



普通高等教育“十一五”国家级规划教材

数学模型

(第四版)

姜启源 谢金星 叶 俊 编

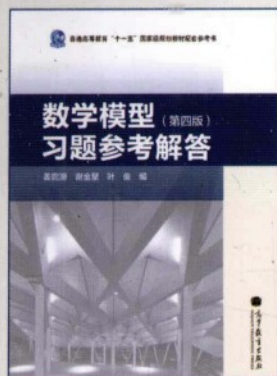
Mathematical Model



高等教育出版社
HIGHER EDUCATION PRESS

本书第一版荣获第二届全国高等学校优秀教材优秀奖

Mathematical Model



ISBN 978-7-04-031150-1



9 787040 311501 >

定价 44.00 元

普通高等教育“十一五”国家级规划教材

数 学 模 型

Shuxue Moxing

(第四版)



高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING



内容提要

本书第一、二、三版分别出版于1987年、1993年和2003年。基于作者20多年来从事数学建模教学、组织数学建模竞赛、开设数学实验课程以及编写相关图书的经验,参考国内外数学建模教材和教学单元,第四版在保持前三版基本结构和风格的基础上,进行补充与修订:增加了一些实用性较强、生活气息浓烈、数学推导简化的案例,改写、合并、调整了若干案例和章节,删除了个别案例,并对习题作了相应的修订。

本书可作为高等学校各专业学生数学建模课程的教材和参加数学建模竞赛的辅导材料,以及科技工作者的参考书。

图书在版编目(CIP)数据

数学模型 / 姜启源, 谢金星, 叶俊编. -- 4版. --
北京: 高等教育出版社, 2011.1
ISBN 978-7-04-031150-1

I. ①数… II. ①姜…②谢…③叶… III. ①数学模型 - 高等学校 - 教材 IV. ①O22

中国版本图书馆CIP数据核字(2010)第234422号

策划编辑	李艳馥	责任编辑	李艳馥	李蕊	封面设计	赵阳
责任绘图	尹文军	版式设计	张岚		责任校对	杨雪莲
责任印制	朱学忠					

出版发行 高等教育出版社
社 址 北京市西城区德外大街4号
邮政编码 100120

购书热线 010-58581118
咨询电话 400-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>

经 销 蓝色畅想图书发行有限公司
印 刷 北京鑫海金澳胶印有限公司

网上订购 <http://www.landraco.com>
<http://www.landraco.com.cn>
畅想教育 <http://www.widedu.com>

开 本 787×960 1/16
印 张 30.5
字 数 570 000

版 次 1987年4月第1版
2011年1月第4版
印 次 2011年1月第1次印刷
定 价 44.00元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 31150-00

第四版序

自本书第一版1987年问世以来,已过了23年。在此期间,“数学建模”这一课程的开设已从无到有地几乎遍及全国高校。这当然是令人高兴的事。

就体例而言,《数学建模》第四版和以前的三版没有大的差异,但在内容上有了较大的扩展。这是考虑到由于使用者增加而产生的多元化要求,同时也是编者多年不断积累经验的结果。

广义来说,数学建模是一个完整的过程,在明确问题以后,从弄清机理,搜集数据,建模,求解,对答案的解释,到修改模型,如是循环。但对一个入门的教材而言,重点似乎仍应放在建模、求解(对现成软件的合理使用)以及答案的解释上。

这是就内容而言。如果选定了内容,就是教与学的方法问题。在这方面,我想再次强调我在第三版序言中所引爱因斯坦的一句话:想象力比知识更重要。除此之外,我还想补充一句:数学建模本质上是一种艺术。教师要教好、学生要学好这门课,多看、多做、多讨论是很有必要的。

萧树铁

2010年8月

数学建模
第四版
PDG

第四版前言

近几十年来,随着科技的发展和社会的进步,数学这门基础科学不仅在工程技术等传统应用领域发挥着越来越重要的作用,而且迅速地向一些新的领域渗透,形成了许多交叉学科,如计量经济学、人口控制论、生物数学、地质数学等等。数学与计算机技术相结合,已产生出一种重要的、可以实现的技术,成为当代高新技术的重要组成部分。“高技术本质上是数学技术”的观点已被越来越多的人所接受。

在用数学方法解决现实生活中碰到的问题,或与其他学科相结合形成新学科的过程中,首要的和关键的一步是用数学的语言来表述所研究的对象,即建立数学模型。在此基础上才能运用数学的理论和方法进行分析和计算,为解决实际问题给出定量的结果或定性的数量依据。计算和建模正在成为数学科学技术转化的主要途径。

教育,特别是高等教育应该及时反映并满足社会发展的实际需要。数学建模进入大学课堂,既顺应时代发展的潮流,也符合教育改革的要求。对于数学教育而言,既应该让学生掌握准确快捷的计算方法和严密的逻辑推理,也需要培养他们用数学工具分析和解决实际问题的意识和能力。传统的数学教学体系和内容无疑偏重于前者,开设数学建模课程则是加强后者的一种有益的、成功的尝试。

数学建模课程是20世纪80年代初进入我国大学的。本书第一版出版于1987年,是我国第一本数学建模教材,当时只有少数几所学校的数学专业开设这门课程。本书第二版出版于1993年,其时开设数学建模课程的学校增加到几十所,并且开始推向非数学专业。本书第三版出版于2003年,其时数学建模教学推广到几百所学校,教材也有几十本。到目前,内容、形式不尽相同的数学建模课程已在千余所高校开设,正式出版的教材和参考书达200多本。回顾数学建模进入我国大学课堂的20多年,可以说已从星星之火迅速发展成燎原之势,这期间产生重要影响的因素,至少有以下几项:

1. 1992年开始由教育部高等教育司和中国工业与应用数学学会主办的、每年一届的全国大学生数学建模竞赛,得到广大同学的热烈欢迎,以及教育部门和教师们的热情关心和支持,成为我国高校规模最大的课外科技活动。在全国竞赛的推动下,许多学校、地区也纷纷组织竞赛。不少学校的同学自发成立数学建模协会等组织,举办各种形式的课外活动。竞赛促进了数学建模教学的开展,教

学又扩大了受益面,为竞赛奠定了坚实的基础。

2. 数学建模教学和竞赛的开展,对数学教学体系、内容和方法改革起了积极的推动作用,得到众多教育界人士和教师们的认可。将数学建模的思想和方法有机地融入大学数学主干课程中去的研究与实践已经起步,教数学建模课和教数学主干课的教师互相结合与交流,在教学上得以相互借鉴与促进。

3. 两年一届的全国数学建模教学与应用会议及各个地区、各种类型的研讨班、培训班吸引了越来越多的教师,经验交流、专家讲学促进了各校数学建模教学质量的不断提高。

4. 从21世纪初开始,越来越多的学校开设了数学实验课程,对数学建模的教学内容和方法既有很大启示,也提出新的挑战。

5. 计算机技术及数学软件的飞速发展和普及,为改善、丰富数学建模课程的内容提供了良好的外部条件。

本书编者基于长期从事数学建模教学、数学建模竞赛的组织,以及近年讲授数学实验课的经验,考虑到上述几项因素,听取了一些同行的意见,又参考国内外一些数学建模教材、教学单元和辅助材料,决定推出第四版。该版在保持第三版基本结构和风格的基础上,有以下几项较大的补充与修订:

1. 增加了14个实用性较强、生活气息浓烈、数学推导简化的案例,其中有如何施救药物中毒、光盘的数据容量、天气预报的评价、CT技术的图像重建、学生作弊现象的调查和估计,还有进攻与撤退的抉择、“一口价”的战略等博弈模型。

2. 改写、合并、调整了若干案例和章节,其中有公平的席位分配、消费者的选择、生产者的决策、人口的预测和控制、原子弹爆炸的能量估计等。

3. 出于内容更新和全书篇幅的考虑,删去了8个案例,其中有录像机计数器的用途、汽车刹车距离、动物的身长和体重、广告中的学问等。

4. 修订了原有的习题和综合题目,修订了习题参考解答(另行同期出版)。习题中标以*号的是需要读者自己作出假设及建模的题目。综合题目可供教师选为期中或期末的作业。

5. 修订了全部章节的多媒体课件(另行同期出版)。数学建模是一门内容活泼、信息量大,基本不涉及新的数学概念,不需要太多数学推导的课程,有时还可以用数学软件作数值计算和图形演示,特别适合多媒体教学。目前提供的课件可能存在诸多不完善之处,今后将不定期地推出修订版本。

最后,对使用本书作为教材或参考书的教师提几点建议:

1. 数学建模课基本上是案例式教学,内容连贯性不强,书中的章节可以跳跃式地选用,未在课堂上讲授的内容可作为课外阅读材料。从培养学生数学建模的意识、方法和能力的角度看,目前这种阅读材料不是太多,而是少了,因此本

书有意包含了一些可能不会在课上讲授的内容。

2. 对于课程在 40 学时以上的非数学类本科生,可讲授第 1~6,8~10 章的部分内容,40 学时以下及高职高专学生,可讲授第 1~3,5~6,8 章的部分内容。

3. 对于课程在 40 学时以上的数学类本科生,可讲授第 1~12 章的部分内容,其中第 4,10 章是否讲授,可考虑运筹、统计课的设置及其内容与此有无重复。

本书第 4 章和第 11 章 11.1~11.4 由谢金星编写,第 10 章和第 9 章 9.7 由叶俊编写,其余各章由姜启源编写、增补和修订,全书由姜启源统稿。

1983 年,萧树铁先生在清华大学首次为本科生讲授数学模型课程,他是我国高校开设数学模型课程的创始人之一,20 多年来一直关心着课程的发展及教材建设,曾为本书第一、二、三版作序,现在又为第四版撰写序言,在此向他表示衷心的感谢。

我们还要向许多在使用本书第一、二、三版的过程中提出宝贵建议的教师们致谢,希望大家共同努力,为数学建模教学和竞赛活动取得更大成绩继续奋斗。

编者
2010.9



目 录

第1章 建立数学模型	1
1.1 从现实对象到数学模型	1
1.2 数学建模的重要意义	4
1.3 建模示例之一 椅子能在不平的地面上放稳吗	6
1.4 建模示例之二 商人们怎样安全过河	8
1.5 建模示例之三 如何施救药物中毒	9
1.6 数学建模的基本方法和步骤	13
1.7 数学模型的特点和分类	16
1.8 数学建模能力的培养	18
习题	20
第2章 初等模型	23
2.1 光盘的数据容量	23
2.2 双层玻璃窗的能效	27
2.3 划艇比赛的成绩	29
2.4 实物交换	31
2.5 污水均流池的设计	34
2.6 交通流与道路通行能力	38
2.7 核军备竞赛	43
2.8 扬帆远航	47
2.9 天气预报的评价	49
习题	55
第3章 简单的优化模型	58
3.1 存贮模型	59
3.2 生猪的出售时机	63
3.3 森林救火	65
3.4 消费者的选择	67
3.5 生产者的决策	72
3.6 血管分支	75
3.7 冰山运输	77
习题	82

第4章 数学规划模型	85
4.1 奶制品的生产与销售	85
4.2 自来水输送与货机装运	95
4.3 汽车生产与原油采购	101
4.4 接力队的选拔与选课策略	108
4.5 饮料厂的生产与检修	117
4.6 钢管和易拉罐下料	122
习题	130
第5章 微分方程模型	136
5.1 传染病模型	136
5.2 经济增长模型	145
5.3 正规战与游击战	149
5.4 药物在体内的分布与排除	154
5.5 香烟过滤嘴的作用	158
5.6 人口的预测和控制	163
5.7 烟雾的扩散与消失	173
5.8 万有引力定律的发现	176
习题	180
第6章 代数方程与差分方程模型	184
6.1 投入产出模型	184
6.2 CT技术的图像重建	187
6.3 原子弹爆炸的能量估计与量纲分析	191
6.4 市场经济中的蛛网模型	201
6.5 减肥计划——节食与运动	206
6.6 按年龄分组的人口模型	208
习题	213
第7章 稳定性模型	215
7.1 捕鱼业的持续收获	215
7.2 军备竞赛	219
7.3 种群的相互竞争	222
7.4 种群的相互依存	228
7.5 食饵-捕食者模型	230
7.6 差分形式的阻滞增长模型	236
7.7 微分方程与差分方程稳定性理论简介	242
习题	247

第8章 离散模型	249
8.1 层次分析模型	249
8.2 循环比赛的名次	269
8.3 社会经济系统的冲量过程	273
8.4 公平的席位分配	278
8.5 存在公正的选举规则吗	286
8.6 价格指数	292
习题	296
第9章 概率模型	300
9.1 传送系统的效率	300
9.2 报童的诀窍	302
9.3 随机存贮策略	304
9.4 轧钢中的浪费	307
9.5 随机人口模型	310
9.6 航空公司的预订票策略	313
9.7 学生作弊现象的调查和估计	317
习题	323
第10章 统计回归模型	325
10.1 牙膏的销售量	325
10.2 软件开发人员的薪金	332
10.3 酶促反应	338
10.4 投资额与生产总值和物价指数	346
10.5 教学评估	352
10.6 冠心病与年龄	356
习题	365
第11章 博弈模型	373
11.1 进攻与撤退的抉择	374
11.2 让报童订购更多的报纸	379
11.3 “一口价”的战略	382
11.4 不患寡而患不均	386
11.5 效益的合理分配	389
11.6 加权投票中权力的度量	397
习题	410
第12章 马氏链模型	414
12.1 健康与疾病	414

12.2	钢琴销售的存贮策略	419
12.3	基因遗传	422
12.4	等级结构	425
12.5	资金流通	433
	习题	436
第 13 章	动态优化模型	439
13.1	速降线与短程线	439
13.2	生产计划的制订	444
13.3	国民收入的增长	446
13.4	渔船出海	448
13.5	赛跑的速度	451
13.6	多阶段最优生产计划	456
	习题	462
	综合题目	464
	参考文献	470



第1章

建立数学模型

随着科学技术的迅速发展,数学模型这个词汇越来越多地出现在现代人的生产、工作和社会活动中.电气工程师必须建立所要控制的生产过程的数学模型,用这个模型对控制装置作出相应的设计和计算,才能实现有效的过程控制.气象工作者为了得到准确的天气预报,一刻也离不开根据气象站、气象卫星汇集的气压、雨量、风速等资料建立的数学模型.生理医学专家有了药物浓度在人体内随时间和空间变化的数学模型,就可以分析药物的疗效,有效地指导临床用药.城市规划工作者需要建立一个包括人口、经济、交通、环境等大系统的数学模型,为领导层对城市发展规划的决策提供科学根据.厂长经理们要是能够根据产品的需求状况、生产条件和成本、贮存费用等信息,筹划出一个合理安排生产和销售数学模型,一定可以获得更大的经济效益.就是在日常活动如访友、采购当中,人们也会谈论找一个数学模型,优化一下出行的路线.对于广大的科学技术人员和应用数学工作者来说,建立数学模型是沟通摆在面前的实际问题与他们掌握的数学工具之间联系的一座必不可少的桥梁.

本章作为全书的导言和数学模型的概述,主要讨论建立数学模型的意义、方法和步骤,给读者以建立数学模型的全面的、初步的了解.1.1节介绍现实对象和它的模型的关系,给出一些模型形式,说明什么是数学模型;1.2节阐述建立数学模型的重要意义;1.3~1.5节通过几个示例说明用数学语言和数学方法表述和解决实际问题,即建立数学模型的过程;1.6节阐述建立数学模型的一般方法和步骤;1.7节介绍数学模型的特点及数学模型的分类;1.8节讨论建立数学模型能力的培养.

1.1 从现实对象到数学模型

人类生活在丰富多彩、变化万千的现实世界里,无时无刻不在运用智慧和力量去认识、利用、改造这个世界,从而不断地创造出日新月异、五彩缤纷的物质文明和精神文明.博览会常常是集中展示这些成果的场所之一,那些五光十色、精美绝伦的展品给我们留下了深刻的印象.工业博览会上,豪华、舒适的新型汽车

叫人赞叹不已；农业博览会上，硕大、娇艳的各种水果令人流连忘返；科技展厅里，大型水电站模型雄伟壮观，人造卫星模型高高耸立，清晰的数字和图表显示着电力工业的迅速发展，和整面墙壁一样大的地图上鲜明地标出了新建的铁路和新辟的航线，核电站工程的彩色巨照前，手持原子结构模型的讲解员深入浅出地介绍反应堆的运行机理；电影演播室里，播放着现代化炼钢厂实现生产自动控制的科技影片，其中既有火花四溅的钢坯浇铸情景，也有展示计算机管理和控制的框图、公式和程序。

参观博览会，像汽车、水果那些原封不动地从现实世界搬到展厅里的物品固然给人以亲切真实的感受，可是从开阔眼界、丰富知识的角度看，电站、卫星、铁路、钢厂……这些在现实世界被人们认识、建造、控制的对象，以它们的各种形式的模型——实物模型、照片、图表、公式、程序……汇集在人们面前，这些模型在短短几小时里所起的作用，恐怕是置身现实世界多少天也无法做到的。

与形形色色的模型相对应，它们在现实世界里的原始参照物通称为原型。本节先讨论原型和模型，特别是和数学模型的关系，再介绍数学模型的意义。

原型和模型 原型(Prototype)和模型(Model)是一对对偶体。原型指人们在现实世界里关心、研究或者从事生产、管理的实际对象。在科技领域通常使用系统(System)、过程(Process)等词汇，如机械系统、电力系统、生态系统、生命系统、社会经济系统，又如钢铁冶炼过程、导弹飞行过程、化学反应过程、污染扩散过程、生产销售过程、计划决策过程等。本书所述的现实对象、研究对象、实际问题等均指原型。模型则指为了某个特定目的将原型的某一部分信息简缩、提炼而构造的原型替代物。

这里特别强调构造模型的目的性。模型不是原型原封不动的复制品，原型有各个方面和各种层次的特征，而模型只要求反映与某种目的有关的那些方面和层次。一个原型，为了不同的目的可以有许多不同的模型。如放在展厅里的飞机模型应该在外形上逼真，但是不一定会飞。而参加航模竞赛的模型飞机要具有良好的飞行性能，在外观上不必苛求。至于在飞机设计、试制过程中用到的数学模型和计算机模拟，则只要求在数量规律上真实反映飞机的飞行动态特性，毫不涉及飞机的实体。所以模型的基本特征是由构造模型的目的决定的。

我们已经看到模型有各种形式。按照模型替代原型的方式来分类，模型可以分为物质模型(形象模型)和理想模型(抽象模型)。前者包括直观模型、物理模型等，后者包括思维模型、符号模型、数学模型等。

直观模型 指那些供展览用的实物模型，以及玩具、照片等，通常是把原型的尺寸按比例缩小或放大，主要追求外观上的逼真。这类模型的效果是一目了然的。

物理模型 主要指科技工作者为了一定目的根据相似原理构造的模型,它不仅可以显示原型的外形或某些特征,而且可以用来进行模拟实验,间接地研究原型的某些规律.如波浪水箱中的舰艇模型用来模拟波浪冲击下舰艇的航行性能,风洞中的飞机模型用来试验飞机在气流中的空气动力学特性.有些现象直接用原型研究非常困难,更可借助于这类模型,如地震模拟装置、核爆炸反应模拟设备等.应注意验证原型与模型间的相似关系,以确定模拟实验结果的可靠性.物理模型常可得到实用上很有价值的结果,但也存在成本高、时间长、不灵活等缺点.

思维模型 指通过人们对原型的反复认识,将获取的知识以经验形式直接贮存于人脑中,从而可以根据思维或直觉作出相应的决策.如汽车司机对方向盘的操纵、一些技艺性较强的工种(如钳工)的操作,大体上是靠这类模型进行的.通常说的某些领导者凭经验作决策也是如此.思维模型便于接受,也可以在一定条件下获得满意的结果,但是它往往带有模糊性、片面性、主观性、偶然性等缺点,难以对它的假设条件进行检验,并且不便于人们的相互沟通.

符号模型 是在一些约定或假设下借助于专门的符号、线条等,按一定形式组合起来描述原型.如地图、电路图、化学结构式等,具有简明、方便、目的性强及非量化等特点.

本书要专门讨论的**数学模型**则是由数字、字母或其他数学符号组成的,描述现实对象数量规律的数学公式、图形或算法.

什么是数学模型 其实你早在学习初等代数的时候就已经碰到过数学模型了.当然其中许多问题是老师为了教会学生知识而人为设置的.譬如你一定解过这样的所谓“航行问题”:

甲乙两地相距 750 km,船从甲到乙顺水航行需 30 h,从乙到甲逆水航行需 50 h,问船速、水速各若干?

用 x, y 分别代表船速和水速,可以列出方程

$$(x+y) \cdot 30 = 750, \quad (x-y) \cdot 50 = 750$$

实际上,这组方程就是上述航行问题的数学模型.列出方程,原问题已转化为纯粹的数学问题.方程的解 $x = 20$ km/h, $y = 5$ km/h, 最终给出了航行问题的答案.

当然,真正实际问题的数学模型通常要复杂得多,但是建立数学模型的基本内容已经包含在解这个代数应用题的过程中了.那就是:根据建立数学模型的目的和问题的背景作出必要的简化假设(航行中设船速和水速为常数);用字母表示待求的未知量(x, y 代表船速和水速);利用相应的物理或其他规律(匀速运动的距离等于速度乘以时间),列出数学式子(二元一次方程);求出数学上的解答($x = 20, y = 5$);用这个答案解释原问题(船速和水速分别为 20 km/h 和

5 km/h);最后还要用实际现象来验证上述结果.

一般地说,数学模型可以描述为,对于现实世界的一个特定对象,为了一个特定目的,根据特有的内在规律,作出一些必要的简化假设,运用适当的数学工具,得到的一个数学结构.

需要指出,本书的重点不在于介绍现实对象的数学模型(Mathematical Model)是什么样子,而是要讨论建立数学模型(Mathematical Modelling)的全过程.建立数学模型下面简称为数学建模或建模.

与数学模型有密切关系的数学模拟,主要指运用数字式计算机的计算机模拟(Computer Simulation).它根据实际系统或过程的特性,按照一定的数学规律用计算机程序语言模拟实际运行状况,并依据大量模拟结果对系统或过程进行定量分析.例如通过各种工件在不同机器上按一定工艺顺序加工的模拟,能够识别生产过程中的瓶颈环节;通过高速公路上交通流的模拟,可以分析车辆在路段上的分布特别是堵塞的状况.与用物理模型的模拟实验相比,计算机模拟有明显的优点:成本低、时间短、重复性高、灵活性强.有人把计算机模拟作为建立数学模型的手段之一,但是数学模型在某种意义上描述了对象内在特性的数量关系,其结果容易推广,特别是得到了解析形式答案时,更易推广.而计算机模拟则完全模仿对象的实际演变过程,难以从得到的数字结果分析对象的内在规律.当然,对于那些因内部机理过于复杂,目前尚难以建立数学模型的实际对象,用计算机模拟获得一些定量结果,可称是解决问题的有效手段.

1.2 数学建模的重要意义

数学,作为一门研究现实世界数量关系和空间形式的科学,在它产生和发展的历史长河中,一直是和人们生活的实际需要密切相关的.作为用数学方法解决实际问题的第一步,数学建模自然有着与数学同样悠久的历史.两千多年以前创立的欧几里得几何,17世纪牛顿发现的万有引力定律,都是科学发展史上数学建模的成功范例.

进入20世纪以来,随着数学以空前的广度和深度向一切领域的渗透,以及电子计算机的出现与飞速发展,数学建模越来越受到人们的重视,可以从以下几方面来看数学建模在现实世界中的重要意义.

1) 在一般工程技术领域,数学建模仍然大有用武之地.

在以声、光、热、力、电这些物理学科为基础的诸如机械、电机、土木、水利等工程技术领域中,数学建模的普遍性和重要性不言而喻.虽然这里的基本模型是已有的,但是由于新技术、新工艺的不断涌现,提出了许多需要用数学方法解决的新问题;高速、大型计算机的飞速发展,使得过去即便有了数学模型也无法求

解的课题(如大型水坝的应力计算,中长期天气预报等)迎刃而解;建立在数学模型和计算机模拟基础上的 CAD 技术,以其快速、经济、方便等优势,很大程度上替代了传统工程设计中的现场实验、物理模拟等手段。

2) 在高新技术领域,数学建模几乎是必不可少的工具.

无论是发展通信、航天、微电子、自动化等高新技术本身,还是将高新技术用于传统工业去创造新工艺、开发新产品,计算机技术支持下的建模和模拟都是经常使用的有效手段.数学建模、数值计算和计算机图形学等相结合形成的计算机软件,已经被固化于产品中,在许多高新技术领域起着核心作用,被认为是高新技术的特征之一.在这个意义上,数学不再仅仅作为一门科学,是许多技术的基础,而且直接走向了技术的前台.国际上一位学者就提出了“高技术本质上是一种数学技术”的观点.

3) 数学迅速进入一些新领域,为数学建模开拓了许多新的处女地.

随着数学向诸如经济、人口、生态、地质等所谓非物理领域的渗透,一些交叉学科如计量经济学、人口控制论、数学生态学、数学地质学等应运而生.这里一般地说不存在作为支配关系的物理定律,当用数学方法研究这些领域中的定量关系时,数学建模就成为首要的、关键的步骤和这些学科发展与应用的基础.在这些领域里建立不同类型、不同方法、不同深浅程度的模型的余地相当大,为数学建模提供了广阔的新天地.马克思说过:“一门科学只有成功地运用数学时,才算达到了完善的地步”.21 世纪,数学必将大踏步地进入所有学科,数学建模将迎来蓬勃发展的新时期.

今天,在国民经济和社会活动的以下诸多方面,数学建模都有着非常具体的应用.

分析与设计 例如描述药物浓度在人体内的变化规律以分析药物的疗效;建立跨音速流和激波的数学模型,用数值模拟设计新的飞机翼型.

预报与决策 生产过程中产品质量指标的预报、气象预报、人口预报、经济增长预报等等,都要有预报模型;使经济效益最大的价格策略、使费用最少的设备维修方案,都是决策模型的例子.

控制与优化 电力、化工生产过程的最优控制、零件设计中的参数优化,要以数学模型为前提.建立大系统控制与优化的数学模型,是迫切需要和十分棘手的课题.

规划与管理 生产计划、资源配置、运输网络规划、水库优化调度,以及排队策略、物资管理等,都可以用数学规划模型解决.

数学建模与计算机技术的关系密不可分.一方面,像新型飞机设计、石油勘探数据处理中数学模型的求解当然离不开巨型计算机,而微型电脑的普及更使数学建模逐步进入人们的日常活动.比如,当一位公司经理根据客户提出的产品

数量、质量、交货期等要求,用笔记本电脑与客户进行价格谈判时,您不会怀疑他的电脑中贮存了由公司的各种资源、产品工艺流程及客户需求等数据研制的数学模型——快速报价系统和生产计划系统. 另一方面,以数字化为特征的信息正以爆炸之势涌入计算机,去伪存真、归纳整理、分析现象、显示结果……,计算机需要人们给它以思维的能力,这些当然要求助于数学模型. 所以把计算机技术与数学建模在知识经济中的作用比喻为如虎添翼,是恰如其分的.

美国国家科学院一位院士总结了将数学科学转化为生产力过程中的成功和失败,得出了“数学是一种关键的、普遍的、可以应用的技术”的结论,认为数学“由研究到工业领域的技术转化,对加强经济竞争力具有重要意义”,而“计算和建模重新成为中心课题,它们是数学科学技术转化的主要途径”.

1.3 建模示例之一 椅子能在地面上放稳吗

本节和下面两节将给出三个数学建模的例子,重点说明如何作出合理的、简化的假设,用数学语言确切地表述实际问题,以及怎样用模型的结果解释实际现象.

本节讨论的问题来源于日常生活中一个普通的事实:把椅子往不平的地面上一放,通常只有三只脚着地,放不稳,然而只需稍挪动几次,就可以使四只脚同时着地,放稳了. 这个看来似乎与数学无关的现象能用数学语言给以表述,并用数学工具来证实吗? 让我们试试看^[50].

模型假设 对椅子和地面应该作一些必要的假设:

1. 椅子四条腿一样长,椅脚与地面接触处可视为一个点,四脚的连线呈正方形.
2. 地面高度是连续变化的,沿任何方向都不会出现间断(没有像台阶那样的情况),即地面可视为数学上的连续曲面.
3. 对于椅脚的间距和椅腿的长度而言,地面是相对平坦的,使椅子在任何位置至少有三只脚同时着地.

假设 1 显然是合理的. 假设 2 相当于给出了椅子能放稳的条件,因为如果地面高度不连续,譬如在有台阶的地方是无法使四只脚同时着地的. 至于假设 3 是要排除这样的情况:地面上与椅脚间距和椅腿长度的尺寸大小相当的范围内,出现深沟或凸峰(即使是连续变化的),致使三只脚无法同时着地.

模型构成 中心问题是用数学语言把椅子四只脚同时着地的条件和结论表示出来.

首先要用变量表示椅子的位置. 注意到椅脚连线呈正方形,以中心为对称点,正方形绕中心的旋转正好代表了椅子位置的改变,于是可以用旋转角度这一

变量表示椅子的位置. 在图 1 中椅脚连线为正方形 $ABCD$, 对角线 AC 与 x 轴重合, 椅子绕中心点 O 旋转角度 θ 后, 正方形 $ABCD$ 转至 $A'B'C'D'$ 的位置, 所以对角线 AC 与 x 轴的夹角 θ 表示了椅子的位置.

其次要把椅脚着地用数学符号表示出来. 如果用某个变量表示椅脚与地面的竖直距离, 那么当这个距离为零时就是椅脚着地了. 椅子在不同位置时椅脚与地面的距离不同, 所以这个距离是椅子位置变量 θ 的函数.

虽然椅子有四只脚, 因而有四个距离, 但是由于正方形的中心对称性, 只要设两个距

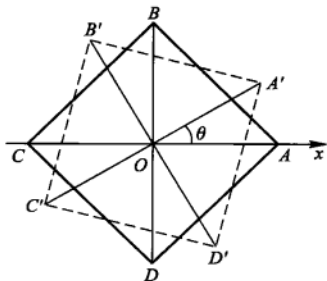


图 1 变量 θ 表示椅子的位置

离函数就行了. 记 A, C 两脚与地面距离之和为 $f(\theta)$, B, D 两脚与地面距离之和为 $g(\theta)$ ($f(\theta), g(\theta) \geq 0$). 由假设 2, f 和 g 都是连续函数. 由假设 3, 椅子在任何位置至少有三只脚着地, 所以对于任意的 θ , $f(\theta)$ 和 $g(\theta)$ 中至少有一个为零. 当 $\theta = 0$ 时不妨设 $g(\theta) = 0, f(\theta) > 0$. 这样, 改变椅子的位置使四只脚同时着地, 就归结为证明如下的数学命题:

已知 $f(\theta)$ 和 $g(\theta)$ 是 θ 的连续函数, 对任意 $\theta, f(\theta) \cdot g(\theta) = 0$, 且 $g(0) = 0, f(0) > 0$. 证明存在 θ_0 , 使 $f(\theta_0) = g(\theta_0) = 0$.

模型求解 上述命题有多种证明方法, 这里介绍其中比较简单, 但是有些粗糙的一种.

将椅子旋转 $90^\circ (\pi/2)$, 对角线 AC 与 BD 互换. 由 $g(0) = 0$ 和 $f(0) > 0$ 可知 $g(\pi/2) > 0$ 和 $f(\pi/2) = 0$.

令 $h(\theta) = f(\theta) - g(\theta)$, 则 $h(0) > 0$ 和 $h(\pi/2) < 0$. 由 f 和 g 的连续性知 h 也是连续函数. 根据连续函数的基本性质, 必存在 $\theta_0 (0 < \theta_0 < \pi/2)$ 使 $h(\theta_0) = 0$, 即 $f(\theta_0) = g(\theta_0)$.

最后, 因为 $f(\theta_0) \cdot g(\theta_0) = 0$, 所以 $f(\theta_0) = g(\theta_0) = 0$.

由于这个实际问题非常直观和简单, 模型解释和验证就略去了.

评注 这个模型的巧妙之处在于用一元变量 θ 表示椅子的位置, 用 θ 的两个函数表示椅子四脚与地面的距离, 进而把模型假设和椅脚同时着地的结论用简单、精确的数学语言表达出来, 构成了这个实际问题的数学模型.

模型假设中“四脚连线呈正方形”不是本质的, 读者可以考虑四脚连线呈长方形的情况 (习题 4).

如果读者对上面如此简单的建模特别是证明过程仔细考虑一下, 可能会发现不严谨之处. 如椅子的旋转轴在哪里, 它在旋转过程中怎样变化?

更严谨些的一种考虑如下: 定义与椅子的四只脚所在的平面垂直、且经过其

中心(O 点)的直线为旋转轴.取 A, B, D 脚同时着地的位置为初始位置,经过 O 点且与 A, B, D 脚所在平面垂直的直线为初始旋转轴.此后将椅子沿该轴高高举起,并在与该轴垂直的平面内逆时针旋转后再慢慢放下.放下的过程中,保持新的旋转轴始终通过 O 点但允许发生倾斜,使得 A, B 脚先着地,再让 D, C 两脚中至少一个着地(尽可能先让 D 脚着地).

1.4 建模示例之二 商人们怎样安全过河

三名商人各带一个随从乘船渡河,一只小船只能容纳二人,由他们自己划行.随从们密约,在河的任一岸,一旦随从的人数比商人多,就杀人越货.但是如何乘船渡河的大权掌握在商人们手中.商人们怎样才能安全渡河呢?

对于这类智力游戏,经过一番逻辑思索是可以找出解决办法的.这里用数学模型求解,一是为了给出建模的示例,二是因为这类模型可以解决相当广泛的一类问题,比逻辑思索的结果容易推广.

由于这个虚拟的问题已经理想化了,所以不必再作假设.安全渡河问题可以视为一个多步决策过程.每一步,即船由此岸驶向彼岸或从彼岸驶回此岸,都要对船上的人员(商人、随从各几人)作出决策,在保证安全的前提下(两岸的随从数都不比商人数多),在有限步内使全部人员过河.用状态(变量)表示某一岸的人员状况,决策(变量)表示船上的人员状况,可以找出状态随决策变化的规律.问题转化为在状态的允许变化范围内(即安全渡河条件),确定每一步的决策,达到渡河的目标.

模型构成 记第 k 次渡河前此岸的商人数为 x_k ,随从数为 $y_k, k=1, 2, \dots, x_k, y_k=0, 1, 2, 3$.将二维向量 $s_k=(x_k, y_k)$ 定义为状态.安全渡河条件下的状态集合称为允许状态集合,记作 S .

$$S = \{(x, y) | x=0, y=0, 1, 2, 3; x=3, y=0, 1, 2, 3; x=y=1, 2\} \quad (1)$$

不难验证, S 对此岸和彼岸都是安全的.

记第 k 次渡船上的商人数为 u_k ,随从数为 v_k .将二维向量 $d_k=(u_k, v_k)$ 定义为决策.允许决策集合记作 D ,由小船的容量可知

$$D = \{(u, v) | 1 \leq u + v \leq 2, u, v = 0, 1, 2\} \quad (2)$$

因为 k 为奇数时船从此岸驶向彼岸, k 为偶数时船由彼岸驶回此岸,所以状态 s_k 随决策 d_k 变化的规律是

$$s_{k+1} = s_k + (-1)^k d_k \quad (3)$$

(3)式称状态转移律.这样,制订安全渡河方案归结为如下的多步决策模型:

求决策 $d_k \in D (k=1, 2, \dots, n)$,使状态 $s_k \in S$ 按照转移律(3),由初始状态

$s_1 = (3, 3)$ 经有限步 n 到达状态 $s_{n+1} = (0, 0)$.

模型求解 根据(1)~(3)式编一段程序,用计算机求解上述多步决策问题是可行的. 不过对于商人和随从人数不大的简单状况,用图解法解这个模型更为方便.

在 xOy 平面坐标系上画出图 1 那样的方格,方格点表示状态 $s = (x, y)$. 允许状态集合 S 是用圆点标出的 10 个格子点. 允许决策 d_k 是沿方格线移动 1 或 2 格, k 为奇数时向左、下方移动, k 为偶数时向右、上方移动. 要确定一系列的 d_k , 使由 $s_1 = (3, 3)$ 经过那些圆点最终移至原点 $(0, 0)$.

图 1 给出了一种移动方案,经过决策 d_1, d_2, \dots, d_{11} , 最终有 $s_{12} = (0, 0)$. 这个结果很容易翻译成渡河的方案.

评注 这里介绍的是一种规格化的方法,所建立的多步决策模型可以用计算机求解,从而具有推广的意义. 譬如当商人和随从人数增加或小船的容量加大时,靠逻辑思考就困难了,而用这种模型则仍可方便地求解. 读者不妨考虑四名商人各带一个随从的情况(小船同前).

适当地设置状态和决策,确定状态转移律,建立多步决策模型,是有效地解决很广泛的一类问题的方法,在以后的章节中还要用到.

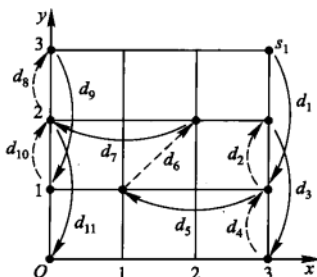


图 1 安全渡河问题的图解法

1.5 建模示例之三 如何施救药物中毒

一天夜晚,你作为见习医生正在医院内科急诊室值班,两位家长带着一个孩子急匆匆进来,诉说两小时前孩子一口气误吞下 11 片治疗哮喘病的、剂量为每片 100 mg 的氨茶碱片,已经出现呕吐、头晕等不良症状. 按照药品使用说明书,氨茶碱的成人用量一次是 100 ~ 200 mg, 儿童是 3 ~ 5 mg/kg, 如果过量服用,可使血药浓度(单位血液容积中的药量)过高,当血药浓度达到 100 $\mu\text{g/ml}$ 时,会出现严重中毒,达到 200 $\mu\text{g/ml}$ 则可致命^[67].

作为一位医生你清楚地知道,由于孩子服药是在两小时前,现在药物已经从胃进入肠道,无法再用刺激呕吐的办法排除. 当前需要作出判断的是,孩子的血药浓度会不会达到 100 $\mu\text{g/ml}$ 甚至 200 $\mu\text{g/ml}$, 如果会达到,则临床上应采取紧急方案来救治孩子.

问题的调查与分析 人体服用一定量的药物后,血药浓度与人体的血液总量有关. 一般来说,血液总量约为体重的 7% ~ 8%, 即体重 50 ~ 60 kg 的成年人

有 4 000 ml 左右的血液. 目测这个孩子的体重约为成年人的一半, 可认为其血液总量约为 2 000 ml. 由此, 血液系统中的血药浓度与药量之间可以相互转换.

药物口服后迅速进入胃肠道, 再由胃肠道的外壁进入血液循环系统, 被血液吸收. 胃肠道中药物的转移率, 即血液系统的吸收率, 一般与胃肠道中的药量成正比. 药物在被血液吸收的同时, 又通过代谢作用由肾脏排出体外, 排除率一般与血液中的药量成正比. 如果认为整个血液系统内药物的分布, 即血药浓度是均匀的, 可以将血液系统看作一个房室, 建立所谓一室模型.

血液系统对药物的吸收率和排除率可以由半衰期确定, 从药品说明书可知, 氨茶碱吸收的半衰期约 5 h, 排除的半衰期约 6 h.

如果血药浓度达到危险的水平, 临床上施救的一种办法是采用口服活性炭来吸附药物, 可使药物的排除率增加到原来(人体自身)的 2 倍, 另一种办法是进行体外血液透析, 药物排除率可增加到原来的 6 倍, 但是安全性不能得到充分保证, 建议尽量少用.

模型的假设和建立 为了判断孩子的血药浓度会不会达到危险的水平, 需要寻求胃肠道和血液系统中的药量随时间变化的规律. 记胃肠道中的药量为 $x(t)$, 血液系统中的药量为 $y(t)$, 时间 t 以孩子误服药的时刻为起点 ($t=0$). 根据前面的调查与分析可作以下假设:

1. 胃肠道中药物向血液系统的转移率与药量 $x(t)$ 成正比, 比例系数记作 $\lambda (>0)$, 总剂量 1 100 mg 的药物在 $t=0$ 瞬间进入胃肠道.

2. 血液系统中药物的排除率与药量 $y(t)$ 成正比, 比例系数记作 $\mu (>0)$, $t=0$ 时血液无药物.

3. 氨茶碱被吸收的半衰期为 5 h, 排除的半衰期为 6 h.

4. 孩子的血液总量为 2 000 ml.

根据假设对胃肠道中药量 $x(t)$ 和血液系统中药量 $y(t)$ 建立如下模型.

由假设 1, $x(0) = 1\ 100$ mg, 随着药物从胃肠道向血液系统的转移, $x(t)$ 下降的速度与 $x(t)$ 本身成正比(比例系数 $\lambda > 0$), 所以 $x(t)$ 满足微分方程

$$\frac{dx}{dt} = -\lambda x, x(0) = 1\ 100 \quad (1)$$

由假设 2, $y(0) = 0$, 药物从胃肠道向血液系统的转移相当于血液系统对药物的吸收, $y(t)$ 由于吸收作用而增长的速度是 λx , 由于排除而减少的速度与 $y(t)$ 本身成正比(比例系数 $\mu > 0$), 所以 $y(t)$ 满足微分方程

$$\frac{dy}{dt} = \lambda x - \mu y, y(0) = 0 \quad (2)$$

方程(1), (2)中的参数 λ 和 μ 可由假设 3 中的半衰期确定.

模型求解 微分方程(1)是可分离变量方程, 容易得到

$$x(t) = 1\,100e^{-\lambda t} \quad (3)$$

表明胃肠道中的药量 $x(t)$ 随时间单调减少并趋于 0. 为了确定 λ , 利用药物吸收的半衰期为 5 h, 即 $x(5) = 1\,100e^{-5\lambda} = x(0)/2 = 1\,100/2$, 得 $\lambda = (\ln 2)/5 = 0.138\,6(1/h)$.

将(3)代入方程(2), 得到一阶线性微分方程, 求解得

$$y(t) = \frac{1\,100\lambda}{\lambda - \mu}(e^{-\mu t} - e^{-\lambda t}) \quad (4)$$

表明血液系统中的药量 $y(t)$ 随时间先增后减并趋于 0.

为了根据药物排除的半衰期为 6 h 来确定 μ , 考虑血液系统只对药物进行排除的情况, 这时 $y(t)$ 满足方程 $\frac{dy}{dt} = -\mu y$, 若设在某时刻 τ 有 $y(\tau) = a$, 则 $y(t) = ae^{-\mu(t-\tau)}$, $t \geq \tau$. 利用 $y(\tau+6) = a/2$, 可得 $\mu = (\ln 2)/6 = 0.115\,5(1/h)$.

将 $\lambda = 0.138\,6$ 和 $\mu = 0.115\,5$ 代入(3), (4), 得 (t 的单位: h; x, y 的单位: mg)

$$x(t) = 1\,100e^{-0.138\,6t} \quad (5)$$

$$y(t) = 6\,600(e^{-0.115\,5t} - e^{-0.138\,6t}) \quad (6)$$

结果分析 用 MATLAB 软件对(5), (6)作图, 得图 1.

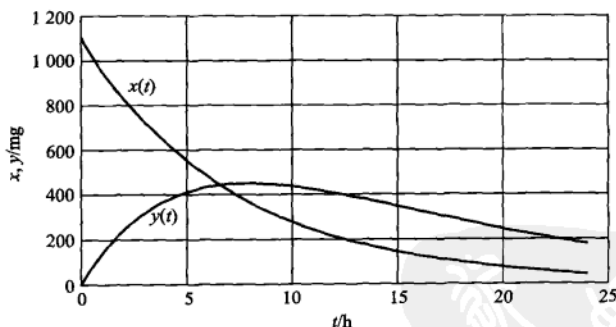


图 1 胃肠道中药量 $x(t)$ 和血液系统中药量 $y(t)$

根据假设 4, 孩子的血液总量为 2 000 ml, 出现严重中毒的血药浓度 $100\,\mu\text{g/ml}$ 和致命的血药浓度 $200\,\mu\text{g/ml}$ 分别相当于血液中药量 y 达到 200 mg 和 400 mg. 由图 1 看出, 药量 y 在约 2 h 达到 200 mg, 即孩子到达医院时已经出现严重中毒; 如不及时施救, 药量 y 将在约 5 h (到医院后 3 h) 达到 400 mg.

由(6)容易精确地算出孩子到达医院时血液中药量 $y(2) = 236.5\,\text{mg}$, 而计算药量达到 400 mg 的时间 (记作 t_1), 则需要解非线性方程 $6\,600(e^{-0.115\,5t_1} - e^{-0.138\,6t_1}) = 400$, 用 MATLAB 软件计算可以得到 $t_1 = 4.87\,\text{h}$.

由图 1 还可以看出,血液中药量 $y(t)$ 达到最大值的时间约在 $t=8$ h,即到达医院后 6 h,其精确值可由方程(2)或解(4)计算,记作 $t_2, t_2 = \frac{\ln(1 + \lambda/6\mu)}{\lambda - \mu} = 7.89$ h,且 $y(t_2) = 442.1$ mg.

利用这个模型还可以确定对于孩子及成人服用氨茶碱能引起严重中毒和致命的最小剂量(习题 6).

施救方案 根据模型计算的结果,如不及时施救,孩子会有生命危险.根据调查,如采用口服活性炭来吸附药物的办法施救,药物的排除率可增加到 μ 的 2 倍,即 0.231 0.让我们计算一下,采用这种施救方案血液中药量 $y(t)$ 的变化情况.

设孩子到达医院时刻($t=2$)就开始施救,前面已经算出 $y(2) = 236.5$,由(2),(3),新的模型为(血液中药量记作 $z(t)$)

$$\frac{dz}{dt} = \lambda x - \mu z, t \geq 2, x = 1100e^{-\lambda t}, z(2) = 236.5 \quad (7)$$

仍是一阶线性微分方程,只不过初始时刻为 $t=2$,当 $\lambda = 0.1386$ (不变)而 $\mu = 0.2310$ 时,(7)的解为

$$z(t) = 1650e^{-0.1386t} - 1609.5e^{-0.2310t}, t \geq 2 \quad (8)$$

用 MATLAB 软件对(8)作图,如图 2.

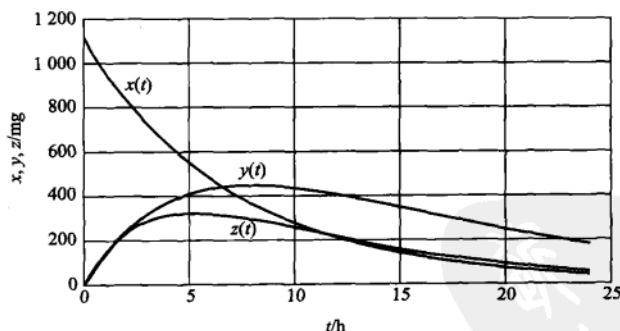


图 2 施救后血液系统中药量 $z(t)$ ($x(t), y(t)$ 同图 1)

由图 2 可看出,施救后血液中药量 $z(t)$ 达到最大值的时间约在 $t=5$ h,即到达医院施救后 3 h,其精确值可由(8)算出,记作 $t_3, t_3 = 5.26$ h,且 $z(t_3) = 318.4$ mg,远低于 $y(t)$ 的最大值和致命水平.

图 2 还表明,虽然采用了口服活性炭来吸附药物的办法施救,血液中药量 $z(t)$ 仍有一段时间在上升,说明用这种方法药物的排除率增加还不够大.不妨计算一下,如果要使 $z(t)$ 在施救后($t \geq 2$)立即下降,排除率 μ 至少应该多大.

$z(t)$ 在 $t=2$ 取得极大值, 相当于 $t=2$ 时 (7) 式满足

$$\left. \frac{dz}{dt} \right|_{t=2} = (\lambda x - \mu z) \Big|_{t=2} = 0 \quad (9)$$

由 (5) 算出 $x(2) = 833.7$, 再利用前面已有的 $z(2) = 236.5$ 和 $\lambda = 0.1386$, 立即得到 $\mu = 0.4885$, 约为原来 (人体自身 $\mu = 0.1155$) 的 4.2 倍。

如果采用体外血液透析的办法, 药物排除率可增加到 $\mu = 0.1155 \times 6 = 0.693$, 血液中药量下降更快, 读者可以用这个 μ 重新求解 (7) 并作图 (习题 7)。至于临床上究竟是否需要采取这种办法, 当由医生综合考虑并征求病人和家属意见后确定。

1.6 数学建模的基本方法和步骤

数学建模面临的实际问题是多种多样的, 建模的目的不同、分析的方法不同、采用的数学工具不同, 所得模型的类型也不同, 我们不能指望归纳出若干条准则, 适用于一切实际问题的数学建模方法。下面所谓基本方法不是针对具体问题而是从方法论的意义上讲的。

数学建模的基本方法

一般说来, 建模方法大体上可分为机理分析和测试分析两种。机理分析是根据对客观事物特性的认识, 找出反映内部机理的数量规律, 建立的模型常有明确的物理或现实意义。前面几个示例都是用的机理分析。测试分析是将研究对象看作一个“黑箱”系统 (意思是它的内部机理看不清楚), 通过对系统输入、输出数据的测量和统计分析, 按照一定的准则找出与数据拟合得最好的模型。

面对一个实际问题用哪一种方法建模, 主要取决于人们对研究对象的了解程度和建模目的。如果掌握了一些内部机理的知识, 模型也要求具有反映内在特征的物理意义, 建模就应以机理分析为主。而如果对象的内部规律基本上不清楚, 模型也不需要反映内部特性 (例如仅用于对输出作预报), 那么就可以用测试分析。

对于许多实际问题还常常将两种方法结合起来建模, 即用机理分析建立模型的结构, 用测试分析确定模型的参数。

机理分析当然要针对具体问题来做, 不可能有统一的方法, 因而主要是通过实例研究 (Case Studies) 来学习。测试分析有一套完整的数学方法, 第 10 章统计回归模型是其中的一小部分。以动态系统为主的测试分析称为系统辨识 (System Identification), 是一门专门学科。本书以后所说的数学建模主要指机理分析。

数学建模的一般步骤

建模要经过哪些步骤并没有一定的模式,通常与问题性质、建模目的等有关.下面介绍的是机理分析方法建模的一般过程,如图1所示.

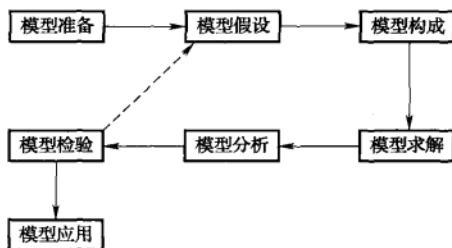


图1 数学建模步骤示意图

模型准备 了解问题的实际背景,明确建模目的,搜集必要的信息如现象、数据等,尽量弄清对象的主要特征,形成一个比较清晰的“问题”,由此初步确定用哪一类模型.情况明才能方法对.在模型准备阶段要深入调查研究,虚心向实际工作者请教,尽量掌握第一手资料.

模型假设 根据对象的特征和建模目的,抓住问题的本质,忽略次要因素,作出必要的、合理的简化假设.对于建模的成败这是非常重要和困难的一步.假设作得不合理或太简单,会导致错误的或无用的模型;假设作得过分详细,试图把复杂对象的众多因素都考虑进去,会使你很难或无法继续下一步的工作.常常需要在合理与简化之间作出恰当的折中.通常,作假设的依据,一是出于对问题内在规律的认识,二是来自对现象、数据的分析,以及二者的综合.想象力、洞察力、判断力以及经验,在模型假设中起着重要作用.

模型构成 根据所作的假设,用数学的语言、符号描述对象的内在规律,建立包含常量、变量等的数学模型,如优化模型、微分方程模型、差分方程模型、图模型等.这里除了需要一些相关学科的专门知识外,还常常需要较为广阔地应用数学方面的知识.要善于发挥想象力,注意使用类比法,分析对象与熟悉的其他对象的共性,借用已有的模型.建模时还应遵循的一个原则是:尽量采用简单的数学工具,因为你的模型总是希望更多的人了解和使用,而不是只供少数专家欣赏.

模型求解 可以采用解方程、画图形、优化方法、数值计算、统计分析等各种数学方法,特别是数学软件和计算机技术.

模型分析 对求解结果进行数学上的分析,如结果的误差分析、统计分析、模型对数据的灵敏性分析、对假设的强健性分析等.

模型检验 把求解和分析结果翻译回到实际问题,与实际的现象、数据比较,检验模型的合理性和适用性.如果结果与实际不符,问题常常出在模型假设

上,应该修改、补充假设,重新建模,如图 1 中的虚线所示.这一步对于模型是否真的有用非常关键,要以严肃认真的态度对待.有些模型要经过几次反复,不断完善,直到检验结果获得某种程度上的满意.

模型应用 应用的方式与问题性质、建模目的及最终的结果有关,一般不属于本书讨论的范围.

应当指出,并不是所有问题的建模都要经过这些步骤,有时各步骤之间的界限也不那么分明,建模时不要拘泥于形式上的按部就班,本书的实例就采用了灵活的表述形式.

数学建模的全过程

从前面几个建模示例以及一般步骤的分析,可以将数学建模的过程分为表述、求解、解释、验证几个阶段,并且通过这些阶段完成从现实对象到数学模型,再从数学模型回到现实对象的循环,如图 2 所示.

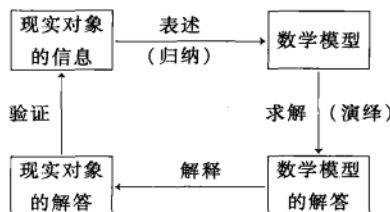


图 2 数学建模的全过程

表述是将现实问题“翻译”成抽象的数学问题,属于归纳法.数学模型的求解则属于演绎法.归纳是依据个别现象推出一一般规律,演绎是按照普遍原理考察特定对象,导出结论.因为任何事物的本质都要通过现象来反映,必然要透过偶然来表露,所以正确的归纳不是主观、盲目的,而是有客观基础的,但也往往是不精细的、带感性的,不易直接检验其正确性.演绎则利用严格的逻辑推理,对解释现象、作出科学预见具有重要意义,但是它要以归纳的结论作为公理化形式的前提,只能在这个前提下保证其正确性.因此,归纳和演绎是辩证统一的过程:归纳是演绎的基础,演绎是归纳的指导^[86].

解释是把数学模型的解答“翻译”回到现实对象,给出分析、预报、决策或者控制的结果.最后,作为这个过程的重要一环,这些结果需要用实际的信息加以验证.

图 2 也揭示了现实对象和数学模型的关系.一方面,数学模型是将现象加以归纳、抽象的产物,它源于现实,又高于现实;另一方面,只有当数学建模的结果经受住现实对象的检验时,才可以用来指导实际,完成实践—理论—实践这一循环.

1.7 数学模型的特点和分类

数学建模是利用数学工具解决实际问题的重要手段,得到的模型有许多优点,也有一些弱点.下面归纳出数学模型的若干特点,以期读者在学习过程中逐步领会^[42].

数学模型的特点

模型的逼真性和可行性 一般说来总是希望模型尽可能逼近研究对象,但是是一个非常逼真的模型在数学上常常是难以处理的,因而不容易达到通过建模对现实对象进行分析、预报、决策或者控制的目的,即实用上不可行.另一方面,越逼真的模型常常越复杂,即使数学上能处理,这样的模型应用时所需要的“费用”也相当高,而高“费用”不一定与复杂模型取得的“效益”相匹配.所以建模时往往需要在模型的逼真性与可行性,“费用”与“效益”之间作出折中和抉择.

模型的渐进性 稍微复杂一些的实际问题的建模通常不可能一次成功,要经过上一节描述的建模过程的反复迭代,包括由简到繁,也包括删繁就简,以获得越来越满意的模型.在科学发展过程中,随着人们认识和实践能力的提高,各门学科中的数学模型也存在着一个不断完善或者推陈出新的过程.从19世纪力学、热学、电学等许多学科由牛顿力学的模型主宰,到20世纪爱因斯坦相对论模型的建立,是模型渐进性的明显例证.

模型的强健性 模型的结构和参数常常是由模型假设及对象的信息(如观测数据)确定的,而假设不可能太准确,观测数据也是允许有误差的.一个好的模型应该具有下述意义的强健性:当模型假设改变时,可以导出模型结构的相应变化;当观测数据有微小改变时,模型参数也只有相应的微小变化.

模型的可转移性 模型是现实对象抽象化、理想化的产物,它不为对象的所属领域所独有,可以转移到另外的领域.在生态、经济、社会等领域内建模就常常借用物理领域中的模型.模型的这种性质显示了它的应用的极端广泛性.

模型的非预知性 虽然已经发展了许多应用广泛的模型,但是实际问题是各种各样、变化万千的,不可能要求把各种模型做成预制品供你在建模时使用.模型的这种非预知性使得建模本身常常是事先没有答案的问题(Open-end Problem).在建立新的模型的过程中甚至会伴随着新的数学方法或数学概念的产生.

模型的条理性 从建模的角度考虑问题可以促使人们对现实对象的分析更全面、更深入、更具条理性,这样即使建立的模型由于种种原因尚未达到实用的程度,对问题的研究也是有利的.

模型的技艺性 建模的方法与其他一些数学方法如方程解法、规划问题解

法等是根本不同的,无法归纳出若干条普遍适用的建模准则和技巧。有人说,建模目前与其说是一门技术,不如说是一种艺术,是技艺性很强的技巧。经验、想象力、洞察力、判断力以及直觉、灵感等在建模过程中起的作用往往比一些具体的数学知识更大。

模型的局限性 这里有几方面的含义。第一,由数学模型得到的结论虽然具有通用性和精确性,但是因为模型是现实对象简化、理想化的产物,所以一旦将模型的结论应用于实际问题,就回到了现实世界,那些被忽视、简化的因素必须考虑,于是结论的通用性和精确性只是相对的和近似的。第二,由于人们认识能力和科学技术包括数学本身发展水平的限制,还有不少实际问题很难得到有着实用价值的数学模型。如一些内部机理复杂、影响因素众多、测量手段不够完善、技艺性较强的生产过程,像生铁冶炼过程,常常需要开发专家系统,与建立数学模型相结合才能获得较满意的应用效果。**专家系统**是一种计算机软件系统,它总结专家的知识和经验,模拟人类的逻辑思维过程,建立若干规则和推理途径,主要是定性地分析各种实际现象并作出判断。专家系统可以看成计算机模拟的新发展。第三,还有些领域中的问题今天尚未发展到用建模方法寻求数量规律的阶段,如中医诊断过程,目前所谓计算机辅助诊断也是属于总结著名中医的丰富临床经验的专家系统。

数学模型的分类

数学模型可以按照不同的方式分类,下面介绍常用的几种。

1. 按照模型的应用领域(或所属学科)分。如人口模型、交通模型、环境模型、生态模型、城镇规划模型、水资源模型、再生资源利用模型、污染模型等。范畴更大一些则形成许多边缘学科,如生物数学、医学数学、地质数学、数量经济学、数学社会学等。

2. 按照建立模型的数学方法(或所属数学分支)分。如初等模型、几何模型、微分方程模型、统计回归模型、数学规划模型等。

3. 按照模型的表现特性又有几种分法:

确定性模型和随机性模型 取决于是否考虑随机因素的影响。近年来随着数学的发展,又有所谓突变性模型和模糊性模型。

静态模型和动态模型 取决于是否考虑时间因素引起的变化。

线性模型和非线性模型 取决于模型的基本关系,如微分方程是否是线性的。

离散模型和连续模型 指模型中的变量(主要是时间变量)取为离散还是连续的。

虽然从本质上讲大多数实际问题是随机性的、动态的、非线性的,但是由于确定性、静态、线性模型容易处理,并且往往可以作为初步的近似来解决问题,所

以建模时常先考虑确定性、静态、线性模型。连续模型便于利用微积分方法求解析解,作理论分析,而离散模型便于在计算机上作数值计算,所以用哪种模型要看具体问题而定。在具体的建模过程中将连续模型离散化,或将离散变量视作连续的,也是常采用的方法。

4. 按照建模目的分。有描述模型、预报模型、优化模型、决策模型、控制模型等。

5. 按照对模型结构的了解程度分。有所谓白箱模型、灰箱模型、黑箱模型。这是把研究对象比喻成一只箱子里的机关,要通过建模来揭示它的奥妙。白箱主要包括用力学、热学、电学等一些机理相当清楚的学科描述的现象以及相应的工程技术问题,这方面的模型大多已经基本确定,还需深入研究的主要是优化设计和控制等问题了。灰箱主要指生态、气象、经济、交通等领域中机理尚不十分清楚的现象,在建立和改善模型方面都还不同程度地有许多工作要做。至于黑箱则主要指生命科学和社会科学等领域中一些机理(数量关系方面)很不清楚的现象。有些工程技术问题虽然主要基于物理、化学原理,但由于因素众多、关系复杂和观测困难等原因,也常作为灰箱或黑箱模型处理。当然,白、灰、黑之间并没有明显的界限,而且随着科学技术的发展,箱子的“颜色”必然是逐渐由暗变亮的。

1.8 数学建模能力的培养

在详细分析了建立数学模型的全过程和数学模型的特点以后,我们看到用建模方法解决实际问题,首先是用数学语言表述问题即构造模型,其次才是用数学工具求解构成的模型。绝大多数数学课程如微积分、线性代数、概率论、计算方法等都是讲授某一专门知识和培养数学运算、逻辑推理能力的,这些数学技巧主要用来求解数学模型。用数学语言表述问题,包括模型假设、模型构造等,除了要有广博的知识(包括数学知识和各种实际知识)和足够的经验之外,特别需要丰富的想象力和敏锐的洞察力。

想象力指人们在原有知识的基础上,将新感知的形象与记忆中的形象相互比较、重新组合、加工处理,创造出新的形象,是一种形象思维活动。**洞察力**指人们在充分占有资料的基础上,经过初步分析能迅速抓住主要矛盾,舍弃次要因素,简化问题的层次,对可以用哪些方法解决面临的问题,以及不同方法的优劣作出判断^[86]。

类比方法和理想化方法是建模中常用的方法,它们的运用与想象力、洞察力有密切关系。类比法注意到研究对象与已熟悉的另一对象具有某些共性,比较二者相似之处以获得对研究对象的新认识。选择什么对象进行类比,比较哪些相似的属性,在一定程度上是靠想象进行的。将交通流与水流类比起来建立交通流模型

是这方面的例子。理想化方法是从观察和经验中通过想象和逻辑思维,把对象简化、纯化,使其升华到理想状态,以期更本质地揭示对象的固有规律。在一定条件下把物体看作质点,把实际位置看作数学上的点、线等都是理想化的结果。

建模过程是一种创造性思维过程,除了想象、洞察、判断这些属于形象思维、逻辑思维范畴的能力之外,直觉和灵感往往也起着不可忽视的作用。直觉是人们对新事物本质的极敏锐的领悟、理解或推断,灵感指在人们有意识或下意识思考过程中迸发出来的猜测、思路或判断。二者都具有突发性,且思维者本人往往说不清它的来路和道理。当由于各种限制利用已有知识难以对研究对象作出有效的推理和判断时,凭借相似、类比、猜测、外推等思维方式,以及不完整、不连续、不严密的、带启发性的直觉和灵感,去“战略性”地认识对象,是人类创造性思维的特点之一,也是人脑比按程序逻辑工作的计算机、机器人的高明之处。历史上不乏在科学家的直觉和灵感的火花中诞生的假说、论证和定律。当然,直觉和灵感不是凭空产生的,它要求人们具有丰富的背景知识,对问题进行反复思考和艰苦探索,对各种思维方法运用娴熟。相互讨论和思想交锋,特别是不同专业的成员之间的探讨,是激发直觉和灵感的重要因素。所以由各种专门人才组成的所谓团队工作方式(Team Work)越来越受到重视。

前面说过,建模可以看成一门艺术。艺术在某种意义上是无法归纳出几条准则或方法的。一名出色的艺术家需要大量的观摩和前辈的指教,更需要亲身的实践。类似地,掌握建模这门艺术,培养想象力和洞察力,不外乎认真做好这样两条:第一,学习、分析、评价、改造别人做过的模型。首先是弄懂它,分析为什么这么做,然后找出它的优缺点,并尝试改进的方法。第二,要亲自动手,踏实地做几个实际题目。为了这个目的,本书采用实例研究方法,一方面给出在各个应用领域不同数学方法建模的大量实例,另一方面通过习题提供若干实际题目让读者自己练习。实例研究方法虽然不能按照严密的逻辑结构去讨论问题,不能划定这些方法的实用范围,其得到的结果也并非无可置疑,但它却是我们学习建模以解决实际问题的一种生动、有效的方法^[47]。

对于亲自动手做几个实际题目,这里特别强调一下参加大学生数学建模竞赛的意义和作用。

全国大学生数学建模竞赛是1992年开始举办的,规模从最初的几十所学校、几百个队,发展到2010年的1197所院校、17317个队,竞赛之所以如此受到广大同学的欢迎,主要是由于它的内容、形式和评判标准,适合培养有创新精神和综合素质人才的需要,并且具有明显区别于大家熟悉的数学等自然科学学科性竞赛的特点:赛题由工程技术、管理科学及社会热点问题简化而成;要求用数学建模方法和计算机技术,完成一篇包括模型的假设、建立和求解,结果的分析 and 检验,以及自我评价优缺点等方面的学术论文;赛题没有标准答案,评判以

假设的合理性、建模的创造性、结果的正确性(指与模型相符合)及表述的清晰性为标准;在3天时间内,3名大学生为1队共同完成,可以使用任何资料、软件、互联网等,惟一的限制是不能与队外的同学、老师讨论赛题.在全国竞赛的推动下,许多学校在组织培训过程中,让同学做有类似特点的练习题,举办模拟赛或选拔赛,显著地扩大了竞赛的参与面和受益面.

只要认真参加这些不同规模的数学建模竞赛,同学们的收获和提高是多方面的.

首先,运用数学建模方法分析和解决实际问题的能力会得到实际的锻炼,数学建模的意识(特指碰到实际问题能从数学建模的角度去思考)会有一定的提高.赛题通常要用到几门数学和计算机课程及有关的多种知识,对于一直是一门课、一门课学习的同学来说,训练运用综合知识的机会是难得的,这些对于后续课程的学习和独立研究能力的培养,都有很大的好处.

其次,合作精神与团队意识会得到培养和提高.竞赛需要3个人先是相互启发、争辩,然后相互妥协、达成一致的基础上分工合作、奋力攻关,这与同学们毕业后经常面临的集体工作方式十分相近.对于十几年来在读书、做题、考试等一系列个人奋斗的环境中成长起来的同学们来说,竞赛提供了一个既充分展示个人的智商,又有助于培养与人合作的情商的平台.

还有,竞赛需要快捷地搜集、整理、消化与题目有关的资料(主要依靠互联网),使之为我所用,对于尚处于学习阶段的同学来说,这是少有的机会;一篇清晰、通畅地阐明建模思路、假设、方法、结果等内容的论文,是参赛成果的集中体现,文字表述能力也是人们工作能力的一个重要部分,竞赛有益于这方面的锻炼;赛题的实用性有助于培养同学们关注社会生活、理论联系实际学风;3天的时间显然不可能做得完美,从理论、方法到实际应用,都可以在赛后阶段继续与老师和同学一起,给以充实和提高;既充分开放、又有规则约束的竞赛方式,可以培养慎独、自律的良好道德品质.

近20年来已经有数十万同学参加了全国大学生数学建模竞赛,而参加赛前培训、选拔赛、校内赛的当有数百万.许多同学表示,不管最后竞赛的成绩如何,只要认真参加了培训、自学、讨论、竞赛的全过程,都会有丰硕的收获,他们用“一次参赛、终身受益”来总结亲身体会.

习 题

1. 举出两三个实例说明建立数学模型的必要性,包括实际问题的背景,建模目的,需要大体上什么样的模型以及怎样应用这种模型等.

2. 从下面不太明确的叙述中确定要研究的问题,要考虑哪些有重要影响的变量^[32]:

(1) 一家商场要建一个新的停车场,如何规划照明设施.

(2) 一农民要在—块土地上作出农作物的种植规划.

(3) 一制造商要确定某种产品的产量及定价.

(4) 卫生部门要确定一种新药对某种疾病的疗效.

(5) 一滑雪场要进行山坡滑道和上山缆车的规划.

3. 怎样解决下面的实际问题, 包括需要哪些数据资料, 要做些什么观察、试验以及建立什么样的数学模型等^[32, 42].

(1) 估计一个人体内血液的总量.

(2) 为保险公司制定人寿保险金计划(不同年龄的人应缴纳的金额和公司赔偿的金额).

(3) 估计一批日光灯管的寿命.

(4) 确定火箭发射至最高点所需的时间.

(5) 决定十字路口黄灯亮的时间长度.

(6) 为汽车租赁公司制订车辆维修、更新和出租计划.

(7) 一高层办公楼有 4 部电梯, 早晨上班时间非常拥挤, 试制订合理的运行计划.

4. 在 1.3 节“椅子能在不平的地面上放稳吗”的假设条件中, 将四脚的连线呈正方形改为呈长方形, 其余不变. 试构造模型并求解.

5. 模仿 1.4 节商人过河问题中的状态转移模型, 做下面这个众所周知的智力游戏: 人带着猫、鸡、米过河, 除需要人划船之外, 船至多能载猫、鸡、米三者之一, 而当人不在场时猫要吃鸡, 鸡要吃米. 试设计一个安全过河方案, 并使渡河次数尽量少.

6. 利用 1.5 节药物中毒施救模型确定对于孩子(血液总量为 2 000 ml)及成人(血液总量为 4 000 ml)服用氨茶碱能引起严重中毒和致命的最小剂量.

7. 对于 1.5 节的模型, 如果采用的是体外血液透析的办法, 求解药物中毒施救模型的血液中药量的变化并作图.

8. 为了培养想象力、洞察力和判断力, 考察对象时除了从正面分析外, 还常常需要从侧面或反面思考. 试尽可能迅速地回答下面的问题:

(1) 某甲早 8:00 从山下旅店出发, 沿一条路径上山, 下午 5:00 到达山顶并留宿. 次日早 8:00 沿同一路径下山, 下午 5:00 回到旅店. 某乙说, 甲必在两天中的同一时刻经过路径中的同一地点. 为什么?

(2) 37 支球队进行冠军争夺赛, 每轮比赛中出场的每两支球队中的胜者及轮空者进入下一轮, 直至比赛结束. 问共需进行多少场比赛, 共需进行多少轮比赛. 如果是 n 支球队比赛呢?

(3) 甲乙两站之间有电车相通, 每隔 10 min 甲乙两站相互发一趟车, 但发车时刻不一定相同. 甲乙之间有一中间站丙, 某人每天在随机的时刻到达丙站, 并搭乘最先经过丙站的那趟车, 结果发现 100 天中约有 90 天到达甲站, 约有 10 天到达乙站. 问开往甲乙两站的电车经过丙站的时刻表是如何安排的.

(4) 某人家住 T 市在他乡工作, 每天下班后乘火车于 6:00 抵达 T 市车站, 他的妻子驾车准时到车站接他回家. 一日他提前下班搭早一班火车于 5:30 抵 T 市车站, 随即步行回家, 他的妻子像往常一样驾车前来, 在半路上遇到他, 即接他回家, 此时发现比往常提前了 10 min.

问他步行了多长时间.

(5) 一男孩和一女孩分别在离家 2 km 和 1 km 且方向相反的两所学校上学, 每天同时放学后分别以 4 km/h 和 2 km/h 的速度步行回家. 一小狗以 6 km/h 的速度由男孩处奔向女孩, 又从女孩处奔向男孩, 如此往返直至回到家中. 问小狗奔波了多少路程.

如果男孩和女孩上学时小狗也往返奔波在他们之间, 问当他们到达学校时小狗在何处^[50].



第2章

初等模型

如果研究对象的机理比较简单,一般用静态、线性、确定性模型描述就能达到建模目的时,我们基本上可以用初等数学的方法来构造和求解模型.通过本章介绍的若干实例读者能够看到,用很简单的数学方法已经可以解决一些饶有兴味的实际问题.

需要强调的是,衡量一个模型的优劣全在于它的应用效果,而不是采用了多么高深的数学方法.进一步说,如果对于某个实际问题我们用初等的方法和所谓高等的方法建立了两个模型,它们的应用效果相差无几,那么受到人们欢迎并采用的,一定是前者而非后者.

2.1 光盘的数据容量

20 世纪 80 年代初激光唱片(CD, Compact Disc 的缩写)问世,通常把它与激光视盘(LD, Laser Disc 的缩写)统称为光盘,用于储存数字声频、视频信号和计算机数据文件等.20 世纪 90 年代推出了 CD 的后继产品——数字视频光盘(DVD, Digital Video Disk 的缩写,后又称 Digital Versatile Disc,即数字多用途光盘).进入 21 世纪,不断改进的光盘制品集计算机、光学记录 and 影视技术为一体,极力满足人们对高性能、大容量数据的存储需求,不仅在音频、视频领域得到广泛应用,而且带动了出版、广播、通信、互联网等行业的发展.CD、DVD 及一系列音像制品已经进入千家万户的日常生活.

CD 的信息容量(Data Capacity)一般为单层 650 MB(MegaBytes,兆字节的缩写),DVD 的信息容量一般为单层 4.7 GB(GigaBytes,千兆字节的缩写).这些数值是怎么确定的,在一定条件下怎样使其最大化,让我们用尽量少的专业知识,从数学模型的角度研究这个问题^[1].

问题分析 光盘的外观尺寸是由一些大公司成立的联盟决定的,如 CD、DVD 等盘片的外径为 120 mm,并且沿外边缘留有 2 mm 宽的环形区域不存储信息,CLV 光盘(CLV 的解释见后)存储信息的内圈直径为 45 mm,如图 1.在内外圈之间的环形区域,经过编码的数字信息,以一定深度和宽度、不同长度的凹坑

的形式,用烧蚀技术存储在光盘表面呈螺旋线形状的信道上,相邻两条螺旋线间的距离称为信道间距(信道间距很小,图中是放大的)。当盘片上环形区域面积一定时,数据信息容量的大小取决于信道的总长度与信道上存储数据的线密度(即单位长度上的字节数),而决定信道长度和线密度大小的主要因素是所用激光的波长和驱动光盘的机械形式。

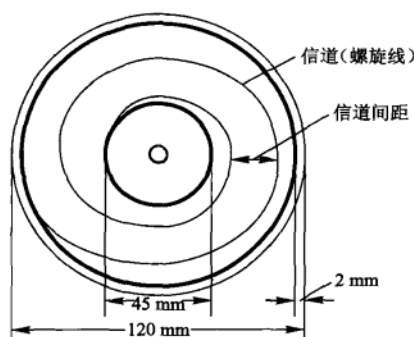


图1 光盘外观示意图

当光盘运转时激光束要能识别出信道上的凹坑所携带的信息,必须精确地聚焦。由于光的衍射作用,激光束不可能聚焦到一点,而是形成一个个称为光斑的圆。为了提高存储数据的线密度,应该使光斑尽量小,而光斑的大小与激光波长成正比。CD用的是红外激光器,DVD用的是红色激光,后又发展到蓝色激光,波长越来越短。对应于这些激光器的激光波长、光斑直径、信道间距(一般是光斑直径的0.8倍)和存储数据的线密度等基本数据见表1。

表1 3种光盘的基本数据

激光器	激光波长/ μm	光斑直径/ μm	信道间距/ μm	数据线密度/ ($\text{B} \cdot \text{mm}^{-1}$)
红外(CD)	0.78	2	1.6	121
红色(DVD)	0.64	0.925	0.74	387
蓝色(DVD)	0.41	0.4	0.32	800

驱动光盘的机械形式分两种:恒定角速度 CAV (Constant Angular Velocity) 和恒定线速度 CLV (Constant Linear Velocity)。CAV 光盘的角速度即转速固定,每一圈螺旋线上要存储同等数量的数据信息(字节),随着由内而外各圈螺旋线长度的增加,数据的线密度递减,信息容量取决于螺旋线最内圈的长度、线密度

以及总圈数. CLV 光盘的线速度固定, 各圈螺旋线上数据的线密度不变, 信息容量取决于固定的线密度和螺旋线总长度.

只从光盘的信息容量比较, CLV 显然优于 CAV, 但是在使用时还要考虑数据的读取时间. 由于 CLV 每圈的转速不同, 当读出磁头在光盘内外圈移动地读取数据时, 需要有等待光盘加速或减速的时间, 而 CAV 则不需要. 实际上, 对于音乐、影像、计算机文件等按顺序播放的信息, 多用 CLV; 而对于词典、数据库、人机交互等常要随机查找的信息, 多用 CAV.

模型建立 根据以上分析分别对 CLV 光盘和 CAV 光盘的信息容量建模.

记光盘上存储信息的环形区域内圆半径为 R_1 , 外圆半径为 R_2 , 信道间距为 d , 存储数据的线密度为 ρ . CLV 光盘的信息容量记作 C_{CLV} , 则

$$C_{CLV} = \rho L_{CLV} \quad (1)$$

其中 L_{CLV} 是信道总长度. 精确计算 L_{CLV} 需要用到曲线 (螺旋线) 积分, 然而从图 1 中光盘的内、外圆半径和表 1 中信道间距的数值可知, 由于信道间距 d 与环形区域宽度 $R_2 - R_1$ 相比很小, 螺旋线总圈数在 10^4 量级, 螺旋线的长度可用一系列同心圆周长之和来近似, 即

$$L_{CLV} \approx \sum_{k=0}^{n-1} 2\pi(R_1 + kd), n = \frac{R_2 - R_1}{d} \quad (2)$$

利用数项级数求和容易得到

$$L_{CLV} \approx \frac{\pi(R_2 - R_1)(R_2 + R_1 - d)}{d} \quad (3)$$

实际上, 还可以建立更简单的模型来近似计算螺旋线的长度 L_{CLV} , 比如用环形区域面积除以信道间距, 即

$$L_{CLV} \approx \frac{\pi(R_2^2 - R_1^2)}{d} \quad (4)$$

或者用同心圆的平均周长乘以总圈数, 即

$$L_{CLV} \approx 2\pi \frac{R_2 + R_1}{2} \frac{R_2 - R_1}{d} \quad (5)$$

显然, (4) 和 (5) 的结果是一样的, (3) 与它们只有微小的差别.

对于 CAV 光盘, 其螺旋线最内圈的长度可近似为 $2\pi R_1$, 总圈数可视为 $\frac{R_2 - R_1}{d}$, 于是 CAV 光盘的信道有效长度 L_{CAV} 和信息容量 C_{CAV} 为

$$L_{CAV} = 2\pi R_1 \frac{R_2 - R_1}{d}, C_{CAV} = \rho L_{CAV} \quad (6)$$

当线密度 ρ 、信道间距 d 和外径 R_2 给定后, 环形区域的内圆半径 R_1 有一个最佳选择, 使信息容量最大。

模型求解 对于 CLV 光盘, 信道长度 L_{CLV} 由 (4) 计算, 信息容量 C_{CLV} 由 (1) 计算。由图 1, 取 $R_2 = 58 \text{ mm}$, $R_1 = 22.5 \text{ mm}$, 对于表 1 的 3 种激光器得到的结果见表 2。

表 2 CLV 光盘的信息容量

激光器	激光波长/ μm	信道长度/mm	信息容量/MB	影像时间/min
红外 (CD)	0.78	5 611 179	679	18
红色 (DVD)	0.64	12 132 279	4 695	126
蓝色 (DVD)	0.41	28 055 895	22 445	603

可以看出, 对于红外激光器 (CD), 信道总长度在 5 km 以上, 信息容量约 680 MB。红色和蓝色激光器 (DVD) 信道更长, 信息容量在 GB 量级。

表 2 最后一列是按照每秒钟影像需要占用 0.62 MB, 计算出光盘可以存储多少分钟的影像。

对于由 (6) 给出的 CAV 光盘的有效长度 L_{CAV} 和信息容量 C_{CAV} , 初等数学的知识告诉我们, 当 $R_1 = R_2/2$, 即存储信息的环形区域的内圆半径是外圆半径一半时, L_{CAV} 和 C_{CAV} 最大, 为

$$L_{\text{CAV}} = \frac{\pi R_2^2}{2d}, C_{\text{CAV}} = \rho L_{\text{CAV}} \quad (7)$$

取 $R_2 = 58 \text{ mm}$, 对于表 1 的 3 种激光器用 (7) 计算的 CAV 光盘信道有效长度 L_{CAV} 和信息容量 C_{CAV} 见表 3。

表 3 CAV 光盘的信息容量

激光器	激光波长/ μm	信道长度/mm	信息容量/MB	影像时间/min
红外 (CD)	0.78	3 302 599	400	11
红色 (DVD)	0.64	7 140 755	2 764	74
蓝色 (DVD)	0.41	16 512 996	13 210	355

可以看出, 即使在内圆半径的最佳选择下, CAV 光盘的信息容量也小于 CLV 光盘。

附注 信道长度 L_{CLV} 的精确计算 (供学过极坐标和对弧长的曲线积分的读者参考)

首先建立螺旋线的方程. 对于自内径 R_1 始, 到外径 R_2 止, 间距 d 不变的螺旋线, 其在极坐标下的方程为(r 为半径, θ 为相角)

$$r(\theta) = R_1 + \frac{d}{2\pi}\theta, 0 \leq \theta \leq \frac{2\pi(R_2 - R_1)}{d} \quad (8)$$

根据极坐标下曲线弧长的微分 $ds = \sqrt{(dr)^2 + (rd\theta)^2}$, 螺旋线长度

$$L_{LCV} = \int_0^{2\pi(R_2 - R_1)/d} \sqrt{\left(R_1 + \frac{d}{2\pi}\theta\right)^2 + \left(\frac{d}{2\pi}\right)^2} d\theta \quad (9)$$

对(9)式作变量代换 $u = R_1 + \frac{d}{2\pi}\theta$ 后再作定积分, 可得

$$\begin{aligned} L_{LCV} &= \frac{2\pi}{d} \int_{R_1}^{R_2} \sqrt{u^2 + \left(\frac{d}{2\pi}\right)^2} du \\ &= \frac{2\pi}{d} \left[\frac{u}{2} \sqrt{u^2 + \left(\frac{d}{2\pi}\right)^2} + \frac{1}{2} \left(\frac{d}{2\pi}\right)^2 \ln \left(u + \sqrt{u^2 + \left(\frac{d}{2\pi}\right)^2} \right) \right] \Big|_{R_1}^{R_2} \quad (10) \end{aligned}$$

对于表 1 的 3 种激光器用(10)计算的结果为: 红外激光器(CD)的 $L_{CLV} = 5\,611\,300$ mm, 红色激光器(DVD)的 $L_{CLV} = 12\,132\,335$ mm, 蓝色激光器(DVD)的 $L_{CLV} = 28\,055\,919$ mm. 与表 2 的 L_{CLV} 比较, 红外激光器(CD)的近似计算误差仅约 0.02%, 其他的更小.

2.2 双层玻璃窗的能效

你是否注意到北方城镇的有些建筑物的窗户是双层的, 即窗户上装两层玻璃且中间留有一定空隙, 如图 1 左图所示, 两层厚度为 d 的玻璃夹着一层厚度为 l 的空气. 据说这样做是为了保暖, 即减少室内向室外的热量流失. 我们要建立一个模型来描述热量通过窗户的传导(即流失)过程, 并将双层玻璃窗与用同样多材料做成的单层玻璃窗(如图 1 右图, 玻璃厚度为 $2d$)的热量传导进行对比, 对双层玻璃窗能够减少多少热量损失给出定量分析结果^[55].

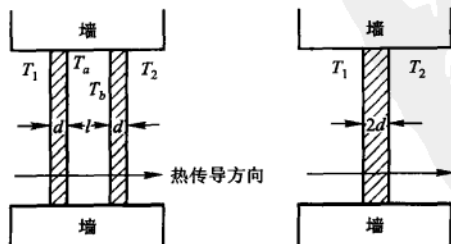


图 1 双层玻璃窗与单层玻璃窗

模型假设

1. 热量的传播过程只有传导,没有对流.即假定窗户的密封性能很好,两层玻璃之间的空气是不流动的.

2. 室内温度 T_1 和室外温度 T_2 保持不变,热传导过程已处于稳定状态,即沿热传导方向,单位时间通过单位面积的热量是常数.

3. 玻璃材料均匀,热传导系数是常数.

模型构成 在上述假设下热传导过程遵从下面的物理定律:

厚度为 d 的均匀介质,两侧温度差为 ΔT ,则单位时间由温度高的一侧向温度低的一侧通过单位面积的热量 Q 与 ΔT 成正比,与 d 成反比,即

$$Q = k \frac{\Delta T}{d} \quad (1)$$

k 为热传导系数.

记双层窗内层玻璃的外侧温度是 T_a ,外层玻璃的内侧温度是 T_b ,如图 1,玻璃的热传导系数为 k_1 ,空气的热传导系数为 k_2 ,由(1)式单位时间单位面积的热量传导(即热量流失)为

$$Q_1 = k_1 \frac{T_1 - T_a}{d} = k_2 \frac{T_a - T_b}{l} = k_1 \frac{T_b - T_2}{d} \quad (2)$$

从(2)式中消去 T_a, T_b ,可得

$$Q_1 = \frac{k_1(T_1 - T_2)}{d(s+2)}, \quad s = h \frac{k_1}{k_2}, \quad h = \frac{l}{d} \quad (3)$$

对于厚度为 $2d$ 的单层玻璃窗,容易写出其热量传导为

$$Q_2 = k_1 \frac{T_1 - T_2}{2d} \quad (4)$$

二者之比为

$$\frac{Q_1}{Q_2} = \frac{2}{s+2} \quad (5)$$

显然 $Q_1 < Q_2$. 为了得到更具体的结果,我们需要 k_1 和 k_2 的数据. 从有关资料可知,常用玻璃的热传导系数 $k_1 = 4 \times 10^{-3} \sim 8 \times 10^{-3} \text{ J/cm} \cdot \text{s} \cdot \text{kw} \cdot \text{h}$, 不流通、干燥空气的热传导系数 $k_2 = 2.5 \times 10^{-4} \text{ J/cm} \cdot \text{s} \cdot \text{kw} \cdot \text{h}$, 于是

$$\frac{k_1}{k_2} = 16 \sim 32$$

在分析双层玻璃窗比单层玻璃窗可减少多少热量损失时,我们作最保守的估计,即取 $k_1/k_2 = 16$,由(3),(5)式可得

$$\frac{Q_1}{Q_2} = \frac{1}{8h+1}, \quad h = \frac{l}{d} \quad (6)$$

比值 Q_1/Q_2 反映了双层玻璃窗在减少热量损失上的功效,它只与 $h = l/d$ 有关,图 2 给出了 $Q_1/Q_2 \sim h$ 的曲线,当 h 增加时, Q_1/Q_2 迅速下降,而当 h 超过一定值(比如 $h > 4$)后 Q_1/Q_2 下降变缓,可见 h 不必选择过大。

模型应用 这个模型具有一定应用价值.制作双层玻璃窗虽然工艺复杂会增加一些费用,但它减少的热量损失却是相当可观的.通常,建筑规范要求 $h = l/d \approx 4$.按照这个模型, $Q_1/Q_2 \approx$

3%,即双层窗比用同样多的玻璃材料制成的单层窗节约热量 97% 左右.不难发现,之所以有如此高的功效主要是由于层间空气的极低的热传导系数 k_2 ,而这要求空气是干燥、不流通的.作为模型假设的这个条件在实际环境下当然不可能完全满足,所以实际上双层窗户的功效会比上述结果差一些.另外,应该注意到,一个房间的热量散失,通过玻璃窗常常只占一小部分,热量还要通过天花板、墙壁、地面等流失.

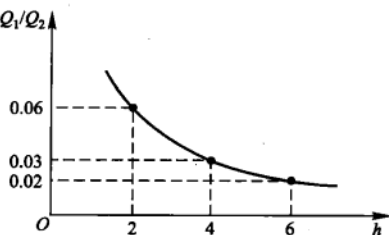


图 2 热量损失比 $\frac{Q_1}{Q_2}$ 与 $h = \frac{l}{d}$ 的关系

2.3 划艇比赛的成绩

赛艇是一种靠桨手划桨前进的小船,分单人艇、双人艇、四人艇、八人艇四种.各种艇虽大小不同,但形状相似. T. A. McMahon 比较了各种赛艇 1964—1970 年四次 2 000 m 比赛的最好成绩(包括 1964 年和 1968 年的两次奥运会和两次世界锦标赛),见表 1 第 1 至 6 列,发现它们之间有相当一致的差别,他认为比赛成绩与桨手数量之间存在着某种联系,于是建立了一个模型来解释这种关系^[10].

表 1 各种赛艇的比赛成绩和规格

艇种	2 000 m 成绩 t/min					艇长 l /m	艇宽 b /m	l/b	艇重 w_0/kg 桨手数 n
	1	2	3	4	平均				
单人	7.16	7.25	7.28	7.17	7.21	7.93	0.293	27.0	16.3
双人	6.87	6.92	6.95	6.77	6.88	9.76	0.356	27.4	13.6
四人	6.33	6.42	6.48	6.13	6.32	11.75	0.574	21.0	18.1
八人	5.87	5.92	5.82	5.73	5.84	18.28	0.610	30.0	14.7

问题分析 赛艇前进时受到的阻力主要是艇浸没部分与水之间的摩擦力. 艇靠桨手的力量克服阻力保持一定的速度前进. 桨手越多划艇前进的动力越大. 但是艇和桨手总质量的增加会使艇浸没面积加大, 于是阻力加大, 增加的阻力将抵消一部分增加的动力. 建模目的是寻求桨手数量与比赛成绩(航行一定距离所需时间)之间的数量规律. 如果假设艇速在整个赛程中保持不变, 那么只需构造一个静态模型, 使问题简化为建立桨手数量与艇速之间的关系. 注意到在实际比赛中桨手在极短的时间内使艇加速到最大速度, 然后把这个速度保持到终点, 那么上述假设也是合理的.

为了分析所受阻力的情况, 调查了各种艇的几何尺寸和质量, 表1第7至10列给出了这些数据. 可以看出, 桨手数 n 增加时, 艇的尺寸 l, b 及艇重 w_0 都随之增加, 但比值 l/b 和 w_0/n 变化不大. 若假定 l/b 是常数, 即各种艇的形状一样, 则可得到艇浸没面积与排水体积之间的关系. 若假定 w_0/n 是常数, 则可得到艇和桨手的总质量与桨手数之间的关系. 此外还需对桨手体重、划桨功率、阻力与艇速的关系等方面作出简化且合理的假定, 才能运用合适的物理定律建立需要的模型.

模型假设

1. 各种艇的几何形状相同, l/b 为常数; 艇重 w_0 与桨手数 n 成正比. 这是艇的静态特性.

2. 艇速 v 是常数, 前进时受的阻力 f 与 sv^2 成正比(s 是艇浸没部分面积). 这是艇的动态特性.

3. 所有桨手的体重都相同, 记作 w ; 在比赛中每个桨手的划桨功率 p 保持不变, 且 p 与 w 成正比.

假设1是根据所给数据作出的必要且合理的简化. 根据物理学的知识, 运动速度中等大小的物体所受阻力 f 符合假设2中 f 与 sv^2 成正比的情况. 假设3中 w, p 为常数属于必要的简化, 而 p 与 w 成正比可解释为: p 与肌肉体积、肺的体积成正比, 对于身材匀称的运动员, 肌肉、肺的体积与体重 w 成正比.

模型构成 有 n 名桨手的艇的总功率 np 与阻力 f 和速度 v 的乘积成正比, 即

$$np \propto fv \quad (1)$$

由假设2,3,

$$f \propto sv^2, \quad p \propto w$$

代入(1)式可得

$$v \propto \left(\frac{n}{s} \right)^{\frac{1}{3}} \quad (2)$$

由假设1, 各种艇几何形状相同, 若艇浸没面积 s 与艇的某特征尺寸 c 的平方成正比($s \propto c^2$), 则艇排水体积 A 必与 c 的立方成正比($A \propto c^3$), 于是有

$$s \propto A^{\frac{2}{3}} \quad (3)$$

又根据艇重 w_0 与桨手数 n 成正比, 所以艇和桨手的总质量 $w' = w_0 + nw$ 也与 n 成正比, 即

$$w' \propto n \quad (4)$$

而由阿基米德定律, 艇排水体积 A 与总质量 w' 成正比, 即

$$A \propto w' \quad (5)$$

(3), (4), (5) 式给出

$$s \propto n^{\frac{2}{3}} \quad (6)$$

将(6)式代入(2)式, 当 w 是常数时得到

$$v \propto n^{\frac{1}{9}} \quad (7)$$

因为比赛成绩 t (时间) 与 v 成反比, 所以

$$t \propto n^{-\frac{1}{9}} \quad (8)$$

(8) 式就是根据模型假设和几条物理规律得到的各种艇的比赛成绩与桨手数之间的关系。

模型检验 为了用表 1 中各种艇的平均成绩检验(8)式, 设 t 与 n 的关系为

$$t = \alpha n^{\beta} \quad (9)$$

其中 α, β 为待定常数. 由(9)式

$$\log t = \alpha' + \beta \log n, \alpha' = \log \alpha \quad (10)$$

利用最小二乘法, 根据所给数据拟合上式^①, 得到

$$t = 7.21 n^{-0.111} \quad (11)$$

可以看出(8)式与这个结果吻合得相当好.

评注 这个模型建立在一些不太精细的假设的基础上, 因为我们只关心各种艇之间的相对速度, 所以数学工具只用到比例方法. 用这种方法建模虽然不能得到关于艇速的完整的表达式, 但是对于我们的建模目的来说已经足够了. 最后的结果与实际数据吻合得如此之好, 恐怕有很大成分的巧合.

2.4 实物交换

甲有面包若干, 乙有香肠若干. 二人共进午餐时希望相互交换一部分, 达到双方满意的结果. 这种实物交换问题可以出现在个人之间或国家之间的各种类型的贸易市场上. 显然, 交换的结果取决于双方对两种物品的偏爱程度, 而偏爱程度很难给出确切的定量关系, 我们用作图的方法对双方将如何交换实物建立

^① n 只有 4 个整数点, 一般情况下不适宜作拟合, 这只是非常粗糙的办法.

一个模型^[10].

设交换前甲占有物品 X 的数量为 x_0 , 乙占有物品 Y 的数量为 y_0 , 交换后甲占有物品 X 和 Y 的数量分别为 x 和 y . 于是乙占有 X, Y 的数量为 $x_0 - x$ 和 $y_0 - y$. 这样在 xOy 平面直角坐标系上, 长方形 $0 \leq x \leq x_0, 0 \leq y \leq y_0$ 内任一点的坐标 (x, y) 都代表了一种交换方案(图 1).

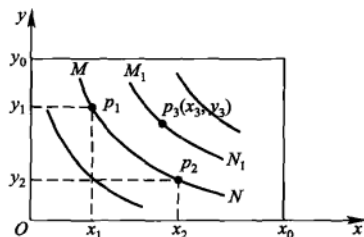


图 1 甲的无差别曲线

用无差别曲线描述甲对物品 X 和 Y 的偏爱程度. 如果占有 x_1 数量的 X 和 y_1 数量的 Y (图 1 中的 p_1 点)与占有 x_2 的 X 和 y_2 的 Y (p_2 点), 对甲来说是同样满意的话, 称 p_1 和 p_2 对甲是无差别的. 或者说 p_2 与 p_1 相比, 甲愿意以 Y 的减少 $y_1 - y_2$ 换取 X 的增加 $x_2 - x_1$. 所有与 p_1, p_2 具有同样满意程度的点组成一条甲的无差别曲线 MN , 而比这些点的满意程度更高的点如 $p_3(x_3, y_3)$ 则位于另一条无差别曲线 M_1N_1 上. 这样, 甲有无数条无差别曲线, 不妨将这族曲线记作

$$f(x, y) = c_1 \quad (1)$$

c_1 称满意度, 随着 c_1 的增加, 曲线向右上方移动. 按照常识, 无差别曲线应是单调减的(x 增加时 y 减小)、下凸的(本节最后作解释)和互不相交的(否则交点处有不同的满意度).

同样, 乙对物品 X 和 Y 也有一族无差别曲线, 记作

$$g(x, y) = c_2 \quad (2)$$

不管无差别曲线 f, g 是否有解析表达式, 每个人都可根据对两种物品的偏爱程度用曲线表示它们, 为用图解法确定交换方案提供了依据.

为得到双方满意的交换方案, 将双方的无差别曲线族画在一起. 图 2 中甲的无差别曲线族 $f(x, y) = c_1$ 如图 1, 而乙的无差别曲线族 $g(x, y) = c_2$ 原点在 O' , x, y 轴均反向, 于是当乙的满意度 c_2 增加时无差别曲线向左下移动. 这两族曲线的切点连成一条曲线 AB , 图中用点线表示. 可以断言: 双方满意的交换方案应在曲线 AB 上, AB 称交换路径. 这是因为, 假设交换在 AB 以外的某一点 p' 进行, 若通过 p' 的甲的无差别曲线与 AB 的交点为 p , 甲对 p 和 p' 的满意度相同, 而乙对 p 的满意度高于 p' , 所以双方满意的交换不可能在 p' 进行.

2.5 污水均流池的设计

城市社区的生活污水在进行净化处理之前,通常需要先进入一个集中储存的大池子,再通过水泵和输水管流向净化处理设备.这是因为生活污水的流量是时刻在变化的,而让进入净化设备的污水保持恒定的流量,才能提高净化处理的效率.这个集中储存污水的池子能够起到均衡调节流量的作用,不妨称之为均流池.

这里要讨论的是,怎样根据社区污水的流量来设计均流池的容积及水泵和输水管的规格,并且在一定条件下怎样按照施工成本最小的原则来确定均流池的具体尺寸^[44].

问题的调查与分析 除了节假日等特殊情况以外,社区生活污水进入均流池的流量可以看作是以天为周期变化的,通过典型调查可以得到至少以小时为单位间隔、一天的污水流量.现在已经有表1的调查结果.我们看到,约在凌晨3时流量最低,其后一路上升,流量的高峰出现在下午5时至7时,之后呈下降趋势.

表1 社区一天以小时为单位间隔的生活污水流量(单位: m^3/s)

时间/h	0	1	2	3	4	5	6	7
流量/ $(\text{m}^3 \cdot \text{s}^{-1})$	0.041 7	0.032 1	0.023 6	0.018 5	0.018 9	0.019 9	0.022 8	0.036 9
时间/h	8	9	10	11	12	13	14	15
流量/ $(\text{m}^3 \cdot \text{s}^{-1})$	0.051 4	0.063 0	0.068 5	0.069 7	0.072 5	0.075 4	0.076 1	0.077 5
时间/h	16	17	18	19	20	21	22	23
流量/ $(\text{m}^3 \cdot \text{s}^{-1})$	0.081 0	0.083 9	0.086 3	0.080 7	0.078 1	0.069 0	0.058 4	0.051 9

由表1不难得到污水一天进入均流池的平均流量,显然,这个平均流量就应该是从均流池用水泵打入净化设备的恒定流量,水泵和输水管的规格需按照这个流量并考虑留有一定裕量来设计.

根据表1以小时为单位间隔的污水流入量和从均流池到净化设备的恒定流出量,可以得到均流池中污水(以小时为单位间隔,随时间变化)的容量,均流池的容积应该按照污水的最大容量并考虑留有一定裕量来设计.

均流池的形状一般为矩形,其深度通常按照工程需要(底部需安装设备、进行清理等)确定,于是均流池的面积可以由它的容积和深度得到.均流池的施工成本中除了底部单位面积的成本以外,由于具体地理、地形条件的限制,四条边上单位长度的施工成本也会有不同.均流池的深度以及施工成本均可经调查得到.

模型假设与建立 由问题分析及调查结果作如下假设:

1. 以表1社区一天的生活污水流量为依据,并留有25%的裕量进行设计.

2. 均流池的深度确定为 3 m, 底部施工面积的成本为 340 元/ m^2 , 两条长边及一条短边的施工长度的成本为 250 元/m, 另一条短边的施工长度的成本为 450 元/m.

按照问题的要求, 需要建立确定均流池的恒定流出量和最大容量, 以及均流池的具体尺寸两个模型.

模型 1 均流池的恒定流出量和最大容量模型

首先将表 1 的流量乘以 3 600 s/h, 流量单位由 m^3/s 换算成 m^3/h , 得表 2.

表 2 社区一天以小时为单位间隔的生活污水流量(单位: m^3/h)

时间/h	0	1	2	3	4	5	6	7
流量/ $(\text{m}^3 \cdot \text{h}^{-1})$	150.12	115.56	84.96	66.60	68.04	71.64	82.08	132.84
时间/h	8	9	10	11	12	13	14	15
流量/ $(\text{m}^3 \cdot \text{h}^{-1})$	185.04	226.80	246.60	250.92	261.00	271.44	273.96	279.00
时间/h	16	17	18	19	20	21	22	23
流量/ $(\text{m}^3 \cdot \text{h}^{-1})$	291.60	302.04	310.68	290.52	281.16	248.40	210.24	186.84

记表 2 中每小时污水流入均流池的流量为 $f(t)$, $t=0, 1, 2, \dots, 23$, 计算一天的平均流量, 记作 g ,

$$g = \frac{1}{24} \sum_{t=0}^{23} f(t) \quad (1)$$

由表 2 计算出 $g=203.67 \text{ m}^3/\text{h}$. g 就是从均流池到净化设备的恒定流量. 若考虑 25% 的裕量, 可按照 $255 \text{ m}^3/\text{h}$ 的流量来设计水泵和输水管的规格.

图 1 是均流池的流入量 $f(t)$ 和恒定流出量(即平均流入量) g 的图形, 每天在大约 9 时到 23 时 $f(t) > g$, 其余时段 $f(t) < g$.

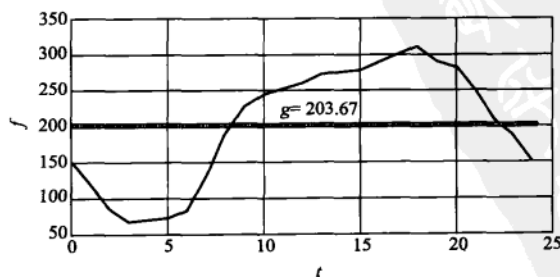


图 1 均流池流入量 $f(t)$ 和恒定流出量 g

记均流池中污水的容量为 $c(t)$, $t=0, 1, 2, \dots, 23$, 显然 $c(t)$ 与流入量 $f(t)$ 和恒定流出量 g 之间的关系为

$$c(t+1) = c(t) + f(t) - g, t = 0, 1, 2, \dots, 23 \quad (2)$$

为了按照(2)式由表2的 $f(t)$ 和 $g = 203.67 \text{ m}^3/\text{h}$ 计算容量 $c(t)$, 需要初值 $c(0)$, 但是我们暂时不知道它, 如果简单地设 $c(0) = 0$, 可以得到 $c(t)$ 如表3(由周期性, $c(24) = c(0)$).

表3 均流池中污水一天以小时为单位间隔的容量(设 $c(0) = 0$)

时间/h	0	1	2	3	4	5
容量/ m^3	0	-53.55	-141.66	-260.37	-397.44	-533.07
时间/h	6	7	8	9	10	11
容量/ m^3	-665.10	-786.69	-857.52	-876.15	-853.02	-810.09
时间/h	12	13	14	15	16	17
容量/ m^3	-762.84	-705.51	-637.74	-567.45	-492.12	-404.19
时间/h	18	19	20	21	22	23
容量/ m^3	-305.82	-198.81	-111.96	-34.47	10.26	16.83

表3中均流池中污水的容量多为负数, 显然是初值 $c(0) = 0$ 设置不当. 我们看到, 容量的最小值出现在 $c(9) = -876.15$. 如果将最小容量设置为0, 即初值设为 $c(0) = 876.15$, 得到的 $c(t)$ 如表4, 这是一个合理的结果.

表4 均流池中污水一天以小时为单位间隔的容量(设 $c(0) = 876.15$)

时间/h	0	1	2	3	4	5	6	7
容量/ m^3	876.15	822.60	734.49	615.78	478.71	343.08	211.05	89.46
时间/h	8	9	10	11	12	13	14	15
容量/ m^3	18.63	0	23.13	66.06	113.31	170.64	238.41	308.70
时间/h	16	17	18	19	20	21	22	23
容量/ m^3	384.03	471.96	570.33	677.34	764.19	841.68	886.41	892.98

图2是均流池中污水容量 $c(t)$ 的图形, 显然, 上面的曲线(设 $c(0) = 876.15$) 是下面曲线(设 $c(0) = 0$) 向上平移的结果.

将图2与图1比较可以看出, 当流量由 $f(t) < g$ 变为 $f(t) > g$ (约9时), 容量 $c(t)$ 达到最小值; 当流量由 $f(t) > g$ 变为 $f(t) < g$ (约23时), 容量 $c(t)$ 达到最大值. 请读者根据(2)式给予解释.

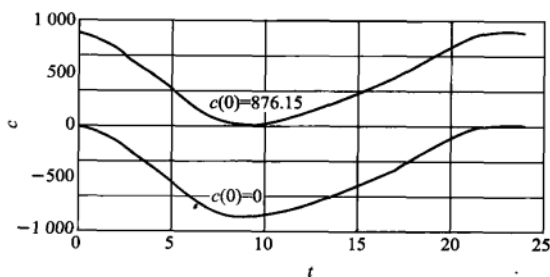


图2 均流池中污水容量 $c(t)$ 的图形

由表4得到,最大容量为 892.98 m^3 ,若考虑25%的裕量,可按照 1116 m^3 来设计均流池的具体尺寸.

我们建立模型(2)并得到表4和图2,是为了考察均流池容量的变化.实际上并不需要(2)式,只要将大于平均流量 g 的各个流量(表2中从 $f(9)$ 到 $f(22)$)减去平均流量相加,即可得到最大容量.

模型2 均流池的具体尺寸模型

根据上面得到的均流池容积 1116 m^3 及假设2,按照施工成本最小的原则来建立模型,确定均流池的具体尺寸.记均流池的长边长度为 l ,短边长度为 w ,施工成本应为

$$S(l, w) = 340lw + 250(2l + w) + 450w \quad (3)$$

根据深度为3m可知均流池的面积 $lw = 1116/3 = 372 \text{ m}^2$,于是 $w = 372/l$, (3)式简化为

$$\begin{aligned} S(l) &= 340 \times 372 + 250(2l + 372/l) + 450 \times 372/l \\ &= 126480 + 500l + 260400/l \end{aligned} \quad (4)$$

利用初等数学知识可以得到,当 $l = \sqrt{260400/500} = 22.82 \text{ m}$ 时 $S(l)$ 最小,为149301元,且 $w = 372/22.82 = 16.30 \text{ m}$.实际上可以建造一个 $23 \text{ m} \times 16.5 \text{ m}$ 的均流池,成本约15万元.

进一步考虑的问题

1. 表1是一天的调查结果,为了增强可靠性,应该调查若干天,用统计方法确定社区的污水流量.

2. 流量的测量数据是离散的,实际的流量是连续的,可以作插值和数值积分,按照连续模型考虑均流池的容量(习题3).

3. 将最小容量设置为0是假定均流池的出水管紧贴底部,实际情况虽不一定如此,但是只要将图2的曲线再向上平移,或者直接将3m的深度适当增加即可.

2.6 交通流与道路通行能力

现代城市生活中交通拥堵是普遍存在的现象,在许多平面交叉路口,红灯后面总是排着长长的汽车队伍等待放行.道路的扩建似乎永远赶不上汽车增加的速度,通过信号灯控制等管理手段提高道路通行能力,已经成为城市交通工程面临的重要课题之一.本节首先介绍交通流的几个基本参数及它们之间的关系,然后讨论一般道路及信号灯控制的十字路口的通行能力^[95].

交通流的基本参数及其特性

为明确和简单起见,这里的交通流均指由标准长度的小型汽车在单方向的道路上行驶而形成的车流,没有外界因素如岔路、信号灯等的影响.

借用物理学的概念,将交通流近似看作一辆辆汽车组成的连续的流体,可以用流量、速度、密度这3个参数描述交通流的基本特性.

流量 q 指某时刻单位时间内通过道路指定断面的车辆数,通常以辆/h为单位.

速度 v 指某时刻通过道路指定断面的车辆速度,通常以 km/h 为单位.

密度 k 指某时刻通过道路指定断面单位长度内的车辆数,通常以辆/km 为单位.

虽然一般说来流量、速度和密度都是时间和地点的函数,但是在讨论指定时段(如早高峰)、指定路段或路口的交通状况时,可以认为交通流是稳定的,即流量、速度和密度都是常数,与时间和地点无关.

根据物理学的基本常识,流量 q 、速度 v 和密度 k 显然满足

$$q = vk \quad (1)$$

例如在高速公路上以 100 km/h 的车速、200 m 的车距行驶的车流,其流量为 500 辆/h. (1) 式是这 3 个参数之间的基本关系.

经验和观测告诉我们,速度与密度之间存在密切关系,当道路上车辆增多,车流密度加大时,司机就被迫降低车速. 1935 年, Greenshields 通过对观测数据的统计分析,提出车速与密度之间的一个线性模型

$$v = v_f(1 - k/k_j) \quad (2)$$

其中 v_f 是密度 $k=0$ 时的车速,即理论上的最高车速,称畅行车速(自由流), k_j 是速度 $v=0$ 时的密度,称阻塞密度. 实际上,模型(2)可以在一些仔细的分析与合理的假设下推导出来,将在后面作为附注给出.

研究和经验表明,线性模型适合通常的车流密度适中的情况. 此后有人提出了密度较大时适用的对数模型

$$v = v_1 \ln(k_j/k) \quad (3)$$

及密度较小时适用的指数模型

$$v = v_f \exp(-k/k_j) \quad (4)$$

(3) 中的 v_i 理论上是密度 $k = k_j/e$ 时的车速, 实际上要由观测数据确定。

将常用的线性模型(2)代入(1), 得到流量与密度的关系

$$q = v_f k (1 - k/k_j) \quad (5)$$

这是一条抛物线, 车流密度 k 由小变大时流量增加, 当 $k = k_j/2$, 即阻塞密度的一半时, 流量最大, 密度 k 继续变大, 流量减小。

由(1)、(2)可以导出流量与车速之间的关系

$$q = k v (1 - v/v_f) \quad (6)$$

也是一条抛物线, 最大流量出现在车速 $v = v_f/2$, 即畅行车速的一半处。

交通工程中常将由(2)、(5)、(6)式确定的流量、速度和密度之间的关系用图1表示, 其中 q_m 是最大流量, k_m, v_m 分别是对应于最大流量的密度和速度。

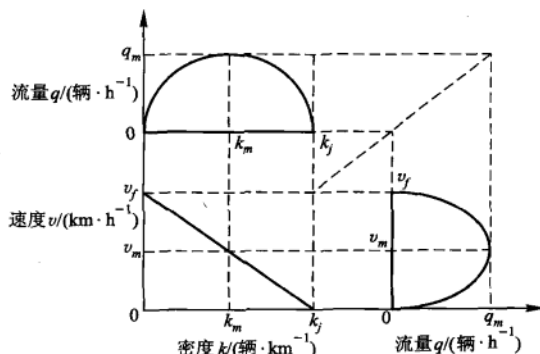


图1 交通流的流量、速度、密度关系图

城市干道的通行能力

道路通行能力指单位时间内通过道路上指定断面的最大车辆数, 是度量道路疏导交通能力的指标, 它由道路设施、交通服务、环境、气候等诸多条件决定, 其数值是相对稳定的。在正常运行状况下, 当交通流量远小于通行能力时, 车速高, 呈自由流状态; 随着车辆的增加, 运行状况恶化, 当流量接近通行能力时, 车速降低, 呈强制流状态, 出现交通拥堵。

通行能力表示道路的容量, 反映服务水平; 交通流量表示道路的负荷, 反映交通需求。流量与通行能力的比值可用来表征道路的负荷程度, 称饱和度或利用率。

下面讨论城市干道的基本通行能力, 它指在理想的道路和交通条件下, 当具有标准长度和技术指标的车辆, 以前后两车最小车头间隔连续行驶时, 单位时间

内通过道路上指定断面的最大车辆数,记作 N (辆/h)。

记车速为 v (km/h),前后两车最小车头间隔为 d (m),显然

$$N = 1\,000v/d \quad (7)$$

最小车头间隔主要由刹车距离决定,而刹车距离又与车速密切相关。交通工程中常用如下公式计算最小车头间隔 d

$$d = d_1 + d_2 + d_3 + d_4 = vt_0 + cv^2 + d_3 + d_4 \quad (8)$$

其中 d_1 是刹车时司机在反应时间 t_0 内汽车行驶的距离, d_2 是刹车时从制动器开始起作用到汽车完全停止行驶的距离,称制动距离, c 是与车辆自重、路面阻力、湿度、坡度等诸多因素有关的系数, d_3 是两车之间的安全距离, d_4 是车辆的标准长度。这里,制动距离 d_2 与车速 v 的关系 $d_2 = cv^2$,可以在合理的假设下推导出来,将在下面给出。

将(8)代入(7)并注意 v (km/h)和 d (m), t_0 (s)的单位换算,可得

$$N = \frac{1\,000}{\frac{t_0}{3.6} + cv + \frac{d_3 + d_4}{v}} \quad (9)$$

参考交通工程的专业教材,在一般情况下可取司机刹车的反应时间 $t_0 = 1$ s,系数 $c = 0.01$ ①,安全距离 $d_3 = 2$ m,小型车辆的标准长度 $d_4 = 5$ m,按照(9)计算的结果如表 1。

表 1 城市干道的基本通行能力 N 与车速 v 的关系

$v/(\text{km} \cdot \text{h}^{-1})$	10	20	30	40	50	60	70	80	90	100
$N/(\text{辆} \cdot \text{h}^{-1})$	958	1 208	1 233	1 173	1 090	1 006	928	858	797	742

利用初等数学知识,由(9)式可以得到,当车速 $v = \sqrt{\frac{d_3 + d_4}{c}}$ 时通行能力 N

达到最大值 $N_m = \frac{1\,000}{\frac{t_0}{3.6} + 2\sqrt{c(d_3 + d_4)}}$, 当 t_0, c, d_3, d_4 变大时最大通行能力 N_m

减小。

制动距离与车速的模型

下面从机理分析的角度建立制动距离模型 $d_2 = cv^2$ 。

制动距离与车重、车速、制动器作用力以及道路、气候等因素有关。制动器

① 由 v (km/h)和 d (m), c 的单位应为 mh^2/km^2 ;按照[95], c 还与车速 v 有关, $c = 0.01$ 相当于次高级路面潮湿状态下 $v = 40$ km/h 时的取值,当 v 提高时 c 变大, $v = 100$ km/h 时 $c = 0.013$,若如此,表 1 中的通行能力 N 将略有变化。

是一个能量耗散装置,刹车时制动力作的功用以抵消汽车动能的减少.设计制动器的一个合理原则是,最大制动力大体上与车的质量成正比,使汽车的减速度基本上是常数,这样,司机和乘客少受剧烈的冲击.这里不考虑道路、气候等因素.

可作以下假设:刹车时使用最大制动力 F , F 作的功等于汽车动能的改变,且 F 与车的质量 m 成正比.

在制动力 F 作用下汽车行驶距离 d_2 作的功为 Fd_2 ,而车速从 v 变成 0,动能的变化为 $mv^2/2$,根据假设有

$$Fd_2 = mv^2/2 \quad (10)$$

又因为刹车时的加速度 a 为常数,由牛顿第二定律知 $F = ma$,代入(10)即得

$$d_2 = cv^2 \quad (11)$$

其中 c 是比例系数.

信号灯控制的十字路口的通行能力

交通高峰时段,在许多城市主干道的十字路口,长长的车队等待着绿灯放行是司空见惯的现象.图 2 是一个典型的十字路口,东西方向有 3 条车道,分别为左转、直行、直右混行,南北方向有 2 条车道,分别为左转、直右混行.

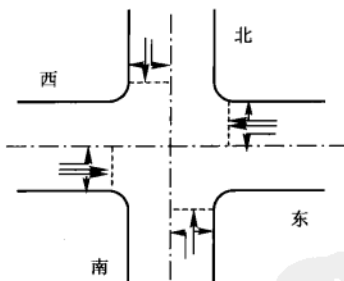


图 2 十字路口车道示意图

城市交通中信号灯通常采用周期控制,在一个周期内,通过不同颜色的信号灯组合,控制不同方向车辆的通行(绿灯)与禁行(红灯).每一种信号灯组合称为一个相位,对于图 2 这样的十字路口,信号灯控制可以采用图 3 所示的 4 相位方案.

在图 3 的相位 A,东西向直行和右转通行(绿灯),其余方向禁行(红灯);在相位 B,东西向左转通行(绿灯),其余方向禁行(红灯);相位 C 和相位 D 有类似解释.在两个相位转换之间有相应的黄灯时间,只准已越过停止线的车辆继续通行,黄灯时间长度与路口的具体条件有关,这里不予讨论.

信号灯控制的道路通行能力(以下均指单行车道),是指在某一相位下单位

时间通过停止线的最大车辆数,它是在这样的假设下计算的:红灯时车辆已在停止线后排队,绿灯后第1辆车立即启动通过停止线,其余车辆按照固定的(或视为平均的)时间间隔通过停止线。

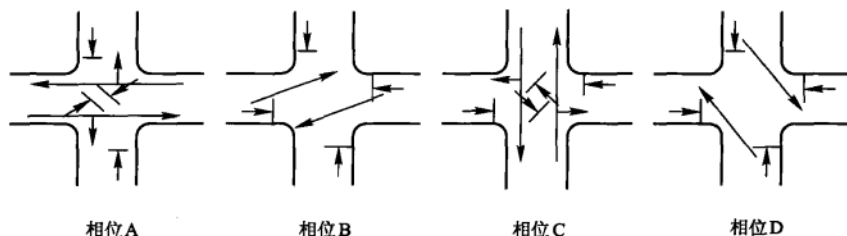


图3 十字路口信号灯控制的4相位方案

记信号灯周期为 $T(s)$,某相位的绿灯时间为 $t_g(s)$,绿灯后第1辆车通过停止线的时间为 $t_0(s)$,直行或右转车辆通过停止线的时间为 $t_s(s)$,反映车辆通过路口不均匀性的折减系数为 φ ,某相位下每小时通过停止线的最大车辆数为 S (辆/h),则有

$$S = \varphi \frac{3600}{T} \left(\frac{t_g - t_0}{t_s} + 1 \right) \quad (12)$$

按照[95], t_0 取 2.3 s, t_s 取 2.5 s(小型车辆)~3.5 s(大型车辆),对直行或右转 φ 取 0.9,左转 φ 更小。某相位的绿灯时间与信号灯周期之比称绿信比,记作 $G = t_g/T$,由(12)可以看出, S 大致等于按照每 $t_s(s)$ 通过一辆车计算的小时流量($3600/t_s$),乘以绿信比 G 和折减系数 φ ,这是与人们的常识相符的。

为了利用图3信号灯控制的4相位方案,使十字路口各方向的交通尽量通畅和均衡,应该先实地调查高峰时段各方向车辆的实际流量,如4个相位通行的实际流量记为 q_A, q_B, q_C, q_D (对相位A,当东西两向的流量不同时,可取两向平均或其中最大者,其他相位同此),将其与最大流量($3600/t_s$)之比作为各相位的交通饱和度,记为 p_A, p_B, p_C, p_D ,然后调整4个相位的绿信比 G_A, G_B, G_C, G_D ,使之与饱和度 p_A, p_B, p_C, p_D ,大致成比例,并适当考虑不同方向通行的折减系数。

评注 下面从机理分析的角度建立车速与车流密度的线性模型 $v = v_f(1 - k/k_j)$ (供学过微分方程的读者参考)。

让我们先分析行驶过程中车辆的位置、速度与受力之间的关系。将车流中第 n 辆车的位置和速度分别记作 $x_n(t)$ 和 $v_n(t)$,在稳定状态下每辆车的速度和车流密度都是常数。若前面第 $n-1$ 辆车突然减速,稳定状态被破坏,则第 n 辆车将施加制动力,随之减速。当两车速度差越大时制动力越大,同时,当车流拥

挤,即两车间隔越小时制动力也越大,经制动作用后稳定状态恢复.若前面第 $n-1$ 辆车突然加速,第 n 辆车施加驱动力的情况与此类似.可作如下简化假设:

1. 车速 v 是车流密度 k 的函数, $k=0$ 时 $v=v_f$ (畅行车速), $k=k_j$ (阻塞密度)时 $v=0$.

2. 在稳定状态下车速 v 及相邻两车的车头间隔 d 都相同,因而车流密度 $k=1/d$ 是常数.

3. 当第 $n-1$ 辆车减速或加速致使稳定状态被破坏时,第 n 辆车施加的制动力或驱动力与两车速度差成正比,与两车间隔的平方成反比,制动或驱动后稳定状态恢复.

根据牛顿第二定律和假设 3 可以写出微分方程

$$\frac{dv_n}{dt} = \lambda \frac{v_n - v_{n-1}}{(x_n - x_{n-1})^2} \quad (13)$$

其中 λ 是比例系数.注意到 $v_n(t)$ 和 $x_n(t)$ 之间的导数关系, (13) 可写作

$$\frac{dv_n}{dt} = -\lambda \frac{d}{dt} \left(\frac{1}{x_n - x_{n-1}} \right) \quad (14)$$

对 (14) 两边积分可得

$$v_n(t) = -\lambda \frac{1}{x_n(t) - x_{n-1}(t)} + c \quad (15)$$

其中 c 是待定常数.

根据假设 2, 稳定状态恢复后 $v_n(t) = v, x_n(t) - x_{n-1}(t) = d = 1/k$, 于是 (15) 式为

$$v = -\lambda k + c \quad (16)$$

利用假设 1 的条件确定 (16) 式中的 λ 和 c , 即得到车速与车流密度的线性模型.

对数模型 (3) 也可以类似地从机理分析的角度推导出来 (习题 4).

2.7 核军备竞赛

在 20 世纪六七十年代的冷战时期,美苏两个核大国都声称为了保卫自己的安全,而实行所谓核威慑战略,核军备竞赛不断升级.随着苏联的解体和冷战的结束,双方通过了一系列的核裁军协议,2010 年 4 月 8 日,美国与俄罗斯领导人在捷克首都布拉格签署新的削减战略武器条约.根据这项新条约,美国和俄罗斯将在 7 年内将各自部署的战略核武器削减到不超过 1 550 枚,并把各自的战略核武器运载工具削减到不超过 700 枚.

在什么情况下双方的核军备竞赛才不会无限扩张而存在暂时的平衡状态,处于这种平衡状态下双方拥有最少的核武器数量是多大,这个数量受哪些因素影响,当一方采取诸如加强防御、提高武器精度、发展多弹头导弹等措施时,平衡状态会发生什么变化?本节将介绍一个定性的模型,在给核威慑战略作出一些合理、简化的假设下,对双方核武器的数量给以图形(结合式子)的描述,粗略地回答上述问题^[10,32].

模型假设 以双方的(战略)核导弹数量为对象,描述双方核军备的大小,假定双方采取如下同样的核威慑战略:

1. 认为对方可能发起所谓第一次核打击,即倾其全部核导弹攻击己方的核导弹基地。

2. 己方在经受第一次核打击后,应保存有足够的核导弹,给对方的工业、交通中心等目标以毁灭性的打击。

在任一方实施第一次核打击时,假定一枚核导弹只能攻击对方的一个核导弹基地,且摧毁这个基地的可能性是常数,它由一方的攻击精度和另一方的防御能力所决定。

图的模型 记 $y=f(x)$ 为甲方拥有 x 枚核导弹时,乙方采取核威慑战略所需的最小核导弹数, $x=g(y)$ 为乙方拥有 y 枚核导弹时,甲方采取核威慑战略所需的最小核导弹数,不妨让我们看看曲线 $y=f(x)$ 应该具有什么性质。

当 $x=0$ 时 $y=y_0$, y_0 是甲方在实施第一次核打击后已经没有核导弹时,乙方为毁灭甲方的工业、交通中心等目标所需的核导弹数,以下简称乙方的威慑值;当 x 增加时 y 应随之增加,并且由于甲方的一枚核导弹最多只能摧毁乙方的一个核导弹基地,所以 $y=f(x)$ 不会超过直线

$$y = y_0 + x \quad (1)$$

这样,曲线 $y=f(x)$ 应在图 1 所示的范围内,可以猜想它的大致形状如图 2。

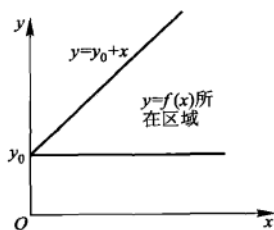


图 1 曲线 $y=f(x)$ 的范围

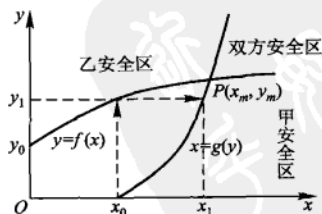


图 2 安全区、安全线和平衡点

曲线 $x=g(y)$ 应有类似的性质 ($y=0$ 时 $x=x_0$, $x=g(y)$ 不超过直线 $x=x_0+y$), 图 2 中将两条曲线画在一起, 可以知道它们会相交于一点, 记交点为 $P(x_m, y_m)$, 我们讨论 P 点的含义。

根据 $y=f(x)$ 的定义, 当 $y \geq f(x)$ 时乙方是安全的 (在核威慑战略意义下), 不妨称该区域为乙安全区, 曲线 $y=f(x)$ 为 (临界情况下的) 乙安全线. 类似地, $x \geq g(y)$ 的区域为甲安全区, $x=g(y)$ 为甲安全线. 两个安全区的公共部分即为双方安全区, 是核军备竞赛的稳定区域, 而 P 点的坐标 x_m 和 y_m 则为稳定状态下甲乙双方分别拥有的最小核导弹数, P 点是平衡点.

平衡点怎样达到呢? 不妨假定甲方最初只有 x_0 枚导弹 (威慑值), 乙方为了自己的安全至少要拥有 y_1 枚导弹, 见图 2, 而甲方为了安全需要将导弹数量增加到 x_1 , 如此下去双方的导弹数量就会趋向 x_m, y_m .

模型的精细化 为了研究 x_m 和 y_m 的大小与哪些因素有关, 这些因素改变时平衡点如何变动, 我们尝试寻求 $y=f(x)$ 和 $x=g(y)$ 的具体形式.

若 $x < y$, 当甲方以全部 x 枚核导弹攻击乙方的 y 个核基地中的 x 个时, 记每个基地未被摧毁的概率为 s , 以下简称乙方的残存率, 则乙方 (平均) 有 sx 个基地未被摧毁, 且有 $y-x$ 个基地未被攻击, 二者之和即为乙方经受第一次核打击后保存下来的核导弹数, 它应该就是图的模型中的威慑值 y_0 , 即 $y_0 = sx + y - x$, 于是

$$y = y_0 + (1-s)x \quad (2)$$

由 $0 < s < 1$ 知直线 (2) 的斜率小于直线 (1) 的斜率, 如图 3.

当 $x=y$ 时显然有 $y_0 = sy$, 所以有

$$y = \frac{y_0}{s} \quad (3)$$

若 $y < x < 2y$, 当甲方以全部 x 枚核导弹攻击乙方的 y 个核基地时, 乙方的 $x-y$ 个将被攻击 2 次, 其中 $s^2(x-y)$ 个未被摧毁, 且有 $y-(x-y)=2y-x$ 个被攻击 1 次, 其中 $s(2y-x)$ 个未被摧毁, 二者之和即为图的模型中的 y_0 , 即 $y_0 = s^2(x-y) + s(2y-x)$, 于是

$$y = \frac{y_0}{s(2-s)} + \frac{1-s}{2-s}x \quad (4)$$

直线 (4) 的斜率小于直线 (2) 的斜率, 如图 3.

当 $x=2y$ 时显然有 $y_0 = s^2y$, 所以有

$$y = \frac{y_0}{s^2} \quad (5)$$

虽然上述过程可以继续下去, 但是如果我们允许 x, y 取连续值, 考察 $x=ay$, a 为大于零的任意实数, 表示乙 (临界) 安全条件下甲乙双方导弹数量之比, 那么

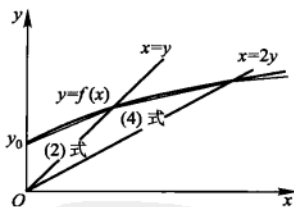


图 3 曲线 $y=f(x)$ 的形成

由 $x=y$ 时的(3)式和 $x=2y$ 时的(5)式可以设想 $y=f(x)$ 的形式为^①

$$y = \frac{y_0}{s^a} = \frac{y_0}{s^{x/y}}, 0 < s < 1 \quad (6)$$

它应该是图3中的光滑曲线,利用微积分的知识可以证明这是一条上凸的曲线.

$x=g(y)$ 有类似的形式,曲线是向右凸的,当然,其中的 s 应为甲方的残存率.

由此可知,这样两条曲线 $y=f(x)$ 和 $x=g(y)$ 必定相交,并且交点惟一.

进一步研究由(6)式表示的乙安全线 $y=f(x)$ 的性质(习题5):

若威慑值 y_0 变大,则曲线整体上移,且变陡;若残存率 s 变大,则曲线变平.

甲安全线 $x=g(y)$ 有类似的性质.利用这些性质可以用上述模型解释核军备竞赛中平衡点 $P(x_m, y_m)$ 的变化.

模型解释

1. 若甲方增加经费保护及疏散工业、交通中心等目标,则乙方的威慑值 y_0 将变大,而其他因素不变,那么乙安全线 $y=f(x)$ 的上移会使平衡点变为 $P'(x'_m, y'_m)$,如图4.显然有 $x'_m > x_m, y'_m > y_m$,说明虽然甲方的防御是被动的,但也会使双方的军备竞赛升级.

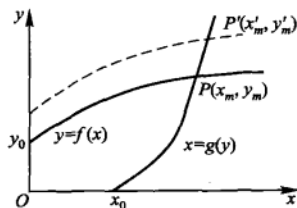


图4 模型解释1的示意图

2. 若甲方将原来的固定核导弹基地改进为可移动发射架,则乙安全线 $y=f(x)$ 不变(试说明其威慑值 y_0 、残存率 s 均不变),而甲方的残存率变大(威慑值 x_0 不变),于是甲安全线 $x=g(y)$ 向 y 轴靠近,平衡点变为 $P'(x'_m, y'_m)$,如图5.显然有 $x'_m < x_m, y'_m < y_m$,说明甲方的这种单独行为,会使双方的核导弹减少.

3. 若双方都发展多弹头导弹,每个弹头可以独立地摧毁目标,设 x, y 仍为双方核导弹的数量,则双方的威慑值 x_0, y_0 和残存率均减小.乙安全线由于 y_0 的减小而下移且变平,又由于残存率的变小,使曲线变陡.甲安全线有类似的变化,

① 用它近似代替(2)、(4)等表示的分段直线,虽然由下面的(6)式并无法解出显函数 $y=f(x)$ 的形式.

二者的综合影响则可能使平衡点变为 $P'(x'_m, y'_m)$ 或 $P''(x''_m, y''_m)$, 出现图6所示的两种情况, 究竟会使双方的核导弹增加还是减少, 需要更多的信息及更详细的分析.

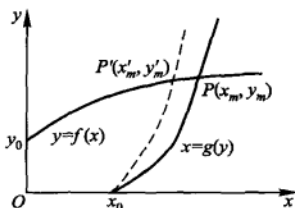


图5 模型解释2的示意图

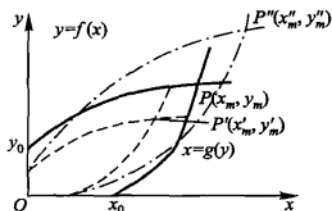


图6 模型解释3的示意图

评注 核军备竞赛问题初看起来似乎与数学无缘, 但是如果把所谓核威慑战略作一些合理、简化的假设, 就能够用一个简单的图的模型, 来描述双方核武器数量相互制约、达到平衡的过程. 本节还通过更精细的分析找到影响安全曲线的参数: 威慑值和残存率, 给出了曲线的表达式, 并由此对核军备竞赛中的一些现象作出解释, 这种由粗及细、从定性到定量(在一定程度上)的建模方法是值得借鉴的.

2.8 扬帆远航

海面上东风劲吹, 帆船要从A点驶向正东方的B点(图1). 稍有常识的人都知道, 为了借助风力, 船应该先朝东北方向前进, 然后再转向东南方, 才能到达B点. 这里要解决的问题是, 确定起航时的航向 θ (与正东方向的夹角)及帆的朝向 α (帆面与航向的夹角)^[89].

模型分析 帆船在航行过程中既受到风通过帆对船的推力, 又受到风对船体的阻力, 需要对这两种力作合理、简化的分解, 找出它们在航向的分力(图2).

风的推力 w 分解为 $w = w_1 + w_2$, 其中 w_1 垂直于帆, w_2 平行于帆. w_1 又分解为 $w_1 = f_1 + f_2$, f_1 即为风在航向的推力. 风的阻力 p 分解为 $p = p_1 + p_2$, 其中 p_1 为风在航向的阻力. 因为 f_1 和 p_1 的方向相反, 所以船受到的净推力为 $f_1 - p_1$.

按照流体力学的知识, 在航行速度不大的情况下, 航速与净推力成正比. 于是航向 θ 和帆的朝向 α 的确定, 应该使船在正东方向的速度, 即净推力在正东方向的分力达到最大.

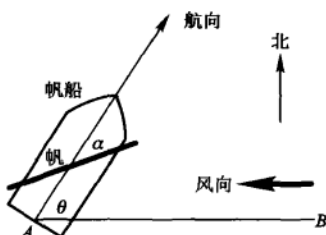


图1 帆船航行示意图

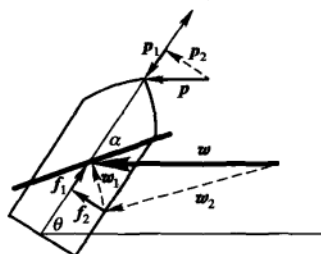


图2 帆船受力分析图

模型假设 记帆的迎风面积为 s_1 , 船的迎风面积为 s_2 .

1. 风通过帆对船的推力 w 与 s_1 成正比, 风对船体的阻力 p 与 s_2 成正比, 比例系数相同 (记作 k), 且 s_1 远大于 s_2 .

2. w 的分力 w_2 与帆面平行, 可以忽略.

3. 分力 f_2 和 p_2 垂直于船身, 可以被船舵抵消, 不予考虑.

4. 航速 v 与净推力 $f = f_1 - p_1$ 成正比, 比例系数记作 k_1 .

模型建立 根据模型假设和图2表示的各个力之间的几何关系, 容易得到

$$w = ks_1, \quad p = ks_2 \quad (1)$$

$$w_1 = w \sin(\theta - \alpha), \quad f_1 = w_1 \sin \theta = w \sin(\theta - \alpha) \sin \theta \quad (2)$$

$$p_1 = p \cos \theta \quad (3)$$

$$v = k_1 (f_1 - p_1) \quad (4)$$

记船在正东方向的速度分量为 v_1 , 则

$$v_1 = v \cos \theta = k_1 (f_1 - p_1) \cos \theta \quad (5)$$

问题是确定 θ 和 α , 使 v_1 最大.

模型求解 这本来是一个二元函数的极值问题, 但是由(3), (5)式知, p_1 与 α 无关, 首先只需在 θ 固定时使 f_1 最大, 解出 α , 然后再求 θ 使 v_1 最大. 用初等数学的办法即可求解.

由(2)式 f_1 可化为

$$f_1 = w [\cos(\theta - 2\alpha) - \cos \theta] / 2 \quad (6)$$

所以当

$$\alpha = \theta / 2 \quad (7)$$

时

$$f_1 = w(1 - \cos \theta) / 2 \quad (8)$$

最大.

将(7), (8), (3)代入(5)式得

$$v_1 = k_1 [w(1 - \cos \theta) / 2 - p \cos \theta] \cos \theta$$

$$= (k_1 w/2) [1 - (1 + 2p/w) \cos \theta] \cos \theta \quad (9)$$

注意到(1)式,记

$$k_2 = k_1 w/2, \quad t = 1 + 2p/w = 1 + 2s_2/s_1 \quad (10)$$

则(9)式为

$$v_1 = k_2 (1 - t \cos \theta) \cos \theta = k_2 t \left[\frac{1}{4t^2} - \left(\cos \theta - \frac{1}{2t} \right)^2 \right] \quad (11)$$

显然

$$\cos \theta = 1/2t \quad (12)$$

时 v_1 最大. 由(10)式,且 s_1 远大于 s_2 ,可知

$$1/4 < \cos \theta < 1/2, \quad 60^\circ < \theta < 75^\circ \quad (13)$$

结果分析 航向 θ 角应在 60° 和 75° 之间(具体数值取决于 s_1 和 s_2 的比值),帆的朝向 α 角为 θ 的一半,这是从 A 点出发时的航向及帆的朝向. 行驶中 B 点将不在船的正东方,上述结论不再成立,所以应该不断调整 θ 和 α ,才能尽快地到达 B 点. 有兴趣的读者可以考虑 B 点不在船的正东方时如何确定航向及帆的朝向.

2.9 天气预报的评价

问题背景 明天是否下雨的天气预报常以有雨概率的形式给出,假如第一种预报方法告诉明天有雨概率是 80%,第二种预报方法告诉明天有雨概率是 60%,要是明天果真下雨了,能说第一种预报方法一定比第二种好吗? 如果好,好多少? 要是明天没有下雨呢?

判断预报方法的优劣不能根据一次预报与实际观测的符合程度下结论,假定得到了某地一个月 4 种预报方法的有雨概率预报,和实际上有雨或无雨的观测结果,见表 1,怎样根据这些数据对这 4 种预报方法给以评价呢?

表 1 31 天 4 种预报方法的有雨概率预报及实际观测结果

日期	预报 A/%	预报 B/%	预报 C/%	预报 D/%	实测(有雨=1,无雨=0)
1	90	30	90	60	1
2	40	30	50	80	1
3	60	30	80	70	1
4	60	30	90	70	1
5	60	30	0	20	0

续表

日期	预报 A/%	预报 B/%	预报 C/%	预报 D/%	实测(有雨=1,无雨=0)
6	30	30	10	50	1
7	80	30	10	40	0
8	70	30	20	30	0
9	80	30	40	30	0
10	60	30	60	40	0
11	80	30	20	80	1
12	40	30	30	40	0
13	90	30	90	40	1
14	50	30	60	20	0
15	10	30	20	10	0
16	60	30	50	80	1
17	20	30	10	30	0
18	0	30	0	50	0
19	90	30	60	40	0
20	70	30	10	0	0
21	20	30	0	30	0
22	40	30	20	30	0
23	40	30	10	10	0
24	80	30	50	40	0
25	30	30	0	20	0
26	30	30	10	30	0
27	30	30	20	0	0
28	0	30	60	40	1
29	60	30	0	20	0
30	20	30	10	10	0
31	80	30	50	10	0

下面给出三种评价这些预报方法的模型^[69]。

计数模型 如果我们听到有雨概率大于 50% 的预报,就认为明天有雨,听

到有雨概率小于 50% 的预报,就认为明天无雨,并且依照明天是否有雨的实际观测,规定预报是否正确,从而可以统计得到预报的正确率。

用这种办法按照实测有雨和无雨、预报有雨和无雨,分成 4 种情况,对表 1 中 4 种预报的结果计数(天数),得到图 1. 对于有雨概率等于 50% 的预报,可以认为是毫无意义的,这里不予统计。

预报A			预报B		
实测 预报	有雨	无雨	实测 预报	有雨	无雨
	有雨	6 10		有雨	0 0
无雨	3 11		无雨	9 22	

预报C			预报D		
实测 预报	有雨	无雨	实测 预报	有雨	无雨
	有雨	5 3		有雨	6 0
无雨	2 17		无雨	3 21	

图 1 4 种预报的结果计数(天数)

将图 1 中每种预报两个对角数字之和,除以总的预报天数(全部 4 个数字之和),得到 4 种预报的正确率依次为 0.57, 0.71, 0.81, 0.93. 它们可以看作正确预报的概率,是评价预报的一种指标,可是不难看出,预报 B 虽然有高达 0.71 的正确率,但这完全是由该地实际上有雨、无雨的天数决定的,实际上预报 B 毫无用途。

从实用角度看,对预报的使用者来说,更重要的是以下的条件概率:在预报无雨的条件下实测有雨的概率,及在预报有雨的条件下实测无雨的概率。前一个事件出现,可能由于预防不足而受灾导致损失;后一个事件出现,则会造成预防费用的浪费。这两个条件概率可以用图 1 中的数字来估计,如对预报 A,在预报无雨的条件下实测有雨的概率是 $3/(3+11)$ 。预报的使用者可以根据这两种后

果的轻重,将两个条件概率加权综合得到一个指标,不妨称为误报率.设两种后果的损失之比为 2:1,则可计算预报 A 的误报率为 $(2/3) \cdot (3/14) + (1/3) \cdot (10/16) = 0.35$, 预报 C, D 的误报率分别为 0.20, 0.06.

从上面计算的正确率和误报率看,都是预报 D 最好,预报 C 次之.

计数模型的缺点是明显的,它只在一定条件下区分预报正确与否,没有考虑预报有雨概率的具体数字,例如对于 90% 和 60% 的预报,不论实测有雨或无雨,结果都是一样的.

记分模型

将预报有雨概率的大小与实测结果(有雨或无雨)比较,给予记分.不同的记分规则形成不同的模型,如

模型 1 在实测有雨的情况,预报有雨概率大于 0.5 的得到相应的正分,预报有雨概率小于 0.5 的得到相应的负分;在实测无雨的情况,预报有雨概率小于 0.5 的得到相应的正分,预报有雨概率大于 0.5 的得到相应的负分.

具体地,记第 k 天某种预报有雨概率为 p_k ,第 k 天实测有雨为 $v_k = 1$,无雨为 $v_k = 0$,令第 k 天的某种预报得分为

$$s_k = \begin{cases} p_k - 0.5, & v_k = 1 \\ 0.5 - p_k, & v_k = 0 \end{cases}$$

将 s_k 对 k 求和得到某种预报在模型 1 下的分数,记作 S_1 , S_1 越大越好.由表 1 数据计算得到预报 A, B, C, D 的分数 S_1 分别为 1.0, 2.6, 7.0, 6.7, 预报 C 最好.

模型 2 直接用 p_k 和 v_k 之差定义分数为

$$s_k = |p_k - v_k|$$

将 s_k 对 k 求和得到某种预报在模型 2 下的分数,记作 S_2 , S_2 越小越好.由表 1 数据计算得到预报 A, B, C, D 的分数 S_2 分别为 14.5, 12.9, 8.5, 8.8, 预报 C 最好.

模型 3 用 p_k 和 v_k 之差的平方定义分数为

$$s_k = (p_k - v_k)^2$$

将 s_k 对 k 求和得到某种预报在模型 3 下的分数,记作 S_3 , S_3 越小越好.由表 1 数据计算得到预报 A, B, C, D 的分数 S_3 分别为 8.95, 6.39, 4.23, 3.21, 预报 D 最好.

分析这些模型及结果可以看出,模型 1, 2 对 4 种预报优劣的排序,甚至相对的分差都是相同的.实际上,读者不难证明,这两个模型是等价的.我们还注意到,模型 3 与模型 2 的结果有很大不同,那么究竟采用哪个模型较好呢?

略去下标 k , 用 p 和 v 分别表示预报的有雨概率和实测值(1 或 0), 用 f 表示理论上的有雨概率(其含义参见下面的讨论), 注意到 v 取 1 和 0 的概率分别为 f 和 $1-f$, 则模型 3 下的期望分数为

$$E(S) = E[(p - v)^2] = f(p - 1)^2 + (1 - f)p^2$$

经过简单的代数运算可得

$$E(S) = f(1 - f) + (p - f)^2$$

显然,当预报的有雨概率 p 等于理论上的有雨概率 f 时,模型 3 下的期望分数最小,并且 f 越接近 1 或 0,期望分数越小.

如果更一般地,考察模型 $S = (|p - v|)^n$,那么通过求期望分数 $E(S)$ 的极值可以发现,仅当 $n = 2$ 时,才有 p 等于 f 使得期望分数最小.在这个意义下模型 3 是最佳的.

图形模型

模型 1 以预报有雨概率 p 为横轴,实测值 v 为纵轴,将表 1 数据在图上用符号 * 标出,如图 2,其中 * 上面()中的数字是坐标在 * 的天数.

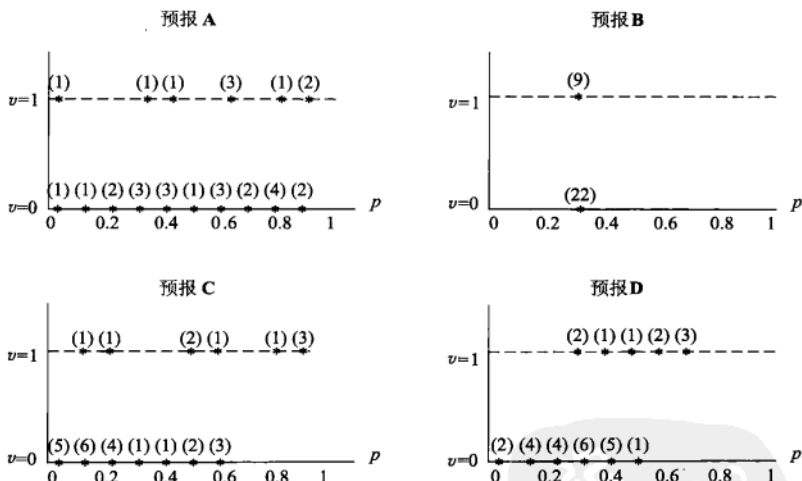


图 2 4 种预报的图形模型 1

从图 2 可以直观地看出,预报 A 的符号 * 几乎像是随机分布的,预报效果很差;预报 B 的符号 * 横坐标 p 没有变化,自然毫无用途;预报 C 中 $v=0$ 的符号 * 都在 $p=0.6$ 左边,表明对无雨的预报较好,但是 $v=1$ 的符号 * 相当分散,表明有雨预报较差;预报 D 中 $v=0$ 的符号 * 都在 $p=0.5$ 左边, $v=1$ 的符号 * 都在 $p=0.4$ 右边,表明对无雨、有雨预报都较好.

一个好的预报应该 $v=0$ 的符号 * 都在 $p=0.4$ 左边, $v=1$ 的符号 * 都在 $p=0.6$ 右边.最理想的当然是符号 * 都集中在 $(0,0)$ 和 $(1,1)$ 处.

模型 2 对于有雨概率 $p = 0.8$ 的预报的一种理解是,如果这种预报有 10

天,而其中实测有 8 天有雨,那么这个预报就是好的.按照这种想法,对每个不同的预报有雨概率 p ,统计实测有雨的天数占预报这个 p 的全部天数的比例,记作 q .显然, p 和 q 越接近越好.

以预报有雨概率 p 为横轴,实测有雨天数的比例 q 为纵轴,将表 1 数据进行统计后在图上用符号 * 标出,如图 3,图中的斜线是 $q=p$,好的预报的符号 * 应该分布在这条线附近.

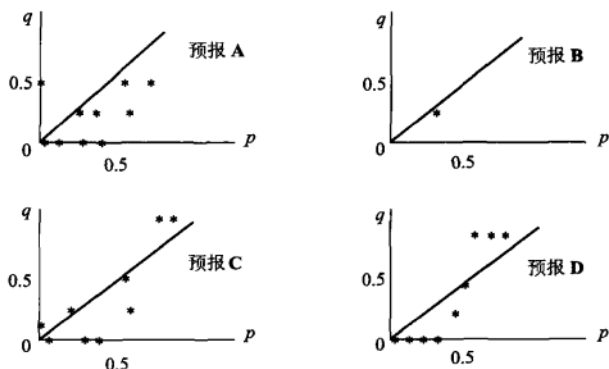


图 3 4 种预报的图形模型 2

从图 3 看出,预报 A 的 * 几乎均匀分布,明显不好;预报 B 只有一个 *, 并且几乎在 $q=p$ 上;预报 C, D 比 A 要好一些,但预报 D 并未显示出优势.从预报 B, D 的结果看,这种模型有一定缺陷,特别它不能用于预报 B 的情况.数据量较小可能是预报 D 未能得到正确评价的一个原因.

用图中符号 * 与直线 $q=p$ 的竖直距离(或距离的平方)度量模型的优劣,并考虑各个 * 的权重,图形模型 2 可以量化为分数模型,请有兴趣的读者一试.

更深入的讨论

评价一个预报的优劣,最重要的是制定评价标准,在这方面没有完全统一的想法,主要提出的有三类不同层次、不同内涵、但相互关联的标准^[58].

第一类标准是预报者本身的一致性,指预报者根据知识、信息和经验对待预报的事件作出的判断,与他对外发布的预报之间的关系.虽然看来要求预报者的判断和预报一致是合理的,但是也会有不完全一致的情况,如预报者没有利用全部判断作预报,而只用了从预报的使用者来说需要的那些信息,或者,出于预报效益等方面的考虑,预报者对判断作了适当改变给出预报.这类标准的一致性受预报者控制,而外界通常难以掌握.在预报以概率形式给出的情况下,当预报与预报者的判断一致时,才会得到与实际观测最相符的结果.这里预报者的判断就

是记分模型 3 中的概率 f .

第二类标准是根据预报和实际观测之间的关系,评价预报的品质.建立这种关系的最全面的方法是利用预报(记作随机变量 x)与观测(记作随机变量 y)的联合分布,记作 $F(x, y)$.从联合分布 $F(x, y)$ 可以得到各种条件分布和边际分布,而根据这些分布作计算,能够评价预报的品质有:

可靠性 指在特定预报 x 下观测 y 的条件均值与预报 x 的一致性,将二者之差对所有预报作平均,可作为衡量预报可靠性的数量指标,它越小越好,可由条件分布 $F(y|x)$ 和边际分布 $F(x)$ 计算得到.

决定性 指在特定预报 x 下观测 y 的条件均值与 y 的无条件均值的相异性,将二者之差对所有预报作平均,可作为衡量预报决定性的数量指标,它越大越好,也可由 $F(y|x)$ 和 $F(x)$ 计算得到.

分辨率 将在特定观测 y 下预报 x 的条件均值与观测 y 之差,对所有观测作平均,是衡量预报分辨率的一种数量指标,它越小越好;将这个条件均值与 y 的无条件均值之差,对所有观测作平均,是衡量预报分辨率的又一种数量指标,它越大越好.这两种指标都可由条件分布 $F(x|y)$ 和边际分布 $F(y)$ 得到.

敏锐性 衡量预报本身的敏锐,与实际事件无关,如在本节天气预报的情况下,若预报有雨概率多数接近 1 或 0,就是敏锐的.显然,敏锐性由边际分布 $F(x)$ 决定.

不确定性 指实际事件发生的不确定,与预报无关,但当事件发生的不确定性较高时,会给预报带来困难.不确定性由边际分布 $F(y)$ 决定.

前面的计数模型、记分模型、图形模型都是从某一个侧面反映第二类标准的.

第三类标准是看利用预报所实现的效益或带来的费用的大小,现在常用的是事前预测,用决策分析法估计预报的效益或费用的期望值,与不用预报(实际上是用先验估计)相比.这类标准虽然不能由预报者完全控制,但是与预报的品质,即第二类标准密切相关.在实际应用中,第二、三类标准间的关系是复杂的,例如前面记分模型 3 的分数与预报效益之间的函数关系可能不是单值的,从而使得二者的关系不一定是单调的^[57, 58].

这类标准在诸如谷物种植、耕种计划、水果保护等实际问题中有广泛应用.

习 题

1. 在 2.3 节中考虑八人艇分重量级组(桨手体重不超过 86 kg)和轻量级组(桨手体重不超过 73 kg),建立模型说明重量级组的成绩比轻量级组大约好 5%.

2. 用 2.4 节实物交换模型中介绍的无差别曲线的概念,讨论以下雇员和雇主之间的协议关系:

(1) 以雇员一天的工作时间 t 和工资 w 分别为横坐标和纵坐标,画出雇员无差别曲线族的示意图.解释曲线为什么是你画的那种形状.

(2) 如果雇主付计时工资,对不同的工资率(单位时间的工资)画出计时工资线族.根据雇员的无差别曲线族和雇主的计时工资线族,讨论双方将在怎样的一条曲线上达成协议.

(3) 雇员和雇主已经达成了一项协议(工作时间 t_1 和工资 w_1).如果雇主想使雇员的工作时间增加到 t_2 ,他有两种办法:一是提高计时工资率,在协议线的另一点 (t_2, w_2) 达成新的协议;二是实行超时工资制,即对工时 t_1 仍付原计时工资,对工时 $t_2 - t_1$ 付给更高的超时工资.试用作图方法分析哪种办法对雇主更有利,指出这个结果的条件^[10].

3. 根据 2.5 节中的流量数据(表 2)和(2)式作插值和数值积分,按照连续模型考虑均流池的容量(用到微积分的极值条件).

4. 在 2.6 节中从机理分析的角度建立车速与车流密度的对数模型 $v = v_1 \ln(k_2/k)$.

5. 在 2.7 节核武器竞赛模型中,证明由(6)式表示的乙安全线 $y = f(x)$ 的性质.

6. 在 2.7 节核武器竞赛模型中,讨论以下因素引起的平衡点的变化^[32]:

(1) 甲方提高导弹导航系统的性能.

(2) 甲方增加导弹爆破的威力.

(3) 甲方发展电子干扰系统.

(4) 双方建立反导弹系统.

7. 在超市购物时你注意到大包装商品比小包装商品便宜这种现象了吗?比如洁银牙膏 50 g 装的每支 1.50 元,120 g 装的每支 3.00 元,二者单位质量的价格比是 1.2:1.试用比例方法构造模型解释这个现象.

(1) 分析商品价格 C 与商品质量 w 的关系.价格由生产成本、包装成本和其他成本等决定,这些成本中有的与质量 w 成正比,有的与表面积成正比,还有与 w 无关的因素.

(2) 给出单位质量价格 c 与 w 的关系,画出它的简图,说明 w 越大 c 越小,但是随着 w 的增加 c 减小的程度变小.解释实际意义是什么.

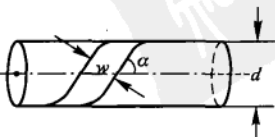
*8. 一垂钓俱乐部鼓励垂钓者将钓上的鱼放生,打算按照放生的鱼的质量给予奖励,俱乐部只准备了一把软尺用于测量,请你设计按照测量的长度估计鱼的质量的方法.假定鱼池中只有一种鲈鱼,并且得到 8 条鱼的如下数据(胸围指鱼身的最大周长):

身长/cm	36.8	31.8	43.8	36.8	32.1	45.1	35.9	32.1
质量/g	765	482	1 162	737	482	1 389	652	454
胸围/cm	24.8	21.3	27.9	24.8	21.6	31.8	22.9	21.6

先用机理分析建立模型,再用数据确定参数^[32].

9. 用宽 w 的布条缠绕直径 d 的圆形管道,要求布条不重叠,问布条与管道轴线的夹角 α 应多大(如图).若知道管道长度,需用多长布条(可考虑两端的影响)?如果管道是其他形状呢^[26]?

*10. 用已知尺寸的矩形板材加工半径一定的圆盘,给出几种简便、有效的排列方法,使加工出尽可能多的圆盘^[26].



11. 雨滴匀速下降,空气阻力与雨滴表面积和速度平方的乘积成正比,试确定雨速与雨滴质量的关系.

* 12. 动物园里的成年热血动物靠饲养的食物维持体温基本不变,在一些合理、简化的假设下建立动物的饲养食物量与动物的某个尺寸之间的关系^[32].

* 13. 生物学家认为,对于休息状态的热血动物,消耗的能量主要用于维持体温,能量与从心脏到全身的血流量成正比,而体温主要通过身体表面散失,建立一个动物体重(单位:g)与心率(单位:次/min)之间关系的模型,并用下面的数据加以检验^[32].

动物	体重	心率
田鼠	25	670
家鼠	200	420
兔	2 000	205
小狗	5 000	120
大狗	30 000	85
羊	50 000	70
人	70 000	72
马	450 000	38

* 14. 举重比赛按照运动员的体重分组,你能在一些合理、简化的假设下建立比赛成绩与体重之间的关系吗?下面是一届奥运会的竞赛成绩,可供检验你的模型^[32].

组别	最大体重/kg	抓举/kg	挺举/kg	总成绩/kg
1	54	132.5	155	287.5
2	59	137.5	170	307.5
3	64	147.5	187.5	335
4	70	162.5	195	357.5
5	76	167.5	200	367.5
6	83	180	212.5	392.5
7	91	187.5	213	402.5
8	99	185	235	420
9	108	195	235	430
10	> 108	197.5	260	457.5

第3章

简单的优化模型

优化问题可以说是人们在工程技术、经济管理和科学研究等领域中最常遇到的一类问题。设计师要在满足强度要求等条件下选择材料的尺寸,使结构总质量最轻;公司经理要根据生产成本和市场需求确定产品价格,使所获利润最高;调度人员要在满足物资需求和装载条件下安排从各供应点到各需求点的运量和路线,使运输总费用最低;投资者要选择一些股票、债券“下注”,使收益最大,而风险最小。

有些人习惯于依赖过去的经验解决面临的优化问题,认为这样切实可行,并且没有太大的风险。但是这种处理过程常常会融入决策者太多的主观因素,从而无法确认结果的最优性。也有些人习惯于作大量的试验反复比较,认为这样真实可靠。但是显然需要花费很多资金和人力,而且得到的最优结果基本上跑不出原来设计的试验范围。

第3,4两章讨论的是用数学建模的方法来处理优化问题,即建立和求解所谓优化模型。虽然由于建模时要作适当的简化,可能使得结果不一定完全可行或达到实际上的最优,但是它基于客观规律和数据,又不需要多大的费用。如果在建模的基础上再辅之以适当的经验和试验,就可以期望得到实际问题的一个比较圆满的回答。在决策科学化、量化的呼声日益高涨的今天,这无疑是符合时代潮流和形势发展需要的。

本章介绍较简单的优化模型,归结为微积分中的函数极值问题,可以直接用微分法求解。

当你打算用数学建模方法来处理一个优化问题的时候,首先要确定优化的目标是什么,寻求的决策是什么,决策受到哪些条件的限制(如果有限制的话),然后用数学工具(变量、常数、函数等)表示它们。当然,在这个过程中要对实际问题作若干合理的简化假设。最后,在用微分法求出最优决策后,要对结果作一些定性、定量的分析和必要的检验。让我们通过下面的实例说明建立优化模型的过程。

3.1 存贮模型

工厂定期订购原料,存入仓库供生产之用;车间一次加工出一批零件,供装配线每天生产之需;商店成批购进各种商品,放在货柜里以备零售;水库在雨季蓄水,用于旱季的灌溉和发电.显然,这些情况下都有一个贮存量多大才合适的问题.贮存量过大,贮存费用太高;贮存量太小,会导致一次性订购费用增加,或不能及时满足需求.

本节在需求量稳定的前提下讨论两个简单的存贮模型:不允许缺货模型和允许缺货模型.前者适用于一旦出现缺货会造成重大损失的情况(如炼铁厂对原料的需求),后者适用于像商店购货之类的情形,缺货造成的损失可以允许和估计^[73,76].

不允许缺货的存贮模型

先考察这样的问题:配件厂为装配线生产若干种部件,轮换生产不同的部件时因更换设备要付生产准备费(与生产数量无关),同一部件的产量大于需求时因积压资金、占用仓库要付贮存费.今已知某一部件的日需求量 100 件,生产准备费 5 000 元,贮存费每日每件 1 元.如果生产能力远大于需求,并且不允许出现缺货,试安排该产品的生产计划,即多少天生产一次(称为生产周期),每次产量多少,可使总费用最小.

问题分析 让我们试算一下:

若每天生产一次,每次 100 件,无贮存费,生产准备费 5 000 元,每天费用 5 000 元;

若 10 天生产一次,每次 1 000 件,贮存费 $900 + 800 + \cdots + 100 = 4\,500$ 元,生产准备费 5 000 元,总计 9 500 元,平均每天费用 950 元;

若 50 天生产一次,每次 5 000 件,贮存费 $4\,900 + 4\,800 + \cdots + 100 = 122\,500$ 元,生产准备费 5 000 元,总计 127 500 元,平均每天费用 2 550 元.

虽然从以上结果看,10 天生产一次比每天和 50 天生产一次的费用少,但是,要得到准确的结论,应该建立生产周期、产量与需求量、生产准备费、贮存费之间的关系,即数学建模.

从上面的计算看,生产周期短、产量少,会使贮存费小,准备费大;而周期长、产量多,会使贮存费大,准备费小.所以必然存在一个最佳的周期,使总费用最小.显然,应该建立一个优化模型.

一般地,考察这样的不允许缺货的存贮模型:产品需求稳定不变,生产准备费和产品贮存费为常数、生产能力无限、不允许缺货,确定生产周期和产量,使总费用最小.

模型假设 为了处理的方便,考虑连续模型,即设生产周期 T 和产量 Q 均为连续量. 根据问题性质作如下假设:

1. 产品每天的需求量为常数 r .
2. 每次生产准备费为 c_1 , 每天每件产品贮存费为 c_2 .
3. 生产能力为无限大(相对于需求量), 当贮存量降到零时, Q 件产品立即生产出来供给需求, 即不允许缺货.

模型建立 将贮存量表示为时间 t 的函数 $q(t)$, $t=0$ 生产 Q 件, 贮存量 $q(0)=Q$, $q(t)$ 以需求速率 r 递减, 直到 $q(T)=0$, 如图 1. 显然有

$$Q = rT \quad (1)$$

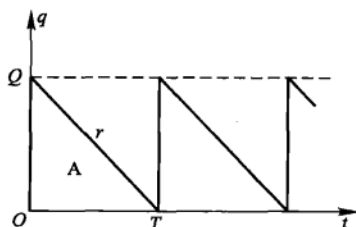


图 1 不允许缺货模型的贮存量 $q(t)$

一个周期内的贮存费是 $c_2 \int_0^T q(t) dt$, 其中积分恰等于图 1 中三角形 A 的面积 $QT/2$. 因为一个周期的准备费是 c_1 , 再注意到(1)式, 得到一个周期的总费用为

$$\bar{C} = c_1 + c_2 QT/2 = c_1 + c_2 r T^2 / 2 \quad (2)$$

于是每天的平均费用是

$$C(T) = \bar{C}/T = c_1/T + c_2 r T / 2 \quad (3)$$

(3) 式为这个优化模型的目标函数.

模型求解 求 T 使(3)式的 C 最小. 容易得到

$$T = \sqrt{\frac{2c_1}{c_2 r}} \quad (4)$$

代入(1)式可得

$$Q = \sqrt{\frac{2c_1 r}{c_2}} \quad (5)$$

由(3)式算出最小的总费用为

$$C = \sqrt{2c_1 c_2 r} \quad (6)$$

(4), (5) 式是经济学中著名的经济订货批量公式 (EOQ 公式).

结果解释 由(4),(5)式可以看到,当准备费 c_1 增加时,生产周期和产量都变大;当贮存费 c_2 增加时,生产周期和产量都变小;当需求量 r 增加时,生产周期变小而产量变大. 这些定性结果都是符合常识的. 当然,(4),(5)式的定量关系(如平方根、系数2等)凭常识是无法猜出的,只能由数学建模得到.

用得到的模型计算本节开始的问题:以 $c_1 = 5\,000$, $c_2 = 1$, $r = 100$ 代入(4),(6)式可得 $T = 10$ 天, $C = 1\,000$ 元. 这里得到的费用 C 与前面计算的950元有微小的差别,你能解释吗?

敏感性分析 讨论参数 c_1, c_2, r 有微小变化时对生产周期 T 的影响.

用相对改变量衡量结果对参数的敏感程度, T 对 c_1 的敏感度记作 $S(T, c_1)$, 定义为

$$S(T, c_1) = \frac{\Delta T/T}{\Delta c_1/c_1} \approx \frac{dT}{dc_1} \frac{c_1}{T} \quad (7)$$

由(4)式容易得到 $S(T, c_1) = 1/2$. 作类似的定义并可得到 $S(T, c_2) = -1/2$, $S(T, r) = -1/2$. 即 c_1 增加1%, T 增加0.5%, 而 c_2 或 r 增加1%, T 减少0.5%. c_1, c_2, r 的微小变化对生产周期 T 的影响是很小的.

思考

1. 建模中未考虑生产费用(这应是最大的一笔费用),在什么条件下才可以不考虑它(习题1)?

2. 建模时作了“生产能力为无限大”的简化假设,如果生产能力有限,是大于需求量的一个常数,如何建模(习题2)?

评注 EOQ 公式是近百年前得到的,至今仍是研究批量生产计划问题的理论基础之一,实际上也还有一些用处.

考察用户向供方订货的情况:如果订货时要付一笔订货费(与订货量无关),贮存费和用户需求量的假设与上面模型一样,并且当贮存量降到零时所订货物立即到达,那么只需将订货费类比为生产准备费,就会得到完全相同的模型. 实际上,EOQ 公式原本就是针对这种订货情况的.

允许缺货的存贮模型

在某些情况下用户允许短时间的缺货,虽然这会造成一定的损失,但是如果损失费不超过不允许缺货导致的准备费和贮存费的话,允许缺货就应该是可以采取的策略.

模型假设 下面讨论一种较简单的允许缺货模型:不允许缺货模型的假设1,2不变,假设3改为:

3a. 生产能力为无限大(相对于需求量),允许缺货,每天每件产品缺货损失费为 c_3 ,但缺货数量需在下次生产(或订货)时补足.

模型建立 因贮存量不足造成缺货时,可认为贮存量函数 $q(t)$ 为负值,如

图2. 周期仍记作 T , Q 是每周期的贮存量, 当 $t = T_1$ 时 $q(t) = 0$, 于是有

$$Q = rT_1 \quad (8)$$

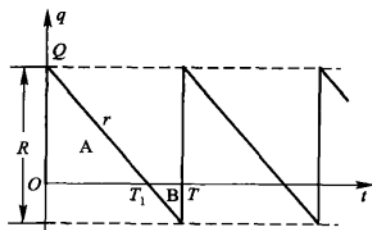


图2 允许缺货模型的贮存量 $q(t)$

在 T_1 到 T 这段缺货时段内需求率 r 不变, $q(t)$ 按原斜率继续下降. 由于规定缺货量需补足, 所以在 $t = T$ 时数量为 R 的产品立即到达, 使下周期的贮存量恢复到 Q .

与建立不允许缺货模型时类似, 一个周期内的贮存费是 c_2 乘以图2中三角形A的面积, 缺货损失费则是 c_3 乘以图2中三角形B的面积. 计算这两块面积, 并加上准备费 c_1 , 得到一个周期的总费用为

$$\bar{C} = c_1 + c_2 QT_1/2 + c_3 r(T - T_1)^2/2 \quad (9)$$

利用(8)式将模型的目标函数——每天的平均费用——记作 T 和 Q 的二元函数

$$C(T, Q) = \frac{c_1}{T} + \frac{c_2 Q^2}{2rT} + \frac{c_3 (rT - Q)^2}{2rT} \quad (10)$$

模型求解 利用微分法求 T 和 Q 使 $C(T, Q)$ 最小, 令 $\partial C/\partial T = 0$, $\partial C/\partial Q = 0$, 可得(为了与不允许缺货模型相区别, 最优解记作 T' , Q')

$$T' = \sqrt{\frac{2c_1}{c_2 r} \frac{c_2 + c_3}{c_3}}, \quad Q' = \sqrt{\frac{2c_1 r}{c_2} \frac{c_3}{c_2 + c_3}} \quad (11)$$

注意到每个周期的供货量 $R = rT'$, 有

$$R = \sqrt{\frac{2c_1 r}{c_2} \frac{c_2 + c_3}{c_3}} \quad (12)$$

记

$$\lambda = \sqrt{\frac{c_2 + c_3}{c_3}} \quad (13)$$

与不允许缺货模型的结果(4), (5)式比较不难得到

$$T' = \lambda T, \quad Q' = Q/\lambda, \quad R = \lambda Q \quad (14)$$

结果解释 由(13)式, $\lambda > 1$, 故(14)式给出 $T' > T, Q' < Q, R > Q$, 即允许缺货时周期及供货量应增加, 周期初的贮存量减少. 缺货损失费 c_3 越大(相对于贮存费 c_2), λ 越小, T' 越接近 T, Q', R 越接近 Q . 当 $c_3 \rightarrow \infty$ 时 $\lambda \rightarrow 1$, 于是 $T' \rightarrow T, Q' \rightarrow Q, R \rightarrow Q$. 这个结果合理吗(考虑 $c_3 \rightarrow \infty$ 的意义)? 由此不允许缺货模型可视为允许缺货模型的特例.

3.2 生猪的出售时机

一饲养场每天投入 4 元资金用于饲料、设备、人力, 估计可使一头 80 kg 重的生猪每天增加 2 kg. 目前生猪出售的市场价格为 8 元/kg, 但是预测每天会降低 0.1 元, 问该场应该什么时候出售这样的生猪. 如果上面的估计和预测有出入, 对结果有多大影响^[53]?

问题分析 投入资金可使生猪体重随时间增长, 但售价(单价)随时间减少, 应该存在一个最佳的出售时机, 使获得利润最大. 这是一个优化问题, 根据给出的条件, 可作如下的简化假设.

模型假设 每天投入 4 元资金使生猪体重每天增加常数 r ($= 2$ kg); 生猪出售的市场价格每天降低常数 g ($= 0.1$ 元).

模型建立 给出以下记号: t ~ 时间(天); w ~ 生猪体重(kg); p ~ 单价(元/kg); R ~ 出售的收入(元); C ~ t 天投入的资金(元); Q ~ 纯利润(元).

按照假设, $w = 80 + rt$ ($r = 2$), $p = 8 - gt$ ($g = 0.1$). 又知道 $R = pw, C = 4t$, 再考虑到纯利润应扣掉以当前价格(8 元/kg)出售 80 kg 生猪的收入, 有 $Q = R - C - 8 \times 80$, 得到目标函数(纯利润)为

$$Q(t) = (8 - gt)(80 + rt) - 4t - 640 \quad (1)$$

其中 $r = 2, g = 0.1$. 求 $t(\geq 0)$ 使 $Q(t)$ 最大.

模型求解 这是求二次函数最大值问题, 用代数或微分法容易得到

$$t = \frac{4r - 40g - 2}{rg} \quad (2)$$

当 $r = 2, g = 0.1$ 时, $t = 10, Q(10) = 20$, 即 10 天后出售, 可得最大纯利润 20 元.

敏感性分析 由于模型假设中的参数(生猪每天体重的增加 r 和价格的降低 g)是估计和预测的, 所以应该研究它们有所变化时对模型结果的影响.

1. 设每天生猪价格的降低 $g = 0.1$ 元不变, 研究 r 变化的影响, 由(2)式可得

$$t = \frac{40r - 60}{r}, \quad r \geq 1.5 \quad (3)$$

t 是 r 的增函数, 表 1 和图 1 给出它们的关系.

表 1 r 与 t 的关系

r	1.5	1.6	1.7	1.8	1.9	2.0	2.1	2.2
t	0	2.5	4.7	6.7	8.4	10.0	11.4	12.7
r	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
t	13.9	15.0	16.0	16.9	17.8	18.6	19.3	20.0

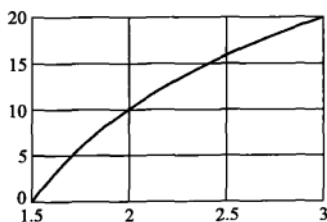
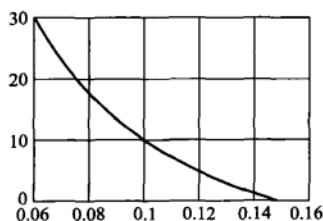
2. 设每天生猪体重的增加 $r=2$ kg 不变, 研究 g 变化的影响, 由(2)式可得

$$t = \frac{3-20g}{g}, \quad 0 \leq g \leq 0.15 \quad (4)$$

t 是 g 的减函数, 表 2 和图 2 给出它们的关系。

表 2 g 与 t 的关系

g	0.06	0.07	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15
t	30.0	22.9	17.5	13.3	10.0	7.3	5.0	3.1	1.4	0

图 1 r 与 t 的关系图 2 g 与 t 的关系

可以用相对改变量衡量结果对参数的敏感程度. t 对 r 的敏感度记作 $S(t, r)$, 定义为

$$S(t, r) = \frac{\Delta t/t}{\Delta r/r} \approx \frac{dt}{dr} \frac{r}{t} \quad (5)$$

由(3)式, 当 $r=2$ 时可算出

$$S(t, r) \approx \frac{60}{40r - 60} = 3 \quad (6)$$

即生猪每天的体重 r 增加 1%, 出售时间推迟 3%。

类似地定义 t 对 g 的敏感度 $S(t, g)$, 由(4)式, 当 $g=0.1$ 时可算出

$$S(t, g) = \frac{\Delta t/t}{\Delta g/g} \approx \frac{dt}{dg} \frac{g}{t} = -\frac{3}{3-20g} = -3 \quad (7)$$

即生猪价格每天的降低 g 增加 1%, 出售时间提前 3%. r 和 g 的微小变化对模型结果的影响并不算大.

强健性 (Robustness) 分析 建模过程中假设了生猪体重的增加和价格的降低都是常数, 由此得到的 w 和 p 都是线性函数, 这无疑是对现实情况的简化. 更实际的模型应考虑非线性和不确定性, 如记 $w = w(t)$, $p = p(t)$, 则 (1) 式应为

$$Q(t) = p(t)w(t) - 4t - 640 \quad (8)$$

用微分法求解 (8) 式的极值问题, 可知最优解应满足

$$p'(t)w(t) + p(t)w'(t) = 4 \quad (9)$$

(9) 式左端是每天收入的增值, 右端是每天投入的资金. 于是出售的最佳时机是保留生猪直到每天收入的增值等于每天投入的资金为止. 本例中 $p' = -0.1$, $w' = 2$ 是根据估计和预测确定的, 只要它们的变化不大, 上述结论就是可用的.

另外, 从敏感性分析知, $S(t, r) = 3$, 所以若 $1.8 \leq w' \leq 2.2$ (10% 以内), 则结果应为 $7 \leq t \leq 13$ (30% 以内). 若设 $p' = -0.1$ 是最坏的情况, 如果这个 (绝对) 值更小, t 就应更大. 所以最好的办法是: 过大约一周后重新估计 p, p', w, w' , 再作计算.

评注 这个问题本身及其建模过程都非常简单, 我们着重介绍的是它的敏感性分析和强健性分析, 这种分析对于一个模型, 特别是优化模型是否真的能用, 或者用的效果如何, 是很重要的.

3.3 森林救火

森林失火了! 消防站接到报警后派多少消防队员前去救火呢? 派的队员越多, 森林的损失越小, 但是救援的开支会越大, 所以需要综合考虑森林损失费和救援费与消防队员人数之间的关系, 以总费用最小来决定派出队员的数目^[10].

问题分析 损失费通常正比于森林烧毁的面积, 而烧毁面积与失火、灭火 (指火被扑灭) 的时间有关, 灭火时间又取决于消防队员数目, 队员越多灭火越快. 救援费既与消防队员人数有关, 又与灭火时间长短有关. 记失火时刻为 $t = 0$, 开始救火时刻为 $t = t_1$, 灭火时刻为 $t = t_2$. 设在时刻 t 森林烧毁面积为 $B(t)$, 则造成损失的森林烧毁面积为 $B(t_2)$. 建模要对函数 $B(t)$ 的形式作出合理的简单假设.

研究 $\frac{dB}{dt}$ 比 $B(t)$ 更为直接和方便. $\frac{dB}{dt}$ 是单位时间烧毁面积, 表示火势蔓延的程度. 在消防队员到达之前, 即 $0 \leq t \leq t_1$, 火势越来越大, 即 $\frac{dB}{dt}$ 随 t 的增加而增加; 开始救火以后, 即 $t_1 < t < t_2$, 如果消防队员救火能力足够强, 火势会越来越小, 即 $\frac{dB}{dt}$ 应减小, 并且当 $t = t_2$ 时 $\frac{dB}{dt} = 0$.

救援费可分为两部分:一部分是灭火器材的消耗及消防队员的薪金等,与队员人数及灭火所用的时间均有关;另一部分是运送队员和器材等的一次性支出,只与队员人数有关。

模型假设 需要对烧毁森林的损失费、救援费及火势蔓延程度 $\frac{dB}{dt}$ 的形式作出假设。

1. 损失费与森林烧毁面积 $B(t_2)$ 成正比,比例系数 c_1 为烧毁单位面积的损失费。

2. 从失火到开始救火这段时间 ($0 \leq t \leq t_1$) 内,火势蔓延程度 $\frac{dB}{dt}$ 与时间 t 成正比,比例系数 β 称火势蔓延速度。

3. 派出消防队员 x 名,开始救火以后 ($t > t_1$) 火势蔓延速度降为 $\beta - \lambda x$,其中 λ 可视为每个队员的平均灭火速度。显然应有 $\beta < \lambda x$ 。

4. 每个消防队员单位时间的费用为 c_2 ,于是每个队员的救火费用是 $c_2(t_2 - t_1)$;每个队员的一次性支出是 c_3 。

第2条假设可作如下解释:火势以失火点为中心,以均匀速度向四周呈圆形蔓延,所以蔓延的半径 r 与时间 t 成正比。又因为烧毁面积 B 与 r^2 成正比,故 B 与 t^2 成正比,从而 $\frac{dB}{dt}$ 与 t 成正比。这个假设在风力不大的条件下是大致合理的。

模型构成 根据假设条件2,3,火势蔓延程度 $\frac{dB}{dt}$ 在 $0 \leq t \leq t_1$ 线性地增加,在 $t_1 < t \leq t_2$ 线性地减小。 $\frac{dB}{dt} \sim t$ 的图形如图1所示。记 $t = t_1$ 时 $\frac{dB}{dt} = b$ 。烧毁面积 $B(t_2) = \int_0^{t_2} \frac{dB}{dt} dt$ 恰是图中三角形的面积,显然有 $B(t_2) = \frac{1}{2} b t_2$,而 t_2 满足

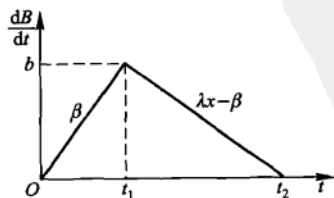


图1 $\frac{dB}{dt} \sim t$ 关系

$$t_2 - t_1 = \frac{b}{\lambda x - \beta} = \frac{\beta t_1}{\lambda x - \beta} \quad (1)$$

于是

$$B(t_2) = \frac{\beta t_1^2}{2} + \frac{\beta^2 t_1^2}{2(\lambda x - \beta)} \quad (2)$$

根据假设条件 1,4, 森林损失费为 $c_1 B(t_2)$, 救援费为 $c_2 x(t_2 - t_1) + c_3 x$. 将 (1), (2) 代入, 得到救火总费用为

$$C(x) = \frac{c_1 \beta t_1^2}{2} + \frac{c_1 \beta^2 t_1^2}{2(\lambda x - \beta)} + \frac{c_2 \beta t_1 x}{\lambda x - \beta} + c_3 x \quad (3)$$

$C(x)$ 即为这个优化模型的目标函数.

模型求解 为求 x 使 $C(x)$ 达到最小, 令 $\frac{dC}{dx} = 0$, 可以得到应派出的队员人数为

$$x = \frac{\beta}{\lambda} + \beta \sqrt{\frac{c_1 \lambda t_1^2 + 2c_2 t_1}{2c_3 \lambda^2}} \quad (4)$$

结果解释 首先, 应派出队员数目由两部分组成, 其中一部分 β/λ 是为了把火扑灭所必需的最少队员数. 因为 β 是火势蔓延速度, 而 λ 是每个队员的平均灭火速度, 所以这个结果是明显的. 从图 1 也可以看出, 只有当 $x > \beta/\lambda$ 时, 斜率为 $\lambda x - \beta$ 的直线才会与 t 轴有交点 t_2 .

其次, 派出队员数的另一部分, 即在最低限度之上的队员数, 与问题的各个参数有关. 当队员灭火速度 λ 和救援费用系数 c_3 增大时, 队员数减少; 当火势蔓延速度 β 、开始救火时刻 t_1 及损失费用系数 c_1 增加时, 队员数增加. 这些结果与常识是一致的. (4) 式还表明, 当救援费用系数 c_2 变大时队员数也增加, 请读者考虑为什么会有这样的结果.

实际应用这个模型时, c_1, c_2, c_3 是已知常数, β, λ 由森林类型、消防队员素质等因素决定, 可以预先制成表格以备查用. 由失火到救火的时间 t_1 则要根据现场情况估计.

评注 建立这个模型的关键是对 $\frac{dB}{dt}$ 的假设. 比较合理而又简化的假设条件 2,3 只能符合风力不大的情况. 在风势的影响下应考虑另外的假设. 再者, 有人对队员灭火的平均速度 λ 是常数的假设提出异议, 认为 λ 应与开始救火的时刻 t_1 有关, t_1 越大 λ 越小. 这时要对函数 $\lambda(t_1)$ 作出合理的假设, 再得到进一步的结果.

3.4 消费者的选择

作为一个消费者, 站在琳琅满目的市场里, 如何分配手里一定数量的钱, 选

择购买若干种需要的商品。“消费者追求最大效用”是经济学最优化原理中的一条,可以根据这条原理,用数学建模的方法帮助消费者决定他在市场里的选择^[6]。

为简单起见,假定只有甲乙两种商品供消费者购买,下面建立的模型可以推广到任意多种商品的情况。

效用函数 当消费者购得数量分别为 x_1, x_2 的甲乙两种商品时,给消费者带来的效用可以用一个数值来度量(本节最后有对“效用”一词的解释),它是 x_1, x_2 的函数,记作 $u(x_1, x_2)$,称为效用函数(Utility Function)。为了用效用函数讨论消费者的选择,利用等高线的概念在 x_1, x_2 平面上画出效用函数 $u(x_1, x_2)$ 的等值线,称为等效用线,如图 1。实际上,等效用线就是 2.4 节实物交换中的无差别曲线,效用就是那里的满意度。根据 2.4 节的分析,等效用线 $u(x_1, x_2) = c$ 是一族单调减、下凸、互不相交的曲线,随着效用值 c 的增加曲线向右上方移动,曲线的具体形状由甲乙两种商品对消费者带来的效用,或者消费者对甲乙两种商品的偏爱程度决定,后面将给出几个常用的效用函数。现在假定对于消费者来说,效用函数 $u(x_1, x_2)$ 已经确定了。

效用最大化模型 设甲乙两种商品的单价分别为 p_1, p_2 , 消费者准备付出的钱为 y , 则他购得的甲乙两种商品的数量 x_1, x_2 满足

$$p_1 x_1 + p_2 x_2 = y \quad (1)$$

所谓效用最大化,就是在(1)式条件下确定购买甲乙两种商品的数量 x_1, x_2 , 使效用函数 $u(x_1, x_2)$ 达到最大。

先从几何上分析、求解这个模型,如图 2。条件(1)在 x_1, x_2 平面上是一条直线,称消费线,如图中的 AB 。根据等效用线族 $u(x_1, x_2) = c$ 具有单调减、下凸、互不相交的性质,直线 AB 必与一条等效用线(图 2 中的 l)相切,记切点的 $Q(x_1, x_2)$, 称消费点,当购买甲乙两种商品的数量为 x_1, x_2 时效用函数达到最大。因为在 AB 与其他等效用线的交点,其效用函数值小于 Q 点的 $u(x_1, x_2)$, 且 Q 点是惟一的。

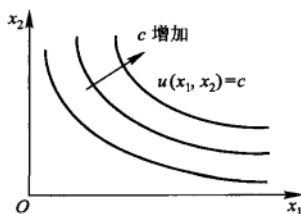


图 1 等效用线示意图

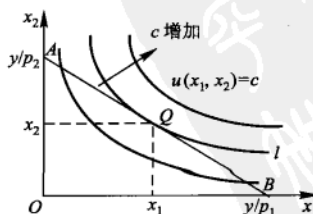


图 2 效用最大化模型的几何解法

如果知道了效用函数 $u(x_1, x_2)$ 的解析表达式, 那么可以按照二元函数的条件极值求解上述问题. 引入拉格朗日乘子 λ 构造函数 $L(x_1, x_2, \lambda) = u(x_1, x_2) + \lambda(y - p_1x_1 - p_2x_2)$, 由 $\frac{\partial L}{\partial x_1} = 0, \frac{\partial L}{\partial x_2} = 0$, 可得最优解 (x_1, x_2) 满足

$$\frac{\frac{\partial u}{\partial x_1}}{\frac{\partial u}{\partial x_2}} \bigg|_{x_1, x_2} = \frac{p_1}{p_2} \quad (2)$$

容易证明几何解法得到的 $Q(x_1, x_2)$ 点与 (2) 的结果是一致的. 因为消费线 AB 的斜率是 $-p_1/p_2$, 而等效用线 $u(x_1, x_2) = c$ 的斜率是 $\frac{dx_2}{dx_1} = -\frac{\partial u / \partial x_1}{\partial u / \partial x_2}$.

数学中的“导数”在经济学中一般称为“边际”, 于是 $\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}$ 称为**边际效用**, 即商品数量增加一个单位时效用函数的增量. (2) 式表明, 当商品的边际效用之比等于它们的价格之比时, 效用函数达到最大. 这是数量经济学的一个重要结论.

效用函数的构造 从上述讨论可以看出, 要对效用最大化模型作定量的分析, 需要有效用函数的解析表达式. 对于等效用线族 $u(x_1, x_2) = c$ 必须具有的性质: 单调减、下凸、互不相交, 可以给出一个便于构造和检验的充分条件:

$$\frac{\partial u}{\partial x_1} > 0, \frac{\partial u}{\partial x_2} > 0, \frac{\partial^2 u}{\partial x_1^2} < 0, \frac{\partial^2 u}{\partial x_2^2} < 0, \frac{\partial^2 u}{\partial x_1 \partial x_2} > 0 \quad (3)$$

请读者解释 (3) 中几个导数正负号的实际意义.

下面列出几个常用的效用函数, 并从效用最大化的角度分析它们的特性.

1. 设效用函数为

$$u(x_1, x_2) = \left(\frac{\alpha}{x_1} + \frac{\beta}{x_2} \right)^{-1}, \alpha, \beta > 0 \quad (4)$$

由效用最大化模型 (2) 可得最优值 x_1, x_2 满足

$$\frac{p_1 x_1}{p_2 x_2} = \eta \sqrt{\frac{p_1}{p_2}}, \eta = \sqrt{\frac{\alpha}{\beta}} \quad (5)$$

即按照效用最大化购买两种商品所用钱的比例, 与商品价格比的平方根成正比, 比例系数是参数 α 与 β 之比的平方根, 效用函数 (4) 中 α, β 分别度量甲乙两种商品对消费者的效用, 或者消费者对甲乙两种商品的偏爱.

2. 设效用函数为

$$u(x_1, x_2) = x_1^\lambda x_2^\mu, 0 < \lambda, \mu < 1 \quad (6)$$

由效用最大化模型 (2) 可得最优值 x_1, x_2 满足

$$\frac{p_1 x_1}{p_2 x_2} = \frac{\lambda}{\mu} \quad (7)$$

即按照效用最大化购买两种商品所用钱的比例,只取决于参数 λ 与 μ 之比,与商品价格无关,效用函数(6)中 λ, μ 分别度量甲乙两种商品对消费者的效用,或者消费者对甲乙两种商品的偏爱。

3. 设效用函数为

$$u(x_1, x_2) = (a\sqrt{x_1} + b\sqrt{x_2})^2, a, b > 0 \quad (8)$$

对(8)的求解和讨论留作习题4。

实际应用时可以根据上面最优解的分析,决定采用哪种效用函数,并由经验数据确定其参数。

效用最大化模型的应用举例^[87]

例1 征销售税还是征收入税

政府从消费者身上征税一般有两种办法,一种是征销售税,即根据消费者在购买若干种商品时所花钱的数额征税,另一种是征收入税,即根据消费者的收入征收所得税。下面利用图形从效用函数和效用最大化的角度讨论这两种办法。

征税前仍设甲乙两种商品的单价为 p_1, p_2 ,消费者从收入中准备花的钱为 y ,等效用线为 $u(x_1, x_2) = c$,消费点为 $Q(x_1, x_2)$,如图2。如果政府只对甲商品征销售税,且税率为 p_0 ,则相当于迫使甲商品的单价由 p_1 增至 $p_1 + p_0$,于是原来的条件(1)变为

$$(p_1 + p_0)x_1 + p_2x_2 = y \quad (9)$$

消费线 AB 变为图3中的 AB_1 , B_1 显然在 B 的左边, AB_1 与等效用线 l 下边的 l_1 相切于 $Q_1(x_1^*, x_2^*)$ 点,消费者购买的商品数量为 x_1^*, x_2^* ,政府得到的销售税额为 $p_0x_1^*$ 。

若政府将征销售税改为征收入税,从消费者收入 y 中征收的税额与销售税额 $p_0x_1^*$ 相同,于是商品价格不变,而原来的条件(1)变为

$$p_1x_1 + p_2x_2 = y - p_0x_1^* \quad (10)$$

消费线 AB 变为图3中 AB 下方的一条平行线 A_2B_2 ,可以证明 B_2 在 B_1 的右边,但是与 A_2B_2 相切的等效用线 l_2 是在 l_1 的上方(如图3),还是在 l_1 的下方(图上没有画出来),仅从图形上难以分辨,需要从效用函数的解析表达式等作进一步分析。而如果 l_2 在 l_1 上方,切点 Q_2 的效用函数值将大于 Q_1 ,那么对消费者来说征收入税比征销售税好。

例2 价格补贴给生产者还是消费者

价格补贴是许多国家的政府为了鼓励某些商品的生产,或者减少消费者的负担所采取的一项政策,通常有两种办法,一种是政府把补贴款直接给生产者,另一种是政府把补贴款发给消费者,而让商品涨价。仍然利用图形从效用函数和效用最大化的角度讨论这两种办法。

如果政府把补贴款直接给某些商品的生产者,自然会鼓励这些商品的生产,而对消费者没有影响,于是仍按图2分析的那样,在条件(1)下效用函数值达到最大的消费点,这里记作 $Q(x_1^*, x_2^*)$, 如图4.

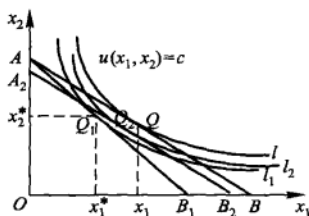


图3 征销售税还是征收入税的讨论

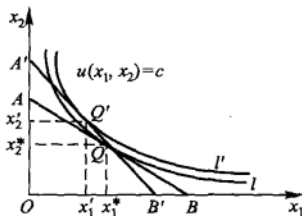


图4 价格补贴给生产者还是消费者的讨论

若政府让甲商品的单价由 p_1 涨到 $p_1 + p_0$ (乙商品不变), 把消费者购买甲商品多花的钱 $p_0 x_1^*$ 补贴给消费者, 使之仍然能够达到原来的消费点 $Q(x_1^*, x_2^*)$, 则条件(1)将变为

$$(p_1 + p_0)x_1 + p_2 x_2 = y + p_0 x_1^* \quad (11)$$

消费线 AB 变为图4中的直线 $A'B'$, $A'B'$ 过 $Q(x_1^*, x_2^*)$ 点, 且容易证明 A' 在 A 的上边, B' 在 B 的左边. 按照效用最大化模型, 与 $A'B'$ 相切的等效曲线 l' 将在 l 的上方, 于是消费点 $Q'(x_1', x_2')$ 的效用函数值将大于 $Q(x_1^*, x_2^*)$, 而甲乙商品的购买量 $x_1' < x_1^*$, $x_2' > x_2^*$. 这个结果表明, 政府的这种补贴办法对消费者更有利(效用函数值增加), 而涨价的甲商品(如猪肉)购买量减少, 不涨价的乙商品(甲的替代品如鸡、鱼)购买量增加, 显然不利于甲商品的生产者.

评注 起初一些经济学家把“效用”解释为人们对商品的偏爱程度, 是主观的、感性的东西, 后来人们把它与对福利的贡献、满足经济目标的物品的性质、商品的“效率”等联系起来, 力图使“效用”从心理学的概念转化为合理的经济行为. 现在消费效用最大化原理已经被广泛接受, 认为它在一定程度上刻画了人们消费的合理性. 但是对于效用是否是一个数值函数仍有疑问. 一些人主张用“偏爱”代替“效用”, 偏爱只有顺序的先后, 没有数值大小的区分, 从数学概念上说, 就是用“序数效用”代替“基数效用”(如在英语语法中序数词是 first, second, ..., 基数词是 one, two, ...), 而这样一来, 数学上就要重起炉灶, 用到如实变函数论、拓扑学等工具了.

对于征销售税还是征收入税的讨论, 只能作为应用效用函数和效用最大化的一个练习, 在现实生活中, 其结论的合理性很大程度上取决于所涉及的是什么商品(生活必需品还是奢侈品), 以及政府要照顾哪些人的利益. 至于价格补贴

问题,我国近年来有几次把副食品的价格补贴直接发给消费者(主要是生活困难群体),按照上面的分析,这样做虽然会减少涨价商品的消费,但总体上对消费者是有利的。可是实际问题是复杂的,比如让猪肉涨价而把补贴一视同仁地发给消费者时,爱吃肉的觉得补贴太少,不吃肉的(吃替代品如鸡、鱼)则白得了一笔钱。另外,猪肉和替代品的生产者更会因为这种补贴办法导致的销售量减少和增加,而得到不公平的待遇。解决这些问题远不是用上面简单的数学模型能完成的。

3.5 生产者的决策

作为一个生产者或供销商,在市场经济中如何根据产品的成本和产值决定他的投入,又怎样按照商品的销售情况确定价格。可以与3.4中“消费者追求最大效用”相提并论,“生产者追求最大利润”是经济学的又一条最优化原理。下面根据这条原理,用数学建模的方法帮助生产者或供销商作出决策。

最大利润模型 众所周知,生产者的利润等于产品的产值减去成本,当然这里假定产品可以全部销售出去变成收入。记生产者对产品的投入量为 x ,产值和成本都是 x 的函数,分别记作 $f(x)$ 和 $c(x)$,则利润 $r(x)$ 为

$$r(x) = f(x) - c(x) \quad (1)$$

使利润达到最大的投入量 x^* 可以从 $r'(x^*) = 0$ 得到,由(1)即有

$$f'(x^*) = c'(x^*) \quad (2)$$

按照经济学中将“导数”称为“边际”的习惯, $f'(x)$ 称为边际产值, $c'(x)$ 称为边际成本,分别是投入量 x 变化一个单位时产值和成本的改变量。最佳投入量 x^* 由(2)式确定,它表明:最大利润在边际产值等于边际成本时达到。这是经济学的一条著名定律。

下面在这条定律的基础上讨论几个具体模型。

最优定价模型 在产品可以全部售出的情况下,产品的投入量 x 就是商品的销售量,产值 $f(x)$ 就是收入。可以一般地假定收入等于销售量乘以价格 p ,即 $f(x) = px$ 。而对于成本函数 $c(x)$,除去投产初期及生产饱和阶段成本与投入量的关系比较复杂以外,在正常生产时期可以简化地假设成本与投入量成正比,即 $c(x) = cx$,等式右边的参数 c 是边际成本,实际上就是单位成本。下面讨论如果对产品投入量没有限制,怎样根据成本和销售情况确定商品的价格,使利润最大。

在市场经济中除了少数生活必需品外,大多数商品的销售量与价格直接相关,价格 p 越高,售量 x 越小,简化地假设二者呈线性关系

$$x = a - bp, a, b > 0 \quad (3)$$

将 $f(x) = px, c(x) = cx$ 及(3)代入(1)式,得

$$r(x) = (p - c)(a - bp) \quad (4)$$

当 a, b, c 给定后容易求出使利润 r 达到最大的定价 p^* 为

$$p^* = \frac{c}{2} + \frac{a}{2b} \quad (5)$$

为了分析(5)的意义,需要考察参数 a, b . 在(3)中 a 可理解为这种商品免费供应时($p=0$)社会的需求,称“绝对需求量”, $b = -\frac{dx}{dp}$ 表示这种商品价格上涨一个单位时销售量下降的幅度,称“市场弹性系数”,反映需求对价格的敏感程度. (5)式表明最优定价由两部分组成,一部分是单位成本 c 的一半,另一部分与绝对需求量 a 成正比,与市场弹性系数 b 成反比,在定性上这是与常识相符的. 至于 a, b 的具体大小,可以由价格 p 和销售量 x 的统计数据,用最小二乘法对(3)式作拟合得到.

投资费用一定下的产值最大模型

这个模型讨论生产两种产品的情况,并可以推广到生产任意多种产品.

设生产者生产甲乙两种产品,其投入量分别为 x_1, x_2 , 单位成本分别为 c_1, c_2 , 投资费用为 s , 得到的产值记作 $f(x_1, x_2)$, 称为产值函数, 则最大利润模型可转化为在投资费用一定的条件

$$c_1 x_1 + c_2 x_2 = s \quad (6)$$

下求投入量 x_1, x_2 , 使产值函数 $f(x_1, x_2)$ 达到最大.

具体地求解这类模型需要知道产值函数的实际意义和解析表达式, 在第4章将给出一些实例. 这里只从图形上直观地分析一下.

与3.4节引入的效用函数和等效用线类似, 我们在 x_1, x_2 平面上画出产值函数 $f(x_1, x_2)$ 的等值线, 称为等产值线, 如图1. 等产值线 $f(x_1, x_2) = v$ 也是一族单调减、下凸、互不相交的曲线, 随着产值 v 的增加曲线向右上方移动, 曲线的具体形状由甲乙两种产品对产值的贡献决定. 等产值线的单调减和互不相交性质是明显的, 而下凸形状可以用稀缺的那种产品的产值比另一种产品的产值更高来解释.

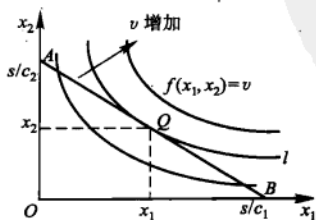


图1 等产值线和产值最大模型

与 3.4 节的效用最大模型类似,条件(6)是图 1 中的直线 AB ,与一条等产值线(图 1 中的 l)的切点为 $Q(x_1, x_2)$,当生产甲乙两种产品的数量为 x_1, x_2 时,产值函数 $f(x_1, x_2)$ 达到最大,而最优解 (x_1, x_2) 满足

$$\frac{\frac{\partial f}{\partial x_1} \Big|_{x_1, x_2}}{\frac{\partial f}{\partial x_2} \Big|_{x_1, x_2}} = \frac{c_1}{c_2} \quad (7)$$

表明当两种产品的边际产值 $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}$ 之比等于它们的价格之比时,产值函数达到最大. 这是数量经济学的又一个重要结论.

产值最大与费用最小的对偶关系^[87] 采用向量符号 $x = (x_1, x_2)^T, c = (c_1, c_2)$, 上面的产值最大模型可以简练地记作

$$g(s, c) = \max \{f(x) \mid cx \leq s\} \quad (8)$$

$g(s, c)$ 的含义是在指定的单位成本 c 下,投资费用不超过 s 时的最大产值. 这里将(6)式的等号写成不等号是为数学上处理的方便,实际上很多情况的最大产值都是在等式条件下取得的.

我们可以对产值最大问题换一种提法:在产值一定的条件下使费用最小,模型记作

$$s(v, c) = \min \{cx \mid f(x) \geq v\} \quad (9)$$

$s(v, c)$ 的含义是在指定的单位成本 c 下,产值不低于 v 时的最小费用. (8)和(9)是对偶的极值问题,正如“求周长一定、面积最大的四边形”与“求面积一定、周长最小的四边形”一样. 对于对偶极值问题,只要解决其中之一,另一个就迎刃而解.

在对偶关系(8)和(9)中,成本函数是简单的线性函数,产值函数则不是(如图 1 的等产值线),实际生产过程中产值函数也常常难以给定. 作为产值最大与费用最小的对偶关系的一个应用,我们用图解法给出从成本函数确定产值函数的一种办法.

给定 v 和 c ,由(9)求得最小费用 $s(v, c) = s$,于是可以画出一条直线 $AB: cx = s$,如图 2. 可以知道,产值不低于 v 的点在 AB 的上方,且 AB 上一定有一点 Q 位于等产值线 $l: f(x) = v$ 上. 改变 c ,重复上述过程,得到一系列不同斜率的直线. 在这些直线上方的区域 $f(x) \geq v$,而区域的边界就是等产值线 $l: f(x) = v$,数学上称为这些直线的包络. 再改变 v ,重复上述过程,就得到其他的等产值线.

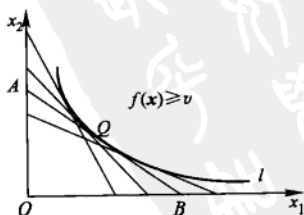


图 2 确定产值函数的图解法

3.6 血管分支

血液在动物的血管中一刻不停地流动,为了维持血液循环,动物的机体要提供能量. 能量的一部分用于供给血管壁以营养,另一部分用来克服血液流动受到的阻力. 消耗的总能量显然与血管系统的几何形状有关. 在长期的生物进化过程中,高级动物血管系统的几何形状应该已经达到消耗能量最小原则下的优化标准了.

我们不可能讨论整个血管系统的几何形状,这会涉及太多的生理学知识. 下面的模型只研究血管分支处粗细血管半径的比例和分岔角度,在消耗能量最小原则下应该取什么样的数值^[10,78].

模型假设

1. 一条粗血管在分支点处分成两条细血管,分支点附近三条血管在同一平面上,有一对称轴. 因为如果不在一个平面上,血管总长度必然增加,导致能量消耗增加,不符合最优原则. 这是一条几何上的假设.

2. 在考察血液流动受到的阻力时,将这种流动视为粘性流体在刚性管道中的运动. 这当然是一种近似,实际上血管是有弹性的,不过这种近似的影响不大. 这是一条物理上的假设.

3. 血液对血管壁提供营养的能量随管壁内表面积及管壁所占体积的增加而增加. 管壁所占体积又取决于管壁厚度,而厚度近似地与血管半径成正比. 这是一条生理上的假设.

根据假设 1,血管分支示意图如图 1 所示. 一条粗血管与两条细血管在 C 点分岔,并形成对称的几何形状. 设粗细血管半径分别是 r 和 r_1 ,分岔处夹角是 θ . 考察长度为 l 的一段粗血管 AC 和长度为 l_1 的两条细血管 CB 和 CB' , $ACB(ACB')$ 的水平 and 竖直距离为 L 和 H ,如图所示. 再设血液在粗细血管中单位时间的流量分别为 q 和 q_1 ,显然 $q = 2q_1$.

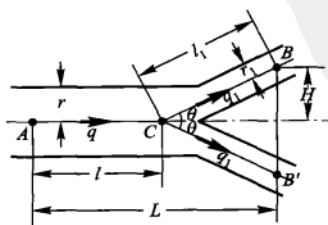


图 1 血管分支示意图

假设2 使我们可以利用流体力学关于粘性流体在刚性管道中流动时能量消耗的定律. 按照 Poiseuille 定律, 血液流过半径 r 、长 l 的一段血管 AC 时, 流量

$$q = \frac{\pi r^4 \Delta p}{8\mu l} \quad (1)$$

其中 Δp 是 A, C 两点的压力差, μ 是血液的粘性系数. 在血液流动过程中, 机体克服阻力所消耗能量为 $E_1 = q \cdot \Delta p$, 将(1)式中的 Δp 代入, 得

$$E_1 = \frac{8\mu q^2 l}{\pi r^4} \quad (2)$$

假设3 比较复杂, 需要作进一步简化. 对于半径为 r 、长度为 l 的血管, 管壁内表面积 $s = 2\pi r l$, 管壁所占体积 $v = s' l$, 其中 s' 是管壁截面积. 记壁厚为 d , 则 $s' = \pi[(r+d)^2 - r^2] = \pi(d^2 + 2rd)$. 设壁厚 d 近似地与半径 r 成正比, 可知 v 近似地与 r^2 成正比. 又因为 s 与 r 成正比, 综合考虑管壁内表面积 s 和管壁所占体积 v 对能量消耗的影响, 可设血液流过长度 l 的血管的过程中, 为血管壁提供营养消耗的能量为

$$E_2 = br^\alpha l \quad (3)$$

其中 $1 \leq \alpha \leq 2$, b 是比例系数.

模型建立 根据上述假设及对假设的进一步分析得到的(2), (3)式, 血液从粗血管 A 点流动到细血管 B, B' 两点的过程中, 机体为克服阻力和供养管壁所消耗的能量为 E_1, E_2 两部分之和, 即有

$$E = \left(\frac{kq^2}{r^4} + br^\alpha \right) l + \left(\frac{kq_1^2}{r_1^4} + br_1^\alpha \right) 2l_1 \quad (4)$$

由图1所示的几何关系不难得到

$$l = L - \frac{H}{\tan \theta}, \quad l_1 = \frac{H}{\sin \theta} \quad (5)$$

将(5)式代入(4)式, 并注意到 $q_1 = q/2$, 能量 E 可表示为 r, r_1 和 θ 的函数, 即

$$E(r, r_1, \theta) = \left(\frac{kq^2}{r^4} + br^\alpha \right) \left(L - \frac{H}{\tan \theta} \right) + \left(\frac{kq^2}{4r_1^4} + br_1^\alpha \right) \frac{2H}{\sin \theta} \quad (6)$$

按照最优化原则, r/r_1 和 θ 的取值应使(6)式表示的 $E(r, r_1, \theta)$ 达到最小.

由 $\frac{\partial E}{\partial r} = 0$ 和 $\frac{\partial E}{\partial r_1} = 0$ 可以得到

$$\begin{cases} -\frac{4kq^2}{r^5} + b\alpha r^{\alpha-1} = 0 \\ -\frac{kq^2}{r_1^5} + b\alpha r_1^{\alpha-1} = 0 \end{cases} \quad (7)$$

从方程(7)可解出

$$\frac{r}{r_1} = 4^{\frac{1}{\alpha+4}} \quad (8)$$

再由 $\frac{\partial E}{\partial \theta} = 0$, 并利用(8)式可得

$$\cos \theta = 2 \left(\frac{r}{r_1} \right)^{-4} \quad (9)$$

将(8)代入(9)式, 则

$$\cos \theta = 2^{\frac{\alpha-4}{\alpha+4}} \quad (10)$$

(8), (10) 两式就是在能量消耗最小原则下血管分岔处几何形状的结果, 由 $1 \leq \alpha \leq 2$, 可以算出 $\frac{r}{r_1}$ 和 θ 的大致范围为

$$1.26 \leq \frac{r}{r_1} \leq 1.32, \quad 37^\circ \leq \theta \leq 49^\circ \quad (11)$$

结果解释 生物学家认为, 上述结果与经验观察吻合得相当好. 由此还可以导出一个有趣的推论.

记动物的大动脉和最细的毛细血管的半径分别为 r_{\max} 和 r_{\min} , 设从大动脉到毛细血管共有 n 次分岔, 将(8)式反复利用 n 次可得

$$\frac{r_{\max}}{r_{\min}} = 4^{\frac{n}{\alpha+4}} \quad (12)$$

r_{\max}/r_{\min} 的实际数值可以测出, 例如对狗而言有 $r_{\max}/r_{\min} \approx 1\,000 \approx 4^5$, 由(12)式可知 $n \approx 5(\alpha+4)$. 因为 $1 \leq \alpha \leq 2$, 所以按照这个模型, 狗的血管应有 25 ~ 30 次分岔. 又因为当血管有 n 次分岔时血管总条数为 2^n , 所以估计狗应约有 $2^{25} \sim 2^{30}$, 即 $3 \times 10^7 \sim 10^9$ 条血管. 这个估计不可过于认真看待, 因为血管分支很难是完全对称的.

3.7 冰山运输

在以盛产石油著称的波斯湾地区, 浩瀚的沙漠覆盖着大地, 水资源十分贫乏, 不得不采用淡化海水的办法为国民提供用水, 成本大约是每立方米淡水 0.1 英镑. 有些专家提出从相距 9 600 km 之遥的南极用拖船运送冰山到波斯湾, 以取代淡化海水的办法. 这个模型要从经济角度研究冰山运输的可行性^[22].

为了计算用拖船运送冰山获得每立方米水所花的费用, 我们需要关于拖船的租金、运量、燃料消耗及冰山运输过程中融化速率等方面的数据, 以此作为建模必需的准备工作的.

模型准备 根据建模的需要搜集到以下数据.

1. 三种拖船的日租金和最大运量.

表1 日租金和最大运量

船 型	小	中	大
日租金/英镑	4.0	6.2	8.0
最大运量/ m^3	5×10^5	10^6	10^7

2. 燃料消耗(英镑/km). 主要依赖于船速和所运冰山的体积, 船型的影响可以忽略.

表2 燃料消耗(单位:英镑/km)

冰山体积/ m^3	10^5	10^6	10^7
船速/ $(\text{km} \cdot \text{h}^{-1})$			
1	8.4	10.5	12.6
3	10.8	13.5	16.2
5	13.2	16.5	19.8

3. 冰山运输过程中的融化速率(即 m/天). 指在冰山与海水、大气接触处每天融化的深度. 融化速率除与船速有关外, 还和运输过程中冰山与南极的距离有关, 这是由于冰山要从南极运往赤道附近的缘故.

表3 融化速率(单位:m/天)

与南极距离/km	0	1 000	>4 000
船速/ $(\text{km} \cdot \text{h}^{-1})$			
1	0	0.1	0.3
3	0	0.15	0.45
5	0	0.2	0.6

建立模型的目的是选择拖船的船型和船速, 使冰山到达目的地后, 可得到的每立方米水所花的费用最低, 并与海水淡化的费用相比较.

根据建模目的和搜集到的有限的资料, 需要作如下的简化假设.

模型假设

1. 拖船航行过程中船速不变, 航行不考虑天气等任何因素的影响. 总航行距离 9 600 km.

2. 冰山形状为球形, 球面各点的融化速率相同. 这是相当无奈的假设, 在冰

山上各点融化速率相同的条件下,只有球形的形状不变,体积的变化才能简单地计算。

3. 冰山到达目的地后, 1 m^3 冰可融化成 0.85 m^3 水。

模型构成 首先需要知道冰山体积在运输过程中的变化情况,然后是计算航行中的燃料消耗,由此可以算出到达目的地后的冰山体积和运费。在计算过程中需要根据搜集到的数据拟合出经验公式。模型构成可分为以下几步。

1. 冰山融化规律

根据假设 2,先确定冰山球面半径的减小,就可以得到冰山体积的变化规律。

记冰山球面半径融化速率为 $r \text{ m/天}$,船速为 $u \text{ km/h}$,拖船与南极距离为 $d \text{ km}$ 。根据表 3 中融化速率的数据,可设 r 是船速 u 的线性函数,且当 $0 \leq d \leq 4000 \text{ km}$ 时 r 与 d 成正比,而当 $d > 4000 \text{ km}$ 时 r 与 d 无关,即设

$$r = \begin{cases} a_1 d(1 + bu), & 0 \leq d \leq 4000 \\ a_2(1 + bu), & d > 4000 \end{cases} \quad (1)$$

其中 a_1, a_2, b 为待定参数。这可以解释为, $0 \leq d \leq 4000 \text{ km}$ 相当于从南极到赤道以南,海水温度随 d 增加而上升,使融化速率 r 也随 d 的增加而变大,而 $d > 4000 \text{ km}$ 后海水温度变化较小,可以忽略。

利用表 3 所给数据确定出^①

$$a_1 = 6.5 \times 10^{-5}, \quad a_2 = 0.2, \quad b = 0.4 \quad (2)$$

当拖船从南极出发航行第 t 天时,与南极的距离为

$$d = 24ut \quad (3)$$

记第 t 天冰山球面半径融化速率为 r_t ,将(2),(3)代入(1)式得

$$r_t = \begin{cases} 1.56 \times 10^{-3} u(1 + 0.4u)t, & 0 \leq t \leq \frac{1000}{6u} \\ 0.2(1 + 0.4u), & t > \frac{1000}{6u} \end{cases} \quad (4)$$

记第 t 天冰山半径为 R_t , 体积为 V_t , 则

$$R_t = R_0 - \sum_{k=1}^t r_k \quad (5)$$

$$V_t = \frac{4\pi}{3} R_t^3, \quad V_0 = \frac{4\pi}{3} R_0^3 \quad (6)$$

其中 R_0, V_0 为从南极启运时冰山的初始半径和体积。由(4)~(6)式可知,冰山

① 由于数据太少,这个结果实际上是凑出来的。读者可从(1),(2)式计算 r ,与表 3 的实际值比较。

体积是船速 u 、初始体积 V_0 和航行天数 t 的函数, 记作 $V(u, V_0, t)$, 有

$$V(u, V_0, t) = \frac{4\pi}{3} \left(\sqrt[3]{\frac{3V_0}{4\pi}} - \sum_{k=1}^t r_k \right)^3 \quad (7)$$

其中 r_k 由(4)式表示.

2. 燃料消耗费用

分析表 2 给出的燃料消耗 (英镑/km, 记作 \bar{q}) 的数据可以看出, q 对船速 u 和冰山体积 V 的对数 $\lg V$ 均按线性关系变化, 所以可设

$$\bar{q} = c_1(u + c_2)(\lg V + c_3) \quad (8)$$

其中 c_1, c_2, c_3 为待定参数. 利用表 2 所给数据可以确定^①

$$c_1 = 0.3, \quad c_2 = 6, \quad c_3 = -1 \quad (9)$$

由(7)~(9)式可将拖船航行第 t 天的燃料消耗记作 $q(u, V_0, t)$ (英镑/天), 且有

$$\begin{aligned} q(u, V_0, t) &= 24u \cdot c_1(u + c_2)[\lg V(u, V_0, t) + c_3] \\ &= 7.2u(u + 6) \left[\lg \frac{4\pi}{3} \left(\sqrt[3]{\frac{3V_0}{4\pi}} - \sum_{k=1}^t r_k \right)^3 - 1 \right] \\ &= 7.2u(u + 6) \left[3 \lg \left(\sqrt[3]{\frac{3V_0}{4\pi}} - \sum_{k=1}^t r_k \right) - 0.378 \right] \end{aligned} \quad (10)$$

3. 运送冰山费用

费用由拖船的租金和燃料消耗两部分组成. 由表 1 知船的日租金取决于船型, 船型又由冰山的初始体积 V_0 决定, 记日租金为 $f(V_0)$, 显然有

$$f(V_0) = \begin{cases} 4.0, & V_0 \leq 5 \times 10^5 \\ 6.2, & 5 \times 10^5 < V_0 \leq 10^6 \\ 8.0, & 10^6 < V_0 \leq 10^7 \end{cases} \quad (11)$$

又因为当船速为 u (km/h) 时冰山抵达目的地所需天数为 $T = \frac{9600}{24u} = \frac{400}{u}$, 所以

租金费用为 $\frac{400f(V_0)}{u}$. 而整个航程的燃料消耗为 $\sum_{i=1}^T q(u, V_0, t)$, 由(10)式得运送冰山的总费用为

$$S(u, V_0) = \frac{400f(V_0)}{u} + 7.2u(u + 6) \left[3 \sum_{i=1}^T \lg \left(\sqrt[3]{\frac{3V_0}{4\pi}} - \sum_{k=1}^i r_k \right) - \frac{151}{u} \right] \quad (12)$$

4. 冰山运抵目的地后可获得水的体积

将 $t = T$ 代入(7)式知, 冰山运抵目的地后的体积为

① 与(2)式一样, 这个结果实际上也是凑出来的.

$$V(u, V_0, T) = \frac{4\pi}{3} \left(\sqrt{\frac{3V_0}{4\pi}} - \sum_{i=1}^T r_i \right)^3 \quad (13)$$

注意到假设(3), 则得到水的体积为

$$W(u, V_0) = \frac{3.4\pi}{3} \left(\sqrt{\frac{3V_0}{4\pi}} - \sum_{i=1}^T r_i \right)^3 \quad (14)$$

5. 每立方米水所需费用

记冰山运抵目的地后每立方米所需费用为 $Y(u, V_0)$, 由(12), (14)式显然有

$$Y(u, V_0) = \frac{S(u, V_0)}{W(u, V_0)} \quad (15)$$

模型求解 这个模型归结为选择船速 u 和冰山初始体积 V_0 , 使(15)式表示的费用 $Y(u, V_0)$ 最小, 其中 $S(u, V_0)$ 由(12), (4)式给出, $W(u, V_0)$ 由(14), (4)式给出. 由于 $f(V_0)$ 是分段函数, 只能固定一系列 V_0 值对 u 求解. 又因为由调查数据(表2, 表3)得到的经验公式是非常粗糙的, 对船速 u 的选取也不用太精细, 所以没有必要用微分法求解这个极值问题. 表4是对几组 (V_0, u) 值的计算结果, 可知若选取最大的冰山初始体积 $V_0 = 10^7 \text{ m}^3$ (当然要租用大型拖船), 船速 $u = 4 \sim 5 \text{ km/h}$, 每立方米水的费用约 0.065 英镑.

表4 不同 V_0, u 下每立方米水的费用(单位: 英镑)

$V_0 \backslash u$	3	3.5	4	4.5	5
10^7	0.072 3	0.068 3	0.064 9	0.066 3	0.065 8
5×10^6	0.225 1	0.201 3	0.183 4	0.184 2	0.179 0
10^6	78.903 2	9.822 0	6.213 8	5.464 7	4.510 2

结果分析 得到的结果虽然小于海水淡化的费用(每立方米 0.1 英镑), 但是模型中未考虑影响航行的种种不利因素, 会拖长航行时间致使冰山抵达目的地后的体积显著地小于模型中的 $V(u, V_0, T)$, 并且没有计算空船费等其他费用. 专家们认为, 只有当用这个模型计算出来的费用显著地小于海水淡化的费用时(譬如小一个数量级), 才有理由考虑采用冰山运输的办法获得淡水.

评注 这个模型的思路是简单的, 建模方法有两点值得注意: 一是根据有限的数(表2, 表3)建立了经验公式(1), (2)和(8), (9), 为整个计算过程提供了基础; 二是假定冰山呈球形, 简化了计算. 读者可以考虑, 如果假定冰山为其他规则的形状, 将如何处理.



1. 在 3.1 节存贮模型的总费用中增加购买货物本身的费用,重新确定最优订货周期和订货批量. 证明在不允许缺货模型和允许缺货模型中结果都与原来的一样.

2. 建立不允许缺货的生产销售存贮模型. 设生产速率为常数 k , 销售速率为常数 r , $k > r$. 在每个生产周期 T 内, 开始的一段时间 ($0 < t < T_0$) 一边生产一边销售, 后来的一段时间 ($T_0 < t < T$) 只销售不生产, 画出贮存量 $q(t)$ 的图形. 设每次生产准备费为 c_1 , 单位时间每件产品贮存费为 c_2 , 以总费用最小为目标确定最优生产周期. 讨论 $k \gg r$ 和 $k \approx r$ 的情况.

3. 在 3.3 节森林救火模型中, 如果考虑消防队员的灭火速度 λ 与开始救火时的火势 b 有关, 试假设一个合理的函数关系, 重新求解模型.

4. 在 3.4 节消费者的选择模型中,

(1) 证明若条件(3)成立, 则 $u(x_1, x_2) = c$ 是单调减、下凸的曲线.

(2) 验证(4), (6), (8)式给出的效用函数是否满足条件(3).

(3) 若消费者的效用函数为(8)式, 求最优比例 $p_1 q_1 / p_2 q_2$, 并分析参数 a, b 的意义.

(4) 若商品甲的价格 p_1 增加, 其余条件不变, 讨论消费点 Q 的变化.

(5) 若消费者购买商品的钱 s 增加, 其余条件不变, 讨论消费点 Q 的变化.

(6) 推广到消费者购买 m ($m > 2$) 种商品的情况.

5. 在 3.5 节最优定价模型中, 如果考虑到成本 q 随着投入 x 的增加而降低, 试作出合理的假设, 重新求解模型.

6. 在考虑最优定价时设销售期为 T , 由于商品的损耗, 成本 q 随时间增长, 设 $q = q_0 + \beta t$, β 为增长率. 又设单位时间的销售量为 $x = a - bp$ (p 为价格). 今将销售期分为 $0 < t < T/2$ 和 $T/2 < t < T$ 两段, 每段的价格固定, 记作 p_1, p_2 . 求 p_1, p_2 的最优值, 使销售期内的总利润最大. 如果要求销售期 T 内的总销售量为 Q_0 , 再求 p_1, p_2 的最优值.

*7. 要在雨中从一处沿直线跑到另一处, 若雨速为常数且方向不变, 试建立数学模型讨论是否跑得越快, 淋雨量越少.

将人体简化成一个长方体, 高 $a = 1.5$ m (颈部以下), 宽 $b = 0.5$ m, 厚 $c = 0.2$ m. 设跑步距离 $d = 1000$ m, 跑步最大速度 $v_m = 5$ m/s, 雨速 $u = 4$ m/s, 降雨量 $w = 2$ cm/h, 记跑步速度为 v . 按以下步骤进行讨论^[26]:

(1) 不考虑雨的方向, 设降雨淋遍全身, 以最大速度跑步, 估计跑完全程的总淋雨量.

(2) 雨从迎面吹来, 雨线与跑步方向在同一平面内, 且与人体的夹角为 θ , 如图 1. 建立总淋雨量与速度 v 及参数 a, b, c, d, u, w, θ 之间的关系, 问速度 v 多大, 总淋雨量最少. 计算 $\theta = 0, \theta = 30^\circ$ 时的总淋雨量.

(3) 雨从背面吹来, 雨线方向与跑步方向在同一平面内, 且与人体的夹角为 α , 如图 2. 建立总淋雨量与速度 v 及参数 a, b, c, d, u, w, α 之间的关系, 问速度 v 多大, 总淋雨量最少. 计算 $\alpha = 30^\circ$ 时的总淋雨量.

(4) 以总淋雨量为纵轴, 速度 v 为横轴, 对(3)作图(考虑 α 的影响), 并解释结果的实际意义.

(5) 若雨线方向与跑步方向不在同一平面内, 模型会有什么变化?

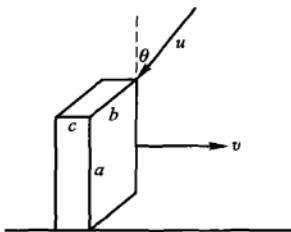


图 1

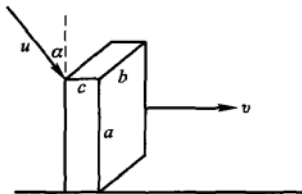


图 2

* 8. 甲乙两公司通过广告来竞争销售商品的数量, 广告费分别是 x 和 y . 设甲乙公司商品的销售量在两公司总销售量中占的份额, 是它们的广告费在总广告费中所占份额的函数 $f\left(\frac{x}{x+y}\right)$ 和 $f\left(\frac{y}{x+y}\right)$. 又设公司的收入与销售量成正比, 从收入中扣除广告费后即为公司利润. 试构造模型的图形, 并讨论甲公司怎样确定广告费才能使利润最大.

(1) 令 $t = \frac{x}{x+y}$, 则 $f(t) + f(1-t) = 1$. 画出 $f(t)$ 的示意图.

(2) 写出甲公司利润的表达式 $p(x)$. 对于一定的 y , 使 $p(x)$ 最大的 x 的最优值应满足什么关系. 用图解法确定这个最优值^[10].

* 9. 人行走时作的功是抬高人体重心所需势能与两腿运动所需动能之和. 试建立模型讨论在作功最小的准则下每秒走几步最合适(匀速行走).

(1) 设腿长 l , 步长 s , 证明人体重心在行走时升高 $\delta \approx s^2/8l (s < l)$.

(2) 将腿看作均匀直杆, 行走看作腿绕腰部的转动. 设腿的质量 m , 行走速度 v , 证明单位时间所需动能为 $mv^2/6s$.

(3) 设人体质量 M , 证明在速度 v 一定时每秒行走 $n = \sqrt{\frac{3Mg}{4ml}}$ 步作功最小. 实际上,

$\frac{M}{m} \approx 4, l \approx 1 \text{ m}$, 分析这个结果合理吗.

(4) 将(2)的假设修改为: 腿的质量集中在脚部, 行走看作脚的直线运动. 证明结果应为 $n = \sqrt{\frac{Mg}{4ml}}$ 步. 分析这个结果是否合理^[66].

* 10. 观察鱼在水中的运动发现, 它不是水平游动, 而是突发性、锯齿状地向上游动和向下滑行. 可以认为这是在长期进化过程中鱼类选择的消耗能量最小的运动方式.

(1) 设鱼总是以常速 v 运动, 鱼在水中净重 w , 向下滑行时的阻力是 w 在运动方向的分力; 向上游动时所需的力是 w 在运动方向分力与游动所受阻力之和, 而游动的阻力是滑行阻力的 k 倍. 水平方向游动时的阻力也是滑行阻力的 k 倍. 写出这些力.

(2) 证明当鱼要从 A 点到达处于同一水平线上的 B 点时(见图 3), 沿折线 ACB 运动消耗的能量与沿水平线 AB 运动消耗的能量之比为(向下滑行不消耗能量) $\frac{k \sin \alpha + \sin \beta}{k \sin(\alpha + \beta)}$.

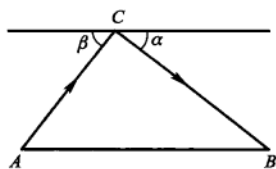


图 3

(3) 据实际观察 $\tan \alpha \approx 0.2$, 试对不同的 k 值(1.5, 2, 3), 根据消耗能量最小的准则估计最佳的 β 值^[10].



第4章

数学规划模型

在上一章中我们看到,建立优化模型要确定优化的目标和寻求的决策.用 x 表示决策变量, $f(x)$ 表示目标函数.实际问题一般对决策变量 x 的取值范围有限制,不妨记作 $x \in \Omega$, Ω 称为可行域.优化问题的数学模型可表示为

$$\min(\text{或} \max)f(x), x \in \Omega$$

在第3章, x 通常是1维或2维变量, Ω 通常是1维或2维的非负域.

实际中的优化问题通常有多个决策变量,用 n 维向量 $x = (x_1, x_2, \dots, x_n)^T$ 表示,目标函数 $f(x)$ 是多元函数,可行域 Ω 比较复杂,常用一组不等式(也可以有等式) $g_i(x) \leq 0 (i=1, 2, \dots, m)$ 来界定,称为约束条件.一般地,这类模型可表述成如下形式

$$\begin{aligned} \min_x \quad & z = f(x) \\ \text{s. t.} \quad & g_i(x) \leq 0, i = 1, 2, \dots, m \end{aligned}$$

这里的 s. t. (subject to) 是“受约束于”的意思.

显然,上述模型属于多元函数的条件极值问题的范围,然而许多实际问题归结出的这种形式的优化模型,其决策变量个数 n 和约束条件个数 m 一般较大,并且最优解往往在可行域的边界上取得,这样就不能简单地用微分法求解,数学规划是解决这类问题的有效方法.

需要指出的是,本章无意涉及数学规划(或运筹学)的具体计算方法,仍然着重于从数学建模的角度,介绍如何建立若干实际优化问题的模型,并且在用现成的数学软件求解后,对结果作一些分析.

4.1 奶制品的生产与销售

企业内部的生产计划有各种不同的情况.从空间层次看,在工厂级要根据外部需求和内部设备、人力、原料等条件,以最大利润为目标制订产品的生产计划,在车间级则要根据产品生产计划、工艺流程、资源约束及费用参数等,以最小成本为目标制订生产作业计划.从时间层次看,若在短时间内认为外部需求和内部资源等不随时间变化,可制订单阶段生产计划,否则就要制订多阶段生产计划.

本节选择几个单阶段生产计划的实例,说明如何建立这类问题的数学规划模型,利用软件求解并对输出结果作一些分析.

例1 加工奶制品的生产计划

问题 一奶制品加工厂用牛奶生产 A_1, A_2 两种奶制品,1 桶牛奶可以在甲类设备上用 12 h 加工成 3 kg A_1 ,或者在乙类设备上用 8 h 加工成 4 kg A_2 . 根据市场需求,生产的 A_1, A_2 全部能售出,且每千克 A_1 获利 24 元,每千克 A_2 获利 16 元. 现在加工厂每天能得到 50 桶牛奶的供应,每天正式工人总的劳动时间为 480 h,并且甲类设备每天至多能加工 100 kg A_1 ,乙类设备的加工能力没有限制. 试为该厂制订一个生产计划,使每天获利最大,并进一步讨论以下 3 个附加问题:

1) 若用 35 元可以买到 1 桶牛奶,应否作这项投资? 若投资,每天最多购买多少桶牛奶?

2) 若可以聘用临时工人以增加劳动时间,付给临时工人的工资最多是每小时几元?

3) 由于市场需求变化,每千克 A_1 的获利增加到 30 元,应否改变生产计划?

问题分析 这个优化问题的目标是使每天的获利最大,要做的决策是生产计划,即每天用多少桶牛奶生产 A_1 ,用多少桶牛奶生产 A_2 (也可以是每天生产多少千克 A_1 ,多少千克 A_2),决策受到 3 个条件的限制:原料(牛奶)供应、劳动时间、甲类设备的加工能力. 按照题目所给,将决策变量、目标函数和约束条件用数学符号及式子表示出来,就可得到下面的模型.

基本模型

决策变量: 设每天用 x_1 桶牛奶生产 A_1 ,用 x_2 桶牛奶生产 A_2 .

目标函数: 设每天获利为 z 元. x_1 桶牛奶可生产 $3x_1$ kg A_1 ,获利 $24 \times 3x_1$, x_2 桶牛奶可生产 $4x_2$ kg A_2 ,获利 $16 \times 4x_2$,故 $z = 72x_1 + 64x_2$.

约束条件:

原料供应 生产 A_1, A_2 的原料(牛奶)总量不得超过每天的供应,即 $x_1 + x_2 \leq 50$;

劳动时间 生产 A_1, A_2 的总加工时间不得超过每天正式工人总的劳动时间,即 $12x_1 + 8x_2 \leq 480$;

设备能力 A_1 的产量不得超过甲类设备每天的加工能力,即 $3x_1 \leq 100$;

非负约束 x_1, x_2 均不能为负值,即 $x_1 \geq 0, x_2 \geq 0$.

综上可得

$$\max z = 72x_1 + 64x_2 \quad (1)$$

$$\text{s. t. } x_1 + x_2 \leq 50 \quad (2)$$

$$12x_1 + 8x_2 \leq 480 \quad (3)$$

$$3x_1 \leq 100 \quad (4)$$

$$x_1 \geq 0, x_2 \geq 0 \quad (5)$$

这就是该问题的基本模型. 由于目标函数和约束条件对于决策变量而言都是线性的, 所以称为线性规划 (Linear Programming, 简记作 LP).

模型分析与假设

从本章下面的实例可以看到, 许多实际的优化问题的数学模型都是线性规划 (特别是在像生产计划这样的经济管理领域), 这不是偶然的. 让我们分析一下线性规划具有哪些特征, 或者说, 实际问题具有什么性质, 其模型才是线性规划.

- **比例性** 每个决策变量对目标函数的“贡献”, 与该决策变量的取值成正比; 每个决策变量对每个约束条件右端项的“贡献”, 与该决策变量的取值成正比.

- **可加性** 各个决策变量对目标函数的“贡献”, 与其他决策变量的取值无关; 各个决策变量对每个约束条件右端项的“贡献”, 与其他决策变量的取值无关.

- **连续性** 每个决策变量的取值是连续的.

比例性和可加性保证了目标函数和约束条件对于决策变量的线性性, 连续性则允许得到决策变量的实数最优解.

对于本例, 能建立上面的线性规划模型, 实际上是事先作了如下的假设:

1. A_1, A_2 两种奶制品每千克的获利是与它们各自产量无关的常数, 每桶牛奶加工出 A_1, A_2 的数量和所需的时间是与它们各自的产量无关的常数;

2. A_1, A_2 每千克的获利是与它们相互间产量无关的常数, 每桶牛奶加工出 A_1, A_2 的数量和所需的时间是与它们相互间产量无关的常数;

3. 加工 A_1, A_2 的牛奶的桶数可以是任意实数.

这 3 条假设恰好保证了上面的 3 条性质. 当然, 在现实生活中这些假设只是近似成立的, 比如, A_1, A_2 的产量很大时, 自然会使它们每千克的获利有所减少.

由于这些假设对于书中给出的、经过简化的实际问题是如此明显地成立, 本章下面的例题就不再一一列出类似的假设了. 不过, 读者在打算用线性规划模型解决现实生活中的实际问题时, 应该考虑上面 3 条性质是否近似地满足.

模型求解

图解法 这个线性规划模型的决策变量为 2 维, 用图解法既简单, 又便于直观地把握线性规划的基本性质.

将约束条件 (2) ~ (5) 中的不等号改为等号, 可知它们是 x_1, x_2 平面上的 5 条直线, 依次记为 $L_1 \sim L_5$, 如图 1. 其中 L_4, L_5 分别是 x_1 轴和 x_2 轴, 并且不难判断, (2) ~ (5) 式界定的可行域是 5 条直线上的线段所围成的 5 边形 $OABCD$. 容

易算出, 5 个顶点的坐标为: $O(0, 0)$, $A(0, 50)$, $B(20, 30)$, $C(100/3, 10)$, $D(100/3, 0)$ 。

目标函数(1)中的 z 取不同数值时, 在图 1 中表示一组平行直线(虚线), 称等值线族。如 $z = 0$ 是过 O 点的直线, $z = 2400$ 是过 D 点的直线, $z = 3040$ 是过 C 点的直线, ……。可以看出, 当这族平行线向右上方移动到过 B 点时, $z = 3360$, 达到最大值, 所以 B 点的坐标 $(20, 30)$ 即为最优解: $x_1 = 20, x_2 = 30$ 。

我们直观地看到, 由于目标函数和约束条件都是线性函数, 在 2 维情形, 可行域为直线段围成的凸多边形, 目标函数的等值线为直线, 于是最优解一定在凸多边形的某个顶点取得。

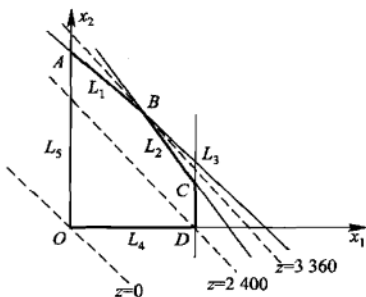


图 1 模型的图解法

推广到 n 维情形, 可以猜想, 最优解会在约束条件所界定的一个凸多面体(可行域)的某个顶点取得。线性规划的理论告诉我们, 这个猜想是正确的。

软件实现 求解线性规划有不少现成的数学软件, 比如用 LINGO 软件就可以很方便地实现。在 LINGO 下新建一个模型文件(即 LINGO 程序, 一般以“LG4”为后缀名), 像书写模型(1)~(5)一样, 直接输入:

```
model:
max = 72 * x1 + 64 * x2;
[milk] x1 + x2 < 50;
[time] 12 * x1 + 8 * x2 < 480;
[cpct] 3 * x1 < 100;
end
```

注: LINGO 程序总是以“model:”开始, 最后以“end”结束(也可以省略不写); 字母不区分大小写; 每个语句都必须以分号“;”结束(注意必须是英文的分号)。LINGO 中已规定所有决策变量均为非负, 故(5)式不必输入; 模型中符号 \leq 、 \geq 用“<=”、“>=”形式输入, 它们与“<”、“>”等效。输入模型中第 1 行为目标函数, [milk]、[time]、[cpct] 是为了对各约束条件命名, 便于从输出结果中查找相应信息(也可以不对约束命名, 此时 LINGO 会自动用数字按顺序对约束条件命名)。

将文件存储并命名后, 选择菜单“LINGO|Solve”执行, 即可得到如下输出:

Global optimal solution found.

Objective value:

3360.000

Total solver iterations:

2

Variable	Value	Reduced Cost
X1	20.000000	0.000000
X2	30.000000	0.000000
Row	Slack or Surplus	Dual Price
1	3360.000	1.000000
MILK	0.000000	48.00000
TIME	0.000000	2.000000
CPCT	40.00000	0.000000

上面结果的前3行告诉我们, LINGO 求出了模型的全局最优解(Global optimal solution), 最优值为3360(即最大利润为3360元), 迭代次数为2次。接下来的3行告诉我们, 这个线性规划的最优解为 $x_1 = 20, x_2 = 30$ (即用20桶牛奶生产 A_1 , 30桶牛奶生产 A_2)。对其中“Reduced Cost”的含义, 将在例2中结合问题3)的讨论进行说明。

结果分析

上面的输出中除了告诉我们问题的最优解和最优值以外, 还有许多对分析结果有用的信息, 下面结合题目中提出的3个附加问题, 并利用图解法的直观给予说明。

(1) 3个约束条件的右端不妨看作3种“资源”: 原料、劳动时间、甲类设备的加工能力。输出第8~11行 Slack or Surplus 给出这3种资源在最优解下是否有剩余: 原料[MILK]、劳动时间[TIME]的剩余均为0, 甲类设备[CPCT]尚余40 kg加工能力。这与图解法的如下结果一致: 最优解在B点(图1中约束条件2, 3所定义的直线 L_1 和 L_2 的交点)取得, 表明原料、劳动时间已用完, 而甲类设备的能力有余。一般称“资源”剩余为0的约束为紧约束(有效约束)。

(2) 目标函数可以看作“效益”, 成为紧约束的“资源”一旦增加, “效益”必然跟着增长。输出第8~11行 Dual Prices 给出这3种资源在最优解下“资源”增加1个单位时“效益”的增量: 原料[MILK]增加1个单位(1桶牛奶)时利润增长48元, 劳动时间[TIME]增加1个单位(1h)时利润增长2元, 而增加非紧约束[CPCT]甲类设备的能力显然不会使利润增长。这里, “效益”的增量可以看作“资源”的潜在价值, 经济学上称为影子价格, 即1桶牛奶的影子价格为48元, 1h劳动的影子价格为2元, 甲类设备的影子价格为0。

读者可以用直接求解的办法验证上面的结论, 即将输入文件中原料约束[MILK]右端的50改为51, 看看得到的最优值(利润)是否恰好增长48元。

用影子价格的概念很容易回答附加问题1): 用35元可以买到1桶牛奶, 低于1桶牛奶的影子价格, 当然应该做这项投资。类似地, 可以回答附加问题2):

聘用临时工人以增加劳动时间,付给的工资低于劳动时间的影子价格才可以增加利润,所以工资最多是每小时 2 元。

(3) 目标函数的系数发生变化时(假定约束条件不变),最优解和最优值会改变吗?这个问题不能简单地回答。从图 1 看,目标函数的系数决定了等值线族的斜率,原题中该斜率(取绝对值,下同)为 $72/64 = 9/8$,介于直线 L_1 的斜率 1 与 L_2 的斜率 $3/2$ 之间,最优解自然在 L_1 和 L_2 的交点 B 取得。并且只要目标函数系数的变化使得等值线族的斜率仍然在 $(1, 3/2)$ 范围内,这个最优解就不会改变。而当目标函数系数的变化使得等值线族的斜率小于 1 时,最优解将在 A 点取得,大于 $3/2$ 时,最优解将在 C 点取得。

这种对目标函数系数变化的影响的讨论,通常称为对目标函数系数的敏感性分析。LINGO 在缺省设置中不会给出这种敏感性分析结果,但可以通过修改 LINGO 选项得到。具体作法是:选择“LINGO|Options”菜单,在弹出的选项卡中选择“General Solver”,然后找到选项“Dual Computations”,在下拉框中选中“Prices & Ranges”,应用或保存设置。重新运行“LINGO|Solve”,然后选择“LINGO|Ranges”菜单,则得到如下输出:

Ranges in which the basis is unchanged:

Objective Coefficient Ranges			
Variable	Current Coefficient	Allowable Increase	Allowable Decrease
X1	72.00000	24.00000	8.000000
X2	64.00000	8.000000	16.00000
Righthand Side Ranges			
Row	Current RHS	Allowable Increase	Allowable Decrease
MILK	50.00000	10.00000	6.666667
TIME	480.0000	53.33333	80.00000
CPCT	100.0000	INFINITY	40.00000

上面输出的第 2~6 行“Current Coefficient”(当前系数)对应的“Allowable Increase”和“Allowable Decrease”给出了最优解不变条件下目标函数系数的允许变化范围: x_1 的系数为 $(72 - 8, 72 + 24)$,即 $(64, 96)$; x_2 的系数为 $(64 - 16, 64 + 8)$,即 $(48, 72)$ 。注意: x_1 系数的允许范围需要 x_2 系数 64 不变,反之亦然。

用这个结果很容易回答附加问题 3):若每千克 A_1 的获利增加到 30 元,则 x_1 系数变为 $30 \times 3 = 90$,在允许范围内,所以不应改变生产计划。

(4) 对“资源”的影子价格作进一步的分析。从图 1 看,随着原料(牛奶)的增加,直线 L_1 向右上方平移, L_1 与 L_2 的交点 B (它仍是最优解)向 A 点靠近,在

这个过程中,每增加 1 桶牛奶利润增长 48 元(影子价格).但是,当 B 点与 A 点重合后再增加牛奶就不可能使利润增长了.这就是说,影子价格的作用(即在最优解下“资源”增加 1 个单位时“效益”的增量)是有限制的.这种对影子价格在什么条件下才有意义的讨论,通常称为对资源约束右端项的敏感性分析.上面输出的第 7~12 行“Current RHS”(当前右端项)对应的“Allowable Increase”和“Allowable Decrease”给出了影子价格有意义条件下约束右端项的限制范围:原料[MILK]最多增加 10 桶牛奶,劳动时间[TIME]最多增加 53.3 h.

现在可以回答附加问题 1) 的第 2 问:虽然应该批准用 35 元买 1 桶牛奶的投资,但每天最多购买 10 桶牛奶.类似地,可以用低于 2 元/h 的工资聘用临时工人以增加劳动时间,但最多增加 53.3 h.

需要注意的是:一般情况下 LINGO 给出的敏感性分析结果只是充分条件,如上述“最多增加 10 桶牛奶”应理解为“增加 10 桶牛奶”一定是有利可图的,但并不意味着“增加 10 桶以上的牛奶”一定不是有利可图的(对最大可增加的劳动时间也应该类似地理解),只是此时无法通过敏感性分析直接得到结论,而需要重新求解新的模型进行判断.以后我们对此不再特别进行说明(同样,对目标函数系数给出的敏感性分析结果也只是充分条件).

评注 本例在产品利润、加工时间等参数均可设为常数的情况下,建立了线性规划模型.线性规划模型的三要素是:决策变量、目标函数和约束条件.线性规划模型可以方便地用 LINGO 软件求解,得到内容丰富的输出,而且利用其中的影子价格和敏感性分析,可对模型结果作进一步的研究,它们对实际问题常常是十分有益的.

例 2 奶制品的生产销售计划

问题 例 1 给出的 A_1, A_2 两种奶制品的生产条件、利润及工厂的“资源”限制全都不变.为增加工厂的获利,开发了奶制品的深加工技术:用 2 h 和 3 元加工费,可将 1 kg A_1 加工成 0.8 kg 高级奶制品 B_1 ,也可将 1 kg A_2 加工成 0.75 kg 高级奶制品 B_2 ,每千克 B_1 能获利 44 元,每千克 B_2 能获利 32 元.试为该厂制订一个生产销售计划,使每天的净利润最大,并讨论以下问题:

(1) 若投资 30 元可以增加供应 1 桶牛奶,投资 3 元可以增加 1 h 劳动时间,应否作这些投资?若每天投资 150 元,可赚回多少?

(2) 每千克高级奶制品 B_1, B_2 的获利经常有 10% 的波动,对制订的生产销售计划有无影响?若每千克 B_1 的获利下降 10%,计划应该变化吗?

(3) 若公司已经签订了每天销售 10 kg 的合同并且必须满足,该合同对公司的利润有什么影响?

问题分析 要求制订生产销售计划,决策变量可以像例 1 那样,取作每天用多少桶牛奶生产 A_1, A_2 ,再添上用多少千克 A_1 加工 B_1 ,用多少千克 A_2 加工 B_2 ,

但是由于问题要分析 B_1, B_2 的获利对生产销售计划的影响, 所以决策变量取作 A_1, A_2, B_1, B_2 每天的销售量更方便. 目标函数是工厂每天的净利润—— A_1, A_2, B_1, B_2 的获利之和扣除深加工费用. 约束条件基本不变, 只是要添上 A_1, A_2 深加工时间的约束. 在与例 1 类似的假定下用线性规划模型解决这个问题.

基本模型

决策变量: 设每天销售 x_1 kg A_1, x_2 kg A_2, x_3 kg B_1, x_4 kg B_2 , 用 x_5 kg A_1 加工 B_1, x_6 kg A_2 加工 B_2 (增设 x_5, x_6 可使下面的模型简单).

目标函数: 设每天净利润为 z , 容易写出 $z = 24x_1 + 16x_2 + 44x_3 + 32x_4 - 3x_5 - 3x_6$.

约束条件:

原料供应 每天生产 $A_1, x_1 + x_5$ kg, 用牛奶 $(x_1 + x_5)/3$ 桶, 每天生产 $A_2, x_2 + x_6$ kg, 用牛奶 $(x_2 + x_6)/4$ 桶, 二者之和不得超过每天的供应量 50 桶;

劳动时间 每天生产 A_1, A_2 的时间分别为 $4(x_1 + x_5)$ 和 $2(x_2 + x_6)$, 加工 B_1, B_2 的时间分别为 $2x_3$ 和 $2x_4$, 二者之和不得超过总的劳动时间 480 h;

设备能力 A_1 的产量 $x_1 + x_5$ 不得超过甲类设备每天的加工能力 100 kg;

非负约束 x_1, x_2, \dots, x_6 均为非负.

附加约束 1 kg A_1 加工成 0.8 kg B_1 , 故 $x_3 = 0.8x_5$, 类似地 $x_4 = 0.75x_6$.

由此得基本模型:

$$\max z = 24x_1 + 16x_2 + 44x_3 + 32x_4 - 3x_5 - 3x_6 \quad (6)$$

$$\text{s. t. } \frac{x_1 + x_5}{3} + \frac{x_2 + x_6}{4} \leq 50 \quad (7)$$

$$4(x_1 + x_5) + 2(x_2 + x_6) + 2x_3 + 2x_4 \leq 480 \quad (8)$$

$$x_1 + x_5 \leq 100 \quad (9)$$

$$x_3 = 0.8x_5 \quad (10)$$

$$x_4 = 0.75x_6 \quad (11)$$

$$x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \quad (12)$$

这仍然是一个线性规划模型.

模型求解 用 LINGO 软件求解, 输入文件时为方便起见将 (7) 式改写为

$$4x_1 + 3x_2 + 4x_5 + 3x_6 \leq 600 \quad (7')$$

(8) 式改写为

$$4x_1 + 2x_2 + 6x_5 + 4x_6 \leq 480 \quad (8')$$

输入并求解, 可得如下输出:

Global optimal solution found.

Objective value: 3460.800

Total solver iterations: 2

Variable	Value	Reduced Cost
X1	0.000000	1.680000
X2	168.0000	0.000000
X3	19.20000	0.000000
X4	0.000000	0.000000
X5	24.00000	0.000000
X6	0.000000	1.520000
Row	Slack or Surplus	Dual Price
1	3460.800	1.000000
MILK	0.000000	3.160000
TIME	0.000000	3.260000
CPCT	76.00000	0.000000
5	0.000000	44.00000
6	0.000000	32.00000

Ranges in which the basis is unchanged:

Objective Coefficient Ranges			
Variable	Current Coefficient	Allowable Increase	Allowable Decrease
X1	24.00000	1.680000	INFINITY
X2	16.00000	8.150000	2.100000
X3	44.00000	19.75000	3.166667
X4	32.00000	2.026667	INFINITY
X5	-3.000000	15.80000	2.533333
X6	-3.000000	1.520000	INFINITY

Righthand Side Ranges			
Row	Current RHS	Allowable Increase	Allowable Decrease
MILK	600.0000	120.0000	280.0000
TIME	480.0000	253.3333	80.00000
CPCT	100.0000	INFINITY	76.00000
5	0.0	INFINITY	19.20000
6	0.0	INFINITY	0.0

最优解为 $x_1 = 0$, $x_2 = 168$, $x_3 = 19.2$, $x_4 = 0$, $x_5 = 24$, $x_6 = 0$, 最优值为 $z = 3460.8$, 即每天生产销售 168 kg A_2 和 19.2 kg B_1 (不出售 A_1, B_2), 可获净利润 3460.8 元. 为此, 需用 8 桶牛奶加工成 A_1 , 42 桶加工成 A_2 , 并将得到的 24 kg

A_1 全部加工成 B_1 .

和例 1 一样,原料(牛奶)、劳动时间为紧约束.

结果分析

利用输出中的影子价格和敏感性分析讨论以下问题:

(1) 上述结果给出,约束[MILK]、[TIME]的影子价格分别为 3.16 和 3.26,注意到约束[MILK]的影子价格为(7')右端增加 1 个单位时目标函数的增量,由(7)式可知,增加 1 桶牛奶可使净利润增长 $3.16 \times 12 = 37.92$ 元,约束[TIME]的影子价格则说明:增加 1 h 劳动时间可使净利润增长 3.26 元.所以应该投资 30 元增加供应 1 桶牛奶,或投资 3 元增加 1 h 劳动时间.若每天投资 150 元,增加供应 5 桶牛奶,可赚回 $37.92 \times 5 = 189.6$ 元.但是通过投资增加牛奶的数量是有限制的,输出结果表明,约束[MILK]右端的允许变化范围为 $(600 - 280, 600 + 120)$,相当于(7)式右端的允许变化范围为 $(50 - 23.3, 50 + 10)$,即最多增加供应 10 桶牛奶.

(2) 上述结果给出,最优解不变条件下目标函数系数的允许变化范围: x_3 的系数为 $(44 - 3.17, 44 + 19.75)$; x_4 的系数为 $(32 - \infty, 32 + 2.03)$.所以当 B_1 的获利向下波动 10%,或 B_2 的获利向上波动 10%时,上面得到的生产销售计划将不再一定是最优的,应该重新制订.如若每千克 B_1 的获利下降 10%,应将原模型(6)式中 x_3 的系数改为 39.6,重新计算,得到的最优解为 $x_1 = 0, x_2 = 160, x_3 = 0, x_4 = 30, x_5 = 0, x_6 = 40$,最优值为 $z = 3400$,即 50 桶牛奶全部加工成 200 kg A_2 ,出售其中 160 kg,将其余 40 kg 加工成 30 kg B_2 出售,获净利润 3400 元,可见计划变化很大,这就是说,(最优)生产计划对 B_1 或 B_2 获利的波动是很敏感的.

(3) 上述结果给出,变量 x_1 对应的“Reduced Cost”严格大于 0(为 1.68),首先表明目前最优解中 x_1 的取值一定为 0;其次,如果限定 x_1 的取值大于等于某个正数,则 x_1 从 0 开始增加一个单位时,(最优的)目标函数值将减少 1.68.因此,若公司已经签订了每天销售 10 kg 的合同并且必须满足,该合同将会使公司利润减少 $1.68 \times 10 = 16.8$ 元,即最优利润为 $3460.8 - 16.8 = 3444$ 元.也可以反过来理解:如果将目标函数中 x_1 对应的费用系数增加不小于 1.68,则在最优解中 x_1 将可以取到严格大于 0 的值.

有两点需要注意:一是与敏感性分析结果类似,这只是一个充分条件,即如果一个变量对应的“Reduced Cost”大于 0,则当前最优解中 x_1 的取值一定为 0;反之不成立,如上面最优解中 x_4 的取值为 0,对应的“Reduced Cost”也等于 0 而不是大于 0(此时的“Reduced Cost”就不能按上面的解释来理解).二是“Reduced Cost”有意义也是有条件的,但条件不能通过上述结果直接得到,例如,如果将 x_1 限定为不小于 100,则问题的最优值为 3040,而不再是 $3460.8 - 1.68 \times$

100 = 3 292.8.

评注 与例1相比,例2多了两种产品 B_1, B_2 , 它们的销售量与 A_1, A_2 的加工程量之间存在等式关系(10), (11), 虽然可以据此消掉2个变量, 但是会增加人工计算, 并使模型变得复杂. 我们建模的原则是尽可能利用原始的数据信息, 而把尽量多的计算留给计算机去作.

4.2 自来水输送与货机装运

钢铁、煤炭、水电等生产、生活物资从若干供应点运送到一些需求点, 怎样安排输送方案使运费最小, 或者利润最大? 各种类型的货物装箱, 由于受体积、质量等的限制, 如何相互搭配装载, 使获利最高, 或者装箱数量最少? 本节将通过两个例子讨论用数学规划模型解决这类问题的方法.

例1 自来水输送问题

问题 某市有甲、乙、丙、丁四个居民区, 自来水由A, B, C三个水库供应. 四个区每天必须得到保证的基本生活用水量(单位: 10^3 t)分别为30, 70, 10, 10, 但由于水源紧张, 三个水库每天最多只能分别供应自来水50, 60, 50. 由于地理位置的差别, 自来水公司从各水库向各区送水所需付出的引水管理费不同(见表1, 其中C水库与丁区之间没有输水管道), 其他管理费用(单位: 元/ 10^3 t)都是450. 根据公司规定, 各区用户按照统一标准900收费. 此外, 四个区都向公司申请了额外用水量, 分别为每天50, 70, 20, 40. 该公司应如何分配供水水量, 才能获利最多?

为了增加供水量, 自来水公司正在考虑进行水库改造, 使三个水库每天的最大供水量都提高一倍, 问那时供水方案应如何改变? 公司利润可增加到多少?

表1 从水库向各区送水的引水管理费

引水管理费	甲	乙	丙	丁
A	160	130	220	170
B	140	130	190	150
C	190	200	230	/

问题分析 分配供水水量就是安排从三个水库向四个区送水的方案, 目标是获利最多. 而从题目给出的数据看, A, B, C三个水库的供水量160, 不超过四个区的基本生活用水量与额外用水量之和300, 因而总能全部卖出并获利, 于是自来水公司每天的总收入是 $900 \times (50 + 60 + 50) = 144\,000$ 元, 与送水方案无关.

同样,公司每天的其他管理费用 $450 \times (50 + 60 + 50) = 72\,000$ 元也与送水方案无关. 所以,要使利润最大,只需使引水管理费最小即可. 另外,送水方案自然要受三个水库的供应量和四个区的需求量的限制.

模型建立

很明显,决策变量为 A, B, C 三个水库 ($i = 1, 2, 3$) 分别向甲、乙、丙、丁四个区 ($j = 1, 2, 3, 4$) 的供水量. 设水库 i 向 j 区的日供水量为 x_{ij} . 由于 C 水库与丁区之间没有输水管道, 即 $x_{34} = 0$, 因此只有 11 个决策变量.

由上分析,问题的目标可以从获利最多转化为引水管理费最少,于是有

$$\begin{aligned} \min z = & 160x_{11} + 130x_{12} + 220x_{13} + 170x_{14} + 140x_{21} + 130x_{22} + \\ & 190x_{23} + 150x_{24} + 190x_{31} + 200x_{32} + 230x_{33} \end{aligned} \quad (1)$$

约束条件有两类:一类是水库的供应量限制,另一类是各区的需求量限制.

由于供水量总能卖出并获利,水库的供应量限制可以表示为

$$x_{11} + x_{12} + x_{13} + x_{14} = 50 \quad (2)$$

$$x_{21} + x_{22} + x_{23} + x_{24} = 60 \quad (3)$$

$$x_{31} + x_{32} + x_{33} = 50 \quad (4)$$

考虑到各区的基本生活用水量与额外用水量,需求量限制可以表示为

$$30 \leq x_{11} + x_{21} + x_{31} \leq 80 \quad (5)$$

$$70 \leq x_{12} + x_{22} + x_{32} \leq 140 \quad (6)$$

$$10 \leq x_{13} + x_{23} + x_{33} \leq 30 \quad (7)$$

$$10 \leq x_{14} + x_{24} \leq 50 \quad (8)$$

模型求解

(1) ~ (8) 构成一个线性规划模型 (当然要加上 x_{ij} 的非负约束). 输入 LINGO:

```
! water supply; water shortage (这是说明语句);
min = 160 * x11 + 130 * x12 + 220 * x13 + 170 * x14
      + 140 * x21 + 130 * x22 + 190 * x23 + 150 * x24
      + 190 * x31 + 200 * x32 + 230 * x33;
x11 + x12 + x13 + x14 = 50;
x21 + x22 + x23 + x24 = 60;
x31 + x32 + x33 = 50;
x11 + x21 + x31 > = 30;
x11 + x21 + x31 < = 80;
x12 + x22 + x32 > = 70;
x12 + x22 + x32 < = 140
```

```

x13 + x23 + x33 > =10;
x13 + x23 + x33 < =30;
x14 + x24 > =10;
x14 + x24 < =50;
end

```

求解得到送水方案为(输出结果略):A 水库向乙区供水 50,B 水库向乙、丁区分别供水 50,10,C 水库向甲、丙分别供水 40,10.引水管理费为 24 400 元,利润为 $144\ 000 - 72\ 000 - 24\ 400 = 47\ 600$ 元.

讨论 如果 A,B,C 三个水库每天的最大供水量都提高一倍,则公司总供水能力为 320,大于总需求量 300,水库供水量不能全部卖出,因而不能像前面那样,将获利最多转化为引水管理费最少.此时我们首先需要计算 A,B,C 三个水库分别向甲、乙、丙、丁四个区供应每 10^3 t 水的净利润,即从收入 900 元中减去其他管理费 450 元,再减去表 1 中的引水管理费,得表 2.

表 2 从水库向各区送水的净利润

净利润	甲	乙	丙	丁
A	290	320	230	280
B	310	320	260	300
C	260	250	220	/

于是决策目标为

$$\max z = 290x_{11} + 320x_{12} + 230x_{13} + 280x_{14} + 310x_{21} + 320x_{22} + 260x_{23} + 300x_{24} + 260x_{31} + 250x_{32} + 220x_{33} \quad (9)$$

由于水库供水量不能全部卖出,所以上面约束(2)~(4)的右端增加一倍的同时,应将 $=$ 改成 \leq ,即

$$x_{11} + x_{12} + x_{13} + x_{14} \leq 100 \quad (10)$$

$$x_{21} + x_{22} + x_{23} + x_{24} \leq 120 \quad (11)$$

$$x_{31} + x_{32} + x_{33} \leq 100 \quad (12)$$

约束(5)~(8)不变.将(5)~(12)构成的线性规划模型输入 LINGO 求解得到送水方案为(详细程序和输出结果略):A 水库向乙区供水 100,B 水库向甲、乙、丁区分别供水 30、40、50,C 水库向甲、丙区分别供水 50、30,总利润为 88 700 元.

其实,由于每个区的供水量都能完全满足,所以上面(5)~(8)每个式子左边的约束可以去掉,右边的 \leq 可以改写成 $=$.作这样的简化后得到的解没有任何

变化。

评注 本题考虑的是将某种物质从若干供应点运往一些需求点,在供需量约束条件下使总费用最小或总利润最大。这类问题一般称为运输问题,是线性规划应用最广泛的领域之一。在标准的运输问题中,供需量通常是平衡的,即供应点的总供应量等于需求点的总需求量。本题中供需量不平衡,但这并不会引起本质的区别,同样可以方便地建立线性规划模型求解。

例2 货机装运

问题 某架货机有三个货舱:前仓、中仓、后仓。三个货舱所能装载的货物的最大质量和体积都有限制,如表3所示。并且为了保持飞机的平衡,三个货舱中实际装载货物的质量必须与其最大容许质量成比例。

表3 三个货舱装载货物的最大容许质量和体积

	前仓	中仓	后仓
质量限制/t	10	16	8
体积限制/ m^3	6 800	8 700	5 300

现有四类货物供该货机本次飞行装运,其有关信息如表4,最后一列指装运后所获得的利润。

表4 四类装运货物的信息

	质量/t	体积/ $(\text{m}^3 \cdot \text{t}^{-1})$	利润/ $(\text{元} \cdot \text{t}^{-1})$
货物1	18	480	3 100
货物2	15	650	3 800
货物3	23	580	3 500
货物4	12	390	2 850

应如何安排装运,使该货机本次飞行获利最大?

模型假设 问题中没有对货物装运提出其他要求,我们可作如下假设:

1. 每种货物可以分割到任意小。
2. 每种货物可以在一个或多个货舱中任意分布。
3. 多种货物可以混装,并保证不留空隙。
4. 所给出的数据都是精确的,没有误差。

模型建立

决策变量:用 x_{ij} 表示第 i 种货物装入第 j 个货舱的质量(t),货舱 $j=1,2,3$ 分别表示前仓、中仓、后仓。

已知参数:货舱 j 的质量限制 WET_j , 体积限制 VOL_j ; 第 i 种货物的质量 w_i , 单位质量的体积 v_i , 利润 p_i . 用行向量表示, 即 $WET = (10, 16, 8)$, $VOL = (6\ 800, 8\ 700, 5\ 300)$; $w = (18, 15, 23, 12)$, $v = (480, 650, 580, 390)$, $p = (3\ 100, 3\ 800, 3\ 500, 2\ 850)$.

决策目标是最大化总利润, 即

$$\max z = \sum_{i=1}^4 p_i \left(\sum_{j=1}^3 x_{ij} \right) \quad (13)$$

约束条件包括以下四个方面(除对 x_{ij} 的非负约束外):

1) 供装载的四种货物的总质量约束, 即

$$\sum_{j=1}^3 x_{ij} \leq w_i, i = 1, 2, 3, 4 \quad (14)$$

2) 三个货舱的质量限制, 即

$$\sum_{i=1}^4 x_{ij} \leq WET_j, j = 1, 2, 3 \quad (15)$$

3) 三个货舱的空间限制, 即

$$\sum_{i=1}^4 v_i x_{ij} \leq VOL_j, j = 1, 2, 3 \quad (16)$$

4) 三个货舱装入质量的平衡约束, 即

$$\sum_{i=1}^4 x_{ij} / WET_j = \sum_{i=1}^4 x_{ik} / WET_k, j, k = 1, 2, 3; j \neq k \quad (17)$$

模型求解

将以上模型输入 LINGO(注意这里通过集合定义变量, 程序简洁、清晰, 而且很容易推广):

! 定义集合及变量;

sets:

cang /1..3 /:WET,VOL;

wu /1..4 /:w,v,p;

link(wu,cang):x;

endsets

! 对已知变量赋值;

data:

WET=10,16,8;VOL=6800,8700,5300;

w=18,15,23,12;v=480,650,580,390;

p=3100,3800,3500,2850;

enddata

max=@sum(wu(i):p(i)*@sum(cang(j):x(i,j)));

```

@ for(wu(i):@ sum(cang(j):x(i,j)) < w(i));
@ for(cang(j):@ sum(wu(i):x(i,j)) < WET(j));
@ for(cang(j):@ sum(wu(i):v(i) * x(i,j)) < VOL(j));
@ for(cang(j):
@ for(cang(k) | k #GT# j: ! #GT#是大于等于的含义;
@ sum(wu(i):x(i,j)/WET(j)) = @ sum(wu(i):x(i,k)
/WET(k));
);
end

```

求解可以得到(以下略去了一些显然的结果):

Global optimal solution found.

Objective value: 121515.8

Total solver iterations: 12

Variable	Value	Reduced Cost
X(1,1)	0.000000	400.0000
X(1,2)	0.000000	57.89474
X(1,3)	0.000000	400.0000
X(2,1)	7.000000	0.000000
X(2,2)	0.000000	239.4737
X(2,3)	8.000000	0.000000
X(3,1)	3.000000	0.000000
X(3,2)	12.94737	0.000000
X(3,3)	0.000000	0.000000
X(4,1)	0.000000	650.0000
X(4,2)	3.052632	0.000000
X(4,3)	0.000000	650.0000

实际上,不妨将所得最优解作四舍五入,结果为货物 2 装入前仓 7 t、装入后仓 8 t;货物 3 装入前仓 3 t、装入中仓 13 t;货物 4 装入中仓 3 t. 最大利润约 121 516 元。(注意:这个问题的最优解并不惟一,但 LINGO 只能给出一个解)

评注 初步看来,本例与运输问题类似,似乎可以把四种货物看成四个供应点,三个货舱看成三个需求点(或者反过来,把货舱看成供应点,货物看成需求点)。但是,这里对供需量的限制包括两个方面:质量限制和空间限制,且有装载平衡要求。因此它只能看成是运输问题的一种变形和扩展。

4.3 汽车生产与原油采购

在 4.1 和 4.2 节的例题中研究的对象都是连续可分的,于是决策变量是连续的,建立的模型是线性规划.在本节的例子中将会遇到不同的情况.

例 1 汽车厂生产计划

问题 一汽车厂生产小、中、大三种类型的汽车,已知各类型每辆车对钢材、劳动时间的需求,利润以及每月工厂钢材、劳动时间的现有量如表 1 所示.试制订月生产计划,使工厂的利润最大.

进一步讨论:由于各种条件限制,如果生产某一类型汽车,则至少要生产 80 辆,那么最优的生产计划应作何改变?

表 1 汽车厂的生产数据

	小型	中型	大型	现有量
钢材/t	1.5	3	5	600
劳动时间/h	280	250	400	60 000
利润/万元	2	3	4	

模型建立与求解

设每月生产小、中、大型汽车的数量分别为 x_1, x_2, x_3 , 工厂的月利润为 z , 在题目所给参数均不随生产数量变化的假设下,立即可得线性规划模型:

$$\max z = 2x_1 + 3x_2 + 4x_3 \quad (1)$$

$$\text{s. t. } 1.5x_1 + 3x_2 + 5x_3 \leq 600 \quad (2)$$

$$280x_1 + 250x_2 + 400x_3 \leq 60\,000 \quad (3)$$

$$x_1, x_2, x_3 \geq 0 \quad (4)$$

用 LINGO 求解,可得最优解 $x_1 = 64.516\,129, x_2 = 167.741\,928, x_3 = 0$, 出现小数,显然不合适.通常的解决办法有以下几种:

(1) 简单地舍去小数,取 $x_1 = 64, x_2 = 167$, 它会接近最优的整数解,可算出相应的目标函数值 $z = 629$, 与 LP 得到的最优值 $z = 632.258\,1$ 相差不大.

(2) 在上面这个解的附近试探:如取 $x_1 = 65, x_2 = 167; x_1 = 64, x_2 = 168$ 等. 因为从输出可知,约束都是紧的,所以若试探的 x_1, x_2 大于 LP 最优解时,必须检验它们是否满足约束条件(2),(3),然后计算函数值 z ,通过比较可能得到更优解.

(3) 在线性规划模型中增加约束条件:

$$x_1, x_2, x_3 \text{ 均为整数} \quad (5)$$

这样得到的(1)~(5)式称为**整数规划**(Integer Programming, 简记作 IP), IP 可以用 LINGO 直接求解, 输入文件:

```
max = 2 * x1 + 3 * x2 + 4 * x3;
1.5 * x1 + 3 * x2 + 5 * x3 < 600;
280 * x1 + 250 * x2 + 400 * x3 < 60000;
@gin(x1); @gin(x2); @gin(x3);
```

注:最后一行中的“@gin”是将变量限定为整数的函数。

求解得到输出(只列出需要的结果):

Global optimal solution found.

Objective value: 632.0000

Extended solver steps: 0

Total solver iterations: 3

Variable	Value	Reduced Cost
X1	64.00000	-2.000000
X2	168.0000	-3.000000
X3	0.000000	-4.000000

IP 的最优解 $x_1 = 64, x_2 = 168, x_3 = 0$, 最优值 $z = 632$, 即问题要求的月生产计划为生产小型车 64 辆、中型车 168 辆, 不生产大型车。

讨论 对于问题中提出的“如果生产某一类型汽车, 则至少要生产 80 辆”的限制, 上面得到的 IP 的最优解不满足这个条件。这种类型的要求是实际生产中经常提出的。下面以本问题为例说明解决这类要求的办法。

对于原 LP 模型(1)~(4), 需将(4)式改为

$$x_1, x_2, x_3 = 0 \text{ 或 } \geq 80 \quad (6)$$

下面是求解模型(1)~(4), (6)的 3 种方法:

1) 分解为多个 LP 子模型

(6)式可分解为 8 种情况:

$$x_1 = 0, x_2 = 0, x_3 \geq 80 \quad (6-1)$$

$$x_1 = 0, x_2 \geq 80, x_3 = 0 \quad (6-2)$$

$$x_1 = 0, x_2 \geq 80, x_3 \geq 80 \quad (6-3)$$

$$x_1 \geq 80, x_2 = 0, x_3 = 0 \quad (6-4)$$

$$x_1 \geq 80, x_2 \geq 80, x_3 = 0 \quad (6-5)$$

$$x_1 \geq 80, x_2 = 0, x_3 \geq 80 \quad (6-6)$$

$$x_1 \geq 80, x_2 \geq 80, x_3 \geq 80 \quad (6-7)$$

$$x_1, x_2, x_3 = 0 \quad (6-8)$$

(6-8)显然不可能是问题的解。可以检查, (6-3)和(6-7)不满足约束条件

(2),也不可能是问题的解.对其他5个LP子模型逐一求解,比较目标函数值,可知最优解在(6-5)情形得到: $x_1=80, x_2=150.399\ 994, x_3=0$,最优值 $z=611.2$.若加上对 x_1, x_2, x_3 的整数约束,可得: $x_1=80, x_2=150, x_3=0$,最优值 $z=610$.

注:可以不检查是否满足约束条件,解所有LP子模型,结果同上.

2) 引入0-1变量,化为整数规划

设 y_1 只取0,1两个值,则“ $x_1=0$ 或 ≥ 80 ”等价于

$$80y_1 \leq x_1 \leq My_1, y_1 \in \{0,1\} \quad (7-1)$$

其中 M 为相当大的正数,本例可取1000(x_1 不可能超过1000).类似地有

$$80y_2 \leq x_2 \leq My_2, y_2 \in \{0,1\} \quad (7-2)$$

$$80y_3 \leq x_3 \leq My_3, y_3 \in \{0,1\} \quad (7-3)$$

于是(1)~(3),(5),(7-1)~(7-3)构成一个特殊的整数规划模型(既有一般的整数变量,又有0-1变量),用LINGO直接求解时,输入的最后要加上0-1变量的限定语句:

@bin(y1);@bin(y2);@bin(y3);

求解可得到与第1种方法同样的结果.

3) 化为非线性规划

条件(4),(6)可表示为

$$x_1(x_1-80) \geq 0 \quad (8-1)$$

$$x_2(x_2-80) \geq 0 \quad (8-2)$$

$$x_3(x_3-80) \geq 0 \quad (8-3)$$

式子左端是决策变量的非线性函数,(1)~(4),(8-1)~(8-3)构成非线性规划(Non-Linear Programming,简记作NLP).该模型可如下输入LINGO:

```
max = 2 * x1 + 3 * x2 + 4 * x3;
1.5 * x1 + 3 * x2 + 5 * x3 < 600;
280 * x1 + 250 * x2 + 400 * x3 < 60000;
x1 * (x1 - 80) > 0;
x2 * (x2 - 80) > 0;
x3 * (x3 - 80) > 0;
@gin(x1);@gin(x2);@gin(x3);
```

求解可得到与第2种方法同样的结果.

评注 像汽车这样的对象自然是整数变量,应该建立整数规划模型,但是求解整数规划比线性规划要难得多(即使使用数学软件),所以当整数变量取值很大时,常作为连续变量用线性规划处理.

一般来说,非线性规划的求解比线性规划困难得多,特别是问题规模较大或者要求得到全局最优解时更是如此.为了考虑(6)式这样的条件,通常是引入

0-1变量,建立整数规划模型,而一般尽量不用非线性规划。

例2 原油采购与加工

问题 某公司用两种原油(A和B)混合加工成两种汽油(甲和乙)。甲、乙两种汽油含原油A的最低比例分别为50%和60%,售价分别为4 800元/t和5 600元/t。该公司现有原油A和B的库存量分别为500 t和1 000 t,还可以从市场上买到不超过1 500 t的原油A。原油A的市场价为:购买量不超过500 t时的单价为10 000元/t;购买量超过500 t但不超过1 000 t时,超过500 t的部分8 000元/t;购买量超过1 000 t时,超过1 000 t的部分6 000元/t。该公司应如何安排原油的采购和加工?^[72]

问题分析 安排原油采购、加工的目标只能是利润最大,题目中给出的是两种汽油的售价和原油A的采购价,利润为销售汽油的收入与购买原油A的支出之差。这里的难点在于原油A的采购价与购买量的关系比较复杂,是分段函数关系,能否以及如何用线性规划、整数规划模型加以处理是关键所在。

模型建立

设原油A的购买量为 x ,根据题目所给数据,采购的支出 $c(x)$ 可表为如下的分段线性函数(以下价格以千元/t为单位):

$$c(x) = \begin{cases} 10x & (0 \leq x \leq 500) \\ 1\,000 + 8x & (500 \leq x \leq 1\,000) \\ 3\,000 + 6x & (1\,000 \leq x \leq 1\,500) \end{cases} \quad (9)$$

设原油A用于生产甲、乙两种汽油的数量分别为 x_{11} 和 x_{12} ,原油B用于生产甲、乙两种汽油的数量分别为 x_{21} 和 x_{22} ,则总的收入为 $4.8(x_{11} + x_{21}) + 5.6(x_{12} + x_{22})$ 。于是本例的目标函数——利润为

$$\max z = 4.8(x_{11} + x_{21}) + 5.6(x_{12} + x_{22}) - c(x) \quad (10)$$

约束条件包括加工两种汽油用的原油A、原油B库存量的限制,和原油A购买量的限制,以及两种汽油含原油A的比例限制,分别表示为

$$x_{11} + x_{12} \leq 500 + x \quad (11)$$

$$x_{21} + x_{22} \leq 1\,000 \quad (12)$$

$$x \leq 1\,500 \quad (13)$$

$$\frac{x_{11}}{x_{11} + x_{21}} \geq 0.5 \quad (14)$$

$$\frac{x_{12}}{x_{12} + x_{22}} \geq 0.6 \quad (15)$$

$$x_{11}, x_{12}, x_{21}, x_{22}, x \geq 0 \quad (16)$$

由于(9)式中的 $c(x)$ 不是线性函数,(9)~(16)给出的是一个非线性规划。而且,对于这样用分段函数定义的 $c(x)$,一般的非线性规划软件也难以输入和

求解. 能不能想办法将该模型化简, 从而用现成的软件求解呢?

模型求解 下面介绍 3 种解法.

第 1 种解法 一个自然的想法是将原油 A 的采购量 x 分解为三个量, 即用 x_1, x_2, x_3 分别表示以价格 10 千元/t, 8 千元/t, 6 千元/t 采购的原油 A 的数量, 总支出为 $c(x) = 10x_1 + 8x_2 + 6x_3$, 且

$$x = x_1 + x_2 + x_3 \quad (17)$$

这时目标函数(10)变为线性函数:

$$\max z = 4.8(x_{11} + x_{21}) + 5.6(x_{12} + x_{22}) - (10x_1 + 8x_2 + 6x_3) \quad (18)$$

应该注意到, 只有当以 10 千元/t 的价格购买 $x_1 = 500$ t 时, 才能以 8 千元/t 的价格购买 $x_2 (x_2 > 0)$, 这个条件可以表示为

$$(x_1 - 500)x_2 = 0 \quad (19)$$

同理, 只有当以 8 千元/t 的价格购买 $x_2 = 500$ t 时, 才能以 6 千元/t 的价格购买 $x_3 (x_3 > 0)$, 于是

$$(x_2 - 500)x_3 = 0 \quad (20)$$

此外, x_1, x_2, x_3 的取值范围是

$$0 \leq x_1, x_2, x_3 \leq 500 \quad (21)$$

由于有非线性约束(19)和(20), (11) ~ (21) 构成非线性规划模型. 将该模型输入 LINGO 软件如下:

```
model:
max = 4.8 * x11 + 4.8 * x21 + 5.6 * x12 + 5.6 * x22 - 10 * x1 - 8 *
x2 - 6 * x3;
x11 + x12 <= x + 500;
x21 + x22 < 1000;
0.5 * x11 - 0.5 * x21 > 0;
0.4 * x12 - 0.6 * x22 > 0;
x = x1 + x2 + x3;
(x1 - 500) * x2 = 0;
(x2 - 500) * x3 = 0;
x1 < 500;
x2 < 500;
x3 < 500;
end
```

注: 因为(17)式和(21)式保证了(13)式, 所以上面输入中不需要(13)式.

将文件存储并命名后, 选择菜单“LINGO|Solve”, 运行该程序得到:

Local optimal solution found.

Objective value: 4 800.000

Total solver iterations: 14

Variable	Value	Reduced Cost
X11	500.0000	0.000000
X21	500.0000	0.000000
X12	0.000000	0.2666667
X22	0.000000	0.000000
X1	0.000000	0.4000000
X2	0.000000	0.000000
X3	0.000000	0.000000
X	0.000000	0.000000

最优解是用库存的 500 t 原油 A、500 t 原油 B 生产 1 000 t 汽油甲,不购买新的原油 A,利润为 4 800 000 元。

但是 LINGO 得到的结果只是一个局部最优解 (Local Optimal Solution),还能得到更好的解吗? 除线性规划外, LINGO 在缺省设置下一般只给出局部最优解,但可以通过修改 LINGO 选项要求计算全局最优解。具体作法是:选择“LINGO|Options”菜单,在弹出的选项卡中选择“General Solver”,然后找到选项“Use Global Solver”将其选中,并应用或保存设置。重新运行“LINGO|Solve”,可得到如下输出:

Global optimal solution found.

Objective value: 5000.000

Extended solver steps: 1

Total solver iterations: 43

Variable	Value	Reduced Cost
X11	0.000000	0.000000
X21	0.000000	0.9000000
X12	1500.000	0.000000
X22	1000.000	0.000000
X1	500.0000	0.000000
X2	500.0000	0.000000
X3	0.000000	0.000000
X	1000.000	0.000000

全局最优解是购买 1 000 t 原油 A,与库存的 500 t 原油 A 和 1 000 t 原油 B 一起,共生产 2 500 t 汽油乙,利润为 5 000 000 元,高于局部最优解对应的利润。

第 2 种解法 引入 0-1 变量将(19)和(20)转化为线性约束。

令 $y_1 = 1, y_2 = 1, y_3 = 1$ 分别表示以 10 千元/t、8 千元/t、6 千元/t 的价格采购原油 A, 则约束 (19) 和 (20) 可以替换为

$$500y_2 \leq x_1 \leq 500y_1 \quad (22)$$

$$500y_3 \leq x_2 \leq 500y_2 \quad (23)$$

$$x_3 \leq 500y_3 \quad (24)$$

$$y_1, y_2, y_3 = 0 \text{ 或 } 1 \quad (25)$$

(11) ~ (18), (21) ~ (25) 构成整数 (线性) 规划模型, 将它输入 LINGO 软件如下:

```
model;
max = 4.8 * x11 + 4.8 * x21 + 5.6 * x12 + 5.6 * x22 - 10 * x1 - 8 *
x2 - 6 * x3;
x11 + x12 <= x + 500;
x21 + x22 <= 1000;
0.5 * x11 - 0.5 * x21 > 0;
0.4 * x12 - 0.6 * x22 > 0;
x = x1 + x2 + x3;
x1 <= 500 * y1;
x2 <= 500 * y2;
x3 <= 500 * y3;
x1 >= 500 * y2;
x2 >= 500 * y3;
@ bin(y1); @ bin(y2); @ bin(y3);
end
```

运行该程序得到的最优解与第 1 种解法得到的结果 (全局最优解) 相同.

第 3 种解法 直接处理分段线性函数 $c(x)$. (9) 式表示的 $c(x)$ 如图 1.

记 x 轴上的分点为 $b_1 = 0, b_2 = 500, b_3 = 1\,000, b_4 = 1\,500$. 当 x 在第 1 个小区间 $[b_1, b_2]$ 时, 记 $x = z_1 b_1 + z_2 b_2, z_1 + z_2 = 1, z_1, z_2 \geq 0$, 因为 $c(x)$ 在 $[b_1, b_2]$ 是线性的, 所以 $c(x) = z_1 c(b_1) + z_2 c(b_2)$. 同样, 当 x 在第 2 个小区间 $[b_2, b_3]$ 时, $x = z_2 b_2 + z_3 b_3, z_2 + z_3 = 1, z_2, z_3 \geq 0, c(x) = z_2 c(b_2) + z_3 c(b_3)$. 当 x 在第 3 个小区间 $[b_3, b_4]$ 时, $x = z_3 b_3 + z_4 b_4, z_3 + z_4 = 1, z_3, z_4 \geq 0, c(x) = z_3 c(b_3) +$

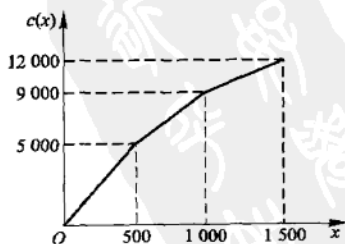


图 1 分段线性函数 $c(x)$ 图形

$z_4 c(b_4)$.

为了表示 x 在哪个小区间, 引入 0-1 变量 $y_k (k=1, 2, 3)$, 当 x 在第 k 个小区间时, $y_k=1$, 否则, $y_k=0$. 这样, $z_1, z_2, z_3, z_4, y_1, y_2, y_3$ 应满足

$$z_1 \leq y_1, z_2 \leq y_1 + y_2, z_3 \leq y_2 + y_3, z_4 \leq y_3 \quad (26)$$

$$z_1 + z_2 + z_3 + z_4 = 1, z_k \geq 0 (k=1, 2, 3, 4) \quad (27)$$

$$y_1 + y_2 + y_3 = 1, y_1, y_2, y_3 = 0 \text{ 或 } 1 \quad (28)$$

此时 x 和 $c(x)$ 可以统一地表示为

$$x = z_1 b_1 + z_2 b_2 + z_3 b_3 + z_4 b_4 = 500z_2 + 1\,000z_3 + 1\,500z_4 \quad (29)$$

$$\begin{aligned} c(x) &= z_1 c(b_1) + z_2 c(b_2) + z_3 c(b_3) + z_4 c(b_4) \\ &= 5\,000z_2 + 9\,000z_3 + 12\,000z_4 \end{aligned} \quad (30)$$

(10) ~ (16), (26) ~ (30) 也构成一个整数规划模型, 将它输入 LINGO 软件求解, 得到的结果与第 2 种解法相同.

评注 这个问题的关键是处理分段线性函数, 我们推荐化为整数规划模型的第 2, 3 种解法, 第 3 种解法更具一般性, 其作法如下:

设一个 n 段线性函数 $f(x)$ 的分点为 $b_1 \leq \dots \leq b_n \leq b_{n+1}$, 引入 z_k , 将 x 和 $f(x)$ 表示为

$$x = \sum_{k=1}^{n+1} z_k b_k \quad (31)$$

$$f(x) = \sum_{k=1}^{n+1} z_k f(b_k) \quad (32)$$

z_k 和 0-1 变量 y_k 满足

$$z_1 \leq y_1, z_2 \leq y_1 + y_2, \dots, z_n \leq y_{n-1} + y_n, z_{n+1} \leq y_n \quad (33)$$

$$y_1 + y_2 + \dots + y_n = 1, y_k = 0 \text{ 或 } 1 \quad (34)$$

$$z_1 + z_2 + \dots + z_{n+1} = 1, z_k \geq 0 (k=1, 2, \dots, n+1) \quad (35)$$

4.4 接力队的选拔与选课策略

实际生活中可能遇到这样的分派问题: 若干项任务分给一些候选人来完成, 因为每个人的专长不同, 他们完成每项任务取得的效益或需要的资源就不一样, 如何分派这些任务使获得的总效益最大, 或付出的总资源最少? 也会遇到这样的选择问题: 有若干种策略供你选择, 不同的策略得到的收益或付出的成本不同, 各个策略之间可以有相互制约关系, 如何在满足一定条件下作出抉择, 使得收益最大或成本最小? 本节将通过几个实例说明怎样用数学规划模型解决这种问题.

例 1 混合泳接力队的选拔

问题 某班准备从 5 名游泳队员中选择 4 人组成接力队, 参加学校的 4 ×

100 m 混合泳接力比赛. 5 名队员 4 种泳姿的百米平均成绩如表 1 所示, 问应如何选拔队员组成接力队?

如果最近队员丁的蛙泳成绩有较大退步, 只有 1'15"2; 而队员戊经过艰苦训练自由泳成绩有所进步, 达到 57"5, 组成接力队的方案是否应该调整?

表 1 5 名队员 4 种泳姿的百米平均成绩

	甲	乙	丙	丁	戊
蝶泳	1'06"8	57"2	1'18"	1'10"	1'07"4
仰泳	1'15"6	1'06"	1'07"8	1'14"2	1'11"
蛙泳	1'27"	1'06"4	1'24"6	1'09"6	1'23"8
自由泳	58"6	53"	59"4	57"2	1'02"4

问题分析 从 5 名队员中选出 4 人组成接力队, 每人一种泳姿, 且 4 人的泳姿各不相同, 使接力队的成绩最好. 容易想到的一个办法是穷举法, 组成接力队的方案共有 $5! = 120$ 种, 逐一计算并作比较, 即可找出最优方案. 显然这不是解决这类问题的好办法, 随着问题规模的变大, 穷举法的计算量将是无法接受的.

可以用 0-1 变量表示一个队员是否入选接力队, 从而建立这个问题的 0-1 规划模型, 借助现成的数学软件求解.

模型的建立与求解

记甲乙丙丁戊分别为队员 $i = 1, 2, 3, 4, 5$; 记蝶泳、仰泳、蛙泳、自由泳分别为泳姿 $j = 1, 2, 3, 4$. 记队员 i 的第 j 种泳姿的百米最好成绩为 c_{ij} (s), 即有

c_{ij}	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$
$j = 1$	66.8	57.2	78	70	67.4
$j = 2$	75.6	66	67.8	74.2	71
$j = 3$	87	66.4	84.6	69.6	83.8
$j = 4$	58.6	53	59.4	57.2	62.4

引入 0-1 变量 x_{ij} , 若选择队员 i 参加泳姿 j 的比赛, 记 $x_{ij} = 1$, 否则记 $x_{ij} = 0$. 根据组成接力队的要求, x_{ij} 应该满足两个约束条件:

第一, 每人最多只能入选 4 种泳姿之一, 即对于 $i = 1, 2, 3, 4, 5$, 应有

$$\sum_{j=1}^4 x_{ij} \leq 1;$$

第二, 每种泳姿必须有 1 人而且只能有 1 人入选, 即对于 $j = 1, 2, 3, 4$, 应有

$$\sum_{i=1}^5 x_{ij} = 1.$$

当队员 i 入选泳姿 j 时, $c_{ij}x_{ij}$ 表示他(她)的成绩, 否则 $c_{ij}x_{ij} = 0$. 于是接力队的成绩可表示为 $z = \sum_{j=1}^4 \sum_{i=1}^5 c_{ij}x_{ij}$, 这就是该问题的目标函数.

综上, 这个问题的 0-1 规划模型可写作

$$\min z = \sum_{j=1}^4 \sum_{i=1}^5 c_{ij}x_{ij} \quad (1)$$

$$\text{s. t. } \sum_{j=1}^4 x_{ij} \leq 1, i = 1, 2, 3, 4, 5 \quad (2)$$

$$\sum_{i=1}^5 x_{ij} = 1, j = 1, 2, 3, 4 \quad (3)$$

$$x_{ij} = \{0, 1\} \quad (4)$$

利用题目所给数据, 将这一模型输入 LINGO:

model:

sets:

person/1..5/;

position/1..4/;

link(person, position):c,x;

endsets

data:

c = 66.8, 75.6, 87.5, 58.6,

57.2, 66.6, 66.4, 53,

78.6, 78.8, 84.6, 59.4

70, 74.2, 69.6, 57.2,

67.4, 71, 83.8, 62.4;

enddata

min=@sum(link:c*x);

@for(person(i):@sum(position(j):x(i,j))<=1);

@for(position(i):@sum(person(j):x(j,i))=1);

@for(link:@bin(x));

end

求解得到结果为: $x_{14} = x_{21} = x_{32} = x_{43} = 1$, 其他变量为 0, 成绩为 253.2 s = 4'13"2. 即应当选派甲乙丙丁 4 人组成接力队, 分别参加自由泳、蝶泳、仰泳、蛙泳的比赛.

讨论 考虑到丁、戊最近的状态, c_{43} 由原来的 69.6 s 变为 75.2 s, c_{54} 由原来的 62.4 s 变为 57.5 s, 讨论对结果的影响. 这类似于 4.1 节中的敏感性分析, 但是对于整数规划模型, 一般没有与线性规划相类似的理论, 此时 LINGO 中所输出的敏感性分析结果通常是没有意义的. 于是我们只好用 c_{43} , c_{54} 的新数据重新输入模型, 用 LINGO 求解得到: $x_{21} = x_{32} = x_{43} = x_{54} = 1$, 其他变量为 0, 成绩为 257.7 s = 4'17"7. 即应当选派乙丙丁戊 4 人组成接力队, 分别参加蝶泳、仰泳、蛙泳、自由泳的比赛.

评注 本例属于这样一类分派问题: 有若干项任务, 每项任务必须有一人且只能有一人承担, 每人也只能承担其中一项, 不同人员承担不同任务的收益 (或成本) 不同, 问题是怎样分派各项任务使总收益最大 (或总成本最小). 它又称为指派问题 (Assignment). 建立 0-1 规划模型是解决这类问题的常用方法. 典型的指派问题中, 任务的数量与能够承担的人员数量相等, 但是二者不相等的情况也常见, 本例是人数多于任务数, 如果任务数多于人数呢 (这时虽然不是上述意义下的分派问题, 但能建立类似的模型)?

例 2 选课策略

某学校规定, 运筹学专业的学生毕业时必须至少学习过两门数学课、三门运筹学课和两门计算机课. 这些课程的编号、名称、学分、所属类别和先修课要求如表 2 所示. 那么, 毕业时学生最少可以学习这些课程中的哪些课程?

如果某个学生既希望选修课程的数量少, 又希望所获得的学分多, 他可以选修哪些课程?

表 2 课程情况

课程编号	课程名称	学分	所属类别	先修课要求
1	微积分	5	数学	
2	线性代数	4	数学	
3	最优化方法	4	数学; 运筹学	微积分; 线性代数
4	数据结构	3	数学; 计算机	计算机编程
5	应用统计	4	数学; 运筹学	微积分; 线性代数
6	计算机模拟	3	计算机; 运筹学	计算机编程
7	计算机编程	2	计算机	
8	预测理论	2	运筹学	应用统计
9	数学实验	3	运筹学; 计算机	微积分; 线性代数

模型的建立与求解

用 $x_i = 1$ 表示选修表 2 中按编号顺序的 9 门课程 ($x_i = 0$ 表示不选; $i = 1, 2, \dots, 9$). 问题的目标为选修的课程总数最少, 即

$$\min z = \sum_{i=1}^9 x_i \quad (5)$$

约束条件包括两个方面:

第一, 每人最少要学习 2 门数学课、3 门运筹学课和 2 门计算机课. 根据表中对每门课程所属类别的划分, 这一约束可以表示为

$$x_1 + x_2 + x_3 + x_4 + x_5 \geq 2 \quad (6)$$

$$x_3 + x_5 + x_6 + x_8 + x_9 \geq 3 \quad (7)$$

$$x_4 + x_6 + x_7 + x_9 \geq 2 \quad (8)$$

第二, 某些课程有先修课程的要求. 例如“数据结构”的先修课是“计算机编程”, 这意味着如果 $x_4 = 1$, 必须 $x_7 = 1$, 这个条件可以表示为 $x_4 \leq x_7$ (注意: $x_4 = 0$ 时对 x_7 没有限制). “最优化方法”的先修课是“微积分”和“线性代数”的条件可表为 $x_3 \leq x_1, x_3 \leq x_2$, 而这两个不等式可以用一个约束表示为 $2x_3 - x_1 - x_2 \leq 0$. 这样, 所有课程的先修课要求可表为如下的约束:

$$2x_3 - x_1 - x_2 \leq 0 \quad (9)$$

$$x_4 - x_7 \leq 0 \quad (10)$$

$$2x_5 - x_1 - x_2 \leq 0 \quad (11)$$

$$x_6 - x_7 \leq 0 \quad (12)$$

$$x_8 - x_5 \leq 0 \quad (13)$$

$$2x_9 - x_1 - x_2 \leq 0 \quad (14)$$

由上得到以 (5) 为目标函数、以 (6) ~ (14) 为约束条件的 0-1 规划模型. 将这一模型输入 LINGO (注意加上 x_i 为 0-1 的约束), 求解得到结果为 $x_1 = x_2 = x_3 = x_6 = x_7 = x_9 = 1$, 其他变量为 0. 对照课程编号, 它们是微积分、线性代数、最优化方法、计算机模拟、计算机编程、数学实验, 共 6 门课程, 总学分为 21.

下面将会看到, 这个解并不是惟一的, 还可以找到与以上不完全相同的 6 门课程, 也满足所给的约束.

讨论 如果一个学生既希望选修课程数少, 又希望所获得的学分数尽可能多, 则除了目标 (5) 之外, 还应根据表 2 中的学分数写出另一个目标, 即

$$\max w = 5x_1 + 4x_2 + 4x_3 + 3x_4 + 4x_5 + 3x_6 + 2x_7 + 2x_8 + 3x_9 \quad (15)$$

我们把只有一个优化目标的规划问题称为单目标规划, 而将多于一个目标的规划问题称为多目标规划. 多目标规划的目标函数相当于一个向量, 如目标 (5) 和 (15) 可以表示为对一个向量进行优化:

$$V = \min(z, -w) \quad (16)$$

上面符号“V-min”是“向量最小化”的意思,注意其中已经通过对 w 取负号而将(15)中的最大化变成了最小化问题。

要得到多目标规划问题的解,通常需要知道决策者对每个目标的重视程度,称为偏好程度。下面通过几个例子讨论处理这类问题的方法。

1. 同学甲只考虑获得尽可能多的学分,而不管所修课程的多少,那么他可以以(15)为目标,不用考虑(5),这就变成了一个单目标优化问题。显然,这个问题不必计算就知道最优解是选修所有9门课程。

2. 同学乙认为选修课程数最少是基本的前提,那么他可以只考虑目标(5)而不管(15),这就是前面得到的,最少为6门。如果这个解是惟一的,则他已别无选择,只能选修上面的6门课,总学分为21。但是LINGO无法告诉我们一个优化问题的解是否惟一,所以他还可能在选修6门课的条件下,使总学分多于21。为探索这种可能,应在上面的规划问题中增加约束

$$\sum_{i=1}^9 x_i = 6 \quad (17)$$

得到以(15)为目标函数、以(6)~(14)和(17)为约束条件的另一个0-1规划模型。求解后会得到不同于前面6门课程的最优解 $x_1 = x_2 = x_3 = x_5 = x_7 = x_9 = 1$,其他变量为0,即3学分的“计算机模拟”换成了4学分的“应用统计”,总学分由21增至22。注意这个模型的解仍然不是惟一的,如 $x_1 = x_2 = x_3 = x_5 = x_6 = x_7 = 1$,其他变量为0,也是最优解。

3. 同学丙不像甲、乙那样,只考虑学分最多或以课程最少为前提,而是觉得学分数和课程数这两个目标大致上应该三七开。这时可以将目标函数 z 和 $-w$ 分别乘以0.7和0.3,组成一个新的目标函数 y ,有

$$\begin{aligned} \min y &= 0.7z - 0.3w \\ &= -0.8x_1 - 0.5x_2 - 0.5x_3 - 0.2x_4 - 0.5x_5 - 0.2x_6 + 0.1x_7 + 0.1x_8 - 0.2x_9 \end{aligned} \quad (18)$$

得到以(18)为目标、以(6)~(14)为约束的0-1规划模型。输入LINGO求解得到结果为: $x_1 = x_2 = x_3 = x_4 = x_5 = x_6 = x_7 = x_9 = 1$,即只有“预测理论”不需要选修,共28学分。

实际上,0.7和0.3是 z 和 $-w$ 的权重。一般地,将权重记作 λ_1, λ_2 ,且令 $\lambda_1 + \lambda_2 = 1, 0 \leq \lambda_1, \lambda_2 \leq 1$,则0-1规划模型的新目标为

$$\min y = \lambda_1 z - \lambda_2 w \quad (19)$$

前面同学甲的考虑相当于 $\lambda_1 = 0, \lambda_2 = 1$,同学乙的考虑相当于 $\lambda_1 = 1, \lambda_2 = 0$,是两种极端情况。通过选取许多不同的 λ_1, λ_2 进行计算,可以发现当 $\lambda_1 < 2/3$ 时,结果与同学甲相同;而当 $\lambda_1 > 3/4$ 时,结果与同学乙相同。这是偶然的吗?我们根据给出的数据分析一下。

当 $\lambda_1 < 2/3$ 时,(19)式中 y 的所有系数都小于0,因此为了使 y 取最小值, x_i ,

x_6, x_7, x_8, x_9 应尽可能取 1, 这与 $\lambda_1 = 0, \lambda_2 = 1$ 的情况, 即学分数最多是一样的。

当 $\lambda_1 > 3/4$ 时, (19) 式中 y 的系数中至少有 5 个大于 0, 它们分别是 x_4, x_6, x_7, x_8, x_9 的系数, 因此为了使 y 取最小值, x_4, x_6, x_7, x_8, x_9 应尽可能取 0, 而根据前面的计算知道约束条件已经保证至少要选修 6 门课, 所以 x_4, x_6, x_7, x_8, x_9 中最多只能有 3 个同时取 0, 这与 $\lambda_1 = 1, \lambda_2 = 0$ 的情况, 即选修的课程数最少是一样的。

评注 用 0-1 变量表示选择策略是常用的方法, 而对于“要选甲必选乙”这样的约束, 可以用类似于 (10) 式 $x_4 \leq x_7$ 来描述。有些选择问题, 如从众多球员中选拔上场队员时, 由于相互配合或相互制约的关系, 还会遇到诸如“甲乙二人至多选一人”、“甲乙二人至少选一人”、“要选甲必不能选乙”等约束。

本节还讨论了多目标规划问题的处理办法, 其基本思想是通过加权组合形成一个新的目标, 从而化为单目标规划。优先考虑一个目标不过是这种办法的极端情况, 而像前面同学乙那样, 把一个目标作为约束条件 (17), 解另一个目标的规划模型, 也是处理多目标规划的方法。

例 3 销售代理的开发与中断

问题 某公司正在考虑在某城市开发一些销售代理业务。经过预测, 该公司已经确定了该城市未来 5 年的业务量, 分别为 400, 500, 600, 700 和 800。该公司已经初步物色了 4 家销售公司作为其代理候选企业, 表 3 给出了该公司与每个候选企业建立代理关系的一次性费用, 以及每个候选企业每年所能承揽的最大业务量和年运行费用。该公司应该与哪些候选企业建立代理关系^[72]。

表 3 代理情况

	候选代理 1	候选代理 2	候选代理 3	候选代理 4
年最大业务量	350	250	300	200
一次性费用/万元	100	80	90	70
年运行费用/万元	7.5	4.0	6.5	3.0

如果该公司目前已经与上述 4 个代理建立了代理关系并且都处于运行状态, 但每年初可以决定临时中断或重新恢复代理关系, 每次临时中断或重新恢复代理关系的费用如表 4 所示。该公司应如何对这些代理进行业务调整?

表 4 代理调整费用

	代理 1	代理 2	代理 3	代理 4
临时中断费用/万元	5	3	4	2
重新恢复费用/万元	5	4	1	9

模型的建立与求解

首先考虑本题前半部分建立代理关系的决策问题.

决策变量

题中没有要求必须从第一年开始建立代理关系,因此我们假定公司可以从未来5年中的任意一年开始与某些候选代理建立代理关系.用 $x_{it}=1$ ($i=1,2,3,4; t=1,2,3,4,5$) 分别表示在第 t 年初(首次)与候选代理 i 建立代理关系(0 则表示不建立这一关系).

决策目标

问题中没有明确说明建立代理关系需要考虑的全部因素,只给出了费用,因此可以假设5年的总费用最小是需要考虑的惟一目标.总费用由建立代理关系的一次性费用和每年的运行费用组成,其中建立代理关系的一次性费用为

$$z_1 = 100 \sum_{i=1}^5 x_{1i} + 80 \sum_{i=1}^5 x_{2i} + 90 \sum_{i=1}^5 x_{3i} + 70 \sum_{i=1}^5 x_{4i} \quad (20)$$

问题中没有说明是否可以临时中断代理关系,因此我们假定公司一旦与候选代理建立代理关系,则这一关系将长期保持,于是有

$$\sum_{i=1}^5 x_{it} \leq 1, \quad i=1,2,3,4 \quad (21)$$

根据 x_{it} 的定义,对候选代理 i 来说, $x_{i1} + \dots + x_{i5}$ 表示的是第 t 年时公司是否与该候选代理已经建立了代理关系.例如,对候选代理1来说,5年的总运行费用为

$$\begin{aligned} & 7.5[x_{11} + (x_{11} + x_{12}) + (x_{11} + x_{12} + x_{13}) + (x_{11} + x_{12} + x_{13} + x_{14}) + \\ & (x_{11} + x_{12} + x_{13} + x_{14} + x_{15})] \\ & = 7.5(5x_{11} + 4x_{12} + 3x_{13} + 2x_{14} + x_{15}) \end{aligned}$$

于是,对所有候选代理来说,5年的总运行费用为

$$\begin{aligned} z_2 = & 7.5(5x_{11} + 4x_{12} + 3x_{13} + 2x_{14} + x_{15}) + 4.0(5x_{21} + 4x_{22} + 3x_{23} + 2x_{24} + x_{25}) \\ & + 6.5(5x_{31} + 4x_{32} + 3x_{33} + 2x_{34} + x_{35}) + 3.0(5x_{41} + 4x_{42} + 3x_{43} + 2x_{44} + x_{45}) \quad (22) \end{aligned}$$

所以,问题的决策目标为

$$\begin{aligned} \min z = & z_1 + z_2 \\ = & 137.5x_{11} + 130x_{12} + 122.5x_{13} + 115x_{14} + 107.5x_{15} + \\ & 100x_{21} + 96x_{22} + 92x_{23} + 88x_{24} + 84x_{25} + \\ & 122.5x_{31} + 116x_{32} + 109.5x_{33} + 103x_{34} + 96.5x_{35} + \\ & 85x_{41} + 82x_{42} + 79x_{43} + 76x_{44} + 73x_{45} \quad (23) \end{aligned}$$

约束条件

问题的约束条件只有一个要求:每年公司的业务量必须能够由足够的代理承担.对于第1年,这一条件为

$$350x_{11} + 250x_{21} + 300x_{31} + 200x_{41} \geq 400 \quad (24)$$

对于第2年,这一条件为

$$350(x_{11} + x_{12}) + 250(x_{21} + x_{22}) + 300(x_{31} + x_{32}) + 200(x_{41} + x_{42}) \geq 500 \quad (25)$$

类似地,对于第3~5年,这一条件为

$$350(x_{11} + x_{12} + x_{13}) + 250(x_{21} + x_{22} + x_{23}) + 300(x_{31} + x_{32} + x_{33}) + 200(x_{41} + x_{42} + x_{43}) \geq 600 \quad (26)$$

$$350(x_{11} + x_{12} + x_{13} + x_{14}) + 250(x_{21} + x_{22} + x_{23} + x_{24}) + 300(x_{31} + x_{32} + x_{33} + x_{34}) + 200(x_{41} + x_{42} + x_{43} + x_{44}) \geq 700 \quad (27)$$

$$350(x_{11} + x_{12} + x_{13} + x_{14} + x_{15}) + 250(x_{21} + x_{22} + x_{23} + x_{24} + x_{25}) + 300(x_{31} + x_{32} + x_{33} + x_{34} + x_{35}) + 200(x_{41} + x_{42} + x_{43} + x_{44} + x_{45}) \geq 800 \quad (28)$$

模型求解

将模型(21), (23)~(28)输入LINGO(当然,加上0-1约束),求解得到 $x_{11} = x_{21} = x_{44} = 1$,其他变量为0,最小总费用为313.5万元。也就是说,公司应在第1年初与代理1、代理2建立代理关系,而在第4年初与代理4建立代理关系。

讨论 现在来考虑本题后半部分代理关系的运作决策问题。虽然代理关系目前已经建立,但由于每年初可以决定临时中断或重新恢复代理关系,使得问题的决策变量增加。用 $x_{it} = 1 (i = 1, 2, 3, 4; t = 1, 2, 3, 4, 5)$ 分别表示在第 t 年里公司允许代理 i 从事代理业务(0则表示不允许);用 $y_{it} = 1 (i = 1, 2, 3, 4; t = 1, 2, 3, 4, 5)$ 分别表示在第 t 年初公司与代理 i 中断代理业务(0则表示不中断或不需中断);用 $z_{it} = 1 (i = 1, 2, 3, 4; t = 1, 2, 3, 4, 5)$ 分别表示在第 t 年初公司与代理 i 恢复代理业务(0则表示不恢复或不需恢复)。

目标函数仍然是5年的总费用,包括运行费、业务中断费、业务恢复费三项。因此,问题的决策目标为

$$\begin{aligned} \min z = & 7.5 \sum_{i=1}^5 x_{1i} + 4.0 \sum_{i=1}^5 x_{2i} + 6.5 \sum_{i=1}^5 x_{3i} + 3.0 \sum_{i=1}^5 x_{4i} + \\ & 5 \sum_{i=1}^5 y_{1i} + 3 \sum_{i=1}^5 y_{2i} + 4 \sum_{i=1}^5 y_{3i} + 2 \sum_{i=1}^5 y_{4i} + \\ & 5 \sum_{i=1}^5 z_{1i} + 4 \sum_{i=1}^5 z_{2i} + \sum_{i=1}^5 z_{3i} + 9 \sum_{i=1}^5 z_{4i} \end{aligned} \quad (29)$$

为了建立模型,我们剩下的工作只需要列出 x_{it} , y_{it} 和 z_{it} 之间需要满足的约束关系。这些约束关系包括三类:业务量约束、业务中断约束、业务恢复约束。

业务量约束是指每年公司的业务量必须能够由足够的代理承担。相应的约束可以表达为

$$350x_{11} + 250x_{21} + 300x_{31} + 200x_{41} \geq 400 \quad (30)$$

$$350x_{12} + 250x_{22} + 300x_{32} + 200x_{42} \geq 500 \quad (31)$$

$$350x_{13} + 250x_{23} + 300x_{33} + 200x_{43} \geq 600 \quad (32)$$

$$350x_{14} + 250x_{24} + 300x_{34} + 200x_{44} \geq 700 \quad (33)$$

$$350x_{15} + 250x_{25} + 300x_{35} + 200x_{45} \geq 800 \quad (34)$$

业务中断约束是指对每个代理而言,如果某年该代理处于运行状态,而下一年不处于运行状态,则下一年初公司必须与之临时中断代理关系。也就是说,如果 $x_{it} = 1$ 而 $x_{i,t+1} = 0$, 则 $y_{i,t+1} = 1$ 。相应的约束为

$$x_{it} - x_{i,t+1} \leq y_{i,t+1} \quad (i = 1, 2, 3, 4; t = 0, 1, 2, 3, 4) \quad (35)$$

注意到当前所有代理都处于运行状态,所以在(35)中 $t=0$ 时 $x_{i0} = 1$ 。

业务恢复约束是指对每个代理而言,如果某年该代理不是处于运行状态,而下一年处于运行状态,则下一年初公司必须与之恢复代理关系。也就是说,如果 $x_{it} = 0$ 而 $x_{i,t+1} = 1$, 则 $z_{i,t+1} = 1$ 。相应的约束为

$$x_{i,t+1} - x_{it} \leq z_{i,t+1} \quad (i = 1, 2, 3, 4; t = 1, 2, 3, 4) \quad (36)$$

注意到当前所有代理都处于运行状态,所以在(36)中 $t=0$ 时没有限制。

可以直接将模型(29)~(36)输入 LINGO 求解(加上所有 0-1 约束)。但是,这一模型中有 60 个 0-1 变量,而整数变量比较多时,LINGO 求解速度会减慢。此外,LINGO 的免费试用版对整数变量的个数的限制很严格(通常不超过 50 个),能否减少这一模型中整数变量的个数呢?

考虑到在(35)中, $t=0$ 时 $x_{i0} = 1$, 即 $1 \leq x_{i1} + y_{i1}$ 。进一步分析,此不等式约束可以写成等式约束 $1 = x_{i1} + y_{i1}$, 即 $y_{i1} = 1 - x_{i1}$ 。将此式代入(29),模型中可以消去四个变量 $y_{i1} (i = 1, 2, 3, 4)$ 。

考虑到在(36)中, $t=0$ 时没有限制,因此 z_{i1} 没有任何限制,而它们在(29)中对应的费用系数大于 0,所以在最优解中一定有 $z_{i1} = 0$,从而可以直接消去四个变量 $z_{i1} (i = 1, 2, 3, 4)$ 。

但即使作以上处理,整数变量的个数仍然有 52 个。进一步分析该问题的特点,实际上并没有必要要求 52 个变量都是 0-1 变量,而只要要求 x_{it} 为 0-1 变量、 y_{it} 和 z_{it} 为非负实数也就可以了。当最优解中 x_{it} 为 0-1 变量时, y_{it} 和 z_{it} 自然也是 0-1 变量。

对模型(29)~(36)作以上处理后输入 LINGO,求解得到 x_{it} 的值为 $x_{11} = x_{12} = x_{13} = x_{14} = x_{15} = x_{23} = x_{24} = x_{25} = x_{41} = x_{42} = x_{43} = x_{44} = x_{45} = 1$, 其他为 0, 最小总费用为 86.5 万元。也就是说:公司应在第 1 年初临时中断与代理 2、代理 3 的代理关系,而在第 3 年初重新恢复与代理 2 的代理关系。

4.5 饮料厂的生产与检修

在 4.1 节的例子中我们讨论的是单阶段的生产销售计划,实际上由于生产、

需求在时间上的连续性,从长期效益出发,应该制订多阶段生产计划。

实际生产中要考虑的除了成本费、存贮费等与产量有关的费用外,有时还要考虑与产量无关的固定费用,如生产准备费,这会给优化模型的求解带来新的困难。

本节通过两个实例研究以上问题。

例1 饮料厂的生产与检修计划

问题 某饮料厂生产一种饮料用以满足市场需求.该厂销售科根据市场预测,已经确定了未来4周该饮料的需求量.计划科根据本厂实际情况给出了未来4周的生产能力和生产成本,如表1所示.每周当饮料满足需求后有剩余时,要支出存贮费,为每周每千箱饮料0.2千元.问应如何安排生产计划,在满足每周市场需求的条件下,使四周的总费用(生产成本与存贮费之和)最小?

如果工厂必须在未来四周的某一周中安排一次设备检修,检修将占用当周15千箱的生产能力,但会使检修以后每周的生产能力提高5千箱,则检修应该安排在哪一周?

表1 饮料的生产和需求数据

周次	需求量/千箱	生产能力/千箱	每千箱成本/千元
1	15	30	5.0
2	25	40	5.1
3	35	45	5.4
4	25	20	5.5
合计	100	135	

问题分析 从表1的数据看,除第4周外每周的生产能力都超过每周的需求,且总生产能力超过总需求,故可以满足每周市场需求.如果第1周、第2周按需生产,第3周多生产5千箱以弥补第4周生产能力对需求的不足,就可以使总的存贮费最小.但是我们注意到:生产成本在逐周上升,所以从总费用最小的角度考虑,前几周多生产一些备用,可能是更好的生产方案.于是应该建立数学规划模型来寻找最优的生产与存贮策略.

模型假设

设饮料厂在第1周开始时没有库存;从费用最小考虑,自然地假定第4周末也不能有库存;周末有库存时需支出一周的存贮费;每周末的库存量就是下周初的库存量.

模型建立与求解

显然,问题的决策变量是未来4周饮料的生产量,分别记作 x_1, x_2, x_3, x_4 ;由于存贮费取决于库存量,记第1,2,3周末的库存量分别为 y_1, y_2, y_3 ;4周的总费

用生产成本与存贮费之和(记作 z)是问题的目标函数,如下面模型中的(1)式。

每周的产量、需求与库存之间存在着平衡关系,比如对第2周有 $y_1 + x_2 - 25 = y_2$,再注意到第1周初和第4周末的库存为0,就得到下面模型的约束条件(2-1)~(2-4)。(3),(4)则分别是产量限制和非负约束。综上可得如下的线性规划模型

$$\min z = 5.0x_1 + 5.1x_2 + 5.4x_3 + 5.5x_4 + 0.2(y_1 + y_2 + y_3) \quad (1)$$

$$\text{s. t.} \quad x_1 - y_1 = 15 \quad (2-1)$$

$$x_2 + y_1 - y_2 = 25 \quad (2-2)$$

$$x_3 + y_2 - y_3 = 35 \quad (2-3)$$

$$x_4 + y_3 = 25 \quad (2-4)$$

$$x_1 \leq 30, x_2 \leq 40, x_3 \leq 45, x_4 \leq 20 \quad (3)$$

$$x_1, x_2, x_3, x_4, y_1, y_2, y_3 \geq 0 \quad (4)$$

将以上模型输入 LINGO 求解,可以得到最优解 $(x_1, x_2, x_3, x_4, y_1, y_2, y_3) = (15, 40, 25, 20, 0, 15, 5)$,即四周的产量分别为15千箱、40千箱、25千箱、20千箱,这样第1周没有库存,第2,3周各有15千箱、5千箱库存。这个生产计划的总费用为528千元,达到最小。

讨论 在需要安排检修时,如果希望以上生产计划保持不变,可以看出:第1周和第2周都有不小于15千箱的生产能力剩余,因此这两周都可以安排检修:究竟将检修安排在哪一周呢?

实际上,由于检修后生产能力发生变化,把检修安排在第1或第2周不一定是最佳选择。从题目给的条件来看,将检修安排在任何一周都是可以的。我们引入0-1变量 w_1, w_2, w_3, w_4 ,用 $w_t = 1$ 表示检修安排在第 t 周($t=1, 2, 3, 4$)。由于检修将占用当周15千箱的生产能力,但会使检修以后每周的生产能力提高5千箱,所以上面模型中关于产量限制的约束条件(3)需要修改为

$$x_1 + 15w_1 \leq 30 \quad (3-1)$$

$$x_2 + 15w_2 - 5w_1 \leq 40 \quad (3-2)$$

$$x_3 + 15w_3 - 5w_2 - 5w_1 \leq 45 \quad (3-3)$$

$$x_4 + 15w_4 - 5w_1 - 5w_2 - 5w_3 \leq 20 \quad (3-4)$$

同时,由于四周中只有一次检修,模型中还需要增加关于 w_1, w_2, w_3, w_4 的约束

$$w_1 + w_2 + w_3 + w_4 = 1 \quad (5)$$

$$w_1, w_2, w_3, w_4 \in \{0, 1\} \quad (6)$$

用 LINGO 对修改的模型直接求解,输入格式为:

$$\begin{aligned} \min &= 5.0 * x1 + 5.1 * x2 + 5.4 * x3 + 5.5 * x4 + 0.2 * y1 + 0.2 * y2 \\ &+ 0.2 * y3; \\ &x1 - y1 = 15; \end{aligned}$$


```

x2 + y1 - y2 = 25;
x3 + y2 - y3 = 35;
x4 + y3 = 25;
x1 + 15 * w1 < 30;
x2 + 15 * w2 - 5 * w1 < 40;
x3 + 15 * w3 - 5 * w1 - 5 * w2 < 45;
x4 + 15 * w4 - 5 * w1 - 5 * w2 - 5 * w3 < 20;
w1 + w2 + w3 + w4 = 1;
@ bin(w1); @ bin(w2); @ bin(w3); @ bin(w4);

```

求解得到输出(只列出需要的结果):

```

Global optimal solution found.
objective value:                      527.00000
Extended solver steps:                  0
Total solver iterations:                5

```

Variable	Value	Reduced Cost
X1	15.000000	0.000000
X2	45.000000	0.000000
X3	15.000000	0.000000
X4	25.000000	0.000000
Y1	0.000000	0.000000
Y2	20.000000	0.000000
Y3	0.000000	0.1000000
W1	1.000000	-0.5000000
W2	0.000000	1.500000
W3	0.000000	0.000000
W4	0.000000	0.000000

即应当将检修安排在第1周,每周的生产量分别为15千箱、45千箱、15千箱、25千箱,最小总费用从528千元下降为527千元。

评注 这一生产计划问题涉及多个生产阶段,各阶段的产量、需求与库存之间有一定的联系,如上面模型的(2-1)~(2-4)所示。虽然由这些式子可以解出库存量 y_1, y_2, y_3 ,代入(1)得到只有产量 x_1, x_2, x_3, x_4 的目标函数,但是通常我们宁愿增加库存量作为决策变量,使模型清晰和便于检查,而尽可能地把计算工作都留给计算机去做。

为安排检修计划引入 0-1 变量,得到混合整数规划模型,求解的结果总费用只下降了一点,是由于检修导致的生产能力的提高,在短短的 4 周内其作用远未得到充分体现。

例 2 饮料的生产批量问题

问题 某饮料厂使用同一条生产线轮流生产多种饮料以满足市场需求。如果某周开工生产其中一种饮料,就要清洗设备和更换部分部件,于是需支出生产准备费 8 千元。现在只考虑一种饮料的生产,假设其未来四周的需求量、生产能力、生产成本和存贮费与例 1 给出的完全相同。问应如何安排这种饮料的生产计划,在按时满足市场需求的前提下,使生产该种饮料的总费用最小?

问题分析 这一问题与例 1 的主要差别在于:除了考虑随产品数量变化的费用(如生产成本和存贮费)外,还要考虑与产品数量无关的费用,即生产准备费,只要某周开工生产时就有这项费用。

模型建立与求解

我们先用一般的数学符号表述这类问题。设需要考虑的时间跨度为 T 个时段(在本例中为 4 个时段,1 周是 1 个时段),第 t 时段的市场需求为 d_t ,生产能力为 M_t ($t=1,2,\dots,T$)。如果第 t 时段开工生产,则需付出生产准备费为 $s_t \geq 0$,单件产品的生产成本为 $c_t \geq 0$ 。在 t 时段末,如果有产品库存,单件产品 1 个时段的存贮费为 $h_t \geq 0$ 。目标函数是生产准备费、生产成本和存贮费之和。

假设在 t 时段,产品的生产量为 x_t (≥ 0), t 时段末产品的库存(即下一个时段的初始库存)为 y_t (≥ 0),合理地假设 $y_0 = y_T = 0$ 。产量、需求与库存之间的平衡关系同前。

为了表述在 t 时段是否生产这种饮料,从而确定是否要支付生产准备费,引入 0-1 变量 w_t , $w_t = 1$ 表示生产, $w_t = 0$ 表示不生产。于是这一问题可以用如下数学规划模型描述:

$$\min z = \sum_{t=1}^T (s_t w_t + c_t x_t + h_t y_t) \quad (7)$$

$$\text{s. t.} \quad y_{t-1} + x_t - y_t = d_t, \quad t = 1, 2, \dots, T \quad (8)$$

$$w_t = \begin{cases} 1, & x_t > 0, \\ 0, & x_t = 0, \end{cases} \quad t = 1, 2, \dots, T \quad (9)$$

$$x_t \leq M_t, \quad t = 1, 2, \dots, T \quad (10)$$

$$y_0 = y_T = 0, \quad (11)$$

$$x_t, y_t \geq 0, \quad t = 1, 2, \dots, T \quad (12)$$

(9) 式不便于计算,可以将(9),(10)合并地表为 $x_t \leq M_t w_t$,或用下面的线性不等式代替

$$x_t - M_t w_t \leq 0, \quad t = 1, 2, \dots, T \quad (13)$$

将本题所给参数代入这一模型,并将这一模型输入 LINGO 如下:

```
sets;
periods/1..4/:s,c,h,d,m,x,y,w;
endsets
data;
s=8 8 8 8;          ! 每次生产准备费用;
c=5.0 5.1 5.4 5.5;   ! 单件生产费用;
h=0.2 0.2 0.2 0.2;   ! 单件生产库存费用;
d=15 25 35 25;       ! 产品需求数量;
m=30 40 45 20;       ! 生产能力;
enddata
min=asum(periods:s*w+c*x+h*y);
x(1)-d(1)=y(1);
afor(periods(t)|t#gt#1:y(t-1)+x(t)-d(t)=y(t));
afor(periods:x<m*w;abin(w));
```

求解得到生产计划为:4 周的生产量分别为 15 千箱、40 千箱、45 千箱、0 千箱,最小总费用为 554.0 千元。与例 1 每周都有生产的结果相比,这里只在前 3 周生产,虽然存贮费增加了,但是节省了生产准备费。

评注 这种既包含可变费用(生产成本、存贮费)又包含固定费用(生产准备费)的生产计划问题一般称为生产批量(Lot-sizing)问题,通常引入 0-1 变量来处理。如果对产量和库存量没有提出整数要求(即认为它们可以取任意非负实数),那么这种模型是同时包含连续变量和整数变量的数学规划,一般称为混合整数规划。

如果把例 2 中对每周生产能力的限制去掉,则将约束(13)中的 M_i 换成一个充分大的常数(比如 4 周的总需求量)就可以了。读者不妨解一下这个问题,会发现其最优解为,4 周的生产量分别为 15 千箱、85 千箱、0 千箱、0 千箱。与上面生产能力有限的结果相比,这里只有第 1,2 周生产,生产准备费会更加节省。

4.6 钢管和易拉罐下料

生产中常会遇到通过切割、剪裁、冲压等手段,将原材料加工成所需尺寸这种工艺过程,称为原料下料问题。按照进一步的工艺要求,确定下料方案,使用料最省或利润最大,是典型的优化问题。本节通过两个实例讨论用数学规划模型解决这类问题的方法。

例1 钢管下料

问题 某钢管零售商从钢管厂进货,将钢管按照顾客的要求切割后售出,从钢管厂进货时得到的原料钢管都是19 m.

(1) 现有一客户需要50根4 m、20根6 m和15根8 m的钢管.应如何下料最节省?

(2) 零售商如果采用的不同切割模式太多,将会导致生产过程的复杂化,从而增加生产和管理成本,所以该零售商规定采用的不同切割模式不能超过3种.此外,该客户除需要(1)中的三种钢管外,还需要10根5 m的钢管.应如何下料最节省^[29]?

问题(1)的求解

问题分析 首先,应当确定哪些切割模式是可行的.所谓一个切割模式,是指按照客户需要在原料钢管上安排切割的一种组合.例如:我们可以将19 m的钢管切割成3根4 m的钢管,余料为7 m;或者将19 m的钢管切割成4 m、6 m和8 m的钢管各1根,余料为1 m.显然,可行的切割模式是很多的.

其次,应当确定哪些切割模式是合理的.通常假设一个合理的切割模式的余料不应该大于或等于客户需要的钢管的最小尺寸.例如:将19 m的钢管切割成3根4 m的钢管是可行的,但余料为7 m,可以进一步将7 m的余料切割成4 m钢管(余料为3 m),或者将7 m的余料切割成6 m钢管(余料为1 m).在这种合理性假设下,切割模式一共有7种,如表1所示.

表1 钢管下料的合理切割模式

	4 m 钢管根数	6 m 钢管根数	8 m 钢管根数	余料/m
模式1	4	0	0	3
模式2	3	1	0	1
模式3	2	0	1	3
模式4	1	2	0	3
模式5	1	1	1	1
模式6	0	3	0	1
模式7	0	0	2	3

问题化为在满足客户需要的条件下,按照哪些种合理的模式,切割多少根原料钢管,最为节省.而所谓节省,可以有两种标准:一是切割后剩余的总余料量最小,二是切割原料钢管的总根数最少.下面将对这两个目标分别讨论.

模型建立

决策变量 用 x_i 表示按照第 i 种模式 ($i=1, 2, \dots, 7$) 切割的原料钢管的根数, 显然它们应当是非负整数.

决策目标 以切割后剩余的总余料量最小为目标, 则由表 1 可得

$$\min z_1 = 3x_1 + x_2 + 3x_3 + 3x_4 + x_5 + x_6 + 3x_7 \quad (1)$$

以切割原料钢管的总根数最少为目标, 则有

$$\min z_2 = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 \quad (2)$$

下面分别在这两种目标下求解.

约束条件 为满足客户的需求, 按照表 1 应有

$$4x_1 + 3x_2 + 2x_3 + x_4 + x_5 \geq 50 \quad (3)$$

$$x_2 + 2x_4 + x_5 + 3x_6 \geq 20 \quad (4)$$

$$x_3 + x_5 + 2x_7 \geq 15 \quad (5)$$

模型求解

1. 将(1), (3), (4), (5)构成的整数线性规划模型(加上整数约束)输入 LINGO 求解, 可以得到最优解如下: $x_2 = 12, x_5 = 15$ (其余变量为 0). 即按照模式 2 切割 12 根原料钢管, 按照模式 5 切割 15 根原料钢管, 共 27 根, 总余料量为 27 m. 显然, 在总余料量最小的目标下, 最优解将是使用余料尽可能小的切割模式(模式 2 和 5 的余料为 1 m), 这会导致切割原料钢管的总根数较多.

2. 将(2)~(5)构成的整数线性规划模型(加上整数约束)输入 LINGO 求解, 可以得到最优解如下: $x_2 = 15, x_5 = x_7 = 5$ (其余变量为 0). 即按照模式 2 切割 15 根原料钢管、按照模式 5 切割 5 根、按照模式 7 切割 5 根, 共 25 根, 可算出总余料量为 35 m. 与上面得到的结果相比, 总余料量增加了 8 m, 但是所用的原料钢管的总根数减少了 2 根. 在余料没有什么用途的情况下, 通常选择总根数最少为目标.

问题(2)的求解

问题分析 按照(1)的思路, 可以通过枚举法首先确定哪些切割模式是可行的. 但由于需求的钢管规格增加到 4 种, 所以枚举法的工作量较大. 下面介绍的整数非线性规划模型, 可以同时确定切割模式和切割计划, 是带有普遍性的方法.

同(1)类似, 一个合理的切割模式的余料不应该大于或等于客户需要的钢管的最小尺寸(本题中为 4 m), 切割计划中只使用合理的切割模式, 而由于本题中参数都是整数, 所以合理的切割模式的余量不能大于 3 m. 此外, 这里我们仅选择总根数最少为目标进行求解.

模型建立

决策变量 由于不同切割模式不能超过 3 种, 可以用 x_i 表示按照第 i 种模式 ($i=1, 2, 3$) 切割的原料钢管的根数, 显然它们应当是非负整数. 设所使用的第

i 种切割模式下每根原料钢管生产 4 m, 5 m, 6 m 和 8 m 的钢管数量分别为 r_{1i} , r_{2i} , r_{3i} , r_{4i} (非负整数).

决策目标 切割原料钢管的总根数最少, 目标为

$$\min z = x_1 + x_2 + x_3 \quad (6)$$

约束条件 为满足客户的需求, 应有

$$r_{11}x_1 + r_{12}x_2 + r_{13}x_3 \geq 50 \quad (7)$$

$$r_{21}x_1 + r_{22}x_2 + r_{23}x_3 \geq 10 \quad (8)$$

$$r_{31}x_1 + r_{32}x_2 + r_{33}x_3 \geq 20 \quad (9)$$

$$r_{41}x_1 + r_{42}x_2 + r_{43}x_3 \geq 15 \quad (10)$$

每一种切割模式必须可行、合理, 所以每根原料钢管的成品量不能超过 19 m, 也不能少于 16 m (余量不能大于 3 m), 于是

$$16 \leq 4r_{11} + 5r_{21} + 6r_{31} + 8r_{41} \leq 19 \quad (11)$$

$$16 \leq 4r_{12} + 5r_{22} + 6r_{32} + 8r_{42} \leq 19 \quad (12)$$

$$16 \leq 4r_{13} + 5r_{23} + 6r_{33} + 8r_{43} \leq 19 \quad (13)$$

模型求解

在(7)~(10)式中出现决策变量的乘积, 是一个整数非线性规划模型, 虽然用 LINGO 软件可以直接求解, 但也可以增加一些显然的约束条件, 从而缩小可行解的搜索范围, 有可能减少运行时间.

例如, 由于 3 种切割模式的排列顺序是无关紧要的, 所以不妨增加以下约束

$$x_1 \geq x_2 \geq x_3 \quad (14)$$

又例如, 我们注意到所需原料钢管的总根数有着明显的上界和下界. 首先,

无论如何, 原料钢管的总根数不可能少于 $\left\lceil \frac{4 \times 50 + 5 \times 10 + 6 \times 20 + 8 \times 15}{19} \right\rceil + 1 =$

26 根. 其次, 考虑一种非常特殊的生产计划: 第一种切割模式下只生产 4 m 钢管, 一根原料钢管切割成 4 根 4 m 钢管, 为满足 50 根 4 m 钢管的需求, 需要 13 根原料钢管; 第二种切割模式下只生产 5 m、6 m 钢管, 一根原料钢管切割成 1 根 5 m 钢管和 2 根 6 m 钢管, 为满足 10 根 5 m 和 20 根 6 m 钢管的需求, 需要 10 根原料钢管; 第三种切割模式下只生产 8 m 钢管, 一根原料钢管切割成 2 根 8 m 钢管, 为满足 15 根 8 m 钢管的需求, 需要 8 根原料钢管. 于是满足要求的这种生产计划共需 $13 + 10 + 8 = 31$ 根原料钢管, 这就得到了最优解的一个上界. 所以可增加以下约束

$$26 \leq x_1 + x_2 + x_3 \leq 31 \quad (15)$$

将(6)~(15)构成的模型输入 LINGO 如下:

model:

! 给模型取个名字;

```

title cutting - stock problem;
sets:
    ! 定义基本集合 needs 及其属性 length,num;
    needs/1..4/:length,num;
    ! 定义基本集合 cuts 及其属性 x;
    cuts/1..3/:x;
    ! 定义派生集合 patterns(这是一个稠密集合)及其属性 r;
    patterns(needs,cuts):r;
endsets
data:
    length=4 5 6 8;
    num=50 10 20 15;
    capacity=19;
enddata
min=@ sum(cuts(i):x(i));
! 满足需求的约束;
@ for(needs(i):@ sum(cuts(j):x(j)*r(i,j))>num(i));
! 合理切割模式的约束;
@ for(cuts(j):@ sum(needs(i):length(i)*r(i,j))<capacity);
@ for(cuts(j):@ sum(needs(i):length(i)*r(i,j))>capacity
    -@ min(needs:length));
! 人为增加的约束;
@ sum(cuts:x)>26;@ sum(cuts:x)<31;
@ for(cuts(i)|i#lt#@ size(cuts):x(i)>x(i+1));
@ for(cuts:@ gin(x));
@ for(patterns:@ gin(r));
end

```

求解可得到输出如下(只列出需要的结果):

```

Local optimal solution found.
objective value:                28.00000
Extended solver steps:          52
Total solver iterations:        2108
Model Title:cutting - stock problem

```

Variable	Value	Reduced Cost
X(1)	10.00000	1.000000

X(2)	10.00000	1.000000
X(3)	8.000000	1.000000
R(1,1)	3.000000	0.000000
R(1,2)	2.000000	0.000000
R(1,3)	0.000000	0.000000
R(2,1)	0.000000	0.000000
R(2,2)	1.000000	0.000000
R(2,3)	0.000000	0.000000
R(3,1)	1.000000	0.000000
R(3,2)	1.000000	0.000000
R(3,3)	0.000000	0.000000
R(4,1)	0.000000	0.000000
R(4,2)	0.000000	0.000000
R(4,3)	2.000000	0.000000

即按照模式 1,2,3 分别切割 10,10,8 根原料钢管,使用原料钢管总根数为 28 根. 第一种切割模式下一根原料钢管切割成 3 根 4 m 钢管和 1 根 6 m 钢管;第二种切割模式下一根原料钢管切割成 2 根 4 m 钢管、1 根 5 m 钢管和 1 根 6 m 钢管;第三种切割模式下一根原料钢管切割成 2 根 8 m 钢管. 但这个模型的解并不惟一,你能找出一个与此不同的解吗?

例 2 易拉罐下料

问题 某公司采用一套冲压设备生产一种罐装饮料的易拉罐,这种易拉罐是用镀锡板冲压制成的. 易拉罐为圆柱形,包括罐身、上盖和下底,罐身高 10 cm,上盖和下底的直径均为 5 cm. 该公司使用两种不同规格的镀锡板原料:规格 1 的镀锡板为正方形,边长 24 cm;规格 2 的镀锡板为长方形,长、宽分别为 32 cm 和 28 cm. 由于生产设备和生产工艺的限制,对于规格 1 的镀锡板原料,只可以按照图 1 中的模式 1、模式 2 或模式 3 进行冲压;对于规格 2 的镀锡板原料只能按照模式 4 进行冲压. 使用模式 1、模式 2、模式 3、模式 4 进行每次冲压所需要的时间分别为 1.5 s、2 s、1 s、3 s.

该工厂每周工作 40 h,每周可供使用的规格 1、规格 2 的镀锡板原料分别为 5 万张和 2 万张. 目前每只易拉罐的利润为 0.10 元,原料余料损失为 0.001 元/cm² (如果周末有罐身、上盖或下底不能配套组装成易拉罐出售,也看做是原料余料损失).

问工厂应如何安排每周的生产?

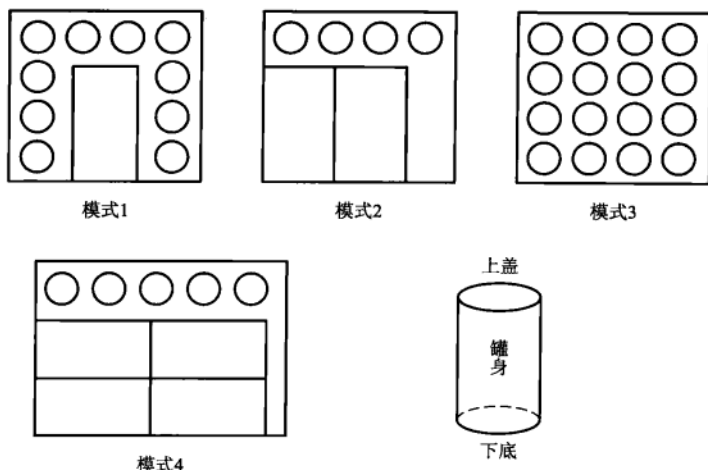


图1 易拉罐下料模式

问题分析 与钢管下料问题不同的是,这里的切割模式已经确定,只需计算各种模式下的余料损失. 已知上盖和下底的直径 $d = 5 \text{ cm}$, 可得其面积为 $s = \pi d^2/4 \approx 19.6 \text{ cm}^2$, 周长为 $L = \pi d \approx 15.7 \text{ cm}$; 已知罐身高 $h = 10 \text{ cm}$, 可得其面积为 $S = hL \approx 157.1 \text{ cm}^2$. 于是模式1下的余料损失为 $24^2 - 10s - S \approx 222.6 \text{ cm}^2$. 同理计算其他模式下的余料损失, 并可将4种冲压模式的特征归纳如表2.

表2 4种冲压模式的特征

	罐身个数	底、盖个数	余料损失/ cm^2	冲压时间/s
模式1	1	10	222.6	1.5
模式2	2	4	183.3	2
模式3	0	16	261.8	1
模式4	4	5	169.5	3

问题的目标显然应是易拉罐的利润扣除原料余料损失后的净利润最大, 约束条件除每周工作时间和原料数量外, 还要考虑罐身和底、盖的配套组装.

模型建立

决策变量 用 x_i 表示按照第 i 种模式的冲压次数 ($i = 1, 2, 3, 4$), y_1 表示一周生产的易拉罐个数. 为计算不能配套组装的罐身和底、盖造成的原料损失, 用 y_2 表示不配套的罐身个数, y_3 表示不配套的底、盖个数. 虽然实际上 x_i 和 $y_1, y_2,$

y_3 应该是整数. 但是由于生产量相当大, 可以把它们看成是实数, 从而用线性规划模型处理.

决策目标 假设每周生产的易拉罐能够全部售出, 公司每周的销售利润是 $0.1y_1$. 原料余料损失包括两部分: 4 种冲压模式下的余料损失和不配套的罐身和底、盖造成的原料损失. 按照前面的计算及表 2 的结果, 总损失为 $0.001(222.6x_1 + 183.3x_2 + 261.8x_3 + 169.5x_4 + 157.1y_2 + 19.6y_3)$.

于是, 决策目标为

$$\max z = 0.1y_1 - 0.001(222.6x_1 + 183.3x_2 + 261.8x_3 + 169.5x_4 + 157.1y_2 + 19.6y_3) \quad (16)$$

约束条件

时间约束: 每周工作时间不超过 $40 \text{ h} = 144\,000 \text{ s}$, 由表 2 最后一列得

$$1.5x_1 + 2x_2 + x_3 + 3x_4 \leq 144\,000 \quad (17)$$

原料约束: 每周可供使用的规格 1、规格 2 的镀锡板原料分别为 50 000 张和 20 000 张, 即

$$x_1 + x_2 + x_3 \leq 50\,000 \quad (18)$$

$$x_4 \leq 20\,000 \quad (19)$$

配套约束: 由表 2 一周生产的罐身个数为 $x_1 + 2x_2 + 4x_4$, 一周生产的底、盖个数为 $10x_1 + 4x_2 + 16x_3 + 5x_4$, 因为应尽可能将它们配套组装成易拉罐销售. 所以 y_1 满足

$$y_1 = \min \{ x_1 + 2x_2 + 4x_4, (10x_1 + 4x_2 + 16x_3 + 5x_4)/2 \} \quad (20)$$

这时不配套的罐身个数 y_2 和不配套的底、盖个数 y_3 应为

$$y_2 = x_1 + 2x_2 + 4x_4 - y_1 \quad (21)$$

$$y_3 = 10x_1 + 4x_2 + 16x_3 + 5x_4 - 2y_1 \quad (22)$$

(16) ~ (22) 就是我们得到的模型, 其中 (20) 是一个非线性关系, 不易直接处理, 但是它可以等价于以下两个线性不等式

$$y_1 \leq x_1 + 2x_2 + 4x_4 \quad (23)$$

$$y_1 \leq (10x_1 + 4x_2 + 16x_3 + 5x_4)/2 \quad (24)$$

模型求解

模型 (16) ~ (19) 和 (21) ~ (24) 可以直接输入 LINGO 求解, 但注意到约束 (17) ~ (19) 中右端项的数值过大 (与左端的系数相比较), 模型中数据之间的数量级不匹配, 此时 LINGO 在计算中容易产生比较大的误差. 我们可以先进行预处理, 缩小数据之间的差别, 例如可以将所有决策变量扩大 10 000 倍 (相当于 x_i 以万次为单位, y_i 以万件为单位). 此时, 目标 (16) 可以保持不变 (记住得到的结果单位为万元就可以了), 而约束 (17) ~ (19) 改为

$$1.5x_1 + 2x_2 + x_3 + 3x_4 \leq 14.4 \quad (25)$$

$$x_1 + x_2 + x_3 \leq 5 \quad (26)$$

$$x_4 \leq 2 \quad (27)$$

将模型(16)和(21)~(27)输入 LINGO 求解得到:

Global optimal solution found.

Objective value: 0.4298337

Total solver iterations: 4

Variable	Value	Reduced Cost
Y1	16.02500	0.000000
X1	0.000000	0.5000000E-04
X2	4.012500	0.000000
X3	0.3750000	0.000000
X4	2.000000	0.000000
Y2	0.000000	0.2233312
Y3	0.000000	0.3648437E-01

即模式 1 不使用,模式 2 使用 40 125 次,模式 3 使用 3 750 次,模式 4 使用 20 000 次,可生产易拉罐 160 250 个,罐身和底、盖均无剩余,净利润为 4 298 元。

评注 下料问题的建模主要有两部分组成,一是确定下料模式,二是构造优化模型。确定下料模式尚无通用的方法,对于钢管下料这样的一维问题,当需要下料的规格不太多时,可以枚举出下料模式,建立整数线性规划模型;否则就要构造整数非线性规划模型,而这种模型求解比较困难。本节介绍的增加约束条件的方法是将原来的可行域“割去”一部分,但要保证剩下的可行域中仍存在原问题的最优解。而像易拉罐下料这样的二维问题,就要复杂多了,读者不妨试一下,看看还有没有比图 1 给出的更好的模式。至于构造优化模型,则要根据问题的要求和限制具体处理,其中应特别注意配套组装的情况。

习 题

1. 某银行经理计划用一笔资金进行有价证券的投资,可供购进的证券以及其信用等级、到期年限、收益如表 1 所示。按照规定,市政证券的收益可以免税,其他证券的收益需按 50% 的税率纳税。此外还有以下限制:

- (1) 政府及代办机构的证券总共至少要购进 400 万元;
- (2) 所购证券的平均信用等级不超过 1.4 (信用等级数字越小,信用程度越高);
- (3) 所购证券的平均到期年限不超过 5 年。

表1 证券信息

证券名称	证券种类	信用等级	到期年限	到期税前收益/%
A	市政	2	9	4.3
B	代办机构	2	15	5.4
C	政府	1	4	5.0
D	政府	1	3	4.4
E	市政	5	2	4.5

问:(1) 若该经理有1 000万元资金,应如何投资?

(2) 如果能够以2.75%的利率借到不超过100万元资金,该经理应如何操作?

(3) 在1 000万元资金情况下,若证券A的税前收益增加为4.5%,投资应否改变?若证券C的税前收益减少为4.8%,投资应否改变^[14]?

2. 一家出版社准备在某市建立两个销售代理点,向7个区的大学生售书,每个区的大学生数量(单位:千人)已经表示在图1上. 每个销售代理点只能向本区和一个相邻区的大学生售书,这两个销售代理点应该建在何处,才能使所能供应的大学生的数量最大? 建立该问题的整数线性规划模型并求解^[72].

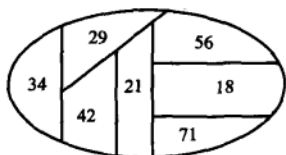


图1

3. 某储蓄所每天的营业时间是上午9:00到下午5:00. 根据经验,每天不同时间段所需要的服务员数量如表2:

表2 不同时间段所需要的服务员数量

时间段/时	9~10	10~11	11~12	12~1	1~2	2~3	3~4	4~5
服务员数量	4	3	4	6	5	6	8	8

储蓄所可以雇佣全时和半时两类服务员. 全时服务员每天报酬100元,从上午9:00到下午5:00工作,但中午12:00到下午2:00之间必须安排1h的午餐时间. 储蓄所每天可以雇佣不超过3名的半时服务员,每个半时服务员必须连续工作4h,报酬40元. 问该储蓄所应如何雇佣全时和半时两类服务员? 如果不能雇佣半时服务员,每天至少增加多少费用? 如果雇佣半时服务员的数量没有限制,每天可以减少多少费用?

4. 一家保姆服务公司专门向顾主提供保姆服务. 根据估计,下一年的需求是:春季6 000人日,夏季7 500人日,秋季5 500人日,冬季9 000人日. 公司新招聘的保姆必须经过5天的培训才能上岗,每个保姆每季度工作(新保姆包括培训)65天. 保姆从该公司而不是从顾主那里得到报酬,每人每月工资800元. 春季开始时公司拥有120名保姆,在每个季度结束后,将有15%的保姆自动离职.

(1) 如果公司不允许解雇保姆,请你为公司制定下一年的招聘计划;哪些季度需求的增加不影响招聘计划? 可以增加多少?

(2) 如果公司在每个季度结束后允许解雇保姆,请为公司制定下一年的招聘计划^[14]。

5. 在甲乙双方的一场战争中,一部分甲方部队被乙方部队包围长达4个月。由于乙方封锁了所有水陆交通通道,被包围的甲方部队只能依靠空中交通维持供给。运送4个月的供给分别需要2次,3次,3次,4次飞行,每次飞行编队由50架飞机组成(每架飞机需要3名飞行员),可以运送10万t物资。每架飞机每个月只能飞行一次,每名飞行员每个月也只能飞行一次。在执行完运输任务后的返回途中有20%的飞机会被乙方部队击落,相应的飞行员也因此牺牲或失踪。在第1个月开始时,甲方拥有110架飞机和330名熟练的飞行员。在每个月开始时,甲方可以招聘新飞行员和购买新飞机。新飞机必须经过一个月的检查后才可以投入使用,新飞行员必须在熟练飞行员的指导下经过一个月的训练才能投入飞行。每名熟练飞行员可以作为教练每个月指导20名飞行员(包括他自己在内)进行训练。每名飞行员在完成一个月的飞行任务后,必须有一个月的带薪假期,假期结束后才能再投入飞行。已知各项费用(单位略去)如表3所示,请你为甲方安排一个飞行计划。

表3 费用表

	第1个月	第2个月	第3个月	第4个月
新飞机价格	200.0	195.0	190.0	185.0
闲置的熟练飞行员报酬	7.0	6.9	6.8	6.7
教练和新飞行员报酬(包括培训费用)	10.0	9.9	9.8	9.7
执行飞行任务的熟练飞行员报酬	9.0	8.9	9.8	9.7
休假期间的熟练飞行员报酬	5.0	4.9	4.8	4.7

如果每名熟练飞行员可以作为教练每个月指导不超过20名飞行员(包括他自己在内)进行训练,模型和结果有哪些改变^[2]?

6. 某公司将4种不同含硫量的液体原料(分别记为甲、乙、丙、丁)混合生产两种产品(分别记为A、B)。按照生产工艺的要求,原料甲、乙、丁必须首先倒入混合池中混合,混合后的液体再分别与原料丙混合生产A、B。已知原料甲、乙、丙、丁的含硫量分别是3%,1%,2%,1%,进货价格分别为6,16,10,15(千元/t);产品A、B的含硫量分别不能超过2.5%,1.5%,售价分别为9,15(千元/t)。根据市场信息,原料甲、乙、丙的供应没有限制,原料丁的供应量最多为50t;产品A、B的市场需求量分别为100t、200t。问应如何安排生产^[29]?

7. 某钢管零售商从钢管厂进货,将钢管按照顾客的要求切割后售出。从钢管厂进货时得到的原料钢管长度都是1850mm。现有一客户需要15根290mm、28根315mm、21根350mm和30根455mm的钢管。为了简化生产过程,规定所使用的切割模式的种类不能超过4种,使用频率最高的一种切割模式按照一根原料钢管价值的1/10增加费用,使用频率次之的切割模式按照一根原料钢管价值的2/10增加费用。以此类推,且每种切割模式下的切割次数不能太多(一根原料钢管最多生产5根产品)。此外,为了减少余料浪费,每种切割模式下的余料浪费不能超过100mm。为了使总费用最小,应如何下料^[29]?

8. 某电力公司经营两座发电站,发电站分别位于两个水库上,位置如图2所示。

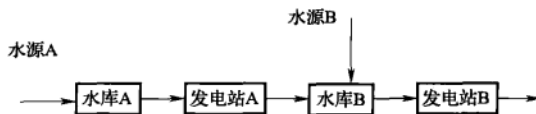


图 2

已知发电站 A 可以将水库 A 的 $10\,000\text{ m}^3$ 的水转换为 400 千度电能,发电站 B 只能将水库 B 的 $10\,000\text{ m}^3$ 的水转换为 200 千度电能.发电站 A,B 每个月的最大发电能力分别是 60 000 千度,35 000 千度.每个月最多有 50 000 千度电能够以 200 元/千度的价格售出,多余的电能只能以 140 元/千度的价格售出.水库 A,B 的其他有关数据如表 4(单位: 10^4 m^3)

表 4

		水库 A	水库 B
水库最大蓄水量		2 000	1 500
水源流入水量	本月	200	40
	下月	130	15
水库最小蓄水量		1 200	800
水库目前蓄水量		1 900	850

请你为该电力公司制订本月和下月的生产经营计划.(千度是非国际单位制单位, $1\text{ 千度} = 10^3\text{ kW}\cdot\text{h}$)

9. 如图 3 所示,有若干工厂的污水经排污口流入某江,各口有污水处理站,处理站对面是居民点.工厂 1 上游江水流量和污水质量浓度,国家标准规定的水的污染浓度,以及各个工厂的污水流量和污水质量浓度均已知.设污水处理费用与污水处理前后的质量浓度差和污水流量成正比,使每单位流量的污水下降一个质量浓度单位需要的处理费用(称处理系数)为已知.处理后的污水与江水混合,流到下一个排污口之前,自然状态下的江水也会使污水质量浓度降低一个比例系数(称自净系数),该系数可以估计.试确定各污水处理站出口的污水质量浓度,使在符合国家标准规定的条件下总的处理费用最小.

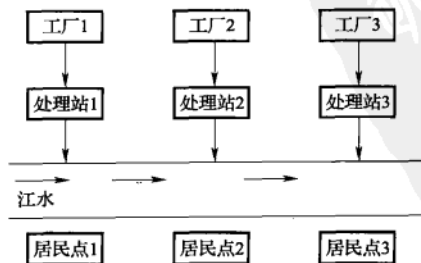


图 3

先建立一般情况下的数学模型,再求解以下具体问题:

设上游江水流量为 $1\,000 \times 10^{12} \text{ L/min}$, 污水质量浓度为 0.8 mg/l , 3 个工厂的污水流量均为 $5 \times 10^{12} \text{ L/min}$, 污水质量浓度(从上游到下游排列)分别为 $100, 60, 50 (\text{mg/l})$, 处理系数均为 $1 \text{ 万元}/((10^{12} \text{ L/min}) \times (\text{mg/l}))$, 3 个工厂之间的两段江面的自净系数(从上游到下游)分别为 $0.9, 0.6$. 国家标准规定水的污染质量浓度不能超过 1 mg/l .

(1) 为了使江面上所有地段的水污染达到国家标准,最少需要花费多少费用?

(2) 如果只要求三个居民点上游的水污染达到国家标准,最少需要花费多少费用^[93]?

10. 生产裸铜线和塑包线的工艺如图 4 所示:

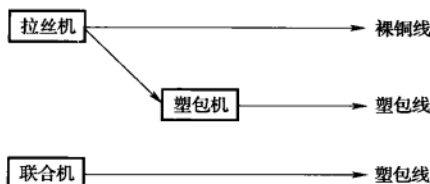


图 4

某厂现有 I 型拉丝机和塑包机各一台,生产两种规格的裸铜线和相应的两种塑包线,没有拉丝塑包联合机(简称联合机).由于市场需求扩大和现有塑包机设备陈旧,计划新增 II 型拉丝机或联合机(由于场地限制,每种设备最多 1 台),或改造塑包机,每种设备选用方案及相关数据如表 5:

表 5

	拉丝机		塑包机		联合机
	原有 I 型	新购 II 型	原有	改造	新购
方案代号	1	2	3	4	5
所需投资/万元	0	20	0	10	50
每小时运行费用/元	5	7	8	8	12
每年固定费用/万元	3	5	8	10	14
规格 1 生产效率/ $(\text{m} \cdot \text{h}^{-1})$	1 000	1 500	1 200	1 600	1 600
规格 2 生产效率/ $(\text{m} \cdot \text{h}^{-1})$	800	1 400	1 000	1 300	1 200
废品率/%	2	2	3	3	3
每 km 废品损失/元	30	30	50	50	50

已知市场对两种规格裸铜线的需求分别为 $3\,000 \text{ km}$ 和 $2\,000 \text{ km}$, 对两种规格塑包线的需求分别为 $10\,000 \text{ km}$ 和 $8\,000 \text{ km}$. 按照规定,新购及改进设备按每年 5% 提取折旧费,老设备不提;每台机器每年最多只能工作 $8\,000 \text{ h}$. 为了满足需求,确定使总费用最小的设备选用方案

和生产计划^[80].

11. 有 4 名同学到一家公司参加三个阶段的面试:公司要求每个同学都必须首先找公司秘书初试,然后到部门主管处复试,最后到经理处参加面试,并且不允许插队(即在任何一个阶段 4 名同学的顺序是一样的).由于 4 名同学的专业背景不同,所以每人在三个阶段的面试时间也不同,如表 6 所示(单位:min):

表 6

	秘书初试	主管复试	经理面试
同学甲	13	15	20
同学乙	10	20	18
同学丙	20	16	10
同学丁	8	10	15

这 4 名同学约定他们全部面试完以后一起离开公司.假定现在是早 8:00,问他们最早何时能离开公司^[80]?



第5章

微分方程模型

当我们描述实际对象的某些特性随时间(或空间)而演变的过程、分析它的变化规律、预测它的未来性态、研究它的控制手段时,通常要建立对象的动态模型.建模时首先要根据建模目的和对问题的具体分析作出简化假设,然后按照对象内在的或可以类比的其他对象的规律列出微分方程,求出方程的解并将结果翻译回实际对象,就可以进行描述、分析、预测或控制了.我们在1.5节建立的胃肠道和血液中药量变化的模型,大致就是这个过程.

事实上,在微分方程课程中,解所谓应用题时我们已经遇到简单的建立动态模型问题,例如“一质量为 m 的物体自高 h 处自由下落,初速是0,设阻力与下落速度的平方成正比,比例系数为 k ,求下落速度随时间的变化规律”,又如“容器内有盐水100 l,内含盐10 kg,今以3 l/min的速度从一管放进净水,以2 l/min的速度从另一管抽出盐水,设容器内盐水浓度始终是均匀的,求容器内含盐量随时间变化的规律”.本章讨论的动态模型与这些问题的主要区别是,所谓微分方程应用题大多是物理或几何方面的典型问题,假设条件已经给出,只需用数学符号将已知规律表示出来,即可列出方程,求解的结果就是问题的答案,答案是惟一的,已经确定的.而本章的模型主要是非物理领域的实际问题,要分析具体情况或进行类比才能给出假设条件.作出不同的假设,就得到不同的方程,所以是事先没有答案的.求解结果还要用来解释实际现象并接受检验.

5.1 传染病模型

随着卫生设施的改善、医疗水平的提高以及人类文明的不断发展,诸如霍乱、天花等曾经肆虐全球的传染性疾病已经得到有效的控制.但是一些新的、不断变异着的传染病毒却悄悄向人类袭来.20世纪80年代,十分险恶的艾滋病毒开始肆虐全球,至今仍在蔓延;2003年春来历不明的SARS病毒突袭人间,给人们的生命财产带来极大的危害.长期以来,建立传染病的数学模型来描述传染病的传播过程,分析受感染人数的变化规律,探索制止传染病蔓延的手段等,一直是各国有关专家和官员关注的课题.

不同类型传染病的传播过程有其各自不同的特点,弄清这些特点需要相当多的病理知识,这里不可能从医学的角度——分析各种传染病的传播,而只是按照一般的传播机理建立几种模型^[15,16,46].

模型 1 在这个最简单的模型中,设时刻 t 的病人人数 $x(t)$ 是连续、可微函数,并且每天每个病人有效接触(足以使人致病的接触)的人数为常数 λ ,考察 t 到 $t + \Delta t$ 病人人数的增加,就有

$$x(t + \Delta t) - x(t) = \lambda x(t) \Delta t$$

再设 $t=0$ 时有 x_0 个病人,即得微分方程

$$\frac{dx}{dt} = \lambda x, \quad x(0) = x_0 \quad (1)$$

方程(1)的解为

$$x(t) = x_0 e^{\lambda t} \quad (2)$$

结果表明,随着 t 的增加,病人人数 $x(t)$ 无限增长,这显然是不符合实际的.

建模失败的原因在于:在病人有效接触的人群中,有健康人也有病人,而其中只有健康人才可以被传染为病人,所以在改进的模型中必须区别这两种人.

模型 2 (SI 模型) 假设条件为

1. 在疾病传播期内所考察地区的总人数 N 不变,既不考虑生死,也不考虑迁移.人群分为易感染者 (Susceptible) 和已感染者 (Infective) 两类(取两个词的第 1 个字母,称之为 SI 模型),以下简称健康者和病人.时刻 t 这两类人在总人数中所占的比例分别记作 $s(t)$ 和 $i(t)$.

2. 每个病人每天有效接触的平均人数是常数 λ , λ 称为日接触率.当病人与健康者有效接触时,使健康者受感染变为病人.

根据假设,每个病人每天可使 $\lambda s(t)$ 个健康者变为病人,因为病人数为 $Ni(t)$,所以每天共有 $\lambda N s(t) i(t)$ 个健康者被感染,于是 $\lambda N s i$ 就是病人数 Ni 的增加率,即有

$$N \frac{di}{dt} = \lambda N s i \quad (3)$$

又因为

$$s(t) + i(t) = 1 \quad (4)$$

再记初始时刻($t=0$)病人的比例为 i_0 ,则

$$\frac{di}{dt} = \lambda i(1 - i), \quad i(0) = i_0 \quad (5)$$

方程(5)是 logistic 模型(详细描述参见 5.6 节).它的解为

$$i(t) = \frac{1}{1 + \left(\frac{1}{i_0} - 1\right)e^{-\lambda t}} \quad (6)$$

$i(t) \sim t$ 和 $\frac{di}{dt} \sim i$ 的图形如图 1 和图 2 所示.

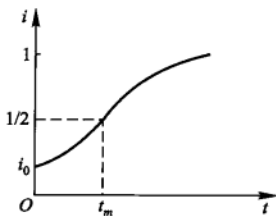


图 1 SI 模型的 $i \sim t$ 曲线

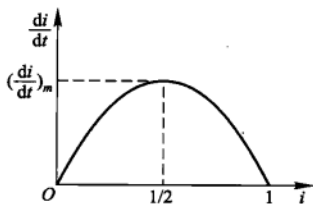


图 2 SI 模型的 $\frac{di}{dt} \sim i$ 曲线

由(5),(6)式及图 1、图 2 可知,第一,当 $i = 1/2$ 时 $\frac{di}{dt}$ 达到最大值 $\left(\frac{di}{dt}\right)_m$, 这个时刻为

$$t_m = \lambda^{-1} \ln\left(\frac{1}{i_0} - 1\right) \quad (7)$$

这时病人增加得最快,可以认为是医院的门诊量最大的一天,预示着传染病高潮的到来,是医疗卫生部门关注的时刻。 t_m 与 λ 成反比,因为日接触率 λ 表示该地区的卫生水平, λ 越小卫生水平越高. 所以改善保健设施、提高卫生水平可以推迟传染病高潮的到来. 第二,当 $t \rightarrow \infty$ 时 $i \rightarrow 1$,即所有人终将被传染,全变为病人,这显然不符合实际情况. 其原因是模型中没有考虑到病人可以治愈,人群中的健康者只能变成病人,病人不会再变成健康者.

为了修正上述结果,必须重新考虑模型的假设,下面两个模型中我们讨论病人可以治愈的情况.

模型 3(SIS 模型) 有些传染病如伤风、痢疾等愈后免疫力很低,可以假定无免疫性,于是病人被治愈后变成健康者,健康者还可以被感染再变成病人,所以这个模型称 SIS 模型.

SIS 模型的假设条件 1,2 与 SI 模型相同,增加的条件为

3. 每天被治愈的病人数占病人总数的比例为常数 μ ,称为日治愈率. 病人治愈后成为仍可被感染的健康者. 显然 $1/\mu$ 是这种传染病的平均传染期.

不难看出,考虑到假设 3,SI 模型的(3)式应修正为

$$N \frac{di}{dt} = \lambda N s i - \mu N i \quad (8)$$

(4) 式不变, 于是(5)式应改为

$$\frac{di}{dt} = \lambda i(1-i) - \mu i, \quad i(0) = i_0 \quad (9)$$

我们不去求解方程(9) (虽然它的解可以解析地表出), 而是通过图形分析 $i(t)$ 的变化规律. 定义

$$\sigma = \lambda/\mu \quad (10)$$

注意到 λ 和 $1/\mu$ 的含义, 可知 σ 是整个传染期内每个病人有效接触的平均人数, 称为接触数.

利用 σ , 方程(9)可以改写作

$$\frac{di}{dt} = -\lambda i \left[i - \left(1 - \frac{1}{\sigma} \right) \right] \quad (11)$$

由方程(11)容易先画出 $\frac{di}{dt} \sim i$ 的图形 (图 3, 图 5), 再画出 $i \sim t$ 的图形 (图 4, 图 6).

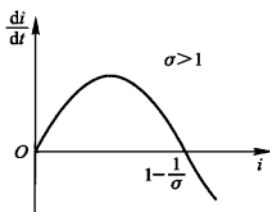


图 3 SIS 模型的 $\frac{di}{dt} \sim i$ 曲线 ($\sigma > 1$)

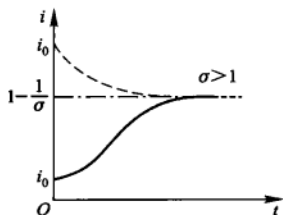


图 4 SIS 模型的 $i \sim t$ 曲线 ($\sigma > 1$),
其中虚线是 $i_0 > 1 - \frac{1}{\sigma}$ 的情况

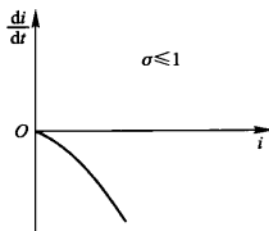


图 5 SIS 模型的 $\frac{di}{dt} \sim i$ 曲线 ($\sigma \leq 1$)



图 6 SIS 模型的 $i \sim t$ 曲线 ($\sigma \leq 1$)

不难看出, 接触数 $\sigma = 1$ 是一个阈值. 当 $\sigma > 1$ 时 $i(t)$ 的增减性取决于 i_0 的大小 (见图 4), 但其极限值 $i(\infty) = 1 - \frac{1}{\sigma}$ 随 σ 的增加而增加 (试从 σ 的含义给

予解释);当 $\sigma \leq 1$ 时病人比例 $i(t)$ 越来越小,最终趋于 0,这是由于传染期内经有效接触从而使健康者变成的病人不超过原来病人数的缘故。

SI 模型可视为本模型的特例,请读者考虑它相当于本模型中 μ 或 σ 取何值的情况。

模型 4(SIR 模型) 大多数传染病如天花、流感、肝炎、麻疹等治愈后均有很强的免疫力,所以病愈的人既非健康者(易感染者),也非病人(已感染者),他们已经退出传染系统。这种情况比较复杂,下面将详细分析建模过程。

模型假设

1. 总人数 N 不变. 人群分为健康者、病人和病愈免疫的移出者(Removed)三类,称 SIR 模型. 时刻 t 三类人在总人数 N 中占的比例分别记作 $s(t)$, $i(t)$ 和 $r(t)$ 。

2. 病人的日接触率为 λ ,日治愈率为 μ (与 SI 模型相同),传染期接触数为 $\sigma = \lambda/\mu$ 。

模型构成

由假设 1 显然有

$$s(t) + i(t) + r(t) = 1 \quad (12)$$

根据条件 2 方程(8)仍成立. 对于病愈免疫的移出者而言应有

$$N \frac{dr}{dt} = \mu Ni \quad (13)$$

再记初始时刻的健康者和病人的比例分别是 s_0 ($s_0 > 0$) 和 i_0 ($i_0 > 0$) (不妨设移出者的初始值 $r_0 = 0$),则由(8),(12),(13)式,SIR 模型的方程可以写作

$$\begin{cases} \frac{di}{dt} = \lambda si - \mu i, & i(0) = i_0 \\ \frac{ds}{dt} = -\lambda si, & s(0) = s_0 \end{cases} \quad (14)$$

方程(14)无法求出 $s(t)$ 和 $i(t)$ 的解析解,我们先作数值计算。

数值计算 在方程(14)中设 $\lambda = 1, \mu = 0.3, i(0) = 0.02, s(0) = 0.98$,用 MATLAB 软件编程:

```
function y = ill1(t,x)
a=1;b=0.3;
y=[a*x(1)*x(2)-b*x(1),-a*x(1)*x(2)]';

ts=0;50;
x0=[0.02,0.98];
[t,x]=ode45('ill1',ts,x0);[t,x]
```

```
plot(t,x(:,1),t,x(:,2)),grid,pause
plot(x(:,2),x(:,1)),grid,
```

输出的简明计算结果列入表1, $i(t), s(t)$ 的图形见图7, 图8是 $i \sim s$ 的图形, 称为相轨线, 初值 $i(0) = 0.02, s(0) = 0.98$ 相当于图8中的 P_0 点, 随着 t 的增加, (s, i) 沿轨线自右向左运动. 由表1、图7、图8可以看出, $i(t)$ 由初值增长至约 $t = 7$ 时达到最大值, 然后减少, $t \rightarrow \infty, i \rightarrow 0$; $s(t)$ 则单调减少, $t \rightarrow \infty, s \rightarrow 0.0398$.

表1 $i(t), s(t)$ 的数值计算结果

t	0	1	2	3	4	5	6	7	8
$i(t)$	0.020 0	0.039 0	0.073 2	0.128 5	0.203 3	0.279 5	0.331 2	0.344 4	0.324 7
$s(t)$	0.980 0	0.952 5	0.901 9	0.816 9	0.692 7	0.543 8	0.399 5	0.283 9	0.202 7
t	9	10	15	20	25	30	35	40	45
$i(t)$	0.286 3	0.241 8	0.078 7	0.022 3	0.006 1	0.001 7	0.000 5	0.000 1	0
$s(t)$	0.149 3	0.114 5	0.054 3	0.043 4	0.040 8	0.040 1	0.039 9	0.039 9	0.039 8

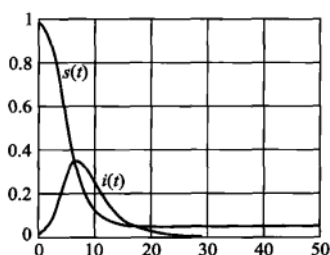


图7 $i(t), s(t)$ 图形

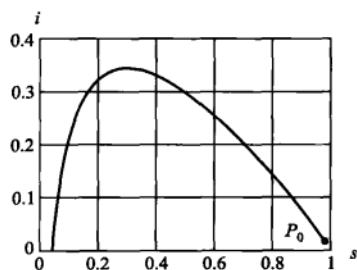


图8 $i \sim s$ 图形(相轨线)

为了分析 $i(t), s(t)$ 的一般变化规律, 需要进行相轨线分析.

相轨线分析 我们在数值计算和图形观察的基础上, 利用相轨线讨论解 $i(t), s(t)$ 的性质.

$s \sim i$ 平面称为相平面, 相轨线在相平面上的定义域 $(s, i) \in D$ 为

$$D = \{(s, i) | s \geq 0, i \geq 0, s + i \leq 1\} \quad (15)$$

在方程(14)中消去 dt , 并注意到 σ 的定义(10), 可得

$$\frac{di}{ds} = \frac{1}{\sigma s} - 1, \quad i|_{s=s_0} = i_0 \quad (16)$$

容易求出方程(16)的解为

$$i = (s_0 + i_0) - s + \frac{1}{\sigma} \ln \frac{s}{s_0} \quad (17)$$

在定义域 D 内, (17) 式表示的曲线即为相轨线, 如图 9 所示, 其中箭头表示了随着时间 t 的增加 $s(t)$ 和 $i(t)$ 的变化趋向.

下面根据 (14), (17) 式和图 9 分析 $s(t)$, $i(t)$ 和 $r(t)$ 的变化情况 ($t \rightarrow \infty$ 时它们的极限值分别记作 s_{∞} , i_{∞} 和 r_{∞}).

1. 不论初始条件 s_0, i_0 如何, 病人终将消失, 即

$$i_{\infty} = 0 \quad (18)$$

其证明如下.

首先, 由 (14), $\frac{ds}{dt} \leq 0$, 而 $s(t) \geq 0$, 故 s_{∞} 存

在; 由 (13), $\frac{dr}{dt} \geq 0$, 而 $r(t) \leq 1$, 故 r_{∞} 存在; 再

由 (12) 知 i_{∞} 存在.

其次, 若 $i_{\infty} = \varepsilon > 0$, 则由 (12), 对于充分大的 t 有 $\frac{dr}{dt} > \mu \frac{\varepsilon}{2}$, 这将导致 $r_{\infty} = \infty$, 与 r_{∞} 存在相矛盾.

从图形上看, 不论相轨线从 P_1 或从 P_2 点出发, 它终将与 s 轴相交 (t 充分大).

2. 最终未被感染的健康者的比例是 s_{∞} , 在 (17) 式中令 $i = 0$ 得到, s_{∞} 是方程

$$s_0 + i_0 - s_{\infty} + \frac{1}{\sigma} \ln \frac{s_{\infty}}{s_0} = 0 \quad (19)$$

在 $(0, 1/\sigma)$ 内的根. 在图形上, s_{∞} 是相轨线与 s 轴在 $(0, 1/\sigma)$ 内交点的横坐标.

3. 若 $s_0 > 1/\sigma$, 则 $i(t)$ 先增加, 当 $s = 1/\sigma$ 时, $i(t)$ 达到最大值

$$i_m = s_0 + i_0 - \frac{1}{\sigma} (1 + \ln \sigma s_0) \quad (20)$$

然后 $i(t)$ 减小且趋于 0, $s(t)$ 则单调减小至 s_{∞} , 如图 9 中由 $P_1(s_0, i_0)$ 出发的轨线.

4. 若 $s_0 \leq 1/\sigma$, 则 $i(t)$ 单调减小至零, $s(t)$ 单调减小至 s_{∞} , 如图 9 中由 $P_2(s_0, i_0)$ 出发的轨线. (3, 4 的证明留作习题 1)

可以看出, 如果仅当病人比例 $i(t)$ 有一段增长的时期才认为传染病在蔓延, 那么 $1/\sigma$ 是一个阈值, 当 $s_0 > 1/\sigma$ (即 $\sigma > 1/s_0$) 时传染病就会蔓延. 而减小传染期接触数 σ , 即提高阈值 $1/\sigma$, 使得 $s_0 \leq 1/\sigma$ (即 $\sigma \leq 1/s_0$), 传染病就不会蔓延 (健康者比例的初始值 s_0 是一定的, 通常可认为 s_0 接近 1).

并且, 即使 $s_0 > 1/\sigma$, 从 (19), (20) 式可以看出, σ 减小时, s_{∞} 增加 (通过作图分析), i_m 降低, 也控制了蔓延的程度. 我们注意到, 在 $\sigma = \lambda/\mu$ 中, 人们的卫生水

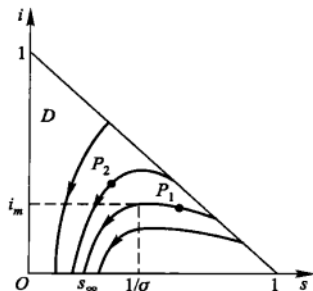


图 9 SIR 模型的相轨线

平越高,日接触率 λ 越小;医疗水平越高,日治愈率 μ 越大,于是 σ 越小,所以提高卫生水平和医疗水平有助于控制传染病的蔓延。

从另一方面看, $\sigma s = \lambda s \cdot 1/\mu$ 是传染期内一个病人传染的健康者的平均数,称为交换数,其含义是一个病人被 σs 个健康者交换。所以当 $s_0 \leq 1/\sigma$, 即 $\sigma s_0 \leq 1$ 时,必有 $\sigma s \leq 1$ 。既然交换数不超过 1,病人比例 $i(t)$ 绝不会增加,传染病不会蔓延。

群体免疫和预防 根据对 SIR 模型的分析,当 $s_0 \leq 1/\sigma$ 时传染病不会蔓延。所以为制止蔓延,除了提高卫生和医疗水平,使阈值 $1/\sigma$ 变大以外,另一个途径是降低 s_0 ,这可以通过比如预防接种使群体免疫的办法做到。

忽略病人比例的初始值 i_0 ,有 $s_0 = 1 - r_0$ 。于是传染病不会蔓延的条件 $s_0 \leq 1/\sigma$ 可以表为

$$r_0 \geq 1 - \frac{1}{\sigma} \quad (21)$$

这就是说,只要通过群体免疫使初始时刻的移出者比例(即免疫者比例) r_0 满足(21)式,就可以制止传染病的蔓延。

这种办法生效的前提条件是免疫者要均匀分布在全体人口中,实际上这是很难做到的。据估计,当时印度等国天花传染病的接触数 $\sigma = 5$,由(21)式至少要有 80% 的人接受免疫才行。据世界卫生组织报告,即使花费大量资金提高 r_0 ,也因很难做到免疫者的均匀分布,使得天花直到 1977 年才在全世界根除。而有些传染病的 σ 更高,根除就更加困难。

数值验证与估量 根据上面的分析,制止传染病蔓延有两种手段,一是提高卫生水平和医疗水平,即降低日接触率 λ ,提高日治愈率 μ ;二是群体免疫,即提高移出者比例的初值 r_0 (相当于降低健康者比例的初值 s_0)。下面作一点数值计算,验证并估量这两种办法的效果。不妨用最终未感染的健康者的比例 s_∞ 和病人比例的最大值 i_m ,作为传染病蔓延程度的度量指标。

给定不同的 λ, μ, s_0, i_0 ,用(19)式计算 s_∞ ,用(20)式计算 i_m (当 $s_0 > 1/\sigma$),结果列入表 2。

表 2 s_∞ 和 i_m 的计算结果

λ	μ	$1/\sigma$	s_0	i_0	s_∞	i_m
1.0	0.3	0.3	0.98	0.02	0.039 8	0.344 9
0.6	0.3	0.5	0.98	0.02	0.196 5	0.163 5
0.5	0.5	1.0	0.98	0.02	0.812 2	0.020 0
0.4	0.5	1.25	0.98	0.02	0.917 2	0.020 0

续表

λ	μ	$1/\sigma$	s_0	i_0	s_∞	i_∞
1.0	0.3	0.3	0.70	0.02	0.084 0	0.168 5
0.6	0.3	0.5	0.70	0.02	0.305 6	0.051 8
0.5	0.5	1.0	0.70	0.02	0.652 8	0.020 0
0.4	0.5	1.25	0.70	0.02	0.675 5	0.020 0

可以看出,对于一定的 s_0 ,降低 λ ,提高 μ ,会使 s_∞ 变大, i_∞ 变小;对于一定的 λ, μ ,降低 s_0 (即提高 r_0),也会使 s_∞ 变大(但是 $s_0 \leq 1/\sigma$ 时 s_∞ 反而小了,你能解释吗?), i_∞ 变小.当然, $s_0 \leq 1/\sigma$ 时, i_∞ 始终等于 i_0 ,即传染病不会蔓延.

我们看到在 SIR 模型中, $\sigma = \lambda/\mu$ 是一个重要参数,实际上 λ, μ 很难估计,而当一次传染病结束以后,可以获得 s_0 和 s_∞ ,在(19)式中略去很小的 i_0 ,即有

$$\sigma = \frac{\ln s_0 - \ln s_\infty}{s_0 - s_\infty} \quad (22)$$

当同样的传染病到来时,如果估计 λ, μ 没有多大变化,那么就可以用上面得到的 σ 分析这次传染病的蔓延过程.

模型验证 上世纪初在印度孟买发生的一次瘟疫中几乎所有病人都死亡了.死亡相当于移出传染系统,有关部门记录了每天移出者的人数,即有了 $\frac{dr}{dt}$ 的实际数据,Kermack 等人用这组数据对 SIR 模型作了验证.

首先,由方程(12),(14)可以得到(习题2)

$$s(t) = s_0 e^{-\sigma r(t)} \quad (23)$$

$$\frac{dr}{dt} = \mu(1 - r - s_0 e^{-\sigma r}) \quad (24)$$

当 $r \leq 1/\sigma$ 时,取(24)式右端 $e^{-\sigma r}$ 的 Taylor 展开的前3项,在初始值 $r_0 = 0$ 下的解为

$$r(t) = \frac{1}{s_0 \sigma^2} \left[(s_0 \sigma - 1) + \alpha \tanh\left(\frac{\alpha \mu t}{2} - \varphi\right) \right] \quad (25)$$

其中 $\alpha^2 = (s_0 \sigma - 1)^2 + 2s_0 i_0 \sigma^2$, $\tanh \varphi = \frac{s_0 \sigma - 1}{\alpha}$.从(25)式容易算出

$$\frac{dr}{dt} = \frac{\alpha^2 \mu}{2s_0 \sigma^2 \coth^2\left(\frac{\alpha \mu t}{2} - \varphi\right)} \quad (26)$$

然后取定参数 s_0, σ 等,画出(26)式的图形,如图10中的曲线,实际数据在图中用圆点表示,可以看出,理论曲线与实际数据吻合得相当不错.

被传染比例的估计 在一次传染病的传播过程中,被传染人数的比例是健康者人数比例的初始值 s_0 与 s_∞ 之差,记作 x ,即

$$x = s_0 - s_\infty \quad (27)$$

当 i_0 很小, s_0 接近于 1 时,由 (19) 式可得

$$x + \frac{1}{\sigma} \ln \left(1 - \frac{x}{s_0} \right) \approx 0 \quad (28)$$

取对数函数 Taylor 展开的前 2 项,有

$$x \left(1 - \frac{1}{s_0 \sigma} - \frac{x}{2s_0^2 \sigma} \right) \approx 0 \quad (29)$$

记 $s_0 = \frac{1}{\sigma} + \delta$, δ 可视该地区人口比例超过阈值 $1/\sigma$ 的部分. 当 $\delta \ll 1/\sigma$ 时, (29) 式给出

$$x \approx 2s_0 \sigma \left(s_0 - \frac{1}{\sigma} \right) \approx 2\delta \quad (30)$$

这个结果表明,被传染人数比例约为 δ 的 2 倍. 对一种传染病,当该地区的卫生和医疗水平不变,即 σ 不变时,这个比例就不会改变. 而当阈值 $1/\sigma$ 提高时, δ 减小,于是这个比例就会降低.

评注 本节介绍的传染病模型从几个方面很好地体现了模型的改进,建模的目的性,以及方法的配合.

第一,最初建立的模型 1 基本上不能用,修改假设后得到的模型 2 虽有所改进,但仍不符合实际. 进一步修改假设,并针对不同情况建立的模型 3,4 才是比较成功的.

第二,模型 3,4 的可取之处在于它们比较全面地达到了建模的目的,即描述传播过程,分析感染人数的变化规律,预测传染病高潮到来时刻,度量传染病蔓延的程度并探索制止蔓延的手段.

第三,对于比较复杂的模型 4,采用了数值计算、图形观察与理论分析相结合的方法,先有感性认识(表 1,图 7,图 8),再用相轨线作理论分析,最后进行数值验证和估算. 可以看作计算机技术与建模方法的巧妙配合.

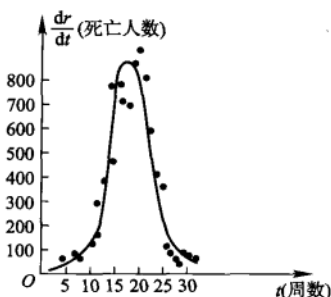


图 10 SIR 模型的理论曲线
与实际数据

5.2 经济增长模型

发展经济、提高生产力主要有以下手段:增加投资、增加劳动力、技术革新. 这里暂不考虑技术革新的作用,一是因为在经济发展的初期(如资本主义早期社会)或者在不太长的时期内,技术相对稳定,二是由于技术革新量化比较困难.

本节的模型将首先建立产值与资金、劳动力之间的关系,然后研究资金与劳动力的最佳分配,使投资效益最大,最后讨论如何调节资金与劳动力的增长率,使劳动生产率得到有效的增长^[55].

1. Douglas 生产函数

用 $Q(t)$, $K(t)$, $L(t)$ 分别表示某一地区或部门在时刻 t 的产值、资金和劳动力,它们的关系可以一般地记作

$$Q(t) = F(K(t), L(t)) \quad (1)$$

其中 F 为待定函数. 对于固定的时刻 t , 上述关系可写作

$$Q = F(K, L) \quad (2)$$

为寻求 F 的函数形式, 引入记号

$$z = Q/L, \quad y = K/L \quad (3)$$

z 是每个劳动力的产值, y 是每个劳动力的投资. 如下的假设是合理的: z 随着 y 的增加而增长, 但增长速度递减. 进而简化地把这个假设表示为

$$z = cg(y), \quad g(y) = y^\alpha, \quad 0 < \alpha < 1 \quad (4)$$

显然函数 $g(y)$ 满足上面的假设, 常数 $c > 0$ 可看成技术的作用. 由 (3), (4) 即可得到 (2) 式中 F 的具体形式为

$$Q = cK^\alpha L^{1-\alpha}, \quad 0 < \alpha < 1 \quad (5)$$

由 (5) 式容易知道, Q 有如下性质

$$\frac{\partial Q}{\partial K}, \frac{\partial Q}{\partial L} > 0, \quad \frac{\partial^2 Q}{\partial K^2}, \frac{\partial^2 Q}{\partial L^2} < 0 \quad (6)$$

请读者解释 (6) 式的含义.

记 $Q_K = \frac{\partial Q}{\partial K}$, Q_K 表示单位资金创造的产值; $Q_L = \frac{\partial Q}{\partial L}$, Q_L 表示单位劳动力创造的产值, 则从 (5) 式可得

$$\frac{KQ_K}{Q} = \alpha, \quad \frac{LQ_L}{Q} = 1 - \alpha, \quad KQ_K + LQ_L = Q \quad (7)$$

(7) 式可解释为: α 是资金在产值中占有的份额, $1 - \alpha$ 是劳动力在产值中占有的份额. 于是 α 的大小直接反映了资金、劳动力二者对于创造产值的轻重关系.

(5) 式是经济学中著名的 Cobb-Douglas 生产函数, 它经受了资本主义社会一些实际数据的检验. 更一般形式的生产函数表为

$$Q = cK^\alpha L^\beta, \quad 0 < \alpha, \beta < 1 \quad (8)$$

2. 资金与劳动力的最佳分配

这里将根据生产函数 (5) 式, 讨论怎样分配资金和劳动力, 使生产创造的效益最大.

假定资金来自贷款,利率为 r ,每个劳动力需付工资 w ,于是当资金 K 、劳动力 L 产生产值 Q 时,得到的效益为

$$S = Q - rK - wL \quad (9)$$

问题化为求资金与劳动力的分配比例 K/L (即每个劳动力占有的资金),使效益 S 最大.

这个模型用微分法即可解得

$$\frac{Q_K}{Q_L} = \frac{r}{w} \quad (10)$$

再利用(7)式,有

$$\frac{K}{L} = \frac{\alpha}{1-\alpha} \frac{w}{r} \quad (11)$$

这就是资金与劳动力的最佳分配.从(11)式可以看出,当 α, w 变大、 r 变小时,分配比例 K/L 变大,这是符合常识的.

3. 劳动生产率增长的条件

常用的衡量经济增长的指标,一是总产值 $Q(t)$,二是每个劳动力的产值 $z(t) = Q(t)/L(t)$,这个模型讨论 $K(t), L(t)$ 满足什么条件,才能使 $Q(t), z(t)$ 保持增长.

首先需要对资金和劳动力的增加作出合理的简化假设:

- 1) 投资增长率与产值成正比,比例系数 $\lambda > 0$,即用一定比例扩大再生产;
- 2) 劳动力的相对增长率为常数 μ , μ 可以是负数,表示劳动力减少.

这两个条件的数学表达式分别为

$$\frac{dK}{dt} = \lambda Q, \quad \lambda > 0 \quad (12)$$

$$\frac{dL}{dt} = \mu L \quad (13)$$

方程(13)的解是^①

$$L(t) = L_0 e^{\mu t} \quad (14)$$

将(4),(5)代入(12)式,得

$$\frac{dK}{dt} = c\lambda L y^\alpha \quad (15)$$

注意到(3)式,有 $K = Ly$,再用(13)式可得

$$\frac{dK}{dt} = L \frac{dy}{dt} + \mu Ly \quad (16)$$

比较(15),(16),得到关于 $y(t)$ 的方程

① 以下凡字母下标为 0 者均表示该变量的初值.

$$\frac{dy}{dt} + \mu y = c\lambda y^\alpha \quad (17)$$

这是著名的 Bernoulli 方程, 它的解是

$$y(t) = \left\{ \frac{c\lambda}{\mu} \left[1 - \left(1 - \mu \frac{K_0}{\dot{K}_0} \right) e^{-(1-\alpha)\mu t} \right] \right\}^{\frac{1}{1-\alpha}} \quad (18)$$

以下根据(18)式研究 $Q(t)$, $z(t)$ 保持增长的条件.

1) $Q(t)$ 增长, 即 $\frac{dQ}{dt} > 0$, 由 $Q = cLy^\alpha$ 及(13), (17)式可算得

$$\frac{dQ}{dt} = cL\alpha y^{\alpha-1} \frac{dy}{dt} + c\mu Ly^\alpha = cLy^{2\alpha-1} [c\lambda\alpha + \mu(1-\alpha)y^{1-\alpha}] \quad (19)$$

将其中的 y 以(18)式代入, 可知条件 $\frac{dQ}{dt} > 0$ 等价于

$$\left(1 - \mu \frac{K_0}{\dot{K}_0} \right) e^{-(1-\alpha)\mu t} < \frac{1}{1-\alpha} \quad (20)$$

因为上式右端大于1, 所以当 $\mu \geq 0$ (即劳动力不减少) 时(20)式恒成立; 而当 $\mu < 0$ 时, (20)式成立的条件是

$$t < \frac{1}{(1-\alpha)\mu} \ln \left[(1-\alpha) \left(1 - \mu \frac{K_0}{\dot{K}_0} \right) \right] \quad (21)$$

说明如果劳动力减少, $Q(t)$ 只能在有限时间内保持增长. 但应注意, 若(21)式中的 $(1-\alpha) \left(1 - \mu \frac{K_0}{\dot{K}_0} \right) \geq 1$, 则不存在这样的增长时段.

2) $z(t)$ 增长, 即 $\frac{dz}{dt} > 0$, 由 $z = cy^\alpha$ 知, 相当于 $\frac{dy}{dt} > 0$, 由方程(17)知, 当 $\mu \leq 0$

时该条件恒成立; 而当 $\mu > 0$ 时, 由(18)式可得, $\frac{dy}{dt} > 0$ 等价于

$$\left(1 - \mu \frac{K_0}{\dot{K}_0} \right) e^{-(1-\alpha)\mu t} > 0 \quad (22)$$

显然, 此式成立的条件为 $\mu \frac{K_0}{\dot{K}_0} < 1$, 即

$$\mu < \dot{K}_0 / K_0 \quad (23)$$

这个条件的含义是, 劳动力增长率小于初始投资增长率.

评注 Douglas 生产函数是计量经济学中重要的数学模型, 本节给出它的一种简洁的建模过程. 在此基础上讨论的资金与劳动力的最佳分配, 是一个静态模

型. 而利用微分方程研究的劳动生产率增长的条件, 是一个动态模型, 虽然它的推导过程稍繁, 但其结果却相当简明, 并且可以给出合理的解释.

5.3 正规战与游击战

早在第一次世界大战期间, F. W. Lanchester 就提出了几个预测战争结局的数学模型, 其中有描述传统的正规战争的, 也有考虑稍微复杂的游击战争的, 以及双方分别使用正规部队和游击部队的所谓混合战争的. 后来人们对这些模型作了改进和进一步的解释, 用以分析历史上一些著名的战争, 如二次世界大战中的美日硫磺岛之战和 1975 年结束的越南战争.

Lanchester 提出的模型是非常简单的, 他只考虑双方兵力的多少和战斗力的强弱, 并且, 当时使用的只是枪炮之类的武器. 兵力因战斗减员和非战斗减员而减少, 又可由后备力量的增援而增加; 战斗力即杀伤对方的能力, 则与射击率 (单位时间的射击次数)、射击命中率以及战争的类型 (正规战、游击战) 等有关. 这些模型当然没有考虑交战双方的政治、经济、社会等因素. 而仅靠战场上兵力的优劣是很难估计战争胜负的, 所以我们认为用这些模型判断整个战争的结局是不可能的, 但是对于局部战役来说或许还有参考价值. 更重要的是, 建模的思路和方法为我们借助数学模型讨论社会科学领域中的实际问题提供了可以借鉴的示例^[15, 49].

一般战争模型 用 $x(t)$ 和 $y(t)$ 表示甲乙交战双方时刻 t 的兵力, 不妨视为双方的士兵人数. 假设:

1. 每一方的战斗减员率取决于双方的兵力和战斗力, 甲乙方的战斗减员率分别用 $f(x, y)$ 和 $g(x, y)$ 表示.
 2. 每一方的非战斗减员率 (由疾病、逃跑等因素引起) 只与本方的兵力成正比.
 3. 甲乙双方的增援率是给定的函数, 分别用 $u(t)$ 和 $v(t)$ 表示.
- 由此可以写出关于 $x(t), y(t)$ 的微分方程为

$$\begin{cases} \dot{x}(t) = -f(x, y) - \alpha x + u(t), & \alpha > 0 \\ \dot{y}(t) = -g(x, y) - \beta y + v(t), & \beta > 0 \end{cases} \quad (1)$$

下面针对不同的战争类型讨论战斗减员率 f, g 的具体表示形式, 并分析影响战争结局的因素.

正规战争模型 甲乙双方都用正规部队作战. 我们只需分析甲方的战斗减员率 $f(x, y)$.

甲方士兵公开活动, 处于乙方每一个士兵的监视和杀伤范围之内, 一旦甲方某个士兵被杀伤, 乙方的火力立即集中在其余士兵身上, 所以甲方的战斗减员率

只与乙方兵力有关,可以简单地设 f 与 y 成正比,即 $f = ay$. a 表示乙方平均每个士兵对甲方士兵的杀伤率(单位时间的杀伤数),称乙方的战斗有效系数. a 可以进一步分解为 $a = r_y p_y$, 其中 r_y 是乙方的射击率(每个士兵单位时间的射击次数), p_y 是每次射击的命中率.

类似地,有 $g = bx$, 且甲方的战斗有效系数 $b = r_x p_x$, r_x 和 p_x 是甲方的射击率和命中率. 于是,在这个模型中方程(1)化为

$$\begin{cases} \dot{x} = -ay - \alpha x + u(t) \\ \dot{y} = -bx - \beta y + v(t) \end{cases} \quad (2)$$

在分析战争结局时忽略非战斗减员一项(与战斗减员相比,这项很小),并且假设双方都没有增援. 记双方的初始兵力分别是 x_0 和 y_0 , 方程(2)简化为

$$\begin{cases} \dot{x} = -ay \\ \dot{y} = -bx \\ x(0) = x_0, y(0) = y_0 \end{cases} \quad (3)$$

由方程(3)可知,双方的兵力 $x(t)$, $y(t)$ 都是单调减函数,不妨认为兵力先减至 0 的一方为负方. 为了得到双方胜负的条件,不必直接求解方程(3),而在相平面上讨论相轨线的变化规律. 由方程(3)可得

$$\frac{dy}{dx} = \frac{bx}{ay} \quad (4)$$

其解为

$$ay^2 - bx^2 = k \quad (5)$$

注意到方程(3)的初始条件,有 $k = ay_0^2 - bx_0^2$ (6)

由(5)式确定的相轨线是双曲线族,如图 1. 箭头表示随时间 t 的增加, $x(t)$, $y(t)$ 的变化趋势. 可以看出,如果 $k > 0$, 轨线将与 y 轴相交. 这就是说存在 t_1 , 使 $x(t_1) = 0, y(t_1) = \sqrt{\frac{k}{a}} > 0$, 即当甲方兵力为 0 时乙方兵力为正值,表明乙方获胜. 同理可知, $k < 0$ 时甲方获胜,而当 $k = 0$ 时双方战平.

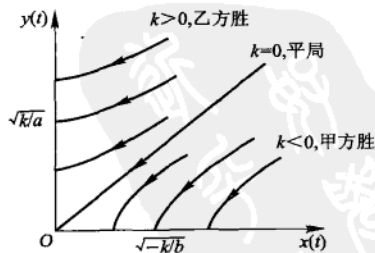


图 1 正规战争模型的相轨线

进一步分析某一方比如乙方取胜的条件. 由(6)式并注意到 a, b 的含义, 乙方获胜的条件可表为

$$\left(\frac{y_0}{x_0}\right)^2 > \frac{b}{a} = \frac{r_x p_x}{r_y p_y} \quad (7)$$

(7) 式说明, 双方初始兵力之比 y_0/x_0 以平方关系影响着战争的结局. 例如, 若乙方兵力增加到原来的 2 倍 (甲方不变), 则影响战争结局的能力增加到 4 倍. 或者说, 若甲方的战斗力比如射击率 r_x 增加到原来的 4 倍 (p_x, r_y, p_y 均不变), 那么为了与此相抗衡, 乙方只需将初始兵力 y_0 增加到原来的 2 倍. 由于这个原因, 正规战争模型称为平方律模型.

游击战争模型 双方都用游击部队作战.

甲方士兵在乙方士兵看不到的某个面积为 s_x 的隐蔽区域内活动, 乙方士兵不是向甲方士兵开火, 而是向这个隐蔽区域射击, 并且不知道杀伤情况. 这时甲方战斗减员率不仅与乙方兵力有关, 而且随着甲方兵力的增加而增加. 因为在有限区域内, 士兵越多, 被杀伤的就越多. 这样可以简单地假设 $f = cxy$, 且乙方战斗有效系数 c 可表为 $c = r_y p_y = r_y \frac{s_{ry}}{s_x}$, 其中 r_y 仍为射击率, 而命中率 p_y 等于乙方一次射击的有效面积 s_{ry} 与甲方活动面积 s_x 之比.

类似地, 有 $g = dxy, d = r_x p_x = r_x \frac{s_{rx}}{s_y}$. 于是在这个模型中方程 (1) 化为

$$\begin{cases} \dot{x} = -cxy - \alpha x + u(t) \\ \dot{y} = -dxy - \beta y + v(t) \end{cases} \quad (8)$$

忽略 αx 和 βy 并设 $u = v = 0$, 在初始条件下 (8) 式为

$$\begin{cases} \dot{x} = -cxy \\ \dot{y} = -dxy \\ x(0) = x_0, y(0) = y_0 \end{cases} \quad (9)$$

与正规战争模型中方程 (3) 的解法类似, 方程 (9) 的解为

$$cy - dx = m \quad (10)$$

$$m = cy_0 - dx_0 \quad (11)$$

(10) 式确定的相轨线是直线族, 如图 2. 像分析正规战争模型一样, 可知 $m > 0$ 时乙方胜, $m < 0$ 时甲方胜, $m = 0$ 时战平.

乙方获胜的条件还可以表为

$$\frac{y_0}{x_0} > \frac{d}{c} = \frac{r_x s_{rx} s_x}{r_y s_{ry} s_y} \quad (12)$$

即初始兵力之比 y_0/x_0 以线性关系影响战争结局, 并且当射击率和射击有效面积一定时, 增加活动面积 s_y 与增加初始兵力 y_0 起着同样的作用. 这个模型又称线性律模型.

混合战争模型 甲方为游击部队, 乙方为正规部队.

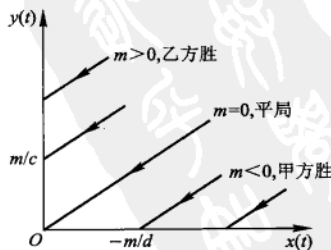


图 2 游击战争模型的相轨线

根据对正规战争和游击战争模型的分析 and 假设, $f = cxy$, $g = bx$, 在同样的忽略和假设下, 方程为

$$\begin{cases} \dot{x} = -cxy \\ \dot{y} = -bx \\ x(0) = x_0, y(0) = y_0 \end{cases} \quad (13)$$

它的相轨线

$$cy^2 - 2bx = n \quad (14)$$

$$n = cy_0^2 - 2bx_0 \quad (15)$$

是抛物线, 如图 3. 可以知道, $n > 0$ 时乙方胜, $n < 0$ 甲方胜, $n = 0$ 时双方战平. 并且乙方 (正规部队一方) 取胜的条件可表为

$$\left(\frac{y_0}{x_0}\right)^2 > \frac{2b}{cx_0} \quad (16)$$

以 $b = r_x p_x$, $c = r_y \frac{s_y}{s_x}$ 代入得

$$\left(\frac{y_0}{x_0}\right)^2 > \frac{2r_x p_x s_x}{r_y s_y x_0} \quad (17)$$

假定以正规部队作战的乙方火力较强, 以游击部队作战的甲方虽火力较弱, 但活动范围较大. 利用 (17) 式可以估计出乙方为了取胜需投入多大的初始兵力. 为确定起见, 不妨设甲方兵力 $x_0 = 100$, 命中率 $p_x = 0.1$, 火力 r_x 是乙方火力 r_y 的一半, 活动区域面积 $s_x = 0.1 \text{ km}^2$, 乙方每次射击的有效面积 $s_y = 1 \text{ m}^2$, 那么由 (17) 式, 乙方取胜的条件为

$$\left(\frac{y_0}{x_0}\right)^2 > \frac{2 \cdot 0.1 \cdot 0.1 \times 10^6}{2 \cdot 1 \cdot 100} = 100 \quad (18)$$

即 $y_0/x_0 > 10$, 乙方必须 10 倍于甲方的兵力方可取胜.

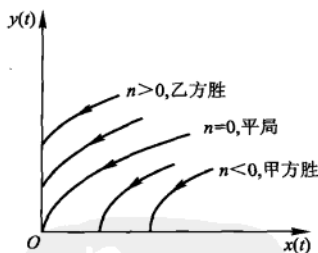


图 3 混合战争模型的相轨线

美国人曾用这个模型分析 20 世纪 70 年代的越南战争 (甲方为越南, 乙方为美国). 根据类似于上面的计算以及在这之前发生在马来西亚、菲律宾、印尼、老挝等地的混合战争的实际情况估计出: 正规部队一方要想取胜, 必须至少投入 8 倍于游击部队一方的兵力, 而美国最多只能派出 6 倍于越南的兵力. 越南战争的结局是美国不得不接受和谈并撤军, 越南人民取得最后胜利.

硫磺岛战役 J. H. Engel 用二次大战末期美日硫磺岛战役中的美军战地记录, 对正规战争模型进行了验证, 发现模型结果与实际数据吻合得很好.

硫磺岛位于东京以南 660 英里 (1 英里 = 1.609 km) 的海面上, 是日军的重

要空军基地. 美军在 1945 年 2 月 19 日开始进攻, 激烈的战斗持续了一个月, 双方伤亡惨重, 日方守军 21 500 人全部阵亡或被俘, 美方投入兵力 73 000 人, 伤亡 20 265 人. 战争进行到 28 天时美军宣布占领该岛, 实际战斗到 36 天才停止. 美军的战地记录有按天统计的战斗减员和增援情况. 日军没有后援, 战地记录则全部遗失.

用 $A(t)$ 和 $J(t)$ 表示美军和日军第 t 天的人数, 在正规战争模型 (2) 式中忽略非战斗减员, 且 $v=0$, 再加上初始条件, 有

$$\begin{cases} \frac{dA}{dt} = -aJ(t) + u(t) \\ \frac{dJ}{dt} = -bA(t) \\ A(0)=0, J(0)=21\,500 \end{cases} \quad (19)$$

美军战地记录给出增援率 $u(t)$ 为

$$u(t) = \begin{cases} 54\,000, & 0 \leq t < 1 \\ 6\,000, & 2 \leq t < 3 \\ 13\,000, & 5 \leq t < 6 \\ 0, & \text{其他} \end{cases} \quad (20)$$

并可由每天伤亡记录得到实际兵力 $A(t)$, $t=1, 2, \dots, 36$ (见图 4 中虚线). 下面利用这些实际数据代入 (19) 式, 算出 $A(t)$ 的理论值, 并与实际值比较.

对方程 (19) 用求和代替积分, 可得

$$A(t) = A(0) - a \sum_{\tau=1}^t J(\tau) + \sum_{\tau=1}^t u(\tau) \quad (21)$$

$$J(t) = J(0) - b \sum_{\tau=1}^t A(\tau) \quad (22)$$

为估计 b , 在 (22) 式中 $t=36$, 由 $A(t)$ 的实际数据可得 $\sum_{\tau=1}^{36} A(\tau) = 2\,037\,000$, 于

是由 $J(36)=0, J(0)=21\,500$, 估计出 $b = \frac{21\,500}{2\,037\,000} = 0.010\,6$. 再把这个值代入

(22) 式, 即可算出 $J(t)$, $t=1, 2, \dots, 36$.

然后从 (21) 式估计 a . 令 $t=36$, 得

$$a \approx \frac{\sum_{\tau=1}^{36} u(\tau) - A(36)}{\sum_{\tau=1}^{36} J(\tau)} \quad (23)$$

其中分子是美军的总伤亡人数, 为 20 265 人, 分母可由已经算出的 $J(t)$ 得到, 为

372 500 人, 于是从 (23) 式有 $a = \frac{20\,265}{372\,500} = 0.054\,4$. 把这个值代入 (21) 式, 得

$$A(t) = -0.0544 \sum_{\tau=1}^t J(\tau) + \sum_{\tau=1}^t u(\tau) \quad (24)$$

由(24)式就能够算出美军人数 $A(t)$ 的理论值,图4中用实线画出.与虚线表示的实际值相比,可以看出吻合的情况.

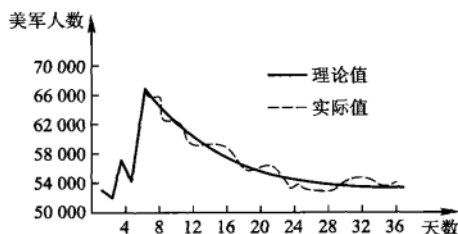


图4 美军兵力实际数据与理论结果的比较

5.4 药物在体内的分布与排除

药物进入机体后,在随血液输送到各个器官和组织的过程中,不断地被吸收、分布、代谢,最终排出体外.药物在血液中的质量浓度,即单位体积血液(单位:ml)中药物含量(单位:mg 或 μg),称**血药浓度**,随时间和空间(机体的各部分)而变化.血药浓度的大小直接影响到药物的疗效,浓度太低不能达到预期的效果,浓度太高又可能导致药物中毒、副作用太强或造成浪费.因此研究药物在体内吸收、分布和排除的动态过程,及这些过程与药理反应间的定量关系,对于新药研制、剂量确定、给药方案设计等药理学和临床医学的发展都具有重要的指导意义和实用价值.这个学科分支称**药物动力学**.

建立**房室模型**(Compartment Model)是药物动力学研究上述动态过程的基本步骤之一.所谓房室是指机体的一部分,药物在一个房室内呈均匀分布,即血药浓度是常数,而在不同房室之间则按照一定规律进行药物的转移.一个机体分为几个房室,要看不同药物的吸收、分布、排除过程的具体情况,以及研究对象所要求的精度而定.本节只讨论二室模型,即将机体分为血液较丰富的**中心室**(包括心、肺、肾等器官)和血液较贫乏的**周边室**(四肢、肌肉组织等).药物的动态过程在每个房室内是一致的,转移只在两个房室之间以及某个房室与体外之间进行.二室模型的建立和求解方法可以推广到多室模型.显然,将一个机体划分为若干房室是人们为了研究目的所作的简化.值得庆幸的是,这种简化在一定条件下已由临床试验证明是正确的,为医学界和药理学界所接受^[96].

模型假设 可以想到,对于二室模型,我们将建立关于两个血药浓度的微分方程描述其动态特性.为了将问题进一步简化,得到线性常系数方程,作如下

假设:

1. 机体分为中心室(Ⅰ室)和周边室(Ⅱ室),两个室的容积(即血液体积或药物分布容积)在过程中保持不变.

2. 药物从一室向另一室的转移速率,以及向体外的排除速率,与该室的血药浓度成正比.

3. 只有中心室与体外有药物交换,即药物从体外进入中心室,最后又从中心室排出体外.与转移和排除的数量相比,药物的吸收可以忽略.

在这些假设下的一种二室模型示意图如图1所示. $c_i(t)$, $x_i(t)$ 和 V_i 分别表示第 i 室($i=1,2$)的血药浓度、药量和容积, k_{12} 和 k_{21} 是两室之间药物转移速率系数, k_{13} 是药物从Ⅰ室向体外排除的速率系数. $f_0(t)$ 是给药速率,由给药方式和剂量确定(下面将详细讨论). 这种速率系数为常数的房室模型称乳突状模型.

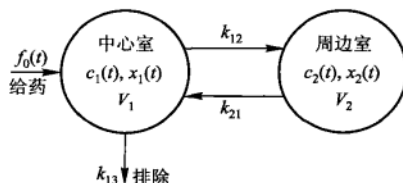


图1 常用的一种二室模型

模型建立 根据假设条件和图1,可以写出两个房室中药量 $x_1(t)$, $x_2(t)$ 满足的微分方程. $x_1(t)$ 的变化率由Ⅰ室向Ⅱ室的转移 $-k_{12}x_1$ 、Ⅰ室向体外的排除 $-k_{13}x_1$ 、Ⅱ室向Ⅰ室的转移 $k_{21}x_2$ 及给药 $f_0(t)$ 组成; $x_2(t)$ 的变化率由Ⅰ室向Ⅱ室的转移 $k_{12}x_1$ 及Ⅱ室向Ⅰ室的转移 $-k_{21}x_2$ 组成. 于是有

$$\begin{cases} \dot{x}_1(t) = -k_{12}x_1 - k_{13}x_1 + k_{21}x_2 + f_0(t) \\ \dot{x}_2(t) = k_{12}x_1 - k_{21}x_2 \end{cases} \quad (1)$$

$x_i(t)$ 与血药浓度 $c_i(t)$ 、房室容积 V_i 之间显然有关系式

$$x_i(t) = V_i c_i(t), i=1,2 \quad (2)$$

(2) 代入(1)式可得

$$\begin{cases} \dot{c}_1(t) = -(k_{12} + k_{13})c_1 + \frac{V_2}{V_1}k_{21}c_2 + \frac{f_0(t)}{V_1} \\ \dot{c}_2(t) = \frac{V_1}{V_2}k_{12}c_1 - k_{21}c_2 \end{cases} \quad (3)$$

这是线性常系数非齐次方程,它的解由齐次方程的通解和非齐次方程的特解组成. 其对应齐次方程的通解为(习题5)

$$\begin{cases} \bar{c}_1(t) = A_1 e^{-\alpha t} + B_1 e^{-\beta t} \\ \bar{c}_2(t) = A_2 e^{-\alpha t} + B_2 e^{-\beta t} \end{cases} \quad (4)$$

其中 α, β 由

$$\begin{cases} \alpha + \beta = k_{12} + k_{21} + k_{13} \\ \alpha\beta = k_{21}k_{13} \end{cases} \quad (5)$$

确定. 为了得到非齐次方程的特解从而解出(3), 需要设定给药速率 $f_0(t)$ 和初始条件. 我们考察下面几种常见的给药方式.

1. 快速静脉注射

这种注射可简化为在 $t=0$ 的瞬间将剂量 D_0 的药物输入中心室, 血药浓度立即上升为 D_0/V_1 , 于是 $f_0(t)$ 和初始条件为

$$f_0(t) = 0, \quad c_1(0) = \frac{D_0}{V_1}, \quad c_2(0) = 0 \quad (6)$$

方程(3)在条件(6)下的解为

$$c_1(t) = Ae^{-\alpha t} + Be^{-\beta t}, \quad A = \frac{D_0(k_{21} - \alpha)}{V_1(\beta - \alpha)}, \quad B = \frac{D_0(\beta - k_{21})}{V_1(\beta - \alpha)} \quad (7)$$

$$c_2(t) = \frac{D_0 k_{12}}{V_2(\beta - \alpha)}(e^{-\alpha t} - e^{-\beta t}) \quad (8)$$

其中 α, β 由(5)确定. 可以看出当 $t \rightarrow \infty$ 时, $c_1(t) \rightarrow 0, c_2(t) \rightarrow 0$.

2. 恒速静脉滴注

当静脉滴注的速率为常数 k_0 时, $f_0(t)$ 和初始条件为

$$f_0(t) = k_0, \quad c_1(0) = 0, \quad c_2(0) = 0 \quad (9)$$

方程(3)在条件(9)下的解可表示为

$$\begin{cases} c_1(t) = A_1 e^{-\alpha t} + B_1 e^{-\beta t} + \frac{k_0}{k_{13} V_1} \\ c_2(t) = A_2 e^{-\alpha t} + B_2 e^{-\beta t} + \frac{k_{12} k_0}{k_{21} k_{13} V_2} \\ A_2 = \frac{V_1(k_{12} + k_{13} - \alpha)}{k_{21} V_2} A_1, B_2 = \frac{V_1(k_{12} + k_{13} - \beta)}{k_{21} V_2} B_1 \end{cases} \quad (10)$$

其中常数 A_1, B_1 由初始条件 $c_1(0) = c_2(0) = 0$ 确定.

当 t 充分大时, $c_1(t), c_2(t)$ 将趋向于(10)式右端第3项表示的常数值. 实际上, 若 $t = T$ 后停止滴注, 那么 $c_1(t), c_2(t)$ 在 $t > T$ 以后将按指数规律衰减并趋于0.

3. 口服或肌肉注射

这种给药方式相当于在药物输入中心室之前先有一个将药物吸收入血液的过程, 可以简化为有一个吸收室, 如图2. $x_0(t)$ 为吸收室的药量, 药物由吸收室进入中心室的转移速率系数为 k_{01} , 于是 $x_0(t)$ 满足

$$\begin{cases} \dot{x}_0(t) = -k_{01}x_0 \\ x_0(0) = D_0 \end{cases} \quad (11)$$

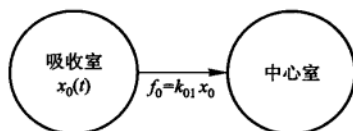


图2 药物经吸收室进入中心室

D_0 是给药量, 而药物进入中心室的速率为

$$f_0(t) = k_{01}x_0(t) \quad (12)$$

将方程(11)的解代入(12)式, 得

$$f_0(t) = D_0 k_{01} e^{-k_{01}t} \quad (13)$$

在这种情况下, 方程(3)的解 $c_1(t)$ 的一般形式为

$$c_1(t) = A e^{-\alpha t} + B e^{-\beta t} + E e^{-k_{01}t} \quad (14)$$

(设 $k_{01} \neq \alpha, \beta$), 其中系数 A, B, E 由初始条件 $c_1(0) = c_2(0) = 0$ 确定.

从以上的讨论可以看出, 中心室的血药浓度 $c_1(t)$ 取决于转移速率系数 k_{12} , k_{21} , k_{13} , 房室容积 V_1, V_2 以及输入参数 D_0, k_0 等因素, 而房室模型的用途恰是通过 $c_1(t)$ 的量测, 确定对于药理学和临床医学最为重要的参数, 如转移速率系数, 特别是从中心室向体外排除的速率系数 k_{13} . 下面介绍在快速静脉注射给药方式下估计诸参数的方法.

参数估计 在 $t=0$ 瞬时快速注射剂量为 D_0 的药物以后, 在一系列时刻 $t_i (i=1, 2, \dots, n)$ 从中心室采取血样并获得血药浓度 $c_1(t_i)$, 根据这些数据利用(7)式和(5)式估计参数 k_{12}, k_{21}, k_{13} 的过程可分两步: 先计算(7)式中的 α, β, A, B , 再确定 k_{12}, k_{21}, k_{13} .

1. 计算 α, β, A, B

不妨设 $\alpha < \beta$, 于是当 t 充分大时(7)式近似为

$$c_1(t) = A e^{-\alpha t} \quad (15)$$

或

$$\ln c_1(t) = \ln A - \alpha t \quad (16)$$

对于适当的 t_i 和相应的 $c_1(t_i)$, 用最小二乘法不难估计出 $\alpha, \ln A$ 和 A .

然后计算

$$\tilde{c}_1(t) = c_1(t) - A e^{-\alpha t} \quad (17)$$

再利用(7)式得

$$\ln \tilde{c}_1(t) = \ln B - \beta t \quad (18)$$

对于较小的 t_i 和由 (17) 式算出的 $\bar{c}_1(t_i)$, 仍用最小二乘法即可得到 β 和 B .

2. 确定 k_{12}, k_{21}, k_{13}

因为 $t \rightarrow \infty$ 时, $c_1(t), c_2(t) \rightarrow 0$, 进入中心室的药物全部被排除, 所以

$$D_0 = k_{13} V_1 \int_0^{\infty} c_1(t) dt \quad (19)$$

将 (7) 代入 (19) 式, 可得

$$D_0 = k_{13} V_1 \left(\frac{A}{\alpha} + \frac{B}{\beta} \right) \quad (20)$$

又因为

$$c_1(0) = \frac{D_0}{V_1} = A + B \quad (21)$$

联合 (20), (21) 式解出

$$k_{13} = \frac{\alpha\beta(A+B)}{\alpha B + \beta A} \quad (22)$$

再利用 (5) 式即可确定

$$k_{21} = \frac{\alpha\beta}{k_{13}} \quad (23)$$

$$k_{12} = \alpha + \beta - k_{13} - k_{21} \quad (24)$$

这就完成了根据中心室血药浓度的量测数据, 估计转移和排除速率系数的过程.

评注 建立房室模型的目的是研究体内血药浓度的变化过程, 确定诸如转移和排除速率系数等参数, 为制订给药方案和剂量大小提供数量依据. 建模过程是将机理分析和测试分析相结合, 先由机理分析确定方程形式, 再由测试数据估计参数.

选用几个房室建模是一个重要问题, 可以先选择一室模型 (习题 6), 其计算非常简单. 不满意时再采用二室或多室模型, 甚至非线性房室模型. 常见的一种非线性模型 (以一室为例) 是 $\dot{c}_1(t) = -\frac{k_1 c_1}{k_2 + c_1}$, 当 c_1 较小时, 它近似于线性模型, 称为一级排除过程, 而当 c_1 较大时, $\dot{c}_1(t)$ 近似于常数, 称为零级排除过程, 所以它表示了一种混合型的排除过程.

5.5 香烟过滤嘴的作用

尽管科学家们对于吸烟的危害提出了许多无可辩驳的证据, 不少国家的政府和有关部门也一直致力于减少或禁止吸烟, 但是仍有不少人不愿放弃对香烟的嗜好. 香烟制造商既要满足瘾君子的需要, 又要顺应减少吸烟危害的潮流, 还

要获取丰厚的利润,于是普遍地在香烟上安装了过滤嘴.过滤嘴的作用到底有多大,与使用的材料和过滤嘴的长度有什么关系?要从定量的角度回答这些问题,就要建立一个描述吸烟过程的数学模型,分析人体吸入的毒物数量与哪些因素有关,以及它们之间的数量表达式^[52].

吸烟时毒物吸入人体的过程大致是这样的:毒物基本上均匀地分布在烟草中,吸烟时点燃处的烟草大部分化为烟雾,毒物由烟雾携带着一部分直接进入空气,另一部分沿香烟穿行.在穿行过程中又部分地被未点燃的烟草和过滤嘴吸收而沉积下来,剩下的进入人体.被烟草吸收而沉积下来的那一部分毒物,当香烟燃烧到那里的时候又通过烟雾部分进入空气,部分沿香烟穿行,这个过程一直持续到香烟燃烧至过滤嘴处为止.于是我们看到,原来分布在烟草中的毒物除了进入空气和被过滤嘴吸收的一部分外,剩下的全都被人体吸入.

实际的吸烟过程非常复杂并且因人而异.点燃处毒物随烟雾进入空气和沿香烟穿行的数量比例,与吸烟的方式、环境等多种因素有关;烟雾穿过香烟的速度随着吸烟动作的变化而不断地改变;过滤嘴和烟草对毒物的吸收作用也会随烟雾穿行速度等因素的影响而有所变化.如果要考虑类似于上面这些复杂情况,将使我们寸步难行.为了能建立一个初步的模型,可以设想一个机器人在典型的环境下吸烟,“他”吸烟的动作、方式及外部环境在整个过程中不变,于是可以认为毒物随烟雾进入空气和沿香烟穿行的数量比例、烟雾穿行的速度、过滤嘴和烟草对毒物的吸收率等在吸烟过程中都是常数.

模型假设 基于上述分析,这个模型的假设条件如下:

1. 烟草和过滤嘴的长度分别是 l_1 和 l_2 ,香烟总长 $l = l_1 + l_2$,毒物 M (单位:mg) 均匀分布在烟草中,密度为 $w_0 = M/l_1$.
2. 点燃处毒物随烟雾进入空气和沿香烟穿行的数量比例是 a' : $a' + a = 1$.
3. 未点燃的烟草和过滤嘴对随烟雾穿行的毒物的吸收率(单位时间内毒物被吸收的比例)分别是常数 b 和 β .
4. 烟雾沿香烟穿行的速度是常数 v ,香烟燃烧速度是常数 u ,且 $v \gg u$.

将一支烟吸完后毒物进入人体的总量(不考虑从空气的烟雾中吸入的)记作 Q ,在建立模型以得到 Q 的数量表达式之前,让我们先根据常识分析一下 Q 应与哪些因素有关,采取什么办法可以降低 Q .

首先,提高过滤嘴吸收率 β 、增加过滤嘴长度 l_2 、减少烟草中毒物的初始含量 M ,显然可以降低吸入毒物量 Q .其次,当毒物随烟雾沿香烟穿行的比例 a 和烟雾速度 v 减小时,预料 Q 也会降低.至于在假设条件中涉及的其他因素,如烟草对毒物的吸收率 b 、烟草长度 l_1 、香烟燃烧速度 u 对 Q 的影响就不容易估计了.

下面通过建模对这些定性分析和提出的问题作出定量的验证和回答.

模型建立 设 $t=0$ 时在 $x=0$ 处点燃香烟, 坐标系如图 1 所示. 吸入毒物量 Q 由毒物穿过香烟的流量确定, 后者又与毒物在烟草中的密度有关, 为研究这些关系, 定义两个基本函数:

毒物流量 $q(x, t)$ 表示时刻 t 单位时间内通过香烟截面 x 处 ($0 \leq x \leq l$) 的毒物量;

毒物密度 $w(x, t)$ 表示时刻 t 截面 x 处单位长度烟草中的毒物含量 ($0 \leq x \leq l_1$). 由假设 1, $w(x, 0) = w_0$.

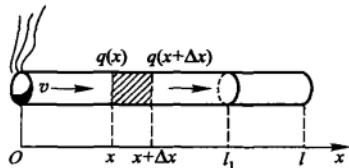


图 1 $x=0$ 处点燃的香烟

如果知道了流量函数 $q(x, t)$, 吸入毒物量 Q 就是 $x=l$ 处的流量在吸一支烟时间内的总和. 注意到关于烟草长度和香烟燃烧速度的假设, 我们得到

$$Q = \int_0^T q(l, t) dt, \quad T = l_1/u \quad (1)$$

下面分 4 步计算 Q .

1. 求 $t=0$ 瞬间由烟雾携带的毒物单位时间内通过 x 处的数量 $q(x, 0)$. 由假设 4 中关于 $v \gg u$ 的假定, 可以认为香烟点燃处 $x=0$ 静止不动.

为简单起见, 记 $q(x, 0) = q(x)$, 考察 $(x, x + \Delta x)$ 一段香烟 (图 1), 毒物通过 x 和 $x + \Delta x$ 处的流量分别是 $q(x)$ 和 $q(x + \Delta x)$, 根据守恒定律, 这两个流量之差应该等于这一段未点燃的烟草或过滤嘴对毒物的吸收量, 于是由假设 2 和 4 有

$$q(x) - q(x + \Delta x) = \begin{cases} bq(x)\Delta\tau, & 0 \leq x \leq l_1, \\ \beta q(x)\Delta\tau, & l_1 < x \leq l, \end{cases} \quad \Delta\tau = \frac{\Delta x}{v}$$

其中 $\Delta\tau$ 是烟雾穿过 Δx 所需时间. 令 $\Delta\tau \rightarrow 0$, 得到微分方程

$$\frac{dq}{dx} = \begin{cases} -\frac{b}{v}q(x), & 0 \leq x \leq l_1 \\ -\frac{\beta}{v}q(x), & l_1 < x \leq l \end{cases} \quad (2)$$

在 $x=0$ 处点燃的烟草单位时间内放出的毒物量记作 H_0 , 根据假设 1, 3, 4 可以写出方程 (2) 的初始条件为

$$q(0) = aH_0, \quad H_0 = uw_0 \quad (3)$$

求解 (2)、(3) 式时, 先解出 $q(x)$ ($0 \leq x \leq l_1$), 再利用 $q(x)$ 在 $x=l_1$ 处的连续性确定 $q(x)$ ($l_1 \leq x \leq l$). 其结果为

$$q(x) = \begin{cases} aH_0 e^{-\frac{bx}{v}}, & 0 \leq x \leq l_1 \\ aH_0 e^{-\frac{bl_1}{v}} e^{-\frac{\beta(x-l_1)}{v}}, & l_1 < x \leq l \end{cases} \quad (4)$$

2. 在香烟燃烧过程的任意时刻 t , 求毒物单位时间内通过 $x=l$ 的数量

$q(l, t)$.

因为在时刻 t 香烟燃至 $x = ut$ 处, 记此时点燃的烟草单位时间放出的毒物量为 $H(t)$, 则

$$H(t) = uw(ut, t) \quad (5)$$

根据与第 1 步完全相同的分析和计算, 可得

$$q(x, t) = \begin{cases} aH(t)e^{-\frac{b(x-ut)}{v}}, & ut \leq x \leq l_1 \\ aH(t)e^{-\frac{b(l_1-ut)}{v}}e^{-\frac{\beta(l_1-l_1)}{v}}, & l_1 < x \leq l \end{cases} \quad (6)$$

实际上, 在(4)式中将坐标原点平移至 $x = ut$ 处即可得到(6)式. 由(5)、(6)式能够直接写出

$$q(l, t) = auw(ut, t)e^{-\frac{b(l_1-ut)}{v}}e^{-\frac{\beta l_2}{v}} \quad (7)$$

3. 确定 $w(ut, t)$.

因为在吸烟过程中未点燃的烟草不断地吸收烟雾中的毒物, 所以毒物在烟草中的密度 $w(x, t)$ 由初始值 w_0 逐渐增加. 考察烟草截面 x 处 Δt 时间内毒物密度的增量 $w(x, t + \Delta t) - w(x, t)$, 根据守恒定律它应该等于单位长度烟雾中的毒物被吸收的部分, 按照假设 2, 4, 有

$$w(x, t + \Delta t) - w(x, t) = b \frac{q(x, t)}{v} \Delta t$$

令 $\Delta t \rightarrow 0$, 并将(5)、(6)式代入得

$$\begin{cases} \frac{\partial w}{\partial t} = \frac{abu}{v} w(ut, t) e^{-\frac{b(x-ut)}{v}} \\ w(x, 0) = w_0 \end{cases} \quad (8)$$

方程(8)的解为(推导见^[52]).

$$\begin{cases} w(x, t) = w_0 \left[1 + \frac{a}{a'} e^{-\frac{bx}{v}} \left(e^{\frac{bat}{v}} - e^{\frac{abut}{v}} \right) \right] \\ w(ut, t) = \frac{w_0}{a'} \left(1 - ae^{-\frac{a'bat}{v}} \right) \end{cases} \quad (9)$$

其中 $a' = 1 - a$ (假设 2).

4. 计算 Q .

将(9)代入(7)式, 得

$$q(l, t) = \frac{auw_0}{a'} e^{-\frac{bl_1}{v}} e^{-\frac{\beta l_2}{v}} \left(e^{-\frac{bat}{v}} - ae^{-\frac{abut}{v}} \right) \quad (10)$$

最后将(10)代入(1)式, 作积分得到

$$Q = \int_0^{l_1/ut} q(l, t) dt = \frac{aw_0v}{a'b} e^{-\frac{\beta l_2}{v}} \left(1 - e^{-\frac{a'bl_1}{v}} \right) \quad (11)$$

为便于下面的分析,将上式化作

$$Q = aMe^{-\frac{\beta l_2}{v}} \cdot \frac{1 - e^{-\frac{a'bl_1}{v}}}{\frac{a'bl_1}{v}} \quad (12)$$

记

$$r = \frac{a'bl_1}{v}, \quad \varphi(r) = \frac{1 - e^{-r}}{r} \quad (13)$$

则(12)式可写作

$$Q = aMe^{-\frac{\beta l_2}{v}} \varphi(r) \quad (14)$$

(13)、(14)式是我们得到的最终结果,表示了吸入毒物量 Q 与 $a, M, \beta, l_2, v, b, l_1$ 等诸因素之间的数量关系。

结果分析

1. Q 与烟草含毒物量 M 、毒物随烟雾沿香烟穿行比例 a 成正比^①。设想将毒物 M 集中在 $x=l$ 处,则吸入量为 aM 。

2. 因子 $e^{-\frac{\beta l_2}{v}}$ 体现了过滤嘴减少毒物进入人体的作用,提高过滤嘴吸收率 β 和增加长度 l_2 能够对 Q 起到负指数衰减的效果,并且 β 和 l_2 在数量上增加一定比例时起的作用相同。降低烟雾穿行速度 v 也可减少 Q 。设想将毒物 M 集中在 $x=l_1$ 处,利用上述建模方法不难证明,吸入毒物量为 $aMe^{-\frac{\beta l_2}{v}}$ 。

3. 因子 $\varphi(r)$ 表示的是由于未点燃烟草对毒物的吸收而起到的减少 Q 的作用。虽然被吸收的毒物还要被点燃,随烟雾沿香烟穿行而部分地进入人体,但是因为烟草中毒物密度 $w(x, t)$ 越来越高,所以按照固定比例跑到空气中的毒物增多,相应地减少了进入人体的毒物量。

根据实际资料 $r = \frac{a'bl_1}{v} \ll 1$, 在(13)式 $\varphi(r)$ 中的 e^{-r} 取 Taylor 展开的前 3 项可得 $\varphi(r) \approx 1 - r/2$, 于是(14)式为

$$Q \approx aMe^{-\frac{\beta l_2}{v}} \left(1 - \frac{a'bl_1}{2v} \right) \quad (15)$$

可知,提高烟草吸收率 b 和增加长度 l_1 (毒物量 M 不变)对减少 Q 的作用是线性的,与 β 和 l_2 的负指数衰减作用相比,效果要小得多。

4. 为了更清楚地了解过滤嘴的作用,不妨比较两支香烟,一支是上述模型讨论的,另一支长度为 l , 不带过滤嘴,参数 w_0, b, a, v 与第一支相同,并且吸到 $x=l_1$ 处就扔掉。

^① 这里忽略 $\varphi(r)$ 中的 $a'(=1-a)$, 因为 $\varphi(r)$ 起的作用较小,见结果分析 3。

吸第一支和第二支烟进入人体的毒物量分别记作 Q_1 和 Q_2 , Q_1 当然可由 (11) 式给出, Q_2 也不必重新计算, 只需把第二支烟设想成吸收率为 b (与烟草相同) 的假过滤嘴香烟就行了, 这样由 (11) 式可以直接写出

$$Q_2 = \frac{aw_0 v}{a'b} e^{-\frac{bl_2}{v}} (1 - e^{-\frac{a'bl_1}{v}}) \quad (16)$$

与 (11) 式给出的 Q_1 相比, 我们得到

$$\frac{Q_1}{Q_2} = e^{-\frac{(\beta-b)l_2}{v}} \quad (17)$$

所以只要 $\beta > b$, 就有 $Q_1 < Q_2$, 过滤嘴是起作用的. 并且, 提高吸收率之差 $\beta - b$ 与加长过滤嘴长度 l_2 , 对于降低比例 Q_1/Q_2 的效果相同. 不过提高 β 需要研制新材料, 将更困难一些.

评注 这个模型在基本合理的简化假设下, 运用精确的数学工具解决了一个粗看起来不易下手的实际问题. 从提出假设、引入两个基本函数 $q(x, t)$ 和 $w(x, t)$, 到运用物理学上常用的守恒定律建立微分方程, 从而构造出动态模型, 最后到对结果的分析, 整个过程可以说是用建模方法解决实际问题的一个范例.

5.6 人口的预测和控制

人类社会进入 20 世纪以来, 在科学技术和生产力飞速发展的同时, 世界人口也以空前的规模增长, 统计数据如表 1.

表 1 世界人口

年	1625	1804	1927	1960	1974	1987	1999
人口/亿	5	10	20	30	40	50	60

可以看出, 人口每增加 10 亿的时间, 由 100 年缩短为 30 年和十几年, 我们赖以生存的地球, 已经携带着它的 60 亿子民踏入 21 世纪.

长期以来, 人类的繁殖一直在自发地进行着. 只是由于人口数量的迅速膨胀和环境质量的急剧恶化, 人们才猛然醒悟, 开始研究人类和自然的关系、人口数量的变化规律, 以及如何进行人口控制等问题.

我国是世界第一人口大国, 地球上每 5 个人中就有一个中国人. 在 20 世纪的一段时间内, 我国人口的增长速度过快, 见表 2.

表 2 中国人口

年	1901	1929	1953	1965	1982	1990	2000
人口/亿	4.26	5.48	6.02	7.25	10.32	11.30	12.95

有效地控制我国人口的增长,不仅是深入贯彻落实科学发展观,到 2020 年实现全面建成小康社会的需要,而且对于全人类的美好理想来说,也是我们义不容辞的责任。

认识人口数量的变化规律,建立人口模型,作出较准确的预报,是有效控制人口增长的前提。长期以来,人们在这方面作了不少工作,下面先介绍两个最基本的人口模型,并利用表 3 给出的近两个世纪的美国人口统计数据作模型参数估计、检验和预报,最后介绍考虑年龄结构和生育模式的人口模型。

表 3 美国人口

年	1790	1800	1810	1820	1830	1840	1850	1860
人口/百万	3.9	5.3	7.2	9.6	12.9	17.1	23.2	31.4
年	1870	1880	1890	1900	1910	1920	1930	1940
人口/百万	38.6	50.2	62.9	76.0	92.0	106.5	123.2	131.7
年	1950	1960	1970	1980	1990	2000		
人口/百万	150.7	179.3	204.0	226.5	251.4	281.4		

一、指数增长模型

最简单的人口增长模型是人所共知的:记今年人口为 x_0 , k 年后人口为 x_k , 年增长率为 r , 则

$$x_k = x_0(1+r)^k \quad (1)$$

显然,这个公式的基本条件是年增长率 r 保持不变。

200 多年前英国人口学家 T. Malthus (1766—1834) 调查了英国 100 多年的人口统计资料,得出了人口增长率不变的假设,并据此建立了著名的人口指数增长模型。

记时刻 t 的人口为 $x(t)$, 当考察一个国家或一个较大地区的人口时, $x(t)$ 是一个很大的整数。为了利用微积分这一数学工具,将 $x(t)$ 视为连续、可微函数。记初始时刻 ($t=0$) 的人口为 x_0 。假设人口增长率为常数 r , 即单位时间内 $x(t)$ 的增量 $\frac{dx}{dt}$ 等于 r 乘以 $x(t)$, 于是得到 $x(t)$ 满足微分方程

$$\frac{dx}{dt} = rx, x(0) = x_0 \quad (2)$$

由这个方程很容易解出

$$x(t) = x_0 e^{rt} \quad (3)$$

$r > 0$ 时, (3) 式表示人口将按指数规律随时间无限增长, 称为指数增长模型。

思考 说明我们常用的预报公式(1)就是指数增长模型(3)的离散近似形式.

历史上,指数增长模型与19世纪以前欧洲一些地区人口统计数据可以很好地吻合,迁往加拿大的欧洲移民后代人口也大致符合这个模型.另外,用它作短期人口预测可以得到较好的结果.显然,这是因为在这些情况下,人口增长率是常数这个基本假设大致成立.

但是长期来看,任何地区的人口都不可能无限增长,即指数模型不能描述、也不能预测较长时期的人口演变过程.这是因为,人口增长率事实上是在不断地变化着.排除灾难、战争等特殊时期,一般说来,当人口较少时,增长较快,即增长率较大;人口增加到一定数量以后,增长就会慢下来,即增长率变小.如果根据表3的数据计算一下美国人口的年增长率,可以看到增长率从19世纪开始就基本上在缓慢下降.如果用一个平均的年增长率作为 r ,用指数增长模型描述美国人口的变化,会发现结果与表3的实际数据相差很大.

看来,为了使人口预报特别是长期预报更好地符合实际情况,必须修改指数增长模型关于人口增长率是常数这个基本假设^[34].

二、阻滞增长模型——logistic 模型

分析人口增长到一定数量后增长率下降的主要原因,人们注意到,自然资源、环境条件等因素对人口的增长起着阻滞作用,并且随着人口的增加,阻滞作用越来越大.所谓阻滞增长模型就是考虑到这个因素,对指数增长模型的基本假设进行修改后得到的.

阻滞作用体现在对人口增长率 r 的影响上,使得 r 随着人口数量 x 的增加而下降.若将 r 表示为 x 的函数 $r(x)$,则它应是减函数.于是方程(2)写作

$$\frac{dx}{dt} = r(x)x, \quad x(0) = x_0 \quad (4)$$

对 $r(x)$ 的一个最简单的假定是,设 $r(x)$ 为 x 的线性函数,即

$$r(x) = r - sx (r, s > 0) \quad (5)$$

这里 r 称固有增长率,表示人口很少时(理论上是 $x=0$)的增长率.为了确定系数 s 的意义,引入自然资源和环境条件所能容纳的最大人口数量 x_m ,称人口容量.当 $x=x_m$ 时人口不再增长,即增长率 $r(x_m)=0$,代入(5)式得 $s=r/x_m$,于是 $r(x)=r(1-x/x_m)$,将 $r(x)$ 代入方程(4),得

$$\frac{dx}{dt} = rx \left(1 - \frac{x}{x_m}\right), \quad x(0) = x_0 \quad (6)$$

方程(6)右端的因子 rx 体现人口自身的增长趋势,因子 $(1 - \frac{x}{x_m})$ 则体现了环境和资源对人口增长的阻滞作用.显然, x 越大,前一因子越大,后一因子越小,人

口增长是两个因子共同作用的结果,(6)式称为阻滞增长模型.

如果以 x 为横轴,以 $\frac{dx}{dt}$ 为纵轴作出方程(6)的图形(图1),可以分析人口增长速度 $\frac{dx}{dt}$ 随着 x 的增加而变化的情况,从而大致地看出 $x(t)$ 的变化规律.

练习 根据图1 $\frac{dx}{dt}$ 随 x 的变化,分析 x 随 t 的变化规律, x 多大时人口增长最快, $t \rightarrow \infty$ 时 $x \rightarrow ?$, 请读者画出 $x(t)$ 的大致图形.

实际上,方程(6)可以很方便地用分离变量法求解得到

$$x(t) = \frac{x_m}{1 + \left(\frac{x_m}{x_0} - 1 \right) e^{-rt}} \quad (7)$$

用计算机软件画出(7)式的图形,它是一条S形曲线(图2), x 增加得先快后慢, $t \rightarrow \infty$ 时 $x \rightarrow x_m$, 拐点在 $x = x_m/2$, 应该与上面练习中画出的 $x(t)$ 的图形一致.

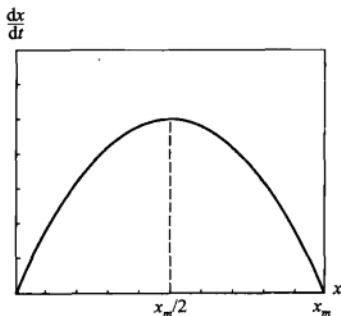


图1 阻滞增长模型 $\frac{dx}{dt} \sim x$ 曲线

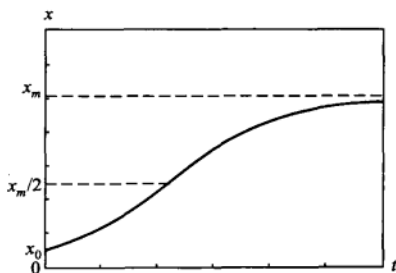


图2 阻滞增长模型 $x \sim t$ 曲线

由方程(6)表示的阻滞增长模型,是荷兰生物数学家 Verhulst 19 世纪中叶提出的.它不仅能够大体上描述人口及许多物种数量(如森林中的树木、鱼塘中的鱼群等)的变化规律,而且在社会经济领域也有广泛的应用,例如耐用消费品的销售量就可以用它来描述.基于这个模型能够描述一些事物符合逻辑的客观规律,人们常称它为 logistic 模型,本书以后的章节将多次用到它^[49].

三、模型的参数估计、检验和预报

用指数增长模型或阻滞增长模型进行人口预报,先要作参数估计.除了初始人口 x_0 外,指数增长模型要估计 r ,阻滞增长模型则要估计 r 和 x_m .它们可以用人口统计数据拟合得到,也可以辅之以专家的估计.

为了估计指数增长模型(2)或(3)中的参数 r 和 x_0 , 需将(3)式取对数, 得

$$y = rt + a, y = \ln x, a = \ln x_0 \quad (8)$$

以美国人口实际数据为例(将表3数据列为表4第1,2列),对(8)式作数据拟合,如用1790年至1900年的数据,得到 $r = 0.2743/(10\text{年})$, $x_0 = 4.1884$; 如用全部数据可得 $r = 0.2022/(10\text{年})$, $x_0 = 6.0450$ 。也可以令 $x_0 = 3.9$ (1790年实际人口),只计算 r 。

用得到的 r 和 x_0 代入(3)式,将计算结果与实际数据作比较。表4中计算人口 x_1 是用1790年至1900年数据拟合的结果, x_2 是用全部数据拟合的结果,图3(a)和图3(b)是它们的图形表示(+号是实际数据,曲线是计算结果)。

表4 指数增长模型和阻滞增长模型对美国人口数据拟合的结果

年	实际人口 /百万	计算人口 x_1 (指数增长模型)	计算人口 x_2 (指数增长模型)	计算人口 x (阻滞增长模型)
1790	3.9	4.2	6.0	3.9
1800	5.3	5.5	7.4	5.0
1810	7.2	7.2	9.1	6.5
1820	9.6	9.5	11.1	8.3
1830	12.9	12.5	13.6	10.7
1840	17.1	16.5	16.6	13.7
1850	23.2	21.7	20.3	17.5
1860	31.4	28.6	24.9	22.3
1870	38.6	37.6	30.5	28.3
1880	50.2	49.5	37.3	35.8
1890	62.9	65.1	45.7	45.0
1900	76.0	85.6	55.9	56.2
1910	92.0		68.4	69.7
1920	106.5		83.7	85.5
1930	123.2		102.5	103.9
1940	131.7		125.5	124.5
1950	150.7		153.6	147.2
1960	179.3		188.0	171.3
1970	204.0		230.1	196.2
1980	226.5		281.7	221.2
1990	251.4		344.8	245.3
2000	281.4		422.1	

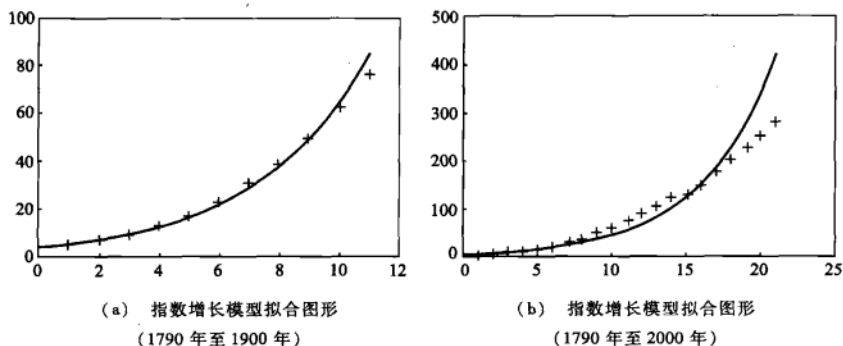


图 3

为了估计阻滞增长模型(6)或(7)中的参数 r 和 x_m , 我们不用(7)式而将方程(6)表为

$$\frac{\frac{dx}{dt}}{x} = r - sx, s = \frac{r}{x_m} \quad (9)$$

(9) 式左端可以从实际人口数据用数值微分算出, 右端对参数 r, s 是线性的. 我们利用 1860 年至 1990 年的数据(去掉个别异常数据)计算得到 $r = 0.2557/(10 \text{ 年})$, $x_m = 392.0886$. 将它们代入(7)式, 计算结果见表 4 最后一列 x_3 和图 4.

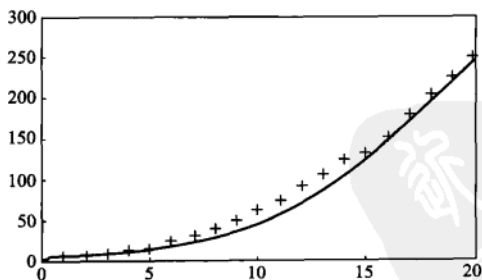


图 4 阻滞增长模型拟合图形(1790 年为起点)

可以看出, 这个模型虽然中间一段(19 世纪中叶到 20 世纪中叶)拟合得不太好, 但是最后一段(20 世纪中叶以后)拟合得不错.

在估计阻滞增长模型的参数时没有用 2000 年的实际数据, 是为了用它作模型检验. 我们用模型计算 2000 年的人口, 与已知的实际数据比较, 来检验模型是否合适.

为简单起见,可利用 $x(1990)$ 和方程(6)作如下计算:

$$x(2000) = x(1990) + \Delta x = x(1990) + rx(1990)[1 - x(1990)/x_m]$$

得到 $x(2000) \approx 274.5$ 百万,与实际数据 281.4 百万的误差约 2.5%,可以认为该模型是相当满意的。

为了预报美国 2010 年的人口,应将 2000 年的实际数据加进去重新估计参数,可得 $r = 0.2490/(10 \text{ 年})$, $x_m = 433.9886$ 。然后再用模型检验中的计算方法进行预报,得到 $x(2010) \approx 306.0$ 百万。这个预报结果的准确性如何? 据美国人口普查局 2010 年 12 月 21 日公布,截止到 2010 年 4 月 1 日,美国总人口为 3.087 亿,预报误差不到 1%。

四、考虑年龄结构和生育模式的人口模型^[88]

指数增长模型和阻滞增长模型都是针对人口总数和总的增长率,不涉及年龄结构。事实上,在人口预测中人口按年龄的分布状况是十分重要的,因为不同年龄人的生育率和死亡率有着很大的差别,两个国家或地区目前人口一样,如果一个国家或地区年轻人的比例明显高于另一国家或地区,那么二者人口的发展状况将大不一样。在考虑年龄结构的人口模型中,除了时间变量外,年龄是另一个变量。

人口发展方程 使人口数量和结构变化的因素不外乎出生、死亡和迁移。为简化起见,只考虑自然的出生与死亡,不计迁移等社会因素的影响。

为研究任意时刻不同年龄的人口数量,引入人口的分布函数和密度函数。时刻 t 年龄小于 r 的人口称为**人口分布函数**,记作 $F(r, t)$,其中 $t, r (\geq 0)$ 均为连续变量,设 F 是连续、可微的。时刻 t 的人口总数记作 $N(t)$,最高年龄记作 r_m ,理论推导时设 $r_m \rightarrow \infty$ 。于是对于非负非降函数 $F(r, t)$,有

$$F(0, t) = 0, \quad F(r_m, t) = N(t) \quad (10)$$

人口密度函数定义为

$$p(r, t) = \frac{\partial F}{\partial r} \quad (11)$$

$p(r, t) dr$ 表示时刻 t 年龄在区间 $[r, r + dr]$ 内的人数。

记 $\mu(r, t)$ 为时刻 t 年龄 r 的人的死亡率,其含义是, $\mu(r, t)p(r, t) dr$ 表示时刻 t 年龄在 $[r, r + dr]$ 内单位时间死亡的人数。

为了得到 $p(r, t)$ 满足的方程,考察时刻 t 年龄在 $[r, r + dr]$ 内的人到时刻 $t + dt$ 的情况。他们中活着的那一部分人的年龄变为 $[r + dr_1, r + dr + dr_1]$, 这里 $dr_1 = dt$ 。而在 dt 这段时间内死亡的人数为 $\mu(r, t)p(r, t) dr dt$ 。于是

$$p(r, t) dr - p(r + dr_1, t + dt) dr = \mu(r, t)p(r, t) dr dt \quad (12)$$

上式可写作

$$[p(r+dr_1, t+dt) - p(r, t+dt)]dr + [p(r, t+dt) - p(r, t)]dr \\ = -\mu(r, t)p(r, t)drdt$$

注意到 $dr_1 = dt$, 就可得到

$$\frac{\partial p}{\partial r} + \frac{\partial p}{\partial t} = -\mu(r, t)p(r, t) \quad (13)$$

这是人口密度函数 $p(r, t)$ 的一阶偏微分方程, 其中死亡率 $\mu(r, t)$ 为已知函数.

方程(13)有两个定解条件: 初始密度函数记作 $p(r, 0) = p_0(r)$; 单位时间出生的婴儿数记作 $p(0, t) = f(t)$, 称婴儿出生率. $p_0(r)$ 可由人口调查资料得到, 是已知函数; $f(t)$ 则对预测和控制人口起着重要作用, 后面将对它进一步分析. 将方程(13)及定解条件写作

$$\begin{cases} \frac{\partial p}{\partial r} + \frac{\partial p}{\partial t} = -\mu(r, t)p(r, t), & t, r > 0 \\ p(r, 0) = p_0(r) \\ p(0, t) = f(t) \end{cases} \quad (14)$$

这个连续型人口发展方程描述了人口的演变过程, 从这个方程确定出密度函数 $p(r, t)$ 以后, 立即可以得到各个年龄的人口数, 即人口分布函数

$$F(r, t) = \int_0^r p(s, t) ds \quad (15)$$

方程(14)的求解过程比较复杂, 这里只给出一种特殊情况下的结果. 在社会安定的局面下和不太长的时间内, 死亡率大致与时间无关, 于是可近似地假设 $\mu(r, t) = \mu(r)$. 这时(14)式的解为

$$p(r, t) = \begin{cases} p_0(r-t)e^{-\int_{r-t}^r \mu(s) ds}, & 0 \leq t \leq r \\ f(t-r)e^{-\int_0^{t-r} \mu(s) ds}, & t > r \end{cases} \quad (16)$$

读者可以验证(16)式满足方程(14). 这个解在 $tO r$ 平面上有一个浅显的解释: 图5中对角线 $r=t$ 将 $tO r$ 平面 ($t, r \geq 0$) 分为两部分, 在 $t < r$ 区域, $p(r, t)$ 完全由年龄为 $r-t$ 的人口初始密度 $p_0(r-t)$ 和这些人的死亡率 $\mu(s)$ ($r-t \leq s < t$) 决定; 而在 $t > r$ 区域, $p(r, t)$ 则由未来的生育状况 $f(t-r)$ 及死亡率 $\mu(s)$ ($0 \leq s < r$) 决定.

生育率和生育模式 在方程(14)或解(16)中 $p_0(r)$ 和 $\mu(r)$ 可从人口统计数据得到, $\mu(r, t)$ 也可由 $\mu(r, 0)$ 粗略估计. 这样, 为了预测和控制人口的发展状况, 人们主要关注和可以用作控制手段的就是婴儿出生率 $f(t)$ 了. 下面对 $f(t)$ 作进一步分解.

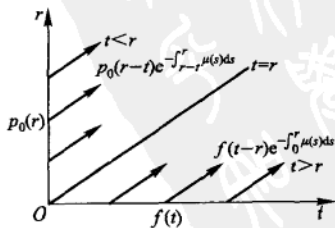


图5 tOr 平面上的 $p(r, t)$

记女性性别比函数为 $k(r, t)$, 即时刻 t 年龄在 $[r, r + dr)$ 的女性人数为 $k(r, t)p(r, t)dr$, 将这些女性在单位时间内平均每人的生育数记作 $b(r, t)$, 设育龄区间为 $[r_1, r_2]$, 则

$$f(t) = \int_{r_1}^{r_2} b(r, t)k(r, t)p(r, t)dr \quad (17)$$

再将 $b(r, t)$ 定义为

$$b(r, t) = \beta(t)h(r, t) \quad (18)$$

其中 $h(r, t)$ 满足

$$\int_{r_1}^{r_2} h(r, t)dr = 1 \quad (19)$$

于是

$$\beta(t) = \int_{r_1}^{r_2} b(r, t)dr \quad (20)$$

$$f(t) = \beta(t) \int_{r_1}^{r_2} h(r, t)k(r, t)p(r, t)dr \quad (21)$$

由(20)式可以看出, $\beta(t)$ 的直接含义是时刻 t 单位时间内平均每个育龄女性的生育数. 如果所有育龄女性在她育龄期所及的时刻都保持这个生育数, 那么 $\beta(t)$ 也表示平均每个女性一生的总和生育数, 所以 $\beta(t)$ 称为总和生育率(简称生育率)或生育胎次.

从(18), (19)两式及 $b(r, t)$ 的含义可以看出, $h(r, t)$ 是年龄为 r 女性的生育加权因子, 称生育模式. 在稳定环境下可以近似地认为它与 t 无关, 即 $h(r, t) = h(r)$. $h(r)$ 表示了在哪些年龄生育率高, 哪些年龄生育率低. 图6给出了 $h(r)$ 的示意图, 表明 $r = r_c$ 附近生育率最高. 由人口统计资料可以知道当前实际的 $h(r, t)$. 作理论分析时, 人们常采用的 $h(r)$ 的一种形式是借用概率论中的 Γ 分布

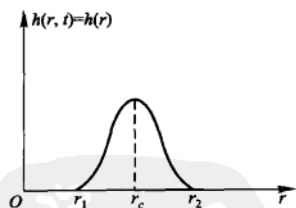


图6 生育模式 $h(r)$ 示意图

$$h(r) = \frac{(r - r_1)^{\alpha-1} e^{-\frac{r-r_1}{\theta}}}{\theta^\alpha \Gamma(\alpha)}, \quad r > r_1 \quad (22)$$

并取 $\theta = 2, \alpha = n/2$, 这时有

$$r_c = r_1 + n - 2 \quad (23)$$

可以看出, 提高 r_1 意味着晚婚, 而增加 n 意味着晚育.

这样, 人口发展方程(14)和单位时间出生的婴儿数 $f(t)$ 的表达式(21), 构成了我们的连续型人口模型. 模型中死亡率函数 $\mu(r, t)$ 、性别比函数 $k(r, t)$ 和初始密度函数 $p_0(r)$ 可以由人口统计资料直接得到, 或在资料的基础上估计, 而生育率 $\beta(t)$ 和生育模式 $h(r, t)$ 则是可以用于控制人口发展过程的两种手段.

$\beta(t)$ 可以控制生育的多少, $h(r,t)$ 可以控制生育的早晚和疏密.我国的计划生育政策正是通过这两种手段实施的(习题9).

从控制论观点看,在方程(14)描述的人口系统中, $p(r,t)$ 可视为状态变量, $p(0,t)=f(t)$ 视为控制变量,是分布参数系统的边界控制函数.(21)式表明控制输入中含有状态变量,形成状态反馈, $\beta(t)$ 视为反馈增益,并且通常是一种正反馈,即人口密度函数 $p(r,t)$ 的增加,通过婴儿出生率 $f(t)$ 又使 $p(r,t)$ 进一步增大(图7).方程的解(16)式中因子 $f(t-r)$ 表明这种反馈还有相当大的滞后作用,所以一旦人口政策失误,使 $p(r,t)$ 在一段时间内增长得过多过快,再想通过控制手段 $\beta(t)$ 和 $h(r,t)$ 把人口增长的势头降下来,就很困难并且常常需要相当长(几代人)的时间.

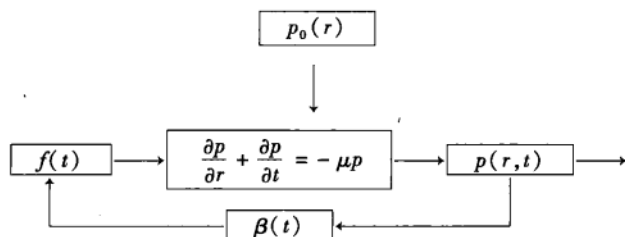


图7 人口正反馈控制系统示意图

人口指数 在上面的模型中,密度函数 $p(r,t)$ 或分布函数 $F(r,t)$ 固然是人口发展过程最完整的描述,但是使用起来并不方便.在人口统计学中,常用一些所谓人口指数来简明扼要地表示一个国家或地区的人口特征.下面是一些人口指数的定义及它们与 $p(r,t)$ 等数量之间的关系.

1. 人口总数

$$N(t) = \int_0^m p(r,t) dr \quad (24)$$

2. 平均年龄

$$R(t) = \frac{1}{N(t)} \int_0^m r p(r,t) dr \quad (25)$$

3. 平均寿命 它表示时刻 t 出生的人不论活到什么时候,死亡率都按时刻 t 的 $\mu(r,t)$ 计算,这些人的平均存活时间,记作 $S(t)$.

$$S(t) = \int_t^{\infty} e^{-\int_0^{\tau-t} \mu(r,t) dr} d\tau \quad (26)$$

$S(t)$ 实际上是预估寿命.通常说目前平均寿命已达到多少岁,是指今年出生婴儿的预估寿命,即 $S(0)$.根据统计资料得到当前的死亡率 $\mu(r,0)$ 后就可以算出 $S(0)$.

4. 老龄化指数 定义为

$$\omega(t) = \frac{R(t)}{S(t)} \quad (27)$$

显然,平均年龄 $R(t)$ 越大, $\omega(t)$ 越大;对于 $R(t)$ 相同的两个国家或地区,平均寿命 $S(t)$ 大的,表示健康水平高,一个人能工作的时间在一生中占的比例大,于是老龄化指数 $\omega(t)$ 较小。

鉴于我国目前人口的年龄结构状况,在控制生育率、降低人口增长速度的同时,需适当考虑不要使老龄化指数变得过高。

5. 依赖性指数

$$\rho(t) = \frac{N(t) - L(t)}{L(t)} \quad (28)$$

$$L(t) = \int_{l_1}^{l_2} [1 - k(r, t)] p(r, t) dr + \int_{l'_1}^{l'_2} k(r, t) p(r, t) dr \quad (29)$$

其中 $[l_1, l_2]$ 和 $[l'_1, l'_2]$ 分别是男性和女性有劳动能力的年龄区间, $L(t)$ 是全体人口中有劳动能力的人数,所以依赖性指数 $\rho(t)$ 表示平均每个劳动者要供养的人数。

5.7 烟雾的扩散与消失

广阔的平原上风日和丽,当一颗炮弹在蔚蓝色的天空中爆炸时我们看到,放出的烟雾以爆炸点为中心向四周迅速扩散,形成一个近于圆形的不透光区域。起初这个区域逐渐增大,后来它的边界变得明亮起来,不透光区域渐渐变小,最后烟雾完全消失,又是晴空万里。本节要建立一个模型描述观察到的烟雾扩散和消失过程,分析消失的时间与哪些因素有关,以及怎样预报消失的时刻^[82]。

问题分析 炮弹爆炸引起的烟雾传播可以看作在无穷空间由瞬时点源导致的扩散过程,能够由二阶抛物型偏微分方程描述烟雾浓度的变化规律。我们用仪器或肉眼观察到的烟雾扩散和消失过程既与烟雾浓度的变化规律有关,又与烟雾对光线的吸收过程有关,还涉及仪器或肉眼对亮暗的灵敏程度。比如用肉眼观察认为烟雾已经消失(由暗变亮),而用灵敏度高的仪器则仍能观测到烟雾(尚未变亮)。

整个建模过程应当包含:烟雾浓度的变化规律;穿过烟雾的光的发光强度的变化规律;仪器辨别亮暗的灵敏度的描述;不透光区域边界的变化过程等。

模型假设

1. 炮弹的爆炸看作在空中某一点向四周等强度地瞬时释放烟雾,烟雾在无穷空间扩散,不计风力和大地的影响。

2. 烟雾的传播服从扩散定律,即单位时间通过单位法向面积的流量与它的浓度梯度成正比。

3. 光线穿过烟雾时其发光强度由于烟雾的吸收而减少,单位距离上发光强度的相对减少量与烟雾浓度成正比;没有烟雾的大气对光线的吸收作用忽略不计。

4. 在烟雾扩散过程中,不穿过烟雾直接进入观测仪器的标准发光强度 I_0 保持不变;对于穿过烟雾进入仪器的发光强度 I ,观测结果只有亮暗之分,仅当 $\frac{I}{I_0} > 1 - \mu$ 时观测结果为亮。 μ 称为仪器的灵敏度, μ 越小仪器越灵敏,通常 $\mu \ll 1$ 。

模型建立

1. 烟雾浓度的变化规律

将爆炸时刻记作 $t = 0$, 爆炸点选为坐标原点。时刻 t 无穷空间中任一点 (x, y, z) 的烟雾浓度记为 $C(x, y, z, t)$ 。根据假设 2, 单位时间通过单位法向面积 S 的流量

$$q = -k \cdot \text{grad } C \quad (1)$$

k 是扩散系数, grad 表示梯度, 负号表示由浓度高向浓度低的地方扩散。考察空间域 Ω , Ω 的体积为 V , 包围 Ω 的曲面为 S , S 的外法线向量为 n , 则在 $[t, t + \Delta t]$ 内通过 Ω 的流量为

$$Q_1 = \int_t^{t+\Delta t} \iint_S q \cdot n d\sigma dt \quad (2)$$

而 Ω 内烟雾的增量为

$$Q_2 = \iiint_V [C(x, y, z, t) - C(x, y, z, t + \Delta t)] dV \quad (3)$$

由质量守恒定律

$$Q_1 = Q_2 \quad (4)$$

根据曲面积分的奥-高公式

$$\iint_S q \cdot n d\sigma = \iiint_V \text{div } q dV \quad (5)$$

其中 div 是散度记号。由 (1) ~ (5) 式, 再利用积分中值定理不难得到

$$\frac{\partial C}{\partial t} = k \text{div}(\text{grad } C) = k \left(\frac{\partial^2 C}{\partial x^2} + \frac{\partial^2 C}{\partial y^2} + \frac{\partial^2 C}{\partial z^2} \right), t > 0, -\infty < x, y, z < \infty \quad (6)$$

这是无界区域的抛物型偏微分方程。根据假设 1, 初始条件为作用在坐标原点的点源函数, 可记作

$$C(x, y, z, 0) = Q\delta(x, y, z) \quad (7)$$

Q 表示炮弹爆炸施放的烟雾总量, $\delta(x, y, z)$ 是单位发光强度的点源函数。

方程 (6) 满足条件 (7) 的解为

$$C(x, y, z, t) = \frac{Q}{(4\pi kt)^{3/2}} e^{-\frac{x^2 + y^2 + z^2}{4kt}} \quad (8)$$

这个结果表明, 对于任意时刻 t , 烟雾浓度 C 的等值面是球面 $x^2 + y^2 + z^2 = R^2$, 并且

随着球面半径 R 的增加, C 的值是连续减少的; 当 $R \rightarrow \infty$ 或 $t \rightarrow \infty$ 时, $C(x, y, z, t) \rightarrow 0$.

2. 穿过烟雾的发光强度的变化规律

考察沿一定方向穿过烟雾的光线, 此方向的长度坐标记为 l , 烟雾浓度为 $C(l)$, 发光强度为 $I(l)$. 按照假设 3, 应有

$$\frac{dI}{dl} = -\alpha C(l)I(l) \quad (9)$$

α 是烟雾对光线的吸收系数. 光线未进入烟雾时 ($l = l_0$) 的发光强度记作 I_0 , 即

$$I(l_0) = I_0 \quad (10)$$

方程(9)在条件(10)下的解为

$$I(l) = I_0 e^{-\alpha \int_{l_0}^l C(s) ds} \quad (11)$$

3. 仪器的灵敏度与不透光区域的边界

从上面分析可知烟雾浓度在空间是连续变化的, 穿过烟雾进入仪器的发光强度也是连续变化的, 那么之所以会观察到烟雾扩散时不透光区域的边界有一个先变大、后缩小、最终消失的过程, 则完全是因为仪器的观测结果只有亮暗之分, 而亮暗分界线由灵敏度 μ 决定. 根据假设 4, 仅当

$$\frac{I}{I_0} > 1 - \mu \quad (12)$$

观察结果为亮, 这时可以认为烟雾已经消失.

按照假设 3, 发光强度 I_0 穿过没有烟雾的大气时, 其衰减可以忽略, 所以不必对它与直接进入仪器的标准发光强度加以区分. 条件(12)中的发光强度 I 可由(11)式计算.

为方便起见, 取沿着 z 轴的光线, 不妨设光源(太阳)在 $z \rightarrow -\infty$ 处而仪器在 $z \rightarrow \infty$ 处. 由(11)式, 条件(12)可以写作

$$e^{-\alpha \int_{-\infty}^{\infty} C(x, y, z, t) dz} > 1 - \mu \quad (13)$$

因为 $C(x, y, z, t)$ 的等值面是球面, 所以仪器观察到的投影在 xy 平面上的不透光区域的边界是圆周, 记作

$$x^2 + y^2 = r^2 \quad (14)$$

这个圆周的半径随时间的变化规律 $r(t)$ 应由条件(13)决定.

4. 不透光区域边界的变化规律

对(13)式取对数并利用近似关系 $\ln(1+x) \approx x$ ($x \ll 1$), (13)式可化为

$$\int_{-\infty}^{\infty} C(x, y, z, t) dz < \frac{1}{\alpha} \ln \frac{1}{1-\mu} \approx \frac{\mu}{\alpha}$$

于是不透光区域的边界由

$$\int_{-\infty}^{\infty} C(x, y, z, t) dx = \frac{\mu}{\alpha} \quad (15)$$

确定. 将(8)式表示的 $C(x, y, z, t)$ 代入(15)进行积分, 并利用公式 $\int_{-\infty}^{\infty} e^{-\frac{x^2}{a}} dx = \sqrt{\pi a}$ 可得

$$\frac{Q}{4\pi kt} e^{-\frac{x^2+y^2}{4kt}} = \frac{\mu}{\alpha} \quad (16)$$

以(14)式代入(16)式, 即可得到不透光区域的半径为

$$r(t) = \sqrt{4kt \ln \frac{\alpha Q}{4\pi k \mu t}} \quad (17)$$

结果分析 由(17)式可以画出 $r(t)$ 的图形如图1, 并且不难算出, 当

$$t = t_1 = \frac{\alpha Q}{4\pi k \mu e} \quad (18)$$

时不透光区域的半径 r 达到最大值 r_m , 当

$$t = t_2 = \frac{\alpha Q}{4\pi k \mu} \quad (19)$$

时 $r=0$, 即 t_2 是观察到烟雾完全消失的时刻.

(18), (19) 式表明, t_1, t_2 与烟雾施放量 Q 和烟雾对光线的吸收系数 α 成正比, 与扩散系数 k 和仪器灵敏度 μ 成反比. 请读者解释这个结果的合理性.

最后, 从(18), (19)式可得

$$t_2 = t_1 \cdot e \approx 2.7 t_1 \quad (20)$$

所以当观察到烟雾扩散区域达到最大的时刻 t_1 以后, 就可以预报烟雾完全消失的时刻 t_2 .

(18) 或 (19) 式的另一个用途是用来确定大气中的扩散系数 k .

评注 烟雾在大气中的扩散, 发光强度通过烟雾后的衰减, 都服从一定的物理定律. 仅从这些规律得到的结果不能解释观察到的烟雾扩散和消失过程. 这个模型引入仪器灵敏度的指标, 成功地解释了上述现象, 得到了基本上符合实际的结论. 而仪器灵敏度本身是客观存在的, 并非人们所虚构.

5.8 万有引力定律的发现

万有引力定律的发现是伟大科学家牛顿的重大贡献之一. 牛顿在研究力学的过程中发明了微积分, 又成功地在开普勒三定律的基础上运用微积分推出了万有引力定律. 这一创造性的成就可以看做是历史上最著名的数学模型之一.

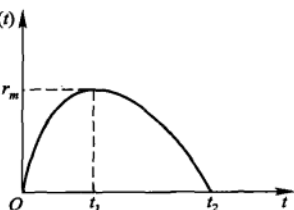


图1 不透光区域半径 $r(t)$ 的示意图

历史背景 15 世纪下半叶开始,欧洲商品经济的繁荣促进了航海业的发展.哥伦布新大陆的发现,麦哲伦的环球远航,引起了社会的普遍关注.当时远洋航船的方位全靠星球的位置来确定.在强大的社会需求推动下,天文观测的精确程度不断提高.在大量的实际观测数据面前,一直处于天文学统治地位的“地心说”开始动摇了.

波兰天文学家哥白尼(1473—1543)在天文观测的基础上,冲破宗教统治和“地心说”的束缚,提出了“日心说”.这是天文学乃至整个科学的一大革命.但是由于历史条件和科学水平的限制,哥白尼的理论还有一些缺陷.他接受了圆周运动是最完善的天体运动形式的概念,认为行星绕太阳的运行轨道是圆形的.

意大利物理学家伽利略(1564—1642)不仅用观察方法证实了哥白尼的学说,而且用实验方法发现了落体定律和惯性原理,揭示了物体在不受阻挠时作匀速直线运动的规律.

德国天文学家、数学家开普勒(1571—1630)在第谷·布拉赫对于行星运动大量观测资料的基础上,用数学方法研究发现,火星的实际位置与按哥白尼理论计算的位置相差 8 弧分($1^\circ = 60$ 弧分).经过对观测数据长期深入的分析,开普勒终于归纳出著名的所谓行星运动三定律,即

各颗行星分别在不同的椭圆轨道上绕太阳运行,太阳位于这些椭圆的一个焦点上;

每颗行星运行过程中单位时间内太阳—行星向径扫过的面积是常数;

各颗行星运行周期的平方与其椭圆轨道长半轴的 3 次方成正比.

在伽利略、开普勒的基础上,17、18 世纪许多科学家致力于行星沿椭圆轨道运行时受力状况的研究.从开普勒定律可以看出,行星运行速度是变化的,而在当时尚没有计算变速运动的动力学方法.英国物理学家胡克(1635—1703)和荷兰物理学家惠更斯(1629—1695)等人虽然都取得了一些成果,但终未得到有关引力的定律.

卓越的英国物理学家、数学家牛顿(1642—1727)认为一切运动都有其力学原因,开普勒三定律的背后必定有力学定律起作用.他在研究变速运动过程中发明了微积分(当时称流数法),又以微积分为工具在开普勒三定律和牛顿第二定律的基础上,用演绎方法得到所谓万有引力定律,于 1687 年汇编入《自然科学之数学原理》出版.这一发现成功地解释了许多自然现象,并为一系列观测和实验进一步证实,直到今天仍是物理学中的一条基本定律.

下面介绍的万有引力定律的推导过程是在牛顿使用的流数法的基础上改写的.

模型假设 开普勒三定律和牛顿第二定律是导出万有引力定律的基础,所以需要将它们表述为这个模型的假设条件.

对于任意一颗行星的椭圆运行轨道,建立极坐标系 (r, θ) ,以太阳为坐标原点 $r=0$,以椭圆长半轴方向为 $\theta=0$,用向径 $\mathbf{r}(t)$ ($r(t), \theta(t)$)表示行星的位置,见图1^①.

1. 轨道方程为

$$r = \frac{p}{1 + e \cos \theta}, \quad p = \frac{b^2}{a}, \quad b^2 = a^2(1 - e^2) \quad (1)$$

a, b 为椭圆的长、短半轴, e 为离心率.

2. 单位时间内向径 \mathbf{r} 扫过的面积是常数 A , 即

$$\frac{1}{2} r^2 \dot{\theta} = A \quad (2)$$

3. 行星运行周期 T 满足

$$T^2 = \lambda a^3 \quad (3)$$

其中 λ 是绝对常数, 与是哪一颗行星无关.

4. 行星运行时受的作用力 \mathbf{f} 等于行星加速度 $\ddot{\mathbf{r}}$ 和质量 m 的乘积, 即

$$\mathbf{f} = m \ddot{\mathbf{r}} \quad (4)$$

模型建立 万有引力定律研究的是力 \mathbf{f} 的方向和大小. 由(4)式可知, 它取决于行星的运动加速度 $\ddot{\mathbf{r}}$, 从而必须研究行星的位置 \mathbf{r} 和速度 $\dot{\mathbf{r}}$. 首先引入基向量(见图1)

$$\begin{cases} \mathbf{u}_r = \cos \theta \mathbf{i} + \sin \theta \mathbf{j} \\ \mathbf{u}_\theta = -\sin \theta \mathbf{i} + \cos \theta \mathbf{j} \end{cases} \quad (5)$$

向径 \mathbf{r} 可表为

$$\mathbf{r} = r \mathbf{u}_r \quad (6)$$

因为由(5)式可以算出

$$\begin{cases} \dot{\mathbf{u}}_r = \dot{\theta} \mathbf{u}_\theta \\ \dot{\mathbf{u}}_\theta = -\dot{\theta} \mathbf{u}_r \end{cases} \quad (7)$$

所以由(6), (7)式得到行星运动的速度和加速度为

$$\dot{\mathbf{r}} = \dot{r} \mathbf{u}_r + r \dot{\theta} \mathbf{u}_\theta \quad (8)$$

$$\ddot{\mathbf{r}} = (\ddot{r} - r \dot{\theta}^2) \mathbf{u}_r + (r \ddot{\theta} + 2 \dot{r} \dot{\theta}) \mathbf{u}_\theta \quad (9)$$

根据(2)式可得

^① 本节用粗体字母表示向量, 以区别于标量.

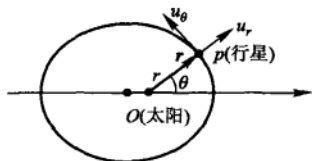


图1 极坐标系中的行星轨道

$$\dot{\theta} = \frac{2A}{r^2}, \quad \ddot{\theta} = \frac{-4A\dot{r}}{r^3} \quad (10)$$

将 $\ddot{\theta}$ 代入(9)式右端, 可知第二项 $r\ddot{\theta} + 2\dot{r}\dot{\theta} = 0$, 于是(9)式简化为

$$\ddot{\mathbf{r}} = (\ddot{r} - r\dot{\theta}^2)\mathbf{u}_r, \quad (11)$$

对(1)式求导并利用(10)式 $\dot{\theta}$ 的结果, 得到

$$\dot{r} = \frac{2Ae}{p} \sin \theta \quad (12)$$

$$\ddot{r} = \frac{4A^2e}{pr^2} \cos \theta = \frac{4A^2(p-r)}{pr^3} \quad (13)$$

将(10)和(13)代入(11)式, 得

$$\ddot{\mathbf{r}} = \frac{-4A^2}{pr^2}\mathbf{u}_r, \quad (14)$$

最后把(14)和(6)代入(4)式, 有

$$\mathbf{f} = -\frac{4A^2m}{pr^2}\mathbf{r}_0, \quad \mathbf{r}_0 = \frac{\mathbf{r}}{r} \quad (15)$$

这里 \mathbf{r}_0 是单位向径, 指示向径方向。(15)式表明, 行星运行时受的力 \mathbf{f} 的方向与它的向径方向 \mathbf{r}_0 相反, 即在太阳—行星连线方向, 指向太阳; f 的大小与行星质量 m 成正比, 与太阳—行星距离 r 的平方成反比。 f 为太阳对行星的引力。

为了完成万有引力的推导, 只需进一步证明(15)式中的 $\frac{A^2}{p}$ 是绝对常数, 即它与是哪一颗行星无关 (A 和 p 不是绝对常数)。

因为 A 是单位时间内向径扫过的面积, 行星运行一个周期 T 向径扫过的面积恰是以 a, b 为长、短半轴的椭圆面积, 所以

$$TA = \pi ab \quad (16)$$

由(1), (3), (16)式容易算出

$$\frac{A^2}{p} = \frac{\pi^2}{\lambda} \quad (17)$$

π 和 λ 均为绝对常数。将(17)代入(15)式有

$$\mathbf{f} = -\frac{4\pi^2m}{\lambda r^2}\mathbf{r}_0 \quad (18)$$

(18)式表明, 太阳对行星的作用力 \mathbf{f} 的大小除了与行星质量 m 成正比、与相互距离平方 r^2 成反比以外, 余下的因子 $4\pi^2/\lambda$ 与行星无关, 它就只可能由太阳本身确定了。将这个结果与我们熟知的形式

$$\mathbf{f} = -k \frac{Mm}{r^2}\mathbf{r}_0 \quad (19)$$

相比较,应该有 $\frac{4\pi^2}{\lambda} = kM$ (k 为万有引力常数, M 为太阳质量). 请读者验证这个关系(习题 10).

评注 从发现万有引力定律的过程可以看出,在正确假设的基础上运用数学演绎方法建模,对自然科学的发展能够发挥多么巨大的作用. 虽然我们大多数人发明不了什么定律,但是学习前辈如何创造性地运用数学方法,对于培养解决实际问题的能力是大有好处的.

习 题

1. 对于 5.1 节传染病的 SIR 模型,证明:

(1) 若 $s_0 > 1/\sigma$, 则 $i(t)$ 先增加, 在 $s = 1/\sigma$ 处最大, 然后减小并趋于 0; $s(t)$ 单调减小至 s_∞ .

(2) 若 $s_0 \leq 1/\sigma$, 则 $i(t)$ 单调减小并趋于 0, $s(t)$ 单调减小至 s_∞ .

2. 对于传染病的 SIR 模型证明(23)~(25)式.

3. 在 5.2 节经济增长模型中,为了适用于不同的对象,可将产量函数 $Q(t)$ 折算成现金,仍用 $Q(t)$ 表示. 考虑到物价上升因素,我们记物价上升指数为 $p(t)$ (设 $p(0) = 1$), 则产品的表面价值 $y(t)$ 、实际价值 $Q(t)$ 和物价指数 $p(t)$ 之间满足 $y(t) = Q(t)p(t)$.

(1) 导出 $y(t)$, $Q(t)$, $p(t)$ 的相对增长率之间的关系,并作出解释.

(2) 设雇用工人数目为 $L(t)$, 每个工人工资 $w(t)$, 企业的利润简化为从产品的收入 $y(t)$ 中扣除工人工资和固定成本. 利用道格拉斯生产函数讨论,企业应雇用多少工人能使利润最大.

4. 在 5.3 节正规战争模型(3)中,设乙方与甲方战斗有效系数之比为 $a/b = 4$, 初始兵力 x_0 与 y_0 相同.

(1) 问乙方取胜时的剩余兵力是多少,乙方取胜的时间如何确定?

(2) 若甲方在战斗开始后有余部部队以不变的速率 r 增援,重新建立模型,讨论如何判断双方的胜负.

5. 在 5.4 节的房室模型中,证明方程(3)对应的齐次方程通解如(4),(5)式所示,说明方程的两个特征根 α 和 β 一定是负实根.

6. 模仿 5.4 节建立的二室模型来建立一室模型(只有中心室),在快速静脉注射、恒速静脉滴注(持续时间为 τ) 和口服或肌肉注射 3 种给药方式下求解血药浓度,并画出血药浓度曲线的图形.

7. 利用上题建立的一室模型,讨论按固定时间间隔 T , 每次给予固定剂量 D 的多次重复给药方式. 为了维持药品的疗效和保证机体的安全,要求血药浓度 c 控制在 (c_1, c_2) 范围内. 设已知中心室容积为 V .

(1) 在快速静脉注射的多次重复给药方式下,写出血药浓度表达式并作图,讨论怎样确定 T 和 D , 使血药浓度的变化满足上述要求;实际上为了简化起见,常采取加大首次剂量给药的方式,给出这种情况下的给药方案.

(2) 在恒速静脉滴注和口服(或肌肉注射)的多次重复给药方式下,给出血药浓度变化

的简图,并选择一种方式讨论确定 T 和 D 的问题.

8. 在 5.5 节香烟过滤嘴模型中,

(1) 设 $M = 800 \text{ mg}$, $l_1 = 80 \text{ mm}$, $l_2 = 20 \text{ mm}$, $b = 0.02$, $\beta = 0.08$, $v = 50 \text{ mm/s}$, $\alpha = 0.3$, 求 Q 和 Q_1/Q_2 .

(2) 若有一支不带过滤嘴的香烟,参数同上. 比较全部吸完和只吸到 l_1 处的情况下,进入人体毒物量的区别.

9. 在 5.6 节人口的预测和控制模型中,总和生育率 $\beta(t)$ 和生育模式 $h(r, t)$ 是两种控制人口增长的手段. 试说明我国目前的人口政策,如提倡一对夫妇只生一个孩子、晚婚晚育,以及生育第 2 胎的一些规定,可以怎样通过这两种手段加以实施.

10. 将 5.8 节得到的万有引力定律(18)式与熟知的形式(19)进行比较. 查询太阳质量、地球运行轨道(椭圆)的长半轴、引力常数等数据,说明二者是一致的.

* 11. 对于技术革新的推广,在下列几种情况下分别建立模型:

(1) 推广工作通过已经采用新技术的人进行,推广速度与已采用新技术的人数成正比,推广是有限的.

(2) 总人数有限,因而推广速度还会随着尚未采用新技术人数的减少而降低.

(3) 在(2)的前提下考虑广告等媒介的传播作用^[16, 49].

* 12. 建立耐用消费品市场销售量的模型. 如果知道了过去若干时期销售量的情况,如何确定模型的参数?

* 13. 人工肾是帮助人体从血液中带走废物的装置,它通过一层薄膜与需要带走废物的血管相通,如图 1. 人工肾中通以某种液体,其流动方向与血液在血管中的流动方向相反,血液中的废物透过薄膜进入人工肾.

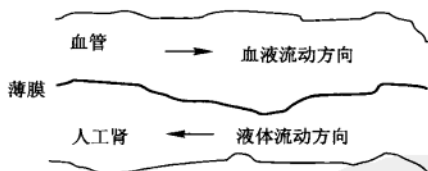


图 1

设血液和人工肾中液体的流速均为常数,废物进入人工肾的数量与它在这两种液体中的浓度差成正比. 人工肾总长 l . 建立单位时间内人工肾带走废物数量的模型^[16].

* 14. 在鱼塘中投放 n_0 尾鱼苗,随着时间的增长,尾数将减少而每尾的质量将增加.

(1) 设尾数 $n(t)$ 的(相对)减少率为常数;由于喂养引起的每尾鱼质量的增加率与鱼表面积成正比,由于消耗引起的每尾鱼质量的减少率与质量本身成正比. 分别建立尾数和每尾鱼质量的微分方程,并求解.

(2) 用控制网眼的办法不捕小鱼,到时刻 T 才开始捕捞,捕捞能力用尾数的相对减少量 $|\dot{n}/n|$ 表示,记作 E ,即单位时间捕获量是 $En(t)$. 问如何选择 T 和 E ,使从 T 开始的捕获量最大^[16].

* 15. 侦察机搜索潜艇. 设 $t = 0$ 时艇在 O 点,飞机在 A 点, $OA = 6 \text{ n mile}$ ($1 \text{ n mile} = 1.852 \text{ km}$).

此时艇潜入水中并沿着飞机不知道的某一方向以直线形式逃去,艇速 20 n mile/h. 飞机以速度 40 n mile/h 按照待定的航线搜索潜艇,当且仅当飞到艇的正上方时才可发现它.

(1) 以 O 为原点建立极坐标系 (r, θ) , A 点位于 $\theta=0$ 的向径上,见图 2. 分析图 2 中由 P, Q, R 组成的小三角形,证明在有限时间内飞机一定可以搜索到潜艇的航线,是先从 A 点沿直线飞到某点 P_0 ,再从 P_0 沿一条对数螺线飞行一周,而 P_0 是一个圆周上的任一点. 给出对数螺线的表达式,并画出一条航线的示意图.

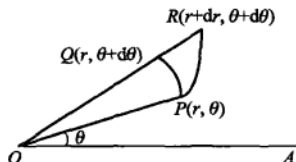


图 2

(2) 为了使整条航线是光滑的,直线段应与对数螺线在 P_0 点相切. 找出这条光滑的航线.

(3) 在所有一定可以发现潜艇的航线中哪一条航线最短,长度是多少? 光滑航线的长度又是多少^[50]?

* 16. 建立铅球掷远模型. 不考虑阻力,设铅球初速为 v ,出手高度为 h ,出手角度为 α (与地面夹角),建立投掷距离与 v, h, α 的关系式,并在 v, h 一定的条件下求最佳出手角度.

* 17. 建立肿瘤生长模型. 通过大量医疗实践发现,肿瘤细胞的生长有以下现象: (1) 当肿瘤细胞数目超过 10^{11} 时才是临床可观察的; (2) 在肿瘤生长初期,几乎每经过一定时间肿瘤细胞就增加一倍; (3) 由于各种生理条件限制,在肿瘤生长后期肿瘤细胞数目趋向某个稳定值.

(1) 比较 logistic 模型与 Gompertz 模型: $\frac{dn}{dt} = -\lambda n \ln \frac{n}{N}$, 其中 $n(t)$ 是细胞数, N 是极限值, λ 是参数.

(2) 说明上述两个模型是 Usher 模型: $\frac{dn}{dt} = \frac{\lambda n}{\alpha} \left(1 - \left(\frac{n}{N} \right)^\alpha \right)$ 的特例^[40].

18. 与 17 题中 Usher 模型类似的是 θ -logistic 模型: $\frac{dx}{dt} = rx \left(1 - \left(\frac{x}{N} \right)^\theta \right)$, 当 $\theta=1$ 时即为普通的 logistic 模型. 讨论 $\theta < 1$ 和 $\theta > 1$ 时模型的性质^[56].

* 19. 药物动力学中的 Michaelis-Menton 模型为 $\frac{dx}{dt} = -\frac{kx}{a+x}$ ($k, a > 0$), $x(t)$ 表示人体内药物在时刻 t 的血药浓度. 研究这个方程的解的性质^[56].

(1) 对于很多药物 (如可卡因), a 比 $x(t)$ 大得多, Michaelis-Menton 方程及其解如何简化?

(2) 对于另一些药物 (如酒精), $x(t)$ 比 a 大得多, Michaelis-Menton 方程及其解如何简化?

* 20. 建立一个模型说明要用三级火箭发射人造卫星的道理.

(1) 设卫星绕地球作匀速圆周运动,证明其速度为 $v = R \sqrt{g/r}$, R 为地球半径, r 为卫星与地心距离, g 为地球表面重力加速度. 要把卫星送上离地面 600 km 的轨道,火箭末速 v 应是多少?

(2) 设火箭飞行中速度为 $v(t)$, 质量为 $m(t)$, 初速为 0, 初始质量 m_0 , 火箭喷出的气体相对于火箭的速度为 u , 忽略重力和阻力对火箭的影响. 用动量守恒原理证明 $v(t) = u \ln \frac{m_0}{m(t)}$.

由此你认为要提高火箭的末速应采取什么措施.

(3) 火箭质量包括 3 部分:有效载荷(卫星) m_p ;燃料 m_f ;结构(外壳、燃料仓等) m_s ,其中 m_s 在 $m_f + m_s$ 中的比例记作 λ ,一般 λ 不小于 10%. 证明若 $m_p = 0$ (即火箭不带卫星),则燃料用完时火箭达到的最大速度为 $v_m = -u \ln \lambda$. 已知目前的 $u = 3 \text{ km/s}$,取 $\lambda = 10\%$,求 v_m . 这个结果说明什么?

(4) 假设火箭燃料燃烧的同时,不断丢弃无用的结构部分,即结构质量与燃料质量以 λ 和 $1 - \lambda$ 的比例同时减少,用动量守恒原理证明 $v(t) = (1 - \lambda) u \ln \frac{m_0}{m(t)}$. 问燃料用完时火箭末速为多少,与前面的结果有何不同?

(5) (4)是个理想化的模型,实际上只能用建造多级火箭的办法一段段地丢弃无用的结构部分. 记 m_i 为第 i 级火箭质量(燃料和结构), λm_i 为结构质量(λ 对各级是一样的). 有效载荷仍用 m_p 表示. 当第 1 级的燃料用完时丢弃第 1 级的结构,同时第 2 级点火. 再设燃烧级的初始质量与其负载质量之比保持不变,比例系数为 k . 证明 3 级火箭的末速 $v_3 = 3u \ln \frac{k+1}{\lambda k+1}$. 计算要使 $v_3 = 10.5 \text{ km/s}$,发射 1 t 重的卫星需要多重的火箭(u, λ 用以前的数据)? 若用 2 级或 4 级火箭,结果如何? 由此得出使用 3 级火箭发射卫星的道理^[94].



第6章

代数方程与差分方程模型

我们都学过用列代数方程的办法解简化的应用题,实际上,对于工程技术和社会经济领域中的许多问题,当不考虑时间因素的变化,作为静态问题处理时,常常可以建立代数方程模型,如 6.1, 6.2, 6.3 通过一个案例所介绍的量纲分析方法也是一个代数方程模型。

第 5 章建立了一些问题的微分方程模型,如果出于计算或应用上的考虑,将微分方程离散化就得到差分方程。而也有一些实际问题,直接建立差分方程模型更为方便,6.4 ~ 6.6 就是这种情况。

把代数方程模型与差分方程模型合在一章,除了篇幅上的原因外,它们都用类似的矩阵、向量的数学形式表达;以及求解过程中的相互联系,也是一个考虑因素。

6.1 投入产出模型

国民经济各个部门之间存在着相互依存和制约关系,每个部门在运转中将其他部门的产品或半成品(称为投入)经过加工变为自己的产品(称为产出),如何根据各部门间投入和产出的平衡关系,确定各部门的产出水平以满足社会的需求,是投入产出模型要研究的问题。

投入产出是 20 世纪 30 年代由美国经济学家 W. Leontief 首先提出和研究的,几十年来在理论和实践方面得到很大发展。投入产出模型已经从静态扩展到动态,并且随着与数量经济分析方法的日益融合,其应用领域不断扩大, W. Leontief 也因此获得了 1973 年诺贝尔经济学奖。

本节在介绍投入产出表和直接消耗系数表的基础上,建立静态的投入产出数学模型,并给出一个简单、具体的应用^[92]。

投入产出表 投入产出表描述了国民经济各部门之间的生产和消耗、投入和产出的数量关系。用一个例子说明它的结构。表 1 是一张简化的中国 2002 年投入产出表,表中国民经济由农业、工业、建筑业、运输邮电业、批零餐饮业和其他服务业 6 个部门构成,对每个部门有初始投入(工资、税收、进口等)和总投入,以及外部需求(消费、积累、出口等)和总产出。

表1 中国2002年投入产出表(产值单位:亿元)

产出 投入	农业	工业	建筑业	运输 邮电业	批零 餐饮业	其他 服务业	外部 需求	总产出
农业	464	788	229	13	127	13	1 284	2 918
工业	499	8 605	1 444	403	557	1 223	4 083	16 814
建筑业	5	9	3	20	23	124	2 691	2 875
运输邮电业	62	527	128	163	67	146	477	1 570
批零餐饮业	79	749	140	43	130	273	927	2 341
其他服务业	146	1 285	272	225	219	542	2 725	5 414
初始投入	1 663	4 851	659	703	1 218	3 093		
总投入	2 918	16 814	2 875	1 570	2 341	5 414		

表中数字均以产值计算,6个部门的横行表示该部门的产品供给各部门生产使用的数量,6个部门的纵列表示该部门生产中消耗的各部门产品的数量。比如对于第2行工业部门和第4列运输邮电部门交叉处的数字403亿元,既表示工业部门提供403亿元的产品给运输邮电部门生产使用,也表示运输邮电部门生产中消耗403亿元的工业产品。同行、同列交叉处的数字表示产品提供给本部门使用,亦即生产中消耗的本部门产品数量。最后一列为各部门的总产出,如第2行的16 814亿元是工业部门的总产出,它等于工业部门分配给6个部门生产使用的数量,加上4 083亿元的外部需求。最后一行为各部门的总投入,如第2列的16 814亿元是对工业部门的总投入,它等于6个部门对工业部门的投入,再加上4 851亿元的初始投入。可以看出,为了保持国民经济的平衡,每个部门的总投入和总产出是相等的。

直接消耗系数表 一般说来,一张投入产出表中的数字是随时间变化的。在分析各部门之间的投入产出关系时,为了利用一些相对稳定的指标来寻求其中的内在规律,引入直接消耗系数,其定义是一个部门的单位产出对各个部门的直接消耗量,它可以由投入产出表直接得到。如表1中运输邮电部门消耗403亿元工业部门的产品,总产出为1 570亿元,于是运输邮电部门的单位产出对工业部门的直接消耗是 $403/1\,570=0.257$,如此得到的直接消耗系数如表2。由于每个部门的总产出等于总投入,计算时将每行数字相应地除以最后一行数字即可。

表 2 中国 2002 年直接消耗系数表

产出 投入	农业	工业	建筑业	运输邮电业	批零餐饮业	其他服务业
农业	0.159	0.047	0.080	0.008	0.054	0.002
工业	0.171	0.512	0.502	0.257	0.238	0.226
建筑业	0.002	0.001	0.001	0.013	0.010	0.023
运输邮电业	0.021	0.031	0.045	0.104	0.029	0.027
批零餐饮业	0.027	0.045	0.049	0.027	0.056	0.050
其他服务业	0.050	0.076	0.095	0.143	0.094	0.100

以表 2 第 1 列为例说明这些系数的经济意义:农业部门每 1 亿元产出要直接消耗 0.159 亿元本部门产品、0.171 亿元工业产品、0.002 亿元建筑业产品、0.021 亿元运输邮电业服务、0.027 亿元批零餐饮业服务和 0.050 亿元其他服务。直接消耗系数反映了国民经济各个部门之间的投入产出关系。

投入产出数学模型 根据投入产出表及各部门间的直接消耗系数可以建立以下模型。

设有 n 个部门,记一定时期内第 i 部门的总产出为 x_i ,其中对第 j 部门的投入为 x_{ij} ,外部需求为 d_i ,则

$$x_i = \sum_{j=1}^n x_{ij} + d_i, i = 1, 2, \dots, n \quad (1)$$

表 1 的每一行都满足(1)式。这里 x_{ij} 也表示在第 j 部门的总产出中对第 i 部门的直接消耗。

记直接消耗系数为 a_{ij} ,根据定义,它是第 j 部门的单位产出对第 i 部门的直接消耗,于是

$$a_{ij} = x_{ij}/x_j, \quad i, j = 1, 2, \dots, n \quad (2)$$

注意到每个部门的总产出等于总投入,所以 x_j 也是第 j 部门的总投入。将(2)代入(1)得

$$x_i = \sum_{j=1}^n a_{ij}x_j + d_i, \quad i = 1, 2, \dots, n \quad (3)$$

记直接消耗系数矩阵 $A = (a_{ij})_{n \times n}$,产出向量 $x = (x_1, \dots, x_n)^T$,需求向量 $d = (d_1, \dots, d_n)^T$,则(3)式可写作

$$x = Ax + d \quad (4)$$

或

$$(I - A)x = d \quad (5)$$

其中 I 是单位矩阵。(5) 是一个代数方程模型, 下面给出它的一个应用。

在技术和管理水平没有明显提高的情况下, 可以假设直接消耗系数矩阵 A 在一个时期内是不变的。这样, 已知各部门的外部需求 d , 就可以从代数方程 (5) 求解出各部门的总产出 x , 并且讨论当外部需求改变时总产出应有什么变化。例如, 假定直接消耗系数矩阵 A 由表 2 给出, 如果某年对农业、工业、建筑业、运输邮电业、批零餐饮业和其他服务业的外部需求分别为 1 500, 4 200, 3 000, 500, 950, 3 000 亿元, 问这 6 个部门的总产出分别应为多少? 如果 6 个部门的外部需求分别增加 1 个单位, 问它们的总产出应分别增加多少?

对于第一个问题, 由表 2 数字构成矩阵 A , 令 $d = (1\ 500, 4\ 200, 3\ 000, 500, 950, 3\ 000)^T$, 由方程 (5) 解出 x , 可得到 6 个部门的总产出分别应为 3 277, 17 872, 3 210, 1 672, 2 478, 5 888 亿元。

对于第二个问题, 从方程组 (5) 可得

$$x = (I - A)^{-1}d \quad (6)$$

表明总产出 x 对外部需求 d 是线性的, 当 d 增加 1 个单位 (记作 Δd) 时, x 的增量为 $\Delta x = (I - A)^{-1} \Delta d$ 。若农业的外部需求增加 1 个单位, $\Delta d = (1, 0, 0, 0, 0, 0)^T$, Δx 为 $(I - A)^{-1}$ 的第 1 列; 工业和建筑业的外部需求增加 1 个单位, Δx 为 $(I - A)^{-1}$ 的第 2, 3 列; 以此类推。经过计算可知, 当农业的需求增加 1 个单位时, 农业、工业、建筑业、运输邮电业、批零餐饮业和其他服务业的总产出应分别增加 1.226 6, 0.562 4, 0.007 5, 0.054 9, 0.070 9, 0.132 5 单位, 其余类似。这些数字称部门联系系数, 在分析生产对于需求变化的灵敏性时是很有用的参数。

6.2 CT 技术的图像重建

CT (计算机断层成像, Computed Tomography 的缩写) 技术是 20 世纪 50 至 70 年代由美国科学家 A. M. Cormack 和英国科学家 G. N. Hounsfield 通过核物理、核医学等领域的一系列研究和实验发明的, 他们因此共同获得 1979 年诺贝尔医学奖。从 1971 年第一代供临床应用的 CT 设备问世以来, 随着电子技术的飞速发展, CT 技术不断改进, 诸如螺旋式 CT 机、电子束扫描机等新型设备逐渐被医疗机构普遍采用。此外, CT 技术在工业无损探测、资源勘探、生态监测等领域也得到了广泛的应用^[37, 89]。

什么是 CT, 它与传统的 X 射线成像 (如人们在医院拍的 X 光片) 有什么区别? 让我们看一个一般的概念图示, 如图 1。

假定有一个半透明物体, 其中嵌入 5 个不同透明度的球。如果按照图 1(a) 那样单方向地观察, 因为有 2 个球被前面的球遮挡, 我们可能错误地认为只有 3

个球,尽管重叠球的透明度较低,仍无法确定球的数目,更不要说各个球的透明度了。而如果按照图 1(b)那样让物体旋转起来,从多角度观察,就能够分辨出 5 个球以及它们各自的透明度。到医院作射线检查时,人体的内脏就像上面的半透明物体,传统的 X 射线成像原理就像图 1(a),X 射线和胶片相当于光源和人眼;CT 技术原理就像图 1(b),只不过旋转的不是人体,而是 X 光管和探测器。

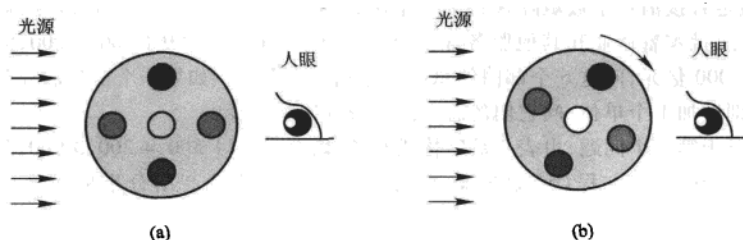


图 1 传统的 X 射线成像和 CT 的一般概念图示

概括地说,传统的 X 射线成像将人体器官和组织前后重叠地直接投影到胶片上,呈现出具有一定分辨率、但仍不够清晰的图像,CT 则在不同深度的断面上,从各个角度用探测器接收旋转的 X 光管发出、并由于穿过人体而使强度衰减的射线,再经过测量和计算,将人体器官和组织的影像重新构建出来,称为**图像重建**。本节将在简要介绍 X 射线强度衰减与图像重建的数学原理之后,给出图像重建的一个代数模型。

X 射线强度衰减与图像重建的数学原理 X 射线在穿过均匀材料的物质时,其强度的衰减率与强度本身成正比,即

$$\frac{dI}{dl} = -\mu I \quad (1)$$

其中 I 为射线强度, l 为物质在射线方向的厚度, μ 为物质对射线的衰减系数。由此可得

$$I = I_0 e^{-\mu l} \quad (2)$$

其中 I_0 为入射强度。当 X 射线的能量一定时,衰减系数 μ 随射线穿过的材料不同而改变,如骨骼的 μ 比软组织的大, X 射线的强度在骨骼中衰减得更快。(2) 式称 Beer-Lambert 定律。

当 X 射线穿过由不同衰减系数的材料组成的非均匀物体,如人体内部的某一断面时,(1)式中的 μ 为某平面坐标 x, y 的函数 $\mu(x, y)$,当射线沿 xy 平面内直线 L 穿行时,(2)式变为

$$I = I_0 e^{-\int_L \mu(x, y) dl} \quad (3)$$

其中 $\int_L \mu(x, y) dl$ 是 $\mu(x, y)$ 沿 L 的线积分, 如图 2. 由 (3) 可得

$$\int_L \mu(x, y) dl = \ln \frac{I_0}{I} \quad (4)$$

(4) 式右端的数值可从 CT 的 X 光管和探测器的测量数据得到.

如果根据 (4) 式得到了沿许多条直线 L 的线积分 $\int_L \mu(x, y) dl$, 能不能确定被积函数 $\mu(x, y)$ 呢? 如果能, 就可以根据人体内各个断面对 X 射线的衰减系数, 得到反映人体器官和组织的大小、形状、密度的图像, 即图像重建.

1917 年奥地利数学家 Radon 给出以下积分变换的逆变换的表达式, 为图像重建提供了理论基础.

定义函数 $f(x, y)$ 在平面上沿直线 L 的线积分为

$$P_f(L) = \int_L f(x, y) dl \quad (5)$$

对任一点 $Q(x, y)$, 作与 Q 相距为 $q (> 0)$ 的直线 L 的线积分 $P_f(L)$, 对所有的 q , 取 $P_f(L)$ 的平均值, 记作 $F_q(q)$, 则 Q 点的函数值 f 为

$$f(Q) = -\frac{1}{\pi} \int_0^\infty \frac{dF_q(q)}{q} \quad (6)$$

(5)、(6) 式分别称 $f(x, y)$ 的 Radon 变换和逆变换^①.

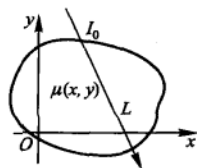


图 2 X 射线沿直线 L 穿过衰减系数为 $\mu(x, y)$ 的区域

从理论上, (5)、(6) 式给出了图像重建问题的完整解答, 而从应用的角度看这仅仅是一个开始. Radon 逆变换需要无穷多条直线 L 的线积分, 实际上只能在有限条直线上得到投影 (即线积分), 在逆变换的数学方法及其稳定性、重建图像的清晰度以及积分的离散化等方面有许多工作要做. 从反投影算法到滤波反投影算法和迭代算法, 从平行光束到扇形光束、锥形光束的重建算法, 图像重建问题在数学方法上的进展, 为 CT 技术在各个领域成功的和不断拓广的应用提供了必要的条件.

下面只从离散化的角度, 讨论图像重建的一个代数模型.

图像重建的代数模型 将待测的平面图形分成若干小正方形, 称为像素, 一定宽度的射线束从各个方向穿过这些像素, 不同像素对射线的衰减系数可以不同, 但合理地假定每一个像素对射线的衰减系数是常数. 这个模型要通过对多条穿过待测平面图形的射线强度的量测, 确定各个像素的衰减系数.

① 逆变换存在的条件是 $f(x, y)$ 连续且有紧支集, 积分 (5) 对任一直线 L 有定义.

设有 m 个像素 (记 $j=1, 2, \dots, m$), n 束射线 (记 $i=1, 2, \dots, n$), 第 i 束射线记作 L_i , 由 L_i 的强度测量数据计算出 (4) 式的右端, 记作 $\ln(I_0/I)_i$, 则 (4) 式的积分形式可以写作如下的和式:

$$\sum_{j \in J(L_i)} \mu_j \Delta l_j = \ln(I_0/I)_i, i = 1, 2, \dots, n \quad (7)$$

其中 μ_j 是像素 j 的衰减系数, Δl_j 是射线在像素 j 中的穿行长度 (具体计算见下), $J(L_i)$ 表示射线束 L_i 穿过的像素 j 的集合。

实际应用时, 对 (7) 的计算定义了几种常用的方法. 为简单起见, 记 $\ln(I_0/I)_i$ 为 b_i , μ_j 为 x_j . 设每个像素的边长和射线束的宽度均为 σ , 如图 3.

1. 中心线法 记射线束 L_i 的中心线在像素 j 内的长度 l_{ij} 与像素边长之比为 a_{ij} , (7) 式化为 (L_i 的中心线不经过像素 j 时 $a_{ij}=0$, 并不计常数因子)

$$\sum_{j=1}^m a_{ij} x_j = b_i, i = 1, 2, \dots, n \quad (8)$$

或写成向量-矩阵形式

$$Ax = b \quad (9)$$

其中 $A = (a_{ij})_{n \times m}$, $x = (x_1, x_2, \dots, x_m)^T$, $b = (b_1, b_2, \dots, b_n)^T$.

2. 面积法 记射线束 L_i 在像素 j 内的面积 s_{ij} 与像素面积之比为 a_{ij} , (7) 式同样化为 (9) 式 (L_i 不经过像素 j 时 $a_{ij}=0$, 并不计常数因子).

3. 中心法 当射线束 L_i 经过像素 j 的中心点时记 $a_{ij}=1$, 否则 $a_{ij}=0$, (7) 式同样化为 (9) 式.

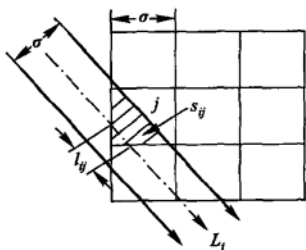


图 3 射线束 L_i 经过像素 j 示意图
(像素边长和射线束宽度均为 σ)

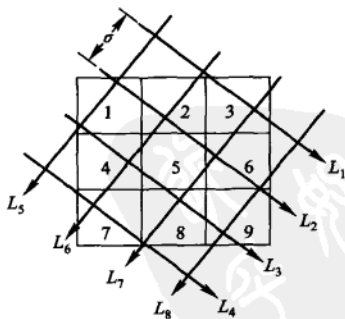


图 4 射线 $L_1 \sim L_8$ 经过像素 1~9 示意图
(像素边长和射线间距均为 σ)

作为中心法的另一种简化形式, 假定射线束的宽度为 0, 间距 σ , 当 L_i 经过像素 j 内任一点时记 $a_{ij}=1$, 否则 $a_{ij}=0$. 图 4 作为这种形式的一个具体例子 ($m=9, n=8$), 可以写出矩阵

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

不论采用上面 3 种方法的哪一个,都可以利用代数方程 $Ax = b$ 来重建图像,即根据已知的矩阵 A 和测量数据计算出的向量 b ,确定像素的衰减系数向量 x 。实际上,像素数量 m 和射线数量 n 都很大,且一般 $m > n$,使方程有无穷多解。另外,测量误差和噪声影响不可忽视,于是可将(9)式修正为

$$Ax + e = b \quad (10)$$

其中 e 是误差向量。于是问题化为在 x 和 e 满足指定的最优准则下,根据 A 和 b 估计 x ,这一过程常用迭代算法进行,称为代数重建技术(ART),详见[97]。

6.3 原子弹爆炸的能量估计与量纲分析

1945 年 7 月 16 日,美国科学家在新墨西哥州的沙漠中试爆了全球第一颗原子弹(图 1)。这一事件令世界为之震惊,并从某种程度上改变了第二次世界大战以及战后世界的历史。但在当时,有关原子弹爆炸的资料都是保密的,一般人无法得到任何有关的数据或影像资料,因此无法比较准确地了解这次爆炸的威力究竟有多大。两年以后,美国政府首次公开了这次爆炸的录像带,但没有发布任何其他有关的资料。英国物理学家 Taylor(1886—1975)通过研究这次爆炸的录像带,建立数学模型对这次爆炸所释放的能量进行了估计,得到的估计值为 $19.2 \times 10^3 \text{ t}$ (10^3 t 相当于 $1\,000 \text{ t}$ TNT 的核子能量)。后来正式公布的信息显示,这次爆炸实际释放的能量为 $21 \times 10^3 \text{ t}$,与 Taylor 的估计相当接近。

除了公开的录像带,Taylor 不掌握这次原子弹爆炸的其他任何信息,他如何估计爆炸释放的能量呢?物理常识告诉我们,爆炸产生的冲击波以爆炸点为中心呈球面向四周传播,爆炸的能量越大,在一定时刻冲击波传播得越远,而冲击波又通过爆炸形成的“蘑菇云”反映出来。Taylor 研究这次爆炸的录像带,测量出了从爆炸开始,不同时刻爆炸所产生的“蘑菇云”的半径。表 1 是他测量出的时刻 t 所对应的“蘑菇云”的半径 r 。

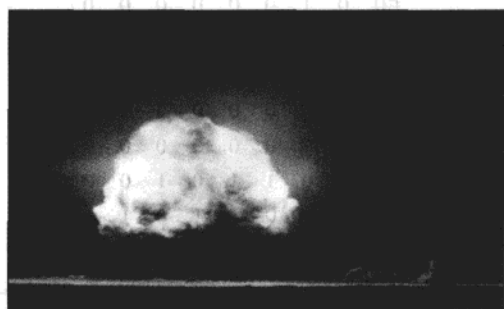


图1 原子弹爆炸示意图

表1 时刻 t 所对应的“蘑菇云”的半径 r

t/ms	r/m	t/ms	r/m	t/ms	r/m	t/ms	r/m	t/ms	r/m
0.10	11.1	0.80	34.2	1.50	44.4	3.53	61.1	15.0	106.5
0.24	19.9	0.94	36.3	1.65	46.0	3.80	62.9	25.0	130.0
0.38	25.4	1.08	38.9	1.79	46.9	4.07	64.3	34.0	145.0
0.52	28.8	1.22	41.0	1.93	48.7	4.34	65.6	53.0	175.0
0.66	31.9	1.36	42.8	3.26	59.0	4.61	67.3	62.0	185.0

Taylor 是首先用量纲分析方法建立数学模型,然后辅以小型试验,又利用表1的数据,对原子弹爆炸的能量进行估计的^[84]. 我们先对量纲分析方法作简单介绍.

量纲齐次原则 量纲分析(Dimensional Analysis)是20世纪初提出的在物理和工程等领域建立数学模型的一种方法,它在经验和实验的基础上利用物理定律的量纲齐次原则,确定各物理量之间的关系.

许多物理量是有量纲的,在物理研究中把若干物理量的量纲作为基本量纲.它们是相互独立的,另一些物理量的量纲则可根据其定义或物理定律由基本量纲推导出来,称为导出量纲.例如,在研究力学问题时,通常将长度 l 、质量 m 和时间 t 的量纲作为基本量纲,记以相应的大写字母 L 、 M 和 T . 在量纲分析中,物理量 q 的量纲记作 $[q]$,于是有 $[l] = L$, $[m] = M$, $[t] = T$. 而速度 v 、加速度 a 的量纲可以按照其定义为 $[v] = LT^{-1}$, $[a] = LT^{-2}$, 力 f 的量纲则应根据牛顿第二定律用质量和加速度的乘积表示,即 $[f] = LMT^{-2}$, 这些就是导出量纲.

有些物理常数也有量纲,如在万有引力定律 $f = k \frac{m_1 m_2}{r^2}$ 中引力常数

$k = \frac{fr^2}{m_1 m_2}$, k 的量纲可从力 f 、长度 r 和质量 m 的量纲得到: $[k] = \text{LMT}^{-2} \cdot \text{L}^2 \cdot \text{M}^{-2} = \text{L}^3 \text{M}^{-1} \text{T}^{-2}$. 对于无量纲的量 λ , 记 $[\lambda] = \text{L}^0 \text{M}^0 \text{T}^0 = 1$.

用数学公式表示一些物理量之间的关系时, 公式等号两端必须有相同的量纲, 称为量纲齐次性 (Dimensional Homogeneity). 量纲分析就是利用量纲齐次原则来建立物理量之间的数学模型^[10, 32]. 先看一个简单的例子.

例 单摆运动 这是一个大家熟知的物理现象. 质量 m 的小球系在长度 l 的线的一端, 稍偏离平衡位置后小球在重力 mg 作用下 (g 为重力加速度) 作往复摆动, 忽略阻力, 求摆动周期 t 的表达式.

在这个问题中出现的物理量有 t, m, l, g , 设它们之间的关系是

$$t = \lambda m^{\alpha_1} l^{\alpha_2} g^{\alpha_3} \quad (1)$$

其中 $\alpha_1, \alpha_2, \alpha_3$ 是待定常数, λ 是无量纲的比例系数. (1) 式的量纲表达式为

$$[t] = [m]^{\alpha_1} [l]^{\alpha_2} [g]^{\alpha_3} \quad (2)$$

将 $[t] = \text{T}$, $[m] = \text{M}$, $[l] = \text{L}$, $[g] = \text{LT}^{-2}$ 代入得

$$\text{T} = \text{M}^{\alpha_1} \text{L}^{\alpha_2 + \alpha_3} \text{T}^{-2\alpha_3} \quad (3)$$

按照量纲齐次原则应有

$$\begin{cases} \alpha_1 & = 0 \\ \alpha_2 + \alpha_3 & = 0 \\ -2\alpha_3 & = 1 \end{cases} \quad (4)$$

(4) 的解为 $\alpha_1 = 0, \alpha_2 = 1/2, \alpha_3 = -1/2$, 代入 (1) 式得

$$t = \lambda \sqrt{\frac{l}{g}} \quad (5)$$

我们看到, 用非常简单的方法得到的 (5) 式与用较深入的力学知识推出的结果是一样的^①.

为了导出用量纲分析建模的一般方法, 将这个例子中各个物理量之间的关系写作

$$f(t, m, l, g) = 0 \quad (6)$$

这里没有因变量与自变量之分. 进而假设 (6) 式形如

$$t^{\gamma_1} m^{\gamma_2} l^{\gamma_3} g^{\gamma_4} = \pi \quad (7)$$

其中 $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ 是待定常数, π 是无量纲常数. 将 t, l, m, g 的量纲用基本量纲 $\text{L}, \text{M}, \text{T}$ 表示为 $[t] = \text{L}^0 \text{M}^0 \text{T}^1$, $[m] = \text{L}^0 \text{M}^1 \text{T}^0$, $[l] = \text{L}^1 \text{M}^0 \text{T}^0$, $[g] = \text{L}^1 \text{M}^0 \text{T}^{-2}$, 则 (7) 式的量纲表达式可写作

① 这个问题的完整结果是 $t = 2\pi \sqrt{l/g}$, 用量纲分析得不到 $\lambda = 2\pi$, 因为它是无量纲的.

$$(L^0 M^0 T^1)^{y_1} (L^0 M^1 T^0)^{y_2} (L^1 M^0 T^0)^{y_3} (L^1 M^0 T^{-2})^{y_4} = L^0 M^0 T^0 \quad (8)$$

即

$$L^{y_3+y_4} M^{y_2} T^{y_1-2y_4} = L^0 M^0 T^0 \quad (9)$$

由量纲齐次原则有

$$\begin{cases} y_3 + y_4 = 0 \\ y_2 = 0 \\ y_1 - 2y_4 = 0 \end{cases} \quad (10)$$

方程组(10)有一个基本解

$$\mathbf{y} = (y_1, y_2, y_3, y_4)^T = (2, 0, -1, 1)^T \quad (11)$$

代入(7)式得

$$t^2 l^{-1} g = \pi \quad (12)$$

(6)式可以等价地表示为

$$F(\pi) = 0 \quad (13)$$

(12), (13)两式是量纲分析方法从(6)式导出的一般结果,前面的(5)式只是它的特殊表达形式.

把(6)到(13)式的推导过程一般化,就是著名的白金汉 π 定理(或记为 Pi 定理)^①.

π 定理 设 m 个有量纲的物理量 q_1, q_2, \dots, q_m 之间存在与量纲单位的选取无关的物理规律,数学上可表示为

$$f(q_1, q_2, \dots, q_m) = 0 \quad (14)$$

若基本量纲记作 $X_1, X_2, \dots, X_n (n \leq m)$, 而 q_1, q_2, \dots, q_m 的量纲可表为

$$[q_j] = \prod_{i=1}^n X_i^{a_{ij}}, j = 1, 2, \dots, m \quad (15)$$

矩阵 $A = (a_{ij})_{n \times m}$ 称量纲矩阵. 若 A 的秩

$$\text{Rank } A = r \quad (16)$$

设线性齐次方程组

$$A\mathbf{y} = \mathbf{0}, \mathbf{y} = (y_1, y_2, \dots, y_m)^T \quad (17)$$

的 $m-r$ 个基本解记作

$$\mathbf{y}^{(s)} = (y_1^{(s)}, y_2^{(s)}, \dots, y_m^{(s)})^T, s = 1, 2, \dots, m-r \quad (18)$$

则存在 $m-r$ 个相互独立的无量纲量

$$\pi_s = \prod_{j=1}^m q_j^{y_j^{(s)}}, s = 1, 2, \dots, m-r \quad (19)$$

且

① 该定理由美国科学家 E. Buckingham (1867—1940) 于 1914 年给出, 本书第二版有简要证明.

$$F(\pi_1, \pi_2, \dots, \pi_{m-r}) = 0 \quad (20)$$

(19), (20) 与 (14) 式等价, F 是一个未定的函数关系.

原子弹爆炸能量估计的量纲分析方法建模

记原子弹爆炸能量为 E , 将“蘑菇云”的形状近似看成一个球形, 记时刻 t 球的半径为 r , 与 r 有关的物理量还可能有“蘑菇云”周围的空气密度 (记为 ρ) 和大气压强 (记为 P), 于是 r 作为 t 的函数还与 E, ρ, P 有关, 要寻求的关系是

$$r = \varphi(t, E, \rho, P) \quad (21)$$

更一般的形式记作

$$f(r, t, E, \rho, P) = 0 \quad (22)$$

其中有 5 个物理量, (22) 相当于 π 定理的 (14) 式. 下面利用 π 定理解决这个问题.

取长度 L , 质量 M 和时间 T 为基本量纲, (22) 中各个物理量的量纲分别是

$$[r] = L, [t] = T, [E] = L^2 M T^{-2}, [\rho] = L^{-3} M, [P] = L^{-1} M T^{-2} \quad (23)$$

由此得到量纲矩阵

$$A_{3 \times 5} = \begin{bmatrix} 1 & 0 & 2 & -3 & -1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & -2 & 0 & -2 \end{bmatrix} \quad (24)$$

因为 A 的秩是 3, 所以齐次方程

$$A\mathbf{y} = \mathbf{0}, \mathbf{y} = (y_1, y_2, y_3, y_4, y_5)^T \quad (25)$$

有 $5 - 3 = 2$ 个基本解.

令 $y_1 = 1, y_3 = 0$, 得到一个基本解 $\mathbf{y} = (1, -2/5, -1/5, 1/5, 0)^T$; 令 $y_1 = 0, y_3 = 1$, 得到另一个基本解 $\mathbf{y} = (0, 6/5, -2/5, -3/5, 1)^T$. 由这 2 个基本解可以得到 2 个无量纲量

$$\pi_1 = r t^{-2/5} E^{-1/5} \rho^{1/5} = r \left(\frac{\rho}{t^2 E} \right)^{1/5} \quad (26)$$

$$\pi_2 = t^{6/5} E^{-2/5} \rho^{-3/5} P = \left(\frac{t^6 P^5}{E^2 \rho^3} \right)^{1/5} \quad (27)$$

且存在某个函数 F , 使得

$$F(\pi_1, \pi_2) = 0 \quad (28)$$

与 (22) 式等价.

为了得到形如 (21) 式的关系, 取 (28) 式的特殊形式 $\pi_1 = \psi(\pi_2)$ (其中 ψ 是某个函数), 由 (26), (27) 式即

$$r \left(\frac{\rho}{t^2 E} \right)^{1/5} = \psi \left(\frac{t^6 P^5}{E^2 \rho^3} \right)^{1/5} \quad (29)$$

于是

$$r = \left(\frac{t^2 E}{\rho} \right)^{1/5} \psi \left(\frac{t^6 P^5}{E^2 \rho^3} \right)^{1/5} \quad (30)$$

函数 ψ 的具体形式需要采用其他方式确定。(30)式就是用量纲分析方法建立的、估计原子弹爆炸能量的数学模型。

原子弹爆炸能量估计的数值计算

为了利用表 1 中 t 和 r 的数据,由(30)式确定原子弹爆炸的能量 E ,必须先估计无量纲量 $\psi(\pi_2)$ 的大小。

Taylor 认为,对于原子弹爆炸来说,所经历的时间非常短,而所释放的能量非常大。仔细分析(27)式可知, $\pi_2 = \left(\frac{t^6 P^5}{E^2 \rho^3} \right)^{1/5} \approx 0$ 。于是 $\psi(0)$ 可看作一个比例系数 λ ,将(30)式记作

$$r = \lambda \left(\frac{t^2 E}{\rho} \right)^{1/5} \quad (31)$$

为了确定 λ 的大小,Taylor 借助一些小型爆炸试验的数据,最终决定取 $\lambda \approx 1$,这样就得到能量 E 的近似估计

$$E = \frac{\rho r^5}{t^2} \quad (32)$$

利用表 1 中时刻 t 所对应的“蘑菇云”的半径 r 作拟合来估计能量 E ,相当于取(32)式右端的平均值。取空气密度为 $\rho = 1.25 \text{ kg/m}^3$,可得到 $E = 8.2825 \times 10^{13} \text{ J}$ 。查表可知 $10^3 \text{ t} = 4.184 \times 10^{12} \text{ J}$,所以爆炸能量是 $19.7957 \times 10^3 \text{ t}$,与实际值 $21 \times 10^3 \text{ t}$ 相差不大(Taylor 是直接由(31)式作拟合,得到爆炸能量为 $19.1863 \times 10^3 \text{ t}$)。

(31)或(32)式还表明,当 E, ρ 一定时, r 与 $t^{2/5}$ 成正比,我们可以用表 1 的数据检验一下这个关系。设

$$r = at^b \quad (33)$$

其中 a, b 是待定系数,对(33)式取对数后可以用线性最小二乘法拟合,根据表 1 中 t 和 r 的资料确定。经过计算得到 $b = 0.4058$,与量纲分析得到的结果 ($b = 2/5$) 非常接近。(33)式与实际数据拟合的情况如图 2。

原子弹爆炸的能量估计被看做是量纲分析方法建模的一个成功范例。

量纲分析方法的评注 量纲齐次原则和 π 定理是具有普遍意义、又相当初等的方法,它不需要掌握研究对象那个领域中专门的物理知识,也用不到高深的数学工具,得到的结果有的与用专门方法推出的结果相同,如单摆运动,有的是用其他方法难以得到的,如原子弹爆炸的能量。一般地说,从 m 个原始物理量具有的规律 $f(q_1, q_2, \dots, q_m) = 0$,到它的等价形式 $F(\pi_1, \pi_2, \dots, \pi_{m-r}) = 0$,不仅变量个数减少了 r 个,而且组成了一些非常有用的无量纲量。在 π 定理的应用过程中有几点值得注意:

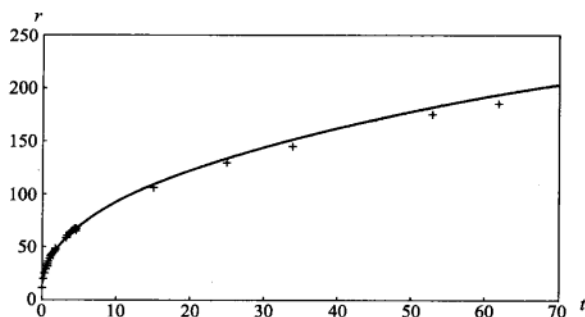


图2 (33)式(曲线)与实际数据(+)的拟合

1. 正确确定各个物理量 面对一个实际问题,将哪些原始物理量 q_1, q_2, \dots, q_m 包含在量纲分析之内,对所得结果的合理性是至关重要的.在原子弹爆炸的能量问题中,如果忽略空气密度 ρ ,就得不到上面的结果.原始物理量的确定主要依靠经验和物理知识,无法绝对保证所得结果的正确或有用.

2. 合理选取基本量纲 基本量纲选少了,无法给出各物理量的量纲表达,当然不行;选多了也会使问题复杂化.一般情况下,力学问题选 L, M, T 即可.当然,这不是惟一的,比如可以将 T 换成速度量纲 V .热学问题可考虑温度量纲,电学问题可考虑电量纲.

3. 恰当构造基本解 线性齐次方程组的基本解可以有许许多多的构造方法,从而有不同的无量纲量,从数学上看它们是等价的,如原子弹爆炸的能量问题中可以得到无量纲量 $\pi_3 = r^3 E^{-1} P (= \pi_1^2 \pi_2^{-1})$, $\pi_4 = r^2 t^{-2} \rho P^{-1} (= \pi_1^3 \pi_2)$ 等,但是为了特定的建模目的,恰当地选取无量纲量,能够更直接地得到我们期望的结果.

从 π 定理的表述和原子弹爆炸能量问题得到的结果可以看到,用量纲分析方法建立的数学模型是有局限的、“不彻底的”,(28)~(30)式中的函数, F, ψ 都无法用量纲分析确定.在一些物理公式中常会出现的三角函数、指数函数等都不可能用量纲分析方法得到,因为这些函数的自变量和因变量都是无量纲的.

量纲分析在物理模拟中的应用 当直接研究实际生活中的原型遇到困难时,一种解决办法是在实验室条件下,按照一定的比例尺构造它的物理模型,通过对模型的研究得出原型的结果,称为物理模拟.量纲分析可以指导物理模拟中比例尺的确定,下面通过一个例子来说明.

为研究大海波浪对航船的阻力,建造航船模型进行物理模拟.根据经验和物理知识,相关的物理量除阻力 f 外,有航船速度 v 、船体尺寸 l 、浸没面积 s 、海水密度 ρ 和重力加速度 g ,将它们的关系记作

$$\varphi(f, v, l, s, \rho, g) = 0 \quad (34)$$

利用 π 定理可以得到(请读者自行推导)

$$\pi_1 = fl^{-3}\rho^{-1}g^{-1}, \pi_2 = vl^{-1/2}g^{-1/2}, \pi_3 = l^{-2}s, F(\pi_1, \pi_2, \pi_3) = 0 \quad (35)$$

由(35)式写出阻力 f 的显式表达式

$$f = l^3 \rho g \psi(\pi_2, \pi_3) \quad (36)$$

将上述物理量和(35), (36)式对应于原型船, 用 $f', v', l', s', \rho', g'$ 及

$$\pi'_1 = f'l'^{-3}\rho'^{-1}g'^{-1}, \pi'_2 = v'l'^{-1/2}g'^{-1/2}, \pi'_3 = l'^{-2}s', F(\pi'_1, \pi'_2, \pi'_3) = 0 \quad (37)$$

$$f' = l'^3 \rho' g' \psi(\pi'_2, \pi'_3) \quad (38)$$

对应于模型船, 注意(36)和(38)中的函数 ψ 是一样的. 于是当无量纲量

$\pi_2 = \pi'_2, \pi_3 = \pi'_3$, 即

$$\frac{v}{\sqrt{l}g} = \frac{v'}{\sqrt{l'}g'}, \quad \frac{s}{l^2} = \frac{s'}{l'^2} \quad (39)$$

由(36)和(38)可得

$$\frac{f}{f'} = \frac{l^3 \rho g}{l'^3 \rho' g'} \quad (40)$$

海水密度 $\rho = \rho'$ 和重力加速度 $g = g'$ 的条件容易满足, 所以由(39)式, 在进行物理模拟时只需保证

$$\frac{v}{v'} = \sqrt{\frac{l}{l'}}, \quad \frac{s}{s'} = \left(\frac{l}{l'}\right)^2 \quad (41)$$

就有

$$\frac{f}{f'} = \left(\frac{l}{l'}\right)^3 \quad (42)$$

这样, 确定了原型和模型船体的比例 $l:l'$, 只要模拟时使得(41)式成立, 在测得模型船所受阻力 f' 后, 就可由(42)式计算原型船受的阻力 f 了.

无量纲化 我们不拟对无量纲化方法作一般阐述, 而是通过一个例子介绍这种方法如何用来对模型进行简化^[47, 48].

抛射问题 在星球表面以初速 v 竖直向上发射火箭^①, 记星球半径为 r , 星球表面重力加速度为 g , 忽略阻力, 讨论发射高度 x 随时间 t 的变化规律.

设 x 轴竖直向上, 在发射时刻 $t=0$ 火箭高度 $x=0$ (星球表面). 火箭和星球的质量分别记作 m_1 和 m_2 , 则由牛顿第二定律和万有引力定律可得

$$m_1 \ddot{x} = -k \frac{m_1 m_2}{(x+r)^2} \quad (43)$$

以 $x=0$ 时 $\ddot{x} = -g$ 代入(43)式, 并注意到初始条件, 抛射问题满足如下方程

① 为使结果更一般化, 这里取星球表面而不是地球表面.

$$\begin{cases} \ddot{x} = -\frac{r^2 g}{(x+r)^2} \\ x(0) = 0 \\ \dot{x}(0) = v \end{cases} \quad (44)$$

(44)的解可以表示为

$$x = x(t; r, v, g) \quad (45)$$

即发射高度 x 是以 r, v, g 为参数的时间 t 的函数. 这里的目的不是研究这个函数的具体形式 (虽然可以通过求解方程 (44) 直接得到), 而是讨论用无量纲化方法简化它的途径.

(45) 式包含 3 个独立参数 r, v, g , 由 (45) 式得到的进一步的结果, 如火箭到达最高点的时间 $t_M = t|_{\dot{x}=0}$ 必定是这 3 个参数的函数 $t_M = \varphi(r, v, g)$. 如果方程 (44) 变得稍微复杂以致必须用数值法求解时, 对不同的参数 r, v 和 g, t_M 的数值就要用 3 维表格给出. 用无量纲化的方法可以减少独立参数的个数, 达到简化模型的目的.

以长度和时间量纲 L, T 为基本量纲, 问题中的变量 x, t 和参数 r, v, g 的量纲表达式为

$$[x] = L, [t] = T, [r] = L, [v] = LT^{-1}, [g] = LT^{-2}$$

所谓无量纲化是指: 对于变量 x 和 t 分别构造具有相同量纲的参数组合 x_c 和 t_c , 使新变量

$$\bar{x} = \frac{x}{x_c}, \bar{t} = \frac{t}{t_c}$$

为无量纲量. x_c 称特征长度, t_c 称特征时间, 统称特征尺度 (Characteristic Scale) 或参考尺度. 利用新变量 \bar{x} 和 \bar{t} , 表达式 (45) 可以简化.

特征尺度 x_c 和 t_c 由参数 r, v, g 构成, 并应与 x 和 t 有相同的量纲, 即 $[x_c] = L, [t_c] = T$. 这样的 x_c 和 t_c 有多种构造方法, 下面举出几种.

1. 令 $x_c = r, t_c = rv^{-1}$, 则 $\bar{x} = x/r, \bar{t} = t/rv^{-1}$, 利用求导数规则可以算出

$$\dot{x} = v \frac{d\bar{x}}{d\bar{t}}, \ddot{x} = \frac{v^2}{r} \frac{d^2\bar{x}}{d\bar{t}^2}$$

\bar{x} 对 \bar{t} 的导数以下简记作 $\dot{\bar{x}}$ 和 $\ddot{\bar{x}}$, 方程 (39) 在新变量 (无量纲量) \bar{x}, \bar{t} 下的表达式为

$$\begin{cases} \ddot{\bar{x}} = -\frac{1}{(\bar{x}+1)^2}, \varepsilon = \frac{v^2}{rg} \\ \bar{x}(0) = 0 \\ \dot{\bar{x}}(0) = 1 \end{cases} \quad (46)$$

(46)的解可表示为

$$\bar{x} = \bar{x}(\bar{t}; \varepsilon) \quad (47)$$

它只含一个独立参数 ε , 不难验证 ε 是无量纲量. 原方程的解 $x = x(t; r, v, g)$ 中的 3 个参数以无量纲组合形式 ε 出现在表达式 (47) 中, 简化了原来的结果.

2. 令 $x_e = r, t_e = \sqrt{rg}^{-1}$, 类似地计算可将方程 (44) 化为

$$\begin{cases} \ddot{\bar{x}} = -\frac{1}{(\bar{x}+1)^2} \\ \bar{x}(0) = 0 \\ \dot{\bar{x}}(0) = \sqrt{\varepsilon}, \varepsilon = \frac{v^2}{rg} \end{cases} \quad (48)$$

其解的表达式仍为 (47) 式 (当然 \bar{x}, \bar{t} 已经变了).

3. 令 $x_e = v^2 g^{-1}, t_e = v g^{-1}$, 方程 (44) 化为

$$\begin{cases} \ddot{\bar{x}} = -\frac{1}{(\varepsilon \bar{x} + 1)^2}, \varepsilon = \frac{v^2}{rg} \\ \bar{x}(0) = 0 \\ \dot{\bar{x}}(0) = 1 \end{cases} \quad (49)$$

其解的表达式也是 (47) 式.

还可以构造其他形式的特征尺度 x_e 和 t_e , 得到其他形式的方程. 以上 3 种构造特征尺度的方法虽然在把方程 (44) 的解 (45) 式简化为 (47) 式这一点上是共同的, 但是进一步分析发现, 它们之间仍有重要差别.

我们知道, 按照今天的技术, 在地球表面发射火箭时, 初速 v 将满足

$$v \ll \sqrt{rg} = \sqrt{6\,370 \times 10^3 \times 9.8} = 8\,000 \text{ m/s}$$

所以必然有 $\varepsilon = v^2/rg \ll 1$. 既然 ε 如此之小, 能不能在方程 (46), (48), (49) 中舍弃以 ε 为因子的项, 从而得到方程的近似解呢? 让我们看看这样做有什么问题.

如果在方程 (46) 中令 $\varepsilon = 0$, 则 (46) 变为

$$\frac{1}{(\bar{x}+1)^2} = 0, \quad \bar{x}(0) = 0, \quad \dot{\bar{x}}(0) = 1 \quad (50)$$

(50) 式显然无解. 所以不能在方程 (46) 中舍弃含 ε 的项.

如果在方程 (48) 中令 $\varepsilon = 0$, 则 (48) 变为

$$\ddot{\bar{x}} = -\frac{1}{(\bar{x}+1)^2}, \quad \bar{x}(0) = 0, \quad \dot{\bar{x}}(0) = 0 \quad (51)$$

(51) 的解显然满足 $\bar{x}(t) < 0 (t > 0)$, 而原方程 (44) 的解 $x(t) \geq 0$, 所以不能从方程 (48) 得到原方程的近似解, 即不能在 (48) 中舍弃含 ε 的项.

如果在方程 (49) 中令 $\varepsilon = 0$, 则 (49) 变为

$$\ddot{\bar{x}} = -1, \quad \bar{x}(0) = 0, \quad \dot{\bar{x}}(0) = 1 \quad (52)$$

(52)的解显然为

$$\bar{x}(\bar{t}) = -\frac{\bar{t}^2}{2} + \bar{t} \quad (53)$$

代回原变量 x 和 t , (53) 式等价于

$$x(t) = -\frac{1}{2}gt^2 + vt \quad (54)$$

不难看出,如果在原抛射问题中假定:火箭发射过程中所受星球引力 m_1g 不变,那么微分方程为

$$\ddot{x} = -g, x(0) = 0, \dot{x}(0) = v \quad (55)$$

(54) 式正是方程 (55) 的解。而将 (55) 与原方程 (44) 对比,因为发射高度 $x \ll r$, 所以 (55) 是 (44) 的近似方程。这就说明可以在方程 (49) 中舍弃含 ε 的项,得到近似解。

第 3 种构造特征尺度 x_c 和 t_c 的方法之所以能够忽略含 ε 的项,成功地得到原问题的近似解的原因,在于 x_c 和 t_c 选得合适。从物理学容易知道,当初速 v 较小时(相对于 \sqrt{rg}),火箭在定常引力 m_1g 作用下到达最高点的时间为 v/g , 达到的最高距离为 $v^2/2g$, 所以选择 $t_c = vg^{-1}$ 和 $x_c = v^2g^{-1}$, 与 t 和 x 的大小相当。这样,无量纲量 $\bar{t} = t/t_c$ 和 $\bar{x} = x/x_c$ 大体上具有单位尺度(读者可以计算 $\dot{\bar{x}}, \ddot{\bar{x}}$, 可知它们大体上也具有单位尺度)。于是在 $\bar{t}, \bar{x}, \dot{\bar{x}}, \ddot{\bar{x}}$ 组成的新方程中,含有因子 $\varepsilon (\ll 1)$ 的项相对于不含 ε 的项而言,就可以舍弃了。

关于选择特征尺度的一般讨论可参看[47]。

评注 无量纲化是用数学工具研究物理问题时常用的方法。恰当地选择特征尺度不仅可以减少独立参数的个数,而且可以帮助人们决定舍弃哪些次要因素。像量纲分析中运用 π 定理时有正确确定各物理量和合理选取基本量纲的问题一样,无量纲化的关键是选择合适的特征尺度,而这也主要依赖于物理知识和经验。

6.4 市场经济中的蛛网模型

在自由贸易市场上你注意过这样的现象吗? 一个时期以来某种消费品如猪肉的上市量远大于需求,由于销售不畅导致价格下降,生产者发现养猪赔钱,于是转而经营其他农副业;过一段时间猪肉上市量就会大减,供不应求将导致价格上涨;生产者看到有利可图,又重操旧业,这样下一个时期会重现供大于求、价格下降的局面。在没有外界干预的情况下,这种现象将如此循环下去。

在完全自由竞争的市场经济中,上述现象通常是不可避免的。因为商品的价格是由消费者的需求关系决定的,商品数量越多价格越低。而下一时期商品的数

量由生产者的供应关系决定,商品价格越低生产的数量就越少.这样的需求和供应关系决定了市场经济中商品的价格和数量必然是振荡的.在现实世界里这样的振荡会出现不同的形式,有的振幅渐小趋向平稳,有的则振幅越来越大,如果没有外界如政府的干预,将导致经济崩溃.

本节先用图形方法建立所谓“蛛网模型”,对上述现象进行分析,给出市场经济趋于稳定的条件;再用差分方程建模,对结果进行解释,并讨论当市场经济不稳定时政府可以采取什么样的干预措施;最后对上述模型作适当推广^[54].

蛛网模型 记第 k 时段商品的数量为 x_k , 价格为 $y_k, k=1, 2, \dots$. 这里我们把时间离散化为时段, 1 个时段相当于商品的 1 个生产周期, 如蔬菜、水果是一个种植周期, 肉类是牲畜的饲养周期.

同一时段商品的价格 y_k 取决于数量 x_k , 设

$$y_k = f(x_k) \quad (1)$$

它反映消费者对这种商品的需求关系, 称需求函数. 因为商品的数量越多价格越低, 所以在图 1 中用一条下降曲线 f 表示它, f 称需求曲线.

下一时段商品的数量 x_{k+1} 由上一时段价格 y_k 决定, 设

$$x_{k+1} = h(y_k) \text{ 或 } y_k = g(x_{k+1}) \quad (2)$$

这里 g 是 h 的反函数. h 或 g 反映生产者的供应关系, 称供应函数. 因为价格越高生产量(即下一时段的商品数量)越大, 所以在图中供应曲线 g 是一条上升曲线.

图中两条曲线相交于 $P_0(x_0, y_0)$ 点. P_0 是平衡点, 其意义是, 一旦在某一时段 k 有 $x_k = x_0$, 则由 (1), (2) 可知 $y_k = y_0, x_{k+1} = x_0, y_{k+1} = y_0, \dots$, 即 k 以后各时段商品的数量和价格将永远保持在 $P_0(x_0, y_0)$ 点. 但是实际生活中的种种干扰使得数量和价格不可能停止在 P_0 点, 不妨设 x_1 偏离 x_0 (如图 1). 我们分析随着 k 的增加, x_k 和 y_k 的变化.

商品数量 x_1 给定后, 价格 y_1 由曲线 f 上的 P_1 点决定, 下一时段的数量 x_2 由曲线 g 上的 P_2 点决定, y_2 又由 f 上的 P_3 点决定, 这样得到一系列的点 $P_1(x_1, y_1), P_2(x_2, y_1), P_3(x_2, y_2), P_4(x_3, y_2), \dots$, 在图 1 上这些点将按箭头所示方向趋向 $P_0(x_0, y_0)$, 表明 P_0 是稳定的平衡点, 意味着市场经济(商品的数量和价格)将趋向稳定.

但是如果需求函数和供应函数由图 2 的曲线所示, 则类似的分析发现, 市场经济将按照 $P_1, P_2, P_3, P_4, \dots$ 的规律变化而远离 P_0 , 即 P_0 是不稳定的平衡点, 意味着商品数量和价格将出现越来越大的振荡.

图 1 和图 2 中折线 $P_1P_2P_3P_4 \dots$ 形似蛛网, 所以这种用需求曲线和供应曲线分析市场经济稳定性的图示法在经济学中称蛛网模型. 实际上, 需求曲线 f 和供应曲线 g 的具体形式通常是根据各个时段商品的数量和价格的一系列统计资料

得到的.一般地说, f 取决于消费者对这种商品的需要程度和他们的消费水平, g 则与生产者的生产能力、经营水平等因素有关.比如当消费者收入增加时, f 会向上移动;当生产能力提高时, g 将向右移动.

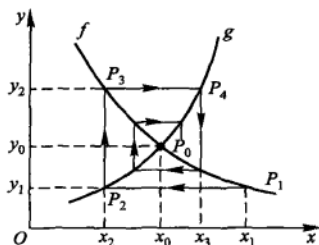


图1 需求曲线 f 和供应曲线 g , P_0 是稳定平衡点

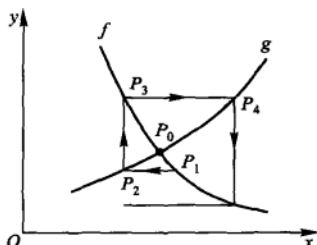


图2 P_0 是不稳定平衡点

一旦需求曲线和供应曲线被确定下来,商品数量和价格是否趋向稳定,就完全由这两条曲线在平衡点 P_0 附近的形状决定.只要分析一下图1和图2的不同之处就会发现,在 P_0 附近,图1的 f 比 g 平缓,而图2的 f 比 g 陡峭.记 f 在 P_0 点斜率的绝对值(因为它是下降的)为 K_f , g 在 P_0 点的斜率为 K_g ,图形的直观告诉我们,当

$$K_f < K_g \quad (3)$$

时, P_0 点是稳定的(图1),当

$$K_f > K_g \quad (4)$$

时, P_0 点是不稳定的(图2).由此可见,需求曲线越平,供应曲线越陡,越有利于经济稳定.为了进一步分析这种现象,下面给出蛛网模型的另一种表达形式——差分方程.

差分方程模型 在 P_0 点附近可以用直线来近似曲线 f 和 h ,设(1),(2)式分别近似为

$$y_k - y_0 = -\alpha(x_k - x_0), \alpha > 0 \quad (5)$$

$$x_{k+1} - x_0 = \beta(y_k - y_0), \beta > 0 \quad (6)$$

从二式中消去 y_k ,可得

$$x_{k+1} - x_0 = -\alpha\beta(x_k - x_0), k = 1, 2, \dots \quad (7)$$

(7)是一阶线性常系数差分方程,对 k 递推不难得到

$$x_{k+1} - x_0 = (-\alpha\beta)^k (x_1 - x_0) \quad (8)$$

容易看出,当 $k \rightarrow \infty$ 时 $x_k \rightarrow x_0$,即 P_0 点稳定的条件是

$$\alpha\beta < 1 \text{ 或 } \alpha < \frac{1}{\beta} \quad (9)$$

而 $k \rightarrow \infty$ 时 $x_k \rightarrow \infty$,即 P_0 点不稳定的条件是

$$\alpha\beta > 1 \text{ 或 } \alpha > \frac{1}{\beta} \quad (10)$$

注意到(5),(6)式中 α, β 的定义,有 $K_f = \alpha, K_g = 1/\beta$,所以条件(9),(10)与蛛网模型中的直观结果(3),(4)式是一致的。

模型解释 首先考察参数 α, β 的含义。由(5)式可知, α 表示商品供应量减少1个单位时价格的上涨幅度;由(6)式可知, β 表示价格上涨1个单位时(下一时期)商品供应的增加量。所以 α 的数值反映消费者对商品需求的敏感程度,如果这种商品是生活必需品,消费者处于持币待购状态,商品数量稍缺,人们立即蜂拥抢购,那么 α 会比较大;反之,若这种商品非必需品,消费者购物心理稳定,或者消费水平低下,则 α 较小。 β 的数值反映生产经营者对商品价格的敏感程度,如果他们目光短浅,热衷于追逐一时的高利润,价格稍有上涨就大量增加生产,那么 β 会比较大;反之,若他们素质较高,有长远的计划,则 β 较小。

根据 α, β 的意义,很容易对市场经济稳定与否的条件(9),(10)作出解释。当供应函数 g 即 β 固定时, α 越小,需求曲线越平,表明消费者对商品需求的敏感程度越小,(9)式越容易成立,有利于经济稳定。当需求函数 f 即 α 固定时, β 越小,供应曲线越陡,表明生产者对价格的敏感程度越小,(9)式也容易成立,有利于经济稳定。反之,当 α, β 较大,表明消费者对商品的需求和生产者对商品的价格都很敏感,则会导致(10)式成立,经济不稳定。

应该指出, α 和 β 都是有量纲的,它们的大小都应在同一量纲单位下比较。同时, α 和 β 的量纲互为倒数,所以 $\alpha\beta$ 是无量纲量,就可以与1比较大小了。

经济不稳定时的干预办法 基于上述分析我们还可以看到,当市场经济趋向不稳定时,政府有两种干预办法。

一种办法是使 α 尽量小,不妨考察极端情况 $\alpha = 0$,即需求曲线水平(图3),这时不论供应曲线如何(即不管 β 多大),(9)式总成立,经济总是稳定的。实际上这种办法相当于政府控制物价,无论商品数量多少,命令价格不得改变。

另一种办法是使 β 尽量小,极端情况是 $\beta = 0$,即供应曲线竖直(图4),于是不论需求曲线如何(不管 α 多大),也总是稳定的。实际上这相当于控制市场上的商品数量,当供应量少于需求时,从外地收购或调拨,投入市场;当供过于求时,收购过剩部分,维持商品上市量不变。显然,这种办法需要政府具有相当强的经济实力。

模型的推广 如果生产者的管理水平和素质更高一些,他们在决定商品生产数量 x_{k+1} 时,不是仅仅根据前一时期的价格 y_k ,而是根据前两个时期的价格 y_k 和 y_{k-1} 。为简单起见,不妨设根据二者的平均值 $(y_k + y_{k-1})/2$,于是供应函数(2)式表为

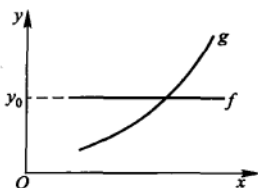


图3 第一种干预办法示意图

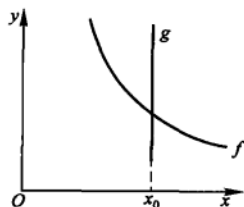


图4 第二种干预办法示意图

$$x_{k+1} = g\left(\frac{y_k + y_{k-1}}{2}\right) \quad (11)$$

相应地, (2)式的线性近似表达式(6)修改为

$$x_{k+1} - x_0 = \frac{\beta}{2}(y_k + y_{k-1} - 2y_0) \quad (12)$$

其中 β 是平均价格上涨1个单位时 x_{k+1} 的增量. 又设需求函数仍由(1), (5)式表示, 则由(5), (12)式得到

$$2x_{k+2} + \alpha\beta x_{k+1} + \alpha\beta x_k = (1 + \alpha\beta)x_0, k = 1, 2, \dots \quad (13)$$

(13)是二阶线性常系数差分方程. 为寻求 $k \rightarrow \infty$ 时 $x_k \rightarrow x_0$, 即 P_0 点稳定的条件, 不必解方程(13), 只需利用判断稳定的条件——方程特征根均在单位圆内(见7.7节(28)式).

由方程(13)的特征方程

$$2\lambda^2 + \alpha\beta\lambda + \alpha\beta = 0$$

容易算出其特征根为

$$\lambda_{1,2} = \frac{-\alpha\beta \pm \sqrt{(\alpha\beta)^2 - 8\alpha\beta}}{4} \quad (14)$$

当 $\alpha\beta > 8$ 时, 显然有

$$\lambda_2 = \frac{-\alpha\beta - \sqrt{(\alpha\beta)^2 - 8\alpha\beta}}{4} < -\frac{\alpha\beta}{4}$$

从而 $|\lambda_2| > 2$, λ_2 在单位圆外. 下面设 $\alpha\beta < 8$, 由(14)式可以算出

$$|\lambda_{1,2}| = \sqrt{\frac{\alpha\beta}{2}} \quad (15)$$

要使特征根均在单位圆内, 即 $|\lambda_{1,2}| < 1$, 必须有

$$\alpha\beta < 2 \quad (16)$$

这就是 P_0 点稳定的条件. 与原有模型中 P_0 点稳定的条件(9)式相比, 参数 α, β 的范围放大了. 可以想到, 这是因为生产者的管理水平和素质提高, 对市场经济的稳定起着有利影响的必然结果.

6.5 减肥计划——节食与运动

您的体重正常吗?不妨用联合国世界卫生组织颁布的所谓**体重指数**(简记 BMI)衡量一下, BMI 定义为体重(单位: kg)除以身高(单位: m)的平方, 规定 BMI 在 18.5 至 25 为正常, 大于 25 为超重, 超过 30 则为肥胖. 据悉, 我国有关机构针对东方人的特点, 拟将上述规定中的 25 改为 24, 30 改为 29.

在国人初步过上小康生活以后, 不少自感肥胖的人纷纷奔向减肥食品的柜台. 可是大量事实说明, 多数减肥食品达不到减肥的目的, 或者即使能减肥一时, 也难以维持下去. 许多医生和专家的意见是, 只有通过控制饮食和适当的运动, 才能在不伤害身体的条件下, 达到减轻体重并维持下去的目的. 本节要建立一个简单的体重变化规律的模型, 并由此通过节食与运动制订合理、有效的减肥计划.

模型分析 通常, 当体内能量守恒被破坏时就会引起体重的变化. 人们通过饮食吸收热量, 转化为脂肪等, 导致体重增加; 又由于代谢和运动消耗热量, 引起体重减少. 只要作适当的简化假设就可得到体重变化的关系.

减肥计划应以不伤害身体为前提, 这可以用吸收热量不要过少、减少体重不要过快来表达. 当然, 增加运动量是加速减肥的有效手段, 也要在模型中加以考虑.

通常, 制订减肥计划以周为时间单位比较方便, 所以这里用离散时间模型——差分方程模型来讨论.

模型假设 根据上述分析, 参考有关生理数据, 作出以下简化假设:

1. 体重增加正比于吸收的热量, 平均每 8 000 kcal (kcal 为非国际单位制单位, $1 \text{ kcal} = 4.2 \text{ kJ}$) 增加体重 1 kg.
2. 正常代谢引起的体重减少正比于体重, 每周每 kg 体重消耗热量一般在 200 kcal 至 320 kcal 之间, 且因人而异, 这相当于体重 70 kg 的人每天消耗 2 000 kcal ~ 3 200 kcal.
3. 运动引起的体重减少正比于体重, 且与运动形式有关.
4. 为了安全与健康, 每周体重减少不宜超过 1.5 kg, 每周吸收热量不要小于 10 000 kcal.

基本模型 记第 k 周末体重为 $w(k)$, 第 k 周吸收热量为 $c(k)$, 热量转换系数 $\alpha = 1/8\,000 \text{ kg/kcal}$, 代谢消耗系数 β (因人而异), 则在不考虑运动情况下体重变化的基本方程为

$$w(k+1) = w(k) + \alpha c(k+1) - \beta w(k), k=0, 1, 2, \dots \quad (1)$$

增加运动时, 只需将 β 改为 $\beta + \beta_1$, β_1 由运动的形式和时间决定.

减肥计划的提出 通过制订一个具体的减肥计划讨论模型(1)的应用.

某甲身高 1.7 m, 体重 100 kg, BMI 高达 34.6. 自述目前每周吸收 20 000 kcal 热量, 体重长期不变. 试为他按照以下方式制订减肥计划, 使其体重减至 75 kg 并维持下去:

1) 在基本上不运动的情况下安排一个两阶段计划, 第一阶段: 每周减肥 1 kg, 每周吸收热量逐渐减少, 直至达到安全的下限 (10 000 kcal); 第二阶段: 每周吸收热量保持下限, 减肥达到目标.

2) 若要加快进程, 第二阶段增加运动, 重新安排第二阶段计划.

3) 给出达到目标后维持体重的方案.

减肥计划的制订

1) 首先应确定某甲的代谢消耗系数 β . 根据他每周吸收 $c = 20\ 000$ kcal 热量, 体重 $w = 100$ kg 不变, 由(1)式得

$$w = w + \alpha c - \beta w, \quad \beta = \alpha c / w = 20\ 000 / 8\ 000 / 100 = 0.025$$

相当于每周每 kg 体重消耗热量 $20\ 000 / 100 = 200$ kcal. 从假设 2 可以知道, 某甲属于代谢消耗相当弱的人. 他又吃得那么多, 难怪如此之胖.

第一阶段要求体重每周减少 $b = 1$ kg, 吸收热量减至下限 $c_{\min} = 10\ 000$ kcal, 即

$$w(k) - w(k+1) = b, \quad w(k) = w(0) - bk$$

由基本模型(1)式可得

$$c(k+1) = \frac{1}{\alpha} [\beta w(k) - b] = \frac{\beta}{\alpha} w(0) - \frac{b}{\alpha} (1 + \beta k)$$

将 α, β, b 的数值代入, 并考虑下限 c_{\min} , 有

$$c(k+1) = 12\ 000 - 200k, \quad k \geq c_{\min} = 10\ 000$$

得 $k \leq 10$, 即第一阶段共 10 周, 按照

$$c(k+1) = 12\ 000 - 200k, \quad k = 0, 1, \dots, 9 \quad (2)$$

吸收热量, 可使体重每周减少 1 kg, 至第 10 周末达到 90 kg.

第二阶段要求每周吸收热量保持下限 c_{\min} , 由基本模型(1)式可得

$$w(k+1) = (1 - \beta)w(k) + \alpha c_{\min} \quad (3)$$

为了得到体重减至 75 kg 所需的周数, 将(3)式递推可得

$$\begin{aligned} w(k+n) &= (1 - \beta)^n w(k) + \alpha c_{\min} [1 + (1 - \beta) + \dots + (1 - \beta)^{n-1}] \\ &= (1 - \beta)^n [w(k) - \alpha c_{\min} / \beta] + \alpha c_{\min} / \beta \end{aligned} \quad (4)$$

已知 $w(k) = 90$, 要求 $w(k+n) = 75$, 再以 α, β, c_{\min} 的数值代入(4)式, 有

$$75 = 0.975^n (90 - 50) + 50 \quad (5)$$

得到 $n = 19$, 即每周吸收热量保持下限 10 000 kcal, 再有 19 周体重可减至 75 kg.

2) 为加快进程,第二阶段增加运动. 通过调查资料得到以下各项运动每小时每千克体重消耗的热量:

运动	跑步	跳舞	乒乓球	自行车 (中速)	游泳 (50 m/min)
热量消耗/kcal	7.0	3.0	4.4	2.5	7.9

记表中热量消耗 γ , 每周运动时间 t , 为利用基本模型(1)式, 只需将 β 改为 $\beta + \alpha\gamma t$, 即

$$w(k+1) = w(k) + \alpha c(k+1) - (\beta + \alpha\gamma t)w(k) \quad (6)$$

试取 $\alpha\gamma t = 0.003$, 即 $\gamma t = 24$, 则(4)式中的 $\beta = 0.025$ 应改成 $\beta + \alpha\gamma t = 0.028$, (5)式为

$$75 = 0.972^n (90 - 44.6) + 44.6 \quad (7)$$

得到 $n = 14$, 即若增加 $\gamma t = 24$ 的运动(如每周跳舞 8 h 或自行车 10 h), 就可将第二阶段的时间缩短为 14 周.

3) 最简单的维持体重 75 kg 的方案, 是寻求每周吸收热量保持某常数 c , 使 $w(k)$ 不变. 由(6)式得

$$\begin{aligned} w &= w + \alpha c - (\beta + \alpha\gamma t)w \\ c &= (\beta + \alpha\gamma t)w/\alpha \end{aligned} \quad (8)$$

若不运动, 容易算出 $c = 15\,000$ kcal; 若运动(内容同上), 则 $c = 16\,800$ kcal.

评注 人体体重的变化是有规律可循的, 减肥也应科学化、量化. 这个模型虽然只考虑了一个非常简单的情况, 但是它对专门从事减肥这项活动(甚至作为一项事业)的人来说也不无参考价值.

体重的变化与每个人特殊的生理条件有关, 特别是代谢消耗系数 β , 不仅因人而异, 而且即使同一个人不同环境下也会有所改变. 从上面的计算中我们看到, 当 β 由 0.025 增加到 0.028 时(变化约 12%), 减肥所需时间就从 19 周减少到 14 周(变化约 26%), 所以应用这个模型时要对 β 作仔细的核对.

6.6 按年龄分组的人口模型

微分方程模型一章曾经介绍过连续型(时间和年龄均为连续量)的人口模型, 在实际的人口统计、预测和控制等方面的工作及研究中, 更方便、更常用的是离散型(时间和年龄均为离散量)的模型——差分方程组.

从数量关系看, 人口的增减与一般动物种群数量的变化规律基本上是相同的, 只不过人类在某种程度上可以约束和控制自身数量的改变, 而动物数量的变

化或者如野生种群那样,依赖于自然环境的制约及种群间的相互竞争和依存,或者如圈养牲畜那样,依赖于饲养者在一定社会经济环境下所作的决策.下面我们先研究动物种群数量的一般变化规律,介绍 Leslie 模型,再在此基础上,将考虑年龄结构和生育模式的连续型人口模型离散化,建立按年龄分组的人口模型.

按年龄分组的种群增长模型——Leslie 模型 种群是直接通过雌性个体的繁殖而增长的,所以用雌性个体数量的变化为研究对象比较方便.下面提到的种群数量均指其中的雌性,总体数量可按一定的性别比算出.

将种群按年龄大小等间隔地分成 n 个年龄组,如每 1 岁或 5 岁为 1 组.与之相对应,时间也分成与年龄组区间大小相等的时段,如 1 年或 5 年为 1 个时段.记时段 k 第 i 年龄组的种群数量为 $x_i(k)$, $k=0,1,2,\dots, i=1,2,\dots,n$.

在稳定的环境下和不太长的时期内,合理地假设种群的繁殖率和死亡率不随时段 k 变化,只与年龄组有关.记第 i 年龄组的繁殖率为 b_i ,即每个(雌性)个体在 1 个时段内繁殖的数量;记第 i 年龄组的死亡率为 d_i ,即 1 个时段内死亡数量(占总量)的比例. $s_i = 1 - d_i$ 称为存活率.通常, b_i 和 s_i 可由统计资料获得,且有以下性质:

$$b_i \geq 0, i=1,2,\dots,n, \text{ 且至少有一个 } b_i > 0; \quad 0 < s_i \leq 1, i=1,2,\dots,n-1$$

种群数量 $x_i(k)$ 的变化规律由 2 个基本关系得到:时段 $k+1$ 第 1 年龄组的数量是各年龄组在时段 k 的繁殖数量之和;时段 $k+1$ 第 $i+1$ 年龄组 ($i=1,2,\dots,n-1$) 的数量是时段 k 第 i 年龄组存活下来的数量.由此得到

$$x_1(k+1) = \sum_{i=1}^n b_i x_i(k), k=0,1,2,\dots \quad (1)$$

$$x_{i+1}(k+1) = s_i x_i(k), k=0,1,2,\dots, i=1,2,\dots,n-1 \quad (2)$$

(1),(2)是差分方程组.记种群数量在时段 k 按年龄组的分布向量为

$$x(k) = [x_1(k), x_2(k), \dots, x_n(k)]^T, k=0,1,2,\dots \quad (3)$$

由繁殖率 b_i 和存活率 s_i 构成的矩阵

$$L = \begin{bmatrix} b_1 & b_2 & \cdots & b_{n-1} & b_n \\ s_1 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & s_{n-1} & 0 \end{bmatrix} \quad (4)$$

则(1),(2)可表为

$$x(k+1) = Lx(k), k=0,1,2,\dots \quad (5)$$

当矩阵 L 和按年龄组的初始分布 $x(0)$ 已知时,可以预测种群数量在时段 k 按年龄组的分布为

$$x(k) = L^k x(0), k = 1, 2, \dots \quad (6)$$

有了 $x(k)$, 不难算出种群在时段 k 的总数.

显然, 种群数量的变化规律完全由矩阵 L 确定, 这个矩阵是 20 世纪 40 年代由 Leslie 提出的, 称 Leslie 矩阵 (简称 L 矩阵), 由 (3) ~ (5) 给出的模型称 Leslie 模型.

Leslie 模型的稳态分析 下面研究时间充分长以后 (即 $k \rightarrow \infty$), 种群的年龄结构和数量的变化.

首先不加证明地叙述关于 L 矩阵的两个定理.

定理 1 L 矩阵有惟一的单重正特征根 λ_1 和正特征向量

$$x^* = [1, s_1/\lambda_1, s_1 s_2/\lambda_1^2, \dots, s_1 s_2 \dots s_{n-1}/\lambda_1^{n-1}]^T \quad (7)$$

其他特征根 λ_k 满足

$$|\lambda_k| \leq \lambda_1, k = 2, 3, \dots, n \quad (8)$$

该定理表明, L 矩阵的特征方程

$$\lambda^n - (b_1 \lambda^{n-1} + b_2 s_1 \lambda^{n-2} + \dots + b_{n-1} s_1 \dots s_{n-2} \lambda + b_n s_1 \dots s_{n-1}) = 0 \quad (9)$$

只有一个正单根 λ_1 , 且容易验证 $Lx^* = \lambda_1 x^*$.

定理 2 若 L 矩阵第 1 行有 2 个顺次的 b_i, b_{i+1} 大于 0, 则定理 1(8) 式中仅不等号成立, 且 (6) 式表示的 $x(k)$ 满足

$$\lim_{k \rightarrow \infty} \frac{x(k)}{\lambda_1^k} = cx^* \quad (10)$$

其中 c 是由 b_i, s_i 和 $x(0)$ 决定的常数.

对于 Leslie 模型, 定理 2 的条件通常是成立的, 由此可对 k 充分大以后种群的年龄结构和数量 $x(k)$ 作出如下分析 (为简便起见, λ_1 以下记作 λ):

1) 由 (10) 式直接有

$$x(k) \approx c\lambda^k x^*, k \text{ 充分大} \quad (11)$$

表明种群的年龄结构趋向稳定, 各年龄组的数量占总量的比例与 (7) 的特征向量 x^* 一致, 故 x^* 可称为稳定分布, 它与初始分布无关.

2) 由 (11) 式又有

$$x(k+1) \approx \lambda x(k), k \text{ 充分大} \quad (12)$$

即

$$x_i(k+1) \approx \lambda x_i(k), k \text{ 充分大}, i = 1, 2, \dots, n \quad (13)$$

表明种群各年龄组的数量也趋向稳定, 都是上时段同一年龄组的 λ 倍. 显然, $\lambda > 1$ 时种群各年龄组的数量递增, $\lambda < 1$ 时递减, λ 称为固有增长率. 不妨将 (12) 给出的近似关系与原模型 (5) 加以对比.

3) 当固有增长率 $\lambda = 1$ 时, 由 (11) 和 (7) 式得

$$x(k) \approx cx^* = c[1, s_1, s_1 s_2, \dots, s_1 s_2 \dots s_{n-1}]^T, k \text{ 充分大} \quad (14)$$

这个结果表明

$$x_{i+1}(k) \approx s_i x_i(k), k \text{ 充分大}, i = 1, 2, \dots, n-1 \quad (15)$$

即存活率 s_i 等于同一时段相邻年龄组 (i 与 $i+1$) 的种群数量之比. 如果将 (15) 给出的近似关系与原模型 (2) 进行对比, 能说明什么?

实际上, 由 (9) 式, 固有增长率 $\lambda = 1$ 等价于

$$b_1 + b_2 s_1 + \dots + b_{n-1} s_1 \dots s_{n-2} + b_n s_1 \dots s_{n-1} = 1 \quad (16)$$

若将 (16) 的左端记作

$$R = b_1 + b_2 s_1 + \dots + b_{n-1} s_1 \dots s_{n-2} + b_n s_1 \dots s_{n-1} \quad (17)$$

可知 R 表示一个 (雌性) 个体在整个存活期内繁殖的平均数量, 称为总和繁殖率, 容易证明, $R > 1$ 时种群数量增长, $R < 1$ 时种群数量减少. 对于人工饲养的动物, 可以通过调节各年龄组的繁殖率 b_i 和存活率 s_i 来改变总和繁殖率 R .

按年龄分组的人口模型 按照 Leslie 模型的基本思路, 将考虑年龄结构和生育模式的连续型人口模型离散化, 即可得到离散形式的人口模型^[88]. 这里仍不考虑迁移等社会因素的影响.

用 $x_i(t)$ 表示第 t 年 i 岁 (指满 i 岁但不到 $i+1$ 岁) 的总人数, $t = 0, 1, 2, \dots$, $i = 0, 1, 2, \dots, n-1$ (设 n 为最高年龄), $b_i(t)$ 表示第 t 年 i 岁女性生育率 (每位女性平均生育的婴儿数), 育龄区间为 $[i_1, i_2]$. 简化地假设女性比与时间无关, 用 k_i 表示 i 岁人口的女性比. 于是第 t 年出生的婴儿数为

$$f(t) = \sum_{i=i_1}^{i_2} b_i(t) k_i x_i(t) \quad (18)$$

引入生育模式 h_i , 将 $b_i(t)$ 分解为

$$b_i(t) = \beta(t) h_i, \sum_{i=i_1}^{i_2} h_i = 1 \quad (19)$$

这里简化地假设生育模式只与年龄有关, 其具体形式可取连续型人口模型给出的 Γ 分布 (见 5.6 节 (22), (23) 式). 由 (18), (19) 式有

$$f(t) = \beta(t) \sum_{i=i_1}^{i_2} h_i k_i x_i(t) \quad (20)$$

$$\beta(t) = \sum_{i=i_1}^{i_2} b_i(t) \quad (21)$$

$\beta(t)$ 是第 t 年所有育龄女性平均生育的婴儿数. 若女性在整个育龄期内保持生育率不变, 则 $\beta(t)$ 就是第 t 年 i_1 岁的每位女性一生平均生育的婴儿数, 即总和生育率 (简称生育率) 或生育胎次, 是控制人口数量的主要参数.

简化地假设死亡率与时间无关, 记 i 岁人口的死亡率为 d_i , 存活率为 $s_i = 1 - d_i$, $i = 0, 1, 2, \dots, n-1$, 则

$$x_{i+1}(t+1) = s_i x_i(t), t = 0, 1, 2, \dots, i = 1, 2, \dots, n-1 \quad (22)$$

而 $x_1(t+1)$ 是第 t 年出生的婴儿中存活下来的数量, 即 $s_0 f(t)$ (这里 $f(t) = x_0(t)$). 于是

$$x_1(t+1) = \beta(t) \sum_{i=1}^{i_2} r_i x_i(t), r_i = s_0 h_i k_i \quad (23)$$

引入按年龄分组的人口分布向量

$$\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T, t = 0, 1, 2, \dots \quad (24)$$

为了清楚地表明 $\beta(t)$ 的作用, 将(4)式的 L 矩阵分解成两个矩阵, 记

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ s_1 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & 0 \\ 0 & 0 & \cdots & s_{n-1} & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & \cdots & 0 & r_{i_1} & \cdots & r_{i_2} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \quad (25)$$

则模型(23), (24)式可表为

$$\mathbf{x}(t+1) = A\mathbf{x}(t) + \beta(t)B\mathbf{x}(t) \quad (26)$$

由统计资料可以得到人口的初始分布 $\mathbf{x}(0)$ 及当前的存活率 s_i 、女性比 k_i , 只需合理地给定今后的 s_i 、 k_i 及生育模式 h_i , 就可以设定不同的总和生育率 $\beta(t)$, 来预测或控制未来的人口总量和年龄结构。

在控制理论中, $\mathbf{x}(t)$ 为状态变量, $\beta(t)$ 为控制变量, (26)式对于 $\mathbf{x}(t)$ 和 $\beta(t)$ 都是线性的, 称双线性方程。

在离散型模型中, $\mathbf{x}(t)$ 是人口发展过程的完整描述, 像连续型模型一样, 也常用一些人口指数简要地表示一个国家或地区的人口特征, 如

人口总数

$$N(t) = \sum_{i=0}^n x_i(t) \quad (27)$$

平均年龄

$$R(t) = \frac{1}{N(t)} \sum_{i=0}^n i x_i(t) \quad (28)$$

平均寿命

$$S(t) = \sum_{j=0}^n \exp \left(- \sum_{i=0}^j d_i(t) \right) \quad (29)$$

离散形式模型的优点是便于用计算机作数值计算, 上个世纪 70 年代末曾用这类模型, 根据 1978 年的统计资料对 21 世纪我国人口总数作预测。在不同的总和生育率 β 下得到了 1980—2080 年一系列结果, 图 1 是这些结果的略图。计算结果表明:

1) 若 $\beta = 3.0$ (20 世纪 70 年代中期水平), 则 2000 年将达 14.2 亿, 2080 年

达 43.1 亿,近于那时全世界人口总和。

2) 若 $\beta = 2.3$ (约为 1980 年水平), 则 2000 年将达 12.9 亿, 2080 年为 21.2 亿。

3) 若 $\beta = 2.0$ (大约是保持人口长期稳定的水平), 则 2000 年为 12.2 亿, 72 年后达到最大值, 此后略有下降。

4) 若 $\beta = 1.5$, 则在 2007 年达到最大值, 到 2080 年降至 7.8 亿 (1968 年水平)。

5) 若 $\beta = 1.0$, 即全国严格执行一对夫妇只生一个孩子的政策, 则在 2004 年达到最大值 10.6 亿, 50 年后降至 9.5 亿 (1978 年水平)。

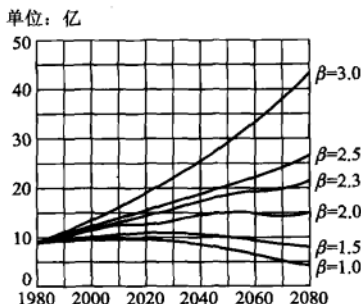


图 1 我国人口预测

习 题

1. 速度为 v 的风吹在迎风面积为 s 的风车上, 空气密度是 ρ . 用量纲分析方法确定风车获得的功率 P 与 v, s, ρ 的关系。

2. 雨滴的速度 v 与空气密度 ρ 、粘滞系数 μ 和重力加速度 g 有关, 其中粘滞系数的定义是: 运动物体在流体中受的摩擦力与速度梯度和接触面积的乘积成正比, 比例系数为粘滞系数。用量纲分析方法给出速度 v 的表达式。

3. 用量纲分析方法研究人体浸在匀速流动的水里时损失的热量。记水的流速 v , 密度 ρ , 比热 c , 粘性系数 μ , 热传导系数 k , 人体尺寸 d . 证明人体与水的热交换系数 h 与上述各物理量的关系可表为 $h = \frac{k}{d} \varphi\left(\frac{v d}{\mu}, \frac{\mu c}{k}\right)$, φ 是未定函数, h 定义为单位时间内人体的单位面积在人体与水的温差为 1°C 时的热量交换^[39]。

4. 用量纲分析方法研究两带电平行板间的引力。板的面积为 s , 间距为 d , 电位差为 v , 板间介质的介电常数为 ϵ , 证明两板之间的引力 $f = \epsilon v^2 \varphi(s/d^2)$. 如果又知道 f 与 s 成正比, 写出 f 的表达式。这里介电常数 ϵ 的定义是 $f = \frac{q_1 q_2}{\epsilon d^2}$, 其中 q_1, q_2 是两个点电荷的电量, d 是点电荷的距离, f 是点电荷间的引力^[39]。

5. 考察阻尼摆的周期, 即在单摆运动中考虑阻力, 并设阻力与摆的速度成正比。给出周期的表达式, 然后讨论物理模拟的比例模型, 即怎样由模型摆的周期计算原型摆的周期。

6. 考察模拟水下爆炸的比例模型。爆炸物质量 m , 在距爆炸点距离 r 处设置仪器, 接收到的冲击波压强为 p , 记大气初始压强 p_0 , 水的密度 ρ , 水的体积弹性模量 k , 用量纲分析方法已经得到 $p = p_0 \varphi\left(\frac{p_0}{k}, \frac{\rho r^3}{m}\right)$. 设模拟实验与现场的 p_0, ρ, k 相同, 而爆炸物模型的质量为原型的 $1/1000$. 为了使实验中接收到与现场相同的压强 p , 问实验时应如何设置接收冲击波的仪器, 即求实验仪器与爆炸点之间的距离是现场的多少倍^[39]。

7. 质量为 m 的小球以速度 v 竖直上抛, 阻力与速度成正比, 比例系数 k . 设初始位置为 $x=0$, x 轴竖直向上, 则运动方程为

$$m\ddot{x} + k\dot{x} + mg = 0, x(0) = 0, \dot{x}(0) = v$$

方程的解可表为 $x = x(t; v, g, m, k)$. 试选择两种特征尺度将问题无量纲化, 并讨论 k 很小时求近似解的可能性^[48].

8. 对于 6.4 节蛛网模型讨论下列问题:

(1) 因为一个时段上市的商品不能立即售完, 其数量也会影响到下一时段的价格, 所以第 $k+1$ 时段的价格 y_{k+1} 由第 $k+1$ 和第 k 时段的数量 x_{k+1} 和 x_k 决定. 如果设 x_{k+1} 仍只取决于 y_k , 给出稳定平衡的条件, 并与 6.4 节的结果进行比较.

(2) 若除了 y_{k+1} 由 x_{k+1} 和 x_k 决定之外, x_{k+1} 也由前两个时段的价格 y_k 和 y_{k-1} 确定, 试分析稳定平衡的条件是否还会放宽.

9. 在 6.6 节按年龄分组的种群增长模型中, 设一群动物最高年龄为 15 岁, 每 5 岁一组, 分成 3 个年龄组, 各组的繁殖率为 $b_1 = 0$, $b_2 = 4$, $b_3 = 3$, 存活率为 $s_1 = 1/2$, $s_2 = 1/4$, 开始时 3 组各有 1 000 只. 求 15 年后各组分别有多少只, 以及时间充分长以后种群的增长率 (即固有增长率) 和按年龄组的分布.

10. 在 6.6 节按年龄分组的种群增长模型中, 证明: 当时间充分长以后, 若总和繁殖率 $R > 1$, 则种群增长, 若 $R < 1$, 则种群减少.

* 11. 在按年龄分组的种群增长模型基础上, 建立种群的稳定收获模型.

(1) 设年龄组区间、时段长度都正好等于种群的繁殖周期, 种群的按年龄组分布、Leslie 矩阵及增长规律仍用 6.6 节的 $x(k+1) = Lx(k)$ 表示. 如果时段 k 第 i 年龄组种群的增加量就是这个时段的收获量, 表示为 $x_i(k) - x_i(k-1) = h_i(k)x_i(k)$, $i = 1, 2, \dots, n$, $k = 1, 2, \dots$, 其中 $h_i(k)$ 为收获系数.

所谓稳定收获是指, 各个时段同一年龄组的收获量不变, 即 $h_i(k)$ 和 $x_i(k)$ (在收获之后) 与 k 无关. 用 H 表示以 h_i 为对角元素的对角阵, 证明稳定收获模型可表为 $Lx - x = HLx$, 其中 x 是种群的按年龄组的稳定分布.

(2) 证明获得稳定收获的充要条件是: h_i 满足

$$(1 - h_1)[b_1 + b_2s_1(1 - h_2) + \dots + b_ns_1s_2\cdots s_{n-1}(1 - h_2)\cdots(1 - h_n)] = 1$$

且 $x = cx^*$ (c 是大于零的常数), 其中

$$x^* = [1, s_1(1 - h_2), \dots, s_1s_2\cdots s_{n-1}(1 - h_2)\cdots(1 - h_n)]^T$$

(3) 利用第 9 题的数据至少给出 H 和 x 的两组解, 并计算按年龄组稳定收获的分布.

* 12. 讨论稳定收获模型 (第 11 题) 的两个特例.

(1) 有些种群最年幼的级别具有较大的经济价值, 所以饲养者只收获这个年龄组的种群, 于是 $h_1 = h (> 0)$, $h_2 = \dots = h_n = 0$. 给出这种情况下稳定收获的充要条件, 并在第 9 题数据下求收获系数、种群的稳定分布和收获量按年龄组的稳定分布.

(2) 对于随机捕获的种群, 区分年龄是困难的, 不妨假定 $h_1 = \dots = h_n = h$. 讨论与 (1) 同样的问题.

第7章

稳定性模型

虽然动态过程的变化规律一般要用微分方程建立的动态模型来描述,但是对于某些实际问题,建模的主要目的并不是要寻求动态过程每个瞬时的性态,而是研究某种意义上稳定状态的特征,特别是当时间充分长以后动态过程的变化趋势.比如,在什么条件下描述过程的变量会越来越接近某些确定的数值,在什么情况下又会越来越远离这些数值而导致过程不稳定.为了分析这种稳定与不稳定的规律,常常不要求解微分方程(并且我们将会看到,即使对于不太复杂的方程,解析解也不是总能得到的),而可以利用微分方程稳定性理论,直接研究平衡状态的稳定性就行了.

在下面建模过程中,大多数情况下我们将直接引用微分方程稳定性理论的结果.不熟悉这方面内容的读者可以参阅 7.7 节.

差分方程的稳定性与微分方程有很多相似之处,也有明显的差别,7.6 和 7.7 节将介绍这方面的内容.

7.1 捕鱼业的持续收获

可持续发展是一项基本国策,对于像渔业、林业这样的再生资源,一定要注意适度开发,不能为了一时的高产去“竭泽而渔”,应该在持续稳产的前提下追求产量或效益的最优化.

考察一个渔场,其中的鱼量在天然环境下按一定规律增长,如果捕捞量恰好等于增长量,那么渔场鱼量将保持不变,这个捕捞量就可以持续下去.本节要建立在捕捞情况下渔场鱼量遵从的方程,分析鱼量稳定的条件,并且在稳定的前提下讨论如何控制捕捞使持续产量或经济效益达到最大.最后研究所谓捕捞过度的问题^[21,46].

产量模型 记时刻 t 渔场中鱼量为 $x(t)$,关于 $x(t)$ 的自然增长和人工捕捞作如下假设:

1. 在无捕捞条件下 $x(t)$ 的增长服从 logistic 规律(见 5.6 节的阻滞增长模型),即

$$\dot{x}(t) = f(x) = rx\left(1 - \frac{x}{N}\right) \quad (1)$$

r 是固有增长率, N 是环境容许的最大鱼量, 用 $f(x)$ 表示单位时间的增长量.

2. 单位时间的捕捞量 (即产量) 与渔场鱼量 $x(t)$ 成正比, 比例常数 E 表示单位时间捕捞率, 又称为**捕捞强度**, 可以用比如捕鱼网眼的大小或出海渔船数量来控制其大小. 于是单位时间的捕捞量为

$$h(x) = Ex \quad (2)$$

根据以上假设并记

$$F(x) = f(x) - h(x)$$

得到捕捞情况下渔场鱼量满足的方程

$$\dot{x}(t) = F(x) = rx\left(1 - \frac{x}{N}\right) - Ex \quad (3)$$

我们并不需要解方程(3)以得到 $x(t)$ 的动态变化过程, 只希望知道渔场的稳定鱼量和保持稳定的条件, 即时间 t 足够长以后渔场鱼量 $x(t)$ 的趋向, 并由此确定最大持续产量. 为此可以直接求方程(3)的平衡点并分析其稳定性.

令

$$F(x) = rx\left(1 - \frac{x}{N}\right) - Ex = 0$$

得到两个平衡点

$$x_0 = N\left(1 - \frac{E}{r}\right), \quad x_1 = 0 \quad (4)$$

不难算出

$$F'(x_0) = E - r, \quad F'(x_1) = r - E$$

所以若

$$E < r \quad (5)$$

有 $F'(x_0) < 0, F'(x_1) > 0$, 故 x_0 点稳定, x_1 点不稳定 (判断平衡点稳定性的准则见 7.7 节); 若 $E > r$, 则结果正好相反.

E 是捕捞率, r 是最大的增长率, 上述分析表明, 只要捕捞适度 ($E < r$), 就可使渔场鱼量稳定在 x_0 , 从而获得持续产量 $h(x_0) = Ex_0$; 而当捕捞过度时 ($E > r$), 渔场鱼量将趋向 $x_1 = 0$, 当然谈不上获得持续产量了.

进一步讨论渔场鱼量稳定在 x_0 的前提下, 如何控制捕捞强度 E 使持续产量最大的问题. 用图解法可以非常简单地得到

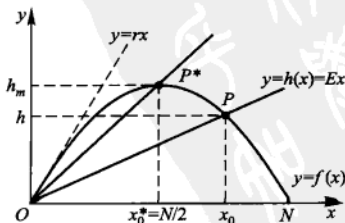


图 1 最大持续产量的图解法

结果.

根据(1),(2)式作抛物线 $y=f(x)$ 和直线 $y=h(x)=Ex$, 如图 1. 注意到 $y=f(x)$ 在原点的切线为 $y=rx$, 所以在条件(5)下 $y=Ex$ 必与 $y=f(x)$ 有交点 P , P 的横坐标就是稳定平衡点 x_0 .

根据假设 2, P 点的纵坐标 h 为稳定条件下单位时间的持续产量. 由图 1 立刻知道, 当 $y=Ex$ 与 $y=f(x)$ 在抛物线顶点 P^* 相交时可获得最大的持续产量, 此时的稳定平衡点为

$$x_0^* = \frac{N}{2} \quad (6)$$

且单位时间的最大持续产量为

$$h_m = \frac{rN}{4} \quad (7)$$

而由(4)式不难算出保持渔场鱼量稳定在 x_0^* 的捕捞率为

$$E^* = \frac{r}{2} \quad (8)$$

综上所述, 产量模型的结论是将捕捞率控制在固有增长率 r 的一半, 更简单一些, 可以说使渔场鱼量保持在最大鱼量 N 的一半时, 能够获得最大的持续产量.

效益模型 从经济角度看不应追求产量最大, 而应考虑效益最佳. 如果经济效益用从捕捞所得的收入中扣除开支后的利润来衡量, 并且简单地假设: 鱼的销售单价为常数 p , 单位捕捞率(如每条出海渔船)的费用为常数 c , 那么单位时间的收入 T 和支出 S 分别为

$$T = ph(x) = pEx, \quad S = cE \quad (9)$$

单位时间的利润为

$$R = T - S = pEx - cE \quad (10)$$

在稳定条件 $x = x_0$ 下, 以(4)代入(10)式, 得

$$R(E) = T(E) - S(E) = pNE \left(1 - \frac{E}{r} \right) - cE \quad (11)$$

用微分法容易求出使利润 $R(E)$ 达到最大的捕捞强度为

$$E_R = \frac{r}{2} \left(1 - \frac{c}{pN} \right) \quad (12)$$

将 E_R 代入(4)式, 可得最大利润下的渔场稳定鱼量 x_R 及单位时间的持续产量 h_R 为

$$x_R = \frac{N}{2} + \frac{c}{2p} \quad (13)$$

$$h_R = rx_R \left(1 - \frac{x_R}{N} \right) = \frac{rN}{4} \left(1 - \frac{c^2}{p^2 N^2} \right) \quad (14)$$

将(12)~(14)式与产量模型中的(6)~(8)式相比较可以看出,在最大效益原则下捕捞率和持续产量均有所减少,而渔场应保持的稳定鱼量有所增加,并且减少或增加的部分随着捕捞成本 c 的增长而变大,随着销售价格 p 的增长而变小. 请读者分析这些结果是符合实际情况的.

捕捞过度 上面的效益模型是以计划捕捞(或称封闭式捕捞)为基础的,即渔场由单独的经营者有计划地捕捞,可以追求最大利润. 如果渔场向众多盲目的经营者开放,比如在公海上无规则地捕捞,那么即使只有微薄的利润,经营者也会蜂拥而去,这种情况称为盲目捕捞(或开放式捕捞). 这种捕捞方式将导致捕捞过度,下面讨论这个模型.

(11)式给出了利润与捕捞强度的关系 $R(E)$, 令 $R(E) = 0$ 的解为 E_s , 可得

$$E_s = r \left(1 - \frac{c}{pN} \right) \quad (15)$$

当 $E < E_s$ 时, 利润 $R(E) > 0$, 盲目的经营者们会加大捕捞强度; 若 $E > E_s$, 利润 $R(E) < 0$, 他们当然要减小强度. 所以 E_s 是盲目捕捞下的临界强度.

E_s 也可由图解法确定. 在图2中以 E 为横坐标, 按(11)式画出 $T(E)$ 和 $S(E)$, 它们交点的横坐标即为 E_s (图2中的 E_{s1} 或 E_{s2}). 由(15)式或图2容易知道, E_s 存在的必要条件(即 $E_s > 0$) 是

$$p > \frac{c}{N} \quad (16)$$

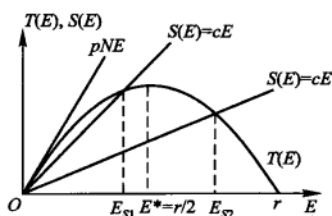


图2 盲目捕捞强度的图解法

即售价大于(相对于总量而言)成本. 并且由(15)式可知, 成本越低, 售价越高, 则 E_s 越大.

将(15)代入(4)式, 得到盲目捕捞下的渔场稳定鱼量为

$$x_s = \frac{c}{p} \quad (17)$$

x_s 完全由成本-价格比决定, 随着价格的上升和成本的下降, x_s 将迅速减少, 出现捕捞过度.

比较(12)和(15)式可知, $E_s = 2E_R$, 即盲目捕捞强度比最大效益下捕捞强度大一倍.

从(15)式和图2还可以得到, 当 $\frac{c}{N} < p < 2\frac{c}{N}$ 时, $(E_R <) E_s < E^*$, 如图2中 E_{s1} , 称经济学捕捞过度; 当 $p > 2\frac{c}{N}$ 时, $E_s > E^*$, 如图2中 E_{s2} , 称生态学捕捞过度.

评注 为了研究渔业的产量、效益及捕捞过度问题,首先在对鱼的自然增长和捕捞情况的合理假设下,建立渔场鱼量的基本方程(3),并利用平衡点稳定性分析确定了保持渔场鱼量稳定的条件.产量、效益和捕捞过度3个模型在稳定的前提下步步深入,数学推导过程十分简单,却得到了在定性关系上与实际情况完全符合的结果.

如果改变对鱼的自然增长和人工捕捞的假设,模型及结果将随之变化(习题1,2).

7.2 军备竞赛

两个国家或两个国家集团之间由于相互不信任和各种矛盾的存在、发展而不断增加自己的军事力量,防御对方可能发动的战争.能否用一个数学模型描述这种军备竞赛的过程,从定性和定量的角度对竞赛的结果作出解释或预测.本节介绍 L. F. Richardson 1939 年提出的一个模型.当然影响军备竞赛的因素是错综复杂的,无法用数学工具给以恰当的圆满的描述.这个模型只不过告诉我们,一个复杂的实际过程可以被合理地简化到什么程度,得到的结果又怎样用来解释实际现象,并且如果允许军备竞赛沿着机械的本能和固定的模式发展的话,会导致什么样的结局^[15,16,32].

模型假设与构成 为了方便起见,用军备这个词表示军事力量的总和,如兵力、装备、军事预算等.甲乙双方在时刻 t 的军备分别记作 $x(t)$ 和 $y(t)$,假定它们的变化只取决于下面3个因素:

1. 由于相互不信任及矛盾的发展,一方军备越大,另一方军备增加得越快.
2. 由于各方本身经济实力的限制,任一方军备越大,对军备增长的制约作用越大.
3. 由于相互敌视或领土争端,每一方都存在着增加军备的固有潜力.

进一步假定前两个因素的影响是线性的,第3个因素的影响是常数,那么 $x(t)$ 和 $y(t)$ 的变化过程可用微分方程组

$$\begin{cases} \dot{x}(t) = -\alpha x + ky + g \\ \dot{y}(t) = lx - \beta y + h \end{cases} \quad (1)$$

表示,其中的系数均大于或等于0. k, l 是对方军备刺激程度的度量; α, β 是己方经济实力制约程度的度量; g, h 是己方军备竞赛的固有潜力.

如果我们感兴趣的是军备竞赛的结局由什么因素决定,而不关心竞赛的过程,那么只需用微分方程稳定性理论讨论时间充分长以后 $x(t), y(t)$ 的变化趋势,即方程(1)的平衡点的稳定情况.

令(1)式右端等于0,容易算出平衡点 (x_0, y_0) 为

$$x_0 = \frac{kh + \beta g}{\alpha\beta - kl}, \quad y_0 = \frac{lg + \alpha h}{\alpha\beta - kl} \quad (2)$$

方程(1)的系数矩阵为

$$A = \begin{bmatrix} -\alpha & k \\ l & -\beta \end{bmatrix}$$

于是按照判断平衡点稳定性的方法计算(见 7.7 节(9)~(13)式)

$$p = -(a_{11} + a_{22}) = \alpha + \beta > 0 \quad (3)$$

$$q = \det A = \alpha\beta - kl \quad (4)$$

由稳定性准则(见 7.7 节(15)式),当

$$\alpha\beta > kl \quad (5)$$

时,平衡点 (x_0, y_0) 是稳定的;反之,是不稳定的。

这就是说,在(5)式的条件下,时间足够长以后双方的军备将分别趋向一个有限值,军备竞赛是稳定的。

模型的定性解释 根据方程(1)和平衡点稳定性的分析,可以解释几个简单而又重要的现象。

1. 条件(5)表明,当双方的经济制约程度 $\alpha\beta$ 大于双方的军备刺激程度 kl 时,军备竞赛才会趋向稳定。反之, $x(t), y(t)$ 将趋向无穷,竞赛无限地进行下去,可能导致战争。

2. 由(2)式,如果 $g = h = 0$,则 $x_0 = 0, y_0 = 0$ 是方程(1)的平衡点,并且在条件(5)下它是稳定的。于是如果在某个时候 t_0 ,有 $x(t_0) = y(t_0) = 0$, x, y 就永远保持为0。这种情况可以解释为双方不存在任何敌视和争端,通过裁军可以达到持久和平。两个友好的邻国正是这样。

3. 如果 $g, h \neq 0$,即使由于某种原因(如裁军协定)在某个时候双方军备大减,不妨设 $x(t_0) = y(t_0) = 0$,那么因为 $\dot{x} = g, \dot{y} = h$,也将使双方重整军备。这说明未经和解的裁军(即不消除敌视或领土争端)是不会持久的。

4. 如果由于某种原因(如战败或协议),在某个时候一方的军备大减,不妨设 $x(t_0) = 0$,那么因为 $\dot{x} = ky + g$,也将使该方重整军备。这说明存在不信任($k \neq 0$)或固有争端($g \neq 0$)的单方面裁军也不会持久。

模型参数的估计 为了利用(5)式判断军备竞赛是否会趋于稳定,需要知道 α, β, k, l 的数值。估计这些参数无疑是很困难的,下面是 Richardson 提出的一种方法。

1. k, l 的估计

设 $x(0) = 0$,当 t 较小时,忽略 g 和 $-\alpha x$ 的作用,并近似地假定 $y = y_1$ 不变,由方程(1)得

$$\dot{x} = ky_1 \quad (6)$$

如果当 $t = \tau$ 时 $x = y_1$, 则由(6)式得到

$$k^{-1} = \tau \quad (7)$$

这说明 k^{-1} 是甲方军备从 0 到赶上乙方军备 y_1 所需的时间.

例如, 德国从 1933 年开始重整军备, 只用了约 3 年的时间就赶上了它的邻国. 假设它增加军备的固有潜力 g 被制约效应 αx 所抵消, 那么可以认为德国的 $k^{-1} \approx 3$ 年, 即 $k \approx 0.3$.

l 可以类似地估计, 或者合理地假定它与国家的经济实力成正比. 这样若乙国的经济实力是德国的 2 倍, 则可以估计 $l \approx 0.6$.

2. α, β 的估计

设 $g = 0, y = 0$, 由方程(1)可得

$$x(t) = x(0)e^{-\alpha t}$$

以 $t = \alpha^{-1}$ 代入, 算出

$$x(\alpha^{-1}) = \frac{x(0)}{e}$$

这表示 α^{-1} 是在乙方无军备时甲方军备减少到原来的 $\frac{1}{e}$ 所需的时间. Richardson 认为这大概是一个国家议会的任期, 对于议会任期 5 年的国家来说, $\alpha \approx 0.2$.

对模型和参数的粗略检验 考察第一次世界大战前夕, 欧洲的两个国家同盟——法俄同盟和德奥匈同盟的军备竞赛情况.

两个同盟的经济实力大致相等, 且约为德国的 3 倍, 因为德国的 $k \approx 0.3$, 所以这两个同盟的 $k = l \approx 0.9$. 同时假定 $\alpha = \beta \approx 0.2$, 那么由于 $\alpha\beta < kl$, (5) 式不成立, 它们的军备竞赛不会趋向稳定.

事实上, 当时两个同盟之间既有军备竞赛也有贸易往来. 用 x_1, y_1 表示双方的军事预算, x_2, y_2 表示双方的贸易往来, 从军事预算中扣除贸易往来作为双方的军备, 即 $x = x_1 - x_2, y = y_1 - y_2$. 以 $k = l, \alpha = \beta$ 代入方程(1), 并将两式相加得到

$$\frac{d}{dt}(x + y) = (k - \alpha)(x + y) + g + h \quad (8)$$

或写作

$$\frac{d}{dt}(x_1 + y_1) = (k - \alpha) \left[(x_1 + y_1) - (x_2 + y_2) + \frac{1}{k - \alpha} \frac{d}{dt}(x_2 + y_2) + \frac{g + h}{k - \alpha} \right] \quad (9)$$

(9) 式表明, $x_1 + y_1$ 与它的变化率的关系是线性的. 为了与实际数据比较, 表 1 列出了两个同盟从 1909 年到 1913 年的军事预算, 表中第 5 行 $\Delta(x_1 + y_1)$ 是 $x_1 + y_1$ 的年增加量, 最后一行 $\overline{x_1 + y_1}$ 是相应的年平均. 用 $\Delta(x_1 + y_1)$ 和 $\overline{x_1 + y_1}$ 的数据作图 (图 1 中圆点), 对这组数据用线性最小二乘拟合得到 (图中直线)

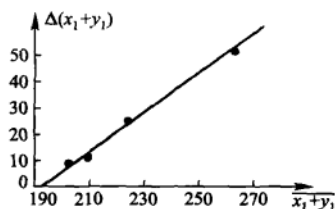


图1 $\overline{x_1 + y_1}$ 与 $\Delta(x_1 + y_1)$ 的关系

$$\Delta(x_1 + y_1) = 0.73(\overline{x_1 + y_1} - 195) \quad (10)$$

表1 两个同盟的军事预算(以百万英镑为单位)

	1909	1910	1911	1912	1913
法俄 x_1	115.3	119.4	127.8	145.0	166.7
德奥匈 y_1	83.9	85.4	87.1	93.7	122.3
$x_1 + y_1$	199.2	204.8	214.9	238.7	289.0
$\Delta(x_1 + y_1)$	5.6	10.1	23.8	50.3	
$\overline{x_1 + y_1}$	202.0	209.8	226.8	263.8	

比较(9)和(10)式,(10)式的线性关系粗略地说明本节介绍的模型具有一定的合理性.同时得到的 $k - \alpha = 0.73$ 也与前面给出的估值 $k \approx 0.9, \alpha \approx 0.2$ 相符,由军事预算体现的军备将继续增加.事实上,两个同盟间的军备竞赛终于引发了第一次世界大战.

评注 用如此简单的模型描述错综复杂的军备竞赛过程也许难以令人信服,但是如果你没有更深入、可靠的知识去建立更满意的模型,那就不妨先做像本节这样的尝试.这种简化模型既是进一步研究的基础,也是建模方法的练习.

7.3 种群的相互竞争

当某个自然环境中只有一种生物的群体(生态学上称为种群)生存时,人们常用 logistic 模型来描述这个种群数量的演变过程,即

$$\dot{x}(t) = rx \left(1 - \frac{x}{N}\right) \quad (1)$$

$x(t)$ 是种群在时刻 t 的数量, r 是固有增长率, N 是环境资源容许的种群最大数

量,在 5.6 节和 7.1 节我们曾应用过这种模型.由方程(1)可以直接得到, $x_0 = N$ 是稳定平衡点,即 $t \rightarrow \infty$ 时 $x(t) \rightarrow N$. 从模型本身的意义看这是明显的结果.

如果一个自然环境中有两个或两个以上种群生存,那么它们之间就要存在着或是相互竞争,或是相互依存,或是弱肉强食(食饵与捕食者)的关系.本节和下面两节将从稳定状态的角度分别讨论这些关系.

当两个种群为了争夺有限的同一种食物来源和生活空间而进行生存竞争时,最常见的结局是竞争力较弱的种群灭绝,竞争力较强的种群达到环境容许的最大数量.人们今天可以看到自然界长期演变成的这样的结局.例如,一个小岛上虽然有四种燕子栖息,但是它们的食物来源各不相同,一种只在陆地上觅食,另两种分别在浅水的海滩上和离岸稍远的海中捕鱼,第四种则飞越宽阔的海面到远方攫取海味,每一种燕子在它各自生存环境中的竞争力明显地强于其他几种.本节要建立一个模型解释类似的现象,并分析产生这种结局的条件^[6,15,49,60,62].

模型建立 有甲乙两个种群,当它们独自在一个自然环境中生存时,数量的演变均遵从 logistic 规律.记 $x_1(t)$, $x_2(t)$ 是两个种群的数量, r_1 , r_2 是它们的固有增长率, N_1 , N_2 是它们的最大容量.于是,对于种群甲有

$$\dot{x}_1(t) = r_1 x_1 \left(1 - \frac{x_1}{N_1} \right)$$

其中因子 $\left(1 - \frac{x_1}{N_1} \right)$ 反映由于甲对有限资源的消耗导致的对它本身增长的阻滞作用, $\frac{x_1}{N_1}$ 可解释为相对于 N_1 而言,单位数量的甲消耗的供养甲的食物量(设食物总量为 1).

当两个种群在同一自然环境中生存时,考察由于乙消耗同一种有限资源对甲的增长产生的影响,可以合理地在因子 $\left(1 - \frac{x_1}{N_1} \right)$ 中再减去一项,该项与种群乙的数量 x_2 (相对于 N_2 而言)成正比,于是得到种群甲增长的方程为

$$\dot{x}_1(t) = r_1 x_1 \left(1 - \frac{x_1}{N_1} - \sigma_1 \frac{x_2}{N_2} \right) \quad (2)$$

这里 σ_1 的意义是:单位数量乙(相对 N_2 而言)消耗的供养甲的食物量为单位数量甲(相对 N_1)消耗的供养甲的食物量的 σ_1 倍.

类似地,甲的存在也影响了乙的增长,种群乙的方程应该是

$$\dot{x}_2(t) = r_2 x_2 \left(1 - \sigma_2 \frac{x_1}{N_1} - \frac{x_2}{N_2} \right) \quad (3)$$

对 σ_2 可作相应的解释.

在两个种群的相互竞争中, σ_1, σ_2 是两个关键指标. 从上面它们的解释可知, $\sigma_1 > 1$ 表示在消耗供养甲的资源中, 乙的消耗多于甲, 因而对甲增长的阻滞作用乙大于甲, 即乙的竞争力强于甲. 对 $\sigma_2 > 1$ 可作相应的理解.

一般地说, σ_1 与 σ_2 之间没有确定的关系, 但是可以把这样一种特殊情况作为较常见的一类实际情况的典型代表, 即两个种群在消耗资源中对甲增长的阻滞作用与对乙增长的阻滞作用相同. 具体地说就是: 因为单位数量的甲和乙消耗的供养甲的食物量之比是 $1:\sigma_1$, 消耗的供养乙的食物量之比是 $\sigma_2:1$, 所谓阻滞作用相同即 $1:\sigma_1 = \sigma_2:1$, 所以这种特殊情形可以定量地表示为

$$\sigma_1 \sigma_2 = 1 \quad (4)$$

即 σ_1, σ_2 互为倒数. 可以简单地理解为, 如果一个乙消耗的食物是一个甲的 $\sigma_1 = k$ 倍, 则一个甲消耗的食物是一个乙的 $\sigma_2 = 1/k$.

下面我们仍然讨论 σ_1, σ_2 相互独立的一般情况, 而将条件(4)下对问题的分析留给读者(习题3).

稳定性分析 为了研究两个种群相互竞争的结局, 即 $t \rightarrow \infty$ 时 $x_1(t), x_2(t)$ 的趋向, 不必要解方程(2), (3)^①, 只需对它的平衡点进行稳定性分析.

首先根据微分方程(2), (3), 解代数方程组

$$\begin{cases} f(x_1, x_2) \equiv r_1 x_1 \left(1 - \frac{x_1}{N_1} - \sigma_1 \frac{x_2}{N_2} \right) = 0 \\ g(x_1, x_2) \equiv r_2 x_2 \left(1 - \sigma_2 \frac{x_1}{N_1} - \frac{x_2}{N_2} \right) = 0 \end{cases} \quad (5)$$

得到4个平衡点:

$$P_1(N_1, 0), P_2(0, N_2), P_3\left(\frac{N_1(1-\sigma_1)}{1-\sigma_1\sigma_2}, \frac{N_2(1-\sigma_2)}{1-\sigma_1\sigma_2}\right), P_4(0, 0)$$

因为仅当平衡点位于平面坐标系的第一象限时($x_1, x_2 \geq 0$)才有实际意义, 所以对 P_3 而言, 要求 σ_1, σ_2 同时小于1, 或同时大于1.

按照判断平衡点稳定性的方法(见7.7节(18), (19)式), 计算

$$A = \begin{bmatrix} f_{x_1} & f_{x_2} \\ g_{x_1} & g_{x_2} \end{bmatrix} = \begin{bmatrix} r_1 \left(1 - \frac{2x_1}{N_1} - \frac{\sigma_1 x_2}{N_2} \right) & -\frac{r_1 \sigma_1 x_1}{N_2} \\ -\frac{r_2 \sigma_2 x_2}{N_1} & r_2 \left(1 - \frac{\sigma_2 x_1}{N_1} - \frac{2x_2}{N_2} \right) \end{bmatrix}$$

$$p = -(f_{x_1} + g_{x_2})|_{p_i}, \quad i = 1, 2, 3, 4$$

$$q = \det A|_{p_i}, \quad i = 1, 2, 3, 4$$

将4个平衡点 p, q 的结果及稳定条件列入表1.

^① 事实上也无法求出 $x_1(t), x_2(t)$ 的解析表达式.

表 1 种群竞争模型的平衡点及稳定性

平衡点	p	q	稳定条件
$P_1(N_1, 0)$	$r_1 - r_2(1 - \sigma_2)$	$-r_1 r_2(1 - \sigma_2)$	$\sigma_1 < 1, \sigma_2 > 1$
$P_2(0, N_2)$	$-r_1(1 - \sigma_1) + r_2$	$-r_1 r_2(1 - \sigma_1)$	$\sigma_1 > 1, \sigma_2 < 1$
$P_3\left(\frac{N_1(1-\sigma_1)}{1-\sigma_1\sigma_2}, \frac{N_2(1-\sigma_2)}{1-\sigma_1\sigma_2}\right)$	$\frac{r_1(1-\sigma_1) + r_2(1-\sigma_2)}{1-\sigma_1\sigma_2}$	$\frac{r_1 r_2(1-\sigma_1)(1-\sigma_2)}{1-\sigma_1\sigma_2}$	$\sigma_1 < 1, \sigma_2 < 1$
$P_4(0, 0)$	$-(r_1 + r_2)$	$r_1 r_2$	不稳定

注意:按照 7.7 节 (15) 式给出的 $p > 0, q > 0$ 得到的 P_1 的稳定条件只有 $\sigma_2 > 1$, 表 1 中的 $\sigma_1 < 1$ 是根据以下用相轨线分析的结果添加的. P_2 的稳定条件 $\sigma_2 < 1$ 有类似的情况.

对于由非线性方程 (2), (3) 描述的种群竞争, 人们关心的是平衡点的全局稳定 (即不论初始值如何, 平衡点是稳定的), 这需要在上面得到的局部稳定性的基础上辅之以相轨线分析.

在代数方程组 (5) 中记

$$\varphi(x_1, x_2) = 1 - \frac{x_1}{N_1} - \sigma_1 \frac{x_2}{N_2}$$

$$\psi(x_1, x_2) = 1 - \sigma_2 \frac{x_1}{N_1} - \frac{x_2}{N_2}$$

对于 σ_1, σ_2 的不同取值范围, 直线 $\varphi = 0$ 和 $\psi = 0$ 在相平面上的相对位置不同, 图 1 给出了它们的 4 种情况.

下面分别对这 4 种情况进行分析.

1. $\sigma_1 < 1, \sigma_2 > 1$. 图 1(a) 中, $\varphi = 0$ 和 $\psi = 0$ 两条直线将相平面 ($x_1, x_2 \geq 0$) 划分为 3 个区域:

$$S_1: \dot{x}_1 > 0, \dot{x}_2 > 0 \quad (6)$$

$$S_2: \dot{x}_1 > 0, \dot{x}_2 < 0 \quad (7)$$

$$S_3: \dot{x}_1 < 0, \dot{x}_2 < 0 \quad (8)$$

可以证明, 不论轨线从哪个区域的任一点出发, $t \rightarrow \infty$ 时都将趋向 $P_1(N_1, 0)$.

若轨线从 S_1 的某点出发, 由 (6) 可知, 随着 t 的增加轨线向右上方运动, 必然进入 S_2 ;

若轨线从 S_2 的某点出发, 由 (7) 可知轨线向右下方运动, 那么它或者趋向 P_1 点, 或者进入 S_3 . 但是进入 S_3 是不可能的, 因为如果设轨线在某时刻 t_1 经直线 $\varphi = 0$ 进入 S_3 , 则 $\dot{x}_1(t_1) = 0$, 由方程 (2) 不难算出

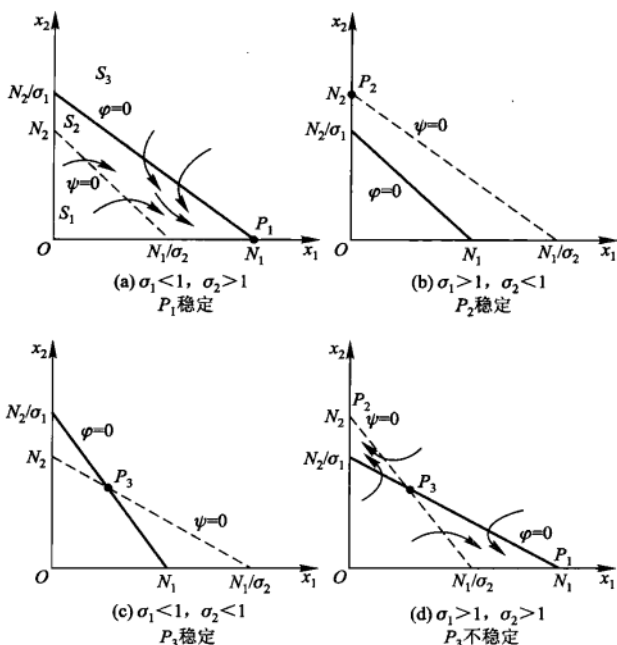


图1 平衡点稳定性的相平面分析

$$\ddot{x}_1(t_1) = -\frac{r_1 \sigma_1}{N_2} x_1(t_1) \dot{x}_2(t_1)$$

由(7),(8)知 $\dot{x}_2(t_1) < 0$, 故 $\ddot{x}_1(t_1) > 0$, 表明 $x_1(t)$ 在 t_1 达到极小值, 而这是不可能的, 因为在 S_2 中 $\dot{x}_1 > 0$, 即 $x_1(t)$ 一直是增加的;

若轨线从 S_3 的某点出发, 由(8)可知轨线向左下方运动, 那么它或者趋向 P_1 点, 或者进入 S_2 . 而进入 S_2 后, 根据上面的分析最终也将趋向 P_1 .

综上所述可以画出轨线示意图(图1(a)). 因为直线 $\varphi = 0$ 上 $dx_1 = 0$, 所以在 $\varphi = 0$ 上轨线方向垂直于 x_1 轴; 在 $\psi = 0$ 上 $dx_2 = 0$, 轨线方向平行于 x_1 轴.

2. $\sigma_1 > 1, \sigma_2 < 1$. 类似的分析可知 $P_2(0, N_2)$ 稳定.

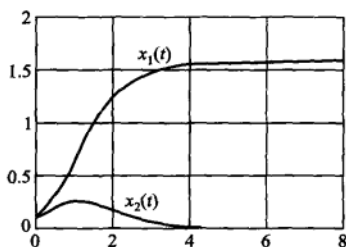
3. $\sigma_1 < 1, \sigma_2 < 1$. $\varphi = 0$ 和 $\psi = 0$ 将相平面分成 4 个区域, 请读者分析, 不论轨线从哪个区域的任一点出发, 都将趋向 P_3 (习题 4).

4. $\sigma_1 > 1, \sigma_2 > 1$. 由表 1 知, 对于 P_3 点 $q < 0$, 故 P_3 不稳定(鞍点). 轨线或者趋向 P_1 , 或者趋向 P_2 , 由轨线的初始位置决定, 示意图见图 1(d). 在这种情况下, P_1 和 P_2 都不能说是全局稳定的(它们只是局部稳定). 正因为这样, 所以 P_1 全局稳定的条件需要加上 $\sigma_1 < 1$. P_2 全局稳定的条件加上 $\sigma_2 < 1$.

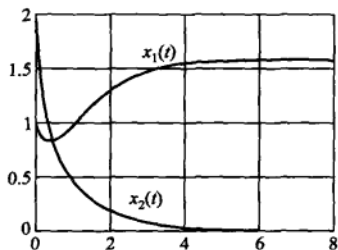
计算与验证 下面用计算机求方程(2),(3)的数值解,画出相轨线的图形,可使我们对两种群的数量变化有进一步的认识,并对以上的分析给以验证.

仅讨论上述第1种情况. 设 $\sigma_1 = 0.5, \sigma_2 = 1.6, r_1 = 2.5, r_2 = 1.8, N_1 = 1.6, N_2 = 1$, 用 MATLAB 软件计算, 为简便起见, 只给出图形结果: 图 2(a), 图 2(b) 是 $x_1(t), x_2(t)$ 的图形 (初值分别取 $(x_1(0) = 0.1, x_2(0) = 0.1)$ 和 $(x_1(0) = 1, x_2(0) = 2)$), 图 2(c) 是两条相轨线的图形.

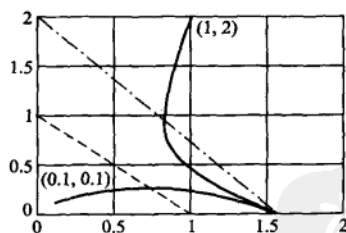
将图 2(c) 与图 1(a) 比较可知二者的一致. 从图 2(a), 图 2(b) 还能看出, 虽然 $t \rightarrow \infty$ 时 $x_1(t) \rightarrow N_1, x_2(t) \rightarrow 0$ 是可预见的, 但是, 若初值较小, $x_2(t)$ 会有一段长时间增长; 若初值较大, $x_1(t)$ 会有一段长时间减少, 却是由计算结果得到的, 你能解释这个现象吗?



(a) $x_1(t), x_2(t)$ 的图形



(b) $x_1(t), x_2(t)$ 的图形



(c) 两条相轨线的图形

图 2

结果解释 根据建模过程中 σ_1, σ_2 的含义, 说明 P_1, P_2, P_3 点稳定在生态学上的意义.

1. $\sigma_1 < 1, \sigma_2 > 1$. $\sigma_1 < 1$ 意味着在对供养甲的资源的竞争中乙弱于甲, $\sigma_2 > 1$ 意味着在对供养乙的资源的竞争中甲强于乙, 于是种群乙终将灭绝, 种群甲趋向最大容量, 即 $x_1(t), x_2(t)$ 趋向平衡点 $P_1(N_1, 0)$.

2. $\sigma_1 > 1, \sigma_2 < 1$. 情况与 1 正好相反.

3. $\sigma_1 < 1, \sigma_2 < 1$. 因为在竞争甲的资源中乙较弱,而在竞争乙的资源中甲较弱,于是可以达到一个双方共存的稳定的平衡状态 P_3 . 这是种群竞争中很少出现的情况.

4. $\sigma_1 > 1, \sigma_2 > 1$. 请读者作出解释.

生态学中有一个竞争排斥原理:若两个种群的单个成员消耗的资源差不多相同,而环境能承受的种群甲的最大容量比种群乙大,那么种群乙终将灭亡. 用本节的模型很容易解释这个原理.

将方程(2),(3)改写为

$$\begin{aligned}\dot{x}_1(t) &= r_1 x_1 \left(1 - \frac{x_1 + \sigma_1 \frac{N_1}{N_2} x_2}{N_1} \right) \\ \dot{x}_2(t) &= r_2 x_2 \left(1 - \frac{\sigma_2 \frac{N_2}{N_1} x_1 + x_2}{N_2} \right)\end{aligned}$$

原理的两个条件相当于

$$\sigma_1 \frac{N_1}{N_2} = 1, \quad \sigma_2 \frac{N_2}{N_1} = 1, \quad N_1 > N_2$$

从这3个式子显然可得 $\sigma_1 < 1, \sigma_2 > 1$, 这正是 P_1 稳定, 即种群乙灭绝的条件.

7.4 种群的相互依存

自然界中处于同一环境下两个种群相互依存而共生的现象是很普遍的. 植物可以独立生存, 昆虫的授粉作用又可以提高植物的增长率, 而以花粉为食物的昆虫却不能离开植物单独存活. 人类与人工饲养的牲畜之间也有类似的关系. 这种共生现象可以描述如下^[6, 62].

设种群甲可以独立存在, 按 logistic 规律增长, 种群乙为甲提供食物, 有助于甲的增长, 类似于 7.3 节的方程(2), 种群甲的数量演变规律可以写作 (r_1, N_1, N_2 的意义同前)

$$\dot{x}_1(t) = r_1 x_1 \left(1 - \frac{x_1}{N_1} + \sigma_1 \frac{x_2}{N_2} \right) \quad (1)$$

7.3 节中 σ_1 前面的 - 号这里变成 + 号, 表示乙不是消耗甲的资源而是为甲提供食物. σ_1 的含义是, 单位数量乙 (相对于 N_2) 提供的供养甲的食物量为单位数量甲 (相对于 N_1) 消耗的供养甲食物量的 σ_1 倍.

种群乙没有甲的存在会灭亡, 设其死亡率为 r_2 , 则乙单独存在时有

$$\dot{x}_2(t) = -r_2 x_2 \quad (2)$$

甲为乙提供食物,于是(2)式右端应加上甲对乙增长的促进作用,有

$$\dot{x}_2(t) = r_2 x_2 \left(-1 + \sigma_2 \frac{x_1}{N_1} \right) \quad (3)$$

显然仅当 $\frac{\sigma_2 x_1}{N_1} > 1$ 时,种群乙的数量才会增长.与此同时乙的增长又会受到自身的阻滞作用,所以(3)式右端还要添加 logistic 项,方程变为

$$\dot{x}_2(t) = r_2 x_2 \left(-1 + \sigma_2 \frac{x_1}{N_1} - \frac{x_2}{N_2} \right) \quad (4)$$

方程(1)、(4)构成相互依存现象的数学模型.下面利用平衡点的稳定性分析,讨论时间足够长以后两个种群的变化趋向.

类似于 7.3 节的作法,将方程(1)、(4)的平衡点及其稳定性分析的结果列入表 1.

表 1 种群依存模型的平衡点及稳定性

平衡点	p	q	稳定条件
$P_1(N_1, 0)$	$r_1 - r_2(\sigma_2 - 1)$	$-r_1 r_2(\sigma_2 - 1)$	$\sigma_2 < 1, \sigma_1 \sigma_2 < 1$
$P_2\left(\frac{N_1(1-\sigma_1)}{1-\sigma_1\sigma_2}, \frac{N_2(\sigma_2-1)}{1-\sigma_1\sigma_2}\right)$	$\frac{r_1(1-\sigma_1) + r_2(\sigma_2-1)}{1-\sigma_1\sigma_2}$	$\frac{r_1 r_2(1-\sigma_1)(\sigma_2-1)}{1-\sigma_1\sigma_2}$	$\sigma_1 < 1, \sigma_2 > 1, \sigma_1 \sigma_2 < 1$
$P_3(0, 0)$	$-r_1 + r_2$	$-r_1 r_2$	不稳定

显然, P_2 点稳定才表明两个种群在同一环境里相互依存而共生,我们着重分析 P_2 稳定的条件.

由 P_2 点的表达式容易看出,要使平衡点 P_2 有实际意义,即位于相平面第一象限($x_1, x_2 \geq 0$),必须满足下面两个条件中的一个:

$$A_1: \sigma_1 < 1, \quad \sigma_2 > 1, \quad \sigma_1 \sigma_2 < 1$$

$$A_2: \sigma_1 > 1, \quad \sigma_2 < 1, \quad \sigma_1 \sigma_2 > 1$$

而由表 1 中 P_2 点的 p, q 可知,仅在条件 A_1 下 P_2 才是稳定的(而在 A_2 下 P_2 是鞍点,不稳定).图 1 画出了条件 A_1 下相轨线的示意图,其中 $\varphi = 1 - \frac{x_1}{N_1} + \sigma_1 \frac{x_2}{N_2}$,

$\psi = -1 + \sigma_2 \frac{x_1}{N_1} - \frac{x_2}{N_2}$. 直线 $\varphi = 0$ 和 $\psi = 0$ 将相平面($x_1, x_2 \geq 0$)划分成 4 个区域:

$S_1: \dot{x}_1 > 0, \dot{x}_2 < 0; S_2: \dot{x}_1 > 0, \dot{x}_2 > 0; S_3: \dot{x}_1 < 0, \dot{x}_2 > 0; S_4: \dot{x}_1 < 0, \dot{x}_2 < 0$. 从 4 个区域中 \dot{x}_1, \dot{x}_2 的正负不难看出其相轨线的趋向如图 1 所示.

分析条件 A_1 的实际意义,其关键部分是 $\sigma_2 > 1$. 考虑到 σ_2 的含义,这表示种群甲要为乙提供足够的食物维持其生长. 而 $\sigma_1 \sigma_2 < 1$ 则是在 $\sigma_2 > 1$ 条件下,为使 P_2 位于相平面第一象限所必需的,当然这要求 σ_1 很小 ($\sigma_1 < 1$ 是必要条件). 注意到 σ_1 的含义,这实际上是对乙向甲提供食物加以限制,以防止甲的过分增长.

在种群依存模型 (1)、(4) 中,如果平衡点 $P_1(N_1, 0)$ 稳定,那么种群乙灭绝,没有种群的共存. 请读者分析导致 P_1 稳定的条件及在生态学上的意义 (习题 5).

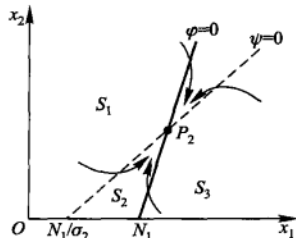


图 1 P_2 稳定的相轨线图

评注 模型 (1)、(4) 是种群相互依存的一种类型,即种群甲可独立生存,而种群乙不能. 依存模型还有其他类型,如两个种群均能独立生存,及均不能独立生存的情况. 这些情况的稳定结果如何,留作习题 6,7.

7.5 食饵 - 捕食者模型

自然界中不同种群之间还存在着一种非常有趣的既有依存、又有制约的生存方式:种群甲靠丰富的自然资源生长,而种群乙靠捕食种群甲为生,食用鱼和鲨鱼、美洲兔和山猫、落叶松和蚜虫等都是这种生存方式的典型. 生态学上称种群甲为食饵 (Prey), 种群乙为捕食者 (Predator), 二者共处组成食饵 - 捕食者系统 (简称 P - P 系统). 近百年来许多数学家和生态学家对这一系统进行了深入的研究,建立了一系列数学模型,本节着重介绍 P - P 系统最初的、最简单的一个模型,它的由来还有一段历史背景.

意大利生物学家 D'Ancona 曾致力于鱼类各种群间相互依存、相互制约关系的研究,从第一次世界大战期间地中海各港口捕获的几种鱼类占总捕获量百分比的资料中,发现鲨鱼 (捕食者) 的比例有明显的增加. 他知道,捕获的各种鱼的比例基本上代表了地中海渔场中各种鱼的比例. 战争中捕获量大幅度下降,应该使渔场中食用鱼 (食饵) 和以此为生的鲨鱼同时增加,但是,捕获量的下降为什么会使鲨鱼的比例增加,即对捕食者更加有利呢? 他无法解释这种现象,于是求助于他的朋友、著名的意大利数学家 Volterra. Volterra 建立了一个简单的数学模型,回答了 D'Ancona 的问题^[15,49,56,60].

Volterra 食饵 - 捕食者模型

食饵 (食用鱼) 和捕食者 (鲨鱼) 在时刻 t 的数量分别记作 $x(t), y(t)$, 因为大海中资源丰富,假设当食饵独立生存时以指数规律增长, (相对) 增长率为 r , 即 $\dot{x} = rx$, 而捕食者的存在使食饵的增长率减小, 设减小的程度与捕食者数量成

正比,于是 $x(t)$ 满足方程

$$\dot{x}(t) = x(r - ay) = rx - axy \quad (1)$$

比例系数 a 反映捕食者掠取食饵的能力.

捕食者离开食饵无法生存,设它独自存在时死亡率为 d ,即 $\dot{y} = -dy$,而食饵的存在为捕食者提供了食物,相当于使捕食者的死亡率降低,且促使其增长.设这种作用与食饵数量成正比,于是 $y(t)$ 满足

$$\dot{y}(t) = y(-d + bx) = -dy + bxy \quad (2)$$

比例系数 b 反映食饵对捕食者的供养能力.

方程(1),(2)是在自然环境中食饵和捕食者之间依存和制约的关系,这里没有考虑种群自身的阻滞增长作用,是 Volterra 提出的最简单的模型.

模型分析 方程(1),(2)没有解析解,我们分两步对这个模型所描述的现象进行分析.首先,利用数学软件求微分方程的数值解,通过对数值结果和图形的观察,猜测它的解析解的构造;然后,从理论上研究其平衡点及相轨线的形状,验证前面的猜测.

1. 数值解

记食饵和捕食者的初始数量分别为

$$x(0) = x_0, y(0) = y_0 \quad (3)$$

为求微分方程(1),(2)及初始条件(3)的数值解 $x(t), y(t)$ (并作图)及相轨线 $y(x)$, 设 $r=1, d=0.5, a=0.1, b=0.02, x_0=25, y_0=2$, 用 MATLAB 软件编制程序如下:

```
function xdot = shier(t,x)
r=1;d=0.5;a=0.1;b=0.02;
xdot=[(r-a*x(2)).*x(1);(-d+b*x(1)).*x(2)];

ts=0:0.1:15;
x0=[25,2];
[t,x]=ode45('shier',ts,x0);[t,x],
plot(t,x),grid,gtext('x(t)'),gtext('y(t)'),
pause,
plot(x(:,1),x(:,2)),grid,
```

可得 $x(t), y(t)$ 及相轨线 $y(x)$ 如图 1、图 2 (数值结果从略). 可以猜测, $x(t), y(t)$ 是周期函数, 与此相应地, 相轨线 $y(x)$ 是封闭曲线. 从数值解近似地定出周期为 10.7, x 的最大、最小值分别为 99.3 和 2.0, y 的最大、最小值分别为 28.4 和 2.0, 并且用数值积分容易算出 $x(t), y(t)$ 在一个周期的平均值为 $\bar{x} = 25, \bar{y} = 10$.

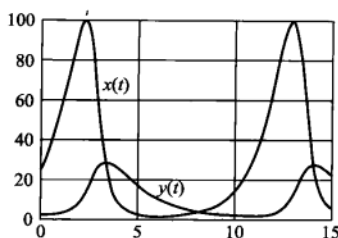


图1 数值解 $x(t), y(t)$ 的图形

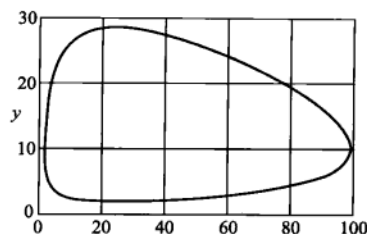


图2 相轨线 $y(x)$ 的图形

2. 平衡点及相轨线

首先求得方程(1), (2)的两个平衡点为

$$P\left(\frac{d}{b}, \frac{r}{a}\right), P'(0, 0) \quad (4)$$

计算它们的 p, q 发现, 对于 $P', q < 0, P'$ 不稳定; 对于 $P, p = 0, q > 0$, 处于临界状态, 不能用判断线性方程平衡点稳定性的准则研究非线性方程(1), (2)的平衡点 P 的情况. 下面用分析相轨线的方法解决这个问题.

从方程(1), (2)消去 dt 后得到

$$\frac{dx}{dy} = \frac{x(r - ay)}{y(-d + bx)} \quad (5)$$

这是可分离变量方程, 写作

$$\frac{-d + bx}{x} dx = \frac{r - ay}{y} dy \quad (6)$$

两边积分得到方程(5)的解, 即方程(1), (2)的相轨线为

$$(x^d e^{-bx})(y^r e^{-ay}) = c \quad (7)$$

其中常数 c 由初始条件确定.

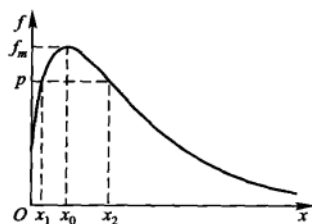
为了从理论上证明相轨线(7)是封闭曲线, 记

$$f(x) = x^d e^{-bx}, \quad g(y) = y^r e^{-ay} \quad (8)$$

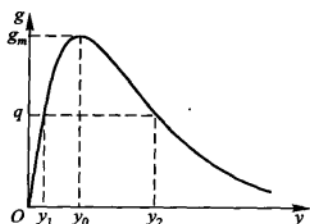
可以用软件作出它们的图形(图3(a), (b)), 将它们的极值点记为 x_0, y_0 , 极大值记为 f_m, g_m , 则不难知道 x_0, y_0 满足

$$\begin{aligned} f(x_0) &= f_m, & x_0 &= d/b \\ g(y_0) &= g_m, & y_0 &= r/a \end{aligned} \quad (9)$$

与(4)相比可知, x_0, y_0 恰好是平衡点 P .



(a) $f(x)$ 的图形



(b) $g(y)$ 的图形

图 3

下面对于给定的 c 值考察相轨线(7)的形状.

当 $c = f_m g_m$ 时, $x = x_0, y = y_0$, 相轨线在图 4 中退化为平衡点 P .

为了考察 $0 < c < f_m g_m$ 时相轨线的形状, 不妨设 $c = p g_m$ ($0 < p < f_m$). 若令 $y = y_0$, 则由 (8), (9) 可得 $f(x) = p$, 而从图 3(a) 知道, 必存在 x_1, x_2 , 使 $f(x_1) = f(x_2) = p$, 且 $x_1 < x_0 < x_2$. 于是相轨线应通过图 4 中的 $Q_1(x_1, y_0)$, $Q_2(x_2, y_0)$ 两点.

接着分析区间 (x_1, x_2) 内的任一点 x , 因为 $f(x) > p$, 由 $f(x)g(y) = p g_m$ 可知, $g(y) < g_m$. 记 $g(y) = q$, 从图 3(b) 知道, 存在 y_1, y_2 , 使 $g(y_1) = g(y_2) = q$, 且 $y_1 < y_0 < y_2$, 于是这条轨线又通过图 4 中的 $Q_3(x, y_1)$, $Q_4(x, y_2)$ 两点. 而因为 x 是区间 (x_1, x_2) 内的任意点, 所以轨线在 Q_1, Q_2 之间对于每个 x 总要通过纵坐标为 y_1, y_2 (且 $y_1 < y_0 < y_2$) 的两点, 这就证明了图 4 中的相轨线是一条封闭曲线.

这样, 当 c 由最大值 $f_m g_m$ 变小时, 相轨线是一族从 P 点向外扩展的封闭曲线 (图 5). P 点称为中心.

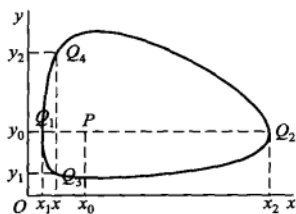


图 4 相轨线的图形

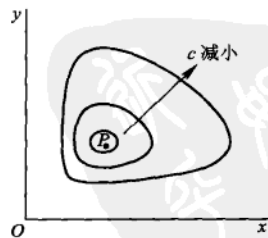


图 5 相轨线族

为确定相轨线的方向, 考察相平面上被 $x = x_0, y = y_0$ 两条直线分成的 4 个区域内 \dot{x}, \dot{y} 的正负号, 由此就决定了相轨线是逆时针方向运动的, 如图 6.

相轨线是封闭曲线等价于 $x(t), y(t)$ 是周期函数 (图 7), 记周期为 T . 在图 7

中周期 T 分为 4 段: $T_1 \sim T_4$, 它们恰好与图 6 中 4 个区域内的 4 段轨线相对应. 结合图 6、图 7 可以看出, $x(t), y(t)$ 的周期变化存在着相位差, $x(t)$ 领先于 $y(t)$, 如 $x(t)$ 领先 T_2 达到最大值.

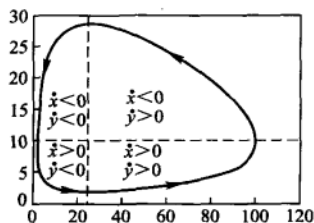


图 6 相轨线的方向

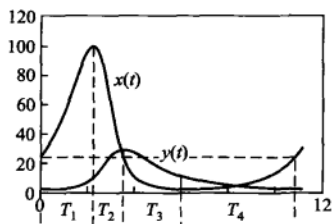


图 7 $x(t), y(t)$ 的相位差

3. $x(t), y(t)$ 在一个周期内的平均值

在数值解中我们看到, $x(t), y(t)$ 一个周期的平均值为 $\bar{x} = 25, \bar{y} = 10$, 这个数值与平衡点 $x_0 = d/b = 0.5/0.02, y_0 = r/a = 1/0.1$ 刚好相等. 实际上, 可以用解析的办法求出它们在一个周期的平均值 \bar{x}, \bar{y} .

将方程(2)改写作

$$\dot{x}(t) = \frac{1}{b} \left(\frac{\dot{y}}{y} + d \right) \quad (10)$$

(10)式两边在一个周期内积分, 注意到 $y(T) = y(0)$, 容易算出平均值为

$$\bar{x} = \frac{1}{T} \int_0^T x(t) dt = \frac{1}{T} \left[\frac{\ln y(T) - \ln y(0)}{b} + \frac{dT}{b} \right] = \frac{d}{b} \quad (11)$$

类似地可得

$$\bar{y} = \frac{r}{a} \quad (12)$$

将(11), (12)与(9)比较可知

$$\bar{x} = x_0, \quad \bar{y} = y_0 \quad (13)$$

即 $x(t), y(t)$ 的平均值正是相轨线中心 P 点的坐标.

模型解释

注意到 r, d, a, b 在生态学上的意义, 上述结果表明, 捕食者的数量(用一个周期的平均值 \bar{y} 代表)与食饵增长率 r 成正比, 与它掠取食饵的能力 a 成反比; 食饵的数量(用一个周期的平均值 \bar{x} 代表)与捕食者死亡率 d 成正比, 与它供养捕食者的能力 b 成反比. 这就是说: 在弱肉强食情况下降低食饵的繁殖率, 可使捕食者减少, 降低捕食者的掠取能力却会使之增加; 捕食者的死亡率上升导致食饵增加, 食饵供养捕食者的能力增强会使食饵减少.

Volterra 用这个模型这样来解释生物学家 D'Ancona 提出的问题: 战争期间

捕获量下降为什么会使鲨鱼(捕食者)的比例有明显的增加.

上面的结果是在自然环境下得到的,为了考虑人为捕获的影响,可以引入表示捕获能力的系数 e ,相当于食饵增长率由 r 下降为 $r-e$,而捕食者死亡率由 d 上升为 $d+e$,用 \bar{x}_1, \bar{y}_1 表示这种情况下食用鱼(食饵)和鲨鱼(捕食者)的(平均)数量,由(11),(12)式可知

$$\bar{x}_1 = \frac{d+e}{b}, \quad \bar{y}_1 = \frac{r-e}{a} \quad (14)$$

显然, $\bar{x}_1 > \bar{x}, \bar{y}_1 < \bar{y}$.

战争期间捕获量下降,即捕获系数为 $e' (< e)$,于是食用鱼和鲨鱼的数量变为

$$\bar{x}_2 = \frac{d+e'}{b}, \quad \bar{y}_2 = \frac{r-e'}{a} \quad (15)$$

显然, $\bar{x}_2 < \bar{x}_1, \bar{y}_2 > \bar{y}_1$. 这正说明战争期间鲨鱼的比例会有明显的增加.

杀虫剂的影响

用 Volterra 模型还可以对某些杀虫剂的影响作出似乎出人意料 of 的解释. 自然界中不少吃农作物的害虫都有它的天敌——益虫,以害虫为食饵的益虫是捕食者,于是构成了一个食饵-捕食者系统. 如果一种杀虫剂既杀死害虫又杀死益虫,那么使用这种杀虫剂就相当于前面讨论的人为捕获的影响,即有 $\bar{x}_1 > \bar{x}, \bar{y}_1 < \bar{y}$. 这说明,从长期效果看(即平均意义下),用这种杀虫剂将使害虫增加,而益虫减少,与使用者的愿望正好相反.

Volterra 模型的局限性

尽管 Volterra 模型可以解释一些现象,但是它作为近似反映现实对象的一个数学模型,必然存在不少局限性.

第一,许多生态学家指出,多数食饵-捕食者系统都观察不到 Volterra 模型显示的那种周期震荡,而是趋向某种平衡状态,即系统存在稳定平衡点. 实际上,只要在 Volterra 模型中加入考虑自身阻滞作用的 logistic 项(为了与 7.3, 7.4 节的模型对比,将这个模型写作)

$$\dot{x}_1(t) = r_1 x_1 \left(1 - \frac{x_1}{N_1} - \sigma_1 \frac{x_2}{N_2} \right) \quad (16)$$

$$\dot{x}_2(t) = r_2 x_2 \left(-1 + \sigma_2 \frac{x_1}{N_1} - \frac{x_2}{N_2} \right) \quad (17)$$

就可以描述这种现象(习题9).

第二,一些生态学家认为,自然界里长期存在的呈周期变化的生态平衡系统应该是结构稳定的,即系统受到不可避免的干扰而偏离原来的周期轨道后,其内部制约作用会使系统自动恢复原状,如恢复原有的周期和振幅. 而 Volterra 模型描述的周期变化状态却不是结构稳定的,因为根据图5,一旦离开某一条闭轨

线,就进入另一条闭轨线(其周期和振幅都会改变),不可能恢复原状.为了得到能反映周期变化的结构稳定的模型,要用到极限环的概念,这超出了本书的范畴.

评注 用数学模型描述、分析食饵-捕食者系统的动态过程和稳定状态,不仅对生态学的研究有重要意义,而且因为它与微分方程定性理论有着密切联系,所以也引起了许多数学家的关注,对此有兴趣的读者可以参看进一步的文献.

7.3~7.5 三节用三个模型研究了两个种群在同一自然环境下生存时相互关系的三种形式.将上面的(16),(17)式与前两节的模型比较,可以看出:这些模型的结构相似,只是相互作用项和自身规律项(繁殖或死亡)的正负号不同.从统一的角度研究两个种群相互关系的模型可参看文献[6,75].

7.6 差分形式的阻滞增长模型

前面几章中我们曾不止一次地用微分方程

$$\dot{x}(t) = rx\left(1 - \frac{x}{N}\right) \quad (1)$$

描述受到环境约束的所谓“阻滞增长”的规律,即 logistic 规律,这种约束随着对象本身数量 x 的增加而增加.人口或其他生物在有限资源环境下的增长、传染病在封闭地区的传播、耐用消费品在有限市场上的销售等现象,都可以合理地、简化地用这个模型描述.

现实对象有时用离散化的时间研究起来比较方便,例如,有些生物每年在固定的时间繁殖,我们用繁殖周期作为时段来研究其增长规律就比用连续时间方便,于是需要阻滞增长的离散模型.将方程(1)中的微分用差分形式表示,就有

$$y_{k+1} - y_k = ry_k\left(1 - \frac{y_k}{N}\right), k=0,1,2,\dots \quad (2)$$

这里用 y_k 而不用 x_k 是为了下面记号的方便. r 和 N 的含义同前,仍分别是固有增长率和最大容量.

(2)式可进一步写作

$$y_{k+1} = (r+1)y_k\left[1 - \frac{r}{(r+1)N}y_k\right] \quad (3)$$

令

$$b = r+1 \quad (4)$$

$$x_k = \frac{r}{(r+1)N}y_k \quad (5)$$

则(3)式可化简为

$$x_{k+1} = bx_k(1 - x_k), k=0,1,2,\dots \quad (6)$$

(6)式是一阶非线性差分方程.在实际应用中没有必要找出方程(6)的一般解,因为给定初值 x_0 后,利用计算机可以方便地由(6)递推地算出 $x_k, k=1, 2, \dots$.

事实上,在应用差分形式的阻滞增长模型(2)或(6)时,人们最关心的通常是 $k \rightarrow \infty$ 时 y_k 或 x_k 的收敛情况,即方程平衡点的稳定性问题.本节主要讨论这个问题^[6, 62].

我们知道,对于微分方程(1), $x^* = N$ 是稳定平衡点, $x^* = 0$ 是不稳定平衡点,即不论 $r(>0)$ 和 $N(>0)$ 为何值,当 $t \rightarrow \infty$ 时都有 $x(t) \rightarrow N$.那么对于方程(1)的差分形式(2),是否也有同样的性质,即 $k \rightarrow \infty$ 时 $y_k \rightarrow N$ 呢?下面将会看到,回答这个问题并不简单,而且将引出一个十分有趣的现象.

平衡点及稳定性 代替(2)式,我们讨论方程(6)的平衡点及其稳定性(非线性差分方程平衡点稳定性判据见7.7节(32)式).因为 $r > 0$,由(4)知 $b > 1$.为了求(6)的平衡点,解代数方程

$$x = f(x) = bx(1-x) \quad (7)$$

容易得到(6)的非零平衡点为^①

$$x^* = 1 - \frac{1}{b} \quad (8)$$

利用(4),(5)式可以验证, x^* 相当于原方程(2)的非零平衡点 $y^* = N$.为分析 x^* 的稳定性,计算

$$f'(x^*) = b(1-2x^*) = 2-b \quad (9)$$

根据 x^* 稳定的条件 $|f'(x^*)| < 1$,立即得到

$$1 < b < 3 \quad (10)$$

由此可知,仅当(10)成立时 x^* 才是稳定平衡点.由(4)式可知,它相当于仅当 $r < 2$ 时 $y^* = N$ 才是方程(2)的稳定平衡点.这与不论 r 多大 $x^* = N$ 都是微分方程(1)的稳定平衡点是不同的.

在条件(10)下, x_k 收敛于 x^* 的状况可以通过方程(6)的图解法清楚地表示出来.以 x 为横坐标作 $y=f(x)=bx(1-x)$ 和 $y=x$ 的图形(图1),曲线 $y=f(x)$ 和直线 $y=x$ 交点的横坐标为平衡点 x^* .对于初值 x_0 ,由方程(6)求 x_1, x_2, \dots 的过程表示为图上带箭头的折线.当 $1 < b < 2$ 时 $x^* < \frac{1}{2}$, $x_k \rightarrow x^*$ 的过程基本上是单调的(见图1(a));而当 $2 < b < 3$ 时 $x^* > \frac{1}{2}$, $x_k \rightarrow x^*$ 的过程则会出现形如6.4节蛛网模型图1那样的衰减振荡(图1(b)).

① $x=0$ 也是方程(6)的一个平衡点,但是容易验证它是不稳定的,以后不再讨论它.

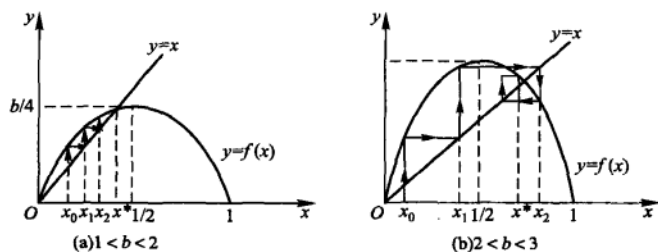


图1 方程(6)的图解法($x_k \rightarrow x^*$)

当 $b > 3$ 时,虽然方程(6)仍可形式地求解,但 x^* 不稳定,其图解法如图2所示,出现形如6.4节蛛网模型图2那样的发散振荡($x_k \rightarrow x^*$).

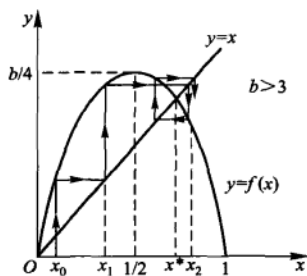


图2 方程(6)的图解法($x_k \rightarrow x^*$)

事情到此并未完结,让我们对不同的 b 值,用方程(6)作一些计算,观察 x_k 的变化趋势.

数值计算 b 由小到大取不同的数值,用方程(6)作计算(初值均取 $x_0 = 0.2$),结果 $x_k (k = 1, 2, \dots, 100)$ 见表1.

表1 不同 b 值下方程(6)计算结果 x_k

k	$b = 1.7$	$b = 2.6$	$b = 3.3$	$b = 3.45$	$b = 3.55$	$b = 3.57$
0	0.200 0	0.200 0	0.200 0	0.200 0	0.200 0	0.200 0
1	0.272 0	0.416 0	0.528 0	0.552 0	0.568 0	0.571 2
2	0.336 6	0.631 7	0.822 4	0.853 2	0.871 1	0.874 4
3	0.379 6	0.604 9	0.482 0	0.432 2	0.398 7	0.392 1
4	0.400 4	0.621 4	0.823 9	0.846 6	0.851 0	0.850 9
5	0.408 1	0.611 7	0.478 7	0.448 0	0.450 0	0.452 9

续表

k	$b = 1.7$	$b = 2.6$	$b = 3.3$	$b = 3.45$	$b = 3.55$	$b = 3.57$
6	0.410 7	0.617 6	0.823 5	0.853 2	0.878 6	0.884 6
7	0.411 4	0.614 1	0.479 6	0.432 2	0.378 5	0.364 5
8	0.411 7	0.616 2	0.823 6	0.846 6	0.835 1	0.827 0
9	0.411 7	0.614 9	0.479 4	0.447 9	0.488 8	0.510 9
10	0.411 8	0.615 7	0.823 6	0.853 1	0.887 1	0.892 1
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
81	0.411 8	0.615 4	0.479 4	0.447 5	0.506 0	0.475 4
82	0.411 8	0.615 4	0.823 6	0.853 0	0.887 4	0.890 3
83	0.411 8	0.615 4	0.479 4	0.432 6	0.354 8	0.348 6
84	0.411 8	0.615 4	0.823 6	0.846 8	0.812 7	0.810 6
85	0.411 8	0.615 4	0.479 4	0.447 5	0.540 5	0.548 0
86	0.411 8	0.615 4	0.823 6	0.853 0	0.881 7	0.884 3
87	0.411 8	0.615 4	0.479 4	0.432 7	0.370 3	0.365 4
88	0.411 8	0.615 4	0.823 6	0.846 9	0.827 8	0.827 8
89	0.411 8	0.615 4	0.479 4	0.447 4	0.506 0	0.508 9
90	0.411 8	0.615 4	0.823 6	0.853 0	0.887 4	0.892 2
91	0.411 8	0.615 4	0.479 4	0.432 7	0.354 8	0.343 3
92	0.411 8	0.615 4	0.823 6	0.846 9	0.812 7	0.804 9
93	0.411 8	0.615 4	0.479 4	0.447 4	0.540 5	0.560 7
94	0.411 8	0.615 4	0.823 6	0.853 0	0.881 7	0.879 3
95	0.411 8	0.615 4	0.479 4	0.432 7	0.370 3	0.378 8
96	0.411 8	0.615 4	0.823 6	0.846 9	0.827 8	0.840 0
97	0.411 8	0.615 4	0.479 4	0.447 4	0.506 0	0.479 7
98	0.411 8	0.615 4	0.823 6	0.853 0	0.887 4	0.891 0
99	0.411 8	0.615 4	0.479 4	0.432 7	0.354 8	0.346 6
100	0.411 8	0.615 4	0.823 6	0.846 9	0.812 7	0.808 5

可以看出,对于 $b = 1.7$ 和 $b = 2.6$, x_k 单调地和振荡地趋向极限 0.4118 和 0.6154 , 与图 1 分析的现象一致, 这两个极限值也与 (8) 式得到的平衡点 x^* 相同. 对于 $b = 3.3$, x_k 好像有两个收敛的子序列, 分别趋向极限值 0.4794 和 0.8236 . 对于 $b = 3.45$ 和 $b = 3.55$, x_k 似乎分别有 4 个和 8 个收敛的子序列. 而对于 $b = 3.57$, x_k 的变化就没有什么规律了. 下面让我们从理论上对 $b > 3$ 的情况作进一步的分析.

倍周期收敛 如果称 $b < 3$ 时 $x_k \rightarrow x^*$ 为单周期收敛 (生物繁殖周期), 那么存在两个收敛的子序列就可以称为 2 倍周期收敛. 一般地, 把方程 (6) 表示为

$$x_{k+1} = f(x_k) \quad (11)$$

在讨论 2 倍周期收敛时, 应考察

$$x_{k+2} = f(x_{k+1}) = f(f(x_k)) = f^{(2)}(x_k) \quad (12)$$

为了求方程 (12) 的平衡点, 对于模型 (6) 要解代数方程

$$x = f(f(x)) = b \cdot bx(1-x)[1-bx(1-x)] \quad (13)$$

因为方程 (12) 的平衡点 x^* 满足 $x^* = f^{(2)}(x^*)$, 所以除了零点和原来的 $x^* = 1 - \frac{1}{b}$ 是它的平衡点外, 满足

$$x_1^* = f(x_2^*), \quad x_2^* = f(x_1^*) \quad (14)$$

的点 x_1^*, x_2^* 也是 (12) 的平衡点, $x_{1,2}^*$ 可由 (13) 解得

$$x_{1,2}^* = \frac{b+1 \pm \sqrt{b^2-2b-3}}{2b} \quad (15)$$

不难验证, 当 $b > 3$ 时

$$0 < x_1^* < x^* < x_2^* < 1 \quad (16)$$

下面在 $b > 3$ 下讨论这些平衡点的稳定性. x^* 显然是不稳定的 (请读者验证). 对于 x_1^* 和 x_2^* , 因为

$$(f^{(2)}(x))'|_{x=x_1^*} = f'(x_2^*)f'(x_1^*), \quad (f^{(2)}(x))'|_{x=x_2^*} = f'(x_1^*)f'(x_2^*)$$

故 x_1^*, x_2^* 的稳定性相同, 再由

$$(f^{(2)}(x))'|_{x=x_1^*, x_2^*} = b^2(1-2x_1^*)(1-2x_2^*) \quad (17)$$

和稳定判据 $|(f^{(2)}(x_1^*, x_2^*))'| < 1$, 并将 (15) 代入 (17), 可得 $x_{1,2}^*$ 的稳定条件为

$$b < 1 + \sqrt{6} \approx 3.449 \quad (18)$$

由上述计算可知, 当 $3 < b < 3.449$ 时, 虽然 x^* 不稳定, 但是 $x_{1,2}^*$ 是方程 (12) 的稳定平衡点, 即 $x_k, x_{k+2}, \dots \rightarrow x_1^*$ (或 x_2^*). 于是对于原方程 (6), $x_{1,2}^*$ 是序列 $\{x_k\}$ 的两个子序列的极限, 即 x_{2k} 和 x_{2k+1} ($k = 1, 2, \dots$) 分别趋向于 x_1^* 或 x_2^* . 以

① 这里上角数字 (2) 表示函数 f 两次作用于自变量, 下同.

$b=3.3$ 代入(15)式,可得 $x_1^*=0.479\ 4$, $x_2^*=0.823\ 6$,与数值计算中的结果相同. 以上的迭代过程也可以从方程(12)的图解法中看到(图3).

作为生物数量阻滞增长的离散模型,以上结果表明,当固有增长率 $2 < r < 2.449$ 时,从一个繁殖周期(即一代)的角度看,其数量增长是不稳定的,即没有极限. 但从两个繁殖周期(即两代)的角度看,增长却是稳定的. 这就是所谓的2倍周期收敛.

读者不难想到,当 $b > 3.449$ 时, $x_{1,2}^*$ 不再是方程(12)的稳定平衡点,从而对于方程(6)来说2倍周期也不收敛了,但是可以讨论4倍周期收敛的情况. 进一步考察方程

$$x_{k+4} = f^{(4)}(x_k) \quad (19)$$

用类似的方法可得,当

$$3.449 < b < 3.544 \quad (20)$$

时,(19)有4个稳定平衡点,数值计算中 $b=3.45$ 就是这种情况. 于是对于原来的模型(6),从4个繁殖周期(4代)的角度看,增长是稳定的.

按照这样的规律,我们可以对模型(6)的增长序列 $\{x_k\}$ 讨论 2^n 倍周期收敛问题, $n=0,1,2,\dots$. 收敛性完全由参数 b 的取值确定. 若记 b_n 为使 2^n 倍周期收敛的 b 的上限,则上面的结果给出: $b_0=3$, $b_1=3.449$, $b_2=3.544$, $b_3=3.564$. 更深入的研究表明,当 $n \rightarrow \infty$ 时, $b_n \rightarrow 3.569$. 而当 $b > 3.569$ 时,就不再存在任何 2^n 倍周期收敛,出现所谓混沌(Chaos)现象,图4给出了模型(6)的收敛、分岔和混沌情况($b=2.5 \sim 4$).

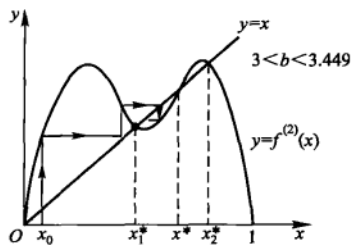


图3 方程(12)的图解法
(图中只画出收敛于 x_1^* 的情况)

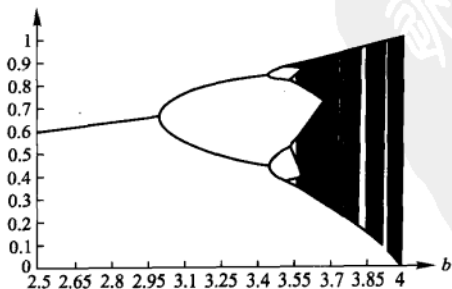


图4 模型(6)的收敛、分岔和混沌

混沌的典型特征之一是对初值的极度敏感:如令 $b=3.7$, 对两个非常靠近的初值 $x_0^{(1)}=0.2$ 和 $x_0^{(2)}=0.200\ 01$ 分别计算, 会发现 $x_{100}^{(1)}=0.481\ 4$, 而 $x_{100}^{(2)}=0.257\ 2$, 可谓差之毫厘, 失之千里, 这就是所谓的蝴蝶效应.

其实, 在混沌区域内也并非乱成一片, 比如, 令 $b=3.83$ 计算一下, 你会发现 x_n 呈 3 倍周期收敛. 在 $3.569\ 9 < b < 4$ 范围内有许多周期为 $p=3, 5, 6, \dots$ 及 $2^np (n=1, 2, \dots)$ 的窗口^[81].

评注 本节虽然没有讨论具体的实际问题, 但是作为描述阻滞增长规律的 logistic 模型, 无论是微分方程形式还是差分形式都是有着广泛应用的, 请读者举出一些应用的例子.

另外, 从数学角度看, 方程(6)是非常简单的非线性差分方程, 可以方便地递推求解. 但是正如我们所看到的, 它的收敛性的研究却引出相当复杂和有趣的现象, 可以作为分岔(Bifurcation)理论和混沌现象的导入. 当然这些问题已经超出本书的范围了.

7.7 微分方程与差分方程稳定性理论简介

本节的介绍仅限于前面几节用到的有关内容.

一阶微分方程的平衡点及稳定性

设有微分方程

$$\dot{x}(t) = f(x) \quad (1)$$

方程右端不显含自变量 t , 称为自治方程. 代数方程

$$f(x) = 0 \quad (2)$$

的实根 $x = x_0$ 称为方程(1)的平衡点(或奇点). 它也是方程(1)的解(奇解).

如果存在某个邻域, 使方程(1)的解 $x(t)$ 从这个邻域内的某个 $x(0)$ 出发, 满足

$$\lim_{t \rightarrow \infty} x(t) = x_0 \quad (3)$$

则称平衡点 x_0 是稳定的(稳定性理论中称渐近稳定); 否则, 称 x_0 是不稳定的(不渐近稳定).

判断平衡点 x_0 是否稳定通常有两种方法. 利用定义即(3)式称间接法. 不求方程(1)的解 $x(t)$, 因而不利用(3)式的方法称直接法. 下面介绍直接法.

将 $f(x)$ 在 x_0 点作 Taylor 展开, 只取一次项, 方程(1)近似为

$$\dot{x}(t) = f'(x_0)(x - x_0) \quad (4)$$

(4)称为(1)的近似线性方程, x_0 也是方程(4)的平衡点. 关于 x_0 点稳定性有如下的结论:

若 $f'(x_0) < 0$, 则 x_0 对于方程(4)和(1)都是稳定的;

若 $f'(x_0) > 0$, 则 x_0 对于方程(4)和(1)都是不稳定的.

x_0 对于方程(4)的稳定性很容易由定义(3)证明, 因为若记 $f'(x_0) = a$, 则(4)的一般解是

$$x(t) = ce^{at} + x_0 \quad (5)$$

其中 c 是由初始条件决定的常数, 显然, 当 $a < 0$ 时(3)式成立.

二阶微分方程的平衡点和稳定性

二阶微分方程可用两个一阶微分方程表为

$$\begin{cases} \dot{x}_1(t) = f(x_1, x_2) \\ \dot{x}_2(t) = g(x_1, x_2) \end{cases} \quad (6)$$

右端不显含 t , 是自治方程. 代数方程组

$$\begin{cases} f(x_1, x_2) = 0 \\ g(x_1, x_2) = 0 \end{cases} \quad (7)$$

的实根 $x_1 = x_1^0, x_2 = x_2^0$ 称为方程(6)的平衡点, 记作 $P_0(x_1^0, x_2^0)$.

如果存在某个邻域, 使方程(6)的解 $x_1(t), x_2(t)$ 从这个邻域内的某个 $(x_1(0), x_2(0))$ 出发, 满足

$$\lim_{t \rightarrow \infty} x_1(t) = x_1^0, \quad \lim_{t \rightarrow \infty} x_2(t) = x_2^0 \quad (8)$$

则称平衡点 P_0 是稳定的(渐近稳定); 否则, 称 P_0 是不稳定的(不渐近稳定).

为了用直接法讨论方程(6)的平衡点的稳定性, 先看线性常数系数方程

$$\begin{cases} \dot{x}_1(t) = a_1 x_1 + a_2 x_2 \\ \dot{x}_2(t) = b_1 x_1 + b_2 x_2 \end{cases} \quad (9)$$

系数矩阵记作

$$A = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \quad (10)$$

为研究方程(9)的惟一平衡点 $P_0(0, 0)$ 的稳定性, 假定 A 的行列式

$$\det A \neq 0 \quad (11)$$

$P_0(0, 0)$ 的稳定性由(9)的特征方程

$$\det(A - \lambda I) = 0 \quad (12)$$

的根 λ (特征根)决定. 方程(12)可以写成更加明晰的形式

$$\begin{cases} \lambda^2 + p\lambda + q = 0 \\ p = -(a_1 + b_2) \\ q = \det A \end{cases} \quad (13)$$

将特征根记作 λ_1, λ_2 , 则

$$\lambda_1, \lambda_2 = \frac{1}{2}(-p \pm \sqrt{p^2 - 4q}) \quad (14)$$

方程(9)的一般解具有形式 $c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$ ($\lambda_1 \neq \lambda_2$) 或 $c_1 e^{\lambda_1 t} + c_2 t e^{\lambda_1 t}$ ($\lambda_1 = \lambda_2$), c_1, c_2 为任意常数. 按照稳定性的定义(8)式可知, 当 λ_1, λ_2 为负数或有负实部时, $P_0(0,0)$ 是稳定平衡点; 而当 λ_1, λ_2 有一个为正数或有正实部时, $P_0(0,0)$ 是不稳定平衡点. 在条件(11)下 λ_1, λ_2 不可能为 0.

微分方程稳定性理论将平衡点分为结点、焦点、鞍点、中心等类型, 完全由特征根 λ_1, λ_2 或相应的 p, q 取值决定. 表 1 简明地给出了这些结果, 表中最后一列指按照定义(8)式得到的关于稳定性的结论.

表 1 由特征方程决定的平衡点的类型和稳定性

λ_1, λ_2	p, q	平衡点类型	稳定性
$\lambda_1 < \lambda_2 < 0$	$p > 0, q > 0, p^2 > 4q$	稳定结点	稳定
$\lambda_1 > \lambda_2 > 0$	$p < 0, q > 0, p^2 > 4q$	不稳定结点	不稳定
$\lambda_1 < 0 < \lambda_2$	$q < 0$	鞍点	不稳定
$\lambda_1 = \lambda_2 < 0$	$p > 0, q > 0, p^2 = 4q$	稳定退化结点	稳定
$\lambda_1 = \lambda_2 > 0$	$p < 0, q > 0, p^2 = 4q$	不稳定退化结点	不稳定
$\lambda_{1,2} = \alpha \pm \beta i, \alpha < 0$	$p > 0, q > 0, p^2 < 4q$	稳定焦点	稳定
$\lambda_{1,2} = \alpha \pm \beta i, \alpha > 0$	$p < 0, q > 0, p^2 < 4q$	不稳定焦点	不稳定
$\lambda_{1,2} = \alpha \pm \beta i, \alpha = 0$	$p = 0, q > 0$	中心	不稳定

由表 1 可以看出, 根据特征方程的系数 p, q 的正负很容易判断平衡点的稳定性, 准则如下: 若

$$p > 0, q > 0 \quad (15)$$

则平衡点稳定; 若

$$p < 0 \text{ 或 } q < 0 \quad (16)$$

则平衡点不稳定.

以上是对线性方程(9)的平衡点 $P_0(0,0)$ 稳定性的结论. 对于一般的非线性方程(6), 可以用近似线性方法判断其平衡点 $P_0(x_1^0, x_2^0)$ 的稳定性. 在 P_0 点将 $f(x_1, x_2)$ 和 $g(x_1, x_2)$ 作 Taylor 展开, 只取一次项, 得(6)的近似线性方程

$$\begin{cases} \dot{x}_1(t) = f_{x_1}(x_1^0, x_2^0)(x_1 - x_1^0) + f_{x_2}(x_1^0, x_2^0)(x_2 - x_2^0) \\ \dot{x}_2(t) = g_{x_1}(x_1^0, x_2^0)(x_1 - x_1^0) + g_{x_2}(x_1^0, x_2^0)(x_2 - x_2^0) \end{cases} \quad (17)$$

系数矩阵记作

$$A = \begin{bmatrix} f_{x_1} & f_{x_2} \\ g_{x_1} & g_{x_2} \end{bmatrix} \bigg|_{P_0(x_1^0, x_2^0)} \quad (18)$$

特征方程系数为

$$p = -(f_{x_1} + g_{x_2}) \bigg|_{P_0}, \quad q = \det A \quad (19)$$

显然, P_0 点对于方程(17)的稳定性由表 1 或准则(15), (16)决定, 而且已经证明了如下结论:

若方程(17)的特征根不为 0 或实部不为 0, 则 P_0 点对于方程(6)的稳定性与对于近似方程(17)的稳定性相同, 即由准则(15), (16)决定.

最后, 提出以下几点值得注意:

1. 平衡点及其稳定性的概念只是对自治方程(1), (6)而言才有意义.

2. 非线性方程(1), (6)的平衡点的稳定性, 与相应的近似线性方程(4), (17)的平衡点的稳定性一致, 是在非临界情况下(即 $a \neq 0$, 或 $p, q \neq 0$)得到的, 在临界情况下(即 $a = 0$ 或 $p, q = 0$)二者可以不一致.

3. 在讨论平衡点稳定性时, 对初始点的要求是存在一个邻域, 这是局部稳定的定义. 如果要求对任意的初始点, (3), (8)式成立, 称为全局稳定. 对于线性方程, 局部稳定与全局稳定是等价的; 对于非线性方程, 二者不同.

4. 对于临界情况和非线性方程的全局稳定, 可以用相轨线分析方法讨论.

差分方程的平衡点及其稳定性.

差分方程的平衡点及其稳定性的概念与微分方程的有关概念是一致的, 例如, 一阶线性常数系数差分方程

$$x_{k+1} + ax_k = b, \quad k = 0, 1, \dots \quad (20)$$

的平衡点由 $x + ax = b$ 解得 $x^* = \frac{b}{1+a}$. 当 $k \rightarrow \infty$ 时, 若 $x_k \rightarrow x^*$, 则 x^* 是稳定的, 否则是不稳定的^①. 容易看出, 可以用变量代换方法将方程(20)的平衡点稳定性问题转换为

$$x_{k+1} + ax_k = 0 \quad (21)$$

的平衡点 $x^* = 0$ 的稳定性问题. 而对于方程(21), 因为其解显然可表为

$$x_k = (-a)^k x_0, \quad k = 1, 2, \dots \quad (22)$$

① 实质上这里是指渐近稳定和不断近稳定.

所以立即可知当且仅当

$$|a| < 1 \quad (23)$$

时, 方程(21)的平衡点(从而方程(20)的平衡点)才是稳定的.

顺便指出, 对于 n 维向量 $x(k)$ 和 $n \times n$ 常数矩阵 A 构成的方程组

$$x(k+1) + Ax(k) = 0 \quad (24)$$

其平衡点稳定的条件是 A 的特征根 $\lambda_i (i=1, 2, \dots, n)$ 均有

$$|\lambda_i| < 1 \quad (25)$$

即均在复平面上的单位圆内. 这个结果可由将 A 化为对角矩阵(或 Jordan 矩阵)得到.

对于二阶线性常系数差分方程, 我们考察

$$x_{k+2} + a_1 x_{k+1} + a_2 x_k = 0 \quad (26)$$

的平衡点 ($x^* = 0$) 的稳定性. 为求 (26) 的通解, 写出它的特征方程 $\lambda^2 + a_1 \lambda + a_2 = 0$, 记这个一元二次代数方程的根为 λ_1, λ_2 , 不难验证, (26) 的通解可表为

$$x_k = c_1 \lambda_1^k + c_2 \lambda_2^k \quad (27)$$

其中常数 c_1, c_2 由初始条件 x_0, x_1 确定. 由 (27) 立即得到, 当且仅当特征根 λ_1, λ_2 满足

$$|\lambda_1| < 1, |\lambda_2| < 1 \quad (28)$$

时, 方程(26)的平衡点才是稳定的.

与一阶线性方程一样, 非齐次方程

$$x_{k+2} + a_1 x_{k+1} + a_2 x_k = b \quad (29)$$

的平衡点的稳定性和方程(26)相同.

二阶方程的上述结果可以推广到 n 阶线性方程, 即稳定平衡点的条件是特征根—— n 次代数方程的根 $\lambda_i (i=1, 2, \dots, n)$ 均有 $|\lambda_i| < 1$. 考虑到高阶方程和方程组的相互转化, 这个条件与 (24), (25) 给出的结论是完全一致的.

最后讨论一阶非线性差分方程

$$x_{k+1} = f(x_k) \quad (30)$$

的平衡点的稳定性. 其平衡点 x^* 由代数方程 $x = f(x)$ 解出. 为分析 x^* 的稳定性, 将方程(30)的右端在 x^* 点作 Taylor 展开, 只取一次项, (30) 近似为

$$x_{k+1} = f'(x^*)(x_k - x^*) + f(x^*) \quad (31)$$

(31) 是 (30) 的近似线性方程, x^* 也是 (31) 的平衡点. 关于线性方程 (31) 的稳定平衡点的讨论已由 (20) ~ (23) 给出, 而当 $|f'(x^*)| \neq 1$ 时, 方程 (30) 与 (31) 平衡点的稳定性相同. 于是得到, 当

$$|f'(x^*)| < 1 \quad (32)$$

时, 对于非线性方程 (30), x^* 是稳定的; 当

$$|f'(x^*)| > 1 \quad (33)$$

时,对于方程(30), x^* 是不稳定的.

习 题

1. 在 7.1 节捕鱼模型中,如果渔场鱼量的自然增长仍服从 logistic 规律,而单位时间捕捞量为常数 h .

(1) 分别就 $h > rN/4$, $h < rN/4$, $h = rN/4$ 这 3 种情况讨论渔场鱼量方程的平衡点及其稳定状况.

(2) 如何获得最大持续产量,其结果与 7.1 节的产量模型有何不同?

2. 与 logistic 模型不同的另一种描述种群增长规律的是 Gompertz 模型: $\dot{x}(t) = rx \ln \frac{N}{x}$, 其中 r 和 N 的意义与 logistic 模型相同.

设渔场鱼量的自然增长服从这个模型,且单位时间捕捞量为 $h = Ex$. 讨论渔场鱼量的平衡点及其稳定性,求最大持续产量 h_m 及获得最大产量的捕捞强度 E_m 和渔场鱼量水平 x_0^* .

3. 在 7.3 节种群竞争模型中设 $\sigma_1 \sigma_2 = 1$ ($\sigma_1 \neq \sigma_2$), 求平衡点并分析其稳定性.

4. 对于 7.3 节种群竞争模型的第 3 种情况: $\sigma_1 < 1, \sigma_2 < 1$ (图 1(c)), 分析相轨线的趋势并画出示意图,解释平衡点 P_3 稳定的意义.

5. 在 7.4 节种群相互依存模型中,按以下 4 种情况作相轨线示意图,并解释平衡点稳定的意义:

(1) $\sigma_2 < 1, \sigma_1 \sigma_2 < 1$.

(2) $\sigma_2 > 1, \sigma_1 \sigma_2 > 1$.

(3) $\sigma_1 > 1, \sigma_2 < 1, \sigma_1 \sigma_2 > 1$.

(4) $\sigma_1 < 1, \sigma_2 > 1, \sigma_1 \sigma_2 < 1$ (相轨线已由图 1 给出,可作为参考).

6. 与 7.4 节的模型稍有不同,如果两个种群都能独立生存,共处时又能相互提供食物,试建立种群依存模型并讨论平衡点的稳定性,解释稳定的意义.

7. 对于第 6 题,如果两个种群都不能独立生存,但共处时可以相互提供食物,试建模以讨论共处的可能性.

8. 在 7.5 节的食饵-捕食者系统中,如果在食饵方程(1)中增加自身阻滞作用的 logistic 项,方程(2)不变,讨论平衡点及稳定性,解释其意义.

9. 如果在 7.5 节的食饵和捕食者方程中都增加 logistic 项,即方程(16), (17), 讨论平衡点及稳定性.

10. 如果食饵-捕食者系统中,捕食者掠食的对象只是成年的食饵,而未成年的食饵因体积太小免遭捕获.在适当的假设下建立这三者之间关系的模型,求平衡点^[15].

11. 验证阻滞增长模型(7.6 节)的(15)~(18)式.

12. 用计算机解阻滞增长模型(7.6 节方程(6)). 比如令 b 从 1.8 逐渐增加,考查序列 x_n 收敛、2 倍周期收敛、4 倍周期收敛、……,直至一片混乱的情况. 试以 b 为横坐标,收敛点为纵坐标作图(与 7.6 节图 4 比较).

* 13. 一个岛屿上栖居着食肉爬行动物和哺乳动物,又长着茂盛的植物.爬行动物以哺乳动物为食物,哺乳动物又依赖植物生存.在适当假设下建立三者之间关系的模型,求平衡点^[15].

* 14. 大陆上物种数目可以看作常数,各物种独立地从大陆向附近一岛屿迁移.岛上物种数量的增加与尚未迁移的物种数目有关,而随着迁移物种的增加又导致岛上物种的减少.在适当假设下建立岛上物种数的模型,并讨论稳定状况^[49].

* 15. 人体注射葡萄糖溶液时,血液中葡萄糖浓度 $g(t)$ 的增长率与注射速率 r 成正比,与人体血液容积 V 成反比,而由于人体组织的吸收作用, $g(t)$ 的减少率与 $g(t)$ 本身成正比.分别在以下几种假设下建立模型,并讨论稳定情况:

(1) 人体血液容积 V 不变.

(2) V 随着注入溶液而增加.

(3) 由于排泄等因素, V 的增加有极限值^[49].

* 16. 讨论资金积累、国民收入与人口增长的关系:

(1) 若国民平均收入 x 与按人口平均资金积累 y 成正比,说明仅当总资金积累的相对增长率 k 大于人口的相对增长率 r 时,国民平均收入才是增长的.

(2) 作出 $k(x)$ 和 $r(x)$ 的示意图,说明二曲线交点是平衡点,讨论它的稳定性.

(3) 分析人口激增会引起什么后果^[10].

* 17. 讨论另一种捕渔业持续收获的效益模型.设渔场鱼量方程仍为 7.1 节(3)式,但捕捞强度为变量 $E(t)$,其变化规律是:当单位时间收入 T 大于支出 S 时(见 7.1 节(9)式) E 增加, T 小于 S 时 E 减少, E 的变化率与 $T-S$ 成正比.

(1) 建立关于 $E(t)$ 的方程,求 $x(t)$, $E(t)$ 的平衡点并讨论其稳定性.

(2) 将所得结果与 7.1 节的效益模型和捕捞过度模型进行比较^[62].

第8章

离散模型

一般地说,确定性离散模型包括的范围很广,除第6章的差分方程模型外,用整数规划、图论、对策论、网络流等数学工具都可以建立离散模型.本章选择了几个在实际中应用较广、涉及的数学知识又不太深的模型.“层次分析模型”和“冲量过程模型”是对社会经济系统进行系统分析的有力工具,“循环比赛的名次”和“公正的选举”讨论了排序问题,“席位分配”是社会政治领域中一个令人关注的问题.

从应用的角度看,这些模型只用到基本的代数、集合及一点点图论的知识.

8.1 层次分析模型

人们在日常生活中常常碰到许多决策问题:买一件衬衫,你要在棉的、丝的、涤纶的……及花的、白的、方格的……之中作出抉择;请朋友吃饭,要筹划是办家宴或去饭店,是吃中餐还是西餐或自助餐;假期旅游,是去风光绮丽的苏杭,还是去迷人的北戴河海滨,或者去山水甲天下的桂林.如果以为这些日常小事不必作为决策问题认真对待的话,那么当你面临报考学校、挑选专业,或者选择工作岗位的时候,就要慎重考虑、反复比较,尽可能地作出满意的决策了.

从事各种职业的人也经常面对决策:一个厂长要决定购买哪种设备,上马什么产品;科技人员要选择研究课题;医生要为疑难病症确定治疗方案;经理要从若干应试者中选拔秘书;各地区各部门的官员则要对人口、交通、经济、环境等领域的发展规划作出决策.

人们在处理上面这些决策问题的时候,要考虑的因素有多有少,有大有小,但是一个共同的特点是它们通常都涉及经济、社会、人文等方面的因素.在作比较、判断、评价、决策时,这些因素的重要性、影响力或者优先程度往往难以量化,人的主观选择(当然要根据客观实际)会起着相当主要的作用,这就给用一般的数学方法解决问题带来本质上的困难.

T. L. Saaty 等人在 20 世纪 70 年代提出了一种能有效地处理这样一类问题的实用方法,称层次分析法 (Analytic Hierarchy Process, 简记 AHP),这是一种定

一、层次分析法的基本步骤

层次分析法的基本思路与人对一个复杂的决策问题的思维、判断过程大体上是一样的。不妨用假期旅游为例,假如有 P_1, P_2, P_3 3 个旅游胜地供你选择,你会根据诸如景色、费用和居住、饮食、旅途条件等一些准则去反复比较那 3 个候选地点。首先,你会确定这些准则在你的心目中各占多大比重,如果你经济宽绰、醉心旅游,自然特别看重景色条件,而平素俭朴或手头拮据的人则会优先考虑费用,中老年旅游者还会对居住、饮食等条件寄以较大关注。其次,你会就每一个准则将 3 个地点进行对比,譬如 P_1 景色最好, P_2 次之; P_2 费用最低, P_3 次之; P_3 居住等条件较好等等。最后,你要将这两个层次的比较判断进行综合,在 P_1, P_2, P_3 中确定哪个作为最佳地点。

上面的思维过程可以加工整理成以下几个步骤:

1. 将决策问题分解为 3 个层次,最上层为目标层,即选择旅游地,最下层为方案层,有 P_1, P_2, P_3 3 个供选择地点,中间层为准则层,有景色、费用、居住、饮食、旅途 5 个准则,各层间的联系用相连的直线表示(图 1 中费用和饮食的连线从略)。

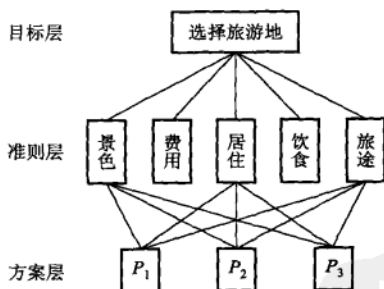


图 1 选择旅游地的层次结构

2. 通过相互比较确定各准则对于目标的权重,及各方案对于每一准则的权重。这些权重在人的思维过程中通常是定性的,而在层次分析法中则要给出得到权重的定量方法。

3. 将方案层对准则层的权重及准则层对目标层的权重进行综合,最终确定方案层对目标层的权重。在层次分析法中要给出进行综合的计算方法。

层次分析法将定性分析与定量计算结合起来完成上述步骤,给出决策结果。下面我们来说明如何比较同一层各因素对上层因素的影响(或在其中的重要性),从而确定它们在上层因素中占的权重。

成对比较矩阵和权向量 涉及到社会、经济、人文等因素的决策问题的主要困难在于, 这些因素通常不易定量地测量. 人们凭自己的经验和知识进行判断, 当因素较多时给出的结果往往是不全面和不准确的, 如果只是定性的结果, 则常常不容易被别人接受. Saaty 等人的作法, 一是不把所有因素放在一起比较, 而是两两相互对比, 二是对比时采用相对尺度, 以尽可能地减少性质不同的诸因素相互比较的困难, 提高准确度.

假设要比较某一层 n 个因素 C_1, C_2, \dots, C_n 对上层一个因素 O 的影响, 如旅游决策问题中比较景色等 5 个准则在选择旅游地这个目标中的重要性. 每次取两个因素 C_i 和 C_j , 用 a_{ij} 表示 C_i 和 C_j 对 O 的影响之比, 全部比较结果可用成对比较矩阵

$$A = (a_{ij})_{n \times n}, \quad a_{ij} > 0, \quad a_{ji} = \frac{1}{a_{ij}} \quad (1)$$

表示. 由于(1)式给出的 a_{ij} 的特点(a_{ji} 是 a_{ij} 的倒数, 即互倒数), A 称为正互反矩阵. 显然必有 $a_{ii} = 1$. 如用 C_1, C_2, \dots, C_5 依次表示景色、费用、居住、饮食、旅途 5 个准则, 设某人用成对比较法(做 $C_5^2 = \frac{5 \times 4}{2}$ 次对比)得到的成对比较阵(正互反阵)为

$$A = \begin{bmatrix} 1 & 1/2 & 4 & 3 & 3 \\ 2 & 1 & 7 & 5 & 5 \\ 1/4 & 1/7 & 1 & 1/2 & 1/3 \\ 1/3 & 1/5 & 2 & 1 & 1 \\ 1/3 & 1/5 & 3 & 1 & 1 \end{bmatrix} \quad (2)$$

(2)中 $a_{12} = 1/2$ 表示景色 C_1 与费用 C_2 对选择旅游地这个目标 O 的重要性之比为 1:2; $a_{13} = 4$ 表示景色 C_1 与居住条件 C_3 之比为 4:1; $a_{23} = 7$ 表示费用 C_2 与居住条件 C_3 之比为 7:1. 可以看出此人在选择旅游地时, 费用因素最重, 景色次之, 居住条件再次. 怎样由成对比较阵确定诸因素 C_1, C_2, \dots, C_n 对上层因素 O 的权重呢?

仔细分析一下(2)式给出的成对比较阵 A 可以发现, 既然 C_1 与 C_2 之比为 1:2, C_1 与 C_3 之比为 4:1, 那么 C_2 与 C_3 之比应为 8:1 而不是 7:1, 才能说明成对比较是一致的. 但是, n 个因素要作 $\frac{n(n-1)}{2}$ 次成对比较, 全部一致的要求是太苛刻了. Saaty 等人给出了在成对比较不一致的情况下计算各因素 C_1, C_2, \dots, C_n 对因素 O 的权重的方法, 并且确定了这种不一致的容许范围, 为了说明这点我们先看成对比较完全一致的情况.

设想把一块单位重量的大石头 O 砸成 n 块小石头 C_1, C_2, \dots, C_n , 如果精确地称出它们的重量为 w_1, \dots, w_n , 在作成对比较时令 $a_{ij} = w_i/w_j$, 那么得到

$$A = \begin{bmatrix} w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\ \vdots & \vdots & & \vdots \\ w_n/w_1 & w_n/w_2 & \cdots & w_n/w_n \end{bmatrix} \quad (3)$$

这些比较显然是一致的。 n 块小石头对大石头的权重(即在大石头中的重量比)可用向量 $w = (w_1, w_2, \dots, w_n)^T$ 表示, 且 $\sum_{i=1}^n w_i = 1$ 。显然, A 的各个列向量与 w 仅相差一个比例因子。

一般地, 如果一个正互反阵 A 满足

$$a_{ij} \cdot a_{jk} = a_{ik}, \quad i, j, k = 1, 2, \dots, n \quad (4)$$

则 A 称为一致性矩阵, 简称一致阵。(3) 式给出的 A 显然是一致阵。容易证明 n 阶一致阵 A 有下列性质(习题 1)。

1. A 的秩为 1, A 的惟一非零特征根为 n ;
2. A 的任一列向量都是对应于特征根 n 的特征向量。

如果得到的成对比较阵是一致阵, 像(3)式的 A , 自然应取对应于特征根 n 的、归一化的特征向量(即分量之和为 1)表示诸因素 C_1, \dots, C_n 对上层因素 O 的权重, 这个向量称为权向量。如果成对比较阵 A 不是一致阵, 但在不一致的容许范围内(下面将说明如何确定这个范围), Saaty 等人建议用对应于 A 最大特征根(记作 λ)的特征向量(归一化后)作为权向量 w , 即 w 满足

$$Aw = \lambda w \quad (5)$$

直观地看, 因为矩阵 A 的特征根和特征向量连续地依赖于矩阵的元素 a_{ij} , 所以当 a_{ij} 离一致性的要求不远时, A 的特征根和特征向量也与一致阵的相差不大。(5)式表示的方法称为由成对比较阵求权向量的特征根法。求 λ 和 w 的简便算法和特征根法更深的意义, 以及其他求权向量的方法见本节第三小节。

比较尺度 当比较两个可能具有不同性质的因素 C_i 和 C_j 对于一个上层因素 O 的影响时, 采用什么样的相对尺度 a_{ij} 较好呢? Saaty 等人提出用 1-9 尺度, 即 a_{ij} 的取值范围是 1, 2, \dots , 9 及其互反数 $1, 1/2, \dots, 1/9$ 。理由如下。

1. 在进行定性的成对比较时, 人们头脑中通常有 5 种明显的等级, 用 1-9 尺度可以方便地表示如表 1。

表 1 1-9 尺度 a_{ij} 的含义

尺度 a_{ij}	含 义
1	C_i 与 C_j 的影响相同
3	C_i 比 C_j 的影响稍强

尺度 a_{ij}	含 义
5	C_i 比 C_j 的影响强
7	C_i 比 C_j 的影响明显的强
9	C_i 比 C_j 的影响绝对的强
2, 4, 6, 8	C_i 与 C_j 的影响之比在上述两个相邻等级之间
1, 1/2, ..., 1/9	C_j 与 C_i 的影响之比为上面 a_{ij} 的互反数

2. 心理学家认为,进行成对比较的因素太多,将超出人的判断能力,最多大致在 7 ± 2 范围. 如以 9 个为限,用 1-9 尺度表示它们之间的差别正合适.

3. Saaty 曾用 $1-3, 1-5, \dots, 1-17, \dots, (d+0.1)-(d+0.9) (d=1, 2, 3, 4), 1^p-9^p (p=2, 3, 4, 5)$ 等共 27 种比较尺度,对在不同距离处判断某光源的亮度等实例构造成对比较阵,并算出权向量. 把这些权向量与按照光强定律等物理知识得到的,或实际测量出的权向量进行对比发现,1-9 尺度不仅在较简单的尺度中最好,而且结果并不劣于较复杂的尺度.

目前在层次分析法的应用中,大多数人都用 1-9 尺度,(2)式中的 A 就是这个尺度. 关于不同尺度的讨论也一直存在着.

一致性检验 成对比较阵通常不是一致阵,但是为了能用它的对应于特征根 λ 的特征向量作为被比较因素的权向量,其不一致程度应在容许范围内. 怎样确定这个范围呢?

前面已经给出 n 阶一致阵的特征根是 n ,在本节第三小节将证明的一个重要定理表明, n 阶正互反阵 A 的最大特征根 $\lambda \geq n$,而当 $\lambda = n$ 时 A 是一致阵.

根据这个定理和 λ 连续地依赖于 a_{ij} 的事实可知, λ 比 n 大得越多, A 的不一致程度越严重,用特征向量作为权向量引起的判断误差越大. 因而可以用 $\lambda - n$ 数值的大小来衡量 A 的不一致程度. Saaty 将

$$CI = \frac{\lambda - n}{n - 1} \quad (6)$$

定义为一致性指标. $CI=0$ 时 A 为一致阵; CI 越大 A 的不一致程度越严重. 注意到 A 的 n 个特征根之和恰好等于 n (为什么?),所以 CI 相当于除 λ 外其余 $n-1$ 个特征根的平均值(取绝对值).

为了确定 A 的不一致程度的容许范围,需要找出衡量 A 的一致性指标 CI 的标准. Saaty 又引入所谓随机一致性指标 RI ,计算 RI 的过程是:对于固定的 n ,随机地构造正互反阵 A' (它的元素 $a'_{ij} (i < j)$ 从 1-9, 1-1/9 中随机取值),然后计算 A' 的一致性指标 CI . 可以想到, A' 是非常不一致的,它的 CI 相当大. 如此构

造相当多的 A' , 用它们的 CI 的平均值作为随机一致性指标. Saaty 对于不同的 n , 用 100 ~ 500 个样本 A' 算出的随机一致性指标 RI 的数值如表 2^①.

表 2 随机一致性指标 RI 的数值

n	1	2	3	4	5	6	7	8	9	10	11
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51

表中 $n=1, 2$ 时 $RI=0$, 是因为 1, 2 阶的正互反阵总是一致阵.

对于 $n \geq 3$ 的成对比较阵 A , 将它的一致性指标 CI 与同阶 (指 n 相同) 的随机一致性指标 RI 之比称为一致性比率 CR , 当

$$CR = \frac{CI}{RI} < 0.1 \quad (7)$$

时认为 A 的不一致程度在容许范围之内, 可用其特征向量作为权向量. (7) 式中 0.1 的选取是带有一定主观信度的.

对于 A 利用 (6), (7) 式和表 2 进行检验称为一致性检验. 当检验不通过时, 要重新进行成对比较, 或对已有的 A 进行修正.

对于 (2) 式给出的 A 可以算出^②, $\lambda = 5.073$, 归一化的特征向量 $w = (0.263, 0.475, 0.055, 0.099, 0.110)^T$. 由 (6) 式 $CI = \frac{5.073 - 5}{5 - 1} = 0.018$, 在表 2 中查出 $RI = 1.12$. 按 (7) 式计算, $CR = \frac{0.018}{1.12} = 0.016 < 0.1$, 一致性检验通过, 上述 w 可作为权向量.

组合权向量 在旅游决策问题中我们已经得到了第 2 层 (准则层) 对第 1 层 (目标层, 只有一个因素) 的权向量, 记作 $w^{(2)} = (w_1^{(2)}, \dots, w_5^{(2)})^T$ (即由 (2) 式的 A 算出的 w). 用同样的方法构造第 3 层 (方案层, 见图 1) 对第 2 层的每一个准则的成对比较阵, 不妨设它们为

$$B_1 = \begin{bmatrix} 1 & 2 & 5 \\ 1/2 & 1 & 2 \\ 1/5 & 1/2 & 1 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 1 & 1/3 & 1/8 \\ 3 & 1 & 1/3 \\ 8 & 3 & 1 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 1 & 3 \\ 1/3 & 1/3 & 1 \end{bmatrix},$$

$$B_4 = \begin{bmatrix} 1 & 3 & 4 \\ 1/3 & 1 & 1 \\ 1/4 & 1 & 1 \end{bmatrix}, \quad B_5 = \begin{bmatrix} 1 & 1 & 1/4 \\ 1 & 1 & 1/4 \\ 4 & 4 & 1 \end{bmatrix}$$

这里矩阵 $B_k (k=1, \dots, 5)$ 中的元素 $b_{ij}^{(k)}$ 是方案 (旅游地) P_i 与 P_j 对于准则 C_k (景

① 由于随机性, 不同的人用不同的样本得到的 RI 的数值会与表 2 稍有出入.

② 计算用的是简便算法——和法 (见本节第三小节).

色、费用等)的优越性的比较尺度。

由第3层的成对比较阵 B_k 计算出权向量 $w_k^{(3)}$, 最大特征根 λ_k 和一致性指标 CI_k , 结果列入表3。

表3 旅游决策问题第3层的计算结果

k	1	2	3	4	5
$w_k^{(3)}$	0.595	0.082	0.429	0.633	0.166
	0.277	0.236	0.429	0.193	0.166
	0.129	0.682	0.142	0.175	0.668
λ_k	3.005	3.002	3	3.009	3
CI_k	0.003	0.001	0	0.005	0

不难看出, 由于 $n=3$ 时随机一致性指标 $RI=0.58$ (表2), 所以上面的 CI_k 均可通过一致性检验。

下面的问题是由各准则对目标的权向量 $w^{(2)}$ 和各方案对每一准则的权向量 $w_k^{(3)}$ ($k=1, \dots, 5$), 计算各方案对目标的权向量, 称为组合权向量, 记作 $w^{(3)}$ 。对于方案 P_1 , 它在景色等5个准则中的权重用 $w_k^{(3)}$ 的第1个分量表示(表3中 $w_k^{(3)}$ 的第1行), 而5个准则对于目标的权重又用权向量 $w^{(2)}$ 表示, 所以方案 P_1 在目标中的组合权重应为它们相应项的两两乘积之和, 即

$$0.595 \times 0.263 + 0.082 \times 0.475 + 0.429 \times 0.055 + \\ 0.633 \times 0.099 + 0.166 \times 0.110 = 0.300$$

同样可以算出 P_2, P_3 在目标中的组合权重为0.246和0.456, 于是组合权向量 $w^{(3)} = (0.300, 0.246, 0.456)^T$ 。结果表明方案 P_3 在旅游地选择中占的权重近于 $1/2$, 远大于 P_1, P_2 , 应作为第1选择地点。

由上述计算可知, 对于3个层次的决策问题, 若第1层只有1个因素, 第2, 3层分别有 n, m 个因素, 记第2, 3层对第1, 2层的权向量分别为

$$w^{(2)} = (w_1^{(2)}, \dots, w_n^{(2)})^T \\ w_k^{(3)} = (w_{k1}^{(3)}, \dots, w_{km}^{(3)})^T, \quad k=1, 2, \dots, n$$

以 $w_k^{(3)}$ 为列向量构成矩阵

$$W^{(3)} = [w_1^{(3)}, \dots, w_n^{(3)}]$$

则第3层对第1层的组合权向量为

$$w^{(3)} = W^{(3)} w^{(2)} \quad (8)$$

更一般地, 若共有 s 层, 则第 k 层对第1层(设只有1个因素)的组合权向量满足

$$w^{(k)} = W^{(k)} w^{(k-1)}, \quad k=3, 4, \dots, s \quad (9)$$

其中 $W^{(k)}$ 是以第 k 层对第 $k-1$ 层的权向量为列向量组成的矩阵。于是最下层

(第 s 层)对最上层的组合权向量为

$$w^{(s)} = W^{(s)} W^{(s-1)} \dots W^{(3)} w^{(2)} \quad (10)$$

组合一致性检验 在应用层次分析法作重大决策时,除了对每个成对比较阵进行一致性检验外,还要进行所谓组合一致性检验,以确定组合权向量是否可以作为最终的决策依据。

组合一致性检验可逐层进行。若第 p 层的一致性指标为 $CI_1^{(p)}, \dots, CI_n^{(p)}$ (n 是第 $p-1$ 层因素的数目),随机一致性指标为 $RI_1^{(p)}, \dots, RI_n^{(p)}$, 定义

$$CI^{(p)} = [CI_1^{(p)}, \dots, CI_n^{(p)}] w^{(p-1)} \quad (11)$$

$$RI^{(p)} = [RI_1^{(p)}, \dots, RI_n^{(p)}] w^{(p-1)} \quad (12)$$

则第 p 层的组合一致性比率为

$$CR^{(p)} = \frac{CI^{(p)}}{RI^{(p)}}, p = 3, 4, \dots, s \quad (13)$$

第 p 层通过组合一致性检验的条件为 $CR^{(p)} < 0.1$ 。

定义最下层(第 s 层)对第 1 层的组合一致性比率为

$$CR^* = \sum_{p=2}^s CR^{(p)} \quad (14)$$

对于重大项目,仅当 CR^* 适当地小时,才认为整个层次的比较判断通过一致性检验。

在旅游决策问题中可以算出 $CI^{(3)} = 0.00176, RI^{(3)} = 0.58, CR^{(3)} = 0.003$ 前面已经有 $CR^{(2)} = 0.016$, 于是 $CR^* = 0.019$, 组合一致性检验通过,前面得到的组合权向量 $w^{(3)}$ 可以作为最终决策的依据。

可将层次分析法的基本步骤归纳如下。

1. 建立层次结构模型 在深入分析实际问题的基础上,将有关的各个因素按照不同属性自上而下地分解成若干层次。同一层的诸因素从属于上一层的因素或对上层因素有影响,同时又支配下一层的因素或受到下层因素的作用,而同一层的各因素之间尽量相互独立。最上层为目标层,通常只有 1 个因素,最下层通常为方案或对象层,中间可以有 1 个或几个层次,通常为准则或指标层。当准则过多时(比如多于 9 个)应进一步分解出子准则层。

2. 构造成对比较阵 从层次结构模型的第 2 层开始,对于从属于(或影响及)上一层每个因素的同一层诸因素,用成对比较法和 1~9 比较尺度构造成对比较阵,直到最下层。

3. 计算权向量并做一致性检验 对于每一个成对比较阵计算最大特征根及对应特征向量(计算方法见本节第三小节),利用一致性指标、随机一致性指标和一致性比率做一致性检验。若检验通过,特征向量(归一化后)即为权向量;若不过,需重新构造成对比较阵。

4. 计算组合权向量并做组合一致性检验 利用(10)式计算最下层对目标的组合权向量,并酌情作组合一致性检验.若检验通过,则可按照组合权向量表示的结果进行决策,否则需重新考虑模型或重新构造那些一致性比率 CR 较大的成对比较阵.

二、层次分析法的广泛应用

层次分析法在 T. L. Saaty 正式提出来之后,由于它在处理复杂的决策问题上的实用性和有效性,很快就在世界范围内得到普遍的重视和广泛的应用.二三十年来它的应用已遍及经济计划和管理、能源政策和分配、行为科学、军事指挥、运输、农业、教育、人才、医疗、环境等领域.从处理问题的类型看,主要是决策、评价、分析、预测等.这个方法在 20 世纪 80 年代初引入我国,也很快为广大的应用数学工作者和有关领域的技术人员所接受,得到了成功的应用.

从上面介绍的层次分析法的基本步骤看,建立层次结构模型是关键的一步,下面给出应用实例时即以这一步为主.构造成对比较阵是整个工作的数量依据,当然是重要的,应当由经验和知识丰富、判断力强的专家给出,还不妨采用群体判断的方式.至于第 3,4 步的计算工作,数学工作者是容易完成的.

例 1 管理信息系统综合评价

当今任何部门每天都会接触到大量的信息,信息管理水平的高低直接关系到工作效率,甚至生存条件.财务、库存、销售、行政……各种各样的管理信息系统(MIS)开发完成或准备推广时,通常要作全面的检查、测试和分析,AHP 是进行综合评价的方法之一.

某一类管理信息系统的综合评价指标体系如下:

1. 系统建设 B_1

- 科学性 C_{11} 规划目标的科学性,经济、技术、管理上的可行性;
- 实现程度 C_{12} 是否达到系统分析阶段提出的目标;
- 先进性 C_{13} 融合了先进的管理科学知识,有较强的适应性;
- 经济性 C_{14} 投资—功能比;
- 资源利用率 C_{15} 对软硬件、信息资源的利用程度;
- 规范性 C_{16} 遵循国际标准、国家标准或行业标准,易于使用、维护和扩充.

2. 系统性能 B_2

- 可靠性 C_{21} 主要是软硬件系统的可靠性;
- 系统效率 C_{22} 系统响应时间、周转时间、吞吐量等;
- 可维护性 C_{23} 确定、修正系统的错误所需的代价;
- 可扩充性 C_{24} 系统结构、硬件设备、软件功能的可扩充程度;

- 可移植性 C_{25} 将系统移植到另一种软硬件环境的代价;
- 安全性 C_{26} 当自然或人为故障造成系统破坏时的有效对策.

3. 系统应用 B_3

- 经济效益 C_{31} 降低成本、增加利润、提高竞争力、改进服务质量等;
- 社会效益 C_{32} 提高科技水平、合理利用资源、增进社会福利、保护生态环境等;
- 用户满意度 C_{33} 人机界面友好、操作方便、容错性强、有帮助功能等;
- 功能应用程度 C_{34} 是否达到预期的技术指标.

用以上各评价指标构造层次结构,形成目标层 A 、准则层 B 、子准则层 C 和方案层 D ,如图 2.

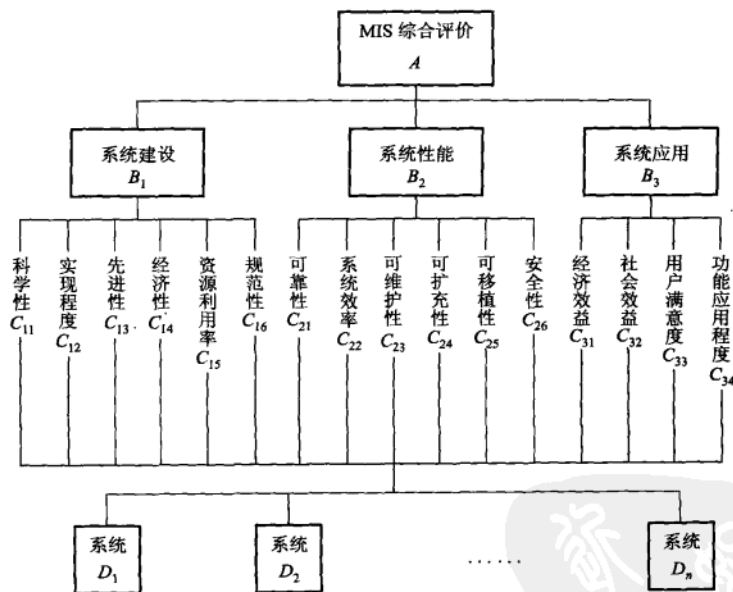


图 2 MIS 综合评价的层次结构

由专家和用户组成的小组对 3 个 MIS 系统 D_1, D_2, D_3 进行综合评价,将成对比较阵略去,得到的权向量及一致性检验的结果如下:

准则层 B 对目标层 A 的权向量 $w^{(2)} = (0.162, 0.309, 0.529)^T$, 一致性指标 $CI^{(2)} = 0.0056$. 子准则层 C 对 B_1, B_2, B_3 的权向量分别为 $w^{(31)} = (0.101, 0.177, 0.177, 0.312, 0.056, 0.177)$, $w^{(32)} = (0.350, 0.126, 0.230, 0.126, 0.043, 0.126)$, $w^{(33)} = (0.336, 0.161, 0.420, 0.082)$, 一致性指标分别为 $CI^{(31)} = 0.0043$, $CI^{(32)} = 0.0048$, $CI^{(33)} = 0.0061$.

方案层 D 对子准则层 C (共 16 个因素) 的权向量 $w_k^{(4)}$ 和一致性指标 $CI_k^{(4)}$ ($k=1, 2, \dots, 16$) 列入表 4, 其中 C 对 A 的权向量 $w^{(3)} = W^{(3)} w^{(2)}$, 而 $W^{(3)}$ 是以 $\tilde{w}^{(31)}, \tilde{w}^{(32)}, \tilde{w}^{(33)}$ 为列向量的 16×3 矩阵 (见 (8) 式), $\tilde{w}^{(31)} = (w^{(31)}, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)^T$, $\tilde{w}^{(32)} = (0, 0, 0, 0, 0, 0, w^{(32)}, 0, 0, 0, 0, 0, 0, 0, 0)^T$, $\tilde{w}^{(33)} = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, w^{(33)}, 0, 0, 0)^T$.

以表 4 中的 16 个权向量 $w_k^{(4)}$ 为列向量构成 3×16 矩阵 $W^{(4)}$, 则方案层 D 对目标层 A 的组合权向量为 $w^{(4)} = W^{(4)} w^{(3)} = (0.315, 0.478, 0.207)^T$.

各层的一致性检验及组合一致性检验全部通过, 上面得到的组合权向量可以作为 3 个 MIS 系统综合评价的依据, 即系统 D_2 最优, D_1 次之.

表 4 MIS 综合评价中方案层 D 对于准则层 C 的计算结果

	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}	C_{16}	C_{21}	C_{22}
$w^{(3)}$	0.016	0.029	0.029	0.051	0.009	0.029	0.108	0.039
$w_k^{(4)}$	0.462	0.344	0.462	0.162	0.535	0.462	0.333	0.462
	0.369	0.535	0.369	0.309	0.344	0.369	0.476	0.369
	0.169	0.121	0.169	0.529	0.121	0.169	0.190	0.169
$CI_k^{(4)}$	0.011 1	0.012 7	0.011 1	0.005 6	0.012 7	0.011 1	0.030 4	0.011 1
	C_{23}	C_{24}	C_{25}	C_{26}	C_{31}	C_{32}	C_{33}	C_{34}
$w^{(3)}$	0.071	0.039	0.013	0.039	0.178	0.085	0.223	0.043
$w_k^{(4)}$	0.109	0.309	0.309	0.109	0.462	0.231	0.274	0.309
	0.570	0.529	0.529	0.570	0.369	0.554	0.632	0.162
	0.321	0.162	0.162	0.321	0.169	0.215	0.095	0.529
$CI_k^{(4)}$	0.002 7	0.005 6	0.005 6	0.002 7	0.011 1	0.010 3	0.013 6	0.005 6

例 2 横渡江河、海峡方案的抉择

1970 年南京长江大桥的建成结束了津浦铁路轮渡长江的历史, 穿越英吉利海峡的隧道为英法两国的交通带来了巨大的方便, 跨越琼州海峡、连接海南岛和雷州半岛的轮渡已经实现, 有人甚至在酝酿横越台湾海峡的海底隧道了. 渡江越海的办法主要有建桥梁、修隧道、轮渡三种, 进行抉择时不外乎要从效益和代价两方面考虑, 这两方面又各有若干准则加以度量, 用 AHP 方法处理应将效益和代价作为两个目标, 分别建立层次结构, 图 3 和图 4 中表述的是某部门对准则的选择, 仅供参考, 因为它们的含义都容易从字面上理解, 这里就不一一解释了.

构造成对比阵和计算权向量的部分从略.

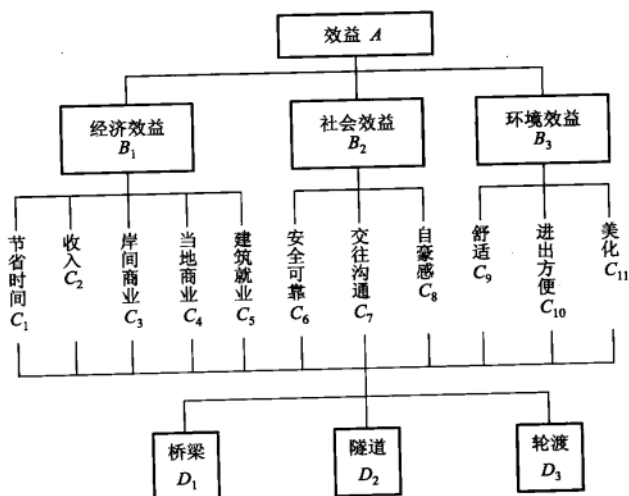


图3 渡江越海评价效益的层次结构

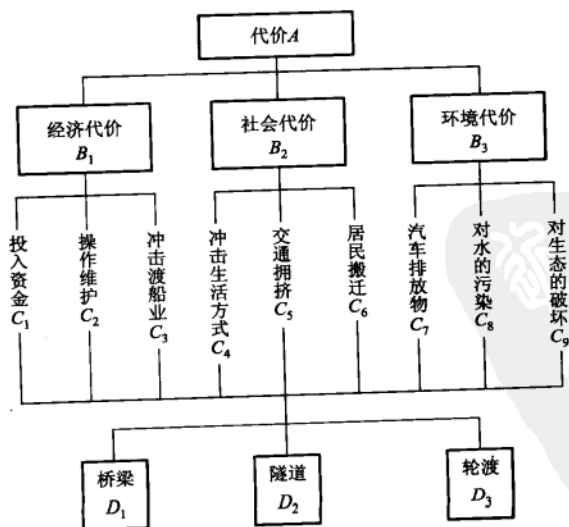


图4 渡江越海评价代价的层次结构

例3 科技成果的综合评价

科技成果涉及的领域很广,种类很多,这里指的是直接应用于国民经济的某个生产部门后,可迅速转化为生产力,带来可定量计算的经济效益的那一类成果.评价准则先分为效益 C_1 、水平 C_2 、规模 C_3 共3类,再在每类中确定若干具体指标,如此构造的层次结构由图5给出.

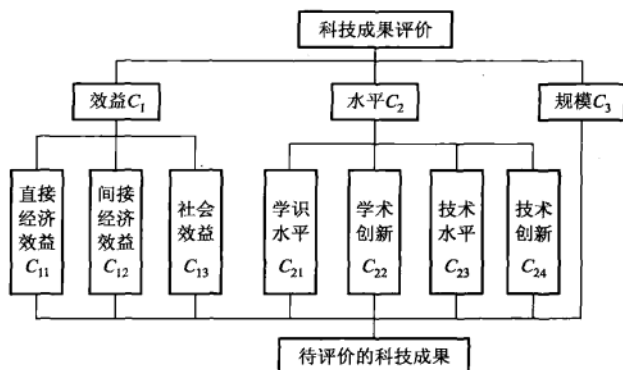


图5 科技成果评价的层次结构

当对科技成果进行相对评价时,可直接利用层次分析法确定出它们对于综合评价的优劣顺序.当对科技成果进行绝对评价时,应先用层次分析法得到 C_{11} , C_{12} ...各项具体指标在综合评价中的相对权重,再给出这些指标的等级标准,如对于 C_{11} ,年经济效益在1 000万元以上为1等(9分);100万元以上为2等(7分);...1万元以下为5等(1分).对于 C_{23} ,达到国际水平为1等(9分);部分达到或全面接近国际水平为2等(7分);国内先进水平为3等(5分);国内水平为4等(3分);一般水平为5等(1分).当某项成果在各指标中的等级被主管部门认定后,将各个分值乘以各指标在综合评价中的权重并求和,即为这项成果的综合绝对评价的分值.

例4 工作选择

一个刚获得学位的大学毕业生面临选择工作岗位,他将要考虑的准则有:能够发挥自己的才干为国家作贡献;丰厚的收入;适合个人的兴趣及发展;良好的声誉;人际关系;地理位置等,于是他可以构造如图6的层次结构,用层次分析法确定可供选择的工作的优先顺序.

你认为这些准则合适吗?试给出准则层对目标的成对比较阵.

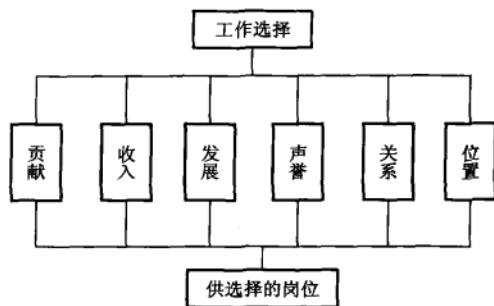


图6 工作选择的层次结构

例5 国家实力分析

一些高层研究人员要对美、俄、中、英、法、日、德等大国的国家综合实力进行分析判断,层次分析法为其提供了一种手段.这里的关键是确定合适的准则及进行实事求是的对比.一个供参考的层次结构如图7所示.

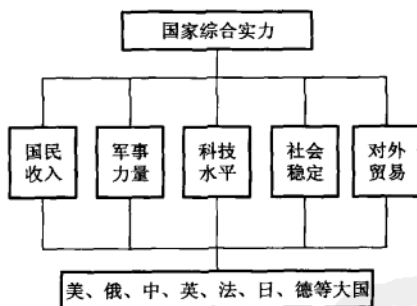


图7 国家综合实力分析的层次结构

通过以上列举的几个实例可以大体上看出层次分析法的应用模式和涉及范围.顺便指出,在这个方法提出和完善的20世纪70年代,Saaty等人曾用它解决过一些国际或国家级的重大课题,如1985年世界石油价格的预测,苏丹运输系统的研究,美国未来高等教育(1985—2000)的规划等^[65,90].

三、层次分析法的若干问题

层次分析法问世以来不仅得到广泛的应用,而且在理论体系、计算方法以及建立更复杂的层次结构等方面都有很快的发展.本节将着重从应用的角度讨论几个问题,对它的公理化体系等方面有兴趣的读者可参看[65,90].

1. 正互反阵最大特征根和对应特征向量的性质

成对比较阵是正互反阵. 层次分析中用对应它的最大特征根的特征向量作为权向量, 用最大特征根定义一致性指标(6)式进行一致性检验. 这里人们首先碰到的问题是: 正互反阵是否存在正的最大特征根和正的特征向量; 一致性指标的大小是否反映它接近一致阵的程度, 特别, 当一致性指标为零时, 它是否就变为一一致阵. 下面两个定理可以回答这些问题.

定理 1 对于正矩阵 A (A 的所有元素为正数),

- 1) A 的最大特征根是正单根 λ ;
- 2) λ 对应正特征向量 w (w 的所有分量为正数);
- 3) $\lim_{k \rightarrow \infty} \frac{A^k \mathbf{1}}{\mathbf{1}^T A^k \mathbf{1}} = w$, 其中 $\mathbf{1} = (1, 1, \dots, 1)^T$, w 是对应 λ 的归一化特征向量.

定理的 1), 2) 是著名的 Perron-Frobenius 定理的一部分, 3) 可通过将 A 化为标准形证明(略).

定理 2 n 阶正互反阵 A 的最大特征根 $\lambda \geq n$; 当 $\lambda = n$ 时 A 是一致阵.

证明见[83]

定理 2 和前面所述的一致阵的性质表明, n 阶正互反阵 A 是一致阵的充要条件为, A 的最大特征根 $\lambda = n$.

上述结论为特征根法用于层次分析提供了一定的理论依据.

2. 正互反阵最大特征根和特征向量的实用算法

众所周知, 用定义计算矩阵的特征根和特征向量是相当困难的, 特别是矩阵阶数较高的时候. 另一方面, 因为成对比较阵是通过定性比较得到的比较粗糙的量化结果, 对它作精确计算是不必要的, 所以完全可以用简便的近似方法计算其特征根和特征向量, 下面介绍几种.

(1) 幂法 步骤如下:

- a. 任取 n 维归一化初始向量 $w^{(0)}$;
- b. 计算 $\tilde{w}^{(k+1)} = A w^{(k)}$, $k = 0, 1, 2, \dots$;
- c. $\tilde{w}^{(k+1)}$ 归一化, 即令 $w^{(k+1)} = \frac{\tilde{w}^{(k+1)}}{\sum_{i=1}^n \tilde{w}_i^{(k+1)}}$;
- d. 对于预先给定的精度 ε , 当 $|w_i^{(k+1)} - w_i^{(k)}| < \varepsilon$ ($i = 1, 2, \dots, n$) 时, $w^{(k+1)}$ 即为所求的特征向量; 否则返回 b;
- e. 计算最大特征根 $\lambda = \frac{1}{n} \sum_{i=1}^n \frac{\tilde{w}_i^{(k+1)}}{w_i^{(k)}}$.

这是求最大特征根对应特征向量的迭代方法, 其收敛性由定理 1 的 3) 保

证. $w^{(0)}$ 可任选或取为下面方法得到的结果.

(2) 和法 步骤如下:

a. 将 A 的每一列向量归一化得 $\tilde{w}_{ij} = \frac{a_{ij}}{\sum_{i=1}^n a_{ij}};$

b. 对 \tilde{w}_{ij} 按行求和得 $\tilde{w}_i = \sum_{j=1}^n \tilde{w}_{ij};$

c. 将 \tilde{w}_i 归一化 $w_i = \frac{\tilde{w}_i}{\sum_{i=1}^n \tilde{w}_i}, w = (w_1, w_2, \dots, w_n)^T$ 即为近似特征向量^①;

d. 计算 $\lambda = \frac{1}{n} \sum_{i=1}^n \frac{(Aw)_i}{w_i}$, 作为最大特征根的近似值.

这个方法实际上是将 A 的列向量归一化后取平均值, 作为 A 的特征向量. 因为当 A 为一致阵时它的每一列向量都是特征向量, 所以若 A 的不一致性不严重, 则取 A 的列向量(归一化后)的平均值作为近似特征向量是合理的.

(3) 根法 步骤与和法基本相同, 只是将步骤 b 改为对 \tilde{w}_{ij} 按行求积并开 n 次方, 即 $\tilde{w}_i = \left(\prod_{j=1}^n \tilde{w}_{ij} \right)^{1/n}.$

根法是将和法中求列向量的算术平均值改为求几何平均值.

以上 3 个方法中以和法最为简便. 试用它计算一个例子:

$$A = \begin{bmatrix} 1 & 2 & 6 \\ 1/2 & 1 & 4 \\ 1/6 & 1/4 & 1 \end{bmatrix} \xrightarrow{\text{列向量归一化}} \begin{bmatrix} 0.6 & 0.615 & 0.545 \\ 0.3 & 0.308 & 0.364 \\ 0.1 & 0.077 & 0.091 \end{bmatrix} \xrightarrow{\text{按行求和}} \begin{bmatrix} 1.760 \\ 0.972 \\ 0.268 \end{bmatrix} \xrightarrow{\text{归一化}} \begin{bmatrix} 0.587 \\ 0.324 \\ 0.089 \end{bmatrix} = w, Aw = \begin{bmatrix} 1.769 \\ 0.974 \\ 0.268 \end{bmatrix}$$

$$\lambda = \frac{1}{3} \left(\frac{1.769}{0.587} + \frac{0.974}{0.324} + \frac{0.268}{0.089} \right) = 3.009$$

精确计算给出 $w = (0.588, 0.322, 0.090)^T, \lambda = 3.010$. 二者相比, 相差甚微.

3. 为什么用成对比较阵的特征向量作为权向量

我们知道, 当成对比较阵 A 是一致阵时, a_{ij} 与权向量 $w = (w_1, \dots, w_n)^T$ 的关系满足 $a_{ij} = \frac{w_i}{w_j}$, 那么当 A 不是一致阵时, 权向量 w 的选择应使得 a_{ij} 与 $\frac{w_i}{w_j}$ 相差

^① 在和法中 $\sum_{i=1}^n \tilde{w}_i = n$, 于是 $w_i = \tilde{w}_i / n$. 但对根法则不具有这种方便性.

(对所有的 i, j) 尽量地小. 这样, 如果从拟合的角度看, 确定 w 可以化为如下的最小二乘问题:

$$\min_{w_i (i=1, \dots, n)} \sum_{i=1}^n \sum_{j=1}^n \left(a_{ij} - \frac{w_i}{w_j} \right)^2 \quad (15)$$

由 (15) 式得到的最小二乘权向量一般与特征根法得到的不同. 因为 (15) 式将导致求解关于 w_i 的非线性方程组, 计算复杂, 且不能保证得到全局最优解, 没有实用价值.

如果改为对数最小二乘问题

$$\min_{w_i (i=1, \dots, n)} \sum_{i=1}^n \sum_{j=1}^n \left(\ln a_{ij} - \ln \frac{w_i}{w_j} \right)^2 \quad (16)$$

则化为求解关于 $\ln w_i$ 的线性方程组. 可以验证, 如此解得的 w_i 恰是前面根法计算的结果 (习题 3).

特征根法解决这个问题的途径可通过对定理 2 的证明看出^[83].

由上可知, 用不同标准确定的权向量是不同的 (当然, 若 A 为一致阵, 则用所有标准确定的权向量应相同). 那么, 相对其他方法而言特征根法有什么优越性呢?

当比较 C_1, C_2, \dots, C_n n 个因素对上层某因素的影响时, a_{ij} 是 C_i 对 C_j (直接比较) 的强度, 不妨称为 1 步强度. 若记 $A^2 = (a_{ij}^{(2)})$, 则不难得到 $a_{ij}^{(2)} = \sum_{s=1}^n a_{is} \cdot a_{sj}$, 即 $a_{ij}^{(2)}$ 是 C_i 通过 $C_s (s=1, 2, \dots, n)$ 对 C_j 比较的强度之和, 称 2 步强度, 它已包含了 1 步强度 a_{ij} (因为和式中包括 $s=i, j$). 显然 $a_{ij}^{(2)}$ 比 a_{ij} 更能反映 C_i 对 C_j 的强度. 类似地, 记 $A^k = (a_{ij}^{(k)})$, $a_{ij}^{(k)}$ 是 k 步强度, 它包含了 1 步至 $k-1$ 步强度. k 越大, $a_{ij}^{(k)}$ 越能全面地反映 C_i 对 C_j 的强度. 可以认为 $a_{ij}^{(k)}$ 体现了相互比较的多步累积效应.

更进一步可以证明, 对于正互反阵 A 和每一对 (i, j) , 存在 k_0 , 当 $k > k_0$ 时, $a_{is}^{(k)} \geq a_{js}^{(k)}$ 或 $a_{is}^{(k)} \leq a_{js}^{(k)}$ 对所有 $s (1 \leq s \leq n)$ 成立. 这表明对于足够大的 k , A^k 的第 i 行元素给出了 C_i 在全部因素中排序权重的信息. 可以用这行元素之和作为 C_i 的权重的度量, 即以 $\frac{A^k \mathbf{1}}{\mathbf{1}^T A^k \mathbf{1}} (\mathbf{1} = (1, 1, \dots, 1)^T)$ 为诸因素的权向量, 其中分母是归一化的需要. 回顾本小节定理 1 的 3). 当 $k \rightarrow \infty$ 时, 这个权向量正是 A 的特征向量 w , 即

$$w = \lim_{k \rightarrow \infty} \frac{A^k \mathbf{1}}{\mathbf{1}^T A^k \mathbf{1}} \quad (17)$$

由 (17) 式用级数理论还不难证明

$$w = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{k=1}^m \frac{A^k \mathbf{1}}{\mathbf{1}^T A^k \mathbf{1}} \quad (18)$$

以上分析表明,无论从全面反映因素间强度对比的多步累积效应的意义上((17)式),还是从各个多步累积效应的平均的意义上((18)式),用特征向量作权向量优于用其他方法得到的权向量。

4. 不完全层次结构中组合权向量的计算

在前两节列举的大多数层次结构模型中,上一层的每个因素都支配着下一层的所有因素,或被下一层所有因素影响,如图1、6、7,这种层次结构称为完全的。但是也有层次结构不是这样,如图2、3、4、5,那里准则层中的一个因素,只支配子准则层的一部分因素,这种层次结构称为不完全的。不过,这类只出现在各准则层中的不完全性容易处理,如第二小节例1中,我们将不支配的那些因素的权向量分量简单地置0,就可以用完全层次结构的办法处理,这显然也是合理的。但是如果不完全结构出现在准则层与方案层之间,事情就有些麻烦,试看下例。

学校要评价教师的贡献,粗略地只考虑教学与科研两个指标,若 P_1, P_2, P_3, P_4 4位教师中 P_1, P_2 只从事教学, P_4 只搞科研, P_3 则二者兼顾,那么层次结构模型如图8。 C_1, C_2 支配因素的数量不等。

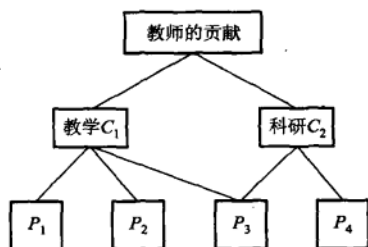


图8 评价教师贡献的层次结构

先看看将不支配因素的权向量分量简单置0有什么后果。设 C_1, C_2 对第1层的权向量 $w^{(2)} = (w_1^{(2)}, w_2^{(2)})^T$ 已经确定, C_1 支配第3层的因素 P_1, P_2, P_3, C_2 支配 P_3, P_4 ,记两个权向量为 $w_1^{(3)} = (w_{11}^{(3)}, w_{12}^{(3)}, w_{13}^{(3)}, 0)^T$ 和 $w_2^{(3)} = (0, 0, w_{23}^{(3)}, w_{24}^{(3)})^T$,按照(8)式应有

$$w^{(3)} = W^{(3)} w^{(2)}, W^{(3)} = (w_1^{(3)}, w_2^{(3)}) \quad (19)$$

考察一个特殊情况:教学与科研两个准则的重要性相同,即 $w^{(2)} = (1/2, 1/2)^T$,4位教师不论从事教学或科研,能力都相同,即 $w_1^{(3)} = (1/3, 1/3, 1/3, 0)^T, w_2^{(3)} = (0, 0, 1/2, 1/2)^T$ 。公正的评价应是,被安排只搞教学或科研的 P_1, P_2, P_4 3人的贡献相同,而 P_3 的贡献为他们的一倍。但是按照(19)式得到的是 $w^{(3)} = (1/6, 1/6, 5/12, 1/4)^T$ 。

怎样才能得到合理的结果呢?一种办法是用支配因素的数量对权向量 $w^{(2)}$ 进行加权,修正为 $\tilde{w}^{(2)}$,再计算 $w^{(3)}$ 。 C_1, C_2 支配因素的数量分别记为 n_1, n_2 ,令

$$\tilde{w}^{(2)} = (n_1 w_1^{(2)}, n_2 w_2^{(2)})^T / (n_1 w_1^{(2)} + n_2 w_2^{(2)}) \quad (20)$$

$$w^{(3)} = W^{(3)} \tilde{w}^{(2)} \quad (21)$$

其中,(20)右端的分母是归一化的需要。

利用上面 $w^{(2)}$, $W^{(3)}$ 的数据, 并注意到 $n_1 = 3, n_2 = 2$, 由 (20), (21) 式可得 $w^{(3)} = (1/5, 1/5, 2/5, 1/5)^T$, 与公正的评价吻合。

从实际考虑, 这种支配因素越多权重越大的修正办法, 只适合于教师从事教学和(或)科研完全由上级安排的情况, 在能力相同的条件下, 承担双份工作的 P_3 的贡献自然大一倍。

但是如果教师从事教学和科研完全靠发挥个人的积极性, 而且上级希望每位教师都二者兼顾, 并鼓励从事人数较少的那一类工作, 就可以用支配因素数量的倒数对 $w^{(2)}$ 加权, (20) 式变为

$$\tilde{w}^{(2)} = \left(\frac{w_1^{(2)}}{n_1}, \frac{w_2^{(2)}}{n_2} \right)^T / \left(\frac{w_1^{(2)}}{n_1} + \frac{w_2^{(2)}}{n_2} \right) \quad (22)$$

不妨用上面的数据按照 (22) 和 (21) 式算一下, 看看这种情况下 4 位教师的贡献如何。

5. 成对比较阵残缺时的处理

专家或有关人士由于某种原因会无法或不愿对某两个因素给出相互对比的结果 a_{ij} , 于是成对比较阵出现残缺 (不能补 0, 因为要求 $a_{ij} > 0$)。如何对此作修正, 以便继续进行权向量的计算呢, 下面通过简例介绍一种办法。

设一成对比较阵为 $A = \begin{bmatrix} 1 & 2 & \theta \\ 1/2 & 1 & 2 \\ \theta & 1/2 & 1 \end{bmatrix}$, 其中符号 θ 表示残缺。记由 A 要计

算的权向量为 $w = (w_1, w_2, w_3)^T$, 用 w_1/w_3 代替残缺的 a_{13} 是合理的, 所以构造一个

辅助矩阵 $C = \begin{bmatrix} 1 & 2 & w_1/w_3 \\ 1/2 & 1 & 2 \\ w_3/w_1 & 1/2 & 1 \end{bmatrix}$, 则 (5) 式 $Aw = \lambda w$ 可以代之以

$$Cw = \lambda w \quad (23)$$

但是 C 中包含未知量 w_1, w_3 , (23) 式无法求解, 而如果将 A 修正为 $\tilde{A} =$

$$\begin{bmatrix} 2 & 2 & 0 \\ 1/2 & 1 & 2 \\ 0 & 1/2 & 2 \end{bmatrix}, \text{ 不难验证}$$

$$\tilde{A}w = \lambda w \quad (24)$$

与 (23) 式等价。由 (24) 式可以得到权向量为 $w = (0.5714, 0.2857, 0.1439)^T$ 。

一般地, 由残缺阵 $A = (a_{ij})$ 构造修正阵 $\tilde{A} = (\tilde{a}_{ij})$ 的方法是令

$$\tilde{a}_{ij} = \begin{cases} a_{ij}, & a_{ij} \neq \theta, i \neq j \\ 0, & a_{ij} = \theta, i \neq j \\ m_i + 1, & m_i \text{ 为第 } i \text{ 行 } \theta \text{ 的个数}, i = j \end{cases} \quad (25)$$

在上面的例子中虽然因为元素 a_{13} 残缺, 没有比较 1, 3 两个因素的直接信息, 但是二者的比较可以通过 a_{12} 和 a_{23} 这样的间接信息获得. 一个应该提出的问题是, 怎样的残缺阵才是可以接受的, 即其残缺元素都能够由已有元素的关系得到.

已经证明^[70], 可以接受的残缺阵 A 的充分必要条件是 A 为不可约矩阵^①.

6. 递阶层次结构和更复杂的层次结构

以上讨论的所有层次结构模型有两个共同的特点, 一是模型所涉及的各项因素可以组合为属性基本相同的若干层次, 层次内部因素之间不存在相互影响或支配作用, 或者这种影响作用可以忽略; 二是层次之间存在自上而下、逐层传递的支配关系, 没有下层对上层的反馈作用, 或层间的循环影响. 具有这些特点的称为递阶层次结构, 前面介绍的全部算法都是针对这种层次结构的.

更复杂的层次结构有以下几种情况.

(1) 层次内部因素之间存在相互影响. 例如以行驶性能为目标对各种型号汽车作评价时, 准则层有刹车、转向、运行、加速等, 这些准则之间就是相关的, 如图 9.

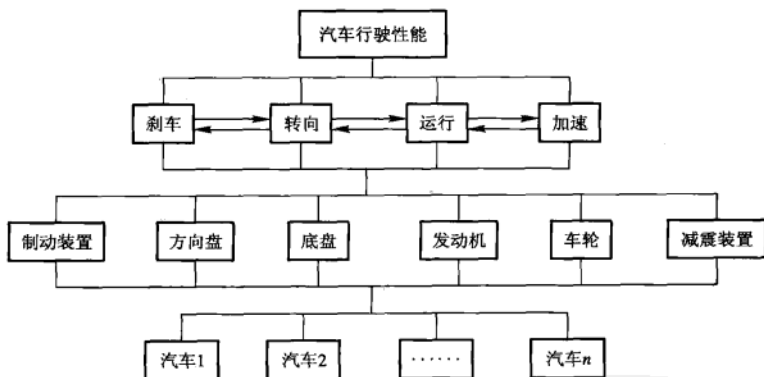


图9 汽车行驶性能的层次结构

(2) 下层反过来对上层有支配作用, 形成循环, 从而无法区分上下层. 例如可以用教学、科研等每一项指标评价几位教师, 也可以反过来对于每一位教师比较他的教学、科研等哪一方面表现最为出色, 从而在指标层和对象层之间形成循环.

(3) 既在层次内部因素之间存在相互影响, 又在层次之间存在反馈作用. 复杂的社会经济系统的层次结构就是这种情况, 它的一个简化模型如图 10. 产业、

① 非负方阵 A 若能通过行列置换化为 $\begin{bmatrix} A_1 & O \\ A_3 & A_4 \end{bmatrix}$ 形式 (其中 A_1, A_4 为方阵), 称 A 是可约矩阵; 否

则, 称 A 是不可约矩阵.

需求、政策等6个层次(或称子系统)之间存在复杂的相互关系(用带箭头的直线表示),在每层内部各因素(如产业包括农业、工业、第三产业,需求包括生活资料、社会发展资料、社会福利、国家安全等等)之间也有相互影响(用带箭头的弧线表示)。

用层次分析法研究这些更复杂的层次结构,需要引入超矩阵、极限相对权向量、极限绝对权向量等概念,并建立相应的算法^[65,90]。

评注 从层次分析法的原理、步骤、应用等方面的讨论不难看出它有以下优点。

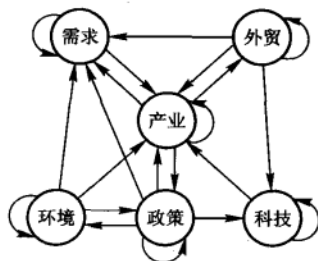


图10 社会经济系统的层次结构

(1) **系统性** 层次分析把研究对象作为一个系统,按照分解、比较判断、综合的思维方式进行决策,成为继机理分析、统计分析之后发展起来的系统分析的重要工具。

(2) **实用性** 层次分析把定性和定量方法结合起来,能处理许多用传统的优化技术无法着手的实际问题,应用范围很广。同时,这种方法将决策者与决策分析者相互沟通,决策者甚至可以直接应用它,这就增加了决策的有效性。

(3) **简洁性** 具有中等文化程度的人即可了解层次分析的基本原理和掌握它的基本步骤,计算也非常简便,并且所得结果简单明确,容易为决策者了解和掌握。

层次分析法的局限性可以用陈旧、粗略、主观等词来概括。就是说,第一,它只能从原有方案中选优,不能生成新方案;第二,它的比较、判断直到结果都是粗糙的,不适于精度要求很高的问题;第三,从建立层次结构模型到给出成对比较矩阵,人的主观因素的作用很大,这就使得决策结果可能难以为众人接受。当然,采取专家群体判断的办法是克服这个缺点的一种途径。

8.2 循环比赛的名次

若干支球队参加单循环比赛,各队两两交锋,假设每场比赛只计胜负,不计比分,且不允许平局。在循环赛结束后怎样根据他们的比赛结果排列名次呢^[12]。

有几种表述比赛结果的办法,较直观的一种是用图的顶点表示球队,而用连接两个顶点的、以箭头标明方向的边表示两支球队的比赛结果。图1给出了6支球队的比赛结果,即1队战胜2,4,5,6队,而输给了3队;5队战胜3,6队,而输给了1,2,4队等等。

根据比赛结果排名次的一个办法是在图中顺箭头方向寻找一条通过全部6个顶点的路径,如3→1→2→4→5→6,这表示3队胜1队,1队胜2队,……,于是3

队为冠军,1 队为亚军等等.但是还可以找出其他路径,如 $1 \rightarrow 4 \rightarrow 6 \rightarrow 3 \rightarrow 2 \rightarrow 5$, $4 \rightarrow 5 \rightarrow 6 \rightarrow 3 \rightarrow 1 \rightarrow 2$ 等.所以用这种方法显然不能决定谁是冠军.

排名次的另一个办法是计算得分,即每支球队获胜的场数.上例中 1 队胜 4 场,2,3 队各胜 3 场,4,5 队各胜 2 场,6 队胜 1 场.由此虽可决定 1 队为冠军,但 2,3 队之间与 4,5 队之间无法决出高低.如果只因为 $3 \rightarrow 2, 4 \rightarrow 5$,就将 3 排在 2 之前,4 排在 5 之前,则未考虑它们与其他队的比赛结果,是不恰当的.

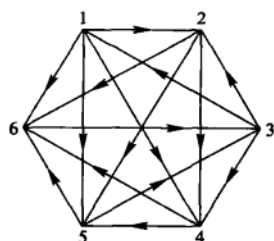


图 1 6 支球队的比赛结果

下面利用图论的有关知识解决这个问题.

竞赛图及其性质 在每条边上都标出方向的图称为有向图(Digraph).每对顶点之间都有一条边相连的有向图称为竞赛图(Tournament).只计胜负、没有平局的循环比赛的结果可用竞赛图表示,如图 1.问题归结为如何由竞赛图排出顶点的名次.

2 个顶点的竞赛图排名次不成问题.

3 个顶点的竞赛图只有图 2 的两种形式(不考虑顶点的标号).对于(1),3 个队的名次排序显然应是 $\{1, 2, 3\}$;对于(2),则 3 个队名次相同,因为他们各胜一场.

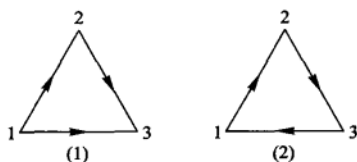


图 2 3 个顶点的竞赛图

4 个顶点的竞赛图共有图 3 所示的 4 种形式,下面分别进行讨论.

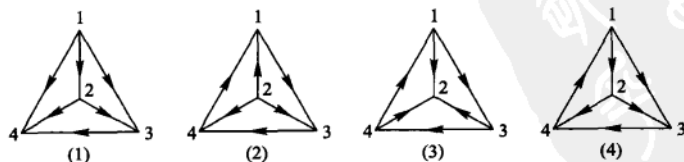


图 3 4 个顶点的竞赛图

(1) 有唯一的通过全部顶点的有向路径 $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$,这种路径称完全路径;4 个队得分为 $(3, 2, 1, 0)$.名次排序无疑应为 $\{1, 2, 3, 4\}$.

(2) 点 2 显然应排在第 1,其余 3 点如图 2(2)形式,名次相同;4 个队得分

为(1,3,1,1).名次排序记作{2,(1,3,4)}.

(3) 点2排在最后,其余3点名次相同;得分为(2,0,2,2).名次排序记作{(1,3,4),2}.

(4) 有不只一条完全路径,如 $1 \rightarrow 2 \rightarrow 3 \rightarrow 4, 3 \rightarrow 4 \rightarrow 1 \rightarrow 2$ 无法排名次;得分为(2,2,1,1).由得分只能排名为{(1,2),(3,4)},如果由 $1 \rightarrow 2, 3 \rightarrow 4$ 就简单地排名为{1,2,3,4}是不合适的.这种情形是研究的重点.

还可以注意到,(4)具有(1)~(3)所没有的性质:对于任何一对顶点,存在两条有向路径(每条路径由一条或几条边组成),使两顶点可以相互连通,这种有向图称为双向连通的(Biconnected).

5个顶点以上的竞赛图虽然更加复杂,但基本类型仍如图3所给出的3种:第1种类型:有唯一完全路径的竞赛图,如(1);第2种类型:双向连通竞赛图,如(4);第3种类型:不属于以上类型,如(2),(3).

一般的 n 个顶点的竞赛图具有以下性质:

1) 竞赛图必存在完全路径(可用归纳法证明).

2) 若存在唯一的完全路径,则由完全路径确定的顶点的顺序,与按得分多少排列的顺序相一致.这里一个顶点的得分指由它按箭头方向引出的边的数目.

显然,性质2给出了第1种类型竞赛图的排名次方法,第3种类型竞赛图无法全部排名,下面只讨论第2种类型.

双向连通竞赛图的名次排序 3个顶点的双向连通竞赛图,如图2(2),名次排序相同.以下讨论 $n(\geq 4)$ 个顶点的双向连通竞赛图.

为了用代数方法进行研究,定义竞赛图的邻接矩阵 $A = (a_{ij})_{n \times n}$ 如下:

$$a_{ij} = \begin{cases} 1, & \text{存在从顶点 } i \text{ 到 } j \text{ 的有向边} \\ 0, & \text{否则} \end{cases} \quad (1)$$

依此,图3(4)的邻接矩阵为

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (2)$$

若记顶点的得分向量为 $s = (s_1, s_2, \dots, s_n)^T$,其中 s_i 是顶点 i 的得分,则由(1)不难知道

$$s = A\mathbf{1}, \quad \mathbf{1} = (1, 1, \dots, 1)^T \quad (3)$$

由(2),(3)式容易算出双向连通的图3(4)的得分向量是 $s = (2, 2, 1, 1)^T$,正如前面已经给出的.由 s 无法排出全部名次.

记 $s = s^{(1)}$,称为1级得分向量,进一步计算

$$s^{(2)} = As^{(1)} \quad (4)$$

称为2级得分向量.每支球队(顶点)的2级得分是他战胜的各个球队的(1级)得分之和,与1级得分相比,2级得分更有理由作为排名次的依据.对于图3(4),
 $s^{(1)} = (2, 2, 1, 1)^T$, $s^{(2)} = (3, 2, 1, 2)^T$. 继续这个程序,得到 k 级得分向量.

$$s^{(k)} = A s^{(k-1)} = A^k \mathbf{1}, \quad k = 1, 2, \dots \quad (5)$$

对于图3(4)有

$$\begin{aligned} s^{(3)} &= (3, 3, 2, 3)^T, & s^{(4)} &= (5, 5, 3, 3)^T, \\ s^{(5)} &= (8, 6, 3, 5)^T, & s^{(6)} &= (9, 8, 5, 8)^T, \\ s^{(7)} &= (13, 13, 8, 9)^T, & s^{(8)} &= (21, 17, 9, 13)^T, \end{aligned}$$

k 越大,用 $s^{(k)}$ 作为排名次的依据越合理,如果 $k \rightarrow \infty$ 时, $s^{(k)}$ 收敛于某个极限得分向量(为了不使它无限变大,应进行归一化),那么就可以用这个向量作为排名次的依据.

极限得分向量是否存在呢? 答案是肯定的. 因为对于 $n (\geq 4)$ 个顶点的双向连通竞赛图,存在正整数 r ,使得邻接矩阵 A 满足 $A^r > 0$. 这样的 A 称为素阵.

再利用著名的 Perron-Frobenius 定理,素阵 A 的最大特征根为正单根 λ , λ 对应正特征向量 s ,且有^①

$$\lim_{k \rightarrow \infty} \frac{A^k \mathbf{1}}{\lambda^k} = s \quad (6)$$

与(5)式比较可知, k 级得分向量 $s^{(k)}$, $k \rightarrow \infty$ 时(归一化后)将趋向 A 的对应于最大特征根的特征向量 s , s 就是作为排名次依据的极限得分向量. 如对图3(4),算出其邻接矩阵 A ((2)式)的最大特征根 $\lambda = 1.4$ 和对应特征向量 $s = (0.323, 0.280, 0.167, 0.230)^T$,从而确定名次排列为 $\{1, 2, 4, 3\}$. 可以看出,虽然3胜了4,但由于4战胜了最强大的1,所以4排名在3之前.

对于本节开始提出的6支球队循环比赛的结果(图1),不难看出这个竞赛图是双向连通的. 写出其邻接矩阵

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad (7)$$

由(5)式可以算出各级得分向量为

$$\begin{aligned} s^{(1)} &= (4, 3, 3, 2, 2, 1)^T, & s^{(2)} &= (8, 5, 9, 3, 4, 3)^T, \\ s^{(3)} &= (15, 10, 16, 7, 12, 9)^T, & s^{(4)} &= (38, 28, 32, 21, 25, 16)^T, \end{aligned}$$

① 这里的结果与8.1节第三小节定理1是一致的.

进一步算出 A 的最大特征根 $\lambda = 2.232$ 和特征向量 $s = (0.238, 0.164, 0.231, 0.113, 0.150, 0.104)^T$, 排出名次为 $\{1, 3, 2, 5, 4, 6\}$ 。

8.3 社会经济系统的冲量过程

本节用一个能源利用系统预测的例子说明用冲量过程建模的方法。

考察某地区的能源利用状况。先界定系统的范围,比如只考虑能源利用量、价格、生产率、环境质量、工业产值、就业机会及人口总数 7 个因素,它们之间相当复杂的关系可以简化为一个因素对另外因素直接的促进或促退作用。要研究的问题是,当某个因素突然发生改变时,预测系统各因素的演变过程和趋势^[50]。

定性模型与定量模型 能源利用系统的每个因素用图的一个顶点表示,因素间的直接影响用带方向的边表示,构成 8.2 节中提到的有向图。为了表明因素间的影响作用是促进的(正面的)还是促退的(负面的),在箭头旁分别标以 + 号或 - 号。于是整个系统可以用带符号的有向图(Signed Digraph) G_1 表示,如图 1。这里有两点需要说明:第一,两顶点之间的有向边表示两因素间的直接影响,如 $v_1 v_3$ 带 + 号表示某时段能源利用量 v_1 的增加导致下一时段能源生产率 v_3 的增长, $v_1 v_4$ 带 - 号则表示 v_1 的增加导致下一时段环境质量 v_4 的降低。因素间的间接影响是由几条相连的同向边反映出来的,如 v_1 的增加导致下时段 v_3 的增长,又引起再下时段工业产值 v_5 的增长。有些因素间的影响是双向的,如 v_1 的增加导致下时段能源价格 v_2 的降低, v_2 的降低又引起再下时段 v_1 的增长。第二,像能源利用这样的社会经济系统,因素之间的影响关系十分复杂,应该合理、简化地确定哪些因素间有直接影响,这里除了主要根据客观规律作出决定外,方针政策有时也是判断的依据。如能源利用量 v_1 增加时能源价格 v_2 是降低($v_1 v_2$ 为 -)还是提高($v_1 v_2$ 为 +),是由鼓励能源利用还是限制能源利用这样的价格政策决定的。

为了用代数方法研究带符号的有向图,定义它的邻接矩阵 $A = (a_{ij})$ 如下:

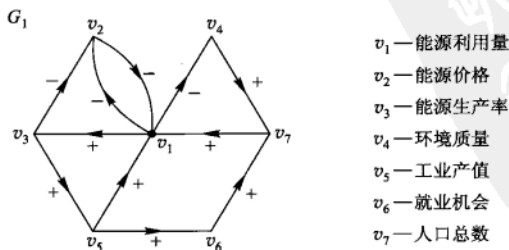


图 1 能源利用系统带符号的有向图

$$a_{ij} = \begin{cases} 1, & \text{若 } v_i v_j \text{ 为 } + \\ -1, & \text{若 } v_i v_j \text{ 为 } - \\ 0, & \text{若 } v_i v_j \text{ 不存在} \end{cases} \quad (1)$$

图 1 的邻接矩阵为

$$A = \begin{bmatrix} 0 & -1 & 1 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (2)$$

用图 G_1 和矩阵 A 表示的是定性模型,如果将系统各因素加以量化,并定量地描述因素间的相互影响,那么可以得到定量模型,用加权的有向图 (Weighted Digraph) G_2 描述,如能源利用系统又可用图 2 表示。这里能源利用量 v_1 和生产率 v_3 均以变化 10% 为 1 个单位,图 2 中表示 v_1 增加 1 个单位将引起 v_3 0.8 个单位的增长。像 0.8 这样标在图中箭头旁的数字称为权。用权代替 (1) 式中 a_{ij} 的 ± 1 取值,记作 w_{ij} (当 $a_{ij}=0$ 时 $w_{ij}=0$),就得到加权有向图的邻接矩阵 W 。图 2 的邻接矩阵如 (3) 式

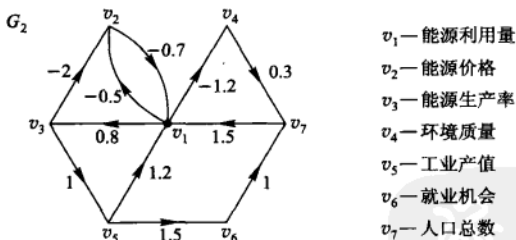


图 2 能源利用系统加权的有向图

$$W = \begin{bmatrix} 0 & -0.5 & 0.8 & -1.2 & 0 & 0 & 0 \\ -0.7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.3 \\ 1.2 & 0 & 0 & 0 & 0 & 1.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1.5 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (3)$$

显然,带符号的有向图的邻接矩阵 A 可视为加权有向图邻接矩阵 W 的特例,以下用 w_{ij} 代替 a_{ij} 表示 v_i 对下一时段 v_j 的影响。

冲量过程 (Pulse Process) 为了研究系统的某因素在初始时段的一个突然变化所引起的整个系统在其后各时段的演变过程,用 $v_i(t)$ 表示因素 v_i 在时段 t 的值, $p_i(t)$ 表示 v_i 在时段 t 的改变量(冲量)。设系统共有 n 个因素,根据 w_{ij} 的含义显然有^①

$$v_i(t+1) = v_i(t) + p_i(t+1), \quad i = 1, 2, \dots, n, \quad t = 0, 1, 2, \dots \quad (4)$$

$$p_j(t+1) = \sum_{i=1}^n w_{ij} p_i(t), \quad j = 1, 2, \dots, n, \quad t = 0, 1, 2, \dots \quad (5)$$

记行向量 $\mathbf{v}(t) = (v_1(t), v_2(t), \dots, v_n(t))$, $\mathbf{p}(t) = (p_1(t), p_2(t), \dots, p_n(t))$, 则(4), (5)可表为

$$\mathbf{v}(t+1) = \mathbf{v}(t) + \mathbf{p}(t+1) \quad (6)$$

$$\mathbf{p}(t+1) = \mathbf{p}(t) \mathbf{W}, \quad t = 0, 1, 2, \dots \quad (7)$$

如果只考虑系统在初始状态基础上的变化,不妨设

$$\mathbf{v}(0) = \mathbf{p}(0) \quad (8)$$

于是当系统在初始时段的冲量 $\mathbf{p}(0)$ 给定后,就可以用(6)~(8)计算系统在任意时段 t 的冲量 $\mathbf{p}(t)$ 和值 $\mathbf{v}(t)$, 即对系统的演变作出预测。

对于能源利用系统用(2)式给出的 A 作为 W , 并设初始时段能源利用量 v_i 有一个突然增加, 即 $\mathbf{p}(0) = (1, 0, 0, 0, 0, 0, 0)$, 按照(6)~(8)得到的 $\mathbf{p}(t)$, $\mathbf{v}(t)$ 如表1所示。

表1 能源利用系统的 $\mathbf{p}(t)$ 和 $\mathbf{v}(t)$

t	p_1	p_2	p_3	p_4	p_5	p_6	p_7	v_1	v_2	v_3	v_4	v_5	v_6	v_7
0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
1	0	-1	1	-1	0	0	0	1	-1	1	-1	0	0	0
2	1	-1	0	0	1	0	-1	2	-2	1	-1	1	0	-1
3	1	-1	1	-1	0	1	0	3	-3	2	-2	1	1	-1
\vdots					

这些结果可以直接从图1得到解释。比如因为 $v_1 v_4$ 为 -, 所以 $p_4(1) = -1$; 因为没有从 v_1 经两条边到 v_4 的路径, 所以 $p_4(2) = 0$; 因为从 v_1 有经过3条符号均为 - 的边到 v_4 的路径 ($v_1 \rightarrow v_2 \rightarrow v_1 \rightarrow v_4$), 所以 $p_4(3) = -1$ 。顺便指出, 如果 a_{ij} 中的 ± 1 没有定量含义的话, 那么表中的具体数字也没有数量意义, 只能表示正或负面的影响所及。

① 这里假设系统是封闭的, 即不考虑系统与外界环境的相互影响。

这种由某些因素在初始时段的变化(冲量)引起的系统的演变过程称为冲量过程. 当初始冲量 $p(0)$ 中只有一个分量是 1, 其余为 0 时, 称为简单冲量过程. 显然, 冲量过程可以分解为若干简单冲量过程的叠加. 以下只讨论简单冲量过程, 记作 S .

简单冲量过程的稳定性 如果系统的任一因素在 $t=0$ 的变化, 引起的系统各因素在任意时段的冲量和值都不会无限增长(或无限减少), 那么这个系统的冲量过程 S 是稳定的. 我们要得到判断 S 是否稳定的条件, 并且当 S 不稳定时希望找到一种途径, 通过改变某些因素之间可以控制的制约关系, 把不稳定过程调整成稳定的.

对于所有的 i 和任意的 t , 若 $|p_i(t)|$ 有界, 称 S 是冲量稳定的 (Pulse Stable), 若 $|v_i(t)|$ 有界, 称 S 是值稳定的 (Value Stable). 由(4)式可知, 若 S 是值稳定的, 必是冲量稳定的, 反之不然.

因为 $p(t), v(t)$ 由一阶差分方程(6), (7)确定, 所以粗略地看, 它们是否有界取决于 W 的特征根的模小于 1 还是大于 1. 更精确的判断由下面的定理给出^①. 设 λ 是 W 的非零特征根.

定理 1 S 冲量稳定的必要条件是 $|\lambda| \leq 1$.

特别, 当 w_{ij} 只取整数值如 $\pm 1, 0$ 时, 条件化为 $|\lambda| = 1$. 必要条件不满足时必存在某个简单冲量过程, 其冲量不稳定.

定理 2 S 冲量稳定的充分条件是 $|\lambda| \leq 1$ 且均为单根.

定理 3 S 值稳定的充要条件是 S 冲量稳定且 $\lambda \neq 1$.

用上述定理检查用带符号的有向图(图 1)表示的能源利用系统. 由(2)式给出的 A 的特征多项式为

$$f(\lambda) = \lambda^2(\lambda^5 - \lambda^3 - \lambda^2 - 1)$$

因为 $f(1) = -2, f(2) = 76$, 所以 $f(\lambda)$ 在 $(1, 2)$ 内有根, 据定理 1, 必有某个简单冲量过程的冲量不稳定.

不稳定过程的调整办法 为了把不稳定的过程调整成稳定的必须改变 W (或 A) 的某些元素, 使特征根满足定理 2, 3 的要求, 但是矩阵元素与特征根之间的关系是复杂的. 在如何改变系统的邻接矩阵的结构使其冲量过程稳定方面, 目前尚没有一般的结论. 下面介绍的方法适用于一类用特殊的有向图描述的系统.

如果一个带符号的有向图是双向连通的(参看 8.2 节), 并且存在一个位于所有闭路上的中心顶点, 那么称它是改进的玫瑰形图 (Advanced Rosette). 所谓闭路是指从某顶点出发, 沿有向边前进回到该顶点, 且途中不经过重复顶点的路

^① 定理中 S 指所有的简单冲量过程, 即 $p(0)$ 中的元素 1 可以是任何一个分量. 定理证明见[50].

径.不难看出,图1满足这个要求, v_1 是中心顶点,由 v_1v_3, v_3v_5, v_5v_1 构成的闭路记作 $v_1v_3v_5v_1$.构成闭路的有向边的数目称为闭路的长度;当闭路包含奇数个带-号的有向边时,定义闭路的符号为-1,否则为+1;用 a_k 记长度为 k 的闭路的符号和(当没有长度为 k 的闭路时令 $a_k=0$),设 r 是使 $a_r \neq 0$ 的最大整数.这样,对于每个用改进的玫瑰形图描述的冲量过程(记作 S^*),存在一个序列 $\{a_1, a_2, \dots, a_r\}$, S^* 的稳定性完全由这个序列决定.我们不加证明地引述下面的结论.

首先,若对于所有的 $k=1, 2, \dots$ 都有 $a_k=0$ (即 $r=0$),则 S^* 是冲量稳定和值稳定的.以下设 $r>0$.

定理4 S^* 冲量稳定的必要条件是

$$a_r = \pm 1 \quad (9)$$

$$a_k = -a_r a_{r-k}, \quad k=1, 2, \dots, r-1 \quad (10)$$

定理5 若 S^* 冲量稳定,则 S^* 值稳定的充要条件是

$$\sum_{k=1}^r a_k \neq 1 \quad (11)$$

定理4和5可以用来指导我们改变邻接矩阵 A 的某些元素 a_{ij} ,以便把不稳定的冲量过程调整为稳定的.对于图1所示的能量利用系统,为了找出导致不稳定的原因,先检查定理4中哪些必要条件不成立.

在图1中,没有长度为1的闭路, $a_1=0$;有1条长度为2的闭路 $v_1v_2v_1$,两条有向边均为-号, $a_2=1$;有3条长度为3的闭路: $v_1v_3v_2v_1$ 的符号是+1, $v_1v_3v_5v_1$ 是+1,而 $v_1v_4v_7v_1$ 是-1,所以 $a_3=1+1-1=1$;没有长度为4的闭路, $a_4=0$;有1条长度为5的闭路 $v_1v_3v_5v_6v_7v_1$,符号是+1, $a_5=1$;对于 $k>5$, $a_k=0$,于是 $r=5$.这样我们得到序列 $\{a_1, a_2, a_3, a_4, a_5\} = \{0, 1, 1, 0, 1\}$.

用定理4的(9),(10)两式检查上述结果发现, $a_2 = -a_5 a_3$ 和 $a_3 = -a_5 a_2$ 不成立.由此可见,为了满足冲量稳定的必要条件,需要将 a_2, a_3, a_5 中的一个由1变为-1.由图1知道,将 v_1v_2 的符号由-改成+可以达到上述目的(a_2 由1变为-1,而 a_3, a_5 不变).这个改动的实际含义是,将原来的“能源利用量增加引起价格下降”,调整为“能源利用量增加导致价格上升”,即从鼓励能源的利用调整为限制利用,这种价格政策的变动一般属于人们可以控制的范围.

当 v_1v_2 的符号改为+号后,邻接矩阵 A 的特征多项式变为

$$f(\lambda) = \lambda^2(\lambda^5 + \lambda^3 - \lambda^2 - 1) = \lambda^2(\lambda - 1)(\lambda^2 + 1)(\lambda^2 + \lambda + 1)$$

特征根是 $\lambda=0, 0, 1, \pm i, \frac{1}{2}(-1 \pm \sqrt{3}i)$.由定理2知冲量稳定,即能源利用系统在任何一个因素突然变化的影响下,各因素在后继时段的变化量都是有界的.

如果进一步检查 S^* 值稳定的充要条件(11)式,就会发现 $\sum_{k=1}^5 a_k = 0 - 1 + 1 + 0 + 1 = 1$. 为了在不破坏(9),(10)的基础上又使(11)式成立,必须使 $a_3 = a_5 = -1$. 这要求将长度为3的一条闭路和长度为5的闭路的公共有向边改变符号. 我们由图1看到,只能是把 $v_3 v_5$ 由+改为-. 其实际含义是,将原来的“能源生产率增加导致工业产值增长”,改变为“能源生产率增加导致工业产值减少”. 这似乎违反客观规律,不是人们能够控制的. 实际上,在能源利用量增加的刺激下,按照客观规律,能源生产率、工业产值、就业机会等因素会越来越大,要求值稳定不一定是合适的.

当然,还有其他的把原来不稳定的能源利用系统的冲量过程调整成冲量稳定和值稳定的办法(习题9).

评注 本节以能源利用系统的预测及其稳定性为题,介绍了一种用带符号(及加权)的有向图和冲量过程建立模型的方法. 它是定性和定量相结合的系统分析方法,适合于社会经济领域复杂大系统的宏观研究. 解决问题的关键是确定研究的对象和范围,即系统的边界,以及各因素间的相互关系. 在这个过程中合理的简化是必不可少的,有时还要划分为若干子系统,建立子模型,用分解、协调和综合的方法处理.

过去,用数学工具研究、解决实际问题主要有机理分析、统计分析方法,近几十年来发展起来的系统分析是又一种方法,8.1节的层次分析模型和本节的冲量过程模型属于这种方法,显然,它们是研究复杂的社会经济系统的有力工具.

8.4 公平的席位分配

某学校有3个系共200名学生,其中甲系100名,乙系60名,丙系40名. 若学生代表会议设20个席位,公平而又简单的席位分配办法是按学生人数的比例分配,显然甲乙丙三系分别应占有10,6,4个席位.

现在丙系有6名学生转入甲乙两系,各系人数如表1第2列所示. 仍按比例(表1第3列)分配席位时出现了小数(表1第4列). 在将取得整数的19席分配完毕后,三系同意参照所谓惯例分给比例中小数最大的丙系,于是三系仍分别占有10,6,4个席(表1第5列).

因为有20个席位的代表会议在表决提案时可能出现10:10的局面,会议决定下一届增加1席. 他们按照上述办法重新分配席位,计算结果见表1第6,7列. 显然这个结果对丙系太不公平了,因为总席位增加1席,而丙系却由4席减为3席.

表 1 按照比例并参照惯例的席位分配

系别	学生人数	学生人数 的比例/%	20 个席位的分配		21 个席位的分配	
			按比例分 配的席位	参照惯 例的结果	按比例分 配的席位	参照惯 例的结果
甲	103	51.5	10.3	10	10.815	11
乙	63	31.5	6.3	6	6.615	7
丙	34	17.0	3.4	4	3.570	3
总和	200	100	20	20	21	21

看来问题出在所谓按照惯例分配上,实际上由 A. Hamilton 提出的这种办法在美国国会 1850—1900 年的众议员席位分配(按照人口比例每个州应分得几个席位)中就多次被采用,也同时被质疑^[27,51],称之为最大剩余法(GR: Greatest Remainders)或最大分数法(LF: Largest Fractions),也称 Hamilton 法或 Vinton 法. 它被质疑的一个理由就是上例出现的所谓席位悖论——总席位增加反而可能导致某州席位的减少^①. 1880 年美国众议员席位分配中亚拉巴马(Alabama)州就曾遇到这种情况,所以这个悖论又称亚拉巴马悖论.

最大剩余法的另一个重大缺陷是所谓人口悖论——某州人口增加反而可能导致该州席位的减少. 如上例若三系学生变为 114, 64, 34 名,按照最大剩余法 21 席的分配结果将是 11, 6, 4 席,乙系学生人数增加却比原来少了 1 席,丙系学生数量未变反而多了 1 席.

为了寻求新的、公平的席位分配方法,先讨论衡量公平的数量指标^[5,17].

不公平度指标 为简单起见考虑 A, B 两方分配席位的情况. 设两方人数分别为 p_1, p_2 , 占有席位分别为 n_1, n_2 , 则比值 $p_1/n_1, p_2/n_2$ 为两方每个席位所代表的人数. 显然仅当 $p_1/n_1 = p_2/n_2$ 时分配才是完全公平的, 但是因为人数和席位都是整数, 所以通常 $p_1/n_1 \neq p_2/n_2$, 分配不公平, 并且是对比值较大的一方不公平.

不妨设 $p_1/n_1 > p_2/n_2$, 不公平程度可用数值 $p_1/n_1 - p_2/n_2$ 衡量. 如设 $p_1 = 120, p_2 = 100, n_1 = n_2 = 10$, 则 $p_1/n_1 - p_2/n_2 = 12 - 10 = 2$, 它衡量不公平的绝对程度, 常常无法区分不公平程度明显不同的情况. 如当双方人数增至 $p_1 = 1\ 020, p_2 = 1\ 000$, 而 n_1, n_2 不变时, $p_1/n_1 - p_2/n_2 = 102 - 100 = 2$, 即不公平的绝对程度不变, 但是常识告诉我们, 后面这种情况的不公平程度比起前面来已经大为改善了.

① 美国国会众议院的总席位从 1787 年的 65 逐渐增加到 1920 年的 435, 此后一直固定不变.

为了改进上述的绝对标准,自然想到用相对标准. 仍设 $p_1/n_1 > p_2/n_2$, 定义

$$r_A(n_1, n_2) = \frac{p_1/n_1 - p_2/n_2}{p_2/n_2} \quad (1)$$

为对 A 的相对不公平度. 若 $p_2/n_2 > p_1/n_1$, 定义

$$r_B(n_1, n_2) = \frac{p_2/n_2 - p_1/n_1}{p_1/n_1} \quad (2)$$

为对 B 的相对不公平度.

建立了衡量分配不公平程度的指标 r_A, r_B 后, 制定席位分配的原则是使它们尽可能小.

新的分配方法 假设 A, B 两方已分别占有席位 n_1, n_2 , 利用相对不公平度 r_A, r_B 讨论当总席位增加 1 席时, 应该分配给 A 还是 B.

不失一般性可设 $p_1/n_1 \geq p_2/n_2$, 即对 A 不公平. 若增加的 1 席分配给 A, n_1 就变为 $n_1 + 1$, 分配给 B 就有 $n_2 + 1$, 原不等式可能出现以下 3 种情况 (只需讨论不等号的情况, 一旦等号出现, 按等式状况分配即可):

$$1. \frac{p_1}{n_1 + 1} > \frac{p_2}{n_2}, \text{说明即使 A 增加 1 席仍对 A 不公平, 这一席显然应分配给 A.}$$

$$2. \frac{p_1}{n_1 + 1} < \frac{p_2}{n_2}, \text{说明 A 增加 1 席将对 B 不公平, 参照 (2) 计算出对 B 的相对不公平度为}$$

$$r_B(n_1 + 1, n_2) = \frac{p_2(n_1 + 1)}{p_1 n_2} - 1 \quad (3)$$

$$3. \frac{p_1}{n_1} > \frac{p_2}{n_2 + 1}, \text{说明 B 增加 1 席将对 A 不公平, 参照 (1) 计算出对 A 的相对不公平度为}$$

$$r_A(n_1, n_2 + 1) = \frac{p_1(n_2 + 1)}{p_2 n_1} - 1 \quad (4)$$

(不可能出现 $\frac{p_1}{n_1} < \frac{p_2}{n_2 + 1}$ 的情况).

在使相对不公平度尽量小的分配原则下, 如果

$$r_B(n_1 + 1, n_2) < r_A(n_1, n_2 + 1) \quad (5)$$

则增加的 1 席应分配给 A, 反之, 则增加的 1 席应分配给 B (等号成立时可分给任一方). 根据 (3), (4) 两式, (5) 式等价于

$$\frac{p_2^2}{n_2(n_2 + 1)} < \frac{p_1^2}{n_1(n_1 + 1)} \quad (6)$$

还不难证明, 上述第 1 种情况 $\frac{p_1}{n_1 + 1} > \frac{p_2}{n_2}$ 也会导致 (6) 式. 于是我们的结论是:

当(6)式成立时增加的1席应分配给A,反之应分配给B.

这种方法可推广到有 m 方分配席位的情况. 设第 i 方人数为 p_i , 已占有 n_i 个席位, $i=1, 2, \dots, m$, 当总席位增加1席时, 计算

$$Q_i = \frac{p_i^2}{n_i(n_i + 1)}, \quad i = 1, 2, \dots, m \quad (7)$$

增加的1席应分配给 Q 值最大的一方, 此方法暂称之为 Q 值法.

下面用 Q 值法重新讨论本节开头提出的甲乙丙三系分配21个席位的问题.

对本例, Q 值法可以从 $n_1 = n_2 = n_3 = 1$ 开始按总席位每增1席计算, 但对这个问题的19席的分配结果是 $n_1 = 10, n_2 = 6, n_3 = 3$, 与最大剩余法得到的整数部分相同. 再每次增加1席计算

$$\text{第20席: } Q_1 = \frac{103^2}{10 \times 11} = 96.45, Q_2 = \frac{63^2}{6 \times 7} = 94.50, Q_3 = \frac{34^2}{3 \times 4} = 96.33,$$

Q_1 最大, 增加的1席应分配给甲系.

$$\text{第21席: } Q_1 = \frac{103^2}{11 \times 12} = 80.37, Q_2, Q_3 \text{ 同上, } Q_3 \text{ 最大, 增加的1席应分配给}$$

丙系.

这样, 21个席位的分配结果是三系分别占有11, 6, 4席, 看来丙系保住了按照最大剩余法分配将会失掉的1席. 可是, 如果你注意一下上面的计算过程就会发现, 当总席位是20席时, 结果为11, 6, 3席, 与最大剩余法的10, 6, 4席不同, 所以很难说这个方法对丙系有利还是不利.

虽然从 Q 值法与最大剩余法对这个具体问题不同的分配结果看, 难以对二者进行评判, 可是 Q 值法不仅有明确的不公平度指标, 而且由于它是每增加1席地计算 Q 值, 所以不会出现席位悖论和人口悖论. 实际上, 这个方法是19世纪20年代由哈佛大学数学家E. V. Huntington提出和推荐的一系列席位分配方法中的一个, 下面对这类方法作简单的介绍与比较^[5, 51].

Huntington 除数法 设共有 m 方分配 N 个席位, 第 i 方人数为 p_i , 记 $p = (p_1, p_2, \dots, p_m)$, $P = \sum_{i=1}^m p_i$. 席位分配是要寻求 $n = (n_1, n_2, \dots, n_m)$, 其中 n_i 是第 i 方分配的席位, 满足 $\sum_{i=1}^m n_i = N$, 且 n_i 均为非负整数.

按照最大剩余法(GR)分配席位时, 首先计算第 i 方精确的席位份额, 记作 q_i , 显然 $q_i = N \frac{p_i}{P}$, $i = 1, 2, \dots, m$, 若 q_i 均为整数, 令 $n_i = q_i$ 即可完成分配, 否则, 记 $[q_i]$ 为 q_i 的整数部分, 先将 $[q_i]$ 分给第 i 方, 然后将尚未分配的 $r = N - \sum_{i=1}^m [q_i]$ 个席位分给剩余 $q_i - [q_i]$ 中最大的 r 方, 每方1席.

Huntington 除数法是首先对于非负整数 n 定义一个非负单调增函数 $d(n)$, 当总席位为 s 时, 第 i 方分配的席位是 p, s 的函数, 记作 $f_i(p, s)$, 而且 $f_i(p, 0) = 0, i = 1, 2, \dots, m$, 让 s 每次 1 席地递增至 N , 按照以下准则分配^①:

设 $n_i = f_i(p, s)$, 若对于某个 k , 有

$$\frac{p_k}{d(n_k)} = \max_{i=1,2,\dots,m} \frac{p_i}{d(n_i)} \quad (8)$$

则令

$$f_k(p, s+1) = n_k + 1, \quad f_i(p, s+1) = n_i (i \neq k) \quad (9)$$

可以看出, $p_i/d(n_i)$ 是一种得到新席位的优先级的度量, 除数 $d(n)$ 则表示各方当前席位数的权重. 若定义除数 $d(n) = \sqrt{n(n+1)}$, 由 (6), (7) 式可知, Q 值法正是按照上述准则分配席位的.

取不同的除数就得到不同的方法, Huntington 推荐的 5 种除数法见表 2, 第 1 列是它们的名称和英文缩写, 前面讨论过的 Q 值法正式的名称是相等比例法 (EP). 第 2 列为采用的除数, 最后一列以人名命名的称谓也经常在文献中出现.

表 2 Huntington 推荐的 5 种除数法

Huntington 除数法	除数 $d(n)$	不公平度的度量指标 (设 $\frac{p_i}{n_i} \geq \frac{p_j}{n_j}$)	以人名命名的称谓
最大除数法 (GD; Greatest Divisors)	$n+1$	$\frac{n_j p_i}{p_j} - n_i$	Jefferson, d'Hondt
主要分数法 (MF; Major Fraction)	$n + \frac{1}{2}$	$\frac{n_j}{p_j} - \frac{n_i}{p_i}$	Webster
相等比例法 (EP; Equal Proportions)	$\sqrt{n(n+1)}$	$\frac{n_j p_i}{n_i p_j} - 1$	Hill
调和平均法 (HM; Harmonic Mean)	$\frac{2n(n+1)}{2n+1}$	$\frac{p_i}{n_i} - \frac{p_j}{n_j}$	Dean
最小除数法 (SD; Smallest Divisors)	n	$n_i - \frac{n_i p_j}{p_i}$	Adams

① 若 $d(0) = 0$, 则隐含在各方自动分得 1 席的基础上再作分配, 通常这是可行的, 除非是各方人数相差悬殊且总席位较少的情况.

我们已经看到,EP 方法可以使相对不公平度最小,由相对不公平度的定义

(1),(2)可知这相当于对任意两方 i, j ,使得 $\frac{p_i/n_i - p_j/n_j}{p_i/n_i} = \frac{n_j p_i}{n_i p_j} - 1$ 最小. 表 2 的第 3 列给出了 5 种除数法对不公平度的度量指标,其中 HM 与 EP 类似,但用的是绝对不公平度,而 MF 方法采用 $n_j/p_j - n_i/p_i$,即用单位人数(如千人)的席位数之差来度量不公平度. 不公平度的度量指标也可以看做一种均衡检验,即是否存在任意两方的席位转移(将一方的席位给另一方),使得转移后的度量指标更优.

用这些方法做的一个数值例子见表 3,其中第 2,3 列分别是 5 方的人数 p_i 和精确的席位份额 q_i ,其余各列是用 5 种方法得到的结果,请读者验证(留作习题 13,在计算过程中可以利用 q_i 从适当的总席位数 s 递推).

容易看出,5 种除数法的除数 $d(n)$ 满足 $n \leq d(n) \leq n+1$,并且在表 2 中是按照从大到小排列的. 一般情况下 GD 偏向人数 p_i 较大的一方,偏向程度也按照表 2 的顺序,SD 偏向人数较小的一方. Huntington 特别推荐偏向适中的 EP,该法 1930 年后一直在美国国会众议员席位分配中采用^①.

表 3 用 5 种除数法做的一个数值例子

	p_i	q_i	GD	MF	EP	HM	SD
A	9 061	9.061	10	9	9	9	9
B	7 179	7.179	7	8	7	7	7
C	5 259	5.259	5	5	6	5	5
D	3 319	3.319	3	3	3	4	3
E	1 182	1.182	1	1	1	1	2
总和	26 000	26	26	26	26	26	26

从除数法分配席位的过程(8),(9)可以得到其结果 $\mathbf{n} = (n_1, n_2, \dots, n_m)$ 满足以下 min-max 不等式:

$$\min_{k=1,2,\dots,m} \frac{p_k}{d(n_k - 1)} \geq \max_{i=1,2,\dots,m} \frac{p_i}{d(n_i)}, \quad \sum_{i=1}^m n_i = N \quad (10)$$

从优化的角度出发,席位分配问题可以归结为,给定 \mathbf{p} 和 N ,在 $\sum_{i=1}^m n_i = N$ 且 n_i 均为非负整数的条件下,寻求 n_1, n_2, \dots, n_m 使某一事先定义的目标函数最优.

① 1830 年前采用 GD,1850—1900 年采用 GR(有时辅以调整),1840 年和 1910 年采用 MF.

容易验证,本节开头用的最大剩余法 (GR) 得到的是 $\min \sum_{i=1}^m (n_i - q_i)^2$ 的解. 一些除数法也都可看作与某个理想状态偏差的加权最小二乘解,如相等比例法 (EP) 给出了 $\min \sum_{i=1}^m n_i \left(\frac{p_i}{n_i} - \frac{P}{N} \right)^2$ (等价于 $\min \sum_{i=1}^m \frac{(n_i - q_i)^2}{n_i}$) 的解,而主要分
数法 (MF) 给出了 $\min \sum_{i=1}^m p_i \left(\frac{n_i}{p_i} - \frac{N}{P} \right)^2$ (等价于 $\min \sum_{i=1}^m \frac{(n_i - q_i)^2}{q_i}$) 的解 (习题 14).

存在公平的席位分配方法吗 约两个世纪以来,出于美国和欧洲诸如议会席位分配等社会政治活动的需要,一些人包括数学家们先后提出了许多分配方法,这些方法对同一个问题常常给出不同的结果,还会出现违反人们意愿甚至违背常识的现象,这更引起数学家们深入研究的兴趣. 所谓公理化方法就是先提出公平的席位分配应该具有的若干性质,再找出满足这些性质的分配方法. 如果不存在这样的方法,那就讨论现有的方法分别满足其中的哪些,再做改进,使之满足更多的性质;或者改变、减少原来提出的性质,再作探寻.

作为初步的、简明的介绍,这里只给出公平的席位分配明显应该具备的 3 条主要性质 (仍用前面定义的符号, $n_i = f_i(p, s)$ 表示人数为 p 、总席位为 s 时分配给第 i 方的席位, $i = 1, 2, \dots, m$):

1. $[q_i] \leq n_i \leq [q_i] + 1$, 即 n_i 必是精确的席位份额 q_i 向下或向上取整得到,称为份额性.

2. $f_i(p, s) \leq f_i(p, s+1)$, 即总席位增加时各方的席位都不会减少,称为席位单调性.

3. 若对于任意的 $i, j = 1, 2, \dots, m, j \neq i, p'_i/p'_j \geq p_i/p_j$, 则 $f_i(p', s) \geq f_i(p, s)$ 或 $f_j(p', s) \leq f_j(p, s)$, 即当 i 方相对于 j 方人数增加时 (其他各方人数及总席位均不变), 不会导致 i 方席位减少而 j 方席位增加 (不排除 i, j 两方席位都增加或都减少)^①, 称为人口单调性.

我们已经看到,最大剩余法满足性质 1, 但会出现席位悖论和人口悖论从而不满足性质 2, 3. 而对于 5 种除数法, 从递推计算过程可知它们自然满足性质 2, 3, 是否能满足性质 1 呢? 试看表 4 给出的例子, 5 种除数法都不满足性质 1, 且其偏向十分明显. 当然, 从这个具体问题看, 因为第 2~6 方人数都差不多, 满足性质 1 的 GR (最后一列) 给第 2~3 方各 2 席, 给第 4~6 方各 1 席, 也不见得公平, 反而是 HM 或 GD, SD 的结果更值得考虑.

^① 在 $p'_i/p'_j = p_i/p_j$ 成立时可能出现 i 方席位减少而 j 方席位增加, 如原来人口相等的两方分奇数个席位, 哪一方多一席都是公平的, 后来两方人口成比例变化, 不论哪方多一席也是公平的.

表 4 5 种除数法都不满足性质 1 的数值例子

i	p_i	q_i	GD	MF	EP	HM	SD	GR
1	91 490	91.49	94	93	90	89	88	92
2	1 660	1.66	1	2	2	2	2	2
3	1 460	1.46	1	1	2	2	2	2
4	1 450	1.45	1	1	2	2	2	1
5	1 440	1.44	1	1	2	2	2	1
6	1 400	1.40	1	1	1	2	2	1
7	1 100	1.10	1	1	1	1	2	1
总和	100 000	100	100	100	100	100	100	100

是否存在满足所有 3 条性质的分配方法呢?事实上已经证明^[5]:对于 $m \geq 4, N \geq m+3$, 不存在满足上述 3 条性质的分配方法^①.

也许因为份额性和席位单调性一直是美国议会席位分配争论的主题之一, Balinski 和 Young 将 GR 与 GD 相结合, 提出了满足性质 1, 2 的份额法 (QM: Quota Method)^[3, 51] (习题 15).

在席位分配问题中, 影响分配结果的因素有人口 p , 总席位数 s 和参与分配的单位数 m , 性质 3 对于人口单调性的定义能扩充到 s 和 m 均可改变的情况, 即若对于任意的 $i, j = 1, 2, \dots, m, j \neq i, i', j' = 1, 2, \dots, m', j' \neq i', p'_i/p'_j \geq p_i/p_j$, 则 $f_i(p', s') \geq f_i(p, s)$ 或 $f_j(p', s') \leq f_j(p, s)$ ^②. 这样的定义已经蕴含了性质 2 席位单调性.

对于本节提出的分配问题一些更深入的性质, 诸如相容性、稳定性、无偏性等及其他方法的讨论, 还有在应用中可能遇到的给每方分配的席位事先设置最小值和最大值时的处理办法等, 读者可以从 [51] 及其参考文献中查到.

评注 起初, 对于出现在社会政治领域中的席位分配, 人们认为是一个简单的数学问题, 用初等的方法处理, 但是在应用过程中这样的分配方案得出许多难以接受的结果, 人们发现所有的方法都有不合理之处. 直到 20 世纪 70 年代 Balinski 和 Young 采用公理化方法进行研究, 才使这一问题的基本原理得以明晰.

虽然不存在同时满足 3 条性质的“理想”的分配方法, 但是在实际应用中“反例”毕竟是很少出现的, Balinski 和 Young 就极力推荐 MF, 指出自 1791 年以来用这个方法分配美国议会席位从未违反份额性, 并且按照模拟结果违反份额

① 对于 $m=2$, 所有除数法满足性质 1; 对于 $m=3$, MF 满足性质 1.

② 在 $p'_i/p'_j = p_i/p_j$ 成立时可能出现 i 方席位减少而 j 方席位增加, 如原来人口相等的两方分奇数个席位, 哪一方多一席都是公平的, 后来两方人口成比例变化, 不论哪方多一席也是公平的.

性的概率仅为 $1/1\ 600^{[3]}$. 实际上,只要掌握了各种方法的特性,可以根据具体情况决定采用哪一个,如只是一次分配,总席位和人口都不再变化,就不妨采用满足份额性的 GR,以示公正,而若是每隔一段时间进行重新分配,并且总席位和人口也在不断改变,那就需要考虑满足席位单调性和人口单调性的除数法了;此外,5种除数法对人口多少的偏向也是决定采用哪一种方法的重要因素。

8.5 存在公正的选举规则吗

你参加过评选优秀运动员、优秀电影、优秀小说这一类活动吗?你知道组织者们是如何根据许多人对评选对象的投票情况决定评选结果的吗?你考虑过怎样决定评选结果才是公正的吗?当然这又引出了新的问题:公正的标准是什么?在普遍赞同的一组标准下是否存在所谓公正的评选结果?

这种根据若干人对某些对象的决策结果,综合出这个群体的决策结果的过程称为**群体决策**(Group Decisionmaking).许多国家都有一些专门的机构用民意测验的方式,调查国民对社会福利、内外政策以及领导阶层的态度,然后用群体决策方法归纳出全体国民的倾向.事实上,在社会经济领域内许多问题都可以归结为群体决策.比如有相当多的产品的质量无法用仪器设备精确测量,而要靠专家们作出带有一定主观成分的评定.像在啤酒质量评定会上,由若干专家对各种啤酒进行品尝,每人对这些啤酒按优劣排序.最后由会议组织者综合专家们给出的排序,确定一个统一的排序结果.

为了叙述方便起见,我们把根据每个人对评选对象所作的排序来确定这个群体对评选对象排序的决策过程,描述为一次选举,有若干选民和若干候选人,每个选民的一次投票是他按照自己的标准对各候选人优劣顺序的一次排序,而选举规则要根据所有选民的排序确定选举结果(也是对各候选人的一个排序).本节首先给出这种排序的一般性质,讨论简单的选举规则并指出它们的缺点,然后论证在一定意义上所谓公正的选举规则是不存在的,最后介绍一个特殊情况下的选举规则以及另一种确定规则的方法^[11,64,76].

简单的选举规则 用 $I = (1, 2, \dots, n)$ 表示选民集合,由 m 个字母构成的 $A = (x, y, z, u, v, \dots)$ 表示候选人集合.选举要求每个选民 $i \in I$ 对全体候选人投票,即对 A 的一个排序,记作 p_i .所谓选举规则是根据 $p_i (i = 1, 2, \dots, n)$ 确定选举结果,即群体对 A 的排序,记作 p ,这种由 (p_1, p_2, \dots, p_n) 到 p 的对应关系在群体决策中称为**群体一致函数**(Group Consensus Function)^①.

① 一些国家用它来衡量人民对社会福利状况优劣的评价,称为**社会福利函数**(Social Welfare Function).

作为对集合 A 的任何一个排序 p_i 和 p , 必须具备以下两个性质, 也可以称为公理:

1. 对于任意的 $x, y \in A$, 下面三种关系必有且仅有一种成立: x 优于 y , 记 $x > y$; x 等同 y , 记 $x \sim y$; x 劣于 y , 记 $x < y$. 用 \geq 记优于或等同, \leq 记劣于或等同^①.

2. 对于 $x, y, z \in A$, 若 $x \geq y, y \geq z$, 则 $x \geq z$; 且仅当前两式等同号成立时, 后一式的等同号才成立. 这是排序的可传递性.

在现实生活中人们常用的简单的选举规则有以下两种:

1) 简单多数规则 当且仅当超过半数的选民投票 $x > y$ 时, 选举结果为 $x > y$ (\sim 和 $<$ 的关系类似).

以下用 $(x > y)_i$ 表示在排序 p_i 中 x 优于 y ; 如不特别指明, 用 $x > y$ 表示在排序 p 中 x 优于 y . 简单多数规则规定, 当且仅当含 $(x > y)_i$ 的排序 p_i 的数目超过 $n/2$ 时, $x > y$ 成立. 试看如下的一次选举: $I = (1, 2, 3)$, $A = (x, y, u, v)$, 投票结果为

$$p_1: x > y > u \sim v, \quad p_2: y > x > u > v, \quad p_3: x \sim u > v > y \quad (1)$$

按照简单多数规则选举结果应为

$$p: x > y > u > v \quad (2)$$

这个规则用起来十分方便, 但是有时用它确定的选举结果 p 违反排序的可传递性, 如考察

$$p_1: x > y > z, \quad p_2: y > z > x, \quad p_3: z > x > y \quad (3)$$

则得到的 p 中应有 $x > y, y > z, z > x$, 于是无法确定选举结果.

2) 记分规则 设 $B_i(x)$ 为排序 p_i 中劣于 x 的候选人的数目, 定义

$$B(x) = \sum_{i=1}^n B_i(x) \quad (4)$$

为 x 在选举中所得的分数, 称为 **Borda 数**. 记分规则规定, 当且仅当 $B(x) > B(y)$ 时, $x > y$ 成立 (\sim 和 $<$ 的关系类似).

在(1)式给出的投票结果中 $B_1(x) = 3, B_2(x) = 2, B_3(x) = 2, B(x) = 3 + 2 + 2 = 7$; 同样可得 $B(y) = 5, B(u) = 3, B(v) = 1$. 于是按记分规则得到的选举结果与用简单多数规则得到的(2)式一致. 而对于(3)式给出的投票, 因为 $B(x) = B(y) = B(z) = 3$, 所以按记分规则得到的结果为 $p: x \sim y \sim z$.

应该指出, 记分规则与对选民投票的要求是有矛盾的. 投票时要求选民只考虑候选人的优劣顺序, 不管其优劣程度, 而记分规则却显然以优劣程度为依据, 并且用同一尺度——分数来衡量 p_i 内部和 p_i 之间的优劣程度. 这样就可能出现一些不合理的情况, 试看下例: $I = (1, 2, 3, 4)$, $A = (x, y, z, u, v)$, 投票情况为

$$p_1, p_2, p_3: x > y > z > u > v, \quad p_4: y > z > u > v > x \quad (5)$$

① 本节中“ $<$ ”, “ \leq ”, “ $>$ ”, “ \geq ”及“ \sim ”均是排序符号.

算出 $B(x) = 12, B(y) = 13$, 按记分规则得到的结果应 $y > x$. 但在 (5) 中有三位选民认为 $x > y$, 只有一位认为 $y > x$, 只是因为他把 y 和 x 分别排在首位和末位, 致使选举结果服从了他个人的意愿, 这是不能为多数人接受的.

上面两种简单的选举规则在多数情况下可以得到合理的、相同的结果, 但在有些情况下则不然. 有没有适合一般情况的、公正合理的选举规则呢? 这首先涉及所谓公正合理的标准. 从数学建模的角度看, 应该先订下几条普遍认可的公理, 然后用逻辑推理方法寻求满足这些公理的规则 (如果这种规则存在的话). K. Arrow 作了这样的工作, 他给出了一组公理, 但是却证明了一个反面的结果: 不存在满足这组公理的选举规则. 他的这种建模方法及所得到的结果对我们是有启发的, 下面作简要介绍.

Arrow 公理与 Arrow 定理 Arrow 提出了 5 条公理^①:

公理 1 (选举的完全性)

选民对候选人的任何一种排序都是允许的.

公理 2 (选举结果与选民投票的正相关)

若对于某次投票 $p_i (i = 1, 2, \dots, n)$, 选举规则确定的选举结果 p 中包括 $x > y$, 而在另一次投票 $p'_i (i = 1, 2, \dots, n)$ 中 x 与 y 的顺序或者与 p_i 相同, 或者 x 提前, 而其他候选人的顺序不变. 那么在选举规则确定的另一次选举结果 p' 中也应包括 $x > y$.

这个公理表明若所有选民对候选人 x 的排序没有向后移动, 则选举结果中 x 对其他候选人的优先性不应改变.

公理 3 (无关候选人的独立性)

设 A_1 是候选人集合 A 的子集. 若在两次投票 p_i 和 $p'_i (i = 1, 2, \dots, n)$ 中 A_1 内各候选人的排序相同, 那么选举规则确定的两次选举结果 p 和 p' 中 A_1 内各候选人的排序也应相同.

该公理表明, 尽管 A_1 以外的候选人在 p_i 和 p'_i 中的排序可以变化, 但这不能影响 A_1 内各候选人在选举结果中的排序.

公理 4 (选民的主权性)

对于任意一对候选人 $x, y \in A$, 存在一种投票 $p_i (i = 1, 2, \dots, n)$, 使得选举规则能由 p_i 确定选举结果中有 $x > y$.

如果这条公理不成立, 那么无论选民怎样投票, 即使所有 p_i 中都有 $(x > y)_i$, 选举规则也不能确定 $x > y$, 这样的规则不尊重全体选民的一致愿望.

公理 5 (选民的独裁性)

不存在这样的选民 i , 使得对于任意一对候选人 x, y , 只要 p_i 中有 $(x > y)_i$,

^① 这里基本上采用了 [11] 的叙述, 与 Arrow 原来的提法稍有出入.

选举规则就确定 $x > y$.

因为这样的选民 i 事实上垄断了选举结果, 所以该公理表明不允许这种独裁者存在.

是否存在满足以上 5 条公理的选举规则呢?

当 $m=1$ 或 $n=1$ 时选举无意义, 不予讨论.

当 $m=2, n \geq 2$ 时简单多数规则就满足上述 Arrow 公理, 请读者验证 (习题 17).

当 $m \geq 3, n \geq 2$ 时 Arrow 证明了^①

定理 当至少有 3 位候选人和 2 位选民时, 不存在满足 Arrow 公理的选举规则.

面对这样的结果人们首先想到的是重新审查 Arrow 公理. 仔细研究这些看来合理的公理, 有人对公理 3 (无关候选人的独立性) 提出质疑, 试看下例.

设 $I = (1, 2), A = (x, y, z)$, 选举规则是, 若 $(x > y)_1, (x > y)_2$, 则 $x > y$; 若 $(x > y)_1, (x < y)_2$ 则 $x \sim y$. 这显然是公正的, 但是将要看到利用公理 3 会导致矛盾的结果. 考察 4 次投票, 前 3 次投票情况及按这个选举规则确定的结果为

$$\begin{cases} p_1^{(1)} : x > y > z \\ p_2^{(1)} : x > y > z \end{cases} \Rightarrow p^{(1)} : x > y > z \quad (6)$$

$$\begin{cases} p_1^{(2)} : x > z > y \\ p_2^{(2)} : z > x > y \end{cases} \Rightarrow p^{(2)} : x \sim z > y \quad (7)$$

$$\begin{cases} p_1^{(3)} : y > z > x \\ p_2^{(3)} : z > y > x \end{cases} \Rightarrow p^{(3)} : y \sim z > x \quad (8)$$

第 4 次投票情况为

$$\begin{cases} p_1^{(4)} : x > y > z \\ p_2^{(4)} : z > x > y \end{cases} \quad (9)$$

比较 (9) 与 (6), 候选人子集 $A_1 = (x, y)$ 在 $p_i (i=1, 2)$ 中的排序不变, 按照公理 3, $p^{(4)}$ 中 x, y 的排序与 $p^{(1)}$ 一致, 即 $x > y$;

比较 (9) 与 (7), $A_1 = (x, z)$ 的排序不变, 按照公理 3, $p^{(4)}$ 中 x, z 的排序与 $p^{(2)}$ 一致, 即 $x \sim z$;

比较 (9) 与 (8), $A_1 = (y, z)$ 的排序不变, 与上面类似地有 $y \sim z$.

上面三个结果显然是矛盾的, 其原因是公理 3 中虽然子集 A_1 内候选人排序不变, 但是子集 A_1 以外候选人 (即无关候选人) 的插入影响了 A_1 内排序的优劣程度. 公理 3 不考虑这种影响, 致使产生了无法统一的结果.

① 定理证明见 [11, 83]

如果修改公理 3 以考虑优劣程度呢,那么就会涉及对选民投票时只考虑候选人优劣顺序这一基本假定,必须另起炉灶.下面转向寻求另外的解决途径.

联合尺度下的选举规则 如果“候选人”是葡萄酒,那么可以按照它们的甜度在从 0 到 1 的尺度上客观地确定各自的位置.“选民”的口味不一,每个人可以在这个尺度上确定自己最理想的“候选人”应在的位置.在图 1 中 4 种葡萄酒 x, y, v, u 的甜度分别居于 0, 0.1, 0.7, 0.9 处, 3 名“选民”1, 2, 3 认为自己最理想的葡萄酒的位置(就甜度而言)分别位于 1, 0.4 和 0.6.

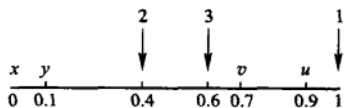


图 1 联合尺度示意图

图 1 这种表示称为选民和候选人的**联合尺度**,联合尺度的选举方法可以用在如颜色的深浅、口味的咸淡等物品质量的评定上.“候选人”按照客观标准确定在尺度上的位置,而“选民”则按照主观爱好分别在同一尺度上给出理想“候选人”的位置.

由联合尺度很容易得到各位选民的一次投票结果,如对图 1 有

$$p_1: u > v > y > x, \quad p_2: y \sim v > x > u, \quad p_3: v > u > y > x \quad (10)$$

如果按照简单多数规则确定选举结果,则有

$$p: v > u > y > x \quad (11)$$

p 与 p_3 完全一致,而选民 3 是三位选民在联合尺度上居中的那一位,这个结果不是偶然的,事实上有

结论 设有奇数个 ($2k+1$ 个) 选民,投票结果 $p_i (i=1, 2, \dots, 2k+1)$ 由选民和候选人的联合尺度得到, j 是联合尺度上居中的那位选民,则简单多数规则确定的选举结果 p 与 p_j 一致,并且简单多数规则符合 Arrow 公理^①.

我们只对这个结论作些解释.如果 $(x > y)_j$, 那么在联合尺度上 x, y 和 j 的位置可以出现图 2 所示的情况.由于 j 是居中者,他的左面(包括他所在位置)正好有 k 个选民,这些选民的投票结果必是 $x > y$, 于是至少有 $k+1$ 个选民的投票为 $x > y$, 按照简单多数规则确定的选举结果应为 $x > y$, 与 $(x > y)_j$ 一致.

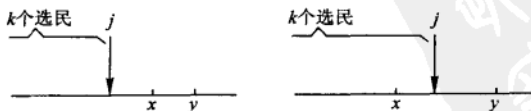


图 2 联合尺度上居中选民 j 与 x, y 的位置

① 这里的 p_i 由联合尺度得到,其范围比候选人的任意排序小得多. Arrow 公理涉及的投票应限制在这个小范围内.

由此可以看出,因为用联合尺度对投票情况作了限制,才可能存在满足 Arrow 公理的选举规则,并且这个规则简单易行——联合尺度上居中的选民的投票结果就是选举结果。

最小距离意义下的选举规则 这是一种与 Arrow 公理完全无关的方法. 仍用 $I = (1, 2, \dots, n)$ 和 $A = (x, y, z, \dots)$ 记选民和候选人集合, 用 P 记 A 的所有可能排序的集合(在排序基本性质 1 的意义下). 选民 i 的一次投票 p_i 可以看作集合 P 中的一个点. 如果能够合理地定义两点 p_i 和 p_j 之间的距离来衡量它们的接近程度, 那么从 I 的一次投票 (p_1, p_2, \dots, p_n) 确定选举结果, 就可以归结为在集合 P 中寻求一个点 p , 使它到 n 个点 p_1, p_2, \dots, p_n 的距离之和最小.

仍然可以用公理化方法来定义距离. 在一组合适的公理下, 能够构造性地证明, 如下定义的距离是唯一的(公理及证明见[11]).

任一对候选人 x, y 在选民 i, j 的一次投票 p_i, p_j 中的距离为:

$$\delta_{(x,y)}(p_i, p_j) = \begin{cases} 0, & p_i, p_j \text{ 中 } x, y \text{ 的排序相同} \\ 1, & p_i, p_j \text{ 中一个含 } x \sim y, \text{ 另} \\ & \text{一个含 } x > y, \text{ 或 } x < y \\ 2, & p_i, p_j \text{ 中 } x, y \text{ 的排序相反} \end{cases} \quad (12)$$

p_i, p_j 之间的距离为

$$d(p_i, p_j) = \sum_{(x,y) \in A} \delta_{(x,y)}(p_i, p_j) \quad (13)$$

其中求和的含义是候选人成对地跑遍集合 A . 例如对于

$$p_1: x > y > z, \quad p_2: x \sim z > y$$

按照定义(12), (13)有

$$d(p_1, p_2) = \delta_{(x,y)}(p_1, p_2) + \delta_{(x,z)}(p_1, p_2) + \delta_{(y,z)}(p_1, p_2) = 0 + 1 + 2 = 3$$

对于一次投票 (p_1, p_2, \dots, p_n) , 确定选举结果 p 的原则通常有两种, 一是使

$\sum_{i=1}^n d(p, p_i)$ 最小; 二是使 $\sum_{i=1}^n d^2(p, p_i)$ 最小. 前者平均地照顾各个选民的意见, 后者对于与多数选民看法不同的少数选民的意见予以更多的考虑.

这种最小距离意义下的选举规则有两个主要缺点, 一是尚没有在上述原则下求 p 的有效方法, 基本上只能利用穷举法, 当候选人数目稍多时计算量很大; 二是可能出现 p 不惟一的情况(不论在哪种原则下). 读者可以验证, 对于本节(3)式给出的投票, 用第一种原则确定的 p 是不惟一的(习题 20).

评注 我们用选民投票、选举规则等词汇叙述的群体决策问题在社会经济领域有着很强的实际背景和广泛的应用. 本节是用先规定排序的基本性质, 再提出一组公理, 通过逻辑推理, 寻求选举规则这样的过程来建立模型的. 可以知道, 如果提出的公理不合适(过多、相互矛盾、过于严格等), 则可能得不到满足这些

公理的结果;另一方面,如果提出的公理不充分(过少、过于宽松),则又可能无法推出结果或者结果不惟一。所以用这种方法建模常常需要多次的反复。我们在叙述了 Arrow 得到的反面结果后又介绍了两种方法,联合尺度下的选举规则是以缩小应用范围为代价换取一定结果的,最小距离意义下的选举规则应用上也存在诸多不便之处。

8.6 价格指数

消费品价格的变化在任何国家都是普通百姓十分关心的问题之一。由于政局不稳或市场失控,物价几倍、几十倍地上涨,固然会引起经济崩溃、民不聊生,而由政府把所有的价格定死,长期不准变动,也会导致生产停滞、比例失调,阻碍经济的正常发展。为了逐步理顺各个经济部门和各种生产品之间的关系,既要允许各种商品的价格有升有降,又要将消费品价格总的上涨幅度控制在一定范围之内。怎样衡量价格变化的趋势和程度呢?在资本主义社会发展的几百年历史中经济学家们提出了许多种所谓价格指数。目前我国有关部门常用的价格指数也是其中的一种。本节要讨论的问题是,如何评价已经存在的这些价格指数;根据客观的经济规律人们有理由要求价格指数满足哪些性质;能否找到满足这些性质的价格指数^[11,94]。

可以看出,这个问题的提法与 8.5 节的选举规则有相似之处,我们可以把价格指数应该满足的性质归纳为若干条公理,然后用逻辑推理的方法讨论这些公理的相容性和独立性,这样既能够从一个角度评价已有的一些价格指数,又可以回答满足上述性质的价格指数的存在性问题。

各种形式的价格指数 首先,对于一种固定的商品,若原来的价格为 p^0 ,现在的价格为 p ,那么可以简单地用

$$I = \frac{p}{p^0} \quad (1)$$

衡量价格的变动。如果两种商品原价是 p_1^0 和 p_2^0 ,现价是 p_1 和 p_2 ,简单地用如下的平均

$$I = \frac{p_1}{p_1^0} + \frac{p_2}{p_2^0} \quad \text{或} \quad \frac{p_1 + p_2}{p_1^0 + p_2^0} \quad \text{或} \quad \frac{p_1 p_2}{p_1^0 p_2^0} \quad (2)$$

显然都不能表示总的价格变化,因为若第一种商品是大米,第二种是钢琴,那么(2)式无法反映人们对大米涨价远比对钢琴降价更为关切的实际情况。我们自然想到应该用加权平均的办法。

与人民生活息息相关的消费品有成百上千种,各个国家和地区要根据具体情况选出具有代表性的若干种,作为制订价格指数的依据。价格指数通常是衡量

基准年(基年)和考察年(现年)价格的总的变动. 设 n 种代表性消费品在基年的价格为 $p_1^0, p_2^0, \dots, p_n^0$, 在现年的价格为 p_1, p_2, \dots, p_n , 按照它们对人民生活、国家财政等方面的影响综合考虑, 其权重分别为 $q_1^0, q_2^0, \dots, q_n^0$ 和 q_1, q_2, \dots, q_n , 可以用它们的销售量、销售额或其他指标作为权重. 记向量 $p^0 = (p_1^0, p_2^0, \dots, p_n^0)^T$, $q^0 = (q_1^0, q_2^0, \dots, q_n^0)^T$, $p = (p_1, p_2, \dots, p_n)^T$, $q = (q_1, q_2, \dots, q_n)^T$, 价格指数可记作 $I(p, q | p^0, q^0)$. 下面列举经济学家们提出的一些价格指数.

$$I_1(p, q | p^0, q^0) = \frac{p \cdot q^0}{p^0 \cdot q^0} = \frac{\sum_{i=1}^n p_i q_i^0}{\sum_{i=1}^n p_i^0 q_i^0} \quad (3)$$

$$I_2(p, q | p^0, q^0) = \frac{p \cdot q}{p^0 \cdot q} = \frac{\sum_{i=1}^n p_i q_i}{\sum_{i=1}^n p_i^0 q_i} \quad (4)$$

$$I_3(p, q | p^0, q^0) = \frac{p \cdot q}{p^0 \cdot q^0} = \frac{\sum_{i=1}^n p_i q_i}{\sum_{i=1}^n p_i^0 q_i^0} \quad (5)$$

$$I_4(p, q | p^0, q^0) = \frac{p \cdot a}{p^0 \cdot a} = \frac{\sum_{i=1}^n p_i a_i}{\sum_{i=1}^n p_i^0 a_i}, \quad a_i > 0 \quad (6)$$

$$I_5(p, q | p^0, q^0) = \left(\frac{p \cdot q^0}{p^0 \cdot q^0} \frac{p \cdot q}{p^0 \cdot q} \right)^{\frac{1}{2}} \quad (7)$$

$$I_6(p, q | p^0, q^0) = \prod_{i=1}^n \left(\frac{p_i}{p_i^0} \right)^{\alpha_i}, \quad \alpha_i > 0, \quad \sum_{i=1}^n \alpha_i = 1 \quad (8)$$

$$I_7(p, q | p^0, q^0) = \prod_{i=1}^n \left(\frac{p_i}{p_i^0} \right)^{\beta_i}, \quad \beta_i = \frac{q_i^0}{\sum_{i=1}^n q_i^0} \quad (9)$$

$$I_8(p, q | p^0, q^0) = \prod_{i=1}^n \left(\frac{p_i}{p_i^0} \right)^{\gamma_i}, \quad \gamma_i = \frac{q_i}{\sum_{i=1}^n q_i} \quad (10)$$

可以看出, 在 I_1 中权重均用基年的数据, 统计和计算比较简单, 很多国家都用这种指数. 在 I_2 中权重均用现年的数据, 增加了统计和计算量, 但是较确切地反映了价格的变化对当前人民生活的影响, 我国有关部门采用这种指数. I_4 的权重是固定的, 与基年和现年无关. I_5 是 I_1 和 I_2 的几何平均. I_6 与 I_7, I_8 的区别在

于 α_i 为固定常数, 与 q_i^0 或 q_i 无关.

为了进一步比较、评价这些指数, 并研究更合适的价格指数的存在性, 下面从直观上价格指数应满足的性质出发, 引入若干公理.

价格指数的公理化 首先列出大多数人认可的价格指数应具有的性质:

1. 只要有一种商品的价格上涨, 其他商品的价格不下降, 价格指数就应该上升 (对价格的单调性).

2. 若所有商品的价格不变, 价格指数不随权重的改变而改变 (对权重的不变性).

3. 若所有商品的价格均上升 k 倍, 价格指数也上升 k 倍 (对价格的齐次性).

4. 价格指数介于单种商品价格比值的最小值和最大值之间.

5. 价格指数与货币单位的选取无关, 即只要商品的实际价格不变, 仅仅货币单位改变, 价格指数不应改变 (对货币单位的独立性).

6. 价格指数与商品计量单位的选取无关 (对计量单位的独立性, 这里隐含着用商品数量表示权重).

7. 两年的价格指数之比与基年的选取无关.

8. 价格指数不因某种商品被淘汰而失去意义.

这 8 条性质可以用数学语言表述为如下的公理.

对于 $p_i, q_i, p_i^0, q_i^0 > 0 (i = 1, 2, \dots, n)$, 价格指数 $I(p, q | p^0, q^0) > 0$, 应满足以下 8 条公理:

1. 若 $\tilde{p} > p$ (指对所有的 i , $\tilde{p}_i \geq p_i$, 且至少有一个 i , $\tilde{p}_i > p_i$), 则 $I(\tilde{p}, q | p^0, q^0) > I(p, q | p^0, q^0)$.

2. $I(p^0, q | p^0, q^0) = 1$.

3. $I(kp, q | p^0, q^0) = kI(p, q | p^0, q^0), k > 0$.

4. $\min_i \frac{p_i}{p_i^0} \leq I(p, q | p^0, q^0) \leq \max_i \frac{p_i}{p_i^0}$.

5. $I(\lambda p, q | \lambda p^0, q^0) = I(p, q | p^0, q^0), \lambda > 0$.

6. $I(\Lambda p, \Lambda^{-1} q | \Lambda p^0, \Lambda^{-1} q^0) = I(p, q | p^0, q^0)$, 其中 $\Lambda = \text{Diag}[\lambda_1, \lambda_2, \dots, \lambda_n], \lambda_i > 0 (i = 1, 2, \dots, n)$. ①

7. $\frac{I(\tilde{p}, \tilde{q} | p^0, q^0)}{I(p, q | p^0, q^0)} = \frac{I(\tilde{p}, \tilde{q} | \bar{p}^0, \bar{q}^0)}{I(p, q | \bar{p}^0, \bar{q}^0)}$

8. $\lim_{p_i \rightarrow 0} I(p, q | p^0, q^0) > 0$. ②

① 请读者说明为什么这个式子表述了前面的第 6 条性质.

② 用 $p_i \rightarrow 0$ 表示第 i 种商品被淘汰.

用这8条公理——检验(3)~(10)列举的价格指数 $I_1 \sim I_8$,我们发现没有一个价格指数满足所有公理.譬如 I_1, I_2, I_3 不满足公理7, I_6, I_7, I_8 不满足公理8, I_3 不满足公理2, I_4 不满足公理6.那么,是否可以找到其他的、由 p, q, p^0, q^0 决定的价格指数 I 满足所有公理呢?可惜的是Eichhorn证明了如下的定理^[46].

定理 不存在同时满足公理2,3,6,7,8的价格指数 $I(p, q | p^0, q^0)$.

[证明] 其过程是证明任何一个满足公理2,3,6,7的指数 $I(p, q | p^0, q^0)$ 必然不满足公理8.

记 $e = (1, 1, \dots, 1)^T$, $C = \text{Diag}[c_1, c_2, \dots, c_n]$, $D = \text{Diag}[d_1, d_2, \dots, d_n]$, $(c_i, d_i > 0, i = 1, 2, \dots, n)$. 容易推出

$$\begin{aligned} I(CDe, C^{-1}D^{-1}e | e, e) &= \frac{I(CDe, C^{-1}D^{-1}e | e, e)}{I(Ce, C^{-1}e | e, e)} \cdot I(Ce, C^{-1}e | e, e) \\ &\stackrel{\text{公理7}}{=} \frac{I(CDe, C^{-1}D^{-1}e | Ce, C^{-1}e)}{I(Ce, C^{-1}e | Ce, C^{-1}e)} \cdot I(Ce, C^{-1}e | e, e) \\ &\stackrel{\text{公理2,6}}{=} I(De, D^{-1}e | e, e) \cdot I(Ce, C^{-1}e | e, e) \end{aligned} \quad (11)$$

$$\begin{aligned} I(p, e | e, e) &= \frac{I(p, e | e, e)}{I(p, p^{-1} | e, e)} \cdot I(p, p^{-1} | e, e) \\ &\stackrel{\text{公理7}}{=} \frac{I(p, e | p, e)}{I(p, p^{-1} | p, e)} \cdot I(p, p^{-1} | e, e) \stackrel{\text{公理2}}{=} I(p, p^{-1} | e, e) \end{aligned} \quad (12)$$

记 $A_i = \text{Diag}[1, \dots, \lambda, \dots, 1]$ (第 i 对角元为 $\lambda, \lambda > 0$, 其余为1), 于是

$$\prod_{i=1}^n A_i = \lambda E \quad (E \text{ 为单位阵}).$$

设 $I(p, q | p^0, q^0)$ 是满足公理2,3,6,7的价格指数,令

$$s = \prod_{i=1}^n I(A_i e, e | e, e) \quad (13)$$

则

$$\begin{aligned} s &\stackrel{(12)\text{式}}{=} \prod_{i=1}^n I(A_i e, A_i^{-1} e | e, e) \\ &\stackrel{(11)\text{式}}{=} I\left(\prod_{i=1}^n A_i e, \prod_{i=1}^n A_i^{-1} e | e, e\right) \\ &= I(\lambda e, \lambda^{-1} e | e, e) \\ &\stackrel{\text{公理3}}{=} \lambda I(e, \lambda^{-1} e | e, e) \stackrel{\text{公理2}}{=} \lambda \end{aligned} \quad (14)$$

当 $\lambda \rightarrow 0$ 时 $s \rightarrow 0$, 于是在(13)式右端必存在某个因子有

$$\lim_{\lambda \rightarrow 0} I(A_i e, \lambda^{-1} e | e, e) = 0 \quad (15)$$

这个结果与公理8矛盾.

[证毕]

顺便指出,该定理没有涉及公理1,4,5的一个原因是,可以证明(习题24,25):

若指数 I 满足公理 1, 2, 3, 则 I 满足公理 4;

若指数 I 满足公理 2, 3, 7, 则 I 满足公理 5.

对常用的价格指数的分析 既然不存在满足所有公理的价格指数, 我们只好回到前面列举的 $I_1 \sim I_8$ ((3) ~ (10) 式). 首先, I_4 和 I_6 需要构造另外的参数 a_i 和 α_i , 不便应用, I_5 可由 I_1, I_2 直接得到, I_3 不满足公理 2 (这是非常基本的要求), 所以不再对它们作进一步的分析.

对于 I_7, I_8 , 不难验证它们不满足公理 6, 7, 8 (习题 22), 并且计算量较大, 而对于 I_1 和 I_2 , 可以验证它们满足除公理 7 以外的其余公理 (习题 21), 计算也较方便, 所以 I_1 和 I_2 是目前常用的价格指数. 进一步分析它们与公理 7 的矛盾还可以发现, 在一般情况下公理 7 能够近似地成立. 例如对于价格指数 I_2 和两个不同基年的价格向量 \bar{p}^0, \bar{p}^1 , 公理 7 成立的条件是存在正数 k ,

$$\bar{p}^1 = k\bar{p}^0 \quad (16)$$

即对于所有商品一个基年的价格都是另一基年价格的 k 倍. 这个条件虽然不会绝对成立, 但实际上是近似满足的. 于是为人们采用 I_1 和 I_2 作为实用的价格指数提供了更强的论据 (关于 I_1 见习题 23).

习 题

1. 证明 8.1 节层次分析模型中定义的 n 阶一致阵 A 有下列性质:

- (1) A 的秩为 1, 惟一非零特征根为 n ;
- (2) A 的任一列向量都是对应于 n 的特征向量.

2. 对于 n 阶成对比较阵 $A = (a_{ij})$, 设 $a_{ij} = \frac{w_i}{w_j} \varepsilon_{ij}$, $\varepsilon_{ij} = 1 + \delta_{ij}$, 其中 $w = (w_1, \dots, w_n)^T$ 是对应于最大特征根的特征向量, δ_{ij} 表示 a_{ij} 在一致性附近的扰动. 若 δ_{ij} 为方差 σ^2 的随机变量, 证明一致性指标 $CI \approx \sigma^2/2$.

3. 证明 8.1 节中用对数最小二乘法得到的权向量 ((16) 式) 与实用算法中根法的计算结果相同.

4. 用层次分析法解决一两个实际问题, 例如:

(1) 学校评选优秀学生或优秀班级, 试给出若干准则, 构造层次结构模型. 可分为相对评价和绝对评价两种情况讨论.

(2) 你要购置一台个人电脑, 考虑功能、价格等的因素, 如何作出决策.

(3) 为大学毕业的青年建立一个选择志愿的层次结构模型.

(4) 你的家乡准备集资兴办一座小型饲养场, 是养猪, 还是养鸡、养鸭、养兔…….

*5. 为减少层次分析法中的主观成分, 可请若干专家每人构造成对比较阵. 试给出一种由若干个成对比较阵确定权向量的方法.

*6. 用另一种方法构造成对比较阵 $A = (a_{ij})$, a_{ij} 表示因素 C_i 与 C_j 的影响之差, $a_{ji} = -a_{ij}$, 于是 A 为反对称阵, 并且, 当 $a_{ik} + a_{kj} = a_{ij}$ ($i, k, j = 1, 2, \dots, n$) 时 A 是一致阵. 规定权向量 $w =$

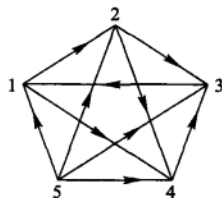
$(w_1, \dots, w_n)^T$ 应满足 $\sum_{i=1}^n w_i = 0$, a_{ij} 可记作 $a_{ij} = (w_i - w_j) + \varepsilon_{ij}$ (对一致阵 $\varepsilon_{ij} = 0$). 试给出一种由 A 确定权向量 w 的方法. 与 1-9 尺度对应, 这里用 0-8 尺度, 即 a_{ij} 取值范围是 0, 1, \dots , 8 及 -1, \dots , -8.

7. 右图是 5 位网球选手循环赛的结果. 作为竞赛图, 它是双向连通的吗? 找出几条完全路径, 用适当方法排出 5 位选手的名次.

8. 排名次的另一方法是考察“失分向量”以代替得分向量 (选手输掉场次的数目为他的失分), 按失分由小到大排列名次.

(1) 证明: 这相当于把竞赛图中各有向边反向后, 按得分向量排列名次, 再把名次倒过来.

(2) 用失分向量方法对 8.2 节图 3(4) 的竞赛图排列名次. 结果与用得分向量方法一致吗?



9. 利用 8.3 节的定理 4,5 说明, 为了把 8.3 节图 1 不稳定的能源利用系统变为冲量和值稳定的, 如果限制只能改变两条有向边的符号, 那么只有 3 种可行方案: 改变 $v_1 v_2$ 和 $v_3 v_5$ (如 8.3 节所述); 改变 $v_2 v_1$ 和 $v_3 v_6$; 改变 $v_2 v_1$ 和 $v_6 v_7$.

10. 考察由野兔 R 和狐狸 F 组成的生态系统, 在野兔的食物资源充足的情况下, 其带符号的有向图如右所示.

(1) 解释图中 +, - 号的意义.

(2) 若初始时段野兔有一增量, 且设 $v(0) = (10, 10)$, 计算 $v(1)$, $v(2)$.

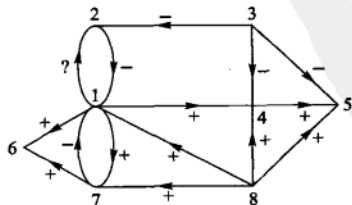


(3) 证明该系统对所有简单冲量过程都是冲量和值稳定的, 从生态意义上进行解释.

(4) 说明若用权描述二者的相互作用, 则在某种加权的情况下系统不再稳定, 并从生态意义上进行解释.

11. 食肉动物 C、食草动物 H 和草 P 组成生态系统, 因为草地有限, 草过密会使草的生长减慢. 用带符号的有向图建立这个系统的冲量过程模型, 并证明冲量过程是不稳定的.

12. 公共汽车系统用带符号的有向图表示如下, 其中只有单位距离票价随着乘客行程的增加应该提高还是降低尚未确定 (图中有向边 1→2 的标以 ? 号). 讨论这个符号应为 + 还是 - 才能使冲量过程稳定.



1—乘客的行程; 2—单位距离票价; 3—节油量; 4—燃料消耗;

5—污染; 6—事故; 7—晚点; 8—居民人数

13. 验证 8.4 节表 3 用 5 种除数法得到的结果.

14. 证明主要分数法 (MF) 给出了 $\min \sum_{i=1}^n p_i \left(\frac{n_i}{p_i} - \frac{N}{P} \right)^2$ 的解.

15. 份额法 QM 可简述如下: 定义第 i 方分配第 $s+1$ 席位“合格”是指 $n_i < q_i = (s+1) \frac{p_i}{P}$, 即不违反份额性的上限, 记 $E(n, s+1) = \{ \text{第 } i \text{ 方分配第 } s+1 \text{ 席位合格}, i=1, 2, \dots, m \}$, 当总席位为 s 时第 i 方的席位分配记作 $n_i = f_i(p, s)$, 且有 $f_i(p, 0) = 0$, 让 s 每次 1 席地递增, 若对于所有 $i \in E(n, s+1)$ 及某个 k 有 $\frac{p_k}{n_k+1} \geq \frac{p_i}{n_i+1}$, 则令 $f_k(p, s+1) = n_k + 1, f_i(p, s+1) = n_i (i \neq k)$.

现有 5 方人口分别为 5 117, 4 400, 162, 161, 160, 试分别用 5 种除数法及 GR 和 QM 分配总共 100 个席位.

份额法不满足人口单调性, 你能举出例子吗?

16. 已知 A, B, C, D 的人口 p_i 如下表第 2 列, 分配 35 个席位, 算出份额 q_i 如下表第 3 列. 验证按照 MF 和 EP 法分配的结果如下表第 4 列, 不满足份额性. 如果从 B 拿出 1 席给 D, 以满足份额性, 验证当 A, B, C, D 的人口由 p_i 变成 p'_i (下表第 5 列) 时, 将违反人口单调性.

	p_i	q_i	MF, EP	p'_i
A	70 653	1. 552	2	86 228
B	117 404	2. 579	3	113 908
C	210 923	4. 633	5	232 778
D	1 194 456	26. 236	25	1 160 522
总和	1 593 436	35	35	1 593 436

17. 证明 8.5 节中当候选人数目 $m=2$ 且选民人数 $n \geq 2$ 时, 简单多数规则满足 Arrow 公理.

18. 举例说明 8.5 节中确定选举结果的记分规则不满足 Arrow 公理 3.

*19. 设 $I = \{1, 2, \dots, n\}$ 和 $A = (x, y, \dots)$ 分别是选民和候选人集合, (p_1, p_2, \dots, p_n) 是 I 对 A 的一次投票 (为简单起见, 不考虑两候选人等同的情况). 选举结果 p 不是对 A 的一个排序, 而只是决定一名优胜者 (第一名). 现举出以下几种选举规则:

- (1) (p_1, p_2, \dots, p_n) 中排在第一名最多的那位候选人为优胜者.
- (2) 若多于或等于半数的选民将 x 排在其他候选人之前, x 是优胜者.
- (3) 若多于或等于半数的选民将 x 排在第一位, 则 x 是优胜者; 若没有这样的 x , 就把排在第一位最多的两个候选人 x, y 进行比较, 当多于或等于半数的选民将 x 排在 y 前面时, x 是优胜者.
- (4) 得分 (Borda 数) 居第一位的为优胜者.

问这些规则都能确定优胜者吗? 对于同一次投票这些规则决定的优胜者相同吗? 你还

能提出一些决定优胜者的选举规则吗?

20. 用最小距离意义下的选举规则研究 8.5 节(3)式给出的投票.

(1) 如果以 $\sum_{i=1}^3 d(p, p_i)$ 最小为原则确定选举结果 p , 说明 p 可以是 p_1, p_2 或 p_3 中任一个.

(2) 如果以 $\sum_{i=1}^3 d^2(p, p_i)$ 最小为原则确定选举结果 p , 说明 $p: x \sim y \sim z$.

21. 在 8.6 节中验证价格指数 I_2 ((4)式) 满足公理 7 以外的其余公理; 而且当条件(16)式成立时它也满足公理 7.

22. 验证价格指数 I_8 (8.6 节(10)式) 满足公理 1~5, 但不满足公理 6~8.

23. 给出价格指数 I_1 (8.6 节(3)式) 满足公理 7 的条件, 并解释该条件的实际含义.

24. 证明: 若价格指数 I 满足公理 1~3, 则满足公理 4.

25. 证明: 若价格指数 I 满足公理 2, 3, 7, 则满足公理 5.

26. 在定义价格指数时如何构造权重 q 和 q^0 是关键之一, 你认为应该如何确定权重. 当然, 你也可以定义不同于 $I_1 \sim I_8$ 的价格指数, 并说明其含义和性质.

鄧平龍
PDG

第9章

概率模型

现实世界的变化受着众多因素的影响,这些因素根据其本身的特性及人们对它们的了解程度,可分为确定的和随机的.虽然我们研究的对象通常包含随机因素,但是如果从建模的背景、目的和手段看,主要因素是确定的,而随机因素可以忽略,或者随机因素的影响可以简单地以平均值的作用出现,那么就能够建立确定性的数学模型,本书前面的部分正是这样的.

如果随机因素对研究对象的影响必须考虑,就应该建立随机性的数学模型,简称**随机模型**.本章通过几个实例讨论,如何用随机变量和概率分布描述随机因素的影响,建立比较简单的随机模型——**概率模型**,其中要用到概率的运算,以及概率分布、期望、方差等基本知识,读者可以从任何一本概率论教科书上找到.

9.1 传送系统的效率

在机械化生产车间里你可以看到这样的情景:排列整齐的工作台旁工人们紧张地生产同一种产品,工作台上有一条传送带在运转,带上设置着若干钩子,工人们将产品挂在经过他上方的钩子上带走,如图1.当生产进入稳定状态后,每个工人生产出一件产品所需时间是不变的,而他要挂产品的时刻却是随机的.衡量这种传送系统的效率可以看它能否及时地把工人们生产的产品带走,显然在工人数目不变的情况下传送带速度越快,带上钩子越多,效率会越高.我们要构造一个衡量传送系统效率的指标,并且在一些简化假设下建立一个模型来描述这个指标与工人数目、钩子数量等参数的关系.

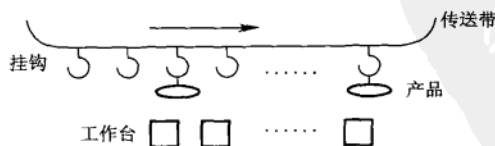


图1 传送系统示意图

模型分析 为了用传送带及时带走的产品数量来表示传送系统的效率,在工人们生产周期(即生产一件产品的时间)相同的情况下,需要假设工人们在生产出一件产品后,要么恰好有空钩子经过他的工作台,使他可以将产品挂上带走,要么没有空钩子经过,迫使他将产品放下并立即投入下一件产品的生产,以保持整个系统周期性地运转。

工人们的生产周期虽然相同,但是由于各种随机因素的干扰,经过相当长时间后,他们生产完一件产品的时刻就不会一致,可以认为是随机的,并且在一个生产周期内任一时刻的可能性是一样的。

由上分析,传送系统长期运转的效率等价于一周期的效率,而一周期的效率可以用它在一周期内能带走的产品数与一周期内生产的全部产品数之比来描述。

为了将问题简化到能用简单的概率方法来解决,我们做出如下的假设。

模型假设

1. 有 n 个工人,他们的生产是相互独立的,生产周期是常数, n 个工作台均匀排列。

2. 生产已进入稳态,即每个工人生产出一件产品的时刻在一周期内是等可能的。

3. 在一周期内有 m 个钩子通过每一工作台上,钩子均匀排列,到达第一个工作台上方的钩子都是空的。

4. 每个工人在任何时刻都能触到一只钩子,也只能触到一只钩子,于是在他生产出一件产品的瞬间,如果能触到的那只钩子是空的,则可将产品挂上带走;如果那只钩子非空(已被他前面的工人挂上了产品),则他只能将这件产品放在地上。而产品一旦放在地上,就永远退出这个传送系统。

模型建立 将传送系统效率定义为一周期内带走的产品数与生产的全部产品数之比,记作 D 。设带走的产品数为 s ,生产的全部产品数显然为 n ,于是 $D = s/n$ 。只需求出 s 就行了。

如果从工人的角度考虑,分析每个工人能将自己的产品挂上钩子的概率,那么这个概率显然与工人所在的位置有关(如第 1 个工人一定可以挂上),这样就使问题复杂化。我们从钩子的角度考虑,在稳态下钩子没有次序,处于同等的地位。若能对一周期内的 m 只钩子求出每只钩子非空(即挂上产品)的概率 p ,则 $s = mp$ 。

得到 p 的步骤如下:(均对一周期内而言)

任一钩子被一名工人触到的概率是 $1/m$;

任一钩子不被一名工人触到的概率是 $1 - 1/m$;

由工人生产的独立性,任一钩子不被所有 n 个工人挂上产品的概率,即任

一只钩子为空钩的概率是 $\left(1 - \frac{1}{m}\right)^n$;

任一钩子非空的概率是 $p = 1 - \left(1 - \frac{1}{m}\right)^n$.

这样, 传送系统效率指标为

$$D = \frac{mp}{n} = \frac{m}{n} \left[1 - \left(1 - \frac{1}{m}\right)^n \right] \quad (1)$$

为了得到比较简单的结果, 在钩子数 m 相对于工人数 n 较大, 即 $\frac{n}{m}$ 较小的情况

下, 将多项式 $\left(1 - \frac{1}{m}\right)^n$ 展开后只取前 3 项, 则有

$$D \approx \frac{m}{n} \left[1 - \left(1 - \frac{n}{m} + \frac{n(n-1)}{2m^2}\right) \right] = 1 - \frac{n-1}{2m} \quad (2)$$

如果将一周期内未带走的产品数与全部产品数之比记作 E , 再假定 $n \gg 1$, 则

$$D = 1 - E, \quad E \approx \frac{n}{2m} \quad (3)$$

当 $n=10, m=40$ 时, (3) 式给出的结果为 $D=87.5\%$, (1) 式得到的精确结果为 $D=89.4\%$.

评注 这个模型是在理想情况下得到的, 它的一些假设, 如生产周期不变, 挂不上钩子的产品退出传送系统等可能是不现实的. 但是模型的意义在于, 一方面利用基本合理的假设将问题简化到能够建模的程度, 并用很简单的方法得到结果; 另一方面所得的简化结果 (3) 式具有非常简明的意义: 指标 $E = 1 - D$ (可理解为相反意义的“效率”) 与 n 成正比, 与 m 成反比. 通常工人数目 n 是固定的, 一周期内通过的钩子数 m 增加 1 倍, 可使“效率” E (未被带走的产品数与全部产品数之比) 降低 1 倍.

另外一种通过增加钩子数目来使“效率”降低的方法请见习题 1.

9.2 报童的诀窍

报童每天清晨从报社购进报纸零售, 晚上将没有卖掉的报纸退回. 设报纸每份的购进价为 b , 零售价为 a , 退回价为 c , 应该自然地假设为 $a > b > c$. 这就是说, 报童售出一份报纸赚 $a - b$, 退回一份赔 $b - c$. 报童每天如果购进的报纸太少, 不够卖的, 会少赚钱; 如果购进太多, 卖不完, 将要赔钱. 请你为报童筹划一下, 他应如何确定每天购进报纸的数量, 以获得最大的收入^[40].

众所周知, 应该根据需求量确定购进量. 需求量是随机的, 假定报童已经通过自己的经验或其他的渠道掌握了需求量的随机规律, 即在他的销售范围内每

天报纸的需求量为 r 份的概率是 $f(r)$ ($r=0,1,2,\dots$). 有了 $f(r)$ 和 a, b, c , 就可以建立关于购进量的优化模型了.

假设每天购进量为 n 份, 因为需求量 r 是随机的, r 可以小于 n 、等于 n 或大于 n , 致使报童每天的收入也是随机的, 所以作为优化模型的目标函数, 不能是报童每天的收入, 而应该是他长期(几个月, 一年)卖报的日平均收入. 从概率论的观点看, 这相当于报童每天收入的期望值, 以下简称平均收入.

记报童每天购进 n 份报纸时的平均收入为 $G(n)$, 如果这天的需求量 $r \leq n$, 则他售出 r 份, 退回 $n-r$ 份; 如果这天的需求量 $r > n$, 则 n 份将全部售出. 考虑到需求量为 r 的概率是 $f(r)$, 所以

$$G(n) = \sum_{r=0}^n [(a-b)r - (b-c)(n-r)]f(r) + \sum_{r=n+1}^{\infty} (a-b)nf(r) \quad (1)$$

问题归结为在 $f(r)$, a, b, c 已知时, 求 n 使 $G(n)$ 最大.

通常需求量 r 的取值和购进量 n 都相当大, 将 r 视为连续变量更便于分析和计算, 这时概率 $f(r)$ 转化为概率密度函数 $p(r)$, (1) 式变成

$$G(n) = \int_0^n [(a-b)r - (b-c)(n-r)]p(r)dr + \int_n^{\infty} (a-b)np(r)dr \quad (2)$$

计算

$$\begin{aligned} \frac{dG}{dn} &= (a-b)np(n) - \int_0^n (b-c)p(r)dr - (a-b)np(n) + \int_n^{\infty} (a-b)p(r)dr \\ &= -(b-c) \int_0^n p(r)dr + (a-b) \int_n^{\infty} p(r)dr \end{aligned}$$

令 $\frac{dG}{dn} = 0$, 得到

$$\frac{\int_0^n p(r)dr}{\int_n^{\infty} p(r)dr} = \frac{a-b}{b-c} \quad (3)$$

使报童日平均收入达到最大的购进量 n 应满足 (3) 式. 因为 $\int_0^{\infty} p(r)dr = 1$, 所以 (3) 式又可表为

$$\int_0^n p(r)dr = \frac{a-b}{a-c} \quad (4)$$

根据需求量的概率密度 $p(r)$ 的图形很容易从 (3) 式确定购进量 n . 在图 1 中用 P_1, P_2 分别表示曲线 $p(r)$ 下的两块面积, 则 (3) 式可记作

$$\frac{P_1}{P_2} = \frac{a-b}{b-c} \quad (5)$$

因为当购进 n 份报纸时, $P_1 = \int_0^n p(r) dr$ 是需求量 r 不超过 n 的概率, 即卖

不完的概率; $P_2 = \int_n^\infty p(r) dr$ 是需求量 r 超过 n

的概率, 即卖完的概率, 所以 (3) 式表明, 购进的份数 n 应该使卖不完与卖完的概率之比, 恰好等于卖出一份赚的钱 $a-b$ 与退回一份赔的钱 $b-c$ 之比. 显然, 当报童与报社签订的合同使报童每份赚钱与赔钱之比越大时, 报童购进的份数就应该越多.

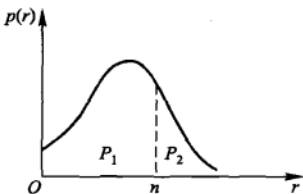


图 1 由 $p(r)$ 确定 n 的图解法

9.3 随机存贮策略

商店在一周中的销售量是随机的. 每逢周末经理要根据存货的多少决定是否订购货物, 以供下周的销售. 适合经理采用的一种简单的策略是制订一个下界 s 和一个上界 S , 当周末存货不少于 s 时就不订货; 当存货少于 s 时则订货, 且订货量使得下周初的存量达到 S . 这种策略称为 (s, S) 随机存贮策略.

为使问题简化起见, 像确定性存贮模型 (3.1 节) 一样, 仍然只考虑费用: 订货费、贮存费、缺货费和商品购进价格, 存贮策略的优劣以总费用为标准. 显然, 总费用 (在平均意义下) 与 (s, S) 策略、销售量的随机规律以及单项费用的大小有关.

模型假设 为了叙述的方便, 时间以周为单位, 商品数量以件为单位.

1. 每次订货费为 c_0 (与数量无关), 每件商品购进价为 c_1 , 每件商品一周的贮存费为 c_2 , 每件商品的缺货损失为 c_3 . c_3 相当于售出价, 所以应有 $c_1 < c_3$.

2. 一周的销售量 r 是随机的. r 的取值很大, 可视为连续变量, 其概率密度函数为 $p(r)$.

3. 记周末的存货量为 x , 订货量为 u , 并且立即到货, 于是周初的存货量为 $x+u$.

4. 一周的销售是集中在周初进行的, 即一周的贮存量为 $x+u-r$, 不随时间改变. 这条假设是为了计算贮存费用的方便, 以后我们将考虑修改它.

建模与求解 按照制订 (s, S) 策略的要求, 当周末存货量 $x \geq s$ 时, 订货量 $u=0$; 当 $x < s$ 时 $u > 0$, 且令 $x+u=S$. 确定 s, S 应以“总费用”最小为标准, 因为销售量 r 的随机性, 贮存量和缺货量也是随机的, 致使一周的贮存费和缺货费也是随机的, 所以目标函数应取一周总费用的期望值, 以下称平均费用.

根据假设条件容易写出平均费用为

$$J(u) = \begin{cases} c_0 + c_1 u + L(x+u), & u > 0 \\ L(x), & u = 0 \end{cases} \quad (1)$$

其中

$$L(x) = c_2 \int_0^x (x-r)p(r)dr + c_3 \int_x^\infty (r-x)p(r)dr \quad (2)$$

我们先在 $u > 0$ 的情况下, 求 u 使 $J(u)$ 达到最小值, 从而确定 S . 为此计算

$$\frac{dJ}{du} = c_1 + c_2 \int_0^{x+u} p(r)dr - c_3 \int_{x+u}^\infty p(r)dr \quad (3)$$

令 $\frac{dJ}{du} = 0$, 记 $x+u = S$, 并注意到 $\int_0^\infty p(r)dr = 1$, 可得

$$\frac{\int_0^S p(r)dr}{\int_S^\infty p(r)dr} = \frac{c_3 - c_1}{c_2 + c_1} \quad (4)$$

这就是说, 令订货量 u 加上原来的存量 x 达到 (4) 式所示的 S , 可使平均费用最小^①.

从 (4) 式可以看出, 当商品购进价 c_1 一定时, 贮存费 c_2 越小, 缺货费 c_3 越大, S 应越大. 这是符合常识的.

下面讨论确定 s 的方法. 当存货量为 x 时, 若订货则由 (1) 式在 S 策略下平均费用为

$$J_1 = c_0 + c_1(S-x) + L(S)$$

若不订货则平均费用为 $J_2 = L(x)$. 显然, 当 $J_2 \leq J_1$ 即

$$L(x) \leq c_0 + c_1(S-x) + L(S) \quad (5)$$

时不应订货. 记

$$I(x) = c_1 x + L(x) \quad (6)$$

则不订货的条件 (5) 式表为

$$I(x) \leq c_0 + I(S) \quad (7)$$

(7) 式右端为已知数. 于是 s 应为方程

$$I(x) = c_0 + I(S) \quad (8)$$

的最小正根.

方程 (8) 可以用图形求解. 注意到 $I(x)$ 与 $J(u)$ 表达式的相似性 (见 (1) 与 (6) 式), 可知 $I(x)$ 是下凸的, 且在 $x = S$ 时达到极小值, 如图 1. 在极小值 $I(S)$ 上叠加 c_0 , 按图中箭头方向即可得到 s .

① 读者容易验证 $\frac{d^2 J}{du^2}$ 恒大于零.

综上所述,根据模型(1),(2)所确定的 (s, S) 策略由(4),(6),(8)式给出,当 c_0, c_1, c_2, c_3 及 $p(r)$ 给定后, s, S 可以惟一地解出.

评注 在这个模型中贮存费用的计算是比较困难的. 因为一般地说贮存费应与贮存时间有关,所以必须对一周内贮存量的变化情况作出适当的假定. 按照模型假设第4条,贮存量 q 在 $0 \leq t \leq 1$ 内的变化可用图2表示(为简单起见设原存量 x 为0),即在可以忽略的短时间内贮存量就降为 $u-r$ ($u > r$ 时)或0 ($u \leq r$ 时). 我们已经看到在这个假定下计算及其结果都十分简单.

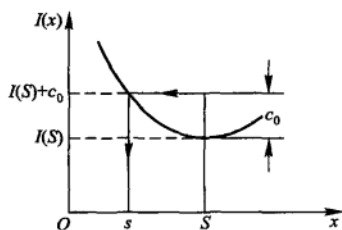


图1 求 s 的图解法

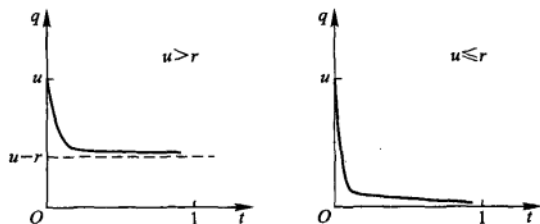


图2 模型假设4的图示

关于贮存量 q 的更合理的假定似乎应该如图3所示,即一周内的销售是均匀的,因而贮存量 q 呈直线下降. 在这种情况下贮存费的计算就比较麻烦了,而且得不到简洁的结果(习题3).

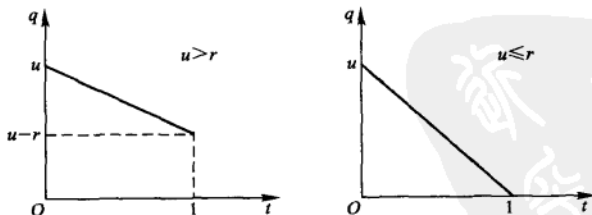


图3 模型假设4的改进

这个模型只考虑一周的存贮与需求,称单时段随机存贮策略,我们将在13.6节讨论多时段的情形.

9.4 轧钢中的浪费

你到过轧钢厂吗?把粗大的钢坯变成合格的钢材(如钢筋、钢板)通常要经过两道工序,第一道是粗轧(热轧),形成钢材的雏形;第二道是精轧(冷轧),得到规定长度的成品材。粗轧时由于设备、环境等方面众多因素的影响,得到的钢材的长度是随机的,大体上呈正态分布,其均值可以在轧制过程中由轧机调整,而均方差则是由设备的精度决定的,不能随意改变。如果粗轧后的钢材长度大于规定长度,精轧时把多出的部分切掉,造成浪费^①;如果粗轧后的钢材已经比规定长度短,则整根报废,造成更大的浪费。显然,应该综合考虑这两种情况,使得总的浪费最少^[10]。

上面的问题可叙述为:已知成品材的规定长度 l 和粗轧后钢材长度的均方差 σ ,确定粗轧后钢材长度的均值 m ,使得当轧机调整到 m 进行粗轧,再通过精轧以得到成品材时总的浪费最少。

问题分析 粗轧后钢材长度记作 x , x 是均值 m 、均方差 σ 的正态随机变量, x 的概率密度记作 $p(x)$,如图 1 所示,其中 σ 已知, m 待定。当成品材的规定长度 l 给定后,记 $x \geq l$ 的概率为 P ,即 $P = P(x \geq l)$, P 是图中阴影部分面积。

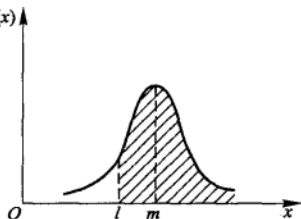


图 1 钢材长度 x 的概率密度

轧制过程中的浪费由两部分构成。一是当 $x \geq l$ 时,精轧时要切掉长 $x - l$ 的钢材;二是当 $x < l$ 时,长 x 的整根钢材报废。由图可以看出, m 变大时曲线右移,概率 P 增加,第一部分的浪费随之增加,而第二部分的浪费将减少;反之,当 m 变小时曲线左移,虽然被切掉的部分减少了,但是整根报废的可能将增加。于是必然存在一个最佳的 m ,使得两部分的浪费综合起来最小。

这是一个优化模型,建模的关键是选择合适的目标函数,并用已知的和待确定的量 l, σ, m 把目标函数表示出来。一种很自然的想法是直接写出上面分析的两部分浪费,以二者之和作为目标函数,于是容易得到总的浪费长度为^②

$$W = \int_l^{\infty} (x - l)p(x)dx + \int_{-\infty}^l xp(x)dx \quad (1)$$

① 精轧设备的精度很高,轧出的成品材可以认为是完全符合规定长度要求的。

② 实际上,钢材长度 x 不可能取负值,但是因为通常 $l, m \gg \sigma$, x 取负值的概率极小,(1)式中积分下限取 $-\infty$ 是为了下面表示和计算的方便。

利用 $\int_{-\infty}^{\infty} p(x) dx = 1$, $\int_{-\infty}^{\infty} xp(x) dx = m$ 和 $\int_l^{\infty} p(x) dx = P$, (1)式可化简为

$$W = m - lP \quad (2)$$

其实, (2)式可以用更直接的办法得到. 设想共粗轧了 N 根钢材 (N 很大), 所用钢材总长为 mN , N 根中可以轧出成品材的只有 PN 根, 成品材总长为 lPN , 于是浪费的总长度为 $mN - lPN$, 平均每粗轧一根钢材浪费长度为

$$W = \frac{mN - lPN}{N} = m - lP \quad (3)$$

问题在于以 W 为目标函数是否合适呢?

轧钢的最终产品是成品材, 如果粗轧车间追求的是效益而不是产量的话, 那么浪费的多少不应以每粗轧一根钢材的平均浪费量为标准, 而应该用每得到一根成品材浪费的平均长度来衡量. 为了将目标函数从前者 (即 (3) 式所表示的) 改成后者, 只需将 (3) 式中的分母 N 改为成品材总数 PN 即可.

建模与求解 以每得到一根成品材所浪费钢材的平均长度为目标函数. 因为当粗轧 N 根钢材时浪费的总长度是 $mN - lPN$, 而只得到 PN 根成品材, 所以目标函数为

$$J_1 = \frac{mN - lPN}{PN} = \frac{m}{P} - l \quad (4)$$

因为 l 是已知常数, 所以目标函数可等价地只取上式右端第一项, 记作

$$J(m) = \frac{m}{P(m)} \quad (5)$$

式中 $P(m)$ 表示概率 P 是 m 的函数. 实际上, $J(m)$ 恰是平均每得到一根成品材所需钢材的长度.

下面求 m 使 $J(m)$ 达到最小. 对于表达式

$$P(m) = \int_l^{\infty} p(x) dx, \quad p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad (6)$$

作变量代换

$$y = \frac{x - m}{\sigma} \quad (7)$$

并令

$$\mu = \frac{m}{\sigma}, \quad \lambda = \frac{l}{\sigma} \quad (8)$$

则 (5) 式可表为

$$J(\mu) = \frac{\sigma\mu}{\Phi(\lambda - \mu)} \quad (9)$$

其中 $\Phi(z)$ 是标准正态变量的分布函数, 即^①

$$\Phi(z) = \int_z^{\infty} \varphi(y) dy, \quad \varphi(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \quad (10)$$

$\varphi(y)$ 是标准正态变量的密度函数. 再设

$$z = \lambda - \mu \quad (11)$$

(9) 式化为用微分法解函数

$$J(z) = \frac{\sigma(\lambda - z)}{\Phi(z)} \quad (12)$$

的极值问题. 注意到 $\Phi'(z) = -\varphi(z)$, 不难推出最优值 z^* 应满足方程

$$\frac{\Phi(z)}{\varphi(z)} = \lambda - z \quad (13)$$

记

$$F(z) = \frac{\Phi(z)}{\varphi(z)} \quad (14)$$

$F(z)$ 可根据标准正态分布的函数值 Φ 和 φ 制成表格 (简表见表 1) 或绘出图形 (略图如图 2). 由表或图可以得到方程 (13) 的根 z^* , 再代回 (11) 和 (8) 式即得到 m 的最优值 m^* .

表 1 $F(z) = \Phi(z)/\varphi(z)$ 简表

z	-3.0	-2.5	-2.0	-1.5	-1.0	-0.5
$F(z)$	227.0	56.79	18.10	7.206	3.477	1.680
z	0	0.5	1.0	1.5	2.0	2.5
$F(z)$	1.253	0.876	0.656	0.516	0.420	0.355

值得指出的是, 对于给定的 $\lambda > F(0) = 1.253^{\text{②}}$, 方程 (13) 不只一个根, 但是可以证明, 只有唯一负根 $z^* < 0$, 才使 $J(z)$ 取得极小值 (习题 6).

试看下面的例子, 设要轧制长 $l = 2.0$ m 的成品钢材, 由粗轧设备等因素决定的粗轧冷却后钢材长度的均方差 $\sigma = 20$ cm, 问这时钢材长度的均值 m 应调整到多少才使浪费最少.

以 (5) 式给出的 $J(m)$ 为目标函数, 由 (8) 式

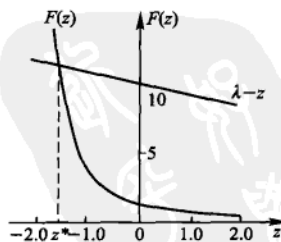


图 2 $F(z)$ 的图形
及 (13) 的图解法

① 通常, 分布函数定义为 $\Phi_1(z) = \int_{-\infty}^z \varphi(y) dy$, 这时有 $\Phi(z) = 1 - \Phi_1(z)$.

② 通常 $l \gg \sigma$, 故对于 $\lambda = l/\sigma$, 条件 $\lambda > F(0) = 1.253$ 是容易满足的.

算出 $\lambda = l/\sigma = 10$, 解出方程(13)的负根为 $z^* = -1.78$ (需要用更精细的 $F(z)$ 表), 如图2所示. 由(11)和(8)式算得 $\mu^* = 11.78$, $m^* = 2.36$, 即最佳的均值应调整为 2.36 m. 还可以算出 $P(m^*) = 0.9625$, 按照(4)式每得到一根成品材浪费钢材的平均长度为 $J_1 = \frac{m^*}{P(m^*)} - l = 0.45$ m. 为了减小这个相当可观的数字, 应该设法提高粗轧设备的精度, 即减小 σ (习题7).

评注 模型中假定当粗轧后钢材长度 x 小于规定长度 l 时就整根报废, 实际上这种钢材还常常能轧成较小规格如长 l_1 ($< l$) 的成品材. 只有当 $x < l_1$ 时才报废. 或者当 $x < l$ 时可以降级使用 (对浪费打一折扣). 这些情况下的模型及求解就比较复杂了 (习题8).

在日常生产活动中类似的问题很多, 如用包装机将某种物品包装成 500 g 一袋出售, 在众多因素的影响下包装封口后一袋的重量是随机的, 不妨仍认为服从正态分布, 均方差已知, 而均值可以在包装时调整. 出厂检验时精确地称量每袋的重量, 多于 500 g 的仍按 500 g 一袋出售, 厂方吃亏; 不足 500 g 的降价处理, 或打开封口返工, 或直接报废, 将给厂方造成更大的损失. 问如何调整包装时每袋重量的均值使厂方损失最小. 生活中类似的现象也很多, 常常难以完全用数量描述, 如你从家中出发去车站赶火车, 由于途中各种因素的干扰, 到达车站的时间是随机的. 到达太早白白浪费时间, 到达晚了则赶不上火车, 损失巨大. 你如何权衡两方面的影响来决定出发时间呢?

9.5 随机人口模型

我们在5.6节讨论的人口模型是确定性的, 已知初始人口并且给定了生育率、死亡率等数据后, 可以确切地预测未来的人口. 但是事实上, 一个人的出生和死亡应该说是随机事件, 无法准确预测. 之所以能用确定性模型描述人口的发展, 是因为考察的是一个国家或地区的数量很大的人口, 用对总数而言的平均生育率、死亡率代替出生、死亡的概率, 将人口作为连续变量处理. 如果研究对象是一个自然村落或一个家族的人口, 数量不大, 需作为离散随机变量看待时, 就要利用随机人口模型来描述其变化过程了^[52, 60].

时刻 t 的人口用随机变量 $X(t)$ 表示, $X(t)$ 只取整数值. 记 $P_n(t)$ 为 $X(t) = n$ 的概率, $n = 0, 1, 2, \dots$. 下面要在对出生和死亡的概率作出适当假设的基础上, 寻求 $P_n(t)$ 的变化规律, 并由此得出人口 $X(t)$ 的期望和方差, 用它们在随机意义下描述人口的发展状况.

模型假设 若 $X(t) = n$, 对人口在 t 到 $t + \Delta t$ 的出生和死亡作如下假设 (Δt 很小):

1. 出生一人的概率与 Δt 成正比, 记作 $b_n \Delta t$; 出生二人及二人以上的概率为 $o(\Delta t)$ ①.

2. 死亡一人的概率与 Δt 成正比, 记作 $d_n \Delta t$; 死亡二人及二人以上的概率为 $o(\Delta t)$.

3. 出生与死亡是相互独立的随机事件.

4. 进一步设 b_n 和 d_n 均与 n 成正比, 记 $b_n = \lambda n, d_n = \mu n, \lambda$ 和 μ 分别是单位时间内 $n=1$ 时一个人出生和死亡的概率.

建模与求解 为了得到 $P_n(t)$ 的方程, 考察随机事件 $X(t + \Delta t) = n$. 将它分解为以下一些互不相容的事件之和, 并且根据假设 1~3, 可以得到这些事件的概率:

1. $X(t) = n-1$, 且 Δt 内出生 1 人, 概率为 $P_{n-1}(t)b_{n-1}\Delta t$;
2. $X(t) = n+1$, 且 Δt 内死亡 1 人, 概率为 $P_{n+1}(t)d_{n+1}\Delta t$;
3. $X(t) = n$, 且 Δt 内没有人出生或死亡, 概率为 $P_n(t)(1 - b_n\Delta t - d_n\Delta t)$;
4. $X(t) = n-k (k \geq 2)$, Δt 内出生 k 人, 或 $X(t) = n+k (k \geq 2)$, Δt 内死亡 k 人, 或 $X(t) = n$, Δt 内出生且死亡 k 人 ($k \geq 1$), 这些事件的概率均为 $o(\Delta t)$.

按照全概率公式有

$$P_n(t + \Delta t) = P_{n-1}(t)b_{n-1}\Delta t + P_{n+1}(t)d_{n+1}\Delta t + P_n(t)(1 - b_n\Delta t - d_n\Delta t) + o(\Delta t) \quad (1)$$

由此可得关于 $P_n(t)$ 的微分方程

$$\frac{dP_n}{dt} = b_{n-1}P_{n-1}(t) + d_{n+1}P_{n+1}(t) - (b_n + d_n)P_n(t) \quad (2)$$

特别地, 在假设 4 下方程为

$$\frac{dP_n}{dt} = \lambda(n-1)P_{n-1}(t) + \mu(n+1)P_{n+1}(t) - (\lambda + \mu)nP_n(t) \quad (3)$$

若初始时刻 ($t=0$) 人口为确定数量 n_0 , 则 $P_n(t)$ 的初始条件为

$$P_n(0) = \begin{cases} 1, & n = n_0 \\ 0, & n \neq n_0 \end{cases} \quad (4)$$

(3) 式对于不同的 n 是一组递推方程, 在条件 (4) 下的求解过程非常复杂, 并且没有简单的结果. 幸而, 通常人们对 (3) 式的解 $P_n(t)$ 并不关心, 感兴趣的只是 $X(t)$ 的期望 $E(X(t))$ (以下简记作 $E(t)$) 和方差 $D(t)$, 而它们可以由 (3), (4) 直接得到. 因为按照期望的定义,

$$E(t) = \sum_{n=1}^{\infty} nP_n(t) \quad (5)$$

① $o(\Delta t)$ 表示 Δt 的高阶小量.

对(5)求导并将(3)代入得

$$\frac{dE}{dt} = \lambda \sum_{n=1}^{\infty} n(n-1)P_{n-1}(t) + \mu \sum_{n=1}^{\infty} n(n+1)P_{n+1}(t) - (\lambda + \mu) \sum_{n=1}^{\infty} n^2 P_n(t) \quad (6)$$

注意到

$$\begin{aligned} \sum_{n=1}^{\infty} n(n-1)P_{n-1}(t) &= \sum_{k=1}^{\infty} k(k+1)P_k(t) \\ \sum_{n=1}^{\infty} n(n+1)P_{n+1}(t) &= \sum_{k=1}^{\infty} k(k-1)P_k(t) \end{aligned}$$

代入(6)式并利用(5)式,则有

$$\frac{dE}{dt} = (\lambda - \mu) \sum_{n=1}^{\infty} nP_n(t) = (\lambda - \mu)E(t) \quad (7)$$

由(4)可以写出 $E(t)$ 的初始条件

$$E(0) = n_0 \quad (8)$$

显然,方程(7)在(8)下的解为

$$E(t) = n_0 e^{rt}, \quad r = \lambda - \mu \quad (9)$$

这个结果与 5.6 节(3)式表示的指数模型.

$$x(t) = x_0 e^{rt} \quad (10)$$

形式上完全一致.从含义上看,随机性模型(9)中出生概率 λ 与死亡概率 μ 之差 r 可称为净增长概率,人口的期望值 $E(t)$ 呈指数增长.在人口数量很多的情况下如果将 r 视为平均意义上的净增长率,那么 $E(t)$ 就可以看成确定性模型(10)中的人口总数 $x(t)$ 了.

对于方差 $D(t)$,按照定义

$$D(t) = \sum_{n=1}^{\infty} n^2 P_n(t) - E^2(t) \quad (11)$$

用类似求 $E(t)$ 的方法可以推出(习题 9)

$$D(t) = n_0 \frac{\lambda + \mu}{\lambda - \mu} e^{(\lambda - \mu)t} [e^{(\lambda - \mu)t} - 1] \quad (12)$$

$D(t)$ 的大小表示了人口 $X(t)$ 在期望值 $E(t)$ 附近的波动范围.(12)式说明这个范围不仅随着时间的延续和净增长概率 $r = \lambda - \mu$ 的增加而变大,而且即使当 r 不变时,它也随着 λ 和 μ 的上升而增长.这就是说,当出生和死亡频繁出现时,人口的波动范围变大.

评注 从模型假设和得到的人口期望值的结果可以看出,这个随机模型在确定性人口模型中相对应的,只不过是简单的指数增长模型.但是即使这样,由方程(3),(4)求解 $P_n(t)$ 已经相当复杂了.读者可以建立与确定性的阻滞增长(logistic)模型相对应的随机模型(习题 10).由这个模型,不仅 $P_n(t)$ 而且 $E(X(t))$ 都难以求出,甚至不知道 $E(X(t))$ 是否与确定性阻滞增长模型的结果

(5.6节(7)式)一致. 所以本节的讨论作为人口模型并没有多大的意义,但是作为一般的生灭过程,特别是从假设1~3得到的模型(2)式有着广泛的用途,如电梯的升降、交通路口的通过以及各种排队现象,都可以在适当的假设下用生灭过程的模型描述.

9.6 航空公司的预订票策略

在激烈的市场竞争中,航空公司为争取更多的客源而开展的一个优质服务项目是预订票业务. 公司承诺,预先订购机票的乘客如果未能按时前来登机,可以乘坐下一班机或退票,无需附加任何费用. 当然也可以订票时只订座,登机时才付款,这两种办法对于下面的讨论是等价的.

开展预订票业务时,对于一次航班,若公司限制预订票的数量恰好等于飞机的容量,那么由于总会有一些订了机票的乘客不按时前来登机,致使飞机因不满员飞行而利润降低,甚至亏本. 而如果不限制预订票数量,那么当持票按时前来登机的乘客超过飞机容量时,必然会引起那些不能飞走的乘客的抱怨,公司不管以什么方式补救,也会导致声誉受损和一定的经济损失,如客源减少、挤掉以后班机的乘客、公司无偿供应食宿、付给一定的赔偿金等. 所以航空公司需要综合考虑经济利益和社会声誉,确定预订票数量的最佳限额^[38].

问题分析 公司的经济利益可以用机票收入扣除飞行费用和赔偿金后的利润来衡量,社会声誉可以用持票按时前来登机、但因满员不能飞走的乘客(以下称被挤掉者)限制在一定数量为标准,注意到这个问题的关键因素——预订票的乘客是否按时前来登机——是随机的,所以经济利益和社会声誉两个指标都应该在平均意义下衡量,这是个两目标的优化问题,决策变量是预订票数量的限额.

模型假设

1. 飞机容量为常数 n , 机票价格为常数 g , 飞行费用为常数 r , r 与乘客数量无关(实际上关系很小), 机票价格按照 $g = r/\lambda n$ 来制订, 其中 λ (< 1) 是利润调节因子, 如 $\lambda = 0.6$ 表示飞机 60% 满员率就不亏本.
2. 预订票数量的限额为常数 m ($> n$), 每位乘客不按时前来登机的概率为 p , 各位乘客是否按时前来登机是相互独立的, 这适合于单独行动的商人、游客等.
3. 每位被挤掉者获得的赔偿金为常数 b .

模型建立

1. 公司的经济利益可以用平均利润 S 来衡量, 每次航班的利润 s 为从机票收入中减去飞行费用和可能发生的赔偿金. 当 m 位乘客中有 k 位不按时前来登

机时

$$s = \begin{cases} (m-k)g-r, & m-k \leq n \\ ng-r-(m-k-n)b, & m-k > n \end{cases} \quad (1)$$

由假设 2, 不按时前来登机的乘客数 k 服从二项分布, 于是概率

$$p_k = P(K=k) = C_m^k p^k q^{m-k}, \quad q = 1-p \quad (2)$$

平均利润 S (即 s 的期望) 为

$$S(m) = \sum_{k=0}^{m-n-1} [ng-r-(m-k-n)b]p_k + \sum_{k=m-n}^m [(m-k)g-r]p_k \quad (3)$$

化简(3)式, 并注意到 $\sum_{k=0}^m kp_k = mp$, 可得

$$S(m) = qmg-r-(g+b) \sum_{k=0}^{m-n-1} (m-k-n)p_k \quad (4)$$

当 n, g, r, p 给定后, 可以求 m 使 $S(m)$ 最大.

2. 公司从社会声誉和经济利益两方面考虑, 应该要求被挤掉的乘客不要太多, 而由于被挤掉者的数量是随机的, 可以用被挤掉的乘客数超过若干人的概率作为度量指标. 记被挤掉的乘客数超过 j 人的概率为 $P_j(m)$, 因为被挤掉的乘客数超过 j 人, 等价于 m 位预订票的乘客中不按时前来登机的不超过 $m-n-j-1$ 人, 所以

$$P_j(m) = \sum_{k=0}^{m-n-j-1} p_k \quad (5)$$

对于给定的 n, j , 显然当 $m = n+j$ 时被挤掉的乘客不会超过 j 人, 即 $P_j(m) = 0$. 而当 m 变大时 $P_j(m)$ 单调增加.

综上, $S(m)$ 和 $P_j(m)$ 虽然是这个优化问题的两个目标, 但是可以将 $P_j(m)$ 不超过某给定值作为约束条件, 以 $S(m)$ 为单目标函数来求解.

模型求解 为了减少 $S(m)$ 中的参数, 取 $S(m)$ 除以飞行费用 r 为新的目标函数 $J(m)$, 其含义是单位费用获得的平均利润, 注意到假设 1 中有 $g = r/\lambda n$, 由(4)式可得

$$J(m) = \frac{S(m)}{r} = \frac{1}{\lambda n} \left[qm - \left(1 + \frac{b}{g} \right) \sum_{k=0}^{m-n-1} (m-k-n)p_k \right] - 1 \quad (6)$$

其中 b/g 是赔偿金占机票价格的比例. 问题化为给定 $\lambda, n, p, b/g$, 求 m 使 $J(m)$ 最大, 而约束条件为

$$P_j(m) = \sum_{k=0}^{m-n-j-1} p_k \leq \alpha \quad (7)$$

其中 α 是小于 1 的正数.

模型(6), (7)无法解析地求解, 我们设定几组数据, 用 MATLAB 软件作数值计算, 结果如下.

设 $n=300$, $\lambda=0.6$, $p=0.05$ 和 0.1 , $b/g=0.2$ 和 0.4 , 计算 $J(m)$, $P_5(m)$, $P_{10}(m)$, 得表 1.

表 1 $n=300$ 时的计算结果

m	$p=0.05$				$p=0.1$			
	J		P_5	P_{10}	J		P_5	P_{10}
	$b/g=0.2$	$b/g=0.4$			$b/g=0.2$	$b/g=0.4$		
300	0.583 3	0.583 3	0	0	0.500 0	0.500 0	0	0
302	0.593 9	0.593 9	0	0	0.510 0	0.510 0	0	0
304	0.604 4	0.604 4	0	0	0.520 0	0.520 0	0	0
306	0.615 0	0.615 0	0.000 0	0	0.530 0	0.530 0	0.000 0	0
308	0.625 4	0.625 4	0.000 0	0	0.540 0	0.540 0	0.000 0	0
310	0.635 3	0.635 1	0.000 7	0	0.550 0	0.550 0	0.000 0	0
312	0.643 9	0.643 4	0.006 6	0.000 0	0.560 0	0.560 0	0.000 0	0.000 0
314	0.650 3	0.649 2	0.034 1	0.000 2	0.570 0	0.570 0	0.000 0	0.000 0
316	0.654 0	0.651 7	0.112 3	0.002 3	0.580 0	0.580 0	0.000 0	0.000 0
318	0.655 1	0.651 2	0.261 2	0.016 0	0.589 9	0.589 9	0.000 1	0.000 0
320	0.654 3	0.648 5	0.463 0	0.065 0	0.599 8	0.599 7	0.000 6	0.000 0
322	0.652 3	0.644 5	0.666 6	0.178 0	0.609 3	0.609 1	0.002 7	0.000 0
324	0.649 9	0.639 8	0.825 0	0.358 3	0.618 1	0.617 8	0.009 7	0.000 2
326	0.647 2	0.635 0	0.922 4	0.568 1	0.625 8	0.625 2	0.028 7	0.001 3
328	0.644 4	0.630 0	0.970 8	0.753 3	0.632 0	0.630 7	0.069 9	0.005 2
330	0.641 7	0.625 0	0.990 7	0.881 0	0.636 2	0.633 9	0.143 9	0.017 1
332					0.638 4	0.634 8	0.254 8	0.045 8
334					0.638 8	0.633 6	0.395 6	0.102 4
336					0.637 7	0.630 6	0.548 4	0.194 9
338					0.635 6	0.626 5	0.691 7	0.322 4
340					0.632 9	0.621 7	0.808 6	0.471 9
342					0.629 8	0.616 5	0.892 3	0.622 5
344					0.626 6	0.611 0	0.945 1	0.754 2

设 $n = 150$, 其他不变, 得表 2.

表 2 $n = 150$ 时的计算结果

m	$p = 0.05$				$p = 0.1$			
	J		P_5	P_{10}	J		P_5	P_{10}
	$b/g = 0.2$	$b/g = 0.4$			$b/g = 0.2$	$b/g = 0.4$		
150	0.583 3	0.583 3	0	0	0.500 0	0.500 0	0	0
152	0.604 4	0.604 4	0	0	0.520 0	0.520 0	0	0
154	0.624 5	0.624 4	0	0	0.540 0	0.540 0	0	0
156	0.640 8	0.639 9	0.000 5	0	0.560 0	0.560 0	0.000 0	0
158	0.650 0	0.647 0	0.018 2	0	0.579 7	0.579 7	0.000 0	0
160	0.651 9	0.645 7	0.125 6	0	0.598 5	0.598 2	0.000 5	0
162	0.649 0	0.638 9	0.372 9	0.004 1	0.614 8	0.613 9	0.005 6	0.000 0
164	0.644 3	0.629 8	0.663 8	0.054 8	0.626 7	0.624 5	0.030 7	0.000 1
166	0.638 9	0.620 0	0.867 8	0.234 4	0.633 0	0.628 5	0.106 0	0.001 9
168	0.633 3	0.610 0	0.961 5	0.522 8	0.634 0	0.626 3	0.254 5	0.014 0
170	0.627 8	0.600 0	0.991 5	0.780 9	0.631 1	0.619 6	0.460 2	0.060 0
172					0.625 9	0.610 3	0.669 4	0.170 9
174					0.619 8	0.599 8	0.830 8	0.353 0
176					0.613 3	0.588 8	0.927 9	0.568 2

结果分析

1. 对于所取的各个 $n, p, b/g$, 平均利润 $J(m)$ 随着 m 的变大都是先增加再减少, 但是在最大值附近变化很小, 而被挤掉的乘客数超过 5 人和 10 人的概率 $P_5(m)$ 和 $P_{10}(m)$ 增加得相当快, 所以应该参考 $J(m)$ 的最大值, 给定约束条件 (7) 式中可以接受的 α , 确定合适的 m .

2. 对于一定的 n, p , 当 b/g 由 0.2 增加到 0.4 时, $J(m)$ 的减少不超过 2%, 所以不妨付给被挤掉的乘客以较多的赔偿金, 赢得社会声誉.

3. 综合考虑经济效益和社会声誉, 可给定 $P_5(m) < 0.2, P_{10}(m) < 0.05$, 由表 1、表 2 知, 对于 $n = 300$, 若估计 $p = 0.05$, 取 $m = 316$; 若估计 $p = 0.1$, 取 $m = 330$. 对于 $n = 150$, 若估计 $p = 0.05$, 取 $m = 160$; 若估计 $p = 0.1$, 取 $m = 166$: (表 1, 2 中黑体数字)

预订票策略的改进 考虑不同客源的实际需要, 如商业界、文艺界人士喜欢上述这种无约束的预订票业务, 他们宁愿接受较高的票价, 而不按时前来登机的可能性较大; 游客及准时上下班的雇员, 会愿意以不能按时前来登机则机票失效

为代价,换取较低额的票价.航空公司为降低风险,可以把上述第2类乘客作为基本客源,对他们降低票价,但购票时即付款,不按时前来登机则机票作废.

设预订票数量 m 中有 t 张是专门预售给第2类乘客的,其折扣票价为 $\beta g (\beta < 1)$, 当 $m-t$ 位第1类乘客中有 k 位不按时前来登机时每次航班的利润 s 为

$$s = \begin{cases} t\beta g + (m-t-k)g - r, & m-k \leq n \\ t\beta g + (n-t)g - r - (m-k-n)b, & m-k > n \end{cases} \quad (8)$$

k 位乘客不按时前来登机的概率为

$$p_k = C_{m-t}^k p^k q^{m-t-k}, \quad q = 1 - p \quad (9)$$

平均利润 S 为

$$\begin{aligned} S(m) &= \sum_{k=0}^{m-n-1} [t\beta g + (n-t)g - r - (m-k-n)b] p_k + \\ &\quad \sum_{k=m-n}^{m-t} [t\beta g + (m-t-k)g - r] p_k \\ &= qmg - r - (g+b) \sum_{k=1}^{m-n-1} (m-k-n)p_k - (q-\beta)tg \end{aligned} \quad (10)$$

正常票价 g , 折扣票价 βg , 利润调节因子 λ 与飞行费用 r 间的关系为

$$\lambda [t\beta g + (n-t)g] = r$$

于是,单位费用获得的平均利润为

$$J(m) = \frac{1}{\lambda [n - (1-\beta)t]} \left[qm - (q-\beta)t - \left(1 + \frac{b}{g} \right) \sum_{k=0}^{m-n-1} (m-k-n)p_k \right] - 1 \quad (11)$$

约束条件——被挤掉的乘客数超过 j 人的概率 $P_j(m)$ 不变((7)式).

取 $\beta = 0.75, t = 50, 100, 150$, 其他参数同上, 计算结果表明, 当 t 增加时 $J(m)$ 和 $P_j(m)$ 均有所减少.

评注 本节在基本合理的假设下对一个两目标的优化问题作了简化处理, 即使这样, 得到的模型也无法解析地求解, 幸而, 数值计算的结果已满足我们对问题进行分析的需要.

与航空公司的预订票策略相类似的事情在日常商务活动中并不少见, 旅馆、汽车出租公司(指将汽车租给顾客使用)等为争夺顾客也可以如此处理.

9.7 学生作弊现象的调查和估计

统计调查中可能会遇到一些因涉及个人隐私或利害关系, 而不受被调查对象欢迎或感到尴尬的所谓敏感问题, 例如是否有考试作弊、赌博、偷税漏税、酒后

驾车等. 这时即使做无记名的直接调查, 也很难消除被调查者的顾虑, 极有可能拒绝应答或故意做出错误的回答, 从而难以保证数据的真实性, 使得调查的结果存在很大的误差, 成为统计调查中的难题. 本节以对学生考试作弊现象的调查和估计为例, 介绍这类敏感问题的调查方法.

问题及分析 学生在考试中的作弊行为是一个严重困扰学校的学风问题, 为了对这种现象的严重程度有一个定量的认识, 需要通过调查来估计有过作弊行为的学生到底占多大的比例. 作弊行为是不光彩的, 很难在学生中做直接调查以得到比较可靠的数据, 因此需要设计合理的调查方案, 来提高应答率并降低不真实应答率.

调查方案设计的基本思想是, 让被调查者从包含是否作过弊的若干问题中, 随机地选答其中一个, 同时让调查者也不知道被调查者回答的是哪一个问题, 从而保护被调查者的隐私, 消除他们的顾虑, 能够对自己所选的问题真实作答. 这种“随机化选答”技术最早是由美国统计学家 Warner 在 1965 年提出的, 目前已经成为敏感问题调查中的常用方法.

下面以 400 名学生对作弊问题的问卷调查结果为依据, 通过 3 个数学模型估计考试中有过作弊行为学生的比例, 讨论估计的精度, 并以此对这些模型作分析、比较.

Warner 模型 Warner 在 1965 年提出了正反问题选答的方法^[71], 要调查的问题在问卷上以正、反两种形式叙述, 比如在学生作弊的调查, 设计以下两个相反的问题供学生选答其中一个:

问题 A: 你在考试中作过弊吗?

问题 B: 你在考试中没有作过弊吗?

对选择的问题只需回答“是”或者“否”. 调查者准备一套 13 张同一花色(如红心)的扑克, 在选答上述问题前, 要求被调查的学生随机抽取一张, 看后还原, 并使调查者不能知道抽取情况. 约定如下: 如果学生抽取的是不超过红心 10 的数(红心 A 看作 1)则回答问题 A, 如果抽取的是红心 J, Q 或 K, 则回答问题 B.

采用这种方式调查, 被调查者回答的是哪个问题不会为其他人(包括调查者)所知, 可以很大地消除被调查者的顾虑, 所以假设一旦被调查者选定了应回答的问题, 他将真实作答.

假定调查结果是收回 $n = 400$ 张有效答卷, 其中有 $n_1 = 112$ 个学生回答“是”, $n - n_1 = 288$ 个学生回答“否”, 即对问题 A, B 两题选答“是”的概率 π 的估计值为 $\hat{\pi} = n_1/n = 7/25$.

我们要估计的是有过作弊行为学生的比例, 可以看作一个被调查学生作过弊的概率, 即对问题 A 回答“是”或对问题 B 回答“否”的概率, 这两个概率相同, 为统一起见将其记为 π_A . 记选答 A 题的学生比例为 p , 由选答规则可知

$p = 10/13$. 引入随机变量

$$X_i = \begin{cases} 1, & \text{若第 } i \text{ 个被调查学生回答“是”,} \\ 0, & \text{若第 } i \text{ 个被调查学生回答“否”,} \end{cases} \quad i = 1, 2, \dots, n$$

X_1, X_2, \dots, X_n 独立同分布, $E(X_i) = \pi$, $D(X_i) = \pi(1 - \pi)$, 且按照假定所有被调查学生的回答都是真实的, 于是由全概率公式知

$$\pi = P(X_i = 1) = p\pi_A + (1 - p)(1 - \pi_A) \quad (1)$$

当 $p \neq 1/2$ 时, 由(1)可得 π_A 的估计值为

$$\hat{\pi}_A = \frac{\hat{\pi} - (1 - p)}{2p - 1} \quad (2)$$

$\hat{\pi}_A$ 是无偏估计, 即

$$E(\hat{\pi}_A) = \frac{1}{2p - 1} E\left[\frac{1}{n} \sum_{i=1}^n X_i - (1 - p)\right] = \pi_A \quad (3)$$

其方差为

$$D(\hat{\pi}_A) = \frac{1}{(2p - 1)^2} D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{\pi(1 - \pi)}{(2p - 1)^2 n} \quad (4)$$

为了对方差作进一步的分析, 将(4)的右端分解可得

$$D(\hat{\pi}_A) = \frac{\pi_A(1 - \pi_A)}{n} + \frac{p(1 - p)}{(2p - 1)^2 n} \quad (5)$$

从(5)式可以看出, $\hat{\pi}_A$ 的方差由两部分组成, 第一部分是进行直接调查并假定能得到被调查者真实回答情况下 $\hat{\pi}_A$ 的方差(令 $p = 1$ 或 $p = 0$ 就是这种情况), 而第二部分是由于随机选答机制所带来的方差. 另外, 由(5)式可以通过选取合适的 n 和 p , 使得估计达到预先确定的方差 δ (看作精度的度量), 因为 $\pi_A(1 - \pi_A) \leq 1/4$, 故欲使 $D(\hat{\pi}_A) \leq \frac{1}{4n} + \frac{p(1 - p)}{(2p - 1)^2 n} \leq \delta$, 在选定合适的 p 后只需取 n 为大

于 $\frac{1}{4\delta} + \frac{p(1 - p)}{\delta(2p - 1)^2}$ 的整数即可.

由(2)和(4)式及调查结果可以求得, 有过作弊行为学生的比例 π_A 的估计值与其标准差分别为 $\hat{\pi}_A = \frac{\hat{\pi} - (1 - p)}{2p - 1} = 0.091$ 和 $\sqrt{D(\hat{\pi}_A)} = \sqrt{\frac{\hat{\pi}(1 - \hat{\pi})}{(2p - 1)^2 n}} = 0.042$.

如果用 2 倍标准差作为估计值的精度, 可以说有过作弊行为学生的比例为 $9.1\% \pm 8.4\%$.

评注 Warner 模型中对 p 的选择设定是有限制的, 即 $p \neq 1/2$, 虽然从被调查者的角度考虑, $p = 1/2$ 最能消除他们的顾虑, 但对调查者而言, 此时无法得到 π_A 合理的估计. 而当 p 接近于 0 或 1 时, 尽管 $D(\hat{\pi}_A)$ 会变得较小, 但从分析的角度看, 对被调查者的保护程度会降低, 因此, 使用这一模型时一般 p 的取值在

0.7 至 0.8 之间比较合适.

Simmons 模型 Warner 模型虽然比直接调查要好,但所提的两个问题都还具有敏感性,被调查者或多或少还有顾虑,而且要求回答问题 A 的人数比例不能为 $1/2$ (接近 $1/2$ 也不好). 1967 年 Simmons 等人对 Warner 模型进行了改进,其最大的不同点在于调查人员提出的是两个不相关的问题,其中一个为敏感性问题,另一个为非敏感性的一般问题,这样的处理能使被调查者的合作态度进一步提高^[33].

现采用 Simmons 的无关问题选答技术来估计有过作弊行为学生的比例,假定被调查学生的人数仍为 $n=400$,问题选答的规则同 Warner 模型,只不过可供选答的两个问题变为

问题 A:你在考试中作过弊吗?

问题 B':你生日的月份是偶数吗?

假定收回的 $n=400$ 张有效答卷中,有 $n_2=80$ 个人回答“是”, $n-n_2=320$ 个人回答“否”,即对问题 A, B' 两题选答“是”的概率 π 的估计值为 $\hat{\pi}=n_2/n=1/5$.

由于被调查的学生人数较大,所以在问题 B' 中回答“是”的概率可认为是 $\pi'_B=1/2$. 以下记 π'_A 为问题 A 回答“是”的概率, π, p 和 $X_i (i=1, 2, \dots, n)$ 的意义同 Warner 模型. 由全概率公式知

$$\pi = P(X_i = 1) = p\pi'_A + (1-p)\pi'_B \quad (6)$$

可得 π'_A 的估计为

$$\hat{\pi}'_A = \frac{\hat{\pi} - (1-p)\pi'_B}{p} \quad (7)$$

$\hat{\pi}'_A$ 为 π'_A 的无偏估计,即

$$E(\hat{\pi}'_A) = \frac{1}{p} E\left[\frac{1}{n} \sum_{i=1}^n X_i - (1-p)\pi'_B\right] = \pi'_A \quad (8)$$

且其方差为

$$D(\hat{\pi}'_A) = \frac{1}{p^2} D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{\pi(1-\pi)}{np^2} \quad (9)$$

对于 $\pi'_B=1/2$, (9) 的右端可分解为

$$D(\hat{\pi}'_A) = \frac{\pi'_A(1-\pi'_A)}{n} + \frac{1-p^2}{4np^2} \quad (10)$$

将(10)式与(5)式进行比较可以发现,当且仅当 $p > 1/3$ 时有 $D(\hat{\pi}'_A) < D(\hat{\pi}_A)$. 进一步的研究可以证明,对任意给定的 π'_B , 只要 $p > 1/3$ 就可以保证 $D(\hat{\pi}'_A) < D(\hat{\pi}_A)$, 所以从估计值的方差的角度看,可以说 Simmons 模型的精度比 Warner 模型高.

由(7)和(9)式及调查结果可以求得,有过作弊行为学生的比例 π'_A 的估计值与标准差分别为 $\hat{\pi}'_A = \frac{\hat{\pi} - (1-p)\hat{\pi}'_B}{p} = 0.11$ 和 $\sqrt{D(\hat{\pi}'_A)} = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{np^2}} = 0.026$. 仍用 2 倍标准差作为估计值的精度,有过作弊行为学生的比例为 $11\% \pm 5.2\%$.

评注 在上述无关问题选答机制中,问题 B' 中回答“是”的概率 π'_B 是已知的,如果 π'_B 本身也是未知的,比如设计的问题 B' 为“你喜欢红色吗?”,而且调查者并不知道喜欢红色的学生的比例,那么就需要设计另外的调查方案,并寻求相应的估计方法(留作习题 14).

Christofides 模型 在 Warner 模型和 Simmons 模型中,被调查者均需要直接回答“是”或“否”.为进一步减少被调查者的顾虑,是否可以考虑以回答数字的方法来替代回答“是”或“否”呢? Christofides 在 2003 年设计了如下的模型^[20].

调查者准备一套外形完全相同的卡片,每张卡片上写有 $1, 2, \dots, L$ 中某一数字,数字为 k 的卡片在卡片总数中所占的比例为 $p_k (k=1, 2, \dots, L)$,并且使 p_k 不全相等.做调查时被调查者随机抽取一张卡片,看后放回,显然抽到数字为 k 的卡片的概率为 p_k .现约定调查规则如下:若被调查的学生考试中作过弊,则回答 $L+1$ 与他抽取的数字之差,反之,则回答他抽取的数字.

假设被调查的学生仍为 $n=400$ 人,调查者准备的一套卡片中,写有 $1, 2, \dots, 6$ 的卡片的数量占卡片总数的比例分别为 $0.5, 0.25, 0.15, 0.03, 0.05$ 和 0.02 ,在约定的规则下,被调查的学生回答数字 $1, 2, \dots, 6$ 的人数分别为 $176, 96, 40, 40, 28$ 和 20 ,如何由此估计有过作弊行为学生的比例呢?

为分析方便引入随机变量

$$Z_i = \begin{cases} L+1, & \text{若第 } i \text{ 个被调查学生有过作弊行为,} \\ 0, & \text{若第 } i \text{ 个被调查学生从未有作弊行为,} \end{cases} \quad i=1, 2, \dots, n$$

记 $\pi_A = P(Z_i = L+1)$, 则 π_A 即为要估计的学生作弊的概率.

记随机变量 Y_i 为第 i 个被调查学生抽到的数字,显然 $Y_i (i=1, 2, \dots, n)$ 是独立同分布的,且也与 Z_i 相互独立,根据调查机制可知,第 i 个被调查学生所回答的数字为 $d_i = |Y_i - Z_i|$,由全概率公式得

$$P(d_i = k) = (1 - \pi_A)p_k + \pi_A p_{L+1-k}, \quad k = 1, 2, \dots, L \quad (11)$$

由此可以计算 d_i 的期望为

$$E(d_i) = \sum_{k=1}^L k P(d_i = k) = \sum_{k=1}^L k p_k + \pi_A (L+1 - 2 \sum_{k=1}^L k p_k)$$

这里 $\sum_{k=1}^L k p_k$ 恰好为 Y_i 的期望,记为 $E(Y)$, 于是

$$E(d_i) = E(Y) + \pi_A (L+1 - 2E(Y)) \quad (12)$$

12. 考察一种既不同于指数模型、也不同于阻滞增长模型的情况:人口为 $x(t)$, 最大允许人口为 x_m , t 到 $t + \Delta t$ 时间内人口增长量与 $x_m - x(t)$ 成正比.

(1) 建立确定性模型, 将结果作图, 与指数模型和阻滞增长模型的结果进行比较.

(2) 作出适当的假设, 建立相应的随机性模型, 求出入口的期望, 并解释其与 (1) 中的 $x(t)$ 在形式上完全一致的意义.

13. 为克服 Warner 模型中需要回答的两个问题都是敏感问题的缺点, 设计如下调查方案: 调查者先制作一套卡片, 由三部分组成, 第一部分卡片 (1 号卡片) 上写有“若你曾在考试中作过弊, 请回答数字 1; 若你从未在考试中作过弊, 请回答数字 0”, 第二部分卡片 (2 号卡片) 上写有“请直接回答 1”, 第三部分卡片 (3 号卡片) 上写有“请直接回答 0”, 将这三部分卡片充分混合后放入一盒中, 其中第 i 部分的卡片的比例为 p_i . 被调查的学生从盒中随机抽一张卡片, 做真实回答后放回. 已知被调查学生的总人数为 n , 回答结果是“1”的人数为 n_1 , 试估计学生中有过作弊行为的比例 $\hat{\pi}_A$, 讨论 $\hat{\pi}_A$ 的无偏性并给出它的方差.

14. 在 Simmons 模型中, 假设调查中设计的问题 B' 为“你喜欢红色吗?”, 而且调查者并不知道喜欢红色的学生的比例, 试设计调查方案, 估计学生有过作弊行为的比例, 并求该估计的方差.

15. 作弊与赌博是两个不相关的敏感问题, 调查者的目的是估计学生中曾有作弊和赌博行为的比例, 为此设计了如下的问题:

问题 A: 你在考试中作过弊吗?

问题 B: 你从未参加过赌博吗?

这样设计提问也能为被调查者提供足够的保护. 为实现此调查方案, 选取两组学生独立进行调查, 并设计两套外形相同的卡片, 其中第 i 套卡片中写有问题 A 的比例为 p_i ($i=1, 2$), 写有问题 B 的比例为 $1-p_i$, 且 $p_1 \neq p_2$, 第 i 组被调查学生的人数为 m_i , 他们从第 i 套卡片中随机选择一张, 真实作答后放回, 其中回答“是”的人数为 $m_i^{(r)}$.

(1) 分别估计学生中曾有作弊和赌博行为的比例, 并给出它们的估计方差.

(2) 假定曾作过弊学生真实回答问题 A 的概率为 T_A , 参加过赌博的学生真实回答问题 B 的概率为 T_B , 且 T_A, T_B 均已知, 而其他情形均真实作答. 试分别重新估计学生中曾有作弊和赌博行为的比例以及它们的估计方差.

16. 对于 9.7 节的 Christofides 模型证明, 当 $L=2$ 时 (14) 和 (15) 式分别退化为 (2) 和 (4), 即 Warner 模型是它的一个特例.

模型最优, Simmons 模型次之, Warner 模型最差. 当然, 如果选择的参数 L, p_1, \dots, p_L 不合适, Christofides 模型不一定最好, 但可以肯定的是, 我们总可以选择合适的参数来提高估计的精度.

习 题

1. 在 9.1 节传送带效率模型中, 设工人数 n 固定不变. 若想提高传送带效率 D , 一种简单的办法是增加一个周期内通过工作台的钩子数 m , 比如增加一倍, 其他条件不变. 另一种办法是在原来放置一只钩子的地方放置两只钩子, 其他条件不变, 于是每个工人在任何时刻可以同时触到两只钩子, 只要其中有一只是空的, 他就可以挂上产品, 这种办法用的钩子数量与第一种办法一样. 试推导这种情况下传送带效率的公式, 从数量关系上说明这种办法比第一种办法好.

2. 利用 9.2 节的模型计算, 若每份报纸的购进价为 0.75 元, 售出价为 1 元, 退回价为 0.6 元, 需求量服从均值 500 份, 均方差 50 份的正态分布, 报童每天应购进多少份报纸才能使平均收入最高, 这个最高收入是多少?

3. 在 9.3 节中将假设条件 4 改为一周的销售是均匀进行的, 如图 3 所示. 试确定使平均费用达到最小的策略 S .

4. 某商店要订购一批商品零售, 设购进价 c_1 , 售出价 c_2 , 订购费 c_0 (与数量无关), 随机需求量 r 的概率密度为 $p(r)$, 每件商品的贮存费为 c_3 (与时间无关). 问如何确定订购量才能使商店的平均利润最大, 这个平均利润是多少. 为使这个平均利润为正值, 需要对订购费 c_0 加什么限制?

5. 建立交货时间为随机变量的存贮模型. 设商品订货费为 c_1 , 每件商品单位时间的贮存费为 c_2 , 缺货费为 c_3 , 单位时间需求量为 r . 下图 n 称订货点. 当贮存量降至 L 时订货, 而交货时间 x 是随机的, 如图中的 x_1, x_2, \dots . 设 x 的概率密度函数为 $p(x)$. 订货量使下一周期初的贮存量达到固定值 Q . 为了使总费用最小, 选择合适的目标函数建立模型, 确定最佳订货点 L .

6. 在 9.4 节中证明方程 (13) 仅有一个负根 z^* , 并且 z^* 给出 (12) 式 $J(z)$ 的极小值.

7. 在 9.4 节给出的例子中, 若 $l = 2.0$ m 不变, 而均方差减为 $\sigma = 10$ cm, 问均值 m 应为多大, 每得到一根成品材的浪费量多大 (与原来的数值相比较).

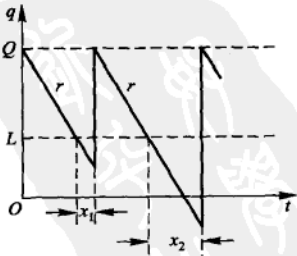
8. 在 9.4 节中若钢材粗轧后, 长度在 l_1 与 l 之间的可降级使用, 长度小于 l_1 的才整根报废. 试选用合适的目标函数建立优化模型, 使某种意义上下的浪费量最小.

9. 推导 9.5 节的 (12) 式.

10. 作出与确定性阻滞增长模型相应的假设, 建立随机性的阻滞增长模型.

11. 假设在 9.5 节的模型 (2), (3) 式中, 只考虑出生、忽略死亡, 验证 $P_n(t)$ 满足负二项分布, 即

$$P_n(t) = C_{n-1}^{n_0} (e^{-\lambda t})^{n_0} (1 - e^{-\lambda t})^{n-n_0}, n = n_0, n_0 + 1, \dots$$



将 Y_i 的方差记为 $D(Y)$, 经过计算还可得到

$$D(d_i) = D(Y) + \pi_A(1 - \pi_A)(L + 1 - 2E(Y))^2 \quad (13)$$

计算 n 个被调查学生所回答数字 d_1, d_2, \dots, d_n 的均值 $\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i$, 作为 (12) 式中 $E(d_i)$ 的估计值, 则可得 π_A 的估计为

$$\hat{\pi}_A^* = \frac{\bar{d} - E(Y)}{L + 1 - 2E(Y)}, L + 1 - 2E(Y) \neq 0 \quad (14)$$

易知, $\hat{\pi}_A^*$ 为 π_A 的无偏估计, 且具有方差

$$\begin{aligned} D(\hat{\pi}_A^*) &= \frac{1}{(L + 1 - 2E(Y))^2} D(\bar{d}) \\ &= \frac{\pi_A^*(1 - \pi_A^*)}{n} + \frac{D(Y)}{n(L + 1 - 2E(Y))^2} \end{aligned} \quad (15)$$

当 $L=2$ 时, 可以证明 (14) 和 (15) 式分别退化为 (2) 和 (4), 即 Warner 模型是 Christofides 模型的一个特例 (留作习题 16)。

在本次调查中已知 $(p_1, \dots, p_6) = (0.5, 0.25, 0.15, 0.03, 0.05, 0.02)$, 可计算出 $E(Y) = 1.94, D(Y) = 1.5364$, 由被调查的学生回答数字 1, 2, \dots , 6 的人数分别为 176, 96, 40, 40, 28 和 20, 可计算出其均值和方差分别为 $\bar{d} = 2.27$ 和 $D(\bar{d}) = 0.0058$, 再由 (14) 和 (15) 式求得有过作弊行为学生的比例 π_A 的估计值与其标准差分别为 $\hat{\pi}_A^* = \frac{2.27 - 1.94}{6 + 1 - 2 \times 1.94} = 0.106$ 和 $\sqrt{D(\hat{\pi}_A^*)} = \sqrt{\frac{0.0058}{(6 + 1 - 2 \times 1.94)^2}} = 0.0244$. 如果仍用 2 倍标准差作为估计值的精度, 有过作弊行为学生的比例为 $10.6\% \pm 4.88\%$.

评注 虽然 Christofides 模型没有像 Simmons 模型那样提出与调查的敏感问题无关的其他问题, 但它不要求被调查者直接回答“是”或“否”, 而是可以灵活地选择合适的 $L, p_1, \dots, p_L (L \geq 3)$, 使之能得到比 Warner 模型甚至比 Simmons 模型具有更高精度的估计. 这一点不难从这些模型的估计值的方差公式中发现, 因为在式 (5), (10) 和 (15) 中, 它们的第一部分均为对敏感问题直接调查时所产生的方差, 所以当被调查的人数相同时, 若要比 3 个模型的精度, 只需比较相应公式中第二部分的大小. 比如在我们的问题中, 对 Warner 模型和 Simmons 模型, 由于 $p = 10/13$, Warner 模型中的第二部分为 $\frac{p(1-p)}{n(2p-1)^2} = 0.00153$, Simmons 模型中的第二部分为 $\frac{1-p^2}{4np^2} = 0.00043$, 而在 Christofides 模型中, 第二部分为 $\frac{D(Y)}{n(L+1-2E(Y))^2} = 0.00039$. 因此, 从估计的精度来看, 对于这个问题 Christofides

第10章

统计回归模型

当人们对研究对象的内在特性和各因素间的关系有比较充分的认识时,一般用机理分析方法建立数学模型,本书前面讨论的绝大多数模型都是如此.如果由于客观事物内部规律的复杂性及人们认识程度的限制,无法分析实际对象内在的因果关系,建立合乎机理规律的数学模型,那么通常的办法是搜集大量的数据,基于对数据的统计分析去建立模型,本章只介绍用途非常广泛的一类随机模型——统计回归模型.

与专门讲述统计方法的教材不同,这里将不涉及回归分析的数学原理和方法,而是通过几个实例讨论如何选择不同类型的模型,以及怎样对软件得到的结果进行分析.没有学过这部分数学知识的读者只要不追究这些结果的统计学上的意义,就仍然可以在这里学到用回归模型解决实际问题的基本方法.

10.1 牙膏的销售量

问题 某大型牙膏制造企业为了更好地拓展产品市场,有效地管理库存,公司董事会要求销售部门根据市场调查,找出公司生产的牙膏销售量与销售价格、广告投入等之间的关系,从而预测出在不同价格和广告费用下的销售量.为此,销售部的研究人员收集了过去30个销售周期(每个销售周期为4周)公司生产的牙膏的销售量、销售价格、投入的广告费用,以及同期其他厂家生产的同类牙膏的市场平均销售价格,见表1.试根据这些数据建立一个数学模型,分析牙膏销售量与其他因素的关系,为制订价格策略和广告投入策略提供数量依据^[13].

表1 牙膏销售量与销售价格、广告费用等数据
(其中价格差指其他厂家平均价格与公司销售价格之差)

销售 周期	公司销售价 格/元	其他厂家平 均价格/元	广告费用 /百万元	价格差/元	销售量 /百万支
1	3.85	3.80	5.50	-0.05	7.38
2	3.75	4.00	6.75	0.25	8.51

续表

销售 周期	公司销售价 格/元	其他厂家平 均价格/元	广告费用 /百万元	价格差/元	销售量 /百万支
3	3.70	4.30	7.25	0.60	9.52
4	3.70	3.70	5.50	0	7.50
5	3.60	3.85	7.00	0.25	9.33
6	3.60	3.80	6.50	0.20	8.28
7	3.60	3.75	6.75	0.15	8.75
8	3.80	3.85	5.25	0.05	7.87
9	3.80	3.65	5.25	-0.15	7.10
10	3.85	4.00	6.00	0.15	8.00
11	3.90	4.10	6.50	0.20	7.89
12	3.90	4.00	6.25	0.10	8.15
13	3.70	4.10	7.00	0.40	9.10
14	3.75	4.20	6.90	0.45	8.86
15	3.75	4.10	6.80	0.35	8.90
16	3.80	4.10	6.80	0.30	8.87
17	3.70	4.20	7.10	0.50	9.26
18	3.80	4.30	7.00	0.50	9.00
19	3.70	4.10	6.80	0.40	8.75
20	3.80	3.75	6.50	-0.05	7.95
21	3.80	3.75	6.25	-0.05	7.65
22	3.75	3.65	6.00	-0.10	7.27
23	3.70	3.90	6.50	0.20	8.00
24	3.55	3.65	7.00	0.10	8.50
25	3.60	4.10	6.80	0.50	8.75
26	3.65	4.25	6.80	0.60	9.21
27	3.70	3.65	6.50	-0.05	8.27
28	3.75	3.75	5.75	0	7.67
29	3.80	3.85	5.80	0.05	7.93
30	3.70	4.25	6.80	0.55	9.26

分析与假设 由于牙膏是生活必需品,对大多数顾客来说,在购买同类产品的牙膏时更多地会在意不同品牌之间的价格差异,而不是它们的价格本身。因

此,在研究各个因素对销售量的影响时,用价格差代替公司销售价格和其他厂家平均价格更为合适。

记牙膏销售量为 y ,其他厂家平均价格与公司销售价格之差(价格差)为 x_1 ,公司投入的广告费用为 x_2 ,其他厂家平均价格和公司销售价格分别为 x_3 和 x_4 , $x_1 = x_3 - x_4$ 。基于上面的分析,我们仅利用 x_1 和 x_2 来建立 y 的预测模型。

基本模型 为了大致地分析 y 与 x_1 和 x_2 的关系,首先利用表 1 的数据分别作出 y 对 x_1 和 x_2 的散点图(见图 1 和图 2 中的圆点)。

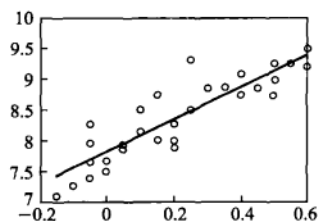


图 1 y 对 x_1 的散点图

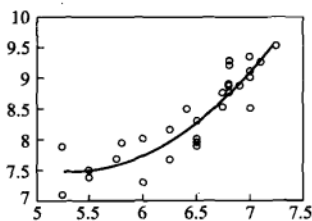


图 2 y 对 x_2 的散点图

从图 1 可以发现,随着 x_1 的增加, y 的值有比较明显的线性增长趋势,图中的直线是用线性模型

$$y = \beta_0 + \beta_1 x_1 + \varepsilon \quad (1)$$

拟合的(其中 ε 是随机误差)。而在图 2 中,当 x_2 增大时, y 有向上弯曲增加的趋势,图中的曲线是用二次函数模型

$$y = \beta_0 + \beta_1 x_2 + \beta_2 x_2^2 + \varepsilon \quad (2)$$

拟合的。

综合上面的分析,结合模型(1)和(2)建立如下的回归模型

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2 + \varepsilon \quad (3)$$

(3)式右端的 x_1 和 x_2 称为回归变量(自变量), $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2$ 是给定价格差 x_1 、广告费用 x_2 时,牙膏销售量 y 的平均值,其中的参数 $\beta_0, \beta_1, \beta_2, \beta_3$ 称为回归系数,由表 1 的数据估计,影响 y 的其他因素作用都包含在随机误差 ε 中。如果模型选择得合适, ε 应大致服从均值为 0 的正态分布。

模型求解 直接利用 MATLAB 统计工具箱中的命令 regress 求解,使用格式为

$$[b, bint, r, rint, stats] = \text{regress}(y, x, \alpha)$$

其中输入 y 为模型(3)中 y 的数据(n 维向量, $n = 30$), x 为对应于回归系数 $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)$ 的数据矩阵 $[1 \ x_1 \ x_2 \ x_2^2]$ ($n \times 4$ 矩阵,其中第 1 列为全 1 向量), α 为置信水平 α (缺省时 $\alpha = 0.05$);输出 b 为 β 的估计值,常记作 $\hat{\beta}$,

bint 为 b 的置信区间, r 为残差向量 $y - x\hat{\beta}$, rint 为 r 的置信区间, stats 为回归模型的检验统计量, 有 4 个值, 第 1 个是回归方程的决定系数 R^2 (R 是相关系数), 第 2 个是 F 统计量值, 第 3 个是与 F 统计量对应的概率值 p , 第 4 个是剩余方差 s^2 .

得到模型(3)的回归系数估计值及其置信区间(置信水平 $\alpha = 0.05$)、检验统计量 R^2, F, p, s^2 的结果见表 2.

表 2 模型(3)的计算结果

参数	参数估计值	参数置信区间
β_0	17.324 4	[5.728 2, 28.920 6]
β_1	1.307 0	[0.682 9, 1.931 1]
β_2	-3.695 6	[-7.498 9, 0.107 7]
β_3	0.348 6	[0.037 9, 0.659 4]
$R^2 = 0.905\ 4 \quad F = 82.940\ 9 \quad p < 0.000\ 1 \quad s^2 = 0.049\ 0$		

结果分析 表 2 显示, $R^2 = 0.905 4$ 指因变量 y (销售量) 的 90.54% 可由模型确定, F 值远远超过 F 检验的临界值, p 远小于 α , 因而模型(3)从整体来看是可用的.

表 2 的回归系数给出了模型(3)中 $\beta_0, \beta_1, \beta_2, \beta_3$ 的估计值, 即 $\hat{\beta}_0 = 17.324 4, \hat{\beta}_1 = 1.307 0, \hat{\beta}_2 = -3.695 6, \hat{\beta}_3 = 0.348 6$. 检查它们的置信区间发现, 只有 β_2 的置信区间包含零点(但区间右端点距零点很近), 表明回归变量 x_2 (对因变量 y 的影响)不是太显著的, 但由于 x_2^2 是显著的, 我们仍将变量 x_2 保留在模型中.

销售量预测 将回归系数的估计值代入模型(3), 即可预测公司未来某个销售周期牙膏的销售量 y , 预测值记作 \hat{y} , 得到模型(3)的预测方程

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_2^2 \quad (4)$$

只需知道该销售周期的价格差 x_1 和投入的广告费用 x_2 , 就可以计算预测值 \hat{y} .

值得注意的是公司无法直接确定价格差 x_1 , 而只能制定公司该周期的牙膏售价 x_4 , 但是同期其他厂家的平均价格 x_3 一般可以通过分析和预测当时的市场情况及原材料的价格变化等估计出. 模型中引入价格差 $x_1 = x_3 - x_4$ 作为回归变量, 而非 x_3, x_4 的好处在于, 公司可以更灵活地来预测产品的销售量(或市场需求量), 因为 x_3 的值不是公司所能控制的. 预测时只要调整 x_4 达到设定的回归变量 x_1 的值, 比如公司计划在未来的某个销售周期中, 维持产品的价格差为 x_1

=0.2 元,并将投入 $x_2 = 6.5$ 百万元的广告费用,则该周期牙膏销售量的估计值为 $\hat{y} = 17.324 4 + 1.307 0 \times 0.2 - 3.695 6 \times 6.5 + 0.348 6 \times 6.5^2 = 8.292 8$ 百万支。

回归模型的一个重要应用是,对于给定的回归变量的取值,可以以一定的置信度预测因变量的取值范围,即预测区间。比如当 $x_1 = 0.2, x_2 = 6.5$ 时可以算出^①,牙膏销售量的置信度为 95% 的预测区间为 $[7.823 0, 8.763 6]$,它表明在将来的某个销售周期中,如公司维持产品的价格差为 0.2 元,并投入 650 万元的广告费用,那么可以有 95% 的把握保证牙膏的销售量在 7.823 到 8.763 6 百万支之间。实际操作时,预测上限可以用来作为库存管理的目标值,即公司可以生产(或库存)8.763 6 百万支牙膏来满足该销售周期顾客的需求;预测下限则可以用来较好地把握(或控制)公司的现金流,理由是该公司对该周期销售 7.823 百万支牙膏十分自信,如果在该销售周期中公司将牙膏售价定为 3.70 元,且估计同期其他厂家的平均价格为 3.90 元,那么董事会可以有充分的依据知道公司的牙膏销售额应在 $7.823 \times 3.7 \approx 29$ 百万元以上。

模型改进 模型(3)中回归变量 x_1 和 x_2 对因变量 y 的影响是相互独立的,即牙膏销售量 y 的均值与广告费用 x_2 的二次关系由回归系数 β_2 和 β_3 确定,而不依赖于价格差 x_1 ,同样, y 的均值与 x_1 的线性关系由回归系数 β_1 确定,不依赖于 x_2 。根据直觉和经验可以猜想, x_1 和 x_2 之间的交互作用会对 y 有影响,不妨简单地用 x_1, x_2 的乘积代表它们的交互作用,于是将模型(3)增加一项,得到

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2 + \beta_4 x_1 x_2 + \varepsilon \quad (5)$$

在这个模型中, y 的均值与 x_2 的二次关系为 $(\beta_2 + \beta_4 x_1) x_2 + \beta_3 x_2^2$,由系数 $\beta_2, \beta_3, \beta_4$ 确定,并依赖于价格差 x_1 。

下面让我们用表 1 的数据估计模型(5)的系数。利用 MATLAB 的统计工具箱得到的结果见表 3。

表 3 模型(5)的计算结果

参数	参数估计值	参数置信区间
β_0	29.113 3	[13.701 3, 44.525 2]
β_1	11.134 2	[1.977 8, 20.290 6]
β_2	-7.608 0	[-12.693 2, -2.522 8]
β_3	0.671 2	[0.253 8, 1.088 7]
β_4	-1.477 7	[-2.851 8, -0.103 7]
$R^2 = 0.920 9 \quad F = 72.777 1 \quad p < 0.000 1 \quad s^2 = 0.042 6$		

① 具体计算参见[91](4.77)~(4.79)式,用 MATLAB 统计工具箱中现成的程序结果与此不同。

表3与表2的结果相比, R^2 有所提高, 说明模型(5)比模型(3)有所改进. 并且, 所有参数的置信区间, 特别是 x_1, x_2 的交互作用项 $x_1 x_2$ 的系数 β_4 的置信区间不包含零点, 所以有理由相信模型(5)比模型(3)更符合实际.

用模型(5)对公司的牙膏销售量作预测. 仍设在某个销售周期中, 维持产品的价格差 $x_1 = 0.2$ 元, 并将投入 $x_2 = 6.5$ 百万元的广告费用, 则该周期牙膏销售量 y 的估计值为 $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_2^2 + \hat{\beta}_4 x_1 x_2 = 29.1133 + 11.1342 \times 0.2 - 7.6080 \times 6.5 + 0.6712 \times 6.5^2 - 1.4777 \times 0.2 \times 6.5 = 8.3253$ 百万支, 置信度为95%的预测区间为 $[7.8953, 8.7592]$, 与模型(3)的结果相比, \hat{y} 略有增加, 而预测区间长度短一些.

在保持广告费用 $x_2 = 6.5$ 百万元不变的条件下, 分别对模型(3)和(5)中牙膏销售量的均值 \hat{y} 与价格差 x_1 的关系作图, 见图3和图4.

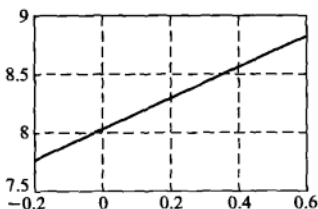


图3 模型(3) \hat{y} 与 x_1 的关系

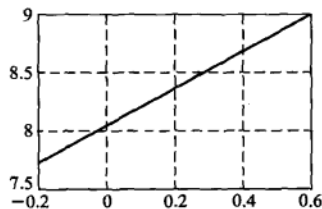


图4 模型(5) \hat{y} 与 x_1 的关系

在保持价格差 $x_1 = 0.2$ 元不变的条件下, 分别对模型(3)和(5)中牙膏销售量的均值 \hat{y} 与广告费用 x_2 的关系作图, 见图5和图6.

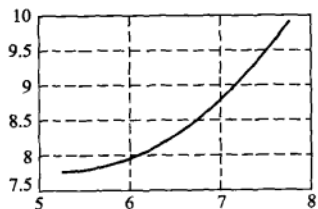


图5 模型(3) \hat{y} 与 x_2 的关系

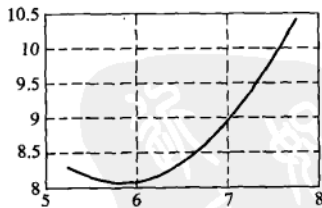


图6 模型(5) \hat{y} 与 x_2 的关系

可以看出, 交互作用项 $x_1 x_2$ 加入模型, 对 \hat{y} 与 x_1 的关系稍有影响, 而 \hat{y} 与 x_2 的关系有较大变化, 当 $x_2 < 6$ 时 \hat{y} 出现下降, $x_2 > 6$ 以后 \hat{y} 上升则快得多.

进一步的讨论 为进一步了解 x_1 和 x_2 之间的交互作用, 考察模型(5)的预测方程

$$\hat{y} = 29.1133 + 11.1342x_1 - 7.6080x_2 + 0.6712x_2^2 - 1.4777x_1x_2 \quad (6)$$

如果取价格差 $x_1 = 0.1$ 元,代入(6)可得

$$\hat{y} \Big|_{x_1=0.1} = 30.2267 - 7.7558x_2 + 0.6712x_2^2 \quad (7)$$

再取 $x_1 = 0.3$ 元,代入(6)得

$$\hat{y} \Big|_{x_1=0.3} = 32.4536 - 8.0513x_2 + 0.6712x_2^2 \quad (8)$$

它们均为 x_2 的二次函数,其图形见图 7,且

$$\hat{y} \Big|_{x_1=0.3} - \hat{y} \Big|_{x_1=0.1} = 2.2269 - 0.2955x_2 \quad (9)$$

由(9)式可得,当 $x_2 < 7.5360$ 时,总有 $\hat{y} \Big|_{x_1=0.3} > \hat{y} \Big|_{x_1=0.1}$,即若广告费用不超过大约 7.5 百万元,价格差定在 0.3 元时的销售量,比价格差定在 0.1 元的大,也就是说,这时的价格优势会使销售量增加。

由图 7 还可以发现,虽然广告投入的增加会使销售量增加(只要广告费用超过大约 6 百万元),但价格差较小时增加的速率要更大些。这些现象都是由于引入了交互作用项 x_1x_2 后产生的。价格差较大时,许多消费者是受价格的驱动来购买公司的产品,所以可以较少地依赖广告投入的增加来提高销售量。价格差较小时,则更需要靠广告来吸引更多的顾客。

另外,当公司牙膏的售价在市场中明显处于弱势时, x_1 和 x_2 之间的交互作用项不见得就是乘积项 x_1x_2 了,可能出现其他形式的组合。

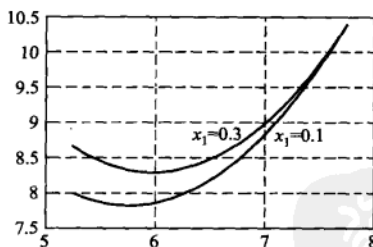


图 7 \hat{y} 与 x_2 的关系((7)与(8)的图形)

完全二次多项式模型 与 x_1 和 x_2 的完全二次多项式模型

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \beta_4x_1^2 + \beta_5x_2^2 + \varepsilon \quad (10)$$

相比,模型(5)只少 x_1^2 项,我们不妨增加这一项,建立模型(10)。这样做的好处之一是 MATLAB 统计工具箱中有直接的命令 `rstool` 求解,并且以交互式画面给出 y 的估计值 \hat{y} 和预测区间。这个命令的输出如图 8,从左下方的输出 Export 可以得到模型(10)的回归系数的估计值为

$$\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \hat{\beta}_5) = (32.0984, 14.7436,$$

-8.636 7, -2.103 8, 1.107 4, 0.759 4)

用鼠标移动交互式画面中的十字线,或在图下方的窗口内输入,可改变 x_1 和 x_2 的数值,图中当 $x_1 = 0.2, x_2 = 6.5$ 时,左边的窗口显示 $\hat{y} = 8.3029$,预测区间为 $8.3029 \pm 0.2558 = [8.0471, 8.5587]$. 这些结果与模型(5)相差不大.

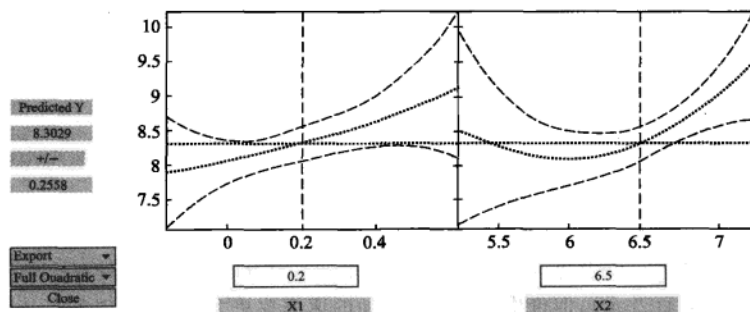


图8 完全二次多项式模型(10)的输出

评注 从这个实例我们看到,建立回归模型可以先根据已知的数据,从常识和经验进行分析,辅以作图(如图1,图2的散点图),决定取哪几个回归变量,及它们的函数形式(如线性的、二次的).用软件(如MATLAB统计工具箱)求解后,作统计分析: R^2 , F , p 值, s^2 的大小是对模型整体的评价,每个回归系数置信区间是否包含零点,可以用来检验对应的回归变量对因变量的影响是否显著(若包含零点则不显著).如果对结果不够满意,则应改进模型,如添加二次项、交互项等.

对因变量进行预测,经常是建立回归模型的主要目的之一,本节提供了预测的方法,以及对结果作进一步讨论的实例.

10.2 软件开发人员的薪金

问题 一家高技术公司人事部门为研究软件开发人员的薪金与他们的资历、管理责任、教育程度等因素之间的关系,要建立一个数学模型,以便分析公司人事策略的合理性,并作为新聘用人员薪金的参考.他们认为目前公司人员的薪金总体上是合理的,可以作为建模的依据,于是调查了46名软件开发人员的档案资料,如表1,其中资历一列指从事专业工作的年数,管理一列中1表示管理人员,0表示非管理人员,教育一列中1表示中学程度,2表示大学程度,3表示更高程度(研究生)^[19].

表1 软件开发人员的薪金与他们的资历、管理责任、教育程度之间的关系

编号	薪金	资历	管理	教育	编号	薪金	资历	管理	教育
1	13 876	1	1	1	24	22 884	6	1	2
2	11 608	1	0	3	25	16 978	7	1	1
3	18 701	1	1	3	26	14 803	8	0	2
4	11 283	1	0	2	27	17 404	8	1	1
5	11 767	1	0	3	28	22 184	8	1	3
6	20 872	2	1	2	29	13 548	8	0	1
7	11 772	2	0	2	30	14 467	10	0	1
8	10 535	2	0	1	31	15 942	10	0	2
9	12 195	2	0	3	32	23 174	10	1	3
10	12 313	3	0	2	33	23 780	10	1	2
11	14 975	3	1	1	34	25 410	11	1	2
12	21 371	3	1	2	35	14 861	11	0	1
13	19 800	3	1	3	36	16 882	12	0	2
14	11 417	4	0	1	37	24 170	12	1	3
15	20 263	4	1	3	38	15 990	13	0	1
16	13 231	4	0	3	39	26 330	13	1	2
17	12 884	4	0	2	40	17 949	14	0	2
18	13 245	5	0	2	41	25 685	15	1	3
19	13 677	5	0	3	42	27 837	16	1	2
20	15 965	5	1	1	43	18 838	16	0	2
21	12 366	6	0	1	44	17 483	16	0	1
22	21 352	6	1	3	45	19 207	17	0	2
23	13 839	6	0	2	46	19 346	20	0	1

分析与假设 按照常识,薪金自然随着资历(年)的增长而增加,管理人员的薪金应高于非管理人员,教育程度越高薪金也越高. 薪金记作 y , 资历(年)记作 x_1 , 为了表示是否管理人员, 定义

$$x_2 = \begin{cases} 1, & \text{管理人员} \\ 0, & \text{非管理人员} \end{cases}$$

为了表示 3 种教育程度, 定义

$$x_3 = \begin{cases} 1, & \text{中学} \\ 0, & \text{其他} \end{cases} \quad x_4 = \begin{cases} 1, & \text{大学} \\ 0, & \text{其他} \end{cases}$$

这样, 中学用 $x_3 = 1, x_4 = 0$ 表示, 大学用 $x_3 = 0, x_4 = 1$ 表示, 研究生则用 $x_3 = 0, x_4 = 0$ 表示。

为简单起见, 我们假定资历(年)对薪金的作用是线性的, 即资历每加一年, 薪金的增长是常数; 管理责任、教育程度、资历诸因素之间没有交互作用, 建立线性回归模型。

基本模型 薪金 y 与资历 x_1 , 管理责任 x_2 , 教育程度 x_3, x_4 之间的多元线性回归模型为

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + \varepsilon \quad (1)$$

其中 a_0, a_1, \dots, a_4 是待估计的回归系数, ε 是随机误差。

利用 MATLAB 的统计工具箱可以得到回归系数及其置信区间(置信水平 $\alpha = 0.05$)、检验统计量 R^2, F, p, s^2 的结果, 见表 2。

表 2 模型(1)的计算结果

参数	参数估计值	参数置信区间
a_0	11 032	[10 258, 11 807]
a_1	546	[484, 608]
a_2	6 883	[6 248, 7 517]
a_3	-2 994	[-3 826, -2 162]
a_4	148	[-636, 931]
$R^2 = 0.957 \quad F = 226 \quad p < 0.0001 \quad s^2 = 1.057 \times 10^6$		

结果分析 从表 2 知 $R^2 = 0.957$, 即因变量(薪金)的 95.7% 可由模型确定, F 值远远超过 F 检验的临界值, p 远小于 α , 因而模型(1)从整体来看是可用的。比如, 利用模型可以估计(或预测)一个大学毕业、有 2 年资历、非管理人员的薪金为

$$\hat{y} = \hat{a}_0 + \hat{a}_1 \times 2 + \hat{a}_2 \times 0 + \hat{a}_3 \times 0 + \hat{a}_4 \times 1 = 12\,272$$

模型中各个回归系数的含义可初步解释如下: x_1 的系数为 546, 说明资历每增加 1 年, 薪金增长 546; x_2 的系数为 6 883, 说明管理人员的薪金比非管理人员

多 6 883; x_3 的系数为 -2 994, 说明中学程度的薪金比研究生少 2 994; x_4 的系数为 148, 说明大学程度的薪金比研究生多 148, 但是应该注意到 a_4 的置信区间包含零点, 所以这个系数的解释是不可靠的。

需要指出, 以上解释是就平均值来说, 并且, 一个因素改变引起的因变量的变化量, 都是在其他因素不变的条件才成立的。

进一步的讨论 a_4 的置信区间包含零点, 说明基本模型(1)存在缺点. 为寻找改进的方向, 常用残差分析方法(残差 ε 指薪金的实际值 y 与用模型估计的薪金 \hat{y} 之差, 是模型(1)中随机误差 ε 的估计值, 这里用了同一个符号). 我们将影响因素分成资历与管理-教育组合两类, 管理-教育组合的定义如表 3.

表 3 管理-教育组合

组合	1	2	3	4	5	6
管理	0	1	0	1	0	1
教育	1	1	2	2	3	3

为了对残差进行分析, 图 1 给出 ε 与资历 x_1 的关系, 图 2 给出 ε 与管理 x_2 -教育 x_3 , x_4 组合间的关系。

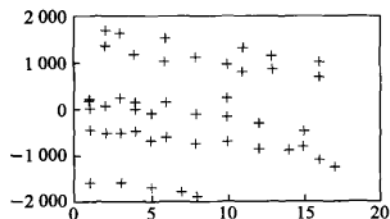


图 1 模型(1) ε 与 x_1 的关系

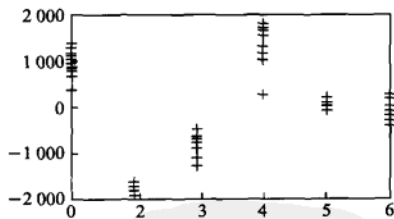


图 2 模型(1) ε 与 x_2-x_3, x_4 组合的关系

从图 1 看, 残差大概分成 3 个水平, 这是由于 6 种管理-教育组合混在一起, 在模型中未被正确反映的结果; 从图 2 看, 对于前 4 个管理-教育组合, 残差或者全为正, 或者全为负, 也表明管理-教育组合在模型中处理不当。

在模型(1)中管理责任和教育程度是分别起作用的, 事实上, 二者可能起着交互作用, 如大学程度的管理人员的薪金会比二者分别的薪金之和高一点。

以上分析提示我们, 应在基本模型(1)中增加管理 x_2 与教育 x_3, x_4 的交互项, 建立新的回归模型。

更好的模型 增加 x_2 与 x_3, x_4 的交互项后, 模型记作

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_2x_3 + a_6x_2x_4 + \varepsilon \quad (2)$$

利用 MATLAB 的统计工具箱得到的结果如表 4.

表 4 模型(2)的计算结果

参数	参数估计值	参数置信区间
a_0	11 204	[11 044, 11 363]
a_1	497	[486, 508]
a_2	7 048	[6 841, 7 255]
a_3	-1 727	[-1 939, -1 514]
a_4	-348	[-545, -152]
a_5	-3 071	[-3 372, -2 769]
a_6	1 836	[1 571, 2 101]
$R^2 = 0.9988 \quad F = 5.545 \quad p < 0.0001 \quad s^2 = 3.0047 \times 10^4$		

由表 4 可知,模型(2)的 R^2 和 F 值都比模型(1)有所改进,并且所有回归系数的置信区间都不含零点,表明模型(2)是完全可用的.

与模型(1)类似,作模型(2)的两个残差分析图(图 3,图 4),可以看出,已经消除了图 1、图 2 中的不正常现象,这也说明了模型(2)的适用性.

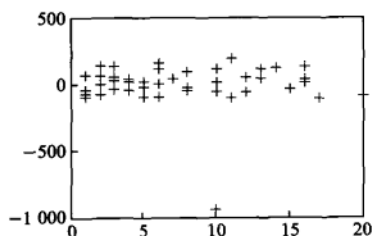


图 3 模型(2) ε 与 x_1 的关系

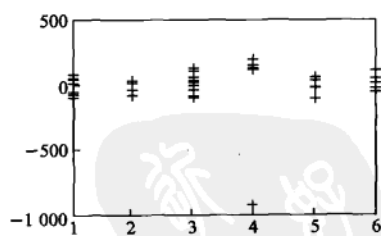


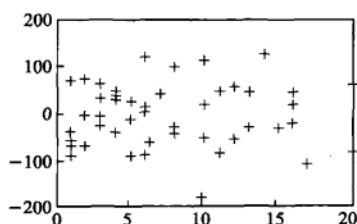
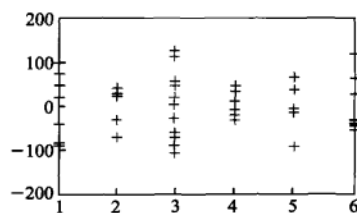
图 4 模型(2) ε 与 x_2, x_3, x_4 组合的关系

从图 3、图 4 还可以发现一个异常点:具有 10 年资历、大学程度的管理人员(从表 1 可以查出是 33 号),他的实际薪金明显地低于模型的估计值,也明显低于与他有类似经历的其他人的薪金.这可能是由我们未知的原因造成的.为了使个别的数据不致影响整个模型,应该将这个异常数据去掉,对模型(2)重新估计回归系数,得到的结果如表 5,残差分析图见图 5,图 6.可以看出,去掉异常数据后结果又有改善.

表 5 模型(2)去掉异常数据后的计算结果

参数	参数估计值	参数置信区间
a_0	11 200	[11 139, 11 261]
a_1	498	[494, 503]
a_2	7 041	[6 962, 7 120]
a_3	-1 737	[-1 818, -1 656]
a_4	-356	[-431, -281]
a_5	-3 056	[-3 171, -2 942]
a_6	1 997	[1 894, 2 100]

$R^2 = 0.9998 \quad F = 36\,701 \quad p < 0.0001 \quad s^2 = 4.347 \times 10^3$

图 5 模型(2)去掉异常数据后
 ε 与 x_1 的关系图 6 模型(2)去掉异常数据后
 ε 与 $x_2 - x_3, x_4$ 组合的关系

模型应用 对于回归模型(2),用去掉异常数据(33号)后估计出的系数,得到的结果是满意的.作为这个模型的应用之一,不妨用它来“制订”6种管理-教育组合人员的“基础”薪金(即资历为零的薪金,当然,这也是平均意义上的).利用模型(2)和表5容易得到表6.

表 6 6种管理-教育组合人员的“基础”薪金

组合	管理	教育	系数	“基础”薪金
1	0	1	$a_0 + a_3$	9 463
2	1	1	$a_0 + a_2 + a_3 + a_5$	13 448
3	0	2	$a_0 + a_4$	10 844
4	1	2	$a_0 + a_2 + a_4 + a_6$	19 882
5	0	3	a_0	11 200
6	1	3	$a_0 + a_2$	18 241

可以看出,大学程度的管理人员的薪金比研究生程度的管理人员的薪金高,而大学程度的非管理人员的薪金比研究生程度的非管理人员的薪金略低.当然,这是根据这家公司实际数据建立的模型得到的结果,并不具普遍性.

评注 从建立回归模型的角度我们通过本例介绍了以下内容:

1. 对于影响因变量的定性因素(管理、教育),可以引入0-1变量来处理,0-1变量的个数可比定性因素的水平少1(如教育程度有3个水平,引入2个0-1变量).

2. 用残差分析方法可以发现模型的缺陷,引入交互作用项常常能够给予改善.

3. 若发现异常值应剔除,有助于结果的合理性.

在本例中我们由简到繁,先分别引进管理和教育因素,再进入交互项.实际上,可以直接对6种管理-教育组合引入5个0-1变量,读者不妨试一下,看看结果如何.

10.3 酶促反应

背景和问题 酶是一种具有特异性的高效生物催化剂,绝大多数的酶是活细胞产生的蛋白质.酶的催化条件温和,在常温、常压下即可进行.酶催化的反应称为酶促反应,要比相应的非催化反应快 $10^3 \sim 10^{17}$ 倍.酶促反应动力学简称酶动力学,主要研究酶促反应的速度与底物(即反应物)浓度以及其他因素的关系.在底物浓度很低时酶促反应是一级反应;当底物浓度处于中间范围时,是混合级反应;当底物浓度增加时,向零级反应过渡.

某生化系学生为了研究嘌呤霉素在某项酶促反应中对反应速度与底物浓度之间关系的影响,设计了两个实验,一个实验中所使用的酶是经过嘌呤霉素处理的,而另一个实验所用的酶是未经嘌呤霉素处理过的,所得的实验数据见表1.试根据问题的背景和这些数据建立一个合适的数学模型,来反映这项酶促反应的速度与底物浓度以及嘌呤霉素处理与否之间的关系^[9].

表1 嘌呤霉素实验中的反应速度与底物浓度数据

底物浓度/ppm		0.02		0.06		0.11		0.22		0.56		1.10	
反应 速度	处理	76	47	97	107	123	139	159	152	191	201	207	200
	未处理	67	51	84	86	98	115	131	124	144	158	160	—

注:1 ppm = 0.001‰.

分析与假设 记酶促反应的速度为 y ,底物浓度为 x ,二者之间的关系写作 $y=f(x, \beta)$,其中 β 为参数.由酶促反应的基本性质可知,当底物浓度较小时,反

应速度大致与浓度成正比(即一级反应);而当底物浓度很大,渐进饱和时,反应速度将趋于一个固定值——最终反应速度(即零级反应).下面的两个简单模型具有这种性质:

Michaelis-Menten 模型

$$y = f(x, \beta) = \frac{\beta_1 x}{\beta_2 + x} \quad (1)$$

指数增长模型

$$y = f(x, \beta) = \beta_1 (1 - e^{-\beta_2 x}) \quad (2)$$

图 1 和图 2 分别是表 1 给出的经过嘌呤霉素处理和未经处理的反应速度 y 与底物浓度 x 的散点图,可以知道,模型(1),(2)与实际数据得到的散点图是符合的.下面只对模型(1)进行详细的分析,将模型(2)留给有兴趣的读者(习题 4).

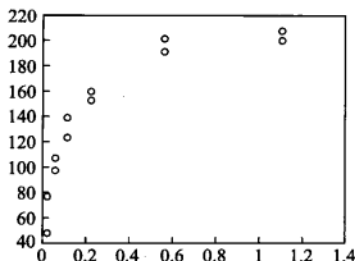


图 1 y 对 x (经处理) 的散点图

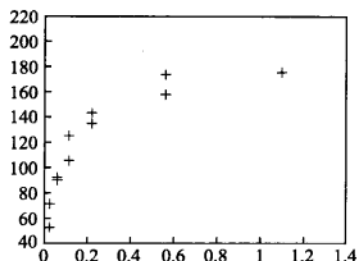


图 2 y 对 x (未经处理) 的散点图

首先对经过嘌呤霉素处理的实验数据进行分析(未经处理的数据可同样分析),在此基础上,再来讨论是否有更一般的模型来统一刻画处理前后的数据,进而揭示其中的联系.

线性化模型 模型(1)对参数 $\beta = (\beta_1, \beta_2)$ 是非线性的,但是可以通过下面的变量代换化为线性模型

$$\frac{1}{y} = \frac{1}{\beta_1} + \frac{\beta_2}{\beta_1} \frac{1}{x} = \theta_1 + \theta_2 u \quad (3)$$

模型(3)中的因变量 $1/y$ 对新的参数 $\theta = (\theta_1, \theta_2)$ 是线性的.

对经过嘌呤霉素处理的实验数据,作出反应速度的倒数 $1/y$ 与底物浓度的倒数 $u = 1/x$ 的散点图(图 3),可以发现在 $1/x$ 较小时有很好的线性趋势,而 $1/x$ 较大时则出现很大的起落.

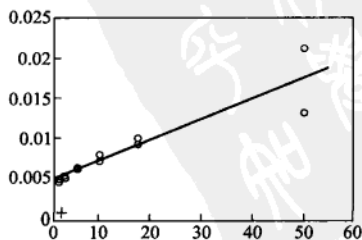


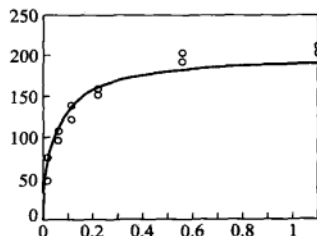
图 3 $1/y$ 与 $1/x$ 的散点图和回归直线

如果单从线性回归模型的角度作计算,很容易得到线性化模型(3)的参数 θ_1, θ_2 的估计和其他统计结果(见表2)以及 $1/y$ 与 $1/x$ 的拟合图(图3)。再根据(3)式中 β 与 θ 的关系得到 $\beta_1 = 1/\theta_1, \beta_2 = \theta_2/\theta_1$, 从而可以算出 β_1 和 β_2 的估计值分别为 $\hat{\beta}_1 = 195.8020$ 和 $\hat{\beta}_2 = 0.04840$ 。

表2 线性化模型(3)参数的估计结果

参数	参数估计值 ($\times 10^{-3}$)	参数置信区间 ($\times 10^{-3}$)
θ_1	5.107 2	[3.538 6, 6.675 8]
θ_2	0.247 2	[0.175 7, 0.318 8]
$R^2 = 0.8557 \quad F = 59.2975 \quad p < 0.0001 \quad s^2 = 3.5806 \times 10^{-6}$		

将经过线性化变换后最终得到的 β 值代入原模型(1),得到与原始数据比较的拟合图(图4)。可以发现,在 x 较大时 y 的预测值要比实际数据小,这是因为在对线性化模型作参数估计时,底物浓度 x 较低($1/x$ 很大)的数据在很大程度上控制了回归参数的确定,从而使得对底物浓度 x 较高数据的拟合,出现较大的偏差。



为了解决线性化模型中拟合欠佳的问题,我们直接考虑非线性模型(1)。

图4 用线性化得到的原始数据拟合图

非线性模型及求解 可以用非线性回归的方法直接估计模型(1)中的参数 β_1 和 β_2 。模型的求解可利用 MATLAB 统计工具箱中的命令进行,使用格式为

`[beta,R,J]=nlinfit(x,y,'model',beta0)`

其中输入 x 为自变量数据矩阵,每列一个变量; y 为因变量数据向量;`model` 为模型的 M 函数文件名,M 函数形式为 $y = f(\text{beta}, x)$, beta 为待估计参数;`beta0` 为给定的参数初值。输出 beta 为参数的估计值, R 为残差, J 为用于估计预测误差的 Jacobi 矩阵。参数 beta 的置信区间用命令 `nlparci(beta,R,J)` 得到。

我们用线性化模型(3)得到的 β 作为非线性模型参数估计的初始迭代值,将实际数据 x, y 输入后执行以下程序:

```
beta0=[195.802 0.0484];
[beta,R,J]=nlinfit(x,y,'huaxue',beta0);
betaci=nlparci(beta,R,J);
beta,betaci
```

```
yy = beta(1) * x ./ (beta(2) + x);
plot(x,y,'o',x,yy,'+'),pause
nlintool(x,y,'huaxue',beta)
```

```
function yhat = huaxue(beta,x)
yhat = beta(1) * x ./ (beta(2) + x);
```

得到的数值结果见表 3。

表 3 模型(1)参数的估计结果

参数	参数估计值	参数置信区间
β_1	212.681 8	[197.202 8, 228.160 8]
β_2	0.064 12	[0.046 7, 0.082 57]

拟合的结果直接画在原始数据图(图 5)上。程序中的 nlintool 用于给出一个交互式画面(图 6),拖动画面中的十字线可以改变自变量 x 的取值,直接得到因变量 y 的预测值和预测区间,同时通过左下方 Export 下拉式菜单,可输出模型的统计结果,如剩余标准差等,本例中剩余标准差 $s = 10.933\ 7$ 。

从上面的结果可以知道,对经过嘌呤霉素处理的实验数据,在用 Michaelis-Menten 模型(1)进行回归分析时,最终反应速度为 $\hat{\beta}_1 = 212.681\ 8$ 。还容易得到,反应的“半速度

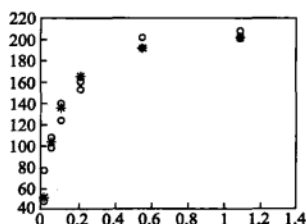


图 5 模型(1)的预测图
(o—原始数据;+—拟合结果)

点”(达到最终反应速度一半时的底物浓度 x 值)恰为 $\hat{\beta}_2 = 0.064\ 12$ 。以上结果对这样一个经过设计的实验(每个底物浓度做两次实验)已经很好地达到了要求。

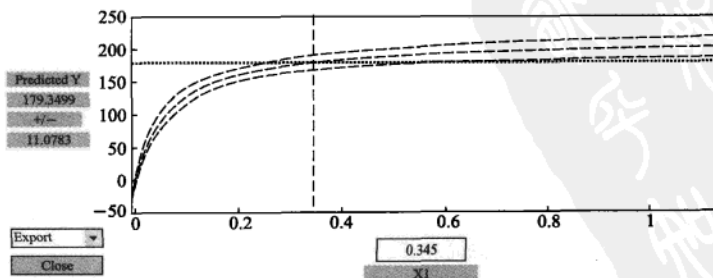


图 6 模型(1)的预测及结果输出

混合反应模型 酶动力学知识告诉我们,酶促反应的速度依赖于底物浓度,并且可以假定,嘌呤霉素的处理会影响最终反应速度参数 β_1 ,而基本不影响半速度参数 β_2 。表1的数据(图1、图2更为明显)也印证了这种看法。模型(1)的形式可以分别描述经过嘌呤霉素处理和未经处理的反应速度与底物浓度的关系(两个模型的参数 β 会不同),为了在同一个模型中考虑嘌呤霉素处理的影响,我们采用对未经嘌呤霉素处理的模型附加增量的方法,考察混合反应模型

$$y = f(x, \beta) = \frac{(\beta_1 + \gamma_1 x_2) x_1}{(\beta_2 + \gamma_2 x_2) + x_1} \quad (4)$$

其中自变量 x_1 为底物浓度(即模型(3)中的 x), x_2 为一示性变量(0-1变量),用来表示是否经嘌呤霉素处理,令 $x_2 = 1$ 表示经过处理, $x_2 = 0$ 表示未经处理;参数 β_1 是未经处理的最终反应速度, γ_1 是经处理后最终反应速度的增长值, β_2 是未经处理的反应的半速度点, γ_2 是经处理后反应的半速度点的增长值(为一般化起见,这里假定嘌呤霉素的处理也会影响半速度点)。

混合模型的求解和分析 仍用MATLAB统计工具箱中的命令nlinfit来计算模型(4)的回归系数 $\beta_1, \beta_2, \gamma_1$ 和 γ_2 。为了给出合适的初始迭代值,从实验数据我们注意到,未经处理的反应速度的最大实验值为160,经处理的最大实验值为207,于是可取参数初值 $\beta_1^0 = 170, \gamma_1^0 = 60$;又从数据可大致估计未经处理的半速度点约为0.05,经处理的半速度点约为0.06,我们取 $\beta_2^0 = 0.05, \gamma_2^0 = 0.01$ 。

与模型(1)的编程计算相似,得到混合模型(4)的回归系数的估计值与其置信区间(表4)、拟合结果(图7)、残差图(图8),及预测和结果输出图(图9),模型的剩余标准差 $s = 10.4000$ 。

表4 模型(4)参数的估计结果

参数	参数估计值	参数置信区间
β_1	160.280 2	[145.846 6, 174.713 7]
β_2	0.047 7	[0.030 4, 0.065 0]
γ_1	52.403 5	[32.413 0, 72.394 1]
γ_2	0.016 4	[-0.007 5, 0.040 3]

然而,从表4可以发现, γ_2 的置信区间包含零点,这表明参数 γ_2 对因变量 y 的影响并不显著,这一结果与前面的说法(即嘌呤霉素的作用不影响半速度参数)是一致的。因此,可以考虑简化模型

$$y = f(x, \beta) = \frac{(\beta_1 + \gamma_1 x_2) x_1}{\beta_2 + x_1} \quad (5)$$

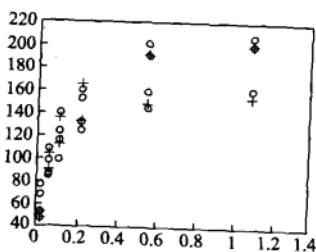


图7 模型(4)的预测图

(○—原始数据; +—拟合结果)

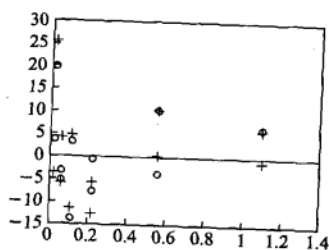


图8 模型(4)残差图

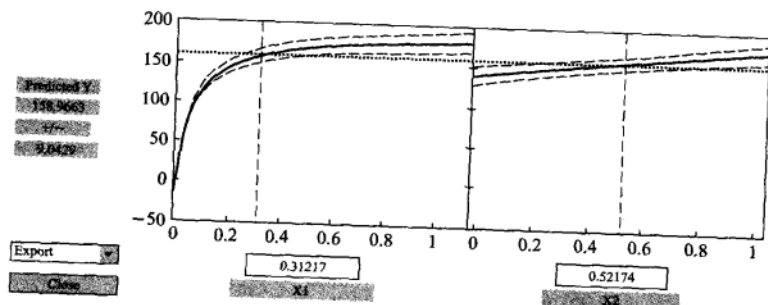


图9 模型(4)的预测及结果输出

采用与模型(4)类似的计算、分析方法,模型(5)的结果概括在表5和表6,以及图10、图11和图12中,模型(5)的剩余标准差 $s = 10.5851$.

表5 模型(4)参数的估计结果

参数	参数估计值	参数置信区间
β_1	166.6025	[154.4886, 178.7164]
β_2	0.0580	[0.0456, 0.0703]
γ_1	42.0252	[28.9419, 55.1085]

表6 模型(4)与模型(5)预测值与预测区间的比较(预测区间为预测值 $\pm \Delta$)

实际数据	模型(4)预测值	Δ (模型(4))	模型(5)预测值	Δ (模型(5))
67	47.3443	9.2078	42.7358	5.4446
51	47.3443	9.2078	42.7358	5.4446
84	89.2856	9.5710	84.7356	7.0478
86	89.2856	9.5710	84.7356	7.0478

续表

实际数据	模型(4)预测值	$\Delta(\text{模型}(4))$	模型(5)预测值	$\Delta(\text{模型}(5))$
98	111.793 8	7.754 6	109.105 3	7.028 1
115	111.793 8	7.754 6	109.105 3	7.028 1
131	131.716 6	7.500 7	131.858 6	7.587 8
124	131.716 6	7.500 7	131.858 6	7.587 8
144	147.697 3	10.372 9	150.974 3	9.442 3
158	147.697 3	10.372 9	150.974 3	9.442 3
160	153.617 6	12.119 7	158.262 3	10.562 1
76	50.566	7.691 4	53.515 8	6.740 9
47	50.566	7.691 4	53.515 8	6.740 9
97	102.811	9.564 3	106.110 1	8.236 8
107	102.811	9.564 3	106.110 1	8.236 8
123	134.361 6	8.252 2	136.627 0	7.422 3
139	134.361 6	8.252 2	136.627 0	7.422 3
159	164.684 7	7.029 4	165.119 7	7.059 5
152	164.684 7	7.029 4	165.119 7	7.059 5
191	190.832 9	9.148 4	189.057 4	8.843 8
201	190.832 9	9.148 4	189.057 4	8.843 8
207	200.968 8	11.044 7	198.183 7	10.181 2
200	200.968 8	11.044 7	198.183 7	10.181 2

混合模型(4)和(5)不仅有类似于模型(1)的实际解释,同时把嘌呤霉素处理前后酶促反应的速度之间的变化体现在模型之中,因此它们比单独的模型具有更实际的价值.另外,虽然模型(5)的某些统计指标可能没有模型(4)的好,比如模型(5)的剩余标准差略大于模型(4),但由于它的形式更简单明了,易于实际中的操作和控制,而且从表6中数据可以发现,虽然两个模型的预测值相差不大,但模型(5)预测区间的长度明显比模型(4)的短.因此,总体来说模型(5)更为优良.

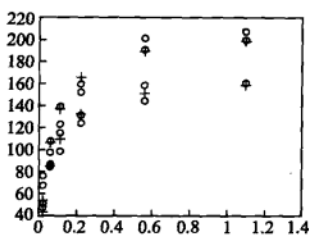


图 10 模型(5)的预测图

(○—原始数据; +—拟合结果)

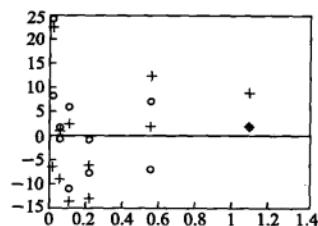


图 11 模型(5)残差图

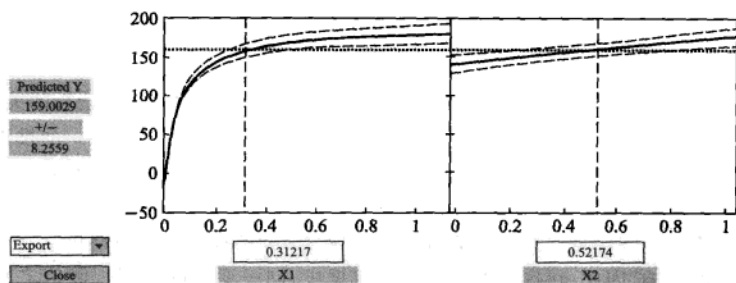


图 12 模型(5)的预测及结果输出

可进一步研究的模型 假如在实验中当底物浓度增加到一定程度后,反应速度反而有轻微的下降(在本例中只有一个数据点如此),那么可以考虑模型

$$y = f(x, \beta) = \frac{\beta_1 x}{\beta_2 + x + \beta_3 x^2} \quad (6)$$

或引入混合模型

$$y = f(x, \beta) = \frac{(\beta_1 + \gamma_1 x_2) x_1}{(\beta_2 + \gamma_2 x_2) + x_1 + (\beta_3 + \gamma_3 x_2) x_1^2} \quad (7)$$

有兴趣的读者可以尝试一下,会发现用这些模型可以改善模型(4)和(5)的残差图(图8、图11)中表现出来的在各个浓度下残差散布不均匀的现象。

评注 无论从机理分析,还是从实验数据看,酶促反应中反应速度与底物浓度及嘌呤霉素的作用之间的关系都是非线性的,本节我们先用线性化模型来简化参数估计,如果这样能得到满意的结果,当然很好,但是由于变量的代换已经隐含了误差扰动项的变换,因此,除非变换后的误差项仍具有常数方差,一般情况下我们还需要采用原始数据做非线性回归,而把线性化模型的参数估计结果作为非线性模型参数估计的迭代初值。

应该指出,在非线性模型参数估计中,用不同的参数初值进行迭代,可能得到差别很大的结果(它们都是拟合误差平方和的局部极小点),也可能出现收敛速度等问题,因此,合适的初值是非常重要的。

另外,评价线性回归模型拟合程度的统计检验无法直接用于非线性模型。例如, F 统计量不能用于非线性模型拟合程度的显著性检验,因为即使误差项服从均值为0的正态分布,也无法从回归残差得到误差方差的一个无偏估计。但是 R^2 和剩余标准差 s 仍然可以在通常意义下用于非线性回归模型拟合程度的度量。

从本例还可以看到,通过引入示性变量,能够描述定性上不同的处理水平对模型参数的影响,这是一种直接明了的建模方法。

10.4 投资额与生产总值和物价指数

问题 为研究某地区实际投资额与国民生产总值(GNP)及物价指数的关系,收集了该地区连续20年的统计数据(见表1),目的是由这些数据建立一个投资额的模型,根据对未来国民生产总值及物价指数的估计,预测未来的实际投资额。

表1的数据是以时间为序的,称**时间序列**。由于投资额、国民生产总值、物价指数等许多经济变量均有一定的滞后性,比如,前期的投资额对后期投资额一般有明显的影响。因此,在这样的时间序列数据中,同一变量的顺序观测值之间的出现相关现象(称**自相关**)是很自然的。然而,一旦数据中存在这种自相关序列,如果仍采用普通的回归模型直接处理,将会出现不良后果,其预测也会失去意义,为此,我们必须先来诊断数据是否存在自相关,如果存在,就要考虑自相关关系,建立新的回归模型^[59,79]。

表1 某地区实际投资额(亿元)与国民生产总值(亿元)及物价指数数据

年份 序号	投资额	国民生 产总值	物价 指数	年份 序号	投资额	国民生 产总值	物价 指数
1	90.9	596.7	0.716 7	11	229.8	1 326.4	1.057 5
2	97.4	637.7	0.727 7	12	228.7	1 434.2	1.150 8
3	113.5	691.1	0.743 6	13	206.1	1 549.2	1.257 9
4	125.7	756.0	0.767 6	14	257.9	1 718.0	1.323 4
5	122.8	799.0	0.790 6	15	324.1	1 918.3	1.400 5
6	133.3	873.4	0.825 4	16	386.6	2 163.9	1.504 2
7	149.3	944.0	0.867 9	17	423.0	2 417.8	1.634 2
8	144.2	992.7	0.914 5	18	401.9	2 631.7	1.784 2
9	166.4	1 077.6	0.960 1	19	474.9	2 954.7	1.951 4
10	195.0	1 185.9	1.000 0	20	424.5	3 073.0	2.068 8

基本的回归模型 首先仿照前几节的方法建立普通的回归模型. 记该地区第 t 年的投资额为 y_t , 国民生产总值为 x_{1t} , 物价指数为 x_{2t} (以第 10 年的物价指数为基准, 基准值为 1), $t = 1, 2, \dots, n$ ($n = 20$). 因变量 y_t 与自变量 x_{1t} 和 x_{2t} 的散点图见图 1 和图 2.

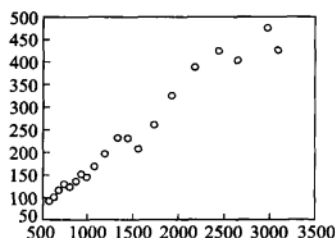


图 1 y_t 对 x_{1t} 的散点图

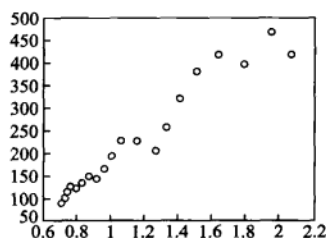


图 2 y_t 对 x_{2t} 的散点图

可以看出,随着国民生产总值的增加,投资额增大,而且两者有很强的线性关系,物价指数与投资额的关系也类似,因此可建立多元线性回归模型

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t \quad (1)$$

模型(1)中除了国民生产总值和物价指数外,影响 y_t 的其他因素的作用都包含在随机误差 ε_t 内,这里假设 ε_t (对 t) 相互独立,且服从均值为 0 的正态分布, $t = 1, 2, \dots, n$. 与前几节不同的是,为了后面模型记号的考虑,这里的变量均加了下标 t .

根据表 1 的数据,对模型(1)直接利用 MATLAB 统计工具箱求解,得到的回归系数估计值及其置信区间(置信水平 $\alpha = 0.05$)、检验统计量 R^2, F, p, s^2 的结果见表 2.

表 2 模型(1)的计算结果

参数	参数估计值	参数置信区间
β_0	322.725 0	[224.338 6, 421.111 4]
β_1	0.618 5	[0.477 3, 0.759 6]
β_2	-859.479 0	[-1 121.475 7, -597.482 3]
$R^2 = 0.990 8$	$F = 919.852 9$	$p < 0.000 1$
		$s^2 = 161.7$

将参数估计值代入(1)得到

$$\hat{y}_t = 322.725 + 0.618 5 x_{1t} - 859.479 x_{2t} \quad (2)$$

用 MATLAB 中 `rstool` 命令得到的交互式画面见图 3,由此可以给出不同水平下的预测值及其置信区间,通过左下方 Export 下拉式菜单,可输出模型的统计结果,如剩余标准差 $s = 12.716 4$.

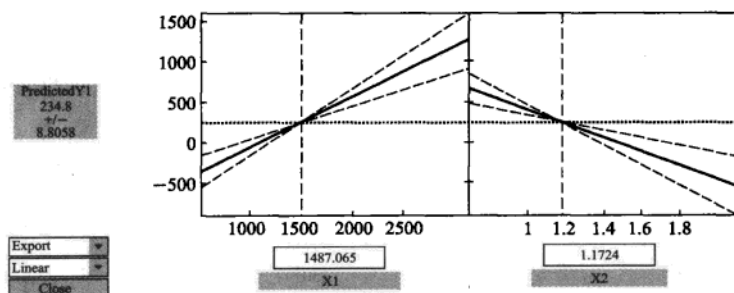


图3 模型(2)的输出

自相关性诊断与处理方法 从表面上看得到的基本模型(2)的拟合度非常高($R^2=0.9908$),应该很满意了.但是,这个模型并没有考虑到我们的数据是一个时间序列(将表1的年份序号打乱,不影响模型(2)的结果).实际上,在对时间序列数据做回归分析时,模型的随机误差项 ε_t 有可能存在相关性,违背模型关于 ε_t (对时间 t) 相互独立的基本假设.如在投资额模型中,国民生产总值和物价指数之外的因素(比如政策等因素)对投资额的影响包含在随机误差 ε_t 中,如果它的影响成为 ε_t 的主要部分,则由于政策等因素的连续性,它们对投资额的影响也有时间上的延续,即随机误差 ε_t 会出现(自)相关性.

残差 $e_t = y_t - \hat{y}_t$ 可以作为随机误差 ε_t 的估计值,画出 $e_t - e_{t-1}$ 的散点图,能够从直观上判断 ε_t 的自相关性.模型(2)的残差 e_t 可在计算过程中得到,如表3,数据 $e_t - e_{t-1}$ 散点图如图4,可以看到,大部分点子落在第1,3象限,表明 ε_t 存在正的自相关.

表3 模型(2)的残差 e_t

t	1	2	3	4	5
e_t	15.130 6	5.728 1	2.468 2	-4.842 1	-14.567 7
t	6	7	8	9	10
e_t	-20.171 0	-11.306 2	-6.473 3	2.411 9	-1.673 7
t	11	12	13	14	15
e_t	-4.346 8	8.072 9	6.400 6	10.101 0	18.690 0
t	16	17	18	19	20
e_t	18.425 0	9.531 1	-14.934 9	2.008 5	-20.652 1

为了对 ε_t 的自相关性作定量诊断,并在确诊后得到新的结果,我们考虑如下的模型:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t, \quad \varepsilon_t = \rho \varepsilon_{t-1} + u_t \quad (3)$$

其中 ρ 是自相关系数, $|\rho| \leq 1$, u_t 相互独立且服从均值为零的正态分布, $t=1, 2, \dots, n$.

模型(3)中若 $\rho=0$, 则退化为普通的回归模型;若 $\rho>0$, 则随机误差 ε_t 存在正的自相关;若 $\rho<0$, 存在负的自相关. 大多数与经济有关的时间序列数据,在经济规律作用下,一般随着时间的推移有一种向上或向下变动的趋势,其随机误差表现出正相关趋势.

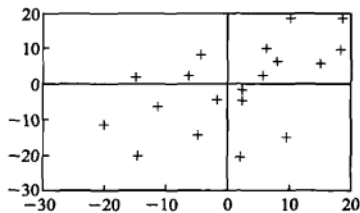


图4 模型(2) $e_t - e_{t-1}$ 的散点图

Durbin-Watson 检验(以下简称 D-W 检验)是一种常用的诊断自相关现象的统计方法. 首先根据模型(1)得到的残差计算 DW 统计量:

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=2}^n e_t^2} \quad (4)$$

经过简单的运算可知,当 n 较大时,

$$DW \approx 2 \left(1 - \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=2}^n e_t^2} \right) \quad (5)$$

而(5)式右端的 $\sum_{t=2}^n e_t e_{t-1} / \sum_{t=2}^n e_t^2$ 正是自相关系数 ρ 的估计值 $\hat{\rho}$, 于是

$$DW \approx 2(1 - \hat{\rho}) \quad (6)$$

由于 $-1 \leq \hat{\rho} \leq 1$, 所以 $0 \leq DW \leq 4$, 并且,若 $\hat{\rho}$ 在 0 附近,则 DW 在 2 附近, ε_t 的自相关性很弱(或不存在自相关);若 $\hat{\rho}$ 在 ± 1 附近,则 DW 接近 0 或 4, ε_t 的自相关性很强.

要根据 DW 的具体数值确定 ε_t 是否存在自相关,应该在给定的检验水平下,依照样本容量和回归变量数目,查 D-W 分布表,得到检验的临界值 d_L 和 d_U , 然后由图 5 中 DW 所在的区间来决定^[79].

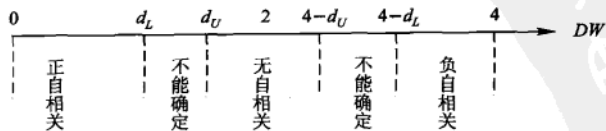


图5 与 DW 值对应的自相关状态

如果检验结果判定存在自相关,就应该采用模型(3),其中 ρ 可由(6)式估计,即

$$\hat{\rho} = 1 - \frac{DW}{2} \quad (7)$$

作变换

$$y_i^* = y_i - \rho y_{i-1}, x_{1i}^* = x_{1i} - \rho x_{1i-1}, i = 1, 2 \quad (8)$$

则模型(3)化为

$$y_i^* = \beta_0^* + \beta_1 x_{1i}^* + \beta_2 x_{2i}^* + u_i, \quad \beta_0^* = \beta_0(1 - \rho) \quad (9)$$

其中 u_i 相互独立且服从均值为0的正态分布,所以(9)是普通的回归模型。

加入自相关后的模型 利用表3给出的残差 e_i ,根据(4)式计算出 $DW = 0.8754$,对于显著性水平 $\alpha = 0.05, n = 20, k = 3$ (回归变量包括常数项的数目),查D-W分布表^[79],得到检验的临界值 $d_L = 1.10$ 和 $d_U = 1.54$ 。现在 $DW < d_L$,由图5可以认为随机误差存在正自相关,且 ρ 可由(7)式估计得 $\hat{\rho} = 0.5623$ 。

以 ρ 的估计值代入(8)式作变换,利用变换后的数据 $y_i^*, x_{1i}^*, x_{2i}^*$ 估计模型(9)的参数,得到的结果见表4。

表4 模型(9)的计算结果

参数	参数估计值	参数置信区间
β_0^*	163.490 5	[1 265.459 2, 2 005.217 8]
β_1	0.699 0	[0.575 1, 0.824 7]
β_2	-1 009.033 3	[-1 235.939 2, -782.127 4]
$R^2 = 0.977 2$		$F = 342.898 8$
		$p < 0.000 1$
		$s^2 = 96.58$

当然应该对模型(9)也作一次自相关检验,即诊断随机误差 u_i 是否还存在自相关。从模型(9)的残差可以计算出 $DW = 1.5751$,对于显著性水平 $\alpha = 0.05, n = 19, k = 3$,检验的临界值为 $d_L = 1.08$ 和 $d_U = 1.53$ 。现在 $d_U < DW < 4 - d_U$,由图5可以认为随机误差不存在自相关。因此,经变换(8)得到的回归模型(9)是适用的。

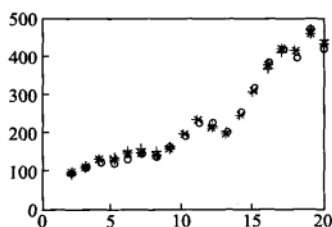
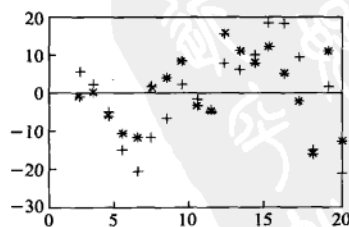
最后,将模型(9)中的 $y_i^*, x_{1i}^*, x_{2i}^*$ 还原为原始变量 y_i, x_{1i}, x_{2i} ,得到的结果为

$$\hat{y}_i = 163.490 5 + 0.562 3 y_{i-1} + 0.699 x_{1i} - 0.393 0 x_{1i-1} - 1 009.033 3 x_{2i} + 567.379 4 x_{2i-1} \quad (10)$$

结果分析及预测 从机理上看,对于带滞后性的经济规律作用下的时间序列数据,加入自相关的模型(10)更为合理,而且在本例中,衡量与实际数据拟合程度的指标——剩余标准差从原模型(2)的12.7164减小到9.8277。我们将模型(10)、模型(2)的计算值 \hat{y}_i 与实际数据 y_i 的比较,以及两个模型的残差 e_i ,表示在表5和图6、图7上,可以看出模型(10)更合适些。

表5 模型(10)、模型(2)的计算值 \hat{y}_t 与残差 e_t

t	y_t (实际数据)	\hat{y}_t (模型(10))	\hat{y}_t (模型(2))	e_t (模型(10))	e_t (模型(2))
2	97.400 0	98.220 3	91.699 6	-0.820 3	5.700 4
3	113.500 0	113.286 4	111.061 8	0.213 6	2.438 2
4	125.700 0	131.522 8	130.574 9	-5.822 8	-4.874 9
5	122.800 0	133.343 5	137.402 4	-10.543 5	-14.602 4
6	133.300 0	144.754 9	153.508 9	-11.454 9	-20.208 9
7	149.300 0	147.630 1	160.647 2	1.669 9	-11.347 2
8	144.200 0	140.015 1	150.716 4	4.184 9	-6.516 4
9	166.400 0	157.781 3	164.034 8	8.618 7	2.365 2
10	195.000 0	198.212 5	196.725 1	-3.212 5	-1.725 1
11	229.800 0	234.560 9	234.204 4	-4.760 9	-4.404 4
12	228.700 0	212.746 2	220.689 3	15.953 8	8.010 7
13	206.100 0	195.016 3	199.766 6	11.083 7	6.333 4
14	257.900 0	249.779 2	247.873 5	8.120 8	10.026 5
15	324.100 0	311.944 6	305.493 2	12.155 4	18.606 8
16	386.600 0	381.233 5	368.268 8	5.366 5	18.331 2
17	423.000 0	424.995 5	413.573 7	-1.995 5	9.426 3
18	401.900 0	417.601 0	416.949 0	-15.701 0	-15.049 0
19	474.900 0	463.847 4	473.019 6	11.052 6	1.880 4
20	424.500 0	437.053 4	445.285 3	-12.553 4	-20.785 3

图6 模型(10)、(2)的预测图
($\circ - y_t$; $* - \hat{y}_t(10)$; $+- \hat{y}_t(2)$)图7 模型(10)、(2)的残差图
($* - e_t(10)$; $+- e_t(2)$)

当用模型(10)或(2)对未来的投资额 y_t 作预测时,需先估计未来的国民生产总值 x_{1t} , 物价指数 x_{2t} , 比如, 设 $t=21$ 时 $x_{1t} = 3\ 312$, $x_{2t} = 2.193\ 8$, 容易由模型(10)得到 $\hat{y}_t = 469.763\ 8$, 由模型(2)得到 $\hat{y}_t = 485.672\ 0$. 模型(10)的 \hat{y}_t 较小, 是由于上一年的实际数据 $y_{t-1} = 424.5$ 过小的作用所致, 而在模型(2)中它不出现.

评注 经 D-W 检验认为普通回归模型(2)的随机误差存在自相关, 由(4), (7)式估计出自相关系数 ρ 后, 采用变换(8)的方法得到模型(9), 称为广义差分法. 这种方法消除了原模型随机误差的自相关性, 得到的(9)式是一阶自相关模型.

D-W 检验和广义差分法在经济数据建模中有着广泛的应用, 但是也存在明显的不足: 若 DW 的数值落在无法确定自相关性的区间, 则只能设法增加数据量, 或选用其他方法; 如果原始数据序列存在高阶自相关, 则需反复使用 D-W 检验和广义差分法, 直至判定不存在自相关为止. 另外, D-W 分布表中数据容量 n 的下限是 15.

10.5 教学评估

问题 为了考评教师的教学质量, 教学研究部门设计了一个教学评估表, 对学生进行一次问卷调查, 要求学生对 12 位教师的 15 门课程(其中 3 位教师有两门课)按以下 7 项内容打分, 分值为 1~5 分(5 分最好, 1 分最差):

X_1 ——课程内容组织的合理性; X_2 ——主要问题展开的逻辑性; X_3 ——回答学生问题的有效性; X_4 ——课下交流的有助性; X_5 ——教科书的帮助性; X_6 ——考试评分的公正性; Y ——对教师的总体评价.

收回问卷调查表后, 得到了学生对 12 位教师、15 门课程各项评分的平均值, 见表 1.

表 1 12 位教师、15 门课程各项评分的平均值

教师 编号	课程 编号	X_1	X_2	X_3	X_4	X_5	X_6	Y
1	201	4.46	4.42	4.23	4.10	4.56	4.37	4.11
2	224	4.11	3.82	3.29	3.60	3.99	3.82	3.38
3	301	3.58	3.31	3.24	3.76	4.39	3.75	3.17
4	301	4.42	4.37	4.34	4.40	3.63	4.27	4.39
5	301	4.62	4.47	4.53	4.67	4.63	4.57	4.69

续表

教师编号	课程编号	X_1	X_2	X_3	X_4	X_5	X_6	Y
6	309	3.18	3.82	3.92	3.62	3.50	4.14	3.25
7	311	2.47	2.79	3.58	3.50	2.84	3.84	2.84
8	311	4.29	3.92	4.05	3.76	2.76	4.11	3.95
9	312	4.41	4.36	4.27	4.75	4.59	4.11	4.18
10	312	4.59	4.34	4.24	4.39	2.64	4.38	4.44
11	333	4.55	4.45	4.43	4.57	4.45	4.40	4.47
12	424	4.67	4.64	4.52	4.39	3.48	4.21	4.61
3	351	3.71	3.41	3.39	4.18	4.06	4.06	3.17
4	411	4.28	4.45	4.10	4.07	3.76	4.43	4.15
9	424	4.24	4.38	4.35	4.48	4.15	4.50	4.33

教学研究部门认为,所列各项具体内容 $X_1 \sim X_6$ 不一定每项都对教师总体评价 Y 有显著影响,并且各项内容之间也可能存在很强的相关性,他们希望得到一个总体评价与各项具体内容之间的模型,这个模型应尽量简单和有效,并且由此能给教师一些合理的建议,以提高总体评价^[23]。

关于逐步回归 虽然给出了6个自变量,但是我们希望从中挑选出对因变量 Y 影响显著的那些来建立回归模型。变量选择的标准,应该是将所有对因变量影响显著的自变量都选入模型,而影响不显著的自变量都不选入模型,从便于应用的角度应使模型中自变量个数尽可能少。逐步回归就是一种从众多自变量中有效地选择重要变量的方法。

逐步回归的基本思路是,先确定一个包含若干自变量的初始集合,然后每次从集合外的变量中引入一个对因变量影响最大的,再对集合中的变量进行检验,从变得不显著的变量中移出一个影响最小的,依此进行,直到不能引入和移出为止。引入和移出都以给定的显著性水平为标准。

MATLAB 统计工具箱中的逐步回归命令是 `stepwise`,它提供人机交互式画面,研究者可以在画面上自由地引入和移出变量,进行统计分析,其通常用法是

`stepwise(x,y,inmodel,penter,premove)`

其中 x 是自变量数据,排成 $n \times m$ 矩阵(m 为自变量个数, n 为每个变量的数据量), y 是因变量数据,排成 n 维向量, $inmodel$ 是自变量初始集合的指标(即矩阵 x 中哪些列进入初始集合),缺省时设定为没有选取任何 x 的列向量, $penter$ 是引

入变量时设定的最大 p 值,缺省时为 0.05, remove 是移出变量时设定的最小 p 值,缺省时为 0.10. remove 值不能小于 penter 值.

回归模型的建立 将表 1 的数据用 stepwise(x,y)命令(6 个自变量都没有进入初始模型)得到 Stepwise Regression 的初始界面窗口,见图 1.

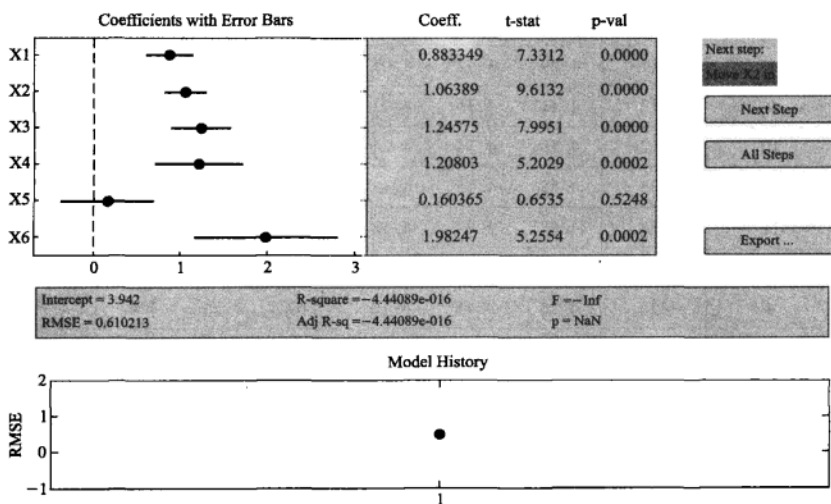


图 1 Stepwise Regression 初始界面

界面的左上方给出了所有 6 个变量的回归系数的估计及误差界(用水平线表示),其中在计算机屏幕上彩色的水平线表示置信度为 90% 的置信区间,灰色的为其 95% 的置信区间.水平线若为红色,则表示该水平线所对应的变量没有被选入模型中.界面上方中间部分的表格显示的是一旦某个变量被选入到模型中时,该变量的回归系数的估计值、检验的 t 统计量值以及 p 值.一般来说,在每一步中选出的是具有最小 p 值或最大统计量值的项,在图 1 中即为自变量 X_2 .一般只需按 Next Step 按钮进行下一步操作,程序就会自动选择所要引入或移出的自变量,并在界面的右上方给出相应的结果.当然也可以手动操作,用鼠标点击表中的一行,改变其状态,即目前不在模型中的一个变量(红色的行)被引入(变蓝),目前在模型中的一个变量(蓝色的行)被移出(变红),直到界面提示 Move no terms 为止.通常我们可以直接按 All Steps 按钮来完成整个模型的逐步回归过程.如在图 1 中按 All Steps 按钮,可得到逐步回归的最终结果,见图 2.

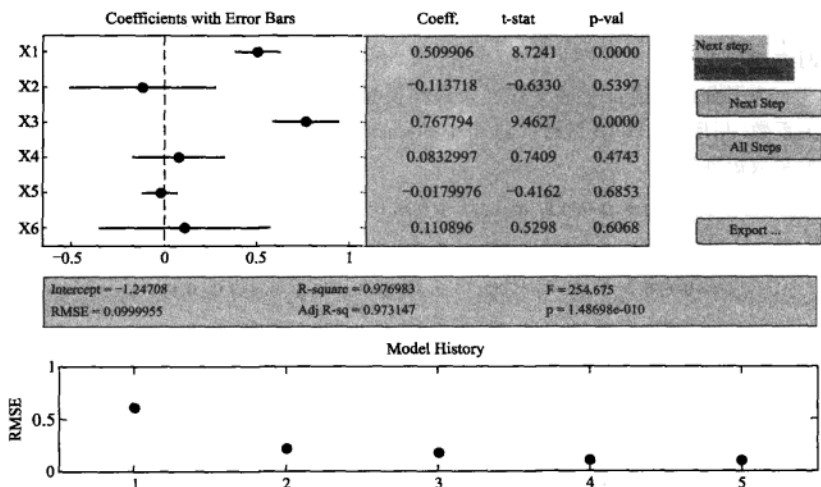


图2 Stepwise Regression 最终界面(按 All Steps 按钮所得)

Stepwise Regression 界面分为上中下三个部分,界面的上面部分已经做过介绍.界面的中间部分的表格给出了该回归模型的所有计算结果,包括 Intercept (截距,即回归常数),决定系数 R^2 ,检验的 F 值, RMSE (剩余标准差),调整的决定系数 R^2 以及检验的 p 值,这里调整的决定系数 R^2 定义为 $1 - \frac{n-1}{n-k-1}(1-R^2)$,其中 k 为选入模型的自变量个数.

界面的下面部分 Model History,给出了逐步回归中每步所对应的模型的剩余标准差的点图(蓝色的点),将鼠标移至某步对应的蓝色点,会显示此步对应的模型中所包含的自变量,点击该蓝色的点可追踪当前模型所对应的界面.

对于我们的问题,从图2可以看出,最终入选的自变量只有 X_1, X_3 (显示为蓝色),通过分析可知,随着逐步回归的进行,每步模型所对应剩余标准差(RMSE)逐渐减少到0.1,虽然在个别步中 R^2 略有下降,但模型的 F 值却大大提高.这些都表明,仅含 X_1, X_3 的模型是合适的,由 X_1, X_3 的回归系数和回归常数利用逐步回归最终得到的模型为

$$Y = 0.5099X_1 + 0.7678X_3 - 1.2471 \quad (1)$$

模型解释 在最终模型里回归变量只有 X_1, X_3 ,是一个简单易用的模型,据此可把课程内容组织的合理性(X_1)和回答学生问题的有效性(X_3)列入考评的重点.模型(1)表明, X_1 的分值每增加一分,对教师的总体评价就增加约0.5分;

X_3 的分值每增加一分,对教师的总体评价就增加约 0.77 分. 应建议教师注重这两方面的工作.

为了分析其他自变量没有进入最终模型的原因,可以计算 $X_1 \sim X_6, Y$ 的相关系数. 利用 MATLAB 统计工具箱中的 `corrcoef` 命令直接得到这 7 个变量的相关系数矩阵

```
1.000 0 0.900 8 0.675 2 0.736 1 0.291 0 0.647 1 0.897 3
0.900 8 1.000 0 0.850 4 0.739 9 0.277 5 0.802 6 0.936 3
0.675 2 0.850 4 1.000 0 0.749 9 0.080 8 0.849 0 0.911 6
0.736 1 0.739 9 0.749 9 1.000 0 0.437 0 0.704 1 0.821 9
0.291 0 0.277 5 0.080 8 0.437 0 1.000 0 0.187 2 0.178 3
0.647 1 0.802 6 0.849 0 0.704 1 0.187 2 1.000 0 0.824 6
0.897 3 0.936 3 0.911 6 0.821 9 0.178 3 0.824 6 1.000 0
```

一般认为,两个变量的相关系数超过 0.85 时才具有显著的相关关系. 由上面结果知道,与 Y 相关关系显著的只有 X_1, X_2, X_3 , 而 X_2 未进入最终模型,是由于它与 X_1, X_3 的相关关系显著(相关系数 $r_{12} = 0.9008, r_{32} = 0.8504$),可以说,模型中有了 X_1, X_3 以后,变量 X_2 是多余的,应该去掉.

评注 如果初步看来影响因变量的因素较多,并得到了大量的数据,为了建立一个有效的、便于应用的模型,可以利用逐步回归只选择那些影响显著的变量“入围”.

如果怀疑原有变量的平方项、交互项等也会对因变量有显著影响,也可以将这些项作为新的自变量加入到候选行列,用逐步回归处理. 这样,10.1 ~ 10.3 节中模型的变量选择,甚至 10.4 节时间序列中自相关项的选择,都可以用逐步回归的方法去做,有兴趣的读者不妨一试.

10.6 冠心病与年龄

问题 冠心病(Coronary Heart Disease,简称 CHD)是一种常见的心脏疾病,严重地危害着人类的健康. 到目前为止,其病因尚未完全研究清楚,医学界普遍认同的、重要的易患因素是高龄、高血脂、高血压、糖尿病、动脉粥样硬化及家族史等. 多项研究表明,冠心病发病率随着年龄的增加而上升,在冠心病的流行病学研究中,年龄也是最常见的混杂因素之一.

为了更好地说明冠心病发病率与年龄的关系,医学家们对 100 名不同年龄的人进行观察,表 1 中给出了这 100 名被观察者的年龄及他们是否患冠心病的

数据(其中冠心病一栏中,1 代表被观察者患冠心病,0 代表不患冠心病). 本节的目的是根据这些数据建立数学模型,来分析冠心病发病率与年龄的关系,并进行统计预测^[36].

表 1 100 名被观察者的年龄与是否患冠心病的观察数据

序号	年龄	冠心病	序号	年龄	冠心病	序号	年龄	冠心病	序号	年龄	冠心病
1	20	0	26	35	0	51	44	1	76	55	1
2	23	0	27	35	0	52	44	1	77	56	1
3	24	0	28	36	0	53	45	0	78	56	1
4	25	0	29	36	1	54	45	1	79	56	1
5	25	1	30	36	0	55	46	0	80	57	0
6	26	0	31	37	0	56	46	1	81	57	0
7	26	0	32	37	1	57	47	0	82	57	1
8	28	0	33	37	0	58	47	0	83	57	1
9	28	0	34	38	0	59	47	1	84	57	1
10	29	0	35	38	0	60	48	0	85	57	1
11	30	0	36	39	0	61	48	1	86	58	0
12	30	0	37	39	1	62	48	1	87	58	1
13	30	0	38	40	0	63	49	0	88	58	1
14	30	0	39	40	1	64	49	0	89	59	1
15	30	0	40	41	0	65	49	1	90	59	1
16	30	1	41	41	0	66	50	0	91	60	0
17	32	0	42	42	0	67	50	1	92	60	1
18	32	0	43	42	0	68	51	0	93	61	1
19	33	0	44	42	0	69	52	0	94	62	1
20	33	0	45	42	1	70	52	1	95	62	1
21	34	0	46	43	0	71	53	1	96	63	1
22	34	0	47	43	0	72	53	1	97	64	0
23	34	1	48	43	1	73	54	1	98	64	1
24	34	0	49	44	0	74	55	0	99	65	1
25	34	0	50	44	0	75	55	1	100	69	1

分析与假设 假设这 100 名被观察者是独立选取的,记 x 为被观察者的年龄, Y 为被观察者患冠心病的情况 ($Y=1$ 表示患冠心病, $Y=0$ 表示未患冠心病),显然, Y 是一个 0-1 变量. 利用表 1 的数据作出 Y 对 x 的散点图(见图 1).

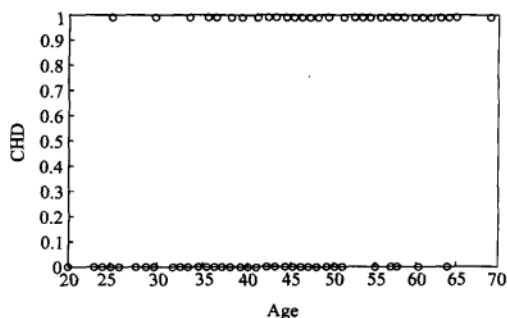


图 1 冠心病患病人数与年龄的散点图

从图 1 容易看出,直接对上述数据建立像前面几节那样的回归模型是行不通的,需要对数据进行预处理.

数据预处理的一种方法是将被观察者按年龄段进行分组,并统计各年龄段中患冠心病的人数,及患病人数占该年龄段总人数的比例(以下简称患病比例).为方便起见,我们将年龄分成 8 个年龄段,分段后的数据见表 2.

表 2 各年龄段的冠心病患病人数及比例

年龄段	年龄段中点	人数	患冠心病人数	患病比例
20 ~ 29	24.5	10	1	0.1
30 ~ 34	32	15	2	0.13
35 ~ 39	37	12	3	0.25
40 ~ 44	42	15	5	0.33
45 ~ 49	47	13	6	0.46
50 ~ 54	52	8	5	0.63
55 ~ 59	57	17	13	0.76
60 ~ 69	64.5	10	8	0.80
合计		100	43	0.43

为考察患病比例与年龄的关系,首先根据表 2 数据做出患病比例对各年龄段中点的散点图(见图 2,为方便起见,散点的横坐标均简单地取各年龄段的中点).

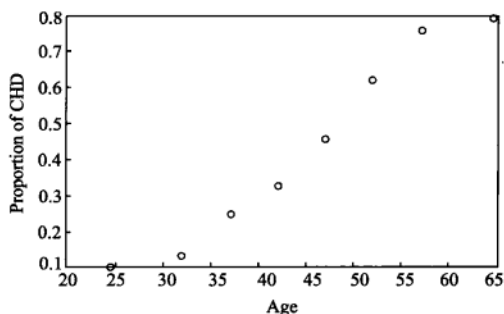


图2 冠心病患病比例对年龄段中点的散点图

从图2可以看出,冠心病患病比例随年龄的增大而递增,大致是一条介于0与1之间的S型曲线,这条曲线应该怎样用回归方程来确定呢?表2和图2中的患病比例实际上就是年龄为 x 时(以下均取年龄段的中点) Y 的平均值,用(条件)期望的符号记作

$$y = E(Y | x) \quad (1)$$

患病比例 y 是年龄 x 的函数,其取值在区间 $[0, 1]$ 上.如果用普通的方法建立回归方程,那么很容易求得其线性回归曲线或更接近于S型曲线的3次多项式回归曲线(分别见图3和图4),其回归模型的形式为

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \varepsilon \quad (2)$$

其中随机误差 ε 服从均值为0的正态分布,特别地,当 $\beta_2 = \beta_3 = 0$ 时为线性回归模型.

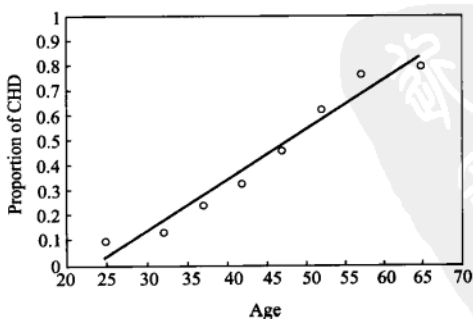


图3 y 对 x 的线性回归曲线

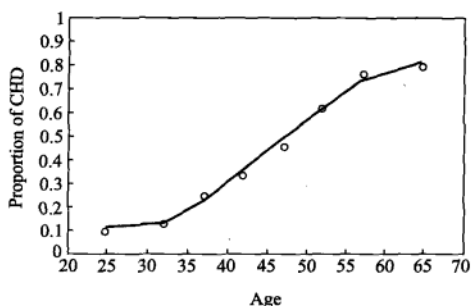


图4 y 对 x 的三次多项式回归曲线

然而在这个问题中, (2) 式回归方程中 y 的取值不一定在 $[0, 1]$ 中. 进一步说, 即使 y 的值在 $[0, 1]$ 中, 由于在给定 x 时, 误差项 ε 也只能取 0, 1 两个值, 显然 ε 不具有正态性, 而且 ε 的方差依赖于 x , 具有异方差性, 这些都违反了普通回归分析的前提条件. 因此, 当 Y 为一个二分类 (或多分类) 变量而不是连续变量时, 用前几节介绍的基于最小二乘法的回归分析是不适合的, 需要用到新的回归模型.

logit 模型 下面用 $\pi(x)$ 表示年龄为 x 的被观察者患冠心病的概率, 即

$$\pi(x) = P(Y = 1 | x) \quad (3)$$

显然 Y 的 (条件) 期望为 $E(Y | x) = \pi(x)$, (条件) 方差为 $D(Y | x) = \pi(x)(1 - \pi(x))$. 由 (1) 式可知, $\pi(x)$ 即为该年龄段的患病比例 y .

为了寻求患病概率 $\pi(x)$ 与年龄 x 之间、形如图 2 的 S 型曲线的函数关系, 并注意 $\pi(x)$ 在 $[0, 1]$ 区间取值, 可以建立如在第 5 章、第 7 章多次用到的 logistic 模型

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad (4)$$

(4) 的反函数写作

$$\ln \frac{\pi(x)}{1 - \pi(x)} = \beta_0 + \beta_1 x \quad (5)$$

(5) 式左端可看作 $\pi(x)$ 的变换, 记作 $\text{logit}(\pi(x)) = \ln \frac{\pi(x)}{1 - \pi(x)}$, 称为 logit 模型或 logistic 回归模型. 当 $\pi(x)$ 在 $[0, 1]$ 取值时, $\text{logit}(\pi(x))$ 取值为 $(-\infty, +\infty)$.

在数据预处理时, 将被观察者的年龄分成 $k = 8$ 组, 记第 i 组 ($i = 1, 2, \dots, k$) 年龄为 x_i , 被观察人数为 n_i , 患病人数为 m_i , 每位被观察者患病概率为

$\pi_i = m_i/n_i$, 这时 logit 模型具有如下形式^①:

$$\text{logit}(\pi_i) = \ln \frac{\pi_i}{1 - \pi_i} = \beta_0 + \beta_1 x_i \quad (6)$$

其中 β_0, β_1 是回归系数. 合理地设 m_i 服从二项分布 $B(n_i, \pi_i)$, β_0, β_1 可用最大似然法估计得到^[36].

模型求解 logit 模型是一种广义线性模型 (Generalized Linear Model), 可利用 MATLAB 统计工具箱中的命令 `glmfit` 求解, 通常的使用格式为:

```
b = glmfit(x,y,'distr','link')
```

或 `[b,dev,stats] = glmfit(x,y,'distr','link')`

其中输入 x 为自变量数据矩阵, 缺省时会自动添加一列 1 向量作为 x 的第 1 列; y 为因变量数据向量; `distr` 为估计系数时所用的分布, 可以是 `binomial`, `poisson` 等, 缺省时为 `normal`; 特别当 `distr` 取 `binomial` 时, y 可取一个 2 列矩阵, 第 1 列为观察“成功”的次数, 第 2 列为观察次数; `link` 取 `logit`, `probit` (Probit 模型见下面) 等, 缺省时为 `logit`. 输出 b 为回归系数的估计值; `dev` 为拟合偏差, 是一般的残差平方和的推广; `stats` 输出一些统计指标, 详见 MATLAB 的帮助文件.

用表 2 的数据输入并执行以下的程序:

```
Age = [24.5 32 37 42 47 52 57 64.5]';  
Chd = [1 2 3 5 6 5 13 8]';  
Total = [10 15 12 15 13 8 17 10]';  
Proport = Chd./Total;  
[b,dev,stats] = glmfit(Age,[Chd Total],'binomial','logit');  
logitFit = glmval(b,Age,'logit');  
plot(Age,Proport,'o',Age,logitFit,'r-');  
xlabel('Age');ylabel('Proportion of CHD')  
b,bi = stats.se,dev
```

得到 logit 模型中的参数 β_0, β_1 的最大似然估计值与它的标准差 (见表 3), 拟合偏差为 0.524 2. 图 5 给出了 logistic 回归曲线与散点图.

利用命令

```
[yhat,dylo,dyhi] = glmval(b,x,'logit')
```

还可以给出自变量为 x 时因变量 y 的预测值及置信度为 95% 的置信区间, 结果如表 4.

① 当 $m_i = 0$ 或 $m_i = n_i$ 时, 可用修正公式 $\pi_i = \frac{m_i + 0.5}{n_i + 1}$ 来计算.

表 3 模型参数的估计结果

参数	参数估计值	标准差
β_0	-5.038 2	1.086 3
β_1	0.105 0	0.023 1

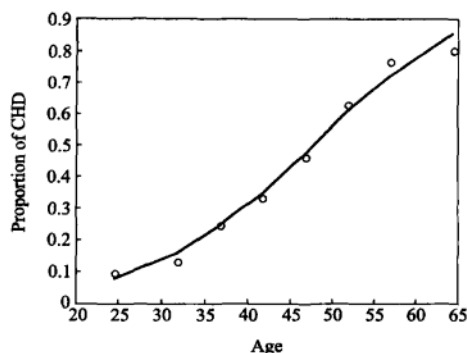


图 5 logistic 回归曲线与散点图

表 4 患病比例的预测值与预测区间

年龄段	年龄 x	患病比例	预测值 $\hat{\pi}(x)$	置信区间
20 ~ 29	24.5	0.1	0.078 3	[0.028 2, 0.199 2]
30 ~ 34	32	0.13	0.157 4	[0.078 8, 0.289 8]
35 ~ 39	37	0.25	0.240 0	[0.146 1, 0.368 3]
40 ~ 44	42	0.33	0.348 1	[0.245 8, 0.466 6]
45 ~ 49	47	0.46	0.474 4	[0.362 5, 0.589 1]
50 ~ 54	52	0.63	0.604 1	[0.472 0, 0.722 7]
55 ~ 59	57	0.76	0.720 7	[0.566 8, 0.835 7]
60 ~ 69	64.5	0.80	0.850 1	[0.685 5, 0.936 6]

模型评价与结果分析 首先我们看到, logit 模型(5)的右端是年龄 x 的线性函数, 如果加入 x 的二次项后, 是否能显著地提高模型的拟合程度呢, 即考虑模型

$$\text{logit}(\pi(x)) = \ln \frac{\pi(x)}{1 - \pi(x)} = \beta_0 + \beta_1 x + \beta_2 x^2 \quad (7)$$

为了检验是否需要引入 x^2 项, 在 logistic 回归中要用似然比检验统计量^[36]. 执行程序

```
[b2, dev2] = glmfit([Age Age.^2], [Chd Total], 'binomial', 'logit');
```

```
b2, pval = 1 - chi2cdf(dev - dev2, 1)
```

计算得到 $\beta_0, \beta_1, \beta_2$ 的估计值分别为 -5.350 6, 0.119 4 和 -0.000 2, p 值为

0.937 1,表示模型中引入 x^2 项并不能显著提高拟合程度。

其次,处理这类问题的另一种广义线性模型是 probit 模型,其形式为

$$\pi(x) = \Phi(\beta_0 + \beta_1 x) \quad (8)$$

$$\text{probit}(\pi(x)) = \Phi^{-1}(\pi(x)) = \beta_0 + \beta_1 x \quad (9)$$

其中 Φ 是正态概率分布函数,它也是 S 型曲线。利用 MATLAB 统计工具箱中的命令 glmfit 求解时,只需将 logit 改为 probit 即可。

用表 2 数据计算得到 probit 模型系数 β_0, β_1 的估计值分别为 -2.993 3 和 0.062 4,标准差分别为 0.601 1 和 0.012 8,拟合偏差为 0.652 9。与 logit 模型的结果比较见表 5,拟合曲线比较见图 6,可以发现这两个模型的拟合程度不相上下。

表 5 probit 模型与 logit 模型的比较

年龄段	年龄 x	患病比例	预测值 (logit)	预测值 (probit)	置信区间 (logit)	置信区间 (probit)
20 ~ 29	24.5	0.1	0.078 3	0.071 5	[0.028 2, 0.199 2]	[0.019 7, 0.192 4]
30 ~ 34	32	0.13	0.157 4	0.159 5	[0.078 8, 0.289 8]	[0.076 0, 0.287 6]
35 ~ 39	37	0.25	0.240 0	0.246 9	[0.146 1, 0.368 3]	[0.151 3, 0.367 7]
40 ~ 44	42	0.33	0.348 1	0.354 8	[0.245 8, 0.466 6]	[0.255 2, 0.465 6]
45 ~ 49	47	0.46	0.474 4	0.475 9	[0.362 5, 0.589 1]	[0.369 0, 0.584 6]
50 ~ 54	52	0.63	0.604 1	0.599 4	[0.472 0, 0.722 7]	[0.474 2, 0.715 0]
55 ~ 59	57	0.76	0.720 7	0.713 5	[0.566 8, 0.835 7]	[0.566 6, 0.831 4]
60 ~ 69	64.5	0.80	0.850 1	0.848 9	[0.685 5, 0.936 6]	[0.685 4, 0.943 0]

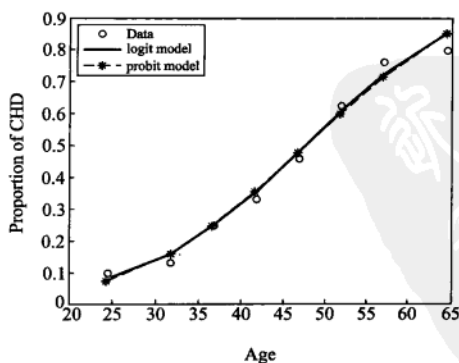


图 6 probit 模型与 logit 模型的拟合比较
(实线为 logit 模型,带“*”虚线为 probit 模型)

模型预测与进一步分析 通过上述分析可知,对于我们的问题和观察数据, logit 模型

$$\text{logit}(\hat{\pi}(x)) = \ln \frac{\hat{\pi}(x)}{1 - \hat{\pi}(x)} = -5.0382 + 0.1050x \quad (10)$$

是一个合适的模型,从(10)式能够给出任何年龄的人患冠心病的概率及相应的置信区间.例如,图7给出了年龄分别为20,30,40,50,60,70,80的人患冠心病的概率,以及置信度为95%的置信区间.

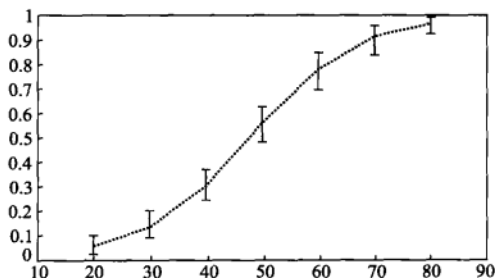


图7 不同年龄的人患冠心病的概率的预测(竖线为置信区间)

logit 模型的另一个好处是其中的回归系数 β_1 有很好的直观解释. logit 模型与统计中 odds(发生比或优势)的概念有密切的联系,所谓 odds 就是事件的发生概率与不发生概率之比.本节中,若记 $\text{odds}(x)$ 为年龄 x 的人患与不患冠心病的概率之比,则

$$\text{odds}(x) = \frac{\pi(x)}{1 - \pi(x)} \quad (11)$$

于是 logit 模型可以表示为

$$\text{odds}(x) = e^{\beta_0 + \beta_1 x} \quad (12)$$

当年龄增加1岁时 odds 比(发生比率或优势比)为

$$\frac{\text{odds}(x+1)}{\text{odds}(x)} = \frac{e^{\beta_0 + \beta_1(x+1)}}{e^{\beta_0 + \beta_1 x}} = e^{\beta_1} \quad (13)$$

于是

$$\beta_1 = \ln \frac{\text{odds}(x+1)}{\text{odds}(x)} \quad (14)$$

即 β_1 为自变量增加1个单位时 odds 比的对数. $\beta_1 > 0$ 时, $e^{\beta_1} > 1$, x 每增加1个单位, odds 比会相应增加,且对任意正整数 k ,有

$$\text{odds}(x+k) = e^{k\beta_1} \text{odds}(x) \quad (15)$$

在模型(10)中 $\beta_1 = 0.1050$, 可以算出一个20岁的青年人患冠心病的概率

仅为 $\hat{\pi}(20) = 0.0503$, 且发生比 (患与不患冠心病的概率之比) 为 $\text{odds}(20) = 0.0593$, 说明这个年龄的人患冠心病几乎是不太可能的, 年龄增加 1 岁患病概率的变化很小. 10 年后, 30 岁人的发生比就变成 $\text{odds}(30) = e^{10 \times \beta_1} \times 0.0593 = 0.1695$, 发生比 (可解释作危险率) 增大到 20 岁时的 2.8577 倍, 而到 60 岁时, $\text{odds}(60) = e^{40 \times \beta_1} \times 0.0593 = 3.9545$, 危险率是 20 岁的 $e^{40 \times \beta_1} = 66.6863$ 倍. 可见回归系数 β_1 在 logit 模型中有着重要的意义. 这一点在 probit 模型中是无法体现的.

最后, 在 logit 模型中, 人们常常感兴趣的是, x 取何值时 $\pi(x) = 0.5$. 由模型 (10) 求解 $\hat{\pi}(x^*) = 0.5$, 可得 $x^* = 47.98$. 这就是说, 当你到 48 岁时, 患冠心病的概率就会大于不患冠心病的概率, 要格外小心了!

评注 因变量是定性变量的回归分析作为一种有效的数据处理方法已被广泛应用, 尤其在医学、社会调查、生物信息处理等领域. 这类回归模型属于广义线性模型的研究范畴. 本节只涉及因变量是 0-1 变量且只有一个自变量的情形, 对多个自变量 x_1, \dots, x_m 的情形, logit 模型和 probit 模型的一般形式分别为

$$\text{logit}(\pi(x)) = \ln \frac{\pi(x)}{1 - \pi(x)} = \beta_0 + \sum_{i=1}^m \beta_i x_i \quad (16)$$

$$\text{probit}(\pi(x)) = \Phi^{-1}(\pi(x)) = \beta_0 + \sum_{i=1}^m \beta_i x_i \quad (17)$$

其中 x_1, \dots, x_m 可以是数值变量, 也可以是分类变量 (如 $x_1 = 1$ 表示男性, $x_1 = 0$ 表示女性等), 其分析与处理方式类似于一元的情形.

在建立多元 logit 模型和 probit 模型时, 可以借鉴 10.5 节介绍的逐步回归的思想, 在初始模型中一个个地加入自变量, 包括某个自变量的二次或高次项 (如 (7) 式), 也包括某些自变量的交叉变量, 并且实时地进行模型比较检验, 以便选择与数据拟合较好的模型.

另外, 当因变量是 (特别是有序) 多分类指标变量时, 如观察结果为“无、轻、中、重”不同等级的数据, 可以采用 (有序) 多分类 logit 模型^[36].

习 题

1. 下表列出了某城市 18 位 35~44 岁经理的年平均收入 x_1 千元, 风险偏好度 x_2 和人寿保险额 y 千元的数据, 其中风险偏好度是根据发给每个经理的问卷调查表综合评估得到的, 它的数值越大就越偏爱高风险. 研究人员想研究此年龄段中的经理所投保的人寿保险额与年均收入及风险偏好度之间的关系. 研究者预计, 经理的年均收入和人寿保险额之间存在着二次关系, 并有把握地认为风险偏好度对人寿保险额有线性效应, 但对风险偏好度对人寿保险额是否有二次效应以及两个自变量是否对人寿保险额有交互效应, 心中没底.

请你通过表中的数据来建立一个合适的回归模型, 验证上面的看法, 并给出进一步的分析^[79].

序号	y	x_1	x_2	序号	y	x_1	x_2
1	196	66.290	7	10	49	37.408	5
2	63	40.964	5	11	105	54.376	2
3	252	72.996	10	12	98	46.186	7
4	84	45.010	6	13	77	46.130	4
5	126	57.204	4	14	14	30.366	3
6	14	26.852	5	15	56	39.060	5
7	49	38.122	4	16	245	79.380	1
8	49	35.840	6	17	133	52.766	8
9	266	75.796	9	18	133	55.916	6

2. 一个医药公司的新药研究部门为了掌握一种新止痛剂的疗效,设计了一个药物实验,给患有同种病痛的病人使用这种新止痛剂的以下4个剂量中的某一个:2 g, 5 g, 7 g 和 10 g,并记录每个病人病痛明显减轻的时间(以分钟计).为了解新药的疗效与病人性别和血压有什么关系,试验过程中研究人员把病人按性别及血压的低、中、高三档平均分配来进行测试.通过比较每个病人血压的历史数据,从低到高分成3组,分别记作0.25, 0.50 和 0.75. 实验结束后,公司的记录结果见下表(性别以0表示女,1表示男).

请你为公司建立一个模型,根据病人用药的剂量、性别和血压组别,预测出服药后病痛明显减轻的时间.

病人序号	病痛减轻时间/min	用药剂量/g	性别	血压组别
1	35	2	0	0.25
2	43	2	0	0.50
3	55	2	0	0.75
4	47	2	1	0.25
5	43	2	1	0.50
6	57	2	1	0.75
7	26	5	0	0.25
8	27	5	0	0.50
9	28	5	0	0.75
10	29	5	1	0.25

病人序号	病痛减轻时间/min	用药剂量/g	性别	血压组别
11	22	5	1	0.50
12	29	5	1	0.75
13	19	7	0	0.25
14	11	7	0	0.50
15	14	7	0	0.75
16	23	7	1	0.25
17	20	7	1	0.50
18	22	7	1	0.75
19	13	10	0	0.25
20	8	10	0	0.50
21	3	10	0	0.75
22	27	10	1	0.25
23	26	10	1	0.50
24	5	10	1	0.75

3. 某地人事部门为研究中学教师的薪金与他们的资历、性别、教育程度及培训情况等因素之间的关系,要建立一个数学模型,分析人事策略的合理性,特别是考察女教师是否受到不公正的待遇,以及她们的婚姻状况是否会影响收入.为此,从当地教师中随机选了3414位进行观察,然后从中保留了90个观察对象,得到了下表给出的相关数据.尽管这些数据具有一定的代表性,但是仍有统计分析的必要.现将表中数据的符号介绍如下:

Z ——月薪(单位:元); X_1 ——工作时间(以月计); $X_2 = 1$ ——男性, $X_2 = 0$ ——女性; $X_3 = 1$ ——男性或单身女性, $X_3 = 0$ ——已婚女性; X_4 ——学历(取值0~6,值越大表示学历越高); $X_5 = 1$ ——受雇于重点中学, $X_5 = 0$ ——其他; $X_6 = 1$ ——受过培训的毕业生, $X_6 = 0$ ——未受过培训的毕业生或受过培训的肄业生; $X_7 = 1$ ——已两年以上未从事教学工作, $X_7 = 0$ ——其他.注意组合 $(X_2, X_3) = (1, 1), (0, 1), (0, 0)$ 的含义.

(1) 进行变量选择,建立变量 $X_1 \sim X_7$ 与 Z 的回归模型(不一定包括每个自变量),说明教师的薪金与那些变量的关系密切,是否存在性别和婚姻状况上的差异.为了数据处理上的方便,建议对薪金取对数后作为因变量.

(2) 除了变量 $X_1 \sim X_7$ 本身之外,尝试将它们平方项或交互项加入到模型中,建立更好的模型^[23].

	Z	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇		Z	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇
1	998	7	0	0	0	0	0	0	30	1 424	78	1	1	0	1	0	0
2	1 015	14	1	1	0	0	0	0	31	1 424	79	0	1	0	0	0	0
3	1 028	18	1	1	0	1	0	0	32	1 347	91	1	1	0	1	0	0
4	1 250	19	1	1	0	0	0	0	33	1 343	92	0	0	0	0	0	1
5	1 028	19	0	1	0	1	0	0	34	1 310	94	0	0	0	1	0	0
6	1 028	19	0	0	0	0	0	0	35	1 814	103	0	0	2	1	1	0
7	1 018	27	0	0	0	0	0	1	36	1 534	103	0	0	0	0	0	0
8	1 072	30	0	0	0	0	0	0	37	1 430	103	1	1	0	0	0	0
9	1 290	30	1	1	0	0	0	0	38	1 439	111	1	1	0	1	0	0
10	1 204	30	0	1	0	0	0	0	39	1 946	114	1	1	3	1	1	0
11	1 352	31	0	1	2	0	1	0	40	2 216	114	1	1	4	1	1	0
12	1 204	31	0	0	0	1	0	0	41	1 834	114	1	1	4	1	1	1
13	1 104	38	0	0	0	0	0	0	42	1 416	117	0	0	0	0	0	1
14	1 118	41	1	1	0	0	0	0	43	2 052	139	1	1	0	1	0	0
15	1 127	42	0	0	0	0	0	0	44	2 087	140	0	0	2	1	1	1
16	1 259	42	1	1	0	1	0	0	45	2 264	154	0	0	2	1	1	1
17	1 127	42	1	1	0	0	0	0	46	2 201	158	1	1	4	0	1	1
18	1 127	42	0	0	0	1	0	0	47	2 992	159	1	1	5	1	1	1
19	1 095	47	0	0	0	0	0	1	48	1 695	162	0	1	0	0	0	0
20	1 113	52	0	0	0	0	0	1	49	1 792	167	1	1	0	1	0	0
21	1 462	52	0	1	2	0	1	0	50	1 690	173	0	0	0	0	0	1
22	1 182	54	1	1	0	0	0	0	51	1 827	174	0	0	0	0	0	1
23	1 404	54	0	0	0	1	0	0	52	2 604	175	1	1	2	1	1	0
24	1 182	54	0	0	0	0	0	0	53	1 720	199	0	1	0	0	0	0
25	1 594	55	1	1	2	1	1	0	54	1 720	209	0	0	0	0	0	0
26	1 459	66	0	0	0	1	0	0	55	2 159	209	0	1	4	1	0	0
27	1 237	67	1	1	0	1	0	0	56	1 852	210	0	1	0	0	0	0
28	1 237	67	0	1	0	1	0	0	57	2 104	213	1	1	0	1	0	0
29	1 496	75	0	1	0	0	0	0	58	1 852	220	0	0	0	0	0	1

续表

	Z	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇		Z	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇
59	1 852	222	0	0	0	0	0	0	75	1 942	319	0	0	0	1	0	0
60	2 210	222	1	1	0	0	0	0	76	2 027	325	1	1	0	0	0	0
61	2 266	223	0	1	0	0	0	0	77	1 942	326	1	1	0	1	0	0
62	2 027	223	1	1	0	0	0	0	78	1 720	329	1	1	0	1	0	0
63	1 852	227	0	0	0	1	0	0	79	2 048	337	0	0	0	0	0	0
64	1 852	232	0	0	0	0	0	1	80	2 334	346	1	1	2	1	1	1
65	1 995	235	0	0	0	0	0	1	81	1 720	355	0	0	0	0	0	1
66	2 616	245	1	1	3	1	1	0	82	1 942	357	1	1	0	0	0	0
67	2 324	253	1	1	0	1	0	0	83	2 117	380	1	1	0	0	0	1
68	1 852	257	0	1	0	0	0	1	84	2 742	387	1	1	2	1	1	1
69	2 054	260	0	0	0	0	0	0	85	2 740	403	1	1	2	1	1	1
70	2 617	284	1	1	3	1	1	0	86	1 942	406	1	1	0	1	0	0
71	1 948	287	1	1	0	0	0	0	87	2 266	437	0	1	0	0	0	0
72	1 720	290	0	1	0	0	0	1	88	2 436	453	0	1	0	0	0	0
73	2 604	308	1	1	2	1	1	0	89	2 067	458	0	1	0	0	0	0
74	1 852	309	1	1	0	1	0	1	90	2 000	464	1	1	2	1	1	0

4. 在 10.3 节酶促反应中,如果用指数增长模型 $y = \beta_1 (1 - e^{-\beta_2 x})$ 代替 Michaelis-Menten 模型对经过嘌呤霉素处理的实验数据作非线性回归分析,其结果将如何?更进一步,若选用模型 $y = \beta_1 (e^{-\beta_3 x} - e^{-\beta_2 x})$ 来拟合相同的数据,其结果是否比指数增长模型有所改进?试作出模型的残差图进行比较。

5. logistic 增长曲线模型和 Gompertz 增长曲线模型是计量经济学等学科中的两个常用模型,可以用来拟合销售量的增长趋势。

记 logistic 增长曲线模型为 $y_t = \frac{L}{1 + ae^{-kt}}$, 记 Gompertz 增长曲线模型为 $y_t = Le^{-be^{-kt}}$, 这两个模型中 L 的经济学意义都是销售量的上限。下表中给出的是某地区高压锅的销售量(单位:万台),为给出此两模型的拟合结果,请考虑如下的问题:

(1) logistic 增长曲线模型是一个可线性化模型吗?如果给定 $L = 3\,000$, 是否是一个可线性化模型,如果是,试用线性化模型给出参数 a 和 k 的估计值。

(2) 利用(1)所得到的 a 和 k 的估计值和 $L = 3\,000$ 作为 logistic 模型的拟合初值,对 logistic 模型做非线性回归。

(3) 取初值 $L^{(0)} = 3\ 000$, $b^{(0)} = 30$, $k^{(0)} = 0.4$, 拟合 Gompertz 模型. 并与 logistic 模型的结果进行比较.

年份	t	y	年份	t	y
1981	0	43.65	1988	7	1 238.75
1982	1	109.86	1989	8	1 560.00
1983	2	187.21	1990	9	1 824.29
1984	3	312.67	1991	10	2 199.00
1985	4	496.58	1992	11	2 438.89
1986	5	707.65	1993	12	2 737.71
1987	6	960.25			

6. 某公司想用全行业的销售额作为自变量来预测公司的销售额, 下表给出了 1977—1981 年公司销售额和行业销售额的分季度数据(单位: 百万元).

(1) 画出数据的散点图, 观察用线性回归模型拟合是否合适.

(2) 建立公司销售额对全行业销售额的回归模型, 并用 D-W 检验诊断随机误差项的自相关性.

(3) 建立消除了随机误差项自相关性后的回归模型^[59].

年	季	t	公司销 售额 y	行业销 售额 x	年	季	t	公司销 售额 y	行业销 售额 x
1977	1	1	20.96	127.3	1979	3	11	24.54	148.3
	2	2	21.40	130.0		4	12	24.30	146.4
	3	3	21.96	132.7	1980	1	13	25.00	150.2
	4	4	21.52	129.4		2	14	25.64	153.1
1978	1	5	22.39	135.0	1981	3	15	26.36	157.3
	2	6	22.76	137.1		4	16	26.98	160.7
	3	7	23.48	141.2	1981	1	17	27.52	164.2
	4	8	23.66	142.8		2	18	27.78	165.6
1979	1	9	24.10	145.5	1981	3	19	28.24	168.7
	2	10	24.01	145.3		4	20	28.78	171.7

7. 下表给出了某工厂产品的生产批量与单位成本(单位: 元)的数据, 从散点图可以明显地发现, 生产批量在 500 以内时, 单位成本对生产批量服从一种线性关系, 生产批量超过 500 时服从另一种线性关系, 此时单位成本明显下降. 希望你构造一个合适的回归模型全面地描述生产批量与单位成本的关系.

生产批量	650	340	400	800	300	600	720	480	440	540	750
单位成本	2.48	4.45	4.52	1.38	4.65	2.96	2.18	4.04	4.20	3.10	1.50

8. 将 10.6 节表 1 的冠心病数据按每 5 岁为一年龄段进行分组,重新建立患病比例与年龄的 logit 模型和 probit 模型,并进行分析和比较.

9. 在一项调查降价折扣券对顾客的消费行为影响的研究中,商家对 1 000 个顾客发放了商品折扣券和宣传资料,折扣券的折扣比例分别为 5%,10%,15%,20%,30%,每种比例的折扣券均发放了 200 人,现记录他们在一个月内使用折扣券购物的人数和比例数据如下表:

折扣比例/%	持折扣券人数	使用折扣券人数	使用折扣券人数比例
5	200	32	0.160
10	200	51	0.255
15	200	70	0.350
20	200	103	0.515
30	200	148	0.740

(1) 对使用折扣券人数比例先作 logit 变换,再对使用折扣券人数比例与折扣比例,建立普通的一元线性回归模型.

(2) 直接利用 MATLAB 统计工具箱中的 glmfit 命令,建立使用折扣券人数比例与折扣比例的 logit 模型.与(1)做比较,并估计若想要使用折扣券人数比例为 25%,则折扣券的折扣比例应该为多大?

10. 人类的性别是由基因决定的,乌龟的性别主要是由什么因素决定的呢?科学研究表明,决定幼龟性别的最关键的因素是乌龟蛋孵化时的温度.为了研究温度是如何影响幼龟的雌雄比例,美国科学家对某一类乌龟的孵化过程作了试验.试验在 5 个不同的恒定温度下进行,每个温度下分别观察 3 批乌龟蛋的孵化过程,得到的数据如下:

温度/℃	乌龟蛋个数	雄龟个数	雌龟个数	雄龟比例
27.2	10	1	9	10%
	8	0	8	0%
	9	1	8	11.1%
27.7	10	7	3	70%
	6	4	2	66.7%
	8	6	2	75%

续表

温度/℃	乌龟蛋个数	雄龟个数	雌龟个数	雄龟比例
28.3	13	13	0	100%
	9	6	3	66.7%
	8	7	1	87.5%
28.4	10	7	3	70%
	8	5	3	62.5%
	9	7	2	77.8%
29.9	11	10	1	90.9%
	8	8	0	100%
	9	9	0	100%

建立幼龟性别比和孵化温度之间的 logit 模型,并求出在孵化温度多大时,孵化出的幼龟的性别比例恰好为 1:1. 分析温度每升高 1℃,幼龟性别的变化情况.



第 11 章

博弈模型

在第 4 章中,我们介绍了解决决策问题的优化模型(数学规划模型),这类模型的三个基本要素是:决策变量(寻求的决策是什么)、目标函数(需要优化的目标是什么)、约束条件(决策面临哪些限制条件)。对单一决策者(个体或组织)如何作出最有利于自己的决策的问题,优化建模是一类重要和有效的方法。但当存在多个决策者,每个决策者有自己的决策变量和目标函数,并且一个决策者的决策变量以某种形式出现在另一个决策者的目标函数中时,决策者之间的决策行为相互影响,就不能用一般的优化模型进行建模和求解了。这种决策主体的决策行为发生直接相互作用的多人决策问题一般称为博弈或对策(Game),研究博弈模型及其均衡问题的理论称为博弈论或对策论(Game Theory)。博弈论对这类问题进行建模和求解提供了有效方法,其基本假定是所有决策主体是完全理性的,每个决策主体都希望最优化自己的个人目标,与整个系统的整体目标不一定一致。

博弈可以分为合作博弈和非合作博弈,其主要区别在于决策者的决策行为相互作用时,当事人能否达成一种有约束力的协议。如果有,就是合作博弈;如果没有,就是非合作博弈。对于合作博弈,决策者面临的主要是如何分享合作带来的成果。对于非合作博弈,每个决策者都面临如何选择自己的行动,即决策变量应该取什么值,更一般地说,每个决策者要制定自己的行动战略,即选择自己的行动规则,这一规则决定在什么情况下自己应该采取什么行动。根据所有决策者的决策是同时作出的、还是按一定先后顺序作出的,非合作博弈可以分为静态博弈和动态博弈;根据决策者在决策时所掌握的信息的多少,非合作博弈可以分为完全信息博弈和不完全信息博弈。本章前 4 节属于非合作博弈,后 2 节属于合作博弈。

博弈论的内容非常丰富,目前已经成为微观经济学的基本分析工具,在军事、政治、企业管理和社会科学中的应用也日益广泛。与本书中其他章节的思路类似,我们在本章不准备介绍博弈论的完整理论,而是通过一些具体例子介绍几个基本的博弈模型。

11.1 进攻与撤退的抉择

在一场局部战役中,指挥员能否审时度势地作出正确的决策,往往决定着这场战役的成败.下面以第二次世界大战期间的一场战役为例,分析德军在面临进攻还是撤退的抉择时,最后的选择是否正确.

问题背景 第二次世界大战进行到1944年6月初,盟军在诺曼底登陆成功.到8月初,盟军(美国第三集团军)突破滩头阵地并迅速向东南方向推进,但与位于其西北方向的盟军(美国第一集团军)之间出现了缺口,而德国军队的北面 and 西北方向已被盟军(加拿大第一集团军和英国第二集团军)占领(见图1,其中虚线表示当时盟军和德军控制区的分界线).此时德军指挥官面临两个选择:一是向西面的盟军缺口发动进攻,二是向东撤退避免被包围;而盟军指挥官对于如何使用其预备队面临三个选择:一是立即将预备队开赴缺口处强化其防御,二是原地待命,三是立即东进与美第三集团军会合,希望能对德军形成包围并消灭之.在这种情况下,双方应该如何决策^[31]?

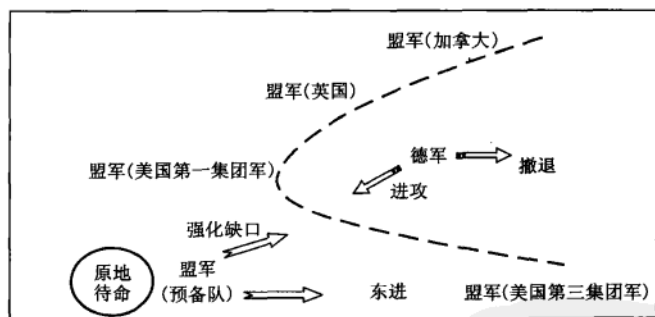


图1 双方兵力部署及可能的决策示意图

问题分析及模型假设 根据双方决策的不同,有两场可能发生的战斗:德军向西进攻盟军缺口;盟军围攻德军.按照当时双方的战斗力,估计不同情况下战争的结果大致是:如果德军进攻盟军预备队强化后的缺口,德军必败;如果德军进攻盟军缺口而盟军预备队向东包围德军,则德军进攻成功从而避免了被包围,因此盟军两场战斗都会失败;如果德军进攻盟军缺口而盟军预备队原地待命,则盟军预备队可以灵活使用,盟军两场战斗都会胜利;如果德军向东撤退而盟军预备队向东包围德军,则盟军将赢得东面的包围战;如果德军向东撤退而盟军预备队不向东包围德军,则不会发生任何战斗.

将这个实例合理地描述为如下的博弈问题:

1. 博弈参与者为两方(盟军和德军)。

2. 博弈中盟军有 3 种使用其预备队的决策行动:强化缺口、原地待命、东进;德军有 2 种决策行动:向西进攻或向东撤退。

3. 博弈双方的目的都是要获得最后的胜利,不妨简单地假定目的都是使战斗中己方获得的净胜场次(指胜利场次减去失败场次)尽可能多。

博弈双方的决策行动及其产生的结果可以清晰地用表 1 描述(盟军胜的场次即德军败的场次)。

表 1 盟军和德军的决策行动及其产生的结果

盟军 \ 德军	向西进攻	向东撤退
强化缺口	盟军胜 1 场	无战斗
原地待命	盟军胜 2 场	无战斗
东进	盟军败 2 场	盟军胜 1 场

假设双方在决策时都知道以上信息,并且必须同时作出抉择,这种博弈称为完全信息静态博弈。博弈中双方都希望通过决策行动获得最大的胜利。

模型建立 参与博弈的双方用 $N = \{1, 2\}$ 表示, 1 为盟军, 2 为德军。盟军可能的决策行动记作 $a_1 \in A_1 = \{1, 2, 3\}$, 分别表示强化缺口、原地待命和东进; 德军可能的决策行动记作 $a_2 \in A_2 = \{1, 2\}$, 分别表示进攻和撤退。

对于双方每一种可能的决策 (a_1, a_2) , 用 $u_1(a_1, a_2)$ 表示对盟军产生的结果, 即净胜场次, 称为盟军的效用函数, 由表 1, $u_1(a_1, a_2)$ 可以用矩阵

$$M = \{m_{ij}\}_{3 \times 2} = \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ -2 & 1 \end{pmatrix} \quad (1)$$

表示, 即 $u_1(i, j) = m_{ij}$, M 称为盟军的赢得矩阵或支付矩阵(Payoff Matrix)。类似地, 德军的效用函数用 $u_2(a_1, a_2)$ 表示, 显然其赢得矩阵是 $-M$ 。这种具有完全竞争性质, 即一方所得正是对手所失的博弈, 一般称为零和博弈。

在博弈中双方都力求通过决策行动使己方的效用函数最大化, 用 a_1^*, a_2^* 分别表示盟军和德军会选择的决策行动, 则 a_1^*, a_2^* 必须满足

$$\begin{aligned} u_1(a_1^*, a_2^*) &\geq u_1(a_1, a_2^*), \\ u_2(a_1^*, a_2^*) &\geq u_2(a_1^*, a_2), a_1 \in \{1, 2, 3\}, a_2 \in \{1, 2\} \end{aligned} \quad (2)$$

在博弈论中这样的策略组合 $a^* = (a_1^*, a_2^*)$ 称为(纯)纳什均衡(Nash Equilibrium), 其含义是, 每一方的决策选择对于他方的决策选择而言都是最优的, 称为最优反应, 因而每一方都不会偏离该决策。

对于由(1)给出的效用函数,容易验证(如通过枚举)不存在纯纳什均衡.

如果不是用净胜场次而是采取某种意义下的综合得分来定义盟军和德军的效用函数,如表2所示(其中两个数字分别是盟军和德军的综合得分,对于表1中的“无战斗”情况,双方的得分完全是虚拟的),那么就可以验证,对于策略组合 $a^* = (a_1^*, a_2^*) = (2, 2)$, 双方的效用函数 $u_1(a_1^*, a_2^*)$ 和 $u_2(a_1^*, a_2^*)$ 满足(2)式,即 a^* 为纯纳什均衡.

表2 盟军和德军的决策行动的综合得分

盟军 \ 德军	德军	
	向西进攻	向东撤退
强化缺口	1, -1	1, 1
原地待命	2, -2	2, 1
东进	-2, 2	1, -1

对于不存在纯纳什均衡的博弈问题,可以考虑双方随机地采取行动,即双方都对每一种决策行动赋予一定的概率,形成混合策略. 设盟军采取行动 i 的概率为 $p_i (i=1, 2, 3)$, 德军采取行动 j 的概率为 $q_j (j=1, 2)$, 记盟军和德军的(混合)策略集分别为

$$S_1 = \{p = (p_1, p_2, p_3) \mid 0 \leq p_i \leq 1, \sum_{i=1}^3 p_i = 1\},$$

$$S_2 = \{q = (q_1, q_2) \mid 0 \leq q_i \leq 1, \sum_{i=1}^2 q_i = 1\} \quad (3)$$

显然, A_1 中的每一个行动 i 也是盟军的一个混合策略(采取行动 i 的概率为1), 称为纯策略(相对于混合策略而言). A_2 也类似.

在混合策略下双方的效用函数用期望效用定义, 记作

$$U_1(p, q) = pMq^T = \sum_{i=1}^3 \sum_{j=1}^2 p_i m_{ij} q_j, U_2(p, q) = -U_1(p, q) \quad (4)$$

类似地可以定义(混合策略)纳什均衡, 只要将(2)式中的效用改成期望效用即可.

盟军希望最大化它的期望效用, 所面临的决策问题是

$$\max_{p \in S_1} U_1(p, q) = pMq^T \quad (5)$$

而德军面临的决策等价于

$$\min_{q \in S_2} U_1(p, q) = pMq^T \quad (6)$$

与一般的优化模型不同, 这里双方各有一个优化问题, 而且决策存在相互影响, 双方各自控制部分决策变量, 难以用第3章和第4章介绍的方法直接求解.

注意到双方的纯策略是有限的,其纯策略下的效用可以用矩阵描述,且双方效用之和为0,这样的博弈称为二人零和矩阵博弈.对这种特殊类型的博弈,可以采用线性规划求解.

模型求解 盟军怎样在不能控制 q 的情况下使 pMq^T 最大? 德军怎样在不能控制 p 的情况下使 pMq^T 最小? 因为双方都希望优化自己的赢得,他们都会得到如下结论:不管己方怎么做,他方总是采取策略使己方的赢得尽量低,所以己方在采用一定的策略时得到的赢得,总是可能得到的赢得当中最小的那个,而最优策略应该使得最小的赢得达到最大.

由此,盟军所面对的决策问题(5)可以转化为

$$\max \min pM \quad (7)$$

其中 \min 是对 pM 的所有元素取极小.

类似地,德军所面对的决策问题(6)转化为

$$\min \max Mq^T \quad (8)$$

对于二人零和矩阵博弈可以证明,以上两个优化问题的最优值是一样的.

可以想到,对于双方效用之和为常数(不一定是0)的情况,二者仍然是严格竞争的,这种求解方法仍然有效.

为求解模型(7)可编写 LINGO 程序如下:

```
model:
sets:
k/1..3/:p;
n/1..2/:q;
pay(k,n):M;
endsets
data:
M=1 0
    2 0
    -2 1;
enddata
max=allies;
@ for(n(j):
    allies<@ sum(k(i):p(i)*m(i,j));
);
@ sum(k;p)=1;
End
```

得到的最优解为 $p_1 = 0, p_2 = 3/5, p_3 = 2/5$, 最优值为 $2/5$. 类似地,可得到

(8)的最优解为 $q_1 = 1/5, q_2 = 4/5$, 最优值也是 $2/5$.

这个最优解(策略组合)具有如下重要性质:没有任何一方能够通过单方面地偏离这一组合中自己的策略,来提高期望效用.也就是说,它就是(混合策略)纳什均衡.请读者注意,对于博弈问题,我们称其解为纳什均衡,而一般不称为最优解.

评注 1. 上面的博弈中是以净胜的战斗场次为依据建模的,可能战争中评价胜败的结局以最后德军能否被包围歼灭更合适,这时只需要将(1)式中的赢得矩阵改为

$$M = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ -1 & 1 \end{pmatrix} \quad (9)$$

就可以类似地讨论,结论也相似.此外,注意到(1)和(9)中矩阵 M 的第2行的数值总是不小于第1行,这说明盟军的行动2占优于行动1,因此实际上盟军肯定不会采用行动1,即行动1可以根本不考虑(按照混合策略求解也得到 $p_1 = 0$).一般地,通过这种占优性分析,可以降低问题的规模.

2. 当决策机会只有一次的时候,现实中采用混合策略似乎不太可行,总不能让决策者采取类似投掷硬币的办法来指挥战争吧!但至少可以看出,盟军应该以较大的可能性让预备队原地待命(行动2),而德军应该以较大的可能性选择撤退(行动2).二战中的实际情况是,盟军选择了让预备队原地待命(行动2),而德军则选择了向西进攻(行动1),结果是德军两场战斗均告失败,最后被包围歼灭.据说德军的指挥官们曾向其最高统帅希特勒建议选择撤退,但希特勒一意孤行,命令部队向西进攻,导致了德军覆灭的后果.也许希特勒另有关于整个战局的全盘考虑,或者是想出奇制胜背水一战,那就不是我们假设的完全信息博弈问题了.

3. 如果决策不是同时作出,例如德军先决定撤退而且该信息被盟军提前得到,那么盟军的最佳选择就变成了一个普通的决策问题而不是博弈,显然这时盟军的最佳选择是立即东进包围而不是原地待命错失良机.这说明在博弈问题中博弈规则是至关重要的,如参与人分别作出自己决策的时间顺序、决策时拥有哪些信息等.

4. 对于多人参与的或者效用函数之和不是常数的博弈问题,一般不能用上面的线性规划方法求解,但通过纳什均衡的定义,仍然可以得到纳什均衡解.

11.2 让报童订购更多的报纸

问题背景 在9.2节我们讨论过报童模型,那里只从报童本身的利益出发,研究每天应该从报社购进多少份报纸可以使他的期望收入最大.这一节要将报童和报社的利益放到一起考虑,研究怎样通过双方的博弈,让报童订购更多的报纸,使报童和报社的利益得以协调.

首先对报童模型作简单的重述和补充(数学符号有变化).设报纸每份的订购价为 w ,零售价为 p ,(未卖出的)处理价为 $v(p > w > v > 0)$,需求量是密度函数为 $f(x)$ 、分布函数为 $F(x)$ 的连续随机变量,且 $F(0) = 0$.如果报童订购 Q 份报纸,则期望销售量为

$$\begin{aligned} S(Q) &= \int_0^Q xf(x)dx + \int_Q^\infty Qf(x)dx \\ &= xF(x) \Big|_0^Q - \int_0^Q F(x)dx + Q(1 - F(Q)) = Q - \int_0^Q F(x)dx \end{aligned} \quad (1)$$

需要处理的期望存货量为

$$I(Q) = Q - S(Q) = \int_0^Q F(x)dx \quad (2)$$

于是报童的期望利润为

$$G(Q) = pS(Q) + vI(Q) - wQ = (p - v)S(Q) - (w - v)Q \quad (3)$$

报童的最优订购量 Q_r 满足 $G'(Q_r) = (p - v)S'(Q_r) - (w - v) = 0$,即

$$F(Q_r) = \frac{p - w}{p - v} \quad (4)$$

与9.2节的(4)式相同.

通常报纸是报童从报社订购的,假设报社每份报纸的成本价为 c ,可以合理地假定报童的订购价 w (即报社的批发价)满足 $w \geq c > v$.仍然假设信息是完全的(即双方均拥有以上信息),此时报社和报童的决策可以通过如下的两阶段动态博弈模型来描述:第一阶段,由报社决定批发价 w ;第二阶段,由报童决定订货量 Q .

动态博弈一般通过反向归纳法求解.首先考虑第二个阶段的问题,即对于给定的 w ,报童的最优订货量 Q_r 是 w 的函数,由(4)式确定,不妨记为 $Q_r(w)$,称为报童的最优反应,即报童的战略.然后,对于第一阶段,报社按照利润最大决定批发价 w ,其决策问题是

$$\max_{w \geq c} (w - c)Q_r(w) = (w - c)F^{-1}\left(\frac{p - w}{p - v}\right) \quad (5)$$

求解问题(5)可以得到最优批发价 w^* .此时,战略组合 $(w^*, Q_r(w^*))$ 就构成了这

个动态博弈的均衡,称为子博弈完美均衡。

在这个均衡结果下,报童和报社的整体利益是否得到了最大化?也就是说,如果将报童和报社的利益放到一起考虑,例如报社直销(或者等价地说,报社与报童是一个整体)并最大化自己的期望利润,报社和报童对没有卖完的报纸的处理价均为 v ,则报社印刷报纸的最优数量 Q^* 应满足(只需要将(4)中的 w 换成 c)

$$F(Q^*) = \frac{p-c}{p-v} \quad (6)$$

因为一般来讲 $w^* > c$,所以由(4)式给出的报童的最优订购量 $Q_r(w^*) < Q^*$.这表明从整体来看,报童的订购量不足.报社和报童双方能否想办法使报童的订购量提高到 Q^* 呢?如果能,我们就说整个系统达到了协调。

目前报社的决策变量只有批发价格 w ,只有让 $w=c$ 才能使报童把订购量提高到 Q^* ,而这相当于报社不赚钱,所以仅仅简单地降低批发价格一般不能使整个系统达到协调。

下面将以(6)式确定的 Q^* 作为报童的最优订购量,寻求使报社和报童双方利益达到协调的途径.最常用的办法之一是通过谈判达成对双方具有约束力的协议,对双方的行为进行协调.这类协调协议很多,我们下面只介绍几种简单的模型^[30]:一类协议是价格折扣协议,报童订购量越多单价越低;另一类协议是回收协议,报社以高于处理价(低于批发价)的价格回收全部卖不完的报纸,或者以批发价回收一定比例的卖不完的报纸。

价格折扣协议模型 价格折扣是一种常见的营销策略,在这一协议下,报童订购的数量 Q 越多,则单价越低,即批发价 $w_d(Q)$ 是 Q 的一个减函数.对于给定的折扣方案 $w_d(Q)$,由(3)式报童的期望利润为

$$U_r(w_d(Q)) = (p-v)S(Q) - (w_d(Q) - v)Q \quad (7)$$

而当报社与报童作为一个整体联合决策时,批发价 w 等于成本价 c ,利润函数为

$$U_{s,r}(Q) = (p-v)S(Q) - (c-v)Q \quad (8)$$

如果报社和报童经协商同意,报童按照批发价 $w_d(Q)$ 订购时,其利润占整体联合时利润的比例为 λ ,即

$$U_r(w_d(Q)) = \lambda U_{s,r}(Q), 0 < \lambda \leq 1 \quad (9)$$

这样,报童的决策就会与整体决策一致.将(7),(8)代入(9)得

$$w_d(Q) = \lambda c + (1-\lambda)[v + (p-v)S(Q)/Q] \quad (10)$$

这就是可以协调整体的价格折扣方案. λ 越大,报童的利润越大,报社的利润越小;当 λ 在 $0 < \lambda \leq 1$ 变化时,总利润在双方之间的任意分配比例都可以达到.容易证明, $w_d(Q)$ 确实是关于 Q 的减函数,但一般是非线性的,使用起来不一定方便。

回收协议模型一——回收价格协议 报童订购量不足的主要原因是害怕报纸订多了卖不出去,因此报社可以通过以一定的价格 b 回收卖不出去的报纸,使报童把自己的订购量提高到(6)式给出的 Q^* . 可以合理地假定回收价 b 满足 $w > b > v$.

具体来说,这是一个两阶段的博弈:首先报社给出批发价和回收价 (w, b) ,然后报童决定订购量 Q . 对于给定的 (w, b) ,只需要将(4)中的 v 换成 b ,即得最优的订购量 Q , 满足

$$F(Q_r) = \frac{p-w}{p-b} \quad (11)$$

根据(6)和(11)式,如果想得到 $Q_r = Q^*$, 则 (w, b) 应满足

$$\frac{p-c}{p-v} = \frac{p-w}{p-b} \quad (12)$$

由(12)式解出批发价 w , 可得

$$w = w_b(b) = b + \frac{c-v}{p-v}(p-b) \quad (13)$$

可以看出,批发价 w_b 是回收价 b 的增函数. 在协议参数为 (w_b, b) 时,报童的订货量将达到 Q^* , 整体的期望利润达到最优. 当然,对于不同的 b , 双方的利润会不同. 由(3)式报童的期望利润 $U_r(b)$ 为

$$U_r(b) = (p-b)S(Q^*) - (w_b(b) - b)Q^* \quad (14)$$

而由最优订货量 Q^* 和期望回收量 $I(Q^*)$ 及相应的价格,可得报社的期望利润 $U_s(b)$ 为

$$U_s(b) = (w_b(b) - c)Q^* - (b-v)I(Q^*) \quad (15)$$

将(13)代入(14), (15)并注意 $Q^* = S(Q^*) + I(Q^*)$ 得到

$$U_r(b) = \frac{p-b}{p-v} [U_r(b) + U_s(b)], \quad U_s(b) = \frac{b-v}{p-v} [U_r(b) + U_s(b)]$$

$$U_r(b) + U_s(b) = (p-v)S(Q^*) - (c-v)Q^* \quad (16)$$

可见, b 越大,报社的利润越大,报童的利润越小;当 b 在 $p > b > v$ 变化时,总利润在双方之间的任意分配比例都可以达到.

回收协议模型二——回收数量协议 回收协议的另一种形式是,如果报童有卖不出去的报纸,其数量不超过订货量的一定比例的部分报社按批发价 w 回收,超过这个比例的部分由报童自己处理.

设按批发价回收的比例为 α , 当报童订货量为 Q 时,未卖出的期望存量为 $I(Q) = \int_0^Q F(x) dx$ (见(2)式),它可以分成报社全价回收和报童自己处理两部分,分别表示为

$$I_1(Q) = \int_0^{(1-\alpha)Q} \alpha Q f(x) dx + \int_{(1-\alpha)Q}^Q (Q-x)f(x) dx = \int_{(1-\alpha)Q}^Q F(x) dx \quad (17)$$

$$I_2(Q) = I(Q) - I_1(Q) = \int_0^{(1-\alpha)Q} F(x) dx \quad (18)$$

对于给定的 (w, α) , 报童的期望利润为

$$\begin{aligned} U_r(w, \alpha, Q) &= pS(Q) + wI_1(Q) + vI_2(Q) - wQ \\ &= (p-w) \left[Q - \int_{(1-\alpha)Q}^Q F(x) dx \right] - (p-v) \int_0^{(1-\alpha)Q} F(x) dx \end{aligned} \quad (19)$$

最优的订货量 Q_r 满足 $\left. \frac{\partial U_r(w, \alpha, Q)}{\partial Q} \right|_{Q_r} = 0$, 即

$$(p-w)[1-F(Q_r)] - (w-v)(1-\alpha)F((1-\alpha)Q_r) = 0 \quad (20)$$

根据(20)和(6)式, 如果想得到 $Q_r = Q^*$, 则 (w, α) 应满足

$$(p-w) \left(1 - \frac{p-c}{p-v} \right) = (w-v)(1-\alpha)F((1-\alpha)Q^*) \quad (21)$$

解得

$$w = w_q(\alpha) = v + \frac{(p-v)(c-v)}{(c-v) + (p-v)(1-\alpha)F((1-\alpha)Q^*)} \quad (22)$$

可以看出, 批发价 w_q 是回收比例 α 的增函数. 在协议参数为 (w_q, α) 时, 整个系统的期望总利润达到了最优. 还可以证明, α 变大时报社的利润变大, 报童的利润变小; 当 α 在 $0 \leq \alpha \leq 1$ 变化时, 总利润在双方之间的任意分配比例都可以达到.

评注 上面对协议的讨论中, 都是通过报童的最优反应直接得到了相应的协调条件(10), (13)和(22), 即只利用了求解动态博弈模型第二阶段的结果. 由于通过调整协议参数 $\lambda, (w_q(b), b), (w_q(\alpha), \alpha)$ 可使总利润在双方之间任意分配, 所以协议参数不能让报社或报童单方面决定, 而一般是由双方在第一阶段通过谈判决定(谈判可以看成是合作博弈的内容, 参见 11.5, 11.6 节).

还有另外一种协调报社和报童双方利益的途径: 报社直接以成本价向报童批发报纸, 同时向报童收取一定的费用如加盟费作为自己的利润, 当然, 如何确定加盟费也需要双方谈判.

11.3 “一口价”的策略

问题背景 外出旅游时人们常常为了买一点纪念品, 与商店或小商贩们反复地讨价还价, 很浪费时间. 当然, 也有人把这当作一种乐趣, 又另当别论. 有家纪念品商店为了节省顾客和商家双方的宝贵时间, 推出了一项新的销售策略: 双

方同时给出报价,如果顾客的出价不低于商家的卖价,则成交,并且成交价等于双方报价的平均值;否则不成交.在这种“一口价”的情况下,双方应该如何报价?

模型假设 1. 商家知道商品对自己的真实价值 v_s ,也就是可以卖出的最低价格;顾客知道商品对自己的真实价值 v_b ,也就是可以支付的最高价格.

2. 商家不知道商品对顾客的真实价值 v_b ,但知道其概率分布;顾客不知道商品对商家的真实价值 v_s ,但也知道其概率分布.

3. 不妨假设 v_s, v_b 都服从 $[0, 1]$ 上的均匀分布.

4. 对一组给定的 (v_s, v_b) 如果以价格 p 成交,该交易对商家和顾客的效用分别为 $p - v_s, v_b - p$;如果不成交,双方的效用均为 0. 商家和顾客都希望最大化自己的期望效用.

5. 以上信息为双方所共有.

模型建立 记商家的战略为 $p_s(v_s)$,即当商家认为商品的值为 v_s 时,他给出卖价 $p_s(v_s)$;记顾客的战略为 $p_b(v_b)$,即当顾客认为商品的值为 v_b 时,他给出报价 $p_b(v_b)$. 自然地, $p_b(v_b)$ 和 $p_s(v_s)$ 都应该是定义在 $[0, 1]$ 区间上、取值也在 $[0, 1]$ 区间上的非减函数.

对于任意给定的 $v_s \in [0, 1]$,商家的报价 $p_s(v_s)$ 应该使其期望利润最大. 因为只有 $p_b(v_b) \geq p_s(v_s)$ 时才能成交,成交后商家的利润为 $(p_s(v_s) + p_b(v_b))/2 - v_s$,而不成交时利润为 0,所以 $p_s(v_s)$ 应满足

$$\max_{p_s} \left\{ \frac{p_s + E[p_b(v_b) | p_b(v_b) \geq p_s]}{2} - v_s \right\} P\{p_b(v_b) \geq p_s\} \quad (1)$$

这里 $E[\cdot]$ 表示的是条件 $p_b(v_b) \geq p_s$ 下 $p_b(v_b)$ 的条件期望, $P\{\cdot\}$ 表示事件的概率.

类似地,对于任意给定的 $v_b \in [0, 1]$,顾客的报价 $p_b(v_b)$ 应该使其期望赢得最大,成交后顾客的赢得为 $v_b - (p_s(v_s) + p_b(v_b))/2$,不成交时赢得为 0,所以 $p_b(v_b)$ 应满足

$$\max_{p_b} \left\{ v_b - \frac{p_b + E[p_s(v_s) | p_b \geq p_s(v_s)]}{2} \right\} P\{p_b \geq p_s(v_s)\} \quad (2)$$

如果战略组合 $(p_s(v_s), p_b(v_b))$ 同时满足 (1) 和 (2), 则是双方的一个均衡. 对于这个博弈问题存在很多均衡,下面介绍其中两个比较简单的均衡.

单一价格均衡 设定 $(0, 1)$ 区间上一个数 x ,商家如果认为商品的值 $v_s \leq x$, 则报价 x , 否则报价为 1; 顾客如果认为商品的值 $v_b \geq x$, 则报价 x , 否则报价为 0. 这种价格战略可表示为

$$p_s(v_s) = \begin{cases} x, & v_s \leq x \\ 1, & v_s > x \end{cases} \quad (3)$$

$$p_b(v_b) = \begin{cases} x, & v_b \geq x \\ 0, & v_b < x \end{cases} \quad (4)$$

战略组合 $(p_s(v_s), p_b(v_b))$ 是否同时满足(1)和(2)呢? 答案是肯定的。

首先, 可以注意到成交价格只能发生在价格 x 。

此外, 从商家的角度看, 如果顾客坚持战略(4), 则商家在 $v_s \leq x$ 时报价 x 是他的最优反应。因为报价低于 x 显然使自己的利润降低(假设能成交); 而报价高于 x 则不能成交, 自己本来可以从成交中获得的利润不能实现。如果 $v_s > x$, 则成交会使商家利润为负, 商家当然不希望成交, 而报价为 1 可以保证不成交(连续分布下双方都报价 1 的可能性为 0, 可以不考虑)。因此战略(3)是商家对顾客的战略(4)的最优反应。

同理, 如果商家坚持战略(3), 战略(4)是顾客的最优反应。因此, (3)和(4)给出的战略组合是一个均衡, 称为单一价格均衡。

对一组给定的 (v_s, v_b) , 当 $v_s < v_b$ 时称为交易是有利的, 因为此时一定存在 $p \in (v_s, v_b)$, 当双方以价格 p 交易时, 对双方都是有利的(由模型假设 4), 交易给双方带来的效用之和(即 $v_b - v_s$)称为交易价值。在给定的战略组合下, 能够实际发生的交易的期望价值与有利的全部交易的期望价值的比值称为该战略的交易效率。

下面分析单一价格战略的交易效率。显然, 当且仅当 $v_s \leq x \leq v_b$ 时交易实际上才能发生。若 v_s, v_b 都服从 $[0, 1]$ 上的均匀分布, 图 1 中对角线上的三角形是交易有利的区域, 而只有标出“交易”的矩形才是交易实际发生的区域, 所以交易效率为

$$\eta = \frac{\int_x^1 \int_0^x (v_b - v_s) dv_s dv_b}{\int_0^1 \int_0^{v_b} (v_b - v_s) dv_s dv_b}$$

$$= 3x(1-x) \leq \frac{3}{4} \quad (5)$$

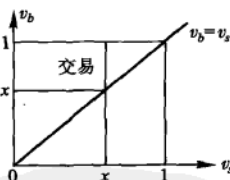


图 1 单一价格战略的交易效率

显然当 $x = 0.5$ 时交易效率最大, 但最大效率也只有 $3/4$ 。

线性价格均衡 假设商家和顾客的报价分别是商品对二者价值的线性函数, 表示为

$$p_s(v_s) = a_s + c_s v_s \quad (6)$$

$$p_b(v_b) = a_b + c_b v_b \quad (7)$$

让我们看看能否确定其中的系数(不妨假设均为正数) a_s, c_s, a_b, c_b , 使这个战略组合 $(p_s(v_s), p_b(v_b))$ 同时满足(1)和(2), 即构成一个均衡。

假设商家的战略为(6), 由假设 3 知 p_s 服从 $[a_s, a_s + c_s]$ 上的均匀分布。此时对于给定的 v_b , 顾客的最优反应就是寻找满足(2)式的 p_b 。当 $p_b \in [a_s, a_s + c_s]$

时, $P\{p_b \geq p_s(v_s)\} = (p_b - a_s)/c_s$, $E[p_s | p_b \geq p_s] = (a_s + p_b)/2$, 于是(2)式为

$$\max_{p_b} \left\{ v_b - \frac{p_b + (a_s + p_b)/2}{2} \right\} \cdot \frac{p_b - a_s}{c_s} \quad (8)$$

这是一个二次函数的优化, 其最优解为

$$p_b = \frac{2}{3}v_b + \frac{1}{3}a_s \quad (9)$$

类似地, 假设顾客的战略为(7), 则对于给定的 v_s , 当 $p_s \in [a_b, a_b + c_b]$ 时, 由(1)式可得商家的最优反应为

$$p_s = \frac{2}{3}v_s + \frac{1}{3}(a_b + c_b) \quad (10)$$

比较(6), (7), (9), (10), 可以解出

$$a_b = \frac{1}{12}, a_s = \frac{1}{4}, c_b = c_s = \frac{2}{3} \quad (11)$$

即线性价格战略(6), (7)为

$$p_s(v_s) = \frac{2}{3}v_s + \frac{1}{4} \quad (12)$$

$$p_b(v_b) = \frac{2}{3}v_b + \frac{1}{12} \quad (13)$$

理论上来说, 在考虑端点条件时, 上述(12)式应该只对 $p_s(v_s) \in [a_b, a_b + c_b] = [\frac{1}{12}, \frac{3}{4}]$, 即 $v_s \leq \frac{3}{4}$ 有效, (13)式应该只对 $p_b(v_b) \in [a_s, a_s + c_s] = [\frac{1}{4}, \frac{11}{12}]$, 即 $v_b \geq \frac{1}{4}$ 有效, 否则双方总有一方不会愿意成交. 考虑到 $v_s > \frac{3}{4}$ 或 $v_b < \frac{1}{4}$ 时即使按(12), (13)报价, 也一定不会成交, 因此, 由(12), (13)确定的战略 $(p_s(v_s), p_b(v_b))$ 是整个 $[0, 1]$ 区间上的一个均衡的战略组合(参见图2).

下面分析线性价格战略的交易效率. 显然, 当且仅当 $p_s(v_s) \leq p_b(v_b)$ 时交易实际上才能发生, 将(12), (13)代入得到交易条件为 $v_b \geq v_s + \frac{1}{4}$, 在图3上标出了有利的交易中实际发生的区域(小三角形), 所以其交易效率为

$$\eta = \frac{\int_{\frac{1}{4}}^1 \int_0^{v_b - \frac{1}{4}} (v_b - v_s) dv_s dv_b}{\int_0^1 \int_0^{v_b} (v_b - v_s) dv_s dv_b} = \frac{27}{32} > \frac{3}{4} \quad (14)$$

可见, 线性价格战略的交易效率大于单一价格战略的交易效率. 更有意义的是, 比较图1和图3可以看出, 线性价格战略中包含了所有交易价值大于 $1/4$ 的交易, 交易有利但不能成交的都是交易价值不大的. 而在单一价格战略中, 有些交

易价值很小的交易成交了,也有些交易价值很大的却未能成交(即使取 $x = 1/2$, 也可能漏掉交易价值接近 $1/2$ 的交易)。

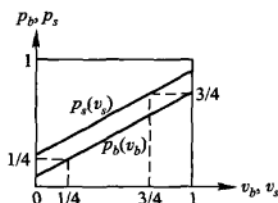


图2 均衡的线性价格战略

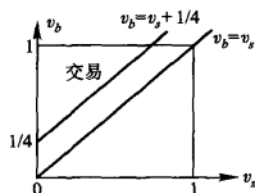


图3 线性价格战略的交易效率

评注 1. 这里讨论的模型一般称为双向拍卖^[18],是一个同时出价的博弈(静态博弈),而且信息不完全(双方的真实价值是各自的私有信息,对方只知道其分布),所以是一个不完全信息静态博弈.可以看出,这个问题不仅是对某个具体交易(给定的 v_s, v_b) 提供报价决策,而是要对所有可能的 v_s, v_b 提供一个报价方案 $(p_s(v_s), p_b(v_b))$,这才是一个均衡(战略组合).对不完全信息静态博弈,这样的均衡一般称为贝叶斯均衡或贝叶斯-纳什均衡.

2. 是否存在比线性战略均衡的交易效率更高的贝叶斯均衡?学者们已经证明,答案是否定的.这也意味着,不存在使所有有利的交易都发生的均衡战略组合(而且已经证明这一结论对一般的连续分布也成立).也就是说,与信息完全(对称信息)的情形相比,信息的不完全(非对称信息)降低了交易效率.

11.4 不患寡而患不均

问题背景 互不相识的甲乙两人获得了总额一定的一笔财富(如100元),假设双方决定按如下程序分配:首先由甲(“提议者”)拿出一个分配提议——分给乙 s 元钱,剩下的 $1-s$ 留给甲自己;其次,由乙(“反应者”)决定是否接受这个提议,如果接受,则按此分配,否则双方什么也得不到(如100元钱被其他人没收).

分配程序是公开的,两人决策有先后,所以是一个很简单的完全信息动态博弈.假设双方都只关心自己的所得,即各自的效用等于自己所得,按照经典博弈论的理论,轮到乙决策时,乙应该接受甲给出的任何分配提议,因为他如果不接受提议,就什么也得不到.于是甲应该提议 $s=0$,即均衡结果是: $s=0$,乙接受.如果对乙来讲要求他接受时的效用严格大于不接受时的效用,则 $s=1$ 分钱也就可以了.

这个博弈称为最后通牒博弈(Ultimatum Game),学者们为了检验现实中人们是否会真的按照经典博弈论导出的均衡进行决策,在世界各地对于不同性别、不同文化、不同富裕程度的人进行了大量实验.结果表明,绝大多数人的决策与

上述均衡相差很远. 首先, 甲提议的分配比例的一般位于 40% ~ 50% 之间, 而不是接近于 0; 其次, 乙经常会拒绝甲给出的低于 20% 的提议, 比例越小, 越容易被乙拒绝.

对于实验结果与理论预测不一致的矛盾, 有很多学者提出了各种解释, 如认为这是因为实验者不懂博弈论, 但这很难解释为什么这么多人都会犯同样的错误. 一种比较有说服力的解释是, 经典博弈论把每个参与者的赢得作为他的效用函数, 这相当于认为人们只关心自己的实际所得, 即人是“绝对自私”而且“完全理性”的, 而现实中人们在决策时不仅仅考虑自己的得失, 还关注自己的感受, 例如人们可能还具有“利他”与“互惠”思想, 特别会关注分配是否公平. 中国有句老话“不患寡而患不均”, 就是表达类似的道理. 下面介绍一个考虑这种公平性的模型^[61].

模型假设与建立 设甲乙二人按如下程序分配总额为 1 的财富: 甲提议分给乙 s ($0 \leq s \leq 1$), 自己留 $1-s$; 乙如果接受提议, 则甲乙二人所得分别为 $x_1 = 1-s$, $x_2 = s$, 否则 $x_1 = x_2 = 0$.

在构造两人的效用函数时, 假定他们除了考虑自己的所得 x_1, x_2 以外, 还都偏爱公平, 具体表现为: 如果某人所得比对方少, 他因“愤怒”使效用降低; 如果某人所得比对方多, 他因“愧疚”也使效用降低. 用 α_1, α_2 (≥ 0) 分别表示两人的“愤怒”系数, β_1, β_2 (≥ 0) 分别表示两人的“愧疚”系数, 并且不妨假设 $\alpha_i \geq \beta_i, i = 1, 2$. 建立两人的效用函数为

$$U_i(x_1, x_2) = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}, i = 1, 2, j = 3 - i \quad (1)$$

还可以进一步假设 $\beta_i < 1/2$, 否则, 当 i 的所得比对手 j 多, 即 $x_i > x_j = 1 - x_i$ 时, i 的效用函数是 $x_i - \beta_i(x_i - x_j) = \beta_i - (2\beta_i - 1)x_i$, 关于 i 自己的所得 x_i 的系数非正, 因此 i 宁愿将自己多得的部分 (即使是很小的一部分) 全部让给对方, 这种过分“愧疚”的情形一般也是不符合实际的.

由于财富的分配 x_1, x_2 实际上只与 s 有关, 所以下面将效用函数 (1) 简记为 $U_1(s)$ 和 $U_2(s)$.

模型求解 首先讨论乙的最优反应. 对于给定的 s , 如果他不接受, 则 $x_1 = x_2 = 0, U_1(s) = U_2(s) = 0$.

如果乙接受, 即 $x_1 = 1-s, x_2 = s$. 若 $s \geq 1/2$, 则 $x_2 \geq x_1$, 于是

$$U_2(s) = s - \beta_2(2s - 1) \quad (2)$$

由 $\beta_2 < 1/2$ 可知 (2) 式中的 $U_2(s) \geq 1/2 > 0$, 所以乙的最优反应是接受.

若 $s \leq 1/2$, 则 $x_2 \leq x_1$, 于是

$$U_2(s) = s - \alpha_2(1 - 2s) = (1 + 2\alpha_2)s - \alpha_2 \quad (3)$$

仅当 (3) 式中的 $U_2(s) \geq 0$, 即 $s \geq \alpha_2 / (1 + 2\alpha_2)$ 时, 乙的最优反应才是接受; 否

则,乙不会接受.记 $\bar{s}(\alpha_2) = \alpha_2 / (1 + 2\alpha_2)$, 容易看出 $0 \leq \bar{s}(\alpha_2) < 1/2$.

可以发现,当 $s = 1/2$ 时(2)和(3)是一致的,所以特例 $s = 1/2$ 放到上面的哪种情况讨论都是一样的,以下也类似,不再特别说明.

现在讨论甲的决策.由于乙不接受时双方的效用都是0,而甲显然有可行的提议可以让乙接受并且使自己的效用为正(如 $s = 1/2$),所以只需要考虑乙接受提议的情形.分两种情况讨论:

情况1 甲知道乙的“愤怒”系数 α_2 .

若 $s \geq 1/2$, 则 $x_2 \geq x_1$, 于是

$$U_1(s) = 1 - s - \alpha_1(2s - 1) \quad (4)$$

$U_1(s)$ 在 $s^* = 1/2$ 时达到最大值 $1/2$, 所以只需要讨论 $s \leq 1/2$ 且 $s \geq \bar{s}(\alpha_2)$ 的情况. 此时 $x_2 \leq x_1$, 于是

$$U_1(s) = 1 - s - \beta_1(1 - 2s) = 1 - \beta_1 + (2\beta_1 - 1)s, \quad \bar{s}(\alpha_2) \leq s \leq 1/2 \quad (5)$$

$U_1(s)$ 是 s 的线性函数, 由 $\beta_1 < 1/2$ 知最优值在左端点取得, 于是甲的最佳决策 s^* 应该是

$$s^* = \bar{s}(\alpha_2) = \alpha_2 / (1 + 2\alpha_2) \quad (6)$$

可见,甲提议给乙的比例为 $\bar{s}(\alpha_2)$, 严格小于 50%. 此外, $\bar{s}(\alpha_2)$ 是乙的“愤怒”系数 α_2 的增函数, α_2 越大, 甲提议分给乙的份额就会越高, 这是符合人们直觉的.

情况2 甲不知道乙的“愤怒”系数 α_2 , 但知道 α_2 的概率分布.

若 $s \geq 1/2$, 则 $x_2 \geq x_1$, 因为此时乙一定会接受甲的提议, 甲的效用仍如(4)式所示, 在 $s^* = 1/2$ 时达到最大值 $1/2$, 所以仍只需要讨论 $s \leq 1/2$ 的情况.

甲知道 $s \geq \bar{s}(\alpha_2)$ 时乙才会接受甲的提议, 且 $s \geq \bar{s}(\alpha_2) = \alpha_2 / (1 + 2\alpha_2)$ 等价于 $\alpha_2 \leq s / (1 - 2s)$. 若设 α_2 的概率分布函数为 $F(\alpha_2)$, 且 $F(\underline{\alpha}) = 0, F(\bar{\alpha}) = 1$ ^①, 则甲可以推测乙接受甲的提议 s 的概率 p 为

$$p = \begin{cases} 0, & s \leq \bar{s}(\underline{\alpha}) \\ F(s/(1-2s)), & \bar{s}(\underline{\alpha}) < s < \bar{s}(\bar{\alpha}) \\ 1, & s \geq \bar{s}(\bar{\alpha}) \end{cases} \quad (7)$$

这个概率是关于 s 的非减函数, 即甲提议的 s 越大, 越可能被乙接受, 这与实验中观察到的现象完全吻合.

于是由(5)和(7)式, 如果甲提议 s , 其期望效用为

^① 严格地说, $\underline{\alpha} = \max\{\alpha | F(\alpha) = 0\}$, $\bar{\alpha} = \min\{\alpha | F(\alpha) = 1\}$.

$$EU_1(s) = \begin{cases} 0, & s \leq \bar{s}(\underline{\alpha}) \\ [1 - \beta_1 + (2\beta_1 - 1)s]F(s/(1-2s)), & \bar{s}(\underline{\alpha}) < s < \bar{s}(\bar{\alpha}) \\ 1 - \beta_1 + (2\beta_1 - 1)s, & s \geq \bar{s}(\bar{\alpha}) \end{cases} \quad (8)$$

甲应该最大化(8)式表示的期望效用。考虑到 $\beta_1 < 1/2$ 及(8)式中 $EU_1(s)$ 是连续函数,只需要考虑在区间 $\bar{s}(\underline{\alpha}) < s \leq \bar{s}(\bar{\alpha})$ 上最大化 $[1 - \beta_1 + (2\beta_1 - 1)s]F(s/(1-2s))$ 即可,其最优解 s^* 就是甲的最优战略。

与情况 1 类似,甲提议给乙的比例不超过 $\bar{s}(\bar{\alpha})$, 严格小于 50%。

评注 这个模型对公平性进行了量化,很好地解释了实验结果与理论预测不一致的矛盾。通过建立数学模型解释存在的现象,这是建模的重要功能之一。可以相信,如果一个模型能够很好地解释存在的现象,这个模型应该对未来的现象 also 具有很强的预测能力。

二人分配财富的模型可以推广到有 n 个参与人,记参与人 i 得到的财富为 x_i , $\mathbf{x} = (x_1, x_2, \dots, x_n)$, 定义 i 的效用函数为

$$U_i(\mathbf{x}) = x_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i} \max\{x_j - x_i, 0\} - \beta_i \frac{1}{n-1} \sum_{j \neq i} \max\{x_i - x_j, 0\}, i = 1, 2, \dots, n \quad (9)$$

其中 α_i, β_i 的含义与前面相同。与二人情形假设 $\beta_1 < 1/2$ 类似,对于多人情形一般假设 $\beta_i < 1$, 主要理由是:当 i 的所得比其他人多而其他人所得相同,即对任意 $j \neq i, x_i > x_j = (1 - x_i)/(n-1)$ 时, i 的效用函数是

$$\begin{aligned} U_i(\mathbf{x}) &= x_i - \beta_i \frac{1}{n-1} \sum_{j \neq i} (x_i - x_j) = x_i - \beta_i \left(x_i - \frac{1 - x_i}{n-1} \right) \\ &= \frac{\beta_i}{n-1} + \left(1 - \frac{n\beta_i}{n-1} \right) x_i \end{aligned} \quad (10)$$

为了避免该效用函数关于 i 自己的所得 x_i 的系数非正(过分“愧疚”的情形),一般假设 $1 - \frac{n\beta_i}{n-1} > 0$, 即 $\beta_i < \frac{n-1}{n} < 1$ 。

11.5 效益的合理分配

在经济或社会活动中若干实体(如个人、公司、党派、国家等)相互合作结成联盟或集团,常能比他们单独行动获得更多的经济或社会效益。确定合理地分配这些效益的方案是促成合作的前提。先看一个简单例子。

甲乙丙三人经商。若单干,每人仅能获利 1 元;甲乙合作可获利 7 元;甲丙合

作可获利 5 元;乙丙合作可获利 4 元;三人合作则可获利 11 元.问三人合作时怎样合理地分配 11 元的收入.

人们自然会想到的一种分配方法是:设甲乙丙三人各得 x_1, x_2, x_3 元,满足

$$x_1 + x_2 + x_3 = 11 \quad (1)$$

$$x_1, x_2, x_3 \geq 1, \quad x_1 + x_2 \geq 7, \quad x_1 + x_3 \geq 5, \quad x_2 + x_3 \geq 4 \quad (2)$$

(2)式表示这种分配必须不小于单干或二人合作时的收入.但是容易看出(1),(2)有许多组解,如 $(x_1, x_2, x_3) = (5, 3, 3), (4, 4, 3), (4, 3.5, 3.5)$ 等.于是应该寻求一种圆满的分配方法.

上例提出的这类问题称为 n 人合作对策 (Cooperative n -person Game). L. S. Shapley^①1953 年给出了解决该问题的一种方法,称 **Shapley 值**^[64,66].

n 人合作对策和 Shapley 值 n 个人从事某项经济活动,对于他们之中若干人组合的每一种合作(特别,单人也视为一种合作),都会得到一定的效益,当人们之间的利益是非对抗性时,合作中人数的增加不会引起效益的减少.这样,全体 n 个人的合作将带来最大效益. n 个人的集合及各种合作的效益就构成 n 人合作对策,Shapley 值是分配这个最大效益的一种方案.正式的定义如下.

设集合 $I = \{1, 2, \dots, n\}$, 如果对于 I 的任一子集 s 都对应着一个实值函数 $v(s)$, 满足

$$v(\emptyset) = 0 \quad (3)$$

$$v(s_1 \cup s_2) \geq v(s_1) + v(s_2), \quad s_1 \cap s_2 = \emptyset \quad (4)$$

称 $[I, v]$ 为 n 人合作对策, v 为对策的特征函数.

在上面所述经济活动中, I 定义为 n 人集合, s 为 n 人集合中的任一种合作, $v(s)$ 为合作 s 的效益.

用 x_i 表示 I 的成员 i 从合作的最大效益 $v(I)$ 中应得到的一份收入. $x = (x_1, x_2, \dots, x_n)$ 叫做合作对策的分配 (Imputation), 满足

$$\sum_{i=1}^n x_i = v(I) \quad (5)$$

$$x_i \geq v(i), \quad i = 1, 2, \dots, n \quad (6)$$

请读者解释(6)式的含义.显然,由(3),(4)定义的 n 人合作对策 $[I, v]$ 通常有无穷多个分配.

Shapley 值由特征函数 v 确定,记作 $\Phi(v) = (\varphi_1(v), \varphi_2(v), \dots, \varphi_n(v))$. 对于任意的子集 s , 记 $x(s) = \sum_{i \in s} x_i$, 即 s 中各成员的分配.对一切 $s \subset I$, 满足 $x(s) \geq$

① L. S. Shapley (1923—) 美国数学家、经济学家,在数学经济学、对策论方面有突出贡献.二战期间,他在哈佛大学求学时入伍,曾来中国参战.1948 年哈佛大学毕业,1953 年在普林斯顿大学获博士学位,1954—1981 年在兰德公司任职,1981 年至今为加州大学洛杉矶分校教授.

$v(s)$ 的 x 组成的集合称 $[I, v]$ 的核心 (Core). 当核心存在时, 即所有 s 的分配都不小于 s 的效益, 可以将 Shapley 值作为一种特定的分配, 即 $\varphi_i(v) = x_i$.

Shapley 首先提出看来毫无疑义的几条公理, 然后用逻辑推理的方法证明, 存在唯一的满足这些公理的分配 $\Phi(v)$, 并把它构造出来. 这里只给出 $\Phi(v)$ 的结果, Shapley 公理可参看 [11, 83].

Shapley 值 $\Phi(v) = (\varphi_1(v), \varphi_2(v), \dots, \varphi_n(v))$ 为

$$\varphi_i(v) = \sum_{s \in S_i} w(|s|) [v(s) - v(s \setminus i)], \quad i = 1, 2, \dots, n \quad (7)$$

$$w(|s|) = \frac{(n - |s|)! (|s| - 1)!}{n!} \quad (8)$$

其中 S_i 是 I 中包含 i 的所有子集, $|s|$ 是子集 s 中的元素数目 (人数), $w(|s|)$ 是加权因子, $s \setminus i$ 表示 s 去掉 i 后的集合.

我们用这组公式计算本节开始给出的三人经商问题的分配, 以此解释公式的用法和意义.

甲乙丙三人记为 $I = \{1, 2, 3\}$, 经商获利定义为 I 上的特征函数, 即 $v(\emptyset) = 0, v(1) = v(2) = v(3) = 1, v(1, 2) = 7, v(1, 3) = 5, v(2, 3) = 4, v(I) = 11$. 容易验证 v 满足 (3), (4). 为计算 $\varphi_1(v)$ 首先找出 I 中包含 1 的所有子集 $S_1: \{1\}, \{1, 2\}, \{1, 3\}, I$, 然后令 s 跑遍 S_1 , 将计算结果记入表 1. 最后将表中末行相加得 $\varphi_1(v) = 13/3$. 同法可计算出 $\varphi_2(v) = 23/6, \varphi_3(v) = 17/6$. 它们可作为按照 Shapley 值方法计算的甲乙丙三人应得的分配.

让我们通过此例对 (7) 式作些解释. 对表 1 中的 s , 比如 $\{1, 2\}$, $v(s)$ 是有甲 (即 $\{1\}$) 参加时合作 s 的获利, $v(s \setminus 1)$ 是无甲参加时合作 s (只剩下乙) 的获利, 所以 $v(s) - v(s \setminus 1)$ 可视为甲对这一合作的“贡献”. 用 Shapley 值计算的甲的分配 $\varphi_1(v)$ 是, 甲对他所参加的所有合作 (S_1) 的贡献的加权平均值, 加权因子 $w(|s|)$ 取决于这个合作 s 的人数. 通俗地说就是按照贡献取得报酬.

表 1 三人经商中甲的分配 $\varphi_1(v)$ 的计算

s	1	$\{1, 2\}$	$\{1, 3\}$	I
$v(s)$	1	7	5	11
$v(s \setminus 1)$	0	1	1	4
$v(s) - v(s \setminus 1)$	1	6	4	7
$ s $	1	2	2	3
$w(s)$	1/3	1/6	1/6	1/3
$w(s) [v(s) - v(s \setminus 1)]$	1/3	1	2/3	7/3

Shapley 值方法可以有效地处理经济和社会合作活动中的利益分配问题. 请看下面的例子.

污水处理费用的合理分担 沿河有三城镇 1, 2 和 3, 地理位置如图 1 所示. 污水需处理后才能排入河中. 三城镇既可以单独建立污水处理厂, 也可以联合建厂, 用管道将污水集中处理(污水应由河流的上游城镇向下游城镇输送). 用 Q 表示污水量(单位: t/s), L 表示管道长度(单位: km), 按照经验公式, 建立处理厂的费用为 $P_1 = 73Q^{0.712}$ 千元, 铺设管道费用为 $P_2 = 0.66Q^{0.51}L$ 千元. 已知三城镇污水量为 $Q_1 = 5, Q_2 = 3, Q_3 = 5, L$ 的数值如图 1 所示. 试从节约总投资的角度为三城镇制定污水处理方案. 如果联合建厂, 各城镇如何分担费用^[51].

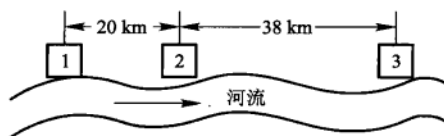


图 1 三城镇地理位置示意图

三城镇污水处理共有以下 5 种方案, 计算出投资费用以作比较.

1) 分别建厂. 投资分别为

$$C(1) = 73 \times 5^{0.712} = 230, C(2) = 160, C(3) = 230, \text{总投资 } D_1 = C(1) + C(2) + C(3) = 620.$$

2) 1, 2 合作, 在城 2 建厂. 投资为

$$C(1, 2) = 73 \times (5 + 3)^{0.712} + 0.66 \times 5^{0.51} \times 20 = 350, \text{总投资 } D_2 = C(1, 2) + C(3) = 580.$$

3) 2, 3 合作, 在城 3 建厂. 投资为

$$C(2, 3) = 73 \times (3 + 5)^{0.712} + 0.66 \times 3^{0.51} \times 38 = 365, \text{总投资 } D_3 = C(1) + C(2, 3) = 595,$$

4) 1, 3 合作, 在城 3 建厂. 投资为

$$C(1, 3) = 73 \times (5 + 5)^{0.712} + 0.66 \times 5^{0.51} \times 58 = 463, \text{这个费用超过了 1, 3 分别建厂的费用 } C(1) + C(3) = 460. \text{合作没有效益, 不可能实现.}$$

$$5) \text{ 三城合作, 在城 3 建厂. 总投资为 } D_5 = C(1, 2, 3) = 73 \times (5 + 3 + 5)^{0.712} + 0.66 \times 5^{0.51} \times 20 + 0.66 \times (5 + 3)^{0.51} \times 38 = 556.$$

比较结果以 $D_5 = 556$ 千元最小, 所以应选择联合建厂方案. 下面的问题是如何分担费用 D_5 .

总费用 D_5 中有 3 部分: 联合建厂费 $d_1 = 73 \times (5 + 3 + 5)^{0.712} = 453$; 城 1 至 2

的管道费 $d_2 = 0.66 \times 5^{0.51} \times 20 = 30$; 城 2 至 3 的管道费 $d_3 = 0.66 \times (5 + 3)^{0.51} \times 38 = 73$. 城 3 提出, d_1 由三城按污水量比例 5:3:5 分担, d_2, d_3 是为城 1, 2 铺设的管道费, 应由他们担负; 城 2 同意, 并提出 d_3 由城 1, 2 按污水量之比 5:3 分担, d_2 则应由城 1 自己担负; 城 1 提不出反对意见, 但他们计算了一下按上述办法各城应分担的费用:

$$\text{城 3 分担费用为 } d_1 \times \frac{5}{13} = 174;$$

$$\text{城 2 分担费用为 } d_1 \times \frac{3}{13} + d_3 \times \frac{3}{8} = 132;$$

$$\text{城 1 分担费用为 } d_1 \times \frac{5}{13} + d_3 \times \frac{5}{8} + d_2 = 250.$$

结果表明城 2, 3 分担的费用均比他们单独建厂费用 $C(2), C(3)$ 小, 而城 1 分担的费用却比 $C(1)$ 大. 显然, 城 1 不能同意这种分担总费用的办法.

为了促成三城联合建厂以节约总投资, 应该寻求合理分担总费用的方案. 三城的合作节约了投资, 产生了效益, 是一个 n 人合作对策问题, 可以用 Shapley 值方法圆满地分配这个效益.

把分担费用转化为分配效益, 就不会出现城 1 联合建厂分担的费用反比单独建厂费用高的情况. 将三城镇记为 $I = (1, 2, 3)$, 联合建厂比单独建厂节约的投资定义为特征函数. 于是有

$$v(\emptyset) = 0, v(1) = v(2) = v(3) = 0$$

$$v(1, 2) = C(1) + C(2) - C(1, 2) = 230 + 160 - 350 = 40$$

$$v(2, 3) = C(2) + C(3) - C(2, 3) = 160 + 230 - 365 = 25$$

$$v(1, 3) = 0$$

$$\begin{aligned} v(I) &= C(1) + C(2) + C(3) - C(1, 2, 3) \\ &= 230 + 160 + 230 - 556 = 64 \end{aligned}$$

三城联合建厂的效益为 64 千元. 用 Shapley 值作为这个效益的分配, 城 1 应分得的份额 $\varphi_1(v)$ 的计算结果列入表 2, 得到 $\varphi_1(v) = 19.7$. 类似地算出 $\varphi_2(v) = 32.1, \varphi_3(v) = 12.2$. 可以验证, $\varphi_1(v) + \varphi_2(v) + \varphi_3(v) = 64 = v(I)$. 看来, 城 2 从总效益 64 千元中分配的份额最大, 你能从城 2 的地理位置与合作对策的角度解释这个结果吗.

表 2 污水处理问题中 $\varphi_1(v)$ 的计算

s	1	$\{1, 2\}$	$\{1, 3\}$	I
$v(s)$	0	40	0	64
$v(s \setminus 1)$	0	0	0	25

续表

$v(s) - v(s \setminus 1)$	0	40	0	39
$ s $	1	2	2	3
$w(s)$	1/3	1/6	1/6	1/3
$w(s)[v(s) - v(s \setminus 1)]$	0	6.7	0	13

最后,在联合建厂方案总投资额 556 千元中各城的分担费用为:城 1 是 $C(1) - \varphi_1(v) = 230 - 19.7 = 210.3$; 城 2 是 $C(2) - \varphi_2(v) = 127.9$; 城 3 是 $C(3) - \varphi_3(v) = 217.8$.

Shapley 值方法的缺点及其他解决办法 Shapley 值方法以严格的公理为基础,在处理合作对策的分配问题时具有公正、合理等优点,但是它需要知道所有合作的获利,即要定义 $I = \{1, 2, \dots, n\}$ 的所有子集(共 2^n 个)的特征函数,这在实际上常常做不到.如 n 个单位合作治理污染,第 i 方单独治理的投资 y_i 和 n 方合作治理的投资 Y ,通常是已知的.为了度量第 i 方在合作中的“贡献”,还常要设法知道第 i 方不参加合作时其余 $n-1$ 方所需的投资 z_i . 特征函数应定义为合作的获利,即节约的投资,有 $v(i) = 0 (i=1, 2, \dots, n)$, $v(I) = \sum_{i=1}^n y_i - Y$, $v(I \setminus i) = \sum_{j \neq i} y_j - z_i$,显然除此之外还有许多 $v(s)$ 不知道,无法用 Shapley 值方法求解.

下面仍以本节开始提出的三人经商问题为例,介绍几种其他解决办法.

我们只知道全体合作的获利,记作 $v(I) = B$,及无 i 参加时其余 $n-1$ 方合作的获利,记作 $v(I \setminus i) = b_i (i=1, 2, \dots, n)$,且记 $\mathbf{b} = (b_1, b_2, \dots, b_n)$. 试确定各方对全体合作获利的分配,记作 $\mathbf{x} = (x_1, x_2, \dots, x_n)$. 在三人经商问题中^① $B = 11$, $\mathbf{b} = (4, 5, 7)$,求 $\mathbf{x} = (x_1, x_2, x_3)$.

1. 协商解

分配按以下两步进行. 先从 n 个 $n-1$ 方合作的获利得出各方分配的下限 $\underline{\mathbf{x}} = (\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n)$,即求解

$$\begin{cases} \sum_{i=1}^n x_i - \underline{x}_1 = b_1 \\ \dots\dots\dots \\ \sum_{i=1}^n x_i - \underline{x}_n = b_n \end{cases} \quad (9)$$

得到

^① 这里不需要知道每人单干的获利.

$$\underline{x}_i = \frac{1}{n-1} \sum_{i=1}^n b_i - b_i, \quad i=1, 2, \dots, n \quad (10)$$

再计算按下限 \underline{x} 分配后全体合作获利的剩余为 $B - \sum_{i=1}^n \underline{x}_i$, 它通常是较小的部分, 经协商将其平均分配, 于是最终的分配结果为

$$x_i = \underline{x}_i + \frac{1}{n} \left(B - \sum_{i=1}^n \underline{x}_i \right) = \frac{B}{n} + \frac{1}{n} \sum_{i=1}^n b_i - b_i \quad (11)$$

剩余 $B - \sum_{i=1}^n \underline{x}_i \geq 0$, 它等价于 $B \geq \frac{1}{n-1} \sum_{i=1}^n b_i$, 请读者考察这个假定的含义.

对三人经商问题, $\underline{x} = (4, 3, 1)$, $\underline{x} = (5, 4, 2)$.

2. 均衡解

设各方能够接受的现状点为 $d = (d_1, d_2, \dots, d_n)$, 可看作谈判时的威慑点, 在此基础上均衡地分配全体合作的获利 B . 根据 n 个数的和一定, 当它们相等时乘积最大的原理, 该模型为

$$\begin{aligned} & \max \prod_{i=1}^n (x_i - d_i) \\ \text{s. t. } & \sum_{i=1}^n x_i = B, \quad x_i \geq d_i \quad (i=1, 2, \dots, n) \end{aligned} \quad (12)$$

$$\text{得到} \quad x_i = d_i + \frac{1}{n} \left(B - \sum_{i=1}^n d_i \right) \quad (13)$$

$d=0$ 时, 相当于各方平均分配 B ; $d=\underline{x}$ 时, 均衡解等价于协商解.

3. 最小距离解

设存在一个各方理想的分配上限, 记作 $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$, 追求分配结果与这个上限的距离最小, 模型为

$$\begin{aligned} & \min \sum_{i=1}^n (x_i - \bar{x}_i)^2 \\ \text{s. t. } & \sum_{i=1}^n x_i = B, \quad x_i \leq \bar{x}_i \quad (i=1, 2, \dots, n) \end{aligned} \quad (14)$$

$$\text{得到} \quad x_i = \bar{x}_i - \frac{1}{n} \left(\sum_{i=1}^n \bar{x}_i - B \right) \quad (15)$$

i 方的理想上限若取为 $\bar{x}_i = B - b_i$, 看作 i 方对全体合作的“贡献”或 i 方的边际效益, 将其代入 (15) 式可得 $x_i = \frac{B}{n} + \frac{1}{n} \sum_{i=1}^n b_i - b_i$, 与 (11) 式相同, 即最小距离解等价于协商解. 对三人经商问题, $\bar{x} = (7, 6, 4)$, $\underline{x} = (5, 4, 2)$.

4. 满意解

i 方分配的满意度定义为 $u_i = \frac{x_i - d_i}{e_i - d_i}$, 其中 d_i 是现状点, e_i 是理想点. 为追求各方的满意度都高, 用最小最大模型

$$\max(\min_i u_i) \quad (16)$$

$$\text{s. t. } \sum_{i=1}^n x_i = B$$

得到

$$x_i = d_i + u^* (e_i - d_i), \quad u^* = \frac{B - \sum_{i=1}^n d_i}{\sum_{i=1}^n e_i - \sum_{i=1}^n d_i} \quad (17)$$

可以验证, 当 $d_i = \underline{x}_i, e_i = \bar{x}_i$ 时, 满意解等价于协商解. 当 $d_i = 0, e_i = \bar{x}_i$ 时, $x_i = \frac{\bar{x}_i}{\sum_{i=1}^n \bar{x}_i} B$, 即按照各方理想上限的比例进行分配.

5. Raiffa 解

Howard Raiffa 提出的解决办法按以下步骤进行:

1) 按照 n 个 $n-1$ 方合作的获利得到各方分配的下限, 即协商解中的 \underline{x} (见 (10) 式), 作为分配的基础;

2) 当 j 方加入 (原来无 j 的) $n-1$ 方合作时计算获利的增加, 即 j 方的边际效益, 是最小距离解中的上限 $\bar{x}_j = B - b_j$;

3) 按两步分配 \bar{x}_j : 先由 j 方和无 j 的 $n-1$ 方平分, 然后 $n-1$ 方再等分, 即

$$x_j = \frac{\bar{x}_j}{2}, \quad x_i = \underline{x}_i + \frac{\bar{x}_j}{2(n-1)}, \quad i = 1, 2, \dots, n, i \neq j \quad (18)$$

其中 $n-1$ 方是在 \underline{x} 的基础上分配;

4) j 取 $1, 2, \dots, n$, 重复第 3 步, 然后求和、平均, 得到最终分配为

$$x_i = \frac{n-1}{n} \underline{x}_i + \frac{1}{n} \left[\frac{\bar{x}_i}{2} + \frac{1}{2(n-1)} \sum_{j \neq i} \bar{x}_j \right], \quad i = 1, 2, \dots, n \quad (19)$$

将 \underline{x}, \bar{x} 代入, (19) 式又可表为

$$x_i = \frac{B}{n} + \frac{2n-3}{2(n-1)} \left[\frac{1}{n} \sum_{i=1}^n b_i - b_i \right], \quad i = 1, 2, \dots, n \quad (20)$$

对三人经商问题, $\underline{x} = (4, 3, 1), \bar{x} = (7, 6, 4), x = \left(4 \frac{2}{3}, 3 \frac{11}{12}, 2 \frac{5}{12} \right)$.

几种方法的比较 上面介绍的方法中, 协商解、均衡解、最小距离解和满意解比较简单、容易理解, 并且在许多情况下是等价的, 不妨并为一类. 这样, 连同

Shapley 值方法我们共讨论了 3 类方法:Shapley 值方法;协商解等;Raiffa 解. 下面结合一个较为极端的例子说明它们的特点.

例 有一资方(甲)和二劳方(乙、丙),当且仅当资方与至少一劳方合作时才获利 10 元,应如何分配该获利?

解 甲、乙、丙三方记作 1,2,3,

1) Shapley 值方法 特征函数定义为获利,则子集 $\{1,2\}, \{1,3\}, \{1,2,3\}$ 的特征函数为 10,其余均为 0,容易算出 Shapley 值,将其作为一种分配,即得 $x = (20/3, 5/3, 5/3)$.

2) 协商解等 由 $B = 10, b = (0, 10, 10)$ 得到 $\underline{x} = \bar{x} = (10, 0, 0)$, 于是 $x = (10, 0, 0)$.

3) Raiffa 解 将 $x = \bar{x} = (10, 0, 0)$ 代入(19)式,即得 $x = (25/3, 5/6, 5/6)$.

3 种方法得到的结果不同,协商解等显然对劳方不公平,Raiffa 解在一定程度上照顾了劳方的利益.

一般地,这 3 类方法有以下特点:

Shapley 值方法公正、合理,但是需要的信息太多, n 较大的实际问题难以提供. 协商解等计算简单,便于理解,但通常偏袒强者,可用于各方实力相差不大的情况. Raiffa 解考虑了分配的上下限,又吸取了 Shapley 的思想,在一定程度上保护弱者.

11.6 加权投票中权力的度量

在许多经济或政治机构中,为了保证每个参与者有平等的权力,在进行投票选举和表决提案等活动时,通常采取“一人一票”的方式,以显示投票和表决的公正性. 然而还有不少不宜采用按人头计票的情况,如在股份制公司的一些机构中,每位股东投票和表决权的大小常由他们占有的股份多少决定. 又如一些国家、地区的议会、政府,甚至总统的产生,是由这些国家、地区所属的州、县等各个区域推出的代表投票决定的,而这些代表投票的权重又取决于他们代表的那个区域的人口,这种看似公平的办法在一些情况下也会出现矛盾,引起质疑.

美国的总统选举是一个典型案例. 根据美国宪法,总统选举不是全民普选,而是实行选举人制度. 总统候选人获得全国 50 个州和华盛顿特区共 538 张选举人票数的一半以上即可当选,各州拥有的选举人票数与该州在国会拥有的参、众议员人数相等. 参议员每州 2 位,众议员人数则根据各州人口比例来确定,而各州人口的悬殊使得各州选举人票的数量相差很大,如人口众多的加利福尼亚州的选举人票多达 55 张,人口较少的阿拉斯加州只有 3 张.

各州的选举人如何投票呢?原来,总统候选人先在各州内实行普选,获得相对多数选票的候选人将得到该州的全部选举人票,这就是在 48 个州和华盛顿特区实行的“胜者全得”原则.这样,在加利福尼亚州以微弱多数普选获胜的总统候选人可得到全部 55 张选举人票.如果有几个人口多的州出现这种情况,总统选举的结果就可能违反全国多数人的意愿,即在各州累计得票最多的总统候选人在选举人投票中反而不能获胜.美国历史上曾多次发生这种情况,在 2000 年布什与戈尔进行的竞选中,戈尔最终败给布什就是一例.

选举人制度在美国已经历了 200 多年的发展与演变.近年来,虽然要求改革的呼声不断,但由于多种因素的制约,改革始终无法进行.

为了叙述的方便,本节考察的对象可以解释为:由若干区域(如省、县、镇等)组成的某个机构中,每区都有代表任职,其数量按照各区人口的比例分配,在机构进行投票选举和表决时,每区的全体代表投相同的票.这种情况等价于每区各派一位代表(也称投票人),按照他们所代表的各区人口比例赋予他们以投票的权重,这就是加权投票的含义.股份制公司中股东以所占有的股份为权重来投票也属于这种情况.

如何度量每位代表的投票对最终结果的影响力(不妨称为投票人的权力),本节根据对策论的方法引入一个模型,介绍两种合理的数量指标,并且通过实例给出它们的应用,最后研究如何调整投票人的权重,使他们的权力大致与他们所代表的各区人口成比例^[51].

加权投票与获胜联盟 本节研究的对象称为加权投票,让我们先看一个虚拟的例子.

例 1 假设一个县有 5 个区,记作 A, B, C, D, E, 人口(单位:千人)分别为 60, 20, 10, 5, 5, 每区出一位代表,按照人口比例分配各区代表的投票权重为 12, 4, 2, 1, 1, 总权重 20. 按照简单多数规则决定投票结果, 即当投赞成票代表的权重之和超过 10(总权重的一半)时,投票结果即为赞成(以下表述为决议通过).显然, A 区的代表是一位独裁者,他可以决定投票结果,其他各区代表都成了摆设.

为了改变这种情况,将 A 区分成人口相等的 3 个子区 A₁, A₂, A₃, 于是一共 7 个区,按人口比例分配,每区一位代表的投票权重为 4, 4, 4, 4, 2, 1, 1. 如果仍按简单多数规则决定投票结果,那么可以决定投票结果的区域集合除了 [A₁, A₂, A₃] 外,还有 [A₁, A₂, B], [A₁, A₃, C, D], [A₁, B, C, E], [A₁, A₃, B, D] 等等.

加权投票可描述如下:一个系统由 n 个投票人的集合 $N = \{A, B, C, \dots\}$ 及他们的投票权重(均为正数) w_1, w_2, \dots, w_n 构成,投票规则规定,当投赞成票的投票人权重之和达到或超过事先给定的正数 q 时,决议通过.具有这种构造及规则的系统称为加权投票系统,记作 S , 用 $S = [q; w_1, w_2, \dots, w_n]$ 表示,其中 q 称为定

额 (Quota), 若记 $w = w_1 + w_2 + \cdots + w_n$, 一般假定 $w/2 < q \leq w$. 在简单多数规则并且权重取整数的情况, q 为大于 $w/2$ 的最小整数.

像例 1 那样, 若 n 个区的人口比例为 $p = (p_1, p_2, \cdots, p_n)$, 按照这个比例直接赋予各位代表的投票权重, 称为直接比例加权投票系统. 例 1 中的 2 个系统分别记作 $S^{(1)} = [11; 12, 4, 2, 1, 1]$ 和 $S^{(2)} = [11; 4, 4, 4, 4, 2, 1, 1]$.

权重之和达到或超过定额 q 的投票人子集称为获胜联盟 (Winning Coalition). 获胜联盟如果没有它的一个真子集也是获胜联盟, 那么称它为极小获胜联盟. 在系统 $S^{(2)}$ 中, $[A1, A2, B]$, $[A1, A3, C, D]$ 都是极小获胜联盟, 而 $[A1, A3, B, D]$ 是获胜联盟, 却不是极小的. 所有获胜联盟的集合称获胜联盟集, 记作 W , 极小获胜联盟集记作 W_m . 一个加权投票系统 $S = [q; w_1, w_2, \cdots, w_n]$ 也可以用投票人集合 N 和获胜联盟集 W 表示, 记作 $G = (N, W)$, 在对策论中是一个合作对策.

极小获胜联盟在加权投票系统中起着重要作用. 再看一个例子.

例 2 某系的一个委员会由主任、教授、学生各一人组成, 依次记作 A, B, C , 投票权重为 w_1, w_2, w_3 . 考察以下几种加权投票系统:

1. $S^{(1)} = [3; 3, 1, 1]$. 只有一个极小获胜联盟, $W_m = ([A])$, 主任是独裁者, 教授、学生的投票对结果没有任何影响, 可以说他们没有权力, 这样的投票人称为傀儡 (Dummy).

2. $S^{(2)} = [2; 1, 1, 1]$. 任何 2 人组成极小获胜联盟, $W_m = ([AB], [AC], [BC])$, 3 人的权力相同.

3. $S^{(3)} = [4; 2, 2, 1]$. 只有一个极小获胜联盟, $W_m = ([AB])$, 学生不在其中, 是傀儡, 而主任和教授的权力相同.

4. $S^{(4)} = [3; 2, 1, 1]$. 有 2 个极小获胜联盟, $W_m = ([AB], [AC])$, 主任虽不是独裁者, 但每一个极小获胜联盟都有他, 表明主任有否决权, 而教授和学生的权力相同.

可以看出, 给定一个加权投票系统 S , 能够写出极小获胜联盟集 W_m , 由 W_m 也容易写出获胜联盟集 W . 实际上, 我们常常用获胜规则描述一个投票过程, 如某系务会由一位主任和两位教师组成, 投票结果由主任和至少一位教师决定, 主任有否决权. 由此首先写出的是 $N = \{A, B, C\}$, $W_m = ([AB], [AC])$, $W = ([AB], [AC], [ABC])$, 这里 A 是主任, B, C 是教师. 而根据这样的 N, W_m 或 W 可以得到不同的加权投票系统, 如上面的 $S^{(4)} = [3; 2, 1, 1], [5; 3, 2, 2]$ 等. 加权投票系统 S 与 N, W_m 或 W 关系的讨论请参看习题 13.

分析例 2 的加权投票系统及其极小获胜联盟可以看出, 每个投票人的投票对系统投票结果的影响, 并不直接依赖于他的投票权重, 如 $S^{(1)}$ 中主任的权重不过是教授或学生的 3 倍, 却能完全操纵结果, $S^{(3)}$ 中学生的权重是主任或教授的

一半,但对结果毫无影响力。

显然,每个投票人对投票结果的影响力才是他在系统中权力或地位最重要的度量.若用 $k = (k_1, k_2, k_3)$ 描述例 2 加权投票系统中 3 位投票人的(相对)权力,作初步分析就可看出,在 $S^{(1)}, S^{(2)}, S^{(3)}$ 中投票人要么有权,且权力相同,可以用 1 表示,要么无权,用 0 表示,于是对于 $S^{(1)}$ 有 $k^{(1)} = (1, 0, 0)$, $S^{(2)}$ 有 $k^{(2)} = (1, 1, 1)$, $S^{(3)}$ 有 $k^{(3)} = (1, 1, 0)$,而对于 $S^{(4)}$, $k^{(4)}$ 如何呢? 因为教授和学生的权力相同,可以令 $k_2 = k_3 = 1$,显然 k_1 应该大一些,等于多少合适呢?

实际上在加权投票系统中,除了很简单的情况外,找到公平、合理的度量投票人权力的办法并不那么容易.下面介绍两种度量投票人权力的数量指标——Shapley 指标和 Banzhaf 指标,统称权力指标(Power Index)。

Shapley 权力指标 对于前面讨论的加权投票系统 $S = [q; w_1, w_2, \dots, w_n]$ 或其对策论的表述 $G = (N, W)$, 必有空集 $\emptyset \notin W, N \in W$, 且若 $R, T \in N, R \subset T, R \in W$, 则 $T \in W$. 下面先列出作为度量投票人权力的数量指标应该具有的性质:

1. 对于每个投票人 i 有一个非负实数 k_i 作为他的权力指标;
2. 当且仅当 $i \notin W_m$ 即 i 是傀儡时, $k_i = 0$;
3. 若权重 $w_i > w_j$, 则 $k_i \geq k_j$;
4. 若投票人 i 和 j 在 W 中“对称”(通过下面例子解释), 则 $k_i = k_j$;
5. $\sum_{i \in N} k_i = 1$ (这种归一化不是必须的)。

满足这些性质的数量指标并不唯一, Shapley 权力指标(或称 Shapley-Shubik 权力指标)是常用的一个. 下面先通过例 2 由主任 A、教授 B、学生 C 构成的加权投票系统 $S^{(4)} = [3; 2, 1, 1]$, 看看这个指标是如何得到的。

首先写出 3 位投票人的全排列 ABC, ACB, BAC, BCA, CAB, CBA. 对第 1 个排列 ABC, 由左向右依次检查: 从 A 增至 AB 时, [AB] 变为获胜联盟, B 下划以横线; 对第 2 个排列 ACB, 从 A 增至 AC 时, [AC] 变为获胜联盟, C 下划以横线; …… 对第 6 个排列 CBA, 从 CB 增至 CBA 时, [CBA] 变为获胜联盟, A 下划以横线. 这样得到

ABC ACB BAC BCA CAB CBA

A 下有 4 条横线, B, C 下各有 1 条横线, 于是系统 $S^{(4)}$ 的 Shapley 指标是 $(4, 1, 1)$, 可归一化为 $(4/6, 1/6, 1/6)$ 。

一般地, 对于有 n 个投票人的加权投票系统, 写出投票人的共 $n!$ 个全排列, 然后对每一个排列由左向右依次检查, 若某位投票人加入时该集合变成获胜联盟, 称该投票人为决定者(Pivot). 将每位投票人在所有排列中的成为决定者的次数 ξ 除以排列数 $n!$, 定义为他们系统中的 Shapley 权力指标, 记作 $\varphi = \xi/n!, \varphi = (\varphi_1, \varphi_2, \dots, \varphi_n)$. 再看一个例子。

例 3 某股份有限公司的 4 个股东分别持有 40%、30%、20%、10% 的股份, 公司的决策必须经持有半数以上股份的股东的同意才可通过, 求这 4 个股东在公司决策中的 Shapley 指标.

4 个股东依次记作 A, B, C, D, 上述问题可归结为加权投票系统 $S = [6; 4, 3, 2, 1]$, 为计算系统的 Shapley 指标, 写出 4 个股东的 $4! = 24$ 个全排列, 对每一个排列找出决定者, 下划横线, 如下:

<u>AB</u> CD	<u>AB</u> DC	<u>AC</u> BD	<u>AC</u> DB	<u>AD</u> BC	<u>AD</u> CB
<u>BAC</u> D	<u>BAD</u> C	<u>BCA</u> D	<u>BCA</u> D	<u>BD</u> AC	<u>BD</u> CA
<u>CAB</u> D	<u>CAD</u> B	<u>CBA</u> D	<u>CBA</u> D	<u>CD</u> AB	<u>CD</u> BA
<u>DABC</u>	<u>DACB</u>	<u>DBAC</u>	<u>DBCA</u>	<u>DCAB</u>	<u>DCBA</u>

统计出 A, B, C, D 成为决定者的次数 $\xi = (10, 6, 6, 2)$, Shapley 指标为 $\varphi = (5/12, 3/12, 3/12, 1/12)$.

可以通过例 2 的 $S^{(4)}$ 和例 3 检查 Shapley 指标满足上面的 5 条性质, 其中 1, 2, 3, 5 是显然的, 对于性质 4, 考察投票人在获胜联盟集 W 或极小获胜联盟集 W_m 中的对称性. 例 2 的 $S^{(4)}$ 中 $W = ([AB], [AC], [ABC])$, B 和 C 对称, $\varphi_2 = \varphi_3$; 例 3 中 $W_m = ([AB], [AC], [BCD])$, B 和 C 对称, $\varphi_2 = \varphi_3$. 而如果事先判断出 B 和 C 对称, 则上面 24 个排列中可以只保留 B 在 C 之前的那些, 减至如下 12 个:

<u>AB</u> CD	<u>AB</u> DC	<u>AD</u> BC	<u>BAC</u> D	<u>BAD</u> C	<u>BCA</u> D
<u>BCA</u> D	<u>BD</u> AC	<u>BD</u> CA	<u>DABC</u>	<u>DBAC</u>	<u>DBCA</u>

统计出 A, B (和 C), D 成为决定者的次数 $\xi = (5, 6, 1)$, 其中 6 为 B 和 C 分享, 所以 Shapley 指标与上相同.

Banzhaf 权力指标 这是另一个常用的权力指标, 让我们仍然通过例 2 的加权投票系统 $S^{(4)} = [3; 2, 1, 1]$, 看看它是如何得到的.

首先写出 $S^{(4)}$ 的获胜联盟集 $W = ([AB], [AC], [ABC])$, 对于 W 中的每个获胜联盟, 检查每位投票人是否决定者, 即这个联盟是否由于他的加入才获胜的, 若是, 下划以横线, 于是得到

<u>AB</u>	<u>AC</u>	<u>ABC</u>
-----------	-----------	------------

A 下有 3 条横线, B, C 下各有 1 条横线, 于是系统 $S^{(4)} = [3; 2, 1, 1]$ 的 Banzhaf 指标是 (3, 1, 1), 可归一化为 $(3/5, 1/5, 1/5)$. 注意到它与 $S^{(4)}$ 的 Shapley 指标 $(4/6, 1/6, 1/6)$ 不同, 并且这两个指标都与投票人的权重 $(2, 1, 1)$ 不一样.

两个指标虽然都是由投票人成为决定者的次数决定, 但是 Shapley 指标是在投票人的排列中检查, 而 Banzhaf 指标是在投票人的组合 (获胜联盟) 中检查.

一般地, 对于有 n 个投票人的加权投票系统, 写出它的获胜联盟集 W , 然后对于每一个联盟检查每位投票人是否决定者, 将每位投票人在 W 中成为决定者

的次数 η 归一化, 定义为他们在系统中的 Banzhaf^① 权力指标, 记作 $\beta = (\beta_1, \beta_2, \dots, \beta_n)$.

让我们计算例 3 加权投票系统 $S = [6; 4, 3, 2, 1]$ 的 Banzhaf 指标.

写出 S 的获胜联盟集 $W = ([AB], [AC], [ABC], [ABD], [ACD], [BCD], [ABCD])$, 在 W 中的每一个检查每位投票人, 是决定者的下划以横线, 得到

$\underline{AB} \quad \underline{AC} \quad \underline{ABC} \quad \underline{ABD} \quad \underline{ACD} \quad \underline{BCD} \quad \underline{ABCD}$

统计出 A, B, C, D 成为决定者的次数 $\eta = (5, 3, 3, 1)$, 归一化后 Banzhaf 指标为 $\beta = (5/12, 3/12, 3/12, 1/12)$, 与前面得到的 Shapley 指标 φ 相同. 从这里也可以检查 Banzhaf 指标满足上面的 5 条性质.

加权投票与权力指标的应用 给出两个应用案例.

例 4 在某些国际拳击比赛中, 设两个 5 人裁判组, 每位裁判一票, 先由第 1 组判, 若第 1 组以 5:0 或 4:1 判选手甲胜, 则甲胜; 若第 1 组以 3:2 判选手甲胜, 则由第 2 组再判, 除非第 2 组以 0:5 或 1:4 判选手甲负, 其他情况最终都判甲胜. 试将以上裁判规则用加权投票系统表示, 并计算 Shapley 指标和 Banzhaf 指标.

设两组 10 人同时裁判, 组成集合 $N = \{A, A, A, A, A, B, B, B, B, B\}$, 其中 A 是第 1 组, B 是第 2 组. 根据裁判规则这一系统的极小获胜联盟为: 第 1 组的 4 位裁判; 第 1 组的 3 位裁判加上第 2 组的 2 位裁判; 第 1 组的 2 位裁判加上第 2 组的 4 位裁判, 记作 $W_m = ([4A], [3A2B], [2A4B])$. 可以合理地加权投票系统取为 $S = [q; a, a, a, a, a, 1, 1, 1, 1, 1]$, 其中 a, q 为大于 1 的整数, 且 a, q 应满足 $4a \geq q, 3a + 2 \geq q, 2a + 4 \geq q, 3a + 1 < q, \dots$. 容易看出, 只需简单地取 $a = 2, q = 8$ 即可. 于是上面的两组裁判办法等价于: 令第 1 组 5 位裁判的权重各为 2, 第 2 组 5 位裁判的权重各为 1, 一起裁判, 按简单多数规则执行.

按照这样的规则裁判 A 的权力刚好是裁判 B 的 2 倍吗? 下面计算系统的 Shapley 指标和 Banzhaf 指标时, 将 5 个 A 看作可分辨的, 但仍记作 A, 5 个 B 也是如此.

对于 Shapley 指标, 5 个 A 和 5 个 B 的全排列共 $10!$ 个. 为了得到某个 B 在所有排列中成为决定者的次数, 只需计算 $(3A1B)\underline{B}$ 和 $(2A3B)\underline{B}$ 的排列数, 其中下划横线的为决定者 B, 横线前括号内应取全排列, 横线后字母省略, 仍取全排列. 它们的排列数是 $C_3^5 C_1^4 4! 5!$ 和 $C_2^5 C_3^4 4! 5!$. 于是某个 B 的 Shapley 指标为

① J. F. Banzhaf III (1940—) 获 MIT 电机工程学士学位 (1962) 和哥伦比亚大学法学博士学位 (1965), 现为乔治华盛顿大学法学教授. 早在攻读博士学位期间就发现了直接比例加权投票系统的缺点, 并研究了已存在的 Shapley 权力指标. 于 1965 年发表文章提出后来以他的名字命名的 Banzhaf 指标, 以 1964 年新泽西州参议院选举为例说明其应用. 稍后, 芝加哥大学著名社会学家 J. S. Coleman 在另外的背景下独立地提出同样的指标, 有时该指标也称 Banzhaf-Coleman 指标.

$\frac{4!}{10!} \frac{5!}{10!} (C_5^3 C_4^1 + C_5^2 C_4^3) = \frac{4}{63} = 0.0635$, 而某个 A 的 Shapley 指标就很容易得到, 为 $\frac{1}{5} \left(1 - \frac{4}{63} \times 5 \right) = 0.1365$. 系统的 Shapley 指标为 $\varphi = (0.1365, 0.1365, 0.1365, 0.1365, 0.1365, 0.0635, 0.0635, 0.0635, 0.0635, 0.0635)$.

对于 Banzhaf 指标, 需要写出 A, B 可能成为决定者的那些获胜联盟类型和个数, 由此计算 A, B 为决定者(下划横线)的次数(表 1).

表 1 Banzhaf 指标中计算决定者次数

获胜联盟类型	<u>4A</u>	<u>4A1B</u>	<u>3A2B</u>	<u>3A3B</u>	<u>2A4B</u>	<u>2A5B</u>
联盟个数	5	25	100	100	50	10
A 为决定者次数	20	100	300	300	100	20
B 为决定者次数	0	0	200	0	200	0

在获胜联盟中 A 为决定者的次数与 B 为决定者的次数之比 840:400, 归一化后的 Banzhaf 指标为 $\beta = (0.1355, 0.1355, 0.1355, 0.1355, 0.1355, 0.0645, 0.0645, 0.0645, 0.0645, 0.0645)$.

两个权力指标相差无几, 也与权重基本一致.

从计算过程可以看出, 由于 Shapley 指标自然就是归一的, 可以在 A, B 中选一个较容易的计算(读者不妨试算 A, 看看有多繁), Banzhaf 指标没有这种归一性, A, B 都要计算.

例 5 “团结就是力量”吗? 在加权投票系统中通过结盟能加强权力吗? 下面虚拟一个议会的例子来分析. 为了叙述的简便我们采用民主党、共和党这样的词汇.

在 40 位议员组成的议会中民主党(记作 M)占 11 席, 共和党(记作 G)占 14 席, 其余 15 席属于独立人士(记作 D). 投票采取简单多数规则, 21 票通过. 用 Shapley 指标来度量议员的权力.

最初是每位议员都独立投票, 系统记作 $S^{(1)} = [21; 1, 1, \dots, 1]$, 其中有 40 个 1, 不必计算就知道每位议员(投票人)的 Shapley 指标相等($\varphi_i = 1/40, i = 1, \dots, 40$), 民主党议员的 Shapley 指标之和为 $\varphi_M = 11/40 = 0.275$, 共和党议员为 $\varphi_G = 14/40 = 0.350$, 独立人士为 $\varphi_D = 15/40 = 0.375$.

如果民主党 11 位议员结盟, 推出 1 位代表作为投票人, 系统记作 $S^{(2)} = [21; 11, 1, \dots, 1]$, 除民主党 M 的权重为 11 外, 其中 29 人均均为 1. 对于这种特殊的系统, 为计算 M 的 Shapley 指标 φ_M , 不必写出所有 30! 个排列, 而只考察 M 在 30 个投票人中的位置, 如图 1 的直线. 当 M 在第 11 到第 21 的任一位置加入排

列时, M 成为决定者, $21 - 11 + 1 = 11$ 个位置在总共 30 个位置中的比例即为 $\varphi_M = 11/30 = 0.367$. 共和党议员的 Shapley 指标之和由余下的 $19/30$ 中按照 14:15 与独立人士分享, 于是 $\varphi_G = (19/30) \times (14/29) = 0.306$, 独立人士 $\varphi_D = 0.327$. 显然, 民主党议员通过结盟增加了权力, 共和党 and 独立人士的权力则有所减少.

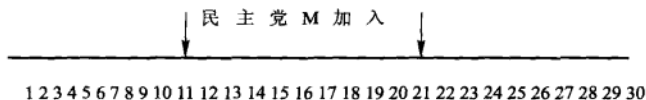


图 1 计算民主党投票人 M 的 Shapley 指标的图示

如果民主党议员的结盟诱使共和党 14 位议员也结盟, 推出 1 位代表作为投票人, 系统记作 $S^{(3)} = [21; 11, 14, 1, \dots, 1]$, 其中有 15 个 1, 共 17 位投票人. 将图 1 用直线表示民主党投票人 M 的加入, 推广到在平面上用 x 轴和 y 轴分别表示民主党投票人 M 和共和党投票人 G 的加入, 如图 2. 图上格子点的坐标 (i, j) 表示 M 在第 i 位置加入, G 在第 j 位置加入, 因为 $i \neq j$, 所以共 $17 \times 17 - 17 = 272$ 个点, 点 (i, j) 与它左下方的方格相对应, 用方格表述更为清晰.

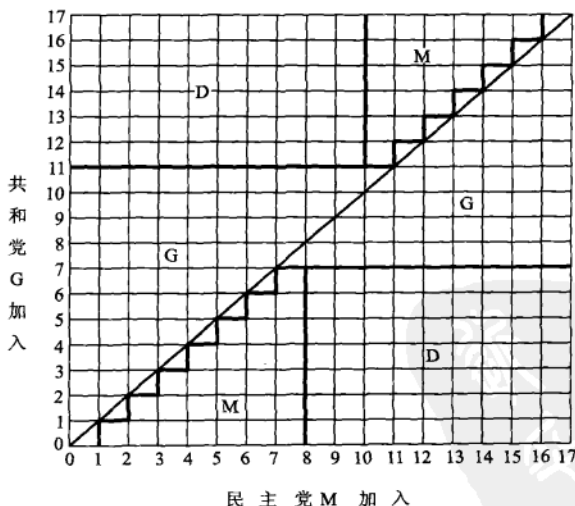


图 2 计算民主党 M 和共和党 G 的 Shapley 指标的图示

图中对角线以下的方格表示 G 在 M 之前加入, 注意到 M 和 G 的权重 $w_M = 11, w_G = 14$ 和定额 $q = 21$, 可知当 $j \leq 7, i \leq 8$ 时 M 成为决定者, $j \leq 7, i > 8$ 时独立人士 D 为决定者, 若 $j > 7$ 则 G 为决定者. 对角线以上的方格表示 M 在 G 之前加

入,做类似的分析得到 M, G, D 为决定者的区域,如图中字母所示. 数一下各区域的方格分别为 49, 100, 123, 于是 $\varphi_M = 49/272 = 0.180$, $\varphi_G = 100/272 = 0.368$, $\varphi_D = 0.452$.

顺便指出,用图 2 的方式进行计算的一个好处是,当除两个主要投票人(M 和 G)之外第 3 方(D)的人数很大时,可以用各个区域面积的比例近似得到 Shapley 指标(参看习题 18)

将上面几种情况下得到的民主党 M 和共和党 G 的 Shapley 指标用表 2 列出,其中共和党结盟、民主党不结盟的数据是按照类似图 1 的方法得到的.

表 2 民主党和共和党不同结盟情况下的 Shapley 指标

	共和党 G 不结盟	共和党 G 结盟
民主党 M 不结盟	$\varphi_M = 0.275$ $\varphi_G = 0.350$	$\varphi_M = 0.204$ $\varphi_G = 0.519$
民主党 M 结盟	$\varphi_M = 0.367$ $\varphi_G = 0.306$	$\varphi_M = 0.180$ $\varphi_G = 0.368$

可以看出,不论民主党是否结盟,共和党议员联合起来总比单干要好. 而共和党一旦结盟,民主党议员不联合更好. 不过从民主党的角度看,应该尽量保持原来大家都是单干的局面,不要率先结盟,贪图一时的胜利,反而会诱使共和党也结盟,结果会败得很惨——远不如全都单干的情况.

有意思的是,从独立人士的角度看,虽然若只有民主党或只有共和党结盟自己都有损失,但如果两个党均结盟,反而可得渔翁之利.

两种权力指标的公理化 从前面计算两种权力指标的过程看,二者都以投票人成为决定者的次数来度量权力大小,其区别是:Shapley 指标是检查投票人加入已有的联盟(排列)时能否使联盟获胜,这样的排列共 $n!$ 个,在每个排列中有且只有一个决定者;Banzhaf 指标是检查投票人在获胜联盟(组合)中是否起决定作用,在每个组合中可以有不止一个(也可以没有)决定者.

前面说过,加权投票系统的这两种权力指标都满足应该具有的 5 条性质,在应用中都有其合理性,甚至不少加权投票系统的两种指标的结果相同,但是它们毕竟是度量投票人对结果影响力的两种不同的指标,追求严密逻辑的数学家们不会停留于此,他们要在一组明显的、不需事先证实的公理下,证明研究对象的存在与惟一性,即公理化. Shapley 指标于 1954 年提出^[68]、1975 年给以公理化^[24]证明;Banzhaf 指标于 1965 年提出^[7]、1979 年给以公理化^[25]证明. 需要从公理化的角度理解两种指标区别的读者,可参阅上面指出的文献,下面只用对策论的术语将公理化得到的这两个指标表述如下^[45].

设有投票人集合 $I = \{1, 2, \dots, n\}$ 及加权投票系统 $S = [q; w_1, w_2, \dots, w_n]$, I 的任一子集 s 对应一个实值函数 v , 满足单调性: 对于 $T \supseteq s$, 有 $v(T) \geq v(s)$ 及

$v(\emptyset) = 0$. 当 s 为获胜联盟时, $v(s) = 1$, 否则 $v(s) = 0$. Shapley 指标可表为

$$Sh_i(v) = \sum_{i \in s} \frac{(|s| - 1)!(n - |s|)!}{n!} [v(s) - v(s \setminus i)], i = 1, 2, \dots, n \quad (1)$$

其中 $|s|$ 是 s 中投票人的数量, $s \setminus i$ 是 s 去掉 i 后的集合. Banzhaf 指标可表为

$$Bz_i(v) = \frac{1}{2^{n-1}} \sum_{i \in s} [v(s) - v(s \setminus i)], i = 1, 2, \dots, n \quad (2)$$

可以看出, $v(s) - v(s \setminus i) = 1$ 表示投票人 i 在 s 中是决定者, 所以对 $v(s) - v(s \setminus i)$ 求和相当于计算 i 成为决定者的次数. 虽然 (1), (2) 都是对所有包含 i 的共 2^{n-1} 个子集 (组合) s 求和 (与前面所述只对包含 i 的获胜联盟求和一样, 因为对非获胜联盟 s 有 $v(s) - v(s \setminus i) = 0$), 但是 (1) 中的系数是按排列计算的, 其分子的 $(|s| - 1)!$ 和 $(n - |s|)!$ 分别表示一个排列中排在 i 前面的和后面的投票人的全排列数. 可以注意到, (1) 式与 11.5 节效益的合理分配中 (7), (8) 式是一样的, 不过这里的函数 v 只取值 0 和 1.

下面用 (1), (2) 式计算一遍例 2 的 $S^{(4)} = [3; 2, 1, 1]$ 和例 3 的 $S = [6; 4, 3, 2, 1]$ 两个加权投票系统的 Banzhaf 指标与 Shapley 指标, 计算过程见表 3 和表 4. 为简单起见, 表的第 2 行只列出那些使 $v(s) - v(s \setminus i) = 1$ 的包含 i 的子集.

表 3 $S^{(4)} = [3; 2, 1, 1]$ Banzhaf 指标 Bz 和 Shapley 指标 Sh 的计算过程

	$i = A$			$i = B$	$i = C$
	[AB]	[AC]	[ABC]	[AB]	[AC]
$1/2^2$	1/4	1/4	1/4	1/4	1/4
Bz	3/4			1/4	1/4
$ s $	2	2	3	2	2
$(s - 1)!(3 - s)!/3!$	1/6	1/6	2/6	1/6	1/6
Sh	4/6			1/6	1/6

表 4 $S = [6; 4, 3, 2, 1]$ Banzhaf 指标 Bz 和 Shapley 指标 Sh 的计算过程

	$i = A$					$i = B$			$i = C$			$i = D$
	[AB]	[AC]	[ABC]	[ABD]	[ACD]	[AB]	[ABD]	[BCD]	[AC]	[ACD]	[BCD]	[BCD]
$1/2^3$	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8
Bz	5/8					3/8			3/8			1/8
$ s $	2	2	3	3	3	2	3	3	2	3	3	3
$(s - 1)! \cdot (4 - s)!/4!$	2/24	2/24	2/24	2/24	2/24	2/24	2/24	2/24	2/24	2/24	2/24	2/24
Sh	5/12					3/12			3/12			1/12

计算结果与前面用统计投票人在组合或排列中成为决定者次数所得到的相同,只是需要指出,(2)式给出的称为绝对 Banzhaf 指标,是投票人在 W 中成为决定者的次数 η 除以 2^{n-1} 得到的,为了与归一化的 Banzhaf 指标 β 相区别,将绝对 Banzhaf 指标记作 β' .

有些情况下绝对 Banzhaf 指标 β' 比归一化的 Banzhaf 指标 β 更能反映投票人权力的真实性,我们用 $S^{(1)} = [2; 1, 1, 1]$ 和 $S^{(2)} = [3; 1, 1, 1]$ 来说明. 显然它们的 β 都是 $(1/3, 1/3, 1/3)$, 但 $S^{(1)}$ 是多数规则,由获胜联盟集可得

$$\underline{AB} \quad \underline{AC} \quad \underline{BC}$$

于是 $\eta = (2, 2, 2)$, $\beta' = (1/2, 1/2, 1/2)$, 而 $S^{(2)}$ 是一致规则,由获胜联盟集可得 \underline{ABC} , 于是 $\eta = (1, 1, 1)$, $\beta' = (1/4, 1/4, 1/4)$. $S^{(1)}$ 的 β' 是 $S^{(2)}$ 的一倍,这是合理的,因为 $S^{(1)}$ 中一个投票人只需说服另外一个人站在他一边,而 $S^{(2)}$ 中一个投票人需要同时说服两个人.

绝对 Banzhaf 指标 β' 的又一个用途是对于常见的 n 人投票系统 $S = [q; 1, 1, \dots, 1]$, 其中有 n 个 1, q 取简单多数,可以证明(习题 15),当 n 很大时, β' 趋于 $1/\sqrt{n}$, 而不是直观想象的 $1/n$.

两种权力指标的概率解释 其实,加权投票系统中投票人对结果的影响力,可以用概率的语言描述为投票人能左右结果的概率,Shapley 指标和 Banzhaf 指标都可以在某种假设下解释为这种概率. 我们再一次用例 2 中由投票人 A, B, C 组成的系统 $S^{(4)} = [3; 2, 1, 1]$ 给以说明.

Banzhaf 指标的概率解释比较直接. 考察“投票人 A 能左右结果”这一事件(记作 R_A), 因为当 B, C 至少一人投赞成票时, A 能左右投票结果,所以 $R_A = \overline{B}\overline{C} + \overline{B}C + B\overline{C}$, 其中 B 表示 B 投赞成票, \overline{B} 表示 B 投反对票. 在 B, C 均以 $1/2$ 的概率独立投赞成或反对票时, R_A 的概率 $P(R_A) = 3/4$. 类似地, $R_B = A\overline{C}$, $R_C = A\overline{B}$, $P(R_B) = P(R_C) = 1/4$. 与前面用(2)式得到的绝对 Banzhaf 指标 β' 一样. 一般地, β' 可解释为在各位投票人独立地、以 $1/2$ 的概率投赞成或反对票(事先不知道投票人对决议的态度, 概率 $1/2$ 是一个合理的假设)的条件下, 每位投票人能左右结果的概率. 注意这些概率之和不一定等于 1.

Shapley 指标的概率解释就不那么明显. 设每位投票人独立地投赞成票的概率均为 p , 投反对票的概率 $q = 1 - p$, 则 $P(R_A) = 2pq + p^2 = 2p - p^2$, $P(R_B) = P(R_C) = pq = p - p^2$, 如果 p 在 $[0, 1]$ 区间均匀分布, 可以算出 A, B, C 能左右结果的概率为 $\int_0^1 P(R_A) dp = 2/3$, $\int_0^1 P(R_B) dp = \int_0^1 P(R_C) dp = 1/6$. 与前面在全排列中计算决定者次数及用(1)式得到的 Shapley 指标 φ 一样. 一般地, φ 可解释为在各位投票人独立且同概率分布($[0, 1]$ 均匀分布)地投赞成票的条件下, 每

位投票人能左右结果的概率. 注意这些概率之和一定等于 1.

两种权力指标的概率解释能帮助我们进一步理解 φ 和 β' 的进一步理解. 一方面, 排列、组合、决定者等手段和定义不是必须的, 这些指标可以在很自然的假设下用概率方法得到. 另一方面, 提供了一种区分 φ 和 β' 的途径. 如果投票人的行为是完全独立、随意 (1/2 的概率) 地进行, 那么适于用 β' ; 如果投票人对决议的态度有共同性 (持同样的 $[0, 1]$ 均匀分布的概率 p), 用 φ 是合适的.

调整加权投票系统 在加权投票系统中我们已经多次看到各位投票人的权重与他们的两种权力指标不相符的情况. 如果各位投票人的权重是直接按照各区域的人口比例赋予的, 这种情况就表明投票人对投票结果的权力与他所代表的人口比例出现失调. 如在例 1 中将 C, D, E 这 3 个区合并为 F, 则 A, B, F 的人口 (单位: 千人) 为 60, 20, 20, 人口比例为 $p = (3, 1, 1)$, 若按人口比例赋予 3 个区投票人的权重, 并用简单多数规则 (定额 $q = 3$), 得到的系统与例 2 的 $S^{(1)} = [3; 3, 1, 1]$ 相同, 其 Banzhaf 指标 $\beta = (1, 0, 0)$ 与人口比例相差甚远 (因为考虑的是比例, 用 β 与 β' 是等价的).

要使 Banzhaf 指标 β 与 p 一致, 只需将 $[3; 3, 1, 1]$ 调整为 $[3; 2, 1, 1]$ 即可, 因为前面已得到它的 $\beta = (3/5, 1/5, 1/5)$. 这就是说, 虽然 A 区的人口是 B 区的 3 倍, 但是赋予 A 区投票人的权重只应是 B 区的 2 倍, 在简单多数规则下可使投票人的权力正好与他们的人口比例相同 (在归一化的意义下).

当然, 这样的加权投票系统不是惟一的. 比如, 若仍按人口比例赋予权重, 而将定额改为 $q = 4$, 则系统 $[4; 3, 1, 1]$ 仍有 $\beta = (3/5, 1/5, 1/5)$. 其实, 还可以写出很多个 β 相同的系统, 如 $[17; 9, 8, 8]$, $[51; 50, 49, 1]$, $[51; 48, 26, 26]$ 等, 虽然其权重和定额各不相同, 系统内部权重有的还相差很大, 但是它们有相同的极小获胜联盟集 $W_m = ([AB], [AC])$, 而正是极小获胜联盟的结构确定了 Banzhaf 指标.

若 n 个区的人口比例为 $p = (p_1, p_2, \dots, p_n)$, 调整加权投票系统的目的是, 寻求一组权重和定额, 使加权投票系统 $S = [q; w_1, w_2, \dots, w_n]$ 的 Banzhaf 指标 β 与人口比例 p (在归一化的意义下) 相近似, 且当 n 较大时近似程度很高. 下面用一个例子说明调整的方法.

例 6 回到例 1, 5 个区 A, B, C, D, E 的人口 (单位: 千人) 分别为 60, 20, 10, 5, 5, 比例 $p = (12, 4, 2, 1, 1)$. 若以 p 为权重, 用简单多数规则 $q = 11$, 加权投票系统 $S = [11; 12, 4, 2, 1, 1]$ 的 Banzhaf 指标显然是 $\beta = (1, 0, 0, 0, 0)$. 下面在权重不变而增大定额 q 的情况下, 借助分析极小获胜联盟的办法, 寻找 β 与 p 相近似的加权投票系统.

令 $q = 12$, 则 $S = [12; 12, 4, 2, 1, 1]$, 仍然有 $\beta = (1, 0, 0, 0, 0)$.

令 $q = 13$, 则 $S = [13; 12, 4, 2, 1, 1]$, 极小获胜联盟集 $W_m = ([AB], [AC])$,

$[AD], [AE]$), B, C, D, E 的权力相同, 即 $\beta_2 = \beta_3 = \beta_4 = \beta_5$, 与要求相差甚远.

令 $q = 14$, 则 $S = [14; 12, 4, 2, 1, 1]$, 极小获胜联盟集 $W_m = ([AB], [AC], [ADE])$, B, C 的权力相同, 即 $\beta_2 = \beta_3$, 与要求相差也远.

令 $q = 15$, 则 $S = [15; 12, 4, 2, 1, 1]$, 极小获胜联盟集 $W_m = ([AB], [ACD], [ACE])$, 从这些集合的结构可以知道, A, B, C, D, E 之间的权力大小的次序为 $\beta_1 > \beta_2 > \beta_3 > \beta_4 = \beta_5$, 与人口比例 $p = (12, 4, 2, 1, 1)$ 的大小次序相同, 是一个有希望的选择, 可仔细计算它的 Banzhaf 指标.

令 $q = 16$, 则 $S = [16; 12, 4, 2, 1, 1]$, 极小获胜联盟集 $W_m = ([AB], [ACDE])$, C, D, E 的权力相同, 即 $\beta_3 = \beta_4 = \beta_5$, 结果又变坏了.

根据以上的分析, 对最有希望的 $S = [15; 12, 4, 2, 1, 1]$ 计算 Banzhaf 指标 β , 结果是 $\beta = (11/21, 5/21, 3/21, 1/21, 1/21)$, 可以说是人口比例 $p = (12, 4, 2, 1, 1)$ 的一个不错的近似.

当 n 较大时, 调整加权投票系统的过程可以借助计算机, 通过人机交互方式完成. 一种可行的方法是:

首先定义 Banzhaf 指标 β 与人口比例 p 之间的“距离”(β, p 可看作 n 维空间的两个点), 作为衡量二者一致程度的指标. 按照实际需要确定一个“阈值”, 然后按以下步骤迭代地进行:

1) 给出权重 w_1, w_2, \dots, w_n 和定额 q 的一个初始值, 通常权重可取 p 或其近似, q 取稍大于 $w/2$ 的数;

2) 编程计算 Banzhaf 指标 β 及 β 与 p 的距离, 当距离小于阈值时停止, 否则转 3;

3) 改变 w_1, w_2, \dots, w_n 和 q , 转 2.

迭代过程中的难点在尚没有调整 w_1, w_2, \dots, w_n 和 q 的好方法. 每调整一次权重和定额, 必须使极小获胜联盟的结构有所变化, β 才有可能改进, 而怎样检查极小获胜联盟结构的变化则是需要仔细考虑的.

评注 任何一个构造和规则有明确定义的投票系统, 都可用极小获胜联盟来描述, 进而常常可以表示成加权投票系统, 例 4 就是这种情况.

有的投票系统, 即使确定了极小获胜联盟, 也无法表为加权投票系统. Hilliard 给出了区别加权与非加权投票系统的方法, 并提供算法, 以找到权重和定额, 或者指明不存在权重和定额的矛盾^[35].

两种权力指标常常给出相同或近似的结果, 而从理论上区分它们的数学公理既不直观, 在使用时也不具说服力, 所以在实际应用中公理化方法并不能解决选择哪个指标的问题. 一般认为, Banzhaf 指标在道理上更浅显, 容易口头解释, 因而更易为实际工作者接受; Shapley 指标可以作为对策论中早已著名的 Shapley 值方法的副产品, 在数学界更有市场. 还有的文章指出, Banzhaf 指标

适于设计投票系统,在代表尚未选出之前,假定所有投票意愿的等可能性是合理的;Shapley 指标适于评价投票系统,代表已经选出,他们的立场已为众人所知。

在加权投票系统中定义量化的权力指标,是将数学应用于社会政治领域的一个有意义的范例。虽然这些模型没有涉及实际的社会政治生活中既有联盟又有对抗、既有党内统一行动又有外界各种干扰、还有弃权、缺席、妥协甚至背叛等种种复杂因素,因而受到一些人的质疑,但是当对投票过程建立起一套规范、标准以及合理要求时,这些研究还是有用的,可以成为社会政治问题在量化分析方面公认的常识,权力指标作为判断社会政治结构中平等性的一种手段,也会显示出它的价值。近年来已经有人提出与“计量经济学”类似的新学科——“计量政治学”。

习 题

1. “田忌赛马”是一个家喻户晓的故事:战国时期,齐国将军田忌经常与齐王赛马,设重金赌注。孙膑发现田忌与齐王的马脚力都差不多,可分为上、中、下三等。于是孙膑对田忌说:“您只管下大赌注,我能让您取胜。”田忌相信并答应了他,与齐王用千金来赌胜。比赛即将开始,孙膑对田忌说:“现在用您的下等马对付他的上等马,拿您的上等马对付他的中等马,拿您的中等马对付他的下等马。”三场比赛完后,田忌只有一场不胜而另两场胜,最终赢得齐王的千金赌注。

(1) 分析这个故事中还隐含了哪些信息,并思考何时可以建模为一个博弈问题,何时只是一个简单的单人决策问题。

(2) 如果齐王和田忌约定比赛开始前双方同时决定马的出场顺序,并且以后不可改变,这个博弈是否存在纯战略纳什均衡? 如果不存在,求出该博弈模型的混合战略纳什均衡。

2. 1943年2月,第二次世界大战中的新几内亚战争处于关键阶段,日军决定从新不列颠附近的岛屿调派援兵。日军运输船可以沿新不列颠北侧航行,但是可能会遇上下雨,能见度也较差;或者沿岛屿的南侧航行,天气会比较好。不论哪种路线都需要三天时间。如果他们希望有个好天气,当然应该选择沿南部走的路线。但是战争期间日军指挥部希望运输船暴露由西南太平洋盟军空军司令肯尼将军指挥的美军攻击火力下的时间尽可能少。在这样的条件下日军应该选择哪条路线?

和日军一样,此刻肯尼将军也面临着困难的选择。盟军情报部门已经侦察到日本护卫舰队在新不列颠远侧集结。肯尼将军当然希望轰炸日军船队的天数达到最大。但是美军没有足够的侦察机兼顾南北两条路线,而无法尽早侦察到日本运输船的航行路线。因此,肯尼将军只能将大量的侦察机集中在南部或者北部路线上。肯尼将军应该怎么做呢?

如果盟军将侦察机集中在南部路线上,日军也选择南部路线,则盟军可以轰炸日军三天;而若日军选择北部路线,则盟军只能轰炸日军一天。如果盟军将侦察机集中在北部路线上,则无论日军选择哪条路线,盟军可以轰炸日军两天。

(1) 建立博弈模型描述双方指挥官的决策问题.

(2) 求出该博弈模型的纯战略纳什均衡,并查阅当时的历史,看看双方的行动是否确实与此一致.

3. 2004 年美国总统选举即将开始前,两位候选人布什和克里都把拉票的重点转移到了竞争异常激烈的宾夕法尼亚、俄亥俄、佛罗里达三个州.民意调查显示,当时布什在这三个州赢得选举的可能性分别是 20%、60% 和 80%.为了赢得整个选举,布什必须至少赢得其中两个州.假设如果两人同时到某个州拉票,则对每个州获胜的概率没有影响;如果两人到不同州拉票,则候选人在其所到访的州获胜的概率将增加 10%.由于剩余的时间只能允许每位候选人到其中一个州拉票,那么他们应该分别选择到哪个州^[31]?

(1) 建立博弈模型描述两位候选人的决策问题.

(2) 求出博弈模型的纯战略纳什均衡.

4. 我们经常见到媒体报道:一些不文明现象或违法行为发生在众目睽睽之下,却无人出面阻止或干预.如果不考虑这类事件的复杂的社会、道德等因素,你能否完全从数学的角度通过建立博弈模型来定量分析一下这种“人多未必势众”的现象?具体来说,希望你的模型回答下面的问题:假设有多个正在目睹某个不文明现象或违法行为,那么当目睹人数增加时,有人出面阻止或干预的可能性是增加了还是减少了?

5. 同类型的商家经常会出现“扎堆”现象,形成各式各样的商品城,如“书城”、“灯具城”等.人们有时不得不跑很远的路去这类商品城,于是会抱怨;如果他们大致均匀地分布到城市的不同地点,难道不是对商家更为有利可图,也更方便顾客?请你以下面的问题为例,作出适当的假设,进行建模分析:某海滨浴场准备设立两个售货亭,以供海滩上游泳和休闲的人购买饮用水和小食品等.那么,这两个售货亭的店主会分别将售货亭设立在哪里?

6. 分析两个完全类似的国家对某种商品的关税税率设定问题,假设每个国家对该商品的需求与价格间成线性减函数关系,两国博弈的顺序为:(1)两个国家的政府各自制定关于该商品的进口关税税率;(2)两个国家各有一家企业决定生产该商品的数量(设两家企业的单件生产成本相同,并且不考虑固定成本),其中一部分供国内消费,一部分供出口(设运输成本可忽略不计).每家企业关注的是最大化自己的利润,而每个国家关注的是最大化本国的社会总福利,包括本国消费者享受到的消费者剩余(如果消费者用价格 p 购得一件他愿意出价 v 购买的商品,他的剩余为 $v - p$)、企业利润和关税收入^[18].

(1) 建模并求出这个博弈的均衡.

(2) 如果两国政府联合起来决策,希望最大化两国的社会总福利之和,结果有什么不同,特别是,关税税率是会提高还是会下降?

7. (1) 在 11.2 节中曾假设回收协议模型的批发价、回收价以及按批发价回收的比例都没有上下限的限制,也说明通过协议参数的变化,可以达到系统协调并可在报社与报童间任意分配利润.如果法律或行规对这些参数有上下限的限制而不能任意取值,回收价格协议和回收数量协议是否还能在满足这些限制的条件下,通过选择参数来达到系统协调?

(2) 若上述参数有上下限的限制,当回收价格协议和回收数量协议都不能达到系统协调时,你能否构造新的协议,使之仍能达到系统协调?

8. 有两家企业同时向政府申请投资某个项目,对企业 $i(i=1,2)$ 这个项目申请得到批准

后的价值(利润)为 v_i . 假设政府的态度是:如果两家企业中没有任何一家放弃申请,则一直不批准这个项目,并且申请期间每单位时间向两家企业各收取 1 个单位的申请费,直到其中一家放弃申请后,将项目批给未放弃的企业. 如果同时放弃,可以通过公平的抽签过程让两家企业中的一家得到项目,机会都是 50%.

(1) 如果 $v_i (i=1,2)$ 是共同信息,问两家企业应该采用什么策略?

(2) 如果 v_i 是企业 i 的私有信息,对方只知道其服从 $[0,1]$ 上的均匀分布,问两家企业应该采用什么策略?

9. 按照 11.4 节介绍的公平性概念,建立考虑公平性的博弈模型,分析如下具有多个(至少两个)“反应者”的最后通牒博弈:首先由一个惟一指定的“提议者”提出一个分配提议——从总量为 1 的财富中分给反应者 s ,剩下的 $1-s$ 留给提议者自己;其次,由 $n-1$ 个“反应者”同时决定自己是否接受这个提议,如果没有人接受,则所有参与者什么也得不到;如果至少有一个人接受,则所有接受的反应者以等概率地(如通过抓阄)得到 s ,提议者得到 $1-s$. 假设 11.4 节(9)式定义的 $\beta_i < (n-1)/n$,给出这个博弈的均衡^[28].

10. 在 11.5 节中证明由(7),(8)给出的 Shapley 值 $\Phi(v)$ 满足 $\varphi_i(v) \geq v(i), i=1,2,\dots,n$.

11. 某甲(农民)有一块土地,若从事农业生产可收入 1 万元;若将土地租给某乙(企业家)用于工业生产,可收入 2 万元;若租给某丙(旅店老板)开发旅游业,可收入 3 万元;当旅店老板请企业家参与经营时,收入达 4 万元. 为促成最高收入的实现,试用 Shapley 值方法分配各人的所得.

12. 奇数个席位的理事会由三派组成,议案表决实行过半数通过方案. 证明在任一派都不能操纵表决的条件下,三派占有的席位不论多少,他们在表决中的权重都是一样的.

13. 记投票人集合为 $\{A, B, C, \dots\}$, 研究加权投票系统 S 与极小获胜联盟集 W_m 的关系:

(1) 设 S 为 $[4; 2, 2, 1, 1], [5; 3, 2, 2, 1], [3; 2, 1, 1, 1]$, 写出相应的 W_m .

(2) 设 W_m 为 $([ABC], [ABD], [ACD]), ([AB], [AC], [AD]), ([ABC], [ABD]),$ 写出相应的 S .

(3) 由 S 到 W_m 是存在且惟一的吗? 由 W_m 到 S 是存在且惟一的吗? 考察 $W_m = ([AB], [BC], [CD])$.

14. 在 11.6 节例 5, 由民主党 M(11 席)、共和党 G(14 席)和独立人士 D(15 席)并按简单多数规则投票组成的议会中,分别就 M 和 G 结盟与不结盟各种情况计算 Banzhaf 指标,得出类似 11.6 节表 2 的结果.

15. 对于 n 人投票系统 $S = [q; 1, 1, \dots, 1]$, 其中有 n 个 1, 设 n 为奇数, $q = (n+1)/2$, 证明当 n 很大时, Banzhaf 指标 β' 趋于 $1/\sqrt{n}$.

16. 对于加权投票系统 $S = [q; w_1, w_2, w_3]$, 设 $w_1 \geq w_2 \geq w_3, q > w/2$, 说明只有 5 种极小获胜联盟集, 对应于 4 种不同的 Banzhaf 指标 β .

17. 联合国安理会有 5 个常任理事国和 10 个非常任理事国组成, 仅当全部常任理事国和至少 4 个非常任理事国投票赞成时决议方能通过. 试将这种规则表示为一个加权投票系统, 并计算、比较它的 Shapley 指标和 Banzhaf 指标.

18. (1) 设某公司一个大股东控制股份的 40%, 其余 60% 由 6 位小股东平分, 按照控股的简单多数规则运作, 计算大股东的 Shapley 指标. 如果 60% 由 60 位、600 位…小股东平分的

话,大股东的 Shapley 指标如何变化,讨论小股东无穷多的极限情况.

(2) 若有两个主要股东分别控制股份的 $3/9$ 和 $2/9$,其余 $4/9$ 由很多小股东平分,仍按控股的简单多数规则运作,计算两个主要股东的 Shapley 指标,分析结果.

(3) 作为(2)的一般化,两个主要股东 A,B 分别控制股份的 x 和 y (都小于 $1/2$),其余由很多小股东平分,按控股的简单多数规则运作,用作图的方法计算 A,B 的 Shapley 指标.



第12章

马氏链模型

在考察有随机因素影响的动态系统时,常常碰到这样的情况:系统在每个时期所处的状态是随机的,从这个时期到下个时期的状态按照一定的概率进行转移,并且下个时期的状态只取决于这个时期的状态和转移概率,与以前各时期的状态无关.这种性质称为无后效性,或马尔可夫(Markov)性,通俗地说就是:已知现在,将来与历史无关.具有无后效性的,时间、状态均为离散的随机转移过程通常用马氏链(Markov Chain)模型描述.

马氏链模型在经济、社会、生态、遗传等许多领域中有着广泛的应用,值得提出的是,虽然它是解决随机转移过程的工具,但是一些确定性系统的状态转移问题也能用马氏链模型处理.

12.1节通过实例对本章模型用到的马氏链的基本知识作了简单介绍,不熟悉马氏链的读者请先阅读这一节.

12.1 健康与疾病

本节通过一个虽有实际背景、但数据纯系虚构的例子来介绍马氏链的基本概念,及其两种主要类型——正则链和吸收链——的性质.

人寿保险公司对投保人的健康状况特别关注,他们欢迎年轻力壮的人投保,患病者和高龄人则需付较高的保险金,甚至被拒之门外.人的健康状态随着时间的推移会发生转变,转变是随机的,保险公司要通过大量数据对状态转变的概率作出估计,才可能确定出不同年龄、不同健康状况的人的保险金和理赔金数额,下面分两种情况进行讨论.

1. 粗略地把人的健康状况分为健康和疾病两种状态,不妨以一年为一个时段研究状态的转变.假定对某一年龄段的人来说,今年健康、明年保持健康状态的概率为0.8,即明年转为疾病状态的概率为0.2;而今年患病、明年转为健康状态的概率为0.7,即明年保持疾病状态的概率为0.3.如果一个人投保时处于健康状态,我们研究以后若干年他分别处于这两种状态的概率.

用随机变量 X_n 表示第 n 年的状态, $X_n = 1$ 表示健康, $X_n = 2$ 表示疾病, $n = 0,$

1, 2, ... 用 $a_i(n)$ 表示第 n 年处于状态 i 的概率, $i=1, 2$, 即 $a_i(n) = P(X_n = i)$. 用 p_{ij} 表示已知今年处于状态 i , 来年处于状态 j 的概率, $i, j=1, 2$, 即 $p_{ij} = P(X_{n+1} = j | X_n = i)$. $a_i(n)$ 称为状态概率, p_{ij} 称为状态转移概率. 这种状态及其转移情况可以用图 1 表示.

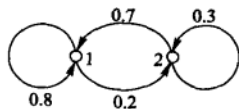


图 1 两种状态及其转移概率

显然, 第 $n+1$ 年的状态 X_{n+1} 只取决于第 n 年的状态 X_n 和转移概率 p_{ij} , 而与以前的状态 X_{n-1}, X_{n-2}, \dots 无关, 即状态转移具有无后效性. 第 $n+1$ 年的状态概率可由全概率公式得到

$$\begin{cases} a_1(n+1) = a_1(n)p_{11} + a_2(n)p_{21} \\ a_2(n+1) = a_1(n)p_{12} + a_2(n)p_{22} \end{cases} \quad (1)$$

由前 $p_{11} = 0.8, p_{12} = 0.2, p_{21} = 0.7, p_{22} = 0.3$, 投保人开始时处于健康状态, 即 $a_1(0) = 1, a_2(0) = 0$, 利用(1)式立即可以算出以后各年他处于两种状态的概率 $a_1(n), a_2(n), n=1, 2, \dots$, 如表 1.

表 1 投保人开始处于健康状态时两种状态概率的变化

n	0	1	2	3	4	...	∞
$a_1(n)$	1	0.8	0.78	0.778	0.7778	...	7/9
$a_2(n)$	0	0.2	0.22	0.222	0.2222	...	2/9

表中最后一列是根据计算数值的趋势猜测的.

如果投保人开始时处于疾病状态, 即 $a_1(0) = 0, a_2(0) = 1$, 类似地可得表 2.

表 2 投保人开始处于疾病状态时两种状态概率的变化

n	0	1	2	3	4	...	∞
$a_1(n)$	0	0.7	0.77	0.777	0.7777	...	7/9
$a_2(n)$	1	0.3	0.23	0.223	0.2223	...	2/9

显然表中最后一列和表 1 相同.

可以将众多投保人处于两种状态的比例, 视为典型的投保人处于两种状态的概率, 比如若健康人占 3/4, 病人占 1/4, 则可设初始状态概率为 $a_1(0) = 0.75, a_2(0) = 0.25$, 读者计算一下就会发现 $n \rightarrow \infty$ 时 $a_1(n), a_2(n)$ 的趋向也和表 1、表 2 相同.

可以看到, 对于给定的状态转移概率, $n \rightarrow \infty$ 时状态概率 $a_1(n), a_2(n)$ 趋向于稳定值, 该值与初始状态无关, 这是一种主要的马氏链类型的重要性质.

2. 把人的死亡作为第 3 种状态, 用 $X_n = 3$ 表示. 今年健康、明年可能因突发

疾病或偶然事故而死亡,今年患病、明年更可能转为死亡,而一旦死亡当然就不能再转为健康或疾病状态.将3种状态的转移表示为图2.

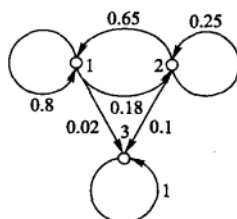


图2 三种状态及其转移概率

仍用 $a_i(n)$ 表示状态概率, $i=1,2,3$, 用 p_{ij} 表示状态转移概率, $i,j=1,2,3$, 具体数值见图2, 特别注意, $p_{31}=p_{32}=0, p_{33}=1$. 第 $n+1$ 年的状态概率可类似于(1)式用

$$\begin{cases} a_1(n+1) = a_1(n)p_{11} + a_2(n)p_{21} + a_3(n)p_{31} \\ a_2(n+1) = a_1(n)p_{12} + a_2(n)p_{22} + a_3(n)p_{32} \\ a_3(n+1) = a_1(n)p_{13} + a_2(n)p_{23} + a_3(n)p_{33} \end{cases} \quad (2)$$

计算,若投保人开始时处于健康状态,则由(2)式算出以后各年他处于三种状态的概率,如表3.

表3 投保人处于健康状态时三种状态概率的变化

n	0	1	2	3	...	30	...	50	...	∞
$a_1(n)$	1	0.8	0.757	0.728 5	...	0.269 8	...	0.129 3	...	0
$a_2(n)$	0	0.18	0.189	0.183 5	...	0.068 0	...	0.032 6	...	0
$a_3(n)$	0	0.02	0.054	0.088 0	...	0.662 1	...	0.838 1	...	1

表中最后一列是根据计算数值的趋势猜测的. 如果设初始状态概率为 $a_1(0) = 0.75, a_2(0) = 0.25, a_3(0) = 0$, 会发现 $n \rightarrow \infty$ 时 $a_1(n), a_2(n), a_3(n)$ 的趋向和表3相同.

可以看到,不论初始状态如何,最终都要转到状态3,这代表了另一种主要的马氏链类型.

通过这个例子容易了解下面给出的马氏链的基本概念.

马氏链及其基本方程 按照系统的发展,时间离散化为 $n=0,1,2,\dots$, 对每个 n , 系统的状态用随机变量 X_n 表示, 设 X_n 可以取 k 个离散值 $X_n=1,2,\dots,k$, 且记 $a_i(n) = P(X_n = i)$, 即状态概率, 从 $X_n = i$ 到 $X_{n+1} = j$ 的概率记 $p_{ij} = P(X_{n+1} = j | X_n = i)$, 即转移概率. 如果 X_{n+1} 的取值只取决于 X_n 的取值及转移概

率,而与 X_{n-1}, X_{n-2}, \dots 的取值无关,那么这种离散状态按照离散时间的随机转移过程称为马氏链. 由状态转移的无后效性和全概率公式可以写出马氏链的基本方程为

$$a_i(n+1) = \sum_{j=1}^k a_j(n)p_{ji}, \quad i=1, 2, \dots, k \quad (3)$$

并且 $a_i(n)$ 和 p_{ij} 应满足

$$\sum_{i=1}^k a_i(n) = 1, \quad n=0, 1, 2, \dots \quad (4)$$

$$p_{ij} \geq 0, \quad i, j=1, 2, \dots, k \quad (5)$$

$$\sum_{j=1}^k p_{ij} = 1, \quad i=1, 2, \dots, k \quad (6)$$

引入状态概率向量(行向量)和转移概率矩阵(简称转移矩阵)

$$\mathbf{a}(n) = (a_1(n), a_2(n), \dots, a_k(n)), \mathbf{P} = \{p_{ij}\}_{k \times k} \quad (7)$$

则基本方程(3)可以表为

$$\mathbf{a}(n+1) = \mathbf{a}(n)\mathbf{P} \quad (8)$$

由此还可得到

$$\mathbf{a}(n) = \mathbf{a}(0)\mathbf{P}^n \quad (9)$$

(5) 式表明转移矩阵 \mathbf{P} 是非负阵, (6) 式表示 \mathbf{P} 的行和为 1, 称为随机矩阵.

对于上例的两种情况, 转移矩阵分别为

$$\mathbf{P} = \begin{bmatrix} 0.8 & 0.2 \\ 0.7 & 0.3 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 0.8 & 0.18 & 0.02 \\ 0.65 & 0.25 & 0.1 \\ 0 & 0 & 1 \end{bmatrix}$$

容易看出, 对于马氏链模型最基本的问题是构造状态 X_n 及写出转移矩阵 \mathbf{P} . 一旦有了 \mathbf{P} , 那么给定初始状态概率 $\mathbf{a}(0)$ 就可以用(9)或(8)式计算任意时段 n 的状态概率 $\mathbf{a}(n)$.

应该指出, 这里的转移概率 p_{ij} 与时段 n 无关, 这种马氏链称为时齐的. 本章讨论的均为这种情况.

从上面的计算结果可以看出这两个马氏链之间有很大差别, 事实上它们属于马氏链的两个重要类型, 下面分别作简单介绍.

正则链 这类马氏链的特点是, 从任意状态出发经过有限次转移都能达到另外的任意状态, 有如下定义.

定义 1 一个有 k 个状态的马氏链如果存在正整数 N , 使从任意状态 i 经 N 次转移都以大于零的概率到达状态 j ($i, j=1, 2, \dots, k$), 则称为正则链.

用下面的定理容易检验一个马氏链是否是正则链.

定理 1 若马氏链的转移矩阵为 \mathbf{P} , 则它是正则链的充要条件是, 存在正整

数 N , 使 $P^N > 0$ (指 P^N 的每一元素大于零).

上面第 1 种情况的转移矩阵显然满足定理 1, 即它是正则链. 我们已经看到, 从任意初始状态 $a(0)$ 出发, $n \rightarrow \infty$ 时状态概率 $a(n)$ 趋于与 $a(0)$ 无关的稳定值. 事实上有如下的定理.

定理 2 正则链存在唯一的极限状态概率 $w = (w_1, w_2, \dots, w_k)$, 使得当 $n \rightarrow \infty$ 时状态概率 $a(n) \rightarrow w$, w 与初始状态概率 $a(0)$ 无关. w 又称稳态概率, 满足

$$wP = w \quad (10)$$

$$\sum_{i=1}^k w_i = 1 \quad (11)$$

定理中的 (10) 式可以从 (8) 式直接得到. P 给定后 (10) 和 (11) 式构成求解 w 的线性方程组. 表 1, 2 中的 $w = (7/9, 2/9)$ 即可直接从相应的代数方程求出.

从 (9) 或 (10) 式还不难看出, $\lim_{n \rightarrow \infty} P^n$ 存在, 记作 P^* , 并且 P^* 的每一行都是稳态概率 w . 如果记 $P^* = \{p_{ij}^{(*)}\}$, 那么有 $p_{ii}^{(*)} = w_i$.

从状态 i 出发经 n 次转移, 第一次到达状态 j 的概率称为 i 到 j 的首达概率, 记作 $f_{ij}(n)$. 于是

$$\mu_{ij} = \sum_{n=1}^{\infty} n f_{ij}(n) \quad (12)$$

为从状态 i 第一次到达状态 j 的平均转移次数. 特别地, μ_{ii} 是状态 i 首次返回的平均转移次数. μ_{ii} 与稳态概率 w 有密切关系, 即

定理 3 对于正则链

$$\mu_{ii} = 1/w_i \quad (13)$$

吸收链 上面第 2 种情况状态 3 的转移概率 $p_{33} = 1$, 于是系统一旦进入状态 3 就再不会离开它, 可以把它看作“吸收”其他状态的一个状态. 并且从状态 1 或 2 出发, 可以经有限次转移到达状态 3, 有如下定义.

定义 2 转移概率 $p_{ii} = 1$ 的状态 i 称为吸收状态. 如果马氏链至少包含一个吸收状态, 并且从每一个非吸收状态出发, 能以正的概率经有限次转移到达某个吸收状态, 那么这个马氏链称为吸收链.

吸收链的转移矩阵可以写成简单的标准形式. 若有 r 个吸收状态, $k-r$ 个非吸收状态, 则转移矩阵 P 可表为

$$P = \begin{bmatrix} I_{r \times r} & O \\ R & Q \end{bmatrix} \quad (14)$$

其中 $k-r$ 阶子方阵 Q 的特征值 λ 满足 $|\lambda| < 1$. 这要求子阵 $R_{(k-r) \times r}$ 中必含有非零元素, 以满足从任一非吸收状态出发经有限次转移可到达某吸收状态的条件. 这样 Q 就不是随机矩阵, 它至少存在一个小于 1 的行和, 且如下定理成立.

定理 4 对于吸收链 P 的标准形式 (14), $I-Q$ 可逆,

$$M = (I - Q)^{-1} = \sum_{n=0}^{\infty} Q^n \quad (15)$$

记元素全为 1 的列向量为 $e = (1, 1, \dots, 1)^T$, 则

$$y = Me \quad (16)$$

的第 i 分量是从第 i 个非吸收状态出发, 被某个吸收状态吸收的平均转移次数。

设状态 i 是非吸收状态, j 是吸收状态, 那么首达概率 $f_{ij}(n)$ 实际上是 i 经 n 次转移被 j 吸收的概率, 而

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n) \quad (17)$$

则是从非吸收状态 i 出发终将被吸收状态 j 吸收的概率。记 $F = \{f_{ij}\}_{(k-r) \times r}$, 下面的定理给出了计算 f_{ij} 的方法。

定理 5 设吸收链的转移矩阵 P 表为标准形式 (14), 则

$$F = MR \quad (18)$$

12.2 钢琴销售的存贮策略

像钢琴这样的奢侈品销售量很小, 商店里一般不会有这么大的库存量让它积压资金。这里通过一个简单的实例来分析、评价一种存贮策略的效果。

一家商店根据以往经验, 平均每周只能售出 1 架钢琴。现在经理制订的存贮策略是, 每周末检查库存量, 仅当库存量为零时, 才订购 3 架供下周销售; 否则, 不订购。试估计在这种策略下失去销售机会的可能性有多大, 以及每周的平均销售量是多少^[53]。

问题分析 对于钢琴这种商品的销售, 顾客的到来是相互独立的, 在服务系统中通常认为需求量近似服从泊松分布, 其参数可由均值为每周销售 1 架得到, 由此可以算出不同需求量的概率。周末的库存可能是 0, 1, 2, 3 (架), 而周初的库存量则只有 1, 2, 3 这 3 种状态, 每周不同的需求将导致周初库存状态的变化, 于是可用马氏链来描述这个过程。当需求超过库存时就会失去销售机会, 可以计算这种情况发生的概率。在动态过程中这个概率每周是不同的, 每周的销售量也不同, 通常采用的办法是在时间充分长以后, 按稳态情况进行分析和计算。

模型假设

1. 钢琴每周需求量服从泊松分布, 均值为每周 1 架。
2. 存贮策略是: 当周末库存量为零时, 订购 3 架, 周初到货; 否则, 不订购。
3. 以每周初的库存量作为状态变量, 状态转移具有无后效性。
4. 在稳态情况下计算该存贮策略失去销售机会的概率和每周的平均销售量。

模型建立 记第 n 周的需求量为 D_n , 由假设 1, D_n 服从均值为 1 的泊松分布, 即

$$P(D_n = k) = e^{-1}/k! \quad (k=0, 1, 2, \dots) \quad (1)$$

记第 n 周初的库存量为 S_n , $S_n \in \{1, 2, 3\}$ 是这个系统的状态变量, 由假设 2, 状态转移规律为

$$S_{n+1} = \begin{cases} S_n - D_n, & D_n < S_n \\ 3, & D_n \geq S_n \end{cases} \quad (2)$$

由(1)式不难算出 $P(D_n = 0) = 0.368$, $P(D_n = 1) = 0.368$, $P(D_n = 2) = 0.184$, $P(D_n = 3) = 0.061$, $P(D_n > 3) = 0.019$, 由此计算状态转移矩阵

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

其中

$$p_{11} = P(S_{n+1} = 1 | S_n = 1) = P(D_n = 0) = 0.368$$

$$p_{12} = P(S_{n+1} = 2 | S_n = 1) = 0$$

$$p_{13} = P(S_{n+1} = 3 | S_n = 1) = P(D_n \geq 1) = 0.632$$

$$p_{21} = P(S_{n+1} = 1 | S_n = 2) = P(D_n = 1) = 0.368$$

$$p_{22} = P(S_{n+1} = 2 | S_n = 2) = P(D_n = 0) = 0.368$$

$$p_{23} = P(S_{n+1} = 3 | S_n = 2) = P(D_n \geq 2) = 0.264$$

$$p_{31} = P(S_{n+1} = 1 | S_n = 3) = P(D_n = 2) = 0.184$$

$$p_{32} = P(S_{n+1} = 2 | S_n = 3) = P(D_n = 1) = 0.368$$

$$p_{33} = P(S_{n+1} = 3 | S_n = 3) = P(D_n = 0) + P(D_n \geq 3) = 0.448$$

得到

$$P = \begin{bmatrix} 0.368 & 0 & 0.632 \\ 0.368 & 0.368 & 0.264 \\ 0.184 & 0.368 & 0.448 \end{bmatrix} \quad (3)$$

记状态概率 $a_i(n) = P(S_n = i)$, $i = 1, 2, 3$, $\alpha(n) = (a_1(n), a_2(n), a_3(n))$, 根据状态转移具有无后效性的假设, 有 $\alpha(n+1) = \alpha(n)P$. 用 12.1 节的定理 1 对照得到的转移矩阵 P , 可知这是一个正则链, 具有稳态概率 w , w 可由 12.1 节 (10), (11) 式得到:

$$w = (w_1, w_2, w_3) = (0.285, 0.263, 0.452) \quad (4)$$

该存贮策略(第 n 周)失去销售机会的概率为 $P(D_n > S_n)$, 按照全概率公式有

$$P(D_n > S_n) = \sum_{i=1}^3 P(D_n > i | S_n = i) P(S_n = i) \quad (5)$$

其中的条件概率 $P(D_n > i | S_n = i)$ 容易由(1)式计算. 当 n 充分大时, 可以认为

$$P(S_n = i) = w_i, \quad i = 1, 2, 3$$

最终得到

$$P(D_n > S_n) = 0.264 \times 0.285 + 0.080 \times 0.263 + 0.019 \times 0.452 = 0.105$$

即从长期看,失去销售机会的可能性大约 10%.

在计算该存贮策略(第 n 周)的平均销售量 R_n 时,应注意到,当需求超过存量时只能销售掉存量,于是

$$R_n = \sum_{i=1}^3 \left[\sum_{j=1}^{i-1} jP(D_n = j | S_n = i) + iP(D_n \geq i | S_n = i) \right] P(S_n = i) \quad (6)$$

同样地,当 n 充分大时用稳态概率 w_i 代替 $P(S_n = i)$,得到

$$R_n = 0.632 \times 0.285 + 0.896 \times 0.263 + 0.976 \times 0.452 = 0.857$$

即从长期看,每周的平均销售量为 0.857 架.你能解释为什么这个数值略小于模型假设中给出的每周平均需求量为 1 架吗?

敏感性分析 这个模型用到的惟一个原始数据是,平均每周售出 1 架钢琴,这个数值会有波动.为了计算当平均需求在 1 附近波动时,最终结果有多大变化,设 D_n 服从均值为 λ 的泊松分布,即有 $P(D_n = k) = \lambda^k e^{-\lambda} / k!$ ($k = 0, 1, 2, \dots$),由此得状态转移矩阵为

$$P = \begin{bmatrix} e^{-\lambda} & 0 & 1 - e^{-\lambda} \\ \lambda e^{-\lambda} & e^{-\lambda} & 1 - (1 + \lambda)e^{-\lambda} \\ \lambda^2 e^{-\lambda} / 2 & \lambda e^{-\lambda} & 1 - (\lambda + \lambda^2 / 2)e^{-\lambda} \end{bmatrix} \quad (7)$$

对于不同的平均需求 λ (在 1 附近),类似于上面的计算过程,记 $P = P(D_n > S_n)$,可得到以下结果:

λ	0.8	0.9	1.0	1.1	1.2
P	0.073	0.089	0.105	0.122	0.139

即当平均需求增长(或减少)10%时,失去销售机会的概率将增长(或减少)约 15%,这是可以接受的.

类似地可以作每周平均销售量的敏感性分析.

评注 本节介绍的是对已经制订的存贮策略,用两个指标加以评价,还可以给出其他的策略和指标,作进一步的研究,见习题 1,2.

动态随机存贮策略是马氏链的典型应用,关键之一是在无后效性的前提下恰当地定义系统的状态.本例中以每周初的库存量作为状态变量即可,但是如果制订的存贮策略不仅与本周的销售有关,还要考虑上周销售的情况,那么状态变量就要扩充,包含 S_n 和 S_{n-1} ,相应的状态转移概率也要改变.

12.3 基因遗传

豆科植物茎的颜色有绿有黄,生猪的毛有黑有白,有粗有光,人类会出现先天性疾病如色盲等,这些都是基因遗传的结果.基因从一代到下一代的转移是随机的,并且具有无后效性,因此马氏链模型是研究遗传学的重要工具之一.本节给出的简单模型属于完全优势基因遗传理论的范畴.

生物的外部表征,如豆科植物茎的颜色,人的皮肤或头发的色素,由生物体内相应的基因决定.基因分优势基因和劣势基因两种,分别用 d 和 r 表示.每种外部表征由体内的两个基因决定,而每个基因都可以是 d 或 r 中的一个,于是有三种基因类型,即 $D(dd)$, $H(dr)$ 和 $R(rr)$,分别称为优种、混种和劣种.含优种 D 和混种 H 基因类型的个体,外部表征呈优势,如豆科植物的茎呈绿色,人的皮肤或头发有色素;含劣种 R 基因类型的个体,外部表征呈劣势,如豆科植物的茎呈黄色,人的皮肤或头发无色素.

生物繁殖时,一个后代随机地继承父亲两个基因中的一个和母亲两个基因中的一个^①,形成它的两个基因.一般两个基因中哪一个遗传下去是等概率的,所以父母的基因类型就决定了每一后代基因类型的概率.父母基因类型有全是优种 DD ,全是劣种 RR ,一优种—混种 DH (父为 D ,母为 H 或父为 H ,母为 D),及 DR , HH , HR 共 6 种组合,对每种组合简单的计算可以得到其后代各种基因类型的概率,如表 1 所示.

表 1 父母基因类型决定后代各种基因类型的概率

父母基因类型 后代基因类型	DD	RR	DH	DR	HH	HR
D	1	0	1/2	0	1/4	0
H	0	0	1/2	1	1/2	1/2
R	0	1	0	0	1/4	1/2

下面我们以马氏链为工具讨论两个具体的基因遗传模型^[64,85].

随机交配 这是自然界中生物群体的一种常见的、最简单的交配方式.假设一个群体中雄性和雌性的比例相等,并且有相同的基因类型分布,即雄性中 D , H , R 的比例和雌性中 D , H , R 的比例相等.所谓随机交配是指:对于每一个(不论属于 D , H 或 R)的雌性(或雄性)个体,都以 D , H , R 的数量比例为概率,与一

① 本节的模型主要针对人类以外的生物,但是为了叙述的方便,仍沿用父亲、母亲这样的称谓.

个雄性(或雌性)个体交配,其后代则按照前面所说的方式,等概率地继承其父母亲的各一个基因,形成它的基因类型.假定在初始一代的群体中,三种基因类型的数量比例是 $D(dd):H(dr):R(rr) = a:2b:c$, 满足 $a+2b+c=1$. 记 $p=a+b, q=b+c$, 则群体中优势基因 d 与劣势基因 r 的数量比例为 $d:r=p:q$, 且 $p+q=1$.

下面讨论随机交配方式产生的一系列后代群体中的基因类型分布.

用 $X_n=1, 2, 3$ 分别表示第 n 代的一个体属于 D, H 及 R 基因类型, 即 3 种状态, $n=0, 1, 2, \dots$. $a_i(n)$ 表示个体属于第 i 种状态的概率, $i=1, 2, 3$, 可视为第 n 代的群体属于第 i 种基因类型的比例. 转移概率 p_{ij} 可用 $p_{ij}=P(\text{一个后代具有基因类型 } j | \text{ 母亲具有基因类型 } i)$ 计算^①. 在已知母亲基因类型的条件下, 后代的基因类型取决于父亲的基因类型. 值得指出的是, 在计算 p_{ij} 时与其考虑被随机选择为父亲的 3 种不同基因类型的比例 $a:2b:c$, 不如直接考察从雄性群体中以 $p:q$ 的比例获得优势基因 d 和劣势基因 r . 比如 $p_{11}=P(\text{后代为 } D(dd) | \text{ 母亲为 } D(dd))=p$, 因为后代只需从雄性群体中以概率 p 获得 d . 类似地有 $p_{12}=P(\text{后代为 } H(dr) | \text{ 母亲为 } D(dd))=q, p_{13}=P(\text{后代为 } R(rr) | \text{ 母亲为 } D(dd))=0$. 而 $p_{21}=P(\text{后代为 } D(dd) | \text{ 母亲为 } H(dr))=p/2$, 因为后代需以 $1/2$ 的概率从母体获得 d , 同时以 p 的概率从雄性群体中获得 d . 同理有 $p_{22}=P(\text{后代为 } H(dr) | \text{ 母亲为 } H(dr))=p/2+q/2=1/2, p_{23}=q/2$. 用同样的方法算出 p_{31}, p_{32}, p_{33} 后得到转移矩阵

$$P = \begin{bmatrix} p & q & 0 \\ p/2 & 1/2 & q/2 \\ 0 & p & q \end{bmatrix} \quad (1)$$

若初始一代是从 3 种基因类型比例为 $a:2b:c$ 的群体中随机选取的, 那么初始状态概率为 $a(0)=(a, 2b, c)$, 其中 $a, 2b, c$ 满足

$$p=a+b, \quad q=b+c \quad (2)$$

利用马氏链基本方程(12.1 节(8)式)可以得到

$$a(1)=a(0)P=(p^2, 2pq, q^2) \quad (3)$$

$$a(2)=a(1)P=(p^2, 2pq, q^2) \quad (4)$$

显然这个分布将保持下去. 这表明在随机交配方式中第一代继承者的基因类型分布为 $D:H:R=p^2:2pq:q^2$, 并永远不变. 这个结果在遗传学中称 Hardy-Weinberg 平稳定律.

由 12.1 节定理 1 容易知道这是一个正则链, 由 12.1 节(10)式算出其稳态概率也是 $w=(p^2, 2pq, q^2)$. 表明即使初始一代不是从群体中随机选取, 在随机交配方式下, 经过足够长时间, 3 种基因类型的分布也趋向上述稳定分布.

^① p_{ij} 中的条件也可用“父亲具有基因类型 i ”, 结果相同.

这个模型得到的结果的正确性已由观察和实验证明。如自然界中通常有 $p=q=1/2$, 于是 3 种基因类型的稳定分布为 $D:H:R=1/4:1/2:1/4$, 而优种 D 和混种 H 的外部表征呈优势。据观察, 豆科植物茎呈绿色(优势表征)的约占 $3/4$, 与上面的结果相一致。

最后考察在随机交配下 3 种基因类型的首次返回平均转移次数, 即平均经过多少代每种基因类型首次回到原来的类型。据 12.1 节定理 3, D, H, R 类型的首次返回平均换代数目

$$\mu_{11} = \frac{1}{p^2}, \quad \mu_{22} = \frac{1}{2pq}, \quad \mu_{33} = \frac{1}{q^2} \quad (5)$$

即一个群体中基因 d 越多(p 越大), 基本类型 $D(dd)$ 的平均换代数目越小。特别, 当 $p=q=1/2$ 时, D, H, R 的平均换代数目分别为 4(代), 2(代)和 4(代)。

近亲繁殖 这是指这样一种繁殖方式, 从同一对父母的大量后代中, 随机地选取一雄一雌进行交配, 产生后代, 如此继续下去, 考察一系列后代的基因类型的演变情况。

与前面的模型不同的是, 那里讨论后代群体中基因类型的分布, 只需设置 D, H, R 三个状态即可。这里则需按照随机选取的雄雌配对, 分析后代配对中基因类型的变化。于是状态应取雄雌 6 种基因类型组合, 设 $X_n = 1, 2, 3, 4, 5, 6$ 依次定义为 DD, RR, DH, DR, HH, HR 。

构造马氏链模型的关键是写出转移概率 p_{ij} , 它可根据本节开始给出的表 1 算出。显然 $p_{11} = 1, p_{1j} = 0 (j \neq 1), p_{22} = 1, p_{2j} = 0 (j \neq 2)$, 因为父母全为优种 D (或劣种 R) 时, 后代全是优种 (或劣种), 随机选取的雄雌配对当然也是。 $p_{31} = 1/4$, 因为配对 DH (状态 3) 的后代中 D 和 H 各占 $1/2$, 所以随机选取的配对为 DD (状态 1) 的概率是 $1/2 \cdot 1/2 = 1/4$, 而 $p_{33} = P(\text{后代配对为 } DH | \text{父母配对为 } DH) = P(\text{后代雄性为 } D, \text{雌性为 } H | \text{父母配对为 } DH) + P(\text{后代雄性为 } H, \text{雌性为 } D | \text{父母配对为 } DH) = 1/2 \cdot 1/2 + 1/2 \cdot 1/2 = 1/2$, 同理有 $p_{35} = 1/4$ 。又因配对 DH 的后代中没有 R , 故对于含有 R 的状态 2, 4, 6, 有 $p_{32} = p_{34} = p_{36} = 0$ 。其他的 p_{ij} 可以类似地计算 (留作习题 3), 最后得到转移矩阵为

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1/4 & 0 & 1/2 & 0 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1/16 & 1/16 & 1/4 & 1/8 & 1/4 & 1/4 \\ 0 & 1/4 & 0 & 0 & 1/4 & 1/2 \end{bmatrix} \quad (6)$$

容易看出, 状态 1(DD) 和状态 2(RR) 是吸收状态, 这是一个吸收链 (12.1 节定义 2)。它表明不论最初选取的配对是哪种基因类型组合, 经过若干代近亲

繁殖, 终将变为 DD 或 RR , 即变成全是优种或全是劣种, 而且一旦如此, 就永远保持下去。

为了计算从任一非吸收状态 3, 4, 5, 6 出发, 平均经过多少代就会被吸收状态 1, 2 吸收, 我们首先将 (6) 式表示的 P 化为 12.1 节给出的转移矩阵的标准形式 (14) 式, 得到

$$Q = \begin{bmatrix} 1/2 & 0 & 1/4 & 0 \\ 0 & 0 & 1 & 0 \\ 1/4 & 1/8 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/2 \end{bmatrix}, R = \begin{bmatrix} 1/4 & 0 \\ 0 & 0 \\ 1/16 & 1/16 \\ 0 & 1/4 \end{bmatrix} \quad (7)$$

然后按 12.1 节 (15), (16), (18) 式计算

$$M = (I - Q)^{-1} = \begin{bmatrix} 8/3 & 1/6 & 4/3 & 2/3 \\ 4/3 & 4/3 & 8/3 & 4/3 \\ 4/3 & 1/3 & 8/3 & 4/3 \\ 2/3 & 1/6 & 4/3 & 8/3 \end{bmatrix} \quad (8)$$

$$y = Me = \left(4 \frac{5}{6}, 6 \frac{2}{3}, 5 \frac{2}{3}, 4 \frac{5}{6} \right)^T \quad (9)$$

$$F = MR = \begin{bmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix} \quad (10)$$

M 的第 1 行至第 4 行依次代表非吸收状态 DH, DR, HH 和 HR , 根据 12.1 节定理 4 对于向量 y 的各个分量的解释, 从 DH 配对的状态出发, 在近亲繁殖的情况下平均经过 $4 \frac{5}{6}$ 代就会被状态 DD 或 RR 吸收, 即全变成优种或劣种。而根据定理 5, 被吸收状态吸收的概率为矩阵 F 的第 1 行元素, 即变成优种和劣种的概率分别为 $3/4$ 和 $1/4$ 。从其他状态 DR, HH 和 HR 出发, 可以得到相应的结论。

上述结果的实用价值在于, 在农业和畜牧业中常常是纯种 (优种或劣种) 生物的某些品质 (如抗病性) 不如混种, 所以在近亲繁殖情况下大约经过 5~6 代就应该重新选种, 以防止品质的下降。

12.4 等级结构

在社会系统中常常按照人们的职务或地位划分出许多等级, 如大学教师分为教授、讲师和助教, 工厂技术人员分为高级工程师、工程师和技术员, 军队里有将、校、尉, 学生也有研究生、大学生、中学生等。不同等级人员的比例形成一个等

级结构,合适的、稳定的等级结构有利于教学、科研、生产等各方面工作的顺利进行.因此希望建立一个模型来描述等级结构的变化状况,根据已知条件和当前的结构预报未来的结构,以及寻求为了达到某个理想的等级结构而应采取的策略^[8].

引起等级结构变化的因素有两种,一是系统内部等级间的转移,即提升或降级;二是系统内外的交流,即调入或退出(包括调离、退休、死亡等).系统各个等级的人员每个时期按一定的比例变化,本是一个确定性的转移问题,但是如果我们把这种比例视为各等级的每个成员提升、降级或退出的概率,就能利用处理随机转移的马氏链模型描述等级结构的变化.当然在这种观点下各等级成员的数量应理解为相应的平均值.为了叙述的统一和方便,以下均采用比例、比例分布等词汇,不称概率、概率分布等.

下面先定义若干基本量,建立基本方程,然后讨论如何调节调入成员在各等级的比例,保持等级结构的稳定,以及怎样尽快地达到或接近某个理想的等级结构.

基本量与基本方程 设一个社会系统由低到高地分为 k 个等级,如大学教师有助教、讲师、教授 3 个等级.时间以年为单位离散化,即每年进行且只进行一次调级.等级记作 $i = 1, 2, \dots, k$, 时间记作 $t = 0, 1, 2, \dots$. 引入以下的定义和记号:

成员按等级的分布向量 $\mathbf{n}(t) = (n_1(t), n_2(t), \dots, n_k(t))$, 其中 $n_i(t)$ 为 t 年属于等级 i 的人数; $N(t) = \sum_{i=1}^k n_i(t)$ 为系统 t 年的总人数.

成员按等级的比例分布 $\mathbf{a}(t) = (a_1(t), a_2(t), \dots, a_k(t))$, 其中 $a_i(t) = \frac{n_i(t)}{N(t)}$, 于是有 $a_i(t) \geq 0$, $\sum_{i=1}^k a_i(t) = 1$, $\mathbf{a}(t)$ 又称为等级结构.

转移比例矩阵 $\mathbf{Q} = \{p_{ij}\}_{k \times k}$, 其中 p_{ij} 为每年从等级 i 转移至等级 j 的成员(在等级 i 中占的)比例.

退出比例向量 $\mathbf{w} = (w_1, w_2, \dots, w_k)$, 其中 w_i 为每年从等级 i 退出系统的成员(在等级 i 中占的)比例; t 年退出系统总人数为

$$W(t) = \sum_{i=1}^k w_i n_i(t) = \mathbf{n}(t) \mathbf{w}^T \quad (1)$$

容易看出, p_{ij}, w_i 满足

$$p_{ij}, w_i \geq 0, \sum_{j=1}^k p_{ij} + w_i = 1 \quad (2)$$

调入比例向量 $\mathbf{r} = (r_1, r_2, \dots, r_k)$, 其中 r_i 为每年调入等级 i 的成员(在总调入人数中占的)比例; 记 t 年调入总人数为 $R(t)$, 则 t 年调入等级 i 的人数为 $r_i R(t)$. r_i 满足

$$r_i \geq 0, \quad \sum_{i=1}^k r_i = 1 \quad (3)$$

等级结构的基本方程 为了导出成员按等级的分布 $n(t)$ 的变化规律,先写出总人数 $N(t)$ 的方程

$$N(t+1) = N(t) + R(t) - W(t) \quad (4)$$

和每个等级人数的转移方程

$$n_j(t+1) = \sum_{i=1}^k p_{ij} n_i(t) + r_j R(t), \quad j=1, 2, \dots, k \quad (5)$$

请读者考虑,为什么在形式上方程(5)中没有像(4)中那样,减去退出的人数。

用向量、矩阵符号可将(5)式表为

$$n(t+1) = n(t)Q + R(t)r \quad (6)$$

从 t 到 $t+1$ 年总人数的增长量记为 $M(t)$,再由(4),(1)式可得

$$R(t) = W(t) + M(t) = n(t)w^T + M(t) \quad (7)$$

将(7)代入(6)式得到

$$n(t+1) = n(t)(Q + w^T r) + M(t)r \quad (8)$$

记

$$P = Q + w^T r \quad (9)$$

从(2),(3)式可知 P 的行和为1,是一个随机矩阵。(8)式记为

$$n(t+1) = n(t)P + M(t)r \quad (10)$$

当已知系统内部转移比例矩阵 Q ,调入比例 r ,初始的成员等级分布 $n(0)$,以及每年调入总人数 $R(t)$ 或每年总增长量 $M(t)$ 时,可以用(6)式或(9),(10)式计算成员等级分布的变化情况 $n(t)$ 。(6)或(9),(10)即为等级结构的基本方程。

基本方程的特殊形式 当每年系统总人数以固定的百分比 α 增长时,即 $M(t) = \alpha N(t)$,可用成员的等级结构 $a(t)$ 代替 $n(t)$ 而将(10)式表示为

$$a(t+1) = (1 + \alpha)^{-1} [a(t)P + \alpha r] \quad (11)$$

如果每年进出系统的人数大致相等,可以简化地假定系统总人数 $N(t)$ 保持不变,即 $M(t) = 0$ (或 $\alpha = 0$),这样(10)或(11)式化为相当简单的形式

$$a(t+1) = a(t)P = a(t)(Q + w^T r) \quad (12)$$

与马氏链的基本方程(12.1节(8)式)完全一致,等级结构 $a(t)$ 相当于状态概率,只是要注意到这里的转移矩阵 P 的 i, j 元素为 $p_{ij} + w_i r_j$,即系统内部转移比例 p_{ij} 加上系统内外交流的比例 $w_i r_j$ 。

下面在方程(12)的基础上进行讨论。

用调入比例进行稳定控制 我们的中心问题是通过对调入比例 r 的调节,尽快地达到或接近给定的理想等级结构 a^* 。但是等级结构 $a(t)$ 是按照(12)式的规律变化的,人们自然希望 a^* 一旦达到,就能够通过选取适当的调入比例使

a^* 保持不变. 下面将看到并不是任何一个等级结构都可以用调入比例控制不变的. 本段的目的是: 给定了内部转移比例矩阵 $Q = \{p_{ij}\}$ (由(2)式, 退出向量 w 也完全被确定), 研究哪些等级结构用合适的调入比例可以保持不变, 称为调入比例对等级结构的稳定控制.

根据方程(12), 对于某个等级结构 a , 如果存在调入比例 r 使得

$$a = a(Q + w^T r) \quad (13)$$

则称 a 为稳定结构, 注意这里的 r 必须满足基本关系(3)式, 即 $r_i \geq 0$, $\sum_{i=1}^k r_i = 1$. 由(13)式不难得到

$$r = \frac{a - aQ}{aw^T} \quad (14)$$

可以验证(14)式给出的 r 满足 $\sum_{i=1}^k r_i = 1$. 为满足 $r_i \geq 0$ 的要求, 可知稳定结构 a 的范围由

$$a \geq aQ \quad (15)$$

确定, 称为等级结构的稳定域. 下面举一个例子, 看看如何得到这种稳定域.

例 设大学教师的 3 个职称 (助教、讲师和教授) 依次记为等级 $i = 1, 2, 3$. 每年等级之间的转移比例矩阵为

$$Q = \begin{bmatrix} 0.5 & 0.4 & 0 \\ 0 & 0.6 & 0.3 \\ 0 & 0 & 0.8 \end{bmatrix} \quad (16)$$

这表示只有提升, 没有降级, 而且只能升一级, 如每年由助教升入讲师的比例为 40%. 求等级结构 a 的稳定域.

将(16)代入(15)式得

$$\begin{cases} a_1 \geq 0.5a_1 \\ a_2 \geq 0.4a_1 + 0.6a_2 \\ a_3 \geq 0.3a_2 + 0.8a_3 \end{cases}$$

即

$$\begin{cases} a_2 \geq a_1 \\ a_3 \geq 1.5a_2 \end{cases} \quad (17)$$

这就是 a 的稳定域. 我们先从几何上把(17)式给定的区域表示出来.

任何一个等级结构 $a = (a_1, a_2, a_3)$ 可看作三维空间的一个点, 并且位于第一象限的平面

$$a_1 + a_2 + a_3 = 1, \quad a_1, a_2, a_3 \geq 0 \quad (18)$$

上. 这是一个以 $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ 为顶点的等边三角形, 将它铺在平面

上,如图1,记作 A ,称可行域。

为了在 A 中找到(17)式所示的 a 的稳定域,画出 $a_2 = a_1, a_3 = 1.5a_2$ 两条直线,它们相交于 s_1 点,容易看出,不等式(17)界定的是以 s_1 和 $s_2(0, 0.4, 0.6)$, $s_3(0, 0, 1)$ 为顶点的三角形,记作 B , B 即稳定域,且可以算出 s_1 的坐标为 $(0.286, 0.286, 0.428)$ 。

在这个例子中,稳定域 B 是以可行域 A 的顶点 s_3 为一个顶点、以 A 的一条边的部分线段为一边的三角形,这是有代表性的。为进一步得到一般情况下稳定域的构造,我们需要从(14)式出发加以研究。

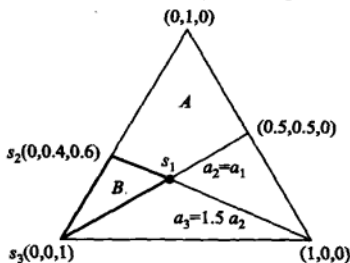


图1 教师等级结构的可行域和稳定域

稳定域的构造 设至少有一个 i 使 $\sum_{j=1}^k p_{ij} < 1$, 则 $I - Q$ 可逆, 记 $M = \{m_{ij}\} = (I - Q)^{-1}$, 则由(14)式可得

$$a = (aw^T) rM \quad (19)$$

记第 i 元素为1、其余元素为0的单位行向量为 e_i , r 可表示为 $r = \sum_{i=1}^k r_i e_i$. 又记 M 的第 i 行向量为 $m_i = (m_{i1}, m_{i2}, \dots, m_{ik})$, $\mu_i = \sum_{j=1}^k m_{ij}$, 则 $rM = \sum_{i=1}^k r_i e_i M = \sum_{i=1}^k r_i m_i$. 对向量表达式(19)的诸分量求和, 注意到其左端为1, 可以得到

$$aw^T = \left(\sum_{j=1}^k r_j \mu_j \right)^{-1} \quad (20)$$

$$a = \sum_{i=1}^k \frac{r_i m_i}{\sum_{j=1}^k r_j \mu_j} \quad (21)$$

为了得到便于应用的结果, 将(21)式进一步表示为

$$a = \sum_{i=1}^k b_i s_i, \quad b_i = \frac{r_i \mu_i}{\sum_{j=1}^k r_j \mu_j}, \quad s_i = \frac{m_i}{\mu_i} \quad (22)$$

对于 b_i , 容易知道 $\sum_{i=1}^k b_i = 1$; 又因为 M 的元素 m_{ij} 非负(为什么?), μ_i 非负, $\sum_{j=1}^k r_j \mu_j$ 非负(为什么?), 所以当且仅当 $r_i \geq 0$ 时, $b_i \geq 0$. 对于 $s_i = (s_{i1}, s_{i2}, \dots, s_{ik})$, 有 $s_{ij} = \frac{m_{ij}}{\mu_i}$, 所以 $s_{ij} \geq 0$, $\sum_{j=1}^k s_{ij} = 1$. 上述分析表明:

当且仅当 a 能够表示为以 b_i 为系数的 s_i 的线性组合 (22) 式, 且 b_i 满足

$$b_i \geq 0, \quad \sum_{i=1}^k b_i = 1 \quad (23)$$

时, a 是稳定结构, 即存在 r , 满足 $r_i \geq 0$, $\sum_{i=1}^k r_i = 1$, 使 (13) 式成立.

回到上面教师等级结构的例子, 先从 (16) 式算出

$$M = (I - Q)^{-1} = \begin{bmatrix} 2 & 2 & 3 \\ 0 & 2.5 & 3.75 \\ 0 & 0 & 5 \end{bmatrix}$$

于是 $\mu_1 = 7, \mu_2 = 6.25, \mu_3 = 5$. 然后由 (22) 式得到 $s_1 = (0.286, 0.286, 0.428)$, $s_2 = (0, 0.4, 0.6)$, $s_3 = (0, 0, 1)$. 稳定结构 a 可表示为 $a = b_1 s_1 + b_2 s_2 + b_3 s_3$, $b_1, b_2, b_3 \geq 0, b_1 + b_2 + b_3 = 1$, 这正是图 1 中的稳定域 B .

从计算过程可以看到, 当等级转移只有提升一级使 Q 的非零元素只能位于 (16) 式所示的位置时, M 必是上三角形矩阵, 从而必然有 $s_3 = (0, 0, 1)$, $s_2 = (0, *, *)$, $s_1 = (*, *, *)$ ($*$ 表示非零元素). 注意到 s_3 表示全部是教授的职称结构, s_2 表示全部是讲师或教授的结构, 由稳定域 B 的构造可知, 较高级职称所占比例较大的结构才是稳定的. 这是职称只升不降的必然结果.

上面这个 3 等级例题的结果可以推广到一般包括 k 个等级的情况.

第一, 等级结构 $a = (a_1, a_2, \dots, a_k)$ 的可行域 A 是 k 维空间中由下式决定的 $k-1$ 维超平面,

$$a_1 + a_2 + \dots + a_k = 1, \quad a_1, a_2, \dots, a_k \geq 0 \quad (24)$$

而稳定域 B 则是 A 中以 s_1, s_2, \dots, s_k 为顶点的凸域.

第二, 当系统内部只有提升一级的转移, 即转移比例矩阵具有

$$Q = \begin{bmatrix} p_{11} & p_{12} & & 0 \\ & p_{22} & \ddots & \\ & & \ddots & p_{k-1,k} \\ 0 & & & p_{kk} \end{bmatrix} \quad (25)$$

形式时, B 的顶点为 $s_k = (0, \dots, 0, 1)$, $s_{k-1} = (0, \dots, 0, *, *)$, \dots , $s_2 = (0, *, \dots, *)$, $s_1 = (*, *, \dots, *)$.

用调入比例进行动态调节 下面讨论我们最关心的问题. 设理想的等级结构是 a^* , 并合理地假设 $a^* \in B$, 已知转移比例矩阵 Q 和初始等级结构 $a(0)$, 求调入比例 r 使 $a(1)$ 达到或尽量接近 a^* . 如果没有达到 a^* , 接着可以将 $a(1)$ 作为新的 $a(0)$, 再求调入比例 r (不一定与上面的 r 相同), 使 $a(2)$ 达到或尽量接近 a^* . 如此继续下去直到某个 $a(t)$ 达到 a^* 或与之接近到满意的程度. 不妨将这个过

首先,要衡量 $a(1)$ 接近 a^* 的程度,必须定义两个等级结构之间的所谓距离. 对于 $a^{(1)} = (a_1^{(1)}, \dots, a_k^{(1)})$ 和 $a^{(2)} = (a_1^{(2)}, \dots, a_k^{(2)})$, 可以合理地定义二者之间的距离

$$D(a^{(1)}, a^{(2)}) = \sum_{i=1}^k \lambda_i (a_i^{(1)} - a_i^{(2)})^2 \quad (26)$$

其中 $\lambda_i (\geq 0)$ 为对等级 i 的加权因子.

我们的目的归结为求解问题 E_1 :

$$\begin{aligned} \min D(a(1), a^*) \\ \text{s. t. } a(1) = a(0)(Q + w^T r), \quad r_i \geq 0, \quad \sum_{i=1}^k r_i = 1 \end{aligned} \quad (27)$$

E_1 的一种解法如下.

$$a^* - a(1) = a^* - a(0)(Q + w^T r) = a(0)w^T \left[\frac{a^* - a(0)Q}{a(0)w^T} - r \right]$$

式中的因子 $a(0)w^T$ 为常数,若记

$$y = \frac{a^* - a(0)Q}{a(0)w^T}, \quad y = (y_1, y_2, \dots, y_k) \quad (28)$$

则 $a^* - a(1)$ 与 $y - r$ 成正比,其中 y_i 由已知数据算出,并且不难验证

$$\sum_{i=1}^k y_i = 1 \quad (29)$$

于是由(26)式问题 E_1 转化为问题 E_2 :

$$\begin{aligned} \min \sum_{i=1}^k \lambda_i (y_i - r_i)^2 \\ \text{s. t. } r_i \geq 0, \quad \sum_{i=1}^k r_i = 1 \end{aligned} \quad (30)$$

若 $y_i \geq 0, i = 1, 2, \dots, k$, 则 E_2 的解显然为

$$r_i = y_i, \quad i = 1, 2, \dots, k \quad (31)$$

否则,将 y_i 分为 $y_j > 0$ 和 $y_i \leq 0$ 两部分. 与 y 的分量(下标)相对应, r 也分为 r_j 和 r_i 两部分. 因为

$$\sum_{i=1}^k \lambda_i (y_i - r_i)^2 = \sum_j \lambda_j (y_j - r_j)^2 + \sum_i \lambda_i (|y_i| + r_i)^2 \quad (32)$$

右端第一项随 r_j 由 0 至 y_j 的增加而减少,第二项随 r_i 的增加而增加,又 $\sum_j y_j > 1$, 所以为使和式最小,应该有 $r_j \leq y_j$, 及

$$r_i = r_i^* = 0 \quad (33)$$

(用 r_i^* 表示最优值), 而 r_j 由问题 E_3 :

$$\min \sum_j \lambda_j (y_j - r_j)^2$$

$$\text{s. t. } 0 \leq r_j \leq y_j, \quad \sum_j r_j = 1 \quad (34)$$

解出. 用乘子法解这个条件极值问题, 令

$$G(r_j) = \sum_j \lambda_j (y_j - r_j)^2 + 2\alpha \sum_j r_j \quad (35)$$

解 $\frac{\partial G}{\partial r_j} = 0$, 可得

$$r_j = y_j - \frac{\alpha}{\lambda_j} \quad (36)$$

由 $\sum_j r_j = 1$ 可以算出

$$\alpha = \frac{\sum_j y_j - 1}{\sum_j \lambda_j^{-1}} \quad (37)$$

于是 E_3 的解为^①

$$r_j^* = \begin{cases} y_j - \frac{\alpha}{\lambda_j}, & y_j > \frac{\alpha}{\lambda_j} \\ 0, & y_j \leq \frac{\alpha}{\lambda_j} \end{cases} \quad (38)$$

用前面教师职称等级结构的例子, 说明用调入比例进行动态调节的方法. 设 Q 如(16)式所示, $a(0) = (0, 0, 1)$, $a^* = (0.286, 0.286, 0.428)$, 即图1中的 s_1 , 求 r 使 $a(1)$ 尽量接近 a^* .

在距离函数(26)中设权 $\lambda_i = 1 (i=1, 2, 3)$. 按照(28)式可以算出 $y_1 = 1.43$, $y_2 = 1.43$, $y_3 = -1.86$. 根据(33)式应令 $r_3^* = 0$. 由(37)式算出 $\alpha = 0.93$, 再由(38)式得 $r_1^* = 0.5$, $r_2^* = 0.5$, 即调入比例为 $r^* = (0.5, 0.5, 0)$. 还可以算出 $a(1) = a(0)(Q + w^T r^*) = (0.1, 0.1, 0.8)$, 这是由 $a(0)$ 出发一年后达到的最接近 a^* 的等级结构(在(26)式定义的距离的意义下).

记上面的 $r^* = r(1)$, 由 $a(1)$ 出发用同样的方法可以得到 $r^* = r(2)$ 和 $a(2)$, 如此下去的计算结果如表1.

$$a(0) = (0, 0, 1), a^* = (0.286, 0.286, 0.428)$$

由表1可以看出, $a(7)$ 已经相当接近 a^* .

评注 这个模型不仅能描述一个社会系统中的等级结构, 也可以研究不同部门间成员的迁移, 如在人才自由流动的情况下, 从商、从政、从工、从教的人员

① 如果有某个 j' 使 $y_{j'} \leq \frac{\alpha}{\lambda_{j'}}$, 则由(38)式有 $r_{j'}^* = 0$, 然后用 $\sum_{j \neq j'} r_j^* = 1$ 重新确定 α , 直到剩下的 j 全有

$y_j > \frac{\alpha}{\lambda_j}$, 从而得到 r_j^* . 这样的 j 至少有一个.

结构的变化,或电子、钢铁、机械、建筑及第三产业中劳动力构成的演变等。

表 1 调入比例 r 对 $a(t)$ 的调节

t	1	2	3	4	5	6	7
$r(t)$	0.5	0.639	0.747	0.827	0.883	0.922	0.949
	0.5	0.361	0.253	0.173	0.117	0.078	0.051
	0	0	0	0	0	0	0
$a(t)$	0.1	0.165	0.207	0.235	0.253	0.264	0.272
	0.1	0.165	0.207	0.235	0.253	0.264	0.272
	0.8	0.670	0.586	0.531	0.495	0.472	0.457

本节讨论的是最基本的、也是最简单的问题,即在总人数和内部转移比例不变的情况下,用调入比例来控制等级结构的变化。为了解决实际问题可以有各方面的推广,例如总人数按一定比例增长的情况(基本方程如(11)式);调入比例有上下界的情况;调入比例固定而用转移比例控制等级结构的问题,事实上许多社会系统是从外部向最低等级调入成员,如学生入学,军人入伍,新教师也基本上由助教当起,即可认为调入比例固定为 $r = (1, 0, \dots, 0)$ 。而转移比例,比如只考虑提升一级的比例,却是按照人们制订的政策可以改变的。用转移比例对等级结构进行控制的问题要比用调入比例复杂得多,因为调入比例是 k 维向量(实际上只有 $k-1$ 个分量可变),而转移比例,即使只许提升一级也有 $2k-1$ 个可以调节的量(如果退出比例 w 也可以改变)。控制变量多了,对等级结构的调节更为灵活,计算也更为复杂。上述各种推广问题的讨论可参阅[6]。

12.5 资 金 流 通

若干地区之间资金每年按一定比例相互流动,各个地区还有一部分资金流出这些地区,并且不再回来。银行为了使这些地区的资金分布趋向给定的稳定分布,计划每年向各地区投放或收回一定的资金。本节要建立一个模型描述各地区资金分布的变化规律,讨论在什么条件下可以趋近稳定分布,并确定银行应投放或收回多少资金^[64]。

这个问题与上节的等级结构有相似之处。地区间的资金流通可类比等级间的成员转移,资金流出这些地区可类比成员退出系统,而银行投放或收回资金相当于成员的调入。也有几点不同之处,一是进入各地区的资金可正(投放)可负(收回),而成员的调入比例不能为负;二是各地区资金总和每年是变化的,而上节是在系统总人数不变的假定下进行讨论的。

下面先建立资金分布的基本方程,再研究趋向稳定分布的问题.

设有 k 个地区,第 t 年地区 i 的资金为 $c_i(t)$, $i=1,2,\dots,k, t=0,1,2,\dots$, 每年从地区 i 流入地区 j 的资金的比例为 p_{ij} , 每年银行向地区 i 投放的资金为 d_i (当 d_i 为负时表示从地区 i 收回). 这些量满足 $c_i(t) \geq 0, p_{ij} \geq 0, \sum_{j=1}^k p_{ij} \leq 1$ (总有某些地区每年有一定比例资金流出该系统). 记 $c(t) = (c_1(t), c_2(t), \dots, c_k(t))$, $Q = \{p_{ij}\}$, $d = (d_1, d_2, \dots, d_k)$, 类似于上节的(6)式, 容易得到

$$c(t+1) = c(t)Q + d \quad (1)$$

经递推可得

$$c(t) = c(0)Q^t + d \sum_{i=0}^{t-1} Q^i \quad (2)$$

如果 k 个地区的资金视为系统的 k 个状态,并增加一个状态 0 表示资金流出这个系统,资金流通的无后效性表明可以用马氏链模型描述其变化过程. 暂不考虑资金投放,资金在 $k+1$ 个状态间的转移矩阵可表为

$$P = \begin{bmatrix} 1 & 0 \\ R & Q \end{bmatrix} \quad (3)$$

其中第 1 行对应于状态 0, 因为资金一旦流出系统, 就不再回来, 所以状态 0 是一个吸收状态, 不妨假定各地区均对应于非吸收状态, 并且从这些状态出发可以到达状态 0, 即形成一个吸收链 (12.1 节定义 2). 于是由 12.1 节转移矩阵 P 的标准形式和定理 4 可知, $I - Q$ 可逆, 且 $(I - Q)^{-1} = \sum_{i=0}^{\infty} Q^i$. 这隐含着 $Q^t \rightarrow 0 (t \rightarrow \infty)$. 这样对(2)式令 $t \rightarrow \infty$ 就有

$$c(\infty) = d(I - Q)^{-1} \quad (4)$$

设银行希望各地区资金趋向于稳定分布 c^* , 在(4)式中令 $c(\infty) = c^*$ 可以得到

$$d = c^*(I - Q) = c^* - c^*Q \quad (5)$$

这就是说, 对于给定的 c^* 和 Q , 由(5)式算出的 d 可使 $c(t) \rightarrow c^* (t \rightarrow \infty)$. 但是我们必须检查当(5)代入(2)式后得到的

$$c(t) = c(0)Q^t + (c^* - c^*Q) \sum_{i=0}^{t-1} Q^i \quad (6)$$

是否对于 $t=1,2,\dots$ 都有 $c(t) \geq 0$ (指每个 $c_i(t) \geq 0$).

分两种情况讨论上述问题.

1. 因为 $c(0) \geq 0, Q \geq 0$ (指每个元素不小于零), 所以由(6)式可知, 若

$$c^* \geq c^*Q \quad (7)$$

则对于任意的初始分布 $c(0)$ 都有 $c(t) \geq 0 (t=1,2,\dots)$. 这时由(5)式给出的 d

就是使 $c(t) \rightarrow c^* (t \rightarrow \infty)$ 的银行资金投放量,不妨称 c^* 是可达到的.

2. (7)式只是 $c(t) \geq 0$ 的充分条件. 当(7)式不成立时可以进一步将(6)式化为

$$\begin{aligned} c(t) &= c(0)Q^t + c^*(I - Q)(I + Q + \cdots + Q^{t-1}) \\ &= c(0)Q^t + c^*(I - Q^t) \\ &= c^* - [c^* - c(0)]Q^t \end{aligned} \quad (8)$$

记

$$h(t) = [c^* - c(0)]Q^t \quad (9)$$

由(8)式可得 $c(t) \geq 0$ 的充要条件为 E_t :

$$c^* \geq h(t), \quad t = 1, 2, \dots \quad (10)$$

条件 E_t 可以方便地用来检验 c^* 不能达到,因为只要存在一个 t ,使 E_t 不满足即可;但是无法判断 c^* 可以达到,因为我们不能对所有的 $t = 1, 2, \dots$ 都来验证 E_t 的正确性.

有一个判断 E_t 成立的充分条件,可以与条件 E_t 结合起来使用. 下面不加证明地引述一个定理.

定理 设 $c^* > 0$, $h(s)$ 由(9)式定义,记^①

$$\bar{h}(s) = \sum_{i=1}^k |h_i(s)|, \quad h(s) = (h_1(s), \dots, h_k(s)) \quad (11)$$

1) 若存在某个 $s(s = 0, 1, 2, \dots)$ 使条件 F_s :

$$\min_i c_i^* \geq \bar{h}(s) \quad (12)$$

成立,则条件 E_t 对 $t \geq s$ 均成立.

2) 必存在某个 s_0 使条件 F_{s_0} 成立.

根据以上的分析(7)~(12)式,对于给定的 c^* , Q 和 $c(0)$,判断 c^* 能否达到的程序如图1所示.

按照图中箭头方向先检验(7)式是否成立,然后交替检验条件 $F_s (s = 0, 1, 2, \dots)$ 和条件 $E_t (t = 1, 2, \dots)$. 一旦 F_s 成立,则 c^* 可达到;一旦 E_t 不成立,则 c^* 不可达到. 而根据上述定理,判断程序不会无限地进行下去.

例 3个地区的资金流通比例矩阵为

$$Q = \begin{bmatrix} \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

^① 实际应用中常能满足 c_i^* 不为零的要求.

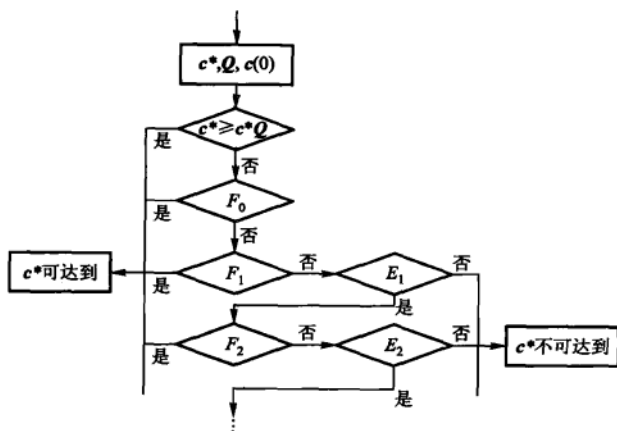


图1 判断 c^* 是否可达到程序示意图

初始分布为 $c(0) = (9, 3, 6)$. 判断稳定分布 $c^* = (12, 6, 3)$ 能否达到; 若能达到, 问银行每年应如何投放资金.

按照图1的程序, 步骤如下:

1. 检验 $c^* \geq c^*Q$. 计算 $c^*Q = (6, 4, 7)$, $c^* \geq c^*Q$ 不成立.

2. 检验 F_0 . 计算 $h(0) = c^* - c(0) = (3, 3, -3)$, $\bar{h}(0) = 9$, 而 $\min_i c_i^* = 3$, $F_0: \min_i c_i^* \geq \bar{h}(0)$ 不成立.

3. 检验 F_1 . 计算 $h(1) = h(0)Q = (2, -1, 1)$, $\bar{h}(1) = 4$, $F_1: \min_i c_i^* \geq \bar{h}(1)$ 不成立.

4. 检验 E_1 . $c^* \geq h(1)$ 成立.

5. 检验 F_2 . 计算 $h(2) = h(1)Q = \left(\frac{1}{3}, \frac{1}{3}, \frac{2}{3}\right)$, $\bar{h}(2) = \frac{4}{3}$, $F_2: \min_i c_i^* \geq \bar{h}(2)$ 成立.

检验完毕, c^* 可达到. 银行应投放的资金 d 由(5)式计算, 有

$$d = c^* - c^*Q = (6, 2, -4)$$

即每年向地区1, 2分别投放6和2个资金单位, 从地区3收回4个资金单位.

习 题

1. 在12.2节中将钢琴销售的存贮策略修改为: 当周末库存量为0或1时, 订购, 使下周初的库存达到3架; 否则, 不订购. 建立马氏链模型, 计算稳态下失去销售机会的概率和每周

的平均销售量.

2. 在 12.2 节中将钢琴销售的存贮策略修改为:当周末库存量为 0 时,订购量为本周销售量加 2 架;否则,不订购. 建立马氏链模型,计算稳态下失去销售机会的概率和每周的平均销售量.

3. 对于 12.3 节中的近亲繁殖模型,推导转移概率矩阵 $P((6)$ 式)的第 4,5 行.

4. 在基因遗传过程中,考虑 3 种基因类型:优种 $D(dd)$,混种 $H(dr)$ 和劣种 $R(rr)$. 对于任意的个体,每次用一混种与之交配,所得后代仍用混种交配,如此继续下去. 构造马氏链模型,说明它是正则链,求稳态概率及由优种和混种出发的首次返回平均转移次数. 如果改为每次用优种交配,再构造马氏链模型,说明它是吸收链,求由混种和劣种出发变为优种的平均转移次数.

5. 色盲具有遗传性,由两种基因 c 和 s 的遗传规律决定. 男性只有一个基因 c 或 s ;女性有两个基因 cc, cs 或 ss . 当某人具有基因 c 或 cc 时则呈色盲表征. 基因遗传关系是:男孩等概率地继承母亲两个基因中的一个;女孩继承父亲的那个基因,并等概率地继承母亲的一个基因. 由此可以看出,当母亲是色盲时男孩一定色盲,女孩却不一定. 用马氏链模型研究非常极端的近亲结婚情况下的色盲遗传,即同一对父母的后代婚配. 父母的基因组合共有 6 种类型,形成马氏链模型的 6 种状态,问哪些是吸收状态. 若父亲非色盲而母亲为色盲,问平均经过多少代其后代就会变成全为色盲或全不为色盲的状态,变成这两种状态的概率各为多大?

6. 两种不同的外部表征是由两种不同基因决定的,这两种基因的遗传关系是相互独立的. 例如猪的毛有颜色表征(黑和白)与质地表征(粗和光). 对于每一种表征仍分为优种 $D(dd)$,混种 $H(dr)$ 和劣种 $R(rr)$ 3 种基因类型,两种表征的组合则有 9 种基因类型. 在完全优势遗传中,优种和混种的猪毛颜色黑、质地粗,劣种则颜色白、质地光,这样共有 4 种外部表征组合,即黑粗、黑光、白粗、白光. 假设群体的两种外部表征对应的基因中 d 和 r 的比例相同(即均为 $1/2$),在随机交配情况下构造马氏链模型. 证明在稳定情况下上述 4 种外部表征组合的比例为 $9:3:3:1$.

7. 设等级结构的转移比例矩阵 Q 仍由 12.4 节(16)式给出,理想的结构为 $a^* = (0.2, 0.3, 0.5)$,证明 $a^* \in B(\text{稳定域})$,若初始结构为 $a(0) = (0.2, 0.8, 0)$,用 12.4 节介绍的问题 E_1, E_2, E_3 的解法求调入比例 r ,使 $a(1)$ 尽量接近 a^* .

8. 如果在等级结构模型中将距离函数的定义由 12.4 节(26)式改为 $D(a^{(1)}, a^{(2)}) = \sum_{i=1}^k \lambda_i |a_i^{(1)} - a_i^{(2)}|$,试给出求解问题 E_1 的方法.

9. 一个服务网络由 k 个工作站 v_1, v_2, \dots, v_k 依次串接而成,当某种服务请求到达工作站 v_i 时, v_i 能够处理的概率为 p_i ,转往下一站 v_{i+1} 处理的概率为 $q_i (i=1, 2, \dots, k-1, \text{设 } q_k=0)$,拒绝处理的概率为 r_i ,满足 $p_i + q_i + r_i = 1$. 试构造马氏链模型,确定到达 v_i 的请求平均经过多少工作站才能获得接受处理或拒绝处理的结果,被接受和拒绝的概率各多大?

10. 用马氏链模型讨论空气污染问题. 有 k 个城市 v_1, v_2, \dots, v_k ,每一时刻 $t=0, 1, 2, \dots, v_i$ 的空气中污染物浓度为 $c_i(t)$,从 t 到 $t+1, v_i$ 空气中污染物扩散到 v_j 去的比例是 p_{ij} ,有 $\sum_{j=1}^k p_{ij} \leq 1 (i=1, 2, \dots, k)$,而扩散到 k 个城市之外的那部分污染物永远不再回来. 在每个时刻各城市的污染源都排出一定的污染物,记 v_i 排出的为 d_i . 按照环境管理条例要求,对充分大的 t 必须

$c_i(t) \leq c_i^*$. 试建立马氏链模型, 在已知 p_{ij} 和 c_i^* 的条件下确定 d_i 的限制范围, 满足管理

条例的要求. 设 $k=3$, p_{ij} 由矩阵 $Q = \begin{bmatrix} 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 0 & 2/3 & 1/3 \end{bmatrix}$ 给出, $c_i^* = 25 (i=1, 2, 3)$, 求 d_i 的

范围^[64].



第13章

动态优化模型

本章讨论动态优化问题,其优化目标仍然是一个数值,而最优策略是函数.对于连续过程可归结为求泛函的极值,常用的方法是古典变分法和最优控制论,后者要用到现代控制理论的内容,本书无意涉及,这里选择的案例都是可以用变分法解决的.对于离散过程我们选择了用动态规划模型求解的例子.

13.1 速降线与短程线

速降线与短程线都是17世纪末提出的问题,这里讨论它们的目的是介绍变分法的基本概念,并给出以下几节要用的主要结果^[74].

速降线问题 铅直平面内给定不在一条垂直线上的两个点 A, B ,如图1,求连接它们的光滑曲线,使质点在重力作用下沿该曲线以最短时间从 A 点滑到 B 点(摩擦力不计).

尽管 A, B 两点间的最短距离是连接它们的直线,但是沿直线运动时速度增长较慢,如果沿一条陡峭的曲线下滑,虽路径加长,而运动速度增长很快.

为了求这条运动时间最短的曲线,在图1中将 A 取为坐标原点 $(0,0)$, B 点坐标为 (x_1, y_1) ,连接 A, B 的曲线记为 $y(x)$,于是曲线上的弧长为 $ds = \sqrt{1+y'^2}dx$.根据能量守恒定律,质点在曲线 $y(x)$ 上任一点的速度 $\frac{ds}{dt}$ 满足 $\frac{1}{2}m\left(\frac{ds}{dt}\right)^2 = mgy$,其中 m 是质点的质量, g 是重力加速度.将上面 ds 的关系代入,得到 $dt = \sqrt{\frac{1+y'^2}{2gy}}dx$,于是质点沿曲线 $y(x)$ 从 A 点滑到 B 点的时间可表为

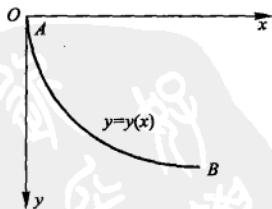


图1 最速降线

$$J(y(x)) = \int_0^{x_1} \sqrt{\frac{1+y'^2}{2gy}} dx \quad (1)$$

$y(x)$ 在 A, B 两个端点应有

$$y(0) = 0, \quad y(x_1) = y_1 \quad (2)$$

速降线问题归结为求 $y(x)$, 在满足(2)的条件下, 使(1)的 $J(y(x))$ 达到最小.

短程线问题 给定任意曲面上的两个点 A, B , 如图 2, 求连接它们的长度最短的光滑曲线. 地球近似于一个椭球体, 由甲地飞往乙地的最短航线是椭球表面上连接甲乙两地的短程线. 由于北极上空对民航的开放, 从北京飞往北美的航线比原来需要飞越太平洋时缩短了许多, 就是因为可以采用接近于短程线的航线.

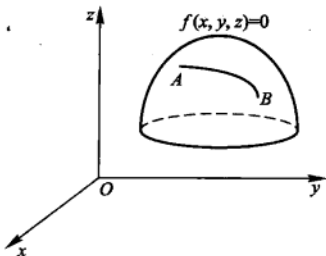


图 2 短程线

这个问题数学上可表述如下: 给定曲面方程 $f(x, y, z) = 0$, 已知曲面上两个点的坐标为 $A(x_0, y_0, z_0), B(x_1, y_1, z_1)$, 在曲面上求 A, B 两点的曲线 $y = y(x), z = z(x)$, 使该曲线的长度最短.

因为曲线的弧长为 $ds = \sqrt{1 + y'^2 + z'^2} dx$, 所以曲线的长度是

$$J(y(x), z(x)) = \int_{x_0}^{x_1} \sqrt{1 + y'^2 + z'^2} dx \quad (3)$$

短程线问题归结为在曲面 $f(x, y, z) = 0$ 上求曲线 $y = y(x), z = z(x)$, 即满足

$$f(x, y(x), z(x)) = 0 \quad (4)$$

条件下, 使(3)的 $J(y(x), z(x))$ 达到最小.

以上两个例子都是求函数 $y(x)$ (或 $y(x), z(x)$), 使由它及其导数决定的某一实数 J 最小, 这类问题称为泛函的极值, 以下简单介绍它的概念与解法. 为了以下几节应用的方便, 将自变量记作 t , 函数记作 $x(t), y(t)$.

泛函、泛函的变分和极值 为了简明地叙述泛函、泛函的变分和极值的概念, 我们将它与函数、函数的微分和极值作如下比较:

函数、函数的微分和极值	泛函、泛函的变分和极值
1. 若对于变量 t 在某域中的任一个值, 都有变量 y 的一个值与之对应, 则称 y 是 t 的函数, 记作 $y = f(t)$.	1. 若对于某类函数集合中的每一个函数 $x(t)$, 都有变量 J 的一个值与之对应, 则称 J 是 $x(t)$ 的泛函, 记作 $J(x(t))$.
2. 自变量 t 在 t_0 的增量记作 $\Delta t = t - t_0$, 且微分 $dt = \Delta t$.	2. 函数 $x(t)$ 在 $x_0(t)$ 的增量记作 $\delta x(t) = x(t) - x_0(t)$, $\delta x(t)$ 又称 $x(t)$ 的变分.

函数、函数的微分和极值	泛函、泛函的变分和极值
3. 函数 y 在 t_0 的增量记作 $\Delta f = f(t_0 + \Delta t) - f(t_0)$, Δf 的线性主部是函数的微分, 记作 dy , $dy = f'(t_0) dt$.	3. 泛函 $J(x(t))$ 在 $x_0(t)$ 的增量记作 $\Delta J = J(x_0(t) + \delta x(t)) - J(x_0(t))$, ΔJ 的线性主部称泛函的变分, 记作 $\delta J(x_0(t))$.
4. 若函数 y 在域内 t 点达到极值, 则在 t 点的微分 $dy(t) = 0$.	4. 若泛函 $J(x(t))$ 在函数集合内部的 $x(t)$ 达到极值, 则在 $x(t)$ 的变分 $\delta J(x(t)) = 0$.
5. 函数 y 在 t 的微分的另一表达式为 $dy(t) = \frac{\partial}{\partial \alpha} f(t + \alpha \Delta t) _{\alpha=0}$	5. 泛函 $J(x(t))$ 在 $x(t)$ 的变分可以表为 ^[74, 83] $\delta J(x(t)) = \frac{\partial}{\partial \alpha} J(x(t) + \alpha \delta x(t)) _{\alpha=0}$

由此得到泛函 $J(x(t))$ 在 $x(t)$ 达到极值的必要条件为

$$\frac{\partial}{\partial \alpha} J(x(t) + \alpha \delta x(t)) |_{\alpha=0} = 0 \quad (5)$$

最简泛函极值的必要条件——欧拉方程 最简单的一类泛函表为

$$J(x(t)) = \int_{t_1}^{t_2} F(t, x(t), \dot{x}(t)) dt \quad (6)$$

其中 F 具有二阶连续偏导数, $x(t)$ 为二阶可微函数, 且满足

$$x(t_1) = x_1, \quad x(t_2) = x_2 \quad (7)$$

称(6), (7)为固定端点条件下的泛函, $J(x(t))$ 在 $x(t)$ 达到极值的必要条件为: $x(t)$ 是二阶微分方程

$$F_x - \frac{d}{dt} F_{\dot{x}} = 0 \quad (8)$$

或

$$F_x - F_{\dot{x}\dot{x}} - F_{\dot{x}x} \dot{x} - F_{xx} \ddot{x} = 0 \quad (9)$$

的通解, 其中两个任意常数由端点条件(7)式确定^[74, 83]. (8), (9)称欧拉方程.

用欧拉方程求解速率降线问题, 考察(1)式知

$$F(\gamma, \gamma') = \sqrt{\frac{1 + \gamma'^2}{\gamma}}$$

F 不显含自变量 x (相当于(6)中的 t), 注意到这里的 γ 相当于(6)中的 x , 欧拉方程(9)化为

$$F_\gamma - F_{\gamma\gamma'} \gamma' - F_{\gamma'\gamma'} \gamma'' = 0$$

等价于

$$\frac{d}{dy}(F - y'F_{y'}) = 0$$

作一次积分得

$$F - y'F_{y'} = c$$

c 为任意常数. 将 F 代入上式得

$$\sqrt{\frac{1+y'^2}{y}} - \frac{y'^2}{\sqrt{y(1+y'^2)}} = c$$

即

$$y(1+y'^2) = \frac{1}{c^2}$$

化为参数方程可表为

$$\begin{cases} x = c_1(t - \sin t) + c_2 \\ y = c_1(1 - \cos t) \end{cases} \quad (10)$$

这是摆线方程, 利用端点条件(2), 由 $y(0) = 0$ 可知 $c_2 = 0$, c_1 由 $y(x_1) = y_1$ 确定.

虽然欧拉方程的解(10)只是泛函(1)达到极值的必要条件, 但是正如求解函数极值一样, 对于实际问题通常不用检验它是否满足泛函达到极小值的充分条件. 因为极小值曲线(即最速降线)一定存在, 它满足欧拉方程, 所以解(10)就是极小值曲线(极大值曲线不存在).

泛函极值的必要条件可以推广到含有两个及两个以上未知函数的情况, 如

$$J(x(t), u(t)) = \int_{t_1}^{t_2} F(t, x, \dot{x}, u, \dot{u}) dt \quad (11)$$

的欧拉方程是

$$\begin{cases} F_x - \frac{d}{dt}F_{\dot{x}} = 0 \\ F_u - \frac{d}{dt}F_{\dot{u}} = 0 \end{cases} \quad (12)$$

横截条件 如果容许函数 $x(t)$ 的一个端点如 $t = t_2$ 不固定, 而是在一条给定的曲线 $x = \psi(t)$ 上变动, 如图 3, 于是端点条件表示为

$$x(t_1) = x_1, \quad x(t_2) = \psi(t_2) \quad (t_2 \text{ 可变}) \quad (13)$$

可以导出, 泛函在可变端点下的极值条件为

$$[F + (\dot{\psi} - \dot{x})F_{\dot{x}}] \big|_{t=t_2} = 0 \quad (14)$$

这是确定欧拉方程通解中另一常数的定解条

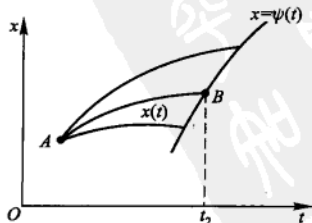


图 3 右端点在 $x = \psi(t)$ 上

件,称为横截条件.

横截条件有两种常用的特殊情况:

(1) 当 $x = \psi(t)$ 垂直于横轴时, t_2 固定但 $x(t_2)$ 自由, 称 t_2 为自由端点, 可得

$$F_x \big|_{t=t_2} = 0 \quad (15)$$

这是自由端点的定解条件.

(2) 当 $x = \psi(t)$ 平行于横轴时,

$$(F - \dot{x}F_x) \big|_{t=t_2} = 0 \quad (16)$$

条件极值 Hamilton 函数 前面讨论的是无条件泛函极值问题. 最优控制涉及的通常是条件极值, 典型形式为

$$J(u(t)) = \int_{t_1}^{t_2} F(t, x(t), u(t)) dt \quad (17)$$

$$\dot{x}(t) = f(t, x(t), u(t)) \quad (18)$$

其中 $u(t)$ 是控制函数, $x(t)$ 是状态函数或称轨线. 问题的提法是求 $u(t) \in U$ 使泛函(17)在条件(18)下达到极值, 并且 $x(t) \in S$. 这里 U 和 S 分别是 $u(t)$ 和 $x(t)$ 的容许集合.

用拉格朗日乘子法化条件极值为无条件极值问题. 引入乘子函数 $\lambda(t)$ 构造泛函

$$I(x(t), u(t)) = \int_{t_1}^{t_2} [F(t, x, u) + \lambda(t)(f(t, x, u) - \dot{x})] dt \quad (19)$$

记

$$H(t, x, u) = F(t, x, u) + \lambda(t)f(t, x, u) \quad (20)$$

H 称 Hamilton 函数, (19)式可写作

$$I(x(t), u(t)) = \int_{t_1}^{t_2} (H - \lambda \dot{x}) dt \quad (21)$$

条件极值(17), (18)与含有两个函数 $x(t), u(t)$ 的无条件极值问题(21)等价. 根据(12)式, (21)的欧拉方程为

$$\begin{cases} (H - \lambda \dot{x})_x - \frac{d}{dt}(H - \lambda \dot{x})_x = 0 \\ (H - \lambda \dot{x})_u - \frac{d}{dt}(H - \lambda \dot{x})_u = 0 \end{cases} \quad (22)$$

将(20)给出的 H 的表达式代入(22)可得

$$\begin{cases} \frac{\partial H}{\partial x} + \dot{\lambda}(t) = 0 \\ \frac{\partial H}{\partial u} = 0 \end{cases}$$

于是, 条件极值问题归结为求解未知函数 $x(t), u(t), \lambda(t)$ 的微分方程组

$$\begin{cases} \dot{\lambda}(t) = -\frac{\partial H}{\partial x} \\ \frac{\partial H}{\partial u} = 0 \\ \dot{x} = f(t, x, u) \end{cases} \quad (23)$$

最优控制函数 $u(t)$ 和最优状态函数(最优轨线) $x(t)$ 由方程组(23)在 $x(t)$ 的端点条件下解出。

13.2 生产计划的制订

工厂与客户签订了一项在某时刻提交一定数量产品的合同,在制订生产计划时要考虑生产和贮存两种费用.生产费用通常取决于生产率(单位时间的产量),生产率越高费用越大;贮存费用自然由已经生产出来的产品数量决定,数量越多费用越大.所谓生产计划这里简单地看做是到每一时刻为止的累积产量,它与每单位时间(如每天)的产量可以互相推算.建模目的是寻求最优的生产计划,使完成合同所需的总费用(生产与贮存费用之和)最小^[41].

假设 开始生产时刻记为 $t=0$,按照合同应在 $t=T$ 提交数量为 Q 的产品.到时刻 t 为止的产量记作 $x(t)$, $x(t)$ 即生产计划.因为时刻 t 的生产率表示为 $\dot{x}(t)$,所以单位时间的生产费用可以一般地记作 $f(\dot{x}(t))$,而单位时间的贮存费用则应记为 $g(x(t))$.于是从 $t=0$ 到 $t=T$ 的总费用 $C(x(t))$ 是

$$C(x(t)) = \int_0^T [f(\dot{x}(t)) + g(x(t))] dt \quad (1)$$

为了确定 f 和 g 的具体形式作如下假设:

1. 单位时间内生产率提高一个单位所需的生产费用与这时的生产率成正比.在需求饱满、生产率很高的工厂里这个假设是合理的.

2. 贮存费与贮存量(即累积产量)成正比.这是关于贮存费的最常用的假设.

假设 1 表明,生产费 f 对生产率 \dot{x} 的变化率与 \dot{x} 成正比,即

$$\frac{df}{d\dot{x}} \propto \dot{x}$$

于是

$$f(\dot{x}(t)) = k_1 \dot{x}^2(t) \quad (2)$$

k_1 是比例系数.由假设 2 则可以直接写出

$$g(x(t)) = k_2 x(t) \quad (3)$$

k_2 是单位时间单位数量产品的贮存费.

建模 将(2),(3)代入(1)式并注意到 $x(t)$ 在 $t=0$ 和 $t=T$ 时的值,我们有

$$C(x(t)) = \int_0^T [k_1 \dot{x}^2(t) + k_2 x(t)] dt \quad (4)$$

$$x(0) = 0, \quad x(T) = Q \quad (5)$$

制订最优生产计划归结为在固定端点条件(5)下,求 $x(t)$ 使(4)式定义的泛函 $C(x(t))$ 取得最小值。

用变分法求解。记 $F(t, x, \dot{x}) = k_1 \dot{x}^2 + k_2 x$, 根据 13.1 节的欧拉方程(8)得

$$F_x(t, x, \dot{x}) - \frac{d}{dt} F_{\dot{x}}(t, x, \dot{x}) = 0$$

可得关于 $x(t)$ 的二阶微分方程

$$k_2 - 2k_1 \ddot{x}(t) = 0 \quad (6)$$

方程(6)在端点条件(5)下的解为

$$x(t) = \frac{k_2}{4k_1} t^2 + \frac{4k_1 Q - k_2 T^2}{4k_1 T} t \quad (7)$$

这就是使总费用 $C(x(t))$ 达到最小的生产计划。

由(7)式不难画出 $x(t)$ 的示意图(图1), 显然, $x(0) = 0, x(T) = Q$, 所以它是过 $(0, 0)$ 和 (T, Q) 两点的抛物线, 且因 $\ddot{x}(t) > 0$ 而呈下凸状。随着参数 k_1, k_2, T, Q 的不同, 曲线 $x(t)$ 可能有 S_1 和 S_2 两种形状。但是对于生产计划 $x(t)$ 应该有明显的限制条件

$$x(t) \geq 0, \quad 0 \leq t \leq T \quad (8)$$

这就是说, 只有当 $x(t)$ 呈 S_1 形状时才有实际意义。

容易看出, 对于(7)式表示的 $x(t)$, 条件(8)等价于

$$\dot{x}(0) \geq 0 \quad (9)$$

由(7)式算出 $\dot{x}(0)$, 可知(9)式又表示为

$$Q \geq \frac{k_2 T^2}{4k_1} \quad (10)$$

于是仅当(10)式成立时(7)式确定的 $x(t)$ 才是最优生产计划。当 k_1, k_2 固定时条件(10)表明, 在一定交货期 T 内要完成的产量 Q 相当大, 需要从 $t=0$ 就开始生产, 如图中曲线 S_1 。

但是, 当

$$Q < \frac{k_2 T^2}{4k_1} \quad (11)$$

即在 T 内要完成的产量 Q 较小时最优生产计划是什么呢?

直观的想法是为了节省贮存费用, 到 $t=t_1$ 才开始生产, 如图2的曲线 S_3 所

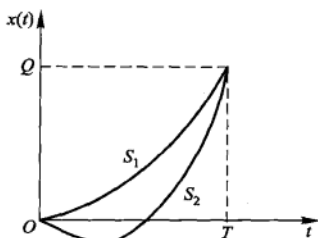


图1 $x(t)$ 的两种形式

示. S_3 是否就是如(7)式所示、图1中曲线 S_2 在 $x \geq 0$ 的那一部分呢? 如果不是, 时刻 t_1 和曲线 S_3 又如何确定(习题1).

解释 为了对最优生产计划作出解释, 考察它满足的方程(6)式, (6)式可以表示为

$$k_2 = \frac{d}{dt} \left(\frac{d}{dx} (k_1 \dot{x}^2) \right) = \frac{d}{dt} \left(\frac{df}{d\dot{x}} \right) \quad (12)$$

式中 $\frac{df}{d\dot{x}}$ 是单位时间内生产率提高一个单位所

需的生产费用, 经济理论中称为边际成本. 而 k_2 (单位时间单位数量产品的贮存费) 称为边

际贮存. 于是(12)式表明, 使边际成本的变化率等于边际贮存的生产计划是最优的.

评注 优化模型通常包括目标函数和约束条件两部分. 在这个模型中目标函数只考虑了两种最基本的费用, 并对它们作了相当简化的假设. 至于约束条件, 由于我们要用古典变分法求解, 所以 $x(t)$ 除了要满足端点条件(5)以外, 还需假定它是二阶可微函数. 这个条件并不影响(7)式的最优意义, 因为一般说来不会存在一个不满足二阶可微条件的、比(7)式更优的解. 但是在另一些约束条件下问题就较难处理了. 如这个模型应该要求 $x(t) \geq 0$, 我们看到当参数满足条件(11)时最优解已经需要仔细考虑. 若还要对生产率加以限制, 比如规定一个范围即 $A \leq \dot{x}(t) \leq B$, 则问题的求解更加困难. 实际上, 对控制函数施加的这类闭集约束, 可能导致古典变分法的失败.

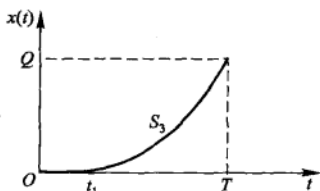


图2 条件(11)下的 $x(t)$

13.3 国民收入的增长

国民经济收入主要用于两个方面: 扩大再生产的积累资金和满足人民生活需要的消费资金. 如何安排积累和消费资金的比例使国民收入得到最快的增长, 是一个重大的理论和实践问题, 本节仅从最优控制的角度介绍一个十分简化的模型.

一般模型 将时刻 t 的国民经济收入记作 $x(t)$, 其中用于积累资金的部分记作 $y(t)$, 积累资金在国民经济收入中占的比例 $u(t) = \frac{y(t)}{x(t)}$ 称为积累率. 建模的目的是寻求最优积累率 $u(t)$ 使国民收入 $x(t)$ 增长最快. 下面给出模型的一般形式.

国民收入的增长率 $\dot{x}(t)$ 取决于当时的收入总值 $x(t)$ 和积累率 $u(t)$, 可以表示为

$$\dot{x}(t) = f(t, x, u) \quad (1)$$

考虑一段时间 T (一个五年或十年计划), 使 $x(t)$ 从初值 x_0 达到尽可能大的 $x(T)$, 即

$$x(0) = x_0 \quad (2)$$

$$\max x(T) \quad (3)$$

问题归结为在条件(1), (2)下求 $u(t)$ 满足(3). 而这又等价于它的对偶问题: 在固定端点条件 $x(T) = x_1$ 下使 T 最小. 于是若记

$$x(0) = x_0, \quad x(T) = x_1 \quad (4)$$

$$J(u(t)) = \int_0^T dt \quad (5)$$

则模型的一般提法是在条件(1), (4)下求 $u(t)$ 使(5)式的 J 达到最小. 这是我们熟悉的泛函极值问题的形式.

按照泛函的条件极值问题的解法(见 13.1 节)构造 Hamilton 函数

$$H = 1 + \lambda f(t, x, u) \quad (6)$$

并得到方程组

$$\begin{cases} \dot{\lambda}(t) = -\lambda f_x(t, x, u) \\ \lambda f_u(t, x, u) = 0 \\ \dot{x}(t) = f(t, x, u) \\ x(0) = x_0, \quad x(T) = x_1 \end{cases} \quad (7)$$

由此求解最优控制函数 $u(t)$ 和最优状态(即最优轨线) $x(t)$.

简化模型 为了确定函数 f 具体的、简化的形式我们粗略地考虑这样的经济规律: 当积累率 u 较小时, 国民收入的相对增长率 $\frac{\dot{x}}{x}$ 随着 u 的增加而增加, 这是因为积累资金的扩大再生产起着促进作用; 随着 u 的变大 $\frac{\dot{x}}{x}$ 增加得越来越慢; 当 u 增到一定程度后 $\frac{\dot{x}}{x}$ 反而会减小, 这是由于消费资金比例太小, 各种制约因素导致国民收入下降. 描述这种规律的最简单的模型是

$$\frac{\dot{x}}{x} = u(a - bu) \quad (8)$$

其中 a, b 为常数, 可以根据统计数据或经验估计.

比较(1), (8)式得到 $f = u(a - bu)x$, 代入方程(7)得

$$\begin{cases} \dot{\lambda} = -\lambda u(a - bu) \\ \lambda(a - 2bu)x = 0 \\ \dot{x} = u(a - bu)x \\ x(0) = x_0, \quad x(T) = x_1 \end{cases} \quad (9)$$

方程(9)的解为

$$u(t) = \frac{a}{2b} \quad (10)$$

$$x(t) = x_0 e^{\frac{a^2}{4b}t} \quad (11)$$

即最优积累率是常数 $a/2b$, 这时国民收入以指数形式(11)增长, 用最短时间

$$T = \frac{4b}{a^2} \ln \frac{x_1}{x_0} \quad (12)$$

从 x_0 上升到 x_1 .

评注 一般情况下国民收入依赖于积累率的最优增长是典型的快速控制问题(1), (4), (5), 最优解要从非线性方程组(7)得到. 而在我们对国民收入的简化假设(8)下, 得到的是平凡解——常数 $a/2b$. 实际上, 对于这种简单情况不必求解泛函极值问题. 由(8)式可以直接看出, 当 $u = a/2b$ 时 \dot{x} 最大, 所以 $u(t) = a/2b$ 必定是模型的最优控制函数.

13.4 渔船出海

这一节继续讨论开发渔业资源的最大经济效益模型, 与 7.1 节的模型不同的是, 这里用出海渔船的数量作为控制函数. 实际上, 捕鱼业的具体作法是等渔场中鱼量增长到相当大以后, 才派出一定数量的渔船进行捕捞. 于是我们的控制函数可以取与这种作法相应的特殊形式, 从而将本来属于动态优化模型的泛函极值问题简化为普通的函数极值问题^[55].

模型假设 1. 渔场鱼量 $x(t)$ 的自然增长服从 logistic 规律, 单位时间捕捞量 h 与渔船数量 $u(t)$ 和渔场鱼量 $x(t)$ 成正比, 在捕捞条件下满足

$$\dot{x}(t) = f(x) - h(u, x) \quad (1)$$

$$f(x) = rx \left(1 - \frac{x}{N} \right) \quad (2)$$

$$h(u, x) = qu(t)x(t) \quad (3)$$

r, N 同前, q 是每只渔船单位时间(如每天)的捕捞率(相对于 x 而言). $u(t)$ 视为连续变量, 非整数部分理解为在部分时间内进行捕捞.

2. 初始时刻渔场鱼量

$$x(0) = \frac{N}{K}, \quad K \gg 1 \quad (4)$$

$x(0)$ 很小. 在时间 $0 \leq t \leq \tau$ 内不派渔船出海. $t > \tau$ 以后出海渔船的数量保持常数 U , 即 $u(t)$ 的形式为

$$u(t) = \begin{cases} 0, & 0 \leq t \leq \tau \\ U, & t > \tau \end{cases} \quad (5)$$

而 τ, U 为待定参数. 捕捞期间 ($t > \tau$) 渔场鱼量 x 保持稳定.

3. 鱼的出售单价为 p , 每只渔船单位时间(天)的费用为 c . 通货膨胀率或称折扣因子为 δ .

建模与求解 在假设 1, 3 下, 单位时间的利润(折合到初始时刻)为 $e^{-\delta t}(ph - cu)$, 模型的目标函数应是以 $u(t)$ 为控制函数的长期效益, 即归纳为如下的泛函极值问题:

$$\begin{aligned} J(u(t)) &= \int_0^{\infty} e^{-\delta t} [ph(u(t), x(t)) - cu(t)] dt \\ &= \int_0^{\infty} e^{-\delta t} [pqx(t) - c] u(t) dt \end{aligned} \quad (6)$$

$$\dot{x}(t) = rx \left(1 - \frac{x}{N} \right) - qu(t)x \quad (7)$$

因为假设 2 给出了控制函数 $u(t)$ 的形式(5), 所以(6), (7)可转化为函数极值问题.

当 $0 \leq t \leq \tau$ 时 $u=0$, $x(t)$ 容易由方程(7)在初始条件(4)下解出; 当 $t > \tau$ 时 $u=U$, $x(t)$ 要保持在某一常量不变(假设 2), 这个常量可由(7)式令 $\dot{x}=0$ 得到. 于是有

$$x(t) = \begin{cases} \frac{N}{1 + (K-1)e^{-t}}, & 0 \leq t \leq \tau \\ N \left(1 - \frac{qU}{r} \right), & t > \tau \end{cases} \quad (8)$$

由 $x(t)$ 在 $t=\tau$ 时的连续性可以写出

$$\frac{1}{1 + (K-1)e^{-\tau}} = 1 - \frac{qU}{r}$$

由此解得

$$\tau = \frac{1}{r} \ln \left[(K-1) \left(\frac{r}{qU} - 1 \right) \right] \quad (9)$$

即 $u(t)$ 中的两个参数 τ, U 中只有一个是独立的, 以下取 U 为独立变量, $\tau(U)$ 由(9)式确定.

将(5), (8)代入(6)式, 目标泛函 $J(u(t))$ 变为 U 的函数, 记作 $F(U)$, 则

$$\begin{aligned} F(U) &= \int_{\tau}^{\infty} U e^{-\delta t} \left[pqN \left(1 - \frac{qU}{r} \right) - c \right] dt \\ &= \frac{pqNU}{\delta} e^{-\delta \tau(U)} \left(1 - \frac{qU}{r} - b \right), \quad b = \frac{c}{pqN} \end{aligned} \quad (10)$$

注意到 c, p, q, N 的含义, 可知无量纲量 b 是费用—价格比的下界 (因为渔场鱼量取最大值 N). 显然应该有 $b < 1$, 否则成本高于售价, 渔船不会出海. 并且由 (10) 式可知, 效益 $F(U)$ 为正值的条件是 $1 - \frac{qU}{r} - b > 0$, 或记作

$$0 < U < \frac{r(1-b)}{q} \quad (11)$$

用微分法求出在条件 (11) 下 $F(U)$ 的最大值点 U^* 为

$$U^* = \frac{r}{4q} \left[3 - b + \frac{\delta}{r} - \sqrt{\left(1 + b - \frac{\delta}{r} \right)^2 + \frac{8b\delta}{r}} \right] \quad (12)$$

将 (12) 的结果代入 (9) 式即得

$$\tau^* = \tau(U^*) \quad (13)$$

U^*, τ^* 为渔船出海的最佳数量与时刻.

模型解释 按照经济学的观点, 最优解应该在边际得益恰好等于边际损失时达到, 称为边际解释. 为了得到这种解释的表达式, 考察单位时间的利润

$$R(u(t)) = [pqx(t) - c]u(t) \quad (14)$$

当 $t > \tau$ 时以 (5), (8) 代入 (14) 式得

$$R(U) = U \left[pqN \left(1 - \frac{qU}{r} \right) - c \right] \quad (15)$$

与 (10) 式比较可知 $F(U)$ 又可表为

$$F(U) = \int_{\tau(U)}^{\infty} e^{-\delta t} R(U) dt = \frac{1}{\delta} e^{-\delta \tau(U)} R(U) \quad (16)$$

容易算出

$$F'(U) = e^{-\delta \tau(U)} \left[\frac{R'(U)}{\delta} - R(U) \tau'(U) \right]$$

对于最优解 U^* 有 $F'(U^*) = 0$, 故 U^* 必满足

$$-\frac{R'(U^*)}{\delta} = -R(U^*) \tau'(U^*) \quad (17)$$

由此可对最优解 U^* 作出如下的边际解释:

从 (9) 式知道派出渔船的时刻 τ 是渔船数量 U 的减函数, 多派出一只船 (从 U 到 $U+1$) 短期利润的增加是

$$R(U) [\tau(U) - \tau(U+1)] \doteq -R(U) \tau'(U) \quad (18)$$

而长期效益的减少是

$$\int_0^{\infty} e^{-\delta t} [R(U) - R(U+1)] dt \doteq -\frac{R'(U)}{\delta} \quad (19)$$

比较 (18), (19) 与 (17) 式可知, 派出渔船的最佳数量 U^* 应使短期的边际得益恰好被长期的边际损失所平衡.

评注 本节提出的以渔船数量 $u(t)$ 为控制函数的最大效益模型 (6), (7) 式, 其特点是对 $u(t)$ 规定了特殊形式 (5). 这种规定的合理性如何呢? 事实上, 如果取消对 $u(t)$ 的约束, 求解泛函极值 (6), (7) 的话, 则 $u(t)$ 的最优解必然取 (5) 的形式. $u(t)$ 在 $0 \leq t \leq \tau$ 内取零值是为了让渔场鱼量水平尽快地从初值 $x(0)$ 达到稳定值 x^* . 于是, 在形如 (5) 式的最优函数控制下, 达到长期最大效益的最佳渔场鱼量水平 $x(t)$, 如图 1 所示.

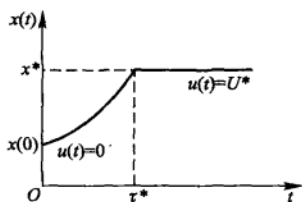


图 1 由最优的 $u(t)$ 控制的 $x(t)$

经济学上关于最优结果的边际解释, 是从正反两方面使短期利益与长期利益取得折中. 想一想人们在日常活动中对相当广泛的优化问题作决策时的思维过程, 不难发现常常就是这类折中办法的具体体现.

13.5 赛跑的速度

参加赛跑的运动员要根据自己的生理状况对赛程中各阶段的速度作出最恰当的安排, 以期获得最好的成绩. 寻求速度安排的最佳策略是一个涉及生理力学的复杂问题. T. B. Keller 提出了一个简单模型, 根据 4 个生理参数从最优控制的角度确定赛程中的速度函数, 并可以预测比赛成绩.

按照 Keller 的模型, 短跑比赛应该用最大冲力跑完全程, 对于中长跑则要将距离分为 3 段, 先用最大冲力起跑, 然后匀速跑过大部分赛程, 最后把贮存在体内的能量用完, 靠惯性冲过终点^[43].

问题分析 运动员在赛跑过程中克服体内外的阻力以达到或保持一定速度, 需要发挥出向前的冲力, 为冲力作功提供能量有两个来源, 一是呼吸和循环系统通过氧的新陈代谢作用产生的与吸入氧量等价的能量, 二是赛跑前贮存在身体内的供赛跑用的能量. 对于前者可以合理地假设在赛跑过程中保持常数, 而后者则有一个如何将贮存能量分配到赛程的各个阶段, 并恰在到达终点前将其用完的问题.

模型需要确定三个关系. 一是冲力与速度的关系, 二是冲力作功与上述两个能量来源的关系, 三是速度与比赛成绩的关系. 虽然比赛成绩指的是一定赛程所用的时间, 但是在用数学工具处理时可以等价为一定时间所跑的距离. 于是最佳成绩将归结为以距离为目标泛函的, 与速度、冲力、贮存能量等函数有关的极值问题. 这个问题的一般解过于复杂, Keller 把它简化了.

由以上分析还可以看出模型需要 4 个生理参数: 运动员能发挥的最大

冲力,体内外的阻力系数,氧的新陈代谢作用单位时间所提供的能量,体内贮存能量的初值. 在 Keller 的模型中这些参数是用世界纪录创造者的成绩估计出来的.

模型假设 需要对赛跑中的阻力作出假设,以确定冲力与速度的关系,还要对氧的代谢作用提供的能量作出假设,以建立能量供给与消耗间的平衡.

1. 赛跑时体内外的阻力与速度成正比,比例系数 τ^{-1} . 运动员能发挥的最大冲力为 F , 初速度为 0.

2. 呼吸和循环系统在氧的代谢作用下单位时间提供的能量是常数 σ , 初始时刻体内贮存的供赛跑用的能量是 E_0 .

实际上,上述参数因人而异,特别是与运动员的体重有关. 为了消除这个因素的影响,我们对运动员的单位质量建模,即在下面各式中均设质量 $m = 1 \text{ kg}$.

一般模型 设运动员以速度函数 $v(t)$ 跑完赛程 D 的时间为 T , 则

$$D(v(t)) = \int_0^T v(t) dt \quad (1)$$

D 已知时求 $v(t)$ 使 T 达到最小, 等价于 T 固定求 $v(t)$ 使 D 达到最大, 后一问题较便于研究.

运动员的冲力记作 $f(t)$, 由假设 1, 根据牛顿第二定律可以得到

$$\dot{v}(t) + \frac{v}{\tau} = f(t) \quad (2)$$

$$v(0) = 0 \quad (3)$$

$$0 \leq f(t) \leq F \quad (4)$$

运动员体内贮存的能量记作 $E(t)$, 其变化率为单位时间提供的能量 σ 与消耗的能量 fv 之差, 即

$$\dot{E}(t) = \sigma - fv \quad (5)$$

$$E(0) = E_0 \quad (6)$$

$$E(t) \geq 0 \quad (7)$$

这样构成了在条件 (2) ~ (7) 下以 (1) 式 $D(v(t))$ 为目标函数的泛函极值问题, 其中 F, τ, σ, E_0 视为已知参数.

这个一般模型的求解是困难的, 因为一旦最优解出现在条件 (4), (7) 的边界上 (从下面的分析可知实际情况正是这样), 变分法就不再适用. 此外, 将 (1), (2), (5) 式合并后得到的泛函表达式

$$D(E(t)) = \sqrt{2} \int_0^T e^{-\frac{t}{\tau}} \left[\int_0^t (\sigma - \dot{E}(s)) e^{\frac{2s}{\tau}} ds \right]^{\frac{1}{2}} dt \quad (8)$$

过于复杂,即使利用其他方法求解也十分不便。

下面将 Keller 提出的简化方法分成两个模型叙述。

短跑模型 当赛程较短时可以用最大冲力 F 跑完全程,这必然会取得最佳成绩。至于多长的赛程才能用这种方法跑,应以贮存于体内的能量 $E(t)$ 不小于零为标准,由参数 F, τ, σ, E_0 决定。

将 $f(t) = F$ 代入方程(2),在初始条件(3)下的解为

$$v(t) = F\tau(1 - e^{-\frac{t}{\tau}}) \quad (9)$$

可知速度是递增的。将(9)及 $f = F$ 代入(5)式得

$$\dot{E}(t) = \sigma - F^2\tau(1 - e^{-\frac{t}{\tau}}) \quad (10)$$

方程(10)在初始条件(6)下的解为

$$E(t) = E_0 - (F^2\tau - \sigma)t + F^2\tau^2(1 - e^{-\frac{t}{\tau}}) \quad (11)$$

由(10),(11)式可以画出 $E(t)$ 的示意图(图1),起跑后在很短的时间 $0 \leq t \leq t_e$

内,由于速度 v 很小, σ 的一部分补充给 $E(t)$ (见(5)式),所以 $E(t)$ 增加;稍后,随着 v 的迅速变大, $E(t)$ 下降,当 $t = t_e$ 时 $E(t) = 0$ 。容易得到

$$t_e = t \mid_{E(t)=0},$$

$$t_e = t \mid_{\dot{E}(t)=0} = \tau \ln \frac{F^2\tau}{F^2\tau - \sigma}$$

并且,在这种情况下所能跑得最远距离为

$$D_e = \int_0^{t_e} v(t) dt = F\tau^2 \left(e^{-\frac{t_e}{\tau}} + \frac{t_e}{\tau} - 1 \right) \quad (12)$$

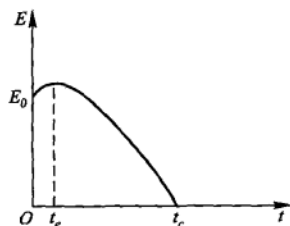


图1 体内能量 $E(t)$ 的示意图

为了估计模型中的参数,并算出 t_e 和 D_e ,

Keller 用若干个当时的短跑世界纪录拟合由(1),(9)式计算的理论结果,得到 F, τ 的估计值为 $F \approx 12.2 \text{ N/kg}$ (对于质量 $m = 1 \text{ kg}$ 而言), $\tau \approx 0.892 \text{ s}$, 算出 $t_e \approx 27.63 \text{ s}$, $D_e \approx 291 \text{ m}$ 。即当小于 291 m 时用最大冲力跑完全程是可行的,并且能够取得最佳成绩(当然是对世界纪录创造者而言)。

表1 约翰逊的成绩 t_1 和刘易斯的成绩 t_2

x/m	10	20	30	40	50	60	70	80	90	100
t_1/s	1.84	2.86	3.80	4.67	5.53	6.38	7.23	8.10	8.96	9.83
t_2/s	1.94	2.96	3.91	4.78	5.64	6.50	7.36	8.22	9.07	9.93

后来,有人用1987年世界田径锦标赛上约翰逊和刘易斯的100 m成绩(每10 m的记录如表1)对 F, τ 的估计值进行修正.由(9)式知最大速度约为 $v_m = F\tau$ ($t \gg \tau$),且距离函数为 $x(t) = F\tau[t - \tau(1 - e^{-t/\tau})]$,当 $t \gg \tau$ 时 $x \approx v_m(t - \tau)$.从表1可得 $v_m = 11.76 \text{ m/s}$, $\tau \approx 1.28 \text{ s}$ 或 1.42 s ,考虑到起跑的反应时间,取 $\tau = 1.06 \text{ s}$ 或 1.16 s ,而 $F \approx 11.1$ 或 10.1 N/kg .

中长跑模型 当赛程大于 D_c 时,将全程分为3段:初始阶段 $0 \leq t \leq t_1$ 以最大冲力 $f(t) = F$ 跑,以便在短时间内获得尽可能的高速度;最后阶段 $t_2 \leq t \leq T$ 把体内贮存能量用完,即 $E(t) = 0$,靠惯性冲刺;中间阶段 $t_1 \leq t \leq t_2$ 保持匀速^①.可以看出,由于在第1,3段分别把控制函数 $f(t), E(t)$ 确定在约束条件(4),(7)的边界上,且取常值,从而这两阶段的速度(分别记作 $v_1(t)$ 和 $v_3(t)$)即可确定.由此不难进一步求出中间阶段的 $v_2(t) = v_2$ (常数).

第1段 $0 \leq t \leq t_1$ (t_1 以后确定) $v_1(t)$ 由(9)式表示(将 $v(t)$ 写作 $v_1(t)$).

第3段 $t_2 \leq t \leq T$ (t_2 以后确定) 将 $E(t) = 0$ 代入(5),(2)式得到

$$\frac{1}{2} \frac{d}{dt} v_3^2 + \frac{v_3^2}{\tau} = \sigma \quad (13)$$

方程(13)在条件 $v_3(t_2) = v_2$ 下的解为

$$v_3(t) = \left[(v_2^2 - \sigma t) e^{-\frac{2(t-t_2)}{\tau}} + \sigma \tau \right]^{\frac{1}{2}} \quad (14)$$

$v_3(t)$ 是单调减和下凸的($\dot{v}_3(t) < 0, \ddot{v}_3(t) > 0$).

第2段 $t_1 \leq t \leq t_2$ v_2 由目标泛函(1)式 $D(v(t))$ 达到最大来确定,这里

$$D(v(t)) = \int_0^{t_1} v_1(t) dt + v_2(t_2 - t_1) + \int_{t_2}^T v_3(t) dt \quad (15)$$

为了得到 $E(t_2) = 0$ 的条件,由方程(2),(5)及初始条件(6)解 $E(t)$,得到

$$E(t) = E_0 + \sigma t - \frac{v^2(t)}{2} - \frac{1}{\tau} \int_0^t v^2(t) dt \quad (16)$$

令 $t = t_2$ 并将(9)式代入得

$$E(t_2) = E_0 + \sigma t_2 - \frac{v_2^2}{2} - \int_0^{t_1} F^2 \tau (1 - e^{-t/\tau})^2 dt - \frac{v_2^2(t_2 - t_1)}{\tau} \quad (17)$$

问题归结为在条件 $E(t_2) = 0$ 下求 v_2, t_2 使(15)式表示的 $D(v(t))$ 最大.

用待定常数 λ 构造函数

$$I(v(t), t_2) = D(v(t)) + \frac{\lambda}{2} E(t_2) \quad (18)$$

不妨将右端与 v_2 和 t_2 无关的项略去,写成

① 如果设速度 $v_2(t)$,泛函问题的解也是常数 v_2 ,详见[43].

$$I(v_2, t_2) = \int_{t_2}^T v_3(t) dt + \frac{\lambda \sigma}{2} t_2 - \frac{\lambda}{4} v_2^2 + \left(v_2 - \frac{\lambda v_2^2}{2\tau} \right) (t_2 - t_1) \quad (19)$$

v_2, t_2 为最优解的必要条件是

$$v_2 = \frac{\tau}{\lambda} \quad (20)$$

$$2 \int_{t_2}^T \left[(v_2^2 - \sigma \tau) e^{-\frac{2(t-t_2)}{\tau}} + \sigma \tau \right]^{-\frac{1}{2}} e^{-\frac{2(t-t_2)}{\tau}} dt = \lambda \quad (21)$$

至此, 3 个阶段的 $v(t)$ 分别由 (9), (14) 和 (20) 给出, 剩下的问题是确定 t_1 , t_2 和 λ .

t_1, t_2, λ 的确定 利用 $v(t)$ 在 $t = t_1$ 的连续性, 由 (9) 和 (20) 式得

$$\lambda F(1 - e^{-\frac{t_1}{\tau}}) = 1 \quad (22)$$

将 (20) 代入 (17) 后令其为 0, 得

$$E_0 + \sigma t_2 - \frac{\tau^2}{2\lambda^2} - \int_0^{t_1} F^2 \tau (1 - e^{-\nu\tau})^2 dt - \frac{\tau}{\lambda^2} (t_2 - t_1) = 0 \quad (23)$$

将 (20) 代入 (21) 并作积分得到

$$2 \left[(\tau^2 - \lambda^2 \sigma \tau) e^{-\frac{2(t-t_2)}{\tau}} + \lambda^2 \sigma \tau \right]^{\frac{1}{2}} - 2\tau = \lambda^2 \sigma - \tau \quad (24)$$

t_1, t_2, λ 由 (22) ~ (24) 确定.

以上模型中的参数 F, τ 已在前面估计出, Keller 又用一些中长跑世界纪录得到 σ 和 E_0 的估计值: $\sigma = 41.5, E_0 = 2403.5$.

模型解释 中长跑模型的速度函数由 3 段组成, 示意图如图 2. 对于最后一段 (通常有一两秒钟) 速度的下降, Keller 的解释是: 像汽车比赛到终点前将燃料用完, 靠惯性冲过终点一样, 赛跑的最佳策略应该是把贮存在体内供赛跑用的能量全部耗尽, 借助惯性冲刺, 这必导致短暂的速度下降. 单从赛跑所用的时间来看, 比如一名运动员测验自己的成绩, 这样做是最优的. 而在实际比赛中当运动员与对手势均力敌时, 从击败对手取得好名次的目的出发, 需要按照实际情况巧妙地安排自己的速度, 这已不是本模型讨论的范围了.

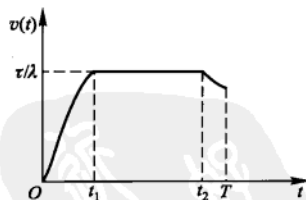


图 2 中长跑模型的最佳速度示意图

评注 Keller 对赛跑的一般模型 (1) ~ (7) 提出了分段解法, 虽然没有严格证明它的解就是 (1) ~ (7) 的最优解, 但是从分析过程和实际检验可以看出这种简化方法是合理的. 另一方面, 模型本身也存在着一一些不合适的地方, 例如对短跑和中长跑的所有赛程用了同一组生理参数, 而短跑和中长跑运动员的这些参数可能有较大差别. 还有, 运动员达到一定的高速度后不可能持续地发挥自己的

最大冲力,短跑模型和中长跑模型中最初阶段的速度(9)式需要修正(习题2).尽管如此,这个模型在把动力学与生理学相结合,用建模方法讨论体育运动问题上,为人们提供了示范.后来,不断有人继续研究赛跑的数学模型,对生理参数的估计值作出修正,并且考虑空气阻力、海拔高度等因素的影响^[63].

13.6 多阶段最优生产计划

4.5节曾用混合整数规划求解过多阶段生产计划,实际上,这是一类典型的动态优化问题,与本章前几节用变分法建立连续动态优化模型不同的是,多阶段生产计划属于离散动态优化问题,动态规划模型是解决这类问题的有效方法.本节先讨论确定需求下的最优生产计划,并将它转化为典型的动态优化模型——最短路问题,然后研究随机需求下如何求解最优生产计划.

问题的提出 已知时段 t 某产品的需求量为 $d_t(t=1,2,\dots,T)$,任一时段若生产该产品,需付出生产准备费 c_0 ,且生产每单位产品的生产成本为 k ,若满足本时段需求后有剩余,每时段每单位产品需付出存贮费 h_0 .设每时段最大生产能力为 X_m ,最大存贮量为 I_m ,且第1时段初有库存量 i_1 ,试制订产品的生产计划,即每时段的产量,使 T 个时段的总费用最小.

为了通过具体的计算说明解决这类问题的方法,设 $T=3, d_1=2, d_2=1, d_3=2$ 单位, $c_0=3$ 千元, $k=2$ 千元/单位, $h_0=1$ 千元/单位·时段, $X_m=4$ 单位, $I_m=3$ 单位, $i_1=1$ 单位.

问题分析与求解 记时段 $t(t=1,2,3)$ 的产量为 x_t ,当 $x_t>0$ 时生产费用为 $c(x_t)=c_0+kx_t=3+2x_t$,而当 $x_t=0$ 时 $c(0)=0$.记时段 t 初的存贮量为 i_t ,满足时段 t 的需求量 d_t 后,时段 $t+1$ 初(即时段 t 末)的存贮量为 $i_{t+1}=i_t+x_t-d_t$,于是时段 t 的存贮费为 $h(i_t)=h_0(i_t+x_t-d_t)=i_t+x_t-d_t$,且 $x_t\leq X_m=4, i_t\leq I_m=3$.

为了简化这个多阶段生产计划问题,我们将它从后向前地分解为一个单个时段问题.

首先看最后一个时段(时段3),对于时段3初的存贮量 i_3 ,记时段3的最小费用为 $f_3(i_3)$,产量为 $x_3(i_3)$.为了使3个时段的总费用最小,时段3末的存贮量显然应为0,且时段3的产量只需满足需求 $d_3=2$ 即可,所以可只考虑 $i_3=0,1,2$ 的情况,容易计算出

$$\begin{aligned} f_3(0) &= c(2) = 3 + 2 \times 2 = 7, x_3(0) = 2; f_3(1) = c(1) = 5, x_3(1) = 1; \\ f_3(2) &= c(0) = 0, x_3(2) = 0 \end{aligned} \quad (1)$$

然后考察倒数第2时段(时段2),对于时段2初的存贮量 i_2 ,产量为 x_2 ,因为 $d_2=1$,时段2末的存贮量为 i_2+x_2-1 ,于是时段2的费用为生产费与存贮费之和 $c(x_2)+h(i_2)$,其中 $h(i_2)=i_2+x_2-1$.时段2与时段3的最小费用之和记

作 $f_2(i_2)$, 注意到时段 3 的最小费用为 $f_3(i_3) = f_3(i_2 + x_2 - 1)$, 所以

$$f_2(i_2) = \min_{x_2} \{c(x_2) + h(i_2) + f_3(i_2 + x_2 - 1)\} \quad (2)$$

$f_2(i_2)$ 的计算见表 1, 由 $d_2 = 1, d_3 = 2$ 知, 应有 $1 \leq i_2 + x_2 \leq 3$, 自然满足 $x_2 \leq X_m = 4$, $i_2 \leq I_m = 3$. 对于每个 i_2 , 表中 $c + h + f_3$ 的最小值标以 * 号, 即为 $f_2(i_2)$, 对应的 x_2 记作 $x_2(i_2)$.

表 1 $f_2(i_2)$ 的计算

i_2	x_2	$c(x_2)$	$h(i_2)$	$f_3(i_2 + x_2 - 1)$	$c + h + f_3$	$f_2(i_2), x_2(i_2)$
0	1	5	0	7	12	$f_2(0) = 11$ $x_2(0) = 3$
0	2	7	1	5	13	
0	3	9	2	0	11*	
1	0	0	0	7	7*	$f_2(1) = 7$ $x_2(1) = 0$
1	1	5	1	5	11	
1	2	7	2	0	9	
2	0	0	1	5	6*	$f_2(2) = 6$ $x_2(2) = 0$
2	1	5	2	0	7	
3	0	0	2	0	2*	$f_2(3) = 2$ $x_2(3) = 0$

最后考察时段 1, 对于时段 1 初的存贮量 $i_1 = 1$, 产量为 x_1 , 因为 $d_1 = 2$, 时段 1 末的存贮量为 $i_1 + x_1 - 2$, 于是时段 1 的费用为 $c(x_1) + h(i_1)$, 其中 $h(i_1) = i_1 + x_1 - 2$. 从时段 1 至时段 3 的最小费用之和记作 $f_1(i_1)$, 注意到时段 2 到时段 3 的最小费用为 $f_2(i_2) = f_2(i_1 + x_1 - 2)$, 所以

$$f_1(i_1) = \min_{x_1} \{c(x_1) + h(i_1) + f_2(i_1 + x_1 - 2)\} \quad (3)$$

$f_1(i_1)$ 的计算见表 2, 由 $d_1 = 2, d_2 = 1, d_3 = 2$ 知, 应有 $2 \leq i_1 + x_1 \leq 5$, 自然满足 $x_1 \leq X_m = 4, i_1 \leq I_m = 3$. 对于 $i_1 = 1$, 表中 $c + h + f_2$ 的最小值标以 * 号, 即为 $f_1(i_1)$, 对应的 x_1 记作 $x_1(i_1)$.

表 2 $f_1(i_1)$ 的计算

i_1	x_1	$c(x_1)$	$h(i_1)$	$f_2(i_1 + x_1 - 2)$	$c + h + f_2$	$f_1(i_1), x_1(i_1)$
1	1	5	0	11	16	$f_1(1) = 15$ $x_1(1) = 2$
1	2	7	1	7	15*	
1	3	9	2	6	17	
1	4	11	3	2	16	

由表 2 可知,3 个时段总费用的最小值为 $f_1(1) = 15$, 达到这个最小值的生产计划, 即 3 个时段的产量可如下得到: 从表 2, $x_1(1) = 2$; 时段 2 初的存贮量为 $i_2 = i_1 + x_1 - 2 = 1 + 2 - 2 = 1$, 从表 1, $x_2(i_2) = x_2(1) = 0$; 时段 3 初的存贮量为 $i_3 = i_2 + x_2 - 1 = 1 + 0 - 1 = 0$, 从(1)式, $x_3(0) = 2$. 即最优生产计划是: 3 个时段的产量依次为 $x_1 = 2, x_2 = 0, x_3 = 2$.

结果表明, 最优生产计划是时段 1 生产 2 单位, 加上原有存贮量 1 单位, 用以满足时段 1, 2 的需求, 时段 2 不生产, 时段 3 生产 2 单位, 满足时段 3 的需求. 显然, 不是每个阶段都产生满足本阶段的需求, 是由于生产准备费的影响.

最短路问题 为了直观地理解上面的算法, 将多阶段生产计划化为下面的最短路问题, 如图 1.

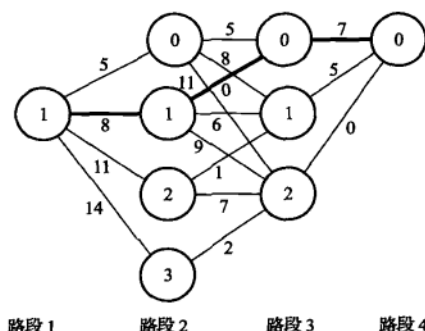


图 1 多阶段生产计划化成的最短路问题

每时段初的存贮量看作各个路段的不同站点, 路段 1 只有站点 1, 路段 2 有 0, 1, 2, 3 共 4 个站点, 路段 3 有 0, 1, 2 共 3 个站点, 而时段 3 末的存贮量为 0 看作路段 4 的站点 0, 图中这些站点用圆圈里的数字表示. 从一个路段的每一站点可以到达下一路段的哪个站点, 由时段初的存贮量、本时段的生产量及需求量确定, 图中用站点间的连线表示, 再将本时段的生产费与存贮费之和作为两站点间的距离, 标在站点间的连线上, 这样, 求各时段生产计划使总费用最小, 就化为寻找从路段 1 的站点 1 到路段 4 的站点 0 的一条最短路径.

最短路问题首先从后向前求解: 路段 3 的每个站点到路段 4 的站点 0 的最短距离相当于(1)式中的 $f_3(i_3)$, 相应的路径为 $x_3(i_3)$, $i_3 = 0, 1, 2$; 路段 2 的每个站点到路段 4 的站点 0 的最短距离相当于表 1 中的 $f_2(i_2)$, 相应的路径为 $x_2(i_2)$, $i_2 = 0, 1, 2, 3$; 路段 1 的站点 1 到路段 4 的站点 0 的最短距离相当于表 2 中的 $f_1(1)$, 相应的路径为 $x_1(1)$. 然后再从前向后确定最短路径: $i_1 = 1, x_1(1) =$

$2 \rightarrow i_2 = 1, x_2(1) = 0 \rightarrow i_3 = 0, x_3(0) = 2 \rightarrow i_4 = 0$, 即图 1 中的粗线标出的路径. 可以看出, 这种做法不仅找到了从起点到终点的最短距离和最短路径, 而且表 1、表 2 中计算的 $f_3(i_3), x_3(i_3), f_2(i_2), x_2(i_2)$ 给出了从路段 3, 2 的各站点到终点的最短距离和最短路径.

上述求解方法来源于这样一个事实: 如果 $i_1 = 1 \rightarrow i_2 = 1 \rightarrow i_3 = 0 \rightarrow i_4 = 0$ 是从 $i_1 = 1$ 到 $i_4 = 0$ 的最短路径, 那么它的任一段子路径, 如 $i_2 = 1 \rightarrow i_3 = 0 \rightarrow i_4 = 0$, 必然是从 $i_2 = 1$ 到 $i_4 = 0$ 的最短路径.

确定需求下多阶段生产计划的一般模型 对于寻求 T 个时段生产计划使总费用最小的问题, 仍沿用上面关于需求量、产量、存贮量、生产费、存贮费及最大生产能力、最大存贮量的记号, 解决过程如下:

1. 根据对时段 T 末存贮量的要求, 确定 $f_{T+1}(i_{T+1})$ (它上面问题中取零值);

2. 时段从后向前地计算最小费用, 按照以下公式递推:

$$f_t(i_t) = \min_{x_t} \{c(x_t) + h(i_t) + f_{t+1}(i_{t+1})\}, \quad i_{t+1} = i_t + x_t - d_t, \quad (4)$$

$$i_t \leq I_m, \quad x_t \leq X_m, \quad t = T, T-1, \dots, 1$$

得到从时段 t 到时段 T 的最小费用 $f_t(i_t)$ 及相应的 $x_t(i_t)$, 若时段 1 初的存贮量为 i_1 , 则 $f_1(i_1)$ 为 T 个时段总费用的最小值;

3. 时段从前向后地确定最优生产计划, 已知 i_1 , 由 $x_t(i_t)$ 及 $i_{t+1} = i_t + x_t(i_t) - d_t$ 得到 $x_t, t = 1, 2, \dots, T$.

随机需求下的多阶段生产计划 如果每个时段的需求量是随机的, 那么对于确定的生产量, 各时段的存贮量也是随机的, 于是存贮费乃至总费用都是随机的, 优化问题的目标应是总费用的期望值最小, 这个随机优化问题可以用随机动态规划求解.

下面通过具体数值计算说明随机动态规划模型的建立和求解. 仍然考察 $T=3$ 个时段, 沿用前面的记号, 设需求量 $d_t = 1$ 的概率为 $1/3$, $d_t = 2$ 的概率为 $2/3$ ($t=1, 2, 3$), 每个时段的需求必须满足, 生产费、存贮费、每时段最大生产能力、最大存贮量、第 1 时段初库存量均同上. 因为计划结束时 (时段 3 末) 存贮量是随机的 (不一定为零), 我们假定, 这时剩余的存贮量能够以 1.5 千元/单位的价格出售.

将随机需求表示为 $P(d_t = 1) = 1/3, P(d_t = 2) = 2/3$, 生产费用为 $c(x_t) = c_0 + kx_t = 3 + 2x_t$ ($x_t > 0$), $c(0) = 0$ ($x_t = 0$), 存贮量的转移仍为 $i_{t+1} = i_t + x_t - d_t$, 由随机需求得到存贮费的期望值

$$\begin{aligned} Eh(i_t) &= h_0 E(i_t + x_t - d_t) \\ &= (i_t + x_t - 1)P(d_t = 1) + (i_t + x_t - 2)P(d_t = 2) \end{aligned}$$

$$= (i_t + x_t - 1)/3 + 2(i_t + x_t - 2)/3$$

$$= i_t + x_t - 5/3$$

当时段3初的存贮量为 i_3 时,计划结束时出售剩余量得到的回报记作 $s(i_3)$,回报的期望值为 $Es(i_3) = 1.5[(i_3 + x_3 - 1)/3 + 2(i_3 + x_3 - 2)/3] = 1.5(i_3 + x_3) - 2.5$,产量、存贮量的限制仍为 $x_t \leq X_m = 4, i_t \leq I_m = 3$.

记时段3期望费用的最小值为 $f_3(i_3)$,产量为 $x_3(i_3)$.在满足随机需求 d_3 的前提下,容易算出(注意:出售剩余量的回报要从费用中扣除)

$$\begin{aligned} f_3(0) &= c(2) - Es(0) = 7 - 1/2 = 13/2, & x_3(0) &= 2; \\ f_3(1) &= c(1) - Es(1) = 5 - 1/2 = 9/2, & x_3(1) &= 1; \\ f_3(2) &= c(0) - Es(2) = 0 - 1/2 = -1/2, & x_3(2) &= 0; \\ f_3(3) &= c(0) - Es(3) = 0 - 2 = -2, & x_3(3) &= 0 \end{aligned} \quad (5)$$

计算时段2与时段3期望费用的最小值 $f_2(i_2)$ 时,(2)式中的 $h(i_2)$ 应改为 $Eh(i_2) = i_2 + x_2 - 5/3, f_3(i_3)$ 应改为 $f_3(i_2 + x_2 - 1)P(d_1=1) + f_3(i_2 + x_2 - 2)P(d_1=2) = f_3(i_2 + x_2 - 1)/3 + 2f_3(i_2 + x_2 - 2)/3$,即

$$f_2(i_2) = \min_{x_2} \{c(x_2) + Eh(i_2) + f_3(i_2 + x_2 - 1)/3 + 2f_3(i_2 + x_2 - 2)/3\} \quad (6)$$

$f_2(i_2)$ 的计算见表3,为满足需求应有 $2 \leq i_2 + x_2 \leq 4$,且 $x_2 \leq X_m = 4, i_2 \leq I_m = 3$.对于每个 i_2 ,表中 $c + Eh + f_3$ 的最小值标以*号,即为 $f_2(i_2)$,对应的 x_2 记作 $x_2(i_2)$.

表3 随机需求下 $f_2(i_2)$ 的计算

i_2	x_2	$c(x_2)$	$Eh(i_2)$	$f_3(i_2 + x_2 - 1)/3 + 2f_3(i_2 + x_2 - 2)/3$	$c + Eh + f_3$	$f_2(i_2), x_2(i_2)$
0	2	7	1/3	35/6	79/6	$f_2(0) = 37/3$ $x_2(0) = 4$
0	3	9	4/3	17/6	79/6	
0	4	11	7/3	-1	37/3*	
1	1	5	1/3	35/6	67/6	$f_2(1) = 31/3$ $x_2(1) = 3$
1	2	7	4/3	17/6	67/6	
1	3	9	7/3	-1	31/3*	
2	0	0	1/3	35/6	37/6*	$f_2(2) = 37/6$ $x_2(2) = 0$
2	1	5	4/3	17/6	55/6	
2	2	7	7/3	-1	25/3	
3	0	0	4/3	17/6	25/6*	$f_2(3) = 25/6$ $x_2(3) = 0$
3	1	5	7/3	-1	19/3	

类似地计算从时段 1 至时段 3 期望费用的最小值 $f_1(i_1)$, $i_1 = 1$, 公式为

$$f_1(i_1) = \min_{x_1} \{c(x_1) + Eh(i_1) + f_2(i_1 + x_1 - 1)/3 + 2f_2(i_1 + x_1 - 2)/3\} \quad (7)$$

$f_1(i_1)$ 的计算见表 4, 为满足需求且要使 $i_2 \leq I_m = 3$, 应有 $2 \leq i_1 + x_1 \leq 4$. 对于 $i_1 = 1$, 表中 $c + Eh + f_2$ 的最小值标以 * 号, 即为 $f_1(i_1)$, 对应的 x_1 记作 $x_1(i_1)$.

表 4 随机需求下 $f_1(i_1)$ 的计算

i_1	x_1	$c(x_1)$	$Eh(i_1)$	$f_2(i_1 + x_1 - 1)/3 + 2f_2(i_1 + x_1 - 2)/3$	$c + Eh + f_2$	$f_1(i_1), x_1(i_1)$
1	1	5	1/3	105/9	306/18	$f_1(1) = 303/18$
1	2	7	4/3	161/18	311/18	$x_1(1) = 3$
1	3	9	7/3	33/6	303/18*	

由表 4 可知, 3 个时段期望费用的最小值为 $f_1(1) = 303/18 \approx 16.8$, 达到这个最小值的生产计划, 即 3 个时段的产量应如下确定:

从表 4, $x_1(1) = 3$. 由于时段 1 的需求 d_1 是随机的, 时段 2 初的存贮量 $i_2 = i_1 + x_1 - d_1$ 也是随机的: 若 $d_1 = 1$, 则 $i_2 = 1 + 3 - 1 = 3$, 从表 3, $x_2(i_2) = x_2(3) = 0$; 若 $d_1 = 2$, 则 $i_2 = 1 + 3 - 2 = 2$, 从表 3, $x_2(i_2) = x_2(2) = 0$. 类似地分析由 d_2 的随机性得到的 i_3 和 $x_3(i_3)$.

在确定性需求下得到的最优生产计划在开始时就完全确定了: 3 个时段的产量 $x_1 = 2, x_2 = 0, x_3 = 2$, 而随机需求下的最优生产计划只有当每个时段初的存贮量知道后才能确定, 这是二者最重要的区别之一. 可以将随机需求下的最优计划用图 2 表示.

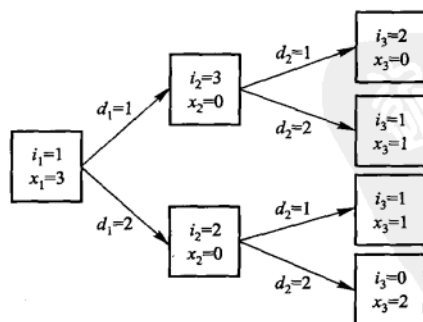


图 2 随机需求下的最优生产计划

评注 上面寻求多阶段生产计划的方法称为动态规划, 它是求解多阶段优化决策问题的有效工具, 建立动态规划模型的主要步骤为: 划分阶段, 定义状态

(如存贮量)和决策(如产量),建立状态转移律(如 $i_{t+1} = i_t + x_t - d_t$),确定允许状态集合和允许决策集合(如 $i_t \leq I_m, x_t \leq X_m$),列出最优方程并确定终端条件(如(4)式及 $f_{T+1}(i_{T+1})$).其中如何选定状态是关键的一步,状态应能描述过程的特征,可以直接或间接观测,并且具有无后效性,即当某阶段的状态给定后,过程以后的演变与该阶段以前的状态无关.

动态规划模型常用来求解经济管理中的货物存贮、设备更新、资源分配、任务均衡、水库调度、系统可靠性等问题,在离散系统最优控制中也有广泛应用.

许多优化决策问题既可以建立静态规划模型,如第4章那样,也可以建立动态规划模型,比较起来,后者的优越性在于:

1. 能够得到全局最优解. 动态规划把全过程化为一系列结构相似、互相关联的子过程,每个子过程的变量个数大大减少,约束集合也简单得多,即使在目标函数和状态转移律无解析表达式时,也可用穷举法求解. 而非线性规划当目标函数及约束集合稍微复杂时,就很难得到全局最优解.

2. 可以得到一族最优解. 动态规划得到的是全过程和所有后部子过程的各个状态的最优解,这在讨论最优决策和最优值对于状态的稳定性,或者实际问题要寻找次优解时是很有用的. 而静态规划一般只能得到一个最优解.

3. 由于动态规划方法反映了动态过程演变的联系和特征,在计算时可以利用实际知识和经验提高求解效率.

动态规划模型的主要缺点是:

1. 没有统一的标准模型,也没有万能的构造模型的方法,需要对每类问题具体分析. 在定义状态、建立状态转移律等方面,需要灵活的技巧,这就带来了应用上的局限性.

2. 用数值方法求解时存在维数灾. 由于状态个数随维数呈指数增长,对高维问题求解十分困难.

习 题

1. 在13.2节生产计划制订模型中,当 $Q < \frac{k_2 T^2}{4k_1}$ 时求最优解. 图2中 t_1 的确定可视为曲线 S_3 始端在直线 $x=0$ 上变动的泛函极值问题.

2. 经研究发现在短跑比赛中,运动员由于生理条件的限制在达到一定的高速度后不可能持续发挥自己的最大冲力. 假设运动员克服生理限制后能发挥的冲力 $f(t)$ 满足 $\dot{f}(t)/f(t) = -\frac{1}{k}$, k 是冲力限制系数, $f(0) = F$ 为最大冲力.

将上述关系代入13.5节赛跑模型的(2)式,求出短跑比赛时速度 $u(t)$ 和距离 $s(t)$ 的表达式,及达到最高速度的时间,作出 $v(t)$ 的示意图.

某届奥运会男子百米决赛前 6 名在比赛中到达距离 s 处所用的时间 t 和当时的速度 v 如下表所示(平均值):

s/m	0	5	15	25	35	45	55	65	75	85	95
t/s	0	0.955	2.435	3.435	4.355	5.230	6.085	6.945	7.815	8.690	9.575
$v/(\text{m} \cdot \text{s}^{-1})$	0	5.24	9.54	10.52	11.19	11.62	11.76	11.49	11.47	11.36	11.22

试从这组数据估计出参数 τ, k, F , 算出 $v(t)$ 的理论值与实际数据比较.

你对这个模型有什么解释和评价^[77]?

3. 用 13.6 节的动态规划方法建立如下的设备更新模型:

(1) 一台设备(如卡车、机床等)每年初要作出“继续使用”还是“更新”的决定. 假定随着设备使用年龄的增长, 设备一年创造的收入、维修费用及更新的净费用均已知, 制订一个有限年度的设备更新计划, 使总的净利润最大.

(2) 在(1)的基础上增加一个可以采取的决策——修复, 即一次全面整修, 能使旧设备的收入增加, 维修费降低, 修复费用又低于更新费用. 同时, 设备一年创造的收入、维修费用及更新的净费用, 除与设备使用年龄有关外, 还与修复的时间有关. 重新建立设备更新模型, 使总的净利润最大.

*4. 用 13.6 节的动态规划方法解决如下的系统可靠性问题: 系统由若干部件串接而成, 只要一个部件出现故障, 系统就不能正常工作. 为提高系统可靠性每个部件都装有备件, 一旦原部件出现故障, 备件就自动进入系统. 显然, 备件越多系统可靠性越大, 但费用也越高. 问题是在一定的总费用下, 如何配置各部件的备件使系统的可靠性最大.

(1) 由 n 个部件串接的系统, 当部件 k 配置 j 个备件时, 该部件正常工作的概率及费用均已知, 在总费用不超过给定值的条件下, 建立使系统的可靠性最大的模型.

(2) 设 $n=3$ 且每个部件至多配置 3 个备件, 部件 k 配置 j 个备件时正常工作的概率 p_{kj} 及费用 c_{kj} 如下:

p_{kj}		备件数量 j			c_{kj}		备件数量 j		
		1	2	3			1	2	3
部件号 k	1	0.5	0.7	0.9	部件号 k	1	2	4	5
	2	0.7	0.8	0.9		2	3	5	6
	3	0.6	0.8	0.9		3	1	2	3

总费用不超过 10, 如何配置各部件的备件数使系统的可靠性最大?

*5. 一家集生产、销售于一体的公司, 希望生产率 and 贮存量都尽量稳定在预先设定的水平上, 如果销售量可以预测, 公司需要制订一个根据贮存量控制生产率的策略.

(1) 以在一定时间 T 内生产率 and 贮存量与设定值误差的(加权)平方和最小为目标, 给出泛函极值问题.

(2) 设销售量为常数, 求出最优解, 并在 T 很大的情况下给出生产率 and 贮存量之间的关系.

综合题目^①

1. 影院座位设计

图1为影院的剖面示意图,座位的满意程度主要取决于视角 α 和仰角 β .视角 α 是观众眼睛到屏幕上、下边缘视线的夹角, α 越大越好;仰角 β 是观众眼睛到屏幕上边缘视线与水平线的夹角, β 太大使人的头部过分上仰,引起不舒适感,一般要求 β 不超过 30° .

记影院屏幕高 h ,上边缘距地面高 H ,地板线倾角 θ ,第一排和最后一排座位与屏幕水平距离分别为 d 和 D ,观众平均坐高为 c (指眼睛到地面的距离).已知参数 $h=1.8\text{ m}$, $H=5\text{ m}$, $d=4.5\text{ m}$, $D=19\text{ m}$, $c=1.1\text{ m}$.

- (1) 地板线倾角 $\theta=10^\circ$,问最佳座位在什么地方?
- (2) 求地板线倾角 θ (一般不超过 20°),使所有观众的平均满意程度最大.
- (3) 地板线设计成什么形状可以进一步提高观众的满意程度.

2. 降落伞的选择

为向灾区空投救灾物资共 $2\ 000\text{ kg}$,需选购一些降落伞.已知空投高度为 500 m ,要求降落伞落地时的速度不能超过 20 m/s .降落伞面为半径 r 的半球面,用每根长 l 共16根绳索连接的载重 m 位于球心正下方球面处,如图2.

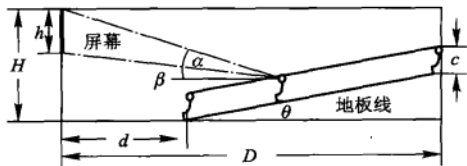


图1 影院剖面示意图

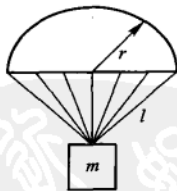


图2

每个降落伞的价格由三部分组成.伞面费用 C_1 由伞的半径 r 决定,见表1;绳索费用 C_2 由绳索总长度及单价 4 元/m 决定;固定费用 C_3 为 200 元 .

降落伞在降落过程中受到的空气阻力,可以认为与降落速度和伞面积的乘

^① 与第3版比较,删去全部全国大学生数学建模竞赛的题目,因为这样的题目和参考解答可以从许多书籍、期刊、网站上得到;更新美国大学生数学建模竞赛的题目;对其他题目也做了删除和增补.

积成正比. 为了确定阻力系数, 用半径 $r=3\text{ m}$ 、载重 $m=300\text{ kg}$ 的降落伞从 500 m 高度作降落试验, 测得各时刻 t 的高度 x , 见表 2.

表 1

r/m	2	2.5	3	3.5	4
$C_1/\text{元}$	65	170	350	660	1 000

表 2

t/s	0	3	6	9	12	15	18	21	24	27	30
x/m	500	470	425	372	317	264	215	160	108	55	1

试确定降落伞的选购方案, 即共需多少个, 每个伞的半径多大 (在表 1 中选择), 在满足空投要求的条件下, 使费用最低.

3. 血样的分组检验^[26]

在一个很大的人群中通过血样检验普查某种疾病, 假定血样为阳性的先验概率为 p (通常 p 很小). 为减少检验次数, 将人群分组, 一组人的血样混合在一起化验. 当某组的混合血样呈阴性时, 即可不经检验就判定该组每个人的血样都为阴性; 而当某组的混合血样呈阳性时, 则可判定该组至少有一人血样为阳性, 于是需要对这组的每个人再作检验.

(1) 当 p 固定时 (如 0.01% , \dots , 0.1% , \dots , 1% , \dots) 如何分组, 即多少人一组, 可使平均总检验次数最少, 与不分组的情况比较.

(2) 当 p 多大时不应分组检验.

(3) 当 p 固定时如何进行二次分组 (即把混合血样呈阳性的组再分成小组检验, 重复一次分组时的程序).

(4) 讨论其他分组方式, 如二分法 (人群一分为二, 阳性组再一分为二, 继续下去)、三分法等.

4. 铅球掷远^[38]

铅球掷远比赛的场地是直径 2.135 m 的圆, 要求运动员从场地中将 7.257 kg (男子) 重的铅球投掷在 45° 的扇形区域内, 如图 3. 观察运动员比赛的录像发现, 他们的投掷角度变化较大, 一般在 $38^\circ \sim 45^\circ$, 有的高达 55° , 试建立模型讨论以下问题:

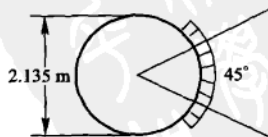


图 3

(1) 以出手速度、出手角度、出手高度为参数, 建立铅球掷远的数学模型.

(2) 给定出手高度, 对于不同的出手速度, 确定最佳出手角度. 比较掷远结

果对出手速度和出手角度的灵敏性.

(3) 考虑运动员推铅球时用力展臂的动作,改进上面的模型.

5. 牧场管理^[38]

有一块一定面积的草场放牧羊群,管理者要估计草场能放牧多少羊,每年保留多少母羊羔,夏季要贮存多少草供冬季之用.

为解决这些问题调查了如下的背景材料:

1) 本地环境下这一品种草的日生长率为

季节	冬	春	夏	秋
日生长率/($g \cdot m^{-2}$)	0	3	7	4

2) 羊的繁殖率 通常母羊每年产1~3只羊羔,5岁后被卖掉.为保持羊群的规模可以买进母羊,或者保留一定数量的母羊羔.每只母羊的平均繁殖率为

年龄	0~1	1~2	2~3	3~4	4~5
产羊羔数	0	1.8	2.4	2.0	1.8

3) 羊的存活率 不同年龄的母羊的自然存活率(指存活一年)为

年龄	1~2	2~3	3~4
存活率	0.98	0.95	0.80

4) 草的需求量 母羊和羊羔在各个季节每天需要的草的数量(单位:kg)为

季节	冬	春	夏	秋
母羊	2.10	2.40	1.15	1.35
羊羔	0	1.00	1.65	0

按照以下假设建模:

(1) 只考虑羊的数量,而不管它们的体重.

(2) 母羊只在春季产羊羔,公、母羊羔各占一半,当年秋季将全部公羊羔和一部分母羊羔卖掉,以保持母羊(每个年龄的)数量不变.

6. 一类答卷评阅结果的处理

在数学建模竞赛、语文作文考试等一类没有标准答案的答卷评阅中,不同评阅人对同一份答卷给出的分数,出现一定范围内的差别是正常的.但是,由于众多客观、主观因素的影响,某些评阅人的打分会存在以下异常现象:

• 打分普遍偏高或偏低,导致他评阅的所有答卷的平均分明显高于或低于总体的平均分(总体指全体评阅人对所有答卷的打分);

• 打分范围过窄,区分度太小,导致他评阅的所有答卷的分数范围明显小于总体的分数范围.

在评阅过程中组织者可以通过一定的程序,让每位评阅人随机地评阅若干份答卷,并且同一份答卷也随机地由若干位评阅人评阅.而在评阅结束后,组织者要根据所有评阅结果筛选出打分存在上述异常现象的评阅人,并且确定每一份答卷的最终分数.需要

(1) 给出筛选打分存在上述异常现象的评阅人的数学模型和求解方法.

(2) 给出确定每份答卷最终分数的数学模型和求解方法.

模拟产生数据进行计算并检验模型:将 150 份答卷随机分配给 9 位评阅人,每份答卷由 3 人评阅(每位评阅人评阅 50 份),评阅人中一人打分偏高,一人打分偏低,一人打分范围过窄.首先模拟产生一组数据作为答卷的真实分数,再在真实分数上增加扰动,作为评阅人的打分.然后将模型用于这组数据,应能把三位出现异常的评阅人筛选出来,并且使确定的每份答卷最终分数与其真实分数相近.

7. 设计环岛(2009 年美国大学生数学建模竞赛 A 题)

许多城市和社区都设有交通环岛——从有几条行车道的大型环岛(如法国巴黎的凯旋门和泰国曼谷的胜利纪念碑处)到只有一或两条行车道的小型环岛.有些环岛在每条进入环岛的车道路口设置停车标志或让行标志,给已经驶入环岛的车辆以行车优先权;有的在每条进入环岛的车道路口设置让行标志,给正在驶入环岛的车辆以行车优先权;还有一些在每条进入环岛的车道路口设置交通信号灯(红灯时不能右转弯).还可能有其他的设计.

本题的目的是要你们用模型来确定进入环岛、环岛内以及从环岛出去的交通流的最优控制.要清楚地叙述为了作出最优选择而在你的模型中用到的目标函数,以及影响这种选择的因素.论文还应包括不超过 2 页 2 倍行距打印的技术报告,向交通工程师解释怎样用你们的模型对任何特定的环岛选择适当的交通流控制方法,即说明应用每种交通流控制方法的条件.如果推荐使用红绿灯的话,则要说明确定绿灯要亮几秒钟(可以按照每天不同的时间以及其他因素而变化)的方法.说明你们的模型怎样能用来解决一些特殊的环岛实例.

8. 遭受巨大损失(2008 年美国大学生数学建模竞赛 A 题)

考虑由于预计全球温度会上升而导致的北极冰盖的融化对陆地的影响.特别要对由于冰盖融化在今后 50 年中每 10 年对佛罗里达州沿岸,尤其是大城市地区的影响进行建模.试提出适当的应对措施来处理这个问题.对所用数据的仔细讨论是回答本问题的重要组成部分.

9. 机场中轮椅的使用(2006年美国大学生数学建模竞赛B题)

乘飞机旅行令人头疼的事情之一是需要多个机场转机,而且每到一个机场通常要求旅客去换乘另外一架飞机.对那些行动有困难的旅客而言,从一个候机区走到另外一个候机区就特别困难了.航空公司能够使中转更加方便的办法之一是,为请求帮助的旅客提供轮椅和陪同人员.通常都能预先知道哪些乘客需要帮助,但也常有旅客在到机场登记时才请求帮助.在很少情况下,直到飞机就要降落前,航空公司可能还没有接到需要帮助的旅客的请求.

航空公司面临着降低成本的持续压力.轮椅很昂贵,会用坏,也需要管理和维护.提供陪同人员也需要费用.另外,为给需要帮助的旅客在他们的航班到达机场时能及时提供帮助,轮椅和陪同人员还要不断在机场移动.在一些大机场,人员和设备在机场内部移动所花费的时间也是不容忽视的.轮椅还需要有存放的地方,但是候机大厅场地的租赁昂贵而且极其有限.还有,把轮椅留在客流繁忙的通道,过往旅客试图绕过它们也会造成不便.最后,最大的代价之一是,如果某位旅客必须等候陪同人员的到来而导致飞机为等他而延误航班.这种代价特别令人烦恼,因为它有可能影响到航空公司的平均航班延误时间,导致有些潜在乘客会避开这个公司的航班,造成该公司机票销售的减少.

Epsilon航空公司决定请第三方帮助他们就为旅客提供轮椅和陪同人员服务的管理和维护中的各种问题和成本进行详细的分析.这家公司希望得到一个讲究成本效益的每天的轮椅调度方法,并找出和定义短期和长期的预算规划所需的各种成本.

Epsilon航空公司要求你们的咨询小组汇集你们的分析形成一个投标,以帮助该公司解决他们的问题.你们的投标书应该包括对实际情况的概述和分析,以便这家航空公司能够确定你们是否已经完全了解他们的问题.他们需要你们提供将要执行的算法的详细叙述,该算法要能确定轮椅和陪同人员应该安置在哪里,以及每天应该怎样移动,其目标是使总的成本尽可能低.你们的投标书是Epsilon航空公司将会考虑的许多投标书之一,你们必须提供一个强有力的案例来说明为什么你们的解决方案是最佳的,而且能够处理各种环境下的各类机场的问题.

你们的投标书还应该包括该算法如何处理大型(至少4个候机大厅)、中型(至少两个候机大厅)和小型(1个候机大厅)机场在客流高峰和低谷时段的各种例子.你们应该确定所有潜在的成本并权衡它们各自的权重.最后,因为在旅客总数中老年人开始占有更大的比例,他们有更多的时间外出旅行,但也可能提出更多的帮助要求,所以你们的报告还应该包括对未来潜在成本和乘客需求的规划,以及怎样满足未来需求的建议.

10. 风和喷泉(2002年美国大学生数学建模竞赛A题)

在一个楼群环绕的宽阔的露天广场上,装饰喷泉把水喷向高空.刮风的日子,风把水花从喷泉吹向过路行人.喷泉射出的水流受到一个与风速计(用于测量风的速度和方向)相连的机械装置的控制,风速计安装在一幢邻近楼房的顶上.这个控制的目标是为行人在赏心悦目的景象和淋水浸湿之间提供可以接受的平衡:风刮得越猛,水量和喷射高度就越低,从而较少的水花落在水池范围以外.

你的任务是设计一个算法,随着风力条件的变化,运用风速计给出的数据来调整由喷泉射出的水流.



参考文献

- [1] Ackerson B, Bertholf D, Choike J, et al. *Red and Blue Laser CDs; How Much Data Can They Hold?* UMAP/ILAP Modules, 2001 - 02: 155 - 177.
- [2] Alevras D, Padberg M W. *Linear Optimization and Extensions: Problems and solutions*. Springer, 2001.
- [3] Balinski M L, Young H P. *The Quota Method of Apportionment*. American Mathematical Monthly, 1975, 82(7).
- [4] Balinski M L, Young H P. *Criteria for Proportional Representation*, Operations Research, 1979, 27(1)
- [5] Balinski M L, Young H P. *Fair Representation*. Washington, D. C. : Brookings Institution Press, 2001.
- [6] Ball M A. *Mathematics in the Social and Life Sciences*. John Wiley & Sons, 1985.
- [7] Banzhaf J F. *Weighted Voting doesn't Work: A Mathematical Analysis*. Rutgers Law Rev. , 1965, 19.
- [8] Bartholomew D J. *Stochastic Models for Social Processes*. 1973.
- [9] Bates D M, Watts D G. *Nonlinear Regression Analysis and Its Applications*. John Wiley & Sons, 1988.
- [10] Bender EA. 数学模型引论. 朱尧辰, 徐伟宣, 译. 北京: 科学普及出版社, 1982.
- [11] Berry J S. *Teaching and Applying Mathematical Modeling*. John Wiley & Sons, 1984.
- [12] Bondy J A. 图的理论及应用. 吴望名等, 译. 北京: 科学出版社, 1984.
- [13] Bowerman B L, O'Connell R T. *Applied Statistics; Improving Business Processes*. McGraw-Hill Comp. Inc. , 1997.
- [14] Bradley S P, Hax A C, Magnanti T L. *Applied Mathematical Programming*. Addison-Wesley Pub. Comp. , 1977.
- [15] Braun M. 微分方程及其应用. 张鸿林, 译. 北京: 人民教育出版社, 1980.
- [16] Burghes D N, Borrie M S. *Modeling with Differential Equations*. Horwood, 1981.
- [17] Burghes D N, Huntley I, McDonald J. *Applying Mathematics; A Course in Mathematical Modeling*. John Wiley & Sons, 1982.
- [18] Cachon G. Supply Chain Coordination With Contracts. In S. Graves, T. de Kok, eds. , *The Handbook of Operations Research and Management Science; Supply Chain Management*. Amsterdam: Kluwer, 2003.
- [19] Chatterjee S, Hadi A S, Price B. *Regression Analysis by Example*. 3rd ed. John Wiley & Sons, 2000.

- [20] Christofides T C. *A Generalized Randomized Response Technique*. *Metrika*, 2003, 57: 195 - 200.
- [21] Clark C W. *Mathematical Bioeconomics*. NY: John Wiley & Sons, 1990.
- [22] Cross M, Moscardini A O. *Learning the Art of Mathematical Modeling*. John Wiley & Sons, 1985.
- [23] Draper N R, Smith H. *Applied Regression Analysis*. 3rd ed. John Wiley & Sons, 1998.
- [24] Dubey P. *On the Uniqueness of the Shapley Value*. *Int. J. Game Theory*, 1975, (4).
- [25] Dubey P, Shapley L S. *Mathematical Properties of the Banzhaf Power Index*. *Math. Operations Res.*, 1979, (4).
- [26] Edwards D, Hamson M. *Guide to Mathematical Modeling*. CRC Press, 1990.
- [27] Ernst L R. *Apportionment Methods for the House of Representatives and the Court Challenges*. *Management Science*, 1994, 40(10).
- [28] Fehr E, Schmidt K M. *A Theory of Fairness, Competition, and Cooperation*. *The Quarterly Journal of Economics*, 1999, 114(3): 817 - 868.
- [29] Floudas C A, et al. *Handbook of Test Problems in Local and Global Optimization*. Kluwer Academic Publishers, 1999.
- [30] Gibbons R. *Game Theory for Applied Economists* (also called *A Primer in Game Theory*). Princeton University Press, 1992. (中文版: 中国社会科学出版社, 1999)
- [31] Gillman R, Housman D. *Models of Conflict and Cooperation*. Providence: American Mathematical Society, 2009.
- [32] Giordano F R, Fox W P, Horton S B, et al. 数学建模. 叶其孝, 姜启源, 等, 译. 北京: 机械工业出版社, 2009.
- [33] Greenberg B G, Abul-Ela A A, Simmons W R, et al. *The Unrelated Question Randomized Response Model: Theoretical Framework*. *Journal of the American Statistical Association*, 1969, 64 (326): 520 - 539.
- [34] Harberman R. *Mathematical Models*. NJ: Prentice - Hall, Englewood Cliffs, 1977.
- [35] Hilliard M R. *Weighted Voting: Theory and Applications*. Ph. D. Thesis, Operations Research. Ithaca, NY: Cornell University, 1983.
- [36] Hosmer D W, Lemeshow S. *Applied logistic Regression*. John Wiley & Sons, 1989.
- [37] Hsieh J. 计算机断层成像技术. 张朝宗, 等, 译. 北京: 科学出版社, 2006.
- [38] Huntley I D, James D J G. *Mathematical Modeling: A Source Book of Case Studies*. Oxford University Press, 1990.
- [39] Isaacson E de St Q, Isaacson M de St Q. *Dimensional Methods in Engineering and Physics*. Edward Arnold, 1975.
- [40] James D J G, McDonald J J. *Case Studies in Mathematical Modeling*. Stanley Thornes Ltd. 1981.
- [41] Kamien M I, Schwartz N L. *Dynamic Optimization: The Calculus and Variations and Optimal Control in Economics and Management*. North-Holland, 1981.

- [42] Kapur J N. *Mathematical Modeling*. John Wiley & Sons, 1988.
- [43] Keller J B. *Optimal Velocity in a Race*. *The American Mathematical Monthly*, 1974, 81: 474.
- [44] Kobylski G., S. Joly, D. Bailey. *Water Purification*. UNAP ILAP Modules, 2007.
- [45] Laruelle A, Valenciano F. *Shapley-Shubik and Banzhaf Indices Revisited*. *Math. Operations Res.*, 2001, 26(1): 89 ~ 104.
- [46] Levin S A. *Applied Mathematical Ecology*. Springer - Verlag, 1989.
- [47] Lin J Q (林家翘), Segel L A. 自然科学中确定性问题的应用数学. 赵国英, 等, 译. 北京: 科学出版社, 1986.
- [48] Logan J D. *Applied Mathematics—A Contemporary Approach*. NY: John Wiley & Sons, 1987.
- [49] Lucas W F. 微分方程模型. 朱煜民, 等, 译. 长沙: 国防科技大学出版社, 1988.
- [50] Lucas W F. 离散与系统模型. 长沙: 国防科技大学出版社, 1996.
- [51] Lucas W F. 政治及有关模型. 长沙: 国防科技大学出版社, 1996.
- [52] Lucas W F. 生命科学模型. 长沙: 国防科技大学出版社, 1996.
- [53] Meerschaert M M. *Mathematical Modeling*. 2nd ed. Academic Press, 1999.
- [54] Meger W J. *Concept of Mathematical Modeling*. McGraw - Hill Book Company, 1985.
- [55] Michael Mesterton-Gibbons. *A Concrete Approach to Mathematical Modeling*. Addison - Wesley, 1989.
- [56] Mooney D D, Swift R J. *A Course in Mathematical Modeling*. The Mathematical Association of America, 1999.
- [57] Murphy A H, Ehrendorfer M. *On the Relationship between the Accuracy and Value of Forecasts in the cost-loss Ratio Situation*. *Weather Forecasting*, 1987, 2.
- [58] Murphy A. H. *What is a Good Forecast? An Essay on Nature of Goodness in Weather Forecasting*. *Weather Forecasting*, 1993, 8.
- [59] Neter J, Wasserman W, Kutner M H. *Applied Linear regression models*. R. D. Irwin, 1983.
- [60] Oliveira-Pinto F, Conolly B W. *Applicable Mathematics of Non - physical Phenomena*. Ellis Horwood, 1982.
- [61] Osborne M J. *An Introduction to Game Theory*. NY: Oxford Univ. Press, 2004. (影印版: 上海财经大学出版社, 2005)
- [62] Pielou E C. 数学生态学. 卢泽愚, 译. 北京: 科学出版社, 1988.
- [63] Pritchard W G. *Mathematical Models of Runing*. *SIAM Review*, 1993, 35(3): 359.
- [64] Roberts F S. *Discrete Mathematical Models*. Prentice-Hall, Englewood Cliffs, 1976.
- [65] Saaty T L. *The Analytic Hierarchy Process*. McGraw-Hill Company, 1980.
- [66] Saaty T L, Alexander J M. *Thinking with Models*. Oxford: Pergamon Press, 1981.
- [67] Selco J I, Beery J L. *Saving a Drug Poisoning Victim*, UMAP ILAP Modules, 2000.
- [68] Shapley L S, Shubik M. *A Method for Evaluating the Distribution of Power in a Committee System*. *American Political Science Review*, 1954, 48.

- [69] Teague D, Doyle D. *The Weather Forecaster Problem*. The Newsletter of the Consortium for Mathematics and Its Applications, Number 90 Spring/Summer 2006.
- [70] Wan F Y M (温耀明). *Mathematical Models and their Analysis*. Harper & Row Pub., 1989.
- [71] Warner S L. *Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias*. *Journal of the American Statistical Association*, 1965, 60 (309): 63 - 69.
- [72] Winston W L. *Operations Research*. 影印版. 北京:清华大学出版社, 2004.
- [73] 《运筹学》编写组. 运筹学. 北京:清华大学出版社, 2005.
- [74] 艾利斯哥尔兹. 变分法. 北京:人民教育出版社, 1958.
- [75] 陈兰荪. 数学生态学模型与研究方法. 北京:科学出版社, 1991.
- [76] 陈珏. 决策分析. 北京:科学出版社, 1987.
- [77] 储志俊, 蒋勇. 短跑跑速变化的数学模型. 体育数学与体育系统工程, 第一卷第二期, 1991.
- [78] 冯元桢. 血液循环. 长沙:湖南科学技术出版社, 1986.
- [79] 何晓群, 刘文卿. 应用回归分析. 北京:中国人民大学出版社, 2001.
- [80] 胡运权. 运筹学习题集. 修订版. 北京:清华大学出版社, 1995.
- [81] 黄润生. 混沌及其应用. 武汉大学出版社, 2000.
- [82] 吉洪诺夫, 萨马尔斯基. 数学物理方程:下册. 黄克欧, 等, 译. 北京:高等教育出版社, 1957.
- [83] 姜启源, 谢金星, 叶俊. 数学模型. 2 版. 北京:高等教育出版社, 1993.
- [84] 姜启源, 谢金星. 数学建模案例选集. 北京:高等教育出版社, 2006.
- [85] 蒋庆琅. 随机过程原理与生命科学模型. 方积乾, 译. 上海:上海翻译出版公司, 1987.
- [86] 刘元亮, 等. 科学认识论与方法论. 北京:清华大学出版社, 1987.
- [87] 史树中. 数学与经济. 大连:大连理工大学出版社, 2008.
- [88] 宋健, 于景元. 人口控制论. 北京:科学出版社, 1985.
- [89] 谭永基, 蔡志杰, 俞文毓. 数学模型. 上海:复旦大学出版社, 2005.
- [90] 王莲芬, 许树柏. 层次分析法引论. 北京:中国人民大学出版社, 1990.
- [91] 吴翊, 李永乐, 胡庆军. 应用数理统计. 长沙:国防科技大学出版社, 1995.
- [92] 姜启源, 谢金星, 邢文训, 等. 大学数学实验. 2 版. 北京:清华大学出版社, 2011.
- [93] 谢云松, 张志让. 数学实验. 北京:科学出版社, 1999.
- [94] 杨启帆, 边馥萍. 数学模型. 杭州:浙江大学出版社, 1990.
- [95] 尤晓伟, 张恩杰, 张青富. 现代道路交通工程学. 北京:清华大学出版社, 北京交通大学出版社, 2008.
- [96] 周怀梧. 医用生物数学. 北京:人民卫生出版社, 1990.
- [97] 庄天戈. CT 原理与算法. 上海:上海交通大学出版社, 1992.