

"十二五"普通高等教育本科国家级规划教材

数学模型

（第五版）

姜启源　谢金星　叶　俊　编

高等教育出版社

本书第一版荣获
第二届全国高等学校优秀教材优秀奖





数字课程网站

网址：http://abook.hep.com.cn/1239629
http://abook.hep.edu.cn/1239629

数字课程账号　使用说明详见书内数字课程说明页

数字课程资源目录

ISBN 978-7-04-049222-4

定价：56.00 元



“十二五”普通高等教育本科国家级规划教材

数学模型

Shuxue Moxing

（第五版）

姜启源 谢金星 叶 俊 编

高等教育出版社·北京

内容提要

本书第一版至第四版分别出版于1987年、1993年、2003年和2011年。基于编者长期从事数学建模和数学实验教学、数学建模竞赛组织和辅导，始终关注国内外数学建模教学案例收集与研究的经验，第五版在保持前四版基本结构和风格的基础上，进行增删与修订，新增和改编的案例生动新颖、内涵丰富，接近案例总数的一半。全书纸质内容与数字化资源一体设计、紧密配合，数字课程包括案例精讲、基础知识、拓展阅读、更多案例、程序文件、数据文件等教学资源，在提升课程教学效果的同时，为学生学习提供了更广的思维与探索空间，便于学生自主学习。

本书可作为高等学校各专业学生学习数学建模课程的教材和参加数学建模竞赛的辅导材料，也可供科技工作者参考。

图书在版编目(CIP)数据

数学模型 / 姜启源，谢金星，叶俊编. --5版. --北京：高等教育出版社，2018.5

ISBN 978-7-04-049222-4

Ⅰ.①数… Ⅱ.①姜… ②谢… ③叶… Ⅲ.①数学模型 Ⅳ.①O141.4

中国版本图书馆CIP数据核字(2018)第003860号

策划编辑 李艳馥 李晓鹏　责任编辑 李晓鹏　封面设计 王 鹏　版式设计 范晓红
插图绘制 杜晓丹　责任校对 刘丽娴　责任印制 耿 轩

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
印　　刷 北京鑫海金澳胶印有限公司
开　　本 787 mm×1092 mm　1/16
印　　张 26
字　　数 590千字
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
网上订购 http://www.hepmall.com.cn
　　http://www.hepmall.com
　　http://www.hepmall.cn
版　　次 1987年4月第1版
　　2018年5月第5版
印　　次 2018年5月第1次印刷
定　　价 56.00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 49222-00

数学模型

（第五版）

姜启源

谢金星

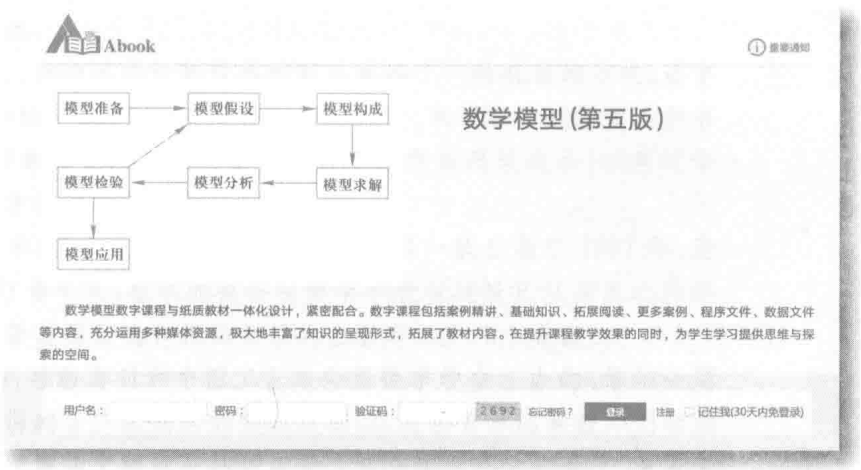
叶　俊

1 计算机访问http://abook.hep.com.cn/1239629，或手机扫描二维码、下载并安装Abook应用。

2 注册并登录，进入“我的课程”。

3 输入封底数字课程账号（20位密码，刮开涂层可见），或通过Abook应用扫描封底数字课程账号二维码，完成课程绑定。

4 单击“进入课程”按钮，开始本数字课程的学习。



课程绑定后一年为数字课程使用有效期。受硬件限制，部分内容无法在手机端显示，请按提示通过计算机访问学习。

如有使用问题，请发邮件至abook@hep.com.cn。



李大潜院士
谈建模

案例精讲

更多案例

拓展阅读

http://abook.hep.com.cn/1239629

第五版前言

半个多世纪以来，由于数学科学与计算机技术的紧密结合，形成了一种普遍的、可以实现的关键技术——数学技术，成为当代高新技术的一个重要组成部分和突出标志，“高技术本质上是一种数学技术”的提法，已经得到越来越多人们的认同，建模与算法正在成为数学这门基础学科从科学向技术转化的主要途径。与此同时，时代发展和科技进步的大潮推动数学科学迅速进入了经济、人口、生物、医学、环境、地质等领域，一些交叉学科如计量经济学、人口控制论、生物数学、数学地质学等应运而生，为数学建模开拓了许多新的、广阔的用武之地。

将数学建模引入大学教育，为数学和外部世界的联系提供了一种有效的方式，让学生能亲自参加将数学用于实际的尝试，参与发现和创造的过程，取得在传统课堂里和书本上无法获得的宝贵经验和切身感受，在知识、能力及素质等方面迅速成长，这是近年来规模最大、最成功的一项数学教学改革实践。

数学建模课程于20世纪80年代初进入我国大学，本书第一版出版于1987年，是我国第一本数学建模教材。30多年来，数学建模课程教学和教材建设已从星星之火迅速发展成燎原之势，这期间起着重要影响和推动作用的主要有以下几个方面：

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，得到了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，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，讲授数学建模课和数学主干课教师之间的不断交流，使得在教学上得以相互借鉴与促进。

3. 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。以计算机为教学工具的数学实验课程的开设，对数学建模的教学内容和方法既有很大启示，也提出了新的挑战。

4. 信息技术在课堂教学中的广泛应用，给传统的教学模式带来了一场革命，也为改进、丰富数学建模课程的教师授课方式和学生学习方式提供了新的平台。

本书编者长期从事数学建模和数学实验教学、数学建模竞赛组织和辅导，并经常关注国内外数学建模教学案例的收集与研究，在保持前四版基本结构和风格的基础上，第五版充分考虑学生学习方式改变的新趋势，与数字化资源进行有机结合，以“纸质教材+数字课程”的方式，对教材的内容和形式进行重新设计，进行了以下几项较大的增补、修订与删减：

1. 案例研究是数学建模教学的主要形式，精心选取并适时更新一些生动新颖、内

涵丰富的案例，才能吸引学生浓厚的学习兴趣，有效地培养学生数学建模的意识、方法和能力，是建设数学建模教材的关键。第五版增加了33个新案例（包括8个融入案例的全国大学生数学建模竞赛赛题），改编了7个老案例，合计接近全书案例总数89个的一半。

2. 为保持全书的容量基本不变，删减了第四版的29个案例（其中7个案例删减了部分内容），全部放入数字课程的“更多案例”中。与此相配合，撤销第四版第7章稳定性模型，将其中3个案例分别放入第五版第5、第6章，撤销第四版第12章马氏链模型，将其中2个案例放入第五版第8章，撤销第四版第13章动态优化模型，将其中1个案例放入第五版第5章。

3. 学习数学建模既要阅读、消化别人做的模型，更要自己动手、做一些实际的练习。第五版对习题进行了相应的增减与修订，并分为复习题与训练题两部分。前者放在每一个案例（节）的后面，供该案例复习所用；后者放在每一类型案例（章）的后面，供这类案例的训练所用。习题参考解答也作了相应的修订，将另行出版。

4. 作为纸质教材内容的拓展和补充，数字课程包括案例精讲、基础知识、拓展阅读、更多案例、程序文件和数据文件等教学资源，在正文中用 标出。其中，案例精讲共10篇，为相关教学案例的视频讲授；基础知识共10篇，是学习本书一些案例所用到的数学原理、方法的简单介绍；拓展阅读共10篇，是对书中相关案例的进一步研究；更多案例共29篇，全部来自本书第四版被删减的内容；程序文件和数据文件为案例和习题中求解模型用到的部分计算程序及过于庞大的原始数据。数字课程的内容将定期进行调整和补充。

本书第4章、7.8节及10.1至10.4节由谢金星编写，9.2至9.7节由叶俊编写，其余各章节由姜启源编写，全书由姜启源统稿。

萧树铁先生1983年在清华大学首次为本科生讲授数学模型课，接着在20世纪80年代中期他先后主持召开了两次数学建模教学、教材研讨会，主持举办了3届教师培训班，培养了我国最早一批数学建模骨干教师，是我国数学模型课程的开拓者和奠基人。萧先生直接筹划了本书第一版的出版，并为第一、二、三、四版作序。萧树铁先生为我国教育事业的发展做出了重大贡献，为了表达对他不幸离世的深切追思和沉痛悼念，这里将萧先生为本书第一版作的序言重新刊出。纵观30多年数学建模课程教学、教材建设和竞赛活动发展的历程，我们从这篇序言可以清晰地看到，萧先生在将数学建模引进我国大学课堂的初创时期，就展现出令人敬佩的远见卓识。李大潜院士十分关注数学建模课程建设和竞赛活动，多次就数学建模对推动科技发展和社会进步、促进教育改革的重要意义发表演讲。本次修订，李先生将他在“走近数学——数学建模篇”MOOC中的一段讲解赠予我们，作为本书数字课程的开篇。在此谨向李大潜院士致以谢意。

我们向在使用本书过程中提出宝贵建议的教师们致谢，向通过学习本书从而喜爱数学建模的同学们致意，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努力。

编　者

2017年9月

第一版序


第二版序


第三版序


第四版序

近十年来，我国高等理工科学校所设置的应用数学专业（系）在数量上有较大的增长，然而直到现在，在如何培养这个专业学生的问题上，却还没有来得及进行认真的研究。

通过一些传统的数学基础课对应用数学专业的学生进行比较严格的数学训练仍然是重要的。但如果要把他们培养成为“会应用”的应用数学工作者，似乎还应该在另外一些方面使他们得到训练。例如，第一，面对一个不太复杂的现象，经过一番初步分析后，能较快地抓住问题的主要方面；可以名之为洞察力。第二，能从两个表面看上去没有多少联系的现象中找出共同点而加以类比；这是一种想象力。在传统的大学数学基础课里面，并非没有对这些方面的要求，然而似乎还没有哪一门课程把进行这类训练作为自己的主要任务，只是近年来国外出现的“数学模型”一课也许不在此列。

为了在这方面进行探索，1982 年秋天我们（姜启源、谭泽光、葛玉安和我）组织了一个“数学模型”讨论班，围绕 E. A. Bender 的书：“*An Introduction to Mathematical Modeling*”，搜集资料，进行分析讨论。1983 年春和暑假中，我分别为清华应用数学系三、四年级和大连的“数学模型讲习班”试讲了“数学模型”课。此后国内有些应用数学系也开设了这门课，其中姜启源同志在清华应用数学系讲了两次。他在讲授中又补充了一些新的材料，尤其是他花了很大的功夫把讲的这些材料以学生较易接受的形式编成讲义，成了这本书的基础。

要培养学生的洞察力和想象力，靠一本教科书当然是不够的，至少还需通过教师根据不同的背景和学生的实际情况来灵活运用书中的内容。因此，在多种不同风格的教材出现以前，一本教材的弹性是不可少的。编者在这方面也作了不少的努力，其中包括选了一批有趣的习题。它们除了可供学生动手练习（这是学习这门课必不可少的环节）以外，有一些也可以作为讲授或讨论的内容。

姜启源同志的这本书是一个很有益的尝试，希望通过它的出版，能促使更多这方面不同风格的教材出现。

萧树铁

1986 年 6 月于清华园

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　包饺子中的数学 …… 5
1.4　建模示例之二　路障间距的设计 …… 7
1.5　建模示例之三　椅子能在不平的地面上放稳吗 …… 9
1.6　数学建模的基本方法和步骤 …… 10
1.7　数学模型的特点和分类 …… 12
1.8　怎样学习数学建模课程和参加竞赛 …… 14
第 1 章训练题 …… 18
第 2 章　初等模型 …… 19
2.1　双层玻璃窗的功效 …… 19
2.2　划艇比赛的成绩 …… 21
2.3　实物交换 …… 23
2.4　汽车刹车距离与道路通行能力 …… 25
2.5　估计出租车的总数 …… 28
2.6　评选举重总冠军 …… 33
2.7　解读 CPI …… 39
2.8　核军备竞赛 …… 47
2.9　扬帆远航 …… 50
2.10　节水洗衣机 …… 52
第 2 章训练题 …… 55
第 3 章　简单的优化模型 …… 57
3.1　存贮模型 …… 57
3.2　森林救火 …… 61
3.3　倾倒的啤酒杯 …… 63
3.4　铅球掷远 …… 67
3.5　不买贵的只买对的 …… 71
3.6　血管分支 …… 78
3.7　冰山运输 …… 81
3.8　影院里的视角和仰角 …… 85
3.9　易拉罐形状和尺寸的最优设计 …… 90
第 3 章训练题 …… 93

第3章简单的优化模型
3.1存贮模型…
3.2森林救火
3.3倾倒的啤酒杯
3.4铅球掷远
3.5不买贵的只买对的……
3.6血管分支
3.7冰山运输3.8影院里的视角和仰角3.9易拉罐形状和尺寸的最优设计第3章训练题

第 4 章 数学规划模型 …… 96
4.1 奶制品的生产与销售 …… 96
4.2 自来水输送与货机装运 …… 104
4.3 汽车生产与原油采购 …… 108
4.4 接力队的选拔与选课策略 …… 113
4.5 饮料厂的生产与检修 …… 117
4.6 钢管和易拉罐下料 …… 120
4.7 广告投入与升级调薪 …… 125
4.8 投资的风险与收益 …… 130
第 4 章训练题 …… 135
第 5 章 微分方程模型 …… 141
5.1 人口增长 …… 141
5.2 药物中毒急救 …… 150
5.3 捕鱼业的持续收获 …… 153
5.4 资金、劳动力与经济增长 …… 156
5.5 香烟过滤嘴的作用 …… 159
5.6 火箭发射升空 …… 163
5.7 食饵与捕食者模型 …… 168
5.8 赛跑的速度 …… 173
5.9 万有引力定律的发现 …… 177
5.10 传染病模型和 SARS 的传播 …… 180
第 5 章训练题 …… 188
第 6 章 差分方程与代数方程模型 …… 192
6.1 贷款购房 …… 192
6.2 管住嘴迈开腿 …… 195
6.3 市场经济中的物价波动 …… 199
6.4 动物的繁殖与收获 …… 203
6.5 信息传播 …… 208
6.6 原子弹爆炸的能量估计 …… 212
6.7 CT 技术的图像重建 …… 217
6.8 等级结构 …… 220
6.9 中国人口增长预测 …… 225
第 6 章训练题 …… 230
第 7 章 离散模型 …… 233
7.1 汽车选购 …… 233
7.2 职员晋升 …… 240
7.3 厂房新建还是改建 …… 247
7.4 循环比赛的名次 …… 251
7.5 公平的席位分配 …… 253

7.6 存在公平的选举吗 …… 259
7.7 价格指数 …… 267
7.8 钢管的订购和运输 …… 271
第 7 章训练题 …… 275
第 8 章 概率模型 …… 279
8.1 传送系统的效率 …… 279
8.2 报童的诀窍 …… 281
8.3 航空公司的超额售票策略 …… 284
8.4 作弊行为的调查与估计 …… 287
8.5 轧钢中的浪费 …… 290
8.6 博彩中的数学 …… 293
8.7 钢琴销售的存贮策略 …… 300
8.8 基因遗传 …… 302
8.9 自动化车床管理 …… 306
第 8 章训练题 …… 309
第 9 章 统计模型 …… 312
9.1 孕妇吸烟与胎儿健康 …… 312
9.2 软件开发人员的薪金 …… 321
9.3 酶促反应 …… 325
9.4 投资额与生产总值和物价指数 …… 331
9.5 冠心病与年龄 …… 336
9.6 蠓虫分类判别 …… 342
9.7 学生考试成绩综合评价 …… 348
9.8 艾滋病疗法的评价及疗效的预测 …… 357
第 9 章训练题 …… 363
第 10 章 博弈模型 …… 369
10.1 点球大战 …… 369
10.2 拥堵的早高峰 …… 373
10.3 “一口价”的战略 …… 379
10.4 不患寡而患不均 …… 382
10.5 效益的合理分配 …… 385
10.6 加权投票中权力的度量 …… 389
第 10 章训练题 …… 398

参考文献 …… 401

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做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的数学模型，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 节给出一些怎样学习数学建模的建议，并对正在蓬勃发展的大学生数学建模竞赛及如何参加竞赛作简要介绍.

1.1　从现实对象到数学模型



什么是数学建模

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；科技展览厅里，港珠澳大桥模型雄伟壮观，长征系列运载火箭模型高高耸立，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高速铁路和新辟的航线，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；体验厅里，智能机器人通过语音识别、触摸交互、移动互联，帮助人们完成各种生活服务活动；“数字故宫”让珍贵文物随着手指滑动屏幕，从沉睡中活了过来，实现 360 度全视角展示.

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，火箭、电站、大桥、铁路……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、视频、图表、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做出相应的决策.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做出必要的简化假设(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**,做出一些必要的**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**.

与数学建模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下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.

复习题

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[24,35].

(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.

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.

半个多世纪以来,随着数学以空前的广度和深度向一切领域的渗透,以及电子计算机和信息技术的飞速发展,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大多是已有的,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.

无论是发展互联网通信、航天、微电子、人工智能等高新技术本身,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,在临床医疗、无损探测、资源勘探等领域被广泛采用的CT(计算机断层成像)技术就是典型的成功范例.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.**"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、物流管理等,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,以数字化为特征的大数据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".

复习题

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.

1.3 建模示例之一 包饺子中的数学

不少人认为需要用数学方法解决的基本上是高新技术、科学研究或者生产建设、经济管理中的重大问题,带着一些神秘色彩的数学离人们的日常生活很远.其实,通过数学建模可以分析我们身边的许多现象和问题.为了让数学走进生活,使大家更容易地了解什么是数学建模,本节和下面两节展示日常生活中三个实例的建模过程,并简单介绍数学建模的方法和步骤.

在最平凡不过的包饺子当中还有什么数学问题吗?让我们从一个具体例子说起[70].

通常,你家用 1 kg 面和 1 kg 馅包 100 个饺子.某次,馅做多了而面没有变,为了把馅全包完,问应该让每个饺子小一些,多包几个,还是每个饺子大一些,少包几个? 如果回答是包大饺子,那么如果 100 个饺子能包 1 kg 馅,问 50 个饺子可以包多少馅呢?

讨论 有人对饺子数量减少一倍就能多包约 40% 的馅这一结果表示怀疑,认为饺子越大面皮应该越厚,建模中关于面皮厚度不变的假设值得探讨,这是有道理的!(复习题 2)

问题分析 很多人都会根据"大饺子包的馅多"的直观认识,觉得应该包大饺子.但是这个理由不足以令人信服,因为大饺子虽然包的馅多,但用的面皮也多,这就需要比较馅多和面多二者之间的数量关系.利用数学方法不仅可以确有道理地回答应该包大饺子,而且能够给出数量结果,回答比如"50 个饺子可以包多少馅"的问题.

首先,把包饺子用的馅和面皮与数学概念联系起来,那就是物体的体积和表面积.用 V 和 S 分别表示大饺子馅的体积和面皮面积,v 和 s 分别表示小饺子馅的体积和面皮面积,如果一个大饺子的面皮可以做成 n 个小饺子的面皮,那么我们需要比较的是,V 与 nv 哪个大? 大多少?

模型假设 容易想到,进行比较的前提是所有饺子的面皮一样厚,虽然这不可能严格成立,但却是一个合理的假定.在这个条件下,大饺子和小饺子面皮面积满足

$$S=ns \tag{1}$$

为了能比较不同大小饺子馅的体积,所需要的另一个假设是所有饺子的形状一样,这是又一个既近似又合理的假定.

模型建立 能够把体积和表面积联系起来的是半径.虽然球体的体积和表面积与半径才存在我们熟悉的数量关系,但是对于一般形状的饺子,仍然可以引入所谓"特征半径"R 和 r,使得

$$V=k_1R^3,\quad S=k_2R^2 \tag{2}$$

$$v=k_1r^3,\quad s=k_2r^2 \tag{3}$$

成立.注意:在所有饺子形状一样的条件下,(2)和(3)中的比例系数 k_1 相同、k_2 也相同.

在(2)和(3)中消去 R 和 r,得

$$V=kS^{\frac{3}{2}},\quad v=ks^{\frac{3}{2}} \tag{4}$$

其中 k 由 k_1 和 k_2 决定,并且两个 k 相同.现在只需在(1)和(4)的 3 个式子中消去 S 和 s,就得到

$$V=n^{\frac{3}{2}}v=\sqrt{n}\,(nv) \tag{5}$$

(5)式就是包饺子问题的数学模型.

小结 回顾整个建模过程,关键的有以下几点:

1. 用数学语言(体积和表面积)表示现实对象(馅和面皮).

2. 做出简化、合理的假设(面皮厚度一样,饺子形状一样).

3. 利用问题蕴含的内在规律(体积、表面积与半径间的几何关系).

实际上,在数学建模中这样几条都是基本的和关键的步骤.

结果解释 模型(5)不仅定性地说明 V 比 nv 大(对于 $n>1$),大饺子比小饺子包的馅多,而且给出了定量结果,即 V 是 nv 的 $\sqrt{n}$ 倍.由此能够回答前面提出的"100 个饺子能包 1 kg 馅,50 个饺子可以包多少馅"的问题,因为饺子数量由 100 变成 50,所以 50 个饺子能包 $\sqrt{100/50}=\sqrt{2}\,(\approx 1.4)$ kg 馅.不用数学建模,你想不到这个结果吧.

复习题

1. 利用模型(5)式说明:如果 n_1 个饺子包 m kg 馅,那么 n_2 个饺子能包多少馅? 由此给出本节中 $\sqrt{2}$ 的结果.

2. 将所有饺子面皮一样厚的假设改为饺子越大面皮越厚,并对此给以简化、合理的数学描述,重

新建模，得出 V 与 nv 之间的关系，讨论“饺子数量减少一倍能多包多少馅”与什么因素有关.

1.4 建模示例之二 路障间距的设计

在校园、机关、居民小区的道路中间，常常设置用于限制汽车速度的路障.看到路障你会想到下面这个问题能够用数学解决吗？

路障之间相距太远，起不到限制车速的作用，相距太近又会引起行车的不便，所以应该有一个合适的间距.不妨向设计者提出这样的问题：如果要求限制车速不超过 40 km/h，路障的间距应该是多少？[70]

问题分析 设计者可以设想，当汽车通过一个路障时，速度近乎零，过了路障，司机就会加速，当车速达到 40 km/h 时，让司机因为前面有下一个路障而减速，至路障处车速又近乎零，如此循环，即可达到限速的目的.

按照这种分析，如果认为汽车在两个相邻路障之间一直在作等加速运动和等减速运动，那么只要确定了加速度和减速度这两个数值，根据基本的物理知识，就很容易算出两个相邻路障之间应有的间距.

收集数据 要得到汽车的加速度和减速度，一个办法是查阅资料，通常可以查到：某牌号的汽车在若干秒内（或多远距离内）可以从静止加速到多快，或者某牌号的汽车在若干秒内（或多远距离内）可从多大车速紧急刹车.由这样的数据推算出的是最大加速度和最大减速度，不能直接用于这里的问题.还有资料会给出加速度的一个范围，如从 1 m/s^2 到 10 m/s^2，也不方便使用.

比较实用的方法是进行测试，办法是请驾驶普通牌号汽车的司机在与设计路障的环境相似的道路上，模拟有路障的情况作加速行驶和减速行驶，记录行驶中的车速和对应的时间（需要坐在副驾位置的助手辅助）.假定我们已经得到了如表 1、表 2 的数据.

表 1 加速行驶的测试数据

速度/($km \cdot h^{-1}$)	0	10	20	30	40
时间/s	0	1.6	3.0	4.2	5.0

表 2 减速行驶的测试数据

速度/($km \cdot h^{-1}$)	40	30	20	10	0
时间/s	0	2.2	4.0	5.5	6.8

模型假设 汽车通过路障时车速为零，其后作等加速运动，当车速达到限速时立即作等减速运动，到达下一个路障时车速为零.

模型建立 记汽车加速行驶的距离为 s_1，时间为 t_1，加速度为 a_1，减速行驶的距离为 s_2，时间为 t_2，减速度为 a_2，限速为 v_{max}.根据熟知的物理定律有

$$s_1=\frac{1}{2}a_1t_1^2,\quad s_2=\frac{1}{2}a_2t_2^2 \tag{1}$$

$$v_{max}=a_1t_1,\quad v_{max}=a_2t_2 \tag{2}$$

汽车在两相邻路障间行驶的总距离 $s=s_1+s_2$，从(1)，(2)中消去 t_1，t_2 得到

$$s=\frac{v_{\max}^2}{2}\left(\frac{1}{a_1}+\frac{1}{a_2}\right) \tag{3}$$

(3)式为路障间距设计的数学模型.对于某个给定限速 $v_{\max}$ 的具体问题，由测试数据估计出加速度 a_1 和减速度 a_2 后，即可计算路障间距 s.

模型计算 以速度 v 为横坐标、时间 t 为纵坐标，将表1、表2的数据作散点图(图1、图2中的圆点)，可以看出速度与时间大致为线性关系，这也能用于验证汽车在路障间作等加速运动和等减速运动的假设是否基本正确.记等加速和等减速运动速度与时间的关系分别为 $t=c_1v+d_1$ 和 $t=c_2v+d_2$.作为粗略估计，不妨手工在图1上尽量靠近数据点画一条直线，因为直线在 $v=0$ 和 $v=40$ 的 t 的坐标差约为 $5.3-0.3=5$，所以图1直线的斜率为 $c_1=\dfrac{5\times3.6}{40}=0.45(\mathrm{s^2/m})$①.类似地处理图2，得到直线的斜率为 $c_2=-\dfrac{7\times3.6}{40}=-0.63(\mathrm{s^2/m})$.由图1、图2可知 d_1，d_2 很小，视为0，于是加速度 $a_1=\dfrac{1}{c_1}$，减速度 $a_2=-\dfrac{1}{c_2}$.

又按问题要求

$$v_{\max}=40\ \mathrm{km/h}=11.1\ \mathrm{m/s}$$

将 a_1，a_2，$v_{\max}$ 代入(3)式计算得

$$s=\frac{11.1^2\times(0.45+0.63)}{2}\approx66.5(\mathrm{m})$$

可将路障间距设计为67 m.

小结 本节建模过程的关键除了做出简化、合理的假设(等加速和等减速行驶)及利用问题蕴含的内在规律(时间、距离、速度、加速度之间的物理关系)以外，还有根据测试数据估计模型的参数(加速度和减速度)，这也是建模中常用的方法.

实际确定路障间距时还要考虑路口及居民楼、教室的位置等因素，那就不是这个模型所能讨论的了.

如果根据表1、表2的数据利用最小二乘法编程计算，可得 $c_1=0.4536$，$c_2=-0.6084$，$s=65.5556$，与手算结果相差不大.

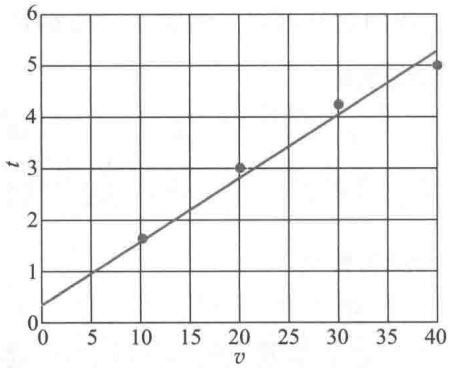


图1 加速行驶测试数据图形

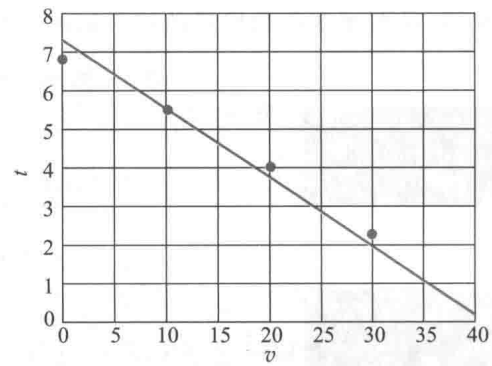


图2 减速行驶测试数据图形

复习题

1. 通过资料调查或实地测试，自己获取汽车加速度和减速度的数据，用来求解模型.
2. 关于路障，你还能想到有哪些问题可以用数学建模分析和解决吗?

① 式中的3.6是速度单位由 km/h 换算成 m/s 的结果.

1.5 建模示例之三 椅子能在不平的地面上放稳吗

这个问题来源于日常生活中一个普通的事实:把四只脚的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[40].

模型假设 对椅子和地面应该做一些必要的假设:

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 是对椅子本身合理的简化.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,比如在有台阶的地方是无法使四只脚同时着地的.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.

模型建立 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.

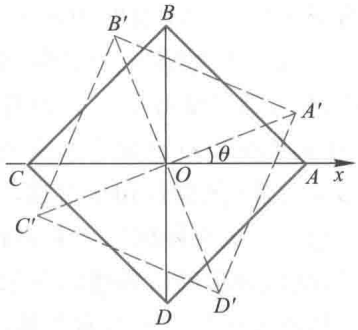


图 1 变量 θ 表示椅子的位置

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 $g(\theta)$($f(\theta),g(\theta)\geqslant 0$).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 $g(0)=0,f(0)>0$,而当椅子旋转 90°后,对角线 AC 与 BD 互换,于是$f(\pi/2)=0,g(\pi/2)>0$.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 θ,$f(\theta)\cdot g(\theta)=0$,且 $g(0)=f(\pi/2)=0,f(0)>0,g(\pi/2)>0$.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

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.

一种.

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.

复习题

将假设条件1中的“四脚的连线呈正方形”改为“四脚的连线呈长方形”，试建立模型并求解.

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9章统计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.

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.

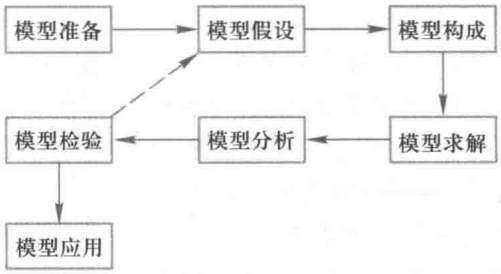


图 1 数学建模步骤示意图

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,做出必要的、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.假设做得不合理或太简单,会导致错误的或无用的模型;假设做得过分详细,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做出恰当的折中.通常,做假设的依据,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 根据所做的假设,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、图的模型等.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的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.

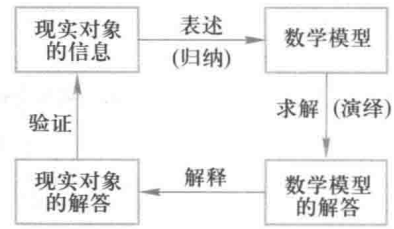


图2 数学建模的全过程

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、做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[75].

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.

复习题

对于1.1节的复习题，考虑建立模型的基本方法，并对建模的具体步骤做出计划.

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[35].

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”之间做出折中和抉择.

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.

模型的非预制性 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.模型的这种非预制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.建模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 怎样学习数学建模——学习课程和参加竞赛



怎样学习数学建模

有人说,数学建模与其说是一门技术,不如说是一门艺术.大家知道,技术一般是有章可循的,许多工程领域都有专门的技术规范,只要严格按照规范去做,事情就可以完成得八九不离十.而艺术通常无法归纳出几条一般的准则或方法,一位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.从这样的意义上说,数学建模更接近艺术,其含义是,目前尚不能找到若干法则或规律,用以完成不同领域各种问题的建模.

这样看来,学习数学建模与学习一般的数学课程会有较大的不同.多数建模课程和建模教材主要讲解的是数学建模过程的一头一尾,即如何将实际问题表述成数学模型和如何用模型的结果分析、解释实际现象,至于建模过程的中段,即模型的求解部分,由于所用的数学方法基本上是大家已经学过的,就不会是讲授的重点.当然,如果用到数学公共课没有涉及的内容,则会有适当的补充.

培养数学建模的意识和能力

在学习数学建模的过程中,与掌握一些建模方法、补充一些数学知识相比,更为重要、也许更加困难的是培养数学建模的意识和能力.

所谓数学建模的意识是指,对于我们日常生活和工作中那些需要或者可以用数学工具分析、解决的实际问题,能够敏锐地发现并从建模的角度去积极地思考、研究.这些问题可能有几种不同的情况,一种是必须用数学方法才能解决的,需要做的大概就是如何建模了;第二种是虽然已经用工程的或经验的办法处理,但是再用上数学方法可能解

决得更好,这就需要你勇于和善于利用数学建模这个工具了;第三种是依靠经验和常识就能得到满意的处理,不一定要用建模解决的问题,而你尝试从数学的角度去考虑,可以起到提高数学建模能力的作用,1.3 节中包饺子和 1.4 节中设计路障就是这样的问题.总之,希望你的头脑里时刻保持从数学建模角度对实际问题做定量分析的一根弦.

至于数学建模的能力,内容很广泛,大体上包含以下内容:

想象力　在已有形象的基础上,根据新的信息在头脑中创造出新形象的能力,是一种形象思维活动,可以通过细心观察、善于联想、勇于突破思维定势(如运用逆向思维、发散思维)等方式来培养.

洞察力　透过现象看到本质,对复杂事物进行分析和判断的能力,需要集中注意力、刻苦工作、勤于思考,适应和创造浓厚的学术环境,培养对科学研究敏锐的"感觉".

类比法　由一类事物具有的某种属性,推测类似事物也具有这种属性的推理方法.因为同一个数学模型可以描述不同领域的对象,所以联想、类比是建模中常用的方法.当然,类比的结果是否正确,需要由实践来检验.

较广博的数学知识　对于已经学过微积分、线性代数、概率论等基础课的大学生来说,还需要了解数值计算、数理统计、数学规划、数值模拟等学科中那些直接与应用相关的内容,如怎样用这些方法建模、怎样在计算机上用数学软件实现、怎样解释模型求解的结果等,并不需要深入了解这些学科的数学原理.

深入实际调查研究的决心和能力　与埋头撰写纯数学论文不同,做数学建模首先要有到实际中去的决心和勇气,并且善于调查研究,掌握第一手材料.

从培养意识、提高能力的角度来学习数学建模,基本上是"学别人的"和"做自己的"两条途径,下面给出一些具体建议.

案例研究——学习、分析、评价、改进和推广

案例研究是学习数学建模课程的主要方法之一,这与法学、管理学的学习方法有些相近.学习课堂上讲的和教材里写的别人做过的模型,不能听过、看过就算学了,应该通过学习、分析、评价、改进和推广几个步骤来逐步深化.首先是弄懂它,再分析它为什么这样做,哪些地方是关键,评价有什么优缺点,最后,如果可能的话,去尝试改进、推广它.

让我们以 1.5 节的案例"椅子能在不平的地面上放稳吗"为例,看看在学懂的基础上可以做哪些研究(有些是应该做的).

1. 对模型假设进行分析

假设 1 对椅子本身的要求大致是合理的,椅脚与地面点接触可看作几何抽象.假设 2 相当于要求地面高度是数学上的连续函数.假设 3 排除了三只脚不能着地的情况,对地面的要求更高.

由建模过程可以知道,假设 2 的地面连续性和假设 3 的三只脚着地是证明命题的必要条件,从常识看这对放稳椅子也是必需的,而假设 1 的椅脚正方形只是为了得到简单的中心对称,对证明命题不是必要的.事实上,很容易将这个条件放宽到椅脚连线为长方形或梯形的情况.还可以对椅子四脚连线的几何形状能放宽到什么程度进行讨论.

2. 对建模的关键进行分析

建模的关键是用数学符号、数学语言把椅子的位置及四只脚着地的条件和结论表

示出来.利用椅脚的中心对称性,用单变量 θ 表示椅子的位置,用两个函数 $f(\theta)$, $g(\theta)$ 表示椅脚与地面的距离,从而得到四脚着地的条件和结论的数学关系.

3. 对建模过程做深入研究

如果读者对上面如此简单的建模过程仔细考虑一下,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: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,首先保持旋转轴始终通过 O 点但允许发生倾斜,使得 A,D 脚先着地,然后再让 B,C 两脚中至少一个着地(此时保持 A,D 脚位置不变,但旋转轴可能不再通过 O 点).

参加数学建模竞赛,做自己的模型

学习数学建模,如果只停留在学别人的模型上,是远远不够的,一定要亲自动手,踏踏实实地做几个实际题目.不妨从包饺子这样的简单问题开始,读者也可以利用本书训练题中提供的那些需要自己做出假设、构造模型的题目,我们更提倡读者在实际生活中发现、提出问题,建立模型.

参加数学建模竞赛为同学们通过亲手做课题,更快地提高用数学建模方法分析、解决实际问题的能力,搭建了广阔的平台.大学生数学建模竞赛是近年来我国高校蓬勃开展的一项学科竞赛活动,有全国性的、地区性的,更有院校自己组织的.各种层次建模竞赛的内容、组织方式、评判标准等都大致相同.下面对竞赛的特点、参赛的过程和收获作简单介绍.

竞赛内容 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,结果的分析和检验以及自我评价优缺点等方面的学术论文.

竞赛方式 采取通信竞赛办法,三名大学生为一队,在三天时间内完成,可以使用任何资料和软件,唯一的限制是不能与队外的同学、老师讨论赛题(包括在网上).

评判标准 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、表述的清晰性为标准.其中结果正确指的是与做出的假设和建立的模型相符合.

全国大学生数学建模竞赛从 1992 年开始举办,规模由最初的几十所学校、几百个队,发展到 2017 年的 1 400 多所院校、36 000 多个队,参加地区、院校竞赛的人数更多.数学建模竞赛之所以受到广大同学的欢迎,主要是由于它的内容、形式和评判标准具有明显区别于大家熟悉的数学等学科竞赛的特点,适合培养有创新精神和综合素质人才的需要.

参加数学建模竞赛的过程一般可以分为三个阶段:

赛前准备 通过教师讲授、学生自学、讨论等方式了解数学建模的基本概念、方法和步骤,学习求解模型的一些数学方法(数值计算、数理统计、数学规划、数值模拟等)及数学软件,还可以演练往年的赛题和模拟题.尽可能地选择不同专业、不同特质(认真踏实、思路灵活、写作好)的同学组队,并在准备过程中进行充分磨合.

三天竞赛 选定赛题既要抓紧时间又需充分讨论,一定要把题目的要求吃透,并拟定一个三人、三天工作的初步计划.开始时不要把目标定得过高,不妨考虑用多数时间、

多数人力保证基本模型的完成,有余力再锦上添花.论文写作要尽早开始,用清晰、完整的文章结构指导全队分工合作、有条不紊地完成任务.

赛后继续 应该结合三天的参赛总结一下整个竞赛准备过程的收获及经验教训.如果有兴趣和精力对赛题作进一步钻研,或者对赛题进行扩展研究,可以寻求教师和其他同学共同参加.更希望通过竞赛激发同学参与数学建模相关实践活动的兴趣.

20 多年来的事实说明,只要认真参加竞赛,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切实的锻炼.赛题通常要用到几门数学和计算机课程及多方面的知识,对于长期一门课一门课学习的同学来说,这种训练运用综合知识的机会是难得的,对于同学独立工作能力的培养有很大的好处.

其次,合作精神与团队意识会得到培养和提高.竞赛需要 3 个人相互启发、争辩和相互妥协、合作,与以后经常面临的集体工作方式十分相近,对于一直在读书、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、通畅的阐明建模思路、假设、方法、结果等内容的论文,是参赛成果的集中体现,竞赛有益于文字表述能力的锻炼;赛题的实用性有助于培养同学们关注社会生活、理论联系实际的学风;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.

许多同学表示,不管三天竞赛的成绩如何,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"来总结竞赛经历.

复习题

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 抵达 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.

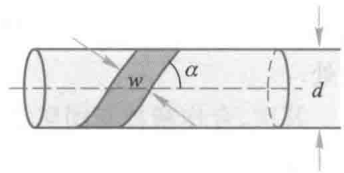
(5) 一男孩和一女孩分别在离家 2 km 和 1 km 且方向相反的两所学校上学,每天同时放学后分

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[40].

第1章训练题

1. 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(如右图).若知道管道长度,需用多长布条(不考虑或考虑两端的影响).如果管道是其他形状呢?[17]



2. 雨滴匀速下降,假定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.

3. 参考更多案例 1-1 商人们怎样安全过河中的状态转移模型,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,并使渡河次数尽量地少.

4. 大包装商品比小包装商品便宜是人所共知的事实(当然指单位质量或体积),你能仅从包装成本的对比对此做出解释吗?

更多案例……

1-1 商人们怎样安全过河

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，其中只有 2.5 节涉及简单的随机现象.

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　双层玻璃窗的功效



案例精讲 2-1
双层玻璃窗的功效

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[45].

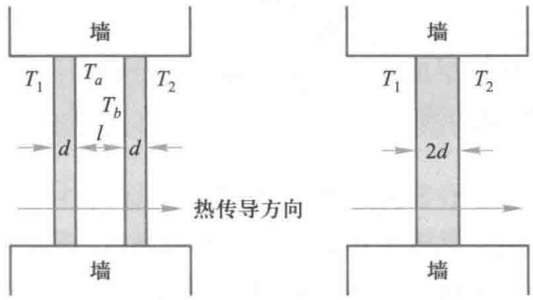


图 1　双层玻璃窗与单层玻璃窗

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

$$Q = k\frac{\Delta T}{d} \tag{1}$$

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

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

从(2)式中消去 T_a，T_b，可得

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

对于厚度为 $2d$ 的单层玻璃窗，容易写出其热量传导为

$$Q_2 = k_1\frac{T_1-T_2}{2d} \tag{4}$$

二者之比为

$$\frac{Q_1}{Q_2} = \frac{2}{s+2} \tag{5}$$

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 $k_1 = 4\times10^{-3}\sim8\times10^{-3}$ J/(cm·s·K)，不流通、干燥空气的热传导系数 $k_2 = 2.5\times10^{-4}$ J/(cm·s·K)，于是

$$\frac{k_1}{k_2} = 16\sim32$$

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

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

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.

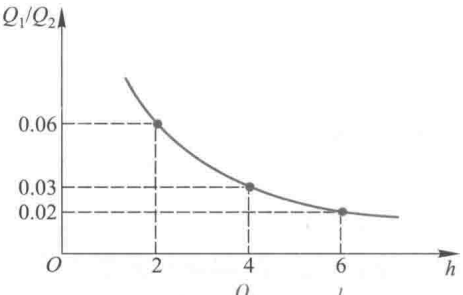


图 2 热量损失比 $\frac{Q_1}{Q_2}$ 与 $h=\frac{l}{d}$ 的关系

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.

复习题

北方旧宅改造时,为了增强保暖效果常用保温材料在外墙外面再加一层墙.假定仍然只考虑热传导,试通过建模对改造后减少的热量损失给出定量分析,并获取相关数据作简单计算.

2.2 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[7].

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	6.34	11.75	0.574	20.5	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做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 是根据所给数据做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

$$np \propto fv \tag{1}$$

由假设 2,3，

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

代入(1)式可得

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

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

$$s \propto A^{\frac{2}{3}} \tag{3}$$

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

$$w' \propto n \tag{4}$$

而由阿基米德定律，艇排水体积 A 与总质量 w' 成正比，即

$$A \propto w' \tag{5}$$

(3),(4),(5)式给出

$$s \propto n^{\frac{2}{3}} \tag{6}$$

将(6)式代入(2)式，当 w 是常数时得到

$$v \propto n^{\frac{1}{9}} \tag{7}$$

因为比赛成绩 t（时间）与 v 成反比，所以

$$t \propto n^{-\frac{1}{9}} \tag{8}$$

(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.

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

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

$$t = \alpha n^{\beta} \tag{9}$$

其中 α,β 为待定常数.由(9)式

$$\ln t = \alpha' + \beta \ln n, \alpha' = \ln\alpha \tag{10}$$

利用最小二乘法，根据所给数据拟合上式①，得到

$$t=7.21n^{-0.111} \tag{11}$$

可以看出(8)式与这个结果吻合得相当好.

结果与实际数据吻合得如此之好，恐怕有很大成分的巧合.

复习题

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的大约好 5%.

2.3 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一个模型[7].

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 坐标系中长方形 $0\leqslant x\leqslant x_0$，$0\leqslant y\leqslant y_0$ 内任一点的坐标 (x,y) 都代表了一种交换方案(图 1).

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

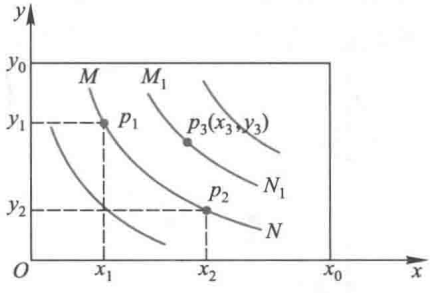


图 1 甲的无差别曲线

$$f(x,y)=c_1 \tag{1}$$

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，无差别曲线应是下降的(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

$$g(x,y)=c_2 \tag{2}$$

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

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.

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.

有了双方的无差别曲线,交换方案的范围可从整个长方形缩小为一条曲线 AB,但仍不能确定交换究竟应在曲线 AB 上的哪一点进行.显然,越靠近 B 端甲的满意度越高而乙的满意度越低,靠近 A 端则反之.要想把交换方案确定下来,需要双方协商或者依据双方同意的某种准则,如等价交换准则.

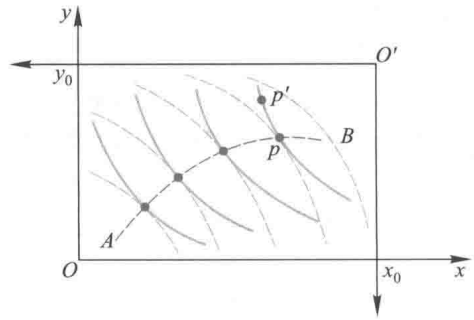


图 2 双方的无差别曲线和交换路径

等价交换准则是指:两种物品用同一种货币衡量其价值,进行等价交换.不妨设交换前甲占有的 x_0(物品 X)与乙占有的 y_0(物品 Y)具有相同的价值,x_0,y_0 分别相应于图 3 中 x 轴,y 轴上的 C,D 两点,那么在直线 CD 上的点进行交换,都符合等价交换准则(为什么?).最后,在等价交换准则下,双方满意的交换方案必是 CD 与 AB 的交点 p.

评注 提出无差别曲线的概念是用图形方法建立实物交换模型的基础,确定这种曲线需要收集大量的数据,还可以研究曲线的解析表达式,在 3.5 节我们将用微积分的方法再次讨论它.另外,如果双方交换的物品是三种呢,可否提出无差别曲面的概念,这时交换路径还是一条曲线吗?

最后,对无差别曲线呈下凸形状作如下解释:当人们占有的 x 较小时(p_1 点附近),他宁愿以较多的 Δy 交换较少的 Δx(图 4),而当占有的 x 较大时(p_2 点附近),就要用较多的 Δx 换取较少的 Δy,正是"物以稀为贵"这一哲理的写照.满足这种特性的曲线是下凸的.

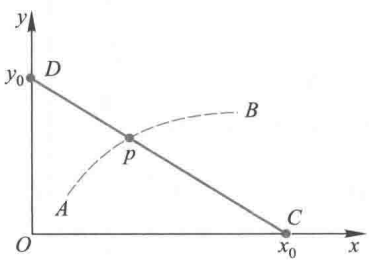


图 3 等价交换准则确定的交换方案

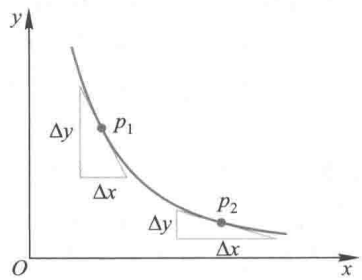


图 4 无差别曲线下凸形状的解释

复习题

用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个协议(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[7].

2.4 汽车刹车距离与道路通行能力

在现代城市生活中,提高道路通行能力是城市交通工程面临的重要课题之一.常识告诉我们,车辆速度越高、密度越大,道路通行能力就越大.但是速度高了,刹车距离必然变大,车辆密度将受到制约,所以需要对影响通行能力的因素进行综合分析.本节首先引入交通流的主要参数及基本规律,然后介绍汽车刹车距离与道路通行能力两个模型[83].

交通流的主要参数及基本规律

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向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

$$q=vk \tag{1}$$

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 100km/h×(1 辆/0.2 km)= 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,司机考虑到刹车距离必然被迫降低车速.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

$$v=v_f(1-k/k_j) \tag{2}$$

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.

将线性模型(2)代入(1),得到流量与密度的关系

$$q=v_fk(1-k/k_j) \tag{3}$$

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

$$q=k_jv(1-v/v_f) \tag{4}$$

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)~(4)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.

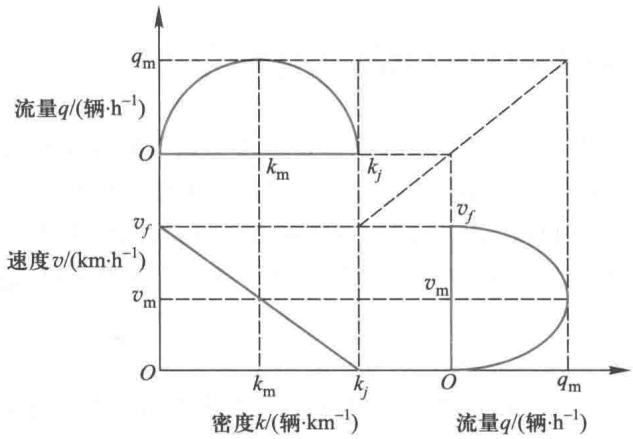


图1 交通流的流量、速度、密度关系图

汽车刹车距离模型

汽车司机在行驶过程中发现前方出现突发事件，会紧急刹车，人们把从司机决定刹车到车完全停止这段时间内汽车行驶的距离，称为刹车距离.车速越快，刹车距离越长.刹车距离与车速之间是线性关系吗？让我们先做一组实验：用固定牌子的汽车，由同一司机驾驶，在不变的道路、气候等条件下，对不同的车速测量其刹车距离，得到的数据如表1和图2所示.

表1 车速和刹车距离的一组数据

车速/(km·h^{-1})	20	40	60	80	100	120	140
刹车距离/m	6.5	17.8	33.6	57.1	83.4	118.0	153.5

从图2可以看到，刹车距离与车速之间并非简单的线性关系，需要对刹车过程的机理进行分析，建立刹车距离与车速之间的数学模型.

问题分析 刹车距离由反应距离和制动距离两部分组成，前者指从司机决定刹车到制动器开始起作用这段时间内汽车行驶的距离，后者指从制动器开始起作用到汽车完全停止所行驶的距离.

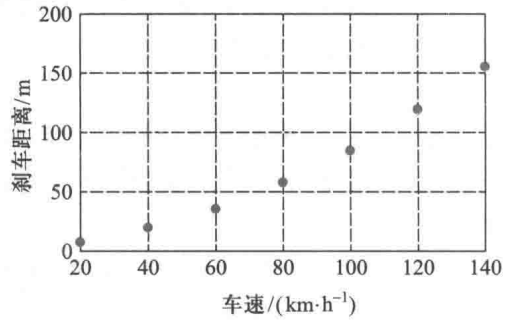


图2 表1的数据(实际刹车距离)

反应距离由反应时间和车速决定，反应时间取决于司机个人状况(灵巧、机警、视野等)和制动系统的灵敏性(从司机脚踏刹车板到制动器真正起作用的时间)，对于固定牌子的汽车和同一类型的司机，反应时间可以视为常数，并且在这段时间内车速尚未改变.

制动距离与制动器作用力、车重、车速以及道路、气候等因素有关，制动器是一个能量耗散装置，制动力作的功被汽车动能的改变所抵消.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，使汽车大致做匀减速运动，司机和乘客少受剧烈的冲击.这里不考虑道路、气候等因素.

模型假设 基于上述分析,作以下假设:

1. 刹车距离 d 等于反应距离 d_1 与制动距离 d_2 之和.

2. 反应距离 d_1 与车速 v 成正比,比例系数为反应时间.

3. 刹车时使用最大制动力 F,F 做的功等于汽车动能的改变,且 F 与车的质量 m 成正比.

模型建立 由假设 2,

$$d_1 = c_1 v \tag{5}$$

c_1 为司机反应时间.由假设 3,在 F 作用下行驶距离 d_2 做的功 Fd_2 使车速从 v 变成 0,动能的变化为 $mv^2/2$,有 $Fd_2 = mv^2/2$,又 F 与 m 成正比,按照牛顿第二定律可知,刹车时的减速度 a 为常数,于是

$$d_2 = c_2 v^2 \tag{6}$$

c_2 为比例系数,实际上 $c_2 = 1/(2a)$.由假设 1,刹车距离为

$$d = c_1 v + c_2 v^2 \tag{7}$$

即刹车距离 d 与车速 v 之间是二次函数关系.

参数估计 估计模型(7)式中的参数 c_1,c_2 通常有两种方法,一是查阅交通工程学的相关资料,二是根据实测数据对(7)式作拟合.按照交通工程学提供的数据,司机反应时间 c_1 为 0.7~1 s,系数 c_2 约为 0.01 m · $h^2/km^2$①.数据拟合的方法将留作复习题 1.

道路通行能力模型

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,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.通行能力表示道路的容量,交通流量表示道路的负荷,交通流量与通行能力的比值反映了道路的负荷程度,称饱和度或利用率.

按照流量、速度和密度的基本关系(1)式,当车辆以一定速度行驶时,道路通行能力应该在安全条件下车辆密度最大,即前后两车车头之间的距离(下称车头间隔)达到最小来计算,而车头的最小间隔主要由刹车距离决定.

记车速为 v (km/h),最小车头间隔为 D (m),道路通行能力为 N (辆/h),显然

$$N = 1000v/D \tag{8}$$

交通工程学中通常将最小车头间隔分解为 $D = d + d_0$,其中 d(m)是刹车距离,由(7)式给出,d_0(m)是车身的标准长度与两车间的安全距离之和,取一固定值.于是

$$N = \frac{1000v}{c_1 v + c_2 v^2 + d_0} = \frac{1000}{c_1 + c_2 v + \dfrac{d_0}{v}} \tag{9}$$

由(9)式可知,当车速 v 一定时道路通行能力 N 与 c_1,c_2 和 d_0,即道路、车辆、司机等的状况有关.那么,在道路等状况不变时多大的车速可以使通行能力 N 达到最大呢? 显然只需(9)最后一个等号分式的分母最小.利用初等数学知识就能够得到,当车速 v =

① $c_2 = 0.01$ m · h^2/km^2 相当于次高级路面潮湿状态且车速 40 km/h 时的数值,当路面状态、车速改变时,c_2 的大小将有变化.

$\sqrt{\frac{d_0}{c_2}}$时，通行能力 N 达到最大值 N_m 为

$$N_m=\frac{1000}{c_1+2\sqrt{c_2 d_0}} \tag{10}$$

在利用(9)和(10)式计算时需注意 v, c_1, c_2, d_0 这些量的单位.

复习题

1. 由表1的数据用最小二乘法拟合模型(7)，计算参数 c_1, c_2，对数据和拟合曲线作图，并估计刹车时的减速度.

2. 采用第1题得到的 c_1, c_2，或者按照交通工程学提供的数据，适当地设定 d_0(可取车身标准长度的1.5~2倍)，对于不同的车速(20~100 km/h)，利用(9)式计算道路通行能力，并由(10)式分析各参数对最大通行能力的影响.

2.5 估计出租车的总数

站在城市的街头，看着一辆辆汽车驶过身旁，眼疾手快地记下一些汽车号码，拿着这些杂乱无章的数字，有人用它来选择处于两难境地的决策(如看到第一辆车是单号今天就去炒股，是双号就不炒)，有人把它当作与朋友打赌的“骰子”(如单号你请客买单，双号朋友买单).做这样的事情人们脑子里想着的基本前提是：在这座城市的千万辆汽车中，出现任何一辆的机会都一样，所以看到的汽车号码是随机的.其实这也是大家的共识.

根据这样的共识，我们能做一些有一定意义的事情.例如，已经知道一座小城市出租车的牌号是从某一个数字0101开始按顺序发放的，你随意地记下了驶过的10辆出租车牌号：0421, 0128, 0702, 0410, 0598, 0674, 0712, 0529, 0867, 0312，根据这些号码能否估计这座城市出租车的总数[33]？

问题分析 把随意记下的10个号码从小到大重新排列，记作 $x_1, x_2, \cdots, x_{10}$，已知的起始号码记作 x_0，未知的终止号码记作 x，可以在数轴上表示出来(图1).$x_1, x_2, \cdots, x_{10}$ 可以看作是从 $[x_0, x]$ 区间内全部整数值(称为总体)中取出的一个样本，问题是根据样本和 x_0 对总体的 x 做出估计.显然，出租车总数为 $x-x_0+1$.

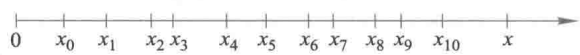


图1 号码在数轴上的图示

模型建立 为了记号的简单、方便，将起始号码 x_0 设定为0001，未知的终止号码仍记作 x，全部号码为模型的总体，记作 $\{0001, 0002, \cdots, x\}$，$x$ 即是出租车总数.从总体中随机抽取的 n 个号码为模型的样本，从小到大排列后记作 $\{x_1, x_2, \cdots, x_n\}$.需要建立由 $x_1, x_2, \cdots, x_n$ 估计 x 的模型，其基本假定是，样本中每个 $x_i (i=1,2,\cdots,n)$ 取自总体中任一号码的概率相等.由样本估计总体的方法有很多，下面给出几个简单的模型.

模型 1——平均值模型

记样本 x_1, x_2, …, x_n 的平均值为 $\bar{x}$,总体 0001, 0002, …, x 的平均值为 $\bar{X}$,显然,$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$,$\bar{X}=\frac{1}{x}\sum_{j=1}^{x}j=\frac{x+1}{2}$,用 $\bar{x}$ 作为 $\bar{X}$ 的估计值即 $\bar{X}=\bar{x}$,于是得到

$$x=2\bar{x}-1 \tag{1}$$

$\bar{x}$较大时可以忽略(1)式右端的 1,所以总数 x 可近似用样本平均值$\bar{x}$的 2 倍来估计.

模型 2——中位数模型

记样本和总体的中位数分别为$\tilde{x}$,$\tilde{X}$,用$\tilde{x}$作为$\tilde{X}$的估计值,由于$\tilde{X}=\bar{X}=\frac{x+1}{2}$,所以

$$x=2\tilde{x}-1 \tag{2}$$

x 也可近似用样本中位数$\tilde{x}$的 2 倍估计.

模型 3——两端间隔对称模型

样本中最小号码与总体的起始号码的间隔(即其中包含的除样本号码之外的号码个数)是 x_1-1,样本中最大号码与总体的终止号码的间隔是 $x-x_n$,假定样本的最小值与最大值在总体中对称,即两个间隔相等:$x-x_n=x_1-1$,所以

$$x=x_n+x_1-1 \tag{3}$$

x 可近似用样本最大值 x_n 与最小值 x_1 之和来估计.

模型 4——平均间隔模型

将利用两端间隔的想法给以推广,把起始号码和样本排成数列:1, x_1, x_2, …, x_n, 考察相邻两数之间的 n 个间隔 x_1-1, x_2-x_1-1, …, $x_n-x_{n-1}-1$,取这些间隔的平均值,作为 x_n 与终止号码 x 间隔 $x-x_n$ 的估计,即 $x-x_n=\frac{1}{n}\left[(x_1-1)+\sum_{i=2}^{n}(x_i-x_{i-1}-1)\right]=\frac{1}{n}(x_n-n)$,由此可得

$$x=\left(1+\frac{1}{n}\right)x_n-1 \tag{4}$$

x 可近似用样本最大值的 $1+\frac{1}{n}$倍来估计.

模型 5——区间均分模型

将总体所在区间[1, x]平均分成 n 份,每个小区间长度为$\frac{x-1}{n}$,假定样本中每个 x_i ($i=1,2,\cdots,n$)都位于小区间的中点,将区间均分,那么 x_n 与 x 的距离$x-x_n$应是小区间长度的一半,即 $x-x_n=\frac{x-1}{2n}$,由此可得

$$x=\left(1+\frac{1}{2n-1}\right)\left(x_n-\frac{1}{2n}\right) \tag{5}$$

x 可近似用样本最大值的 $1+\frac{1}{2n}$倍来估计.

利用以上模型(1)~(5)计算后,都应取靠近的整数作为出租车总数 x 的估计值.

计算与分析　对于你最初随意地记下的 10 辆出租车牌号,将起始号码 x_0 由 0101

平移为0001后,得到的样本(称第1样本)为{0321, 0028, 0602, 0310, 0498, 0574, 0612, 0429, 0767, 0212},用5个模型分别计算出的 x 列入表1第2行,估计的出租车总数最少是模型3得到的794,最多是模型2得到的926,差别很大.哪个准确呢?

让我们再随机记下10个牌号,平移0001后得到的第2样本为{0249, 0739, 0344, 0148, 0524, 0284, 0351, 0089, 0206, 0327},同样计算的 x 列入表1第3行,发现不仅各个模型估计的不同,而且与第1样本相比,同一模型估计的也相差较大.特别值得注意的是,模型1、模型2估计出的 x 为651和610,都小于样本中的最大值739,显然是不合理的.

表1 根据2个样本用5个模型估计的出租车总数

	模型1	模型2	模型3	模型4	模型5
第1样本	870	926	794	843	807
第2样本	651	610	827	812	778

但是,因为不知道出租车总数实际上到底是多少,所以无法从这样的结果判断哪个模型更准确.并且,样本是从总体中随机抽取的,从少数几个样本得出的结果随意性很大.下面我们通过大量的数值模拟来对这些模型的准确度进行评价.

数值模拟 这里的数值模拟是指给定一个总体,从中抽取若干样本,根据5个模型分别对样本进行计算,估计总体,将估计结果与给定总体作对比,根据各个模型的估计值与总体比较的结果做出评价.

对于总体{1, 2, …, x},设定 $x=1\,000$,从总体中随机取 $n=10$ 个数为一个样本,对每个样本用5个模型分别估计 x(估计值记作 $\hat{x}$),如此取 $m=200$ 个样本,计算由 m 个样本估计的 $\hat{x}$ 的平均值和标准差,以及平均值与真值(1 000)间的误差.为了进一步分析 m 个样本估计的 $\hat{x}$ 的分布情况,可以画出 $\hat{x}$ 的直方图.

这样的模拟做了2次,数值结果如表2和表3,直方图如图2和图3①.

表2 第1次模拟的数值结果

	模型1	模型2	模型3	模型4	模型5
$\hat{x}$ 的平均值	1 023.2	1 037.4	1 010.0	1 005.6	962.3
平均值的误差	23.2	37.4	10.0	5.6	-37.7
$\hat{x}$ 的标准差	170.1	261.0	126.3	90.9	87.0

表3 第2次模拟的数值结果

	模型1	模型2	模型3	模型4	模型5
$\hat{x}$ 的平均值	986.5	985.4	980.8	992.9	950.1
平均值的误差	-13.5	-14.6	-19.2	-7.1	-49.9
$\hat{x}$ 的标准差	181.4	271.1	107.9	86.6	82.8

① 由于模拟中的样本是通过伪随机数程序产生的,每次运行都会得到不同的样本,所以表2、表3、图2、图3的结果是不可重复的.

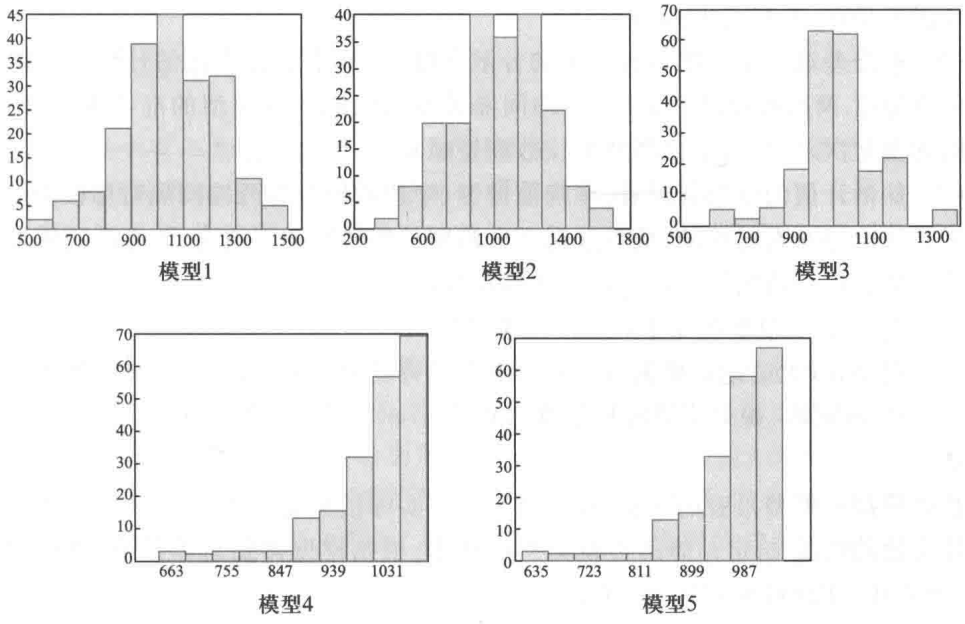


图 2 第 1 次模拟的直方图

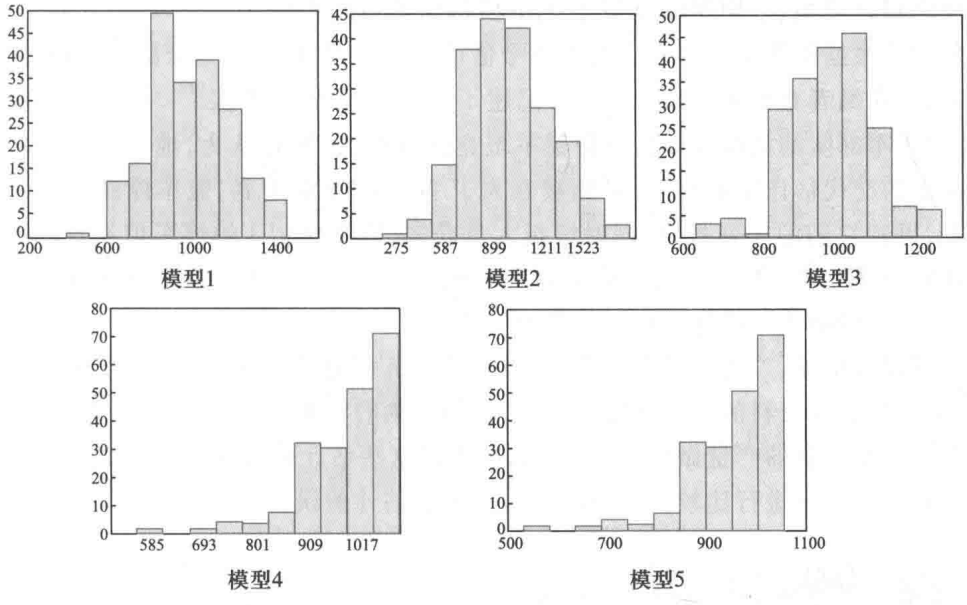


图 3 第 2 次模拟的直方图

根据两次模拟的结果可以得出以下的初步结论：

(1) 平均值模型和中位数模型是容易想到、容易理解的方法，并且从理论上讲，对于均匀分布的总体，样本平均值和中位数都是总体平均值和中位数的无偏估计.但是，从有限个样本得到的总体平均值的估计误差却很大.两端间隔对称模型是无偏的，估计误差也不小.区间均分模型不仅估计误差较大，而且有过低估计的趋势.平均间隔模型

的估计误差最小，并且是无偏的①.

(2) 标准差表示了多次估计结果的分散程度，在一定意义上比估计误差更能反映模型的准确度.两次模拟结果表明，平均间隔模型和区间均分模型的准确度最高，两端间隔对称模型次之，平均值模型和中位数模型最差.

(3) 从估计值的分布情况看，平均值模型、中位数模型和两端间隔对称模型的估计值对于平均值是左右对称的，显然这与3个模型的构造形式相符合.平均间隔模型和区间均分模型的估计值则呈左低右高的非对称型.

综合地看，5个模型中以平均间隔模型最优.

需要指出，以上结论是根据编者的两次模拟得到的，不排除读者通过模拟得出不同结论的可能.特别地，如果模拟时增加样本的大小 n 和样本的数量 m，应该能得到更可信的结论.

评注 为了解决本节开始提出的问题，建立了5个模型，除平均值模型和中位数模型用了一点统计学的知识外，其他3个模型大体上来自直观的常识，而有点出人意料的是，后者较前者更优.相信读者还会提出自己的模型，不一定需要充足的理由.

在建模时将起始号码作为已知条件，5个模型((1)~(5)式)适用于从1开始取整数值的总体，这就限制了应用的范围.请读者思考，其中哪些建模方法可以推广到起始号码未知的情况.

数值模拟是模型检验的重要方法，特别是在实际信息缺乏、需要由少数样本估计未知总体参数的情况.给定总体通过模拟产生样本，根据模型得到总体参数，进行比较和评价，很有些“实践检验真理”的味道.

估计值标准差的大小对于应用的实际价值影响较大，文献[33]给出了(1)~(4)式估计值的(理论)方差(标准差的平方)，从大到小的顺序是：中位数模型、平均值模型、两端间隔对称模型、平均间隔模型.我们的模拟结果与此相符合.

后记 根据随意记下的几个汽车牌号估计出租车数量，似乎是个臆造的问题，而有意思的是，在编者查到的一篇文献中，记述了这样一个历史事实[20].第二次世界大战中，一支盟军部队的指挥部急需掌握德军坦克的数量，指挥官认为，他们根据德军的宣传和战俘的交代估计出的数字，可能被夸大了.随着战争的进展，盟军俘获了若干辆德军坦克，得到它们的序列号码，同时情报人员获知，这支标记I的德军坦克号码是按顺序编排的.于是，就有数学家以俘获的坦克号码为样本，用统计方法估计出标记I的德军坦克总量，帮助盟军指挥部修正了此前的估计.

不仅仅是坦克，二次世界大战期间，英美情报机构通过对捕获德军武器装备上序列编号、生产日期等资料的分析研究，对于从军用车辆的轮胎、枪支、装甲车，直到飞弹、火箭等众多武器装备的产量做出估计.战后，有人将这些估计值与从敌人档案中得到的这些装备的实际产量进行比较，令人惊讶的是，多数估计的误差竟然在10%以内[53].

复习题

1. 对总体{1, 2, …, x}仍然设定 $x=1\,000$，增加样本大小 n 和样本数量 m，用5个模型作模拟计算，估计总体，对估计结果进行分析研究.n 和 m 的增加对结果有什么影响.

2. 如果汽车的起始号码未知，5种建模方法中有哪种方法可加以改进来解决这样的问题？给出具体的估计模型并作模拟.

---

① 样本平均值是总体平均值的无偏估计是指，样本平均值的期望(意思是取无穷多个样本)等于总体平均值.因为现实应用中只能获取有限个(甚至少数几个)样本，所以无偏性没有多大实际意义.平均间隔模型无偏性的证明见文献[33].

2.6 评选举重总冠军

体育运动中有一些项目按照运动员的体重划分级别进行比赛，如举重、赛艇、拳击、摔跤等，因为这些项目在相当程度上依靠运动员全身的力量来完成，而人体的力量与体重有密切的关系，让体重大体相同的运动员相互比赛才能维护竞赛的公平.本节讨论大家比较熟悉、划分级别较多、规则较为简单的举重比赛.

男子举重比赛按照运动员体重(上限)分为 8 个级别:56 kg,62 kg,69 kg,77 kg,85 kg,94 kg,105 kg 和 105 kg 以上.每个级别都设 3 个项目:抓举、挺举和总成绩(抓举与挺举成绩之和).每个级别、每个项目都产生一个冠军，所以一次比赛共有 24 个冠军.大级别冠军的成绩一般都高于小级别冠军，问题是，从同一项目(如抓举)的 8 个冠军中能评选出一个“总冠军”吗？当然不能直接比较他们的成绩，但是如果把不同级别冠军的成绩，都按照各自的体重合理地“折合”成某个标准级别的成绩，对折合成绩进行比较，选出最高的作为总冠军，应该是可以被人们接受的[70].

问题分析　构造折合成绩的前提是建立体重与举重成绩之间的数学模型.如果有了这样的模型，就可以计算出各个级别冠军举重成绩的理论值.当在一次比赛中得到了各个级别冠军成绩的实际值后，计算实际值与理论值的比值.再根据这个比值构造一个简单、合适的数量指标作为折合成绩，各个级别冠军中折合成绩最大者，即可评选为总冠军.

数据收集与分析　大家知道，举重成绩除了与体重有关以外，还在相当大程度上取决于运动员的选拔、训练等因素.在建立举重成绩的数学模型时，为了只保留体重而尽量排除其他因素的影响，一个合适的办法是利用举重比赛的世界纪录.因为作为世界顶级运动员，他们通过严格的选拔和艰苦的训练，在举重技巧的发挥方面已趋完善(当然是在目前的环境条件下，随着时代的发展，运动水平还将提高)，不同级别成绩的差别基本上是由运动员体重决定的.另外，多年积累下来的世界纪录与某一次比赛成绩(如奥运会、世锦赛等)相比，更能避免偶然性.

男子举重比赛世界纪录的情况(截至 2013 年)如表 1 所示，可以看到，中国运动员在小级别比赛中占有较大的优势，保持着 7 项世界纪录.

表 1　男子举重比赛世界纪录(截至 2013 年)

级别	项目	纪录	纪录保持者	日期
56 kg 级	抓举	138 kg	穆特鲁(土耳其)	2001.11.4
	挺举	168 kg	穆特鲁(土耳其)	2001.4.24
	总成绩	305 kg	穆特鲁(土耳其)	2000.9.16
62 kg 级	抓举	153 kg	石智勇(中国)	2002.6.28
	挺举	182 kg	乐茂盛(中国)	2002.10.2
	总成绩	327 kg	金恩国(朝鲜)	2012.7.31

续表

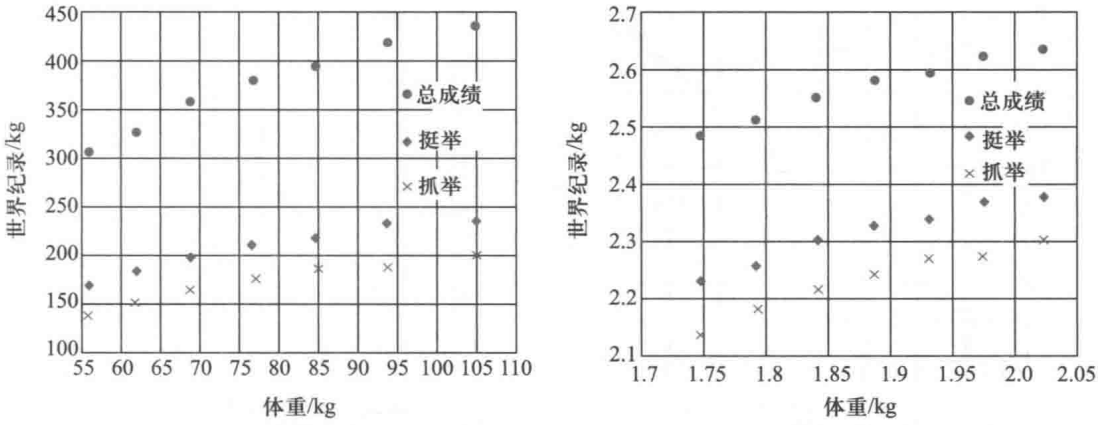
级别	项目	纪录	纪录保持者	日期
69 kg 级	抓举	165 kg	马尔科夫(保加利亚)	2000. 9. 20
	挺举	198 kg	廖辉(中国)	2013. 10. 23
	总成绩	358 kg	廖辉(中国)	2013. 10. 23
77 kg 级	抓举	176 kg	吕小军(中国)	2013. 10. 25
	挺举	210 kg	佩雷佩切诺夫(俄罗斯)	2001. 4. 27
	总成绩	380kg	吕小军(中国)	2013. 10. 25
85 kg 级	抓举	187 kg	雷巴科夫(白俄罗斯)	2007. 9. 22
	挺举	218 kg	张勇(中国)	1998. 4. 25
	总成绩	394 kg	雷巴科夫(白俄罗斯)	2008. 8. 15
94 kg 级	抓举	188 kg	阿卡基奥斯(希腊)	1999. 11. 27
	挺举	233 kg	伊利亚(哈萨克斯坦)	2012. 8. 4
	总成绩	418 kg	伊利亚(哈萨克斯坦)	2012. 8. 4
105 kg 级	抓举	200 kg	安德烈(白俄罗斯)	2008. 8. 18
	挺举	237 kg	特萨加耶夫(保加利亚)	2004. 4. 25
	总成绩	436 kg	安德烈(白俄罗斯)	2008. 8. 18
105 kg 以上级	抓举	214 kg	萨利米(伊朗)	2011. 11. 13
	挺举	263 kg	拉扎扎德(伊朗)	2004. 8. 25
	总成绩	472 kg	拉扎扎德(伊朗)	2000. 9. 26

在处理世界纪录与体重数据时还应注意到,虽然我们不掌握创造纪录的运动员的实际体重,但是基于体重越大、举得越重的常识,比赛时运动员的体重都会调整到非常接近各级别的上限,所以可用每个级别的上限代表运动员的实际体重①.并且,由于105 kg以上级未设上限,更不知道运动员的实际体重,只好略去这个级别,在其余7个级别中评选总冠军.

在定量地研究举重成绩与运动员体重的模型之前,将表1中抓举、挺举和总成绩3个项目的世界纪录与体重数据,以散点图形式用图1表示出来,直观地看一下二者之间大致呈现怎样的关系.图1(a)显示,世界纪录与体重大致上呈线性关系,但是大级别运动员成绩的增加有变慢的趋势.如果将世界纪录和体重先取对数,再用图1(b)表示,可以看出取对数后二者的线性关系有所改进,这启示我们,用幂函数(幂次小于1)比线性函数表示举重成绩与体重的关系可能更为合适.

从表1和图1还可以知道,虽然举重总成绩记录并不等于抓举和挺举记录之和(由于是不同时间、不同运动员创造的),但只是略小一点,3个记录随体重变化的趋势相

① 虽然在同一级别比赛中获得相同最高成绩的运动员以体重最轻者为冠军,但是这与评选各个级别的总冠军无关.



(a) 普通坐标　　(b) 对数坐标

图 1　抓举、挺举和总成绩世界纪录与体重数据图示

同,且总成绩没有像 94 kg 级抓举那样个别点的变异现象.下面只利用举重总成绩的记录来建立模型,其方法完全可用于抓举和挺举.

下面通过 3 个模型研究举重成绩与运动员体重的关系[13].

模型建立　记举重总成绩为 y,运动员体重为 w,建立 y 与 w 的线性模型、幂函数模型和幂函数改进模型.

模型一——线性模型

将举重总成绩 y 与体重 w 的数据重新画一个散点图(图 2,其中 7 个点与图 1 最上方 7 个点相同,只是坐标轴有别),然后尽量接近这些点作一条直线,发现直线与横坐标相交于-60 左右.据此建立如下的线性模型

$$y=k(w+60) \tag{1}$$

式中的系数 k 可以由图 2 直线的斜率粗略地估计:$k\approx\dfrac{430}{160}=2.69$,也可由表 1 的数据利用最小二乘法编程计算:$k=2.703\ 9$,于是(1)式表为

$$y=2.703\ 9(w+60) \tag{2}$$

模型二——幂函数模型

建立这个模型的关键是确定幂函数的幂次,可以从运动生理学的角度研究.

合理地假定举重成绩 y 与运动员身体肌肉的截面积 s 成正比,而 s 与身体的特定尺寸 l 的平方成正比,再假定体重 w 与 l 的立方成正比,即有

$$y=k_1 s,s=k_2 l^2,w=k_3 l^3 \tag{3}$$

k_1, k_2, k_3 为比例系数.从(3)式可以得到

$$y=kw^{\frac{2}{3}} \tag{4}$$

将举重总成绩 y 与 $w^{\frac{2}{3}}$ 的数据画在散点图(图 3)上,然后尽量接近这些点作一条直线,(4)式中的系数 k 可以由图 3 直线的斜率粗略地估计:$k\approx\dfrac{140}{7.5}=18.7$,也可由表 1 数据利用最小二乘法编程计算出 $k=20.471\ 1$,得到

$$y=20.471\ 1w^{\frac{2}{3}} \tag{5}$$

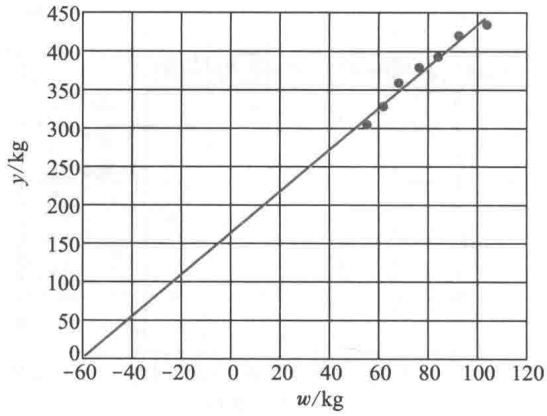


图2 总成绩 y 与体重 w 及直线

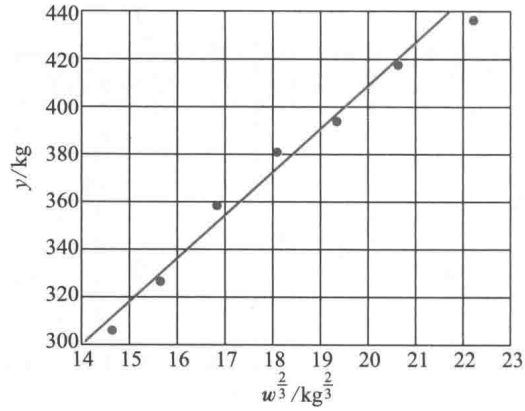


图3 总成绩 y 与 $w^{\frac{2}{3}}$ 及直线

模型三——幂函数改进模型

1967年,Carroll对幂函数模型作了如下改进.首先,他从生理学的角度认为,考虑到举重过程中力量的损失以及身体尺寸的各种变化,(3)式中 y,s,l 之间的关系应改写为

$$y=k_1 s^{\alpha}(\alpha<1),s=k_2 l^{\beta}(\beta<2) \tag{6}$$

并且将体重 w 分为肌肉与非肌肉两部分,记非肌肉部分为 w_0,(3)式中 w 与 l 的关系改写为

$$w=k_3 l^3+w_0 \tag{7}$$

从(6),(7)两式得到

$$y=k(w-w_0)^{\gamma},\gamma<2/3 \tag{8}$$

然后,Carroll利用1964年以前世界范围内50名各级别顶尖运动员的成绩对(8)式作统计分析,得到 $w_0=35$ kg,$\gamma=1/3$,于是有

$$y=k(w-35)^{\frac{1}{3}} \tag{9}$$

略去作图粗略估计(9)式中系数 k 的方法,直接由表1数据利用最小二乘法编程计算,得到

$$y=108.110\,6(w-35)^{\frac{1}{3}} \tag{10}$$

(2)、(5)、(10)式表示的曲线见图4~图6,计算结果以及与实际记录的相对误差分别列入表2第3、4、5列.

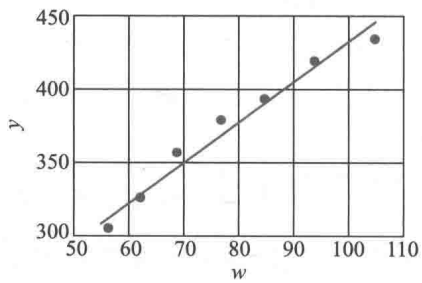


图4 总成绩与体重的线性模型

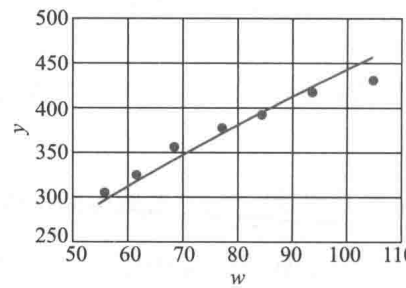


图5 总成绩与体重的幂函数模型

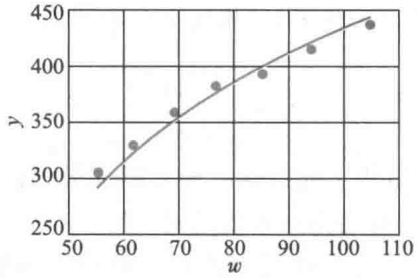


图 6 总成绩与体重的幂函数改进模型

表 2 三个模型的计算结果及与实际记录的相对误差(定义为(实际值-计算值)/计算值)

级别	总成绩纪录	线性模型(2)	幂函数模型(5)	幂函数改进模型(10)
56 kg 级	305 kg	313.648 6(-2.76%)	299.640 5(1.79%)	298.268 9(2.26%)
62 kg 级	327 kg	329.871 8(-0.87%)	320.678 4(1.97%)	324.331 7(0.82%)
69 kg 级	358 kg	348.798 8(2.64%)	344.382 7(3.95%)	350.236 3(2.22%)
77 kg 级	380 kg	370.429 8(2.58%)	370.512 1(2.56%)	375.795 2(1.12%)
85 kg 级	394 kg	392.060 7(0.49%)	395.750 2(-0.44%)	398.282 7(-1.08%)
94 kg 级	418 kg	416.395 5(0.39%)	423.214 4(-1.23%)	420.874 0(-0.68%)
105 kg 级	436 kg	446.138 0(-2.27%)	455.618 6(-4.31%)	445.554 5(-2.14%)

为根据表 2 的结果比较三个模型与实际记录的拟合程度,将每个模型 7 个级别的相对误差取绝对值后加以平均,称为总平均误差,计算得到:线性模型(2)、幂函数模型(5)、幂函数改进模型(10)的总平均误差依次为 1.71%、2.32%、1.47%.看来两个幂函数模型的结果比线性模型并没有多大改进.

观察图 5、图 6 可以发现,如果降低这两个幂函数的幂次,使曲线上升得更缓慢一点,有可能提高曲线与数据点的拟合程度.不妨作如下试验:将(4)式的幂次 2/3 降为 0.6,(9)式的幂次 1/3 降为 0.3,用同样的数据可得

$$y=27.432\ 9w^{0.6} \tag{11}$$

$$y=122.627\ 8(w-35)^{0.3} \tag{12}$$

模型(11)、(12)的总平均误差分别为 1.23%和 0.66%,相比有很大改进.

评选总冠军 以上面的线性模型(1)、幂函数模型(4)和幂函数改进模型(9)为基础,建立评选总冠军的模型,并以模型(1)为例说明建模过程.

将 7 个级别从轻量级到重量级记作 $i=1,2,\cdots,7$,体重(上限)记作 w_i,按照模型(1)各个级别冠军的理论成绩应为 $\hat{y}_i=k(w_i+60)$,在一次比赛中各个级别冠军的实际成绩记作 y_i,比值 $y_i/\hat{y}_i$ 表示第 i 级别冠军在评选总冠军中的实力.取与这个比值成正比(比例系数 λ)的数量指标 z_i 为

$$z_i=\frac{\lambda y_i}{\hat{y}_i}=\frac{\lambda y_i}{k(w_i+60)} \tag{13}$$

为了确定(13)式中的系数 λ/k,任意取一个级别如 77 kg 级($i=4$)为标准,使 $w_4=77$ 时

z_4 正好等于这个级别冠军的实际成绩 y_4，将这个关系代入(13)式得$\lambda/k=77+60=137$，于是(13)式化为

$$z_i=y_i\frac{137}{w_i+60} \tag{14}$$

z_i 是将体重折合成 77 kg 级后第 i 级别冠军的实际成绩，称为折合成绩.按照比赛中 7 个级别冠军的折合成绩排名，排名第一者为总冠军.

(14)式是以线性模型(1)为基础建立的评选总冠军的模型，可以注意到，(14)式的折合成绩 z_i 完全由第 i 级别冠军的实际成绩 y_i 决定(体重上限 w_i 对每个级别是固定的)，与模型(1)中随着世界纪录的不断刷新而改变的系数 k(如(2)式那样)无关，所以(14)式是一个简单的、便于应用的模型.

评注 举重成绩与体重之间的数学模型是评选总冠军方法的基础，我们通过分析举重世界纪录数据，建立了线性模型(1)、幂函数模型(4)和幂函数改进模型(9)，其中模型(1)基于对世界纪录数据的直观观察，(4)和(9)则是从机理分析角度出发，结合数据分析完成建模的.读者可能已经注意到，3 个模型中都只有一个以因子形式出现的系数 k，这是构造折合成绩的需要，大家已经看到，在总冠军模型(14)、(15)、(16)中系数 k 消失了.

按照通常的建模方法，线性模型等 3 个模型的一般形式为 $y=aw+b$，$y=aw^b$，$y=a(w-w_0)^b$，其中系数 a,b,w_0 用最小二乘法由世界纪录确定后，可以得到与实际数据拟合得更好的模型.但是如果这样做，在构造折合成绩时难以将这些系数全部消去，包含系数的总冠军模型不便于应用，因为系数会随世界纪录的刷新而改变.

从幂函数模型(4)和幂函数改进模型(9)出发，仍以 77 kg 级为标准，可以完全类似地推出相应的评选总冠军的模型

$$z_i=y_i\left(\frac{77}{w_i}\right)^{\frac{2}{3}} \tag{15}$$

$$z_i=y_i\left(\frac{42}{w_i-35}\right)^{\frac{1}{3}} \tag{16}$$

让我们用模型(14)、(15)、(16)评选 2008 年北京奥运会男子举重比赛的总冠军.表 3 的前 3 列是各级别举重总成绩冠军的实际资料(105 kg 以上级未列入)，后 3 列是用这 3 个模型计算的折合成绩及排名.可以看到，由模型(15)、(16)评选的前 3 名完全相同，与模型(14)的前 2 名只是稍有出入，综合 3 个模型的结果，可以评选出总冠军是 69 kg 级的廖辉(中国).

表 3 2008 年北京奥运会男子举重比赛各级别总成绩冠军及由模型得到的折合成绩及排名

级别	冠军获得者	总成绩	折合成绩及名次		
			线性模型(14)	幂函数模型(15)	幂函数改进模型(16)
56 kg 级	龙清泉(中国)	292 kg	344.862 1(7)	361.064 4(5)	367.896 9(4)
62 kg 级	张湘祥(中国)	319 kg	358.221 3(6)	368.572 9(3)	369.617 5(3)
69 kg 级	廖辉(中国)	348 kg	369.581 4(2)	374.403 9(1)	373.395 7(1)
77 kg 级	史才秀(韩国)	366 kg	366.000 0(3)	366.000 0(4)	366.000 0(6)
85 kg 级	陆永(中国)	394 kg	372.262 1(1)	368.873 5(2)	371.754 3(2)
94 kg 级	伊利亚(哈萨克斯坦)	406 kg	361.181 8(5)	355.441 3(6)	362.514 3(7)
105 kg 级	阿拉姆诺夫(白俄罗斯)	436 kg	362.012 1(4)	354.558 1(7)	367.736 6(5)

从表 3 还可以看出，利用线性模型得到的第一名与最后一名的折合成绩相差约 30 kg，利用幂函数模型相差约 20 kg，而利用幂函数改进模型相差仅约 10 kg.如果这届奥运会男子举重比赛的各级别冠军都是顶尖选手，那么这也能够从一个方面说明，幂函数改进模型比其他两个模型更合理一些.

复习题

1. 搜集下列举重比赛的实际数据，利用(14)~(16)式计算折合成绩及排名：

(1) 截至目前的男子举重比赛世界纪录.

(2) 最近一届奥运会男子举重比赛成绩.

(3) 截至目前的女子举重比赛世界纪录.

2. 研究一般形式的幂函数模型 $y=aw^b$，利用表 1 的男子举重比赛世界纪录或第 1 题搜集的数据，确定系数 a，b(提示：通过取对数将幂函数化为线性函数)，与幂函数模型(5)式比较与实际记录的拟合程度，再构造计算折合成绩的公式，对搜集的某次举重比赛各级别总成绩冠军进行排名.

2.7 解读 CPI

“你看，最近菜价又涨啦，可电视上说，CPI 回落了，怎样回事?”

“哟！大妈，您还懂 CPI 啊，人家说的是 CPI 涨幅回落，是比过去涨得少点，可还是涨!”

眼下，居民消费价格指数(consumer price index，简称 CPI)差不多全民普及了.每个月的 9 日左右国家统计局会发布上个月全国的 CPI 数据，经过多种媒体的宣传和各界人士的解读，引起老百姓的普遍关心和纷纷议论.

CPI 是反映居民购买生活消费品和服务项目时价格变动趋势的相对数字，是观察通货膨胀水平的重要指标.这里，我们既不从国家经济政策的角度，也不从与其他经济指标的关系方面研讨 CPI 的变化，而是沿着数学建模的思路，按照数据分析的方法去解读 CPI[70].

按照时间顺序解读 CPI

根据国家统计局消费价格统计方法，目前我国 CPI 按照对比基期的不同，有环比价格指数、同比价格指数、累计价格指数等，基期指当前与过去的哪个时期对比.

通常，国家统计局直接公布的不是价格指数，而是价格指数的增长率(以%表示)，为方便人们更直观地了解价格上涨的幅度，表 1 和表 2 给出的就是全国 2011 年和 2012 年 CPI 每个月份 3 个指数的增长率.由这些数据很容易得到价格指数，如表 1 显示 2011 年 1 月至 3 月的环比增长率分别是 1.0%，1.2%，-0.2%，由于基期的价格指数一般取作 100，所以这 3 个月的环比指数分别是 100×(1+0.01)= 101.0(可以简单地将 1.0 加上 100 即得)，100×(1+0.012)= 101.2，100×(1−0.002)= 99.8.同比指数与累计指数可以类似地推算.

表 1 全国 2011 年 CPI 各月份的 3 个价格指数增长率

	月份											
	1	2	3	4	5	6	7	8	9	10	11	12
环比增长率/%	1.0	1.2	-0.2	0.1	0.1	0.3	0.5	0.3	0.5	0.1	-0.2	0.3
同比增长率/%	4.9	4.9	5.4	5.3	5.5	6.4	6.5	6.2	6.1	5.5	4.2	4.1
累计增长率/%	4.9	4.9	5.0	5.1	5.2	5.4	5.5	5.6	5.7	5.6	5.5	5.4

表 2 全国 2012 年 CPI 各月份的 3 个价格指数增长率

	月份											
	1	2	3	4	5	6	7	8	9	10	11	12
环比增长率/%	1.5	−0.1	0.2	−0.1	−0.3	−0.6	0.1	0.6	0.3	−0.1	0.1	0.8
同比增长率/%	4.5	3.2	3.6	3.4	3.0	2.2	1.8	2.0	1.9	1.7	2.0	2.5
累计增长率/%	4.5	3.9	3.8	3.7	3.5	3.3	3.1	2.9	2.8	2.7	2.7	2.6

下面通过表 1 和表 2 的数据来解读 3 个价格指数.

1. 环比价格指数 以上月为基期进行对比的价格指数,简称环比,是最能反映当前价格波动的指标.如表 1 中 2011 年 1 月环比增长率 1.0%,表示 2011 年 1 月比 2010 年 12 月价格上涨 1.0%,2011 年 3 月环比增长率−0.2%,表示 2011 年 3 月比 2011 年 2 月价格下降 0.2%.

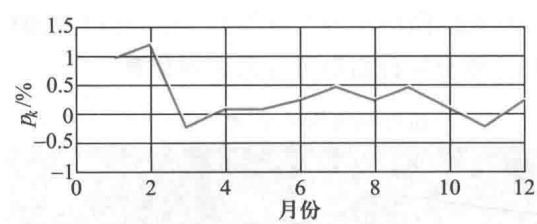
记 p_k 为某年 k 月的环比增长率($k=1,2,\cdots,12$),表 1、表 2 的第 3 行即 2011 年和 2012 年的 p_k,图 1(a)、图 2(a)是按照 2011 年和 2012 年的 p_k 画出的.由表 1 和图 1(a)可知,2011 年只有 3 月和 11 月的环比增长率为负值,即价格比上月下降,而 8 月和 10 月虽然图上曲线下降,表示增长率比上月减少,但是价格仍然是上涨的,公布 CPI 数据时会说"价格涨幅回落".根据表 2 和图 2(a),2012 年有 5 个月价格比上月下降,还有 1 个月价格涨幅回落,物价形势比 2011 年明显好转.

为了得到更直接、清晰的全年价格走势,以上一年 12 月为基期(价格指数 100),可以根据 p_k 计算这一年逐月的价格指数 P_k($k=1,2,\cdots,12$)为

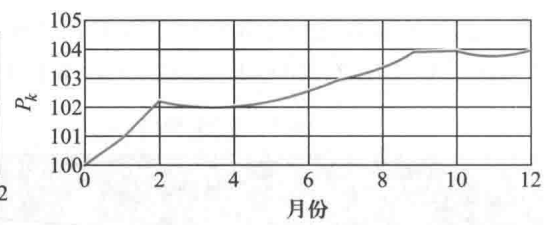
$$P_k = 100\prod_{i=1}^{k}(1+p_i) \tag{1}$$

图 1(b)、图 2(b)是由(1)式计算的 2011 年和 2012 年的 1 至 12 月的价格指数 P_k 的曲线.将图 1、图 2 的(a)图与(b)图对比,可以发现,当 p_k 为负值时 P_k 下降,当 p_k 为正值但下降时 P_k 上升但增量减少.

由图 1(b)、图 2(b)可以看出,2011 年全年的价格走势比 2012 年明显偏高,2011 年 12 月 P_{12} 约为 104,即比 2010 年 12 月约增长 4%,而 2012 年 12 月 P_{12} 约为 102.5,即比 2011 年 12 月约增长 2.5%.



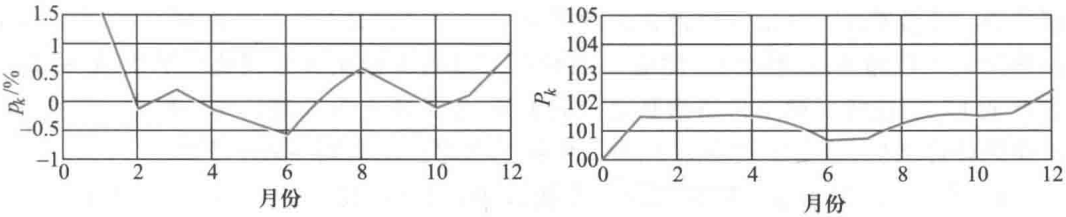
(a) 全国2011年各月份环比增长率p_k



(b) 全国2011年各月份价格指数P_k(上一年12月为100)

图 1

2. 同比价格指数 以上年同月为基期进行对比的价格指数,简称同比.为了消除季节变化和春节、国庆节等节日对价格的影响,今年与上年同月份的价格做比较,也是



(a) 全国2012年各月份环比增长率p_k　　(b) 全国2012年各月份价格指数P_k(上一年12月为100)

图 2

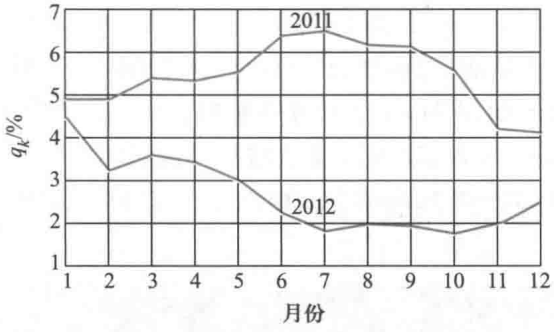


图 3　2011 年和 2012 年各月份同比增长率

有意义的.

记 q_k 为某年 k 月的同比增长率($k=1,2,\cdots,12$), 表 1、表 2 的第 4 行即 2011 年和 2012 年的 q_k,由这些数据画出图 3 的两条曲线.可以明显地看到,虽然这两年每个月的价格比上年同月都有所上涨(增长率为正值),但是 2012 年比 2011 年每月价格上涨的幅度明显减少.

环比和同比之间应该有数量关系.记 $p_k(j)$, $q_k(j)$ 分别为 j 年 k 月的环比指数和同比指数(注意是指数,不是增长率),$x_k(j)$ 为 j 年 k 月的价格指数(基于同一基期,例如以 $j-2$ 年 12 月为基期),按照环比和同比的定义,有

$$p_k(j)=\frac{x_k(j)}{x_{k-1}(j)},q_{k-1}(j)=\frac{x_{k-1}(j)}{x_{k-1}(j-1)},q_k(j)=\frac{x_k(j)}{x_k(j-1)},p_k(j-1)=\frac{x_k(j-1)}{x_{k-1}(j-1)}$$

由此可得

$$p_k(j)q_{k-1}(j)=q_k(j)p_k(j-1) \tag{2}$$

于是若已知(2)式的 $p_k(j-1)$ 和 $q_{k-1}(j)$, $q_k(j)$,可以算出 $p_k(j)$,如由 2011 年的环比指数(表 1 第 3 行数字加上 100)和 2012 年的同比指数(表 2 第 4 行数字加上 100),可计算 2012 年的环比指数(读者可检验,若忽略计算误差,结果应与表 2 第 3 行数字加上 100 相符).

另外,由同比的定义和按照(1)式计算的价格指数 P_k 可以知道,每年 12 月的同比指数与每年的 P_{12}是一回事,从图 1(b)、图 2(b)可以看到,2011 年的 P_{12}等于表 1 给出的 12 月的同比指数 104.1,2012 年的 P_{12}等于表 2 给出的 12 月的同比指数 102.5.

3. 累计价格指数　从 1 月起到发布月份为止累计的、以上年同一期间价格为基期进行对比的指数,简称累计.因为若干个月的价格指数是以其每个月价格指数的平均值

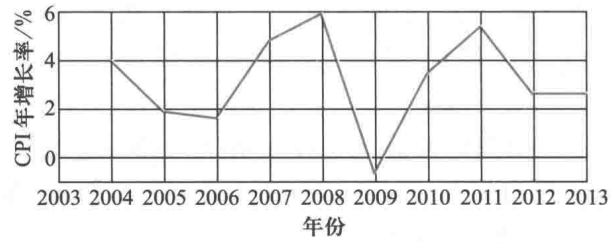
度量的，所以累计可以由同比价格指数得到.1月份的累计与1月份的同比相同，2月份的累计是1月和2月同比的平均值，3月份的累计是1月至3月同比的平均值，依此下去直至12月.这既是两个价格指数之间的关系，也是两个指数增长率之间的关系，对于这样的计算结果，读者不妨用表1、表2中同比和累计的数据核对一下.

4. 年价格增长指数 环比和同比是按月进行统计、比较的，如果要得到年价格增长指数，只需注意到，每年12月的累计为1月至12月同比的平均值，由此可得每一年与上年比较的价格指数增长率.表1、表2右下角的那个数字表示，2011年比2010年价格增长5.4%，2012年比2011年增长2.6%.

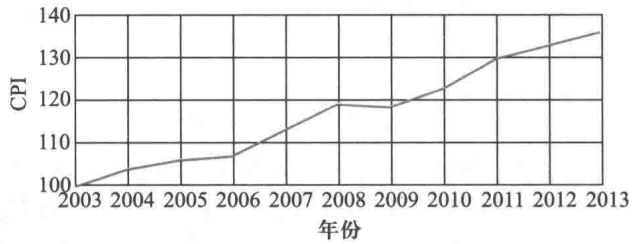
全国2004年至2013年CPI年增长率列入表3第3行，利用(1)式(其中p_i取为年增长率)并以2003年为基期(价格指数100)，计算2004年至2013年的CPI，得表3第4行.这两组数据的图形如图4(a)(b).可以看出，这10年中价格起伏不定，2007年、2008年和2011年上升得较高，2009年呈负增长，可能是对2008年过高增长的调整.总体上看，从2003年到2013年10年全国CPI增长了35%，平均年增长率3.5%.

表3 全国2004年至2013年CPI的增长

	年份										
	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
CPI年增长率/%		3.9	1.8	1.5	4.8	5.9	−0.7	3.3	5.4	2.6	2.6
CPI	100	103.9	105.8	107.4	112.5	119.1	118.3	122.2	128.8	132.2	135.6



(a) 全国2004年至2013年CPI年增长率



(b) 全国2004年至2013年CPI

图4

按照分类结构解读CPI

在按照时间顺序解读我国CPI的过程中我们看到，2011年和2012年的月环比增长率基本上不超过1%，还有近三分之一的月份是负增长，近10年的平均年增长率也不过3.5%.这些数据可能与你和你的家庭对物价的亲身感受有较大差距.原因是多方面

的,其中一个重要因素是,国家统计局公布的 CPI 是对全国居民家庭衣食住行各类消费品和服务价格,采用抽样、统计、综合加工得到的,而个人和家庭可能只关注了个别种类商品的价格变化.

目前,我国在编制 CPI 过程中将消费品和服务项目分为 8 大类,263 个基本分类,约 700 个代表品种.国家统计局在按照《流通和消费价格统计调查方案》采样取得各个品种的价格后,加权计算出各个分类的指数及各个大类的指数,再由各个大类的指数加权计算 CPI,所以 CPI 是由各种消费品和服务项目的价格及其权重二者共同决定的.

权重是国家统计局根据居民家庭用于各种消费品和服务项目的开支占总消费支出的比重确定的,随着人们生活水平的提高及消费结构的变化,权重会进行调整,据称我国每 5 年、10 年会有较大的调整.表 4 是我国现行的消费品和服务项目的大类和若干中类的类别,以及 2011 年调整后的 8 大类的权重①.

表 4 我国消费品和服务项目的类别及权重(2011 年)

	大 类	中 类	权重
1	食品	粮食、油脂、肉禽及其制品、水产品、蛋、鲜菜、鲜果、液体乳及乳制品	31.79%
2	烟酒及用品	烟草、酒	3.49%
3	衣着	服装、鞋	8.52%
4	家庭设备及维修服务	耐用消费品、家庭服务及加工维修服务	5.64%
5	医疗保健个人用品	中药材及中成药、西药、医疗保健服务	9.64%
6	交通和通信	交通工具、车用燃料及零配件、通信工具、通信服务	9.95%
7	娱乐教育文化用品及服务	教育服务、文娱用耐用消费品及服务、文化娱乐类、旅游	13.75%
8	居住	建房及装修材料、住房租金、水、电、燃料	17.22%

由表 4 可以看出,当前在 8 大类消费品和服务项目中以食品权重最大,居住和娱乐教育文化用品及服务次之.一直有不少舆论认为前者的权重仍太大,而后者偏小.实际上,我国 CPI 中的食品权重已由 20 世纪 80 年代的 60%左右下降了很多,近年来每次调整还在下降.另外应该注意到,居住中并不包含近年来飞涨的购买房屋的支出,官方的解释是购房属于投资而非消费②.

用 v 表示 CPI 总水平,v_i, w_i 分别表示第 i 大类消费品和服务项目的价格指数及权重($i=1,2,\cdots,8$),且权重之和等于 1,则 CPI 总水平与 8 大类消费品和服务项目的价格指数及权重的数学模型可表为

$$v=\sum_{i=1}^{8} w_i v_i, \qquad \sum_{i=1}^{8} w_i = 1 \tag{3}$$

① 网上可查到不同的版本,表 4 中的权重来自比较可靠的一个.

② 有的国家如美国将购房支出计入 CPI,致使居住一类的权重在 40%以上.

用 Δv 和 Δv_i 表示 v 和 v_i 的增长率，即 $\Delta v=\dfrac{v-100}{100}$，$\Delta v_i=\dfrac{v_i-100}{100}$，由于模型(3)是线性的，所以有

$$\Delta v=\sum_{i=1}^{8} w_i\Delta v_i,\qquad \sum_{i=1}^{8} w_i=1 \tag{4}$$

(3)或(4)适用于计算每个月 CPI 总水平的环比及同比指数或指数增长率.

一方面，权重 w_i 对 CPI 总水平的大小有很大影响，引起民间对权重数值合理性的关注和研究，另一方面，权重随时调整的具体情况并不能为民间及时掌握，于是，利用每个月官方公布的 CPI 的相关数据，校核原来的权重是否有变化，以及估算调整后的权重，便成为一些研究者和关心者的课题.下面我们根据模型(4)讨论几种校核与估算的方法.

1. 利用每个月公布的环比和同比分类增长率 Δv_i 及原来权重 w_i，按照(4)计算环比和同比 CPI 总水平增长率，检查与每个月公布的 Δv 是否相符.

表 5、表 6 分别是国家统计局公布的 2013 年环比和同比各月份的 Δv_i 和 Δv，用其中的 Δv_i 及表 4 的权重 w_i，按照模型(4)计算各月份 CPI 总水平的环比和同比，将计算值分别列入表 5、表 6 的最后一列，以便与倒数第 2 列公布的 Δv 比较.

这是一个粗糙的方法.如果计算值与公布的 Δv 相符，并不能说明所有的 w_i 没有改变；如果不符，无法确认是否是数字的舍入误差所致，因为公布的指数只有 2 位有效数字，对计算结果可能有很大影响.

表 5 2013 年各月份 CPI 环比分类和总水平增长率 单位：%

月份	食品 Δv_1	烟酒及用品 Δv_2	衣着 Δv_3	家庭设备及维修服务 Δv_4	医疗保健个人用品 Δv_5	交通和通信 Δv_6	娱乐教育文化用品及服务 Δv_7	居住 Δv_8	CPI 总水平 Δv	CPI 总水平计算值
1	2.8	0	-0.4	0.3	0.1	0.1	0.5	0.2	1.0	0.995 7
2	2.7	-0.1	-0.7	0.2	0.3	0.5	0.6	0.3	1.1	1.019 3
3	-2.9	0.1	0.6	0	0	-0.3	-0.4	0.5	-0.9	-0.866 1
4	0.4	0	0.7	0.1	0	-0.4	0.4	0.2	0.2	0.242 1
5	-1.6	-0.1	0.1	0.1	0.1	-0.4	-0.3	0.1	-0.6	-0.552 2
6	0	-0.1	-0.3	0.1	0	0	0	0.1	0	-0.006 2
7	0	0	-0.6	0.1	0	0.3	0.8	0.3	0.1	0.146 0
8	1.2	0	-0.2	0.1	0.1	0.2	0	0.2	0.5	0.434 1
9	1.5	-0.2	1.2	0	0.2	0.2	1.0	0.3	0.8	0.800 5
10	-0.4	0	1.0	0.1	0.1	-0.3	0.8	0.2	0.1	0.087 9
11	-0.2	0.1	0.6	0.1	0.1	-0.2	-0.6	0.2	-0.1	-0.061 7
12	0.6	0	0.1	0.1	0	0.3	-0.1	0.3	0.3	0.272 7

表 6 2013 年各月份 CPI 同比分类和总水平增长率 单位:%

月份	食品 Δv_1	烟酒及用品 Δv_2	衣着 Δv_3	家庭设备及维修服务 Δv_4	医疗保健个人用品 Δv_5	交通和通信 Δv_6	娱乐教育文化用品及服务 Δv_7	居住 Δv_8	CPI 总水平 Δv	CPI 总水平计算值
1	2.9	1.4	2.5	1.5	1.8	-0.3	0.5	2.9	2.0	1.980 2
2	6.0	1.1	2.1	1.6	1.8	0.2	2.0	2.8	3.2	3.165 5
3	2.7	1.0	2.3	1.6	1.7	-0.3	1.7	2.9	2.1	2.046 6
4	4.0	0.8	2.5	1.6	1.5	-1.1	1.5	2.9	2.4	2.343 5
5	3.2	0.5	2.5	1.6	1.5	-1.2	1.3	3.0	2.1	2.058 5
6	4.9	0.3	2.3	1.5	1.4	-0.7	1.4	3.1	2.7	2.640 4
7	5.0	0.2	2.2	1.4	1.2	-0.1	1.3	2.8	2.7	2.629 5
8	4.7	0	2.2	1.4	1.2	0	1.2	2.6	2.6	2.488 9
9	6.1	-0.2	2.3	1.4	1.1	-0.2	1.9	2.6	3.1	3.002 2
10	6.5	-0.2	2.4	1.5	1.0	-0.6	2.5	2.6	3.2	3.176 6
11	5.9	-0.2	2.0	1.3	1.0	-0.5	2.8	2.6	3.0	2.991 7
12	4.1	-0.4	2.1	1.4	0.9	-0.1	2.9	2.8	2.5	2.505 0

2. 利用每个月公布的环比和同比分类增长率 Δv_i 及其对 CPI 总水平 Δv 的影响,根据模型(4)计算权重,检查与原来的 w_i 是否相符.

在模型(4)中只考虑第 i 大类价格的变化 Δv_i(即其他大类价格不变)对总水平的影响 Δv,可以得到

$$w_i=\frac{\Delta v}{\Delta v_i},\quad i=1,2,\cdots,8 \tag{5}$$

这里,w_i 相当于多元函数 v 对自变量 v_i 的偏导数.

据国家统计局公布,2013 年 1 月份食品价格环比上涨 2.8%,影响居民消费价格总水平环比上涨约 0.92%,食品价格同比上涨 2.9%,影响居民消费价格总水平同比上涨约 0.95%.表 7 是 2013 年 1 月至 12 月公布的食品价格同比增长率 Δv_1 和影响居民消费价格总水平同比增长率 Δv,以及利用(5)式计算食品的权重 w_1,比表 4 给出的 w_1 = 0.317 9 稍大.

表 7 2013 年 1 月至 12 月的食品同比增长率 Δv_1 和影响价格总水平同比增长率 Δv

	月份											
	1	2	3	4	5	6	7	8	9	10	11	12
Δv_1	2.9	6.0	2.7	4.0	3.2	4.9	5.0	4.7	6.1	6.5	5.9	4.1
Δv	0.95	1.98	0.88	1.33	1.05	1.59	1.61	1.54	1.98	2.11	1.92	1.33
w_1	0.328	0.330	0.326	0.333	0.328	0.325	0.322	0.328	0.325	0.325	0.325	0.324

这个方法的优点是,如果数据完整,可以对各个权重 w_i 分别计算、校核.当然,由于公布数据的有效数字所限,舍入误差对结果有不小影响.

3. 利用公布的较长时期的数据 Δv_i 和 Δv,对模型(4)作拟合,对权重进行重新估计.设已收集到 n 个月的分类增长率 Δv_i 和总水平增长率 Δv(包括环比和同比),分别记作 Δv_{ik} 和 Δv_k($k=1,2,\cdots,n, i=1,2,\cdots,8$),则模型(4)可写作

$$\boldsymbol{Aw}=\boldsymbol{b}, \boldsymbol{A}=\begin{bmatrix} \Delta v_{11} & \Delta v_{21} & \cdots & \Delta v_{81} \\ \Delta v_{12} & \Delta v_{22} & \cdots & \Delta v_{82} \\ \vdots & \vdots & & \vdots \\ \Delta v_{1n} & \Delta v_{2n} & \cdots & \Delta v_{8n} \\ 1 & 1 & \cdots & 1 \end{bmatrix}, \boldsymbol{w}=\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_8 \end{bmatrix}, \boldsymbol{b}=\begin{bmatrix} \Delta v_1 \\ \Delta v_2 \\ \vdots \\ \Delta v_8 \\ 1 \end{bmatrix} \tag{6}$$

一次代数方程组(6)式含 8 个未知数($w_1, w_2, \cdots, w_8$)、$n+1$ 个方程,用MATLAB的左除命令 $\boldsymbol{w}$=A\b 可得到方程组 $\boldsymbol{Aw}=\boldsymbol{b}$ 的最小二乘解.

用这个方法估计权重原则上没有问题,但是实际计算遇到了困难.由于原始数据精度太低(只有 1 或 2 位有效数字),计算得到的与表 4 列出的权重有较大差距,甚至出现负值.表 8 第 2 列是用 2013 年逐月环比和同比的 24 组数据计算的,第 3 列用了 2011 年至 2013 年逐月环比和同比的 72 组数据,第 4 列加了对权重的非负约束,与最后一列权重的公布数据(表 4 给出)比较,总的情况大致符合,但是个别数值相差较大.

表 8 用逐月环比、同比和总水平增长率对权重的最小二乘估计

权重	$n=24$(2013 年)	$n=72$(2011—2013 年)	$n=72$ 加非负约束	公布数据
w_1	0.328 3	0.314 2	0.314 3	0.317 9
w_2	−0.048 6	−0.003 5	0	0.034 9
w_3	0.077 6	0.095 6	0.095 3	0.085 2
w_4	0.032 3	0.083 4	0.081 7	0.056 4
w_5	0.235 3	0.137 8	0.133 7	0.096 4
w_6	0.115 7	0.069 3	0.069 8	0.099 5
w_7	0.102 1	0.147 5	0.148 7	0.137 5
w_8	0.159 4	0.142 6	0.143 9	0.172 2

评注 CPI 是当今社会的热门词汇,各种媒体特别是互联网上有大量的报道和评论,但大多是经济政策方面的论述.从数据分析和数学建模角度看,资料较少且不够完整,编者只能根据查到的有限数据作了以上的解读,一定会有若干不足之处,读者可按照自己的理解和兴趣继续研究.

以上几种校核与估算各大类权重的方法也适用于各中类的权重,只要能得到各中类的环比和同比数据.

按照地区差别解读 CPI

我国不同地区的经济发展和人们生活水平的差异较大,全国 CPI 总水平和各大类环比、同比指数会与你所在地区的情况有一定差距.国家统计局公布的全国 CPI 数据还区分“城市”和“农村”,31 个省、市、自治区的统计局也逐月公布当地的 CPI 数据,其统计、分析方法与全国类似.

复习题

1. 表5、表6显示，2013年12月全国CPI环比增长率比11月上升0.4%，而同比增长率比11月下降0.5%，如何解释？

2. 按照你所在省（市、自治区）统计局公布的当地CPI数据，进行以下分析：

（1）取1至2年逐月的环比、同比数据作图，观察、分析指数增长率和指数本身之间的关系.

（2）取最近10年的年价格指数作图，观察、分析指数增长率和指数本身之间的关系.

（3）利用得到的数据分析表4列出的全国CPI 8大类的权重是否与你当地的情况一样？

2.8 核军备竞赛



案例精讲2-2
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，在给核威慑战略做出一些合理、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[7,24].

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

$$y=y_0+x \tag{1}$$

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.

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.

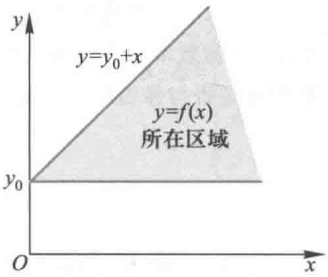


图 1 曲线 $y=f(x)$ 的范围

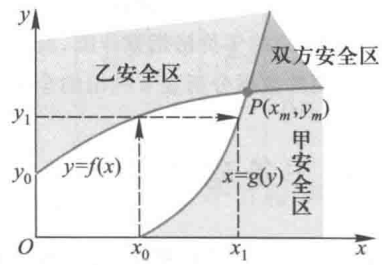


图 2 安全区、安全线和平衡点

根据 $y=f(x)$ 的定义，当 $y\geqslant f(x)$ 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，$x\geqslant g(y)$ 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

$$y=y_0+(1-s)x \tag{2}$$

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

$$y=\frac{y_0}{s} \tag{3}$$

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

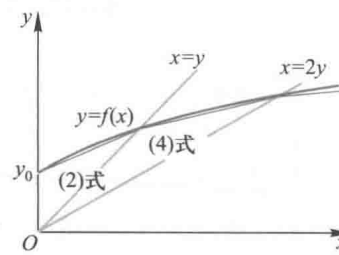


图 3 曲线 $y=f(x)$ 的形成

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

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

$$y=\frac{y_0}{s^2} \tag{5}$$

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

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

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,并且交点唯一.

进一步研究由(6)式表示的乙安全线 $y=f(x)$ 的性质(复习题 1):

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.

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.

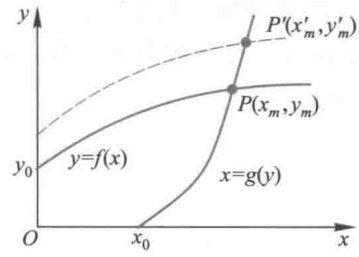


图 4 模型解释 1 的示意图

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,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.

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

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.

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做出解释,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.

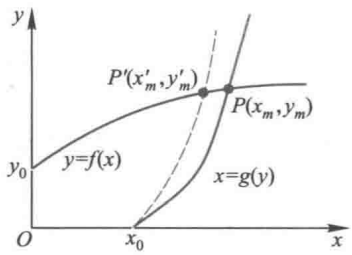


图5 模型解释2的示意图

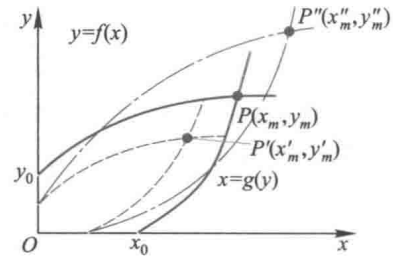


图6 模型解释3的示意图

复习题

1. 在核军备竞赛模型中,证明由(6)式表示的乙安全线 $y=f(x)$ 的性质.
2. 在核军备竞赛模型中,讨论以下因素引起的平衡点的变化[24]:
(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.

2.9 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)[76].

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).

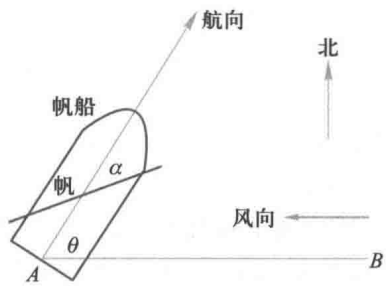


图1 帆船航行示意图

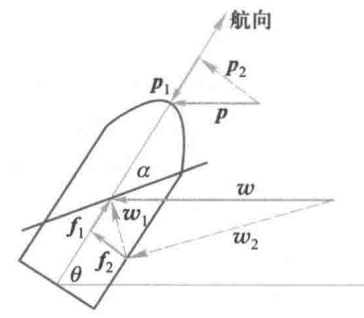


图2 帆船受力分析图

风的推力 $\boldsymbol{w}$ 分解为 $\boldsymbol{w}=\boldsymbol{w}_1+\boldsymbol{w}_2$,其中 $\boldsymbol{w}_1$ 垂直于帆,$\boldsymbol{w}_2$ 平行于帆.$\boldsymbol{w}_1$ 又分解为 $\boldsymbol{w}_1=\boldsymbol{f}_1+\boldsymbol{f}_2$,$\boldsymbol{f}_1$ 即为风在航向的推力.风的阻力 $\boldsymbol{p}$ 分解为 $\boldsymbol{p}=\boldsymbol{p}_1+\boldsymbol{p}_2$,其中 $\boldsymbol{p}_1$ 为风在航向的阻力.因为 $\boldsymbol{f}_1$ 和 $\boldsymbol{p}_1$ 的方向相反,所以船受到的净推力为 $\boldsymbol{f}_1-\boldsymbol{p}_1$.

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.

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. $\boldsymbol{w}$ 的分力 $\boldsymbol{w}_2$ 与帆面平行,可以忽略.

3. 分力 $\boldsymbol{f}_2$ 和 $\boldsymbol{p}_2$ 垂直于船身,可以被船舵抵消,不予考虑.

4. 航速 $\boldsymbol{v}$ 与净推力 $\boldsymbol{f}=\boldsymbol{f}_1-\boldsymbol{p}_1$ 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

$$w=ks_1,\quad p=ks_2 \tag{1}$$

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

$$p_1=p\cos\theta \tag{3}$$

$$v=k_1(f_1-p_1) \tag{4}$$

记船在正东方向的速度分量为 v_1,则

$$v_1=v\cos\theta=k_1(f_1-p_1)\cos\theta \tag{5}$$

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

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

所以当

$$\alpha=\theta/2 \tag{7}$$

时

$$f_1=w(1-\cos\theta)/2 \tag{8}$$

最大.

将(7),(8),(3)代入(5)式得

$$\begin{aligned} v_1&=k_1[w(1-\cos\theta)/2-p\cos\theta]\cos\theta\\ &=(k_1w/2)[1-(1+2p/w)\cos\theta]\cos\theta \end{aligned} \tag{9}$$

注意到(1)式,记

$$k_2=k_1w/2,\quad t=1+2p/w=1+2s_2/s_1 \tag{10}$$

则(9)式为

$$v_1=k_2(1-t\cos\theta)\cos\theta=k_2t\left[\frac{1}{4t^2}-\left(\cos\theta-\frac{1}{2t}\right)^2\right] \tag{11}$$

显然

$$\cos\theta=1/2t \tag{12}$$

时 v_1 最大.由(10)式,且 s_1 远大于 s_2,可知

$$1/4<\cos\theta<1/2,\quad 60^\circ<\theta<75^\circ \tag{13}$$

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.

复习题

若风向不变，行驶中 B 点将不在船的正东方，应该如何确定航向及帆的朝向.

2.10 节水洗衣机

我国淡水资源有限，节约用水人人有责.洗衣在家庭用水中占有相当大的份额，目前洗衣机已非常普及，节约洗衣机用水十分重要.假设在放入衣物和洗涤剂后洗衣机的运行过程为：加水—漂洗—脱水—加水—漂洗—脱水……（称“加水—漂洗—脱水”为运行一轮）.请为洗衣机设计一种程序（包括运行多少轮，每轮加水量等），使在满足一定洗涤效果的条件下，总用水量最少.选用合理的数据进行计算.对照目前常用的洗衣机的运行情况，对你的模型做出评价（引自 1996 年全国大学生数学建模竞赛 B 题）.

问题分析　不论是用手洗衣、洗菜，还是使用洗衣机，甚至工业品、矿产品的清洗，都存在一个在洗净污物的前提下，尽量节省水的问题.因为机器比人能够更精确地控制用水量和洗涤过程，洗衣机又是人人熟悉的家用电器，所以用所谓“节水洗衣机”来研究这样一类应用广泛的实际问题，是具有相当代表性的.

洗衣机运行的基本过程可以简述为：开始时将待洗的衣物和洗涤剂一起放入缸内，加入水后启动洗衣机.在漂洗中通过洗涤剂的物理化学作用，将附着在衣物上的污物溶于水中，再脱去含有污物的污水，构成一个“加水—漂洗—脱水”环节，称为一轮.一轮后残留在衣物上的污物自然有所减少，但若尚未达到洗净的效果，就需要再来一轮.如此循环，直到衣物上的污物减少到相对清洁、可以为人们接受的程度.在现实生活中，洗衣机大多是运行 2 轮或 3 轮，建模过程中应该考虑这个实际情况.至于通常洗衣机运行的最后一步——甩干或烘干，不在本文的讨论范围之内.

需要指出的是，开始洗衣时一次性加入的洗涤剂固然能帮助衣物上的污物溶解于水，但它本身也是不希望留在衣物上的东西，因此不妨将“污物”视为衣物上原有的污物与留在衣物上的洗涤剂的总和，下文中的污物就照此理解.

此外，洗涤剂溶解污物的过程涉及物理化学的微观机制，超出这个问题的研究范围，我们只需从宏观层面上认为，在每一轮运行中污物都已充分溶于水中，形成一定的浓度.通过一轮一轮地加水和脱水，使污物浓度不断降低.

模型假设　根据对洗衣机运行过程的分析，结合实际情况做出以下合理的简化假设：

1. 洗衣机的运行以“加水—漂洗—脱水”为一轮，每轮漂洗后衣物上的污物全部均匀地溶于水中.

2. 与每轮的加水量相比，每轮脱水后衣物仍含少量的水，且每轮的含水量为常数.

3. 每轮脱水前和脱水后污物在水中的浓度保持不变.

4. 洗衣机通常运行 2 轮或 3 轮，最多不超过 4 轮.

5. 相对于初始(洗涤前)衣物含有的污物,洗涤后只能达到相对的清洁,所以规定,最后一轮脱水后衣物的污物含量与衣物的初始污物含量之比(简称污物比),需不超过某个给定的数值.

6. 建立模型的目标是,在满足衣物污物比的条件下,确定洗衣机运行多少轮(在2至4轮范围内)及每轮的加水量,使总的用水量最少.

需要说明的是,由于洗衣机缸体容积的限制及加水需浸没衣物的要求,每轮加水量应该有一个上限和一个下限.为了模型求解的简便,我们先不考虑这个限制条件来建立模型,再在对模型结果的解释中给以讨论.

模型建立 设洗衣机共运行 n 轮($n=2,3,4$).记 u_k 为第 k 轮加水量,x_k 为第 k 轮脱水后污物含量($k=1,2,\cdots,n$),x_0 为初始污物含量,常数 c 为每轮脱水后衣物的含水量,ε 为给定的污物比.

污物浓度是单位容积水中的污物含量,根据假设1、2和假设3关于每轮脱水前后污物浓度保持不变的特性,有

$$\frac{x_0}{u_1}=\frac{x_1}{c},\frac{x_1}{u_2+c}=\frac{x_2}{c},\cdots,\frac{x_{n-1}}{u_n+c}=\frac{x_n}{c} \tag{1}$$

经代数运算可得

$$\frac{x_n}{x_0}=\frac{c^n}{u_1(u_2+c)\cdots(u_n+c)} \tag{2}$$

根据假设5应有 $x_n/x_0\leqslant\varepsilon$,即

$$\frac{c^n}{u_1(u_2+c)\cdots(u_n+c)}\leqslant\varepsilon \tag{3}$$

总的用水量等于每轮加水量之和,记作 z,即

$$z=\sum_{k=1}^{n}u_k \tag{4}$$

按照假设6,我们建模的目标函数是,在(3)式条件下确定洗衣机运行轮数 n($=2,3,4$)和每轮加水量 u_k($k=1,2,\cdots,n$),使(4)式的总用水量 z 最少,其中每轮脱水后衣物含水量 c 和污物比 ε 是已知常数.

模型简化 根据假设2,c 与 u_k 相比较小,(3)式可简化为

$$\frac{c^n}{\prod_{k=1}^{n}u_k}\leqslant\varepsilon \tag{5}$$

要使(4)式的 z 最小,每轮加水量 u_k 应该尽量小,所以(5)式必为等式,可以写成

$$\prod_{k=1}^{n}u_k=a \tag{6}$$

其中 $a=c^n/\varepsilon$.注意到对于每个固定的 n($=2,3,4$),(4)式右端与(6)式左端分别相当于 n 个 u_k 的算术平均值与几何平均值(分别相差因子 $1/n$ 与幂次 $1/n$,不影响下面的结论),根据论据:"n 个数的几何平均值小于或等于算术平均值,当且仅当 n 个数相等时等号成立",可知(4)式和(6)式的解 u_k($k=1,2,\cdots,n$)必全相等,即每轮加水量是一个固定值,记作 u.于是(4)式和(6)式分别简化为

$$z=nu \tag{7}$$

$$u^n = c^n/\varepsilon \tag{8}$$

问题则简化为求 u 和 n(在 2,3,4 中取值)使 z 最小.

回顾从(3)式到(5)式的简化,不妨认为,由(7)式和(8)式得到的 u 是第 1 轮加水量,此后各轮的加水量应减掉一个 c(脱水后衣物的含水量).

模型求解 简化后的(7)式和(8)式能够用初等方法求解.

由于 n 取有限个离散值,可以先固定 n,从(8)式解出 $u=\dfrac{c}{\varepsilon^{1/n}}$,再代入(7)式得 $z=n\dfrac{c}{\varepsilon^{1/n}}$.对于给定的 c,ε, 令 $n=2,3,4$,通过总用水量 z 的比较确定 u 和 n 的解.

注意到 u 和 n 的表达式中 c 以独立因子出现,我们可以将其记为

$$u=\mu_n c,\quad \mu_n=\frac{1}{\varepsilon^{1/n}} \tag{9}$$

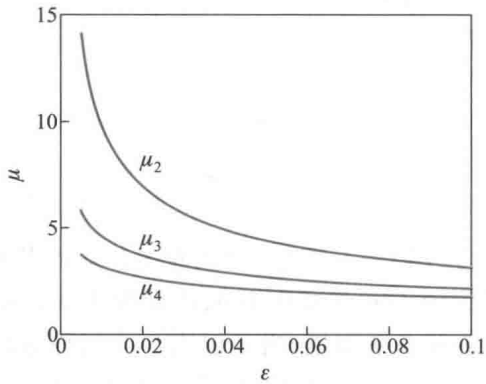
$$z=\lambda_n c,\quad \lambda_n=\frac{n}{\varepsilon^{1/n}} \tag{10}$$

对于固定的 c, 数值 μ_n 和 λ_n 直接反映了每轮加水量和总用水量的大小.

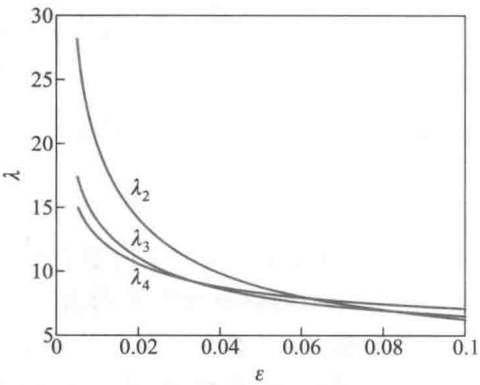
不妨设定若干个污物比 ε(从 0.5%到 10%),对于 $n=2,3,4$,按照(9)式和(10)式计算 μ_n 和 λ_n,结果如表 1 和图 1.

表 1 不同污物比 ε 下的 μ_n 和 λ_n($n=2,3,4$)

ε	0.5%	1%	2%	5%	10%
μ_2	14.2	10.0	7.1	4.5	3.2
μ_3	5.8	4.6	3.7	2.7	2.2
μ_4	3.8	3.2	2.7	2.1	1.8
λ_2	28.3	20.0	14.2	8.9	6.3
λ_3	17.5	13.9	11.1	8.1	6.5
λ_4	15.0	12.6	10.6	8.5	7.1



(a) 不同污物比ε下的μ_n(n=2,3,4)



(b) 不同污物比ε下的λ_n(n=2,3,4)

图 1

从表 1 和图 1 可以看出，当对衣物的清洁程度要求较高时（污物比 ε 不超过 2%），洗衣机运行 4 轮的总用水量最少，而当清洁程度要求较低时（污物比 ε 大于 5%），洗衣机总用水量都差不多，可能运行 3 轮更合适.

模型讨论　在模型假设中曾经提到，由于洗衣机缸体容积的限制，每轮加水量应该有一个上限，记作 u_{max}，又由于加水需浸没衣物便于漂洗的要求，每轮加水量应该有一个下限，记作 u_{min}.那么，上面的模型简化后得到的加水量 u 是否满足上下限的限制呢？

从(9)式可知，加水量 u 正比于脱水后衣物的含水量 c，而 c 的大小与洗涤前衣物本身的重量和质地有关.记洗涤前衣物重量为 w，可以合理地假设脱水后衣物的含水量 $c=aw$，a 是单位重量衣物脱水后的含水量.加水量下限 u_{min} 由浸没衣物以便漂洗的需要决定，类似地可以假设 $u_{min}=bw$，b 是单位重量衣物浸泡所需的水量.a 和 b 主要取决于衣物的质地，但其变化范围不会太大.a 和 b 的数值可以通过实验大致确定.

从以上分析及(9)式可知，要使加水量 u 满足下限 u_{min} 的限制，只需 $\mu_n a \geqslant b$.按照表 1 中洗衣机运行 3 轮的数据 μ_3，对衣物的清洁程度要求较高时，这个条件大体上能够满足.

按照上面的分析不难看出，只需对洗涤衣物的重量加以控制，加水量上限 u_{max} 的限制就可以满足.

实际上，有些洗衣机开始运行时先通过空转对衣物重量进行自动量测，这相当于对含水量 c 和下限 u_{min} 做出估计，再按照衣物的重量范围采用不同的程序运行，如不同的加水量，甚至不同的轮数.

至于题目中提出的选用合理的数据进行计算，对照目前洗衣机运行情况对模型做出评价的工作就留给读者了.

复习题

1. 对于模型(3)，(4)式考虑 $n=2$ 的情况，不作简化直接用初等方法求解，与(9)，(10)式的结果比较.

2. 对于模型(3)，(4)式考虑 $n=3$ 的情况，设第 2，3 轮的加水量相同，不作简化直接求解（可用微积分），与(9)，(10)式的结果比较.

3. 如果将模型假设 1 的“每轮漂洗后衣物上的污物全部均匀地溶于水中”，改为“每轮漂洗后衣物上只有一定比例的污物均匀地溶于水中”，重新建立模型，讨论如何求解.

第 2 章训练题

1. 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？比如佳洁士牙膏 120 g 装的每支 10.80 元，200 g 装的每支 15.80 元，二者单位质量的价格比是 1.14∶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.

2. 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[24].

3. 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[17].

4. 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[24].

5. 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[24].

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

6. 你能针对估计出租车数量问题提出自己的模型吗？如果有，通过模拟与2.5节5个模型作比较.

更多案例……

2-1 光盘的数据容量

2-2 污水均流池的设计

2-3 信号灯控制的十字路口的通行能力与车速线性模型

2-4 天气预报的评价

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、定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



案例精讲 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[63,84].

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,可使平均每天总费用最小.

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+…+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+…+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,使平均每天总费用最小.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,使每天费用最小.

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

$$Q=rT \tag{1}$$

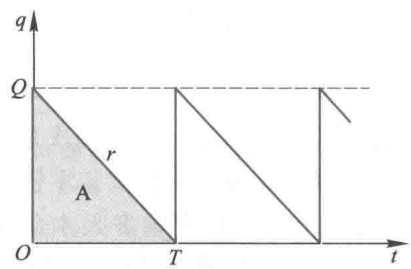


图1 不允许缺货模型的贮存量 $q(t)$

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

$$\bar{C}=c_1+c_2QT/2=c_1+c_2rT^2/2 \tag{2}$$

于是每天的平均费用是

$$C(T)=\bar{C}/T=c_1/T+c_2rT/2 \tag{3}$$

(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

$$T=\sqrt{\frac{2c_1}{c_2r}} \tag{4}$$

代入(1)式可得

$$Q=\sqrt{\frac{2c_1r}{c_2}} \tag{5}$$

由(3)式算出最小的每天费用为

$$C=\sqrt{2c_1c_2r} \tag{6}$$

(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.

评注　EOQ 公式是大约一百年前得到的,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于生产准备费,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.

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

$$S(T,c_1)=\frac{\Delta T/T}{\Delta c_1/c_1}\approx\frac{\mathrm{d}T}{\mathrm{d}c_1}\frac{c_1}{T} \tag{7}$$

由(4)式容易得到 $S(T,c_1)=1/2$.做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(复习题 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(复习题 2)?

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 是每周期初的贮存量,当 $t=T_1$ 时 $q(t)=0$,于是有

$$Q=rT_1 \tag{8}$$

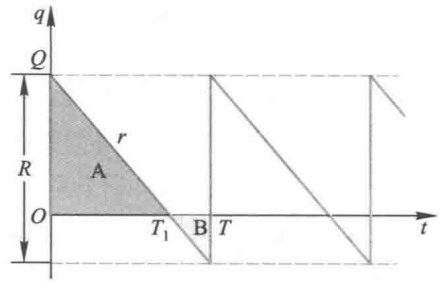


图 2 允许缺货模型的贮存量 $q(t)$

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,使下周期初的贮存量恢复到 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

$$\bar{C}=c_1+c_2QT_1/2+c_3r(T-T_1)^2/2 \tag{9}$$

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

$$C(T,Q)=\frac{c_1}{T}+\frac{c_2Q^2}{2rT}+\frac{c_3(rT-Q)^2}{2rT} \tag{10}$$

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')

$$T'=\sqrt{\frac{2c_1}{c_2r}\frac{c_2+c_3}{c_3}},\quad Q'=\sqrt{\frac{2c_1r}{c_2}\frac{c_3}{c_2+c_3}} \tag{11}$$

注意到每个周期的供货量 $R=rT'$,有

$$R=\sqrt{\frac{2c_1r}{c_2}\frac{c_2+c_3}{c_3}} \tag{12}$$

记

$$\lambda=\sqrt{\frac{c_2+c_3}{c_3}} \tag{13}$$

与不允许缺货模型的结果(4),(5)式比较不难得到

$$T'=\lambda T,\quad Q'=Q/\lambda,\quad R=\lambda Q \tag{14}$$

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 $c_3\to\infty$ 时 $\lambda\to1$,于是 $T'\to T$,$Q'\to Q$,$R\to Q$.这个结果合理吗(考虑 $c_3\to\infty$ 的意义)? 由此不允许缺货模型可视为允许缺货模型的特例.

复习题

1. 在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，以平均每天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.

3.2 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[7].

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 $\dfrac{\mathrm{d}B}{\mathrm{d}t}$ 比 $B(t)$ 更为直接和方便.$\dfrac{\mathrm{d}B}{\mathrm{d}t}$ 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 $0\leqslant t\leqslant t_1$，火势越来越大，即 $\dfrac{\mathrm{d}B}{\mathrm{d}t}$ 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 $\dfrac{\mathrm{d}B}{\mathrm{d}t}$ 应减小，并且当 $t=t_2$ 时 $\dfrac{\mathrm{d}B}{\mathrm{d}t}=0$.

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 $\dfrac{\mathrm{d}B}{\mathrm{d}t}$ 的形式做出假设.

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（$0\leqslant t\leqslant t_1$）内，火势蔓延程度 $\dfrac{\mathrm{d}B}{\mathrm{d}t}$ 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 $0\leqslant t\leqslant t_1$ 线性地增加,在 $t_1<t<t_2$ 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 $B(t_2)=\frac{1}{2}bt_2$,而 t_2 满足

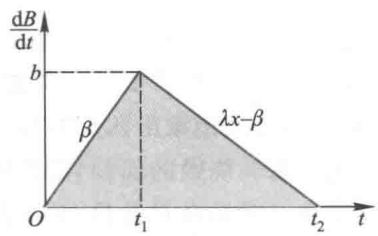


图1 $\frac{dB}{dt}\sim t$ 关系

$$t_2-t_1=\frac{b}{\lambda x-\beta}=\frac{\beta t_1}{\lambda x-\beta} \tag{1}$$

于是

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

根据假设条件1,4,森林损失费为 $c_1B(t_2)$,救援费为 $c_2x(t_2-t_1)+c_3x$.将(1),(2)代入,得到救火总费用为

$$C(x)=\frac{c_1\beta t_1^2}{2}+\frac{c_1\beta^2 t_1^2}{2(\lambda x-\beta)}+\frac{c_2\beta t_1 x}{\lambda x-\beta}+c_3x \tag{3}$$

$C(x)$ 即为这个优化模型的目标函数.

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

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

$$x=\frac{\beta}{\lambda}+\beta\sqrt{\frac{c_1\lambda t_1^2+2c_2t_1}{2c_3\lambda^2}} \tag{4}$$

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.

救火时的火势 b 有关,b 越大 λ 越小.这时要对函数 $\lambda(b)$ 做出合理的假设,再得到进一步的结果.(复习题)

复习题

在森林救火模型中,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.

3.3 倾倒的啤酒杯

一个春意盎然的午后,你和家人来到郊野公园的草地上,铺下一领凉席,打开带来的食品袋,斟满一杯啤酒随手放到席子上,哎哟,啤酒杯倾倒了,金黄色的啤酒洒了一摊.又一个夏日炎炎的傍晚,你和朋友来到碧沙起伏的海边,朋友刚要把从吧台买来的啤酒撒在沙滩上,你突然想起那次在草地上的教训,赶紧接过来一口气把啤酒喝得只剩下不多,才放心地放下,哎呀,啤酒杯又倒了.满杯也倒,空杯也倒,怎么回事?

上面的场景说明,在不大平坦的地方盛着啤酒的杯子不容易放稳,这是因为重心太高了.直观地看,满杯和空杯的重心都大约在杯子中央稍下一点的位置.而在你饮用一满杯啤酒的过程中,随着杯中啤酒液面的下降,重心应该先降低,后又升高(从常识角度想想为什么会升高),那么重心必然有一个最低点.如果你在啤酒杯的重心大致处于最低点时停止饮用,放下来,就是最不容易倾倒的情况.

饮酒时啤酒杯重心的变化过程与倒酒时的正好相反,重心的最低点是一样的,而讨论倒酒过程比较方便.本节要建立一个数学模型,描述啤酒杯的重心随啤酒液面上升而变化的规律,找出重心最低点的位置,并讨论最低点由哪些因素决定[70].

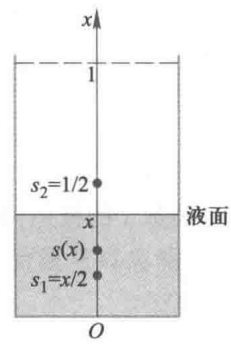


图 1 啤酒杯及其重心坐标

模型分析与假设 将最简单的啤酒杯的形状看作一个圆柱体,杯子的高度可以假定为 1(个单位).由于在啤酒液面上升过程中,重心始终在杯子的中轴线上变化,故以中轴线的底端为坐标原点,沿中轴线建立坐标轴 x,倒酒时液面高度从 $x=0$ 增加到 $x=1$,而重心是 x 的函数,记作 $s(x)$,见图 1.

装着酒的啤酒杯的重心位置由啤酒和空杯的质量和重心决定.假设啤酒和杯子材料都是均匀的,一满杯啤酒(不计杯子)的质量记作 w_1,空杯侧壁的质量记作 w_2,空杯底面的质量记作 w_3.因为啤酒杯是高度为 1 的圆柱,所以液面高度为 x 的啤酒质量是 $w_1 x$,其重心记作 s_1,是高度的一半,即 $s_1 = \dfrac{x}{2}$.倒酒过程中空杯的质量不变,其重心与 x 无关,只由 w_2 和 w_3 的大小决定.为简化起见,

下面先忽略 w_3，建立啤酒杯重心 $s(x)$ 的模型，求解重心最低点的位置，并讨论其性质，然后再将 w_3 加入模型.

模型一的建立与求解 在忽略空杯底面质量而只考虑侧壁的情况下，空杯的重心位置记作 s_2，显然 $s_2=\frac{1}{2}$，如图 1 所示.

液面高度为 x 时啤酒杯的质量为 w_1x+w_2，其重心 $s=s(x)$ 由啤酒质量 w_1x、重心 s_1 与空杯质量 w_2、重心 s_2 的力矩平衡关系得到，根据物理学关于重心的计算公式，有

$$(w_1x+w_2)s=w_1xs_1+w_2s_2 \tag{1}$$

将 $s_1=\frac{x}{2}$、$s_2=\frac{1}{2}$ 代入，并记 $a=\frac{w_2}{w_1}$，由(1)式可解出 s 得

$$s(x)=\frac{x^2+a}{2(x+a)} \tag{2}$$

(2)式表明啤酒杯的重心随液面高度的变化规律 $s(x)$ 只与空杯和满杯啤酒(不计杯子)的质量比 a 有关.当 a 取不同数值 0.1，0.3，0.5，1 时，s 随 x 的变化曲线如图 2.由图可见对于每一个 a，s 都有一最小值.如若 $a=0.3$，则液面 x 在 0.35 左右 s 最小，即重心最低.

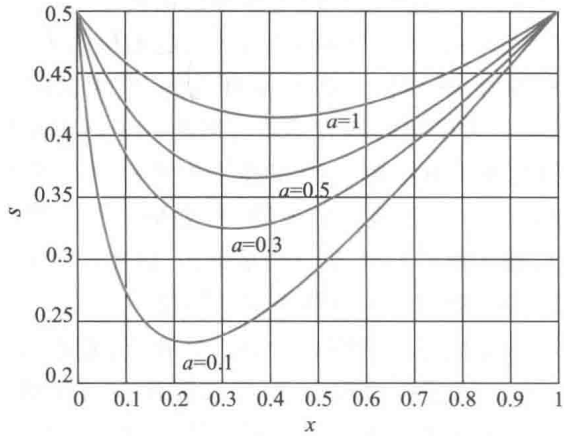


图 2 重心 s 随 x 的变化曲线(对不同的 a)

函数 s 的极值问题可以用微分法求解.对(2)式求导数得

$$\frac{ds}{dx}=\frac{4x(x+a)-2(x^2+a)}{4(x+a)^2} \tag{3}$$

令 $\frac{ds}{dx}=0$ 得到 $x^2+2ax-a=0$，这个二次代数方程的正根是 $s(x)$ 的最小点，记作 x^*，有

$$x^*=\sqrt{a^2+a}-a \tag{4}$$

当液面高度为 x^* 时，啤酒杯重心处于最低位置.由(4)式表示的曲线如图 3，给出了 x^* 与质量比 $a=w_2/w_1$ 之间的关系.将(4)代入(2)即可得到重心的最低点.

模型一结果分析 由(4)式可知，啤酒杯重心最低时的液面高度只取决于空杯和满杯啤酒的质量比 a.如果考察半升啤酒的杯子，啤酒的密度与水差不多，满杯啤酒的

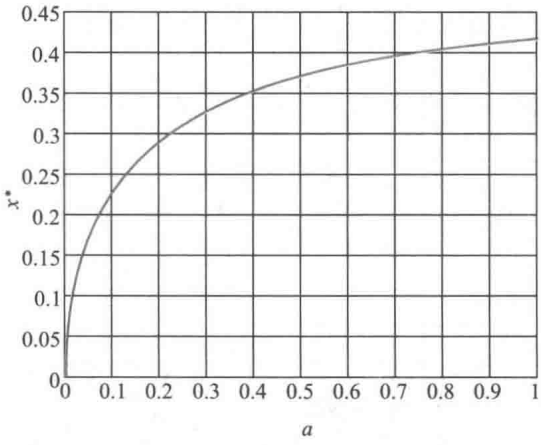


图 3 重心最低时的液面位置随 a 的变化曲线

质量是 500 g,空杯的质量则依材料的不同(纸制品、化学制品、玻璃制品等)可在相当大的范围(如从 50 g 到 500 g)内变化,于是质量比 a 取 0.1 到 1 的数值.若将 $a=0.3$(空杯质量150 g)代入(4)式可算出 $x^*=0.324\ 5$,表明一杯啤酒喝到差不多只剩$\frac{1}{3}$时重心最低,最不容易倾倒.

由图 2、图 3 可以看出,随着质量比 a 的变大,啤酒杯重心最低时的液面高度 x^* 在升高.对于固定的 w_1 而言,空杯质量越大,重心最低时的液面越高,这显然是符合常识的.比如你用质量 500 g、半升装的厚玻璃杯饮酒时 $a=1$,按照(4)式重心最低的液面高度应为$\sqrt{2}-1=0.414$.

从数学角度还可以研究,质量比 a 无限变大时重心最低的液面高度变化趋势,即(4)式中 $a\to\infty$ 的极限值,可如下计算:

$$\lim_{a\to\infty}(\sqrt{a^2+a}-a)=\lim_{a\to\infty}\frac{a}{\sqrt{a^2+a}+a}=\lim_{a\to\infty}\frac{1}{\sqrt{1+\frac{1}{a}}+1}=\frac{1}{2}$$

说明用(与啤酒质量相比)非常重的杯子饮酒,喝到一半($x=\frac{1}{2}$)重心最低.还可注意到,此时啤酒杯的重心也是 $s=\frac{1}{2}$,正好在液面的位置.显然是由于啤酒质量可以忽略,啤酒杯重心与空杯重心相合.

意料之外与情理之中 将(4)式的 x^* 代入重心 s 随 x 的变化规律(2)式,得到

$$s(x^*)=\frac{(\sqrt{a^2+a}-a)^2+a}{2[(\sqrt{a^2+a}-a)+a]}=\frac{2(a^2+a-a\sqrt{a^2+a})}{2\sqrt{a^2+a}}=\sqrt{a^2+a}-a \tag{5}$$

与(4)式的 x^* 完全相同,表明当啤酒杯重心与液面高度相重合时重心最低.这是否在你的意料之外呢?

这个结果其实也在情理之中.考察向杯中倒酒的过程,液面高度 $x=0$ 时重心 $s=s_2=$

$\frac{1}{2}$，随着 x 的升高，由于啤酒重心 $s_1=\frac{x}{2}$ 的“向下作用”，重心 s 下降，直至 x 与 s 重合.此后，x 继续升高，$s_1=\frac{x}{2}$ 起“向上作用”，重心 s 上升，直到 $x=1$ 时 $s=\frac{1}{2}$.

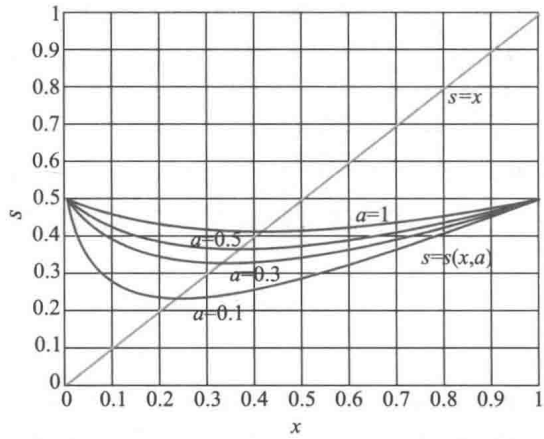


图4 直线 $s=x$ 通过曲线 $s(x,a)$ 的最低点

上述过程可以通过数学模型给以清晰的表示.对(2)、(3)两式作代数运算可得

$$\frac{\mathrm{d}s}{\mathrm{d}x}=\frac{x-s}{x+a} \tag{6}$$

表示 s 对 x 的导数与差 $x-s$ 同号，且当 $x=s$ 时导数为0.这说明，随着 x 变大，$x<s$ 时，s 对 x 的导数为负，s 下降；$x>s$ 时，s 对 x 的导数为正，s 上升；$x=s$ 时，s 达到最小值.

将图2的坐标范围稍加改动，画成图4，可以看出，对于任意的 a，直线 $s=x$ 通过每条曲线 $s(x,a)$ 的最低点，这是上述结果的图示.

模型二的建立与求解 考虑空杯底面的质量 w_3，重心记作 s_3，由于与杯子高度相比底面的厚度很小，可以忽略，所以可认为 $s_3=0$.仍然利用重心的计算公式，与模型一(1)式类似地有

$$(w_1x+w_2+w_3)s=w_1xs_1+w_2s_2+w_3s_3 \tag{7}$$

将 $s_1=\frac{x}{2}$、$s_2=\frac{1}{2}$、$s_3=0$ 代入，仍记 $a=w_2/w_1$，再记 $b=w_3/w_1$，由(7)式可解出 s 得

$$s(x)=\frac{x^2+a}{2(x+a+b)} \tag{8}$$

$b=0$ 时，与模型一(2)式相同.

用与模型一相同的方法求解，得到(2)式的最小点与最小值为

$$x^*=s(x^*)=\sqrt{(a+b)^2+a}-(a+b) \tag{9}$$

再次出现与模型一同样的结果：当啤酒杯重心与液面高度相重合时重心最低.

粗略估计一下考虑空杯底面的质量对最低重心位置的影响.设杯子侧壁和底面的厚度和材质相同，记侧壁高度和底面直径分别为 h，d，且不妨假定 $h=2d$，则容易得到 $\frac{w_3}{w_2}=\frac{d}{4h}=\frac{1}{8}$.若设 $a=\frac{w_2}{w_1}=0.3$(空杯质量150 g)，则 $b=\frac{w_3}{w_1}=\frac{1}{8}\times0.3=0.0375$.将 a，b 代入

评注 对于一个虽然实际意义不大、却饶有生活情趣的现象，用数学建模的方法加以表述，对得到的结果给予解释.在对杯子做适当的简化假设后，应用物理学的基本知识，建立了函数极值问题的优化模型，通过求导数、极限、作图等数学手段，给出了精确的数量关系.

我们得到的“啤酒杯重心与液面高度相重合时重心最低”这个既在意料之外又在情理之中的结果，不通过数学建模

(9)式得 $x^*=0.3059$.与模型一在 $a=0.3$ 时的结果 $x^*=0.3245$ 相比,相差不到 6%.

大概是难以想象的.从数学角度看,这相当于函数 $s=s(x)$ 的最小点 x^*(满足 $s'(x^*)=0$)是函数的不动点,即 $x^*=s(x^*)$.

模型推广 由于设定啤酒杯是最简单的圆柱形,所以建立的模型及求解的结果也非常简明.如果啤酒杯呈圆台或球台等形状,建模的方法是一样的,但是推导会比较复杂.一个有趣的现象是,只要啤酒杯是旋转体(侧壁由任意曲线绕中轴线旋转而成),那么上面那个意料之外、情理之中的结果就会成立.

复习题

设啤酒杯的侧壁是由任意曲线绕中轴线旋转形成的,建立啤酒杯重心随啤酒液面上升而变化的数学模型,证明啤酒杯重心与液面高度相重合时重心最低.

3.4 铅球掷远

你投掷过铅球或者看过投掷铅球吧.运动员手指握球、置于颌下、上体微屈、背对投向,左腿摆动、右腿蹬伸、滑步向前、送出右髋,伸直右臂,将球迅速、有力地推出,极大的爆发力使铅球画出一条优美的弧线,落向远方.

铅球掷远起源于 14 世纪欧洲炮兵闲暇期间推掷炮弹的游戏和比赛,当时炮弹是圆形铁球,重量为 16 磅(合 7.26 kg),这一重量一直沿袭成为现在男子铅球正式比赛的标准.见证这一历史渊源的还有,英语中“铅球”与“炮弹”是同一单词“shot”.男子铅球早在 1896 年第一届奥运会上就被列为比赛项目,女子铅球也在 1948 年进入奥运会.

虽然运动会上我们看到的铅球运动员都是人高马大、胳膊粗壮,但并不是只要有力气就能把铅球掷得远.从常识判断,运动员能够对铅球施加作用的因素只有初始速度和出手角度,显然初始速度越大掷得越远,出手角度则存在一个使投掷距离最远的最佳值.另外,教练员在挑选运动员时还会注意他(她)的身高,因为出手高度越大也会让铅球掷得越远.本节要通过数学建模,定量地分析投掷距离与铅球的初始速度、出手角度、出手高度等因素的关系,找出最佳出手角度,并研究初始速度、出手角度、出手高度的微小改变对投掷距离的影响[70].

问题分析 男子铅球的直径只有 11 cm 至 13 cm,在短暂的飞行中所受的阻力可以忽略.于是可将铅球视为一个质点,以一定的初始速度和出手角度投出后,在重力作用下作斜抛运动.中学物理知识告诉我们,飞行轨迹是一条抛物线,如图 1,投掷距离除了取决于铅球的初始速度 v、出手角度 θ 外,还与出手高度 h 有关.下面先暂不考虑出手高度,建立一个简单模型,分析初始速度 v 和出手角度 θ 对投掷距离的影响,再将出手高度 h 加入,改进原来的模型[13].

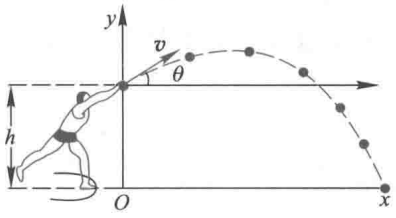


图 1 铅球飞行轨迹示意图

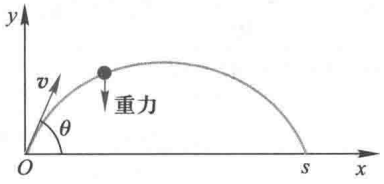


图 2 不考虑出手高度的简单模型

模型一 不考虑铅球出手高度，建立坐标系如图2，初始速度 v 方向与 x 轴的夹角为 θ.v 在 x 轴和 y 轴上的投影分别为 $v_x=v\cos\theta$，$v_y=v\sin\theta$.忽略空气阻力，铅球飞行过程中在 x 轴方向不受力，在 y 轴方向只受重力作用(与 y 轴反向)，记重力加速度为 g.设铅球在时刻 $t=0$ 从坐标原点投出，按照运动的基本规律，铅球在任意时刻 t 的位置坐标 (x, y) 满足

$$x=vt\cos\theta,\ y=vt\sin\theta-\frac{gt^2}{2} \tag{1}$$

记投掷距离为 s，注意到铅球落地即 $y=0$，此时的 x 坐标等于 s，(1)式化为

$$s=vt\cos\theta,\ 0=vt\sin\theta-\frac{gt^2}{2} \tag{2}$$

消去 t 得到

$$s=\frac{v^2\sin 2\theta}{g} \tag{3}$$

(3)式给出投掷距离 s 与铅球初始速度 v 和出手角度 θ 的关系.对模型一可作以下初步分析：

1. 当出手角度 $\theta=\frac{\pi}{4}$ 时，$\sin 2\theta=1$，投掷距离 $s=\frac{v^2}{g}$ 达到最大，最佳出手角度 $\frac{\pi}{4}$ 与初始速度无关，这也就是大家熟知的“物体以45度角抛出的距离最远”.

2. 对任何出手角度 θ，投掷距离 s 与铅球初始速度 v 的平方成正比，表明初始速度的提高能使投掷距离大幅度地增加.

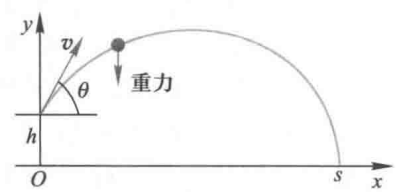


图3 考虑出手高度的改进模型

模型二 设铅球出手高度为 h，建立坐标系如图3，铅球在时刻 $t=0$ 从坐标 $(0, h)$ 投出，(1)式改为

$$x=v_xt,\ y=h+v_yt-\frac{gt^2}{2} \tag{4}$$

将 $y=0$，$x=s$ 代入(4)式得到

$$s=vt\cos\theta,\ 0=h+vt\sin\theta-\frac{gt^2}{2} \tag{5}$$

消去 t 解出

$$s=\frac{v^2\sin\theta\cos\theta}{g}+\frac{v\cos\theta}{g}\sqrt{v^2\sin^2\theta+2hg} \tag{6}$$

显然，当 $h=0$ 时(6)式等同于模型一的(3)式.

由(6)式计算 $\frac{ds}{d\theta}=0$ 可以得到最佳出手角度，但是这个计算太繁了！我们另辟一条捷径：先由(5)式消去 t 得

$$h+s\tan\theta-\frac{gs^2}{2v^2\cos^2\theta}=0 \tag{7}$$

对(7)式求导数并令 $\frac{ds}{d\theta}=0$ 得

$$\frac{s}{\cos^2\theta}-\frac{gs^2\sin\theta}{v^2\cos^3\theta}=0 \tag{8}$$

由此解出

$$s=\frac{v^2}{g\tan\theta} \tag{9}$$

将(9)式代入(7)式有

$$h+\frac{v^2}{g}-\frac{v^2}{2g\sin^2\theta}=0 \tag{10}$$

于是

$$\sin\theta=\frac{v}{\sqrt{2(v^2+gh)}} \tag{11}$$

(11)式的 θ 为最佳出手角度.将(11)式代入(9)式得

$$s=\frac{v}{g}\sqrt{v^2+2gh} \tag{12}$$

容易验证,当 $h=0$ 时(11),(12)式给出与模型一相同的结果.

(6)式表示的是投掷距离 s 与铅球初始速度 v、出手角度 θ 和出手高度 h 之间的一般关系,(11),(12)式则给出了由初始速度 v、出手高度 h 决定的最佳出手角度和最远投掷距离.对模型二可作以下初步分析:

1. 最佳出手角度 $\theta<\frac{\pi}{4}$,且 θ 与 v 和 h 有关,v 越大 θ 越大,h 越大 θ 越小.

2. 以最佳角度 θ 投出的最远距离 s 与 v 和 h 有关,v 越大 s 越大,h 越大 s 越大.由于大致上 s 与 v 为 2 次方关系,与 h 为 $\frac{1}{2}$ 次方关系,所以提高 v 远比提高 h 对 s 的增加有效.

表 1 是出手高度为 1.8 m~2.2 m、初始速度为 8 m/s~15 m/s 下的最佳出手角度及最远投掷距离.出手高度可以按照身高加 20 cm 左右估计,对普通人初始速度大概在 8 m/s~10 m/s,一般运动员可达 10 m/s~13 m/s,最佳出手角度应在 40°左右.目前男子铅球(重 7.26 kg)世界纪录为 23.12 m(巴恩斯,美国,1990 年),女子铅球(重 4 kg)世界纪录为 22.63 m(里索斯卡娅,苏联,1987 年).按照这个模型,他(她)投掷的初始速度应接近15 m/s.

表 1 按照模型二计算的最佳出手角度及最远投掷距离

		初始速度/(m·s^{-1})							
		8	9	10	11	12	13	14	15
最佳角度/°	$h=1.8$ m	38.76	39.85	40.69	41.35	41.87	42.29	42.63	42.92
	$h=2.0$ m	38.22	39.38	40.28	40.99	41.56	42.02	42.39	42.70
	$h=2.2$ m	37.70	38.93	39.89	40.65	41.26	41.75	42.16	42.49
最远距离/m	$h=1.8$ m	8.13	9.90	11.87	14.03	16.40	18.96	21.73	24.69
	$h=2.0$ m	8.29	10.07	12.04	14.21	16.57	19.14	21.91	24.88
	$h=2.2$ m	8.45	10.23	12.21	14.38	16.75	19.32	22.09	25.06

根据表1中模型二的计算结果，设 $h=2.0$ m，$v=10$ m/s 时 $s=12.04$ m，$v=12$ m/s 时 $s=16.57$ m.按照模型一中 $\theta=\frac{\pi}{4}$ 的 $s=\frac{v^2}{g}$ 计算，得 $v=10$ m/s 时 $s=10.20$ m，$v=12$ m/s 时 $s=14.69$ m.对比发现，模型二的最远投掷距离 s 比模型一大约增加了2 m，差不多正是一个出手高度 h.这个结果可以由(12)式近似地推出(复习题1).

敏感性分析 研究铅球初始速度 v、出手角度 θ 和出手高度 h 的微小改变对投掷距离 s 的影响.让我们先看数值计算结果，再利用微分法作理论分析.

模型一的数值计算 若 v 提高5%，用 $1.05v$ 代替(3)式中的 v，则 s 将变为1.102 5 s，即 s 增加约10%，表明初始速度微小的提高能使投掷距离成倍地增加.

如果出手角度 θ 变化5%，由最佳角度45°变为42.75°(或47.25°)，则(3)式中 $\sin 2\theta=\sin 85.5°=0.996\ 9$，投掷距离 s 仅减少约0.3%，说明出手角度在最佳角度处的微小改变对投掷距离的影响很小.

模型一的理论分析 可以用 $\frac{\Delta s}{s}$，$\frac{\Delta v}{v}$ 分别表示 s 和 v 的相对微小改变量，并用微分 $\mathrm{d}s$，$\mathrm{d}v$ 代替增量 Δs，Δv，由(3)式计算 s 的微分，得到

$$\frac{\mathrm{d}s}{s}=2\frac{\mathrm{d}v}{v} \tag{13}$$

即 s 的相对(微小)改变是 v 的相对(微小)改变的两倍.

类似的计算可得

$$\frac{\mathrm{d}s}{s}=\frac{2\theta}{\tan 2\theta}\frac{\mathrm{d}\theta}{\theta} \tag{14}$$

这个结果不能分析 θ 正好等于 $\pi/4$ 时的情况，不妨取略小于 $\pi/4$ 的角度如42.75°，可得 $\frac{\mathrm{d}s}{s}=0.12\frac{\mathrm{d}\theta}{\theta}$，所以 θ 的微小改变对 v 的影响很小.

模型二的数值计算 从表1可以看出，出手高度增加0.2 m(约10%)，最佳出手角度减小不到0.5°，最远投掷距离增加近0.2 m(约1.5%)，对于激烈的比赛来说这个数值还是相当可观的，所以运动员选材时应该对身高给予关注.

初始速度提高1 m/s(约10%)，最佳出手角度增加稍多于0.5°，最远投掷距离可提高2 m以上(15%以上)，这是一个非常重要的结果，说明在注重运动员身高与出手角度的基础上，训练中应调动全身力量，形成巨大的投掷爆发力，使铅球达到尽可能大的初始速度.

模型二的理论分析 根据(12)式计算 s 的微分 $\mathrm{d}s$ 与 v 的微分 $\mathrm{d}v$(h 视为常数)，及 $\mathrm{d}s$ 与 h 的微分 $\mathrm{d}h$(v 视为常数)，

$$\mathrm{d}s=\frac{2(v^2+gh)}{g\sqrt{v^2+2gh}}\mathrm{d}v,\quad \mathrm{d}s=\frac{v}{\sqrt{v^2+2gh}}\mathrm{d}h \tag{15}$$

再由(15)式和(12)式计算出

$$\frac{\mathrm{d}s}{s}=\left(1+\frac{v^2}{v^2+2gh}\right)\frac{\mathrm{d}v}{v},\quad \frac{\mathrm{d}s}{s}=\frac{gh}{v^2+2gh}\frac{\mathrm{d}h}{h} \tag{16}$$

评注 投掷铅球大概是体育运动

可以看出，$\frac{ds}{s}>\frac{dv}{v}$，而$\frac{ds}{s}<\frac{dh}{h}$，且因通常 v^2 比 $2gh$ 大许多，所以 v 的微小改变对投掷距离 s 的影响比 h 的微小改变对 s 的影响显著得多.

若以 $v=12$ m/s，$h=2.0$ m 代入(16)式，可得$\frac{ds}{s}=1.78\frac{dv}{v}$，$\frac{ds}{s}=0.11\frac{dh}{h}$，说明在 $v=12$ m/s，$h=2.0$ m 附近，v 的1%改变可引起 s 多达1.78%的变化，而 h 的1%改变只引起 s 有0.11%的变化.

理论上严格分析出手角度 θ 的微小改变对 s 的影响需计算(6)式的微分，因过于烦琐不作推导，不过根据(6)式的近似(3)式及模型一的(14)式可以估计到，θ 的微小改变对 s 的影响也会很小.

本节利用微分对模型一、二所做的敏感性分析方法可以一般地表示如下：

已知函数 $y=f(x)$，计算微分 $dy=f'(x)dx$，在 $x=x_0$ 有

$$\frac{dy}{y}=g(x_0)\frac{dx}{x},\quad g(x_0)=\frac{x_0f'(x_0)}{f(x_0)} \tag{17}$$

式中 $g(x_0)$ 是 $x=x_0$ 处函数的相对(微小)改变$\frac{dy}{y}$与自变量的相对(微小)改变$\frac{dx}{x}$之比，$g(x_0)$ 越大，x 的变化对 y 的影响越大，即 y 对 x 越敏感.

对多元函数用(17)式计算时，应将其他自变量作为常数处理.

中最适于用数学模型描述的项目之一，我们应用简单的物理知识建立了两个模型，对影响投掷距离的主要因素——初始速度、出手角度和出手高度进行定性和定量研究.

当一个实际问题归结为自变量与因变量之间函数形式的数学模型时，由于自变量不易精确控制到预先设定的或理想的数值(或者根本不能控制)，常常需要研究自变量在这个数值附近的微小改变对因变量的影响，因此敏感性分析是数学建模的重要环节之一.

复习题

1. 从(12)式出发，利用近似关系$(1+x)^{\frac{1}{2}}\approx 1+\frac{x}{2}$说明，模型二的最远投掷距离比模型一大约增加了一个出手高度.

2. 根据(6)式，用类似的方法说明，对任意出手角度，模型二的投掷距离比模型一大约增加了多少，并从图3与图2给以几何解释.

3.5 不买贵的只买对的

作为一个消费者，你在琳琅满目的市场里徘徊，如何花掉口袋里的钱，选择购买若干件商品.“不买贵的，只买对的”！买哪些商品、买多少才是“对的”？当然应该买最需要的、最有用的.“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你决定在市场里的选择.让我们从商品的效用说起[70].

效用函数 刚打完一场篮球，你又累又饿.朋友买来一袋切片面包，你狼吞虎咽地吃下一片又一片.开始的几片使饥饿感得到极大的缓解，生理上和心理上都很满足，继续吃下去，每一片面包产生的快感会逐渐减少，饱胀以后就不想再吃了.

为了定量地研究随着吃下越来越多的面包，所得到的满足感的变化趋势，用 $U(x)$ 表示吃了 x 片面包后获得的满足程度，可看作面包产生的效用，如表1中吃一片面包的效用为10(单位)，2片面包的效用为20，3片的效用为28，……，9片的效用为46，与8片相同，等于白吃，再吃，效用反而减少，说明你吃到第8片、最多第9片面包就应该停

止了.以下我们就不再考虑 $x=10$ 的情况. 表 1 中 $\Delta U(x)$ 表示每多吃一片面包所产生效用的增量,即 $\Delta U(x)=U(x)-U(x-1)$.图 1 是 $U(x)$ 和 $\Delta U(x)$ 的图形.由表 1、图 1 可以看到,$U(x)$ 是递增的,但是增长越来越慢,曲线呈上凸形状;$\Delta U(x)$ 大于或等于零,但它是递减的.

表 1 x 片面包产生的效用 $U(x)$ 及其增量 $\Delta U(x)$

面包数量 x	0	1	2	3	4	5	6	7	8	9	10
$U(x)$	0	10	20	28	35	40	43	45	46	46	44
$\Delta U(x)$		10	10	8	7	5	3	2	1	0	

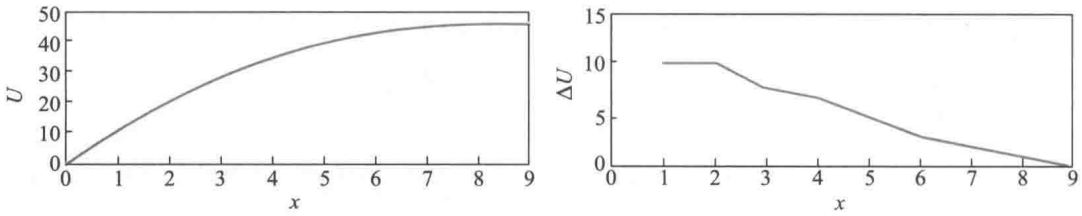


图 1 面包的 $U(x)$ 及 $\Delta U(x)$ 图形

像购买面包以满足人们生活需要这样的例子,在经济社会里比比皆是.经济学中把人们商品消费、服务消费所获得的生理、心理上的满足程度称为效用,引入效用函数(utility function)$U(x)$,表示数量为 x 的某种商品产生的效用,用约定的某个单位来度量①.为了数学研究的方便,将 $U(x)$ 视为连续、可微函数,其变化率 $\dfrac{\mathrm{d}U(x)}{\mathrm{d}x}$ 表示商品数量 x 增加 1 个单位时效用函数 $U(x)$ 的增量,称为边际效用(marginal utility).一种典型的效用函数数学表达式为

$$U(x)=ax^{\alpha},a>0,0<\alpha<1 \tag{1}$$

图 2 是(1)式中 $a=2$, $\alpha=\dfrac{1}{3}$ 的 $U(x)$ 和 $\dfrac{\mathrm{d}U(x)}{\mathrm{d}x}$ 的曲线.在理性消费情况(指消费增加导致效用也增加)下效用函数和边际效用具有以下性质:

1. 效用函数 $U(x)>0$,随着 x 的增加 $U(x)$ 递增,但增量越来越小.

2. 边际效用 $\dfrac{\mathrm{d}U(x)}{\mathrm{d}x}>0$,随着 x 的增加 $\dfrac{\mathrm{d}U(x)}{\mathrm{d}x}$ 递减,即边际效用的变化率 $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(\dfrac{\mathrm{d}U}{\mathrm{d}x}\right)$ 为负值.

"边际效用递减"是经济学中一条普遍的、重要的法则.除了人们的商品消费、服务消费以外,还可以举出一些用这条法则解释的例子.月收入 5 000 元的职员拿到 500 元的奖金,会十分高兴地盘算给亲人买些什么礼品,而月入 5 万元的小老板对于 500 元的外快就不会有多大兴趣.一些中老年人回忆起童年的困难生活时,常常对一碗红烧肉带

① 有些文献用 util, utils(复数)作为效用函数的单位,中文译名"尤特尔".

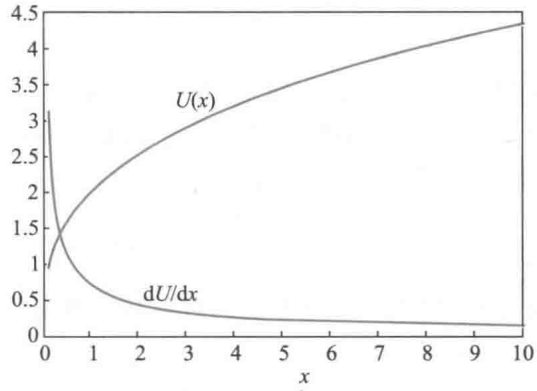


图 2 $U(x)=ax^{\alpha}\left(a=2,\alpha=\frac{1}{3}\right)$ 及 $\mathrm{d}U(x)/\mathrm{d}x$ 的曲线

来的满足和快感津津乐道,而当前餐桌上再好的美味佳肴也引不起太大兴趣.

效用函数和边际效用的这种特性可用两个数学式子简明地表示为

$$\frac{\mathrm{d}U}{\mathrm{d}x}>0,\frac{\mathrm{d}^2U}{\mathrm{d}x^2}<0 \tag{2}$$

容易验证(1)式的效用函数 $U(x)$ 满足(2)式.

当消费者面对两种商品时效用函数会是什么样子呢?

无差别曲线 这次朋友买来的除了面包还有香肠,你可以搭配着吃.如果 1 片面包加 4 根香肠,或者 4 片面包加 1 根半香肠,又或者 7 片面包加 1 根香肠,都能使你获得同样的满足,那么这几种面包加香肠组合的效用函数相等.用 $U(x,y)$ 表示 x 片面包和 y 根香肠组合的效用函数.在图 3 的 $x-y$ 平面中,上面 3 种组合分别用坐标点 A_1, A_2, A_3 表示,将它们连接成一条曲线,线上所有点(如 2 片面包加 2 根半香肠)对应的效用函数都等于某常数 u_1,记作 $U(x,y)=u_1$, 这条连线称为无差别曲线(指效用相同)或等效用线,与 2.3 节实物交换中介绍的曲线一样.

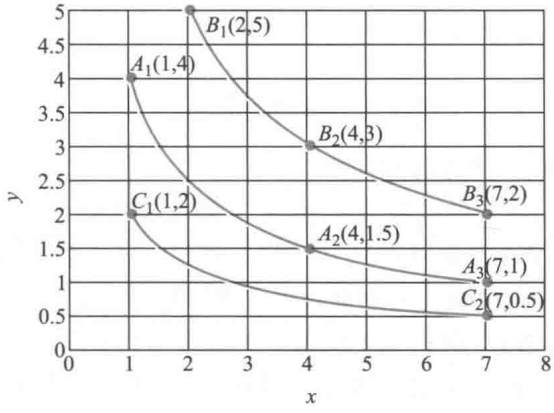


图 3 面包和香肠的无差别曲线

如果你拿取更多的面包和香肠(吃不完留下次吃),像图 3 的点 B_1(2 片面包加 5 根香肠)以及与 B_1 有同样效用函数值的 B_2, B_3,这 3 个点连成另一条无差别曲线,记作

$U(x, y)=u_2$，显然 $u_2>u_1$.可以画出更多的无差别曲线，例如由点 C_1（1片面包加2根香肠），C_2 连成的 $U(x, y)=u_3$，有 $u_3<u_1$.

两种商品的效用函数 $U(x, y)$ 是二元函数，当 $U(x, y)$ 的一个变量 y（或 x）固定时，看作另一个变量 x（或 y）的一元函数，所以 $U(x, y)$ 仍然具有前面关于 $U(x)$ 的“边际效用递减”性质，与(2)式相应，对二元函数 $U(x, y)$ 应该用偏导数表示为

$$\frac{\partial U}{\partial x}>0,\ \frac{\partial U}{\partial y}>0,\ \frac{\partial^2 U}{\partial x^2}<0,\ \frac{\partial^2 U}{\partial y^2}<0 \tag{3}$$

作为(1)式 $U(x)$ 形式的推广，一种典型的两种商品效用函数的数学表达式为

$$U(x,y)=ax^{\alpha}y^{\beta},\ a>0,0<\alpha,\beta<1 \tag{4}$$

图4是(4)式中 $a=1$，$\alpha=\frac{1}{3}$，$\beta=\frac{1}{2}$的无差别曲线 $U(x, y)=u(u=1, 2, 3)$.

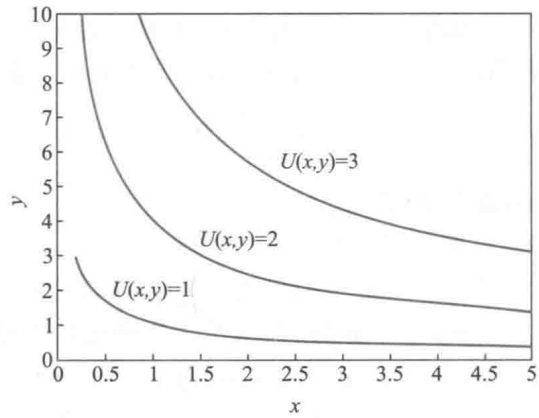


图4 $U(x,y)=ax^{\alpha}y^{\beta}\left(a=1,\alpha=\frac{1}{3},\beta=\frac{1}{2}\right)$ 的无差别曲线

无差别曲线 $U(x, y)=u$（u 为常数）在 $x-y$ 平面上表示为 y 对 x 的函数关系 $y=y(x)$，数学上 $U(x, y)=u$ 是 $y=y(x)$ 的隐函数，即有 $U(x, y(x))=u$.如(4)式中 $U(x, y)=u$ 可表示为

$$y=y(x)=\left(\frac{u}{a}\right)^{\frac{1}{\beta}}x^{-\frac{\alpha}{\beta}} \tag{5}$$

图4的3条曲线正是按照(5)式画出的.

无差别曲线是效用函数的几何表示.由图3、图4的几何直观可以看出，无差别曲线是下降（斜率为负）、下凸（凸向原点）、互不相交的，2.3节实物交换中曾给出简单解释，现在从经济学和数学的角度对无差别曲线的这些性质进行论证.

（1）下降（斜率为负） 既然无差别曲线上效用函数 $U(x, y)=u$ 不变，那么 x 的增加必导致 y 的减少.在经济学中对于两种可以相互替代的商品，x 增加1个单位引起 y 的减少量称为 x 对 y 的边际替代率（指绝对值），在图5无差别曲线 $U(x, y)=u$ 上的 P_1 点，边际替代率用 $-\frac{\Delta y}{\Delta x}$ 表示（$\Delta x>0$ 时 $\Delta y<0$），Δx 趋于0时 $\frac{\Delta y}{\Delta x}$ 趋于曲线斜率 $\frac{\mathrm{d}y}{\mathrm{d}x}$，边际替代率为 $-\frac{\mathrm{d}y}{\mathrm{d}x}$.注意到 x，y 的边际效用分别为 $\frac{\partial U}{\partial x}$，$\frac{\partial U}{\partial y}$，用 Δx 替代 $-\Delta y$ 后效用不变，即有

$\frac{\partial U}{\partial x}\Delta x=-\frac{\partial U}{\partial y}\Delta y$，于是有

$$\frac{\mathrm{d}y}{\mathrm{d}x}=-\frac{\partial U/\partial x}{\partial U/\partial y}<0 \tag{6}$$

即边际替代率$-\frac{\mathrm{d}y}{\mathrm{d}x}$等于边际效用$\frac{\partial U}{\partial x}$与$\frac{\partial U}{\partial y}$之比.

其实，在数学上，(6)式可以由隐函数 $U(x, y(x))=u$ 的求导公式直接得到.

(2) 下凸(凸向原点) 随着 x 增加，x 对 y 的"边际替代率递减"，也是经济学中一条重要法则，其数学意义为边际替代率$-\frac{\mathrm{d}y}{\mathrm{d}x}$对 x 的导数是负值，即$\frac{\mathrm{d}}{\mathrm{d}x}\left(-\frac{\mathrm{d}y}{\mathrm{d}x}\right)<0$，于是

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2}>0 \tag{7}$$

这正是曲线下凸(凸向原点)的条件.

(3) 互不相交 两条无差别曲线 $U(x, y)=u_1$，$U(x, y)=u_2(u_1\neq u_2)$ 如果相交，交点的效用函数将取两个不同的数值 u_1，u_2，这是不可能的.

当 $U(x, y)=u$ 取不同的效用函数值 u 时，得到一族无差别曲线.随着 u 的增加曲线向右上方移动，如图 5.至于曲线的具体形状，则由两种商品对消费者带来的效用以及消费者对两种商品的偏爱程度决定.如对(4)式表示的效用函数，无差别曲线的具体形状取决于参数 α，β 的大小.

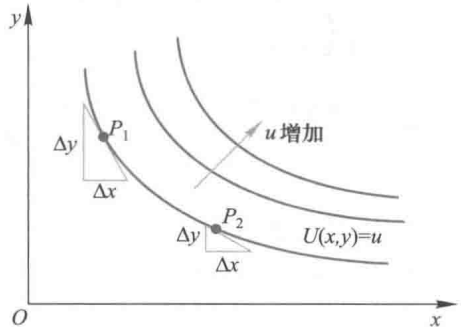


图 5 无差别曲线的下凸解释

效用最大化模型

如果消费者已经确定了对甲乙两种商品效用函数 $U(x, y)$ 和无差别曲线，那么他用一定数额的钱会购买多少数量的商品甲和乙呢？从效用最大化的角度出发，当然应该使 $U(x, y)$ 达到最大.

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 s，则他购买甲乙两种商品的数量 x，y 应满足

$$\begin{aligned}&\max U(x,y)\\&\text{s.t.}\quad p_1x+p_2y=s\end{aligned} \tag{8}$$

式中 s.t.是受约束于(subject to)的符号.这就是效用最大化模型.

先从几何图形上分析、求解这个模型.在图 6 中，模型(8)的目标函数值由一族下

降、下凸、互不相交的无差别曲线 $U(x, y)=u$ 给出，而由钱 s 和商品单价 p_1，p_2 构成的约束条件是直线 AB，称消费线，在 x 轴、y 轴上的截距分别为 $\frac{s}{p_1}$，$\frac{s}{p_2}$. 直线 AB 与某一条无差别曲线 l 相切，记切点为 $Q(x, y)$，称消费点，当购买甲乙两种商品的数量为 x，y 时，效用函数达到最大. 因为在 AB 与其他无差别曲线 l_1 的交点 Q_1，其效用函数值小于 Q 点的效用函数值.

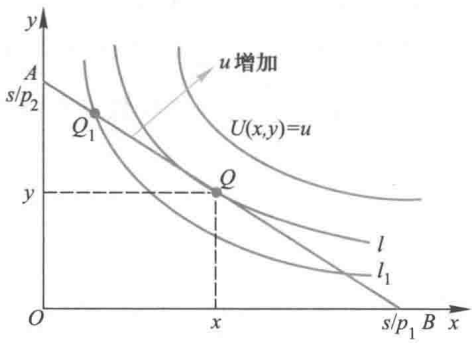


图 6 效用最大化模型的几何解法

如果知道效用函数 $U(x, y)$ 的数学表达式（如(4)式），可以按照二元函数条件极值的方法求解模型(8). 引入拉格朗日乘子 λ，构造函数

$$L(x,y,\lambda)=U(x,y)+\lambda(s-p_1x-p_2y)$$

由 $L(x, y, \lambda)$ 对 x，y 的偏导数等于 0，即

$$\frac{\partial L}{\partial x}=\frac{\partial U}{\partial x}-\lambda p_1=0,\frac{\partial L}{\partial y}=\frac{\partial U}{\partial y}-\lambda p_2=0$$

可得

$$\frac{\partial U/\partial x}{p_1}=\frac{\partial U/\partial y}{p_2}=\lambda \tag{9}$$

即最优解在甲乙两种商品的边际效用与二者价格之比相等时取得. (9)式又可记作

$$\frac{\partial U/\partial x}{\partial U/\partial y}=\frac{p_1}{p_2} \tag{10}$$

(10)式表明，当两种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，称效用最大化原理，是经济学中又一条重要法则. 这个结果与总钱数 s 无关，但可以给出购买商品甲和乙的数量之比 $\frac{x}{y}$ 或钱数之比 $\frac{p_1x}{p_2y}$.

几何解法得到的 $Q(x, y)$ 点与(10)式是一致的. 因为消费线 AB 的斜率是 $-\frac{p_1}{p_2}$，而由(6)式无差别曲线 $U(x, y)=u$ 的斜率是 $-\frac{\partial U/\partial x}{\partial U/\partial y}$.

如果某消费者对甲乙两种商品的效用函数如(4)式所示，则 $\frac{\partial U}{\partial x}=a\alpha x^{\alpha-1}y^{\beta}$，$\frac{\partial U}{\partial y}=$

$a\beta x^{\alpha}y^{\beta-1}$，由(10)式，最优解 x，y 应满足$\frac{\alpha y}{\beta x}=\frac{p_1}{p_2}$，或记作

$$\frac{p_1x}{p_2y}=\frac{\alpha}{\beta} \tag{11}$$

这表示按照效用最大化原理，他购买两种商品的钱数之比，等于效用函数(4)中的参数 α 与 β 之比，与商品价格无关，说明 α 与 β 分别是甲乙两种商品对消费者效用的度量，或者代表消费者对两种商品的偏爱.

模型(8)可以推广到购买 n 种商品的情况.记消费者的效用函数为 $U(x_1, x_2,\cdots, x_n)$，n 种商品的单价分别为 $p_1, p_2,\cdots, p_n$，消费者准备付出的钱为 s，则他购买 n 种商品的数量 $x_1, x_2,\cdots, x_n$ 满足

$$\begin{aligned}&\max \quad U(x_1,x_2,\cdots,x_n)\\&\text{s.t.}\sum_{i=1}^{n}p_ix_i=s\end{aligned} \tag{12}$$

用多元函数条件极值的方法求解模型(12)，其结果是(9)式或(10)式的推广，即

$$\frac{\partial U/\partial x_1}{p_1}=\frac{\partial U/\partial x_2}{p_2}=\cdots=\frac{\partial U/\partial x_n}{p_n} \tag{13}$$

式中边际效用与价格比值的含义是单位金额的边际效用.于是效用最大化原理又可表述为，当每种商品单位金额的边际效用均相等时，效用函数达到最大.

怎样才能“不买贵的，只买对的”！

市场上有草莓、芒果和橘子 3 种水果，每千克价格依次为 15 元、10 元和 5 元，李君准备花 100 元从中选择、采购.他应该怎样分配这笔钱呢？

按照效用最大化原理，李君应该先确定自己对这 3 种水果的效用函数或边际效用，然后根据(13)式决定 100 元钱的采购方案.记 3 种水果的价格分别为 p_1, p_2, p_3，购买数量分别为 x_1, x_2, x_3.有两种办法确定效用函数或边际效用.

第一种办法是采用现成的效用函数数学表达式，如(4)式类型的 $U(x_1, x_2, x_3)=ax_1^{\alpha}x_2^{\beta}x_3^{\gamma}$，$a>0$，$0<\alpha,\beta,\gamma<1$，李君需根据 3 种水果带来的效用或他对 3 种水果的偏爱程度确定参数 α,β,γ，然后计算边际效用$\frac{\partial U}{\partial x_1},\frac{\partial U}{\partial x_2},\frac{\partial U}{\partial x_3}$，由(13)式得到效用函数达到最大的条件为 $p_1x_1:p_2x_2:p_3x_3=\alpha:\beta:\gamma$，即按照 $\alpha:\beta:\gamma$ 的比例分配100 元钱.若李君认定 $\alpha=\frac{6}{10},\beta=\frac{1}{10},\gamma=\frac{3}{10}$，则他应该用 60 元买 4 kg 草莓、10 元买 1 kg 芒果、30 元买 6 kg 橘子.

采用这样的效用函数购买量 x_1, x_2, x_3 可以是连续的(若实际情况允许)，如果李君准备只花 60 元，那么他应该用 36 元买 2.4 kg 草莓、6 元买 0.6 kg 芒果、18 元买 3.6 kg 橘子.

第二种办法是李君根据 3 种水果带来的效用或他对 3 种水果的偏爱程度，分别给出每多购买 1 kg 草莓、1 kg 芒果和 1 kg 橘子时效用函数的增加，即边际效用，如他给出表 2 的数据，相当于$\frac{\partial U}{\partial x_1},\frac{\partial U}{\partial x_2},\frac{\partial U}{\partial x_3}$在 $x_1, x_2, x_3=1, 2, 3, \cdots$的取值.为了应用(13)式，先计算 3 种水果单位金额的边际效用(应将边际效用分别除以 $p_1=15$，$p_2=10$，$p_3=5$，由

于比例关系,等价于分别除以3, 2, 1),如此得到表3.

表2 草莓、芒果和橘子的边际效用

水果	数量/kg							
	1	2	3	4	5	6	7	8
草莓	24	21	20	18	16	15	13	10
芒果	12	10	9	8	6	5	3	2
橘子	15	12	10	9	7	6	5	5

表3 草莓、芒果和橘子单位金额的边际效用

水果	数量/kg							
	1	2	3	4	5	6	7	8
草莓	8	7	20/3	6	16/3	5	13/3	10/3
芒果	6	5	9/2	4	3	5/2	3/2	1
橘子	15	12	10	9	7	6	5	5

评注 无差别曲线的引入使得人们可以通过图形直观、定性地讨论效用最大化原理以及实际应用中的问题,至于效用函数的数学表达式虽然能够给出一些具体的定量结果,但是这些表达式(包括其中的参数)与商品本身的特性以及消费者的需要和偏爱密切相关,是很难确定的.

根据表3,可以按照3种水果单位金额边际效用从大到小的顺序,每次增加1 kg的购买量,直到花完准备付出的钱.如果李君准备花100元,那么第一笔钱买1 kg橘子(表3中15最大),第二笔钱又买1 kg橘子(表3中12次大),如此下去,结果仍是一共买4 kg草莓、1 kg芒果和6 kg橘子(表3中大于或等于6);而若他准备只花70元,那么将只能买3 kg草莓和5 kg橘子,不买芒果.

在第二种办法中,由于表2的边际效用是在离散点 x_1, x_2, $x_3=1, 2, 3, \cdots$ 得到的,所以隐含了购买量也是离散值的假定,这样,准备付出的钱(假如是60元)就不一定恰好花完.当然,可以先将边际效用拟合为适当的连续函数,再按第一种办法处理.

在这个例子中,芒果比橘子贵,但芒果买的比橘子少(钱不多时甚至不买芒果),这是因为芒果对李君的效用比橘子小,效用最大化的结果正是"不买贵的,只买对的"!

复习题

1. 检验(4)式表示的效用函数是否满足(6)、(7)式.

2. 如果商品甲的单价 p_1 下降,而商品乙单价 p_2、消费者准备付出的钱 s 和两种商品的效用函数 $U(x, y)$ 均不变,讨论效用最大化模型的最优解如何变化.

3. 如果收入 s 增加,而两种商品的单价 p_1, p_2 和效用函数 $U(x, y)$ 均不变,讨论效用最大化模型的最优解如何变化.

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叉角度,在消耗能量最小原则下应该取什么样的数值[7,64].

案例精讲 3-2
血管分支

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,将这种流动视为黏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 点分叉,并形成对称的几何形状.设粗细血管半径分别是 r 和 r_1,分叉处夹角是 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',$ACB(ACB')$ 的水平和竖直距离为 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 $q=2q_1$.

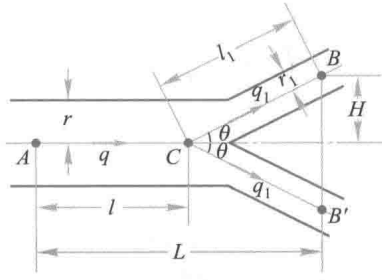


图 1 血管分支示意图

假设 2 使我们可以利用流体力学关于黏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

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

其中 Δp 是 A,C 两点的压力差,μ 是血液的黏性系数.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

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

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 $s=2\pi rl$,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

$$E_2 = br^{\alpha} l \tag{3}$$

其中 $1 \leqslant \alpha \leqslant 2$，$b$ 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

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

由图 1 所示的几何关系不难得到

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

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

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

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

$$\begin{cases} -\dfrac{4kq^2}{r^5} + b\alpha r^{\alpha-1} = 0 \\ -\dfrac{kq^2}{r_1^5} + b\alpha r_1^{\alpha-1} = 0 \end{cases} \tag{7}$$

从方程(7)可解出

$$\frac{r}{r_1} = 4^{\frac{1}{\alpha+4}} \tag{8}$$

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

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

将(8)代入(9)式，则

$$\cos\theta = 2^{\frac{\alpha-4}{\alpha+4}} \tag{10}$$

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

$$1.26 \leqslant \frac{r}{r_1} \leqslant 1.32, \quad 37^{\circ} \leqslant \theta \leqslant 49^{\circ} \tag{11}$$

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 次分叉，将(8)式反复利用 n 次可得

$$\frac{r_{\max}}{r_{\min}}=4^{\frac{n}{\alpha+4}} \tag{12}$$

$r_{\max}/r_{\min}$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 $1\leqslant\alpha\leqslant 2$，所以按照这个模型，狗的血管应有 25~30 次分叉.又因为当血管有 n 次分叉时血管总条数为 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.

复习题

如果模型假设中不作(3)式的简化，而是在消耗的能量中既有与 r^2 成正比、又有与 r 成正比的部分，试分析对建模和求解有什么影响.

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[16].

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.

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).

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.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

船速/(km·h^{-1})	冰山体积/m^3		
	10^5	10^6	10^7
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.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

船速/(km·h^{-1})	与南极距离/km		
	0	1 000	>4 000
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 (m/天),船速为 u (km/h),拖船与南极距离为 d (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 $0\leqslant d\leqslant 4\ 000$ km时 r 与 d 成正比,而当 $d>4\ 000$ km 时 r 与 d 无关,即设

$$r=\begin{cases}a_1 d(1+bu), & 0\leqslant d\leqslant 4\ 000\\ a_2(1+bu), & d>4\ 000\end{cases} \tag{1}$$

其中 a_1,a_2,b 为待定参数.这可以解释为,$0\leqslant d\leqslant 4\ 000$ 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 $d>4\ 000$ 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

$$a_1=6.5\times10^{-5},\quad a_2=0.2,\quad b=0.4 \tag{2}$$

当拖船从南极出发航行第 t 天时,与南极的距离为

$$d=24ut \tag{3}$$

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

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.

$$r_t=\begin{cases}1.56\times10^{-3}u(1+0.4u)t, & 0\leqslant t\leqslant\dfrac{1\,000}{6u}\\ 0.2(1+0.4u), & t>\dfrac{1\,000}{6u}\end{cases}\tag{4}$$

记第 t 天冰山半径为 R_t，体积为 V_t，则

$$R_t=R_0-\sum_{k=1}^{t}r_k\tag{5}$$

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

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

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

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

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

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

$$c_1=0.3,\quad c_2=6,\quad c_3=-1\tag{9}$$

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

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

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

$$f(V_0)=\begin{cases}4.0, & V_0\leqslant5\times10^5\\ 6.2, & 5\times10^5<V_0\leqslant10^6\\ 8.0, & 10^6<V_0\leqslant10^7\end{cases}\tag{11}$$

又因为当船速为 u(km/h)时冰山抵达目的地所需天数为 $T=\dfrac{9\,600}{24u}=\dfrac{400}{u}$，所以租金费用为 $\dfrac{400f(V_0)}{u}$.而整个航程的燃料消耗为 $\sum_{t=1}^{T}q(u,V_0,t)$，由(10)式得运送冰山的总费用为

① 与(2)式一样，这个结果实际上也是凑出来的.

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

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

$$V(u,V_0,T)=\frac{4\pi}{3}\left(\sqrt[3]{\frac{3V_0}{4\pi}}-\sum_{t=1}^{T}r_t\right)^3 \tag{13}$$

注意到假设(3),则得到水的体积为

$$W(u,V_0)=\frac{3.4\ \pi}{3}\left(\sqrt[3]{\frac{3V_0}{4\pi}}-\sum_{t=1}^{T}r_t\right)^3 \tag{14}$$

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

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

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 $V_0=10^7\ \mathrm{m}^3$(当然要租用大型拖船),船速 $u=4\sim5$ 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	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

评注 这个模型的思路是简单的,建模方法有两点值得注意:一是根据有限的数据(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.

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.

复习题

将冰山看作长方体更切合实际,试对冰山融化状况作适当假定后建立模型,用本节的数据求解.

3.8 影院里的视角和仰角

到影院看电影时有人愿意坐在前排，有人愿意坐在后面，更多的人选择中间的座位.除了个人的原因之外，前后排的主要差别在于视角和仰角.视角是观众眼睛到屏幕上、下边缘视线的夹角，视角大画面看起来饱满；仰角是观众眼睛到屏幕上边缘视线与水平线的夹角，仰角太大会使人的头部过分上仰，引起不舒适感.从影院设计者的角度出发，应该在总体上使观众的视角尽可能大①，仰角则应有一定限制.这个模型简单地讨论影院屏幕和座位设计中的一些问题，对于观众选择座位也有一定意义[70].

作为一种简化情况，图 1 是垂直于屏幕和地面的影院纵向剖面示意图，即只考虑前后座位观看屏幕的差别，不管左右座位及屏幕横向方面的因素.影院座位的地板线与水平面有一个角度 θ，主要是避免前排观众遮挡后排的视线.

图 1 中画出了某一排观众的视角 α 和仰角 β，可以看到，影响视角 α 和仰角 β 的因素很多，除了观众在第几排座位以外，有屏幕高度 h、屏幕下边缘距地面高度 b、第一排座位与屏幕水平距离 d、座位地板线与地面夹角 θ、两排座位间距② q、观众眼睛到地板的距离 c 等，其中 h，q，c 基本上是固定的，如果总排数一定，d 的改变余地也不大，于是在影院设计时只有 b 和 θ 可以在一定范围内调整.本节将 h，d，q，c 的数值以及座位总排数 n 固定，只讨论 b 和 θ 的变化对观众满意程度的影响，研究 b 和 θ 取什么数值时（在设定范围内）全体观众的满意程度最高.

从设计者的角度，按照以下几个方面考虑观众的满意程度：

1. 全体观众视角的平均值尽量大，而各排座位视角的分散程度尽量小.
2. 各排座位的仰角基本上不要超过 30°，可以允许 1~2 排的例外.
3. 前排观众不应遮挡后排观众的视线.

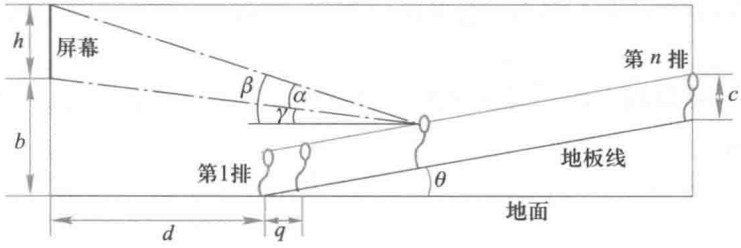


图 1 影院纵向剖面示意图

问题分析 记影院各排座位序号为 $k=1,2,\cdots,n$，从图 1 可以直观地看出，随着座位序号 k 的增加，视角 α 和仰角 β 均在减小.全体观众视角的平均值为 $k=1$ 到 n 排视角的均值，各排座位视角的分散程度不妨借用概率统计中这 n 个视角的均方差来度量.因为均值越大且均方差越小，全体观众满意程度就越高，所以观众的满意程度可以定义为

① 虽然视角 α 也不宜过大，但实际上还远未到过大的情况，所以在一定范围内仍是越大越好.

② 虽然地板线是倾斜的，但座椅放在台阶形成的平面上，所以排间距仍为水平间距.

各排视角的均值与均方差之比，概率统计中称为变异系数①.这个优化问题的目标函数就取作各排座位视角的变异系数.

问题的第一个约束条件是仰角 $\beta \leqslant 30°$，且允许 1~2 排不满足.由于 β 随着 k 的增加而减小，所以只需检查前 3 排仰角 β 的数值即可.

第二个约束条件是前排观众不遮挡后排的视线，需要在观众眼睛到地板距离 c 之上设定眼睛到头顶的高度，使后排观众眼睛到屏幕下边缘的视线，在前排观众头顶之上.直观地看，只要最后一排满足这个条件，所有座位都会满足.

目标函数的自变量即设计者的控制变量，是屏幕下边缘距地面高度 b 及座位地板线与地面夹角 θ.

模型假设 作为固定参数，设屏幕高度 $h=2.5$ m，第 1 排座位与屏幕水平距离 $d=6$ m，两排座位间距 $q=0.8$ m，观众眼睛到地板的距离 $c=1.1$ m，观众眼睛到头顶的高度 $c_1=0.1$ m，座位总排数 $n=16$.作为控制变量，设屏幕下边缘距地面高度 b 在 2 m 至 3 m 变化，座位地板线与地面夹角 θ 在 10°至 20°变化.

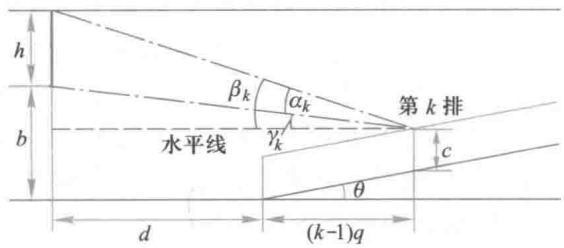


图 2 仰角和视角的计算

记观众眼睛到屏幕下边缘的视线与水平线的夹角为 γ，称下仰角，由图 1 可见，视角 $\alpha=\beta-\gamma$.需要注意的是，当这条视线在水平线之下时（可能有这种情况）下仰角 γ 应取负值.对仰角 β（以下称上仰角）也应作此规定，不过实际上不大可能出现 β 为负的情况.

模型建立 截取图 1 的一部分如图 2，计算第 k 排座位的下仰角 γ_k、上仰角 β_k 和视角 α_k.利用三角关系容易得到

$$\tan \gamma_k = \frac{b-c-(k-1)q\tan\theta}{d+(k-1)q} \tag{1}$$

$$\tan \beta_k = \frac{b-c-(k-1)q\tan\theta+h}{d+(k-1)q} \tag{2}$$

$$\alpha_k = \beta_k - \gamma_k \tag{3}$$

记各排视角的均值为 $m(\alpha)$，均方差为 $s(\alpha)$，变异系数为 $v(\alpha)$，则

$$m(\alpha)=\frac{1}{n}\sum_{k=1}^{n}\alpha_k, s(\alpha)=\sqrt{\frac{1}{n-1}\sum_{k=1}^{n}[\alpha_k - m(\alpha)]^2}, v(\alpha)=\frac{m(\alpha)}{s(\alpha)} \tag{4}$$

$v(\alpha)$是这个优化模型的目标函数，其中 h，d，q，c 和 n 为常数，要确定 b 和 θ（在设定区间内）使 $v(\alpha)$最大.

① 均方差和变异系数在概率统计中都是对随机变量而言，不过这里的变量是确定性的.

按照前面的分析，模型对于仰角 β 的约束条件为

$$\beta_k \leqslant 30^\circ, k=3,\cdots,n \tag{5}$$

对于前排不遮挡后排视线的约束条件可作以下计算.图 3 表示的是最后一排与前一排之间的遮挡关系，图中 γ_n 是从第 n 排观众眼睛 A 点到屏幕下边缘视线的下仰角，δ 是连接 A 点和第 $n-1$ 排观众头顶 B 点连线与水平线的夹角，当连线在水平线以下时 δ 为负.视线不被遮挡的条件是 $\gamma_n>\delta$，或者

$$\tan\gamma_n>\tan\delta \tag{6}$$

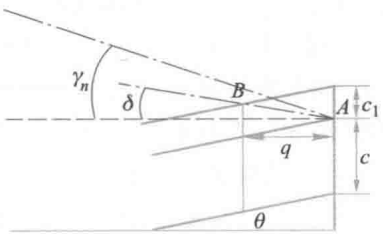


图 3 前排不遮挡后排视线的计算

其中 $\tan\gamma_n$ 由(1)式以 $k=n$ 代入得到，$\tan\delta$ 可根据图 3 写出

$$\tan\delta=\frac{c_1-q\tan\theta}{q} \tag{7}$$

于是(6)式等价于

$$\frac{b-c-(n-1)q\tan\theta}{d+(n-1)q}>\frac{c_1-q\tan\theta}{q} \tag{8}$$

不难看出，由于 γ_k 是 k 的减函数，而 δ 是与 k 无关的常数，所以(6)式或(8)式就是全体观众视线不被遮挡的条件.

综上，建立了以(1)~(4)式为目标函数、(5)式和(6)式为约束条件、b 和 θ 为自变量的优化模型.

模型分析 在求解之前，让我们对模型作一个定性分析，讨论 b 和 θ 的变化对目标函数和约束条件会有什么影响.

由(1)式和(2)式容易知道，b 增加使 γ_k 和 β_k 均增加，而 θ 增加使 γ_k 和 β_k 均减少，这也可以从图 2 上直观地看出.视角 α_k 是 γ_k 与 β_k 之差((3)式)，b 和 θ 的增加对 α_k 的影响如何呢？根据三角公式 $\tan(\beta-\gamma)=\dfrac{\tan\beta-\tan\gamma}{1+\tan\beta\tan\gamma}$，从(1)~(3)式可得

$$\tan\alpha_k=\frac{h[d+(k-1)q]}{[d+(k-1)q]^2+[b-c+h-(k-1)q\tan\theta][b-c-(k-1)q\tan\theta]} \tag{9}$$

对(9)式分析，b 增加会使 α_k 减少，而 θ 增加会使 α_k 增加.你的直观感觉是这样吗？

再看(4)式，各排视角 α_k 的增减当然导致均值 $m(\alpha)$ 的增减.多数情况下，均方差 $s(\alpha)$ 也会有同样的增减.但是变异系数 $v(\alpha)$ 作为二者之比，将有什么样的变化就无法作定性分析了.

对于仰角的约束条件(5)式，只能说 b 越小、θ 越大，越容易满足.

对于遮挡的约束条件(8)式，作代数运算可化为

$$b+d\tan\theta>c+c_1(n-1)+\frac{c_1d}{q} \tag{10}$$

可知，b 和 θ 越大，这个条件越容易满足，并且当模型的已知常数 c，c_1，n 变小及 q 变大时，也会使条件(10)容易满足，这些都是与直观相符的.

模型求解 用导数为零的解析方法求解上述模型显然是难以实现的，我们转向数

值搜索方法，在 b 和 θ 的设定区间内取一系列离散值，先按照(1)~(4)式计算目标函数 $v(\alpha)$，找出最优解，再检验约束条件(5)式和(6)式是否满足，如果不满足，将最优解放宽后再作检验，直到找出满足约束条件的最优解.

根据(1)~(4)式和给定的 h, d, q, c, n 数值，对屏幕下边缘距地面高度 b 在 2 m 至 3 m 区间、座位地板线与地面夹角 θ 在 10°至 20°区间离散化，编程计算得到视角 α_k 的均值 $m(\alpha)$、均方差 $s(\alpha)$ 及变异系数 $v(\alpha)$，表 1 只给出了模型的目标函数 $v(\alpha)$ 的结果.

计算结果显示，随着 b 的增加，均值 $m(\alpha)$ 和均方差 $s(\alpha)$ 都在减少，但 $m(\alpha)$ 只减少约 5%，而 $s(\alpha)$ 减少约 15%，结果导致 $v(\alpha)$ 一直在增加，最大值位于 b 的区间端点 3.0 m，如表 1 所示；随着 θ 的增加，$m(\alpha)$ 和 $s(\alpha)$ 都在增加，导致 $v(\alpha)$ 先增后减(当 $b>2.3$ m)，最大值在 θ 区间[10°, 20°]内部，表 1 中 $b=3.0$ m 时 $v(\alpha)$ 最大值在 $\theta=13°$ 达到(表中黑体数字).

表 1 各排座位视角变异系数的计算结果

b/m	θ										
	10°	11°	12°	13°	14°	15°	16°	17°	18°	19°	20°
2.0	3.145 2	3.144 3	3.142 3	3.139 2	3.135 1	3.130 0	3.123 7	3.116 3	3.107 7	3.098 1	3.087 2
2.1	3.179 5	3.178 9	3.177 2	3.174 5	3.170 7	3.165 8	3.159 8	3.152 6	3.144 3	3.134 8	3.124 0
2.2	3.215 4	3.215 2	3.213 9	3.211 5	3.208 1	3.203 4	3.197 7	3.190 7	3.182 6	3.173 2	3.162 6
2.3	3.253 0	3.253 2	3.252 3	3.250 3	3.247 1	3.242 8	3.237 3	3.230 6	3.222 6	3.213 4	3.202 9
2.4	3.292 4	3.293 0	3.292 5	3.290 8	3.288 0	3.284 0	3.278 7	3.272 2	3.264 5	3.255 4	3.245 1
2.5	3.333 6	3.334 7	3.334 6	3.333 2	3.330 7	3.327 0	3.322 0	3.315 8	3.308 2	3.299 4	3.289 1
2.6	3.376 7	3.378 2	3.378 5	3.377 5	3.375 4	3.372 0	3.367 3	3.361 3	3.353 9	3.345 2	3.335 1
2.7	3.421 7	3.423 6	3.424 3	3.423 8	3.422 0	3.418 9	3.414 5	3.408 7	3.401 6	3.393 1	3.383 1
2.8	3.468 6	3.471 0	3.472 2	3.472 0	3.470 6	3.467 9	3.463 8	3.458 3	3.451 3	3.443 0	3.433 1
2.9	3.517 6	3.520 4	3.522 1	3.522 4	3.521 3	3.518 9	3.515 1	3.509 9	3.503 2	3.495 0	3.485 3
3.0	3.568 6	3.572 0	3.574 1	**3.574 8**	3.574 2	3.572 1	3.568 6	3.563 7	3.557 2	3.549 2	3.539 6

表 2 在 $b=3.0$ m，$\theta=13°$ 处各排座位仰角的计算结果

k	1	2	3	4	5	6	7	8
β_k	36.253 8	31.794 7	27.939 0	24.600 5	21.700 5	19.170 1	16.951 2	14.995 1
k	9	10	11	12	13	14	15	16
β_k	13.261 5	11.717 3	10.335 0	9.091 6	7.968 4	6.949 4	6.021 4	5.173 0

接着需要检查在 $b=3.0$ m，$\theta=13°$ 处的约束条件是否满足.仰角 β_k 的计算结果如表 2，可知除 β_1, β_2 外均满足(5)式.按照(1)式计算出 $\tan\gamma_n=-2.768\ 5$，按照(7)式计算出 $\tan\delta=-6.043\ 3$，可知满足(6)式.于是 $b=3.0$ m，$\theta=13°$ 确是整个模型的最优解.

表 3 是在最优解 $b=3.0$ m，$\theta=13°$ 处视角 α_k 及均值 $m(\alpha)$、均方差 $s(\alpha)$ 的计算结果.表 2、表 3 的仰角 β_k 和视角 α_k 还表示在图 4 中.

表 3 在 $b=3.0$ m，$\theta=13°$ 处各排座位视角的计算结果

k	1	2	3	4	5	6	7	8	9
α_k	18.682 6	17.637 1	16.552 1	15.497 5	14.506 7	13.592 7	12.757 9	11.999 0	11.310 3
k	10	11	12	13	14	15	16	$m(\alpha)$	$s(\alpha)$
α_k	10.685 5	10.117 8	9.601 2	9.130 1	8.699 3	8.304 4	7.941 5	12.313 5	3.444 5

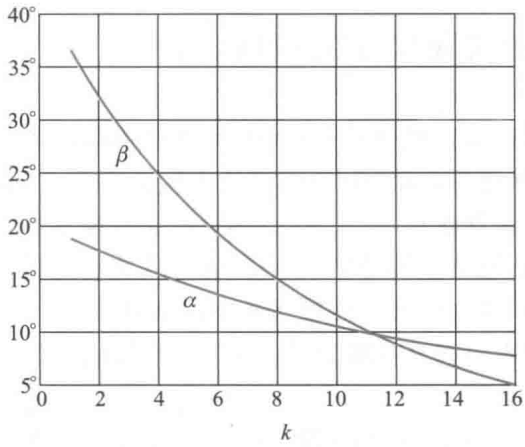


图 4 最优解处各排座位的仰角 β 和视角 α

结果分析 首先根据表 1 作敏感性分析.在最优解 $b=3.0$ m，当 $\Delta b=0.1$ m 即 $\frac{\Delta b}{b}\approx 3.3\%$ 时，$\frac{\Delta v}{v}\approx 1.5\%$，相当于自变量 b 的 1%的改变引起目标函数不到 0.5%的变化.在最优解 $\theta=13°$，当 $\Delta\theta=1°$ 即 $\frac{\Delta\theta}{\theta}\approx 8\%$ 时，$\frac{\Delta v}{v}\approx 0.2\%$，相当于自变量 θ 的 1%的改变引起目标函数不到 0.03%的变化.b 对目标函数的影响比 θ 的影响大十几倍.

由表 2 看到，在最优解 $b=3.0$ m，$\theta=13°$ 处，第 1，2 排座位的仰角大于 30°，刚刚满足约束(5)式.如果对仰角的要求更严格，需要调整 b 和 θ(复习题 1).

由表 3 看到，最优解处各排视角都在 8°到 19°之间，相邻两排视角相差不到 1°.

图 4 直观地给出各排座位仰角和视角的变化情况，观众可以据此选择适合自己的座位.从图上看，仰角下降很快，而视角变化不大，所以建议一般观众可以选择仰角下降变缓的第 10 排左右.

从结果分析看出，地板线与地面夹角的变化对目标函数的影响较小，不妨在设定范围内固定某个数值，而将其他因素如第 1 排座位与屏幕水平距离作为新的控制变量，重新研究这个优化问题.

在影院屏幕和座位设计中，控制变量、目标函数、约束条件都可以因考虑问题的目的不同、角度不同而改变.此外，这个模型仅仅讨论了影院座位的纵向设计，读者不妨拓宽眼界和思路，研究座位的横向设计，甚至座位的二维平面设计.

评注 影院屏幕和座位设计中要考虑的因素很多，作为初步的简化，这里只选取各排座位的视角和仰角为控制目标，屏幕下边缘距地面高度和座位地板线与地面夹角为控制变量，根据设计者对全体观众的视角及其分散程度的要求，用视角均值和均方差之比构造目标函数，并按照对仰角大小和视线遮挡的限制给定约束条件，建立了完整的优化模型.鉴于模型表述的复杂性，采用数值搜索方法通过编程计算求得最优解.

在模型的定性分析中，研究了控制变量的变化对目标函数和约束条件的影响，所得结论与直观和常识相符合，可作为检验模型正确性的一个方面，是建模过程的重要环节之一.模型定量结果与定性分析的相互印证、控制变量的敏感性分析，以及对各排座位仰角和视角的讨论，更丰富了建模的成果，拓广了模型的应用.

复习题

1. 如果严格限制各排座位的仰角不超过30°(误差0.5°),如何调整 b 和 θ?

2. 将 θ 设定在15°,而第一排座位与屏幕水平距离 d 可在5 m至7 m间调整,先定性分析 d 的变化对仰角和视角的影响,再重新求解模型,得出 b 和 d 的最优解.

3.9 易拉罐形状和尺寸的最优设计

我们只要稍加留意就会发现销量很大的饮料(例如饮料量为355 mL的可口可乐、青岛啤酒等)的饮料罐(即易拉罐)的形状和尺寸几乎都是一样的.看来,这并非偶然,这应该是某种意义下的最优设计.当然,对于单个的易拉罐来说,这种最优设计可以节省的钱可能是很有限的,但是如果是生产几亿,甚至几十亿个易拉罐的话,可以节约的钱就很可观了(引自全国大学生数学建模竞赛2006年C题).

现在就请你们小组来研究易拉罐的形状和尺寸的最优设计问题.具体说,请你们完成以下的任务:

1. 取一个饮料量为355 mL的易拉罐,例如355 mL的可口可乐饮料罐,测量你们认为验证模型所需要的数据,例如易拉罐各部分的直径、高度、厚度等,并把数据列表加以说明;如果数据不是你们自己测量得到的,那么你们必须注明出处.

2. 设易拉罐是一个正圆柱体.什么是它的最优设计?其结果是否可以合理地说明你们所测量的易拉罐的形状和尺寸?例如说,半径和高之比,等等.



图1 易拉罐纵断面

3. 设易拉罐的中心纵断面如图1所示,即上面部分是一个正圆台,下面部分是一个正圆柱体.什么是它的最优设计?其结果是否可以合理地说明你们所测量的易拉罐的形状和尺寸?

4. 利用你们对所测量的易拉罐的洞察和想象力,做出你们自己的关于易拉罐形状和尺寸的最优设计.

本节以发表在《工程数学学报》2006年增刊上参赛学生的优秀论文和评述文章为基本材料,加以整理、归纳和提高,介绍建模的全过程.

问题分析 很多微积分教材在导数应用的章节中都有这样的极值问题:设计一个容积固定的、有盖的圆柱形容器,若侧壁及底、盖的厚度都相同,问容器高度与底面半径之比为多少时,所耗的材料最少?

其求解过程简述如下:由于侧壁及底、盖的厚度相同,容器所耗材料可用总面积表示.记容器底面半径为 r,高度为 h,侧壁及盖、底的总面积为 S,容器容积为 V,则

$$S=2\pi rh+2\pi r^2 \tag{1}$$

$$V=\pi r^2 h \tag{2}$$

问题化为在 V 固定的条件下求 r,h 满足什么关系可使 S 最小.

从(2)式解出 h 代入(1)式右端,有

$$S=\frac{2V}{r}+2\pi r^2 \tag{3}$$

S 对 r 求导并令其等于 0,得

$$\frac{\mathrm{d}S}{\mathrm{d}r}=-\frac{2V}{r^2}+4\pi r=0 \tag{4}$$

不必解出 r 与 V 的关系,可以直接从(4)式得到 $\frac{V}{\pi r^2}=2r$,代入 $h=\frac{V}{\pi r^2}$ 即得

$$h=2r \tag{5}$$

表明容器高度与底面直径相等时所耗材料最少.

当用这个结果审视日常所见的易拉罐时,不相符合之处十分明显.通常易拉罐的高度比底面直径大得多.究其原因,抛开设计美学等因素的考虑,单纯从节省材料的角度分析,如果罐底、盖的厚度比侧壁大,就应该增加高度、减少底面直径.本题要求先测量易拉罐的直径、高度、厚度等数据,正是让大家在实际的感性认识的基础上,注意到易拉罐各个表面的材料厚度的不同,建模时应考虑这个因素的影响.

粗略地看,易拉罐是一个正圆柱体.利用简单的几何知识,容易写出它的体积和表面积,在各表面厚度不同的情况下,建立类似于(1)、(2)式的模型.精细一些,易拉罐的形状是圆柱体上面有一个小的圆台,由于独立变量数量的增加,建立的模型是多元函数,求解会复杂一些.至于如何发挥洞察力和想象力,做出易拉罐形状和尺寸的最优设计,容易想到又比较符合实际的是将顶部的小圆台改为球台,看看能否有所改进.

数据测量 对形如图 1 的易拉罐的各项尺寸用适当的仪器进行测量,为精确起见,取 5 只罐子测量后计算平均值,得表 1①.

表 1 易拉罐(可口可乐)的各项尺寸

罐高 /mm	圆柱高 /mm	圆柱直径 /mm	圆台高 /mm	顶盖直径 /mm	罐壁厚 /mm	顶盖厚 /mm	罐底厚 /mm	罐内容积 /cm^3
120.6	110.5	66.1	10.1	60.1	0.103	0.306	0.300	364.8

表 1 数据中值得注意的有两点:一是易拉罐底、盖的厚度约为罐壁的 3 倍,相差很大,这大概是制作工艺和使用方便的需要,正因为这样,易拉罐尺寸的优化设计才与前面简述的微积分中的极值问题有差别;二是圆台高不到圆柱体高度的 10%,顶盖与圆柱的直径相差也是 10%,所以把圆台近似作圆柱处理误差很小.另外,表中罐内容积是用量筒测出的,不妨用罐的各项尺寸粗略地核算一下:罐内圆柱直径为 $66.1-0.1\times2=65.9$(mm),罐内高度为 $120.6-0.3\times2=120$(mm),所以罐内体积应是 $\pi\times33^2\times120=410.3\times10^3$($mm^3$),比测量数据大 12%,这大概是测量误差造成的.

下面按照对题目中第 2, 3, 4 个问题的分析,依次讨论圆柱模型、圆台模型和球台模型.

圆柱模型 将易拉罐顶部的小圆台近似于圆柱,与下面圆柱体的直径相同.

---

① 《工程数学学报》2006 年增刊上郭文飞等的论文给出的结果.从操作的方便出发,认为是从罐的外部进行测量的(原文未注明).

记圆柱半径为 r,高度为 h,侧壁厚度为 b,底、盖的厚度分别为 kb 和 k_1b.在罐壁与盖、底厚度不同的情况下,假定所耗材料用侧壁、底、盖的面积乘以厚度得到的体积表示,记作 SV_1,罐的容积记作 V_1,则①

$$SV_1=2\pi rhb+\pi r^2(kb+k_1b) \tag{6}$$

$$V_1=\pi r^2h \tag{7}$$

其中 b, k 和 k_1 为已知常数.易拉罐尺寸的最优设计是在 V_1 固定的条件下,求 r,h 满足什么关系可使 SV_1 最小.经过与(3)式至(5)式完全相似的推导可得

$$h=(k+k_1)r \tag{8}$$

这个结果表明,易拉罐高度是底面半径的多少倍,取决于罐底和罐盖的厚度比侧壁厚度大多少.显然,这与人们的直观认识是一致的.

如果根据表1中的测量数据,k 和 k_1 都接近3,按照(8)式将有 $h=6r$,圆柱高度为直径的3倍②.但是表1的数据又显示,罐高大致是直径的2倍,这比较符合我们日常看到的情景.对于这里出现的矛盾,编者认为可能有两种原因:一是表1中侧壁及底、盖厚度的测量存在较大误差③,二是罐的实际加工制作还有除节省材料外的其他考虑.

圆台模型 实际生活中多数易拉罐顶部有一个小圆台,圆台下面与圆柱相接.设圆柱半径、高度、侧壁和底部厚度仍采用圆柱模型的记号 r, h, b 和 kb,圆台上底面(即罐盖)的半径、高度和厚度分别记作 r_1, h_1 和 k_1b,再设圆台侧壁的厚度与圆柱相同,由圆柱和圆台组成的易拉罐所耗材料的体积记作 SV_2,罐的容积记作 V_2,则④

$$SV_2=2\pi rhb+\pi r^2kb+\pi r_1^2k_1b+\pi\sqrt{(r-r_1)^2+h_1^2}\,(r+r_1)b \tag{9}$$

$$V_2=\pi r^2h+\frac{\pi h_1(r^2+r_1^2+rr_1)}{3} \tag{10}$$

其中 b, k 和 k_1 为已知常数.易拉罐尺寸的最优设计是在 V_2 固定的条件下,求 r,h, r_1, h_1 满足什么关系可使 SV_2 最小,这个模型实际上只有3个独立变量⑤.

这是一个多元函数的条件极值问题,理论上可用拉格朗日乘子法求解.令

$$F(r,h,r_1,h_1,\lambda)=SV_2(r,h,r_1,h_1)+\lambda V_2(r,h,r_1,h_1) \tag{11}$$

(9)、(10)式的最优解由

$$\frac{\partial F}{\partial r}=0,\frac{\partial F}{\partial h}=0,\frac{\partial F}{\partial r_1}=0,\frac{\partial F}{\partial h_1}=0,\frac{\partial F}{\partial \lambda}=0 \tag{12}$$

确定.但是由于这些方程的复杂性,由(11)、(12)式难以求出解析解,转而根据(9)、(10)式直接求如下约束极小问题的数值解:

---

① 若半径 r、高度 h 为罐内尺寸,下式中侧壁体积 $2\pi rhb$ 的精细表示是 $[\pi(r+b)^2-\pi r^2][h+k_1b+kb]$,展开后,$b$ 的一次项即 $2\pi rhb$,而由于 b 比 r、h 小得多,b 的二次、三次项可略去.

② 《工程数学学报》2006年增刊上郭文飞等的论文中,假设侧壁和罐底的厚度均为 b,顶盖厚度为 $3b$,得到 $h=4r$,圆柱高度为直径的2倍.但是这样做与测量数据不统一.

③ 《工程数学学报》2006年增刊上叶其孝教授的评论文章中,也提供了侧壁、底、盖厚度的数据,其中 k 和 k_1 都比3小一些.

④ 圆台以及下面用到的球台的侧面积和体积公式,都可以在数学手册上查到.

⑤ 《工程数学学报》2006年增刊上郭文飞等的论文中,用圆台侧壁与底面的夹角 θ 代替高度 h_1,这样做是等价的.

$$\begin{aligned}\min\quad & SV_2(r,h,r_1,h_1)\\ \text{s.t.}\quad & V_2(r,h,r_1,h_1)=V_0\\ & r>0,h>0,r_1>0,h_1>0\end{aligned}\tag{13}$$

其中 b, k, k_1 和容积 V_0 可由测量数据给出.

对(13)式用 LINGO 软件编程计算,得到的结果是 $r=31.43$, $h=108.34$, $r_1=0$, $h_1=28.10$(单位:mm,下同),即易拉罐的顶部从圆台退化为一个圆锥.分析出现这种结果的原因,是因为假定圆台侧壁与下面圆柱的厚度 b 一样,而圆台上底面的厚度是 k_1b,为 b 的 3 倍,从节省材料的角度,自然要尽量减少圆台上底面的面积.

由于罐盖要安装拉环以及工艺、美观等方面的考虑,圆台上底面的半径 r_1 应该有一个下限,不妨设 $r_1\geqslant 20$.将(13)式中的 $r_1>0$ 改成这个约束后,得到的结果是 $r=31.62$, $h=104.52$, $r_1=20$, $h_1=17.29$,目标函数值(材料体积)为3 732(mm^3).

另外,圆台侧壁的厚度可能与圆柱侧壁厚度不同,圆台的高度也应该有所限制,这些问题留作复习题 2.

球台模型 将圆台模型中顶部的小圆台改为球台,可以与下面的圆柱连接得更为光滑,并且符合体积一定下球体表面积最小的原则.设圆柱半径、高度、侧壁和底部厚度仍采用圆柱模型的记号 r, h, b 和 kb,球台上底面(即罐盖)的半径、高度和厚度分别记作 r_1, h_1 和 k_1b,再设球台侧壁的厚度与圆柱相同,由圆柱和圆台组成的易拉罐所耗材料的体积记作 SV_3,罐的容积记作 V_3,则

$$SV_3=2\pi rhb+\pi r^2kb+\pi r_1^2k_1b+\pi b\sqrt{4r^2h_1^2+(r^2-r_1^2-h_1^2)^2}\tag{14}$$

$$V_3=\pi r^2h+\frac{\pi h_1(3r^2+3r_1^2+h_1^2)}{6}\tag{15}$$

其中 b, k 和 k_1 为已知常数.易拉罐尺寸的最优设计是在 V_3 固定的条件下,求 r, h, r_1, h_1 满足什么关系可使 SV_3 最小,这个模型实际上仍然是 3 个独立变量①.

可以像圆台模型那样,对球台上底面的半径 r_1 加以限制,求类似于(13)式的约束极小问题数值解.由于易拉罐顶部只占整个罐体的一小部分,估计求解结果不会有太大变动.

评注 从像易拉罐尺寸这样司空见惯的日常事物中,发现与数学课本类似问题的结果有不同之处,并在亲自实践(测量数据)的基础上,用数学方法给以分析和解决,对于培养数学建模的意识和能力是大有益处的.

本节的建模经过了从圆柱到圆台和球台、从材料厚度相同到不同、从解析解到数值解的过程,对于学习数学建模方法也有启示和指导意义.

实际问题是复杂的,即使像易拉罐的形状和尺寸的优化设计这样看来相对简单的事物,单靠数学手段也不能得到圆满的解决,必须考虑工艺、美观以及使用方便等方面的因素,才能满足人们的需要.

复习题

1. 对于圆柱和圆台形状的易拉罐,在容积一定的情况下要求焊缝最短,建立数学模型确定易拉罐的尺寸.

2. 实际测量易拉罐的尺寸,特别注意上盖、下底、圆柱和圆台侧壁的厚度,并考虑对上盖和圆台侧壁高度的限制,按照圆柱模型重新计算.

第 3 章训练题

1. 一饲养场每天投入 a 元资金用于饲料、设备、人力,估计可使一头重 w_0 kg 的生猪每天增加

① 《工程数学学报》2006 年增刊上郭文飞等的论文中,用球台半径和两个夹角代替 h_1,又增加了约束,这样做虽然等价,但是表达较繁.

b kg.目前生猪出售的市场价格为 c 元/kg,但是预测每天会降低 d 元,建立最佳出售生猪时间的模型.若 $a=4, b=2, c=8, d=0.1, w_0=80$,问应饲养几天后出售.如果上面估计和预测的 b 和 d 有出入,对结果有多大影响?

2. 研究生入学考试科目为数学、外语和专业课三门,王君已经报考,尚有 12 周复习时间.下表是他每门课的复习时间和预计得分,问在下面两种情况下他应该如何分配 12 周的复习时间?预计最高可得多少分?

(1) 只考虑总分最多.

(2) 在三门课都及格的条件下总分最多.

科目	周数										
	0	1	2	3	4	5	6	7	8	9	10
数学	20	40	55	65	72	77	80	82	83	84	85
外语	30	45	53	58	62	65	68	70	72	74	75
专业课	50	70	85	90	93	95	96	96	96	96	96

3. 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.

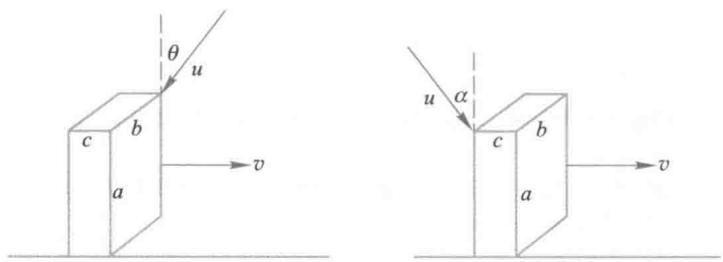
将人体简化成一个长方体,高 $a=1.5$ m(颈部以下),宽 $b=0.5$ m,厚 $c=0.2$ m.设跑步距离 $d=1\,000$ m,跑步最大速度 $v_m=5$ m/s,雨速 $u=4$ m/s,降雨量 $w=2$ cm/h,记跑步速度为 v.按以下步骤进行讨论[17]:

(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,雨线方向与跑步方向在与地面垂直的同一平面内,且与人体的夹角为 θ,如下左图.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 $\theta=0°$, $\theta=30°$ 时的总淋雨量.

(3) 雨从背面吹来,雨线方向与跑步方向在与地面垂直的同一平面内,且与人体的夹角为 α,如下右图.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,以总淋雨量为纵轴,速度 v 为横轴作图(考虑 α 的影响),问速度 v 多大,总淋雨量最少.计算 $\alpha=30°$ 时的总淋雨量,并解释结果的实际意义.

(4) 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?



4. 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,是他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的利润.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 的最优解应满足什么关系?用图解法确定这个最优解[7].

5. 人行走时做的功是抬高人体重心所需势能与两腿运动所需动能之和.试建立模型讨论在做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,且 $s<l$.证明人体重心在行走时升高 $\delta\approx s^2/(8l)$.

(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 $mv^3/(6s)$.

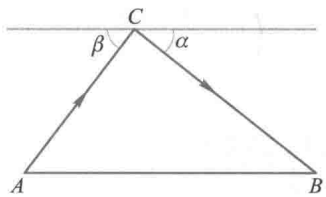
(3) 设人体质量 M,证明在速度 v 一定时每秒行走 $n=\sqrt{\dfrac{3Mg}{4ml}}$ 步做功最小.实际上,$\dfrac{M}{m}\approx 4$,$l\approx 1$ m,分析这个结果合理吗?

(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 $n=\sqrt{\dfrac{Mg}{4ml}}$ 步.分析这个结果是否合理[57].

6. 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(见下图),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)$\dfrac{k\sin\alpha+\sin\beta}{k\sin(\alpha+\beta)}$.



(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[7].

更多案例……

3-1 生产者的决策

3-2 生猪的出售时机

3-3 效用函数的构造及应用

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 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 表示，目标函数 $f(\boldsymbol{x})$ 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）$g_i(\boldsymbol{x}) \leqslant 0 (i=1,2,\cdots,m)$ 来界定，称为约束条件.一般地，这类模型可表述成如下形式

$$\min_{x} z=f(\boldsymbol{x})$$
$$\text{s.t.} \quad g_i(\boldsymbol{x}) \leqslant 0, i=1,2,\cdots,m$$

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



基础知识 4-1
线性规划的基本原理和解法

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 $x_1+x_2\leqslant 50$；

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 $12x_1+8x_2\leqslant 480$；

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 $3x_1\leqslant 100$；

非负约束 x_1，x_2 均不能为负值，即 $x_1\geqslant 0$，$x_2\geqslant 0$.

综上可得

$$\max\ z=72x_1+64x_2 \tag{1}$$

$$\text{s.t.}\quad x_1+x_2\leqslant 50 \tag{2}$$

$$12x_1+8x_2\leqslant 480 \tag{3}$$

$$3x_1\leqslant 100 \tag{4}$$

$$x_1\geqslant 0,x_2\geqslant 0 \tag{5}$$

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，实际上是事先做了如下的假设：

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_1Ox_2 平面上的5条直线，依次记为 L_1~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，$z=2\,400$ 是过 D 点的直线，$z=3\,040$ 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，$z=3\,360$，达到最大值，所以 B 点的坐标(20,30) 即为最优解：$x_1=20$,$x_2=30$.

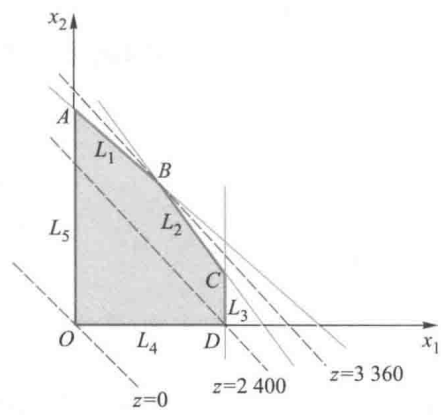


图1　模型的图解法

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.

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：



程序文件 4-1
加工奶制品生产计划
prog0401a.lg4

```
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≤、≥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        20.00000            0.000000
                           X2        30.00000            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 3 360(即最大利润为 3 360 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 Price 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(1 h)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,1 h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.具体做法是: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
                 Current        Allowable      Allowable
Variable         Coefficient    Increase       Decrease
      X1         72.00000       24.00000       8.000000
      X2         64.00000       8.000000       16.00000
                        Righthand Side Ranges
      Row        Current        Allowable      Allowable
                   RHS          Increase       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30×3=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 例 1 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 A_1 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 x_1+x_5 kg A_1,用牛奶 $(x_1+x_5)/3$ 桶,每天生产 x_2+x_6 kg A_2,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 $2x_5$ 和 $2x_6$，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　$x_1,x_2,\cdots,x_6$ 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：

$$\max z=24x_1+16x_2+44x_3+32x_4-3x_5-3x_6 \tag{6}$$

$$\text{s.t.}\quad \frac{x_1+x_5}{3}+\frac{x_2+x_6}{4}\leqslant 50 \tag{7}$$

$$4(x_1+x_5)+2(x_2+x_6)+2x_5+2x_6\leqslant 480 \tag{8}$$

$$x_1+x_5\leqslant 100 \tag{9}$$

$$x_3=0.8x_5 \tag{10}$$

$$x_4=0.75x_6 \tag{11}$$

$$x_1,x_2,x_3,x_4,x_5,x_6\geqslant 0 \tag{12}$$

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

$$4x_1+3x_2+4x_5+3x_6\leqslant 600 \tag{7'}$$

(8)式改写为

$$4x_1+2x_2+6x_5+4x_6\leqslant 480 \tag{8'}$$

输入并求解，可得如下输出：

程序文件 4-2
奶制品生产销售计划
prog0401b.lg4

```
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
                          Current     Allowable     Allowable
```

Variable	Coefficient	Increase	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 $z=3\,460.8$，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 3 460.8 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 $3.16\times12=37.92$ 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 $37.92\times5=189.6$ 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 $z=3\,400$，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 3 400 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

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做.

了每天销售10 kg A_1 的合同并且必须满足,该合同将会使公司利润减少 $1.68\times10=16.8$ 元,即最优利润为 $3\,460.8-16.8=3\,444$ 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3 040,而不再是 $3\,460.8-1.68\times100=3\,292.8$.

复习题

考虑以下"食谱问题"[72]:某学校为学生提供营养套餐,希望以最小费用来满足学生对基本营养的需求.按照营养学家的建议,一个人一天对蛋白质、维生素A和钙的需求如下:50 g蛋白质、4 000 IU(国际单位)维生素A和1 000 mg钙.我们只考虑以下食物构成的食谱:苹果、香蕉、胡萝卜、枣汁和鸡蛋,其营养含量见下表.确定每种食物的用量,以最小费用满足营养学家建议的营养需求,并考虑:

(1) 对维生素A的需求增加一个单位时,是否需要改变食谱?成本增加多少?如果对蛋白质的需求增加1 g呢?如果对钙的需求增加1 mg呢?

(2) 胡萝卜的价格增加1角时,是否需要改变食谱?成本增加多少?

食物	单位	蛋白质/g	维生素A/IU	钙/mg	价格/角
苹果	138 g/个(中等大小)	0.3	73	9.6	10
香蕉	118 g/个(中等大小)	1.2	96	7	15
胡萝卜	72 g/个(中等大小)	0.7	20 253	19	5
枣汁	178 g/杯	3.5	890	57	60
鸡蛋	44 g/个(中等大小)	5.5	279	22	8

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 单位:元/10^3t

	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×(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×(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

$$\min z = 160x_{11}+130x_{12}+220x_{13}+170x_{14}+140x_{21}+130x_{22}+190x_{23}+150x_{24}+190x_{31}+200x_{32}+230x_{33} \tag{1}$$

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

$$x_{11}+x_{12}+x_{13}+x_{14}=50 \tag{2}$$

$$x_{21}+x_{22}+x_{23}+x_{24}=60 \tag{3}$$

$$x_{31}+x_{32}+x_{33}=50 \tag{4}$$

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

$$30\leqslant x_{11}+x_{21}+x_{31}\leqslant 80 \tag{5}$$

$$70\leqslant x_{12}+x_{22}+x_{32}\leqslant 140 \tag{6}$$

$$10\leqslant x_{13}+x_{23}+x_{33}\leqslant 30 \tag{7}$$

$$10\leqslant x_{14}+x_{24}\leqslant 50 \tag{8}$$

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).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.



程序文件 4-3
自来水输送
prog0402a.lg4

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 单位:元

净利润	甲	乙	丙	丁
A	290	320	230	280
B	310	320	260	300
C	260	250	220	/

评注 例 1 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.

于是决策目标为

$$\max z=290x_{11}+320x_{12}+230x_{13}+280x_{14}+310x_{21}+320x_{22}+260x_{23}+300x_{24}+260x_{31}+250x_{32}+220x_{33} \tag{9}$$

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≤,即

$$x_{11}+x_{12}+x_{13}+x_{14}\leqslant 100 \tag{10}$$

$$x_{21}+x_{22}+x_{23}+x_{24}\leqslant 120 \tag{11}$$

$$x_{31}+x_{32}+x_{33}\leqslant 100 \tag{12}$$

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≤可以改写成=.作这样的简化后得到的解没有任何变化.

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/($m^3\cdot t^{-1}$)	利润/(元$\cdot t^{-1}$)
货物 1	18	480	3 100

续表

	质量/t	体积/($m^3 \cdot t^{-1}$)	利润/(元·t^{-1})
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(单位: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 $WET=(10,16,8)$,$VOL=(6\ 800,8\ 700,5\ 300)$;$\boldsymbol{w}=(18,15,23,12)$,$\boldsymbol{v}=(480,650,580,390)$,$\boldsymbol{p}=(3\ 100,3\ 800,3\ 500,2\ 850)$.

决策目标是最大化总利润,即

$$\max z=\sum_{i=1}^{4} p_i\left(\sum_{j=1}^{3} x_{ij}\right) \tag{13}$$

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

$$\sum_{j=1}^{3} x_{ij}\leqslant w_i, i=1,2,3,4 \tag{14}$$

2. 三个货舱的质量限制,即

$$\sum_{i=1}^{4} x_{ij}\leqslant WET_j, j=1,2,3 \tag{15}$$

3. 三个货舱的空间限制,即

$$\sum_{i=1}^{4} v_i x_{ij}\leqslant VOL_j, j=1,2,3 \tag{16}$$

4. 三个货舱装入质量的平衡约束,即

$$\sum_{i=1}^{4} x_{ij}/WET_j=\sum_{i=1}^{4} x_{ik}/WET_k, j,k=1,2,3; j\neq k \tag{17}$$

模型求解　将以上模型输入 LINGO,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唯一,但 LINGO 只能给出一个解.)



程序文件 4-4
货机装运
prog0402b.lg4

评注　初步看来,例 2 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.

复习题

考虑以下“运输问题”[72]:某公司有6个建筑工地要开工,每个工地的位置(用平面坐标 x, y 表示,距离单位:km)及水泥日用量 d(单位:t)由下表给出.目前有两个临时料场位于A(5, 1),B(2, 7),日储量各有20 t.假设从料场到工地之间均有直线道路相连,试制定每天的供应计划,即从A,B两料场分别向各工地运送多少吨水泥,使总的吨千米数最小.

	1	2	3	4	5	6
x/km	1.25	8.75	0.5	5.75	3	7.25
y/km	1.25	0.75	4.75	5	6.5	7.75
d/t	3	5	4	7	6	11

4.3 汽车生产与原油采购

基础知识 4-2
非线性规划的基本原理和解法

在4.1节和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

基础知识 4-3
整数规划的基本原理和解法

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:

$$\max\ z=2x_1+3x_2+4x_3 \tag{1}$$

$$\text{s.t.}\quad 1.5x_1+3x_2+5x_3\leqslant 600 \tag{2}$$

$$280x_1+250x_2+400x_3\leqslant 60\,000 \tag{3}$$

$$x_1,x_2,x_3\geqslant 0 \tag{4}$$

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:

$$x_1, x_2, x_3 \quad \text{均为整数} \tag{5}$$

这样得到的(1)~(5)式称为整数规划(integer programming,简记作 IP),可以用 LINGO 直接求解,输入文件中需要用"@ gin"函数将变量限定为整数.

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.



程序文件 4-5
汽车生产
prog0403a.lg4

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

$$x_1, x_2, x_3 = 0 \text{ 或} \geqslant 80 \tag{6}$$

下面是求解模型(1)~(3),(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:

$$x_1=0, x_2=0, x_3\geqslant 80 \tag{6-1}$$
$$x_1=0, x_2\geqslant 80, x_3=0 \tag{6-2}$$
$$x_1=0, x_2\geqslant 80, x_3\geqslant 80 \tag{6-3}$$
$$x_1\geqslant 80, x_2=0, x_3=0 \tag{6-4}$$
$$x_1\geqslant 80, x_2\geqslant 80, x_3=0 \tag{6-5}$$
$$x_1\geqslant 80, x_2=0, x_3\geqslant 80 \tag{6-6}$$
$$x_1\geqslant 80, x_2\geqslant 80, x_3\geqslant 80 \tag{6-7}$$
$$x_1, x_2, x_3=0 \tag{6-8}$$

(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



程序文件 4-6
汽车生产(加产量限制)0-1 规划
prog0403b.lg4

设 y_1 只取 0, 1 两个值,则"$x_1=0$ 或$\geqslant 80$"等价于

$$80y_1\leqslant x_1\leqslant My_1, y_1\in\{0,1\} \tag{7-1}$$

其中 M 为相当大的正数,本例可取 1 000(x_1 不可能超过 1 000).类似地有

$$80y_2\leqslant x_2\leqslant My_2, y_2\in\{0,1\} \tag{7-2}$$
$$80y_3\leqslant x_3\leqslant My_3, y_3\in\{0,1\} \tag{7-3}$$

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);
```

程序文件 4-7
汽车生产(加产量限制)非线性规划
prog0403c.lg4

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

$$x_1(x_1-80)\geqslant 0 \tag{8-1}$$

$$x_2(x_2-80)\geqslant 0 \tag{8-2}$$

$$x_3(x_3-80)\geqslant 0 \tag{8-3}$$

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(nonlinear programming,简记作 NLP).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?[61]

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):

$$c(x)=\begin{cases}10x, & 0\leqslant x\leqslant 500\\ 1\ 000+8x, & 500\leqslant x\leqslant 1\ 000\\ 3\ 000+6x, & 1\ 000\leqslant x\leqslant 1\ 500\end{cases} \tag{9}$$

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

$$\max z=4.8(x_{11}+x_{21})+5.6(x_{12}+x_{22})-c(x) \tag{10}$$

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

$$x_{11}+x_{12}\leqslant 500+x \tag{11}$$

$$x_{21}+x_{22}\leqslant 1\ 000 \tag{12}$$

$$x\leqslant 1\ 500 \tag{13}$$

$$\frac{x_{11}}{x_{11}+x_{21}}\geqslant 0.5 \tag{14}$$

$$\frac{x_{12}}{x_{12}+x_{22}} \geqslant 0.6 \tag{15}$$

$$x_{11}, x_{12}, x_{21}, x_{22}, x \geqslant 0 \tag{16}$$

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

$$x = x_1 + x_2 + x_3 \tag{17}$$

这时目标函数(10)变为线性函数:

$$\max z = 4.8(x_{11} + x_{21}) + 5.6(x_{12} + x_{22}) - (10x_1 + 8x_2 + 6x_3) \tag{18}$$

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

$$(x_1 - 500)x_2 = 0 \tag{19}$$

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

$$(x_2 - 500)x_3 = 0 \tag{20}$$

此外,x_1, x_2, x_3 的取值范围是

$$0 \leqslant x_1, x_2, x_3 \leqslant 500 \tag{21}$$

由于有非线性约束(19)和(20),(11)~(21)构成非线性规划模型.

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.具体做法是: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.



程序文件 4-8
原油采购与加工
第 1 种解法
prog0403d.lg4

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

$$500y_2 \leqslant x_1 \leqslant 500y_1 \tag{22}$$

$$500y_3 \leqslant x_2 \leqslant 500y_2 \tag{23}$$

$$x_3 \leqslant 500y_3 \tag{24}$$

$$y_1, y_2, y_3 = 0 \text{ 或 } 1 \tag{25}$$

(11)~(18),(21)~(25)构成整数(线性)规划模型,最优解与第 1 种解法得到的结果



程序文件 4-9
原油采购与加工
第 2 种解法
prog0403e.lg4

(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 $x=z_1b_1+z_2b_2, z_1+z_2=1, z_1, z_2\geqslant 0$,因为 $c(x)$ 在 $[b_1, b_2]$ 是线性的,所以 $c(x)=z_1c(b_1)+z_2c(b_2)$. 同样,当 x 在第2个小区间 $[b_2, b_3]$ 时,$x=z_2b_2+z_3b_3, z_2+z_3=1, z_2, z_3\geqslant 0, c(x)=z_2c(b_2)+z_3c(b_3)$. 当 x 在第3个小区间 $[b_3, b_4]$ 时,$x=z_3b_3+z_4b_4, z_3+z_4=1, z_3, z_4\geqslant 0, c(x)=z_3c(b_3)+z_4c(b_4)$.

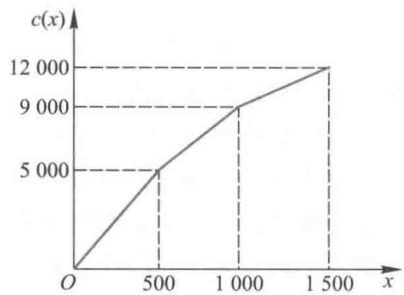


图1 分段线性函数 $c(x)$ 图形

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

$$z_1\leqslant y_1, z_2\leqslant y_1+y_2, z_3\leqslant y_2+y_3, z_4\leqslant y_3 \tag{26}$$

$$z_1+z_2+z_3+z_4=1, z_k\geqslant 0\ (k=1,2,3,4) \tag{27}$$

$$y_1+y_2+y_3=1, y_1, y_2, y_3=0\text{ 或 }1 \tag{28}$$

此时 x 和 $c(x)$ 可以统一地表示为

$$x=z_1b_1+z_2b_2+z_3b_3+z_4b_4=500z_2+1\ 000z_3+1\ 500z_4 \tag{29}$$

$$\begin{aligned} c(x)&=z_1c(b_1)+z_2c(b_2)+z_3c(b_3)+z_4c(b_4)\\ &=5\ 000z_2+9\ 000z_3+12\ 000z_4 \end{aligned} \tag{30}$$

程序文件 4-10
原油采购与加工第3种解法
prog0403f.lg4

(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.

设一个 n 段线性函数 $f(x)$ 的分点为 $b_1\leqslant\cdots\leqslant b_n\leqslant b_{n+1}$,引入 z_k,将 x 和 $f(x)$ 表示为

$$x=\sum_{k=1}^{n+1}z_kb_k \tag{31}$$

$$f(x)=\sum_{k=1}^{n+1}z_kf(b_k) \tag{32}$$

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.

z_k 和0-1变量 y_k 满足

$$z_1\leqslant y_1, z_2\leqslant y_1+y_2, \cdots, z_n\leqslant y_{n-1}+y_n, z_{n+1}\leqslant y_n \tag{33}$$

$$y_1+y_2+\cdots+y_n=1, y_k=0\text{ 或 }1 \tag{34}$$

$$z_1+z_2+\cdots+z_{n+1}=1, z_k\geqslant 0\ (k=1,2,\cdots,n+1) \tag{35}$$

复习题

1. 考虑4.2节的复习题,为了进一步减少吨千米数,打算舍弃两个临时料场,改建两个新的,日储量仍各为20 t,应建在何处?节省的吨千米数有多大?[72]

2. 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(单位:千元/t)分别为6,16,10,15;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.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 50 t；产品 A，B 的市场需求量分别为 100 t，200 t. 问应如何安排生产[21]？

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做出抉择，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

续表

	c_{ij}				
	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
$j=3$	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$\sum_{j=1}^{4} x_{ij} \leqslant 1$;

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$\sum_{i=1}^{5} x_{ij}=1$.



程序文件 4-11
接力队选拔
prog0404a.lg4

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

$$\min z=\sum_{j=1}^{4}\sum_{i=1}^{5} c_{ij}x_{ij} \tag{1}$$

$$\text{s.t.}\sum_{j=1}^{4} x_{ij}\leqslant 1, i=1,2,3,4,5 \tag{2}$$

$$\sum_{i=1}^{5} x_{ij}=1, j=1,2,3,4 \tag{3}$$

$$x_{ij}=\{0,1\} \tag{4}$$

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 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指派问题,但能建立类似的模型.

例1 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.

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;$i=1,2,\cdots,9$).问题的目标为选修的课程总数最少,即

$$\min z=\sum_{i=1}^{9} x_i \tag{5}$$

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

$$x_1+x_2+x_3+x_4+x_5\geqslant 2 \tag{6}$$

$$x_3+x_5+x_6+x_8+x_9\geqslant 3 \tag{7}$$

$$x_4+x_6+x_7+x_9\geqslant 2 \tag{8}$$

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 $x_4\leqslant x_7$(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 $x_3\leqslant x_1$,$x_3\leqslant x_2$,而这两个不等式可以用一个约束表示为 $2x_3-x_1-x_2\leqslant 0$.这样,所有课程的先修课要求可表为如下的约束:

$$2x_3-x_1-x_2\leqslant 0 \tag{9}$$

$$x_4-x_7\leqslant 0 \tag{10}$$

$$2x_5-x_1-x_2\leqslant 0 \tag{11}$$

$$x_6-x_7\leqslant 0 \tag{12}$$

$$x_8-x_5\leqslant 0 \tag{13}$$

$$2x_9-x_1-x_2\leqslant 0 \tag{14}$$



程序文件 4-12
选课策略
prog0404b.lg4

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,这个解并不是唯一的,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.



基础知识 4-4
多目标规划和目标规划的基本原理与解法

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

$$\max\ w=5x_1+4x_2+4x_3+3x_4+4x_5+3x_6+2x_7+2x_8+3x_9 \tag{15}$$

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:

$$\text{V-min}(z,-w) \tag{16}$$

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.



程序文件 4-13
两目标选课策略
prog0404c.lg4

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.如果这个解是唯一的,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唯一,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

$$\sum_{j=1}^{9} x_i=6 \tag{17}$$

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.求解后发现会得到不同于前面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唯一的,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

$$\begin{aligned}\min\ y &= 0.7z-0.3w\\ &= -0.8x_1-0.5x_2-0.5x_3-0.2x_4-0.5x_5-0.2x_6+0.1x_7+0.1x_8-0.2x_9\end{aligned} \tag{18}$$

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 $\lambda_1+\lambda_2=1$,

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 $x_4\leqslant x_7$ 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

$0 \leqslant \lambda_1, \lambda_2 \leqslant 1$，则 0-1 规划模型的新目标为

$$\min\ y = \lambda_1 z - \lambda_2 w \tag{19}$$

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 x 的所有系数都小于 0，因此为了使 y 取最小值，所有决策变量应尽可能取 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 x 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.

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(多目标规划的更多例子参见本书 4.7 节和 4.8 节).

复习题

某公司指派 n 个员工到 n 个城市工作(每个城市单独一人)，希望使所花费的总电话费用尽可能少.n 个员工两两之间每个月通话的时间表示在下面的矩阵的上三角形部分(假设通话的时间矩阵是对称的，没有必要写出下三角形部分)，n 个城市两两之间通话费率表示在下面的矩阵的下三角形部分(同样道理，假设通话的费率矩阵是对称的，没有必要写出上三角形部分).试求解该二次指派问题.(如果你的软件解不了这么大规模的问题，那就只考虑最前面的若干员工和城市.)[21]

$$\begin{pmatrix}
0 & 5 & 3 & 7 & 9 & 3 & 9 & 2 & 9 & 0 \\
7 & 0 & 7 & 8 & 3 & 2 & 3 & 3 & 5 & 7 \\
4 & 8 & 0 & 9 & 3 & 5 & 3 & 3 & 9 & 3 \\
6 & 2 & 10 & 0 & 8 & 4 & 1 & 8 & 0 & 4 \\
8 & 6 & 4 & 6 & 0 & 8 & 8 & 7 & 5 & 9 \\
8 & 5 & 4 & 6 & 6 & 0 & 4 & 8 & 0 & 3 \\
8 & 6 & 7 & 9 & 4 & 3 & 0 & 7 & 9 & 5 \\
6 & 8 & 2 & 3 & 8 & 8 & 6 & 0 & 5 & 5 \\
6 & 3 & 6 & 2 & 8 & 3 & 7 & 8 & 0 & 5 \\
5 & 6 & 7 & 6 & 6 & 2 & 8 & 8 & 9 & 0
\end{pmatrix}$$

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

$$\min z=5.0x_1+5.1x_2+5.4x_3+5.5x_4+0.2(y_1+y_2+y_3) \tag{1}$$

$$\text{s.t.}\quad x_1-y_1=15 \tag{2-1}$$

$$x_2+y_1-y_2=25 \tag{2-2}$$

$$x_3+y_2-y_3=35 \tag{2-3}$$

$$x_4+y_3=25 \tag{2-4}$$

$$x_1\leqslant 30,x_2\leqslant 40,x_3\leqslant 45,x_4\leqslant 20 \tag{3}$$

$$x_1,x_2,x_3,x_4,y_1,y_2,y_3\geqslant 0 \tag{4}$$



程序文件 4-14
饮料厂生产计划
prog0405a.lg4

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?



程序文件 4-15
饮料厂生产与检修计划
prog0405b.lg4

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

$$x_1+15w_1\leqslant 30 \tag{3-1}$$

$$x_2+15w_2-5w_1\leqslant 40 \tag{3-2}$$

$$x_3+15w_3-5w_2-5w_1\leqslant 45 \tag{3-3}$$

$$x_4+15w_4-5w_1-5w_2-5w_3\leqslant 20 \tag{3-4}$$

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

$$w_1+w_2+w_3+w_4=1 \tag{5}$$

$$w_1,w_2,w_3,w_4\in\{0,1\} \tag{6}$$

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.

用LINGO对修改的模型直接求解,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.

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条件下,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 $M_t(t=1,2,\cdots,T)$.如果第 t 时段开工生产,则需付出生产准备费为 $s_t\geqslant 0$,单件产品的生产成本为 $c_t\geqslant 0$.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 $h_t\geqslant 0$.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 $x_t(\geqslant 0)$,t 时段末产品的库存(即下一个时段的初始库存)为 $y_t(\geqslant 0)$,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的数学规划模型

描述：

$$\min z = \sum_{t=1}^{T} (s_t w_t + c_t x_t + h_t y_t) \tag{7}$$

$$\text{s.t.} \quad y_{t-1} + x_t - y_t = d_t, \quad t = 1,2,\cdots,T \tag{8}$$

$$w_t = \begin{cases} 1, x_t > 0, \\ 0, x_t = 0, \end{cases} \quad t = 1,2,\cdots,T \tag{9}$$

$$x_t \leqslant M_t, t = 1,2,\cdots,T \tag{10}$$

$$y_0 = y_T = 0, \tag{11}$$

$$x_t, y_t \geqslant 0, \quad t = 1,2,\cdots,T \tag{12}$$

程序文件 4-16
饮料厂生产批量计划
prog0405c.lg4

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**.

(9)式不便于计算,可以将(9),(10)合并地表为 $x_t \leqslant M_t w_t$,或用下面的线性不等式代替

$$x_t - M_t w_t \leqslant 0, \quad t = 1,2,\cdots,T \tag{13}$$

将本题所给参数代入这一模型,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.

如果把对每周生产能力的限制去掉,则将约束(13)中的 M_t 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制的结果相比,这里只有第1,2周生产,生产准备费会更加节省.

复习题

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下表:

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1 h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4 h,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?

4.6 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[21]?

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($i=1,2,\cdots,7$)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

$$\min z_1 = 3x_1 + x_2 + 3x_3 + 3x_4 + x_5 + x_6 + 3x_7 \tag{1}$$

以切割原料钢管的总根数最少为目标,则有

$$\min z_2 = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 \tag{2}$$

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

$$4x_1+3x_2+2x_3+x_4+x_5\geqslant 50 \tag{3}$$

$$x_2+2x_4+x_5+3x_6\geqslant 20 \tag{4}$$

$$x_3+x_5+2x_7\geqslant 15 \tag{5}$$



程序文件 4-17
钢管下料问题1
prog0406a.lg4

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

$$\min z=x_1+x_2+x_3 \tag{6}$$

约束条件 为满足客户的需求,应有

$$r_{11}x_1+r_{12}x_2+r_{13}x_3\geqslant 50 \tag{7}$$

$$r_{21}x_1+r_{22}x_2+r_{23}x_3\geqslant 10 \tag{8}$$

$$r_{31}x_1+r_{32}x_2+r_{33}x_3\geqslant 20 \tag{9}$$

$$r_{41}x_1+r_{42}x_2+r_{43}x_3\geqslant 15 \tag{10}$$

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

$$16\leqslant 4r_{11}+5r_{21}+6r_{31}+8r_{41}\leqslant 19 \tag{11}$$

$$16\leqslant 4r_{12}+5r_{22}+6r_{32}+8r_{42}\leqslant 19 \tag{12}$$

$$16 \leqslant 4r_{13}+5r_{23}+6r_{33}+8r_{43} \leqslant 19 \tag{13}$$

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

$$x_1 \geqslant x_2 \geqslant x_3 \tag{14}$$

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 $\left[\dfrac{4\times50+5\times10+6\times20+8\times15}{19}\right]+1=26$ 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

$$26 \leqslant x_1+x_2+x_3 \leqslant 31 \tag{15}$$

将(6)~(15)构成的模型输入 LINGO 求解得到: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唯一,你能找出一个与此不同的解吗?



程序文件 4-18
钢管下料问题 2
prog0406b.lg4

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/cm^2(如果周末有罐身、上盖或下底不能配套组装成易拉罐出售,也看作是原料余料损失).

问工厂应如何安排每周的生产?

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 $d=5$ cm,可得其面积为 $s=\pi d^2/4\approx19.6$ cm^2,周长为 $L=\pi d\approx15.7$ cm;已知罐身高 $h=10$ cm,可得其面积为 $S=hL\approx157.1$ cm^2.于是模式 1 下的余料损失为 24^2 $cm^2-10s-S\approx222.6$ 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.

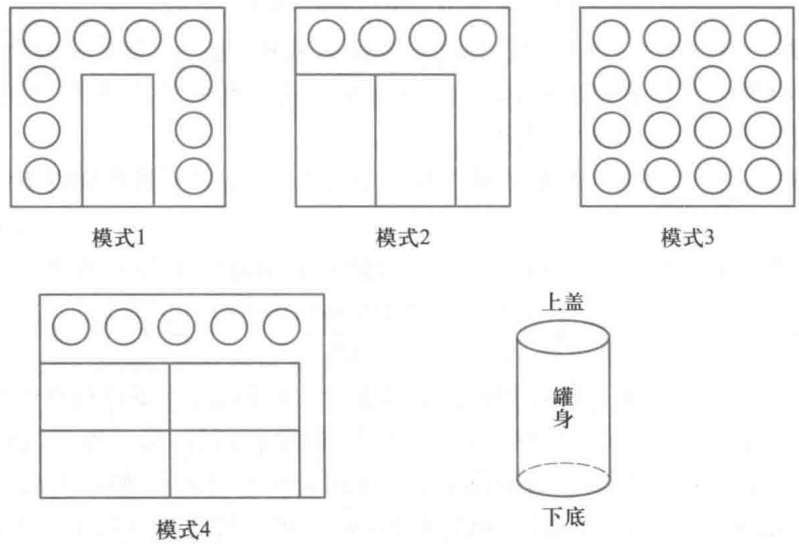


图1 易拉罐下料模式

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

$$\max z = 0.1y_1-0.001(222.6x_1+183.3x_2+261.8x_3+169.5x_4+157.1y_2+19.6y_3) \quad (16)$$



程序文件 4-19
易拉罐下料
prog0406c.lg4

约束条件

时间约束:每周工作时间不超过 40 h = 144 000 s,由表2最后一列得

$$1.5x_1+2x_2+x_3+3x_4 \leqslant 144\ 000 \quad (17)$$

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

$$x_1+x_2+x_3\leqslant 50\ 000 \tag{18}$$

$$x_4\leqslant 20\ 000 \tag{19}$$

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

$$y_1=\min\{x_1+2x_2+4x_4,(10x_1+4x_2+16x_3+5x_4)/2\} \tag{20}$$

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

$$y_2=x_1+2x_2+4x_4-y_1 \tag{21}$$

$$y_3=10x_1+4x_2+16x_3+5x_4-2y_1 \tag{22}$$

(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，但是它可以等价为以下两个线性不等式

$$y_1\leqslant x_1+2x_2+4x_4 \tag{23}$$

$$y_1\leqslant(10x_1+4x_2+16x_3+5x_4)/2 \tag{24}$$

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

$$1.5x_1+2x_2+x_3+3x_4\leqslant 14.4 \tag{25}$$

$$x_1+x_2+x_3\leqslant 5 \tag{26}$$

$$x_4\leqslant 2 \tag{27}$$

将模型(16)和(21)~(27)输入LINGO求解得到：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由两部分组成，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.

复习题

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长度都是1 850 mm.现有一客户需要15根290 mm、28根315 mm、21根350 mm和30根455 mm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100 mm.为了使总费用最小，应如何下料[21]？

4.7 广告投入与升级调薪

从4.4节中的例2可以看出，很多实际决策问题往往需要考虑多个目标，本节例1

再给出一个多目标规划的例子.在本节的例2中,决策者不仅要考虑多个目标,而且这些目标以“软约束”的形式出现且有优先顺序,这时可以建立目标规划模型.

例1 广告投入

问题 某公司计划在某地通过5种媒体(手机、网络、电视、报纸、电台)对公司产品做广告,并将目标人群分为7类(如按年龄分为小孩、老人、青年人、中年人,某些年龄段的人再按性别细分,等等).根据过去的经验以及当前的市场预测,每花费1万元广告费,能够新吸引到的目标人群的数量(万人)如表1的2~6行所示(空格处为0),反映不同媒体对不同人群的吸引能力,表1最后两行还分别列出了公司希望吸引到的目标人群的最小数量,和可能吸引到的目标人群的最大数量.问公司应该在5种媒体上分别花费多少广告费?如果公司还希望吸引到尽可能多的目标人群,并得到花费的广告费与所能吸引的目标人群之间的数量关系,应该怎么做呢?[59]

表1 广告与目标人群的信息

	人群1	人群2	人群3	人群4	人群5	人群6	人群7
手机		10	4	50	5		2
网络		10	30	5	12		
电视	20					5	3
报纸	8					6	10
电台		6	5	10	11	4	
最小数量	25	40	60	120	40	11	15
最大数量	60	70	120	140	80	25	55

模型建立 公司当然希望花费尽量少的广告费而吸引到尽可能多的目标人群,这是两个互相冲突的目标.这个问题仅给出了希望吸引到和可能吸引到的目标人群的最小数量和最大数量,并没有给出公司所能花费的广告费的上下限,我们首先在满足吸引到的目标人群的最低要求的情况下,建立使广告费最少的模型.

将5种媒体依次记为$i=1,2,\cdots,5$,7类目标人群依次记为$j=1,2,\cdots,7$,表1中2~6行的目标人群数量记为a_{ij},最后2行的数据分别记为l_j和u_j.

用x_i表示投入第i种媒体的广告费(万元),则总的广告费投入为

$$z_1=x_1+x_2+x_3+x_4+x_5 \tag{1}$$

记通过广告吸引到的7类目标人群的数量分别为y_j,$j=1,2,\cdots,7$,则由表1第2列数据可知

$$y_1=\min(20x_3+8x_4,60) \tag{2}$$

类似地,可得$y_j(j=2,3,\cdots,7)$的表达式.一般可表为

$$y_j=\min\left(\sum_{i=1}^{5}a_{ij}x_i,u_j\right),\quad j=1,2,\cdots,7 \tag{3}$$

由于吸引到的目标人群最小数量为l_j(最大数量限制u_j已列在(3)中),所以

$$y_j\geqslant l_j,\quad j=1,2,\cdots,7 \tag{4}$$

还有对广告费的非负约束

$$x_i \geqslant 0,\ i=1,2,\cdots,5 \tag{5}$$

因此,如果仅考虑广告费最少,问题归结为在约束(3)~(5)下求(1)的最小值,是一个单目标的线性规划模型.按照这个模型,虽然花费的广告费最少,但可以预见,吸引到的人群在满足最低要求的条件下也会比较少.

现在考虑公司的第二个目标——吸引到尽可能多的目标人群,目标函数为

$$z_2 = y_1 + \cdots + y_7 \tag{6}$$

如果直接在约束(3)~(5)下求(6)的最大值,也是一个单目标的线性规划模型,但显然最优解是将所有潜在的目标人群(最大数量 u_j)全部吸引过来,这时花费的总广告费相应地也会很高.

综合考虑两个目标(1)和(6),我们可以把这个问题表示为一个多目标规划模型:

$$\begin{aligned} &\min \quad z=(z_1,-z_2) \\ &\text{s.t.} \quad (3)\sim(5) \end{aligned} \tag{7}$$

注意到 z_1 希望最小,而 z_2 希望最大,所以(7)中对 z_2 取了负号.利用求解多目标规划的方法,可以得到花费的广告费与所能吸引的目标人群的数量关系.

模型求解 如果公司希望在吸引到最小数量目标人群的条件下使广告费最少,可直接在约束(3)~(5)下求(1)的最小值,得到最少广告费为 6.327 万元,吸引到 328.7 万人;如果希望将所有潜在的目标人群全部吸引过来,可再增加约束 $y_j=u_j(j=1,2,\cdots,7)$,求(1)的最小值,可得到最少广告费为 13.798 万元,所有 550 万人都被吸引了.



程序文件 4-20
广告投入
prog0407a.lg4

下面介绍求解多目标规划问题(7)的一种方法:分别以广告费不超过 6.5,7.0,…,14(万元)作为约束,计算对应的所能吸引到的目标人群的最大数量.注意到目标人群人数较多,这里将其作为连续变量处理.

为此,将(1)式的目标改为约束(以广告费不超过 6.5 万元为例,其他类似)

$$x_1+x_2+x_3+x_4+x_5 \leqslant 6.5 \tag{8}$$

则可以在约束(3)~(5)和(8)下求(6)的最大值,同时得到最优解 x_i.

LINGO 软件(10.0 以上版本)提供了子模型功能,可以方便地处理诸如(8)这类约束(右端项取不同数值进行计算的情形).

编程求解可得到结果如表 2 所示.

表 2 广告费与吸引到的总人数的关系

广告费/万元	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
总人数/万人	343.8	373.1	400.8	428.5	447.5	462.4	475.0	486.2
广告费/万元	10.5	11.0	11.5	12.0	12.5	13.0	13.5	14.0
总人数/万人	497.0	507.8	518.3	525.7	533.0	540.0	546.2	550.0

将这个结果画图可得到有效前沿如图 1 所示.

评注 例 1 中决策者希望得到广告费和广告吸引到的目标人群之间的数量关系,即由非劣解构成的有效前沿,采用(8)式将广告费目标作为约束处理是很自然和方便的.当然,也可以采用与 4.4 节例 2 中类似的加权法,通过设置多个不同的权系数求解得到结果.需要指出的是,对于多个目标量纲不同的情形,加权后的新目标函数往往并没有什么物理意义.

例 2 升级调薪

问题 某单位的员工薪资分成 3 个等级(Ⅰ、Ⅱ、Ⅲ),分别为年薪 20、15、10 万元;上级核定的编制人数分别为 12、15、15 人,而实际现有 36 名员工,其中Ⅰ、Ⅱ、Ⅲ级的人数分别为 9、12、15 人.该单位计划最近进行升级调薪,按规定员工不得越级提升.此外,

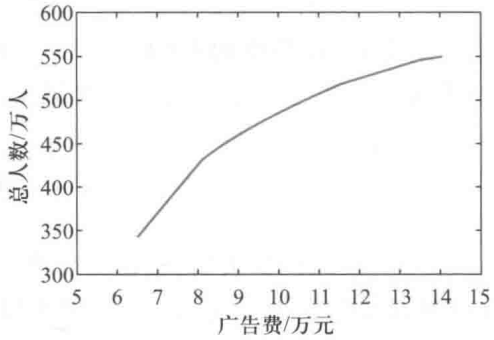


图1 例1的有效前沿

需要依次考虑以下要求：

1）该单位年工资总额尽可能不超过600万元.

2）每级的人数尽可能不超过定编规定的人数.

3）Ⅱ、Ⅲ级的升级面尽可能达到现有相应等级人数的20%.

4）Ⅲ级不足编制的人数可录用新员工.

请为该单位拟定一个满意的升级调薪方案.[84]

模型建立 记 x_1,x_2,x_3 分别为Ⅱ、Ⅲ级提升到Ⅰ、Ⅱ级和新录用到Ⅲ级的员工数量，是这个问题的决策变量（非负整数）.

首先考虑要求1）：升级调薪后，该单位年工资总额为 $20(9+x_1)+15(12-x_1+x_2)+10(15-x_2+x_3)=510+5x_1+5x_2+10x_3$，这个数额应该尽可能不超过600万元.怎么描述这个“尽可能”的约束呢？可以采用目标规划的思想建模，把600万元理解为工资总额的目标值，并引入非负变量 d_1^+ 和 d_1^-，其中 d_1^+ 表示决策值超过目标值的部分，d_1^- 表示决策值低于目标值的部分，则

$$510+5x_1+5x_2+10x_3-d_1^++d_1^-=600 \tag{9}$$

目标规划中将 d_1^+ 和 d_1^- 分别称为（工资总额的）正、负偏差变量（均为非负变量）.由于决策值不可能既超过又低于目标值，所以 d_1^+ 和 d_1^- 至少有一个为0.要求1）希望薪资总额尽可能不超过600万元，也就是希望超过600万元的部分 d_1^+ 尽可能小.

类似地考虑要求2）：记升级调薪后与Ⅰ、Ⅱ、Ⅲ级员工定编数对应的正、负偏差变量分别为 d_2^+ 和 d_2^-，d_3^+ 和 d_3^-，d_4^+ 和 d_4^-，则

$$9+x_1-d_2^++d_2^-=12 \tag{10}$$

$$12-x_1+x_2-d_3^++d_3^-=15 \tag{11}$$

$$15-x_2+x_3-d_4^++d_4^-=15 \tag{12}$$

要求2）希望尽可能不超编，也就是总的超编人数 $d_2^++d_3^++d_4^+$ 尽可能小（这里我们不再对超编属于哪一个级别进行区分）.

对于要求3），记Ⅱ、Ⅲ级员工升级人数对应的正、负偏差变量分别为 d_5^+ 和 d_5^-，d_6^+ 和 d_6^-，则

$$x_1-d_5^++d_5^-=12\times0.2 \tag{13}$$

$$x_2-d_6^++d_6^-=15\times0.2 \tag{14}$$

要求3)希望升级面尽可能达到现有相应等级人数的20%,也就是希望低于20%升级面的人数 $d_5^-+d_6^-$ 尽可能小.

要求4)只是说Ⅲ级不足编制的人数可录用新员工,而没有说录用的新员工越多越好还是越少越好.我们假设在满足前面三个要求的前提下尽可能多地录用新员工,即希望 x_3 尽可能大.

该单位需要"依次"考虑上述4个要求,其含义应该是重要性按照要求1)到要求4)排序.记 P_1,P_2,P_3,P_4 分别为要求1),2),3),4)的优先级,则可以认为 $P_1 \gg P_2 \gg P_3 \gg P_4 > 0$.此时,可将该单位最后的决策目标表达成

$$\min z = P_1 d_1^+ + P_2(d_2^+ + d_3^+ + d_4^+) + P_3(d_5^- + d_6^-) - P_4 x_3 \tag{15}$$

自然,需要加上升级人数的整数约束以及所有变量的非负约束,即

$$x_1, x_2, x_3 \geqslant 0, \text{且为整数}; \ d_i^-, d_i^+ \geqslant 0, \ i = 1,2,3,4,5,6 \tag{16}$$

于是问题归结为优化模型(9)~(16),这是一个目标规划.



程序文件 4-21
升级调薪
prog0407b.lg4

读者可能会问:d_i^-, d_i^+ 至少有一个为0这个条件为什么不写在约束中?仔细分析上述与 d_i^-, d_i^+ 相关的约束会发现,当它们同时增加或减少时也是满足上述约束的,因此只要最小化其中一个偏差变量,另一个自然会是0,所以没有必要再增加 d_i^-, d_i^+ 至少有一个为0这个条件.

模型求解 按照(15),决策总目标是使 $z = P_1 d_1^+ + P_2(d_2^+ + d_3^+ + d_4^+) + P_3(d_5^- + d_6^-) - P_4 x_3$ 最小,而各目标的权重 $P_1 \gg P_2 \gg P_3 \gg P_4 > 0$,所以首先以 d_1^+ 最小为目标,(9)~(14)和(16)为约束求解.可得 d_1^+ 最小值为0,即可以保证达到第一目标——工资总额不超过600万元.

将 $d_1^+ = 0$ 加入约束,以 $d_2^+ + d_3^+ + d_4^+$ 最小为目标,可求得其最小值也为0,可以同时保证达到第二目标——各级人员人数不超编.

将 $d_2^+ + d_3^+ + d_4^+ = 0$ 也加入约束,以 $d_5^- + d_6^-$ 最小为目标,可求得其最小值也为0,可以同时保证达到第三目标——各级提职人员人数不少于20%.

最后,将 $d_5^- + d_6^- = 0$ 也加入约束,以 x_3 最大为目标,可求得其最大值为5,对应的最优解为 $x_1 = 3, x_2 = 5, x_3 = 5$(其他略).即从Ⅱ级提升3人至Ⅰ级,从Ⅲ级提升5人至Ⅱ级,新录用Ⅲ级人员为5名.

评注 例2中决策者希望达到的目标不是强制性的"硬约束"(刚性约束),而是以"尽可能大于(小于)"、"尽可能接近"等"软约束"(柔性约束)的形式出现,不同目标的优先顺序已经明确给定,通过引入偏差变量建立目标规划模型是处理这类问题的合适方法,求解时可以按照目标的优先顺序依次进行求解.

复习题

某公司3个工厂生产的产品供应给4个客户,各工厂生产量、各用户需求量及从各工厂到各用户的单位产品的运输费用由下表给出.

用户	1	2	3	4	生产量
工厂1	5	2	6	7	300
工厂2	3	5	4	6	200
工厂3	4	5	2	3	400
需求量	200	100	450	250	

(1) 如果只考虑总运费最小,制定所有产品的最优运输方案.

(2) 由于总生产量小于总需求量,公司经研究确定了调配方案的8项目标并规定了重要性的次序:

第一目标:用户4为重要客户,需求量必须全部满足;

第二目标:供应用户1的产品中,工厂3的产品不少于100个单位;

第三目标:每个用户的满足率不低于80%;

第四目标:应尽量满足各用户的需求;

第五目标:总运费不超过(1)中运输方案的10%;

第六目标:工厂2到用户4的路线应尽量避免运输任务;

第七目标:用户1和用户3的需求满足率应尽量保持平衡;

第八目标:力求减少总运费.

请列出相应的目标规划模型并进行求解.[79]

4.8 投资的风险与收益

随着生产力的发展和生活水平的提高、经济条件的改善,个人和社会的财富积累大幅增长,投资需求应运而生.投资收益往往具有很大的不确定性,投资者需要具备很强的控制风险意识.在投资的风险和收益之间寻求平衡,本质上是一个多目标规划问题.本节给出投资方面的两个例子.

例1 投资组合

问题 某3种股票A, B, C从1943年到1954年价格每年的增长(包括分红在内)如表1所示.例如,表中第一个数据1.300的含义是股票A在1943年的年末价值是其年初价值的1.300倍,即年收益率为30%,其余数据的含义依此类推.假设你在1955年有一笔资金准备投资这3种股票,希望年收益率至少达到15%,问应当如何投资?并考虑以下问题:

1) 假设除了上述3种股票外,投资人还有一种无风险的投资方式,如购买国库券.假设国库券的年收益率为5%,应如何调整投资计划?

2) 在1)的条件下,如果希望年收益率至少达到10%而不是15%,应如何调整投资计划?[79]

表1 股票收益数据

数据文件4-1 投资组合股票收益

年份	股票A	股票B	股票C
1943	1.300	1.225	1.149
1944	1.103	1.290	1.260
1945	1.216	1.216	1.419
1946	0.954	0.728	0.922
1947	0.929	1.144	1.169
1948	1.056	1.107	0.965
1949	1.038	1.321	1.133

续表

年份	股票 A	股票 B	股票 C
1950	1.089	1.305	1.732
1951	1.090	1.195	1.021
1952	1.083	1.390	1.131
1953	1.035	0.928	1.006
1954	1.176	1.715	1.908

问题分析　上面提出的问题称为投资组合(portfolio),早在 1952 年,现代投资组合理论的开创者 Markowitz 就给出了这个模型的基本框架,并于 1990 年获得了诺贝尔经济学奖.

一般来说,人们投资股票的收益是不确定的,可视为一个随机变量,其大小很自然地可以用年收益率的期望来度量.收益的不确定性当然会带来风险,风险怎样度量呢?Markowitz 建议,风险可以用收益的方差(或标准差)来衡量:方差越大,风险越大;方差越小,风险越小.在一定的假设下,用收益的方差来衡量风险确实是合适的.为此,我们可用表 1 给出的 1955 年以前 12 年的数据,计算出 3 种股票收益的均值和方差(包括协方差).

于是,一种股票收益的均值衡量的是这种股票的平均收益状况,而收益的方差衡量的是这种股票收益的波动幅度,方差越大则波动越大,收益越不稳定.两种股票收益的协方差表示的则是它们之间的相关程度:

- 协方差为 0 时两者不相关.
- 协方差为正表示两者正相关,协方差越大则正相关性越强(越有可能一赚皆赚,一赔俱赔).
- 协方差为负表示两者负相关,绝对值越大则负相关性越强(越有可能一个赚,另一个赔).

记股票 A, B, C 每年的收益率分别为 R_1, R_2 和 R_3(注意表中的数据减去 1 以后才是年收益率),则 $R_i(i=1,2,3)$是一个随机变量.用 E 和 D 分别表示随机变量的期望和方差,用 cov 表示两个随机变量的协方差(covariance),根据概率论的知识和表 1 给出的数据,可以计算出股票 A, B, C 年收益率的期望为

$$ER_1=0.089\ 1,\quad ER_2=0.213\ 7,\quad ER_3=0.234\ 6 \tag{1}$$

年收益率的协方差矩阵为

$$COV=\begin{pmatrix}0.010\ 8 & 0.012\ 4 & 0.013\ 1\\ 0.012\ 4 & 0.058\ 4 & 0.055\ 4\\ 0.013\ 1 & 0.055\ 4 & 0.094\ 2\end{pmatrix} \tag{2}$$

即 $DR_1=\mathrm{cov}(R_1,R_1)=0.010\ 8$, $DR_2=\mathrm{cov}(R_2,R_2)=0.058\ 4$, $DR_3=\mathrm{cov}(R_3,R_3)=0.094\ 2$, $\mathrm{cov}(R_1,R_2)=0.012\ 4$, $\mathrm{cov}(R_1,R_3)=0.013\ 1$, $\mathrm{cov}(R_2,R_3)=0.055\ 4$.

模型建立　用决策变量 x_1, x_2, x_3 分别表示投资人投资股票 A, B, C 的比例.假设市场上没有其他投资渠道,且手上资金(不妨假设只有 1 个单位的资金)必须全部用于

投资这3种股票,则

$$x_1,x_2,x_3\geqslant 0,\quad x_1+x_2+x_3=1 \tag{3}$$

年投资收益率 $R=x_1R_1+x_2R_2+x_3R_3$ 也是一个随机变量.根据概率论的知识,投资的年期望收益率为

$$ER=x_1ER_1+x_2ER_2+x_3ER_3 \tag{4}$$

年投资收益率的方差为

$$\begin{aligned}V&=D(x_1R_1+x_2R_2+x_3R_3)\\&=D(x_1R_1)+D(x_2R_2)+D(x_3R_3)+\\&\quad 2\mathrm{cov}(x_1R_1,x_2R_2)+2\mathrm{cov}(x_1R_1,x_3R_3)+2\mathrm{cov}(x_2R_2,x_3R_3)\\&=x_1^2DR_1+x_2^2DR_2+x_3^2DR_3+\\&\quad 2x_1x_2\mathrm{cov}(R_1,R_2)+2x_1x_3\mathrm{cov}(R_1,R_3)+2x_2x_3\mathrm{cov}(R_2,R_3)\\&=\sum_{j=1}^{3}\sum_{i=1}^{3}x_ix_j\mathrm{cov}(R_i,R_j)\end{aligned} \tag{5}$$

实际的投资者可能面临许多约束条件,这里只考虑题中要求的年收益率的期望不低于15%,即

$$x_1ER_1+x_2ER_2+x_3ER_3\geqslant 0.15 \tag{6}$$

最后的优化模型是在约束(3)和(6)下使(5)最小.这里目标函数 V 是决策变量的二次函数,而约束都是线性函数,称为**二次规划**.这是一类特殊的非线性规划.



程序文件 4-22
投资组合股票收益
prog0408a.lg4

模型求解 计算结果表明,投资3种股票的比例大致是:A占53.01%,B占35.64%,C占11.35%.风险(年收益率的方差)为0.022 4,即收益率的标准差为0.149 7.

下面分析本题需要进一步研究的两个问题.对于问题1),由于无风险的投资方式(如国库券、银行存款等)是有风险的投资方式(如股票)的一种特例,所以模型(3)~(6)仍然是适用的.只不过无风险的投资方式的收益是固定的(5%),所以方差以及与其他投资方式收益的协方差都是0.

相应地可以得到,在年收益率仍然达到15%的条件下,投资A约占8.69%,B占42.85%,C占14.34%,国库券占34.12%,风险(方差)为0.020 8.与上面的结果比较,无风险资产的存在可以使得投资风险减小.虽然国库券的收益率只有5%,比希望得到的总投资的收益率15%小很多,但只要国库券的投资占到34.12%,同时将投资A的比例从53.01%降到8.69%(B,C变化不大),就可以减少风险.

对于问题2),只需把模型中的期望收益减少到10%,结果表明投资A约占4.34%,B占21.43%,C占7.17%,国库券占67.06%,此时风险(方差为0.005 2)进一步下降.仔细观察问题1)和2)的两个结果,可以发现:后者投资在有风险资产(股票A,B,C)上的比例大约都是前者相应的比例的一半!也就是说,无论你的期望收益大小如何,你手上所握有的风险资产相互之间的比例居然是不变的!变化的只是投资于风险资产与无风险资产之间的比例.有趣的是,这一现象在一般情况下也是成立的,一般称为"分离定理",即风险资产之间的相对投资比例与期望收益无关.发现这一重要规律的Tobin教授于1981年获得了诺贝尔经济学奖.

评注 例1介绍的均值-方差模型最重要的思想是用收益的方差刻画风险,它的提出被认为是"华尔街的第一次革命".在实际投资决策中,还存在其他度量风险的指标,如常用的还有下半方差(lower semi-variance)、平均绝对离差(MAD: mean absolute deviation)、风险价值(VaR: value at risk)、条件风险价值(CVaR: conditional VaR)等.

例 2 投资的收益和风险(1998 年全国大学生数学建模竞赛 A 题)

问题 市场上有 n 种资产(如股票、债券、……)$S_i(i=1,\cdots,n)$ 供投资者选择,某公司有数额为 M 的一笔相当大的资金可用作一个时期的投资.公司财务分析人员对这 n 种资产进行了评估,估算出在这一时期内购买 S_i 的平均收益率为 r_i,并预测出购买 S_i 的风险损失率为 q_i.考虑到投资越分散,总的风险越小,公司确定,当用这笔资金购买若干种资产时,总体风险可用所投资的 S_i 中最大的一个风险来度量.

购买 S_i 要付交易费,费率为 p_i,并且当购买额不超过给定值 u_i 时,交易费按购买 u_i 计算(不买当然无须付费).另外,假定同期银行存款利率是 $r_0(r_0=5\%)$,且既无交易费又无风险.

1) 已知 $n=4$ 时的相关数据如表 2 所示:

表 2 四种资产的相关数据

S_i	r_i/%	q_i/%	p_i/%	u_i/元
S_1	28	2.5	1	103
S_2	21	1.5	2	198
S_3	23	5.5	4.5	52
S_4	25	2.6	6.5	40

试给该公司设计一种投资组合方案,即用给定的资金 M,有选择地购买若干种资产或存银行生息,使净收益尽可能大,而总体风险尽可能小.



数据文件 4-2
投资收益和风险

2) 试就一般情况对以上问题进行讨论,并利用数据文件 4-2 中的数据进行计算.[71]

模型建立 设购买资产 S_i 的金额为 $x_i(i=1,\cdots,n)$,存银行的金额为 x_0,投资组合向量记为 $\boldsymbol{x}=(x_0,x_1,\cdots,x_n)$,它是这个问题的决策变量.题目中明确指出要考虑两类目标,即投资的净收益和风险,分别记为 $V(\boldsymbol{x})$ 和 $Q(\boldsymbol{x})$.

由于购买 S_i 要付交易费,费率为 p_i,并且当购买额不超过给定值 u_i 时,交易费按购买 u_i 计算(不买当然无须付费),所以购买 S_i 的交易费为

$$c_i(x_i)=\begin{cases}0, & x_i=0,\\ p_iu_i, & 0<x_i<u_i,\\ p_ix_i, & x_i\geqslant u_i,\end{cases}\quad i=1,2,\cdots,n \tag{7}$$

对于存银行的金额 x_0,显然 $c_0(x_0)=0$.

题目给出购买 S_i 的平均收益率为 r_i,因此投资 S_i 的平均净收益为

$$V_i(x_i)=r_ix_i-c_i(x_i),\quad i=0,1,\cdots,n \tag{8}$$

于是投资组合 $\boldsymbol{x}$ 的平均净收益为

$$V(\boldsymbol{x})=\sum_{i=0}^{n}V_i(x_i) \tag{9}$$

题目给出购买 S_i 的风险损失率为 q_i,且投资组合 $\boldsymbol{x}$ 的风险用 S_i 中最大的一个风险来度量,所以投资组合 $\boldsymbol{x}$ 的风险为

$$Q(\boldsymbol{x})=\max_{0\leqslant i\leqslant n}q_ix_i \tag{10}$$

投资所需的总资金为

$$I(\boldsymbol{x}) = \sum_{i=0}^{n} (x_i + c_i(x_i)) \tag{11}$$

题目给出公司提供的资金是 M(M 相当大).

为了使投资的净收益尽量高和风险尽量低,这个问题归结为以下多目标规划模型:

$$\begin{aligned} &\min \quad (Q(\boldsymbol{x}), -V(\boldsymbol{x})) \\ &\text{s.t.} \quad I(\boldsymbol{x}) = M \\ &\qquad \boldsymbol{x} \geqslant \boldsymbol{0} \end{aligned} \tag{12}$$

模型求解 设定风险和净收益的权重分别为 w, $1-w$, 即目标函数如下:

$$\min \quad wQ(\boldsymbol{x}) - (1-w)V(\boldsymbol{x}) \tag{13}$$

程序文件 4-23
投资收益和风险
prog0408b.lg4

我们只针对第一组数据($n=4$)计算,假设总资金 M 为 1 万元,对 w 在 0 到 1 之间按 0.1 的间隔取值计算,结果如表 3 所示.

表 3 权值按 0.1 的间隔对应的风险与收益

	权值										
	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
风险/元	247.5	247.5	247.5	247.5	247.5	247.5	247.5	247.5	92.25	59.40	0.00
收益/元	2 673	2 673	2 673	2 673	2 673	2 673	2 673	2 673	2 165	2 016	500

程序文件 4-24
投资收益和风险
的有效前沿
prog0408c.m

可以看出,在权值 w 小于 0.7 时,计算结果相同,此时收益最大,风险也大,对应的投资方案是将所有资金购买资产 S_1;当 w 从 0.7 增加到 1.0 时,风险和收益逐步下降;收益最低为 500 元(即收益率为 5%),此时风险为 0,对应的投资方案显然是将所有资金存入银行.

我们可以在 0.7 到 1.0 之间取更多权值计算,更仔细地观察结果的变化过程.例如,对 w 按 0.001 的间隔取值计算,可得表 4 结果(只列出部分结果).

表 4 权值按 0.001 的间隔对应的风险与收益(部分结果)

	权值										
	0.766	0.767	0.810	0.811	0.824	0.825	0.962	0.963	0.964	0.965	0.966
风险/元	247.5	92.29	92.25	78.49	78.49	59.40	59.40	2.97	2.97	2.86	0.00
收益/元	2 673	2 165	2 165	2 106	2 106	2 016	2 016	576	576	573	500

将这个结果画图可得到有效前沿如图 1 所示.图中曲线存在一个明显的拐点,其坐标大约是(59,2 016).对于风险和收益没有明显偏好的投资者来说,这一点应该是一个满意的选择,其对应的投资方案是购买 $S_1 \sim S_4$ 的资金(不含手续费)分别为 2 376 元、3 960 元、1 080 元、2 284 元,没有资金存入银行.

对于问题中给出的第二组数据,可类似计算.

评注 本节介绍的两个例子只是基本模型,现实中的投资决策需要考虑的因素还有很多.例如,我们在例 1 评注中已经指出还存在其他的风险指标.又如,本节都假设投资(比例)非负,这仅适用于不考虑融资融券等买空卖空行为的投资活动.实际中可供投资的资产品种成千上万,投资者通常还会为投资的资产总数量设定上下限,为单一资产的投资比例设定上下限等.另外,还可以考虑投资人通常并不拥有完全的信息、投资人之间存在复杂的博弈行为等因素.

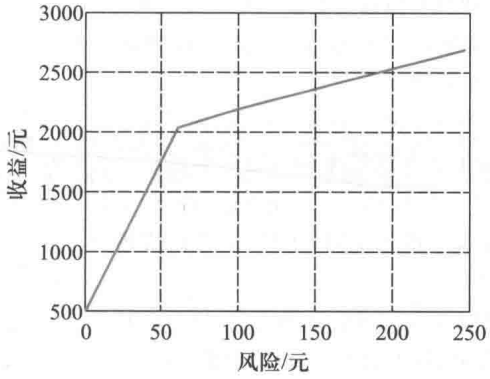


图 1　例 2 的有效前沿

复习题

1. 对一个随机变量 R,其下半方差定义为 $\mathrm{Var}^-(R)=E(\max(ER-R,0))^2$.请对例 1 采用下半方差作为风险指标建立模型,求解并与例 1 的结果比较.你觉得采用哪个风险指标更合适?

2. 例 2 中的约束(7)是一个分段线性函数,请你分别按照以下两种思路对这个约束进行预处理,并对模型进行求解.

(1) 由于总的投资金额 M 较大,而购买各种资产的阈值 u_i 都不太大,可将约束(7)用线性约束 $c_i(x_i)=p_ix_i$ 近似,使模型得到简化.请进一步分析这一近似的效果与 M 的大小有什么关系?

(2) 对每种资产,引入两个 0-1 变量,对投资额是否为 0、是否大于阈值 u_i 进行刻画,得到约束(7)的另一种等价描述.

第 4 章训练题

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、到期税前收益如下表所示.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.

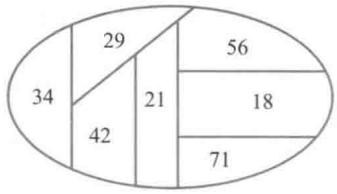
证券名称	证券种类	信用等级	到期年限	到期税前收益/%
A	市政	2	9	4.3
B	代办机构	2	15	5.4
C	政府	1	4	5.0
D	政府	1	3	4.4
E	市政	5	2	4.5

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[10]?

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)表示在右图上.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[61].



3. 一家保姆服务公司专门向顾主提供保姆服务.公司向顾主收取费用后,统一给保姆发工资,每人每月工资固定.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.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,请你为公司制订下一年的招聘计划;哪些季度需求的增加不影响招聘计划?可以增加多少?

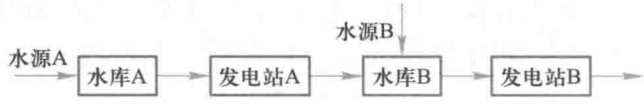
(2) 如果公司在每个季度结束后允许解雇保姆,请为公司制订下一年的招聘计划[10].

4. 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10^5 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.只有不执行当月飞行任务的熟练飞行员可以作为当月的教练指导新飞行员进行训练,每名教练每个月正好指导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)如下表所示,请你为甲方安排一个飞行计划.

	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

如果每名教练每个月可以指导不超过20名新飞行员(不包括他自己在内)进行训练,模型和结果有哪些改变[1]?

5. 某电力公司经营两座发电站,发电站分别位于两个水库上,位置如下图所示.



已知发电站A可以将水库A的10 000 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 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,多余的电能只能够以 140 元/千度的价格售出.水库 A,B 的其他有关数据如下表(单位:10^4 m^3).

		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 千度 = 10^3kW · h).

6. 如下图所示,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及各个工厂的污水流量和污水质量浓度均已知道.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.



先建立一般情况下的数学模型,再求解以下的具体问题:

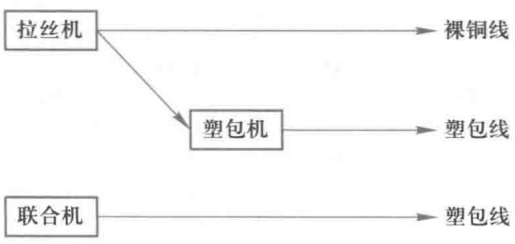
设上游江水流量为 1 000×10^{12}L/min,污水质量浓度为 0.8 mg/L,3 个工厂的污水流量均为 5×10^{12}L/min,污水质量浓度(从上游到下游排列)(单位:mg/L)分别为 100,60,50,处理系数均为 1 万元/(((10^{12}L/min)×(mg/L)),3 个工厂之间的两段江面的自净系数(从上游到下游)分别为0.9 和0.6.国家标准规定水的污染浓度不能超过 1 mg/L.

(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 3 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?[80]

(提示:可以进行适当简化和近似,建立线性规划模型.)

7. 生产裸铜线和塑包线的工艺如下图所示:



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下表:

	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/($m \cdot h^{-1}$)	1 000	1 500	1 200	1 600	1 600
规格2生产效率/($m \cdot h^{-1}$)	800	1 400	1 000	1 300	1 200
废品率/%	2	2	3	3	3
每千米废品损失/元	30	30	50	50	50

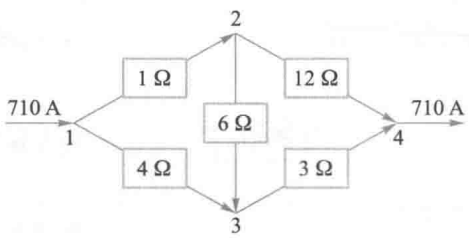
已知市场对两种规格裸铜线的需求分别为3 000 km和2 000 km,对两种规格塑包线的需求分别为10 000 km和8 000 km.按照规定,新购及改进设备按每年5%提取折旧费,老设备不提;每台机器每年最多只能工作8 000 h.为了满足需求,确定使总费用最小的设备选用方案和生产计划[66].

8. 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,如下表所示(单位:min):

	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.假定现在时间是早8:00,问他们最早何时能离开公司[66]?

9. 在如下图所示的电网中,需要从节点1传送710 A的电流到节点4.当电流通过电网传送时,存在功率损失,而电流在传送时将"自然而然"地使总功率损失达到最小.请根据这种自然特性,确定流过各个电阻的电流,并与按照电路定律列出的代数方程组的解相比较.[72]



10. 某电子系统由 3 种元件组成，系统正常运转需要 3 种元件都正常工作.如一种或多种元件安装几个备用件，将提高元件的可靠性.系统运转的可靠性是各元件可靠性的乘积，而每一元件的可靠性是备用件数量的函数，具体数值如下表中 2~7 行所示.3 种元件的价格和质量如表中 8~9 行所示.已知全部备用件的预算限制为 150 元，质量限制为 20 kg，应如何安装备用件，可使系统运转的可靠性最大?[66]

		元件序号		
		1	2	3
备用件数	0	0.5	0.6	0.7
	1	0.6	0.75	0.9
	2	0.7	0.95	1.0
	3	0.8	1.0	1.0
	4	0.9	1.0	1.0
	5	1.0	1.0	1.0
单件价格/元		20	30	40
单件质量/kg		2	4	6

11. 某公司利用钢材和铝材作为原材料，生产两种产品(A 和 B).单件产品 A 需消耗钢材 6 kg，铝材 8 kg，劳动力 11 h，利润 5 000 元(不含工人加班费)；单件产品 B 需消耗钢材12 kg，铝材 20 kg，劳动力 24 h，利润 11 000 元(不含工人加班费).该企业目前可提供钢材200 kg，铝材 300 kg，劳动力 300 h.如果要求工人加班，每小时加班费 100 元.请制订生产计划，最大化公司的利润和最小化工人加班时间.

12. 某银行营业部设立 3 个服务窗口，分别为个人业务、公司业务和特殊业务(如外汇和理财等).现有 3 名服务人员，每人处理不同业务的效率(每天服务的最大顾客数)，以及每人处理不同业务的质量(如顾客的满意度)见下表.如何为服务人员安排相应的工作(服务窗口)?

最大顾客数

	个人业务	公司业务	特殊业务
员工 1	20	12	10
员工 2	12	15	9
员工 3	6	5	10

顾客满意度

	个人业务	公司业务	特殊业务
员工 1	6	8	10
员工 2	6	5	9
员工 3	9	10	8

13. 某农场有 3 万亩农田(1 亩 = 666.67 m^2)，准备种植玉米、大豆和小麦 3 种农作物.每种作物每

亩需施化肥分别为 0.12 t,0.20 t,0.15 t.预计秋后玉米每亩可收获500 kg,售价为 1.2 元/kg;大豆每亩可收获 200 kg,售价为 6 元/kg;小麦每亩可收获 300 kg,售价为3.5 元/kg.农场年初规划时依次考虑以下几个方面:

(1) 年终总收益尽量不低于 1 650 万元.

(2) 总产量尽量不低于 1.25×10^4 t.

(3) 小麦产量以 0.5×10^4 t 为宜.

(4) 大豆产量尽量不低于 0.2×10^4 t.

(5) 玉米产量尽量不超过 0.6×10^4 t.

(6) 农场目前能提供 5 000 t 化肥;若不够,可额外购买,但希望额外购买量越少越好.请为该农场制订种植规划.[66]

更多案例……

4-1 销售代理的开发与中断

第5章　微分方程模型

微分方程是包含连续变化的自变量、未知函数及其变化率(函数的微分或导数)的方程式,当我们的研究对象涉及某个过程或物体随时间连续变化的规律时,通常会建立微分方程模型.例如,进行人口预测时应建立包含人口数量及增长率的微分方程,研究发射火箭的高度时要建立燃料燃烧的推力所提供的火箭加速度与速度、高度关系的微分方程.

建立微分方程模型通常采用机理分析方法,如果研究对象来自工程技术、科学研究,大多归属于力、热、光、声、电等物理领域,那么牛顿定律、热传导定律、电路原理等物理规律可能是必不可少的理论依据,而若研究对象属于人口、经济、医药、生态等非物理领域,则要具体分析该领域特有的机理,找出研究对象所遵循的规律.

事实上,高等数学课程中解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".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,答案是唯一的,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.做出不同的假设,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.

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,而是研究某种意义下稳定状态的特征,特别是当时间充分长以后动态过程的变化趋势.这时常常不需要求解微分方程(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.

5.1　人口增长

2011年10月31日,在俄罗斯海港城市加里宁格勒出生的一名婴儿被联合国人口基金会认定为世界上第70亿个居民.全球人口在哪一天真正达到70亿,真正的"第70亿个人"诞生在哪里,都只具有象征意义,而人口迅猛增长给自然资源和生态环境带来的压力,给各国发展与稳定带来的挑战却是严峻的现实,需要全人类积极、妥善地应对.

联合国人口基金会的统计显示(表1),世界人口翻番的时间已由过去的一个多世纪缩短为20世纪的40多年.从1960年开始,每十三四年就增长10亿.随着文化普及和妇女社会地位的提高,21世纪的女性平均生育率已显著下降,但庞大的人口基数仍将驱动"人口时钟"快速前行.据预测,如果目前的生育率不变,21世纪中期世界人口将突破90亿,此后人口增速才会放缓,到21世纪末超过100亿.

表1 世界人口

年份	1625	1804	1927	1960	1974	1987	1999	2011
人口/亿	5	10	20	30	40	50	60	70

中国是世界第一人口大国,地球上每5个人中就有一位中国同胞.在20世纪的一段时间内,我国人口的增长速度过快,从20世纪70年代开始实施计划生育政策,经过40多年的努力,有效控制了人口快速增长,缓解了资源和环境的压力(见表2).我国人口的年净增量由最多时的2 000多万降到目前的600多万.但是,随着计划生育政策实施周期拉长,其积极效应正在递减,人口老龄化提速、性别比失调等结构性矛盾却日益凸显,对人口政策进行适当调整的措施已经开始实施.

表2 中国人口①

年份	1949	1953	1965	1982	1990	2000	2010
人口/亿	5.42	5.88	7.25	10.17	11.43	12.67	13.40

认识人口数量的变化过程,建立数学模型描述人口发展规律,做出较准确的增长预测,是制定积极、稳妥的人口政策的前提.下面引入两个最基本的人口模型,利用表3给出的美国人口统计数据进行参数估计,并作模型检验和增长预测.

表3 美国人口②

年份	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
增长率/10年	0.294 9	0.311 3	0.298 6	0.296 9	0.290 7	0.301 2	0.308 2	0.245 2
年份	1870	1880	1890	1900	1910	1920	1930	1940
人口/百万	38.6	50.2	62.9	76.0	92.0	105.7	122.8	131.7
增长率/10年	0.243 5	0.242 0	0.205 1	0.191 4	0.161 4	0.145 7	0.105 9	0.105 9
年份	1950	1960	1970	1980	1990	2000	2010	
人口/百万	150.7	179.3	203.2	226.5	248.7	281.4	308.7	
增长率/10年	0.157 9	0.146 4	0.116 1	0.100 4	0.110 4	0.134 9		

一、指数增长模型

1. 一个常用的人口预测公式

如果已知今年人口为 x_0,年增长率为 r,你一定会用下面的公式预测 k 年后的人口 x_k,

$$x_k = x_0(1+r)^k \tag{1}$$

显然,这个公式的基本前提是年增长率 r 在 k 年内保持不变.

① 来自《中国统计年鉴》,不含港澳台人口.

② 数据来自参考文献[24]第117页,小数点后取1位数字.2010年的数据是美国人口普查局于2010年12月21日公布的.之所以采用美国的人口数据,是因为有长达200多年、每10年一次完整的统计资料,且美国是世界上人口数量较少受到战争影响的大国.

利用(1)式还可以根据人口统计数据估计年增长率,例如,若经过 n 年人口翻了一番,那么这期间的年平均增长率约为 $(70/n)\%$.如世界人口从 1960 年到 1999 年翻了一番(表 1),则年平均增长率约为 $(70/39)\% = 1.8\%$;中国人口从 1953 年到 1990 年将近翻了一番(表 2),那么年平均增长率也不低于 1.8%.

2. 人口指数增长模型的建立

当考察一个国家或一个较大地区的人口随着时间延续而变化的规律时,为了利用微积分这一数学工具,可以将人口看作连续时间 t 的连续可微函数 $x(t)$.记初始时刻($t=0$)的人口为 x_0.假设单位时间人口增长率为常数 r,$rx(t)$ 就是单位时间内 $x(t)$ 的增量 $\frac{dx}{dt}$,于是得到 $x(t)$ 满足的微分方程和初始条件

$$\frac{dx}{dt}=rx, x(0)=x_0 \tag{2}$$

由(2)式容易解出

$$x(t)=x_0e^{rt} \tag{3}$$

$r>0$ 时,人口将按指数规律无限增长,(2)或(3)式称为指数增长模型.

请读者解释,我们常用的公式(1)是(3)式的离散近似形式(复习题 2).

指数增长模型又被称为 Malthus 人口模型.Malthus 是 18 世纪末、19 世纪初的英国著名人口学家、经济学家,他在 1798 年发表的“人口原理”中断言:“正常情况下,人口每 25 年以几何级数增加,而生活资料只以算术级数增长”,得到“占优势的人口繁殖力为贫困和罪恶所抑制,因而使现实的人口得以与生活资料保持平衡”的所谓人口均衡原理,继而做出“平等制度不可能实现、济贫法的作用适得其反”等反对社会改革的推论.Malthus 的人口理论从问世开始,就一直遭到猛烈的抨击,但是它在客观上提醒人们注意人口与生活资料的协调,防止人口过快增长,从而成为现代人口理论的开端.

3. 指数增长模型的参数估计

根据数据对模型的参数进行估计又称数据拟合,对(3)式中参数 r 和 x_0 的估计主要有两种方法.

方法一　直接用人口数据和线性最小二乘法,将(3)式取对数得

$$y=rt+a, y=\ln x, a=\ln x_0 \tag{4}$$

根据表 3 中 1790 年至 2000 年的美国人口数据,用 MATLAB 编程计算,1790 年取作 $t=0$(本节以后均作此规定),得到 $r=0.202\ 0/10$ 年,$x_0=6.049\ 6$.将 r 和 x_0 代入(3)式,计算结果列入表 4 第 3 列,可与实际数据(表 4 第 2 列,同表 3)比较.

方法二　先对人口数据作数值微分,再计算增长率并将其平均值作为 r 的估计;x_0 直接采用原始数据.

数值微分的中点公式如下:设函数 $x(t)$ 在分点 t_0, t_1, …, t_n(等间距 Δt)的离散值为 x_0, x_1, …, x_n,则函数在各分点的导数近似值为

$$x'(t_k)=\frac{x_{k+1}-x_{k-1}}{2\Delta t},\quad k=1,2,\cdots,n-1$$

$$x'(t_0)=\frac{-3x_0+4x_1-x_2}{2\Delta t}, x'(t_n)=\frac{x_{n-2}-4x_{n-1}+3x_n}{2\Delta t} \tag{5}$$

根据表3中1790年至2000年的美国人口数据，利用(5)式的 $x'(t_k)$ $(k=0,1,\cdots,n)$ 计算增长率 $\dfrac{x'(t_k)}{x(t_k)}$（以10年为间距 Δt），记作 r_k，表3第3行的增长率就是这样得到的.将 $r_k(k=0,1,\cdots,n)$ 的均值作为平均增长率 r 的估计，得到 $r=0.205\ 2/10$ 年，再取 $x_0=3.9$（1790年的人口），将 r 和 x_0 代入(3)式的计算结果列入表4第4列.

图1是用这两种方法估计的参数计算的美国人口增长曲线（图中圆点是实际人口）.由表4和图1看到，用指数增长模型计算的美国人口与实际数据相差很大，这主要是由于在长达200多年的时间内，假设增长率 r 为常数与实际情况不符合的缘故.

表4 用指数增长模型计算的美国人口

年份	实际人口/百万	指数增长模型（方法一）	指数增长模型（方法二）	改进的指数增长模型
1790	3.9	6.0	3.9	3.9
1800	5.3	7.4	4.8	5.4
1810	7.2	9.1	5.9	7.3
1820	9.6	11.1	7.2	9.8
1830	12.9	13.6	8.9	13.1
1840	17.1	16.6	10.9	17.2
1850	23.2	20.3	13.4	22.3
1860	31.4	24.9	16.4	28.7
1870	38.6	30.4	20.1	36.5
1880	50.2	37.3	24.7	45.8
1890	62.9	45.6	30.4	56.9
1900	76.0	55.8	37.3	69.9
1910	92.0	68.3	45.8	84.8
1920	105.7	83.6	56.2	101.8
1930	122.8	102.3	69.0	120.8
1940	131.7	125.2	84.7	141.6
1950	150.7	153.3	104.0	164.2
1960	179.3	187.6	127.6	188.3
1970	203.2	229.6	156.7	213.4
1980	226.5	281.0	192.4	239.1
1990	248.7	343.8	236.2	264.8
2000	281.4	420.8	290.0	290.0
误差平方和/10^4		3.474 2	2.204 8	0.113 3

4. 改进的指数增长模型

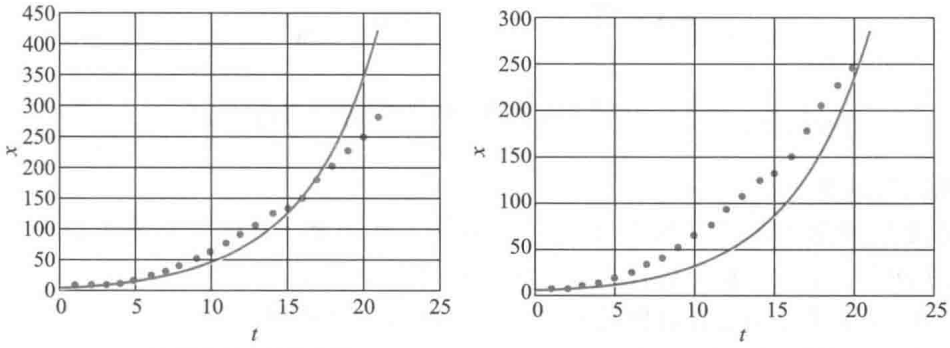
将表3第3行美国人口10年增长率 r 对时间 t 作图，如图2中圆点所示，可以看

到，r 从 19 世纪上半叶的 0.3/10 年左右，下降到 20 世纪末只略大于 0.1/10 年.针对这么大的变化，应该改进指数增长模型(4)中人口增长率为常数的假设，将 r 视为 t 的函数 $r(t)$.若认为图 2 中圆点大致在一条直线附近，则可假设$r(t)=r_0-r_1t$，将指数增长模型(2)改写为

$$\frac{dx}{dt}=r(t)x=(r_0-r_1t)x,x(0)=x_0 \tag{6}$$

这是可分离变量的微分方程，其解为

$$x(t)=x_0e^{(r_0t-r_1t^2/2)} \tag{7}$$



(a) 用方法一估计参数　　(b) 用方法二估计参数

图 1　用指数增长模型计算的美国人口

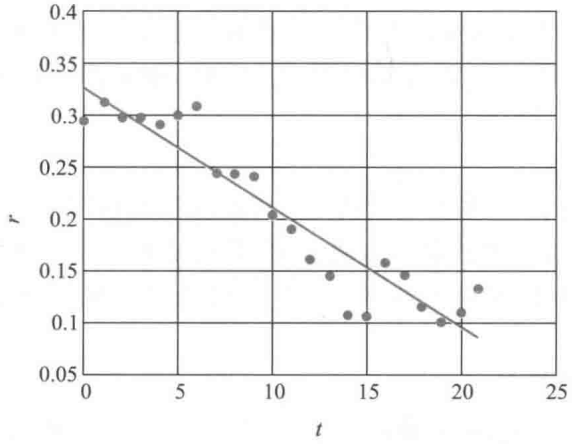


图 2　美国人口 10 年增长率 t-r 散点图及拟合直线

实际应用中，方程(6)中的 $r(t)$ 可以根据人口增长率的变化情况，选取简单、合适的其他函数形式.

由美国人口 10 年增长率数据，利用最小二乘法得到 $r_0=0.325\ 2$，$r_1=0.011\ 4$，再取 $x_0=3.9$(1790 年的人口)，将 r_0，r_1 和 x_0 代入(7)式，计算结果列入表 4 第 5 列，其图形见图 3.

比较图 3 和图 1，可以明显地看到这个模型的改进效果.定量上常用误差平方和(计算值与实际值之差的平方和)来度量模型或方法与实际数据拟合的程度，表 4 最后一行给出了指数增长模型的两种参数估计方法及改进的指数增长模型的误差平方和，

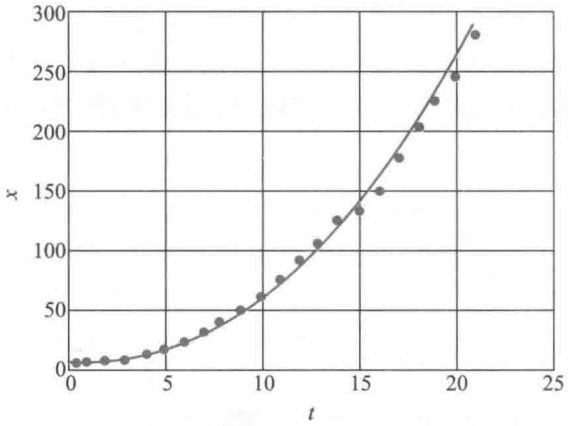


图3 用改进的指数增长模型计算的美国人口

其差别相当显著.

历史上,指数增长模型与19世纪以前欧洲一些国家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,虽然指数增长模型不适于描述、也不能预测较长时期的人口演变过程,但是用它做短期人口预测可以得到较好的结果.显然是因为在这些情况下,模型的基本假设——人口增长率是常数——大致成立.

改进的指数增长模型考虑到增长率随时间的变化,计算结果有所改善,但是它没有反映增长率下降的成因,增长率函数形式也呈不确定性,给应用带来不便.为了建立更符合实际情况的人口增长模型,必须考察人口增长率下降的机理,修改原来的模型假设.

二、logistic 模型

在分析人口增长到一定数量后增长率下降的主要原因时,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.logistic 模型就是考虑到这些因素,对指数增长模型的基本假设进行修改后得到的[39].

1. logistic 模型的建立

资源和环境对人口增长的阻滞作用体现在对增长率 r 的影响上,使 r 随着人口数量 x 的增加而下降.将 r 表示为 x 的函数 $r(x)$,并且取既简单又便于应用的线性减函数 $r(x)=a+bx$.为了赋予增长率函数 $r(x)$ 中的系数 a,b 以实际含义,引入两个参数:

1) 内禀增长率 r r 是(理论上)$x=0$ 的增长率,即 $r(0)=r$,于是 $a=r$;

2) 人口容量 x_m x_m 是资源和环境所能容纳的最大人口数量,当 $x=x_m$ 时人口不再增长,即 $r(x_m)=r+bx_m=0$,得到 $b=-\dfrac{r}{x_m}$.

由此导出的增长率函数为 $r(x)=r\left(1-\dfrac{x}{x_m}\right)$.用 $r(x)$ 代替模型(2)的 r,得到

$$\frac{dx}{dt}=rx\left(1-\frac{x}{x_m}\right),x(0)=x_0 \tag{8}$$

方程(8)右端的因子 rx 体现人口自身的增长趋势,因子$\left(1-\frac{x}{x_m}\right)$则体现了资源和环境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.

以 x 为横轴、$\frac{dx}{dt}$为纵轴对(8)式作图,得到一条抛物线(图 4),$x=\frac{x_m}{2}$时,$\frac{dx}{dt}$最大.根据图 4 中$\frac{dx}{dt}$随 x 增加而变化的情况,可以对曲线 $x(t)$ 的形状作如下分析.

设 $t=0$ 时 $x_0<\frac{x_m}{2}$,随着 t 的增加$\frac{dx}{dt}$变大,所以 x 增长越来越快,曲线 $x(t)$ 呈下凸状上升;$x>\frac{x_m}{2}$后$\frac{dx}{dt}$变小,x 增长越来越慢,$x(t)$ 呈上凸状升高;$x=\frac{x_m}{2}$为曲线的拐点;$x\to x_m$ 时 $\frac{dx}{dt}\to 0$,于是 $x=x_m$ 是曲线 $x(t)$ 的渐近线.由以上分析可以大致画出 $x(t)$ 的图形如图 5,它是一条 S 形曲线.

实际上,微分方程(8)可以用分离变量法求解得到

$$x(t)=\frac{x_m}{1+\left(\frac{x_m}{x_0}-1\right)e^{-rt}} \tag{9}$$

读者可以随意设定参数 r,x_0,x_m,编程画出(9)式的图形,与图 5 比较.(8)式称为 logistic 方程,(9)式称为 logistic 曲线,统称为 logistic 模型.

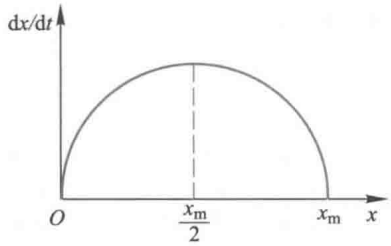


图 4 logistic 模型的 $x-dx/dt$ 曲线

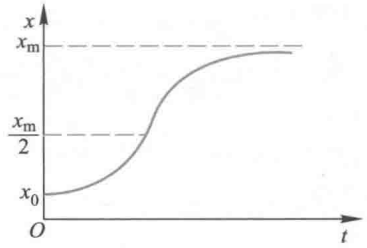


图 5 logistic 模型的 $t-x$ 曲线

2. logistic 模型的参数估计

可以用两种方法估计(9)式的参数 r,x_0,x_m.

方法一 将方程(8)改写为

$$\frac{dx/dt}{x}=r-\frac{r}{x_m}x,x(0)=x_0 \tag{10}$$

像指数增长模型参数估计的方法二一样,对人口数据 $x(t)$ 作数值微分后计算增长率,即(10)式左端;对(10)式右端 x 用线性最小二乘法估计 r,x_m;x_0 直接采用原始数据.

根据表 3 中 1790 年至 2000 年的美国人口数据,利用数值微分计算的 10 年增长率已列入表 3 第 3 行,将其对 x 作图,如图 6 中圆点所示.圆点大致在一条直线附近,用线性最小二乘法编程计算的结果为 $r=0.280\ 5,x_m=352.054\ 8$.再取 $x_0=3.9$(1790 年的人口),将 r,x_m 和 x_0 代入(9)式,计算结果列入表 5 第 3 列.

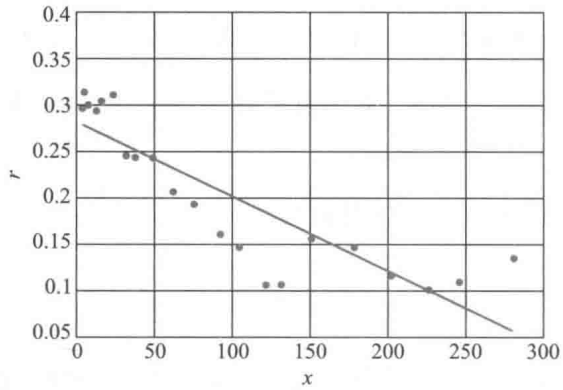


图6 美国人口增长率 x-r 散点图及拟合直线

方法二 直接用人口数据和非线性最小二乘法估计(9)式的 r, x_0, x_m.

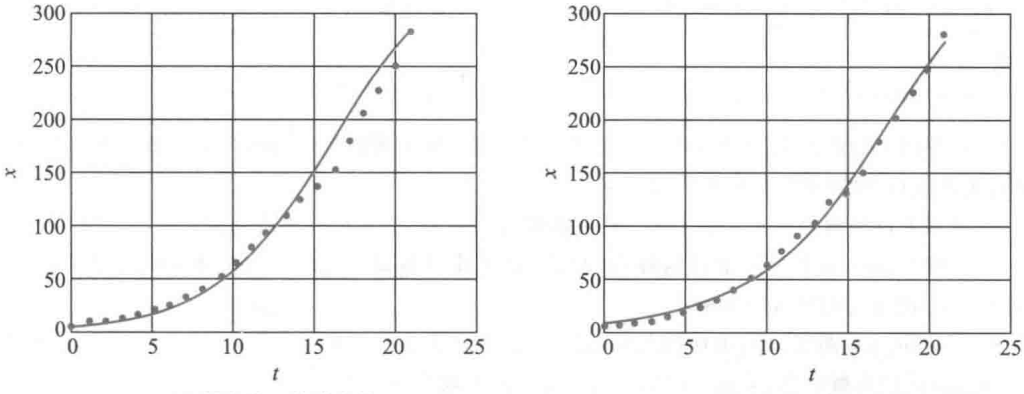
根据表3中1790年至2000年的美国人口数据,用非线性最小二乘法对(9)式编程计算得到 $r=0.215\ 5$, $x_0=7.696\ 2$, $x_m=443.993\ 1$,再代入(9)式,计算结果列入表5第4列.

表5最后一行给出了logistic模型的两种参数估计方法的误差平方和.图7是用这两种方法估计的参数计算的美国人口(图中圆点是实际人口).

表5 用logistic模型计算的美国人口

年份	实际人口/百万	logistic模型(方法一)	logistic模型(方法二)
1790	3.9	3.9	7.7
1800	5.3	5.1	9.5
1810	7.2	6.8	11.7
1820	9.6	8.9	14.5
1830	12.9	11.7	17.8
1840	17.1	15.3	21.9
1850	23.2	20.0	26.8
1860	31.4	26.0	32.8
1870	38.6	33.6	40.0
1880	50.2	43.2	48.5
1890	62.9	55.0	58.6
1900	76.0	69.3	70.5
1910	92.0	86.3	84.2
1920	105.7	105.8	99.9
1930	122.8	127.6	117.6
1940	131.7	151.2	137.1
1950	150.7	175.8	158.4
1960	179.3	200.3	180.9
1970	203.2	223.9	204.4
1980	226.5	245.8	228.3
1990	248.7	265.4	252.0
2000	281.4	282.4	275.1
误差平方和/10^3		2.810 4	0.458 2

对比表5与表4最后一行的误差平方和以及图7、图1和图3，不难看出，就1790年至2000年美国人口数据的拟合效果而言，logistic模型比指数增长模型有很大改善，特别是用非线性最小二乘估计参数的结果，效果很好(图7(b)).



(a) 用方法一估计参数　　(b) 用方法二估计参数

图7　用logistic模型计算的美国人口

三、模型检验和人口预测

表4、表5和图1、图3、图7给出的结果，都是利用1790年至2000年美国人口数据估计的参数代入两个模型计算得到的，将这些结果与同期的实际数据比较，能够反映模型与这些数据的拟合程度，还不是真正意义上的模型检验.

你可能已经注意到，在估计指数增长模型和logistic模型的参数过程中，我们没有用表3中2010年的美国人口，留下这个实际数据正是为了用来做模型检验的.用这两个模型及前面估计的参数计算2010年的美国人口，列入表6第2行，与实际人口308.7(百万)的误差见表6第3行.用改进的指数增长模型和logistic模型计算结果的误差是可以接受的.

最后，我们用改进的指数增长模型和第2种方法估计参数的logistic模型，对2020年的美国人口作预测.将2010年美国实际人口加入，重新估计模型参数①，然后计算2020年的人口，得到表6第4行的两个数据.至于预测的准确性如何，需要等到2020年美国人口调查结果的公布，让我们拭目以待.

表6　指数增长模型和logistic模型的检验和预测

	实际人口	指数增长模型(方法一)	指数增长模型(方法二)	改进的指数增长模型	logistic模型(方法一)	logistic模型(方法二)
2010年人口/百万(误差)	308.7	515.0 (66.8%)	356.0 (15.3%)	314.0 (1.7%)	296.8 (-3.9%)	297.0 (-3.8%)
预测2020年人口/百万	?			327.8		326.8

小结　用数学工具描述人口变化规律，关键是对人口增长率做出合理、简化的假定.指数增长模型关于人口增长率是常数的假设过于简化，不适于作长期人口预测.logistic模型对增长率随人口增加而线性减少的假设，是一个合理、简化的修正.

logistic方程(曲线)是比利时生物数学家Verhulst 1845年提出的，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认为受到已售量的阻滞增长作用.

评注　参数估计是建模的重要步骤，最小二乘法是参数估计的基本方法，并且应该尽量采用线性最小二乘法.当然，随着数学软件的发展，非线性最小二乘法的应用也是很方便的.对于微分方程模型来说，参数估计既可直接从方程的解出发，也可针对方程本身，应灵活掌握.

做模型检验时需要注意的一点是，用作检验的数据不应该参与建模过程的参数估计，正像裁判员不宜做运动员一样.

① 改进的指数增长模型重新估计的参数为 $r_0=0.326\ 6$，$r_1=0.011\ 6$，第2种方法估计参数的logistic模型为 $r=0.208\ 0$，$x_0=8.192\ 0$，$x_m=486.156\ 6$.

拓展阅读 5-1 logistic 方程的发展进程

本节介绍的两个模型都是以人口总数、人口平均增长率为研究对象，实际上，一个国家或地区的人口发展规律与年龄结构密切相关，考虑年龄结构的人口模型可参考更多案例 5-3.

复习题

1. 利用(1)式解释：若经过 n 年人口翻了一番，那么这期间的年平均增长率约为 $(70/n)\%$. 由表 3 估计美国人口 1950 年后的年平均增长率.

2. 解释人口预测公式 $x_k=x_0(1+r)^k$ 是指数增长模型 $x(t)=x_0e^{rt}$ 的离散近似形式.

3. 求解 logistic 方程(8)，验证其解为(9)式，设定几组参数 r,x_0,x_m，用软件编程画出解的曲线，分析 r 和 x_m 的变化对曲线的影响.

4. 写出 logistic 曲线 $x(t)$ 出现拐点时刻的表达式，分析这个时刻与参数 r,x_0,x_m 的关系. 按照本节两种方法估计的参数计算，美国人口增长曲线的拐点出现在哪一年？

5. 求解 logistic 方程(8)，将 t 表示为 x 的函数. 说明 logistic 有 log-like(像对数)的含义(参考拓展阅读 5-1 logistic 方程的发展进程).

5.2 药物中毒急救

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，氨茶碱的每次用量成人是 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 μg/mL 时，会出现严重中毒，达到 200 μg/mL 则可致命[60].

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 μg/mL 甚至 200 μg/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 的成年人血液总量约为 4 000 mL. 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$\mu(>0)$,$t=0$时血液中无药物.

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

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

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

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

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

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

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

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

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$y(t)=ae^{-\mu(t-\tau)}$,$t\geqslant\tau$.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)

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

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

结果分析 用MATLAB软件对(5),(6)作图,得图1.

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 μg/mL和

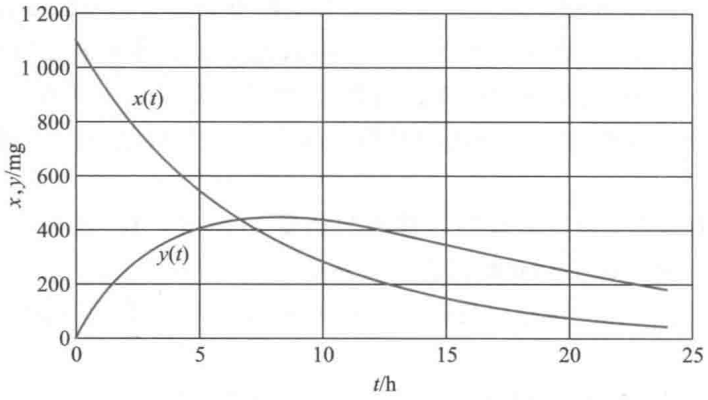


图1 胃肠道中药量 $x(t)$ 和血液系统中药量 $y(t)$

致命的血药浓度 200 μg/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 $y(2)=236.5$ 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 $t_1=4.87$ 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,$t_2=\dfrac{\ln(1+\lambda/6\mu)}{\lambda-\mu}=7.89$ 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(复习题 1).

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 $z(t)$)

$$\frac{dz}{dt}=\lambda x-\mu z,t\geqslant 2,x=1\ 100e^{-\lambda t},z(2)=236.5 \tag{7}$$

仍是一阶线性微分方程,只不过初始时刻为 $t=2$,当 $\lambda=0.138\ 6$(不变)而 $\mu=0.231\ 0$ 时,(7)的解为

$$z(t)=1\ 650e^{-0.138\ 6t}-1\ 609.5e^{-0.231\ 0t},t\geqslant 2 \tag{8}$$

用 MATLAB 软件对(8)作图,如图 2.

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

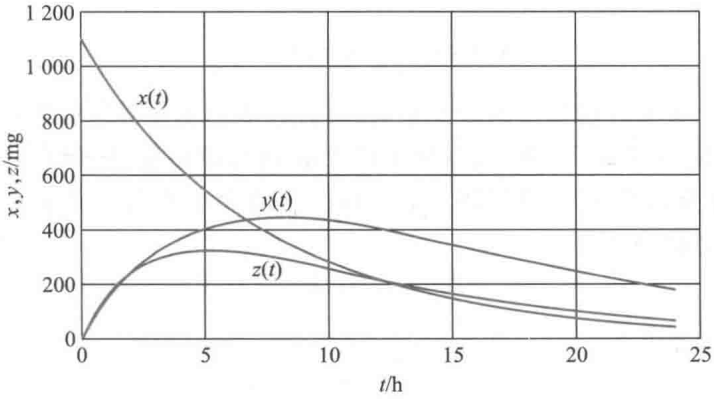


图 2 施救后血液系统中药量 $z(t)$($x(t)$,$y(t)$同图 1)

使 $z(t)$在施救后($t\geqslant 2$)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

$$\left.\frac{\mathrm{d}z}{\mathrm{d}t}\right|_{t=2}=(\lambda x-\mu z)\Big|_{t=2}=0 \tag{9}$$

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(复习题 2).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.

复习题

1. 利用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.

2. 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规律并作图.

5.3 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、林业这样的可再生资源,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[15,38].

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 方程(见 5.1 节),即

$$\dot{x}(t)=f(x)=rx\left(1-\frac{x}{N}\right) \tag{1}$$

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强度. 于是单位时间的捕捞量为

$$h(x)=Ex \tag{2}$$

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

$$\dot{x}(t)=F(x)=rx\left(1-\frac{x}{N}\right)-Ex \tag{3}$$

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(可参考基础知识 5-1).

基础知识 5-1 微分方程稳定性简介

令

$$F(x)=rx\left(1-\frac{x}{N}\right)-Ex=0$$

得到两个平衡点

$$x_0=N\left(1-\frac{E}{r}\right),\quad x_1=0 \tag{4}$$

不难算出

$$F'(x_0)=E-r,\quad F'(x_1)=r-E$$

所以若

$$E<r \tag{5}$$

有 $F'(x_0)<0$，$F'(x_1)>0$，故 x_0 点稳定，x_1 点不稳定；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.

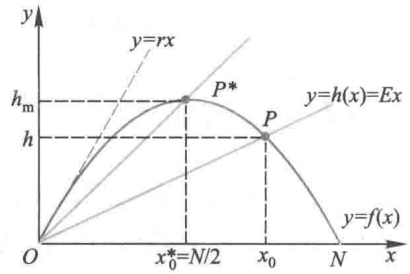


图 1 最大持续产量的图解法

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

$$x_0^* = \frac{N}{2} \tag{6}$$

且单位时间的最大持续产量为

$$h_m = \frac{rN}{4} \tag{7}$$

而由(4)式不难算出保持渔场鱼量稳定在 x_0^* 的捕捞率为

$$E^* = \frac{r}{2} \tag{8}$$

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

$$T = ph(x) = pEx, \quad S = cE \tag{9}$$

单位时间的利润为

$$R = T - S = pEx - cE \tag{10}$$

在稳定条件 $x = x_0$ 下,以(4)代入(10)式,得

$$R(E) = T(E) - S(E) = pNE\left(1 - \frac{E}{r}\right) - cE \tag{11}$$

用微分法容易求出使利润 $R(E)$ 达到最大的捕捞强度为

$$E_R = \frac{r}{2}\left(1 - \frac{c}{pN}\right) \tag{12}$$

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

$$x_R = \frac{N}{2} + \frac{c}{2p} \tag{13}$$

$$h_R = rx_R\left(1 - \frac{x_R}{N}\right) = \frac{rN}{4}\left(1 - \frac{c^2}{p^2N^2}\right) \tag{14}$$

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

$$E_S = r\left(1-\frac{c}{pN}\right) \tag{15}$$

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

$$p>\frac{c}{N} \tag{16}$$

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.

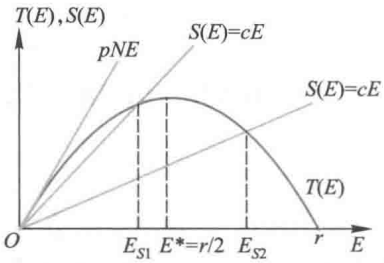


图2 盲目捕捞强度的图解法

将(15)代入(4)式，得到盲目捕捞下的渔场稳定鱼量为

$$x_S = \frac{c}{p} \tag{17}$$

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(复习题1，2).

复习题

1. 如果渔场鱼量的自然增长仍服从 logistic 方程，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，其结果与本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^*.

5.4 资金、劳动力与经济增长

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[45].

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

$$Q(t)=F(K(t),L(t)) \tag{1}$$

其中 F 为待定函数.对于固定的时刻 t,上述关系可写作

$$Q=F(K,L) \tag{2}$$

为寻求 F 的函数形式,引入记号

$$z=Q/L,\quad y=K/L \tag{3}$$

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

$$z=cg(y),\quad g(y)=y^{\alpha},\quad 0<\alpha<1 \tag{4}$$

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

$$Q=cK^{\alpha}L^{1-\alpha},\quad 0<\alpha<1 \tag{5}$$

由(5)式容易知道,Q 有如下性质

$$\frac{\partial Q}{\partial K},\frac{\partial Q}{\partial L}>0,\quad \frac{\partial^2 Q}{\partial K^2},\frac{\partial^2 Q}{\partial L^2}<0 \tag{6}$$

请读者解释(6)式的含义.

记 $Q_K=\dfrac{\partial Q}{\partial K}$,$Q_K$ 表示单位资金创造的产值;$Q_L=\dfrac{\partial Q}{\partial L}$,$Q_L$ 表示单位劳动力创造的产值,则从(5)式可得

$$\frac{KQ_K}{Q}=\alpha,\quad \frac{LQ_L}{Q}=1-\alpha,\quad KQ_K+LQ_L=Q \tag{7}$$

(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示为

$$Q=cK^{\alpha}L^{\beta},\quad 0<\alpha,\beta<1 \tag{8}$$

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

$$S=Q-rK-wL \tag{9}$$

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

$$\frac{Q_K}{Q_L}=\frac{r}{w} \tag{10}$$

再利用(7)式,有

$$\frac{K}{L}=\frac{\alpha}{1-\alpha}\frac{w}{r} \tag{11}$$

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做出合理的简化假设:

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

$$\frac{\mathrm{d}K}{\mathrm{d}t}=\lambda Q,\quad \lambda>0 \tag{12}$$

$$\frac{\mathrm{d}L}{\mathrm{d}t}=\mu L \tag{13}$$

方程(13)的解是①

$$L(t)=L_0\mathrm{e}^{\mu t} \tag{14}$$

将(4),(5)代入(12)式,得

$$\frac{\mathrm{d}K}{\mathrm{d}t}=c\lambda Ly^{\alpha} \tag{15}$$

注意到(3)式,有 $K=Ly$,再用(13)式可得

$$\frac{\mathrm{d}K}{\mathrm{d}t}=L\frac{\mathrm{d}y}{\mathrm{d}t}+\mu Ly \tag{16}$$

比较(15),(16),得到关于 $y(t)$ 的方程

$$\frac{\mathrm{d}y}{\mathrm{d}t}+\mu y=c\lambda y^{\alpha} \tag{17}$$

这是著名的 Bernoulli 方程,它的解是

$$y(t)=\left\{\frac{c\lambda}{\mu}\left[1-\left(1-\mu\frac{K_0}{\dot{K}_0}\right)\mathrm{e}^{-(1-\alpha)\mu t}\right]\right\}^{\frac{1}{1-\alpha}} \tag{18}$$

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 $\frac{\mathrm{d}Q}{\mathrm{d}t}>0$,由 $Q=cLy^{\alpha}$ 及(13),(17)式可算得

$$\frac{\mathrm{d}Q}{\mathrm{d}t}=cL\alpha y^{\alpha-1}\frac{\mathrm{d}y}{\mathrm{d}t}+c\mu Ly^{\alpha}=cLy^{2\alpha-1}[c\lambda\alpha+\mu(1-\alpha)y^{1-\alpha}] \tag{19}$$

---

① 以下凡字母下标为0者均表示该变量的初值.

将其中的 y 以(18)式代入,可知条件$\frac{dQ}{dt}>0$ 等价于

$$\left(1-\mu\frac{K_0}{\dot{K}_0}\right)e^{-(1-\alpha)\mu t}<\frac{1}{1-\alpha} \tag{20}$$

因为上式右端大于 1,所以当 $\mu\geqslant 0$(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

$$t<\frac{1}{(1-\alpha)\mu}\ln\left[(1-\alpha)\left(1-\mu\frac{K_0}{\dot{K}_0}\right)\right] \tag{21}$$

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$(1-\alpha)\left(1-\mu\frac{K_0}{\dot{K}_0}\right)\geqslant 1$,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 $\mu\leqslant 0$ 时该条件恒成立;而当 $\mu>0$ 时,由(18)式可得,$\frac{dy}{dt}>0$ 等价于

$$\left(1-\mu\frac{K_0}{\dot{K}_0}\right)e^{-(1-\alpha)\mu t}>0 \tag{22}$$

显然,此式成立的条件为 $\mu\frac{K_0}{\dot{K}_0}<1$,即

$$\mu<\dot{K}_0/K_0 \tag{23}$$

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.

复习题

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,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,并做出解释.

(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 Douglas 生产函数讨论,企业应雇用多少工人能使利润最大.

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[42].

案例精讲 5-1
香烟过滤嘴的作用

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，这个过程一直继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，它吸烟的动作、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 $a'：a$，$a'+a=1$.

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做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（$0\leqslant x\leqslant l$）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（$0\leqslant x\leqslant l_1$）.由假设 1，$w(x,0)=w_0$.

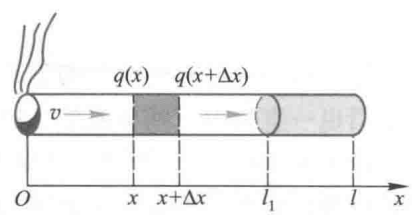


图 1　$x=0$ 处点燃的香烟

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

$$Q=\int_0^T q(l,t)\,\mathrm{d}t,\quad T=l_1/u \tag{1}$$

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

$$q(x)-q(x+\Delta x)=\begin{cases} bq(x)\Delta\tau, & 0\leqslant x\leqslant l_1,\\ \beta q(x)\Delta\tau, & l_1<x\leqslant l,\end{cases}\qquad \Delta\tau=\frac{\Delta x}{v}$$

其中 $\Delta\tau$ 是烟雾穿过 Δx 所需时间.令 $\Delta\tau\to 0$,得到微分方程

$$\frac{\mathrm{d}q}{\mathrm{d}x}=\begin{cases} -\dfrac{b}{v}q(x), & 0\leqslant x\leqslant l_1\\ -\dfrac{\beta}{v}q(x), & l_1<x\leqslant l\end{cases} \tag{2}$$

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

$$q(0)=aH_0,\quad H_0=uw_0 \tag{3}$$

求解(2)、(3)式时,先解出 $q(x)(0\leqslant x\leqslant l_1)$,再利用 $q(x)$ 在 $x=l_1$ 处的连续性确定 $q(x)(l_1\leqslant x\leqslant l)$.其结果为

$$q(x)=\begin{cases} aH_0\mathrm{e}^{-\frac{bx}{v}}, & 0\leqslant x\leqslant l_1\\ aH_0\mathrm{e}^{-\frac{bl_1}{v}}\mathrm{e}^{-\frac{\beta(x-l_1)}{v}}, & l_1<x\leqslant l\end{cases} \tag{4}$$

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

$$H(t)=uw(ut,t) \tag{5}$$

根据与第 1 步完全相同的分析和计算,可得

$$q(x,t)=\begin{cases} aH(t)\mathrm{e}^{-\frac{b(x-ut)}{v}}, & ut\leqslant x\leqslant l_1\\ aH(t)\mathrm{e}^{-\frac{b(l_1-ut)}{v}}\mathrm{e}^{-\frac{\beta(x-l_1)}{v}}, & l_1<x\leqslant l\end{cases} \tag{6}$$

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

$$q(l,t)=auw(ut,t)\mathrm{e}^{-\frac{b(l_1-ut)}{v}}\mathrm{e}^{-\frac{\beta l_2}{v}} \tag{7}$$

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 $\Delta t\to 0$,并将(5),(6)式代入得

$$\begin{cases}\dfrac{\partial w}{\partial t}=\dfrac{abu}{v}w(ut,t)\,\mathrm{e}^{-\frac{b(x-ut)}{v}}\\ w(x,0)=w_0\end{cases}\tag{8}$$

方程(8)的解为(推导见[42]).

$$\begin{cases}w(x,t)=w_0\left[1+\dfrac{a}{a'}\mathrm{e}^{-\frac{bx}{v}}\left(\mathrm{e}^{\frac{but}{v}}-\mathrm{e}^{\frac{abut}{v}}\right)\right]\\ w(ut,t)=\dfrac{w_0}{a'}\left(1-a\mathrm{e}^{-\frac{a'but}{v}}\right)\end{cases}\tag{9}$$

其中 $a'=1-a$(假设 2).

4. 计算 Q.

将(9)代入(7)式,得

$$q(l,t)=\frac{auw_0}{a'}\mathrm{e}^{-\frac{bl_1}{v}}\mathrm{e}^{-\frac{\beta l_2}{v}}\left(\mathrm{e}^{-\frac{but}{v}}-a\mathrm{e}^{-\frac{abut}{v}}\right)\tag{10}$$

最后将(10)代入(1)式,作积分得到

$$Q=\int_0^{l_1/u}q(l,t)\,\mathrm{d}t=\frac{aw_0v}{a'b}\mathrm{e}^{-\frac{\beta l_2}{v}}\left(1-\mathrm{e}^{-\frac{a'bl_1}{v}}\right)\tag{11}$$

为便于下面的分析,将上式化作

$$Q=aM\mathrm{e}^{-\frac{\beta l_2}{v}}\cdot\frac{1-\mathrm{e}^{-\frac{a'bl_1}{v}}}{\frac{a'bl_1}{v}}\tag{12}$$

记

$$r=\frac{a'bl_1}{v},\quad \varphi(r)=\frac{1-\mathrm{e}^{-r}}{r}\tag{13}$$

则(12)式可写作

$$Q=aM\mathrm{e}^{-\frac{\beta l_2}{v}}\varphi(r)\tag{14}$$

(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,β,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.

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.

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 $r=\dfrac{a'bl_1}{v}\ll 1$，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

$$Q\approx aMe^{-\frac{\beta l_2}{v}}\left(1-\frac{a'bl_1}{2v}\right) \tag{15}$$

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.

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

$$Q_2=\frac{aw_0v}{a'b}e^{-\frac{bl_2}{v}}\left(1-e^{-\frac{a'bl_1}{v}}\right) \tag{16}$$

与(11)式给出的 Q_1 相比，我们得到

$$\frac{Q_1}{Q_2}=e^{-\frac{(\beta-b)l_2}{v}} \tag{17}$$

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.

复习题

在香烟过滤嘴模型中，

(1) 设 $M=800$ mg，$l_1=80$ mm，$l_2=20$ mm，$b=0.02$ s^{-1}，$\beta=0.08$ s^{-1}，$v=50$ mm/s，$a=0.3$，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.

5.6 火箭发射升空

“10，9，8，…，3，2，1，点火”，随着指令员的口令，巨大火箭的尾端喷出炽烈的

火焰,在震天动地的轰鸣声中,搭载着宇宙飞船拔地而起,冉冉上升."助推器分离""一二级分离""船箭分离",随着第一级火箭关机、分离,第二级、第三级火箭点火、关机、分离等步骤,火箭不断加速飞行,达到预定的高度和速度,将飞船送上地球引力作用下的惯性运行轨道,运载火箭完成了它的使命.一次次在屏幕上目睹这绚丽、壮观的情景,我们注意到火箭是垂直地面发射的,这样才能以很短的距离穿越稠密的大气层,尽量减少空气的阻力;三级火箭接力棒式地助推,把燃料耗尽的火箭结构残骸一一丢弃,以便在产生巨大推力的同时,减轻火箭本身的质量,得到尽可能大的有效载荷.本节建立的是单级小型火箭发射、上升过程的数学模型,并由此讨论提高火箭上升高度的办法[70].

单级小型火箭的发射

考察火箭垂直于地面发射、上升的过程:火箭垂直向上发射后,燃料以一定的速率燃烧,火焰向后喷射,对火箭产生向前的推力,在克服地球引力和空气阻力的同时,推动火箭加速飞行.燃料燃尽后火箭依靠获得的速度继续上升,但在引力和阻力的作用下速度逐渐减小,直到速度等于零,火箭达到最高点.建立数学模型研究火箭上升高度、速度和加速度的变化规律以及与火箭质量、燃料推力等因素的关系.

显然,这个虚拟的场景只是实际发射火箭最初的一个阶段,若不施加其他手段,火箭将从最高点自由下落.

火箭发射、上升过程所遵循的基本规律是牛顿第二定律,火箭在运动中受到燃料燃烧的推力、地球引力和空气阻力的联合作用,其中推力与引力的作用是明确和易于处理的,空气阻力随着火箭速度的增加而变大,但二者之间的数量关系则不易确定.下面先不考虑空气阻力,建立相对简单的模型,再通过简化、合理的假定将阻力引入模型.

不考虑空气阻力的简单模型

问题与假设　设小型火箭初始质量为 $m_0=1\ 600$ kg,其中包括 $m_1=1\ 080$ kg 燃料.火箭从地面垂直向上发射时,燃料以 $r=18$ kg/s 的速率燃烧,对火箭产生 $F=27\ 000$ N 的恒定推力.燃料燃尽后火箭继续上升,直到达到最高点.因为火箭上升高度与地球半径相比很小,所以可认为整个过程中受到的地球引力不变,重力加速度取 $g=9.8\ \mathrm{m/s^2}$.建立火箭上升高度、速度和加速度随时间变化的数学模型,给出燃料燃尽时火箭的高度、速度和加速度以及火箭到达最高点的时间和高度.

模型建立　设火箭 $t=0$ 时从地面 $x=0$ 向上发射,火箭高度为 $x(t)$,速度为 $\dot{x}(t)$,加速度为 $\ddot{x}(t)$.火箭的质量记为 $m(t)$,随燃料燃烧而减少,有 $m(t)=m_0-rt$.燃料燃尽的时间记为 t_1,显然 $t_1=\dfrac{m_1}{r}=60$ s,t_1 以后火箭质量保持为 $m_0-m_1=520$ kg.火箭到达最高点的时间记为 t_2,t_2 由 $v(t_2)=0$ 确定.

火箭从 $x=0$ 处零初速地发射,上升过程中燃料燃烧阶段受到推力 F 和重力 $m(t)g$ 的作用,按照牛顿第二定律可以写出

$$(m_0-rt)\ddot{x}=F-(m_0-rt)g,0\leqslant t\leqslant t_1,t_1=\frac{m_1}{r},\quad x(0)=\dot{x}(0)=0 \tag{1}$$

燃料燃尽后火箭只受重力作用,于是

$$(m_0-m_1)\ddot{x}=-(m_0-m_1)g,\quad t_1<t\leqslant t_2 \tag{2}$$

方程(2)的初始条件由(1)的结果 $x(t_1)$,$\dot{x}(t_1)$ 给出.

模型求解 (1)式虽然表示为 $x(t)$ 的二阶微分方程的形式,但是由于方程不含 x 和 $\dot{x}$,可以用积分法直接求解.将加速度记作 $a(t)$,速度记作 $v(t)$,方程(1)可写作

$$a(t)=-g+\frac{F}{m_0-rt} \tag{3}$$

对(3)式积分并利用 $v(0)=0$,得到

$$v(t)=\int_0^t a(t)\,\mathrm{d}t=-gt+\frac{F}{r}\ln\frac{m_0}{m_0-rt} \tag{4}$$

对(4)式积分并利用 $x(0)=0$,得到

$$\begin{aligned}x(t)&=\int_0^t v(t)\,\mathrm{d}t\\&=-\frac{gt^2}{2}+\frac{F(m_0-rt)[\ln(m_0-rt)-1]}{r^2}+\frac{(F\ln m_0)t}{r}-\frac{Fm_0[\ln m_0-1]}{r^2}\end{aligned} \tag{5}$$

由(3)、(4)、(5)式可计算燃料燃烧阶段($0\leqslant t\leqslant t_1$)的加速度、速度和高度的变化,并得到 $a(t_1)$, $v(t_1)$ 和 $x(t_1)$.

方程(2)可写作

$$a(t)=-g \tag{6}$$

对(6)式积分并利用 $v(t_1)$,得到

$$v(t)=\int_{t_1}^t a(t)\,\mathrm{d}t=-g(t-t_1)+v(t_1) \tag{7}$$

对(7)式积分并利用 $x(t_1)$,得到

$$x(t)=\int_{t_1}^t v(t)\,\mathrm{d}t=-\frac{g(t-t_1)^2}{2}+v(t_1)(t-t_1)+x(t_1) \tag{8}$$

由(6)、(7)、(8)式可计算燃料燃尽后($t_1<t\leqslant t_2$)的加速度、速度和高度的变化.在(7)式中令 $v(t_2)=0$ 可得火箭到达最高点的时间 t_2,再由(8)式得到最高点的高度 $x(t_2)$.

将问题中给出的数据代入(3)至(8)式编程计算,结果见图 1,燃料燃尽时($t_1=60$ s)火箭的高度 $x(t_1)=2.365\,6\times10^4$ m,速度 $v(t_1)=1\,098$ m/s,加速度 $a(t_1)=42.12$ m/s^2,火箭到达最高点的时间 $t_2=172$ s,高度 $x(t_2)=8.515\,5\times10^4$ m.

可以看出,燃料燃烧阶段($0\leqslant t\leqslant t_1$)火箭的高度 $x(t)$、速度 $v(t)$ 和加速度 $a(t)$ 都在快速地增大,在燃料燃尽后($t_1<t\leqslant t_2$),$a(t)$ 突降并保持为 $-g$,$v(t)$ 呈直线下降,$x(t)$ 则上升减缓.

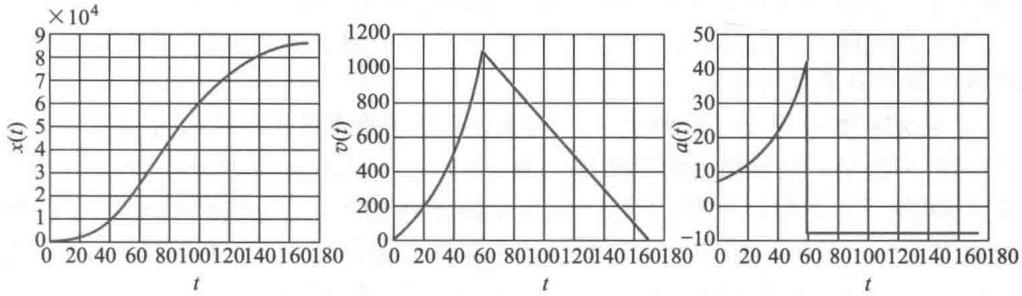


图 1 不考虑空气阻力时火箭上升的高度 $x(t)$、速度 $v(t)$ 和加速度 $a(t)$

考虑空气阻力的模型

运动的物体都会受到空气的阻力，运动速度越大阻力就越大，但二者之间没有确定的数量关系.按照相关知识和经验，低速时阻力与速度成正比，高速时阻力与速度的平方甚至三次方成正比.对于小型火箭，通常可设阻力与速度的平方成正比，比例系数记作 k，于是在考虑空气阻力情况下方程(1)、(2)应改写为

$$(m_0-rt)\ddot{x}=F-k\dot{x}^2-(m_0-rt)g,\ 0\leqslant t\leqslant t_1,t_1=m_1/r,\ x(0)=\dot{x}(0)=0 \tag{9}$$

$$(m_0-m_1)\ddot{x}=-k\dot{x}^2-(m_0-m_1)g,\ t_1<t\leqslant t_2 \tag{10}$$

方程(10)的初始条件由(9)的结果 $x(t_1)$，$\dot{x}(t_1)$ 给出.

方程(9)无法解析地求解，转而求数值解.设方程(1)中已有的数据不变，阻力系数一般为 0.2 kg/m 至 0.4 kg/m，不妨取 $k=0.3$ kg/m.需要注意的是，用 MATLAB 软件求解高阶微分方程时必须先转化为方程组，设 $x=x_1$，$\dot{x}=x_2$，方程(9)化为

$$\dot{x}_1=x_2,\dot{x}_2=-g+\frac{F-kx_2^2}{m_0-rt},\ 0\leqslant t\leqslant t_1,t_1=m_1/r,x_1(0)=x_2(0)=0 \tag{11}$$

方程(10)可类似地计算.对方程(9)、(10)编程计算的结果见图 2，燃料燃尽时($t_1=60$ s)火箭的高度 $x(t_1)=1.054\,6\times10^4$ m，速度 $v(t_1)=266$ m/s，加速度 $a(t_1)=1.296$ m/s^2，火箭到达最高点的时间 $t_2=75$ s，高度 $x(t_2)=1.196\,9\times10^4$ m.

对比图 2 与图 1，抛掉数值上的差别(这与给定的数据有关)，我们看到，燃料燃烧阶段火箭的加速度已经由增变减，速度增加渐缓，致使上升高度大大减少.

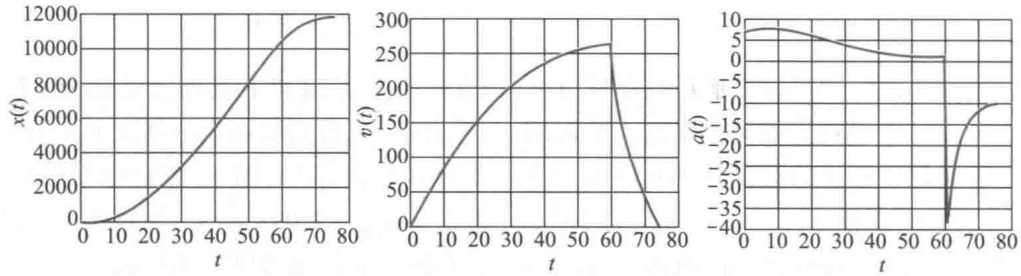


图 2 考虑空气阻力时火箭上升的高度 $x(t)$、速度 $v(t)$ 和加速度 $a(t)$

提升火箭高度的办法 提高上升高度是火箭发射的中心问题之一，按照前面的模型，从火箭燃料方面可以有两种办法实现，一是增加燃料的数量 m_1，从而延长燃烧时间 t_1，二是改进燃料的效能，提高产生的恒定推力 F.下面通过实例计算，考察这两种办法对提高火箭上升高度的效果.

1) 将燃料数量增加一倍即 $m_1=2\,160$ kg，火箭初始质量相应地增为 $m_0=2\,680$ kg，其他参数不变.于是燃料燃尽的时间延长为 $t_1=m_1/r=120$ s，按照方程(9)、(10)编程计算，结果见图 3 中标记#1 的曲线，及表 1 中标记#1 的数值，与原来的结果(标记#0 的曲线和表 1 中标记#0 的数值)比较，虽然燃料燃尽时的速度 $v(t_1)$ 和加速度 $a(t_1)$ 都相同，但高度 $x(t_1)$ 及火箭到达最高点的高度 $x(t_2)$ 均有较大提高.

2) 将燃料产生的推力增加一倍即 $F=54\,000$ N，其他参数不变.燃料燃尽的时间仍为 $t_1=60$ s，按照方程(9)、(10)编程计算，结果见图 4 中标记#2 的曲线，及表 1 中

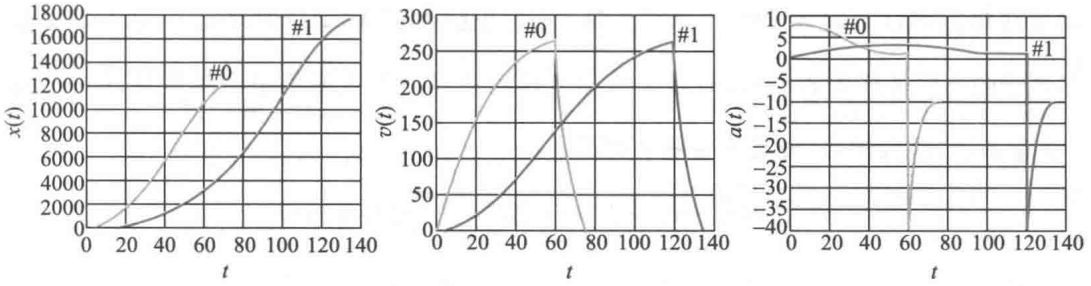


图 3 燃料数量增加一倍时火箭上升的高度 $x(t)$、速度 $v(t)$ 和加速度 $a(t)$(标记#1)

标记#2 的数值,与原来的结果比较,燃料燃尽时的加速度$a(t_1)$减小,速度 $v(t_1)$增大,高度 $x(t_1)$及火箭到达最高点的高度 $x(t_2)$均有很大提高.

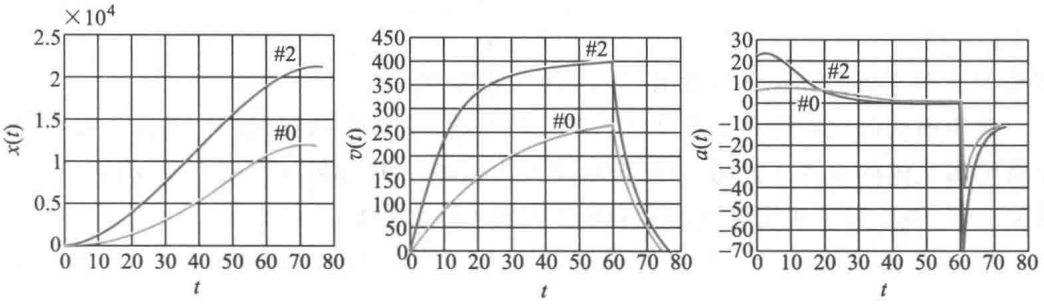


图 4 燃料推力增加一倍时火箭上升的高度 $x(t)$、速度 $v(t)$ 和加速度 $a(t)$(标记#2)

将燃料数量和产生的推力同时增加一倍的计算结果见表 1 中标记#3 的数值.

表 1 将燃料数量和推力增加一倍时的计算结果

		t_1/s	$x(t_1)$/(10^4m)	$v(t_1)$/(m·s^{-1})	$a(t_1)$/(m·s^{-2})	t_2/s	$x(t_2)$/(10^4m)
#0	m_1=1 080 kg F=27 000 N	60	1.054 6	266	1.30	75	1.196 9
#1	m_1=2 160 kg F=27 000 N	120	1.598 3	266	1.30	135	1.740 5
#2	m_1=1 080 kg F=54 000 N	60	1.933 4	402	0.80	77	2.137 3
#3	m_1=2 160 kg F=54 000 N	120	3.660 7	402	0.80	137	3.864 6

下面讨论火箭燃料燃烧产生的推力是如何确定的.仍记火箭上升过程中的质量为 $m(t)$,速度为 $v(t)$,燃料燃烧时向后喷出气体产生对火箭的推力,记喷出气体相对于火

箭的速度为 u,则相对地球的(绝对)速度为 $v(t)-u$.在忽略引力和阻力的情况下,根据动量守恒定律,t 到 $t+\Delta t$ 时间内火箭质量、速度变化所引起的动量的改变,应等于 Δt 内燃料燃烧喷出气体产生的动量,而因为喷出气体的质量就是火箭质量的减少量,即 $-\frac{dm}{dt}\Delta t$(注意 $m(t)$ 是减函数,导数为负值),所以

$$m(t)v(t)-m(t+\Delta t)v(t+\Delta t)=-\frac{dm}{dt}\Delta t(v(t)-u) \tag{12}$$

两端除以 Δt 且令 $\Delta t\to 0$,再利用 $\frac{d}{dt}(m(t)v(t))=\frac{dm}{dt}v(t)+m(t)\frac{dv}{dt}$,由(12)式可得

$$m(t)\frac{dv}{dt}=-u\frac{dm}{dt} \tag{13}$$

根据牛顿第二定律,这就是火箭的推力 F.当燃料以不变的速率 r 燃烧时,$\frac{dm}{dt}=-r$,于是

$$F=ru \tag{14}$$

即推力 F 等于燃烧速率 r 与喷出气体速度 u 的乘积.

对于前面讨论的提高火箭高度的第二种办法(表1中的标记#2),在 r 不变的情况下将喷出气体的速度 u 提高一倍,即可使推力增加一倍.如果 u 保持不变,将燃烧速率 r 提高一倍,对于提高火箭高度的影响有多大呢(复习题)?

将(13)式左端的 $m(t)$ 移到右端作积分,设 $v(0)=0$,$m(0)=m_0$,可得

$$v(t)=u\ln\frac{m_0}{m(t)} \tag{15}$$

与(4)式的结果一致(忽略引力).(15)式表明增加喷出气体速度 u 和减少质量 $m(t)$ 可以提高火箭速度,而要减少 $m(t)$,一种办法是上面所说的提高燃烧速率 r,但提高是有限的;另一种办法是分级携带燃料,每一级的燃料一旦燃尽就将这一级的结构部分抛弃,以尽量减轻火箭的质量,这就是构造多级火箭的理由.那为什么常用三级火箭呢?有兴趣的读者可参考文献[76].

评注 本节利用物理定律建立的火箭发射、上升过程的数学模型,在一定意义下具有一般性,可以作为微分方程建模的练习,不过所给的数据(特别是阻力系数)都是虚拟的,模型求解的数值结果没有实际意义.

高度 $x(t)$、速度 $v(t)$ 和加速度 $a(t)$ 互为微分、积分的关系,用数值方法得到的3个函数的图形(图1—图4)显示,在 $t=t_1$ 点 $a(t)$ 是间断的,$v(t)$ 连续但为尖点(不可导),$x(t)$ 则是光滑的(可导),对这些现象的观察有助于微积分概念的理解.

复习题

在考虑空气阻力的火箭发射模型中,如果将燃料燃烧速率 r 提高一倍,而喷出气体速度 u 保持不变,在以下两种情况下讨论对于提高火箭高度的影响:

(1) 燃料质量 m_1 不变,燃料燃尽时间 t_1 减半.

(2) 燃料质量 m_1 也增加一倍,燃料燃尽时间 t_1 不变.

5.7 食饵与捕食者模型

自然界中不同种群之间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[11,39,46,49].

基础知识 5-1 微分方程稳定性简介

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少,设减少率与捕食者数量成正比,于是 $x(t)$ 满足方程

$$\dot{x}(t)=x(r-ay)=rx-axy \tag{1}$$

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.设增长率与食饵数量成正比,于是 $y(t)$ 满足

$$\dot{y}(t)=y(-d+bx)=-dy+bxy \tag{2}$$

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

$$x(0)=x_0,y(0)=y_0 \tag{3}$$

为求微分方程(1),(2)满足初始条件(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 软件编程计算,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 $\bar{x}=25,\bar{y}=10$.

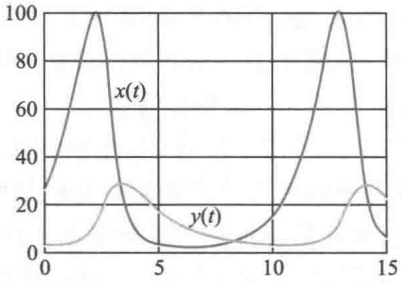


图 1 数值解 $x(t)$,$y(t)$的图形

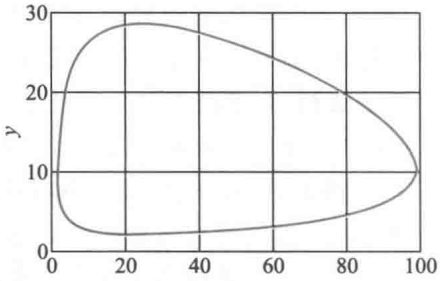


图 2 相轨线 $y(x)$的图形

2. 平衡点及相轨线

首先求得方程(1),(2)的两个平衡点为

$$P\left(\frac{d}{b},\frac{r}{a}\right),P'(0,0) \tag{4}$$

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 $\mathrm{d}t$ 后得到

$$\frac{\mathrm{d}x}{\mathrm{d}y}=\frac{x(r-ay)}{y(-d+bx)} \tag{5}$$

这是可分离变量方程,写作

$$\frac{-d+bx}{x}\mathrm{d}x=\frac{r-ay}{y}\mathrm{d}y \tag{6}$$

两边积分得到方程(5)的解,即方程(1),(2)的相轨线为

$$(x^d\mathrm{e}^{-bx})(y^r\mathrm{e}^{-ay})=c \tag{7}$$

其中常数 c 由初始条件确定.

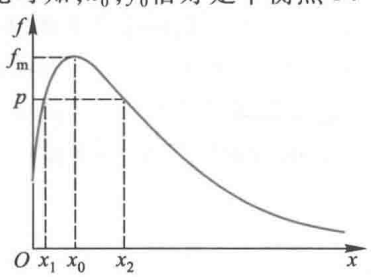
为了从理论上证明相轨线(7)是封闭曲线,记

$$f(x)=x^d\mathrm{e}^{-bx},\quad g(y)=y^r\mathrm{e}^{-ay} \tag{8}$$

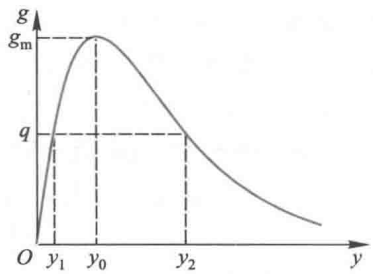
可以用软件做出它们的图形(图 3(a),(b)),将它们的极值点记为 x_0,y_0,极大值记为 f_m,g_m,则不难知道 x_0,y_0 满足

$$\begin{aligned} f(x_0)&=f_m,\quad x_0=d/b\\ g(y_0)&=g_m,\quad y_0=r/a \end{aligned} \tag{9}$$

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 $c=pg_m(0<p<f_m)$.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 $f(x)g(y)=pg_m$ 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.

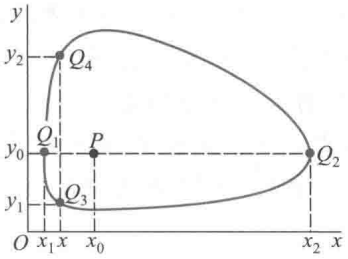


图 4 相轨线的图形

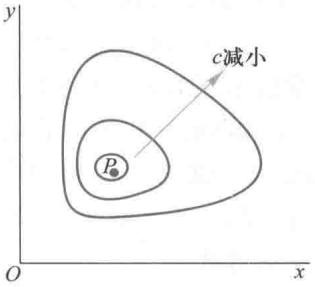


图 5 相轨线族

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.

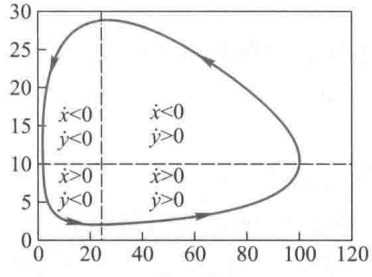


图 6 相轨线的方向

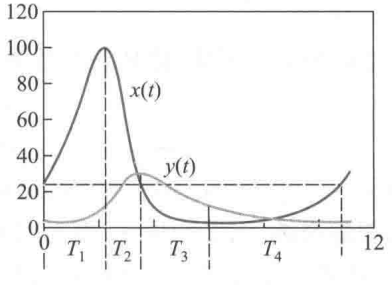


图 7 $x(t)$,$y(t)$ 的相位差

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

$$x(t)=\frac{1}{b}\left(\frac{\dot{y}}{y}+d\right) \tag{10}$$

(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

$$\bar{x}=\frac{1}{T}\int_0^T x(t)\,\mathrm{d}t=\frac{1}{T}\left[\frac{\ln y(T)-\ln y(0)}{b}+\frac{dT}{b}\right]=\frac{d}{b} \tag{11}$$

类似地可得

$$\bar{y}=\frac{r}{a} \tag{12}$$

将(11),(12)与(9)比较可知

$$\bar{x}=x_0,\quad \bar{y}=y_0 \tag{13}$$

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

$$\bar{x}_1=\frac{d+e}{b},\quad \bar{y}_1=\frac{r-e}{a} \tag{14}$$

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

$$\bar{x}_2=\frac{d+e'}{b},\quad \bar{y}_2=\frac{r-e'}{a} \tag{15}$$

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做出似乎出人意料的解释.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振荡,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,用方程

$$\dot{x}_1(t)=r_1x_1\left(1-\frac{x_1}{N_1}-\sigma_1\frac{x_2}{N_2}\right) \tag{16}$$

$$\dot{x}_2(t)=r_2x_2\left(-1+\sigma_2\frac{x_1}{N_1}-\frac{x_2}{N_2}\right) \tag{17}$$

就可以描述这种现象,其中各个参数的含义可参看更多案例 5-6,5-7.

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.

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.

复习题

1. 在食饵-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.

2. 如果在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.

5.8 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做出最恰当的安排,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[36].

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,为冲力做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,二是冲力做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,氧的新陈代谢作用在单位时间提供的能量,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做出假设,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做出假设,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 $m=1$ 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

$$D(v(t))=\int_0^T v(t)\,\mathrm{d}t \tag{1}$$

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

$$\dot{v}(t)+\frac{v}{\tau}=f(t) \tag{2}$$

$$v(0)=0 \tag{3}$$

$$0\leqslant f(t)\leqslant F \tag{4}$$

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

$$\dot{E}(t)=\sigma-fv \tag{5}$$

$$E(0)=E_0 \tag{6}$$

$$E(t)\geqslant 0 \tag{7}$$

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),(3),(5)式合并后得到的泛函表达式

$$D(E(t))=\sqrt{2}\int_0^T \mathrm{e}^{-\frac{t}{\tau}}\left[\int_0^t (\sigma-\dot{E}(s))\mathrm{e}^{\frac{2s}{\tau}}\mathrm{d}s\right]^{\frac{1}{2}}\mathrm{d}t \tag{8}$$

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

$$v(t)=F\tau\left(1-\mathrm{e}^{-\frac{t}{\tau}}\right) \tag{9}$$

可知速度是递增的.将(9)及 $f=F$ 代入(5)式得

$$\dot{E}(t)=\sigma-F^2\tau\left(1-\mathrm{e}^{-\frac{t}{\tau}}\right) \tag{10}$$

方程(10)在初始条件(6)下的解为

$$E(t)=E_0-(F^2\tau-\sigma)t+F^2\tau^2\left(1-\mathrm{e}^{-\frac{t}{\tau}}\right) \tag{11}$$

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 $0\leqslant t\leqslant t_e$ 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 $t=t_c$ 时 $E(t)=0$.容易得到 $t_c=t\big|_{E(t)=0}$，

$$t_e=t\Big|_{\dot{E}(t)=0}=\tau\ln\frac{F^2\tau}{F^2\tau-\sigma}$$

并且，在这种情况下所能跑的最远距离为

$$D_c=\int_0^{t_c}v(t)\,\mathrm{d}t=F\tau^2\left(\mathrm{e}^{-\frac{t_c}{\tau}}+\frac{t_c}{\tau}-1\right) \tag{12}$$

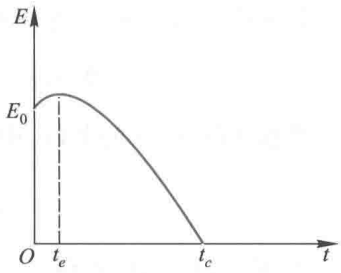


图 1 体内能量 $E(t)$ 的示意图

为了估计模型中的参数，并算出 t_c 和 D_c，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 $F\approx 12.2$ N/kg(对于质量 $m=1$ kg 而言)，$\tau\approx 0.892$ s，算出 $t_c\approx 27.63$ s，$D_c\approx 291$ 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).

后来，有人用 1987 年世界田径锦标赛上 Johnson 和 Lewis 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 $x(t)=F\tau[t-\tau(1-\mathrm{e}^{-t/\tau})]$，当 $t\gg\tau$ 时 $x\approx v_m(t-\tau)$.从表 1 可得 $v_m=11.76$ m/s，$\tau\approx 1.28$ s 或 1.42 s，考虑到起跑的反应时间，取 $\tau\approx 1.06$ s 或 1.16 s，而 $F\approx 11.1$ N/kg 或 10.1 N/kg.

表 1 Johnson 的成绩 t_1 和 Lewis 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

中长跑模型 当赛程大于 D_c 时，将全程分为 3 段：初始阶段 $0\leqslant t\leqslant t_1$ 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 $t_2\leqslant t\leqslant T$ 把体内贮存能量

用完，即$E(t)=0$，靠惯性冲刺；中间阶段$t_1\leqslant t\leqslant t_2$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 $0\leqslant t\leqslant t_1$($t_1$以后确定)　$v_1(t)$由(9)式表示(将$v(t)$写作$v_1(t)$).

第3段 $t_2\leqslant t\leqslant T$($t_2$以后确定)　将$E(t)=0$代入(5)，(2)式得到

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}v_3^2+\frac{v_3^2}{\tau}=\sigma \tag{13}$$

方程(13)在条件$v_3(t_2)=v_2$下的解为

$$v_3(t)=\left[(v_2^2-\sigma\tau)\mathrm{e}^{-\frac{2(t-t_2)}{\tau}}+\sigma\tau\right]^{\frac{1}{2}} \tag{14}$$

$v_3(t)$是单调减和下凸的($\dot{v}_3(t)<0$，$\ddot{v}_3(t)>0$).

第2段 $t_1\leqslant t\leqslant t_2$　v_2由目标泛函(1)式$D(v(t))$达到最大来确定，这里

$$D(v(t))=\int_0^{t_1}v_1(t)\,\mathrm{d}t+v_2(t_2-t_1)+\int_{t_2}^{T}v_3(t)\,\mathrm{d}t \tag{15}$$

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

$$E(t)=E_0+\sigma t-\frac{v^2(t)}{2}-\frac{1}{\tau}\int_0^t v^2(t)\,\mathrm{d}t \tag{16}$$

令$t=t_2$并将(9)式代入得

$$E(t_2)=E_0+\sigma t_2-\frac{v_2^2}{2}-\frac{1}{\tau}\int_0^{t_1}v_1^2(t)\,\mathrm{d}t-\frac{v_2^2(t_2-t_1)}{\tau} \tag{17}$$

问题归结为在条件$E(t_2)=0$和$v_1(t_1)=v_2$下求v_2,t_1,t_2使$D(v(t))$最大.

由于在条件$v_1(t_1)=v_2$下$D(v(t))$和$E(t_2)$对t_1的偏导数均为零，所以在用待定常数λ构造的函数

$$I(v(t),t_2)=D(v(t))+\frac{\lambda}{2}E(t_2) \tag{18}$$

中可以将右端不显含v_2和t_2的项略去，写成

$$I(v_2,t_2)=\int_{t_2}^{T}v_3(t)\,\mathrm{d}t+\frac{\lambda\sigma}{2}t_2-\frac{\lambda}{4}v_2^2+\left(v_2-\frac{\lambda v_2^2}{2\tau}\right)(t_2-t_1) \tag{19}$$

v_2,t_2为最优解的必要条件是

$$v_2=\frac{\tau}{\lambda} \tag{20}$$

$$2\int_{t_2}^{T}\left[(v_2^2-\sigma\tau)\mathrm{e}^{-\frac{2(t-t_2)}{\tau}}+\sigma\tau\right]^{-\frac{1}{2}}\mathrm{e}^{-\frac{2(t-t_2)}{\tau}}\,\mathrm{d}t=\lambda \tag{21}$$

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

$$\lambda F(1-\mathrm{e}^{-\frac{t_1}{\tau}})=1 \tag{22}$$

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些不合适的地方，例如对短跑和中长

① 如果设速度$v_2(t)$，泛函问题的解也是常数v_2，详见[36].

将(20)代入(17)后令其为0,得

$$E_0+\sigma t_2-\frac{\tau^2}{2\lambda^2}-\int_0^{t_1}F^2\tau(1-e^{-t/\tau})^2dt-\frac{\tau}{\lambda^2}(t_2-t_1)=0 \tag{23}$$

将(20)代入(21)并作积分得到

$$2[(\tau^2-\lambda^2\sigma\tau)e^{-\frac{2(T-t_2)}{\tau}}+\lambda^2\sigma\tau]^{\frac{1}{2}}-2\tau=\lambda^2\sigma-\tau \tag{24}$$

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:$\sigma=41.5$,$E_0=2\,403.5$.

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.

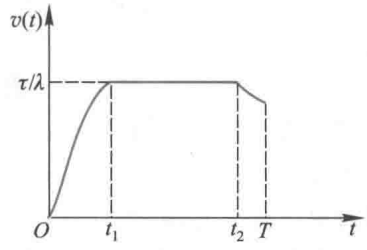


图2 中长跑模型的最佳速度示意图

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(复习题).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做出修正,并且考虑空气阻力、海拔高度等因素的影响[52].

复习题

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 $\dot{f}(t)/f(t)=-1/k$,其中 k 是冲力限制系数,$f(0)=F$ 为最大冲力.将上述关系代入(2)式,求出短跑比赛时速度 $v(t)$ 和距离 $s(t)$ 的关系式,及达到最高速度的时间,并做出 $v(t)$ 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/(m·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,并与实际数据比较.

5.9 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作是历史上最著名的数学模型之一.

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.哥伦布发现的新大陆、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做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°=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),以太阳为极点$r=0$,以椭圆长半轴方向为$\theta=0$,用向径[①]$\boldsymbol{r}(t)(r(t),\theta(t))$表示行星的位置,见图1.

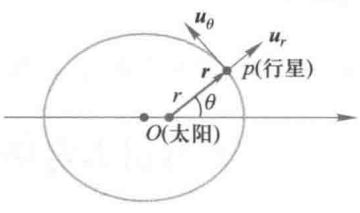


图1 极坐标系中的行星轨道

1. 轨道方程为

$$r=\frac{p}{1+e\cos\theta},\quad p=\frac{b^2}{a},\quad b^2=a^2(1-e^2) \tag{1}$$

① 本节用粗体字母表示向量,以区别于标量.

a,b 为椭圆的长、短半轴，e 为离心率.

2. 单位时间内向径 $\boldsymbol{r}$ 扫过的面积是常数 A，即

$$\frac{1}{2}r^2\dot{\theta}=A \tag{2}$$

3. 行星运行周期 T 满足

$$T^2=\lambda a^3 \tag{3}$$

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 $\boldsymbol{f}$ 等于行星加速度 $\ddot{\boldsymbol{r}}$ 和质量 m 的乘积，即

$$\boldsymbol{f}=m\ddot{\boldsymbol{r}} \tag{4}$$

模型建立 万有引力定律研究的是力 $\boldsymbol{f}$ 的方向和大小.由(4)式可知，它取决于行星的运动加速度 $\ddot{\boldsymbol{r}}$，从而必须研究行星的位置 $\boldsymbol{r}$ 和速度 $\dot{\boldsymbol{r}}$.首先引入基向量(图 1)

$$\begin{cases}\boldsymbol{u}_r=\cos\theta\boldsymbol{i}+\sin\theta\boldsymbol{j}\\ \boldsymbol{u}_\theta=-\sin\theta\boldsymbol{i}+\cos\theta\boldsymbol{j}\end{cases} \tag{5}$$

向径 $\boldsymbol{r}$ 可表为

$$\boldsymbol{r}=r\boldsymbol{u}_r \tag{6}$$

因为由(5)式可以算出

$$\begin{cases}\dot{\boldsymbol{u}}_r=\dot{\theta}\boldsymbol{u}_\theta\\ \dot{\boldsymbol{u}}_\theta=-\dot{\theta}\boldsymbol{u}_r\end{cases} \tag{7}$$

所以由(6)，(7)式得到行星运动的速度和加速度为

$$\dot{\boldsymbol{r}}=\dot{r}\boldsymbol{u}_r+r\dot{\theta}\boldsymbol{u}_\theta \tag{8}$$

$$\ddot{\boldsymbol{r}}=(\ddot{r}-r\dot{\theta}^2)\boldsymbol{u}_r+(r\ddot{\theta}+2\dot{r}\dot{\theta})\boldsymbol{u}_\theta \tag{9}$$

根据(2)式可得

$$\dot{\theta}=\frac{2A}{r^2},\quad \ddot{\theta}=\frac{-4A\dot{r}}{r^3} \tag{10}$$

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

$$\ddot{\boldsymbol{r}}=(\ddot{r}-r\dot{\theta}^2)\boldsymbol{u}_r \tag{11}$$

对(1)式求导并利用(10)式 $\dot{\theta}$ 的结果，得到

$$\dot{r}=\frac{2Ae}{p}\sin\theta \tag{12}$$

$$\ddot{r}=\frac{4A^2e}{pr^2}\cos\theta=\frac{4A^2(p-r)}{pr^3} \tag{13}$$

将(10)和(13)代入(11)式，得

$$\ddot{\boldsymbol{r}}=\frac{-4A^2}{pr^2}\boldsymbol{u}_r \tag{14}$$

最后把(14)和(6)代入(4)式，有

$$\boldsymbol{f}=-\frac{4A^2m}{pr^2}\boldsymbol{r}_0,\quad \boldsymbol{r}_0=\frac{\boldsymbol{r}}{r} \tag{15}$$

这里 $\boldsymbol{r}_0$ 是单位向径，指示向径方向.(15)式表明，行星运行时受的力 $\boldsymbol{f}$ 的方向与它的向

径方向 $\boldsymbol{r}_0$ 相反，即在太阳—行星连线方向，指向太阳；$\boldsymbol{f}$ 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.$\boldsymbol{f}$ 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

$$TA=\pi ab \tag{16}$$

由(1)，(3)，(16)式容易算出

$$\frac{A^2}{p}=\frac{\pi^2}{\lambda} \tag{17}$$

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.

π 和 λ 均为绝对常数.将(17)代入(15)式有

$$\boldsymbol{f}=-\frac{4\pi^2 m}{\lambda r^2}\boldsymbol{r}_0 \tag{18}$$

(18)式表明，太阳对行星的作用力 $\boldsymbol{f}$ 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

$$\boldsymbol{f}=-k\frac{Mm}{r^2}\boldsymbol{r}_0 \tag{19}$$

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(复习题).

复习题

将万有引力定律(18)式与(19)式进行比较，查询太阳质量、地球运行轨道(椭圆)的长半轴、万有引力常数等数据，说明二者是一致的.

5.10 传染病模型和SARS的传播

2002年冬到2003年春，一种名为SARS(severe acute respiratory syndrome，严重急性呼吸道综合征，民间俗称非典)的传染病肆虐全球.SARS首发于中国广东，迅速扩散到东南亚乃至30多个国家和地区，包括医务人员在内的多名患者死亡，引起社会恐慌、媒体关注，以及各国政府和联合国、世界卫生组织的高度重视、积极应对，直至最终控制住疫情的蔓延.突如其来的SARS及其迅猛的传播和严重的后果，成为21世纪初的第一大社会热点事件.

SARS被控制住不久，2003年的全国大学生数学建模竞赛就以“SARS的传播”命名当年的A题和C题，题目中要求“建立你们自己的模型；特别要说明怎样才能建立一个真正能够预测以及能为预防和控制提供可靠、足够的信息的模型，这样做的困难在哪

里？对于卫生部门所采取的措施做出评论，如：提前或延后 5 天采取严格的隔离措施，对疫情传播所造成的影响做出估计."题目还给出北京市从 2003 年 4 月 20 日至 6 月 23 日逐日的疫情数据供参考(表 1 是部分数据).

表 1 北京市的疫情数据

日期	已确诊病例累计	现有疑似病例	死亡累计	治愈出院累计
4 月 20 日	339	402	18	33
4 月 21 日	482	610	25	43
4 月 22 日	588	666	28	46
⋮	⋮	⋮	⋮	⋮
6 月 22 日	2521	2	191	2257
6 月 23 日	2521	2	191	2277



数据文件 5-1
SARS 的传播

本节先介绍数学医学领域中基本的传染病模型，然后结合赛题，介绍几个描述、分析 SARS 传播过程的模型及求解结果.

基本的传染病模型　虽然不同类型传染病的传播过程在医学上有其各自的特点，但是可以从数学建模的角度，按照一般的传播机理建立以下几种基本模型[11,12,38].

SI 模型　这个模型将人群分为两类：易感染者(susceptible)和已感染者(infective)，分别简称健康人和患者.在传染病模型中取两个词的英文字头称为 SI 模型.设所考察地区的总人数(记作 N)不变，时刻 t 健康人和患者在总人数中的比例分别记作 $s(t)$ 和 $i(t)$，显然有 $s(t)+i(t)=1$.

假设每个患者每天有效接触的人数是常数 λ，称接触率，且当健康人被有效接触后立即被感染成为患者，λ 也称感染率.

根据上述假设，每个患者每天有效接触的健康人数是 $\lambda s(t)$，全部患者 Ni 每天有效接触的健康人数是 $N\lambda s(t)i(t)$，这些健康人立即被感染成为患者，于是患者比例 $i(t)$ 满足微分方程(约去方程两端的 N)

$$\frac{\mathrm{d}i}{\mathrm{d}t}=\lambda si \tag{1}$$

将 $s=1-i$ 代入，并记初始时刻的患者比例为 i_0，得

$$\frac{\mathrm{d}i}{\mathrm{d}t}=\lambda i(1-i),\ i(0)=i_0 \tag{2}$$

这是我们熟悉的 logistic 模型(参见 5.1 节).

logistic 方程(2)的解为 S 形曲线，患者比例 $i(t)$ 从 i_0 迅速上升，通过曲线的拐点后上升变缓，当 $t\to\infty$ 时 $i\to 1$，即所有健康人终将被感染成为患者.这显然不符合实际情况.究其原因，是 SI 模型只考虑了健康人可以被感染，而没有考虑到患者可以被治愈的情况.下面的模型将增加关于患者被治愈的假设.

SIS 模型　有些传染病如伤风、痢疾等，虽然可以治愈，但愈后基本上没有免疫力，于是患者愈后又变成易感染的健康人.由此将这个模型称为 SIS 模型.

SIS 模型增加的假设条件是：患者每天被治愈的人数比例是常数 μ，称治愈率.

容易知道,增加了这个假设后,方程(1)的右端应该减去每天被治愈的患者数 $N\mu i$,于是有(同样约去方程两端的 N)

$$\frac{\mathrm{d}i}{\mathrm{d}t}=\lambda si-\mu i \tag{3}$$

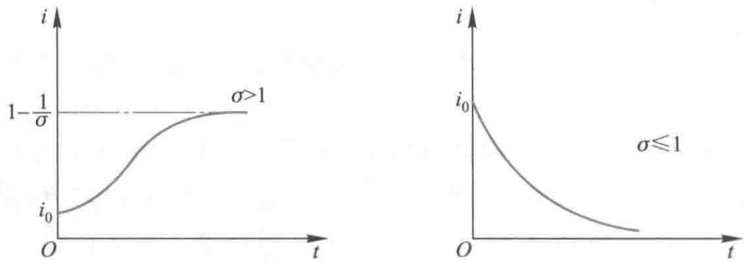
定义

$$\sigma=\lambda/\mu \tag{4}$$

将 $s=1-i$ 和 $\mu=\lambda/\sigma$ 代入(3)式,可得

$$\frac{\mathrm{d}i}{\mathrm{d}t}=\lambda i[(1-1/\sigma)-i] \tag{5}$$

若 $\sigma>1$,(5)式仍然是 logistic 方程,$i(t)$ 呈 S 形曲线上升,当 $t\to\infty$ 时 $i\to 1-1/\sigma$,图形如图 1(a)(通常初始时刻的患者比例 i_0 很小,可设 $i_0<1-1/\sigma$).而若 $\sigma\leqslant 1$,则方程(5)的右端恒为负,曲线 $i(t)$ 将单调下降,当 $t\to\infty$ 时 $i\to 0$,图形如图 1(b).



(a) SIS模型的$i(t)$曲线($\sigma>1$)　　(b) SIS模型的$i(t)$曲线($\sigma\leqslant 1$)

图 1

由此可知,σ 是一个重要参数,$\sigma>1$ 还是 $\sigma\leqslant 1$,决定了患者比例是持续增加,还是持续减少.检查 σ 的定义(4)式,因为 μ 是治愈率,$1/\mu$ 可以视为平均传染期(指患者被治愈所需的平均时间),而 λ 是感染率,所以 σ 表示整个感染期内每个患者有效接触而感染的平均(健康)人数,可称感染数.于是直观上容易理解,若每个患者在生病期间因有效接触而感染的人数大于 1,那么患者比例自然会增加,反之,患者比例会减少.

SIR 模型 许多传染病如天花、流感、麻疹等,治愈后免疫力很强,于是患者愈后不会成为易感染的健康人,当然也不再是患者.传染病模型中将病愈免疫后和因病死亡的人称为移除者(removed),由此将这个模型称为 SIR 模型.仍设总人数 N 不变,时刻 t 移除者在总人数中的比例记作 $r(t)$,有

$$s(t)+i(t)+r(t)=1 \tag{6}$$

关于感染率 λ、治愈率 μ(因包含死亡,实际上是移除率)和感染数 σ 的假设和定义与 SIS 模型相同.容易看出,SIS 模型中患者比例 $i(t)$ 满足的方程(3)不变,即

$$\frac{\mathrm{d}i}{\mathrm{d}t}=\lambda si-\mu i$$

但是由于 SIR 模型中增加了移除者 $r(t)$,在条件(6)下 3 个函数 $s(t)$,$i(t)$ 和 $r(t)$ 中有 2 个是独立的,还需要列出健康人 $s(t)$ 或移除者 $r(t)$ 的微分方程.根据接触率 λ 的假设和定义,$s(t)$ 将单调减少,每天被患者有效接触的健康人数为 $N\lambda s(t)i(t)$,所以 $s(t)$ 满足方程(同样约去方程两端的 N)

$$\frac{ds}{dt}=-\lambda si \tag{7}$$



拓展阅读 5-2
SIR 模型的相轨线分析与群体免疫

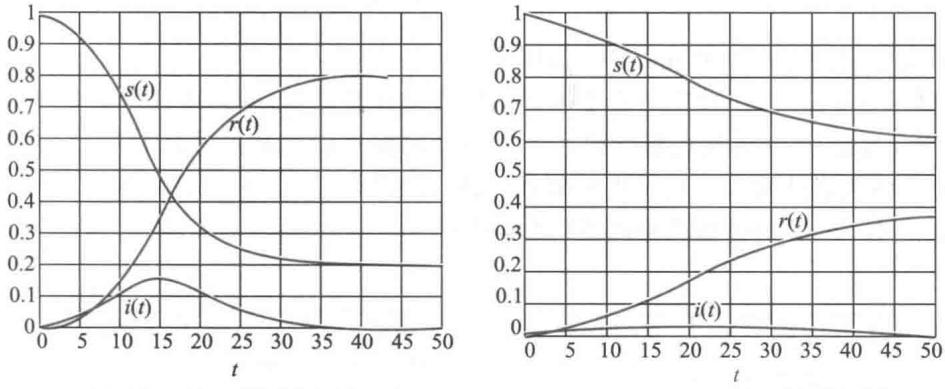
而 $r(t)$ 的方程为

$$\frac{dr}{dt}=\mu i \tag{8}$$

当参数 λ 和 μ 确定后，给定 $s(0)$，$i(0)$，$r(0)$ 中的任意 2 个初始值，即可由上面 3 个微分方程及条件(6)中的任意 3 个式子求解，得到 SIR 模型中 3 类人群的人数比例 $s(t)$，$i(t)$ 和 $r(t)$.但是，这个非线性微分方程组没有解析解(虽然形式上看来很简单)，只能求数值解.图 2 是设定 λ，μ，$s(0)$，$i(0)$ 后用 MATLAB 编程计算得到的 $s(t)$，$i(t)$，$r(t)$ 曲线.

由图 2 可以看到，健康人比例 $s(t)$ 单调减少，移除者比例 $r(t)$ 单调增加，都趋向稳定值，而患者比例 $i(t)$ 先增后减趋向于 0.$t\to\infty$ 时 $s(t)$ 的稳定值表示在传染病传播过程中最终没有被感染的人数比例，$i(t)$ 的最大点和最大值表示传染病高潮(患者最多)到来的时刻和人数比例，这些数值可以衡量传染病传播的强度和速度.

感染率 λ 和治愈率 μ(可以假定死亡率很低)是影响传播过程的重要参数.社会的卫生水平越高感染率 λ 越小，医疗水平越高治愈率 μ 越大，于是感染数 $\sigma(=\lambda/\mu)$ 越小，应该有助于控制传染病的传播.对比不同 σ 下的图 2(a)($\sigma=2$)和图 2(b)($\sigma=1.25$)，后者的健康人比例 $s(t)$ 更大，患者比例 $i(t)$ 更小.



(a) $\lambda=0.6$, $\mu=0.3$, $s(0)=0.99$, $i(0)=0.01$　　(b) $\lambda=0.5$, $\mu=0.4$, $s(0)=0.99$, $i(0)=0.01$

图 2　SIR 模型的 $s(t)$，$i(t)$，$r(t)$

进一步的研究表明，当 $\sigma s(0)>1$ 时 $i(t)$ 先增后减，可认为传染病有一个蔓延过程；而若 $\sigma s(0)\leqslant 1$ 则 $i(t)$ 单调减少，传染病没有蔓延开来.一般情况下 $i(0)$ 和 $r(0)$ 都很小，$s(0)$ 接近 1，所以控制传染病蔓延需要 σ 小于 1，可以通过提高卫生水平和医疗水平达到.

控制传染病蔓延的另一途径是降低 $s(0)$，这可以通过提高 $r(0)$，即预防接种使群体免疫来实现(参考拓展阅读 5-2).

上面所有的模型都假定参数 λ 和 μ 是常数，而在实际的传播过程中，它们会随着预防措施的加强与医疗水平的提高而发生较大变化，下面对于 SARS 传播的讨论就是这种情况.

评注　3 个传染病模型(SI，SIS，SIR)体现了建模过程的不断深化.SI 模型描述了传染病的蔓延，但不符合实际；SIS 和 SIR 模型则针对愈后是否免疫这两种情况，描述了传播过程，得到患者比例的变化规律.SIR 模型特别值得注意，它是研究更复杂、也更实用的许多传染病模型的基础.

SARS的传播 2003年SARS在我国爆发初期，限于卫生部门和公众对这种疾病的认识，使其处于几乎不受制约的自然传播形式，后期随着卫生部门预防和治疗手段的不断加大，SARS的传播受到严格控制.虽然影响这个过程的因素众多，也不只有健康人、患者、移除者3个人群，但是仍然可以用愈后免疫的SIR模型来描述.并且，越复杂的模型包含的参数越多，为确定这些参数所需要的疫情数据就越全面，而实际上能够得到的数据是有限的.下面介绍3个模型.

模型一——参数时变的SIR模型

模型构造 为了计算的方便，在总人数不变的条件下用$s(t)$，$i(t)$，$r(t)$分别表示第t天健康人、患者、移除者(治愈与死亡之和)的数量，$\lambda(t)$，$\mu(t)$分别表示第t天的感染率和移除率(治愈率与死亡率之和).SIR模型表示为如下参数时变方程：

$$\frac{ds}{dt}=-\lambda(t)s(t)i(t) \tag{9}$$

$$\frac{di}{dt}=\lambda(t)s(t)i(t)-\mu(t)i(t) \tag{10}$$

$$\frac{dr}{dt}=\mu(t)i(t) \tag{11}$$

由于s远大于i和r，$s(t)$几乎是常数，所以方程(10)可简化为

$$\frac{di}{dt}=\lambda(t)i(t)-\mu(t)i(t) \tag{12}$$

下面用方程(11)，(12)研究SARS的传播.

参数估计与拟合 首先对表1提供的数据进行预处理：第1列日期从$t=1$开始到$t=65$(天)为止，第4列死亡累计和第5列治愈出院累计之和是移除者数量$r(t)$，第2列已确诊病例累计减去$r(t)$是患者数$i(t)$，原始数据$i(t)$，$r(t)$见图3，可以看出，在$t=25$(天)左右SARS的传播达到高潮(患者最多).

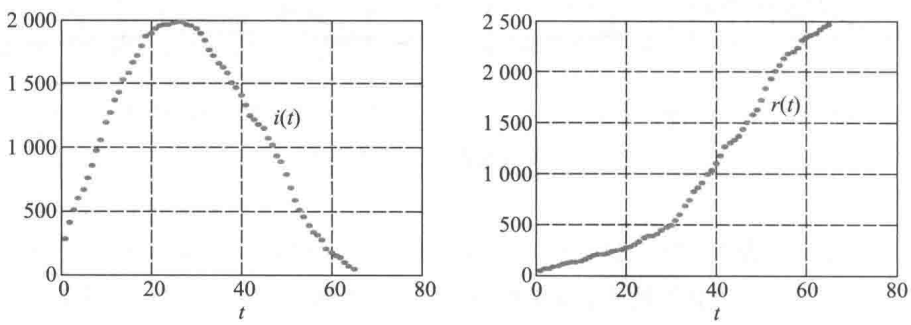


图3 原始数据的患者数$i(t)$和移除数$r(t)$

为了估计感染率$\lambda(t)$和移除率$\mu(t)$，取$i(t)$，$r(t)$的差分$\Delta r(t)$，$\Delta i(t)$作为方程(11)，(12)左端导数的近似值.先用(11)式估计$\mu(t)=\Delta r(t)/i(t)$，再将这个结果代入(12)式得$\lambda(t)=(\Delta i(t)+\Delta r(t))/i(t)$.$\lambda(t)$，$\mu(t)$如图4中的圆点所示，可以看出，感染率迅速下降，$t=15$(天)以后已经很小，说明疫情传播受到有力制约：移除率在$t=40$(天)之前变化不大，以后较快上升，表明经过约1个月的治疗期后，传播达到高潮时的

大量患者被治愈.

为了求解并检验模型,用 $\lambda(t)$, $\mu(t)$ 的部分数据拟合这两条曲线.图 4 中圆点的变化趋势表明,用指数函数作拟合比较合适.对 $\lambda(t)$ 用 $t=1\sim20$ 的数据拟合的结果是 $\hat{\lambda}(t)=0.2612e^{-0.1160t}$,对 $\mu(t)$ 用 $t=20\sim50$ 的数据拟合的结果是 $\hat{\mu}(t)=0.0017e^{-0.0825t}$,这两条指数曲线在图 4 中用实线表示,拟合较好.

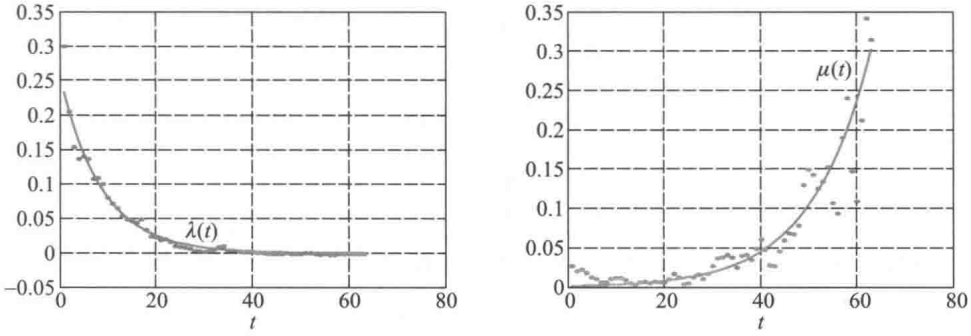


图 4 感染率 $\lambda(t)$ 和移除率 $\mu(t)$ 的估计与拟合

模型求解与检验 将拟合得到的 $\hat{\lambda}(t)$ 和 $\hat{\mu}(t)$ 代入(11),(12)式,是变量分离的方程组,解析解比较复杂,我们仍求 $i(t)$, $r(t)$ 的数值解,结果如图 5 实线所示.与 $i(t)$, $r(t)$ 的原始数据(图中圆点)比较,$i(t)$ 的计算值整体较小,且 $t=50$ 后下降过快,在模型构造、参数拟合等方面仍需改进.

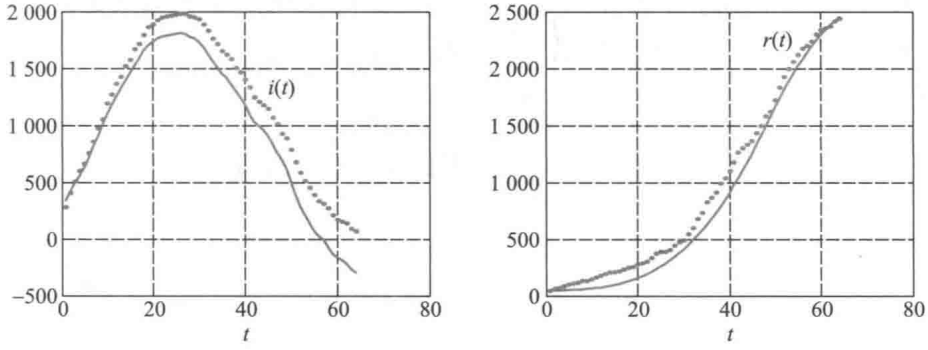


图 5 $i(t)$, $r(t)$ 的计算值与原始数据的比较

模型二——引入不可控带菌者和疑似已感染者的模型 这是在 SARS 的传播受到严格控制条件下的模型[70].

$s(t)$, $i(t)$, $r(t)$, μ 的含义与 SIR 模型相同.增加两个人群:不可控带菌者和疑似已感染者,其比例分别记作 $c(t)$ 和 $e(t)$.每个不可控带菌者发病后(收治前)每天有效感染的人数为 λ(相当于 SIR 模型的感染率),λ 中可以控制的比例为 α,不可控带菌者每天转化为已感染者的比例为 ε.疑似已感染者每天被排除的比例为 β,每天被确诊的比例为 γ.5 个人群之间的转移关系如图 6 所示.

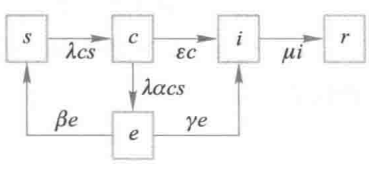


图 6 模型二中人群的构造和转移

仿照 SIR 模型并按照图 6 的转移关系可建立如下方程:

$$s(t)+c(t)+e(t)+i(t)+r(t)=1 \tag{13}$$

$$\frac{ds}{dt}=\beta e-\lambda cs \tag{14}$$

$$\frac{dc}{dt}=\lambda(1-\alpha)cs-\varepsilon c \tag{15}$$

$$\frac{de}{dt}=-(\beta+\gamma)e+\lambda\alpha cs \tag{16}$$

$$\frac{di}{dt}=\varepsilon c-\mu i+\gamma e \tag{17}$$

$$\frac{dr}{dt}=\mu i \tag{18}$$

当参数 $\lambda,\mu,\varepsilon,\alpha,\beta,\gamma$ 确定后，由(13)~(18)式的任意5个方程及 $s(0),i(0),c(0),e(0),r(0)$ 中任意4个初始值，可以计算传播过程中5类人群的比例 $s(t),c(t),e(t),i(t)$ 和 $r(t)$.

这个模型中各个参数的估计可通过两种方式进行，一种是直接利用实际数据做出估计，另一种是先由经验估计的初值，代入模型计算，再根据计算值与实际值的偏差调整估计值①.

模型三——引入尚未隔离和已经隔离已感染者的模型　下面再介绍一个关于SARS在中国的传播与控制的模型[62].

我们结合图7说明各个人群的构造和转移关系(图中没有画出易感染者 $S(t)$)②.

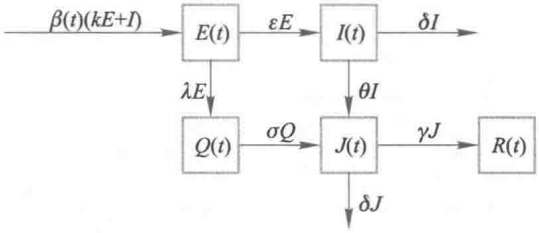


图7　模型三人群的构造和转移

$E(t)$ 表示处于潜伏期的已感染者，有弱传染性；

$I(t)$ 表示已发病但尚未隔离的已感染者，有强传染性；

$Q(t)$ 表示已发病且已经隔离、尚未治疗的已感染者(无传染性)；

$J(t)$ 表示已发病且已经隔离、正在治疗的已感染者(无传染性)；

$R(t)$ 表示治愈者(无传染性，有免疫性)；

$\beta(t)$ 表示感染率(每个有强传染性的已感染者每天传染的人数)；

k 表示一个处于潜伏期(E)的已感染者折合成未隔离(I)的已感染者的分值(指传染性)；

ε,λ 表示处于潜伏期(E)的已感染者分别转入未隔离(I)和已隔离(Q)的已感染者的比例；

---

① 各个参数的具体估计方法见《工程数学学报》2003年第7期邹宇庭等的论文.

② 这里只介绍其中建模及数值计算部分，并将原文的差分形式改写为微分形式.

θ,σ 表示未隔离(I)和已隔离(Q)的已感染者分别转入正在治疗(J)的人群的比例;

γ 表示正在治疗的已感染者(J)治愈的比例;

δ 表示未隔离(I)和正在治疗(J)的已感染者的死亡率.

按照图7各个人群的转移关系建立如下方程:

$$\frac{dE}{dt}=\beta(t)(kE+I)-(\varepsilon+\lambda)E \tag{19}$$

$$\frac{dI}{dt}=\varepsilon E-(\theta+\delta)I \tag{20}$$

$$\frac{dQ}{dt}=\lambda E-\sigma Q \tag{21}$$

$$\frac{dJ}{dt}=\theta I+\sigma Q-(\gamma+\delta)J \tag{22}$$

$$\frac{dR}{dt}=\gamma J \tag{23}$$

方程中的参数满足 $0<k<1,0<\varepsilon+\lambda<1,0<\theta+\delta<1,0<\sigma<1,0<\gamma+\delta<1$,可以利用公布的资料和试验的方法确定.

设潜伏期为6天,其中后3天才有强传染性,于是 $\varepsilon+\lambda=1/3$,设潜伏期的已感染者(E)转入未隔离(I)和已隔离(Q)的比例为2∶3,可得 $\varepsilon=1/3\times2/5=2/15$,$\lambda=1/3\times3/5=1/5$;设 I 和 Q 均约3天后转入治疗(J),于是 $\theta=\sigma=1/3$;设治疗者(J)平均3周治愈,于是 $\gamma=1/21$;死亡率取为 $\delta=1/21\times15/100=1/140$;取 $k=0.1$.

感染率 $\beta(t)$ 是模型中最重要的参数,从实际数据可知,在严格控制措施下 β 是 t 的减函数.经过多次数值计算确定 $\beta(t)=\dfrac{31+t}{22+5t}$.

根据公布的数据,以2003年4月21日各个人群的数量为初始值:$E(0)=477,I(0)=286,Q(0)=191,J(0)=848,R(0)=1213$,将上面确定的参数代入模型(19)—(23),结果如图8,实际数据与计算值相当接近(原文还有 θ 和 k 取其他数值的计算结果,与实际数据相差较大).

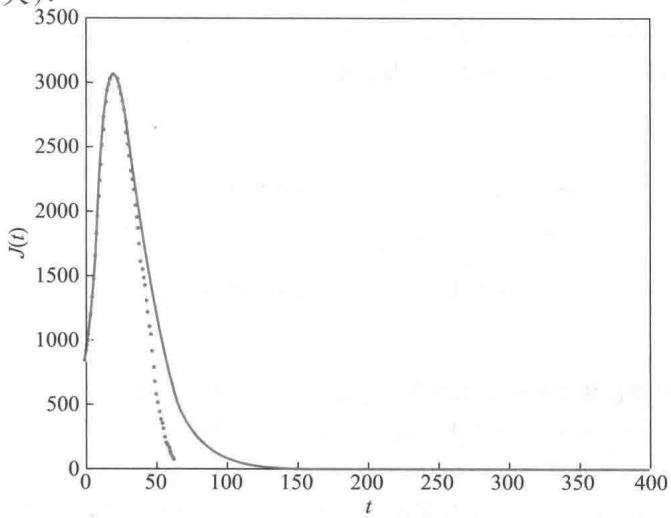


图8 模型三 $J(t)$ 的实际数据与计算结果

评注 上面3个模型以及许多参赛论文都在SIR模型的基础上建立各具特色的微分方程(或差分形式)模型,差别主要是人群的划分和参数的定义.模型求解的结果是否与实际数据相吻合,关键之一在于参数估计.而一些参数能否得到比较恰当的估计,又依赖于是否有充分的数据.人群的细分必然要引进更多的参数,如果因为受限于能够得到的实际数据,致使某些参数无法或很难估计,那么即便模型很精细,也得不到好的结果.

复习题

1. 在SIR模型中设定不同的参数 λ,μ 和初始值 $s(0)$,$i(0)$,编程计算 $s(t)$ 和 $i(t)$ 并作图.需包含 $\sigma s(0)>1$ 及 $\sigma s(0)<1$ 的设定,分析这两种情况下患者比例 $i(t)$ 的不同变化.

2. 在SIS模型中若 $\sigma>1$,且 $i_0>1-1/\sigma$,画出 $i(t)$ 的略图,分析曲线的特性(单调、凹凸、极限值等).

3. 尝试对描述SARS传播的参数时变的SIR模型进行改进,重新计算,如改变感染率 $\lambda(t)$ 和移除率 $\mu(t)$ 的拟合函数、取不同时期的数据作拟合、将治愈和死亡人群分开等.

第5章训练题

1. 为治理湖水的污染,引入一条较清洁的河水,河水与湖水混合后又以同样的流量由另一条河排出.设湖水容积为 V,河水单位时间流量为 Q,河水的污染浓度为常数 c_h,湖水的初始污染浓度为 c_0.

(1) 建立湖水污染浓度 c 随时间 t 变化的微分方程,并求解.

(2) 若测量出引入河水10天后湖水的污染浓度0.9 g/m^3,40天后湖水的污染浓度0.5 g/m^3,且河水的污染浓度 $c_h=0.1$ g/m^3,问引入河水后多少天,湖水的污染浓度可以降到标准值0.2 g/m^3.

(3) 若由于水的蒸发等原因湖水容积每天减少 b,湖水污染浓度如何变化?

2. 凌晨某地发生一起凶杀案,警方于早晨6时到达现场,测得尸温26 ℃,室温10℃.早晨8时,又测得尸温18 ℃.

(1) 若近似认为室温不变,估计凶杀案发生的时间.

(2) 若早晨8时室温升至12 ℃,如何处理这个问题?

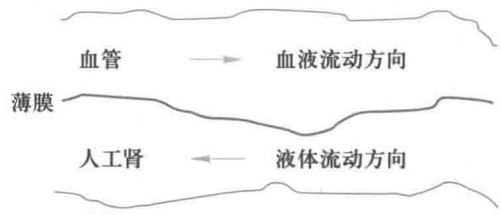
3. 对于技术革新的推广,在下列几种情况下分别建立模型:

(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,推广是无限的.

(2) 总人数有限,因而推广速度还会随着尚未采用新技术的人的人数的减少而降低.

(3) 在(2)的前提下考虑广告等媒介的传播作用[12,39].

4. 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,如下图.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.



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[12].

5. 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.

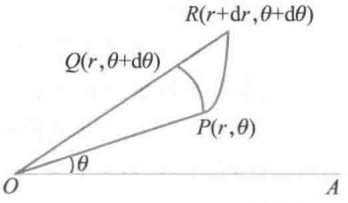
(1) 设尾数 $n(t)$ 的(相对)减少率为常数;由于喂养引起的每尾鱼质量的增加率与鱼表面积成正比,由于消耗引起的每尾鱼质量的减少率与质量本身成正比.分别建立尾数和每尾鱼质量的微分方

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[12].

6. 侦察机搜索潜艇.设 $t=0$ 时艇在 O 点,飞机在 A 点,$OA=6$ n mile(1 n mile = 1.852 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,见下图.分析图中由 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.



(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[40]?

7. 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:$\dfrac{\mathrm{d}n}{\mathrm{d}t}=-\lambda n\ln\dfrac{n}{N}$,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:$\dfrac{\mathrm{d}n}{\mathrm{d}t}=\dfrac{\lambda n}{\alpha}\left[1-\left(\dfrac{n}{N}\right)^{\alpha}\right]$ 的特例[32].

8. 与 7 题中 Usher 模型类似的是 θ-logistic 模型:$\dfrac{\mathrm{d}x}{\mathrm{d}t}=rx\left[1-\left(\dfrac{x}{N}\right)^{\theta}\right]$,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[46].

9. 药物动力学中的 Michaelis-Menton 模型为 $\dfrac{\mathrm{d}x}{\mathrm{d}t}=-\dfrac{kx}{a+x}(k,a>0)$,$x(t)$ 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[46].

(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?

10. 建立一个模型说明要用三级火箭发射人造卫星的道理.

(1) 设卫星绕地球做匀速圆周运动,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 应为多少?

(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 $v(t)=u\ln\dfrac{m_0}{m(t)}$.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 $u=3$ 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 $v(t)=(1-\lambda)u\ln\dfrac{m_0}{m(t)}$.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 $v_3=3u\ln\dfrac{k+1}{\lambda k+1}$.计算要使 $v_3=10.5$ 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[82].

11. 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[11].

12. 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[39].

13. 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[39].

14. 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) 做出 $k(x)$ 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.

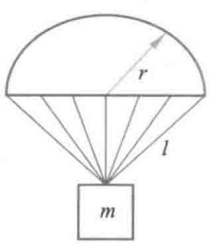
(3) 分析人口激增会引起什么后果[7].

15. 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5.3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5.3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5.3节的效益模型和捕捞过度模型进行比较[51].

16. 降落伞的选择 为向灾区空投救灾物资共2 000 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,如右图.



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,见下表;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.

r/m	2	2.5	3	3.5	4
C_1/元	65	170	350	660	1 000

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积成正比.为了确定阻力系数,用半径 $r=3$ m、载重 $m=300$ 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,见下表.

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（在第 1 个表中选择），在满足空投要求的条件下，使费用最低.

更多案例……

5-1 正规战与游击战

5-2 药物在体内的分布与排除

5-3 考虑年龄结构和生育模式的人口模型

5-4 烟雾的扩散与消失

5-5 军备竞赛

5-6 种群的相互竞争

5-7 种群的相互依存

第6章　差分方程与代数方程模型

基础知识 6-1
差分方程的类型、求解及稳定性

差分方程是在离散的时间点上描述研究对象动态变化规律的数学表达式.有的实际问题本身就是以离散形式出现的,如存贷款利息的计算;也有的是将现实世界中随时间连续变化的过程离散化,例如湖泊中的污水浓度是因上游河流的入流量和污水浓度、湖泊的天然处理污物能力等因素的不同,随时间连续变化的,而管理者采用每周对湖泊和上游河流监测一次、获取数据的方法,建立以周为离散时间单位的湖泊污水浓度的变化规律,即差分方程模型.还有一些实际问题,当不考虑时间因素作为静态问题处理时,常常可以建立代数方程模型.把代数方程与差分方程合在一章,是因为它们都用类似的矩阵、向量的数学形式表达,求解过程也有相似之处.

6.1　贷款购房

当今许多年轻人把购置一套属于自己的住房,当作是成家立业的一个标志或者前提.面对大中城市飙升的房价,大多数人选择向银行贷款购房,虽然要节衣缩食、月月上供,但是有一个温馨的小家,还是会感到无比甜蜜.买多大的房子,一共贷多少钱,每月还款几何,是首要考虑的问题.当然,网上有现成的房贷计算器,输入必要的信息,轻击鼠标立即可得结果.然而,作为一位学习数学建模、对数学的应用有兴趣的读者,了解计算原理,对结果作一些数量上的分析,会是有趣的、有帮助的.贷款购房的数学模型是最简单的差分方程模型[70].

单利和复利　你有1万元长期不用,到银行存5年定期,年利率为4.75%①,到期后得到的本金加利息(下称本息)为 $10\,000\times(1+0.047\,5\times5)=12\,375$ 元,这样计算的利息称为单利. 如果想灵活一点取用,存1年定期,年利率为3%,并申报若到期不取则自动转存,那么一年后本息为 $10\,000\times(1+0.03)$,仍按年利率3%再存1年,若如此共存5年,得到的本息为 $10\,000\times(1+0.03)^5=11\,593$ 元,这样计算的利息称为**复利**,俗称"利滚利".按照复利算出的结果比单利还少,是因为年利率不同造成的.如果用同一利率 r 计算,存期为 n,那么单位本金按照单利计算的本息是 $1+nr$,而按照复利计算的本息是 $(1+r)^n$,显然后者大于前者.

银行的零存整取业务是按单利计算的.零存整取是指每月固定存额,集零成整,约定存款期限,到期一次支取本息的定期储蓄,一般5元起存,多存不限.存期分为1年、3年、5年.走上工作岗位的年轻人如果计划贷款购房,就要开始准备首付款,其办法之一是每月从收入中留下一部分办理零存整取,既能培养勤俭节约、科学理财的习惯,也可得到一点回报.如果你打算每月存3 000元,5年能有多少利息呢?

用网上的零存整取计算器,输入存期5年、每月存入金额3 000元和5年期的年利率3.5%,得到的结果是,累计存入金额180 000元,到期本息总额196 012.50元.这是

① 按照2012.7.6调整的利率,下同.

怎么计算的?

记每月存入金额为 a,月利率为 $r=\dfrac{0.035}{12}$,存入 k 个月本息总额为 x_k,按单利计算时只有本金产生利息,没有"利滚利",设存期 $n=12\times5=60$ 个月,有

$$x_1=a+ar,x_2=x_1+a+a2r,\cdots,x_k=x_{k-1}+a+akr,k=2,3,\cdots,n \tag{1}$$

将(1)式由 $k=n$ 递推至 $k=1$ 得到整个存期的本息总额为

$$x_n=na+ar(1+2+\cdots+n)=na+ar\frac{n(n+1)}{2} \tag{2}$$

(2)式是零存整取计算到期本息总额的一般公式,将上面的 a, r, n 代入,即得计算器给出的 196 012.50 元.(2)式右端 na 为累计存入金额,后一项是利息.其实,由于每份本金 a 分别存了 n, $n-1$, …,2, 1 个月,利息可以直接算出.

等额本息贷款和等额本金贷款 下面讨论贷款购房,先利用网上的房贷计算器计算每月的还款金额,这需要先作一些选择,并输入必要的信息:

- 贷款类别:商业贷款、公积金、组合型.区别在于年利率不同,组合型指商业贷款和公积金的组合.以下选择商业贷款.
- 计算方法:根据贷款总额计算或根据面积、单价及按揭成数计算.以下选择根据贷款总额计算.
- 按揭年数:可选 1 至 30 年.以下选择 **20** 年.
- 银行利率:可选基准利率、利率上限或下限.以下选择商业贷款的基准利率 **6.55%**.
- 还款方式:等额本息还款或等额本金还款.这是我们讨论的重点,下面将对二者分别研究,进行分析、比较.

例 1 在房贷计算器上选择等额本息还款,并输入:商业贷款总额 100 万元,期限 20 年,年利率 6.55%.点击"开始计算"得到:还款总额 1 796 447.27 元,月均还款 7 485.2元.建立等额本息还款方式的数学模型,并作数值计算.

建模 所谓等额本息还款,是指每月归还本息的金额相同,记每月还款金额为 a,设贷款总额为 x_0,月利率为 r,第 k 月还款后尚欠金额为 x_k,贷款期限(月)为 n,则

$$x_k=x_{k-1}(1+r)-a,\ k=1,2,\cdots,n \tag{3}$$

将(3)式由 $k=n$ 递推至 $k=1$ 得①

$$x_n=x_0(1+r)^n-a[1+(1+r)+\cdots+(1+r)^{n-1}]=x_0(1+r)^n-a\frac{(1+r)^n-1}{r} \tag{4}$$

当贷款到期时,(4)式左端的 $x_n=0$,于是得到每月还款金额为

$$a=x_0r\frac{(1+r)^n}{(1+r)^n-1} \tag{5}$$

记等额本息还款总额为 A_1,显然

$$A_1=na=x_0rn\frac{(1+r)^n}{(1+r)^n-1} \tag{6}$$

---

① (3)式是一阶线性常系数差分方程,其解可以直接由基础知识 6-1 的求解公式得到.

(3) ~(6)式为等额本息还款方式的数学模型.

将 $x_0=100$ 万元,$r=0.065\ 5/12$,$n=12\times20=240$ 月代入(5)、(6)式,即得计算器给出的 $a=7\ 485.2$ 元,$A_1=1\ 796\ 447.27$ 元.

例 2 在房贷计算器上选择等额本金还款,并输入:商业贷款总额 100 万元,期限 20 年,年利率 6.55%.点击"开始计算"得到:还款总额 1 657 729.17 元,每月还款金额由第 1 月的 9 625 元逐月递减,最后 1 月为 4 189.41 元.建立等额本金还款方式的数学模型,并作数值计算.

建模 所谓等额本金还款,是指每月归还同等数额的本金,加上所欠本金的利息.由于所欠本金逐月减少,所以每月还款金额递减,记第 k 月还款金额为 x_k.仍设贷款总额为 x_0,月利率为 r,贷款期限(月)为 n,则每月归还本金为$\frac{x_0}{n}$,且还款金额逐月减少归还本金$\frac{x_0}{n}$所产生的利息$\frac{x_0 r}{n}$,于是

$$x_k=x_{k-1}-\frac{x_0 r}{n},\quad k=2,3,\cdots,n \tag{7}$$

(7)式递推至 $k=2$,并注意到第 1 月的还款金额为 $x_1=\frac{x_0}{n}+x_0 r$,得第 k 月还款金额为

$$x_k=\frac{x_0}{n}+x_0\left(1-\frac{k-1}{n}\right)r,\quad k=1,2,\cdots,n \tag{8}$$

记等额本金还款总额为 A_2,则

$$A_2=\sum_{k=1}^{n}x_k=x_0+x_0 r\frac{n+1}{2} \tag{9}$$

(7) ~(9)式为等额本金还款方式的数学模型.

将 $x_0=100$ 万元,$r=0.065\ 5/12$,$n=12\times20=240$ 月代入,即得计算器给出的 $x_1=9\ 625$ 元,$x_{240}=4\ 189.41$ 元,每月还款金额递减$\frac{x_0 r}{n}=22.743\ 1$ 元,$A_2=1\ 657\ 729.17$ 元.

等额本息与等额本金方式的比较 从以上建模过程和数值计算可以归纳出两种还款方法的特点如下:

- 等额本息还款方式简单,便于安排收支;
- 等额本金方式每月还款金额前期高于等额本息方式,后期低于等额本息方式,适合当前收入较高人群;
- 等额本息方式的还款总额大于等额本金方式.

对于最后一点,例 1 和例 2 的数值计算表明等额本息还款总额 A_1 比等额本金还款总额 A_2 多了近 14 万元.直观地看,等额本金方式将每月利息还清,而等额本息方式还款额(前期)比等额本金少,所欠本息的利息归入以后逐月归还,产生的利息总额将大于等额本金方式.从数量关系上,可以根据(6)、(9)式证明 $A_1>A_2$(复习题 2).

评注 本节讨论了贷款购房中等额本息和等额本金这两种基本的还款方式,虽然建模过程采用目前流行的每月还款一次的办法,但是所得到的模型(3)~(9)式适合于任何还款周期,如半个月、一季度等,只需将公布的年利率折换为一个还款周期的利率.对于不同的还款周期,一次还款金额和还款总额都不会一样.直观地看,周期越短,还款总额应该越小.读者可以从数量关系上进行分析、比较.

复习题

1. 在网上用当前的还款利率在等额本息还款、商业贷款 100 万元、期限 20 年条件下,查出还款

总额和月均还款,然后用本节的模型计算,予以核对.

2. 根据(6)、(9)式证明等额本息还款总额 A_1 大于等额本金还款总额 A_2.

6.2 管住嘴迈开腿

您的体重正常吗?目前人们公认的测评体重的标准是联合国世界卫生组织颁布的体重指数(body mass index,简记 BMI),定义为 $\mathrm{BMI}=\frac{h}{l^2}$,其中 h 是体重(单位:kg),l 是身高(单位:m),显然身材矮胖者比瘦高者的 BMI 要大.标准身材的 BMI=22,如身高 $l=1.70$ m 的人的标准体重为 $h=63.5$ kg.按照 BMI 的大小,世界卫生组织和我国给出的体重指数分级标准如表 1.

表 1 体重指数分级标准

	偏瘦	正常	超重	肥胖
世界卫生组织标准	<18.5	18.5~24.9	25.0~29.9	≥30.0
我国参考标准	<18.5	18.5~23.9	24.0~27.9	≥28.0

20 世纪 80 年代以后,随着我国人民物质生活水平的迅速提高,越来越多自感肥胖,甚至害怕肥胖的人纷纷奔向减肥药品的柜台.可是大量事实说明,大多数减肥药品并达不到减肥的效果,或者即使成功一时,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,也就是人们常说的"管住嘴、迈开腿",才能在不伤害身体的前提下,达到减轻体重并使体重得以控制的目的.本节要建立一个简单的体重变化规律的数学模型,并由此通过节食与运动制定合理、有效的减肥计划[70].

模型分析 在正常情况下,人体通过食物摄入的热量与代谢和运动消耗的热量大体上是平衡的,于是体重基本保持不变,而当体内能量守恒被破坏时就会引起体重的变化.因此需要从人体对热量的吸收和消耗两方面进行分析,在适当简化的假设下建立体重变化规律的数学模型.

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限制.另外,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——差分方程来讨论.

模型假设 根据上述分析,参考有关生理数据,做出以下简化假设:

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 kcal=4.2 kJ)增加体重 1 kg;

2. 身体正常代谢引起的体重减少正比于体重,每周每千克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 至 3 200 kcal;

3. 运动引起的体重减少正比于体重,且与运动形式和运动时间有关;

4. 为了安全与健康,每周吸收热量不要小于 10 000 kcal,且每周减少量不要超过 1 000 kcal,每周体重减少不要超过 1.5 kg.

调查资料经综合后,得到若干食物每百克所含热量及各项运动每小时每千克体重消耗热量,如表2和表3,供参考.

表2 食物每百克所含热量 单位:kcal

食物	米饭	豆腐	青菜	苹果	瘦肉	鸡蛋
所含热量/(kcal·(100 g)$^{-1}$)	120	100	20~30	50~60	140~160	150

表3 运动每小时每千克体重消耗热量

运动	步行(4 km/h)	跑步	跳舞	乒乓	自行车(中速)	游泳(50 m/min)
热量消耗/(kcal·h^{-1}·kg^{-1})	3.1	7.0	3.0	4.4	2.5	7.9

基本模型 记第k周(初)体重为$w(k)$(kg),第k周吸收热量为$c(k)$(kcal),$k=1,2,\cdots$.设热量转换(体重的)系数为α,身体代谢消耗系数为β,根据模型假设,正常情况下(不考虑运动)体重变化的基本方程为

$$w(k+1)=w(k)+\alpha c(k)-\beta w(k),\quad k=1,2,\cdots \tag{1}$$

由假设1,$\alpha=\dfrac{1}{8\ 000}$kg/kcal,当确定了每个人的代谢消耗系数β后,就可按(1)式由每周吸收的热量$c(k)$推导他(她)体重$w(k)$的变化.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 通过制定一个具体的减肥计划讨论模型(1)的应用.

某人身高1.70 m,体重100 kg,BMI高达34.6.自述目前每周吸收20 000 kcal热量,体重长期未变.试为他按照以下方式制订减肥计划,使其体重减至75 kg(即BMI=26)并维持下去:

1)在正常代谢情况下安排一个两阶段计划,第一阶段:吸收热量由20 000 kcal每周减少1 000 kcal,直至达到安全下限(每周10 000 kcal);第二阶段:每周吸收热量保持下限,直至达到减肥目标.

2)为加快进程而增加运动,重新安排两阶段计划.

3)给出达到目标后维持体重不变的方案.

减肥计划的制订 首先应确定某人的代谢消耗系数β.根据他每周吸收$c=20\ 000$ kcal热量,体重$w=100$ kg不变,在(1)式中令$w(k+1)=w(k)=w$,$c(k)=c$,得

$$w=w+\alpha c-\beta w$$

于是

$$\beta=\frac{\alpha c}{w}=\frac{20\ 000/8\ 000}{100}=0.025 \tag{2}$$

相当于每周每千克体重消耗热量$\dfrac{20\ 000}{100}=200$ kcal.从假设2可以知道,某人属于正常代谢消耗相当弱的人.他又吃得那么多,难怪如此之胖.

1)正常代谢情况下的两阶段计划.第一阶段要求吸收热量由20 000 kcal每周减少1 000 kcal(由表2,如每周减少米饭和瘦肉各350 g),达到下限$c_{\min}=10\ 000$ kcal,即

$$c(k)=20\ 000-1\ 000k,\quad k=1,2,\cdots,10 \tag{3}$$

将 $c(k)$ 及 $\alpha=\dfrac{1}{8\ 000}$,$\beta=0.025$ 代入(1)式,可得

$$\begin{aligned} w(k+1)&=(1-\beta)w(k)+\alpha(20\ 000-1\ 000\,k)\\ &=0.975w(k)+2.5-0.125k,\quad k=1,2,\cdots,10 \end{aligned} \tag{4}$$

(4)式虽是一阶线性常系数差分方程,但因右端含有时间变量 k,不能直接应用基础知识 6-1 的求解公式.更方便的办法是,以 $w(1)=100$ kg 为初始值,按照(4)式编程计算,可得第 11 周(初)体重 $w(11)=93.615\ 7$ kg.

第二阶段要求每周吸收热量保持下限 $c_{\min}$,由(1)式可得

$$\begin{aligned} w(k+1)&=(1-\beta)w(k)+\alpha c_{\min}\\ &=0.975w(k)+1.25,\quad k=11,12,\cdots \end{aligned} \tag{5}$$

以第一阶段的终值 $w(11)$ 为第二阶段的初值,按照(5)式编程计算,直至 $w(11+n)\leqslant 75$ kg 为止,可得 $w(11+22)=74.988\ 8$ kg,即每周吸收热量保持下限10 000 kcal,再有 22 周体重可减至 75 kg.两阶段共需 32 周.

2) 为加快进程而增加运动.记表 3 中热量消耗为 γ,每周运动时间为 t h,在(1)式中将 β 改为 $\beta+\alpha\gamma t$,即

$$w(k+1)=w(k)+\alpha c(k)-(\beta+\alpha\gamma t)w(k) \tag{6}$$

试取 $\alpha\gamma t=0.005$,则 $\beta+\alpha\gamma t=0.03$,(4)、(5)式分别变为

$$w(k+1)=0.97w(k)+2.5-0.125k,\quad k=1,2,\cdots,10 \tag{7}$$

$$w(k+1)=0.97w(k)+1.25,\quad k=11,12,\cdots \tag{8}$$

类似的计算可得,$w(11)=89.331\ 9$ kg,$w(11+12)=74.738\ 8$ kg,即若增加 $\alpha\gamma t=0.005$ 的运动,就可将第二阶段的时间缩短为 12 周.由 $\alpha=\dfrac{1}{8\ 000}$ 可知,增加的运动内容应满足 $\gamma t=40$,可从表 3 选择合适的运动形式和时间,如每周步行7 h 加乒乓 4 h.两阶段共需 22 周,增加运动的效果非常明显.

将正常代谢和增加运动两种情况下的体重 $w(k)$ 作图,得到的曲线如图 1.

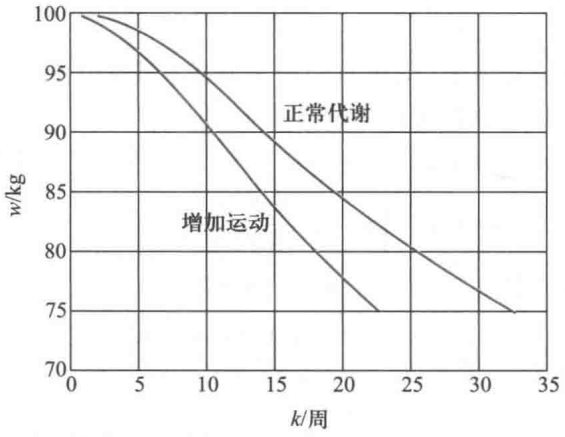


图 1 正常代谢和增加运动下的体重 $w(k)$ 曲线

经检查,两种情况下每周体重的减少都不超过 1.5 kg.

3) 达到目标后维持体重不变的方案.最简单的是寻求每周吸收热量保持某常数值 c,使体重 $w=75$ kg 不变.类似于(2)式,在(6)式中令 $w(k+1)=w(k)=w$, $c(k)=c$, 得

$$w=w+\alpha c-(\beta+\alpha\gamma t)w$$

于是

$$c=\frac{(\beta+\alpha\gamma t)w}{\alpha} \tag{9}$$

由(9)式,在正常代谢下($\gamma=0$)可算出 $c=15\ 000$ kcal;若增加 $\gamma t=40$ 的运动,则 $c=18\ 000$ kcal.

实际上,如果在减肥计划的开始,就让某人每周吸收热量由 20 000 kcal 直接减至 $c=15\ 000$ kcal,那么他的体重也将逐渐下降,很长时间以后(理论上是无穷长)才会降到 75 kg. 而若每周吸收热量减至 14 000 kcal,13 000 kcal,12 000 kcal,体重下降速度将加快,直接按照(1)式计算,得到的曲线如图 2.可以看到,达到目标体重 75 kg 的时间分别约为 70 周、50 周和 40 周,都比上面的两阶段计划长,且吸收热量的突减也对身体不利.

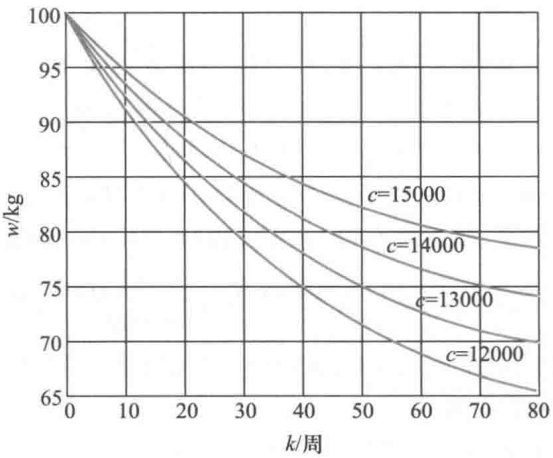


图 2 吸收不同热量 c 下的体重 $w(k)$ 曲线

评注 人体体重的变化是有规律可循的,减肥也应科学化、定量化.这个模型虽然只考虑了一个非常简单的情况,但是它对自愿进行科学减肥,甚至打算将其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,在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 0.03 时(变化 20%),减肥所需时间就从 32 周减少到 22 周(变化约 30%),所以应用这个模型时要对 β 做仔细的核对.

如果要寻求达到目标体重(记作 w^*)所需时间(周数记作 n)与每周吸收热量 c 之间的关系,将基本模型(1)式的 k 由 1 递推至 n,得

$$\begin{aligned} w(n+1)&=(1-\beta)^n w(1)+\alpha c[1+(1-\beta)+\cdots+(1-\beta)^{n-1}] \\ &=(1-\beta)^n\left[w(1)-\frac{\alpha c}{\beta}\right]+\frac{\alpha c}{\beta} \end{aligned} \tag{10}$$

令 $w^*=w(n+1)$ 并记 $w_1=w(1)$,由(10)式得

$$n=\frac{\ln(w^*-\alpha c/\beta)-\ln(w_1-\alpha c/\beta)}{\ln(1-\beta)} \tag{11}$$

以 $w^*=75$, $w_1=100$, $\alpha=\dfrac{1}{8\ 000}$, $\beta=0.025$ 及 $c=14\ 000$, $13\ 000$, $12\ 000$ 代入(11)式算出 n 分别为 70.770 7, 49.481 5, 38.740 7,与图 2 所示基本相同.

复习题

若将模型假设修改为“每周体重减少不要超过 1 kg”，检查所制订的减肥计划是否满足.如不满足.重新制订计划.

6.3 市场经济中的物价波动

在自由竞争的市场经济中常会出现这样的现象：一个时期以来某种商品如鸡蛋的供应量大于需求，由于销售不畅致使价格下跌，生产者发现养鸡赔钱，于是转而经营其他农副业.过一段时间，鸡蛋供应量就会大减，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，下一个时期又会重现供过于求、价格下跌的局面.

商品数量和价格主要由供求关系决定，供求平衡时，商品的数量和价格基本稳定，一旦出现不平衡，商品的数量和价格就会发生振荡.市场经济中这种振荡是绝对的，不可避免的，它可以有两种完全不同的形式.常见的一种是振幅逐渐减小，最终趋向平稳，这是老百姓能够接受的形式.另一种是振幅越来越大，如果没有外界如政府的干预，将导致一定意义下的经济崩溃.本节要建立一个简化的数学模型描述这种振荡现象，研究趋向平稳的条件，并讨论不稳定时政府可能采取的干预方式[44].

模型假设　研究对象是市场上某种商品的数量和价格随时间的变化规律.时间离散化为时段，1 个时段相当于 1 个生产周期，对肉、禽等指的是牲畜饲养周期，对蔬菜、水果等指的是种植周期.记第 k 时段的商品数量为 x_k，价格为 y_k.模型假设如下：

1. 当商品的供求关系处于平衡状态时，其数量为 x_0，价格为 y_0，均保持不变.

2. 按照经济规律，第 k 时段的商品价格 y_k 由消费者的需求关系决定，当商品数量 x_k 大于 x_0 时，供过于求导致价格下跌，y_k 将小于 y_0.商品数量 x_{k+1} 由生产者的供应关系决定，当价格 y_k 小于 y_0 时，低价格导致产量下降，第 $k+1$ 时段商品数量 x_{k+1} 将小于 x_0.

3. 在 x_k, y_k 偏离其平衡值 x_0，y_0 都不太大的范围内，y_k 的偏离 y_k-y_0 与 x_k 的偏离 x_k-x_0 成正比，x_{k+1} 的偏离 $x_{k+1}-x_0$ 与 y_k 的偏离 y_k-y_0 成正比.

差分方程模型　按照假设，y_k-y_0 与 x_k-x_0 成正比，且符号相反，记比例系数为 α，有

$$y_k-y_0=-\alpha(x_k-x_0),\alpha>0,k=1,2,\cdots \tag{1}$$

$x_{k+1}-x_0$ 与 y_k-y_0 成正比，且符号相同，记比例系数为 β，有

$$x_{k+1}-x_0=\beta(y_k-y_0),\beta>0,k=1,2,\cdots \tag{2}$$

若系数 α,β 及平衡值 x_0，y_0 已知，那么由商品数量的初始值 x_1，可用(1)，(2)式计算 x_k, y_k.

(1)，(2)式是关于 x_k, y_k 的差分方程组，从中消去 y_k-y_0 可得对于 x_k 的差分方程

$$x_{k+1}-x_0=-\alpha\beta(x_k-x_0),\ k=1,2,\cdots \tag{3}$$

由(3)式可以递推地解出

$$x_{k+1}-x_0=(-\alpha\beta)^k(x_1-x_0),\ k=1,2,\cdots \tag{4}$$

由(4)式可知，对于给定的商品数量的初始偏离 x_1-x_0，仅当系数 α,β 满足

$$\alpha\beta<1 \text{ 或 } \alpha<\frac{1}{\beta} \tag{5}$$

时，才有 $x_k\to x_0(k\to\infty)$，且 $y_k\to y_0$，平衡值 x_0,y_0 是稳定的；而当

$$\alpha\beta>1 \text{ 或 } \alpha>\frac{1}{\beta} \tag{6}$$

时，$x_k,y_k\to\infty$ $(k\to\infty)$，平衡值 x_0,y_0 不稳定.

模型分析 (1)~(6)式给出市场经济中商品数量和价格随时间的变化规律，显然，这种振荡是否趋向平稳完全取决于系数 α,β 的大小.先看一个数值例子.

若某种商品在市场上处于平衡状态时的数量为 $x_0=100$(单位)，价格为 $y_0=10$(元/单位)，开始时段商品数量为 $x_1=110$.当商品数量减少1个单位时，需求关系使价格上涨0.1元，而当价格上涨1元时，供应关系使下一时段供应量增加5个单位，于是 $\alpha=0.1$，$\beta=5$.按照方程(1)，(2)编程计算，得到的 $x_k,y_k(k=1,2,\cdots,10)$ 如图1，可以看出 x_k,y_k 分别趋向平衡值 x_0,y_0.显然，$\alpha\beta=0.5<1$，满足稳定条件(5)式.

如果 α 提高至0.24而 $\beta=5$ 不变，得到的 x_k,y_k 如图2，因为 $\alpha\beta=1.2>1$，由(6)式，x_0,y_0 不稳定.

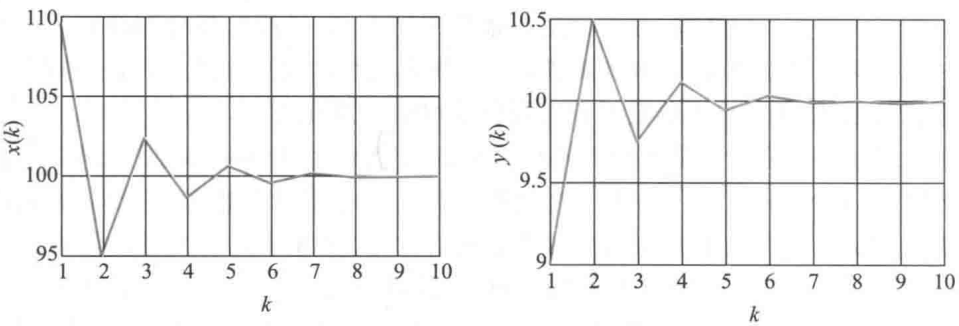


图1 $\alpha=0.1,\beta=5$ 差分方程模型(1)、(2)的 x_k,y_k

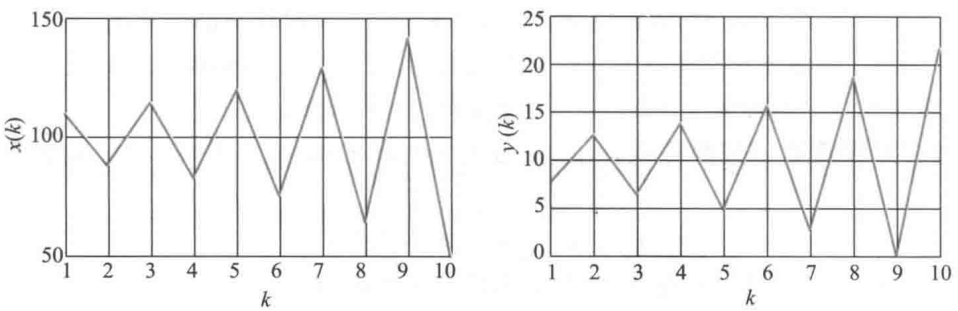


图2 $\alpha=0.24,\beta=5$ 差分方程模型(1)、(2)的 x_k,y_k

考察 α,β 的含义，可以对商品数量和价格是否会趋于稳定做出合理的解释.由(1)式可知，α 表示商品数量减少一个单位时价格的上涨幅度；由(2)式可知，β 表示价格上涨一个单位时(下一时段)供应量的增加.所以 α 反映消费者需求的敏感程度，如果是生活必需品，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，α 就会较大；反之，若消费者购物心理稳定，则 α 较小.β 的数值反映生产经营者对价格的敏感程

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，β 就会较大；反之，若他们素质较高，有长远的计划，则 β 较小.差分方程(4)表明，仅当 α,β 的乘积小于 1 时，商品数量和价格的振荡才会趋向稳定.

蛛网模型　差分方程模型以及振荡趋向稳定的条件可以通过经济学中的蛛网模型进行直观的解释.

以商品数量为横坐标 x，价格为纵坐标 y，x，y 之间的关系通过 x_k,y_k 由方程(1)、(2)给出.图 3 中，方程(1)表示为需求直线 f，方程(2)表示为供应直线 g，两条直线相交于 $P_0(x_0, y_0)$ 点.P_0 点称为平衡点，即一旦某时段 k 商品数量 $x_k = x_0$，则有 $y_k = y_0, x_{k+1} = x_0, y_{k+1} = y_0, \cdots$，永远保持在 $P_0(x_0, y_0)$ 点.而实际上，种种干扰会使得商品数量和价格偏离 P_0，不妨设 x_1 偏离 x_0，如图 3.我们分析 x_k，y_k 的变化.

商品数量 x_1 给定后，价格 y_1 由直线 f 上的 P_1 点决定.数量 x_2 又由 y_1 和直线 g 上的 P_2 点决定，y_2 由 x_2 和 f 上的 P_3 点决定，这样在图 3 上得到一系列的点 P_1，P_2，P_3，P_4，$\cdots$，这些点按图上箭头方向趋向 P_0 点，称 P_0 为稳定平衡点，意味着商品的数量和价格的振荡将趋向稳定.

但是如果直线 f 和 g 如图 4 给出的那样，完全类似的分析可以发现，P_1，P_2，P_3，$\cdots$ 沿着箭头的方向，将越来越远离 P_0 点，称 P_0 为不稳定平衡点，意味着商品的数量和价格将出现越来越大的振荡.由于图 3、图 4 中的折线 $P_1P_2P_3P_4\cdots$ 形似蛛网，所以称为蛛网模型.

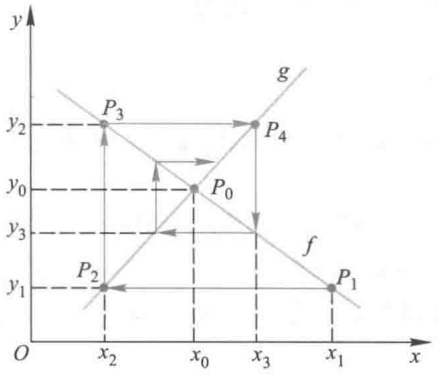


图 3　蛛网模型，P_0 稳定

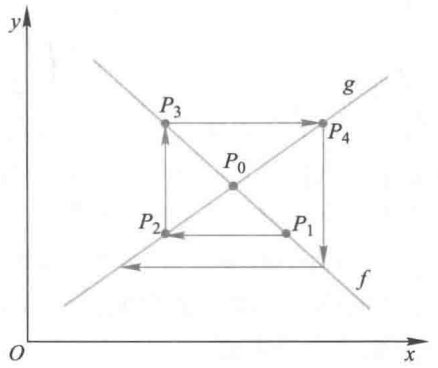


图 4　蛛网模型，P_0 不稳定

为什么会有这样截然相反的现象呢？分析一下图 3 和图 4 的不同之处就可发现，图 3 的 f 比 g 平，而图 4 的 f 比 g 陡.用 K_f 和 K_g 分别表示 f 和 g 的斜率(取绝对值)，可以看出，当 $K_f < K_g$ 时 P_0 稳定；当 $K_f > K_g$ 时 P_0 不稳定.

在差分方程模型中曾经得到平衡值 x_0,y_0 稳定的条件(5)式，注意到蛛网模型中斜率 K_f 和 K_g 与差分方程中系数 α,β 的关系是 $K_f=\alpha$，$K_g=1/\beta$，可以知道蛛网模型的直观分析与差分方程模型推导出的条件(5)、(6)式完全一致.

政府的干预　基于上述分析可以得到政府在市场经济趋向不稳定时的两种干预办法.

一种办法是使 α 尽量小，不妨考察其极端情况 $\alpha=0$，即需求直线 f 变为水平，如图 5.这样不管供应关系如何，即不管 β 多大，(5)式恒成立，平衡点 P_0 总是稳定的.实际上

这种办法相当于政府控制价格，无论商品上市量有多少，命令价格 y_0 不得改变.

另一种办法是使 β 尽量小，极端情况是 $\beta=0$，即供应直线 g 变为竖直，如图 6. 于是不管需求关系如何，即不管 α 多大，(5)式恒成立，平衡点 P_0 也总是稳定的. 实际上这相当于政府控制商品的上市数量，供不应求时，从外地收购，投入市场；供过于求时，收购过剩部分，维持上市量 x_0 不变. 当然，这种办法需要政府具有相当的经济实力.

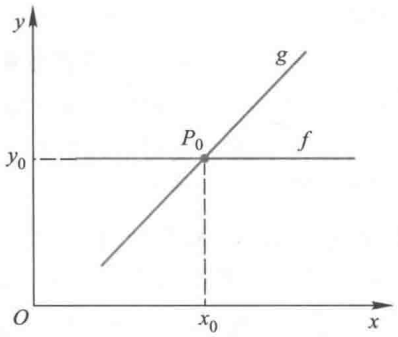


图 5 需求直线水平，经济稳定

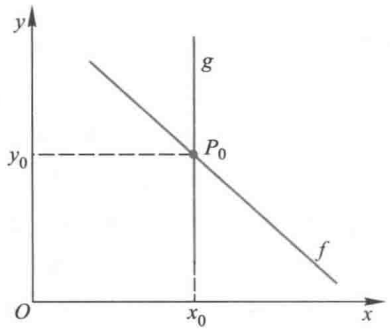


图 6 供应直线竖直，经济稳定

模型推广 在差分方程模型中，商品价格 y_k 由数量 x_k 决决定，用(1)式给出，商品数量 x_{k+1} 由价格 y_k 决定，用(2)式给出. 如果生产经营者的管理水平和素质更高一些，在决定产量 x_{k+1} 时不仅根据价格 y_k，而且考虑更前一时段的价格 y_{k-1}. 在新的模型中，y_k 与 x_k 的关系仍由(1)式给出，而 x_{k+1} 由 y_k 和 y_{k-1} 的平均值决定，(2)式改为

$$x_{k+1}-x_0=\beta\left(\frac{y_k+y_{k-1}}{2}-y_0\right),\ \beta>0,k=2,3,\cdots \tag{7}$$

对于(1)、(7)式的 x_k,y_k 差分方程组，已知系数 α,β 及平衡值 x_0，y_0，需要商品数量的 2 个初始值 x_1，x_2 才可递推地计算 x_k,y_k.

前面曾用模型(1)、(2)计算如下例子：$\alpha=0.24,\beta=5,x_0=100,y_0=10,x_1=110$，由于 $\alpha\beta=1.2>1$，平衡值 x_0,y_0 不稳定，如图 2($x_k,y_k\to\infty$). 现在令 α,β,x_0,y_0,x_1 不变，增加一个初始值 $x_2=88$①，按照新模型(1)、(7)编程计算，得到的 x_k,y_k 如图 7，可以看出 x_k,y_k 分别趋向平衡值 x_0,y_0.

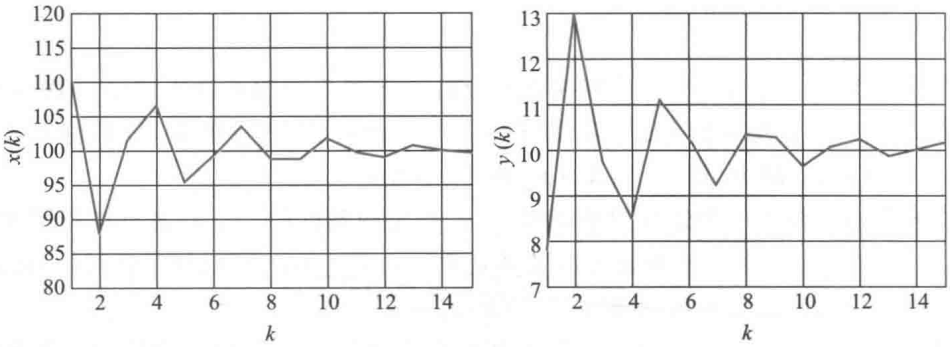


图 7 $\alpha=0.24,\beta=5$ 时差分方程模型(1)、(7)的 x_k,y_k

① 这是用模型(1)、(2)计算的结果. 实际上，x_2 换成任意数值都不影响 $k\to\infty$ 的趋向.

我们看到,对于同样的 $\alpha\beta=1.2>1$,原来不稳定的平衡值变为稳定了.为了分析新模型中平衡值 x_0,y_0 稳定的条件,从(1)、(7)式中消去 y_k 和 y_{k-1},得到对于 x_k 的二阶线性常系数差分方程:

$$2x_{k+2}+\alpha\beta x_{k+1}+\alpha\beta x_k=2(1+\alpha\beta)x_0,\quad k=1,2,\cdots \tag{8}$$

方程(8)的解表示为(参考基础知识 6-1)

$$x_k=c_1\lambda_1^k+c_2\lambda_2^k+x_0,\quad k=1,2,\cdots \tag{9}$$

其中 λ_1,λ_2 是特征方程

$$2\lambda^2+\alpha\beta\lambda+\alpha\beta=0 \tag{10}$$

的特征根,有

$$\lambda_1,\lambda_2=\frac{-\alpha\beta\pm\sqrt{(\alpha\beta)^2-8\alpha\beta}}{4} \tag{11}$$

当 $\alpha\beta<8$ 时得到

$$|\lambda_1|=|\lambda_2|=\sqrt{\frac{\alpha\beta}{2}} \tag{12}$$

根据 $k\to\infty$ 的稳定条件 $|\lambda_1|,|\lambda_2|<1$ 可得,当且仅当

$$\alpha\beta<2 \tag{13}$$

时,平衡值 x_0,y_0 是稳定的,与原模型(1)、(2)的稳定条件 $\alpha\beta<1$ 相比,新模型(1)、(7)的稳定条件放宽了,可以想到,这是因为生产者的管理水平和素质提高,对经济稳定起着有利影响的必然结果.这也自然解释了上面数值例子中 $\alpha\beta=1.2<2$ 致使平衡值 x_0,y_0 稳定的事实.

差分方程平衡点的稳定性本是一个理论问题,在这个模型中却有明显的实际意义,也从一个侧面反映了数学建模与现实世界的密切关系.

评注 在市场经济中,供不应求价格上涨、供过于求价格下跌的现象司空见惯,本节用式子简单的差分方程和形式鲜明的蛛网模型给以描述和解读,模型参数都有明确的经济学含义,参数值的两种极端情况也具有人们熟知或了解的现实背景.

复习题

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 $k+1$ 时段和第 k 时段的数量 x_{k+1} 和 x_k 决定.如果仍设 x_{k+1} 只取决于 y_k,给出稳定平衡的条件,并与本节的结果进行比较.

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.

6.4 动物的繁殖与收获

野生动物种群在自然环境下繁殖、成长、死亡,只要环境条件不发生大的变化,种群的大小就会达到基本稳定,不同年龄动物的数量比例也会保持大致的平衡.饲养动物种群受到人类的控制,为了一些特殊需要,人们可以设法使不同年龄动物的数量达到预期的比例.本节先建立动物种群的自然增长模型,再讨论饲养动物种群的稳定收获问题.

按年龄分组的动物种群增长模型

不同年龄动物的繁殖率、死亡率有较大差别,因此在研究种群数量变化时,需要按

照年龄将动物分组,时间也相应地离散化.动物种群是直接通过雌性繁殖而增长的,用雌性个体数量的变化作为研究对象比较方便,下面讨论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[73].

模型假设 对种群分组和时段划分、种群的繁殖和死亡作如下简化、合理的假设:

1. 将动物种群按年龄大小等间隔地分成 n 个年龄组,如每1岁或5岁为一组.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年为一个时段.

2. 在稳定环境下和不太长的时期内,每个年龄组种群的繁殖率和死亡率不随时段变化,只与年龄组有关.

模型建立 记第 i 年龄组第 k 时段的种群数量为 $x_i(k)$, $i=1,2,\cdots,n$, $k=0,1,2,\cdots$.设第 i 年龄组的繁殖率为 b_i,即每个(雌性)个体在一个时段内繁殖的数量;第 i 年龄组的死亡率为 d_i,即一个时段内死亡数量在总量中的比例.记 $s_i=1-d_i$,称为存活率.通常 b_i 和 s_i 可由统计资料获得,且 $b_i\geqslant 0$, $i=1,2,\cdots,n$,其中至少一个 $b_i>0$,而 $0<s_i\leqslant 1$, $i=1,2,\cdots,n-1$, $s_n=0$(不考虑最高年龄组的存活).

为寻求按年龄分组的种群数量 $x_i(k)$ 由第 k 到 $k+1$ 时段的变化规律,首先注意到,第1年龄组(包含出生的年龄组)第 $k+1$ 时段的数量是各年龄组第 k 时段的繁殖数量之和,即

$$x_1(k+1)=\sum_{i=1}^{n}b_i x_i(k),\ k=0,1,2,\cdots \tag{1}$$

其中繁殖率 b_i 的统计中已经扣除了虽出生但在一个时段内死亡的幼年动物.其次,只有第 k 时段第 i 年龄组种群中存活下来的部分,才能在第 $k+1$ 时段演变到第 $i+1$ 年龄组($i=1,2,\cdots,n-1$),所以

$$x_{i+1}(k+1)=s_i x_i(k),\ k=0,1,2,\cdots,\ i=1,2,\cdots,n-1 \tag{2}$$

(1),(2)式构成包含 n 个变量 $x_i(k)$($i=1,2,\cdots,n$)的差分方程组,表示了按年龄分组的种群数量的变化规律.已知繁殖率 b_i($i=1,2,\cdots,n$)、存活率 s_i($i=1,2,\cdots,n-1$)及种群各年龄组的初始数量 $x_i(0)$,可由方程组(1),(2)计算任意时段各年龄组种群的数量分布,称为按年龄分组的种群增长模型.

将各年龄组第 k 时段的种群数量 $x_i(k)$ 排成如下向量

$$\boldsymbol{x}(k)=[x_1(k),x_2(k),\cdots,x_n(k)]^{\mathrm{T}},\ k=0,1,2,\cdots \tag{3}$$

繁殖率 b_i 和存活率 s_i 排成如下矩阵

$$\boldsymbol{L}=\begin{bmatrix} b_1 & b_2 & \cdots & b_{n-1} & b_n \\ s_1 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & s_{n-1} & 0 \end{bmatrix} \tag{4}$$

则差分方程组(1),(2)可以简洁地表示为如下的向量-矩阵形式

$$\boldsymbol{x}(k+1)=\boldsymbol{L}\boldsymbol{x}(k),\ k=0,1,2,\cdots \tag{5}$$

给定种群各年龄组的初始数量 $\boldsymbol{x}(0)$,种群各年龄组任意时段的数量为

$$\boldsymbol{x}(k)=\boldsymbol{L}^k\boldsymbol{x}(0),\ k=0,1,2,\cdots \tag{6}$$

有了 $\boldsymbol{x}(k)$,就很容易计算种群在任意时段的总量.

有时,人们更关注任意时段种群各年龄组的数量在总量中的比例,记

$$\boldsymbol{x}^*(k)=[x_1^*(k),x_2^*(k),\cdots,x_n^*(k)]^{\mathrm{T}},x_i^*(k)=\frac{x_i(k)}{\sum_{i=1}^{n}x_i(k)} \tag{7}$$

即 $\boldsymbol{x}^*(k)$ 是 $\boldsymbol{x}(k)$ 的归一化向量,称为种群按年龄组的分布向量.

形如(4)式的矩阵 $\boldsymbol{L}$ 是 20 世纪 40 年代由 Leslie 提出的,称 **Leslie 矩阵**,按年龄分组的种群增长规律完全由 Leslie 矩阵 $\boldsymbol{L}$ 决定,(3)~(5)式又称 **Leslie 模型**.

模型求解 对一个数值例子做计算.

设一个种群分成 5 个年龄组,各年龄组的繁殖率为 $b_1=0$, $b_2=0.2$, $b_3=1.8$, $b_4=0.8$, $b_5=0.2$, 存活率为 $s_1=0.5$, $s_2=0.8$, $s_3=0.8$, $s_4=0.1$, 各年龄组现有数量均为 100 只,求任意时段种群各年龄组的数量 $\boldsymbol{x}(k)$ 及分布向量 $\boldsymbol{x}^*(k)$.

按照(3)~(5)式及(7)式编程计算,得到的结果见表 1 和图 1.

表 1 $x(k)$ 和 $x^*(k)$ 的计算结果

	k								
	0	1	2	3	…	27	28	29	30
$x_1(k)$	100	300	220	155	…	403	412	423	434
$x_2(k)$	100	50	150	110	…	196	201	206	211
$x_3(k)$	100	80	40	120	…	152	157	161	165
$x_4(k)$	100	80	64	32	…	120	122	126	129
$x_5(k)$	100	10	8	6	…	12	12	12	13
$x_1^*(k)$	0.200 0	0.576 9	0.456 4	0.365 8	…	0.456 4	0.455 3	0.455 9	0.456 2
$x_2^*(k)$	0.200 0	0.096 2	0.311 2	0.259 9	…	0.222 4	0.222 9	0.221 8	0.222 2
$x_3^*(k)$	0.200 0	0.153 8	0.083 0	0.283 6	…	0.172 6	0.173 7	0.173 7	0.173 0
$x_4^*(k)$	0.200 0	0.153 8	0.132 8	0.075 6	…	0.135 5	0.134 9	0.135 4	0.135 5
$x_5^*(k)$	0.200 0	0.019 2	0.016 6	0.015 1	…	0.013 2	0.013 2	0.013 1	0.013 2

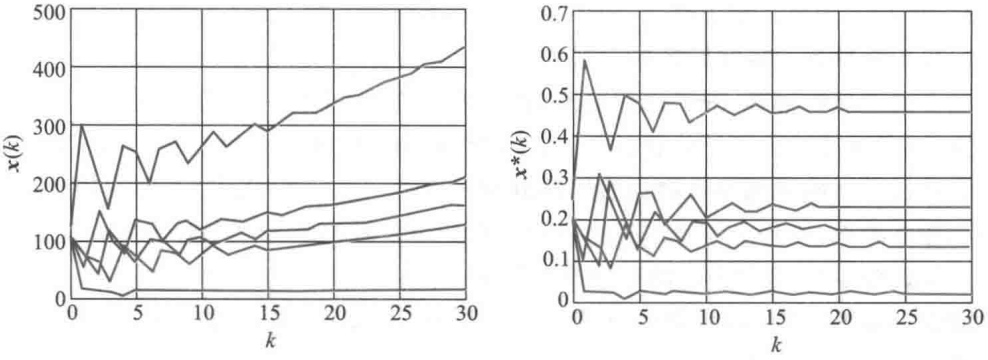


图 1 $\boldsymbol{x}(k)$ 和 $\boldsymbol{x}^*(k)$ 的图形(自上而下为 $x_1(k)$, $x_1^*(k)$ 至 $x_5(k)$, $x_5^*(k)$)

结果分析 从表 1 和图 1 可以看出,时间充分长以后,虽然各年龄组(除最高年龄

组外)种群数量 $\boldsymbol{x}(k)$ 仍在增长,但是分布向量 $\boldsymbol{x}^*(k)$ 却趋向稳定.为了进一步分析 k 充分大以后 $\boldsymbol{x}(k)$ 和 $\boldsymbol{x}^*(k)$ 的变化规律,先不加证明地给出一个数学上的结论①:

矩阵 $\boldsymbol{L}$ 有一个正的最大特征根 λ(单根),λ 对应的特征向量记作 $\boldsymbol{x}_\lambda$,由

$$\boldsymbol{x}_\lambda=\left[1,\frac{s_1}{\lambda},\frac{s_1s_2}{\lambda^2},\cdots,\frac{s_1s_2\cdots s_{n-1}}{\lambda^{n-1}}\right]^{\mathrm{T}} \tag{8}$$

确定,可以验证 $\boldsymbol{L}\boldsymbol{x}_\lambda=\lambda\boldsymbol{x}_\lambda$,且由(6)式表示的 $\boldsymbol{x}(k)$ 满足

$$\lim_{k\to\infty}\boldsymbol{x}(k)/\lambda^k=c\boldsymbol{x}_\lambda \tag{9}$$

其中 c 是常数. 将 $\boldsymbol{x}_\lambda$ 归一化后记作 $\boldsymbol{x}^*$,当 k 充分大以后,由此可得以下几条性质:

1. 种群按年龄组的分布向量 $\boldsymbol{x}^*(k)$ 满足

$$\boldsymbol{x}^*(k)\approx\boldsymbol{x}^* \tag{10}$$

表示按年龄组的分布向量接近于归一化的特征向量 $\boldsymbol{x}^*$,$\boldsymbol{x}^*$ 称为稳定分布.

2. 各年龄组种群数量 $\boldsymbol{x}(k)$ 满足 $\boldsymbol{x}(k+1)\approx\lambda\boldsymbol{x}(k)$,即

$$x_i(k+1)\approx\lambda x_i(k),i=1,2,\cdots,n \tag{11}$$

表示各年龄组种群数量将按同一比例 λ 增减,λ 称固有增长率.当 $\lambda=1$ 时,各年龄组种群数量保持不变.

3. 当 $\lambda=1$ 时,$\boldsymbol{x}_\lambda=[1,s_1,s_1s_2,\cdots,s_1s_2\cdots s_{n-1}]^{\mathrm{T}}$,于是各年龄组种群数量 $x_i(k)$ 满足

$$x_{i+1}(k)\approx s_ix_i(k),i=1,2,\cdots,n \tag{12}$$

即存活率 s_i 等于同一时段相邻年龄组的数量之比.

评注 在按年龄分组的种群模型中,曾假定种群的繁殖率和死亡率(或存活率)只与年龄组有关. 如果这些参数随时段变化,用 $b_i(k)$,$s_i(k)$,$\boldsymbol{L}(k)$ 表示,也可以建立类似于(1)~(5)式的模型,得到种群各年龄组任意时段的数量,只是由于 $\boldsymbol{L}(k)$ 不再具有前面矩阵 $\boldsymbol{L}$ 的性质,所以无法进行 k 充分大以后的稳定性分析.

上面的数值例子可以验证这几条结论:

1. 用 MATLAB 程序 [v,d]=eig(L) 可以得到任意矩阵 $\boldsymbol{L}$ 的全部特征根及特征向量.对例子中的矩阵 $\boldsymbol{L}$ 编程计算,得到特征根中最大的一个为 $\lambda=1.0254$,λ 对应的特征向量归一化后为 $\boldsymbol{x}^*=[0.4559, 0.2223, 0.1734, 0.1353, 0.0132]^{\mathrm{T}}$,与按照(8)式计算的 $\boldsymbol{x}_\lambda$ 归一化后的 $\boldsymbol{x}^*$ 一致. 表1中 k 充分大以后 $\boldsymbol{x}^*(k)$ 的计算结果非常接近 $\boldsymbol{x}^*$,满足(10)式.

2. 可以校核在表1的 $\boldsymbol{x}(k)$ 的计算结果中,k 充分大以后,$x_i(k+1)$ 与 $x_i(k)$ 的比值都在 $\lambda=1.0254$ 附近($i=1,2,\cdots,5$,其中 $x_5(k)$ 数值较小,取整后计算误差较大),满足(11)式.

3. 因 $\lambda=1.0254$ 比1略大,可以由表1对于充分大的 k 近似地验证(12)式.

饲养动物种群的持续稳定收获模型

对于饲养动物种群,除了自然繁殖、自然死亡之外,人们往往希望控制各年龄组的种群数量,实现持续稳定的收获,即同一年龄组种群的收获量在每个时段都相等.

假定自然环境下饲养动物仍然用种群增长模型(1)~(5)式表示,实现持续稳定收获的一种办法是,使每个年龄组每个时段种群的增长量都等于这个时段的收获量,并且种群数量始终不变.这个办法的数学模型如下:

建模 记第 i 年龄组种群的收获系数为 $h_i(0\leqslant h_i\leqslant 1)$,表示收获量占总量的比例,那么增长量等于收获量就可以表为

① 要求 Leslie 矩阵 $\boldsymbol{L}$ 第一行有两个顺序的 b_i 大于零,这是一个实际上很容易满足的条件.

$$x_i(k+1)-x_i(k)=h_i x_i(k+1),i=1,2,\cdots,n,k=1,2,\cdots \tag{13}$$

用 $\boldsymbol{H}$ 表示以 h_i 为对角元素的对角矩阵

$$\boldsymbol{H}=\begin{bmatrix} h_1 & 0 & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & h_n \end{bmatrix} \tag{14}$$

则(13)式可记作

$$\boldsymbol{x}(k+1)-\boldsymbol{x}(k)=\boldsymbol{H}\boldsymbol{x}(k+1),k=1,2,\cdots \tag{15}$$

利用(5)式又可将(15)式写为

$$\boldsymbol{L}\boldsymbol{x}(k)-\boldsymbol{x}(k)=\boldsymbol{H}\boldsymbol{L}\boldsymbol{x}(k),k=1,2,\cdots \tag{16}$$

实现持续稳定收获,要求种群数量 $\boldsymbol{x}(k)$ 对 k 保持不变,记作 $\boldsymbol{x}$,由(16)式得 $\boldsymbol{L}\boldsymbol{x}-\boldsymbol{x}=\boldsymbol{H}\boldsymbol{L}\boldsymbol{x}$,即

$$\boldsymbol{L}'\boldsymbol{x}=\boldsymbol{x},\boldsymbol{L}'=\boldsymbol{L}-\boldsymbol{H}\boldsymbol{L} \tag{17}$$

由(4)式和(14)式可将 $\boldsymbol{L}'$表为

$$\boldsymbol{L}'=\begin{bmatrix} b_1(1-h_1) & b_2(1-h_1) & \cdots & b_{n-1}(1-h_1) & b_n(1-h_1) \\ s_1(1-h_2) & 0 & \cdots & 0 & 0 \\ 0 & s_2(1-h_3) & & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & s_{n-1}(1-h_n) & 0 \end{bmatrix} \tag{18}$$

$\boldsymbol{L}'$也是 Leslie 矩阵.

持续稳定收获的条件 $\boldsymbol{L}'\boldsymbol{x}=\boldsymbol{x}$ 表明,矩阵 $\boldsymbol{L}'$的最大特征根 $\lambda'=1$.用线性代数中求矩阵特征根的方法,可知 $\lambda'=1$ 等价于

$$(1-h_1)[b_1+b_2s_1(1-h_2)+\cdots+b_ns_1\cdots s_{n-1}(1-h_2)\cdots(1-h_n)]=1 \tag{19}$$

$\lambda'=1$ 对应的特征向量为

$$\boldsymbol{x}'=[1,s_1(1-h_2),\cdots,s_1\cdots s_{n-1}(1-h_2)\cdots(1-h_n)]^{\mathrm{T}} \tag{20}$$

设各年龄组的繁殖率 b_i 和存活率 s_i 已知,选择一组满足(19)式的收获系数 $h_i(i=1,2,\cdots,n)$,就可以实现持续稳定收获,且种群按年龄组的稳定分布为(20)式的 $\boldsymbol{x}'$,而收获量的稳定分布为

$$\begin{aligned}\boldsymbol{H}\boldsymbol{L}\boldsymbol{x}'=[&h_1(b_1+b_2s_1(1-h_2)+\cdots+b_ns_1\cdots s_{n-1}(1-h_2)\cdots(1-h_n)),\\ &h_2s_1,h_3s_1s_2(1-h_2),\cdots,h_ns_1\cdots s_{n-1}(1-h_2)\cdots(1-h_{n-1})]^{\mathrm{T}}\end{aligned} \tag{21}$$

求解 对一个数值例子做计算.

设一个种群分成 3 个年龄组,各年龄组的繁殖率为 $b_1=0$, $b_2=5$, $b_3=2$, 存活率为 $s_1=0.8$, $s_2=0.5$, 确定各年龄组的收获系数以实现稳定收获,并求种群及收获量按年龄组的稳定分布.

记各年龄组的收获系数为 h_1, h_2, h_3, 将以上数据代入持续稳定收获的条件(19)式得

$$(1-h_1)[4(1-h_2)+0.8(1-h_2)(1-h_3)]=1 \tag{22}$$

代入(20)式得种群按年龄组的稳定分布为

评注 从数学建模角度看,人口的增减与动物种群数量

的变化规律基本上是相同的，只需将 b_i 视为生育率，(1)~(5)式就可表示为离散型的女性人口模型．当然，作为人口模型还应研究通过控制生育调整人口数量等方面的内容．

$$\boldsymbol{x}'=[1,0.8(1-h_2),0.4(1-h_2)(1-h_3)]^{\mathrm{T}} \tag{23}$$

代入(21)式得收获量的稳定分布为

$$\boldsymbol{HLx}'=[h_1(4(1-h_2)+0.8(1-h_2)(1-h_3)),0.8h_2,0.4h_3(1-h_2)]^{\mathrm{T}} \tag{24}$$

满足(22)式的 h_1，h_2，h_3 有很多，可以根据饲养者的意愿选择．如取 $h_1=0$，$h_2=0.75$，$h_3=1$，即不出售幼畜、出售75%的成年牲畜及全部老年牲畜，则 $\boldsymbol{x}'=[1,0.2,0]^{\mathrm{T}}$，$\boldsymbol{HLx}'=[0,0.6,0.1]^{\mathrm{T}}$；或取 $h_1=0.5$，$h_2=0.5$，$h_3=1$，即出售50%的幼畜和成年牲畜及全部老年牲畜，则 $\boldsymbol{x}'=[1,0.4,0]^{\mathrm{T}}$，$\boldsymbol{HLx}'=[1,0.4,0.2]^{\mathrm{T}}$．

复习题

1. 按年龄分组的种群模型中，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．

2. 讨论饲养动物种群的持续稳定收获模型的两个特例：

(1) 有些种群最年幼组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，可设收获系数为 $h_1=h$，$h_2=\cdots=h_n=0$．仍用第1题中繁殖率和存活率的数据，求收获系数、种群及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 $h_1=\cdots=h_n=h$．讨论与(1)同样的问题．

6.5 信息传播

当今社会处于信息时代，每天都有大量的、正面或负面的信息，甚至谣言，通过各种传统的、近代的、特别是互联网的渠道，在几乎没有限制的人群中传播．信息传播，尤其是在互联网上的传播，已成为当前热门的研究课题．本节讨论传统的信息传播数学模型，可作为进一步研究的基础．

考察一个封闭的、总人数一定的环境，如一所学校、一座办公楼、一个社区内的人群，开始时有极少数人得到了一条与这个人群有关的信息，或者制造了一条谣言，然后每天通过人与人之间的交流在这个人群中传播，使得获知信息的人数越来越多．我们要在一些合理的简化假设下，建立一个数学模型来描述信息传播的规律，研究其发展趋势．

模型假设　在封闭环境中人群的总人数不变．信息通过已获知的人向未获知的人传播．

因为已获知信息的人数越多(传播人群)每天新获知信息的人数就越多；未获知信息的人数越多(潜在人群)新获知信息的人数也越多，故可作如下假设：每天新获知信息的人数与已获知信息的人数和未获知信息的人数的乘积成正比．

模型建立　记总人数为 N，第 k 天已获知信息的人数为 p_k，则未获知信息的人数为 $N-p_k$．按照假设，第 $k+1$ 天新获知信息的人数 $p_{k+1}-p_k$(或记作 Δp_k)与 p_k 和 $N-p_k$ 的乘积成正比，即

$$p_{k+1}-p_k=cp_k(N-p_k) \tag{1}$$

其中 c 是比例系数，可写成 $c=\dfrac{\Delta p_k/p_k}{N-p_k}$，其含义是对于每个未获知信息的人($N-p_k$ 中的

一位)而言,每天新获知信息人数的相对增量(百分比增量).比如 $c=0.000\ 1$,若某天潜在人群尚有 500 人,则下一天新获知信息的人数将增加 $0.000\ 1\times500=5\%$.显然,c 是反映传播速度的参数,c 越大传播速度越快.

模型求解 (1) 式是一阶差分方程,可以改写成便于递推的形式:

$$p_{k+1}=(1+cN)p_k\left(1-\frac{c}{1+cN}p_k\right) \tag{2}$$

设定 $N=1\ 000$,初始获知信息的人数 $p_0=10$,c 取 2 个不同的值 0.000 1 和 0.000 2,按照(2) 式递推地计算 $p_k(k=1,2,\cdots)$,结果如图 1.与 $c=0.000\ 1$ 时 80 天后已获知信息的人数 p_k 才接近总人数 N 相比,$c=0.000\ 2$ 时 p_k 接近总人数 N 只需 40 天,传播速度明显加快.

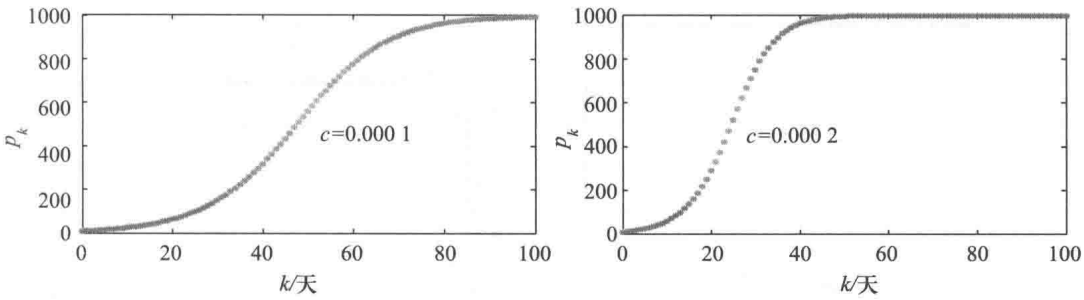


图 1 c 取不同值时方程(2)的解($N=1\ 000$,$p_0=10$)

由图 1 还可以看出,p_k 的图形恰似 5.1 节人口增长中的 S 形连续曲线.实际上,与 5.1 节的 logistic 方程对比可知,方程(1)就是 logistic 微分方程的离散形式——差分方程.

但是与 logistic 微分方程有解析解不同,其对应的差分方程(2)却无法得到 p_k 的显式表达式,对于这类方程人们常常更关注当 $k\to\infty$ 时 p_k 的收敛性.为了研究的方便,对方程(2)作如下变换,令

$$b=1+cN \tag{3}$$

$$x_k=\frac{c}{1+cN}p_k \tag{4}$$

则(2)式简化为只有一个参数 b 的标准形式:

$$x_{k+1}=bx_k(1-x_k) \tag{5}$$

(5) 式是一阶非线性差分方程,尽管看起来足够简单,可是它的解 x_k 当 $k\to\infty$ 时的收敛性却显示出复杂的性质.

方程(5)的平衡点及其稳定性 (参考基础知识 6-1)为了求出方程(5)的平衡点,解代数方程 $x=f(x)=bx(1-x)$,得到非零平衡点

$$x^*=1-1/b \tag{6}$$

为了判断平衡点 x^* 的稳定性,计算 $f(x)=bx(1-x)$ 在 x^* 点的导数 $f'(x^*)=b(1-2x^*)=2-b$,由 x^* 的稳定性条件 $|f'(x^*)|<1$,得到 $1<b<3$.按照(3)式定义的 b,必有 $b>1$.那么,对于 $b<3$ 和 $b>3$,方程(5)的解 x_k 当 $k\to\infty$ 时会发生什么现象呢?

图 2 是在不同 b 值下对方程(5)递推计算的结果(初始值均为 $x_0=0.2$).可以看到,当 $b=1.7$ 和 $b=2.9$ 时 x_k 都趋向平衡点 x^*,即 x^* 是稳定的,二者的差别在于前者 x_k 单



拓展阅读 6-1

差分方程 $x_{k+1}=bx_k(1-x_k)$ 的收敛、分岔和混沌

调收敛，后者 x_k 振荡地（在某个 k 后 x_k 交替地大于和小于 x^*）收敛（进一步分析表明，两种情况的分界点是 $b=2$）；$b=3.15$ 时 x_k 不趋向平衡点 x^*，而有 x_k 的 2 个子序列（x_{2k} 和 x_{2k+1} 序列）分别趋向另外 2 个平衡点；$b=3.525$ 和 $b=3.565$ 时则分别有 x_k 的 4 个子序列（x_{4k}，x_{4k+1}，x_{4k+2}，x_{4k+3}）和 8 个子序列，分别趋向另外 4 个和 8 个平衡点，这种子序列的收敛称为**分岔**；$b=3.7$ 时 x_k 就不趋向任何平衡点，序列变得杂乱无章，出现所谓**混沌**现象.

评注 注意到不论参数值多大，logistic 微分方程的非零平衡点总是稳定的，看来非线性差分方程与对应的非线性微分方程本质上确实有不同之处.

当参数 b 取不同的数值时方程(5)的解 x_k 呈现出不同的特性：单调或振荡收敛于稳定平衡点，出现分岔或混沌，这是非线性差分方程特有的现象.对方程(5)的平衡点及其稳定性的进一步分析，参考拓展阅读 6-1.

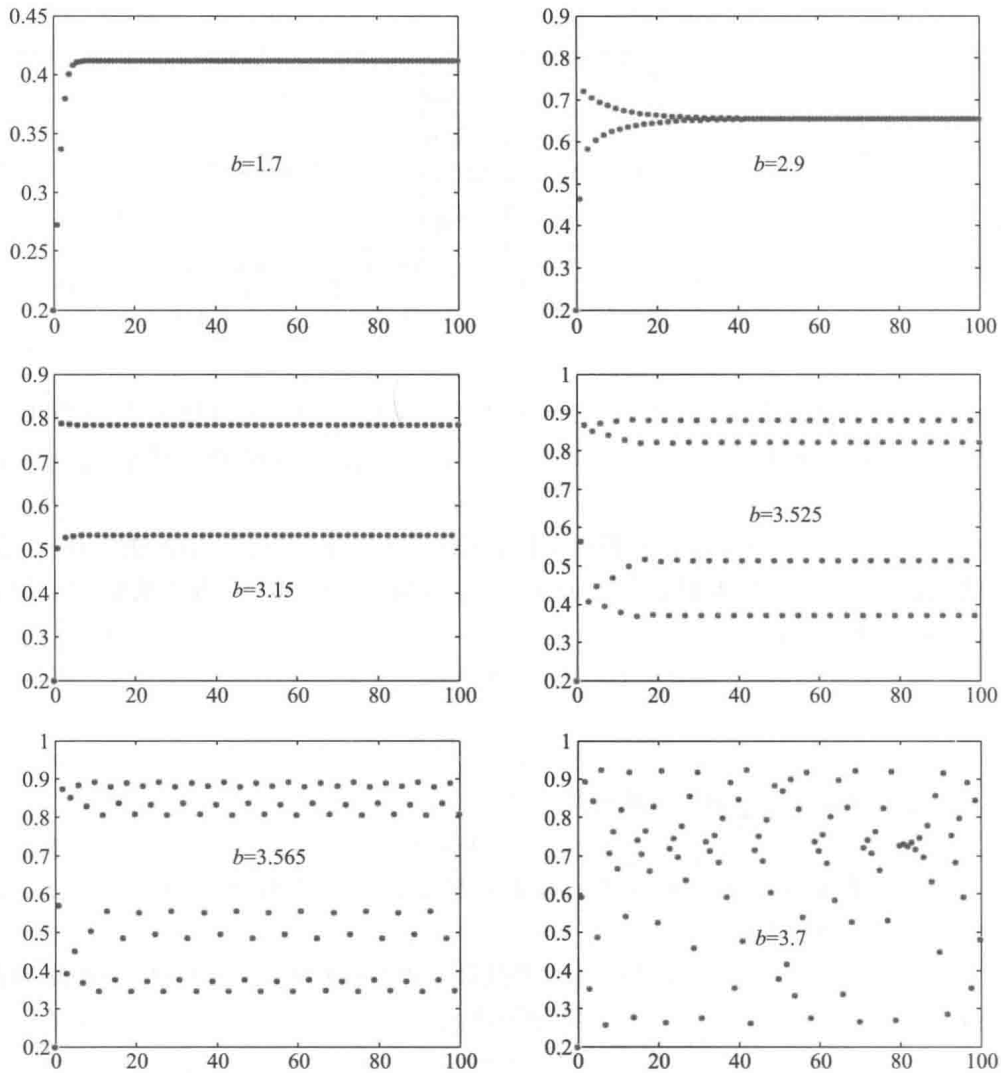


图 2 在不同 b 值下方程(5)的计算结果（初始值均为 $x_0=0.2$）

模型讨论 回到上面的信息传播模型，由(3)，(4)可以验证，方程(5)的平衡点 $x^*=1-1/b$ 等价于原方程(1)的非零平衡点 $p^*=N$，平衡点 x^* 的稳定性条件 $1<b<3$ 等价于原方程(1)平衡点 p^* 的稳定性条件 $cN<2$，而 x_k 单调收敛于 x^* 的条件 $b<2$ 等价于 $cN<1$.我们已经看到，$2<b<3$ 时 x_k 振荡收敛于 x^*，$b>3$ 时更是出现分岔、混沌现象，那么

需要讨论的是,对于信息传播模型的方程(1),会不会出现 $cN>1$ 的情况,导致 p_k 振荡收敛于 $p^*=N$ 甚至出现分岔、混沌呢?

从方程(1)出发,设总人数 N 固定,很容易得出反映传播速度的参数 c 的上限是多少.注意到传播过程中对于任意的 k 都有 $p_{k+1}<N$,于是由(1)式得 $cp_k<1$,而 p_k 可以无限接近 N,所以 c 的上限是 $1/N$,不会出现 $cN>1$ 的情况.

在实际的信息传播过程中,c 不大可能保持不变,如果能够确定随 k 而变的 c_k,用以代替方程(1)的 c,仍然可以递推地计算 p_k,模型的结果会更好.

模型拓展 上面讨论的是信息得以自由传播的过程,如果对谣言的传播加以人为干预呢?

记总人数为 N,第 k 天已听信谣言的人数为 p_k(传播人群),如果对于每天新听信谣言人数的增加仍遵从前面的假设,而对人为干预谣言的传播作如下假定:每天被制止传播谣言的人数比例为常数 a,则方程(1)的右端应增加一项 $-ap_k$,即

$$p_{k+1}-p_k=cp_k(N-p_k)-ap_k \tag{7}$$

(7)式右端的含义还表明,被制止传播谣言的人又加入到听信谣言的潜在人群中.

容易看出,方程(7)的求解与方程(1)没有本质的区别.

如果被制止传播谣言的人既不再传播也不会听信,从此退出这个信息传播系统呢?那就要考虑一个新的人群——被制止传播后退出系统者,记第 k 天退出系统者的人数为 q_k,则有

$$q_{k+1}-q_k=ap_k \tag{8}$$

而上面的方程(7)应改为

$$p_{k+1}-p_k=cp_k(N-q_k-p_k)-ap_k \tag{9}$$

(8),(9)联立组成非线性差分方程组,给定 p_0 和 q_0(通常 q_0 很小),可递推地计算 p_k 和 q_k.

评注 看来信息传播模型的解 p_k 只能具有单调增的收敛形式,完全符合人们的直观认识.引入并介绍这类方程的标准形式(5)的平衡点及其稳定性,不过是说明即使非常简单的非线性差分方程,也会出现比它对应的 logistic 微分方程有着更为复杂的特性.

评注 有不少与信息传播过程类似的实际问题,可以建立形如方程(1)的数学模型,例如受限环境下传染病的蔓延和生物种群的增长.回顾 5.10 节介绍的传染病微分方程模型(SI 模型、SIS 模型和 SIR 模型),如果按照那里的假设条件建立相应的差分方程模型,那么它们将转化为本节的哪几个方程呢?(复习题 1)

复习题

1. 按照 5.10 节传染病模型中 SI 模型、SIS 模型和 SIR 模型的假设条件,分别建立差分方程模型,并说明与本节的哪些方程相对应(特别注意 SIR 模型中变量的选取).

2. 下表是在受限环境下培养的酵母生物量的实验数据(k:小时,p_k:生物量).[24]

(1) 对 $k\sim p_k$ 作散点图,观察变化趋势,估计 p_k 增长的上限 N;

(2) 参考本节的模型对酵母生物量的增长建模;

(3) 用实验数据估计模型参数(提示:对数据 $\dfrac{p_{k+1}-p_k}{p_k}\sim N-p_k$ 作图,用线性最小二乘法估计参数);

(4) 计算模型的理论值并与实验数据 p_k 比较.

k	0	1	2	3	4	5	6	7	8	9
p_k	9.6	18.3	29.0	47.2	71.1	119.1	174.6	257.3	350.7	441.0
k	10	11	12	13	14	15	16	17	18	
p_k	513.3	559.7	594.8	629.4	640.8	651.1	655.9	659.6	661.8	

6.6 原子弹爆炸的能量估计



案例精讲 6-1
原子弹爆炸的能量估计

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,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 19.2×10^3 t(10^3t 相当于 1 000 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 21×10^3 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.

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[69].让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 $[l]=\mathrm{L}$,$[m]=\mathrm{M}$,$[t]=\mathrm{T}$.而速度 v、加速度 a 的量纲可以按照其定义表为 $[v]=\mathrm{LT}^{-1}$,$[a]=\mathrm{LT}^{-2}$,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 $[f]=\mathrm{LMT}^{-2}$,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 $f=k\dfrac{m_1m_2}{r^2}$ 中引力常数 $k=\dfrac{fr^2}{m_1m_2}$，k 的量纲可从力 f、长度 r 和质量 m 的量纲得到：$[k]=\mathrm{LMT^{-2}\cdot L^2\cdot M^{-2}=L^3M^{-1}T^{-2}}$.对于无量纲的量 λ，记 $[\lambda]=\mathrm{L^0M^0T^0}=1$.

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[7,24].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)做往复摆动，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

$$t=\lambda m^{\alpha_1}l^{\alpha_2}g^{\alpha_3} \tag{1}$$

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

$$[t]=[m]^{\alpha_1}[l]^{\alpha_2}[g]^{\alpha_3} \tag{2}$$

将 $[t]=\mathrm{T},[m]=\mathrm{M},[l]=\mathrm{L},[g]=\mathrm{LT^{-2}}$ 代入得

$$\mathrm{T=M^{\alpha_1}L^{\alpha_2+\alpha_3}T^{-2\alpha_3}} \tag{3}$$

按照量纲齐次原则应有

$$\begin{cases}\alpha_1 & & & =0\\ & \alpha_2 & +\alpha_3 & =0\\ & & -2\alpha_3 & =1\end{cases} \tag{4}$$

(4)的解为 $\alpha_1=0,\alpha_2=1/2,\alpha_3=-1/2$，代入(1)式得

$$t=\lambda\sqrt{\frac{l}{g}} \tag{5}$$

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

$$f(t,m,l,g)=0 \tag{6}$$

这里没有因变量与自变量之分.进而假设(6)式形如

$$t^{y_1}m^{y_2}l^{y_3}g^{y_4}=\pi \tag{7}$$

其中 y_1,y_2,y_3,y_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 L,M,T 表示为 $[t]=\mathrm{L^0M^0T^1},[m]=\mathrm{L^0M^1T^0},[l]=\mathrm{L^1M^0T^0},[g]=\mathrm{L^1M^0T^{-2}}$，则(7)式的量纲表达式可写作

$$\mathrm{(L^0M^0T^1)^{y_1}(L^0M^1T^0)^{y_2}(L^1M^0T^0)^{y_3}(L^1M^0T^{-2})^{y_4}=L^0M^0T^0} \tag{8}$$

即

$$\mathrm{L^{y_3+y_4}M^{y_2}T^{y_1-2y_4}=L^0M^0T^0} \tag{9}$$

由量纲齐次原则有

---

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.

$$\begin{cases} y_3 + y_4 = 0 \\ y_2 = 0 \\ y_1 - 2y_4 = 0 \end{cases} \tag{10}$$

方程组(10)有一个基本解

$$\boldsymbol{y} = (y_1, y_2, y_3, y_4)^{\mathrm{T}} = (2, 0, -1, 1)^{\mathrm{T}} \tag{11}$$

代入(7)式得

$$t^2 l^{-1} g = \pi \tag{12}$$

(6)式可以等价地表示为

$$F(\pi) = 0 \tag{13}$$

(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 $q_1, q_2, \cdots, q_m$ 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

$$f(q_1, q_2, \cdots, q_m) = 0 \tag{14}$$

若基本量纲记作 $\mathrm{X}_1, \mathrm{X}_2, \cdots, \mathrm{X}_n (n \leqslant m)$,而 $q_1, q_2, \cdots, q_m$ 的量纲可表为

$$[q_j] = \prod_{i=1}^{n} \mathrm{X}_i^{a_{ij}}, j = 1, 2, \cdots, m \tag{15}$$

矩阵 $\boldsymbol{A} = (a_{ij})_{n \times m}$ 称量纲矩阵.若 $\boldsymbol{A}$ 的秩

$$\mathrm{Rank}\ \boldsymbol{A} = r \tag{16}$$

设线性齐次方程组

$$\boldsymbol{A}\boldsymbol{y} = \boldsymbol{0}, \boldsymbol{y} = (y_1, y_2, \cdots, y_m)^{\mathrm{T}} \tag{17}$$

的 $m-r$ 个基本解记作

$$y^{(s)} = (y_1^{(s)}, y_2^{(s)}, \cdots, y_m^{(s)})^{\mathrm{T}}, s = 1, 2, \cdots, m - r \tag{18}$$

则存在 $m-r$ 个相互独立的无量纲量

$$\pi_s = \prod_{j=1}^{m} q_j^{y_j^{(s)}}, s = 1, 2, \cdots, m - r \tag{19}$$

且

$$F(\pi_1, \pi_2, \cdots, \pi_{m-r}) = 0 \tag{20}$$

(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

$$r = \varphi(t, E, \rho, P) \tag{21}$$

更一般的形式记作

---

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.

$$f(r,t,E,\rho,P)=0 \tag{22}$$

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

$$[r]=\mathrm{L},[t]=\mathrm{T},[E]=\mathrm{L}^2\mathrm{MT}^{-2},[\rho]=\mathrm{L}^{-3}\mathrm{M},[P]=\mathrm{L}^{-1}\mathrm{MT}^{-2} \tag{23}$$

由此得到量纲矩阵

$$\boldsymbol{A}_{3\times5}=\begin{bmatrix}1&0&2&-3&-1\\0&0&1&1&1\\0&1&-2&0&-2\end{bmatrix} \tag{24}$$

因为$\boldsymbol{A}$的秩是3,所以齐次方程

$$\boldsymbol{A}\boldsymbol{y}=\boldsymbol{0},\boldsymbol{y}=(y_1,y_2,y_3,y_4,y_5)^{\mathrm{T}} \tag{25}$$

有5-3=2个基本解.

令$y_1=1,y_5=0$,得到一个基本解$\boldsymbol{y}=(1,-2/5,-1/5,1/5,0)^{\mathrm{T}}$;令$y_1=0,y_5=1$,得到另一个基本解$y=(0,6/5,-2/5,-3/5,1)^{\mathrm{T}}$.由这2个基本解可以得到2个无量纲量

$$\pi_1=rt^{-2/5}E^{-1/5}\rho^{1/5}=r\left(\frac{\rho}{t^2E}\right)^{1/5} \tag{26}$$

$$\pi_2=t^{6/5}E^{-2/5}\rho^{-3/5}P=\left(\frac{t^6P^5}{E^2\rho^3}\right)^{1/5} \tag{27}$$

且存在某个函数F,使得

$$F(\pi_1,\pi_2)=0 \tag{28}$$

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

$$r\left(\frac{\rho}{t^2E}\right)^{1/5}=\psi\left(\frac{t^6P^5}{E^2\rho^3}\right)^{1/5} \tag{29}$$

于是

$$r=\left(\frac{t^2E}{\rho}\right)^{1/5}\psi\left(\frac{t^6P^5}{E^2\rho^3}\right)^{1/5} \tag{30}$$

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,$\pi_2=\left(\frac{t^6P^5}{E^2\rho^3}\right)^{1/5}\approx0$.于是$\psi(0)$可看作一个比例系数$\lambda$,将(30)式记作

$$r=\lambda\left(\frac{t^2E}{\rho}\right)^{1/5} \tag{31}$$

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

$$E = \frac{\rho r^5}{t^2} \tag{32}$$

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 $E = 8.282\ 5 \times 10^{13}$ J.查表可知 10^3 t = 4.184×10^{12} J，所以爆炸能量是 $19.795\ 7 \times 10^3$ t，与实际值21×10^3 t相差不大(Taylor 是直接由(31)式作拟合，得到爆炸能量为 $19.186\ 3 \times 10^3$ 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

$$r = at^b \tag{33}$$

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 $b = 0.405\ 8$，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 1.

原子弹爆炸的能量估计被看作是量纲分析方法建模的一个成功范例.

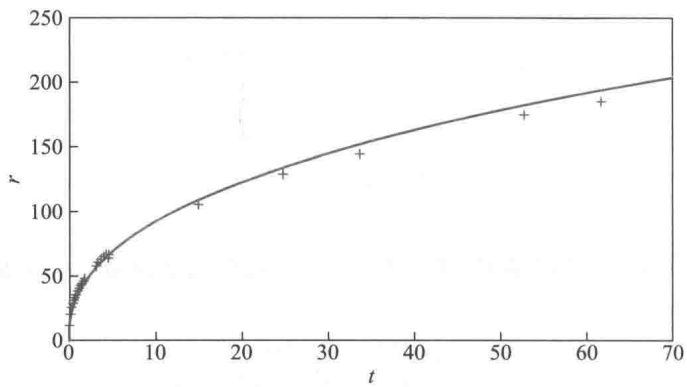


图 1 (33)式(曲线)与实际数据(+)的拟合

量纲分析方法的要点 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 $f(q_1, q_2, \cdots, q_m) = 0$，到它的等价形式 $F(\pi_1, \pi_2, \cdots, \pi_{m-r}) = 0$，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：

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 $q_1, q_2, \cdots, q_m$ 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，这不是唯一

评注 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

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多不同的构造方法,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 $\pi_3=r^3E^{-1}P(=\pi_1^3\pi_2)$,$\pi_4=r^2t^{-2}\rho P^{-1}(=\pi_1^2\pi_2^{-1})$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.

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.

复习题

1. 雨滴的速度 v 与空气密度 ρ、黏滞系数 μ 和重力加速度 g 有关,其中黏滞系数的定义是: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,比例系数为黏滞系数.用量纲分析方法给出速度 v 的表达式.

2. 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,黏性系数 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 $h=\frac{k}{d}\varphi\left(\frac{v\rho d}{\mu},\frac{\mu c}{k}\right)$,$\varphi$ 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 ℃时的热量交换[31].

3. 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 $f=\varepsilon v^2\varphi(s/d^2)$.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 $f=\frac{q_1q_2}{\varepsilon d^2}$,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[31].

6.7 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.Cormark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[28,76].

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.

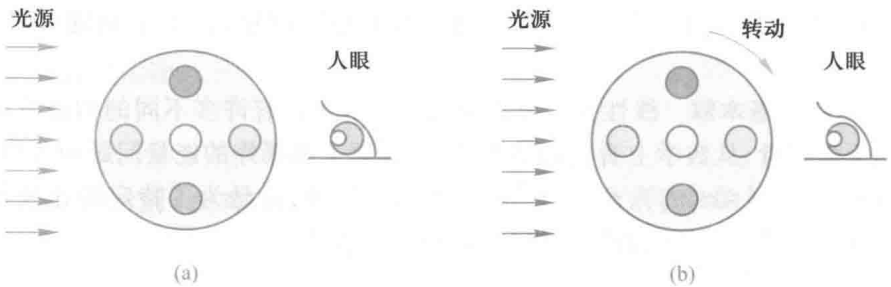


图 1 传统的 X 射线成像和 CT 的一般概念图示

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

$$\frac{\mathrm{d}I}{\mathrm{d}l} = -\mu I \tag{1}$$

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

$$I = I_0 \mathrm{e}^{-\mu l} \tag{2}$$

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

$$I = I_0 \mathrm{e}^{-\int_L \mu(x,y)\,\mathrm{d}l} \tag{3}$$

其中 $\int_L \mu(x,y)\,\mathrm{d}l$ 是 $\mu(x,y)$ 沿 L 的线积分,如图 2.由(3)可得

$$\int_L \mu(x,y)\,\mathrm{d}l = \ln \frac{I_0}{I} \tag{4}$$

(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 $\int_L \mu(x,y)\,\mathrm{d}l$,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.

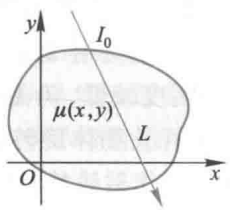


图 2 X 射线沿直线 L 穿过衰减系数为 $\mu(x,y)$ 的区域

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

$$P_f(L)=\int_L f(x,y)\,\mathrm{d}l \tag{5}$$

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

$$f(Q)=-\frac{1}{\pi}\int_0^{\infty}\frac{\mathrm{d}F_Q(q)}{q} \tag{6}$$

(5)、(6)式分别称 $f(x,y)$ 的 **Radon 变换**和逆变换①.

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.

设有 m 个像素(记 $j=1,2,\cdots,m$)，n 束射线(记 $i=1,2,\cdots,n$)，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：

$$\sum_{j\in J(L_i)}\mu_j\Delta l_j=\ln(I_0/I)_i,\ i=1,2,\cdots,n \tag{7}$$

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)

$$\sum_{j=1}^{m}a_{ij}x_j=b_i,\ i=1,2,\cdots,n \tag{8}$$

或写成向量-矩阵形式

$$\boldsymbol{Ax}=\boldsymbol{b} \tag{9}$$

其中 $\boldsymbol{A}=(a_{ij})_{n\times m}$，$\boldsymbol{x}=(x_1,x_2,\cdots,x_m)^{\mathrm{T}}$，$\boldsymbol{b}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}$.

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).

---

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.

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.

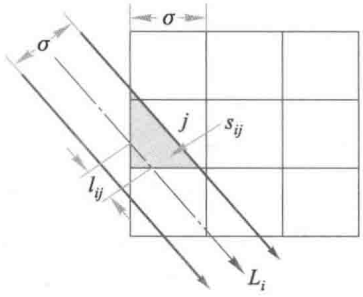


图3 射线束 L_i 经过像素 j 示意图
(像素边长和射线束宽度均为 σ)

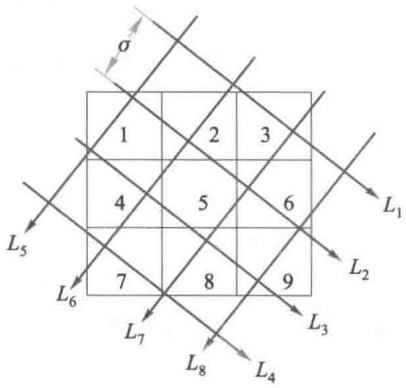


图4 射线 $L_1 \sim L_8$ 经过像素 1~9 示意图
(像素边长和射线间距均为 σ)

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

$$\boldsymbol{A}=\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

不论采用上面3种方法的哪一个，都可以利用代数方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 来重建图像，即根据已知的矩阵 $\boldsymbol{A}$ 和测量数据计算出的向量 $\boldsymbol{b}$，确定像素的衰减系数向量 $\boldsymbol{x}$.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

$$\boldsymbol{Ax}+\boldsymbol{e}=\boldsymbol{b} \tag{10}$$

其中 $\boldsymbol{e}$ 是误差向量.于是问题化为在 $\boldsymbol{x}$ 和 $\boldsymbol{e}$ 满足指定的最优准则下，根据 $\boldsymbol{A}$ 和 $\boldsymbol{b}$ 估计 $\boldsymbol{x}$，这一过程常用迭代算法进行，称为代数重建技术(ART)，详见[85].

6.8 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[5].

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,是一个确定性的转移问题.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.

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 $i=1,2,\cdots,k$,时间记作 $t=0,1,2,\cdots$.引入以下的定义和记号:

成员按等级的分布向量 $\boldsymbol{n}(t)=(n_1(t),n_2(t),\cdots,n_k(t))$,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 $\boldsymbol{a}(t)=(a_1(t),a_2(t),\cdots,a_k(t))$,其中 $a_i(t)=\dfrac{n_i(t)}{N(t)}$,于是有 $a_i(t)\geqslant 0$,$\sum_{i=1}^{k} a_i(t)=1$,$\boldsymbol{a}(t)$ 又称为**等级结构**.

转移比例矩阵 $\boldsymbol{Q}=\{p_{ij}\}_{k\times k}$,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 $\boldsymbol{w}=(w_1,w_2,\cdots,w_k)$,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

$$W(t)=\sum_{i=1}^{k} w_i n_i(t)=\boldsymbol{n}(t)\boldsymbol{w}^{\mathrm{T}} \tag{1}$$

容易看出,p_{ij},w_i 满足

$$p_{ij},w_i\geqslant 0,\ \sum_{j=1}^{k} p_{ij}+w_i=1 \tag{2}$$

调入比例向量 $\boldsymbol{r}=(r_1,r_2,\cdots,r_k)$,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 $r_iR(t)$.r_i 满足

$$r_i\geqslant 0,\quad \sum_{i=1}^{k} r_i=1 \tag{3}$$

等级结构的基本方程 为了导出成员按等级的分布 $\boldsymbol{n}(t)$ 的变化规律,先写出总人数 $N(t)$ 的方程

$$N(t+1)=N(t)+R(t)-W(t) \tag{4}$$

和每个等级人数的转移方程

$$n_j(t+1)=\sum_{i=1}^{k} p_{ij}n_i(t)+r_jR(t),\quad j=1,2,\cdots,k \tag{5}$$

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

$$\boldsymbol{n}(t+1)=\boldsymbol{n}(t)\boldsymbol{Q}+R(t)\boldsymbol{r} \tag{6}$$

从 t 到 $t+1$ 年总人数的增长量记为 $M(t)$,再由(4),(1)式可得

$$R(t)=W(t)+M(t)=\boldsymbol{n}(t)\boldsymbol{w}^{\mathrm{T}}+M(t) \tag{7}$$

将(7)代入(6)式得到

$$\boldsymbol{n}(t+1)=\boldsymbol{n}(t)(\boldsymbol{Q}+\boldsymbol{w}^{\mathrm{T}}\boldsymbol{r})+M(t)\boldsymbol{r} \tag{8}$$

记

$$\boldsymbol{P}=\boldsymbol{Q}+\boldsymbol{w}^{\mathrm{T}}\boldsymbol{r} \tag{9}$$

从(2),(3)式可知 $\boldsymbol{P}$ 的行和为1,(8)式记为

$$\boldsymbol{n}(t+1)=\boldsymbol{n}(t)\boldsymbol{P}+M(t)\boldsymbol{r} \tag{10}$$

当已知系统内部转移比例矩阵 $\boldsymbol{Q}$,调入比例 $\boldsymbol{r}$,初始的成员等级分布 $\boldsymbol{n}(0)$,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 $\boldsymbol{n}(t)$.(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 $\boldsymbol{a}(t)$ 代替 $\boldsymbol{n}(t)$ 而将(10)式表示为

$$\boldsymbol{a}(t+1)=(1+\alpha)^{-1}[\boldsymbol{a}(t)\boldsymbol{P}+\alpha\boldsymbol{r}] \tag{11}$$

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

$$\boldsymbol{a}(t+1)=\boldsymbol{a}(t)\boldsymbol{P}=\boldsymbol{a}(t)(\boldsymbol{Q}+\boldsymbol{w}^{\mathrm{T}}\boldsymbol{r}) \tag{12}$$

注意等级结构 $\boldsymbol{a}(t)$ 的转移矩阵 $\boldsymbol{P}$ 的 i,j 元素为 $p_{ij}+w_ir_j$,即系统内部转移比例 p_{ij} 加上系统内外交流的比例 w_ir_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 假定在实际问题中人们的理想等级结构是 $\boldsymbol{a}^*$,因为等级结构 $\boldsymbol{a}(t)$ 是按照(12)式的规律变化的,人们自然希望 $\boldsymbol{a}^*$ 一旦达到,就能够通过选取适当的调入比例使 $\boldsymbol{a}^*$ 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.我们的目的是:给定了内部转移比例矩阵 $\boldsymbol{Q}=\{p_{ij}\}$(由(2)式,退出向量 $\boldsymbol{w}$ 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 $\boldsymbol{a}$,如果存在调入比例 $\boldsymbol{r}$ 使得

$$\boldsymbol{a}=\boldsymbol{a}(\boldsymbol{Q}+\boldsymbol{w}^{\mathrm{T}}\boldsymbol{r}) \tag{13}$$

则称 $\boldsymbol{a}$ 为**稳定结构**,注意这里的 $\boldsymbol{r}$ 必须满足基本关系(3)式,即 $r_i\geqslant 0$, $\sum_{i=1}^{k} r_i=1$.由(13)式不难得到

$$\boldsymbol{r}=\frac{\boldsymbol{a}-\boldsymbol{a}\boldsymbol{Q}}{\boldsymbol{a}\boldsymbol{w}^{\mathrm{T}}} \tag{14}$$

可以验证(14)式给出的 $\boldsymbol{r}$ 满足 $\sum_{i=1}^{k} r_i=1$.为满足 $r_i\geqslant 0$ 的要求,可知稳定结构 $\boldsymbol{a}$ 的范围由

$$\boldsymbol{a}\geqslant\boldsymbol{a}\boldsymbol{Q} \tag{15}$$

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

$$\boldsymbol{Q}=\begin{bmatrix}0.5 & 0.4 & 0\\ 0 & 0.6 & 0.3\\ 0 & 0 & 0.8\end{bmatrix} \tag{16}$$

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 $\boldsymbol{a}$ 的稳定域.

将(16)代入(15)式得

$$\begin{cases}a_1 \geqslant 0.5a_1\\ a_2 \geqslant 0.4a_1+0.6a_2\\ a_3 \geqslant 0.3a_2+0.8a_3\end{cases}$$

即

$$\begin{cases}a_2 \geqslant a_1\\ a_3 \geqslant 1.5a_2\end{cases} \tag{17}$$

这就是 $\boldsymbol{a}$ 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 $\boldsymbol{a}=(a_1,a_2,a_3)$ 可看作三维空间的一个点,并且位于第一象限的平面

$$a_1+a_2+a_3=1,\quad a_1,a_2,a_3 \geqslant 0 \tag{18}$$

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 $\boldsymbol{a}$ 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).

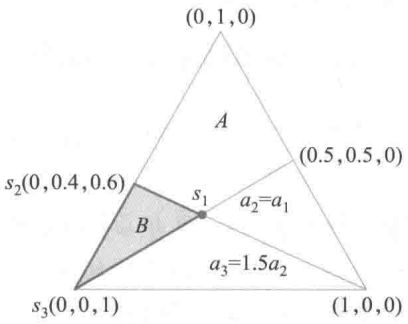


图1 教师等级结构的可行域和稳定域

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.

稳定域的构造 设至少有一个 i 使 $\sum_{j=1}^{k} p_{ij}<1$,则 $\boldsymbol{I}-\boldsymbol{Q}$ 可逆,记 $\boldsymbol{M}=\{m_{ij}\}=(\boldsymbol{I}-\boldsymbol{Q})^{-1}$,则由(14)式可得

$$\boldsymbol{a}=(\boldsymbol{a}\boldsymbol{w}^{\mathrm{T}})\boldsymbol{r}\boldsymbol{M} \tag{19}$$

记第 i 元素为1、其余元素为0的单位行向量为 $\boldsymbol{e}_i$,$\boldsymbol{r}$ 可表示为 $\boldsymbol{r}=\sum_{i=1}^{k} r_i\boldsymbol{e}_i$.又记 $\boldsymbol{M}$ 的第 i 行向量为 $\boldsymbol{m}_i=(m_{i1},m_{i2},\cdots,m_{ik})$,$\mu_i=\sum_{j=1}^{k} m_{ij}$,则 $\boldsymbol{r}\boldsymbol{M}=\sum_{i=1}^{k} r_i\boldsymbol{e}_i\boldsymbol{M}=\sum_{i=1}^{k} r_i\boldsymbol{m}_i$.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

$$\boldsymbol{a}\boldsymbol{w}^{\mathrm{T}}=\left(\sum_{j=1}^{k}r_j\mu_j\right)^{-1} \tag{20}$$

$$\boldsymbol{a}=\sum_{i=1}^{k}\frac{r_i\boldsymbol{m}_i}{\sum_{j=1}^{k}r_j\mu_j} \tag{21}$$

为了得到便于应用的结果,将(21)式进一步表示为

$$\boldsymbol{a}=\sum_{i=1}^{k}b_i\boldsymbol{s}_i,\quad b_i=\frac{r_i\mu_i}{\sum_{j=1}^{k}r_j\mu_j},\quad \boldsymbol{s}_i=\frac{\boldsymbol{m}_i}{\mu_i} \tag{22}$$

对于b_i,容易知道$\sum_{i=1}^{k}b_i=1$;又因为$\boldsymbol{M}$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$r_i\geqslant 0$时,$b_i\geqslant 0$.对于$\boldsymbol{s}_i=(s_{i1},s_{i2},\cdots,s_{ik})$,有$s_{ij}=\frac{m_{ij}}{\mu_i}$,所以$s_{ij}\geqslant 0$,$\sum_{j=1}^{k}s_{ij}=1$.上述分析表明:

当且仅当$\boldsymbol{a}$能够表示为以b_i为系数的$\boldsymbol{s}_i$的线性组合(22)式,且b_i满足

$$b_i\geqslant 0,\quad \sum_{i=1}^{k}b_i=1 \tag{23}$$

时,$\boldsymbol{a}$是稳定结构,即存在$\boldsymbol{r}$,满足$r_i\geqslant 0$,$\sum_{i=1}^{k}r_i=1$,使(13)式成立.

回到上面教师等级结构的例子,先从(16)式算出

$$\boldsymbol{M}=(\boldsymbol{I}-\boldsymbol{Q})^{-1}=\begin{bmatrix}2 & 2 & 3\\ 0 & 2.5 & 3.75\\ 0 & 0 & 5\end{bmatrix}$$

于是$\mu_1=7,\mu_2=6.25,\mu_3=5$.然后由(22)式得到$\boldsymbol{s}_1=(0.286,0.286,0.428)$,$\boldsymbol{s}_2=(0,0.4,0.6)$,$\boldsymbol{s}_3=(0,0,1)$.稳定结构$\boldsymbol{a}$可表示为$\boldsymbol{a}=b_1\boldsymbol{s}_1+b_2\boldsymbol{s}_2+b_3\boldsymbol{s}_3$,$b_1,b_2,b_3\geqslant 0$,$b_1+b_2+b_3=1$,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$\boldsymbol{Q}$的非零元素只能位于(16)式所示的位置时,$\boldsymbol{M}$必是上三角形矩阵,从而必然有$\boldsymbol{s}_3=(0,0,1)$,$\boldsymbol{s}_2=(0,*,*)$,$\boldsymbol{s}_1=(*,*,*)$(*表示非零元素).注意到$\boldsymbol{s}_3$表示全部是教授的职称结构,$\boldsymbol{s}_2$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$\boldsymbol{a}=(a_1,a_2,\cdots,a_k)$的可行域$A$是$k$维空间中由下式决定的$k-1$维超平面,

$$a_1+a_2+\cdots+a_k=1,\quad a_1,a_2,\cdots,a_k\geqslant 0 \tag{24}$$

而稳定域B则是A中以$\boldsymbol{s}_1,\boldsymbol{s}_2,\cdots,\boldsymbol{s}_k$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

评注 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

$$\boldsymbol{Q}=\begin{bmatrix} p_{11} & p_{12} & & 0 \\ & p_{22} & \ddots & \\ & & \ddots & p_{k-1,k} \\ 0 & & & p_{kk} \end{bmatrix} \tag{25}$$

形式时，B 的顶点为 $\boldsymbol{s}_k=(0,\cdots,0,1)$，$\boldsymbol{s}_{k-1}=(0,\cdots,0,*,*)$，$\cdots$，$\boldsymbol{s}_2=(0,*,\cdots,*)$，$\boldsymbol{s}_1=(*,*,\cdots,*)$.

调入比例固定为 $\boldsymbol{r}=(1,0,\cdots,0)$. 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，因为控制变量多了，调节更为灵活，计算也更为复杂.

等级结构的动态调节 实际上人们关心的另一个问题可能是，给定转移比例矩阵 $\boldsymbol{Q}$ 和初始结构 $\boldsymbol{a}(0)$，怎样通过一系列的调入比例 $\boldsymbol{r}$，使等级结构 $\boldsymbol{a}(t)$ 达到或逐渐接近某个理想结构 $\boldsymbol{a}^*$（可以合理地假设 $\boldsymbol{a}^*$ 在稳定域 $\boldsymbol{B}$ 内部）. 在这个过程中调入比例 $\boldsymbol{r}$ 可以变化，称为动态调节. 具体做法可参考拓展阅读 6-2.

复习题

设某校每年助教、讲师和教授 3 个等级之间的转移矩阵如(16)式，目前 3 个等级人数的比例为6∶3∶1.

(1) 若每年只调入助教，问 5 年后各等级人数比例是多少.

(2) 若每年调入 3 个等级的人数之比为 7∶1∶2，问 5 年后各等级人数比例是多少.

(3) (1)和(2)得到的 5 年后等级人数比例是否是稳定结构？

拓展阅读 6-2
等级结构的动态调节

6.9 中国人口增长预测

中国是一个人口大国，人口问题始终是制约我国发展的关键因素之一. 2007 年公布的《国家人口发展战略研究报告》提出，如果我国人口总量峰值 2033 年前后达到峰值 15 亿人左右，全国总和生育率应保持在 1.8 左右，过高或过低都不利于人口与经济社会的协调发展①.

在全国大学生数学建模竞赛 2007 年 A 题“中国人口增长预测”中提出：近年来中国的人口发展出现了一些新的特点，例如，老龄化进程加速、出生人口性别比持续升高，以及乡村人口城镇化等因素，这些都影响着中国人口的增长. 要求参赛者从中国的实际情况和人口增长的特点出发，参考 2005 年人口抽样数据（见表 1），建立中国人口增长的数学模型，并对人口增长的中短期和长期趋势做出预测.

本节以发表在《工程数学学报》2007 年增刊二上参赛学生的优秀论文为基本材料，加以整理、归纳和提高，介绍建模的全过程[70].

① 2015 年中共中央公布“全面二孩”政策，对实施多年的计划生育策略做出重大调整. 据报道，国家发改委正在酝酿新一轮的《国家人口发展战略研究报告》的编写.



数据文件 6-1
中国人口增长

表 1 2005 年人口抽样的部分数据①

	年龄	0	1	…	20	21	…	90+
市	男性比例	0.46	0.46	…	0.69	0.69	…	0.03
	男死亡率	6.38	0.84	…	0.59	0.43	…	289.2
	女性比例	0.4	0.39	…	0.74	0.76	…	0.07
	女死亡率	6.08	0.76	…	0.06	0.26	…	269.87
镇	男性比例	0.56	0.58	…	0.52	0.47	…	0.03
	男死亡率	0.91	0.93	…	0.73	1.31	…	253.8
	女性比例	0.48	0.47	…	0.51	0.53	…	0.07
	女死亡率	10.79	0.54	…	0.28	0.42	…	172.6
乡	男性比例	0.65	0.65	…	0.53	0.49	…	0.03
	男死亡率	13.97	1.45	…	1.91	1.75	…	355.2
	女性比例	0.54	0.51	…	0.53	0.52	…	0.08
	女死亡率	18.55	1.96	…	0.82	0.78	…	237
妇女生育率	市			…	29.01	48.17	…	
	镇			…	68.23	113.45	…	
	乡			…	95.01	150.84	…	

问题分析 用数学建模预测人口增长的方法主要有差分方程、微分方程、回归分析、时间序列等,结合题目所给数据,本节以差分方程组形式表示的 Leslie 模型为基础,考虑不同地区、不同性别人口参数的差别,以及农村人口向城市的迁移等因素,按照地区和性别,建立以时间和年龄为基本变量的中国人口增长模型.利用历史数据用统计方法估计生育率、死亡率及人口迁移等参数,代入模型迭代求解并做出预测.

模型假设 针对问题要求和所给数据作以下简化假设:

① 本数据表示 2005 年的市、镇和乡人口不同性别的人在该类人口中所占的百分比(%),死亡率为所占该类年龄段人口的千分比(‰),生育率为该类年龄段妇女生育子女的千分比(‰).如 2005 年 20 岁的城市男性占城市总人口的 0.69 %,其死亡率为 20 岁城市男性总人口的 0.59‰, 2005 年 20 岁的城市女性占城市总人口的 0.74 %,其死亡率为 20 岁城市女性总人口的 0.06‰, 2005 年 20 岁城市生育子女的妇女占 20 岁城市妇女总人口的 29.01 ‰,2005 年 20 岁乡村生育子女的妇女占 20 岁乡村妇女总人口的 95.01‰.其余类推.本表取自《中国人口统计年鉴》(1995—2006),每年年鉴所载数据为上一年数据.个别数据有异常,原文如此,可酌情处理.

1. 中国人口系统是封闭的，将表1数据中的市、镇合并为城市，与农村（乡）作为两个地区，只考虑农村向城市人口的单向迁移，不考虑与境外的相互移民.

2. 作中短期人口预测时，生育率、死亡率及人口迁移等参数用历史数据估计，作长期人口预测时，考虑总和生育率的控制、城镇化指数的变化趋势等政策因素.

3. 女性每胎生育一个子女.

模型建立　以1年为一个时段，1岁为一个年龄段，记 $x_j^l(i,k)$ 为 k 年、i 岁（满 i 岁但不到 $i+1$ 岁）、地区 j（$j=1$ 为城市，$j=2$ 为农村）、性别 l（$l=m$ 为男性，$l=w$ 为女性）的人口数量，$k=0,1,2,\cdots$，$i=0,1,2,\cdots,n$，n 为最高年龄. $d_j^l(i,k)$ 为人口死亡率（死亡人数占总人数的比例），$s_j^l(i,k)=1-d_j^l(i,k)$ 为存活率. $b_j^w(i,k)$ 为女性人口生育率（k 年每位 i 岁女性平均生育婴儿数），$[i_1,i_2]$ 为育龄区间，$a_j(k)$ 为出生婴儿性别比（男婴占婴儿总数的比例）. $v_j^l(i,k)$ 为人口的迁移数量（迁入为正、迁出为负）. 对于 $k=0,1,2,\cdots$，$j=1,2$，$l=m$，w，一般模型可写作

$$\begin{cases} x_j^m(1,k+1)=s_j^m(0,k)a_j(k)\sum\limits_{i=i_1}^{i_2}b_j^w(i,k)x_j^w(i,k)+v_j^m(0,k) \\ x_j^w(1,k+1)=s_j^w(0,k)(1-a_j(k))\sum\limits_{i=i_1}^{i_2}b_j^w(i,k)x_j^w(i,k)+v_j^w(0,k) \\ x_j^l(i+1,k+1)=s_j^l(i,k)x_j^l(i,k)+v_j^l(i,k),i=1,2,\cdots,n-1,l=m,w \end{cases} \tag{1}$$

对女性生育率 $b_j^w(i,k)$ 作进一步分析. 记

$$\beta_j^w(k)=\sum_{i=i_1}^{i_2}b_j^w(i,k) \tag{2}$$

$\beta_j^w(k)$ 为 k 年每位育龄女性平均生育婴儿数，如果所有女性在育龄期间都保持这个生育数，那么 $\beta_j^w(k)$ 也表示每位女性一生的平均生育数，称总和生育率或生育胎次，是表述与控制人口增长的重要指标. 记

$$h_j^w(i,k)=\frac{b_j^w(i,k)}{\beta_j^w(k)} \tag{3}$$

表示 i 岁女性的生育数在育龄女性中占的比例，称生育模式，满足 $\sum\limits_{i=i_1}^{i_2}h_j^w(i,k)=1$.

记 $v^l(i,k)$ 为农村向城市迁移的人口数量，按照农村向城市单向迁移的假设，有 $v_1^l(i,k)=v^l(i,k)$，$v_2^l(i,k)=-v^l(i,k)$. 记

$$\begin{aligned} \boldsymbol{x}_j^l(k)&=[x_j^l(1,k),x_j^l(2,k),\cdots,x_j^l(n,k)]^{\mathrm{T}} \\ \boldsymbol{v}^l(k)&=[v^l(0,k),v^l(1,k),\cdots,v^l(n-1,k)]^{\mathrm{T}} \end{aligned} \tag{4}$$

$$\boldsymbol{S}_j^l(k)=\begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ s_j^l(1,k) & 0 & \cdots & 0 & 0 \\ 0 & s_j^l(2,k) & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & s_j^l(n-1,k) & 0 \end{bmatrix} \tag{5}$$

$$\boldsymbol{H}_j^w(k)=\begin{bmatrix} 0 & \cdots & 0 & h_j^w(i_1,k) & \cdots & h_j^w(i_2,k) & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{6}$$

上述模型可表为如下向量-矩阵形式

$$\begin{cases} \boldsymbol{x}_1^m(k+1)=\boldsymbol{S}_1^m(k)\boldsymbol{x}_1^m(k)+s_1^m(0,k)a_1(k)\beta_1^w(k)\boldsymbol{H}_1^w(k)\boldsymbol{x}_1^w(k)+\boldsymbol{v}^m(k) \\ \boldsymbol{x}_1^w(k+1)=\boldsymbol{S}_1^w(k)\boldsymbol{x}_1^w(k)+s_1^w(0,k)(1-a_1(k))\beta_1^w(k)\boldsymbol{H}_1^w(k)\boldsymbol{x}_1^w(k)+\boldsymbol{v}^w(k) \\ \boldsymbol{x}_2^m(k+1)=\boldsymbol{S}_2^m(k)\boldsymbol{x}_2^m(k)+s_2^m(0,k)a_2(k)\beta_2^w(k)\boldsymbol{H}_2^w(k)\boldsymbol{x}_2^w(k)-\boldsymbol{v}^m(k) \\ \boldsymbol{x}_2^w(k+1)=\boldsymbol{S}_2^w(k)\boldsymbol{x}_2^w(k)+s_2^w(0,k)(1-a_2(k))\beta_2^w(k)\boldsymbol{H}_2^w(k)\boldsymbol{x}_2^w(k)-\boldsymbol{v}^w(k) \end{cases} \tag{7}$$

(7)式是按地区和性别划分、以年龄为离散变量、随时段演变的人口发展模型，其右端包括从 k 年到 $k+1$ 年人口发展的3个因素：由存活率构成的矩阵 $\boldsymbol{S}_j^l(k)$ 表述人口随年龄增长的演变；由总和生育率、生育模式构成的 $\beta_j^w(k)$ 和矩阵 $\boldsymbol{H}_j^w(k)$ 表述人口生育形成的演变；由农村向城市的人口迁移向量 $\boldsymbol{v}^l(k)$ 引起的演变.

(7)式是 $4n$ 阶差分方程组，因生育过程使得女性人口出现在男性人口的方程中，形成了耦合作用.

参数估计 在利用模型(7)按照时段(年)进行递推，预测人口的演变过程时，应先确定存活率、生育率及迁移向量等参数，并且需要根据不同地区、性别和各年龄人口在 k 年的这些参数及人数，才能计算 $k+1$ 年的人口数量.所以为了预测人口的增长，需要先估计这些参数在未来时段的数值.

存活率的估计 从常识看，死亡率与年龄的关系很大，与地区、性别和时间的关系较小.中国过去几十年死亡率降低较快，从长期看仍需考虑死亡率的持续下降，而短期则可近似地认为不变.所以在作中短期预测时，可以将过去若干年(如数据文件6-1给出的2001—2005年)不同地区、性别和各年龄人口的死亡率简单地加以平均，作为模型中的死亡率参数.作长期预测时，则可利用统计方法对历史数据加以处理，并参考发达国家人口死亡率的演变过程给出估计值.

存活率按照定义由死亡率算出.

生育率的估计 从20世纪80年代以来，随着计划生育政策的实施，中国女性生育率已在持续下降以后大致保持稳定，所以在作中短期预测时也可以将过去若干年(如数据文件6-1给出的2001—2005年)不同地区、各年龄女性人口的生育率简单地加以平均，作为模型中的生育率参数，这时应将模型(7)中的 $\beta_j^w(k)h_j^w(i,k)$ 还原回(1)式的 $b_j^w(i,k)$(参见(3)式).

作长期预测时，不妨设定几个不同水平的总和生育率 β_j^w(不考虑随时段 k 的变化)①，分别计算.至于生育模式 $h_j^w(i)$(也不考虑 k 的变化)，可以采用概率论中的伽马分布

---

① 《工程数学学报》2007年增刊二上郭媛等的论文给出总和生育率的以下几个数值：保持数据文件6-1的近年的数据1.375，2007年A题附录1的建议值1.8，生育更替水平值2.1，调查得到的意愿值(农村1.8，城市1.56).

$$h_j^w(i)=\frac{1}{2^{n/2}\Gamma(n/2)}(i-i_1)^{(n/2-1)}e^{-(i-i_1)/2},n=i_c-i_1+2,i_1\leqslant i\leqslant i_2 \tag{8}$$

其中 i_c 是生育高峰的年龄.如根据数据文件6-1给出的数据,有 $i_1=15,i_2=49$,城市的 i_c 为25,农村的 i_c 为23.于是对城市可设 $h_1^w(i)=\frac{1}{7\,680}(i-15)^5e^{-(i-15)/2}(15\leqslant i\leqslant 49)$①.

模型(7)中的出生婴儿性别比可由数据文件6-1给出的数据中0岁男婴和女婴的比率得出,或者参考近年公开发表的出生婴儿性别比资料.

人口迁移的估计 需要从两个方面进行估计,一是每年农村向城市迁移的人口总数,可以按照城镇化率的增长来估算,二是迁移人口的年龄、性别分布.

记 $X_1(k),X_2(k),X(k)$ 分别为 k 年全国的城镇人口、农村人口和总人口,则

$$X_j(k)=\sum_{l=m,w}\sum_{i=0}^{n}x_j^l(i,k),\quad X(k)=X_1(k)+X_2(k) \tag{9}$$

城镇化率定义为城镇人口占全国人口的比例,记作 $\alpha(k)=\frac{X_1(k)}{X(k)}$.$\alpha(k)$ 可以用合适的模型描述,并根据历史数据对 $\alpha(k)$ 进行预测②.k 年农村向城市迁移的人口总数记作 $V(k)=\sum_{l=m,w}\sum_{i=0}^{n}v^l(i,k)$,可用城镇化率的增长估计为

$$V(k)=[\alpha(k)-\alpha(k-1)]X(k) \tag{10}$$

对于迁移人口的年龄、性别分布,因为难以得到全国的数据,可以用典型城市的资料代替③,得到 i 岁、性别 l 迁移人口占总数的比例 $c^l(i,k)$,则农村向城市迁移的人口数量 $v^l(i,k)$ 为

$$v^l(i,k)=c^l(i,k)V(k),\sum_{l=m,w}\sum_{i=0}^{n}c^l(i,k)=1 \tag{11}$$

模型求解 在确定存活率、生育率及迁移向量等参数后,选定初始年份(如2005年)即可利用模型(7)进行递推,对于任意年份 k,得到 i 岁、地区 j、性别 l 的人口数量 $x_j^l(i,k)$,它全面、完整地描述了人口的演变过程.而为了简明、方便的需要,通常采用一些人口指数来集中反映一个国家或地区的人口特征,主要有:

- **人口总数** 记作 $X(k)$,见(9)式.
- **平均年龄** 记作 $R(k)$,是年龄 i 对人口数量 $x_j^l(i,k)$ 的加权平均,即

$$R(k)=\frac{1}{X(k)}\sum_{l=m,w}\sum_{j=1}^{2}\sum_{i=0}^{n}ix_j^l(i,k) \tag{12}$$

- **平均寿命** 记作 $S(k)$,是按 k 年的死亡率计算的 k 年出生人口的平均存活时间,有

$$S(k)=\sum_{l=m,w}\sum_{j=1}^{2}\sum_{r=0}^{n}\exp\left(-\sum_{i=0}^{r}d_j^l(i,k)\right) \tag{13}$$

评注 差分方程组形式的Leslie模型简明地表述了人口发展过程中生育和死亡这两个自然因素的作用.死亡基本上是不可控制的,在稳定环境下死亡率变化很慢,所以用历史数据能够较

① 《工程数学学报》2007年增刊二上杨海滨等的论文做了计算.

② 《工程数学学报》2007年增刊二上陈鹏等的论文建立了logistic模型.

③ 《工程数学学报》2007年增刊二上陈鹏等的论文取上海市2000年迁移人口中各年龄段所占比例为代表数据,用正态分布密度函数作拟合.

好地对中短期预测做出估计.生育在一定意义上可以控制,作预测时应该对总和生育率和生育模式设定几种方案,如当前水平和模式、理想水平和模式等,得到不同的预测结果,进行分析、比较.人口迁移是比较复杂的因素,历史数据也不多、不完整,上面模型中的人口迁移部分比较粗糙.如果将市、镇分开,与乡组成三个地区,模型结构没有什么本质改变,只是从乡到市、镇的人口迁移如何分配,以及怎样考虑市、镇之间的人口迁移,都更加缺乏数据的支持.

● **老龄化指数** 记作 $\omega(k)$,是平均年龄与平均寿命之比,即

$$\omega(k)=\frac{R(k)}{S(k)} \tag{14}$$

● **抚养指数** 记作 $\rho(k)$,是每个劳动力人口平均抚养的(无劳动力)人数.设劳动力年龄区间男性为$[i_{m1},\ i_{m2}]$,女性为$[i_{w1},\ i_{w2}]$,记劳动力人数为$L(k)$,则

$$L(k)=\sum_{j=1}^{2}\left(\sum_{i=i_{m1}}^{i_{m2}}x_j^m(i,k)+\sum_{i=i_{w1}}^{i_{w2}}x_j^w(i,k)\right),\ \rho(k)=\frac{X(k)-L(k)}{L(k)} \tag{15}$$

其他如60岁、65岁以上人口比例等可以类似于(15)式计算.

表2是发表在《工程数学学报》2007年增刊二上优秀论文关于人口总数的部分计算结果①.其他人口指数的计算结果参看上述期刊的论文.

表2 人口总数(亿)的部分计算结果

		年份				
		2010	2020	2030	2040	2050
1	$\beta=1.375$	13.2	13.5	13.2	12.4	11.4
2	$\beta=1.8$	13.4	14.2	14.2	13.8	12.7
3	$\beta=2.1$	13.6	14.4	14.7	14.7	14.7
4		13.5	14.3	14.8	14.2	13.9

第6章训练题

1. 一老人60岁时将养老金10万元存入基金会,月利率0.4%,他每月取1 000元做生活费,建立差分方程模型计算他每年末尚有多少钱?多少岁时将基金用完?如果想用到80岁,60岁时应存多少钱?

2. 在一处古篝火遗址附近发现了一个猿人头骨,考古学家确信该头骨和古篝火是同时代的.经测试确定在取自篝火的灰烬中,碳14的含量仅留存原来的1%.已知碳14的半衰期是5 730年.建立一个用碳14测定年代的模型,并用以确定古篝火遗址附近发现的猿人头骨的年代[24].

3. 某湖泊每天有$10^4\ m^3$的河水流入,河水中污物浓度为$0.02\ g/m^3$,经渠道排水后湖泊容积保持$200\times10^4\ m^3$不变,现测定湖泊中污物浓度为$0.2\ g/m^3$,建立模型计算湖泊中一年内逐月(每月按30天计)下降的污物浓度,问要多长时间才能达到环保要求的浓度$0.04\ g/m^3$?为了把这个时间缩短为一年,应将河水中污物浓度降低到多少[73]?

4. 当某种药物的浓度大于100 mg/L(有效水平)时才能治疗疾病,且最高浓度不能超过1 000 mg/L(安全水平).从实验知道该药物浓度以每小时现有量15%的速率衰减.

(1) 建立每小时药物浓度的模型.若初始浓度为600 mg/L,确定何时浓度达到100 mg/L.

(2) 设计一个合适的药物处方(包括初始剂量及维持剂量和间隔).

5. 设在一个受限环境中鲸鱼的最大容量是M,最小生存水平是m,用a_n表示n年后的鲸鱼数量,

① 见《工程数学学报》2007年增刊二上的优秀论文,表2中1~3引自郭媛等的论文,4引自杨海滨等的论文.

建立关于 a_n 的差分方程模型.再设 $M=5\ 000$,$m=100$,对模型参数和初始鲸鱼数量取不同数值作计算,分析对结果的影响.

6. 斑点猫头鹰和隼在同一栖息地为生存而竞争.假定只有一个种群存在时可以无限增长,即种群增量与此时的种群数量成正比.而种群的竞争使每个种群增量的减少与两个种群的数量有关.建立猫头鹰和隼数量变化的差分方程模型,改变模型参数进行计算,分析对结果的影响.用这种模型描述的两个相互竞争种群的后果是什么,怎样改进模型才能反映现实世界中在同一栖息地只有一个种群能够生存的现象[24].

7. 据报道,某种山猫在较好、中等及较差的自然环境下,年平均增长率分别为 1.68%, 0.55%和 -4.50%,假定开始时有 100 只山猫,按以下情况讨论山猫数量逐年变化过程及趋势:

(1) 3 种自然环境下 25 年的变化过程(作图).

(2) 如果每年捕获 3 只,会发生什么情况? 山猫会灭绝吗? 如果每年只捕获 1 只呢?

(3) 在较差的自然环境下,如果想使山猫数量稳定在 60 只左右,每年要人工繁殖多少只[73]?

8. 在某种环境下猫头鹰的主要食物来源是田鼠,设田鼠的年平均增长率为 r_1,猫头鹰的存在引起的田鼠增长率的减少与猫头鹰的数量成正比,比例系数为 a_1;猫头鹰的年平均减少率为 r_2;田鼠的存在引起的猫头鹰减少率的增加与田鼠的数量成正比,比例系数为 a_2.建立模型描述田鼠和猫头鹰共处时的数量变化规律,对以下情况作图给出 50 年的变化过程.

(1) 设 $r_1=0.2$,$r_2=0.3$,$a_1=0.001$,$a_2=0.002$,开始时有 100 只田鼠和 50 只猫头鹰.

(2) r_1,r_2,a_1,a_2同上,开始时有 100 只田鼠和 200 只猫头鹰.

(3) 适当改变参数 a_1,a_2(初始值同上).

(4) 求差分方程的平衡点,它们稳定吗[24]?

9. 研究将鹿群放入草场后草和鹿两种群的相互作用.草的生长遵从 logistic 规律,年固有增长率 0.8,最大密度为 3 000(密度单位),在草最茂盛时每只鹿每年可吃掉 1.6(密度单位)的草.若没有草,鹿群的年死亡率高达 0.9,而草的存在可使鹿的死亡得以补偿,在草最茂盛时补偿率为 1.5.做出一些简化假设,用差分方程模型描述草和鹿两种群数量的变化过程,就以下情况进行讨论.

(1) 比较将 100 只鹿放入密度为 1 000 和密度为 3 000 的草场两种情况.

(2) 适当改变参数,观察变化趋势[73].

10. 在 6.1 节的两种还款方式下,如果将每月还款一次的办法改为每年还款一次(每年的最后一月还款),其他条件均不变,比较两种办法的还款额.

11. 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1/1 000.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[31].

12. 牧场管理[29].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 · 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.

更多案例……

6-1 投入产出模型

6-2 量纲分析的应用与无量纲化

第7章　离散模型

所谓离散模型，一般是相对于连续模型而言的，前几章中的差分方程相对于微分方程、整数规划相对于线性规划，是常见的离散模型.本章的建模案例主要取自经济、社会等领域经常出现的决策、排序、分配等方面的问题.

7.1　汽车选购

选购一辆私家车是许多进入稳定社会生活的人们要费心考虑的事情之一.由于经济情况、生活习惯、兴趣追求等方面的差别，他(她)们选购汽车的标准自然不会相同.可以认为主要会考虑经济适用、性能良好、款式新颖3个因素，只是每个人对这3个因素的侧重有所不同.初入社会的年轻人可能以经济适用为重，有一定经济实力的中年人更注重性能良好，所谓“富二代”则更钟情于款式新颖.如果某人对3个因素在汽车选购这一目标中的重要性已经有了大致的比较，也确定了待选的若干种型号的汽车，那么他(她)必然要深入了解每一种待选的汽车，以便对各种汽车在每个因素中的优劣程度做出基本的判断.最后，他(她)要根据以上信息对待选汽车进行综合评价，从而为选购哪种汽车做出决策.

人们在日常生活中常常会碰到类似的决策问题：假期旅游，是去风光绮丽的苏杭，还是迷人的北戴河海滨，抑或山水甲天下的桂林，这与旅游地的景色、旅途的费用、吃住条件等因素在你心目中的重要程度有关；选择工作岗位，是国企、私企还是科研院所，当然会考虑薪酬、地域、发展前景等方面的因素.

从事各种职业的人在工作中也经常面对决策：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；经理要从应聘者中选拔秘书；各地区、各部门的官员要对经济、环境、交通、居住等方面的发展做出规划.

现实世界里诸如此类的众多决策问题有一些共同的特点：需要考虑的因素经常涉及经济、社会、人文、环境等领域，对它们的重要性、影响力作比较、评价时缺乏客观的标准；待选对象对于这些因素的优劣程度也往往难以量化.这就给用数学建模解决一大类实际问题带来困难.**多属性决策**(multiple attribute decision making)和**层次分析法**(analytic hierarchy process)是处理这类决策问题的常用方法.本节以汽车选购为例介绍用多属性决策建模的全过程.[30,81]

多属性决策的几个要素

多属性决策指人们为了一个特定的目的，要在若干备选方案(如几种型号的汽车)中确定一个最优的，或者对这些方案按照优劣程度排序，或者需要给出优劣程度的数量结果，而方案的优劣由若干属性(如汽车的价格、性能、款式等因素)给以定量或定性的表述.一般地说，多属性决策包含以下要素：

- **决策目标、备选方案与属性集合**　通常，决策目标和备选方案是由实际问题本身决定的，少有选择的余地.而选取哪些属性对于决策结果则至关重要，需要决策者在

相关领域有较多的实际经验.大体上,确定属性集合有以下几条原则:全面考虑影响决策目标的因素,注意选取影响力(或重要性)较强的属性;各个属性之间应尽量独立,至少相关性不要太强;尽量选取能够定量的属性,定性的也要能分出明确的优劣程度;如果某个属性对各备选方案的差别很小,根据该属性就难以辨别方案的优劣,那么这个属性就不必选入(即使它对决策目标的影响力很强);当属性数量太多时应该将它们分层,上层的每一属性包含下层的若干子属性.

● **决策矩阵** 指以方案为行、属性为列,以每一方案对每一属性的取值为元素形成的矩阵,表示了方案对属性的优劣(或偏好)程度.当某一属性可以定量描述时(如汽车价格),矩阵这一列元素的数值比较容易得到,而当某一属性只能定性描述时(如汽车款式),这一列元素的数值就需要寻求合适的方法确定.

● **属性权重** 各属性之间的权重分配对于决策结果至关重要,例如当今媒体上出现的不同机构给出的大学排名榜有不小的差别,在很大程度上是由于对属性(如教师水平、毕业生质量、研究经费等)的选取及其权重分配的不同造成的.同时,由于各属性的不同特性,常常难以客观地、定量地确定其权重.

● **综合方法** 指将决策矩阵与属性权重加以综合,得到最终决策的数学方法.

下面结合汽车选购说明如何确定决策矩阵、属性权重以及利用综合方法得到决策结果.

决策矩阵及其标准化

假定有3种型号的汽车(相当于3个方案)供选购,记作 A_1, A_2, A_3,3个属性为价格、性能和款式,依次记作 X_1, X_2, X_3.对于价格 X_1(单位:万元),3种汽车分别为25, 18, 12;对于性能 X_2, 采用打分的办法(满分10分),3种汽车分别为9, 7, 5;对于款式 X_3, 类似地打分为7, 7, 5.将这些数值列入表1,并记作 $d_{ij}(i=1,2,3,j=1,2,3)$,表示方案 A_i 对属性 X_j 的取值,或称原始权重.

表1 汽车采购中的原始权重 $d_{ij}(i=1,2,3,j=1,2,3)$

	X_1	X_2	X_3
A_1	25	9	7
A_2	18	7	7
A_3	12	5	5

一般地,如果一个多属性决策问题有 m 个备选方案 A_1, A_2, $\cdots$, A_m, n 个属性 X_1, X_2, $\cdots$, X_n,方案 A_i 对属性 X_j 的取值为 d_{ij}, 以下称属性值,$\boldsymbol{D}=(d_{ij})_{m\times n}$ 称决策矩阵,表1可以用(原始)决策矩阵表示为

$$\boldsymbol{D}=\begin{bmatrix}25 & 9 & 7\\ 18 & 7 & 7\\ 12 & 5 & 5\end{bmatrix}$$

决策矩阵的获取一般有两种途径,一种是直接通过量测或调查得到,如表1中价格 X_1,是偏于客观的方法,另一种是由决策者或请专家评定,是偏于主观的方法.

决策矩阵的每一列表示各方案对某一属性的属性值,由于通常各属性的物理意义

(包括量纲)各不相同,在下一步分析之前常需要将决策矩阵标准化(或称规范化).

进行标准化时首先需要区分**效益型属性**和**费用型属性**,前者指属性值越大,该属性对决策的重要程度越高,后者正相反.汽车选购中的属性 X_2, X_3是效益型的,而 X_1是费用型的.如果一个多属性决策问题中效益型属性占多数(大多数实际情况如此),在标准化时应先对费用型属性值作变换,通常的方法是取原属性值的倒数,或者用一个大数减去原属性值,将全部属性统一为效益型的①.用取倒数的方法可将汽车选购中的决策矩阵重新表示为

$$\boldsymbol{D}=\begin{bmatrix}\dfrac{1}{25} & 9 & 7\\ \dfrac{1}{18} & 7 & 7\\ \dfrac{1}{12} & 5 & 5\end{bmatrix}$$

以下所说的决策矩阵 $\boldsymbol{D}$ 均对统一为效益型属性而言.

所谓决策矩阵标准化是对 $\boldsymbol{D}=(d_{ij})_{m\times n}$作比例尺度变换,标准化的决策矩阵 $\boldsymbol{R}=(r_{ij})_{m\times n}$有以下几种形式:

$$r_{ij}=\frac{d_{ij}}{\sum\limits_{i=1}^{m} d_{ij}} \tag{1}$$

$\boldsymbol{R}$ 的列向量的分量之和为 1,称归一化;

$$r_{ij}=\frac{d_{ij}}{\max\limits_{i=1,2,\cdots,m} d_{ij}} \tag{2}$$

$\boldsymbol{R}$ 的列向量的分量最大值为 1,称最大化;

$$r_{ij}=\frac{d_{ij}}{\sqrt{\sum\limits_{i=1}^{m} d_{ij}^2}} \tag{3}$$

$\boldsymbol{R}$ 的列向量的模为 1,称模一化.经过这些变换可得 $0\leqslant r_{ij}\leqslant 1$,当且仅当 $d_{ij}=0$ 时才有 $r_{ij}=0$.$\boldsymbol{R}$ 的各个列向量表示了在同一尺度下各属性的属性值.

汽车选购中的决策矩阵 $\boldsymbol{D}$ 经过(1),(2),(3)式标准化后分别化为

$$\boldsymbol{R}=\begin{bmatrix}0.223\,6 & 0.428\,6 & 0.368\,4\\ 0.310\,6 & 0.333\,3 & 0.368\,4\\ 0.465\,8 & 0.238\,1 & 0.263\,2\end{bmatrix}$$

$$\boldsymbol{R}=\begin{bmatrix}0.480\,0 & 1.000\,0 & 1.000\,0\\ 0.666\,7 & 0.777\,8 & 1.000\,0\\ 1.000\,0 & 0.555\,6 & 0.714\,3\end{bmatrix}$$

评注 比例尺度变换还假定属性对决策的重要性是随属性值线性变化的.对于明显的非线性,如边际效益递减的情况,应先对属性值拟合适当的函数再作变换.

① 若遇到既非效益型也非费用型的,而是属性取中间某些数值时最好(如环境温度),标准化方法可参阅[81].

$$\boldsymbol{R}=\begin{bmatrix}0.370\,9 & 0.722\,9 & 0.631\,2\\ 0.515\,1 & 0.562\,3 & 0.631\,2\\ 0.772\,7 & 0.401\,6 & 0.450\,8\end{bmatrix}$$

属性权重的确定

各个属性对决策目标的影响程度(或重要性)称为属性权重,记属性 X_1, X_2, …, X_n的权重为 w_1, w_2, …, w_n,满足 $\sum_{j=1}^{n} w_j = 1$,$\boldsymbol{w}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}$ 称为权向量.属性权重的确定也有偏于主观和偏于客观两种方法.偏于主观的方法可以由决策者根据决策目的和经验先验地给出.

信息熵法属于典型的偏于客观的方法.在信息论中熵是衡量不确定性的指标,一个信息量的(概率)分布越趋向一致,所提供信息的不确定性越大,当信息呈均匀分布时不确定性最大.在多属性决策中将按照归一化(1)式得到的决策矩阵 $\boldsymbol{R}$ 的各个列向量 $(r_{1j},r_{2j},\cdots,r_{mj})^{\mathrm{T}}(j=1,2,\cdots,n)$ 看作信息量的概率分布,按照 Shannon 给出的数量指标——熵的定义,各方案关于属性 X_j的熵为

$$E_j=-k\sum_{i=1}^{m} r_{ij}\ln r_{ij},\quad k=1/\ln m,\quad j=1,2,\cdots,n \tag{4}$$

当各方案对某个 X_j的属性值全部相同,即 $r_{ij}=1/m(i=1,2,\cdots,m)$ 时,$E_j=1$ 达到最大,这样的 X_j对于辨别方案的优劣不起任何作用;当各方案对 X_j的属性值 r_{ij}只有一个 1 其余全为 0 时,$E_j=0$ 达到最小,这样的 X_j最能辨别方案的优劣.一般地,属性值 r_{ij}相差越大,E_j越小,X_j辨别方案优劣的作用越大.于是定义

$$F_j=1-E_j,0\leqslant F_j\leqslant 1 \tag{5}$$

为属性 X_j的区分度.进一步,将归一化的区分度取作属性 X_j的权重 w_j,即

$$w_j=\frac{F_j}{\sum_{j=1}^{n} F_j},\quad j=1,2,\cdots,n \tag{6}$$

对于汽车选购,表 2 给出由归一化的决策矩阵 $\boldsymbol{R}$(即表中的 r_{ij})按照(4)~(6)式计算的熵 E_j、区分度 F_j和权重 $w_j(j=1,2,3)$.

表 2 汽车采购中决策矩阵的熵、区分度和权重

	X_1	X_2	X_3
r_{ij}	0.223 6	0.428 6	0.368 4
	0.310 6	0.333 3	0.368 4
	0.465 8	0.238 1	0.263 2
E_j	0.959 4	0.974 9	0.989 5
F_j	0.040 6	0.025 1	0.010 5
w_j	0.533 0	0.329 3	0.137 7

将各属性的权重按照大小排序为 w_1, w_2, w_3.实际上,观察原始决策矩阵 $\boldsymbol{D}$(或表 1)可以看出:3 种汽车对 X_1(价格)的属性值相差很大,对 X_3(款式)的属性值相差甚小,

根据这样的数据利用信息熵法计算权重,结果自然是 w_1 较大而 w_3 较小.

信息熵法完全由决策矩阵计算属性权重,如果决策矩阵主要是直接通过量测或调查得到的,那么这种获取权重的方法客观性较强.

与信息熵法的思路相似,可以用 $r_{1j},r_{2j},\cdots,r_{mj}$ 的标准差或极差作为区分度 F_j 计算权重,这适用于 m 较大的情况.

如果将偏于主观和客观两种方法得到的权重分别记作 $\boldsymbol{w}^{(1)}=(w_1^{(1)},w_2^{(1)},\cdots,w_n^{(1)})^{\mathrm{T}}$, $\boldsymbol{w}^{(2)}=(w_1^{(2)},w_2^{(2)},\cdots,w_n^{(2)})^{\mathrm{T}}$,那么可以用如下简单的方法将二者综合,得到新的权重

$$w_j=\frac{w_j^{(1)}w_j^{(2)}}{\sum_{j=1}^{n}w_j^{(1)}w_j^{(2)}},\quad j=1,2,\cdots,n \tag{7}$$

式中乘积 $w_j^{(1)}w_j^{(2)}$ 可以改为 $w_j^{(1)\alpha}w_j^{(2)\beta}$ 或 $\alpha w_j^{(1)}+\beta w_j^{(2)}$,其中 α,β 可根据决策者对 $w^{(1)}$, $w^{(2)}$ 的偏好程度进行调节.

综合方法

得到决策矩阵及属性权重以后,数学上可以采用多种方法将它们综合,按照决策者的需要确定一个最优方案,或者各方案按照优劣排序的数量结果,即方案对目标的(综合)权重.下面是几种主要的方法.

- 加权和法 这是人们熟知的、常用的方法.已知标准化决策矩阵 $\boldsymbol{R}=(r_{ij})_{m\times n}$ 及属性权重 $\boldsymbol{w}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}$(满足 $\sum_{j=1}^{n}w_j=1$),则方案 A_i 对目标的权重 v_i 是 r_{ij} 对 w_j 的加权和,即

$$v_i=\sum_{j=1}^{n}r_{ij}w_j,\quad i=1,2,\cdots,m \tag{8}$$

若记向量 $\boldsymbol{v}=(v_1,v_2,\cdots,v_m)^{\mathrm{T}}$,则(8)可写作矩阵形式

$$\boldsymbol{v}=\boldsymbol{R}\boldsymbol{w} \tag{9}$$

应该指出,当对决策矩阵采用不同的标准化方法时,得到的结果会有差别.对于汽车选购问题,设属性权重为用信息熵法得到的表 2 的最后一行 $\boldsymbol{w}$,用归一化和最大化的决策矩阵 $\boldsymbol{R}$ 代入(9)计算,得到的结果见表 3 的第 2,3 列(第 3 列是将直接计算出的数值又归一化的结果,以便与第 2 列比较).

- 加权积法 将加权和(算数平均)改为加权积(几何平均),与(8)式相对应的公式为

$$v_i=\prod_{j=1}^{n}d_{ij}^{w_j},\quad i=1,2,\cdots,m \tag{10}$$

评注 加权积法(10)式可以直接用方案原始的属性值 d_{ij} 代替 r_{ij},为什么?

对于汽车选购直接根据决策矩阵 $\boldsymbol{D}$(统一为效益型属性后)及表 2 最后一行 $\boldsymbol{w}$,利用(10)式计算,得到的结果归一化后见表 3 的第 4 列.

- 接近理想解的偏好排序法(简称 TOPSIS 方法) 将 n 个属性、m 个方案的多属性决策放到 n 维空间中 m 个点的几何系统中处理.用向量模一化(3)式对决策矩阵标准化,以便在空间定义欧氏距离.每个点的坐标由各方案标准化后的加权属性值确定.理论上的最优方案(称正理想解)由所有可能的加权最优属性值构成,最劣方案(称负理想解)由所有可能的加权最劣属性值构成,在确定最优和最劣属性值时应区分效

益型与费用型属性.定义距正理想解尽可能近、距负理想解尽可能远的数量指标——相对接近度,备选方案的优劣顺序按照相对接近度的大小确定.下面通过汽车选购说明TOPSIS方法的一般步骤.

1. 将决策矩阵模一化后的 r_{ij} 乘属性权重 w_j,得 $v_{ij}=r_{ij}w_j$,构成矩阵

$$\boldsymbol{V}=(v_{ij})=\begin{bmatrix}0.197\,7 & 0.238\,1 & 0.086\,9\\ 0.274\,6 & 0.185\,2 & 0.086\,9\\ 0.411\,8 & 0.132\,3 & 0.062\,1\end{bmatrix}$$

2. 正理想解(记作 $\boldsymbol{v}^+$)和负理想解(记作 $\boldsymbol{v}^-$)分别由 $\boldsymbol{V}$ 每一列向量的最大元素和最小元素构成,有 $\boldsymbol{v}^+=(0.411\,8,\ 0.238\,1,\ 0.086\,9)$,$\boldsymbol{v}^-=(0.197\,7,\ 0.132\,3,\ 0.062\,1)$.

3. 方案 A_i 与正理想解的(欧氏)距离按照 $S_i^+=\sqrt{\sum_{j=1}^{3}(v_{ij}-v_j^+)^2}$ 计算(其中 v_j^+ 为 $\boldsymbol{v}^+$ 的第 j 分量),得 $\boldsymbol{S}^+=(S_1^+,S_2^+,S_3^+)=(0.214\,2,\ 0.147\,1,\ 0.108\,7)$;$A_i$ 与负理想解的距离按照 $S_i^-=\sqrt{\sum_{j=1}^{3}(v_{ij}-v_j^-)^2}$ 计算(其中 v_j^- 为 $\boldsymbol{v}^-$ 的第 j 分量),得 $\boldsymbol{S}^-=(S_1^-,S_2^-,S_3^-)=(0.108\,7,0.096\,6,0.214\,2)$.

4. 定义方案 A_i 与正理想解的相对接近度为 $C_i^+=\dfrac{S_i^-}{S_i^++S_i^-}$,$0<C_i^+<1$,计算得 $\boldsymbol{C}^+=(C_1^+,C_2^+,C_3^+)=(0.336\,6,\ 0.396\,3,\ 0.663\,4)$,归一化后的结果见表3的第5列.

评注 简单、直观的加权和法往往是人们的首选,但是应注意这个方法的前提是,属性之间相互独立,并且具有互补性,即一个方案在某一属性上的欠缺可以被在另一属性上的优势补偿.由于实际应用时这个前提常常难以保证和检验,所以采用加权和法需要小心点.

表3 汽车采购问题用3种综合方法的计算结果(备选方案对决策目标的权重)

方案	方法			
	加权和($\boldsymbol{R}$ 归一化)	加权和($\boldsymbol{R}$ 最大化)	加权积	TOPSIS
A_1	0.311 0	0.316 2	0.306 7	0.241 1
A_2	0.326 0	0.327 7	0.336 4	0.283 8
A_3	0.362 9	0.356 2	0.356 8	0.475 1

由表3可以看出,加权和法与加权积法得到的结果差别很小,TOPSIS方法的结果差别稍大,但用这些方法得到的3个方案的优劣顺序均为 A_3,A_2,A_1.对综合方法更多的比较结果可参看[67].

多属性决策应用的步骤

1) 确定决策目标、备选方案与属性集合;

2) 通过量测、调查、专家评定等手段确定决策矩阵和属性权重,推荐用信息熵法由决策矩阵得出属性权重;

3) 采用归一化(分配模式)、最大化(理想模式)或模一化对决策矩阵标准化;

4) 选用加权和、加权积、TOPSIS等综合方法计算方案对目标的权重.

多属性决策应用中的几个问题

- **比例尺度变换的归一化和最大化**

比例尺度变换通过归一化或最大化将原始决策矩阵标准化,用两种方法计算的结果一般不会相同(如表3).但是大多数实际问题中,特别是在方案数量不多时,方案的

优劣排序会大致一样.在实际应用中应该采用哪种方法呢?

以汽车选购为例,如果选购者的目标是综合指标最优、最理想,在确定各种型号汽车的每项属性值时,不受其他型号汽车的某些属性值变化的影响,那么他应该采用最大化方法;如果选购者的目标是综合指标在他的同事们选购的汽车中最突出,承认并接受不同型号汽车某些属性值变化对选择结果的影响,那么他就应该采用归一化方法.最大化和归一化方法又分别称为理想模式和分配模式[58].

评注 一般地说,当决策者关心的是每个方案相对于某些基准指标的性能时,应考虑使用理想模式,而当决策者关心的是每个方案相对其他方案的占优程度时,应考虑使用分配模式.由此看来,在实际应用中,对于从众多候选方案中只选一个最优者的情况,多用理想模式,而那些需要对候选方案的优劣给出定量意义下的比较,或者是对资源按照候选方案的优劣进行分配时,多用分配模式.

- 区间尺度变换

所谓区间尺度变换是指对原始权重 d_{ij} 作如下形式的标准化:

$$r_{ij}=\frac{d_{ij}-\min\limits_{i=1,2,\cdots m} d_{ij}}{\max\limits_{i=1,2,\cdots m} d_{ij}-\min\limits_{i=1,2,\cdots m} d_{ij}} \tag{11}$$

与比例尺度变换的最大化(2)式相似,d_{ij} 的最大值都变为 $r_{ij}=1$,它也可以用于多属性决策.需要注意的是,区间尺度变换后的最小值一定是 $r_{ij}=0$,正是这个 0 可能造成不良后果.

虚拟一个既简单又极端的例子,说明当按照方案的优劣处理资源分配问题时,用归一化的比例尺度变换特别合适,用最大化变换会出现较大的谬误,而若采用区间尺度变换将得到极不合理的结果[68].

奖金分配 将绩效奖金 1 万元按照教学和科研并重的原则分配给 A,B 两位教师,两个准则的权重 w_j 及两位教师的教学和科研原始得分 d_{ij} 如表 4.

表 4 奖金分配中的权重 w_j 及原始得分 $d_{ij}(i,j=1,2)$

	教学 $X_1(w_1=0.5)$	科研 $X_2(w_2=0.5)$
A	51	1
B	49	99

按照常识可以非常简单地大致分配 1 万元奖金:教学与科研各分 5000 元,教学 5000 元由 A,B 平分,科研 5000 元全给 B,于是 A 得 0.25 万元,B 得 0.75 万元.

用比例尺度变换的归一化、最大化及区间尺度变换 3 种方法计算的结果如表 5.

表 5 奖金分配用 3 种方法计算的结果

	归一化			最大化				区间尺度		
	X_1 $(w_1=0.5)$	X_2 $(w_2=0.5)$	综合权重	X_1 $(w_1=0.5)$	X_2 $(w_2=0.5)$	综合权重	综合权重归一化	X_1 $(w_1=0.5)$	X_2 $(w_2=0.5)$	综合权重
A	0.51	0.01	0.26	1	0.01	0.51	0.33	1	0	0.5
B	0.49	0.99	0.74	0.96	1	0.98	0.67	0	1	0.5

评注 区间尺度变换与比例尺度变换的最大化时常被混用,一般情况下可能没有问题,但是需要注意能否接受最小值变为 0 造成的后果.

可以看出,归一化的结果与常识相符,只是更精确些;最大化的结果(再归一化)与常识相差较大,这是由于对应 X_1 的 A,B 权重之和远大于对应 X_2 的 A,B 权重之和,相当于放大了教学的权重,缩小了科研的权重;区间尺度的结果则完全不能接受,这显然是由于把 A,B 非常接近的教学原始分 51 和 49 分别变成 1 和 0 的缘故.

评注 虽然讨论了一些关于决策矩阵归一化和最大化的区别，但是在多属性决策的大多数实际应用中，采用不同的标准化和综合方法对最终决策的影响，总的说来远小于不同的属性集合及属性权重对最终决策的影响，所以通常不需要过度注意前者，而应对后者多花些精力.

● 方案的排序保持与排序逆转

在多属性决策中有时遇到新方案加入或旧方案退出的情况.假定各属性对目标的权重和原有方案对属性的权重都不变，那么方案的加入或退出会影响原有方案的优劣排序吗？直观上似乎不会，但是事实上用理想模式（最大化）在某些条件下原有方案的排序一定保持，而用分配模式（归一化）却可能发生逆转[55,58].本章训练题14给出了一个例子.

用两种模式计算会出现如此不同的后果并不难理解：在分配模式下各方案对每一属性权重 r_{ij} 之和恒为1，新方案的加入会导致原来 r_{ij} 的减少，即“稀释”了原有“资源”，“资源”的重新分配自然可能导致原方案排序逆转；在理想模式下各方案对每一属性权重 r_{ij} 的最大值为1，新方案的加入只要不改变原来的最大值，就不会“稀释”原有“资源”，原方案排序将保持不变.

复习题

1. 在表1汽车采购的原始权重中，对价格 X_1 采用一个大数减去原属性值的方法（如果买车的心理价位是30万元，用它减去 X_1 相当于节省的钱），化为效益型.按照本节的方法，得出标准化决策矩阵、属性权重及表3那样的最终结果.

2. 选择战斗机.待测评或购买的战斗机有4种备选型号 A_1, A_2, A_3, A_4，已确定的属性为：最高速度 X_1（马赫）、航程 X_2（10^3 n mile）、最大载荷 X_3（10^3 lb）、价格 X_4（10^6 美元）、可靠性 X_5、机动性 X_6.4种战斗机对6个属性的定量取值或定性表述如下表[30].

备选方案	属性					
	X_1	X_2	X_3	X_4	X_5	X_6
A_1	2.0	1.5	20	5.5	中	很高
A_2	2.5	2.7	18	6.5	低	中
A_3	1.8	2.0	21	4.5	高	高
A_4	2.2	1.8	20	5.0	中	中

根据以下要求确定最终决策（优劣排序和数值结果）：

（1）对属性 X_5, X_6 的定性表述给以定量化，对“很高”、“高”、“中”、“低”、“很低”分别给以分值9,7,5,3,1，或者分别给以分值5,4,3,2,1.

（2）属性权重主观地给定为0.2,0.1,0.1,0.1,0.2,0.3，或者对决策矩阵用信息熵方法得到.

（3）对决策矩阵归一化、最大化、模一化.

（4）用加权和法、加权积法、TOPSIS方法计算方案对目标的权重.

7.2 职员晋升

职场中公平、公正地实施职员晋升，是管理者的一件非常重要而又困难的工作.一种简单易行、具有一定合理性的办法是，由评委会先订立全面评价一位职员的几条准

则,如工作年限、教育程度、工作能力、道德品质等,并且确定各条准则在职员晋升这个总目标中所占的权重,然后按照每一条准则对每位申报晋升的职员进行比较和评判,最后将准则的权重与比较、评判的结果加以综合,得到各位申报者的最终排序,作为管理者对职员晋升的决策.

可以看出,职员晋升与7.1节的汽车选购是具有相同特点的决策问题,用多属性决策完全可以类似地给以解决.本节要以职员晋升为例,介绍处理这类问题的另一种常用的方法——**层次分析法**(analytic hierarchy process,简记AHP)建模的全过程.

层次分析法是Saaty T. L.于20世纪70年代(稍晚于多属性决策)提出的一种系统化、层次化、定性和定量相结合的思路和手段[40,54,57,77],在实际应用的领域、处理问题的类型以及具体的计算方法等方面,与多属性决策有诸多类似和相通之处.一般地说,层次分析法包含以下几个要素.

层次结构图

许多决策问题都可以自上而下地分为目标、准则、方案3个层次,并且直观地用一个层次结构图表示,如图1是职员晋升的结构图,最上层是目标层(职员晋升),通常只有一个元素,最下层是方案层(3位职员),中间的准则层既影响目标,又支配方案,图中用直线表示上下层元素之间的联系. 与多属性决策相同,层次分析法首先也要确定各个准则(即属性)对目标的权重以及各个方案对每一准则的权重,然后再将二者综合,得到方案对目标的权重.Saaty的贡献之一,在于提出利用成对比较矩阵和特征向量确定下层诸元素对上层元素的权重,并进行一致性检验.

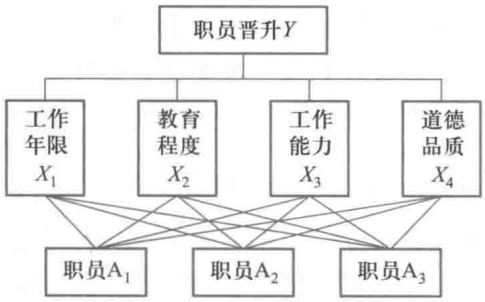


图1 职员晋升的层次结构图

成对比较矩阵和特征向量 当确定某一层的 n 个元素 $X_1, X_2, \cdots, X_n$ 对上层的一个元素 Y 的权重(如准则对目标的权重)时,为了减少在这些性质不同的元素之间相互比较的困难,Saaty建议不把它们放在一起比较,而是两两相互对比,并且对比时采用相对尺度.将 X_i 和 X_j 对 Y 的重要性之比用 a_{ij} 表示,n 个元素 $X_1, X_2, \cdots, X_n$ 两两成对地对 Y 重要性之比的结果用**成对比较矩阵**

$$\boldsymbol{A} = (a_{ij})_{n\times n}, \quad a_{ij} > 0, \quad a_{ji} = 1/a_{ij} \tag{1}$$

表示,数学上 $\boldsymbol{A}$ 又称**正互反阵**.显然 $\boldsymbol{A}$ 的对角线元素 $a_{ii}=1(i=1,2,\cdots,n)$.

例如当确定4个准则(工作年限 X_1、教育程度 X_2、工作能力 X_3、道德品质 X_4)对目标(职员晋升 Y)的权重时,需将 X_1, X_2, X_3, X_4 两两进行对比.假定决策者认为对于职员晋升的重要性来说,工作年限 X_1 与教育程度 X_2 之比是1∶2,即 $a_{12}=1/2$,X_1 与工作能力 X_3 之比是1∶3,即 $a_{13}=1/3$, X_1 与道德品质 X_4 之比是1∶2,即 $a_{14}=1/2$;教育程度 X_2 与 X_3 之比是1∶2,即 $a_{23}=1/2$, X_2 与 X_4 之比是1∶1,即 $a_{24}=1$;X_3 与 X_4 之比是2∶1,即 $a_{34}=2$.这些 $a_{ij}(1<i<j<4)$ 是矩阵 $\boldsymbol{A}$ 的上三角元素,按照(1)式即可写出成对比较阵为

$$A=\begin{bmatrix}1 & 1/2 & 1/3 & 1/2\\ 2 & 1 & 1/2 & 1\\ 3 & 2 & 1 & 2\\ 2 & 1 & 1/2 & 1\end{bmatrix}$$

你可能已经发现,既然 $a_{12}=1/2, a_{23}=1/2$,那么 X_1 与 X_3 比较时应该是 1∶4,即 $a_{13}=1/4$ 而非 1/3,这样才能做到成对比较的一致性.但是,n 个元素需做 $n(n-1)/2$ 次成对比较,要求全部一致是不现实、也不必要的.层次分析法容许成对比较存在不一致,并且给出了不一致情况下计算各元素权重的方法,同时确定了这种不一致的容许范围.

为了说明 Saaty 提出的办法,考察成对比较完全一致的情况.假定 n 个元素 X_1, X_2, …, X_n 对 Y 的重要性之比已经精确地测定为 $w_1 : w_2 : \cdots : w_n$,在成对比较时只需取 $a_{ij}=w_i/w_j$,就一定完全一致,这样的矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 简称一致阵.实际上,只要成对比较的全部取值满足

$$a_{ij}\cdot a_{jk}=a_{ik},\quad i,j,k=1,2,\cdots,n \tag{2}$$

成对比较阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 就是一致阵.容易看出,一致阵的各列均相差一个比例因子,由此可得一致阵 A 的以下代数性质:

A 的秩为 1,唯一非零特征根为 n,任一列向量都是对应于特征根 n 的特征向量.

如果成对比较阵 $\boldsymbol{A}$ 是一致阵,显然特征向量可以取 $\boldsymbol{w}=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}$.不妨设 $w_1,w_2,\cdots,w_n$ 已经归一化,即满足 $\sum_{j=1}^{n} w_j=1$,那么 $\boldsymbol{w}$ 的各个分量正是 n 个元素的权重.如果成对比较阵 $\boldsymbol{A}$ 不一致,但在不一致的容许范围内(稍后说明),Saaty 等人建议[①]用对应于 $\boldsymbol{A}$ 的最大特征根(记作 λ)的特征向量(归一化后)为权向量 $\boldsymbol{w}$,即 $\boldsymbol{w}$ 满足

$$\boldsymbol{A}\boldsymbol{w}=\lambda\boldsymbol{w} \tag{3}$$

对于职业晋升中 4 个准则的成对比较阵 $\boldsymbol{A}$,用 MATLAB 软件算出的最大特征根 $\lambda=4.010\,4$,对应特征向量(归一化后)$\boldsymbol{w}=(0.122\,3, 0.227\,0, 0.423\,6, 0.227\,0)^{\mathrm{T}}$.

注意到成对比较阵是通过定性比较得到的相当粗糙的结果,精确地计算特征向量常常是不必要的,在实际应用时可以用成对比较阵各列向量的平均值近似代替特征向量(复习题 1).

一致性指标和一致性检验 实际应用中成对比较阵 $\boldsymbol{A}$ 通常不是一致阵,为了能用对应于 $\boldsymbol{A}$ 最大特征根 λ 的特征向量作为权向量 $\boldsymbol{w}$,需要对它不一致的范围加以界定.

数学上已经证明,对于 n 阶成对比较阵(正互反阵)$\boldsymbol{A}$,其最大特征根 $\lambda\geqslant n$,且 $\boldsymbol{A}$ 是一致阵的充分必要条件为 $\lambda=n$.

根据这个结果可知,λ 比 n 大得越多,$\boldsymbol{A}$ 与一致阵相差得越远,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 $\boldsymbol{A}$ 的不一致程度.Saaty 将

$$CI=\frac{\lambda-n}{n-1} \tag{4}$$

① 其原因可参看拓展阅读 7-1.

定义为一致性指标①.显然 $CI=0$ 时 $\boldsymbol{A}$ 是一致阵,CI 越大,$\boldsymbol{A}$ 越不一致.于是可以制定一个衡量 CI 数值的标准,来界定 A 不一致的容许范围.

为此 Saaty 又引入随机一致性指标 RI,其产生过程是,对于每个 $n(=3,4,\cdots)$ 随机模拟大量的正互反阵 $\boldsymbol{A}$(其元素 a_{ij} 从 $1,2,\cdots,9$ 及 $1,1/2,\cdots,1/9$ 中等可能地随机取值,原因稍后说明),计算这些 $\boldsymbol{A}$ 的一致性指标 CI 的平均值作为 RI.Saaty 用大量样本得到的 RI 如表 1.

表 1 Saaty 给出的随机一致性指标 RI 的数值②

n	3	4	5	6	7	8	9	10
RI	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

实际应用中计算出 n 阶成对比较阵 $\boldsymbol{A}$ 的一致性指标 CI 以后,与同阶的随机一致性指标 RI 比较,当比值 CR(称为一致性比率)满足

$$CR=\frac{CI}{RI}<0.1 \tag{5}$$

时认为 $\boldsymbol{A}$ 的不一致程度在容许范围之内.(5)式中的 0.1 是可以调整的,对于重要决策问题应该适当减小.

对于成对比较阵 $\boldsymbol{A}$ 利用(4),(5)式及表 1 进行的检验称为一致性检验,若检验通过,可以用 $\boldsymbol{A}$ 的特征向量作为权向量,若检验不通过,需要对 $\boldsymbol{A}$ 作修正,或者重新做成对比较.

对于职业晋升中 4 个准则的成对比较阵 $\boldsymbol{A}$,由 $\lambda=4.010\ 4$ 和(4)式算出 $CI=0.003\ 5$,由表 1 知 $RI=0.90$,由(5)式 $CR=0.003\ 5/0.90<0.1$,一致性检验通过,于是,归一化的特征向量 $\boldsymbol{w}=(0.122\ 3,0.227\ 0,0.423\ 6,0.227\ 0)^{\mathrm{T}}$ 可以作为权向量.

综合权重 对于职业晋升问题用成对比较得到第 2 层 4 个准则对第 1 层目标的权重(记作 $\boldsymbol{w}^{(2)}$)之后(参看图 1),用同样的方法确定第 3 层 3 个方案(职员 A_1, A_2, A_3)对第 2 层每一准则的权重.设决策者给出的第 3 层对 X_1, X_2, X_3, X_4 的成对比较阵依次为

$$\boldsymbol{B}_1=\begin{bmatrix}1 & 1/2 & 1/4\\ 2 & 1 & 1/3\\ 4 & 3 & 1\end{bmatrix},\boldsymbol{B}_2=\begin{bmatrix}1 & 2 & 3\\ 1/2 & 1 & 2\\ 1/3 & 1/2 & 1\end{bmatrix}$$

$$\boldsymbol{B}_3=\begin{bmatrix}1 & 1 & 2\\ 1 & 1 & 2\\ 1/2 & 1/2 & 1\end{bmatrix},\boldsymbol{B}_4=\begin{bmatrix}1 & 3 & 4\\ 1/2 & 1 & 2\\ 1/4 & 1/2 & 1\end{bmatrix}$$

由 $\boldsymbol{B}_j(j=1,2,3,4)$ 计算其最大特征根 λ_j,一致性指标 CI_j 及(归一化)特征向量 $\boldsymbol{w}_j^{(3)}$,结果如表 2.为了下面计算的方便,将已经得到的 $\boldsymbol{w}^{(2)}$ 放到表 2 的最后一列.

因为表 1 的 $RI=0.58$,由表 2 的 CI_j 可知 $CR_j=CI_j/RI<0.1$,4 个成对比较阵均通过一致性检验,4 个(归一化)特征向量 $\boldsymbol{w}_j^{(3)}$ 可作为第 3 层(职员 A_1, A_2, A_3)对第 2 层每一

① 由于 $\boldsymbol{A}$ 的 n 个特征根之和等于 n,CI 的分母 $n-1$ 可以抵消随着 n 的增加分子的变大.

② 由于随机性,不同的人用不同样本得到的 RI 的数值会与表 1 稍有出入.$n=2$ 时 $\boldsymbol{A}$ 总是一致阵,$RI=0$.

准则的权重.

表 2 职业晋升问题第 3 层对第 2 层的计算结果

	j				
	1	2	3	4	$w^{(2)}$
$w_j^{(3)}$	0.136 5	0.539 6	0.400 0	0.625 0	0.122 3
	0.238 5	0.297 0	0.400 0	0.238 5	0.227 0
	0.625 0	0.163 4	0.200 0	0.136 5	0.423 6
					0.227 0
λ_j	3.018 3	3.009 2	3.000 0	3.018 3	
CI_j	0.009 2	0.004 6	0	0.009 2	

注意到表 2 中第 2 行的前 4 个数值分别是职员 A_1 对 4 项准则的权重,将它们与表 2 最后一列 4 个准则对目标的权重 $\boldsymbol{w}^{(2)}$ 对应地相乘再求和,就得到职员 A_1 对目标的权重.职员 A_2, A_3 对目标的权重可以类似地得到.

评注 与多属性决策的加权和法对比,$\boldsymbol{W}^{(3)}$ 相当于归一化的决策矩阵 $\boldsymbol{R}$,$\boldsymbol{w}^{(2)}$ 相当于归一化的属性权重 $\boldsymbol{w}$,$\boldsymbol{w}^{(3)}$ 即方案对目标的权向量 $\boldsymbol{v}$,于是(6)式可看作分层加权和法,与 7.1 节(9)式完全一样.

容易看出,上述运算相当于用 $\boldsymbol{w}_j^{(3)}$ 构成的矩阵 $\boldsymbol{W}^{(3)}=[\boldsymbol{w}_1^{(3)},\boldsymbol{w}_2^{(3)},\boldsymbol{w}_3^{(3)},\boldsymbol{w}_4^{(3)}]$ 与权重向量 $\boldsymbol{w}^{(2)}$ 相乘,从而得到第 3 层对第 1 层的综合权重 $\boldsymbol{w}^{(3)}$,表示为

$$\boldsymbol{w}^{(3)}=\boldsymbol{W}^{(3)}\boldsymbol{w}^{(2)} \tag{6}$$

根据表 2 的数据,按照(6)式计算可得 $\boldsymbol{w}^{(3)}=(0.450\,5,\ 0.320\,2,\ 0.229\,2)^{\mathrm{T}}$,表示如果管理者在职业晋升中所做的成对比较由 $\boldsymbol{A}$ 和 $\boldsymbol{B}_1,\boldsymbol{B}_2,\boldsymbol{B}_3,\boldsymbol{B}_4$ 给出,那么用层次分析法得到的 3 位职员的优劣顺序为 A_1, A_2, A_3.

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,还常要作所谓综合一致性检验,详见文献[58].

1-9 比较尺度 当成对比较 X_i, X_j 对 Y 的重要性时,比较尺度 a_{ij} 采用什么数值表示合适呢? Saaty 等人建议用 1-9 尺度,即 a_{ij} 的取值范围规定为 1,2,…,9 及其互反数 1,1/2,…,1/9,其大致理由如下:

- 进行定性的成对比较时人们脑海中通常有 5 个明显的等级,可以用 1-9 尺度方便地表示如表 3.

表 3 比较尺度 a_{ij} 的含义

X_i 与 X_j 的比较尺度 a_{ij}	含义
1	X_i 与 X_j 的重要性相同
3	X_i 比 X_j 稍微重要
5	X_i 比 X_j 重要
7	X_i 比 X_j 明显重要
9	X_i 比 X_j 绝对重要
2,4,6,8	X_i 与 X_j 重要性在以上两个相邻等级之间
1, 1/2 ,…, 1/9	X_i 与 X_j 的重要性比较和以上结果相反

● 心理学家认为，成对比较的元素太多，会超出人们的判断能力，如以 9 个为限，用 1-9 尺度表示它们之间的差别正合适.

● Saaty 曾用 1-3，1-5，…，1-17，1^p-9^p(p=2,3,4,5)等 27 种尺度，对一些已经知道元素权重的实例(如不同距离处的光源)构造成对比较阵，再由此计算权重，与已知的权重对比发现，1-9 尺度不仅在较简单的尺度中最好，而且不劣于较复杂的尺度.

目前在层次分析法的应用中，大多数人采用 1-9 尺度.

评注 7.1 节提到的区间尺度变换不符合 Saaty 提出的用比例尺度构造正互反阵、得到(正)特征向量的要求，因而不被层次分析法采用，它只是多属性决策可以使用的方法之一.

职员晋升问题的再讨论

上面是对每位申报晋升的职员直接按照每一条准则进行比较和评判，由于这些准则过于笼统、不够具体，比较和评判起来有一定困难，特别是当申报者较多时进行成对比较的一致性难以保证.实际上，可以将每条准则细分为若干等级，如工作年限和教育程度可分别用入职时间和学历分级，工作能力和道德品质大概只好按照优、良、中划分.若评委会确定了每条准则中各个等级的分值，那么每一位申报晋升的职员只需根据在每条准则中所处的等级“对号入座”，再根据 4 条准则在职员晋升中的权重，就可以计算出他(她)的总分，根据总分确定能否晋升.

对 4 个准则不妨作如下分级：工作年限 X_1 分为很长(10 年以上)、长(5 至 10 年)、中(2 至 5 年)、短(2 年以下)4 个等级，依次记为 X_{11}，X_{12}，X_{13}，X_{14}；教育程度 X_2 分为本科以上、本科、专科、中学 4 个等级 X_{21}，X_{22}，X_{23}，X_{24}；工作能力 X_3 分为优、良、中、差 4 个等级 X_{31}，X_{32}，X_{33}，X_{34}；道德品质 X_4 只有优、良、中 3 个等级(品质差的根本不考虑晋升)X_{41}，X_{42}，X_{43}.申报晋升的职员有 A_1，A_2，…，A_k，… .职员晋升细分等级的层次结构如图 2.



图 2 职员晋升细分等级的层次结构图

假定 4 个准则对职员晋升目标的权重仍为前面用成对比较得到的 $\boldsymbol{w}^{(2)}=(w_1, w_2, w_3, w_4)=(0.122\,3, 0.227\,0, 0.423\,6, 0.227\,0)^{\mathrm{T}}$，在每一准则中将最高等级定为 100 分，由评委会决定其他等级的分数 w_{ij}(i=1,2,3, j=1,2,3,4;i=4, j=1,2,3)，其数值如表 4.

表 4 职员晋升问题准则的权重 w_i 及其等级的分数 w_{ij}

	w_1=0.122 3				w_2=0.227 0				w_3=0.423 6				w_4=0.227 0			
	w_{11}	w_{12}	w_{13}	w_{14}	w_{21}	w_{22}	w_{23}	w_{24}	w_{31}	w_{32}	w_{33}	w_{34}	w_{41}	w_{42}	w_{43}	总分
	=100	=80	=60	=30	=100	=90	=60	=30	=100	=80	=40	=10	=100	=80	=40	
⋮																
A_k			√			√			√					√		88
⋮																

评注 最高等级为 100 分的打分方法相当于多属性决策中属性值向量的最大化处理(用 100 分代替 1).用这种方法可以在评定之前确定一个标准分，如 80 分，标准分以上的才可以晋升，这就避免了待评人员直接的相互比较以及一些其他人为因素的干扰.

每个申报的职员只需根据在每个准则中所处等级的位置“对号入座”(画√),就可算出他(她)的总分.如一位工作4年、工作能力优秀、道德品质良好的本科毕业生 A_k,总分为 60×0.122 3+90×0.227 0+100×0.423 6+80×0.227 0=88.29.用符号表示:设 A_k 在4项准则中所处的等级为 $\{j_1, j_2, j_3, j_4\}$,则其总分是 $\sum_{i=1}^{4} w_i w_{ij_i}$.工作10年以上、工作能力和道德品质均优、获得硕士学位的毕业生可得最高总分100.

层次分析法应用的步骤

多属性决策应用的步骤可以归纳如下:

评注 层次分析法的优点可以用“系统性、实用性、简洁性”来表述;其局限性则可用“囿旧、粗糙、主观”来概括.

1) 建立由目标层、准则层、方案层等构成的层次结构;

2) 构造下层各元素对上层每一元素的成对比较阵;

3) 计算各个成对比较阵的特征根和特征向量,作一致性检验,通过后将特征向量取作权向量;

4) 用分层加权和法计算最下层各元素对最上层元素的权重.

层次分析法与多属性决策的比较 两种方法都能用于解决决策问题,从本节和上节的介绍可以看到,二者在步骤、方法上有很多相同之处,也有一些差别.

评注 由于多属性决策和层次分析法发展进程不同,公理化体系不同,致使对这两种方法所得结果(如是否允许原有方案排序发生逆转)的争论长期存在[58],不过这些分歧对应用来说没有太大影响,只需注意和了解采用不同方法的条件及可能出现的后果就行了.

不论是层次分析法还是多属性决策,重点都是要确定准则(属性)对目标的权重和方案对准则(属性)的权重,其手段可分为**相对量测**和**绝对量测**.像层次分析法中进行的成对比较属于前者,而如果能用定量的尺度来描述方案或准则的特征,则属于后者.

对于尚没有太多知识的新问题和模糊、抽象的准则,主要依赖于相对量测,而对于已有充分了解的老问题和明确、具体的准则,应该尽可能地采用绝对量测.如购物选择、旅游地选择中的价格,人员聘用和录取中的工作年限、奖学金评定中的学习成绩、宜居城市评选中的空气质量、大学排行榜制定中的论文数量等,都是可以使用绝对量测的准则.

一般来说,相对量测偏于主观、定性,绝对量测偏于客观、定量,应尽量采用绝对量测.

绝对量测的另一个好处是,当新方案加入或老方案退出时,原有方案的结果不会改变.而若用相对量测就要重新作若干比较,原有方案的结果也可能改变.

在应用中可以将多属性决策和层次分析中的方法结合起来,如用成对比较阵来确定属性权重,用绝对量测确定决策矩阵.

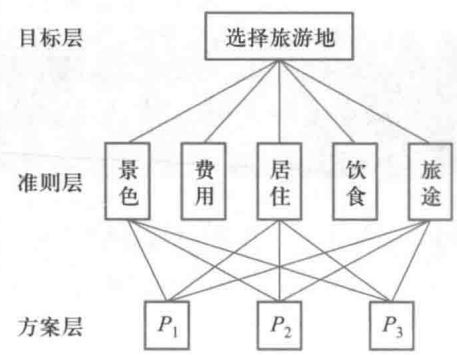
复习题

拓展阅读 7-1
层次分析法的若干问题

1. 层次分析法在实际应用时,可以用成对比较阵 $\boldsymbol{A}$ 的列向量的平均值近似代替特征向量,称为和法,其步骤是:先将 $\boldsymbol{A}$ 的每一列向量归一化,按行求和后再归一化,得到的 $\boldsymbol{w}=(w_1, w_2, \cdots, w_n)^{\mathrm{T}}$ 即为近似特征向量,并将 $\frac{1}{n}\sum_{i=1}^{n}\frac{(\boldsymbol{Aw})_i}{w_i}$ 作为近似最大特征根.

设 $\boldsymbol{A}=\begin{bmatrix} 1 & 2 & 6 \\ 1/2 & 1 & 4 \\ 1/6 & 1/4 & 1 \end{bmatrix}$,用和法计算近似特征向量和近似最大特征根,并与精确值比较.

2. 以选择旅游地为目标的层次结构图如右图,景色、



费用等5个准则构成准则层，P_1，P_2，P_3 3个旅游地构成方案层.已知准则对目标的成对比较阵

$$\boldsymbol{A}=\begin{bmatrix} 1 & 1/2 & 4 & 3 & 3 \\ 2 & 1 & 7 & 5 & 5 \\ 1/4 & 1/7 & 1 & 1/2 & 1/3 \\ 1/3 & 1/5 & 2 & 1 & 1 \\ 1/3 & 1/5 & 3 & 1 & 1 \end{bmatrix}$$

及方案对5个准则的成对比较阵

$$\boldsymbol{B}_1=\begin{bmatrix} 1 & 2 & 5 \\ 1/2 & 1 & 2 \\ 1/5 & 1/2 & 1 \end{bmatrix},\quad \boldsymbol{B}_2=\begin{bmatrix} 1 & 1/3 & 1/8 \\ 3 & 1 & 1/3 \\ 8 & 3 & 1 \end{bmatrix},\quad \boldsymbol{B}_3=\begin{bmatrix} 1 & 1 & 3 \\ 1 & 1 & 3 \\ 1/3 & 1/3 & 1 \end{bmatrix}$$

$$\boldsymbol{B}_4=\begin{bmatrix} 1 & 3 & 4 \\ 1/3 & 1 & 1 \\ 1/4 & 1 & 1 \end{bmatrix},\quad \boldsymbol{B}_5=\begin{bmatrix} 1 & 1 & 1/4 \\ 1 & 1 & 1/4 \\ 4 & 4 & 1 \end{bmatrix}$$

(1) 计算各个成对比较阵的特征向量，作一致性检验，确定权向量.

(2) 计算方案对目标的综合权重，确定用层次分析法选择的旅游地.

3. 证明层次分析模型中定义的 n 阶一致阵 $\boldsymbol{A}$ 有下列性质：

(1) $\boldsymbol{A}$ 的秩为1，唯一非零特征根为 n.

(2) $\boldsymbol{A}$ 的任一列向量都是对应于 n 的特征向量.

7.3 厂房新建还是改建

在经济活动中，管理者经常要对若干备选方案做出决策，有时决策的后果是确定的，像7.1节的汽车选购问题，有时决策的后果是随机的，比如公司面临扩大规模的几个备选方案，其后果取决于未来市场产品的销路，而目前只能估计出未来产品销路处于良好状态还是较差状态的概率.在这种情况下怎样做出决策呢？下面看一个具体的案例.

一家公司为提高某种产品的质量以拓展市场，拟制定一个10年计划.现有新建厂房与改建厂房两种备选方案.如果投资400万元新建厂房，若未来产品销路好，则年收益可达100万元；而若销路差，则年亏损20万元.如果投资100万元改建厂房，那么未来销路好和销路差的年收益分别为40万元和10万元.又据估计，未来产品销路好与产品销路差的可能性是7∶3.从利润最大化的角度，公司应该新建还是改建厂房？

问题的分析与求解　这是一个非常简单的问题，只需计算、比较两种备选方案10年的总利润，就可得到应该新建还是改建厂房的决策.

如果把“未来产品销路好与产品销路差的可能性是7∶3”简单地解释为，未来10年中有7年销路好和3年销路差，那么新建厂房10年的总利润是 $100\times7-20\times3-400=240$（万元），改建厂房10年的总利润是 $40\times7+10\times3-100=210$（万元），决策应该是新建厂房.

如果把“未来产品销路好与产品销路差的可能性是7∶3”解释为，未来10年产品销路好与销路差的概率分别是0.7与0.3，那么新建厂房10年总利润的期望值（即平均值）是 $100\times10\times0.7-20\times10\times0.3-400=240$（万元），改建厂房是 $40\times10\times0.7+10\times10\times0.3-100=210$（万元），以总利润的期望值最大为目标的决策也是新建厂房.这种决策的依据称为**期望值（平均值）准则**.

值得注意的是，期望值可看作随机事件多次重复出现条件下的平均值，将期望值准则用于这里的一次性决策会有较大的风险. 让我们计算一下，若选择新建厂房决策，如果未来10年产品真的销路好，总利润将是100×10-400=600(万元)，而如果未来10年产品真的销路差，总利润将是-20×10-400=-600(万元)，即亏损600万元，正负相差1 200万元，风险相当大. 若选择改建厂房决策，则产品真的销路好的总利润是300万元，而产品真的销路差的总利润是零，风险相对较小.

评注 在本节研究的决策问题中，每个备选方案的后果都至少存在两种状态，且各种状态的概率是可以估计的. 由于大多数实际问题属于一次性而非多次重复进行的决策，采用既有理论依据又便于应用的期望值准则作为选择标准，有时会冒较大的风险. 这类问题称为风险性决策.

在上面的计算中还可以看到，虽然根据期望值准则应选择新建厂房，但是与改建厂房相比总利润(期望值)只多240-210=30(万元). 因为对未来产品销路好与销路差的概率的估计并不准确，让我们分析一下，对这种概率的估计有多大变化，就会导致最终决策的改变.

将新建厂房(方案1)和改建厂房(方案2)10年总利润的期望值分别记作$E(1)$，$E(2)$，未来产品销路好的概率记作p，销路差的概率为$1-p$，像上面的计算一样，有

$$E(1)=100\times10\times p+(-20\times10)\times(1-p)-400=1\ 200p-600$$

$$E(2)=40\times10\times p+10\times10\times(1-p)-100=300p$$

令$E(1)=E(2)$可得$p=2/3$，且当$p<2/3$时$E(1)<E(2)$. 这个结果表明，若目前估计的销路好的概率从0.7降到0.66(下降只约5%)，总利润期望值最大的决策就将由新建厂房变为改建厂房. 看来，决策对概率的变化是相当敏感的.

新方案的提出 新建厂房决策的一次性风险太大，决策对概率的变化十分敏感，促使公司考虑制定能降低风险的新的折中方案：先改建厂房经营3年，3年后视市场情况再定. 如果3年销路好，则未来7年销路仍然好的概率预计将提高到0.9，这时若再投资200万元进行扩建，销路好时年收益将为90万元，销路差时不盈不亏；若不扩建，年收益不变. 如果3年销路差，未来7年预计销路一定也差，则不扩建，年收益也不变. 以总利润期望值最大为目标如何做出新的决策呢？

现在变成一个2次决策问题. 第1次决策：当前是新建厂房还是改建厂房经营3年；第2次决策：3年后扩建还是不扩建. 在进行第1次决策时，为了计算10年总利润的期望值必须先对第2次决策的后果做出估计，也就是要从整个过程的终点向前推进. 下面介绍一种简洁、直观的方法解决这类问题.

用决策树模型求解 决策树(decision tree)是表述、分析、求解风险性决策的有效方法[74]，下面先用开始提出的只有新建与改建厂房两种备选方案的问题说明决策树模型的构造.

决策树由节点和分枝组成，用□表示**决策节点**，○表示**状态节点**，◁表示**结果节点**，在由决策节点引向状态节点的分枝(直线)上标注不同的方案，在由状态节点引向结果节点的分枝(直线)上标注发生的概率，在结果节点处标注取得的收益值(投资标以负值). 对于新建或改建厂房的问题，决策树模型的构造如图1所示.

按照结果节点的收益值和状态节点引出分枝上的概率，计算出每个状态节点的期望值并作比较，取最大值的方案为决策，用//表示“砍掉”该方案分枝. 新建或改建厂房问题决策树模型的求解见图2，结果与前面的直接计算相同.

新建还是改建的问题比较简单，决策树方法的优点并不明显，让我们用决策树模型继续讨论加入改建厂房经营3年新方案(方案3)后的决策问题. 第1次决策在方案1，

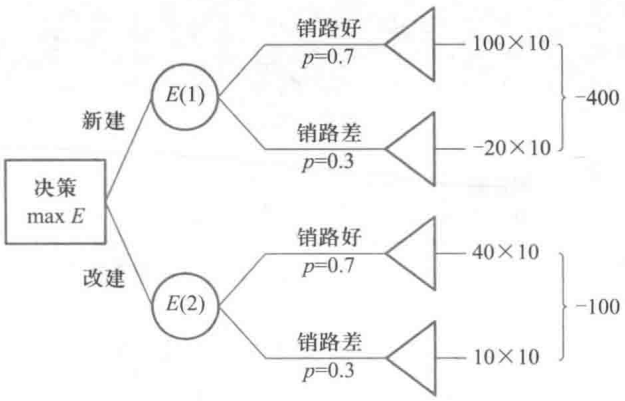


图 1 新建还是改建问题决策树的构造

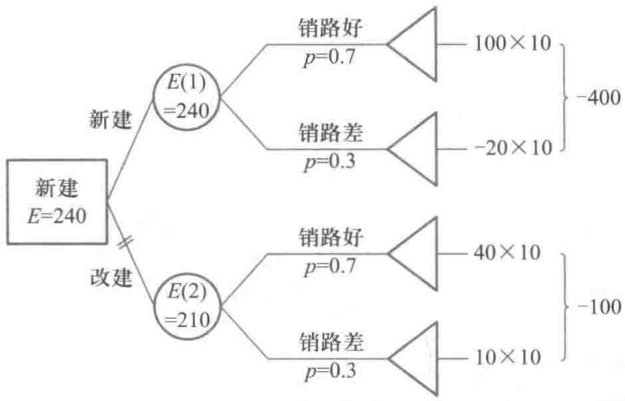


图 2 新建还是改建问题决策树的求解

2,3 中选择(实际上由图 2 已将方案 2 淘汰),方案 1,2 的计算同前,只需分析方案 3;若采用方案 3,需在 3 年后作第 2 次决策——扩建或不扩建.决策树的构造如图 3,其中与图 1 相同的部分(方案 1,2)没有画出.

决策树的计算是由结果节点开始,反向进行.求解过程见图 4.首先计算第 2 次决策扩建与不扩建两种方案利润的期望值,记作 $E(3)_1$, $E(3)_2$:

$$E(3)_1 = 90\times7\times0.9-200 = 367$$

$$E(3)_2 = 40\times7\times0.9+10\times7\times0.1 = 259$$

因为 $E(3)_1 > E(3)_2$,所以第 2 次决策选择扩建,利润期望值 $E = E(3)_1 = 367$(万元).“砍掉”不扩建分支.

然后计算第 1 次决策中方案 3 的总利润期望值 $E(3)$:

$$E(3) = (40\times3\times0.7+10\times3\times0.3-100) + (367\times0.7+10\times7\times0.3) = 270.9$$

其中第 1 项是改建厂房经营 3 年的利润期望值(图 4 最后一行),第 2 项是 3 年以后的利润期望值.因为 $E(3) > E(1) = 240$,所以第 1 次决策选择改建厂房经营 3 年,2 次决策的总利润期望值 $E = E(3) = 270.9$(万元).



拓展阅读 7-2
贝叶斯决策

评注 风险性决策的风险在几个随机状态出现的概率相差不大的情况下尤其严重.一种解决的办法是通过有针对性的试验,获取更多的关于随机状态出现概率的信息,以提高某一状态发生的(后验)概率,可以使决策风险明显降低,称为贝叶斯决策,可参考拓展阅读 7-2.

评注 决策树是求解风险性决策常用的手段,具有直观、简便、逻辑关系清晰等优点,对于 2 次或更多次决策(称为序列决策)过程的表述、分析和求解尤其方便.

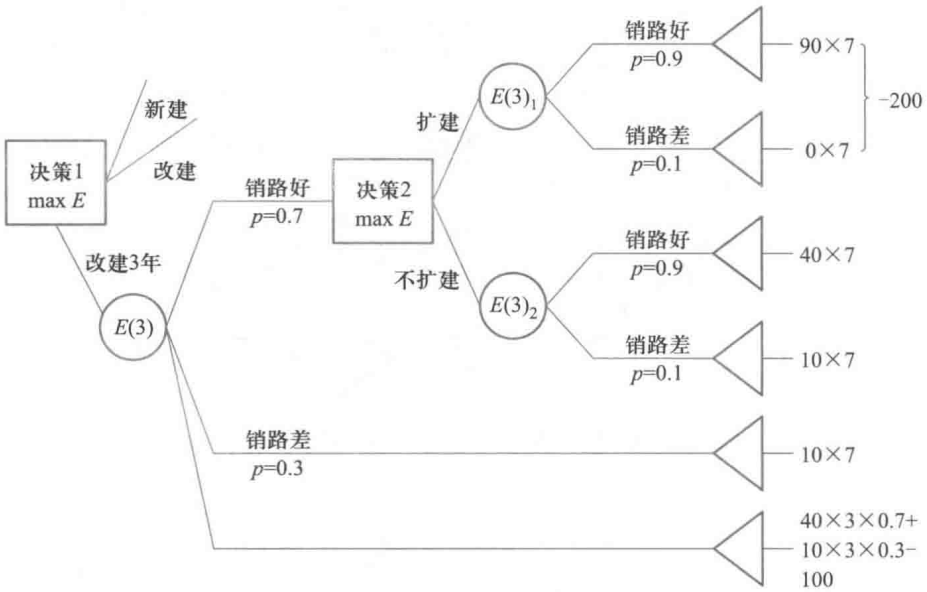


图 3 加入新方案后决策树的构造

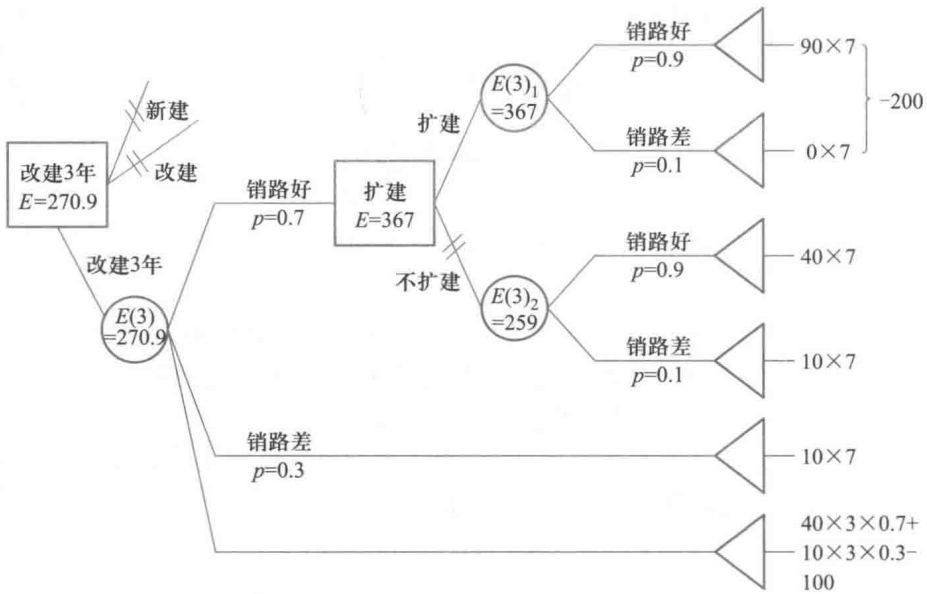


图 4 加入新方案后决策树的求解

复习题

一家公司的管理层打算为扩大产品市场而建造新厂房，现有建造大厂房和小厂房两种方案，对未来市场需求有高、中、低 3 种状态，3 种状态出现的概率及每个方案在各种状态下的收益（单位：万元）如下表[24]. 利用期望值准则和决策树方法确定最优方案，并讨论高需求概率的变化对决策的影响.

方案	状态		
	高需求($p=0.25$)	中需求($p=0.4$)	低需求($p=0.35$)
大厂房	20	10	-12
小厂房	9	7.5	-2

7.4 循环比赛的名次

问题提出 单循环(若干参赛者需要两两交锋且只交锋一次)是很多比赛所采用的赛制,假设每场比赛只计胜负,不计比分,且不允许平局,比赛结束后如何根据比赛结果排列参赛者名次呢?一种常见的表示单循环比赛结果的方式如表 1 所示,这里有 6 支球队,"+"表示对应行的球队战胜了对应列的球队,"-"含义相反.很多比赛规定按照每个队的得分进行排名,如每胜一场计 1 分,败了则计 0 分,表 1 最后 1 列表明 1 队排名第 1,其次是 2、3 队并列,再次 4、5 队并列,6 队排名最后.但 2、3 队之间和 4、5 队之间无法确定排名,如果并列队按他们相互比赛的结果排名,可得到总排名(1,3,2,4,5,6).但这个排名可能不能令人信服:3 队输给了排名最后的 5、6 队,却占据第 2 的名次,2 队很可能不服气;而 3 队认为自己战胜了排名第 1 的 1 队,又战胜了 2 队,排在 2 队前合情合理;4、5 队都只胜两场,除了都战胜排名最后的 6 队外,5 队还战胜了排名第 2 的 3 队,仅仅因为其输给 4 队就将其排在 4 队之后,5 队也很可能不服气.有没有更合理的排名方式呢?[9]

表 1 6 支球队的比赛成绩表

	1	2	3	4	5	6	得分
1	/	+	-	+	+	+	4
2	-	/	-	+	+	+	3
3	+	+	/	+	-	-	3
4	-	-	-	/	+	+	2
5	-	-	+	-	/	+	2
6	-	-	+	-	-	/	1

竞赛图 除了如表 1 所示的表示法外,单循环比赛的结果还可以方便地用有向图表示:图的顶点表示球队,每对顶点间有且只有一条弧,弧的方向表示两支球队的比赛结果(如从胜者指向败者),这样的有向图称为**竞赛图**(tournament).例如,表 1 的比赛结果可表示成图 1.

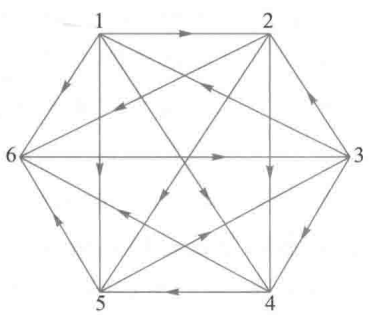


图 1 6 支球队的比赛结果

竞赛图具有以下性质:

1) 必定存在**完全路径**,即从某个顶点出发,沿着弧的方向经过所有顶点的路径,如图 1 中的路径 1→

4→6→3→2→5,3→1→2→4→5→6,4→5→6→3→1→2 等.

2) 若存在唯一的完全路径,则由它确定的顶点顺序与按得分排列的顺序一致.

对于2)的情况,我们已经得到各队唯一的排名;对于2)之外的情况,又可以分为两种情形:a) 竞赛图是双向连通图,即从任意一个顶点出发,至少存在一条有向路径到达另一个任意顶点,如图1;b) 竞赛图不是双向连通图.下面只讨论情形a),因为情形b)发生时无法给出所有队的排名.

双向连通竞赛图的排名

为了便于用代数方法对竞赛图进行研究,采用有向图的邻接矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$($n$ 为顶点数,即球队数):

$$a_{ij}=\begin{cases}1, & \text{存在从 } i \text{ 到 } j \text{ 的有向弧}\\ 0, & \text{否则}\end{cases} \tag{1}$$

图1的邻接矩阵为(实际上就是将表1中对应的"+"变成1,其他元素为0)

$$\boldsymbol{A}=\begin{pmatrix}0&1&0&1&1&1\\0&0&0&1&1&1\\1&1&0&1&0&0\\0&0&0&0&1&1\\0&0&1&0&0&1\\0&0&1&0&0&0\end{pmatrix}$$

如果记顶点的得分向量为 $\boldsymbol{s}=(s_1,s_2,\cdots,s_n)^{\mathrm{T}}$,其中 s_i 是顶点 i 的得分,由邻接矩阵的定义不难得到

$$\boldsymbol{s}=\boldsymbol{A}\boldsymbol{e},\quad \boldsymbol{e}=(1,1,\cdots,1)^{\mathrm{T}} \tag{2}$$

对于6支球队的例子,$\boldsymbol{s}=(4,3,3,2,2,1)^{\mathrm{T}}$,无法由 $\boldsymbol{s}$ 确定全部球队的名次.

得分向量只考虑了球队取胜的场次总数,而前面提到2队、5队可能不服气时,实际上是在进一步关注他们所战胜的那些球队的实力强弱.如何考虑这个因素呢?一种直觉的想法是:如果得分向量是一个队实力的初步估计,那么与战胜一个得分较低的队相比,战胜一个得分较高的队,对自己最后得到的累积分数的贡献应该更大.比如说,可以将所战胜的对手的得分加到自己的累积分数中.按照这样的想法,记 $\boldsymbol{s}^{(1)}=\boldsymbol{s}$,称为1级得分向量,进一步计算

$$\boldsymbol{s}^{(k)}=\boldsymbol{A}\boldsymbol{s}^{(k-1)}=\boldsymbol{A}^k\boldsymbol{e},\quad k=2,3,\cdots \tag{3}$$

称为 k 级得分向量,显然 k 级得分向量比 $k-1$ 级得分向量更能反映球队的实力.

评注 体育比赛的排名是经常遇到的问题,针对不同项目和赛制有多种方法解决.本节讨论的是只计胜负(不设平局、不考虑比分)的单循环比赛,对于这种较简单的比赛形式,竞赛图的排名方法是有充分理论依据的.

对于6支球队的例子,2~4级得分向量分别为

$$\boldsymbol{s}^{(2)}=(8,5,9,3,4,3)^{\mathrm{T}},\quad \boldsymbol{s}^{(3)}=(15,10,16,7,12,9)^{\mathrm{T}}$$

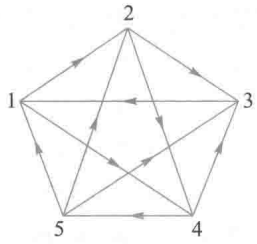
$$\boldsymbol{s}^{(4)}=(38,28,32,21,25,16)^{\mathrm{T}}$$

如果 $k\to\infty$ 时,$\boldsymbol{s}^{(k)}$ 收敛到某个极限得分向量(为了不让它趋向于无穷大,可进行归一化处理),那么就可以用这个向量作为排名的依据.

对于 $n\geqslant 4$ 个顶点的双向连通竞赛图,可以证明存在正整数 r,使得邻接矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^r>\boldsymbol{O}$(表示矩阵的所有元素均大于0),这样的 $\boldsymbol{A}$ 称为素阵.对于素阵 $\boldsymbol{A}$,能够证明其最大特征根为正单根(记为 λ),其对应的特征向量 $\boldsymbol{s}^*$ 满足

$$s^* = \lim_{k\to\infty} \frac{A^k e}{\lambda^k} \tag{4}$$

因此，可以将 s^* 作为排名的依据.对于 6 支球队的例子，用 MATLAB 的 eig 命令算出邻接矩阵 A 的最大特征值为 $\lambda \approx 2.232$，其对应的归一化的特征向量 $s^* = (0.238, 0.164, 0.231, 0.113, 0.150, 0.104)^T$，球队排名为(1,3,2,5,4,6).



复习题

如图为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.

7.5 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.

案例精讲 7-1
公平的席位分配

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分配中小数部分最大的丙系，于是三系仍分别占有 10,6,4 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 年的众议院席位分配(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[18,41]，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 年美国众议院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,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[4,13].

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.

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

$$r_A(n_1,n_2)=\frac{p_1/n_1-p_2/n_2}{p_2/n_2} \tag{1}$$

为对 A 的相对不公平度.若 $p_2/n_2>p_1/n_1$,定义

$$r_B(n_1,n_2)=\frac{p_2/n_2-p_1/n_1}{p_1/n_1} \tag{2}$$

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 $p_1/n_1\geqslant p_2/n_2$,大于号成立时对 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. $\dfrac{p_1}{n_1+1}>\dfrac{p_2}{n_2}$,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. $\dfrac{p_1}{n_1+1}<\dfrac{p_2}{n_2}$,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

$$r_B(n_1+1,n_2)=\frac{p_2(n_1+1)}{p_1n_2}-1 \tag{3}$$

① 美国国会众议院的总席位数从 1787 年的 65 逐渐增加到 1920 年的 435,此后一直固定不变.

3. $\frac{p_1}{n_1} > \frac{p_2}{n_2+1}$，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

$$r_A(n_1, n_2+1) = \frac{p_1(n_2+1)}{p_2 n_1} - 1 \tag{4}$$

(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

$$r_B(n_1+1, n_2) < r_A(n_1, n_2+1) \tag{5}$$

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

$$\frac{p_2^2}{n_2(n_2+1)} < \frac{p_1^2}{n_1(n_1+1)} \tag{6}$$

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，$i=1,2,\cdots,m$，当总席位增加 1 席时，计算

$$Q_i = \frac{p_i^2}{n_i(n_i+1)}, \qquad i = 1,2,\cdots,m \tag{7}$$

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，但对这个问题直到 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

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.

第 21 席：$Q_1 = \frac{103^2}{11\times 12} = 80.37$，$Q_2, Q_3$ 同上，Q_3 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.

Q 值方法是 20 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[4,41].

Huntington 除数法 设共有 m 方分配 N 个席位，第 i 方人数为 p_i，记 $\boldsymbol{p} = (p_1, p_2, \cdots, p_m)$，$P = \sum_{i=1}^{m} p_i$.席位分配是要寻求 $\boldsymbol{n} = (n_1, n_2, \cdots, n_m)$，其中 n_i 是第 i 方分配的

评注 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(也可证明不会出现人口悖论).

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 $q_i = N\frac{p_i}{P}, i = 1,2,\cdots,m$，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 $\boldsymbol{p}, s$ 的函数，记作 $f_i(\boldsymbol{p}, s)$，而且 $f_i(\boldsymbol{p}, 0) = 0, i = 1,2,\cdots, m$，让 s 每次 1 席地递增至 N，按照以下准则分配①：

设 $n_i = f_i(\boldsymbol{p}, s)$，若对于某个 k，有

$$\frac{p_k}{d(n_k)} = \max_{i=1,2,\cdots,m} \frac{p_i}{d(n_i)} \tag{8}$$

则令

$$f_k(\boldsymbol{p}, s+1) = n_k + 1, \quad f_i(\boldsymbol{p}, s+1) = n_i (i \neq k) \tag{9}$$

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 $\frac{p_i}{n_i} \geqslant \frac{p_j}{n_j}$)	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

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.

续表

Huntington 除数法	除数 $d(n)$	不公平度的度量指标（设 $\frac{p_i}{n_i} \geqslant \frac{p_j}{n_j}$）	以人名命名的称谓
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	$n_j - \frac{n_i p_j}{p_i}$	Adams

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 $\frac{p_i/n_i - p_j/n_j}{p_j/n_j} = \frac{n_j p_i}{n_i p_j} - 1$ 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作一种均衡检验，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（留作复习题 1，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 $n \leqslant d(n) \leqslant n+1$，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院席位分配中采用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 $\boldsymbol{n} = (n_1, n_2, \cdots, n_m)$ 满足以下 min-max 不等式：

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.

$$\min_{k=1,2,\cdots,m}\frac{p_k}{d(n_k-1)}\geqslant\max_{i=1,2\cdots,m}\frac{p_i}{d(n_i)},\ \sum_{i=1}^{m}n_i=N \tag{10}$$

从优化的角度出发,席位分配问题可以归结为,给定 $\boldsymbol{p}$ 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 $n_1,n_2,\cdots,n_m$ 使某一事先定义的目标函数最优.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.

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.

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[4].

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,$n_i=f_i(\boldsymbol{p},s)$ 表示人数为 $\boldsymbol{p}$、总席位为 s 时分配给第 i 方的席位,$i=1,2,\cdots,m$):

1. $\lfloor q_i\rfloor\leqslant n_i\leqslant\lceil q_i\rceil$,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. $f_i(\boldsymbol{p},s)\leqslant f_i(\boldsymbol{p},s+1)$,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 $i,j=1,2,\cdots,m,j\neq i$,$p_i'/p_j'\geqslant p_i/p_j$,则 $f_i(\boldsymbol{p}',s)\geqslant f_i(\boldsymbol{p},s)$ 或 $f_j(\boldsymbol{p}',s)\leqslant f_j(\boldsymbol{p},s)$,即当 i 方相对于 j 方人数增加时(总席位不变),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[4]：对于 $m \geqslant 4, N \geqslant m+3$，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)[4,41].

在席位分配问题中，影响分配结果的因素有人口 $\boldsymbol{p}$，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 $i,j=1,2,\cdots,m, j\neq i, i', j'=1,2,\cdots,m', j'\neq i', p'_{i'}/p'_{j'} \geqslant p_i/p_j$，则 $f_{i'}(\boldsymbol{p}',s') \geqslant f_i(\boldsymbol{p},s)$ 或 $f_{j'}(\boldsymbol{p}',s') \leqslant f_j(\boldsymbol{p},s)$ 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[41]及其参考文献中查到.

评注 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性的概率仅为 1/1 600[4]. 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.

复习题

1. 验证表 3 用 5 种除数法得到的结果.

2. 证明主要分数法(MF)给出了 $\min \sum_{i=1}^{m} p_i \left(\frac{n_i}{p_i} - \frac{N}{P} \right)^2$ 的解.

7.6 存在公平的选举吗

在现代社会的公众生活和政治活动中，有很多事情是靠投票选举的办法决定的.从推选班长、队长，到竞选市长、总统；从评选最佳影片、优秀运动员，到推举旅游胜地、宜居城市；国际奥委会经过层层筛选、一轮一轮地投票选出奥运会举办城市的过程更为世人瞩目.无论什么层次的选举，人们都希望公平、公正，可是，世界上存在公平的选举吗？面对这个问题，你当然会问“什么是真正的公平”，这正是本节要讨论的重点之一[37].

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.

我们对投票选举作如下的表述和约定:每位选民对所有候选人按照偏爱程度所做的排序称为一次投票,根据全体选民的投票确定哪位候选人是获胜者的规则称为选举方法;按照公认的民主法则和选民的理性行为,对选举的公平、公正做出的一些规定,称为公平性准则.下面介绍几种选举方法和几条公平性准则,并讨论是否存在满足全部公平性准则的选举方法,由此引入著名的 Arrow 不可能性定理,最后给出选举方法和公平性准则的两个应用实例①.

先看一个虚拟的例子:选举班级喜爱的球队.

一个班 30 名学生从 A,B,C,D 共 4 支球队中投票选举班级喜爱的球队,每个学生按照自己的偏爱程度将球队从第 1 名排到第 4 名的投票结果如表 1.

表 1 30 名学生对喜爱球队的投票结果

票数	11	10	9
第 1 名	B	C	A
第 2 名	D	D	D
第 3 名	C	A	C
第 4 名	A	B	B

如果只比较哪支球队得第 1 名的票数最多,那么获胜者是 B,但是有近 2/3 的学生将 B 排到最后一名;也可以认为获胜者是 C,因为 C 不仅没有排到最后,而且得第 1 名只比 B 少 1 票;还可以考虑 D,因为没有学生把 D 排到第 2 名以后.那么怎样决定获胜者呢?这就取决于采用什么选举方法了.

选举方法集锦 这里介绍 5 种选举方法.

1. 简单多数法 得到第 1 名票数最多的候选人为获胜者.

这个方法易于实施,但是缺点很明显,因为没有考虑到除排在第 1 名的候选人以外的全部排序信息,甚至被大多数选民排在最后的候选人都有可能是获胜者,正如选举喜爱球队例子中的获胜者 B(见表 1).显然,好的选举方法应该考虑更多的排序信息.

2. 单轮决胜法 得到第 1 名票数最多和次多的两位候选人进入单轮决胜投票,由简单多数法决定决胜投票的获胜者,并确定为整个选举的获胜者.

在选举喜爱球队例子中,得到第 1 名票数最多和次多的是 B, C, 按照单轮决胜法,这 2 支球队进入决胜投票.假定所有学生对 B, C 的偏爱没有改变,那么决胜投票的结果如表 2.

表 2 30 名学生对喜爱球队单轮决胜投票的结果

票数	11	10	9
第 1 名	B	C	C
第 2 名	C	B	B

① 规定候选人不少于 2 位.若只有 2 位候选人,得到第 1 名票数占多数(超过半数)的候选人为获胜者;如果 2 人票数相同,不妨用随机抽签(等概率)的办法确定一位获胜者.另外,在本节介绍的各种选举方法中,如果出现 2 位候选人得票数相同的情况,可以针对 2 人进行一次附加投票或随机抽签,从中选出 1 人.

按照单轮决胜法选举的获胜者是 C.

3. 系列决胜法 进行多轮决胜投票,每轮只淘汰得第 1 名票数最少的候选人,当剩下两位候选人时,由简单多数法决定获胜者,并确定为整个选举的获胜者.

在选举喜爱球队的例子中,初次投票得到第 1 名票数最少的是 D,按照系列决胜法进入第 2 轮的球队是 A,B,C.假定所有学生对这些球队的偏爱没有改变(以下均作偏爱不变的假定,不再重复),那么第 2 轮的投票结果如表 3.

表 3 按照系列决胜法对喜爱球队第 2 轮的投票结果

票数	11	10	9
第 1 名	B	C	A
第 2 名	C	A	C
第 3 名	A	B	B

按照系列决胜法应该淘汰 A,进入第 3 轮的是 B,C,投票结果如表 4.

表 4 按照系列决胜法对喜爱球队第 3 轮的投票结果

票数	11	10	9
第 1 名	B	C	C
第 2 名	C	B	B

于是,按照系列决胜法,整个选举的获胜者是 C.

4. Coombs 法 除了每轮不是淘汰第 1 名票数最少的,而是淘汰倒数第 1 名票数最多的候选人以外,其余与系列决胜法相同.这种办法是美国心理学家 Clyde Coombs (1912—1988)提出的.娱乐游戏中常常采用系列决胜法和 Coombs 法.

在选举喜爱球队的例子中,初次投票倒数第 1 名票数最多的是 B,按照 Coombs 法进入第 2 轮的是 A,C,D,第 2 轮的投票结果如表 5.

表 5 按照 Coombs 法对喜爱球队第 2 轮的投票结果

票数	11	10	9
第 1 名	D	C	A
第 2 名	C	D	D
第 3 名	A	A	C

倒数第 1 名票数最多的 A 被淘汰,进入第 3 轮的是 C,D,第 3 轮的投票结果如表 6.

表 6 按照 Coombs 法对喜爱球队第 3 轮的投票结果

票数	11	10	9
第 1 名	D	C	D
第 2 名	C	D	C

于是,按照 Coombs 法,最终获胜者是 D.

评注 开头曾约定,每位选民要对所有候选人按照偏爱程度排序,而从5种选举方法可以看到,简单多数法、单轮决胜法和系列决胜法都只考虑第1名的票数,所以对于实际生活中那些只要求选民投票给最喜爱的一位候选人的情况,这3种选举方法仍然适用.实际上,只有Borda计数法用到了选民的全部排序信息.

5. Borda计数法 对每一张选票,排倒数第1名的候选人得1分,排倒数第2名的得2分,依此下去,排第1名的得分是候选人的总数.将全部选票中各位候选人的得分求和,总分最高的为获胜者.这种方法以法国天文学家、海军军官Jean-Charles de Borda(1733—1799)命名.在美国大学的篮球、橄榄球队的民意调查中,经常采用Borda计数法选出最受欢迎的球队和球星.

在选举喜爱球队的例子中,A排第1名9票,第3名10票,第4名11票,A的总分为$9\times4+10\times2+11\times1=67$(分).同样地计算B的总分为$11\times4+19\times1=63$(分),C的总分为$10\times4+20\times2=80$(分),D的总分为$30\times3=90$(分).按照Borda计数法获胜者是D.

我们看到,在选举喜爱球队的例子中(表1),按照简单多数法获胜者是B,按照单轮决胜法和系列决胜法获胜者是C,按照Coombs法和Borda计数法获胜者是D.这个结果表明,哪位候选人是获胜者不仅取决于选民的偏爱,而且与采用的选举方法有关.

在选民对候选人同样的偏爱下,不同的选举方法可能产生不同的获胜者,这是一个既有趣又令人不安的现象.人们自然要问:哪一种选举办法才是公平的?

直接回答这个问题是困难的.让我们换一个角度思考,讨论在公认的民主法则和选民的理性行为下,选举方法应该满足哪些所谓"公平性准则",上面的几种方法满足或者违反了其中哪些准则.

选举中的公平性准则 下面给出5条准则.

1. 多数票准则 得到第1名票数超过选民半数的候选人应当是获胜者.

一种特定的选举方法或者满足,或者违反这个准则.例如,简单多数法满足多数票准则,因为若一位候选人得到第1名票数超过半数,那么他的票数必定多于其他候选人,按照简单多数法他必定是获胜者.类似地,单轮决胜法和系列决胜法也满足多数票准则.

应该正确理解多数票准则的含义.比如这个准则不能倒过来说成,获胜者必须是第1名得票数超过半数的.多数票准则也不能用于选举喜爱球队的例子,因为没有球队第1名票数超过半数,所以谈不上哪种选举方法是否违反这条准则.

Borda计数法满足还是违反多数票准则呢?请看这样一个例子:27位选民对4位候选人的投票结果如表7.

表7 27位选民对4位候选人的投票结果

票数	9	8	5	4	1
第1名	C	A	C	A	B
第2名	D	D	D	B	A
第3名	A	B	B	D	D
第4名	B	C	A	C	C

按照Borda计数法计算各位候选人的得分:D为$22\times3+5\times2=76$(分),C为69分,B为51分,A为74分,所以获胜者是D.但是C得第1名票数是14,超过选民的半数,按照多数票准则获胜者应该是C.

显然,在这个例子中Borda计数法违反多数票准则,当然,并不是说采用Borda计数法每次选举都会违反这个准则.

2. Condorcet 获胜者准则 如果候选人 X 在与每一位候选人的两两对决中都获胜(按照多数票准则),那么 X 应当是获胜者.这条准则以法国哲学家、政治学家 Marquis de Condorcet (1743—1794)命名,以下简称获胜者准则.

评注 说明某种选举方法违反一条公平性准则,只需举出一个例子,但是论证某种选举方法满足一条公平性准则,必须对于任意的投票,选举结果都不能违反准则,这就不能举例说明,而需要给以证明.

考察选举喜爱球队例子中球队两两对决的结果:由表 1 可知在 D 和 B 的对决中,19 名学生将 D 排在 B 前面,11 名学生将 B 排在 D 前面,D 与 B 的票数之比是 19∶11,D 获胜.类似地,D 与 C 的票数之比是 20∶10,D 与 A 的票数之比是 21∶9.D 在所有的两两对决中都获胜,按照获胜者准则 D 是这次选举的获胜者.对于表 1 给出的投票结果,若某种选举方法的获胜者不是 D,那就一定违反获胜者准则.这样,简单多数法(获胜者是 B)、单轮决胜法和系列决胜法(获胜者是 C)都违反获胜者准则.

虽然在选举喜爱球队的例子中 Coombs 法和 Borda 计数法的获胜者正好是 D,但是据此并不能说明这两种方法满足获胜者准则.

3. Condorcet 失败者准则 如果候选人 Y 在与每一位候选人的两两对决中都未获胜,那么 Y 不应当是获胜者.以下简称失败者准则.

考察选举喜爱球队例子中 B 与其他球队两两对决的结果:由表 1 可知 B 与 D,C,A 的票数之比都是 11∶19,按照失败者准则 B 不应当是获胜者.对于表 1 给出的投票结果,若某种选举方法的获胜者是 B,那就一定违反失败者准则,所以简单多数法(获胜者是 B)违反这条准则.

4. 无关候选人的独立性准则 假定在最终排序中候选人 X 领先于候选人 Y,如果其他一位候选人退出选举,或者一位新的候选人进入选举,那么在最终排序中候选人 X 仍领先于候选人 Y.这条准则的意思是,候选人的最终排序与其他候选人的退出或进入无关,以下简称独立性准则.

考察用系列决胜法确定选举喜爱球队例子中获胜者的过程,根据被淘汰的顺序,4 支球队最终排序应为 C,B,A,D.假定 B 退出选举,投票结果将如表 8.

表 8 学生对喜爱球队的投票结果(B 退出选举)

票数	11	10	9
第 1 名	D	C	A
第 2 名	C	D	D
第 3 名	A	A	C

A 被淘汰,进入第 2 轮的是 C,D,投票结果将如表 9.

表 9 用系列决胜法的第 2 轮投票结果(B 退出选举)

票数	11	10	9
第 1 名	D	C	D
第 2 名	C	D	C

于是在 B 退出选举后用系列决胜法对其余 3 支球队的最终排序是 D,C,A.D 排到了 C,A 前面,而 B 未退出时 D 是排在最后的,所以系列决胜法违反了独立性准则.

5. 单调性准则 假定候选人 X 在一次选举的最终排序中居于某个位置,如果某些

选民只将 X 的顺序提前而其他候选人的排序不变,那么对于新的选举,X 在最终排序中的位置不应在原来位置的后面.这条准则的意思是,若选民对 X 的排序没有后移,那么最终排序中 X 相对其他候选人的优先性不应改变.

可以证明,简单多数法和 Borda 计数法满足单调性准则.虽然违反单调性准则看来似乎是荒谬的,然而后面的实例说明单轮决胜法和系列决胜法都违反这条准则.

上面讨论了 5 种选举方法和 5 条公平性准则,没有一种选举方法满足全部公平性准则①,虽然我们没有考察当今社会上使用的所有选举方法,但是还需要沿着这条路继续研究下去,竭力搜寻满足全部准则的那种方法吗?

评注 回顾从研究选举方法、公平性准则到发现不可能性定理的过程,可以认识到用公理化方法建模的重要性.提出的公平性准则被人们认为是(无须证明的)公理,从公理出发进行逻辑推理,公理不合适(过严、过多甚至相互矛盾)会导致得不到满足公理的结论,公理不充分(过松、过少等)可能无法推出结果或者结果不唯一.所以这种建模方法常常需要多次反复(参看拓展阅读 7-3).

Arrow 不可能性定理

美国经济学家 K. Arrow 1951 年开始研究选举理论时,先列出自己认为的公平性准则,并且为找不到满足所有准则的选举方法而困扰,他不断修正提出的准则,在一再尝试但是没有结果之后改变了思路.后来他这样表述这段经历:"我开始有种想法,也许不存在满足所有我认为是合理条件的选举方法,我着手去证明这点,实际上只用了不过几天时间."

Arrow 用不过几天时间得到的是选举理论历史上最重要的结果,现在称为 Arrow 不可能性定理,正是由于这个成果,Arrow 获得了 1972 年诺贝尔经济学奖.

Arrow 不可能性定理(修正版本②) 任何一种选举方法都至少违反下列 4 条准则之一:多数票准则、获胜者准则、独立性准则、单调性准则.

应该确切地理解这个定理.比如它不是说,按照任何选举方法进行的每一次投票都至少违反 4 条准则之一,而是表明,对于任意的选举方法,总可以发现选民对候选人的一次投票,使得在这种选举方法下投票结果至少违反上述 4 条准则之一.

面对这个"残酷"的事实,人们想到的是重新审查那些所谓公平性准则,其中最受人质疑的是独立性准则,认为这条准则太强势,要求太高.D. Saari 提出一条考虑到选民对候选人偏爱强度的、比独立性更弱的准则,用于代替原来的独立性准则,从而得到了所谓 Saari 可能性定理,即存在满足这些准则的选举方法.事实上,Borda 计数法就是满足可能性定理中全部准则的一种选举方法.关于 Arrow 不可能性定理的原始版本、Saari 可能性定理的内容,以及其他一些选举方法,可参看拓展阅读 7-3.



拓展阅读 7-3
公平选举的再讨论

选举方法和公平性准则的应用实例

● **系列决胜法在推选 2004 年奥运会举办城市中的应用及讨论**

国际奥委会采用系列决胜法选择奥运会举办城市.通常先从申办城市中挑出几座候选城市,然后奥委会成员进行几轮投票,投票时不要求对候选城市排序,只要求投给最偏爱的一座城市,每轮将得票最少的城市淘汰,直至选出获胜城市.

1. 2004 年奥运会举办城市的推选过程

2004 年奥运会的候选城市是雅典、布宜诺斯艾利斯、开普敦、罗马、斯德哥尔摩,在奥委会成员的第 1 轮投票中,5 座城市的票数如表 10 第 2 行.

因为得票最少的布宜诺斯艾利斯和开普敦票数相同,按照规定需对这两座城市进

① 虽然这里没有给出 Coombs 法违反任何准则的例子,然而它确实违反多数票准则和单调性准则.

② 许多讨论 Arrow 的结果的书都采用与 Arrow 自己写的有些差别的表述,这里是其中之一.

行一轮附加投票，得票少的淘汰.投票结果布宜诺斯艾利斯被淘汰，其他 4 座城市进入下一轮.第 2 轮投票结果如表 10 第 3 行.

斯德哥尔摩被淘汰，其他 3 座城市进入下一轮.第 3 轮投票结果如表 10 第 4 行①.

开普敦被淘汰，其他 2 座城市进入最后一轮.第 4 轮投票结果如表 10 第 5 行.

最终获胜者是雅典.

表 10 2004 年奥运会举办城市各轮投票结果

城市	雅典	布宜诺斯艾利斯	开普敦	罗马	斯德哥尔摩
第 1 轮票数	32	16	16	23	20
第 2 轮票数	38		22	28	19
第 3 轮票数	52		20	35	
第 4 轮票数	66			41	

2. 采用系列决胜法在推选过程中可能出现的情况

尽管开普敦在第 1 轮投票中是得票最少的城市之一，但并不是在淘汰布宜诺斯艾利斯后下一个被淘汰的.如果第 2 轮投票中投给斯德哥尔摩的 19 票在第 3 轮投票中都投给开普敦②，那么开普敦将得到 22+19=41 票，第 3 轮淘汰的将是罗马，而开普敦将在最后一轮投票中与雅典对决.二者之中哪个会获胜？如果投给罗马的 35 票在最后一轮投票中都投给开普敦③，那么开普敦将是最终的获胜者.这就是说，采用系列决胜法在最初一轮投票中差点被淘汰的城市也可能最终获胜！

3. 采用系列决胜法可能违反单调性准则

在用系列决胜法的第 1 轮投票中（见表 10），假定原来投给开普敦的票中有一票转投给雅典，只发生有利于雅典的这一点改变，可使第 1 轮投票结果如表 11 第 2 行.

于是第 1 个被淘汰的城市不是布宜诺斯艾利斯，而是开普敦.剩下 4 座城市进入下一轮，我们无法确切知道会有什么结果，但是出现表 11 第 3 行的情况是可能的.

斯德哥尔摩被淘汰，第 3 轮投票结果可能如表 11 第 4 行.

罗马被淘汰，雅典和布宜诺斯艾利斯将进入最后一轮的竞争，而布宜诺斯艾利斯有可能最终获胜.

表 11 2004 年奥运会举办城市各轮投票结果（假定第 1 轮投给开普敦的票中有一票转投给雅典）

城市	雅典	布宜诺斯艾利斯	开普敦	罗马	斯德哥尔摩
第 1 轮票数	33	16	15	23	20
第 2 轮票数	33	31		23	20
第 3 轮票数	33	51		23	

① 从表 10 发现，斯德哥尔摩第 2 轮得票比第 1 轮少 1 票，开普敦第 3 轮得票比第 2 轮少 2 票，这违反了系列决胜法关于选民对候选人偏爱不变的假定，不过这一点改变不会影响结果.

② 实际上只需 19 票中投给开普敦的比投给罗马的多 7 票，这是完全可能的.

③ 这是完全可能的，实际上也不需要这么多.

我们看到,在用系列决胜法的选举过程中,第1轮投票雅典位居第1,而有利于雅典的一点点改变(投给开普敦的票中有一票转投给雅典),在最终排序中却可以使雅典落到第2位.如果真的如此(这是完全可能的),单调性准则就不成立了.

● 单轮决胜法在2002年法国总统选举中的应用及讨论

法国总统选举采用单轮决胜法.在初次投票中不要求选民对候选人排序,只投票给最偏爱的一位候选人,获得票数最多和次多的两位候选人进入决胜投票,决胜投票由简单多数法决定获胜者.

1. 2002年法国总统的选举过程

在2002年法国总统选举中进入初次投票的有16位候选人,其得票百分比如表12.

表12 2002年法国总统选举初次投票中候选人的得票百分比(略去的候选人得票低于6%)

候选人	Chirac	Le Pen	Jospin	Bayrou	…
得票百分比	19.88%	16.86%	16.18%	6.84%	…

得票最多和次多的Chirac和Le Pen进入决胜投票,结果Chirac以82%的绝对优势获胜.

2. 采用单轮决胜法在选举过程中可能出现的情况

在初次投票中Chirac,Le Pen和Jospin 3位候选人得票都超过15%且相差不大,其中Chirac是著名右翼政治家,1995—2007年任法国总统,Le Pen是极右翼政治家,有60%以上选民反对他,Jospin是左翼政治家,1997—2002年任法国总理.本来Chirac和Jospin是被看好能够进入决胜投票的势均力敌的两位候选人,但是Le Pen以稍多一点的票数挤掉了Jospin.然而几乎可以肯定,假若不用单轮决胜法而是采用系列决胜法的话,如果选民的偏爱不变,Chirac和Jospin将有一场决战,结果难料.

3. 采用单轮决胜法可能违反单调性准则

假定在初次投票中有1%的选民将原来投给Le Pen的票转投Chirac,这会使Chirac的得票由19.88%上升到20.88%,而Le Pen的得票由16.86%下降到15.86%,低于Jospin的16.18%.这样在决胜投票中竞选的将是Chirac和Jospin.虽然无法知道谁会在决胜投票中获胜,但Jospin确有胜出的可能.而一旦Jospin获胜,就违反了单调性准则.这次选举中按照选民最初的真实意愿排序是Chirac位居第1,而好像有利于Chirac的1%选民的转投这一点点改变,却可能使Chirac在最终排序中落到第2位.

后记 15年之后的2017年法国总统选举中,进入初次投票的11位候选人的得票百分比如表13.

表13 2017年法国总统选举初次投票中候选人的得票百分比(略去的候选人得票低于7%)

候选人	Macron	Le Pen	Fillon	Melenchon	…
得票百分比	24.01%	21.30%	20.01%	19.58%	…

得票最多和次多的Macron和Le Pen(表12中Le Pen的女儿)进入决胜投票,结果Macron以66%的优势获胜.

对比这两次选举可以发现,旧的一幕几乎重演,采用单轮决胜法同样可能出现违反

评注 本节讨论的选举过程可以看作一种群体决策,即根据若干人

单调性准则的情况.

对某些对象的个体决策结果(如每位选民的投票),综合得到这个群体的决策.社会上有许多专门机构通过民意调查了解部分群众对社会福利、内外政策及领导人的态度,然后用群体决策方法归纳出国民的整体倾向.

4. 单调性准则与虚假投票

上面所说的这个诡异的结果清楚地说明单调性准则的重要性.那些最偏爱 Chirac 将他排到第 1 位的选民,可能会伤害 Chirac 在最终排序中的位置! 对于这些选民来说,更好的策略是把一部分票投给 Le Pen(即使很不喜欢他),增加 Le Pen 领先 Jospin 进入决胜投票的可能性(他们知道在决胜投票中 Chirac 一定能击败 Le Pen),以避免让 Jospin 进入有可能胜过 Chirac 的决胜投票.

这就是说,单轮决胜法对于虚假投票非常敏感.所谓虚假投票是指:投票不反映选民的真实意愿,而是帮助所喜爱的候选人在最终结果中得到某个位置.

复习题

1. 19 位选民对 4 位候选人的投票结果如下表.

票数	7	5	4	3
第 1 名	B	D	C	A
第 2 名	A	C	D	C
第 3 名	D	A	A	D
第 4 名	C	B	B	B

(1) 分别用简单多数法、单轮决胜法、系列决胜法、Coombs 法和 Borda 计数法确定获胜者.

(2) 对于本题,前 4 种方法的结果违反多数票准则、获胜者准则、失败者准则吗? 若是,哪种方法违反? 若不是,做出解释.

(3) 假定 C 退出选举,选民对其他 3 位候选人的偏爱不变,重新用 Borda 计数法确定获胜者.这两个选举结果违反独立性准则、单调性准则吗? 做出解释.

2. 下表是本节已经说明的 5 种选举方法满足/违反 5 条公平性准则的情况,试尽可能完成表格("满足"需说明理由,"违反"需举出例子).

	多数票准则	获胜者准则	失败者准则	独立性准则	单调性准则
简单多数法	满足	违反	违反		满足
单轮决胜法	满足	违反			违反
系列决胜法	满足	违反		违反	违反
Coombs 法					
Borda 计数法	违反				满足

7.7 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[8,82].

可以看出,这个问题的提法与7.6节有相似之处,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

$$I=\frac{p}{p^0} \tag{1}$$

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

$$I=\frac{p_1}{p_1^0}+\frac{p_2}{p_2^0} \quad 或 \quad \frac{p_1+p_2}{p_1^0+p_2^0} \quad 或 \quad \frac{p_1p_2}{p_1^0p_2^0} \tag{2}$$

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 $p_1^0,p_2^0,\cdots,p_n^0$,在现年的价格为 $p_1,p_2,\cdots,p_n$,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 $q_1^0,q_2^0,\cdots,q_n^0$ 和 $q_1,q_2,\cdots,q_n$,可以用它们的销售量、销售额或其他指标作为权重.记向量 $\boldsymbol{p}^0=(p_1^0,p_2^0,\cdots,p_n^0)^{\mathrm{T}}$,$\boldsymbol{q}^0=(q_1^0,q_2^0,\cdots,q_n^0)^{\mathrm{T}}$,$\boldsymbol{p}=(p_1,p_2,\cdots,p_n)^{\mathrm{T}}$,$\boldsymbol{q}=(q_1,q_2,\cdots,q_n)^{\mathrm{T}}$,价格指数可记作 $I(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)$.下面列举经济学家们提出来的一些价格指数.

$$I_1(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)=\frac{\boldsymbol{p}\cdot\boldsymbol{q}^0}{\boldsymbol{p}^0\cdot\boldsymbol{q}^0}=\frac{\sum_{i=1}^{n}p_iq_i^0}{\sum_{i=1}^{n}p_i^0q_i^0} \tag{3}$$

$$I_2(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)=\frac{\boldsymbol{p}\cdot\boldsymbol{q}}{\boldsymbol{p}^0\cdot\boldsymbol{q}}=\frac{\sum_{i=1}^{n}p_iq_i}{\sum_{i=1}^{n}p_i^0q_i} \tag{4}$$

$$I_3(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)=\frac{\boldsymbol{p}\cdot\boldsymbol{q}}{\boldsymbol{p}^0\cdot\boldsymbol{q}^0}=\frac{\sum_{i=1}^{n}p_iq_i}{\sum_{i=1}^{n}p_i^0q_i^0} \tag{5}$$

$$I_4(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)=\frac{\boldsymbol{p}\cdot\boldsymbol{a}}{\boldsymbol{p}^0\cdot\boldsymbol{a}}=\frac{\sum\limits_{i=1}^{n}p_ia_i}{\sum\limits_{i=1}^{n}p_i^0a_i},\quad a_i>0 \tag{6}$$

$$I_5(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)=\left(\frac{\boldsymbol{p}\cdot\boldsymbol{q}^0}{\boldsymbol{p}^0\cdot\boldsymbol{q}^0}\frac{\boldsymbol{p}\cdot\boldsymbol{q}}{\boldsymbol{p}^0\cdot\boldsymbol{q}}\right)^{\frac{1}{2}} \tag{7}$$

$$I_6(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)=\prod_{i=1}^{n}\left(\frac{p_i}{p_i^0}\right)^{\alpha_i},\quad \alpha_i>0,\ \sum_{i=1}^{n}\alpha_i=1 \tag{8}$$

$$I_7(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)=\prod_{i=1}^{n}\left(\frac{p_i}{p_i^0}\right)^{\beta_i},\quad \beta_i=\frac{q_i^0}{\sum\limits_{i=1}^{n}q_i^0} \tag{9}$$

$$I_8(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)=\prod_{i=1}^{n}\left(\frac{p_i}{p_i^0}\right)^{\gamma_i},\quad \gamma_i=\frac{q_i}{\sum\limits_{i=1}^{n}q_i} \tag{10}$$

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 $p_i,q_i,p_i^0,q_i^0>0(i=1,2,\cdots,n)$,价格指数 $I(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)>0$,应满足以下 8 条公理:

1. 若 $\tilde{\boldsymbol{p}}>\boldsymbol{p}$(指对所有的 i,$\tilde{p}_i\geqslant p_i$,且至少有一个 i,$\tilde{p}_i>p_i$),则 $I(\tilde{\boldsymbol{p}},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)>I(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)$.

2. $I(\boldsymbol{p}^0,\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)=1$.

3. $I(k\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)=kI(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)$, $k>0$.

4. $\min\limits_i \dfrac{p_i}{p_i^0} \leqslant I(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0) \leqslant \max\limits_i \dfrac{p_i}{p_i^0}$.

5. $I(\lambda\boldsymbol{p},\boldsymbol{q}|\lambda\boldsymbol{p}^0,\boldsymbol{q}^0)=I(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)$, $\lambda>0$.

6. $I(\boldsymbol{\Lambda p},\boldsymbol{\Lambda}^{-1}\boldsymbol{q}|\boldsymbol{\Lambda p}^0,\boldsymbol{\Lambda}^{-1}\boldsymbol{q}^0)=I(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)$, 其中 $\boldsymbol{\Lambda}=\mathrm{Diag}[\lambda_1,\lambda_2,\cdots,\lambda_n]$, $\lambda_i>0(i=1,2,\cdots,n)$.①

7. $\dfrac{I(\tilde{\boldsymbol{p}},\tilde{\boldsymbol{q}}|\boldsymbol{p}^0,\boldsymbol{q}^0)}{I(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)}=\dfrac{I(\tilde{\boldsymbol{p}},\tilde{\boldsymbol{q}}|\overline{\boldsymbol{p}}^0,\overline{\boldsymbol{q}}^0)}{I(\boldsymbol{p},\boldsymbol{q}|\overline{\boldsymbol{p}}^0,\overline{\boldsymbol{q}}^0)}$.

8. $\lim\limits_{p_i\to 0} I(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)>0$.②

用这8条公理一一检验(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_5 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 $\boldsymbol{p},\boldsymbol{q},\boldsymbol{p}^0,\boldsymbol{q}^0$ 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[38].

定理 不存在同时满足公理2,3,6,7,8的价格指数 $I(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)$.

证明 其过程是证明任何一个满足公理2,3,6,7的指数 $I(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)$ 必然不满足公理8.

记 $\boldsymbol{e}=(1,1,\cdots,1)^{\mathrm{T}}$, $\boldsymbol{C}=\mathrm{Diag}[c_1,c_2,\cdots,c_n]$, $\boldsymbol{D}=\mathrm{Diag}[d_1,d_2,\cdots,d_n]$, $(c_i,d_i>0,i=1,2,\cdots,n)$. 容易推出

$$I(\boldsymbol{CDe},\boldsymbol{C}^{-1}\boldsymbol{D}^{-1}\boldsymbol{e}|\boldsymbol{e},\boldsymbol{e})=\frac{I(\boldsymbol{CDe},\boldsymbol{C}^{-1}\boldsymbol{D}^{-1}\boldsymbol{e}|\boldsymbol{e},\boldsymbol{e})}{I(\boldsymbol{Ce},\boldsymbol{C}^{-1}\boldsymbol{e}|\boldsymbol{e},\boldsymbol{e})}\cdot I(\boldsymbol{Ce},\boldsymbol{C}^{-1}\boldsymbol{e}|\boldsymbol{e},\boldsymbol{e})$$

$$\overset{\text{公理7}}{=\!=}\frac{I(\boldsymbol{CDe},\boldsymbol{C}^{-1}\boldsymbol{D}^{-1}\boldsymbol{e}|\boldsymbol{Ce},\boldsymbol{C}^{-1}\boldsymbol{e})}{I(\boldsymbol{Ce},\boldsymbol{C}^{-1}\boldsymbol{e}|\boldsymbol{Ce},\boldsymbol{C}^{-1}\boldsymbol{e})}\cdot I(\boldsymbol{Ce},\boldsymbol{C}^{-1}\boldsymbol{e}|\boldsymbol{e},\boldsymbol{e})$$

$$\overset{\text{公理2,6}}{=\!=}I(\boldsymbol{De},\boldsymbol{D}^{-1}\boldsymbol{e}|\boldsymbol{e},\boldsymbol{e})\cdot I(\boldsymbol{Ce},\boldsymbol{C}^{-1}\boldsymbol{e}|\boldsymbol{e},\boldsymbol{e}) \tag{11}$$

$$I(\boldsymbol{p},\boldsymbol{e}|\boldsymbol{e},\boldsymbol{e})=\frac{I(\boldsymbol{p},\boldsymbol{e}|\boldsymbol{e},\boldsymbol{e})}{I(\boldsymbol{p},\boldsymbol{p}^{-1}|\boldsymbol{e},\boldsymbol{e})}\cdot I(\boldsymbol{p},\boldsymbol{p}^{-1}|\boldsymbol{e},\boldsymbol{e})$$

$$\overset{\text{公理7}}{=\!=}\frac{I(\boldsymbol{p},\boldsymbol{e}|\boldsymbol{p},\boldsymbol{e})}{I(\boldsymbol{p},\boldsymbol{p}^{-1}|\boldsymbol{p},\boldsymbol{e})}\cdot I(\boldsymbol{p},\boldsymbol{p}^{-1}|\boldsymbol{e},\boldsymbol{e})\overset{\text{公理2}}{=\!=}I(\boldsymbol{p},\boldsymbol{p}^{-1}|\boldsymbol{e},\boldsymbol{e}) \tag{12}$$

记 $\boldsymbol{\Lambda}_i=\mathrm{Diag}[1,\cdots,\lambda,\cdots,1]$ (第 i 对角元为 λ, $\lambda>0$, 其余为1), 于是 $\prod\limits_{i=1}^{n}\boldsymbol{\Lambda}_i=\lambda\boldsymbol{E}$ ($\boldsymbol{E}$ 为单位阵).

设 $I(\boldsymbol{p},\boldsymbol{q}|\boldsymbol{p}^0,\boldsymbol{q}^0)$ 是满足公理2,3,6,7的价格指数, 令

$$s=\prod_{i=1}^{n} I(\boldsymbol{\Lambda}_i\boldsymbol{e},\boldsymbol{e}|\boldsymbol{e},\boldsymbol{e}) \tag{13}$$

则

---

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 $p_i\to 0$ 表示第 i 种商品被淘汰.

$$
\begin{aligned}
s &\overset{(12)式}{=\!=\!=} \prod_{i=1}^{n} I(\boldsymbol{\Lambda}_i \boldsymbol{e}, \boldsymbol{\Lambda}_i^{-1}\boldsymbol{e} \mid \boldsymbol{e}, \boldsymbol{e}) \\
&\overset{(11)式}{=\!=\!=} I\Big(\prod_{i=1}^{n} \boldsymbol{\Lambda}_i \boldsymbol{e},\ \prod_{i=1}^{n} \boldsymbol{\Lambda}_i^{-1}\boldsymbol{e} \mid \boldsymbol{e}, \boldsymbol{e}\Big) \\
&= I(\lambda \boldsymbol{e}, \lambda^{-1}\boldsymbol{e} \mid \boldsymbol{e}, \boldsymbol{e}) \\
&\overset{公理3}{=\!=\!=} \lambda I(\boldsymbol{e}, \lambda^{-1}\boldsymbol{e} \mid \boldsymbol{e}, \boldsymbol{e}) \overset{公理2}{=\!=\!=} \lambda
\end{aligned} \tag{14}
$$

当 $\lambda \to 0$ 时 $s \to 0$，于是在(13)式右端必存在某个因子有

$$
\lim_{\lambda \to 0} I(\boldsymbol{\Lambda}_i \boldsymbol{e}, \boldsymbol{e} \mid \boldsymbol{e}, \boldsymbol{e}) = 0 \tag{15}
$$

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：

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(复习题 2)，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(复习题 1)，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 $\boldsymbol{p}^0, \overline{\boldsymbol{p}}^0$，公理 7 成立的条件是存在正数 k，使

$$
\overline{\boldsymbol{p}}^0 = k\boldsymbol{p}^0 \tag{16}
$$

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.

复习题

1. 验证价格指数 I_2 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.
2. 验证价格指数 I_8 满足公理 1~5，但不满足公理 6~8.

7.8 钢管的订购和运输

要铺设一条 $A_1 \to A_2 \to \cdots \to A_{15}$ 的输送天然气的主管道，如图 1 所示. 经筛选后可以生产这种主管道钢管的钢厂有 $S_1, S_2, \cdots, S_7$. 图中粗线表示铁路，单细线表示公路，双细线表示要铺设的管道(假设沿管道或者原来有公路，或者建有施工公路)，圆圈表示火车站，每段铁路、公路和管道旁的阿拉伯数字表示里程(单位：km). 为方便计，1 km主管道钢管称为 1 单位钢管.

一个钢厂如果承担制造这种钢管，至少需要生产500个单位.钢厂S_i在指定期限内能生产该钢管的最大数量为s_i个单位，钢管出厂销价1单位钢管为p_i万元，如表1；1单位钢管的铁路运价如表2.

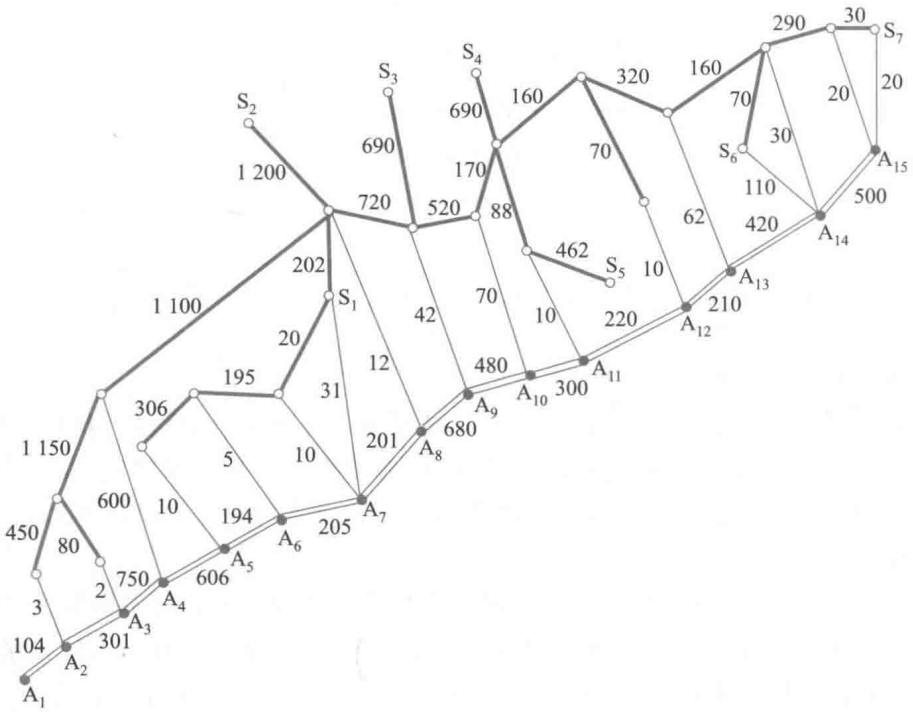


图1

表1 钢管生产的数量与价格

i	1	2	3	4	5	6	7
s_i	800	800	1000	2000	2000	2000	3000
p_i	160	155	155	160	155	150	160

表2 单位钢管的铁路运价①

里程/km	≤300	301~350	351~400	401~450	451~500
运价/万元	20	23	26	29	32
里程/km	501~600	601~700	701~800	801~900	901~1000
运价/万元	37	44	50	55	60

公路运输费用为1单位钢管0.1万元/km(不足整千米部分按整千米计算).

钢管可由铁路、公路运往铺设地点(不只是运到点$A_1,A_2,\cdots,A_{15}$，而是管道全线).

这是2000年全国大学生数学建模竞赛B题的已知条件.该题第一问要求制定一个

① 1000 km以上每增加1 km至100 km运价增加5万元.

主管道钢管的订购和运输计划,使总费用最小.为此需要首先计算从任一钢管厂订购钢管并运输到任一铺设地点的最小费用.如果能够根据图上所给的距离(一种赋权),按照题意转换成相应的费用(另一种赋权),那么这个问题相当于最短路问题.在此基础上,就可以建立钢管的订购和运输计划的数学规划模型.[79] 该题第二问和第三问留作复习题 2.

运费矩阵的计算 考虑从每个钢管厂 $S_1, S_2, \cdots, S_7$(记为 $i=1,2,\cdots,7$)到每个节点 $A_1, A_2, \cdots, A_{15}$(记为 $j=1,2,\cdots,15$)每单位钢管的最小运费 c_{ij}(不妨称为运费矩阵)及其对应的运输方式和路线.

因为题目没有给出装卸成本,可以简单地假设总是采用最经济的运输方式,在运输过程中允许多次装卸(如铁路转公路,再转铁路,等等)而不计费用.自然的想法是运输路线应该是距离最短的路径,但由于有两种运输和计价方式(铁路和公路),公路运输费用为 1 单位钢管 0.1 万元/km(不足整千米部分按整千米计算),运费是路程的线性函数;而铁路运费要通过运输里程查表得到,是一个阶梯函数.这两种运输计价方式混合在一起,使得我们不能直接将整个铁路、公路混合的运输网络上计算得到的距离最短的路径作为运输路线.但可以分别在铁路、公路网上计算距离最短的路径,然后换算成相应的运输费用;最后以两个子网上相应的运输费用为基础在整个网络上赋权,再求一次最短路问题,就可以得到最小运费矩阵.当年竞赛有不少论文对此处理不当,没能正确求出最小运费 c_{ij}.

1) 铁路子网络和公路子网络 假设铁路运输路线应该是距离最短的路径,而且采用连续路径计价方式一定优于分段计价方式(其实题中数据并不符合这一假定,例如题中 650 km 的运价为 44 万元,而分成 300 km 和 350 km 两段计价只需要 43 万元,这是不符合实际的,估计是命题时选择数据的疏忽,我们不再考虑这种情况).这时可以把铁路运输子网络独立出来,在这个子网络上计算任意两个节点 i,j 之间的最短路长度(距离)$d_{ij}^{(1)}$,然后按照这个最短路长度查铁路运价表得到最小运费 $c_{ij}^{(1)}$.

类似地假设公路运输路线应该是距离最短的路径,把公路运输子网络独立出来,在这个子网络上计算任意两个节点 i,j 之间的最短路长度 $d_{ij}^{(2)}$,然后乘以运费 0.1 万元/km 就得到最小运费 $c_{ij}^{(2)}$.

2) 最小费用矩阵的计算 经过上述的预处理,就可以将上面两个子网络(需要分别看成两个完全子图,即任意节点间都有弧相连)组合成一个网络,以每条弧上相应的运费 $c_{ij}^{(1)}$ 或 $c_{ij}^{(2)}$ 为权.需要注意的是,如果从某节点 i 到节点 j 既有铁路可到达,也有公路可到达,则弧 (i,j) 上既有铁路运费 $c_{ij}^{(1)}$,又有公路运费 $c_{ij}^{(2)}$,此时只需取其中较小的一个为权.然后再依此计算从每个钢管厂 $i=1,2,\cdots,7$ 到每个节点 $j=1,2,\cdots,15$ 的最短路,得到的就是每单位钢管的最小运费 c_{ij}.

具体算法的选择 求解最短路的算法主要有 Dijkstra 算法、Bellman-Ford 算法、Floyd-Warshall 算法等.本题弧上的权(距离)全为正数,可以采用 Dijkstra 算法.但由于需要求出任意两点之间最短路,所以采用 Floyd-Warshall 算法计算最短距离 d_{ij}^{1},d_{ij}^{2} 和最小运费 c_{ij} 可能更方便.利用软件或自己编程求解,最后可以得到最小费用 c_{ij} 为(同时可以得到相应的运输路径,从略):

170.7	160.3	140.2	98.6	38.0	20.5	3.1	21.2	64.2	92.0	96.0	106.0	121.2	128.0	142.0
215.7	205.3	190.2	171.6	111.0	95.5	86.0	71.2	114.2	142.0	146.0	156.0	171.2	178.0	192.0
230.7	220.3	200.2	181.6	121.0	105.5	96.0	86.2	48.2	82.0	86.0	96.0	111.2	118.0	132.0
260.7	250.3	235.2	216.6	156.0	140.5	131.0	116.2	84.2	62.0	51.0	61.0	76.2	83.0	97.0
255.7	245.3	225.2	206.6	146.0	130.5	121.0	111.2	79.2	57.0	33.0	51.0	71.2	73.0	87.0
265.7	255.3	235.2	216.6	156.0	140.5	131.0	121.2	84.2	62.0	51.0	45.0	26.2	11.0	28.0
275.7	265.3	245.2	226.6	166.0	150.5	141.0	131.2	99.2	76.0	66.0	56.0	38.2	26.0	2.0

订购和运输计划的制定 题目图中已经给出了从节点 j 到节点 $j+1$ 需要铺设的钢管总长度 $l_j, j=1,2,\cdots,14$. 记钢管厂 i 订购并运输到节点 j 的钢管数量为 x_{ij}，从节点 j 向左（节点 $j-1$ 方向）铺设的钢管数量为 y_j，向右（节点 $j+1$ 方向）铺设的钢管数量为 z_j，引入 0-1 变量 f_i 表示是否从第 i 个钢厂订购钢管，则可以建立如下数学规划模型：

$$\min \quad \sum_{i=1}^{7}\sum_{j=1}^{15}(p_i + c_{ij})x_{ij} + \frac{0.1}{2}\sum_{j=1}^{15}(y_j(y_j+1) + z_j(z_j+1)) \tag{1}$$

$$\text{s.t.} \quad 500f_i \leqslant \sum_{j=1}^{15} x_{ij} \leqslant s_i f_i, \qquad i = 1,2,\cdots,7 \tag{2}$$

$$\sum_{i=1}^{7} x_{ij} = z_j + y_j, \qquad j = 1,2,\cdots,15 \tag{3}$$

$$z_j + y_{j+1} = l_j, \qquad j = 1,2,\cdots,14 \tag{4}$$

$$x_{ij} \geqslant 0, z_j \geqslant 0, y_j \geqslant 0, \qquad i = 1,2,\cdots,7, j = 1,2,\cdots,15 \tag{5}$$

$$y_1 = 0, z_{15} = 0, \tag{6}$$

$$f_i \in \{0,1\}, \qquad i = 1,2,\cdots,7 \tag{7}$$

其中目标函数(1)中的第一项表示从钢管厂 i 订购并运输到节点 j 的总购运费用，第二项表示从节点 j 向左右两个方向铺设的总费用. 由于假设沿管道或者原来有公路，或者建有施工公路，所以第二项中的费用是按公路运费计算的(0.1 万元/km). 例如，从节点 j 向左铺设第 1 km 的运费为 0.1，铺设第 2 km 的运费为 0.2，……，铺设第 y_j km 的运费为 $0.1y_j$，因此向左铺设的总费用为 $0.1+0.2+\cdots+0.1y_j = 0.1y_j(y_j+1)/2$. 如果考虑到不足整千米部分按整千米计算，严格来说这里的 y_j 还需要向上取整数，但实际上可以忽略.

约束(2)和(7)描述了“一个钢厂如果承担制造这种钢管，至少需要生产 500 个单位”的要求. 约束(3)~(6)的含义很容易理解，这里就不再多解释了.

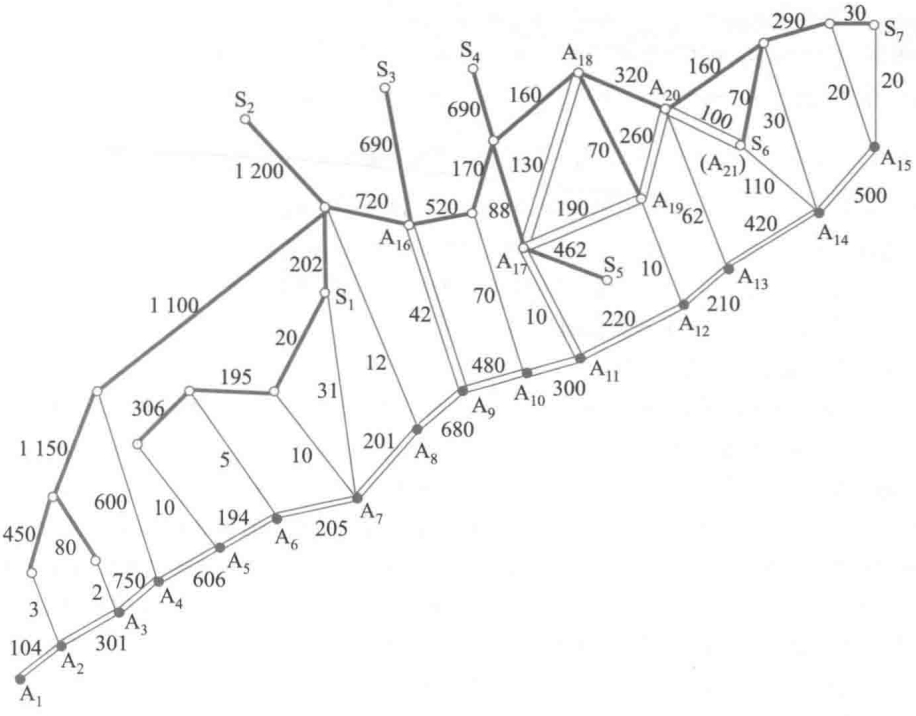
求解上述模型，可以得到问题的最小费用约为 128 亿元（最优方案从略）.

评注 当年不少参赛论文在以下几个方面存在不足：(1) 只是针对题目所给的数据用凑的方法算出结果，没有给出解决这类问题的一般模型；(2) 对铁路和公路混合运输网络处理不当，没能正确求出运费矩阵；(3) 对“一个钢厂如果承担制造这种钢管，至少需要生产 500 个单位”这个要求处理不当，得到的只是局部最优.

复习题

1. 哪个钢厂钢管的销价的变化对购运计划和总费用影响最大？哪个钢厂钢管的产量的上限的变化对购运计划和总费用的影响最大？

2. 如果要铺设的管道不是一条线，而是一个树形图，铁路、公路和管道构成网络如下图所示，请给出相应的模型和计算结果.



第 7 章训练题

1. 奖学金申请.6 位大学毕业生申请某校的研究生奖学金,该校决定考察申请者的 5 项指标:X_1~GRE,X_2~GPA,X_3~毕业学校等级分,X_4~推荐书等级分,X_5~面试等级分.5 项指标的权重及申请者的考查结果如下表,其中 X_1 满分 800,X_2 满分 4.0,其余为 10 点尺度(10 点最优).用加权和法、加权积法、TOPSIS 方法计算,如给 3 位学生奖学金,应给哪 3 位?[67]

指标权重		X_1 $w_1=0.3$	X_2 $w_2=0.2$	X_3 $w_3=0.2$	X_4 $w_4=0.15$	X_5 $w_5=0.15$
申请者	A	690	3.1	9	7	4
	B	590	3.9	7	6	10
	C	600	3.6	8	8	7
	D	620	3.8	7	10	6
	E	700	2.8	10	4	6
	F	650	4.0	6	9	8

2. 利用多属性决策和层次分析法从以下问题中选一两个完成建模过程(简单地给出模型结构、收集数据手段、计算方法和步骤等,或者详细地利用数据、得出结果):

(1) 学校评选优秀学生或优秀班级.

(2) 大学毕业生选择工作岗位.

(3) 民众评选优秀电影、电视剧或优秀运动员.

(4) 民众评选宜居城市.

3. 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.

4. 用另一种方法构造成对比较阵 $\boldsymbol{A}=(a_{ij})$,a_{ij}表示因素 C_i 与 C_j 的影响之差,$a_{ji}=-a_{ij}$,于是 $\boldsymbol{A}$ 为反对称阵,并且,当 $a_{ik}+a_{kj}=a_{ij}(i,k,j=1,2,\cdots,n)$时 $\boldsymbol{A}$ 是一致阵.规定权向量 $\boldsymbol{w}=(w_1,\cdots,w_n)^{\mathrm{T}}$ 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 $\boldsymbol{A}$ 确定权向量 $\boldsymbol{w}$ 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,…, 8 及-1,…, -8.

5. 一家向全国播送节目的地方电视台得到 150 万元投资,准备将新节目推向市场,有 3 种方案:

(1) 在当地试播新节目并作市场调查,根据调查结果决定,将新节目推向全国还是取消新节目.

(2) 不进行试播,直接将新节目推向全国.

(3) 不进行试播,也不推向全国.

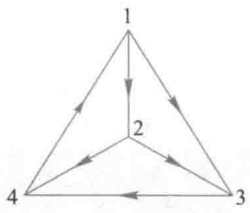
在直接推向全国的情况下,电视台估计有 55%的可能成功,45%的可能失败,如果成功,将得到 300 万元的收益;如果失败,损失 100 万元.

在当地试播进行市场调查的费用为 30 万元,有 60%的概率预测乐观结果(新节目在当地成功),40%的概率预测悲观结果(在当地失败).如果预测当地成功,有 85%的概率推向全国成功;如果预测当地失败,有 10%的概率推向全国成功.利用期望值准则和决策树方法为电视台确定最优方案[24].

6. 一台机器生产零件每批 1 000 个,次品率的概率分布如下表.若一个次品未被检出,损失费 5 元.如添置一套检验装置,每批零件检验费 350 元,可将次品自动检出.从费用期望值最小角度出发,是否应该添置检验装置?如果对零件作一次抽样检查,10 个零件中发现 1 个次品,通过修正先验概率,重新考虑添置检验装置的决策(参考拓展阅读 7-2).

次品率	0.01	0.05	0.10	0.15
概率	0.2	0.4	0.3	0.1

7. 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) 用得分向量和失分向量方法分别对竞赛图(右图)排列名次,两种方法结果一致吗?

8. 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 $n_i < q_i = (s+1)\dfrac{p_i}{P}$,即不违反份额性的上限,记 $E(n,s+1)=\{$第 i 方分配第 $s+1$ 席位合格,$i=1,2,\cdots,m\}$,当总席位为 s 时第 i 方的席位分配记作 $n_i=f_i(\boldsymbol{p},s)$,且有 $f_i(\boldsymbol{p},0)=0$,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 $\dfrac{p_k}{n_k+1}\geqslant\dfrac{p_i}{n_i+1}$,则令 $f_k(\boldsymbol{p},s+1)=n_k+1$,$f_i(\boldsymbol{p},s+1)=n_i(i\neq k)$.

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?

9. 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

10. 构造一个有 5 位候选人的投票,使得按照 5 种投票方法确定的获胜者是这 5 位候选人中的同一个人;再构造一个投票,使得 5 位获胜者正好分别是这 5 位候选人.

11. 下面的选举方法称为成对比较法:每一对候选人进行两两对决,按照多数票规则决定获胜者,并赋予 1 分(若二人平手各得 0.5 分),全部对决结束后,总分最高的为获胜者.

(1) 考察 7.6 节选举喜爱球队的投票结果,用成对比较法确定获胜者.

(2) 说明成对比较法满足多数票准则、获胜者准则、失败者准则.

12. 赞成投票被很多人认为是最简单、实用的选举方法:每位选民可以对任意多的候选人投赞成票,获得最多赞成票的候选人获胜.在现实生活中找一个按照赞成投票方法确定获胜者的例子,写出简述.研究这种方法满足还是违反 5 条公平性准则.

13. 在定义价格指数时如何构造权重 $\boldsymbol{q}$ 和 $\boldsymbol{q}^0$ 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.

14. 选择工作岗位主要考虑发展前景和当前报酬 2 个指标,某人有 3 个岗位 A_1, A_2, A_3 可选,权重 w_j 及原始分 d_{ij} 如下表.

	发展前景 $X_1(w_1=0.6)$	当前报酬 $X_2(w_2=0.4)$
A_1	4	1
A_2	1	5
A_3	2	2

(1) 分别用理想模式(最大化)和分配模式(归一化)计算 3 个岗位的综合权重,给出优劣排序.

(2) 增加一个新的岗位,其发展前景和当前报酬都与 A_1 相同,再用理想模式和分配模式计算,原岗位 A_1, A_2, A_3 的排序是保持不变,还是发生逆转?

(3) 如果增加的新岗位是 A'_4,其发展前景比其他岗位都好,如下表,用理想模式计算,原岗位的排序也会逆转吗? 如果允许新方案的权重大于 1 呢(原方案的最大权重仍为 1)?

	发展前景 $X_1(w_1=0.6)$	当前报酬 $X_2(w_2=0.4)$
A_1	4	1
A_2	1	5
A_3	2	2
A'_4	6	1

从以上的讨论,你能得出什么结论?

更多案例……

7-1 层次分析法的广泛应用

7-2 社会经济系统中的冲量过程模型

第8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.

8.1　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，工作台上方一条传送带在运转，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.

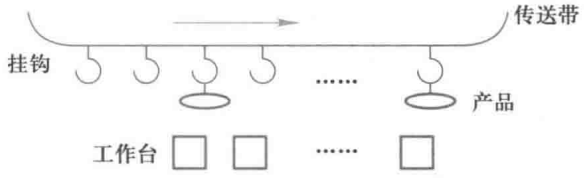


图1　传送系统示意图

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.

由上面的分析，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方，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，如果他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）：

任一只钩子被一名工人触到的概率是 $1/m$；

任一只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，任一只钩子不被所有 n 个工人挂上产品的概率，即任一只钩子为空钩的概率是 $\left(1-\frac{1}{m}\right)^n$；

任一只钩子非空的概率是 $p=1-\left(1-\frac{1}{m}\right)^n$.

这样，传送系统效率指标为

$$D=\frac{mp}{n}=\frac{m}{n}\left[1-\left(1-\frac{1}{m}\right)^n\right] \tag{1}$$

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

$$D\approx\frac{m}{n}\left[1-\left(1-\frac{n}{m}+\frac{n(n-1)}{2m^2}\right)\right]=1-\frac{n-1}{2m} \tag{2}$$

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 $n>>1$，则

$$D=1-E,\quad E\approx\frac{n}{2m} \tag{3}$$

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，若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.

复习题

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.

8.2 报童的诀窍

报童每天清晨用每份 2 元的价格从报社买进一批报纸后,在报亭以每份 4 元零售,到晚上将没有卖掉的报纸退回,得到每份 1 元的补偿.这样,他每售出一份报纸可赚 2 元,每退回一份要赔 1 元.报童每天应该买进多少报纸,当然取决于他每天能卖出多少报纸.显然,卖出的数量是随机的,因而报童每天的收入也是随机的.如果已经由统计资料或经验判断得到了报纸需求量的概率分布.那么,所谓报童的诀窍是指,他每天应该买进多少份报纸,从长期来看可以获得最高的日均利润.[32]

类似的实际问题很多.面包店每天早上烘烤一定数量的面包出售,每卖出一只可获利若干,晚间要将未卖出的面包处理掉而赔钱.如果知道了需求量的概率分布,确定每天烘烤面包的数量,便可得到最高的日均利润.出版社每年都要重印一次教科书,按照过去的销售记录,可以给出今年需求量的概率分布.在确定这次印刷数量时需要考虑的是,如果供过于求当然会因占用资金及库存而蒙受损失,而若供不应求,那么为保证学生用书必须临时加印,则会导致成本的增加.

这些问题的共同点是,已知某种商品在供过于求和供不应求时所带来的收益或损失,在需求量概率分布一定的条件下,确定商品的数量使得平均利润最高.下面以报童售报为例建立这类问题的数学模型.

离散型需求下的报童售报模型

问题 为简单起见,将每天报纸需求量看作离散型随机变量,100 份报纸为 1 单位,由统计资料得到的需求量概率分布如表 1 所示.对于报童每天以一份报纸 2 元买进、4 元售出、退回得到 1 元补偿的情况,可知每售出 1 单位报纸可获利 200 元,因剩余而退回 1 单位的报纸则损失 100 元.问报童每天应该购进多少单位报纸,能够获得最高的日均利润?

表 1 报童售报离散型需求的概率分布

需求量/单位	0	1	2	3	4	5
概率/分布	0	0.1	0.3	0.4	0.1	0.1

建模 记需求量取值 r 时的概率为 $f(r)$($r=0,1,2,\cdots,n$,对上述问题 $n=5$,$f(r)$ 如表 1).已知报童每天售出 1 单位报纸获利为 s_1,因剩余而退回 1 单位报纸损失为 s_2.设报童每天购进 q 单位报纸,若某天的需求量为 r,当供不应求即 $r>q$ 时,其售出的获利

为 s_1q,没有损失;而当供过于求即 $r\leqslant q$ 时,其售出的获利为 s_1r,因剩余而退回的损失为 $s_2(q-r)$.记一天的利润为 $s(r,q)$,则由上分析可得

$$s(r,q)=\begin{cases}s_1r-s_2(q-r), & r\leqslant q\\ s_1q, & r>q\end{cases} \tag{1}$$

日均利润为 $s(r,q)$ 的期望值,记作 $E(q)$,即有

$$E(q)=\sum_{r=0}^{n}s(r,q)f(r)=\sum_{r=0}^{q}[s_1r-s_2(q-r)]f(r)+\sum_{r=q+1}^{n}s_1qf(r) \tag{2}$$

问题归结为在已知 s_1,s_2 和 $f(r)$ 的条件下,求购进报纸的最优数量 q,使日均利润 $E(q)$ 最大.

求解 用 $P(r>q)$ 和 $P(r\leqslant q)$ 分别表示需求量取值 $r>q$ 和 $r\leqslant q$ 时的概率,让 q 从 1, 2, …依次增加,分析从 $E(q)$ 到 $E(q+1)$ 的变化.

若 $q<r$,q 增加到 $q+1$,可以多售出 1 单位,利润增加 s_1;若 $q\geqslant r$,q 增加到 $q+1$,因剩余而多退回 1 单位,利润减少 s_2.于是

$$E(q+1)-E(q)=s_1P(r>q)-s_2P(r\leqslant q) \tag{3}$$

注意到 $P(r>q)=1-P(r\leqslant q)$,(3)式化为

$$E(q+1)-E(q)=s_1[1-P(r\leqslant q)]-s_2P(r\leqslant q)=s_1-(s_1+s_2)P(r\leqslant q) \tag{4}$$

当 $E(q+1)-E(q)$ 由正变负时 $E(q)$ 达到最大,从(4)式可知,这相当于 $P(r\leqslant q)$ 由小于 $s_1/(s_1+s_2)$ 变为大于 $s_1/(s_1+s_2)$.由此得到结论:不等式

$$P(r\leqslant q)\geqslant\frac{s_1}{s_1+s_2} \tag{5}$$

成立时的最小 q 值,就是 $E(q)$ 达到最大的最优点.

对于问题中的 $s_1=200$ 元、$s_2=100$ 元,可得 $s_1/(s_1+s_2)=2/3$,根据表 1 的 $f(r)$,使 $P(r\leqslant q)\geqslant 2/3$ 成立的最小 q 值是 $q=3$,即每天购进 3 单位报纸可使日均利润 $E(q)$ 达到最大.至于 $E(q)$ 的最大值仍需由(2)式计算,其结果为 $E(q)=450$ 元.

评注 这种将决策变量 q 增加 1 单位,考察目标函数 $E(q)$ 增量的分析方法,在经济学中经常用到,称为边际分析.

分析 对于离散型需求下的报童售报(或其他类似问题的)模型(2),当 s_1,s_2 和 $f(r)$ $(r=0,1,2,\cdots,n)$ 已知时,可由(5)式直接得到使日均利润 $E(q)$ 达到最大的 q.容易看出,若获利 s_1 变大则 q 增加,若损失 s_2 变大则 q 减少.显然这是符合常识的.

连续型需求下的报童售报模型

问题 如果报纸的需求以 1 份为 1 单位,当份数很多时离散型随机变量概率分布的表述和计算就很烦琐,而将其视为连续型随机变量,用概率密度函数描述概率分布,就更为方便.

设由统计资料和经验判断认为连续型的报纸需求量大致服从正态分布,设其平均值是 260,标准差是 50,即需求量为 $N(260,50^2)$.报童仍然每天以一份报纸 2 元买进、4 元售出、退回得到 1 元补偿,即每售出 1 份报纸获利 2 元,因剩余而退回 1 份报纸损失 1 元.问报童每天应该购进多少份报纸,才能获得最高的日均利润?

建模 记报纸需求量的概率密度函数为 $p(r)$,s_1,s_2 和 q 的含义同前,用 $p(r)\mathrm{d}r$ 代替(2)式中的 $f(r)$,对 r 的求和变为积分,于是日均利润为①

① 理论上将(6)式的积分限写作 $-\infty$ 和 ∞,但对于实际问题,远离均值的概率密度可以忽略.

$$E(q)=\int_{-\infty}^{q}[s_1 r-s_2(q-r)]p(r)\mathrm{d}r+\int_{q}^{\infty}s_1 q p(r)\mathrm{d}r \tag{6}$$

问题是在已知 s_1,s_2 和 $p(r)$ 的条件下,求购进报纸的最优数量 q,使 $E(q)$ 达到最大.

求解　建立连续模型的好处是,能够用微积分方法求解优化问题.(6)式对 q 求导数①并令 $\frac{\mathrm{d}E}{\mathrm{d}q}=0$,再记 r 的概率分布函数为 $F(q)=\int_{-\infty}^{q}p(r)\mathrm{d}r$,可得 q 的最优解满足

$$F(q)=\frac{s_1}{s_1+s_2} \tag{7}$$

不难看出,这个结果与离散型需求下的模型(5)式是一致的.

当 $s_1=2$,$s_2=1$ 时(7)式右端为 2/3,对于需求概率分布 $N(260,50^2)$,为求解 $F(q)=2/3$,可以利用 MATLAB 软件的逆正态分布函数命令 x = norminv(p,mu, sigma),其中输入 mu, sigma 分别是正态分布的均值和标准差,p 是分布函数值,输出 x 是 p 的分位数(分布函数自变量的值).将以上数据代入计算得到 $x=281.5409$.取 $q=282$,报童每天购进 282 份报纸能获得最高的日均利润.

如果要得到最高日均利润 $E(q)$ 的数值,应将 $q=282$ 代入(6)式计算.

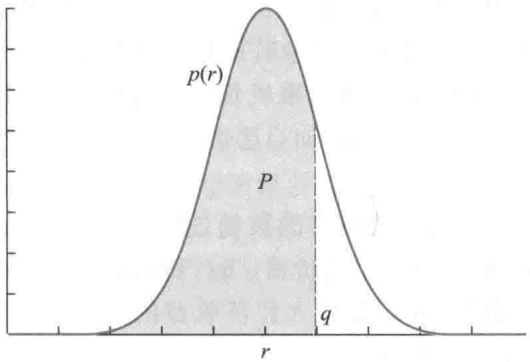


图 1　连续型需求模型求解的图形表示

分析　对于连续型模型的解(7)式,可以用概率密度函数 $p(r)$ 的图形给以直观的表示.图 1 中曲线是正态分布概率密度函数(实际上可以是任意的概率密度函数),曲线下的总面积为 1,图中虚线 $r=q$ 左边曲线下的面积记作 P,显然 $P=\int_{-\infty}^{q}p(r)\mathrm{d}r=F(q)$ 恰是需求量不超过 q 的概率.(7)式表明,$E(q)$ 达到最大的 q 值应使需求量概率密度曲线下的面积满足 $P=\frac{s_1}{s_1+s_2}$.可以看出,若获利 s_1 变大则虚线 $r=q$ 右移,若损失 s_2 变大则虚线 $r=q$ 左移.

评注　对于报童及类似的实际问题,虽然用连续型需求模型表述和计算更为便捷,但是需要知道需求量的概率密度函数,当需求量不大或者不掌握密度函数时,还是要采用离散型模型.

复习题

1. 用微积分求极值的方法由(6)式导出(7)式,并验证(7)式确实为最大值.
2. 若每天报纸的需求量服从正态分布 $N(\mu,\sigma^2)$,证明最高日均利润 $E(q)$ 为

① 注意(6)式右端的被积函数和上下限均含变量 q.

$$E(q)=s_1q-(s_1+s_2)[(q-\mu)F(q)+\sigma^2p(q)]$$

将本节的数据代入，计算 $E(q)$ 的数值.

3. 面包店每天烘烤一定数量的面包出售，每只成本 3 元，以 8 元的价格卖出，晚间关门前将未卖完的面包无偿处理掉，若已知每天面包需求量的概率分布如下表，确定每天烘烤面包的数量，使得能够得到最高的日均收入，这个收入是多少？

需求量/只	50	100	150	200	250
概率	0.2	0.3	0.2	0.2	0.1

8.3 航空公司的超额售票策略

你遇到过预订了机票、赶到机场却因飞机已满员而无法登机的情况吗？实际上，由于不少人是提前很长时间预订机票的，总有旅客因为各种变故不能按时登机（有资料统计这个数在 2%左右），航空公司为了减少按座位定额售票导致空位运行所蒙受的经济损失，通常采用超额售票策略，每个航班超额多售几张票.

航空公司超额售票时必须考虑对因飞机满员而无法登机的正常旅客给以相当力度的补偿.我国民航行业标准《公共航空运输航班超售处置规范》要求，航空公司制定详细的超售实施办法，明确旅客的权利、优先乘机规则和补偿标准；航班超售时在使用优先乘机规则前应寻找放弃登机的自愿者，向自愿者提供免费或减价航空运输、赠送里程等作为补偿①.

由于预订了机票而不按时登机旅客的数量是随机的，航空公司需要针对不同航班分析这个数量的概率分布，再结合机票价格、飞行费用、补偿金额等因素，建立一个数学模型来确定超额售票的数量，在获取最大经济收益的同时，尽量维护公司的社会声誉（避免出现过多旅客无法登机的情况）.

模型分析　实行超额售票时公司的经济收益通常用机票收入扣除飞行费用和补偿金额后的利润来衡量，社会声誉可以用因飞机满员而无法登机的旅客限制在一定数量为标准.注意到这个问题的关键因素——订票旅客是否按时前来登机——是随机的，所以经济收益和社会声誉两个方面的指标应该在平均意义或概率意义下衡量.对于这个两目标的优化问题，我们先分析经济收益，以收益最大为目标确定超额售票的数量，再来考虑如何维护社会声誉.[29]

经济收益最大的超额售票模型

模型假设　根据以上所述超额售票的内容和分析，对于一次航班作以下假设：

1. 飞机容量为 n，超额售票的数量为 q，已订票的 $n+q$ 位旅客中不按时登机的数量为 r，r 是一个随机变量.

2. 每位订票旅客不按时登机的概率为 p，他们是否按时登机的行为相互独立，这个假设适用于单独行动的商人、游客等.

---

① 指自愿放弃本次航班而改签或退票，同时得到一定数额的补偿金及其他优惠.相信总有旅客会“以时间换取金钱”，充当自愿者.所谓优先乘机的因素包括：办理乘机手续的时间、到达登机口前是否安排了座位、支付的票价等.

3. 每张机票价格为 s_1，因飞机满员而无法登机的每位旅客得到的补偿金额为 s_2. 飞行费用与旅客数量关系很小，不予考虑.

模型建立　根据订票而不按时登机的旅客数量 r 的多少，公司的收益有两种情况：

若 $r\leqslant q$，即按时登机的旅客数量 $n+q-r\geqslant n$，则只有 n 位旅客可以登机，机票收入为 $s_1 n$，剩下 $q-r$ 位旅客因无法登机而得到的补偿金额为 $s_2(q-r)$，于是公司收益为 $s_1 n-s_2(q-r)$；

若 $r>q$，即按时登机的旅客数量 $n+q-r<n$，则他们均可登机，于是公司收益为 $s_1(n+q-r)$.

对于订票而不按时登机的旅客数量为 r 的一次航班，记公司的收益为 $s(r,q)$，按照上面的分析有

$$s(r,q)=\begin{cases}s_1 n-s_2(q-r), & r\leqslant q\\ s_1(n+q-r), & r>q\end{cases} \tag{1}$$

已订票的 $n+q$ 位旅客中恰有 r 位不按时登机的概率记作 $f(r)$，$r=0,1,\cdots,n+q$，则航班的平均收益为 $s(r,q)$ 的期望值，记作 $E(q)$，有

$$E(q)=\sum_{r=0}^{n+q}s(r,q)f(r)=\sum_{r=0}^{q}[s_1 n-s_2(q-r)]f(r)+\sum_{r=q+1}^{n+q}s_1(n+q-r)f(r) \tag{2}$$

因为每位订票旅客不来登机的概率为 p，且他们是否前来登机相互独立，所以随机变量 r 服从二项分布，即

$$f(r)=C_{n+q}^{r}p^{r}(1-p)^{n+q-r},\quad r=0,1,\cdots,n+q \tag{3}$$

问题归结为：已知 s_1,s_2,n,p，求超额售票的数量 q，使航班的平均收益 $E(q)$ 最大.

模型求解　注意到平均收益 $E(q)$ 的表达式(2)与 8.2 节离散型需求下报童售报模型(2)式相似，利用同样的方法可以得到与 8.2 节(5)式完全相同的结果，即最优的 q 应为满足不等式

$$P(r\leqslant q)=\sum_{r=0}^{q}C_{n+q}^{r}p^{r}(1-p)^{n+q-r}\geqslant\frac{s_1}{s_1+s_2} \tag{4}$$

的最小 q 值.

若将补偿金额 s_2 与机票价格 s_1 之比记为 $s=s_2/s_1$，则(4)式可表示为

$$P(r\leqslant q)=\sum_{r=0}^{q}C_{n+q}^{r}p^{r}(1-p)^{n+q-r}\geqslant\frac{1}{1+s} \tag{5}$$

如果设定飞机容量 $n=300$，每位订票旅客不来登机的概率 $p=0.05$，补偿金额与机票价格之比 $s=1/2$，(5)式可用 MATLAB 软件逆二项分布函数命令 x = binoinv (y, n, p)求解，其中输入 n 是二项分布的试验总数，对应于模型中的 $n+q$，p 是概率，与模型的 p 相同，y 是分布函数值，对应于模型中的 $1/(1+s)$，输出 x 是分布函数的自变量，对应于模型中的 q. 代入 s,n,p 的数据，从 $q=1$ 开始运行程序 x=binoinv(2/3,300+q,0.05)，直到 x=q 为止，即可得到(5)式的解 $q=17$.

可以看出，在航班平均收益最大的目标下，对于固定的飞机容量 $n(=300)$，影响超额售票数量 q 的参数只有 p 和 s. 表 1 是 p 和 s 取若干数值时 q 的计算结果.

表 1 平均收益最大的超额售票数量 q（$n=300$）

q	$s=1/3$	$s=1/2$	$s=1$
$p=0.01$	4	4	3
$p=0.03$	11	10	9
$p=0.05$	18	17	16

评注 直观地看，当旅客不来登机的概率 p 变大时，超额售票数量 q 可以增加；补偿金额（与机票价格相比）变大时，超额售票数量 q 应该减少. 表 1 定量地反映了这个认识.

考虑社会声誉的超额售票模型

从维护社会声誉的角度，应该对因飞机满员而无法登机的旅客数量加以限制，由于订票旅客按时登机的随机性，所谓限制也只能以概率表示.

设超额售票的数量为 q，将因飞机满员而无法登机的旅客数量超过 j 人的概率记作 $P_j(q)$，j 可视为维护社会声誉的“门槛”，不妨限制 $P_j(q)$ 不超过某个可以接受的数值 α. 因为“无法登机的旅客数量超过 j 人”与“订票旅客不来登机的不超过 $q-(j+1)$ 人”等价，所以这种限制可以表示为

$$P_j(q)=\sum_{r=0}^{q-(j+1)} C_{n+q}^{r} p^{r}(1-p)^{n+q-r} \leqslant \alpha \tag{6}$$

（6）式的 $P_j(q)$ 可用 MATLAB 软件二项分布函数命令 y = binocdf (x, n, p) 计算，其中 x 是模型中的 $q-(j+1)$，y 即 $P_j(q)$，n, p 与 binoinv 中的含义相同. 实际应用时对于给定的 p 和 s，先算出航班平均收益最大的超额售票数量 q，再设定“门槛”j，然后计算 $P_j(q)$，与可以接受的数值 α 比较. 当超额售票数量 q 变大时“门槛”j 应该随之提高. 对于表 1 得到的结果我们取 j 约为 q 的 1/3，得到的 $P_j(q)$ 如表 2 所示.



拓展阅读 8-1
超额售票策略的再讨论

表 2 平均收益最大的 $P_j(q)$（$n=300$）

$P_j(q)$	$s=1/3$	$s=1/2$	$s=1$
$p=0.01$, $j=1$	0.4131	0.4131	0.1932
$p=0.03$, $j=3$	0.2828	0.1767	0.0968
$p=0.05$, $j=5$	0.1931	0.1282	0.0791

评注 直观地看，当旅客不来登机的概率 p 或者设定的“门槛”j 变大时，因超额售票而无法登机旅客数量超过 j 人的概率 $P_j(q)$ 都会减少；而超额售票数量 q 变大时，这个概率将增加. 表 2 大致反映了这个认识.

需要注意的是，概率 $P_j(q)$ 与费用参数 s 无关，在表 2 的同一行中，$P_j(q)$ 数值的变化是由于 q 的不同引起的.

对于航空公司的超额售票策略我们建立了经济收益最大和考虑社会声誉两个模型，应用时可以将二者结合起来. 如果对于经济收益最大的超额售票数量 q，考虑社会声誉的“门槛”j 所对应的概率 $P_j(q)$ 可以接受，问题自然解决；否则，可以适当改变 q，牺牲一定的经济收益，来换取 $P_j(q)$ 的降低.

从上面的计算结果可以看到，订票旅客不来登机的概率 p 对经济收益和社会声誉的两个指标影响都比较大，应用时需要针对不同航班、不同时间（季节、假日等因素），利用统计数据实时调整概率 p，以提高模型的准确度.

复习题

1. 由 $E(q)$ 的表达式（2）推导（4）式；将（2）中的概率 $f(r)$ 换为密度函数，对连续型模型求解.

2. 据统计数据订票旅客不来登机的概率 $p=0.02$，对于 $n=300, j=3$ 和 $s=1/3, 1/2, 1$，计算平均收益最大的超额售票数量 q 及概率 $P_j(q)$. 如果设定 $\alpha=0.1, s=1/3$，问 $P_j(q)$ 是否满足(6)式的限制？若不满足，如何调整 q，调整的后果是什么？

8.4 作弊行为的调查与估计

“考试时你是否有过作弊行为？”如果在学生中做这样的、即使是无记名的直接调查，恐怕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，使得调查结果存在很大的偏差.

社会调查中类似“考试作弊”的所谓敏感问题有不少，例如是否有过赌博、婚外恋、酒后驾车行为等. 如何通过巧妙的调查方案设计，尽可能地获取被调查者的真实回答，得到比较可靠的统计结果，是社会调查工作的难题之一. 本节以学生考试作弊行为的调查和估计为例，讨论这类敏感问题的调查方法及数学模型.

问题分析　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. 需要强调的是，这种调查并不涉及具体是哪些学生有过作弊行为，所以调查方案的设计应该保护被调查者的隐私，消除他们的顾虑，从而能对调查的问题做出真实的回答.

进行这项调查的基本想法是，让被调查的学生从包含“是否有过作弊行为”的若干个问题中，按照一定的随机规律，对其中某一个问题用一个字“是”或“否”作答. 调查者只知道，对于全部调查问题回答“是”的有多少人，回答“否”的有多少人，而根本不了解被调查者回答的是哪一个具体问题.

调查者面临的任务是，设计具体的调查方案，并建立相应的数学模型，使得能从调查结果估计出有过作弊行为学生的比例.

正反问题选答的调查方案及数学模型

所谓正反问题选答，就是将要调查的内容用一正、一反两个问题表述.

方案设计　在学生作弊行为调查中，设计以下两个正反问题供选答其中一个：

问题 A：你在考试中有过作弊行为，是吗？

问题 B：你在考试中从未有过作弊行为，是吗？

对选择的问题只需回答“是”或“否”.

选答问题的规则是让每个被调查的学生按照一定的概率选答问题 A 或 B，比如采取以下方式：调查者准备 A,2,3,…,Q,K 共 13 张不同点数的扑克牌，学生在选答问题前随机抽取一张，并约定，如果抽取的是不超过 10 点的牌(牌 A 看作 1 点)则回答问题 A，如果抽取的是 J,Q,K，则回答问题 B. 学生看完牌后还原，且使调查者不能知道抽取的牌. 这样，学生选答问题 A 和 B 的概率分别为 $\frac{10}{13}$ 和 $\frac{3}{13}$.

模型假设　按照问题分析及方案设计作以下假设：

1. 调查共收回 n 张有效答卷，其中有 m 张回答“是”，$n-m$ 张回答“否”，于是回答“是”的学生的比例 m/n 可视为学生回答“是”的概率.

2. 被调查的学生对每个选定的问题作真实回答，于是对问题 A 回答“是”的学生（在选定问题 A 的所有学生中）所占的比例可视为学生作弊的概率.同样，对问题 B 回答“否”的学生所占的比例也视为学生作弊的概率.

3. 每个被调查的学生选答问题 A 的概率是确定的，记作 $P(A)$，选答问题 B 的概率记作 $P(B)$，$P(A)+P(B)=1$.

模型建立 记事件 A 为学生选答问题 A，事件 B 为学生选答问题 B，事件 C 为学生回答“是”，事件 $\overline{C}$ 为学生回答“否”.按照全概率公式有

$$P(C)=P(C|A)P(A)+P(C|B)P(B) \tag{1}$$

且

$$P(C|B)=1-P(\overline{C}|B) \tag{2}$$

在(1)和(2)式中，由假设1，$P(C)=m/n$ 为调查结果；由假设2，条件概率 $P(C|A)$，$P(\overline{C}|B)$ 正是学生作弊的概率，记作 p；由假设3，$P(A)$，$P(B)$ 已确定，且 $P(B)=1-P(A)$.于是(1)式化为

$$P(C)=pP(A)+(1-p)(1-P(A)) \tag{3}$$

由此容易推出待估计的学生作弊的概率为

$$p=\frac{P(C)+P(A)-1}{2P(A)-1} \tag{4}$$

这个模型是1965年由美国统计学家 Warner 提出的，称为 Warner 模型.请特别注意，(4)式成立的前提是 $P(A)\neq\frac{1}{2}$，也就是说，不能让学生选答问题 A 与问题 B 的概率相等，像掷硬币出现正面和反面那样.请读者检查(1)~(3)式，对此给以解释.

具体算例 假定采用方案设计中抽取扑克牌的办法决定学生选答问题 A 或 B，在400张有效答卷中有112张回答“是”，按照(4)式估计学生作弊的概率.

将 Warner 模型 $P(C)=\frac{112}{400}$，$P(A)=\frac{10}{13}$ 代入(4)式，立即得到 $p=0.091$，即根据这个调查有9.1%的学生在考试中有过作弊行为.

模型分析 对模型(4)式先考察两种极端情况：若学生无人作弊，对问题 A 均答“否”、问题 B 均答“是”，即 $P(C|A)=0$，$P(C|B)=1$，由(1)式得 $P(C)=P(B)$，于是(4)式给出 $p=0$；若学生都作弊，对问题 A 均答“是”、问题 B 均答“否”，即 $P(C|A)=1$，$P(C|B)=0$，由(1)式 $P(C)=P(A)$，于是(4)式给出 $p=1$.

其次，当 $P(A)>\frac{1}{2}$，即选问题 A 的比选问题 B 的人数多，若 $P(C)$ 变大，即回答“是”的增加，其中对问题 A 回答“是”（作弊）占的比例大于对问题 B 回答“是”（未作弊）占的比例，所以学生作弊的概率 p 变大.当 $P(A)<\frac{1}{2}$ 时，情况正好相反.

最后，按照概率统计的观点，由模型(4)确定的 p 是学生作弊概率的估计值，通常

记作$\hat{p}$,这是一个随机变量①,可以推导出 $\hat{p}$ 的期望 $E(\hat{p})$ 和方差 $D(\hat{p})$ 如下:

$$E(\hat{p})=p \tag{5}$$

$$D(\hat{p})=\frac{P(C)(1-P(C))}{(2P(A)-1)^2 n} \tag{6}$$

(5)式右端的 p 是学生作弊概率的真值.(5)式表明,模型(4)确定的 $\hat{p}$ 是 p 的无偏估计,即如果作多次这样的调查,其估计值的平均值将趋于真值.

(6)式表明,被调查的人数 n 越多,方差越小,调查结果的精度越高.$P(A)$ 越接近 $\frac{1}{2}$,方差越大,调查结果的精度越低.当 $P(A)$ 接近 1 或 0 时,方差固然会变小,但是这意味着让绝大多数人都回答问题 A 或问题 B,对被调查者的保护程度会降低,不利于调查的正常进行.

对于上面的算例,将 $P(C)=\frac{112}{400}$,$P(A)=\frac{10}{13}$,$n=400$ 代入(6)式,可得 $\sqrt{D(\hat{p})}=0.042$,即标准差为 4.2%.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%±8.4%(以 95%的概率在此区间内).

评注 利用 Warner 模型制订方案时,选答问题 A 或 B 的概率 $P(A)$、$P(B)$ 是由调查者设计的.根据前面的分析,$P(A)$ 接近 $\frac{1}{2}$ 或者接近于 1(或 0)都不合适.一般地说,在实际采用 Warner 模型作调查时,建议 $P(A)$ 在 0.7 到 0.8 之间取值.

无关问题选答的调查方案及数学模型

Warner 模型虽然比直接调查要好,但正反两个问题涉及的是同一个内容,可能会引起被调查者的猜疑或不快,而且要求回答每个问题的人数比例不要等于或接近 1/2,也会给调查带来不利影响.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者提出的是两个不相关的问题,其中一个为要调查的敏感问题,另一个是与调查无关的非敏感问题,这样的处理能使被调查者的合作态度进一步提高.下面这个模型称为 **Simmons 模型**.

方案设计与建模 在学生作弊行为调查中,问题 A 不变,而问题 B 设计为一个与问题 A 完全无关的问题,如:

问题 A:你在考试中有过作弊行为吗?

问题 B:你生日的月份是偶数吗?

对选择的问题只需回答“是”或“否”.选答规则也是让每个被调查的学生按照一定的概率选答问题 A 或 B.

仍设调查收回 n 张有效答卷,其中 m 张回答“是”,回答“是”的概率为 $P(C)=m/n$.条件概率 $P(C|A)$ 仍是学生作弊的概率,为了区别于 Warner 模型,将其记作 p'.另一个条件概率 $P(C|B)$ 是能够事先确定的,记作 p_0,当 n 较大时,对于“生日的月份是偶数”可以认为 $p_0=\frac{1}{2}$.由全概率公式(1)得到

$$P(C)=p'P(A)+p_0(1-P(A)) \tag{7}$$

由此得到待估计的学生作弊的概率为

$$p'=\frac{P(C)+p_0(P(A)-1)}{P(A)} \tag{8}$$

---

① (4)式中 $P(C)$ 也是随机变量,$\frac{m}{n}$ 是它的估计值.

模型分析与算例 根据(8)式,若 $P(C)$ 变大,不论 $P(A)$ 的大小如何,学生作弊的概率 p' 都会增加,这点与 Warner 模型有别.

按照概率统计的观点,由模型(8)确定的 p' 也是学生作弊概率的估计值,记作 $\hat{p}'$,$\hat{p}'$ 的期望和方差如下:

$$E(\hat{p}')=p' \tag{9}$$

$$D(\hat{p}')=\frac{P(C)(1-P(C))}{(P(A))^2 n} \tag{10}$$

(9)式右端的 p' 是学生作弊概率的真值,$\hat{p}'$ 是 p' 的无偏估计.(10)式表明,被调查的人数 n 越多,方差越小,调查结果的精度越高.$P(A)$ 越接近 1,虽然方差也变小,但是让绝大多数人都回答问题 A 显然是不合适的.

给出一个具体算例:在 400 张有效答卷中有 80 张回答"是",选答问题规则采用与 Warner 模型相同的抽扑克牌的办法,按照(8)式估计学生作弊的概率.

将 $n=400$,$P(C)=\frac{80}{400}$,$p_0=\frac{1}{2}$,$P(A)=\frac{10}{13}$ 代入(8)和(10)式,得到 $\hat{p}'=0.11$,$\sqrt{D(\hat{p}')}=0.026$,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% ±5.2%.

在采用 Simmons 模型时,如何确定选答问题 A 的概率 $P(A)$,才能使估计值的精度比 Warner 模型更高呢?对(6)式与(10)式作进一步的分解,去除其中因 $P(C)$ 不同造成的影响,可以得到,当 $P(A)>1/3$ 时,$D(\hat{p}')<D(\hat{p})$.

评注 不论是 Warner 模型还是 Simmons 模型,全概率公式(1)都是建模的基础.Warner 模型中对问题 B 回答"是"的概率为 $1-p$,Simmons 模型中对问题 B 回答"是"的概率 p_0 为已知,通常设计问题 B 使 $p_0=1/2$.于是从全概率公式(1)可以直接推出学生作弊的概率,如(4)和(8)式.若 Simmons 模型的 p_0 本身也是未知的,如问题 B 是"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法.比如将被调查者分成两批,做两次调查,且选答问题 A 的概率不同,从而用全概率公式得到两个方程,得出学生作弊的概率 p(本章训练题 3).

复习题

为克服 Warner 模型中正反两个问题都是敏感问题的缺点,设计如下调查方案:调查者先制作三种卡片,每种卡片若干张.第一种卡片上写"若你有过作弊行为,请回答数字 1;若你从未有过作弊行为,请回答数字 0";第二种卡片上写"请直接回答数字 1";第三种卡片写"请直接回答数字 0".将三种卡片充分混合放入一盒,使第一、二、三种卡片的比例为 $p_1:p_2:p_3$,$p_1+p_2+p_3=1$.每个被调查学生从盒中随机抽一张卡片,做真实回答后放回.若被调查学生的总数为 n,回答"1"的人数为 m,建立估计有过作弊行为学生比例的模型.

若三种卡片的比例为 5:3:2,200 名被调查学生中有 70 人回答数字 1,估计有过作弊行为学生的比例.

8.5 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[7].

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 $x\geqslant l$ 的概率为 P,即 $P=P(x\geqslant l)$,P 是图中阴影部分面积.

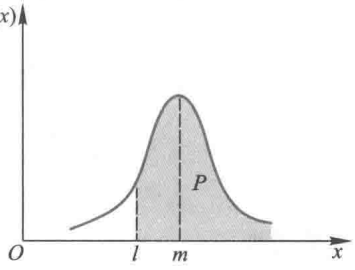


图 1 钢材长度 x 的概率密度

轧制过程中的浪费由两部分构成.一是当 $x\geqslant l$ 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

$$W=\int_{l}^{\infty}(x-l)p(x)\,\mathrm{d}x+\int_{-\infty}^{l}xp(x)\,\mathrm{d}x \tag{1}$$

利用 $\int_{-\infty}^{\infty}p(x)\,\mathrm{d}x=1$,$\int_{-\infty}^{\infty}xp(x)\,\mathrm{d}x=m$ 和 $\int_{l}^{\infty}p(x)\,\mathrm{d}x=P$,(1)式可化简为

$$W=m-lP \tag{2}$$

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

$$W=\frac{mN-lPN}{N}=m-lP \tag{3}$$

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

---

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 $l,m>>\sigma$,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.

$$J_1=\frac{mN-lPN}{PN}=\frac{m}{P}-l \tag{4}$$

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

$$J(m)=\frac{m}{P(m)} \tag{5}$$

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

$$P(m)=\int_l^{\infty} p(x)\,\mathrm{d}x,\quad p(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}} \tag{6}$$

作变量代换

$$y=\frac{x-m}{\sigma} \tag{7}$$

并令

$$\mu=\frac{m}{\sigma},\quad \lambda=\frac{l}{\sigma},\quad z=\lambda-\mu \tag{8}$$

则(5)式可表为

$$J(z)=\frac{\sigma(\lambda-z)}{1-\Phi(z)} \tag{9}$$

其中 $\Phi(z)$ 是标准正态变量的分布函数,即

$$\Phi(z)=\int_{-\infty}^{z}\varphi(y)\,\mathrm{d}y,\ \varphi(y)=\frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{y^2}{2}} \tag{10}$$

$\varphi(y)$ 是标准正态变量的密度函数.

用微分法求解函数 $J(z)$ 的极值问题.注意到 $\Phi'(z)=\varphi(z)$,并记

$$F(z)=\frac{1-\Phi(z)}{\varphi(z)} \tag{11}$$

可以得到 $J(z)$ 的最优值 z^* 应满足方程

$$F(z)=\lambda-z \tag{12}$$

用 MATLAB 软件得到方程(12)的根 z^*,再代回(8)式即得到 m 的最优值 m^*.

值得指出的是,对于给定的 $\lambda>F(0)=1.253$①,方程(12)不止一个根,但是可以证明,只有唯一负根 $z^*<0$,才使 $J(z)$ 取得极小值(复习题 2).

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

① 通常 $l>>\sigma$,故对于 $\lambda=l/\sigma$,条件 $\lambda>F(0)=1.253$ 是容易满足的.

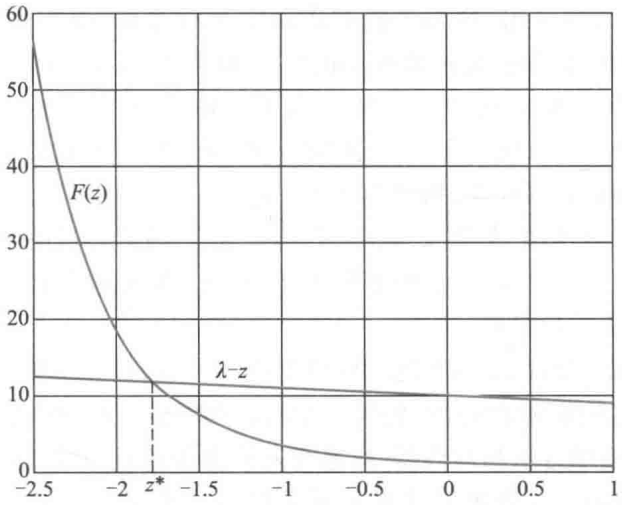


图 2 $F(z)$的图形及(12)的图解法

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)式算出 $\lambda=l/\sigma=10$,解出方程(12)的负根为 $z^*=-1.78$,如图 2 所示.由(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 $P(m^*)=0.962\,5$,按照(4)式每得到一根成品材浪费钢材的平均长度为 $J_1=\dfrac{m^*}{P(m^*)}-l=0.45$ 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.

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.

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?

复习题

1. 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).

2. 证明方程(12)仅有一个负根 z^*,并且在 z^* 取得(9)式 $J(z)$的极小值.

8.6 博彩中的数学

每到福彩、体彩开奖的电视和网站直播时刻,千万彩民手握彩票,双眼紧盯摇奖机

里滚出的一个个小球的号码，迎来的既有幸运中奖的喜悦，也有擦肩而过的遗憾.据2006年的调查结果推算，全国的经常性彩民约7000万人.夜幕下的澳门是赌场的天下，轮盘、老虎机、扑克牌等赌台前挤满了赌客.与北美的拉斯维加斯、欧洲的蒙特卡罗齐名，亚洲的澳门是世界三大赌城之一，政府财政收入的30%、税收的50%都来自赌场.香港赛马场上人头攒动、骏马狂奔，据说600多万港人中有200多万马迷，“马照跑”作为香港回归后实行“一国两制”的重要特征而广为人知.一场精彩的足球赛开赛在即，两队的赔率已经开出，竞猜者犹豫再三、押下赌注，随着赛场形势的跌宕起伏，下注者的心情比一般的球迷更多了一份牵挂.

上面描述的彩票、赌台、动物赛跑和体育竞猜，是当前世界上博彩业的4种主要经营方式.博彩起源于赌博，可以追溯到两人面对某个悬而不决的问题而抓阄，以及市井中用扔铜币看正反面的办法相对而赌，后来有人从中看到了生财之道，为赌客提供物质条件并居中抽头，成为今天赌场的雏形.而随着赌博名声越来越坏，人们为它找到了一个好听的名字——博彩.词典上给博彩下的定义是，对一个不确定的结果做出预测判断，并以押上的钱来为自己的决策承担后果.这里的钱实际上指任何抵押物.可以说，有赌注、有对手、碰运气，是博彩的三大要素[78].

在我国，目前只有一部分彩票发行和体育竞猜是合法的，在政府监督下由相关机构管理，其他如赌台、赛马等都被禁止.不过从数学的角度看，彩票与轮盘、掷骰子等赌台游戏的数学模型基本上是一样的.本节讨论两种博彩方式中的数学问题，首先以叙述简单、便于计算的轮盘为例，从玩家(赌客)和庄家(赌场)两方面，介绍赌台游戏中的建模，然后选择一种简单的球赛竞猜方式，研究开盘、下注以及收益中的模型[25].

轮盘游戏中的数学

轮盘(roulette)是一种赌场中常见的非常刺激的游戏，由一个轮盘、一个白色小球和一张赌桌构成.(美式)轮盘分成38个沟槽，并以转轴为中心转动，沟槽编号1至36，一半红色、一半黑色，还有两个绿色沟槽编号0和00①.玩家可以对单独一个数字、几个数字组合、单双数字、红或黑一种颜色下注，然后庄家让转轮沿逆时针方向转动，并使放在微凸盘面上的小球以顺时针方向旋动，以小球最后停在某个沟槽的那个数字或颜色与下注的数字或颜色是否相同，决定玩家是赢还是输.

概率和赔率 讨论轮盘游戏中玩家的两种下注方式：对1至36其中一个数字下注；对红或黑一种颜色下注.转动后小球最后停在38个沟槽中的哪一个是随机的、等可能性的，于是对一个数字下注时玩家赢钱的概率为$\frac{1}{38}$，对一种颜色下注时赢钱的概率为$\frac{18}{38}$，输钱的概率自然分别为$\frac{37}{38}$和$\frac{20}{38}$.

赌场里很多人玩同一种游戏，每个人还会玩很多次，按照概率论中的大数定律，玩家赢钱次数的频率将逐渐稳定于赢钱的概率，记作p.博彩中常用概率的另一种表示——赔率.理论上的赔率称为真实赔率(true odds)，记作y，其含义是，如果庄家按照赔率y赔给赢钱的玩家，那么很多次游戏后，庄家和玩家都不赔不赚.设玩家下注1元，

① 欧式轮盘只有37个沟槽，少一个绿色沟槽00.

如果玩家赢(概率 p),则得到庄家赔付的 y 元;如果玩家输(概率 $1-p$),则失去下注的 1 元,不赔不赚应表示为 $p\times y=(1-p)\times 1$,由此得到赔率与概率的关系

$$y=\frac{1-p}{p},\text{或 } p=\frac{1}{y+1} \tag{1}$$

按照(1)式计算,对一个数字下注(赢钱概率 $p=\frac{1}{38}$)的真实赔率 $y=37$,即 1 赔 37.对一种颜色下注(赢钱概率 $p=\frac{18}{38}$)的真实赔率 $y=1.11$,即 1 赔 1.11.这样的赔率对双方是公平的,所以又称平率.

实际上,任何一个赌场老板都不会以真实赔率 y 赔付给赢钱的赌客,那样他就一分钱也赚不到还白搭上服务.通常赌场对一种游戏会选择一个比真实赔率 y 稍低一点的赔率赔付,称为支付赔率(payoff odds),记作 x,如对一个数字下注的支付赔率取 $x=35$,即 1 赔 35,对一种颜色下注的支付赔率取 $x=1$.以下如不特别说明,赔率均指支付赔率.

玩家劣势和庄家优势 常识告诉人们,赌场上的赌客碰到好运气,可能赢上几把,但是一直赌下去,最终总是要输的.在轮盘游戏中,让我们算一下对一个数字下注并一直赌下去的玩家会输多少钱.仍设下注 1 元,如果玩家赢($p=\frac{1}{38}$),他得到庄家赔付的 35 元(赔率 $x=35$);如果玩家输,失去下注的 1 元.如此下注很多次玩家的平均收益(即期望收益)为 $35\times\frac{1}{38}+(-1)\times\frac{37}{38}=-\frac{2}{38}=-0.0526$(元),即输掉下注金额的 5.26%.用式子表示这个计算,玩家的期望收益 EV_1(expected value)为

$$EV_1=xp+[-(1-p)]=(x+1)p-1 \tag{2}$$

EV_1 完全由概率 p 和赔率 x 决定,通常以百分比的形式表示.

换一种玩法,对一种颜色下注($p=\frac{18}{38}$,$x=1$),由(2)式得到,玩家的期望收益 EV_1 与上面算出的对一个数字下注时完全一样.实际上这是庄家早就安排好的.

赌场中各种游戏玩家的期望收益 EV_1 一般都是负值,取绝对值后的数值称为玩家**劣势**(赌客劣势).

更一般地,如果玩家在一种游戏中的下注金额为 1 个单位,游戏出现第 i 种结果的收益是 $a_i(i=1,2,\cdots,n)$,出现这种结果的概率是 p_i,则期望收益

$$EV_1=\sum_{i=1}^{n}a_ip_i \tag{3}$$

买彩票的期望收益就可以用(3)式计算.

赌场上玩家输的正是庄家赢的,如果计算庄家的期望收益 EV_2,得到的结果与(2)式的 EV_1 相差一个负号,一般都是正值,称为庄家优势(赌场优势).显然,轮盘游戏中对一个数字下注与对一种颜色下注的庄家优势也都是 5.26%.

玩家的加倍下注策略 赌场上对于赔率 $x=1$(如对轮盘一种颜色下注)的游戏,有些玩家会采用一种自以为能够稳赢不输的下注策略.办法是,从最低金额(假定是 1 元)开始下注,输了就加倍下注,再输再加倍,直到赢了为止.这样只要赢一次,就不仅能把

输的钱全部捞回来，而且净赚1元.真的是这样吗？

玩家手中的钱当然是有限的，不妨假定他有1万元，首次下注1元，够他连输13次，因为 $1+2+2^2+\cdots+2^{12}=2^{13}-1=8\ 191$（元）.对一种颜色下注时赢的概率 $p=\frac{18}{38}=0.473\ 7$，将近 $\frac{1}{2}$，连输13次的概率极小，于是他想当然地认为最终会赢钱.让我们仔细算一下吧.

玩家输的概率 $q=1-p=0.526\ 3$，连输13次、输掉8 191元的概率是 $q^{13}=0.000\ 237\ 8$，而赢其中一次、净赚1元的概率是 $1-q^{13}$，于是期望收益为 $-8\ 191q^{13}+(1-q^{13})=-0.948\ 0$，所以理论上（长期按照这种策略下注）他还是会输的，虽然输得并不多.

这种策略下玩家的期望下注金额是 $1+2q+2^2q^2+\cdots+2^{12}q^{12}=\frac{(2q)^{13}-1}{2q-1}=18.008\ 7$，所以他每下注1元的期望收益为 $-\frac{0.948\ 0}{18.008\ 7}=-0.0526$，与前面按照(2)式计算的一次下注的期望收益分毫不差.这说明与普通的下注方式相比，采用加倍下注策略（在理论上）没有得到任何便宜.请读者自行推导：不论概率 p 多大、玩家手中的钱有多少，加倍下注策略的期望收益与(2)式的 EV_1（赔率 $x=1$）完全相同.

赌场游戏的设计方法和步骤 站在庄家的角度上，对于赌场中的每一种游戏，从制定玩法、确定赔率，到估计收益、规避风险，有一系列的设计方法和步骤.

1. 概率计算 赌场中各种游戏结果的不确定性正是吸引赌客的魅力所在，而隐藏在实际发生的不确定性背后的，是理论上的确定性——出现每种结果的精确概率.概率计算是整个游戏设计的基础，轮盘、掷骰子等游戏的概率计算比较简单，彩票各等级奖项的概率计算要复杂些，而像扑克牌21点游戏的概率通常需要作计算机模拟才能得到.

2. 赔率确定 理论上的真实赔率 y 直接由概率决定，如(1)式，赌场实际上采用的赔率 x（支付赔率）需根据真实赔率向下调整到一个简单的（一般取整数）、合理的（见下面的讨论）数值.

3. 庄家优势的调整 如前所述，庄家优势（即期望收益 EV_2）与玩家期望收益 EV_1 只差负号.对于简单的轮盘游戏(2)式，有 $EV_2=1-(x+1)p$.对一种游戏（玩家赢钱）的概率 p 是一定的，赌场可以采用降低或提高赔率 x 的方法调整庄家优势.

4. 最低和最高下注额的设定 对下注额的限制称为限红.赌场通常采用筹码（如10元兑换1个最小面额的筹码）的下注方式来设定最低限红，防止个别玩家零敲碎打地下注，以节约服务成本.设最高限红（如10万元的筹码）的目的是，防止个别财大气粗的玩家过分高额地下注，给赌场带来瞬间的威胁.比如一个赌客拿1亿元（对轮盘的一种颜色）下注，赌场和赌客双方输的概率都接近 $\frac{1}{2}$，赌场没有必要冒这样的风险，因为只需设一个最高限红，并鼓励这位亿万富翁长期地赌下去，那1亿元早晚会落入赌场老板的囊中.另外，最高、最低限红的区间也限制了玩家的加倍下注策略，赌场毕竟是不欢迎这种下注方式的.

赌场的收益及其波动 赌场对经营的所有游戏都会作一系列的设计或调整，使总的收益控制在一个适当的范围内.收益太高监管部门会干预，也不利于行业内部的竞争；收益太低会影响扣除税收、成本以及慈善捐赠等支出后的利润.

博彩业的收益是随机的，前面讨论的庄家优势 EV_2 是期望值，实际收益的波动范围可用收益的标准差描述.对于一次下注 1 元的轮盘游戏，赌场收益的标准差 SD_2 可以参照(2)式写出

$$SD_2=\sqrt{x^2p+(1-p)-EV_2^2} \tag{4}$$

若对一种颜色下注($x=1,p=\dfrac{18}{38},EV_2=0.052\ 6$)，可算出 $SD_2=0.998\ 6$；若对一个数字下注($x=35,p=\dfrac{1}{38},EV_2=0.052\ 6$)，可算出 $SD_2=5.762\ 8$.一次下注时庄家收益(相对于期望值)的波动太大(当然玩家也一样).可是赌场接待成千上万个赌客，每个赌客要玩好多次，当上述游戏下注 n 次时，收益的平均值仍是 EV_2，而收益的标准差却是

$$SD_2(n)=\frac{SD_2}{\sqrt{n}} \tag{5}$$

假设 $n=10\ 000$(如 1 000 个赌客，每人玩 10 次)，则对一种颜色下注的 $SD_2(n)=0.010\ 0$，对一个数字下注的 $SD_2(n)=0.057\ 6$.虽然后者的标准差仍然不小，可是实际赌场上的 n 会更大.

进而，根据概率论的中心极限定理，当下注次数 n 很大时，其平均收益趋向正态分布.于是可以以 95%的把握预测，赌场的收益在 $EV_2\pm 2SD_2(n)$ 范围内.

玩家应该关注的 赌场中的玩家看到的只是每种游戏的赔率，买彩票也只知道各等级奖项的中奖额.如果不同赌场同一种游戏的赔率不同，玩家自然会选赔率高的.在不同游戏或不同彩票中选择时，如果不管兴趣，只从赢钱的角度出发(应该说是少赔钱)，应该估算游戏中各种结果或彩票各奖项出现的概率，然后根据赔率和概率计算玩家的期望收益 EV_1，进行比较.实际上，如果玩家只是偶尔赌上几次或买几张彩票，期望收益的大小就没有多大意义，更需要关注的是反映收益波动范围的标准差(与(4)、(5)式给出的赌场收益的标准差类似).一般地说，保守型玩家宜选择标准差小的，如在轮盘游戏中对一种颜色下注，冒险型玩家可选择标准差大的，如在轮盘游戏中对一个数字下注.

体育竞猜中的数学

体育竞猜是近些年来迅速发展起来的一项博彩，庄家(博彩公司)在一场体育比赛(各种球赛、拳击等)开始前，组织玩家预测比赛结局并下注，比赛结束后根据实际结果决定玩家的输赢.与彩票、赌台游戏各种结果出现的概率可以精确计算不同的是，体育比赛双方胜负的概率，不论玩家或庄家都只能根据两队的实力、状态等因素做出大致的估计，所以竞猜既对庄家的技术水平、经营技巧和经验积累有很高的要求，是博彩业中风险最大的分支，也给一些比较内行的玩家提供了赢钱的机会，是博彩业中最容易产生职业赌客的地方.

体育竞猜中下注金额的制定称为开盘，国际上开盘的机制有所谓加(让)线盘、加(让)分盘、赔率盘等，下面只介绍在竞猜中计算较为简单、又颇吸引人的加线盘.

开盘、押注、玩家劣势和庄家优势 在A,B两支球队比赛前庄家为竞猜开盘,即给出两队胜负的赔率,让玩家押注.赔率用两个数字表示,如-160/+140,其含义是,当玩家押A队获胜时,赔率为1.6赔1,即一注押160元,结果若A队胜,玩家获赔100元(押金退还),若A队负,押金归庄家;当玩家押B队获胜时,赔率为1赔1.4,即一注押100元,结果若B队胜,玩家获赔140元(押金退还),若B队负,押金归庄家.这就是说,-160/+140中前面带-号和后面带+号的数字分别代表了A队和B队的赔率.很明显,这个开盘说明庄家认为A队是强队(押多赔少)、B是弱队(押少赔多).这种形式的开盘称**加线盘**(后面解释加线的含义),又称**独赢盘**,即不考虑平局,或者将平局归为弱队获胜.

两队强弱对比的定量关系用双方的获胜概率(以下简称胜率)表示,假定庄家已经估计出强队A的胜率为0.6,弱队B的胜率为0.4.让我们先按这个概率算一下,竞猜中玩家和庄家的期望收益有多大.

押强队A胜的玩家的期望收益为$100\times0.6-160\times0.4=-4$,押弱队B胜的玩家的期望收益为$140\times0.4-100\times0.6=-4$,二者相等.对这两种下注方式,单位下注额的玩家劣势分别为$\frac{4}{160}=2.5\%$和$\frac{4}{100}=4\%$,押强队A胜的玩家劣势较小.

评注 我们从数学的角度出发讨论了以轮盘为例的赌台游戏和体育竞猜两种博彩形式,二者都符合本节开始给出的博彩的定义及三大要素,但是也有明显的区别.前者是专为赌博而人为设置的赌局,需要一定的成本,其各种结果的概率可用数学工具精确得到,庄家和玩家基本上都不需要什么"技术",一时的输赢完全是碰运气;后者是博彩公司借用现实生活中不可预知又吸引眼球的事件开设的赌局,成本很小,各种结果的概率主要靠"技术"水平来估计,内行的玩家在这类博彩中有一定优势.有一些引人注目的政治、社会事件如美国总统大选、英国小王子取名等也被博彩公司"免费搭车",来一场全民竞猜.

两种下注方式的玩家劣势正是对这两种下注方式的庄家优势.在押强队A胜和押弱队B胜的玩家数量相等的情况下,单位下注额的整体庄家优势为$\frac{4+4}{160+100}=3.08\%$.

应该注意到,与彩票、赌台在大量、重复的游戏中各种结果出现的频率逐渐稳定于概率不同,体育竞猜中两队在相同条件下的比赛很少重复,庄家估计的胜率实际上并不具有大数定律所赋予的性质,它只是在开盘和计算理论的期望收益中用到,同时,玩家也会估计两队的胜率,并结合庄家的开盘决定押注.

加线盘的设计方法和步骤 在实际操作中,庄家开盘的设计方法和步骤如下:

1. 估计获胜概率、开出公平盘 庄家首先组织专家(称开盘手)估计强队A的胜率,记作p.当两队胜率相等即$p=0.5$时,竞猜没有任何意义,设定开盘为-100/+100.而当强队A胜率为p时,容易验证,当开盘为$-\frac{100p}{1-p}/+\frac{100p}{1-p}$时,押强队A胜和押弱队B胜的玩家劣势均等于零.显然,强队A的胜率p越大,开盘中的数字就越大,如$p=0.6$时开盘为-150/+150,$p=0.8$时开盘为-400/+400.这种开盘的庄家优势也是零,称为**公平盘**.

2. 调整公平盘、开出加线盘 庄家为了有利可图,需要在公平盘基础上加以调整,称为加线.如$p=0.6$时在-150/+150上各加10(线),变成-160/+140,即押强队A胜的下注额加10,押弱队B胜的获赔减10,这一加一减就制造出庄家优势.一般地,设在公平盘$-\frac{100p}{1-p}/+\frac{100p}{1-p}$上强队A和弱队B的加线分别为$a$和$b(a,b>0)$,开出的加线盘为

$$-\left(\frac{100p}{1-p}+a\right)/+\left(\frac{100p}{1-p}-b\right) \tag{6}$$

可以算出押强队A胜和押弱队B胜的玩家的期望收益分别为$-a(1-p)$和$-b(1-p)$,于

是庄家的期望收益分别为 $a(1-p)$ 和 $b(1-p)$.当 $a=b$ 时,二者相等.

在押强队 A 胜和押弱队 B 胜的玩家数量相等的情况下,单位下注额的整体庄家优势为

$$\frac{(a+b)(1-p)}{\frac{100p}{1-p}+a+100}=\frac{(a+b)(1-p)^2}{100+a(1-p)} \tag{7}$$

当 a,b 增加时,整体庄家优势变大①.

3. 加线的调整 在一场竞猜中庄家的实际收益并不能按照期望值计算,而是取决于多少玩家押强队 A 胜、多少玩家押弱队 B 胜,以及结果是 A 还是 B 获胜.设押强队 A 胜与押弱队 B 胜的玩家数量(假定每人押一注)分别为 n_A,n_B,对于开出的加线盘(6),计算庄家的实际收益如下:

若结果是强队 A 胜,则庄家的实际收益为

$$V_A=-100n_A+100n_B=100(n_B-n_A) \tag{8}$$

若结果是弱队 B 胜,则庄家的实际收益为

$$V_B=\left(\frac{100p}{1-p}+a\right)n_A-\left(\frac{100p}{1-p}-b\right)n_B=\frac{100p}{1-p}(n_A-n_B)+an_A+bn_B \tag{9}$$

仍设 $p=0.6,a=b=10$,即开盘为-160/+140,由(9)式得 $V_B=150(n_A-n_B)+10(n_A+n_B)$,对于不同的 n_A,n_B 及胜负结果,计算庄家的实际收益如表 1.

表 1 开盘-160/+140 下的庄家实际收益

	n_A	n_B	获胜队	庄家实际收益
1	100	100	A	0
2	100	100	B	2 000
3	150	50	A	-10 000
4	50	150	A	10 000
5	150	50	B	17 000
6	50	150	B	-13 000

由表 1 可以看出,当 n_A,n_B 相差很大时,庄家可能暴赢也可能暴输.因为比赛结束前庄家并不知道最终哪一队获胜,所以为了规避风险,庄家必须密切关注玩家下注的走势.一旦发现 n_A,n_B 相差变大,就应适时地调整 a,b,尽量使 n_A,n_B 接近.从(8)、(9)式看出,只要 $n_A=n_B$,不论 A 胜还是 B 胜,庄家都不会赔钱.读者可以替庄家考虑一下,如果发现 n_A 比 n_B 大了很多,应该怎样调整、提高 a 和 b.

评注 无论是赌场里的各种游戏,还是博彩公司开盘的体育竞猜,从表面上看玩家都是在与庄家对赌,玩家赢了向庄家讨取赔付,输了下注额归庄家.但是实质上,现代的博彩无一不是玩家与玩家之间的博弈,赌场里每天无数的筹码从大部分玩家的口袋,经庄家之手转入少数赢钱赌客的囊中,庄家只靠玩家对赌而从中抽头牟利,自己绝不涉赌.体育竞猜中博彩公司开盘后一再调整加线,竭力均衡押注两队的玩家之间的角力,也是为了促成玩家的对赌,实现自己抽身而退.

复习题

证明加倍下注策略(赔率为 1)与普通下注方式玩家(单位下注金额)的期望收益相同.

① 其中庄家优势对 a 的单调性要求 $b<100/(1-p)$,这通常是满足的.

8.7 钢琴销售的存贮策略

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 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 **Markov 链模型**描述.8.7 和 8.8 节介绍 2 个 **Markov** 链模型.不熟悉 **Markov** 链的读者可参考基础知识 8-1.



基础知识 8-1 Markov 链的基本概念

像钢琴这样的奢侈品销售量很小，商店里一般不会有多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[43].

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 **Markov** 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

$$P(D_n=k)=\mathrm{e}^{-1}/k! \quad (k=0,1,2,\cdots) \tag{1}$$

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

$$S_{n+1}=\begin{cases} S_n-D_n, & D_n<S_n \\ 3, & D_n\geqslant S_n \end{cases} \tag{2}$$

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

$$\boldsymbol{P}=\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

其中

$$p_{11}=P(S_{n+1}=1\mid S_n=1)=P(D_n=0)=0.368$$
$$p_{12}=P(S_{n+1}=2\mid S_n=1)=0$$
$$p_{13}=P(S_{n+1}=3\mid S_n=1)=P(D_n\geqslant 1)=0.632$$
$$p_{21}=P(S_{n+1}=1\mid S_n=2)=P(D_n=1)=0.368$$
$$p_{22}=P(S_{n+1}=2\mid S_n=2)=P(D_n=0)=0.368$$
$$p_{23}=P(S_{n+1}=3\mid S_n=2)=P(D_n\geqslant 2)=0.264$$
$$p_{31}=P(S_{n+1}=1\mid S_n=3)=P(D_n=2)=0.184$$
$$p_{32}=P(S_{n+1}=2\mid S_n=3)=P(D_n=1)=0.368$$
$$p_{33}=P(S_{n+1}=3\mid S_n=3)=P(D_n=0)+P(D_n\geqslant 3)=0.448$$

得到

$$\boldsymbol{P}=\begin{bmatrix}0.368 & 0 & 0.632\\ 0.368 & 0.368 & 0.264\\ 0.184 & 0.368 & 0.448\end{bmatrix} \tag{3}$$

记状态概率 $a_i(n)=P(S_n=i)$，$i=1,2,3$，$\boldsymbol{a}(n)=(a_1(n),a_2(n),a_3(n))$，根据状态转移具有无后效性的假设，有

$$\boldsymbol{a}(n+1)=\boldsymbol{a}(n)\boldsymbol{P}$$

可以看出，转移矩阵 $\boldsymbol{P}$ 是一个正则链，具有稳态概率 $\boldsymbol{w}$，满足 $\boldsymbol{wP}=\boldsymbol{w}$，可得

$$\boldsymbol{w}=(w_1,w_2,w_3)=(0.285,0.263,0.452) \tag{4}$$

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

$$P(D_n>S_n)=\sum_{i=1}^{3}P(D_n>i\mid S_n=i)P(S_n=i) \tag{5}$$

其中的条件概率 $P(D_n>i\mid S_n=i)$ 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

$$R_n=\sum_{i=1}^{3}\left[\sum_{j=1}^{i-1}jP(D_n=j\mid S_n=i)+iP(D_n\geqslant i\mid S_n=i)\right]P(S_n=i) \tag{6}$$

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　这个模型用到的唯一一个原始数据是，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 $P(D_n=k)=\lambda^k e^{-\lambda}/k!$ $(k=0,1,2,\cdots)$，由此得状态转移矩阵为

$$\boldsymbol{P}=\begin{bmatrix} e^{-\lambda} & 0 & 1-e^{-\lambda} \\ \lambda e^{-\lambda} & e^{-\lambda} & 1-(1+\lambda)e^{-\lambda} \\ \lambda^2 e^{-\lambda}/2 & \lambda e^{-\lambda} & 1-(\lambda+\lambda^2/2)e^{-\lambda} \end{bmatrix} \tag{7}$$

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 动态随机存贮策略是Markov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.

复习题

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Markov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.

8.8 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Markov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,一优种一混种 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.

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.

表 1 父母基因类型决定后代各种基因类型的概率

后代基因类型	父母基因类型					
	DD	*RR*	*DH*	*DR*	*HH*	*HR*
D	1	0	$\frac{1}{2}$	0	$\frac{1}{4}$	0
H	0	0	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$
R	0	1	0	0	$\frac{1}{4}$	$\frac{1}{2}$

下面以 Markov 链为工具讨论混种繁殖、优种繁殖和近亲繁殖 3 个基因遗传模型.

混种繁殖 假设在繁殖过程中用一混种与一个个体交配,所得后代仍用混种交配,如此继续下去,称为混种繁殖.建立 Markov 链模型描述在混种繁殖下各代具有 3 种基因类型的概率,并讨论稳态情况.

状态用基因类型定义,$i=1,2,3$ 依次表示优种 D、混种 H 和劣种 R,状态概率 $a_i(k)$ 表示第 k 代个体具有第 i 种基因类型的概率,记 $\boldsymbol{a}(k)=[a_1(k),\ a_2(k),\ a_3(k)]$.当用混种 H 与优种 D 交配时,其后代的基因类型是 D、H 和 R 的概率如表 1 第 4 列;类似地,当用混种 H 与混种 H 和劣种 R 交配时,后代的基因类型是 D、H 和 R 的概率如表 1 第 6 列和第 7 列.这样就可以写出转移概率矩阵为(表 1 的列转换为矩阵的行)

$$\boldsymbol{P}=\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \tag{1}$$

若初始时混种与优种交配,即 $\boldsymbol{a}(0)=[1,0,0]$,按照 Markov 链的基本方程

$$\boldsymbol{a}(k+1)=\boldsymbol{a}(k)\boldsymbol{P}$$

计算任意时段的状态概率 $\boldsymbol{a}(k)$,结果如表 2.当 $k\to\infty$ 时 $\boldsymbol{a}(k)\to\boldsymbol{a}=[1/4,1/2,\ 1/4]$,表示经过足够多代繁殖以后,优种、混种、劣种的比例接近于 1∶2∶1.

表 2 混种繁殖下 3 种基因类型的状态概率(初始与优种交配)

k	0	1	2	3	4	5	…	∞
$a_1(k)$	1	$\frac{1}{2}$	$\frac{3}{8}$	$\frac{5}{16}$	$\frac{9}{32}$	$\frac{17}{64}$	…	$\frac{1}{4}$
$a_2(k)$	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	…	$\frac{1}{2}$
$a_3(k)$	0	0	$\frac{1}{8}$	$\frac{3}{16}$	$\frac{7}{32}$	$\frac{15}{64}$	…	$\frac{1}{4}$

如果初始时混种与混种交配、混种与劣种交配,当 $k\to\infty$ 时也会有同样的结果.实际上,由(1)式容易检验 $\boldsymbol{P}^2>\boldsymbol{O}$,这个 Markov 链是正则链,存在唯一的稳态概率 $\boldsymbol{a}=[a_1,$

$a_2, a_3]$，而 $\boldsymbol{a}$ 可由 $\boldsymbol{aP}=\boldsymbol{a}$ 解出，得 $\boldsymbol{a}=[1/4, 1/2, 1/4]$，与表 2 的结果一致.

优种繁殖　对状态和状态概率的定义同上，只是将混种交配改为每代都用优种交配，那么当用优种 D 与 D、H、R 交配时，其后代的基因类型是 D、H 和 R 的概率如表 1 第 2 列、第 4 列和第 5 列. 于是优种繁殖的转移概率矩阵为

$$\boldsymbol{P}=\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{2}$$

与吸收链转移矩阵的标准形式比较，可知这个 Markov 链是吸收链，优种 $D(i=1)$ 是吸收态. 不论初始时用优种与优种、混种还是劣种交配，最终都将成为优种.

在农业、畜牧业生产中常常是纯种（优种或劣种）的抗病性等品质不如混种，仅从这个意义上说，根据上面的两个 Markov 链模型，用混种繁殖比用优种繁殖的效果要好.

近亲繁殖　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 Markov 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 $p_{33}=P$（后代配对为 DH | 父母配对为 DH）$=P$（后代雄性为 D，雌性为 H | 父母配对为 DH）$+P$（后代雄性为 H，雌性为 D | 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，最后得到转移矩阵为

$$\boldsymbol{P}=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{2} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{1}{16} & \frac{1}{16} & \frac{1}{4} & \frac{1}{8} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \tag{3}$$

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. 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.

为了计算从任一个非吸收状态 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(3)式表示的 $\boldsymbol{P}$ 化为转移矩阵的标准形式 $\boldsymbol{P}=\begin{bmatrix}\boldsymbol{I} & \boldsymbol{O}\\ \boldsymbol{R} & \boldsymbol{Q}\end{bmatrix}$，得到

$$\boldsymbol{Q}=\begin{bmatrix}\frac{1}{2} & 0 & \frac{1}{4} & 0\\ 0 & 0 & 1 & 0\\ \frac{1}{4} & \frac{1}{8} & \frac{1}{4} & \frac{1}{4}\\ 0 & 0 & \frac{1}{4} & \frac{1}{2}\end{bmatrix},\boldsymbol{R}=\begin{bmatrix}\frac{1}{4} & 0\\ 0 & 0\\ \frac{1}{16} & \frac{1}{16}\\ 0 & \frac{1}{4}\end{bmatrix} \tag{4}$$

然后计算

$$\boldsymbol{M}=(\boldsymbol{I}-\boldsymbol{Q})^{-1}=\begin{bmatrix}\frac{8}{3} & \frac{1}{6} & \frac{4}{3} & \frac{2}{3}\\ \frac{4}{3} & \frac{4}{3} & \frac{8}{3} & \frac{4}{3}\\ \frac{4}{3} & \frac{1}{3} & \frac{8}{3} & \frac{4}{3}\\ \frac{2}{3} & \frac{1}{6} & \frac{4}{3} & \frac{8}{3}\end{bmatrix} \tag{5}$$

$$\boldsymbol{y}=\boldsymbol{M}\boldsymbol{e}=\left(4\frac{5}{6},6\frac{2}{3},5\frac{2}{3},4\frac{5}{6}\right)^{\mathrm{T}} \tag{6}$$

$$\boldsymbol{F}=\boldsymbol{M}\boldsymbol{R}=\begin{bmatrix}\frac{3}{4} & \frac{1}{4}\\ \frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2}\\ \frac{1}{4} & \frac{3}{4}\end{bmatrix} \tag{7}$$

$\boldsymbol{M}$ 的第 1 行至第 4 行依次代表非吸收状态 DH，DR，HH 和 HR，根据对于向量 $\boldsymbol{y}$ 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.而根据基础知识 8-1 的定理 5，被吸收状态吸收的概率为矩阵 $\boldsymbol{F}$ 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.

评注 Markov 链模型是研究离散时间、离散状态随机转移过程的有力工具，在经济、社会、生态、遗传等许多领域有着广泛的应用.状态转移的无后效性是应用 Markov 链模型的先决条件，定义状态和状态概率，并构造转移概率矩阵是建立 Markov 链模型的关键.一些实际问题需要恰当地选择状态，精心计算转移概率，甚至有的随机转移过程粗看起来并不具有无后效性，但通过适当扩充状态可以改变其性质，仍然可以构建 Markov 链模型.

复习题

1. 在基因遗传中，若初始时取一劣种，通过计算三种基因类型状态概率，比较混种繁殖和优种繁殖两种形式，经过足够多代繁殖以后的演变.

2. 推导近亲繁殖的概率转移矩阵 $\boldsymbol{P}$((3)式)的第5行.

8.9 自动化车床管理

自动化车床管理是全国大学生数学建模竞赛1999年A题，原文如下：

一道工序用自动化车床连续加工某种零件，由于刀具损坏等原因该工序会出现故障，其中刀具损坏故障占95%，其他故障仅占5%. 工序出现故障是完全随机的，假定在生产任一零件时出现故障的机会均相同. 工作人员通过检查零件来确定工序是否出现故障.

数据文件8-1
自动化车床管理

现积累有100次刀具故障记录，故障出现时该刀具完成的零件数见数据文件8-1. 现计划在刀具加工一定件数后定期更换新刀具.

已知生产工序的费用参数如下：

故障时产出的零件损失费用 $f=200$ 元/件；

进行检查的费用 $t=10$ 元/次；

发现故障进行调节使恢复正常的平均费用 $d=3\ 000$ 元/次(包括刀具费)；

未发现故障时更换一把新刀具的费用 $k=1\ 000$ 元/次.

1) 假定工序故障时产出的零件均为不合格品，正常时产出的零件均为合格品，试对该工序设计效益最好的检查间隔(生产多少零件检查一次)和刀具更换策略.

2) 如果该工序正常时产出的零件不全是合格品，有2%为不合格品；而工序故障时产出的零件有40%为合格品，60%为不合格品. 工序正常而误认有故障停机产生的损失费用为1 500元/次. 对该工序设计效益最好的检查间隔和刀具更换策略.

3) 在2)的情况，可否改进检查方式获得更高的效益？

本节以发表在《数学的实践与认识》2000年第1期上参赛学生的优秀论文和评述文章为基本材料，加以整理和归纳，介绍建模的全过程.

问题分析 在生产设备运行过程中，工具、部件如本题自动化车床的刀具，一般都会随机地发生故障，需要通过对生产的零件及时检查发现故障，加以维修或更换. 同时，由于刀具的寿命有限，运行到一定时期后，即使尚未出现故障，也进行所谓**预防性更换**，以减少刀具带故障运行造成的损失，经济上可能更为合算. 所以对刀具的检查和更换策略应该包含两个方面，一是确定一个检查间隔(以生产的零件数计量)，每次检查时若发现故障立即更换刀具. 二是确定一个更换周期(也以生产的零件数计量)，到期即便刀具运行正常也需更换. 显然，为了方便起见，应使更换周期是检查间隔的整数倍.

本题要求确定能使效益达到最佳的检查间隔和更换周期. 由于题目给出的是检查、更换、零件损失等费用，所以可以用每生产一个零件付出的总费用最小为目标. 而因为刀具发生故障是随机的，于是应该用总费用的期望值即平均费用来计算. 由此，首先需

要根据题目给出的刀具故障记录，找出刀具无故障时间（以故障出现时完成的零件数计量）的概率分布.

以上叙述的是生产过程中的一种基本的、简单的情况：通过零件检查发现故障，即认定是刀具故障并进行更换，而且零件检查不存在错误判断.这道题目还需考虑两个附加内容：一是其他故障的影响，即通过检查零件发现的除刀具故障外，还有其他故障，二者合称工序故障.二是零件检查存在误判的影响，即工序正常时生产的零件有不合格品，若检查出不合格品会误判为工序故障；工序故障时生产的零件有合格品，若检查出合格品又会误判为工序正常.这两种情况增加了建模的复杂性.下面先建立基本问题的模型，再讨论这两个附加内容.

建模准备 作100次刀具故障记录的直方图，近似于正态分布.计算出均值 $\mu=600$，均方差 $\sigma=196$，经统计检验，可认为刀具的无故障时间或称刀具寿命（以故障出现时完成的零件数计量）服从正态分布 $N(\mu,\sigma^2)$.

模型假设与符号 对刀具检查和更换的基本问题作如下假设：

1. 检查间隔为常数 n 件，预防性更换周期为常数 m 件，且 $m=sn$，其中 s 为整数.

2. 检查时若发现零件为不合格品，则认为是刀具故障，需立即更换.

3. 相对于刀具寿命而言，检查间隔很短，可认为在相邻两次检查之间每个零件出现故障是等可能的.

4. 刀具寿命 $X\sim N(600,196^2)$，其分布函数和密度函数分别记作 $F(x)$ 和 $f(x)$，$F(x)$ 是刀具寿命不超过 x 的概率.

5. 检查费 $t=10$ 元/次，预防性更换刀具费 $k=1\ 000$ 元/次，零件为不合格品的损失费 $f=200$ 元/件，发现故障更换刀具并恢复正常生产的费用 $d=3\ 000$ 元/次.

基本模型 先计算一个更换周期 m 的平均费用，包括预防更换费用和故障更换费用两部分.

由假设1，在刀具预防性更换前共检查 $s-1$ 次，由假设5得预防更换费用为 $(s-1)t+k$.由假设4，预防更换的刀具寿命应超过 m，其概率为 $1-F(m)$.

若在第 i 次检查中发现故障（$i=1,2,\cdots,s-1$），由假设3，在第 $i-1$ 次到第 i 次检查中生产出1个、2个、……、n 个不合格零件的概率相等，平均为 $\frac{n+1}{2}$ 件，由假设5，零件损失费为 $\frac{f(n+1)}{2}$，检查、更换及恢复生产费用共 $it+d$，故障更换总费用为 $it+d+\frac{f(n+1)}{2}$.由假设4，第 i 次检查发现（第 $i-1$ 次检查未发现）故障的概率为 $F(in)-F((i-1)n)$.

记一个更换周期的平均费用为 EC，则

$$EC=c_1[1-F(m)]+\sum_{i=1}^{s-1}c_{2i}[F(in)-F((i-1)n)]$$

$$c_1=(s-1)t+k,\ c_{2i}=it+d+\frac{f(n+1)}{2} \tag{1}$$

为了表达和计算的方便，利用密度函数 $f(x)$ 将和式化为积分（其中 $x=in$），进行整理并取近似可得

$$EC=k[1-F(m)]+\left[d+\frac{f(n+1)}{2}\right]F(m)+\left[\int_0^m xf(x)\mathrm{d}x+m(1-F(m))\right]\frac{t}{n} \tag{2}$$

利用分部积分可将(2)式的一部分简化为

$$\int_0^m xf(x)\,dx + m(1 - F(m)) = m - \int_0^m F(x)\,dx \tag{3}$$

然后计算一个更换周期生产零件的平均数量,记作 ER.只需在(1)式中令 $c_1 = m, c_{2i} = in$,与(1)~(3)式类似推导可得

$$ER = m - \int_0^m F(x)\,dx \tag{4}$$

记 EL 为一个更换周期中生产一个零件的平均费用,定义为①

$$EL = \frac{EC}{ER} \tag{5}$$

检查和更换问题的基本模型为,确定检查间隔 n 和更换周期 m,使(1)~(5)式表示的一个零件的平均费用 EL 达到最小.

模型用搜索法求解,即设定一系列的 n 和 m(m 为 n 的整数倍),计算 EL,选出使 EL 最小的那一组解为 $n = 18, m = 360$②.

其他故障的影响 工序故障包括刀具故障和其他故障两类,检查只能发现工序故障,在基本模型中认为工序故障就是刀具故障.

为了考虑其他故障的影响,先计算其他故障出现的概率,即故障率.记 p, q 分别是刀具故障率和其他故障率,根据两类故障分别占 95%和 5%,我们有 $p:q = 95:5$.因为刀具故障的均值为 $\mu = 600$,所以刀具故障率是 $p = \frac{1}{\mu} = \frac{1}{600}$,于是 $q = \frac{5}{95} \times p = \frac{1}{11\,400}$.

假定生产任一零件时其他故障出现的概率均为 q,并且相互独立,则生产第 j 个零件才出现其他故障的概率 P_j 服从几何分布,即 $P_j = (1-q)^{j-1}q (j = 1, 2, \cdots, m-1)$,而 $P_m = (1-q)^{m-1}$(第 m 个零件后不再生产).与刀具故障下平均费用 EC 的公式(1)类似,若记其他故障下平均费用为 EC_1,则

$$EC_1 = c_1(1 - q)^{m-1} + \sum_{j=1}^{m-1} c_{2j}(1 - q)^{j-1}q, \quad c_{2j} = \left[\frac{j}{n}\right] t + d + \frac{f(n+1)}{2} \tag{6}$$

其中 c_1 同前.再记其他故障下生产零件的平均数量为 ER_1,则有

$$ER_1 = m(1 - q)^{m-1} + \sum_{j=1}^{m-1} j(1 - q)^{j-1}q = \frac{1 - (1 - q)^m}{q} \tag{7}$$

将工序故障下的平均费用记作 C_2,生产零件的平均数量记作 R_2,C_2 和 R_2 可分别由刀具故障和其他故障的平均费用 EC, EC_1 和 ER, ER_1 加权得到,即

$$C_2 = \lambda EC + (1-\lambda) EC_1, \quad R_2 = \lambda ER + (1-\lambda) ER_1 \tag{8}$$

将生产一个零件的平均费用记作 L_2,与(5)式类似有

$$L_2 = C_2 / R_2 \tag{9}$$

如果将刀具故障和其他故障出现的比例 95%和 5%作为权重,那么由(6)~(9)式表示的其他故障的影响不会太大.

---

① 因不合格零件的损失费已经计入 EC,不合格零件的平均数量在故障更换时仅为 $(n+1)/2$,与 ER 相比很小,所以 L 也可看作"生产一个合格零件的平均费用".

② 《数学的实践与认识》2000 年第 1 期于杰等论文的结果是 $n = 18, m = 360$;石敏等论文是 $n = 18, m = 342$.

零件检查存在误判的影响 当零件检查存在两种误判时，讨论基本模型（工序故障中只考虑刀具故障）应做哪些修改.

第一种误判是由于刀具正常时生产的零件有 2%的不合格品，若检查时正好查到不合格品，则会判断刀具故障导致停机，由此产生的损失费用为 1 500 元/次.

在每个检查间隔内生产每个零件时的刀具故障率是 $p=\frac{1}{\mu}=\frac{1}{600}$，所以每次检查时刀具正常的概率为 $(1-p)^n$，因查到不合格品误判停机产生的损失费为 $0.02\times 1\ 500\times(1-p)^n=30(1-p)^n$，于是基本模型（1）式 c_1, c_{2i} 中的检查费 t 应增加为 $t+30(1-p)^n$.

第二种误判是由于刀具故障时生产的零件有 40%合格品，若检查时查到的是合格品，则会判断刀具正常，将继续生产直到下一次检查，导致不合格品数量的增加.经过仔细推导得增加量可近似为 $n\times\frac{0.4}{1-0.4}=\frac{2n}{3}$，于是基本模型（1）式 c_{2i} 中的零件损失费 $\frac{f(n+1)}{2}$ 应增加为 $f\left[\frac{(n+1)}{2}+\frac{2n}{3}\right]$.

模型改进 在上面的模型中检查间隔 n 是常数，但是由于刀具寿命服从正态分布，新刀具发生故障的概率很小，随着生产零件数的增加，发生故障概率逐渐变大，所以直观地看，为了降低检查费用，应该让刀具开始使用阶段的检查间隔大一些，然后逐渐缩小.一种具体做法是根据“每个检查间隔内刀具发生故障的概率相等”的原则，确定一系列的检查点（进行检查的零件顺序号）①，还可以参考本书第三版 13.5 节，将检查间隔视为生产零件数的函数 $n(x)$，以某种定义下的费用函数最小为目标，建立泛函极值问题的优化模型.当然，不等间隔的检查会带来实际操作上的麻烦.

如果零件检查误判造成的损失远大于检查费（本题给出的费用数据就是如此），那么增加检查零件的数量应会减小误判的影响.当然，检查几个零件（不宜过多）、怎样根据零件的合格与否判断刀具故障等，都需要经过仔细演算.

由日本田口玄一博士创立的所谓田口方法，也可以确定生产过程中零部件的检查间隔和更换周期，用于求解本题.②

评注 自动化生产过程中零部件的检查和更换是保证产品质量、降低运行成本的重要手段，而最优的检查间隔和更换周期是由各种费用的构成比例以及零部件的寿命分布决定的.根据基本模型的求解，检查间隔 $n=18$，更换周期 $m=360$.注意到由 100 次刀具故障记录计算出的刀具寿命均值 $\mu=600$，可知在预防性更换时刀具大都未出现故障，这是由于刀具故障后生产的每件不合格零件的损失费 $f(=200)$，相对于更换刀具费 $k(=1\ 000)$ 及发现故障更换刀具并恢复正常生产的费用 $d(=3\ 000)$ 较大的缘故.

第 8 章训练题

1. 出版社每年都要重印某种教科书，按照过去的销售记录知道今年需求量大致为均值 8 000 本、均方差 1 000 本的正态分布.已知每本书的成本 15 元，售价 50 元，如果供过于求则以售价的 2 折处理，问年初应重印多少本使出版社平均收入最大？这个收入是多少？如果供不应求，为保证学生用书必须临时加印以满足需求，每本书成本加倍，售价不变，再问年初应重印多少本使出版社平均收入最大？这个收入是多少？

2. 某单位每年一届会议的组织者要为各地客户在一家有合同约定的宾馆预订住房，每个房间需交 100 元的预订金，若预订房间过多，即使无人入住预订金也不退回.而若预订房间不足，则需为参会

① 见《数学的实践与认识》2000 年第 1 期戚正君等的论文.

② 见《数学的实践与认识》2000 年第 1 期孙山泽教授的文章.

客户在另外的宾馆临时订房，每个房间租金250元.按照往年会议的规模估计今年参会人数在250人至350人之间(95%的把握).请为组织者确定预订房间的数量(按每人一间)，使预期的总费用最低.

3. 在8.4节Simmons模型中假设问题B为“你喜欢红色吗?”，而且调查者不知道喜欢红色的学生的比例，调查方法是将被调查者分成两批，做两次独立调查，且选答问题A的概率不同.试设计具体的调查方案并建立模型，估计有过作弊行为学生的比例.

4. 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$m_i^{(Y)}$.

(1) 分别估计学生中曾有作弊和赌博行为的比例.

(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.

5. 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，使某种意义下的浪费量最小.

6. 双骰是根据掷出的两颗骰子点数之和判断输赢的一种游戏.玩家下注后掷出两颗骰子，称出场掷.输赢规则如下：

出场掷两颗骰子点数之和是7、11，玩家赢，此轮结束.

出场掷两颗骰子点数之和是2、3、12，玩家输，此轮结束.

出场掷两颗骰子点数之和是4、5、6、8、9、10，该数字成为牌点，玩家继续掷骰子，直到牌点或点数之和7出现.如果牌点先出现，玩家赢，此轮结束；如果7先出现，玩家输，此轮结束.

(1) 计算各种结果出现的概率，并由此计算玩家劣势.

(2) 双骰游戏有不同的下注机制，前面给出的是“过线”(pass line)方式，另一种是“不过线”(don't pass)方式，其输赢规则与“过线”方式正好相反，只是出场掷两颗骰子点数之和是12时，仍为玩家输.计算“不过线”方式各种结果出现的概率，并由此计算玩家劣势.

7. 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Markov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.

8. 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Markov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Markov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?(参考基础知识8-1 Markov链的基本概念)

9. 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Markov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.

10. 一个服务网络由 k 个工作站 $v_1,v_2,\cdots,v_k$ 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 $q_i(i=1,2,\cdots,k-1$，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 Markov 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？（参考基础知识 8-1　Markov 链的基本概念）

11. 血样的分组检验[17]. 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%，…，0.1%，…，1%，…）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.

12. 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.

更多案例……

8-1　随机存贮策略

8-2　随机入口模型

8-3　健康与疾病

8-4　资金流通

第9章　统计模型

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，本章主要介绍用途非常广泛的一类统计模型——各种类型的回归模型，包括线性回归(9.1 节，9.2 节)、非线性回归(9.3 节)、自回归(9.4 节)、logit 回归(9.5 节)，以及判别分析模型(9.6 节)和主成分分析模型(9.7 节).

9.1　孕妇吸烟与胎儿健康

“吸烟有害健康”是人们的共识，那么孕妇吸烟是否还会伤害到腹中的胎儿呢？美国公共卫生总署在香烟盒上设置的一条健康警告是：“孕妇吸烟可能导致胎儿受损、早产及新生儿低体重”.他们在一份报告中更详细的叙述是：“对于新生儿体重来说，看来吸烟比妇女怀孕前身高、体重、受孕历史等因素的影响更为显著.孕妇吸烟也会增加早产率，对每个受孕年龄段妇女来说新生儿都更小”[48].



数据文件 9-1
胎儿健康 CHDS 原始数据
data0901.txt

让我们利用美国儿童保健和发展项目(简记 CHDS)提供的一份数据来研究上述说法.这份数据包括 1 236 个出生后至少存活 28 天的男性单胞胎新生儿的体重及其母亲的相关资料，见表 1，全部数据见数据文件 9-1.

表 1　CHDS 的部分原始数据(数据量 $n=1236$)

1. 新生儿体重/oz(1 oz=28.35 g)	120	113	128	123	108	136	138	…
2. 孕妇怀孕期/天	284	282	279	999①	282	286	244	…
3. 新生儿胎次状况(1——第 1 胎，0——非第 1 胎)	1	0	1	0	1	0	0	…
4. 孕妇怀孕时年龄	27	33	28	36	23	25	33	…
5. 孕妇怀孕前身高/in(1 in=2.54 cm)	62	64	64	69	67	62	62	…
6. 孕妇怀孕前体重/lb(1 lb=0.454 kg)	100	135	115	190	125	93	178	…
7. 孕妇吸烟状况(1——吸烟，0——不吸烟)	0	0	1	1	1	1	0	…

CHDS 的数据来自美国奥克兰市参加支付医疗保险项目的家庭，在就业人群中具有广泛的代表性.参加这个项目的妇女在怀孕初期就接受了全面、详细的调查，得到了新生儿父母双方诸多方面的信息，这里给出的数据只是调查记录的一部分.因为关于孕妇的那些数据是怀孕初期得到的，所以不会受到所关心的结果——新生儿体重的影响.

我们的任务是利用这份数据来建立新生儿体重与孕妇怀孕期、吸烟状况等诸因素

① CHDS 原始数据中用 999(二位数用 99，一位数用 9)表示该数据缺失.

之间的数学模型，定量地讨论以下问题：

- 对于新生儿体重来说，孕妇吸烟是否是比孕妇年龄、身高、体重等更为显著的决定因素；
- 孕妇吸烟是否会使早产率增加，怀孕期长短是否对新生儿体重有影响；
- 对每个年龄段来说，孕妇吸烟对新生儿体重和早产率的影响是怎样的.

问题背景及分析　美国公共卫生总署的警告容易受到人们的质疑，因为按照是否吸烟划分人群所做的任何研究，都只能依赖于观测数据，而不可能受研究者的控制，即无法作人为的实验，于是很难确定新生儿体重的差别是由于吸烟的缘故，还是与吸烟相关联的其他因素（如怀孕期长短）有关，又如吸烟孕妇往往年轻，自身体重较轻，会不会由此造成新生儿体重低呢？

显然，“孕妇吸烟可能导致胎儿受损、早产及新生儿低体重”的警告从语气上就不如“吸烟导致肺癌”来得强，这是由于对孕妇吸烟与胎儿健康之间生理学上的关系研究得不够.现在能够解释的理由是：吸烟产生一氧化碳，由于血液中血色素与一氧化碳的亲和力远大于与氧的亲和力，血色素与一氧化碳结合形成的血红蛋白限制了血液携带的氧的流量，于是孕妇吸烟会造成氧对胎儿的供应量减少.虽然这对胎儿生理上的影响目前尚不完全清楚，但已有的研究表明，为了补偿氧供应的减少，孕妇胎盘的表面积和血管数量以及胎儿的血色素水平都要增加，血流会重新分布以供养胎儿的各个组织，导致胎盘直径变大和变细，其后果可能是使胎盘过早地从子宫壁上剥离，从而导致早产甚至胎儿死亡.

关于新生儿体重及孕妇早产的标准，一般认为，新生儿的正常体重为 2 500～4 000 g(88.2～141.1 oz)，小于 2 500 g 视为体重低；正常怀孕期为 40 周左右，小于 37 周视为早产.下面采用这个标准来讨论.

参数估计与假设检验　首先根据 CHDS 的数据对不吸烟与吸烟孕妇的各项指标作参数估计，包括新生儿体重均值的点估计（记作 μ_{y0} 和 μ_{y1}）和区间估计、新生儿体重低的比例（指体重小于2 500 g 的比例）的点估计（记作 r_0 和 r_1）、怀孕期均值的点估计（记作 μ_{x0} 和 μ_{x1}）和区间估计、早产率（指孕期小于 37 周的比例）的点估计（记作 q_0 和 q_1），结果如表 2.

基础知识 9-1
参数估计与假设检验的基本内容

表 2　不吸烟与吸烟孕妇的新生儿体重和怀孕期的参数估计（n 是样本数，下同）

参数估计	不吸烟孕妇（$n=742$）	吸烟孕妇（$n=484$）①
新生儿体重(oz)均值的点估计	$\mu_{y0}=123.047\ 2$	$\mu_{y1}=114.109\ 5$
新生儿体重均值的区间估计	[121.793 2,124.301 1]	[112.493 0,115.726 0]
新生儿体重低的比例的点估计	$r_0=0.031\ 0$	$r_1=0.082\ 6$
怀孕期(天)均值的点估计	$\mu_{x0}=280.186\ 9(n=733)$	$\mu_{x1}=277.979\ 2(n=480)$②
怀孕期均值的区间估计	[278.981 2,281.392 6]	[276.627 3,279.331 1]
早产率的点估计	$q_0=0.076\ 4$	$q_1=0.085\ 4$

程序文件 9-1
胎儿健康的参数估计与假设检验
prog0901a.m

① 在总共 1236 个数据中有 10 位孕妇吸烟状况数据缺失.

② 又有 13 位孕妇怀孕期数据缺失.

从表2可以看出,吸烟孕妇比不吸烟孕妇的新生儿体重平均约低9 oz(约250 g),新生儿体重低的比例则明显地高;吸烟孕妇的怀孕期比不吸烟孕妇平均只约短2天,早产率相差不多.对于不吸烟与吸烟孕妇来说,新生儿体重和怀孕期的这些差别在统计学意义上是不是显著的,需要进行两总体的假设检验,结果如表3.

表3 不吸烟与吸烟孕妇的新生儿体重和怀孕期的假设检验(显著性水平 $\alpha=0.05$)

假设检验	假设	检验结果
新生儿体重均值的假设检验	$H_0:\mu_{y0}\leqslant\mu_{y1}$; $H_1:\mu_{y0}>\mu_{y1}$	拒绝 H_0,接受 H_1
新生儿体重低的比例的假设检验	$H_0:r_0\geqslant r_1$; $H_1:r_0<r_1$	拒绝 H_0,接受 H_1($t=4.0304$)
怀孕期均值的假设检验	$H_0:\mu_{x0}\leqslant\mu_{x1}$; $H_1:\mu_{x0}>\mu_{x1}$	拒绝 H_0,接受 $H_1$①
早产率的假设检验	$H_0:q_0=q_1$; $H_1:q_0\neq q_1$	接受 H_0,拒绝 H_1($t=0.5663$)



基础知识 9-2
线性回归分析的基本内容

评注 "吸烟孕妇比不吸烟孕妇的新生儿体重平均约低9 oz"的结论虽然有显著意义,但是这里面混杂着孕妇的怀孕期、体重等诸多可能影响新生儿体重的因素,利用回归分析能够将这些因素分开.

从表3可以看出,吸烟孕妇的新生儿体重比不吸烟孕妇的低,且新生儿体重低的比例高,这个结论在统计学上是有显著意义的;而吸烟与不吸烟孕妇的怀孕期和早产率的差别则难以肯定是显著的.

为了充分利用所给的数据进行全面的研究,需要借助回归分析方法,建立新生儿体重与孕妇吸烟状况、怀孕期等诸因素的回归模型,分析模型得到的结果,回答提出的问题.

一元线性回归分析 从参数估计和假设检验的结果可以知道,孕妇的吸烟状况对新生儿体重大小有显著影响,但是对怀孕期长短的影响难以确定.那么,新生儿体重与怀孕期之间有怎样的关系呢?显然需要对吸烟孕妇和不吸烟孕妇分别讨论.

对于CHDS数据中的480位吸烟孕妇,将新生儿体重和怀孕期分别记作 y_i 和 $x_i(i=1,2,\cdots,480)$,其散点图如图1.用 y 和 x 分别表示新生儿体重和怀孕期这两个变量,由图1假定二者的关系可以用一次函数模型 $y=b_0+b_1x$ 来描述,按照最小二乘准则,用数据拟合的办法很容易得到模型系数 b_0,b_1,这个一次函数就是图1中的那条直线.

从图1看出,虽然直线大致描述了数据点的变化趋势,但是拟合得并不好.怎样衡量由拟合得到的模型的有效性?模型系数的精确度和模型预测的数值范围有多大?实际应用中的这些重要问题需要用回归分析的方法来回答.

将新生儿体重(因变量 y)与孕妇怀孕期(自变量 x)的一元线性回归模型表为

$$y=b_0+b_1x+\varepsilon \tag{1}$$

其中随机误差 ε 是 x 以外影响 y 的随机因素的总和.假定对于不同的 x,ε 是相互独立的随机变量,服从 $N(0,\sigma^2)$ 分布.



程序文件 9-2
胎儿健康(吸烟孕妇)的一元回归
prog0901b.m

根据480位吸烟孕妇的新生儿体重和怀孕期数据,对回归模型(1)编程计算,结果如表4.

① 若显著性水平 $\alpha=0.01$,则接受 H_0.

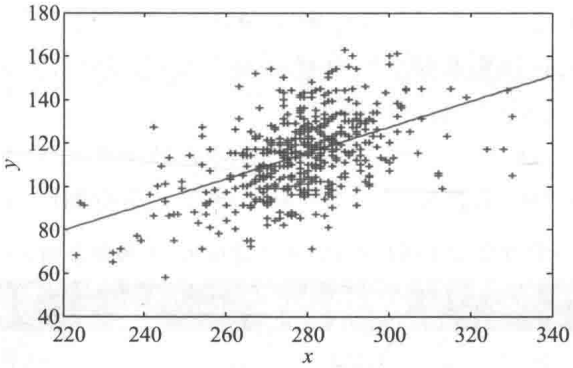


图 1 480 位吸烟孕妇的新生儿体重和怀孕期散点图及拟合的直线

表 4 模型(1)的计算结果(吸烟孕妇 $n=480$)

回归系数	系数估计值	系数置信区间
b_0	−51.298 3	[−77.511 0,−25.085 6]
b_1	0.594 9	[0.500 8,0.689 1]
$R^2=0.243\ 8$ $F=154.146\ 3$ $p<0.000\ 1$ $s^2=249.919\ 2$		

由表 4 可对模型做检验:b_1 的置信区间不含零点,$F=154.146\ 3$ 远大于 $F_{(1,n-2),1-\alpha}=3.861\ 0(\alpha=0.05)$,表示应拒绝 $H_0:b_1=0$ 的假设,模型有效.但是 b_1 的置信区间较长,决定系数 R^2 较小(y 的 24.38%由 x 决定),剩余方差 s^2 较大,说明模型的精度不高.

模型(1)中 y, b_0,b_1 的估计值分别记作 $\hat{y}$,$\hat{b}_0$,$\hat{b}_1$,于是新生儿体重 y 的估计为

$$\hat{y}=\hat{b}_0+\hat{b}_1x=-51.298\ 3+0.594\ 9x \tag{2}$$

$\hat{b}_1=0.594\ 9$ 是这个模型最重要的数值,它表示吸烟孕妇的怀孕期 x 增加一天,新生儿体重平均增加约 0.6(oz).当然,这里没有考虑除怀孕期外的其他因素对新生儿体重的影响.至于 $\hat{b}_0$,决不能理解为 $x=0$ 时 y 的估计值,由于模型依据的数据中吸烟孕妇怀孕期大约在 220 至 340 天(见图 1),因此 x 只能在这个范围之内对 y 作估计.比如,若怀孕期 $x=280$ 天,按(2)式计算,新生儿体重的预测值为 114.593 7(oz),还可以得到预测区间是[88.094 9,141.092 5] ($\alpha=0.05$),这个区间如此之大也是由于模型精度不高的缘故.

y 的实际值和预测值的差 $y-\hat{y}$ 是模型的残差,视为随机误差 ε 的估计值,应服从均值为 0 的正态分布.由程序可以得到残差的图形,如果发现有一些数据的残差的置信区间不含零点,可以认为这些数据偏离整体数据的变化趋势,称为**异常点**,应予以剔除.

在 480 位吸烟孕妇新生儿体重的模型中,剔除异常点后重新计算得到的结果如表 5.

表 5 剔除异常点后模型(1)的计算结果(吸烟孕妇 $n=451$)

回归系数	系数估计值	系数置信区间
b_0	−53.612 6	[−77.060 6,−30.164 5]
b_1	0.600 7	[0.516 4,0.685 0]
$R^2=0.304\ 0$ $F=196.159\ 9$ $p<0.000\ 1$ $s^2=182.785\ 2$		

比较表5和表4可以看出，虽然 b_0 和 b_1 的估计值变化不大，但置信区间变短，且 R^2 和 F 变大，s^2 减小，说明模型精度得到提高.在残差及其置信区间图上又会发现新的异常点，剔除它们后模型又会有所改进.

评注 与数据拟合只计算出模型系数的估计值不同，回归分析可以得到系数的估计值(在最小二乘准则下与数据拟合得到的数值相同)和置信区间，以及 R^2，s^2，F 等，从而对模型的有效性和精度作出判断.

由于采用最小二乘准则，异常点对残差平方和的"贡献"很大，给模型的有效性和精度带来不利影响，应注意剔除.

对于不吸烟孕妇可以建立与(1)式形式相同的另一个一元线性回归模型，用她们的新生儿体重与怀孕期的数据编程计算并剔除异常点后得到的结果表6.

表6 剔除异常点后模型(1)的计算结果(不吸烟孕妇 $n=690$)

回归系数	系数估计值	系数置信区间
b_0	33.5330	[14.9989,52.0671]
b_1	0.3201	[0.2541,0.3860]
$R^2=0.1165$ $F=90.7415$ $p<0.0001$ $s^2=180.8567$		

对表6可作与上类似的分析，值得注意的是，$\hat{b}_1=0.3201$ 表示不吸烟孕妇的怀孕期 x 增加一天，新生儿体重平均只增加约0.32(oz)，与吸烟孕妇差别很大.

能不能将吸烟状况也作为一个自变量，建立新生儿体重与孕妇怀孕期和吸烟状况2个自变量的回归模型，从而利用全体孕妇的相关数据进行计算呢?

程序文件 9-3
胎儿健康(不吸烟孕妇)的一元回归
prog0901c.m

多元线性回归分析 新生儿体重仍记作 y，孕妇怀孕期记作 x_1 吸烟状况记作 x_2，只取2个数值：$x_2=0$ 表示不吸烟，$x_2=1$ 表示吸烟.如果假定 y 与 x_1，x_2 之间的关系都是线性的，则可建立如下的多元线性回归模型：

$$y=b_0+b_1x_1+b_2x_2+\varepsilon \tag{3}$$

其中 ε 仍是服从 $N(0,\sigma^2)$ 分布、相互独立的随机变量.

根据新生儿体重和孕妇怀孕期、吸烟状况的全部数据，按照(3)式编程计算并剔除异常点，得到 b_0，b_1，b_2 的估计值(置信区间等其他结构从略)，对 y 的估计为

$$\hat{y}=\hat{b}_0+\hat{b}_1x_1+\hat{b}_2x_2=0.7698+0.4365x_1-8.7610x_2 \tag{4}$$

程序文件 9-4
胎儿健康的多元回归
prog0901d.m

这里的 $\hat{b}_2=-8.7610$ 表示，对于怀孕期 x_1 相同的孕妇，吸烟比不吸烟的新生儿体重平均约低8.8(oz)，虽然与前面参数估计的数值基本一样，但增加了怀孕期相同的条件；这个数值还补充了2个一元回归模型无法得到的、吸烟与不吸烟孕妇相互比较的结果.而 $\hat{b}_1=0.4365$ 表示的是，对于吸烟状况 x_2 相同的孕妇，怀孕期 x_1 增加一天新生儿体重平均增加约0.44(oz)，介于一元回归模型得到的、吸烟孕妇的0.6与不吸烟孕妇的0.32之间，但是这个数值却无法区分吸烟状况对怀孕期增加引起的新生儿体重的变化.

为了反映怀孕期 x_1 和吸烟状况 x_2 对新生儿体重的综合影响，需要改进模型(3)，增加二者的乘积项 x_1x_2，有

$$y=b_0+b_1x_1+b_2x_2+b_3x_1x_2+\varepsilon \tag{5}$$

这仍是多元线性回归模型，因为它对于待定系数而言是线性的.按照(5)式编程计算并剔除异常点后得到的结果如表7.

表 7 模型(5)的计算结果($n=1143$)

回归系数	系数估计值	系数置信区间
b_0	34.0925	[15.4605,52.7244]
b_1	0.3181	[0.2517,0.3844]
b_2	-87.0738	[-116.9656,-57.1820]
b_3	0.2804	[0.1734,0.3875]
$R^2=0.2766$ $F=145.1357$ $p<0.0001$ $s^2=182.9536$		

评注 多元回归模型中某一个自变量系数估计值的含义,是在其余自变量固定的条件下,所引起的因变量的变化率.因此,对于自变量集合不同的模型,同一个自变量系数估计值的含义并不相同,如模型(1)和(2)中的 b_1.

由表 7,对 y 的估计可表示为

$$\hat{y}=34.0925+0.3181x_1-87.0738\ x_2+0.2804x_1x_2 \tag{6}$$

对不吸烟孕妇和吸烟孕妇,分别以 $x_2=0$ 和 $x_2=1$ 代入(6)式,得

$$\hat{y}=34.0925+0.3181x_1,\quad \hat{y}=-52.9813+0.5985x_1 \tag{7}$$

(7)式中的系数与只含 x_1 的一元模型得到的表 6 和表 5 的数值非常接近,(6)式所表示的多元线性回归模型综合描述了 x_1,x_2 两个自变量与 y 的关系.

表 7 的结果虽然可以验证模型的有效性,但是 R^2 较小,s^2 较大,仍有改进余地.

变量选择与逐步回归 在 CHDS 提供的数据中,除了新生儿体重以及孕妇怀孕期和吸烟状况外,还有孕妇怀孕时的年龄、体重、身高和胎次状况,在以新生儿体重为因变量的模型中,这些因素中是否有几个也应该作为自变量加入呢?这是建立多元线性回归模型时经常遇到的问题.

一般地说,实际问题中影响因变量的因素可能很多,从应用的角度既希望将所有对因变量影响显著的自变量都纳入回归模型,又希望最终的模型尽量简单,即不包含那些对因变量影响不显著的自变量,这就是所谓"变量选择".逐步回归是一种迭代式的变量选择方法,下面利用 CHDS 数据提供的全部信息,通过逐步回归来选择变量,建立新生儿体重的多元线性回归模型.

用 y 表示新生儿体重,x_1, x_2, x_3, x_4, x_5, x_6 依次表示孕妇怀孕期、胎次状况、年龄、身高、体重、吸烟状况(与表 1 的顺序相同),组成候选变量集合.根据前面的结果认为孕妇的怀孕期和吸烟状况是最重要的变量,故选取 x_1, x_6 为初始子集.用 MATLAB 的逐步回归程序计算,显著性水平取缺省值(引入变量 $\alpha=0.05$,移出变量 $\alpha=0.10$).



程序文件 9-5
胎儿健康的逐步回归
prog0901e.m

程序运行的第 1 个输出图形如图 2.

图左上部的窗口用圆点和线段显示各个候选变量的回归系数估计值及其置信区间,屏幕上蓝色表示在模型中的变量(图 2 中也是蓝色),红色表示不在模型中的变量(图 2 中是黑色),粗线段和细线段分别表示 90%和 95%的置信区间.图右上部的窗口中列出一个统计表,包括各个候选变量的回归系数估计值、t 统计量和 p 概率值(图 2 中蓝字、黑字的含义同上).图的中部是当前模型的输出:常数项 b_0 的估计值(intercept),决定系数 R^2,F 值,剩余标准差 s(RMSE),p 值等.图下部的窗口用圆点记录每次选取模型的剩余标准差 s,供比较用,当鼠标指向某个圆点时,就显示那个模型包含的变量,若点击圆点,则重现那个模型的结果.

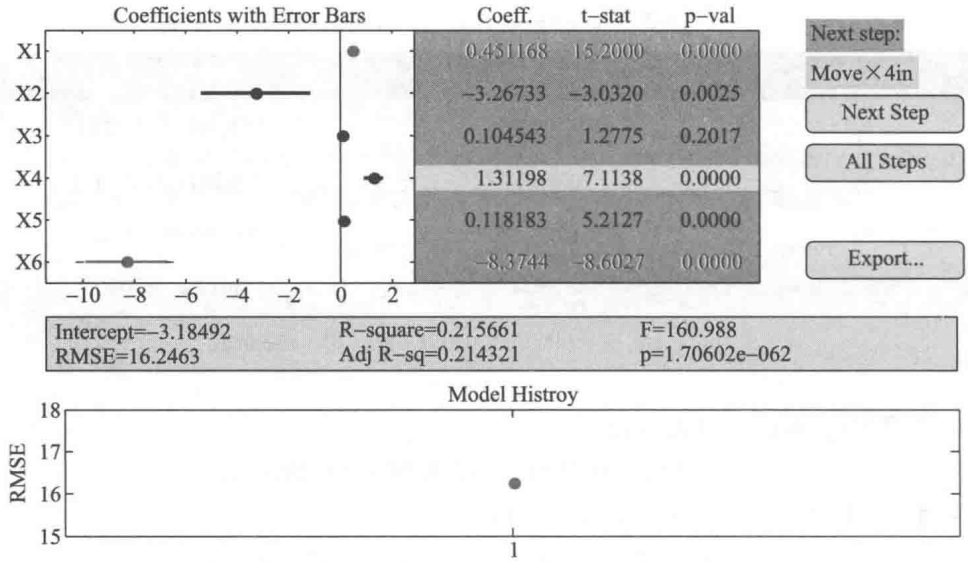


图 2 MATLAB 逐步回归程序运行的第 1 个输出图形

当用鼠标点击左上部的圆点或右上部的字段时，可使之由红变蓝，该变量被引入；或者由蓝变红，该变量被移出.通常用剩余标准差 s 达到最小作为确定最终模型的标准，点击图右侧的按钮 Next Step，可以按照 s 减少的大小自动地引入或移出变量，直至出现 Move no terms 的提示，结束运行.如果点击按钮 All Steps，则立即得到最终结果.

新生儿体重逐步回归程序的最终输出图形如图 3（蓝色、黑色的含义同图 2），包含 x_3 以外的其余 5 个自变量.

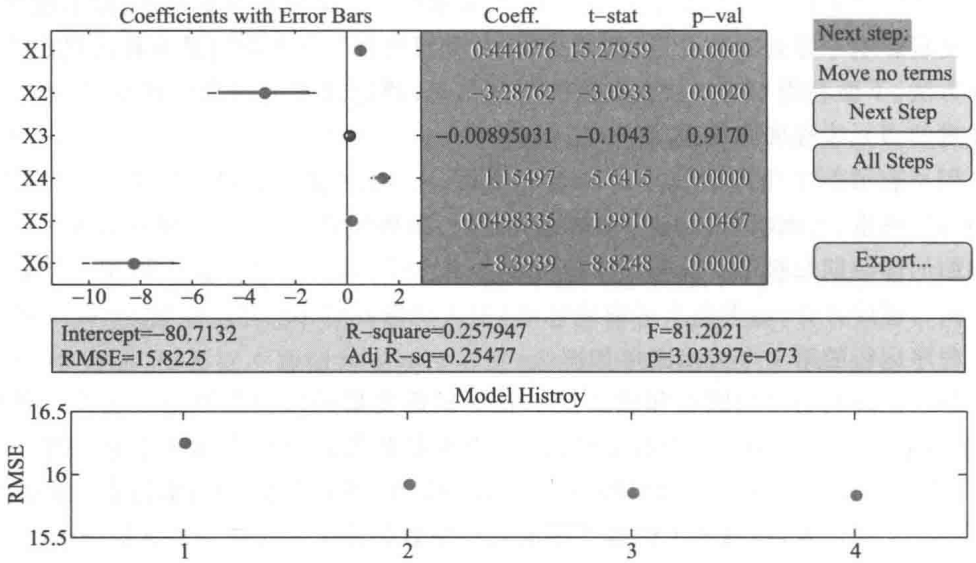


图 3 逐步回归程序运行的最终输出图形

通过图 3 右侧的 Export 菜单可以传送输出数据，得到的结果如表 8.

表 8 逐步回归的最终结果($n=1174$)

回归系数	系数估计值	系数置信区间
b_0	-80.713 2	
b_1	0.444 1	[0.387 0,0.501 1]
b_2	-3.287 6	[-5.372 9,-1.202 4]
b_4	1.155 0	[0.753 3,1.556 6]
b_5	0.049 8	[0.000 7,0.098 9]
b_6	-8.393 9	[-10.260 1,-6.527 7]
$R^2=0.257\ 9$ $F=81.202\ 1$ $p<0.000\ 1$ $s=15.822\ 5$		

评注 当遇到影响因变量的因素众多、又难以辨识哪些因素最为重要的实际问题时，采用逐步回归来选择自变量不失为一种有效的建立多元线性回归模型的方法.除了原始的自变量，变量的平方、两个变量的乘积等(如本节中孕妇怀孕期 x_1 和吸烟状况 x_6 的乘积项 x_1x_6)也可以作为新的变量供模型选择.同时，人机交互式的逐步回归程序更大大提高了建模的“效率”.

最终模型为

$$\hat{y}=-80.713\ 2+0.444\ 1x_1-3.287\ 6x_2+1.155\ 0x_4+0.049\ 8x_5-8.393\ 9x_6 \tag{8}$$

需要指出，由于残差的置信区间无法通过 MATLAB 的逐步回归程序输出，所以未剔除异常点，致使表 8 中的 R^2，s 等与表 7 相比没有什么改进.读者可以对这里选出的变量建立多元线性回归模型，并剔除异常点，看看结果如何.

结果分析 用逐步回归方法得到的新生儿体重模型(表 8 和(8)式)包含孕妇的怀孕期、胎次状况、体重和吸烟状况等自变量，它们的回归系数置信区间不含零点，可以认为对新生儿体重的影响都是显著的，回归系数的数值给出了定量关系.如 $b_6=-8.393\ 9$ 说明当孕妇的怀孕期、胎次状况、身高、体重相同时，吸烟孕妇比不吸烟孕妇的新生儿体重约低 8.4 oz，可与前面的结果(表 2 和(4)式)相比.b_1，b_4，b_5 均为正值表明孕妇的怀孕期、身高、体重对新生儿体重的影响都是正面的，$b_2=-3.287\ 6$ 则表示第 1 胎新生儿体重比非第 1 胎约低 3.3 oz.

在这个问题中，不仅因变量与每个自变量有一定的相关关系，而且许多自变量两两之间也可能存在相关关系.多元随机变量的相关系数矩阵用于描述相关性的正反和强弱，即相关系数的正或负表示正相关或负相关，相关系数绝对值越大相关性越强.y，x_1，x_2，x_3，x_4，x_5，x_6 的相关系数矩阵如表 9.

表 9 $y,x_1,x_2,x_3,x_4,x_5,x_6$ 的相关系数矩阵(对称矩阵只给出上三角元素)

	y	x_1	x_2	x_3	x_4	x_5	x_6
y	1.000 0	0.407 5	-0.043 9	0.027 0	0.203 7	0.155 9	-0.246 8
x_1		1.000 0	0.080 9	-0.053 4	0.070 5	0.023 7	-0.060 3
x_2			1.000 0	-0.351 0	0.043 5	-0.096 4	-0.009 6
x_3				1.000 0	-0.006 5	0.147 3	-0.067 8
x_4					1.000 0	0.435 3	0.017 5
x_5						1.000 0	-0.060 3
x_6							1.000 0

从表9相关系数的数值可以看出，与新生儿体重 y 相关性较强（相关系数大于0.2）的依次是孕妇的怀孕期 x_1、吸烟状况 x_6、身高 x_4.各自变量之间相关性较强的有：孕妇体重 x_5 与身高 x_4 的正相关，这是自然的；孕妇年龄 x_3 与胎次状况 x_2 的负相关，大概因为年龄越大第1胎越少.

评注 当回归模型的自变量之间存在较强的（线性）相关时，求解将出现病态，称为多重共线性，是应该避免的现象，可以通过相关系数矩阵进行筛查.

如果几个自变量之间存在较强的相关性，那么删除多余的只保留一个变量不会对模型的有效性和精确度有多大影响，比如可以去掉孕妇体重 x_5 与身高 x_4 两个变量中任何一个.

表9中孕妇怀孕期 x_1 与吸烟状况 x_6 的相关系数很小，表明二者几乎无关，这与假设检验的结果（表2）一致.

不同年龄段孕妇吸烟对新生儿体重的影响 按照孕妇的年龄将数据分为4组：小于25岁，25~30岁，30~35岁，大于35岁.对每组建立 y 与 x_1，x_2，x_4，x_5，x_6 的回归模型，得到的结果如表10.

表10 按孕妇年龄分组的计算结果

	小于25岁	25~30岁	30~35岁	大于35岁
b_0	-66.389 3	-39.129 6	-157.130 7	-130.174 0
b_1	0.397 2	0.352 1	0.595 1	0.672 8
b_2	-0.997 8	-7.412 4	-0.093 2	-4.183 5
b_4	1.214 4	0.840 9	1.682 8	0.874 7
b_5	-0.002 1	0.095 9	0.055 7	0.073 2
b_6	-8.411 9	-8.265 6	-10.541 1	-6.400 8
R^2	0.254 9	0.233 0	0.339 4	0.313 6
s^2	211.635 9	239.720 1	272.602 1	304.720 8
n	444	362	211	157

评注 本节通过孕妇吸烟与胎儿健康案例的讨论，介绍了利用一元线性回归、多元线性回归和逐步回归方法建立模型、分析结果的全过程.线性回归是回归分析方法的基础，应用也最广泛，希望读者不仅会用软件求解，而且能从应用的角度对它的基本原理和方法有一定的了解.可参考基础知识9-2.

比较表10中各年龄组 b_1 和 b_6 的数值发现，对于孕妇怀孕期 x_1 和吸烟状况 x_6 这两个影响新生儿体重的主要因素，30岁以下的两组结果差别不大，而与30岁以上的两组则有一定的差异.至于这些差别是否具有显著意义，则需要通过假设检验或方差分析作进一步研究.

复习题

1. 对逐步回归选出的变量 x_1，x_2，x_4，x_5，x_6，建立 y 的多元线性回归模型，并剔除异常点，与逐步回归的结果（表8）比较.

2. 在候选变量集合中增加乘积项 x_1x_6，重作逐步回归，与原来的结果（表8）比较，解释增加 x_1x_6 后各个系数的含义.

9.2 软件开发人员的薪金

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，包括薪金（单位：元）、资历（从事专业工作的年数）、管理责任（1 表示管理人员，0 表示非管理人员）、教育程度（1 表示中学，2 表示大学，3 表示研究生），全部数据见数据文件 9-2.

表 1 软件开发人员的薪金与资历、管理责任、教育程度

编号	薪金	资历	管理	教育
1	13 876	1	1	1
2	11 608	1	0	3
⋮	⋮	⋮	⋮	⋮
46	19 346	20	0	1



数据文件 9-2
软件开发人员的薪金
data0902.txt

分析与假设　按照常识，薪金自然随着资历的增长而增加，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，资历记作 x_1，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

$$x_3=\begin{cases}1, \text{中学}\\ 0, \text{其他}\end{cases}\qquad x_4=\begin{cases}1, \text{大学}\\ 0, \text{其他}\end{cases}$$

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

$$y=a_0+a_1x_1+a_2x_2+a_3x_3+a_4x_4+\varepsilon \tag{1}$$

其中 a_0，a_1，…，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]



程序文件 9-6
软件开发人员薪金基本模型
prog0902a.m

续表

参数	参数估计值	参数置信区间
a_2	6 883	[6 248, 7 517]
a_3	-2 994	[-3 826, -2 162]
a_4	148	[-636, 931]
	$R^2=0.957$ $F=226$ $p<0.000\ 1$ $s^2=1.057\times10^6$	

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

$$\hat{y}=\hat{a}_0+\hat{a}_1\times2+\hat{a}_2\times0+\hat{a}_3\times0+\hat{a}_4\times1=12\ 272$$

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下才成立的.

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.

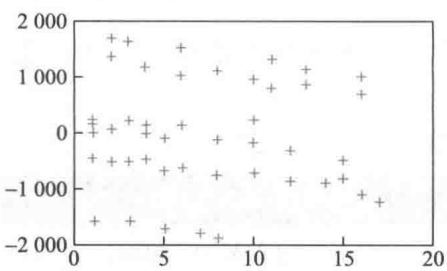


图1 模型(1)ε 与 x_1 的关系

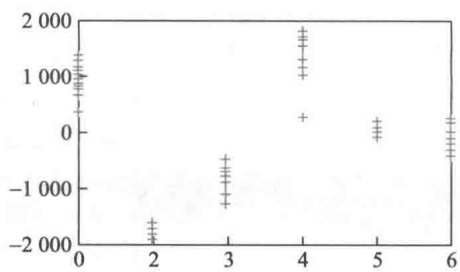


图2 模型(1)ε 与 x_2-x_3,x_4 组合的关系

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

$$y=a_0+a_1x_1+a_2x_2+a_3x_3+a_4x_4+a_5x_2x_3+a_6x_2x_4+\varepsilon \tag{2}$$

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]
$R^2=0.998\ 8$　$F=5\ 545$　$p<0.000\ 1$　$s^2=3.004\ 7\times10^4$		



程序文件 9-7
软件开发人员薪金增加交互项
prog0902b.m

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.

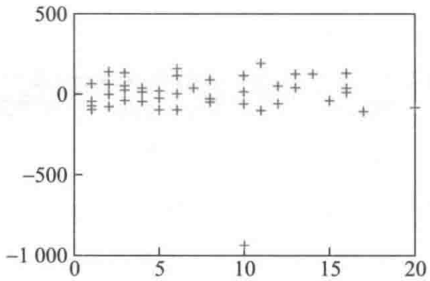


图 3　模型(2)ε 与 x_1 的关系

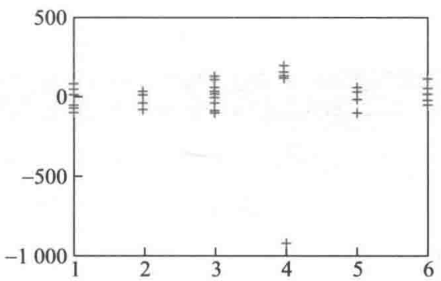


图 4　模型(2)ε 与 x_2-x_3, x_4 组合的关系

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(从数据文件 9-2 可以查出是 33 号),他的实际薪金明显低于模型的估计值,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.



程序文件 9-8
软件开发人员薪金去除异常点
prog0902c.m

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]
$R^2=0.999\ 8$ $F=36\ 701$ $p<0.000\ 1$ $s^2=4.347\times10^3$		

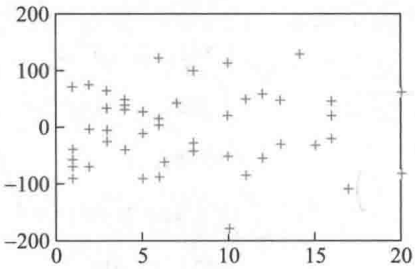


图5 模型(2)去掉异常数据后 ε 与 x_1 的关系

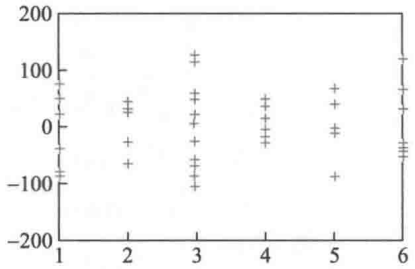


图6 模型(2)去掉异常数据后 ε 与 x_2-x_3,x_4 组合的关系

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.

评注 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).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.实际上,可以直接对6种管理-教育组合引入5个0-1变量.

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.

复习题

1. 用逐步回归重新建立模型,与表 5 的结果比较.

2. 汽车销售商认为汽车销售量与汽油价格、贷款利率有关,两种类型汽车(普通型和豪华型)18个月的调查资料如下表,其中 y_1 是普通型汽车销售量(千辆),y_2 是豪华型汽车销售量(千辆),x_1 是汽油价格(美元/加仑),x_2 是贷款利率(%)[72],全部数据见数据文件 9-3.

序号	y_1	y_2	x_1	x_2
1	22.1	7.2	1.89	6.1
2	15.4	5.4	1.94	6.2
⋮	⋮	⋮	⋮	⋮
18	44.3	15.6	1.68	2.3

(1) 对普通型和豪华型汽车分别建立如下模型:

$$y_1=\beta_0^{(1)}+\beta_1^{(1)}x_1+\beta_2^{(1)}x_2,\quad y_2=\beta_0^{(2)}+\beta_1^{(2)}x_1+\beta_2^{(2)}x_2$$

给出 β 的估计值和置信区间,决定系数 R^2,F 值及剩余方差等.

(2) 用 $x_3=0$, 1 表示汽车类型,建立统一模型:$y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_3$,给出 β 的估计值和置信区间,决定系数 R^2,F 值及剩余方差等.以 $x_3=0$, 1 代入统一模型,将结果与(1)的两个模型的结果比较,解释二者的区别.

(3) 对统一模型就每种类型汽车分别作 x_1 和 x_2 与残差的散点图,有什么现象,说明模型有何缺陷?

(4) 对统一模型增加二次项和交互项,考察结果有什么改进.



数据文件 9-3
9.2 节复习题 2 汽车销售量

9.3 酶促反应

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 $10^3\sim10^{17}$ 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[6].

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/ppm①		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	—

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 $y=f(x,\boldsymbol{\beta})$，其中 $\boldsymbol{\beta}$ 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

$$y=f(x,\boldsymbol{\beta})=\frac{\beta_1 x}{\beta_2+x} \tag{1}$$

指数增长模型

$$y=f(x,\boldsymbol{\beta})=\beta_1(1-\mathrm{e}^{-\beta_2 x}) \tag{2}$$

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大致符合的.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(复习题).

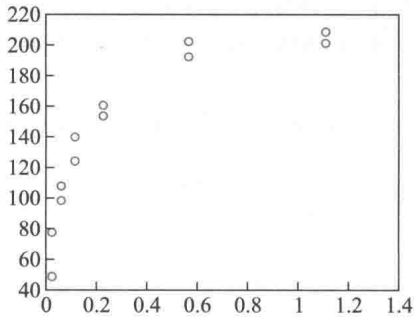


图1 y 对 x(经处理)的散点图

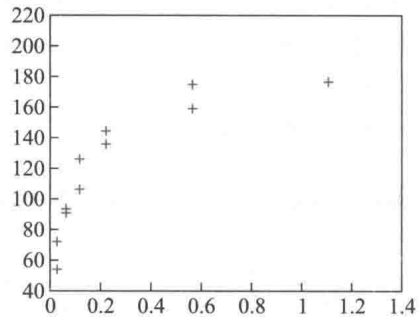


图2 y 对 x(未经处理)的散点图

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 $\boldsymbol{\beta}=(\beta_1,\beta_2)$ 是非线性的，但是可以通过下面的变量代换化为线性模型

$$\frac{1}{y}=\frac{1}{\beta_1}+\frac{\beta_2}{\beta_1}\frac{1}{x}=\theta_1+\theta_2 u \tag{3}$$

模型(3)中的因变量 $1/y$ 对新的参数 $\boldsymbol{\theta}=(\theta_1,\theta_2)$ 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 1 ppm = 0.001‰.

$1/x$ 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.

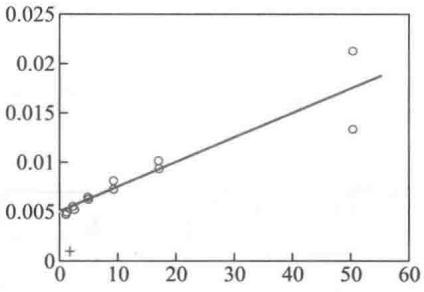


图 3 $1/y$ 与 $1/x$ 的散点图和回归直线

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 $\hat{\beta}_1=195.802\ 0$ 和 $\hat{\beta}_2=0.048\ 4$.


程序文件 9-9
酶促反应线性化模型和非线性模型
prog0903a.m


程序文件 9-10
酶促反应非线性模型子程序
prog0903a1.m

表 2 线性化模型(3)参数的估计结果

参数	参数估计值($\times10^{-3}$)	参数置信区间($\times10^{-3}$)
θ_1	5.107 2	[3.538 6, 6.675 8]
θ_2	0.247 2	[0.175 7, 0.318 8]
$R^2=0.855\ 7$	$F=59.297\ 5$ $p<0.000\ 1$	$s^2=3.580\ 6\times10^{-6}$

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.

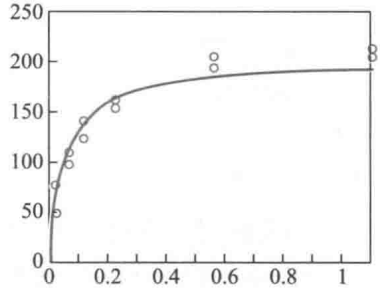


图 4 用线性化得到的原始数据拟合图

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).

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 统计工具箱进行,得到的数值结果见表 3.

表 3 非线性模型(1)参数的估计结果

参数	参数估计值	参数置信区间
β_1	212.683 7	[197.204 5, 228.162 9]
β_2	0.064 1	[0.045 7, 0.082 6]

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 $\hat{\beta}_1=212.683\ 7$.还容易得到,反应的"半速度

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,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.

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.

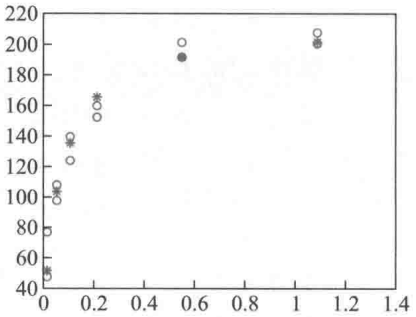


图5 模型(1)的预测图

(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 $\hat{\beta}_2=0.064\ 1$.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.

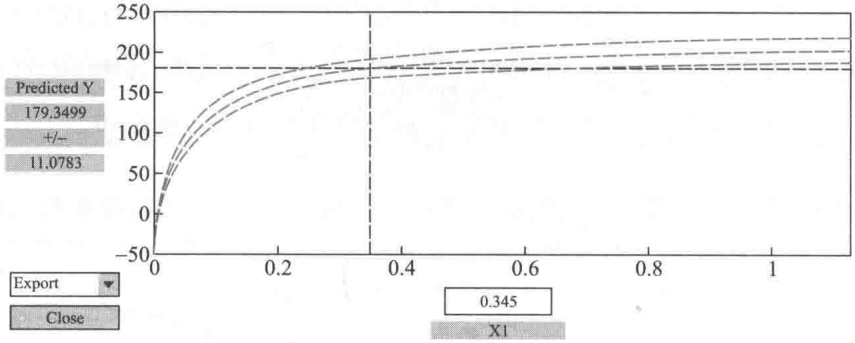


图6 模型(1)的预测及结果输出

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 $\boldsymbol{\beta}$ 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

$$y=f(\boldsymbol{x},\boldsymbol{\beta})=\frac{(\beta_1+\gamma_1 x_2)x_1}{(\beta_2+\gamma_2 x_2)+x_1} \tag{4}$$

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统计工具箱来计算模型(4)的回归系数 β_1,β_2,γ_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)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),及预测和结果输出图(图 8),模型的剩余标准差 $s=10.400\ 0$.



程序文件 9-11
酶促反应混合模型
prog0903b.m

表 4 非线性模型(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

$$y=f(\boldsymbol{x},\boldsymbol{\beta})=\frac{(\beta_1+\gamma_1 x_2)x_1}{\beta_2+x_1} \tag{5}$$

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 9、图 10 中,模型(5)的剩余标准差 $s=10.585\ 1$.

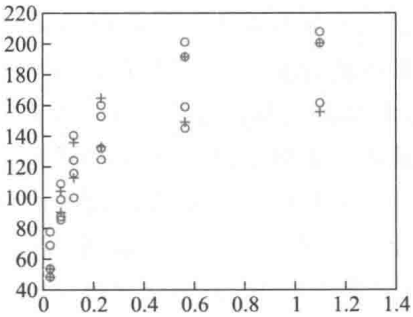


图 7 模型(4)的预测图

(o——原始数据;+——拟合结果)

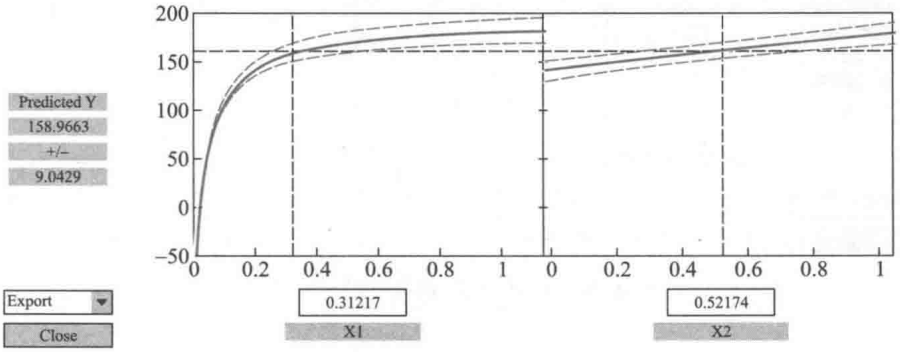


图 8 模型(4)的预测及结果输出

程序文件 9-12
酶促反应混合模型子程序 1
prog0903b1.m



程序文件 9-13
酶促反应混合模型子程序 2
prog0903b2.m

表 5 非线性模型(5)参数的估计结果

参数	参数估计值	参数置信区间
β_1	166.602 5	[154.488 6, 178.716 4]
β_2	0.058 0	[0.045 6, 0.070 3]
γ_1	42.025 2	[28.941 9, 55.108 5]

表 6 模型(4)与(5)预测值与预测区间的比较(预测区间为预测值±Δ,全部数据见运行程序文件 9-11 后的输出结果)

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	47.344 3	9.207 8	42.735 8	5.444 6
51	47.344 3	9.207 8	42.735 8	5.444 6
⋮	⋮	⋮	⋮	⋮
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.

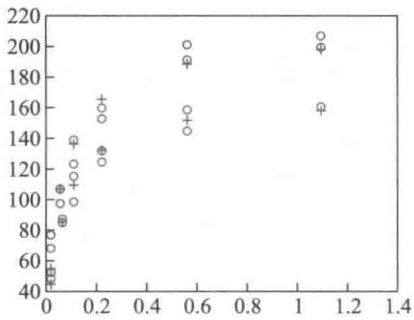


图 9 模型(5)的预测图

(○——原始数据;+——拟合结果)

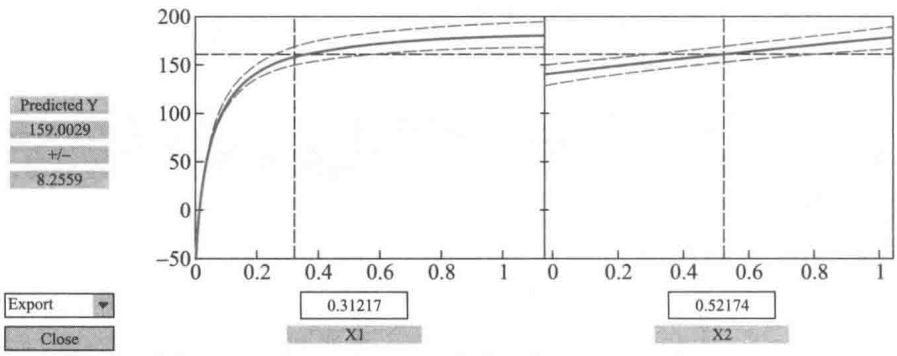


图 10 模型(5)的预测及结果输出

评注 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.

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

$$y=f(x,\boldsymbol{\beta})=\frac{\beta_1 x}{\beta_2+x+\beta_3 x^2} \tag{6}$$

或引入混合模型

$$y=f(\boldsymbol{x},\boldsymbol{\beta})=\frac{(\beta_1+\gamma_1 x_2)x_1}{(\beta_2+\gamma_2 x_2)+x_1+(\beta_3+\gamma_3 x_2)x_2^2} \tag{7}$$

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)的残差图中表现出来的在各个浓度下残差散布不均匀的现象.

复习题

在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.

9.4 投资额与生产总值和物价指数

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[47,65].

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)，$t=1,2,\cdots,n(=20)$.因变量 y_t 与自变量 x_{1t} 和 x_{2t} 的散点图见图 1 和图 2.

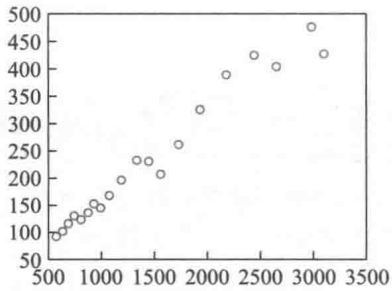


图1 y_t 对 x_{1t} 的散点图

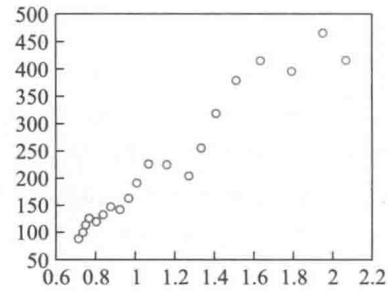


图2 y_t 对 x_{2t} 的散点图

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

$$y_t=\beta_0+\beta_1 x_{1t}+\beta_2 x_{2t}+\varepsilon_t \tag{1}$$

程序文件 9-14
投资额与生产总值和物价指数
prog0904.m

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，$t=1,2,\cdots,n$. 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

$$\hat{y}_t=322.725+0.618\ 5x_{1t}-859.479x_{2t} \tag{2}$$

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 $s=12.716\ 4$.

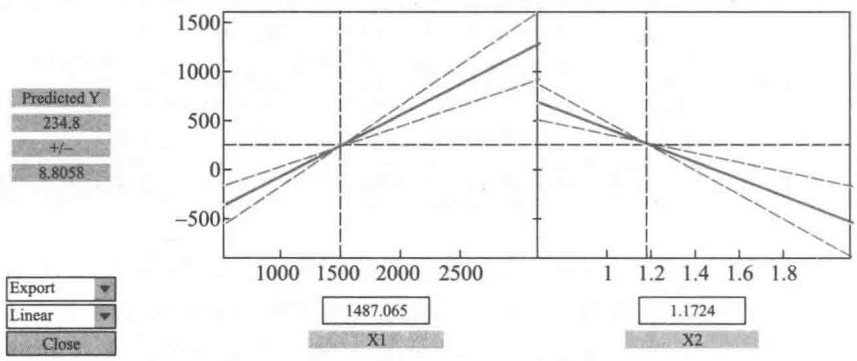


图3 模型(2)的输出

自相关性诊断与处理方法 从表面上看得到的基本模型(2)的拟合度非常之高($R^2=0.990\ 8$),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	e_t	t	e_t	t	e_t	t	e_t
1	15.130 6	6	-20.171 0	11	-4.346 8	16	18.425 0
2	5.728 1	7	-11.306 2	12	8.072 9	17	9.531 1
3	2.468 2	8	-6.473 3	13	6.400 6	18	-14.934 9
4	-4.842 1	9	2.411 9	14	10.101 0	19	2.008 5
5	-14.567 7	10	-1.673 7	15	18.690 0	20	-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:

$$y_t=\beta_0+\beta_1x_{1t}+\beta_2x_{2t}+\varepsilon_t,\quad \varepsilon_t=\rho\varepsilon_{t-1}+u_t \tag{3}$$

其中 ρ 是自相关系数,$|\rho|\leqslant 1$,u_t 相互独立且服从均值为零的正态分布,$t=1,2,\cdots,n$.

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.

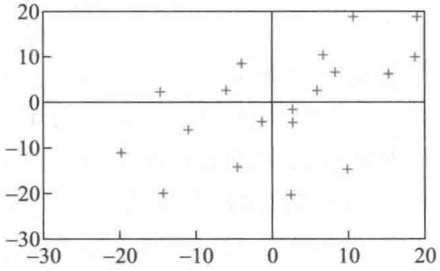


图4 模型(2)e_t-e_{t-1}的散点图

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:

$$DW=\frac{\sum_{t=2}^{n}(e_t-e_{t-1})^2}{\sum_{t=2}^{n}e_t^2} \tag{4}$$

经过简单的运算可知,当 n 较大时,

$$DW \approx 2\left(1-\frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=2}^{n} e_t^2}\right) \tag{5}$$

而(5)式右端的 $\sum_{t=2}^{n} e_t e_{t-1} \Big/ \sum_{t=2}^{n} e_t^2$ 正是自相关系数 ρ 的估计值 $\hat{\rho}$，于是

$$DW \approx 2(1-\hat{\rho}) \tag{6}$$

由于 $-1 \leqslant \hat{\rho} \leqslant 1$，所以 $0 \leqslant DW \leqslant 4$，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[65].

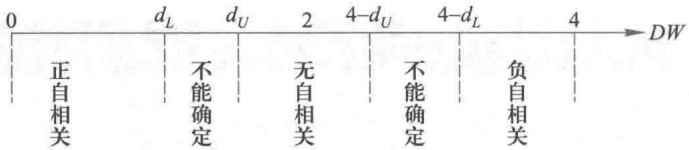


图5 与 DW 值对应的自相关状态

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

$$\hat{\rho} = 1 - \frac{DW}{2} \tag{7}$$

作变换

$$y_t^* = y_t - \rho y_{t-1}, \quad x_{it}^* = x_{it} - \rho x_{i,t-1}, \quad i = 1,2 \tag{8}$$

则模型(3)化为

$$y_t^* = \beta_0^* + \beta_1 x_{1t}^* + \beta_2 x_{2t}^* + u_t, \quad \beta_0^* = \beta_0(1-\rho) \tag{9}$$

其中 u_t 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.

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.

加入自相关后的模型 利用表3给出的残差 e_t，根据(4)式计算出 $DW = 0.875\,4$，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[65]，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 $\hat{\rho} = 0.562\,3$.

以 ρ 的估计值代入(8)式作变换，利用变换后的数据 y_t^*，x_{1t}^*，x_{2t}^* 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_t 是否还存在自相关.从模型(9)的残差可以计算出 $DW=1.575\ 1$,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_t^*,x_{1t}^*,x_{2t}^* 还原为原始变量 y_t,x_{1t},x_{2t},得到的结果为

$$\hat{y}_t = 163.490\ 5+0.562\ 3y_{t-1}+0.699x_{1t}-0.393\ 0x_{1,t-1}-1\ 009.033\ 3x_{2t}+567.379\ 4x_{2,t-1} \tag{10}$$

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 12.716 4 减小到 9.827 7.我们将模型(10)、模型(2)的计算值 $\hat{y}_t$ 与实际数据 y_t 的比较,以及两个模型的残差 e_t,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(全部数据见运行程序文件 9-14 后的输出结果)

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	98.468 2	91.671 9	-1.068 2	5.728 1
3	113.500 0	113.559 6	111.031 8	0.059 6	2.468 2
⋮	⋮	⋮	⋮	⋮	⋮
20	424.500 0	438.188 6	445.152 1	-13.688 6	-20.652 1

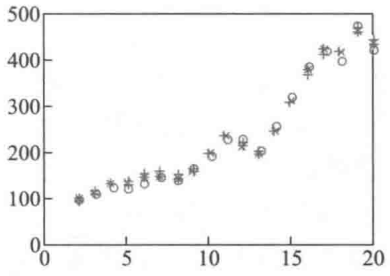


图 6 模型(10),(2)的预测图

(○——y_t;*——$\hat{y}_t$(10);+——$\hat{y}_t$(2))

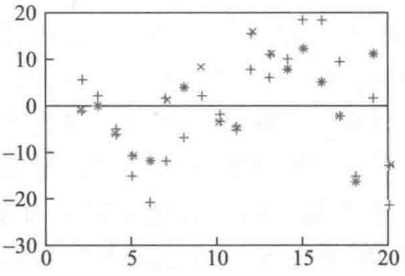


图 7 模型(10),(2)的残差图

(*——e_t(10);+——e_t(2))

评注 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.

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.

复习题

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[47].

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		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		3	19	28.24	168.7
	2	10	24.01	145.3		4	20	28.78	171.7

9.5 冠心病与年龄

冠心病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部分数据(其中冠心病一行中,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[27].



程序文件 9-15
冠心病与年龄
prog0905.m

表 1 100 名被观察者的年龄与是否患冠心病的部分数据(全部数据见程序文件 9-15)

序号	1	2	…	100
年龄	20	23	…	69
冠心病	0	0	…	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).

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

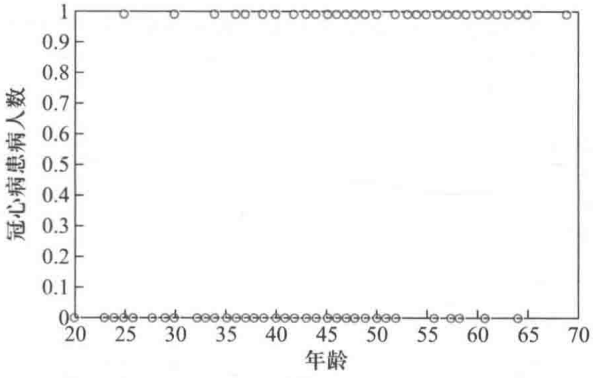


图 1 冠心病患病人数与年龄的散点图

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 数据作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).

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 形曲线,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

$$y = E(Y \mid x) \tag{1}$$

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 形曲线的 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

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \varepsilon \tag{2}$$

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.

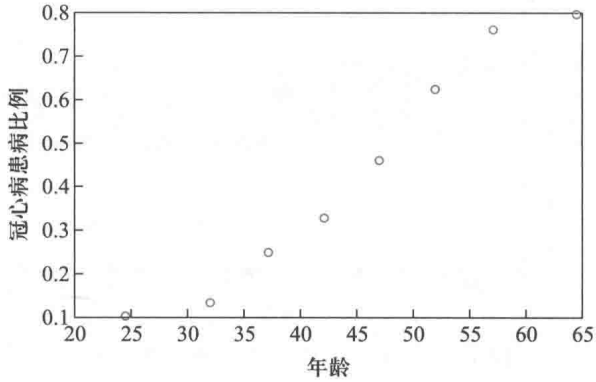


图 2 冠心病患病比例对年龄段中点的散点图

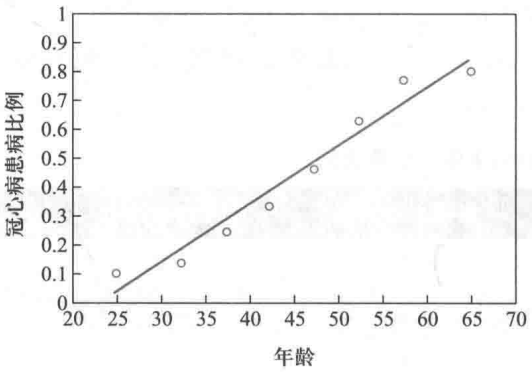


图 3 y 对 x 的线性回归曲线

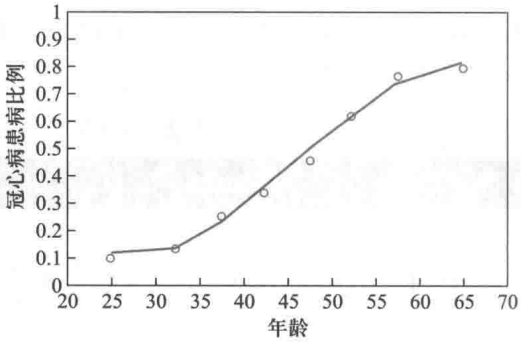


图 4 y 对 x 的三次多项式回归曲线

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，用前几节介绍的普通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

$$\pi(x)=P(Y=1\mid x) \tag{3}$$

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 形曲线的函数关系，并注意到 $\pi(x)$ 在[0,1]区间取值，可以建立如在第 5 章多次用到的 logistic 模型(如 5.1 节(9)式的变形)

$$\pi(x)=\frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}} \tag{4}$$

(4)的反函数写作

$$\ln\frac{\pi(x)}{1-\pi(x)}=\beta_0+\beta_1 x \tag{5}$$

(5)式左端可看作 $\pi(x)$ 的变换,记作 $\text{logit}(\pi(x)) = \ln \dfrac{\pi(x)}{1-\pi(x)}$,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($i=1,2,\cdots,k$)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:

$$\text{logit}(\pi_i) = \ln \frac{\pi_i}{1-\pi_i} = \beta_0 + \beta_1 x_i \tag{6}$$

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[27].

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 的数据编程得到 logit 模型中的参数 β_0, β_1 的最大似然估计值分别为 -5.038 2和 0.105 0.图 5 给出了 logistic 回归曲线与散点图.

表 3 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.

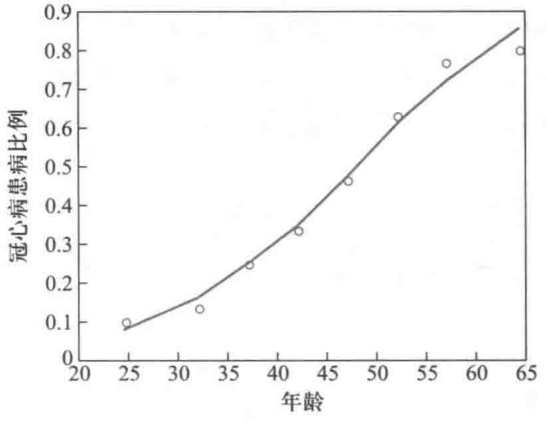


图 5 logistic 回归曲线与散点图

① 当 $m_i=0$ 或 $m_i=n_i$ 时,可用修正公式 $\pi_i = \dfrac{m_i+0.5}{n_i+1}$ 来计算.

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

$$\text{logit}(\pi(x)) = \ln\frac{\pi(x)}{1-\pi(x)} = \beta_0 + \beta_1 x + \beta_2 x^2 \tag{7}$$

计算得到 β_0,β_1,β_2 的估计值分别为 -5.350 6,0.119 4 和 -0.000 2,其中 β_2 的 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

$$\pi(x) = \Phi(\beta_0 + \beta_1 x) \tag{8}$$

$$\text{probit}(\pi(x)) = \Phi^{-1}(\pi(x)) = \beta_0 + \beta_1 x \tag{9}$$

其中 Φ 是正态概率分布函数,它也是 S 形曲线.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.y 的预测值及置信区间与 logit 模型的结果比较见表 3,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 3 probit 模型与 logit 模型预测值的比较

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]

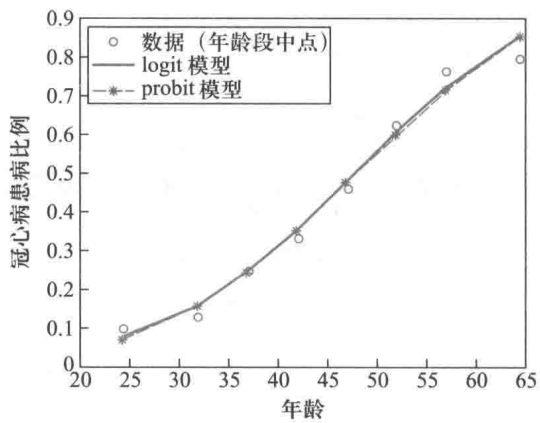


图 6 probit 模型与 logit 模型的拟合比较

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

$$\operatorname{logit}(\hat{\pi}(x)) = \ln \frac{\hat{\pi}(x)}{1-\hat{\pi}(x)} = -5.038\,2 + 0.105\,0x \tag{10}$$

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.

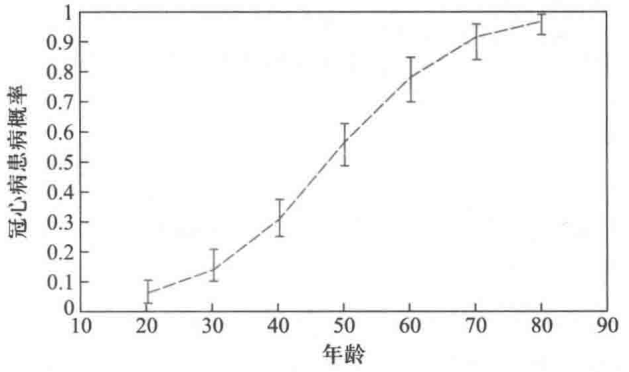


图 7　不同年龄的人患冠心病的概率的预测(竖线为置信区间)

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 odds(x)为年龄 x 的人患与不患冠心病的概率之比，则

$$\operatorname{odds}(x) = \frac{\pi(x)}{1-\pi(x)} \tag{11}$$

于是 logit 模型可以表示为

$$\operatorname{odds}(x) = e^{\beta_0+\beta_1 x} \tag{12}$$

当年龄增加 1 岁时 odds 比(发生比率或优势比)为

$$\frac{\operatorname{odds}(x+1)}{\operatorname{odds}(x)} = \frac{e^{\beta_0+\beta_1(x+1)}}{e^{\beta_0+\beta_1 x}} = e^{\beta_1} \tag{13}$$

于是

$$\beta_1 = \ln \frac{\operatorname{odds}(x+1)}{\operatorname{odds}(x)} \tag{14}$$

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

$$\operatorname{odds}(x+k) = e^{k\beta_1}\operatorname{odds}(x) \tag{15}$$

在模型(10)中 $\beta_1 = 0.105\,0$，可以算出一个 20 岁的青年人患冠心病的概率仅为 $\hat{\pi}(20) = 0.050\,3$，且发生比(患与不患冠心病的概率之比)为 $\operatorname{odds}(20) = 0.059\,3$，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 $\operatorname{odds}(30) = e^{10\times\beta_1}\times 0.059\,3 = 0.169\,5$，发生比(可解释作危险

率)增大到20岁时的2.857 7倍,而到60岁时,$\text{odds}(60)=e^{40\times\beta_1}\times 0.059\ 3=3.954\ 5$,危险率是20岁的$e^{40\times\beta_1}=66.686\ 3$倍.可见回归系数$\beta_1$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!

上面只涉及因变量是0-1变量且只有一个自变量的情形,对多个自变量$x_1,\cdots,x_m$的情形,logit模型和probit模型的一般形式分别为

$$\text{logit}(\pi(\boldsymbol{x}))=\ln\frac{\pi(\boldsymbol{x})}{1-\pi(\boldsymbol{x})}=\beta_0+\sum_{i=1}^{m}\beta_i x_i \tag{16}$$

$$\text{probit}(\pi(\boldsymbol{x}))=\Phi^{-1}(\pi(\boldsymbol{x}))=\beta_0+\sum_{i=1}^{m}\beta_i x_i \tag{17}$$

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.

其中$x_1,\cdots,x_m$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[27].

复习题

将程序文件9-15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.

9.6 蠓虫分类判别

生物学家Grogan和Wirth曾就两种蠓虫Af和Apf的鉴别问题进行研究,他们根据蠓虫的触角长和翅长对蠓虫进行了分类,这种分类方法在生物学上是常用的.研究中他们对已经确定类别的6个Apf蠓虫与9个Af蠓虫的触角长和翅长分别进行了测量,结果见表1.

表1 蠓虫的触角长和翅长原始数据

Apf蠓虫样本		Af蠓虫样本		待判样本x	
触角长	翅长	触角长	翅长	触角长	翅长
1.14	1.78	1.24	1.72	1.24	1.80
1.18	1.96	1.36	1.74	1.29	1.81
1.20	1.86	1.38	1.64	1.43	2.03

续表

Apf 蠓虫样本		Af 蠓虫样本		待判样本 x	
触角长	翅长	触角长	翅长	触角长	翅长
1.26	2.00	1.38	1.82		
1.28	2.00	1.38	1.90		
1.30	1.96	1.40	1.70		
		1.48	1.82		
		1.54	1.82		
		1.56	2.08		

我们的任务是利用这份数据来建立一种区分两种蠓虫的模型，用于对已知触角长和翅长的待判蠓虫样本进行识别，并着重讨论以下问题：

- 根据表 1 中给出的 Apf 蠓虫和 Af 蠓虫样本数据建立模型，以便正确区分这两类蠓虫；
- 用建立的模型对表 1 中已知触角长和翅长的 3 个待判蠓虫样本进行判别；
- 如果 Apf 蠓虫是某种疾病的载体毒蠓，Af 蠓虫是传粉益蠓，在这种情况下是否需要修改所用的分类方法，如何修改？

问题背景 本问题改编自 1989 年美国大学生数学建模竞赛的赛题，实际中经常会遇到这类数学建模问题，比如医生在掌握了以往各种疾病（如流感、肺炎、心肌炎等）指标特征的情况下，根据一个新患者的各项检查指标来判定他患有哪类疾病，又如已知不同类型飞机的雷达反射波的各项特征指标，来判定一架飞机属于哪种类型.这种在已知样本分类的前提下，利用已有的样本数据建立判别模型，用来对未知类别的新样本进行分类的问题属于统计学中的**判别分析**.目前，判别分析已经成为数据挖掘、机器学习、模式识别（语音识别、图像识别、指纹识别、文本识别）等应用领域的重要理论基础.

蠓虫分类判别属于判别分析，可以通过生物学家给出的 9 个 Af 蠓虫与 6 个 Apf 蠓虫作为训练样本，利用这些蠓虫的触角长和翅长的数据，建立关于指标的判别函数和判别准则，由此对新的蠓虫样本进行分类，并且利用判别准则进行回判或交互验证，以估计误判率.下面先利用距离判别方法建立比较简单的模型，再用 Bayes 判别方法处理需要区分毒蠓和益蠓的特殊情况.

问题分析 对表 1 蠓虫样本的数据作散点图（图 1），观察图形发现，Apf 蠓虫的数据点集中在图的左上方，Af 蠓虫的数据点集中在图的右下方，而 3 个待判样本处于两类点之间.这表明每类蠓虫的触角长和翅长两个指标之间是相关的.直观的判别方法应该是，找一条直线把这两类点分开，将它作为 Apf 蠓虫和 Af 蠓虫的分界线，依此构成判别准则，用于确定 3 个待判样本属于哪一类.

距离判别模型的基本思路 假设 Apf 蠓虫和 Af 蠓虫的训练样本分别来自总体 G_1 和 G_2，蠓虫的指标记作 2 维向量 $\boldsymbol{x}=(x^{(1)},x^{(2)})^{\mathrm{T}}$，是平面上的一个点，其中 $x^{(1)},x^{(2)}$ 分别为蠓虫的触角长和翅长.距离判别模型的基本思路是，恰当地定义每个总体的中心，以及每个样本与中心的距离；计算待判样本与两个中心的距离，以距离较近作为判别

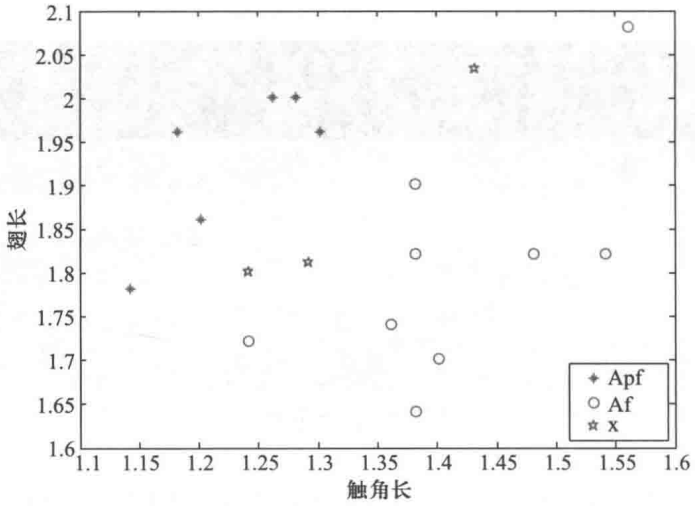


图 1 蠓虫样本数据的散点图(☆表示待判样本)

准则.[34]

设总体 G 的期望值向量为 $\boldsymbol{\mu}$,协方差矩阵为 $\boldsymbol{\Sigma}$.显然,取 $\boldsymbol{\mu}$ 为总体的中心.至于距离,我们不用通常的欧氏距离,而采用统计学的马氏(Mahalanobis)距离.样本 $\boldsymbol{x}$ 到总体 G 的马氏距离定义为

$$d(\boldsymbol{x},G)=\sqrt{(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})} \tag{1}$$

应该指出,马氏距离考虑了指标 $x^{(1)}$(触角长)和 $x^{(2)}$(翅长)之间的相关性及各自取值的分散性,而欧氏距离则把两个指标视为完全独立的变量,所以在制定判别准则时采用马氏距离更加合理.显然,当 $\boldsymbol{\Sigma}$ 为单位矩阵时,两个距离等价.

记总体 G_1 和 G_2 的期望值向量分别为 $\boldsymbol{\mu}_1$ 与 $\boldsymbol{\mu}_2$,协方差矩阵分别为 $\boldsymbol{\Sigma}_1$ 与 $\boldsymbol{\Sigma}_2$.对给定的样本 $\boldsymbol{x}$,分别计算 $\boldsymbol{x}$ 到总体 G_1 和 G_2 的马氏距离的平方 $d^2(\boldsymbol{x},G_1)$ 和 $d^2(\boldsymbol{x},G_2)$.依据距离就近的原则,若 $d^2(\boldsymbol{x},G_1)\leqslant d^2(\boldsymbol{x},G_2)$,则判定 $\boldsymbol{x}\in G_1$;否则判定 $\boldsymbol{x}\in G_2$.由于

$$d^2(\boldsymbol{x},G_2)-d^2(\boldsymbol{x},G_1)=(\boldsymbol{x}-\boldsymbol{\mu}_2)^{\mathrm{T}}\boldsymbol{\Sigma}_2^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_2)-(\boldsymbol{x}-\boldsymbol{\mu}_1)^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_1) \tag{2}$$

若记

$$W(\boldsymbol{x})=\frac{1}{2}(d^2(\boldsymbol{x},G_2)-d^2(\boldsymbol{x},G_1)) \tag{3}$$

则判别准则等价于

$$\begin{cases}\boldsymbol{x}\in G_1, & \text{如果 } W(\boldsymbol{x})\geqslant 0\\ \boldsymbol{x}\in G_2, & \text{如果 } W(\boldsymbol{x})<0\end{cases} \tag{4}$$

称 $W(\boldsymbol{x})$ 为距离判别函数,它一般是 $\boldsymbol{x}$ 的二次函数,而当 $\boldsymbol{\Sigma}_1=\boldsymbol{\Sigma}_2=\boldsymbol{\Sigma}$ 时容易证明

$$W(\boldsymbol{x})=\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{\mu}),\quad \boldsymbol{a}=\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2),\quad \boldsymbol{\mu}=\frac{1}{2}(\boldsymbol{\mu}_1+\boldsymbol{\mu}_2) \tag{5}$$

这时 $W(\boldsymbol{x})$ 是 $\boldsymbol{x}$ 的一次函数,称为线性距离判别函数,并称 $\boldsymbol{a}$ 为判别系数.

蠓虫分类的距离判别模型 在实际应用中,总体的期望值向量 $\boldsymbol{\mu}_1,\boldsymbol{\mu}_2$ 及协方差矩阵 $\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2$ 需要用训练样本数据来估计.设样本 $\boldsymbol{x}_{1k}(k=1,2,\cdots,n_1)$ 取自 G_1,$\boldsymbol{x}_{2k}(k=1,$

$2,\cdots,n_2$)取自 G_2,分别计算样本的均值向量 $\bar{\boldsymbol{x}}_1,\bar{\boldsymbol{x}}_2$

$$\bar{\boldsymbol{x}}_1=\frac{1}{n_1}\sum_{k=1}^{n_1}\boldsymbol{x}_{1k},\quad \bar{\boldsymbol{x}}_2=\frac{1}{n_2}\sum_{k=1}^{n_2}\boldsymbol{x}_{2k} \tag{6}$$

和协方差矩阵 $\boldsymbol{S}_1,\boldsymbol{S}_2$

$$\boldsymbol{S}_1=\frac{1}{n_1-1}\sum_{k=1}^{n_1}(\boldsymbol{x}_{1k}-\bar{\boldsymbol{x}}_1)(\boldsymbol{x}_{1k}-\bar{\boldsymbol{x}}_1)^{\mathrm{T}},\quad \boldsymbol{S}_2=\frac{1}{n_2-1}\sum_{k=1}^{n_2}(\boldsymbol{x}_{2k}-\bar{\boldsymbol{x}}_2)(\boldsymbol{x}_{2k}-\bar{\boldsymbol{x}}_2)^{\mathrm{T}} \tag{7}$$

作为 $\boldsymbol{\mu}_1,\boldsymbol{\mu}_2$ 及 $\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2$ 的估计.

当 $\boldsymbol{\Sigma}_1\neq\boldsymbol{\Sigma}_2$ 时,用 $\boldsymbol{S}_1,\boldsymbol{S}_2$ 代替(2)式的 $\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2$.当 $\boldsymbol{\Sigma}_1=\boldsymbol{\Sigma}_2$ 时,用混合协方差矩阵

$$\boldsymbol{S}_w=\frac{(n_1-1)\boldsymbol{S}_1+(n_2-1)\boldsymbol{S}_2}{n_1+n_2-2} \tag{8}$$

代替(5)式的 $\boldsymbol{\Sigma}$.

为了将上述方法应用到蠓虫分类判别,首先需要确定 Apf 蠓虫总体 G_1 和 Af 蠓虫 G_2 是否具有相同的协方差矩阵,这关系到我们将选用线性判别函数还是二次判别函数,为此需要先作两总体 G_1 和 G_2 协方差矩阵的一致性检验(Box M 检验),即检验假设

$$H_0:\boldsymbol{\Sigma}_1=\boldsymbol{\Sigma}_2,\quad H_1:\boldsymbol{\Sigma}_1\neq\boldsymbol{\Sigma}_2$$

由多元统计的结果可以知道,对于 2 维总体,该假设检验的统计量为①

$$M^*=(1-c)M\sim\chi^2(3)$$

其中 $M=(n_1+n_2-2)\ln|\boldsymbol{S}_w|-(n_1-1)\ln|\boldsymbol{S}_1|-(n_2-1)\ln|\boldsymbol{S}_2|$,$c=\dfrac{13}{18}\left(\dfrac{1}{n_1-1}+\dfrac{1}{n_2-1}-\dfrac{1}{n_1+n_2-2}\right)$.

对给定的检验水平 α,计算概率 $p_0=P(M^*>\chi^2_\alpha(3))$,若 $p_0<\alpha$ 则拒绝 H_0,否则接受 H_0.在蠓虫分类问题中,若取 $\alpha=0.05$,经编程计算,$p_0=0.435\ 9>0.05$,故接受原假设 H_0,可以用线性判别函数进行分类.



程序文件 9-16
蠓虫分类判别
prog0906.m

根据表 1 的蠓虫样本数据,按照(5)~(8)式编程计算②,输出的判别系数 $\boldsymbol{a}=(-58.2364,38.0587)^{\mathrm{T}}$,常数项为 5.8715,故线性距离判别函数为

$$W(\boldsymbol{x})=-58.2364x^{(1)}+38.0587x^{(2)}+5.8715 \tag{9}$$

判别直线 $W(\boldsymbol{x})=0$ 的图形见图 2. 3 个待判蠓虫样本的判别结果见表 2,其结论是:待判蠓虫样本 1 属于 Apf 蠓虫,而样本 2 和 3 均属于 Af 蠓虫.从图 2 也可直接给出这个结果.

表 2 3 个待判蠓虫样本的距离判别结果

待判蠓虫序号	触角长 $x^{(1)}$	翅长 $x^{(2)}$	判别函数值 $W(x)$	判别结果
1	1.24	1.80	2.1640	Apf
2	1.29	1.81	-0.3673	Af
3	1.43	2.03	-0.1475	Af

① 这里假定总体 G_1 和 G_2 服从正态分布.

② MATLAB 中的函数命令 classify 可用于对未知类别样本的距离判别和一些特定的 Bayes 判别,详见帮助系统.

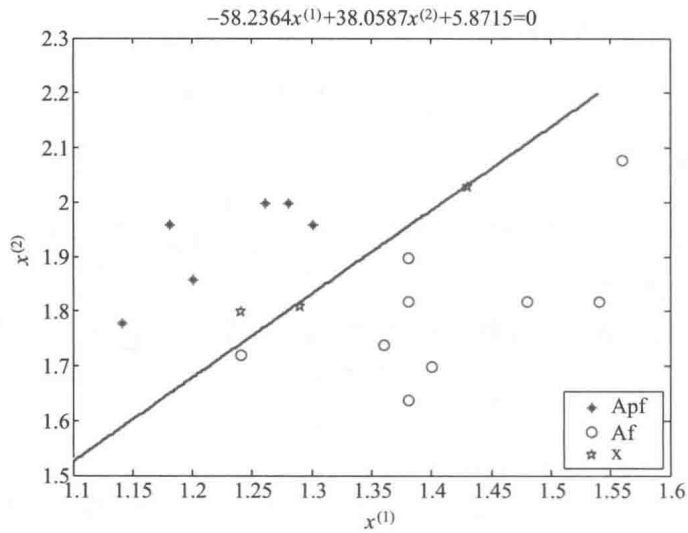


图 2 距离判别直线及样本判别结果图

模型检验 对于距离判别函数比如蠓虫分类判别(9)式的有效性,通常有下面两种检验方法:

方法一 回代误判法

将取自总体 G_1 的 n_1 个训练样本和 G_2 的 n_2 个训练样本,逐个回代到判别函数中并判定其归属.若原本属于 G_1 被误判属于 G_2 的样本个数为 n_{12},原本属于 G_2 被误判属于 G_1 的样本个数为 n_{21},则回代误判率的估计值为 $\hat{p}=\dfrac{n_{12}+n_{21}}{n_1+n_2}$.

对于蠓虫分类判别问题,经编程计算,$n_{12}=n_{21}=0$,故回代误判率的估计值为 0.

方法二 交叉验证法

从来自总体 G_1 的 n_1 个训练样本中,每次拿出一个作为验证样本,其余 n_1-1 个与来自总体 G_2 的 n_2 个一起,作为训练样本用于建立判别准则,用验证样本进行检验.然后换一个作为验证样本.在总共 n_1 次检验中误判的样本个数记为 n_{12}^*.对来自总体 G_2 的 n_2 个训练样本完成同样的步骤,在总共 n_2 次检验中误判的样本个数记为 n_{21}^*,则交叉验证误判率的估计值为 $\hat{p}^*=\dfrac{n_{12}^*+n_{21}^*}{n_1+n_2}$.

对于蠓虫分类判别问题,经编程计算,$n_{12}^*=0$, $n_{21}^*=1$,因为 $n_1+n_2=15$,所以交叉验证误判率的估计值为 1/15.

距离判别方法虽然简单、直观,应用较广,但也有明显的缺点,一是没有考虑在整体环境中两个总体出现的概率会有不同;二是没有涉及误判造成的损失的影响.在蠓虫分类问题中已经指出,Apf 蠓虫是某种疾病的载体毒蠓,Af 蠓虫是传粉益蠓.这两种蠓虫在自然界中出现的概率应该是不一样的,并且,将毒蠓 Apf 误判成益蠓 Af 的危害要比将益蠓 Af 误判成毒蠓 Apf 的危害更大.基于这些考虑,需要在模型中引入总体类别的先验概率和误判造成的损失函数,于是有下面的 Bayes 判别模型.[34]

评注 距离判别模型的思路简单、直观,线性判别函数是最常用的.只是两总体协方差矩阵的一致性检验需用到总体正态分布的假定,实际上有时较难处理.

模型检验用的回代误判法虽然简单,缺点是建立与检验判别准则使用相同的样本,往往会低估误判率.交叉验证法虽计算量较大,但克服了回代法的缺点,是一种较好的检验方法.

Bayes 判别模型的基本思路 假设样本来自总体 G_1 和 G_2 的先验概率分别是 p_1 和 $p_2(p_1+p_2=1)$，并且两总体的概率密度函数分别是 $f_1(\boldsymbol{x})$ 和 $f_2(\boldsymbol{x})$，那么在取到样本 $\boldsymbol{x}$ 以后，它属于总体 $G_i(i=1,2)$ 的后验概率可以根据 Bayes 公式得到

$$P(G_i|\boldsymbol{x})=\frac{p_i f_i(\boldsymbol{x})}{p_1 f_1(\boldsymbol{x})+p_2 f_2(\boldsymbol{x})},\quad i=1,2 \tag{10}$$

因此，在不考虑误判损失的情况下，有以下的判别规则

$$\begin{cases}\boldsymbol{x}\in G_1, & \text{如果 } P(G_1|\boldsymbol{x})\geqslant P(G_2|\boldsymbol{x})\\ \boldsymbol{x}\in G_2, & \text{如果 } P(G_1|\boldsymbol{x})<P(G_2|\boldsymbol{x})\end{cases} \tag{11}$$

若考虑误判损失，用 R_1,R_2 分别表示样本 $\boldsymbol{x}$ 根据某种判别规则被判入总体 G_1 和 G_2 的取值集合($R_1\cup R_2=\Omega$).用 $L(j|i)(i,j=1,2)(i\neq j)$ 表示将来自总体 G_i 的样本 $\boldsymbol{x}$ 误判入 $\boldsymbol{G}_j$ 的损失，造成损失 $L(j|i)(i\neq j)$ 的误判概率为

$$P(j|i)=P(\boldsymbol{x}\in R_j|\boldsymbol{x}\in G_i)=\int_{R_j} f_i(\boldsymbol{x})\mathrm{d}\boldsymbol{x},\quad i,j=1,2 \tag{12}$$

因此，总误判损失的期望(平均误判损失)为

$$ECM(R_1,R_2)=L(2|1)P(2|1)p_1+L(1|2)P(1|2)p_2 \tag{13}$$

一个合理的判别准则是最小化 $ECM(R_1,R_2)$.

可以证明，最小化 $ECM(R_1,R_2)$ 的判别准则为

$$\begin{cases}\boldsymbol{x}\in G_1, & \text{如果 } f_1(\boldsymbol{x})L(2|1)p_1\geqslant f_2(\boldsymbol{x})L(1|2)p_2\\ \boldsymbol{x}\in G_2, & \text{如果 } f_1(\boldsymbol{x})L(2|1)p_1<f_2(\boldsymbol{x})L(1|2)p_2\end{cases} \tag{14}$$

称为 **Bayes 判别准则**.

特别地，当总体 $G_1\sim N(\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1)$，$G_2\sim N(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2)$，且 $\boldsymbol{\Sigma}_1=\boldsymbol{\Sigma}_2=\boldsymbol{\Sigma}$ 时，经过简单的计算，可知(14)式等价于以下的 Bayes 判别准则

$$\begin{cases}\boldsymbol{x}\in G_1, & \text{如果 } W_B(\boldsymbol{x})\geqslant\beta\\ \boldsymbol{x}\in G_2, & \text{如果 } W_B(\boldsymbol{x})<\beta\end{cases} \tag{15}$$

其中

$$W_B(\boldsymbol{x})=\left(\boldsymbol{x}-\frac{\boldsymbol{\mu}_1+\boldsymbol{\mu}_2}{2}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2),\quad \beta=\ln\frac{L(1|2)p_2}{L(2|1)p_1} \tag{16}$$

$W_B(\boldsymbol{x})$ 称为 **Bayes 判别函数**，与(5)式定义的线性距离判别函数完全一致，因此当阈值 $\beta=0$ 时，Bayes 判别准则与线性距离判别准则等价，而这正是既不考虑样本的先验概率(即 $p_1=p_2$)、也不考虑误判损失(即 $L(1|2)=L(2|1)$)的特殊情况.

蠓虫分类的 Bayes 判别模型 假定 Apf 蠓虫总体服从二维正态分布，$G_1\sim N(\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1)$，Af 蠓虫总体 $G_2\sim N(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2)$①.前面已经验证 $\boldsymbol{\Sigma}_1=\boldsymbol{\Sigma}_2=\boldsymbol{\Sigma}$，Bayes 判别函数 $W_B(\boldsymbol{x})$ 与(9)式表示的线性距离判别函数 $W(\boldsymbol{x})$ 完全相同，对待判蠓虫进行判别时只需考虑阈值 β 的影响.

实际上，总体 G_1 和 G_2 的先验概率可以利用历史资料和经验进行估计，一种常见的

① 这个假定可以通过二维正态分布 Q-Q 图检验来验证，但考虑到样本数太小，不妨认为它是基本合理的.

做法是取为训练样本个数的比例，即令 Apf 蠓虫的先验概率 $p_1=\dfrac{n_1}{n_1+n_2}=0.4$，Af 蠓虫的先验概率 $p_2=\dfrac{n_2}{n_1+n_2}=0.6$. 考虑到将毒蠓 Apf 误判成益蠓 Af 的危害更大，可设 $L(2|1)=\alpha L(1|2)$，其中参数 $\alpha>1$，于是阈值 $\beta=\ln\left(\dfrac{3}{2\alpha}\right)$，Bayes 判别函数为

$$W_B(\boldsymbol{x})=-58.2364x^{(1)}+38.0587x^{(2)}+5.8715 \tag{17}$$

而 Bayes 判别准则为

$$\begin{cases}\boldsymbol{x}\in G_1, & \text{如果 } W_B(\boldsymbol{x})\geqslant\ln\left(\dfrac{3}{2\alpha}\right)\\ \boldsymbol{x}\in G_2, & \text{如果 } W_B(\boldsymbol{x})<\ln\left(\dfrac{3}{2\alpha}\right)\end{cases} \tag{18}$$

分别取 $\alpha=1.5,2,2.5$ 等，用 Bayes 判别准则对待判蠓虫进行判别，结果见表 3.

表 3 不同误判损失下的判别结果

待判蠓虫序号	触角长 $x^{(1)}$	翅长 $x^{(2)}$	判别函数值 $W_B(x)$	判别结果 ($\alpha=1.5$)	判别结果 ($\alpha=2.0$)	判别结果 ($\alpha=2.5$)
1	1.24	1.80	2.1640	Apf	Apf	Apf
2	1.29	1.81	−0.3673	Af	Af	Apf
3	1.43	2.03	−0.1475	Af	Apf	Apf

当 $\alpha=1.5$ 时，三个待判样本的 Bayes 判别法的判别结果完全与距离判别法的结果一致，当 $\alpha=2$ 时，待判样本 3 被判别为 Apf 蠓虫，而当 $\alpha=2.5$ 时，待判样本 2 和 3 均被判别为 Apf 蠓虫，这充分反映了考虑到误判造成损失的 Bayes 判别法的作用，也说明了 Bayes 判别法要比距离判别法更切合实际.

经编程计算，对上述三种 α 的不同取值，Bayes 判别的回代误判率的估计值为 0，事实上可以验证，当 $\alpha\in[0.0129,3.6189]$ 时，回代误判率的估计值均为 0. 其交叉验证误判率的估计值，有兴趣的读者可以自行编程得到.

评注 本节以蠓虫分类判别为例介绍了距离判别模型和 Bayes 判别模型，针对的是两分类的判别问题，可以推广到多分类的判别情况. 还有很多其他的判别方法，拓展阅读 9-1 将介绍 Fisher 判别法. 还有，逐步判别法、k 近邻判别法、支持向量机及主成分判别法等，在处理高维数据分类研究中起着非常重要的作用，目前判别分析建模已经成为人工智能和数据挖掘中非常有影响力的方法.

复习题

拓展阅读 9-1 蠓虫分类的 Fisher 判别模型

1. 假定 Apf 蠓虫总体与 Af 蠓虫总体的协方差矩阵不相等，对问题中所给的蠓虫数据，重新估计三个待判样本到两总体的距离，并进行分类判别.

2. 如果已经确认了三个待判蠓虫样本 1,2 和 3 分别属于 Apf 蠓虫、Af 蠓虫和 Af 蠓虫，请再次用距离判别法对新发现的蠓虫样本 $\boldsymbol{x}=(1.35,1.88)^{\mathrm{T}}$ 进行判别，并写出线性距离判别函数.

9.7 学生考试成绩综合评价

某高校数学系为开展研究生的推荐免试工作，对报名参加推荐的 52 名学生已修过的 6 门课的考试分数统计如表 1. 这 6 门课是：数学分析、高等代数、概率论、微分几何、

抽象代数和数值分析，其中前 3 门基础课采用闭卷考试，后 3 门为开卷考试.

表 1 52 名学生的原始考试成绩(全部数据见数据文件 9-4)

学生序号	数学分析	高等代数	概率论	微分几何	抽象代数	数值分析	总分
A_1	62	71	64	75	70	68	410
A_2	52	65	57	67	60	58	359
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
A_{52}	70	73	70	88	79	69	449



数据文件 9-4
学生考试成绩

在以往的推荐免试工作中，该系是按照学生 6 门课成绩的总分进行学业评价，再根据其他要求确定最后的推荐顺序.但是这种排序办法没有考虑到课程之间的相关性，以及开闭卷等因素，丢弃了一些信息.我们的任务是利用这份数据建立一个统计模型，并研究以下问题：

- 如何确定若干综合评价指标来最大程度地区分学生的考试成绩，并在不丢失重要信息的前提下简化对学生的成绩排序；
- 在学生评价中如何体现开闭卷的影响，找到成绩背后的潜在因素，并科学地针对考试成绩进行合理排序.

问题分析 考试成绩是目前衡量学生学业水平最重要的标准，对于多门课的成绩，通常的方法是用总分作为排序的定量依据.这样做虽然简化了问题，但失去了许多有用的信息.若用 x_1,x_2,x_3,x_4,x_5,x_6 分别表示数学分析、高等代数、概率论、微分几何、抽象代数和数值分析的分数，那么 6 维向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_6)^{\mathrm{T}}$ 表示一个学生的 6 门课的分数，平均分相当于各门课分数的等权平均值，是将一个 6 维的数据简单地化为一维指标.能不能不用这样的平均分，而是寻找一组权重 $\boldsymbol{a}_1=(a_{11},a_{12},\cdots,a_{16})^{\mathrm{T}}$，将加权后的平均分数 $y_1=\sum_{j=1}^{6}a_{1j}x_j$ 作为评价一个学生综合成绩的指标呢？

当然希望选取合适的 $\boldsymbol{a}_1$ 使 y_1 能尽可能多地反映原变量 $\boldsymbol{x}$ 的信息，即最大程度地区分学生的成绩.我们知道，一个变量的方差越大，它的区分性就越大，因此取 $\boldsymbol{a}_1$ 使 y_1 的方差最大，并且可以要求 $\boldsymbol{a}_1$ 是单位向量，即 $\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_1=1$.这样选择的原变量 $\boldsymbol{x}$ 的线性组合 y_1 在统计上称为第一主成分(即第一个综合变量).用 y_1 代替 $\boldsymbol{x}$，评价指标就从 6 维降到了 1 维.

如果 y_1 反映的原变量 $\boldsymbol{x}$ 的信息还不够充分，则可以提取新的信息，构造第二主成分 $y_2=\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{x}$，在 y_2 与 y_1 所提供的信息不重叠(即不相关)的条件下，选取 $\boldsymbol{a}_2$ 使 y_2 的方差最大.用 y_1,y_2 代替 $\boldsymbol{x}$，评价指标降到 2 维.依此类推，这种解决问题的方法就是统计中主成分分析的基本思路.

为了直观地描述这样的降维过程，假定只有数学分析和高等代数两门课的成绩，即 2 维数据，全体学生的成绩可用散点图表示，如图 1 中的符号+(其横坐标、纵坐标分别是数学分析和高等代数的分数).这些数据点大多集中在一个向上斜置的椭圆内，表明两门课的分数有较强的正相关性.椭圆的长轴(实线)与短轴(虚线)相互垂直，显然，在长轴方向数据变化较大，在短轴方向数据变化较小，选用长轴方向的 1 维

变量就包含了2维数据的大部分信息，而寻求这样的新变量可以通过变量代换（即坐标旋转）来完成.

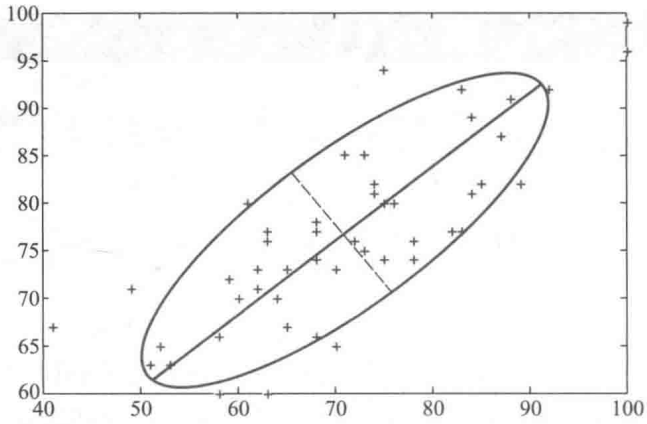


图1 2维数据（数学分析和高等代数的分数）降维示意图

主成分分析和因子分析是用统计学解决这个问题的两种方法[34]，下面先简单介绍它们的基本思路，再用来处理学生成绩的综合评价.

主成分分析的基本思路 按照统计学的观点，将学生各门课（假定是 p 门）的分数视为一个 p 维随机变量，记作 $\boldsymbol{x}=(x_1,x_2,\cdots,x_p)^{\mathrm{T}}$，假设 $\boldsymbol{x}$ 的期望向量 $E(\boldsymbol{x})=\boldsymbol{\mu}$ 和协方差矩阵 $\mathrm{Cov}(\boldsymbol{x})=\boldsymbol{\Sigma}$ 为已知.用一组单位向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_p$ 构造 $\boldsymbol{x}$ 的 p 个线性组合，称为**主成分** $\boldsymbol{y}=(y_1,y_2,\cdots,y_p)^{\mathrm{T}}$，有

$$\begin{cases} y_1=a_{11}x_1+a_{12}x_2+\cdots+a_{1p}x_p=\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{x} \\ y_2=a_{21}x_1+a_{22}x_2+\cdots+a_{2p}x_p=\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{x} \\ \cdots\cdots\cdots\cdots \\ y_p=a_{p1}x_1+a_{p2}x_2+\cdots+a_{pp}x_p=\boldsymbol{a}_p^{\mathrm{T}}\boldsymbol{x} \end{cases} \tag{1}$$

简记作

$$\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x},\quad \boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_p)^{\mathrm{T}} \tag{2}$$

$\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_p$ 称为**主成分（载荷）系数**.要求各个主成分 $y_1,y_2,\cdots,y_p$ 之间互不相关，即协方差矩阵 $\mathrm{Cov}(\boldsymbol{y})$ 为对角阵，记作 $\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_p)$，且第一主成分 y_1 是 $\boldsymbol{x}$ 的一切线性组合中方差最大的，第二主成分 y_2 是与 y_1 不相关的 $\boldsymbol{x}$ 的线性组合中方差最大的，依此类推.

怎样确定主成分系数 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_p$ 呢？按照 $\boldsymbol{y}$ 的协方差矩阵应为对角阵的要求，有

$$\mathrm{Cov}(\boldsymbol{y})=\mathrm{Cov}(\boldsymbol{A}\boldsymbol{x})=\boldsymbol{A}\mathrm{Cov}(\boldsymbol{x})\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_p) \tag{3}$$

因为 $\boldsymbol{x}$ 的协方差矩阵 $\boldsymbol{\Sigma}$ 通常是对称正定矩阵，根据线性代数的基本定理，一定存在一组单位正交特征向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_p$ 构成的正交矩阵 $\boldsymbol{A}$，使（3）式成立，且若 $\boldsymbol{\Sigma}$ 的特征根（也是 $\boldsymbol{\Lambda}$ 的特征根）按照大小排序 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_p\geqslant 0$，则 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_p$ 为其对应的单位正交特征向量.

评注 原始变量 $x_1,x_2,\cdots,x_p$ 之间的相关性越强，主成分包含的信息

记 $\boldsymbol{\Sigma}$ 的特征根之和为 $\lambda=\sum\limits_{j=1}^{p}\lambda_j$，则主成分 $y_1,y_2,\cdots,y_p$ 的方差之和与原始变量

$x_1, x_2, \cdots, x_p$ 的方差之和均等于 λ，可见 p 个互不相关的主成分包含了原始数据中的全部信息，但主成分所包含的信息更为集中.并且，第 j 主成分 y_j 的方差占全部方差的比例为 λ_j/λ，称为 y_j 的方差**贡献率**，显然第一主成分 y_1 的方差贡献率最大，其余的依次递减.前 m 个主成分的方差贡献率之和称为它们的**累积贡献率**.

越集中，用主成分代替原始变量的效果越好.相反，如果 $x_1, x_2, \cdots, x_p$ 相互独立，其协方差阵 $\boldsymbol{\Sigma}$ 为对角阵，主成分就是它自己，所以，原始变量独立性越强，主成分分析的效果越差.

主成分分析的目的是要降维，所以一般不会使用所有的 p 个主成分，在信息损失不太多的情况下，可用少数几个主成分来代替原始变量进行数据分析.究竟需要多少个主成分来代替呢？通常取累积贡献率达到 80%的前 m 个即可.

学生成绩的主成分分析 回到学生成绩的综合评价问题，记 $x_{ij}(i=1,2,\cdots,n,\ j=1,2,\cdots,p)$ 为第 i 位学生第 j 门课的分数，$\boldsymbol{X}=(x_{ij})_{n\times p}$ 为分数数据矩阵（对于表 1，$n=52$，$p=6$）.记 $\boldsymbol{x}_i=(x_{i1},x_{i2},\cdots,x_{ip})^{\mathrm{T}}(i=1,2,\cdots,n)$，是 p 维随机变量 $\boldsymbol{x}$ 的一个观测值，均值向量 $\bar{\boldsymbol{x}}=\dfrac{1}{n}\sum\limits_{i=1}^{n}\boldsymbol{x}_i=(\bar{x}_1,\bar{x}_2,\cdots,\bar{x}_p)^{\mathrm{T}}$ 作为 $E(\boldsymbol{x})=\boldsymbol{\mu}$ 的估计值，$\boldsymbol{X}$ 的协方差矩阵 $\boldsymbol{S}=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(\boldsymbol{x}_i-\bar{\boldsymbol{x}})(\boldsymbol{x}_i-\bar{\boldsymbol{x}})^{\mathrm{T}}$ 作为 $\mathrm{Cov}(\boldsymbol{x})=\boldsymbol{\Sigma}$ 的估计值.

计算 $\boldsymbol{S}$ 的特征根并按大小排序为 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_p\geqslant 0$，其对应的单位正交特征向量就是主成分系数 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_p$.

MATLAB 中用观测数据矩阵 $\boldsymbol{X}$ 进行主成分分析的函数命令是 princomp，调用格式为：

```
[COEFF, SCORE, LATENT] = princomp(X)
```

其中输入参数 X 是 $n\times p$ 阶观测数据矩阵，每一行对应一个观测值，每一列对应一个变量.输出参数 COEFF 是主成分的 $p\times p$ 阶系数矩阵，第 j 列是第 j 主成分的系数向量；SCORE 是 $n\times p$ 阶得分矩阵，其第 i 行第 j 列元素是第 i 观测值、第 j 主成分的得分；LATENT是 X 的特征根按大小排序构成的向量.

值得注意的是，princomp 函数对观测数据进行了中心化处理，即 $\boldsymbol{X}$ 的每一个元素减去其所在列的均值，相应地，其输出也是中心化的主成分系数.

用协方差矩阵或相关系数矩阵进行主成分分析的函数命令是 pcacov，调用格式见 MATLAB 帮助系统.

对于数据文件 9-4 的考试成绩数据矩阵，首先计算均值向量 $\bar{\boldsymbol{x}}$ 和协方差矩阵 $\boldsymbol{S}$，得

$$\begin{aligned}\bar{\boldsymbol{x}}&=(\bar{x}_1,\bar{x}_2,\bar{x}_3,\bar{x}_4,\bar{x}_5,\bar{x}_6)^{\mathrm{T}}\\&=(70.903\,8,76.576\,9,71.807\,7,74.769\,2,67.826\,9,61.384\,6)^{\mathrm{T}}\end{aligned}$$

$$\boldsymbol{S}=(s_{ij})_{6\times 6}=\begin{pmatrix}163.382\,7 & 97.880\,1 & 113.569\,4 & -67.297\,1 & -67.487\,6 & -71.550\,5\\ 97.880\,1 & 86.641\,0 & 82.054\,3 & -35.746\,6 & -39.623\,7 & -37.206\,6\\ 113.569\,4 & 82.054\,3 & 113.296\,5 & -35.986\,4 & -43.877\,1 & -54.356\,0\\ -67.297\,1 & -35.746\,6 & -35.986\,4 & 192.455\,5 & 90.253\,4 & 88.168\,9\\ -67.487\,6 & -39.623\,7 & -43.877\,1 & 90.253\,4 & 88.498\,9 & 67.342\,4\\ -71.550\,5 & -37.206\,6 & -54.356\,0 & 88.168\,9 & 67.342\,4 & 122.672\,7\end{pmatrix}\tag{4}$$



程序文件 9-17
学生考试成绩综合评价
prog0907.m

由 MATLAB 编程计算出协方差矩阵 $\boldsymbol{S}$ 的特征根，进而得到各主成分的方差贡献率，以便确定主成分的个数，具体结果见表 2.从图像上观测主成分的方差累积贡献率也

是常见的办法之一，只要图2中的曲线达到方差解释为80%的位置，即可确定所需主成分的个数.

表2 协方差矩阵的特征根与贡献率

S 的特征根	贡献率	累积贡献率
469.681 6	61.081 2	61.081 2
173.952 5	22.622 2	83.703 4
58.510 0	7.609 1	91.312 5
29.252 7	3.804 3	95.116 7
21.416 3	2.785 1	97.901 9
16.133 4	2.098 1	100.000 0

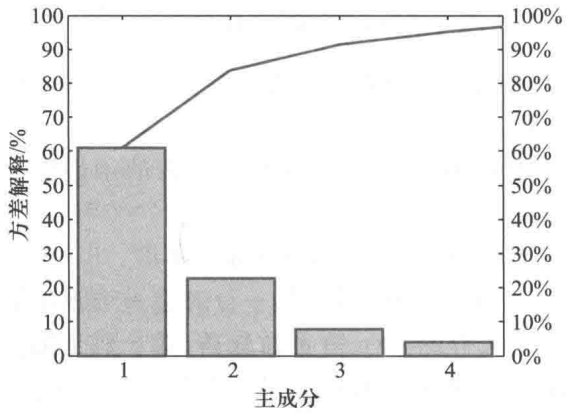


图2 方差解释图

前两个主成分的累积贡献率为83.703 4%，所以只取两个主成分是合适的.根据主成分系数矩阵的输出结果，可得第一主成分与第二主成分分别为

$$\begin{aligned} y_1 &= 0.515\ 7x_1^* + 0.332\ 1x_2^* + 0.387\ 9x_3^* - 0.453\ 4x_4^* - 0.345\ 8x_5^* - 0.385\ 0x_6^* \\ y_2 &= 0.381\ 2x_1^* + 0.348\ 2x_2^* + 0.414\ 7x_3^* + 0.677\ 0x_4^* + 0.222\ 3x_5^* + 0.231\ 8x_6^* \end{aligned} \tag{5}$$

这里 x_j^* 是 x_j 的中心化数据，即 $x_j^* = x_j - \bar{x}_j, j = 1, 2, \cdots, 6.$

结果分析 在(5)式中第一主成分对应的系数符号前3个均为正，后3个均为负，系数绝对值相差不大.由于前3个系数正好对应3门闭卷考试分数，后3个对应开卷考试分数，如果一个学生第一主成分的得分是个很大的正数，说明他更擅长闭卷考试，反之，如果得分是一个绝对值很大的负数，就说明他在开卷科目考试中有很好的表现.如果得分接近于0，则说明开闭卷对该学生无所谓.因此，第一主成分实际上反映了开闭卷考试的差别，可理解为“成绩的开闭卷成分”.第二主成分对应的系数符号均为正，只有微分几何课程对应的系数比其他课程略大，反映了学生各门课程成绩的某种均衡性，可理解为“成绩的均衡成分”.

通过以上分析可知，为了综合评价考试成绩，需要知道每个学生在这两个主成分上

的得分.根据得分矩阵的输出结果,将 52 名学生的第一主成分、第二主成分的得分以及成绩总分列于表 3,第一、第二主成分的得分散点图如图 3.

表 3 第一、第二主成分得分表(全部数据见运行程序文件 9-17 后的输出结果)

学生序号	成绩总分	第一主成分得分	第二主成分得分
A_1	410	-12.874 8	-6.401 1
A_2	359	-11.803 7	-25.162 0
⋮	⋮	⋮	⋮
A_{52}	449	-15.149 4	10.866 3

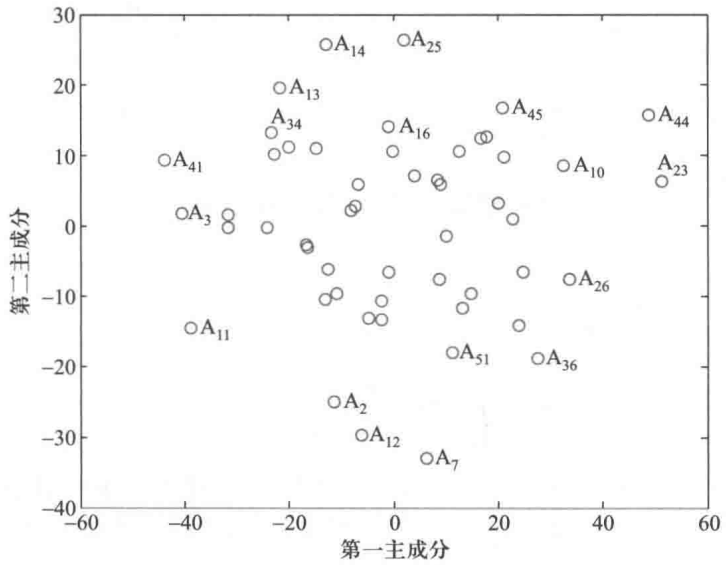


图 3 前两个主成分得分散点图

由图 3 可以直观地发现,从第一主成分来看,学生 A_{23}、A_{44}、A_{26}、A_{10}具有较大的正数,说明他们擅长于闭卷考试,体现在三门基础课数学分析、高等代数和概率论有较好的成绩,学生 A_{41}、A_3、A_{11}有绝对值较大的负数,说明他们更擅长于开卷考试;从第二主成分来看,学生 A_{25}、A_{14}、A_{13}具有较大的正数,说明他们 6 门课程比较均衡,成绩也较好,而学生 A_7、A_{12}、A_2 有绝对值较大的负数,说明各科成绩均不太理想.

因子分析的基本思路 与主成分分析中构造原始变量 $x_1,x_2,\cdots,x_p$ 的线性组合 $y_1,y_2,\cdots,y_p$(见(1)式)不同,因子分析是将原始变量 $x_1,x_2,\cdots,x_p$ 分解为若干个因子的线性组合,表示为

$$\begin{cases} x_1=\mu_1+a_{11}f_1+a_{12}f_2+a_{13}f_3+\cdots+a_{1m}f_m+\varepsilon_1 \\ x_2=\mu_2+a_{21}f_1+a_{22}f_2+a_{23}f_3+\cdots+a_{2m}f_m+\varepsilon_2 \\ \cdots\cdots\cdots \\ x_p=\mu_p+a_{p1}f_1+a_{p2}f_2+a_{p3}f_3+\cdots+a_{pm}f_m+\varepsilon_p \end{cases} \tag{6}$$

简记作

评注 利用主成分分析无须考察 6 门课程的具体成绩,只要对原始分数做恰当的线性组合,就可以找到两个指标(主成分),在不丢失重要信息的前提下,最大程度地区分学生的成绩.还可以直接从第一、第二主成分的得分出发,构造合适的评价函数对学生作进一步的评价,也可以用主成分得分值对学生进行聚类或判别,后者就是目前人工智能中常提及的主成分聚类分析和主成分判别分析.

评注 当各变量数据的量纲不同或取值数量级相差较大时,通常需要先将原始变量作标准化处理,于是协方差矩阵变换为相关系数矩阵.但是用这两个矩阵得到的主成分系数一般是不同的(复习题 1).如果原始数据的量纲相同且数量级相差不悬殊时,建议不要对数据做标准化处理.

$$\boldsymbol{x}=\boldsymbol{\mu}+\boldsymbol{A}\boldsymbol{f}+\boldsymbol{\varepsilon} \tag{7}$$

其中$\boldsymbol{\mu}=(\mu_1,\mu_2,\cdots,\mu_p)^{\mathrm{T}}$是$\boldsymbol{x}$的期望向量，$\boldsymbol{f}=(f_1,f_2,\cdots,f_m)^{\mathrm{T}}$称**公共因子向量**，$\boldsymbol{\varepsilon}=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_p)^{\mathrm{T}}$称**特殊因子向量**，均为不可观测的变量，$\boldsymbol{A}=(a_{ij})_{p\times m}$称为**因子载荷矩阵**，$a_{ij}$是变量$x_i$在公共因子$f_j$上的载荷，反映$f_j$对$x_i$的重要度.通常对模型(6)作如下假设：$f_j$互不相关且具有单位方差；$\varepsilon_i$互不相关且与$f_j$互不相关，$\mathrm{Cov}(\boldsymbol{\varepsilon})=\boldsymbol{\psi}$为对角阵.在这些假设下，由(7)式可得

$$\mathrm{Cov}(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{\psi},\quad \mathrm{Cov}(\boldsymbol{x},\boldsymbol{f})=\boldsymbol{A} \tag{8}$$

对因子模型(6)，每个原始变量x_i的方差都可以分解成共性方差h_i^2与特殊方差σ_i^2之和，其中$h_i^2=\sum_{j=1}^{m}a_{ij}^2$反映全部公共因子对变量$x_i$的方差贡献，$\sigma_i^2=D(\varepsilon_i)$(即$\boldsymbol{\psi}$的对角线上的元素)是特殊因子对$x_i$的方差贡献.显然，$\sum_{i=1}^{p}h_i^2=\sum_{i=1}^{p}\sum_{j=1}^{m}a_{ij}^2$是全部公共因子对$\boldsymbol{x}$总方差的贡献，令$b_j^2=\sum_{i=1}^{p}a_{ij}^2$，则$b_j^2$是公共因子$f_j$对$\boldsymbol{x}$总方差的贡献，$b_j^2$越大，$f_j$越重要，称$\dfrac{b_j^2}{\sum_{i=1}^{p}(h_i^2+\sigma_i^2)}$为$f_j$的**贡献率**.特别地，若$\boldsymbol{x}$的各分量已经标准化，则有$h_i^2+\sigma_i^2=1$，故$f_j$的贡献率为$\dfrac{b_j^2}{p}=\dfrac{\lambda_j}{p}$，其中$\lambda_j$是$\boldsymbol{x}$的相关系数矩阵的第$j$大特征根.

根据模型(7)，(8)式计算因子载荷矩阵$\boldsymbol{A}$的过程比较复杂，并且这个矩阵不唯一，只要$\boldsymbol{T}$为m阶正交矩阵，则$\boldsymbol{AT}$仍为该模型的因子载荷矩阵.矩阵$\boldsymbol{A}$左乘正交矩阵$\boldsymbol{T}$相当于作因子旋转，目的是找到简单结构的因子载荷矩阵，使得每个变量都只在少数的因子上有较大的载荷值，即只受少数几个因子的影响.通常，在因子分析模型建立后，还需要对每个样本估计公共因子的值，即所谓因子得分.对于以上详细的分析过程，读者可参考基础知识9-3.



基础知识9-3
因子分析的基本内容

学生成绩的因子分析模型 学生的分数数据矩阵$\boldsymbol{X}$的均值向量$\bar{\boldsymbol{x}}$和协方差矩阵$\boldsymbol{S}$已由(4)式给出，为了分析因子模型公共因子的存在，一般先计算出$\boldsymbol{X}$的相关系数矩阵

$$\boldsymbol{R}=\begin{pmatrix}1.000\,0 & 0.813\,3 & 0.834\,7 & -0.379\,5 & -0.561\,2 & -0.505\,4\\ 0.813\,3 & 1.000\,0 & 0.818\,8 & -0.273\,7 & -0.447\,4 & -0.356\,8\\ 0.834\,7 & 0.818\,8 & 1.000\,0 & -0.243\,7 & -0.438\,2 & -0.461\,1\\ -0.379\,5 & -0.273\,7 & -0.243\,7 & 1.000\,0 & 0.691\,6 & 0.573\,8\\ -0.561\,2 & -0.447\,4 & -0.438\,2 & 0.691\,6 & 1.000\,0 & 0.646\,3\\ -0.505\,4 & -0.356\,8 & -0.461\,1 & 0.573\,8 & 0.646\,3 & 1.000\,0\end{pmatrix} \tag{9}$$

从$\boldsymbol{R}$中的相关系数可以发现，变量x_1,x_2,x_3之间具有较强的正相关性，相关系数均在0.8以上，变量x_4,x_5,x_6之间也存在较强的正相关性，而这两组之间的相关性就没有组内的大，因此，有理由相信它们的背后都会有一个或多个共同因素(公共因子)在驱动，需要用因子分析方法来解释.

为了确定公共因子个数m，计算相关系数矩阵$\boldsymbol{R}$的特征根.$\boldsymbol{R}$的6个特征根按大小

排列为 $\lambda_1=3.709\ 9,\lambda_2=1.260\ 4,\lambda_3=0.436\ 5,\lambda_4=0.275\ 8,\lambda_5=0.170\ 3,\lambda_6=0.147\ 0$.前2个公共因子的累积贡献率为$(\lambda_1+\lambda_2)/6=0.828\ 4$,超过80%,因此,认为公共因子个数$m=2$是合适的.实际上,一个经验的确定$m$的方法,是将$m$定为$\boldsymbol{R}$中大于1的特征根个数,这与上面得到的结果一致.

MATLAB中利用数据矩阵$\boldsymbol{X}$进行因子分析的函数命令是factoran,调用格式为:

```
[lambda, psi, T, stats, F] = factoran(X, m)
```

其中输入参数 X 与主成分分析命令 princomp 相同,m 为公共因子个数,需满足$(p-m)^2\geqslant p+m$.输出参数 lambda 是$p\times m$的因子载荷矩阵,其第i行第j列的元素是第i变量在第j公共因子上的载荷,默认是用最大方差旋转法计算的;参数 psi 是p维列向量,对应p个特殊方差的最大似然估计;参数 T 为m阶(旋转后的)因子载荷旋转矩阵;参数 stats 是对原假设H_0(给定因子数m)做检验的统计量,其中p值若大于显著性水平α,则接受H_0;参数 F 是$n\times m$因子得分矩阵,每一行对应一个样本的m个公共因子的得分.这个函数命令也可以输入协方差矩阵或相关系数矩阵,调用格式见MATLAB帮助系统.

输入分数数据矩阵$\boldsymbol{X}$和$m=2$,调用 factoran 函数命令,在输出的检验信息中 stats.p=0.5060>0.05,可知在显著性水平$\alpha=0.05$下接受$H_0:m=2$.根据因子载荷矩阵的输出结果可以得到

$$\begin{cases} x_1^*=0.849\ 2f_1-0.362\ 8f_2+\varepsilon_1^* \\ x_2^*=0.863\ 7f_1-0.209\ 3f_2+\varepsilon_2^* \\ x_3^*=0.898\ 7f_1-0.204\ 3f_2+\varepsilon_3^* \\ x_4^*=-0.101\ 4f_1+0.807\ 3f_2+\varepsilon_4^* \\ x_5^*=-0.309\ 3f_1+0.819\ 6f_2+\varepsilon_5^* \\ x_6^*=-0.314\ 7f_1+0.668\ 6f_2+\varepsilon_6^* \end{cases} \tag{10}$$

其中x_i^*为x_i的标准化,即$x_i^*=\dfrac{x_i-\bar{x}_i}{\sqrt{s_{ii}}},i=1,2,\cdots,6,\bar{x}_i,s_{ii}$由(4)式给出.由此不难转换为原始变量$x_i$的因子分析模型.(10)式中特殊方差的估计也可以得到:$D(\boldsymbol{\varepsilon}^*)=(0.147\ 3,0.210\ 1,\ 0.150\ 5,\ 0.338\ 0,0.232\ 6,0.454\ 0)^{\mathrm{T}}$.

结果分析 在(10)式中第一公共因子f_1与数学分析、高等代数、概率论三门课程有很强的正相关,说明f_1对这3门课的解释力非常高,而对其他3门课就没那么重要了;第二公共因子f_2与微分几何、抽象代数和数值分析有很强的正相关,其解释恰好与f_1相反.由于数学分析、高等代数、概率论是数学系学生最重要的基础课,所以我们将f_1取名为"基础课因子",而微分几何、抽象代数与数值分析均为开卷考试,f_2又恰好是解释这3门课,为了区分考试类型的不同,不妨将f_2叫作"开闭卷因子".f_1和f_2的方差贡献率分别为$\lambda_1/6=0.6183$和$\lambda_2/6=0.2101$,f_1的影响要比f_2大得多.

每位学生的因子得分也可以在函数命令 factoran 的输出中得到,其结果列于表4.由于只有2个公共因子,以基础课因子f_1的得分为横轴,开闭卷因子f_2的得分为纵轴,画出因子得分的散点图,见图4.

表 4 公共因子得分表(全部数据见运行程序文件 9-17 后的输出结果)

学生序号	成绩总分	因子 f_1 得分	因子 f_2 得分
A_1	410	-0.775 0	0.057 1
A_2	359	-1.966 7	-1.350 9
⋮	⋮	⋮	⋮
A_{52}	449	0.161 7	1.284 6

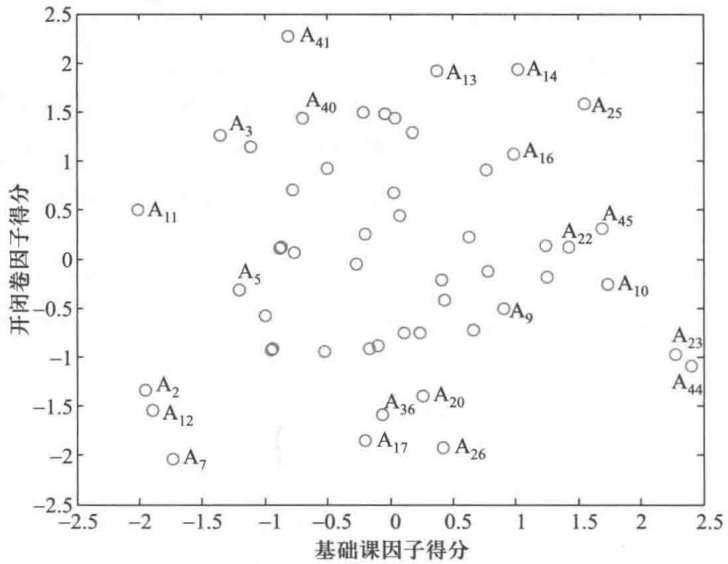


图 4 因子得分散点图

从图 4 可以发现,学生 A_{44}、A_{23}、A_{10}在 f_1 上有较高的得分,说明他们 3 门基础课的成绩表现非常好,而学生 A_{11}、A_2、A_{12}在 f_1 上的得分偏低,3 门基础课的表现不够好.学生 A_{41}、A_{14}、A_{13}在 f_2 上有较高的得分,他们较擅长于开卷考试,而学生 A_7、A_{26}、A_{17}的 f_2 的得分偏低,说明他们在开卷考试中表现不够理想.

以 2 个公共因子 f_1 和 f_2 的方差贡献率所占的比重加权,可以构造一个因子综合得分

$$F(f_1,f_2)=c_1f_1+c_2f_2 \tag{11}$$

这里权重 $c_1=\dfrac{\lambda_1}{\lambda_1+\lambda_2}=0.746\ 4$,$c_2=\dfrac{\lambda_2}{\lambda_1+\lambda_2}=0.253\ 6$,由(11)式计算出每位学生的因子综合得分值,并按得分值的大小对学生进行排序.为便于比较,将考试总分及排序一起列入表 5.

表 5 因子综合得分排名与排序结果(全部数据见运行程序文件 9-17 后的输出结果)

学生序号	成绩总分	总分排名	因子综合得分	因子综合得分排名
A_1	410	34	-0.564 0	39
A_2	359	51	-1.810 5	50
⋮	⋮	⋮	⋮	⋮
A_{52}	449	14	0.446 4	16

评注 多元统计中的主成分分析与因子分析方法的主要思想,都是

从表5可以看到,在总成绩排名前10名的同学中,有8人的因子综合得分的排名也在前10名,在总成绩排名后10名的同学中,有9人的因子综合得分的排名也在后10名;反过来,在因子综合得分排名前10名的同学中,有8人的总成绩的排名也在前10名,在因子综合得分排名后10名的同学中,也有8人的总成绩的排名在后10名;并且这两种排名次序差异不超过5名的比例为61.54%,具有较好的吻合度.

两种排名次序差异较大的如学生 A_3,总分排名为29,综合因子得分排名为44,相差15名,分析发现该学生的基础课因子 f_1 得分排名仅为48,尽管在3门开卷考试中的表现不错(因子 f_2 得分排名为10),由于综合得分中 f_1 占了约75%的权重,虽然总分排名不错,但因子综合得分就要差些了.再看一个极端的例子,如学生 A_{44},其总分排名第7,而因子综合得分排名高居第2,分析该学生的基础课因子 f_1 和开闭卷因子 f_2 的得分情况,发现在 f_1 上的得分排在第1名,而在 f_2 上的得分排在第45,说明他极不擅长开卷考试,好在他有极好的基础课考试成绩,使得因子综合得分跃升到了第2名.看来,利用因子综合得分排名,比传统的排名方法更具有科学性与参考价值.

采取常用的降维手段来降低整个数据的复杂程度.因子分析是从数据的协方差矩阵或相关系数矩阵出发,寻找潜在的起支配作用的因子,和主成分分析相比,由于因子分析可以使用因子旋转技术帮助解释因子,在数据解释方面更加有优势.这两种方法各有优缺点,都是目前数据挖掘与人工智能领域研究的重要方法.

复习题

1. 假定原始数据 $\boldsymbol{X}$ 的协方差阵 $\boldsymbol{S}=\begin{pmatrix}1 & 4\\ 4 & 100\end{pmatrix}$,若将原始数据 $\boldsymbol{X}$ 标准化,得到相关系数阵 $\boldsymbol{R}=\begin{pmatrix}1 & 0.4\\ 0.4 & 1\end{pmatrix}$.分别计算 $\boldsymbol{S}$ 和 $\boldsymbol{R}$ 的特征根和特征向量,构造相应的2个主成分,你会发现二者有很大差别.试做出解释.

2. 在制定服装标准过程中对100名成年男子的身材进行了测量,共6项指标:身高 x_1、坐高 x_2、胸围 x_3、臂长 x_4、肋围 x_5、腰围 x_6,样本相关系数阵为

$$\boldsymbol{R}=\begin{pmatrix}1 & 0.80 & 0.37 & 0.78 & 0.26 & 0.38\\ 0.80 & 1 & 0.32 & 0.65 & 0.18 & 0.33\\ 0.37 & 0.32 & 1 & 0.36 & 0.71 & 0.62\\ 0.78 & 0.65 & 0.36 & 1 & 0.18 & 0.39\\ 0.26 & 0.18 & 0.71 & 0.18 & 1 & 0.69\\ 0.38 & 0.33 & 0.62 & 0.39 & 0.69 & 1\end{pmatrix}$$

试给出主成分分析表达式,并对主成分做出解释.

3. 同第2题数据,试给出因子分析表达式,并对因子做出解释.

9.8 艾滋病疗法的评价及疗效的预测

艾滋病是当前人类社会最严重的瘟疫之一,从1981年发现以来它已经吞噬了数千万人的生命.人们正在艾滋病的研究、预防、治疗等各个领域进行着不懈的努力,作为用数学方法分析和解决实际问题关键途径的数学建模,也在这场艰苦的斗争中发挥着重要作用."艾滋病疗法的评价及疗效的预测"被选为2006年全国大学生数学建模竞赛B题(以下有删节):

艾滋病的医学全名为"获得性免疫缺陷综合征",英文简称AIDS,它是由艾滋病毒

(医学全名为"人体免疫缺陷病毒",英文简称 HIV)引起的.这种病毒破坏人的免疫系统,使人体丧失抵抗各种疾病的能力,从而严重危害人的生命.人类免疫系统的 CD4 细胞在抵御 HIV 的入侵中起着重要作用,当 CD4 被 HIV 感染而裂解时,其数量会急剧减少,HIV 将迅速增加,导致艾滋病发作.

艾滋病治疗的目的是尽量减少人体内 HIV 的数量,同时产生更多的 CD4,至少要有效地降低 CD4 减少的速度,以提高人体免疫能力.迄今为止,人类还没有找到能根治 AIDS 的疗法.

题目给出了美国艾滋病医疗试验机构公布的两组数据(ACTG320, 193A),要求:

1) 利用 ACTG320 数据,预测继续治疗的效果,或者确定最佳治疗终止时间(继续治疗指在测试终止后继续服药,如果认为继续服药效果不好,则可选择提前终止治疗).

2) 利用 193A 数据,评价 4 种疗法的优劣(仅以 CD4 为标准),并对较优的疗法预测继续治疗的效果,或者确定最佳治疗终止时间.

数据文件 9-5
艾滋病疗法 ACTG320 数据

表 1 ACTG320 数据:服用 3 种药物的 300 多名患者每隔几周测试的 CD4 和 HIV 浓度(全部数据见数据文件 9-5)

患者序号	CD4 时间/周	CD4 浓度	HIV 时间/周	HIV 浓度
23 424	0	178	0	5.5
	4	228	4	3.9
	8	126	8	4.7
	25	171	25	4
	40	99	40	5
23 425	0	14	0	5.3
	4	62	4	2.4
	9	110	9	3.7
	23	122	23	2.6
	40	320		
⋮	⋮	⋮	⋮	⋮

数据文件 9-6
艾滋病疗法 193A 数据

表 2 193A 数据:1300 多名患者随机地分为 4 组,每组按一种疗法服药,每隔约 8 周测试的 CD4 浓度(取 lg(CD4+1))(全部数据见数据文件 9-6)

患者序号	疗法	年龄	CD4 时刻/周	CD4 浓度
1	2	36.427 1	0	3.135 5
			7.271 4	3.044 5
			15.271 4	2.772 6
			23.271 4	2.833 2
			32.271 4	3.218 9
			40	3.044 5

续表

患者序号	疗法	年龄	CD4 时刻/周	CD4 浓度
2	4	47.846 7	0	3.068 1
			8	3.891 8
			16	3.970 3
			23	3.610 9
			30.714	3.332 2
			39	3.091 0
⋮	⋮	⋮	⋮	⋮

问题分析 查阅相关资料可知,目前对 AIDS 的治疗以针对 HIV 的高效抗逆转录病毒疗法为主.临床上需要改变治疗方案的原因有:最新临床试验结果提示,感染者正在使用的不是最佳治疗方案;感染者虽然采用高效的治疗方案,但 CD4 细胞数量继续下降;患者有临床进展表现或严重的毒副作用,使之难以坚持治疗.

评注 目前人们对艾滋病的发展和治疗的机理尚不完全掌握,本题主要依靠收集数据和统计方法进行疗法的评价及疗效的预测.

HIV 浓度的测试成本很高,而 CD4 细胞数量的降低是免疫缺陷进展的直接标志,被认为是 HIV 感染状态最重要的参考,也是各类疗法的有效评价指标.

艾滋病疗法的评价标准是降低 HIV 病毒,提升 CD4 细胞.治疗过程中如果 HIV 不再降低,CD4 不再升高,就应及时终止治疗,否则可以继续治疗.

数据处理 分析数据文件 9-5,9-6 的数据可知,大多数是一位病人在 5 或 6 个时间点的测试记录,如果测试时间点过少(只有 2 或 3 个),可将该患者删除.根据前后记录可辨别的、明显的错误数据予以删除.

为了消除患者的初始状态(时间 $t=0$ 的 CD4 和 HIV 浓度)对模型的影响,可以取各位患者每次的测量值与初始值的比值或差值,作为以下分析、建模的依据.也可以先按照患者的初始状态分类(如轻度、中度、重度),然后对于每一类患者进行分析、建模.

预测治疗效果或确定治疗终止时间

数据分析 为了利用 ACTG320 数据预测治疗效果或确定治疗终止时间,随机选取若干患者(如 20 个)画出 CD4 和 HIV 浓度变化图形,如图 1.可以发现,CD4 浓度大致有先增后减的趋势,HIV 浓度有先减后增的趋势,启示我们应建立浓度对时间的二次回归模型:$y=b_0+b_1t+b_2t^2$, 其中 y 是 CD4 (或 HIV) 浓度, t 是时间.

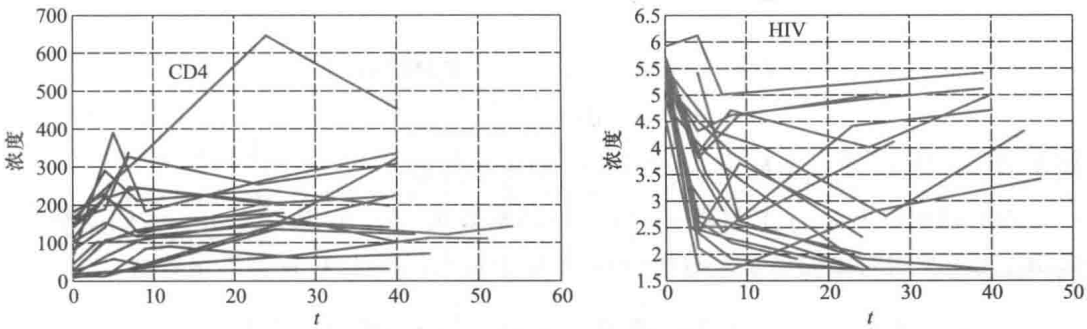


图 1 20 个患者的 CD4 和 HIV 浓度变化

随机取3位患者的CD4浓度拟合二次回归模型，如图2.可以发现，彼此差异较大.

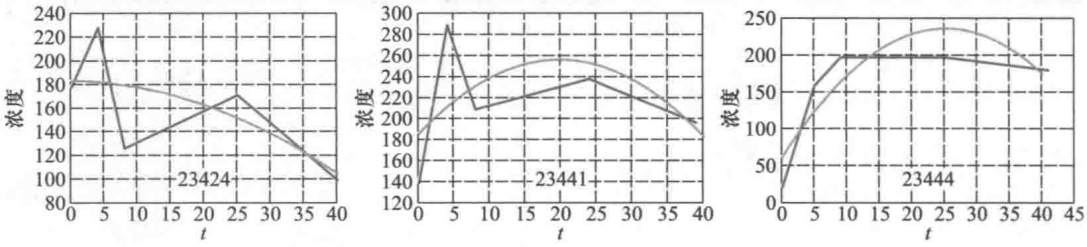


图2 3个患者CD4浓度拟合的二次回归曲线(折线为原始数据)

模型建立 分别用 t_{ij}, y_{ij}表示表1中第 i 个患者第 j 次测量的时间(周)和CD4浓度，二者之间可以建立以下几种形式的回归模型：

1) 纵向数据回归模型 首先写出 y_{ij}对 t_{ij}的二次回归函数

$$y_{ij}=b_{0i}+b_{1i}t_{ij}+b_{2i}t_{ij}^2+\varepsilon_{ij},\quad i=1,\cdots,n,\quad j=1,\cdots,n_i \tag{1}$$

其中 b_{0i}, b_{1i}, b_{2i}是回归系数，ε_{ij}是随机误差，假定服从零均值、方差为常数 σ^2 的正态分布，n 是患者数，n_i 是第 i 个患者的测量次数.需要注意的是，与普通的回归模型相比，模型(1)的 b_{0i}, b_{1i}, b_{2i}多了下标 i，用于描述不同患者的CD4浓度具有不同的二次曲线系数(如图2那样)，它们也应视为随机变量.

为了将患者整体CD4浓度的二次曲线系数从 b_{0i}, b_{1i}, b_{2i}中分离出来，令

$$b_{ki}=b_k+\eta_{ki},\quad k=0,1,2 \tag{2}$$

其中 b_k 是患者整体的固定效应系数(与哪个患者无关)，η_{ki}是随患者 i 变化的随机效应系数，假定服从零均值、方差为常数 d_k^2 的正态分布，且 η_{ki}之间(对 k)相互独立.

将(2)代入(1)式得到回归系数分解后的纵向数据回归模型：

$$y_{ij}=b_0+b_1t_{ij}+b_2t_{ij}^2+\eta_{0i}+\eta_{1i}t_{ij}+\eta_{2i}t_{ij}^2+\varepsilon_{ij} \tag{3}$$

记

$$\boldsymbol{Y}_i=\begin{bmatrix}y_{i1}\\y_{i2}\\\vdots\\y_{in_i}\end{bmatrix},\quad \boldsymbol{X}_i=\begin{bmatrix}1&t_{i1}&t_{i1}^2\\1&t_{i2}&t_{i2}^2\\\vdots&\vdots&\vdots\\1&t_{in_i}&t_{in_i}^2\end{bmatrix},\quad \boldsymbol{b}=\begin{bmatrix}b_0\\b_1\\b_2\end{bmatrix}$$

$$\boldsymbol{D}=\begin{bmatrix}d_0^2&0&0\\0&d_1^2&0\\0&0&d_2^2\end{bmatrix},\quad \boldsymbol{V}_i=\boldsymbol{X}_i\boldsymbol{D}\boldsymbol{X}_i^{\mathrm{T}}+\sigma^2\boldsymbol{I}_{n_i} \tag{4}$$

则 $\boldsymbol{Y}_i$ 服从均值向量为 $\boldsymbol{X}_i\boldsymbol{b}$、方差矩阵为 $\boldsymbol{V}_i$ 的正态分布.

为了利用数据 t_{ij}, y_{ij}估计模型(3)、(4)的系数 b_0, b_1, b_2 及 η_{0i}, η_{1i}, η_{2i}的方差 d_0^2, d_1^2, d_2^2，无法利用简单的最小二乘法，而需要应用最大似然法.$\boldsymbol{Y}_i$ 的似然函数为

$$L(\boldsymbol{b},\boldsymbol{D},\sigma^2)=\prod_{i=1}^{n}\left\{(2\pi)^{-\frac{n_i}{2}}|\boldsymbol{V}_i|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{b})^{\mathrm{T}}\boldsymbol{V}_i^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{b})\right)\right\} \tag{5}$$

求解似然函数 L 的最大值点，即得系数 $\boldsymbol{b}$ 及 $\boldsymbol{D}, \sigma^2$ 的估计值，用 SAS 等软件可以方便地实现.

固定效应系数 b_k 反映的是全体患者的总体效应，利用 b_k 可以预测平均意义下的最佳治疗终止时间(或继续治疗).随机效应系数 η_{ki} 的方差 d_k^2 反映随不同患者而异的分散性，在正态分布假定下可以得到 b_k 属于任意区间的概率，从而在一定置信度下给出最佳治疗终止时间的置信区间.

2) 个体回归模型　形式与(1)式相同，但不分解回归系数，相当于 n 个二次回归模型.对每个患者 $i(=1,\cdots,n)$ 由数据 t_{ij}, y_{ij} 用最小二乘法估计 b_{0i}, b_{1i}, b_{2i}，再计算 n 个患者 b_{0i}, b_{1i}, b_{2i} 的平均值和方差，作为纵向数据回归模型 b_0, b_1, b_2 和 d_0^2, d_1^2, d_2^2 的近似.

因为每位患者只有 5 或 6 个时间点的测试记录，数据量过少使得每个模型的回归系数估计的精度都较低，由此得到的平均值和方差的可靠性也不高.

3) 总体回归模型　若忽略患者之间的差异，可以将(1)式简化为普通的回归模型：

$$y_{ij}=b_0+b_1t_{ij}+b_2t_{ij}^2+\varepsilon_{ij},\quad i=1,\cdots,n,\quad j=1,\cdots,n_i \tag{6}$$

由全部数据 t_{ij}, y_{ij} 用最小二乘法估计(6)式的系数 b_0, b_1, b_2，作为纵向数据回归模型系数的近似值.

由于对(6)式作计算时，将患者的分散性造成的随机效应方差与模型的随机误差混合在一起，因而剩余方差将很大，相当于纵向数据回归模型中 $\boldsymbol{D}$ 和 σ^2 的总和，这就降低了总体回归模型的有效性.

一种代替用全部数据直接估计总体回归模型(6)式的办法是，先在每个时间点上对患者的 CD4 浓度取平均，再由这些平均浓度用最小二乘法估计系数 b_0, b_1, b_2.这种看似简单的方法有什么问题吗(有相当一部分竞赛论文正是这样做的)?

全部数据 t_{ij}, y_{ij} 是 300 多名患者在若干时间点上的 CD4 浓度，观察发现，多数患者的测试时间点是 0,4,8,24,…,40，少数患者是 5,9,23,…，还有其他一些个别的时间点.在这些时间点上对 CD4 浓度取平均后，有的是上百个患者的平均浓度，也有的是几个患者的平均浓度，用这些平均值与用原始数据得到的回归系数可能会有较大的差别.

怎样能够使得利用(对自变量的每个数值取因变量的)平均值得到的回归系数与利用原始数据得到的回归系数相同呢? 办法是对每个平均值用它所“代表”的原始数据的个数加权.可以证明，用加权后的平均值作普通的最小二乘拟合，得到的系数 b_0, b_1, b_2 与用原始数据相同.

结果分析　根据 ACTG320 数据和上述 3 个模型得到的 CD4 与 HIV 浓度随时间的变化规律可用 $y=b_0+b_1t+b_2t^2$ 表示，二次曲线示意图如图 3 所示.

根据得到的系数 b_0, b_1, b_2 可知，对 CD4 有 $b_2<0$, $b_1>0$，当 $t=-b_1/2b_2$ 时浓度达到最大；对 HIV 有 $b_2>0$, $b_1<0$，当 $t=-b_1/2b_2$ 时浓度达到最小.示意图的结果提示，大致应在 25~30 周终止治疗.

与采用原始数据的结果相比，当模型的 y_{ij} 取各位患者每次的测量值与初始值的比值或差值时，所得结果有较大差别，后者合理性更高.

根据个体回归模型的结果，可以将 300 多位患者中 CD4 浓度曲线 $b_2<0$ 且最大点

评注　对于这道题目提供的测试数据，与传统的回归分析相比，纵向数据(或称追踪数据)回归模型是真正有效的处理方法，这个方法在经济、社会、心理、医疗等领域有着广泛的应用，虽然它在理论上需要稍多一些统计知识，但是从应用的角度并不难掌握.

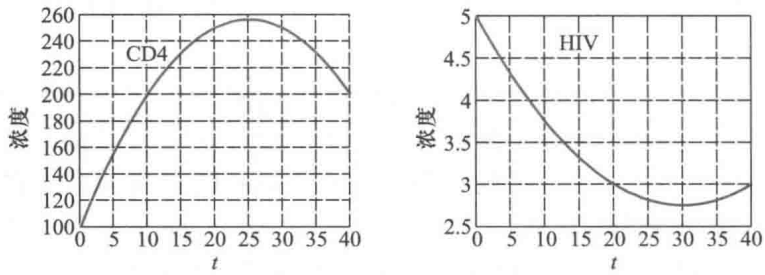


图3 由ACTG320数据得到的CD4和HIV浓度的二次曲线示意图

$t<40$ 的比例，作为及时终止治疗的概率（同时考虑HIV浓度曲线 $b_2>0$ 且最小点 $t<40$）. 相反情况的比例可作为继续治疗的概率.

评价4种疗法的优劣

数据分析 与ACTG320数据（数据文件9-5）相比，193A数据（数据文件9-6）只有CD4浓度，而增加了疗法和患者年龄2个因素.

对于每种疗法随机取若干患者（年龄尽量接近），画出CD4浓度随时间变化的图形（如图1那样），图形显示采用疗法1~3的患者CD4浓度变化不大且略有下降，而疗法4有先增后减的趋势，由此可以对4种疗法建立浓度对时间的一次回归模型，其中疗法4可建立二次模型，估计并比较一次项系数，以便对4种疗法的优劣给出定量的评价.

模型及其求解 为简便起见用总体回归模型给以说明.

对4种疗法分别建模 只需对(6)式作如下修改：去掉 t_{ij} 的二次项（疗法4可保留），增加年龄变量 x_i（患者 i 的年龄）.具体做法不再详述.

对4种疗法统一建模 引入0-1变量 z_{ik}，若患者 i 采用疗法 k，令 $z_{ik}=1$（$k=1,2,3,4$），否则 $z_{ik}=0$.回归模型可表示为

$$y_{ij}=a_0x_i+\sum_{k=1}^{4}a_kz_{ik}+\sum_{k=1}^{4}b_kz_{ik}t_{ij}+b_5z_{i4}t_{ij}^2+\varepsilon_{ij} \tag{7}$$

利用表2的全部数据估计出(7)式的系数，从而得到4种疗法的回归曲线，其中疗法1~3是直线，疗法4是二次曲线，如图4所示.直观地看，4种疗法的优劣按照4,3,2,1排序，并且疗法4在30周之前明显优于其他.

实际的计算结果表明，4种疗法分别建模和统一建模得到的系数 b_k 的估计值相同，但是后者 b_k 的置信区间较短，模型的有效性提高，这是由于统一建模所用的数据量远大于分别建模的缘故.

可以采用假设检验和方差分析的方法，从统计意义上辨别各种疗法是否存在显著差异.

比较两种疗法的假设检验 为比较疗法1和疗法2，应检验它们的一次项系数，即(7)式的 b_1 与 b_2 有无显著性差异，可以仅引入一个0-1变量 z_i，若患者 i 采用疗法1，令 $z_i=1$，否则（即疗法2）$z_i=0$.将回归模型表示为（略去 x_i 项）

$$y_{ij}=c_0+c_1t_{ij}+c_2z_i+c_3z_it_{ij}+\varepsilon_{ij} \tag{8}$$

并作如下假设检验：$H_0:c_3=0$；$H_1:c_3\neq0$.

利用疗法1,2的数据计算 c_3 的置信区间，若区间包含零点，则接受 H_0.显然，(8)式

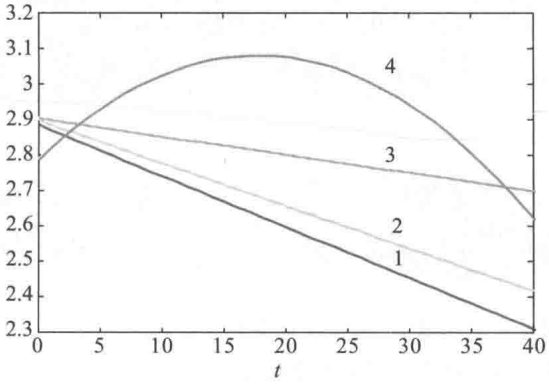


图 4　4 种疗法统一模型的回归曲线

的 $c_3=0$ 等价于(7)式的 b_1 与 b_2 无显著性差异.

同样的方法可用于比较疗法 1 和疗法 3、疗法 2 和疗法 3.

假设检验的结果显示,b_1 与 b_2 无显著性差异,而 b_3 与 b_1,b_2 有显著性差异,即统计意义上的结论是,疗法 3 优于疗法 1 和疗法 2,疗法 1,2 并无差别.

如果要比较疗法 4 和疗法 3,不妨对这两种疗法建立如(8)式那样的一次回归模型,可以得出疗法 4 优于疗法 3 的结论.

4 种疗法一起比较的单因素方差分析　4 种疗法都建立一次回归模型,以疗法为单因素对一次项系数进行方差分析,即作假设检验:H_0:4 个回归系数相等;H_1:4 个回归系数不全相等.如果检验结果是拒绝 H_0,那么仍需进行疗法的两两比较.

评注　以上几种统计方法都是根据一次回归模型的一次项系数,即 CD4 浓度的变化率的正负和大小,来比较 4 种疗法的优劣,对于这个题目提供的数据而言,基本上是合理的.如果数据需要用到二次回归模型,则优劣的比较就复杂多了.

复习题

在一元回归模型中,若对自变量的每个数值 $t_i(i=1,2,\cdots,m)$,因变量都有 k 个取值 $y_{ij}(j=1,2,\cdots,k)$,其平均值记作 y_i,证明:用 t_i 和 y_i 得到的回归系数与用原始数据 t_i 和 y_{ij} 得到的回归系数相同;但若对每个 t_i 因变量 y_{ij} 的个数不同,即 $j=1,2,\cdots,k_i$,而 k_i 不全相等,则得不到这个结论.

第 9 章训练题

1. 电影院调查电视广告费用和报纸广告费用对每周收入的影响,得到数据如下表,建立回归模型并进行检验,诊断异常点的存在并进行处理.

每周收入	96	90	95	92	95	95	94	94
电视广告费用	1.5	2.0	1.5	2.5	3.3	2.3	4.2	2.5
报纸广告费用	5.0	2.0	4.0	2.5	3.0	3.5	2.5	3.0

2. 营养学家为研究食物中蛋白质含量对婴儿生长的影响,按照食物中蛋白质含量的高低,调查了两组两个月到三岁婴儿的身高(cm),见下表.

高蛋白食物组

年龄	0.2	0.5	0.8	1.0	1.0	1.4	1.8	2.0	2.0	2.5	2.5	2.7	3.0
身高	54	55	63	66	69	73	82	83	80	91	93	94	94

低蛋白食物组

年龄	0.2	0.4	0.7	1.0	1.0	1.3	1.5	1.8	2.0	2.0	2.4	2.8	3.0
身高	51	52	55	61	64	65	66	69	68	69	72	76	77

(1) 分别用两组数据建立蛋白质高、低含量对婴儿身高的回归模型,解释所得结果.

(2) 怎样检验蛋白质含量的高低对婴儿的生长有无显著影响? 检验结果如何?

3. 在有氧锻炼中人的耗氧能力 y(单位:mL/(min·kg)) 是衡量身体状况的重要指标,它可能与以下因素有关:年龄 x_1,体重 x_2(单位:kg),1 500 m 跑用的时间 x_3(单位:min),静止时心率 x_4(单位:次/min),跑步后心率 x_5(单位:次/min).对 24 名 38 至 57 岁的志愿者进行了测试,结果如下表(全部数据见数据文件 9-7).试建立耗氧能力 y 与诸因素之间的回归模型.

数据文件 9-7
第9章训练题3 耗氧能力

序号	y	x_1	x_2	x_3	x_4	x_5
1	44.6	44	89.5	6.82	62	178
2	45.3	40	75.1	6.04	62	185
⋮	⋮	⋮	⋮	⋮	⋮	⋮
24	54.7	50	70.9	5.35	48	146

(1) 若 $x_1 \sim x_5$ 中只许选择 1 个变量,最好的模型是什么?

(2) 若 $x_1 \sim x_5$ 中只许选择 2 个变量,最好的模型是什么?

(3) 若不限制变量个数,最好的模型是什么? 你选择哪个作为最终模型,为什么?

(4) 对最终模型观察残差,有无异常点? 若有,剔除后如何?

4. 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,设计了一个药物试验,给 24 名患有同种病痛的患者使用这种新止痛剂的以下 4 个剂量中的某一个:2,5,7 和10(g),并记录每个患者病痛明显减轻的时间(单位:min).为了解新药的疗效与患者性别和血压有什么关系,试验过程中研究人员把患者按性别及血压的低、中、高三档平均分配来进行测试.通过比较每个患者血压的历史数据,从低到高分成 3 组,分别记作 0.25,0.50 和 0.75.试验结束后,公司的记录结果见下表(性别以 0 表示女,1 表示男)(全部数据见数据文件 9-8).请你为公司建立一个模型,根据患者用药的剂量、性别和血压组别,预测出服药后病痛明显减轻的时间.

数据文件 9-8
第9章训练题4 止痛剂疗效

患者序号	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
⋮	⋮	⋮	⋮	⋮
24	5	10	1	0.75

5. 调查了 12 名 6 至 12 岁正常儿童的体重、身高和年龄,如下表.

序号	体重/kg	身高/m	年龄	序号	体重/kg	身高/m	年龄
1	27.1	1.34	8	7	30.9	1.39	10
2	30.2	1.49	10	8	27.8	1.21	9
3	24.0	1.14	6	9	29.4	1.26	10
4	33.4	1.57	11	10	24.8	1.06	6
5	24.9	1.19	8	11	36.5	1.64	12
6	24.3	1.17	7	12	29.1	1.44	9

(1) 建立直接用身高 x_1 和年龄 x_2 预测儿童体重 y 的回归模型.

(2) 考虑 $x_3=x_1^2$, $x_4=x_2^2$, $x_5=x_1x_2$ 等候选变量,用逐步回归建立预测儿童体重的模型.

6. 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.

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

7. 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 3 414 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(全部数据见数据文件 9-9):

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),说明教师的薪金与哪些变量的关系密切,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,尝试将它们的平方项或交互项加入到模型中,建立更好的模型.



数据文件 9-9
第9章训练题7教师薪金

	Z	X_1	X_2	X_3	X_4	X_5	X_6	X_7
1	998	7	0	0	0	0	0	0
2	1 015	14	1	1	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
90	2 000	464	1	1	2	1	1	0

8. 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 $y_t = \dfrac{L}{1+ae^{-kt}}$,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			

9. 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

10. 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?

11. 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
	10	1	9	10%
27.2	8	0	8	0%
	9	1	8	11.1%
	10	7	3	70%
27.7	6	4	2	66.7%
	8	6	2	75%
	13	13	0	100%
28.3	9	6	3	66.7%
	8	7	1	87.5%
	10	7	3	70%
28.4	8	5	3	62.5%
	9	7	2	77.8%
	11	10	1	90.9%
29.9	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.



数据文件 9-10
第 9 章训练题 12
男子径赛纪录

12. 下表列出了 2005 年 54 个国家或地区男子径赛纪录的数据(全部数据见数据文件 9-10).

(1) 求标准化变量的前两个主成分，并给出主成分及由这两个主成分解释的(标准化)样本总方差的累积百分比，并解释这两个主成分.

(2) 把这 54 个国家或地区按它们在第一主成分的得分排序，这种排序与你最初对不同国家或地区的运动水平的看法是否一致？

(3) 对标准化变量作因子分析，解释公共因子的含义.

(4) 根据因子得分说明，哪些国家或地区的短跑项目具有优势，哪些国家或地区的长跑项目更具有优势.

54 个国家或地区男子径赛纪录　　单位：s

序号	国家或地区	100 m	200 m	400 m	800 m	1500 m	5000 m	10000 m	马拉松
1	阿根廷	10.23	20.37	46.18	106.2	220.8	799.8	1659	7774.2
2	澳大利亚	9.93	20.06	44.38	104.4	211.8	775.8	1651.8	7650.6
3	奥地利	10.15	20.45	45.8	106.2	214.8	795.6	1663.2	7933.2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
54	美国	9.78	19.32	43.18	102.6	207.6	778.2	1632.6	7522.8



更多案例……

9-1　牙膏的销售量

9-2　教学评估

第 10 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）如何做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，决策者面临的主要问题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做出的、还是按一定先后顺序做出的，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.

10.1　点球大战

在很多体育比赛项目中，双方队员需要斗智斗勇，相互博弈，此时博弈模型是分析双方策略的重要工具.本节以点球大战为例进行说明.

问题背景　足球比赛中的点球大战紧张刺激、扣人心弦，我们会对扑住点球的守门员交口称赞，也会为未能进球的罚球队员扼腕惋惜.假设罚点球时不考虑罚球队员把球踢向中路以及守门员停在球门中间扑救的情况，那么罚球队员有两种基本策略：把球踢向左侧或者右侧；守门员也有两种基本策略：扑向左侧或者右侧.实战中罚球队员和守门员几乎是同时做出射门和扑球的决策，因为罚球队员踢出的球速非常快，守门员若想看到球的方向后再扑救，肯定来不及.我们可能会认为，罚球队员会完全随机地把球踢向左侧或右侧，即踢向左侧或右侧的概率各为 50%；守门员也会完全随机地向左或向

右扑球,即扑向左侧或右侧的概率各为50%.

果真如此吗?如果不是,射门方向和扑球方向应该有什么规律?[14][24]

问题分析 为了回答这个问题,应该先弄清楚射门方向和扑球方向确定后,进球的概率分别是多少.为方便起见,方向以其中一人如防守队员的位置为基准.通常会认为,只要射门方向和扑球方向一致(都是左侧或者都是右侧),进球的概率就一样;只要方向不一致,进球的概率也一样.如果真是这样,根据直觉就可以知道,罚球队员应该完全随机地把球踢向左侧或右侧,守门员也没有理由不完全随机地向左或向右扑球.但有人通过对实战中大约1400次罚球的统计分析,得到的进球概率(不妨称为经验概率)却不是这样,见表1.

表1 根据实战统计得到的进球概率(依赖于射门和扑球方向)

罚球队员	守门员	
	扑向左侧	扑向右侧
踢向左侧	0.58	0.95
踢向右侧	0.93	0.70

由表1不仅看到"射门和扑球方向不一致时进球概率要比方向一致时大得多"这个完全合乎逻辑的现象,而且可以注意到,射门和扑球方向都是右侧时的进球概率比都是左侧时有显著的增加.基于这样的统计数据,射门方向和扑球方向应该不会是完全随机的.显然,能否进球及进球的概率取决于两个队员(决策者)的决策,二者的决策是相互影响的.假设所有决策者都知道以上信息(博弈论中一般称为**共同知识**),两名队员同时做出决策,这样的博弈称为**完全信息的静态博弈**.

模型建立 完全信息的静态博弈模型包括三个基本要素:**参与人**(player,也译为局中人,即决策者)的集合,每个参与人的**策略空间**(决策变量的取值范围),以及每个参与人的**效用函数**(决策的目标函数).

点球大战的博弈中参与人集合可以用 $N=\{1,2\}$ 表示,1表示罚球队员,2表示守门员.罚球队员可能的策略记作 $a_1 \in A_1=\{1,2\}$,1,2分别表示把球踢向左侧和右侧;守门员可能的策略记作 $a_2 \in A_2=\{1,2\}$,1,2分别表示扑向左侧和右侧.A_1 和 A_2 分别是参与人1和2的策略空间.

对于双方每一种策略组合 (a_1,a_2),用 $u_1(a_1,a_2)$ 表示罚球队员一次射门的期望得分(进球得1分,不进球得0分),实际上等于进球概率,可作为罚球队员的效用函数.由表1,$u_1(a_1,a_2)$ 可以用矩阵

$$\boldsymbol{M}=(m_{ij})_{2\times 2}=\begin{pmatrix}0.58 & 0.95\\ 0.93 & 0.70\end{pmatrix} \tag{1}$$

表示,即 $u_1(i,j)=m_{ij}$,$\boldsymbol{M}$ 称为罚球队员的**支付矩阵**(payoff matrix,也译为收益矩阵、赢得矩阵等).类似地,守门员的效用函数用 $u_2(a_1,a_2)$ 表示,显然有 $u_2(a_1,a_2)=-u_1(a_1,a_2)$,于是其支付矩阵是 $-\boldsymbol{M}$.在对策中双方都力求通过决策行动使己方的效用函数最大化.本例是两人之间的博弈,且具有完全竞争性质,即一方所得正是对手所失,这种对策一般称为零和博弈.

纯策略 Nash 均衡 用 a_1^*,a_2^* 分别表示罚球队员和守门员应该选择的策略,由于

双方都希望通过决策使自己的效用函数达到最大，所以 a_1^*, a_2^* 应该满足

$$u_1(a_1^*, a_2^*) \geqslant u_1(a_1, a_2^*), u_2(a_1^*, a_2^*) \geqslant u_2(a_1^*, a_2), a_1 \in \{1,2\}, a_2 \in \{1,2\} \quad (2)$$

在博弈论中这样的策略组合 $\boldsymbol{a}^* = (a_1^*, a_2^*)$ 称为（纯策略）Nash 均衡（Nash equilibrium），其含义是，如果对方不改变策略，每一方的策略选择都是最优的. 每一方都不会偏离这个策略，因为单方面地偏离不能使自己的效用得到提升.

对于由(1)给出的效用函数，通过枚举不难验证这样的策略组合即(纯策略)Nash均衡不存在：对于(2)式，(1,2)和(2,1)满足关于 u_1 的不等式，即 $u_1(1,2)=0.95>u_1(2,2)=0.70, u_1(2,1)=0.93>u_1(1,1)=0.58$；(1,1)和(2,2)满足关于 u_2 的不等式，即 $u_2(1,1)=-0.58>u_2(1,2)=-0.95, u_2(2,2)=-0.70>u_2(2,1)=-0.93$，但(1,2)，(2,1)，(1,1)，(2,2)都不能同时满足(2)的所有关系. 以上结果与“射门和扑球方向不一致只对罚球队员有利、方向一致只对守门员有利，因而不存在对双方均有利的策略”这样的常识完全符合.

对于点球大战而言，什么情况下才会存在(纯策略)Nash均衡呢？让我们虚拟一个进球概率，将表1中的0.95改为0.65，则双方的效用函数如表2所示(其中两个数字分别是罚球队员和守门员的得分)，可以验证，对于策略组合 $\boldsymbol{a}^* = (a_1^*, a_2^*) = (2,2)$，由于 $u_1(2,2)=0.70>u_1(1,2)=0.65$，使得双方的效用函数 $u_1(a_1^*, a_2^*)$ 和 $u_2(a_1^*, a_2^*)$ 满足(2)式，即 $\boldsymbol{a}^*$ 为纯策略 Nash 均衡. 实际上我们虚拟了这样一种情况：即便守门员扑向右侧，罚球队员踢向左侧时的进球概率还小于踢向右侧时的进球概率(可能是一位向左踢不准的队员)，于是罚球队员踢向右侧时得分更高(不论守门员扑向哪一侧)，而守门员自然是扑向右侧更有利，所以 $\boldsymbol{a}^* = (2,2)$ 是对双方最优的策略.

表2 双方的效用函数(根据虚拟的进球概率)

罚球队员	守门员	
	扑向左侧	扑向右侧
踢向左侧	0.58, -0.58	0.65, -0.65
踢向右侧	0.93, -0.93	0.70, -0.70

混合策略 Nash 均衡 对于不存在纯策略 Nash 均衡的博弈问题，如对(1)式的 $\boldsymbol{M}$ 给出的效用函数，可以考虑双方随机地采取行动，即双方都对每一种策略赋予一定的概率，形成混合策略. 设罚球队员采取策略 i 的概率为 $p_i (i=1,2)$，守门员采取策略 j 的概率为 $q_j (j=1,2)$，记行向量 $\boldsymbol{p}=(p_1,p_2)$，$\boldsymbol{q}=(q_1,q_2)$，满足

$$0 \leqslant p_i \leqslant 1, \ \sum_{i=1}^{2} p_i = 1, \ 0 \leqslant q_i \leqslant 1, \ \sum_{i=1}^{2} q_i = 1 \quad (3)$$

的概率向量 $\boldsymbol{p}, \boldsymbol{q}$ 分别构成罚球队员和守门员的混合策略空间，记为 S_1, S_2.

在混合策略下双方的效用函数用期望效用定义，记作

$$U_1(\boldsymbol{p},\boldsymbol{q}) = \sum_{i=1}^{2}\sum_{j=1}^{2} p_i m_{ij} q_j = \boldsymbol{p}\boldsymbol{M}\boldsymbol{q}^{\mathrm{T}}, \quad U_2(\boldsymbol{p},\boldsymbol{q}) = -U_1(\boldsymbol{p},\boldsymbol{q}) \quad (4)$$

将(2)式中的效用 u_1, u_2 改成期望效用 U_1, U_2，可以类似地定义(混合策略)Nash均衡.

显然，前面 A_1 的每一个策略 $i(=1,2)$ 是罚球队员混合策略的特例(采取策略 i 的概率为1)，称为纯策略，A_1 称为纯策略空间. A_2 类似.

模型求解 罚球队员希望最大化期望效用 $U_1(\boldsymbol{p},\boldsymbol{q})$，所面临的决策问题是

$$\max_{\boldsymbol{p}\in S_1} U_1(\boldsymbol{p},\boldsymbol{q})=\boldsymbol{pMq}^{\mathrm{T}} \tag{5}$$

而守门员希望最大化 $U_2(\boldsymbol{p},\boldsymbol{q})$，面临的决策等价于

$$\min_{\boldsymbol{q}\in S_2} U_1(\boldsymbol{p},\boldsymbol{q})=\boldsymbol{pMq}^{\mathrm{T}} \tag{6}$$

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4章介绍的数学规划方法直接求解.

罚球队员怎样在不能控制 $\boldsymbol{q}$ 的情况下使 $\boldsymbol{pMq}^{\mathrm{T}}$ 最大？守门员怎样在不能控制 $\boldsymbol{p}$ 的情况下使 $\boldsymbol{pMq}^{\mathrm{T}}$ 最小？双方都希望使自己的效用最大化，而且他们都知道，不管己方怎么做，对方总是会采取策略使己方的效用尽量小，所以己方在采用一定的策略时得到的效用，总是可能得到的效用当中最小的那个，而最优策略应该使得己方最小的效用达到最大.

由此，罚球队员所面对的决策问题(5)可以转化为

$$\max_{\boldsymbol{p}\in S_1} \min \boldsymbol{pM} \tag{7}$$

注意 $\boldsymbol{pM}$ 是一个向量，这里 min 是对 $\boldsymbol{pM}$ 的所有元素取极小.

类似地，守门员所面对的决策问题(6)转化为

$$\min_{\boldsymbol{q}\in S_2} \max \boldsymbol{Mq}^{\mathrm{T}} \tag{8}$$

求解得到的最优解(精确到三位小数)为 $p_1=0.383, p_2=0.617$，最优值为 0.796.类似地得到(8)的最优解为 $q_1=0.417, q_2=0.583$，最优值也是 0.796.理论上可以证明，对于双方效用 U_1, U_2 之和为0的零和博弈，(7)(8)两个优化问题的最优值是一样的.

这个最优解(策略组合)具有如下重要性质：没有任何一方通过单方面地偏离这一组合中自己的策略，可以提高自己的期望效用.也就是说，它就是(混合策略)Nash均衡.

模型检验 按照 Nash 均衡的要求，罚球队员应该以 38.3%的概率向左侧踢球，61.7%的概率向右侧踢球，而守门员以 41.7%的概率向左侧扑球，58.3%的概率向右侧扑球.这个结果是否与实战中罚球队员和守门员的行为基本一致？有人收集了400多次实际罚球的数据，统计分析发现罚球队员大约以40%的概率向左侧踢球，60%的概率向右侧踢球，而守门员以42%的概率向左侧扑球，58%的概率向右侧扑球.可见，博弈论模型的结果是实际情况的一个较好的近似.

在上面的零和博弈中，双方的支付矩阵是 $\boldsymbol{M}$ 和 $-\boldsymbol{M}$.更一般地，也可以将 $-\boldsymbol{M}$ 中的每个元素加上一个常数后作为守门员的支付矩阵，如加上常数1时，效用的含义是不进球的概率，可视为守门员的期望得分，这时博弈一般称为常数和博弈.可以想到，由于二者效用之和为常数，双方仍然是严格竞争的，这时上面的求解方法依然有效.

评注 表1是对许多场比赛和很多球员得到的平均进球概率，而每个特定罚球队员和守门员的身体素质和运动水平会有所差异，进球概率自然会不同，Nash均衡也会不同.对于一场特定的点球大战，如果有具体出场的罚球队员和守门员以前对阵的进球概率的统计数据，应该可以预测其应该采用的最佳混合策略.

复习题

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 Nash 均衡？如果不存在，求出该博弈模型的混合战略 Nash 均衡.

2. 一场棒球比赛即将开赛，击球手已经出场，等待投球手上场.准备上场的投球手可能投出快球、弧线球、变速球、叉指快速球，而该击球手可以猜测投球手会投出哪种球而做好应对准备.过去的经验数据显示，不同投球方式和应对方式下该击球手的平均得分不同，并且得分还依赖于投球手是右手投球队员还是左手投球队员(分别参见下表).如果你是教练员，你应该派右手投球手还是左手投球手出场？

击球手	右手/左手投球手			
	快球	弧线球	变速球	叉指快速球
快球	0.337/0.353	0.246/0.185	0.220/0.220	0.200/0.244
叉指快速球	0.283/0.143	0.571/0.333	0.339/0.333	0.303/0.253
弧线球	0.188/0.071	0.347/0.333	0.714/0.353	0.227/0.247
变速球	0.200/0.300	0.227/0.240	0.154/0.254	0.500/0.450

10.2 拥堵的早高峰

交通拥堵是现代城市生活中司空见惯的现象，它可以理解为出行者之间相互博弈的结果.本节通过一个简单的博弈模型对拥堵的早高峰现象进行分析.

问题背景 小王最近高兴地搬进了某居民区的新居，但上班的公司位于新技术开发区，居民区与公司之间没有公共交通，每天只能开私家车上班.居民区与开发区之间有高速公路连接，如果没有拥堵，20 min 就可以到达公司.但高速公路的出口在上班早高峰时段非常拥堵，虽然公司要求早上 8:30 上班，小王刚开始的几天每天 8:00 从居民区出发，8:30 以后才到公司；后来他改为 7:30 出发，结果8:00 以前就到了公司.在路上拥堵的时间耗费更多汽油，而迟到后公司要扣工资，早到了又浪费时间(早到的时间不算在工作时间之内).小王想，自己应该几点出发去上班呢？

问题分析 在不考虑交通事故等突发因素的情况下，道路拥堵本质上是由于一段时间内车辆出行的实际需求超过了道路允许的通行能力(供给)而必然出现的现象.对于纵横交错的复杂道路网络，每条道路的通行能力与车流密度有关，而且由于车流不断分流和汇合，每条道路上的需求也不容易界定，理论上进行分析比较困难.我们这里只考虑一种非常简单的情形，假设从居民区到公司的高速道路没有分叉、没有其他道路与其交汇.这样就可以合理地假设每天早高峰的出行需求是一个常数(不妨设每天通过这条道路去上班的人每人驾驶一辆车)，并且出口是这条道路唯一可能的拥堵点(称为瓶颈)，瓶颈处单位时间最多只能通过一定数量的车辆，超过这个上限的车辆就出现排队现象.即使对于这样简单的情形，小王一个人的决策(何时

出发去上班)也不能简单地决定他何时到达公司.显然,道路是否拥堵、拥堵程度如何还取决于其他人的行为(其他人何时出发去上班).同一时刻出发的人越多,在出口处的排队时间就越长.因此出行者之间的决策是相互影响的,这类问题也可以用**完全信息的静态博弈**进行建模和分析.

按照上述分析,不妨认为处于开发区的公司位于高速公路出口处,忽略从出口到公司所需的时间,而出口是公路上唯一的拥堵点(瓶颈).对小王与住在同一居民区开车到公司上班的所有出行者来说,出发到达公路出口的时间相同,都是 20 min(忽略瓶颈处的车辆排队对此路程和所需时间的微小影响).这个常数对整个问题的分析、求解没有本质的影响,可以不做考虑.这样,每个出行者上班所需的时间就等于他在公路出口处排队等待的时间.

每个出行者需要做的决策是从居民区出发的时刻,使得自己每天的效用函数最大,这等价于每天的总出行成本最小.成本怎样衡量呢? 排队等待时间过长不好,按照小王的经验,早到、晚到公司也不好,所以可认为每人的总出行成本包括等待成本、早到成本和迟到成本.尽管每个人出发和在公路出口排队等待的时刻不同,从而有人会早到公司、有人会正点到达,也有的会迟到,但是在经过一段时间的尝试、磨合之后,所有出行者的决策将处于 Nash 均衡状态,即每个人的总出行成本相同,都达到最小.

在 10.1 节纯策略 Nash 均衡的讨论中,每个参与人的策略只取有限个离散值,而对于拥堵的早高峰问题,每个参与人的决策是出门去上班的时间这样一个连续变量.根据 Nash 均衡(策略组合)的本质是任一参与人单方面地偏离该策略均不能使自己的效用得到提升,可以将 Nash 均衡的概念推广到纯策略空间是无限集合的情形.

本节的问题就是要在 Nash 均衡状态下建立出行者出发时刻的分布规律.[2]

模型假设 在 Nash 均衡状态下作以下具体假设:

1) 公司对所有出行者要求的上班时刻为 t^*,每天早高峰有 n 辆完全相同的车通过高速公路去公司,公路出口的最大通行能力为 s(单位时间最多通过的车辆数).

2) 早高峰时段(从在出口排队的第一辆车开始到最后一辆车为止)出口处一直处于拥堵状态,记从居民区第一辆车的出发时刻为 t_1,最后一辆车的出发时刻为 t_2,正点时刻 t^* 到公司的车的出发时刻为 t_0,$t_1 \leqslant t_0 \leqslant t_2$.

3) 出行者出发时刻的分布规律用时刻 t 累计出发的车辆数描述,记为 $F(t)$ $(t_1 \leqslant t \leqslant t_2)$;时刻 t 出口处的车辆数(排队长度)记为 $Q(t)$,且 $Q(t_1)=Q(t_2)=0$;时刻 t 累计通过出口(即到达公司)的车辆数为 $G(t)$.由于 n 较大,把 $F(t)$,$Q(t)$,$G(t)$ 当成连续量(可取任意非负实数值)处理.

4) 单位时间的等待成本(耗油等)记为 α,早到成本(浪费时间等)记为 β,迟到成本(罚金等)记为 γ,且 $\gamma>\alpha>\beta>0$;总出行成本是这些成本之和,并且每个出行者的总出行成本相同.

模型建立 记时刻 t 出发的车辆在公路出口处排队等待的时间为 $T(t)$,由假设 1)~3)可知

$$T(t)=Q(t)/s \tag{1}$$

$T(t)$也就是时刻t出发的车辆到达公司的时间.

对于正点时刻t^*到达公司的车辆,其出发时刻为

$$t_0=t^*-T(t_0) \tag{2}$$

如果$t<t_0$,时刻t出发的车辆早到的时间记为$E(t)$,则

$$E(t)=t^*-T(t)-t \tag{3}$$

如果$t>t_0$,时刻t出发的车辆晚到的时间记为$L(t)$,则

$$L(t)=t+T(t)-t^* \tag{4}$$

按照假设4),时刻t出发的每个出行者花费的总成本为

$$C(t)=\alpha T(t)+\beta E(t)+\gamma L(t) \tag{5}$$

当$t<t_0$,将(1)和(3)代入(5)得到

$$C(t)=\beta(t^*-t)+\frac{\alpha-\beta}{s}Q(t) \tag{6}$$

根据假设4)所有早到的出行者的成本相同,即$C(t)$与t无关,所以$\mathrm{d}C/\mathrm{d}t=0$.对(6)求导并得到$Q(t)$后,再利用$Q(t_1)=0$可得

$$Q(t)=\frac{\beta s}{\alpha-\beta}(t-t_1) \tag{7}$$

当$t>t_0$,将(1)和(4)代入(5)得到

$$C(t)=\gamma(t-t^*)+\frac{\alpha+\gamma}{s}Q(t) \tag{8}$$

根据假设4)所有晚到的出行者成本相同,类似于(7)式的推导,并利用$Q(t_2)=0$得到

$$Q(t)=\frac{\gamma s}{\alpha+\gamma}(t_2-t) \tag{9}$$

由(7)(9)可知,排队长度$Q(t)$是分段线性函数,并且在$t=t_0$点$Q(t)$是连续的.

由假设3)时刻t累计出发的车辆数$F(t)$等于通过出口车辆数$G(t)$与排队长度$Q(t)$之和,即

$$F(t)=G(t)+Q(t) \tag{10}$$

因为出口处从t_1到t_2一直处于拥堵状态(否则显然不是Nash均衡),时刻t累计通过出口的车辆数$G(t)=s(t-t_1)$,将$G(t)$及(7)(9)式的$Q(t)$一起代入(10)可得

$$F(t)=\begin{cases}s(t-t_1)+\dfrac{\beta s}{\alpha-\beta}(t-t_1)=\dfrac{\alpha s}{\alpha-\beta}(t-t_1), & t_1\leqslant t\leqslant t_0\\ s(t-t_1)+\dfrac{\gamma s}{\alpha+\gamma}(t_2-t)=\dfrac{\alpha s}{\alpha+\gamma}t+\left(\dfrac{\gamma s}{\alpha+\gamma}t_2-st_1\right), & t_0<t\leqslant t_2\end{cases} \tag{11}$$

这就是出行者出发时刻的分布规律.可以看出$F(t)(t_1\leqslant t\leqslant t_2)$也是分段线性函数,且在区间端点有$F(t_1)=0,F(t_2)=s(t_2-t_1)$.在已知成本参数$\alpha,\beta,\gamma$及$t^*,n,s$的条件下,需要确定的只是区间端点$t_1,t_2$以及$t_0$的值.

模型求解 为了求出t_1,t_2,t_0的值,需要找到它们应满足的3个方程.

由假设1),2)容易得到区间$[t_1,t_2]$的长度为

$$t_2-t_1=\frac{n}{s} \tag{12}$$

因为 $Q(t)$ 在时刻 t_0 是连续的，在(7)(9)中以 $t=t_0$ 代入可得

$$\frac{t_2-t_0}{t_0-t_1}=\frac{\beta(\alpha+\gamma)}{\gamma(\alpha-\beta)} \tag{13}$$

这给出了 t_0 对区间$[t_1,t_2]$划分的比值.

根据(1)(2)式并将 $t=t_0$ 代入(7)得(若代入(9)结果是等价的)

$$t_0=\frac{\alpha-\beta}{\alpha}t^*+\frac{\beta}{\alpha}t_1 \tag{14}$$

求解方程组(12)~(14)可以得到

$$t_1=t^*-\frac{\gamma}{\beta+\gamma}\frac{n}{s} \tag{15}$$

$$t_2=t^*+\frac{\beta}{\beta+\gamma}\frac{n}{s} \tag{16}$$

$$t_0=t^*-\frac{\beta\gamma}{\alpha(\beta+\gamma)}\frac{n}{s} \tag{17}$$

(15)~(17)给出了 t_1,t_2,t_0 与参数 $\alpha,\beta,\gamma,t^*,n,s$ 的关系.因为 $\gamma>\beta>0$，所以 $t^*-t_1>t_2-t^*$.

下面计算时刻 t 出发的每个出行者花费的总成本.对于 $t\leqslant t_0$，由(6)(7)(15)可得

$$C(t)=\frac{\beta\gamma}{\beta+\gamma}\frac{n}{s} \tag{18}$$

对于 $t\geqslant t_0$，由(8)(9)(16)得到的结果相同.$C(t)$ 与 t 无关，即任意时刻出发的出行者花费的总成本相同，正是均衡状态的结果.于是 n 个人出行的总成本为

$$\mathrm{TC}=\frac{\beta\gamma}{\beta+\gamma}\frac{n^2}{s} \tag{19}$$

表明总成本与单位时间的等待成本 α 和 t^* 无关(请思考一下这是为什么).

结果分析 在图1中按照(11)式画出时刻 t 累计出发的车辆数 $F(t)$(折线 OBD)，及时刻 t 累计通过出口(即到达公司)的车辆数 $G(t)$(直线 OD)，每一线段的斜率已在图中标出，显然 OB 段的斜率大于 BD 段的斜率.

对于任意时刻 t 由(10)式可知，折线 OBD 与直线 OD 之间的垂直距离表示出口处排队的车辆数 $Q(t)$，而这两条线之间的水平距离表示在出口处的排队等待时间 $T(t)$.时刻 t_0 对应 B 点，$t=t_0$ 出发的人等待时间(BC 段的长度)最长，但是能够正点 t^* 到达公司.$t<t_0$(OB 段)出发的人都会早到，$t>t_0$(BD 段)出发的人都会迟到.

由图所表示的等待时间 $T(t)$ 可以清楚地看出，三角形 OBD 的面积 $S_{\triangle OBD}$ 正好对应于所有人的等待时间之和.根据简单的几何关系：$S_{\triangle OBD}=S_{\triangle OBC}+S_{\triangle BCD}=BC\cdot OA/2+BC\cdot DE/2=BC\cdot DG/2$，又 $BC=t^*-t_0$，$DG=n$，利用(17)式可得所有人的总等待成本是

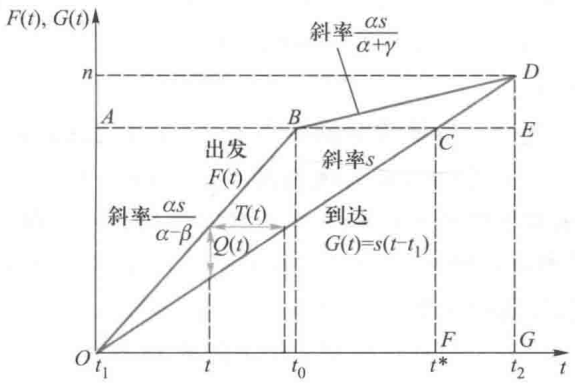


图 1 累计的出发车辆数 $F(t)$ 与到达车辆数 $G(t)$

$$\mathrm{TTC}=\alpha S_{\triangle OBD}=\frac{\beta\gamma}{\beta+\gamma}\frac{n^2}{2s} \tag{20}$$

与(19)式比较可知 TTC = TC/2,即总等待成本正好是总成本的一半,那么另一半应该就等于总的早到和迟到成本之和.在图 1 中,三角形 OCF 的面积 $S_{\triangle OCF}$ 对应于总的早到时间之和,而三角形 CDE 的面积 $S_{\triangle CDE}$ 对应于总的迟到时间之和.读者容易对此进行计算和验证.

根据以上分析,无论小王什么时候从居民区出发,他的出行成本都是一样的.如果他早些出发或者晚些出发,虽然可以减少在出口的排队等待时间,但要么会早到公司浪费时间,要么会迟到被罚款.如果他希望正点到公司,就需要在路上忍受更长时间的车辆排队.

模型应用 为了减少大家在上班路上的拥堵时间,有人建议管理部门对早高峰时段通过高速公路出口的车辆按照拥堵程度收取“拥堵费”.这一建议是否真的有效?如果有效,拥堵费又应该按照什么标准收取为好?

首先我们分析一下,如果有一个权威的计划人员为所有出行车辆进行统一规划(即所谓的“集中决策”),希望使所有人出行的总成本最小(称为“系统最优”),那么车辆应该如何出发?由于排队等待时间完全是浪费,因此应尽量避免;此外,应该使瓶颈资源得到充分利用,即车流应保持连续且瓶颈全程处于满负荷状态;第一辆车早到的成本应等于最后一辆车晚到的成本,否则可以将部分车流从一端移动到另一端而降低总成本.这表明,系统最优决策其实很简单,就是从 t_1 到 t_2 的任意时刻($t_1\leqslant t\leqslant t_2$),单位时间内出发的车辆数正好等于瓶颈的通行能力 s.从图 1 中看,就是使累计的出发车辆数 OBD 与 OCD 线重合.由于完全消除了等待成本,而早到成本和迟到成本不变,因此系统最优值将等于(19)式 TC 的一半,即可以节省一半的成本.

但是在没有权威的计划人员的情况下,是否存在一种在瓶颈处的收费方案,使得出行者的行为与“系统最优”一致呢?如果能够实现这一目标,我们通常说整个系统达到了协调.对社会整体来说,这应该是比较理想的状态.

可以想象,如果管理部门规定对任何排队的车辆收取一个足够高的固定拥堵费,以至于任何人偏离系统最优的出行方案都将增加其成本(因此也是达到了 Nash 均衡状

态),那么将不会有车辆排队(因此管理部门实际上也收不到拥堵费),从而实现系统最优.但这种方案将导致不同时刻出发的车辆的出行成本不同,似乎不太公平,是一种变相的、强制性的计划行为.

考虑到 t 时刻($t_1 \leqslant t \leqslant t_2$)出发的车辆将于 $t+T(t)$ 时刻到达公司,导致的早到成本是 $\beta E(t)=\beta[t^*-t-T(t)]$(当 $t<t_0$)或者晚到成本是 $\gamma L(t)=\gamma[t+T(t)-t^*]$(当 $t>t_0$),为了消除排队(此时要求对于任意 $t(t_1 \leqslant t \leqslant t_2)$,对应的 $T(t)=0$,即出发时间就是到达时间),收费方案可以考虑对每辆车"虚拟地"收取一个固定费用 a,然后对早到成本和晚到成本分别进行补偿,让每辆车的出行成本都相同(等于 a).具体来说,一种精细的收费方案是:管理部门制定一个与到达时刻 t 相关的收费方案,即对时刻 t 到达瓶颈(出口)的车辆所收取的实际费用为

$$p(t)=\begin{cases}0, & \text{当 } t<t_1\\ a-\beta(t^*-t), & \text{当 } t_1 \leqslant t<t^*\\ a-\gamma(t-t^*), & \text{当 } t^* \leqslant t \leqslant t_2\\ 0, & \text{当 } t>t_2\end{cases} \tag{21}$$

其中 a 是一个常数,表示每辆车在收费后的最终出行成本,此时瓶颈处不会出现排队(请读者对此进行证明).例如,如果取常数 $a=C(t)=(n/s)\beta\gamma/(\beta+\gamma)$,则每个出行者的成本与收费前相同,而管理部门却可以收取到拥堵费(读者容易验证此时 $p(t) \geqslant 0$),可以说是一种双赢的方案(车辆不再出现排队等待现象,管理部门收到拥堵费).

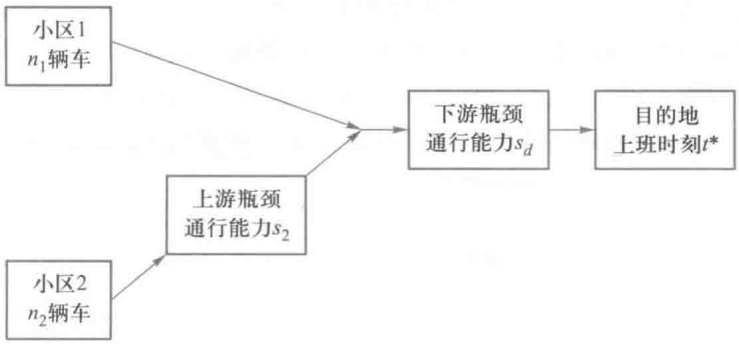
从实际实施的角度出发,(21)的收费方案也存在一定缺陷:因为收费额随时刻 t 实时变化,对收费系统有较高的要求.随着智能交通设施不断完善,这一困难可能会被逐步克服.也可以考虑分为若干离散时段分别收费(同一时段收费相同)来近似实现这一精细的收费方案(如最简单地分为两段,分别对早到和迟到车辆收取不同费用),这在实际中比较容易实施(可以减缓但不能完全消除排队).

评注 本模型作了很多简化,有很多学者对其做了进一步推广,如考虑多个瓶颈、更为复杂的道路网络(如文献[3]);出行者并不是完全相同的,每个人的参数 α,β,γ 可能不相同,上班时间也不一定相同(可能会有错峰上班、弹性上班时间等);瓶颈的通行能力可能也会受到气候等因素的影响而有一定随机性等.

复习题

1. 在收费方案(21)中,如果选择 $a<(n/s)\beta\gamma/(\beta+\gamma)$,每个出行者的实际成本与不收费时相比还有所降低,但此时不能保证 $p(t) \geqslant 0$,即对于靠近 t_1 和 t_2 的某些 t 有 $p(t)<0$,这意味着管理部门对最早出发和最晚出发的一部分车辆应补贴 $|p(t)|$;但只要 a 不是太小,扣除补贴后,管理部门最终仍可能收取到一定的拥堵费.请你计算对于任意的 a,管理部门实际收到的总的拥堵费(扣除补贴后)为多少.为了保证管理部门"不亏本",相应的 a 最小为多少?

2. 考虑如下 Y 字形的道路网,两个小区分别有 n_1,n_2 辆车从各自小区到目的地上班,小区 1 的车辆可以直接进入下游瓶颈,而小区 2 的车辆要先通过上游瓶颈,上下游瓶颈的通行能力分别为每单位时间 s_2,s_d.假设所有车辆都希望不迟到(迟到成本非常高),建立相应的数学模型,求出早高峰时的均衡结果(包括每个小区的车辆出发率,每个瓶颈前的车辆到达率和每个瓶颈后的车辆通过率,每个小区每辆车的出行成本、总出行成本等).根据上述计算结果,说明如果只增加上游瓶颈的通行能力 s_2,什么情况下将会导致总出行成本增加而不是减少.[3]



10.3 “一口价”的战略

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？


案例精讲 10-1
“一口价”的战略

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价值为 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价值为 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 $p_b(v_b) \geqslant p_s(v_s)$ 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

$$\max_{p_s}\left\{\frac{p_s + E[p_b(v_b) \mid p_b(v_b) \geqslant p_s]}{2} - v_s\right\} P\{p_b(v_b) \geqslant p_s\} \tag{1}$$

这里 $E[\,]$ 表示的是条件 $p_b(v_b) \geqslant p_s$ 下 $p_b(v_b)$ 的条件期望，$P\{\,\}$ 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

$$\max_{p_b}\left\{v_b - \frac{p_b + E[p_s(v_s) \mid p_b \geqslant p_s(v_s)]}{2}\right\} P\{p_b \geqslant p_s(v_s)\} \tag{2}$$

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价值$v_s \leqslant x$,则报价x,否则报价为1;顾客如果认为商品的价值$v_b \geqslant x$,则报价x,否则报价为0.这种价格战略可表示为

$$p_s(v_s)=\begin{cases}x, & v_s \leqslant x\\ 1, & v_s > x\end{cases} \tag{3}$$

$$p_b(v_b)=\begin{cases}x, & v_b \geqslant x\\ 0, & v_b < x\end{cases} \tag{4}$$

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$v_s \leqslant x$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$v_s \leqslant x \leqslant v_b$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

$$\eta=\frac{\int_x^1\int_0^x (v_b-v_s)\,\mathrm{d}v_s\mathrm{d}v_b}{\int_0^1\int_0^{v_b} (v_b-v_s)\,\mathrm{d}v_s\mathrm{d}v_b}$$

$$=3x(1-x) \leqslant \frac{3}{4} \tag{5}$$

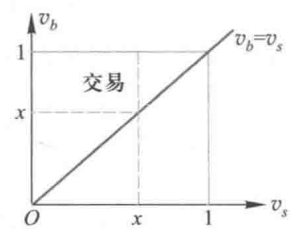


图1 单一价格战略的交易效率

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

$$p_s(v_s)=a_s+c_sv_s \tag{6}$$

$$p_b(v_b)=a_b+c_bv_b \tag{7}$$

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,$P\{p_b \geqslant p_s(v_s)\} = (p_b-a_s)/c_s$,$E[p_s | p_b \geqslant p_s] = (a_s+p_b)/2$,于是(2)式为

$$\max_{p_b}\left\{v_b - \frac{p_b + (a_s + p_b)/2}{2}\right\} \cdot \frac{p_b - a_s}{c_s} \tag{8}$$

这是一个二次函数的优化,其最优解为

$$p_b = \frac{2}{3}v_b + \frac{1}{3}a_s \tag{9}$$

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

$$p_s = \frac{2}{3}v_s + \frac{1}{3}(a_b + c_b) \tag{10}$$

比较(6),(7),(9),(10),可以解出

$$a_b = \frac{1}{12}, a_s = \frac{1}{4}, \quad c_b = c_s = \frac{2}{3} \tag{11}$$

即线性价格战略(6),(7)为

$$p_s(v_s) = \frac{2}{3}v_s + \frac{1}{4} \tag{12}$$

$$p_b(v_b) = \frac{2}{3}v_b + \frac{1}{12} \tag{13}$$

理论上来说,在考虑端点条件时,上述(12)式应该只对$p_s(v_s) \in [a_b, a_b+c_b] = \left[\frac{1}{12}, \frac{3}{4}\right]$,即$v_s \leqslant \frac{3}{4}$有效,(13)式应该只对$p_b(v_b) \in [a_s, a_s+c_s] = \left[\frac{1}{4}, \frac{11}{12}\right]$,即$v_b \geqslant \frac{1}{4}$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$p_s(v_s) \leqslant p_b(v_b)$时交易实际上才能发生,将(12),(13)代入得到交易条件为$v_b \geqslant v_s + \frac{1}{4}$,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

$$\eta = \frac{\int_{\frac{1}{4}}^{1}\int_{0}^{v_b-\frac{1}{4}}(v_b - v_s)\,\mathrm{d}v_s\,\mathrm{d}v_b}{\int_{0}^{1}\int_{0}^{v_b}(v_b - v_s)\,\mathrm{d}v_s\,\mathrm{d}v_b} = \frac{27}{32} > \frac{3}{4} \tag{14}$$

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).

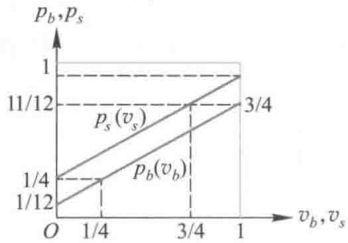


图2 均衡的线性价格战略

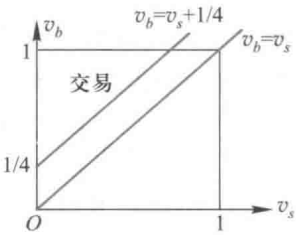


图3 线性价格战略的交易效率

评注 是否存在比线性战略均衡的交易效率更高的Bayes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.

这里讨论的模型一般称为双向拍卖[22],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Bayes均衡或Bayes-Nash均衡.

复习题

1. 两个投标人为获得某件物品竞标,每个投标人知道自己对该物品的估值,但不知道对手的估值,只知道对手的估值服从[0,1]区间上的均匀分布.投标采用最高价格密封拍卖,即两个投标者分别同时向物品的拥有者秘密地、一次性地给出自己的报价,然后物品的拥有者(拍卖者)从中选择报价最高者成交(如果两个投标人给出的报价相同,则等概率地随机选择一个成交),成交的竞标者向拍卖者支付的费用等于自己的报价.试建模分析两个投标人的均衡报价策略,并将结果推广到多于两个投标人的情形.

2. 继续考虑第1题中的问题,但假设投标采用次高价格密封拍卖,即成交的竞标者向拍卖者支付的费用等于除自己的报价外其他人的最高报价.试建模分析投标人的均衡报价策略,并比较两种拍卖形式下拍卖者的期望收益.

10.4 不患寡而患不均

案例精讲10-2 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，特别是还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[19].

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$s(0\leqslant s\leqslant 1)$，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$\alpha_1,\alpha_2(\geqslant 0)$分别表示两人的“愤怒”系数，$\beta_1,\beta_2(\geqslant 0)$分别表示两人的“愧疚”系数，并且不妨假设$\alpha_i\geqslant\beta_i,i=1,2$.建立两人的效用函数为

$$U_i(x_1,x_2)=x_i-\alpha_i\max\{x_j-x_i,0\}-\beta_i\max\{x_i-x_j,0\},i=1,2,j=3-i \tag{1}$$

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$s\geqslant 1/2$，则$x_2\geqslant x_1$，于是

$$U_2(s)=s-\beta_2(2s-1) \tag{2}$$

由$\beta_2<1/2$可知(2)式中的$U_2(s)\geqslant 1/2>0$，所以乙的最优反应是接受.

若$s\leqslant 1/2$，则$x_2\leqslant x_1$，于是

$$U_2(s)=s-\alpha_2(1-2s)=(1+2\alpha_2)s-\alpha_2 \tag{3}$$

仅当(3)式中的$U_2(s)\geqslant 0$，即$s\geqslant\alpha_2/(1+2\alpha_2)$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$0\leqslant\bar{s}(\alpha_2)<1/2$.

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 $s\geqslant 1/2$,则 $x_2\geqslant x_1$,于是

$$U_1(s)=1-s-\alpha_1(2s-1) \tag{4}$$

$U_1(s)$ 在 $s^*=1/2$ 时达到最大值 1/2,所以只需要讨论 $s\leqslant 1/2$ 且 $s\geqslant \underline{s}(\alpha_2)$ 的情况.此时 $x_2\leqslant x_1$,于是

$$U_1(s)=1-s-\beta_1(1-2s)=1-\beta_1+(2\beta_1-1)s,\qquad \underline{s}(\alpha_2)\leqslant s\leqslant 1/2 \tag{5}$$

$U_1(s)$ 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

$$s^*=\underline{s}(\alpha_2)=\alpha_2/(1+2\alpha_2) \tag{6}$$

可见,甲提议给乙的比例为 $\underline{s}(\alpha_2)$,严格小于 50%.此外,$\underline{s}(\alpha_2)$ 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 $s\geqslant 1/2$,则 $x_2\geqslant x_1$,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 $s\leqslant 1/2$ 的情况.

甲知道 $s\geqslant \underline{s}(\alpha_2)$ 时乙才会接受甲的提议,且 $s\geqslant \underline{s}(\alpha_2)=\alpha_2/(1+2\alpha_2)$ 等价于 $\alpha_2\leqslant s/(1-2s)$.若设 α_2 的概率分布函数为 $F(\alpha_2)$,且 $F(\underline{\alpha})=0$,$F(\overline{\alpha})=1$①,则甲可以推测乙接受甲的提议 s 的概率 p 为

$$p=\begin{cases}0, & s\leqslant \underline{s}(\underline{\alpha})\\ F(s/(1-2s)), & \underline{s}(\underline{\alpha})<s<\underline{s}(\overline{\alpha})\\ 1, & s\geqslant \underline{s}(\overline{\alpha})\end{cases} \tag{7}$$

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

$$EU_1(s)=\begin{cases}0, & s\leqslant \underline{s}(\underline{\alpha})\\ [1-\beta_1+(2\beta_1-1)s]F(s/(1-2s)), & \underline{s}(\underline{\alpha})<s<\underline{s}(\overline{\alpha})\\ 1-\beta_1+(2\beta_1-1)s, & s\geqslant \underline{s}(\overline{\alpha})\end{cases} \tag{8}$$

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 $\underline{s}(\underline{\alpha})<s\leqslant \underline{s}(\overline{\alpha})$ 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 $\underline{s}(\overline{\alpha})$,严格小于 50%.

---

① 严格地说,$\underline{\alpha}=\max\{\alpha\mid F(\alpha)=0\}$,$\overline{\alpha}=\min\{\alpha\mid F(\alpha)=1\}$.

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，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$，定义 i 的效用函数为

$$U_i(\boldsymbol{x})=x_i-\alpha_i\frac{1}{n-1}\sum_{j\neq i}\max\{x_j-x_i,0\}-\beta_i\frac{1}{n-1}\sum_{j\neq i}\max\{x_i-x_j,0\},\ i=1,2,\cdots,n \tag{9}$$

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也具有很强的预测能力.

其中 α_i,β_i 的含义与前面相同.与二人情形假设 $\beta_i<1/2$ 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

$$\begin{aligned}U_i(\boldsymbol{x})&=x_i-\beta_i\frac{1}{n-1}\sum_{j\neq i}(x_i-x_j)=x_i-\beta_i\left(x_i-\frac{1-x_i}{n-1}\right)\\&=\frac{\beta_i}{n-1}+\left(1-\frac{n\beta_i}{n-1}\right)x_i\end{aligned} \tag{10}$$

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 $\beta_i<\frac{n-1}{n}<1$.

复习题

按照本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：首先由一个唯一指定的“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（9）式定义的 $\beta_i<(n-1)/n$，给出这个博弈的均衡[19].

10.5 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

$$x_1+x_2+x_3=11 \tag{1}$$

$$x_1,x_2,x_3\geqslant 1,\quad x_1+x_2\geqslant 7,\quad x_1+x_3\geqslant 5,\quad x_2+x_3\geqslant 4 \tag{2}$$

（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.Shapley① 1953 年给出了解决该问题的一种方法,称 **Shapley 值**[57].

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(为统一起见,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 $I=\{1,2,\cdots,n\}$,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

$$v(\varnothing)=0 \tag{3}$$

$$v(s_1\cup s_2)\geqslant v(s_1)+v(s_2),\quad s_1\cap s_2=\varnothing \tag{4}$$

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.$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 叫作合作对策的分配(imputation),满足

$$\sum_{i=1}^{n}x_i=v(I) \tag{5}$$

$$x_i\geqslant v(i),\quad i=1,2,\cdots,n \tag{6}$$

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 $\boldsymbol{\Phi}(v)=(\varphi_1(v),\varphi_2(v),\cdots,\varphi_n(v))$.对于任意的子集 s,记 $x(s)=\sum\limits_{i\in s}x_i$,即 s 中各成员的分配.对一切 $s\subset I$,满足 $x(s)\geqslant v(s)$ 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 $\boldsymbol{\Phi}(v)$,并把它构造出来.这里只给出 $\boldsymbol{\Phi}(v)$ 的结果,Shapley 公理可参看[8].

Shapley 值 $\boldsymbol{\Phi}(v)=(\varphi_1(v),\varphi_2(v),\cdots,\varphi_n(v))$ 为

$$\varphi_i(v)=\sum_{s\in S_i}w(|s|)\,[v(s)-v(s\backslash i)],\quad i=1,2,\cdots,n \tag{7}$$

$$w(|s|)=\frac{(n-|s|)!\,(|s|-1)!}{n!} \tag{8}$$

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

---

① L.S.Shapley(1923—2016) 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,2012 年获诺贝尔经济学奖.第二次世界大战期间,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—2016 年为加州大学洛杉矶分校教授.

子，$s\backslash i$ 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 $v(\varnothing)=0, v(1)=v(2)=v(3)=1, v(1,2)=7, v(1,3)=5, v(2,3)=4, v(I)=11$. 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，$v(s\backslash 1)$ 是无甲参加时合作 s(只剩下乙)的获利，所以 $v(s)-v(s\backslash 1)$ 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
$v(s\backslash 1)$	0	1	1	4
$v(s)-v(s\backslash 1)$	1	6	4	7
$\|s\|$	1	2	2	3
$w(\|s\|)$	1/3	1/6	1/6	1/3
$w(\|s\|)[v(s)-v(s\backslash 1)]$	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[41]？

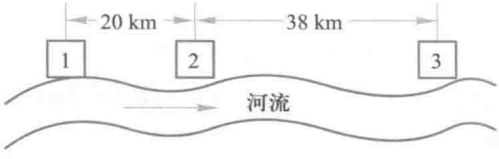


图 1 三城镇地理位置示意图

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

$C(1)=73\times5^{0.712}=230$, $C(2)=160$, $C(3)=230$,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

$C(1,2)=73\times(5+3)^{0.712}+0.66\times5^{0.51}\times20=350$,总投资 $D_2=C(1,2)+C(3)=580$.

3) 2,3 合作,在城 3 建厂.投资为

$C(2,3)=73\times(3+5)^{0.712}+0.66\times3^{0.51}\times38=365$,总投资 $D_3=C(1)+C(2,3)=595$.

4) 1,3 合作,在城 3 建厂.投资为

$C(1,3)=73\times(5+5)^{0.712}+0.66\times5^{0.51}\times58=463$,这个费用超过了 1,3 分别建厂的费用 $C(1)+C(3)=460$.合作没有效益,不可能实现.

5) 三城合作,在城 3 建厂.总投资为 $D_5=C(1,2,3)=73\times(5+3+5)^{0.712}+0.66\times5^{0.51}\times20+0.66\times(5+3)^{0.51}\times38=556$.

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 $d_2=0.66\times5^{0.51}\times20=30$;城 2 至 3 的管道费 $d_3=0.66\times(5+3)^{0.51}\times38=73$.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:

城 3 分担费用为 $d_1\times\frac{5}{13}=174$;

城 2 分担费用为 $d_1\times\frac{3}{13}+d_3\times\frac{3}{8}=132$;

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

$$
\begin{aligned}
&v(\varnothing)=0, v(1)=v(2)=v(3)=0\\
&v(1,2)=C(1)+C(2)-C(1,2)=230+160-350=40\\
&v(2,3)=C(2)+C(3)-C(2,3)=160+230-365=25\\
&v(1,3)=0\\
&v(I)=C(1)+C(2)+C(3)-C(1,2,3)\\
&\quad=230+160+230-556=64
\end{aligned}
$$

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
$v(s\backslash 1)$	0	0	0	25
$v(s)-v(s\backslash 1)$	0	40	0	39
$\lvert s\rvert$	1	2	2	3
$w(\lvert s\rvert)$	1/3	1/6	1/6	1/3
$w(\lvert s\rvert)[v(s)-v(s\backslash 1)]$	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 $C(3)-\varphi_3(v)=217.8$.

评注 Shapley 值最初是从解决合作对策的分配问题提出来的，后来给以公理化的证明，成为一种公正、合理的解决方法，但是定义特征函数时需要知道集合 $N=\{1,2,\cdots,n\}$ 的每一个子集 S（共有 2^n 个）获得的效益，在实际上常常做不到，这就限制了 Shapley 值方法的应用范围. 当掌握的信息较少时处理这类分配问题有另一类方法，参见更多案例 10-3.

复习题

1. 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.

2. 证明由(7)，(8)给出的 Shapley 值 $\boldsymbol{\Phi}(v)$ 满足 $\varphi_i(v)\geqslant v(i)$，$i=1,2,\cdots,n$.

10.6 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[41].

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$N=\{A,B,C,\cdots\}$及他们的投票权重（均为正数）$w_1,w_2,\cdots,w_n$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$S=[q;w_1,w_2,\cdots,w_n]$表示，其中q称为**定额**（quota），若记$w=w_1+w_2+\cdots+w_n$，一般假定$w/2<q\leqslant w$.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$\boldsymbol{p}=(p_1,p_2,\cdots,p_n)$，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.

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.

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.

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,却对结果毫无影响力.

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 $\boldsymbol{k}=(k_1,k_2,k_3)$ 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 $\boldsymbol{k}^{(1)}=(1,0,0)$,$S^{(2)}$ 有 $\boldsymbol{k}^{(2)}=(1,1,1)$,$S^{(3)}$ 有 $\boldsymbol{k}^{(3)}=(1,1,0)$,而对于 $S^{(4)}$,$\boldsymbol{k}^{(4)}$ 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 $S=[q;w_1,w_2,\cdots,w_n]$ 或其对策论的表述 $G=(N,W)$,必有空集 $\varnothing\notin W$,$N\in W$,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 $k_i \geqslant k_j$；

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

$$\underline{\mathrm{A}}\underline{\mathrm{B}}\mathrm{C} \quad \underline{\mathrm{A}}\underline{\mathrm{C}}\mathrm{B} \quad \underline{\mathrm{B}}\underline{\mathrm{A}}\mathrm{C} \quad \underline{\mathrm{B}}\underline{\mathrm{C}}\mathrm{A} \quad \underline{\mathrm{C}}\underline{\mathrm{A}}\mathrm{B} \quad \underline{\mathrm{C}}\underline{\mathrm{B}}\mathrm{A}$$

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 $\boldsymbol{\xi}$ 除以排列数 $n!$，定义为他们在系统中的 **Shapley 权力指标**，记作 $\boldsymbol{\varphi}=\boldsymbol{\xi}/n!$，$\boldsymbol{\varphi}=(\varphi_1,\varphi_2,\cdots,\varphi_n)$.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：

$$\begin{array}{llllll}
\underline{\mathrm{A}}\underline{\mathrm{B}}\mathrm{CD} & \underline{\mathrm{A}}\underline{\mathrm{B}}\mathrm{DC} & \underline{\mathrm{A}}\underline{\mathrm{C}}\mathrm{BD} & \underline{\mathrm{A}}\underline{\mathrm{C}}\mathrm{DB} & \underline{\mathrm{A}}\underline{\mathrm{D}}\mathrm{BC} & \underline{\mathrm{A}}\underline{\mathrm{D}}\mathrm{CB} \\
\underline{\mathrm{B}}\underline{\mathrm{A}}\mathrm{CD} & \underline{\mathrm{B}}\underline{\mathrm{A}}\mathrm{DC} & \underline{\mathrm{B}}\underline{\mathrm{C}}\mathrm{AD} & \underline{\mathrm{B}}\underline{\mathrm{C}}\mathrm{DA} & \underline{\mathrm{B}}\underline{\mathrm{D}}\mathrm{AC} & \underline{\mathrm{B}}\underline{\mathrm{D}}\mathrm{CA} \\
\underline{\mathrm{C}}\underline{\mathrm{A}}\mathrm{BD} & \underline{\mathrm{C}}\underline{\mathrm{A}}\mathrm{DB} & \underline{\mathrm{C}}\underline{\mathrm{B}}\mathrm{AD} & \underline{\mathrm{C}}\underline{\mathrm{B}}\mathrm{DA} & \underline{\mathrm{C}}\underline{\mathrm{D}}\mathrm{AB} & \underline{\mathrm{C}}\underline{\mathrm{D}}\mathrm{BA} \\
\underline{\mathrm{D}}\underline{\mathrm{A}}\mathrm{BC} & \underline{\mathrm{D}}\underline{\mathrm{A}}\mathrm{CB} & \underline{\mathrm{D}}\underline{\mathrm{B}}\mathrm{AC} & \underline{\mathrm{D}}\underline{\mathrm{B}}\mathrm{CA} & \underline{\mathrm{D}}\underline{\mathrm{C}}\mathrm{AB} & \underline{\mathrm{D}}\underline{\mathrm{C}}\mathrm{BA}
\end{array}$$

统计出 A，B，C，D 成为决定者的次数 $\boldsymbol{\xi}=(10,6,6,2)$，Shapley 指标为 $\boldsymbol{\varphi}=(5/12,3/12,3/12,1/12)$.

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：

$$\underline{\text{ABCD}} \quad \underline{\text{ABDC}} \quad \underline{\text{ADBC}} \quad \underline{\text{BACD}} \quad \underline{\text{BADC}} \quad \underline{\text{BCAD}}$$
$$\underline{\text{BCDA}} \quad \underline{\text{BDAC}} \quad \underline{\text{BDCA}} \quad \underline{\text{DABC}} \quad \underline{\text{DBAC}} \quad \underline{\text{DBCA}}$$

统计出 A,B(和 C),D 成为决定者的次数 $\boldsymbol{\xi}=(5,6,1)$,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 $W=([\text{AB}],[\text{AC}],[\text{ABC}])$,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

$$\underline{\text{AB}} \quad \underline{\text{AC}} \quad \underline{\text{ABC}}$$

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(4,1,1)或(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 $\boldsymbol{\eta}$ 归一化,定义为他们在系统中的 **Banzhaf**① 权力指标,记作 $\boldsymbol{\beta}=(\beta_1,\beta_2,\cdots,\beta_n)$.

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 $W=([\text{AB}],[\text{AC}],[\text{ABC}],[\text{ABD}],[\text{ACD}],[\text{BCD}],[\text{ABCD}])$,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

$$\underline{\text{AB}} \quad \underline{\text{AC}} \quad \underline{\text{ABC}} \quad \underline{\text{ABD}} \quad \underline{\text{ACD}} \quad \underline{\text{BCD}} \quad \text{ABCD}$$

统计出 A,B,C,D 成为决定者的次数 $\boldsymbol{\eta}=(5,3,3,1)$,归一化后 Banzhaf 指标为 $\boldsymbol{\beta}=(5/12,3/12,3/12,1/12)$,与前面得到的 Shapley 指标 $\boldsymbol{\varphi}$ 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 $N=\{\text{A},\text{A},\text{A},\text{A},\text{A},\text{B},\text{B},\text{B},\text{B},\text{B}\}$,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

---

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.

$W_m=([4A],[3A2B],[2A4B])$.可以合理地将加权投票系统取为 $S=[q;a,a,a,a,a,1,1,1,1,1]$,其中 a,q 为大于1的整数,且 a,q 应满足 $4a\geqslant q,3a+2\geqslant q,2a+4\geqslant q,3a+1<q,\cdots$.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)$\underline{\text{B}}$和(2A3B)$\underline{\text{B}}$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 $C_5^3C_4^1 4!5!$ 和 $C_5^2C_4^3 4!5!$.于是某个B的Shapley指标为 $\frac{4!\ 5!}{10!}(C_5^3C_4^1+C_5^2C_4^3)=\frac{4}{63}=0.063\ 5$,而某个A的Shapley指标就很容易得到,为 $\frac{1}{5}\left(1-\frac{4}{63}\times 5\right)=0.136\ 5$.系统的Shapley指标为 $\boldsymbol{\varphi}=(0.136\ 5,0.136\ 5,0.136\ 5,0.136\ 5,0.136\ 5,0.063\ 5,0.063\ 5,0.063\ 5,0.063\ 5,0.063\ 5)$.

评注 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.

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	$\underline{4A}$	$\underline{4A}1B$	$\underline{\underline{3A}}2B$	$\underline{3A}3B$	$\underline{\underline{2A}}4B$	$\underline{2A}5B$
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 $\boldsymbol{\beta}=(0.135\ 5,0.135\ 5,0.135\ 5,0.135\ 5,0.135\ 5,0.064\ 5,0.064\ 5,0.064\ 5,0.064\ 5,0.064\ 5)$.

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 $S^{(1)}=[21;1,1,\cdots,1]$,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($\varphi_i=1/40,i=1,\cdots,40$),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 $S^{(2)} = [21;11,1,\cdots,1]$，除民主党 M 的权重为 11 外，其中 29 人均为 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，共和党和独立人士的权力则有所减少.

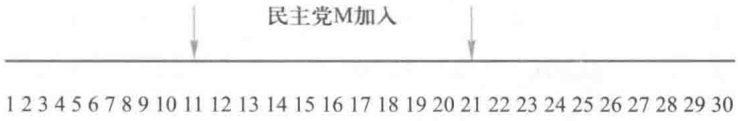


图 1 计算民主党投票人 M 的 Shapley 指标的图示

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 $S^{(3)} = [21;11,14,1,\cdots,1]$，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 17×17−17=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.

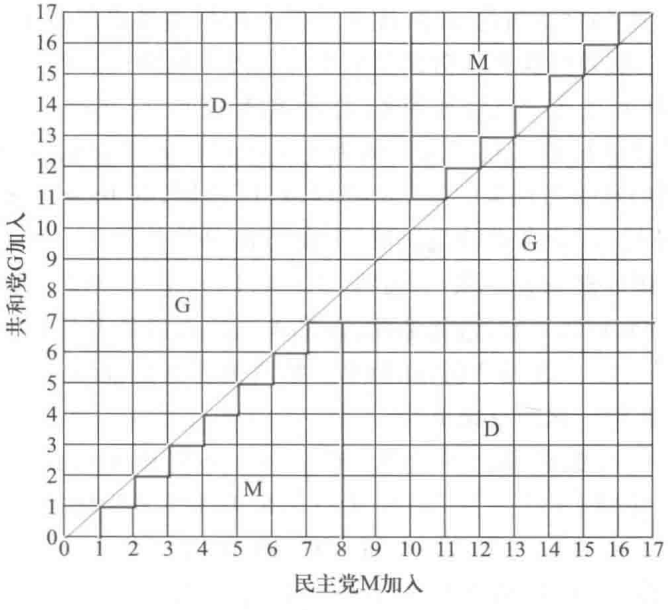


图 2 计算民主党 M 和共和党 G 的 Shapley 指标的图示

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 $j \leqslant 7$，$i \leqslant 8$ 时 M 成为决定者，$j \leqslant 7$，$i > 8$ 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.

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.

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 $\boldsymbol{p}=(3,1,1)$,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 $\boldsymbol{\beta}=(1,0,0)$ 与人口比例相差甚远(因为考虑的是比例,用 $\boldsymbol{\beta}$ 与 $\boldsymbol{\beta}'$ 是等价的).

要使 Banzhaf 指标 $\boldsymbol{\beta}$ 与 $\boldsymbol{p}$ 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 $\boldsymbol{\beta}=(3/5,1/5,1/5)$.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唯一的.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 $\boldsymbol{\beta}=(3/5,1/5,1/5)$.其实,还可以写出很多个 $\boldsymbol{\beta}$ 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 $\boldsymbol{p}=(p_1,p_2,\cdots,p_n)$,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 $S=[q;w_1,w_2,\cdots,w_n]$ 的 Banzhaf 指标 $\boldsymbol{\beta}$ 与人口比例 $\boldsymbol{p}$(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 $\boldsymbol{p}=(12,4,2,1,1)$.若以 $\boldsymbol{p}$ 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 $\boldsymbol{\beta}=(1,0,0,0,0)$.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 $\boldsymbol{\beta}$ 与 $\boldsymbol{p}$ 相近似的加权投票系统.

令 $q=12$,则 $S=[12;12,4,2,1,1]$,仍然有 $\boldsymbol{\beta}=(1,0,0,0,0)$.

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 $\boldsymbol{p}=(12,4,2,1,1)$ 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 $\boldsymbol{\beta}$,结果是 $\boldsymbol{\beta}=(11/21,5/21,3/21,1/21,1/21)$,可以说是人口比例 $\boldsymbol{p}=(12,4,2,1,1)$ 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 $\boldsymbol{\beta}$ 与人口比例 $\boldsymbol{p}$ 之间的"距离"($\boldsymbol{\beta}$,$\boldsymbol{p}$ 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 $w_1,w_2,\cdots,w_n$ 和定额 q 的一个初始值,通常权重可取 $\boldsymbol{p}$ 或其近似,q 取稍大于 $w/2$ 的数;

2) 编程计算 Banzhaf 指标 $\boldsymbol{\beta}$ 及 $\boldsymbol{\beta}$ 与 $\boldsymbol{p}$ 的距离,当距离小于阈值时停止,否则转 3;

3) 改变 $w_1,w_2,\cdots,w_n$ 和 q,转 2.

迭代过程中的难点在尚没有调整 $w_1,w_2,\cdots,w_n$ 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,$\boldsymbol{\beta}$ 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 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".

拓展阅读 10-1

Shapley, Banzhaf 两种权力指标的公理化和概率解释

复习题

1. 记投票人集合为{A,B,C,…},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 是存在且唯一的吗？由 W_m 到 S 是存在且唯一的吗？考察 $W_m=([AB],[BC],[CD])$.

2. 例5中由民主党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，得出类似表2的结果.

第10章训练题

1. 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在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Nash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.

2. 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[23]？

(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Nash均衡.

3. 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：假设有多个人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？

4. 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，做出适当的假设，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?

5. 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[22].

(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?

6. 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?

7. 在一条东西向的笔直河流的南岸,依次分布着三个村庄 A,B,C,村庄 B 位于 A,C 之间,与 A,C 距离相等.现有两个商家,都计划选择三个村庄中的一个开一家杂货店.村庄 A,B,C 分别住着 200 人,200 人,120 人,假设所有人都只到离他最近的杂货店购物(如果他到两家杂货店的距离相等,就完全随机地选择一家).两个商家的杂货店应分别开在哪个村庄?当村庄 B 的人数发生变化时,结果是否会发生变化?

8. 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.

9. 对于 n 人投票系统 $S=[q;1,1,\cdots,1]$,其中有 n 个 1,设 n 为奇数,$q=(n+1)/2$,证明当 n 很大时,每人的绝对 Banzhaf 指标 β' 与 $1/\sqrt{n}$ 成正比(参考拓展阅读 10-1　Shapley,Banzhaf 两种权力指标的公理化和概率解释).

10. 对于加权投票系统 $S=[q;w_1,w_2,w_3]$,设 $w_1\geqslant w_2\geqslant w_3$,$q>w/2$,说明只有 5 种极小获胜联盟集,对应于 4 种不同的 Banzhaf 指标 $\boldsymbol{\beta}$.

11. 联合国安理会由 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 个非常任理事国投赞成票时决议方能通过.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.

12. (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.



更多案例……

10-1 进攻与撤退的抉择

10-2 让报童订购更多的报纸

10-3 合理分配效益的另一类模型

参考文献



郑重声明

高等教育出版社依法对本书享有专有出版权。任何未经许可的复制、销售行为均违反《中华人民共和国著作权法》，其行为人将承担相应的民事责任和行政责任；构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人进行严厉打击。社会各界人士如发现上述侵权行为，希望及时举报，本社将奖励举报有功人员。

反盗版举报电话　(010)58581999　58582371　58582488

反盗版举报传真　(010)82086060

反盗版举报邮箱　dd@ hep.com.cn

通信地址　北京市西城区德外大街 4 号

高等教育出版社法律事务与版权管理部

邮政编码　100120

防伪查询说明

用户购书后刮开封底防伪涂层，利用手机微信等软件扫描二维码，会跳转至防伪查询网页，获得所购图书详细信息。也可将防伪二维码下的 20 位密码按从左到右、从上到下的顺序发送短信至 106695881280，免费查询所购图书真伪。

反盗版短信举报

编辑短信“JB，图书名称，出版社，购买地点”发送至 10669588128

防伪客服电话

(010)58582300

[General Information]
书名=a
SS号=14695964