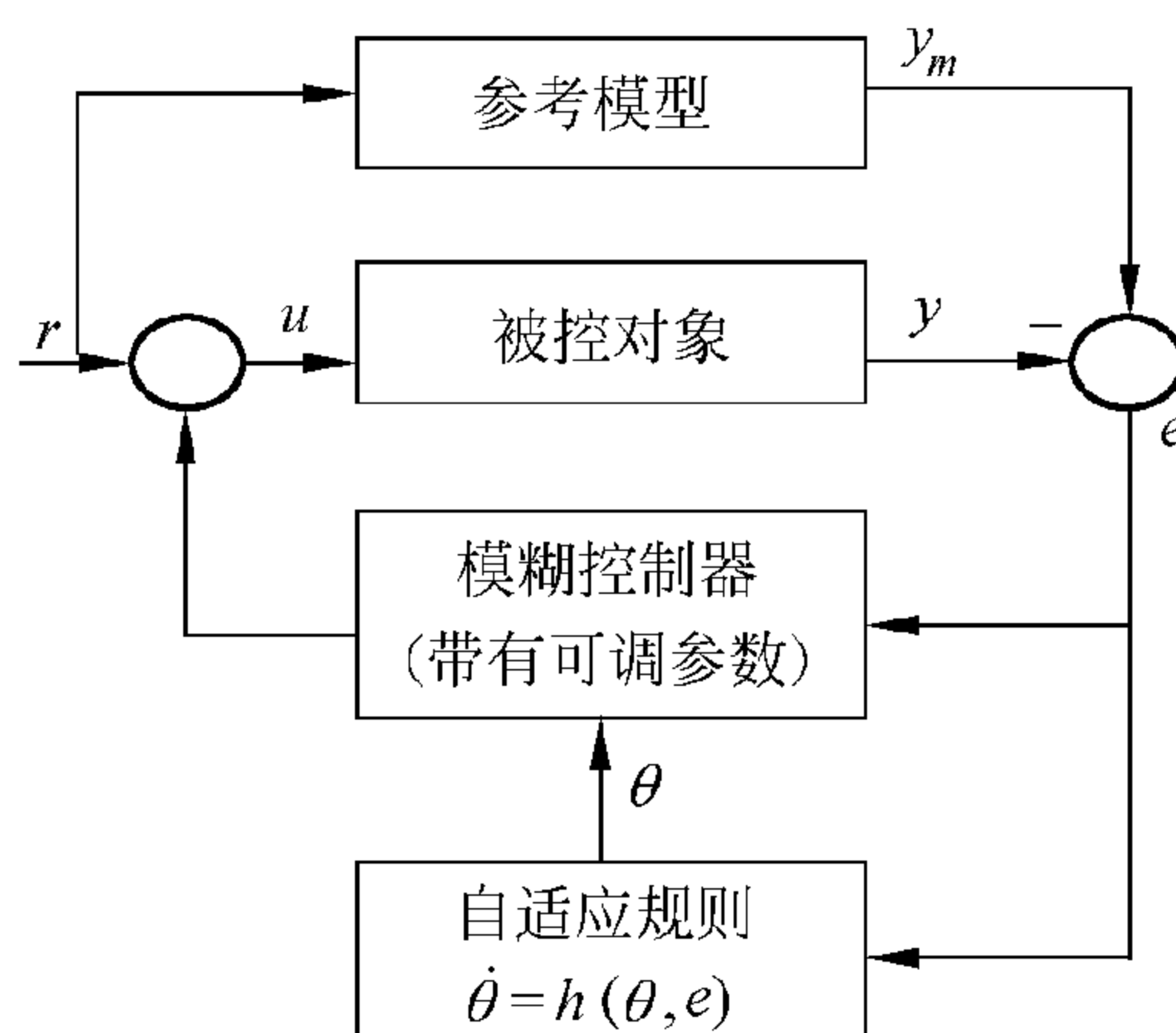


图 23.1 自适应模糊控制系统的基本框图

由图 23.1 可知,自适应模糊控制系统和非自适应模糊控制系统的区别主要在于:(i) 自适应模糊控制系统中的模糊控制器在实时控制过程中始终是变化的,而非自适应模糊控制系统中的模糊控制器在实时控制之前是固定的;(ii) 自适应律作为附加的部件,被引入自适应模糊控制系统中,来调节模糊控制器中的参数。

自适应模糊控制与非自适应模糊控制相比,其优点主要在于:(i) 自适应模糊控制器可以随环境变化进行自行调节,因此自适应模糊控制具有较好的控制性能;(ii) 因为自适应律有助于在实时控制过程中学习被控对象的动态特性,所以自适应模糊控制对被控对象的信息要求不高。

自适应模糊控制与非自适应模糊控制相比,其缺点主要在于:(i) 因为生成的控制系统不仅具有非线性特性,而且还具有时变性,所以更加难于分析;(ii) 实施代价要更高

一些。

第 23 至第 26 章,将给出未知或部分未知的非线性系统的一些自适应模糊控制器。第 23 章,首先把自适应模糊控制器分为间接型、直接型和组合型三类,然后介绍了一种设计基本间接型自适应模糊控制器的方法。第 24 章,给出了直接型和组合型自适应模糊控制器设计的基本原理。尽管在第 23 章和第 24 章中给出的自适应模糊控制器都有较好的收敛性,但还不能保证生成的闭环非线性时变系统具有全局稳定性。所以,第 25 章,将进一步介绍如何用监督控制法和参数投影法来保证系统中所有变量的一致有界性。最后,第 26 章,将运用输入—输出线性化的方法,给出设计一般非线性系统的自适应模糊控制器的方法。

基本自适应模糊控制器 I

23.1 自适应模糊控制器的分类

自适应模糊控制与传统的自适应控制之间既有相同点也有不同点。它们的相同点在于：(i)基本框架与原理或多或少有些相同；(ii)在分析与设计中使用的数学工具非常相似。二者主要不同点在于：(i)模糊控制器具有一种特殊的非线性结构，这种结构对不同的被控对象来说是通用的；而传统自适应控制器的结构因对象的不同而不同。(ii)与被控对象的动态特性和控制策略有关的人类知识可以嵌入自适应模糊控制器，而这些知识在传统的自适应控制系统中是不考虑的。第二点不同表明了自适应模糊控制优于传统自适应控制。

在研究嵌入人类知识的方法之前，须考虑使用什么类型的人类知识。从概念上来看，任何控制系统至少包含一个被控对象和一个控制器。因此，关于控制系统的人类知识可以分成两类：被控对象知识和控制知识。在本书的模糊控制框架中，这两种类型的人类知识是指：

- **被控对象知识。** 模糊 IF-THEN 规则，用来描述未知的被控对象的行为（例如，可以用如下模糊 IF-THEN 规则描述一辆汽车的行为：“如果加大油门，则汽车的速度会增大”。规则中的“加大”和“增大”是用模糊集来表征的）。
- **控制知识。** 模糊控制规则，用来描述在什么情况下应采取什么样的控制行为（例如，可以经常使用下面的模糊 IF-THEN 规则驾车：“如果汽车速度慢了，就加大油门”。规则中的“慢”和“加大”是用模糊集来表征的）。

根据所用的人类知识的不同和模糊控制器结构的不同，可将自适应模糊控制分成以下三种类型：

- **间接型自适应模糊控制。** 该模糊控制器包含许多模糊系统，这些模糊系统最初是根据被控对象的知识构造出来的。
- **直接型自适应模糊控制。** 该模糊控制器只包含一个模糊系统，这个模糊系统最初是根据控制知识构造出来的。
- **组合型自适应模糊控制。** 该模糊控制器是间接型和直接型自适应模糊控制器的加权平均（因此，组合型模糊控制器中的模糊系统既用到了被控对象知识，也用到了控制知识）。

下面,将介绍间接型自适应模糊控制器的设计方法。在第 24 章再介绍直接型和组合型自适应模糊控制器的设计方法。

23.2 间接型自适应模糊控制器的设计

23.2.1 问题描述

考虑如下形式的 n 阶非线性系统:

$$\dot{x}^{(n)} = f(x, \dot{x}, \dots, x^{(n-1)}) + g(x, \dot{x}, \dots, x^{(n-1)})u \quad (23.1)$$

$$y = x \quad (23.2)$$

式中, f 和 g 为未知函数, $u \in R$ 和 $y \in R$ 分别为系统的输入和输出, $\mathbf{x} = (x_1, x_2, \dots, x_n)^T = (x, \dot{x}, \dots, x^{(n-1)})^T \in R^n$ 为系统的状态向量并假设这些向量可以通过测量得到。式(23.1)的能控条件是 $g(\mathbf{x}) \neq 0$ 。不失一般性,假设 $g(\mathbf{x}) > 0$ 。由非线性控制文献(Isidori[1989])可知,这样的系统是标准型系统,相对阶数等于 n 。

控制的目标是基于模糊系统设计一个反馈控制器 $u = u(\mathbf{x}|\theta)$ 和调整参数向量 θ 的自适应律(adaptation law),使得系统输出 y 能够跟踪理想输出 y_m, y_m 及其导数是已知且有界的。

由于系统中函数 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 是非线性的而且假设是未知的,所以这里将其作为一般的单维输入单维输出非线性控制问题处理。这样,控制的目标就不是要求系统输出 y 渐近地收敛于理想的输出 y_m 了,而是要求系统的输出 y 尽可能地跟踪 y_m 。在第 25 章至第 26 章里,我们将介绍更为高级的算法,来保证自适应模糊控制系统的稳定性和收敛性。本章目的在于说明自适应模糊控制的基本思想。

由于本节是要设计一个间接型自适应模糊控制器,所以这里假设关于被控对象的一些专家知识是可以得到的。具体地,假设用于描述 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 输入—输出行为的模糊 IF-THEN 规则集合是可以得到的,其形式为

$$\text{如果 } x_1 \text{ 是 } F_1^r \text{ 且 } \dots \text{ 且 } x_n \text{ 是 } F_n^r, \text{ 则 } f(\mathbf{x}) \text{ 是 } C^r \quad (23.3)$$

和

$$\text{如果 } x_1 \text{ 是 } G_1^s \text{ 且 } \dots \text{ 且 } x_n \text{ 是 } G_n^s, \text{ 则 } g(\mathbf{x}) \text{ 是 } D^s \quad (23.4)$$

其中 F_i^r, C^r, G_i^s 和 D^s 为 R 中模糊集合, $r = 1, 2, \dots, L_f$ 且 $s = 1, 2, \dots, L_g$ 。

23.2.2 模糊控制器的设计

如果非线性函数 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 是已知的,则可以选择控制 u 来消除其非线性的性质,然后再根据线性控制理论设计控制器。具体地,令 $e = y_m - y = y_m - x, e = (e, \dot{e}, \dots, e^{(n-1)})^T, \mathbf{k} = (k_n, \dots, k_1)^T$, 使多项式 $s^n + k_1 s^{(n-1)} + \dots + k_n$ 的所有根都在复平面左半开平面上,然后选

择控制律为

$$u^* = \frac{1}{g(\mathbf{x})} [-f(\mathbf{x}) + y_m^{(n)} + \mathbf{k}^T \mathbf{e}] \quad (23.5)$$

将式(23.5)代入式(23.1),可以得到下面的闭环控制系统的系统方程

$$e^{(n)} + k_1 e^{(n-1)} + \cdots + k_n e = 0 \quad (23.6)$$

由 \mathbf{k} 的选取,可得 $t \rightarrow \infty$ 时 $e(t) \rightarrow 0$,即系统的输出 y 渐近地收敛于理想输出 y_m 。

因为 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 是未知的,所以理想控制器(23.5)根本无法实现。又因为模糊 IF-THEN规则(23.3)~(23.4)描述了 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 输入—输出行为,所以,一种合理的办法就是用规则(23.3)和规则(23.4)所构造的模糊系统 $\hat{f}(\mathbf{x})$ 和 $\hat{g}(\mathbf{x})$ 分别代替 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 。由于规则(23.3)和规则(23.4)给出的只是关于 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 的粗略信息,所以基于这些规则构造的模糊系统 $\hat{f}(\mathbf{x})$ 和 $\hat{g}(\mathbf{x})$ 可能无法很好地逼近 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 。提高 $\hat{f}(\mathbf{x})$ 和 $\hat{g}(\mathbf{x})$ 逼近精度的一种办法就是,允许 $\hat{f}(\mathbf{x})$ 和 $\hat{g}(\mathbf{x})$ 中的一些参数在整个在线操作过程中自由变化,以使 $\hat{f}(\mathbf{x})$ 和 $\hat{g}(\mathbf{x})$ 的逼近精度随时间逐渐提高。令 $\theta_f \in R^{M_f}$ 和 $\theta_g \in R^{M_g}$ 分别为 $\hat{f}(\mathbf{x})$ 和 $\hat{g}(\mathbf{x})$ 的自由参数,则有 $\hat{f}(\mathbf{x}) = \hat{f}(\mathbf{x}|\theta_f)$ 和 $\hat{g}(\mathbf{x}) = \hat{g}(\mathbf{x}|\theta_g)$ 。用模糊系统 $\hat{f}(\mathbf{x}|\theta_f)$ 和 $\hat{g}(\mathbf{x}|\theta_g)$ 分别代替式(23.5)中的 $f(\mathbf{x})$ 和 $g(\mathbf{x})$,就可以得到下面的模糊控制器

$$u = u_f = \frac{1}{\hat{g}(\mathbf{x}|\theta_g)} [-\hat{f}(\mathbf{x}|\theta_f) + y_m^{(n)} + \mathbf{k}^T \mathbf{e}] \quad (23.7)$$

人们将以上控制器称为确定等效控制器(certainty equivalent controller)的原因在于,如果 \hat{f} 和 \hat{g} 分别等于 f 和 g (这意味着 f 和 g 没有不确定性),则控制器 u_f 就是式(23.5)中的理想控制器 u^* 。

要想实现控制器(23.7),必须给出 $\hat{f}(\mathbf{x}|\theta_f)$ 和 $\hat{g}(\mathbf{x}|\theta_g)$ 的具体公式。由于式(23.3)和式(23.4)中的规则数量可能非常少,所以通常情况下,仅使用式(23.3)中的规则 L_f 和式(23.4)中的规则 L_g 是不足以构造 $\hat{f}(\mathbf{x}|\theta_f)$ 和 $\hat{g}(\mathbf{x}|\theta_g)$ 的,应该使用包括(23.3)和(23.4)中的规则在内的规则完备集合来构造 $\hat{f}(\mathbf{x}|\theta_f)$ 和 $\hat{g}(\mathbf{x}|\theta_g)$,具体地, $\hat{f}(\mathbf{x}|\theta_f)$ 和 $\hat{g}(\mathbf{x}|\theta_g)$ 可由以下两步来构造:

步骤 1: 对变量 $x_i (i=1,2,\dots,n)$,定义 p_i 个模糊集合 $A_i^{l_i} (l_i=1,2,\dots,p_i)$,式(23.3)中的模糊集合 $F_i^r (i=1,2,\dots,L_f)$ 作为特例包括在其中;定义 q_i 个模糊集合 $B_i^{l_i} (l_i=1,2,\dots,q_i)$,式(23.4)中的模糊集合 $G_i^s (i=1,2,\dots,L_g)$ 作为特例也包括在其中。

步骤 2: 用以下 $\prod_{i=1}^n p_i$ 条模糊规则来构造模糊系统 $\hat{f}(\mathbf{x}|\theta_f)$:

如果 x_1 是 $A_1^{l_1}$ 且...且 x_n 是 $A_n^{l_n}$, 则 f^\wedge 是 $E^{l_1 \cdots l_n}$ (23.8)

其中, $l_i = 1, 2, \dots, p_i, i = 1, 2, \dots, n$, 且当式(23.8)中的 IF 部分与式(23.3)中的 IF 部分一致时, $E^{l_1 \cdots l_n}$ 等于 C^r , 否则, $E^{l_1 \cdots l_n}$ 等于某个任意模糊集。类似地, 可以用 $\prod_{i=1}^n q_i$ 条模糊规则来构造模糊系统 $\hat{g}(\mathbf{x}|\theta_g)$:

如果 x_1 是 $B_1^{l_1}$ 且...且 x_n 是 $B_n^{l_n}$, 则 \hat{g} 是 $H^{l_1 \cdots l_n}$ (23.9)

其中, $l_i = 1, 2, \dots, q_i, i = 1, 2, \dots, n$, 且如果式(23.9)中 IF 部分与式(23.4)中的 IF 部分一致, 则 $H^{l_1 \cdots l_n}$ 等于 D^s , 否则, $H^{l_1 \cdots l_n}$ 等于某个任意模糊集。具体地, 如果使用乘积推理机、单值模糊器和中心平均解模糊器, 则可得

到

$$f^\wedge(\mathbf{x}|\theta_f) = \frac{\sum_{l_1=1}^{p_1} \cdots \sum_{l_n=1}^{p_n} \bar{y}_f^{l_1 \cdots l_n} \left(\prod_{i=1}^n \mu_{A_i^{l_i}}(x_i) \right)}{\sum_{l_1=1}^{p_1} \cdots \sum_{l_n=1}^{p_n} \left(\prod_{i=1}^n \mu_{A_i^{l_i}}(x_i) \right)} \quad (23.10)$$

$$\hat{g}(\mathbf{x}|\theta_g) = \frac{\sum_{l_1=1}^{q_1} \cdots \sum_{l_n=1}^{q_n} \bar{y}_g^{l_1 \cdots l_n} \left(\prod_{i=1}^n \mu_{B_i^{l_i}}(x_i) \right)}{\sum_{l_1=1}^{q_1} \cdots \sum_{l_n=1}^{q_n} \left(\prod_{i=1}^n \mu_{B_i^{l_i}}(x_i) \right)} \quad (23.11)$$

令 $\bar{y}_f^{l_1 \cdots l_n}$ 和 $\bar{y}_g^{l_1 \cdots l_n}$ 是自由参数, 分别放在集合 $\theta_f \in R^{\prod_{i=1}^n p_i}$ 和 $\theta_g \in R^{\prod_{i=1}^n q_i}$ 中, 则式(23.10) ~ 式(23.11)可改写为

$$f^\wedge(\mathbf{x}|\theta_f) = \theta_f^T \xi(\mathbf{x}) \quad (23.12)$$

$$\hat{g}(\mathbf{x}|\theta_g) = \theta_g^T \eta(\mathbf{x}) \quad (23.13)$$

式中, $\xi(\mathbf{x})$ 是一个 $\prod_{i=1}^n p_i$ 维向量, 其第 $l_1 \cdots l_n$ 个元素为

$$\xi_{l_1 \cdots l_n}(\mathbf{x}) = \frac{\prod_{i=1}^n \mu_{A_i^{l_i}}(x_i)}{\sum_{l_1=1}^{p_1} \cdots \sum_{l_n=1}^{p_n} \left(\prod_{i=1}^n \mu_{A_i^{l_i}}(x_i) \right)} \quad (23.14)$$

式中, $\eta(\mathbf{x})$ 是一个 $\prod_{i=1}^n q_i$ 维向量, 其第 $l_1 \cdots l_n$ 个元素为

$$\eta_{l_1 \dots l_n}(\mathbf{x}) = \frac{\prod_{i=1}^n \mu_{B_{l_i}}(x_i)}{\sum_{l_1=1}^{q_1} \dots \sum_{l_n=1}^{q_n} \left(\prod_{i=1}^n \mu_{B_{l_i}}(x_i) \right)} \quad (23.15)$$

从步骤 2 可知, θ_f 和 θ_g 中的一些参数按照规则(23.3)和规则(23.4)选取, 其余的参数随机选取(或给定结构)。因为在线控制时, 参数 θ_f 和 θ_g 会发生变化, 所以上面设定的参数是初始参数。下一任务是为 θ_f 和 θ_g 设计一种自适应律, 使跟踪误差 e 达到最小。

23.2.3 自适应律的设计

把式(23.7)代入式(23.3), 整理后可得如下模糊控制系统的闭环动态

$$\dot{e}^{(n)} = -\mathbf{k}^T \mathbf{e} + [f^{\wedge}(\mathbf{x}|\theta_f) - f(\mathbf{x})] + [g^{\wedge}(\mathbf{x}|\theta_g) - g(\mathbf{x})]u_l \quad (23.16)$$

令

$$\mathbf{\Lambda} = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ -k_n & -k_{n-1} & \dots & \dots & \dots & \dots & -k_1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0 \\ \dots \\ 0 \\ 1 \end{bmatrix} \quad (23.17)$$

则动态方程(23.16)可以改写为如下向量形式

$$\dot{\mathbf{e}} = \mathbf{\Lambda} \mathbf{e} + \mathbf{b} \{ [f^{\wedge}(\mathbf{x}|\theta_f) - f(\mathbf{x})] + [g^{\wedge}(\mathbf{x}|\theta_g) - g(\mathbf{x})]u_l \} \quad (23.18)$$

定义最优参数为

$$\theta_f^* = \arg \min_{\theta_f \in R^{\prod_{i=1}^n p_i}} [\sup_{x \in R^n} |f^{\wedge}(\mathbf{x}|\theta_f) - f(\mathbf{x})|] \quad (23.19)$$

$$\theta_g^* = \arg \min_{\theta_g \in R^{\prod_{i=1}^n q_i}} [\sup_{x \in R^n} |g^{\wedge}(\mathbf{x}|\theta_g) - g(\mathbf{x})|] \quad (23.20)$$

于是, 对所有的形如式(23.10)和式(23.11)的模糊系统而言, $f^{\wedge}(\mathbf{x}|\theta_f^*)$ 和 $g^{\wedge}(\mathbf{x}|\theta_g^*)$ 分别是 $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 最优的(min-max)逼近器。下面定义最小逼近误差(minimum approximation error)

$$w = [f^{\wedge}(\mathbf{x}|\theta_f^*) - f(\mathbf{x})] + [g^{\wedge}(\mathbf{x}|\theta_g^*) - g(\mathbf{x})]u_l \quad (23.21)$$

由此 w 可将式(23.18)改写为

$$\dot{\mathbf{e}} = \mathbf{\Lambda} \mathbf{e} + \mathbf{b} \{ [f^{\wedge}(\mathbf{x}|\theta_f) - f^{\wedge}(\mathbf{x}|\theta_f^*)] + [g^{\wedge}(\mathbf{x}|\theta_g) - g^{\wedge}(\mathbf{x}|\theta_g^*)]u_l + w \} \quad (23.22)$$

将式(23.12)和式(23.13)代入式(23.22), 可得如下闭环动态方程

$$\dot{\mathbf{e}} = \mathbf{\Lambda} \mathbf{e} + \mathbf{b} [(\theta_f - \theta_f^*)^T \xi(\mathbf{x}) + (\theta_g - \theta_g^*)^T \eta(\mathbf{x})u_l + w] \quad (23.23)$$

该方程清晰地描述了跟踪误差和控制参数 θ_f 、 θ_g 之间的关系。自适应律的任务是为 θ_f 、 θ_g

确定一个调节机理,使得跟踪误差 e 和参数误差 $\theta_f - \theta_f^*$ 、 $\theta_g - \theta_g^*$ 达到最小。

为了完成这一任务,考虑以下候选 Lyapunov 函数

$$V = \frac{1}{2} \mathbf{e}^T P \mathbf{e} + \frac{1}{2\gamma_1} (\theta_f - \theta_f^*)^T (\theta_f - \theta_f^*) + \frac{1}{2\gamma_2} (\theta_g - \theta_g^*)^T (\theta_g - \theta_g^*) \quad (23.24)$$

式中, γ_1, γ_2 是正常数, P 是一个正定矩阵且满足 Lyapunov 方程

$$\Lambda^T P + P \Lambda = -Q \quad (23.25)$$

这里, Q 是一个任意的 $n \times n$ 正定矩阵, Λ 由式(23.17)给出。根据闭环系统轨迹(23.23)可得 V 关于时间的导数为

$$\begin{aligned} \dot{V} = & -\frac{1}{2} \mathbf{e}^T P \mathbf{e} + \mathbf{e}^T P \mathbf{b} w + \frac{1}{\gamma_1} (\theta_f - \theta_f^*)^T [\dot{\theta}_f + \gamma_1 \mathbf{e}^T P \mathbf{b} \xi(\mathbf{x})] \\ & + \frac{1}{\gamma_2} (\theta_g - \theta_g^*)^T [\dot{\theta}_g + \gamma_2 \mathbf{e}^T P \mathbf{b} \eta(\mathbf{x}) u_1] \end{aligned} \quad (23.26)$$

使跟踪误差 e 和参数误差 $\theta_f - \theta_f^*$ 、 $\theta_g - \theta_g^*$ 最小,等价于使 V 最小,为此应该选择自适应律,

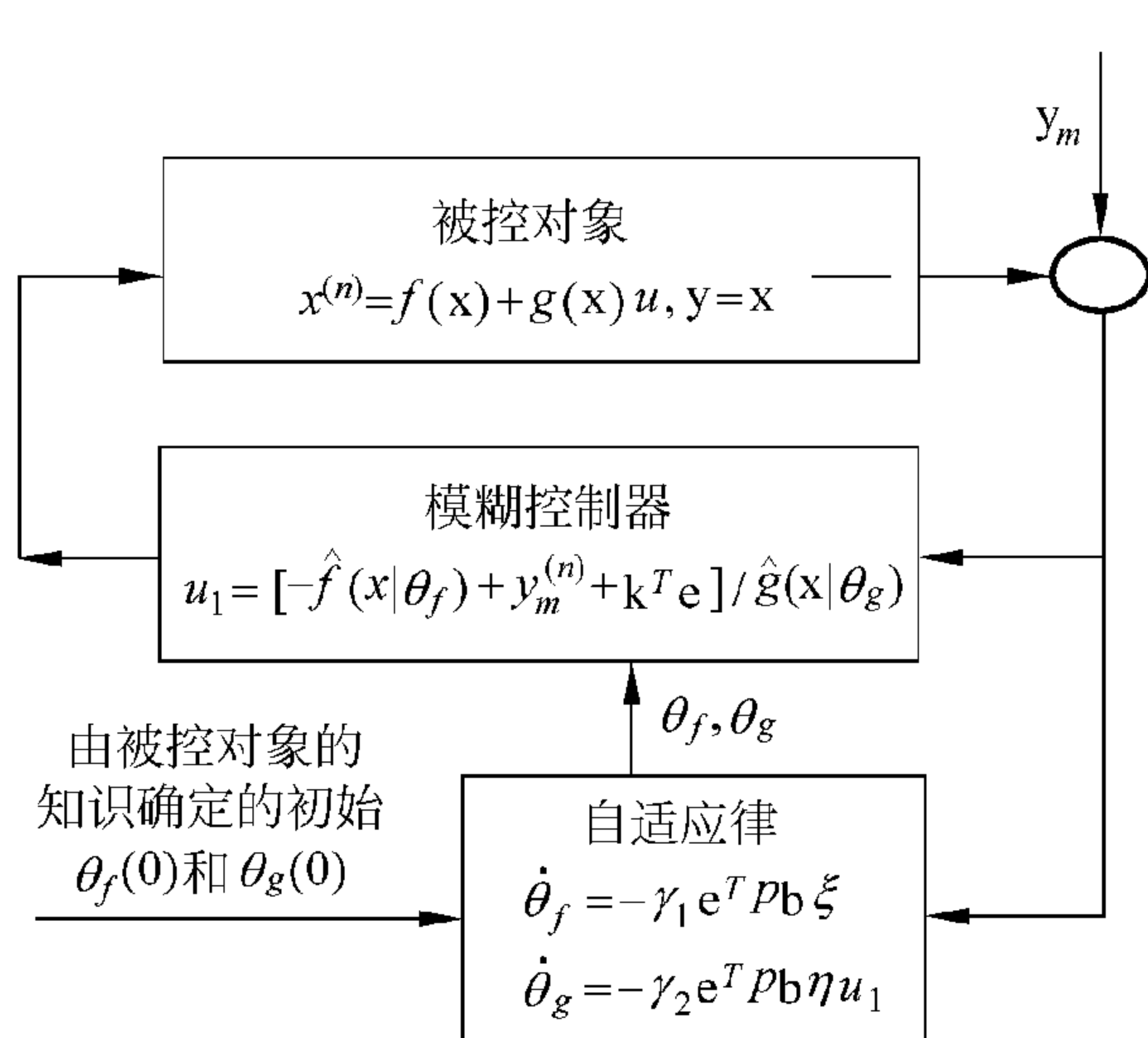


图 23.2 间接型自适应模糊控制系统

使得 \dot{V} 取负值。由于 $-\frac{1}{2} \mathbf{e}^T P \mathbf{e}$ 是负的,且应选择使得最小逼近误差 w 非常小的模糊系统,所以,一个好的策略是选择自适应律,使式(23.26)中的最后两项为零,即选择自适应律

$$\dot{\theta}_f = -\gamma_1 \mathbf{e}^T P \mathbf{b} \xi(\mathbf{x}) \quad (23.27)$$

$$\dot{\theta}_g = -\gamma_2 \mathbf{e}^T P \mathbf{b} \eta(\mathbf{x}) u_1 \quad (23.28)$$

因为其目标是使 Lyapunov 函数 V 最小,所以这种自适应律设计方法被称为 Lyapunov 综合法(synthesis approach)。

作为总结,整个的间接型自适应模糊控制系统框图如图 23.2 所示。这里应明确指出的是,与被控对象有关的知识(模糊“IF-THEN”规则(23.3)和规则(23.4))经由初始参数 $\theta_f(0)$ 和 $\theta_g(0)$ 被嵌入到系统中,这一过程详见

$\hat{f}(\mathbf{x}|\theta_f)$ 和 $\hat{g}(\mathbf{x}|\theta_g)$ 的设计程序中的步骤 2。

23.3 在倒立摆跟踪控制中的应用

现在,将间接型自适应模糊控制器应用于如图 1.9 所示的倒立摆跟踪控制系统中,该系统的动态特性是由式(20.15)~式(20.16)描述的。选择 $k_1 = 2$ 和 $k_2 = 1$ (保证 $s^2 + k_1 s + k_2$ 是稳定的), $Q = \text{diag}(10, 10)$, 求解 Lyapunov 方程(23.25),以得到满足式(23.29)的 P 。现

在考虑两个例子：一个例子中没有任何关于被控对象的知识(例 23.1)；另一个例子中有一些关于被控对象的知识(例 23.2)。

例 23.1 本例假设不存在式(23.3)和(23.4)中的语言规则。取 $p_1 = p_2 = q_1 = q_2 = 5$, $A_1^l = A_2^l = B_1^l = B_2^l$, 其隶属度函数为

$$\mu_{A_1^l}(x_1) = \exp\left[-\left(\frac{x_1 + \pi/6}{\pi/24}\right)^2\right], \mu_{A_2^l}(x_1) = \exp\left[-\left(\frac{x_1 + \pi/12}{\pi/24}\right)^2\right],$$

$$\mu_{A_3^l}(x_1) = \exp\left[-\left(\frac{x_1}{\pi/24}\right)^2\right], \mu_{A_4^l}(x_1) = \exp\left[-\left(\frac{x_1 - \pi/12}{\pi/24}\right)^2\right],$$

$$\mu_{A_5^l}(x_1) = \exp\left[-\left(\frac{x_1 - \pi/6}{\pi/24}\right)^2\right]$$

这些隶属度函数覆盖了 $[-\pi/6, \pi/6]$ 整个区间。由于 $f(x_1, x_2)$ 的取值范围比 $g(x_1, x_2)$ 的取值范围大得多, 所以选 $\gamma_1 = 50, \gamma_2 = 1$ 。图 23.3 和图 23.4 分别给出了初始条件为 $\mathbf{x}(0) = \left(-\frac{\pi}{60}, 0\right)^T$ 和 $\mathbf{x}(0) = \left(\frac{\pi}{60}, 0\right)^T$ 时的闭环系统的 $x_1(t)$ 和输出 $y_m(t)$ 。

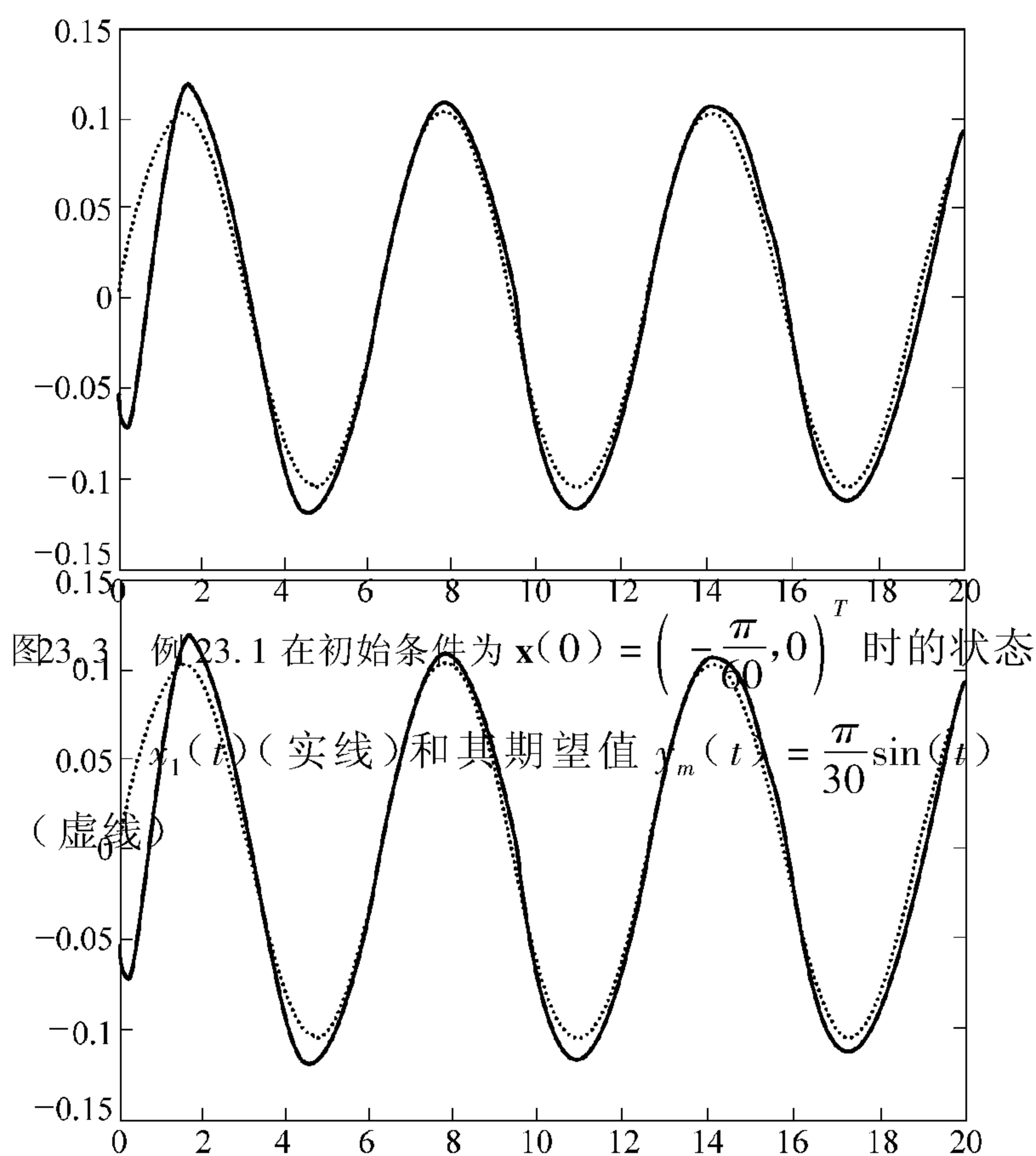


图23.3 例 23.1 在初始条件为 $\mathbf{x}(0) = \left(-\frac{\pi}{60}, 0\right)^T$ 时的状态 $x_1(t)$ (实线) 和其期望值 $y_m(t) = \frac{\pi}{30} \sin(t)$ (虚线)

图23.4 例 23.1 在初始条件为 $\mathbf{x}(0) = \left(\frac{\pi}{60}, 0\right)^T$ 时的状态 $x_1(t)$ (实线) 和其期望值 $y_m(t) = \frac{\pi}{30} \sin(t)$ (虚线)

例 23.2 本例中考虑的情况除一点外,其余完全与例 23.1 相同。这一不同点是,存在某些 $f(x_1, x_2)$ 和 $g(x_1, x_2)$ 的语言规则,这些语言规则是根据直观的物理常识得到的。首先,假设不施加控制,即 $u = 0$ 。在这种情况下,角度 $\theta = x_1$ 的加速度等于 $f(x_1, x_2)$ 。基于物理常识可得如下的观察结果

$$x_1 \text{ 越大, } f(x_1, x_2) \text{ 越大} \tag{23.29}$$

这里的任务是要将观察结果转换成 $f(x_1, x_2)$ 的模糊规则。由于 $(x_1, x_2) = (0, 0)$ 为倒立摆系统的一个(非稳定的)平衡点,因此,第一条规则为

$$R_f^{(1)}: \text{ 如果 } x_1 \text{ 是 } F_1^3 \text{ 且 } x_2 \text{ 是 } F_2^3, \text{ 则 } f(x_1, x_2) \text{ 接近于零} \tag{23.30}$$

式中的 $F_i^j (i = 1, 2; j = 1, 2, \dots, 5)$ 为例 23.1 中定义的模糊集合 A_i^j , 而“接近于零”是指模糊集合的中心为零。从图 1.9 中可以看出, x_1 的加速度与重力 $mg\sin(x_1)$ 成正比,即可近似地得到 $f(x_1, x_2) = \alpha\sin(x_1)$, 其中 α 为常数。显然, $f(x_1, x_2)$ 将在 $x_1 = \pi/2$ 处取得最大值,这样再根据式(20.16),就可以近似地得到 $\alpha = 16$ 。于是,可以得到 $f(x_1, x_2)$ 的如下模糊 IF-THEN 规则:

$$R_f^{(2)}: \text{ 如果 } x_1 \text{ 是 } F_1^1 \text{ 且 } x_2 \text{ 是 } F_2^3, \text{ 则 } f(x_1, x_2) \text{ 接近于 } -8 \tag{23.31}$$

$$R_f^{(3)}: \text{ 如果 } x_1 \text{ 是 } F_1^2 \text{ 且 } x_2 \text{ 是 } F_2^3, \text{ 则 } f(x_1, x_2) \text{ 接近于 } -4 \tag{23.32}$$

$$R_f^{(4)}: \text{ 如果 } x_1 \text{ 是 } F_1^4 \text{ 且 } x_2 \text{ 是 } F_2^3, \text{ 则 } f(x_1, x_2) \text{ 接近于 } 4 \tag{23.33}$$

$$R_f^{(5)}: \text{ 如果 } x_1 \text{ 是 } F_1^5 \text{ 且 } x_2 \text{ 是 } F_2^3, \text{ 则 } f(x_1, x_2) \text{ 接近于 } 8 \tag{23.34}$$

式中, $F_i^j (i = 1, 2; j = 1, 2, \dots, 5)$ 与例 23.1 中对应的模糊集合 A_i^j 完全相同,而“接近于”之后的值则是由 $16\sin(\pi/6) = 8$ 和 $8\sin(\pi/12) = 4$ 来确定的。再由物理常识可知, $f(x_1, x_2)$ 对 x_1 比对 x_2 敏感,因此,可以将规则(23.30)和规则(23.34)推广到 x_2 为任意 $F_2^j (j = 1, 2, \dots, 5)$ 的规则中去。总之,最终由于表征 $f(x_1, x_2)$ 的规则都在图 23.5 中给出了。可以使用这

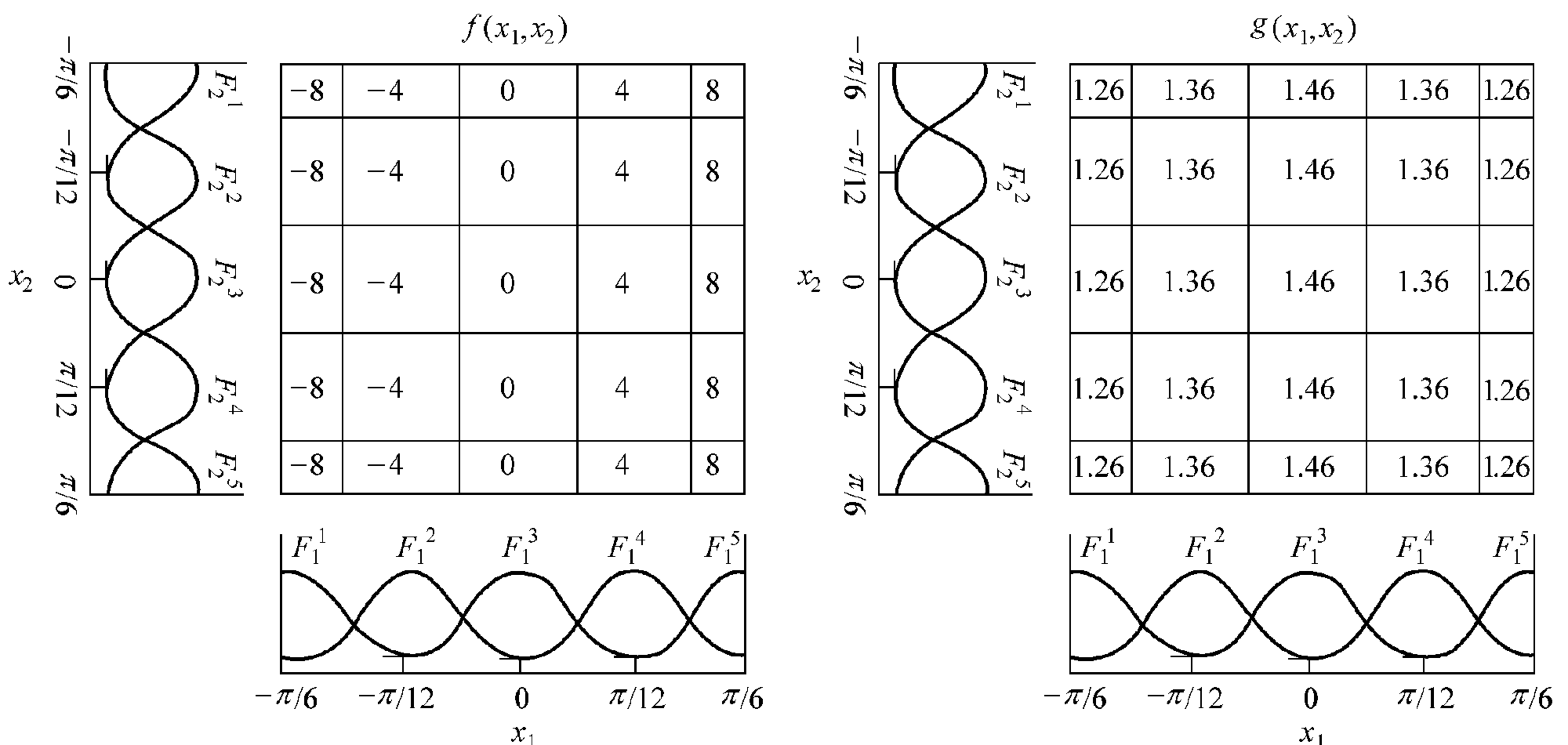


图 23.5 $f(x_1, x_2)$ 的语言模糊 IF-THEN 规则图 23.6 $g(x_1, x_2)$ 的语言模糊 IF-THEN 规则

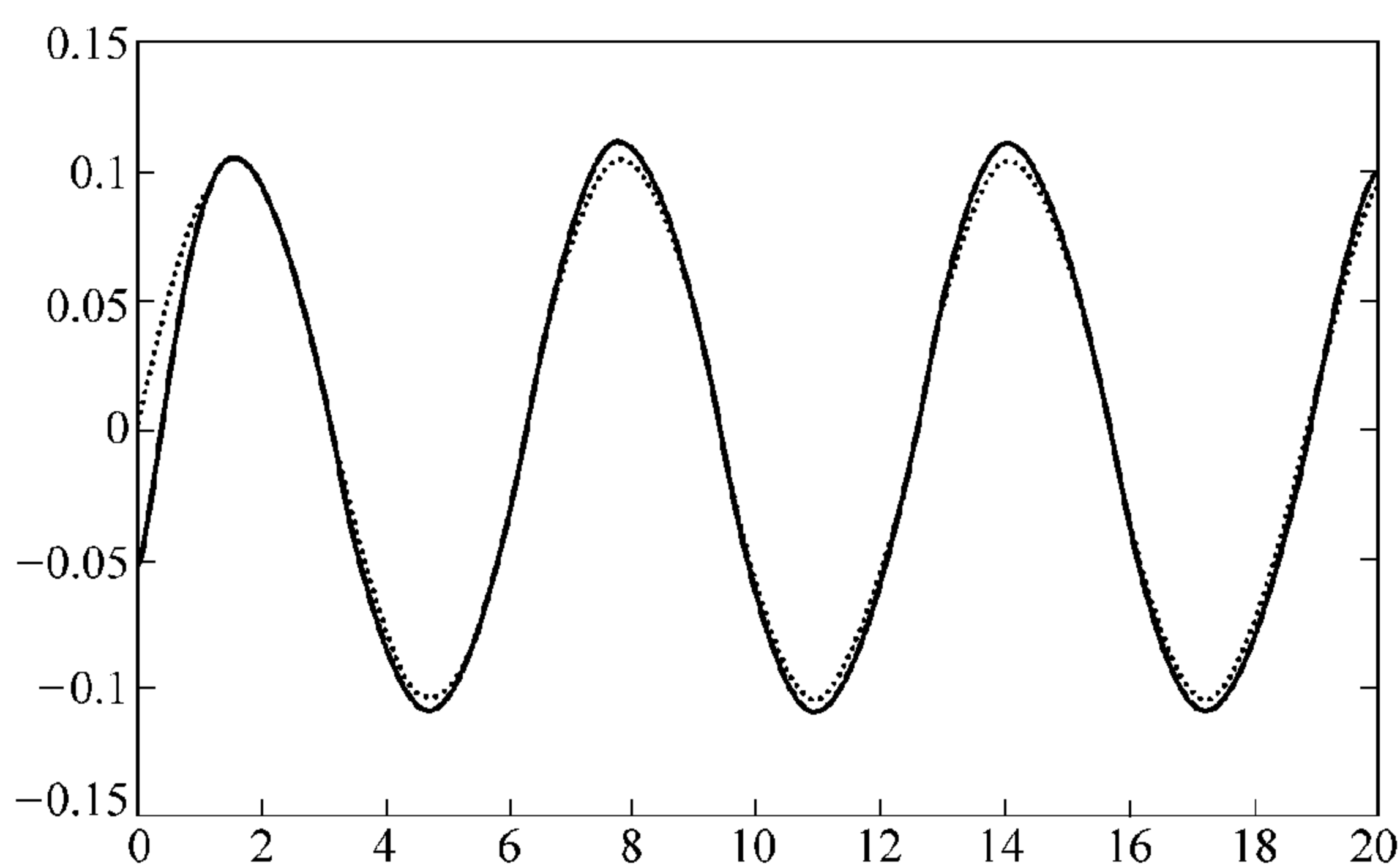
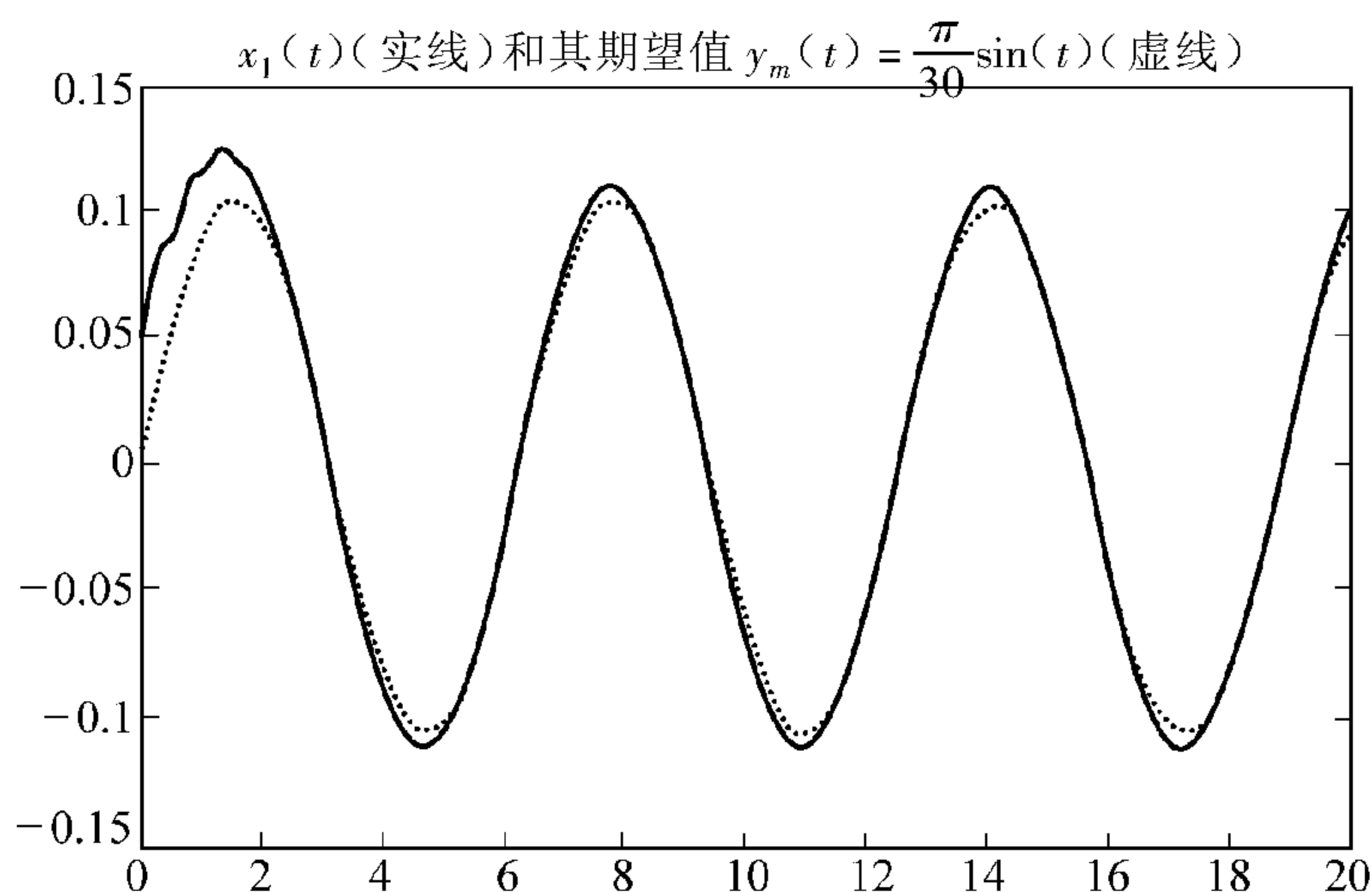
些规则确定初始参数 $\theta_f(0)$ 。

下面, 根据物理常识来确定 $g(x_1, x_2)$ 的模糊 IF-THEN 规则。由于 $g(x_1, x_2)$ 决定控制 u 对系统的控制强度, 且显然, 这一强度在 $x_1 = 0$ 处达到最大, 所以可得如下观察结果:

$$x_1 \text{ 越小, } g(x_1, x_2) \text{ 越大} \quad (23.35)$$

用类似于获取 $f(x_1, x_2)$ 的规则的方法, 可将式(23.35)中的观察结果转换成 25 条 $g(x_1, x_2)$ 的模糊规则(见图 23.6)。

图 23.7 和图 23.8 分别给出了在图 23.5 和图 23.6 中模糊 IF-THEN 规则给定之后, 自适应模糊控制器初始条件为 $\mathbf{x}(0) = \left(-\frac{\pi}{60}, 0\right)^T$ 和 $\mathbf{x}(0) = \left(\frac{\pi}{60}, 0\right)^T$ 时的闭环系统状态

图 23.7 例 23.2 中在初始条件为 $\mathbf{x}(0) = \left(-\frac{\pi}{60}, 0\right)^T$ 时的状态图 23.8 例 23.2 中在初始条件为 $\mathbf{x}(0) = \left(\frac{\pi}{60}, 0\right)^T$ 时的状态

$x_1(t)$ (实线) 和其期望值 $y_m(t) = \frac{\pi}{30} \sin(t)$ (虚线)

$x_1(t)$ 和理想输出 $y_m(t)$ 。与例 23.1 中的结果进行比较,可以发现,早期的控制在给定这些规则后得到明显改善。

23.4 本章小结及导读

本章主要阐述了以下内容:

- 自适应模糊控制器的研究目的及分类;
- 使用确定等效原理和 Lyapunov 综合法设计间接型自适应模糊控制器;
- 如何应用间接型自适应模糊控制器解决倒立摆的跟踪控制问题。

在自适应控制方面有许多好的教材,例如, Narendra 和 Annaswamy[1989], 以及 Åström 和 Wittenmark[1995]。非线性系统自适应控制是当前很活跃的一个研究领域,一些较早的文章可见于 Sastry 和 Isidori[1989], Narendra 和 Parthasarathy[1990], 以及 Sanner 和 Slotine [1991]。本章中的方法摘自 Wang[1994a]。

23.5 习题

习题 23.1 考虑下面控制问题,一物体在光滑平面上受动力 u 控制,被控对象的动态行为满足方程

$$m \ddot{x} = u \quad (23.36)$$

目标是设计一个控制器 u ,使得 $x(t)$ 收敛于参考信号 $y_m(t)$,其中 $y_m(t)$ 遵循如下规律

$$\ddot{y}_m + 2\dot{y}_m + y_m = r(t) \quad (23.37)$$

考虑下面两种情况:

(a) m 是已知的,设计一个非自适应控制器来实现上述目标。

(b) m 是未知的,设计一个自适应控制器来实现上述目标。画出这个自适应控制系统的框图,并讨论其自适应律的直观合理性。

习题 23.2 考虑一阶非线性系统

$$\dot{x} = f(x) + u \quad (23.38)$$

式中, $f(x)$ 是未知的。设计一个自适应模糊控制器 u ,使得当 t 趋于无穷大时, $x(t)$ 能达到目标状态 $x_d(t) = \sin(t)$ 。在什么条件下能够保证 t 趋于无穷时跟踪误差趋于零?

习题 23.3 假设利用被控对象的知识可以很好地估计 f 和 g ,并且在在线自适应中不改变这些估计值。试修改本章的间接型自适应模糊控制器设计方法来实现这一目标。

习题 23.4 考虑自适应律(23.27)和(23.28),从直观上讨论,为什么在不稳定的系统

中进行参数估计要比在稳定的系统中容易些。

习题 23.5 考虑如下系统

$$\dot{e} = -e + \theta w(t) \quad (23.39)$$

$$\dot{\theta} = -ew(t) \quad (23.40)$$

式中, $e, \theta \in R, w(t)$ 有界, 试证明当 $t \rightarrow \infty$ 时 $e(t) \rightarrow 0$ 。原点处的平衡点是渐近稳定的吗?

习题 23.6 重做本章 23.3 节中的仿真, 试对更多的例子进行仿真计算。

24.1 直接型自适应模糊控制器的设计

24.1.1 问题描述

考虑如下方程所描述的研究对象

$$x^{(n)} = f(x, \dot{x}, \dots, x^{(n-1)}) + bu \quad (24.1)$$

$$y = x \quad (24.2)$$

式中, f 为未知函数, b 为未知的正常数。控制的目标和间接型自适应模糊控制的目标是一样的,也是基于模糊系统设计一个反馈控制器 $u = u(\mathbf{x}|\theta)$ 和一个调整参数向量 θ 的自适应律,使得系统输出 y 尽可能地跟踪理想输出 $y_m(t)$ 。二者的主要不同之处在于对可用知识的假设不同。具体地讲,这里已知的是一些控制知识,而不是被控对象知识式(23.3)和式(23.4),即采用下面的模糊 IF-THEN 规则来描述控制知识:

$$\text{如果 } x_1 \text{ 是 } P_1^r \text{ 且 } \dots \text{ 且 } x_n \text{ 是 } P_n^r, \text{ 则 } u \text{ 是 } Q^r \quad (24.3)$$

式中, P_i^r, Q^r 为 R 中模糊集合,且 $r = 1, 2, \dots, L_u$ 。模糊控制器的设计要求能将规则(24.3)自然地包括进来。

24.1.2 模糊控制器的设计

将规则(24.3)包含进模糊控制器的一种自然方法就是选择使用单一模糊系统作为控制器,即此情况下的模糊控制器为

$$u = u_D(\mathbf{x}|\theta) \quad (24.4)$$

式中, u_D 是一个模糊系统, θ 是可调参数的集合。具体地,这个模糊系统 $u_D(\mathbf{x}|\theta)$ 可由以下两步来构造:

步骤 1 对变量 $x_i (i = 1, 2, \dots, n)$, 定义 m_i 个模糊集合 $A_i^{l_i} (l_i = 1, 2, \dots, m_i)$, 式(24.3)中的模糊集合 $P_i^r (i = 1, 2, \dots, L_u)$ 作为特例包含在其中。

步骤 2 用下面的 $\prod_{i=1}^n m_i$ 条模糊规则来构造模糊系统 $u_D(\mathbf{x}|\theta)$

如果 x_1 是 $A_1^{l_1}$ 且...且 x_n 是 $A_n^{l_n}$, 则 u_D 是 $S^{l_1 \cdots l_n}$ (24.5)

其中, $l_i = 1, 2, \dots, m_i, i = 1, 2, \dots, n$, 且当式(24.5)中的 IF 部分与式(24.3)中的 IF 部分相同时, $S^{l_1 \cdots l_n}$ 等于式(24.3)中的 Q^r , 否则, $S^{l_1 \cdots l_n}$ 为任意模糊集。具体地, 使用乘积推理机、单值模糊器和中心平均解模糊器, 可得

$$u_D(\mathbf{x} | \theta) = \frac{\sum_{l_1=1}^{m_1} \cdots \sum_{l_n=1}^{m_n} \bar{y}_u^{l_1 \cdots l_n} \left(\prod_{i=1}^n \mu_{A_i^{l_i}}(x_i) \right)}{\sum_{l_1=1}^{m_1} \cdots \sum_{l_n=1}^{m_n} \left(\prod_{i=1}^n \mu_{A_i^{l_i}}(x_i) \right)} \quad (24.6)$$

选择 $\bar{y}_u^{l_1 \cdots l_n}$ 作为可调节的参数, 把这些参数放在集合 $\theta \in R^{\prod_{i=1}^n m_i}$ 中, 这时模糊控制器变为

$$u_D(\mathbf{x} | \theta) = \theta^T \xi(\mathbf{x}) \quad (24.7)$$

其中, $\xi(\mathbf{x})$ 除了用 m_i 代替 p_i 外, 其余与式(23.14)中的 $\xi(\mathbf{x})$ 相同。

从步骤 2 可知, θ 中的一些参数初始值按规则(24.3)选取, 其余的参数可随机(或按某些战略)选取。因此, 控制知识式(24.3)是通过设置其初始参数而被嵌入到模糊控制器中去的。

24.1.3 自适应律的设计

令 u^* 表示和 23.2 节中式(23.5)一样的理想控制, 其中 $g(\mathbf{x}) = b$ 。把式(24.4)代入式(24.1), 整理可得

$$\dot{e}^{(n)} = -\mathbf{k}^T \mathbf{e} + b[u^* - u_D(\mathbf{x} | \theta)] \quad (24.8)$$

令由 Λ 的定义由式(23.17)给出, $\mathbf{b} = (0, \dots, 0, b)^T$, 则闭环系统动态方程可以改写为如下向量形式

$$\dot{\mathbf{e}} = \Lambda \mathbf{e} + \mathbf{b}[u^* - u_D(\mathbf{x} | \theta)] \quad (24.9)$$

定义最优参数为

$$\theta^* = \arg \min_{\theta \in R^{\prod_{i=1}^n m_i}} \left[\sup_{\mathbf{x} \in R^n} |u_D(\mathbf{x} | \theta) - u^*| \right] \quad (24.10)$$

定义最小逼近误差为

$$w = u_D(\mathbf{x} | \theta^*) - u^* \quad (24.11)$$

由式(24.11)和式(24.7), 可将误差方程(24.9)改写为

$$\dot{\mathbf{e}} = \Lambda \mathbf{e} + \mathbf{b}(\theta^* - \theta)^T \xi(\mathbf{x}) - \mathbf{b}w \quad (24.12)$$

现在来考虑候选 Lyapunov 函数

$$V = \frac{1}{2} \mathbf{e}^T P \mathbf{e} + \frac{b}{2\gamma} (\theta^* - \theta)^T (\theta^* - \theta) \quad (24.13)$$

式中, P 是一个正定矩阵且满足 Lyapunov 方程(23.25), 参数 γ 是正的常数(由假设知 $b > 0$, 所以 V 正定)。由式(24.12)和式(23.25), 可得

$$\dot{V} = -\frac{1}{2}\mathbf{e}^T Q \mathbf{e} + \mathbf{e}^T P \mathbf{b} [(\theta^* - \theta)^T \xi(\mathbf{x}) - w] - \frac{b}{\gamma} (\theta^* - \theta)^T \dot{\theta} \quad (24.14)$$

令 p_n 为 P 的最后一列, 则由 $\mathbf{b} = (0, \dots, 0, b)^T$, 可得 $\mathbf{e}^T P \mathbf{b} = \mathbf{e}^T p_n b$, 于是式(24.14)又变为

$$\dot{V} = -\frac{1}{2}\mathbf{e}^T Q \mathbf{e} + \frac{b}{\gamma} (\theta^* - \theta)^T [\gamma \mathbf{e}^T p_n \xi(\mathbf{x}) - \dot{\theta}] - \mathbf{e}^T p_n b w \quad (24.15)$$

如果选择如下自适应律

$$\dot{\theta} = \gamma \mathbf{e}^T p_n \xi(\mathbf{x}) \quad (24.16)$$

则有

$$\dot{V} = -\frac{1}{2}\mathbf{e}^T Q \mathbf{e} - \mathbf{e}^T p_n b w \quad (24.17)$$

由于 $Q > 0$, w 是最小逼近误差, 所以可以考虑设计嵌入足够多规则的模糊系统 $u_D(\mathbf{x}|\theta)$, 这样 w 将会充分小, 并满足 $|\mathbf{e}^T p_n b w| < \frac{1}{2}\mathbf{e}^T Q \mathbf{e}$, 从而使得 $\dot{V} < 0$ 。

大体上, 整个直接型自适应模糊控制系统的框图如图 24.1 所示。模糊控制规则(24.3) (控制知识)通过模糊控制器初始参数的选择而被包含在模糊控制系统中, 这一过程详见模糊控制器设计程序中的步骤 2。

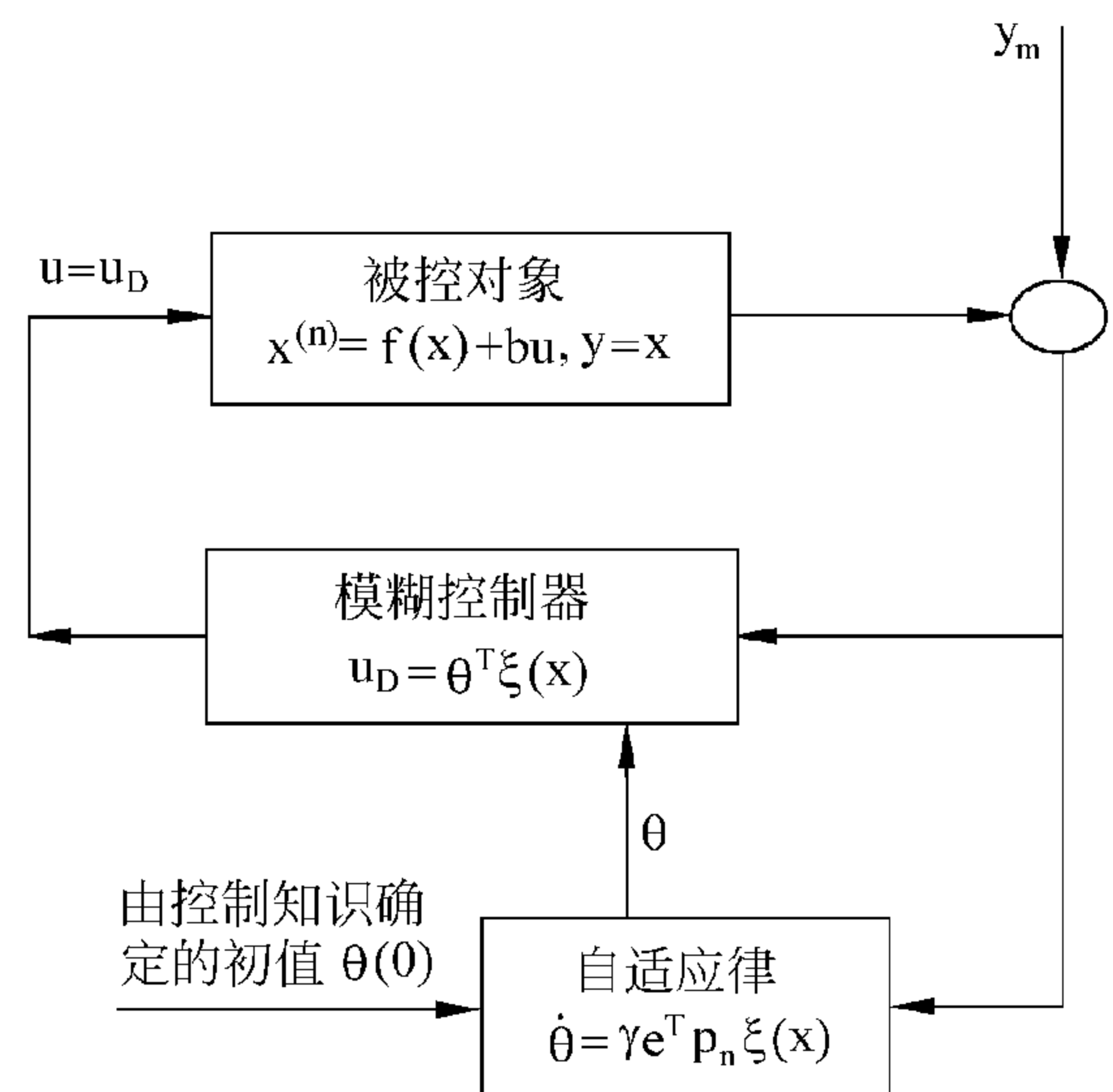


图 24.1 直接型自适应模糊控制系统

24.1.4 仿真

例 24.1 考虑一阶非线性系统(19.31), 此处的目标是, 使用直接型自适应模糊控制器将状态 $x(t)$ 调节至零点, 即 $y_m = 0$ 。显然, 如果控制为零, 则被控对象式(19.31)是不稳定的。实际上, 如果 $u(t) \equiv 0$, 则 $x > 0$ 时, 有 $\dot{x} = \frac{1 - e^{-x}}{1 + e^{-x}} > 0$; 而 $x < 0$ 时, 有 $\dot{x} = \frac{1 - e^{-x}}{1 + e^{-x}} < 0$ 。在仿真中, 选 $\gamma = 1$, 在 $[-3, 3]$ 上定义六个模糊集合, 分别记为 $N_3, N_2, N_1, P_1, P_2, P_3$, 相应的隶属度函数分别为

$$\begin{aligned} \mu_{N_3} &= 1/(1 + \exp(5(x + 2))), \mu_{N_2} = \exp(-(x + 1.5)^2), \mu_{N_1} = \exp(-(x + 0.5)^2), \\ \mu_{P_1} &= \exp(-(x - 0.5)^2), \mu_{P_2} = \exp(-(x - 1.5)^2), \mu_{P_3} = 1/(1 + \exp(-5(x - 2))) \end{aligned}$$

其图形是图 24.2。分别考虑以下两种情况: (i) 不存在任何模糊控制规则, 且初始参数 $\theta_i(0)$ 在区间 $[-2, 2]$ 上随机选取; (ii) 存在如下两种模糊控制规则

如果 x 是 $N2$, 则 $u(x)$ 是 PB (24.18)

如果 x 是 $P2$, 则 $u(x)$ 是 NB (24.19)

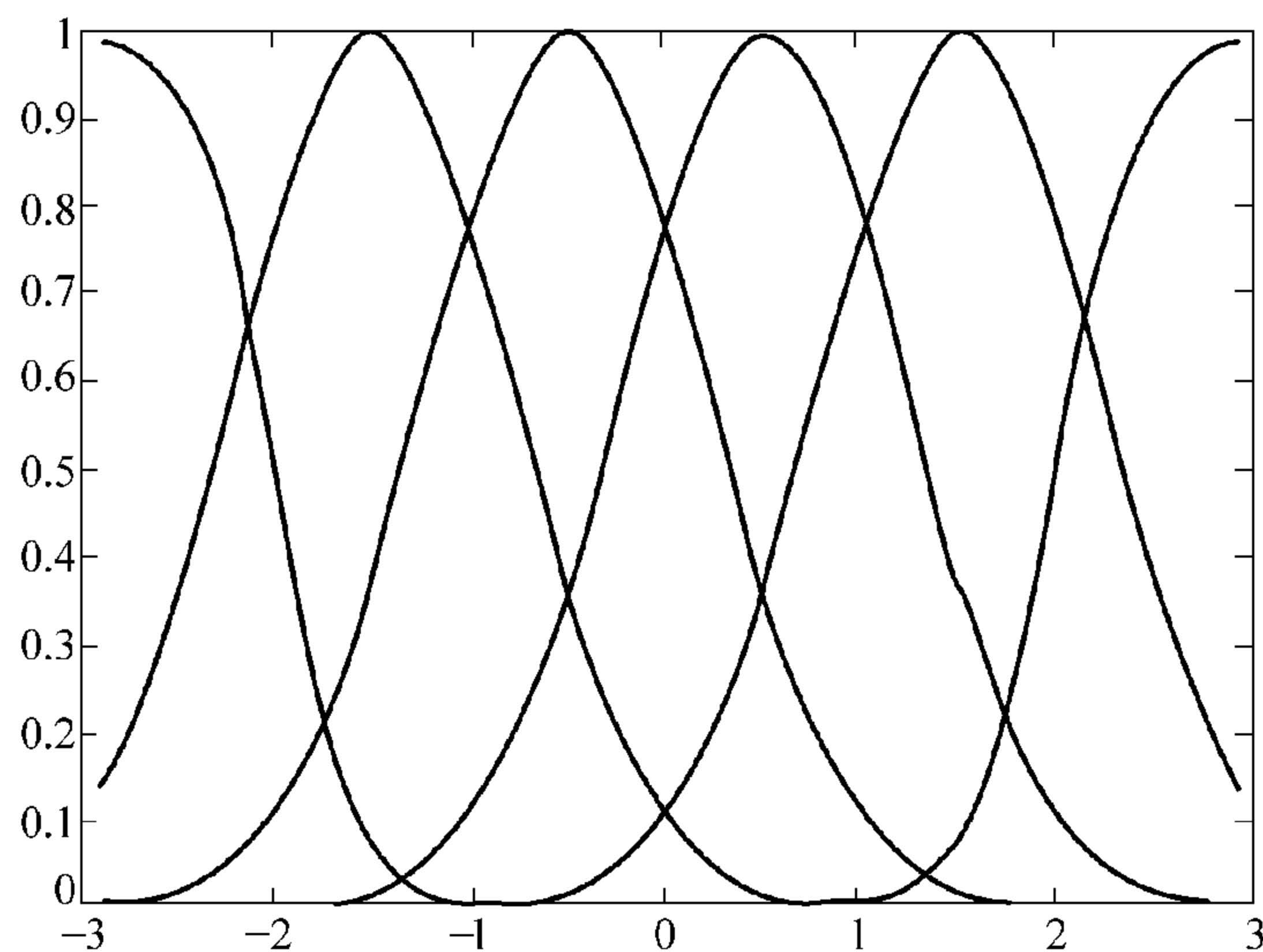


图 24.2 例 24.1 中定义在状态空间上的隶属度函数

其中, $\mu_{PB}(u) = \exp(-(u-2)^2)$, $\mu_{NB}(u) = \exp(-(u+2)^2)$ 。这两条规则是由考虑到要将 $x(t)$ 控制为零而得到的。因此, 若 x 取负值, 则控制 $u(x)$ 应取“较大的正值”(PB) 才能使得 $\dot{x} > 0$ [见式(19.31)]; 反之, 如果 x 取正值, 则控制 $u(x)$ 应该取“较大的负值”(NB) 才能使得 $\dot{x} < 0$ 。图 24.3 和图 24.4 分别给出了不采用和采用式(24.18)和式(24.19)中的语言规则两种情况下 $x(t)$ 的图形, 其初始条件为 $x(0) = 1$ 。由图 24.3 和图 24.4 可以看出: (a) 即

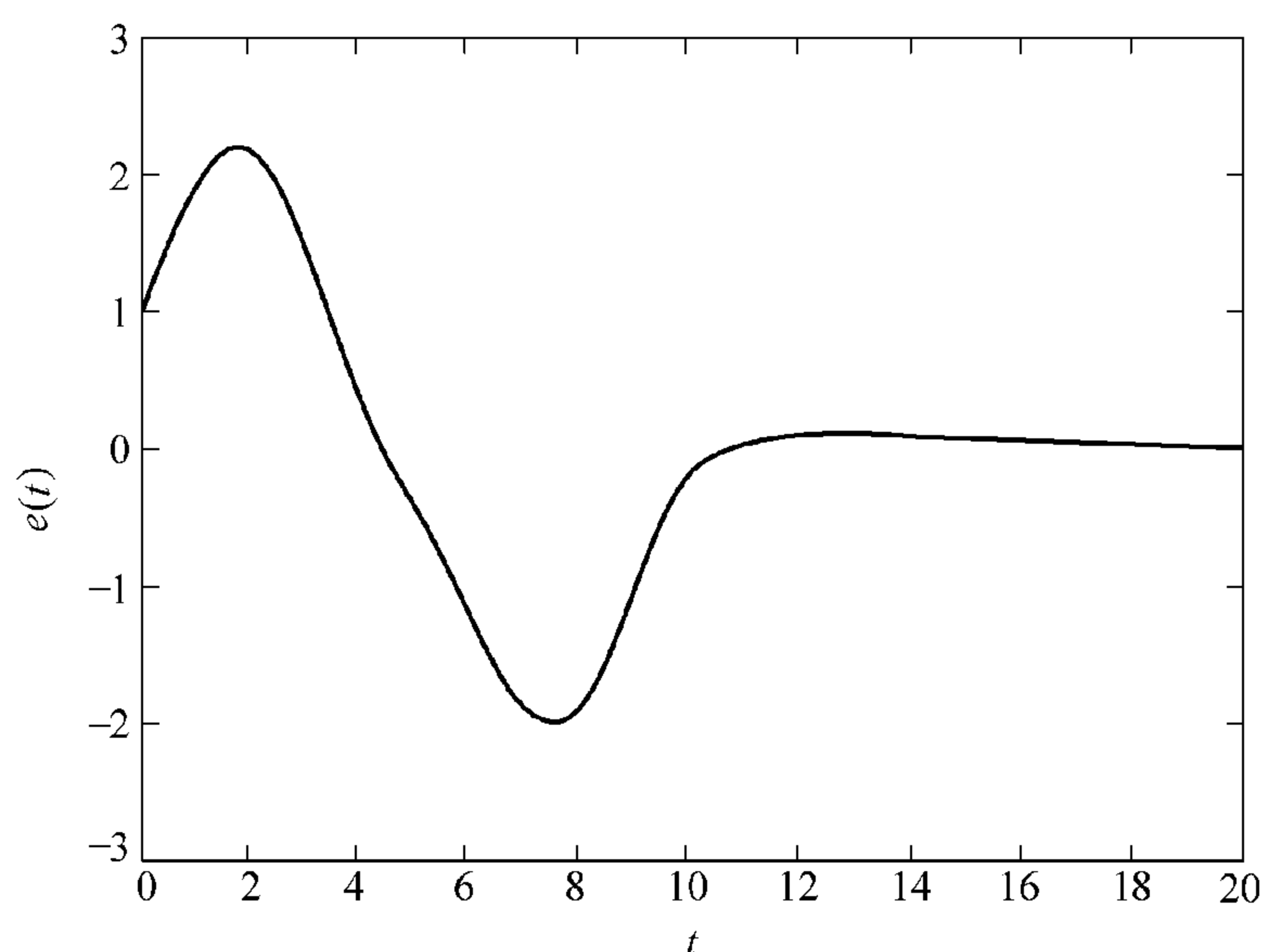


图 24.3 被控对象(19.31)在无模糊控制规则的情况下, 采用直接型自适应模糊控制器所得到的闭环系统的状态 $x(t)$

使没有模糊控制规则(24.18)和(24.19),直接型自适应模糊控制器也能够将被控对象调节至零点;(b)当采用了模糊控制规则后,控制器的收敛速度得到明显提高。

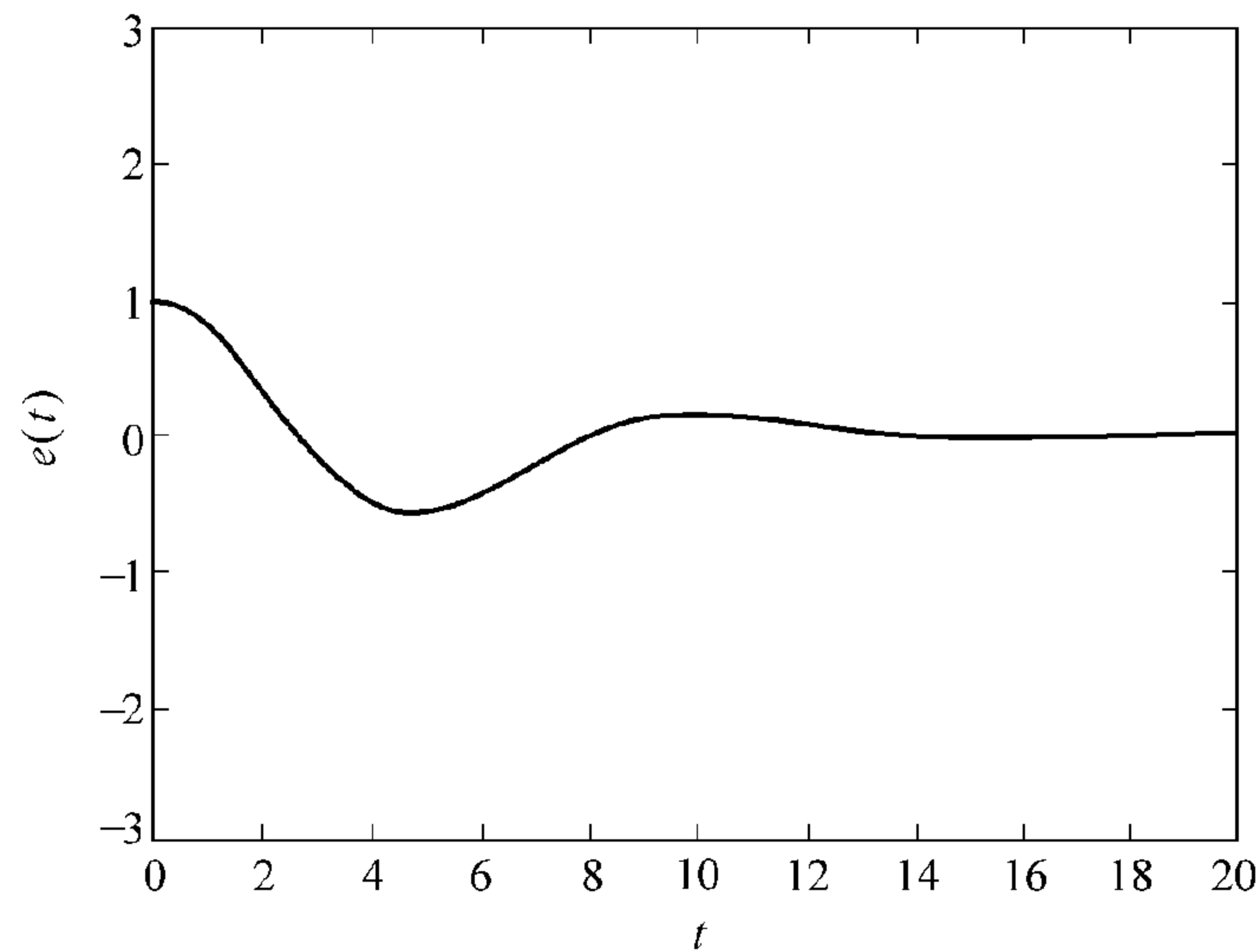


图 24.4 被控对象(19.31)在利用模糊控制规则(24.18)和(24.19)的情况下,采用直接型自适应模糊控制器所得到的闭环系统的状态 $x(t)$

例 24.2 本例研究如下达芬强迫振荡系统(Duffing forced-oscillation system):

$$\dot{x}_1 = x_2 \quad (24.20)$$

$$\dot{x}_2 = -0.1x_2 - x_1^3 + 12\cos(t) + u(t) \quad (24.21)$$

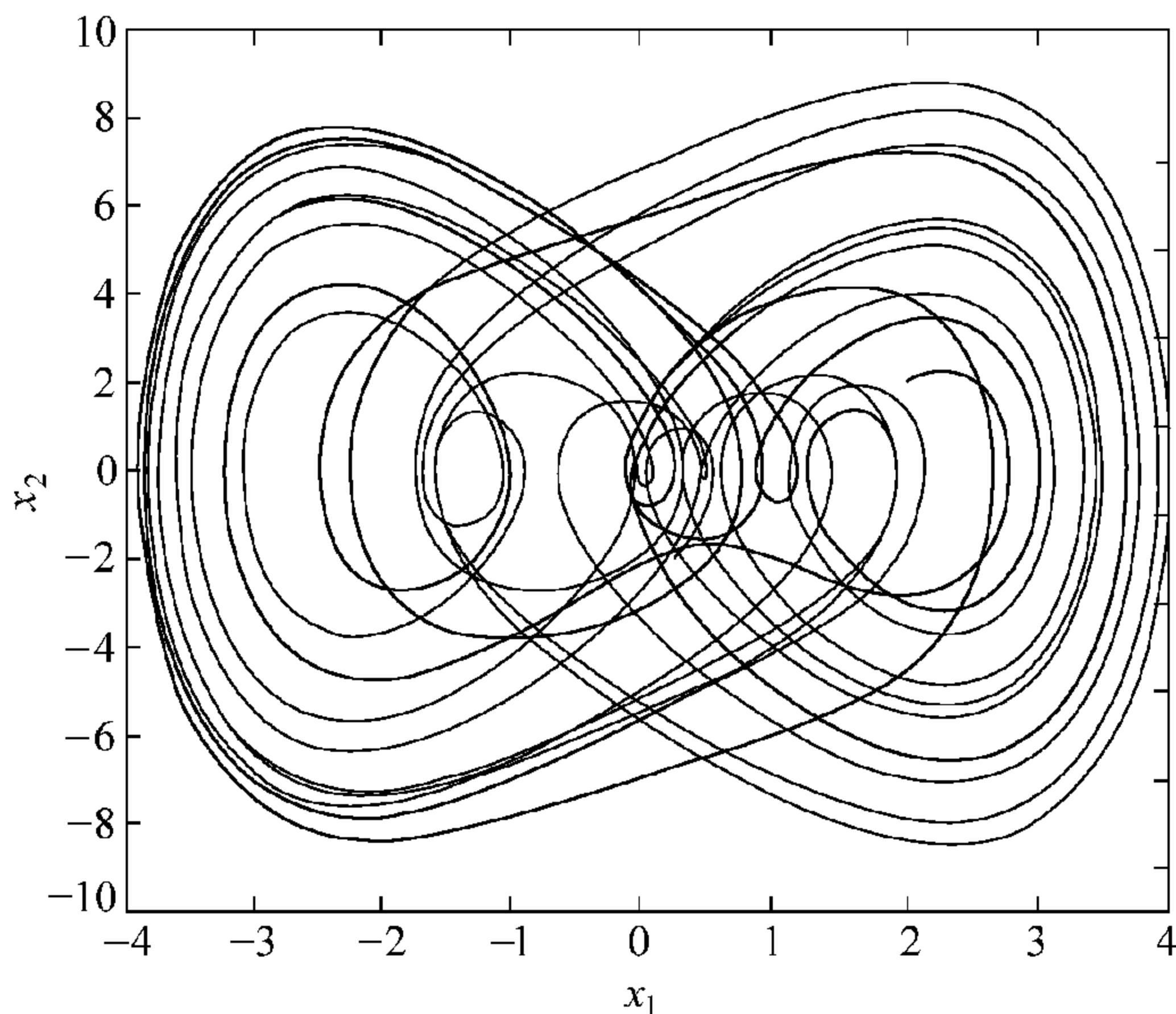


图 24.5 当 $u(t) \equiv 0$ 和 $x_1(0) = x_2(0) = 2$ 时混沌系统(24.20)和(24.21)在相平面 (x_1, x_2) 上的轨迹

如果控制 $u(t)$ 为零,则系统是一个混沌系统。图 24.5 给出了该系统在 $u(t) \equiv 0$, 初始条件为 $x_1(0) = x_2(0) = 2$, 时间周期从 $t_0 = 0$ 到 $t_f = 60$ 时的 (x_1, x_2) 相平面图。现在,采用直接型自适应模糊控制器来控制系统状态 x_1 以跟踪参考轨迹 $y_m(t) = \sin(t)$ 。在相平面中,参考轨迹是一个单位圆 $y_m^2 + \dot{y}_m^2 = 1$ 。仿真中, $k_1 = 2, k_2 = 1, \gamma = 2, Q = \text{diag}(10, 10)$ 。对 x_1 和 x_2 采用图 24.2 所示的 6 个模糊集合,并假设不存在任何模糊控制规则。图 24.6 给出了初始条件为 $x_1(0) = x_2(0) = 2$ 的闭环轨迹,轨迹从 $t_0 = 0$ 开始,到 $t_f = 60$ 结束。

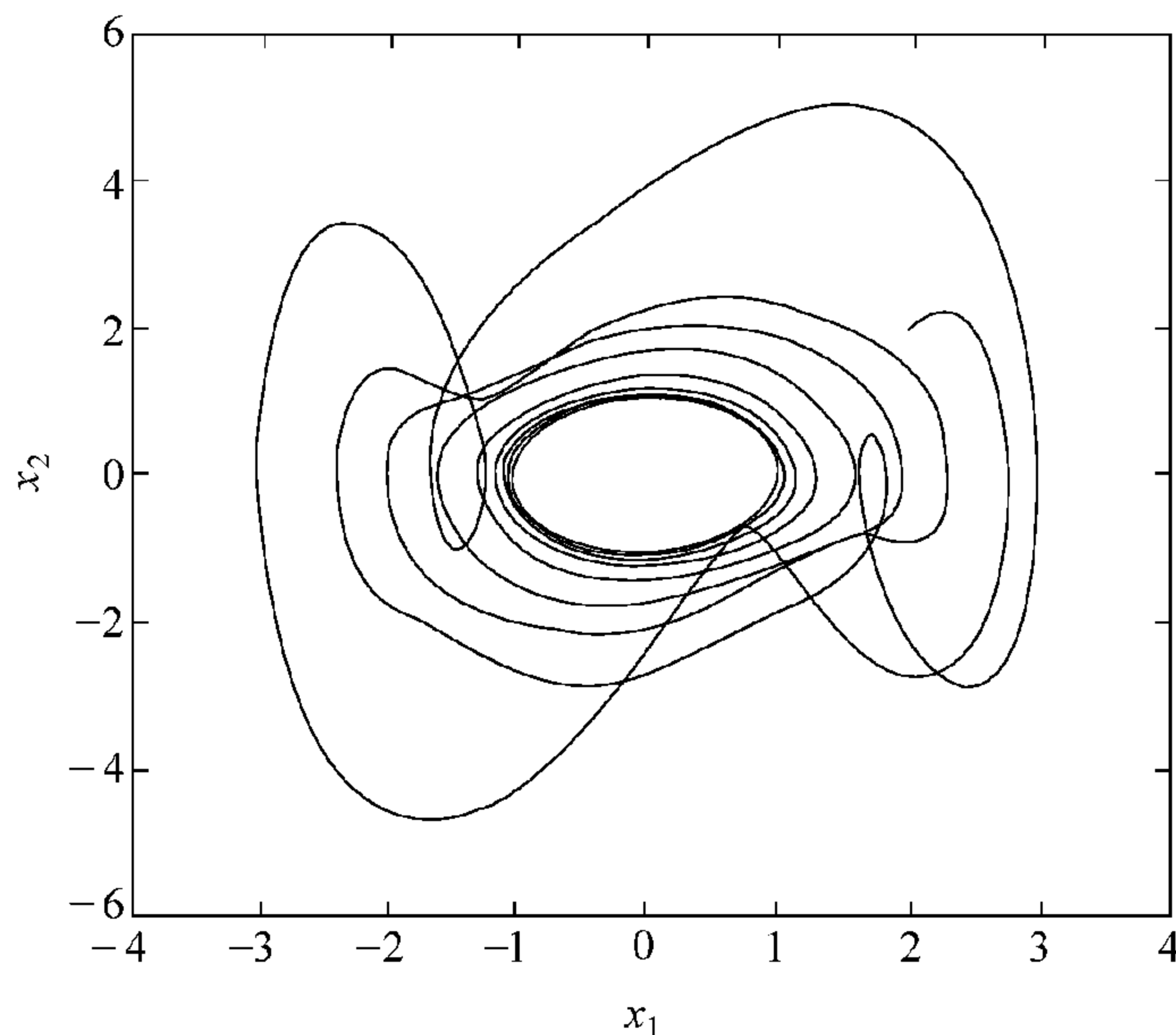


图 24.6 混沌系统(24.20)和(24.21)采用直接型自适应模糊控制器所得到的闭环系统轨迹 $(x_1(t), x_2(t))$

24.2 组合型自适应模糊控制器的设计

间接型自适应模糊控制器可以利用与被控对象有关的语言描述(被控对象知识),直接型自适应模糊控制器可以利用语言控制规则(控制知识),而本节中设计的组合型自适应模糊控制器则可以利用包含了上述两类语言信息的规则。

24.2.1 问题描述

考虑式(24.1)~式(24.2)描述的被控对象,为了简单起见,假设 $b=1$,则可得到如下三条信息:

- 信息 1: 关于被控对象的一个近似模型,即,给定的函数 \hat{f} 是式(24.1)中 f 的估计。
- 信息 2: 给定一个模糊 IF-THEN 规则集合,用其描述在各种条件下 $f - \hat{f}$ 的值;即,有

$$\text{如果 } x_1 \text{ 是 } S_1^j \text{ 且 } \cdots \text{ 且 } x_n \text{ 是 } S_n^j, \text{ 则 } f - \hat{f} \text{ 是 } E^j \quad (24.22)$$

其中, S_i^j, E^j 为 R 中的模糊集合, $j=1, 2, \dots, L_e$ 。

- 信息 3: 给出了形如式(24.3)的模糊 IF-THEN 规则 L_u , 用其描述各种条件下的控制行为。

信息 1 表示被控对象的数学知识(根据系统的结构和物理定律得到), 信息 2 和信息 3 分别为用于逼近的数学模型和不同条件下的控制行为的语言知识。本节的目标是把这三条信息组合嵌入到模糊控制器中, 并为控制器中的可调节参数设计一个自适应律, 使得闭环系统的输出 $y(t)$ 能够跟踪理想输出 $y_m(t)$ 。

24.2.2 模糊控制器的设计

按 23.2.2 小节中的定义, 给出 \mathbf{k} 和 \mathbf{e} 的值。若 $f(\mathbf{x})$ 已知, 则可根据 23.2.2 小节中得到最优控制为

$$u^* = -f(\mathbf{x}) + y_m^{(n)} + \mathbf{k}^T \mathbf{e} \quad (24.23)$$

将上式代入式(24.1)(其中 $b=1$)中就可以保证 $y(t)$ 收敛于 $y_m(t)$ 。令 $\tilde{f}(\mathbf{x}|\theta_l)$ 是由包含规则(24.22)在内的规则所构造的模糊系统, 则根据信息 1 和信息 2 可以得到如下关于 $f(\mathbf{x})$ 最优估计

$$\hat{f}(\mathbf{x}) + \tilde{f}(\mathbf{x}|\theta_l) \quad (24.24)$$

因此, 为利用信息 1 和信息 2, 应该使用下面的控制器

$$u_{12} = -\hat{f}(\mathbf{x}) - \tilde{f}(\mathbf{x}|\theta_l) + y_m^{(n)} + \mathbf{k}^T \mathbf{e} \quad (24.25)$$

信息 3 是由模糊控制规则集合组成的, 为了使用信息 3, 应考虑如下控制器

$$u_3 = u_D(\mathbf{x}|\theta_D) \quad (24.26)$$

式中, u_D 是一个由包含规则(24.3)在内的规则所构造的模糊系统。因此, u_{12} 和 u_3 的加权平均不失为最终控制器的一个好的选择, 即

$$u = \alpha u_{12} + (1 - \alpha) u_3 \quad (24.27)$$

其中 $\alpha \in [0, 1]$ 是权重系数。如果被控对象知识, 即信息 1 和信息 2, 比控制知识信息 3 更重要、更可靠, 则应该选择较大的 α , 否则, 选择较小的 α 。

模糊系统 $\tilde{f}(\mathbf{x}|\theta_l)$ 和 $u_D(\mathbf{x}|\theta_D)$ 可以按与 23.2 小节和 24.1 小节中相同的步骤进行设计, 这里略去细节。可调的参数为 θ_l 和 θ_D , 于是可把 $\tilde{f}(\mathbf{x}|\theta_l)$ 和 $u_D(\mathbf{x}|\theta_D)$ 表示为

$$\tilde{f}(\mathbf{x}|\theta_l) = \theta_l^T \xi(\mathbf{x}) \quad (24.28)$$

$$u_D(\mathbf{x}|\theta_D) = \theta_D^T \eta(\mathbf{x}) \quad (24.29)$$

式中, $\xi(\mathbf{x})$ 和 $\eta(\mathbf{x})$ 是模糊基函数向量。下面的任务和前面一样, 为 θ_l 和 θ_D 设计一个自适应律, 使跟踪误差 \mathbf{e} 尽可能小。

24.2.3 自适应律的设计

把式(24.27)代入式(24.1), 其中 $b=1$, 经过一些直接处理, 可以得到如下闭环系统误差方程

$$\dot{e}^{(n)} = -\mathbf{k}^T \mathbf{e} + \alpha(\hat{f} + \tilde{f} - f) + (1 - \alpha)(u^* - u_D) \quad (24.30)$$

式中, u^* 的定义由式(24.23)给出, $e = y_m - y = y_m - x$, $\mathbf{e} = (e, \dot{e}, \dots, e^{(n-1)})^T$ 。令 \mathbf{A} 和 \mathbf{b} 的定义由式(24.17)给出, 则方程(24.30)可以改写为

$$\dot{\mathbf{e}} = \mathbf{A}\mathbf{e} + \mathbf{b}[\alpha(\hat{f} + \tilde{f} - f) + (1 - \alpha)(u^* - u_D)] \quad (24.31)$$

假定 θ_I 和 θ_D 分别是 M 维和 N 维向量, 其最优值定义为

$$\theta_I^* = \arg \min_{\theta_I \in R^M} [\sup_{\mathbf{x} \in R^n} |\hat{f}(\mathbf{x}) + \tilde{f}(\mathbf{x}|\theta_I) - f(\mathbf{x})|] \quad (24.32)$$

$$\theta_D^* = \arg \min_{\theta_D \in R^N} [\sup_{\mathbf{x} \in R^n} |u^*(\mathbf{x}) - u_D(\mathbf{x}|\theta_D)|] \quad (24.33)$$

最小逼近误差 w 定义为

$$w = \alpha[\hat{f}(\mathbf{x}) + \tilde{f}(\mathbf{x}|\theta_I^*) - f(\mathbf{x})] + (1 - \alpha)[u^*(\mathbf{x}) - u_D(\mathbf{x}|\theta_D^*)] \quad (24.34)$$

则由 w 和式(24.28) ~ 式(24.29), 可将误差方程(24.31)改写为

$$\dot{\mathbf{e}} = \mathbf{A}\mathbf{e} + \mathbf{b}[\alpha(\theta_I - \theta_I^*)^T \xi(\mathbf{x}) + (1 - \alpha)(\theta_D - \theta_D^*)^T \eta(\mathbf{x}) + w] \quad (24.35)$$

考虑候选 Lyapunov 函数

$$V = \frac{1}{2} \mathbf{e}^T P \mathbf{e} + \frac{\alpha}{2\gamma_1} (\theta_I - \theta_I^*)^T (\theta_I - \theta_I^*) + \frac{1 - \alpha}{2\gamma_2} (\theta_D - \theta_D^*)^T (\theta_D - \theta_D^*) \quad (24.36)$$

式中, P 是一个正定矩阵且满足 Lyapunov 方程(23.25), 参数 γ_1 和 γ_2 是正的常数。由式(24.35)和式(23.25), 可得

$$\begin{aligned} \dot{V} = & -\frac{1}{2} \mathbf{e}^T Q \mathbf{e} + \frac{\alpha}{\gamma_1} (\theta_I - \theta_I^*)^T [\dot{\theta}_I + \gamma_1 \mathbf{e}^T P \mathbf{b} \xi(\mathbf{x})] \\ & + \frac{1 - \alpha}{\gamma_2} (\theta_D - \theta_D^*)^T [\dot{\theta}_D + \gamma_2 \mathbf{e}^T P \mathbf{b} \eta(\mathbf{x})] + \mathbf{e}^T P \mathbf{b} w \end{aligned} \quad (24.37)$$

因此, 如果选择如下自适应律

$$\dot{\theta}_I = -\gamma_1 \mathbf{e}^T P \mathbf{b} \xi(\mathbf{x}) \quad (24.38)$$

$$\dot{\theta}_D = -\gamma_2 \mathbf{e}^T P \mathbf{b} \eta(\mathbf{x}) \quad (24.39)$$

则有

$$\dot{V} = -\frac{1}{2} \mathbf{e}^T Q \mathbf{e} + \mathbf{e}^T P \mathbf{b} w \quad (24.40)$$

因为 w 是最小逼近误差, 所以式(24.40)是可以得到的最好结果。

大体上, 整个组合型自适应模糊控制系统的框图如图 24.7 所示。从图中可以看出, 信息 1 是直接用在控制器设计中的, 而信息 2 和信息 3 则是分别通过初始参数 $\theta_I(0)$ 和 $\theta_D(0)$ 而得到应用的。

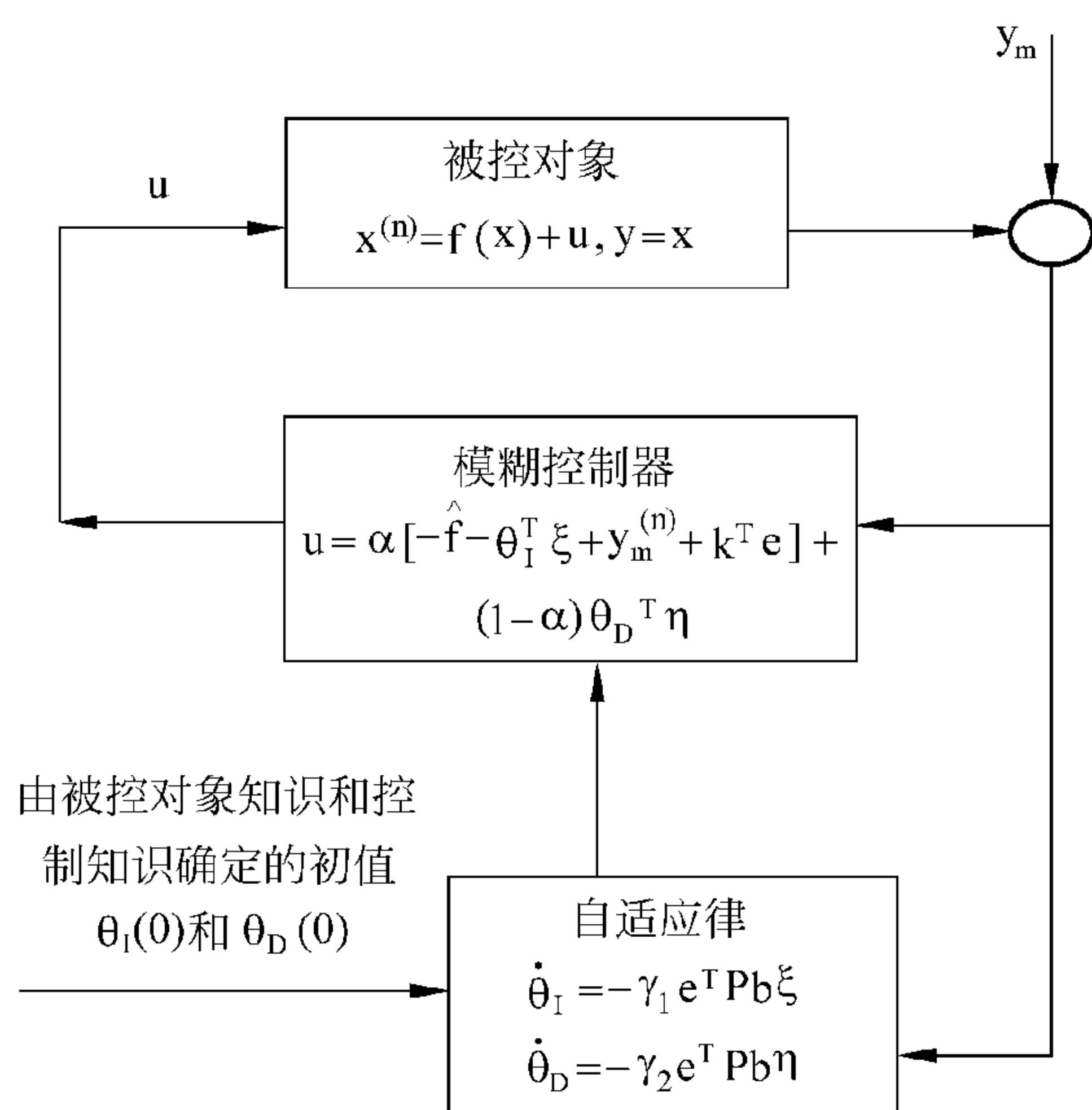


图 24.7 组合型自适应模糊控制系统

24.2.4 收敛性分析

目前,已经介绍了间接型、直接型和组合型自适应模糊控制器的基本设计方法。为了使设计方法在直观上有吸引力,我们采用 Lyapunov 综合法来设计这些自适应模糊控制器。但我们并没有对所生成的自适应模糊控制系统的性能进行详细地分析,只进行两项基本的性能分析:稳定性和收敛性。稳定性指系统中的变量(状态变量和参数变量等)都是有界的,而收敛性指跟踪误差收敛于零。下面的定理将要证明,如果图 24.7 所示模糊控制系统是稳定的,则在适度情况下,跟踪误差将收敛于零。在下一章中,我们将介绍实现稳定性的方法。

定理 24.1 考虑如图 24.7 所示的组合型自适应模糊控制系统,如果状态变量 \mathbf{x} 、参数 θ_i 和 θ_D 及最小逼近误差 w 有界,则

(a) 对任意 $t \geq 0$,跟踪误差都满足

$$\int_0^t |\mathbf{e}(\tau)|^2 d\tau \leq a + b \int_0^t |w(\tau)|^2 d\tau \quad (24.41)$$

式中, a 和 b 是常数, w 是由式(24.34)定义的最小逼近误差。

(b) 如果 w 平方可积,即,若 $\int_0^\infty |w(t)|^2 dt < \infty$,则 $\lim_{t \rightarrow \infty} |\mathbf{e}(t)| = 0$ 。

证明:(a) 令 $\lambda_{Q_{\min}}$ 是 Q 的最小特征值,由式(24.40),可得

$$\begin{aligned} \dot{V} &\leq -\frac{\lambda_{Q_{\min}} - 1}{2} |\mathbf{e}|^2 - \frac{1}{2} (|\mathbf{e}|^2 - 2\mathbf{e}^T P \mathbf{b} w + |P \mathbf{b} w|^2) + \frac{1}{2} |P \mathbf{b} w|^2 \\ &\leq -\frac{\lambda_{Q_{\min}} - 1}{2} |\mathbf{e}|^2 + \frac{1}{2} |P \mathbf{b} w|^2 \end{aligned} \quad (24.42)$$

对式(24.42)的两边同时求积分,并假设 $\lambda_{Q_{\min}} > 1$ (因为 Q 由设计者给出,所以可以找到这样的 Q),可得

$$\begin{aligned} \int_0^t |\mathbf{e}(\tau)|^2 d\tau &\leq \frac{2}{\lambda_{Q_{\min}} - 1} (|V(0)| + |V(t)|) \\ &\quad + \frac{1}{\lambda_{Q_{\min}} - 1} |P \mathbf{b}|^2 \int_0^t |w(\tau)|^2 d\tau \end{aligned} \quad (24.43)$$

定义 $a = \frac{2}{\lambda_{Q_{\min}} - 1} (|V(0)| + \sup_{t \geq 0} |V(t)|)$ 和 $b = \frac{1}{\lambda_{Q_{\min}} - 1} |P \mathbf{b}|^2$,则式(24.43)就变成了式(24.41)。(注意, $\sup_{t \geq 0} |V(t)|$ 是有限的,因为这里假设 \mathbf{e} 、 $\theta_i - \theta_i^*$ 和 $\theta_D - \theta_D^*$ 都是有界的)。

(b) 如果 w 平方可积,则由式(24.41)可知, \mathbf{e} 也是平方可积的。由假设可知,在式(24.35)右端的全部变量都有界,所以 $\dot{\mathbf{e}}$ 有界。由 Barbalat 的引理(Sastry 和 Bodson[1989]) (如果 \mathbf{e} 平方可积且有界, $\dot{\mathbf{e}}$ 也有界,则 $\lim_{t \rightarrow \infty} |\mathbf{e}(t)| = 0$),可得 $\lim_{t \rightarrow \infty} |\mathbf{e}(t)| = 0$ 。

采用同定理 24.1 的证明相同的方法,可以证明,如果如图 23.2 和图 24.1 所示的间接

型、直接型自适应模糊控制系统的全部变量都有界,并且最小逼近误差平方可积,则跟踪误差将收敛到零。

24.3 本章小结及导读

本章主要阐述了以下内容:

- 如何设计直接型自适应模糊控制器;
- 如何把逼近数学模型、语言模型描述和语言控制规则三方面的信息组合成一个单一自适应模糊控制器(组合型自适应模糊控制器);
- 自适应模糊控制系统的收敛性分析。

本章的内容摘自 Wang[1994a],在原文中有更多的仿真结果。其他的自适应模糊控制方法可参考文献有, Spooner 和 Passino [1995], Johansen [1994] 和 Vandegrift, Lewis, Jagannathan 和 Liu[1995]。

24.4 习题

习题 24.1 考虑如图 23.2 所示的间接型自适应模糊控制系统,证明闭环控制系统的跟踪误差满足有界条件(24.41)。

习题 24.2 证明如图 24.1 所示的直接型自适应模糊控制系统的跟踪误差满足有界条件(24.41)。

习题 24.3 假设用描述未知函数 f 的模糊 IF-THEN 规则来代替 24.2 小节中的信息 1 和信息 2,试把这些规则和信息 3 组合起来设计一个自适应模糊控制器。

习题 24.4 假设式(24.6)中模糊控制器 u_D 的隶属度函数 μ_{A_i} 中的参数也是可调参数 (u_D 与参数 θ 不像式(24.7)中那样成线性关系),试为此情况设计一个直接型自适应模糊控制器。

习题 24.5 使用不同的初始条件,重做图 24.3 中的仿真。

习题 24.6 对例 24.1 和例 24.2 中系统的组合型自适应模糊控制器进行仿真。

高级自适应模糊控制器 I

在前面的第 23 章至第 24 章中,我们学习了间接型、直接型和组合型自适应模糊控制器的基本内容。其思路都是使用 Lyapunov 综合法和确定等效原理。尽管我们没有给予更多的理论分析,但在定理 24.1 中已经证明了,如果系统中所有的变量有界且最小逼近误差平方可积,则跟踪误差收敛于零。因此,关键问题是:(i)如何保证所有变量的有界性;(ii)如何使得最小逼近误差平方可积。因为模糊系统是万能逼近器,所以我们可以通过使用更多的规则来构造模糊系统的方法,使得最小逼近误差任意小。本章的主要目标是解决第一个问题。在 25.1 节和 25.2 节中,将分别使用监督控制法和参数投影法来保证系统的状态有界和参数有界。

25.1 保证状态有界的监督控制法

25.1.1 在间接型自适应模糊控制系统中的应用

考虑如图 23.2 所示的间接型自适应模糊控制系统,即被控对象是由式(23.1)~式(23.2)所表示的,模糊系统 \hat{f} 和 \hat{g} 是由式(23.12)~式(23.13)给出的,其控制器是根据式(23.7)设计的,自适应律是由式(23.27)~式(23.28)给出的。现在的任务是使得系统的状态变量 \mathbf{x} 有界。

第 20 章曾用监督控制的思想来达到这个目的。具体地,在由式(23.7)设计的基本控制器 u_l 的顶端增加一个监督控制器 u_s ,于是控制器变为

$$u = u_l + u_s \quad (25.1)$$

u_s 应该设计为:(i)保证 \mathbf{x} 的有界性;(ii)以监督控制的方式运作,即当状态变量 \mathbf{x} 正好在稳态范围内时, u_s 应为零,而当状态变量 \mathbf{x} 趋于远离稳态时, u_s 不为零。现在,我们设计一个这样的 u_s 。

用 $u = u_l + u_s$ 代替 $u = u_l$,这样误差方程(23.18)变为

$$\dot{\mathbf{e}} = \Lambda \mathbf{e} + \mathbf{b} [(\hat{f}(\mathbf{x}|\theta_f) - \hat{f}(\mathbf{x})) + (\hat{g}(\mathbf{x}|\theta_g) - \hat{g}(\mathbf{x}))u_l - g(\mathbf{x})u_s] \quad (25.2)$$

进一步定义

$$V_e = \frac{1}{2} \mathbf{e}^T P \mathbf{e} \quad (25.3)$$

式中, P 是一个正定矩阵且满足 Lyapunov 方程(23.25)。因为此处假设 $\mathbf{e} = (y_m - x_1, \dot{y}_m - x_2, \dots, y_m^{(n-1)} - x_n)^T$ 和 $y_m, \dot{y}_m, \dots, y_m^{(n-1)}$ 全部有界, 所以 V_e 有界也就是 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ 有界。于是任务变为设计 u_s , 使得 $V_e \leq \bar{V}$, 其中 \bar{V} 为一个给定的常数, 其值可根据 \mathbf{x} 的边界

来确定。由式(25.3), 有 $(V_e)^{1/2} \geq \left(\frac{\lambda_{P\min}}{2}\right)^{1/2} |\mathbf{e}| \geq \left(\frac{\lambda_{P\min}}{2}\right)^{1/2} (|\mathbf{x}| - |\mathbf{y}_m|)$, 其中 $\lambda_{P\min}$ 是 P 的最

小特征值, $\mathbf{y}_m = (y_m, \dot{y}_m, \dots, y_m^{(n-1)})^T$ 。从而可知 $V_e \leq \bar{V}$ 等价于 $|\mathbf{x}| \leq |\mathbf{y}_m| + \left(\frac{2\bar{V}}{\lambda_{P\min}}\right)^{1/2}$ 。因此,

要想使 $|\mathbf{x}| \leq M_x$ (其中 M_x 是一个常数), 则可以选择

$$\bar{V} = \frac{\lambda_{P\min}}{2} (M_x - \sup_{t \geq 0} |\mathbf{y}_m|)^2 \quad (25.4)$$

因为 $V_e \geq 0$, 所以保证 $V_e \leq \bar{V}$ 的一个方法是设计 u_s , 使得当 $V_e \geq \bar{V}$ 时, $\dot{V}_e < 0$ 。由式(25.2)和式(23.25), 可得

$$\begin{aligned} \dot{V}_e &= -\frac{1}{2} \mathbf{e}^T Q \mathbf{e} + \mathbf{e}^T P \mathbf{b} [(\hat{f} - f) + (\hat{g} - g)u_l - gu_s] \\ &\leq -\frac{1}{2} \mathbf{e}^T Q \mathbf{e} + |\mathbf{e}^T P \mathbf{b}| [|\hat{f}| + |f| + |\hat{g}u_l| + |gu_l|] - \mathbf{e}^T P \mathbf{b} g u_s \end{aligned} \quad (25.5)$$

要设计满足上式的右端为负的 u_s , 还需知道 f 和 g 的边界。即, 还要做如下假设:

假设 25.1 对所有 $\mathbf{x} \in R^n$, 都存在函数 $f^U(\mathbf{x})$, $g^U(\mathbf{x})$ 和 $g_L(\mathbf{x})$, 使 $|f(\mathbf{x})| \leq f^U(\mathbf{x})$ 和 $0 < g_L(\mathbf{x}) \leq g(\mathbf{x}) \leq g^U(\mathbf{x})$ 成立。

根据上面的 f^U , g^U 和 g_L , 并观察式(25.5), 可选定如下监督控制器 u_s

$$u_s = I^* \operatorname{sgn}(\mathbf{e}^T P \mathbf{b}) \frac{1}{g_L(\mathbf{x})} [|\hat{f}(\mathbf{x}|\theta_f)| + f^U(\mathbf{x}) + |\hat{g}(\mathbf{x}|\theta_g)u_l| + g^U(\mathbf{x})u_l] \quad (25.6)$$

其中, 当 $V_e \geq \bar{V}$ 时; $I^* = 1$; 当 $V_e < \bar{V}$ 时, $I^* = 0$; 当 $y \geq 0$ 时, $\operatorname{sgn}(y) = 1$; 当 $y < 0$ 时, $\operatorname{sgn}(y) = -1$ 。将式(25.6)代入式(25.5), 可知, 如果 $V_e \geq \bar{V}$, 则

$$\begin{aligned} \dot{V}_e &\leq -\frac{1}{2} \mathbf{e}^T Q \mathbf{e} + |\mathbf{e}^T P \mathbf{b}| [|\hat{f}| + |f| + |\hat{g}u_l| + |gu_l| - \frac{g}{g_L} (|\hat{f}| + |f^U| \\ &\quad + |\hat{g}u_l| + |g^U u_l|)] \leq -\frac{1}{2} \mathbf{e}^T Q \mathbf{e} < 0 \end{aligned} \quad (25.7)$$

当 $V_e \geq \bar{V}$ 时, 可以自然地假设 $\mathbf{e} \neq 0$ 。所以, 使用 u_s 由式(25.6)给出的控制器(25.1)可以保证 $V_e \leq \bar{V}$, 从而, 对任意给定常数 M_x , 当 \bar{V} 根据式(25.4)选择时, 有 $|\mathbf{x}| \leq M_x$ 。

根据指示函数 I^* 的定义, 只有当 $V_e \geq \bar{V}$ 时, u_s 才是非零的, 所以此时的 u_s 是一个监督控制器。也就是说, 如果带有模糊控制器 u_l (u_l 由式(23.7)给出) 的闭环系统具有良

好的性能,能使误差在约束集内(即 $V_e \leq \bar{V}$ 或等价地 $|\mathbf{x}| \leq M_x$),则此时的监督控制器 u_s 停止工作;反之,当闭环系统趋于不稳定(即 $V_e \geq \bar{V}$)时,监督控制器 u_s 将开始工作,以迫使 $V_e \leq \bar{V}$ 。

25.1.2 在直接型自适应模糊控制系统中的应用

考虑如图 24.1 所示的基本直接型自适应模糊控制系统,其系统过程模型是由式(24.1)~式(24.2)给出的,控制器是由式(24.7)设计的,自适应律是由式(24.16)给出的。现在的任务是为模糊控制器 u_D 的顶端设计一个监督控制器 u_s ,使得系统的状态变量 \mathbf{x} 有界。

令控制为

$$u = u_D(\mathbf{x}|\theta) + u_s(\mathbf{x}) \quad (25.8)$$

理想控制 u^* 的定义由式(23.5)给出(其中 $g(\mathbf{x}) = b$)。将式(25.8)代入到式(24.1)中,整理可得闭环系统的误差方程为

$$\dot{\mathbf{e}} = \mathbf{A}\mathbf{e} + \mathbf{b}[u^* - u_D(\mathbf{x}|\theta) - u_s(\mathbf{x})] \quad (25.9)$$

式中, \mathbf{A} 和 \mathbf{b} 与式(24.9)所定义的相同。由式(25.3)定义的候选 Lyapunov 函数 V_e ,由式(25.9)和式(23.25),可得

$$\begin{aligned} \dot{V}_e &= -\frac{1}{2}\mathbf{e}^T Q\mathbf{e} + \mathbf{e}^T P\mathbf{b}[u^* - u_D(\mathbf{x}|\theta) - u_s(\mathbf{x})] \\ &\leq -\frac{1}{2}\mathbf{e}^T Q\mathbf{e} + |\mathbf{e}^T P\mathbf{b}|(|u^*| + |u_D|) - \mathbf{e}^T P\mathbf{b}u_s(\mathbf{x}) \end{aligned} \quad (25.10)$$

为了设计 u_s ,以使 $\dot{V}_e < 0$,必须做如下假设,其基本思想与假设 25.1 相同。

假设 25.2 存在函数 $f^U(\mathbf{x})$ 和常数 b_L ,使得 $|f(\mathbf{x})| \leq f^U(\mathbf{x})$ 且 $0 < b_L \leq b$ 。

通过观察式(25.10)和式(23.5),可设计如下监督控制器 u_s

$$u_s(\mathbf{x}) = I^* \operatorname{sgn}(\mathbf{e}^T P\mathbf{b}) \left[|u_D| + \frac{1}{b_L}(f^U + |y_m^{(n)}| + |\mathbf{k}^T \mathbf{e}|) \right] \quad (25.11)$$

式中指示函数 I^* 与式(25.6)中的相同。即,当 $V_e \geq \bar{V}$ 时, $I^* = 1$,当 $V_e < \bar{V}$ 时, $I^* = 0$ 。同式(25.4)中的情况一样 \bar{V} 可由 \mathbf{x} 的取值范围确定。将式(25.11)和式(23.5)代入式(25.10),并考虑 $I^* = 1$ 的情况,可得

$$\begin{aligned} \dot{V}_e &\leq -\frac{1}{2}\mathbf{e}^T Q\mathbf{e} + |\mathbf{e}^T P\mathbf{b}| \left[\frac{1}{b}(|f| + |y_m^{(n)}| + |\mathbf{k}^T \mathbf{e}|) - \frac{1}{b_L}(f^U + |y_m^{(n)}| + |\mathbf{k}^T \mathbf{e}|) \right] \\ &\leq -\frac{1}{2}\mathbf{e}^T Q\mathbf{e} < 0 \end{aligned} \quad (25.12)$$

因此,使用监督控制器 u_s ,可以保证 $V_e \leq \bar{V}$,于是对任意常数 M_x ,若 \bar{V} 根据式(25.4)选取,则有 $|\mathbf{x}| \leq M_x$ 。

组合型自适应模糊控制系统的监督控制器的设计留作习题。

25.2 保证参数有界的投影法

25.2.1 在间接型自适应模糊控制系统中的应用

基本的自适应律(23.27) ~ (23.28)不能保证参数 θ_f 和 θ_g 有界。如果 θ_f 发散到无穷大,则模糊系统 $f^\wedge(\mathbf{x} | \theta_f)$ 将稳步增大,导致控制 u_f 无界,这显然是不能接受的。因此,要想设计一个稳定系统(即所有变量有界),就必须修改自适应律,使参数有界。

定义参数 θ_f 和 θ_g 的约束集 Ω_f 和 Ω_g 为

$$\Omega_f = \left\{ \theta_f \in R^{\prod_{i=1}^n p_i} \mid \|\theta_f\| \leq M_f \right\} \quad (25.13)$$

$$\Omega_g = \left\{ \theta_g \in R^{\prod_{i=1}^n q_i} \mid 0 < \varepsilon \leq \|\theta_g\| \leq M_g \right\} \quad (25.14)$$

式中, p_i 和 q_i 的定义由 23.2.2 小节给出, M_f 、 ε 和 M_g 是常数。因为由式(23.7)可知, $\hat{g} = \theta_g^T \eta$ 必须非零,所以由 $\varepsilon > 0$, 可得 $\|\theta_g\|$ 有界。现在修正基本自适应律(23.27) ~ (23.28), 以保证 $\theta_f \in \Omega_f$ 和 $\theta_g \in \Omega_g$ 。

基本想法如下: 如果参数向量在约束集内部或边界上, 且正向约束集内部移动, 则使用基本自适应律(23.27) ~ (23.28); 如果参数向量在约束集边界上, 且正向约束集外部移动, 则将参数向量的梯度映射到支撑超平面上(见图 25.1)。具体地, 用投影法修正间接型自适应模糊控制系统的自适应律的方法如下。

用投影法确定自适应律:

- 对 θ_f , 使用

$$\dot{\theta}_f = \begin{cases} -\gamma_1 \mathbf{e}^T \mathbf{P} \mathbf{b} \xi(\mathbf{x}) & \text{如果 } (\|\theta_f\| < M_f) \text{ 或 } (\|\theta_f\| = M_f \text{ 且 } \mathbf{e}^T \mathbf{P} \mathbf{b} \theta_f^T \xi(\mathbf{x}) \geq 0) \\ P \{ -\gamma_1 \mathbf{e}^T \mathbf{P} \mathbf{b} \xi(\mathbf{x}) \} & \text{如果 } (\|\theta_f\| = M_f \text{ 且 } \mathbf{e}^T \mathbf{P} \mathbf{b} \theta_f^T \xi(\mathbf{x}) < 0) \end{cases} \quad (25.15)$$

其中, 投影算子 $P\{*\}$ 的定义如下

$$P \{ -\gamma_1 \mathbf{e}^T \mathbf{P} \mathbf{b} \xi(\mathbf{x}) \} = -\gamma_1 \mathbf{e}^T \mathbf{P} \mathbf{b} \xi(\mathbf{x}) + \gamma_1 \mathbf{e}^T \mathbf{P} \mathbf{b} \frac{\theta_f \theta_f^T \xi(\mathbf{x})}{\|\theta_f\|^2} \quad (25.16)$$

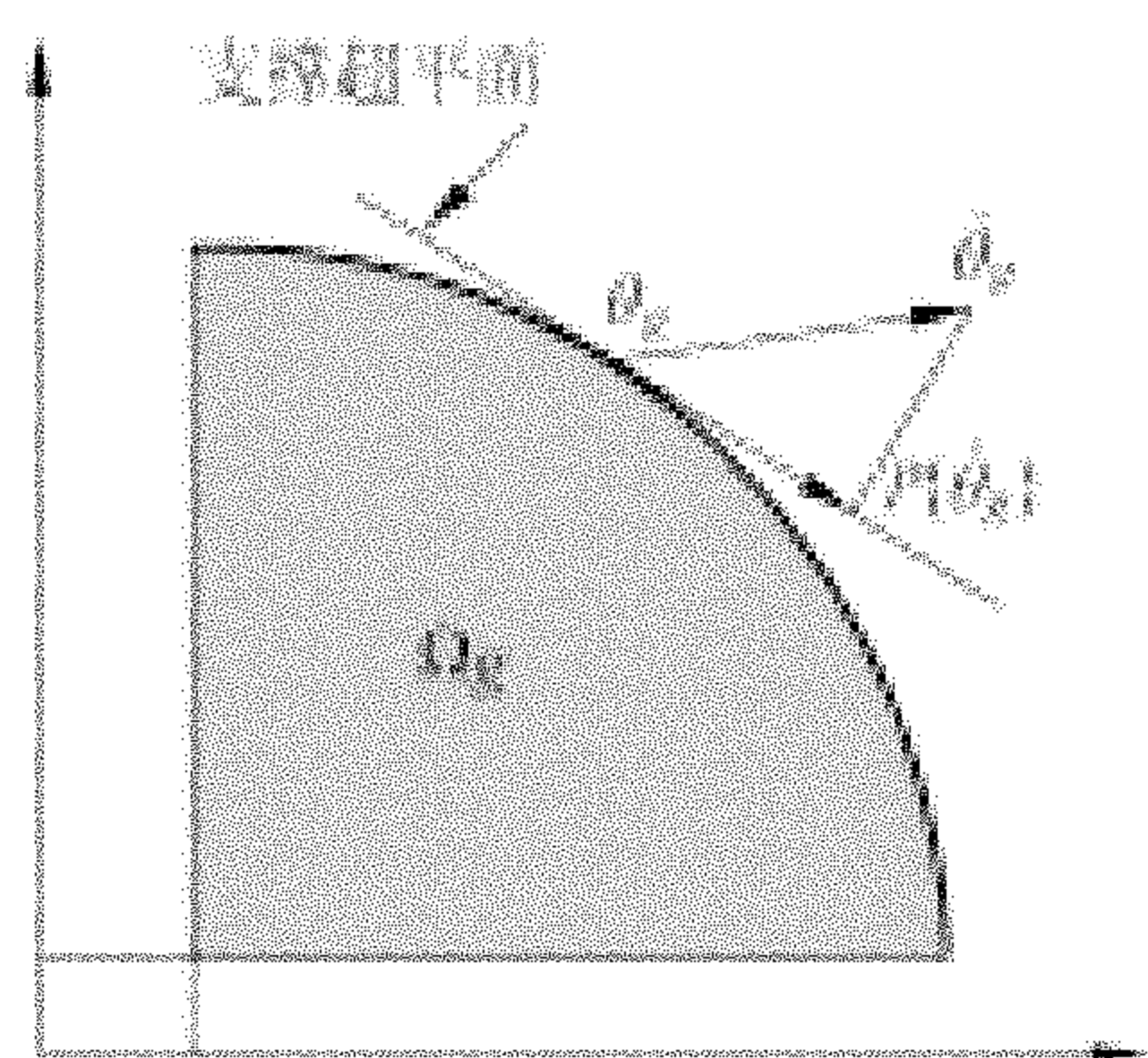


图 25.1 投影算法示意图

• 对 θ_g , 使用

(a) 只要 θ_g 中的某个元素 $\theta_{gi} = \varepsilon$, 就使用

$$\dot{\theta}_{gi} = \begin{cases} -\gamma_2 \mathbf{e}^T P \mathbf{b} \eta_i(\mathbf{x}) u_l & \text{如果 } \mathbf{e}^T P \mathbf{b} \eta_i(\mathbf{x}) u_l < 0 \\ 0 & \text{如果 } \mathbf{e}^T P \mathbf{b} \eta_i(\mathbf{x}) u_l \geq 0 \end{cases} \quad (25.17)$$

式中 $\eta_i(\mathbf{x})$ 为 $\eta(\mathbf{x})$ 第 i 个分量。

(b) 否则, 使用

$$\dot{\theta}_g = \begin{cases} -\gamma_2 \mathbf{e}^T P \mathbf{b} \eta(\mathbf{x}) u_l & \text{如果 } (|\theta_g| < M_g) \text{ 或 } (|\theta_g| = M_g \text{ 且 } \mathbf{e}^T P \mathbf{b} \theta_g^T \eta(\mathbf{x}) u_l \geq 0) \\ P \{-\gamma_2 \mathbf{e}^T P \mathbf{b} \eta(\mathbf{x}) u_l\} & \text{如果 } (|\theta_g| = M_g \text{ 且 } \mathbf{e}^T P \mathbf{b} \theta_g^T \eta(\mathbf{x}) u_l < 0) \end{cases} \quad (25.18)$$

其中, 投影算子 $P\{*\}$ 的定义为

$$P\{-\gamma_2 \mathbf{e}^T P \mathbf{b} \eta(\mathbf{x}) u_l\} = -\gamma_2 \mathbf{e}^T P \mathbf{b} \eta(\mathbf{x}) u_l + \gamma_2 \mathbf{e}^T P \mathbf{b} \frac{\theta_g \theta_g^T \eta(\mathbf{x}) u_l}{|\theta_g|^2} \quad (25.19)$$

下面的定理证明了, 修正的自适应律(25.15) ~ (25.19)可以保证 $\theta_f \in \Omega_f$ 和 $\theta_g \in \Omega_g$ 。

定理 25.1 设约束集 Ω_f 和 Ω_g 的定义由式(25.13) ~ 式(25.14)给出。若参数的初始值满足 $\theta_f(0) \in \Omega_f$ 和 $\theta_g(0) \in \Omega_g$, 则对任意 $t \geq 0$, 自适应律(25.15) ~ (25.19)都能保证 $\theta_f(t) \in \Omega_f$ 和 $\theta_g(t) \in \Omega_g$ 。

证明: 为证明 $|\theta_f| \leq M_f$, 先令 $V_f = \frac{1}{2} \theta_f^T \theta_f$ 。如果式(25.15)的第一行条件成立, 则当 $|\theta_f| = M_f$ 时, 要么有 $|\theta_f| \leq M_f$, 要么有 $\dot{V}_f = -\gamma_1 \mathbf{e}^T P \mathbf{b} \theta_f^T \xi(\mathbf{x}) \leq 0$ 。因此, 在此情况下, 有 $|\theta_f| \leq M_f$, 如果式(25.15)的第二行条件成立, 则有 $|\theta_f| = M_f$ 和 $\dot{V}_f = -\gamma_1 \mathbf{e}^T P \mathbf{b} \theta_f^T \xi(\mathbf{x}) + \gamma_1 \mathbf{e}^T P \mathbf{b} \frac{\theta_f^T \theta_f}{|\theta_f|^2} \theta_f^T \xi(\mathbf{x}) = 0$ 。因此, 在此情况下, 有 $|\theta_f| \leq M_f$ 成立。因为初始条件为 $|\theta_f(0)| \leq M_f$, 所以对 $t \geq 0$, 有 $|\theta_f(t)| \leq M_f$ 。

用同样的方法, 可以证明, 对任意 $t \geq 0$, 都有 $|\theta_g(t)| \leq M_g$ 。下面证明 $|\theta_g| \geq \varepsilon$ 。由式(25.17)可以看出, 如果 $\theta_{gi} = \varepsilon$, 则 $\dot{\theta}_{gi} \geq 0$; 因此, 总有 $\theta_{gi} \geq \varepsilon$, 使得 $|\theta_g| \geq \varepsilon$ 。

25.2.2 在直接型自适应模糊控制系统中的应用

想法是和前面一样的, 令直接型自适应模糊控制器(24.7)中 θ 的约束集为

$$\Omega_D = \left\{ \theta \in R^{\prod_{i=1}^n m_i} \mid \|\theta\| \leq M_D \right\} \quad (25.20)$$

用投影法得到修正自适应律为

$$\dot{\theta} = \begin{cases} \gamma \mathbf{e}^T p_n \xi(\mathbf{x}) & \text{如果 } (|\theta| < M_D) \text{ 或 } (|\theta| = M_D \text{ 且 } \mathbf{e}^T p_n \theta^T \xi(\mathbf{x}) \geq 0) \\ P\{\gamma \mathbf{e}^T p_n \xi(\mathbf{x})\} & \text{如果 } (|\theta| = M_D \text{ 且 } \mathbf{e}^T p_n \theta^T \xi(\mathbf{x}) < 0) \end{cases} \quad (25.21)$$

式中投影算子 $P\{*\}$ 定义为

$$P\{\gamma \mathbf{e}^T p_n \xi(\mathbf{x})\} = \gamma \mathbf{e}^T p_n \xi(\mathbf{x}) - \gamma \mathbf{e}^T p_n \frac{\theta \theta^T \xi(\mathbf{x})}{|\theta|^2} \quad (25.22)$$

下面的定理将证明参数有界。

定理 25.2 设参数约束集 Ω_D 的定义由式(25.20)给出, 如果初始值 $\theta(0) \in \Omega_D$, 则对任意 $t \geq 0$, 自适应律(25.21)都可以保证 $\theta(t) \in \Omega_D$ 。

此定理的证明方法和定理 25.1 的证明方法相同, 留作习题。

25.3 稳定的直接型自适应模糊控制系统

25.3.1 稳定性和收敛性分析

在如图 24.1 所示的基本自适应模糊控制系统上添加监督控制器(25.11), 并采用修正自适应律(25.21), 即可得到如图 25.2 所示的高级直接型自适应模糊控制系统。下面的定理表明了这个自适应模糊控制系统的特性。

定理 25.3 考虑如图 25.2 所示的自适应模糊控制系统, 即被控对象为形如式(24.1) ~ 式(24.2)的系统, 控制器 $u = u_D(\mathbf{x}|\theta) + u_s(\mathbf{x})$ 中的模糊系统根据式(24.6)或式(24.7)设计, 监督控制器 $u_s(\mathbf{x})$ 根据式(25.11)设计, 自适应律由式(25.21) ~ 式(25.22)给出, 则这个自适应模糊控制系统必有如下特性:

(a) 所有的参数和状态变量都有界, 即

$$|\theta(t)| \leq M_D \quad (25.23)$$

$$|\mathbf{x}(t)| \leq M_x \quad (25.24)$$

式中, M_D 和 M_x 是常数, 确定 u_s 中 I^* 的 \bar{V} 根据式(25.4)选择。

(b) 跟踪误差 e 由最小逼近误差 w (其定义是式(24.11))给出边界

$$\int_0^t |\mathbf{e}(\tau)|^2 d\tau \leq a + b \int_0^t |w(\tau)|^2 d\tau \quad \forall t \leq 0 \quad (25.25)$$

式中, a 和 b 为常数。

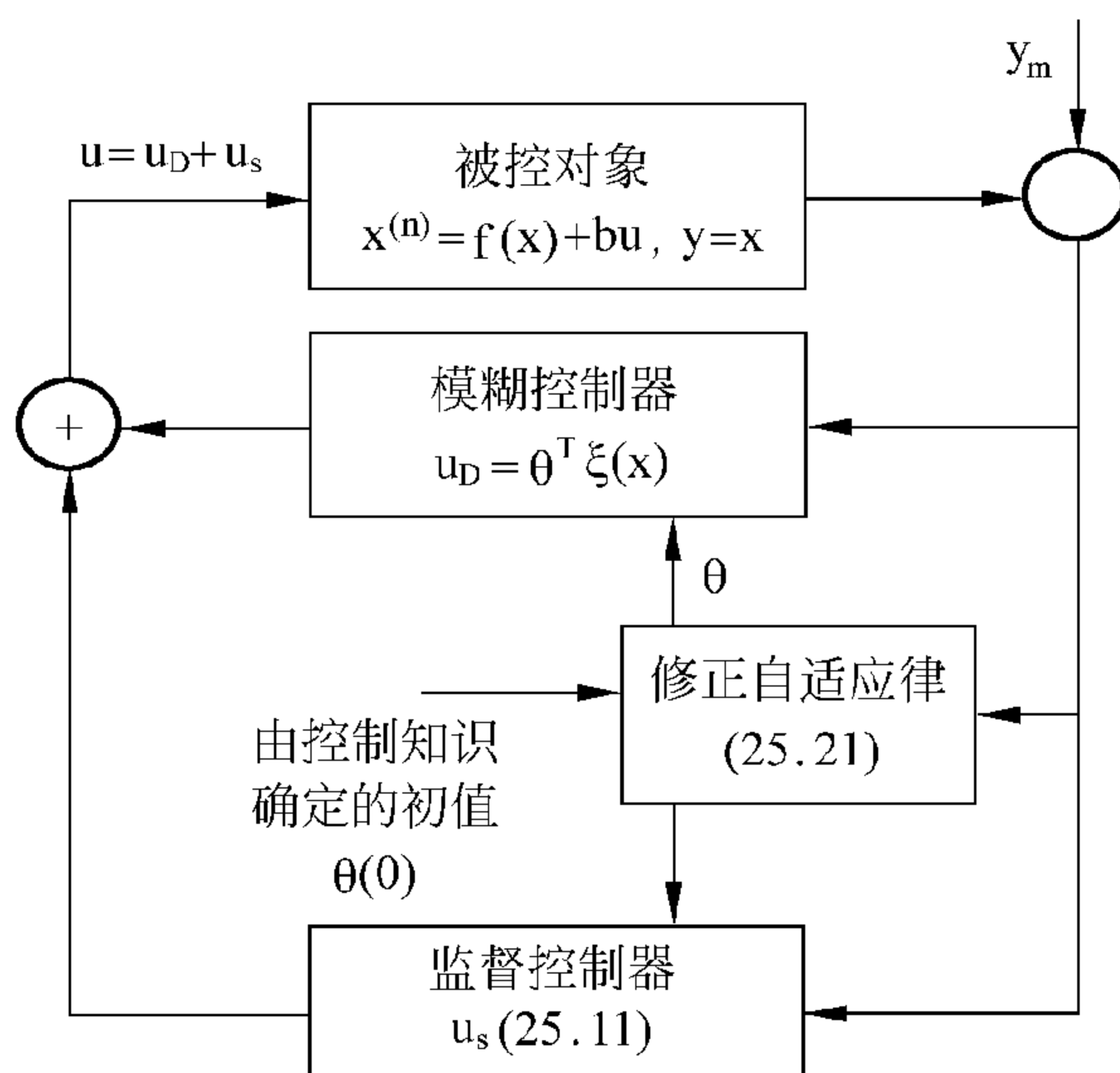


图 25.2 具有监督控制器和修正自适应律的高级直接型自适应模糊控制系统

(c) 如果 w 平方可积, 即如果 $\int_0^\infty |w(t)|^2 dt < \infty$, 则 $\lim_{t \rightarrow \infty} |e(t)| = 0$ 。

证明: (a) 式(25.23)可以由定理 25.2 得到, 式(25.24)可以由式(25.12)(它表明 $V_e \leq \bar{V}$)和式(25.4)(它表明 $V_e \leq \bar{V}$ 等价于 $|x(t)| \leq M_x$)得到。

(b) 添加监督控制器 $u_s(x|\theta)$ 后, 误差方程(24.12)变为

$$\dot{\mathbf{e}} = \mathbf{A}\mathbf{e} + \mathbf{b}(\theta^* - \theta)^T \xi(\mathbf{x}) - \mathbf{b}w - \mathbf{b}u_s(\mathbf{x}) \quad (25.26)$$

考虑候选 Lyapunov 函数(24.13)中的 V 并由式(25.26), 可得

$$\dot{V} = -\frac{1}{2}\mathbf{e}^T Q \mathbf{e} + \frac{b}{\gamma}(\theta^* - \theta)^T [\gamma \mathbf{e}^T p_n \xi(\mathbf{x}) - \dot{\theta}] - \mathbf{e}^T p_n b w - \mathbf{e}^T P \mathbf{b} u_s(\mathbf{x}) \quad (25.27)$$

将自适应律(25.21)代入式(25.27), 并令 $I_\theta^* = 1$, 则可知式(25.21)中的第二行条件得到了满足。再令 $I_\theta^* = 0$ 表示其他情况, 可得

$$\dot{V} = -\frac{1}{2}\mathbf{e}^T Q \mathbf{e} + I_\theta^* \mathbf{e}^T p_n b \frac{(\theta^* - \theta)^T \theta \theta^T \xi(\mathbf{x})}{|\theta|^2} - \mathbf{e}^T p_n b w - \mathbf{e}^T P \mathbf{b} u_s(\mathbf{x}) \quad (25.28)$$

下面来证明式(25.28)中右端的第二项为负。如果 $I_\theta^* = 0$, 结论显然成立。令 $I_\theta^* = 1$, 即表明 $|\theta| = M_D$ 且 $\mathbf{e}^T p_n \theta^T \xi(\mathbf{x}) < 0$, 因为 $|\theta| = M_D \geq |\theta^*|$, 所以有 $(\theta^* - \theta)^T \theta = \frac{1}{2}(|\theta^*|^2 - |\theta|^2 - |\theta - \theta^*|^2) < 0$ 。由此可将式(25.28)简化为

$$\dot{V} \leq -\frac{1}{2}\mathbf{e}^T Q \mathbf{e} - \mathbf{e}^T p_n b w - \mathbf{e}^T P \mathbf{b} u_s(\mathbf{x}) \quad (25.29)$$

因为式(25.11)中的监督控制器 $u_s(\mathbf{x})$ 和 $\mathbf{e}^T P \mathbf{b}$ 有相同的符号, 所以由式(25.29)可得

$$\begin{aligned} \dot{V} &\leq -\frac{1}{2}\mathbf{e}^T Q \mathbf{e} - \mathbf{e}^T p_n b w \\ &\leq -\frac{\lambda_{Q\min} - 1}{2} |\mathbf{e}|^2 - \frac{1}{2} (|\mathbf{e}|^2 - 2\mathbf{e}^T p_n b w + |p_n b w|^2) + \frac{1}{2} |p_n b w|^2 \\ &\leq -\frac{\lambda_{Q\min} - 1}{2} |\mathbf{e}|^2 + \frac{1}{2} |p_n b w|^2 \end{aligned} \quad (25.30)$$

式中, $\lambda_{Q\min}$ 是 Q 的最小特征值且假设 $\lambda_{Q\min} > 1$ 。对式(25.30)两边同时积分, 可得

$$\int_0^t |\mathbf{e}(\tau)|^2 d\tau \leq \frac{2}{\lambda_{Q\min} - 1} (|V(0)| + V(t)) + \frac{1}{\lambda_{Q\min} - 1} |p_n b|^2 \int_0^t |w(\tau)|^2 d\tau \quad (25.31)$$

定义 $a = \frac{2}{\lambda_{Q\min} - 1} (|V(0)| + \sup_{t \geq 0} |V(t)|)$, $b = \frac{1}{\lambda_{Q\min} - 1} |p_n b|^2$, 则式(25.31)就变成了式(25.25)。

(c) 若 $w \in L_2$, 则由式(25.25), 可得 $\mathbf{e} \in L_2$ 。由于式(24.26)右端所有的变量都有界, 所以有 $\dot{\mathbf{e}} \in L_\infty$ 。再由 Barbalat 的引理(Sastry 和 Bodson[1989]: 如果 $\mathbf{e} \in L_2 \cap L_\infty$ 且 $\dot{\mathbf{e}} \in L_\infty$, 则有 $\lim_{t \rightarrow \infty} |\mathbf{e}(t)| = 0$), 可得 $\lim_{t \rightarrow \infty} |\mathbf{e}(t)| = 0$ 。

25.3.2 仿真

例 25.1 考虑和例 24.1 相同的情况, 例外之处在于, 本例系统中使用了式(25.11)所示的监督控制器 u_s 和修正自适应律(25.21)。这里仍然使用如图 24.2 所示的六个模糊集, 并且在 $[-2, 2]$ 上随机选择初值 $\theta_i(0)$ 。差别在于, 这里的目标是要求 $|x| \leq M_x = 1.5$ 且 $|\theta| \leq M_D = 10$ 。当初值为 $x(0) = 1$ 的仿真结果见图 25.3。从图 25.3 可以看到, 监督控制器确实可将状态限定在约束集内。

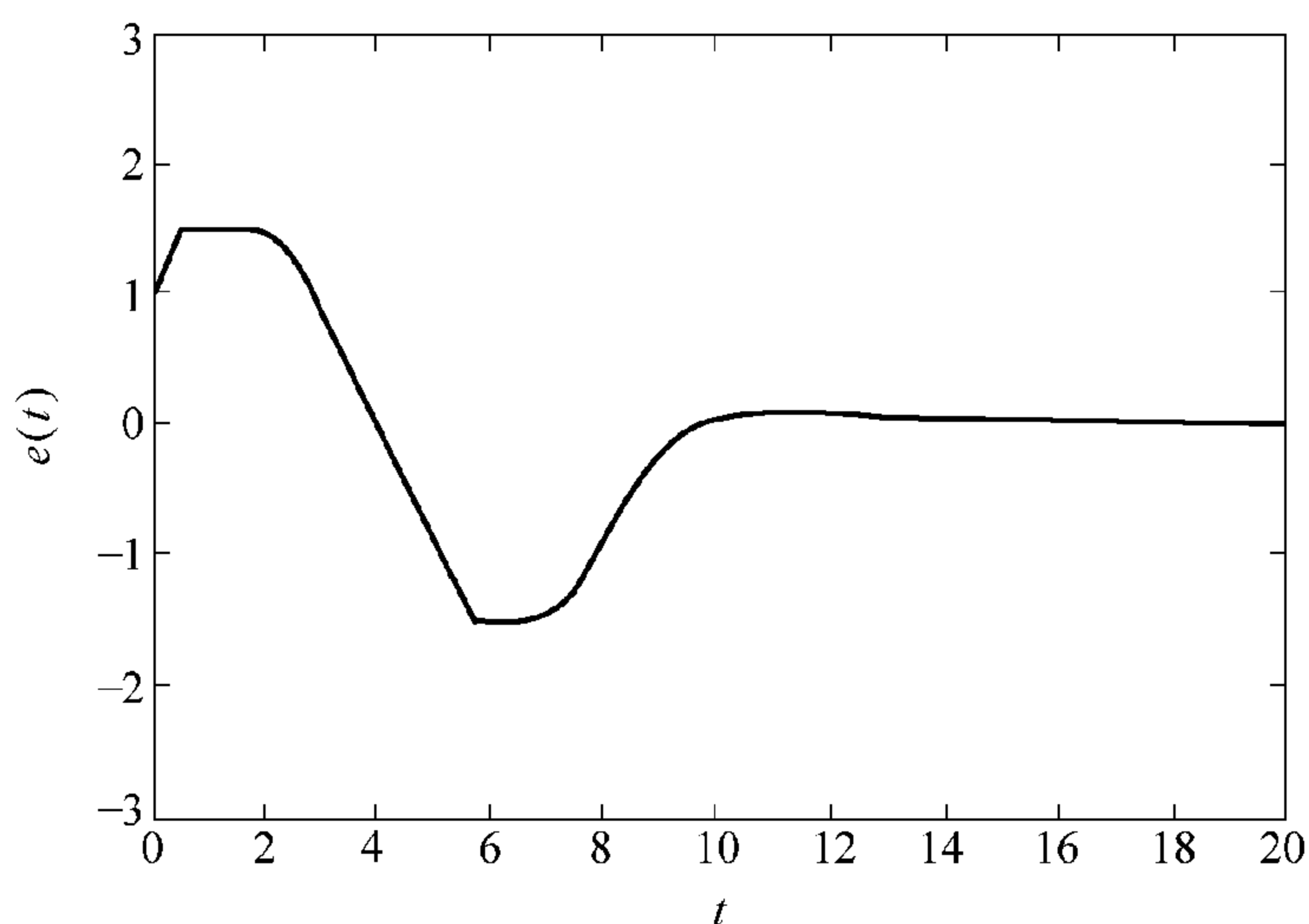


图 25.3 例 25.1 中闭环系统的状态 $x(t) = e(t)$

例 25.2 考虑和例 24.2 相同的情况, 例外之处在于, 本例系统中附加了监督控制器和修正自适应律。除要求 $|x| \leq M_x = 3$ 外, 其他条件同图 24.6 中的一样。图 25.4 给出了本例的仿真结果。从图 25.4 可以再次看到, 监督控制器确实起作用。

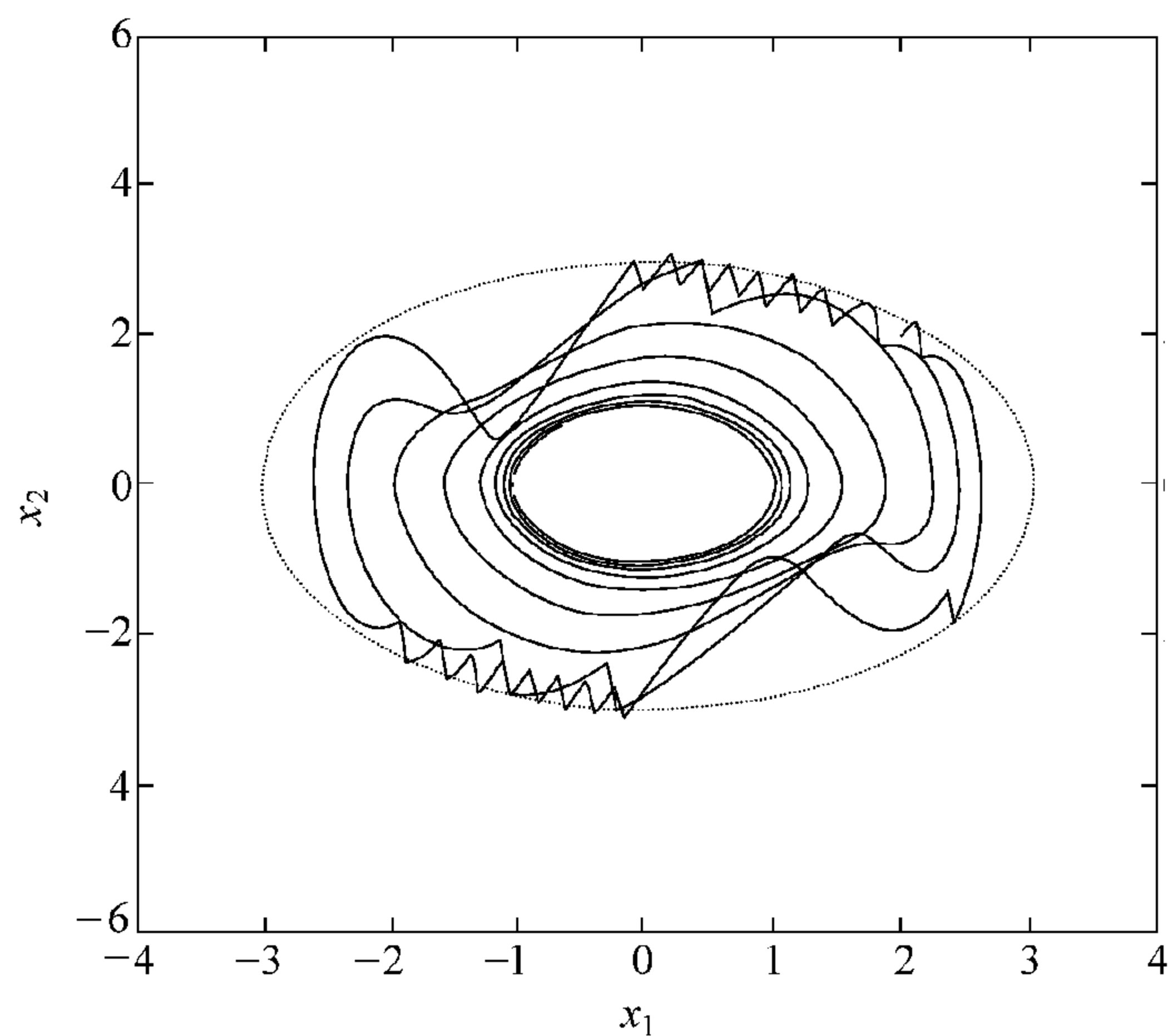


图 25.4 例 25.2 中闭环系统在相平面 $x_1 - x_2$ 上的轨迹

25.4 本章小结及导读

本章主要阐述了以下内容：

- 如何设计监督控制器以保证系统状态有界；
- 如何使用投影法来修正基本的自适应律以保证系统参数有界；
- 稳定的直接型自适应模糊控制方案及其特性。

投影法见 Luenberger[1969]和 Luenberger[1984]。还存在其他的保证参数有界方法，如死区法(Narendra 和 Annaswamy[1989], Sastry 和 Bodson[1989], δ -修正方法(Ioannou 和 Kokotovic[1983])，以及 ε -修正方法(Narendra 和 Annaswamy[1989])。本章中介绍的稳定的直接型自适应模糊控制器摘自 Wang[1993]。

25.5 习题

习题 25.1 为 24.2 节所介绍的组合型自适应模糊控制系统设计一个监督控制器，使得闭环系统的状态有界。证明所设计的监督控制器能够满足目标要求。

习题 25.2 用投影法修正 24.2 节所介绍的组合型自适应模糊控制系统的自适应律，使闭环系统的参数有界。证明所得到的自适应律能够满足目标要求。

习题 25.3 证明式(23.27)确实是一个投影算子。即，证明

(a) $P\{-\gamma_1 \mathbf{e}^T P \mathbf{b} \xi(\mathbf{x})\}$ 垂直于 $P\{-\gamma_1 \mathbf{e}^T P \mathbf{b} \xi(\mathbf{x})\} - (-\gamma_1 \mathbf{e}^T P \mathbf{b} \xi(\mathbf{x}))$ ，且

(b) $P\{-\gamma_1 \mathbf{e}^T P \mathbf{b} \xi(\mathbf{x})\}$ 在 θ_f 点处的约束集 Ω_f 的超平面上。

习题 25.4 证明定理 25.2。

习题 25.5 假定系统初始值 $x(0) = 1.4$ ，重做例 25.1 的仿真。

习题 25.6 假定 $M_x = 4$ ，重做例 25.2 中的仿真。

26.1 稳定的间接型自适应模糊控制系统

26.1.1 稳定性和收敛性分析

接第 25 章的思路,为如图 23.7 所示基本间接型自适应模糊控制系统 u_l 添加式(25.6)所示的监督控制器 $u_s(\mathbf{x})$,并将自适应律改为式(25.15)~式(25.19),这样可得到如图 26.1 所示的高级间接型自适应模糊控制系统。下面的定理将说明这个系统的稳定性和收敛性。

定理 26.1 考虑如图 26.1 所示的高级间接型自适应模糊控制系统,即被控对象为式(23.1)~式(23.2),控制为 $u = u_l + u_s$ 。其中 u_l 由式(23.7)给出, u_s 由式(25.6)给出,自适应律由式(25.15)~式(25.19)给出。这个自适应模糊控制系统必有如下特性:

(a) 所有的参数和状态都在有界集内,即

$$\theta_f \in \Omega_f, \theta_g \in \Omega_g \quad (26.1)$$

$$\|\mathbf{x}\| \leq M_x \quad (26.2)$$

其中, Ω_f 和 Ω_g 的定义分别由式(25.13)和式(25.14)给出, u_s 中 I^* 中的 \bar{V} 根据式(25.4)确定。

(b) 跟踪误差 e 满足式(25.25),其中 w 由式(23.21)定义, a 和 b 是常数。

(c) 若 w 平方可积,则 $\lim_{t \rightarrow \infty} |e(t)| = 0$ 。

本定理的证明和定理 25.3 证明思路一样,留作习题。

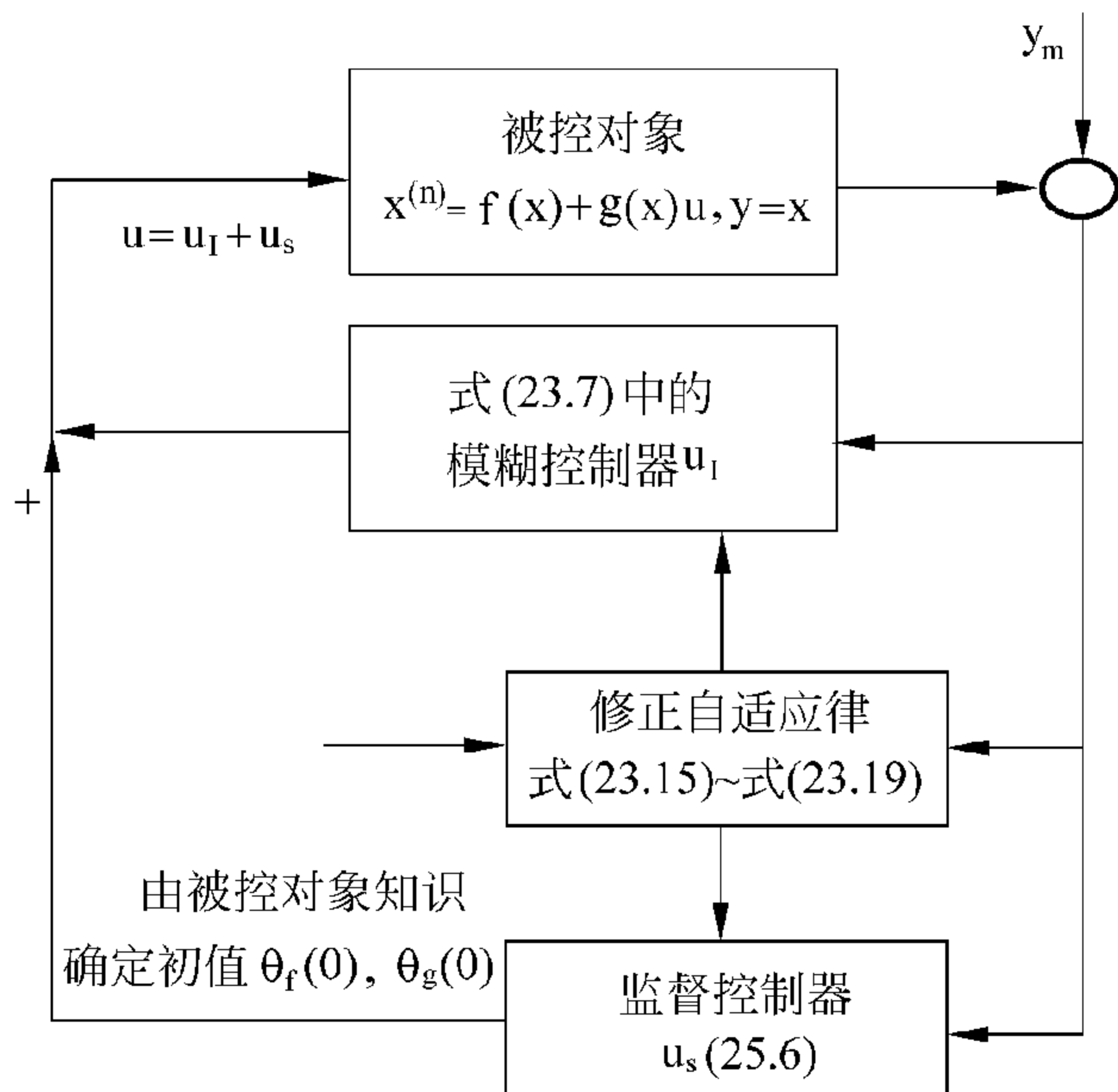


图 26.1 具有监督控制器和修正自适应律的高级间接型自适应模糊控制系统

26.1.2 非线性参数化

在图 23.2 和图 26.1 所示的间接型自适应模糊控制系统的模糊控制器中,模糊系统 $\hat{f}(\mathbf{x}|\theta_f)$ 和 $\hat{g}(\mathbf{x}|\theta_g)$ 是线性参数化的,其表达式是式(23.12)~式(23.13)。要想实现这些参数的线性化,IF-THEN 规则中 IF 部分的隶属度函数必须固定不变,而且自适应律只能调整 THEN 部分的模糊集的中心 θ_f 和 θ_g 。为了使模糊系统 $\hat{f}(\mathbf{x}|\theta_f)$ 和 $\hat{g}(\mathbf{x}|\theta_g)$ 在 \mathbf{x} 的任何取值区域内都具有良好的行为,IF 部分的隶属度函数应该覆盖整个区域。因此,模糊系统 $\hat{f}(\mathbf{x}|\theta_f)$ 和 $\hat{g}(\mathbf{x}|\theta_g)$ 中模糊规则的数目会随 \mathbf{x} 维数的增加而指数性地上升。一种减少模糊规则数目的方法是允许 IF 部分的隶属度函数在自适应过程中发生变化,使得同样的规则能够在不同时刻用于整个状态空间的不同区域。

具体地,选择 $\hat{f}(\mathbf{x}|\theta_f)$ 和 $\hat{g}(\mathbf{x}|\theta_g)$ 为形如式(9.6)的模糊系统(带有乘积推理机、单值模糊器、中心平均解模糊器和高斯隶属度函数的模糊系统),即

$$\hat{f}(\mathbf{x}|\theta_f) = \frac{\sum_{l=1}^M \bar{y}_f^l \left[\prod_{i=1}^n \exp\left(-\left(\frac{x_i - \bar{x}_{f^i}^l}{\sigma_{f^i}}\right)^2\right)\right]}{\sum_{l=1}^M \left[\prod_{i=1}^n \exp\left(-\left(\frac{x_i - \bar{x}_{f^i}^l}{\sigma_{f^i}}\right)^2\right)\right]} \quad (26.3)$$

$$\hat{g}(\mathbf{x}|\theta_g) = \frac{\sum_{l=1}^M \bar{y}_g^l \left[\prod_{i=1}^n \exp\left(-\left(\frac{x_i - \bar{x}_{g^i}^l}{\sigma_{g^i}}\right)^2\right)\right]}{\sum_{l=1}^M \left[\prod_{i=1}^n \exp\left(-\left(\frac{x_i - \bar{x}_{g^i}^l}{\sigma_{g^i}}\right)^2\right)\right]} \quad (26.4)$$

式中, θ_f 和 θ_g 分别为参数 $(\bar{y}_f^l, \bar{x}_{f^i}^l, \sigma_{f^i}^l)$ 和 $(\bar{y}_g^l, \bar{x}_{g^i}^l, \sigma_{g^i}^l)$ 的集合。令最优参数 θ_f^* 和 θ_g^* 由式(23.19)~式(23.20)定义。为了使用与 23.2.3 小节相同的方法设计自适应律,现将 $\hat{f}(\mathbf{x}|\theta_f^*)$ 和 $\hat{g}(\mathbf{x}|\theta_g^*)$ 在 θ_f 和 θ_g 处进行泰勒展开,可得

$$\hat{f}(\mathbf{x}|\theta_f) - \hat{f}(\mathbf{x}|\theta_f^*) = (\theta_f - \theta_f^*)^T \left[\frac{\partial \hat{f}(\mathbf{x}|\theta_f)}{\partial \theta_f} \right] + O(|\theta_f - \theta_f^*|^2) \quad (26.5)$$

$$\hat{g}(\mathbf{x}|\theta_g) - \hat{g}(\mathbf{x}|\theta_g^*) = (\theta_g - \theta_g^*)^T \left[\frac{\partial \hat{g}(\mathbf{x}|\theta_g)}{\partial \theta_g} \right] + O(|\theta_g - \theta_g^*|^2) \quad (26.6)$$

式中, $O(|\theta_f - \theta_f^*|^2)$ 和 $O(|\theta_g - \theta_g^*|^2)$ 表示高阶项。

将式(26.5)和式(26.6)代入到式(23.22)中,可以得到误差方程

$$\dot{\mathbf{e}} = \mathbf{A}\mathbf{e} + \mathbf{b} \left\{ (\theta_f - \theta_f^*)^T \frac{\partial \hat{f}(\mathbf{x}|\theta_f)}{\partial \theta_f} + (\theta_g - \theta_g^*)^T \frac{\partial \hat{g}(\mathbf{x}|\theta_g)}{\partial \theta_g} u_I + v \right\} \quad (26.7)$$

其中

$$v = w + O(|\theta_f - \theta_f^*|^2) + O(|\theta_g - \theta_g^*|^2) \quad (26.8)$$

w 的定义是式(23.21)。比较式(26.7)和式(23.23)易知,可以使用与 23.2.3 小节相同的方法来设计自适应律,不过要将 $\xi(\mathbf{x})$ 用 $\frac{\partial \hat{f}(\mathbf{x}|\theta_f)}{\partial \theta_f}$ 来代替, $\eta(\mathbf{x})$ 用 $\frac{\partial \hat{g}(\mathbf{x}|\theta_g)}{\partial \theta_g}$ 来代替, w 用 v

来代替。即,这种情况下的自适应律为

$$\dot{\theta}_f = -\gamma_1 \mathbf{e}^T P \mathbf{b} \left[\frac{\partial \hat{f}(\mathbf{x}|\theta_f)}{\partial \theta_f} \right] \quad (26.9)$$

$$\dot{\theta}_g = -\gamma_2 \mathbf{e}^T P \mathbf{b} \left[\frac{\partial \hat{g}(\mathbf{x}|\theta_g)}{\partial \theta_g} \right] u_I \quad (26.10)$$

在线性参数化的情况下可以用同样的方法设计带有投影算子的(高级)自适应律。

要想应用自适应律(26.9)和(26.10),必须明确如何计算 $\frac{\partial \hat{f}(\mathbf{x}|\theta_f)}{\partial \theta_f}$, $\frac{\partial \hat{g}(\mathbf{x}|\theta_g)}{\partial \theta_g}$ 。从第

13 章可知,这两个偏导数可以用反向传播算法计算的。具体地讲, $\frac{\partial \hat{f}(\mathbf{x}|\theta_f)}{\partial \theta_f}$ 可以根据下面

三个式子计算

$$\frac{\partial \hat{f}}{\partial \bar{y}_f^l} = \frac{z_f^l}{b_f} \quad (26.11)$$

$$\frac{\partial \hat{f}}{\partial \bar{x}_{f^i}^l} = \frac{\bar{y}_f^l - \hat{f}}{b_f} z_f^l \frac{2(x_i - \bar{x}_{f^i}^l)}{(\sigma_{f^i}^l)^2} \quad (26.12)$$

$$\frac{\partial \hat{f}}{\partial \sigma_{f^i}^l} = \frac{\bar{y}_f^l - \hat{f}}{b_f} z_f^l \frac{2(x_i - \bar{x}_{f^i}^l)^3}{(\sigma_{f^i}^l)^3} \quad (26.13)$$

其中

$$z_f^l = \prod_{i=1}^n \exp \left(- \left(\frac{x_i - \bar{x}_{f^i}^l}{\sigma_{f^i}^l} \right)^2 \right) \quad (26.14)$$

$$b_f = \sum_{l=1}^M z_f^l \quad (26.15)$$

类似地, $\frac{\partial \hat{g}(\mathbf{x}|\theta_g)}{\partial \theta_g}$ 也可以用与式(26.11) ~ 式(26.15)相同的算法来计算, 只不过用 g 来代替 f 。

26.2 一般非线性系统的自适应模糊控制

前面介绍的多种自适应模糊控制器是为式(23.1) ~ 式(23.2)这样的标准型非线性系统设计的。然而在实际中许多非线性系统不具有这样的标准型。一般地讲, 一个单输入—单输出的连续系统可表示为如下形式:

$$\mathbf{x} = F(\mathbf{x}, u), y = h(\mathbf{x}) \quad (26.16)$$

式中, $\mathbf{x} \in R^n$ 为状态向量, $u \in R$ 和 $y \in R$ 分别是系统的输入和输出, F 和 h 为非线性函数。本节将考虑如式(26.16)所示的一般非线性系统; 控制的任务是使得系统输出 $y(t)$ 能够跟踪一期望轨迹 $y_m(t)$ 。

若将式(26.16)和标准型(23.1) ~ (23.2)进行比较会发现, 处理这种一般模型的难点在于, 输出 y 只是通过状态变量 \mathbf{x} 和非线性状态方程间接地与系统的输入 u 相关; 而另一方面, 标准型的输出(等于第一状态变量)却直接地与输入相关。因此, 受第23章至第25章思路的启发, 可以猜想, 若能够在系统的输出 y 和控制输入 u 之间找到某种直接的关系, 则可以减少对式(26.16)这个一般非线性系统的跟踪控制难度。事实上, 正是这种想法构成了建立在直觉基础上的所谓的输入—输出线性化方法(Isidori[1989])。

输入—输出线性化是非线性控制器设计的一种方法, 这种方法近年来在非线形控制领域引起了广泛的兴趣。该方法完全不同于传统的线性化方法(如雅可比线性化方法), 它是通过精确的状态变换和反馈来实现线性化的, 而不是通过非线性动态方程的线性逼近。在26.2.1节中, 我们将简要给出这种线性化方法的基本概念, 在26.2.2节中, 讨论如何根据输入—输出线性化方法为式(26.16)所描述的一般非线性系统设计自适应模糊控制器, 最后在26.2.3节中, 研究所设计的自适应模糊控制器在球—横木(ball-and-beam)系统中的应用。

26.2.1 输入—输出线性化的直观概念

输入—输出线性化的基本思想可以概括如下: 对系统输出 y 反复进行求导, 直到系统输入 u 出现为止, 然后选取特殊的 u , 使非线性消失, 最后根据线性控制来设计控制器。下面举例说明这一基本思想。

考虑如下三阶系统

$$\dot{x}_1 = \sin(x_2) + (x_2 + 1)x_3 \quad (26.17)$$

$$\dot{x}_2 = x_1^5 + x_3 \quad (26.18)$$

$$\dot{x}_3 = x_1^2 + u \quad (26.19)$$

$$y = x_1 \quad (26.20)$$

为了在 y 和 u 之间建立直接关系, 需要对 y 求导

$$\dot{y} = \dot{x}_1 = \sin(x_2) + (x_2 + 1)x_3 \quad (26.21)$$

因为 y 和 u 仍没有直接关系, 所以再对 y 求导

$$\ddot{y} = (\cos x_2 + x_3)(x_1^5 + x_3) + (x_2 + 1)x_1^2 + (x_2 + 1)u \quad (26.22)$$

显然, 式(26.22)表明了 y 和 u 之间有直接关系。如果选

$$u = \frac{1}{x_2 + 1}(v - f_1) \quad (26.23)$$

式中, $f_1 \equiv (\cos x_2 + x_3)(x_1^5 + x_3) + (x_2 + 1)x_1^2$, 则可得

$$\ddot{y} = v \quad (26.24)$$

如果把 v 视为一个新的输入, 则原非线性系统(26.17) ~ (26.20) 就被线性化为线性系统(26.24)了。这一线性化过程如图 26.2 所示。

现在若把新的控制输入选为

$$v = \ddot{y}_m + k_1 \dot{e} + k_2 e \quad (26.25)$$

式中, $e \equiv y_m - y$, 则闭环系统可以表征为

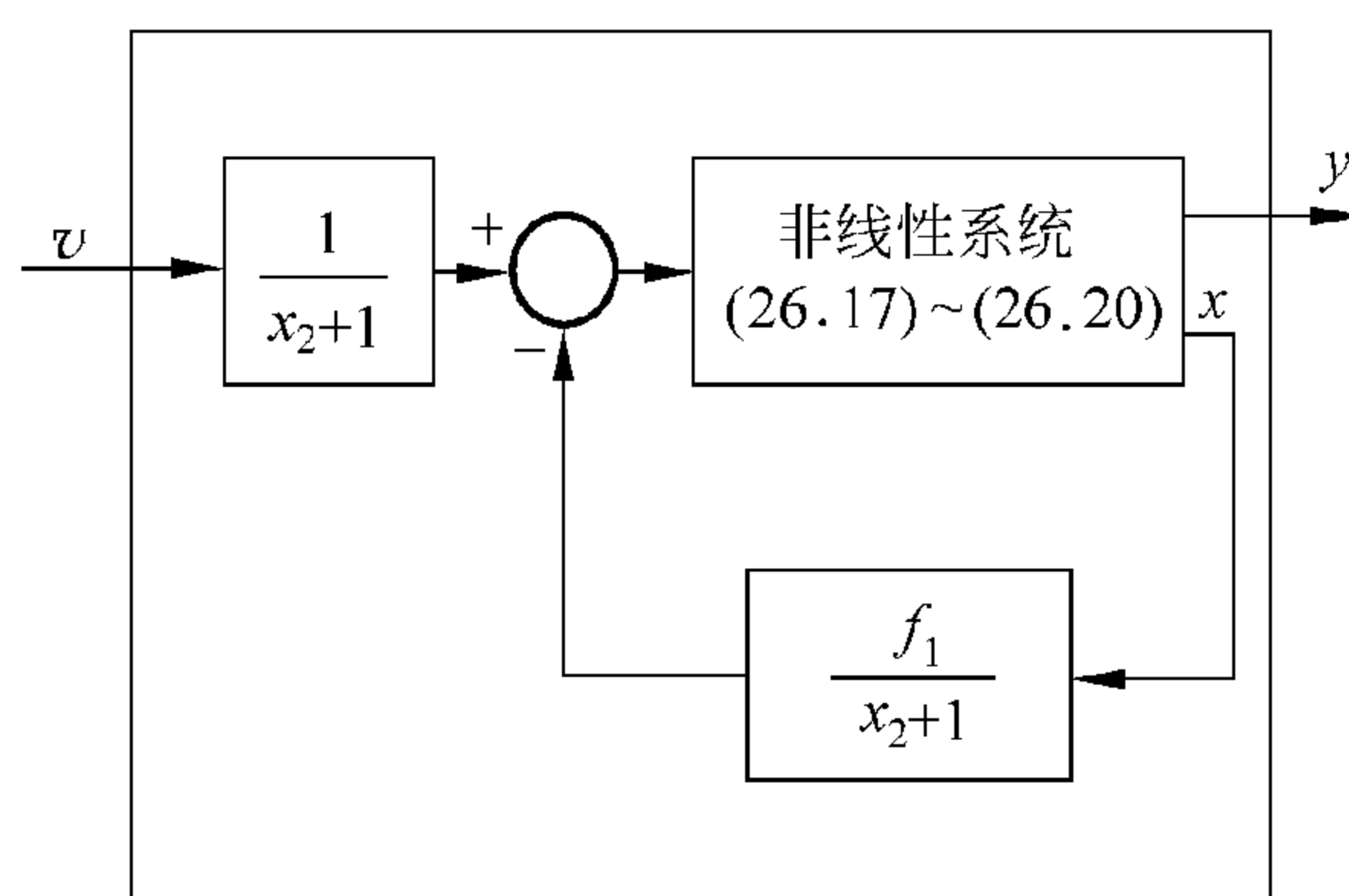
$$\ddot{e} + k_1 \dot{e} + k_2 e = 0 \quad (26.26)$$

若选择 k_1 和 k_2 使得 $s^2 + k_1 s + k_2 = 0$ 的所有根均在复平面的左半开平面上, 则有 $\lim_{t \rightarrow \infty} e(t) = 0$, 此即控制的目标。

若需要对某一个系统的输出求 r 次导数才能在输出 y 和输入 u 之间建立直接关系, 则称这个系统具有 r 阶相对阶数。由此可知, 系统(26.17) ~ (26.20) 的相对阶数为 2。可以严格证明, 对任何 n 阶可控系统的任意输出, 使控制输入出现所需的求导次数最多为 n 。也就是说, 任意 n 阶可控系统的相关阶数一定小于等于 n 。

至此, 也许有人会觉得, 非线性系统(26.17) ~ (26.20) 的跟踪控制问题用控制法则(26.23)和(26.25)就可以解决了。然而必须明确的是, 式(26.26)只考虑了整个系统的局部, 整个系统是三阶系统, 而式(26.26)只是二阶系统。因此, 输入—输出线性化使系统的一部分动态具有不可观测性。因为系统的这一部分不能从外部的输入—输出关系(26.24)中被观测到, 所以称系统的这一部分为“内部动态”(internal dynamics)。对上面的例子, 其内部动态可表示为

$$\dot{x}_3 = x_1^2 + \frac{1}{x_2 + 1}(\ddot{y}_m + k_1 \dot{e} + k_2 e - f_1) \quad (26.27)$$



线性化系统

图 26.2 非线性系统(26.17) ~ (26.20) 的线性化系统示意图

上式是将式(26.23)和式(26.25)代入式(26.19)得到的。若这一内部动态是稳定的,即状态 x_3 有界,则我们的控制任务也就完成了。否则就必须重新设计控制律。

26.2.2 基于输入—输出线性化的自适应模糊控制器的设计

从上一小节可以看出,输入—输出线性化方法要求给定系统的数学模型,否则求导过程就不能进行。而在此处的问题中,假设式(26.16)中的非线性函数 F 和 h 是未知的。那么,输入—输出线性化怎样推广到本问题中呢?

首先,受第23章的确定等效控制器的启发,考虑用模糊系统来代替 F 和 h 。然后再设计一个自适应律来调节模糊系统的参数使之逼近 F 和 h 。这确实是个好方法。然而这种方法生成的自适应控制系统非常复杂。其原因在于,虽然逼近 F 和 h 的原模糊系统中的参数是线性的,但求导会使这些参数在求导过程的最后阶段呈非线性。因此,这里不采用这种方法。

这里采用的方法主要基于如下考虑:从上小节中可以看出,控制设计(26.23)只建立在求导过程最后所得的系统(26.22)上,中间过渡系统(26.21)并没有被直接用到。因此,可以不用模糊系统来逼近 F 和 h ,而是用其来逼近求导过程最后所得到的方程中的非线性函数。然后,设计一个自适应律来调节模糊系统的参数,使得 y 能够跟踪 y_m 。要正规地给出这种方法的详细步骤,还需作如下假设。

假设 26.1 假设:(i)非线性系统(26.16)的相对阶数为 r ; (ii)控制 u 与 $y^{(r)}$ 之间呈线性关系,即

$$y^{(r)} = f(\mathbf{x}) + g(\mathbf{x})u \quad (26.28)$$

式中, f 和 g 为未知函数且 $g(\mathbf{x}) \neq 0$; (iii)在下面的自适应模糊控制器的作用下,系统的内部动态是稳定的。

基于输入—输出线性化的自适应模糊控制器设计:

步骤 1 根据物理常识求出非线性系统(26.16)的相对阶数 r 。具体地讲,可先分析 $y, \dot{y}, \ddot{y}, \dots$ 的物理意义,然后确定与 u 有直接关联的 $y^{(r)}$ 。在下一小节中将以球—横本系统为例说明如何完成这一步骤。

步骤 2 选择形如式(9.6)的模糊系统 $\hat{f}(\mathbf{x}|\theta_f)$ 和 $\hat{g}(\mathbf{x}|\theta_g)$, 即

$$\hat{f}(\mathbf{x}|\theta_f) = \theta_f^T \xi(\mathbf{x}) \quad (26.29)$$

$$\hat{g}(\mathbf{x}|\theta_g) = \theta_g^T \xi(\mathbf{x}) \quad (26.30)$$

步骤 3 设计如下控制器

$$u = \frac{1}{\hat{g}(\mathbf{x}|\theta_g)} [-\hat{f}(\mathbf{x}|\theta_f) + y_m^{(r)} + \mathbf{k}^T \mathbf{e}] \quad (26.31)$$

式中, $\mathbf{k} = (k_r, \dots, k_1)$ 是使 $s^r + k_1 s^{r-1} + \dots + k_r = 0$ 的所有根均在复平面左半开平面上的系数

向量, $\mathbf{e} = (e, \dot{e}, \dots, e^{(r-1)})^T$, 其中 $e = y_m - y$ 。

步骤 4 采用如下自适应律来调节参数

$$\dot{\theta}_f = -\gamma_1 \mathbf{e}^T P \mathbf{b} \xi(\mathbf{x}) \quad (26.32)$$

$$\dot{\theta}_g = -\gamma_2 \mathbf{e}^T P \mathbf{b} \eta(\mathbf{x}) u \quad (26.33)$$

式中, p 和 \mathbf{b} 的定义见第 23 章, 除被控对象改为式(26.16)和 n 改为 r 以外, 整个控制系统与图 23.2 所示的控制系统完全相同。

控制器(26.31)和自适应律(26.32)和(26.33)均是采用第 23 章中的 Lyapunov 综合法得到的。采用此方法是因为, 假设 26.1 保证了式(26.16)与式(26.28)等价。

采用与第 25 章相同的方法, 可以把监督控制器添加到模糊控制器(26.31)中以保证 $\mathbf{y} = (y, \dot{y}, \dots, y^{(r-1)})^T$ 有界。也可以采用投影法来修正自适应律(26.32)和(26.33), 以保证 θ_f 和 θ_g 有界。如果存在关于 f 和 g 的语言信息, 则可以用与第 23 章相同的方法, 在给定初始值 $\theta_f(0)$ 和 $\theta_g(0)$ 时嵌入这些语言信息。

以上的自适应模糊控制器属于间接型自适应模糊控制器, 可以用与第 24 章相同的方法来设计。生成的控制器为

$$u = \theta^T \xi(\mathbf{x}) \quad (26.34)$$

相应的自适应律为

$$\dot{\theta} = \gamma \mathbf{e}^T \mathbf{p}_r \xi(\mathbf{x}) \quad (26.35)$$

式中, \mathbf{p}_r 的定义见第 24 章, 类似地, 也可以为一般非线性系统开发一个组合型自适应模糊控制器。

26.2.3 在球—横木系统中的应用

球—横木系统见图 16.7, 该系统可用式(16.4)~式(16.5)来描述。在本小节中, 将采用上小节中的自适应模糊控制器来控制球的位置 $y = x_1$ 以跟踪轨迹 $y_m(t) = \sin(t)$ 。

步骤 1 根据物理常识确定球—横木系统的相对阶数 r 。首先, 已经知道控制 u 等于横木的角 θ 的加速度; 于是, 任务变为, 确定 y 的哪一阶导数与 θ 的加速度直接相关。显然, 球的位置 $y = x_1 = r$ 和球沿横木方向的速度 \dot{y} 均与 u 无直接关系。根据牛顿定律, 球的位置变化的加速度 \ddot{y} 与 $\sin(\theta)$ 成正比, 与 $u = \ddot{\theta}$ 无直接关系。因此, \ddot{y} 与 u 不直接相关。由于 \ddot{y} 与 $\sin(\theta)$ 成正比, 所以 $y^{(3)}$ 与 $\dot{\theta}$ 直接相关, 而与 $u = \ddot{\theta}$ 不直接相关。最后, 我们发现, $y^{(4)}$ 与 u 直接相关。由此可得球—横木系统的相对阶数为 4。^①

^① 在 26.2.1 小节中没有给出相对阶数在数学上的严格定义。因此, 用这种方法确定的相对阶数仅是直观意义上的相对阶数。关于相对阶数的严格定义见 Isidori[1989], Slotine 和 Li[1991]。

步骤2 为由 x_1 到 x_4 中的每一个变量定义三个模糊集合 $\mu_N(x_i) = 1/(1 + \exp(5(x_i + 1)))$, $\mu_Z(x_i) = \exp(-x_i^2)$ 和 $\mu_P(x_i) = 1/(1 + \exp(-5(x_i - 1)))$, 这样 θ_f, θ_g 和 $\xi(\mathbf{x})$ 的维数就等于 $3^4 = 81$ 。由于没有任何关于 f 和 g 的语言信息, 所以初始参数 $\theta_f(0)$ 和 $\theta_g(0)$ 是在区间 $[-2, 2]$ 上随机选取的。

步骤3 选定 $\mathbf{k} = (1, 4, 6, 4)^T$ 。

步骤4 选定 $\gamma_1 = 2, \gamma_2 = 0.2$ 。图 26.3 和图 26.4 分别给出了在 $\mathbf{x}(0) = (1, 0, 0, 0)^T$ 和 $\mathbf{x}(0) = (-1, 0, 0, 0)^T$ 这两种初始条件下, 采用本自适应模糊控制器所得到的输出 $y(t)$ (实线) 和期望轨迹 $y_m(t)$ (虚线)。

另外, 这里还对直接型自适应模糊控制器(26.34)和(26.35)进行了仿真。其中, $\xi(\mathbf{x})$ 的选取与上述间接型自适应模糊控制器中的相同。这里选 $\gamma = 10$, 图 26.5 和图 26.6 分别给出了在 $\mathbf{x}(0) = (1, 0, 0, 0)^T$ 和 $\mathbf{x}(0) = (-0.4, 0, 0, 0)^T$ 这两种初始条件下, 采用此直接型自适应模糊控制

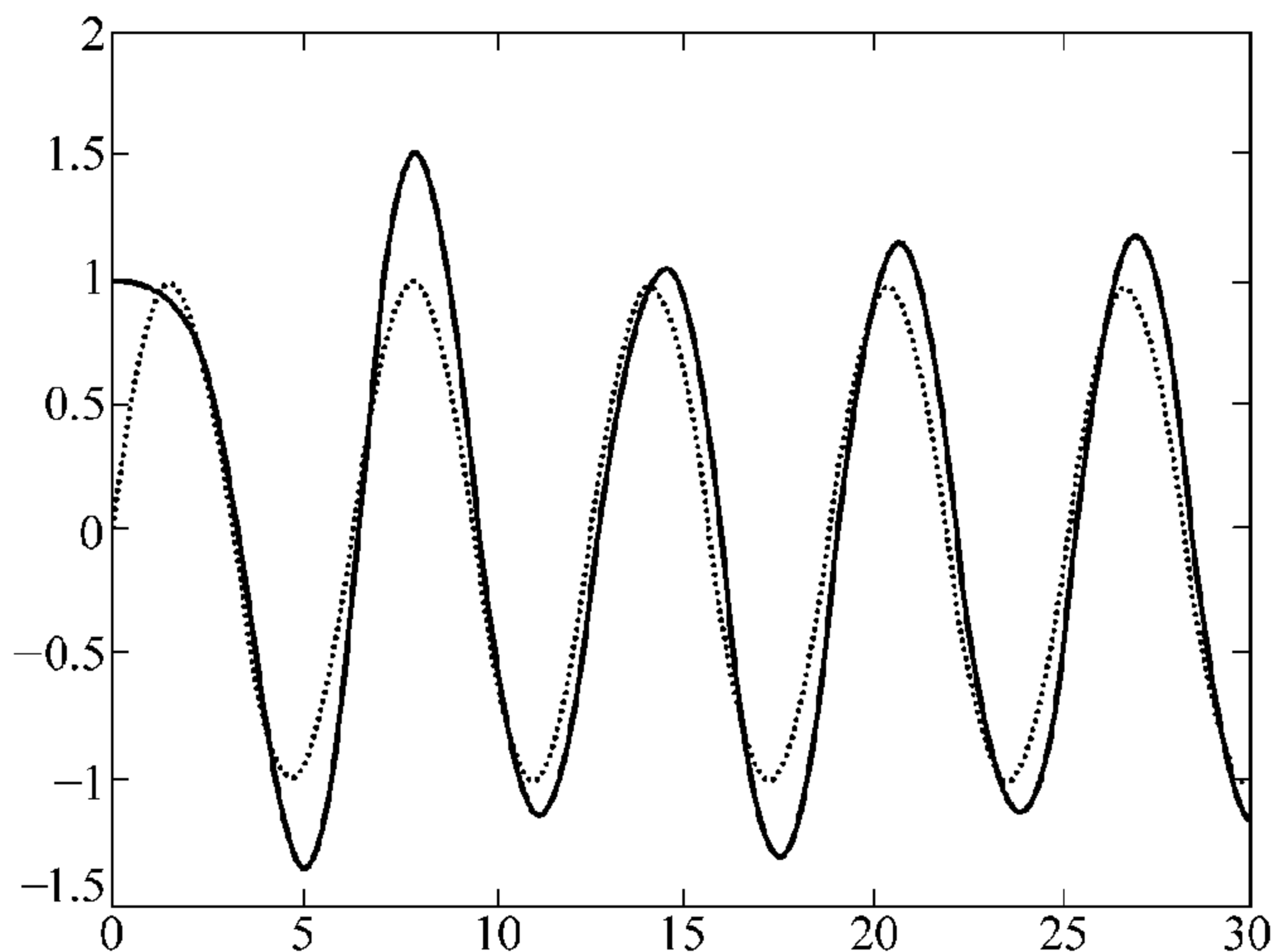


图 26.3 球—横木系统在初始条件为 $\mathbf{x}(0) = (1, 0, 0, 0)^T$ 时, 采用间接型自适应模糊控制器(26.31)~(26.33)所得到的输出 $y(t)$ (实线) 和期望轨迹 $y_m(t)$ (虚线)

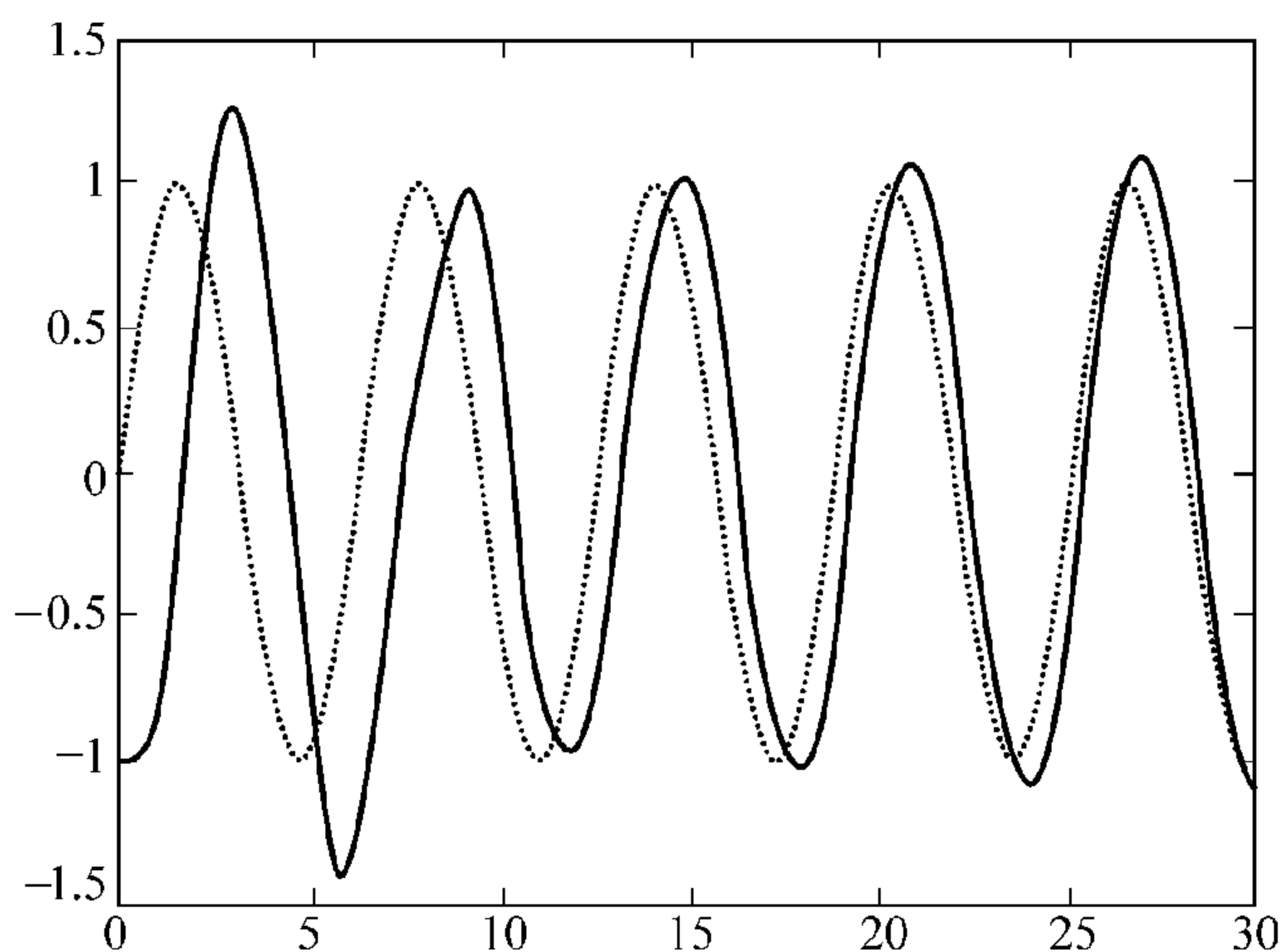


图 26.4 球—横木系统在初始条件为 $\mathbf{x}(0) = (-1, 0, 0, 0)^T$ 时, 采用间接型自适应模糊控制器(26.31)~(26.33)所得到的输出 $y(t)$ (实线) 和期望轨迹 $y_m(t)$ (虚线)

器所得到的输出 $y(t)$ (实线) 和期望轨迹 $y_m(t)$ (虚线)。比较图 26.3 和 26.4 与图 26.5 和图 26.6, 可以看出, 本例中直接型自适应模糊控制器的性能优于间接型自适应模糊控制器。

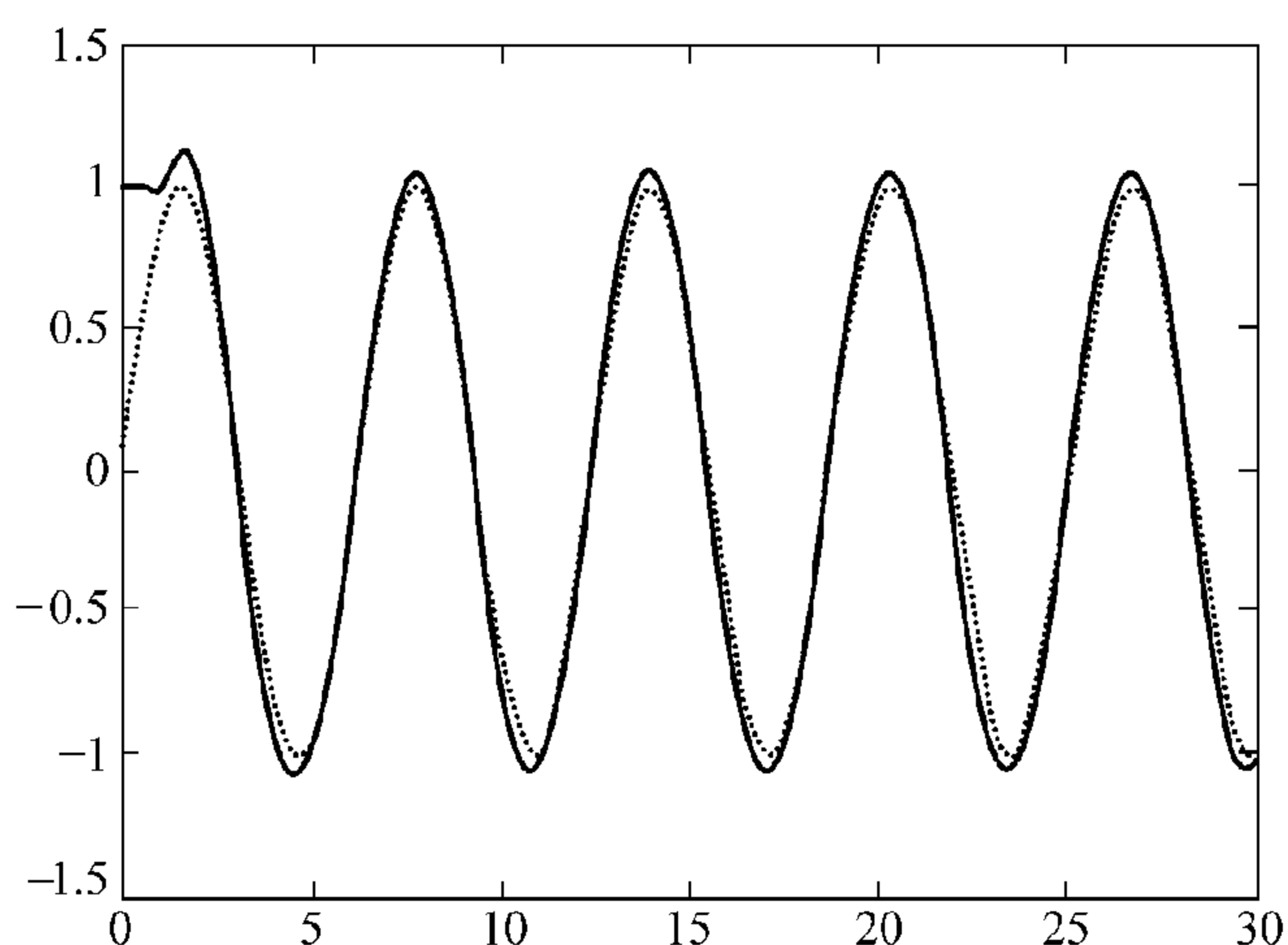


图 26.5 球—横木系统在初始条件为 $\mathbf{x}(0) = (1, 0, 0, 0)^T$ 时, 采用直接型自适应模糊控制器(26.34)~(26.35)所得到的输出 $y(t)$ (实线) 和期望轨迹 $y_m(t)$ (虚线)

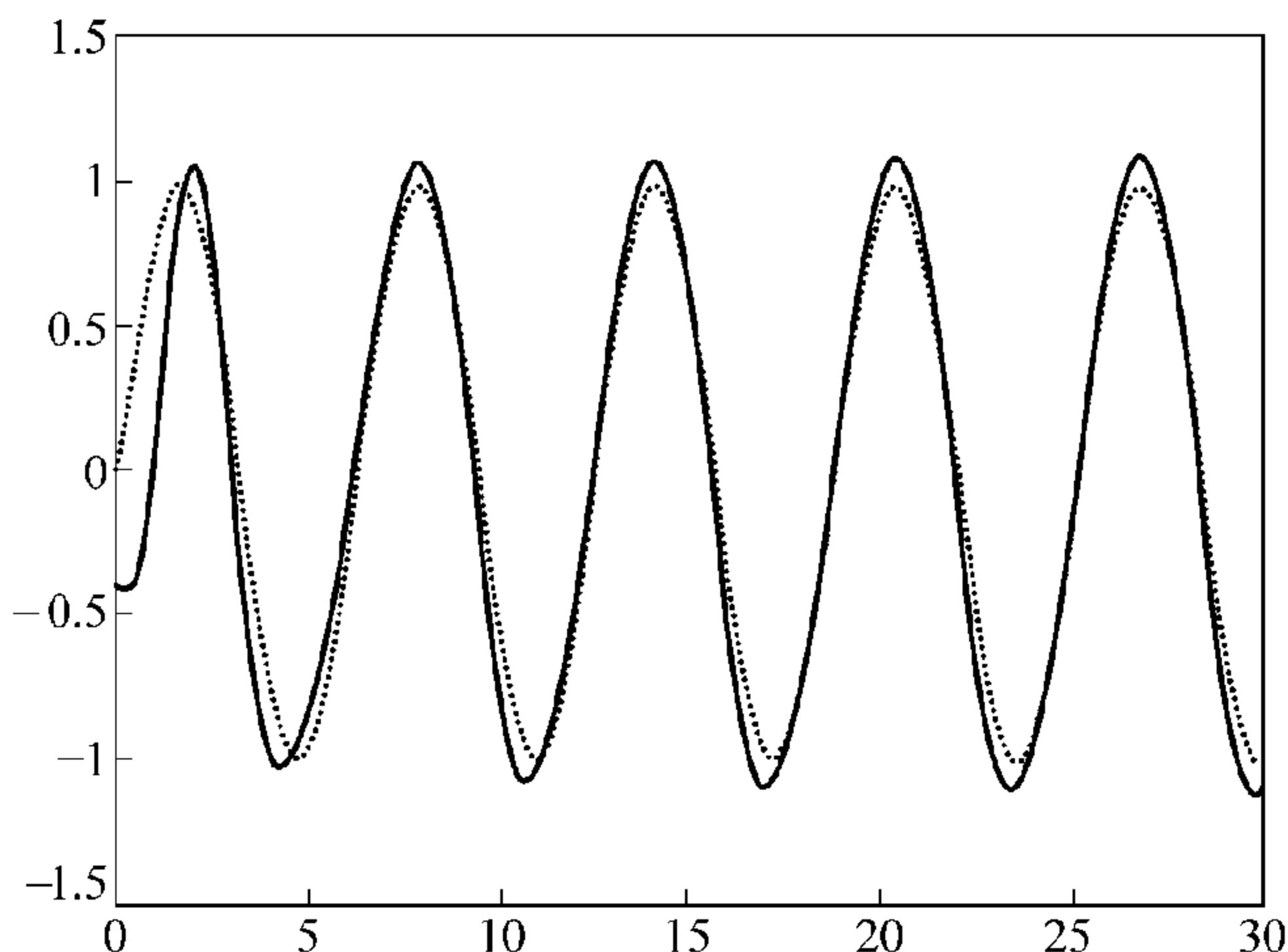


图 26.6 球—横木系统在初始条件为 $\mathbf{x}(0) = (-0.4, 0, 0, 0)^T$ 时, 采用直接型自适应模糊控制器(26.34)~(26.35)得到的输出 $y(t)$ (实线) 和期望轨迹 $y_m(t)$ (虚线)

26.3 本章小结及导读

本章主要阐述了以下内容:

- 在自适应调节过程中, 如果模糊系统中的 IF 部分的隶属度函数发生变化, 那么应如何设计自适应律;
- 输入—输出线性化的基本思想;
- 如何根据输入—输出线性化的概念为一般非线性系统设计自适应模糊控制器。

非线性控制理论是模糊控制进一步发展的重要工具。目前有许多非线性控制方面的好书, 如, Isidori[1989]给出了关于非线性控制的严谨论述; Slotine 和 Li[1991]的论述则更为通俗易懂。本章中的方法摘自 Wang[1994a]。

26.4 习题

习题 26.1 证明定理 26.1。

习题 26.2 假设式(24.9)中的直接型自适应模糊控制器 $u_d(\mathbf{x}|\theta)$ 是一个形如式(9.6)的模糊系统, 其中 $x_i^l = 1$, 试用李亚普诺夫综合法为参数 θ 设计自适应律。

习题 26.3 考虑如下非线性系统

$$\dot{x}_1 = x_1^2 x_2 \quad (26.36)$$

$$\dot{x}_2 = 3x_2 + u \quad (26.37)$$

$$y = -2x_1 - x_2 \quad (26.38)$$

设计一个状态反馈控制器使得闭环控制系统在平衡点 $\mathbf{x} = \mathbf{0}$ 处是局部渐近稳定的。

习题 26.4 为系统

$$\dot{x}_1 = \sin x_2 + \sqrt{t+1} x_2 \quad (26.39)$$

$$\dot{x}_2 = \alpha_1 x_1^4 \cos x_2 + \alpha_2 x_1 x_2 \sin x_2 + u \quad (26.40)$$

设计一个自适应控制器来跟踪任一期望轨迹 $x_{d1}(t)$ 。假设状态 $(x_1, x_2)^T$ 是可测量的, $x_{d1}(t), \dot{x}_{d1}(t), \ddot{x}_{d1}(t)$ 都是已知的, α_1, α_2 是未知常数。

习题 26.5 为一般非线性系统(26.16)设计一个直接型自适应模糊控制器, 并将其应用于球—横木系统进行仿真。

第六部分

其他专题

前言部分曾指出,在编写本书时,先是建立一个模糊系统与模糊控制理论应该遵循的合理框架,然后对其进行详细阐述。框架主要由第1章至第26章构成。但由于模糊系统与模糊控制是一个广阔且有很多分支的领域,所以一些重要的专题不可避免地未被纳入到这一理论框架中去。因此,本书的最后部分(第27章至第31章)将介绍第1章至26章中没有涉及的模糊系统与模糊控制方面的一些重要专题。这些专题和前面的内容同等重要。

第27章,将详细介绍模糊c-均值算法并研究其收敛特性,这种算法也许是模糊模式识别方法中最重要的一种方法。第28章,将介绍模糊关系方程并研究如何得到不同模糊关系方程的精确解和近似解,模糊关系方程在模糊系统与模糊控制中潜在地起着作用。第29章,将介绍模糊数的基本算术运算,包括模糊加、模糊减、模糊乘和模糊除,这些模糊数运算在模糊决策中非常重要。第30章,将介绍模糊决策中最重要专题——模糊线性规划,我们将证明使用模糊线性规划的必要性,并给出一些模糊线性规划问题的求解方法。最后,第31章将简要地复习一下概率论的基础知识,并用模糊(fuzziness)与概率(probability)的比较来结束本书。

27.1 为什么模糊模型可用于模式识别

模式识别(pattern recognition)是一门研究噪声或复杂环境下用机器识别重要的有效规律的学科。简言之,模式识别就是对数据结构的探索。例如,图 27.1 给出了四个在平面上的数据结构的例子。由图 27.1 可以看出,四个例子中的数据都被分成两组,不过不同例子中的“组”的定义是不同的。具体地,图 27.1 (a)中的组是根据数据点间的距离来划分的(即相互间距离小的数据点应分为一组),图 27.1 (b)中的组是根据数据点的连接特性来划分的(即很好地连接在一起的数据点应分为一组),图 27.1 (c)和 (d)中数据点分组的标准是基于距离特性和连接特性的混合尺度来划分的。在模式识别中,一组数据称为一类(cluster)。

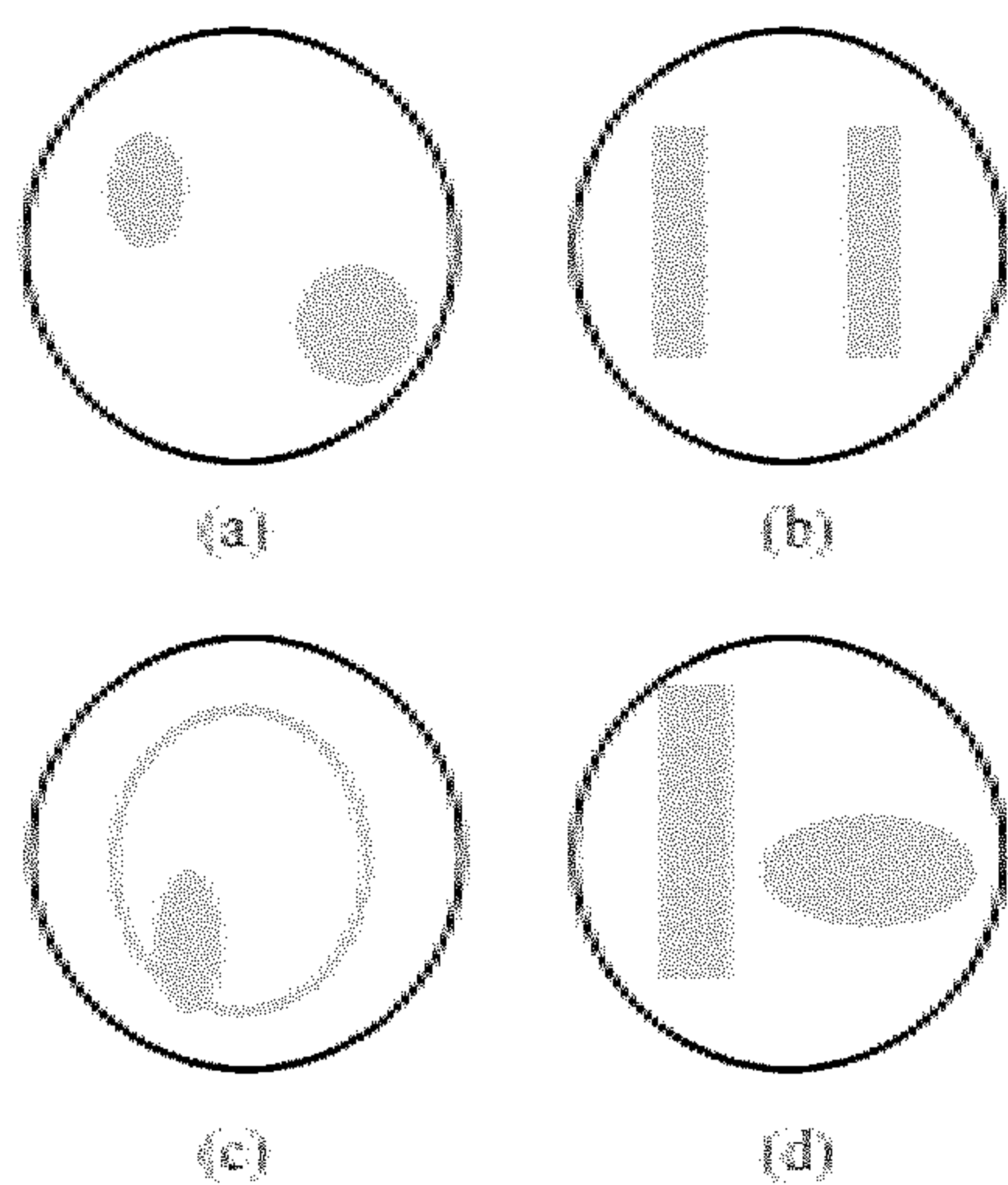


图 27.1 可能的数据结构

在实践中,数据通常不具有好的分布性质,因此,可能无法精确地定义“规律”或“结构”,即模式识别就其本质特性而言是一门不准确的学科。为了处理这种模棱两可的特性,可引入“模糊性”来帮助描述模式识别问题。如,类之间的界限有可能是模糊的、不清晰的,即一个数据点可能以不同的隶属度隶属于两个或更多的类。用这种方式会使问题的描述更贴近实际问题,从而有望实现更好的性能。这是模糊模型用于模式识别的第一个原因:模式识别问题的内在特性要求使用模糊模型(事实上,模糊建模意味着更灵活的建模——通过将 0—1 隶属度关系扩展到 $[0,1]$ 区间上隶属度关系而引入了更多的灵活性)。

模糊模型用于模式识别的第二个原因在于,模糊模型使得模式识别问题的求解运算更容易。求解一个非模糊模型通常要对整个空间进行全面搜索(因为一些关键变量的值只能取 0 和 1),而在模糊模型中所有变量都是连续的,可以进行梯度计算,进而能找到一个正确的搜索方向。在学习了 27.3 节中的模糊 c-均值算法后会加深对该原因的理解。

模式识别中的一个关键问题就是在一系列数据点中寻找所划分的类。现有文献已经提出了许多模糊聚类分析算法。本章将给出 Bezdek[1981]提出的最著名的模糊聚类分析算法——模糊 c-均值算法。感兴趣的读者可查阅 Bezdek 和 Pal[1992]提出的其他方法。目

前,模糊识别是模糊技术应用研究的一个活跃领域,且已取得了大量的成果,本章仅对其做简单的介绍。

27.2 硬 c-划分与模糊 c-划分

给定数据集 $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, 其中 \mathbf{x}_k 为任意元素, 如 $\mathbf{x}_i \in R^p$ 。令 $P(X)$ 为 X 的幂集 (集合 X 的所有子集的集合)。 X 的硬 c -划分是集合族 $\{A_i \in P(X) \mid 1 \leq i \leq c\}$, 它满足 $\bigcup_{i=1}^c A_i = X$ 和 $A_i \cap A_j = \phi (1 \leq i \neq j \leq c)$ 。每个 A_i 就是一类, 所以说, $\{A_1, \dots, A_c\}$ 将 X 划分为 c 类。

硬 c -划分可以用 A_i 中元素 \mathbf{x}_k 的特征 (隶属度) 函数来描述。具体的定义如下

$$u_{ik} = \begin{cases} 1, & \mathbf{x}_k \in A_i \\ 0, & \mathbf{x}_k \notin A_i \end{cases} \quad (27.1)$$

式中, $\mathbf{x}_k \in X, A_i \in P(X), i = 1, 2, \dots, c, k = 1, 2, \dots, n$ 。显然, $u_{ik} = 1$ 时, 表明 \mathbf{x}_k 属于类 A_i 。则给定 u_{ik} 的值, 就能够确定 X 的一个唯一的硬 c -划分, 反之亦然。 u_{ik} 应满足以下三个条件

$$u_{ik} \in \{0, 1\}, 1 \leq i \leq c, 1 \leq k \leq n \quad (27.2)$$

$$\sum_i^c u_{ik} = 1, \forall k \in \{1, 2, \dots, n\} \quad (27.3)$$

$$0 < \sum_{k=1}^n u_{ik} < n, \forall i \in \{1, 2, \dots, c\} \quad (27.4)$$

式(27.2)和式(27.3)说明所有 $\mathbf{x}_k \in X$ 应该属于且只能属于某一类。式(27.4)要求每一类 A_i 至少包含一个、至多含 $n-1$ 个数据点。用元素 $u_{ik} (1 \leq i \leq c, 1 \leq k \leq n)$ 构成矩阵 $U_{c \times n}$, 即可得到硬 c -划分的矩阵形式, 定义如下:

定义 27.1 硬 c -划分: 令 $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ 为任意集合, V_{cn} 为实 $c \times n$ 阶矩阵 $U = [u_{ik}]$ 的集合, 当 c 为整数且 $2 \leq c < n$ 时, X 的硬 c -划分空间为集合

$$M_c = \{U \in V_{cn} \mid \text{式(27.2) ~ 式(27.4)为真}\} \quad (27.5)$$

现在考虑硬 c -划分的一个例子。

例 27.1 设 X 为三个品牌的汽车的集合:

$$X = \{x_1 = \text{福特}, x_2 = \text{丰田}, x_3 = \text{克莱斯勒}\} \quad (27.6)$$

如果 $c = 2$, 则在约束(27.2) ~ (27.4)下, 有 X 的三个硬 c -划分

$$U_1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, U_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}, U_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (27.7)$$

约束(27.3) ~ (27.4)分别剔除了下面的矩阵

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad (27.8)$$

若目标是划分 X 为美国车和日本车, 则 U_3 是最合适的划分。不过, 正如第 2 章讨论过的那样, 美国车和日本车的差别不很明确, 因为美国车的许多部件是从日本进口来的, 而一些日本车却是在美国制造的。显然, 解决该问题的一种办法是允许 u_{ik} 在 $[0, 1]$ 区间上取任意值。

硬 c -划分的另一个问题是空间 M_c 过大。事实上

$$|M_c| = \frac{1}{c!} \left[\sum_{j=1}^c \binom{c}{j} (-1)^{c-j} j^n \right] \quad (27.9)$$

是明确地把 X 划分为 c 个非空子集的划法数。例如, $c = 10$ 且 $n = 25$ 时, 大约有 10^{18} 种方法可以明确地对 25 个数据进行 10-划分。这个问题起因于特征函数 u_{ik} 的离散性。尽管离散的 u_{ik} 生成一个有限的空间 M_c , 但由于 M_c 中元素的数量太多, 以致于寻求“最佳”划分成为了令人生畏的任务。若将 u_{ik} 变为一个连续变量, 使其在 $[0, 1]$ 区间上取任意值, 则就能够算出某个目标函数关于 u_{ik} 的梯度。进而利用梯度找到最好的搜索方向, 从而大大简化了寻求最佳划分的方法。

基于上述两种原因(概念上的适当性和计算上的简便性), 引入下面的模糊 c -划分概念。

定义 27.2 模糊 c -划分: 令 X, V_{cn} 和 c 的含义同定义 27.1, 则 X 的模糊 c -划分空间为集合

$$M_{fc} = \{U \in V_{cn} \mid u_{ik} \in [0, 1], 1 \leq i \leq c, 1 \leq k \leq n; \text{式(27.3)为真}\} \quad (27.10)$$

u_{ik} 是 \mathbf{x}_k 隶属于类 A_i 的隶属度值。注意, 条件(27.4)并没有包含在 M_{fc} 的定义中, 这就是 Bezdek[1981]所说的退化的模糊 c -划分。

重新考虑例 27.1, 使用模糊 c -划分, 可得如下 X 的一个更为合理的划分

$$U = \begin{bmatrix} 0.9 & 0.2 & 0.9 \\ 0.1 & 0.8 & 0.1 \end{bmatrix} \quad (27.11)$$

27.3 硬 c -均值算法和模糊 c -均值算法

27.3.1 目标函数聚类法和硬 c -均值算法

从空间 M_c 或 M_{fc} 中如何选取“最佳”划分呢? 有三种方法: 分级法、图论法和目标函数法。在分级法中, 根据相似性度量, 利用类的“合并”和“分裂”技术来建立新类, 最终生成一个嵌套的分级系统。在图论法中, 根据相似性度量, 令 X 为“相互连接”的节点的集合, 并按照节点集合间的连接性度量这一典型聚类准则进行聚类。在目标函数法中, 目标函数是对每个 c 的所有候选类的可取度的度量, 最优的类就是使目标函数达到局部最小值的类。因

为目标函数法的聚类准则公式是最精确的,所以本章将采用这种方法。

研究得最多的目标函数是总体组内误差平方和(overall within-group sum of squared errors),其定义为

$$J_w(U, V) = \sum_{k=1}^n \sum_{i=1}^c u_{ik} \| \mathbf{x}_k - \mathbf{v}_i \|^2 \quad (27.12)$$

式中 $U = [u_{ik}] \in M_c$ 或 M_{fc} , $V = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c)$, 其中 \mathbf{v}_i 是类 A_i 的中心,其定义为

$$\mathbf{v}_i = \frac{\sum_{k=1}^n u_{ik} \mathbf{x}_k}{\sum_{k=1}^n u_{ik}} \quad (27.13)$$

显然, \mathbf{v}_i 是类 A_i 中所有点的平均值(硬 c-划分)或加权平均值(模糊 c-划分)。从现在起,假设 $\mathbf{x}_k, \mathbf{v}_i \in R^p$ 。若 U 是一个硬 c-划分,则式(27.12)中的 $J_w(U, V)$ 可以改写成

$$J_w(U, V) = \sum_{i=1}^c \left(\sum_{\mathbf{x}_k \in A_i} \| \mathbf{x}_k - \mathbf{v}_i \|^2 \right) \quad (27.14)$$

“总体组内误差平方和” $J_w(U, V)$ 是根据其上面的表达式命名的。由于 $u_{ik} \| \mathbf{x}_k - \mathbf{v}_i \|^2$ 是 \mathbf{v}_i 表示成 \mathbf{x}_k 的误差平方,所以也可把 $u_{ik} \| \mathbf{x}_k - \mathbf{v}_i \|^2$ 看做是对局部密度的度量。如果所有(硬)类 A_i 中的点都紧密地贴近于它们的类中心 \mathbf{v}_i , 则 $J_w(U, V)$ 很小。

寻找使 J_w 最小的 (U, V) 并不容易,这是因为, M_c 的容量虽然有限但非常大[见式(27.9)]。最常见的寻找逼近 J_w 最小值的方法是下面的硬 c-均值算法(也称作 ISODATA 算法)。

硬 c-均值算法

步骤 1 假设给出 n 个数据点 $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, 其中 $\mathbf{x}_i \in R^p$ 。设一初始聚类数 $c (2 \leq c < n)$, 并初始化 $U^{(0)} \in M_c$ 。

步骤 2 当迭代次数为 $l (l=0, 1, 2, \dots)$ 时, 计算 c-均值向量

$$\mathbf{v}_i^{(l)} = \frac{\sum_{k=1}^n \mathbf{x}_k u_{ik}^{(l)}}{\sum_{k=1}^n u_{ik}^{(l)}} \quad (27.15)$$

其中 $[u_{ik}^{(l)}] = U^{(l)}, i=1, 2, \dots, c$ 。

步骤 3 用下式将 $U^{(l)}$ 更新为 $U^{(l+1)} = [u_{ik}^{(l+1)}]$

$$u_{ik}^{(l+1)} = \begin{cases} 1 & \| \mathbf{x}_k - \mathbf{v}_i^{(l)} \| = \min_{1 \leq j \leq c} (\| \mathbf{x}_k - \mathbf{v}_j^{(l)} \|) \\ 0 & \text{其他} \end{cases} \quad (27.16)$$

步骤 4 比较 $U^{(l)}$ 和 $U^{(l+1)}$, 若 $\| U^{(l+1)} - U^{(l)} \| < \varepsilon$ (ε 是一个非常小的常数), 则停止算法; 否则, 令 $l=l+1$, 返回步骤 2。

硬 c-均值算法直观上看十分合理: 猜想 c 的硬分类(步骤 1), 寻找各分类的中心(步骤

2), 重新分配类的隶属度以减少数据和当前中心的误差平方(步骤3), 当循环不再能显著地降低 J_w 时, 停止算法(步骤4)。如果令 J_w 关于每个 \mathbf{v}_i 的斜率均为零, 则可以证明, 用式(27.15)计算 $\{\mathbf{v}_i^{(l)}\}$ 是必要的。

例 27.2 假设 X 含有 R^2 中的 15 个点, 如图 27.2 所示, 这些点看起来像一只蝴蝶, 其中 \mathbf{x}_1 到 \mathbf{x}_7 构成左翅膀, \mathbf{x}_9 到 \mathbf{x}_{15} 构成右翅膀, \mathbf{x}_8 是连接两个翅膀的桥梁。当 $c=2$, 且

$$U^{(0)} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (27.17)$$

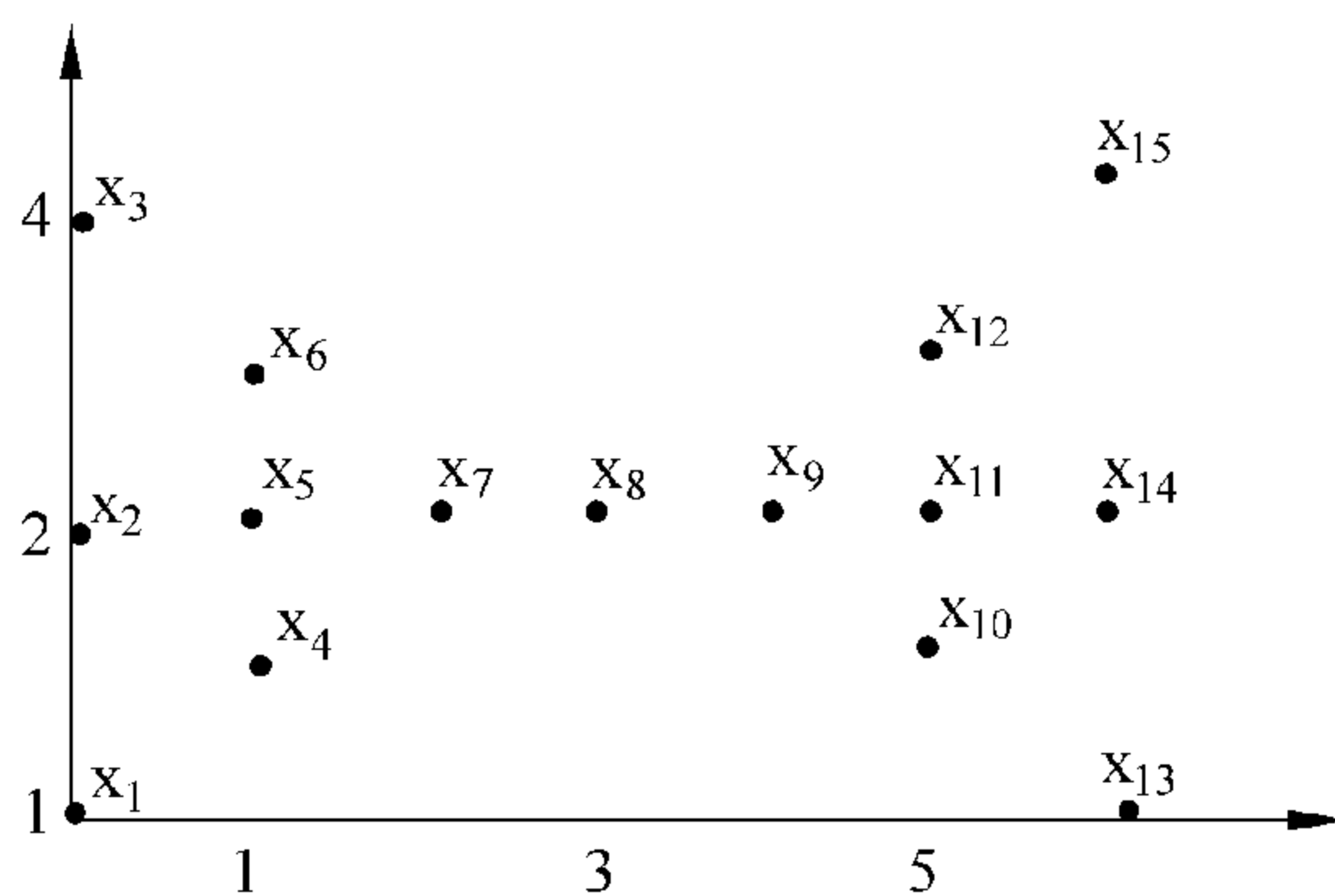


图 27.2 蝴蝶数据

时, 硬 c -均值算法终止于 $l=3$, 且有

$$U^{(3)} = U^{(4)} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (27.18)$$

$U^{(3)}$ 说明了 \mathbf{x}_1 到 \mathbf{x}_7 可以划分为一类 A_1 , 其他的 \mathbf{x}_8 到 \mathbf{x}_{15} 可以划分为另一类 A_2 。需要注意的是, 因为 \mathbf{x}_8 必须完全属于 A_1 或 A_2 , 这使得 A_1 和 A_2 不能关于 \mathbf{x}_8 对称, 而 X 中的数据本身是完全对称的, 所以非对称的分类是不具有直观吸引力的。解决这个问题的一個办法就是使用下面介绍的模糊 c -均值算法。

27.3.2 模糊 c -均值算法

模糊 c -均值算法的目标在于找到 $U = [u_{ik}] \in M_{fc}$ 和 $V = (\mathbf{v}_1, \dots, \mathbf{v}_c) (\mathbf{v}_i \in R^P)$, 使得

$$J_m(U, V) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m \|\mathbf{x}_k - \mathbf{v}_i\|^2 \quad (27.19)$$

最小, 其中 $m \in (1, \infty)$ 是一个加权指数。下面, 首先建立这个最小化问题的必要条件, 然后根据此条件提出模糊 c -均值算法。

定理 27.1 令 $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, $\mathbf{x}_i \in R^P$ 为一给定数据集。设定 $c \in \{2, 3, \dots, n-1\}$ 和 $m \in (1, \infty)$, 假设对所有 $1 \leq k \leq n$ 和 $1 \leq i \leq c$ 有 $\|\mathbf{x}_k - \mathbf{v}_i\| \neq 0$ 。则仅当

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{\|\mathbf{x}_k - \mathbf{v}_i\|}{\|\mathbf{x}_k - \mathbf{v}_j\|} \right)^{\frac{2}{m-1}}}, 1 \leq i \leq c, 1 \leq k \leq n \quad (27.20)$$

和

$$\mathbf{v}_i = \frac{\sum_{k=1}^n (u_{ik})^m \mathbf{x}_k}{\sum_{k=1}^n (u_{ik})^m}, 1 \leq i \leq c \quad (27.21)$$

时, $U = [u_{ik}]$ 和 $V = (\mathbf{v}_1, \dots, \mathbf{v}_c)$ 才是 $J_m(U, V)$ 的局部最小值。

证明 证明式(27.20)成立。假设 \mathbf{v}_i 不变, 则问题变为在约束条件(27.3)下, J_m 关于 u_{ik} 的最小化问题。引入拉格朗日乘子后, 本问题等价于下面的无约束最小化问题

$$L(U, \lambda) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m \|\mathbf{x}_k - \mathbf{v}_i\|^2 - \sum_{k=1}^n \lambda_k \left(\sum_{i=1}^c u_{ik} - 1 \right) \quad (27.22)$$

本问题的必要条件是

$$\frac{\partial L(U, \lambda)}{\partial u_{ik}} = [m(u_{ik})^{m-1} \|\mathbf{x}_k - \mathbf{v}_i\|^2 - \lambda_k] = 0 \quad (27.23)$$

$$\frac{\partial L(U, \lambda)}{\partial \lambda_k} = \sum_{i=1}^c u_{ik} - 1 = 0 \quad (27.24)$$

由式(27.23), 可得

$$u_{ik} = \left(\frac{\lambda_k}{m \|\mathbf{x}_k - \mathbf{v}_i\|^2} \right)^{\frac{1}{m-1}} \quad (27.25)$$

将式(27.25)代入式(27.24), 可得

$$\left(\frac{\lambda_k}{m} \right)^{\frac{1}{m-1}} = \frac{1}{\sum_{i=1}^c \left(\frac{1}{\|\mathbf{x}_k - \mathbf{v}_i\|^2} \right)^{\frac{1}{m-1}}} \quad (27.26)$$

再将式(27.26)代入式(27.25), 即可得到式(27.20)。

为证明式(27.21)成立, 假设 u_{ik} 不变。则本问题变为无约束最小化问题, 其必要条件是

$$\frac{\partial J_m(U, V)}{\partial \mathbf{v}_i} = - \sum_{k=1}^n 2(u_{ik})^m (\mathbf{x}_k - \mathbf{v}_i) = 0 \quad (27.27)$$

由上式可得式(27.21)。

模糊 c-均值算法是建立在必要条件(27.20) ~ (27.21)的基础上的。

模糊 c-均值算法(**fuzzy c-means algorithm**)

步骤 1 给定数据集 $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, $\mathbf{x}_i \in R^p$ 。设定 $c \in \{2, 3, \dots, n-1\}$ 和 $m \in (1, \infty)$, 并初始化 $U^{(0)} \in M_{fc}$ 。

步骤 2 当迭代次数为 $l(l=0, 1, 2, \dots)$ 时, 计算 c-均值向量

$$\mathbf{v}_i^{(l)} = \frac{\sum_{k=1}^n (u_{ik}^{(l)})^m \mathbf{x}_k}{\sum_{k=1}^n (u_{ik}^{(l)})^m}, 1 \leq i \leq c \quad (27.28)$$

步骤3 用下式将 $U^{(l)} = [u_{ik}^{(l)}]$ 更新为 $U^{(l+1)} = [u_{ik}^{(l+1)}]$

$$u_{ik}^{(l+1)} = \frac{1}{\sum_{j=1}^c \left(\frac{\|\mathbf{x}_k - \mathbf{v}_i^{(l)}\|}{\|\mathbf{x}_k - \mathbf{v}_j^{(l)}\|} \right)^{\frac{2}{m-1}}}, 1 \leq i \leq c, 1 \leq k \leq n \quad (27.29)$$

步骤4 若 $\|U^{(l+1)} - U^{(l)}\| < \varepsilon$ (ε 是一个非常小的常数), 则停止算法; 否则, 令 $l = l + 1$, 返回步骤2。

下面将模糊 c -均值算法应用于蝴蝶数据的例子。

例 27.3 令 $X = \{\mathbf{x}_1, \dots, \mathbf{x}_{15}\}$ 为图 27.2 中的蝴蝶数据, 取 $c = 2, m = 1.25, \varepsilon = 0.01$, 且

$$U^{(0)} = \begin{bmatrix} 0.854 & 0.146 & 0.854 & 0.854 & \cdots & 0.854 \\ 0.146 & 0.854 & 0.146 & 0.146 & \cdots & 0.146 \end{bmatrix}_{2 \times 15} \quad (27.30)$$

可得模糊 c -均值算法终止于 $l = 5$, 隶属度值 $u_{ik}^{(5)}$ 见图 27.3。可以看到, 左边翅膀和右边翅膀的数据得到了很好的划分, 其中桥 \mathbf{x}_8 几乎以同一程度隶属于这两类, 这很具有直观吸引力。

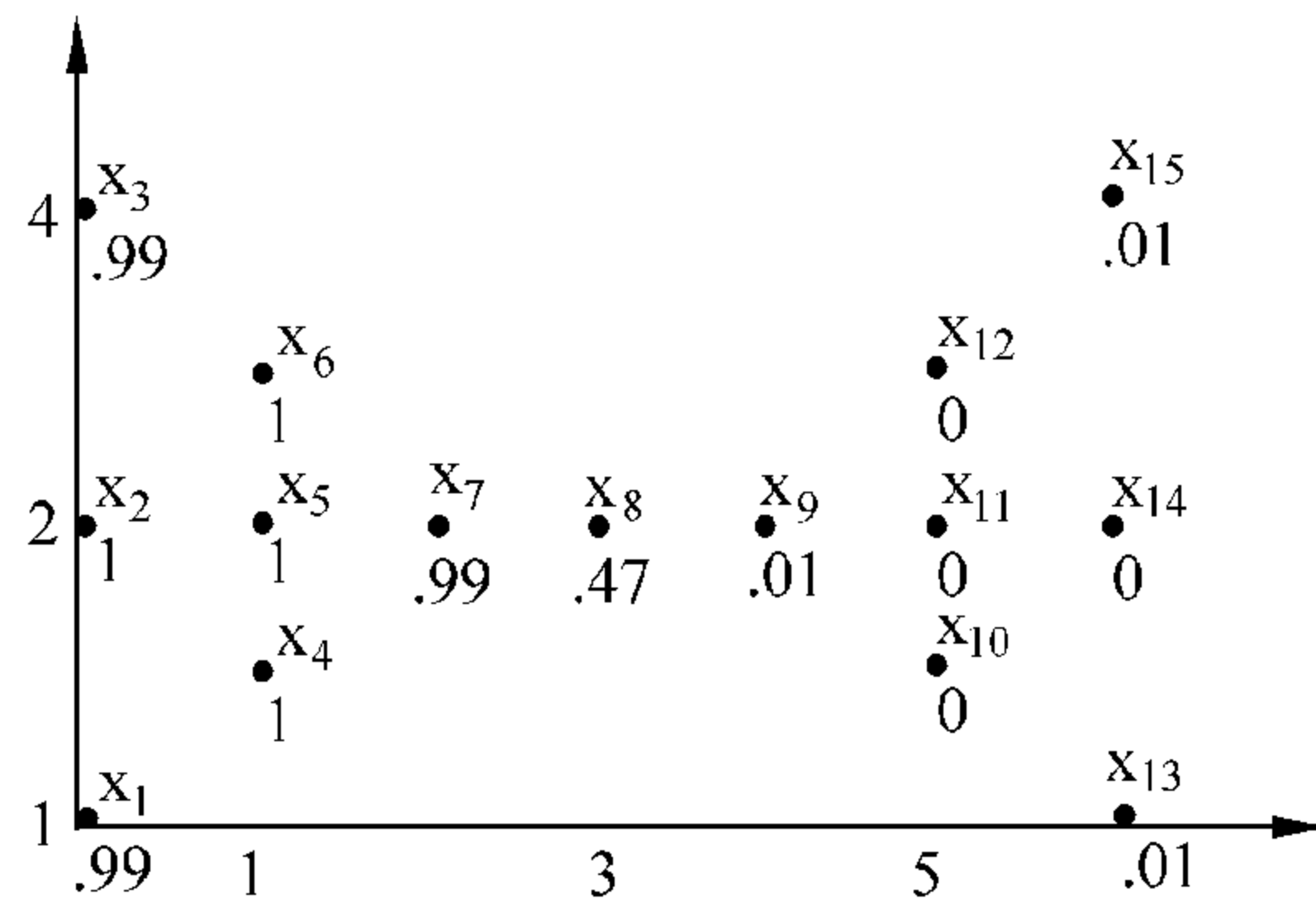


图 27.3 采用模糊 c -均值算法得到的蝴蝶数据点的隶属度值

27.4 模糊 c -均值算法的收敛性

定理 27.1 给出了 (U, V) 可以达到 J_m 的一个局部最小值的必要条件 (27.20) ~ (27.21)。在模糊 c -均值算法中, \mathbf{v}_i 和 u_{ik} 按照必要条件进行迭代计算。因此, 算法是否收敛于 J_m 的一个局部最小值还不清楚。事实上, 这是一个相当复杂的问题。在这一节中, 将只证明模糊 c -均值算法的一条本质特性: 通过迭代计算, 目标函数 $J_m(U, V)$ 将趋于减小或保持不变, 即 $J_m(U^{(l+1)}, V^{(l+1)}) \leq J_m(U^{(l)}, V^{(l)})$ 。

为证明这一性质,下面首先证明当 $J_m(U, V)$ 中的 V 不变时,式(27.20)是 J_m 局部最小化的一个充分条件(引理 27.1)。类似地,当 $J_m(U, V)$ 中的 U 不变时,式(27.21)是 J_m 最小化的一个充分条件(引理 27.2)。

引理 27.1 令 $\phi(U) = J_m(U, V)$, 其中 $V \in R^{p \times c}$ 是不变的, 且对所有 $1 \leq k \leq n$ 和 $1 \leq i \leq c$ 都有 $\|\mathbf{x}_k - \mathbf{v}_i\| \neq 0$, 则 $U = [u_{ik}]$ 达到 $\phi(U)$ 的一个局部最小值的充要条件是, U 是由式(27.20)计算出来的。

证明 必要性已在定理 27.1 证完。下面证明充分性。先根据式(27.20)计算 U , 然后计算引入了拉格朗日乘子的 $\phi(U)$ 的 $cn \times cn$ 阶海森矩阵—— $H(U)$ 。由式(27.23), 有

$$h_{st, ik}(U) = \frac{\partial}{\partial u_{st}} \left[\frac{\partial \phi(U)}{\partial u_{ik}} \right] = \begin{cases} m(m-1)(u_{st})^{m-2} \|\mathbf{x}_t - \mathbf{v}_s\|^2 & s=i, t=k \\ 0 & \text{其他} \end{cases} \quad (27.31)$$

式中, u_{st} 是由式(27.20)计算出来的。因此, $H(U) = [h_{st, ik}(U)]$ 是对角矩阵。由于 $m > 1$ 且对所有 $1 \leq t \leq n$ 和 $1 \leq s \leq c$ 有 $\|\mathbf{x}_t - \mathbf{v}_s\| \neq 0$, 进而有 $m(m-1)(u_{st})^{m-2} \|\mathbf{x}_t - \mathbf{v}_s\|^2 > 0$ 。因此, 海森矩阵 $H(U)$ 是正定的, 从而说明式(27.20)是 $\phi(U)$ 最小化的充分条件。

引理 27.2 令 $\psi(V) = J_m(U, V)$, 其中 $U \in M_{fc}$ 是不变的, 且对所有 $1 \leq k \leq n, 1 \leq i \leq c$ 和 $m > 1$ 都有 $\|\mathbf{x}_k - \mathbf{v}_i\| \neq 0$ 。则 $V = [v_i]$ 达到 $\psi(V)$ 的一个局部最小值的充要条件是, V 是由式(27.21)计算出来的。

证明 必要性已在定理 27.1 证完。下面证明充分性。由式(27.27), 可得

$$\frac{\partial}{\partial v_j} \left[\frac{\partial \psi(V)}{\partial v_i} \right] = \begin{cases} \sum_{k=1}^n 2(u_{jk})^m & j=i \\ 0 & \text{其他} \end{cases} \quad (27.32)$$

因此, 海森矩阵是正定的, 从而说明式(27.21)是 $\psi(V)$ 最小化的充分条件。

由引理 27.1 和引理 27.2 可以证明这里的主要结论: $J_m(U^{(l+1)}, V^{(l+1)}) \leq J_m(U^{(l)}, V^{(l)})$ 。

定理 27.2 令 $\mathbf{v}_i^{(l)}, u_{ik}^{(l)} (l=0, 1, 2, \dots)$ 是根据模糊 c-均值算法(27.28)~(27.29)计算而得的序列。若 $m > 1$, 且对所有的 $k=1, 2, \dots, n, i=1, 2, \dots, c$ 和 $l=0, 1, 2, \dots$, 都有 $\|\mathbf{x}_k^{(l)} - \mathbf{v}_i^{(l)}\| \neq 0$, 则对所有 $l=0, 1, 2, \dots$, 有

$$J_m(U^{(l+1)}, V^{(l+1)}) \leq J_m(U^{(l)}, V^{(l)}) \quad (27.33)$$

证明 由于 $V^{(l)}$ 是在 U 不变的条件下, 由式(27.28)计算出来的, 所以由引理 27.2 可得

$$J_m(U^{(l+1)}, V^{(l+1)}) \leq J_m(U^{(l+1)}, V^{(l)}) \quad (27.34)$$

由于 $U^{(l+1)}$ 是在 V 不变的条件下, 由式(27.29)计算出来的, 所以由引理 27.1 可得

$$J_m(U^{(l+1)}, V^{(l)}) \leq J_m(U^{(l)}, V^{(l)}) \quad (27.35)$$

合并式(27.34)和式(27.35), 即可得到式(27.33)。

令 Ω 集合为

$$\Omega = \{(U^*, V^*) \in M_{fc} \times R^{p \times c} \mid J_m(U^*, V^*) \leq J_m(U, V^*), \forall U \in M_{fc}\};$$

$$J_m(U^*, V^*) < J_m(U^*, V), V \neq V^* \} \quad (27.36)$$

在 Bezdek, Hathaway, Sabin 和 Tucker[1987] 中已经证明了, 模糊 c -均值算法要么终止于 Ω 中的某一点, 要么终止于收敛于 Ω 中某一点的一个子序列。具体证明可参考 Bezdek, Hathaway, Sabin 和 Tucker[1987]。

27.5 本章小结及导读

本章主要阐述了以下内容:

- 模糊模型用于模式识别的动因;
- 硬 c -划分和模糊 c -划分的定义以及硬 c -均值算法的步骤;
- 模糊 c -均值算法的详细步骤以及其理论证明;
- 模糊 c -均值算法的收敛性。

原始文献 Bezdek[1981] 一直是学习模糊 c -均值算法的最好著作。模糊 c -均值算法和用于模式识别的其他模糊算法的有关文章可参考 Bezdek 和 Pal[1992]。传统模式识别的方法可参考 Duda 和 Hart[1973]。

27.6 习题

习题 27.1 证明将 n 个元素划分为 c 个非空子集的划法数是由式(27.9)给出的。

习题 27.2 令 $U \in M_{fc}$, 且

$$F(U) = \sum_{k=1}^n \sum_{i=1}^c \frac{(u_{ik})^2}{n} \quad (27.37)$$

证明: (a) 对任意 $U \in M_{fc}$, 有 $(1/c) \leq F(U) \leq 1$ 。

(b) 当且仅当 U 为硬时, 有 $F(U) = 1$

(c) 当且仅当 $\forall i, k$ 有 $u_{ik} = (1/c)$ 时, $F(U) = (1/c)$ 成立。

习题 27.3 令 $\mathbf{x}_1, \dots, \mathbf{x}_n$ 为 n 个 d 维样本, 且 Σ 为任一 $d \times d$ 阶非奇异矩阵, 证明使

$$\sum_{k=1}^n (\mathbf{x}_k - \mathbf{x})^T \Sigma^{-1} (\mathbf{x}_k - \mathbf{x}) \quad (27.38)$$

最小的向量 \mathbf{x} , 是样本均值 $\frac{1}{n} \sum_{k=1}^n \mathbf{x}_k$

习题 27.4 证明 V 不变时, 式(27.16)是使式(27.12)中的 $J_m(U, V)$ 最小化的必要条件。

习题 27.5 证明式(27.12)中的 $J_w(U, V)$ 随 c 的增大而单调递减。

习题 27.6 取初始划分为

$$U^{(0)} = \begin{bmatrix} (0.8)_{1 \times 8} & (0.2)_{1 \times 7} \\ (0.2)_{1 \times 8} & (0.8)_{1 \times 7} \end{bmatrix} \quad (27.39)$$

重做例 27.3, 这里 $(x)_{1 \times r}$ 是一个 $1 \times r$ 维的向量, 且向量中每个元素均为 x 。

习题 27.7 将模糊 c -均值算法应用于蝴蝶数据, 取 $c = 3, m = 1.25, \varepsilon = 0.01$ 和

$$U^{(0)} = \begin{bmatrix} (0.4)_{1 \times 8} & (0.6)_{1 \times 7} \\ (0.3)_{1 \times 8} & (0.7)_{1 \times 7} \\ (0.3)_{1 \times 8} & (0.7)_{1 \times 7} \end{bmatrix} \quad (27.40)$$

习题 27.8 利用模糊 c -均值算法, 给出一种根据输入和输出数据设计模糊系统的方法。

28.1 引言

给定输入空间 U 上的模糊集 A 和输入—输出乘积空间 $U \times V$ 中的模糊关系 Q , 由推理的合成规则(见第 6 章), 可得 V 上的模糊集 B 为

$$\mu_B(y) = \sup_{x \in U} (\mu_A(x) \star \mu_Q(x, y)) \quad (28.1)$$

令 \circ 表示 Sup- \star 合成, 则式(28.1)可以改写成

$$B = A \circ Q \quad (28.2)$$

如果把模糊关系 Q 看做一个纯模糊系统(见第 1 章), 则式(28.1)或式(28.2)给出的是如何在给定输入 A 和系统本身 Q 的情况下计算系统的输出 B 。有些时候, 思考下面两个问题是有意义的:

- 给定系统 Q 及其输出 B , 确定系统的输入 A , 这同信号处理和通信中反卷积或均衡问题相近, 称之为模糊反卷积问题(fuzzy deconvolution problem)。
- 给定系统的输入 A 和输出 B , 确定系统 Q , 这与控制中的系统辨识问题相近, 称之为模糊识别问题(fuzzy identification problem)。

因此, 求解模糊关系方程(28.2)就等价于求解上述两个问题。本章的目的就是研究如何求解这两个问题。

28.2 求解模糊关系方程

在这一节中, 首先介绍一种有用的算子—— φ -算子, 并研究其性质。然后, 再证明几个引理, 由此可以得到模糊识别问题和模糊反卷积问题的解。

定义 28.1 φ -算子是二元算子 $\varphi: [0, 1] \times [0, 1] \rightarrow [0, 1]$, 其定义是

$$a\varphi b = \sup [c \in [0, 1] | a \star c \leq b] \quad (28.3)$$

式中 \star 是 t -范数算子。

显然, t -范数 \star 不同, φ -算子也不同。具体地, 定义 \star 为最小算子, 则 φ -算子就变成了所谓的 α -算子

$$a \alpha b = \sup [c \in [0, 1] | \min(a, c) \leq b] = \begin{cases} 1 & a \leq b \\ b & a > b \end{cases} \quad (28.4)$$

现在证明 φ -算子的几个有用的性质。

引理 28.1 设 φ -算子由式(28.3)所定义, 则有下列不等式成立

$$a \varphi \max(b, c) \geq \max(a \varphi b, a \varphi c) \quad (28.5)$$

$$a \star (a \varphi b) \leq b \quad (28.6)$$

$$a \varphi (a \star b) \geq b \quad (28.7)$$

式中, $a, b, c \in [0, 1]$ (通常 a, b, c 为隶属度函数)。

证明 由式(28.3)可知, 显然 $a \varphi b$ 是关于第二个变量 b 的非减函数, 即如果 $b_1 \geq b_2$, 则 $a \varphi b_1 \geq a \varphi b_2$, 进而有, $a \varphi \max(b, c) \geq a \varphi b$ 和 $a \varphi \max(b, c) \geq a \varphi c$, 故式(28.5)成立。式(28.6)可直接由定义(28.3)得出。最后, 由 $a \varphi (a \star b) = \sup [c \in [0, 1] | a \star c \leq a \star b]$ 和 t -范数的非减性, 可得式(28.7)成立。

下面研究模糊识别问题, 即, 给定 A 和 B , 求解模糊关系方程(28.2)中的 Q 。首先, 证明其特解为 $\hat{Q} = A \varphi B$ 。为证明这一结果, 需要引入以下两个引理。

引理 28.2 令 A 为定义在 U 上的模糊集合, Q 为定义在 $U \times V$ 中的模糊关系, 则

$$Q \subseteq A \varphi (A \circ Q) \quad (28.8)$$

式中, $A \varphi (A \circ Q)$ 是定义在 $U \times V$ 中的模糊关系, 其隶属度函数为 $\mu_A(x) \varphi \mu_{A \circ Q}(y)$ 。

证明 因为 $a \varphi b$ 是一个关于 b 的非减函数, 所以有

$$\begin{aligned} \mu_{A \varphi (A \circ Q)}(x, y) &= \mu_A(x) \varphi \left[\sup_{x \in U} (\mu_A(x) \star \mu_Q(x, y)) \right] \\ &\geq \mu_A(x) \varphi [\mu_A(x) \star \mu_Q(x, y)] \end{aligned} \quad (28.9)$$

由式(28.9), 可得

$$\mu_{A \varphi (A \circ Q)}(x, y) \geq \mu_Q(x, y) \quad (28.10)$$

故式(28.8)成立。

引理 28.3 设 A 和 B 是分别定义在 U 和 V 上的模糊集, 则有

$$A \circ (A \varphi B) \subseteq B \quad (28.11)$$

式中, $A \varphi B$ 是定义在 $U \times V$ 中的模糊关系, 其隶属度函数为 $\mu_A(x) \varphi \mu_B(y)$ 。

证明 由式(28.6), 可得

$$\begin{aligned} \mu_{A \circ (A \varphi B)}(y) &= \sup_{x \in U} [\mu_A(x) \star (\mu_A(x) \varphi \mu_B(y))] \\ &\leq \sup_{x \in U} \mu_B(y) \\ &= \mu_B(y) \end{aligned} \quad (28.12)$$

上式等价于式(28.11), 故式(28.11)成立。

下面求解模糊识别问题。

定理 28.1 令 \mathcal{Q} 为给定 A 和 B 的条件下模糊关系方程(28.2)所有的解 Q 的集合。如

果 \mathcal{Q} 为非空集合, 则 \mathcal{Q} 中最大的元素(在集合论意义上)为

$$\hat{Q} = A \varphi B \quad (28.13)$$

证明 设 Q 为 \mathcal{Q} 中的任意一个元素, 则有 $B = A \circ Q$ 。由引理 28.2, 可得 $Q \subseteq A \varphi B = \hat{Q}$ 。因为 t -范数是非减函数, 所以有 $B = A \circ Q \subseteq A \circ \hat{Q}$ 。又由引理 28.3, 有 $A \circ \hat{Q} = A \circ (A \varphi B) \subseteq B$, 进而有 $B = A \circ \hat{Q}$, 这表明 \hat{Q} 为 \mathcal{Q} 中的一个元素。因为 $Q \subseteq \hat{Q}$, 且 Q 是 \mathcal{Q} 中的任意一个元素, 所以可得出结论, \hat{Q} 为 \mathcal{Q} 中的最大元素。

下面研究模糊反卷积问题, 即在给定 B 和 Q 的条件下, 求解模糊关系方程(28.2)中的 A 。首先, 再证明两个引理, 然后由此给出模糊关系方程的一个可能的解。

引理 28.4 令 B 为定义在 V 上的模糊集合, Q 为定义在 $U \times V$ 中的模糊关系, 则有

$$(Q \varphi B) \circ Q \subseteq B \quad (28.14)$$

式中, 合成 $(Q \varphi B) \circ Q$ 是 V 上一个模糊集, 其隶属度函数为 $\sup_{x \in U} [\mu_{Q \varphi B}(x, y) \star \mu_Q(x, y)]$ 。

证明: 由式(28.6)可得

$$\begin{aligned} \mu_{(Q \varphi B) \circ Q}(y) &= \sup_{x \in U} [(\mu_Q(x, y) \varphi \mu_B(y)) \star \mu_Q(x, y)] \\ &\leq \sup_{x \in U} [\mu_B(y)] \\ &= \mu_B(y) \end{aligned} \quad (28.15)$$

故式(28.14)成立。

引理 28.5 设 A 为定义在 U 上的模糊集合, Q 为定义在 $U \times V$ 中的模糊关系, 则有

$$A \subseteq Q \varphi (A \circ Q) \quad (28.16)$$

式中, 合成 $Q \varphi (A \circ Q)$ 是 U 上的一个模糊集, 其隶属度函数为 $\sup_{y \in V} [\mu_Q(x, y) \varphi \mu_{A \circ Q}(y)]$ 。

证明: 由式(28.7)和 $a \varphi b$ 关于 b 的非减性, 可得

$$\begin{aligned} \mu_{Q \varphi (A \circ Q)}(x, y) &= \sup_{y \in V} \{(\mu_Q(x, y) \varphi [\sup_{x \in U} (\mu_A(x) \star \mu_Q(x, y))])\} \\ &\geq \sup_{y \in V} [\mu_Q(x, y) \varphi (\mu_A(x) \star \mu_Q(x, y))] \\ &\geq \sup_{y \in V} [\mu_A(x)] = \mu_A(x) \end{aligned} \quad (28.17)$$

故式(28.16)成立。

下面给出模糊反卷积问题的解的一个简洁公式。

定理 28.2 设 \mathcal{A} 是在给定 B 和 Q 条件下, 模糊关系方程(28.2)所有的解 A 的集合, 如果 \mathcal{A} 非空, 则 \mathcal{A} 中最大的元素为

$$\hat{A} = Q \varphi B \quad (28.18)$$

式中, $Q \varphi B$ 是定义在 U 上的一个模糊集, 其隶属度函数为 $\sup_{y \in V} [\mu_Q(x, y) \varphi \mu_B(y)]$ 。

证明 设 A 为 \mathcal{A} 中的任意一个元素, 则有 $B = A \circ Q$ 。由引理 28.5, 可得 $A \subseteq Q \varphi (A \circ Q) = Q \varphi B = \hat{A}$, 这表明 $B = A \circ Q \subseteq \hat{A} \circ Q$ 成立。再由引理 28.4, 可得 $\hat{A} \circ Q = (Q \varphi B) \circ Q$ 。

$Q \subseteq B$, 进而有 $\hat{A} \circ Q = B$, 这表明 \hat{A} 为 \mathcal{A} 中的一个元素。又因为 A 是 \mathcal{A} 中的任意元素且 $A \subseteq \hat{A}$, 所以 \hat{A} 为 \mathcal{A} 中的最大元素。

定理 28.1 和定理 28.2 成立的基本前提是该问题有解。但在许多情况下, 问题的精确解可能不存在。因此, $\hat{Q} = A \varphi B$ 和 $\hat{A} = Q \varphi B$ 可能并不是问题的解, 见下例。

例 28.1 设, $U = \{x_1, x_2\}, V = \{y_1, y_2\}$

$$A = 0.4/x_1 + 0.3/x_2 \quad (28.19)$$

$$B = 0.5/y_1 + 0.4/y_2 \quad (28.20)$$

$$Q = 0.3/(x_1, y_1) + 0.2/(x_1, y_2) + 0.1/(x_2, y_1) + 0/(x_2, y_2) \quad (28.21)$$

令 φ -算子为 α -算子(28.4), 则有

$$\hat{Q} = A \varphi B = 1/(x_1, y_1) + 1/(x_1, y_2) + 1/(x_2, y_1) + 1/(x_2, y_2) \quad (28.22)$$

sup- \star 合成中的 t -范数选用最小算子, 则有

$$A \circ \hat{Q} = 0.4/y_1 + 0.4/y_2 \neq B \quad (28.23)$$

因此, $\hat{Q} = A \varphi B$ 不是问题的解。类似地, 将 φ 算子取为 α 算子, t -范数取为最小算子, 则有

$$\hat{A} = Q \varphi B = 1/x_1 + 1/x_2 \quad (28.24)$$

和

$$\hat{A} \circ Q = 0.3/y_1 + 0.2/y_2 \neq B \quad (28.25)$$

那么, 在什么条件下, 模糊识别问题和模糊反卷积问题的解存在呢? 下一小节将回答这个问题。

28.3 模糊关系方程的可解度

因为 B 和 $A \circ Q$ 是 V 上的模糊集, 所以, 模糊关系方程(28.2)是否有解与 B 和 $A \circ Q$ 这两个模糊集是否相等是等价的。因此, 首先研究如何度量两个模糊集的接近程度。

28.3.1 两个模糊集的贴程度

令 C 和 D 是 U 上的两个模糊集, 则很自然地, C 和 D 的贴程度(Equality Indices)为

$$d_p(C, D) = \left(\int_U |C(x) - D(x)|^p dx \right)^{1/p}, p \geq 1 \quad (28.26)$$

当 $p=1$ 时, 上式为海明距离(Hamming distance), 当 $p=2$ 时, 上式为欧氏距离(Euclidean distance)。定义模糊集 C 和 D 的贴程度为

$$eq_p(C, D) = 1 - d_p(C, D) \quad (28.27)$$

显然,当且仅当 $C = D$ 时, $eq_p(C, D) = 1$, 因此, $eq_p(C, D)$ 可以作为一种贴近度。由于 $d_p(C, D)$ 的计算要用到复杂的积分, 所以, 研究其他度量两个模糊集的贴近度的方法是有意义的。

另一种比较模糊集的方法就是用更多的面向集合理论的工具。首先, 由 φ -算子的定义(28.3)和 t -范数的边界条件(见第3章)可知, 当且仅当 $\inf_{x \in U} [\mu_C(x) \varphi \mu_D(x)] = 1$ 时, 有 $C \subseteq D$ 。又因为 $C = D$ 等价于 $C \subseteq D$ 且 $D \subseteq C$, 所以 C 和 D 的贴近度可定义为

$$eq_t(C, D) = \inf_{x \in U} [\mu_C(x) \varphi \mu_D(x)] \star \inf_{x \in U} [\mu_D(x) \varphi \mu_C(x)] \quad (28.28)$$

显然, 当且仅当 $C = D$ 时, $eq_t(C, D) = 1$, 因此 $eq_t(C, D)$ 也可以作为一个贴近度。又因为, $a \varphi b = 1$ (当且仅当 $a \leq b$), 式(28.28)中第二个 $\inf_{x \in U}$ 可以提前(根据 t -范数的连续性), 所以贴近度可写成

$$eq_t(C, D) = \inf_{x \in U} [(\mu_C(x) \vee \mu_D(x)) \varphi (\mu_C(x) \wedge \mu_D(x))] \quad (28.29)$$

其中, \vee 表示取最大, \wedge 表示取最小。

现在, 已经为定义模糊关系方程的可解度做好了准备。

28.3.2 模糊关系方程的可解度

用两个模糊集的贴近度定义模糊关系方程(28.2)的可解度(Solvability Indices)为

$$\xi = eq_t(A \circ Q, B) \quad (28.30)$$

但因为 A 或 Q 是未知的, 所以 ξ 无法计算。由定理 28.1 和定理 28.2 可知, “模糊识别(模糊反卷积)问题”的一个可能解是 $\hat{Q} = A \varphi B$ ($\hat{A} = Q \varphi B$), 因此, 定义可解度的一个方法是将式(28.30)中的 Q 用 \hat{Q} 代替, A 用 \hat{A} 代替。具体地, 有如下定义。

定义 28.2 定义模糊识别问题的可解度为

$$\xi_I = eq_t[A \circ (A \varphi B), B] \quad (28.31)$$

定义模糊反卷积问题的可解度为

$$\xi_D = eq_t[(Q \varphi B) \circ Q, B] \quad (28.32)$$

显然, $\xi_I = 1$ ($\xi_D = 1$) 不仅表明“模糊识别(模糊反卷积)问题”是可解的, 而且表明 $\hat{Q} = A \varphi B$ ($\hat{A} = Q \varphi B$) 就是它的一个解。因为在定义 ξ_I 和 ξ_D 时所使用的是特解, 所以有人可能想知道, $\xi_I \neq 1$ ($\xi_D \neq 1$) 是否就意味着“模糊识别(模糊反卷积)问题”是不可解的, 会不会此时只存在使 $eq_t(A \circ \hat{Q}', B) = 1$ ($eq_t(\hat{A}' \circ Q, B) = 1$) 成立的 \hat{Q}' (\hat{A}'), 而不存在使其成立的 $A \varphi B$ ($Q \varphi B$)。幸运的是, 这样的 \hat{Q}' (\hat{A}') 并不存在。因为由定理 28.1 和定理 28.2 可知, 如果 $\hat{Q} = A \varphi B$ ($\hat{A} = Q \varphi B$) 不是“模糊识别(模糊反卷积)问题”的一个解, 那么该问题就不存在其他解了。因此, 当且仅当“模糊识别(模糊反卷积)问题”可解时, 才有 $\xi_I = 1$ ($\xi_D = 1$)。

下面的任务是简化 ξ_I 和 ξ_D , 使它们便于计算。为此, 首先介绍下面的引理。

引理 28.6 对任意 $a, b \in [0, 1]$, 有下面两式成立:

$$a \star (a \varphi b) = \min(a, b) \quad (28.33)$$

$$b \varphi (a \star (a \varphi b)) = b \varphi a \quad (28.34)$$

证明 如果 $a \leq b$, 则, $a \varphi b = 1$, $a \star (a \varphi b) = a = \min(a, b)$; 如果 $a > b$, 则由 φ -算子的定义, 可得 $a \star (a \varphi b) = a \star \sup\{c \in [0, 1] \mid a \star c \leq b\} = \sup\{a \star c \mid a \star c \leq b\} = b = \min(a, b)$, 于是式(28.33)得证。由式(28.33)可知, $b \varphi (a \star (a \varphi b)) = b \varphi[\min(a, b)] = b \varphi a$, 所以式(28.34)在 $a \leq b$ 时是显然成立的; 但当 $b < a$ 时, 有 $b \varphi b = 1 = b \varphi a$ 。

下面的定理将证明 ξ_I 和 ξ_D 是 A, B 和 Q 的高度的简单函数。

定理 28.3 设 $hgh(A), hgh(B)$ 和 $hgh(Q)$ 分别是 A, B 和 Q 的高度(前面已经学过 $hgh(A) = \sup_{x \in U} \mu_A(x)$), 则有

$$\xi_I = hgh(B) \varphi hgh(A) \quad (28.35)$$

$$\xi_D = hgh(B) \varphi hgh(Q) \quad (28.36)$$

证明 因为 $A \circ (A \varphi B) \subseteq B$ (由引理 28.3), 所以, 由式(28.31)和式(28.29), 可得

$$\begin{aligned} \xi_I &= \inf_{y \in V} \{ \mu_B(y) \varphi [\mu_{A \circ (A \varphi B)}(y)] \} \\ &= \inf_{y \in V} \{ \mu_B(y) \varphi [\sup_{x \in U} (\mu_A(x) \star (\mu_A(x) \varphi \mu_B(y)))] \} \end{aligned} \quad (28.37)$$

由 φ -算子关于第二个元素的单调性, 可得

$$\xi_I = \inf_{y \in V} \sup_{x \in U} \{ \mu_B(y) \varphi [\mu_A(x) \star (\mu_A(x) \varphi \mu_B(y))] \} \quad (28.38)$$

再由式(28.34), 可得

$$\begin{aligned} \xi_I &= \inf_{y \in V} \sup_{x \in U} [\mu_B(y) \varphi \mu_A(x)] \\ &= \inf_{y \in V} [\mu_B(y) \varphi \sup_{x \in U} \mu_A(x)] \end{aligned} \quad (28.39)$$

由 φ -算子关于第一个元素的单调性可知, 当 $\mu_B(y)$ 取它的最大值时, 即可得到式(28.29)中的 \inf , 于是有

$$\xi_I = \sup_{y \in V} [\mu_B(y)] \varphi \sup_{x \in U} [\mu_A(x)] \quad (28.40)$$

式(28.36)的证明留作习题。

现在回到例 28.1, 研究为什么它不存在精确解。对于式(28.19) ~ 式(28.21)中的 A, B 和 Q , 有 $hgh(A) = 0.4$, $hgh(B) = 0.5$ 和 $hgh(Q) = 0.3$, 由此可得 $\xi_I = 0.5 \varphi 0.4 \neq 1$ 和 $\xi_D = 0.5 \varphi 0.3 \neq 1$ 。如果不存在精确解, 那么所能做的就只有去寻找近似解了, 这是下节要讨论的内容。

28.4 近似解——神经网络算法

由定理 28.3 和 φ -算子的定义可知, 如果 B 的高度大于 A 或 Q 的高度, 则不存在精确解。在这种情况下, 确定 A 或 Q 的一种自然方法是使下式最小化

$$d_2(B, A \circ Q) = \int_V \{ \mu_B(y) - \sup_{x \in U} [\mu_A(x) \star \mu_Q(x, y)] \}^2 dy \quad (28.41)$$

上式给出的是最小二乘法的解。这里,只研究“模糊反卷积问题”。“模糊识别问题”可用类似的方法进行求解。为了使得式(28.41)的最小化易于处理,首先要做两件事:(i)将 $\mu_A(x)$ 参数化为高斯函数(也可以用三角形或梯形函数)

$$\mu_A(x) = \exp\left(-\frac{1}{2} \frac{\|x - \bar{x}\|^2}{\sigma^2}\right) \quad (28.42)$$

式中, \bar{x} 和 σ 是自由参数;(ii)分别在 U 和 V 的整个区域内采样 x_i 和 y_j ($i=1,2,\dots,N$; $j=1,2,\dots,M$),则式(28.41)可由下式逼近

$$\sum_{j=1}^M \{ \mu_B(y_j) - \max_{1 \leq i \leq N} [\mu_A(x_i) \star \mu_Q(x_i, y_j)] \}^2 \quad (28.43)$$

因此,问题变成确定参数 \bar{x} 和 σ ,使式(28.43)达到最小值。

现在就设计一个神经网络算法来解决这个问题。

首先,用 $\mu_A(x)\mu_Q(x,y)$ 表示图(28.1)中的网络,其中

$$a = \exp\left[-\frac{1}{2} \left(\frac{\|x - \bar{x}\|}{\sigma}\right)^2\right], b = \mu_Q(x, y), z = ab.$$

当 $y = y_j$ 不变时,对任意的 $x = x_1, x_2, \dots, x_N$ 计算 z_1, z_2, \dots, z_N ,并确定 $z_{i^*} = \max[z_1, z_2, \dots, z_N]$ 。显然 z_{i^*} 等于式(28.43)中的 $\max_{1 \leq i \leq N} [\mu_A(x_i)\mu_Q(x_i, y_j)]$ 。然后,用梯度下降法调整参数 \bar{x} 和 σ ,使下面的误差平方最小:

$$e_j^2 = \{ \mu_B(y_j) - \max_{1 \leq i \leq N} [\mu_A(x_i)\mu_Q(x_i, y_j)] \}^2 \quad (28.44)$$

下面给出算法的详细步骤。

求解“模糊反卷积问题”的神经网络算法

步骤1 分别在 U 和 V 的整个区域内采样 x_i 和 y_j ($i=1,2,\dots,N$; $j=1,2,\dots,M$)。如果 U 和 V 有界,则我们通常选取在 U 和 V 上均匀分布的样本。确定 \bar{x} 和 σ 的初始值 $\bar{x}(0)$ 和 $\sigma(0)$;有时,专家能够提供这些初值。如果专家不能够提供,则可在 U 上随机选取 $\bar{x}(0)$ 并合理地选取某个正数作为 σ 的取值,如 $\sigma=1$ 。选定初始 $j=1$ 。

步骤2 令图28.1的网络中的 $y = y_j$ 保持不变,沿着网络对 $x = x_1, x_2, \dots, x_N$ 向前计算相应的 z_1, z_2, \dots, z_N 。寻找 i^* ,使其满足 $z_{i^*} = \max[z_1, z_2, \dots, z_N]$ 。

步骤3 采用梯度下降法调整参数 \bar{x} 和 σ

$$\begin{aligned} \bar{x}(k+1) &= \bar{x}(k) - \frac{\alpha \partial e_j^2}{2 \partial x} \\ &= \bar{x}(k) + \alpha e_j b^* a^* \frac{x_{i^*} - \bar{x}(k)}{(\sigma(k))^2} \end{aligned} \quad (28.45)$$

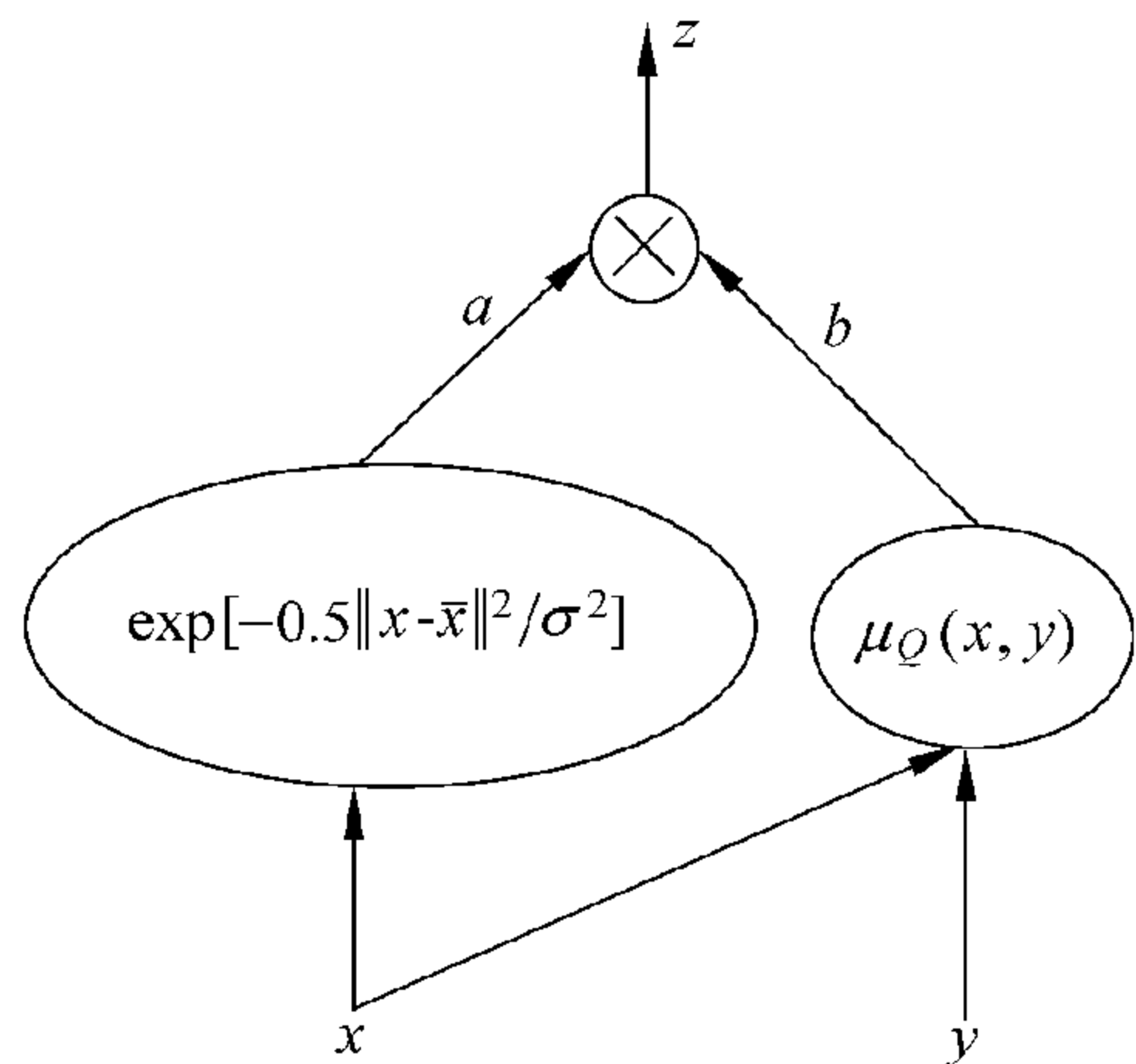


图 28.1 $\mu_A(x)\mu_Q(x,y)$ 的网络示意图

$$\begin{aligned}\sigma(k+1) &= \sigma(k) - \frac{\alpha \partial e_j^2}{2 \partial \sigma} \\ &= \sigma(k) + \alpha e_j b^* a^* \frac{\|x_{i^*} - \bar{x}(k)\|^2}{(\sigma(k))^3}\end{aligned}\quad (28.46)$$

式中,

$$a^* = \exp\left[-\frac{1}{2}\left(\frac{\|x_{i^*} - \bar{x}(k)\|}{\sigma(k)}\right)^2\right]\quad (28.47)$$

$$b^* = \mu_Q(x_{i^*}, y_j)\quad (28.48)$$

e_j^2 由式(28.44)给出, $k=0, 1, 2, \dots$ 。

步骤 4 令 $y = y_{j+1}$, 返回步骤 2, 对 $y = y_1, y_2, \dots, y_M, y_{M+1}, \dots$ ^①, 重复计算, 直至 k 大于 M 时, 有 $|\bar{x}(k+1) - \bar{x}(k)|$ 和 $|\sigma(k+1) - \sigma(k)|$ 的变化小于某个阈值。

学习算法(28.45)和(28.46)是通过下面的复合函数求导规则得到的

$$\frac{\partial e_j^2}{\partial \bar{x}} = 2e_j \frac{\partial e_j}{\partial \bar{x}} = -2e_j \frac{\partial z_{i^*}}{\partial \bar{x}} = -2e_j b^* \frac{\partial a^*}{\partial \bar{x}} = -2e_j b^* a^* \frac{x_{i^*} - \bar{x}(k)}{(\sigma(k))^2}$$

可用同样方法来处理 $\frac{\partial e_j^2}{\partial \sigma}$ 。可以证明, 以上学习算法可以使得总拟合误差 $\sum_{j=1}^M e_j^2$ 在经过 M 步的训练后趋于减小。因为要求 $\sigma > 0$, 所以在运算中需要选择一个小的常数 $\varepsilon > 0$ 。若 $\sigma(k+1) > \varepsilon$, 则根据式(28.46)计算 $\sigma(k+1)$; 若 $\sigma(k+1) \leq \varepsilon$, 则令 $\sigma(k+1) = \varepsilon$ 。这种方法可以保证对所有的 k , 都有 $\sigma(k) \geq \varepsilon$ 。

现在, 应用上面算法来仿真一个例子。

例 28.2 考虑由如下两条模糊 IF-THEN 规则表征的模糊关系

$$\text{规则}^{(1)}: \text{如果 } x \text{ 小, 则 } y \text{ 接近于 } 1\quad (28.49)$$

$$\text{规则}^{(2)}: \text{如果 } x \text{ 大, 则 } y \text{ 接近于 } 0\quad (28.50)$$

规则中的“小”, “大”, “接近于 1”和“接近于 0”都是模糊集, 其隶属度函数分别为 $\mu_{\text{小}}(x) = \frac{1}{1 + e^{10(x-1)}}$, $\mu_{\text{大}}(x) = \frac{1}{1 + e^{10(-x+3)}}$, $\mu_{\text{接近于 } 1}(y) = e^{-10(y-1)^2}$, $\mu_{\text{接近于 } 0}(y) = e^{-10y^2}$ 规则⁽ⁱ⁾ ($i=1, 2$) 为

定义在 $U \times V$ 中的模糊关系, 应用 Mamdani 积含义可得其隶属度函数为

$$\mu_{\text{规则}^{(1)}}(x, y) = \mu_{\text{小}}(x) \mu_{\text{接近于 } 1}(y)\quad (28.51)$$

$$\mu_{\text{规则}^{(2)}}(x, y) = \mu_{\text{大}}(x) \mu_{\text{接近于 } 0}(y)\quad (28.52)$$

最终的模糊关系 Q 为

$$\mu_Q(x, y) = \max[\mu_{\text{规则}^{(1)}}, \mu_{\text{规则}^{(2)}}] = \max\left[\frac{e^{-10(y-1)^2}}{1 + e^{10(x-1)}}, \frac{e^{-10y^2}}{1 + e^{10(-x+3)}}\right]\quad (28.53)$$

① 原文此处为 $y = y_1, y_2, \dots, y_M, y_1, \dots$, 疑为有误一译者。

现在,为简单起见,假设式(28.42)中的 $\sigma \equiv 1$,于是任务就变成了,为不同的 $\mu_B(y)$ 确定 \bar{x} 。考虑 $\mu_B(y)$ 的三种不同情况:(i) $\mu_B(y) = e^{-10(y-0.5)^2}$ (即 y 接近于0.5);(ii) $\mu_B(y) = e^{-10(y-1)^2}$ (即 y 接近于1);(iii) $\mu_B(y) = e^{-10y^2}$ (即 y 接近于0)。这里将区域选定为 $U = [0, 4]$ 和 $V = [0, 1]$ 。分别取 $N = 17$ 和 $M = 9$ 时均匀分布在 U 和 V 上的样点,即 $(x_1, x_2, \dots, x_{17}) = (0, 0.25, 0.5, \dots, 4)$, $(y_1, y_2, y_3, \dots, y_9) = (0, 0.125, 0.25, \dots, 1)$ 。图28.2~图28.4分别给出了对应于 $\mu_B(y)$ 三种不同情况的不同初值条件下的 $\bar{x}(k)$ 的收敛过程,图中横轴表示学习算法中的 k ,纵轴表示 $\bar{x}(k)$ 。从图28.2~图28.4中可以看出,在三种情况下收敛过程都大约需要10步左右。

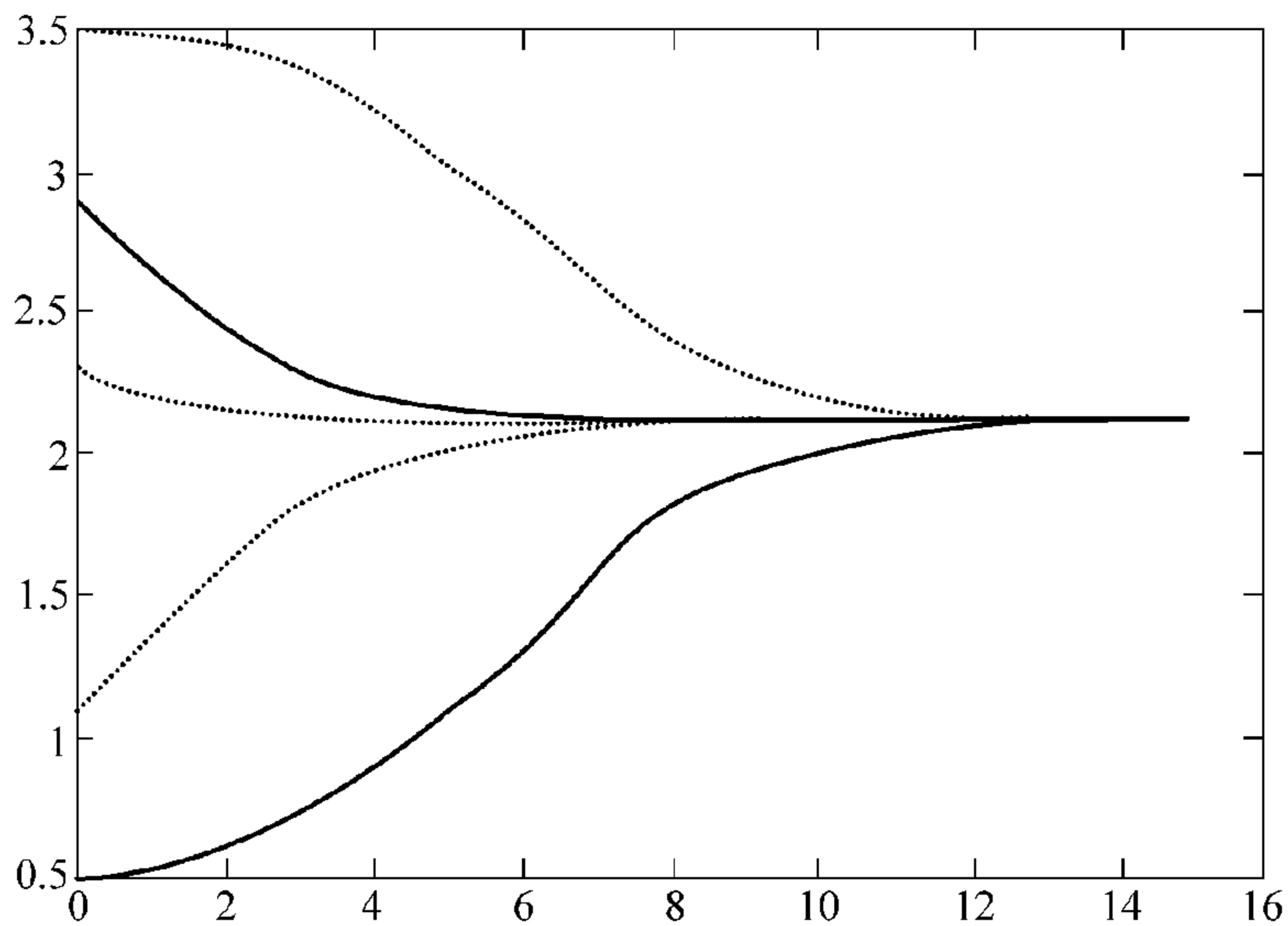
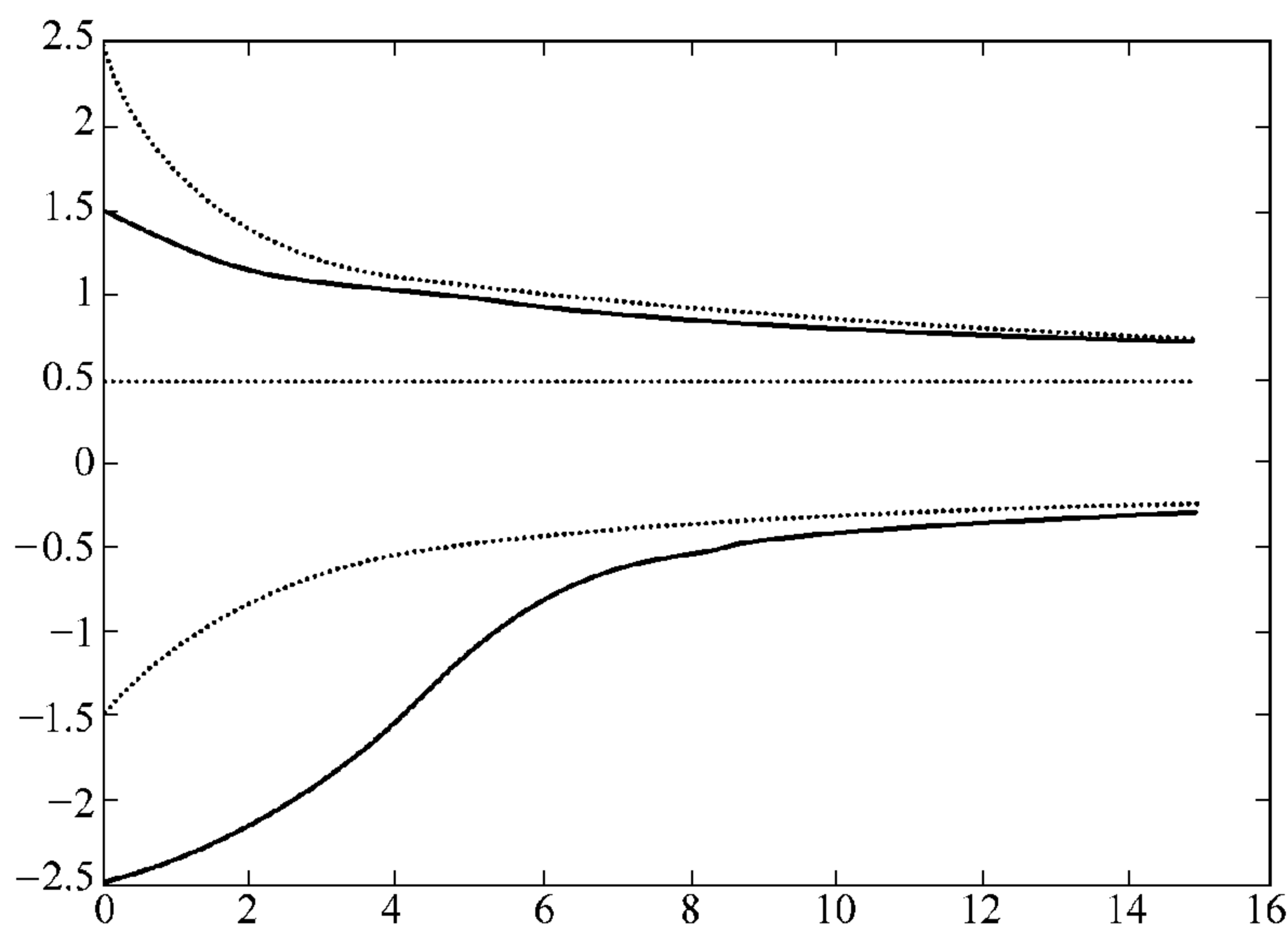
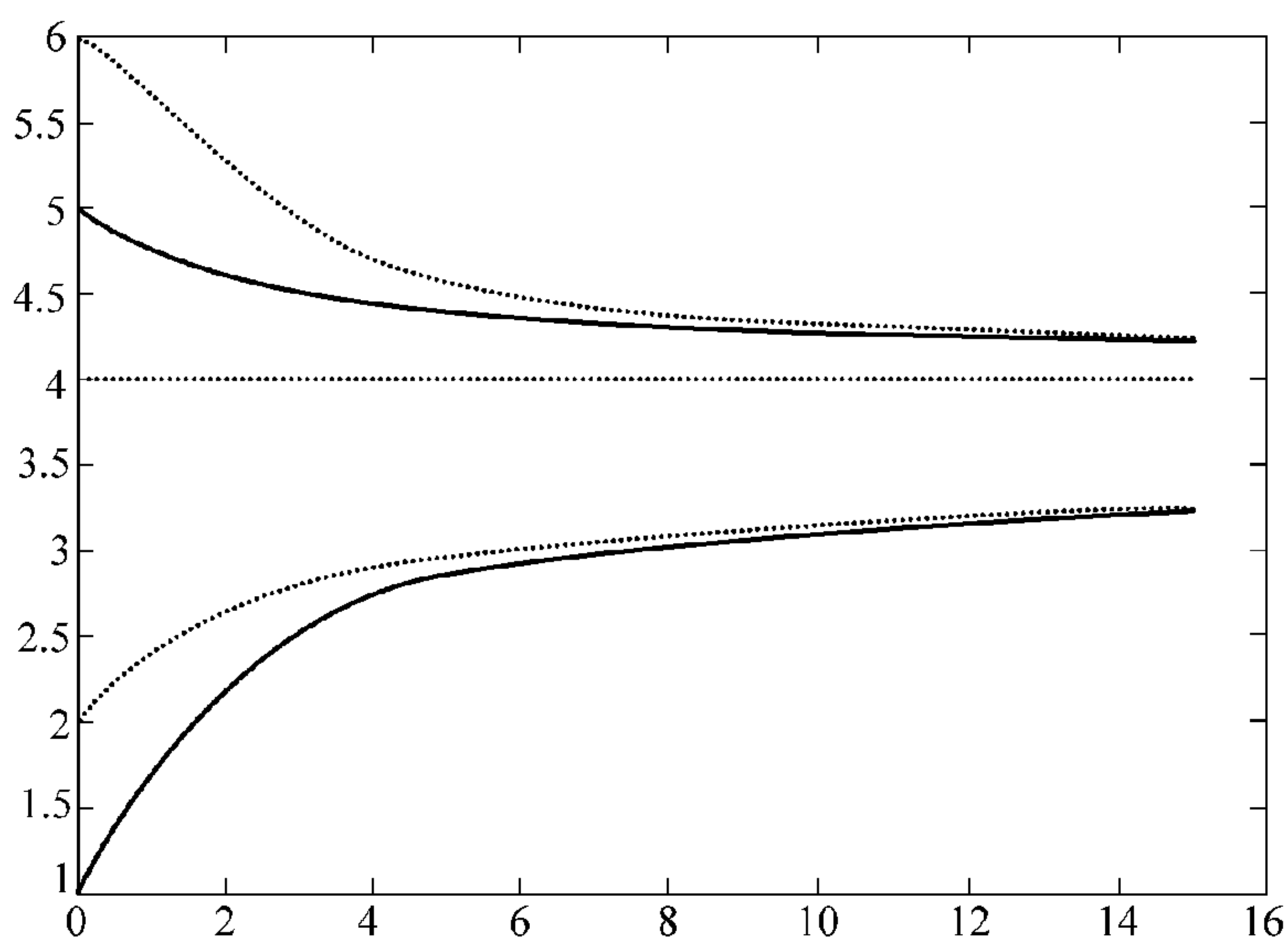


图 28.2 $\mu_B(y) = e^{-10(y-0.5)^2}$ 时 $\bar{x}(k)$ 的收敛过程

由图28.2可以看出,在这个例子中,从不同的初值 $\bar{x}(0)$ 出发, $\bar{x}(k)$ 都会收敛到相同的值,即接近于2。这和人们的直觉是一致的,因为由式(28.49)和式(28.50)可以想像出,当 y 接近于0.5时,即当 y 在0和1的中点时, x 应为“小”和“大”的中点,通过观察隶属度函数 $\mu_{\text{小}}$ 和 $\mu_{\text{大}}$ 可知, x 应该在2附近。图28.3和图28.4说明了,在“ y 接近于1”和“ y 接近于0”的情况下,从不同的初值条件出发, $\bar{x}(k)$ 会收敛到一个值域,而不是一个点。这也符合人们的直觉,因为由式(28.49)可以看出,若“ y 接近于1”,则 x 应该“小”,即图28.3中的-0.3和0.8之间。类似地,由式(28.50)可以看出,如果“ y 接近于0”,则 x 应该“大”,即图28.4中的3.2和4.3之间。总之,三种情况下的解都符合人们的直觉。

28.5 本章小结及导读

本章主要阐述了以下内容:

图 28.3 $\mu_y(y) = e^{-10(y-1)^2}$ 时 $\bar{x}(k)$ 的收敛过程图 28.4 $\mu_y(y) = e^{-10y^2}$ 时 $\bar{x}(k)$ 的收敛过程

- “模糊反卷积和模糊识别问题”的定义；
- 如何得到这些问题的特解；
- 可解度的定义和计算；
- 如何把求解模糊关系方程问题转化为神经网络问题，以及如何为计算近似解设计神经网络学习算法。

Nola, Sessa, Pedrycz 和 Sanchez [1989] 是一部关于模糊关系方程的专著, Pedrycz [1993] 一书也涉及了模糊关系方程的内容。模糊关系方程方面比较有指导意义的文献可见 Pedrycz [1991]。本章中的神经网络算法摘自 wang [1993b]。

28.6 习题

习题 28.1 求解下面的“模糊反卷积问题”(其中合成为 max-min 合成)。

$$(a) \quad A \circ \begin{bmatrix} 0.9 & 0.8 & 0.7 \\ 0.8 & 0.9 & 0.8 \\ 0.7 & 0.8 & 0.9 \end{bmatrix} = [0.6 \quad 0.5 \quad 0.4] \quad (28.54)$$

$$(b) \quad A \circ \begin{bmatrix} 0.5 & 0.7 & 0 & 0.2 \\ 0.4 & 0.6 & 1 & 0 \\ 0.2 & 0.4 & 0.5 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{bmatrix} = [0.5 \quad 0.5 \quad 0.4 \quad 0.2] \quad (28.55)$$

$$(c) \quad A \circ \begin{bmatrix} 0.5 & 0.4 & 0.6 & 0.7 \\ 0.2 & 0 & 0.6 & 0.8 \\ 0.2 & 0.4 & 0.6 & 0.7 \\ 0 & 0.3 & 0.3 & 1 \end{bmatrix} = [0.5 \quad 0.5 \quad 0.4 \quad 0.2] \quad (28.56)$$

$$(d) \quad A \circ \begin{bmatrix} 0.5 & 0 & 0.3 & 0 \\ 0.4 & 1 & 0.3 & 0 \\ 0 & 0.1 & 1 & 0.1 \\ 0.4 & 0.3 & 0.3 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.3 & 0.3 & 0.1 \\ 0.5 & 0.4 & 0.4 & 0.2 \end{bmatrix} \quad (28.57)$$

习题 28.2 求解下面的“模糊识别问题”(其中合成为 max-min 合成)。

$$(a) \quad [1 \quad 0.3 \quad 0.5] \circ Q = [0.6 \quad 0.5 \quad 0.4] \quad (28.58)$$

$$(b) \quad [0.1 \quad 0.3 \quad 0.5 \quad 0.7] \circ Q = [0.6 \quad 0.5 \quad 0.4 \quad 0.3] \quad (28.59)$$

$$(c) \quad [0.6 \quad 0.4 \quad 0.6 \quad 0.8] \circ Q = [0.1 \quad 0.5 \quad 0.4 \quad 0.2] \quad (28.60)$$

$$(d) \quad \begin{bmatrix} 0.6 & 0.2 & 0.3 & 0.9 \\ 0.1 & 0.3 & 0.5 & 0.7 \end{bmatrix} \circ Q = \begin{bmatrix} 0.1 & 0.5 & 0.4 & 0.2 \\ 0.2 & 0.2 & 0.4 & 0.4 \end{bmatrix} \quad (28.61)$$

习题 28.3 证明定理 28.3 中的式(28.36)。

习题 28.4 设 P, Q 和 R 分别是 $U \times V, V \times W$ 和 $U \times W$ 中的模糊关系, $F(U \times V)$ 是 $U \times V$ 中所有模糊关系的集合, 定义模糊关系方程 $P \circ Q = R$ 的可解度为

$$\delta = \sup_{P \in F(U \times V)} eq_t(P \circ Q, R) \quad (28.62)$$

证明

$$\delta \leq \inf_{Z \in W} \left[\sup_{x \in U} \mu_R(x, z) \varphi \sup_{y \in V} \mu_Q(y, z) \right] \quad (28.63)$$

习题 28.5 设计一个神经网络算法来求解“模糊识别问题”。

29.1 模糊数和分解定理

在模糊 IF-THEN 规则中,原模糊命题“ x_i 为 A_i^l ”,通常是对实数域 R 中的数的一种不精确描述,如“接近于 5”,“大约为 7”等等。因为在许多应用中经常会出现定义在 R 上的模糊集,所以有必要对其进行深入研究。粗略地讲,模糊数是 R 中的模糊集,但为使模糊算术有运算意义,须增加一些约束条件。具体地,有如下定义:

定义 29.1 设 A 是 R 中的模糊集,若(i) A 是标准的,(ii) A 是凸集,(iii) A 的支撑集有界,(iv) A 的所有 α -截集都为 R 中的闭区间,则称 A 为一个模糊数(fuzzy number)。

要求模糊数的标准性是因为,从现有概念上讲,实数 r 本身是完全满足“接近于 r 的实数”的,故有 $\mu_A(r) = 1$ 。要求模糊数的凸性和有界性条件,是为了保证所定义的运算对模糊数有意义。

在实际中经常用到两种特殊类型的模糊数,三角形模糊数和梯形模糊数。三角形模糊数 A 是 R 中以三角形为隶属度函数的模糊集:

$$\mu_A(x) = \mu_A(x; a, b, c) = \begin{cases} (x-a)/(b-a) & a \leq x < b \\ (c-x)/(c-b) & b \leq x \leq c \\ 0 & x > c \text{ 或 } x < a \end{cases} \quad (29.1)$$

类似地,若 R 中的模糊集 A 具有下面的梯形隶属度函数,则称 A 为梯形模糊数:

$$\mu_A(x) = \mu_A(x; a, b, c, d) = \begin{cases} (x-a)/(b-a) & a \leq x < b \\ 1 & b \leq x \leq c \\ (d-x)/(d-c) & c < x \leq d \\ 0 & x > d \text{ 或 } x < a \end{cases} \quad (29.2)$$

容易验证三角形模糊数和梯形模糊数满足定义 29.1 中的四个条件。

研究模糊数算术运算的一种有效方法就是用模糊数的 α -截集。回顾一下, R 中模糊集 A 的 α -截集的定义为

$$A_\alpha = \{x \in R \mid \mu_A(x) \geq \alpha\} \quad (29.3)$$

现在就可以证明,模糊集是由其 α -截集唯一确定的。为了便于精确表示,现定义 R 中模糊集 \tilde{A}_α 的隶属度函数为

$$\mu_{A_\alpha}(x) = \alpha I_{A_\alpha}(x) \quad (29.4)$$

式中的 $I_{A_\alpha}(x)$ 是清晰集合 A_α 的指示函数, 即若 $x \in A_\alpha$, 则 $I_{A_\alpha}(x) = 1$; 若 $x \in R - A_\alpha$, 则 $I_{A_\alpha}(x) = 0$ 。

下面介绍著名的分解定理, 并用其说明模糊集 A 等价于模糊集 $\tilde{A}_\alpha (\alpha \in [0, 1])$ 的并集。

定理 29.1 (分解定理) 令 A 和 \tilde{A}_α 为 R 上的模糊集合, 其中 \tilde{A}_α 的定义见式(29.4), 则

$$A = \bigcup_{\alpha \in [0, 1]} \tilde{A}_\alpha \quad (29.5)$$

式中, \bigcup 表示标准模糊并集(即, 取 $\alpha \in [0, 1]$ 时的上确界)。

证明 设 x 为 R 中任意点, $a = \mu_A(x)$, 则有

$$\begin{aligned} \mu_{\bigcup_{\alpha \in [0, 1]} A_\alpha}(x) &= \sup_{\alpha \in [0, 1]} \mu_{A_\alpha}(x) \\ &= \max\left[\sup_{\alpha \in [0, a]} \mu_{A_\alpha}(x), \sup_{\alpha \in (a, 1)} \mu_{A_\alpha}(x)\right] \end{aligned} \quad (29.6)$$

对所有 $\alpha \in (a, 1]$, 都有 $\mu_{A_\alpha}(x) = a < \alpha$, 于是有 $x \notin A_\alpha$, 这表明 $\mu_{A_\alpha}(x) = 0$ 。对所有 $\alpha \in [0, a]$, 都有 $\mu_{A_\alpha}(x) = a \geq \alpha$, 于是有 $x \in A_\alpha$ 。再由式(29.4), 可得 $\mu_{A_\alpha}(x) = \alpha$ 。因此, 可得

$$\mu_{\bigcup_{\alpha \in [0, 1]} A_\alpha}(x) = \sup_{\alpha \in [0, a]} \alpha = a = \mu_A(x) \quad (29.7)$$

因为 x 是任意点, 所以, 由式(29.7)成立, 可推出式(29.5)成立。

由定理 29.1 可以看出, 如果能求出一个模糊集在 $\alpha \in [0, 1]$ 时的 α -截集, 就可以确定该模糊集。因此, 确定一个模糊集就等价于确定它在 $\alpha \in [0, 1]$ 时的 α -截集。在下面两节中, 将使用 α -截集来定义模糊数的算术运算。

29.2 模糊数的加和减

29.2.1 α -截集法

设 A, B 是两个模糊数, $A_\alpha = [a_\alpha^-, a_\alpha^+]$ 和 $B_\alpha = [b_\alpha^-, b_\alpha^+]$ 分别是 A 和 B 的 α -截集。那么, A 和 B 的加, 记作 $A + B$, 也是一个模糊数, 定义其在 $\alpha \in [0, 1]$ 时的 α -截集为

$$(A + B)_\alpha = [a_\alpha^- + b_\alpha^-, a_\alpha^+ + b_\alpha^+] \quad (29.8)$$

因为模糊集可以完全由其 α -截集确定(定理 29.1), 所以模糊数的加 $A + B$ 也可以很好地由其 α -截集(29.8)来确定。

类似地, A 和 B 的减, 记作 $A - B$, 也是一个模糊数, 定义其在 $\alpha \in [0, 1]$ 时的 α -截集为

$$(A - B)_\alpha = [\min(a_\alpha^- - b_\alpha^-, a_\alpha^+ - b_\alpha^+), \max(a_\alpha^- - b_\alpha^-, a_\alpha^+ - b_\alpha^+)] \quad (29.9)$$

例 29.1 计算三角形模糊数 A 和 B 的加和减, 其中 A 和 B 的隶属度函数为

$$\mu_A(x) = \mu_A(x; -1, 0, 1) \quad (29.10)$$

$$\mu_B(x) = \mu_B(x; -1, 1, 3) \quad (29.11)$$

对某一给定的 $\alpha \in [0, 1]$, A 和 B 的 α -截集分别为

$$A_\alpha = [\alpha - 1, 1 - \alpha], B_\alpha = [2\alpha - 1, 3 - 2\alpha] \quad (29.12)$$

于是可得

$$(A + B)_\alpha = [3\alpha - 2, 4 - 3\alpha] \quad (29.13)$$

因为 $3\alpha - 2$ 和 $4 - 3\alpha$ 是关于 α 的线性函数, $A + B$ 又是三角形模糊数, 所以, 令式(29.13)中的 $\alpha = 0$ 和 $\alpha = 1$, 可得

$$\mu_{A+B}(x) = \mu_{A+B}(x; -2, 1, 4) \quad (29.14)$$

其图形见图 29.1。类似地, 可得

$$(A - B)_\alpha = [\alpha - 2, -\alpha] \quad (29.15)$$

和

$$\mu_{A-B}(x) = \mu_{A-B}(x; -2, -1, 0) \quad (29.16)$$

其图形见图 29.1。

另一种计算模糊数的加和减的方法是使用扩展原理。

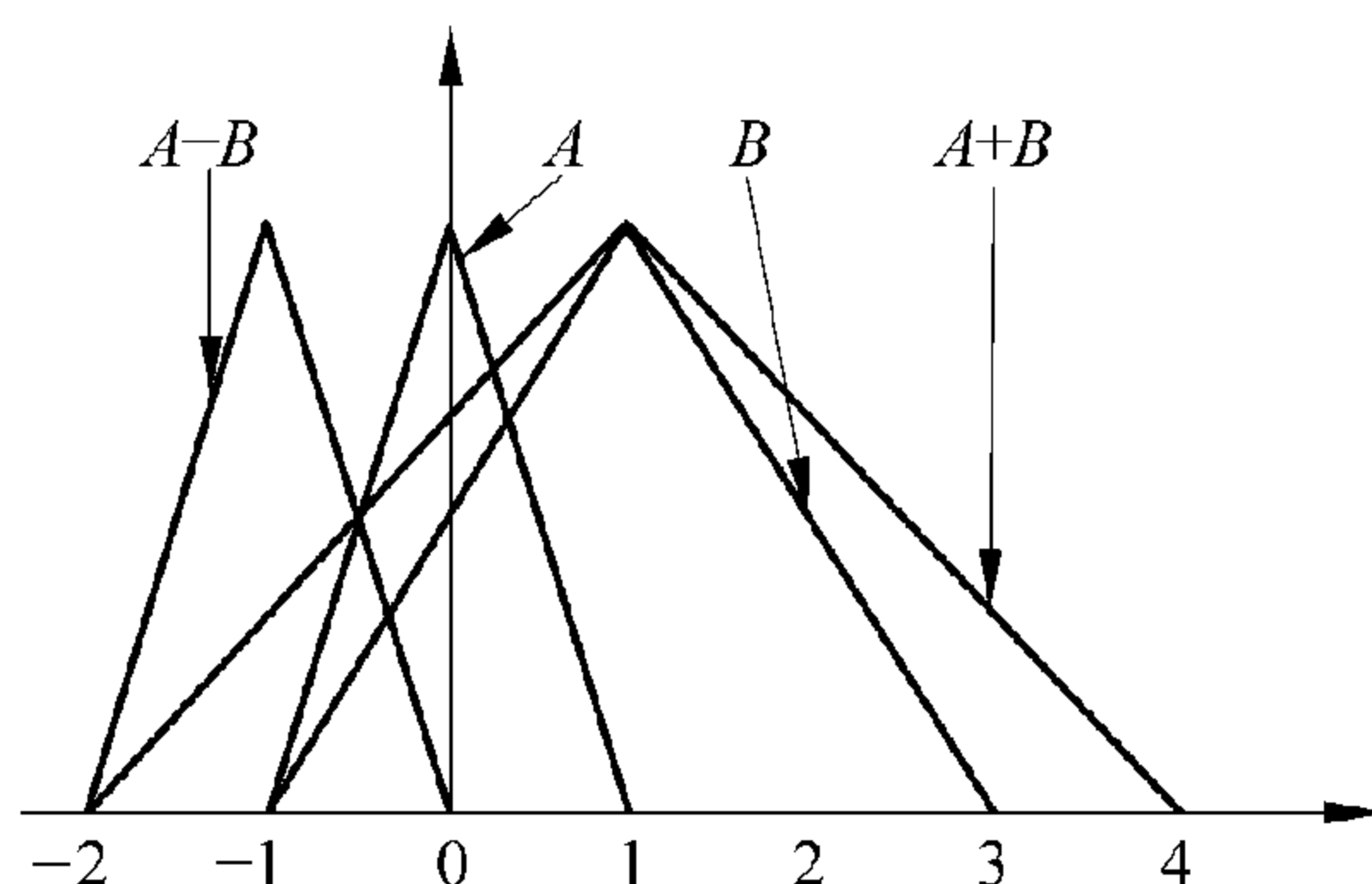


图 29.1 例 29.1 中两个模糊数的加和减

29.2.2 扩展原理法

使用扩展原理(第 4 章), 可以定义模糊数 A 和 B 的加为

$$\mu_{A+B}(z) = \sup_{x+y=z} \min[\mu_A(x), \mu_B(y)] \quad (29.17)$$

式中, $z \in R$, $\sup_{x+y=z}$ 表示对所有满足 $x + y = z$ 的 $x, y \in R$ 取上确界。类似地, 定义

$$\mu_{A-B}(z) = \sup_{x-y=z} \min[\mu_A(x), \mu_B(y)] \quad (29.18)$$

下面证明该定义等价于 α -截集法。

定理 29.2 扩展原理方法(29.17)和(29.18)等价于 α -截集法(29.8)和(29.9), 即它们给出的模糊数 $A + B$ 和 $A - B$ 是相同的。

证明 首先要注意, 由式(29.17)和式(29.18)定义的模糊集 $A + B$ 和 $A - B$ 是模糊数(其证明留作习题)。因此, 根据定理 29.1, 只需证明由式(29.17)和式(29.18)定义的 $A + B$ 和 $A - B$ 的 α -截集分别是由式(29.8)和式(29.9)给出的。由式(29.17), 可以看出, 对任意 $x \in A_\alpha = [a_\alpha^-, a_\alpha^+]$ 和 $y \in B_\alpha = [b_\alpha^-, b_\alpha^+]$, 当 $z = x + y$ 时, 有 $\mu_{A+B}(z) \geq \alpha$; 因为对任意 $z \in [a_\alpha^- + b_\alpha^-, a_\alpha^+ + b_\alpha^+]$, 都能找到满足 $z = x + y$ 的 $x \in [a_\alpha^-, a_\alpha^+]$ 和 $y \in [b_\alpha^-, b_\alpha^+]$, 所以有 $[a_\alpha^- + b_\alpha^-, a_\alpha^+ + b_\alpha^+] \subseteq (A + B)_\alpha$ 。反过来, 如果 $z \in R - [a_\alpha^- + b_\alpha^-, a_\alpha^+ + b_\alpha^+]$, 则对任意 $z = x + y$, 根据式(29.17), 有 $x \notin [a_\alpha^-, a_\alpha^+]$ 或 $y \notin [b_\alpha^-, b_\alpha^+]$ 满足 $\mu_{A+B}(z) = 0 < \alpha$ 。这说明对任意满足 $\mu_{A+B}(z) \geq \alpha$ 的 z , $z \in [a_\alpha^- + b_\alpha^-, a_\alpha^+ + b_\alpha^+]$ 是必然成立的, 即 $(A + B)_\alpha \subseteq [a_\alpha^- + b_\alpha^-, a_\alpha^+ + b_\alpha^+]$, 因此, 式(29.8)成立。同理可证, 式(29.9)成立。

由于实数 a 是一个特殊的模糊数(当 $x = a$ 时,其隶属度函数 $\mu_a = 1$; 当 $x \neq a$ 时,其隶属度函数 $\mu_a(x) = 0$),所以,根据式(29.17),可知, $A + a$ 也是一个模糊数,其隶属度函数为

$$\mu_{A+a}(z) = \mu_A(z - a) \quad (29.19)$$

即, $\mu_{A+a}(z)$ 是将 $\mu_A(z)$ 平移一个距离 a 得到的。类似地, $A - a$ 也是一个模糊数,其隶属度函数为

$$\mu_{A-a}(z) = \mu_A(z + a) \quad (29.20)$$

可以把 0 看做一个特殊的模糊数,由式(29.18),有

$$\mu_{-B}(z) = \mu_{0-B}(z) = \mu_B(-z) \quad (29.21)$$

最后,给出一个计算两个有限模糊集的加和减的例子。

例 29.2 计算下面两个模糊集的加和减:

$$A = 0.1/1 + 0.5/2 + 1/3 + 0.7/4 + 0.3/5 \quad (29.22)$$

$$B = 0.2/1 + 1/2 + 0.5/3 \quad (29.23)$$

由式(29.17)和式(29.18),可得

$$A + B = 0.1/2 + 0.2/3 + 0.5/4 + 1/5 + 0.7/6 + 0.5/7 + 0.3/8 \quad (29.24)$$

$$A - B = 0.1/-2 + 0.5/-1 + 0.5/0 + 1/1 + 0.7/6 + 0.3/3 + 0.2/4 \quad (29.25)$$

29.3 模糊数的乘和除

29.3.1 α -截集法

在传统的区间分析里,区间 $[a, b]$ 乘以 $[c, d]$ 的定义为

$$[a, b] \cdot [c, d] = [\min(ac, ad, bc, bd), \max(ac, ad, bc, bd)] \quad (29.26)$$

如果 $0 \notin [c, d]$, 则区间 $[a, b]$ 和 $[c, d]$ 的除可定义为

$$\begin{aligned} [a, b]/[c, d] &= [a, b] \cdot \left[\frac{1}{c}, \frac{1}{d} \right] \\ &= [\min(a/c, a/d, b/c, b/d), \max(a/c, a/d, b/c, b/d)] \end{aligned} \quad (29.27)$$

下面用模糊数的 α -截集和式(29.26)和式(29.27)来定义模糊数的乘和除。

设 A, B 是两个模糊数, $A_\alpha = [a_\alpha^-, a_\alpha^+]$ 和 $B_\alpha = [b_\alpha^-, b_\alpha^+]$ 是 A 和 B 的 α -截集。则 A 和 B 的乘,记作 $A \cdot B$, 是一个模糊数,定义其在 $\alpha \in [0, 1]$ 时的 α -截集为

$$(A \cdot B)_\alpha = [\min \{a_\alpha^- b_\alpha^-, a_\alpha^- b_\alpha^+, a_\alpha^+ b_\alpha^-, a_\alpha^+ b_\alpha^+\}, \max \{a_\alpha^- b_\alpha^-, a_\alpha^- b_\alpha^+, a_\alpha^+ b_\alpha^-, a_\alpha^+ b_\alpha^+\}] \quad (29.28)$$

如果对任意 $\alpha \in [0, 1]$, 都有 $0 \in [b_\alpha^-, b_\alpha^+]$, 则 A 和 B 的除,记为 A/B , 也是一个模糊数,定义其在 $\alpha \in [0, 1]$ 时的 α -截集为

$$(A/B)_\alpha = [\min \{a_\alpha^-/b_\alpha^-, a_\alpha^-/b_\alpha^+, a_\alpha^+/b_\alpha^-, a_\alpha^+/b_\alpha^+\}, \max \{a_\alpha^-/b_\alpha^-,$$

$$a_{\alpha}^{-}/b_{\alpha}^{+}, a_{\alpha}^{+}/b_{\alpha}^{-}, a_{\alpha}^{+}/b_{\alpha}^{+} \}] \quad (29.29)$$

例 29.3 考虑例 29.1 中的模糊数 A 和由下式定义的模糊数 B

$$\mu_B(x) = \mu_B(x; 0.5, 1, 3) \quad (29.30)$$

现在, 计算 $A \cdot B$ 和 A/B 。首先, 由 α -截集 $A_{\alpha} = [a_{\alpha}^{-}, a_{\alpha}^{+}] = [\alpha - 1, 1 + \alpha]$ 和 $B_{\alpha} = [b_{\alpha}^{-}, b_{\alpha}^{+}] = [(\alpha + 1)/2, 3 - 2\alpha]$, 可得

$$\min \{a_{\alpha}^{-}b_{\alpha}^{-}, a_{\alpha}^{-}b_{\alpha}^{+}, a_{\alpha}^{+}b_{\alpha}^{-}, a_{\alpha}^{+}b_{\alpha}^{+}\} = a_{\alpha}^{-}b_{\alpha}^{+} = -2\alpha^2 + 5\alpha - 3 \quad (29.31)$$

$$\max \{a_{\alpha}^{-}b_{\alpha}^{-}, a_{\alpha}^{-}b_{\alpha}^{+}, a_{\alpha}^{+}b_{\alpha}^{-}, a_{\alpha}^{+}b_{\alpha}^{+}\} = a_{\alpha}^{+}b_{\alpha}^{+} = 2\alpha^2 - 5\alpha + 3 \quad (29.32)$$

且

$$(A \cdot B)_{\alpha} = [-2\alpha^2 + 5\alpha - 3, 2\alpha^2 - 5\alpha + 3] \quad (29.33)$$

再由式(29.33), 可得

$$\mu_{A \cdot B}(x) = \frac{5 - (1 + 8|x|)^{1/2}}{4} \quad (29.34)$$

其图形见图 29.2。类似地, 可得

$$\min \{a_{\alpha}^{-}/b_{\alpha}^{-}, a_{\alpha}^{-}/b_{\alpha}^{+}, a_{\alpha}^{+}/b_{\alpha}^{-}, a_{\alpha}^{+}/b_{\alpha}^{+}\} = a_{\alpha}^{-}/b_{\alpha}^{+} = (\alpha - 1)/(3 - 2\alpha) \quad (29.35)$$

$$\max \{a_{\alpha}^{-}/b_{\alpha}^{-}, a_{\alpha}^{-}/b_{\alpha}^{+}, a_{\alpha}^{+}/b_{\alpha}^{-}, a_{\alpha}^{+}/b_{\alpha}^{+}\} = a_{\alpha}^{+}/b_{\alpha}^{-} = 2(1 - \alpha)/(1 + \alpha) \quad (29.36)$$

及

$$(A/B)_{\alpha} = [(\alpha - 1)/(3 - 2\alpha), 2(1 - \alpha)/(1 + \alpha)] \quad (29.37)$$

由式(29.37)可得

$$\mu_{A/B}(x) = \begin{cases} (3x + 1)/(2x + 1) & x < 0 \\ (2 - x)/(2 + x) & x \geq 0 \end{cases} \quad (29.38)$$

其图形见图 29.2。

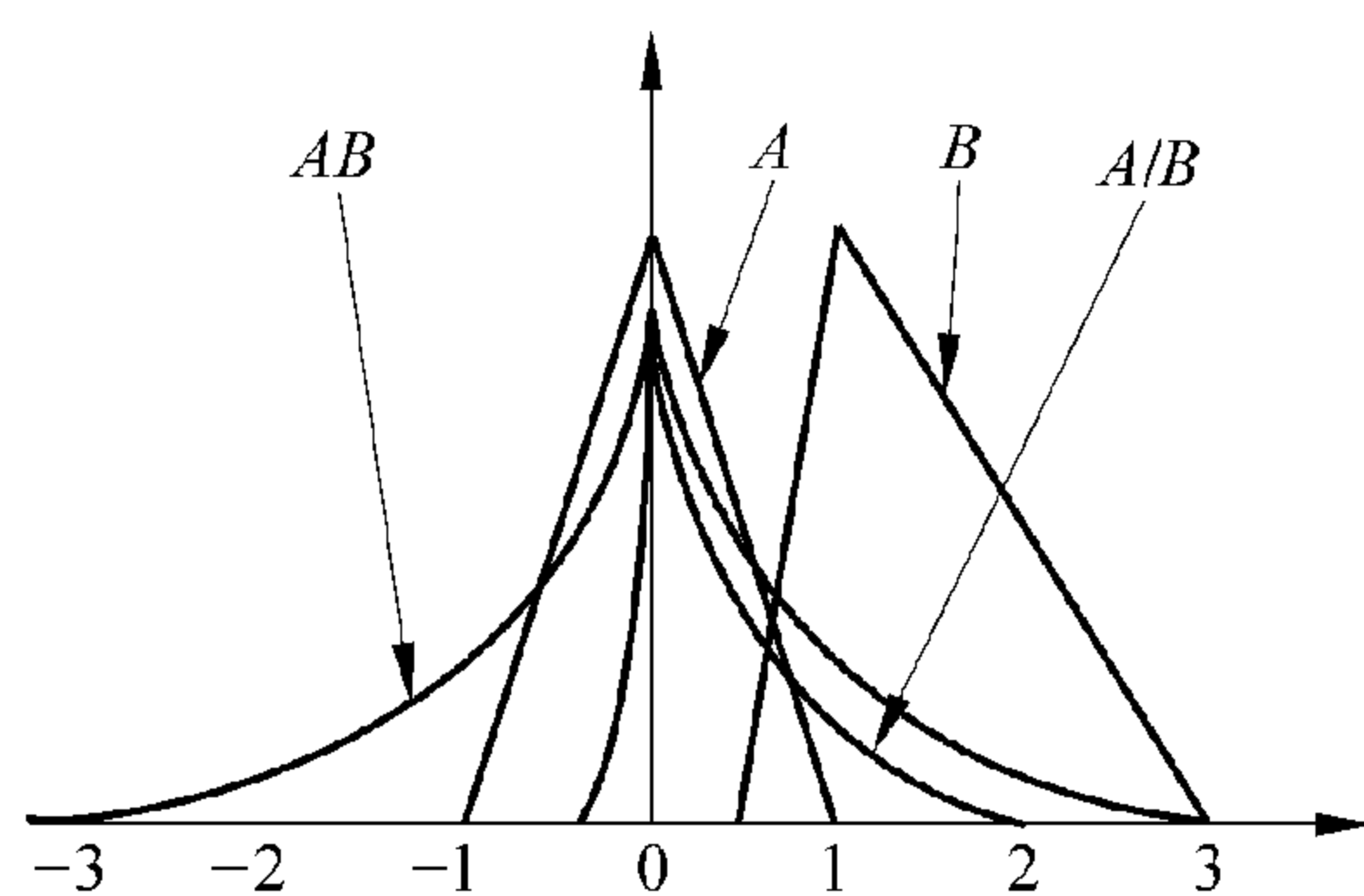


图 29.2 例 29.3 中两个模糊数的乘和除

29.3.2 扩展原理法

和模糊数的加和减一样, 可以用扩展原理法定义 $A \cdot B$ 和 A/B 。具体地, 模糊数 $A \cdot B$ 和 A/B 可以分别定义为

$$\mu_{A \cdot B}(z) = \sup_{xy=z} \min[\mu_A(x), \mu_B(y)] \quad (29.39)$$

和

$$\mu_{A/B}(z) = \sup_{x/y=z} \min[\mu_A(x), \mu_B(y)] \quad (29.40)$$

可以证明, 式(29.39) ~ 式(29.40)分别等价于由式(29.28) ~ 式(29.29)所定义的模糊数。

若 $B = b$ 是一个清晰数, 则由式(29.39), 可得

$$\mu_{bA}(z) = \mu_A(z/b) \quad (29.41)$$

上式给出了实数与模糊数的乘法运算。

29.4 模糊方程

设 A, B 和 X 是模糊数, 求解模糊方程

$$A + X = B \quad (29.42)$$

和

$$A \cdot X = B \quad (29.43)$$

式中, A, B 已知, X 未知。即, 求解模糊方程就是在给定模糊数 A 和 B 的条件下, 求解模糊数 X , 使其满足方程(29.42)和(29.43)。

首先, 研究如何求解方程(29.42)。设 $A_\alpha = [a_\alpha^-, a_\alpha^+]$, $B_\alpha = [b_\alpha^-, b_\alpha^+]$ 和 $X_\alpha = [x_\alpha^-, x_\alpha^+]$, 根据分解定理, 方程(29.42)成立的充要条件是

$$[a_\alpha^-, a_\alpha^+] + [x_\alpha^-, x_\alpha^+] = [b_\alpha^-, b_\alpha^+] \quad (29.44)$$

于是可得方程的一个可能解为

$$X_\alpha = [x_\alpha^-, x_\alpha^+] = [b_\alpha^- - a_\alpha^-, b_\alpha^+ - a_\alpha^+] \quad (29.45)$$

要使式(29.45)中 X_α 为 α -截集, 则必须满足下面两个条件:

- (i) 对所有 $\alpha \in (0, 1]$ 有 $b_\alpha^- - a_\alpha^- \leq b_\alpha^+ - a_\alpha^+$ 成立;
- (ii) 若 $\alpha \leq \beta$, 则 $b_\alpha^- - a_\alpha^- \leq b_\beta^- - a_\beta^- \leq b_\beta^+ - a_\beta^+ \leq b_\alpha^+ - a_\alpha^+$ 。

条件(i)保证, $[b_\alpha^- - a_\alpha^-, b_\alpha^+ - a_\alpha^+]$ 是一个区间; 条件(ii)保证, $[b_\alpha^- - a_\alpha^-, b_\alpha^+ - a_\alpha^+]$ 是使模糊数成为 α -截集的区间套。因此, 如果条件(i)和(ii)成立, 则由分解定理, 可以得到下面的解:

$$X = \bigcup_{\alpha \in (0, 1]} \tilde{X}_\alpha \quad (29.46)$$

式中, \tilde{X}_α 是由下式定义的模糊集:

$$\mu_{\tilde{X}_\alpha}(x) = \alpha I_{[b_\alpha^- - a_\alpha^-, b_\alpha^+ - a_\alpha^+]}(x) \quad (29.47)$$

例 29.4 设 A, B 是例 29.1 中的三角形模糊数。则由式(29.12), 有 $a_\alpha^- = \alpha - 1$, $a_\alpha^+ = 1 - \alpha$, $b_\alpha^- = 2\alpha - 1$ 和 $b_\alpha^+ = 3 - 2\alpha$ 。进而, 有 $b_\alpha^- - a_\alpha^- = \alpha$ 和 $b_\alpha^+ - a_\alpha^+ = 2 - \alpha$ 。因为当 $\alpha \in (0, 1]$ 时, $\alpha \leq 2 - \alpha$; 当 $\alpha \leq \beta \leq 1$ 时, $\alpha \leq \beta \leq 2 - \beta \leq 2 - \alpha$, 所以, 条件(i)和(ii)都是成立的。因此, 可得

$$X_\alpha = [\alpha, 2 - \alpha] \quad (29.48)$$

由此可知, 所得到的解是三角形模糊数

$$\mu_X(x) = \mu_X(x; 0, 1, 2) \quad (29.49)$$

下一步将研究模糊方程(29.43)。为了简单起见, 假设 A, B 是 R^+ 上的模糊数, 则式

(29.43) 等价于下式

$$[a_{\alpha}^{-}, a_{\alpha}^{+}] \cdot [x_{\alpha}^{-}, x_{\alpha}^{+}] = [b_{\alpha}^{-}, b_{\alpha}^{+}] \quad (\alpha \in (0, 1]) \quad (29.50)$$

若下面的两个条件得到满足

(i) 对每个 $\alpha \in (0, 1]$, 有 $b_{\alpha}^{-}/a_{\alpha}^{-} \leq b_{\alpha}^{+}/a_{\alpha}^{+}$

(ii) 若 $\alpha \leq \beta$, 即有 $b_{\alpha}^{-}/a_{\alpha}^{-} \leq b_{\beta}^{-}/a_{\beta}^{-} \leq b_{\beta}^{+}/a_{\beta}^{+} \leq b_{\alpha}^{+}/a_{\alpha}^{+}$

则由其 α -截集可以得到方程(29.43)的解 X 为

$$X_{\alpha} = [x_{\alpha}^{-}, x_{\alpha}^{+}] = [b_{\alpha}^{-}/a_{\alpha}^{-}, b_{\alpha}^{+}/a_{\alpha}^{+}] \quad (29.51)$$

例 29.5 考虑模糊方程(29.43), 其中, $\mu_A(x) = \mu_A(x; 3, 4, 5)$, $\mu_B(x) = \mu_B(x; 12, 20, 32)$ 。 $A_{\alpha} = [\alpha + 3, 5 - \alpha]$, $B_{\alpha} = [8\alpha + 12, 32 - 12\alpha]$ 。容易证明上面的两个条件都成立, 因此有

$$X_{\alpha} = \left[\frac{8\alpha + 12}{\alpha + 3}, \frac{32 - 12\alpha}{5 - \alpha} \right] \quad (29.52)$$

由此得到的解为

$$\mu_x(x) = \begin{cases} 0 & x \leq 4 \text{ 且 } x \geq 32/5 \\ \frac{12 - 3x}{x - 8} & 4 < x \leq 5 \\ \frac{32 - 5x}{12 - x} & 5 \leq x \leq 32/5 \end{cases} \quad (29.53)$$

29.5 模糊排序

在模糊多属性决策中(Chen 和 Hwang [1991]), 各种方案的最终打分是用模糊数表示的。为了在各种方案中做出清晰的选择, 需要用一种方法来比较模糊数。不过, 对模糊数不能像对 R 中的实数那样用 \leq 进行线性排序, 这直观上容易理解。考虑图 29.3 中的模糊数 A 和 B 。一些人认为 B 大于 A , 因为 B 的中心大于 A 的中心, 而另一些人持相反意见, 因为 B 广泛地分散在很小的数上, 而 A 集中在很大的数上。显然, 问题本身的不良性(即无清楚答案)使其不存在完美的模糊排序方法。现有文献中提出了大量的模糊数排序方法, 每种方法都是在某些场合适用, 而在另一些场合不适用。Chen 和 Hwang [1991] 综述了近 20 种排序方法。本节只考虑两种方法: α -截集法和海明距离法。

首先, 应该弄清楚, 在什么条件下可以明确地对模糊数进行排序。如果 α -截集 $A_{\alpha} = [a_{\alpha}^{-}, a_{\alpha}^{+}]$ 和 $B_{\alpha} = [b_{\alpha}^{-}, b_{\alpha}^{+}]$ 的边界对所有的 $\alpha \in (0, 1]$ 都满足 $a_{\alpha}^{-} \leq b_{\alpha}^{-}$ 和

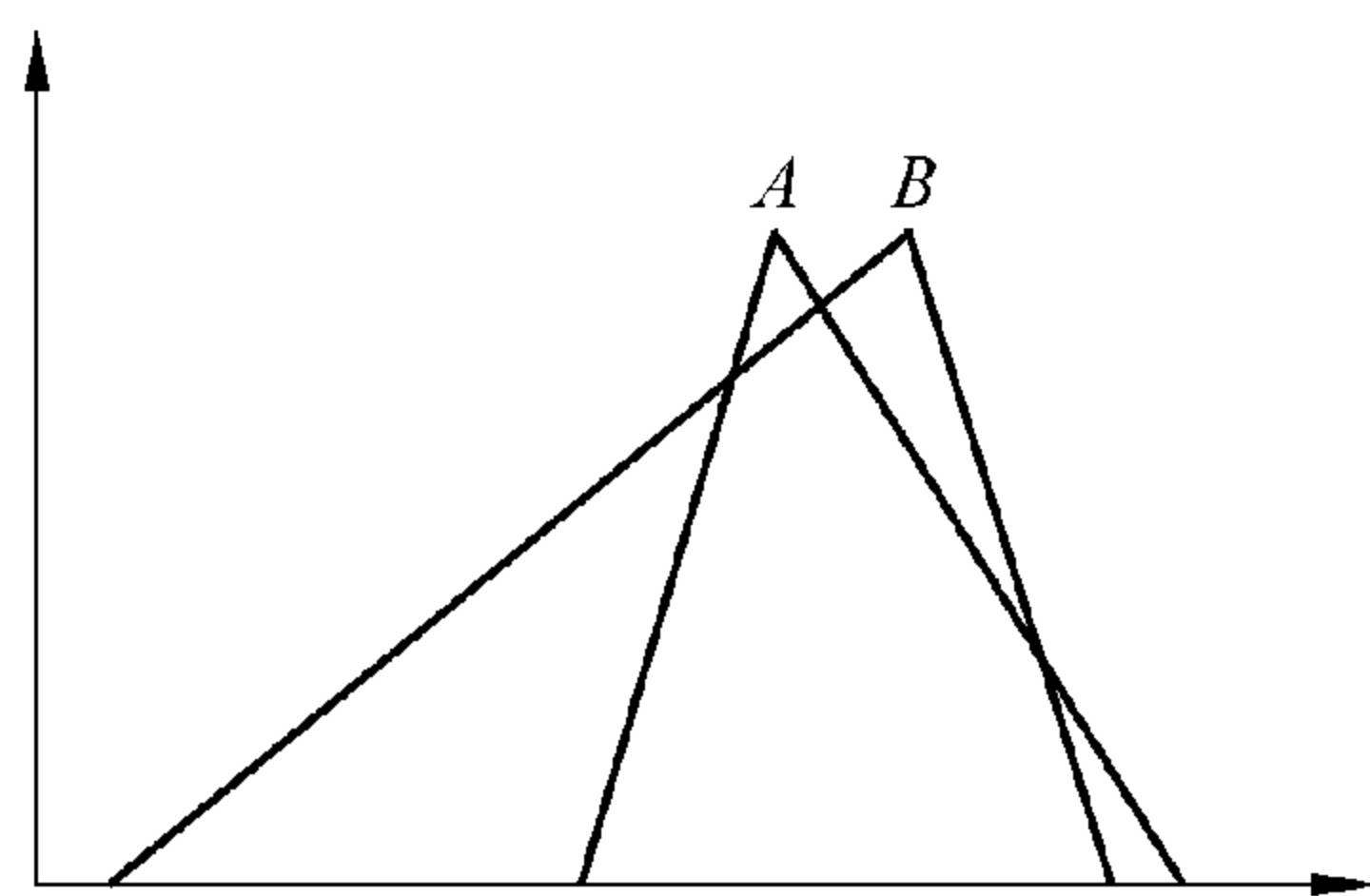


图 29.3 模糊数排序是一个不良性问题

$a_\alpha^+ \leq b_\alpha^+$, 则可认为 A 小于 B 。这就是第一种方法—— α -截集排序法。

- α -截集法 1: 设 A 和 B 是模糊数, 其 α -截集分别为 $A_\alpha = [a_\alpha^-, a_\alpha^+]$ 和 $B_\alpha = [b_\alpha^-, b_\alpha^+]$ 。如果对所有的 $\alpha \in (0, 1]$, 都有 $a_\alpha^- \leq b_\alpha^-$ 和 $a_\alpha^+ \leq b_\alpha^+$, 则认为 A 小于 B , 记为 $A \leq B$ 。这种方法的优点在于, 它得到的结论少有争议。其缺点在于, 它只能应用于一些情况明晰的场合。如图 29.3 中的模糊数 A 和 B 就不能用这种方法进行比较, 原因在于, α 较小时, 有 $b_\alpha^- \leq a_\alpha^-$; 而 α 较大时, 又有 $b_\alpha^- \geq a_\alpha^-$ 。因此, 下面给出了一种弱排序方法。
- α -截集法 2: 如果对所有的 $\alpha \in (c, 1]$, 有 $a_\alpha^+ \leq b_\alpha^+$, 则认为 $A \leq B$, 其中 c 通常为大于 0.5 常数。

显然, 这种方法强调具有较大隶属值的模糊数。根据这种排序方法对图 29.3 中的模糊数 A 和 B 进行排序, 可得 $A \leq B$ 。

下面介绍海明距离法。它的基本思路是根据 α -截集法 1, 寻找一个模糊数 C 使其满足 $A \leq C$ 和 $B \leq C$ (很合理, 但严格)。如果 C 和 A 的海明距离大于 C 和 B 海明距离, 则认为 $A \leq B$ 。显然 C 的选择不惟一。这里把 C 选为模糊数 $MAX(A, B)$, 其定义为

$$\mu_{MAX(A, B)}(z) = \sup_{\max(x, y) = z} \min[\mu_A(x), \mu_B(y)] \quad (29.54)$$

注意, $MAX(A, B)$ 不同于隶属度函数为 $\max(\mu_A, \mu_B)$ 的模糊集。根据 α -截集法 1, 容易证明, $A \leq MAX(A, B)$, $B \leq MAX(A, B)$ 。令海明距离 $d_1(C, D)$ 为式(28.26)中所定义的海明距离 (其中 $p = 1$), 则有如下模糊数的排序方法:

- 海明距离法: 如果 $d_1(MAX(A, B), A) \geq d_1(MAX(A, B), B)$, 则认为 $A \leq B$ 。

例 29.6 考虑模糊数

$$\mu_A(x) = \mu_A(x; 1, 2, 5) \quad (29.55)$$

$$\mu_B(x) = \mu_B(x; 2, 3, 4) \quad (29.56)$$

则由式(29.54), 可得

$$\mu_{MAX(A, B)}(x) = \begin{cases} x - 2 & 2 \leq x < 3 \\ 4 - x & 3 \leq x < 3.5 \\ \frac{5 - x}{3} & 3.5 \leq x \leq 5 \\ 0 & x < 1 \text{ 或 } x > 5 \end{cases} \quad (29.57)$$

其图形是图 29.4 中的粗线部分。容易看到

$$\begin{aligned} & d_1(MAX(A, B), A) \\ &= \int_R \left| \mu_{MAX(A, B)}(x) - \mu_A(x) \right| dx = 1 \end{aligned} \quad (29.58)$$

和

$$d_1(MAX(A, B), B)$$

$$= \int_R \left| \mu_{\text{MAX}(A,B)}(x) - \mu_B(x) \right| dx = 0. \quad (29.59)$$

25

因此,由海明距离法可得 $A \leq B$ 。如果采用 α -截集法 2,并令 $c = 0.5$,则仍可得 $A \leq B$ 。显然, α -截集法 1 不适用于此例。

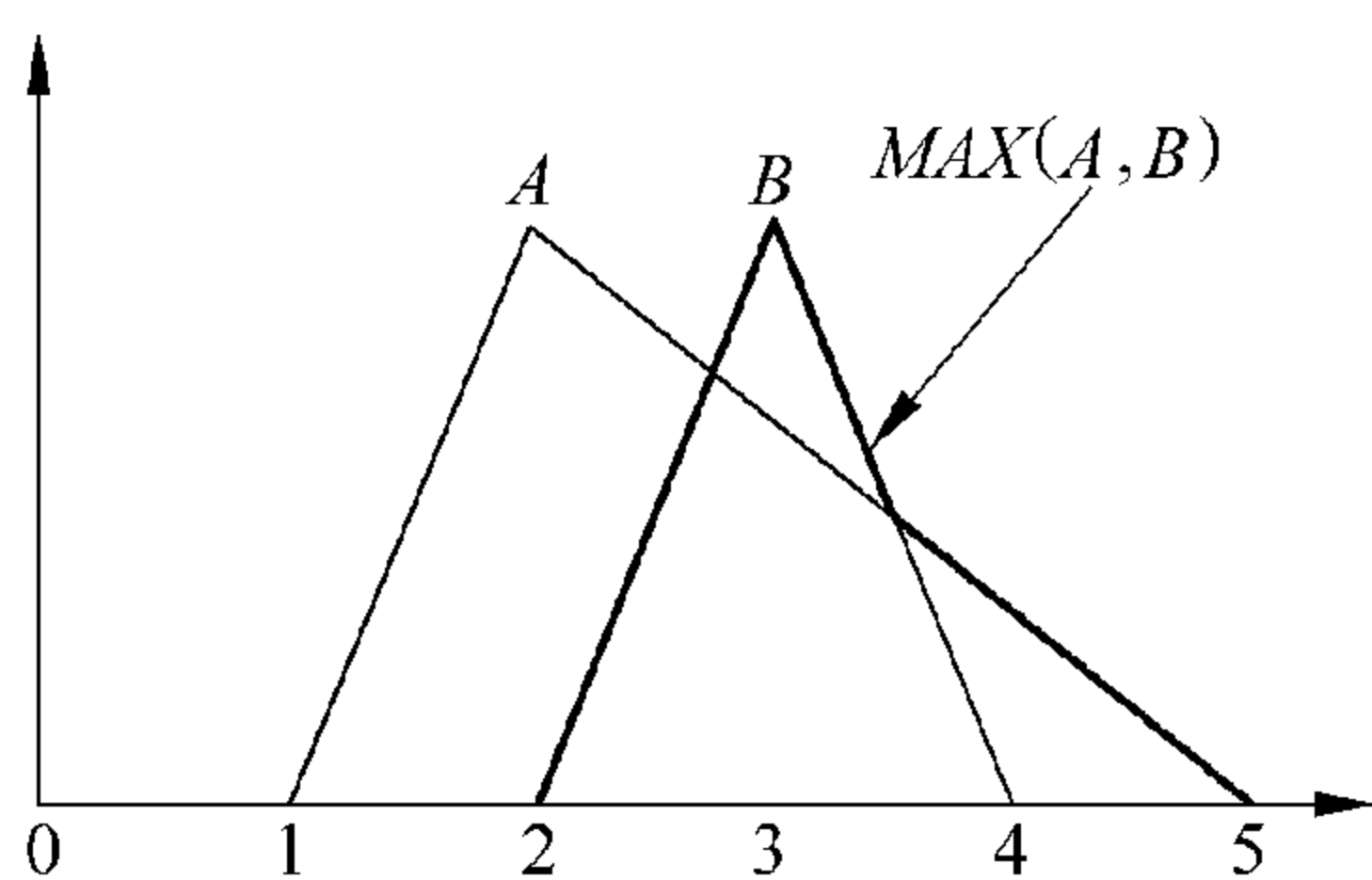


图 29.4 例 29.6 中的两个模糊数

29.6 本章小结及导读

本章中主要阐述了以下内容:

- 模糊数及模糊数的加、减、乘、除运算;
- 模糊数运算的 α -截集法和扩展原理法的等价性;
- 模糊集可由其 α -截集来确定(分解定理);
- 如何求解模糊方程;
- 用于模糊数排序的 α -截集法和海明距离法。

Kaufmann 和 Gupta [1985]是一部关于模糊算术的专著。在模糊数排序中用到的各种方法的综述可见 Chen 和 Hwang [1992]。

29.7 习题

习题 29.1 设 A 为定义在 R 上的模糊集。证明 A 为一个模糊数的充要条件是,存在一个闭区间 $[a, b] \neq \emptyset$,使得下式成立

$$\mu_A(x) = \begin{cases} 1 & \text{当 } x \in [a, b] \\ l(x) & \text{当 } x \in (-\infty, a) \\ r(x) & \text{当 } x \in (b, \infty) \end{cases} \quad (29.60)$$

式中, $l(x)$ 是一个从 $(-\infty, a)$ 映射到 $[0, 1]$ 的函数,它是单调增加且右连续的,所以,当 $x \in (-\infty, w_1)$ 时,有 $l(x) = 0$; $r(x)$ 是一个从 (b, ∞) 映射到 $[0, 1]$ 的函数,它是单调减少且左连续的,所以,当 $x \in (w_2, \infty)$ 时,有 $r(x) = 0$ 。

习题 29.2 对任意模糊集 A ,证明下式成立

$$A_\alpha = \bigcap_{\beta < \alpha} A_\beta \quad (29.61)$$

习题 29.3 证明由式(29.17)和式(29.18)所定义的模糊集 $A + B$ 和 $A - B$ 是模糊数。

习题 29.4 计算 $A + B, A - B, A \cdot B, A/B$ 和 $\text{MAX}(A, B)$,其中模糊数 A 和 B 的定义为

$$(a) \quad \mu_A(x) = \mu_A(x; -1, 1, 3), \mu_B(x) = \mu_B(x; 1, 3, 5);$$

(b) $\mu_A(x) = \mu_A(x; -2, 0, 2), \mu_B(x) = \mu_B(x; 2, 4, 6)$ 。

习题 29.5 证明用于定义模糊数的乘和除的 α -截集法和扩展原理法是等价的。

习题 29.6 设 A 和 B 是习题 29.4 中模糊数[(a)或(b)], C 为三角形模糊数, 且 $\mu_C(x) = \mu_C(x; 6, 8, 10)$ 。求解下面模糊方程中的 X :

(a) $A + X = B$;

(b) $B \cdot X = C$ 。

习题 29.7 采用 29.5 小节中的模糊数排序方法对 $\mu_A(x) = \mu_A(x; 0, 1, 3, 4), \mu_B(x) = \mu_B(x; 3, 4, 5)$ 和 $\mu_C(x) = \mu_C(x; 4, 5, 6)$ 三个模糊数进行排序。

习题 29.8 证明根据 α -截集法 1, 有 $A \leq \text{MAX}(A, B)$ 和 $B \leq \text{MAX}(A, B)$ 成立, 其中 $\text{MAX}(A, B)$ 的定义见式(29.54)。

30.1 模糊线性规划问题分类

在描述信息量庞大、内部联系错综复杂的问题时,线性规划是最常用的、也是最省力的方法。线性规划得以普及的主要原因有两点:(i)许多实际问题可以利用线性规划建模;(ii)求解线性规划问题已有一些有效方法(如单纯形法,见 Luenberger [1984])。经典线性规划问题是指,通过求解未知变量的值,使得线性目标函数在一定的线性不等式或方程约束条件下达到最大值。一般地,标准的线性规划问题可以描述如下:

$$\begin{aligned} \max \quad & \mathbf{c}\mathbf{x} \\ \text{s. t} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \quad (30.1)$$

其中, $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ 是要确定的决策变量, $\mathbf{c} = (c_1, c_2, \dots, c_n)$ 是目标系数, $\mathbf{A} = [a_{ij}] \in R^{m \times n}$ 是约束矩阵, 其元素 a_{ij} 是约束系数, $\mathbf{b} = (b_1, b_2, \dots, b_m)^T$ 是资源向量。

在许多实际情况中,很难将目标函数和约束精确明晰地描述出来。为了理解在什么地方以及如何将模糊性应用到线性规划问题中,先看一个简单例子。

例 30.1 (Lai 和 Hwang [1992]) 一个玩具公司生产两种玩具。玩具 A 是高质量的,其单位利润为 0.4 美元,玩具 B 是低质量的,其单位利润为 0.3 美元。假设每天生产 x_1 个玩具 A 和 x_2 个玩具 B, 这样每天可获利润 $0.4x_1 + 0.3x_2$ 。尽管生产每个玩具 A 可以获得较高的单位利润,但它需要的劳动力小时数是生产玩具 B 的两倍。若该公司每天可用的劳动力小时数为 500 小时,则有约束 $2x_1 + x_2 \leq 500$ 。此外,每天供应的原材料只够生产 400 个玩具(A 和 B 加在一起),因此,还有一个约束 $x_1 + x_2 \leq 400$ 。归纳得出,公司管理部门会建立如下生产计划问题模型:

$$\begin{aligned} \max \quad & 0.4x_1 + 0.3x_2 && \text{(利润)} \\ \text{s. t} \quad & 2x_1 + x_2 \leq 500 && \text{(劳动力小时数)} \\ & x_1 + x_2 \leq 400 && \text{(原材料)} \\ & x_1, x_2 \geq 0 && \end{aligned} \quad (30.2)$$

不过,因为经理可以请求工人加班和要求供应商提供额外的原材料,所以总的劳动力小时数和原材料资源数可能不那么精确。因此,一些容差应该可加入到式(30.2)的约束中。例如,当实际劳动力工时数 $2x_1 + x_2$ 小于 500 时,认为约束 $2x_1 + x_2 \leq 500$ 是可以得到完全满

足的；当实际劳动力工时数 $2x_1 + x_2$ 大于 600 时，认为约束 $2x_1 + x_2 \leq 500$ 被彻底破坏了；而

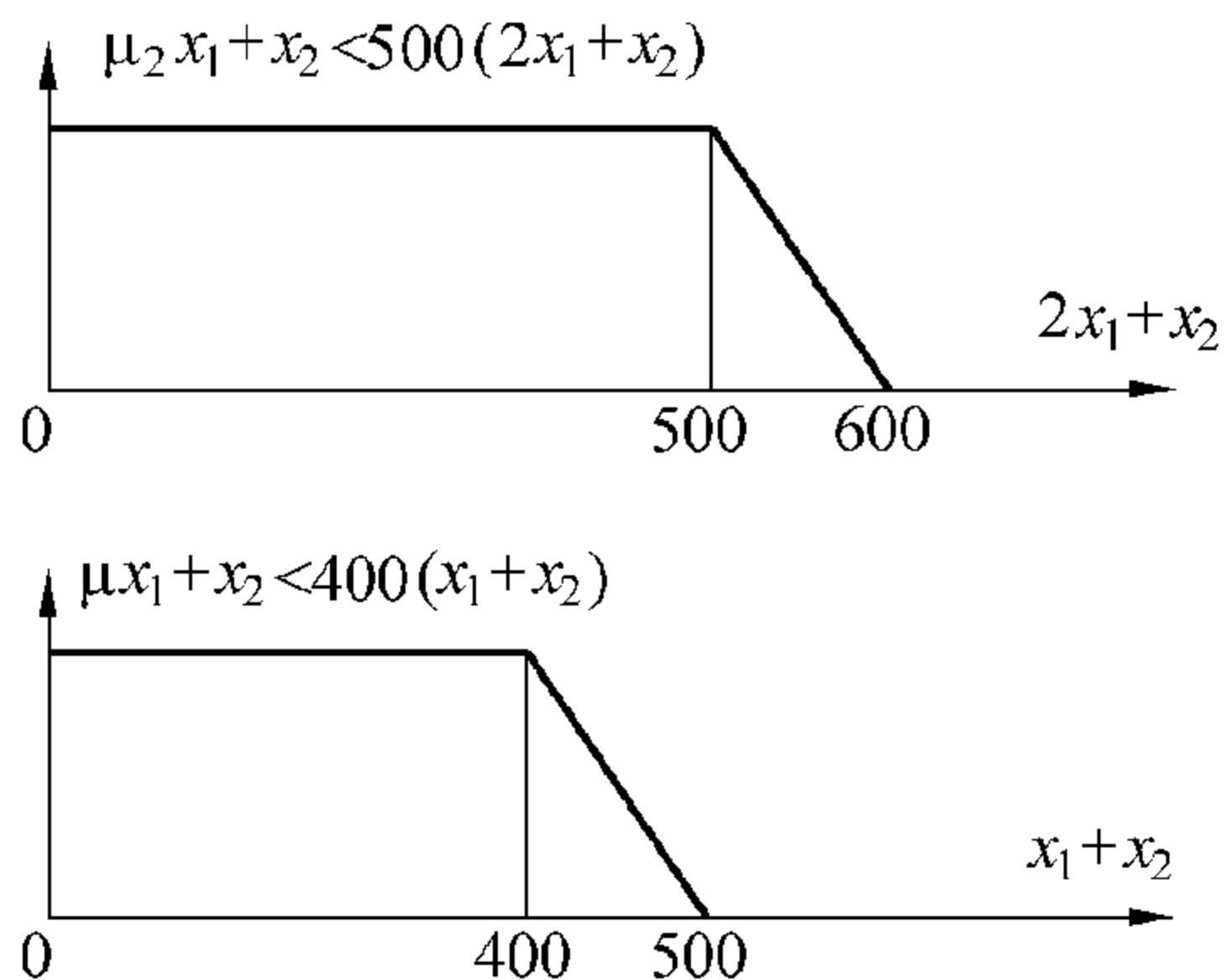


图 30.1 例 30.1 中劳动力小时和原材料的隶属度函数

当实际劳动力工时数 $2x_1 + x_2$ 介于 500 和 600 之间时，可以使用线性单调减函数来表示约束的满足程度。类似地，也可以定义一个隶属度函数来描述约束 $x_1 + x_2 \leq 400$ 被满足的程度。这些隶属度函数的图形见图 30.1。总之，第一类模糊性出现在资源描述中，称此问题为模糊资源型线性规划问题。

第二类模糊性出现在描述目标函数的系数 0.4 和 0.3 中。由于市场是不断变化的，可能无法确定玩具 A 和玩具 B 各自的单位利润就是 0.4 美元和 0.3 美元，他们只能看做是最可能的单位利润值。因此，用模糊数来表示目标函数中的系数是合理的。这是

第二种类型的模糊线性规划问题——模糊目标系数型线性规划问题。

第三类模糊性出现在约束系数中。由于劳动力素质的不一致性，生产玩具 A 和玩具 B 需要的劳动力小时的比率只能认为大约为 2。类似地，生产每个玩具 A 和玩具 B 需要的原材料也可能有些许的差异。因此，用模糊数来表示约束中的系数是合理的。这就给出了第三种类型的模糊规划问题——模糊约束系数型线性规划问题。

总结例 30.1 的分析可以看出，有以下三种类型的模糊线性规划问题：

- 模糊资源型线性规划问题

$$\begin{aligned} \max \quad & \mathbf{c}\mathbf{x} \\ \text{s. t.} \quad & \mathbf{A}\mathbf{x} \lesssim \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \quad (30.3)$$

其中，模糊不等式 \lesssim 是用类似于图 30.1 中的隶属度函数来描述的（将在 30.2 节中给出其确切的定义）。

- 模糊目标系数型线性规划问题

$$\begin{aligned} \max \quad & \tilde{\mathbf{c}}\mathbf{x} \\ \text{s. t.} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \quad (30.4)$$

其中， $\tilde{\mathbf{c}} = (\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_n)$ 是模糊数向量。

- 模糊约束系数型线性规划问题

$$\begin{aligned} \max \quad & \mathbf{c}\mathbf{x} \\ \text{s. t.} \quad & \tilde{\mathbf{A}}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \quad (30.5)$$

其中, $\tilde{A} = [\tilde{a}_{ij}]$ 是由模糊数构成的矩阵。

当然, 这三种类型的模糊线性规划问题可以组合成更多类型的模糊线性规划问题。不过, 如果知道了这三种问题的求解方法, 就可以用类似的方法去求解其他问题了。接下来的三个的小节, 将分别介绍求解这三种问题的一些基本方法。

30.2 模糊资源型线性规划问题

考虑式(30.3)所给出的模糊资源型线性规划问题。令 $t_i (> 0)$ 表示第 i 项资源 b_i 的容差, 则模糊不等式 $(A\mathbf{x})_i \lesssim b_i$ 就可具体化为 $(A\mathbf{x})_i \lesssim b_i + \theta t_i$, 其中 $\theta \in [0, 1]$ 。换言之, 模糊约束 $(A\mathbf{x})_i \lesssim b_i$ 可以定义为一个模糊集 i , 其隶属度函数为

$$\mu_i(\mathbf{x}) = \begin{cases} 1 & \text{若 } (A\mathbf{x})_i < b_i \\ 1 - [(A\mathbf{x})_i - b_i]/t_i & \text{若 } b_i \leq (A\mathbf{x})_i \leq b_i + t_i \\ 0 & \text{若 } (A\mathbf{x})_i > b_i + t_i \end{cases} \quad (30.6)$$

因此, 该问题变成求解 \mathbf{x} , 使得 $\mathbf{c}\mathbf{x}$ 和 $\mu_i(\mathbf{x}) (i = 1, 2, \dots, m)$ 最大, 这是一个多目标优化问题。

Werners [1987] 给出了求解该问题的方法。首先, 求解下面两个标准线性规划问题:

$$\begin{array}{ll} \max & \mathbf{c}\mathbf{x} \\ \text{s. t} & (A\mathbf{x})_i \leq b_i \quad i = 1, 2, \dots, m \\ & \mathbf{x} \geq 0 \end{array} \quad (30.7)$$

和

$$\begin{array}{ll} \max & \mathbf{c}\mathbf{x} \\ \text{s. t} & (A\mathbf{x})_i \leq b_i + t_i \quad i = 1, 2, \dots, m \\ & \mathbf{x} \geq 0 \end{array} \quad (30.8)$$

令 \mathbf{x}^0 和 \mathbf{x}^1 分别为式(30.7)和式(30.8)所示线性规划的解, 并分别定义 $z^0 = \mathbf{c}\mathbf{x}^0$ 和 $z^1 = \mathbf{c}\mathbf{x}^1$, 则可用下面定义的隶属度函数描述最优性程度

$$\mu_0(\mathbf{x}) = \begin{cases} 1 & \text{若 } \mathbf{c}\mathbf{x} > z^1 \\ 1 - \frac{z^1 - \mathbf{c}\mathbf{x}}{z^1 - z^0} & \text{若 } z^0 \leq \mathbf{c}\mathbf{x} \leq z^1 \\ 0 & \text{若 } \mathbf{c}\mathbf{x} < z^0 \end{cases} \quad (30.9)$$

显然, 当 $\mathbf{c}\mathbf{x} \geq z^1$ 时, 有 $\mu_0(\mathbf{x}) = 1$, 即最优性程度为最大; 当 $\mathbf{c}\mathbf{x} \leq z^0$ 时, 有 $\mu_0(\mathbf{x}) = 0$, 即最优性程度为最小; 当 $\mathbf{c}\mathbf{x}$ 介于 z^1 和 z^0 之间时, 最优性程度在 1 到 0 之间变化。

由于约束和目标函数分别用隶属度函数(30.6)和(30.9)表示, 所以可用极大—极小方法求解多目标最优化问题。具体地, 问题变为

$$\max_{\mathbf{x} \geq 0} \min[\mu_0(\mathbf{x}), \mu_1(\mathbf{x}), \dots, \mu_m(\mathbf{x})] \quad (30.10)$$

或等价于

$$\begin{aligned} \max \quad & \alpha \\ \text{s. t.} \quad & \mu_0(\mathbf{x}) \geq \alpha \\ & \mu_i(\mathbf{x}) \geq \alpha \quad i = 1, 2, \dots, m \\ & \alpha \in [0, 1], \quad \mathbf{x} \geq 0 \end{aligned} \quad (30.11)$$

将式(30.6) ~ 式(30.9)代入到式(30.11)中, 模糊资源型线性规划问题(30.3)可通过求解下面的标准线性规划问题而得以解决:

$$\begin{aligned} \max \quad & \alpha \\ \text{s. t.} \quad & \mathbf{c}\mathbf{x} \geq z^1 - (1 - \alpha)(z^1 - z^0) \\ & (A\mathbf{x})_i \leq b_i + (1 - \alpha)t_i \quad i = 1, 2, \dots, m \\ & \alpha \in [0, 1], \quad \mathbf{x} \geq 0 \end{aligned} \quad (30.12)$$

例 30.1 (Lai 和 Hwang [1992]) 考虑下面混合产品选择问题

$$\begin{aligned} \max \quad & 4x_1 + 5x_2 + 9x_3 + 11x_4 \quad (\text{利润}) \\ \text{s. t.} \quad & g_1(\mathbf{x}) = x_1 + x_2 + x_3 + x_4 \leq 15 \quad (\text{劳动力}) \\ & g_2(\mathbf{x}) = 7x_1 + 5x_2 + 3x_3 + 2x_4 \leq 80 \quad (\text{原材料 } Y) \\ & g_3(\mathbf{x}) = 3x_1 + 5x_2 + 10x_3 + 15x_4 \leq 100 \quad (\text{原材料 } Z) \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned} \quad (30.13)$$

其中, 劳动力、原材料 Y 和原材料 Z 的容差分别为 $t_1 = 5, t_2 = 40$ 和 $t_3 = 30$ 。求解式(30.7)和式(30.8), 可得 $z^0 = 99.29$ 和 $z^1 = 130$ 。由(30.12)可得, 原问题的等价问题为

$$\begin{aligned} \max \quad & \theta \\ \text{s. t.} \quad & z = 4x_1 + 5x_2 + 9x_3 + 11x_4 \geq 130 - 30.71\theta \\ & g_1(\mathbf{x}) = x_1 + x_2 + x_3 + x_4 \leq 15 + 5\theta \\ & g_2(\mathbf{x}) = 7x_1 + 5x_2 + 3x_3 + 2x_4 \leq 80 + 40\theta \\ & g_3(\mathbf{x}) = 3x_1 + 5x_2 + 10x_3 + 15x_4 \leq 100 + 30\theta \\ & x_1, x_2, x_3, x_4 \geq 0, \theta \in [0, 1] \end{aligned} \quad (30.14)$$

其中, $\theta = 1 - \alpha$ 。最后得到, 当 $\theta = 0.5$ 时, 解为 $z^* = 114.65$ 。

30.3 模糊目标系数型线性规划问题

考虑式(30.4)所给出的模糊目标系数型线性规划问题。为了简单而又不失一般性, 假设 \tilde{c}_i 为三角形模糊数, 且其隶属度函数为 $\mu_{\tilde{c}_i}(x; c_i^-, c_i^0, c_i^+)$, 记为 $\tilde{c}_i = (c_i^-, c_i^0, c_i^+)$ 。则式

(30.4)变为

$$\begin{aligned} \max \quad & (\mathbf{c}^- \mathbf{x}, \mathbf{c}^0 \mathbf{x}, \mathbf{c}^+ \mathbf{x}) \\ \text{s. t.} \quad & A\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \quad (30.15)$$

其中, $\mathbf{c}^- = (c_1^-, c_2^-, \dots, c_n^-)$, $\mathbf{c}^0 = (c_1^0, c_2^0, \dots, c_n^0)$ 和 $\mathbf{c}^+ = (c_1^+, c_2^+, \dots, c_n^+)$ 。这是一个多目标线性规划问题。现有文献中(见 Lai 和 Hwang [1992])提出了许多求解该问题的方法,下面介绍两种方法。

第一种方法是简单地把三个目标组合为一个目标函数。如, $\mathbf{c}^- \mathbf{x}$, $\mathbf{c}^0 \mathbf{x}$ 和 $\mathbf{c}^+ \mathbf{x}$ 可以组合为所谓的最适当准则(most-likely criterion) $\frac{(4\mathbf{c}^0 + \mathbf{c}^- + \mathbf{c}^+) \mathbf{x}}{6}$ (Lai 和 Hwang [1992])。所以式(30.15)可以转化为下面的标准线性规划问题:

$$\begin{aligned} \max \quad & \frac{4\mathbf{c}^0 + \mathbf{c}^- + \mathbf{c}^+}{6} \mathbf{x} \\ \text{s. t.} \quad & A\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \quad (30.16)$$

也可采用其他的加权求和方法。

第二种方法是通过观察使三角形模糊数($\mathbf{c}^- \mathbf{x}, \mathbf{c}^0 \mathbf{x}, \mathbf{c}^+ \mathbf{x}$)最大化这个目标而得到的。但它并不是同时最大化 $\mathbf{c}^- \mathbf{x}$ 、 $\mathbf{c}^0 \mathbf{x}$ 和 $\mathbf{c}^+ \mathbf{x}$, 而是分别最大化 $\mathbf{c}^0 \mathbf{x}$ (中心), 极小化 $\mathbf{c}^0 \mathbf{x} - \mathbf{c}^- \mathbf{x}$ (左边界)和最大化 $\mathbf{c}^+ \mathbf{x} - \mathbf{c}^0 \mathbf{x}$ (右边界)。这种方式将三角形隶属度函数推向右侧。于是, 问题(30.15)变为另一个多目标线性规划问题:

$$\begin{aligned} \min \quad & z_1 = (\mathbf{c}^0 - \mathbf{c}^-) \mathbf{x} \\ \max \quad & z_2 = \mathbf{c}^0 \mathbf{x} \\ \max \quad & z_3 = (\mathbf{c}^+ - \mathbf{c}^0) \mathbf{x} \\ \text{s. t.} \quad & A\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \quad (30.17)$$

求解这个问题的一种方法是用隶属度函数来描述这三个目标函数, 然后最大化它们的 α -截集。具体地, 首先得到解

$$\begin{aligned} z_1^P &= \min_{\mathbf{x} \in X} (\mathbf{c}^0 - \mathbf{c}^-) \mathbf{x}, z_1^N = \max_{\mathbf{x} \in X} (\mathbf{c}^0 - \mathbf{c}^-) \mathbf{x} \\ z_2^P &= \max_{\mathbf{x} \in X} \mathbf{c}^0 \mathbf{x}, z_2^N = \min_{\mathbf{x} \in X} \mathbf{c}^0 \mathbf{x} \\ z_3^P &= \max_{\mathbf{x} \in X} (\mathbf{c}^+ - \mathbf{c}^0) \mathbf{x}, z_3^N = \min_{\mathbf{x} \in X} (\mathbf{c}^+ - \mathbf{c}^0) \mathbf{x} \end{aligned} \quad (30.18)$$

其中, $X = \{\mathbf{x} | A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq 0\}$ 。解 z_i^P 被称为正理想解, z_i^N 被称为负理想解。然后, 定义下面三个隶属度函数来描述这三个目标。

$$\mu_{z_1}(\mathbf{x}) = \begin{cases} 1 & (\mathbf{c}^0 - \mathbf{c}^-)\mathbf{x} < z_1^P \\ \frac{z_1^N - (\mathbf{c}^0 - \mathbf{c}^-)\mathbf{x}}{z_1^N - z_1^P} & z_1^P \leq (\mathbf{c}^0 - \mathbf{c}^-)\mathbf{x} \leq z_1^N \\ 0 & (\mathbf{c}^0 - \mathbf{c}^-)\mathbf{x} > z_1^N \end{cases} \quad (30.19)$$

$$\mu_{z_2}(\mathbf{x}) = \begin{cases} 1 & \mathbf{c}^0\mathbf{x} > z_2^P \\ \frac{\mathbf{c}^0\mathbf{x} - z_2^N}{z_2^P - z_2^N} & z_2^N \leq \mathbf{c}^0\mathbf{x} \leq z_2^P \\ 0 & \mathbf{c}^0\mathbf{x} < z_2^N \end{cases} \quad (30.20)$$

$$\mu_{z_3}(\mathbf{x}) = \begin{cases} 1 & (\mathbf{c}^+ - \mathbf{c}^0)\mathbf{x} > z_3^P \\ \frac{(\mathbf{c}^+ - \mathbf{c}^0)\mathbf{x} - z_3^N}{z_3^P - z_3^N} & z_3^N \leq (\mathbf{c}^+ - \mathbf{c}^0)\mathbf{x} \leq z_3^P \\ 0 & (\mathbf{c}^+ - \mathbf{c}^0)\mathbf{x} < z_3^N \end{cases} \quad (30.21)$$

最后,通过求解下面的标准线性规划问题来求解原问题

$$\begin{aligned} \max \quad & \alpha \\ \text{s. t} \quad & \mu_{z_i}(\mathbf{x}) \geq \alpha \quad i = 1, 2, 3 \\ & A\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \quad (30.22)$$

30.4 模糊约束系数型线性规划问题

考虑式(30.5)所描述的模糊约束系数型线性规划问题。同样,为了简单和不失一般性,假设 $\tilde{A} = [\tilde{a}_{ij}]$ 是由三角形模糊数组成的,即, $\tilde{a}_{ij} = (a_{ij}^-, a_{ij}^0, a_{ij}^+)$, $\tilde{A} = (A^-, A^0, A^+)$, 其中 $A^- = [a_{ij}^-]$, $A^0 = [a_{ij}^0]$ 和 $A^+ = [a_{ij}^+]$ 。则原问题变为

$$\begin{aligned} \max \quad & \mathbf{c}\mathbf{x} \\ \text{s. t} \quad & (A^-\mathbf{x}, A^0\mathbf{x}, A^+\mathbf{x}) \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \quad (30.23)$$

采用式(30.16)中的最适当准则,可把问题(30.23)转化为下面的标准线性规划问题

$$\begin{aligned} \max \quad & \mathbf{c}\mathbf{x} \\ \text{s. t} \quad & \frac{4A^0 + A^- + A^+}{6}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \quad (30.24)$$

到目前为止,已经介绍了三种基本模糊线性规划问题(30.3)~(30.5)的求解方法。其他类型的模糊线性规划问题在本质上是上述三种基本问题的组合,因此可以使用类似的方法

法进行求解。例如,考虑所有的系数全为模糊数的线性规划问题

$$\begin{aligned} \max \quad & \tilde{\mathbf{c}}\mathbf{x} \\ \text{s. t} \quad & \tilde{A}\mathbf{x} \leq \tilde{\mathbf{b}} \\ & \mathbf{x} \geq 0 \end{aligned} \quad (30.25)$$

令 $\tilde{\mathbf{c}}, \tilde{A}$ 和 $\tilde{\mathbf{b}}$ 为三角形模糊数,即 $\tilde{\mathbf{c}} = (\mathbf{c}^-, \mathbf{c}^0, \mathbf{c}^+)$, $\tilde{A} = (A^-, A^0, A^+)$ 和 $\tilde{\mathbf{b}} = (\mathbf{b}^-, \mathbf{b}^0, \mathbf{b}^+)$, 则式 (30.25) 可以转化为下面的多目标线性规划问题

$$\begin{aligned} \min \quad & z_1 = (\mathbf{c}^0 - \mathbf{c}^-)\mathbf{x} \\ \max \quad & z_2 = \mathbf{c}^0\mathbf{x} \\ \max \quad & z_3 = (\mathbf{c}^+ - \mathbf{c}^0)\mathbf{x} \\ \text{s. t} \quad & A^-\mathbf{x} \leq \mathbf{b}^-, A^0\mathbf{x} \leq \mathbf{b}^0, A^+\mathbf{x} \leq \mathbf{b}^+ \\ & \mathbf{x} \geq 0 \end{aligned} \quad (30.26)$$

可用求解(30.17)的方法来求解此问题。

30.5 随机线性规划和模糊线性规划的比较

随机线性规划主要是用于解决系数 c 、 A 和资源数 b 为不精确数且由随机变量来描述的问题。尽管自 20 世纪 50 年代后期以来,已对随机线性规划进行了广泛研究,但在解决实际问题方面,它的应用还十分有限。正如 Lai 和 Hwang[1992]所指出的那样,随机线性规划主要的问题在于:(i)缺乏计算效率;(ii)刚性的概率论可能无法对决策者的实际不确定性问题进行建模。模糊线性规划在一定程度上克服了这两个问题。本节将对一种特殊的随机规划方法,机会—约束规划模型和相应的模糊规划方法进行比较。

机会—约束规划模型是由 Charnes 和 Cooper[1959]提出的,其定义如下

$$\begin{aligned} \max \quad & z = \sum_j \bar{c}_j x_j \\ \text{s. t.} \quad & P\left\{ \sum_j \bar{a}_{ij} x_j \leq \bar{b}_i \right\} \geq 1 - \alpha_i, \quad \forall i \\ & x_j \geq 0, \quad \forall j \end{aligned} \quad (30.27)$$

其中, \bar{c}_j, \bar{a}_{ij} 和 \bar{b}_i 为随机变量, $P\{\}$ 表示概率密度, α_i 是一个小的正常数。“机会—约束”名称的由来是因为,每个约束 $\sum_j \bar{a}_{ij} x_j \leq \bar{b}_i$ 都是以最小的概率 $1 - \alpha_i$ 来实现的。为方便阐释,假设 \bar{a}_{ij} 和 \bar{b}_i 服从均值和方差已知的高斯分布。令 $\bar{g}_i = \sum_j \bar{a}_{ij} x_j - \bar{b}_i$, 则可把式(30.27)中的概率约束看做

$$P\{\bar{g}_i \leq 0\} = P\left\{ \frac{\bar{g}_i - E[\bar{g}_i]}{\sqrt{\text{Var}[\bar{g}_i]}} \leq \frac{0 - E[\bar{g}_i]}{\sqrt{\text{Var}[\bar{g}_i]}} \right\} \geq 1 - \alpha_i \quad (30.28)$$

或

$$P \{ \bar{g}_i \leq 0 \} = \Phi \left(\frac{0 - E[\bar{g}_i]}{\text{Var}[\bar{g}_i]^{1/2}} \right) \geq 1 - \alpha_i \quad (30.29)$$

其中, Φ 为标准高斯分布的累积分布函数。由于 \bar{a}_{ij} 和 \bar{b}_i 是高斯随机变量, 所以 \bar{g}_i 也是服从如下均值和方差的正态分布

$$E[\bar{g}_i] = \sum_j E[\bar{a}_{ij}]x_j - E[\bar{b}_i] \quad (30.30)$$

$$\text{Var}[\bar{g}_i] = \mathbf{x}^T D_i \mathbf{x} \quad (30.31)$$

其中, $\mathbf{x} = (x_1, \dots, x_n)^T$, 且

$$D_i = \begin{pmatrix} \text{Var}[\bar{a}_{i1}] & \cdots & \text{Cov}[\bar{a}_{i1}, \bar{a}_{in}] & \text{Cov}[\bar{a}_{i1}, \bar{b}_i] \\ \text{Cov}[\bar{a}_{in}, \bar{a}_{i1}] & \cdots & \text{Var}[\bar{a}_{in}] & \text{Cov}[\bar{a}_{in}, \bar{b}_i] \\ \text{Cov}[\bar{b}_i, \bar{a}_{i1}] & \cdots & \text{Cov}[\bar{b}_i, \bar{a}_{in}] & \text{Var}[\bar{b}_i] \end{pmatrix} \quad (30.32)$$

令 S_{α_i} 为满足 $\Phi(S_{\alpha_i}) = 1 - \alpha_i$ 的标准值, 则约束 $P \{ \bar{g}_i \leq 0 \} \geq 1 - \alpha_i$ 能够实现的充要条件为

$$\frac{0 - E[\bar{g}_i]}{\text{Var}[\bar{g}_i]^{1/2}} \geq S_{\alpha_i} \quad (30.33)$$

将式(30.30)和式(30.31)代入到式(30.33)中, 用其均值 c_j 代替 \bar{c}_j , 则可将随机线性规划问题转化为如下确定性规划问题:

$$\begin{aligned} \max \quad & z = \sum_j c_j x_j \\ \text{s. t} \quad & \sum_j E[\bar{a}_{ij}]x_j - E[\bar{b}_i] + S_{\alpha_i} \mathbf{x}^T D_i \mathbf{x} \leq 0, \quad \forall i \\ & x_j \geq 0, \quad \forall j \end{aligned} \quad (30.34)$$

显然, 式(30.34)是一个不容易求解的非线性规划问题。而在模糊线性规划中, 可以用三角形模糊数取代系数和资源向量。由 30.2 小节 ~ 30.4 小节可以看出, 模糊线性规划问题可以转化为比非线性规划问题(30.34)更容易求解的标准线性规划问题。从运算角度看, 模糊线性规划的计算比随机线性规划更有效率。

30.6 本章小结及导读

本章主要阐述了以下内容:

- 模糊性是在什么地方和怎样出现在线性规划问题中的以及模糊线性规划问题的分类;
- 如何求解模糊资源型线性规划问题;
- 如何求解模糊目标系数型线性规划问题;

- 如何求解模糊约束系数型线性规划问题;
- 模糊线性规划与随机线性规划相比所具有的优点。

有关模糊数学规划的一本非常好的书是 Lai 和 Hwang [1992]。Chen 和 Hwang [1992], 以及 Lai 和 Hwang [1994] 的书中也论述了相关专题。经典线性规划可见 Luenberger [1984]。

30.7 习题

习题 30.1 证明线性规划问题(30.1)可以转化为下面问题:

$$\begin{aligned} \max \quad & \mathbf{c}'\mathbf{x}' \\ \text{s. t.} \quad & A'\mathbf{x}' \leq \mathbf{b}' \\ & \mathbf{x}' \geq 0 \end{aligned} \quad (30.35)$$

习题 30.2 用图解法求解线性规划问题(30.2)。

习题 30.3 用图解法求解下面的线性规划问题:

$$\begin{aligned} \min \quad & x_1 - 2x_2 \\ \text{s. t} \quad & 3x_1 - x_2 \geq 1 \\ & 2x_1 + x_2 \leq 6 \\ & x_1 \geq 0, 0 \leq x_2 \leq 2 \end{aligned} \quad (30.36)$$

习题 30.4 证明极大一极小优化问题(30.10)等价于数学规划问题(30.11)。

习题 30.5 一家公司生产的高利润装饰材料,包括 x 和 y 两种。生产装饰材料需要五种不同的配料:金线、绸、丝绒、银线和尼龙。这些配料的价格及其对 x 和 y 的技术贡献见表 30.1。每单位 x 和 y 的边际利润分别为 400 美元和 300 美元。为保持边际利润,公司要求购买配料的金额大体上不超过 2600 美元。请建立模糊资源型线性规划问题的模型,并用 30.2 小节中的方法进行求解。

表 30.1 输入和技术系数

原 料	技术系数		价格/单位	原 料	技术系数		价格/单位
	x	y			x	y	
金 线	4	0	30	银 线	0	3	20
丝 绸	2	6	40	尼 龙	4	4	10
天 鹅 绒	12	4	9.5				

习题 30.6 求解下面的模糊线性规划问题:

(a)

$$\begin{aligned}
 \max \quad & 5x_1 + 4x_2 \\
 \text{s. t} \quad & (4, 2, 1)x_1 + (5, 3, 1)x_2 \leq (24, 5, 8) \\
 & (4, 1, 2)x_1 + (1, 0.5, 1)x_2 \leq (12, 6, 3) \\
 & x_1, x_2 \geq 0
 \end{aligned} \tag{30.37}$$

(b)

$$\begin{aligned}
 \max \quad & 6x_1 + 5x_2 \\
 \text{s. t} \quad & (5, 3, 2)x_1 + (6, 4, 2)x_2 \leq (25, 6, 7) \\
 & (5, 2, 3)x_1 + (2, 1.5, 1)x_2 \leq (13, 7, 4) \\
 & x_1, x_2 \geq 0
 \end{aligned} \tag{30.38}$$

习题 30.7 设计一个或两个不同于 30.3 小节中求解模糊目标系数型线性规划问题的方法。

习题 30.8 设计一个或两个不同于 30.4 小节中求解模糊约束系数型线性规划问题的方法。

习题 30.9 你能找出模糊线性规划与随机线性规划相比的一些缺点吗？解释你的答案。

31.1 引言

Zadeh[1978]始创的可能性理论是对概率论在处理不确定性问题时的补充。正如 Zadeh[1978]在下文中多次强调的:

Wiener 和 Shannon 关于通信的统计理论的先驱性工作使得人们普遍接受了这样一种理念,信息在本质上是通过统计得来的,正因为如此,信息必须用概率论方法来处理。

毫无疑问,统计学的观点为我们深入了解编译码、数据的传送和接收等基本过程做出了贡献,并且在现代通信、探测和遥测系统等领域的发展中起了关键作用。但近年来,许多其他方面的应用成为了前沿问题,如自然语言处理、知识表示、语音识别、机器人学、医疗诊断、稀有事件分析、不确定决策、图像分析、信息修复及相关领域等,这类应用的研究重心由信息的传送转向信息的内涵,即都能够回答与数据库中储存的信息有关的问题。

.....

我们主要关注的是信息的内涵,而不是其度量方法。即,信息分析的正确框架本质上应是可能性理论,而不是概率论,这暗示着这种分析需要的不是概率论,而是一种类似的、却又不同的理论——可能性理论。

自 Zadeh[1978]以来,关于可能性理论的研究很多。对可能性理论的处理有两种方法:其一是由 Zadeh[1978]提出来的,是将可能性理论作为模糊集理论的一个扩展而引入的;其二是 Klir 和 Folger[1988]以及其他一些学者在 Dempster-Shafer 证据理论框架中提出来的。第一种方法直观上看是正确的,而且和引入可能性理论的最初目的——表示信息的含义——密切相关。第二种方法将可能性理论建立在公理的基础上,有助于对可能性理论进行深入研究。在下面两小节中,将分别介绍这两种方法的基本思想。

31.2 可能性的直观方法

31.2.1 可能性分布与可能性测度

可能性的直观方法源于模糊约束的概念。以 U 为论域, x 为在 U 上取值的一个变量, A

为 U 上的一个模糊集。则命题“ x 为 A ”可以解释为对 x 的取值起一种约束的作用,这种约束用隶属度函数 μ_A 来描述。换言之,也可以把 $\mu_A(u)$ 解释为 $x = u$ 时的可能性的程度。例如,设 x 表示人的年龄, A 表示模糊集“年轻”。假设已知“一个人是年轻的”(x 为 A),则 $\mu_A(30)$ 可以看做此人的年龄是 30 的可能性程度。为规范起见,给出如下定义。

定义 31.1 给定 U 上模糊集 A 和命题“ x 为 A ”,则与 x 有关联的可能性分布,记为 π_x ,可在数值上定义为等于 A 的隶属度函数,即

$$\pi_x(u) = \mu_A(u), u \in U \quad (31.1)$$

举个例子,将模糊集“小整数”定义为

$$\text{小整数} = 1/1 + 1/2 + 0.8/3 + 0.6/4 + 0.4/5 + 0.2/6 \quad (31.2)$$

则命题“ x 是小整数”就使得 x 与如下的可能性分布联系在一起

$$\pi_x = 1/1 + 1/2 + 0.8/3 + 0.6/4 + 0.4/5 + 0.2/6 \quad (31.3)$$

其中任一项,例如 $0.8/3$,表明“ x 是 3”确定命题“ x 是小整数”的可能性为 0.8。

现在,令 x 表示一个人的年龄, A 表示模糊集“年轻”。给定“ x 为 A ”时,可知 $x = 30$ 的可能性等于 $\mu_A(30)$ 。或许有人要问:“已知一个人是年轻的,那么这个人的年龄在 25 和 35 之间的可能性是多少呢?”。对该问题的一个合适的答案是 $\sup_{u \in [25,35]} \mu_A(u)$ 。推广这个例子就可以得到可能性测度的概念。

定义 31.2 设 C 为 U 上的一个清晰子集, π_x 是与 x 有关联的可能性分布。则 x 属于 C 的可能性测度,记为 $Pos_x(C)$,可定义为

$$Pos_x(C) = \sup_{u \in C} \pi_x(u) \quad (31.4)$$

例如,考虑由式(31.2)所定义的模糊集“小整数”和命题“ x 是小整数”,若 $C = \{3,4,5\}$,则 x 等于 3,4 或 5 的可能性的测度为

$$\begin{aligned} Pos_x(C) &= \sup_{u \in \{3,4,5\}} \pi_x(u) \\ &= \max[0.8, 0.6, 0.4] = 0.8 \end{aligned} \quad (31.5)$$

31.2.2 边际可能性分布与非关联性

令 $x = (x_1, x_2, \dots, x_n)$ 为在 $U = U_1 \times U_2 \times \dots \times U_n$ 上取值的一个向量, A 为 U 上的一个模糊集(关系),其隶属度函数为 $\mu_A(u_1, \dots, u_n)$ 。则当给定命题“ x 为 A ”时,由定义 31.1 可知, x 的可能性分布为 $\pi_x = \mu_A$,称 π_x 为基本可能性分布(basic distribution)。

现在,令 $q = (i_1, \dots, i_k)$ 为 $(1, \dots, n)$ 的一个子序列且 $x_{(q)} = (x_{i_1}, \dots, x_{i_k})$,则与 $x_{(q)}$ 有关联的边际可能性分布,记为 $\pi_{x_{(q)}}$,可定义为 π_x 在 $U_{(q)} = U_{i_1} \times \dots \times U_{i_k}$ 上的投影,即

$$\pi_{x_{(q)}}(u_{(q)}) = \sup_{u_{(q')} \in U_{(q')}} \pi_x(u) \quad (31.6)$$

其中, $u_{(q)} = (u_{i_1}, \dots, u_{i_k})$, $q' = (j_1, \dots, j_m)$ 是 $(1, \dots, n)$ 的一个子序列,也是 q 的补集,

$u_{(q')} = (u_{j_1}, \dots, u_{j_m})$ 且 $U_{(q')} = U_{j_1} \times \dots \times U_{j_m}$, 如, 当 $n=5$ 和 $q = (i_1, i_2) = (2, 4)$ 时, 有 $q' = (j_1, j_2, j_3) = (1, 3, 5)$ 。

举例说明边际可能性分布的含义。假设 $U = U_1 \times U_2 \times U_3$, 其中 $U_1 = U_2 = U_3 = \{a, b\}$, 且

$$\begin{aligned} \pi_x(u_1, u_2, u_3) = & 0.8/(a, a, a) + 1/(a, a, b) + 0.6/(b, a, a) \\ & + 0.2/(b, a, b) + 0.5/(b, b, b) \end{aligned} \quad (31.7)$$

则 (x_1, x_2) 的边际可能性分布为

$$\begin{aligned} \pi_{(x_1, x_2)}(u_1, u_2) = & \max_{u_3 \in \{a, b\}} \pi_x(u_1, u_2, u_3) \\ = & 1/(a, a) + 0.6/(b, a) + 0.5/(b, b) \end{aligned} \quad (31.8)$$

由随机变量的独立性概念类推, 当且仅当下式成立时, 变量 $x_{(q)}$ 和 $x_{(q')}$ 是非关联的

$$\pi_x(u_1, u_2, \dots, u_n) = \min[\pi_{x_{(q)}}(u_{i_1}, \dots, u_{i_k}), \pi_{x_{(q')}}(u_{j_1}, \dots, u_{j_m})] \quad (31.9)$$

其中, q 和 q' 的定义同上。因为采用的是 \min 运算, 所以 u_1 可能性的增加不能由 u_2 可能性的减少来补偿, 反之亦然。这就是非关联性的直观含义。

31.2.3 条件可能性分布

由条件概率分布的概念类推, 可以得到可能性理论中的条件可能性分布的概念。同上小节一样, 令 $q = (i_1, \dots, i_k)$ 和 $q' = (j_1, \dots, j_m)$ 为 $(1, \dots, n)$ 的互补子序列, $(a_{j_1}, \dots, a_{j_m})$ 为 $U_{j_1} \times \dots \times U_{j_m}$ 上的给定点。则当给定 $x_{(q')} = (a_{j_1}, \dots, a_{j_m})$ 时, 可定义 $x_{(q)}$ 的条件可能性分布为

$$\pi_{x_{(q)}}(u_{i_1}, \dots, u_{i_k} | x_{j_1} = a_{j_1}, \dots, x_{j_m} = a_{j_m}) = \pi_x(u_1, \dots, u_n) |_{u_{j_1} = a_{j_1}, \dots, u_{j_m} = a_{j_m}} \quad (31.10)$$

以可能性分布(31.7)为例, 可得

$$\pi_{(x_2, x_3)}(u_2, u_3 | x_1 = a) = 0.8/(a, a) + 1/(a, b) \quad (31.11)$$

$$\pi_{(x_2, x_3)}(u_2, u_3 | x_1 = b) = 0.6/(a, a) + 0.2/(a, b) + 0.5/(b, b) \quad (31.12)$$

在式(31.10)中, 条件可能性分布是以单值 $x_{(q')} = (a_{j_1}, \dots, a_{j_m})$ 为条件的分布。现假设 $x_{(q')}$ 仅被其可能性分布 $\pi_{x_{(q'')}}$ 施以模糊约束, 那么当给定 $x_{(q')}$ 时, $x_{(q)}$ 的条件可能性分布是什么呢? 在 Zadeh[1978]中, 该条件可能性分布的定义如下。首先, 将 $x_{(q')}$ 由 $U_{j_1} \times \dots \times U_{j_m}$ 柱状扩展至 $U_1 \times U_2 \times \dots \times U_n$, 即

$$\pi_{\bar{x}_{(q')}}(u_1, \dots, u_n) = \pi_{x_{(q')}}(u_{j_1}, \dots, u_{j_m}) \quad (31.13)$$

其中, $\bar{x}_{(q')}$ 是 $x_{(q')}$ 的柱状扩展。于是, 当给定 $x_{(q')}$ 时, $x_{(q)}$ 的条件可能性分布可定义为

$$\pi_{x_{(q)}}(u_{i_1}, \dots, u_{i_k} | x_{(q')}) = \sup_{u_{(q')} \in U_{j_1} \times \dots \times U_{j_m}} \min[\pi_x(u_1, \dots, u_n), \pi_{\bar{x}_{(q')}}(u_1, \dots, u_n)] \quad (31.14)$$

其中, $u_{(q')} = (u_{j_1}, \dots, u_{j_m})$ 。即, 取 π_x 和 $\pi_{\bar{x}_{(q'')}}$ 的交集且定义条件可能性分布为该交集在 $U_{i_1} \times \dots \times U_{i_k}$ 上的投影。

再看式(31.7)中的可能性分布 π_x , 并令

$$\pi_{(x_1, x_2)}(u_1, u_2) = 0.4/(a, a) + 0.8/(b, a) + 1/(b, b) \quad (31.15)$$

则, (x_1, x_2) 的柱状扩展为

$$\begin{aligned} \pi_{(\overline{x_1, x_2})}(u_1, u_2, u_3) = & 0.6/(a, a, a) + 0.4/(a, a, b) + 0.8/(b, a, a) \\ & + 0.8/(b, a, b) + 1/(b, b, a) + 1/(b, b, b) \end{aligned} \quad (31.16)$$

将式(31.16)和式(31.17)代入式(31.14), 可得

$$\pi_{x_3}(u_3 | x_1, x_2) = 0.6/a + 0.5/b \quad (31.17)$$

31.3 可能性的公理化方法

在 Klir 和 Folger[1988] 以及 Klir 和 Yuan[1995] 中, 可能性理论是在 Dempster-Shafer 的证据理论框架内提出的。因为它是建立在公理化基础之上的, 所以可导出可能性的许多数学特性。在证据理论中, 有两个重要的测度: 似真性测度和信任性测度。用这两个测度概念可以定义可能性理论中的基本概念。

31.3.1 似真性测度和信任性测度

在概率测度的定义中, 有一个非常强的要求——可加性公理, 即

$$\text{Pro}(A \cup B) = \text{Pro}(A) + \text{Pro}(B) \quad (31.18)$$

其中, $\text{Pro}(\cdot)$ 表示概率测度, A 和 B 是域 U 的子集且满足 $A \cap B = \phi$ (空集)。从不同方面放松可加性公理的条件, 就可得到似真性测度和信任性测度的定义。具体的定义如下。

定义 31.3 给定一个全集 U 和 U 的一个非空子集族 F , 似真性测度是一个函数 $Pl: F \rightarrow [0, 1]$, 它满足

- (p1) $pl(\phi) = 0$ 和 $Pl(U) = 1$ (边界条件)
- (p2) 对于所有的 $A_i \in F$, 有

$$\begin{aligned} Pl(A_1 \cap \cdots \cap A_n) \leq & \sum_j Pl(A_j) - \sum_{j < k} Pl(A_j \cup A_k) + \cdots + \\ & (-1)^{n+1} Pl(A_1 \cup \cdots \cup A_n) \end{aligned} \quad (31.19)$$

(次可加性)

- (p3) 对 F 中的任意递增序列 $A_1 \subset A_2 \subset \cdots$, 当 $\bigcup_{i=1}^{\infty} A_i \in F$ 时, 有

$$\lim_{i \rightarrow \infty} Pl(A_i) = Pl\left(\bigcup_{i=1}^{\infty} A_i\right) \quad (31.20)$$

(上连续性)

定义 31.4 给定一个全集 U 和 U 的非空子集族 F , 信任性测度是一个函数 $Bel: F \rightarrow [0, 1]$ 使得

- (p1) $Bel(\phi) = 0$ 和 $Bel(U) = 1$ (边界条件)
- (p2) 对所有的 $A_i \in F$, 有

$$Bel(A_1 \cup \cdots \cup A_n) \geq \sum_j Bel(A_j) - \sum_{j < k} Bel(A_j \cap A_k) + \cdots + (-1)^{n+1} Bel(A_1 \cap \cdots \cap A_n) \quad (31.21)$$

(超可加性)

- (p3) 对 F 中的任意递减序列 $A_1 \supset A_2 \supset \cdots$, 当 $\bigcap_{i=1}^{\infty} A_i \in F$ 时, 有

$$\lim_{i \rightarrow \infty} Bel(A_i) = Bel\left(\bigcap_{i=1}^{\infty} A_i\right) \quad (31.22)$$

(下连续性)

令式(31.19)和式(31.21)中的 $n = 2$, $A_1 = A$ 和 $A_2 = \bar{A}$, 可得如下似真性测度和信任性测度的基本不等式

$$Pl(A) + Pl(\bar{A}) \geq 1 \quad (31.23)$$

$$Bel(A) + Bel(\bar{A}) \leq 1 \quad (31.24)$$

此外, 可以证明似真性测度和信任性测度有以下关系:

$$Pl(A) + Bel(\bar{A}) = 1 \quad (31.25)$$

因此, 可以认为, 似真性测度和信任性测度互为补集。此外, 通过证明下式成立, 还可认为, 似真性测度和信任性测度是概率测度的上界和下界

$$Bel(A) \leq Pro(A) \leq Pl(A) \quad (31.26)$$

证据理论已经取得了许多成果, 有兴趣的读者可参考 Shafer[1976] 及 Guan 和 Bell [1991]。

31.3.2 可能性测度和必要性测度

在 31.2 小节给出的可能性的直观方法中, 可能性分布是由模糊集的隶属度函数推导出来的, 可能性测度又是由可能性分布定义的。而这里介绍的公理化方法中, 可能性测度被定义为一种特殊的似真性测度, 且可能性分布是由可能性测度推导出来的。因为由式(31.25)可知, 似真性测度有一个互补的信任性测度, 所以这里引进了所谓的必要性测度作为可能性测度的补集。具体地, 有下面定义。

定义 31.5 定义可能性测度 $Pos: F \rightarrow [0, 1]$ 为一种特殊的似真性测度, 其次可加性条件(31.19)由下式取代

$$Pos\left(\bigcup_{k \in K} A_k\right) = \sup_{k \in K} Pos(A_k) \quad (31.27)$$

其中, K 为任意指标集, 且 $\bigcup_{k \in K} A_k \in F$ 。

定义 31.6 定义必要性测度 $Nec: F \rightarrow [0, 1]$ 为一种特殊的信任性测度, 其超可加性条件(31.21)由下式取代

$$Nec\left(\bigcap_{k \in K} A_k\right) = \inf_{k \in K} Nec(A_k) \quad (31.28)$$

其中, K 为任意指标集, 且 $\bigcap_{k \in K} A_k \in F$ 。

因为可能性测度和必要性测度分别是一种特殊的似真性测度和信任性测度, 所以式(31.23) ~ 式(31.25)中的条件都成立, 即

$$Pos(A) + Pos(\bar{A}) \geq 1 \quad (31.29)$$

$$Nec(A) + Nec(\bar{A}) \leq 1 \quad (31.30)$$

$$Pos(A) + Nec(\bar{A}) = 1 \quad (31.31)$$

可以证明 F 上的每一个可能性测度 Pos , 对于任意 $A \in F$ 都存在一个函数 $\pi: U \rightarrow [0, 1]$ 使得下式成立

$$Pos(A) = \sup_{x \in A} \pi(x) \quad (31.32)$$

定义 π 为可能性分布, 即, 此方法中可能性分布是由可能性测度推导而来的。

除式(31.29) ~ 式(31.31)外, 可能性测度和必要性测度还有其他一些有意义的性质。例如

$$\max[Pos(A), Pos(\bar{A})] = 1 \quad (31.33)$$

$$\min[Nec(A), Nec(\bar{A})] = 0 \quad (31.34)$$

其中, 式(31.33)可由 $1 = Pos(A \cup \bar{A}) = \max[Pos(A), Pos(\bar{A})]$ 得到, 式(31.34)可由 $0 = Nec(A \cap \bar{A}) = \min[Nec(A), Nec(\bar{A})]$ 得到。此外, 由式(31.31)和式(31.33) ~ 式(31.34), 有 $Pos(A) < 1$ (表明 $Nec(A) = 0$) 和 $Nec(A) > 0$ (表明 $Pos(A) = 1$)。 $Pos(A) < 1$ 表明 $Nec(A) = 0$ 是因为, $Pos(A) < 1$ 时, 由式(31.33)可知 $Pos(\bar{A}) = 1$, 再由式(31.31)可得 $Nec(A) = 1 - Pos(\bar{A}) = 0$ 。类似地, $Nec(A) > 0$ 表明 $Pos(A) = 1$ 是因为, $Nec(A) > 0$ 时, 由式(31.34)可知 $Nec(\bar{A}) = 0$, 从而由式(31.31)可得 $Pos(A) = 1 - Nec(\bar{A}) = 1$ 。

因为可能性测度和必要性测度是定义在公理基础之上的, 所以可推导出其许多特性, 见 Klir 和 Yuan[1995] 及 Dobois 和 Prade[1988]。这些文献给出了关于可能性理论的丰富内容。

31.4 可能性与概率

31.4.1 可能性与概率的无休止的论战

看起来, 如果不讨论可能性(模糊性)和概率的差异, 本书就不够完整。的确, 这是围绕模糊理论的最有争议性的问题, 而且以前也曾引起激烈的论战。为了使叙述更为引人入胜, 把论战分为三个回合。

论战始于可能性论者^①认为,概率论不适于处理用自然语言描述的不确定性,概率论者却不以为然。概率论已经有几百年的历史了,其理论十分丰富,应用也非常成功。概率论者认为,可以将模糊集的隶属度函数看做是主观概率(Loginov[1966])。20世纪80年代,概率论者曾借助经典之作 Lindley[1982]、Lindley[1987]和 Cheeseman[1988]对模糊理论进行了一系列的攻击,随后还在统计学杂志上刊登了相关讨论。其结论是“任何模糊理论能做的事情,概率论都能做,甚至会做得更好”,而可能性论者并没有对这些强有力的攻击进行有效反击。因此,在第一轮论战接近尾声时,概率论者占了上风,而模糊理论则在随后的几年里遭到了排斥和忽视。

第一轮论战失败后,工程师占大多数的可能性论者,将其注意力转向了模糊理论的工程应用。的确,一个人可以否认一种理论的原理,但他不能否认这种理论的实际应用。在模糊控制器成功应用于家用电器和工业过程控制之后,可能性论者发起了有力的反击,他们在 IEEE Trans. on Fuzzy Systems(Vol. 2, No. 2, 1994)上对模糊性与概率之间的差异进行了论述。其结论是“两种理论是互补的,它们可以处理的不确定性的类型是不同的”。这一轮论战可能性论者占了上风,这结束了模糊理论一度遭忽视的时代,此后,许多科学家和工程师都开始重视模糊理论。

第三轮论战始于经典之作 Laviolette, Seaman, Barrett 和 Woodall[1995]的文章的发表和对其所进行的六点讨论(这六点讨论大多是由可能性论者提出的)。这篇经典之作并举了模糊和概率二者的模型,并比较了各自的优势和不足。尽管这两个学派的观点仍然存在着重要分歧,但气氛变得更有合作性了。Laviolette, Seaman, Barrett 和 Woodall[1995]在总结性评论中写道:

尽管为了慎重起见,我们还对 FST(模糊集理论)的一些哲学原则持有保留意见,但我们不会宣布这种理论无用。因为有充足的证据表明,FST的应用是成功的,尤其是在控制理论方面。这些证据证明了,如果认真实施,FST是可以起作用的……但我们会持怀疑观点,因为在于,迄今还没有找到例子说明 FST 是唯一有效的,即,使用 FST 的方法得到的解至少用概率和统计方法是得不到的。

目前,这一轮论战仍在继续,毫无疑问,未来还将会有更多的争论。

既然模糊性与概率论之间的论战是无休止的,那么现在能做的也就是列举出这两种理论主要的不同之处,由读者明鉴。

31.4.2 两种理论的主要差别

因为哲学范畴内的任何辩论都会引起争议,所以,这里避开讨论这两种理论的哲学差别,有兴趣的读者可参考前面小节提到的文献。下面,主要从技术角度论述的二者的差别。

^① 这里称模糊理论的支持者为可能性论者,概率论的支持者为概率论者,不过二者间的边界通常也是模糊的。

差别 1 尽管二者处理的都是存在不确定性的问题,但它们倾向解决的实际问题却大不相同。模糊理论着重论述人类的推理和看法,因此,它研究的对象是工业过程控制、模式识别和群决策等领域中的问题,在这些领域中人的因素起着重要的作用。而概率论的应用则主要集中在统计技术、数据分析和通信系统等领域,在这些领域中人的推理和看法不起主要作用。

尽管概率论者声称,如果需要,概率论可以用来描述人的知识,但这方面研究的发展始终没有达到模糊理论所达到的规模。事实上,正如下面 Bernardo 和 Smith[1994]所声称的那样,主流概率论者尽量避开描述人的行为:

有必要指出我们所提出的公理是规范的而非描述的。所以,公理体系不会描述个人在阐述问题或做出选择时的实际行为方式,也不会告诉我们在某些假定的道德规范基础上应遵循的行为方式。它只是简单地规定了一些约束条件,似乎在告诉我们,必须了解这些约束条件,否则在选择备选方案时会出现某种形式的行为不一致性”。

因此,尽管两派声称他们处理的是同一主题——不确定性,但他们真正所解决的问题却大相径庭。

差别 2 模糊集理论舍弃了排中率,而概率论却是建立在经典集合的理论基础之上的,其中排中率是其基石之一。理论基础的差异决定了这两种理论的学术内容的巨大差别。

差别 3 31.3 小节中定义的可能性测度用较弱的次可加性条件(31.19)代替了概率论中的可加性公理(31.18)。这些基本差异也导致了可能性理论和概率论之间的一系列的差别,其中的一部分可见 31.3 节。

差别 4 从应用角度来看,这两种理论产生的算法和实现各自算法所要求的信息也大不相同。例如,在人工智能模型中,使用概率论会出现的很多问题,其根源在于相乘的两个概率非独立,而若使用模糊理论,则根本不需要这个独立性信息。

二者间还存在着许多其他差别和无休止的争论。由于本书主要是针对工程师和未来的工程师的,所以评注如何从工程师的观点来看待这些争议,或许是有益的。

31.4.3 从工程师的观点怎样看待这些争议

模糊理论和概率论间为什么会有如此大的争议?除了哲学上和学术上的差别之外,还因为科学家(大多数科学家是概率论者)和工程师(大多数工程师是可能性论者)之间所存在的职业差别。从某种意义上讲,科学家对所有的断言都持怀疑态度,而一个优秀的工程师则具有开放意识且随时准备接纳所有的备选方案。的确,科学的目的是致力于发现主宰宇宙的根本规律,而工程师的任务是构建宇宙中以前不存在的事物。因此,科学家通常吹毛求疵,把注意力放在发现断言或理论的缺陷上,而工程师则是典型的实用主义者,他致力于使用最好的方法解决问题。因此,从工程师的观点来看,只要模糊理论能够帮助我们生产出好产品,就应该理所当然地使用这种技术,而不管概率论者怎样看待模糊理论。如果将

来, 概率论者为一些现在需要模糊理论才能解决的问题提供了解决工具, 工程师们会理所当然地试用它们, 并将它们的应用结果与相应的模糊理论应用结果进行比较。下面想通过引述 Zadeh[1995]的话来结束本节和本书:

在许多情况下, 将模糊理论和概率论结合起来所带来的益处要比争论哪种方法论更好要好得多。在这一点上, 软计算的概念就是一个例子。软计算不是一种方法论, 而是一个在不精确和不确定的环境中适用的多种方法的协同体, 它通过求解不精确性、不确定性和部分真值的容差来实现可跟踪性、稳健性和低求解成本的目标。软计算主要由模糊逻辑、神经计算和概率推理构成, 后来又将遗传算法、证据推理及部分学习理论和混沌理论包含进来。

31.5 本章小结及导读

本章主要阐述了以下内容:

- 由模糊集的隶属度函数推出可能性分布;
- 以隶属度函数推出的可能性分布为基础, 推出可能性测度、边际可能性分布、条件可能性分布和非关联性等概念;
- 似真性测度和信任性测度的概念及二者的关系;
- 由似真性测度和信任性测度推出可能性测度和必要性测度以及二者的特性;
- 概率论和模糊集理论的相似之处和不同之处。

Dubois 和 Prade[1988]是关于可能性理论的一本专著。可能性的直观方法节选自 Zadeh[1978]。可能性的公理化方法的详细内容可见 Klir 和 Folger [1988]及 Klir 和 Yuan [1995]。关于模糊集和概率论的论战多见于统计学和模糊集杂志上, 例如,《国际统计评论》(1982)、《统计科学》(1987)、《IEEE Trans. on Fuzzy System》[1994] 和《Technometrics》[1995]。

31.6 习题

习题 31.1 考虑式(31.7)中的可能性分布 π_x , 求如下条件下 π_x 的条件可能性分布。

$$(a) \pi_{(x_1, x_3)}(u_1, u_3) = 0.2/(a, a) + 0.9/(a, b) + 0.7(b, b)$$

$$(b) \pi_{(x_2, x_3)}(u_2, u_3) = 0.8/(a, a) + 0.2/(b, a)$$

$$(c) \pi_{x_3}(u_3) = 0.6/a + 0.5/b$$

$$(d) \pi_{x_1}(u_1) = 0.6/a$$

习题 31.2 给定分别定义在模糊集 $X = \{a, b, c\}$ 和 $Y = \{\alpha, \beta\}$ 上的两个非关联的边际可能性分布 $\pi_x = (1, 0.8, 0.5)$ 和 $\pi_y = (1, 0.7)$, 求相应的基本可能性分布。

习题 31.3 在给定下列边际可能性分布条件下,重做习题 31.2。

(a) $X = \{a, b, c\}$ 上的 $\pi_x = (1, 0.7, 0.2)$ 和 $Y = \{\alpha, \beta, \gamma\}$ 上的 $\pi_y = (1, 1, 0.4)$

(b) $X = \{a, b, c, d\}$ 上的 $\pi_x = (1, 0.9, 0.6, 0.2)$ 和 $Y = \{\alpha, \beta\}$ 上的 $\pi_y = (1, 0.6)$

习题 31.4 证明似真性测度和信任性测度是互补的,即证明式(31.25)成立。

习题 31.5 证明概率测度是以似真性测度和信任性测度为边界的,即证明式(31.26)成立。

习题 31.6 证明可能性测度和必要性测度分别是似真性测度和信任性测度的特例。

习题 31.7 令可能性测度 Pos 为定义 31.5,证明存在一个函数 $\pi: U \rightarrow [0, 1]$ 使得式(31.32)成立。

习题 31.8 举例说明可能性和主观概率是不同的。

参 考 文 献

1. Aliev, R. , F. Aliev, and M. Babaev [1991], *Fuzzy Process Control and Knowledge Engineering in Petrochemical and Robotic Manufacturing*, Verlag TÜV Rheinland, Köln.
2. Altrock, C. V. , B. Krause, and H. J. Zimmermann [1992], “ Advanced fuzzy logic control of a model can in extreme situations, ” *Fuzzy Sets and Systems*, 48, no. 2, pp. 41 -52.
3. Anderson, B. D. O. and J. B. Moore [1990], *Optimal Control: Linear Quadratic Methods*, Prentice-Hall, NJ.
4. Aracil, J. , A. Ollero, and A. Garcia-Cerezo [1989], “ Stability indices for the global analysis of expert control systems, ” *IEEE Trans. on Systems, Man , and Cybern.* , 19, no. 5, pp. 998 -1007.
5. Åström, K. J. , and B. Wittenmark [1995], *Adaptive Control*, Addison-Wesley Publishing Company, MA.
6. Baldwin, J. F. [1993], “ Fuzzy reasoning in FRIL for fuzzy control and other knowledge-based applications, ” *Asia-Pacific Engineering J.* , 3, no. 1 -2, pp. 59 -82.
7. Bandler, W. , and L. J. Kohout [1980], “ Semantics of implication operators and fuzzy relational Products, ” *International J. of Man-Machine Studies*, 12, no. 1, pp. 89 -116.
8. Bellman, R. E. , and M. Giertz [1973], “ On the analytic formalism of the theory of fuzzy sets, ” *Information Sciences*, 5, pp. 149 -156.
9. Bellman, R. E. , and L. A. Zadeh [1970], “ Decision-making in a fuzzy environment, ” *Management Science*, 17, no. 4, pp. 141 -164.
10. _____ [1977] , “ Local and fuzzy logics, ” in *Modern Uses of Multiple -Valued Logic*, J. M. Dunn and G. Epstein, eds. , Reidel Publ . , Dordrecht, Netherlands: Reidel Publ. , pp. 103 -165.
11. Berenji, H. R. , and P. Khedkar [1992], “ Learning and tuning fuzzy logic controllers through reinforcements, ” *IEEE Trans. on Neural Networks*, 3, no. 5, pp. 724 -740.
12. Bernard. J. A. [1988], “ Use of rule-based system for process control, ” *IEEE Contr. Syst. Mag.* , 8, no. 5, pp. 3 -13.
13. Bernardo, J. M. , and A. F. M. Smith [1994], *Bayesian Theory*, John Wiley, NY.
14. Bezdek, J. C. [1981], *Pattern Recognition with Fuzzy Objective Function Algorithms* , Plenum Press, NY.
15. _____ [1994], “ Editorial: Fuzziness versus probability —again(!?), ” *IEEE Trans. on Fuzzy Systems*, 2, no. 1, pp. 1 -3.
16. Bezdek, J. C. , R. J. Hathaway, M. J. Sabin and W. T. Tucker [1987], “ Convergence theory for fuzzy c-means: counterexamples and repairs, ” *IEEE Trans. on Systems, Man, and Cybern.* , 17, no. 5, pp. 873 -877.
17. Bezdek, J. C. , and S. K. Pal, eds. [1992], *Fuzzy Models for Pattern Recognition; Methods that Search for Patterns in Data* , IEEE Press, NY.
18. Biglieri, E. , A. Gersho, R. D. Gitlin, and T. L. Lim [1984], “ Adaptive cancellation of nonlinear intersymbol interference for voiceband data transmission, ” *IEEE J. on Selected Areas in Communications*, SAC-2, 5, pp. 765 -777.

19. Bonissone, P. P. , S. Dutta, and N. C. Wood [1994], “ Merging strategic and tactical planning in dynamic and uncertain environments, ” *IEEE Trans. on Systems, Man, and Cybern.* , 24, no. 6, pp. 841-862.
20. Box, G. E. P. , and G. M. Jenkins [1976], “ *Time Series Analysis: Forecasting and Control*, ” Holden-Day Inc. , 1976.
21. Braae, M. , and D. A. Rutherford [1979], “ Theoretical and linguistic aspects of the fuzzy logic controller, ” *Automatica* , 15, pp. 553-577.
22. Brown, M. , and C. Harris [1994], *Neurofuzzy Adaptive Modeling and Control*, Prentice Hall, Englewood Cliffs, NJ.
23. Bryson, A. E. , and Y. C. Ho [1975], *Applied Optimal Control* , Hemisphere, NY.
24. Buckley, J. J. [1992a], “ Solving fuzzy equations, ” *Fuzzy Sets and Systems* , 50, no. 1, pp. 1-14.
25. _____ [1992b], “ Universal fuzzy controllers, ” *Automatica* , 28, no. 6, pp. 1245-1248.
26. _____ [1993], “ Sugeno type controllers are universal controllers, ” *Fuzzy Sets and Systems* , 53, no. 3, pp. 299-303.
27. Buckley, J. J. , and Y. Hayashi [1994], “ Can fuzzy neural nets approximate continuous fuzzy functions? ” *Fuzzy Sets and Systems* , 61, no. 1, pp. 43-52.
28. Chang, C. L. [1968], “ Fuzzy topological spaces, ” *J. of Math. Analysis and Applications* , 24, no. 1, pp. 182-190.
29. Charnes, A. , and W. W. Cooper [1959], “ Chance-constrained programming, ” *Management Science* , 5, pp. 73-79.
30. Cheeseman, P. [1988], “ An inquiry into computer understanding, ” (with discussions) *Computational Intelligence* , 4. , pp. 58-142.
31. Chen, J. Q. , and L. J. Chen [1993], “ Study on stability of fuzzy closed-loop control systems, ” *Fuzzy Sets and Systems* , 57, no. 2, pp. 159-168.
32. Chen, S. , C. F. N. Cowan, and P. M. Grant [1991] , “ Orthogonal least squares learning algorithm for radial basis function networks, ” *IEEE Trans. on Neural Networks* , 2, 2, pp. 302-309.
33. Chen, S. , G. J. Gibson, C. F. N. Cowan, and P. M. Grant [1990], “ Adaptive equalization of finite non-linear channels using multilayer perceptrons, ” *Signal Processing* , 20, pp. 107-119.
34. _____ [1991], “ Reconstruction of binary signals using an adaptive radial-basisfunction equalizer, ” *Signal Processing* , 22, pp. 77-93.
35. Chen, S. J. , and C. L. Hwang [1992], *Fuzzy Multiple Attribute Decision Making: Methods and Applications*, Springer-Verlag, NY.
36. Chiu, S. , S. Chand , D. Moore, and A. Chaudhary [1991], “ Fuzzy logic for control of roll and moment for a flexible wing aircraft, ” *IEEE Control Systems Magazine* , 11, 4, pp. 42-48.
37. Chu, C. K. , and J. M. Mendel [1994], “ First break refraction event picking using fuzzy logic systems, ” *IEEE Trans. on Fuzzy Systems* , 2, no. 4, pp. 255-266.
38. Cox, E. [1994], *The Fuzzy Systems Handbook*, Academic Press, MA.
39. Cowan, C. F. N. , and P. M. Grant (eds.) [1985], “ *Adaptive Filters*, ” Prentice -Hall, Inc. , Englewood Cliffs, NJ.

40. Cybenko, G. [1989], “ Approximation by superpositions of a sigmoidal function, ” *Mathematics of Control, Signals, and Systems*.
41. Czogala, E. , and W. Pedrycz [1982], “ Fuzzy rule generation for fuzzy control, ” *Cybernetics and Systems*, 13, no. 3, pp. 275-293.
42. Di Nola, A. , W. Pedrycz, S. Sessa, and E. Sanchez [1991], “ Fuzzy relation equations theory as a basis for fuzzy modeling : an overview, ” *Fuzzy Sets and Systems*, 40, no. 3, pp. 415-429.
43. Di Nola, A. , S. Sessa, W. Pedrycz, and E. Sanchez [1989], *Fuzzy Relation Equations and Their Applications to Knowledge Engineering*, Kluwer, Boston.
44. Dombi, J. [1982], “ A general class of fuzzy operators, the De Morgan class of fuzzy operators and fuzziness measures induced by fuzzy operators, ” *Fuzzy Sets and Systems*, 8, no. 2, pp. 149-163.
45. Driankov, D. , H. Hellendoorn, and M. Reinfrank [1993], *An Introduction to Fuzzy Control*, Springer-Verlag, Berlin.
46. Dubois, D. , and H. Prade [1980], *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, Inc. , Orlando, Florida.
47. _____ [1985], “ A review of fuzzy set aggregation connectives, ” *Information Sciences*, 36, no. 1-2, pp. 85-121.
48. _____ [1987a], “ Fuzzy numbers: an overview, ” in *Analysis of Fuzzy Information -Vol. 1: Mathematics and Logic*, Bezdek, J. C. , ed. , CRC Press, FL, pp. 3-39.
49. _____ [1987b], “ Necessity measures and the resolution principle, ” *IEEE Trans. on Systems, Man, and Cybern.* , 17, no. 3, pp. 474-478.
50. _____ [1988], *Possibility Theory*, Plenum Press, NY.
51. _____ [1991], “ Fuzzy sets in approximate reasoning, Parts 1 and 2, ” *Fuzzy Sets and Systems*, 40, no. 1, pp. 143-244.
52. _____ [1993], “ Fuzzy sets and probability: Misunderstanding, bridges and gaps, ” *Proc. 2nd IEEE Intern. Conf. on Fuzzy Systems*, San Francisco, pp. 1059-1068.
53. Duda R. O. , and P. E. Hart [1973], *Pattern Classification and Scene Analysis*, John Wiley & Sons, NY.
54. Falconer, D. D. [1978], “ Adaptive equalization of channel nonlinearities in QAM data transmission systems, ” *The Bell System Technical Journal*, 57, 7, pp. 2589-2611.
55. Filev, D. P. , and R. R. Yager [1993], “ Three models of fuzzy logic controllers. ” *Cybernetics and Systems*, 24, no. 2, pp. 91-114.
56. Gegov, A. [1994], “ Multilevel intelligent fuzzy control of oversaturated urban traffic networks, ” *Intern. J. of Systems Science*, 25, no. 6, pp. 967-978.
57. Goguen, J. A. [1967], “ L-fuzzy sets, ” *J. of Math. Analysis and Applications*, 18, no. 1, pp. 145-174.
58. Gottwald, S. [1979], “ Set theory for fuzzy sets of higher level, ” *Fuzzy Sets and Systems*, 2, no. 2, pp. 125-151.
59. Green, M. , and D. J. N. Limebeer [1995], *Linear Robust Control*, Prentice Hall, Englewood Cliffs, NJ.
60. Guan, J. W. and D. A. Bell [1991], *Evidence Theory and Its Applications, Vols. I and II*. North-Holland, NY.
61. Gupta, M. M. , J. B. Kiszka, and G. M. Trojan [1986], “ Multivariable structure of fuzzy control systems, ” *IEEE Trans. on Systems, Man, and Cybern.* , 16, no. 5, pp. 638-656.
62. Gupta. M. M. , and D. H. Rao [1994], “ On the principles of fuzzy neural networks, ” *Fuzzy Sets and Systems*,

- 61, no. 1, pp. 1-18.
63. Hansson, A. , P. Gruber , and J. Tödli [1994], “ Fuzzy anti-reset windup for PID controllers, ” *Control Eng. Practice*, 2, no. 3, pp. 389-396.
64. Harris, C. J. , C. C. Moore, and M. Brown [1993], *Intelligent Control: Aspects of Fuzzy Logic and Neural Nets*, World Scientific , Singapore.
65. Hauser, J. , S. Sastry, , and P. Kokotovic [1992], “ Nonlinear control via approximate input-output linearization: the ball and beam example, ” *IEEE Trans. on Automatic Control*, 37, 3, pp. 392-398.
66. Hellendoorn. H. , and R. Palm [1994], “ Fuzzy system technologies at Siemens R&D, ” *Fuzzy Sets and Systems*, 63, no. 3, pp. 245-269.
67. Hellendoorn H. , and C. Thomas [1993], “ Defuzzification in fuzzy controllers, ” *J. of Intelligent & Fuzzy Systems*, 1, no. 2, pp. 109-123.
68. Higashi, M. , and G. J. Klir [1982], “ On measure of fuzziness and fuzzy complements, ” *Intern. J. of General Systems*, 8, no. 3, pp. 169-180.
69. Hirota, K. , A. Arai, and S. Hachisu [1989], “ Fuzzy controlled robot arm playing two-dimensional ping-pong game, ” *Fuzzy Sets and Systems*, 32, no. 2, pp. 149-159.
70. Holmblad, L. P. , and J. J. Østergaard [1982], “ Control of a cement kiln by fuzzy logic, ” In: Gupta, M. M. , and E. Sanchez, eds. , *Fuzzy Information and Decision Processes*, North-Holland, Amsterdam, pp. 398-409.
71. Hornik, K. , M. Stinchcombe, and H. White [1989], “ Multilayer feedforward networks are universal approximators, ” *Neural Networks*, 2, pp. 359-366.
72. Hwang, Y. R. , and M. Tomizuka [1994], “ Fuzzy smoothing algorithms for variable structure systems, ” *IEEE Trans. on Fuzzy Systems*, 2, no. 4, pp. 277-284.
73. Ioannou, P. A. , and P. V. Kokotovic [1983], *Adaptive systems with Reduced Models*, Springer-Verlag, NY.
74. Isidori, A. [1989], *Nonlinear Control Systems*, Springer-Verlag, Berlin.
75. Itoh, O. , K. Gotoh, T. Nakayama, and S. Takamizawa [1987], “ Application of fuzzy control to activated sludge process, ” *Proc. 2nd IFSA Congress*, Tokyo, Japan, pp. 282-285.
76. Jang, J. R. [1993], “ ANFIS: Adaptive-network-based fuzzy inference system, ” *IEEE Trans. on Systems, Man, and Cybern.* , 23, no. 3, pp. 665-685.
77. Johansen. T. A. [1994], “ Fuzzy model based control: stability, robustness, and performance issues, ” *IEEE Trans. on Fuzzy Systems*, 2, pp. 221-234.
78. Kandel, A. , and G. Langholz. eds. [1994], *Fuzzy Control Systems*, CRC Press, FL.
79. Kaufmann, A. , and M. M. Gupta [1985], *Introduction to Fuzzy Arithmetic: Theory and Applications*, Van Nostrand, NY.
80. Kickert, W. J. M. , and H. R. Van Nauta Lemke [1976], “ Application of a fuzzy controller in a warm water plant, ” *Automatica*, 12, 4, pp. 301-308.
81. Kiszka, J. , M. Gupta, and P. Nikiforuk [1985], “ Energetic stability of fuzzy dynamic systems, ” *IEEE Trans. Systems, Man, and Cybern.* , 15, 5, pp. 783-792.
82. Klir, G. J. , and T. Folger [1988], *Fuzzy Sets, Uncertainty , and Information*, Prentice Hall, Englewood Cliffs, NJ.

83. Klir, G. J. , and B. Yuan [1995], *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, Englewood Cliffs, NJ.
84. Kosko, B. [1991], *Neural Networks and Fuzzy Systems*, Prentice Hall, Englewood Cliffs, NJ.
85. Kosko, B. [1993], *Fuzzy Thinking: The New Science of Fuzzy Logic*, Hyperion, NY.
86. Kruse, R. , J. Gebhardt, and F. Klawonn, *Foundations of Fuzzy Systems*, John Wiley & Sons, 1994.
87. Kuipers, B. , and K. Åström [1994], “ The composition and validation of heterogeneous control laws, ” *Automatica*, 30, no. 2, pp. 233-249.
88. Lai, Y. J. , and C. L. Hwang [1992], *Fuzzy Mathematical Programming*, Springer-Verlag, NY.
89. Lai, Y. J. , and C. L. Hwang [1994], *Fuzzy Multiple Objective Decision Making*, Springer-Verlag, NY.
90. Langari, G. , and M. Tomizuka [1990], “ Stability of fuzzy linguistic control systems, ” *Proc. IEEE Conf. on Decision and Control*, pp. 2185-2190.
91. Larkin, L. I. [1985], “ A fuzzy logic controller for aircraft flight control, ” in *Industrial Applications of Fuzzy Control* , ed. M. Sugeno, North-Holland , Amsterdam, pp. 87-104.
92. Larsen, P. M. [1980], “ Industrial application of fuzzy logic control, ” *Int. J. Man Mach. Studies*. 12, 1, pp. 3-10.
93. Laviolette , M. , J. W. Seaman, Jr. , J. D. Barrett, and W. H. Woodall, “ A probabilistic and statistical view of fuzzy methods, ” *Technometrics*, 37, no. 3, pp. 249-261.
94. Lee, C. C. [1990], “ Fuzzy logic in control systems: fuzzy logic controller, part I and II, ” *IEEE Trans. on Syst. , Man, and Cybern.* , 20, 2, 404-435.
95. Lewis, F. L. , K. Liu and A. Yesildirek [1993], “ Neural net robot controller with guaranteed tracking performance, ” *Proc. 1993 Int. Symp. on Intelligent Control*, pp. 225-231.
96. Lewis, F. L. , S. Q. Zhu, and K. Liu [1995], “ Function approximation by fuzzy systems, ” *Proc. 1995 American Control Conf.* , Seattle , pp. 3760-3764.
97. Lin, C. T. [1994], *Neural Fuzzy Control Systems with Structure and Parameter Learning*, World Scientific, Singapore.
98. Lindley, D. V. [1982], “ Scoring rules and the inevitability of probability, ” (With discussions), *International Statistical Review*, 50, pp. 1-26.
99. _____ [1987], “ The probability approach to the treatment of uncertainty in artificial intelligence and expert systems, ” *Statistical Science*, 2, 1, pp. 17-24.
100. Liu, Y. M. [1985], “ Some properties of convex fuzzy sets, ” *J. of Math. Analysis and Applications*, 111, no. 1, pp. 119-129.
101. Ljung, L. [1991], “ Issues in system identification, ” *IEEE Control Systems Magazine*, pp. 25-29.
102. Loginov, V. J. [1966], “ Probability treatment of Zadeh membership functions and their use in pattern recognition, ” *Engineering Cybern.* , pp. 68-69.
103. Lowen, R. [1980], “ Convex fuzzy sets, ” *Fuzzy Sets and Systems*, 3, no. 3, pp. 291-310.
104. Luenberger, D. G. [1969], *Optimization by Vector Space Methods*, John Wiley, NY.
105. _____ [1984], *Linear and Nonlinear Programming* , Addison - Wesley Publishing Company, Inc. , Reading, MA.

106. Maiers, J. , and Y. S. Sherif [1985], “ Applications of fuzzy sets theory, ” *IEEE Trans. Syst. Man Cybern.* , 15, 6, 175-189.
107. Mamdani, E. H. [1974], “ Applications of fuzzy algorithms for simple dynamic plant, ” *Proc. IEE*, 121, 12, pp. 1585-1588.
108. Mamdani, E. H. , and S. Assilian [1975], “ An experiment in linguistic synthesis with a fuzzy logic controller, ” *Int. J. Man Mach. Studies*, 7, 1, pp. 1-13.
109. Mendel, J. M. [1994], *Lessons in Digital Estimation Theory*, Prentice Hall, Englewood Cliffs, NJ.
110. McNeill, D. , and P. Freiburger [1993], *Fuzzy Logic: The Discovery of a Revolutionary Computer Technology—and How It Is Changing Our World*, Simon & Schuster, NY.
111. Miller, W. T. , R. S. Sutton, and P. J. Werbos eds. [1990], *Neural Networks for Control*, The MIT Press, Cambridge, MA.
112. Narendra, K. S. , and A. M. Annaswamy [1987], “ A new adaptive law for robust adaptation with persistent excitation, ” *IEEE Trans. on Automatic Control*, 32, no. 2, pp. 134-145.
113. _____ [1989], *Stable Adaptive Systems*, Prentice Hall , Englewood Cliffs, NJ.
114. Narendra, K. S. , and K. Parthasarathy [1990], “ Identification and control of dynamical systems using neural networks, ” *IEEE Trans. on Neural Networks*, 1, 1, pp. 4-27.
115. Nguyen, D. , and B. Widrow [1990], “ The truck backer-upper: an example of self-learning in neural networks, ” *IEEE Cont. Syst. Mag.* , 10, 3, pp. 18-23.
116. Nie, J. , and D. A. Linkens [1994], *Fuzzy Neural Control: Principles, Algorithms, and Applications*, Prentice Hall , Englewood Cliffs, NJ.
117. Pal, S. K. , and D. K. Majumder [1986], *Fuzzy Mathematical Approach to Pattern Recognition*, John Wiley, New York, NY.
118. Palm, R. [1992]. “ Sliding mode fuzzy control, ” *Proc. 1st IEEE Inter. Conf. on Fuzzy Systems*, pp. 519-526. San Diego.
119. Pappis, C. P. , and E. H. Mamdani [1977], “ A fuzzy logic controller for a traffic junction, ” *IEEE Trans. Syst. Man Cybern.* , 7, 10, pp. 707-717.
120. Pedrycz, W. [1991], “ Processing in relational structures: fuzzy relational equations, ” *Fuzzy Sets and Systems*, 40, no. 1, pp. 77-106.
121. _____ [1992], “ Fuzzy neural networks with reference neurons as pattern classifiers, ” *IEEE Trans. on Neural Networks*, 3, no. 5, pp. 770-775.
122. _____ [1993], *Fuzzy Control and Fuzzy Systems*, John Wiley & Sons, Inc. , New York, NY.
123. Pitas, I. , and A. N. Venetsanopoulos [1990], “ *Nonlinear Digital Filters*, ” Kluwer Academic Publishers, Boston.
124. Polycarpou, M. M. , and P. A. Ioannou [1991], “ Identification and control of nonlinear systems using neural network models: design and stability analysis, ” USC EE-Report.
125. Powell, M. J. D [1981], *Approximation Theory and Methods*, Cambridge University Press, Cambridge.
126. Raju, G. V. S. , J. Zhou, and R. A. Kisner [1991], “ Hierarchical fuzzy control, ” *Int. J. of Control*, 54, no. 5, pp. 1201-1216.

127. Raju, G. V. S. , and J. Zhou [1993], “ Adaptive hierarchical fuzzy controller, ” *IEEE Trans. on Systems, Man, and Cybern.* , 23, no. 4, pp. 973-980.
128. Rescher, N. [1969], *Many-Valued Logic*, McGraw-Hill, Inc. , New York.
129. Rudin, W. [1976], *Principles of Mathematical Analysis*, McGraw-Hill, Inc. , New York.
130. Ruspini, E. H. [1969], “ A new approach to clustering, ” *Information and Control*, 15, no. 1, pp. 22-32.
131. Sanner, R. M. , and J. E. Slotine [1991], “ Gaussian networks for direct adaptive control, ” *Proc. American Control Conf.* , pp. 2153-2159.
132. Sastry, S. , and M. Bodson [1989], *Adaptive Control: Stability, Convergence, and Robustness*, Prentice Hall, Englewood Cliffs, NJ.
133. Sastry, S. , and Isidori, A. [1989], “ Adaptive control of linearizable systems, ” *IEEE Trans. on Automatic Control*, 34, 11, pp. 1123-1131.
134. Schwartz, A. L. [1992]. “ Comments on “ Fuzzy logic for control of roll and moment for a flexible wing aircraft, ” ” *IEEE Control Systems Magazine*.
135. Shafer, G. [1976], *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ.
136. Simpson, P. K. [1992], “ Fuzzy min-max neural networks, Part 1: Classification, ” *IEEE Trans. on Neural Networks*, 3, no. 5, pp. 776-786.
137. Slotine, J. E. , and W. Li [1991], *Applied Nonlinear Control*, Prentice Hall, Englewood Cliffs, NJ.
138. Spooner, J. T. , and K. M. Passino [1995], “ Stable adaptive control using fuzzy systems and neural networks, ” IVHS OSU Report 95-01.
139. Stallings, W. [1977], “ Fuzzy set theory versus Bayesian statistics, ” *IEEE Trans. Systems, Man, and Cybern.* , 7; 3, pp. 216-219.
140. Sugeno, M. [1977], “ Fuzzy measures and fuzzy integrals: A survey, ” In: Gupta, M. , G. N. Saridis, and B. R. Gaines, eds. , *Fuzzy Automata and Decision Processes*, North-Holland, NY, pp. 329-346.
141. Sugeno, M. , ed. [1985], *Industrial Applications of Fuzzy Control*, North-Holland, NY.
142. Sugeno, M. , and M. Nishida [1985], “ Fuzzy control of model car, ” *Fuzzy Sets and Systems*, pp. 103-113.
143. Sugeno, M. , T. Murofushi, T. Mori, T. Tatematsu, and J. Tanaka [1989], “ Fuzzy algorithmic control of a model car by oral instructions, ” *Fuzzy Sets and Systems*, 32, pp. 207-219.
144. Sugeno, M. , and G. T. Kang [1988], “ Structure identification of fuzzy model, ” *Fuzzy Sets and Systems*, 28, no. 1, pp. 15-33.
145. Sugeno, M. , and K. Tanaka [1991], “ Successive identification of a fuzzy model and its applications to prediction of a complex system, ” *Fuzzy Sets and Systems*, 42, no. 3, pp. 315-334.
146. Takagi, T. , and M. Sugeno [1985], “ Fuzzy identification of systems and its applications to modeling and control, ” *IEEE Trans. on Systems, Man, and Cybern.* , 15, 1, pp. 116-132.
147. Tanaka, K. , and M. Sano [1994], “ A robust stabilization problem of fuzzy control systems and its application to backing up a truck-trailer, ” *IEEE Trans. on Fuzzy Systems*, 2, no. 2, pp. 107-118.
148. Tanaka, K. , and M. Sugeno [1992], “ Stability analysis and design of fuzzy control systems, ” *Fuzzy Sets and Systems* , 45, no. 2, pp. 135-156.
149. Terano, T. , K. Asai, and M. Sugeno, eds. [1994], *Applied Fuzzy Systems*, Academic Press, MA.
150. Tong, R. M. [1980], “ Some properties of fuzzy feedback systems, ” *IEEE Trans. Systems, Man, and Cybern.* ,

- 10,6, pp. 327-330.
151. _____ [1984], "A retrospective view of fuzzy control systems," *Fuzzy Sets and Systems*, 14,3, pp. 199-210.
 152. Tong, R. M., M. B. Beck, and A. Latten [1980], "Fuzzy control of the activated sludge wastewater treatment process," *Automatica*, 16,6, pp. 695-701.
 153. Togai, M., and H. Watanabe [1986], "Expert system on a chip: an engine for real-time approximate reasoning," *IEEE Expert Syst. Mag.*, 1,55-62.
 154. Turksen, I. B., and D. D. W. Yao [1984], "Representations of connectives in fuzzy reasoning: the view through normal form." *IEEE Trans. Systems, Man, and Cybern.*, 14, no. 1, pp. 146-151.
 155. Tzafestas, S. G. (ed.) [1991], *Intelligent Robotic Systems*, Marcel Dekker, Inc., NY.
 156. Umbers, I. G., and P. J. King [1980], "An analysis of human decision making in cement kiln control and the implications for automation," *Int. J. Man Mach. Studies*, 12,1, pp. 11-23.
 157. Uragami, M., M. Mizumoto, and K. Tananka [1976], "Fuzzy robot controls," *Cybern.*, 6, pp. 39-64.
 158. Utkin, V. I. [1978], *Sliding Modes and Their Application to Variable Structure Systems*, MIR Publishers, Moscow.
 159. Valavanis, K. P., and G. N. Saridis [1992], *Intelligent Robotic Systems: Theory, Design and Applications*, Kluwer Academic Publishers, Boston.
 160. Vandegrift, M. W., F. L. Lewis, S. Jagannathan, and K. Liu [1995], "Adaptive fuzzy control of discrete-time dynamical systems," personal communication.
 161. Vidyasager, M. [1993], *Nonlinear Systems Analysis*, Prentice Hall, Englewood Cliffs, NT.
 162. Voxman, W., and R. Goetschel [1983], "A note on the characterization of max and min operations," *Information Sciences*, 30, no. 1, pp. 5-10.
 163. Walsh, G., D. Tilbury, S. Sastry, and J. P. Laumond [1994], "Stabilization of trajectories for systems with nonholonomic constraints," *IEEE Trans. on Automatic Control*, 39, no. 1, pp. 216-222.
 164. Wang, L. X. [1992], "Fuzzy systems are universal approximators," *Proc. IEEE International Conf. on Fuzzy Systems*, San Diego, pp. 1163-1170.
 165. _____ [1993a], "Stable adaptive fuzzy control of nonlinear systems," *IEEE Trans. on Fuzzy Systems*, 1, no. 2, pp. 146-155.
 166. _____ [1993b], "Solving fuzzy relational equations through network training," *Proc. 2nd IEEE Intern. Conf. on Fuzzy Systems*, San Francisco, pp. 956-960.
 167. _____ [1994a], *Adaptive Fuzzy Systems and Control: Design and Stability Analysis*, Prentice Hall, Englewood Cliffs, NJ.
 168. _____ [1994b], "A supervisory controller for fuzzy control systems that guarantees stability," *IEEE Trans. on Automatic Control*, 39, no. 9, pp. 1845-1848.
 169. _____ [1995a], "Analysis and design of fuzzy identifiers of nonlinear dynamic systems," *IEEE Trans. on Automatic Control*, 40, no. 1, pp. 11-23.
 170. _____ [1996], "Universal approximation by hierarchical fuzzy systems," submitted for publication.
 171. Wang, L. X., and J. M. Mendel [1992a], "Fuzzy basis functions, universal approximation, and orthogonal least squares learning," *IEEE Trans. on Neural Networks*, 3, no. 5, pp. 807-814.
 172. _____ [1992b], "Generating fuzzy rules by learning from examples," *IEEE Trans. on Systems, Man and*

- Cybern.* ,22, no. 6, pp. 1414-1427.
173. _____ [1993], "Fuzzy adaptive filters, with application to nonlinear channel equalization," *IEEE Trans. on Fuzzy Systems*, 1, no. 3, pp. 161-170.
174. Wang, P. P. , ed. [1993], *Advances in Fuzzy Theory and Technology, Vol. I*, Bookwrights Press, Durham, NC.
175. Wang, P. Z. [1990], "A factor spaces approach to knowledge representation," *Fuzzy Sets and Systems*, 36, no. 1, pp. 113-124.
176. Werbos, P. [1974], "New tools for prediction and analysis in the behavioral sciences," Ph. D. Thesis, Harvard University Committee on Applied Mathematics.
177. Werner, B. [1987], "An interactive fuzzy programming system," *Fuzzy Sets and Systems*, 23, pp. 131-147.
178. White, D. A. , and D. A. Sofge, eds. [1992], *Handbook of Intelligent Control: Neuro-Fuzzy and Adaptive Approaches*, Van Nostrand, NY.
179. Yager, R. R. [1980], "On a general class of fuzzy connectives," *Fuzzy Sets and Systems*, 4, no. 3, pp. 235-242.
180. _____ [1986], "A characterization of the extension principle," *Fuzzy Sets and Systems*, 18, no. 3, pp. 205-217.
181. Yager, R. R. , and D. P. Filev [1994], *Essentials of Fuzzy Modeling and Control*, John Wiley, New York, NY.
182. Yagishita, O. , O. Itoh, and M. Sugeno [1985], "Application of fuzzy reasoning to the water purification process," in *Industrial Applications of Fuzzy Control*, ed. M. Sugeno, North-Holland, Amsterdam, pp. 19-40.
183. Yamakawa, T. [1989], "Stabilization of an inverted pendulum by a high-speed fuzzy logic controller hardware systems," *Fuzzy Sets and Systems*, 32, pp. 161-180.
184. Yasunobu, S. , S. Miyamoto, , and H. Ihara [1983], "Fuzzy control for automatic train operation system," *Proc. 4th IFAC/IFIP/IFORS Int. Congress on Control in Transportation Systems*, Baden-Baden.
185. Yasunobu, S. , and S. Miyamoto [1985], "Automatic train operation by predictive fuzzy control," in *Industrial Application of Fuzzy Control*, ed. M. Sugeno, Amsterdam: North-Holland, pp. 1-18.
186. Yasunobu, S. , S. Sekino, , and T. Hasegawa [1987], "Automatic train operation and automatic crane operation systems based on predictive fuzzy control," *Proc. 2nd IFSA Congress*, Tokyo, Japan, pp. 835-838.
187. Yasunobu, S. , and T. Hasegawa [1986], "Evaluation of an automatic container crane operation system based on predictive fuzzy control," *Control Theory Adv. Technol.* , 2, 3, pp. 419-432.
188. Ying, H. [1994], "Sufficient conditions on general fuzzy systems as function approximations," *Automatica*, 30, no. 3, pp. 521-525.
189. Zadeh, L. A. [1962], "From circuit theory to systems theory," *Proc. Institution of Radio Engineers*, 50, pp. 856-865.
190. _____ [1965], "Fuzzy sets," *Informat. Control*, 8, pp. 338-353.
191. _____ [1968], "Fuzzy algorithms," *Information and Control* , 12, no. 2, pp. 94-102.
192. _____ [1971a], "Toward a theory of fuzzy systems," in *Aspects of Network and System Theory*, eds. R. E. Kalman and N. DeClaris.
193. _____ [1971b], "Similarity relations and fuzzy ordering," *Information Sciences*, 3, no. 2, pp. 177-200.
194. _____ [1973], "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Trans. on Systems, Man , and Cybern.* , 3, 1, pp. 28-44.

195. _____ [1975], “ The concept of a linguistic variable and its application to approximate reasoning I , II , III , ” *Information Sciences*, 8, pp. 199-251, pp. 301-357; 9, pp. 43-80.
196. _____ [1978], Fuzzy sets as a basis for a theory of possibility, ” *Fuzzy Sets and Systems*, 1, no. 1, pp. 3-28.
197. _____ [1985], “ Syllogistic reasoning in fuzzy logic and its application to usuality and reasoning with dispositions, ” *IEEE Trans. on Systems, Man, and Cybern.* , 15, 6(1985), pp. 754-763.
198. _____ [1994], “ Fuzzy logic , neural networks and soft computing, ” *Communications of the ACM*, 37, pp. 77-84.
199. _____ [1995], “ Discussion: Probability theory and fuzzy logic are complementary rather than competitive, ” *Technometrics*, 37, no. 3, pp. 271-276.
200. Zeng, X. J. , and M. G. Singh [1994], “ Approximation theory for fuzzy systems—SISO case, ” *IEEE Trans. on Fuzzy Systems*, 2, no. 2, pp. 162-176.
201. _____ [1995], “ Approximation accuracy analysis of fuzzy systems as function approximators, ” personal communication.
202. Zhao, Z. Y. , M. Tomizuka, and S. Isaka [1993], “ Fuzzy gain scheduling of PID controllers, ” *IEEE Trans. on Systems, Man, and Cybern.* , 23, no. 5, pp. 1392-1398.
203. Zimmermann, H. J. [1978], “ Results of empirical studies in fuzzy set theory, ” In: Klir, G. J. , ed. , *Applied General Systems Research*, Plenum Press, NY, pp. 303-312.
204. Zimmermann, H. J. [1991], *Fuzzy Set Theory and Its Applications*, Kluwer Academic Publishers, Boston.

重要词汇英汉对照

	A
α -cut	α -截集
definition	α -截集定义
for fuzzy arithmetic	模糊算术的 α -截集
α -operator	α -算子
Adaptive fuzzy control	自适应模糊控制
classification	自适应模糊控制分类
combined indirect/direct	组合型自适应模糊控制
based on input-output linearization	基于输入—输出线性化的自适应模糊控制
direct	直接型自适应模糊控制
indirect	间接型自适应模糊控制
stable	稳定自适应模糊控制
Adaptation law	自适应律
Additivity axiom	加法公理
Algebraic product	代数积
sum	代数和
Algorithm	算法
fuzzy c-means	模糊 c-均值算法
gradient descent	梯度下降算法
hard c-means	硬 c-均值算法
recursive least squares	递推最小二乘法
Approximate reasoning	近似推理
solution	近似解
Approximation accuracy	逼近精度
Approximator, universal	万能逼近器
Arithmetic, fuzzy	算术, 模糊
Associative condition	关联条件
Averaging operator	平均算子
	B
Back-propagation	反向传播

Ball-and-beam system	球—横木系统
Barbalat's Lemma	Barbalat 的引理
Belief measure	置信度
Bifurcation	分歧点
static	静态分歧点
Hopf	Hopf 分歧点
Binary relation	二元关系
fuzzy	二元模糊关系
Boundary relation	边界关系
fuzzy	模糊边界关系
Boundary layer	边界层
C	
Cartesian product	笛卡儿积
Cement kiln	水泥窑
Center	中心
-average defuzzifier	中心平均解模糊器
of fuzzy set	模糊集的中心
-of-gravity defuzzifier	重心解模糊器
Certainty equivalent principle	确定等效原理
Chain rule	复合函数求导法则
Channel equalization	信道均衡
Characteristic function	特征函数
Chaotic	混沌
Duffing forced oscillation	达芬强迫混沌振荡
time series	混沌时间序列
Chattering	振荡
Classical	经典的
logic	经典逻辑
relation	经典关系
set	经典集合
Closed-loop controller	闭环控制器
Cluster	聚类
Clustering	聚类分析
nearest neighborhood	最近邻聚类分析
fuzzy c-means	模糊 c-均值聚类分析

hard c-means	硬 c-均值聚类分析
Comparison	比较
of defuzzifiers	解模糊器的比较
of fuzzifiers	模糊器的比较
of stochastic and fuzzy linear programming	随机线性规划和模糊线性规划的比较
Complement, fuzzy	模糊补集
Complete	完备的
fuzzy sets	完备模糊集
primitives	初始全集
rule base	完备规则库
Composition of fuzzy relations	模糊关系合成
Compositional rule of inference	组合推理
Computing, soft	软计算
Conditional possibility distribution	条件可能性分布
Conflicting rules	冲突规则
Consistency	一致性
of fuzzy rules	模糊规则的一致性
of fuzzy sets	模糊集的一致性
Constraint matrix	约束矩阵
Continuity	连续性
from above	上连续性
from below	下连续性
of defuzzifier	解模糊器的连续性
of fuzzy rules	模糊规则的连续性
Contradiction	永假公式
Control objectives	控制目标
Controllability	能控性
Conventional control, versus fuzzy control	传统控制与模糊控制的比较
Convex set, fuzzy	模糊凸集
Crossover point	交叉点
Curse of dimensionality	维度灾难
Cylindric extension	柱状扩展
D	
Debate on fuzziness versus probability	模糊与概率的论战
Decision making, fuzzy	模糊决策

Decision variable	决策变量
Decomposition Theorem	分解定理
Defuzzifier	解模糊器
center of gravity	重心解模糊器
center average	中心平均解模糊器
maximum	最大值解模糊器
De Morgan law	德·摩根定律
Digital image stabilizer	数字图像均衡器
Direct adaptive fuzzy control	直接型自适应模糊控制
Discrimination function	差别函数
Distance	距离
Euclidean	欧氏距离
Hamming	海明距离
Distribution, possibility	可能性分布
Drastic product	直积
Drastic sum	直和
Dual role of fuzzy system	模糊系统的双重角色
Duffing forced oscillation	达芬强迫振荡
Dynamic system	动态系统
fuzzy	模糊动态系统
TSK	TSK 动态系统
E	
Equality	贴近
index	贴近度
of fuzzy sets	模糊集的贴近度
Equalization	均衡
Equation	方程
fuzzy	模糊方程
fuzzy relation	模糊关系方程
Equilibrium	均衡器
Error equation	误差方程
Evidence theory	证据理论
Extension principle	扩展原理
F	
Feedforward network	前馈网络

Fuzzifier	模糊器
Gaussian	高斯模糊器
singletom	单值模糊器
triangular	三角形模糊器
Fuzziness versus probability	模糊性与概率
Fuzzy	模糊
arithmetic	模糊算术
c-means algorithm	模糊 c-均值算法
c-partition	模糊 c-划分
complement	模糊补集
controller	模糊控制器
equation	模糊方程
IF-THEN rule	模糊 IF-THEN 规则
inference engine	模糊推理机
intersection	模糊交集
logic	模糊逻辑
number	模糊数
proposition	模糊命题
ranking	模糊排序
relation	模糊关系
rule base	模糊规则库
set	模糊集
union	模糊并集
Fuzzy system	模糊系统
with center average defuzzifier	带有中心平均解模糊器的模糊系统
with maximum defuzzifier	带有最大值解模糊器的模糊系统
G	
Gain scheduling of PID controller	PID 控制器的增益调度
Gaussian	高斯
fuzzifier	高斯模糊器
membership function	高斯隶属度函数
noise	高斯噪声
Generalized	广义的
hypothetical syllogism	广义假言推理
mean	广义均值

modus ponens	广义取式推理
modus tollens	广义拒式推理
Geometric formulation of equalization	均衡的几何公式
Geometric interpretation of projection	投影的几何含义
Gödel implication	Gödel 含义
Gradient descent algorithm	梯度下降法
Gradual rule	渐进规则
H	
Hedge, linguistic	限制, 语言变量
Heuristics	启发式规则
Height of fuzzy set	模糊集的高度
Hierarchical fuzzy system	多级模糊系统
History of fuzzy theory and applications	模糊理论及其应用历史
Human	人类
expertise	人类专家
knowledge	人类知识
operator	算子
Hypothetical syllogism	假言推理
I	
Identification of nonlinear system	非线性系统辨识
Implication	含义
Dienes-Rescher	Dienes-Rescher 含义
Gödel	Gödel 含义
Lukasiewicz	Lukasiewicz 含义
Mamdani	Mamdani 含义
Zadeh	Zadeh 含义
Index	指标
equality	贴近度
robustness	稳健性度量
solvability	可解度
Indirect adaptive fuzzy control	间接型自适应模糊控制
Industrial process control	工业过程控制
Inference	推理
composition based	组合推理
individual-rule based	独立推理

Inference engine	推理机
Dienes-Rescher	Dienes-Rescher 推理机
Lukasiewicz	Lukasiewicz 推理机
minimum	最小推理机
product	乘积推理机
Zadeh	Zadeh 推理机
Information	信息
Initial parameter choosing	初始参数选择
Input-output	输入—输出
data	输入—输出数据
linearization	输入—输出线性化
stability	输入—输出稳定性
Internal dynamics	内在动力学
Interpretation of fuzzy IF-THEN rules	模糊 IF-THEN 规则的解释
Intersection, fuzzy	模糊交集
Dombi class	Dombi 的模糊交
Duboise-Prade class	Duboise-Prade 的模糊交
Yager class	Yager 的模糊交
Interval analysis	区间分析
Interval-valued function	区间值函数
Intuition	直觉
Inverted pendulum	倒立摆
Involutive fuzzy complement	自反的模糊补集
	J
Jacobian matrix	雅可比矩阵
	K
Knowledge	知识
conscious	显性知识
control	控制知识
human	人类知识
plant	被控对象知识
subconscious	隐性知识
Knowledge based system	知识库系统
	L
Language, natural	自然语言

Largest	最大
s-norm	最大 s-范数
t-norm	最大 t-范数
Law	法则
De Morgan	德·摩根定律
of the excluded middle	排中律
Learning algorithm	学习算法
back-propagation	反向传播学习算法
gradient descent	梯度下降学习算法
recursive least squares	递推最小二乘法学习算法
Linear programming	线性规划
with fuzzy constraint coefficients	模糊约束系数型线性规划
with fuzzy objective coefficients	模糊目标系数型线性规划
with fuzzy resources	模糊资源型线性规划
Linear in the parameter	线性参数
Linguistic	语言
hedge	语言限制
rule	语言规则
trajectory	语言轨迹
variable	语言变量
Lipschitz condition	Lipschitz 条件
Logic	逻辑
classical	经典逻辑
fuzzy	模糊逻辑
Logic formulas	逻辑公式
Logic function	逻辑函数
Lower probability	概率下界
Lyapunov	Lyapunov
equation	Lyapunov 方程
function	Lyapunov 函数
stability	Lyapunov 稳定性
stability theorem	Lyapunov 稳定性定理
synthesis approach	综合法
M	
Marginal possibility distribution	边际可能性分布

Matrix Inequality Lemma	矩阵不等式引理
Max-min composition	最大—最小合成
Max-product composition	最大—乘积合成
Maximum defuzzifier	最大值解模糊器
Mean Value Theorem	均值定理
Measure	测度
belief	信任性测度
plausibility	似真性测度
necessity	必要性测度
possibility	可能性测度
probability	概率测度
Membership function	隶属度函数
Gaussian	高斯隶属度函数
pseudo-trapezoid	四边形隶属度函数
trapezoid	梯形隶属度函数
triangular	三角形隶属度函数
Minimum approximation error	最小逼近误差
Modus ponens	取式推理
generalized	广义取式推理
Modus tollens	拒式推理
generalized	广义拒式推理
Multiple objective linear programming	多目标线性规划
N	
Natural language	自然语言
Nearest neighborhood clustering	最近邻聚类分析
Necessity measure	必要性测度
Nonfuzzy rule	非模糊规则
Noninteractiveness	非关联性
Nonlinear channel equalization	非线性信道均衡
Nonlinear control theory	非线性控制论
Normal fuzzy set	标准模糊集
O	
Objective coefficient	目标系数
Observability	能观性
Open-loop controller	开环控制器

Optimal	最优
equalizer	最优均衡器
fuzzy controller	最优模糊控制器
fuzzy system	最优模糊系统
Order between fuzzy sets	模糊集的顺序
Or-rule	或规则
P	
φ -operator	φ -算子
Parameter projection	参数投影
Partial rule	部分规则
Partition	划分
hard	硬划分
fuzzy	模糊划分
Pattern recognition	模式识别
Performance criterion	性能指标
Phase plane	相平面
PID controller	PID 控制器
Plausibility measure	似真性测度
Positive real	正实域性
Possibility	可能性
distribution	可能性分布
measure	可能性度量
theory	可能性理论
versus probability	可能性与概率
Pontryagin Minimum Principle	Pontryagin 最小值原理
Prediction, time series	时间序列预测
Primitive	初始
Probability	概率
measure	概率测度
versus possibility	概率与可能性
Product	乘积
algebraic	代数积
Cartesian	笛卡儿积
drastic	直积
Einstein	爱因斯坦积

Product inference engine	乘积推理机
Programming, linear	线性规划
Projection algorithm	投影算法
Projection of fuzzy relation	模糊关系的投影
Proposition, fuzzy	模糊命题
Propositional variable	命题变量
Pure fuzzy system	纯模糊系统
Q	
Questionnaire	调查表
R	
Radius	半径
Ranking, fuzzy	模糊排序
Reasoning	推理
Recursive least squares	递推最小二乘法
Reference model	参考模型
Relation	关系
binary	二元关系
classical	经典关系
fuzzy	模糊关系
Relation equation, fuzzy	模糊关系方程
Relational matrix, fuzzy	模糊关系矩阵
Robust fuzzy control	稳健模糊控制
S	
Sliding control	滑动控制
Smallest	最小
s-norm	最小 s-范数
t-norm	最小 t-范数
Small Gain Theorem	小增益原理
Soft computing	软计算
Solvability index	可解度
Stable	稳定的
direct adaptive fuzzy control	稳定的直接型自适应模糊控制
indirect adaptive fuzzy control	稳定的间接型自适应模糊控制
Stability	稳定性
input-output	输入—输出稳定性

- | | |
|---------------------------------------|----------------------|
| Lyapunov | Lyapunov 稳定性 |
| Stone-Weierstrass Theorem | Stone-Weierstrass 定理 |
| Subadditivity | 次可加性 |
| Subway train | 地铁 |
| Sum | 和 |
| algebraic | 代数和 |
| drastic | 直和 |
| Einstein | 爱因斯坦和 |
| Superadditivity | 超可加性 |
| Supervisory control | 监督控制 |
| Support of fuzzy set | 模糊集的支撑集 |
| Sup-star composition | Sup- \star 合成 |
| Syllogism | 推理 |
| hypothetical | 假言推理 |
| generalized hypothetical | 广义假言推理 |
| T | |
| Table look-up scheme | 查表法 |
| Takagi-Sugeno-Kang (TSK) fuzzy system | TSK 模糊系统 |
| Tautology | 重复 |
| Taylor series expansion | 泰勒展开 |
| T-norm | t-范数 |
| Time series prediction | 时间序列预测 |
| Training | 训练 |
| Trapezoidal | 梯形的 |
| fuzzy number | 梯形模糊数 |
| membership function | 梯形隶属度函数 |
| Trail-and-error approach | 试错法 |
| Triangular | 三角形的 |
| fuzzy number | 三角形模糊数 |
| membership function | 三角形隶属度函数 |
| Truck backer-upper control | 倒车控制 |
| U | |
| Uncertainty | 不确定性 |
| Union, fuzzy | 模糊并集 |
| Dombi class | Dombi 的模糊并 |

Dubois-Prade class	Dubois-Prade 的模糊并
Yager class	Yager 的模糊并
Universal approximator	万能逼近器
Universal set	全集
Universe of discourse	论域
Upper probability	概率上界
V	
Variable	变量
linguistic	语言变量
numerical	数值变量
W	
Washing machine, fuzzy	模糊洗衣机
Wastewater treatment	污水处理
Weighted average	加权平均
What are	什么是
fuzzy control	什么是模糊控制
fuzzy systems	什么是模糊系统
research fields in fuzzy theory	什么是模糊理论的研究领域
Why	为什么
adaptive fuzzy control	为什么研究自适应模糊控制
fuzzy system	为什么研究模糊系统
Y	
Yager class	Yager 的
fuzzy complement	Yager 的模糊补
s-norm	Yager 的 s-范数
t-norm	Yager 的 t-范数
Z	
Zadeh implication	Zadeh 含义
Zadeh inference engine	Zadeh 推理机



教师服务登记表

尊敬的老师：感谢您购买本书。

书名：《 _____ 》

清华大学出版社植根于“清华”这座久负盛名的高等学府，秉承清华人“自强不息，厚德载物”的人文精神，规划出版了大批优秀的经管类教材，受到国家有关部门和全国各院校广大师生的一致好评，树立了清华出版的品牌。为了进一步加强我们与老师之间的联系与沟通，请您协助填写下表并寄回，以便为您提供最新的出版信息，同时我们还会为您的教学工作以及论著或译著的出版提供尽可能的帮助。欢迎访问清华商务出版在线 <http://tupub.com>，请按作者会员注册登陆并修改信息。

姓名		性别		年龄	
职务		职称		学历	
联系电话		手机		传真	
学校				省市	
院系					
通信地址				邮编	
E-mail:					
毕业学校					
您所教授的课程				学生层次	
您的研究领域					
您希望在清华出版的教材或图书?					
最需要我们提供哪些方面的信息					
对清华文泉工作建议					

通信地址：北京清华大学出版社（海淀区双清路学研大厦 A605）文泉公司经管研发部（100084）

办公地址：北京清华大学校内清华文泉公司（南门往北 200 米路西绿楼）

选题策划：徐学军 010-62781829 [E-mail:xuxuejun@mail.tsinghua.edu.cn](mailto:xuxuejun@mail.tsinghua.edu.cn)

市场推广：白杰 010-62794748 [E-mail:baij@tup.tsinghua.edu.cn](mailto:baij@tup.tsinghua.edu.cn)

传真：010-62794748 清华商务出版在线 <http://www.tupub.com>



读者服务登记表

尊敬的读者感谢您购买本书。

书名：《 _____ 》

清华大学出版社植根于“清华”这座久负盛名的高等学府，秉承清华人“自强不息，厚德载物”的人文精神，规划出版了大批优秀的经管类教材，受到国家有关部门和全国各院校广大师生的一致好评，树立了清华出版的品牌。我们很愿意成为您的朋友，请您协助填写下表并寄回，以便为您提供最新的出版信息。欢迎访问清华商务出版在线 <http://tupub.com>，请按读者会员注册登陆并修改信息。

姓名		性别		年龄	
职务		职称		学历	
联系电话		手机		传真	
单位				省市	
部门					
通信地址				邮编	
E-mail:					
毕业学校					
您对本书的评价					
您感兴趣的经济管理类读物有	<input type="checkbox"/> 经济学 <input type="checkbox"/> 管理学 <input type="checkbox"/> 财会 <input type="checkbox"/> 金融 <input type="checkbox"/> 保险 <input type="checkbox"/> 营销 <input type="checkbox"/> 管理决策 <input type="checkbox"/> 电子商务 <input type="checkbox"/> 物流 <input type="checkbox"/> 房地产 <input type="checkbox"/> 经济热点丛书 <input type="checkbox"/> 项目管理(综合) <input type="checkbox"/> 项目管理(软件业) <input type="checkbox"/> 人力资源管理 <input type="checkbox"/> 其它				
您就职单位的行业性质					
您主要从哪些渠道获取经济管理信息	<input type="checkbox"/> 财经类网站 <input type="checkbox"/> 图书类网站 <input type="checkbox"/> 报刊、杂志 <input type="checkbox"/> 图书馆 <input type="checkbox"/> 书店 广播、电视 <input type="checkbox"/> 其他				
您急需的图书是					
对清华文泉工作建议					

通信地址：北京清华大学出版社（海淀区双清路学研大厦 A605）文泉公司经管研发部（100084）

办公地址：北京清华大学校内清华文泉公司（南门往北 200 米路西绿楼）

编辑部咨询电话：010-62794748 baij@tup.tsinghua.edu.cn

发行部咨询电话：010-62776969 e-sale@tup.tsinghua.edu.cn，

编辑部传真：010-62794748 清华商务出版在线 <http://www.tupub.com>