管理决策模型、方法与应用

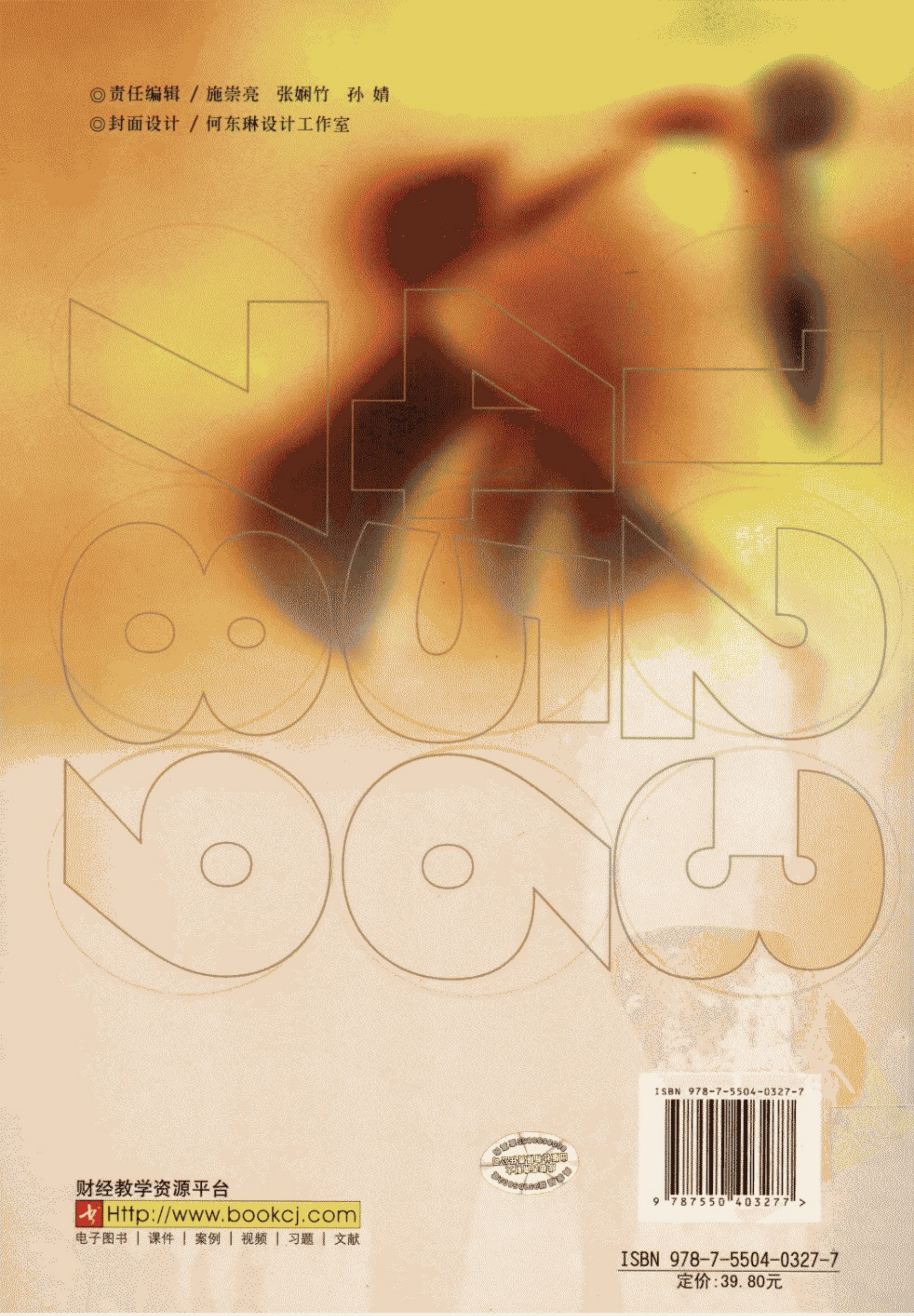
GUANLI JUECE MOXING FANGFA YU YINGYONG

（第三版）

主编 张照贵
副主编 鲁万波

西南财经大学出版社
Southwestern University of Finance & Economics Press

◎责任编辑 / 施崇亮　张娴竹　孙 婧
◎封面设计 / 何东琳设计工作室



财经教学资源平台
Http://www.bookcj.com
电子图书 | 课件 | 案例 | 视频 | 习题 | 文献

ISBN 978-7-5504-0327-7
9 787550 403277 >

ISBN 978-7-5504-0327-7
定价:39.80元

GUANLI JUECE MOXING FANGFA YU YINGYONG

管理决策模型、方法与应用

（第三版）

主　编　张照贵
副主编　鲁万波

西南财经大学出版社
Southwestern University of Finance & Economics Press

图书在版编目(CIP)数据

管理决策模型、方法与应用/张照贵主编. 3 版. —成都:西南财经大学出版社,2011.7

ISBN 978-7-5504-0327-7

Ⅰ.①管… Ⅱ.①张… Ⅲ.①管理决策—研究 Ⅳ.①C934

中国版本图书馆 CIP 数据核字(2011)第 128750 号

管理决策模型、方法与应用(第三版)

主　编:张照贵　副主编:鲁万波

责任编辑:施从亮　张娴竹　孙　婧

封面设计:何东琳设计工作室

责任印制:封俊川

出版发行	西南财经大学出版社(四川省成都市光华村街 55 号)
网　　址	http://www.bookcj.com
电子邮件	bookcj@foxmail.com
邮政编码	610074
电　　话	028-87353785　87352368
照　　排	四川胜翔数码印务设计有限公司
印　　刷	四川森林印务有限责任公司
成品尺寸	185mm×260mm
印　　张	19.25
字　　数	390 千字
版　　次	2012 年 6 月第 3 版
印　　次	2012 年 6 月第 1 次印刷
印　　数	1—2000 册
书　　号	ISBN 978-7-5504-0327-7
定　　价	39.80 元

1. 版权所有,翻印必究。

2. 如有印刷、装订等差错,可向本社营销部调换。

3. 本书封底无本社数码防伪标志,不得销售。

序

决策理论的创始人西蒙认为“管理即决策”，这足以表明决策是管理的核心问题。现代决策理论是一门较为年轻的学科，直到20世纪40年代，Von Neumann－Morgenstern建立了理性行为公理体系及其效用理论，以及运筹学及其优化理论的诞生，现代决策理论才正式形成。现代决策理论有着非常广泛的应用前景，是管理科学专业的重要基础，“管理决策模型与方法”也已成为管理科学专业的一门主干课程。但是目前管理决策模型与方法的教材多偏重于理论研究，这给经济管理类院校的教学带来了一定的困难。能够真正兼顾管理决策基本理论、基本方法和实际应用，并辅以计算机支持的教材还不多见。西南财经大学统计学院管理科学教研室的老师们经过多年的教学实践，编写了这本《管理决策模型、方法与应用》，这是一种有益的探索。

作为一本教材，本书有一些重要的特点：第一，精选了教学内容，对管理决策模型与方法教学内容的深度和宽度作了较为恰当的安排，决策方法既有发挥人的创造性思维的软方法，也有应用数学模型的硬方法，还有硬方法软化的决策模拟。该教材将这三类决策方法融为一体，是很有特色的。第二，书中注重决策的思想方法和实际应用，而不过分强调数学证明和推导，从而使读者不至于望而生畏。第三，在介绍各种决策模型的原理和方法的同时，对每种模型实际应用的例子都给出了计算的途径，并对如何应用计算机软件去实现作了具体的指导，使学生在学习的过程中就能够实际应用，并从中体会到管理决策的成就感。

正由于本书有上述特色，它不仅适合作为高等院校经济管理专业的教材使用，对于实际部门的管理人员来说，也是一本很实用的决策方法的参考书。

庞皓

2006年7月

于西南财经大学光华园

第三版前言

《管理决策模型、方法与应用》前两版出版后，由于其理论结合实践、侧重于实际应用的特点，在读者中引起较好反响。编者特此对第二版进行了较大幅度的修订，以飨读者。

为了使本书更完善，应用型特征更明显，本书第三版作了如下修订：

(1)对第二版本中的出现的一些文字、数据、符号的错误进行了修正。

(2)为了使本书的应用型特征更明显，新增加了案例。

(3)为了使读者更好地阅读，对第二版的内容进行了调整，使其更符合读者的阅读习惯。

由于编者水平有限，书中错误在所难免，望广大读者批评、指正。

编者

2011 年 4 月 2 日

第二版前言

《管理决策模型、方法与应用》出版后，由于其侧重于实际应用并辅以专用软件的支持，受到读者的欢迎，并被一些学校选为教材。

为了使本教材在体系上更加完善，在方法上更适合实际应用，在内容上更加丰富，我们在认真听取读者意见及总结教学经验的基础上，对本书作了如下一些修订：

(1)对第一版中出现的一些文字、数据、符号的错误进行了修正，对某些概念表述不十分确切，符号使用不符合惯例的地方进行了修改。

(2)为了使决策方法更注重于实际应用，对一些章节的内容进行了调整和修改，特别是对第六章的库存决策进行了较大的调整和修改。

(3)本书的参编人员有一些调整，其变动如下：第六章由黎春、张红历编写，第八章由张照贵、刘宏琨编写，第九章由鲁万波、李俭富编写，其余各章仍为第一版的作者。

由于决策的模型和方法发展很快，其应用领域也非常广泛，修订后存在错误也难以避免，期待广大读者的建议、批评和指正。

编者

2008 年 7 月

前　言

怀着忐忑不安的心情,《管理决策模型、方法与应用》一书终于要面市了。这本教材是西南财经大学统计学院管理科学教研室所有教师经过近三年的准备、写作、补充修改以及在教学实践中的体会,最后形成了现在这本比较完善的教材。

全书共分九章,第一章和第八章由张照贵编写,第二章和第三章由周游编写,第四章由吴江编写,第五章、第七章和第九章由鲁万波编写,第六章由黎春编写,每章的WinQSB软件计算部分以及附录A的WinQSB软件操作指南由鲁万波编写。全书由张照贵、鲁万波总纂定稿。

本书适合作为高等院校经济管理专业类本科生管理决策模型课程的教材使用,也适合作为大专院校各专业的现代决策论教材或参考书;本书也可作为MBA的教材或参考书;同时,本书还可作为企业的领导者和管理者决策时参考的工具书。

在本书的写作过程中,始终得到西南财经大学统计学院和西南财经大学出版社的支持,在此表示感谢。本书参考引用了一些专家学者的观点和材料,在这里也向原作者致以诚挚的谢意。同时,我们谨向对本书出版给予帮助的师长和朋友表示衷心的感谢。

由于编者水平有限,书中错误在所难免,恳请广大读者批评、指正。

编　者

2006年7月

目　录

第一章　决策概述 …… (1)
引例:九星连线 …… (1)
第一节　决策概论 …… (1)
第二节　决策模型、方法 …… (8)
习题 …… (13)

第二章　线性规划模型及其应用 …… (14)
引例:玻璃制品公司产品组合问题 …… (14)
第一节　线性规划模型 …… (15)
第二节　线性规划模型的一些概念 …… (17)
第三节　线性规划模型的求解 …… (18)
第四节　线性规划的 WinQSB 应用 …… (30)
习题 …… (38)

第三章　对偶理论与灵敏度分析 …… (41)
引例:家具生产问题 …… (41)
第一节　对偶线性规划模型 …… (41)
第二节　对偶问题的经济学解释——影子价格 …… (43)
第三节　对偶单纯形法 …… (44)
第四节　灵敏度分析 …… (46)
第五节　对偶线性规划的 WinQSB 应用 …… (53)
习题 …… (58)

第四章　线性规划模型的扩展 …… (62)
引例:背包问题 …… (62)
第一节　整数规划 …… (62)
第二节　目标规划 …… (70)
第三节　动态规划 …… (74)
第四节　非线性规划 …… (82)
习题 …… (91)

第五章　图与网络分析 …… (93)
引例:哥尼斯堡七桥问题 …… (93)

第一节　图与网络的基本概念 …… (94)
第二节　最小树问题 …… (95)
第三节　最短路问题 …… (101)
第四节　最大流问题 …… (108)
第五节　运输问题 …… (114)
第六节　指派问题 …… (136)
第七节　旅行推销商问题与中国邮递员问题 …… (140)
第八节　网络计划 …… (144)
习题 …… (160)

第六章　库存决策 …… (165)
引例:毛巾订购量问题 …… (165)
第一节　基本概念 …… (165)
第二节　确定型库存模型 …… (167)
第三节　不确定型库存模型 …… (175)
第四节　WinQSB 应用举例 …… (180)
习题 …… (186)

第七章　排队论决策 …… (188)
引例:高速公路收费站排队问题 …… (188)
第一节　基本概念 …… (188)
第二节　排队系统中的常用分布 …… (192)
第三节　排队模型 …… (194)
第四节　其他服务时间分布模型 …… (204)
第五节　排队系统的经济分析 …… (206)
第六节　排队系统的 WinQSB 软件应用 …… (208)
习题 …… (211)

第八章　风险决策 …… (213)
引例:期刊订购量决策 …… (213)
第一节　先验概率的风险决策 …… (213)
第二节　后验概率的风险决策 …… (225)
第三节　决策树 …… (229)
第四节　完全风险模型 …… (231)
第五节　风险模型的应用 …… (233)
第六节　马尔柯夫决策法 …… (236)
第七节　WinQSB 软件应用 …… (249)
习题 …… (255)

第九章　管理决策模拟 …… (258)
引例:比尔洗车场决策问题 …… (258)
第一节　模拟过程 …… (259)
第二节　建立模型 …… (260)
第三节　随机数的产生 …… (262)
第四节　蒙特卡罗法 …… (265)
第五节　模拟模型的完成与分析 …… (268)
第六节　模拟分析的例子 …… (273)
第七节　模拟的优点和局限性 …… (280)

附录 A　WinQSB 软件操作指南 …… (282)

附录 B　参考答案 …… (285)

参考文献 …… (293)

第一章　决策概论

引例:九星连线

有图 1.1 的九个点,用 4 条直线(一笔画)将其连在一起。

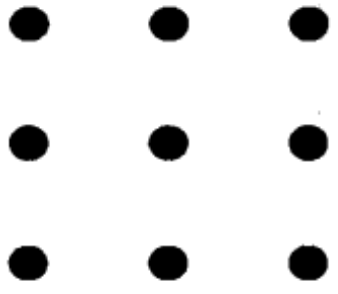


图 1.1

第一节　决策概论

一、决策及决策的发展

“人人做决策,天天做决策。”决策并不神秘,从狭义上讲,决策就是“决定”。日常生活中许多事情需要我们作出决定,如:出门是否需要带雨具;选什么课程;到商店买衣服,买不买,买什么样式,什么颜色等等问题,都需要作出决定。不过,对于这些小问题的决策,一般称为“决定”。决策一般可理解为“决定政策”,主要是指关系到国家发展、企业生存等一些大问题的决定。

管理理论中的决策是广义的,既包括对小问题的“决定”,也包括对大问题的“决策”,但侧重于大问题的决策。

“要胜曹公,需用火攻,万事俱备,只欠东风。”三国赤壁之战采用火攻这一正确的战术决策,为诸葛亮“三分天下”的战略决策的实现奠定了基础。可见,决策思想古已有之。决策是人类固有的行为之一,有人类就有决策。

中国是一个文明古国,在历史上可找出许多“运筹帷幄之中,决胜千里之外”的成功的决策范例:大禹分而治之的治洪方案;三十六计中的“围魏救赵”、“增兵减灶”等策略;战国时的“田忌赛马”,就有了博弈论的思想;宋朝丁渭“一举三得”的修城方案,已具有现代系统工程的思维方法;都江堰李冰父子提出的“深淘沙,低作堰”的治河方案,至今仍有参考价值。

但古时的决策主要是依靠领导者个人的知识和经验来进行的,既有许多成功的范例,也有许多“一失策顿成千古恨”的事例,如:项羽饮恨自刎乌江;赤壁之战的“蒋干中计”等。那时的决策被称为“经验决策”。

“战局瞬息万变。”决策的观念和方法的科学化，首先是在军事领域中逐步发展并开始形成的。随着历史进程的演进，战争的规模越来越大，武器越来越先进，参战的兵种越来越多样化，相互配合、协同作战的要求越来越高。军事指挥者越来越感觉到仅靠个人的经验和才能已不能应付日趋复杂的战争局面。1806 年，普鲁士的军事改革家香霍斯特创建了“参谋本部”体制，它以众多参谋的集体智慧来协助和支持军事长官的决策。

“规章制度是重复的结果。”人们在决策时发现许多决策问题会重复出现，若能寻找出一些共同的规律并加以制度化，则可大大提高决策的效率。

“一种科学只有成功地运用数学时，才算达到真正完善的地步。”20 世纪以来，特别是在“运筹学”产生以后，数学理论、方法的飞速发展，使得数学这门学科显得光彩夺目。第二次世界大战期间，“搜索论”被用于研究飞机如何有效地搜索德军潜艇；“规划论”被用来研究如何用最少的护航舰达到最好的护航效果；“对策论”（博弈论）被用来研究舰只在“神风突击队”（日本空军敢死队）攻击下的各种应急措施；“库存论”被用于研究军备、武器的合理库存量；“网络论”被用于战时的交通调度和管理；在“系统论”指导下建造的“自由轮”，被美国罗斯福总统认为是第二次世界大战获胜的重要工具。

第二次世界大战结束后，许多军事人员开始从事工业和商业经济活动，他们发觉经济活动与军事活动有许多相似之处，因此，许多用于军事领域的数学方法也可用于经济领域。

直到 20 世纪，决策才真正走上科学化的道路，而决策科学化的标志是“数学化、模型化、电子计算机化”。

“危机是革命的前夜。”由于数学方法（特别是运筹学）在军事和经济领域取得了巨大的成功，管理科学一词几乎成了运筹学的代名词，许多公司成立了专门研究数学模型的机构。他们相信，任何管理问题只要给出足够多的约束条件和变量数目，再设计出一个更比一个庞大的矩阵图、网络图及数学模型就都能迎刃而解。这种狂热在西方一直持续到 20 世纪 60 年代。

但是，一个象棋大师异于常规的怪招，往往使得储备了许多常规棋谱的机器人举手无措。第 12 届世界杯足球赛时，人们将 24 支球队的历史、队员的技术条件、身体素质、气候、环境、营养状况及过去的比赛记录等极为详细的情况输入计算机进行分析，希望用计算机来预测比赛的结果。但令人沮丧的是，第一轮比赛过后，计算机预测的结果与实际结果的误差高达 41.7%，而西班牙国王卡洛斯的预测误差却只有 16.7%。决赛的结果与计算机的预测更是相差十万八千里。

随着过分追求数学化、模型化的失败记录的日渐增多，决策方法数学化、模型化的上空笼罩着一片乌云。好走极端大概是人类的通病，许多公司撤销了数学模型机构，决策方法面临着一场危机。

“创造力是人类最高的能力，正是它把新事物引进了人们的想法、观念、方法和系统。”错误和教训使人们变得聪明起来，人们终于认识到，决策的科学化并不单纯是决策方法的数学化 —— 决策是一种复杂的社会行为，决策是人作出的，也要人来执行。决策方案是在特定的政治、经济、文化等条件下确定并实施的，这就必然涉及人

的、社会的、经济政治的、文化的、心理的诸多因素。而有些因素是难以或不能定量的，数学方法对这些因素是无能为力的。

因此，过分地追求决策方法的数学化、模型化将导致管理科学的死亡。于是许多发挥人的创造性思维的决策方法应运而生。如“头脑风暴法”、“强制联想法”等等。相对于决策的数学方法（硬方法），这些方法被称为决策的“软方法”。

美国的管理学家斯塔尔对此深有感触地说：“管理科学逐渐地越来越不像物质科学，而越来越像社会科学。人们渐渐明白了，问题的软的行为科学方面和硬的物质数量测定方面不仅是相互联系的，而且是不可分割的。”

二、决策和决策学

决策是为达到某种预定目标，按照某种条件、形式、程序和标准，从多种方案的比较中，确定最佳方案，帮助方案实施和进行反馈修正的过程。

决策学是为决策提供理论基础的科学，是一门研究决策这种社会行为合理原则的方法论科学，具体内容包括决策的分类、程序、方法、实施原则、信息反馈、经济性等。决策学是一门由科学群组成的“大科学”，是科学综合的结果。它涉及哲学、社会学、经济学、信息论、控制论、系统论、人才学、预测学、行为科学等众多学科。

三、决策程序

西蒙将决策过程分为参谋活动、设计活动和选择活动。具体可分为以下几个步骤：

（一）分析问题

“把一个问题说得清清楚楚，已经解决了一半。”

1. 问题与决策问题

问题是一切决策的开端，没有问题就无所谓决策。问题是主观与客观、理想与现实、应有现象与实际现象之间的矛盾，而矛盾是普遍存在的，所以，问题无所不在。要使问题成为决策问题要具有两个条件。首先，它是客观存在的，且是被发现、认识、理解并能清楚表达的问题；其次，必须是令决策者感到不满的问题才能成为决策问题。后面所说的问题均指决策问题。

2. 正确分析问题

（1）正确确定问题的类型

决策问题一般分为“开创型”和“常规型”两类。

开创型决策一般是涉及一些重大问题的决策，且有时用现存的理论和方法均不能对问题进行清楚的认识和表述，必须用创造性和预见性的逻辑思维进行概念开发。这类问题一般由该部门的最高领导来进行决策。彼得·*F*. 杜拉克认为：“他们的决策，是最高层次的、观念方面的、少数的重大决策。他们需要的是决策的冲击，而不是决策的技巧；是好的决策而不是巧的决策。”

常规型决策也称例行性决策，一般属于战术或战役决策。它一般是经常重复出现的、有共同规律的一些小问题的决策。这类问题一般由下一级的管理者来进行决策。

(2) 正确界定问题

所谓界定问题,就是把问题说清楚。一般要就什么问题、什么时间、什么地点、什么幅度等四个方面进行清楚的说明。界定问题就像医生诊断病情,如果病情都未诊断清楚就开药方,显然是达不到治病目的的。

(3) 以差距的形式表达问题

美国管理学家克普勒和崔可把“应有现象与实际现象的偏差”定义为问题,即“问题(差距) = 标准 - 现状”。用这样的形式表达问题,一是可以把问题界定得更加清楚;二是可以衡量这样的差距是否令决策者感到不满,是否有必要付出代价来消灭或缩小这种差距。正如医生看病是以健康人为标准来判断受诊者是否有病及病的严重程度的。

(4) 找到产生问题的真正原因

确定问题后,还必须找到产生问题的真正原因,才能对症下药。要找到问题产生的真正原因,应从事物的特点与变化中去寻找,并对这些特点与变化进行验证。

所谓特点的寻找,就是将界定问题的资料归为“是问题”与“不是问题”两类,然后就会产生一种对比,对比的结果,就会显示出各项特点来。

变化就是一种新的、不同的事物或状况。这些变化很可能就是产生问题的真正原因。

如果在寻找原因时未发现对比,则表明“界定”的问题并不准确,需再加“界定”;某些情况下,原因的原因构成事物的因果链,必须层层深挖。有时一项特点中的变化还要和另一项或几项变化结合起来,才能成为问题的真正原因。

是否是产生问题的真正原因,还必须加以验证。所谓“验证”,就是对一个问题原因的任何解释,都去寻求例外的一种程序。如案件审理中的“例外原则”,就是验证的一种运用。即即使有若干条理由支持“某人是罪犯”的假设,但只要有一个理由能真正证明他不可能犯罪,则其他的理由都不能成立,如成语中的“疑人偷斧”。因为只可能有一项变化是造成该问题的原因,这一原因会产生确定的事实,而不会有例外,也不会有假设。所谓“例外原则”,就是要对问题产生的各种可能原因进行反面论证。

在分析问题的过程中,一位高明的管理者,应如一位高明的侦探一样,能找出有关的资料并加以运用,一点点地缩小侦查范围,直到查出真正的罪犯。

(二) 确定目标

“有的放矢”和“南辕北辙”给人的启示是,如果目标不清楚甚至目标的方向都错了,则决策方案实施得越好,造成的损失可能越大。管理上流行的“管理效能 = 目标方向 × 工作效率”的公式,充分说明了目标确定的重要性。

1. 决策目标的概念

决策目标是对决策问题的本质概括和抽象。

2. 确定目标的要求

(1) 目标必须具体明确。第一,要确定价值准则。所谓价值准则,是指决策者希望达到的理想、期望和意图的科学的、定性与定量的表述。价值准则一般包括物质性价值准则和精神性价值准则。正确的价值准则是准确确定目标的前提,也是评价决策方案和实施结果的依据。第二,应使目标单义,即任何人对目标都应只有一种理解,因

此，目标应尽可能量化。第三，目标要落实。决策目标是由总目标到具体目标的多层次系统，即分层目标结构。上层次的具体目标应是下层次的总目标，不能照搬总目标作为下层次的具体目标。

（2）目标是可以达到的。有时，由于主、客观条件的限制，不能达到理想的目标，这时需要订立一些“适应性目标”以保证目标是可以达到的。此外，还应考虑决策权与执行权的一致性。

（3）规定实现目标的时间。“机不可失，失不再来。”没有规定时间的目标不是完整的目标。要避免“马后炮”的决策，就必须规定达到目标的时间。

（三）拟订方案

“狡兔三窟”的故事说明了拟订方案的重要性。所谓决策就是在多方案中选优，没有选择就无所谓决策。人们把没有选择的单方案决策讥讽为“霍布森决策”。霍布森是英国剑桥的一个贩马者，他允许买马者可以挑选马匹，但又加上只许挑选靠近门边的那匹马，实际上是不让人选择。管理上有句名言：“如果你感到似乎只有一条路可走，这条路很可能是走不通的。”一般而言，可供选择的方案越多，说明决策的环境越好。

1. 拟订方案的要求

从理论上讲，拟订方案要做到整体的详尽性和方案间的排斥性。

方案的详尽性是指在拟订的方案中应把达到目标的所有可能的方案都包括无遗，即所谓“条条道路通罗马”。

在实际工作中，只可能选择一个方案实施，因此，所拟订的方案间必须是相互排斥的。如果一个方案包含另一个方案，或两个方案可同时采用，或方案虽是从不同角度出发但彼此无法比较优劣，则这些方案均违背了相互排斥性。

2. 拟订方案的阶段

拟订方案一般可分为大胆寻找与精心论证两个阶段。

方案拟定的过程很大程度上是一个创新的过程。大胆寻找阶段就是要充分发挥人的创造性思维，找到好方案。创造性人人都有，但需要开发。创新思维的障碍有社会和思维两方面的，要打破这些障碍，需要经常做一些软化头脑的柔软体操。如丰田汽车公司的“传票卡制”的管理方法就是从小孩倒着放电影中受到启示的。因此，在大胆寻求方案时，波亚说：“寻求真理时，正确答案的作用比错误答案小得不可比拟，而且一项新的假设不管多么荒唐，总是比对现有假设作一点点改进有用。”

如果说方案的寻找需要的是创新精神和想象力，那么在精心设计阶段则需要求实精神和冷静的头脑。“没有缺点就是最不能容忍的缺点”，世界上没有十全十美的事情，决策方案也是如此。由于方案的提出者对自己的方案一般都有一种偏好，故对优点的估计较充分，而对其缺点容易忽视。因此，在设计阶段首先要估计该方案有些什么缺点，并提出一些补救措施。其次，应对方案的可行性进行研究并对各种后果进行估计，这些后果一般包括：直接后果、间接后果、长远后果、社会后果和无形后果等。

（四）评价方案

“兼听则明，偏信则暗。”在评价方案时要充分发扬百花齐放、百家争鸣的精神。任何方案都是优劣并存的，要对这些方案作出实事求是的综合评价，并提出参考性的

方案供领导者选择。

评价方案的内容主要有：审定方案、评价方案、推荐方案。

审定方案要就决策问题的分析是否清楚、原因是否找到、目标是否正确及清楚、方案是否有细节、有无缺点及后果估计等等进行审定。

评价方案时多采用“反面论证法”，要特别注意倾听反面意见，一个反面意见有时会是一个很好的补救措施。

推荐方案就是在对众多方案进行综合评价的基础上，选择几个较好的方案供领导者选择。

评价方案的人员要注意选择“反馈型人才”。这种人才一般具有思维活跃而敏锐、知识面广、吸收新生事物快、综合分析能力强、敢于直言不讳、只有真理精神、没有权力欲望等素质。

（五）选择方案

“成败在此一举。”选择方案就是人们常说的“拍板”，这是领导者的主要职责。指挥员高于参谋和其他人员的地方就在于要能敢于和善于拍板。

在选择方案时会遇到许多困难。选中的方案将会付诸实际，一旦选错方案有时会产生严重的损失甚至不可弥补的后果。一失足而成千古恨的教训会使人犹豫不决；许多方案优劣互见，没有一个方案比其他方案更具明显优势，这也会使领导者难以下定决心。

领导者一般属于“指挥型人才”。这种人才一般具有高瞻远瞩的战略眼光、大智大勇、果敢、善断、永不枯竭的事业进取心等素质。

（六）实施方案

“没有实施就没有决策。”决策方案必须是能够付诸实施的方案，否则就是没有实用价值的方案。实施应该是决策中最有成效的一个环节，也可能是最困难的一个环节。

所谓实施，就是将选中的方案付诸实际运行。为了使方案能够实施，首先，在此之前的各环节都必须树立“一切为了实施”的观念，如目标能否达到、方案是否可行等。其次，方案的顺利实施需要有正确的宣传、上级的支持、群众的参与和正确的组织形式等。有人认为：“一个实施方案就是一次变革，当变革将造成对某种现有的相对满足的社会系统的改变时，如组织结构、观念、地位、待遇、工作条件、领导关系等方面的改变时，如处理不好，甚至考虑不周，就会产生阻力。”

在人员选择方面，需要选择“执行型人才”来实施方案。这种人才具有忠实坚决、善于领会领导意图、埋头苦干、任劳任怨等素质。

（七）信息反馈

“不能只问耕耘，不问收获。”方案实施后，并不是万事大吉。从控制论的角度看，经济管理过程就是一个不断进行控制的过程，控制得越好，管理成效就越高，但控制必须借助信息反馈才能实现。

反馈本是机械系统中的一种作用，指人或机械控制一个运动并要求它达到一定的目标，在此过程中往往会产生偏差，把这个偏差反映回去，使它影响产生这个运动的人或机械的控制活动。

心理学中将信息反馈称为结果的知识。结果的知识对人的活动具有重要的意义。如一个人打靶，若没有报靶人，则无论射击多少次也不能提高射击技术。美国人桑代克曾做过一个让人射击三千多次，但不将射击结果告诉射击人的实验，结果射击人的命中率没有一点提高。

决策方案在实施过程中，一般会出现三种情况。其一是运行正常。这时让它继续运行。其二是偏离目标。其原因在于：一是主观因素，如执行人对决策目标、方案理解等方面的偏差或执行人选择不当等，其措施有沟通或调换执行人等；二是客观因素，如某种设备在实施时已经落后或河水变质等，这时要根据具体原因修改方案以达到目标。其三是达不到目标。其原因可能是目标不合理，也可能是现有方案均达不到目标等，措施有修订目标、重拟方案或对方案进行重大修改。

信息反馈不仅在方案实施阶段进行，而且贯穿于整个决策过程。

四、决策分类

（一）按决策范围分

决策按范围可分为宏观决策和微观决策。

宏观决策是指对整个社会、经济、政治具有普遍重大影响的决策。国家、地区、部门范围内的决策都可称为宏观决策。

微观决策主要是指基层企事业单位的决策。

（二）按决策目标层次分

决策按目标层次可分为战略、战术和战役决策。

战略决策是解决带全局性及与外界有密切联系的重大问题的决策。它具有全局性、方向性、战略性和长期性等特点，如2020年GDP在2000年的基础上翻两番、企业中新产品的开发等。

战术决策是为保证战略决策的实现而制定的。它具有局部性、暂时性和策略性等特点，如要开发产品（战略决策），需要进行设备改装、职工培训等（战术决策）。

战役决策主要是对具体作业中的问题的决策。它具有短期性和技术性等特点，如要完成职工培训（战术决策），战役决策则要确定培训的人员、内容、时间和地点等具体内容。

（三）按决策问题是否会重复出现分

决策按问题是否会重复出现可分为重复性决策和一次性决策。

重复性决策是指在实际中会经常出现的问题的决策。如工厂中物资的购进与库存、广告宣传和新产品的开发等均是重复性决策问题。这种决策一般用决策“硬技术”解决。

一次性决策是对只出现一次或不经常出现的问题的决策。如工厂是否转产、重大项目的投资、兼并、重组等问题的决策。这种问题一般用决策“软技术”解决。

（四）按对决策对象的信息了解程度分

决策按对决策对象的信息了解程度可分为确定型决策、不确定性决策和风险型决策。

确定型决策是对决策对象的信息有完全了解，不存在任何的不确定性的问题的

决策。这种决策一般是通过数学上解最优化的方法来选择方案。如线性规划求解的问题。

不确定型决策又称为完全风险决策，是指对决策问题可能出现若干种情况（随机现象）但不知道各种情况发生的概率的问题的决策。如新产品开发，可能畅销，也可能滞销，但不知道畅销、滞销的概率，这时对是否应该开发这种新产品进行决策。

风险决策是介于确定型与不确定型中间的一种决策，是指对决策问题可能出现若干种情况（随机现象）且知道各种情况发生的概率的问题的决策。

第二节　决策模型、方法

决策方法是指人们进行决策时借助的科学方法，也称为决策技术。决策技术一般分为“硬技术”和“软技术”。决策的“硬技术”主要是指一些数学方法和模型，特别是运筹学的方法。决策的“软技术”主要是指建立在心理学和行为科学基础上的、发挥人的创造性思维的方法。

一、决策的硬技术

为论述方便起见，我们把下面介绍的一些常用的硬技术统称为模型。

（一）规划模型

管理者经常需要面对资源有限（如资金短缺、能源紧张、运输繁忙、信息不足、时间不够和精力有限等）而需求无限（更多的资金、更高的运输能力、更多的信息、更充裕的时间和更充沛的精力等）时如何合理安排资源的问题。

规划模型就是在满足既定条件下，按某一衡量指标（目标函数）来寻求最优方案的方法。

规划模型按对约束条件的要求不同可分为：线性规划模型、非线性规划模型、单目标规划模型、多目标规划模型、整数规划模型、静态规划模型和动态规划模型等。

规划模型可广泛应用于解决运输安排、厂址选择、饲料配合、生产安排、合理下料、航空班次、设备配置和平稳生产等诸多问题中。

（二）风险模型

“你想要的是白猫，抓出来的却是黑猫。”大多数管理者都讨厌风险，但风险问题却又无处不在，这是因为社会经济现象中存在大量的随机现象。

风险模型是根据概率论和数理统计知识来研究如何在风险决策问题中合理选择方案的一种方法。

（三）库存模型

“多多并非益善，太少也非妙事。”库存模型就是研究如何确定一个合理的库存量，从而使涉及相互消长的库存费用（存储费用、订货费用、脱销费用）总和最少的一种数学方法。

库存模型可用于解决工厂原材料的合理库存、图书馆藏书的数量和种类、银行流动资金的合理储备量、军备武器的库存、粮食的合理库存和国家之间长期关系的稳定

研究等诸多问题中。

（四）排队模型

排队和拥挤令人讨厌，但这种现象却又随处可见：到食堂买饭、到图书馆借书、上下班乘车、到银行取钱等，这是“人”排队等待服务；工厂中的原材料和半成品等待加工、坏的机器等待修理、货物等待装卸、信件等待发送、水库中的水等待排泄、布机上的断头等待接头等，这是“物”排队等待服务。以上这些是有形的排队等待服务。此外还有打电话占线、飞机等待航道着陆等无形的排队等待服务。

人们经常批评排队和拥挤的现象，并要求增加服务设施。但是，大量增加服务设施（窗口），则又可能造成“服务”排队等待“顾客”，从而增加服务成本。

排队模型就是在使总费用最少的情况下，合理确定“窗口”数量和服务时间等问题，从而减少排队和拥挤的一种有用的工具。

排队模型与库存模型有相似之处，但排队模型是研究如何控制输出，而库存模型是研究如何控制输入。

（五）对策模型

“兵来将挡，水来土掩。”对策模型就是研究在有竞争对手的条件下，如何选择自己的最优方案的一种决策方法。近年来，由于对策模型的思路已成为主流经济学的基础，对策模型已独立成为一门专门的课程，即博弈论。

对策模型广泛用于体育竞赛，产品、价格竞争，人才选择，民事、经济及政治纠纷等问题的决策。

（六）网络模型

网络模型是利用图形与网络的知识去解决实际工作中“卡脖子”的决策问题。一项复杂的任务往往包括许多各不相同而又密切相关的部分，各部分需要的人力、财力、物力、时间及先后顺序又不平衡，因而容易出现“卡脖子”的环节，从而影响整个任务的完成。网络模型就是解决如何在有许多变化的情况下去计划、组织和控制一项任务的完成，以节约时间和资源的有力助手。

如美国的“阿波罗登月计划”是一个庞大而复杂的计划，该计划总共耗资380亿美元，动员了42万人、120所大学和2万家公司，生产了700万个零件。他们利用网络模型从而提前两年完成了该计划。

网络模型可用于邮递员的路线选择，城市水管、下水道的设计安排，无线电线路图的设计和计划任务的设计等。

（七）马尔柯夫模型

俄国数学家马尔柯夫在研究一种所谓“布朗运动”的现象时发现有这样一种现象：当它在“将来”处于什么状况与且仅与“现在”处于什么状况有关，而与“过去”处于什么状况无关时，就可根据某种现象现在的状况去预测它将来的状况，从而作出相应的决策。

马尔柯夫模型常用于对市场销售、设备的更新和维修、最优工作分配和仓存量的研究等领域的问题的决策。

二、决策的软技术

（一）头脑风暴法

这是美国奥斯本博士所倡导的一种发挥集体创造力的一种方法。

头脑风暴原是医学上的一个词汇，原指精神病人发病时胡言乱语的现象，这里借用这个词汇来比喻放开思想、不受约束和畅所欲言的意思。

这个方法是召开5至10人的专家（是指思维开放、某一方面有独到见解的人）会议。为保证到会的人都能畅所欲言，会上有4条规定：① 不作结论，不许反驳；② 追求数量，不求质量；③ 欢迎奇思异想；④ 不许兜售，但可相互启发。

（二）特尔菲法

这个方法是20世纪50年代初由美国兰德公司首创的。当时美国空军委托兰德公司研究这样一个课题：若苏联对美国发动核袭击，袭击的目标会选在什么地方？将产生什么后果？遇袭后的美国是否有还击的力量等问题。该课题很难用定量的方法进行研究，兰德公司用专家法进行研究，为了保密，取名为“特尔菲项目”。

特尔菲是古希腊的历史遗址，是传说中的阿波罗神殿所在地。该神以预言灵验而著称，之所以预言灵验是因为他经常派人到世界各地搜集聪明人的意见来解答人们的疑难问题。用特尔菲法命名这种方法，含有灵验和集中众人智慧的意思。

特尔菲法也是一种专家分析法，与头脑风暴法不同的是，这种方法要求参加的人员不见面，以免发生权威左右局面的现象。该方法采用信函征求意见的方式，有匿名、反馈和统计处理三个特点，一般反复进行三次。对一些重大问题的决策常采用特尔菲法。如：美国兰德公司用特尔菲法对科学的突破、人口增长、自动化技术、航天技术、战争的可能性和新武器系统等进行了预测（由82位专家组成）；美国和加拿大锻造协会对粉末冶金和冷锻技术的发展及其对锻造业的影响用此法进行了预测；美国马尼托巴大学请40位专家用此法就能源和环境问题进行了预测；中国工业机械部用此法对大功率激光器的发展进行了预测等。

（三）特性列举法

这是美国尼布拉斯加大学克劳福德教授提倡的一种创新思考法，最先用于对产品的改进。

这种方法的实质是从“分解”中找出新的方法。他认为任何一个事物，都可分解为名词特性、形容词特性和动词特性，将这些特性一一列举后，对每一个方面进行检查，看有什么可改进的地方，进而产生新的想法。

如想生产一种新式的水壶，从整体上有时会觉得很难下手。这时可先将其分解如下：

名词特性	组成：壶嘴、把手、壶盖、壶身、壶底等；
	材料：铁皮、铜皮、铝合金等；
	制造工艺：焊接、压制；
形容词特性	颜色：材料本色；
	形状：椭圆形；
	重量：轻、重；

动词特性　　冲水是否方便、壶把是否烫手、能否保温等。

然后，对每一个特性进行分析和改进，则可生产出千万种不同的水壶。

（四）缺点列举法

缺点列举法是从特性列举法中衍生出的一种方法，又称为“吹毛求疵”法。

这种方法的基本思路是：世界上没有十全十美的事物，即任何事物都有缺点，找出这些缺点并加以改进，便可产生新的东西。管理上有句名言：“没有缺点就是最不能容忍的缺点。”据统计，20 世纪真正的发明创造并不多，60%～80% 新产品的问世是来自顾客的意见。

国外一些企业专门成立一个“自己打击自己”的部门，在这个部门工作的人员对本企业的管理、产品和服务等各方面只许说坏，不许说好。因此，许多新的管理方法和新的产品源源不断地产生出来。

（五）形态分析法

这是由美国加州理工学院兹维基教授提出的方法，实质是通过不同的“组合”而产生新的思路。组合包括直接组合与飞跃式组合，如小孩玩积木。该方法最先用于包装的设计和改进以及新产品和新技术的开发。

具体步骤是先分解后组合。某种食品的包装，如可分为：

包装材料 —— 玻璃瓶、塑料、马口铁、特种纸、铝合金；

包装形状 —— 小口瓶、圆形、方盒、球形、扁状、袋式；

包装重量 —— 1 千克、2 千克、0.5 千克、0.1 千克等。

这样，可产生 120 种可能的组合。而某种产品的滞销，有时可能仅是某一方面不满足消费者的需要。

（六）替换措施表列法

这是奥斯本博士发明的一种创新方法。他认为人们头脑中的一些先入为主的、经验性的思想是很顽固的。这种顽固性有碍于创新思想的产生，而通过一系列的启发式提问将有助于打破这种顽固性，从而使人们豁然开朗。比如：

（1）除此之外，还有没有别的办法可用呢？把现在的状况改变一下，如用途、颜色、运动形式、声音、味道、形状或花式的改变，会产生什么结果呢？

（2）能不能用别的方法呢？与此相似的东西还有什么呢？能不能模仿呢？过去有过相类似的东西吗？今后有什么思路可循呢？

（3）能否增加一些东西呢？添加组件、延长时间、增加次数、增加长度、增加附加价值、再粗些、再厚些、再浓些、再夸张一些、再重一些会怎么样呢？

（4）能否减少一些东西呢？减少组件、压缩体积、减轻重量、减少时间、再薄些、再稀释一些、再简单一些会怎么样呢？

（5）能否采用代换品呢？如用别的材料、别的工艺、别的配方、别的能源、别的方法、别的设备、别的人物、别的声音会怎样呢？

（6）能否综合起来呢？加起来、挂起来、拉一拉会产生什么结果？能否分割开来呢？

（7）能否互相替换呢？顺序改变、结构改变、时间改变、因果互换、元件互换、速度变换会怎样呢？

(8) 能否反过来呢?任务颠倒、上下颠倒、左右颠倒、里外颠倒、正反颠倒会怎样呢?

(9) 能否组合呢?方案组合、任务组合、目标组合、材料组合、部件组合、人员组合、设备组合等,这样组合后的结果怎样呢?

(七) 综摄法

这是美国瑟瑟·里特尔公司的威廉·戈登博士发明的一种创新方法,首先用于设计和改进产品。

这种方法的本质是通过"强制类比"从而产生新的思路,即所谓的"变熟悉为陌生"或"不识庐山真面目,只缘身在此山中",对某一领域非常熟悉的人员,反而会产生许多老框框和老经验,容易使思维受到束缚从而妨碍创新思维的产生。这时,采用类比的方法从而产生一些好的主意来。综摄法又称类比法、提喻法。如想设计一种夜视装置,不是让设计师直接设计,而是先由生物学家介绍猫头鹰的眼睛是如何构造的从而类比产生新型的夜视器。

类比的方法有拟人类比、直接类比、象征类比和幻想类比等。

(八) 焦点法

这是美国人怀亭所发明的一种"强制联想"的方法,其特点是"殊途同归"。

这种方法的思路是,为达到预定的目标(焦点),把看来完全风马牛不相及的事物进行强制联想,从而产生新的思维和方法。收录机的快录功能就是与飞机进行强制联想产生的。

(九) 方案前提法

这是国外近期产生的一种软方法,特点是"离题万里的讨论"。

在现实生活或工作中,解决一个问题往往会产生几种方法,若持这几种方法的人数差不多,直接讨论这几种方法的优劣可能会出现"公说公有理,婆说婆有理"而相持不下的局面,且人们很难公正、客观地评价自己方法的优劣。为克服这种弊病,方案前提法的出发点是:每种方法的成立都有一定的假设前提(原因),而方案(结果)是否正确的关键在于前提是否成立,且前提还有前提,如此等等。因此,这种方法是不直接讨论方案的优劣,而是讨论前提是否成立。这种"离题万里"的讨论,目的是使参与人不知道讨论的题目与自己所提方案有什么关系,从而能客观、公正地进行讨论并一步一步倒推回去,从而找到一个较优的方案。

(十) 潜在问题分析法

"谨防因丢失一颗马钉而丧失一个王国。"任何一个方案都不可能是十全十美的,既有优点,也有缺点。对可能出现的问题分析得越充分、越透彻,则决策成功的可能性就越大。潜在问题分析法就是从这个目的出发,充分运用预测的方法对决策方案可能出现的各种问题进行充分研究,对这些问题产生后的严重性、危险性进行分析并制定预防措施。

(十一) 决策模拟

决策模拟严格地讲并不是一种决策的软技术,而是"硬技术"的软化。

决策模拟企图建立一个社会经济现象的实验室。决策模拟是在数学模型上进行实验,即给定不同的数据使模型运转。它不追求理论上的最优解,而注重实际运行的

效果。许多大公司做计划时都采用模拟的方法。

决策模拟有许多种类，如“仿真技术”、“启发式搜索”、“管理演习”和“蒙特卡罗模拟”等。

习题

1. 什么是决策和决策学?决策有哪些分类?
2. 决策的基本程序有哪些?
3. 为什么说决策学是一门综合性学科?
4. 什么是决策问题?如何分析决策问题?
5. 什么是决策目标?如何确定决策目标?
6. 如何拟定决策方案和评价决策方案?
7. 决策实施后，一般会产生哪几种情况?应采取什么措施?
8. 常用的决策模型有哪些?
10. 常用的创新决策方法有哪些?
11. 头脑风暴法有什么规定?
12. 特尔菲法有何特点?
13. 特性列举法认为任何事物都可分解为哪些特性?
14. 综摄法和焦点法的本质特征是什么?
15. 如何在学习、工作中发挥创新能力?

第二章　线性规划模型及其应用

引例:玻璃制品公司产品组合问题

某玻璃制品公司有三个生产车间:

车间 1—— 生产铝框和硬制件;

车间 2—— 生产木框;

车间 3—— 生产玻璃和组装窗和门。

三个车间分别完成产品的一个工序。由于某些产品销售量下降,高层决定调整公司的产品线,将不盈利的产品停止生产并撤出生产能力来生产两个新产品:8 英尺(1 英尺 = 0.304 8 米,全书同) 的铝框玻璃门和 4 英尺 ×6 英尺的双把木框窗。

现在公司管理部门要考虑以下两个问题:

(1) 公司是否该生产这两个新产品?

(2) 若生产,两个新产品的生产组合如何(即每周各生产多少)?

为了回答上面两个问题,首先需要收集以下信息:

(1) 各车间的可得生产能力;

(2) 生产每一种产品一个单位各需要每个车间多少能力;

(3) 每一产品的单位利润。

表 2.1

生产每个单位产品所需时间			
	门	窗	资源:每周可得时间
车间 1	1 小时	0 小时	4 小时
车间 2	0 小时	2 小时	12 小时
车间 3	3 小时	2 小时	18 小时
单位利润(美元)	300	500	

这是一个典型的产品组合问题(Product Mix Problem),需要对公司的资源进行最优分配以生产合适数量的产品,从而获得最大的利润。为此,需建立一个数学模型(线性规划模型),来对问题进行描述和求解。

任何一个组织的管理都面临无数的决策问题,线性规划是帮助管理这些决策的一个功能强大的问题解决工具,它是实现现代科学管理的重要手段之一。其主题就是要找出活动的最佳组合 —— 做哪些以及做到什么程度。

自 1947 年 G. B. Dantzig 提出一般线性规划的单纯形法之后,线性规划在理论上日趋成熟。尤其是在电子计算机能处理成千上万个约束条件和决策变量的线性规划之后,其应用领域就更为广泛了,实用性也更强了。目前,线性规划在工业、农业、商

业、交通运输业和经济管理等领域都有重要应用。最常见的三类线性规划问题是：资源分配(Resource - Allocation)问题、成本收益平衡(Cost - Benefit - Trade - Off)问题以及网络配送(Distribution - Nework)问题。

线性规划应用数学模型对所研究的问题进行表述。所谓线性(Linear)，是指模型中的表达式是各个变量的线性组合。

第一节　线性规划模型

一、线性规划的一般模型

先定义一些代数符号来代替上面问题中的变量。

令：D—— 门的生产数量；W—— 窗的生产数量；P—— 利润。

该工厂的目标是要获得最大利润，假设生产 D 个门和 W 个窗，则利润 P 为：

$$P = 300D + 500W$$

但是，由于每个厂的资源 —— 时间是有限的，所以生产量必须满足：

车间 1 资源约束：$D \times 1 + W \times 0 \leqslant 4$

车间 2 资源约束：$D \times 0 + W \times 2 \leqslant 12$

车间 3 资源约束：$D \times 3 + W \times 2 \leqslant 18$

上面三个约束称为资源约束。

又由于门和窗的生产量不会为负数，所以要求：$D \geqslant 0, W \geqslant 0$。此称为非负约束。

这样，可用下面模型描述上问题：

$$\text{Max } P = 300D + 500W \text{(目标函数)}$$

$$s.t. \begin{cases} D \leqslant 4 & \text{(资源约束 1)} \\ 2W \leqslant 12 & \text{(资源约束 2)} \\ 3D + 2W \leqslant 18 & \text{(资源约束 3)} \\ D, W \geqslant 0 & \text{(非负约束)} \end{cases}$$

一些术语：决策变量，是用符号表示的可控制的因素，如 D、W；可行解，指满足所有约束条件的决策变量的取值，反之为不可行解；可行解的集合称为可行域；最大(小)化目标函数的可行解称为最优解。

可见，要建立一个线性规划模型，首先要提出和分析问题，收集相应数据，明确约束条件。

一个线性规划模型的特征有：

(1) 用一组决策变量如(D, W) 表示某一个方案，这组决策变量的值就代表一个具体方案，如(1,2) 就表示生产 1 个门和 2 个窗。一般地，这些决策变量取值非负。

(2) 存在一定的约束条件，这些条件可以用线性等式或者不等式来表示。

(3) 都有一个要求达到的目标，它可以用决策变量的线性函数来表示。

满足上述三个条件的模型称为线性规划的数学模型。其一般形式为：

目标函数 $\text{Max(Min)} z = c_1 x_1 + c_2 + x_2 + \cdots + c_n x_n$

$$\text{满足约束条件:}\begin{cases} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n \leqslant (=, \geqslant) b_1 \\ c_{21}x_1 + c_{22}x_2 + \cdots + c_{2n}x_n \leqslant (=, \geqslant) b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ c_{m1}x_1 + c_{m2}x_2 + \cdots + c_{mn}x_n \leqslant (=, \geqslant) b_m \\ x_1, x_2, \cdots, x_n \geqslant 0 \end{cases}$$

一般地,约束变量要求非负,但有时候也允许无此非负约束。

二、线性规划模型的标准型

为探讨模型性质和设计算法等的方便,需将上面的线性规划一般形式的模型变化成标准型,规定标准型为:

$$\text{Max } z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

$$\begin{cases} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n = b_1 \\ c_{21}x_1 + c_2x_2 + \cdots + c_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ c_{m1}x_1 + c_{m2}x_2 + \cdots + c_{mn}x_n = b_m \\ x_1, x_2, \cdots, x_n \geqslant 0 \end{cases}$$

标准型中要求各约束条件等式右端值 $b_i \geqslant 0, i = 1, 2, \cdots, m$。若某一等式不满足,则在该等式两端乘以“-1”。

上面标准型也可以简写为:

$$\text{Max } z = \sum_{j=1}^{n} c_jx_j$$

$$\text{s. t.} \begin{cases} \sum_{j=1}^{n} a_{ij}x_j = b_i, i = 1, \cdots, m \\ x_j \geqslant 0, j = 1, \cdots, n \end{cases}$$

用向量和矩阵符号表述时,标准型为:

$$\text{Max } z = CX$$

$$\text{s. t.} \begin{cases} \sum_{j=1}^{n} P_jx_j = b \\ x_j \geqslant 0, j = 1, \cdots, n \end{cases}$$

其中:$C = (c_1, c_2, \cdots, c_n)$　　$X = (x_1, x_2, \cdots, x_n)^T$

$P_j = (a_{1j}, a_{2j}, \cdots, a_{nj})^T$　　$b = (b_1, b_2, \cdots, b_n)^T$

用矩阵表述标准型更简洁,为:

$$\text{Max } z = CX$$

$$AX = b$$

$$X \geqslant 0$$

其中:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m_1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

$= (P_1, P_2, \cdots, P_n), 0 = (0, 0, \cdots, 0)^T$

A 称为约束条件矩阵，是 $m \times n$ 维的系数矩阵。一般有 $m > n, m, n > 0$；b 为资源向量；C 为价值向量；X 为决策变量矩阵。

一般地，实际中的线性规划问题，都要转化成标准型来求解（很多计算机程序求解时不需要这样的转化）。可以通过如下处理将一个一般形式的模型转化成标准型：

（1）若要求目标函数实现最小化，即 Min z，这时只需要令 $z' = -z$，即可将目标函数变为 Max z'。

（2）若约束方程为“≤”不等式，可以在该不等式左端加入一个非负松弛变量，就把原“≤”不等式变为等式。

（3）若约束方程为“≥”不等式，可以在该不等式左端减去一个非负松弛变量，就把原“≥”不等式变为等式。

（4）对取值无约束的变量，可用两个非负的新变量相减来替换。例如，无约束，可令。

例 2.1 将下列线性规划问题转化为标准型

$$\text{Min } z = x_1 - 2x_2 + 3x_3$$

$$s.t. \begin{cases} x_1 + x_2 + 2x_3 \leqslant 5 \\ x_1 - 3x_2 + x_3 \geqslant 3 \\ 2x_1 + 3x_2 - x_3 = 6 \\ x_1, x_2 \geqslant 0 \end{cases}$$

解：（1）令 $x_3 = x_4 - x_5$，其中 $x_4, x_5 \geqslant 0$。

（2）第一个约束条件左边加上非负约束变量 x_6。

（3）第二个约束条件左边减去非负约束变量 x_7。

（4）令 $z' = -z$

如此便得到上问题的标准型：

$$\text{Max } z' = -x_1 + 2x_2 - 3(x_4 - x_5) + 0x_6 + 0x_7$$

$$s.t. \begin{cases} x_1 + x_2 + 2(x_4 - x_5) + x_6 = 5 \\ x_1 - 3x_2 - (x_4 - x_5) - x_7 = 3 \\ 2x_1 + 3x_2 - (x_4 - x_5) = 6 \\ x_1, x_2, x_4, x_5, x_6, x_7 \geqslant 0 \end{cases}$$

第二节　线性规划模型的一些概念

在引例中，已经介绍了决策变量、可行解、不可行解、可行域、最优解的概念。以下介绍其他几个常用的相关概念：基、基变量、非基变量、基可行解、可行基、退化解。

如前所述，线性规划模型的标准型为：

$$\text{Max } z = CX$$

$$s.t. \begin{cases} AX = b \\ X \geqslant 0 \end{cases}$$

一、基、基向量、基变量和非基变量

设 A 为约束方程的 $m \times n$ 系数矩阵，B 是 A 中 $m \times m$ 阶非奇异矩阵（即 $|B| \neq 0$），则称 B 为线性规划问题的一个基。这就是说，矩阵 B 是由 m 个线性独立的列向量组成的，不失一般性，设：

$$B = (P_1, P_2, \cdots, P_m) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}$$

$P_j (j = 1, 2, \cdots, m)$ 称为基向量。与基向量 P_j 相对应的基变量决策变量 x_j 称为基变量，其他决策变量称为非基变量。

二、基解、基可行解、可行基、退化解

标准型下，线性规划的约束条件可以用向量的形式表示为：

$$\begin{cases} \sum_{j=1}^{n} P_j x_j = b \\ x_j \geqslant 0, j = 1, \cdots, n \end{cases}$$

等价于：

$$\begin{cases} \sum_{j=1}^{m} P_j x_j = b - \sum_{j=m+1}^{n} P_j x_j \\ x_j \geqslant 0, j = 1, \cdots, n \end{cases}$$

按上面，设前面的 m 个系数向量 $P_j (j = 1, 2, \cdots, m)$ 为基向量，则对应于 P_j 的决策变量 $X_B = (x_1, x_2, \cdots, x_m)^T$ 为基变量，其余决策变量为非基变量。令非基变量全取零，则上面约束条件方程变为 $\sum_{j=1}^{m} P_j x_j = b$，用高斯消元法，可以求出一个解：

$$X = (x_1, x_2, \cdots, x_m, 0, \cdots, 0)^T$$

我们称此解为一个基解。由此可见，有一个基，就可以求出一个基解。这些基解，对应于可行域的顶点。一个线性规划模型，最多有 C_n^m 个基解。满足非负条件 $x_j \geqslant 0$，$j = 1, \cdots, n$ 的基解，称为基可行解。对应于基可行解的基，称为可行基。在基解中，当非零分量的个数小于 m 个时，这个基解称为退化解。在以下讨论中，我们将假定不出现退化情况。

第三节　线性规划模型的求解

线性规划模型的求解方法很多，最基本和最常用的方法是单纯形法。对于决策变量不超过 3 个的线性规划模型，可以用图解法，对一些特殊的线性规划问题，比如指派问题、运输问题、网络问题等，还发展出了一些专门的解法以提高求解的效率。

基于上面一些解法，人们开发了不少计算机程序，如 WinQSB、Lingo、Lindo、

Excel、Reslove 等,用来求解线性模型。计算机的引入,大大提高了求解的速度,使人们可以处理含有更多变量和更多约束条件的模型,也极大地减轻了决策人员的负担,使得他们可以不用了解具体的算法便能得到模型的求解结果。

一、图解法

对于决策变量小于 3 个的模型,可以用图解法来求解。图解法求解不用将模型转化成标准型。

以下以引例的模型为例,介绍图解法的求解过程。引例模型为:

$$\text{Max } P = 300D + 500W \text{ (目标函数)}$$

$$s.t. \begin{cases} D \leqslant 4 & \text{(资源约束)} \\ 2W \leqslant 12 & \text{(资源约束 2)} \\ 3D + 2W \leqslant 18 & \text{(资源约束 3)} \\ D, W \geqslant 0 & \text{(非负约束)} \end{cases}$$

第一步,在坐标中,先画出各个约束条件线,确定出可行域。

第二步,绘制目标函数线,如图 2.1 所示。

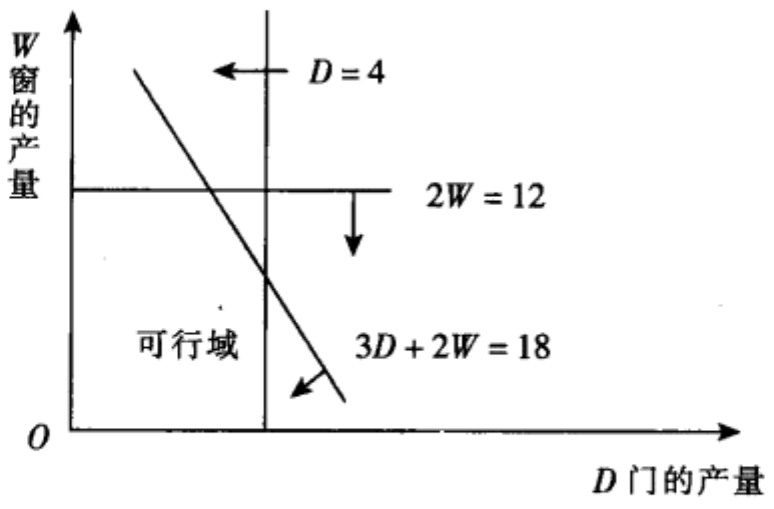


图 2.1

图 2.1 的纵坐标是 W,所以把目标函数 $P = 300D + 500W$ 变为:

$$W = -\frac{300}{500}D + \frac{1}{500}P$$

将目标函数线标到图 2.1,如图 2.2 所示。

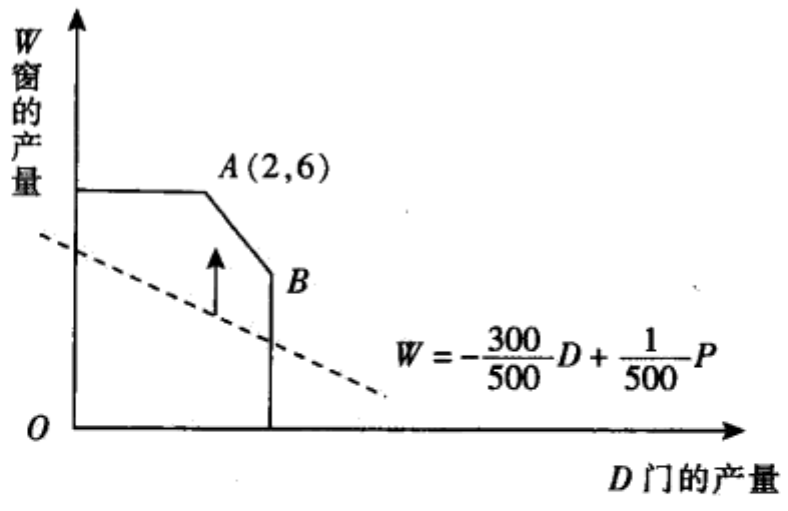


图 2.2

可以看出，目标函数线的纵截距是 $P/500$。当目标函数线向上平行移动时纵截距增大，P 值也增大（总利润提高）。当目标函数线向上平行移动到 A 点时，就不能再向上移动了，否则会移动出可行域。从而知，过 A 点的目标函数线具有最大截距项，也即，具有最大的目标函数值。故点 $A(2,6)$ 是最优解。此时利润为 $P = 300 \times 2 + 500 \times 6 = 3\,600$（美元）。

这样，用图解法得到了一个简单线性规划模型的最优解。下面，借用图解法，简单讨论线性规划模型的解。

1. 唯一最优解 ，

如上例，目标函数线与可行域只有一个交点，这时只有一个解是最优解。

2. 无穷多最优解

若将引例中窗的单位利润调整为 200，目标函数线向上移动到 A 点时，将与直线 AB 重合，这时候，线段 AB 上每一个点都是最优解，具有相等的目标函数值 P，如图 2.3 所示。

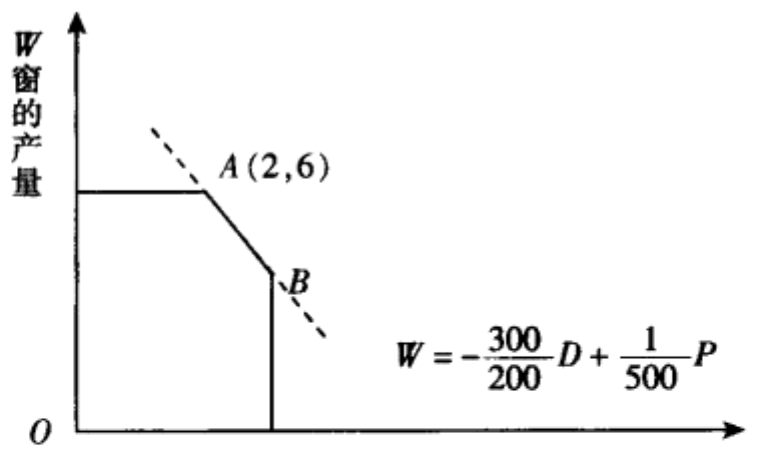


图 2.3

3. 无界解

也不是所有线性规划模型的可行域都是封闭的，如图 2.4 所示。

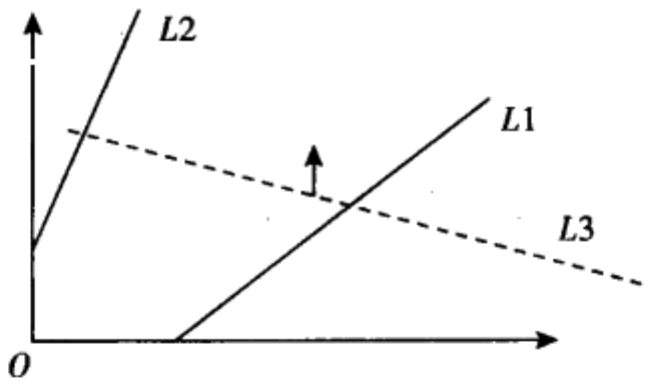


图 2.4

图中 L_1、L_2 为约束条件线，L_3 为目标函数线。

假如目标为取最大值，则目标函数线应向上移动。目标函数线可以向上一直移动下去，目标函数值将增大到无穷大。这种情形称为无界解或者无最优解。出现这种情形时，多数是因为模型缺少必要的约束条件，此时需对模型重新进行设定。

4. 无可行解

有时候，可行域为空集，这时就没有可行解，也就没有最优解。出现这种情形，有

时候是因为模型中出现了相互矛盾的约束条件,此时需重新审视模型的设定。

当线性规划模型的可行域非空时,可行域是一个有界或者无界的凸多边形;若线性规划存在最优解,最优解一定在可行域的某个顶点得到,就如引例中的 A 点;若在两个顶点同时得到最优解,则此两点的连线段一定为可行域的一段边界(如图 2.4 所示),其上的任意一点,都是最优解,此时有无穷多的最优解。

二、单纯形法

为了便于直观地理解单纯形法的基本原理和掌握其计算方法,下面用一个例子来演示其演算过程,并做适当的经济解释。

例 2.2 设有一个家具厂用木材和钢材生产 A、B、C 三种家具。生产一件家具所需要的材料数量、每件家具可以获得的利润以及每月可提供的材料数量如表 2.2 所示。

问:如何安排各种家具每月的生产量,该厂才能在该月获得最大利润?

表 2.2

产品	A 产品	B 产品	C 产品	材料可供量
木材	3	1	4	8 000
钢材	2	1	4	3 000
单位产品利润	4	1	5	

解 先建立线性规划模型。设 x_1、x_2、x_3 分别代表 A、B、C 三种家具的产量(件)。

$$\text{Max } z = 4x_1 + x_2 + 5x_3 \tag{2.1}$$

$$s.t. \begin{cases} 3x_1 + x_2 + 4x_3 \leqslant 8\,000 \\ 2x_1 + x_2 + 4x_3 \leqslant 3\,000 \\ x_j \geqslant 0 \quad j = 1,2,3 \end{cases} \tag{2.2}$$

这是一般形式的线性规划模型。为使用单纯形法求解,需要将其变成标准型。添加松弛变量 x_4、x_5,使约束条件变成等式,得其标准型:

$$\text{Max } z = 4x_1 + x_2 + 5x_3 + 0x_4 + 0x_5 \tag{2.3}$$

$$s.t. \begin{cases} 3x_1 + x_2 + 4x_3 + x_4 = 8\,000 \\ 2x_1 + x_2 + 4x_3 + x_5 = 3\,000 \\ x_j \geqslant 0 \quad j = 1,2,3,4,5 \end{cases} \tag{2.4}$$

约束条件的系数矩阵为:

$$A = (P_1, P_2, P_3, P_4, P_5) = \begin{bmatrix} 3 & 1 & 4 & 1 & 0 \\ 2 & 1 & 4 & 0 & 1 \end{bmatrix}$$

其中 $P_4 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$,$P_5 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 是线性无关的。取这两个向量作为一个基矩阵。

令 $B = (P_4, P_5) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

对应于 B 的变量为基变量,它们分别是:

$$x_4 = 8\,000 - (3x_1 + x_2 + 4x_3) \tag{2.5a}$$

$$x_5 = 3\,000 - (2x_1 + x_2 + 4x_3) \tag{2.5b}$$

将上述两个基变量代入(2.3)式,得:

$$z = 4x_1 + x_2 + 5x_3 + 0 \tag{2.6}$$

当工厂未投入生产时,即 $x_1 = x_2 = x_3 = 0$,木材和钢材都没有被利用,取 $x_4 = 8\,000$,$x_5 = 3\,000$,工厂不生产利润,$z = 0$,此时得到一个基本可行解

$$X^{(0)} = (0,0,0,8\,000,3\,000)^T$$

现在分析目标函数式(2.6)。式中 x_1、x_2、x_3 的系数都是正值,它们都是非基变量,如将非基变量变为基变量,安排三种产品投产,就可以增加工厂的利润。在三种产品中,若选择一种产品先投产,则可以考虑获利最大的产品,本例中 x_3 的系数最大,所以可以考虑先安排生产产品 C。

而由资源约束式(2.2)知,x_3 的取值应该满足:

$$x_4 = 8\,000 - (3x_1 + x_2 + 4x_3) \geqslant 0$$

$$x_5 = 3\,000 - (2x_1 + x_2 + 4x_3) \geqslant 0$$

因 $x_1 = x_2 = 0$,所以有:

$$x_4 = 8\,000 - 4x_3 \geqslant 0$$

$$x_5 = 3\,000 - 4x_3 \geqslant 0$$

需要同时满足上两式,且值越大越好,所以应该有:

$$x_3 = \text{Min}\{8\,000/4, 3\,000/4\} = 750$$

上式表明,生产750件 C 产品,正好将3 000单位的钢材用完,即 $x_5 = 0$。由于生产750件 C 产品,只消耗了的木材 $750 \times 4 = 3\,000$ 单位,还有5 000单位的木材没有得到利用。此时得到一个新的基本可行解:

$X^{(1)} = (0,0,750,5\,000,0)^T$。

从经济意义上看,用3 000单位的木材和3 000单位的钢材生产了750件 C 产品,其他产品没有生产。

比较前后两个基本可行解:x_3 从非基变量变成了基变量,而 x_5 由基变量变成了非基变量,此时称 x_3 置换了 x_5。这个基变换过程可以表示如下:

由式(2.5a)和(2.5b)得:

$$x_4 + 4x_3 = 8\,000 - 3x_1 - x_2 \tag{2.7}$$

$$4x_3 = 3\,000 - 2x_1 - x_2 - x_5 \tag{2.8}$$

用消元法将式(2.7)和(2.8)中的列向量变成单位向量。具体步骤为(2.8) - (2.7),(2.8) ÷ 4,结果为:

$$x_4 = 5\,000 - x_1 - 0x_2 + x_2 \tag{2.9a}$$

$$x_3 = 750 - 1/2x - 1/4x_2 - 1/4x_5 \tag{2.9b}$$

将式(2.9a)和(2.9b)代入式(2.3)得:

$$z = 3\,750 + 3/2x_1 - 1/4x_2 - 5/4x_5 \tag{2.10}$$

令非基变量 $x_1 = x_2 = x_3 = 0$,则有 $z = 3\,750$,并得到一个新的 B:

$$B = (P_4, P_3) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

新的约束方程的系数矩阵为:

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 & -1 \\ 1/2 & 1/4 & 4 & 0 & 1/4 \end{bmatrix}$$

如此就完成了这次置换。

此时目标函数式(2.10)中还有为正值的系数,若安排生产产品A,则还可以增加利润。所以继续进行基变换。与上面过程相同,由(2.9a)和(2.9b)可得:

$$x_4 + x_1 = 5\,000 - 0x_2 + x_5$$

$$1/2x_1 = 750 - 1/4x_2 - x_3 - 1/4x_5$$

即:

$$x_4 = 3\,500 + 1/2x_2 + 2x_3 + 3/2x_5 \tag{2.11a}$$

$$x_1 = 1\,500 - 1/2x_2 - 2x_3 - 1/2x_5 \tag{2.11b}$$

将(2.11a)和(2.11b)代入(2.3)得:

$$z = 6\,000 - x_2 - 3x_3 - 2x_5 \tag{2.12}$$

从式(2.12)可以看出,非基变量x_2、x_3、x_5的系数都是负数,这些变量值的增加,将导致目标利润的减少,故不能再考虑生产B和C产品。故目前的最优方案是:$x_1 = 1\,500$,$x_4 = 3\,500$,$x_2 = x_3 = x_5 = 0$。目标达到最大值$z = 6\,000$元。

最优解为 $X^{(2)} = (1\,500,0,0,3\,500,0)^T$

新的约束条件的系数矩阵为:

$$A = \begin{bmatrix} 0 & -1/2 & -2 & 1 & -3/2 \\ 1/ & 1/2 & 2 & 0 & 1/2 \end{bmatrix}$$

通过上例的求解,可以比较清楚地了解单纯形法的算法思想和计算步骤。但是这种代数运算缺乏严格的计算程序,容易造成差错。细心的读者还会发现,上例中共进行了两次基变换。这两次基变换都只是先分析了将某个非基变量变成基变量,而没有事先分析要哪一个基变量置换出来,使之变成非基变量。

运用下面介绍的特制的单纯形表格可以克服上述缺点,能方便、简捷、准确地计算线性规划的最优解。

三、单纯形法的计算步骤

为便于理解单纯形法的计算关系,研究者设计了一种计算表,我们称之为单纯形表。以下使用单纯形表来给出单纯形法的计算步骤。

第一步,首先将非标准的线性规划模型转化成标准型。设模型的标准型为:

$$\text{Max } z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

$$s.t. \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + c_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots \\ a_{m1} + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \\ x_1, x_2, \cdots, x_n \geqslant 0 \end{cases}$$

例如,上例2.2的标准型为:

$$\text{Max } z = 4x_1 + x_2 + 5x_3 + 0x_4 + 0x_5$$

$$s.t.\begin{cases}3x_1 + x_2 + 4x_3 + x_4 = 8\,000\\ 2x_1 + x_2 + 4x_3 + x_5 = 3\,000\\ x_j \geqslant 0 \quad j = 1,2,3,4,5\end{cases}$$

第二步,确定初始基本可行解,填制初始单纯形表。

根据基本可行解的定义,它必须满足两个条件:一是不能违反任何约束条件,且在不退化的情况下,非零变量的个数等于约束条件方程个数;二是解中基变量的系数列向量不相关。

为表述的方便,不失一般性,假设初始基变量刚好是前 m 个决策变量,则初始基为:

$$B = (P_1, P_2, \cdots, P_m) = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & 1\end{bmatrix}$$

例如上例 2.2 中的初始基矩阵为:

$$B = (P_4, P_5) = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

可见,不是任何解都可以作为初始可行解的。上例2.2 中,取 $x_1 = x_2 = x_3 = 0, x_4 = 8\,000, x_5 = 3\,000$ 作为基本可行解,可以满足上述要求。

确定了初始基本可行解之后,将有关的数据和符号按一定格式编入一个表格中,编制出初始单纯形表。如表 2.3 所示。

表 2.3

$c_j \longrightarrow$			c_1	$\cdots$	c_m	c_{m+1}		c_n	θ_i
C_B	X_B	b	x_1	$\cdots$	x_m	x_{m+1}		x_n	
c_1	x_1	b_1	1	$\cdots$	0	$a_{1,m+1}$	$\cdots$	a_{1n}	θ_1
c_2	x_2	b_2	0	$\cdots$	0	$a_{2,m+1}$	$\cdots$	a_{2n}	θ_2
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots\cdots$	$\cdots$	$\cdots$	$\cdots$
c_m	x_m	b_m	0	$\cdots$	1	$a_{m,m+1}$	$\cdots$	a_{mn}	θ_m
Z_j			0	$\cdots$	0	$\sum_{i=1}^{m} c_i a_{i,m+1}$	$\cdots$	$\sum_{i=1}^{m} c_i a_{in}$	
$c_j - Z_j$			c_1	$\cdots$	c_m	$c_{m+1} - \sum_{i=1}^{m} c_i a_{i,m+1}$		$c_n = \sum_{i=1}^{m} c_i a_{in}$	

第一行是目标函数的系数 $c_j (j = 1, \cdots, n)$;第二行依次是基变量的系数矩阵 C_B,基变量矩阵 X_B,约束条件右端值矩阵 b,接着是各个决策变量 $x_i, i = 1, \cdots, n$;第三行前三列分别是矩阵 C_B、X_B、b 的展开,接下来的 n 列是约束条件方程系数矩阵。第四行中有两个内容:一是 Z_j,称为机会成本,计算公式为 $Z_j = \sum_{i=1}^{m} c_i a_{ij}$;二是 $c_j - Z_j$,称为检验数,这是将基变量的函数式代入目标函数式后,非基变量在目标函数中的系数,它

表示非基变量转换成基变量后对目标函数的贡献(增加或者减少)。接下来 n 列是 Z_j 和 $c_j - Z_j$ 的展开,即按照 Z_j 和 $c_j - Z_j$ 的计算公式,对各个变量计算的具体 Z_j 和 $c_j - Z_j$ 值。此表的最后一列是 $\theta_i(i = 1,2,\cdots,m)$,用于确定基变换的换出变量。

对于例 2.2,确定基本可行解之后,把数据和符号按照上面的格式填入表格如表 2.4 所示。

表 2.4

$c_j \longrightarrow$			4	1	5	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
0	x_4	8 000	3	1	4	1	0	2 000
0	x_5	3 000	2	1	4	0	1	750
		Z_j	0	0	0	0	0	
		$c_j - Z_j$	4	1	5	0	0	

由例 2.2 知,从初始基可行解出发,要经过多次运算才能得到最优解。

第三步,最优性检验与解的判别。

(1) 计算各非基变量的检验数:$\sigma_j = c_j - Z_j = c_j - \sum_{i=1}^{m} c_i a_{ij}, j = m+1,\cdots,n$。若每个非基变量的检验数都是负数,则已经得到唯一最优解;若有非基变量的检验数为零且其余非基变量检验数均为负数,则问题有无穷多最优解且已经得到其中一个最优解。在此两种情形下均可停止计算,否则转入(2)。

(2) 在 $\sigma_j > 0, j = m+1,\cdots,n$ 中,若有某个 σ_k 对应 x_k 的系数列向量 $P_k \leqslant 0$,即 P_k 中每个元素都小于或等于 0,则此问题有无界解,此时应停止计算;否则,目标函数还可以继续优化,应继续迭代,转入第四步。

第四步,根据 $\text{Max}(\sigma_j > 0) = x_k$ 确定 x_k 为换入变量,按照 θ 规则计算:

$$\theta = \text{Min}\left(\frac{b_i}{a_{ik}} \mid a_{ik} > 0\right) = \frac{b_l}{a_{lk}}$$

可确定 x_1 为换出变量。

对例 2.2,从表 2.3 中可以看到,每一个正值的检验数所对应的非基变量的系数都是正数,由此可以判断例 2.2 不是无界的;又,三个正检验数中,x_3 的检验数最大,所以选择 x_3 为这次基变换的换入变量;再由 b_1、b_2 除以 x_3 的系数列向量 $\begin{pmatrix}4\\4\end{pmatrix}$ 中相应的正元素,得到 $\theta_4 = 8\,000/4 = 2\,000$,$\theta_5 = 3\,000/4 = 750$,此两者中最小的是 θ_5,所以选择 x_5 为换出变量。

第五步,以 a_{lk} 为主元素进行迭代(高斯消元法或旋转运算),把 x_k 所对应的列向量变成单位列向量。即:

$$P_k = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \cdots \\ a_{lk} \\ \cdots \\ a_{mk} \end{pmatrix} \xrightarrow{\text{变换成}} \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \\ \cdots \\ 0 \end{pmatrix}$$（第 l 行元素为 1，其余各行元素为 0）。

将 X_B 中的 x_i 换成 x_k，得到新的单纯形表。返回第三步。

对例 2.2，从表 2.3 中可以看到应该换入 x_3，换出 x_5。以下是高斯消元法运算的具体步骤：

第一步，x_5 所在的那一行各列元素，从 b_2 开始，同除以 4，得到单纯形表（表 2.5）。

表 2.5

$c_j \longrightarrow$			4	1	5	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
0	x_4	8 000	3	1	4	1	0	
0	x_5	750	1/2	1/1	1	0	1/4	
		Z_j						
		$c_j - Z_j$						

第二步，将变换后的 x_5 所在的那一行各列元素乘以 4，减去 x_4 所在的那一行各列元素，得到单纯形表（表 2.6）。

表 2.6

$c_j \longrightarrow$			4	1	5	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
0	x_4	5 000	1	0	0	1	-1	
0	x_5	750	1/2	1/4	1	0	1/4	
		Z_j						
		$c_j - Z_j$						

此时，x_3 所在的列已经变成单位列向量了。当然，还有其他办法使该列变成单位列向量，并非非用此法不可。

第三步，将 x_3 换入，x_5 换出，并重新计算各列的 z_j 和 $c_j - Z_j$，调整 C_B，得到此次迭代的最终结果（表 2.7）。

表 2.7

$c_j \longrightarrow$			4	1	5	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
0	x_4	5 000	1	0	0	1	-1	5 000
5	x_3	750	1/2	1/4	1	0	1/4	1 500
		Z_j	5/2	5/4	5	0	5/4	
		$c_j - Z_j$	3/2	-1/4	0	0	-5/4	

表2－7中，还有x_1的检验数为正数，且其对应列向量还有正的元素，故目标函数值还可以继续优化。重复上面第一步到第三步可知，x_1为换入变量，x_3为换出变量，将x_1所在列变成单位列向量，以此类推，可得到最终的单纯形表（表2.8）。

表2.8

$c_j \longrightarrow$			4	1	5	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
0	x_4	3 500	0	－1/2	－2	1	－3/2	
5	x_1	1 500	1	1/2	2	0	1/2	
		Z_j	4	2	8	0	2	
		$c_j - Z_j$	0	－1	－3	0	－2	

表2.8中各个检验数都小于或等于0，表示目标函数不可能再增大了，于是得到问题的最优解：$X = (1\ 500, 0, 0, 3\ 500, 0)^T$。

四、单纯形法的进一步讨论

1. 单纯形法的经济解释——Z_j和$c_j - Z_j$的经济含义

在单纯形表中，Z_j的实质是机会成本。机会成本是指在决策分析过程中，从各个备选方案中选择最优方案而放弃次优方案所丧失的潜在收益。它虽然不是实际支出，但是它必须从被选用的方案中得到补偿，这样才能对最优方案的经济效益作出全面、合理的评价。

从表2.7中可知，不生产A产品（$x_1 = 0$），只生产750件C产品。生产750件C产品消耗了3 000单位的木材和3 000（750×4）单位的钢材，木材还剩下5 000单位，而钢材已经全部用完。若现在要生产一件A产品，则要牺牲C产品的生产，一件A产品需要2个单位的钢材和3个单位的木材，所以生产一件A产品，需要少生产1/2件C产品以得到所需的钢材。木材仍有剩余。所以生产一件A产品要牺牲的利润为$5 \times 1/2 + 0 \times 1 = 5/2$。

这种为生产一个单位某种产品，将所需的资源从原有的用途中抽出来，而牺牲的利润称为此产品在此生产方案下的机会成本，其计算公式为：

$$Z_j = \sum_{i=1}^{m} c_i a_{ij}$$

由于生产一件A产品可以获利4元，而机会成本只有5/2元，利润增加量为$c_j - Z_j = 4 - 5/2 = 3/2$元，说明改变表2.7中的方案是有利的，用$x_1$换出$x_3$可以增加利润。

2. 人工变量法

为了得到一个初始基可行解，一个可行的办法是在模型中加入人工变量。

设原有的线性规划的模型为：

$$\text{Max } z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

$$s.t. \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

在每个约束方程中加入一个人工变量，得：

$$s.t.\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + x_{n+1} = b_1 \\ a_{21}x_{22}x_2 + \cdots + a_{2n}x_n + x_{n+2} = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + x_{n+m} = b_m \\ x_1, x_2, \cdots, x_n \geqslant 0, x_{n+1}, \cdots, x_{n+m} \geqslant 0 \end{cases}$$

以 $x_{n+1}, \cdots, x_{n+m}$ 为基变量，即可得到一个 $m \times m$ 的单位阵，令非基变量 $x_1, x_2, \cdots, x_n$ 为零，便可得到一个初始基可行解：$X^{(0)} = (0, \cdots, 0, b_1, \cdots, b_m)^T$

因为人工变量是后加入到原约束条件中的虚拟变量，要求将它们从基变量中逐一替换出来。若经过基变换，基变量中不再含有非零的人工变量，这表示原问题有解；若在最终表中当所有的检验数 $c_j - Z_j \leqslant 0$，基变量中仍然有某个非零的人工变量时，这表示无可行解。

(1) 大 M 法

在一个线性规划问题的约束条件中加入人工变量后，要求人工变量的目标函数取值不受影响，为此假定人工变量在目标函数中的系数为 $-M$，M 为任意大的正数。这样，目标函数要实现最大化，则必须把人工变量从基变量中换出，否则目标函数不可能实现最大化。

(2) 两阶段法

用电子计算机求解含有人工变量的线性规划模型时，只能用很大的数代替 M，这样可能造成计算上的错误，这时可以用两阶段法求解。

第一阶段，不考虑原问题是否存在基可行解，给原线性问题加入人工变量后，构造仅含有人工变量的目标函数和要求实现最小化。

$$\text{Min } \omega = 0x_1 + 0x_2 + \cdots + 0x_n + x_{n+1} + \cdots + x_{n+m}$$

$$s.t.\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + x_{n+1} = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{1n}x_n + x_{n+1} = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + x_{nm} = b_m \\ x_1, x_2, \cdots, x_n \geqslant 0, x_{n+1}, \cdots, x_{n+m} \geqslant 0 \end{cases}$$

用单纯形法求解，若得到 $\omega = 0$，则说明原问题存在基可行解，可以进行第二阶段的计算。否则原问题无可行解，应停止计算。

第二阶段，将第一阶段计算得到的最终表，删去人工变量和相应的数据，把目标函数行的系数 $c_j (j = 1, \cdots, n)$ 换成原问题的目标函数系数，把 C_B 中的元素也作相应变换，重新计算 Z_j 和 $c_j - Z_j$。其他所有变量和数据不变。这样处理后的表作为第二阶段计算的初始表。接着用单纯形法就可以迭代出最优解。

3. 退化

单纯形法计算中，用规则确定换出变量时，有时存在两个以上相同的最小比值，这样在下一次迭代中就有一个或者几个基变量等于零，且使迭代后目标函数值不变，产生退化现象。在有些情况下，退化现象有可能出现迭代无限循环从而达不到最优

解。

实践表明,退化循环现象很少发生,但也有人构造了特例,证明了循环现象确实存在。1974 年,勃兰特(Bland) 提出了一个简便方法,称为勃兰特规则,可以解决此问题。

勃兰特规则:

(1) 从大于零的检验数 σ_j 中选取下标最小的非基变量 x_k 作为换入变量。

(2) 当按 θ 规则计算存在两个或两个以上相同的最小比值时,选取下标最小的基变量为换入变量。

按此规则,一般能够避免循环现象。

五、单纯形法的矩阵描述

设线性规划模型的矩阵形式为:

$$\text{Max } z = CX$$

$$s.t. \begin{cases} AX \leqslant b \\ X \geqslant 0 \end{cases}$$

在此模型约束方程左端加入松弛变量 $X_s = (X_{n+1}, X_{n+2}, \cdots, X_{n+m})^T$,得到:

$$\text{Max } z = CX + 0X_s$$

$$s.t. \begin{cases} AX + IX_s = b \\ X \geqslant 0, X_s \geqslant 0 \end{cases}$$

这里 I 为 $m \times m$ 阶单位矩阵。

将系数矩阵(A,I),按变量的性质重新划分为两块:(B,N)。其中,基变量X_B的系数矩阵为B,非基变量X_N的系数矩阵为N。对应地把目标函数系数矩阵$(C,0)$也重新分为两块(C_B,C_N)。于是,模型标准型可以表示为:

$$\text{Max } z = C_B X_B + C_N X_N$$

$$s.t. \begin{cases} BX_B + NX_N = b \\ X_B \geqslant 0, X_N \geqslant 0 \end{cases}$$

将约束条件方程移项得:

$$BX_B = b - NX_N$$

B 是基,为非奇异的,存在逆矩阵 B^{-1}。上式两端都左乘 B^{-1},得:

$$X_B = B^{-1}b - B^{-1}NX_N \tag{2.13}$$

将式(2.13) 代入目标函数有:

$$\begin{aligned} Z &= C_B X_B + C_N X_N \\ &= C_B B^{-1} - C_B B^{-1} N X_N + C_N X_N \\ &= C_B B^{-1} b + (C_N - C_B B^{-1} N) X_N \end{aligned} \tag{2.14}$$

令 $X_N = 0$,就可以得到一个基本可行解 $X^{(1)} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix}$,目标函数值为 $Z = C_B B^{-1} b$。

现在讨论单纯形表的矩阵表达式。

1. 检验数

非基变量的检验数矩阵为:$C_N - C_B B^{-1} N$。

基变量的检验数为零,可将其检验数写成 $C_B - C_B B^{-1}$ 的形式。因为 X_s 对应的目标函数系数为 0,所以 X_s 对应的检验数矩阵为 $0 - C_B B^{-1} I$,即 $- C_B B^{-1}$。

所以,所有的检验数矩阵都可以用 $C - C_B B^{-1} A$ 与 $- C_B B^{-1}$ 表示。

2. θ 规则的矩阵表示

$$\theta = \text{Min}\left\{\frac{(B^{-1}b)_i}{(B^{-1}P_j)_i} \mid (B^{-1}P_j)_i > 0\right\} = \frac{(B^{-1}b)_i}{(B^{-1}P_j)_i}$$

其中 $(B^{-1}b)_i$ 是向量 $(B^{-1}b)$ 中的第 i 个元素。$(B^{-1}P_j)_i$ 是向量 $(B^{-1}P_j)$ 中的第 i 个元素。

3. 单纯形表

将非基变量系数矩阵 N 分成 (N_1, I),标准型的目标函数和约束条件可改写为:

$$\begin{cases} -z + C_B X_B + C_{N1} X_{N_1} + 0X_s = 0 \\ BX_B + N_1 X_{N_1} + LX_s = b \end{cases}$$

当确定 X_B 为基变量时,经过基变换,得到以下两式:

$$\begin{cases} X_B + B^{-1} N_1 X_{N_1} + B^{-1} X_s = B^{-1} b \\ -z + (C_{N_1} - C_B B^{-1} N_1) X_{N_1} - C_B B^{-1} X_s = - C_B B^{-1} b \end{cases}$$

用矩阵关系式表示上两式为:

$$\begin{pmatrix} 0 & I & B^{-1}N_1 & B^{-1} \\ I & 0 & C_{N_1} - C_B B^{-1} N_1 & - C_B B^{-1} \end{pmatrix} \begin{pmatrix} -z \\ X_B \\ X_N \\ X_s \end{pmatrix} = \begin{pmatrix} B^{-1}b \\ - C_B B^{-1} b \end{pmatrix}$$

这些分块的系数矩阵用表格表示为表 2.9。

表 2.9

基变量 X_B	非基变量		RHS
	X_{N_1}	X_s	
I	$B^{-1}N_1$	B^{-1}	$B^{-1}b$
0	$C_N - C_B B^{-1} N_1$	$- C_B B^{-1}$	$- C_B B^{-1} b$

单纯形表的分块系数矩阵中,$(0, I)^T$ 不参加运算,表中不填这些数据。表 2.9 即为迭代后的计算表,各部分数据都可以用矩阵运算求得。在表 2.9 中还可以看到初始单位阵的位置,在各运算表中就是 B^{-1} 所在的位置。

第四节　线性规划的 WinQSB 应用

本节结合例题介绍 WinQSB 软件求解线性规划的操作步骤及应用。

例 2.3　用 WinQSB 软件求解下列线性规划问题:

$$\text{Max } Z = 6x_1 + 5x_2 + x_3 + 7x_4$$

$$s.t. \begin{cases} x_1 + 2x_2 + 6x_3 + 9x_4 \leqslant 260 \\ 8x_1 - 5x_2 + 2x_3 - x_4 \geqslant 150 \\ 7x_1 + x_2 + x_3 = 30 \\ x_1 - x_2 \geqslant 0 \\ x_3 - x_4 \geqslant 0 \\ 10 \leqslant x_3 \leqslant 20 \\ x_1, x_2, x_3 \geqslant 0 \end{cases}$$

解:应用 WinQSB 软件求解线性规划问题不必转化为标准型。如果是可以线性化的模型则先线性化,则对于有界变量及无约束变量可以不用转化,只需要修改系统的变量类型即可;对于不等式约束,可以在输入数据时直接输入不等式符号。

(1) 启动线性规划(LP) 和整数规划(ILP) 程序。点击“开始” → “程序” → “WinQSB” → “Linear and Integer Programming”,显示线性规划和整数规划工作界面。

(2) 建立新问题或者打开磁盘中已有的文件。点击“File” → “New Problem” 建立一个新问题。输入本问题的文件名称“lp1”(读者可以任意取名),决策变量个数为4,约束条件个数为5。由于本问题是一个最大化问题,所以选择“Maximization”,同时可以确定数据的输入形式,一种为表单形式,另一种为模型形式。如果我们选择了表单形式,则如图 2.5 所示。

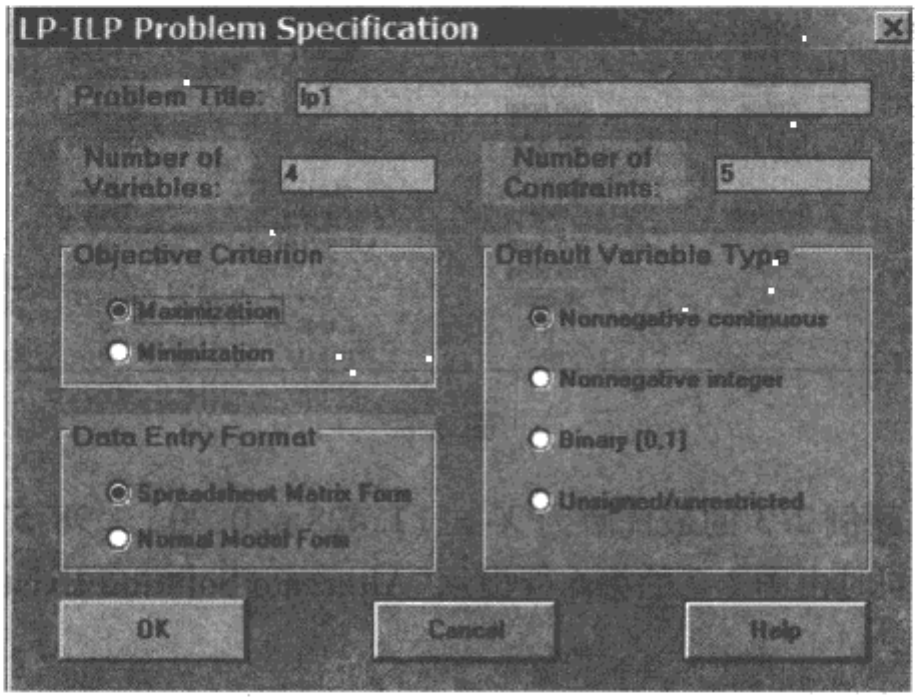


图 2.5

(3) 输入数据。按照例 2.3 输入变量系数和右端常数项数据。

(4) 修改变量类型。图 2.5 中给出了非负连续、非负整数、0 - 1 型和无符号限制或者无约束 4 种变量类型选项,当选择了某一种类型后,系统默认所有变量都属于该种类型。在例 2.3 中,$10 \leqslant x_3 \leqslant 20$,直接将 x_3 中的下界(Lower Bound) 改为 10,上界(Upper Bound) 改为20。把 x_4 设定为无约束(Unrestricted),M 是一个任意大的正数。得到如图 2.6 所示的表格。

Variable ->	X1	X2	X3	X4	Direction	R. H. S.
Maximize	6	5	1	7		
C1	1	2	6	9	<=	260
C2	8	-5	2	-1	>=	150
C3	7	1	1		=	30
C4	1	-1			>=	0
C5			1	-1	>=	0
LowerBound	0	0	10	-M		
UpperBound	M	M	20	M		
VariableType	Continuous	Continuous	Continuous	Unrestricted		

图 2.6

(5) 修改变量名和约束名。系统默认变量名为 $X_1, X_2, \cdots, X_n$，约束名为 $C_1, C_2, \cdots, C_m$。默认名可以修改，点击菜单栏“Edit” 后，下拉菜单有四个修改选项：修改标题名(Problem Name)、变量名(Variable Name)、约束名(Constraint Name) 和目标函数准则(Max 或 Min)。由于 WinQSB 软件支持中文，读者可以输入中文名称。

(6) 求解。点击菜单栏“Solve and Analyze”，下拉菜单有三个选项：求解不显示迭代过程(Solve the Problem)、求解并显示单纯形法迭代步骤(Solve and Display Steps) 及图解法(Graphic Method，限两个决策变量)。如选择“Solve the Problem”，系统直接显示求解的综合报告如图2.7所示，表中的各项含义如图2.11所示。线性规划问题有最优解或无最优解(无可行解或无界解)，系统会给出提示。

	Decision Variable	Solution Value	Unit Cost or Profit c(j)	Total Contribution	Reduced Cost	Basis Status	Allowable Min. c(j)	Allowable Max. c(j)
1	X1	1.4286	6.0000	8.5714	0	basic	-266.0000	49.0000
2	X2	0	5.0000	0	-38.8571	at bound	-M	43.8571
3	X3	20.0000	1.0000	20.0000	0	basic	-5.1429	M
4	X4	-98.5714	7.0000	-690.0000	0	at bound	-M	7.0000
	Objective	Function	(Max.) =	-661.4285	(Note:	Alternate	Solution	Exists!!)
	Constraint	Left Hand Side	Direction	Right Hand Side	Slack or Surplus	Shadow Price	Allowable Min. RHS	Allowable Max. RHS
1	C1	-765.7142	<=	260.0000	1,025.7140	0	-765.7142	M
2	C2	150.0000	>=	150.0000	0	-7.0000	51.4286	M
3	C3	30.0000	=	30.0000	0	8.8571	20.0000	116.2500
4	C4	1.4286	>=	0	1.4286	0	-M	1.4286
5	C5	118.5714	>=	0	118.5714	0	-M	118.5714

图 2.7

由图 2.7 得到例 2.3 的最优解为 $X = (1.428\,6, 0, 20, -98.571\,4)^T$，最优值 $Z = -661.428\,5$。同时由图2.7的第6行提示“Alternate Solution Exists!!” 知原线性规划问题有多重解。

(7) 显示结果分析。点击菜单栏“Result” 或者点击快捷方式图标，存在最优解时，下拉菜单有9个选项，如下① ~ ⑨，无最优解时有两个选项，如下⑩ ~ ⑪：① 只显示最优解(Solution Summary)。② 约束条件摘要(Constraint Summary)，比较约束条件两端的值。③ 对目标函数进行灵敏度分析(Sensitivity Analysis of OBJ)。④ 对约束条件右端常数进行灵敏度分析(Sensitivity Analysis of RHS)。⑤ 求解结果组合报告(Combined Report)，显示详细综合分析报告。⑥ 进行参数分析(Perform Parametric Analysis)，某个目标函数系数或约束条件右端常数带有参数，计算出参数的变化区间及其对应的最优解，属于参数规划内容。⑦ 显示最后一张单纯形表(Final Simplex Tableau)。⑧ 显示另一个基本最优解(Obtain Alternate Optimal)，存在多重解时，系统显示另一个基本最优

解,然后考虑对基本最优解进行组合可以得到最优解的通解。⑨ 显示系统运算时间和迭代次数(Show Run Time and Iteration)。⑩ 不可行性分析(Infeasibility Analysis),线性规划问题无可行解时,系统指出存在无可行解的原因,如将例 2.3 的第五个约束改为 $x_3 - x_4 \leqslant 0$,系统显示无可行解并且给出图 2.8 所示的显示报告。这说明第五个约束不可能小于或等于零,右端常数至少等于 117.142 9 才可行。⑪ 无界性分析(Unboundedness Analysis),线性规划问题存在无界解时,系统指出存在无界解的可能原因。如将目标函数系数 $c_4 = 7$ 改为 $c_4 = -7$,系统显示无界(如图 2.9 所示)。系统提示要使线性规划问题有解,应该改变第二个约束条件。⑫ 保存结果。求解后将结果显示在顶层窗口,点击"File" → "Save As",系统以文本格式存储计算结果。⑬ 将计算表格转换成 Excel 表格。在计算结果界面中点击"File" → "Copy to Clipboard",系统将计算结果复制到剪贴板,再粘贴到 Excel 表格中即可。

Infeasible	solution!!!	Make any of	the following	RHS changes	and solve the	problem again.
03-06-2006 19:52:37	Constraint	Direction	Right Hand Side	Shadow Price	Add More Than This To RHS	Add Up To This To RHS
1	C1	<=	260.0000	0	-107.1429	M
2	C2	>=	150.0000	0	-M	-117.1429
3	C3	=	30.0000	0.8571	102.5000	750.0000
4	C4	>=	0	0	-2.8571	M
5	C5	<=	0	7.0000	-117.1429	-117.1429

图 2.8

Unbounded	solution!!!	Make any of	the following	changes and	solve it again.
03-06-2006 19:55:07	Constraint	Decision Variable	Coefficient A(i,j)	Subtract More Than This From A(i,j)	Or Add More Than This To A(i,j)
	Change	the direction	of constraint	C2	

图 2.9

(8) 单纯形表。选择求解并显示单纯形法迭代步骤,系统显示初始单纯形表,如图 2.10 所示可以发现,系统将 $X4$ 无约束改写成 $X4 - Neg_X4$,即两个非负变量之差;系统将 $10 \leqslant x_3 \leqslant 20$ 改写成约束 $C_6: 0 \leqslant x_3 - 10 \leqslant 10$,令 $x'_3 = x_3 - 10$,则有 $x'_3 \leqslant 10$,将 $x_3 = x'_3 + 10$ 代入约束条件并整理,在图 2.10 中的 x_3 实际上是 x'_3,如约束 $C1$:

$$X1 + 2X2 + 6(X3 + 10) + 9X4 - Neg_X4 + Slack_C1 = 260$$

整理后得到图 2.7 第一行($Slack_C1$)。

约束 $C1, C4, C5, C6$ 加入 4 个松弛变量 $Slack_C1, Slack_C4, Slack_C5$ 以及 $Slack_UB_X3$,约束 $C2$ 减去剩余变量 $Surplus_C2$,然后 $C2$ 与 $C3$ 加入 2 个人工变量 $Artificial_C2$ 和 $Artificial_C3$,共 6 个约束 12 个变量。

图 2.10 最后两行为检验数,如 $X1$ 的检验数 $C(1) - Z(1) * Big\ M = 6 - 15M$。选 $X1$ 进基,图 2.10 最后一列为比值,变量 $Artificial_C3$ 出基,主元素 $A(3,1) = 7$。

下一步点击菜单栏"Simplex Iteration" 选择"Next Iteration" 继续迭代,还可以人工选择进基变量,或直接显示最终单纯形表。

(9) 模型形式转换。点击菜单栏"Format" → "Switch to Normal Model Form",将图 2.6 电子表格转换成表 2.10 的模型形式,再点击一次转换成图 2.6 的电子表格。

(10) 写出对偶模型。点击菜单栏“Format” → “Switch to Dual Form”，系统自动给出线性规划的对偶模型，再点击一次给出原问题模型。

Simplex Tableau — Iteration 1

Basic	Cj	X1	X2	X3	X4	Surplus C1	Artificial C1	Surplus C2	Artificial C2	Surplus C3	Artificial C3	RHS	Ratio
		6.0000	5.0000	1.0000	7.0000	0	M	0	M	0	M		
Slack_C1	0	1.0000	2.0000	6.0000	9.0000	-0.0000	1.0000	0	0	0	0	200.0000	200.0000
Surplus_C2	-M	8.0000	-5.0000	2.0000	-1.0000	1.0000	0	-1.0000	0	0	0	130.0000	16.2500
Artificial_C3	-M	7.0000	1.0000	1.0000	0	0	0	0	0	0	0	20.0000	2.8571
Slack_C4	0	-1.0000	1.0000	0	0	0	0	0	1.0000	0	0	0	0
Slack_C5	0	0	0	-1.0000	1.0000	-1.0000	0	0	0	1.0000	0	10.0000	M
Slack_C6_X3	0	0	0	1.0000	0	0	0	0	0	0	1.0000	10.0000	M
	C(j)-Z(j)	6.0000	5.0000	1.0000	7.0000	-7.0000	0	0	0	0	0	-M	
	*sig N	15.0000	-4.0000	3.0000	-1.0000	1.0000	0	-1.0000	0	0	0	0	

图 2.10

lp1

Maximize	6X1+5X2+1X3+7X4
	OBJ/Constraint/Variable Type/Bound
Maximize	6X1+5X2+1X3+7X4
C1	1X1+2X2+6X3+9X4<=260
C2	8X1-5X2+2X3-1X4>=150
C3	7X1+1X2+1X3=30
C4	1X1-1X2>=0
C5	1X3-1X4>=0
Integer	
Binary	
Unrestricted	X4
X1	>=0, <=M
X2	>=0, <=M
X3	>=10, <=20
X4	>=-M, <=M

图 2.11　标准模型输入格式

表 2.10　线性规划常用熟语词汇及其含义

常用术语	含义	常用术语	含义
Alternative Solution Exists	有多重解	Minimum and Maximum Allowable RHS	最优基不变时，资源限量允许变化范围
Basic and Nonbasic Variable	基变量和非基变量	Objective Function	目标函数
Basis	基	Optimal Solution	最优解
Basis Status	基变量状态	Parametric Analysis	参数分析
Branch-and-Bound Mrthod	分支定界法	Range and Slope of Parametric Analysis	参数分析的区间和斜率
Cj - Zj	检验数	Reduced Cost	约简成本(价值)
Combined Report	组合报告	Range of Feasibility	可行区间
Constraint Summary	约束条件摘要	Range of Optimality	最优区间
Constraint	约束条件	Relaxed Problem	松弛问题
Constraint Direction	约束方向	Relaxed Optimum	松弛最优
Constraint Status	约束状态	Right - hand Side	右端常数
Decision Variable	决策变量	Sensitivity Analysis of OBJ Coefficients	目标函数的灵敏度分析
Dual Problem	对偶问题	Sensitivity Analysis of Right - Hand - sides	右端常数的灵敏度分析
Entering Variable	入基变量	Shadow Price	影子价格
Feasible Area	可行域	Simplex Method	单纯形法
Feasible Solution	可行解	Slack, Surplus or Artificial Variable	松弛变量、剩余变量或人工变量
Infeasible	不可行	Solution Summary	最优解摘要
Infeasibility Analysis	不可行分析	Subtract(Add) More Than This From A(i,j)	减少(增加)约束系数
Leaving Variable	出基变量	Total Contribution	总体贡献
Left - hand side	左端	Unbounded Solution	无界解
Lower or Upper Bound	上界或下界		
Minimum and Maximum	最优解不变时，价值		
Allowable Cj	系数允许变化范围		

案例:超级食品公司的广告组合问题

超级食品公司正在考虑如何才能进入本来就有许多供应商的早点谷类食品市场。幸运的是,他们有一个很好的早点食品产品。公司副总裁决定搞一次促销活动。不过她非常清楚的是,她必须避免上一次产品促销所犯的一个错误,那就是忽略了与那次活动没有多大关系的目标市场:幼年儿童及其父母。同时她还总结出另一个失误是:没有将优惠券包含在杂志和报纸的广告中。

这次,她决定和管理决策小组携手共同来研究和解决问题了。

1. 问题及问题的数据

公司拟雇佣一家第一流的广告公司 G&J 公司来帮助设计全国性的促销活动,并付给广告公司不超过 100 万美元的酬金;另外分配 400 万美元作为广告费用。公司和管理决策小组一致同意以目标市场的消费者感知量最大化代替利润最大化作为总的绩效测度(目标函数)。

G&J 公司确定出这种产品最有效的三种广告媒体:

媒体 1:周六上午儿童节目的电视广告;

媒体 2:食品与家庭导向的杂志上的广告;

媒体 3:主要报纸周日增刊上的广告。

将收集到的数据整理如 2.11 所示。

表 2.11

资源	每种活动的单位资源使用量			可获得的资源数
	电视广告	杂志广告	周日增刊广告	
1. 广告预算	30 万	15 万	10 万	400 万
2. 计划预算	9 万	3 万	4 万	100 万
3. 电视时段	1	0	0	5
单位贡献	130	60	50	

单位贡献 = 单位广告可取得的消费者感知量的单位数

(上面数据解释一下:第一行数值表示各种媒体上每一广告的单位成本;第二行表示 G&J 公司公司用于设计与开发各种媒体广告的估计费用;第三行表示每一广告可以使用可获得的五个电视广告时段之一)

现在的问题是:如何确定各种广告活动的水平以取得最有效的广告组合。

2. 建立代数模型及求解

拟订代数符号:Q—— 目标消费者的感知量;TV—— 电视广告时段数目;M—— 杂志上广告数目;SS—— 周日增刊上的广告数目。

模型:

$$MaxQ = 130TV + 60M + 50SS$$

$$st.\begin{cases}300TV + 150M + 100SS \leqslant 4\,000\\ 90TV + 30M + 40SS \leqslant 1\,000\\ TV \leqslant 5\\ TV, M, SS \geqslant 0\end{cases}$$

容易求解上面模型，最优解是(0,20,10)，即不做任何电视广告，在杂志上刊登20出广告，在周日增刊上刊登10出广告。

3. 模型正确性评价

管理决策小组认识到，模型是在一定假设基础上对实际问题的抽象，应该要求模型和实际有较高的关联度。考查二者是否达到这个较高的关联度，就是对模型正确性的评价。

模型假设之一是允许有分数解。事实上对于做广告这类事情不能有分数解的。所幸的是，我们模型正好都是整数解。

模型假设之二是增加的广告的效果不变。事实上，随着这一媒体上广告数的增加，并达到饱和时，增加的广告效果肯定会减小。

而且，模型还假定了不同媒体不互相影响，这也不是绝对真实的。

模型假设之三是：媒体上广告的开发成本与数量成比例，同种媒体上广告开发成本不会受到其他媒体广告开发投入的多少的影响。事实上，*G&J* 公司最近在另一媒体上做过相同主题的广告而大大减少成本也是可能的。

为了判断“增加的广告的效果不变”的假定是否合理，管理决策小组者们与 *G&J* 公司会谈。*G&J* 公司认为，目前广告远远没有达到饱和水平，大多数读者最多看到两次广告，所以上面假定可以认为是合理的。

接下来，管理决策小组者们询问 *G&J* 公司广告的开发成本问题。*G&J* 公司承认在不同媒体上的广告设计是有重叠部分，特别是均为印刷的媒体时，但是，由于不同媒体截然不同的特色，这种重叠是非常有限的。此外，*G&J* 公司认为，比例性的假设很合理，因为在某一种媒体上设计开发广告的工作量大致相等，基本上等同于在该媒体上第一次设计开发该种广告的工作量。

基于上面的讨论与了解，管理决策小组认为，上面模型是合理的，可以提供给管理层参考使用。

4. 管理者的反应

总裁：居然没有使用电视广告，但好像也有些合理。(你觉得如何?)

市场部副总裁：我对不使用电视广告有些意见。大家知道，能够使儿童向他们父母索要我们的这个产品是极为重要的。这将是我们产品销售的最佳途径。而那些电视广告时段往往都是家长陪同小孩一起看的。因此，我们需要的是，能同时迎合儿童和家长的广告时段，能让儿童立即要求家长去买。我认为，这才是这次促销活动成功的关键。

总裁：是的。我们必须让管理决策小组将这一因素考虑进去，重新计算一遍。你只需要告诉他们这次促销活动所要接触的儿童及其家长的数量就可以了。(还有哪些因素管理决策小组忽略了呢?)

市场部副总裁：没问题，我已经有了那些数据。另外，该方案没有考虑杂志和报纸

上的优惠券的预算。我想,他们把这一预算加进去,而且是不困难的。

5　模型的重新设定与求解

为此,公司提供149万元作为优惠券预算,同时,管理层提出两个目标。

目标1:必须至少有500万儿童看到该广告;

目标2:必须至少有500万儿童的家长看到该广告。

实际上,这两个目标就是广告活动收益的最低可接受水平。该收益由广告接触到的人群数量来衡量。

为了建立相应的收益约束,管理决策小组要求*G&J*公司估计出每一个单位的各种广告能让多少儿童及其家长看到。调查和估计的结果如表(表2.12所示)。

表2.12

收益	每一活动的单位收益			最低可接受水平
	电视广告	杂志广告	周日增刊广告	
儿童	1.2万人	0.1万人	0万人	5万人
家长	0.5万人	0.2万人	0.2万人	5万人

对优惠券,有如表2.13所示的安排。

表2.13

需求	每一活动对需求数量的单位贡献			需求数量
	电视广告	杂志广告	周日增刊广告	
优惠券	0	40 000	120 000	1 490 000

因此,除了原先的模型约束外,我们在这里还增加了两个收益约束,一个确定需求约束:

$$\begin{cases} 1.2TV + 0.1M \geqslant 5 & \text{(目标1,单位是百元)} \\ 0.5TV + 12M \geqslant 5 & \text{(目标2,单位是百万)} \\ 4M + 12SS = 1\,490 & \text{(以10 000美元为单位)} \end{cases}$$

因而总结起来,整个模型变为:

$$MaxQ = 130TV + 60M + 50SS$$

约束条件:

1. 资源约束

$$\begin{cases} 300TV + 150M + 100SS \leqslant 4\,000 \\ 90TV + 30M + 40SS \leqslant 1\,000 \\ TV \leqslant 5 \end{cases}$$

2. 收益约束

$$\begin{cases} 1.2TV + 0.1M \geqslant 5 \\ 0.5TV + 12M \geqslant 5 \end{cases}$$

3. 确定需求约束

$$4M + 12SS = 149$$

4. 非负约束

$TV, M, SS \geqslant 0$

重新求解，即可得到最优解是(3,14,7.75)，即在电视上做3处广告，在杂志上做14处广告，在周日增刊上做7.75处广告(因此第八个增刊广告仅能在75% 的报纸上刊出)。

习题

1. 求解线性规划问题时可能出现哪几种结果?哪些结果反映建模时有错误?

2. 判断正误

(1) 线性规划模型中增加一个约束条件，可行域的范围一般将缩小；减少一个条件，可行域的范围一般将扩大。

(2) 线性规划问题的每一个基解对应可行域的一个顶点。

(3) 用单纯形法求解标准型的线性规划问题时，$\sigma_j > 0$ 与对应的变量都可以被选为换入变量。

(4) 用单纯形法求解标准型的线性规划问题时，选取最大正检验数对应的变量作为换入变量，将使目标函数值得到最快的增长。

(5) 用单纯形法求解标准型的线性规划问题时，单纯形法的迭代计算过程是从一个可行解转换到目标函数值更大的另一个可行解。

3. 用图解法求解下列线性规划问题，并指出各问题是具有唯一最优解、无穷多最优解、无界解或无可行解。

(1) $\text{Min } z = 6x_1 + 4x_2$

$$s.t. \begin{cases} 2x_1 + x_2 \geqslant 1 \\ 3x_1 + 4x_2 \geqslant 1.5 \\ x_1, x_2 \geqslant 0 \end{cases}$$

(2) $\text{Max } z = 4x_1 + 8x_2$

$$s.t. \begin{cases} x_1 + x_2 \leqslant 5 \\ -x_1 + x_2 \geqslant 8 \\ x_1, x_2 \geqslant 0 \end{cases}$$

(3) $\text{Max } z = x_1 + x_2$

$$s.t. \begin{cases} 4x_1 + 3x_2 \geqslant 12 \\ 2x_1 + 3x_2 \geqslant -6 \\ x_2 \geqslant 2 \\ x_1, x_2 \geqslant 0 \end{cases}$$

(4) $\text{Max } z = 3x_1 + 9x_2$

$$s.t.\begin{cases} x_1 + 3x_2 \leqslant 22 \\ -x_1 + x_2 \leqslant 4 \\ x_2 \leqslant 6 \\ 2x_1 - 5x_2 \leqslant 0 \\ x_1, x_2 \geqslant 0 \end{cases}$$

4. 某炼油厂根据计划每季度需要供应合同单位汽油15万吨、煤油12万吨、重油12万吨。该厂从A、B两处运回原油提炼。已知两处原油成分如表2.14所示。又知从A处采购原油每吨价格(含运费等)为200元,B处原油价格(含运费等)为310元。试求该炼油厂采购原油的最优决策。

表2.14　　原油成分

	A(%)	B(%)
含汽油	15	50
含煤油	20	30
含重油	50	15
其他	15	5

5. 将下列线性规划问题变成标准型,并列出初始单纯形表。

(1) $\text{Min } z = 2x_1 - x_2 + x_3 + 2x_3$

$$s.t.\begin{cases} -x_1 + x_2 + x_3 = 4 \\ -x_1 + x_2 - x_3 \leqslant 6 \\ x_1 \leqslant 0, x_2 \geqslant 0 \end{cases}$$

(2) $\text{Max } z = 2x_1 + x_2 + 3x_3 + x_4$

$$s.t.\begin{cases} x_1 + x_2 + 3x_3 + x_4 \leqslant 7 \\ 2x_1 - 3x_2 + 5x_3 = -8 \\ x_1 - 2x_3 + 2x_4 \geqslant 1 \\ x_1, x_3 \geqslant 0, x_2 \leqslant 0 \end{cases}$$

6. 用单纯形法求解下列线性规划问题,并指出问题的解属于哪一类。

(1) $\text{Max } z = 3x_1 + 5x_2$

$$s.t.\begin{cases} x_1 \leqslant 4 \\ x_2 \leqslant 6 \\ 3x_1 + 2x_2 \leqslant 18 \\ x_1, x_2 \geqslant 0 \end{cases}$$

(2) $\text{Max } z = 2x_1 - x_2 + x_3$

$$s.t.\begin{cases} 3x_1 + x_2 + x_3 \leqslant 60 \\ x_1 - x_2 + 2x_3 \leqslant 10 \\ x_1 + x_2 - x_3 \leqslant 20 \\ x_1, x_2, x_3 \geqslant 0 \end{cases}$$

(3) $\text{Max } z = 6x_1 + 2x_2 + 10x_3 + 8x_4$

$$s.t.\begin{cases}6x_1 + 2x_2 + 10x_3 + 8x_4 \leqslant 20\\3x_1 - 3x_2 + 2x_3 + 8x_4 \leqslant 25\\4x_1 - 2x_2 + x_3 + 3x_4 \leqslant 10\\x_1, x_2, x_3, x_4 \geqslant 0\end{cases}$$

(4) $\text{Max } z = x_1 + 6x_2 + 4x_3$

$$s.t.\begin{cases}-x_1 + 2x_2 + 2x_3 \leqslant 13\\4x_1 - 4x_2 + x_3 \leqslant 20\\x_1 + 2x_2 + x_3 \leqslant 17\\x_1 \geqslant 1, x_2 \geqslant 2, x_3 \geqslant 3\end{cases}$$

7. 某医院昼夜24小时各时段需要的护士数量如下:2:00 ~ 6:00,10人; 6:00 ~ 10:00,15人; 10:00 ~ 14:00,25人; 14:00 ~ 18:00,20人; 18:00 ~ 22:00,18人; 22:00 ~ 2:00,12人。护士分别于各个时段的开始时间,分6批上班,并连续工作8个小时。试确定,该医院至少需要多少名护士,才能满足上班的需要。

8. 某人有一笔30万元的资金,在今后三年内可以有以下投资项目:

(1) 三年内每年年初均可投资,每年获利为投资额的20%,其本利可以一起用于下年投资。

(2) 只允许第一年年初投资,第二年年末可收回,本利合计为投资额的150%,但此类投资不能超过15万元。

(3) 于三年内第二年年初允许投资,可于第三年年末回收,本利合计为投资额的160%,投资限额为20万元。

(4) 于三年内的第三年初允许投资,一年回收,可获利40%,投资限额为10万元。

试为此人确定一个最优的投资方案,使得其三年内获利最大。

第三章　对偶理论与灵敏度分析

引例：家具生产问题

设有一个家具厂用木材和钢材生产 A、B、C 三种家具。生产一件家具所需要的材料数量、每件家具可以获得的利润以及每月可提供的材料数量如表 3.1 所示。问：如何安排各种家具每月的生产量，该厂才能在该月获得最大利润？

表 3.1

产品	A 产品	B 产品	C 产品	材料可供量
木材	3	1	4	8 000
钢材	2	1	4	3 000
单位产品利润(单位:元)	4	1	5	

这是上一章的例 2.2。现在换一个角度来思考这个问题。假设该厂除了自己生产之外，还有出售拥有的资源的选择。那么，决策者就应该考虑这些资源的定价问题，这样才能比较是自己生产好，还是出售好。

第一节　对偶线性规划模型

假设木材的价格用 y_1 表示，钢材的价格用 y_2 表示。决策者考虑：

(1)A 产品：3 个单位的木材和 2 个单位的钢材可以生产一个单位的 A 产品，获利 4 元。所以，3 个单位的木材和 2 个单位的钢材的组合价格不能低于 4 元，即：

$$3y_1 + 2y_2 \geqslant 4$$

(2) 对 B 产品和 C 产品做同样的分析，可得到另外两个要求(即约束条件)：

$$\begin{cases} y_1 + y_2 \geqslant 1 \\ 4y_1 + 4y_2 \geqslant 5 \end{cases}$$

(3) 工厂每月共有 8 000 单位的木材和 3 000 单位的钢材，若是全部出售，则总收入为：

$$\omega = 8\,000y_1 + 3\,000y_2$$

从工厂的角度来看，当然是价格越高越好，但是从资源购买者的角度来看，却是价格越低越好，所以工厂决策者只能在满足上面(1) 和(2) 所给出的要求的前提下，使其总收入尽可能地低。

从而得到下面的线性规划问题：

$$\text{Min } \omega = 8\,000y_1 + 3\,000y_2$$

$$s.t.\begin{cases}3y_1+2y_2\geqslant 4\\ y_1+y_2\geqslant 1\\ 4y_1+4y_2\geqslant 5\\ y_1,y_2\geqslant 0\end{cases}$$

我们称这个问题为例 2.2 的对偶问题,把例 2.2 称为原问题。

一、对偶模型

线性规划模型的矩阵形式为:

$$\text{Max } z = CX$$

$$s.t.\begin{cases}AX\leqslant b\\ X\geqslant 0\end{cases} \tag{3.1}$$

称 C_BB^{-1} 为单纯形乘子。令 $Y=C_BB^{-1}$,当所有检验数都为非正数,即 $C-C_BB^{-1}A\leqslant 0$ 与 $-C_BB^{-1}\leqslant 0$ 时,模型(3.1)得到最优解。可见检验数非正,可以作为得到最优解的条件。检验数为非正数时,有 $YA\geqslant C, Y\geqslant 0$。

在 $Y=C_BB^{-1}$ 两边都乘 b,得 $Yb=C_BB^{-1}b$。Y 的上界无限大,所以 Yb 只存在最小值。

于是得到模型(3.1)的对偶模型:

$$\text{Min } \omega = Yb$$

$$s.t.\begin{cases}YA\leqslant C\\ Y\geqslant 0\end{cases} \tag{3.2}$$

比较上原模型和对偶模型的目标函数系数、约束条件系数矩阵、约束条件右端值等可知,只要知道原模型就可以对应写出对偶模型。但上面原模型约束条件只有“≤”不等式,目标函数也只是求最大值,而在实际问题中,原模型约束条件和目标函数可能有多种形式。以下把各种形式的目标函数和约束条件在对偶模型中的对应关系列出,如表 3.2 所示。

表 3.2

原问题(或对偶问题)		对偶问题(或原问题)	
目标函数 Max z		目标函数 Min ω	
变	n 个	n 个	约
	$\geqslant 0$	$\geqslant$	束
	$\leqslant 0$	$\leqslant$	条
量	无约束	$=$	件
约	m 个	m 个	变
束	$\geqslant 0$	$\geqslant$	
条	$\leqslant 0$	$\leqslant$	
件	无约束	$=$	量
约束条件右端项		目标函数变量的系数	
目标函数变量的系数		约束条件右端项	

例 3.1 试求如下线性规划问题的对偶问题：

$$\text{Min } \omega = 2x_1 + 3x_2 - 5x_3 + x_4$$

$$s.t. \begin{cases} x_1 + x_2 - x_3 + x_4 \geqslant 5 & (1) \\ 2x_1 + 2x_3 - x_4 \geqslant 5 & (2) \\ x_2 + x_3 + x_4 = 5 & (3) \\ x_1 \leqslant 0; x_2, x_3 \geqslant 0 \end{cases}$$

解：设对应约束条件(1)(2)(3)的对偶变量分别为 y_1、y_2、y_3，则由表 3.2 原问题和对偶问题的对应关系，可以直接写出上述问题的对偶问题：

$$\text{Max } z = 5y_1 + 4y_2 + 6y_3$$

$$s.t. \begin{cases} y_1 + 2y_2 \geqslant 2 \\ y_1 + y_3 \leqslant 3 \\ -3y_1 + 2y_2 + y_3 \leqslant 5 \\ y_1 - y_2 + y_3 = 1 \\ y_1 \geqslant 0, y_2 \leqslant 0 \end{cases}$$

二、对偶问题性质

(1) 对称性。对偶问题的对偶是原问题。

(2) 对偶性。若 $\bar{X}$ 为原问题的解，$\bar{Y}$ 是对偶问题的可行解，则存在 $C\bar{X} \leqslant \bar{Y}b$。

(3) 无界性。若原问题(对偶问题)为无界解，则其对偶问题(原问题)无可行解。

(4) 可行解是最优解的性质。设 $\hat{X}$ 是原问题的可行解，$\hat{Y}$ 是对偶问题的可行解，当 $C\hat{X} = \hat{Y}b$ 时，$\hat{X}$、$\hat{Y}$ 是最优解。

(5) 对偶定理。若原问题有最优解，那么对偶问题也有最优解，且目标函数值相等。

(6) 互补松弛性。若 $\hat{X}$、$\hat{Y}$ 分别是原问题和对偶问题的可行解，那么 $\hat{Y}X_s = 0$ 和 $Y_s\hat{X} = 0$，当且仅当 $\hat{X}$、$\hat{Y}$ 是最优解。

(7) 设原问题是：$\text{Max } z = CX; AX + X_s = b; X, X_s \geqslant 0$

对偶问题是：$\text{Min } \omega = Y_b; YA - Y_s = C; Y, Y_s \geqslant 0$

则原问题单纯形表的检验数行对应其对偶问题的一个基解，对应关系如表 3.3 所示。

表 3.3

	X_B	X_{N_1}	X_s
检验数	0	$C_N - C_BB^{-1}N$	$-C_BB^{-1}$
对偶问题基解	Y_{s1}	$-Y_{s2}$	$-Y$

第二节 对偶问题的经济学解释——影子价格

前面讲到单纯形乘子 $Y = C_BB^{-1}$，它在目标函数取值 $z = C_BB^{-1}b$ 和检验数 $C_N -$

$C_BB^{-1}N_1$ 中都有出现，那么 Y 的经济意义是什么呢？

设 B 是 $\{\text{Max } z = CX \mid AX \leqslant b, X \geqslant 0\}$ 的最优基，则：$z^* = C_BB^{-1}b = Y^*b$。对此式求关于 b 的一阶有偏导，有：$\dfrac{\partial z^*}{\partial b} = C_BB^{-1} = Y^*$。

所以变量 y_i^* 的经济意义是：在其他条件不变的情况下，第 i 种资源的单位变化所引起的目标函数的最优值的变化，即对目标值的贡献（增加或者减少），称为第 i 种资源的影子价格。这种价格是针对具体工厂的具体产品而言的。

例如，在上一章对例 2.2 进行求解时，最后单纯形成表 3.4。

表 3.4

c_j			4	1	5	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
0	x_4	3 500	0	-1/2	-2	1	-3/2	
5	x_1	1 500	1	1/2	2	0	1/2	
		Z_j	4	2	8	0	2	
		$c_j - Z_j$	0	-1	-3	0	-2	

由表 3.3 可知，原问题单纯形表的检验数行对应其对偶问题的一个基解。在例 2.2 中，$X_s = (x_4, x_5)$，若 x_4 和 x_5 对应的检验数分别为 0 和 -2，则对应对偶问题的变量 y_1 和 y_2 取值为 0 和 2。也即木材和钢材的影子价格为 0 元和 2 元。若工厂不想自己生产，可以在原成本的基础上分别加价 0 元和 2 元出售木材和钢材。可以验证，按最优决策，当自己生产 1 500 单位 A 产品时，消耗 4 500 单位木材和 3 000 钢材单位，可赚 $1\,500 \times 4 = 6\,000$（元）；若是出售这 4 500 单位木材和 3 000 单位钢材，可以赚 $4\,500 \times 0 + 3\,000 \times 2 = 6\,000$（元），两者的结果是一样的。

自己生产和出售原料获利是一样的，这事实上只是一种理想情况。现实生活中，买卖双方往往信息不对称，生产能力和技术条件也不一定一样，所以，不同企业生产同样的产品，获利往往不一样。这就使得有些单位愿意出更高的价格来购买原料。这个事实可以用来指导我们的决策：当市场上某种资源的市场价低于资源成本与影子价格之和时，企业应该买入资源自己生产；而当市场价格高于资源成本与影子价格之和时，决策者就应该把资源卖掉。

第三节　对偶单纯形法

从前面知道，原问题和对偶问题的解有如下对应关系（性质 7）：单纯形表中，在 b 列中得到原问题的基可行解，在检验数行中得到对偶问题的基解。保持 b 列为原问题的基可行解，通过逐步迭代，当检验数都非正（即检验数变成对偶问题的基可行解）时，检验数行变成对偶问题的最优解。此时原问题和对偶问题同时得到最优解。

反之，可以根据原问题和对偶问题的对称性，在求解时这样处理：单纯形表中，保持检验数行为非正，通过逐步迭代将 b 列变成基可行解。这时便同时得到了原问题和

对偶问题的最优解。这样的求解方法称为对偶单纯形法。

这样做的优点在于:第一,当问题的模型中存在“≥”的约束条件时,不需要引入人工变量,只要在约束条件方程的两边乘以“-1”后加入松弛变量,再按单纯形法求解;第二,当约束条件方程个数 m 大于变量个数 n 的时候,将原问题变换成对偶问题求解,计算量一般会小些。

对偶单纯形法的计算步骤为:

第一步,根据模型,列出初始单纯形表。

第二步,若 b 列的数字都非负,检验数都非正,则已经得到最优解,停止计算;若 b 列中至少还有一个负数时,检验数保持非正,转入第三步。

第三步,确定换出变量。

按 $\text{Min}\{(B^{-1}b)_i \mid (B^{-1}b)_i < 0\} = (B^{-1}b)_i$,对应的基变量 x_i 为换出变量。

第四步,确定换入变量。

在单纯形表中,检查 x_i 所在行的系数 a_{ij},若所有的 $a_{ij} \geqslant 0$,则无可行解,停止计算。若存在 $a_{ij} \leqslant 0,(i = 1,2,\cdots,n)$,则计算:

$$\theta = \text{Min}_i\left\{\frac{c_j - Z_j}{a_{ij}} \mid a_{ij} < 0\right\} = \frac{c_k - z_k}{a_{lk}}$$

按 θ 规则,所对应的列的非基变量 x_k 为换入变量。

第五步,以 a_{lk} 为主元素,在原单纯形表中进行迭代运算,得新的单纯形表,转入第二步。

例 3.2 用对偶单纯形法求解:

$$\text{Min } S = 1\,200x_1 + 800x_2 + 1\,600x_3 + 1\,200x_4$$

$$s.t.\begin{cases} 2x_1 + x_2 + 4x_3 \geqslant 2 \\ 2x_1 + 2x_2 + 4x_4 \geqslant 3 \\ x_1, x_2, x_3, x_4 \geqslant 0 \end{cases}$$

解:先将模型化成标准型:目标函数乘以 -1,在模型约束条件两边乘以 -1,加入松弛变量。

$$\text{Max } z = -1\,200x_1 - 800x_2 - 1\,600x_3 - 1\,200x_4$$

$$s.t.\begin{cases} -2x_1 - x_2 - 4x_3 + 0x_4 + x_5 = -2 \\ -2x_1 - 2x_2 + 0x_3 - 4x_4 + x_6 = -3 \\ x_1, x_2, x_3, x_4, x_5, x_6 \geqslant 0 \end{cases}$$

初始单纯表和后续各表如表 3.5、表 3.6、表 3.7、表 3.8 所示。

表 3.5

c_j			-1 200	-800	-1 600	-1 200	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	x_6	
0	x_5	-2	-2	-1	-4	0	1	0	
0	x_6	-3	-2	2	0	-4	0	1	
	Z_j		0	0	0	0	0	0	
	$c_j - Z_j$		-1 200	-800	-1 600	-1 200	0	0	

表 3.6

c_j			-1 200	-800	-1 600	-1 200	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	x_6	
0	x_5	-2	-2	①	-4	0	1	0	800
-1 200	x_4	3/4	1/2	1/2	0	1	0	-1/4	
		Z_j	600	-600	0	-1 200	0	300	
		$c_j - Z_j$	-600	-200	-1 600	0	0	-300	

表 3.7

c_j			-1 200	-800	-1 600	-1 200	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	x_6	
-800	x_2	2	2	1	4	0	1	0	
-1 200	x_4	-1/4	1/2	0	②	1	1/2	-1/4	
		Z_j	1 000	-800	-800	-1 200	200	300	
		$c_j - Z_j$	-200	0	-800	0	-200	-300	

表 3.8

c_j			-1 200	-800	-1 600	-1 200	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	x_6	
-800	x_2	3/2	1	1	0	2	0	-1/2	
-1 600	x_3	1/8	1/4	0	1	-1/2	-1/4	1/8	
		Z_j	-1 200	-800	-1 600	-800	400	200	
		$c_j - Z_j$	0	0	0	-400	-400	-200	

至此，问题得到最优解 $X = (0, 3/2, 1/8, 0, 0, 0)$，目标函数值为 1 400。

第四节　灵敏度分析

线性规划模型是对实际问题的一个粗略抽象。通常，在取得最初版本的模型的最优解之后，需要进行分析才能得到对问题的深入认识。其中，研究参数的变化对最优解的影响，称为灵敏度分析。其主要问题形式是：如果假设的情况发生变动，最优解将会发生怎样的变化。

灵敏度分析的几个基本作用：

（1）线性规划模型的许多数据（也叫参数）在建模时是很难精确确定的，只能对某些数据加以估计。通过灵敏度分析可以表明，参数估计值必须精确到怎样的程度，才能避免得出错误的最优解，而且，可以找到哪些参数是需要重新精确定义的灵敏度参数（Sensitive Parameters）。

（2）如果研究结束后，问题的条件发生变化，通过灵敏度分析，即使不求解，也可以估算模型参数的变化是否会改变最优解。

灵敏度分析还可以进一步进行前摄性分析，即测试各种可能的管理行为会对模型造成什么影响。如果模型的主要参数是管理层的政策而不是管理层所无法控制的外部数值时，可以用这种方法。这便有了下面第三个作用。

（3）当模型特定的参数反映管理政策决策时，灵敏度分析可以表明改变这些决策对结果的影响，从而有效地指导管理者作出最终决策。

一、资源数量变化的分析

资源数量变化指系数 b 发生变化。假设第 r 种资源发生了变动，而其他系数都不变。这时，线性规划模型的最优解可能发生变化。现在希望找到第 r 种资源变动的最大范围，使得最优基能保持不变。

设原最终表中最优基是 B，新的资源向量是 $b' = b + \Delta b$，其中 $\Delta b = (0\cdots0, \Delta b_r, 0\cdots0)^T$。从而可知，原问题的解变为：

$$X'_B = B^{-1}(b + \Delta b) = B^{-1}b + B^{-1}\Delta b = B^{-1} + b^{-1}\begin{bmatrix} 0 \\ \cdots \\ \Delta b_r \\ \cdots \\ 0 \end{bmatrix}$$

$$\text{其中：} B^{-1}\begin{bmatrix} 0 \\ \cdots \\ \Delta b_r \\ \cdots \\ 0 \end{bmatrix} = \begin{bmatrix} \bar{a}_{1r}\Delta b_r \\ \cdots \\ \bar{a}_{ir}\Delta b_r \\ \cdots \\ \bar{a}_{mr}\Delta b_r \end{bmatrix} = \Delta b_r \begin{bmatrix} \bar{a}_{1r} \\ \cdots \\ \bar{a}_{ir} \\ \cdots \\ \bar{a}_{mr} \end{bmatrix}, \begin{bmatrix} \bar{a}_{1r} \\ \cdots \\ \bar{a}_{ir} \\ \cdots \\ \bar{a}_{mr} \end{bmatrix} \text{为 } B^{-1} \text{ 的第 } r \text{ 列向量。}$$

只要 $X'_B \geqslant 0$，最终表中检验数不变，则最优基不变，但最优解的值发生了变化。这时，在最终表中求得的 b 列的所有元素 $\bar{b}_i + \bar{a}_{ir}\Delta b_i \geqslant 0, i = 1,2,\cdots,m$，即：

$\bar{a}_{ir}\Delta b_i \geqslant -\bar{b}_i, i = 1,2,\cdots,m$

当 $\bar{a}_{ir} \geqslant 0$ 时，$\Delta b_r \geqslant -\bar{b}_i/\bar{a}_{ir}$；当 $\bar{a}_{ir} \leqslant 0$ 时，$\Delta b_r \leqslant -\bar{b}_i/\bar{a}_{ir}$，于是得到保持最优基不变的 Δb_r 的变动范围：

$$\text{Max}\{-\bar{b}_i/\bar{a}_{ir} \mid \bar{a}_{ir} > 0\} \leqslant \Delta b_r \leqslant \text{Min}\{-\bar{b}_i/\bar{a}_{ir} \mid \bar{a}_{ir} < 0\}$$

例如，第二章例 2.2 中最后的单纯形表如表 3.9 所示：

表 3.9

c_j			4	1	5	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
0	x_4	3 500	0	-1/2	-2	1	-3/2	
5	x_1	1 500	1	1/2	2	0	1/2	
		Z_j	4	2	8	0	2	
		$c_j - Z_j$	0	-1	-3	0	-2	

那么钢材的变化范围可计算如下：

$$X'_B = B^{-1}(b + \Delta b) = B^{-1}b + B^{-1}\Delta b = B^{-1}\Delta b + B^{-1}\begin{bmatrix}0\\ \Delta b_2\end{bmatrix}$$

$$= \begin{bmatrix}3\ 500\\ 1\ 500\end{bmatrix} + \begin{bmatrix}1 & -3/2\\ 0 & 1/2\end{bmatrix}\begin{bmatrix}0\\ \Delta b_2\end{bmatrix}$$

$$= \begin{bmatrix}3\ 500\\ 1\ 500\end{bmatrix} + \begin{bmatrix}-3/2\\ 1/2\end{bmatrix}\begin{bmatrix}0\\ \Delta b_2\end{bmatrix}$$

$$= \begin{bmatrix}3\ 500\\ 1\ 500 + 1/2\Delta b_2\end{bmatrix} \geqslant 0$$

由 $1\ 500 + 1/2\Delta b_2 \geqslant 0$,得 $\Delta b_2 \geqslant -3\ 000$。

从上面可以看到,若 Δb_r 在保持最优基不变的范围变动,可以不必重新求解原模型,便可以求得最优解的值 $X'_B = B^{-1}(b + \Delta b) = B^{-1}b + B^{-1}\Delta b$,这一点在企业经营管理中很有用。

例如,例2.2中,可用的钢材增加了1 000个单位。这时最优的方案还是只是生产 A 产品。最优解的值为:

$$X'_B = B^{-1}(b + \Delta b) = B^{-1}b + B^{-1}\Delta b$$

$$= \begin{bmatrix}3\ 500\\ 1\ 500 + 1/2\Delta b_2\end{bmatrix}$$

$$= \begin{bmatrix}3\ 500\\ 1\ 500 + 1/2 \times 1\ 000\end{bmatrix}$$

$$= \begin{bmatrix}3\ 500\\ 2\ 000\end{bmatrix}$$

也就是,说生产 A 产品 2 000 个单位是最优决策。

但是,有时候资源的变动可能会超过其保持最优基不变的范围,这时候也可以不必从头计算资源改变后模型。可以在原模型最后的单纯形表上直接做一些调整之后,接着求解单纯形表就可以了。

例 3.3 某工厂在计划期内,要安排生产甲乙两种产品,已知生产单位产品所需的设备台时、材料消耗如表 3.10 所示。

表 3.10

	单位产品消耗资源量		资源可用量
	甲产品	乙产品	
设备台时	1	0	4
A 材料	0	2	12
B 材料	3	2	18
单位利润(元)	2	3	

试问:如何安排生产才能得到最大利润?若是增加 4 台机器,这时原最优生产方案将做什么调整?

解:(1) 根据问题,设甲乙两种产品各生产 x_1、x_2 个单位。则可以建立模型如下:

$$\text{Max } z = 2x_1 + 3x_2$$

$$s.t.\begin{cases} x_1 + 2x_2 \leqslant 8 \\ 4x_1 \leqslant 16 \\ 4x_2 \leqslant 12 \\ x_1, x_2 \geqslant 0 \end{cases}$$

将原模型变成标准型，在约束条件方程中分别加入松弛变量 x_3、x_4、x_5。则新模型为：

$$\text{Max } z = 2x_1 + 3x_2 + 0x_3 + 0x_4 + 0x_5$$

$$s.t.\begin{cases} x_1 + 2x_2 + x_3 = 8 \\ 4x_1 + x_4 = 16 \\ 4x_2 + x_5 = 12 \\ x_1, x_2, x_3, x_4, x_5 \geqslant 0 \end{cases}$$

建立初始单纯形表（见表 3.11）。

表 3.11

c_j			2	3	0	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
0	x_3	8	1	2	1	0	0	4
0	x_4	16	4	0	0	1	0	–
0	x_5	12	4	0	0	0	1	3
$c_j - Z_j$			0	2	3	0	0	

经过若干中间表，得到最终表为（见表 3.12）。

表 3.12

c_j			2	3	0	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
2	x_1	4	1	0	0	1/4	0	
0	x_5	4	0	0	– 2	1/2	1	
3	x_2	2	0	1	1/2	– 1/8	0	
$c_j - Z_j$			0	0	– 1.5	– 1/8	0	

这样便得到了最优的生产方案：$X = (4,2,0,0,4)^T$，即生产甲产品 4 个单位、乙产品 2 个单位，获利 14 元。

（2）当增加 4 个单位的设备台时，最优方案可以直接在上最终表中加以调整即可。

先计算 $B^{-1}\Delta b$：

$$B^{-1}\Delta b = \begin{bmatrix} 0 & 1/4 & 0 \\ -2 & 1/2 & 1 \\ 1/2 & -1/8 & 0 \end{bmatrix}\begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -8 \\ 2 \end{bmatrix}$$

将上述结果反映到上最终表中，有（见表 3.13）。

表 3 - 13

c_j			2	3	0	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
2	x_1	4 + 0	1	0	0	1/4	0	0
0	x_5	4 - 8	0	0	- 2	1/2	1	1
3	x_2	2 + 2	0	1	1/2	- 1/8	0	0
$c_j - Z_j$			0	0	- 1.5	- 1/8	0	

由于表3. 13 中b列有负数，而同时检验数行都非正，所以可以用对偶单纯形法求解上表。结果如表 3. 14 所示。

表 3. 14

c_j			2	3	0	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
2	x_1	4	1	0	0	1/4	1/4	
0	x_3	2	0	0	- 2	1/4	- 1/4	
3	x_2	3	0	1	1/2	0	0	
$c_j - Z_j$			0	0	0	- 1/2	- 3/4	

即该厂最优生产方案应改为：生产甲产品 4 件，乙产品 3 件，获利 17 元。

二、目标函数中价值系数的敏感性分析

不同的价值系数变化会带来不同的影响。以下分别对非基变量和基变量在目标函数中的价值系数的变化做敏感性分析，目的是要求出保持目前最优基和最优解取值不变的前提下，价值系数的最大变化范围。

1. c_j 是非基变量 x_j 的系数

若令 c_j 的变化量为 Δc_j。要保持目前最优基 B 不发生改变，c_j 的变化量 Δc_j 应该满足 $c_j + \Delta c_j - Z_j \leqslant 0$。即：$\Delta c_j \leqslant Z_j - c_j$。这便是保持目前最优基和最优解取值不变的 c_j 变动范围。

2. c_r 是基变量 x_r 的系数

因为 x_r 在最优解内，当 c_r 发生变动时，原检验数将发生变动，这将可能导致最优解发生变动。若要保持最优解不变，可以控制目前的任一非基变量不能成为换入变量。

当 c_r 变化 Δc_r 后，价值系数矩阵变化量为 $\Delta C = (0,\cdots,0,\Delta c_r,0,\cdots,0)$ 最终表中的非基变量检验数为：

$$\begin{aligned}\sigma'_j &= c_j - (C_B + \Delta C_r)B^{-1}A \\ &= c_j - C_B B^{-1}A - \Delta c_r \bar{a}_{rj} \\ &= \sigma_j - \Delta c_r \bar{a}_{rj}\end{aligned}$$

对一切非基变量 x_j，为使非基变量不成为换入变量，就需要满足：$\sigma'_j \leqslant 0$，于是要求：

对一切非基变量 x_j，当 $\bar{a}_{rj} < 0$ 时，$\Delta c_r < \sigma_j / \bar{a}_{rj}$；当 $\bar{a}_{rj} > 0$ 时，$\Delta c_r \geqslant \sigma_j / \bar{a}_{rj}$。从而得到 c_r 的变化范围：

$$\text{Max}_j \{\sigma_j / \bar{a}_{rj} \mid \bar{a}_{rj} > 0\} \leqslant \Delta c_r \leqslant \text{Max}_j \{\sigma_j / \bar{a}_{rj} \mid \bar{a}_{rj} < 0\} \qquad (3-1)$$

例 3.4 例 3.3 中乙产品的单位利润在什么范围变动时，不影响目前的最优决策？

解：假设乙产品的单位利润（即价值系数）的变化量为 Δc_2，按（3－1）式，要确定的变动范围，需要知道最终表的系数 $\bar{a}_{rj}$。例 2.2 的最终表如表 3.15 所示：

表 3.15

c_j			2	3	0	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
2	x_1	4	1	0	0	1/4	0	
0	x_5	4	0	0	－2	1/2	1	
3	x_2	2	0	1	1/2	－1/8	0	
$c_j - Z_j$			0	0	－1.5	－1/8	0	

当发生变化后，最终表变为表 3.16 所示。

表 3.16

c_j			2	$3 + \Delta c_2$	0	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
2	x_1	4	1	0	0	1/4	0	
0	x_5	4	0	0	－2	1/2	1	
$3 + \Delta c_2$	x_2	2	0	1	1/2	－1/8	0	
$c_j - Z_j$			0	0	$-1.5 - 1/2\Delta c_2$	$-1/8 + 1/8\Delta c_2$	0	

为了保持最优解不变，则 x_2 的检验数应该为 0，x_3 和 x_4 的检验数要非正。从上表可以看到，x_2 的检验数已经为 0。故只需要设定 Δc_2 的变动范围，使得 x_3 和 x_4 的检验数非正。

即 $-1.5 - 1/2\Delta c_2 \leqslant 0$ 和 $-1/8 + 1/8\Delta c_2 \leqslant 0$，易解得：$-3 \leqslant \Delta c_2 \leqslant 1$。

即 c_2 可以在[0,4] 的范围变动，不会影响目前的最优决策。

三、技术系数 a_{ij} 的敏感性分析

技术系数的变化对最优基和最优解的影响稍复杂些。为简明起见，以下用例子来对几种常见情况做讨论。

1. 是否应该增加新产品

例3.5 例3.3 中，假设该厂除了生产甲、乙两种产品之外，还有一种新产品丙可以考虑生产。生产单位产品丙，要消耗 A 原料 6 个单位和 B 原料 3 个单位，同时还要消耗 2 个台时。每件产品丙获利 5 元。问该厂是否应该安排生产产品丙；若生产则该生产多少，才能使得该厂利润最大。

解：设生产产品丙 x'_3 件。

产品丙的技术系数向量为 $P'_3 = (2,6,3)^T$,其在单纯形表中的检验数为:

$\sigma'_3 = c'_3 - C_B B^{-1} P'_3 5 - (1.5,0.125,0)(2,6,3)^T = 1.25 > 0$,说明生产产品丙是有利的。

产品丙在最终表中对应 x'_3 的列向量为:

$$B^{-1}P'_3 = \begin{bmatrix} 0 & 0.25 & 0 \\ -2 & 0.5 & 1 \\ 0.5 & -0.125 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 6 \\ 3 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 2 \\ 0.25 \end{bmatrix}$$

将上面计算结果填入例 3.3 的第一个问的最终表(见表 3.17)中,得:

表 3.17

c_j			2	3	0	0	0	5	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	x'_3	
2	x_1	4	1	0	0	1/4	0	1.5	
0	x_5	4	0	0	-2	1/2	1	1.2	
3	x_2	2	0	1	1/2	-1/8	0	0.25	
$c_j - Z_j$			0	0	-1.5	-1/8	0	1.25	

由于 b 列的数字没有变化,原问题是可行解。但检验数行中还有正检验数,说明目标函数还可以改善。

用前面介绍的单纯形表的求解步骤,易得此问题的最终表(见表 3.18)。

表 3.18

c_j			2	3	0	0	0	5	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	x'_3	
2	x_1	1	1	0	1.5	-0.125	-0.75	0	
5	x'_3	2	0	0	-1	0.25	0.5	1	
3	x_2	1.5	0	1	0.75	-0.187 5	-0.125	0	
$c_j - Z_j$			0	0	-0.25	-0.437 5	-0.625	0	

由表可以知道最优方案为:生产甲产品 1 个单位,生产乙产品 1.5 个单位,生产丙产品 2 个单位。

2. 技术系数 a_{ij} 发生变化

例 3.6 例 3.3 中,若产品 A 的生产工艺发生了变化,现在其技术系数向量变为 $P'_1 = (2,5,2)^T$,单位利润变为 4 元,试分析此对原最优安排有何影响。

解:设变化后的甲产品生产量为 x'_1,对应 x'_1 的系数列向量为 P'_1,在最终表中 P'_1 为:

$$B^{-1}P'_1 = \begin{bmatrix} 0 & 0.25 & 0 \\ -2 & 0.5 & 1 \\ 0.5 & -0.125 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 1.25 \\ 0.5 \\ 0.375 \end{bmatrix}$$

x'_1 的检验数为:

$\sigma'_1 = c'_1 - C_B B^{-1} P'_1 = 4 - (1.5,0.125,0)(2,5,2)^T = 0.375$

将以上计算结果填入最终表 x_1 所在的列向量的位置，得表 3.19 所示结果。

表 3.19

c_j			4	3	0	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
4	x'_1	4	1.25	0	0	1/4	0	
0	x_5	4	0.5	0	−2	1/2	1	
3	x_2	2	0.375	1	1/2	−1/8	0	
$c_j - Z_j$			0.375	0	−1.5	−1/8	0	

经过迭代，将 x'_1 所在列单位化成 $(1,0,0)^T$，可以得到最终表（见表 3.20）。

表 3.20

c_j			4	3	0	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
4	x'_1	3.2	1	0	0	0.2	0	
0	x_5	2.4	0	0	−2	0.4	1	
3	x_2	0.8	0	1	1/2	−0.2	0	
$c_j - Z_j$			0	0	−1.5	−0.2	0	

可知最优决策变成：生产甲产品3.2个单位，生产乙产品2.4个单位，生产丙产品0.8个单位。

注意，若是碰到填入新的数据之后，单纯形表上原问题和对偶问题都变成了非可行解时，则需要引进人工变量，重新求解。

以上只是讨论了模型中单个参数的变化，没有涉及多个变量同时变化的情形。读者在掌握上面的分析思路和技巧之后，可以自己设计一些多个参数同时变化的情形进行研究。

读者还可以对某些因素，每次变动一个固定比例，以测定其对最优方案影响的大小，进而对比判断出：哪些因素对最优方案影响大，需要重点估计；哪些对最优方案的影响较小，可以对其参数估计粗略一些。必要时，还可以结合各种变化情形发生的概率来分析各个因素的重要性等。

第五节　对偶线性规划的 WinQSB 应用

本节结合例题介绍 WinQSB 软件求解对偶线性规划的操作步骤及其应用。

例 3.7　已知线性规划：

Max $Z = x_1 + 2x_2 + 4x_3 + x_4$

$$s.t.\begin{cases}3x_1 + 9x_3 + 5x_4 \leqslant 15\\ 6x_1 + 4x_2 + x_3 + 7x_4 \leqslant 30\\ 4x_2 + 3x_3 + 4x_4 \leqslant 20\\ 5x_1 + 3x_2 + 8x_3 + 3x_4 \leqslant 40\\ x_j \geqslant 0, j = 1,2,3,4\end{cases}$$

要求:(1) 写出对偶线性规划,变量用 y 表示;

(2) 求原问题及对偶问题的最优解;

(3) 分别写出价值系数 c_j 及右端常数的最大允许变化范围;

(4) 目标函数系数改为 $C = (4,2,6,1)$,同时常数改为 $b = (20,40,20,40)$,求最优解;

(5) 删除第四个约束,同时删除第三个变量,求最优解;

(6) 增加一个变量 X_5,系数为$(c_5, a_{15}, a_{25}, a_{35}, a_{45}) = (6,5,4,2,3)$,求最优解。

解:启动线性规划与整数规划(Linear and Integer Programming),建立新问题,取名为"dual1"(读者可以任意取名),输入数据得到图 3.1,存盘。

Variable -->	X1	X2	X3	X4	Direction	R. H. S.
Maximize	1	2	4	1		
C1	3		9	5	<=	15
C2	6	4	1	7	<=	30
C3		4	3	4	<=	20
C4	5	3	8	3	<=	40
LowerBound	0	0	0	0		
UpperBound	M	M	M	M		
VariableType	Continuous	Continuous	Continuous	Continuous		

图 3.1

(1) 点击"Format" → "Switch to Dual Form",得到对偶问题的数据表,点击"Format" → "Switch to Normal Model Form",得到对偶模型,点击"Edit" → "Variable Name",分别修改变量名,得到以 y 为变量名的对偶模型,如图 3.2 所示。

dual1	
Minimize	15y1+30y2+20y3+40y4
	OBJ/Constraint/VariableType/Bound
Minimize	15y1+30y2+20y3+40y4
X1	3y1+6y2+5y4>=1
X2	4y2+4y3+3y4>=2
X3	9y1+1y2+3y3+8y4>=4
X4	5y1+7y2+4y3+3y4>=1
Integer:	
Binary:	
Unrestricted:	
y1	>=0, <=M
y2	>=0, <=M
y3	>=0, <=M
y4	>=0, <=M

图 3.2

(2) 再求一次对偶,返回到原问题,求解显示结果如图 3.3 所示,此时最优解为 X

$=(2,4.25,0)^T$,最优值 $Z=14.5$。表中影子价格(*Shadow Price*)对应列的数据就是对偶问题的最优解:$Y=(0.2833,0.025,0.475,0)$。

(3) 由图 3.3 可知:最后两列价值系数 $c_j(j=1,2,3,4)$ 的最大允许变化范围分别是[0.833 3,4.166 7],[1.333,5.777 8],[1.166 7,4.5],(-∞,3.491 7];右端常数 $b_i(i=1,2,3,4)$ 的最大允许变化范围分别是[5,27.471 9],[16.666 7,50],[0,33.333 3],[30.75,+∞)。

	Decision Variable	Solution Value	Unit Cost or Profit c[j]	Total Contribution	Reduced Cost	Basis Status	Allowable Min. c[j]	Allowable Max. c[j]
1	X1	2.0000	1.0000	2.0000	0	basic	0.8333	4.1667
2	X2	4.2500	2.0000	8.5000	0	basic	1.3333	5.7778
3	X3	1.0000	4.0000	4.0000	0	basic	1.1667	4.5000
4	X4	0	1.0000	0	-2.4917	at bound	-M	3.4917
	Objective	Function	(Max.) =	14.5000				
	Constraint	Left Hand Side	Direction	Right Hand Side	Slack or Surplus	Shadow Price	Allowable Min. RHS	Allowable Max. RHS
1	y1	15.0000	<=	15.0000	0	0.2833	5.0000	27.4719
2	y2	30.0000	<=	30.0000	0	0.0250	16.6667	50.0000
3	y3	20.0000	<=	20.0000	0	0.4750	0	33.3333
4	y4	30.7500	<=	40.0000	9.2500	0	30.7500	M

图 3.3

(4) 直接修改图3.1的数据,求解后得到最优解为 $X=(3.6667,4.25,1,0)^T$,最优值 $Z=29.1667$。

(5) 将数据修改回原问题,点击"Edit"→"Delete a Constraint",选择要删除的约束 C4。点击"Edit"→"Delete a Variable",选择要删除的变量 X3。得到如图 3.4 所示的模型,求解得到最优解:$X=(1.6667,5,0)^T$,最优值 $Z=11.6667$。

Variable ->	X1	X2	X4	Direction	R. H. S.
Maximize	1	2	1		
C1	3		5	<=	15
C2	6	4	7	<=	30
C3		4	4	<=	20
LowerBound	0	0	0		
UpperBound	M	M	M		
VariableType	Continuous	Continuous	Continuous		

图 3.4

(6) 调用原问题数据表,点击"Edit"→"Insert a Variable",选择变量名和变量插入的位置,如图 3.5,在显示的电子表格中输入数据(6,5,4,2,3),得到最优解:$X=(0,3.5,0,0,3)^T$,最优值 $Z=25$。

案例:Wyndor 玻璃制品公司的产品组合

1. Wyndor 玻璃制品公司的产品组合案例。

原问题:该公司打算在三个工厂生产两种新产品,有关数据如下表(见表 3.21)。

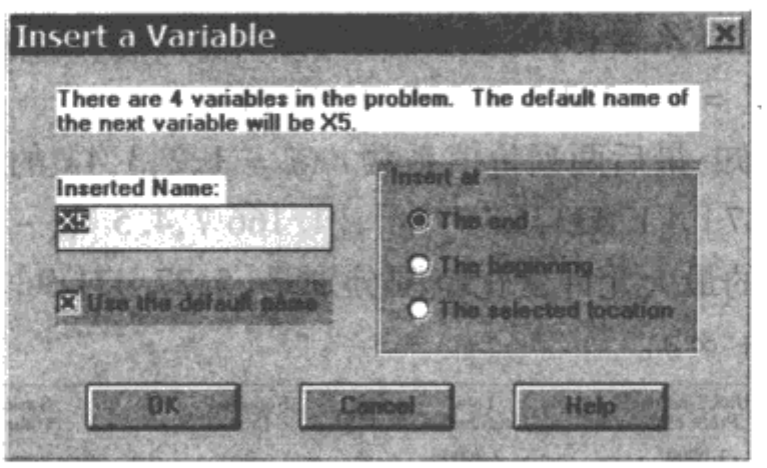


图 3.5

表 3.21

生产每个单位产品所需时间			
	门	窗	每周可得时间
工厂 1	1 小时	0	4 小时
工厂 2	0	2 小时	12 小时
工厂 3	3 小时	2 小时	18 小时
单位利润(美元)	300	500	

易求得的最优解是:每周生产门 2 个,窗 6 个,总利润为 3 600 美元。

对于研究者提出这个方案,管理层通过讨论后,提出以下问题:

(1) 如果新产品中,有一个产品的单位利润估计值不准确,将会发生怎样的情况?

(2) 如果某个工厂的可用时间发生变化,将会对结果产生什么影响?

以下我们逐一回答上面问题。

2. 最优域灵敏度分析

(回答问题 1:如果新产品中,有一个产品的单位利润估计值不准确,将会发生怎样的情况?)

模型中的许多参数,都是对实际数据的大致估计。灵敏度分析可以帮助我们知道,一个估计值需要精确到何种程度,才能避免得出错误的最优解,也即可以在多大程度上偏离实际值,而最优解不变。

问题:现在估计门的价格 Pd 是每个 300 美元,则该价格可以在多大程度上偏离实际值,而最优解不变?(或者说,可以偏离多大而不会产生错误的最优解)

解决:我们可以用 WinQSB 来做这个灵敏度分析。这是非常方便的。因为我们在得到最优解的同时,会出现一个对话框,如图 3.6 所示。(Excel 文件:chap2 的 sheet1)

点击敏感性报告即可得到一个参数的敏感性分析报告,它列出了模型各个参数的可以变化的范围,如表 3.22 所示。

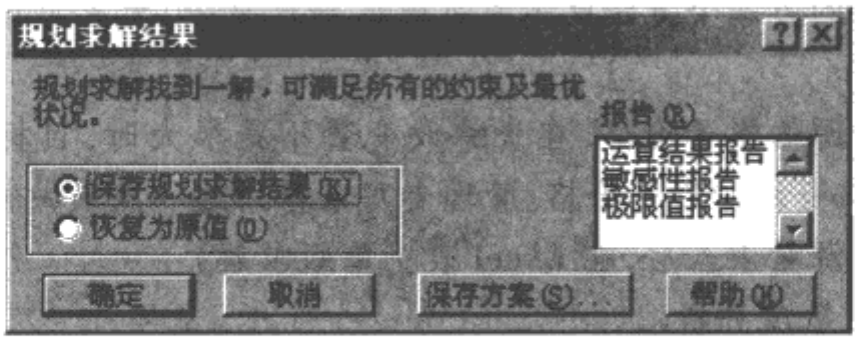


图 3.6

表 3.22

可变单元格

单元格	名字	终值	递减成本	目标式系数	允许的增量	允许的减量
C9	solution Doors	2	0	300	450	300
D9	solution Windows	6	0	500	1E + 30	300

约束

单元格	名字	终值	阴影价格	约束限制值	允许的增量	允许的减量
E5	plant1 Totals	2	0	4	1E + 30	2
E6	plant2 Totals	12	150	12	6	6
E7	plant3 Totals	18	100	18	6	6
E5	plant1 Totals	2	0	0	2	1E + 30
E6	plant2 Totals	12	0	0	12	1E + 30
E7	plant3 Totals	18	0	0	18	1E + 30

注意:1E + 30 表示 e30,Excel 用这个数来表示无穷大。

最优域:能保持最优解不变的 Pd(门的价格) 的变化范围,称为 Pd 的最优域。由上面的敏感性分析报告,我们可以方便地得到各个参数的最优域,这当然也包括 Pd 的最优域:(300 - 300,300 + 450),即(0,750)。

3. 右端值影子价格分析(资源的影子价格分析)

(回答问题 2:如果某个工厂的可用时间发生变化,将会对结果产生什么影响?)

资源量往往都是管理层决定的。建模并求解以后,管理者想要知道如果改变对资源量的决策,是否会对最终收益有影响。为了回答这一点,我们就需要做资源的影子价格分析,为管理者提供这方面的信息。

影子价格也可以如下定义。影子价格(Shadow Price):在给定线性规划模型的最优解和目标函数相应值的条件下,某资源增加一个单位,使得目标函数值发生一个改变量,这个量称为该资源的影子价格。

同样的,寻找影子价格较简易的方法还是使用电子表格。增加资源量一个单位,重新求解,就可以得到目标函数的新值,从而可计算出增量,这便是该资源的影子价格。

寻找影子价格的另一个方法就是看灵敏度分析报告。大家可以在该报告中找到影子价格这一项。

若知道了某资源的影子价格，在资源改变量不是很大时，目标函数值的变动量 = 资源变动量 × 该资源的影子价格。管理者可以借助这一点来做初步决策。(当然也可以将改变后的资源量录入后，在 Excel 表上重新求解)

4. 可行域的灵敏度分析

(还是在回答问题 2：如果某个工厂的可用时间发生变化，将会对结果产生什么影响?)

资源影子价格约束提供许多有价值的信息，但是，这些信息只有在资源变动不大的情况下才有效。现在我们就来讨论，为了保持影子价格有效，约束右端资源的变动范围。(这个范围称为可行域)

可以用电子表格来测试这个范围。方法是：逐步增加某资源的量(每次增加一个单位)，记录目标函数的值、目标值的变化量。其变化量保持不变的资源变化范围就是可行域，即资源影子价格有效的区域。越过这个区域，影子价格将发生变化，变化后的新影子价格又会有个有效区域。

不过我们大可不必用上面这个方法来测试可行域，因为还有一个更加简便的方法 —— 我们可以用敏感性分析报告的数据，做个简单计算便可得到可行域。

例如上面敏感性报告中：

表 3.23　　约束

单元格	名字	终值	阴影价格	约束限制值	允许的增量	允许的减量
E5	plant1 Totals	2	0	4	1E + 30	2
E6	plant2 Totals	12	150	12	6	6
E7	plant3 Totals	18	100	18	6	6
E5	plant1 Totals	2	0	0	2	1E + 30
E6	plant2 Totals	12	0	0	12	1E + 30
E7	plant3 Totals	18	0	0	18	1E + 30

工厂 1 的工时的影子价格可行域为：$(4-2,\infty)$。

习题

1. 判断

(1) 任何线性规划问题都存在并具有唯一的对偶问题。

(2) 若线性规划问题有无穷多最优解，则其对偶问题也一定有无穷多最优解。

(3) 已知 y_i^* 为线性规划问题的对偶问题的最优解，若 $y_i^* \geqslant 0$，说明最优生产计划中第 i 种资源已经完全耗尽。

(4) 若某种资源的影子价格等于 k,在其他条件不变的情形下,当该种资源增加 5 个单位时,相应的目标函数将增大 $5k$。

2. 已知某线性规划问题,用单纯形法计算得到中间某两步的计算表如下,试将表 3.24 中空白处数字填上。

表 3.24

c_j			3	5	4	0	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	x_6	
	x_2	3/8	2/3	1	0	1/3	0	0	
	x_5	14/3	-4/3	0	5	-2/3	1	0	
	x_6	20/3	5/3	0	-4	-2/3	0	1	
$c_j - Z_j$			-1/3	0	4	-5/3	0	-1/2	
	x_2					15/41	8/41	-10/41	
	x_3					-6/41	5/41	4/41	
	x_1					-2/41	-12/41	15/41	
$c_j - Z_j$									

3. 写出下列线性规划问题的对偶问题。

(1) $\text{Min } z = 3x_1 + 2x_2 - 3x_3 + 4x_4$

$$\text{s. t.} \begin{cases} x_1 - 2x_2 + 3x_3 - 4x_4 \leqslant 3 \\ x_2 + 3x_3 + 4x_4 \geqslant -5 \\ 2x_1 - 3x_2 - 7x_3 - 4x_4 = 2 \\ x_1 \geqslant 0, x_4 \geqslant 0 \end{cases}$$

(2) $\text{Min } z = -5x_1 - 6x_2 - 7x_3$

$$\text{s. t.} \begin{cases} -x_1 + 5x_2 - 3x_3 \geqslant 15 \\ -5x_1 - 6x_2 + 10x_3 \leqslant 20 \\ x_1 - x_2 - x_3 \leqslant 17 \\ x_1 \leqslant 0, x_2 \geqslant 0 \end{cases}$$

4. 已知下表为求解某线性规划问题的最终单纯形表(见表 3.25),表中 x_4, x_5 为松弛变量,问题约束为“$\leqslant$”形式。

表 3.25

c_j						0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
	x_3	5/2	0	1/2	1	1/2	0	
	x_1	5/2	1	1/2	0	-1/6	1/3	
$c_j - Z_j$			0	-4	0	-4	-2	

试写出原线性规划问题,并直接根据表中数据写出其对偶问题的最优解。

5. 已知线性规划问题:

$\text{Max } z = 3x_1 + x_2 + 4x_3$

$$s.t.\begin{cases}6x_1+3x_2+5x_3\leqslant 25\\3x_1+4x_2+5x_3\leqslant 20\\x_1,x_2,x_3\geqslant 0\end{cases}$$

用单纯形法求解时,最终表如表 3.26 所示。

表 3.26

c_j			3	1	4	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
3	x_1	5/3	1	-1/3	0	1/3	-1/3	
4	x_3	3	0	1	1	-1/5	2/5	
c_j-Z_j			0	-2	0	-1/5	-3/5	

(1) 直接写出原问题和对偶问题的最优解。

(2) 若问题中 x_2 的系数变为$(3,2,3)^T$,表中的解是否仍然为最优解?

(3) 若增加一个新的变量 x_6,其相应的系数为$(2,3,2)^T$,此时表中的最优解是否发生变化?

6. 已知线性规划问题:

$$\text{Max } z=10x_1+5x_2$$

$$\begin{cases}3x_1+4x_2\leqslant 9\\5x_1+2x_2\leqslant 8\\x_1,x_2\geqslant 0\end{cases}$$

用单纯形法求解的最终表如表 3.27 所示。

表 3.27

c_j			10	5	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	
5	x_2	3/2	0	1	5/14	-3/14	
10	x_1	1	1	0	-1/7	2/7	
c_j-Z_j			0	-2	-5/14	-25/14	

试用灵敏度分析方法判断:

(1) 目标系数 c_1,c_2 分别在什么范围变动时,最优解不变。

(2) 约束条件右端值 b_1,b_2 分别在什么范围内变动时,最优解保持不变。

7. 某厂生产甲、乙、丙三种产品,需消耗 A、B 两种原料。有关数据如表 3.28 所示。试回答下列问题。

表 3.28

	甲	乙	丙	原料拥有量
A	6	3	5	45
B	3	4	5	30
单位利润(单位:元)	4	1	5	

（1）建立线性模型，求使该厂获利最大的生产计划。

（2）若原料 A 市场紧缺，除已拥有量外一时无法购进；而原料 B 则可以很方便地在市场上购买，单价为 0.5。该厂是否应该购买，购进多少为宜？

（3）若由于某种原因，该厂决定暂停生产甲产品，试重新确定该厂的最优生产计划。

第四章　线性规划模型的扩展

引例:背包问题

有一背包,可装入8单位重量,10单位体积物品。四种物品的重量、体积和价值如表4.1所示。问每种物品各取多少件装入背包,使其中物品的总价值最高。

表4.1

物品	名称	重量	体积	价值
1	小说	3	3	20
2	摄像机	5	1	25
3	收音机	1	2	10
4	休闲食品	2	4	15

第一节　整数规划

变量取整数值的规划称为整数规划。所有变量都取整数的规划称为纯整数规划,部分变量取整数的规划称为混合整数规划。所有变量都取0、1两个值的规划称为0-1规划,部分变量取0、1两个值的规划称为0-1混合规划。

线性规划和整数规划的关系,我们用例4.1来说明。

例4.1　设线性规划问题为:

$$\text{Max } z = x_1 + 4x_2$$

$$s.t. \begin{cases} 14x_1 + 42x_2 \leqslant 196 \\ -x_1 + 2x_2 \leqslant 5 \\ x_1, x_2 \geqslant 0 \end{cases}$$

相应的整数规划问题为:

$$\text{Max } z = x_1 + 4x_2$$

$$s.t. \begin{cases} 14x_1 + 42x_2 \leqslant 196 \\ -x_1 + 2x_2 \leqslant 5 \\ x_1, x_2 \geqslant 0, \text{且 } x_1, x_2 \text{ 为非负整数} \end{cases}$$

线性规划的最优解为$(x_1, x_2) = (13/5, 19/5) = (2.6, 3.8)$,线性规划最优解的目标函数值为$z = 89/5 = 17.8$。

而相应的整数规划的可行解为$(x_1, x_2) = (5, 3)$,整数规划最优解的目标函数值为$z = 17$。

由例 4.1 可以看到,简单地将线性规划的非整数的最优解,用四舍五入或舍去尾数的办法得到整数解,一般情况下并不能得到整数规划的最优解。整数规划的求解方法要比线性规划复杂得多。

例 4.2 背包问题的数学模型。

解:x_i 为带第 i 种物品的数量:

$$\text{Max } Z = 20x_1 + 25x_2 + 10x_3 + 15x_4$$

$$s.t. \begin{cases} 3x_1 + 5x_2 + x_3 + 2x_4 \leqslant 8 \\ 3x_1 + x_2 + 2x_3 + 3x_4 \leqslant 10 \\ x_i \text{ 为 } 0,1。 \end{cases}$$

例 4.3 厂址选择问题。

在 4 个地点中选 3 处建生产同一产品的工厂,在这 5 个地点建厂所需投资、占用农田、建成以后的生产能力等数据如表 4.2 所示。

表 4.2

地点	1	2	3	4
所需投资(万元)	32	28	24	21
员工(人)	20	18	15	11
生产能力(吨)	70	55	42	28

现在有总投资 80 万元,所需 60 人。应如何选择厂址,使工厂建成后的总生产能力最大。

设四个 0 - 1 变量 x_1, x_2, x_3, x_4,其中:

$$x_i = \begin{cases} 0 & \text{表示在 } i \text{ 地不建厂} \\ 1 & \text{表示在 } i \text{ 地建厂} \end{cases} \quad i = 1,2,3,4$$

整数规划模型为:

$$\text{Max } z = 70x_1 + 55x_2 + 42x_3 + 28x_4$$

$$s.t. \begin{cases} 32x_1 + 28x_2 + 24x_3 + 21x_4 \leqslant 80 \\ 20x_1 + 18x_2 + 15x_3 + 11x_4 + 8x_5 \leqslant 60 \\ x_1 + x_2 + x_3 + x_4 = 3 \\ x_1, x_2, x_3, x_4, x_5 = 0,1 \end{cases}$$

这是一个 0 - 1 规划问题。

一、数学模型

整数规划的数学模型,一般形式可表示为:

$$\text{Max } Z = \sum_{i=1}^{n} C_i X_i$$

$$s.t. \begin{cases} \sum_{i=1}^{n} a_i x_i \leqslant b_i \\ x_i \geqslant 0 \\ x_i \text{ 为整数} \end{cases}$$

二、整数规划的求解方法

1. 割平面法

割平面法是1958年由Gomory提出的最早求解整数规划问题的一种方法，适用于求解纯整数规划问题。其基本思想是：首先放弃变量的整数要求，求得线性规划的最优解。如果最优解恰是一个整数解，则线性规划的最优解就是相应的整数规划的最优解。如果线性规划的最优解不是整数解，则要求构造一个新的约束（割平面），对线性规划问题的可行域进行切割，切除已经得到的线性规划的最优解，但保留原可行域中所有的整数解，求解新的线性规划问题。如果最优解仍不是整数解，再增加附加的约束将其切除，但仍保持最初可行域中所有的整数解。如此一直进行，直至得到一个整数的最优解为止。

例4.4 用割平面法求解以下整数规划：

$$\text{Min } z = 3x_1 + 4x_2$$

$$s.t. \begin{cases} 3x_1 + x_2 \geqslant 4 \\ x_1 + 2x_2 \geqslant 4 \\ x_1, x_2 \text{ 为整数} \end{cases}$$

先求相应的线性规划问题，得到最优单纯形表（见表4.3）。

表4.3

		x_1	x_2	x_3	x_4
x_1	4/5	1	0	-2/5	1/5
x_2	8/5	0	1	1/5	-3/5
	44/5	0	0	-2/5	-9/5

选择一个非整数的基变量，例如$x_2 = 8/5$，构造约束条件，其中：

$x_2 = 8/5 = 1 + 3/5$，

$x_3 = 1/5 = 0 + 1/5$

$x_4 = -3/5 = -1 + 2/5$，

附加的约束条件为：

$3/5 - (1/5x_3 + 2/5x_4) \leqslant 0$

即$1/5x_3 + 2/5x_4 \geqslant 3/5$。

将这个约束加到线性规划的最优单纯形表中，并增加一个松弛变量x_5，得到表4.4。

表4.4

		x_1	x_2	x_3	x_4	x_5
x_1	4/5	1	0	-2/5	1/5	0
x_2	8/5	0	1	1/5	-3/5	0
x_5	-3/5	0	0	[-1/5]	-2/5	1
	44/5	0	0	-2/5	-9/5	0

用对偶单纯形法，x_5 离基，x_3 进基，得到表 4.5。

表 4.5

		x_1	x_2	x_3	x_4	x_5
x_1	2	1	0	0	1	-2
x_2	1	0	1	0	-1	1
x_3	3	0	0	1	2	-5
	10	0	0	0	-1	-2

已获得整数的最优解。

为了得到切割约束 $1/5x_3 + 2/5x_4 \geqslant 3/5$，在 (x_1, x_2) 平面中的表达式，将其中的松弛变量 x_3, x_4 用 x_1, x_2 表示：

$$x_3 = 3x_1 + x_2 - 4,$$

$$x_4 = x_1 + x_2 - 4$$

代入切割约束，得到：

$$x_1 + x_2 \geqslant 3$$

这个切割的图解如图 4.1 所示。

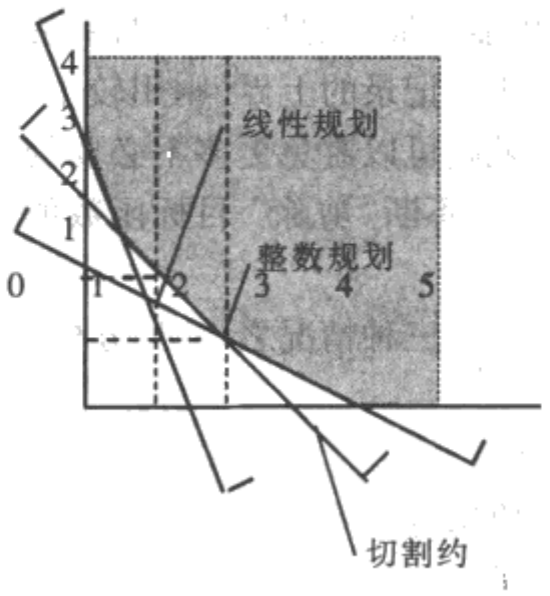


图 4.1

2. 分枝定界法

分枝定界法(Branch and Bound，简称 B&B) 的基本思想如下：

首先不考虑变量的整数约束，求解相应的线性规划问题，得到线性规划的最优解。设线性规划问题：

$$\text{Min}z = C^T X$$

$$s.t. \begin{cases} AX = B \\ X \geqslant 0 \end{cases}$$

如果这个最优解中所有的变量都是整数，则已经得到整数规划的最优解。如果其中某一个变量 X_r 不是整数，则在可行域中除去一块包含这个最优解但不包含任何整数解的区域 $Ir < X_r < Ir + 1$(其中 Ir 是变量 X_r 的整数部分)，线性规划的可行域被划分成不相交的两部分，分别以这两部分区域作为可行域，用原来的目标函数，构造两个子问题 Sub1 和 Sub2：

Subl

$$\text{Min} z = C^{\mathrm{T}}X$$

$$s.t.\begin{cases} AX = B \\ X_r \leqslant I_r \\ X \geqslant 0 \end{cases}$$

Sub2

$$\text{Min}\ z = C^{\mathrm{T}}X$$

$$s.t.\begin{cases} ZX = B \\ X_r \geqslant I_r \\ X \geqslant 0 \end{cases}$$

由于这两个子问题的可行域都是原线性规划问题可行域的子集，这两个子问题的最优解的目标函数值都不会比原线性规划问题的最优解的目标函数值更小。如果这两个问题的最优解仍不是整数解，则继续选择一个非整数的变量，继续将这个子问题分解为两个更下一级的子问题。这个过程称为“分枝(Branch)”。在分枝过程中，每一次分枝得到的子问题最优解的目标函数值，都大于或等于分枝前问题的最优解的目标函数值。

如果某一个子问题的最优解是整数解，就获得了一个整数可行解，这个子问题的目标函数值要记录下来，作为整数规划最优目标函数值的上界。如果某一个子问题的解还不是整数解，但这个非整数解的目标函数值已经超过这个上界，那么这个子问题就不必再进行分枝，因为继续分枝即使得到整数解，这个整数解的目标函数值必定要大于(或等于) 分枝以前问题的目标函数值，因而也大于(或等于) 已经获得的整数规划的目标函数值，因此不可能是最优的整数解。如果在分枝过程中得到新的整数解且该整数解的目标函数值小于已记录的上界，则用较小的整数解的目标函数值代替原来的上界。上界的值越小，就越可以避免更多不必要的分枝。这个确定整数解目标函数值上界、不断更新上界并且不断“剪除”目标函数值超过上界的分枝的过程，称为定界(Bound)。

当最低一层子问题出现以下三种情况之一时，分枝定界算法终止：

(1) 子问题无可行解；

(2) 子问题已获得整数解；

(3) 子问题的目标函数值超过上界。

例 4.5 用分枝定界法求解以下整数规划。

$$\text{Min}\ z = -2x_1 - 3x_2$$

$$s.t.\begin{cases} 5x_1 + 7x_2 \leqslant 35 \\ 4x_1 + 9x_2 \leqslant 36 \\ x_1, x_2 \geqslant 0 \\ x_1, x_2 \text{ 为整数} \end{cases}$$

先求得相应的线性规划的最优解，为：

$x_1 = 3\frac{12}{17}, x_2 = 2\frac{6}{17}, z = -14\frac{8}{17}$，取 $x_2 = 2\frac{6}{17}$ 分割可行域，得到以下两个子问题：

Sub－1

$$\begin{aligned}&\text{Min } z = -2x_1 - 3x_2\\ &\text{s.t.}\begin{cases}5x_1 + 7x_2 \leqslant 35\\ 4x_1 + 9x_2 \leqslant 36\\ x_2 \leqslant 2\\ x_1, x_2 \geqslant 0\end{cases}\end{aligned}$$

Sub－2

$$\begin{aligned}&\text{Min } z = -2x_1 - 3x_2\\ &\text{s.t.}\begin{cases}5x_1 + 7x_2 \leqslant 35\\ 4x_1 + 9x_2 \leqslant 36\\ x_2 \leqslant 3\\ x_1, x_2 \geqslant 0\end{cases}\end{aligned}$$

Sub1 的最优解为：

$x_1 = 4\frac{1}{5}, x_2 = 2, z = -14\frac{2}{5}$

取 $x_1 = 4\frac{1}{5}$ 对可行域进行分割，

得到子问题 Sub3 和 Sub4。

Sub2 的最优解为：

$x_1 = 2\frac{1}{4}, x_2 = 3, z = -13\frac{1}{2}$

$z = -13\frac{1}{2} > \bar{z} = -14$，停止分枝。

Sub－3

$$\begin{aligned}&\text{Min } z = -2x_1 - 3x_2\\ &\text{s.t.}\begin{cases}5x_1 + 7x_2 \leqslant 35\\ 4x_1 + 9x_2 \leqslant 36\\ x_2 \leqslant 2\\ x_1 \leqslant 4\\ x_1, x_2 \geqslant 0\end{cases}\end{aligned}$$

Sub－4

$$\begin{aligned}&\text{Min } z = -2x_1 - 3x_2\\ &\text{s.t.}\begin{cases}5x_1 + 7x_2 \leqslant 35\\ 4x_1 + 9x_2 \leqslant 36\\ x_2 \leqslant 3\\ x_1 \geqslant 5\\ x_1, x_2 \geqslant 0\end{cases}\end{aligned}$$

Sub－3 的最优解为：

$x_1 = 4, x_2 = 2, z = -14$

获得整数解，取得上界 $\bar{z} = -14$

停止分枝。

Sub－4 的最优解为：

$x_1 = 5, x_2 = 1\frac{3}{7}, z = -14\frac{2}{7}$

取 $x_2 = 1\frac{3}{7}$ 对可行域进行分割

得到子问题 Sub－5 和 Sub－6

Sub－5

$$\begin{aligned}&\text{Min } z = -2x_1 - 3x_2\\ &\text{s.t.}\begin{cases}5x_1 + 7x_2 \leqslant 35\\ 4x_1 + 9x_2 \leqslant 36\\ x_2 \leqslant 2\\ x_1 \geqslant 5\\ x_2 \leqslant 1\\ x_1, x_2 \geqslant 0\end{cases}\end{aligned}$$

Sub－6

$$\begin{aligned}&\text{Min } z = -2x_1 - 3x_2\\ &\text{s.t.}\begin{cases}5x_1 + 7x_2 \leqslant 35\\ 4x_1 + 9x_2 \leqslant 36\\ x_2 \leqslant 2\\ x_1 \geqslant 5\\ x_2 \geqslant 2\\ x_1, x_2 \geqslant 0\end{cases}\end{aligned}$$

Sub－5 的最优解为：

$x_1 = 5\frac{3}{5}, x_2 = 1, z = -14\frac{1}{5}$

取 $x_1 = 5\frac{3}{5}$ 对可行域进行分割，

Sub－6 的可行域是空集，停止分枝。

得到子问题 Sub - 7 和 Sub - 8。

Sub - 7

$$\begin{aligned}&\text{Min } z = -2x_1 - 3x_2\\&\text{s.t.}\begin{cases}5x_1 + 7x_2 \leqslant 35\\4x_1 + 9x_2 \leqslant 36\\x_2 \leqslant 2\\x_1 \geqslant 5\\x_2 \leqslant 1\\x_1 \leqslant 5\\x_1, x_2 \geqslant 0\end{cases}\end{aligned}$$

Sub - 8

$$\begin{aligned}&\text{Min } z = -2x_1 - 3x_2\\&\text{s.t.}\begin{cases}5x_1 + 7x_2 \leqslant 35\\4x_1 + 9x_2 \leqslant 36\\x_2 \leqslant 2\\x_1 \geqslant 5\\x_2 \leqslant 1\\x_1 \geqslant 6\\x_1, x_2 \geqslant 0\end{cases}\end{aligned}$$

Sub - 7 的最优解为：

$x_1 = 5, x_2 = 1, z = -13$

获得整数解，停止分枝。

由于 $z = -13 > \bar{z} = -14$，
上界仍保持为 $\bar{z} = -14$。

Sub - 8 的最优解为：

$x_1 = 6, x_2 = \frac{5}{7}, z = -14\frac{3}{7}$

取 $x_2 = \frac{5}{7}$ 对可行域进行分割，

得到子问题 Sub - 9 和 Sub - 10。

Sub - 9

$$\begin{aligned}&\text{Min } z = -2x_1 - 3x_2\\&\text{s.t.}\begin{cases}5x_1 + 7x_2 \leqslant 35\\4x_1 + 9x_2 \leqslant 36\\x_2 \geqslant 3\\x_1 \geqslant 5\\x_2 \geqslant 2\\x_1 \geqslant 6\\x_2 \leqslant 0\\x_1, x_2 \geqslant 0\end{cases}\end{aligned}$$

Sub - 10

$$\begin{aligned}&\text{Min } z = -2x_1 - 3x_2\\&\text{s.t.}\begin{cases}5x_1 + 7x_2 \leqslant 35\\4x_1 + 9x_2 \leqslant 36\\x_2 \geqslant 3\\x_1 \geqslant 5\\x_2 \geqslant 2\\x_1 \geqslant 6\\x_2 \geqslant 1\\x_1, x_2 \geqslant 0\end{cases}\end{aligned}$$

Sub - 9 的最优解为：

$x_1 = 7, x_2 = 0, z = -14$，

得整数解，$z = -14$，上界仍为：

$\bar{z} = -14$。

Sub - 10 的可行域是空集，

停止分枝。

至此已将所有可能分解的子问题都已分解到底，最后得到两个目标函数值相等的最优整数解：$(x_1, x_2) = (4,0)$ 和 $(x_1, x_2) = (0,7)$，他们的目标函数值都是 - 14。

以上的搜索最优解过程可以用一个树状图表示（如图 4.2 所示），由分枝定界算法可以知道，这个搜索树是一个二叉树。

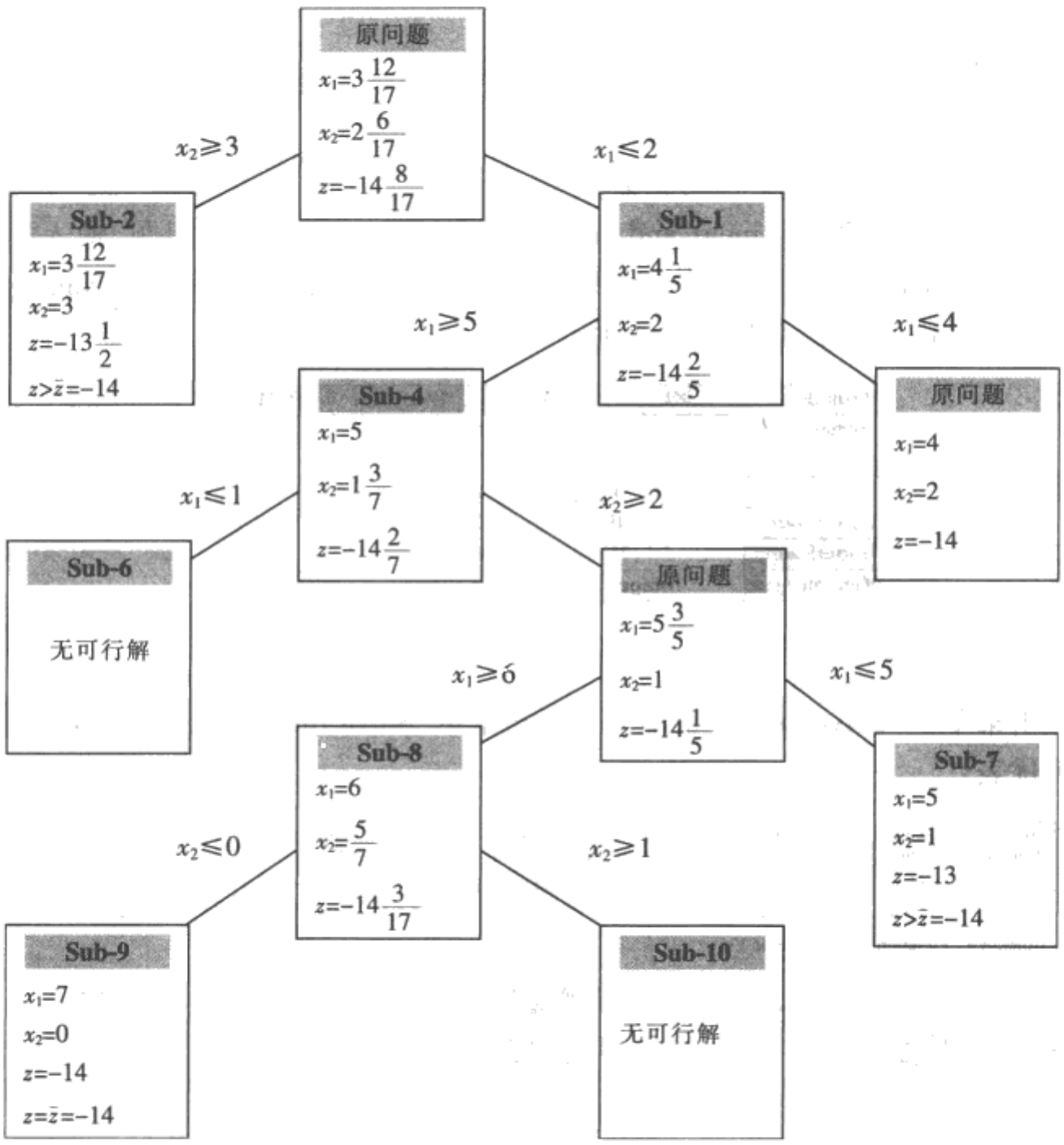


图 4.2

分枝定界法注意事项：

(1) 分枝变量的选择原则。

① 按目标函数系数：选系数绝对值最大者变量先分枝。对目标值升降影响最大。

② 选与整数值相差最大的非整数变量先分枝。

③ 按使用者经验，对各整数变量排定重要性的优先顺序。

(2) 分枝节点的选择方法。

① 深探法(后进先出法)：最后打开的节点最先选，尽快找到整数解。整数解质量可能不高。

② 广探法：选目标函数当前最大值节点，找到的整数解质量高，速度慢。

三、整数规划的 WinQSB 应用

运用 WinQSB 软件求解线性整数规划仍然是调用子程序 Linear and Integer Programming，操作时改变变量类型即可。下面以例 4.6 来说明这个应用。

例 4.6 用 WinQSB 软件求解以下整数规划问题：

$$\text{Max } z = x_1 + 4x_2$$

$$s.t. \begin{cases} 14x_1 + 42x_2 \leqslant 196 \\ -x_1 + 2x_2 \leqslant 5 \\ x_1, x_2 \geqslant 0, \text{且 } x_1, x_2 \text{ 为整数。} \end{cases}$$

解:首先启动子程序 Linear and Integer Programming,建立新问题,输入类似图4.3的选项。本例中,变量数等于2,约束数等于2,变量类型选非负整数(Nonnegative Integer)。然后输入数据,如图4.3所示。

Variable ->	X1	X2	Direction	R. H. S.
Maximize	1	4		
C1	14	42	<=	196
C2	-1	2	<=	5
LowerBound	0	0		
UpperBound	M	M		
VariableType	Integer	Integer		

图4.3

点击菜单栏"Solve and Analyze"的下拉菜单"Solve the Problem"得到图4.4所示的最优表。

	Decision Variable	Solution Value	Unit Cost or Profit c(j)	Total Contribution	Reduced Cost	Basis Status
1	X1	5.0000	1.0000	5.0000	0	basic
2	X2	3.0000	4.0000	12.0000	0	basic
	Objective	Function	(Max.) =	17.0000		
	Constraint	Left Hand Side	Direction	Right Hand Side	Slack or Surplus	Shadow Price
1	C1	196.0000	<=	196.0000	0	0.0714
2	C2	1.0000	<=	5.0000	4.0000	0

图4.4

最优解为:$x_1 = 5, x_2 = 3, x_3 = 0, x_4 = 4, x_5 = 0$,最优值为 $z = 17$。

其他类型的整数规划问题只要改变变量类型即可。

第二节　目标规划

例4.7　某厂生产Ⅰ,Ⅱ两种产品,利润、需要的原材料和设备如表4.6所示。求获利最大的生产方案。

表4.6

	Ⅰ	Ⅱ	拥有量
原材料 设备	2 1	1 2	11 10
利润	8	10	

这是一个单目标规划问题,我们建立如下的线性规划模型:

$$\text{Max } Z = 8x_1 + 10x_2$$

$$s.t.\begin{cases}2x_1 + x_2 \leqslant 11\\ x_1 + 2x_2 \leqslant 10\\ x_1, x_2 \geqslant 0\end{cases}$$

最优方案为：

$x_1 = 4, x_2 = 3, Z = 62$

实际上工厂做决策时要考虑到市场等一系列其他条件或目标，如：

（1）根据市场信息，产品 Ⅰ 的销售有下降，因此考虑产品 Ⅱ 的产量不低于产品 Ⅰ；

（2）尽可能利用设备，但不希望加班；

（3）应尽可能达到计划利润指标 ——56 元。

这样该产品的生产方案决策，成为多目标决策问题，需要通过目标规划方法来解决，其结果是满意解（可以是不到一点，也可以超出一点，但要尽可能地接近目标值）。

因此，目标规划解决需要考虑多个目标的决策问题，它在满足现有的一组约束的条件下，求出尽可能接近理想值的解。我们对目标规划与线性规划进行一个比较，如表 4.7 所示。

表 4.7

	线性规划	目标规划
问题	单目标	多目标
约束	硬约束、矛盾	划分等级
求解	绝对最优	实际满意

一、数学模型

首先，我们给出目标规划数学模型有关概念，设决策变量为：$x_1, x_2, \cdots$

偏差：实际决策值与目标值之间的差异；

正偏差：决策值超过目标值的部分，记为 d^+；

负偏差：决策值低于目标值的部分，记为 d^-；

且：$d^+ + d^- = 0$。

绝对约束：严格满足的等式或不等式约束；

目标约束：把约束右端项看成是要追求的目标值，在达到此目标时允许有正负偏差，线性规划问题的目标函数，在给定目标值和加入正、负偏差后可变换为目标约束，也可将绝对约束变换为目标约束。

例 4.7 中：$Z = 8x_1 + 10x_2$ 可变换为：$8x_1 + 10x_2 + d_1^- - d_1^+ = 56$

$2x_1 + x_2 \leqslant 11$ 可变换为：$2x_1 + x_2 + d_2^- - d_2^+ = 11$

优先等级与权系数：要达到的多个目标之间有主次、轻重、缓急之分，因此各目标之间有优先等级。凡第一位要达到的目标赋予等级系数 p_1，次位的赋予等级系数 p_2，以此类推；并规定若 $p_k \geqslant p_{k+1}$，则 p_k 比 p_{k+1} 有更大的优先权。相同等级的以不同的权系数 ω 加以区别。

目标规划的目标函数：目标规划的目标函数是按各目标约束的正、负偏差变量和赋予相应的优先因子而构造的。当每一目标值确定后，决策者的要求是尽可能缩小偏离目标值，因此目标规划的目标函数只能是：

$\text{Min } Z = f(d^+, d^-)$

(1) 要求恰好达到目标值(正、负偏差都要尽可能地小)，这时：

$\text{Min } Z = f(d^+ + d^-)$

(2) 要求不超过目标值(允许达不到，正偏差要尽可能地小)：

$\text{Min } Z = f(d^+)$

(3) 要求不低于(至少达到) 目标值：

$\text{Min } Z = f(d^-)$

例 4.8 在原材料供应严格限制的条件下，可对引例问题建立如下的目标规划模型：

(1) 产品 Ⅰ 的产量不超过产品 Ⅱ；优先因子 p_1；

(2) 尽可能利用设备，但不希望加班；优先因子 p_2；

(3) 应尽可能达到(超过) 利润指标 56 元；优先因子 p_3。

目标函数：$\text{Min } Z = p_1 d_1^+ + p_2(d_2^- + d_2^+) + p_3 d_3^-$

$$
s.t.\begin{cases} 2x_1 + x_2 \leqslant 11 \\ x_1 - x_2 + d_1^- - d_1^+ = 0 & (\leqslant 0) \\ x_1 + 2x_2 + d_2^- - d_2^+ = 10 & (=10) \\ 8x_1 + 10x_2 + d_3^- - d_3^+ = 56 & (\geqslant 56) \\ x_1, x_2, d_1^-, d_1^+, d_2^-, d_2^+, d_3^-, d_3^+ \geqslant 0 \end{cases}
$$

二、目标规划的求解方法

目标规划的求解方法与线性规划问题求解相同，采用单纯型方法，只是变量增加。其两个决策变量也可用图解法求解，但比线性规划问题求解要复杂得多。

步骤：

(1) 先作绝对约束；

(2) 再作 $d = 0$ 时的目标约束；

(3) 标出 d^+, d^- 的方向；

(4) 按优先的次序，逐个使目标函数中的极小偏差变量尽可能取 0，缩小可行域，找出问题解。

例 4.8 的图解，如图 4.7 所示。

$G(2,4)$，$D(10/3, 10/3)$，线段 GD 上的任一点都是解。

三、目标规划的 WinQSB 应用

目标规划的 WinQSB 软件求解需要调用子程序 Goal Programming(GP)，该程序可以求解线性目标规划、整数线性目标规划和多目标规划问题，下面我们通过例 4.9 来介绍具体的计算操作步骤。

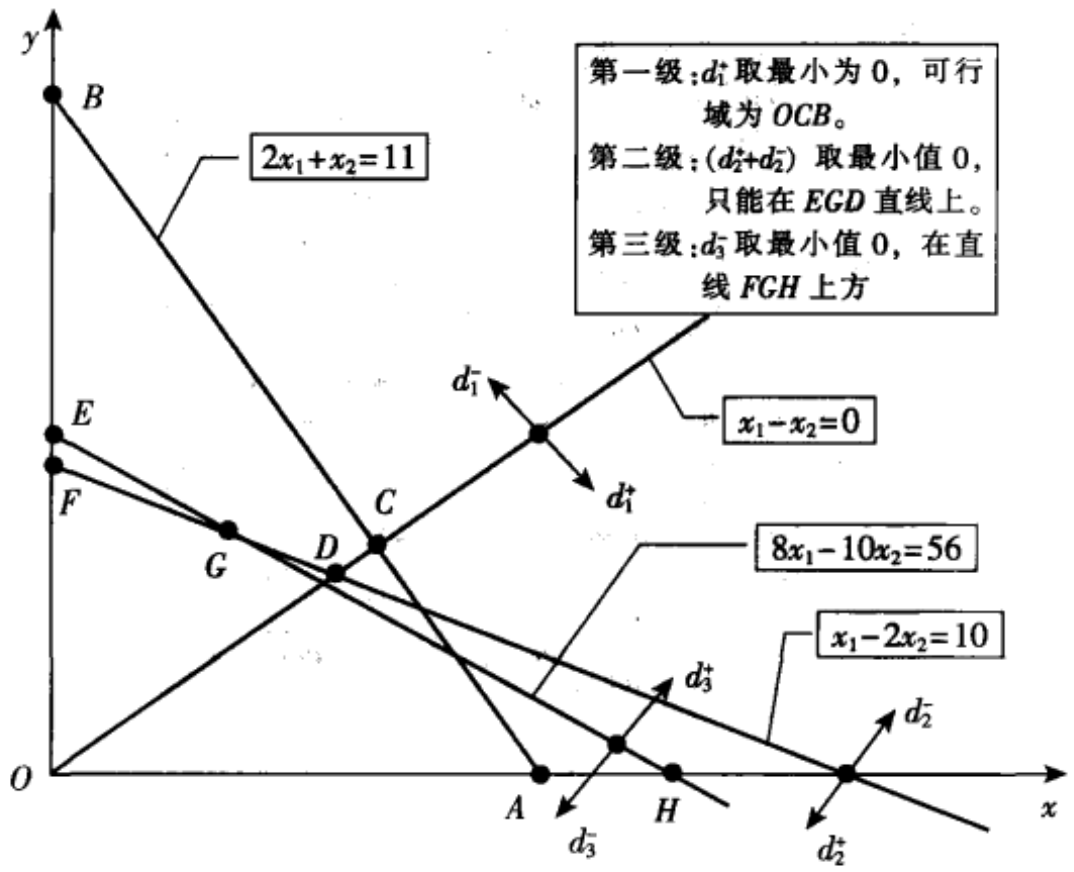


图 4.7

例 4.9 某电视机厂装配黑白和彩色两种电视机，每装配一台电视机需占用装配线 1 小时，装配线每周计划开动 40 小时。预计：市场每周彩电的销量是 24 台；黑白电视机的销量是 30 台。该厂确定的目标为：

第一优先级：充分利用装配线，每周计划开动不低于 40 小时。

第二优先级：允许装配线加班，但加班时间每周尽量不超过 10 小时。

第三优先级：装配电视机的数量尽量满足市场需要，因彩电利润高于黑白电视机，取其权系数为 2/3。

试建立目标规划模型并运用 WinQSB 软件求解。

解：设 X_1, X_2 分别为彩色和黑白电视机的产量：

目标函数：$\text{Min } Z = p_1 d_1^- + p_2 d_2^+ + p_3(2d_3^- + d_4^-)$

$$s.t.\begin{cases} x_1 - x_2 + d_1^- - d_1^+ = 40 & (\geqslant 40) \\ x_1 + x_2 + d_2^- - d_2^+ = 50 & (\leqslant 50) \\ x_1 + d_3^- - d_3^+ = 24 & (\geqslant 24) \\ x_2 + d_4^- - d_4^+ = 30 & (\geqslant 30) \\ x_1, x_2, d_1^-, d_1^+, d_2^-, d_2^+, d_3^-, d_3^+, d_4^-, d_4^+ \geqslant 0 \end{cases}$$

首先启动程序：点击“开始”→“程序”→“WinQSB”→“Goal Programming”；然后建立新问题，如图 4.8 所示，输入标题 gp1、目标数 3、变量数 10 及约束数 4。目标数等于优先级别个数，变量数等于决策变量数加偏差变量数，其他选项同线性规划。

接着，输入数据，系统显示变量为 $x_j(j = 1, 2, \cdots, 10)$，为了便于观察，将变量重新命名为 $x_1, x_2, d_i^-, d_i^+(i = 2, \cdots, 3)$，输入数据后得图 4.9。点击菜单栏“Solve and Analyze”的下拉菜单“Solve the Problem”就可以得到结果，最优解有两个，一个是(24,26)，一个是(20,30)。

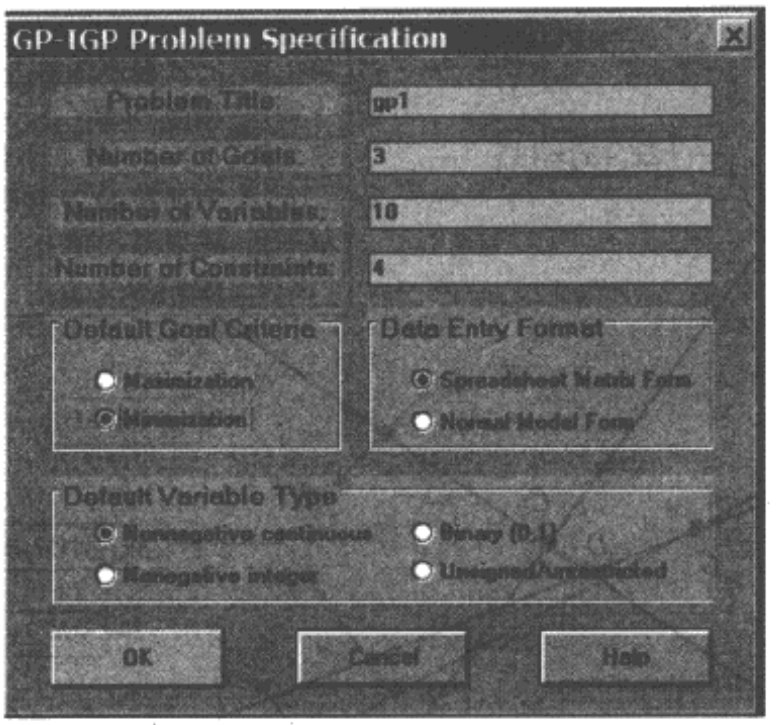


图 4.8

Variable ->	x1	x2	d1-	d1+	d2-	d2+	d3-	d3+	d4-	d4+	Direction	R. H. S.
			1									
						1						
							2		1			
C1	1	1	1	-1							=	40
C2	1	1			1	-1					=	50
C3	1						1	-1			=	24
C4		1							1	-1	=	30
Lower Bound	0	0	0	0	0	0	0	0	0	0		
Upper Bound	M	M	M	M	M	M	M	M	M	M		
Variable Type	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous	Continuous		

图 4.9

第三节　动态规划

例 4.10　一家医院有 5 位护士要分派到 3 个病房，不同数量的护士在每个病房的作用如表 4.8 所示。请思考：如何分配才能使总作用最大？能否用线性规划来做？

表 4.8

护士人数	在病房中的作用		
	X	Y	Z
0	0	0	0
1	4	4	6
2	6	8	9
3	10	13	13
4	14	15	16
5	16	16	17

求解思路：护士分到 X 后就不能再分到 Y 与 Z，因此可考虑先分配给 X，再分配给 Y，最后分配给 Z。每一阶段可分配数是前一阶段做决策时的可分配数减去实际分配数。可转换为求图 4.10 中 S 到 F 的最长距离。

求解结果为：X 分 1 个，Y 分 3 个，Z 分 1 个，总作用为 23。

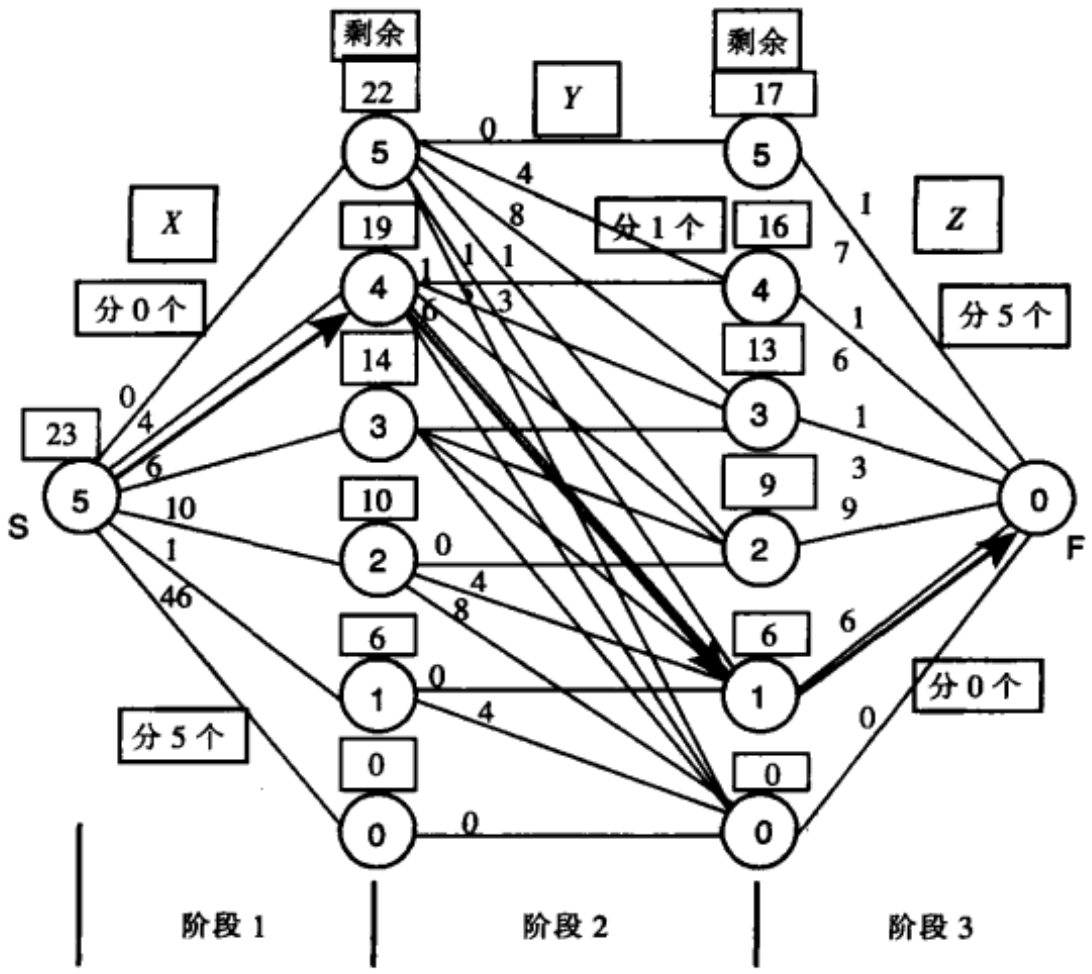


图 4.10

一、数学模型

在生产实际过程中，我们可能会遇到如下一类多阶段决策问题（如图 4.11 所示）：将过程按时间、空间等标志分为若干个阶段；每一个阶段都需要做出决策；阶段决策依赖于当前状态，又影响以后发展。

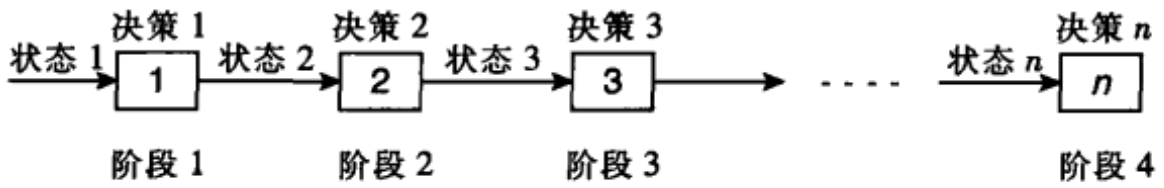


图 4.11

用动态规划求解上述多阶段决策问题，有以下一些基本概念：

阶段 k：表示决策顺序的离散量，阶段可按时间或空间划分。

状态 S_k：能确定地表示决策过程当前特征的量。状态可以是数量也可以是字符。数量状态可以是连续的，也可以是离散的。

状态变量 X_k：表示每一状态可以取不同值的变量。

决策 d_k：从某一状态向下一状态过渡时所做的选择。决策是所在状态变量的函数，记为 $d_k(X_k)$。

决策允许集合 $D_k(x_k)$：在状态 X_k 下，允许采取决策的全体。

状态转移方程 $X_{k+1} = T(X_k, d_k)$：某一状态以及该状态下的决策，与下一状态之间的函数关系。

最优化原理:最佳路径中任一状态(中间某点)到最终状态(最终点)的路径也是该状态到最终状态一切可能中的最短路径。

动态规划问题求解的一般方法:动态规划虽然没有标准模型,没有唯一确定的解法,但可以考虑从以下四个方面进行求解:

(1) 将完整的问题划分成若干个阶段;

(2) 明确最终要达到的目的;

(3) 从最终结果倒推,逐个阶段地做出决策;

(4) 回到出发点,做出最后一个决策,再向前推可得到每一阶段的最优决策。

二、应用举例

例 4.11 图 4.12 表示从起点 A 到终点 E 之间各点间的距离。求 A 到 E 的最短路径。

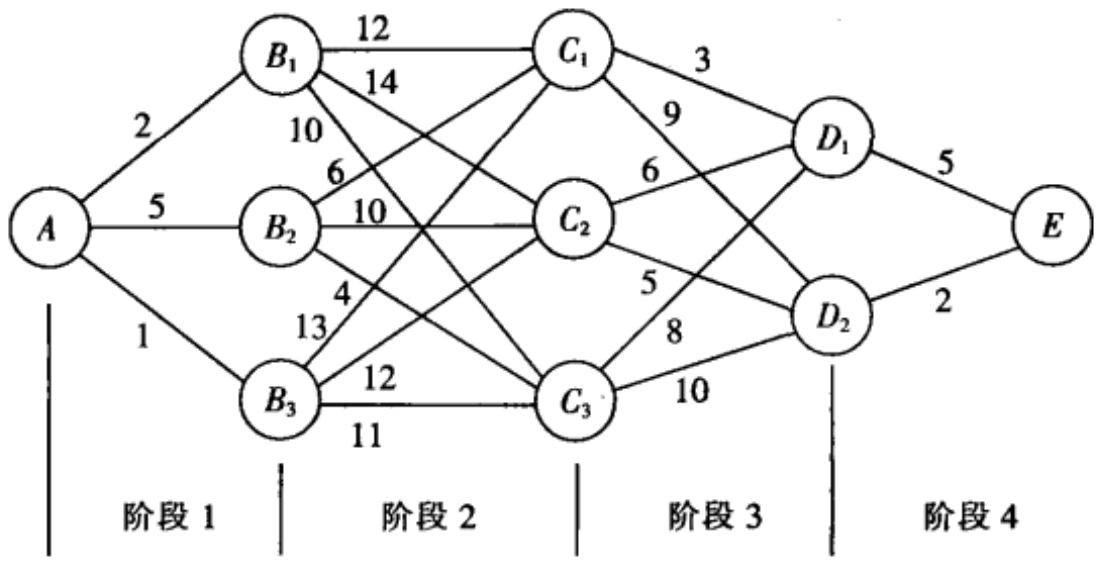


图 4.12

(1) 每一个阶段都至少有一个起始点 —— 初始状态(如 B_i);

(2) 每一个阶段都需要做一个选择 —— 决策本阶段由初始状态应演变到下一个阶段的哪一个起始点(本阶段终点);

(3) 每一个阶段的决策不仅影响到本阶段的效果,还影响到下一阶段的初始状态,从而也影响到此后的演变过程;

(4) 每一个阶段的决策不能只从这一阶段本身考虑,要考虑整个过程的最优效果。

求从 A 到 E 的最短路径问题,可以转化为三个性质完全相同,但规模较小的子问题,即分别从 B_1、B_2、B_3 到 E 的最短路径问题。

记:从 $B_i(i = 1, 2, 3)$ 到 E 的最短路径为 $S(B_i)$,则从 A 到 E 的最短距离 $S(A)$ 可以表示为:

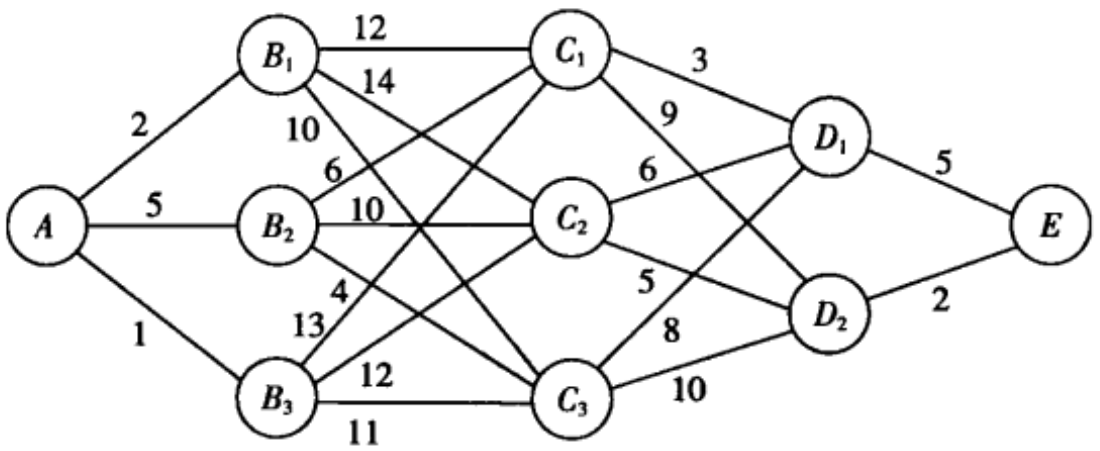
$$S(A) = \text{Min}\begin{Bmatrix} AB_1 + S(B_1) \\ AB_2 + S(B_2) \\ AB_3 + S(B_3) \end{Bmatrix} = \text{Min}\begin{Bmatrix} 2 + S(B_1) \\ 5 + S(B_2) \\ 1 + S(B_3) \end{Bmatrix}$$

状态转移函数在本例中为距离和最小。

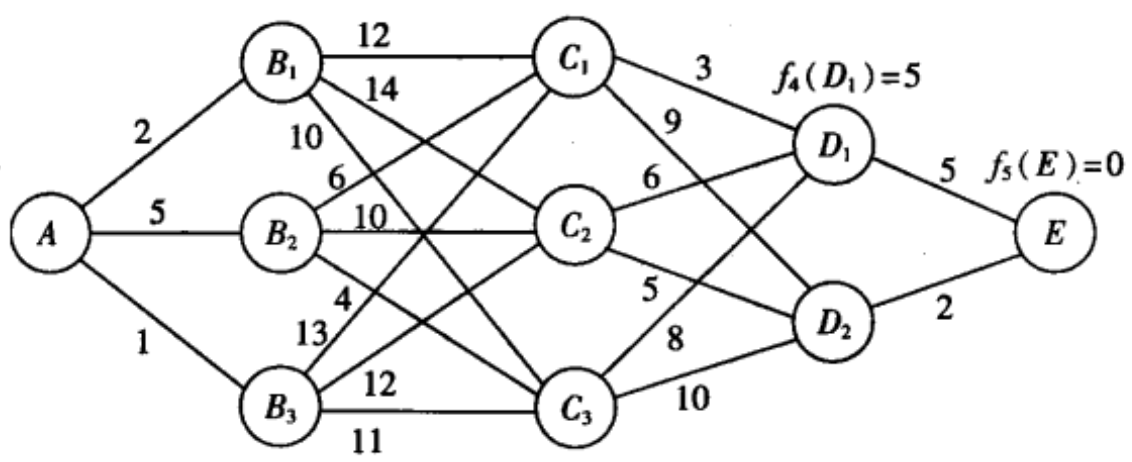
同样，计算 $S(B_1)$ 又可以归结为性质完全相同但规模更小的问题，即分别求 C_1，C_2，C_3 到 E 的最短路径问题 $S(C_i)$（$i = 1, 2, 3$），而求 $S(C_i)$ 又可以归结为求 $S(D_1)$ 和 $S(D_2)$ 这两个子问题。从图4.12可以看出，在这个问题中，$S(D_1)$ 和 $S(D_2)$ 是已知的，它们分别是：

$S(D_1) = 5, S(D_2) = 2$

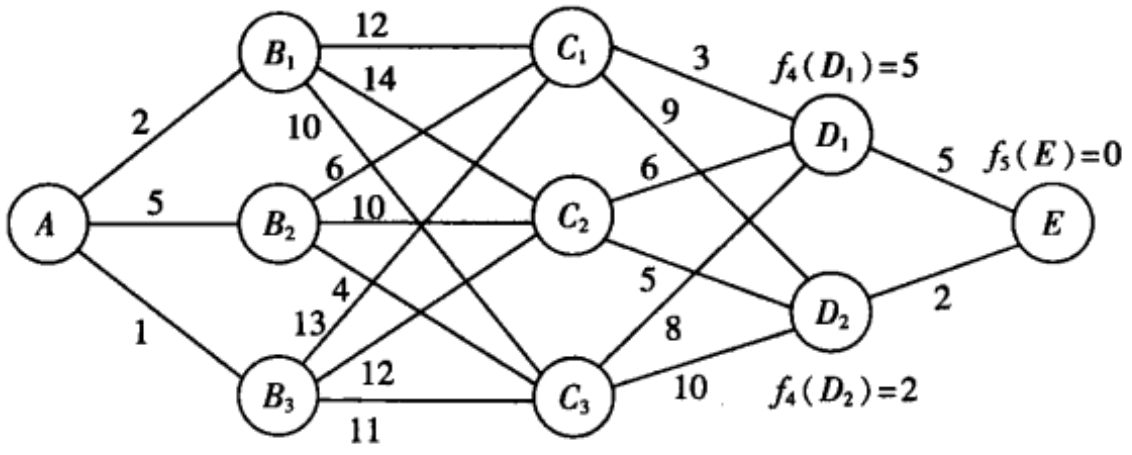
因而，可以从这两个值开始，逆向递归计算 $S(A)$ 的值。如下所示：



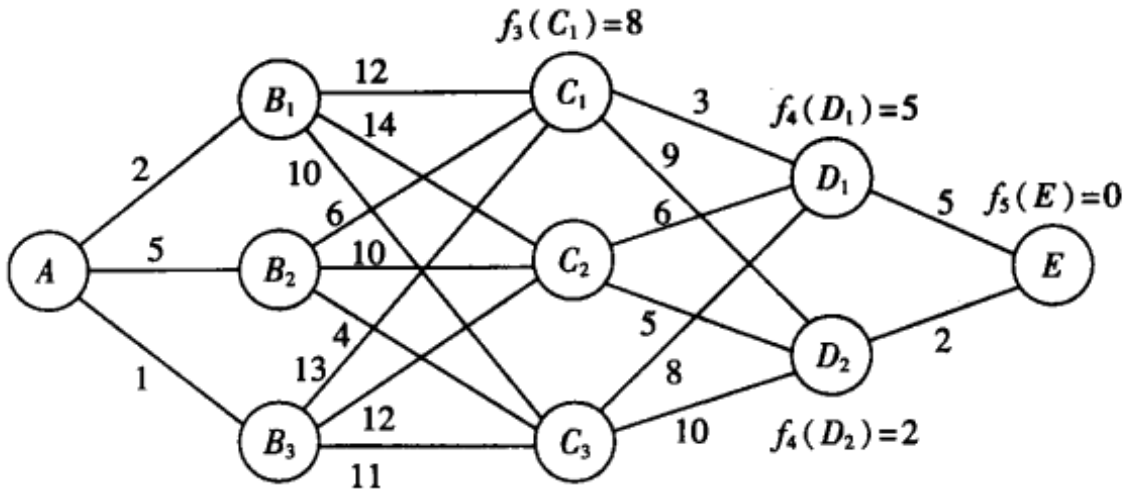
$f_5(E) = 0$



$f_4(D_1) = d(D_1 \rightarrow E) + f_5(E) = 5 + 0 = 5$

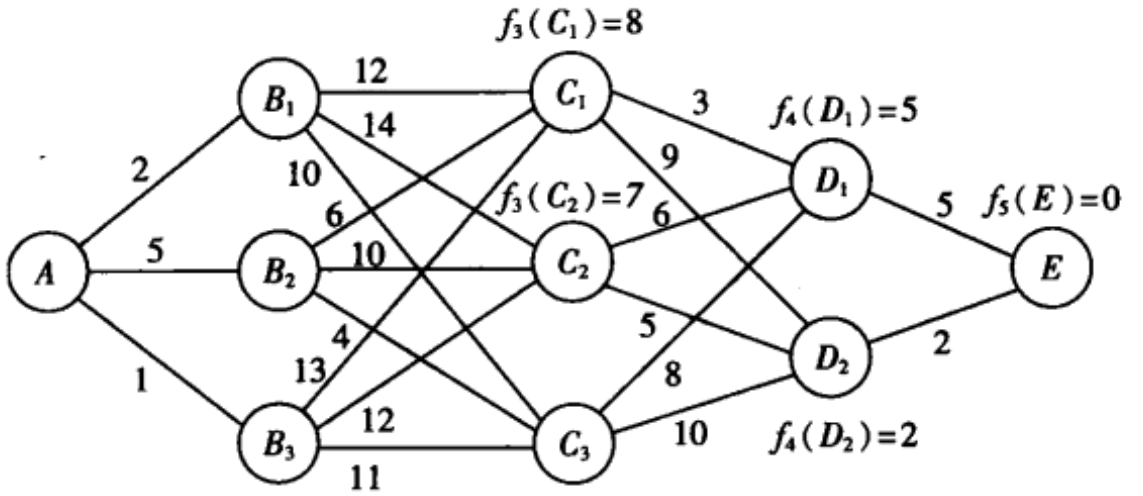


$f_4(D_2) = d(D_2 \rightarrow E) + f_2(E) = 2 + 0 = 2$



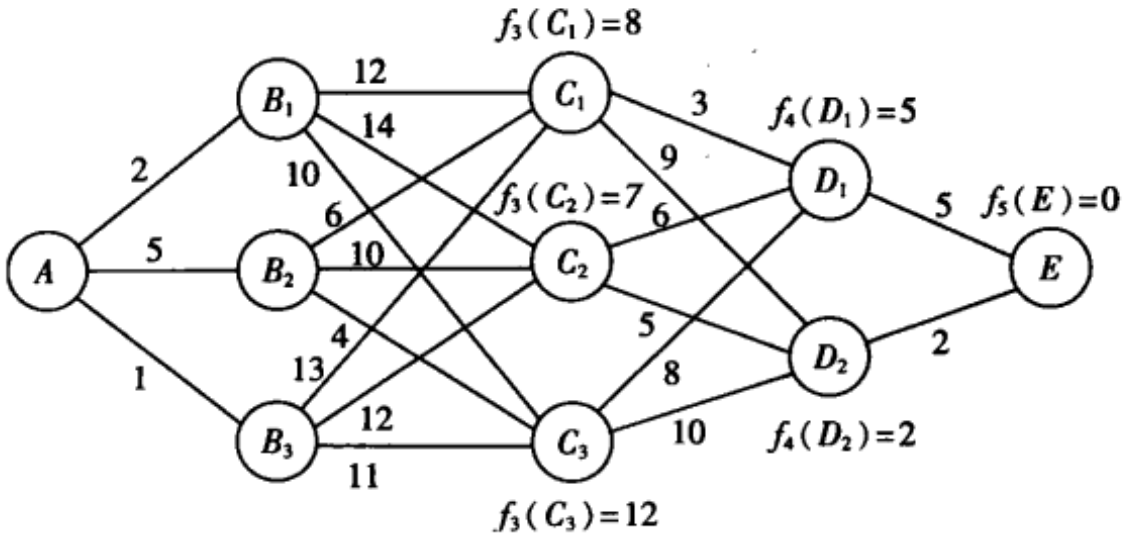
$$f_3(C_1) = \text{Min}\begin{Bmatrix}(C_1,D_1) + f_4(D_1)\\(C_1,D_2) + f_4(D_2)\end{Bmatrix}$$

$$= \text{Min}\begin{Bmatrix}3+5\\9+2\end{Bmatrix} = \text{Min}\begin{Bmatrix}8\\11\end{Bmatrix} = 8 \quad \text{最优决策 } C_1 \rightarrow D_1$$



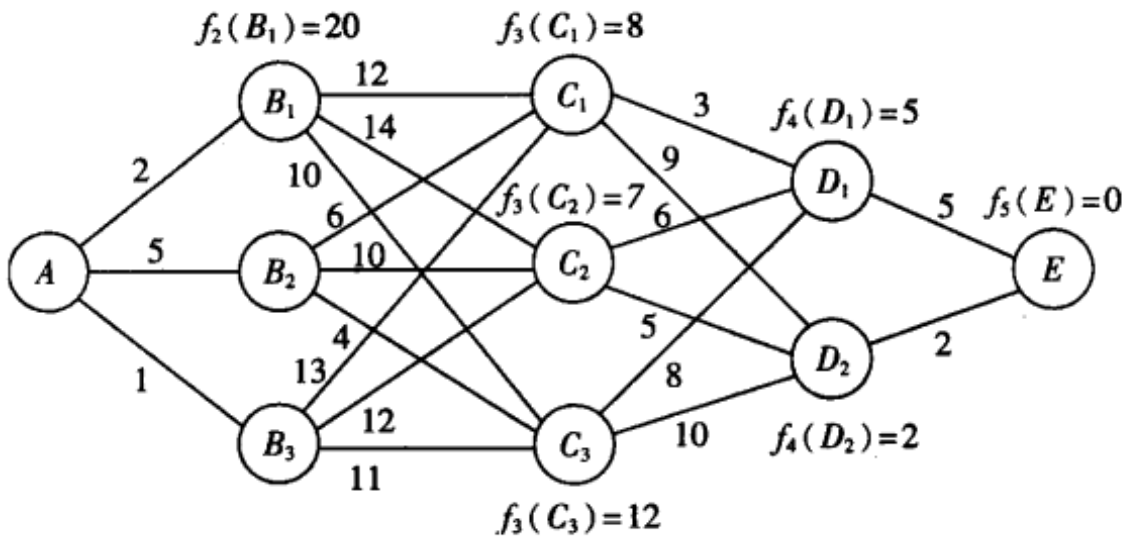
$$f_3(C_2) = \text{Min}\begin{Bmatrix}(C_2,D_1) + f_4(D_1)\\(C_2,D_2) + f_4(D_2)\end{Bmatrix}$$

$$= \text{Min}\begin{Bmatrix}6+5\\5+2\end{Bmatrix} = \text{Min}\begin{Bmatrix}11\\7\end{Bmatrix} = 7 \quad \text{最优决策 } C_2 \rightarrow D_2$$



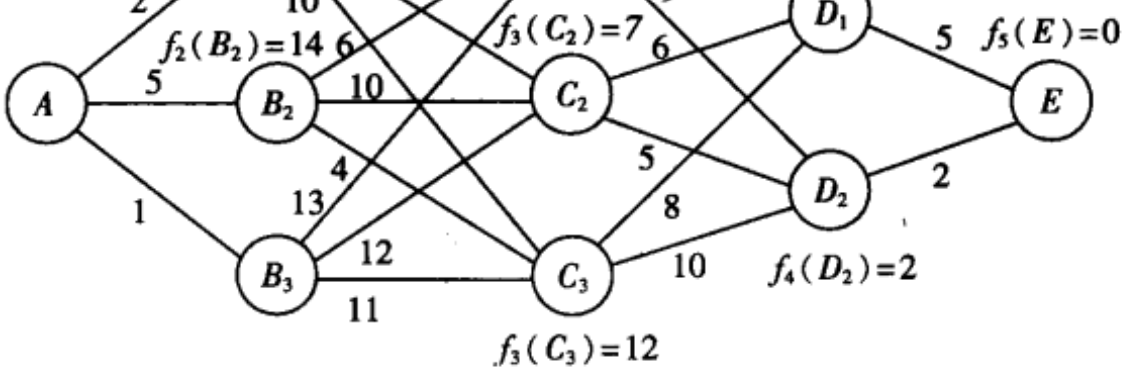
$$f_3(C_3) = \text{Min}\begin{Bmatrix}(C_3,D_1) + f_4(D_1)\\(C_3,D_2) + f_4(D_2)\end{Bmatrix}$$

$$= \text{Min}\begin{Bmatrix}8+5\\10+2\end{Bmatrix} = \text{Min}\begin{Bmatrix}13\\2\end{Bmatrix} = 12 \quad \text{最优决策 } C_3 \rightarrow D_2$$



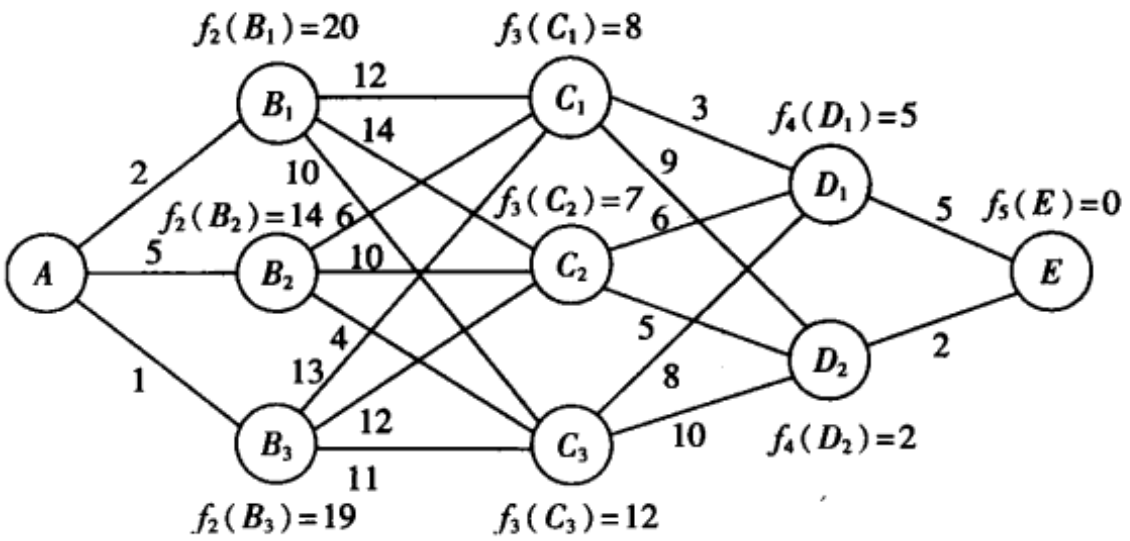
$$f_2(B_1) = \text{Min}\begin{Bmatrix}(B_1,C_1)+f_3(C_1)\\(B_1,C_2)+f_3(C_2)\\(B_1,C_3)+f_3(C_3)\end{Bmatrix} = \text{Min}\begin{Bmatrix}12+8\\14+7\\10+12\end{Bmatrix} = \text{Min}\begin{Bmatrix}20\\21\\22\end{Bmatrix} = 20$$

最优决策 $B_1 \rightarrow C_1$



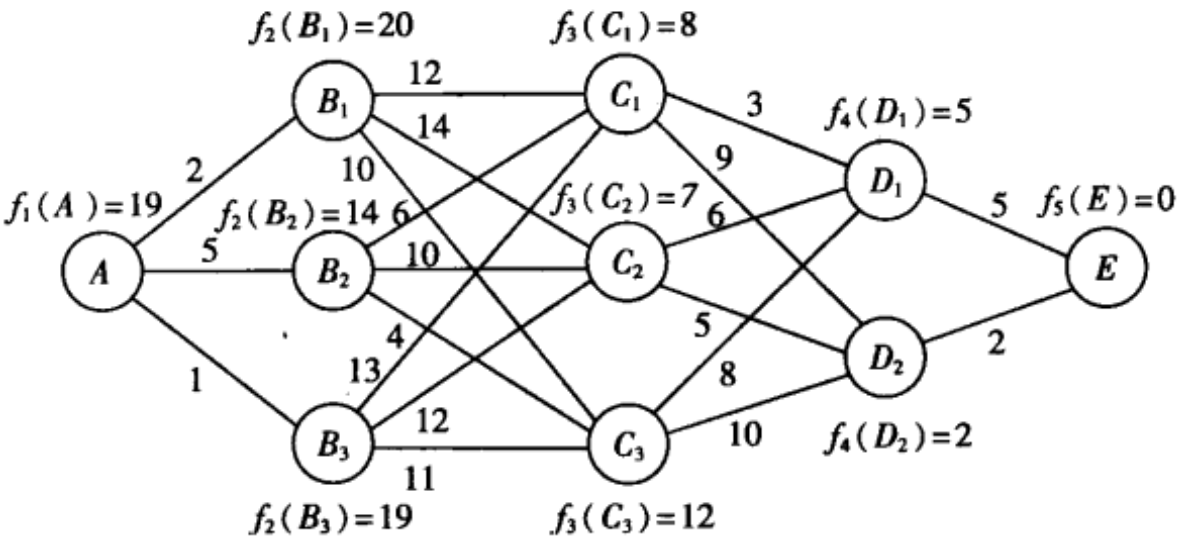
$$f_2(B_2) = \text{Min}\begin{Bmatrix}(B_2,C_1)+f_3(C_1)\\(B_2,C_2)+f_3(C_2)\\(B_2,C_3)+f_3(C_3)\end{Bmatrix} = \text{Min}\begin{Bmatrix}6+8\\10+7\\4+12\end{Bmatrix} = \text{Min}\begin{Bmatrix}14\\17\\16\end{Bmatrix} = 14$$

最优决策 $B_2 \rightarrow C_1$



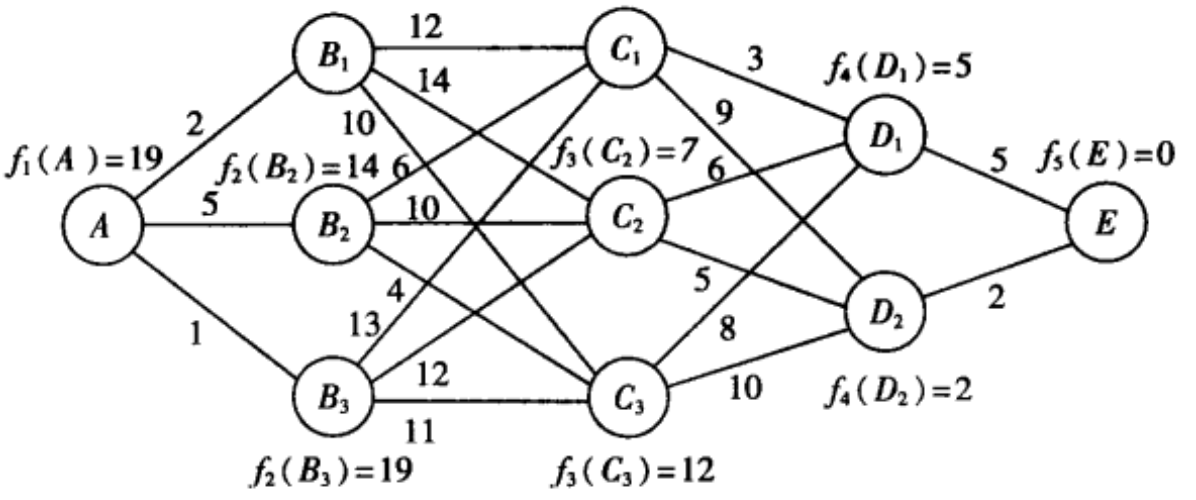
$$f_2(B_3) = \text{Min}\begin{Bmatrix}(B_3,C_1) + f_3(C_1)\\(B_3,C_2) + f_3(C_2)\\(B_3,C_3) + f_3(C_3)\end{Bmatrix} = \text{Min}\begin{Bmatrix}13+8\\12+7\\11+12\end{Bmatrix} = \text{Min}\begin{Bmatrix}21\\19\\23\end{Bmatrix} = 19$$

最优决策 $B_3 \rightarrow C_2$

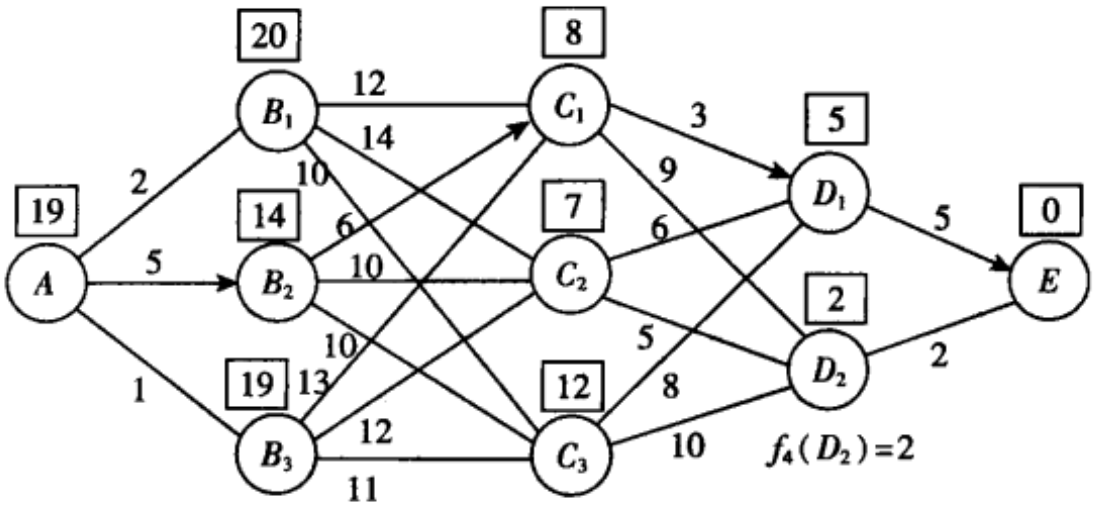


$$f_1(A) = \text{Min}\begin{Bmatrix}(A,B_1) + f_2(B_1)\\(A,B_2) + f_2(B_2)\\(A,B_3) + f_2(B_3)\end{Bmatrix} = \text{Min}\begin{Bmatrix}2+21\\5+14\\1+19\end{Bmatrix} = \text{Min}\begin{Bmatrix}23\\19\\20\end{Bmatrix} = 19$$

最优决策 $A \rightarrow B_2$



从 A 到 E 的最短路径为 19，路线为：$A \rightarrow B_2 \rightarrow C_1 \rightarrow D_1 \rightarrow E$



三、动态规划的 WinQSB 应用

用 WinQSB 软件求解动态规划问题，需要调用子程序 Dynamic Programming(DP)。该程序有 3 个子块：最短路问题(Stagecoach Problem)、背包问题(Knapsack Problem)和生产与存储(Production and Inventory Scheduling)问题。在这里，我们只介绍前面两个问题，有兴趣的读者可以自己练习生产与存储问题。

1. 最短路问题

用 WinQSB 软件求解例 4.11 中 A 到 E 的最短路径。

解：调用子程序 Dynamic Programmin(DP)，建立新问题，如图 4.13 所示，选择第一项，输入标题和节点数。然后，输入数据按图 4.12 两点间连线的方向将距离输入图 4.14，没有线连接时不输入数据。最后，点击菜单栏“Solve and Analyze”的下拉菜单“Solve the Problem”，选择路线 $v_1 \rightarrow v_{10}$，得到图 4.15 的结果，如果需要得到各点之间的最短路，只需要选择相应路线即可。

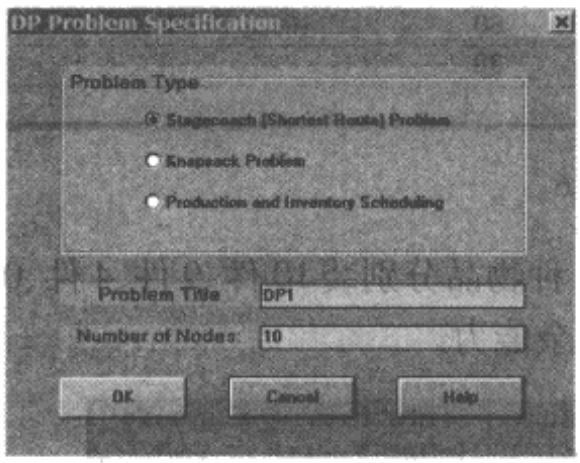


图 4.13

From \ To	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10
v1		2	5	1						
v2					12	14	10			
v3					6	10	10			
v4					13	12	11			
v5						3	9			
v6						6	5			
v7						8	10			
v8										5
v9										2
v10										

图 4.14

04-23-2006 Stage	From Input State	To Output State	Distance	Cumulative Distance	Distance to v10
1	v1	v3	5	5	19
2	v3	v5	6	11	14
3	v5	v8	3	14	8
4	v8	v10	5	19	5
	From v1	To v10	Min. Distance	= 19	CPU = 0

图 4.15

2. 背包问题

用 WinQSB 软件求解下列背包问题。

例 4.12 已知 1 吨集装箱最大载重量为 800 千克，有 5 种物品各 10 件，单位物品重量和价值如表 4.9 所示，求价值最大的装载方案。

表 4.9

物品	1	2	3	4	5
物品限量(件)	10	10	10	10	10
单位物品重量(千克)	20	15	40	60	30
单位物品价值(元)	40	25	60	70	60

解:如图 4.13 所示,选择第二项,输入标题和物品的品种数 5。按图4 - 16 的形式输入数据。第一列为物品名称,第二列为物品限量和集装箱载重量限制,第三列为单位物品重量,最后一列是物品价值函数。如果只输入 40、25 等数据,系统将看成与数量无关的固定值。

Item (Stage)	Item Identification	Units Available	Unit Capacity Required	Return Function (X: Input D) [e.g., 50X, 3X+100, Z-(3X^2+5)]
1	a	10	20	40a
2	b	10	15	25b
3	c	10	40	60c
4	d	10	60	70d
5	e	10	30	50e
Knapsack	Capacity =	800		

图 4.16

求解结果如图 4.17 所示,图 4.17 表示 5 种物品分别装 10 件、9 件、4 件、0 件及 10 件,总价值 1 365 元,集装箱还有 5 千克的剩余能力。

04-23-2006 Stage	Item Name	Depletion Quantity (X)	Return Function	Total Item Return Value	Capacity Left
1	a	10	40a	400	600
2	b	9	25b	225	465
3	c	4	60c	240	305
4	d	0	70d	0	305
5	e	10	50e	500	5
	Total	Return	Value =	1365	CPU = 0.23

图 4.17

第四节　非线性规划

例 4.13　投资问题。

有资本金 5(百万元),存在两个投资项目的回报率为:投资项目 1 为 120%,项目 2 为 116%。项目 1、项目 2 的风险损失为 $2x_1^2 + x_2^2 + (x_1 + x_2)^2$。问如何使期望收益最大,风险损失最小?

$$M1:\text{Min } f(x) = 2x_1^2 + x_2^2 + (x_1 + x_2)^2$$

$$s.t.\begin{cases} g_1(x) = x_1 + x_2 \leqslant 5 \\ g_2(x) = 0.2x_1 + 0.16x_2 \geqslant a \\ x_1 \geqslant 0, x_2 \geqslant 0 \end{cases}$$

$$M2:\text{Max}\, f(x) = 0.2x_1 + 0.16x_2 - \theta[2x_1^2 + x_2^2(x_1 + x_2)^2]$$

$$s.t.\begin{cases} g_1(x) = x_1 + x_2 \leqslant 5 \\ x_1 \geqslant 0, x_2 \geqslant 0 \end{cases}$$

其中,$\theta(\geqslant 0)$ 为常数,为风险系数

例 4.14 选址问题。

建煤气站,第 j 个用户位置:(a_j, b_j) $(j = 1, 2, \cdots, m)$;煤气需求 A_jB_j 为煤气站向 j 用户输气 1 吨公里运价;第 3 位用户距气站不超过 D 公里;第 9 位用户距气站不少于 E 公里试决定煤气站位置,使总运费最低。

解:设煤气站位置为(x, y)

$$\text{Min}f(x,y) = \sum_{j=1}^{m} A_jB_j \sqrt{(x - a_j)^2 + (y - b_j)^2}$$

$$s.t.\begin{cases} (x - a_3)^2 + (y - b_3)^2 \leqslant D^2 \\ (x - a_9)^2 + (y - b_9)^2 \leqslant E^2 \end{cases}$$

以上两例的目标函数或约束条件为非线性函数,则这样的规划问题就属于非线性规划,而线性规划的目标函数和约束条件均为线性函数。

非线性规划的研究始于 1939 年,是由 W. 卡鲁什首次进行的,20 世纪 40 年代后期进入系统研究。1951 年 H. W. 库恩和 A. W. 塔克尔提出最优化的判别条件,从而奠定了非线性规划的理论基础,后来在理论研究和实用算法方面都有很大的发展。非线性规划求解方法可分为无约束问题和约束问题来讨论,前者实际上就是多元函数的极值问题,是后一问题的基础。无约束问题的求解方法有最速下降法、共轭梯度法、变尺度法和鲍威尔直接法等。约束问题情况比较复杂,因为在迭代过程中除了要使目标函数下降外,还要考虑近似解的可行性。

一、数学模型

非线性规划数学模型的一般形式为:

$$\text{Min}f(X)$$

$$s.t.\begin{cases} h_i(X) = 0 & i = 1,2,\cdots,m; \\ g_j(X) \geqslant 0 & j = 1,2,\cdots,n。 \end{cases}$$

目标函数 $f(X)$ 和约束函数 $h_i(X)$、$g_j(X)$ 中至少有一个为 X 的非线性函数。

考虑约束函数 $h_i(X) = 0$ 可以用以下两个不等式约束进行替换:

$$\begin{cases} h_i(X) \geqslant 0 \\ -h_i(X) \leqslant 0 \end{cases}$$

因此,非线性规划也可以写成:

$$\text{Min}f(X)$$

$$s.t.\begin{cases} h_i(X) \geqslant 0 \\ -h_i(X) \leqslant 0 \\ g_j(X) \geqslant 0 \end{cases}$$

如果约束条件出现“$\leqslant$”不等式形式,可以两边同乘以“-1”,即可以将其变为“$\geqslant$”形式。

二、图解法

当只有 2 个自变量时,非线性规划也可以用图形法求解。

例 4.15 用图解法求解下面的非线性规划问题。

$$\text{Min} f(x_1,x_2) = x_1^2 + x_2^2$$

$$s.t.\begin{cases} 1 - x_1 - x_2 \leqslant 0 \\ x_1 - 1 \leqslant 0 \\ x_2 - 1 \leqslant 0 \end{cases}$$

首先,在坐标平面上划出目标函数的等值线,它是以点(0,0) 为圆形的同心圆。然后,根据约束条件划出可行域,它是如图 4.18 所示的三角形。最后,求解最小目标函数值。很容易可以看出,图 4.18 中,点 x^* 为最小目标函数值最小点,$x^* = (1/2,1/2)$,最小值为 1/2。

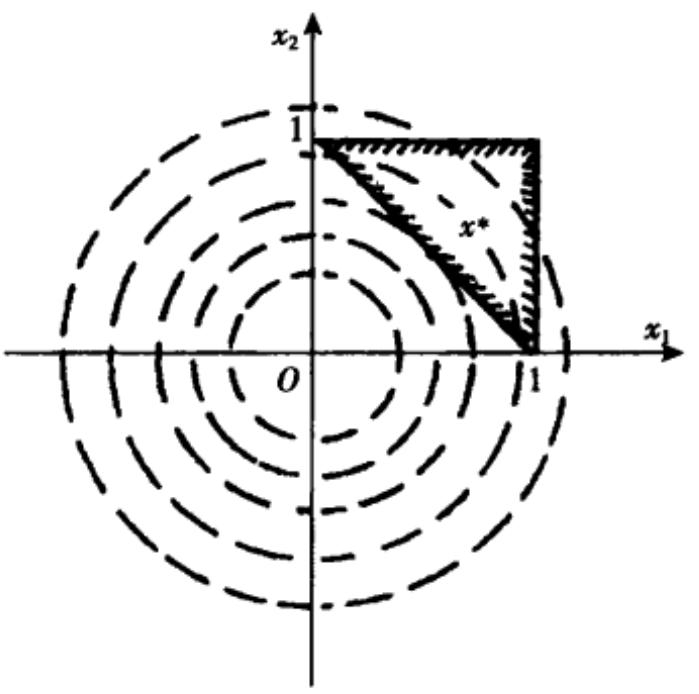


图 4.18

三、无约束问题

多变变量无约束极值问题一般可表示为 $\text{Min} f(X)$ 的形式,其中,$X = (x_1,x_2,\cdots,x_n)^T$,且 $x_n \in R^n$。一个复杂的非线性函数,用一般求驻点[$\nabla f(X^*) = 0$] 的方法求其极值点往往是很困难的,变量很多时就更加困难。因此,人们在实际计算时还是要选用迭代的方法。与单变量极值问题一样,所用的迭代方法通常也分两类:一类用到函数的一阶或二阶导数,称为解析法;另一类不用导数计算,在迭代中仅用各点的函数值,称为直接法。在函数复杂、求导计算量大时,用直接法有利。但对于简单函数,解析法往往快一些。本节将介绍解析法中两个比较常用的算法。

1. 最速下降法(梯度法)

最速下降法的计算步骤为:

第一步,给定初始数据:起始点 $X^{(0)}$,给定终止误差 $\varepsilon > 0$,令 $k = 0$。

第二步,求梯度向量模的值:$\| \nabla f(x^{(k)}) \|$。

若 $\| \nabla f(x^{(k)}) \| < \varepsilon$,停止计算,输出 $X^{(k)}$,作为极小点的近似值,否则转下一步。

第三步,构造负梯度方向:$p^{(k)} = -\nabla f(X^{(k)})$。

第四步,进行一维搜索。无论用哪种方法求得 λ_k 后,令:

$X^{(k+1)} = X^{(k)} + \lambda k p^{(k)} = X^{(k)} - \lambda_k \nabla f(X^{(k)})$,

置 $k = k + 1$,转到第二步。

例 4.16 用最速下降法求 $f(X) = x_1^2 + 5x_2^2$ 的极小点,取允许误差 $\varepsilon = 0.7$。

解:取初始点 $X^{(0)} = (2,1)^T$。

$\nabla f(X) = (2x_1, 10x_2)^T$, $\nabla f(X^{(0)}) = (4,10)^T$。其海赛矩阵为:

$$\nabla^2 f(X) = \begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix}$$

$$\lambda_0 = \frac{(4,10)\begin{pmatrix} 4 \\ 10 \end{pmatrix}}{(4,10)\begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix}\begin{pmatrix} 4 \\ 10 \end{pmatrix}} = 0.112\,4$$

$$X^{(1)} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - 0.112\,4\begin{pmatrix} 4 \\ 10 \end{pmatrix} = \begin{pmatrix} 1.550\,4 \\ -0.124\,0 \end{pmatrix}$$

$$\nabla f(X^{(1)}) = \begin{pmatrix} 3.100\,8 \\ -1.240\,0 \end{pmatrix}, \| \nabla f(X^{(1)}) \|^2 = 11.526 > \varepsilon$$

$$\lambda_1 = \frac{(3.100\,8, -1.240\,0)\begin{pmatrix} 3.100\,8 \\ -1.240\,0 \end{pmatrix}}{(3.100\,8, -1.240\,0)\begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix}\begin{pmatrix} 3.100\,8 \\ -1.240\,0 \end{pmatrix}} = 0.322\,3$$

$$X^{(2)} = \begin{pmatrix} 1.550\,4 \\ -0.124\,0 \end{pmatrix} - 0.322\,3\begin{pmatrix} 3.100\,8 \\ -1.240\,0 \end{pmatrix} = \begin{pmatrix} 0.551\,0 \\ 0.275\,7 \end{pmatrix}$$

$$\nabla f(X^{(2)}) = \begin{pmatrix} 1.102 \\ 2.757 \end{pmatrix}, \| \nabla f(X^{(2)}) \|^2 = 8.815 > \varepsilon$$

$$\lambda_2 = \frac{(1.102,2.757)\begin{pmatrix} 1.102 \\ 2.757 \end{pmatrix}}{(1.102,2.757)\begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix}\begin{pmatrix} 1.102 \\ 2.757 \end{pmatrix}} = 0.112\,4$$

$$X^{(3)} = \begin{pmatrix} 0.551\,0 \\ 0.275\,7 \end{pmatrix} - 0.112\,4\begin{pmatrix} 1.102 \\ 2.757 \end{pmatrix} = \begin{pmatrix} 0.427\,1 \\ -0.341\,9 \end{pmatrix}$$

$$\nabla f(X^{(3)}) = \begin{pmatrix} 0.854\,2 \\ -0.341\,9 \end{pmatrix}, \| \nabla f(X^{(3)}) \|^2 = 0.846\,6 > \varepsilon$$

$$\lambda_3 = \frac{(0.854\,2, -0.341\,9)\begin{pmatrix} 0.854\,2 \\ -0.341\,9 \end{pmatrix}}{(0.854\,2, -0.341\,9)\begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix}\begin{pmatrix} 0.854\,2 \\ -0.341\,9 \end{pmatrix}} = 0.322\,1$$

$$X^{(4)} = \begin{pmatrix} 0.427\,1 \\ -0.341\,9 \end{pmatrix} - 0.322\,1\begin{pmatrix} 0.854\,2 \\ -0.341\,9 \end{pmatrix} = \begin{pmatrix} 0.152 \\ 0.075\,9 \end{pmatrix}$$

$$\nabla f(X^{(4)}) = \begin{pmatrix} 0.304 \\ 0.759 \end{pmatrix}, \| \nabla f(X^{(4)}) \|^2 = 0.668\ 5 < \varepsilon$$

故以 $f(X^{(4)}) = (0.152, 0.075\ 9)^T$ 为近似极小点，此时的函数值 $f(X^{(4)}) = 0.519$。该问题的精确解是 $X^* = (0,0)^T$，$f(X^*) = 0$。可知，要得到真正的精确解，需无限迭代下去。如图 4.19 所示。

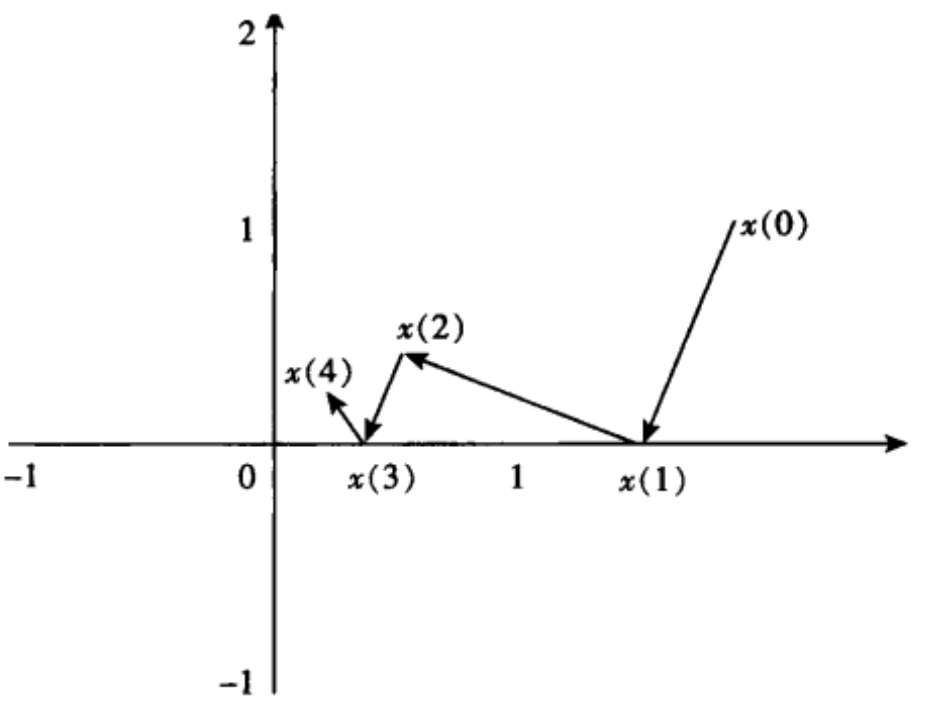


图 4.19

最速下降法对初始点的选择要求不高，每一轮迭代工作量较少，它可以比较快地从初始点达到极小点附近。但在接近极小点时，最速下降法却会出现锯齿现象，收敛速度很慢，因为对一般二元函数，在极小点附近可用极小点的二阶泰勒多项式来近似，而后者为凸函数时，它的等值线是一组共心椭圆，特别是当椭圆比较扁平（与圆相差越大）时，最速下降法的收敛速度越快（见图 4.20）。

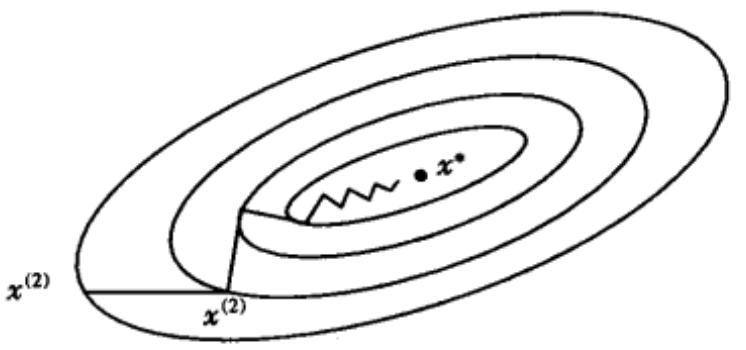


图 4.20

2. 牛顿法

牛顿法的计算步骤：

第一步，选取初始数点：初始点 $X^{(0)}$，终止条件 $\varepsilon > 0$，令 $k = 0$。

第二步，求梯度向量 $\nabla f(X^{(k)})$，并计算 $\| \nabla f(x^{(k)}) \|$，若 $\| \nabla f(x^{(k)}) \| < \varepsilon$，停止迭代，输出 $X^{(k)}$，否则转下一步。

第三步，构造牛顿方向。计算 $[\nabla^2 f(X^{(k)})]^{-1}$，且以下式

$p^{(k)} = -[\nabla^2 f(X^{(k)})]^{-1}\nabla f(X^{(k)})$

计算 $p^{(k)}$。

第四步,以式 $X^{(k+1)} = X^{(k)} + p^{(k)}$ 计算作为下一轮迭代点,令 $k = k+1$,转第二步。

例 4.16 用牛顿法求解上例的极小点。

解:任取初始点 $X^{(0)} = (2,1)^T$,算出 $\nabla f(X^{(0)}) = (4,10)^T$。在本例中:

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix} \quad A^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/10 \end{pmatrix}$$

$$X^* = X^{(0)} - A^{-1}\nabla f(X^{(0)}) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 1/2 & 0 \\ 0 & 1/10 \end{pmatrix}\begin{pmatrix} 4 \\ 10 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$\nabla f(X^*) = (0,0)^T$,可知 X^* 确实为极小点。

将本例与例 4.15 进行比较,可知牛顿法的搜索方向与最速下降法的搜索方向不同。

四、约束极值问题

以上我们介绍了无约束问题的最优化方法。但实际问题中,大多数都是有约束条件的问题,求解带有约束条件的问题比无约束问题要困难得多,也复杂得多。在每次迭代时,不仅要使目标函数值有所下降,而且要使迭代点都落在可行域内(个别算法除外)。求解带有约束的极值问题的常用方法是:将约束问题化为一个或一系列的无约束极值问题;将非线性规划化为近似的线性规划;将复杂问题变为较简单问题等等。

1. 库恩 — 塔克条件

库恩 — 塔克(K - T)条件是非线性规划领域中最重要的理论成果之一,是确定某点为最优点的必要条件,只要是最优点,就必须满足这个条件。但一般来说它并不是充分条件,因而满足这个条件的点不一定就是最优点。

现可将库恩 — 塔克条件叙述如下:

设 X^* 是非线性规划的极小点,而且与 X^* 点处的所有起作用的约束梯度线性无关,则存在数 $\mu_1^*, \mu_2^*, \cdots, \mu_l^*$,使

$$\text{Min } f(x_1,x_2) - \sum_{j=1}^{l} \mu_j^* \nabla g_j(X^*) = 0$$

$$s.t. \begin{cases} \mu_j^* g_j(X^*) = 0 \\ \mu_j^* \geqslant 0 \\ j = 1,2,\cdots,l \end{cases}$$

例 4.17 求下面非线性规划问题的 K - T 点。

$$\text{Min } f(x_1,x_2) = (x_1-1)^2 + (x_2-2)^2$$

$$s.t. \begin{cases} g_1(x) = x_1 + x_2 - 2 \leqslant 0 \\ g_2(x) = -x_1 \leqslant 0 \\ g_3(x) = -x_2 \leqslant 0 \end{cases}$$

解:写出 $K-T$ 条件:

$$\begin{cases} 2(x_1-1)+\lambda_1-\lambda_2-\mu_1=0 \\ 2(x_2-2)+\lambda_1-\lambda_3+\mu_1=0 \\ \lambda_1(x_1+x_2-2)=0 \\ \lambda_2 x_1=0 \\ \lambda_1 x_2=0 \\ \lambda_1,\lambda_2,\lambda_3 \geqslant 0 \end{cases}$$

$K-T$ 点还应该满足原问题的约束条件：

$$\begin{cases} x_1+x_2-2 \leqslant 0 \\ -x_1 \leqslant 0,\ -x_2 \leqslant 0 \\ -x_1 \leqslant 0,\ -x_2 \leqslant 0 \\ -x_1+x_2-1=0 \end{cases}$$

可以求出 $K-T$ 点：

$$x^* = \left(\frac{1}{2},\frac{3}{2}\right)^T$$

极小值为 1/2。

2. 制约函数法

制约函数法是通过构造某种制约函数，并将它加到非线性规划的目标函数上，从而将原来的约束极值问题，转化为无约束极值问题来求解。其中最基本的两种是罚函数法（也称外点法）和障碍函数法（也称内点法）。

（1）罚函数法

罚函数法的迭代步骤如下：

第一步，取第一个罚因子 $M_1>0$（例如取 $M_1=1$），允许误差 $\varepsilon>0$，并令 $k=1$。

第二步，求 Min $P(X,M_k)$ 无约束极值问题的最优解，其中 $P(X,M_k)$ 可取下式的形式：

$$P(X,M)=f(X)+M\sum_{i=1}^{m}[h_i(X)]^2+M\sum_{j=1}^{l}\{\text{Min}[0,g_j(X)]\}^2$$

$$P(X,M)=f(X)+M\sum_{i=1}^{m}[h_i(X)]^3+M\sum_{j=1}^{l}[g_j(X)]^2u_j[g_j(X)]$$

设其极小点为 $X^{(k)}$。

第三步，若存在某一个 $j(1\leqslant j\leqslant l)$，有 $-g_j(X^{(k)})>\varepsilon$ 和存在某一个 $i(1\leqslant i\leqslant m)$，有 $|h_i(X^{(k)})|>\varepsilon$，取 $M_{k+1}>M_k$，并令 $k=k+1$，然后转回第二步。

否则，停止迭代，得到所要的点 $X^{(k)}$。

例 4.18　用外点法求解非线性规划。

$$\text{Min } f(X)=x_1+x_2$$

$$s.t.\begin{cases} g_1(X)=-x_1^2+x_2\geqslant 0 \\ g_2(X)=x_1\geqslant 0 \end{cases}$$

解：构造罚函数

$$P(X,M)=x_1+x_2+M\{[\text{Min}[0,(-x_1^2+x_2)]^2+[\text{Min}(0,x_1)]^2\}$$

$$\frac{\partial P}{\partial x_1} = 1 + 2M\{\text{Min}\,[0,(-x_1^2 + x_2)(-2x_1)]\} + 2M[\text{Min}\,(0,x_1)]$$

$$\frac{\partial P}{\partial x_2} = 1 + 2M\{\text{Min}\,[0,(-x_1^2 + x_2)]\}$$

令$\frac{\partial P}{\partial x_1} = \frac{\partial P}{\partial x_2} = 0$

$Min\ P(X,M)$ 得的解为：

$$X(M) = [-\frac{1}{2(1+M)},(\frac{1}{4(1+M)^2} - \frac{1}{2M})]^T$$

取 $M = 1,2,3,4$，可得到以下结果：

$M = 1: X = (-1/4, -7/16)^T$

$M = 2: X = (-1/6, -2/9)^T$

$M = 3: X = (-1/8, -29/192)^T$

$M = 4: X = (-1/10, -23/200)^T$

可知 $X(M)$ 从 R 的外面逐步逼近的边界，当 $M \to \infty$ 时，$X(M)$ 趋于极小解 $X_{\text{Min}} = (0,0)^T$。

(2) 障碍函数法

障碍函数法的迭代步骤如下：

第一步，取第一个障碍因子 $r_1 > 0$（例如取 $r_1 = 1$），允许误差 $\varepsilon > 0$，并令 $k = 1$。

第二步，构造障碍函数，障碍项可采用倒数函数：

$$\bar{P}(X,r_k) = f(X) + r_k \sum_{j=1}^{l} \frac{1}{g_j(X)}$$

也可采用对数函数：

$$\bar{P}(X,r_k) = f(X) - r_k \sum_{j=1}^{l} \lg(g_j(X))$$

第三步，对障碍函数进行无约束极小化，设所得极小解为 $X^{(k)}$。

第四步，检查是否满足收敛准则：$r_k \sum_{j=1}^{l} \frac{1}{g_j(X^{(k)})} \leqslant \varepsilon$ 或 $\left| r_k \sum_{j=1}^{l} \lg[g_j(X^{(k)})] \right| \leqslant \varepsilon$。

如果满足此准则，则以 $X^{(k)}$ 为原约束问题的近似极小解，停止选代；否则，取 $r_{k+1} < r_k$，令 $k = k + 1$，转回第三步继续进行迭代。

例 4.19 用内点法解非线性规划。

$$\text{Min}\, f(X) = \frac{1}{3}(x_1 + 1)^3 + x_2$$

$$s.t. \begin{cases} g_1(X) = x_1 - 1 \geqslant 0 \\ g_2(X) = x_2 \geqslant 0 \end{cases}$$

解：构造障碍函数

$$\bar{P}(X,r) = \frac{1}{3}(x_1 + 1)^3 + x_2 + \frac{r}{x_1 - 1} + \frac{r}{x_2}$$

$$\frac{\partial \bar{P}}{\partial x_1} = (x_1 + 1)^2 - \frac{r}{(x_1 - 1)^2} = 0$$

$$\frac{\partial \bar{P}}{\partial x_2} = 1 - \frac{r}{x_2^2} = 0$$

联立解上述两个方程,得:

$$x_1(r) = \sqrt{1+\sqrt{r}},\ x_2(r) = \sqrt{r}$$

得最优解:

$$X_{\text{Min}} = \lim_{r\to 0}(\sqrt{1+\sqrt{r}},\sqrt{r})^T = (1,0)^T$$

五、WinQSB 应用举例

用 WinQSB 软件求解非线性规划问题,要调用子程序 Nonlinear Programming(DP)。下面,我们首先给出应用牛顿法计算例 4.20 的过程,然后用 WinQSB 软件来求解。

例 4.20 用牛顿法求解 $\text{Min}\, f(X) = x_1^2 + 25x_2^2$,选取 $X^{(0)} = (2,2)^T$,$\varepsilon = 10^{-6}$。

解:计算

$$\nabla f(X) = \begin{pmatrix} 2x_1 \\ 50x_2 \end{pmatrix},\ \nabla^2 f(X) = \begin{pmatrix} 2 & 0 \\ 0 & 50 \end{pmatrix}$$

故有:

$$\nabla f(X^{(0)}) = \begin{pmatrix} 4 \\ 100 \end{pmatrix},\ \nabla^2 f[X^{(0)}] = \begin{pmatrix} 2 & 0 \\ 0 & 50 \end{pmatrix}$$

得 $\nabla^2 f[X^{(0)}]$ 是正定矩阵,又:

$$[\nabla^2 f(X^{(0)})]^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/50 \end{pmatrix}$$

$$p^{(0)} = -[\nabla^2 f(X^{(0)})]^{-1}\nabla f(X^{(0)}) = -\begin{pmatrix} 1/2 & 0 \\ 0 & 1/50 \end{pmatrix}\begin{pmatrix} 4 \\ 100 \end{pmatrix} = -\begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

$$X^{(1)} = X^{(0)} + p^{(0)} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} + \begin{pmatrix} -2 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

计算 $\nabla f(X^{(1)}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$,故有:$\| \nabla f(X^{(1)}) \| < \varepsilon$

停止迭代,并输出 $X^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$作为极小点。

如果运用 WinQSB 软件求解,我们首先调用子程序 Nonlinear Programming,然后建立一个新问题,如图 4.21 所示,输入标题、变量数目和约束条件数,按照非线性规划问题输入模型和数据。最后,点击菜单栏“Solve and Analyze”的下拉菜单“Solve the Problem”,在弹出的“Solution Setup”对话框中(如图 4.22)输入计算精确度 1.00E - 6 和初始迭代点(0,0),得到如图 4.24 所示的计算结果。

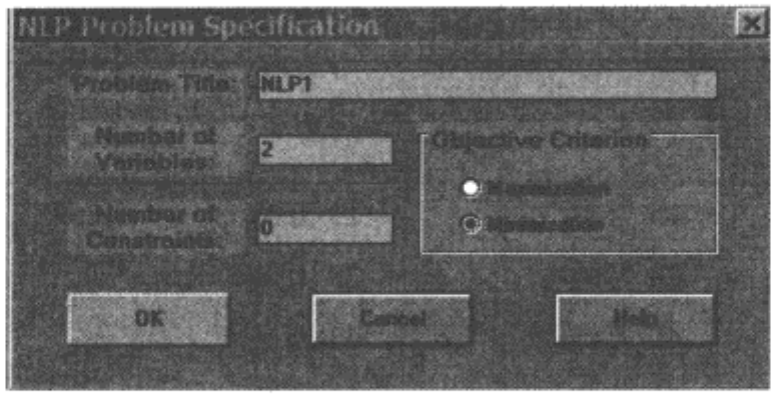


图 4.21

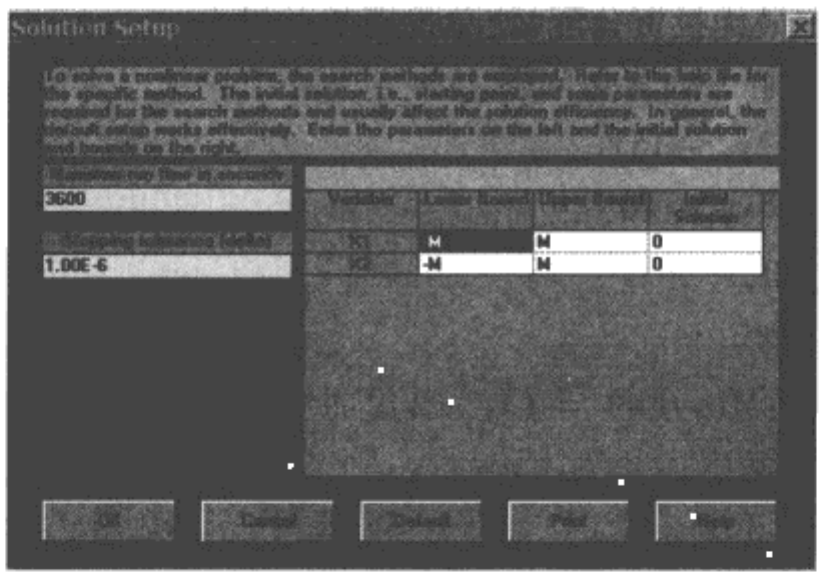


图 4.22

	OBJ / Constraint / Variable Bound
Minimize	x1^2+25x2^2
X1	>=-M, <=M
X2	>=-M, <=M

图 4.23

04-25-2006	Decision Variable	Solution Value
1	X1	0
2	X2	0
Minimized	Objective Function =	0

图 4.24

习题

1. 有资金4万元,投资A、B、C三个项目,每个项目的投资效益与投入该项目的资金有关。三个项目A、B、C的投资效益(万吨)和投入资金(万元)的关系如表4.9所示。

表 4.9

项目 / 投入资金	A	B	C
1 万元	15 万吨	13 万吨	11 万吨
2 万元	28 万吨	29 万吨	30 万吨
3 万元	40 万吨	43 万吨	45 万吨
4 万元	51 万吨	55 万吨	58 万吨

求对三个项目的最优投资分配,使总投资效益最大。

2. 用动态规划求解下列背包问题。

$$\text{Max } z = 12x_1 + 22x_2 + 15x_3$$

$$s.t. \begin{cases} 2x_1 + 4x_2 + 3x_3 \leqslant 10 \\ x_1, x_2, x_3 \geqslant 0, \\ x_1, x_2, x_3 \in Z \end{cases}$$

3. 用梯度法求解:$\text{Min } f(x) = (x_1 - 2)2 + (x_2 - 1)2, \varepsilon = 0.1$。

4. 用 K - T 条件解:

$$\text{Min } f(x) = (x - 3)2$$

$$s.t.\ 0 \leqslant x \leqslant 5$$

5. 用罚函数法求解:

$$\text{Min } f(X) = x_1^2 + x_2^2$$

$$s.t.\ x_2 - 1 = 0$$

6. 用外点法求解:

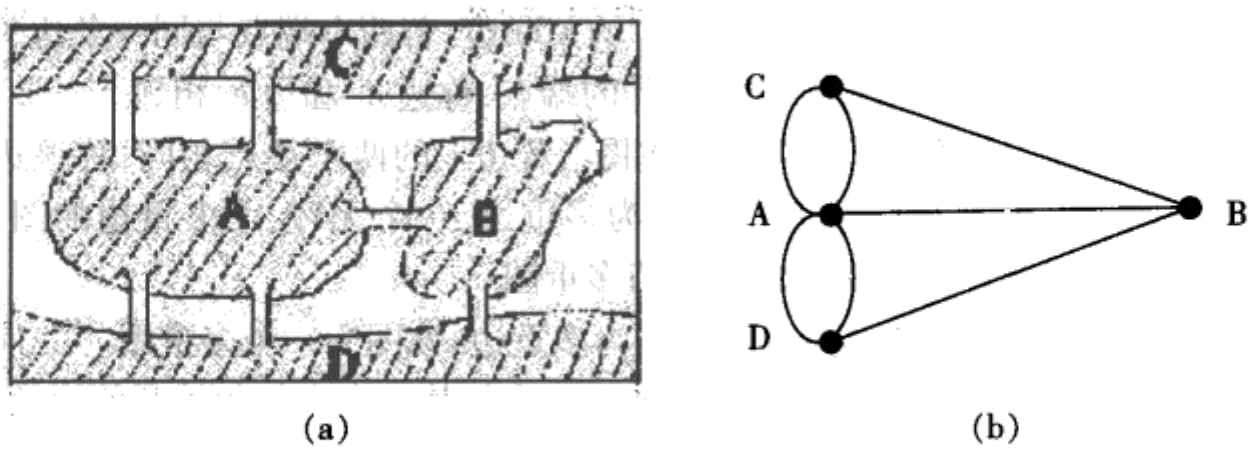
$$\text{Min} f(X) = (x_1 - 2)^2 + x_2^2$$

$$s.t.\ x_2 - 1 = 0$$

第五章　图与网络分析

引例:哥尼斯堡七桥问题

18 世纪,瑞士的哥尼斯堡城中有一条叫普雷格尔(Pregel)的河,该河中有两个小岛,河上有七座桥连接着两岸和这两个小岛,如图 5.1(a)所示。当时那里的居民热衷于这样的游戏:一个散步者能否走过七座桥,且每座桥只走过一次,最后回到出发点。很长一段时间,没有人想出走法,但又不能说明走法确实不存在。欧拉用 C、D、A、B 四点分别表示河的两岸和两个小岛,用两点间的连线表示桥,将此问题归结为如图 5.1(b)所示的一笔画问题,即能否从某一点开始,一笔画出这个图形,最后回到原点,而不重复?欧拉证明了这是不可能的,因为图 5.1(b)中的每个点只与奇数条线相关联,不可能将这个图不重复地一笔画成。这是古典图论中的一个著名问题。



(a)　　(b)

图 5.1

图论(Theory of Graphs,或 Graph Theory)是决策科学,特别是运筹学中应用十分广泛的一个分支,它广泛地应用在经济学、管理学、物理学、化学、控制论、信息论和电子计算机等领域。在实际生活、生产和科学研究中,有很多问题可以用图论的理论和方法来解决。图论的创始人是著名的瑞士数学家欧拉(Euler,1707—1783 年),他在自己 29 岁时,即公元 1736 年发表了关于图论的第一篇论文,解决了著名的“哥尼斯堡七桥问题”(The Konigsberg Bridge Problem),创立了拓扑学。

近年来,随着科学技术的进步以及电子计算机的蓬勃发展,图论得到进一步发展。特别是,图论被应用于网络分析中所产生的最小树问题、最短路问题、最大流问题、运输问题、分配问题以及旅行售货员与中国邮路等问题,解决了很多庞大而复杂的工程系统设计和管理决策的最优化问题,在数学、工程技术及经营管理等各个领域受到了越来越广泛的重视。

第一节　图与网络的基本概念

在实际生活当中，我们经常碰到各种各样的图：公路或铁路交通图、组织机构图、通信网络图等。决策科学中研究的图是上述各种图的抽象概括，因为上述图只反映对象与对象之间的某种特定关系和相互联系。如交通图是表明所研究的不同地点之间的道路沟通情况，机构图表明某机构的构成及组成该机构的各分支机构之间的关系等。

定义 5.1　［图，赋权图］若用点表示研究的对象，用边表示这些对象之间的联系，则图可以定义为点和边的集合，记作 $G = (V,E)$，其中 V 是点的集合，E 是边的集合。若给图中的边赋以具体的含义和权数，如距离、时间、费用、容量等，则这样的图称为赋权图，记为 $N = (V,E,W)$，其中 W 为边上所有全体权数的集合。

一般地，图中的点用 v 表示，边用 e 表示。每条边可用它所连接的点表示，如在图 5.2 中，$e_1 = [v_1, v_1]$，$e_2 = [v_1, v_3]$ 或 $e_2 = [v_3, v_1]$。

定义 5.2　［端点，关联边，相邻］若有边 e 可表示为 $e = [v_i, v_j]$，称点 v_i 和点 v_j 是边 e 的端点，反之，称边 e 为点 v_i 或点 v_j 的关联边。若两个端点与同一条边关联，称点 v_i 和点 v_j 相邻；若边 e_i 和边 e_j 具有公共的端点，称边 e_i 和边 e_j 相邻。

定义 5.3　［环，多重边，简单图、多重图］若边 e 的两个端点相同，称该边为环。若两个点之间有多于一条的边，称这些边为多重边。一个无环、无多重边的图称为简单图，一个无环、但允许有多重边的图称为多重图。

如图 5.2 中，边 e_1 为环，边 e_4 和边 e_5 为多重边，该图为一个多重图。

定义 5.4　［次，奇点，偶点，悬挂点，孤立点］与某一个点 v_i 相关联的边的数目称为点 v_i 的次，记为 $d(v_i)$。次为奇数的点，称为奇点。次为偶数的点，称为偶点。次为 1 的点，称为悬挂点。次为 0 的点，称为孤立点。

如图 5.2 中，$d(v_1) = 4$，$d(v_3) = 5$，$d(v_5) = 1$，$d(v_6) = 0$。

定义 5.5　［链，路，圈，回路，连通图］对于图 G 中交替序列 $\mu = (v_{i_1}, e_{i_1}, v_{i_2}, e_{i_2}, \cdots, v_{i_{k-1}}, e_{i_{k-1}}, v_{i_k})$，若各边 $e_{i_1}, e_{i_2}, \cdots, e_{i_{k-1}}$ 互不相同，且任意 $v_{i_{t-1}}$ 和 $v_{i_t}(2 \leqslant t \leqslant k)$ 均相邻，则称 μ 为链。若链中所有点 $v_{i_1}, v_{i_2}, \cdots, v_{i_k}$ 也不相同，称此链为路。对起点和终点重合的链称作回路。在图 G 中，若每一对点之间至少存在一条链，称此图为连通图，否则称为不连通图。

在图 5.2 中，$\mu_1 = (v_5, e_8, v_3, e_2, v_1, e_3, v_2, e_4, v_3, e_7, v_4)$ 和 $\mu_2 = (v_5, e_8, v_3, e_7, v_4)$ 均是一条链，μ_2 可称作路，但在 μ_1 中，因为点 v_3 重复出现，故不能称作路。

定义 5.6　［子图，部分图］对于图 $G_1 = (V_1, E_1)$ 和图 $G_2 = (V_2, E_2)$，若 $V_1 \subseteq V_2$ 和 $E_1 \subseteq E_2$，称 G_1 是 G_2 的一个子图；若有 $V_1 = V_2$，$E_1 \subset E_2$，则称 G_1 是 G_2 的一个部分图。

如图 5.3(a) 是图 5.2 的一个子图，图 5.3(b) 是图 5.2 的一个部分图。可见，部分图也是子图，但子图不一定是部分图。

定义 5.7　［弧，有向图］在图中，若点与点之间的连线是有方向性的（用箭头表

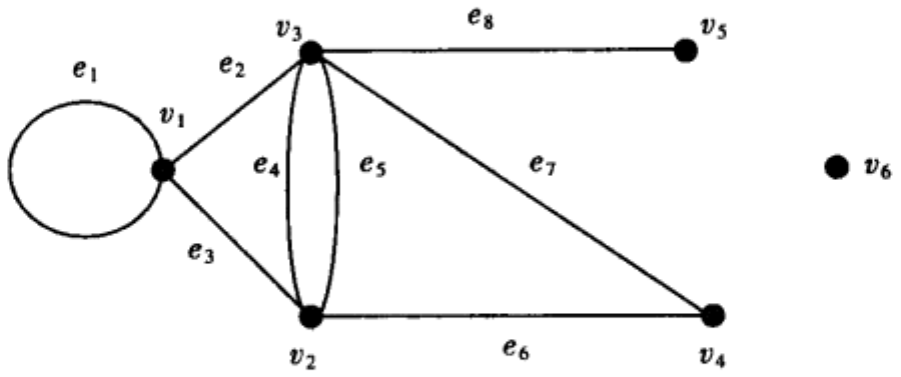
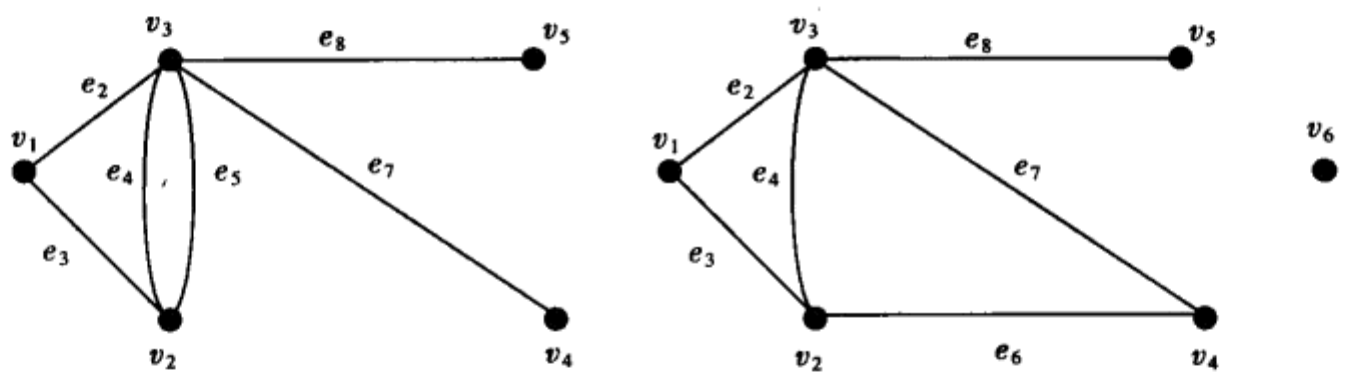


图 5.2



(a)子图　　(b)部分图

图 5.3

示),则这种连线称为弧。一条方向是从 v_i 指向 v_j 的弧记为 $a = (v_i, v_j)$。若一个图是由点及弧所构成的,则称为有向图,记为 $D = (V,A)$,其中 V,A 分别表示有向图 D 的点集合和弧集合。否则,称为无向图。

定义 5.8 [网络] 给定一个有向图 $D = (V,A)$,在其中指定了一点,称为发点(记为 v_s),指定另一点,称为收点(记为 v_t),其余的点称为中间点,并把 D 中的每一条弧的赋权数 c_{ij} 称之为弧 (v_i, v_j) 的容量。这样的赋权有向图,称为网络,记为 $N = (V, A, C)$,其中 C 为弧上所有全体容量的集合。

第二节　最小树问题

一、树的概念

例 5.1　自来水公司铺设管道问题。某自来水公司计划由设在点 v_1 的水厂向分设在其他点共九处的单位供应自来水,v_i 与 v_j 之间的边表示两点间可以铺设管道,边上的权数表示铺设管道的实际费用(单位:万元),如图 5.4 所示。自来水公司现在需要考虑的是,如何架设管道,既保证每个单位都有自来水供应,又使架设管道的总费用最小。

从图 5.4 中可以发现,保证每个单位都有自来水供应的方案不止一个,但我们所要寻找的是最佳方案,即同时使管道的总费用为最小,以降低管道网络的成本。类似这样的问题还有很多,比如在公园景点安装电话线的问题,村镇电缆架设等问题。这

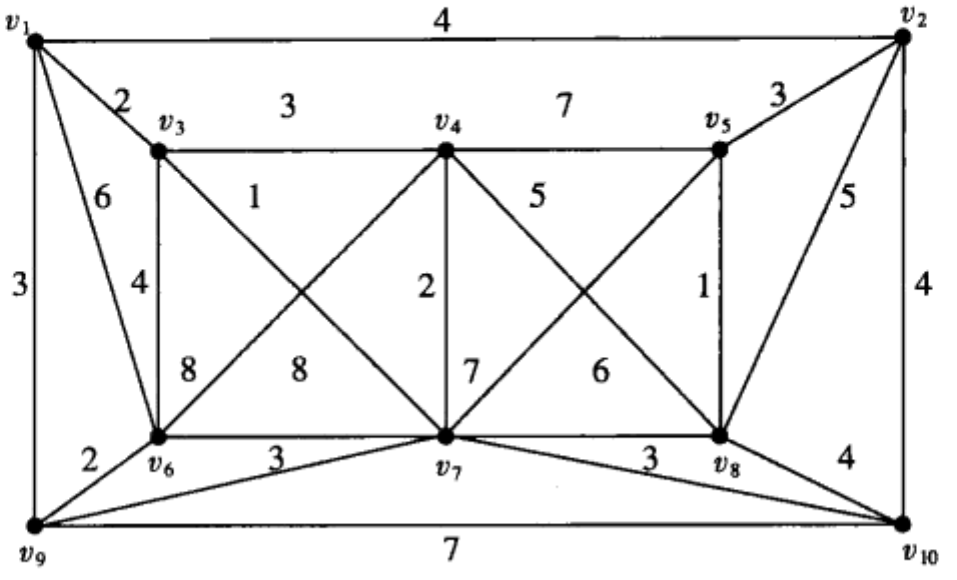


图 5.4

类问题都要求管道或者电缆是互相连通的，并且总希望铺设的距离尽可能短或费用尽可能少。

在生产、生活中常见到的管道、电缆等铺设问题中，我们所面对的工作对象往往是树(*Tree*)，树是图论中的一个重要概念。所谓树就是一个无圈并且连通的无向图，也称为树图。图 5.5 是图 5.4 的一个管道铺设方案路线图，其特征是任意两点之间都由唯一的一条路连通起来，是一棵树。类似地，组织机构、家族谱系、学科分支、因特网、通信网络等都能表示成一个树图。

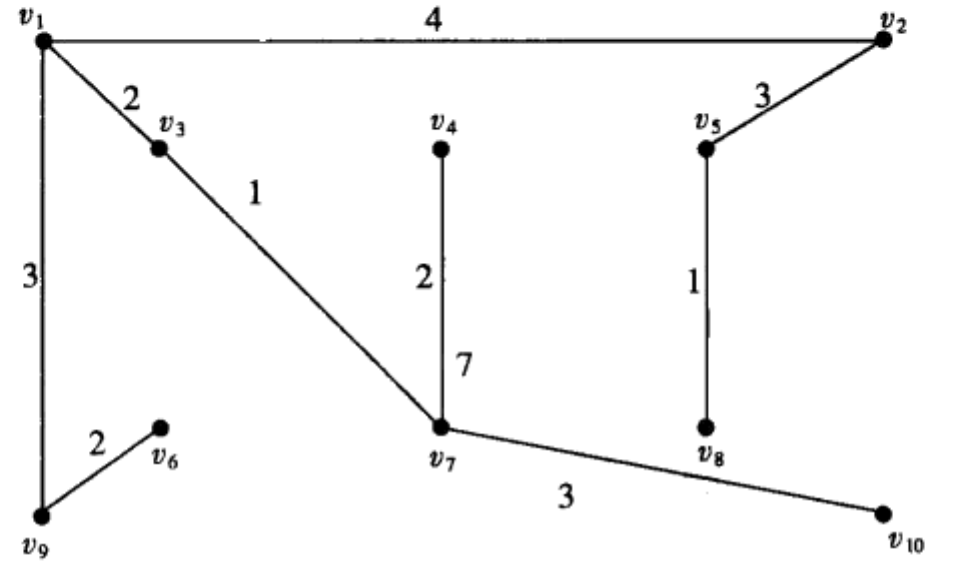


图 5.5

二、最小部分树

如果 G_1 是 G_2 的部分图，又是树图，则称 G_1 是 G_2 的部分树(或支撑树)。树图的各边称为树枝，一般 G_2 含有多个部分树，其中树枝总长最小的部分树，称为该树图的最小部分树(也叫最小支撑树)(Minimal Spanning Tree)。最小树问题就是求取一个树图的部分树(即支撑树)，使其总权数(总距离、总费用等) 为最少的一类优化问题。

(一) 最小树的求解方法

最常见的求解方法是破圈法和避圈法，前者适合于手工计算，后者适合于计算机求解。

1. 破圈法

破圈法也称为管梅谷算法，是由我国复旦大学的管梅谷教授于 1955 年提出的。其基本思想是在保证网络图连通的前提下，用削减长边的方法减少圈的个数，直到网络图无圈为止。破圈法的具体步骤如下：

（1）在给定的赋权的连通图上任找一个圈；

（2）在所找的圈中去掉边权数最大的一条边（如果有两条或两条以上的边都是权数最大的边，则任意去掉其中一条）；

（3）如果所余下的图已不含圈，则计算结束，所余下的图即为最小部分树，否则返回步骤（1）。

例 5.2 学院办公室计算机联网问题。某大学准备对其所属的 7 个学院办公室的计算机联网，这个网络的可能连通的途径如图 5.6 所示，图中 $v_1,\cdots,v_7$ 表示 7 个学院办公室，图中的边为可能联网的途径，边上所赋的权数为这条路线的长度，单位为百米。请设计一个网络能连通 7 个学院办公室，并使总的线路长度为最短。

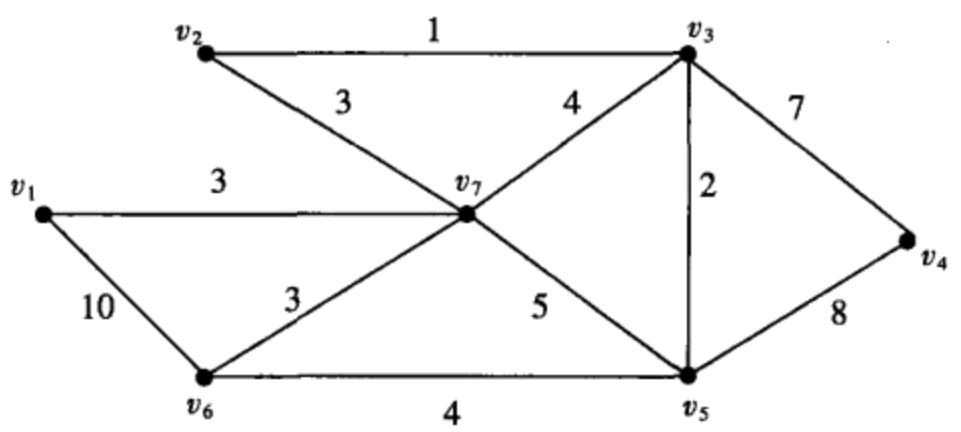


图 5.6

解：此问题实际上是求图 5.6 的一个最小部分树。下面我们用破圈法来计算：

（1）在 G 中找到一个圈 (v_1,v_7,v_6,v_1)，并知在此圈上边 $[v_1, v_6]$ 的权数 10 为最大，在 G 中去掉边 $[v_1, v_6]$ 得 G_1 图，如图 5.6(a) 所示；

（2）在 G_1 中找到一个圈 (v_3,v_4,v_5,v_7,v_3)，去掉其中权数最大的边 $[v_4, v_5]$ 得图 G_3，如图 5.6(b) 所示；

（3）在 G_2 中找到一个圈 (v_2,v_3,v_5,v_7,v_2)，去掉其中权数最大的边 $[v_5, v_7]$ 得图 G_3，如图 5.6(c) 所示；

（4）在 G_3 中找到一个圈 (v_3,v_5,v_6,v_7,v_3)，去掉其中权数最大的边 $[v_5, v_6]$（也可以去掉边 $[v_3, v_7]$）得图，如图 5.6(d) 所示；

（5）在 G_4 中找到一个圈 (v_2,v_3,v_7,v_2)，去掉其中权数最大的边 $[v_3, v_7]$ 得图 G_5，如图 5.6(e) 所示；

（6）在 G_5 中已找不到任何一个圈了，可知图 5.6(e) 即为图 G 的一个最小部分树，这个最小部分树的所有边的总权数为 $3+3+3+1+2+7=19$。

按照图 5.6(e) 的设计，可使此网络的总的线路长度为最短，为 19（百米）。

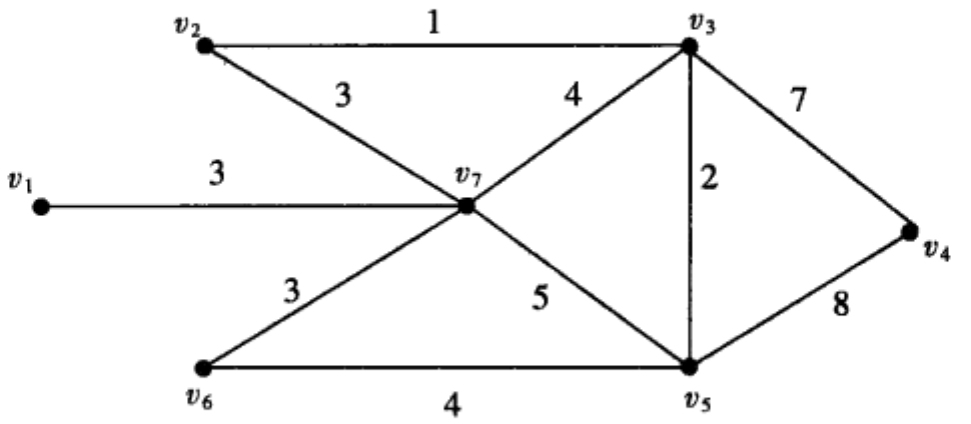


图 5.6(a)

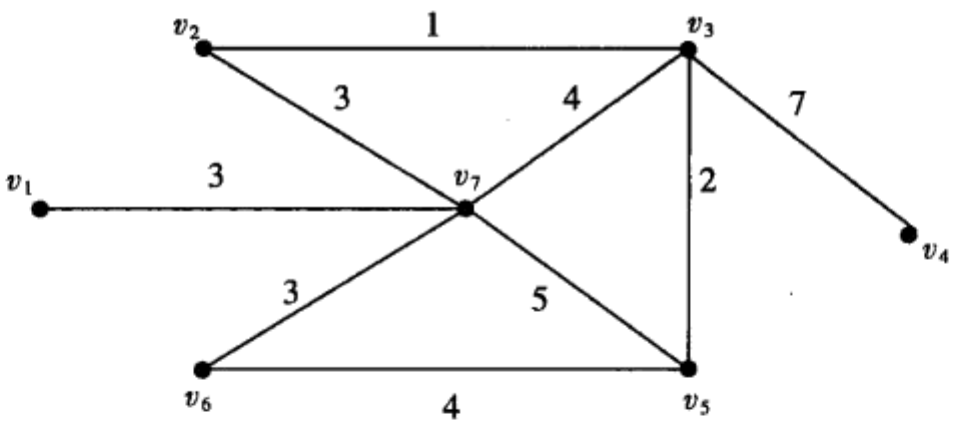


图 5.6(b)

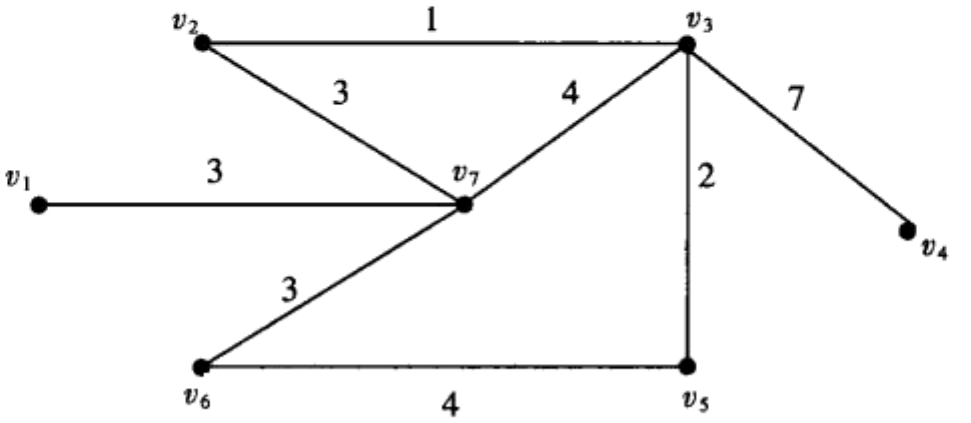


图 5.6(c)

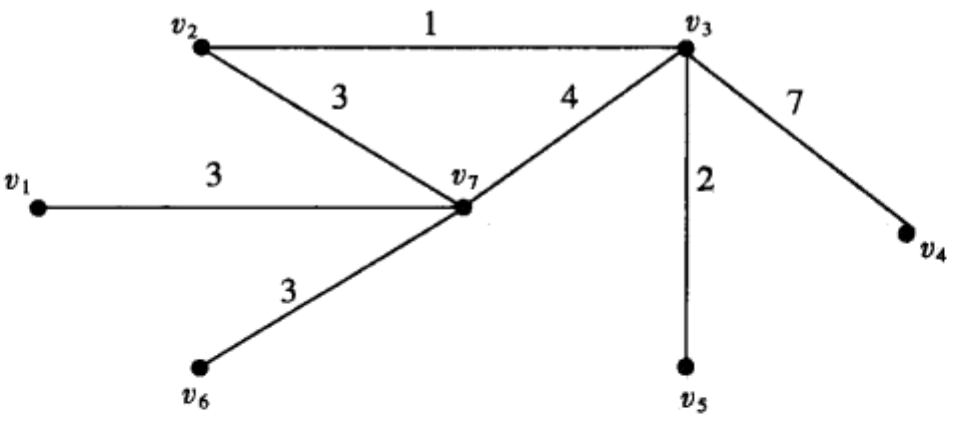


图 5.6(d)

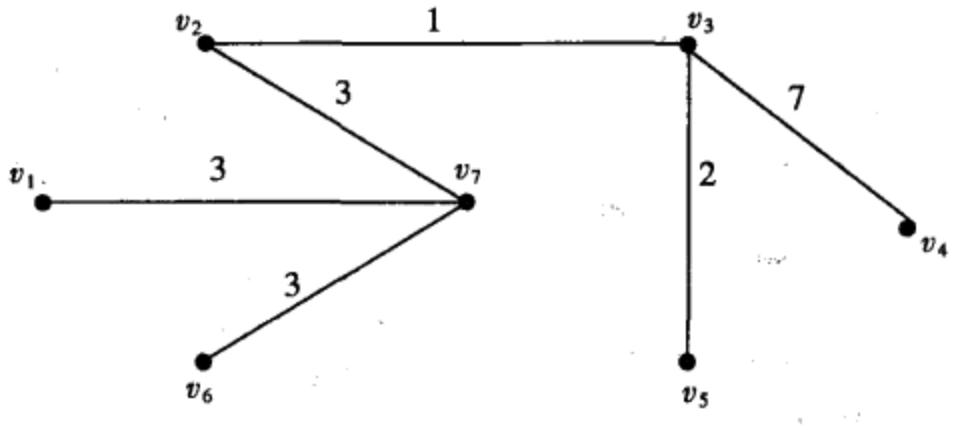


图 5.6(e)

2. 避圈法

避圈法也叫做 Kruskal 算法，其思想是开始选一条最小权的边，以后每一步总从未被选取的边中选一条权最小的边，并使之与已选取的边不构成圈（如果有两条或两条以上的边都是权数最小的边，则从中任选一条）。避圈法的具体步骤如下：

（1）从图中任选一点 v_i，让 $v_i \in V$，图中其余点均包含在 $\bar{V}$ 中；

（2）从 V 与 $\bar{V}$ 的连线中找出一条最小边，若有两条或两条以上的边相等，任选一条即可，则这条边一定包含在最小部分树内；

（3）令 $V \cup v_i \Rightarrow V$，$\bar{V} \backslash v_i \Rightarrow \bar{V}$；

（4）重复（2）（3）两步，直到图中所有点均包含在 V 中为止。

例 5.3 用避圈法求解例 5.2 的学院办公室计算机联网问题。

解：任选点 v_1 开始，离点 v_1 最近，未连通且边的权数最小的是点 v_7，连通 v_1v_7；离点 v_1 或点 v_7 最近未连通且边的权数最小的是点 v_6 和点 v_2（与点 v_7 连通），连通 v_7v_2 和 v_7v_6；离点 v_1、点 v_2、点 v_6 或点 v_7 最近，未连通且边的权数最小的是点 v_3（与点 v_2 连通），连通 v_2v_3；离点 v_1、点 v_2、点 v_3、点 v_6 或点 v_7 最近，未连通且边的权数最小的是点 v_5（与点 v_3 连通），连通 v_3v_5；未连通点只有点 v_4（和点 v_3 最近，未连通且边的权数最小），连通 v_3v_4。最小部分树长为 $3+3+3+1+2+7=19$。与破圈法所得结果相同。

三、最小树的 WinQSB 软件应用举例

在 WinQSB 软件的网络流模块中，最小部分树的生成采用的是上面的破圈法。本章几乎所有 WinQSB 软件计算都要调用子程序 Network Modeling。下面以例 5.1 为示例，说明怎样应用 WinQSB 软件计算最小部分树。

首先，调用 WinQSB 软件的子程序 Network Modeling，建立一个新问题，弹出对话框，如图 5.7 所示界面，选择“Minimal Spanning Tree”，输入问题的文件名“Tree1”（读者自己可以任意取名），节点数为 10。然后，点击“OK”，此时弹出一张需要输入数据的表格，对照图 5.4 输入表 5.1 所示的数据，两点间的权数只输一次（上三角）。点击菜单栏“Solve and Analyze”，输出表 5.2 所示的最小部分树结果，最小树长为 21。点击菜单栏“Results” → “Graphic Solution”，显示最小部分树的图形，见图 5.8，图 5.5 与该结果相同，正好是一个最小部分树。

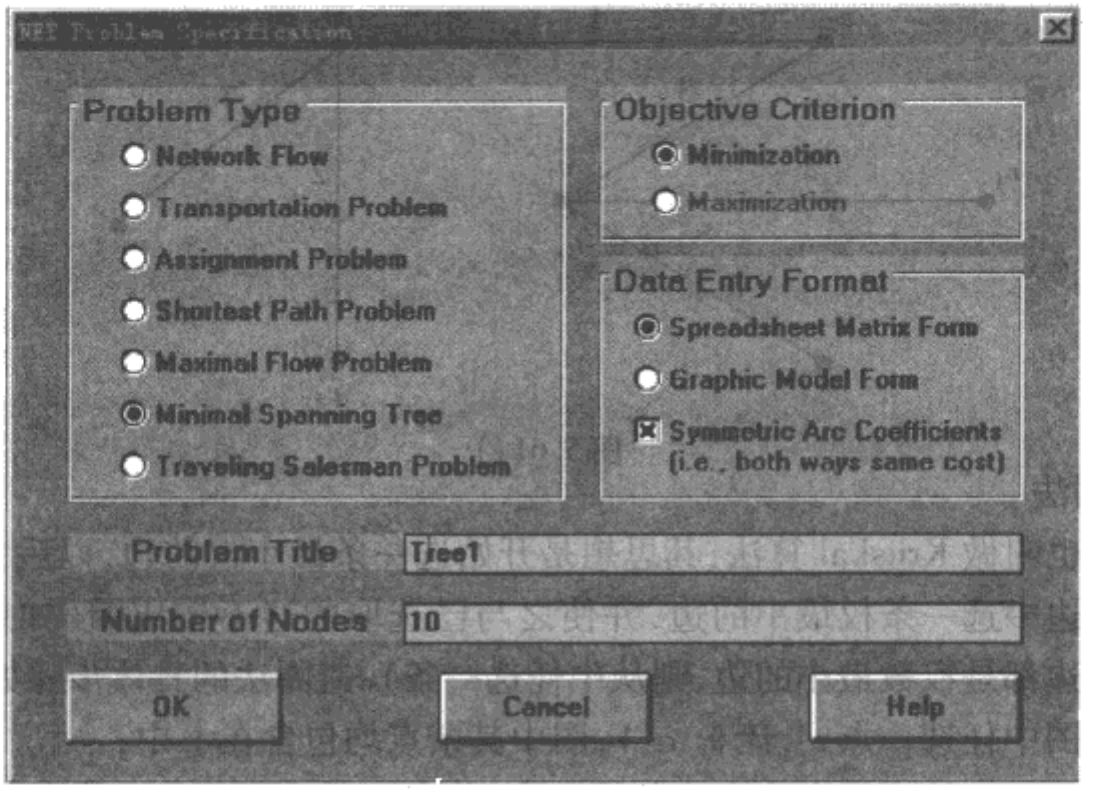


图 5.7

表 5.1

From \ To	Node1	Node2	Node3	Node4	Node5	Node6	Node7	Node8	Node9	Node10
Node1		4	2			6			3	
Node2					3			5		4
Node3				3		4	1			
Node4					7	8	2	5		
Node5							7	1		
Node6							8		2	
Node7								6	3	3
Node8										6
Node9										7
Node10										

表 5.2

01-21-2006	From Node	Connect To	Distance/Cost		From Node	Connect To	Distance/Cost
1	Node1	Node2	4	6	Node3	Node7	1
2	Node1	Node3	2	7	Node5	Node8	1
3	Node4	Node7	2	8	Node1	Node9	3
4	Node2	Node5	3	9	Node7	Node10	3
5	Node6	Node9	2				
	Total	Minimal	Connected	Distance	or Cost	=	21

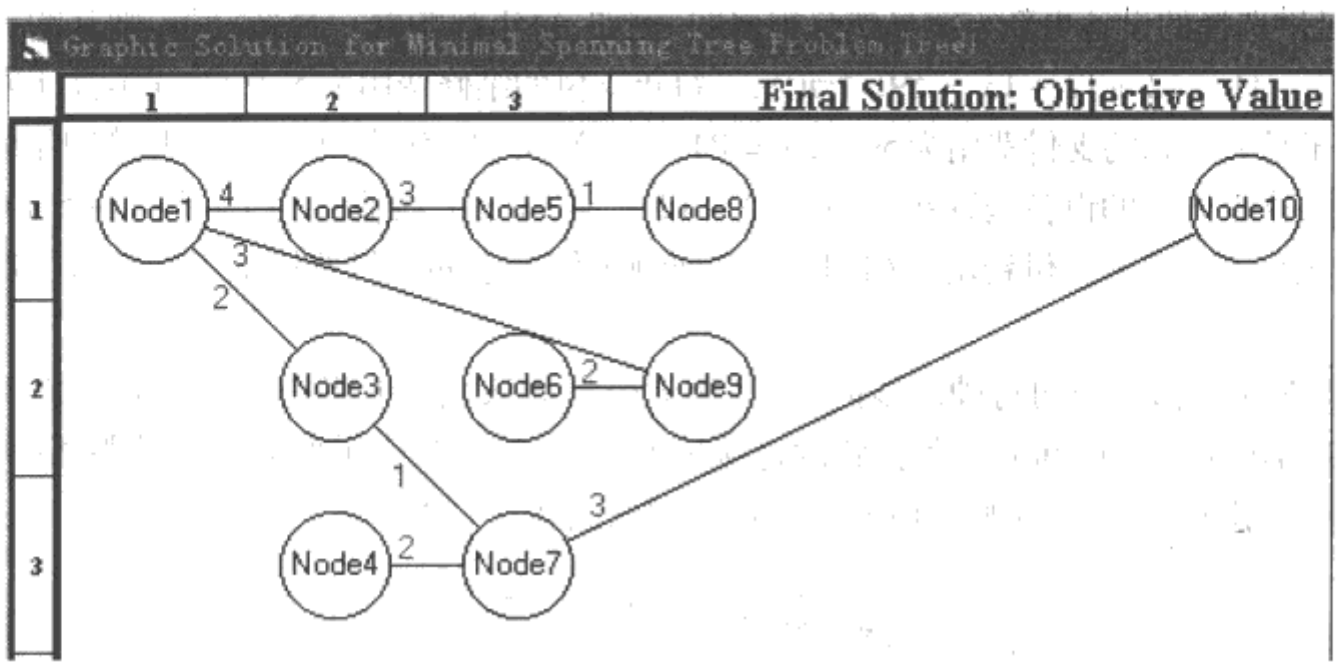


图 5.8

第三节　最短路问题

在网络理论中，最短路问题(Shortest Path Problem)是在实际应用中最广泛的问题之一，如设备更新、管道铺设、线路安排、厂区布局等。最短路问题也是动态规划的重要研究对象，但是，动态规划所能解决的最短路问题有一个重要的前提条件，即过程必须具有明确的阶段性。对于没有明显阶段性的最短路问题，要构造动态规划模型就比较困难，而使用图和网络则比较有效。本节主要介绍用图和网络的方法解决一般的最短路问题。

一、最短路问题的网络模型

例 5.4　设备更新问题。某公司使用一台设备，在每年年初，公司就要决定是购买新的设备还是继续使用旧设备。如果购置新设备，就要支付一定的购置费，当然新设备的维修费用就低。如果继续使用旧设备，这样就可以省去购置费，但维修费用就会增加。现在需要制定一个五年之内的设备更新计划，使得五年内购置费和维修费总的支付费用最少。该设备每年年初的价格所表 5.3 所示。

表 5.3　　单位：千元

年份	1	2	3	4	5
年初价格	11	11	12	12	13

使用不同时间(年)的设备所需要的维修费用如表 5.4 所示：

表 5.4　　单位：千元

年份	0 ~ 1	1 ~ 2	2 ~ 3	3 ~ 4	4 ~ 5
每年维修费用	5	6	8	11	18

可供选择的设备更新方案显然是很多的。例如，每年都购置一台新设备，则其购置费用为 11 + 11 + 12 + 12 + 13 = 59(千元)，而每年支付的维修费用为 5 千元，五年合计为 25 千元，于是五年总的支付费用为 59 + 25 = 84(千元)。又如决定在第一、第二和第四年各购进一台，这个方案的设备购置费为 11 + 11 + 12 = 34(千元)，维修费为 5 + 5 + 6 + 5 + 6 = 27(千元)，五年总的支付费用为 61 千元。如何制定使得总的支付费用最少的设备更新计划呢?

例 5.5 物流公司行驶路线问题。某物流公司欲从城市 1 将一批货物运往城市 7，图 5.9 给出了两城市间的各条公路的连接情况，图中的权数表示两城市间的距离(可以是行驶费用、时间等)，驾驶员该如何选择行驶的路线?

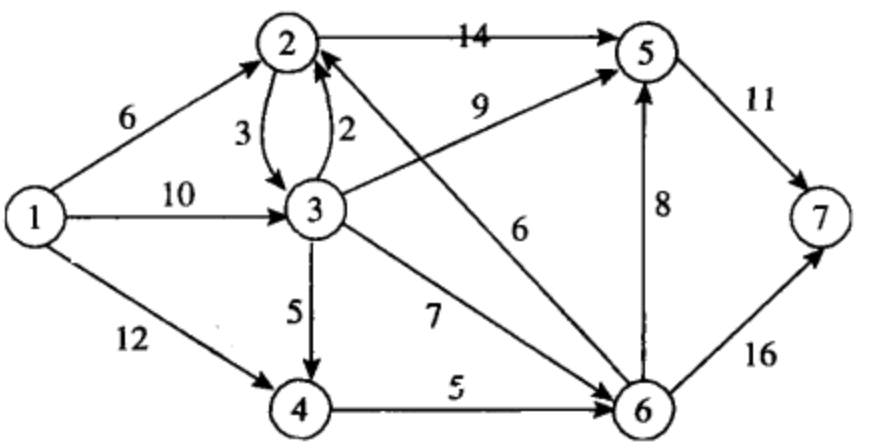


图 5.9

从这两个例子，可以引出一般的最短路问题：给定一个赋权有向图(网络) $N = (V,A,C)$，设 P 是 $N = (V,A,C)$ 中从 v_s 到 v_t 的一条路，定义路 P 的权是 P 中所有弧的权之和，记为 $w(P)$，则最短路问题就是要在所有 v_s 到 v_t 的路中，求一条从 v_s 到 v_t 的路 P_0，使得：

$$w(P_0) = \min_P w(P)$$

则 P_0 就是从 v_s 到 v_t 的最短路。

最短路问题是重要的最优化问题之一，它不仅可以直接应用于解决许多实际问题，如设备更新、管道铺设、线路安排、厂区布局等，而且还经常被用作一个重要的基本工具解决其他的优化问题。

二、最短路的求解方法

本部分主要介绍最短路的求解算法：Dijkstra 算法和 Floyd 算法以及最短路问题在 WinQSB 软件上的计算实现。

(一) 最短路的 Dijkstra 算法

1. 有向图的 Dijkstra 算法

Dijkstra 算法是目前解决网络最短路问题最有效的算法之一，适用于每条弧的赋权数 C_{ij} 都大于或等于零的情况，Dijkstra 算法也称为双标号法。所谓双标号，也就是对网络中的点 v_j 赋予两个标号 (l_j,k_j)，第一个标号 l_j 表示从起点 v_s 到 v_j 的最短路的长度，第二个标号 k_j 表示在 v_s 至 v_j 的最短路上 v_j 前面一个邻点的下标，从而找到 v_s 到 v_t 的最短路及 v_s 与 v_t 的距离。算法的基本思想是从始点出发，逐步顺序地向外探寻，要求每向外延伸一步都要求路是最短的。具体步骤如下：

(1) 给起点 v_1 以标号$(0,s)$,表示从 v_1 到 v_1 的距离为0,v_1 为起点。

(2) 找出已标号的点的集合 I,没标号的点的集合 J,以及弧的集合$\{(v_i,v_j) \mid v_i \in I, v_j \in J\}$,这里弧的集合是指所有从已标号的点到未标号的点的弧的集合。

(3) 若上述弧的集合是空集,则计算结束;若 v_t 已标号(l_t,k_t),则 v_s 到 v_t 的距离即为 l_t,而从 v_s 到 v_t 的最短路可以从 v_t 反向追踪到起点 v_s 得到;若 v_t 未标号,则不存在从 v_s 到 v_t 的有向路。

(4) 对上述弧的集合中的每一条弧,计算 $S_{ij} = l_i + c_{ij}$。在所有的 S_{ij} 中,找到其值为最小的弧,不妨设此弧为(v_c,v_d),则给此弧的终点以双标号(S_{cd},c),返回步骤(2)

若在步骤(4) 中,使得 S_{ij} 值为最小的弧不止一条,则可以任选一个标定,也可以都予以标定,若这些弧中的有些弧的终点为同一点,则此点应有多个双标号,以便最后找到多个最短路。

例 5.6 用 Dijkstra 算法求解图 5.9 所示的城市 1 到城市 7 的最短路及最短路长。

解:(1) 给起始点 v_1 以标号$(0,S)$,表示从 v_1 到 v_1 的距离为0,v_1 为起点。

(2) 这时已标定点集合 $I = \{v_1\}$,未标定点的集合 $J = \{v_2,v_3,v_4,v_5,v_6,v_7\}$,弧集合$\{(v_i,v_j) \mid v_i \in I, v_j \in J\} = \{(v_1,v_2),(v_1,v_3),(v_1,v_4)\}$,并有:

$S_{12} = l_1 + c_{12} = 0 + 6 = 6,$

$S_{13} = l_1 + c_{13} = 0 + 10 = 10,$

$S_{14} = l_1 + c_{14} = 0 + 12 = 12,$

$\text{Min}\{S_{12},S_{13},S_{14}\} = S_{13} = 6$。

这样,我们给弧(v_1,v_2) 的终点标以(6,1),表示从 v_1 到 v_2 的距离为6,并且在 v_1 到 v_2 的最短路径中 v_2 的前面一个点是 v_1。

(3) 这时 $I = \{v_1,v_2\}$,$J = \{v_3,v_4,v_5,v_6,v_7\}$,弧集合$\{(v_i,v_j) \mid v_i \in I, v_j \in J\} = \{(v_1,v_3),(v_1,v_4),(v_2,v_3),(v_2,v_5)\}$,并有:

$S_{23} = l_2 + c_{24} = 6 + 3 = 9,$

$S_{25} = l_2 + c_{25} = 6 + 16 = 20,$

$\text{Min}\{S_{13},S_{14},S_{23},S_{25}\} = S_{23} = 9$。

这样,我们给弧(v_1,v_2) 的终点标以(6,1),表示从 v_1 到 v_2 的距离为6,并且在 v_1 到 v_2 的最短路径中 v_2 的前面一个点是 v_1;我们给弧(v_2,v_3) 的终点标以(9,2),表示从 v_1 到 v_3 的距离为9,并且在 v_1 到 v_3 的最短路径中 v_3 的前面一个点是 v_2。

(4) 这时 $I = \{v_1,v_2,v_3\}$,$J = \{v_4,v_5,v_6,v_7\}$,弧集合$\{(v_i,v_j) \mid v_i \in I, v_j \in J\} = \{(v_1,v_4),(v_2,v_5),(v_3,v_4),(v_3,v_5),(v_3,v_6)\}$,并有:

$S_{34} = l_3 + c_{34} = 9 + 5 = 14,$

$S_{35} = l_3 + c_{35} = 9 + 9 = 18,$

$S_{36} = l_3 + c_{36} = 9 + 7 = 16,$

$\text{Min}\{S_{14},S_{25},S_{34},S_{35},S_{36}\} = S_{14} = 12$。

这样,我们给弧(v_1,v_4) 的终点 v_4 标以(12,1),表示从 v_1 到 v_4 的距离为12,并且在 v_1 到 v_4 的最短路径中 v_4 的前面一个点是 v_1。

(5) 这时 $I = \{v_1,v_2,v_3,v_4\}$,$J = \{v_5,v_6,v_7\}$,弧集合$\{(v_i,v_j) \mid v_i \in I, v_j \in J\}$

$= \{(v_2,v_5),(v_3,v_5),(v_3,v_6),(v_4,v_6)\}$，并有：

$S_{46} = l_4 + c_{46} = 12 + 5 = 17$，

$\text{Min}\{S_{25},S_{35},S_{36},S_{46}\} = S_{36} = 16$。

这样，我们给弧(v_3,v_6)的终点v_6标以(16,3)，表示从v_3到v_6的距离为16，并且在v_3到v_6的最短路径中v_6的前面一个点是v_3。

(6) 这时$I = \{v_1,v_2,v_3,v_4,v_6\}$，$J = \{v_5,v_7\}$，弧集合$\{(v_1,v_j) \mid v_i \in I, v_j \in J\} = \{(v_2,v_5),(v_3,v_5),(v_6,v_5),(v_6,v_7)\}$，并有：

$S_{65} = l_6 + c_{65} = 16 + 8 = 24$，

$S_{67} = l_6 + c_{67} = 16 + 16 = 32$

$\text{Min}\{S_{25},S_{35},S_{65},S_{67}\} = S_{35} = 18$。

这样，我们给弧(v_3,v_5)的终点v_5标以(18,3)，表示从v_3到v_5的距离为18，并且在v_3到v_5的最短路径中v_5的前面一个点是v_3。

(7) 这时$I = \{v_1,v_2,v_3,v_4,v_5,v_6\}$，$J = \{v_7\}$，弧集合$\{(v_i,v_j) \mid v_i \in I, v_j \in J\} = \{(v_6,v_7),(v_5,v_7)\}$，并有：

$S_{57} = l_5 + c_{57} = 18 + 11 = 29$，

$\text{Min}\{S_{57},S_{67}\} = S_{57} = 29$。

这样，我们给弧(v_5,v_7)的终点v_7标以(29,5)，表示从v_5到v_7的距离为18，并且在v_5到v_7的最短路径中v_7的前面一个点是v_5。

(8) 由于终点已经标号，说明已经得到从v_1到v_7的最短路，计算结束。v_1到v_7的最短路为：$p_{17} = \{v_1,v_2,v_3,v_5,v_7\}$，最短路长为$L_{17} = 29$。

2. 无向图的 Dijkstra 算法

如果v_i与v_j之间存在一条无方向的边相关联，这说明v_i与v_j两点之间可以到达。当v_i与v_j之间至少有两条边相关联时，留下一条最短边，去掉其他关联边。对于无向图最短路问题的求解 Dijkstra 算法同样适用。标号方法与有向图相同。

例 5.7 用 Dijkstra 算法求解图 5.10 所示的城市 1 到其他各城市的最短路。

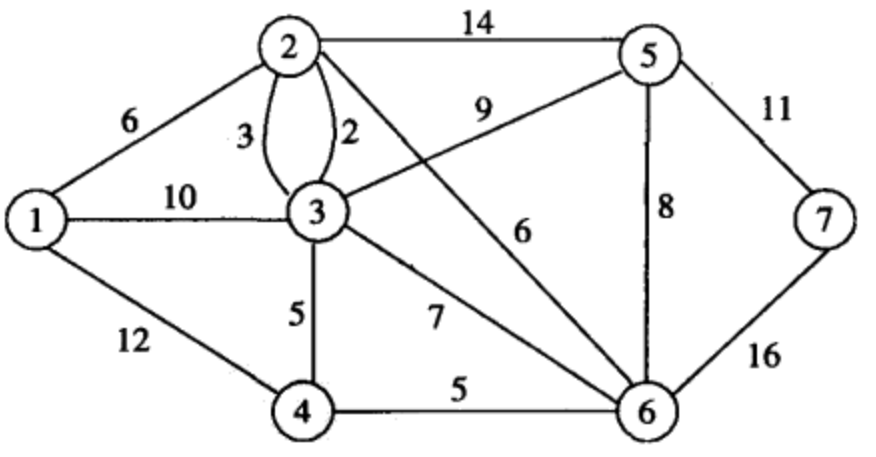


图 5.10

解：由于计算算法相同，我们就省略掉详细计算过程，求解结果为：

$v_1 \to v_2$，最短路长 6；

$v_1 \to v_2 \to v_3$，最短路长 8；

$v_1 \to v_4$，最短路长 12；

$v_1 \to v_2 \to v_3 \to v_5$，最短路长 17；

$v_1 \to v_2 \to v_6$，最短路长 12；

$v_1 \to v_2 \to v_3 \to v_5 \to v_7$，最短路长 28。

（二）最短路的 Floyd 算法

Floyd 算法是更一般的求解算法，该算法是一种矩阵（表格）迭代方法，对于求网络中任意两点间最短路、混合网络（部分边有方向、部分边无方向）的最短路和有负权网络（$c_{ij} < 0$）的最短路等一般网络问题来说比较有效。

Floyd 算法具体步骤如下：

（1）写出 v_i 一步到达 v_j 的距离矩阵 $L_1 = [L_{ij}^{(1)}]$，L_1 也是一步到达的最短距离矩阵。若 v_i 与 v_j 之间没有关联边，则令 $c_{ij} = +\infty$。

（2）计算两步最短距离矩阵。设 v_i 经过一个中间点 v_r 两步到达 v_j，则 v_i 到 v_j 的最短距离为 $L_{ij}^{(2)} = \text{Min}\{c_{ir} + c_{rj}\}$，最短距离矩阵为 $L_2 = [L_{ij}^{(2)}]$。

（3）计算 k 步最短距离矩阵。设 v_i 经过一个中间点 v_r 到达 v_j，v_i 经过 $k-1$ 步到达点 v_r 的最短距离为 $L_{ir}^{(k-1)}$，v_r 经过 $k-1$ 步到达点 v_j 的最短距离为 $L_{rj}^{(k-1)}$，则 v_i 经 k 步到 v_j 的最短距离为 $L_{ir}^{(k)} = \text{Min}\{L_{rj}^{(k-1)}\}$，最短距离矩阵为 $L_k = [L_{ij}^{(k)}]$。

（4）比较矩阵 L_k 与 L_{k-1}，当 $L_k = L_{k-1}$ 时得到任意两点间的最短距离矩阵 L_k。

当图的点数为 n 并且 $c_{ij} \geqslant 0$ 时，迭代次数 k 由公式 $2^{k-1} < n-1 \leqslant 2^k$ 或 $k-1 < \dfrac{\lg(n-1)}{\lg 2} \leqslant k$ 估计得到。

例 5.8 求图 5.11 中任意两点间的最短距离。

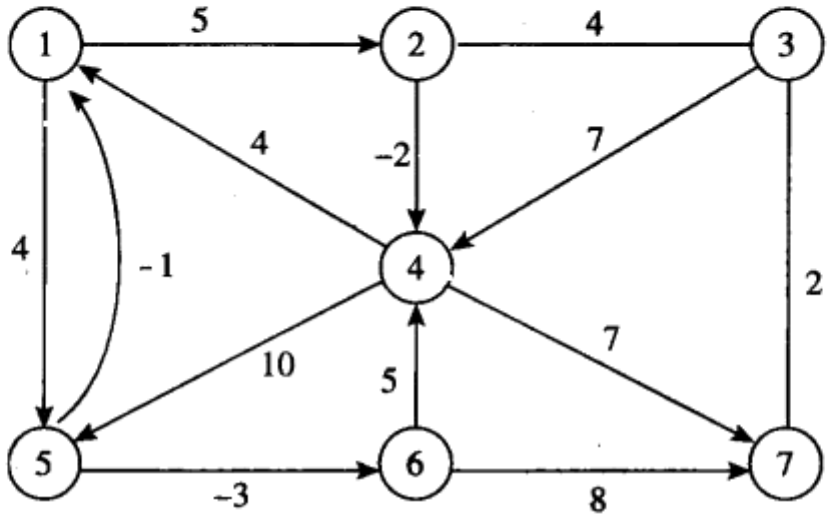


图 5.11

解：图 5.11 是一个混合网络，有 3 条边的权是负数，有两条边无方向。根据图 5.11，写出任意两点间一步到达距离表 L_1，表中第一列的点表示弧的起点，第一行的点表示弧的终点，无方向的边表明可以互相到达，如表5.5 所示。计算过程见表5.6 至表 5.8。

表 5.5

	v_1	v_2	v_3	v_4	v_5	v_6	v_7
v_1	0	5	∞	∞	4	∞	∞
v_2	∞	0	4	-2	∞	∞	∞
v_3	∞	4	0	7	∞	∞	2
v_4	4	∞	∞	0	10	∞	7
v_5	-1	∞	∞	∞	0	-3	∞
v_6	∞	∞	∞	5	∞	0	8
v_7	∞	∞	2	∞	∞	∞	0

表 5.6

	v_1	v_2	v_3	v_4	v_5	v_6	v_7
v_1	0	5	9	3	4	1	∞
v_2	2	0	4	-2	8	∞	5
v_3	11	4	0	2	17	∞	2
v_4	4	9	9	0	8	7	7
v_5	-1	4	∞	2	0	-3	5
v_6	9	∞	10	5	15	0	8
v_7	∞	6	2	4	12	9	0

表 5.7

	v_1	v_2	v_3	v_4	v_5	v_6	v_7
v_1	0	5	9	3	4	1	9
v_2	2	0	4	-2	6	3	5
v_3	6	4	0	2	10	9	2
v_4	4	9	9	0	8	5	7
v_5	-1	4	7	2	0	-3	5
v_6	9	14	10	5	13	0	8
v_7	8	6	2	4	14	16	0

表 5.8

	v_1	v_2	v_3	v_4	v_5	v_6	v_7
v_1	0	5	9	3	4	1	9
v_2	2	0	4	-2	6	3	5
v_3	6	4	0	2	10	7	2
v_4	4	9	9	0	8	5	7
v_5	-1	4	7	2	0	-3	5
v_6	9	14	10	5	13	0	8
v_7	8	6	2	4	14	9	0

经计算 $L_4 = L_5$，L_4 是最优表，从表上可以得到任意两点间的最短距离。

三、最短路的 WinQSB 软件应用举例

在 WinQSB 软件的网络流模块中，最短路问题的求解采用的是上面介绍的标号法。下面我们以例 5.4 的设备更新问题为示例，说明怎样应用 WinQSB 软件计算最短路问题。

对于例 5.4 的设备更新问题，我们先要把它转化为一个标准的最短路问题的网络模型。用点 v_i 表示“第 i 年年初购进一台新设备”这种状态，假设一点 v_6，表示第 5 年年底。从 v_i 到 v_{i+1}，…，v_6 各画一条弧，弧 (v_i, v_j) 表示在第 i 年年初购进的设备一直使用到第 j 年年初（即第 $j-1$ 年年底）。每条弧的权数为第 i 年年初购进的设备使用到第 $j-1$ 年年底所花费的购置费及维修费的总和。例如，弧 (v_1, v_4) 的权数应为第 1 年年初购置设备的价格 11 与从第 2 年年初到第 3 年年底两年的维修费用（5 + 6 = 11）之和，应为 22；而弧 (v_1, v_6) 的权数应为第 1 年年初购置设备的价格 11 与从第 1 年年初到第 5 年年底五年的维修费用（5 + 6 + 8 + 11 + 18 = 48）之和，应为 59；如此等等，就可以计算出弧 (v_i, v_j) 的权数 c_{ij}。这样，设备更新问题就等价于寻求图 5.12 中从 v_1 到 v_6 的最短路问题。

下面我们应用 WinQSB 软件来求解设备更新问题。首先，调用 WinQSB 软件的子程序 Network Modeling，建立一个新问题，弹出对话框，如图 5.13 所示界面，选择“Shortest Path Problem”，输入问题的文件名“Equip1”（读者自己可以任意取名），节点数 6。然后，点击“OK”，此时弹出一张需要输入数据的表格，对照图 5.12 输入图

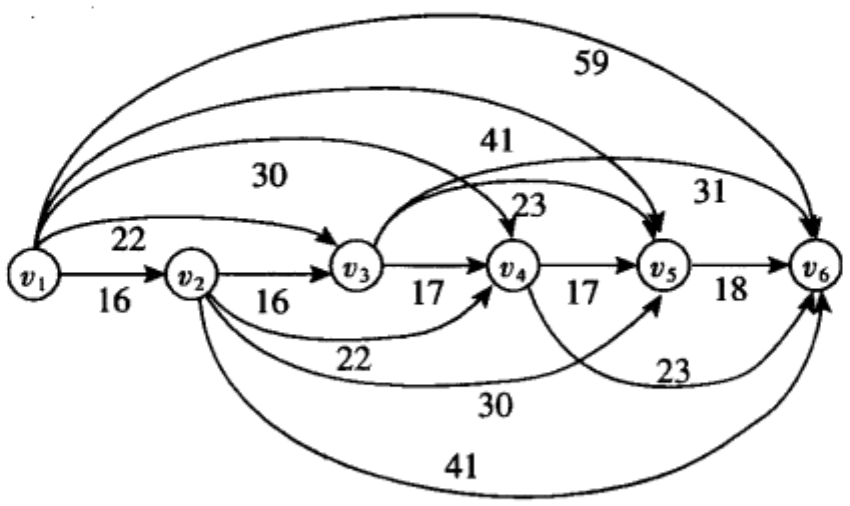


图 5.12

5.15 所示的数据，两点间的权数只输一次（上三角）。在这里，本问题是一个有向图，是按照弧的方向输入数据的，如果是无向图，每一条边就必须输入两次，把无向边改变为两条方向相反的弧。点击菜单栏“Solve and Analyze”后，WinQSB 软件系统提示用户选择图的起点和终点，系统默认第一个点为起点，最后一个点为终点，如图 5.14 所示，根据本问题，起点为节点 1，终点为节点 6。再点击“Solve”按钮，系统输出 v_1 到 v_6 的最短路路径和最短路路长，如表 5.16 所示的最短路求解输出结果，最短路路径为 $v_1 \rightarrow v_3 \rightarrow v_6$，最短路路长为 53。点击菜单栏“Results”→“Graphic Solution”，显示最短路的图形，见图 5.17。这里，请大家注意的是，WinQSB 软件系统只给出本问题的一个求解结果，实际上，$v_1 \rightarrow v_4 \rightarrow v_6$ 也是一条最短路，最短路路长同样为 53。这说明，对于例 5.4 的设备更新问题，使得总的支付费用最少的设备更新方案有两个：一个方案为第 1 年年初购置的新设备使用到第 2 年年底（第 3 年年初），第 3 年年初再购置新设备使用到第 5 年年底（第 6 年年初）；另一个方案为第 1 年年初购置的新设备使用到第 3 年年底（第 4 年年初），第 4 年年初再购置新设备使用到第 5 年年底（第 6 年年初）。这两个方案使得总的支付为最小，均为 53 千元。

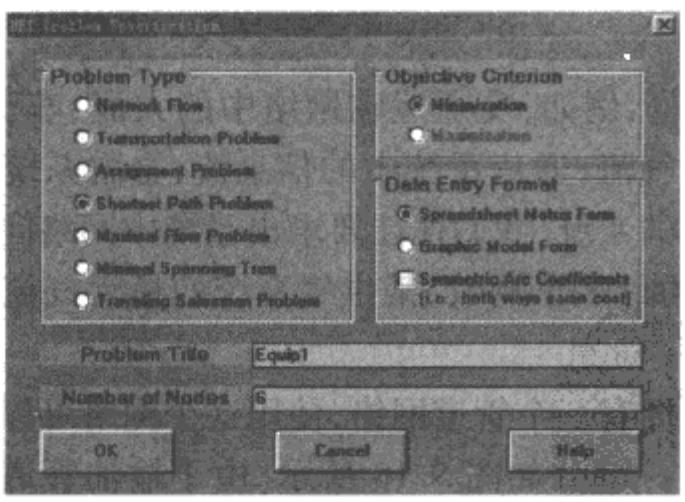


图 5.13

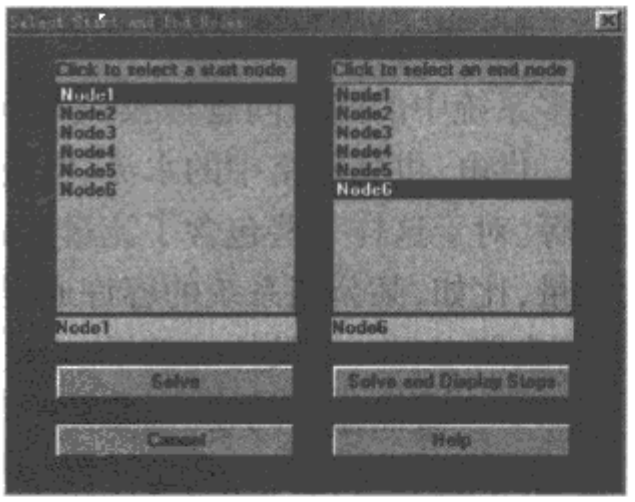


图 5.14

From \ To	Node1	Node2	Node3	Node4	Node5	Node6
Node1		16	22	30	41	59
Node2			16	22	30	41
Node3				17	23	31
Node4					17	23
Node5						18
Node6						

图 5.15

From	To	Distance/Cost	Cumulative Distance/Cost
Node1	Node3	22	22
Node3	Node6	31	53
From Node1	To Node6	Distance/Cost	= 53

图 5.16

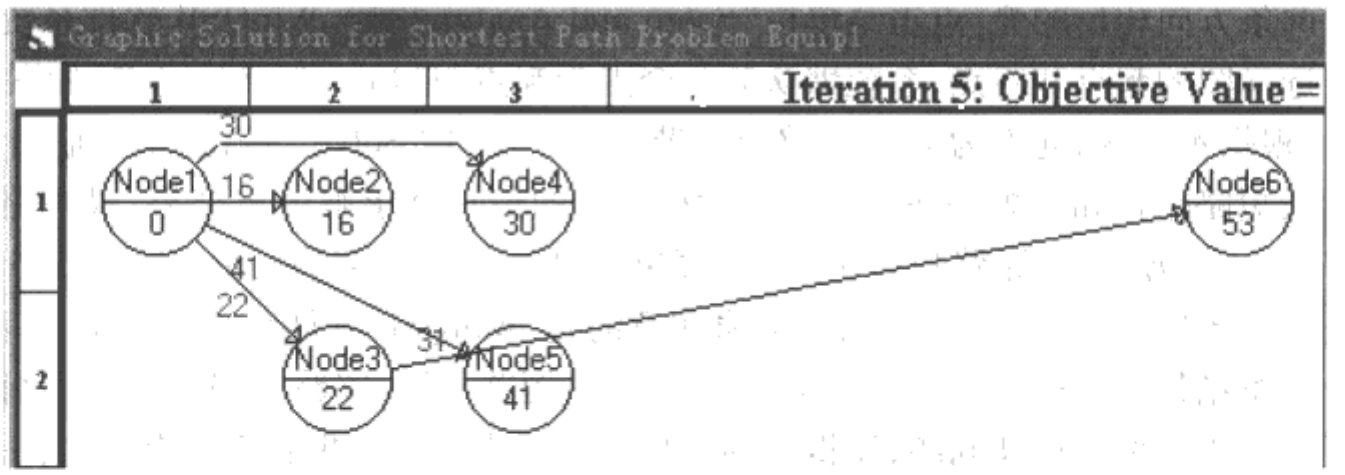


图 5.17

第四节　最大流问题

许多系统中的网络问题涉及流量问题。例如，交通运输网络中的人流、车流、物流，供水、供电、供气网络中的水流、电流、气流，通信系统中的信息流，金融系统中的现金流等。对于这样一些包含了流量问题的系统，我们往往要求出系统中网络问题的最大流量，比如，某公路系统的容许通过的最多车辆数，某供气系统的最大气流量等等。下面我们来看一个具体的例子。

例 5.9　石油输送问题。某石油公司拥有一个管道网络，使用这个网络可以把石油从采地运送到一些销售点，这个网络的一部分如图 5.18 所示。由于输油管道的直径的变化，它的各段管道(v_i, v_j)的流量（容量）c_{ij}也是不一样的，c_{ij}的单位为万加仑/小时。如果使用这个网络系统从采地 v_1 向销地 v_7 运送石油，每小时能运送多少加仑石油？

利用我们在第二章学过的线性规划问题，我们可以建立该问题的线性规划模型。设弧(v_i, v_j)上的流量为f_{ij}，网络上总的流量为 F，则有：

$$\text{Max } F = f_{12} + f_{14}$$

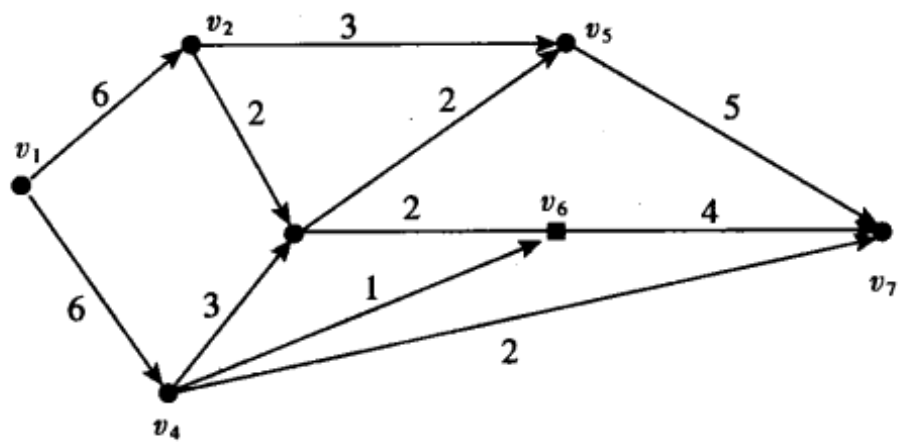


图 5.18

$$
s.t.\begin{cases}
f_{12} = f_{23} + f_{25} \\
f_{14} = f_{43} + f_{46} + f_{47} \\
f_{23} + f_{43} = f_{35} + f_{36} \\
f_{25} + f_{35} = f_{57} \\
f_{36} + f_{46} = f_{67} \\
f_{57} + f_{67} = f_{47} = f_{12} + f_{14} \\
f_{ij} \leqslant c_{ij}(i = 1,2,\cdots,6;\ j = 2,\cdots,7) \\
f_{ij} \geqslant 0(i = 1,2,\cdots,6;\ j = 2,\cdots,7)
\end{cases}
$$

不难发现,这一模型具有如下特点:

(1) 变量个数与有向图的弧的个数相同,约束条件个数等于有向图的点数减 1。

(2) 目标函数为总输出量最大。

(3) 约束条件方程主要是两类:一是对于每一个中间点,输入量 = 输出量;二是对于始点和终点,总输入量 = 总输出量。

(4) 流量的取值以每一弧的容量为上界。

可见这类问题是在满足容量限制条件和中间点平衡条件的情况下求取流量达到最大的可行流的一类优化问题。简言之,是求网络中具有最大流量值的可行流问题,即最大流(Maximum Flow Problem) 问题。

一、基本概念与基本定理

为了求解最大流问题,我们必须先明确一些概念。

定义 5.9 [流、可行流、最大流] 所谓流是指加在网络各条弧上的流量的集合,它是定义在弧集合上的一个函数 $f = \{f(v_i,v_j)\}$,其中 $f(v_i,v_j)$ 是弧 (v_i,v_j) 上的流量,简记为 f_{ij}。若网络上所有的 $f_{ij} = 0$,这个流成为零流。满足下面条件的流称作可行流:

(1) 容量限制条件:对每一条弧,有 $0 \leqslant f_{ij} \leqslant c_{ij}$;

(2) 平衡条件:对于中间点,输入量 = 输出量,即对于每个 $i(i \neq s,t)$,有 $\sum\limits_{(v_i,v_j)} f_{ij} = \sum\limits_{(v_j,v_i)} f_{ji}$;对于发点和收点,总输入量 = 总输出量,即 $\sum\limits_{j} f_{sj} = \sum\limits_{i} f_{it} = F$。

任何网络上一定存在可行流,比如,零流就是一个可行流。

最大流就是能使网络总流量 F 取最大值的可行流。

定义 5.10 ［弧的分类］网络中有各种各样的弧，按流量的特征可分为：

饱和弧 —— 弧上的流量等于容量，即有 $f_{ij} = c_{ij}$；

非饱和弧 —— 弧上的流量小于容量，即有 $f_{ij} < c_{ij}$；

零流弧 —— 弧上的流量等于零，即有 $f_{ij} = 0$；

非零流弧 —— 弧上的流量大于零，即有 $f_{ij} > 0$。

在网络中，从发点到收点的一条链中，按弧的方向可分为：

前向弧（正向弧）—— 与链的方向一致，前向弧的全体记为 μ^+；

后向弧（反向弧）—— 与链的方向相反，后向弧的全体记为 μ^-。

定义 5.11 ［增广链］设 f 是一个可行流，若存在一条从发点到收点的链，满足下列条件，则称该链为增广链，记为 μ。

（1）在所有前向弧上 $f_{ij} < c_{ij}$；

（2）在所有后向弧上 $f_{ij} > 0$。

例 5.10 图 5.19 中，链 $\mu = (v_1, v_2, v_3, v_4, v_6, v_7)$ 是一条增广链，因为 μ^+ 和 μ^- 中的弧满足增广链的条件。比如：$(v_1, v_2) \in \mu^+, f_{12} = 3 < c_{12} = 6$；$(v_4, v_3) \in \mu^-, f_{43} = 2 > 0$ 等。

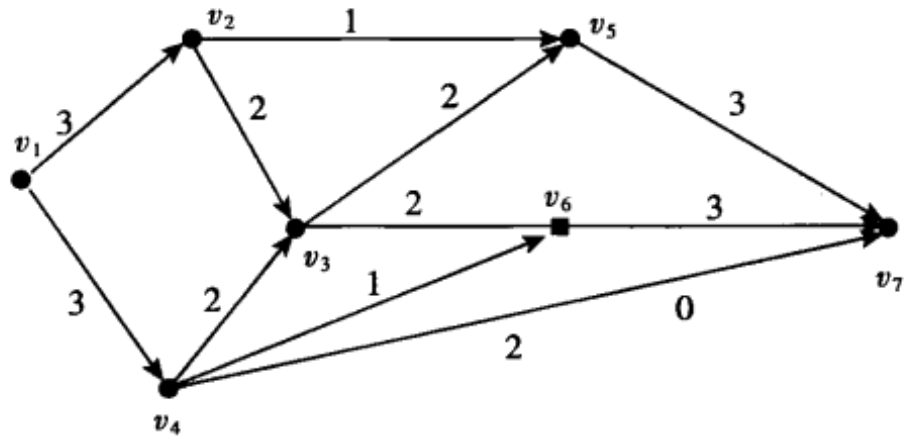


图 5.19

定义 5.12 ［割集、割量、最小割］给定网络 $N = (V, A, C)$，若点集 V 被分为两个非空集合 S 和 T，使得 $v_s \in S, v_t \in T$，且有 $S \cup T = V, S \cap T = \phi$，则把弧集 (S, T) 称为分离 v_s 和 v_t 的割集。割集 (S, T) 中所有弧的容量之和称为这个割集的容量，简称为割量。在所有割集中，容量最小的割集称为最小割。

定理 5.1 ［流量 — 割集定理］任何一个可行流的流量都不会超过任一割集的容量。

定理 5.2 ［最大流量 — 最小割集定理］在任一网络中，从 v_s 到 v_t 的最大流量等于它的最小割集的容量。

定理 5.3 ［充要条件］可行流 f^* 是最大流的充要条件是：网络中不存在关于 f^* 的增广链。

二、最大流的求解方法

常用的最大流的求解方法是 Ford – Fulkerson 标号算法，该方法是由 Ford 和 Fulkerson 于 1956 年提出的，故又称为 Ford – Fulkerson 标号算法。算法的实质是判断

在网络当中是否存在增广链，并设法把增广链找出来。算法的基本步骤如下：

第一步，首先给发点 v_s 标号$[0,\varepsilon(s)]$。括弧中第一个数字是使这个点得到标号的前一个点的代号，因为 v_s 是发点，故记为0。括弧中第二个数字表示 $\varepsilon(s)$ 从上一个标号点到这个标号点流量的最大允许调整值。v_s 为发点，调整值不受限制，故 $\varepsilon(s) = \infty$。

第二步，列出与已标号点相邻的所有标号点：

(1) 考虑从标号 i 点出发的弧(i,j)，若有 $f_{ij} = c_{ij}$，则不给点 j 标号；若有 $f_{ij} < c_{ij}$，则对点 j 标号，记为$[i,\varepsilon(j)]$。括弧中的 i 表示点的标号是从 i 点延伸过来的，$\varepsilon(j) = \text{Min}\{\varepsilon(i),(c_{ij} - f_{ij})\}$；

(2) 考虑所有指向标号点出发的弧(h,i)，若有 $f_{hi} = 0$，则不给点 h 标号；若有 $f_{hi} > 0$，则对点 h 标号，记为$[i,\varepsilon(h)]$，其中 $\varepsilon(h) = \text{Min}\{\varepsilon(i),f_{hi}\}$；

(3) 若某未标号点 k 有两个以上相邻的标号点，为减少迭代次数，可按(1)、(2)分别计算出 $\varepsilon(k)$ 的值，并取其中最大的一个标记。

第三步，重复第二步，可能出现两种情况：

(1) 标号过程中断。v_t 得不到标号，说明该网络中不存在增广链，给定的流量即为最大流。记已标号点的集合为 S，未标号点的集合为 T，(S,T) 为网络的最小割。

(2) v_t 得到标号。这时可用反向追踪法在网络中找出一条从 v_s 到 v_t 的由标号点和相应的弧连接成的增广链。

第四步，修改流量。若图中原有可行流为 f，令：

$$f' = \begin{cases} f + \varepsilon(t) & \text{对增广链上所有前向弧} \\ f - \varepsilon(t) & \text{对增广链上所有后向弧} \\ f & \text{所有非增广链上的弧} \end{cases}$$

这样又得到网络上的一个新的可行流 f'。

第五步，抹掉图上所有标号，重复第一步到第四步，直至图中找不到任何增广链，即出现第三步的情况(1)为止，这时网络中的流量即为最大流。

例 5.11 用标号算法求图 5.20 中 v_s 到 v_t 的最大流，并找出该网络的最小割。

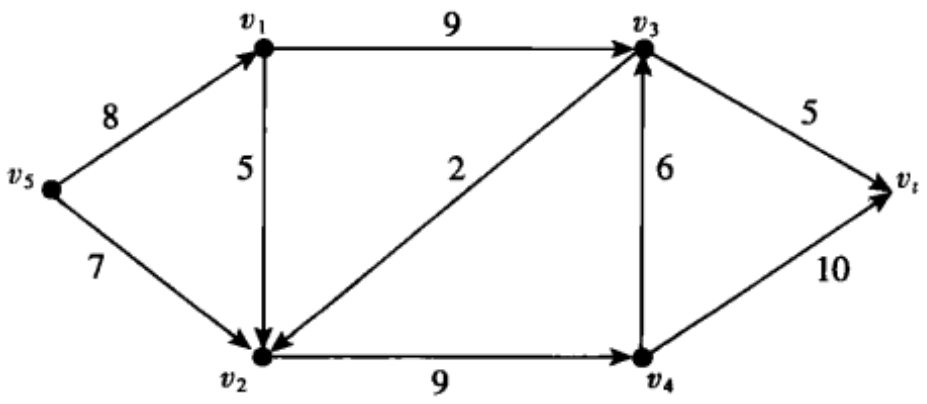


图 5.20

解：(1) 首先给发点 v_s 标号$(0,\infty)$；

(2) 从 v_s 点出发的弧(v_s,v_2)，因有 $f_{s2} < c_{s2}$，故对 v_2 点标号$[s,\varepsilon(2)]$，其中 $\varepsilon(2) = \text{Min}\{\varepsilon(s),(c_{s2} - f_{s2})\} = 2$；

(3) 对弧(v_1,v_2)，因$f_{12}>0$，故对v_1点标号$[v_2,\varepsilon(1)]$，其中$\varepsilon(v_1)=\text{Min}\{\varepsilon(2),f_{12}\}=2$；

(4) 对弧(v_1,v_3)，因$f_{13}<c_{13}$，故对v_3点标号$[v_1,\varepsilon(3)]$，其中$\varepsilon(3)=\text{Min}\{\varepsilon(1),(c_{13}-f_{13})\}=2$；

(5) 对弧(v_4,v_3)，因$f_{43}>0$，故对v_4点标号$[v_3,\varepsilon(4)]$，其中$\varepsilon(4)=\text{Min}\{\varepsilon(3),f_{43}\}=1$；

(6) 对弧(v_4,v_t)，因$f_{4t}<c_{4t}$，故对v_t点标号$[v_4,\varepsilon(t)]$，其中$\varepsilon(t)=\text{Min}\{\varepsilon(4),(c_{4t}-f_{4t})\}=1$；

(7) 因为收点v_t得到标号，可用反向追踪法找出网络中的一条增广链，如图5.21中虚线所示。

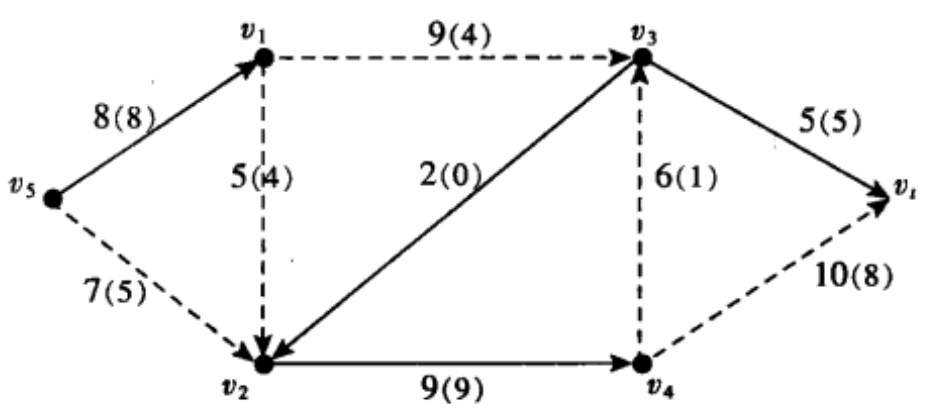


图5.21

(8) 修改增广链上各弧的流量：

$$f'_{s2}=f_{s2}+\varepsilon(t)=5+1=6$$

$$f'_{12}=f_{12}-\varepsilon(t)=4-1=3$$

$$f'_{13}=f_{13}+\varepsilon(t)=4+1=5$$

$$f'_{43}=f_{43}-\varepsilon(t)=1-1=0$$

$$f'_{4t}=f_{4t}+\varepsilon(t)=8+1=9$$

非增广链上的所有弧的流量不变。这样得到网络上的一个新的可行流，如图5.22所示。

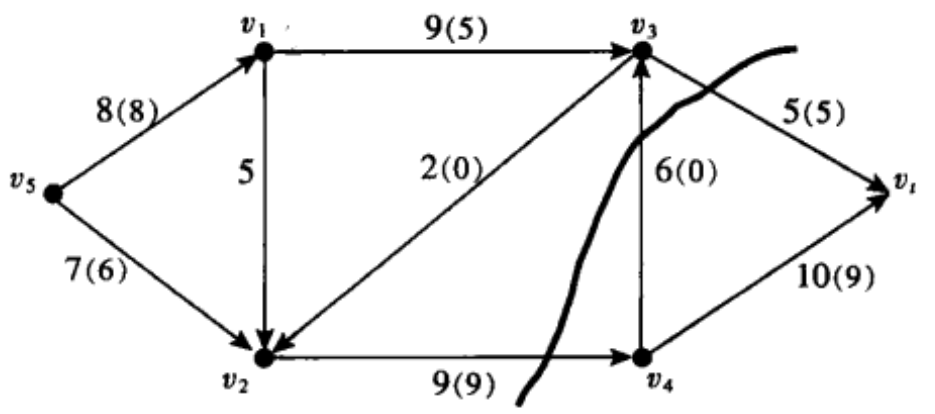


图5.22

在图5.22中重复上述标号过程，由于对v_s,v_2,v_1,v_3点标号后，标号中断，故图中给出的可行流即为该网络的最大流，流量为$f^*=14$。将已标号点v_s、v_2、v_1、v_3的集合记为S，未标号点v_4、v_t的集合为T，$(S,T)=\{(v_3,v_t),(v_2,v_4)\}$即为该网络的最小

割，有 $c^* = 14$。

三、最大流的 WinQSB 软件应用举例

在 WinQSB 软件的网络流模块中，最大流问题的求解采用的是上面介绍的标号法。下面我们以例 5.9 的石油输送问题为示例，说明怎样应用 WinQSB 软件计算最大流问题。

首先，调用 WinQSB 软件的子程序 Network Modeling，建立一个新问题，弹出对话框，如图 5.23 所示界面，选择“Maximal Flow Problem”，输入问题的文件名“Flow1”（读者自己可以任意取名），节点数 7。此时弹出一张需要输入数据的表格，对照图 5.24 输入数据，输入方法与最短路方法相同，求解方法也与最短路方法一样，系统输出如表 5 - 9 所示的最大流求解输出结果，最大流流量为 10。点击菜单栏“Results” → “Graphic Solution”，显示最大流网络图，见图 5.25。

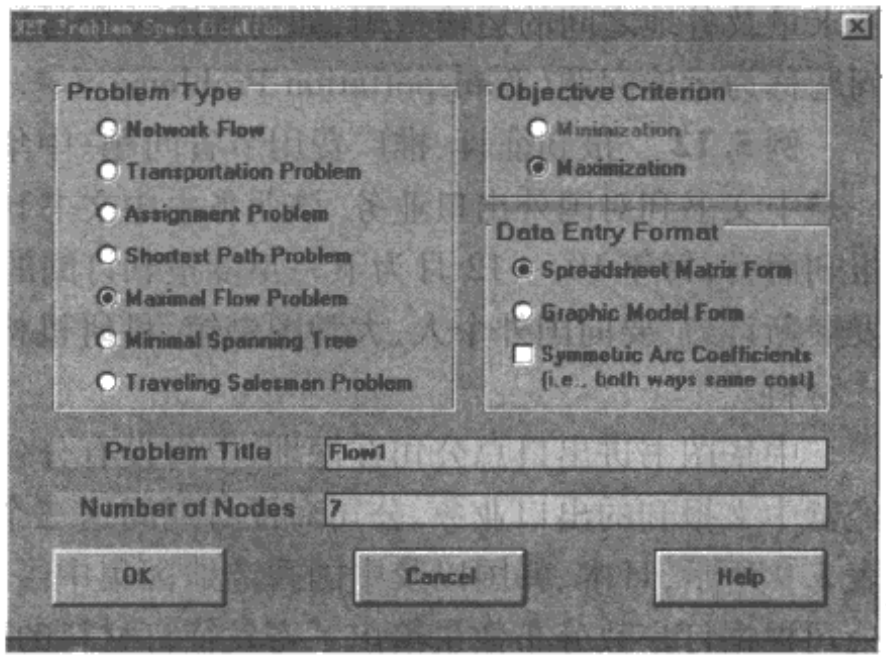


图 5.23

From	To	Net Flow		From	To	Net Flow
Node1	Node2	5	7	Node4	Node3	2
Node1	Node4	5	8	Node4	Node6	1
Node2	Node3	2	9	Node4	Node7	2
Node2	Node5	3	10	Node5	Node7	5
Node3	Node5	2	11	Node6	Node7	3
Node3	Node6	2				
Net Flow	From	Node1	To	Node7	=	10

图 5.24

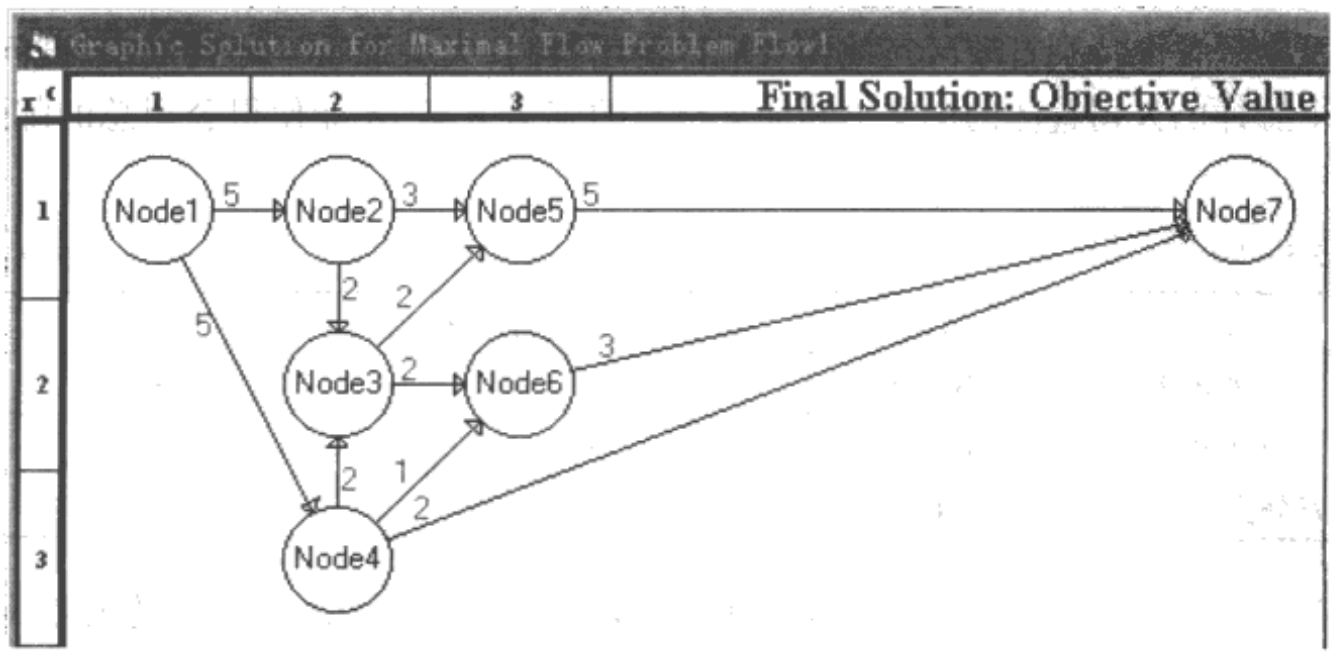


图 5.25

第五节　运输问题

人们在生产活动中，不可避免地要进行各种产品的调运工作。如某时期内将某生产地的煤、钢铁、粮食、水果等各类物资，分别运到若干个销地。根据各地的生产量和需求量及各地之间的运输费用，如何指定一个运输方案，使总运输费用最小，这样的问题称为运输问题(Transportation Problem)。

例 5.12　报刊征订、推广费用节省问题。中华图书进出口总公司的主营业务之一是中文书刊对国外出口业务。该业务由中文书刊出口部及两个分公司负责。就中文报刊而言，每年 10 ~ 12 月为下一年度报刊订阅的征订期，在此期间，为巩固老订户、发展新订户，要向国外个人、大学图书馆、科研机构等无偿寄发小礼品和征订宣传推广材料。

中华图书进出口总公司在深圳、上海设有分公司，同时允许这两家分公司也部分经营中文报刊的出口业务。公司对征订期间这三个部门寄发征订材料的工作安排如表 5.9 所示，日本、韩国以及中国香港地区集中了该公司的绝大部分中文报刊订户，公司根据订户数分布数量给出了寄发征订材料的数量，如表 5.10 所示。

表 5.9

部门	份数(册)
中文书刊出口部	15 000
深圳分公司	7 500
上海分公司	7 500
总计	30 000

表 5.10

国家和地区	份数(册)
日本	15 000
中国香港特别行政区	10 000
韩国	5 000
总计	30 000

一般情况下，这些材料无论由三家中哪个部门寄出，征收订户的效果都大致相同；同时，无论读者向哪个部门订阅，为总公司创造的利益是大致一样的。但由于各部门邮寄距离不同，邮寄方式及人工费用不同，导致从各部门寄往各地的费用也不同，如表 5.11 所示。由于寄发量大且每份材料的寄发费用较高，导致每年征订期日本、韩国以及中国香港特别行政区三地读者征订费用很高，大大加重了经营成本。试问，给出一个怎样的公司整体的中文书刊征订材料的邮运方案，才能使得公司总的邮运费最小？

表 5.11　　单位：元

	日本	中国香港特别行政区	韩国
中文书刊出口部	10.20	7	9
深圳分公司	12.50	4	14
上海分公司	6	8	7.50

类似的问题很多，均可以归纳为一般的运输问题，即在已知每个产地的供应量与每个销地的需求量以及两地之间的运输单价的情况下，解决如何把某种产品从若干个产

地调运到若干个销地，并使得总的运输费用最小。

一、运输问题的数学模型

运输问题实质是一种应用广泛的网络最优化模型。设某种物资有 m 个产地，$A_i(i=1,2,\cdots,m)$ 产量为 a_i，另有 n 个销地(需求地)$B_j(j=1,2,\cdots,n)$，销量(需求量) 分别为 b_j，现在需要把这种物资从各个产地运到各个销地，总的运费(运价) 为 c_{ij}，若产量总数等于销量总数，即 $\sum_{i=}^{m} a_i = \sum_{j=1}^{n} b_j$，那么如何组织调运，才能使总运费最小?

设 $x_{ij}(i=1,2,\cdots,m;j=1,2,\cdots,n)$ 为从第 i 个产地到第 j 个销地的物资运量，则运输问题的数学模型为：

$$\text{Min } Z = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}$$

$$s.t.\begin{cases}\sum_{j=1}^{n} x_{ij} = a_i \quad (i=1,2,\cdots,m)\\ \sum_{i=1}^{m} x_{ij} = b_j \quad (j=1,2,\cdots,n)\\ x_{ij} \geqslant 0 \quad (i=1,2,\cdots,m;j=1,2,\cdots,n)\end{cases}$$

我们注意到产销平衡条件为 $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$，因此这个模型描述的是产销平衡的运输问题，它是一个有 $m \times n$ 个变量和 $m+n$ 个等式约束的线性规划，具有如下特征：

(1) 运输问题是一个具有 $m \times n$ 个变量和 $m+n$ 个等式约束的线性规划问题；

(2) 运输问题的基变量个数是 $m+n-1$；

(3) $m+n-1$ 个变量构成基变量的充要条件是它们不构成闭回路。

闭回路是运输问题求解过程中非常重要的一个基本概念。如表 5.12 所示，$\{x_{11},x_{13},x_{23},x_{21}\}$ 即称为一个闭回路，其中每一个变量称作该闭回路的一个顶点。同样，$\{x_{12},x_{22},x_{24},x_{14}\}$ 和 $\{x_{22},x_{23},x_{33},x_{11},x_{12}\}$ 也都是闭回路。

表 5.12

	B_1	B_2	B_3	B_4
A_1	x_{11}	x_{12}	x_{13}	x_{14}
A_2	x_{21}	x_{22}	x_{23}	x_{24}
A_3	x_{31}	x_{32}	x_{33}	x_{34}

二、表上作业法

表上作业法是直接在运价表上求最优解的一种方法，其实质是单纯形法，所以又称为运输单纯形法(Transportation Simples Method)。在目标函数最小、产销平衡和运价非负的标准运输问题中，其基本的计算步骤如下：

(1) 确定初始基本可行解(初始调运方案)，常用的方法有西北角法(左上角

法)、最小元素法、最大差额法(Vogel 近似法) 等;

(2) 求各非基变量的检验数,判断是否达到最优解,常用的检验方法有闭回路法和位势法,如已是最优解,则停止计算,否则转到下一步;

(3) 确定入基变量与出基变量,在运价表上用闭回路法调整运量,得到新的基本可行解;

(4) 重复(2) 和(3) 直至获得最优解。

以上运算都是在运价表上完成的,下面我们通过一个例子来说明表上作业法的计算步骤。

例 5.13 旺旺食品有限公司有三个生产面包的分厂 A_1、A_2、A_3,有四个销售公司 B_1、B_2、B_3、B_4,其各分厂每日的产量、各销售公司每日的销售量(单位:吨) 以及各分厂到各销售公司的单位运价(单位:百元/吨) 如表5 - 13 所示。试问,该公司应如何调运产品,使得在满足各销售点需求量的前提下,总运费最少?

表 5.13

销地 / 产地	B_1	B_2	B_3	B_4	产量
A_1	3	11	3	10	7
A_2	1	9	2	8	4
A_3	7	4	10	5	9
销量	3	6	5	6	20

下面,我们用表上作业法来解此题,这是一个产销平衡的运输问题。

(一) 确定初始基本可行解

我们在产销平衡与运价表上找出初始基本可行解,为了把初始基本可行解与运价区分开,我们把运价放在每一栏的右上角,每一栏的中间写上初始基本可行解(调运量),如表 5 - 14 所示。

表 5 - 14

销地 / 产地	B_1	B_2	B_3	B_4	产量
A_1	3 3	11 4	3	10	7 4 0
A_2	1	9 2	2 2	8	4 2 0
A_3	7	4	10 3	5 6	9 6 0
销量	3 0	6 2 0	5 0	6 0	

1. 西北角法

先从表的左上角(西北角)的变量 x_{11} 开始分配运输量,并使 x_{11} 取尽可能大的值,这里发点 A_1 的产量为7,B_1 的销量为3,x_{11} 只能取3,即 $x_{11} = \text{Min}(7,3) = 3$。

由于 x_{11} 为3,所以 x_{21} 与 x_{31} 必须为0。令 x_{21} 与 x_{31} 为非基变量,这样我们在 x_{11} 格里填上3,并把 B_1 的销量与 A_1 的产量都减去3填入销量、产量处,把原来的销量、产量划去。新填上的销量、产量表示在 $x_{11} = 3$(即 A_1 运给 B_1 3吨)的情况下,B_1 还需要的销量与 A_1 还能供应的数量,这时 A_1 还能供应4吨,而 B_1 销量为0,已不需要从 A_2 和 A_3 处再运入了,这样我们就可把 B_1 列划去了。这样在产销平衡与运价表上只剩下3×3矩阵,这时 x_{12} 为西北角了,同样我们取 x_{12} 为尽可能大的值,知 $x_{12} = \text{Min}(4,6) = 4$。

取 $x_{12} = 4$ 填入,把 A_1 的产量改为 $4 - 4 = 0$,把 B_2 的销量改为 $6 - 4 = 2$ 并填上,然后划掉 A_1 这一行;同样找到西北角 x_{22},取 $x_{22} = \text{Min}(4,2)$,改写 A_2 产量为2,B_2 销量为0,并划去 B_2 列;继续下去,取 $x_{23} = \text{Min}(2,5) = 2$,改写 A_2 产量为0,B_3 销量为3,划去 A_2 行;再取 $x_{33} = \text{Min}(3,9) = 3$,改写 A_3 产量为6,B_3 销量为0,并划去 B_3 列;最后取 $x_{34} = \text{Min}(6,6) = 6$ 并填上,A_3 的产量和 B_4 的销量都改写为0,并划去 A_3 行。这样就得到了一个初始基本可行解,有 $m + n - 1 = 3 + 4 - 1 = 6$ 个基变量,其中 $x_{11} = 3, x_{12} = 4, x_{22} = 2, x_{23} = 2, x_{33} = 3, x_{34} = 6$,此时,其总运输费用为 $3 + 11 \times 4 + 9 \times 2 + 2 \times 2 + 10 \times 3 + 5 \times 6 = 135$(百元)。

2. 最小元素法

我们再来介绍确定初始基本可行解的另一个方法——最小元素法。

西北角法是对西北角的变量分配运输量,而最小元素法的做法是就近供应,即对单位运价最小的变量分配运输量。我们仍以上例为例确定初始基本可行解,如表5.15所示。

表5.15

产地 \ 销地	B_1	B_2	B_3	B_4	产量
A_1	3	11	3 4	10 3	7 3 0
A_2	1 3	9	2 1	8	4 1 0
A_3	7	4 6	10	5 3	9 3 0
销量	3 0	6 0	5 4 0	6 3 0	

在表5.15上找到单位运价最小的 x_{21},开始分配运输量,并使 x_{21} 取尽可能大的值。这里发点 A_2 产量为4,收点 B_1 销量为3,取 $x_{21} = \text{Min}(4,3) = 3$,把 A_2 的产量改写为 $4 - 3 = 1$,把 B_1 的销量改写为 $3 - 3 = 0$,并把 B_1 列划去。在所剩下的3×3矩阵里找到运价最小的变量 x_{23},取 $x_{23} = \text{Min}(1,5) = 1$,$A_2$ 的产量改为 $1 - 1 = 0$,B_3 的销量

改为 5 − 1 = 4，并把 A_2 行划去。在所剩下的矩阵里找到运价最小的变量，取 x_{13} = Min (7,4) = 4，A_1 的产量改为 3，B_3 销量改为 0，并划去 B_3 列。在剩下的矩阵里找到运价最小的变量 x_{32}，取 x_{32} = Min (9,6) = 6，产量改为 9 − 6 = 3，B_2 的销量改为 6 − 6 = 0，并把 B_2 列划去。在剩下的表中找到运价最小的变量 x_{34}，取 x_{34} = Min (3,6) = 3，A_3 产量改为 3 − 3 = 0，B_4 的销量改为 6 − 3 = 3，并划去 A_3 行。在所剩下的表中找到运价最小的变量 x_{14}，取 x_{14} = Min (3,3) = 3，A_1 产量改为 0，B_4 的销量改为 0，并划去 A_1 行。这就得到了一个初始基本可行解，有 6 个基变量，其中 x_{13} = 4，x_{14} = 3，x_{23} = 1，x_{32} = 6，x_{34} = 3，其总运费为 3 × 4 + 10 × 3 + 1 × 3 + 2 × 1 + 4 × 6 + 5 × 3 = 86。一般用最小元素法求得的初始基本可行解比用西北角法求得的初始基本可行解总费用要少一些。这样从用最小元素法求得的初始基本可行解出发求最优解迭代次数可能少一些。

另外在求初始基本可行解时要注意两个问题：

（1）当我们取定 x_{ij} 的值之后，会出现的 A_i 产量与 B_j 的销量都改为 0 的情况，这时只能划去 A_i 行或 B_j 列，但不能同时划去行与列。

（2）用最小元素法时，可能会出现只剩一行或一列的所有格均未填数或未被划掉的情况，此时在这一行或这一列中除已填上的数外均填上 0，不能按空格划掉，这样可以保证填过数或 0 的格为 $m + n - 1$ 个，即保证基变量的个数为 $m + n - 1$ 个。

（二）最优解的判别

这里介绍两种检验已得的运输方案是否是最优解的方法：一个是闭回路法，另一个是位势法。

1. 闭回路法

所谓闭回路，是在已给出的调运方案的运输表上，从一个非基变量的空格出发，沿水平或垂直方向前进，只有碰到代表基变量的填入数字的格才能向左或右转 90°（当然也可以不改变方向）继续前进，这样继续下去，直至回到出发的那个空格，由此形成的封闭的折线叫做回路，一个空格存在唯一的闭回路。

所谓闭回路法，就是对于代表非基变量的空格（其调运量为 0），把它的调运量调整为 1，由于产销平衡的要求，我们必须对这个空格的闭回路的顶点的调运量加上或减少 1。最后我们计算出由于这些变化给整个运输方案的总运输费带来的变化，其增加值或减少值作为该空格的检验数填入该空格，如果所有代表非基变量的空格的检验数也即非基变量的检验数都大于等于 0，也就是任何一个非基变量变成基变量都会使得总运输费增加（对于求目标函数最大值的线性规划问题是要求所有检验数都小于等于 0），那么原基本可行解就是最优解了，否则要进一步迭代以找出最优解。

我们不妨以上例中用最小元素法求出的初始基本可行解中的非基变量来加以说明，如表 5.16 所示。

我们先从空格（非基变量）x_{11} 出发，找到一个闭回路，如表 5.16 所示，这个闭回路有四个顶点，除 x_{11} 为非基变量外，其余的都是基变量。现在把 x_{11} 的调运量从 0 增加为 1 吨，运费增加了 3 百元，为了使 A_1 产量平衡，x_{13} 就减少 1 吨，运费减少了 3 百元，为了使 B_3 的销量平衡，x_{23} 就增加了 1 吨，运费增加 2 百元。为了使 A_2 产量与 B_1 的销量平衡，x_{21} 就减少 1 吨，运费减少 1 百元。这样调整之后，运费增加了 3 − 3 + 2 − 1 =

1(百元),这就说明了 x_{11} 为非基变量,其值为0是对的选择。如果让 x_{11} 变为基变量,则运费要增加,这时我们把运费增加值1填入此空格作为 x_{11} 的检验数,为了与调整量加以区

表 5.16

产地 \ 销地	B_1	B_2	B_3	B_4	产量
A_1	3 ①		3 4		7
A_2	1 3		2 1		4
A_3					9
销量	3	6	5	6	

别我们在1上加○,记为①。

同样我们可以用闭回路法求出其他非基变量的检验数,如表5.17所示。

表 5.17

空格	闭回路	检验数
x_{11}	$x_{11}——x_{13}——x_{23}——x_{21}——x_{11}$	1
x_{12}	$x_{12}——x_{14}——x_{34}——x_{32}——x_{12}$	2
x_{22}	$x_{22}——x_{23}——x_{13}——x_{14}——x_{34}——x_{32}——x_{22}$	1
x_{24}	$x_{24}——x_{23}——x_{33}——x_{14}——x_{24}$	-1
x_{31}	$x_{31}——x_{34}——x_{14}——x_{13}——x_{23}——x_{21}——x_{31}$	10
x_{33}	$x_{33}——x_{34}——x_{14}——x_{13}——x_{33}$	12

用闭回路法求检验数,需要给每一个空格找一条闭回路,当产销点很多时这种计算很繁杂。

2. 位势法

所谓位势法,就是我们对运输表上的每一行赋予一个数值,对每一列也赋予一个数值,它们的数值是由基变量 x_{ij} 的检验数 $\lambda_{ij} = c_{ij} - u_i - v_j = 0$ 所决定的,则非基变量 x_{ij} 的检验数就可用公式 $\lambda_{ij} = c_{ij} - u_i - v_j$ 求出。

下面我们用位势法对本例用最小元素法求出的初始基本可行解求检验数,对给出的初始基本可行解做一个表,如表5.18所示。把原来表中的最后一列的产量改成 u_i 值,最后一行的产量改成 v_j 值,表中每一栏的右上角的数值仍表示运价,栏中的数值表示调运量,栏中无数值的表示此栏为非基变量,调运量为0。

表 5.18

销地 / 产地	B_1	B_2	B_3	B_4	u_i
A_1	3 ①	11 ②	3 4	10 3	**0**
A_2	1 3	9 ①	2 1	8 (-1)	**-1**
A_3	7 ⑩	4 6	10 ⑫	5 3	**-5**
销量	**2**	**9**	**3**	**10**	

我们先给 u_i 赋予一个任意数值，不妨令 $u_i = 0$，则从基变量 x_{13} 的检验数 $\lambda_{13} = c_{13} - u_1 - v_3$，求得 $v_3 = c_{13} - u_1 = 3 - 0 = 3$。

同样，从 $\lambda_{13} = c_{14} - u_1 - v_4 = 0$，求得 $v_4 = c_{14} - u_1 = 10 - 0 = 10$；从 $\lambda_{23} = c_{23} - u_{21} - v_3 = 0$，求得 $u_2 = c_{23} - v_3 = -1$；从 $\lambda_{34} = c_{34} - u_3 - v_4 = 0$，求得 $u_3 = c_{34} - v_4 = -5$；从 $\lambda_{21} = c_{21} - u_2 - v_1 = 0$，求得 $v_1 = c_{21} - u_1 = 2$；从 $\lambda_{32} = c_{32} - u_3 - v_2 = 0$，求得 $v_2 = c_{32} - u_3 = 9$。把所得的值填入表5 - 19，利用所求得的 u_i 与 v_j 值来计算非基变量的检验数：$\lambda_{11} = 1$，$\lambda_{12} = 2$，$\lambda_{22} = 1$，$\lambda_{24} = -1$，$\lambda_{31} = 10$，$\lambda_{33} = 12$。把检验数填入上表，显然用位势法求得的检验数与用闭回路法求得的检验数是一样的。

（三）改进运输方案的办法 —— 闭回路调整法

我们已知当表中某个非基变量（非基变量所在的空格）的检验数为负值时，表明未获得最优解，要进行调整。我们在所有负值的检验数中，选其中最小的负检验数，以它对应的非基变量为入基变量，如上例中因为 $\lambda_{24} = -1$，选非基变量 x_{24} 为入基变量，并以 x_{24} 所在格为出发点做一个闭回路，如表 5.19 所示。

表 5.19

销地 / 产地	B_1	B_2	B_3	B_4	产量
A_1			4(+1)	3(-1)	**7**
A_2	3		1(-1)	(+1)	**4**
A_3		6		3	**9**
销量	**3**	**6**	**5**	**6**	

由于 $\lambda_{24} = -1$，表明增加 1 个单位的 x_{24} 运输量，就可以使总运输量减少 1，我们应

尽量多地增加 x_{24} 的运输量，但为了保证运输方案的可行性（即所有调运量必须大于等于0），所以在以出发点 x_{24} 所在空格为1的闭回路顶点的序号中，找出所有偶数的顶点的调运量：$x_{14}=3$，$x_{23}=1$，取其中的最小值为 x_{24} 的值，即 $x_{24}=\text{Min}(3,1)=1$。为了使产销平衡，把所有的闭回路为偶数顶点的运输量都减少这个值，而其他的闭回路上的奇数顶点的运输量都增加这个值，即得到了调整后的方案，如表5.20所示。

表5.20

产地＼销地	B_1	B_2	B_3	B_4	产量
A_1			5	2	7
A_2	3			1	4
A_3		6		3	9
销量	3	6	5	6	

对表5.20给出的运输方案，我们用位势法进行检验，如表5.21所示。

表5.21中带圈的数字是非基变量的检验数，可知所有检验数都大于等于0（基变量的检验数都等于0），此解是最优解，这时最小总运输费为85（百万元），具体的运输方案如下：A_1 分厂运5吨给销售公司 B_3，运2吨给销售公司 B_4；A_2 分厂运3吨给销售公司 B_1，运1吨给销售公司 B_4；A_3 分厂运6吨给销售公司 B_2，运3吨给销售公司 B_4。

表5.21

产地＼销地	B_1	B_2	B_3	B_4	u_i
A_1	① 3	② 11	5 3	2 10	0
A_2	3 1	② 9	① 2	1 8	-2
A_3	⑨ 7	6 4	⑫ 10	3 5	-5
销量	3	9	3	10	

（四）如何找多个最优方案

与单纯形法一样，表上作业法求解运输问题也会存在多个最优方案的情况，这对决策者来说是很重要的，他可以考虑与模型无关的其他因素，从而确定最后的方案。

识别是否有多个最优解的方法与单纯形法一样，只需看最优方案中是否存在非基变量的检验数为0，如在上例表中给出的最优运输方案中 x_{11} 的检验数 $\lambda_{11} = 0$，可知此运输问题有多个最优解。为求得另一个最优解，只要把 x_{11} 作为入基变量，调整运输方案，就可得到另一个最优方案，如表5.22所示。

表5.22

销地 / 产地	B_1	B_2	B_3	B_4	产量
A_1	(+2)		5	2(−2)	7
A_2	3(−2)			1(+2)	4
A_3		6		3	9
销量	3	6	5	6	

新的最优方案如表5.23所示。

表5.23

销地 / 产地	B_1	B_2	B_3	B_4	产量
A_1	2		5		7
A_2	1			3	4
A_3		6		3	9
销量	3	6	5	6	

最小运费为 $3 \times 2 + 1 \times 1 + 4 \times 6 + 3 \times 5 + 8 \times 3 + 5 \times 3 = 85$(百万元)。

三、运输模型的扩展

(一) 产销不平衡的运输问题

当总产量与总销量不相等时，我们称其为产销不平衡的运输问题。这类运输问题我们在实际中常常遇到，其求解方法是将不平衡问题化为平衡问题来求解。

(1) 当产大于销，即 $\sum_{i=1}^{m} a_i > \sum_{j=1}^{n} b_j$ 时，则运输问题的数学模型变成：

$$\text{Min } Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

$$s.t.\begin{cases}\sum_{j=1}^{n} x_{ij} \leqslant a_i & i=1,2,\cdots,m \\ \sum_{i=1}^{m} x_{ij} = b_j & j=1,2,\cdots,n \\ x_{ij} \geqslant 0 & i=1,2,\cdots,m;j=1,2,\cdots,n\end{cases}$$

此时，由于总产量大于总销量，就要考虑部分产地的产量不能全部运送完。相应地库存量为 $x_{i,n+1}(i=1,2,\cdots,m)$，则总的库存量为 $b_{n+1}=\sum_{i=1}^{m} x_{i,n+1}=\sum_{i=1}^{m} a_i-\sum_{j=1}^{n} b_j$。我们可以虚设一个销地 B_{n+1}，各地运往此处的销量为 b_{n+1}，各地到此处的运价为0，即 $c_{i,n+1}=0(i=1,2,\cdots,m)$，将产大于销的不平衡问题转化为一个平衡问题：

$$\text{Min } Z=\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}$$

$$s.t.\begin{cases}\sum_{j=1}^{n+1} x_{ij} = a_i, i=1,2,\cdots,m \\ \sum_{i=1}^{m} x_{ij} = b_j & j=1,2,\cdots,n+1 \\ x_{ij} \geqslant 0 & i=1,2,\cdots,m;j=1,2,\cdots,n+1\end{cases}$$

(2) 当销大于产，即 $\sum_{i=1}^{m} a_i < \sum_{j=1}^{n} b_j$ 时，则运输问题的数学模型变成：

$$\text{Min } Z=\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}$$

$$s.t.\begin{cases}\sum_{j=1}^{n} x_{ij} = a_i, i=1,2,\cdots,m \\ \sum_{i=1}^{m} x_{ij} \leqslant b_j & j=1,2,\cdots,n \\ x_{ij} \geqslant 0 & i=1,2,\cdots,m;j=1,2,\cdots,n\end{cases}$$

此时，由于总销量大于总产量，就要考虑部分需求地的需求不能满足。我们可以虚设一个产地 A_{n+1}，其产量为 $a_{n+1}=\sum_{i=1}^{m} x_{m+1,j}=\sum_{j=1}^{n} b_j-\sum_{i=1}^{m} a_i$，$x_{m+1,j}$，是产地 A_{n+1} 到各销地的运量，相应运价也为0，即 $c_{m+1,j}=0(j=1,2,\cdots,n)$，将销大于产的不平衡问题转化为一个平衡问题：

$$\text{Min } Z=\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}$$

$$s.t.\begin{cases}\sum_{j=1}^{n} x_{ij} = a_i & i=1,2,\cdots,m+1 \\ \sum_{i=1}^{m+1} x_{ij} = b_j & j=1,2,\cdots,n \\ x_{ij} \geqslant 0 & i=1,2,\cdots,m+1;j=1,2,\cdots,n\end{cases}$$

例 5.14 农用化肥调运问题。有三个化肥厂供应四个地区的农用化肥。假定等量的化肥在这些地区使用效果相同。各化肥厂年产量、各地区年需求量以及从各化肥

厂到各地区单位化肥的运价如表 5 - 24 所示，试给出总的运费最节省的化肥调运方案。

表 5.24　　运价：万元 / 万吨

需求地区 / 化肥厂	B_1	B_2	B_3	B_4	产量(万吨)
A_1	16	13	22	17	50
A_2	14	13	19	15	60
A_3	19	20	23	—	50
最低需求(万吨)	30	70	0	10	
最高需求(万吨)	50	70	30	不限	

解：这是一个产销不平衡的运输问题，总产量为 160 万吨，四个地区的最低需求为 110 万吨，最高需求为无限。根据现有产量，B_4 地区每年最多能分配到 60 万吨化肥，这样最高需求为 210 万吨，大于总产量。为了转化为平衡问题，在产销平衡表中增加一个假想的化肥厂 A_4，其年产量为 50 万吨。由于各地区的需求量包含两部分，如地区 B_1，其中 30 万吨是最低需求，故不能由假想化肥厂 A_4 供给，令相应运价为 M(任意大正数)，而另一部分 20 万吨满足或不满足均可，故也可以由假想化肥厂 A_4 供给，相应运价为 0。对凡是需求分两种情况的地区，实际上可按照两个地区看待。这样就可以写出这个问题的产销平衡表和单位运价表(见表 5.25)。

根据表上作业法计算，可以求得这个问题的最优方案如表 5.26 所示，最小总费用为 2 460 万元。

表 5.25

销地 / 产地	B_1'	B_1''	B_2	B_3	B_4'	B_4''	产量
A_1	16	16	13	22	17	17	50
A_2	14	14	13	19	15	15	60
A_3	19	19	20	23	M	M	50
A_4	M	0	M	0	M	0	50
销量	30	20	70	30	10	50	210

表 5.26

销地 / 产地	B_1'	B_1''	B_2	B_3	B_4'	B_4''	产量
A_1			50				50
A_2			20		10	30	60
A_3	30	20	0				50
A_4				30		20	50
销量	30	20	70	30	10	50	210

（二）综合生产计划问题

运输问题在生产实践中得到了广泛的应用，多时期的综合生产计划编制问题也可以通过运输模型来解决，下面举一例来说明。

例 5.15 某公司在接下来的三个月内每月都要按照销售合同生产出两种产品。这两种产品使用相同的设备并需要投入相同的生产能力。每个月可供使用的生产和储存设备都会发生变化。所以生产能力、单位生产成本以及单位储存成本每个月都不相同，有必要在某些月中多生产一种或者多种并储存起来以备需要的时候使用。

对于每一个月来说，表 5.27 中给出了在正常时间和加班时间内能够生产这两种产品的总数、按照合同需要生产的数量、在正常时间和加班时间内的单位产品成本以及每件产品储存到下一个月的储存成本。

表 5.27

月份	最大生产总量		销售量		单位生产成本(1 000 元／件)				单位储存成本(1 000 元／件)	
					正常时间		加班时间			
	正常时间	加班时间	产品 1	产品 2	产品 1	产品 2	产品 1	产品 2	产品 1	产品 2
1	10	3	5	3	15	16	18	20	1	2
2	8	2	3	5	17	15	20	18	2	1
3	10	3	4	4	19	17	22	22		

生产管理人员想要制定一个在正常时间（如果正常时间不够，使用加班时间）内进行生产的每一种产品数量的计划进度，目标是在满足合同规定的基础上，使三个月的总生产和储存成本最小。开始和三个月结束后的储存都为 0。问应如何安排一至三月的生产计划。

解：这是一个多时期、多品种、多种生产方式的综合平衡计划问题。不同时期的生产能力、需求量、生产成本和储存成本已知，求每个时期在正常时间和加班时间每种产品各生产多少，第一月和第二月月底两种产品各储存多少，既满足合同又使总成本最少。

设决策变量为 x_{ij}，其具体含义是：

$i, j = 1,3,5$ 表示第一月、第二月和第三月正常时间内生产产品 1 用于第一月、第二月和第三月交货的产量；

$i = 1,3,5, j = 2,4,6$ 表示第一月、第二月和第三月正常时间内生产产品 2 用于第一月、第二月和第三月交货的产量；

$i = 2,4,6, j = 1,3,5$ 表示第一月、第二月和第三月加班时间内生产产品 1 用于第一月、第二月和第三月交货的产量；

$i, j = 2,4,6$ 表示第一月、第二月和第三月加班时间内生产产品 2 用于第一月、第二月和第三月交货的产量。

相应的单位产品的总成本（生产成本 + 储存成本）为 $c_{ij}(i \leqslant j)$，对于 $i > j$ 时的成本用 M 表示。例如 x_{35} 表示第二月正常时间内生产的产品 1 用于第三月交货的数量，第 1 种单位产品的成本是 17 千元，在第三月交货，单位产品的储存成本是 2 千元，因此单位产品总成本 c_{35} 为 19 千元，如此等等，就可以获得单位产品的总成本。注意

本问题生产能力大于合同需求量，是属于产大于销的问题，将成本增加一列，成本等于0，需求量等于12，得到一个平衡运输问题的产销平衡表和单位运价表（见表5.28）。

表5.28

	1月		2月		3月		剩余能力	生产能力
	产品1	产品2	产品1	产品2	产品1	产品2		
1月正常时间	15	16	16	18	18	19	0	10
1月加班时间	18	20	19	22	21	23	0	3
2月正常时间	*M*	*M*	17	15	19	16	0	8
2月加班时间	*M*	*M*	20	18	22	19	0	2
3月正常时间	*M*	*M*	*M*	*M*	19	17	0	10
3月加班时间	*M*	*M*	*M*	*M*	22	22	0	3
需求量	5	3	3	5	4	4	12	36

根据表上作业法，可以求得这个问题的最优方案如表5.29所示，最优生产计划是：第一个月正常时间内生产第1种产品7件，当月交货5件，第二个月交货2件；生产第2种产品3件，当月交货，总产量10件，不加班。第二个月正常时间内生产第1种产品1件，当月交货1件；生产第2种产品7件，当月交货5件，第三个月交货2件，总产量8件，不加班。第三个月正常时间内生产第1种产品4件，当月交货；生产第2种产品2件，当月交货，总产量6件，不加班。三月末无库存，总成本为389千元。

表5.29

	1月		2月		3月		剩余能力	生产能力
	产品1	产品2	产品1	产品2	产品1	产品2		
1月正常时间	5	3	2					10
1月加班时间							3	3
2月正常时间			1	5		2		8
2月加班时间							2	2
3月正常时间					4	2	4	10
3月加班时间							3	3
需求量	5	3	3	5	4	4	12	36

（三）转运问题

所谓转运问题（Transshipment Problem）就是将产地A_i的物资运送到销地B_j，在运送的过程中不一定直接到达销地，而是通过其他产地、销地或者中转地最后到达销地。

例5.16 飞跃电子仪器有限公司在大连、广州和成都有三个分厂，大连分厂每月生产200台某种仪器，广州分厂每月生产300台某种仪器，成都分厂每月生产500台某种仪器。该公司在上海与天津有两个销售公司负责对南京、济南、南昌和青岛四个城市的仪器供应。同时，公司同意广州向成都运送货物，大连和成都可直接分别向

南京和青岛供货，南京的货物也可以运往济南，南昌的货物也可以运往青岛。这些城市间的每台仪器的运输费用标在两个城市间的弧上，单位为百元。问应该如何调运仪器，使得总的运输费用最低？

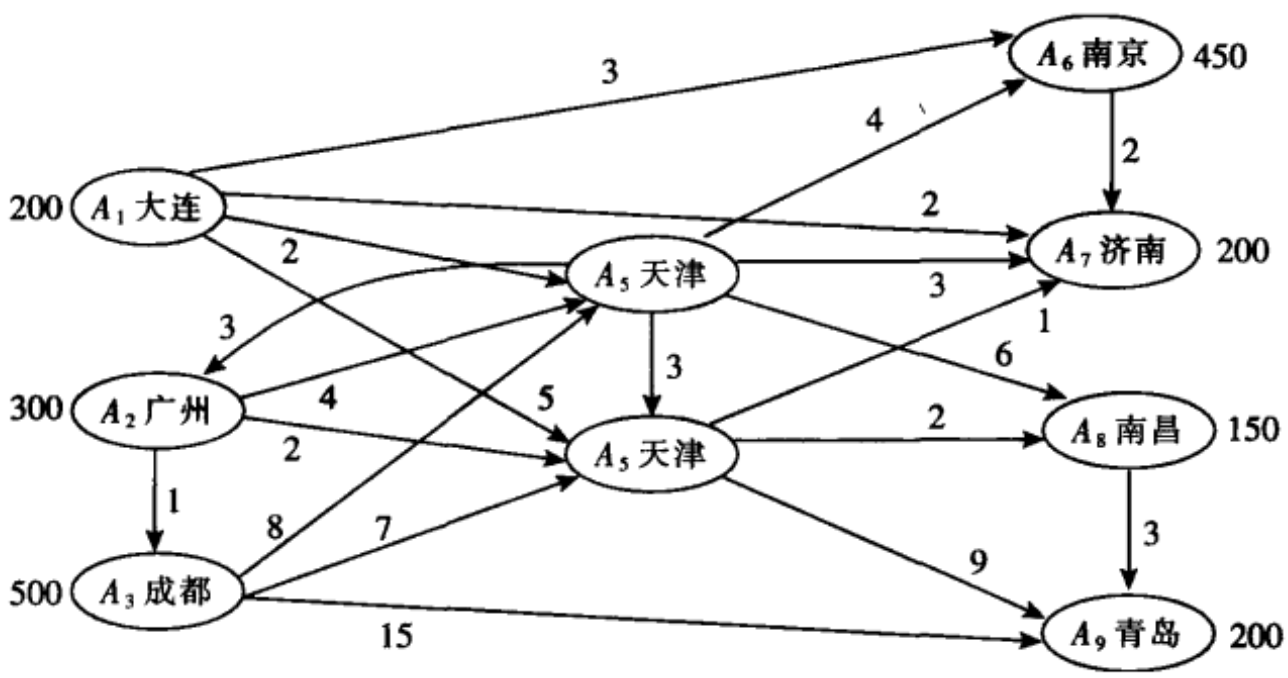


图 5.26

在图 5.26 中，A_1、A_2、A_3 是产地（称为网络的发点），A_4、A_5 是中转地（称为网络的中间点），$A_6 \sim A_9$ 是销地（称为网络的收点）。货物只能沿两条弧的方向流动，弧上的数值为两地间的单位运输费用。我们可以把上述问题一般化，设产地 A_i 为 m 个，销地 A_j 为 n 个，中转地 A_t 为 r 个，则产地和销地总数为 $m+n+r$ 个，产销平衡。记 A_i 到 A_j 的弧为 (i,j)，单位运输费用为 c_{ij}，不可到达的运价为 M，本地到达本地的运价为零。设 x_{ij} 为 A_i 到 A_j 的运量 $(i,j=1,2,\cdots,m+n+r)$，则产销平衡的转运问题的数学模型如下：

$$\text{Min } Z = \sum_{\text{所有的弧}} c_{ij}x_{ij}$$

$$s.t. \begin{cases} \sum\limits_{\text{流出弧}} x_{ij} - \sum\limits_{\text{流入弧}} x_{ij} = a_i \quad i=1,2,\cdots,m & \text{发点 } A_i \\ \sum\limits_{\text{流出弧}} x_{tj} - \sum\limits_{\text{流入弧}} x_{it} = 0 \quad t=1,2,\cdots,r & \text{中间点 } A_t \\ \sum\limits_{\text{流入弧}} x_{ij} - \sum\limits_{\text{流出弧}} x_{ij} = b_j \quad j=1,2,\cdots,n & \text{收点 } A_j \\ x_{ij} \geqslant 0 \quad i,j=1,2,\cdots,m+n+r \end{cases}$$

如果产大于销，则将模型中的第一个约束等式改为“≤”约束；如果销大于产，则将模型中的第二个约束等式改为“≥”约束即可。

运用表上作业法求解转运问题时需要把上述模型转化为产销平衡表，此时，产地的产量等于 a_i 加上总产量，销地和中转地的产量均等于总产量；销地的需求量等于 b_j 加上总需求量，产地和中转地的需求量等于总需求量。这样，我们可以得到例 5.16 的产销平衡表，如图 5.27 所示，根据表上作业法可以求得这个问题的最优方案如图 5.28 所示。

From \ To	A1	A2	A3	A4	A5	A6	A7	A8	A9	Supply
A1	0			2	5	3	2			1200
A2		0	1	4	2					1300
A3			0	8	7				15	1500
A4		3		0	3	4	3	6		1000
A5					0		1	2	9	1000
A6						0	2			1000
A7							0			1000
A8								0	3	1000
A9									0	1000
Demand	1000	1000	1000	1000	1000	1450	1200	1150	1200	

图 5.27

02-07-2006	From	To	Shipment	Unit Cost	Total Cost	Reduced Cost
1	A1	A1	1000	0	0	0
2	A1	A6	200	3	600	0
3	A2	A2	1000	0	0	0
4	A2	A5	300	2	600	0
5	A3	A3	1000	0	0	0
6	A3	A4	250	8	2000	0
7	A3	A5	250	7	1750	0
8	A4	A4	750	0	0	0
9	A4	A6	250	4	1000	0
10	A5	A5	450	0	0	0
11	A5	A7	200	1	200	0
12	A5	A8	350	2	700	0
13	A6	A6	1000	0	0	0
14	A7	A7	1000	0	0	0
15	A8	A8	800	0	0	0
16	A8	A9	200	3	600	0
17	A9	A9	1000	0	0	0
	Total	Objective	Function	Value =	7450	

图 5.28

四、运输问题的 WinQSB 软件应用

(一) 运输模型

在 WinQSB 软件的网络流模块中,一般运输模型的求解采用的是前述介绍的表上作业法。下面我们以例 5.12 的报刊征订、推广费用节省问题为例,说明怎样应用 WinQSB 软件计算产销平衡的运输问题。

在利用 WinQSB 软件进行计算之前,我们先建立该问题的产销平衡和运价表,如表 5.30 所示。

表 5.30

销地 / 产地	日本	中国香港	韩国	产量
中文书刊出口部	10.20	7	9	15 000
深圳分公司	12.50	4	14	7 500
上海分公司	6	8	7.50	7 500
销量	15 000	10 000	5 000	30 000

然后，调用 WinQSB 软件的子程序 Network Modeling，建立一个新问题，弹出对话框，如图 5.29 所示界面，选择“Network Flow”或者“Transportation Problem”（本例我们选择后者），以及“Minimization”，输入问题的文件名“Tran1”（读者自己可以任意取名），产地数目 3 和销地数目 3。接着，点击 OK，此时弹出一张需要输入数据的表格，对照表 5.33 输入数据，输入方法与最短路方法相同，求解方法也与最短路方法一样，并重新命名产地和销地，系统输出如图 5.30 所示的数据表格。点击菜单栏“Solve and Analyze”，下拉菜单有四个求解方法供选择：“Solve the Problem”（只求出最优解）、“Solve the Display Steps – Network”（网络图求解并显示迭代步骤）、“Solve the Display Steps – Tableau”（表格求解并显示迭代步骤）、“Select Initial Solution Method”（选择求初始解方法）。初始解求解方法有八种方法供选择：

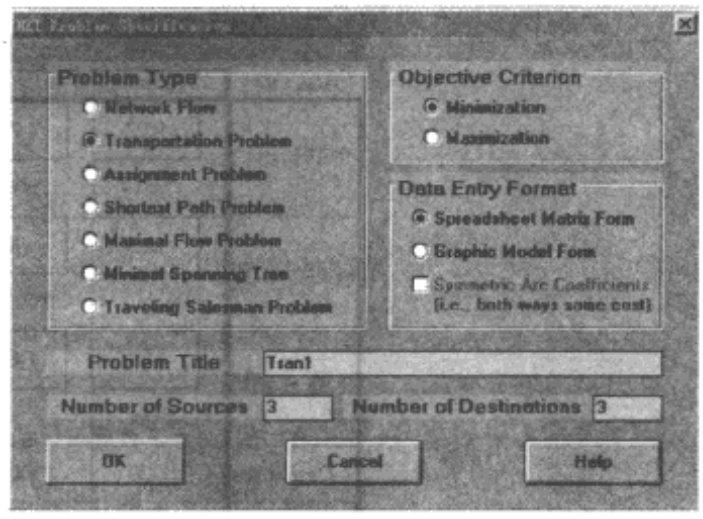


图 5.29

(1) Row Minimum(RM) 逐行最小元素法；

(2) Modified Row Minimum(MRM) 修正的逐行最小元素法；

(3) Column Minimum(CM) 逐列最小元素法；

(4) Modified Column Minimum(MCM) 修正的逐列最小元素法；

(5) NorthWest Corner Method(NWC) 西北角法；

(6) Matrix Minimum(MM) 矩阵最小元素法，即最小元素法；

(7) Vogel's Approximation Method(VAM) Vogel 近似法；

(8) Russell's Approximation Method(RAM) Russell 近似法。

如果不选择，系统的默认方法是逐行最小元素法(RM)。

如果选择矩阵最小元素法(MM) 和“Solve the Display Steps – Tableau”，得到如图 5.31 所示的初始表。由表可以看到入基、出基变量，还可以得到位势即对偶变量[Dual P(i), Dual P(j)]，求出检验数。继续迭代得到最优方案表，如图 5.32 所示。此时，最优调运方案为：中文书刊出口部调运 7 500 册寄往日本、调运 2 500 册寄往中国香港、调运 5 000 册寄往韩国，深圳分公司的 7 500 册全部寄往中国香港，上海分公司的 7 500 册全部寄往日本，总费用为 214 000 元。

最后，点击菜单栏“Results” → “Graphic Solution”，系统将以网络图的形式显示最优调运方案，如图 5.33 所示。

From \ To	日本	中国香港特别行政区	韩国	Supply
中文书刊出口部	10.2	7	9	15000
深圳分公司	12.5	4	14	7500
上海分公司	6	8	7.5	7500
Demand	15000	10000	5000	

图 5.30

From \ To	日本	中国香港特别行政区	韩国	Supply	Dual P(i)
文书刊出口	10.2 7500	7 2500	9 5000	15000	0
深圳分公司	12.5	4 7500	14	7500	-3
上海分公司	6 7500	8	7.5	7500	-4.2
Demand	15000	10000	5000		
Dual P(j)	10.2	7	9		
	Objective Value = 214000 (Minim				

图 5.31

02-07-2006	From	To	Shipment	Unit Cost	Total Cost	Reduced Cost
1	中文书刊出口部	日本	7500	10.20	76500	0
2	中文书刊出口部	中国香港特别行	2500	7	17500	0
3	中文书刊出口部	韩国	5000	9	45000	0
4	深圳分公司	中国香港特别行	7500	4	30000	0
5	上海分公司	日本	7500	6	45000	0
	Total	Objective	Function	Value =	214000	

图 5.32

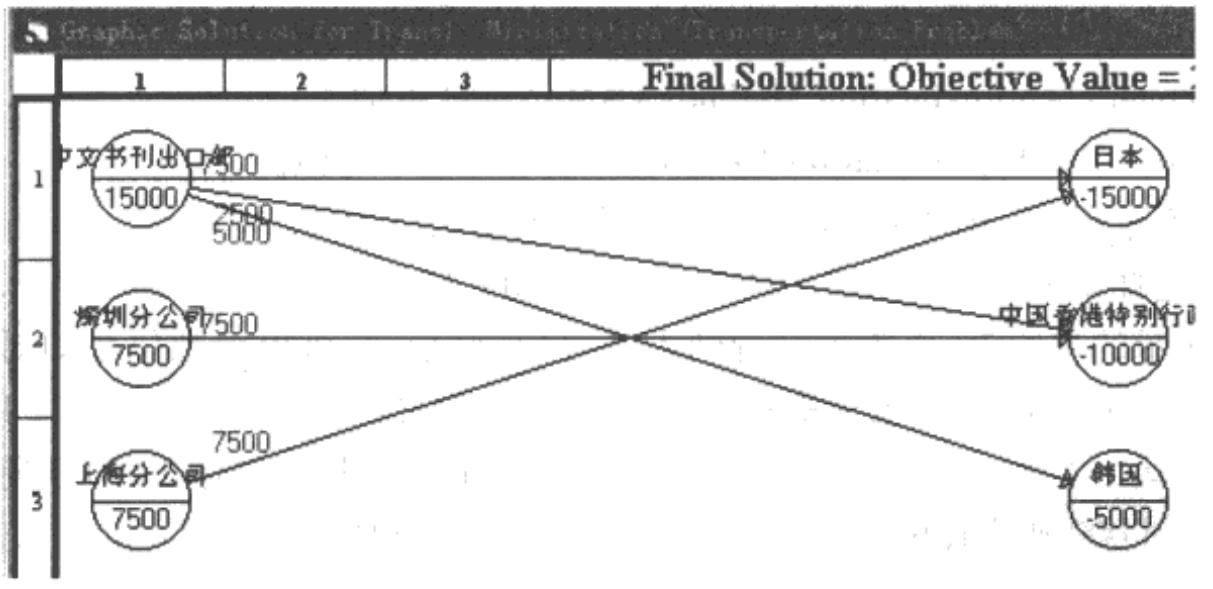


图 5.33

下面,我们给大家介绍怎样运用 WinQSB 软件计算产销不平衡的运输问题,以下例 5.17 水果调运问题为例来说明,这是一个销大于产的问题。

例 5.17 水果调运问题。有三个水果生产基地供应四个地区的某种新鲜水果。假定等量的水果在这些地区受欢迎程度相同。各生产基地年产量,各地区年需求量以及从各生产基地到各地区单位水果的运价如表 5.31 所示,试给出总的运费最节省的水果调运方案。

表 5.31　　运价:万元/万吨

需求地区 / 生产基地	B_1	B_2	B_3	B_4	产量(万吨)
A_1	8	10	5	8	50
A_2	6	—	7	6	40
A_3	12	14	15	10	80
销量(万吨)	60	30	70	20	

用软件求解不用把产销不平衡问题化为平衡问题,令 $c_{22} = M$,软件实施步骤和例 5.12 的一样,我们把文件名取为“Tran2”,输入产地数目 3 和销地数目 4,点击“OK”后按照表 5.31 输入数据,得到图 5.34,如果选择西北角法(NWC)、“Solve the Display Steps - Tableau”,得到如图 5.35 所示的初始表。由表可以看到入基、出基变量,还可以得到位势即对偶变量[Dual P(i),Dual P(j)],求出检验数。继续迭代得到最优方案表,如图5.36 所示。此时,最优调运方案为:A_1 生产基地运送 50 万吨水果供应 B_3 地区;A_2 生产基地分别运送 20 万吨水果供应 B_1 和 B_3 地区;A_3 生产基地运送 40 万吨水果供应 B_1 地区,分别运送 20 万吨水果供应 B_2 和 B_4 地区;B_2 地区有 10 万吨水果需求不能满足。总费用为 1 470 万元。

From \ To	B1	B2	B3	B4	Supply
A1	8	10	5	8	50
A2	6	M	7	6	40
A3	12	14	15	10	80
Demand	60	30	70	20	

图 5.34

	B1	B2	B3	B4	Supply	Dual P(i)
A1	8 50	10	5	8	50	0
A2	6 10	+1M 30	7 0	6	40	-2
A3	12	14	15 70	10 10	80	6
nfilled_Deme	+1M	+1M Cij=2-1M ***	+1M	+1M 10*	-60	-4+1M
Demand	30	70	20	-10		
Dual P(j)	8	2+1M	9	4		
	Objective Value = 40M+1610 (Minimizatio					
	*** Entering: Unfilled_Demand to B2 * Leaving: Unfilled_Demand t					

图 5.35

02-07-2006	From	To	Shipment	Unit Cost	Total Cost	Reduced Cost
1	A1	B3	50	5	250	0
2	A2	B1	20	6	120	0
3	A2	B3	20	7	140	0
4	A3	B1	40	12	480	0
5	A3	B2	20	14	280	0
6	A3	B4	20	10	200	0
7	Unfilled_Demand	B2	10	0	0	0
	Total	Objective	Function	Value =	1470	

图 5.36

最后,点击菜单栏“Results → Graphic Solution”,系统以网络图的形式显示最优调运方案,如图 5.37 所示。

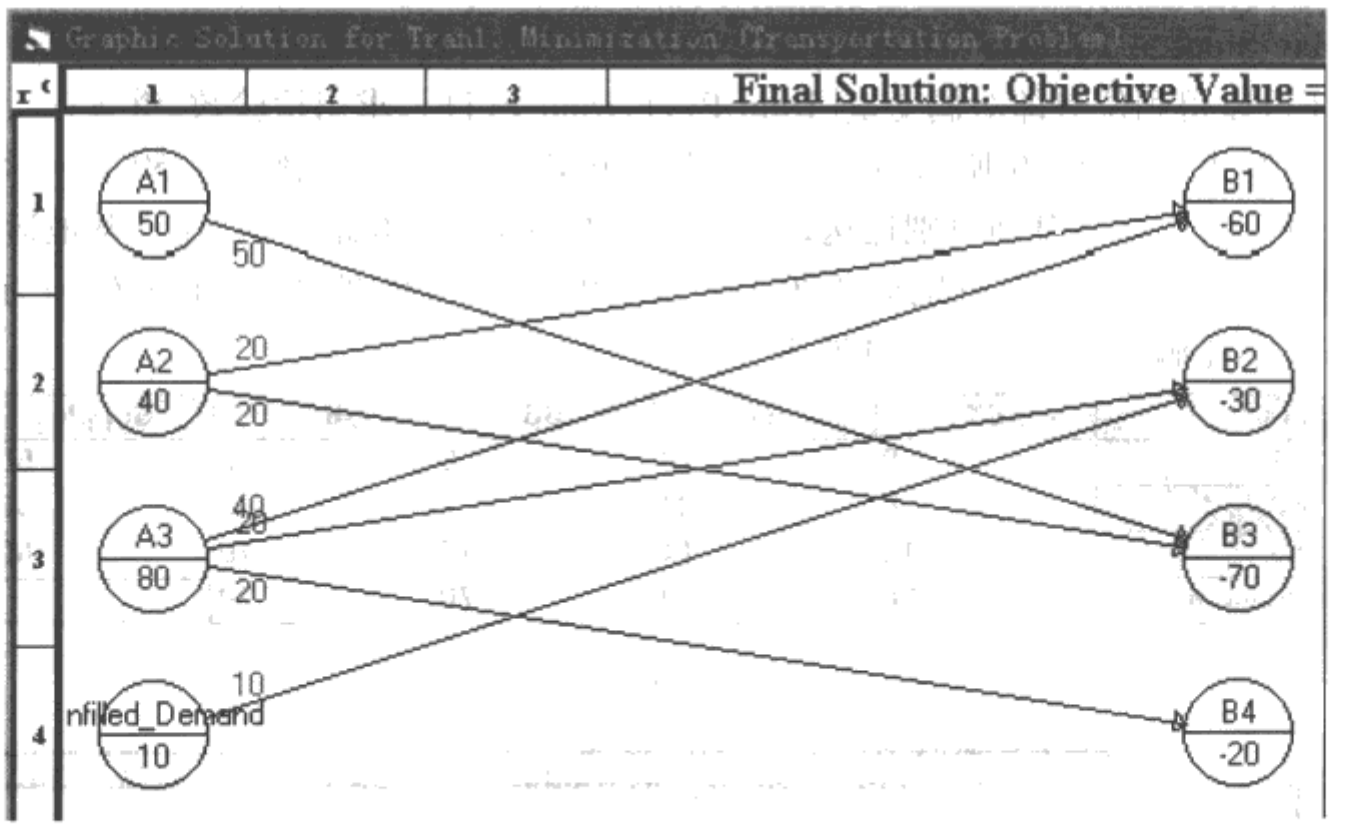


图 5.37

(二) 综合生产计划问题

对于这类问题,读者可以将其转化成平衡运输问题来求解,但 WinQSB 软件提供了此类综合生产计划问题的求解模块。为此,我们举一例介绍 WinQSB 软件的操作方法。

例 5.18 某企业未来四个季度的需求量、生产能力及有关费用如表5.32 所示,试制定全年总费用最小的生产计划。

表 5.32

	第一季度	第二季度	第三季度	第四季度
1. 各时期预测需求量(件)	500	950	1 600	650
2. 正常时间生产能力	400	500	850	450
3. 正常时间生产单位成本(千元)	1.1	1.3	1.2	1.4
4. 加班时间生产能力	150	150	150	90
5. 加班时间生产单位成本(千元)	1.5	1.5	1.5	1.5
6. 期初存量(+)或延期交货量(-)	300			

表 5.32(续)

	第一季度	第二季度	第三季度	第四季度
7. 最小期末存量(安全存量)				350
8. 单位产品每季度贮存费(千元)	0.2	0.2	0.2	0.2
9. 转包(外协)生产能力	300	300	300	300
10. 转包生产单位产品成本(千元)	1.8	1.8	1.8	1.8

调用 WinQSB 软件的子程序 Aggregate Planning,建立新问题,在选项对话框中选中“Transportation Model、Overtime Allowed”及“Subcontracting Allowed”,输入文件名“Aggp1”(读者自己可以任意取名),计划时期数4和期初存量300。如果期初还要补充上期的缺货量(延迟交货,Backorder),则输入负数,如图 5.38 所示。点击“OK”,弹出数据输入对话框,输入数据,重命名计划时期,得到图 5.39。点击菜单栏“Solve and Analyze” → “Solve the Problem”,显示图 5.40 的生产计划表。点击菜单栏“Results → Show Transportation Tableau”,显示类似运输问题运价 —— 运量的最优表,限于篇幅,图 5.41 只显示了一部分内容,这样我们就可以得到完整的生产计划。比如,对于第一季度,期初库存量在第一季度交货;正常时间生产 400 件产品,第一、第二季度分别交货 200 件;加班时间生产 150 件产品用于第二季度交货 100 件,第三季度交货 50 件;转包生产 110 件用于第四季度末库存。总费用为 5 654 千元。

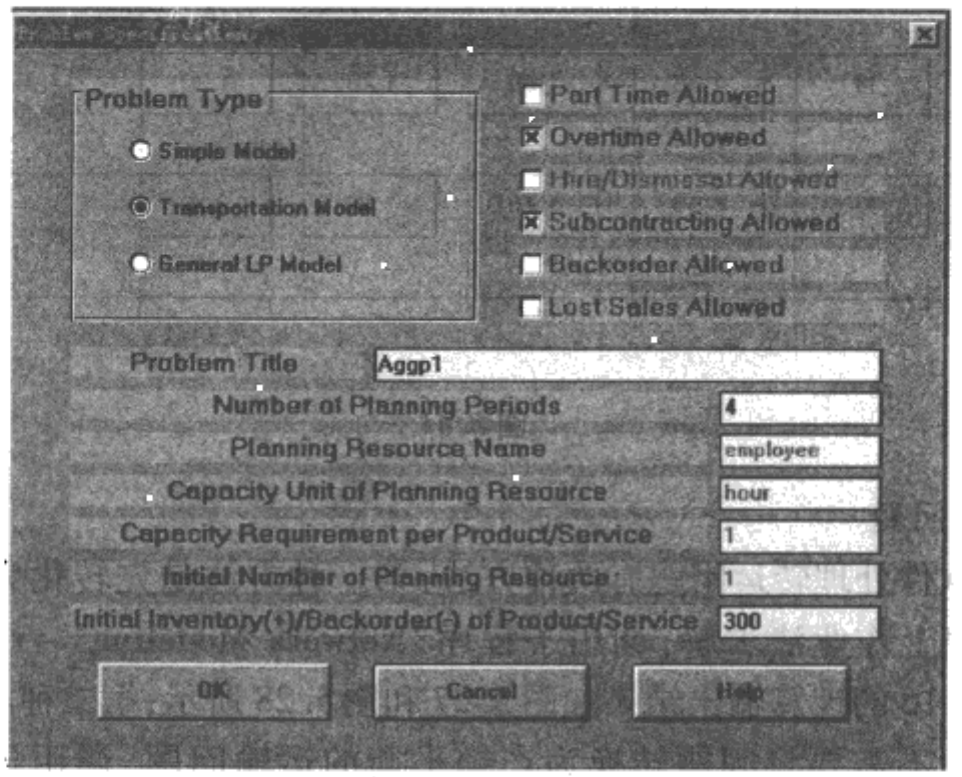


图 5.38

DATA ITEM	第一季度	第二季度	第三季度	第四季度
Forecast Demand	500	950	1600	650
Regular Time Capacity in Unit	400	500	850	450
Regular Time Cost per Unit	1.1	1.3	1.2	1.4
Overtime Capacity in Unit	150	150	150	90
Overtime Cost per Unit	1.5	1.5	1.5	1.5
Initial Inventory (+) or Backorder (-)	300			
Minimum Ending Inventory (Safety Stock)				350
Unit Inventory Holding Cost	0.2	0.2	0.2	0.2
Subcontracting Capacity in Unit	300	300	300	300
Unit Subcontracting Cost	1.8	1.8	1.8	1.8

图 5.39

02-07-2006 20:01:24	Demand	Regular Production	Overtime Production	Subcontracting Production	Total Production	Ending Inventory
Initial						300.00
第一季度	500.00	400.00	150.00	110.00	660.00	460.00
第二季度	950.00	500.00	150.00	300.00	950.00	460.00
第三季度	1,600.00	850.00	150.00	300.00	1,300.00	160.00
第四季度	650.00	450.00	90.00	300.00	840.00	350.00
Total	3,700.00	2,200.00	540.00	1,010.00	3,750.00	1,430.00

图 5.40

	第一季度 Demand	第二季度 Demand	第三季度 Demand	第四季度 Demand	Ending Inventory	Unused Capacity	Total Capacity
Initial Inventory	0.00 300	0.20	0.40	0.60	0.80	0	300
第一季度 Regular time	1.10 200	1.30 200	1.50	1.70	1.90	0	400
第一季度 Overtime	1.50	1.70 100	1.90 50	2.10	2.30	0	150
第一季度 Subcontract	1.80	2.00	2.20	2.40	2.60 110	190	300
第二季度 Regular time	M	1.30 500	1.50	1.70	1.90	0	500
第二季度 Overtime	M	1.50 150	1.70	1.90	2.10	0	150

图 5.41

（三）转运问题

WinQSB 软件处理转运问题有两种方法：第一种方法是先转化为产销平衡运价表，然后运用表上作业法求解，调用子程序“Network Modeling” → “Transportation Problem”，例 5.16 的计算输出结果如图 5.27 和图 5.28 所示；第二种方法是将问题看成一般网络图，不需要将问题转换为产销平衡的运输问题，调用子程序“Network Modeling” → “Network Flow”，输入数据时，中转地与需求地的供应量为 0，供应地与中转地的需求量为 0，运价按实际发生的运价输入，本地到本地和不可到达的，保持空白不需要输入运价。数据输入表格如图 5.42 所示。

From \ To	A1	A2	A3	A4	A5	A6	A7	A8	A9	Supply
A1				2	5	3	2			200
A2			1	4	2					300
A3				8	7				15	500
A4		3			3	4	3	6		0
A5							1	2	9	0
A6							2			0
A7										0
A8									3	0
A9										0
Demand	0	0	0	0	0	450	200	150	200	

图 5.42

点击菜单栏“Solve and Analyze”→“Solve the Problem”,显示图 5.43 的最优运输方案。点击菜单栏“Results”→“Graphic Solution”,得到最优运输网络图(见图 5.44)。

02-07-2006	From	To	Flow	Unit Cost	Total Cost	Reduced Cost
1	A1	A6	200	3	600	0
2	A2	A5	300	2	600	0
3	A3	A4	250	8	2000	0
4	A3	A5	250	7	1750	0
5	A4	A6	250	4	1000	0
6	A5	A7	200	1	200	0
7	A5	A8	350	2	700	0
8	A8	A9	200	3	600	0
	Total	Objective	Function	Value =	7450	

图 5.43

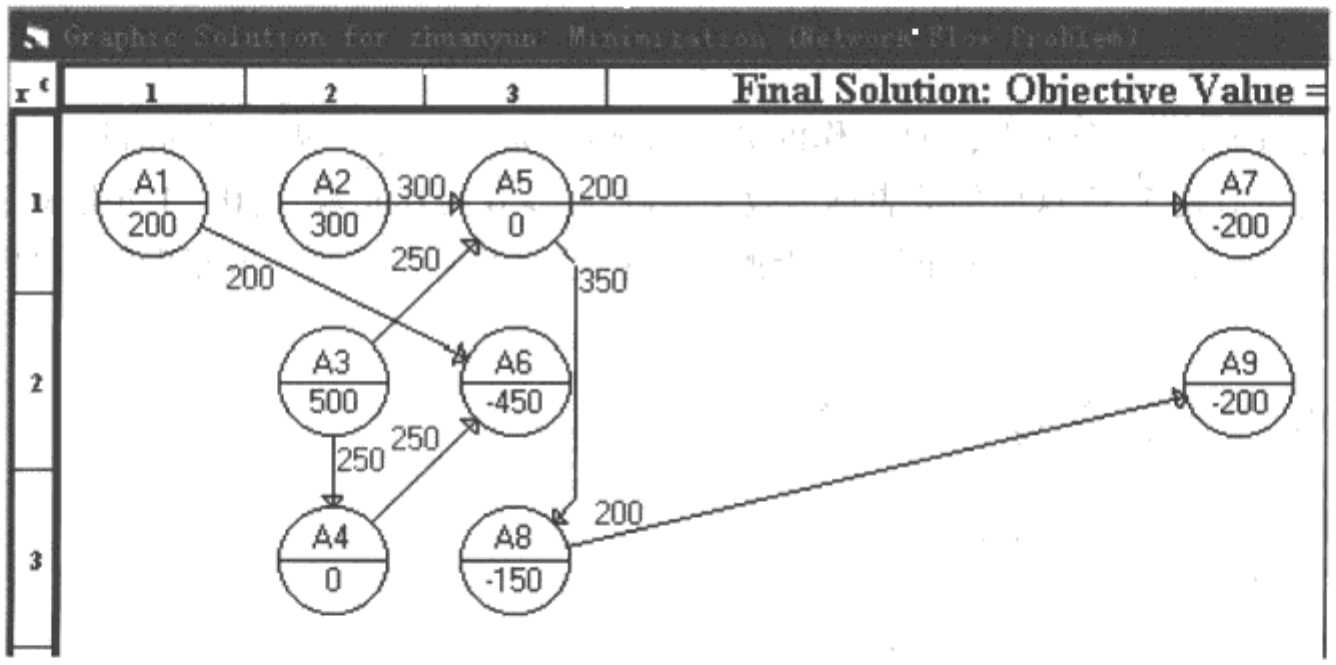


图 5.44

第六节　指派问题

在现实生活中，人们经常遇到这样的问题，某单位有项任务需要完成，每一项任务只能由一个人来承担，由于每个人的专长存在差异，各人完成任务的时间或者费用（效率）也就不同，因而需要明确指派谁去完成哪项任务才能使完成该项任务的总效率最高（所需时间或费用等最小）。这类问题通常被称为指派问题（Assignment Problem），也叫作分配或配置问题。

一、数学模型

例 5.19　项目合同翻译问题。有一份中文项目合同，需要译成英、日、德、意四种语言文本，分别用字母 E、J、G、I 表示这几种语言。现在有甲、乙、丙、丁四个翻译去完成这些工作，由于个人所擅长的语种不同，他们将中文项目合同翻译成不同语种文本所需要的时间，如表 5.33 所示。应指派谁完成哪种翻译工作，才能保证所需总时间最少？

表 5.33

翻译人员 \ 翻译工作	E	J	G	I
甲	2	10	9	7
乙	15	4	14	8
丙	13	14	16	11
丁	4	15	13	9

类似地，有 n 个工人，n 件工作的分配问题；有 n 条航线，n 架飞机的航行问题等。我们把类似表 5.33 这样的数表称为效率矩阵，$c_{ij} > 0$ 为其元素，用 0 - 1 变量 x_{ij} 表示指派第 i 个人去完成第 j 项任务，当指派时取值为 1，否则为 0，则一般指派问题的数学模型为：

$$\text{Min } Z = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij}$$

$$s.t.\begin{cases}\sum\limits_{i=1}^{n} x_{ij} = 1 \quad j = 1,2,\cdots,n \\ \sum\limits_{j=1}^{n} x_{ij} = 1 \quad i = 1,2,\cdots,n \\ x_{ij} = \begin{cases}1 & \text{当指派第 } i \text{ 个人去完成第 } j \text{ 项任务时} \\ 0 & \text{当不指派第 } i \text{ 个人去完成第 } j \text{ 项任务时}\end{cases}\end{cases}$$

二、指派问题的匈牙利解法

指派问题是 0 - 1 整数规划问题的特例，也是运输问题的特例，此时，$n = m$，$a_i =$

$b_j = 1$。可以运用整数规划、0 - 1 规划或运输问题的解法去求解，但这和用单纯形法求解运输问题一样计算量大，划不来，解决指派问题有更简便的解法。

匈牙利数学家康尼克(*D. Konig*) 证明了下面的定理5.4 和定理5.5，数学家库恩(W. W. Kuhn) 基于这两个定理提出了指派问题的解法，这一解法虽然后来几经改进，但一直被称为匈牙利解法，这种方法简单有效，适用的条件是：问题求最小值、人员数目与工作数目相等以及效率矩阵中的元素非负。

定理 5.4 如果从指派问题效率矩阵$[c_{ij}]_{n\times n}$ 的每一行元素中分别减去(或加上) 一个常数p_i(称为该行的位势)，从每一列分别减去(或加上) 一个常数q_j(称为该列的位势)，得到一个新的效率矩阵$[b_{ij}]_{n\times n}$，其中$b_{ij} = c_{ij} - p_i - q_j > 0$，则$[b_{ij}]_{n\times n}$ 的最优解等价于$[c_{ij}]_{n\times n}$ 的最优解。

定理 5.5 系数矩阵中独立的0 元素的最大数目等于覆盖该矩阵中所有0 元素的最小直线数。

匈牙利解法的全部解题过程可分为四个步骤，下面我们结合例5.19 进行说明：

1. 根据效率矩阵获得机会成本矩阵

所谓机会成本即比采用最优方案多支付的成本。比如，如果分配人员甲完成任务J则多支付10 - 2 = 8 的时间，等等。逐行计算后得到机会成本矩阵，如表5.34 所示。由于第三列无0 元素，计算列机会成本，得到表5.35，此时每一列都有0 元素，进入下一步。

表 5.34

	E	*J*	*G*	*I*
甲	0	8	7	5
乙	11	0	10	4
丙	2	3	5	0
丁	0	11	9	5

表 5.35

	E	*J*	*G*	*I*
甲	0	8	2	5
乙	11	0	5	4
丙	2	3	0	0
丁	0	11	4	5

2. 进行最优性检验

检验的方法是：用尽可能少的水平或垂直的直线覆盖所有的0，如果所画的直线数恰好等于矩阵的行数或列数，即为最优解。要使所画的直线数为最少，一般应首先从0 最多的行或列开始划起，直至划完为止。如表5.36 所示，删去全部0 格后所画的最少直线数为3，小于行数或列数4，故未达到最优指派，需要对此方案进行调整。

表 5.36

	E	*J*	*G*	*I*
甲	0	8	2	5
乙	11	0	5	4
丙	2	3	0	0
丁	0	11	4	5

3. 调整方案

调整的方法是:在所有的未划去的元素中间去掉最小元素,然后在所画直线的交叉点处加上这一最小元素,以增加矩阵中0元素的数目。在本例中,相当于在第1行和第4行减最小元素,在第1列加最小元素,调整后得到表5.37。

表5.37

	E	*J*	*G*	*I*
甲	0	6	0	3
乙	13	0	5	4
丙	4	3	0	0
丁	0	9	2	3

调整以后重复2和3,直到满足最优性条件为止。本例检验后的矩阵如表5.38所示。可以看出,所画直线的数目已经与矩阵的行数或列数相等,已经获得最优解。

表5.38

	E	*J*	*G*	*I*
甲	0	6	0	3
乙	13	0	5	4
丙	4	3	0	0
丁	0	9	2	3

4. 进行试指派,确定最优方案

进行试指派时,在 n 很小的情况下,可用观察法或试探法找出 n 个独立的0元素即可;在 n 比较大的情况下,则可以按如下步骤进行:

(1) 先从只有一个0元素的行开始,给这个0元素加上标记,将其所在的列的其他0元素划去,表示这一列的任务已经指派;

(2) 再从只有一个0元素的列开始,给这个0元素加上标记,将其所在的行的其他0元素划去;

(3) 重复(1)和(2),直至所有的0元素都被加上标记或被划去;

(4) 当标记的0元素的数目等于矩阵的行数或列数时,就找到了最优指派方案。

本例中指派的结果见表5.39,最优解为:$x_{13} = x_{22} = x_{34} = x_{41} = 1$,其余变量为0,最优目标函数值为 $Z = 28$。最优指派方案为:甲负责将汉语译为德语、乙负责将汉语译为日语、丙负责将汉语译为意大利语、丁负责将汉语译为英语的翻译工作。

表5.39

	E	*J*	*G*	*I*
甲	0	6	0*	3
乙	13	0*	5	4
丙	4	3	0	0*
丁	0*	9	2	3

三、指派问题的扩展

在标准形式指派问题中，我们遇到的问题是求最小值、人员数目与工作数目相等以及效率矩阵中的元素非负这些条件下的情况，作为一种扩展，我们还有可能遇到以下几个方面的非标准形式的指派问题，但是均可以通过一定的方法将它转换成标准形式的指派问题来求解。

（一）最大值指派问题

对于极大化的问题，即求 $\text{Max } Z = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij}$，设 $b_{ij} = M - c_{ij} > 0$（M 为一足够大的常数，如 c_{ij} 中的最大数），于是目标函数变为 $\text{Min } Z' = \sum_{i=1}^{n}\sum_{j=1}^{n}(M - c_{ij})x_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} Mx_{ij} - \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij} = nM - Z$，因 M 为常数，Z' 的最小化就是 Z 的最大化。

所以，将最大值指派问题转换成标准指派问题的方法是在效率矩阵中，用它的最大元素减去其余各元素，得到初始的机会成本矩阵。

（二）人员数目与工作数目不相等的指派问题

若人少工作多，则可加入一些虚拟的“人”，这些虚拟的“人”做各件工作的费用取值为0，因为实际上，这些费用不会发生；若人多工作少，则可加入一些虚拟的“工作”，这些虚拟的“工作”的费用取值也为0。

（三）一个人可做几件工作的指派问题

若一个人可做几件工作，则可以把该人当成“几个人”来安排，这几个人做同一件工作的费用是一样的。

（四）某工作不能由某人做的指派问题

若某工作要求不能由某人来做，则可将相应工作费用设定为一个足够大的常数 M。

四、指派问题的 WinQSB 软件应用

在 WinQSB 软件的网络流模块中，指派问题的求解采用的是上面介绍的匈牙利解法。下面我们以例 5.19 的项目合同翻译问题为例，说明怎样应用 WinQSB 软件计算指派问题。

首先，调用 WinQSB 软件的子程序 Network Modeling，建立一个新问题，弹出对话框，如图 5.45 所示界面，选择“Assignment Problem”，输入问题的文件名“Assig1”（读者自己可以任意取名），人数及任务数。然后，点击“OK”，此时弹出一张需要输入数据的表格，对照表 5.33 输入数据，

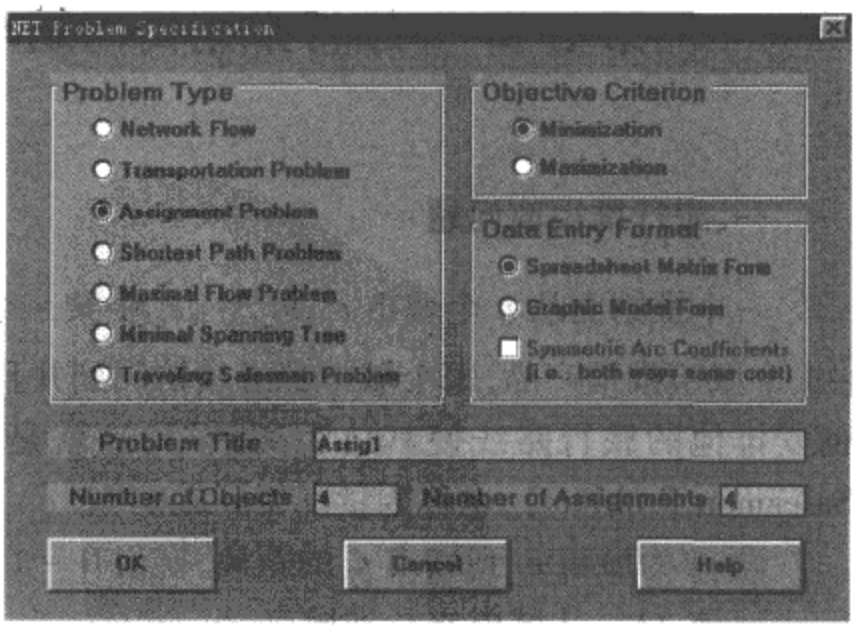


图 5.45

重命名网络节点后得到图5.46。与运输问题的求解方法一样，点击“Solve the Display Steps - Tableau”时，系统输出匈牙利解法的每一步迭代结果，如图5.47到图5.48所示。点击菜单栏“Results”→“Graphic Solution”，以网络图的形式显示结果。计算结果与例题完全相同。

From \ To	E	J	G	I
甲	2	10	9	7
乙	15	4	14	8
丙	13	14	16	11
丁	4	15	13	9

图5.46

Hungarian Method for Assigl - Iteration 1

From \ To	E	J	G	I
甲	0	8	2	5
乙	11	0	5	4
丙	2	3	0	0
丁	0	11	4	5

图5.47

Hungarian Method for Assigl - Iteration 2 (Final)

From \ To	E	J	G	I
甲	0	6	0	3
乙	13	0	5	4
丙	4	3	0	0
丁	0	9	2	3

图5.48

02-07-2006	From	To	Assignment	Unit Cost	Total Cost	Reduced Cost
1	甲	G	1	9	9	0
2	乙	J	1	4	4	0
3	丙	I	1	11	11	0
4	丁	E	1	4	4	0
	Total	Objective	Function	Value =	28	

图5.49

第七节　旅行推销商问题与中国邮递员问题

一、旅行推销商问题

一个推销商从 n 个城市 $v_1, v_2, \cdots, v_n$ 中某一个城市 $v_i(1 \leqslant i \leqslant n)$ 出发，到其他 $n-1$ 个城市推销产品，每个城市都必须访问到并且只经过一次，最后回到 v_i，那么应如何安排他的旅行路线使其总距离最短。这就是著名的旅行推销商问题(Traveling Salesmen Problem)，也称为货郎担问题。

设 c_{ij} 为城市 v_i 到城市 v_j 的距离，定义0-1变量：

$$x_{ij} = \begin{cases} 1 & \text{从城市旅行到城市} \\ 0 & \text{否则} \end{cases}$$

则旅行推销商问题的数学模型可表示成为一个整数规划问题。

$$\text{Min } Z = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij} \qquad i \neq j$$

$$s.t.\begin{cases} \sum_{i=1}^{n} x_{ij} = 1 \quad j = 1,2,\cdots,n(i \neq j) \\ \sum_{j=1}^{n} x_{ij} = 1 \quad i = 1,2,\cdots,n(i \neq j) \\ x_{ij} + x_{ji} \leqslant 1 \quad i \neq j \\ x_{ij} + x_{jk} + x_{ki} \leqslant 2 \quad i \neq j \neq k \\ \cdots \\ x_{ij} - x_{jk} + x_{ki} + \cdots + x_{pi} \leqslant n-2 \quad i \neq j \neq \cdots \neq p \\ x_{ij} = 0 \text{ 或 } 1 \quad i = 1,2,\cdots,n \end{cases}$$

虽然旅行推销商问题能够运用整数规划和动态规划等方法求解，但是当问题较大时求解就会遇到一定的困难，一种可行且有效的方法是求最小的 Hamilton 回路。

定义 5.13 ［Hamilton 回路］设 $v_1, v_2, \cdots, v_n$ 是图 G 中的 n 个顶点，若有一条从某一顶点 $v_i(1 \leqslant i \leqslant n)$ 出发，经过各节点一次且仅一次，最后返回出发点 v_i 的回路，则称此回路为 Hamilton 回路。

定理 5.6 若图 G 中任意两个点 v_i、v_j 满足 $d(v_i) + d(v_j) \geqslant n$（$n$ 为图 G 的点数并且 $n \geqslant 3$），则 G 中存在 Hamilton 回路。

因而，旅行推销商所走的路线是一个由 n 个城市构成的交通图 G 的一个 Hamilton 回路，旅行推销商问题就是寻找一个总距离最小的 Hamilton 回路。下面我们举一例来说明怎样运用 WinQSB 软件求解旅行推销商问题。

例 5.20 某电动汽车公司与某旅游景区合作，打算在旅游景区内开通无污染无噪音的“绿色交通”路线。图 5.50 是旅游景区各景点的分布图，其中 C、F 之间是两条单向通道，边上的数字为汽车通过两点间的正常时间（分钟）。电动汽车公司应该如何设计一条路线，使汽车通过每一处景点一次且总时间最少？

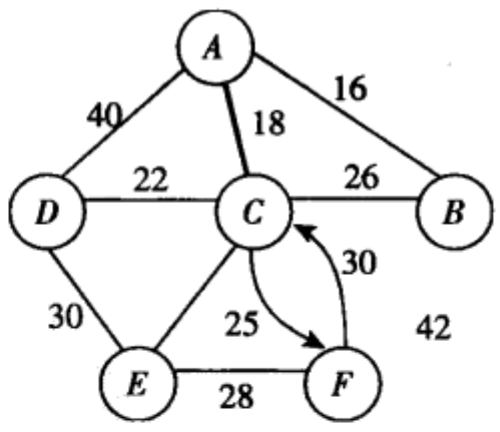


图 5.50

解：在图 5.50 中，我们可以找到 $d(A) + d(D) = 6 > n = 6$，由定理 5.6 可得图 5.50 中存在 Hamilton 回路。将上图表示成距离矩阵，如表 5.40 所示，两点间没有边连接的时间为 ∞。

表 5.40

	A	B	C	D	E	F
A	∞	16	18	40	∞	∞
B	16	∞	26	∞	∞	42
C	18	26		22	15	25
D	40	∞	22	∞	30	∞
E	∞	∞	15	30	∞	28
F	∞	42	30	∞	28	∞

调用 WinQSB 软件的子程序 Network Modeling，建立一个新问题，弹出对话框，如图 5.51 所示界面，选择“Traveling Salesmen Problem”，目标函数选择“Minimization”，输入问题的文件名“Travel1”（读者自己可以任意取名），输入节点数。然后，点击“OK”，此时弹出一张需要输入数据的表格，对照表 5.40 输入数据，输入数据时应注意弧的方向，无向弧对称输入数据，有向弧按弧的方向输入，重命名网络节点后得到图 5.52。

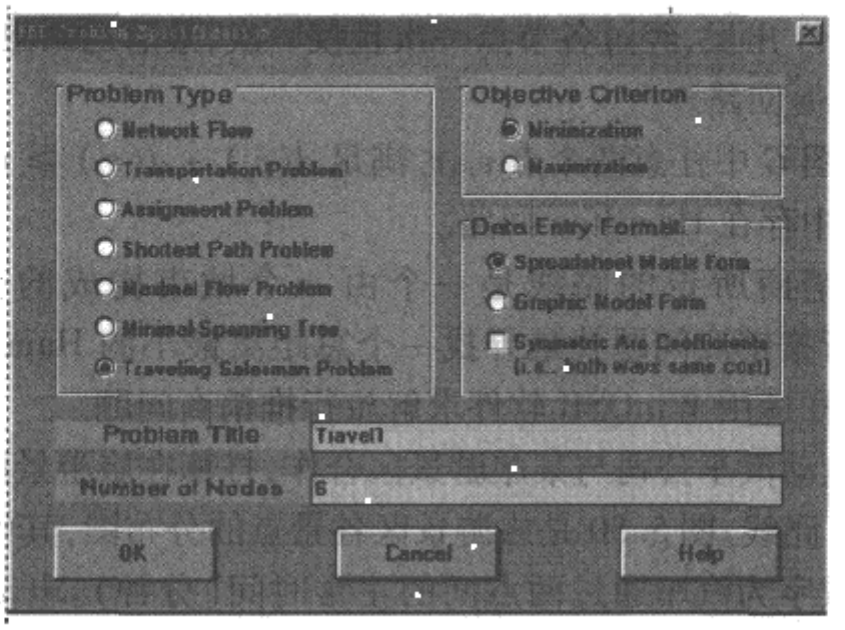


图 5.51

From \ To	A	B	C	D	E	F
A		16	18	40		
B	16		26			42
C	18	26		22	15	25
D	40		22		30	
E			15	30		28
F		42	30		28	

图 5.52

点击菜单栏“Solve and Analyze → Solve the Problem”后，系统提示用户选择求解方法，对于较大的旅行推销商问题目前还没有有效的解法，系统提供了三种近似解法和一种分支定界法。分支定界法能得到最优解，对于大型问题求解时间可能达数小时。三种近似方法为最近城市法（Nearest Neighbor Heuristic）、逐步包围法（Cheapest Insertion Heuristic）、两两交换改进法（Two - way Exchange Improvement Heuristic），

不同的方法得到的结果可能不一样,其原则是取最短回路的解。

对于本问题,用最近城市法无解、用逐步包围法回路长度为168,用两两交换改进法与分支定界法的结果一样,为156,可作为求解的结果。点击"Results → Graphic Solution"得到旅行线路图5.53。

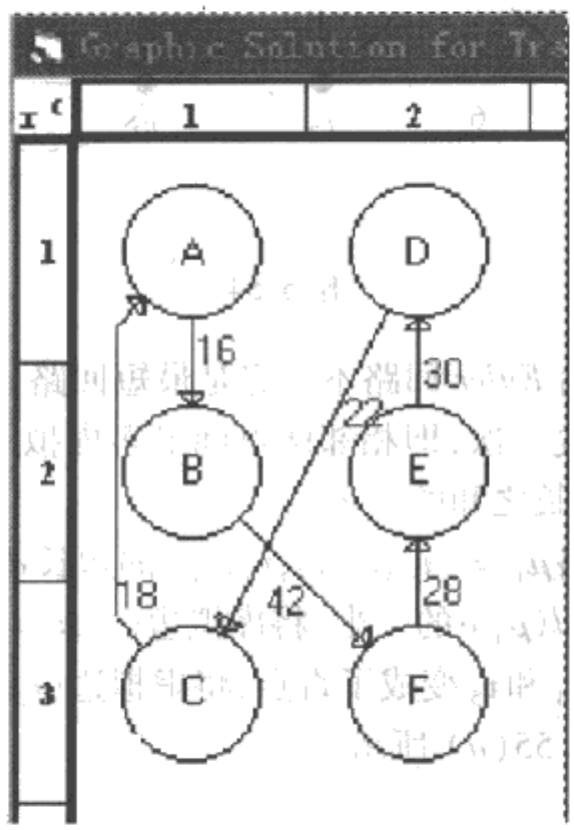


图5.53

二、中国邮递员问题

一个邮递员从邮局出发,遍历他所管辖的每一条街道,将邮件送到后返回邮局,要求所走的路径最短。这个问题是由中国数学家管梅谷教授于1962年提出来的,大家习惯地将其称为中国邮递员问题。该问题与旅行推销商问题的不同之处在于前者是遍历图的所有边,后者是遍历图的所有点。用图论的语言,中国邮递员问题可以表述成为在图 G 中,求一条回路,经过网络的各条边至少一次,使回路中各条边的权数之和最小。

定义5.14 [*Euler* 回路,*Euler* 图] 对于连通的无向图 G,若存在一条回路,不重复包含 G 的每一条边,这条回路称为 *Euler* 回路,具有 *Euler* 回路的图称为 *Euler* 图。

定理5.7 连通无向图 G 存在一条 *Euler* 回路的充要条件是所有顶点的次都是偶数。

例5.21 在图5.54(a)中有4个奇点 v_1、v_2、v_6、v_7,不存在 *Euler* 回路,无论邮局在哪一个点,邮递员要经过每一条边至少会有一条边重复经过。若将其增加4条边变为图5.54(b),4条虚线就等价于邮递员重复经过的边,图5.54(b)中所有点都是偶点,因而是 *Euler* 图,存在 *Euler* 回路。那么,中国邮递员问题就变成在一个具有奇点的图中,如何将奇点连接起来变为偶点,成为 *Euler* 图,使各边长度之和最短。试用这种思想求解图5.54(a)的中国邮递员问题。

解:(1) 虚拟边所有奇点变为偶点,如图5.54(b)所示,虚拟边是邮递员重复经过的街道。

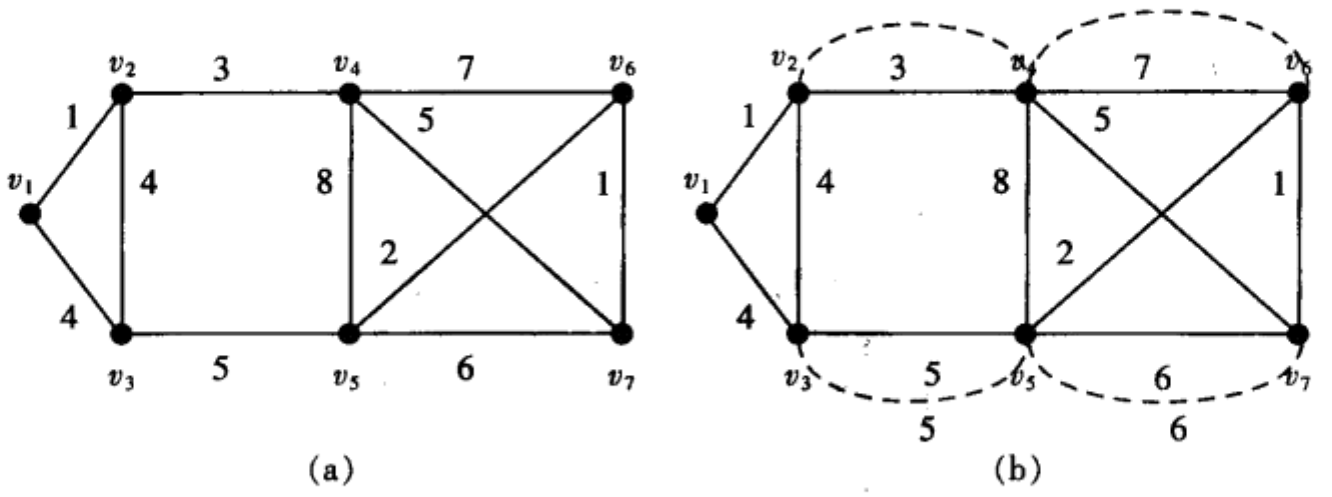


图 5.54

(2) 调整虚拟边。初始 *Euler* 回路不一定是最短回路。判断是否是最短回路的准则是:(a) 每条边最多重复一次,即相邻两点间最多虚拟一条边;(b) 所有回路中虚拟边长之和不超过回路边长之和的一半。

在图5.54(b) 中,回路$\mu_1 = \{v_4, v_5, v_7, v_6, v_4\}$ 的边长$d(\mu_1) = 8 + 6 + 1 + 7 = 22$,其中虚拟边长为 13,超过 $d(\mu_1)$ 的一半,将虚拟边(v_4, v_6) 和(v_5, v_7) 去掉,在 v_6 与 v_7 之间加一条虚拟边。这时v_4 和v_5 变成了奇点,将虚拟边(v_2, v_4) 和(v_3, v_5) 改为虚拟边(v_1, v_2) 和(v_1, v_3),如图 5.55(a) 所示。

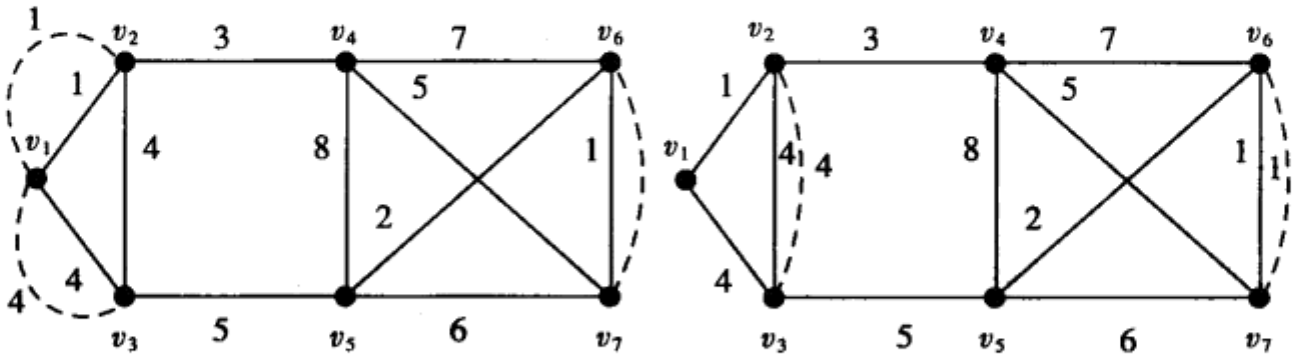


图 5.55

(3) 检查图5.55(a),回路$\mu_2 = (v_1, v_2, v_3, v_1)$ 的边长$d(\mu_2) = 1 + 4 + 4 = 9$,虚拟边长为 15,需要调整,将虚拟边(v_1, v_2) 和(v_1, v_3) 去掉,在(v_2, v_3) 之间添加虚拟边(v_2, v_3),如图 5.55(b) 所示。

(4) 继续检查,所有回路满足最短回路的准则,图5.55(b) 是最短的 Euler 回路,其中边(v_2, v_3) 和(v_6, v_7) 各重复一次。

第八节　网络计划

例 5.22　某项目的工序资料如表 5.41 所示,考虑该项目的网络计划图,并以正常时间为标准,分别计算提前一周完工增加收益 6 万元和 9 万元时成本最低的完成情况;如果给定预算成本 544 万元,项目的完工时间是多少?项目施工成本进度是怎

样的?项目施工到20周时工序完成情况怎样?

表5.41

工序	紧前工序	时间(周)		成本(万元)		时间的最大缩量(周)	应急增加成本(万元/周)
		正常	应急	正常	应急		
A		8	6	50	60	2	5
B		20	16	100	140	4	10
C	*A*	9	7	80	130	2	25
D	*B*	9	6	41	50	3	3
E	*B*,*C*	13	11	60	90	2	15
F	*C*,*D*	8	6	80	120	2	0
G	*E*,*F*	15	12	60	84	3	8
总成本				471	674		

这是一个典型的用网络编制计划的问题,人们又常常把它称为网络计划(Network Programming)。网络计划技术由计划评审技术(Program Evaluation and Review Technique,PERT)和关键路线法(Critical Path Method,CPM)组成。PERT是主要针对完成工作时间的不能确定性,是一个随机的计划编制方法,活动的时间通常用三点法估计,注重计划的评价和审查。CPM以经验数据确定工作时间,将它视为确定的数值,主要研究项目的费用与工期的相互关系。我们通常将这两种方法融为一体,统称为网络计划或网络计划技术(PERT/CPM)。

当前,世界上工业发达国家都非常重视网络计划技术在现代管理中的应用。它已被许多国家公认为当前最为行之有效的管理方法之一。大量的工程项目实践表明,应用网络计划技术管理后可以缩短建设周期20%,降低工程成本10%,而编制网络计划所需费用仅为总费用的0.1%。美国是网络技术的发源地,美国政府于1962年规定,凡与政府签订合同的企业,都必须采用网络计划技术,以保证工程进度和质量。1974年麻省理工学院调查指出"绝大部分美国公司采用网络计划编制施工计划"。目前,美国基本上实现了用计算机绘图、优化计算和资源平衡以及项目进度控制,实现了计划工作自动化。我国应用网络计划技术是从20世纪60年代初期开始的。近几年,网络计划技术的应用逐渐得到工程管理人员的重视,并且获得了一些经济效益。如:上海宝钢炼铁厂1号高炉土建工程施工应用网络法缩短工期21%,降低成本9.8%;广州白天鹅宾馆在建设中,运用网络计划技术,工期比与外商签订的合同提前四个半月,仅投资利息就节约了1 000万港币。

特别是近几年来,随着计算方法的改进和网络计划软件的不断开发,网络计划技术已经主要应用于新产品的研制与开发、大型工程项目的计划和编制以及计划的优化,成为了项目管理和项目安排领域比较科学的一种计划编制方法,它对于计划的控制、管理、调整和优化、更清晰地了解工作之间的相互联系和相互制约的逻辑关系以及掌握关键工作和计划的全盘情况,具有重要的作用。

一、网络计划图

（一）网络计划图的基本术语

网络计划的重要标志是网络图。将项目中所有活动之间的衔接关系用箭线（弧）和节点连接起来，弧边的权是完成该活动的时间，这种描述项目计划的网络图称为计划网络图或项目网络图（Project Network）。

（1）项目，也称为工程。它是一项科研试制项目、施工任务、生产任务以及较复杂项目。一个大项目根据不同部门的任务可以分解成若干个子项目。子项目之间相互独立。

（2）箭线、节点。它是网络计划图的基本组成元素。箭线是一段线段带箭头的实射线（用“→”表示），节点是箭线两端的连接点（用“□”或“○”表示。）

（3）工作（也称工序、活动、作业）。它是将整个项目按需要，以粗细程度分解成若干个需要耗费时间或需要耗费其他资源的子项目单元。

（4）描述工程项目网络计划图的两种表达方式：双代号网络计划图和单代号网络计划图。双代号网络计划图在计算时间参数时，又可分为工作计算法和节点计算法。

（5）双代号网络计划图。在双代号网络计划图中，用箭线表示工作，箭尾的节点表示工作的开始点，箭头的节点表示工作的完成点。用（$i-j$）两个代号及箭线表示一项工作，在箭线上标记必需的信息，箭线之间的连接顺序表示工作之间的先后开工的逻辑关系，如图 5.56 所示。

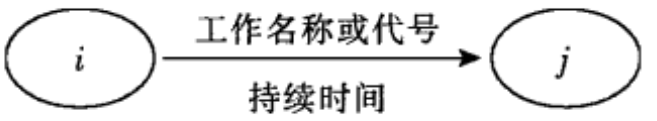


图 5.56

（6）单代号网络计划图。用节点表示工作，箭线表示工作之间的先完成与后完成的关系为逻辑关系。在节点中标记必需的信息，如图 5.57 所示。

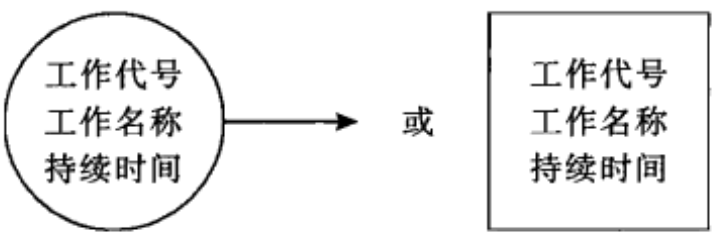


图 5.57

（二）双代号网络计划图

我们在这里主要介绍双代号网络计划图的绘制和按工作计算时间参数的方法。以下通过例 5.23 来说明网络计划图的绘制和时间参数的计算。

例 5.23　开发一个新产品，需要完成的工作和先后关系，各项工作需要的时间汇总在表 5.42 中，要求编制该项目的网络计划图和计算有关参数。我们根据表 5.42 中数据，绘制网络图，如图 5.58 所示。

表 5.42

序号	工作名称	工作代号	工作持续时间(天)	紧后工作
1	产品设计和工艺设计	*A*	60	*B*,*C*,*D*,*E*
2	外购配套件	*B*	45	*L*
3	锻件准备	*C*	10	*F*
4	工装制造 1	*D*	20	*G*,*H*
5	铸件	*E*	40	*H*
6	机械加工 1	*F*	18	*L*
7	工装制造 2	*G*	30	*K*
8	机械加工 2	*H*	15	*L*
9	机械加工 3	*K*	25	*L*
10	装配与调试	*L*	35	/

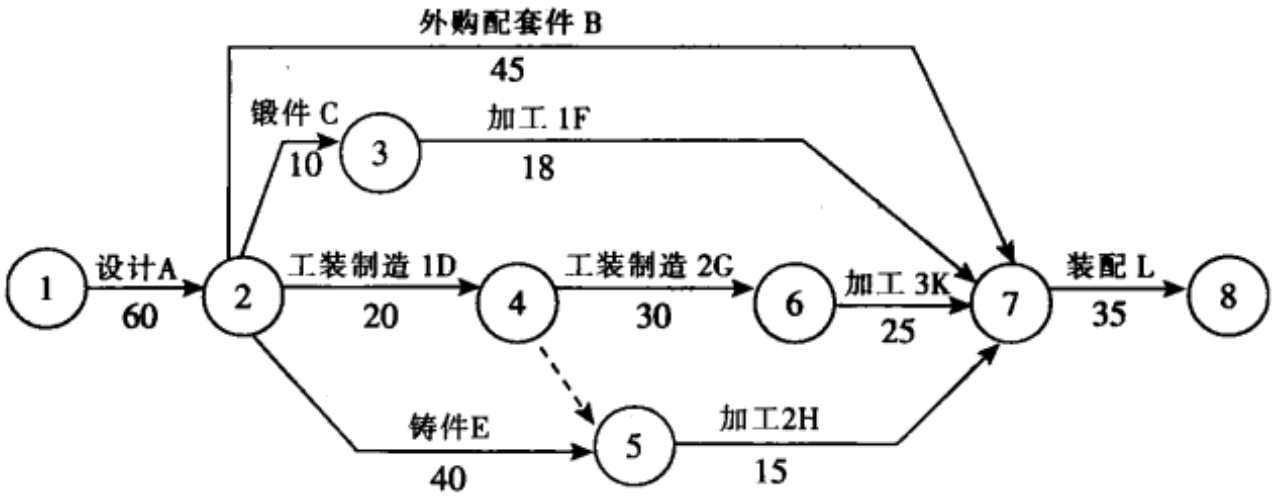


图 5.58

为了正确表述工程项目中各个工作的相互连接关系和正确绘制网络计划图,应该遵循以下规则和术语:

1. 网络计划图的方向、时序和节点编号

网络计划图是有向、有序的赋权图,按项目的工作流程自左向右的绘制。在时序上反映各项工作的先后顺序。节点编号必须按箭尾节点的编号小于箭头节点的编号来标记。在网络图中:只能有一个起始节点,表示工程项目的开始;只有一个终点节点,表示工程项目的完成。

2. 紧前工作和紧后工作

紧前工作是指紧排在本工作之前的工作。紧后工作是指紧排在本工作之后的工作。如在图 5.58 中,只有工作 *A* 完成以后,工作 *B*、*C*、*D*、*E* 才可能开始,工作 *A* 是工作 *B*、*C*、*D*、*E* 的紧前工作;而工作 *B*、*C*、*D*、*E* 则是工作 *A* 的紧后工作。从起始点至本工作之前在同一线路的所有工作,称为先行工作;自本工作到终点节点在同一线路的所有工作,称为后继工作。如在图 5.58 中,工作 *G* 的先行工作有工作 *A* 和工作 *D*;工作 *K* 和工作 *L* 是工作 *G* 的后继工作。

3. 虚工作

在双代号网络计划图中,虚工作是指只表示相邻工作之间的逻辑关系,不占用时

间和不消耗人力、资金的虚设的工作。虚工作用虚箭头表示。如在图 5.58 中，④⇢⑤ 表示工作 D 完成后，工作 H 才能开始。

4. 相邻两点之间只能有一条箭线连接，否则会造成逻辑上的混乱

图 5.59 是错误的画法，为了使两节点之间只有一条箭线，可增加一个节点 ②′，并增加一项虚工作 ②′⇢② 。图 5.60 是正确的画法。

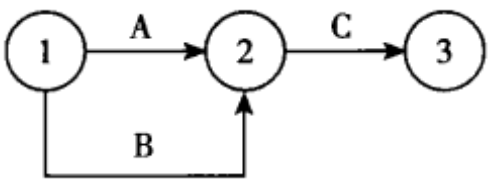


图 5.59

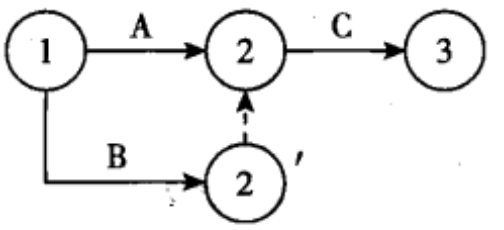


图 5.60

5. 网络计划图中不能有缺口和回路

在网络计划图中严禁出现从一个节点出发，顺箭线方向又回到原出发点，形成回路。回路将表示这工作永远不能完成。网络计划图中出现缺口，表示这些工作永远达不到终点，项目无法完成。

6. 平行工作

可与本工作同时进行的工作。

7. 起始节点和终点节点

在网络计划图中只能有一个起始点和一个终点节点。若工程开始或完成时存在几个平行工作，则可以用虚工作将他们与起始点或终结点连接起来。

8. 线路

网络图中从起点节点沿箭线方向顺序通过一系列箭线与节点，最后到达终点节点的通路。例 5.23 中有五条线路。并可以计算出各线路的持续时间，如表 5.43 所示。

表 5.43

线路	线路的组成	各工作的持续时间之和(天)
1	①→②→⑦→⑧	60 + 45 + 35 = 140
2	①→②→③→⑦→⑧	60 + 10 + 18 + 35 = 123
3	①→②→④→⑥→⑦→⑧	60 + 20 + 30 + 25 + 35 = 170
4	①→②→④→⑤→⑦→⑧	60 + 20 + 15 + 35 = 130
5	①→②→⑤→⑦→⑧	60 + 40 + 15 + 35 = 150

从网络图中可以计算出各线路的持续时间。其中有一条线路的持续时间最长，该线路是关键路线或称为主要矛盾线。关键线路上的各工作为关键工作。关键线路的持续时间决定了整个项目的工期。

9. 网络计划图的布局

尽可能将关键路线布置在网络计划图的中心位置，按工作的先后顺序将联系紧密的工作布置在近邻的位置。为了便于在网络计划图上标注时间等数据，箭线应是水平线或具有一段水平线的折线。在网络计划图上附有时间坐标或日历进程。

10. 网络计划图的类型

(1) 总网络计划图,以整个项目为计划对象,编制网络计划图。供决策领导层使用。

(2) 分级网络计划图,这是按不同管理层次的需要,编制的范围大小不同,详细程度不同的网络计划图,供不同管理部门使用。

(3) 局部网络计划图,将整个项目某部分为对象,编制更详细的网络计划图,供专业部门使用。

二、网络计划图的时间参数计算

网络计划的时间参数计算有几种类型:双代号网络计划有工作计算法和节点计算法;单代号网络计划有节点计算法。下面仅介绍工作计算法,其余的计算法可参考《工程网络计划技术规程》(JGJ/T121 - 99)。

网络计划图中的时间参数包括:工作持续时间(D),工作最早开始时间(ES),工作最早完成时间(EF),工作最迟开始时间(LS),工作最迟完成时间(LF),工作总时差(TF) 和工作自由时差(FF)。

(一) 工作持续时间 D

工作持续时间计算是一项基础工作,它关系到网络计划是否能得到正确实施。为了有效地使用网络计划技术,需要建立相应的数据库。这需要专项讨论。这里简述计算工作持续时间的两类数据和两种方法。

1. 单时估计法(定额法)

每项工作只估计或规定一个确定持续时间值的方法。一般有工作的工作量、劳动定额资料以及投入人力的多少等,计算各工作的持续时间:

$$D = \frac{Q}{P \times S \times n}$$

其中,Q 为工作的工作量,以时间单位表示,如小时,或以体积、质量、长度等代为表示;R 为可投入的人力和设备的数量;S 为每人或每台设备每工作班能完成的工作量;n 为每天正常工作班数。

2. 三时估计法

在不具备有关工作的持续时间的历史资料、在较难估计出工作持续时间时,可对工作进行三种时间值的估计,然后计算其平均值。这三种时间是:

乐观时间:在一切都顺利时,完成工作需要的最少时间,记为 a。

最可能时间:在正常条件下,完成工作所需要的时间,记为 m。

悲观时间:在不顺利条件下,完成工作所需要的最多时间,记为 b。

这三种时间发生都有可能,根据经验,这些时间的概率分布可认为是正态分布。一般情况下,通过专家估计法,给出三时估计的数据。可以认为工作进行时出现最顺利和最不顺利的情况比较少,正常的情况比较多,按平均意义可用下列公式计算平均工作持续时间:$D = \frac{a + 4m + b}{6}$,方差为 $\sigma^2 = (\frac{b - a}{6})^2$。

(二) 其他时间参数的计算

时间参数的关系可以用来表示各工作之间的关系状态,其计算步骤为:

(1) 计算各线路的持续时间(见表 5.43)。

(2) 按网络图的箭线的方向,从起始工作开始,计算各工作的 ES,EF。

(3) 从网络图的终点节点开始,按逆箭线的方向,推算出各工作的 LS,LF。

(4) 确定关键路线(CP)。

(5) 计算 TF,FF。

(6) 平衡资源。

1. 工作最早开始时间(ES) 和工作最早完成时间(EF) 的计算

利用网络计划图,从网络计划图的起始点开始,沿箭线方向依次逐步计算。第一项工作的最早开始时间为0,记作 $ES_{i-j}=0$(起始点 $i=1$)。第一项工作的最早完成时间 $EF_{1-j}=ES_{1-j}+D_{1-j}$。第一项工作完成后,其紧后工作才能开始。第一项工作的最早完成时间 EF 就是其紧后工作最早开始时间 ES。本工作的持续时间 D 表示为 $EF_{1-j}=ES_{i-j}+D_{i-j}$。

计算工作的 ES 时,在有多项紧前工作情况下,只能在这些紧前工作都完成后才能开始。因此本工作的最早开始时间是:$ES=\text{Max}$(紧前工作的 EF),其中 $EF=ES+$ 工作持续时间 D,即:

$$ES_{i-j}=\text{Max}_h(EF_{h-i})=\text{Max}_h(ES_{h-i}+D_{h-i})$$

例 5.23 的 ES 和 EF 的计算结果在表 5.44 的 ③ 和 ④ 列中。

利用双代号的特征,很容易在表中确定某工作的紧前工作和紧后工作。凡是后续工作的箭尾代号与某工作的箭头代号相同者,便是它的紧后工作;凡是先行工作的箭头代号与某工作的箭尾代号相同者,便是它的紧前工作。在表 5.44 中,首先填入 ①、②两列数据,然后由上往下计算 ES 与 EF。若某工作($i-j$) 的先行工作中存在几个($h-i$),从中选择最大的 EF_{h-i} 进行计算 $ES_{i-j}=\text{Max}_h(EF_{h-i})$,即计算 ES_{7-8}。如计算时,从表 5.44 的 ④ 列已有的 EF_{6-7},EF_{5-7},EF_{3-7} 中找到最大的 $EF_{6-7}=135$。将它填入表 5.44 的 ③ 列对应的 $L(7-8)$ 行即可。

表 5.44

工作($i-j$)	持续时间 D_{i-j}	最早开始时间 ES_{i-j}	最早完成时间 EF_{i-j}
①	②	③	④ = ③ + ②
$A(1-2)$	60	$ES_{1-2}=0$	$EF_{1-2}=ES_{1-2}+D_{1-2}=0+60=60$
$B(2-7)$	45	$ES_{2-7}=EF_{1-2}=60$	$EF_{2-7}=ES_{2-7}+D_{2-7}=60+45=105$
$C(2-3)$	10	$ES_{2-3}=EF_{1-2}=60$	$EF_{2-3}=ES_{2-3}+D_{2-3}=60+10=70$
$D(2-4)$	20	$ES_{2-4}=EF_{1-2}=60$	$EF_{2-4}=ES_{2-4}+D_{2-4}=60+20=80$
$E(2-5)$	40	$ES_{2-5}=EF_{1-2}=60$	$EF_{2-5}=ES_{2-5}+D_{2-5}=60+40=100$
$E'(4-5)$	0(虚工作)	$ES_{4-5}=EF_{2-4}=80$	$EF_{4-5}=ES_{4-5}+D_{4-5}=80+0=80$
$F(3-7)$	18	$ES_{3-7}=EF_{2-3}=70$	$EF_{3-7}=ES_{3-7}+D_{3-7}=70+18=88$
$G(4-6)$	30	$ES_{4-6}=EF_{2-4}=80$	$EF_{4-6}=ES_{4-6}+D_{4-6}=80+30=110$
$H(5-7)$	15	$ES_{5-7}=\text{Max}(EF_{2-5},EF_{4-5})=EF_{2-5}=100$	$EF_{5-7}=ES_{5-7}+D_{5-7}=100+15=115$
$K(6-7)$	25	$ES_{6-7}=EF_{4-6}=110$	$EF_{6-7}=ES_{6-7}+D_{6-7}=110+25=135$
$L(7-8)$	35	$ES_{7-8}=\text{Max}(EF_{2-7},EF_{3-7},EF_{6-7},EF_{5-7})=EF_{6-7}=135$	$EF_{7-8}=ES_{7-8}+D_{7-8}=135+35=170$

2. 工作最迟开始时间 LS 与工作最迟完成时间 LF

应从网络图的终点节点开始，采用逆序逐项计算。即按逆箭线方向，依次计算各工作的最迟完成时间 LS 与工作最迟完成时间 LF，直到第一项工作为止。网络图中最后一项工作的最迟完成时间应由工程的计划工期确定。在未给定时，可令其等于其最早完成时间，即 $LF_{i-n} = EF_{i-n}$；而 EF_{i-n} 由表 5.44 中的计算可知，它应当小于或等于计划工期规定的时间 T_r。

LF = Min（紧后工作的 LS），$LS = LF$ - 工作持续时间 D

其他工作的最迟开始时间 $LS_{i-j} = LF_{i-j} - D_{i-j}$；当有多个紧后工作时，最迟完成时间 LF = Min（紧后工作的 LS），即 $LF_{i-j} = \text{Min}_k(LF_{j-k} - D_{j-k}) = \text{Min}_k(LS_{j-k})$。计算结果在表 5.45 中，计算从下到上的进行，直到第一行为止。

表 5.45

工作 $(i-j)$	持续时间 D_{i-j}	最迟完成时间 $LF_{i-j} = \min_k(LS_{j-k})$	最迟开始时间 $LS_{i-j} = LF_{i-j} - D_{i-j}$	总时差 $TF_{i-j} = LS_{i-j} - ES_{i-j}$	自由时差 $FF_{i-j} = ES_{i-j} - EF_{i-j}$
①	②	⑤	⑥ = ⑤ - ②	⑦ = ⑥ - ③	⑧
$A(1-2)$	60	$LF_{1-2} = LS_{2-4} = 60$	60 - 60 = 0	0 - 0 = 0	60 - 60 = 0
$B(2-7)$	45	$LF_{2-7} = LS_{7-8} = 135$	135 - 450 = 90	90 - 60 = 30	70 - 70 = 0
$C(2-3)$	10	$LF_{2-3} = LS_{3-7} = 117$	117 - 10 = 107	107 - 60 = 47	70 - 70 = 0
$D(2-4)$	20	$LF_{2-4} = LS_{4-6} = 80$	80 - 20 = 60	60 - 60 = 0	80 - 80 = 0
$E(2-5)$	40	$LF_{2-5} = LS_{5-7} = 120$	120 - 40 = 80	80 - 60 = 20	100 - 100 = 0
$F(3-7)$	18	$LF_{3-7} = LS_{7-8} = 135$	135 - 18 = 117	117 - 70 = 47	135 - 88 = 47
$G(4-6)$	30	$LF_{4-6} = LS_{6-7} = 110$	110 - 30 = 80	80 - 80 = 0	110 - 110 = 0
$H(5-7)$	15	$LF_{5-7} = LS_{7-8} = 135$	135 - 115 = 20	120 - 100 = 20	135 - 115 = 20
$K(6-7)$	25	$LF_{6-7} = LS_{7-8} = 135$	135 - 25 = 110	110 - 110 = 0	135 - 135 = 0
$L(7-8)$	35	$LF_{7-8} = EF_{7-8} = 170$	170 - 135 = 35	135 - 135 = 0	170 - 170 = 0

3. 工作时差

工作时差是指工作有机动时间。常用有两种时差，即工作总时差和工作自由时差。

（1）工作总时差 TF_{i-j}。TF_{i-j} 是在不影响工期的前提下，工作所具有的机动时间，按工作计算法计算。计算公式为：$TF_{i-j} = EF_{i-j} - ES_{i-j} - D_{i-j} = LS_{i-j} - ES_{i-j}$ 或 $TF_{i-j} = LF_{i-j} - EF_{i-j}$，计算结果见表 5.45 中 ⑦ = ⑥ - ③ 的数据。

（2）工作自由时差 FF_{i-j}。FF_{i-j} 是在不影响其紧后工作最早开始的前提下，工作所具有的机动时间。计算公式为：$FF_{i-j} = ES_{j-k} - ES_{i-j} - D_{i-j}$ 或 $FF_{i-j} = ES_{j-k} - EF_{i-j}$，计算结果见表 5.45 中 ⑧ 的数据。

三、网络计划的优化

绘制网络计划图，计算时间参数和确定关键线路，仅能得到一个初始计划方案。然后根据项目委托方的要求和实际资源的配置，需要对初始方案进行调整和完善，即进行网络计划的优化，目标是综合考虑进度、合理利用资源和降低费用等。

(一) 工期优化

如果网络计划图的计算工期大于项目委托方要求的工期时,则必须根据要求加快计划的进度,缩短工程项目的完工工期。可主要采取以下措施,增加对关键工作的投入,以便缩短关键工作的持续时间,实现工期缩短。

(1) 采取技术措施,提高工效,缩短关键工作的持续时间,使关键线路的时间缩短。

(2) 采取组织措施,充分利用非关键工作的总时差,合理调配人力、物力和资金等资源。

(二) 资源优化

在编制初始网络计划图后,需要进一步考虑尽量利用现有资源的问题。即在项目工期不变的条件下,均衡地利用资源。实际工程项目包括的工作繁多,需要投入的资源种类繁多,均衡地利用资源是件很麻烦的事,需要用计算机来辅助完成。为了简化计算,具体操作如下:

(1) 优先安排关键工作所需要的资源。

(2) 利用非关键工作的总时差,错开各工作的开始时间,避开在同一时区内集中使用同一资源,以免出现资源负荷高峰期。

(3) 在确实受到资源限制或在考虑综合经济效益的条件下,在得到许可时也可以适当地推迟项目的工期,以达到错开资源负荷高峰的目的。

下面通过例 5.23 来说明平衡人力资源的方法。假设在例 5.23 中,现有机械加工工人 65 人,要完成工作 D、F、G、H、K,各工作需要工人人数列于表 5.46 中。

表 5.46

工作	持续时间(天)	需要工人(人数)	总时差(天)
D	20	58	0
F	18	22	47
G	30	42	0
H	15	39	20
K	25	26	0

由于机械加工工人人数的限制,如果上述工作都按最早开始时间安排,那么在完成关键工作的 75 天工期中,每天需要加工工人人数如图 5.61 所示。有 10 天需要 80 人,另 10 天需要 81 人。超过了现有机械工人人数的约束,必须进行调整。以虚线表示的非关键路线上非关键工作 F、H 有机动时间,若将工作 F 延迟 10 天开工,就可以解决第 70 ~ 80 天的超负荷问题;将工作 H 推迟 10 天开工,就可以解决第 100 ~ 110 天的超负荷问题。于是产生了新的负荷图,如图 5.62 所示,能够满足机械工人人数 65 人的约束。这里人力资源的平衡是利用了非关键工作的总时差,避开了资源负荷的高峰。

避开资源负荷高峰时,可以采用将非关键工作分段或采用技术措施减少所需要的资源,也可以根据计划规定适当延长项目的工期。

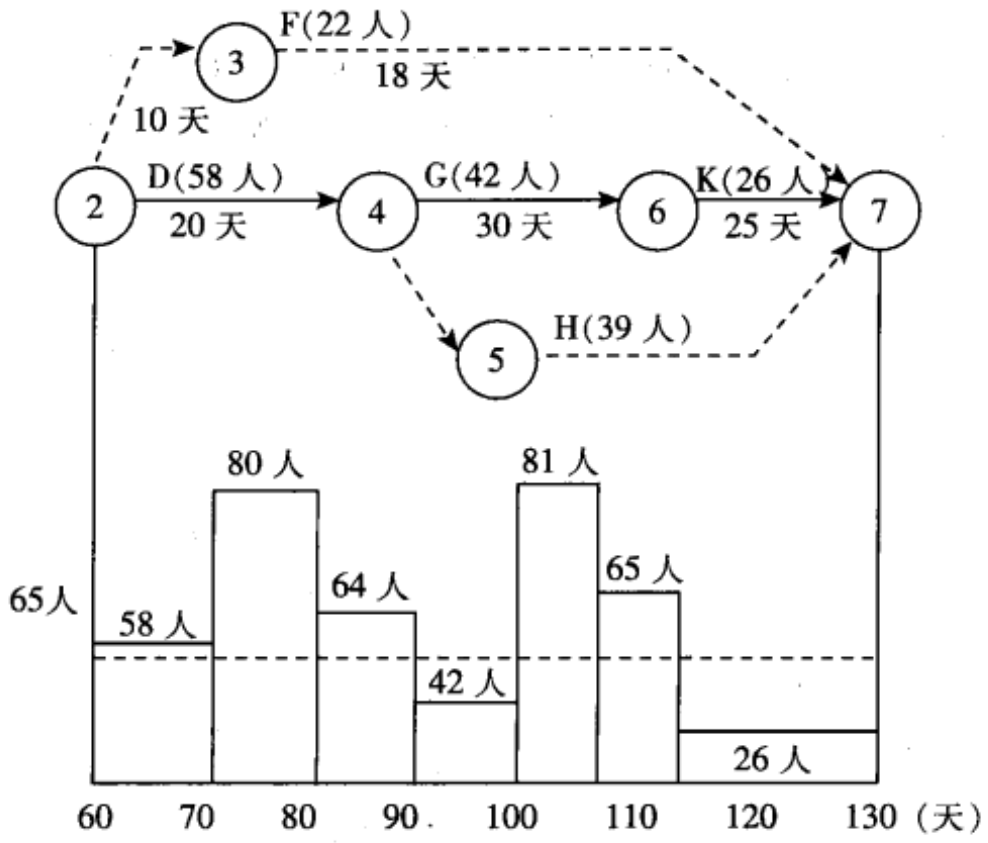


图 5.61

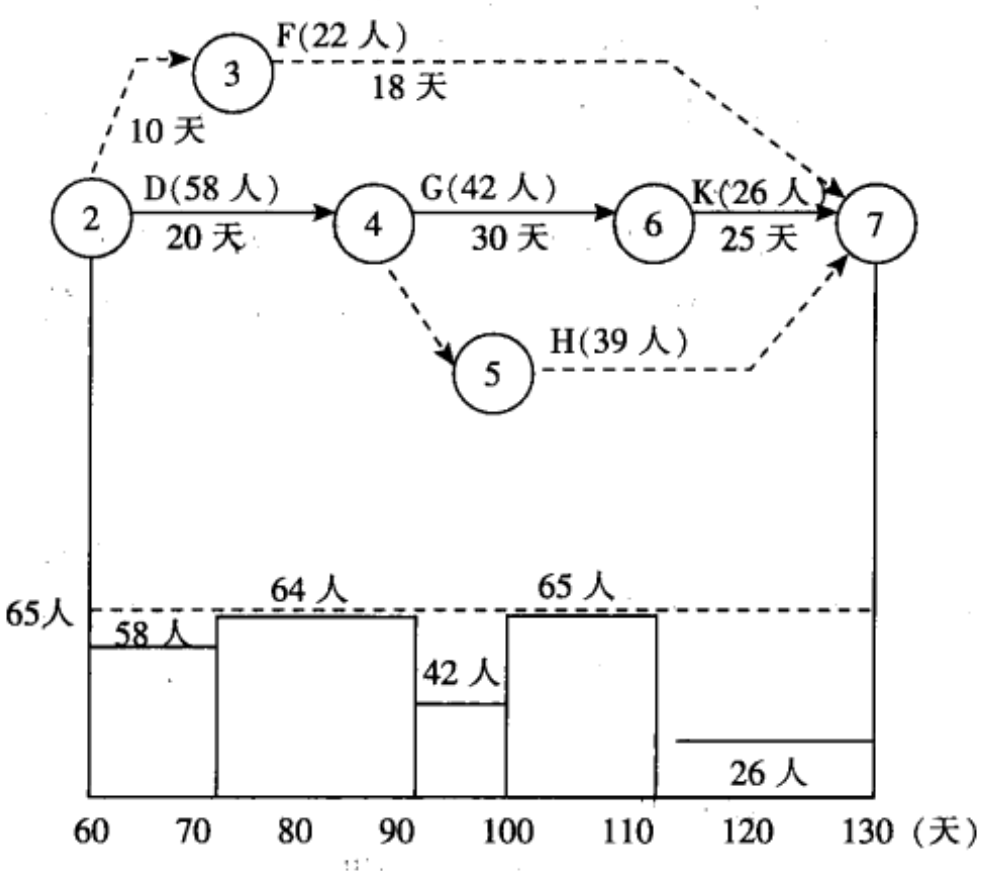


图 5.62

（三）时间—费用优化

编制网络计划时，要研究如何使完成项目的工期尽可能短，费用尽可能少；或在保证既定项目完成时间条件下，所需要的费用最少；或在费用限制的条件下，项目完工的时间最短。这就是时间—费用优化要解决的问题。完成一个项目的费用可以分为两大类：

1. 直接费用

直接与项目的规模有关的费用，包括材料费用和工人的工资等。为了缩短工作的持续时间和工期，就需要增加直接费用。

2. 间接费用

间接费用包括管理费等。一般按项目工期长度进行分摊，工期越短，分摊的间接费用就越少。一般项目的总费用与直接费用、间接费用、项目工期之间存在如图 5.63 的关系。

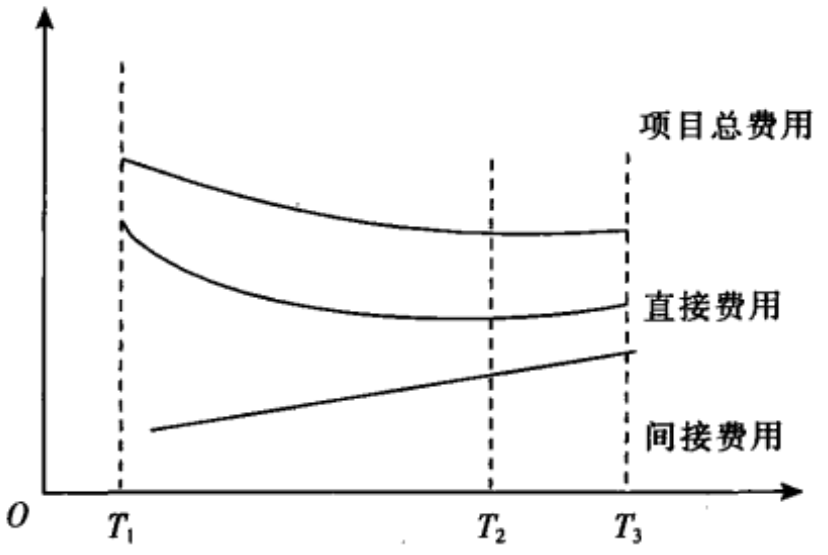


图 5.63　工期与总费用的关系曲线

图中：T_1 为最短工期，项目总费用最高；T_2 为最佳工期；T_3 为正常工期。

当总费用最少，工期短于要求工期时，这就是最佳工期。

进行时间 — 费用优化时，首先要计算出不同工期下最低直接费用率，然后考虑相应的间接费用。费用优化的步骤如下：

（1）计算工作费用率。工作费用率是指缩短工作持续时间每一单位时间（如一天）所需要的增加的费用。按工作的正常持续时间计算各关键工作的费用率，通常可以表示为：

$$\Delta C_{i-j} = \frac{CC_{i-j} - CN_{i-j}}{DN_{i-j} - DC_{i-j}}$$

其中：ΔC_{i-j} 表示工作 $i-j$ 的费用率；CC_{i-j} 表示将工作 $i-j$ 持续时间缩短为最短持续时间后，完成该工作所需要的直接费用；CN_{i-j} 是在正常条件下完成工作 $i-j$ 所需要的直接费用；DN_{i-j} 表示工作 $i-j$ 正常持续时间；DC_{i-j} 表示工作 $i-j$ 最短持续时间。

（2）在网络计划图中找出费用率最低的一项关键工作或一组关键工作作为缩短持续时间的对象。其缩短后的值不能小于最短持续时间，不能成为非关键工作。

（3）同时计算相应增加的总费用，然后考虑由于工期的缩短间接费用的变化，在这基础上计算项目的总费用。

重复以上步骤，直到获得满意的方案为止。

下面举一例来说明。已知项目每天的间接费用为400元，利用表5.47中的已知资料，安排进度，项目正常工期为 170 天，对应的项目直接费用为 68 900 元，间接费用为 $170 \times 400 = 68\,000$ 元，项目总费用为 136 900 元。这是在正常条件下进行的方案，称为 170 天方案。若要缩短这方案的工期，首先从缩短关键路线上直接费用率最小的工作的持续时间，在 170 天方案中关键工作 K 和 G 的直接费用率最低。从表中可见这两项工作的持续时间都只能缩短 10 天。由此总工期可以缩短到 $170 - 10 - 10 = 150$

天。按 150 天工期计算，这时总直接费用增加到 68 900 + (290 × 10 + 350 × 10) = 75 300 元。由于缩短工期，可以减少间接费用 400 × 20 = 8 000，工期为 150 天方案的总费用为 75 300 + 60 000 = 135 300 元。与工期 170 天方案相比，可以节省总费用 1 600 元。

但在 150 天方案中已有两条关键线路，即 ①→②→④→⑥→⑦→⑧ 或 ①→②→⑤→⑦→⑧。

如果再缩短工期，工作的直接费用将大幅度增加。例如在 150 天方案的基础上再缩短工期 10 天，成为 140 天方案。这时应选择工作 D，缩短 10 天；工作 H 缩短 5 天，工作 E 缩短 5 天。这时直接费用成为 75 300 + 400 × 10 + 400 × 5 + 500 × 5 = 83 800 元；间接费用为 140 × 400 = 56 000 元；总费用为 139 800 元。显然 140 天方案的总费用比 150 天方案和 170 天方案的总费用都高。综合考虑 150 天方案为最佳方案。计算结果见表 5.48。

表 5.49

序号	工作代号	正常持续时间(天)	工作直接费用(元)	最短工作时间(天)	工作直接费用(元)	费用率(元/天)
1	A	60	10 000	60	1 000	/
2	B	45	4 500	30	6 300	120
3	C	10	2 800	5	4 300	300
4	D	20	7 000	10	11 000	400
5	E	40	10 000	35	12 500	500
6	F	18	3 600	10	5 440	230
7	G	30	9 000	20	12 500	350
8	H	15	3 750	10	5 750	400
9	K	25	6 250	15	9 150	290
10	L	35	12 000	35	12 000	/

表 5.48

工期方案	170 天方案	150 天方案	140 天方案
缩短关键工作		K,G	D,H,E
缩短工作持续时间(天)		10,10	10,5,5
直接费用(元)	68 900	75 300	83 800
间接费用(元)	68 000	60 000	56 000
总费用(元)	136 000	135 300	139 800

四、网络计划的 WinQSB 软件应用

网络计划的 WinQSB 软件计算调用的程序模块是 PERT_CPM。该程序模块具有自动绘制节点网络图、三时估计法计算时间参数、计算网络参数、最低成本日程、项目完工时期与成本之间的模拟运算、时间坐标网络图和项目施工进度分析等功能。

下面,我们以例 5.22 为例,介绍网络计划的 WinQSB 软件计算。首先,调用子程序 PERT_CPM,建立新的文件,弹出对话框,如图 5.63 所示界面,输入文件标名“Plan1”(读者自己可以任意取名),工序数(Number of Activity) 7 和时间单位(Time Unit)week,然后在图 5.63 显示的选项中进行选择,当有三种估计时间时选择“Probabilistic PERT”,图 5.63 的右下方自动显示工序时间为三点分布,右上方数据不需要选择。本例工序时间为确定型,选择“Deterministic CPM”,在图的右上方选择正常时间(Normal Time)、应急时间(Crash Time)、正常成本(Normal Cost) 和应急成本(Crash Cost)。在图 5.63 左下方给出数据的输入格式有两种:表格形式(Spreadsheet) 和图形形式(Graphic Model),如果选择图形形式,系统就会显示一张 n × n(n 为工序数) 的方阵表格,在对应的方格中双击,系统自动按工序 AB 顺序显示,输入时间数据,画箭线将工序连接起来。本例我们选择表格输入形式,点击“OK”,此时弹出一张需要输入数据的表格,对照表 5.41 输入数据,得到图 5.64。

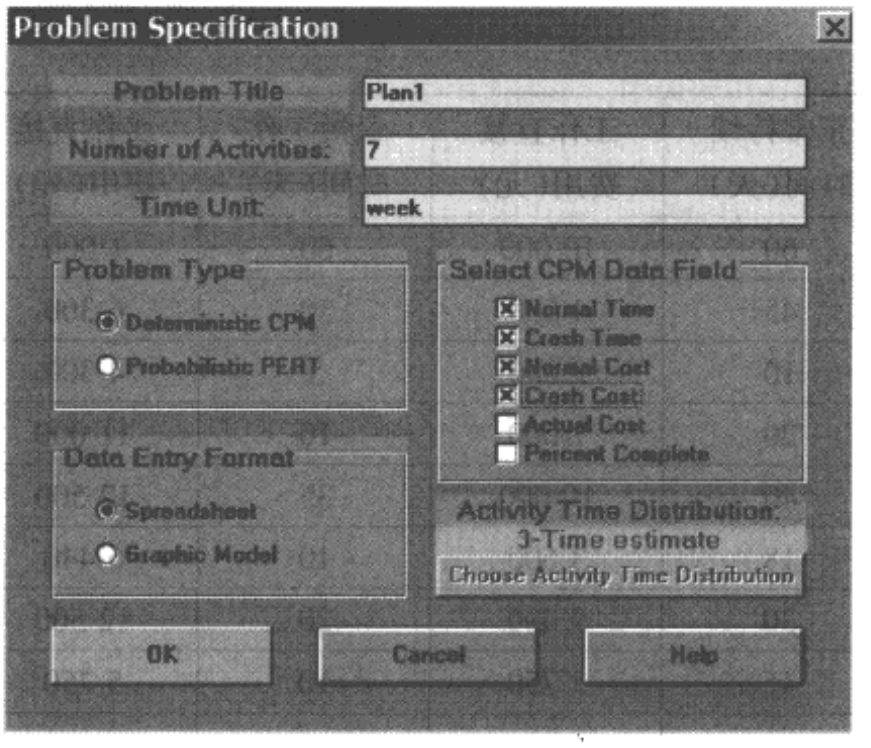


图 5.63

Activity Number	Activity Name	Immediate Predecessor (list number/name, separated by ',')	Normal Time	Crash Time	Normal Cost	Crash Cost
1	A		8	6	50	60
2	B		20	16	100	140
3	C	A	9	7	80	130
4	D	B	9	6	41	50
5	E	B,C	13	11	60	90
6	F	C,D	8	6	80	120
7	G	E,F	15	12	60	84

图 5.64

点击菜单栏“Solve and Analyze”,下拉菜单有三个选项选择求解方法:求解正常时间关键路径(Solve Critical Path Using Normal Time)、应急时间的关键路径(Solve Critical Path Using Crash Time) 和应急分析(Perform Crashing Analysis)。选择正常时间关键路径求解,显示图 5.65 所示的时间参数计算结果。

04-05-2006 17:20:23	Activity Name	On Critical Path	Activity Time	Earliest Start	Earliest Finish	Latest Start	Latest Finish	Slack (LS-ES)
1	A	no	8	0	8	7	15	7
2	B	Yes	20	0	20	0	20	0
3	C	no	9	8	17	15	24	7
4	D	Yes	9	20	29	20	29	0
5	E	no	13	20	33	24	37	4
6	F	Yes	8	29	37	29	37	0
7	G	Yes	15	37	52	37	52	0
	Project	Completion	Time	=	52	weeks		
	Total	Cost of	Project	=	$471	(Cost on	CP =	$281)
	Number of	Critical	Path(s)	=	1			

图 5.65

点击“Graphic Activity Analysis”,显示网络图 5.66,圆圈中的数据分别为工序的最早开始、最早结束、最迟开始和最迟结束时间。点击“Show Critical Path” 显示关键工序、关键路线和完工期,如图 5.67 所示。

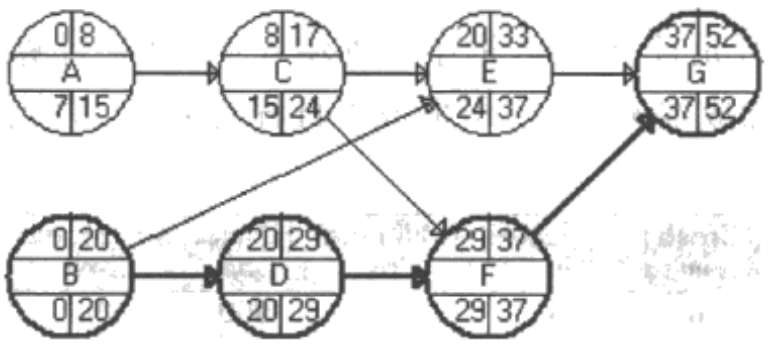


图 5.66

04-05-2006	Critical Path 1
1	B
2	D
3	F
4	G
Completion Time	52

图 5.67

点击菜单栏“Results → Perform Crashing Analysis”,显示图 5.68 的选项,这是网络计划优化的主要部分。图 5.68 右上方显示了正常时间和应急时间的完工期和相应的成本。在 Crashing Option 对话框中有三个选项,第一个选项为给定项目完工期,求总成本;第二个选项为给定预算成本,求项目工期;第三个选项是给定一个完工期,求最低成本日程。第一个和第二个选项是在只有应急成本而没有其他成本和收益时,模拟分析工期与成本之间的关系。图 5.68 的左下方的三个输入框的含义分别为:设定一个完工期 T,延迟一个单位时间(对 T 而言) 完工的损失,缩短一个单位时间(对 T 而言) 的收益。

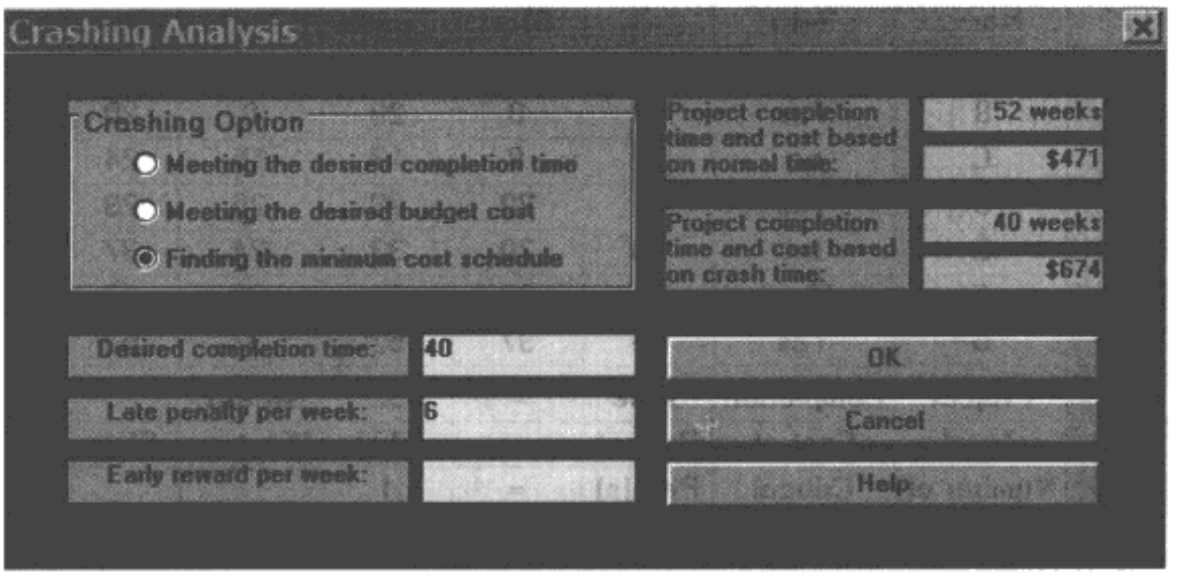


图 5.68

在本例中,选择“Finding the minimum cost schedule”,输入框中分别输入 40 和 6,注意不能在第一个输入框中输入 52,第三个框中输入 6。计算结果如图 5.69 所示。工序 D 缩短 3 周,增加成本 9 万元,最低成本日程为 49 周,总成本为 $471 + 9 - 3 \times 6 = 462$(万元)。

同理,提前一周完工增加收益 9 万元时,在输入框中分别输入 40 和 9,计算结果是工序 D 和工序 G 分别缩短 3 周,增加成本 33 万元,最低成本日程为 46 周,总成本为 $471 + 33 - 6 \times 9 = 450$(万元)。

04-05-2006 20:28:44	Activity Name	Critical Path	Normal Time	Crash Time	Suggested Time	Additional Cost	Normal Cost	Suggested Cost
1	A	no	8	6	8	0	$50	$50
2	B	Yes	20	16	20	0	$100	$100
3	C	no	9	7	9	0	$80	$80
4	D	Yes	9	6	6	$9	$41	$50
5	E	no	13	11	13	0	$60	$60
6	F	Yes	8	6	8	0	$80	$80
7	G	Yes	15	12	15	0	$60	$60
	Late	Penalty:						$54
	Overall	Project:			49	$9	$471	$534

图 5.69

点击图 5.68 选项框中“Meeting the desired budget cost”,输入 544,得到图 5.70 的计算结果。结果显示所有工序都是关键工序,其中工序 B、D 和 G 需要赶工,完工期为 42 周。

04-05-2006 20:37:59	Activity Name	Critical Path	Normal Time	Crash Time	Suggested Time	Additional Cost	Normal Cost	Suggested Cost
1	A	Yes	8	6	8.0000	¥ 0.00	$50	¥ 50.00
2	B	Yes	20	16	16	$40	$100	$140
3	C	Yes	9	7	9	0	$80	$80
4	D	Yes	9	6	6	$9	$41	$50
5	E	Yes	13	11	13	0	$60	$60
6	F	Yes	8	6	8.0000	¥ 0.00	$80	¥ 80.00
7	G	Yes	15	12	12	$24	$60	$84
	Overall	Project:			42.00	¥ 73.00	$471	¥ 544.00

图 5.70

按正常时间求解后，点击菜单栏“Results” → “PERT/Cost - Table”，得到图5.71，表中显示了从第1周到第52周最早开始、最迟开始每周成本和累计成本需求进度表。点击菜单栏“Results” → “PERT/Cost” - “Graphic”，将图5.71 绘制成图 5.72 所示的成本曲线图。点击“Results” → “*Project Completion Analysis*” 并输入20，显示图 5.73 的工序完成进度表。

当工序时间是随机变量时，在图 5.63 的选项中选择“PERT”，系统不仅能够计算时间参数，绘制网络图，还可以进行概率分析，对给定的完工时间求完工概率并进行模拟。读者可以自己尝试练习一下这种操作。

04-05-2006 20:44:06	Project Time in week	Cost Schedule Based on ES	Cost Schedule Based on LS	Total Cost Based on ES	Total Cost Based on LS
1	1	¥ 11.25	$5	¥ 11.25	$5
9	9	¥ 13.89	¥ 11.25	¥ 103.89	¥ 57.50
18	18	$5	¥ 13.89	¥ 220.00	¥ 166.67
21	21	¥ 9.17	¥ 13.44	¥ 239.17	¥ 207.89
30	30	¥ 14.62	¥ 14.62	¥ 327.15	¥ 308.69
34	34	$10	¥ 14.62	¥ 381.00	¥ 367.15
38	38	$4	$4	¥ 415.00	¥ 415.00
52	52	$4	$4	¥ 471.00	¥ 471.00

图 5.71

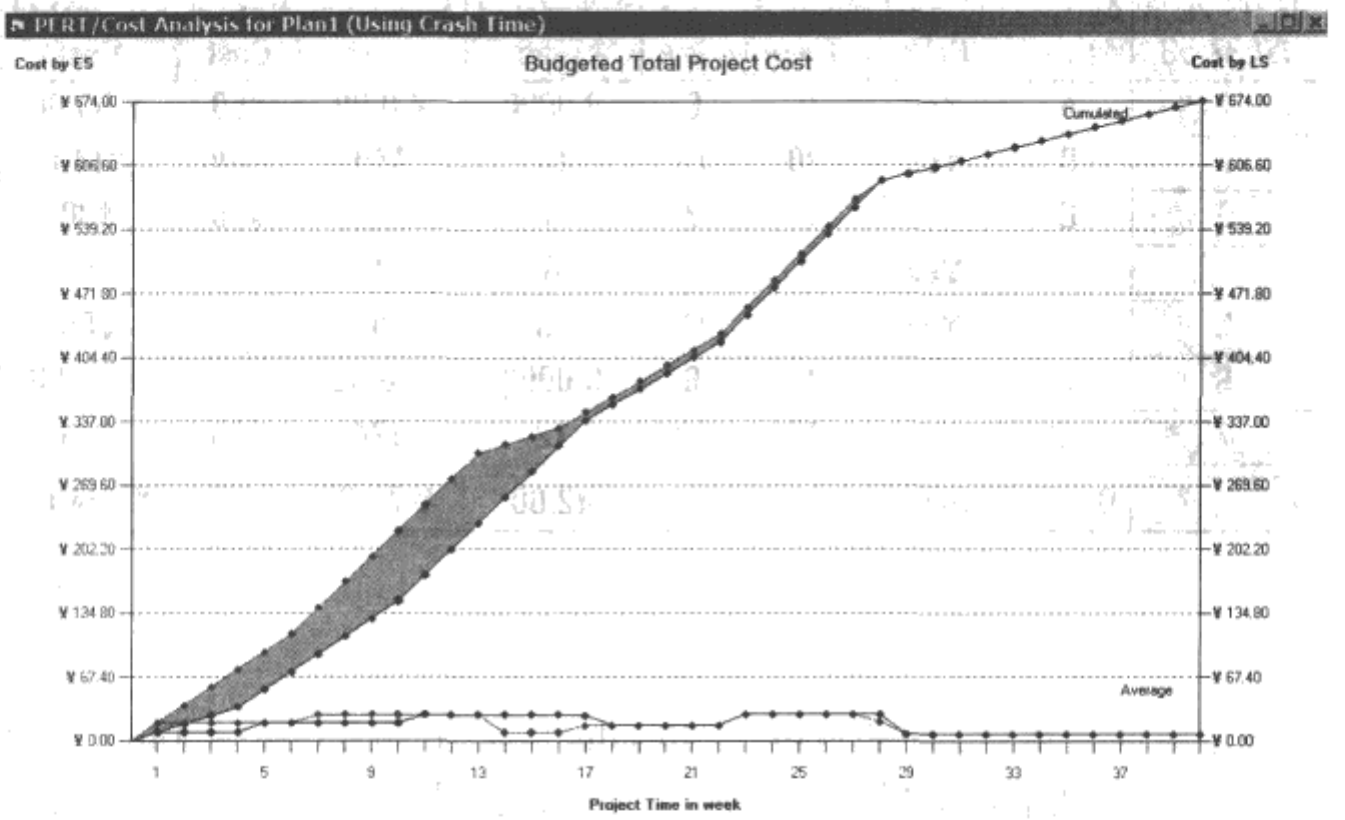


图 5.72

04-05-2006 21:12:12	Activity Name	On Critical Path	Activity Time	Latest Start	Latest Finish	Planned % Completion
1	A	no	8	7	15	100
2	B	Yes	20	0	20	100
3	C	no	9	15	24	55.5556
4	D	Yes	9	20	29	0
5	E	no	13	24	37	0
6	F	Yes	8	29	37	0
7	G	Yes	15	37	52	0
	Overall	Project:		0	52	38.4615

图 5.73

习题

1. 判断下列说法是否正确：

（1）图论中的图不仅反映了研究对象之间的关系，而且是真实图形的写照，因而对图中点与点的相对位置以及点与点连线的长短曲直等都要严格注意。

（2）在任一图 G 中，当点集 V 确定后，数图是 G 中边数最少的连通图。

（3）如果图中某点 v_i 有若干个相邻点，与其距离最远的相邻点为 v_j，则 $[v_i, v_j]$ 边必不包含在最小支撑树内。

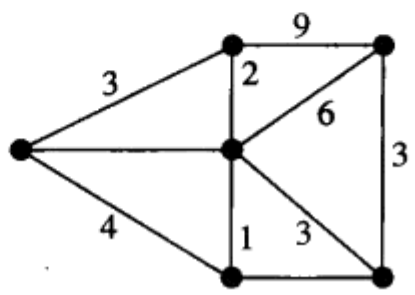
（4）如果图中从 v_1 至各点有唯一的最短路，则连接 v_1 至其他各点的最短路在去

掉重复部分后，恰好构成该图的最小支撑数。

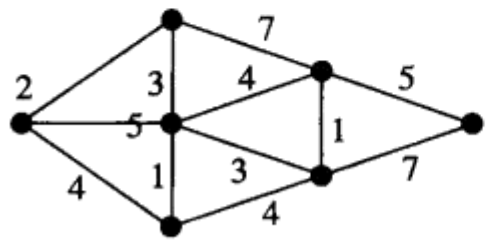
（5）求图的最小支撑数以及求图中一点至另一点的最短路问题，都可以归结为求解整数规划问题。

（6）求网络最大流的问题可归结为求解一个线性规划模型。

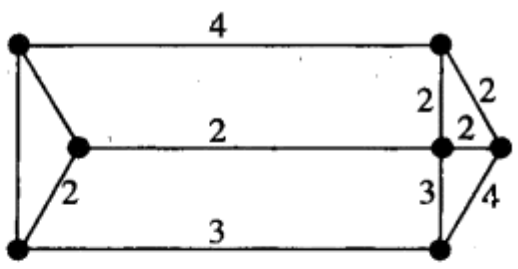
2. 分别用破圈法和避圈法求图 5.74 中各图的最小支撑树（最小部分树）。



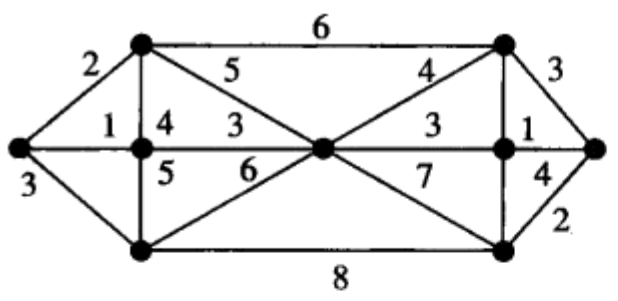
(1)



(2)



(3)



(4)

图 5.74

3. 有六个村庄，相互间道路的距离如图 5.75 所示。六个村拟合建一所小学。已知有小学生：*A* 村 50 人、*B* 村 40 人、*C* 村 60 人、*D* 村 20 人、*E* 村 70 人、*F* 村 90 人。问：小学建在哪个村庄，才能使所有小学生上学走的总路程最短？

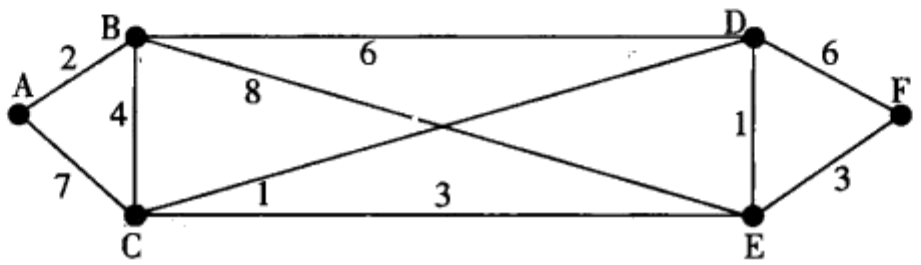


图 5.75

4. 已知有 8 口海上油井，相互间距离如表 5.49 所示。已知 1 号井离海岸最近，为 5 海里。从海岸经 1 号井铺设油管将各油井连接起来，应如何铺设使输油管长度为最短（为便于计量和检修，油管只准在各井位处分叉）？

表 5.49

从＼到	2	3	4	5	6	7	8
1	1.3	2.1	0.9	0.7	1.8	2.0	1.5
2		0.9	1.8	1.2	2.6	2.3	1.1
3			2.6	1.7	2.5	1.9	1.0
4				0.7	1.6	1.5	0.9
5					0.9	1.1	0.8
6						0.6	1.0
7							0.5

5. 天源纸盒厂生产 5 种规格的包装箱，每种包装箱的容积(cm^3)、需求量及生产一个包装箱的可变费用(元／件) 如表 5.50 所示，不同规格包装箱生产时需调整设备，其固定费用为 1 100 元。当某种包装箱不能满足需要时，可用规格大的替代。在满足需求的情况下，如何组织生产，才能使总费用为最小?并将此问题归结于求最短路的问题。

表 5.50

包装箱代号	1	2	3	4	5
容积／cm^3	15 000	25 000	40 000	60 000	90 000
需求量	1 000	1 200	2 600	1 800	1 800
可变费用(元／件)	3.5	4.2	5.0	6.0	6.5

6. 用标号法求图 5.76 所示的网络中从 v_s 到 v_t 的最大流量，图中弧旁数字为容量。

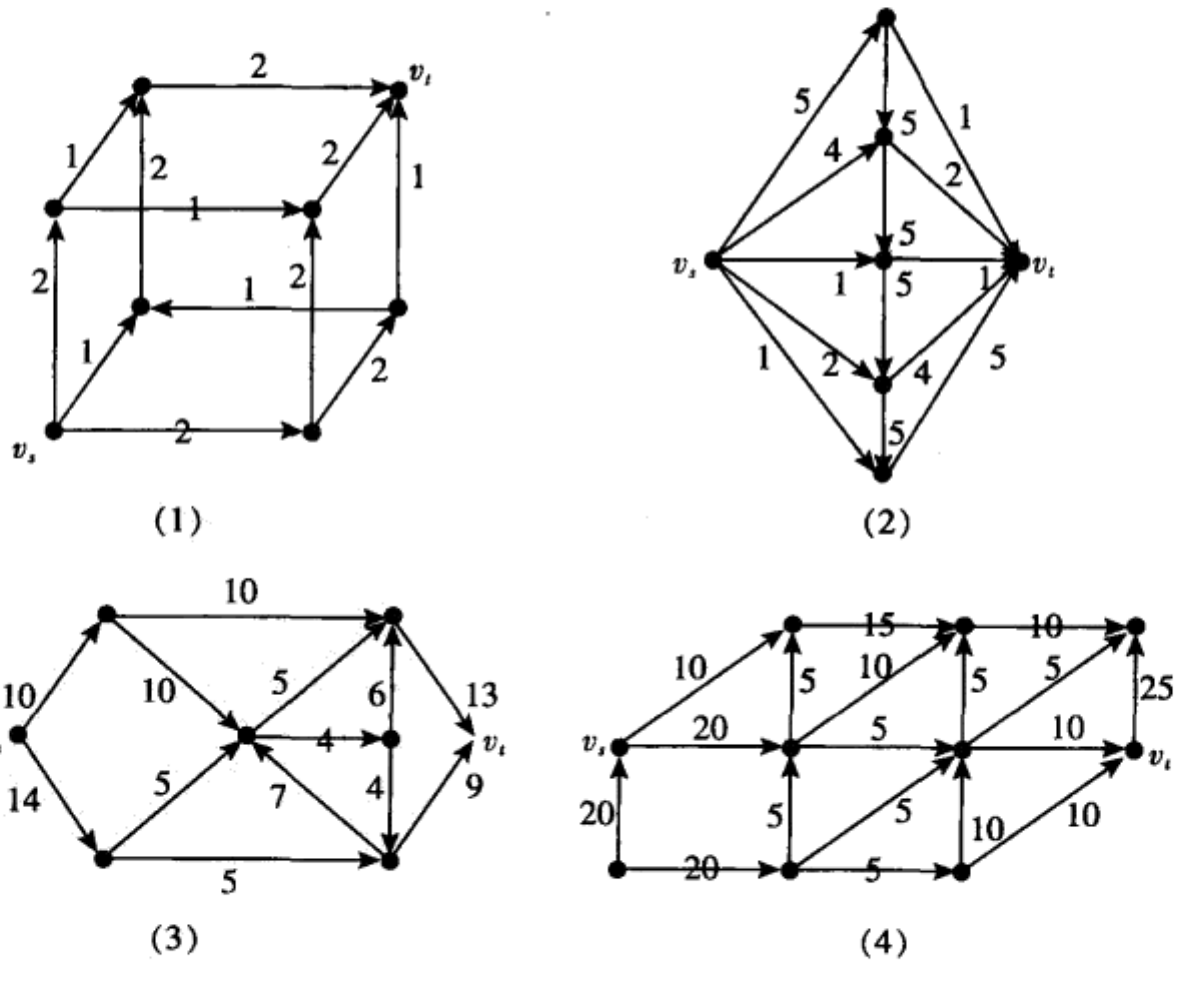


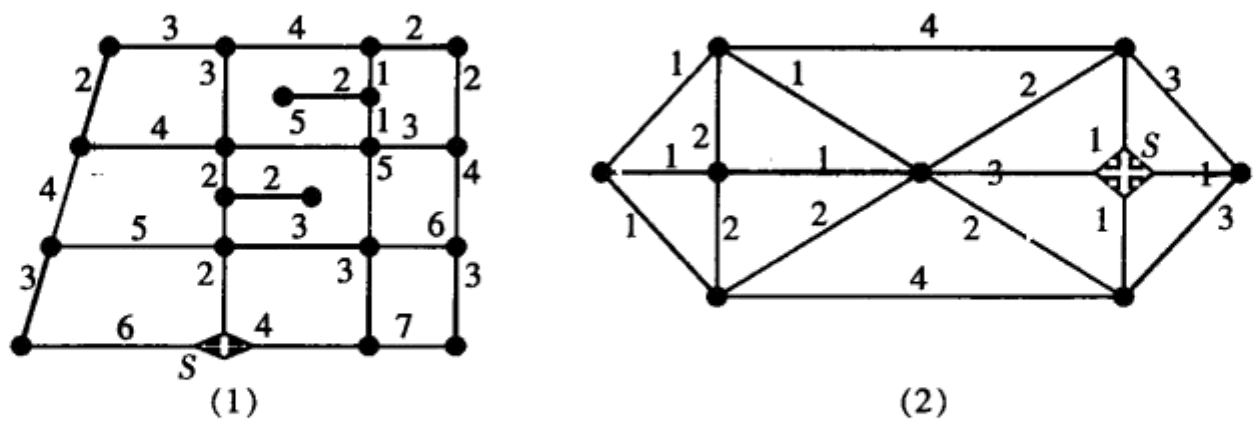
图 5.76

7. 一条流水线有5个岗位，要分别完成某产品装配的5道工序。现分配甲、乙、丙、丁、戊5个工人去操作。由于每人专长不同，各个工人在不同岗位上生产效率不一样，具体数字见表5.51（单位：件／分钟）。问到底怎样分配每个工人的操作岗位，才能使这条流水线的生产能力最大？

表5.51

工序 / 工人	Ⅰ	Ⅱ	Ⅲ	Ⅳ	Ⅴ
甲	2	3	4	1	7
乙	2	4	2	5	6
丙	2	5	3	4	1
丁	5	2	3	2	5
戊	3	7	6	2	4

8. 邮递员投递区域及街道分布如图5.77所示，图中数字为街道长度，S点为邮局所在地。试分别为邮递员设计一条最佳的投递路线。



(1)　　(2)

图5.77

9. 根据表5.52所示作业明细表，绘制网络图。

表5.52

工序	紧前工序	紧后工序
a	—	d
b	—	d,e
c	—	f
d	a,b	f
e	b	f
f	c,d,e	—

10. 已知建设一个汽车库及引道的作业明细表如表5.53所示。求：

（1）该项工程从施工开始到全部结束的最短周期；

（2）若工序l拖期10天，对整个工程进度有何影响；

（3）若工序j的工序时间由12天缩短到8天，对整个工程进度有何影响；

（4）为保证整个工程在最短周期内完成，工序i最迟必须在哪一天开工；

(5) 若要求整个工程在75天完工,要不要采取措施?如要,应从哪些方面采取措施?

表 5.53

工序代号	工序名称	工序时间(天)	紧前工序
a	清理场地,准备施工	10	—
b	备料	8	—
c	车库地面施工	6	a,b
d	预制墙及房顶的桁架	16	b
e	车库混凝土地面保养	24	c
f	立墙架	4	d,e
g	立房顶桁架	4	f
h	装窗及边墙	10	f
i	装门	4	f
j	装提天花板	12	g
k	油漆	16	h,i,j
l	引道混凝土施工	8	c
m	引道混凝土保养	24	l
n	清理场地,交工验收	4	k,m

第六章　库存决策

引例:毛巾订货量问题

“百花”小商店是一个专门经营各类毛巾的商店。每年营业360天,每天平均售出400条毛巾,每条毛巾的批发价为0.70元。每次订货的平均费用为112元,即每次订货,不论购买的数量多少都要支出112元。现在商店是每半年进一次货,一年进两次货。每条毛巾的存储费用一年为0.126元。这个商店的经理感觉到每年订货两次并非一个好的订货方法,他希望能找到一种方法能帮助他确定每年应该订货几次,每次的数量应该为多少,为他节约一笔总的库存费用。

第一节　基本概念

库存管理是现代化企业经营管理中的一个重要环节。必要的库存量是生产和销售得以顺利进行的保证。库存不足将使生产停工待料产生损失;商业销售中常因库存不足而产生脱销,给商店带来机会损失及信誉上的打击。因此,生产部门和销售部门希望有足够多的库存,以避免这种损失。

但是,过多的库存量将会占用大量的资金,占用的资金要支付利息,还需要支付保管费、耗损费等,这将使成本上升。一个大型的工厂,有的要支付与库存相关的费用高达几十万甚至上百万元。据统计,制造业的库存资金占用资金的25%是常有的事,一些管理不善的企业甚至达55%以上。国外许多企业更是因为库存管理不好而遭到失败或破产。从资金占用和支出角度出发,保管部门希望库存越少越好。

物资从订货到入库,每次要花费一定的费用。如差旅费、订货费、电话电报费、验收搬运费等,这些费用与订货次数有关。财务部门希望订货次数越少越好,每次订货的数量越多越好,以减少订货费用。

以上三方面的要求往往是互相矛盾的,科学的库存管理就是权衡这三者之间的关系,以使总的库存费用最少。科学的库存管理在于解决下面两个问题:①合理的库存量是多少? ②合理的订货时间是多长?

就我国的管理现状而言,库存管理尚未引起足够的重视,目前的主要的倾向是库存量过大,库存周期过长,物资积压和浪费现象严重。据统计,某年我国钢材库存量为1 980万吨,而同年德国为380万吨,英国660万吨,法国280万吨,意大利310万吨,我国的库存量比这4个国家的库存总量还多300多万吨。库存钢材占用流动资金达140亿元人民币左右,这与我国的财力和物力是很不相称的。

存储论(Inventory Theory)即是解决这一问题的科学方法论。早在1915年人们

就开始了对存储论的系统研究，它是运筹学最早成功应用的领域之一，是运筹学的重要分支。有效地利用库存模型对库存进行科学管理，将会带来较大的经济效益。

一、库存费用

一般而言，库存涉及以下三种费用：

1. 存储费用

所谓存储费用，是指为保持库存量而耗用的费用。存储费用一般包括：

(1)存储物资占用的流动资金及这些资金的利息；

(2)物资在存储中，因物品的腐蚀、变质、破损、遗失、被窃而产生的损失费用；某些物资因技术发展，新产品向世而降价或淘汰的损失费用；

(3)由于保管而产生的费用，如仓库的租金或税金、保险费、修理费、通风照明、保暖冷藏、水电、办公用品费及保管人员工资等费用。

存储费用是笔相当大的开支，一般为库存物资价值的 12% ~35%。存储费用随库存量的增加而增加。

2. 订货费用

所谓订货费用，是指每次物资从订货到入库所需的费用，统称为订货费用。如发出订单费、催办联络费、差旅费、运输费、进货检查费、调整费、验收费、搬运费等。这些费用一般是以每次订货来计算的。因此，订货费用随着订货次数的增加而增加。

3. 脱销费用

所谓脱销费用，是指因库存不足造成商品脱销或工厂生产停工待料而产生的直接或间接损失费用，统称为脱销费用。如商店因脱销而不能按期执行供货合同而受到罚款产生的直接损失费用，工厂停工待料使产量下降造成的直接损失费用等。同时，商店或工厂因此而受到信誉上的间接损失费用，甚至可能永远失掉一些“顾客”而造成未来业务上的间接损失费用，也称为脱销费用。这种间接损失费用往往难于直接计量而常被人忽视，许多工厂和商店不重视形象和信誉，以至于生意清淡或难以为继。

科学的库存管理就是控制、压缩和平衡这三类相互矛盾的费用，使总的库存费用最少。

二、需求形式

一般而言，需求的形式有如下三种情况：

(1)间断的还是连续的。比如羽绒服、空调、冷饮等这类有季节性的产品其需求是间断的，而日常用品的需求是连续的。

(2)均匀的还是不均匀的。如车间自动生产流水线对原料的需求是均匀的，而一个家庭对电的需求则不是均匀的。

(3)确定性的还是随机的。如生产活动对原材料的需求一般是确定的，而销售活动中对商品的需求量通常具有不确定性，但我们可以通过大量观察试验，掌握其统计规律性。正是根据需求量的确定与否，我们的存储模型分为确定性存储模型与随机性存储模型两大类。

三、供给形式

在一个库存(存储)系统中,有两个基本问题需要我们做出决策——何时订货和订货数量,分别称作"期"和"量"的问题。关于"何时订货"和"订货数量"的决策,就称为存储系统的供给形式,也称为存储策略。管理者就是通过控制这两个变量,来调节存储系统的运行,以达到最优的存储安排。如果将供给形式看成系统的输入,将需求形式看成系统的输出,则在库存系统中,输出是人们不能控制的,人们只能通过控制和调整输入来进行科学的库存管理。而在排队系统(第七章)中,输入是人们不能控制的,人们只能通过控制和调整输出来进行排队系统的科学管理。

不同的存储系统有不同的存储策略,即存储系统的存储策略多种多样,本书仅介绍一些常用的、较简单的存储策略。

(1)t_0 循环策略,即每隔 t_0 时间补充存货,补充量为 Q。这种策略一般适用于需求比较明确,一次进货及时补充的存储系统。

(2)(s,S)策略,即当存储量为 s 时立即补充,订货量为 $Q=S-s$,表明把库存量补充到 S。这种策略一般适用于需求比较明确,允许缺货的存储系统。

(3)(t,s,S)策略,即每隔 t 时间检查库存,当库存量小于 s 时,立即补充库存量到 S;而当库存量大于 s 时,可以暂时不补充。这种策略一般适用于需求不确定的存储系统。

四、补充形式

为了满足不断的需求,存储系统需要经常补充存储的物资品。补充的途径有内部生产和外部订购两种方式。补充的形式有一次补充(瞬时补充)和多次补充等。在一般情况下,由于供货方生产和运输需要一段时间,即从订货到交货之间是有一段滞后的时间,这一时滞可以是某一常量,也可以是随机的。总之,为了在某一时刻获得补充,就必须提前一段时间订货,这段时间称为订货提前期。

第二节　确定型库存模型

一、无折扣的经济订货量模型

在本章引例"毛巾订货量问题"中,我们假定毛巾的需求量不变,即每天都是售出400条毛巾(以后我们将对这种假设提出疑问,但这种假设对于许多均衡生产的工厂是基本合理的)。库存量的补充是一次及时补充,即第一次库存刚好用完时,第二次订货又一次进入库房。这样就得到一个简单而理想的库存图(见图6.1)。

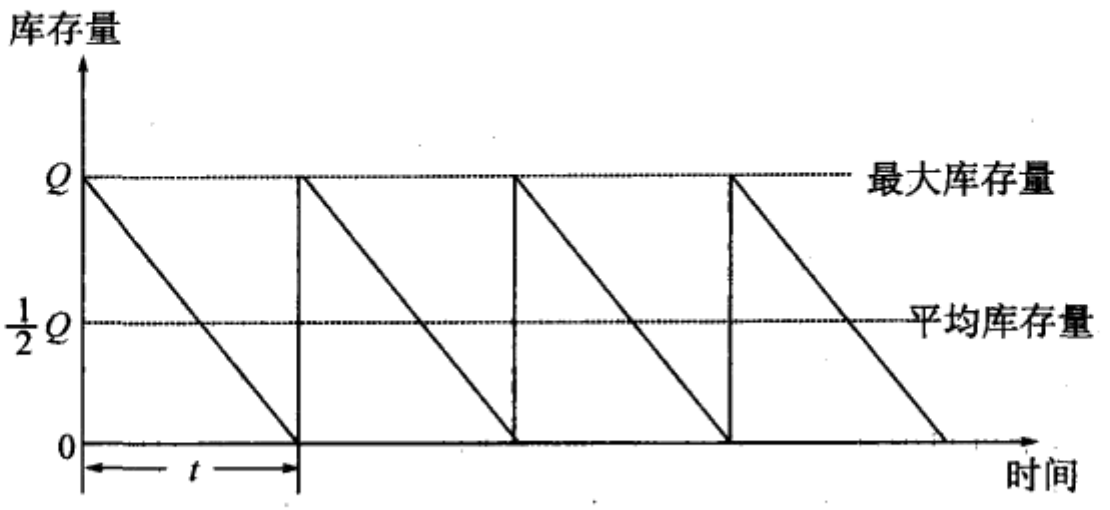


图 6.1

另外,我们假定毛巾的批发价不论购买数量多少,都是每条 0.7 元,这时称为无折扣的订货价格(有时,批发部门为了鼓励订货,对一次购买达到一定数量以上的实行优惠价格,这时称为有折扣的订货价格)。

现在"百花"商店是每年进货两次,每年毛巾的需求量是 $H = 400 \times 360 = 144\ 000$条,则每次订货数量为 144 000/2 = 72 000 条。

这个库存问题是等量需求及时补充的,因此不会产生脱销费用。这时的年度总库存费用 = 年订货费用 + 年存储费用,用公式表示为:

$$A = B + F$$

其中:A 为年总库存费用;F 是年存储费用;B 为年订货费用。

$$B = S \times H/Q$$

式中:S 为每次订货费用;H 为年需求量;Q 为每次订货量。

本例中,$S = 112$ 元,$h = 144\ 000$ 条,$Q = 72\ 000$ 条,则:

$B = S \times H/Q = 112 \times 144\ 000 / 72\ 000 = 224$(元)。而 H/Q 为每年订货次数($N = H/Q$),则年订货费用也可这样计算,即 $B = NS = 2 \times 112 = 224$(元)。

F 为年存储费用,$F = Q/2 \times K$。式中,K 为单位商品的存储费用,$Q/2$ 为平均库存量。本例 $K = 0.126$ 元,则 $F = 72\ 000/2 \times 0.126 = 4\ 536$(元)。

因此"百花"商店每年订货两次,每次订货量为 72 000 张时的总库存费用为 $A = B + F = 224 + 4\ 536 = 4\ 760$(元)。

问题在于是否还有其他更好的订货次数和订货批量能使总的库存费用更省呢?让我们先用列举法来看能否找到一个更好的库存策略。我们考虑年订货次数为 2、4、6、9、12、18 时总库存费用的情况是怎样的呢?(见表 6.1)

表 6.1

年需求量(张)	每次订货量(张)	年订货次数(次)	每次订货费用(元)	单位年存储费用(元)	年订货费用(元)	年存贮费用(元)	年总库存费用(元)
H	Q	$N = H/Q$	S	K	$B = NS$	$F = QK/2$	$A = B + F$
144 000	72 000	2	112	0.126	224	4 536	4 760
144 000	36 000	4	112	0.126	448	2 268	2 716

表 6.1(续)

年需求量（张）	每次订货量（张）	年订货次数（次）	每次订货费用（元）	单位年存储费用（元）	年订货费用（元）	年存贮费用（元）	年总库存费用（元）
H	Q	$N=H/Q$	S	K	$B=NS$	$F=QK/2$	$A=B+F$
144 000	24 000	6	112	0.126	672	1 512	2 184
144 000	16 000	9	112	0.126	1 008	1 008	2 016
144 000	12 000	12	112	0.126	1 344	756	2 100
144 000	8 000	18	112	0.126	2 016	504	2 520

从表 6.1 中我们可以清楚地看到，年订货费用随着订货次数的增加而增加，而年存储费用随着每次订货量的增加而增加。而当年订货费用等于年存储费用时，年总的库存费用最低，这时为 2 016 元，对应的每次订货批量为 16 000 条，年订货次数为 9 次。这就是“百花”商店应该采用的最好的订货策略，较之每年订货 2 次，每次订货量为 72 000 条的年度总库存费用4 760元，节约了 2 744 元 (4 760 - 2 016)。

“百花”商店采用每年订货 9 次，每次订货量为 16 000 条毛巾时，意味着每次订货能供应(16 000/400 = 40 天)。如果每次从订货到入库的时间为 15 天，则每次进货后的第 26 天是下一次开始订货的时间。

从上面的列举法，可以得到订货量与库存费用的关系图(见图 6.2)。

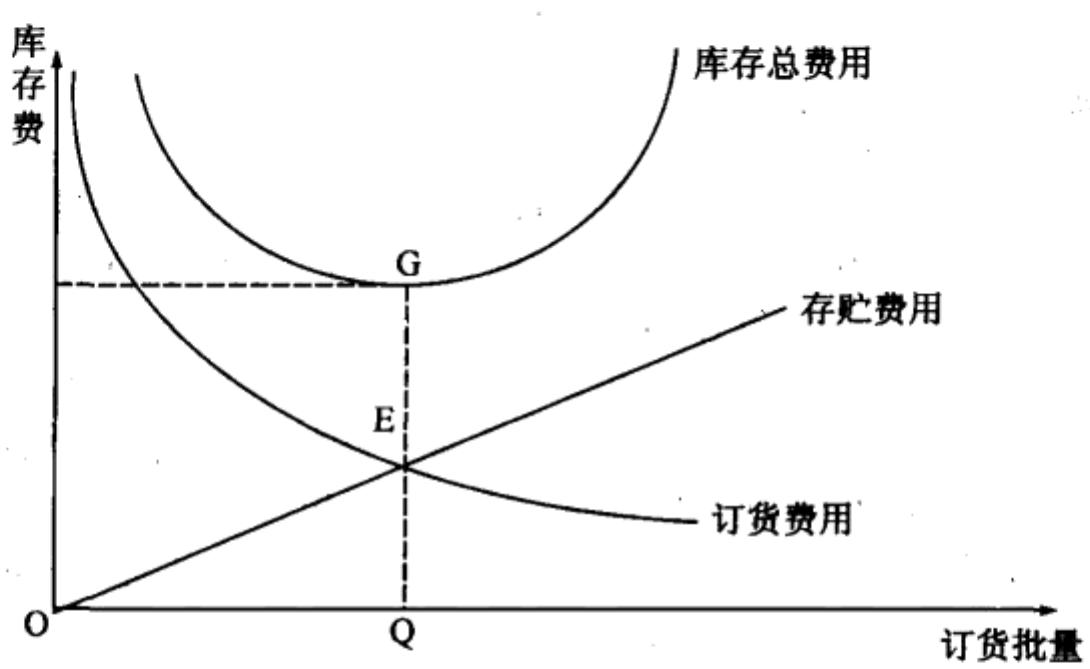


图 6.2 订货批量与库存费用的关系

从图 6.2 中可以看出，当存储费用等于订货费用时，即图中表现为两曲线的交点 E 时，总库存费用最低。图中 E 点对应于总库存费用曲线的最低点 G，这时 G 对应在横轴上的交点 Q 就是最佳的每次订货批量(EOQ)。

实际工作中，当然不会使用列举的笨办法。因年存储费用等于年订货费用的总库存费用最低，我们利用这一等式可以推出经济订货量 Q，即：

当 $B = F$ 时 A 最小，即

$$\frac{HS}{Q}=\frac{KQ}{2}$$

$$Q^2=\frac{2HS}{K}$$

则可推得

$$Q^*=\sqrt{\frac{2HS}{K}}$$

即只要知道了全年需求量 H,每次订货费用 S,单位商品的存储费用 K,就可直接利用经济订货批量公式计算,本例为:

$$\text{订货批量 } Q^*=\sqrt{\frac{2HS}{K}}=\sqrt{\frac{2\times144\,000\times112}{0.126}}=16\,000(\text{条})$$

$$\text{订货次数 } N=\frac{H}{Q^*}=\frac{144\,000}{1\,600}=9(\text{次})$$

即每年订货 9 次,每次订货量为16 000条毛巾时,总库存费用最少。

$$\text{存储费用 } F=\frac{1}{2}Q^*K=\sqrt{\frac{HSK}{2}}=\frac{1}{2}\times16\,000\times0.126$$

$$=\sqrt{\frac{144\,000\times112\times0.126}{2}}=1\,008(\text{元})$$

$$\text{订货费用 } B=\frac{H}{Q^*}\times S=\sqrt{\frac{HSK}{2}}=\frac{144\,000}{1\,600}\times112=1\,008(\text{元})$$

总库存费用 $A=B+F=1\,008+1\,008=2\,016$(元)

二、有折扣的经济订货量模型

“百花”商店的采购人员在向批发部门进货时得知,每次若购买的数量低于18 000条时,每条毛巾的单价为 0.7 元;若每次购买的数量在 18 000 至 70 000 条时,每条毛巾的单价给予 10% 的折扣,即每条毛巾的单价为 $0.7\times0.9=0.63$ 元;若每次购买的数量在 70 000 条及以上时,给予 12% 的折扣,即每条毛巾的单价为 $0.7\times0.88=0.616$ 元。商店负责人想知这在这种有价格折扣的情况下,他是否能够利用这一价格折扣的好处;在有两种以上的价格折扣时,他应该利用哪一种价格折扣更合适。

在很多时候,商品的供应方为了多推销产品,规定每次购买达到一定数量时,给以价格上的折扣优惠,以鼓励买方多购买。这时,购买方面临着有价格折扣的订货决策问题。利用价格折扣的好处,在于可获得较低的单价,并且由于每次订货的数量大,可以减少订货的次数而降低年订货费用。有时因大批购货还可得到运输上的好处,如铁路上的整车发运就比零单发运的价格低。但另一方面,批量过大会使存储费用增加,积压资金。由于有价格折扣这一因素的加入,存储费用与购买物品的费用直接有关,这时决策者是从总费用(包括购买物品的费用)最小的标准来选择订货批量的。

现在让我们来帮助“百花”商店在有两种价格折扣的情况下来进行订货批量的选择。我们已知每条毛巾的年存储费用为 0.126 元 。在有价格折扣的条件下,年度的单位存储费用显然就与单位商品的价格直接有关。因此我们需要将单位存储费用

分为两部分，使 $K=PI$，其中 P 为订货单价，I 为单位存货价格的百分比。我们知 $K=0.126$ 元是在单价为 0.7 元时的单位存储费用，$I=K/P=0.126/0.7=0.18$，表示单位存储费用是单位价值的 18%（实际工作中，单位存储费用的计算较为繁杂，故常用单位存货价值的百分比来表示，且一般是一个不变的比例）。在有价格折扣的条件时，则可计算出不同的单位存储费用，例如：

当 $P_1=0.7$ 元时，$K_1=P_1I_1=0.7\times0.18=0.126$（元）；

当 $P_2=0.63$ 元时，$K_2=P_2I_2=0.63\times0.18=0.113\ 4$（元）；

当 $P_3=0.616$ 元时，$K_3=P_3I_3=0.616\times0.18=0.110\ 88$（元）。

这时经济订货批量的公式可写成：

$$Q^*=\sqrt{\frac{2HS}{PI}}$$

当订货单价不同时，经济订货批量也不一样 。

当 $P_1=0.7$ 元时，

$$Q_1^*=\sqrt{\frac{2HS}{P_1I}}=\sqrt{\frac{2\times144\ 000\times112}{0.7\times0.18}}=16\ 000\text{（条）}$$

当 $P_2=0.63$ 元时，

$$Q_2^*=\sqrt{\frac{2HS}{P_2I}}=\sqrt{\frac{2\times144\ 000\times112}{0.63\times0.18}}=16\ 865\text{（条）}$$

当 $P3=0.616$ 元时，

$$Q_3^*=\sqrt{\frac{2HS}{P_3I}}=\sqrt{\frac{2\times144\ 000\times112}{0.616\times0.18}}=17\ 056\text{（条）}$$

从上面的计算可以看出，当订货单价为 0.63 元时，“百花”商店应该每批订货 16 865条，这时总的库存费用最低，但是订 16 865 条不能取得折扣 10% 的好处。因供应方规定取得折扣好处的最低订货量是 18 000 条。为了取得折扣，“百花”商店必须一次至少订购 18 000 条，但这是否合算呢？当单价为 0.616 时，取得总库存费用最低的订货数量应为 17 056 条，但这个订购数量也不能取得折扣 12% 的好处。要能得到 12% 的折扣，商店必须一次至少进货 70 000 条，但这对商店是否有利呢？

要对这几种订货批量作出评价。这时需要加入年购货费用因素的总费用（A）来加以衡量：总费用 = 年订货费用 + 年存储费用 + 年购货费用。若用 D 表示年购货费用，则 $D=PH$，$A=B+F+D$。

当订货单价为 0.7 元时，订货批量为 16 000 条时，

年订货费用为：$B_1=HS/Q_1^*=144\ 000\times112/16\ 000=1\ 008$（元）

年存储费用为：$F_1=P_1Q_11/2=(0.7\times16\ 000\times0.18)/2=1\ 008$（元）

年购货费用为：$D_1=P_1H=0.7\times144\ 000=100\ 800$（元）

则总费用：$A_1=B_1+F_1+D_1=1\ 008+1\ 008+100\ 800=102\ 816$（元）

当订货单价为 0.63 元时，经济订货批量 $Q^*=16\ 865$ 条，但这不能取得价格折扣 10% 的优惠，迫使要将一次货批量定为 18 000 条。这 18 000 条是带限制性的最优量，计算总费用时应将 $Q^*=18\ 000$ 代入进行计算，这时的年订货费用不等于年存储费用，这时：

年订货费用为：$B_2 = HS/Q_2 = 144\ 000 \times 112/18\ 000 = 896$(元)

年存储费用为：$F_2 = P_2 Q_2 I/2 = (0.63 \times 18\ 000 \times 0.18)/2 = 1\ 021$(元)

年购货费用为：$D_2 = P_2 H = 0.63 \times 144\ 000 = 90\ 720$(元)

则总费用：$A_2 = B_2 + F_2 + D_2 = 892 + 1\ 021 + 90\ 720 = 92\ 637$(元)

当订货单价为0.616元时，经济订货批量为 $Q^* = 17\ 056$ 条，同样不能取得价格折扣12%的优惠，商店一次订货量至少为70 000条才能取得价格优惠，计算总费用时应将 $Q^* = 70\ 000$ 代入进行计算，这时：

年订货费用为：$B_3 = HS/Q_3 = 144\ 000 \times 112/70\ 000 \approx 230$(元)

年存储费用为：$F_3 = P_3 Q_3 I/2 = (0.616 \times 70\ 000 \times 0.18)/2 \approx 3\ 881$(元)

年购货费用为：$D_3 = P_3 H = 0.616 \times 144\ 000 = 88\ 704$(元)

则总费用：$A_3 = B_3 + F_3 + D_3 = 230 + 3\ 881 + 88\ 704 = 92\ 815$(元)

可见，在有三种订货单价下，以每次订货18 000条毛巾的总费用92 637元为最低，商店负责人应选择每年订货8(144 000/18 000)次，这时能获得价格折扣10%的好处，且总费用为最低。

三、边补充边消耗的库存模型

“百花”商店的经理最近得到供应方的通知，由于对方运输和生产上的原因，近几年准备将原来的一次订货量一次送达的供货方式，改为每天送60打一包的毛巾来，即每天送来720条，保证商店的正常销售。经理考虑到由于毛巾是陆续分批到达，商店实际上是处于一边进货一边售货的状况。在这种情况下，是否应该改变商店的订货策略以适应这种变化，而使总的库存费用最低呢？

有的时候，订货由于生产或者运输方面的原因，不能一次到齐，而是分批送到，库存就是陆续补充陆续消耗的，而在货物一次到达时，库存是一次补充陆续消耗的。就企业来说，大多数情况都是边补充边消耗的，如加工车间向装配车间提供零部件，就是陆续送到装配车间的；就装配车间来讲，库存就是边补充边消耗的。边补充边等量消耗的库存情况如图6.3所示。

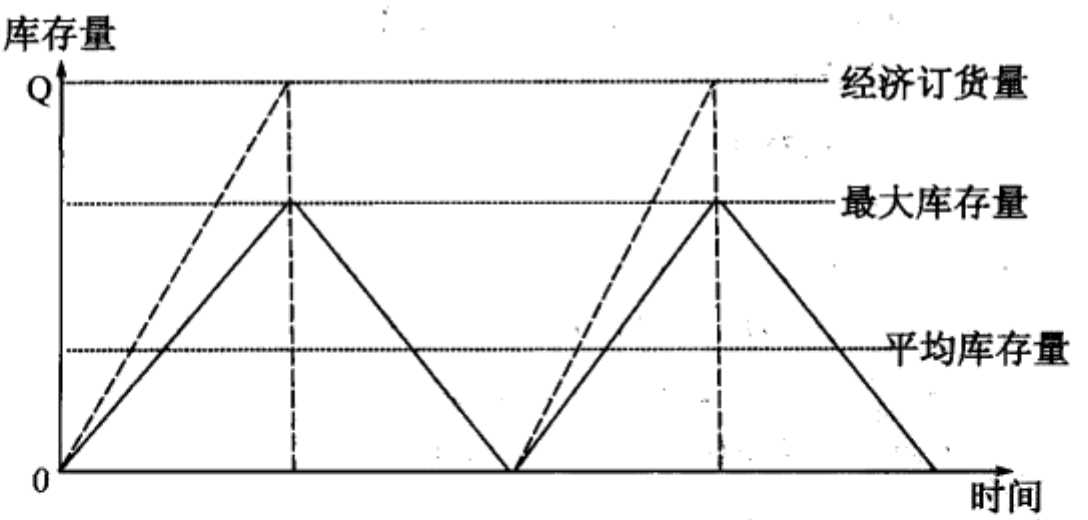


图6.3

从图中可以看出，在边补充边消耗的情况下，最大的库存量低于经济订货批量。而在等量消耗，一次补充的库存图中，最大库存量等于订购批量(如图6.1所示)。

在这种情况下,我们将会找到新的订货策略。

在"百花"商店面临供货方式改变的情况下,让我们来考虑新的订货策略。令商店每天的销售量为 U,则 $U=400$ 条,令每天进货量为 R,则 $R=720$ 条。商店在获得价格折扣 10% 的最优订货是 18 000 条,每年订货 8 次,订货单价为 $P=0.63$ 元,单位存储费用为单位价格的 18%,每次订货费用还是 $S=112$ 元。由于是分批进货,销售与进货同时进行,因此最大库存量小于订货批量。这时的最大库存量是多少呢?

假定在分批进货的订货批量仍是 18 000 条,现在是每天进货 $R=720$ 条。则原来一次进的货现在要 18 000/720 =25(Q/R,进货周期的天数)天才能进货完毕。当 25 天进完 18 000 条毛巾时,由于商店每天售出 400 条,进货完时商店已经售出了 $25\times400=10\ 000$ 条(QU/R,进货周期内的销售量)毛巾,这时仓库的最大库存量为 18 000 - 10 000 =8 000 条$[Q-(QU/R)=Q(1-U/R)]$。因为平均库存量为最大库存量的一半 ,则平均库存量为 4 000 条。则得到分批进货时的平均库存量:$Q^*(1-U/R)/2$

年存储费用为:$F=PIQ^*(1-U/R)/2$

年订货费用为:$B=HS/Q^*$

即有 $HS/Q^*=PIQ^*(1-U/R)/2$

解之得:

$$Q^*=\sqrt{\frac{2HS}{PI(1-U/R)}}$$

在分批进货的条件下,"百花"商店的每次订货量应为:

$$Q^*=\sqrt{\frac{2HS}{PI(1-U/R)}}=\sqrt{\frac{2\times144\ 000\times112}{0.63\times0.18\times(1-400/720)}}=25\ 298$$

但这时的订货次数 $N=H/Q$ =144 000/25 298 =5.7(次)不是整数,用"四舍五入法"得到每年订货次数为 6 次,得到修正后的订货批量 $Q=144\ 000/6=24\ 000$ 条,这时的年订货费用为:

$B=HS/Q^*=144\ 000\times112/24\ 000=672$(元)

年存储费用为:$F=PIQ(1-U/R)/2=0.63\times0.18\times24\ 000(1-400/720)/2=605$(元)

年总库存费用为:$A=B+F=672+605=1\ 277$(元)

边补充边消耗时,较之每年订货 8 次,订货量为 18 000 条,全部订货一次入库的年总库存费用要少 640 元(1 277 - 1 917)。

四、允许缺货的经济订购批量模型

在上面几种情况的库存模型中,都是在假设不允许缺货的条件下来进行讨论的。而现实情况是,由于在一个订货周期(T)内的需求量增加或者是供货方延期到货,这时就会发生缺货(脱销)。但如果企业发生缺货并不因此遭受损失或损失较小,顾客会耐心等到新的补充到来,即称允许缺货。这在一些供不应求的行业中可能出现这种情况,这时企业的存储策略将发生变化。一般情况下,企业在存储量降为零后,不会立即补充,而是等待一段时间后再集中订货,这样既可以减少存储费用又可以减少

订购费用。

可见,允许缺货的经济订购批量模型除了允许缺货而外,其他假设均与经济批量模型一致。其库存图见图 6.4。

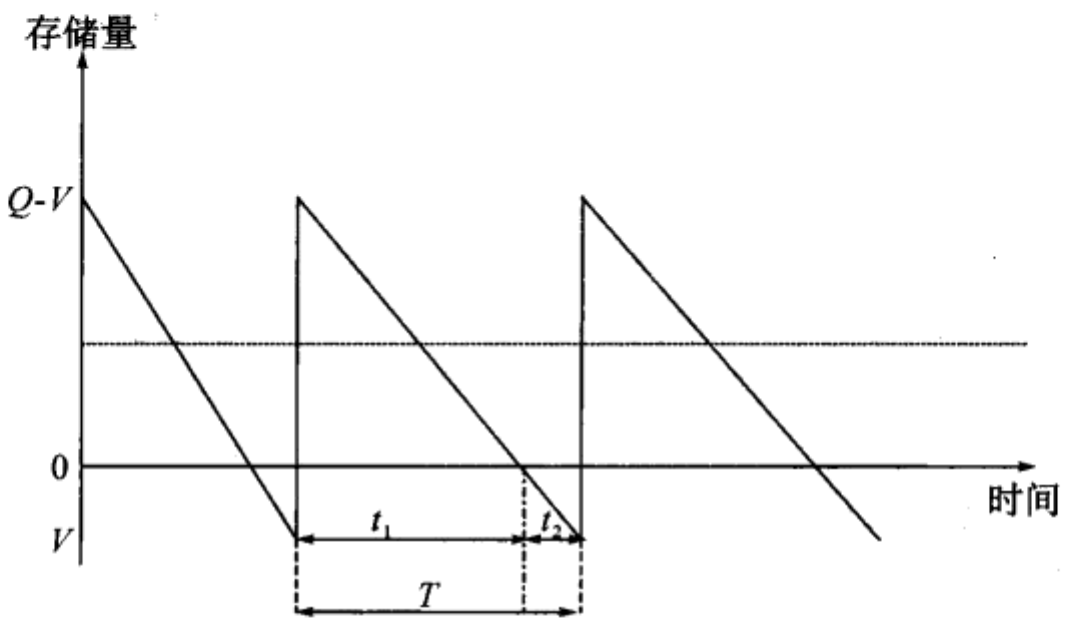


图 6.4

由图 6.4 可看出,在一个订货周期时间(两次订货的时间间隔)T 内,t_1 时间段没有缺货,t_2 时间段是缺货时间,最大缺货量达到 V,因此每次订货量 Q 是最高存储量与最大缺货量 V 之和,则最高存储量为$(Q-V)$。因此,在不缺货时期平均存储量为最高存储量的一半为$(Q-V)/2$,而在缺货时期内存储量为零,这时的年存储费用为 $K(Q-V)/2$。年订货费用为$(H/Q)S$,因年存储费用等于年订货费用时的总库存费用最小,则有:

$$K(Q-V)/2 = (H/Q)S$$

解之得:

$$Q(Q-V) = 2HS/k$$

其中,V 是最大缺货量。若令 C 为单位缺货的费用,可证明 $V=KQ/(K+C)$(证明略),将 $V=KQ/(K+C)$ 代入上式,可解得:

$$Q^* = \sqrt{\frac{2HS(K+C)}{KC}}$$

可见,在允许缺货的经济订购批量的公式中,需要考虑单位缺货的费用 C。仍由引例中的毛巾订货量为例来进行分析。

已知 $H=144\ 000$ 条、$S=112$ 元、$K=0.126$ 元。若令单位缺货的费用 $C=0.2$ 元。则有:

$$Q^* = \sqrt{\frac{2HS(K+C)}{KC}} = \sqrt{\frac{2\times 144\ 000\times 112(0.126+0.2)}{0.126\times 0.2}} = 20\ 427(\text{条})$$

为了对比分析费用起见,在允许缺货的经济订购批量为 20 427 条时的订货次数不是整数(实际工作中订货次数一般应修正为整数),这里不加以修正。经计算,最大缺货量 $V=7\ 895$ 条,年存储费用 $F=K(Q-V)/2=789.5$ 元,年订货费用 $B=(H/Q)S=789.5$ 元,年总库存费用 $A=F+C=789.5+789.5=1\ 579$ 元,比不允许缺货的年总库存费用 2 016 元少。当然,当单位缺货费用 C 较高时,允许缺货的总费用有可

能超过不允许缺货下的总费用，在这种情况下，由于缺货的成本较高，管理者就会避免出现缺货现象，问题也就演变成了不允许缺货条件下的经济订购批量决策。

五、经济订货模型的应用——投产批量决策

火星机械厂成批生产一种类型的家用电器，年产量为 6 000 台。这种类型的家用电器有几种不同的规格和型号，每换一种规格和型号的生产时，需要对设备作一些小的调整，清洗和开始生产的质量控制试验及生产和材料的准备等，这需要花一笔费用，每次为 375 元。每个电器未出厂前在仓库的年存储费用为 2 元。负责生产的厂长希望能确定每批产品的数量和年调整次数，以使每批产品调整的总费用最少 。

经济批量的确定在于使总的库存费用最少。这个模式对于一些分批或间断性生产的工厂如何确定合理的投产批量也是十分有用的。如许多机械厂生产一定数量的某一产品后转而生产另一种产品，这是分批生产的例子；一个化工厂，生产一定数量的美容油脂后，转而生产一种工业用油脂，这是间断生产的例子。为了使一定批量调整的总费用最少，这相当于库存问题中使总库存费用最少。转产时，如设备改装费、调整费、生产和材料准备费等统称为调整费用，相当于订货费用。生产批量则相当于订货批量。要使年度的调整费用降低，一则需减少调整次数（这相当于年订货次数）但调整次数降低，则要加大生产批量，生产批量加大会增加产品在仓库的存储费用。平衡这两种费用的办法就是选择一个经济的投产批量，而使其调整的总费用最少。因此经济订货批量的公式在解决生产批量问题时也是完全适用的。

在家用电器的投产批量中，年产量 $H=6\ 000$ 台，每次调整费用 $S=375$ 元，单位存储费用 $k=2$ 元，则经济批量为：

$$Q^{*}=\sqrt{\frac{2HS}{KC}}=\sqrt{\frac{2\times 6\ 000\times 375}{2}}=1\ 500$$

因此，火星机械厂一种规格和型号的投产批量应为 1 500 台，每年的调整次数为 4 次。这时的年调整费用 $B=NS=4\times 375=1\ 500$ 元，年存储费用 $F=KQ/2=2\times 1\ 500/2=1\ 500$ 元，年总费用为 $A=B+F=1\ 500+1\ 500=3\ 000$（元）。

第三节　不确定型库存模型

一、多周期不确定型库存模型

（一）有安全库存量的订货模型

“火炬”商店经营一种汽车配件，商店一年营业 360 天，平均每天售出 10 件。这种配件是从附近一家汽车配件厂进的货，每次订货费用为 126 元，每件的年存储费用为 2.5 元。这家商店用经济批量模型计算出每次的最优订货量为 600 件，订货后 25 天汽车配件厂一次交货入库。因此商店每当库存量降到 250 件时就开订下一批订货，这样当库存量刚好售完时则另一批货正好运到。

但实际的销售情况并非如此理想，销售 10 件产品是一个平均数，并不是每天都

不多不少正好销出 10 件,而是时多时少,在 10 件左右浮动。有时,在下一批订货还未到时,商店里已经售完了库存的最后一个配件,产生脱销;有时,则是因为汽车配件厂的货没有按时到达而造成脱销。该商店经理估计每脱销一件配件将损失 5 元的利润,但更使经理忧虑的是,当商店的汽车配件脱销后,顾客将到其他的商店购买,并且有可能永远失去这些顾客。他认为这是比直接利润损失更严重的事情。由于这种汽车配件是一种耐用品,在一个订货周期内未售完的货品可以存储到下一个周期甚至下几个周期进行销售。

因此,当由于需求量或交货期不确定,货品能多周期使用时的库存问题称为多周期不确定型库存问题,也称为多周期随机存储问题。

由于需求量或交货期的不确定就可能产生脱销,解决脱销的办法是提前订货或者是增加每批的订货量,这样将得到一个保险的安全库存量,以减少脱销的机会,这将降低或消除由脱销而损失的费用。但提前订货或者增加订货批量将会增加库存而使存储费用增加,我们希望能找到一个能平衡脱销费用和存储费用的方法。

在"百花"商店的毛巾订货问题中,我们就提到对每天平均销售 400 条毛巾提出过疑问,这 400 条毛巾的销量也是一个平均数,而实际的销售情况是变化的,有时减少有时增加,这种情况也将使商店有面临脱销危险而造成脱销损失,同样也有可能由于进货不及时而造成脱销损失,这两种情况的脱销如图 6.5 所示。

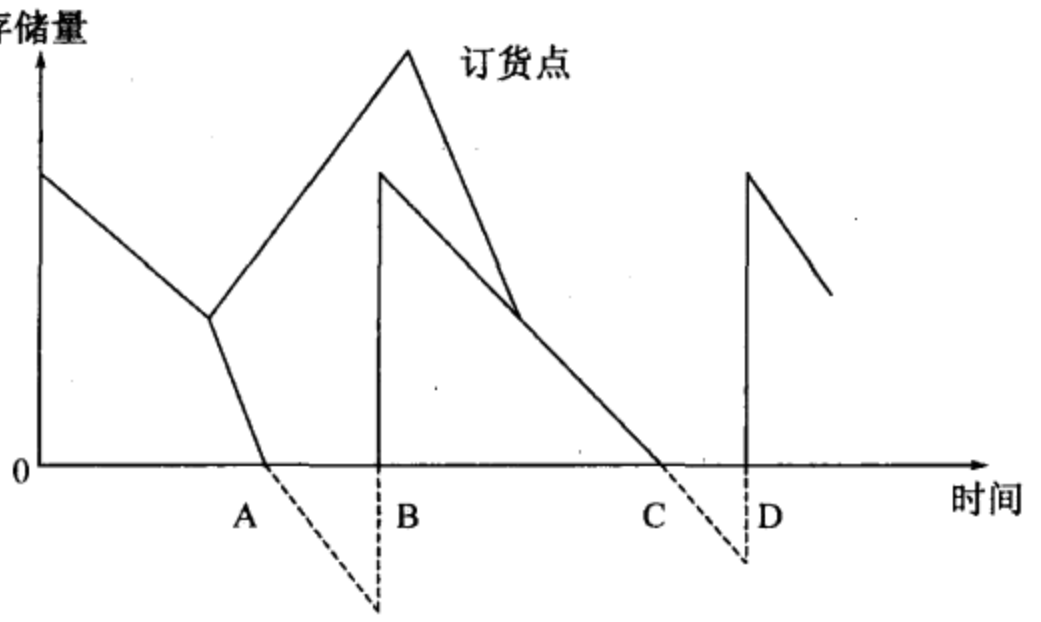


图 6.5

减少脱销费用的一个有效办法是厂商保有一个安全库存量。所谓安全库存量就是一种能防止因意外情况售完正常库存商品而保持的库存量。假定厂商有一个数量为 M 的安全库存量,这时的最大库存量为 $Q+M$,因此,在拥有安全库存量的情况下,当由于再订货后需求量增加而将发生脱销时,或者由于发货方延期交货将发生脱销时,则可动用安全库存量来防止脱销情况的发生(如图 6.6 所示)。

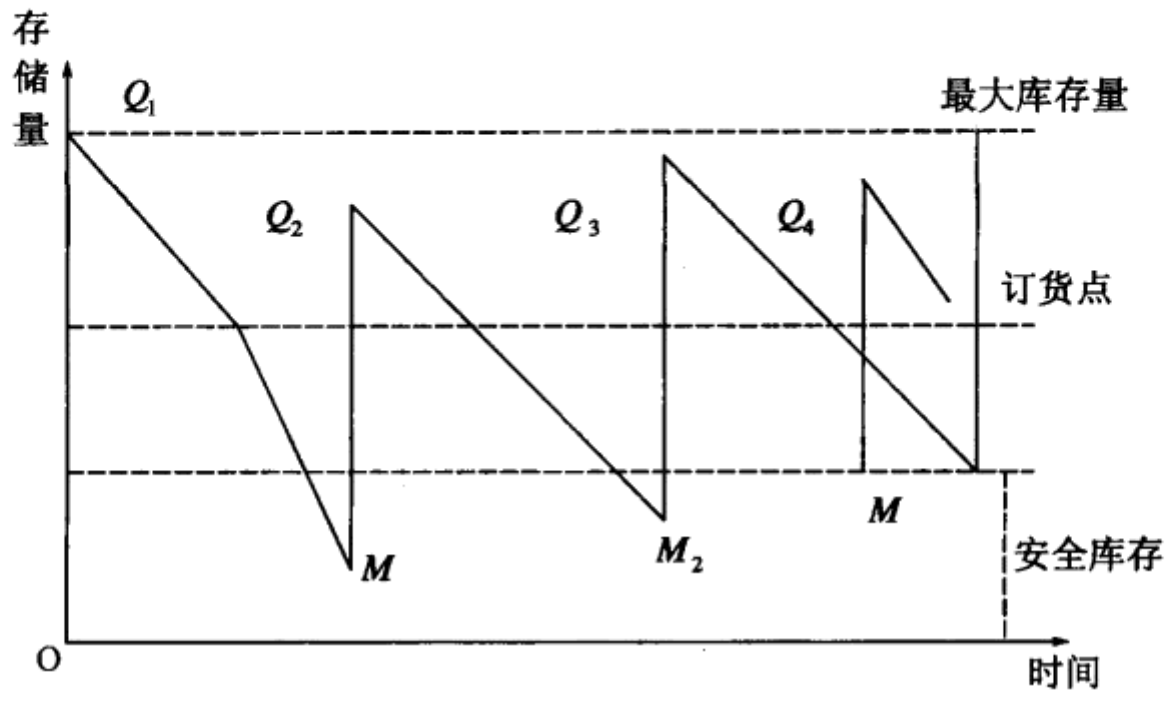


图 6.6

但显然我们不能把安全库存量 M 定得太多来防止全部的脱销情况。因为过多的安全库存量将增加年存储费用,我们需要确定一个合理的安全库存量以使存储费用和脱销货用的总和最小。而这一数值决定于存储费用、脱销费用和发生脱销的概率。

我们现在来为“火炬”商店确定一个合理的安全库存量 M^*。其步骤是:首先确定在一个订货周期内不同需求量下的脱销概率并计算各种安全库存量下的年度期望脱销费用,其次计算出存储费用,最后比较各种安全库存量下的总费用。总费用最少的安全库存量则为最优的(合理的)安全库存量。

一个订货周期是指从订货到货物入库间的时间。“火炬”商店订货周期为 25 天,经济订货批量 $Q^* = 600$ 件,即当库存为 250 件时,应开始进行下一次的订货活动。在这个时期内,若需求量超过 250 件则会发生脱销。商店过去 50 个订货周期内需求量及发生的次数如下(见表 6.2),据此可得到需求量超过库存量时的脱销概率。

表 6.2

订货周期需求量(件)	周期数(次)	概率(P)
150	5	0.1
190	5	0.1
220	10	0.2
250	15	0.3
280	9	0.18
300	4	0.08
340	2	0.04
合计	50	1.00

从表 6.2 可以看出,订货周期内需求量超过库存量的有 280 件、300 件、340 件,

这时将发生脱销。脱销的概率分别为0.18、0.08 和0.04。如果商店想避免出现脱销,则应当有30、60、90 件的安全库存量。

计算期望年度脱销费用的公式是:

期望年度脱销费用($EASC$) = 缺货件数 × 脱销概率 × 每件商品脱销费用 × 年订货次数

可得到不同安全库存量的期望脱销费用(见表6.3)。

表6.3

安全库存量(件)	缺货件数(件)	脱销概率	脱销费用(元)	年订货次数(次)	期望脱销费用(元)
0	30	0.18	5	6	162
	50	0.08	5	6	120
	90	0.04	5	6	108
					390
30	20	0.08	5	6	48
	60	0.04	5	6	72
					120
50	40	0.04	5	6	48
90	0	0	0	6	0

从表6.3 中看到,随着安全库存量的增加脱销费用减少,但增加安全库存量会增加存储费。我们知道火炬商店的单位存储费是2.5 元。因此:

安全库存量 $M=0$ 时,存储费用为0(元);

安全库存量 $M=30$ 时,存储费用 $=30\times2.5=75$(元);

安全库存量 $M=60$ 时,存储费用 $=50\times2.5=125$(元);

安全库存量 $M=90$ 时,存储费用 $=90\times2.5=225$(元)。

将期望脱销费用加上存储费用则得到不同安全库存量下的总费用(见表6.4)。

表6.4

安全库存量 M(件)	期望脱销费用(元)	存储费用(元)	总费用(元)
0	390	0	390
30	120	75	195
50	48	125	173
90	0	225	225

从表6.4 中看到,当安全库存量为50 件时,年度的总费用为173 元,这时总费用最少。因此,“火炬”商店的最优安全库存量应为50 件,以降低脱销情况的发生。这时,每当库存下降到300 件时,就应该进行下一次的订货,而不是原来的250 件时再

进行下一次的订货。

(二)供货水平的库存模型

“时新”服装店是一个经营时装的商店。商店经理知道服装业是竞争激烈的行业,如果发生脱销,对商店的信誉将是一个沉重的打击,他将可能永远失去一些顾客,这比由于脱销而产生的直接损失严重得多,但很难用货币来衡量脱销一件产品而造成的脱销费用。由于脱销费用估计不出来,则难于用上面的方法计算出他应该持有的安全库存量。但经理知道服装在订货期内的需求量接近正态分布,过去的售货记录表明,订货期为 25 天,在过去的订货期内销售的平均值为 10 000 件,标准差为 600。他希望商店在 95% 的时间内不会发生脱销,才能较好地保持商店的信誉。他希望能有其他的办法为商店找到一个合理的安全库存量。

许多情况下,计算单位脱销费用是很复杂的,甚至是不可能的,因而可采用一种供货水平的库存模型来确定安全库存量。所谓一个供货水平的值,是指在一个订货周期内不致造成脱销的概率。“时新”商店经理希望有 95% 的时间在订货周期内不会发生脱销,就是指在 95% 的订货周期内的满足全部需求量,而只有 5% 的订货周期内可能发生脱销,就说该店采用了一个 95% 的供货水平。很显然,一个 95% 的供货水平比一个 80% 的供货水平需要更多的安全库存量,当然也就需要更多的存储费用,但 95% 的供货水平将减少脱销,从而使脱销费用降低。

利用供货水平的概念来确定安全库存量,要用到分布的知识。“时新”商店的服装在订货期内的需求量接近正态分布(见图 6.7)。

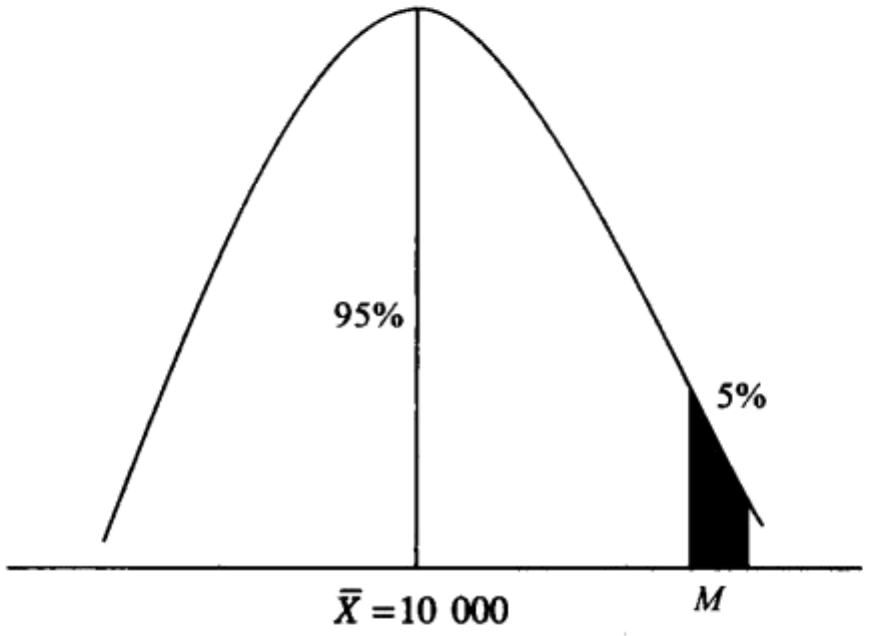


图 6.7

在图 6.7 中,阴影部分表示“时新”商店经理只希望有 5% 的脱销时间。从均值 10 000 件到 M 点间的面积是正态曲线下总面积的 45%,查正态分布表,知 M 点在均值右方 1.65 个标准差的地方,M 点则代表安全库存量的数值,即 $M = 1.65 \times 600 = 990$ 件 。再订货点则为:10 000 + 990 = 10 990件。

因此,“时新”服装店为保证有 95% 的时候不发生脱销,应保有 990 件的安全库存量,当库存量降为 10 990 件时,商店就应进行下一次的订货活动。

二、单周期不确定型库存模型

在现实生活中，有的商品的使用期较短，生产这些商品的厂商在使用期只生产一次，这就迫使订购这些商品的用户在一个周期内只能订货一次。而当需求又是不确定（随机）时，订购者将面临着两种风险损失：一是当订购量大于需求量（供大于求）时，将产生降价销或报废的直接损失；二是当订购量小于需求量（供不应求）时，将产生由于脱销而带来的机会损失。如报纸是有效期为一天的商品，当某报亭某天该报纸的订购量太少，当想再去订购时报社已售罄，这时将失去盈利的机会；若某天该报纸的订购量太多，则售不出的报纸只能当废纸处理，该报亭将会产生直接损失。

因此，所谓单周期不确定型库存问题，是指当需求不确定时，企业在一个周期内只能订货一次时所面临的库存问题，也称单周期随机存储问题。单周期不确定型库存问题广泛存在于需求具有时效性或者易变质、难以保存的产品存储上，比如季节性服装、报纸杂志、日历、易变质食品等。

由于在这类问题中，周期通常较短并且固定，所以一般不考虑订货费用和存储费用问题，而解决的思路在于平衡由于需求的不确定性所面临的两种风险损失，从而确定经济订货量，这实际已是风险决策的问题了（可参见第八章的风险决策）。

第四节　WinQSB 应用举例

WinQSB 软件可以求解表 6.5 所列八个方面的存储问题，其功能强大、内容丰富。

表 6.5

1. Deterministic Demand Economic Order Quantity（EOQ）Problem	确定型需求经济订货批量问题
2. Deterministic Demand Quantity Discount Analysis Problem	确定型需求批量折扣分析问题
3. Single – period Stochastic Demand（Newsboy）Problem	单时期随机需求（报童）问题
4. Multiple – period Dynamic Demand Lot Sizing Problem	多时期动态需求批量问题
5. Continuous Review Fixed – Order – Quantity（s,Q）System	连续盘存的固定订货批量系统
6. Continuous Review Order – Up – To（s,S）System	连续盘存上、下界存量系统
7. Periodic Review Fixed – Order – Interval（R,S）System	定期盘存固定订货区间系统
8. Periodic Review Optional Replenishment（R,s,S）System	定期盘存有选择的再补充订货系统

（1）启动程序。调用的子程序是 Inventory Theory and System（存储论及存储控制

系统)。

(2)建立新问题。系统显示类似表6.5的选项,系统缺损时间单位是年,用户可以修改时间单位,如季度、月、周及天等。当选择第四项时,要求用户输入周期数。下面分别介绍其应用方法。

一、确定性存储模型

对于确定性存储模型,选择表6.5的第一项或第二项就能完成本章第二节所有模型的计算。下面我们举一例子来说明WinQSB软件的计算步骤。

例6.11 已知某汽车车身厂一年钢板的需求量为15 000吨,一次订货成本为5 000元,钢板价格为6 000元/吨,钢板的持有成本为200元/(吨×年),订货提前期为15天,一年按365天计算,则订货提前期为0.0411年。求下列情形的订货策略:

(1)瞬时供货、不允许缺货;

(2)瞬时供货、允许缺货,货到后补充,缺货费200元/吨;

(3)供货速率为每天100吨,允许缺货,缺货费200元/吨;

(4)瞬时供货、不允许缺货。订货量一次达到1 000吨时价格优惠1%,以后每增加2 000吨价格优惠0.5%,最多优惠3%。工厂考虑到库存容量限制,当存储量达到1 000吨时,持有成本将达到500元/(吨×年)。

解:(1)选择表6.5中第一个选项后,输入图6.8所示的数据。在数据选项中,"Demand per year"表示需求量,"Order or Setup Cost per year"表示一次订货成本,"Unit Holding Cost per year"表示持有成本,"Unit Shortage Cost per year"表示缺货成本,"Replenishment or Production rate per year"表示供应速率,"Lead time for a new order in year"表示订货提前期(按年计算),"Unit Acquisition Cost without Discount"表示产品价格。由于是瞬时供货、不允许缺货,可选择系统默认值*M*。

DATA ITEM	ENTRY
Demand per year	15000
Order or setup cost per order	5000
Unit holding cost per year	200
Unit shortage cost per year	M
Unit shortage cost independent of time	
Replenishment or production rate per year	M
Lead time for a new order in year	0.0411
Unit acquisition cost without discount	6000
Number of discount breaks (quantities)	
Order quantity if you known	

图6.8

点击"Solve and Analyze"→"Solve the Problem",求解显示如图6.9所示的计算结果。点击"Results"→"Graphic Inventory Analysis"显示成本变化图;点击"Results"→"Graphic Inventory Profile"显示储量变化图;点击"Results"→"Perform Parametric Analysis"显示参数分析选项,输入所选参数的取值区间和步长,系统会模拟所有取值的经济订货批量及各项成本。

Input Data	Value	Economic Order Analysis	Value
Demand per year	15000	Order quantity	866.0254
Order (setup) cost	¥ 5,000.00	Maximum inventory	866.0254
Unit holding cost per year	¥ 200.00	Maximum backorder	0
Unit shortage cost		Order interval in year	0.0577
per year	M	Reorder point	616.5
Unit shortage cost			
independent of time	0	Total setup or ordering cost	¥ 86,602.54
Replenishment/production		Total holding cost	¥ 86,602.54
rate per year	M	Total shortage cost	0
Lead time in year	0.0411	Subtotal of above	¥ 173,205.08
Unit acquisition cost	¥ 6,000.00		
		Total material cost	¥ 90,000,000.00
		Grand total cost	¥ 90,173,208.00

图 6.9

由图 6.9 可以得到订货策略:每次订货批量为 866 吨,订货间隔期是0.057 7 年,约 21 天订货一次,再订货点是 616 吨,各项成本可见图 6.9。

(2)允许缺货时,将图 6.8 中“Unit Shortage Cost per year”一栏的 M 改为 200。求解后得到订货策略:每次订货批量为 1 225 吨,订货间隔期是0.081 6 年,约 30 天订货一次,再订货点是 4 吨,最大缺货量为 612 吨。

(3)供应速率为 36 500 吨,将图 6.8 中“Unit Shortage Cost per year”一栏的 M 改为 200 及“Replenishment or Production rate per year”一栏的 M 改为 36 500。求解后得到订货策略:每次订货批量为 1 596 吨,订货间隔期是 0.1064 年,约 39 天订货一次,再订货点是 146 吨,最大缺货量为 470 吨。

(4)选择表 6.5 中第 2 个选项后,输入图 6.10 所示的数据,点击“Edit”→“Discount Breaks”,输入图 6.11 中的数据。求解后得到订货策略:每次订货批量为 1 000 吨,订货间隔期是 0.0667 年,约 24 天订货一次。

DATA ITEM	ENTRY
Demand per year	15000
Order or setup cost per order	5000
Unit holding cost per year	500
Unit shortage cost per year	M
Unit shortage cost independent of time	
Replenishment or production rate per year	M
Lead time for a new order in year	0.0411
Unit acquisition cost without discount	6000
Number of discount breaks (quantities)	5
Order quantity if you known	

图 6.10

Discount Breaks

3

Number	Discount Break	Discount %
1	1000	1
2	3000	1.5
3	5000	2
4	7000	2.5
5	9000	3

图 6.11

二、单周期随机存储模型

例 6.12 考虑如例 6.6 的售报问题,此时需求是离散型随机变量。现在某报社为了扩大销售量,报社采取了如下策略:如果报亭每天早上从报社以现金买进报纸,每百份 35 元,零售价每百份 50 元,利润归报亭所有;如果当天没有售完第二天早上

退还报社,报社按每百份报纸 10 元退款。

若某报亭每天对该报纸的需求为随机变量 r,根据销售经验得到 r 的概率分布 $P(r)$ 见表 6.1。应用 WinQSB 软件求解:该报亭每天应该准备多少份报纸最佳,一个月的期望收益是多少?

解:选择表 6.2 中第三项后,系统显示如图 6.12 所示的数据输入格式。首先双击第一行,系统显示如图 6.5 所示的需求分布。

在图 6.5 中选择"Discrete",在图 6.12 的第二行输入随机变量取值个数 6,第三行输入随机变量的取值/概率,共 6 组数据。求解得到图 6.13。由图 6.13 的第 13 行知,最优服务水平为 30.612 2%,由第 11 行和第 14 行知,报亭每天应订购份报纸 5 百份,期望收益是 88.69 元。

DATA ITEM	ENTRY
Demand distribution (in year)	Discrete
Number of discrete values	6
Discrete values	3/0.07,4/0.15,5/0.2,6/0.3,7/0.18,8/0.1
(Not used)	
Order or setup cost	
Unit acquisition cost	35
Unit selling price	50
Unit shortage (opportunity) cost	
Unit salvage value	10
Initial inventory	
Order quantity if you know	
Desired service level (%) if you know	

图 6.12

22-20	Input Data or Result	Value
1	Demand distribution (in day)	Discrete
2	Demand mean	5.67
3	Demand standard deviation	1.3788
4	Order or setup cost	0
5	Unit cost	¥ 35.00
6	Unit selling price	¥ 50.00
7	Unit shortage (opportunity) cost	0
8	Unit salvage value	¥ 1.00
9	Initial inventory	0
10		
11	Optimal order quantity	5
12	Optimal inventory level	5
13	Optimal service level	30.6122%
14	Optimal expected profit	¥ 88.69

图 6.13

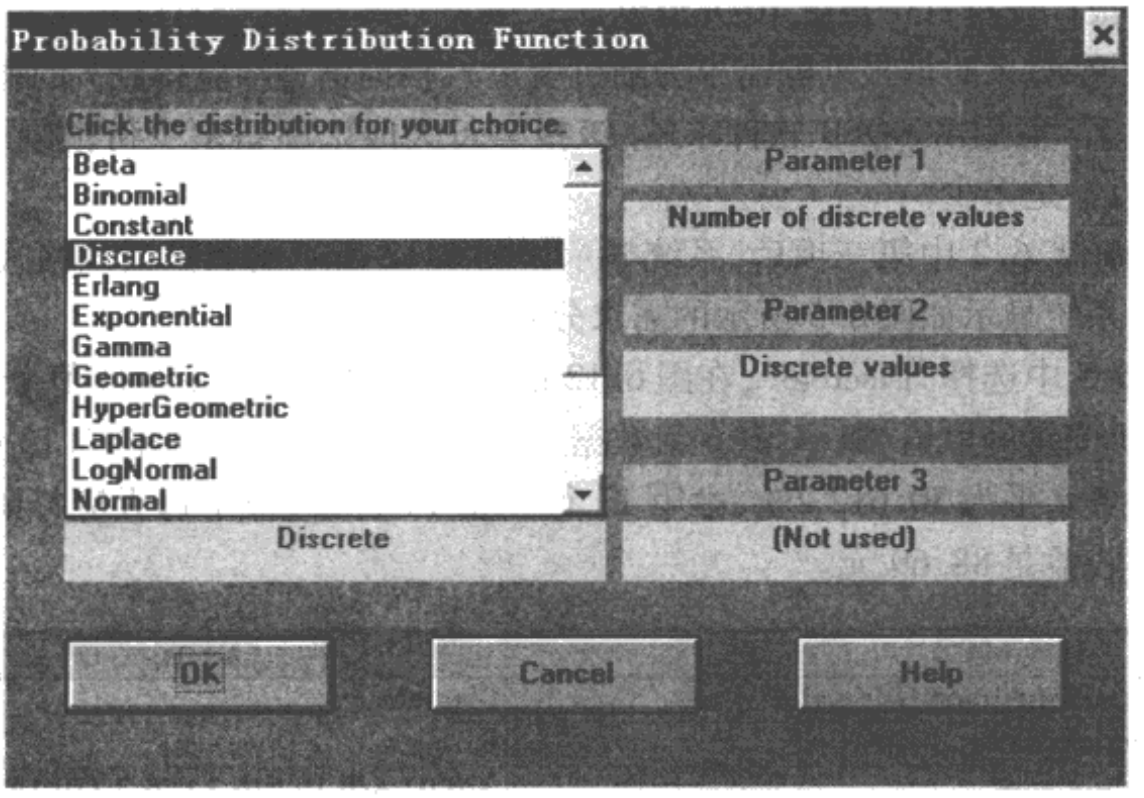


图 6.14

例 6.13 应用 WinQSB 软件求解例 6.8。

解:在图 6.5 中选择“Norm”,输入数据如图 6.14 所示,求解得到图 6.16。因此每天最佳制作量为 79.307 5 杯。

DATA ITEM	ENTRY
Demand distribution (in year)	Normal
Mean (u)	75
Standard deviation (s>0)	10
(Not used)	
Order or setup cost	
Unit acquisition cost	1
Unit selling price	3
Unit shortage (opportunity) cost	
Unit salvage value	0
Initial inventory	
Order quantity if you know	
Desired service level (%) if you know	

图 6.15

22-2(	Input Data or Result	Value
1	Demand distribution (in year)	Normal
2	Demand mean	75
3	Demand standard deviation	10
4	Order or setup cost	0
5	Unit cost	¥ 1.00
6	Unit selling price	¥ 3.00
7	Unit shortage (opportunity) cost	0
8	Unit salvage value	0
9	Initial inventory	0
10		
11	Optimal order quantity	79.3075
12	Optimal inventory level	79.3075
13	Optimal service level	66.6667%
14	Optimal expected profit	¥ 139.09

图 6.16

三、多周期随机存储模型

例 6.14 已知一月初有 30 件库存产品,要求四月末有 10 件库存产品,其他资料如表 6.6 所示,要求制定 4 个月的生产与存储策略。

表 6.6

月份	需求量	准备成本	单位变动成本	单位持有成本	单位缺货成本
1	60	10	2.5	0.5	1
2	80	13	2.6	0.5	1.2
3	90	12	2.7	0.6	1.2
4	100	15	4	0.6	1.3

解:选择表 6.5 中第四项,将表 6.6 的数据输入到图 6.17 中,点击"Solve and Analyze"系统显示如图 6.18 所示的方法选项,共有 10 种求解准则,如选择 EOQ 准则,并输入期初和期末库存量,如果期末允许缺货,则期末库存量输入负数。求解结果如表 6.19 所示。

month	Demand	Setup Cost	Unit Variable Cost	Unit Holding Cost	Unit Backorder Cost
1	60	10	2.5	0.5	1
2	80	13	2.6	0.5	1.2
3	90	12	2.7	0.6	1.2
4	100	15	4	0.6	1.3

图 6.17

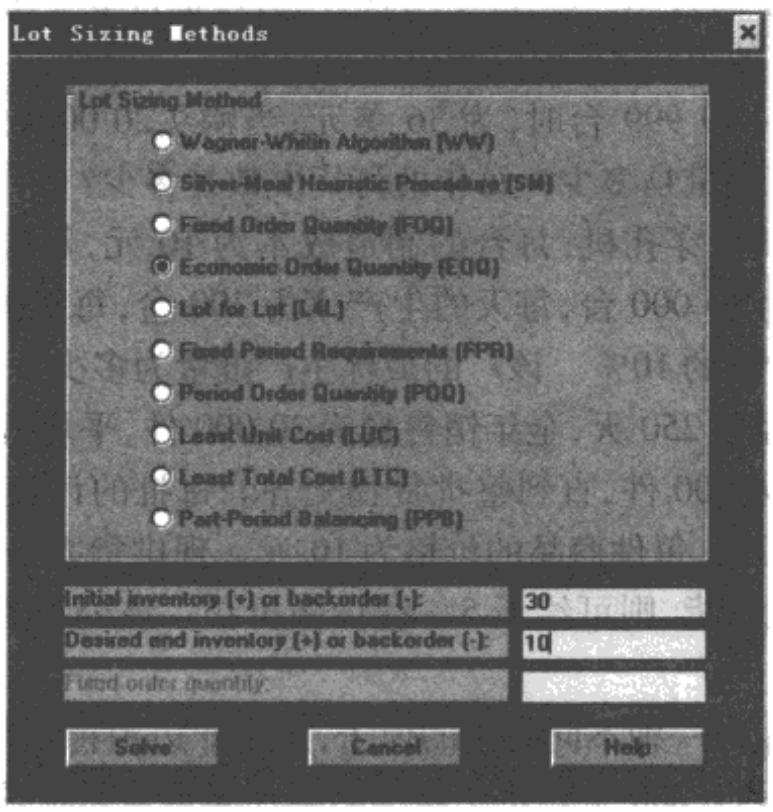


图 6.18

-22-20	Demand	Production (Lot Size)	Setup	Expected Inventory	Expected Backorder	Cumulative Cost
Initial				30.0000		
1	60.0000	72.0000	Yes	42.0000	0	¥ 211.00
2	80.0000	72.0000	Yes	34.0000	0	¥ 428.20
3	90.0000	72.0000	Yes	16.0000	0	¥ 644.20
4	100.0000	94.0000	Yes	10.0000	0	¥ 1,041.20
	Solution	Method:	EOQ		Total Cost =	¥ 1,041.20

图 6.19

习题

1. 某公司出售旅行电熨斗，全年营业 360 天。每年需求量为 6 400 个，单价为 10 美元，每年的单位存储费用为货物价值的 20%，每次订购费用为 100 美元，不允许缺货情况出现，试求其经济订货批量，每年的订货次数，订货间隔时间，及库存的总费用。

2. 科布公司每年购买 40 000 台专用收音机，每批订货费用为 5 000 美元，单位存储费用为货物价值的 10%，每台收音机的单价如下：当购买 1 ~ 12 499 台时，为 40 美元；当购买 12 500 ~ 19 999 台时，为 36 美元；当购买20 000台及以上时，为 32 美元。则最小费用的订购量是多少？每年的订货次数是多少？

3. 某工厂生产纸张穿孔机，每台的制造成本为 10 元，每次的生产调整费用为 500 元，每年的产量为 20 000 台，每天的生产率为 300 台，每天的销售率为 60 台，存储费用为平均库存价值的 10%。该厂的最优生产批量为多少？

4. 某公司一年经营 250 天，全年销售量为 20 000 件，平均每天销售 80 件。每次订货后，进货率为每天 400 件，直到整批货物交齐。每批的订货费用为 640 元，存储费用为货物价值的 20%，每件商品的价格为 10 元。现供货方告之该公司，若能一次整批进货而不是分批送货，则可给予 5% 的价格折扣。该公司是否应该接受供货方的这个建议？

5. 某公司成批生产一种不同规格的窗子，边生产边销售。每天生产 50 个，销售 30 个，年销售量估计为 6 000 个。每个窗子的年存储费为 2 元，每次生产的调整费用为 150 元。试确定该公司在不允许缺货条件下最优的经济生产批量、每年总费用、每年调整生产的次数及最大存储水平。

6. 某商场销售某品牌的方便面，已知该方便面的周销售量为 50 箱，每次订购费用为 100 元，每箱每天的存储费用为 0.2 元，每箱每周的缺货费用为 5 元。试求：(1)若不允许缺货，最优的经济订购批量和最低周费用；(2)若允许缺货发生，最优订购批量，周总平均费用，最大缺货量。

7. 某公司生产某种商品，年生产量为 50 000 件，年需求量为30 000件；每次生产准备费用为 1 000 元，每件产品每年的存储费用为 28 元，试计算在不允许缺货条件下最优的经济生产批量、年总费用，周期时间，以及最大存储水平。

8. 某公司销售某种机电设备，每年订货 6 次，交货期限为 6 天。每天平均销售量为 5 台。每台的年存储费为 100 元，每台的脱销费用为 100 元。过去 200 次订货周期内的需求量为 25、30、35、40 台的次数分别为 60、90、40、10 次。试帮助该公司确定合理的安全库存量。

9. 某服装专卖店统计，在过去的订货周期内，服装的平均需求量为 200 件，标准差为 40 件。该专卖店的订货批量为 800 件，年单位存储费为 10 元。该专卖店原实行 90% 的供货水平，现欲将供货水平提高到 95%。试问由于供货水平的改变，该店的再订货点为多少？增加多少安全库存量？由此增加的存储费是多少？

10. 某书店出售某种专业书籍，订货期内的平均需求量为 400 本，标准差为 80 本，每本书籍的年单位存储费为 10 元。该店希望了解在只允许 5%、10% 缺货时的再订货点、安全库存量及存储费各为多少？

第七章　排队论决策

引例:高速公路收费站排队问题

高速公路收费站设有一个收费通道,汽车到达服从泊松分布,平均到达速率为150辆/小时,收费时间服从负指数分布,平均收费时间为15秒/辆。求收费处空闲和忙的概率各是多少?系统中分别有1、2、3辆车的概率是多少?

类似引例这种现象,我们在日常生活中经常遇到,如顾客到商店去买东西,病人到医院去看病,当售货员、医生的数量满足不了顾客或病人及时服务的需求时,就会出现排队现象。这种现象的出现使人感到厌烦,但由于顾客到达人数和服务时间的随机性,使得排队现象又经常不可避免,一个办法就是增加服务设施(售货员、医生),这样同时会增加投资,但当供大于求时,设施又常常空闲、浪费。作为管理人员就需要考虑排队问题,把排队时间控制在一定的限度内,在服务质量的提高和成本的降低之间找到平衡,找到最合适的方案。

第一节　基本概念

排队论(Queuing Theory)又称随机服务系统理论(Random Service System Theory),是研究排队系统的数学理论和方法,是运筹学的一个重要分支。具体说,它是在研究各种排队论系统概率规律性的基础上,解决相应排队系统的最优设计和最优控制问题。排队论被广泛地用于解决诸如电话的占线,车站、码头、机场等交通枢纽的堵塞与疏导,故障机器的停机待修,水库的存储调节等排队问题。

一、排队系统的描述

上述各种问题虽互不相同,却都有要求得到某种服务的人或物和提供服务的人或机构。排队论里要求服务的对象统称为“顾客”,而把提供服务的人或机构称为“服务台”或“服务员”。不同的顾客与服务组成了各式各样的排队系统。

任何一个排队问题的基本排队过程都可以用图7.1表示。每个顾客由顾客源按一定的方式到达服务系统,首先加入排队队列等待接受服务,服务台按一定规则从队列中选择顾客进行服务,获得服务的顾客立即离开。

从基本的排队系统可以引申出单服务台单队系统、多服务台单队系统、多服务台多队系统、多服务台串联系统,如图7.2至图7.5所示。

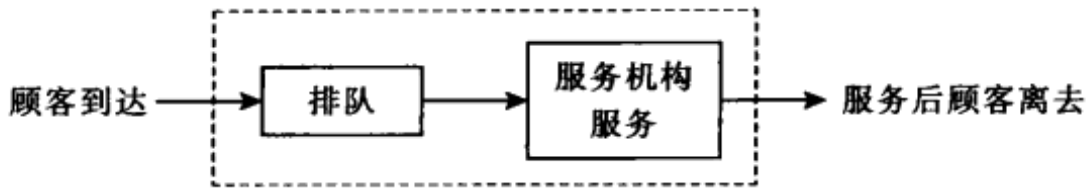


图 7.1

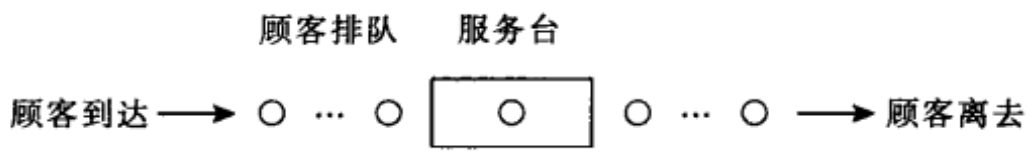


图 7.2

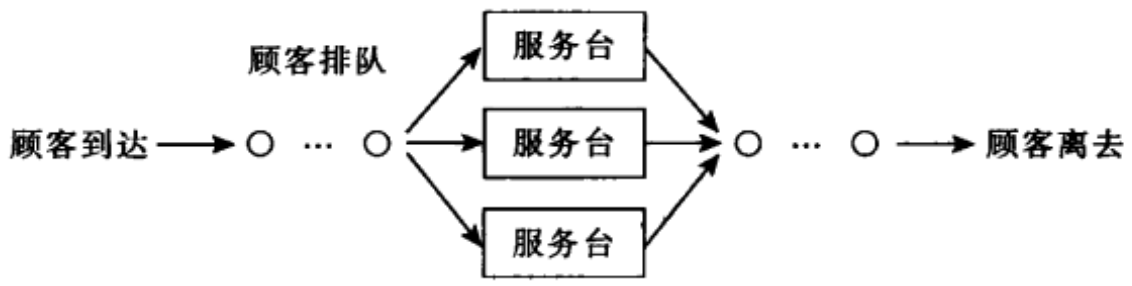


图 7.3

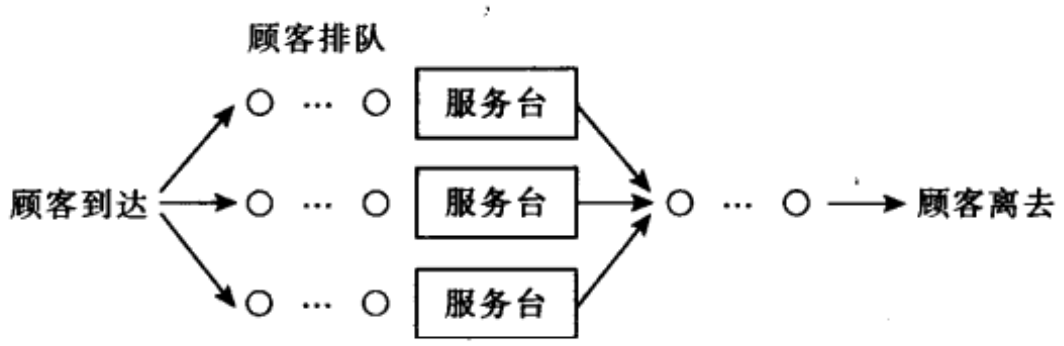


图 7.4

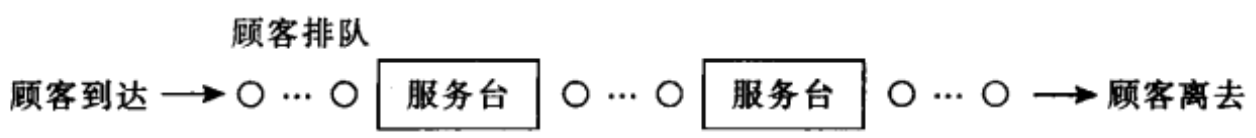


图 7.5

二、排队系统的组成

一般来说,排队系统由输入过程、排队规则和服务机构三个部分组成。

1. 输入过程

输入过程是指要求服务的顾客按怎样的规律到达排队系统的过程,有时也称之为顾客流。一般可以从三个方面来描述一个输入过程:

(1) 顾客总体数,又称顾客源、输入源。顾客源可以是有限的,也可以是无限的。如到售票处购票的顾客总数可以认为是无限的,而某个工厂因故障待修的机床则是有限的。

(2) 顾客到达的形式。这是描述顾客是怎样来到系统的,是单个到达,还是成批到达。大学生到图书馆借书是单个到达的例子,而购买材料入库则可以看成是成批

到达。

(3) 顾客流的概率分布或称顾客相继到达时间间隔分布,这是首先需要确定的指标。顾客流的概率分布一般有定长分布、二项分布、泊松分布和负指数分布等。

2. 排队规则

排队分为有限排队和无限排队两类。前者是指系统的空间是有限的,当系统被占满时,后面再来的顾客将不能进入系统;后者是指系统中的顾客数可以是无限的,队列可以排到无限长,顾客到达后均进入系统排对或接受服务。具体分为以下三种:

(1) 等待制。它指顾客到达系统后,所有服务台都不空,顾客加入排队行列等待服务,一直等到服务完毕以后才离去。如排队等待售票,故障设备等待维修等。等待制中,服务台选择顾客进行服务时通常有如下四种规则:

先到先服务(First Come First Serve,FCFS)。它指按顾客到达的先后顺序进行服务,这是最普遍的情形。

后到先服务(Last Come First Serve,LCFS)。仓库中叠放的钢材,后放上去的先被领走,重大消息先刊登,都属于这种情形。

随机服务(Service in Random Order,SIRO)。当服务台空闲时,不按排队序列而随意指定某个顾客去接受服务,如电话交换台接通呼叫就是一例。

有优先权的服务(Priority)。如老人、小孩先进车站,重病号先就诊,遇到重要数据需要立即中断其他数据的处理等,均属于这种规则。

(2) 损失制。它指当顾客到达系统时,所有服务台都已被占用,顾客不愿等待而离开系统。如电话拨号后出现忙音,顾客不愿等待而挂断电话,如要再打则需要重新拨号。

(3) 混合制。这是等待制与损失制相结合的一种服务规则,一般允许排队,但又不允许队列无限长下去。大体有以下三种:

①队长有限。当等待服务的顾客人数超过规定数量时,后来的顾客就自动离去,另求服务,即系统的等待空间是有限的。

②等待时间有限。即顾客在系统中的等待时间不超过某一给定的长度,当等待时间超过时,顾客将自动离去,并且不再回来。

③逗留时间(等待时间与服务时间之和)有限。

3. 服务机构

服务机构可以从以下三个方面来描述:

(1) 服务机构数量即构成形式。从数量上说。服务台有单台和多台之分。从构成形式上看,有单队单服务台式、单队多服务台并联式、多队多服务台并联式、单队多服务台串联式等,如图 7.2 至图 7.5 所示。

(2) 服务方式。它指在某一时刻接受服务的顾客数,有单个服务和成批服务两种。

(3) 服务时间的分布。在多数情况下,对某一个顾客的服务时间是一随机变量,与顾客到达的时间间隔分布一样,服务时间的分布有定长分布、负指数分布、爱尔朗分布等。

三、排队系统的主要数量指标和符号

1. 排队系统的主要数量指标

研究排队系统的主要目的是通过了解系统运行的状况，对系统进行调整和控制，使系统处于最优运行状态。因此，首先需要弄清系统的运行状况。描述一个排队系统运行状况的主要数量指标有：

(1) 队长和队列长(排队长)。队长是指标系统中的顾客数(排队等待的顾客数之和)；队列长是指系统中正在排队等待服务的顾客数。队长和队列长一般都是随机变量。队长的分布是顾客和服务员都关心的，特别是对系统设计人员来说，如果能知道队长的分布，就能确定队长超过某个数的概率，从而确定合理的等待空间。

(2) 等待时间和逗留时间。从顾客到达时刻起到他开始接受服务止这段时间称为等待时间。等待时间是个随机变量，也是顾客最关心的指标，因为顾客通常是希望等待时间越短越好。从顾客到达时刻起到他接受完服务止这段时间称为逗留时间，也是随机变量，顾客同样非常关心。

(3)忙期和闲期。忙期是从顾客到达空闲着的服务机构起，到服务机构再次成为空闲止的这段时间，也是服务机构连续忙的时间。这是个随机变量，是服务员最关心的指标，因为它关系到服务员的服务强度。与忙期相对的是闲期，即服务机构连续保持空闲的时间。在排队系统中，忙期和闲期总是交替出现的。

除了上述几个基本数量指标外，还会用到其他一些重要指标。如在损失制或系统容量有限的情况下，由于顾客被拒绝而使服务系统受到损失的顾客损失率及服务强度等，也是十分重要的指标。

2. 排队系统中常用记号

$N(t)$：t 时刻系统中的顾客数(又称为系统的状态)，即队长。

$N_q(t)$：t 时刻系统中排队的顾客数，即队列长。

$T(t)$：t 时刻到达系统的顾客在系统中的逗留时间。

$T_q(t)$：t 时刻到达系统的顾客在系统中的等待时间。

上面给出的这些数量指标一般都是和系统运行的时间有关的随机变量，直接求出它们的瞬时分布一般是很困难的。一般情况下，排队系统在运行了一段时间后，都会趋于一个平稳状态，在平稳状态下，队长的分布、等待时间的分布和忙期的分布都和系统所有的时刻无关，而且系统的初始状态的影响也会消失。因此，我们在本章中将主要讨论统计平衡性质。

L：平均队长，即稳态系统任一时刻顾客数的期望值。

L_q：平均等待队长，即稳态系统任一时刻等待服务的顾客数的期望值。

W：平均逗留时间，即在任一时刻进入稳态系统的顾客逗留时间的期望值。

W_q：平均等待时间，即在任一时刻进入稳态系统的顾客等待时间的期望值。

这四项主要性能指标的值越小，说明系统排队越少，因而系统性能越好。它们是顾客与服务系统的管理者都非常关注的。

λ：顾客到达的平均速率，即单位时间内平均到达的顾客数。

$1/\lambda$：平均到达时间间隔。

μ:平均服务速率,即单位时间内服务完毕离去的顾客数。

$1/\mu$:平均服务时间。

s:系统中服务台的个数。

ρ:服务强度,即每个服务台单位时间内的平均服务时间。

N:稳态系统任一时刻的状态(即系统中所有顾客数)。

U:任一顾客在稳态系统中的逗留时间。

Q:任一顾客在稳态系统中的等待时间。

$P_n = P\{N = n\}$:稳态系统任一时刻状态为 n 的概率;特别地,当 $n = 0$ 时,$P_n = P_0$,即稳态系统所有服务台全部空闲的概率。

λ_n:有效平均到达率,即期望每单位时间内来到系统(包括未进入系统)的概率。

3. 排队系统的符号表示

为了区别各种排队系统,根据输入过程、排队规则和服务机构的变化对排队模型进行描述或分类,可得到很多模型。1953 年,肯道尔(Kendall) 提出一个分类方法,即 Kendall 符号,其形式是 $X/Y/Z$。

在 1971 年一次关于排队论符号标准化的国际会议上,Kendall 符号被扩充为 $X/Y/Z/A/B/C$ 或$[X/Y/Z]:[A/B/C]$ 标准形式。

(1) X 表示顾客相继到达时间间隔的概率分布,可取 M、D、E_K、G 等,其中:

M 表示到达过程为泊松过程或负指数分布;

D 表示定长输入;

E_K 表示 K 阶爱尔朗(Erlang) 分布;

G 表示一般互相独立的随机分布。

(2)Y 表示服务时间分布,所用符号与 X 相同。

(3)Z 表示服务台个数,取正整数。1 表示单个服务台,$s(s > 1)$ 表示多个服务台。

(4)A 表示系统中顾客容量限额,或称等待空间容量。若系统中有 K 个等待位子 $(0 \leqslant K \leqslant \infty)$,当 $K = 0$ 时,说明系统不允许等待,即为损失制;当 $K = \infty$ 时为等待制系统;当 K 为有限整数时,表示为混合制系统。

(5)B 表示顾客源限额,可取正整数或 ∞,即有限与无限两种。

(6)C 表示服务规则,如 $FCFS$、$LCFS$ 等。

例如$[M/M/1]:[\infty/\infty/FCFS]$ 表示顾客到达的时间间隔是负指数分布,服务时间是负指数分布,一个服务台,排队系统和顾客源的容量都是无限的,实行先到先服务的一个服务系统。

第二节　排队系统中的常用分布

在排队系统中,顾客相继到达的时间间隔与服务的时间分布主要有负指数分布、泊松分布和爱尔朗分布等。

一、负指数分布

由概率论可知,如果随机变量服从负指数分布,则其分布函数为:

$F_T(t) = 1 - e^{-\lambda t}(t \geqslant 0, \lambda \geqslant 0)$

,密度函数为$f_T(t) = \lambda e^{-\lambda t}(t \geqslant 0, \lambda \geqslant 0)$

T 的期望值为 $E(T) = \int_0^{\infty} t f_T(t)dt = \int_0^{\infty} t\lambda e^{-\lambda t}dt = \frac{1}{\lambda}$

T 的方差为 $Var(T) = \frac{1}{\lambda^2}$

负指数分布具有以下重要性质：

(1) 密度函数$f_T(t)$ 对时间 t 严格递减。

(2) 无记忆性或马尔柯夫性，即 $P\{T > t + l \mid T > l\} = P\{T > t\}$。

该性质说明一个顾客到来所需的时间与过去一个顾客到来所需的时间 l 无关，这种情形下的顾客到达是纯随机的。

(3) 当顾客到达过程是泊松流时，顾客相继到达的间隔时间 T 必服从负指数分布。

二、泊松分布

若随机变量 X 的概率密度为 $P\{X = n\} = \frac{\lambda^n e^{-\lambda}}{n!}(\lambda > 0, n = 0,1,2,\cdots)$ 时，则称 X 服从参数为 λ 的泊松分布，记为 $X \sim P(\lambda)$。其均值和方差分别为：

$E(X) = \lambda, Var(X) = \lambda$

1. 泊松过程的定义

泊松过程是应用最为广泛的一类随机过程，它常用来描述排队系统中顾客到达的过程、城市中的交通事故、保险公司的理赔次数等。泊松过程是构造更复杂的随机过程，是一个非常重要的随机过程。

记 $N(t)$ 表示在时间区间$[0,t](t > 0)$ 内发生的事件数，若 $N(t)$ 是一个随机变量，则$\{N(t) \mid t \in (0,T)\}$ 就称为一个随机过程。

定义 7.1 ［泊松过程］对于随机过程$\{N(t) \mid t \in (0,T)\}$，若满足：

(1) 独立增量性，即对任意 n 个参数 $t_n > t_{n-1} > t_{n-2}\cdots t_1 \geqslant 0$，增量 $N(t_2) - N(t_1), N(t_3) - N(t_2), \cdots, N(t_n) - N(t_{n-1})$ 互相独立。

(2) 增量平稳性，即在长度为 t 的时间区间内恰好到达 k 个顾客的概率仅与区间长度 t 有关，而与区间起始点无关。对任意 $b \in (0,\infty)$，在$(b,b + t)$ 内恰好到达 k 个顾客的概率相等，有：

$P\{N(b + t) - N(t) = k\} = P\{N(t) - N(0) = k\} = P_k(t)$

(3) 普遍性，即当 t 充分小时，有：

$P\{N(t) = 1\} = \lambda t + o(t)$

$P\{N(t) = 0\} = 1 - \lambda t + o(t)$

$P\{N(t) \geqslant 2\} = o(t)$

即称上述过程为泊松过程，其中 λ 为泊松过程的参数，且 $N(t)$ 服从泊松分布。

2. 排队系统与泊松过程

设 $N(t)$ 为时间区间$[0,t](t > 0)$ 内到达系统的顾客数，则 $N(t)$ 是一个随机变量，且$\{N(t) \mid t \in (0,T)\}$ 为一个随机过程。若该随机过程满足：

(1) 在不相重叠的区间内,顾客的到达数是相互独立的。

(2) 在时间区间$[t,t+\Delta t]$内,顾客的到达数只与区间长度Δt有关,而与区间起始点t无关。

(3) 对于充分小的Δt,在时间区间$[t,t+\Delta t]$内有两个或两个以上的顾客到达的概率极小,以至可以忽略,即:

$$\sum_{k=2}^{\infty} P_k(t,t+\Delta t) = o(\Delta t)$$

则认为顾客到达系统的过程是泊松过程,且:

$$P\{N(t)=k\} = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \quad (k=0,1,2,\cdots,t>0)$$

$$E[N(t)] = \lambda t \quad Var[N(t)] = \lambda t$$

式中,λ表示单位时间内到达系统的顾客数。

下面的定理,说明了泊松流与负指数分布之间的关系。

定理7.1 在排队系统中,如果到达的顾客数服从以λt为参数的泊松分布,则顾客相继到达的时间间隔服从以λ为参数的负指数分布。

由定理7.1可以看出,"到达的顾客数是一个以λ为参数的泊松流"与"顾客相继到达的时间间隔服从以λ为参数的负指数分布"是等价的。

三、K阶爱尔朗分布

定理7.2 设$X_1,X_2,\cdots,X_K$是K个互相独立的、具有相同参数μ的负指数分布随机变量,则随机变量$X=X_1+X_2+\cdots+X_K$服从K阶爱尔朗分布,X的密度函数为:

$$f(t) = \frac{K\mu(K\mu t)^{K-1}}{(K-1)!}e^{-K\mu t} \quad (t>0)$$

记为$X \sim E_K(\mu)$或简记为$X \sim E_K$。随机变量X的均值和方差分别为:

$$E(X) = \frac{1}{\mu} \quad Var(X) = \frac{1}{K\mu^2}$$

例如,如果顾客连续接受串联的K个服务台的服务,各服务台的服务时间互相独立,且均服从参数为μ的负指数分布,则顾客接受K个服务台总共所需的时间就服从K阶爱尔朗分布。

第三节 排队模型

一、单服务台模型

在这部分中,我们将讨论顾客到达过程服从泊松过程,服务时间服从负指数分布的单服务台排队模型,分成以下三种情况:

(1) 基本模型,即$[M/M/1]:[\infty/\infty/FCFS]$。

(2) 有限队列模型,即$[M/M/1]:[N/\infty/FCFS]$。

(3) 有限顾客源模型,即$[M/M/1]:[\infty/m/FCFS]$。

（一）基本模型

基本模型满足如下基本假定：

（1）输入过程。顾客源是无限的，顾客的到达过程是泊松过程。

（2）排队规则。单个队列，对队长无限制，先到先服务。

（3）服务机构。单服务台，服务时间服从负指数分布。

（4）服务时间和顾客到达的时间间隔相互独立。

如果单位时间到达排队系统的顾客数为 λ，单位时间服务的顾客数为 μ。在单服务台且顾客源无限的情况下，系统在稳态情况下的状态转移图如图 7.6 所示。

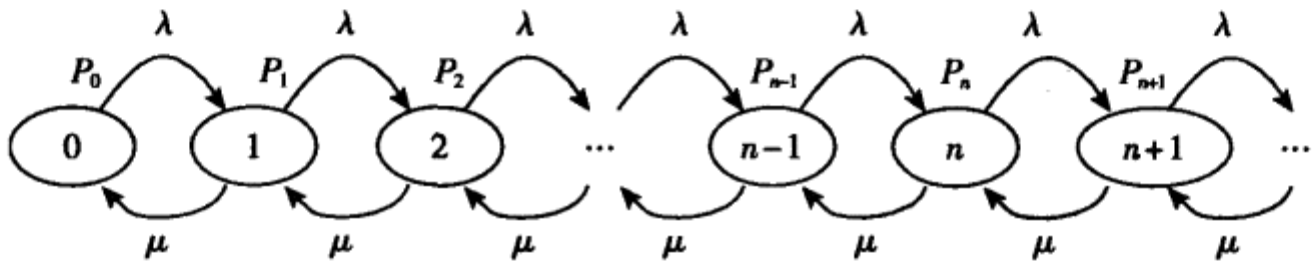


图 7.6

根据以上状态转移图，可以得到如下平衡方程组：

$$\mu P_1 - \lambda P_0 = 0$$

$$\lambda P_{n-1} + \mu P_{n+1} - (\lambda + \mu) P_n = 0 \quad (n = 1,2,\cdots)$$

这样，我们就可以得到：

$$P_1 = \frac{\lambda}{\mu} P_0$$

$$P_2 = -\frac{\lambda}{\mu} P_0 + (1 + \frac{\lambda}{\mu}) P_1 = (\frac{\lambda}{\mu})^2 P_0$$

$$P_n = (\frac{\lambda}{\mu})^n P_0 \quad (n = 1,2,\cdots)$$

设 $\rho = \frac{\lambda}{\mu} < 1$，我们就可以得到：

$$P_1 = \rho P_0, P_2 = \rho^2 P_0, \cdots, P_n = \rho^n P_0$$

而由 $\sum_{n=1}^{\infty} P_n = 1$，有：

$$P_0 = 1 - \rho$$

$$P_n = (1 - \rho)\rho^n \quad n \geqslant 1$$

例 7.1 问题如引例。

解：根据题意，$\lambda = 150$ 辆／小时，$1/\mu = 15$ 秒／辆 $= 1/240$ 小时／辆，即 $\mu = 240$ 辆／小时。$\rho = \frac{\lambda}{\mu} = 150/240 = \frac{5}{8}$，则有：

（1）系统空闲的概率为：$P_0 = 1 - \rho = 0.375$；

（2）系统忙的概率为：$1 - P_0 = 0.625$；

（3）系统中有 1 辆车的概率为：$P_1 = \rho(1 - \rho) = 0.234$；

（4）系统中有 2 辆车的概率为：$P_2 = \rho^2(1 - \rho) = 0.146$；

(5) 系统中有 3 辆车的概率为:$P_3 = \rho^3(1-\rho) = 0.091$。

下面,我们不加证明地给出基本模型系统的数量指标。

(1) 系统的平均顾客数 $L = \frac{\rho}{1-\rho}$;

(2) 队列中的平均顾客数 $L_q = \frac{\rho^2}{1-\rho} = \rho L$;

(3) 顾客在系统里的平均逗留时间 $W = \frac{1}{\mu - \lambda}$;

(4) 顾客在队列里的平均逗留时间 $W_q = \frac{\rho}{\mu - \lambda} = \rho W$。

这四个公式称为Little公式,具有非常直观的含义:如果系统处于稳定状态,则系统中的平均人数就等于顾客在系统中的平均逗留时间乘以系统的平均到达率。这些指标之间存在如下的相互关系:

$$L = \lambda W$$

$$L_q = \lambda W_q$$

$$W = W_q + \frac{1}{\mu}$$

$$L = L_q + \rho = L_q + \frac{\lambda}{\mu}$$

例 7.2 医院看病挂号处有一个挂号窗口,病人到达服从泊松分布,平均到达率为 200 人/小时,挂号时间服从负指数分布,平均挂号时间为 15 秒/人。求挂号的平均顾客数,挂号排队的平均顾客数,病人在系统里的平均逗留时间以及病人在队列里的平均逗留时间。

解:根据题意,$\lambda = 200$ 人/小时,$\mu = 240$ 人/小时。$\rho = \frac{\lambda}{\mu} = \frac{5}{6}$,则:

$$L = \frac{\rho}{1-\rho} = 5$$

$$L_q = \frac{\rho^2}{1-\rho} = \rho L = 4.17$$

$$W = \frac{1}{\mu - \lambda} = 90(\text{秒})$$

$$W_q = \frac{\rho}{\mu - \lambda} = \rho W = 75(\text{秒})$$

(二) 有限队列模型

当队列的容量从无限值变为有限值 N 时,基本模型就转化为有限队列模型。如果系统的最大容量为 N,对于单服务台的情形,排队等待的顾客最多为 $N-1$,在某一时刻有顾客到达时,若系统中已有 N 个顾客,那么这个顾客就被拒绝接纳入系统。有限队列模型系统的状态转移如图 7.7 所示。

根据以上状态转移图,可以得到如下平衡方程组:

$$\mu P_1 = \lambda P_0$$

$$\mu P_{k+1} + \lambda P_{k-1} = (\lambda + \mu) P_k \quad k \leqslant N-1$$

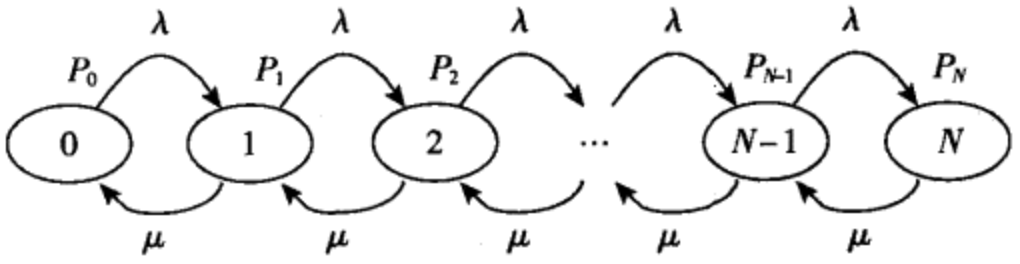


图 7.7

$$\mu P_N = \lambda P_{N-1}$$

由 $\sum_{k=1}^{N} P_k = 1$ 知，当 $\rho = \frac{\lambda}{\mu} < 1$ 时，我们就可以得到：

$$P_0 = \frac{1-\rho}{1-\rho^{N+1}}$$

$$P_k = \rho^k \frac{1-\rho}{1-\rho^{N+1}} \quad (k \leqslant N)$$

当 $\rho = \frac{\lambda}{\mu} = 1$ 时，我们就可以得到：

$$\sum_{k=0}^{N} \rho^k = N+1$$

$$P_k = \rho^k \cdot P_0 = P_0$$

于是：

$$P_0 = \frac{1}{\sum_{k=0}^{N} \rho^k} = \frac{1}{N+1}$$

$$P_k = \rho^k \frac{1}{N+1} \quad (k = 1,2,\cdots,N)$$

有限队列模型系统的数量指标为($\rho \neq 1$)：

(1) 系统的平均顾客数 $L = \sum_{k=0}^{N} kP_k = \frac{\rho}{1-\rho} - \frac{(N+1)\rho^{N+1}}{1-\rho^{N+1}}$。

(2) 队列中的平均顾客数 $L_q = \sum_{k=1}^{N} (k-1)P_k = L - \rho(1-P_N) = L - (1-P_0)$。

令 $\lambda_e = \lambda(1-P_N)$，$\rho_e = \frac{\lambda_e}{\mu}$，其中，$\lambda_e$ 称为有效到达率，表示单位时间内到达能进入队列的平均顾客数；ρ_e 称为有效服务强度，且 $L_q = L - \rho_e$。

(3) 顾客在系统里的平均逗留时间 $W = \frac{L}{\lambda_e} = \frac{1}{\lambda(1-P_N)}$。

(4) 顾客在队列里的平均逗留时间 $W_q = \frac{L_q}{\lambda_e} = \frac{L-\rho_e}{\lambda_e} = \frac{L}{\lambda_e} - \frac{1}{\mu} = W - \frac{1}{\mu}$。

例 7.3 统计咨询中心有一位固定咨询工作人员，每次只能咨询一个人，另外有四个座位供前来咨询的人等候。某顾客到达后发现没有座位，就会离去。前来咨询者的到达服从泊松过程，到达的平均速率为 4 人／小时，咨询人员的平均咨询时间为 10 分钟／人。咨询时间服从负指数分布。求：

(1) 咨询者到达不用等待就可以咨询的概率;

(2) 咨询中心的平均人数以及等待咨询的平均人数;

(3) 咨询者前来咨询中心一次平均花费的时间以及平均等待的时间;

(4) 咨询者到达后因客满而离去的概率;

(5) 增加一个咨询工作人员可以减少的顾客损失率。

解:这是一个有限队列模型系统,其中 $N = 5$,$\lambda = 4$ 人/小时,$\mu = 6$ 人/小时,$\rho = 2/3$。

(1) $P_0 = \dfrac{1-\rho}{1-\rho^{N+1}} = 0.365$, $\lambda_e = \lambda(1 - P_N) = \lambda(1 - \rho^N P_0) = 3.808$;

(2) $L = \dfrac{\rho}{1-\rho} - \dfrac{(N+1)\rho^{N+1}}{1-\rho^{N+1}} = 1.423$, $L_q = L - \dfrac{\lambda_e}{\mu} = 0.788$;

(3) $W = \dfrac{L}{\lambda_e} = 0.374$(小时), $W_q = \dfrac{L_q}{\lambda_e} = 0.207$(小时);

(4) $P^5 = \rho^5 P_0 = 0.048$;

(5) 当 $N = 6$ 时, $P_0 = \dfrac{1-\rho}{1-\rho^{N+1}} = 0.365$, $P^6 = \rho^6 P_0 = 0.032$, $P_5 - P_6 = 0.048 - 0.032 = 0.016$,即增加一个咨询工作人员可以减少的顾客损失率为 16.9%。

(三) 有限顾客源模型

如果顾客总数为 m,当顾客需要服务时,就进入队列等待;服务完毕后,重新回到顾客源中,如此循环往复。这样的模型就是有限顾客源模型。比较典型的有限顾客源问题是机器维修问题。有 m 台机器在运转,单位时间内平均出现故障的机器数可以看成顾客平均到达率 λ,修理工修理一台设备的平均时间为平均服务时间 μ,已经修复的机器可能再出现故障。实际上,在这个问题中,顾客源的数量是有限的,因而队列的长度也是有限的,且队列的长度必定小于顾客总数。

在稳态的情况下,当系统由状态 0 转移到状态 1,每台设备由正常状态变为故障状态,转移率为 λP_0;如果有 m 台设备由正常状态变为有一台设备(无论哪一台)发生故障,转移率为 $m\lambda P_0$;当系统由状态 1 转移到状态 0,转移率为 μP_1,所以在状态 0 时有平衡方程 $m\lambda P_0 = \mu P_1$。有限顾客源模型的状态转移如图 7.8 所示。

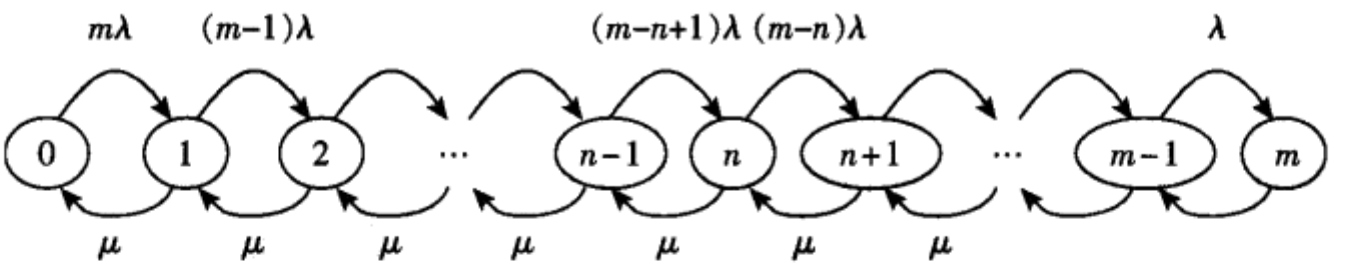


图 7.8

根据以上状态转移图,可以得到如下平衡方程组:

$$\mu P_1 = m\lambda P_0$$

$$\mu P_{n+1} + (m-n+1)\lambda P_{n-1} = [(m-n)\lambda + \mu]P_n \quad (1 \leqslant n \leqslant m-1)$$

$$\mu P_m = \lambda P_{m-1}$$

由 $\sum_{k=0}^{m} P_k = 1$（不要求 $\rho = \frac{\lambda}{\mu} < 1$），我们可以得到：

$$P_0 = \frac{1}{\sum_{i=0}^{m} \frac{m!}{(m-i)!}\rho^i}$$

$$P_n = \frac{m!}{(m-n)!}\rho^n P_0$$

有限顾客源模型系统的数量指标为：

（1）系统的平均顾客数 $L = m - \frac{\mu}{\lambda}(1 - P_0)$；

（2）队列中的平均顾客数 $L_q = m - (1 + \rho)(1 - P_0) = L - (1 - P_0)$；

（3）顾客在系统里的平均逗留时间 $W = \frac{m}{\mu(1 - P_0)} - \frac{1}{\lambda}$；

（4）顾客在队列里的平均逗留时间 $W_q = W - \frac{1}{\mu}$。

在机器维修问题中，L 是等待检修及正在检修的平均机器数，$m - L = \frac{\mu}{\lambda}(1 - P_0)$ 为正常运行的平均机器数。

例 7.4 某工厂生产车间有五台机床，每台机床的连续运转时间服从负指数分布，一天（8 小时）平均连续运行时间 120 分钟。有一个修理工人，每次的修理时间服从负指数分布，工人平均每次工作 96 分钟。求：

（1）修理工忙的概率；

（2）5 台车床都出故障的概率；

（3）出故障的平均机器台数；

（4）平均停工时间；

（5）平均等待修理时间；

（6）评价这个系统的运行情况。

解：首先要统一单位，我们以一天为一个单位时间，即 8 小时为一个单位；一天内来修理的机器数平均为 4 台，修理工一天平均修理机器数为 5 台；则 $m = 5, \lambda = 4, \rho = \frac{\lambda}{\mu} = 0.8$。

（1）因 $P_0 = \frac{1}{\sum_{i=0}^{m} \frac{m!}{(m-i)!}\rho^i} = 0.003\,7$，则修理工忙的概率为 $1 - P_0 = 0.992\,7$

（2）$P_5 = \frac{5!}{0!}(0.8)^5 P_0 = 0.287$

（3）$L = m - \frac{\mu}{\lambda}(1 - P_0) = 3.76$

（4）$L_q = L - (1 - P_0) = 2.77$

（5）$W = \frac{m}{\mu(1 - P_0)} - \frac{1}{\lambda} = 0.742\,7$（天）$= 356$（分钟）

(6) $W_q = W - \frac{1}{\mu} = 0.542\,7$(天) = 260(分钟)

由计算结果可以看出,系统的修理工几乎没有空闲时间,机器的停工时间是平均运行时间的3倍,系统的效率很低。

二、多服务台模型

在本节中,我们将讨论单个队列,多服务台排队模型。满足如下基本假定:

(1) 输入过程:顾客源是无限的,顾客的到达过程是泊松过程;

(2) 排队规则:单个队列,先到先服务;

(3) 服务机构:多服务台,各服务台工作相互独立,服务时间服从负指数分布;服务时间和顾客到达的时间间隔相互独立。

和单服务台模型系统一样,我们分成以下三种情况:

(1) 基本模型,即[$M/M/s$]:[$\infty/\infty/FCFS$];

(2) 有限队列模型,即[$M/M/s$]:[$N/\infty/FCFS$];

(3) 有限顾客源模型,即[$M/M/s$]:[$\infty/m/FCFS$]。

(一) 基本模型

基本的多服务台模型与基本的单服务台模型具有相同的假定,另外还假定各服务台工作相互独立且服务速率相同,即$\mu_1 = \mu_2 = \cdots = \mu_s$。整个系统的平均服务速率为$s\mu$。考虑$\rho = \frac{\lambda}{s\mu}$,当$\rho < 1$时,系统不会排成无限的队列,$\rho$为系统的服务强度或者服务机构的平均利用率。

对于多服务台基本模型系统,该系统的服务速率与系统中的顾客数有关。当系统中的顾客数k小于或者等于服务台个数时,即$1 \leqslant k \leqslant s$时,系统中的顾客全都在服务台中,此时系统的服务速率为$k\mu$;当系统中的顾客数k大于服务台个数时,即$k > s$时,服务台中正在接受服务的顾客数为s个,其余顾客在队列中等待服务,此时系统的服务速率为$s\mu$。该系统的状态转移如图7.9所示。

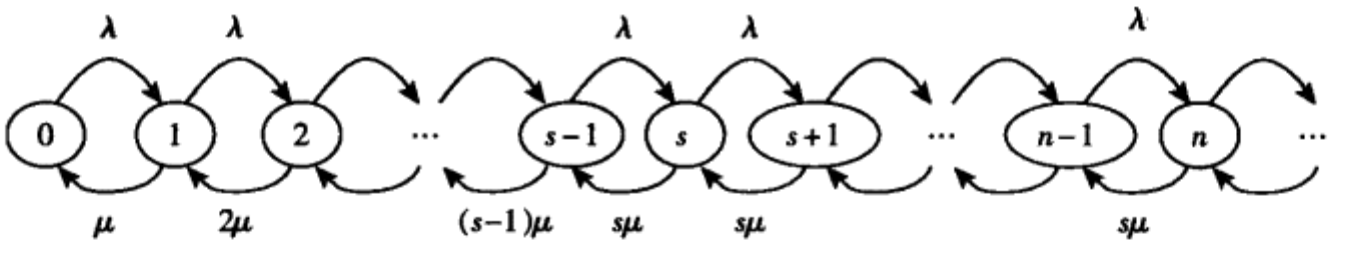


图7.9

根据以上状态转移图,可以得到如下平衡方程组:

$$\begin{cases} \lambda P_0 = \mu P_1 \\ \lambda P_0 + 2\mu P_2 = (\lambda + \mu) P_1 \\ \vdots \\ \lambda P_{s-1} + s\mu P_{s+1} = (\lambda + s\mu) P_s \\ \vdots \\ \lambda P_{n-1} + s\mu P_{n+1} = (\lambda + s\mu) P_n \\ \vdots \end{cases}$$

由 $\sum_{n=1}^{\infty} P_n = 1$，我们可以得到：

$$P_0 = [(\sum_{n=0}^{s-1} \frac{\lambda^n}{\mu^n n!}) + \frac{1}{s!}(\frac{\lambda}{\mu})^s (\frac{1}{1-\rho})]^{-1}$$

$$P_n = \begin{cases} \frac{\lambda^n}{\mu^n n!} P_0 & 1 \leqslant n \leqslant s \\ \frac{\lambda^n}{\mu^n s! s^{n-s}} P_0 & n > s \end{cases}$$

多服务台基本模型的数量指标为：

$$L_q = \frac{\lambda^s \rho P_0}{\mu^s s! (1-\rho)^2} \quad L = L_q + \frac{\lambda}{\mu} \quad W = \frac{L}{\lambda} \quad W_q = \frac{L_q}{\lambda}$$

例 7.5 某银行有三个窗口办理个人储蓄业务，顾客到达服从泊松过程，到达速率为 0.9 人／分钟，办理业务时间服从负指数分布，每个窗口的平均服务速率为 0.4 人／分钟。顾客到达后取得一个排队号，依次由空闲窗口按号码顺序办理储蓄业务。求：

（1）所有窗口都空闲的概率；

（2）平均队长；

（3）平均等待时间及逗留时间；

（4）顾客到达后必须等待的概率。

解：这是一个 $[M/M/s]:[\infty/\infty/FCFS]$ 系统，$\lambda/\mu = 2.25$，$\rho = \lambda/s\mu = 0.75$。

（1）所有窗口都空闲的概率：

$$P_0 = [(\sum_{n=0}^{s-1} \frac{\lambda^n}{\mu^n n!}) + \frac{1}{s!}(\frac{\lambda}{\mu})^s (\frac{1}{1-\rho})]^{-1} = 0.074\ 8$$

（2）平均队长：由于 $L_q = \frac{\lambda^s \rho P_0}{\mu^s s! (1-\rho)^2} = 1.70$，所以 $L = L_q + \frac{\lambda}{\mu} = 3.95$

（3）平均等待时间 $W_q = \frac{L_q}{\lambda} = 1.89$（分钟），平均逗留时间 $W = \frac{L}{\lambda} = 4.39$（分钟）

（4）顾客到达后必须等待的概率，即为 $n \geqslant 3$ 的概率：

$$P\{n \geqslant 3\} = \frac{(2.25)^3}{3!(1-0.75)} \times 0.074\ 8 = 0.57$$

（二）有限队列模型

当队列的容量从无限值变为有限值 $N(N \geqslant s)$ 时，基本模型就转化为有限队列模

型。当系统中的顾客数 $n \leqslant N$ 时，到达的顾客进入系统；当 $n > N$ 时，到达的顾客就被拒绝。如果顾客到达的速率为 λ，每个服务台服务的速率为 μ，$\rho = \lambda/s\mu$。有限队列模型系统的状态转移图如图 7.10 所示。

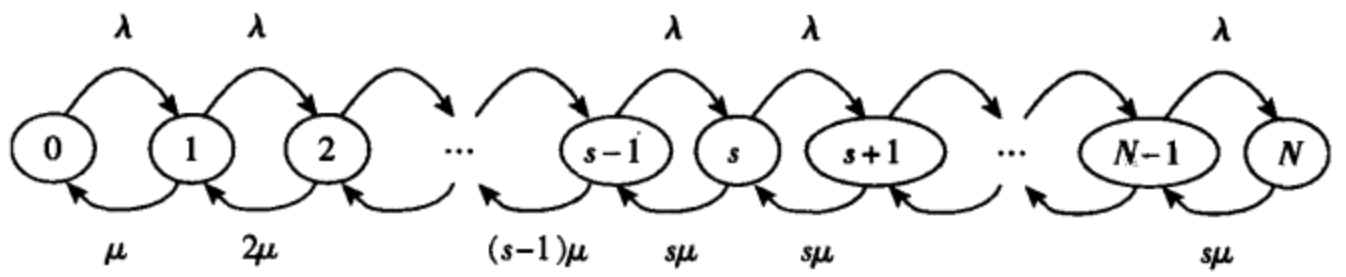


图 7.10

根据以上状态转移图，可以得到如下平衡方程组：

$$\begin{cases} \lambda P_0 = \mu P_1 \\ \lambda P_0 + 2\mu P_2 = (\lambda + \mu) P_1 \\ \lambda P_{s-2} + s\mu P_s = [\lambda + (s-1)\mu] P_{s-1} \\ \lambda P_{s-1} + \mu P_{s+1} = (\lambda + s\mu) P_s \\ \lambda P_s + s\mu P_{s+2} = (\lambda + s\mu) P_{s+1} \\ \vdots \\ \lambda P_{N-1} = s\mu P_N \end{cases}$$

由 $\sum_{n=1}^{\infty} P_n = 1$，我们可以得到：

$$P_0 = \left[\sum_{k=0}^{s} \frac{(s\rho)^k}{k!} + \frac{s^s}{s!} \cdot \frac{\rho(\rho^s - \rho^N)}{1-\rho} \right]^{-1} \quad (\rho \neq 1)$$

当 $\rho = 1$ 时，P_0 的表达式为：

$$P_n = \begin{cases} \dfrac{(s\rho)^n}{n!} P_0 & (0 \leqslant n \leqslant s) \\ \dfrac{s^s \rho^n}{s!} P_0 & (s < n \leqslant N) \end{cases}$$

多服务台有限队列模型的数量指标为：

$$L_q = \frac{\rho(s\rho)^s}{s!(1-\rho)^2}[1 - \rho^{N-s} - (N-s)\rho^N - s(1-\rho)]P_0$$

$$L = L_q + s\rho(1 - P_N) \quad W_q = \frac{L_q}{\lambda(1-P_N)} \quad W = W_q + \frac{1}{\mu}$$

例 7.6 某旅馆有 10 个床位，旅客到达服从泊松过程，平均速率为 6 人／天，旅客平均逗留时间为 2 天，求旅客客满的概率。

解：这是一个 $[M/M/10]:[10/\infty/FCFS]$ 系统，其中 $N = s = 10$，$\lambda = 6$，$\frac{1}{\mu} = 2$，$\mu = 0.5$，$s\rho = \frac{\lambda}{\mu} = 12$，$\frac{\rho(\rho^s - \rho^N)}{1-\rho} = 0$，所以：

$$P_0 = \left[\sum_{k=0}^{s} \frac{(s\rho)^k}{k!} + \frac{s^s}{s!} \cdot \frac{\rho(\rho^s - \rho^N)}{1-\rho} \right]^{-1} = \left[\sum_{k=0}^{s} \frac{(s\rho)^k}{k!} \right]^{-1} = 0.001\,8$$

$$P_{10} = \frac{(s\rho)^N}{N!}P_0 = 0.3019$$

则旅馆10个床位全满的概率为0.3019。

（三）有限顾客源模型

设顾客源为有限数m，服务台个数为s，并且$m > s$。与单服务台系统一样，顾客到达率是按每个顾客考虑的，为λ；系统中的顾客数为n，而且服务台各自独立；服务时间服从参数为μ的负指数分布。这个模型的典型例子是机器维修问题，此时机器数量为m台，修理工人数量为s人，每个顾客的到达率λ是每台机器在单位运行时间内发生故障的期望次数，当正常运行的机器数为m时，发生故障的机器数为$m\lambda$；系统中的顾客数n就是发生故障的机器数，当$n \leqslant s$时，所有发生故障的机器都在修理，有$s - n$个修理工空闲；当$s < n \leqslant m$时，有$n - s$台机器在停机等待修理，修理工均处于繁忙状态；同时假定这s个修理工的技术相同，修理时间服从参数为μ的负指数分布，并且故障的修复时间和正在生产的机器是否发生故障是相互独立的。

由于多服务台有限顾客源模型状态转移图、状态概率和数量指标的推导较为繁琐，我们在这里省略状态转移图，直接给出状态概率和数量指标的结果。

$$P_0 = \frac{1}{m!} \cdot \frac{1}{\sum_{k=0}^{s} \frac{1}{k!(m-k)!}\left(\frac{s\rho}{m}\right)^k + \frac{s^s}{s!}\sum_{k=s+1}^{m} \frac{1}{(m-k)!}\left(\frac{\rho}{m}\right)^k}$$，其中，$\rho = \frac{m\lambda}{s\mu}$

$$P_n = \begin{cases} \frac{m!}{(m-n)!}\left(\frac{\lambda}{\mu}\right)^n P_0 & 0 \leqslant n \leqslant s \\ \frac{m!}{(m-n)!s!s^{n-s}}\left(\frac{\lambda}{\mu}\right)^n P_0 & s+1 \leqslant n \leqslant m \end{cases}$$

$$L = \sum_{n=1}^{m} nP_n, L_q = \sum_{n=s+1}^{m} (n-s)P_n$$

$L = L_q + \frac{\lambda_e}{\mu}$，其中，$\lambda_e = \lambda(m - L)$为单位时间内出现故障的机器数，可以看作有效到达速率。

$$W = \frac{L}{\lambda_e}, W_q = \frac{L_q}{\lambda_e}$$

例7.7 某车间有5台机器，每台机器的故障率为1次／小时，有两个修理工负责这5台机器，工作效率相同，为4台／小时。求：

（1）等待修理的平均机器数；

（2）等待修理及正在修理的平均机器数；

（3）每小时发生故障的平均机器数；

（4）平均等待修理的时间；

（5）平均停工时间。

解：这是一个[*M/M/2*]:[∞ */5/ FCFS*] 模型，其中$m = 5, \lambda = 1, \mu = 4, s = 2$，$\rho = \frac{m\lambda}{s\mu} = \frac{5}{8}$，因而：

$$P_0 = \frac{1}{m!} \cdot \frac{1}{\sum_{k=0}^{s} \frac{1}{k!(m-k)!}(\frac{s\rho}{m})^k + \frac{s^s}{s!}\sum_{k=s+1}^{m} \frac{1}{(m-k)!}(\frac{\rho}{m})^k} = 0.314\,9$$

$$P_1 = \frac{5!}{(5-1)!1!}(\frac{1}{4})P_0 = 0.394, P_2 = \frac{5!}{(5-2)!2!}(\frac{1}{4})^2 P_0 = 0.197,$$

$$P_3 = \frac{5!}{(5-3)!3!3^{5-2}}(\frac{1}{4})^3 P_0 = 0.074, P_4 = 0.018, P_5 = 0.002,$$

$$(1) L_q = \sum_{n=s+1}^{m}(n-s)P_n = P_3 + 2P_4 + 3P_5 = 0.118$$

$$(2) L = \sum_{n=1}^{m} nP_n = P_1 + 2P_2 + 3P_3 + 4P_4 + 5P_5 = 1.092$$

$$(3) \lambda_e = \lambda(m - L) = 3.908$$

$$(4) W_q = \frac{L_q}{\lambda_e} = 0.03 \text{ 小时} = 1.8(\text{分钟})$$

$$(5) W = \frac{L}{\lambda_e} = 0.28 \text{ 小时} = 16.8(\text{分钟})$$

第四节　其他服务时间分布模型

一、任意服务时间分布模型

这种模型表示为[$M/G/1$]:[$\infty/\infty/FCFS$],表示单服务台到达、任意服务时间的排队模型,该排队系统满足:

(1) 输入过程。顾客源是无限的,顾客的到达过程是泊松过程。

(2) 排队规则。单个队列,对队长无限制,先到先服务。

(3) 服务机构。单服务台,G 表示服务时间的分布为任意的概率分布,但已知期望和方差。

仍设 λ 为平均到达率,μ 为平均服务率,则平均服务时间为$\frac{1}{\mu}$,假设服务时间的标准差为 σ,则任意服务时间分布模型的数量指标为:

(1) 系统里没有顾客的概率 $P_0 = 1 - \frac{\lambda}{\mu}$;

(2) 队列中的平均顾客数 $L_q = \frac{\lambda^2\sigma^2 + (\lambda/\mu)^2}{2(1-\lambda/\mu)}$;

(3) 系统的平均顾客数 $L = L_q + \frac{\lambda}{\mu}$;

(4) 顾客在队列里的平均逗留时间 $W_q = \frac{L_q}{\lambda}$;

(5) 顾客在系统里的平均逗留时间 $W = W_q + \frac{1}{\mu}$。

例 7.8 某杂货店只有一名售货员，已知顾客到达过程服从泊松过程，平均到达率为每小时 20 人。不清楚这个系统的服务时间服从什么分布，但已知售货员平均服务一名顾客的时间为 2 分钟，服务时间的标准差为 1.5，求这个排队系统的各数量指标。

解：这是一个[*M/G/*1]:[∞/∞/ *FCFS*] 排队系统，其中 $\lambda = \frac{1}{3}, \mu = \frac{1}{2}, \sigma = 1.5$，则：

(1) 系统里没有顾客的概率 $P_0 = 1 - \frac{\lambda}{\mu} = 0.333$；

(2) 队列中的平均顾客数人 $L_q = \frac{\lambda^2\sigma^2 + (\lambda/\mu)^2}{2(1 - \lambda/\mu)} = 1.041$(人)；

(3) 系统的平均顾客数人 $L = L_q + \frac{\lambda}{\mu} = 1.708$(人)；

(4) 顾客在队列里的平均逗留时间 $W_q = \frac{L_q}{\lambda} = 3.124$(分钟)；

(5) 顾客在系统里的平均逗留时间 $W = W_q + \frac{1}{\mu} = 5.124$(分钟)。

二、定长服务时间模型

这个模型是[*M/G/*1]:[∞/∞/ *FCFS*] 的一种特殊情况，在这个模型中，服务时间是个常量，我们把这个模型表示成为[*M/D/*1]:[∞/∞/ *FCFS*]，表示单服务台泊松到达、定长服务时间的排队模型。这种模型在一些自动控制的生产设备和装配线常常出现。因为服务时间为常量，只需要将[*M/G/*1]:[∞/∞/ *FCFS*] 模型中的表准差改为 0，就可以得到定长排队模型的各项数量指标。

例 7.9 某汽车冲洗站有一套自动冲洗设备，冲洗每辆汽车所需时间为 6 分钟，到此冲洗站来的汽车的到达时间服从泊松过程，每小时平均到达 6 辆。求这个排队系统的各数量指标。

解：这是一个[*M/D/*1]:[∞/∞/ *FCFS*] 模型，其中 $\lambda = 6, \mu = 60/6 = 10, P_0 = 1 - \frac{\lambda}{\mu} = 0.4, L_q = \frac{(\lambda/\mu)^2}{2(1 - \lambda/\mu)} = 0.45, L = L_q + \frac{\lambda}{\mu} = 1.05, W_q = \frac{L_q}{\lambda} = 0.075, W = W_q + \frac{1}{\mu} = 0.175$。

三、爱尔朗分布模型

这种模型表示为[$M/E_K/1$]:[∞/∞/ *FCFS*]。每一个顾客依次经过 K 个服务台，接受 K 次服务，每个服务台的服务时间 T_i 服从相同的负指数分布，则总的服务时间 $T = \sum_{i=1}^{K} T_i$ 服从 K 阶爱尔朗分布。该模型的数量指标为：

$$P_0 = 1 - \frac{\lambda}{\mu}, \rho = \frac{\lambda}{\mu}, W_q = \frac{(K+1)\rho}{2K\mu(1-\rho)} = \frac{\lambda}{\mu(\mu-\lambda)} - \frac{(K-1)\lambda}{2K\mu(\mu-\lambda)}$$

$$L_q = \lambda W_q = \frac{(K+1)\rho^2}{2K(1-\rho)}, L = L_q + \frac{\lambda}{\mu} = \frac{(K+1)\rho^2}{2K(1-\rho)} + \rho, W = W_q + \frac{1}{\mu}$$

例 7.10 一个质量检验员平均每小时收到 2 件送来检验的样品，每件样品要依次完成 5 项检验才能判断是否合格。根据统计，每项检验所需时间的期望值都是 4 分钟，每项检验的时间和送检产品的到达时间间隔都服从负指数分布。求检验过程的各项数量指标。

解：这个检验系统是一个$[M/E_K/1]:[\infty/\infty/FCFS]$排队系统，其中，$\lambda = \frac{2}{60} = \frac{1}{30}, \frac{1}{\mu} = 20, \rho = \frac{2}{3}$，所以：

$$P_0 = 1 - \frac{\lambda}{\mu} = \frac{1}{3}, L_q = \frac{(K+1)\rho^2}{2K(1-\rho)} = \frac{4}{5}, L = L_q + \frac{\lambda}{\mu} = \frac{22}{15}$$

$$W_q = \frac{L_q}{\lambda} = 24(\text{分钟}), W = W_q + \frac{1}{\mu} = 44(\text{分钟})$$

第五节 排队系统的经济分析

从经济角度考虑，排队系统的费用包含两个：一个是服务费用，它是服务水平的递增函数；另一个是顾客等待的机会损失，它是服务水平的递减函数。所以在一般情形下，提高服务水平可减少顾客的等待费用，但会增加服务机构的成本。因而系统优化的目标是使两者的总费用之和为最小，并确定达到最优时的服务水平。这实际上是一个排队系统的优化问题，这类问题经常分为两类：一类称为静态最优设计，目的在于使设备达到最大效益，也就是说，在保证一定服务质量指标的前提下，要求服务机构最为经济；另一类称为动态最优控制，是考虑对于一个排队系统如何营运可以使得目标函数最优。

总的来说，排队系统常见的经济分析问题有：

（1）确定最优服务率μ^*；

（2）确定最佳服务台数s^*；

（3）选择最为合适的服务规则；

（4）确定上述几个量的最佳组合。

在本节里，考虑到有些问题需要较多的数学知识，这里我们只介绍简单模型当中最优服务率和最佳服务台数的确定方法，而不涉及动态排队最优控制问题。

一、确定最优服务率

假定顾客到达服从参数为 λ 的泊松分布，服务台的平均服务率为 μ，服从负指数分布；设 C_1 为单位时间内与 μ 值大小有关的费用，C_2 为每名顾客单位等待（含服务）时间的费用；用 $TC(\mu)$ 表示给定 μ 值时顾客等待和服务设施的费用之和，则有：

$$TC(\mu) = C_1\mu + C_2 L$$

对于$[M/M/1]:[\infty/\infty/FCFS]$排队系统，$TC(\mu) = C_1\mu + C_2\frac{\lambda}{\mu-\lambda}$，因而确定

最优服务率：

$$\mu^* = \lambda + \sqrt{\frac{C_2\lambda}{C_1}}$$

例 7.11 某地欲新建一座港口码头，但只有一个装卸船只的位置，装卸能力用每天装卸的船只数表示。已知单位装卸能力每天平均生产成本为 2 000 元；船只到港后若不及时装卸，则停留一天损失运输费 1 500 元。预计船只的平均到达率为 3 只／天。设船只到达的时间间隔和装卸时间都服从负指数分布。问港口装卸能力为多大时，每天的总支出最少？

解：这是一个典型的[*M/M/*1]：[∞／∞／ *FCFS*] 排队系统的经济优化问题。其中，C_1 = 2 000 元／天，C_2 = 1 500 元／天，λ = 3 只／天。

$$\mu^* = \lambda + \sqrt{\frac{C_2\lambda}{C_1}} = 4.5(\text{只／天})$$

所以，最优装卸能力为 4.5 只／天。

二、确定最佳服务台数

若用 C_3 表示每增加一个服务台系统单位时间开支的费用，$L(s)$ 表示 s 个服务台时系统中的平均顾客数，$TC(s)$ 为给定 s 值时，系统中顾客等待费用与服务台开支费用的总和。于是对于[*M/M/s*]：[∞／∞／ *FCFS*] 排队系统有：

$$TC(s) = C_2L(s) + C_3s$$

由于在这个式子中 s、$L(s)$ 等是离散值，无法用求导数的方法求得最优解，故可以用比较法确定。我们可以将不同的 s 值代入上式，当计算得到 $TC(s-1) \geqslant TC(s)$ 和 $TC(s+1) \geqslant TC(s)$ 成立时，此时的 s 值即是所求的最优解。如果应用 $L(s)$ 来计算，s 的最优值则应该满足：

$$L(s-1) - L(s) \geqslant \frac{C_3}{C_2} \geqslant L(s) - L(s+1)$$

例 7.12 某车间有一个工具维修部，要求维修的工具按泊松过程到达，平均每小时 17.5 件。维修部工人每人每小时平均维修 10 件，维修时间服从负指数分布。已知每名工人每小时工资为 6 元，因工具维修使机器停产的损失为每台每小时 30 元，试确定该维修部的最佳工人数。

解：这是一个[*M/M/s*]：[∞／∞／ *FCFS*] 排队系统，C_2 = 30，C_3 = 6，故 C_3/C_2 = 0.2。分别计算不同 s 值时的 $L(s)$ 值，并计算 $L(s-1) - L(s)$ 的值，如表 7.1 所示。

表 7.1

s	$L(s)$	$L(s-1) - L(s)$
1	7.467	—
2	2.217	∞
3	1.842	5.25
4	1.769	0.375
5	1.754	0.073
6		0.015

因为 $L(4) - L(5) = 0.073 < 0.2 < 0.375 = L(3) - L(4)$，即该工具维修部最佳应配备 4 名工人。

第六节　排队系统的 WinQSB 软件应用

排队论的 WinQSB 软件计算都要调用子程序 Queuing Analysis，该程序具有各种排队模型的求解与性能分析、灵明度分析、服务能力分析和成本分析等功能。

首先要建立一个新问题，系统会显示如图 7.11 的选项，系统默认时间的单位为小时。输入数据的格式有两种：一种是简单排队系统（Simple M/M System），此时顾客到达的时间和服务时间服从负指数分布，系统会显示如图 7.12 所示的数据输入格式；另一种是一般当选择此种服务系统（General Queuing System）时，系统显示如图 7.13 的数据输入格式，图 7.13 中的服务时间和到达时间分布系统默认为负指数分布，若要改变分布，双击空格系统显示如图 7.14 所示的分布选项，具体含义可参见表 7.2。

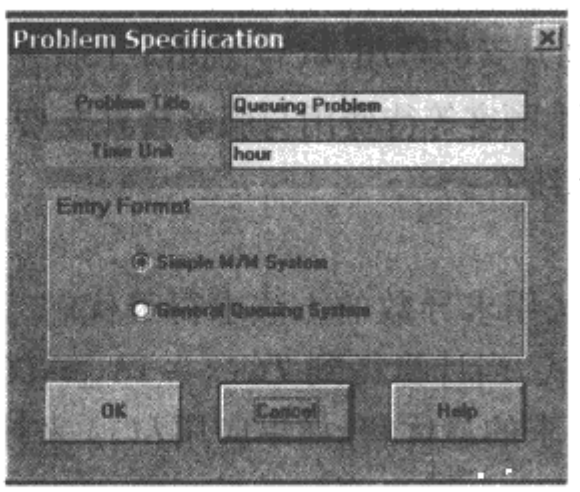


图 7.11

Data Description	ENTRY
Number of servers	服务台数
Service rate (per server per hour)	平均服务率
Customer arrival rate (per hour)	平均到达率
Queue capacity (maximum waiting space)	队列容量
Customer population	顾客源
Busy server cost per hour	每小时忙的成本
Idle server cost per hour	每小时空闲的成本
Customer waiting cost per hour	每小时顾客等待成本
Customer being served cost per hour	每小时顾客服务成本
Cost of customer being balked	损失顾客的成本
Unit queue capacity cost	单位队列容量成本

图 7.12

Data Description	ENTRY
Number of servers	服务台数
Service time distribution (in hour)	Exponential 服务时间分布，双击选择
Location parameter (a)	输入第1个参数
Scale parameter (b>0) (b=mean if a=0)	输入第2个参数
(Not used)	输入第3个参数（如果有）
Service pressure coefficient	服务强度系数
Interarrival time distribution (in hour)	Exponential 到达时间分布，双击选择
Location parameter (a)	输入第1个参数
Scale parameter (b>0) (b=mean if a=0)	输入第2个参数
(Not used)	输入第3个参数（如果有）
Arrival discourage coefficient	到达阻尼系数
Batch (bulk) size distribution	Geberal/Arvitrary，批量分布，双击选择
Constant value	1
(Not used)	
(Not used)	
Queue capacity (maximum waiting space)	对列容量，无限为M
Customer population	顾客源，无限为M
Busy server cost per hour	单位时间忙时的成本
Idle server cost per hour	单位时间空闲的成本
Customer waiting cost per hour	单位时间顾客等待的成本
Customer being served cost per hour	单位时间服务顾客的成本
Cost of customer being balked	损失顾客的成本
Unit queue capacity cost	队列容量单位成本

图 7.13

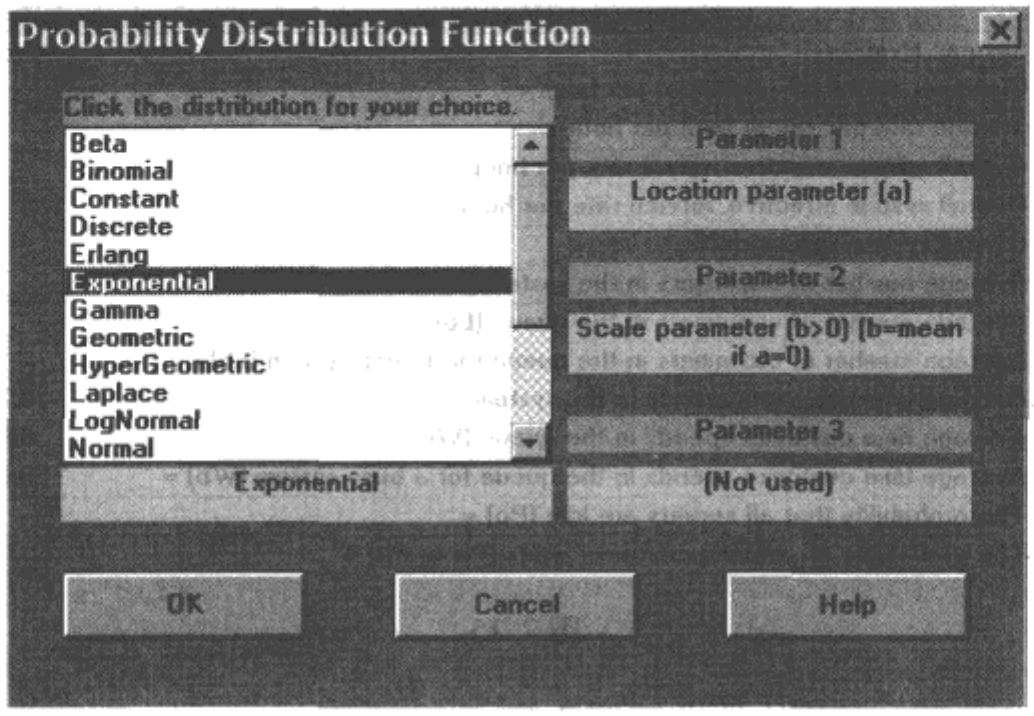


图 7.14

表 7.2

Beta	贝塔分布	Binomial	二项分布
Constant	常数	Discrete	离散分布
Erlang	爱尔朗分布	Exponential	指数分布
Gamma	伽玛分布	Geometric	几何分布
Hyper Geometric	超几何分布	Laplace	拉普拉斯分布
Log Normal	对数正态分布	Normal	正态分布
Pareto	帕累托分布	Poisson	泊松分布
Power Function	功效函数	Triangular	三角分布
Uniform	一致分布	Weibull	威布尔分布
General/Arbitrary	一般分布 / 任意分布		

下面,我们举例来说明怎样用 WinQSB 软件求解排队问题。

例 7.13　用 WinQSB 软件求解例 7.4。

解:顾客到达时间和服务时间都服从负指数分布,在图 7.11 中选择简单排队系统(Simple M/M System),在图 7.12 中输入数据,图 7.14。求解得到与例 7.4 相同的结果,如图 7.15 所示。

Data Description	ENTRY
Number of servers	1
Service rate (per server per hour)	5
Customer arrival rate (per hour)	4
Queue capacity (maximum waiting space)	M
Customer population	5
Busy server cost per hour	
Idle server cost per hour	
Customer waiting cost per hour	
Customer being served cost per hour	
Cost of customer being balked	
Unit queue capacity cost	

图 7.14

03-19-2006	Performance Measure	Result
1	System: M/M/1//5	From Formula
2	Customer arrival rate (lambda) per hour =	4.0000
3	Service rate per server (mu) per hour =	5.0000
4	Overall system effective arrival rate per hour =	4.9635
5	Overall system effective service rate per hour =	4.9635
6	Overall system utilization =	99.2701 %
7	Average number of customers in the system (L) =	3.7591
8	Average number of customers in the queue (Lq) =	2.7664
9	Average number of customers in the queue for a busy system (Lb) =	2.7868
10	Average time customer spends in the system (W) =	0.7574 hours
11	Average time customer spends in the queue (Wq) =	0.5574 hours
12	Average time customer spends in the queue for a busy system (Wb) =	0.5615 hours
13	The probability that all servers are idle (Po) =	0.7300 %
14	The probability an arriving customer waits (Pw or Pb) =	99.2700 %

图 7.15

例 7.14　用 WinQSB 软件求解例 7.9。

解:由于 λ = 6 辆/小时,且到达时间服从负指数分布,故 $1/\lambda$ = 0.166 66,μ = 60/6 = 10 辆/小时,服务时间分布为常数,$1/\mu$ = 0.1,在图 7.11 中选择一般服务系统(General Queuing System),输入数据上述数据,如图 7.16 所示。求解结果如图 7.17 所示。

Data Description	ENTRY
Number of servers	1
Service time distribution (in hour)	Constant
Constant value	0.1
(Not used)	
(Not used)	
Service pressure coefficient	
Interarrival time distribution (in hour)	Exponential
Location parameter (a)	0
Scale parameter (b>0) (b=mean if a=0)	0.16666

图 7.16

03-19-2006	Performance Measure	Result
1	System: M/D/1	From Formula
2	Customer arrival rate (lambda) per hour =	6.0002
3	Service rate per server (mu) per hour =	10.0000
4	Overall system effective arrival rate per hour =	6.0002
5	Overall system effective service rate per hour =	6.0002
6	Overall system utilization =	60.0024 %
7	Average number of customers in the system (L) =	1.0501
8	Average number of customers in the queue (Lq) =	0.4501
9	Average number of customers in the queue for a busy system (Lb) =	0.7501
10	Average time customer spends in the system (W) =	0.1750 hours
11	Average time customer spends in the queue (Wq) =	0.0750 hours
12	Average time customer spends in the queue for a busy system (Wb) =	0.1250 hours
13	The probability that all servers are idle (Po) =	39.9976 %
14	The probability an arriving customer waits (Pw or Pb) =	60.0024 %

图 7.17

在图 7.16 和图 7.17 中,有三项指标解释如下:

P_b 或 P_w 为系统忙的概率。具体含义为系统所有服务台都在服务的概率,或系统的顾客大于或等于服务台数,或某个顾客到来时系统已有 s 个顾客的概率,或者顾客到达系统时需要等待的概率 $P_b = 1 - P_0$。

L_b 为系统忙时队列中顾客的平均数，与 L_q 的关系为 $L_q = L_b P_b$。

W_b 为系统忙时顾客在队列中等待的平均时间，与 W_q 的关系为 $W_q = W_b P_b$。

习题

1. 判断下列说法是否正确：

（1）若到达排队系统的顾客源为泊松过程，则依次到达的两名旅客之间的间隔时间服从负指数分布。

（2）假如到达排队系统的顾客来自两个方面，分别服从泊松分布，则这两部分顾客合起来的顾客流仍为泊松分布。

（3）若两两顾客一次到达的时间间隔服从负指数分布，又将顾客按到达先后顺序排序，则第 1，3，5，7，… 名顾客到达的间隔时间也服从负指数分布。

（4）对 *M/M/*1 或 *M/M/C* 的排队系统，服务完毕离开系统的顾客流也为泊松流。

（5）在排队系统中，一般假定对顾客服务时间的分布为负指数分布，这是因为通过对大量实际系统的统计研究，这样的假定比较合理。

（6）一个排队系统中，不管顾客到达和服务时间的情况如何，只要运行足够长的时间后，系统将进入稳定状态。

（7）排队系统中，顾客等待时间的分布不受排队服务规则的影响。

（8）在顾客到达及机构服务时间的分布相同的情况下，对容量有限的排队系统，顾客的平均等待时间将少于允许队长无限的系统。

（9）在顾客到达的分布相同的情况下，顾客的平均等待时间同服务时间分布的方差大小有关，当服务时间分布的方差越大时，顾客的平均等待时间将越长。

（10）在机器发生故障的概率及工人修复一台机器的时间分布不变的情况下，由一名工人看管5台机器，或由3名工人联合看管15台机器时，机器因故障等待工人维修的平均时间不变。

2. 指出下列排队系统中的顾客和服务员。

（1）机场起飞的客机。

（2）十字路口红灯前的车辆。

（3）超级市场收款台前的顾客。

（4）高速公路收费站。

（5）汽车加油站。

（6）邮局。

3. 某市消费者协会一年365天接受顾客对产品质量的申诉。设申诉以 $\lambda = 4$ 件／天的泊松流到达，该协会处理申诉的定额为5件／天，当天处理不完的将移交专门小组处理，不影响每天业务。试求：（1）一年内有多少天无申诉；（2）一年内有多少天处理不完当天的申诉。

4. 某地区的人口出生数服从 $1/\lambda = 2$ 小时的负指数分布，试计算：（1）某一天内无婴儿诞生的概率；（2）某天内恰好诞生20名婴儿的概率；（3）在4小时内诞生5名

婴儿的概率;(4) 在4小时内诞生5名以上婴儿的概率。

5. 来到一个汽车加油站加油的汽车服从泊松分布,平均每五分钟到达一辆。设加油站对每辆汽车的加油时间为十分钟,问在这段时间内发生以下情况的概率:(1) 没有一辆汽车到达;(2) 有两辆汽车到达;(3) 不少于五辆汽车到达。

6. 假定一个排队系统中有两个服务员,到达该系统服从泊松分布,平均1人/小时;又每个服务员对顾客的服务时间服从负指数分布,平均1人/小时。如有一名顾客于中午12点到达该排队系统的情况下,试求:(1) 下一名顾客在下午一点前、一点到两点之间、两点以后到达的概率为多大?(2) 假定下午一点前没有别的顾客到达,则下一名顾客于一点到两点间到达的概率为多大?(3) 在一点到两点间顾客到达数分别为0、1和不少于2的概率是多少?(4) 假定两个服务员于下午一点都正在为顾客服务,则两个被服务的顾客于下午两点前,一点十分前,一点零一分前都没有被服务完的概率分别为多大?

7. 一个有2名服务员的排队系统,该系统最多容纳4名顾客。当系统处于稳定状态时,系统中恰好有n名顾客的概率为:$P_0=\frac{1}{16},P_1=\frac{4}{16},P_2=\frac{6}{16},P_3=\frac{4}{16},P_4=\frac{1}{16}$。试求:(1) 系统中的平均顾客数$L$;(2) 系统中平均排队的顾客数$L_q$;(3) 某一时刻正在被服务的顾客的平均数;(4) 顾客的平均到达率为2人/小时,求顾客在系统中的平均逗留时间W;(5) 若2名服务员具有相同的服务率,利用(4) 的结果求服务员为一名顾客服务的平均时间$1/\mu$。

8. 某排队系统有两名服务员,到达该系统的顾客流服从泊松分布,平均1人/小时,又每个服务员服务一名顾客的时间平均都为1小时,服从负指数分布。求:(1) 计算该系统的平均队长;(2) 计算服务员的平均忙期。

9. 两个各有一名理发员的理发店,每个店内只能容纳4名顾客。两个理发店具有相同顾客到达率$\lambda=10$人/小时,服从泊松分布。当店内顾客满员时,新来顾客自动离去。已知第一个理发店内,对顾客服务时间平均每人15分钟,收费11元;第二个理发店内,对顾客服务时间平均每人10分钟,收费7.5元。以上时间服从负指数分布。若两个理发店每天均营业12小时,问哪个店内的理发员收入高一些。

10. 某个只有一名服务员的排队系统,其等待空间总共可容纳3名顾客,即系统中总顾客数不允许超过4人。已知对每名顾客的服务时间服从负指数分布,平均为5分钟,顾客到达服从泊松分布,平均10人/小时。试求:(1) 系统中顾客数n分别为0,1,2,3,4的概率;(2) 系统中顾客的平均数;(3) 一名新到的顾客需要排队等候服务的概率;(4) 一名新到的顾客未能进入该排队系统的概率。

11. 某企业有5台运货车辆,已知每台车每运行100小时平均需维修2次,每次需时20分钟,以上分别服从泊松和负指数分布,求该企业内部车辆正常运行的概率,及分别有1,2,3辆车不能正常运行时的概率。

12. 某工具室有k名工人,到达工具室要求得到服务的顾客流为$1/\lambda=1.5$(分钟) 的负指数分布,每名工人对顾客的服务时间为$1/\mu=0.8$(分钟) 的负指数分布。假如工具室工人费用为9(元/小时),生产工人(顾客) 等待损失为18(元/小时),试决定工具室工人的最佳数量k。

第八章　风险决策

引例:期刊订购量决策

某书亭销售一种月刊,每份购进价为4元,正常售价为5元,当月售不出的月刊可以2元的价格处理。书亭每月底要确定下一个月订购这种月刊的数量。据统计,过去50个月中,售出100份、110份、120份、130份、140份的分别是5、10、5、20、10个月。试问,该书亭每月底应订购多少份这种月刊,其月利润才会最大?

第--节　先验概率的风险决策

"你要的是白猫,抓出来的却是黑猫。"大多数管理者都讨厌风险,但风险问题又无处不在。如企业想开发一种新产品,新产品可能畅销,也可能滞销。畅销时企业可获得丰厚的利润,但滞销时则会造成亏损,企业是否决定开发这种新产品呢?又如海底钻探石油,可能出油,也可能不出油,你是否在此开钻呢?再如建筑项目的招投标,投标需要一笔较大的准备费用,但投标后可能中标,也可能不中标,中标后可获得一笔可观的利润,不中标则造成投标准备费的损失,你是否决定去投标呢?如此等等。风险的存在是因为社会经济现象中存在大量的随机现象。对这些随机现象进行决策时就面临着风险。

一、预备知识

风险决策是在概率论和数理统计的基础上来研究如何在风险决策问题中合理选择方案的一种决策方法。故需要用到概率论和数理统计的一些基本概念。

(一)基本概念

随机现象。所谓随机现象是指在一定的条件下,某种现象可能出现这种结果,也可能出现另一种结果,事前谁也不能肯定会出现某一结果的现象。如掷一枚均匀的骰子,可能出现1、2、3、4、5、6点六种可能结果,事前谁也不能肯定会出现某一点,则掷骰子是一个随机现象;又如建筑项目的招投标,可能中标,也可能不中标,则建筑项目的招投标也是一个随机现象。

随机事件。它是指随机现象在一次实验中,可能发生也可能不发生的结果。如建筑项目的招投标,中标是一个结果,不中标也是一个结果,也就是说建筑项目的招投标这个随机现象有两个随机事件。决策中常将随机事件称为自然状况,一般用 S_i 表示。

概率。某一随机事件发生的可能性大小,称为概率。概率可反映随机事件的规律

性。在决策中，一般用 $P(S_i)$ 表示某一自然状况发生的概率。

频率。在 n 次重复的随机实验中，某一随机事件发生了 m 次，则比率 m/n 称为这个随机事件发生的频率。如某个建筑公司，过去进行了20次投标，6次中标，14次未中标，则中标的频率为 $P(S_1) = 6/20 = 0.3$，未中标的频率为 $P(S_2) = 14/20 = 0.7$。在大量重复实验中，频率会趋近于概率，故实际决策问题中，常用频率来代替概率，也称为统计概率。

客观概率。根据有明确的历史资料或历史先例（推理）得到的概率，称为客观概率。如根据统计资料得到的频率，是根据历史资料得到的；52张牌中抽出一张是红桃的概率为1/4，是基于推理得到的。

主观概率。根据个人的经验和知识估计出某一事件发生的概率，称为主观概率。如某个销售人员，估计某种新产品畅销的可能性为0.4，滞销的可能性为0.6，就是根据他的经验来估计的。在风险决策中，由于有时会得不到客观概率，则经常会利用主观概率来进行决策。

先验概率。没有得到补充信息前分配给某随机事件的概率。

后验概率。在得到补充信息后，对先验概率修正后分配给某随机事件的概率。如一个仆人被主人陷害被判处死刑，根据英国一个古老的规定，死囚在被处死前可以再进行一次生死选择。即在一个装有生牌、死牌的袋中抽出一牌来定生死。主人为置仆人于必死境地，将两牌都换成死牌，仆人巧妙地利用这一补充信息，从而将“生”的先验概率0.5变为后验概率的1。

随机变量。某些随机现象的结果可用数量来表示，但这些数值的取值不能事先确定，受随机因素的影响而变化，则称这种受随机因素的影响而变化的随机现象的取值，称为随机变量。如电话总机在一分钟内接到的呼唤次数，可能是0，1，2，…

离散变量。如果随机变量在其取值区间只能取有限个或无穷可数个的值，这样的随机变量称为离散型随机变量，简称离散变量或非连续变量。如电话总机在一分钟内接到的呼唤次数是离散变量。

连续变量。如果随机变量在其取值区间能取任何数值，这样的随机变量称为连续型随机变量，简称连续变量。如某天的降雨量是一个连续变量。

（二）风险决策问题应具备的条件

风险决策问题一般要满足以下几个条件：

（1）存在决策者希望达到的目标（收益最大或损失最小）；

（2）有两个及以上可供选择的方案（常用 a_j 来表示）；

（3）有两个及以上的自然状况（随机事件，常用 S_i 表示）；

（4）能估计或计算各自然状况的概率 $P(S_i)$；

（5）能计算不同方案在不同自然状况下的损益值（V_{ij}）。

当一个决策问题满足以上五个条件时，则可判断该问题属于风险决策问题。当不满足第四个条件时，则称为完全风险决策。

二、离散变量先验概率的风险决策

在引例中，某书亭确定每月底应购进多少份这种月刊目标是希望其利润最大。每

月能按正常售价售出的份数可能是100份、110份、120份、130份或140份，故销售份数是离散随机变量，意味着该问题有五种自然状况。每种自然状况发生的概率可用统计资料计算的频率来代替，且该概率是未获得任何补充信息时的先验概率。该例有五种可供选择的决策方案，即每月底购进月刊的份数可以是100份、110份、120份、130份或140份。根据所给资料，可计算出不同方案在不同自然状况下的损益值。可见，该问题是一个离散变量先验概率的风险决策问题。

（一）选择方案的标准

所谓决策，就是在多种方案中选择好的方案。而什么是好的方案，取决于选择方案的标准。不同的标准，意味着好方案的不同含义，在某种标准下的好方案，在另一种标准下则可能不是好的方案。不同的标准可能选择出不同的方案，而每一种标准都有其适用条件和各自的优缺点，故在实际决策中选用什么标准来选择方案是非常重要的，通常也是困难的。

在风险决策问题中，选择方案的标准一般有四种：一是期望值标准，这是风险决策中常用的一种选择方案的标准，具体又可分为期望收益（费用）值标准、期望后悔值标准和边际分析标准；二是最大概率标准；三是满意标准；四是效用标准。这些标准的含义、适用条件和优缺点将在后面的运用中加以介绍。

（二）期望值标准

1. 期望收益（费用）值标准

期望值标准中的一种表现形式是期望收益（费用）值标准，这种标准是先求出每种方案在各种自然状况下的条件收益（如利润、产值等）值，然后将条件收益值乘以先验概率得到每种方案的期望收益值，期望收益值最大的方案就是这种标准下最好的方案。如果决策问题是以费用（如成本、管理费等）最小为目标，则先求出每种方案在各种自然状况下的条件费用值，然后将条件费用值乘以先验概率得到每种方案的期望费用值，期望费用值最小的方案就是这种标准下最好的方案。在期刊订购量决策中，是希望利润最大，则该例的条件及期望利润值如表8.1所示。

表8.1　　单位：元

需求量 S_i（份）	概率 $P(S_i)$	订购量（份）									
		100		110		120		130		140	
		条件利润	期望利润	条件利润	期望利润	条件利润	期望利润	条件利润	期望利润	条件利润	期望利润
100	0.1	100	10	80	8	60	6	40	4	20	2
110	0.2	100	20	110	22	90	18	70	14	50	10
120	0.1	100	10	110	11	120	12	100	10	80	8
130	0.4	100	40	110	44	120	48	130	52	110	44
140	0.2	100	20	110	22	120	24	130	26	140	28
合计	1.0		100		107		108		106		92

从表8.1中可看出，按期望利润值标准，每月初订购120份的期望利润为108元，高于订购其他份数的期望利润，故每月初应订购120份，其平均利润最高。

2. 期望后悔值标准

俗话讲“有钱难买后悔药”,但人们又经常因为未达到理想的要求而后悔。如果这种后悔的感觉能用数量值(货币)来表示,则这种数量值称为后悔值。

在商业行为中,通常会产生两种损失:一种是由于供大于求而产生的实际损失;另一种是由于供不应求而产生的机会损失。无论哪种损失都会使人后悔。

期望后悔值标准就是从损失的角度出发来选择方案的一种方法。这种标准是先求出每种方案在各种自然状况下的条件后悔值,然后将条件后悔值乘以先验概率得到每种方案的期望后悔值,期望后悔值最小的方案就是这种标准下最好的方案。期刊订购量决策中的条件及期望后悔值如表 8.2 所示。

在表 8.2 中,当在自然状况与方案数相同的情况下,条件后悔值表为对角线为零的方阵,零的上方的数值为实际损失,零的下方的数值为机会损失。期望后悔值标准是从损失的角度来分析问题,故期望后悔值最小的方案就是最佳方案。本例中订购量为 120 份的期望后悔值最小(16 元),故月初订购量为 120 份。与用期望收益值标准得到的方案是一致的。这种标准可得到一般被人忽视的机会损失。

表 8.2　　单位:元

需求量 S_i(份)	概率 $P(S_i)$	订购量(份) 100		110		120		130		140	
		条件后悔值	期望后悔值	条件后悔值	期望后悔值	条件后悔值	期望后悔值	条件后悔值	期望后悔值	条件后悔值	期望后悔值
100	0.1	0	0	20	2	40	4	60	6	80	8
110	0.2	10	2	0	0	20	4	40	8	60	12
120	0.1	20	2	10	1	0	0	20	2	40	4
130	0.4	30	12	20	8	10	4	0	0	20	8
140	0.2	40	8	30	6	20	4	10	2	0	0
合计	1.0		24		17		16		18		32

3. 边际分析标准

在上面的期望利润值和期望后悔值标准中,均是利用决策矩阵表的形式来进行的。但当自然状况或决策方案很多时(有时甚至为无穷多),编制决策矩阵表就很复杂(甚至不可能)。这时,利用西方经济学中的“边际原理”将大大简化决策的计算和分析过程。

通俗地讲,“边际”就是最后一个的意思。边际费用是指增加最后一个单位产品所需要增加的费用;边际收入是指增加最后一个单位产品的生产并顺利售出所得到的收入增量。在期刊订购量中,边际利润是指增加一个单位(这里的一个单位是广义的,可表示一批的意思。本例为论述方便,以 10 本期刊为一个单位)的订购量并顺利售出所获得的利润,一般用 MP 表示。本例中,$MP = 50 - 40 = 10$(元)。边际损失指增加一个单位的订购量而售不出去而造成的损失,一般用 ML 表示。本例中,$ML = 40 - 20 = 20$(元)。

经济学证明,当边际费用等于边际收入时的生产量是最优的,这时的利润最大。在风险决策问题中,可证明当期望边际费用等于期望边际收入时的生产量是最优的。

边际分析方法作为风险决策问题选择方案的标准,首先是可以简化计算;其次,边际分析法在选择方案时是从"停止"的角度出发的。有人说,假定在一千级阶梯中有一级是最好的,与其去寻找最优的,不如去寻找应该在哪一级"止步"。期刊订购量决策的期望边际利润和期望边际损失如表 8.3 所示。

表 8.3

需求量(十本)	能卖出的边际概率(P)	卖不出的边际概率($1-P$)	边际利润(元) MP	边际损失(元) ML	期望边际利润(元) EMP	期望边际损失(元) EMP
10	1.0	0	10	20	10	0
11	0.9	0.1	10	20	9	2
12	0.7	0.3	10	20	7	6
13	0.6	0.4	10	20	6	8
14	0.2	0.8	10	20	2	16

边际原理表明,当期望边际利润大于期望边际损失时($EMP > EML$),应当增加订购量;当期望边际利润小于期望边际损失时($EMP < EML$),应减少订购量;当 $EMP = EML$ 时,应在此停止。本例中,当订购量为 10、11、12 时,$EMP > EML$;而当订购量为 13 时发生了转折,即 $EMP < EML$(且 $|7-6| = 1 < |6-8| = 2$),这时给出一个信号,即应该在订购量为 12 时止步,再增加订购量将是危险的。

实际运用这个标准时,并不需要将每种需求量下的期望边际利润和期望边际损失都计算出来加以比较,只需要计算一个特定的概率(决策所需的最低概率)P' 与能卖出的边际概率相比较,则可选出所需的方案。

由 $EMP = EML$

有:$P(MP) = (1-P)(ML)$

则:$P(MP) = ML - P(ML)$ $\quad P(MP + ML) = ML$

得:$P' = ML / MP + ML$

本例中,$P' = 20/20 + 10 = 0.67$,该概率意味着,订购量一个单位一个单位的增加,当增加到最后一个单位时,能卖出的概率不能低于 0.67。用计算出来的这个特定的概率与能卖掉的边际概率相比较,当某一边际概率正好包含这个特定概率所对应的需求量时就是最优的订购量。本例中,需求量为 12 时能卖出的边际概率为 0.7,正好包含 0.67,故最优订购量为 12。

4. 完全(完善)信息期望值($EVPI$)

在风险决策中,有人认为之所以存在风险是因为没有获得全部的信息,如果能获得全部信息则是在确定条件下进行决策,这时将不存在风险。

完全信息期望值是获得完全(全部百分之百的准确可靠)的补充信息的最高价值。完全信息期望值是完全信息条件下的期望值($EVWPI$)与先验概率下最优方案的期望值(EV)之差。

在收益情况下：$EVPI = EVWPI - EV$

在费用情况下：$EVPI = EV - EVWPI$

完全信息期望值的大小，反映了决策者对信息的了解程度。在风险决策中，任何信息都不可能是完全的，但*EVPI*给出了购买信息价值的最高限额。在决策中，必须研究信息的经济性，并把购买信息的决策纳入总的决策中去。

在期刊的最优进货量决策中，由于需求的随机性，某月的需求量可能是100份，也可能是110份、120份、130份、140份，故存在风险。现若有一市场分析家他说能准确地预测到某一月的需求量，即他预测某一月的需求量是130份，则购进130份期刊到月底会全部顺利售出，一份不多一份不少，供求正好相等。若果真如此，经销商就是在确定条件下进行决策，不会产生任何风险。但市场分析家认为他提供了信息应得到相应的报酬。经销商为此付给该信息的最高报酬是多少呢？

假定市场分析家提供的信息是完全的，这时的期望利润应是多少呢？

$EVWPI = 100 \times 0.1 + 110 \times 0.2 + 120 \times 0.1 + 130 \times 0.4 + 140 \times 0.2 = 124$(元)

$EV = 108$(元)

$EVPI = EVWPI - EV = 124 - 108 = 16$(元)

计算表明，经销商若能获得完全的信息，平均每月的期望利润为124元，比在不完全信息(先验概率)下可多获利16元。同时也表明，支付信息的最高限额为16元。

海底石油钻探决策：某一专门从事海底石油钻探的公司，据统计，钻出石油的概率为0.4，钻不出石油的概率为0.6。若钻出石油将获利100万元，若钻不出石油将亏损50万元。按期望利润值标准，该公司是选择开钻还是不开钻呢？其决策矩阵如表8.4所示。

表8.4

自然状况 S_i	先验概率 $P(S_i)$	钻(a_1)		不钻(a_2)	
		V_{ij}(万元)	EV(万元)	V_{ij}(万元)	EV(万元)
出油	0.4	100	40	0	0
不出油	0.6	-50	-30	0	0
合计	1.0		10		0

从表8.4中可看出，在先验概率条件下，应选择钻，这时的期望利润为10万元。

现若有一海底石油勘探公司建议先勘探后再决定钻与不钻。为此石油钻探公司愿意支付给勘探公司的最高报酬应是多少呢？

$EVWPI = 100 \times 0.4 + 0 \times 0.6 = 40$(万元)，$EV = 10$(万元)

$EVPI = 40 - 10 = 30$(万元)

即该决策问题的完全信息期望值为30万元。

可见，不同的决策问题的*EVPI*是不相同的，有时相差甚大，是否值得购买补充信息，需从实际问题出发，不能一概而论。

(三) 最大概率标准

最大概率标准就是将概率最大的自然状况所对应的方案作为中选方案的一种选择方案的标准。

在期刊最优进货量决策中，需求量为130份的概率为0.4，比其他需求量发生的概率都大，按以需定产的方针，则选中订购量为130份的方案。而用期望值标准选中的方案是120份。

最大概率决策标准，实际上是把概率最大的可能状况当成是必然事件来对待，不考虑还有其他状况发生的可能性。因此，利用这种标准来选择方案时，要求某种可能状况发生的概率比其他可能状况发生的概率要大得多，这时运用这种标准可能是简单和实用的。如乘公共汽车，实际上是风险决策，但因为出事故的概率相当小（如为千分之一），而不出事故的概率很大（如千分之九百九十九）。人们将不出事故当成是必然的现象，因此人们在乘车时才会蜂拥而上。若不出事故的概率为0.5，可能乘车时就会是另一番景象。

在运用最大概率标准时，应记住这样的警句："若某件事情发生的概率低于0.5，无论前景多么诱人，也不应当去冒这个风险。"

（四）满意标准

满意标准是决策学的创始人西蒙提出的一种选择方案的标准。西蒙认为，在选择方案时应考虑以下的一些情况：第一，有些决策问题的备选方案非常之多，甚至无穷多，以至于不可能将所有方案一一列举出来。这时只可能在列举出来的方案中选优，如果这种情况发生，就可能出现最优方案本身就未在列举方案之中，选中的方案何谈是最优方案呢?第二，要考虑决策的成本问题。决策的每一方案的设计，构思要花费大量的人力、财力和物力，使得决策者有时基于成本上的考虑，只设计若干备选方案进行选择，这时也可能将最优方案给遗漏了。第三，决策所追求的最优是在实施阶段的最优，而不是在选择阶段的理论上的最优。因为决策是人做出的，也要人来执行，一个理论上的最优方案，如果在实施中得不到人们的配合，有可能成为糟糕的方案。而不断地追求满意的结果，有可能达到在实施中的最佳效果。

如人们在购买衣服时，同种样式、颜色的衣服有一千件，假如服务员允许你在这一千件中任意挑选，理论上讲，其中只有一件是最好的。但人们一般是只挑选几件后，觉得满意就购买了。

在运用满意标准时，没有铁定的金科玉律，要根据实际问题正确确定满意的概念及程度。

（五）效用标准

1. 期望值标准的缺点

效用标准实质上是满意标准的一种表现形式，但由于在风险决策中是一种用得较多的选择方案的标准，且在技术处理上有其独特的地方，故一般均将它单独作为一种选择方案的标准加以介绍。

期望值标准是风险决策中最常用的一种选择方案的标准，但这种标准只适合于重复决策的情况，而对于以下的情况则不适合用期望值标准来选择方案。

第一，人们常常对期望值标准提出批评，就在于期望值是一个统计平均数，它描述了事物的一般趋势，而忽视了一切不同的、特殊的和个别的情况。如一位病人躺在手术台上显得战战兢兢，尽管医生告诉他这种手术的成功率为98%，但病人似乎对这个成功率持怀疑态度。因为这个成功率对个别病人而言，并没有任何作用。对个别

病人而言，手术要么成功，要么失败，并不存在有 98% 的成功，2% 的失败。若这 2% 的失败概率落在这个病人身上则是 100% 的失败。

第二，当风险决策中的某一可能结果将给决策者带来严重后果和极端威胁时，决策者一般是不会按期望值标准来选择方案的。如一个小偷偷了一匹马，按照一种古老的私法，将会在第二天太阳升起之前被砍头。但马的主人提出若小偷能在一周内用诚实的劳动挣得 10 美金作为罚金，则将免去死刑。该小偷在一周内辛苦劳动挣到 10 美金后，在准备送给马主人的路上碰到往日的同伙，同伙建议用掷硬币的方法进行一次赌博，若正面向上，小偷将赢得 100 块美金，若反面向上，小偷只输 10 块美金。若按期望值标准，这是对小偷有利的赌局，但这一次，小偷连头都不回的走掉了。

第三，在相同的期望值下，由于决策者不同的气质，不同的性格及所处的环境、条件、时间等不相同时，决策者将作出不同的选择。如“朝三暮四”是《庄子·齐物论》中的一个寓言，现意是形容做事三心二意。其原意是齐国有一个人养了一只猕猴，主人每天给猴子七个饼子，但其分配的形式不一样而猴子的反应也大不相同，即“朝四暮三则喜，朝三暮四则怒”。据心理学家分析，猴子的这种表现，反映出猴子有一种只顾眼前利益，不顾长远利益的倾向。又如，同样数量的资金，穷人和富人的感觉会很不一样，如此等等。

第四，某些决策问题不能仅从经济利益出发时（如某些公共产品的供给），也不宜用期望货币值作标准。

综上所述，当决策是一次性决策时，当决策的某种后果对决策者构成极端威胁时，当决策者所处环境、条件、时间不同及决策者具有不同性格时，当决策问题不能仅从经济利益考虑时，决策者往往不是根据期望值标准来选择方案的。

2. 效用的概念及度量

效用是西方经济学中的一个概念，本义是指商品满足人们欲望程度的一种主观量度。效用概念引入风险决策中是指决策者对风险的主观态度。如某些决策者生来就是冒险家，而某些决策者则生来就胆小怕事，不喜欢冒风险。

效用一般可用三种尺度加以衡量：

（1）名义尺度（分类尺度）

这种尺度是用一种对立状态的形容词或动词对效用进行表述。如满意、不满意；拒绝、接受；合格、不合格；得力、不得力等等。这种尺度适宜于一些难以用数量表现的事物，如人的工作态度、产品装潢设计以及产品质量等。

（2）比较尺度（序数尺度）

这种尺度是用相对的顺序或排列来表述效用的等次。这种尺度在多目标决策中经常使用，如购买汽车，需要考虑汽车的售价、耗油量、速度、维修费用、舒适性等，这些目标在决策者心目中的重要程度不一样，则可按重要程度加以排序（如 1、2、3、4、5）。

（3）计数尺度（区间尺度）

这种尺度对于许多有程度变化的事物，如光亮、声音、时间、温度等，则可根据程度的变化截取一定数量区间来衡量程度的变化。如在用期望货币值作风险决策标准时，可用效用值与货币值相结合得到能计数的货币效用值。这种尺度在风险决策中是经常运用到的一种效用决策分析标准。

3. 常见的效用曲线的类型

把效用与货币值结合得到的货币效用值，可对决策者的效用作出定量的描述并绘制成效用曲线。

不同的决策者对同一货币的主观效用可能大不一样；同一决策者在不同的时间、环境下对同一货币的主观效用也可能不一样，因此就会出现多种多样的效用曲线。但一般可分为冒险型、客观型和保守型三种效用曲线（如图 8.1 所示）。

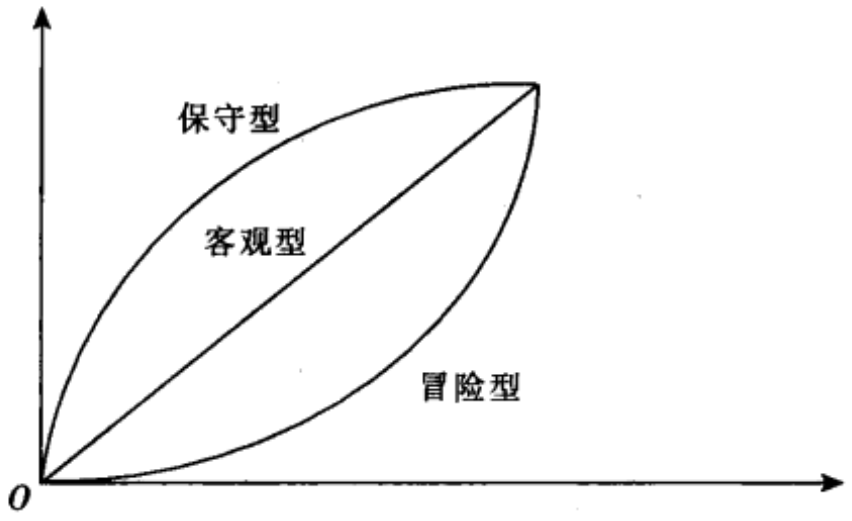


图 8.1

图 8.1 是三种基本的效用曲线，由这三种基本形状可衍生出许多其他形状的效用曲线。

4. 单目标效用决策

例 8.1 新产品开发决策

时新服装厂是一个资金不雄厚的乡镇企业，现工厂准备推出一种新式服装，工厂面临大量生产还是小量生产两种可供选择的方案，其期望利润值如表 8.5 所示。

表 8 - 5

可能状况 S_i	概率 $P(S_i)$	大量生产（$a1$）		小量生产（$a2$）	
		条件利润（万元）	期望利润（万元）	条件利润（万元）	期望利润（万元）
畅销	0.6	10	6	4	2.4
滞销	0.4	-3	-1.2	1	0.4
合计	1.0		4.8		2.8

从表 8.5 中可看出，若按期望利润值标准，应选大量生产的方案。但该厂近来经济状况不佳，故厂长对大量生产时有 40% 的可能造成 3 万元的亏损十分敏感，迟迟不敢对大量生产这一方案拍板。看来该厂长对按期望利润值标准选中的方案持怀疑态度。下面我们将利用效用值标准进行决策。

用效用值标准进行决策的步骤：

首先，用“心理测验法”绘制效用曲线。其做法是：

① 先利用已有资料确定效用曲线的起点、终点坐标。一般令效用值的区间为 (0,1)，用 V 表示货币值，用 $u(V)$ 表示效用值，效用值是货币值的函数，即决策者对货币值的效用。横轴为货币值，纵轴为效用值。该例中，最坏的情况是亏损 3 万元，令

其效用值为0,即 $u(V) = u(-3) = 0$,则起点坐标为(-3,0)。最好的情况是盈利10万元,令其效用值为1,即 $u(V) = u(10) = 1$,则终点坐标为(10,1)。

② 其次,用"心理测验法"确定中间各点坐标。在确定了起始点后,以这两点的纵、横坐标为基础,用"心理测验法"来确定中间各点的坐标,然后连线形成效用曲线来判断决策者的类型。"心理测验法"来确定效用值的方法是:先确定一个随机方案 A_1,如有50%的机会盈利10万元,有50%的机会盈利-3万元。再确定一个确定方案 A_2,即有100%的机会盈利3.5万元(其数值为随机方案的期望值)。然后让决策者在两方案中进行选择,若决策者选 A_2 而不选 A_1,说明 A_2 的效用大于 A_1,也表明该决策者不是按期望值标准在选择方案。这时,变动 A_2 的值,一直到决策者认为两方案等价为止。如当 $A_2 = 0$ 万元时,决策者认为两方案等价,即可认为这时两方案的效用值相等,有:$u(A_1) = 0.5 \times u(10) + 0.5 \times u(-3) = 0.5 \times 1 + 0.5 \times 0 = 0.5 = u(A_2) = u(0)$ 这样就得到第三点的坐标(0,0.5)。再假定有方案 A_3,50%获利10万元,50%获利0万元,方案 A_4 是100%获利5万元,重复上述的询问过程,当 A_4 变动为3万元时,决策者认为 A_3 与 A_4 等价;则有:$u(A_3) = 0.5 \times u(10) + 0.5 \times u(0) = 0.5 \times 1 + 0.5 \times 0.5 = 0.75 = u(A_4) = u(3)$ 这样得到第四点的坐标(3,0.75)。反复重复这样的测验过程,可得到若干点的坐标,在直角坐标系中标出这些点并连线,则可得到该决策者的效用曲线,如图8.2所示。

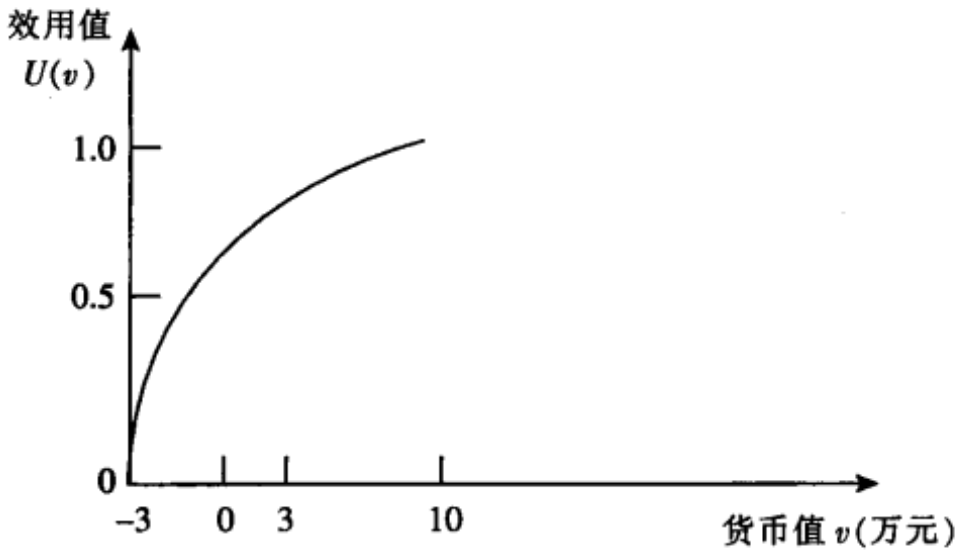


图8.2

从图8.2上看出,该厂长有一保守型的效用曲线。从该曲线可查到1万元的效用值约为0.58,4万元的效用值约为0.82。

最后,利用期望效用值进行决策。

有了决策者的效用曲线图后,则可用期望效用值标准来选择方案,如表8.6所示。

表8.6

可能状况 S_i	概率 $P(S_i)$	大量生产(a1)		小量生产(a2)	
		效用值	期望效用	效用值	期望效用
畅销	0.6	1	0.6	0.82	0.492
滞销	0.4	0	0	0.58	0.232
合计	1.0		0.6		0.724

从表 8.6 中可看到，大量生产的期望效用值为 0.6，小于小量生产的期望效用值 0.724，则按期望效用值标准应选小量生产，而按期望利润值标准应选大量生产（4.8 万元 > 2.8 万元）。可见，不同的选择方案的标准导致选中不同的方案。

5. 多目标效用决策

例 8.2 电冰箱购买决策。

某家庭准备购买一台电冰箱，基本要求是双门直冷式，容积为 170 立升，购价在 1 800 元以下，日耗电量不超过 2 度，发动机保修期 1 年以上，外观要求尽可能美观。现有 *A*、*B*、*C*、*D* 四种牌子满足该家庭的要求，这 4 种牌子均为 170 立升的双门直冷式，但售价、日耗电量、保修期、外观则不尽相同（外观采用百分制直接打分）。具体数据如表 8.7 所示。

表 8.7

冰箱牌子（方案）	购价（元）	日耗电量（度）	保修期（年）	外观（分）
A	1 750	1	1	80
B	1 600	1.2	1.5	65
C	1 650	1.4	2	70
D	1 580	1.5	1	55
限额	1 800 以下	2 以下	1 以上	无限额

该问题是一个多目标决策问题，该家庭认为其重要程度为日耗电量、售价、保修期、外观，其权数分别为 8、7、4、1。为该家庭推荐一综合性能最好的冰箱。在该多目标决策中，我们采用的是比较尺度。

多目标效用决策的步骤：

首先，计算各牌子各目标的效用得分。效用分的计算有多种方法，我们采用下式计算效用分：

有上限限制的目标效用分 = 100 - （目标实际值 / 目标上限值）× 100

有下限限制的目标效用分 = 100 - （目标下限值 / 目标实际值）× 100

对没有限额的目标效用分可采取直接百分制打分（如外观）。

如 *A* 牌子各目标的效用分为：

耗电量效用分 = 100 - (1/2) × 100 = 50（分）

售价效用分 = 100 - (1 750 / 1 800) × 100 = 3（分）

保修期效用分 = 100 - (1 / 1.5) × 100 = 25（分）

外观效用分 = 80（分）

四种冰箱各目标的效用得分如表 8.8 所示。

表 8.8

冰箱牌子	效用得分(分)			
	日耗电量	售价	保修期	外观
A	50	3	25	80
B	40	11	50	65
C	30	8	67	70
D	25	12	25	55

其次,计算加权效用总分进行方案选择。由于各目标在决策者的心目中的重要程度不一样,则以不同的权数来体现,可得加权效用总分,如表 8.9 所示。

表 8.9

冰箱牌子	效用得分(分)				总效用分(分)
	日耗电量	售价	保修期	外观	
A	50 × 8 = 400	3 × 7 = 21	25 × 4 = 100	80	601
B	40 × 8 = 320	11 × 7 = 77	50 × 4 = 200	65	662
C	30 × 8 = 240	8 × 7 = 56	67 × 4 = 268	70	634
D	25 × 8 = 200	12 × 7 = 84	25 × 4 = 100	55	439

从表 8.9 中可以看出,*B* 牌子的综合效用总分最高,故应向该家庭推荐购买 *B* 牌子的冰箱。

三、先验概率连续变量的风险决策

在离散变量的决策时,主要是利用决策矩阵表进行决策。但当可能状况或决策方案无穷多(连续变量)时,显然不可能用决策矩阵表进行决策。特别是当可能状况或决策方案很多时,可用连续变量的分布近似代替不连续变量的分布进行决策,这将大大简化决策的分析过程,减少计算量。

(一) 利用正态分布进行决策

对连续变量进行决策是利用其分布曲线来进行的,连续变量的分布有许多类型,在风险决策问题中,很多随机变量的分布属于 γ 或 β 分布。这里仅以大家熟悉的正态分布为例加以论述(实际运用时,需要对采用的分布类型加以检验)。

例 8.3 公共交通线的开辟决策。

某公共汽车运输公司准备开辟一条新的客运线,开辟新线路每年的成本约为 60 万元,平均每位乘客的票价为 1.5 元,每年营运天数为 365 天,据有经验的司乘人员估计,有 80% 的可能每天的乘客在 1 000 ~ 1 400 人之间。该公司经理若仅从本公司是否盈利出发,他希望一年内若有 90% 的把握盈利,则开辟该线路,否则不开辟。按照这种要求,是否该开辟这条线路呢?

该线路是否开发很显然取决于每天乘客人数(可能状况)的多少,而乘客数显然是一个连续变量(从理论上讲是零到无穷多)。方案是开辟和不开辟。

假定乘客人数近似服从正态分布(需检验),则可根据所给资料求得正态分布的

均值和标准差。

$$\bar{X} = \frac{1\,000 + 1\,400}{2} = 1\,200(\text{人})$$

$$P(1\,200 \pm 200) = P(\bar{X} \pm 1.28\sigma) = 80\%$$

$$\sigma = \frac{200}{1.28} = 156.25 \approx 156$$

每一乘客每天的平均票价为 1.5 元,则每位乘客一年的车费为 $1.5 \times 365 \approx 548$ 元,这意味着每年每损失一位乘客将损失 548 元,称为单位损失。很显然存在着一个每天的乘客数 S_b,使得该公司的收支相抵,正好保本。即 $600\,000 = 548 \times S_b$,则 $S_b = 1\,095$ 人。该点称为盈亏平衡点,即平均每天的乘客数低于 1 095 人时将发生亏损,高于此数将有盈利,如图 8.3 所示。

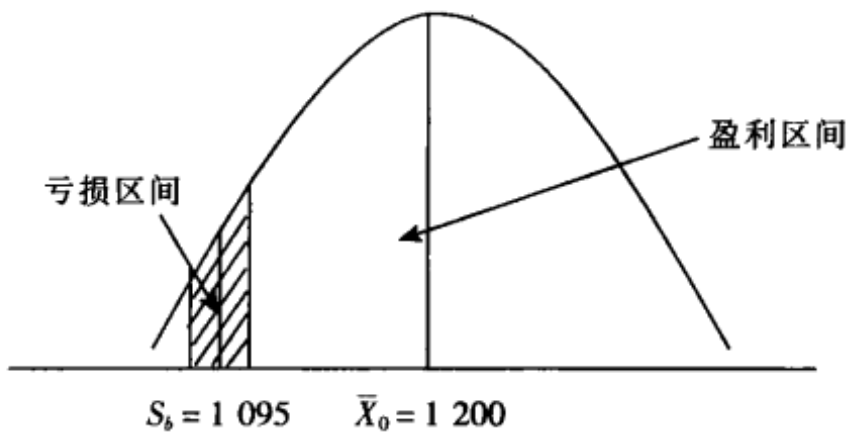


图 8.3

图 8.3 中的阴影部分为发生亏损的区间,阴影部分的右边为盈利区间。发生盈利的概率计算如下:

将 $S_b = 1\,095$ 至 $\bar{X} = 1\,200$ 间的距离标准化,

$$\frac{|S - \bar{X}|}{\sigma^x} = \frac{|1\,095 - 1\,200|}{156} = 0.67$$

查正态分布表得 0.67 个标准差对应的面积为 74.86%,即盈利的把握为 74.86%,低于公交公司经理规定的盈利概率 90%,故不开辟该线路。

第二节　后验概率的风险决策

在风险条件下,好的决策取决于好的概率估计,而良好的概率估计取决于对补充信息的有效利用。如一个补充信息使死囚犯死里逃生的故事为许多人熟知;又如在日常生活和工作中,人们经常利用补充信息来修正其行动。如一个感冒病人并不立即进医院就诊,往往会等上一段时间看病情如何发展再采取行动;一个医生对病情初诊后并不马上开出处方,往往会进行一些检查后,以获得补充信息来支持或否定他的诊断等等。在实际的经济问题中,获得补充信息的手段很多,如市场调查、用户座谈、广告、展销会、专家咨询等。

一、后验概率离散变量的风险决策

(一) 对先验概率的修正

对先验概率的修正是利用概率论中的条件概率公式(贝叶斯公式)来进行的。莎士比亚的喜剧《威尼斯商人》中的一个情节是:少女鲍细霞的婚姻被押在她的亡父为她设计的一个猜匣子的赌博中。亡父留下三个外观完全一样的金、银、铅匣子,其中一个匣子装有女儿的肖像,谁能根据匣子上的一句话猜中装肖像的匣子谁就将成为女儿的丈夫。这里我们做如下修改:金匣子中装有 7 个红球,2 个白球;银匣子装有 1 个红球,3 个白球;铅匣子装有 1 个红球,5 个白球。现找到其中一个匣子,问这个匣子是金匣子的先验概率是多少?

若金、银、铅匣子分别用 A、B、C 来表示。很显然,在未得到任何补充信息时,找到金匣子的先验概率为:

$$P(A) = 1/3 = P(B) = P(C)$$

现若允许从匣子中抽出一球再进行判断,则意味着获取了一个补充信息。假定从匣子中抽出的一球为红球,这时找到金匣子的后验概率又是多少?

设 D 代表抽中红球,$P(D)$ 为抽到红球的概率。在抽到红球的条件下计算是金匣子的后验概率,实际上是用贝叶斯公式求条件概率,即:

$$P(A/D) = \frac{P(AD)}{P(D)} = \frac{P(A) \cdot P(D/A)}{P(D)}$$

$$= \frac{1/3 \times 7/9}{1/3 \times 7/9 + 1/3 \times 1/4 + 1/3 \times 1/6} = \frac{7/27}{43/108} = 0.65$$

即在获得补充信息时,找到金匣子的概率为 0.65。

(二) 二项分布的后验概率决策

例 8.4 质检决策。

某工厂的产品 100 件装一箱。若装箱前不对产品进行质量检查,当顾客发现不合格品时必须更换或修理,每件产品的更换或修理的平均费用为 12 元。若装箱前对产品进行质量检查,每件产品的质检费为 1 元。据过去统计,产品不合格品率为 5% 的概率为 0.5,产品不合格品率为 10% 的概率为 0.4,产品不合格品率为 15% 的概率为 0.1。若该厂仅从每箱产品的质检费用的多少出发,是否应在装箱前进行质量检查呢?

该问题中,不合格品率的概率是根据以前的统计确定的,没有其他的补充信息,得到的是先验概率,用期望费用值为标准选择方案,如表 8.10 所示。

表 8.10

不合格品率 (%) S_i	先验概率 $P(S_i)$	先检验($a1$)		先不检验($a2$)	
		条件费用 (元)	期望费用 (元)	条件费用 (元)	期望费用 (元)
5	0.5	100	50	60	30
10	0.4	100	40	120	48
15	0.1	100	10	180	18
合计	1.0		100		96

从表 8.10 中可看到，先检验的期望费用为 100 元，大于先不检验的期望费用 96 元，故按期望费用值标准应选先不检验。

若按期望后悔值标准，如何确定条件后悔值？

先验概率的完全信息期望值的计算。

本例中，选择先检验还是先不检验的依据是根据不合格品率的大小来确定的。显然存在着一个不合格品率(S_b) 使先检验还是先不检验所花费用是一样的，即：

$100 = (100 \times S_b) \times 12$，解之得，$S_b = 0.083\ 3$。

这意味着，如果每箱产品的不合格品率均为 8.33% 时，则选择先检验还是先不检验的费用是一样的。

也就是说，如果能够准确地知道每箱的不合格品率的话，则当 $S > S_b$ 时，则先检验；当 $S < S_b$ 时，则先不检验。这时就意味着获得了完全信息而不存在风险了。

则：$EVWPI = 5 \times 12 \times 0.5 + 100 \times 0.4 + 100 \times 0.1 = 80$(元)

$EVPI = E(V) - EVWPI = 96 - 80 = 16$(元)

即若需获得补充信息，其价值最高为 16 元。

厂长认为，不合格品率的概率是过去的记录，为增加对不合格率的了解，决定装箱前在 100 个产品中抽两个产品进行检验，再决定是先检验装箱，还是不检验装箱。经抽检两个产品，发现有一个不合格品。这时是应该先检验装箱，还是不检验装箱呢(假设抽检两产品的费用为 3 元)？

1. 贝努里过程与二项分布

贝努里过程：

(1) 每次实验只有对立的两种结果，且必然出现其中的一种；

(2) 多次实验中的任何一次实验，出现某一特定结果的概率是稳定的；

(3) 各次实验都是独立的；

(4) 实验次数和结果次数取整数值。

满足以上四个特征的随机实验称为贝努里过程(实验)。

对于满足贝努里过程的随机事件的概率分布服从于二项分布。对于贝努里过程，事件 A 在 n 次实验中出现 k 次的概率为：

$$P_n(k) = C_n^k p^k q^{n-k}$$

2. 对先验概率的修正

在本例中，从一批产品中进行质量抽检是贝努里过程。这里 $n = 2$，令 k 为抽中不合格品的次数，$k = 1$，p 为不合格品率，这里我用 S_i 来代替，$i = 1$、2、3。则合格品率为 $(1 - S_i)$。

本例中，若这两个产品是从不合格品率为 5% 中抽出的，则出现一个不合格品的概率可表示为 $P(k/S_1)$，实质上是计算二项分布的概率，即：

$$P(k/S_1) = \frac{n!}{k!(n-k)!} S_1^k (1 - S_1)^{n-k} = \frac{2!}{1!(2-1)!} \times 0.05^1 \times (1 - 0.005)^{2-1}$$

$$= 0.095$$

$$P(k/S_2) = \frac{n!}{k!(n-k)!} S_2^k (1 - S_2)^{n-k} = 0.18$$

$$P(k/S_3) = \frac{n!}{k!(n-k)!}S_3^k(1-S_3)^{n-k} = 0.255$$

如此,再利用条件概率公式可计算后验概率,如表 8.11 所示。

表 8.11

Si	$P(Si)$	$P(k/Si)$	$P(Si)P(k/Si)$	$P(Si/k)$
0.05	0.5	0.095	0.047 5	0.328
0.1	0.4	0.18	0.072	0.496
0.15	0.1	0.255	0.025 5	0.176
合计	1.0	—	0.145	1.0

在得到抽检两个产品中有一个不合格品的补充信息后,不合格品率为 5%、10%、15% 的后验概率分别为 0.328、0.496、0.176。

3. 利用后验概率进行决策

得到后验概率后,则可利用期望费用进行决策,如表 8.12 所示。

表 8.12

不合格品率(%)S_i	后验概率 $P(Si/k)$	先检验(a_1)		先不检验(a_2)	
		条件费用(元)	期望费用(元)	条件费用(元)	期望费用(元)
5	0.328	100	32.8	60	19.68
10	0.496	100	49.6	120	59.52
15	0.176	100	17.6	180	31.68
合计	1.0	—	100	—	110.88

在后验概率条件下,先检验的期望费用为 100 元,先不检验的期望费用为 110.88 元。应选先检验的方案。

一般而言,当先验概率与后验概率的方案不一致时,应以后验概率选中的方案为准。

(三) 后验完全信息期望值与补充信息价值

在风险条件下,任何信息都不可能是完全的。

后验信息的获取有可能多次获得或多种方法获取,若还需要获取信息,则需要计算后验完全信息期望值。后验完全信息期望值的计算方法与先验完全信息期望值的计算方法完全一样。本例中,后验的 $EVWPI = 5 \times 12 \times 0.328 + 100 \times 0.496 + 100 \times 0.176 = 86.88$(元),则后验的 $EVPI = E(V) - EVWPI = 100 - 86.88 = 13$(元),即还要获得补充信息的话,其费用不能高于 13 元。

补充信息价值 = 先验 $EVPI$ − 后验 $EVPI = 16 - 13 = 3$(元)

意味着这次抽检两个产品获得的补充信息的价值为 3 元。

二、后验概率连续变量的决策

(一) 正态分布下的完全信息期望值($EVPI$)

在公交线路的开辟决策中,其乘客人数是根据有经验的司乘人员估计的。为了对

该线路乘客有更多的了解，该公司经理拟试运行七天。但需花费一笔费用，这笔费用的最高限额是多少呢？

这实际是计算连续变量的完全信息期望值（*EVPI*），在正态分布时，有如下计算公式：

$$EVPI = A \cdot \sigma \cdot N(D)$$

其中：A 为单位损失，本例中为 548 元；

σ 为标准差，本例中为 156 人；

$N(D)$ 为正态损失系数，D 为标准化距离，$D = 0.67$，

查 $N(0.67) = 0.150\ 5$。

则先验的 $EVPI = 548 \times 156 \times 0.150\ 5 = 12\ 866$（元）

即七天的费用不应超过 12 866 元。

（二）后验正态分布参数的确定

公司经理认为七天试运行的费用不会超过 12 866 元，决定试运行以获取补充信息。七天试运行的结果是，平均每天的乘客数是 1 300 人（$\bar{x}$），标准差（S）为 134。这相当于采用抽样调查获得了一个补充信息。经理认为应将司乘人员的先验信息与补充信息获得的乘客人数加以综合后，得到后验信息的乘客人数再进行决策更为妥当。

在正态分布下，一般用 $\bar{X}_0$ 和 σ_0 代表先验正态分布的均值和标准差，用 $\bar{x}$ 和 S 代表抽样正态分布的均值和标准差，用 $\bar{X}_1$ 和 σ_1 代表后验正态分布的均值和标准差。本例中，利用先验正态分布参数与抽样正态分布参数得到后验正态分布参数的一种方法是：

$$\bar{X}_0 = 1\ 200 \quad \sigma_0 = 156 \quad \bar{x} = 1\ 300 \quad S = 134$$

$$\bar{X}_1 = \frac{\bar{X}_0 \cdot \frac{1}{\sigma_0^2} + \bar{x} \cdot \frac{1}{s^2}}{\frac{1}{\sigma_0^2} + \frac{1}{s^2}} \qquad \sigma_1 = 1/\sqrt{1/\sigma_0^2 + 1/s^2}$$

$$\bar{X}_1 = \frac{1\ 200/156^2 + 1\ 300/134^2}{1/156^2 + 1/134^2} = 1\ 258$$

$$\sigma_1 = 1/\sqrt{1/\sigma_0^2 + 1/s^2} = 1/\sqrt{1/156^2 + 1/134^2} = 102$$

（三）利用后验分布进行决策

将 S_b 到 X_1 的距离标准化：$\dfrac{|S_b - \bar{X}_1|}{\sigma_1} = \dfrac{|1\ 095 - 1\ 258|}{102} = 1.6$

查正态分布表1.6下的面积为94.52%，即意味着，在后验概率条件下，公司盈利的把握达到94.52%，应开辟这条公交线路。

第三节　决策树

一、决策树的概念

“千言万语不如一张图。”决策树是决策局面的一种图解表示。顾名思义，这种图

像一颗被砍倒的树，有干有枝。决策树按逻辑关系从左到右展开，而决策则从右到左进行。图 8.4 中的方框为决策结，由方框引出的直线称为决策枝；圆框为概率结，由圆框引出的直线称为概率枝；树的终止结可用三角框表示。

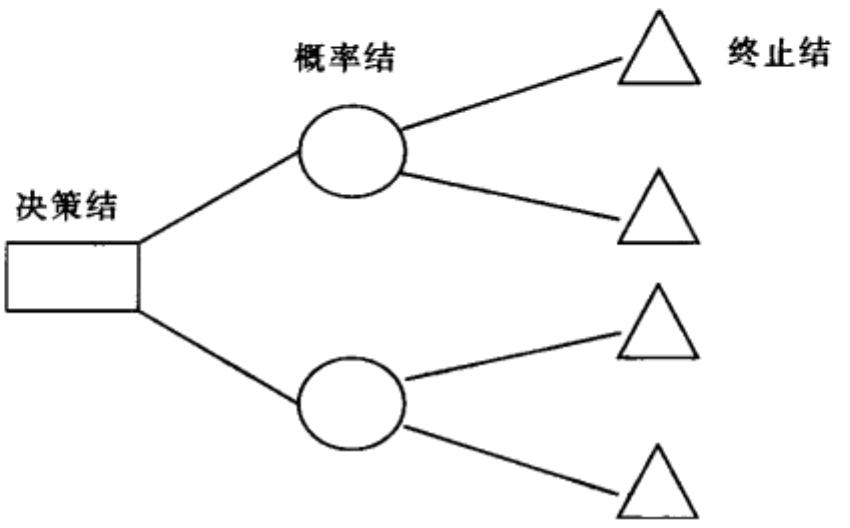


图 8.4

决策树为决策者构造出一个合理而有序的决策过程，并具有形象直观的优点。一般多用于多阶段决策。

二、利用决策树进行决策的一般步骤

（1）确定决策目标及决策阶段数；

（2）按逻辑关系从左到右绘制决策树；

（3）按决策顺序从右到左进行决策。

例 8.5　投标决策。

华光建筑公司正考虑是否为一项建设工程进行投标，投标准备费约为 1 万元，中标的可能性为 40%，不中标的可能性为 60%。若中标后能按期保质完成，则可盈利 40 万元。若延期一月，罚款 5 万；延期二月，罚款 10 万。公司有两种修建方法：一是传统方法，这种方法按期完成的可能性为 60%，延期一月的可能性为 20%，延期两月的可能性为 20%；另一是改进方法，但需花 1.5 万元购置一种设备。改进方法按期完成的可能性为 85%，延期一月的可能性为 10%，延期二月的可能性为 5%。公司是否应该为这项工程进行投标呢？中标后应选择哪一种修建方法呢？

分析：很显然，该决策为两阶段决策，用决策树更能反映决策的有序过程，如图 8.5 所示。

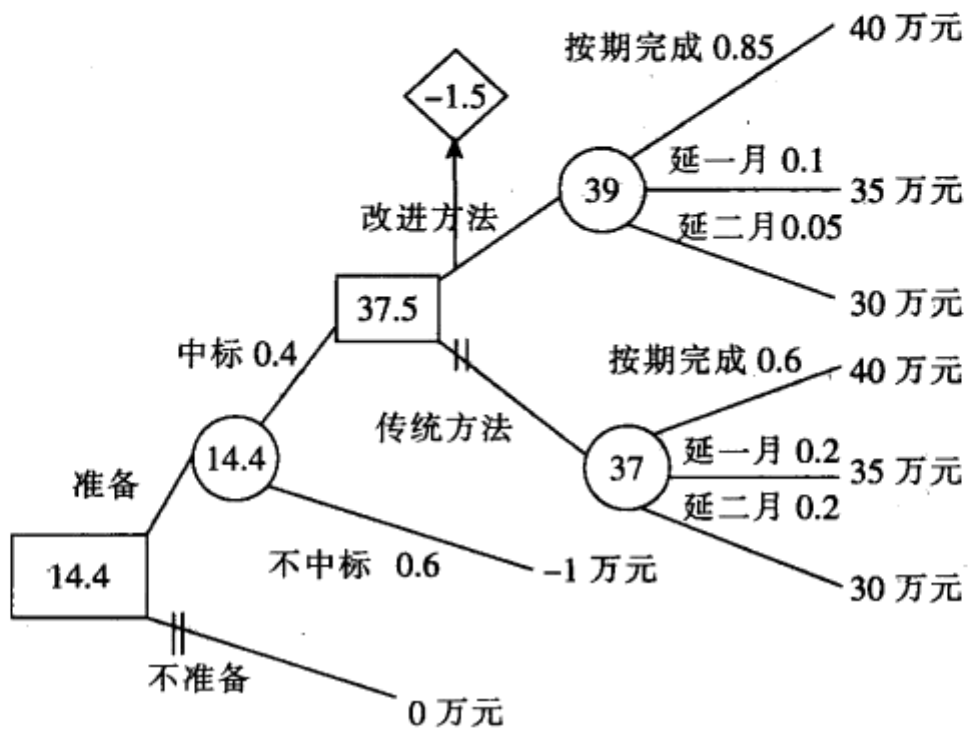


图 8.5

第四节　完全风险模型

一、完全风险决策的概念

在风险决策中，当决策者不能获得自然状况的概率时，这样的决策称为完全风险决策。

二、完全风险决策选择方案的标准

在完全风险决策中，没有一个统一的标准。但有一些可供决策者参考的标准，常用的有：

1. 乐观标准

决策者对未来持乐观态度，认为一切都会好上加好，这种标准又称为“大中大标准”（Maximax）。在收益情况下，这种标准是先选择每个方案各可能状况下的最大值，在这些最大值中再选最大值所对应的方案。

2. 悲观标准

决策者对未来持悲观态度，希望在最坏的情况下使自己的损失最小（盈利最大），又称为“小中大标准”（Minimax）。这种标准是先选择每个方案各可能状况下的最小值，在这些最小值中再选最大值所对应的方案。

3. 现实主义标准

决策者对未来既不能过于乐观，也不宜过分悲观，应采取现实态度，这种标准是估计一个乐观系数 α，则 $1-\alpha$ 为悲观系数，用乐观系数乘以某方案在各可能状况下的最大值加上用悲观系数乘以该方案在各可能状况下的最小值，得现实值。各方案现实值比较来选择方案。

4. 拉普拉斯标准

拉普拉斯从“不充足理由原则”出发,认为各可能状况的概率既然没有充足的理由认为哪一种更大或更小,则就不能否定它们出现的概率是相等的。实际上是给各可能状况以相同的概率来计算各方案的期望值来选择方案。

5. 后悔标准

这种标准是先用某一可能状况下各方案的最大值减各方案的条件值得到后悔值,再选出各方案的最大后悔值,在这些最大后悔值中最小的后悔值对应的方案为该标准所选中的方案(在收益情况下)。

例 8.6 时新服装厂是一个资金不雄厚的乡镇企业,最近准备推出一种新式服装,估计推出后可能出现畅销、一般和滞销三种情况,但不能估计出相应的概率。厂里现有大量生产、中量生产、小量生产和不生产四种意见。四种方案在各可能状况下的盈亏如表 8.13 所示。

表 8.13　　单位:万元

	大量生产 (a_1)	中量生产 (a_2)	小量生产 (a_3)	不生产 (a_4)
畅销	20	15	12	0
一般	10	8	4	0
滞销	-12	-4	-1	0

该厂应该选择哪一种方案呢?

分析:该问题由于没有各可能状况的概率,故是完全风险决策。用前面所讲的标准选择的方案如下:

(1) 乐观标准。先选出每一方案在各可能状况下的最大利润值,即 $a_1 = 20$、$a_2 = 15$、$a_3 = 12$、$a_4 = 0$。然后在这些值中再选最大值,即 $a_1 = 20$,按该标准选中大量生产。

乐观标准又称为最冒险的标准,实际中采用得并不多。

(2) 悲观标准。先选出每一方案在各可能状况下的最小利润值,即 $a_1 = -12$、$a_2 = -4$、$a_3 = -1$、$a_4 = 0$。然后在这些值中再选最大值,即 $a_4 = 0$,按该标准选中不生产。

这种标准往往被称为最保守的标准。是从损失最小的角度出发,虽然在最坏的情况下不会造成大的损失,但往往选中“无所作为”的方案。一般为一些十分保守的决策者和一些经不起打击的小企业所采用。

(3) 现实主义标准。本例中若令乐观系数 $\alpha = 0.4$,则悲观系数 $1 - \alpha = 0.6$。各方案的现实值如下:

a_1 的现实值 $= 20 \times 0.4 + (-12) \times 0.6 = 3.2$(万元)

a_2 的现实值 $= 15 \times 0.4 + (-4) \times 0.6 = 3.6$(万元)

a_3 的现实值 $= 12 \times 0.4 + (-1) \times 0.6 = 4.2$(万元)

a_4 的现实值 $= 0 \times 0.4 + 0 \times 0.6 = 0$(万元)

按这种标准应选 a_3,即小量生产。

这种标准由于较接近实际而经常用到。但确定乐观系数是个难题，有较大的主观随意性。

(4) 拉普拉斯标准。按等概率计算各方案的期望值：

a_1 的期望值 $= 20 \times 1/3 + 10 \times 1/3 + (-12) \times 1/3 = 6$(万元)

a_2 的期望值 $= 15 \times 1/3 + 8 \times 1/3 + (-4) \times 1/3 = 6.33$(万元)

a_3 的期望值 $= 12 \times 1/3 + 4 \times 1/3 + (-1) \times 1/3 = 5$(万元)

a_4 的期望值 $= 0 \times 1/3 + 0 \times 1/3 + 0 \times 1/3 = 0$(万元)

按这种标准应选 a_2，即中量生产。

由不充足理由而推论各可能状况的概率是一样的结论在现实中显得非常勉强，在实际中采用得并不多。

(5) 后悔标准。先计算各方案在各可能状况下的后悔值(表 8.14 中括号内的数值)。

表 8.14　　单位：万元

	大量生产 (a_1)	中量生产 (a_2)	小量生产 (a_3)	不生产 (a_4)
畅销	20 (0)	15(5)	12(8)	0(20)
一般	10 (0)	8(2)	4(6)	0(10)
滞销	-12(12)	-4(4)	-1(1)	0(0)

然后找出各方案的最大后悔值，即 $a_1 = 12$、$a_2 = 5$、$a_3 = 8$、$a_4 = 20$。再在这些最大后悔值中找最小后悔值，即 $a_2 = 5$，则中量生产为该标准下的最优方案。

在完全风险的决策中，由于各可能状况的概率不能确定，以上五种标准并无真正的优劣可言。最终决策往往取决于决策者个人对风险的态度、决策目标的要求、决策单位的条件及环境因素等。这就需要决策者根据具体决策问题的实际，加以灵活选择。

第五节　风险模型的应用

实际工作中的许多决策问题都是具有随机现象特征的风险决策问题，如设备更新、采购成本、投标决策、建厂问题、新产品开发、订购量、生产量等等。这些问题均可用风险模型加以解决。在重复决策时，期望值是一个较合理的选择方案的标准。

例 8.7　设备(零部件) 更新决策。

某化工厂在生产过程中需 500 支温度计来控制液体温度，损坏时必须更新。据统计，温度计在第 1、2、3、4、5、6 月的损坏率分别为 0.05、0.1、0.15、0.2、0.25、0.25。若大批更新时(即到一定时间将全部 500 支温度计全部换新)，每支更新费用为 1 元，而个别更新的费用为 4 元。试为该厂选择一个更新策略，使月平均更新费用最少。该问题应如何考虑呢？

分析：该问题由于温度计的寿命不一样，故是一风险决策问题。可供选择的方案有随坏随换、每一月、每二月、每三月、每四月、每五月、每六月更换七个方案。

采用随坏随换方案时,先计算温度计的平均寿命:

$$\bar{x} = \sum x \frac{f}{\sum f} = 1 \times 0.05 + 2 \times 0.1 + 3 \times 0.15 + 4 \times 0.2 + 5 \times 0.25 \quad +6 \times 0.25$$

$$= 4.25$$

然后再计算每月平均更新支数及每月费用:

月平均更新支数为 500/4.25 = 118(支)

每月更新费用为 118 × 4 = 472(元)

采用定期大批更新方案时,涉及大批更新(固定的 500 支温度计的更新费用)和个别更新两种费用。在计算个别更新费用时,要先计算每个月温度计的损坏数,如表 8.15 所示。

表 8.15

第几月	计算过程	更新数(支)
0		500
1	500 × 0.05	25
2	500 × 0.1 + 25 × 0.05	51
3	500 × 0.15 + 25 × 0.1 + 51 × 0.05	81
4	500 × 0.2 + 25 × 0.15 + 51 × 0.1 + 81 × 0.05	113
5	500 × 0.25 + 25 × 0.2 + 51 × 0.15 + 81 × 0.1 + 113 × 0.05	152
6	500 × 0.25 + 25 × 0.25 + 51 × 0.2 + 81 × 0.15 + 113 × 0.1 + 152 × 0.05	172

得到每个月温度计的损坏数后,则可得到各方案的更新费用,如表8.16 所示。

表 8.16　　单位:元

更新方案	大批更新费用	个别更新费用		总费用	月平均费用
		支数(支)	费用		
随坏随换	—	118	472	472	472
每一月	500	25	100	600	600
每二月	500	76	304	804	402
每三月	500	157	628	1 128	376
每四月	500	270	1 080	1 580	395
每五月	500	422	1 688	2 188	438
每六月	500	594	2 376	2 876	479

可见,采用每三个月整批更新的月平均费用最少,故应每三个月更新一次。

例 8.8　采购成本决策。

某工厂生产某种产品时需要一种化学添加剂,该添加剂有效期仅为一月,故需每月初采购一次。该厂对这种添加剂过去每月耗费的情况是:使用数量(千克)为 0、1、2、3、4 的概率分别为 0.05、0.1、0.4、0.3 和 0.15。生产这种化学添加剂的化工厂有两种购买方式:一是预定,在购买 5 千克及以下时,固定费用 500 元,每千克价格为 100 元;二是临时购买,每千克的价格为 400 元。为该厂推荐一购买策略,使其采购成本最低。

分析:该厂在使用这种添加剂时,由于每月使用的数量是随机的,故为一风险决策。且该厂每月都面临按什么方式及购买多少数量的问题,可用期望值标准选择方案(表 8.17)。

表 8.17

月末订购量(千克)	采购成本期望值(元)	
0	400 × 0.1 + 800 × 0.4 + 1 200 × 0.3 + 1 600 × 0.15	= 960
1	500 + 100 + 400 × 0.4 + 800 × 0.3 + 1 200 × 0.15	= 1 180
2	500 + 200 + 400 × 0.3 + 800 × 0.15	= 940
3	500 + 300 + 400 × 0.15	= 860
4	500 + 400	= 900

分析表明,该厂应采用每月预订 3 千克的采购策略,月平均成本为 860 元。

例 8.9 设备调整决策。

华尔斯机器零件公司是制造大型工业炉的精密零件公司。每次换班时,必须对设备进行调整,这是精细而又费钱的事。根据过去的经验,设备一次调整正确的概率为 80%,这时生产 95% 的合格品;设备一次调整不正确的概率为 20%,这时生产 30% 的合格品。工厂决定,在是否正式生产前,先生产 3 个产品进行质检。如果设备调整正确的概率高于 95% 时,则转入正式生产;如果设备调整正确的概率低于 50% 时,则再调整一次;如果设备调整正确的概率介于二者之间,则再生产 3 个产品检验后再作决定。现对 3 个产品检验后,发现 2 个合格品,1 个不合格品,这时机器调整正确的概率为多少?应采取哪种行动?

分析:这是利用补充信息再进行决策的例子,即后验概率的决策问题。

令:G 为抽中合格品,D 为抽中不合格品,则可计算后验概率,如表8.18 所示。

表 8.18

可能状况 Si	先验概率 $P(Si)$	$P(G/Si)$	$P(D/Si)$	$P(GGD/Si)$	$P(Si \cdot GGD)$	后验概率 $P(Si/GGD)$
正确	0.8	0.95	0.05	0.135	0.108	0.74
不正确	0.2	0.3	0.7	0.189	0.037 8	0.26
合计	1.0	—	—	—	0.146	1.00

故在抽3个产品出现2个合格品和1个不合格品的条件下,机器调整正确的后验概率为 74%,这时应再抽 3 个产品检验后再作决策。

例 8.10 矿产权购买决策。

皮克林公司考虑用 10 万美元购买一片地的矿产权。这一价格包括地震试验以探明这片地是 A 型还是 B 型地质结构。购买后才能知道该地的地质结构类型。估计该地区40% 是 A 型,60% 是 B 型。若公司准备钻井,将花费20 万元。可能钻到油井、气井或干井。经验表明,在 A 型地上钻到油井的概率为0.4,利润为90 万美元;钻到气井的概率为0.2,利润为70 万美元。在 B 型地上钻到油井的概率为0.1,利润为90 万美元;钻到气井的概率为 0.3,利润为 70 万美元。利润是除了矿产权和钻井费外的利润。问皮

克林公司是否应该购买这项矿产权?

分析:该问题是一个多阶段决策问题,可用决策树决策,如图 8.6 所示。

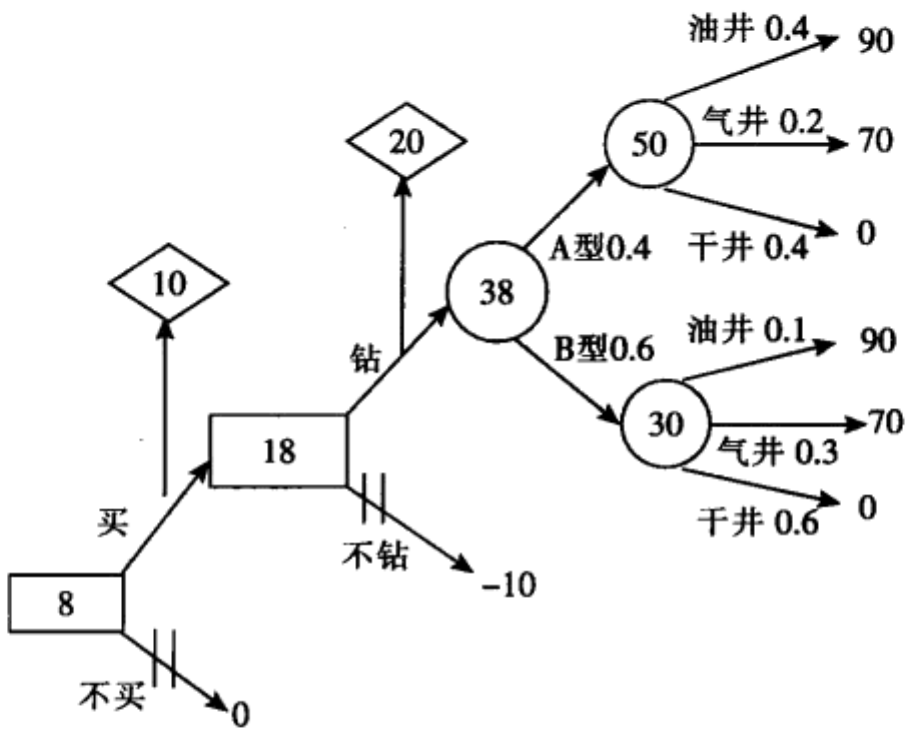


图 8.6

第六节　马尔柯夫决策法

"商场如战场。"市场竞争,风云多变,商家明争暗斗,各出奇招,均不满足于既得的市场份额。归根结底,都想获得更多的顾客。某些商家发现,自家商店留住的老顾客少,流到别家商店的多,而从别的商店转到自己商店的顾客少,长此下去,会使自己占有的市场份额越来越少。更糟糕的是,如果顾客只出不进,最终会被挤出市场。顾客的流动有什么规律可循吗?市场份额减少趋势怎样?经过一段时间,到底会失去多少市场份额?对此情况,有什么办法可以挽救呢?

农贸市场上,某季某种蔬菜的数量过多,则价格下跌,许多农民在下季则转移去种别的蔬菜,这种转移的趋势怎样呢?

工厂中,某些零件或设备由新到旧,由旧到坏,需要更新,这种变化有无规律可循呢?如何考虑一项合理的更新计划,使所花的费用最少呢?

上述问题是否有一定的规律并能否预测呢?面对这种情况应如何决策呢?

一、马尔柯夫决策法的基本概念

马尔柯夫决策法就是利用某一变量的现在状态和动向去预测该变量未来状态及动向,并据此采取相应措施的一种预测和决策方法。

马尔柯夫决策法在经济管理上主要应用在市场销售领域,也可用于维修、更新策略的选择,决定最优工作分配,分析排队系统及库存管理分析的辅助手段等方面。

(一) 一阶马尔柯夫链及特性

俄国数学家马尔柯夫在研究一种称为布朗运动的物理现象时,通过一系列的实

验观察，发现某些事件从一种结果（状态）转移到另一种结果（状态），其后一次的结果仅取决于前一次的结果，而与更前各次的结果是无关的，这样的一种转移过程称为一阶马尔柯夫链。即进行一系列随机实验，如果已知在第 K 次实验中实现什么事件，而在第 $K+1(K=1,2,\cdots)$ 次实验中实现事件的条件概率只与第 K 次实验中发生的事件有关，而与更早实验中发生的事件无关，则称这个实验系列形成一个简单马尔柯夫链，也称为一阶马尔柯夫链。

马尔柯夫链的一个重要特性是无后效性。所谓无后效性，简单来说，就是已知系统的“现在”，那么“将来”与“过去”无关，仅与“现在”有关。对于一阶马尔柯夫链可这样理解：首先，一阶马尔柯夫链是指后一事件的概率仅取决于前一次事件，而与前两次，前三次，乃至更前若干次事件无关；其次，一阶马尔柯夫链是指事件或时间参数为离散变量。

如青蛙跳荷叶，可以认为一只青蛙现在在哪张荷叶上，仅取决于前一次它在哪张荷叶上，而与它前两次，前三次及前更多次在哪张荷叶上没有关系。换句话说，如果要推测这只青蛙下一次会跳到（转移到）哪张荷叶上，只需考虑现在青蛙在哪张荷叶上，而不必考虑它前若干次在哪张荷叶上。

经济现象中，农民今年决定种什么蔬菜，往往是根据去年种什么蔬菜来决定的，而与前年，上前年以及更前各年种什么蔬菜没有多大关系。

（二）高阶马尔柯夫链

许多实例表明，有些现象后一事件的概率不仅与前一次事件有关，而且与前两次、前三次乃至与前 n 次事件有关，而与更前各次事件无关，这样的马尔柯夫链则分别称为二阶、三阶和 n 阶马尔柯夫链。

（三）马尔柯夫过程

若事件或时间参数为连续变量时，则称为马尔柯夫过程。从这个意义上讲，马尔柯夫链可称为离散型马尔柯夫过程，习惯上称为马尔柯夫链，以示与连续变量的马尔柯夫过程相区别。

本课程只讨论用一阶马尔柯夫链进行预测、决策的问题。这是因为：首先，许多研究表明，一阶马尔柯夫链能得出相当可靠的结果，特别是转移过程的概率始终保持稳定时，更是如此；其次，高阶马尔柯夫过程用到的数学更艰难烦琐，计算量太大而不利于实际应用。

（四）状态及初始状态

状态是指在一系列随机实验中，每 一次实验的结果是出现有限个或可数无穷个两两互斥的事件 $S_1,S_2,\cdots$ 中的一个而且仅出现一个，则称这些事件为状态。

在实际应用中，根据不同的问题，应对状态作出不同的规定或合理的解释。如对市场占有率进行预测时，假定某一顾客买某种商品可随机的选择 A、B、C 三个商店，他到某一个商店去购买后，就不会到另一个商店去购买。这里的状态就指 A、B、C 三个商店，也可记为 S_1、S_2、S_3。

初始状态：如果以 S_i^k 表示第 k 次实验的结果是处于 S_i 状态，则实验开始（或从研究时刻）时所处的状态叫做初始状态，记为 S_i^0。

（五）初始概率与转移概率

1. 初始概率

各初始状态发生的概率称为初始概率，可记为 $p(S_i^0)(i = 1, 2, 3\cdots)$ 或简记为 p_i^0。在实际应用中，初始概率可用统计资料计算的频率近似代替。如市场占有率预测中，从研究时刻（实验开始），顾客到 A，B，C 三家（初始状态）去购买某种商品的可能性根据统计资料计算的百分比（频率）分别为 50%、30%、20%，则可认为初始概率分别为 $p(S_1^0) = 50\% = p_1^0; p(S_2^0) = 30\% = p_2^0; p(S_3^0) = 20\% = p_3^0$。这个概率表明，研究时刻开始，$A$ 家占有市场 50% 的顾客，B 家占有 30% 的顾客，C 家占有 20% 的顾客。这个概率也表明，随机挑选一个顾客，他愿意到 A 家购买的意愿为 50 %，到 B 家购买的意愿为 30 %，到 C 家购买的意愿为 20%。

2. 转移概率

若有随机序列 ζ，系统在时刻 m 处于状态 $S_i(m)$，经过 K 步转移后，系统处于 $S_j(m + k)$ 的条件概率为 $P(\zeta_{m+k} = S_j/\zeta_m = S_i)$，记为 $P_{ij}(m,k)$，称为转移概率。可见转移概率是条件概率，实际应用中，转移概率的具体含义要根据实际问题来确定。其值可根据统计资料计算的频率来近似代替。如在市场占有率预测中，转移概率分别以保留率，转出率和转入率三种形式来表示并形成转移概率矩阵，系统的 k 步转移概率矩阵可记为 $P(m,k)$。

$$P(m,k) = \begin{pmatrix} p_{11}(m,k)p_{12}(m,k)\cdots p_{1n}(m,k) \\ p_{21}(m,k)p_{22}(m,k)\cdots p_{2n}(m,k) \\ \cdots \quad \cdots \quad \cdots \quad \cdots \\ p_{n1}(m,k)p_{n2}(m,k)\cdots p_{nn}(m,k) \end{pmatrix}$$

特别称 $P(m,1)$ 为一步转移概率矩阵。

（六）齐次马尔柯夫链

如果马尔柯夫链的转移概率矩阵对任意 k 步转移概率矩阵 $P(m,k)$ 都与 m 无关，则可记 $P(m,k)$ 为 $P(k)$，称此马尔柯夫链为齐次马尔柯夫链。这意味着具有齐次性质的马尔柯夫链由状态 S_i 转为状态 S_j 的概率与状态 S_i 所处的时刻无关，或者说转移概率值与时间原点的选取无关，即转移概率矩阵始终是稳定的。

我们只限于研究齐次马尔柯夫链。对于齐次马尔柯夫链，任意 k 步转移概率矩阵 $P(k)$，均可由一步转移概率矩阵 $P(1)$（简记为 P）来确定，由数学归纳法可得：

$$P(n) = [P(1)]^n = P^n$$

这一特性在理论上与应用上均有重要意义，进一步可推出：

$$p^{(n)} = p^{(0)}P^n$$

可见，系统第 n 个时刻处于各状态的概率取决于初始概率 $p^{(0)}$ 及转移概率矩阵 P。显然，也可看成系统由初始时刻 0 经过 n 步转移后处于各状态的概率，此式在预测中有重要的作用。

（七）马尔柯夫链的遍历性

设有齐次马尔柯夫链，若对于一切状态 S_i 及 S_j，存在不依赖于 i 的常数 Πj，使：

$$\lim_{n\to\infty} P_{ij}(n) = \pi_j$$

则称该齐次马尔柯夫链具有遍历性。

遍历性的含义是指不论系统自哪一状态出发，当转移的“步长”n 充分大时，转移到状态 S_j 的概率接近于某个常数 Π_j，这时有：

$p(S_j)(n-1) = p(S_j)(n)$。

这一结论反映了系统的平衡性，或说系统达到了平衡状态。即事物经过 n 次（理论上 n 趋于无穷大）转移后，则后一次状态的概率与前一次状态的概率完全相同，不再因时间的推移而改变。若马尔柯夫链是遍历的，则 π_j 是如下方程组：

$$\pi_j = \sum_{i=1}^{n} \pi_j P_{ij} (j = 1,2,\cdots,n)$$

在满足条件：$\pi_j > 0$，$\sum_{j=1}^{n} \pi_j = 1$ 的唯一解。

利用这一方程组，可求得马尔柯夫链处于平衡状态的概率来预测事物发展的终极结果。

许多实例表明，某些经济过程是满足或基本满足具有遍历性的齐次马尔柯夫链的变化过程的，对于这些经济过程，在掌握了初始概率与转移概率矩阵资料时，其过程的未来状态就是可预测的。

二、马尔柯夫法预测市场占有率

（一）市场占有率的概念

市场占有率是指一段时间内企业某种产品的销售量占该产品全部销售量的百分比。如全国某年洗衣机的销售量为 100 万台，某厂生产的洗衣机该年销售量为 35 万台，则该厂洗衣机的市场占有率为 35%。

市场占有率是厂商经常追逐的目标，有资料表明，投资利润率高低与市场占有率高低有很大的关系。国外一份研究报告指出，公司的市场占有率低于 10%，其平均投资利润率约为 9%；市场占有率超过 40%，其平均投资利润率则高达 30%。

厂商追逐市场占有率的真正目的是增加利润。一般而言，市场占有率高，利润也高，市场占有率低，利润也低。但有时会发生这样的情况，企业采取增加广告宣传，加强售后服务等手段来提高市场占有率。若采用这些手段的费用高于提高市场占有率增加的利润，则市场占有率的提高反而使利润降低。结合各项支出与利润的综合考虑，市场占有率多高才是恰当的呢？

所谓市场竞争，就是对顾客占有率的竞争。因此，很多情况下可用顾客占有率来近似代替市场占有率。由于各厂商均有提高市场占有率（或顾客占有率）的欲望，并不断采取相应措施，从而使各厂商的顾客占有率经常发生变化。这种变化是否有什么规律没有？能否预测呢？有些厂商面对日益减少的顾客流，希望知道这种减少的趋势如何？有无办法挽救等等问题都迫使厂商要了解、重视、预测市场占有率及市场占有率与利润之间的关系，以便能及时采取相应的措施保证企业的生存和发展。

分析和预测市场占有率的方法有很多，马尔柯夫法是最常用的方法之一。

（二）马尔柯夫决策法的假设条件

马尔柯夫预测决策法作为一种数学模型，在运用这种模型时均有一定的假设条件。实际问题对这些假设条件的满足程度越高，则预测的效果越好。不同的实际问题，

某些假设条件也有所不同。

下面的假设条件是以市场占有率预测为对象的。对于其他领域的应用，这些假设条件也是基本适用的。

（1）仅以一阶马尔柯夫过程为研究对象：

（2）转移概率矩阵逐期保持不变；

（3）顾客按固定的时间间隔购货，且每次购货的数量相等；

（4）广大顾客相对均一，使转移概率矩阵适用于所有顾客；

（5）销售总额的大小逐期保持不变；

（6）市场上竞争对手的数目保持不变，既没有新的卖主加入竞争，也没有一个卖主停业退出竞争。

现实经济生活中，完全满足这些假设条件的经济问题是不多的。如很多实例表明，一个顾客的购买行为并非仅受前一周期购买行为的影响，而常常是或多或少与他的全部购货经历或至少与前两，前三个周期的经历有关，这时用二阶，三阶或 n 阶马尔柯夫链来描述更切合实际。即很多顾客的购买行为并不符合一阶马尔柯夫链。但若其购买行为虽与前几期有关，但仅与前一个周期的关系最密切，以前各期的影响较小，则可近似地用一阶马尔柯夫链来描述。

上例表明，这些假设条件对我们在解决实际问题中应用马尔柯夫法带来一定的局限性。解决的方法一是研究能否近似的满足这些假设条件；二是选用更准确的模型，从而使马尔柯夫法更切合实际。我们介绍的是最简单的一阶马尔柯夫模型，选用更复杂的马尔柯夫模型，一般会增加数学上的复杂程度及计算工作量。

（三）马尔柯夫法预测市场占有率的基本步骤

（1）分析所预测的问题是否适用于马尔柯夫模型及是否满足模型的假设条件，在后面的例子中，我们均假设是满足这些条件的；

（2）分析该问题有几种状态并计算初始状态的概率；

（3）计算转移概率并用矩阵表示；

（4）预测所需周期的市场占有率；

（5）计算平衡状态下的市场占有率；

（6）根据市场占有率预测进行相应的决策。

（四）案例：市场争夺之战

某地区某种化肥市场，一直销售 A 、B 、C 三个厂家的化肥，为该地区50 000 户农户服务。历史上，A 厂家一直是该地区最大的厂商，其市场占有率为50%，即为该地区 50% 的顾客服务；B 厂家为该地区 30% 的顾客服务；C 厂家为该地区 20% 的顾客服务。

一般来讲，该地区的农户一季度购买一次化肥。

最近，A 厂感到 C 厂采取了一种把 A 厂和 B 厂的顾客吸引到 C 厂的营销策略，使得 A 厂的顾客大为减少。市场调查表明，在今年第一季度（研究时刻，可记为周期 0），A 厂有 25 000 户顾客，B 厂有 15 000 户顾客，C 厂有10 000 户顾客。但在第二季度（周期 1），A 厂保住了 17 500 户老顾客，损失了 7 500 户顾客，其中有 2 500 户转向 B 厂，有 5 000 户转向 C 厂；B 厂保留住了12 000 户老顾客，损失的 3 000 户顾客中，有 1 500

户转向A厂,另1 500户转向C厂;C厂保留住了9 000户老顾客,只损失了1 000户 顾客,其中有500户转向A厂,另500户转向B厂。

A厂厂长对顾客的明显减少深感忧虑,他希望了解这种状况如果继续下去,本厂在今年第四季度(周期3)及明年第一季度(周期4)的市场占有率将会下降到什么程度?这种下降势头是否会有尽头?如果有,大约会发生在什么时间?那时的市场占有率是多少?面对这种状况,有什么措施可以补救吗?

分析:对于这个问题,可用马尔柯夫法进行预测和决策。

1. 马尔柯夫链的适用性及假设条件分析

在这个问题中,顾客下一个季度购买哪家的化肥往往是根据上一个季度的购买经历来作出决定的;顾客每隔一个季度购买一次化肥,且每次购买的数量大致相同;每季度购买化肥的总量大致相同;每季度各厂的保留率,转出率和转入率大致相同;顾客愿意到哪个厂购买化肥的意愿是随机的;该地区的化肥市场由3个厂家占领,估计今后一段时间没有厂家加入,也没有厂家退出。因此,该问题基本满足一阶马尔柯夫模型的假设条件,故可用一阶马尔柯夫模型进行分析。

2. 状态及初始状态概率的确定

这个问题一共有三种状态,即市场份额分别由A、B、C三家瓜分,即$S_1 = A, S_2 = B, S_3 = C$。今年第一季度的各厂家的市场占有率(初始状态概率)分别是:$p(S_1) = 50\%$、$p(S_2) = 30\%$、$p(S_3) = 20\%$,则初始概率用向量形式可表示为:

$$p^{(0)} = (0.5 \quad 0.3 \quad 0.2)$$

3. 转移概率的计算

在第二季度(周期1),顾客流动情况如表8.19所示。

表8.19 单位:户

厂家(状态)	第一季度顾客数	第二季度供应厂家			第二季度顾客合计数
		$A(S_1)$	$B(S_2)$	$C(S_3)$	
$A(S_1)$	25 000	17 500	2 500	5 000	25 000
$B(S_2)$	15 000	1 500	12 000	1 500	15 000
$C(S_3)$	10 000	500	500	9 000	10 000
合计	50 000	19 500	15 000	15 500	50 000

表8.19的资料表明,第二季度(周期2)与第一季度的顾客总数没有变化,都为50 000户。这符合我们的基本假设,以后周期也是如此。

第一行表明在周期1,A厂保留住25 000户中的17 500户,2 500户转向B厂,5 000户转向C厂;

第二行表明在周期1,B厂保留了15 000户中的12 000户,1 500户转向A厂,1 500户转向C厂;

第三行表明在周期1,C厂保留了10 000户中的9 000户,500户转向A厂,500户转向B厂。

第一列表明,A家保留住老顾客17 500户,从B厂吸引了1 500户新顾客,从C厂吸引了500户新顾客,A厂周期1共有19 500户顾客,比周期0减少5 500户;

第二列表明，B厂保留住12 000户老顾客，从A厂吸引了2 500户新顾客，从C厂吸引了500户新顾客，顾客总数与周期0相同，均为15 000户。

第三列表明，C厂保留住9 000户老顾客，从A厂吸引了5 000户新顾客，从B厂吸引了1 500户新顾客。C厂周期1的顾客数为15 500户，比周期0增加了5 500户顾客。

利用表8.19的资料，可以计算出顾客的转移频率。如A厂的保留率为$P(S_{11}) = 0.7(17\ 500/25\ 000)$；转到$B$厂的转出率为$P(S_{12}) = 0.1\ (2\ 500/25\ 000)$，这也是$B$厂的转入率；转到$C$厂的转出率为$P(S_{13}) = 0.2(5\ 000/25\ 000)$，这也是$C$厂的转入率。

同理可计算B厂，C厂的保留率，转出率。得到该问题的转移频率，如表8.20所示。

表8.20

厂家	A	B	C	合计
A	0.7	0.1	0.2	1.0
B	0.1	0.8	0.1	1.0
C	0.05	0.05	0.9	1.0

用此全部转移频率来代替转移概率，则可得到转移概率矩阵，这就是第一次（步）转移概率矩阵：

$$P(1) = P = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix} = \begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.1 & 0.8 & 0.1 \\ 0.05 & 0.05 & 0.0 \end{bmatrix}$$

矩阵中的每一行表示各厂家在下一个周期能保住老顾客的百分比（概率）及将要流失给竞争对手的顾客的百分比，即表示各厂家的留与失的概率。表中的每一列表示各厂家在下一个周期能够保住老顾客的概率及将要从其他竞争对手那里获得顾客的概率，即表示各厂家留与得的概率。

转移概率矩阵还表示随机挑选一个顾客，从一个周期到下一个周期他愿意购买哪家产品的可能性（即概率）。如从原A厂的顾客中随机挑选一名顾客，他在下一季度仍在A厂购买的可能性是70%，转到B厂购买的可能 为10%，转到C厂购买的可能性为20%。因此，转移概率矩阵也描绘了顾客随机流动的规律性。

4. 未来时期市场占有率预测

如果顾客流动的规律在以后的周期均按此转移概率流动，则可利用公式$p^{(n)} = p^{(0)}P^n$来预测以后周期各厂的市场占有率。即只要知道初始时刻的市场占有率$p^{(0)}$及一步转移概率矩阵P，就可递推出以后各期各厂家的市场占有率。

即：当$n = 1$时，$p^{(1)} = p^{(0)}P^1 = p^{(0)}P$；

即：当$n = 2$时，$p^{(2)} = p^{(0)}P^2$；

即：当$n = 3$时，$p^{(3)} = p^{(0)}P^3$。

具体而言，本年第二季度（周期1）各厂的市场占有率为：

周期0的市场占有率 (A B C) × 转移概率矩阵 = 周期1的市场占有率 (A B C)

$$(0.5 \quad 0.3 \quad 0.2) \times \begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.1 & 0.8 & 0.1 \\ 0.05 & 0.05 & 0.9 \end{bmatrix} = (0.39 \quad 0.3 \quad 0.31)$$

上面计算表明，*A*厂的市场占有率由第一季度(周期0)的50%下降到第二季度(周期1)的39%；*B*厂第二季度(周期1)的市场占有率与第一季度(周期0)相同，均为30%；*C*厂的市场占有率由第一季度(周期0)的20%上升到第二季度(周期1)的31%。

现以*A*厂第二季度的市场占有率的计算过程为例说明上述计算过程及其含义。

将周期0的市场占有率乘以转移概率矩阵的第一列，则有：

0.5×0.7 = 0.35，即*A*的市场占有率乘以*A*厂保留率，结果表示*A*厂在新周期留住的老顾客占全部顾客数的百分比，可验证，50 000 × 0.35 = 17 500(户)；

0.3×0.1 = 0.03，即*B*厂的市场占有率乘以*A*厂从*B*厂的转入率，结果表示*A*厂在新周期从*B*厂获得的新顾客占全部顾客数的百分比，可验证，50 000 × 0.03 = 1 500(户)；

0.2×0.05 = 0.01，即*C*厂的市场占有率乘以*A*厂从*C*厂的转入率，结果表示*A*厂在新周期从*C*厂获得的新顾客占全部顾客数的百分比，可验证，50 000 × 0.01 = 500(户)；

0.35 + 0.03 + 0.01 = 0.39 三者相加，表示*A*厂在第二季度的市场占有率，顾客数为50 000 × 0.39 = 19 500(户)。

第三季度(周期2)的市场占有率计算如下：

周期1的市场占有率 (A B C) × 转移概率矩阵 = 周期2的市场占有率 (A B C)

$$(0.39 \quad 0.3 \quad 0.31) \times \begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.1 & 0.8 & 0.1 \\ 0.05 & 0.05 & 0.9 \end{bmatrix} = (0.319 \quad 0.294 \quad 0.387)$$

计算结果表明，如果按目前顾客转移的趋势继续下去(转移概率不变)，到第三季度(周期2)，*A*厂的市场占有率将下降到31.9%，*B*厂下降到为29.4%，*C*厂将上升为38.7%。

按此方法逐期递推，则可继续推出以后周期的市场占有率，以此递推出该地区第十个季度化肥市场的市场占有率，如表8.21所示。

表 8.21

周期	市场占有率(%)		
	A	*B*	*C*
0	50.0	30.0	20.0
1	39.0	30.0	31.0
2	31.9	29.4	38.7
3	27.2	28.6	44.2
4	24.1	27.9	48.0
5	22.1	27.1	50.8
6	20.7	26.4	52.9
7	19.8	25.8	54.4
8	19.1	25.4	55.5
9	18.7	25.0	56.3

表 8.21 的资料表明,如果以后周期继续按相同的转移概率转移,到第十季度(转移 9 次),*A* 厂的市场占有率将由 50% 下降为 18.7%,*B* 厂的市场占有率将由 30% 下降为 25%,唯有 *C* 厂的市场占有率将由 20% 上升为 56.3%。

表 8.21 的资料是根据逐步递推而来的,计算比较繁琐,特别是当我们只需要知道未来某一个周期的状态时,则可用下式直接计算:

$$p^{(n)} = p^{(0)}P^n$$

例如要预测本年第四季(周期 3)各厂的市场占有率,这时 $n = 3$,则用公式 $p^{(3)} = p^{(0)}P^3$ 计算如下:

$$(0.5 \quad 0.3 \quad 0.2) \times \begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.1 & 0.8 & 0.1 \\ 0.05 & 0.05 & 0.9 \end{bmatrix}^3 = (0.272 \quad 0.286 \quad 0.442)$$

若要预测下一年第一季度(周期 4)各厂的市场占有率,这时 $n = 4$,则用公式 $p^{(4)} = p^{(0)}P^4$ 计算如下:

$$(0.5 \quad 0.3 \quad 0.2) \times \begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.1 & 0.8 & 0.1 \\ 0.05 & 0.05 & 0.9 \end{bmatrix} = (0.241 \quad 0.279 \quad 0.48)$$

5. 平衡状态时市场占有率预测

A 厂厂长对前面推算出的该厂到第十个季度(周期 9)的市场占有率由第一个季度的 50% 下降到 18.7% 深感不安。他希望了解这种顾客减少的趋势是否会有尽头?如果有,会在什么时间发生?到时该厂的市场占有率会是多少?

用上面递推的办法,经计算,到第 24 个季度(周期 23),*A* 厂的市场占有率下降到 17.65 %,*B* 厂为 23.53%,*C* 厂为 58.82 %。这时的市场占有率将稳定下来,以后周期不会再发生变化。可以验证,*A* 、*B* 、*C* 厂到第 25 个季度的市场占有率与第 24 个季度完全一样,即:

第23个周期市场占有率 (A B C) × 转移概率矩阵 = 第24个周期市场占有率 (A B C)

$$(0.1765 \quad 0.2353 \quad 0.5882) \times \begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.1 & 0.8 & 0.1 \\ 0.05 & 0.05 & 0.9 \end{bmatrix} = (0.1765 \quad 0.2353 \quad 0.5882)$$

这时我们称市场占有率经过 n 次（本例为 $n = 23$）转移后达到平衡状态。即这三个厂家的市场占有率以后逐期将保持变。

用逐期递推达到平衡状态的好处在于能具体知道这种变化趋势持续多长时期，但计算量过大。若只需了解最后的平衡状态，可利用公式：

$$\pi_j = \sum_{i=1}^{n} \pi_j P_{ij} (j = 1,2,\cdots,n)$$

在满足条件 $\pi_j > 0, \sum_{j=1}^{n} \pi_j = 1$ 时，建立联立方程组直接求得平衡状态的市场占有率。

$$P(1) = P = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix} = \begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.1 & 0.8 & 0.1 \\ 0.05 & 0.05 & 0.9 \end{bmatrix}$$

在市场占有率预测中，π_j 表示 j 个厂商（本例 $j = 1,2,3$）在平衡状态时的市场占有率，$P_{ij}(i,j = 1,2,3)$ 为转移概率，则可建立如下方程式组：

$$\begin{cases} \pi_1 = 0.7\pi_1 + 0.1\pi_2 + 0.05\pi_3 & (1) \\ \pi_2 = 0.1\pi_1 + 0.8\pi_2 + 0.05\pi_3 & (2) \\ \pi_3 = 0.2\pi_1 + 0.1\pi_2 + 0.9\pi_3 & (3) \\ \pi_1 + \pi_2 + \pi_3 = 1 & (4) \end{cases}$$

上面方程式组中，(1) 式表示，在平衡条件下，A 厂在 $n+1$ 个周期的市场占有率为 n 期 A 厂 70% 市场占有率加上 n 期 B 厂市场占有率的 10%，再加上 n 期 C 厂市场占有率的 5%。(2)、(3) 式的经济意义与 (1) 式相同，(4) 式表示三家的市场占有率之和必为 100%，即该市场由这三个厂家瓜分。

上面方程组中有 4 个方程，但只有三个未知数，可用其中三个方程来求解。但须注意，必须利用第 4 个方程（$\pi_1 + \pi_2 + \pi_3 = 1$）和其他三个方程中的任意两个。如不用第 4 个方程，则可能出现对本问题的不可行解。下面用后三式求解如下：

$$\begin{cases} \pi_2 = 0.1\pi_1 + 0.8\pi_2 + 0.05\pi_3 & (2) \\ \pi_3 = 0.2\pi_1 + 0.1\pi_2 + 0.9\pi_3 & (3) \\ \pi_1 + \pi_2 + \pi_3 = 1 & (4) \end{cases}$$

解方程组得：

$\pi_1 = 0.1765 = A, \pi_2 = 0.2353 = B, \pi_3 = 0.5882 = C$

这一结果与前面递推得到的结果完全一致，但计算简便得多。

6. 针对市场占有率预测结果的市场决策

利用马尔柯夫方法预测到未来市场占有率的状况使 A 厂厂长深感问题严重，他

希望能找到一些办法来扭转这一局势，而不能任其发展。可供 A 厂厂长选择的策略一般有保留策略和争取策略。

保留策略的目标在于力图保留住老顾客，使保留率较高。其办法有提供更好的服务，加强售后服务，价格折扣优惠等。

争取策略的目标在于尽量从对手中争取新顾客，使转入率较高。其办法有加大广告宣传，直接联络等。

现 A 厂若采用对连续两期购货的顾客折价优待的留住更多老顾客的保留策略，估计 A 厂的保留率可从 70% 提高到 85%，则这时的转移概率矩阵为：

$$P = \begin{bmatrix} 0.85 & 0.1 & 0.05 \\ 0.1 & 0.8 & 0.1 \\ 0.05 & 0.05 & 0.9 \end{bmatrix}$$

按此转移概率矩阵计算得到的平衡状态的市场占有率分别为：A 厂 31.6%、B 厂 26.3%、C 厂 42.1%，但需要花费 5 万元。

如 A 厂准备通过广告宣传从 B 厂和 C 厂吸引新顾客，估计从某一周期 B 厂和 C 厂的转入率均为 15%，则新的转移概率矩阵如下：

$$P = \begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.15 & 0.75 & 0.1 \\ 0.15 & 0.05 & 0.8 \end{bmatrix}$$

按此转移概率矩阵计算出的平衡状态的市场占有率分别为：A 厂 33.3%、B 厂 22.2%、C 厂 44.5%。但需广告费 8 万元。

若每多争取到 1% 的市场份额，可获利 1 万元。争取策略比保留策略的市场占有率高 1.7%，可多获利 1.7 万元。但花费上比保留策略多花 1.3(8 − 5 − 1.7) 万元。故从费用角度出发，以选保留策略花费较少。

实际工作中，这两种策略是可以同时采用的，但也要进行费用与所得的比较。

三、马尔柯夫法的其他应用

市场占有率的预测和决策虽然是马尔柯夫法主要的应用领域，但马尔柯夫法在其他领域也有广泛的应用。如把某些排队问题当成马尔柯夫过程来进行分析，特别是当“到达”或“服务”的概率分布是某种不常见分布时，更显出其优点；又如设备的维修策略选择等；此外，对环境保护、生态平衡等问题的预测与决策时，应用马尔柯夫法，更能获得良好效果；在对确定污染物份额及其变化趋势，污染物份额的平衡条件，选择治理环境的最佳策略等问题的分析与处理，应用马尔柯夫分析方法更为适宜等等。

例 8.11 机器维修策略的选择。某工厂每季对一种机器进行检查，机器状况可分为 5 种状态：

(1) S_1 为优良状态；

(2) S_2 为良好状态；

(3) S_3 为及格状态；

(4) S_4 为可用状态；

(5)S_5 为不可使用状态。

目前,该厂对这种机器实行“一种状态”的维修方法,即当检查到机器属于状态5时,则进行维修,维修后的机器则处于第1种状态。每台机器的维修费为500元。

以往经验表明,在某给定检查周期内处于状态1的机器,在下一个检查周期分别属于状态2、3、4、5的概率为0.6、0.2、0.1、0.1;处于状态2的机器,在下一个周期分别处于状态2、3、4、5的概率为0.3、0.4、0.2、0.1;处于状态3的机器,在下一个周期分别处于状态3、4、5的概率为0.4、0.4、0.2;处于状态4的机器,在下一个周期分别处于状态4、5的概率为0.5、0.5;处于状态5的机器因要全部送去修理,故在下一个周期均处于状态1,即处于状态1的概率为1。

该厂现在关心的是,这种“一种状态”的维修方案,从经济上是否是最合算的。能否有一种使总维修费用最少的方案?

分析:目前该厂考虑有两种可供增加选择的方案:一是“二种状态”的维修方案,即将处于状态4、5的机器全部送去维修;另一个是“三种状态”的维修方案,即将处于状态3、4、5的机器全部送去维修。估计的平均修理费用对处于状态3的为200元,处于状态4的为250元,处于状态5的仍为500元。

经分析,机器下周期处于什么状态是根据上一个周期来确定的,变量(处于不同状态的机器数)为离散变量,各状态间存在一个转移概率矩阵,故可用一阶马尔柯夫链来进行分析。

在“一种状态”的维修时,转移概率矩阵为:

$$
\begin{array}{c}
 \\
S_1 \\
S_2 \\
S_3 \\
S_4 \\
S_5
\end{array}
\begin{array}{c}
\begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\
\begin{bmatrix}
0 & 0.6 & 0.2 & 0.1 & 0.1 \\
0 & 0.3 & 0.4 & 0.2 & 0.1 \\
0 & 0.0 & 0.4 & 0.4 & 0.2 \\
0 & 0.0 & 0.0 & 0.5 & 0.5 \\
1 & 0.0 & 0.0 & 0.0 & 0.0
\end{bmatrix}
\end{array}
$$

根据这一转移概率矩阵可建立方程组,求得平衡状态下处于各种状态的机器所占的百分比。

方程组为:

$$
\begin{cases}
S_1 = S_5 \\
S_2 = 0.6S_1 + 0.3S_2 \\
S_3 = 0.2S_1 + 0.4S_2 + 0.4S_3 \\
S_4 = 0.1S_1 + 0.2S_2 + 0.4S_3 + 0.5S_4 \\
S_5 = 0.1S_1 + 0.1S_2 + 0.2S_3 + 0.5S_4 \\
S_1 + S_2 + S_3 + S_4 + S_5 = 1
\end{cases}
$$

利用该方程组中的5个方程(注意必须用第6个方程)解方程组得:

$S_1 = 0.199, S_2 = 0.17, S_3 = 0.18, S_4 = 0.252, S_5 = 0.199$。

表明某一批机器经过若干周期后,处于各种状态的机器将达到一种稳定的百分比状态,即处于状态1的机器占机器总数的19.9%,处于状态2、3、4、5的机器分别占

机器总数17%、18%、25.2%、19.9%。

在“一种状态”维修方案下，则平均每季维修的机器数为19.9%，由于状态5每台的维修费为500元，则每季每台机器的期望维修费为 $EMC_1 = 500 \times 0.199 = 99.5$(元)。

在“二种状态”的维修时，转移概率矩阵为：

$$\begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.6 & 0.2 & 0.1 & 0.1 \\ 0 & 0.3 & 0.4 & 0.2 & 0.1 \\ 0 & 0.0 & 0.4 & 0.4 & 0.2 \\ 1 & 0.0 & 0.0 & 0.0 & 0.0 \\ 1 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix} \end{array}$$

这个转移概率矩阵与“一种状态”维修方案的不同处是第4行，因为处于状态4的机器也要维修，故下一个周期均属于状态1。

根据新的转移概率矩阵计算出的平衡状态为：$S_1 = 0.266$，$S_2 = 0.228$，$S_3 = 0.241$，$S_4 = 0.168$，$S_5 = 0.097$。

由于处于状态4和状态5的机器均要修理，这一方案每季每台机器的期望维修费为：

$EMC_2 = 250 \times 0.168 + 500 \times 0.097 = 90.5$(元)。

在“三种状态”的维修时，由于处于状态3、4、5的机器均需维修，故下一周期均属于状态1，则其转移概率矩阵为：

$$\begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & S_5 \end{array} \\ \begin{bmatrix} 0 & 0.6 & 0.2 & 0.1 & 0.1 \\ 0 & 0.3 & 0.4 & 0.2 & 0.1 \\ 1 & 0.0 & 0.0 & 0.0 & 0.2 \\ 1 & 0.0 & 0.0 & 0.0 & 0.0 \\ 1 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix} \end{array}$$

据此计算的平衡状态为：$S_1 = 0.35$，$S_2 = 0.3$，$S_3 = 0.19$，$S_4 = 0.095$，$S_5 = 0.065$。

则“三种状态”维修方案每季每台机器的期望维修费为：

$EMC_3 = 200 \times 0.19 + 250 \times 0.095 + 500 \times 0.065 = 94.25$(元)。

三种可供选择的维修方案的期望维修费分别为99.5元，90.5元，94.25元。故“两种状态”的维修方案是维修费用最少的方案，比目前的“一种状态”的维修方案更好。

例8.12 佳维修点的选择。某公园游船管理处在湖的周围设有甲、乙、丙三个租船还船点，游人可在三处中的任何一处租船，也可在三处中的任何一处还船，假设游人在甲、乙、丙三处租船还船的概率为表8.22所列的数据。试用马尔柯夫分析法选择在甲、乙、丙三处中，何处设置游船维修站较佳。

表 8.22

租船处	还船处		
	甲	乙	丙
甲	0.8	0.2	0.0
乙	0.2	0.0	0.8
丙	0.2	0.2	0.6

分析:据表 8.22 可知,甲、乙、丙三处船只转移概率矩阵为:

$$P = \begin{bmatrix} 0.8 & 0.2 & 0.0 \\ 0.2 & 0.0 & 0.8 \\ 0.2 & 0.2 & 0.6 \end{bmatrix}$$

根据这一转移概率矩阵可建立方程组,可求得平衡状态甲、乙、丙三处船只的百分比。

$$\begin{cases} S_{甲} = 0.8S_{甲} + 0.2S_{乙} + 0.2S_{丙} \\ S_{乙} = 0.2S_{甲} + 0.2S_{丙} \\ S_{丙} = 0.8S_{乙} + 0.6S_{丙} \\ S_{甲} + S_{乙} + S_{丙} = 1 \end{cases}$$

利用该方程组中的3个方程(必须用第4个方程),解得 $S_{甲} = 1/2, S_{乙} = 1/6, S_{丙} = 1/3$。即经过较长时间经营后,甲处拥有全部船只的1/2,乙有1/6,丙有1/3,故维修点建于甲处为最佳。

第七节　WinQSB 软件应用

WinQSB 软件用于解决风险决策问题的子程序是 Decision Analysis(DA),主要功能包括贝叶斯分析、效益表分析、决策树等问题的求解。马尔柯夫决策需要另外调用子程序 MarKov Process(MKP)。下面,我们通过例子来说明各种决策分析的具体操作和计算。

一、完全风险决策

效益表分析是已知策略各状态的效益和概率,分析 7 种决策准则下的决策结果。

例 8.13　某公司为经营业务的需要,决定在现有生产条件不变的情况下,生产一种新产品,现可供开发的产品有 Ⅰ、Ⅱ、Ⅲ、Ⅳ 四种不同产品,对应的方案为 A_1、A_2、A_3、A_4。由于缺乏相关资料背景,对产品的市场需求只能估计为大、中、小三种状态,而且对于每种状态出现的概率也无法预测。每种方案在各种自然状态下的效益值表如表 8.23 所示,试用 WinQSB 软件求解。

表 8.23　　单位:万元

自然状态 / 供选方案	需求量大 S_1	需求量中 S_2	需求量小 S_3
A_1:生产产品 Ⅰ	800	320	-250
A_2:生产产品 Ⅱ	600	300	-200
A_3:生产产品 Ⅲ	300	150	50
A_4:生产产品 Ⅳ	400	250	100

解:首先要建立一个新问题,系统会显示如图 8.7 的对话框,选择输入标题“DP1”(读者也可以自己取名)、自然状态数 3 和供选方案数 4,将表8.23 的数据输入到图 8.9 中,第一行要求输入先验概率,本例没有先验概率,按各状态发生的概率相等计算。

点击“Solve the Problem”,系统显示如图8.8所示的界面,提示即将用到的各种决策准则,输入乐观系数 0.3,得到对应的决策结果,求解结果见图 8.10。点击“Results” → “Show Payoff Table Analysis”,显示各决策准则的详细分析结果,见图 8.10。点击“Show Regret Table”显示后悔值表。点击“Show Decision Tree Graph”显示决策树图。

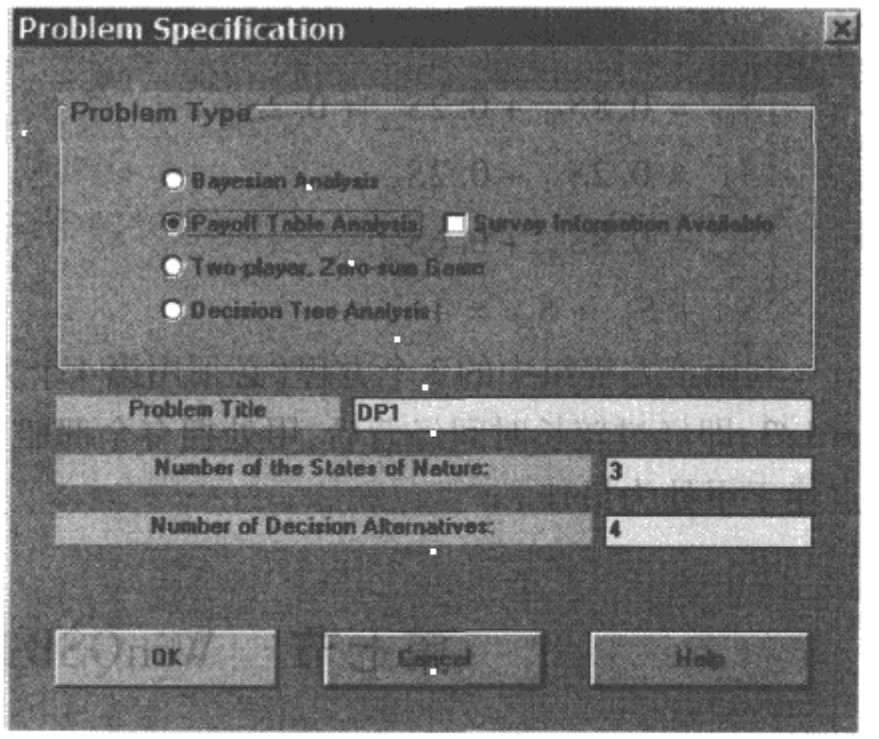


图 8.7

由于表中没有输入状态发生的概率,图 8.9 和图 8.10 的期望收益(后悔)值按各状态发生的概率相等计算。

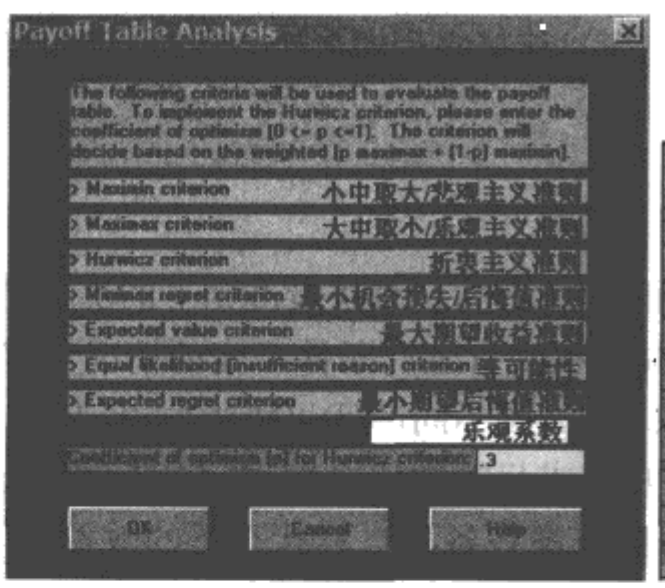


图 8.8

表 8 – 24

04-11-2006 Criterion	Best Decision	Decision Value	
Maximin	Alternative4	$100	
Maximax	Alternative1	$800	
Hurwicz (p=0.3)	Alternative4	$190	
Minimax Regret	Alternative2	$300	
Expected Value	Alternative1	¥ 289.95	
Equal Likelihood	Alternative1	¥ 290.00	
Expected Regret	Alternative1	¥ 116.69	
Expected Value	without any	Information =	¥ 289.95
Expected Value	with Perfect	Information =	¥ 406.64
Expected Value	of Perfect	Information =	¥ 116.69

图 8.9

[64-bit] [Alternative]	Optimistic Value	Maximax Value	Hurwicz (a=0.2) Value	Minimax Regret Value	Equal Likelihood Value	Expected Value	Expected Regret
Alternative1	($250)	$800**	¥ 65.00	$350	¥ 290.00**	¥ 289.95**	¥ 116.69**
Alternative2	($200)	$600	¥ 40.00	$300**	¥ 233.33	¥ 233.29	¥ 173.35
Alternative3	$50	$300	$125	$500	¥ 166.67	¥ 166.65	¥ 239.98
Alternative4	$100**	$400	$190**	$400	¥ 250.00	¥ 249.98	¥ 156.65

图 8.10

二、决策树

例 8.14 某公司由于市场需求增加决定扩大公司规模，供选方案有三种：第一种方案，新建一个大工厂，需投资250万元；第二种方案，新建一个小工厂，需投资150万元；第三种方案，新建一个小工厂，2年后若产品销路好再考虑扩建，扩建需追加120万元，后3年收益与新建大工厂相同。根据预测，该产品前3年畅销和滞销的概率分别为0.6、0.4。若前2年畅销，则后3年畅销和滞销概率为0.8、0.2；若前2年滞销，则后3年一定滞销。如表8.24所示，请运用决策树方法对方案作出选择。

表 8.24

自然状态	概率		供选方案与效益			
					先小后大	
	前2年	后3年	大工厂（万元）	小工厂（万元）	前2年（万元）	后3年（万元）
畅销	0.6	畅销 0.8 滞销 0.2	150	80	80	150
滞销	0.4	畅销 0 滞销 1	-50	20	20	-50

解：这是一个多级决策问题，在 WinQSB 中将节点分为决策点（Decision Node）和分枝节点（Chance Node）两类，同时将事件的终点也看成分枝节点。

根据问题，我们发现可以建立一个有23个节点的决策树，其中决策点设为节点1和节点9，分枝节点有21个。运用 WinQSB，首先要建立一个新问题，在图8.7中选择“Decision Tree Analysis”，输入标题“DP2”（读者也可以自己取名）和节点数23。然后输入表8.26中的数据，如图8.11。图8.11中第二列为节点类型，决策点输入字母“D”，分枝节点（状态）输入字母“C”。节点1和节点9输入“D”，其他节点输入“C”。第三列输入紧后节点；第四列输入成本或收益，其中节点2、3、4、11输入成本，节点5到节点10为2年的收益，将图8.11的效益乘以2后再输入，节点13到节点23为3年的效益，乘以3后再输入。

Node/Event Number	Node Name or Description	Node Type (enter D or C)	Immediate Following Node (numbers separated by ";")	Node Payoff (+ profit, - cost)	Probability (if available)
1	Event1	D	2,3,4		
2	Event2	C	5,6	-250	
3	Event3	C	7,8	-150	
4	Event4	C	9,10	-150	
5	Event5	C	13,14	300	0.6
6	Event6	C	15	-100	0.4
7	Event7	C	16,17	160	0.6
8	Event8	C	18	40	0.4
9	Event9	D	11,12	160	0.6
10	Event10	C	23	40	0.4
11	Event11	C	19,20	-120	
12	Event12	C	21,22		
13	Event13	C		450	0.8
14	Event14	C		-150	0.2
15	Event15	C		-150	1
16	Event16	C		240	0.8
17	Event17	C		60	0.2
18	Event18	C		60	1
19	Event19	C		450	0.8
20	Event20	C		-150	0.2
21	Event21	C		240	0.8
22	Event22	C		60	0.2
23	Event23	C		60	1

图 8.11

点击“Solve and Analyze”→“Solve the Problem”,显示如图8.12所示的计算结果。点击“Results”→“Show Decision Tree Graph”,显示如图8.13所示的决策树。

04-12-2006	Node/Event	Type	Expected value	Decision
1	Event1	Decision node	$112	Event4
2	Event2	Chance node	$278	
3	Event3	Chance node	¥ 258.40	
4	Event4	Chance node	$262	
5	Event5	Chance node	$330	
6	Event6	Chance node	($150)	
7	Event7	Chance node	$204	
8	Event8	Chance node	$60	
9	Event9	Decision node	$210	Event11
10	Event10	Chance node	$60	
11	Event11	Chance node	$330	
12	Event12	Chance node	$204	
13	Event13	Chance node	0	
14	Event14	Chance node	0	
15	Event15	Chance node	0	
16	Event16	Chance node	0	
17	Event17	Chance node	0	
18	Event18	Chance node	0	
19	Event19	Chance node	0	
20	Event20	Chance node	0	
21	Event21	Chance node	0	
22	Event22	Chance node	0	
23	Event23	Chance node	0	
Overall	Expected	Value =	$112	

图 8.12

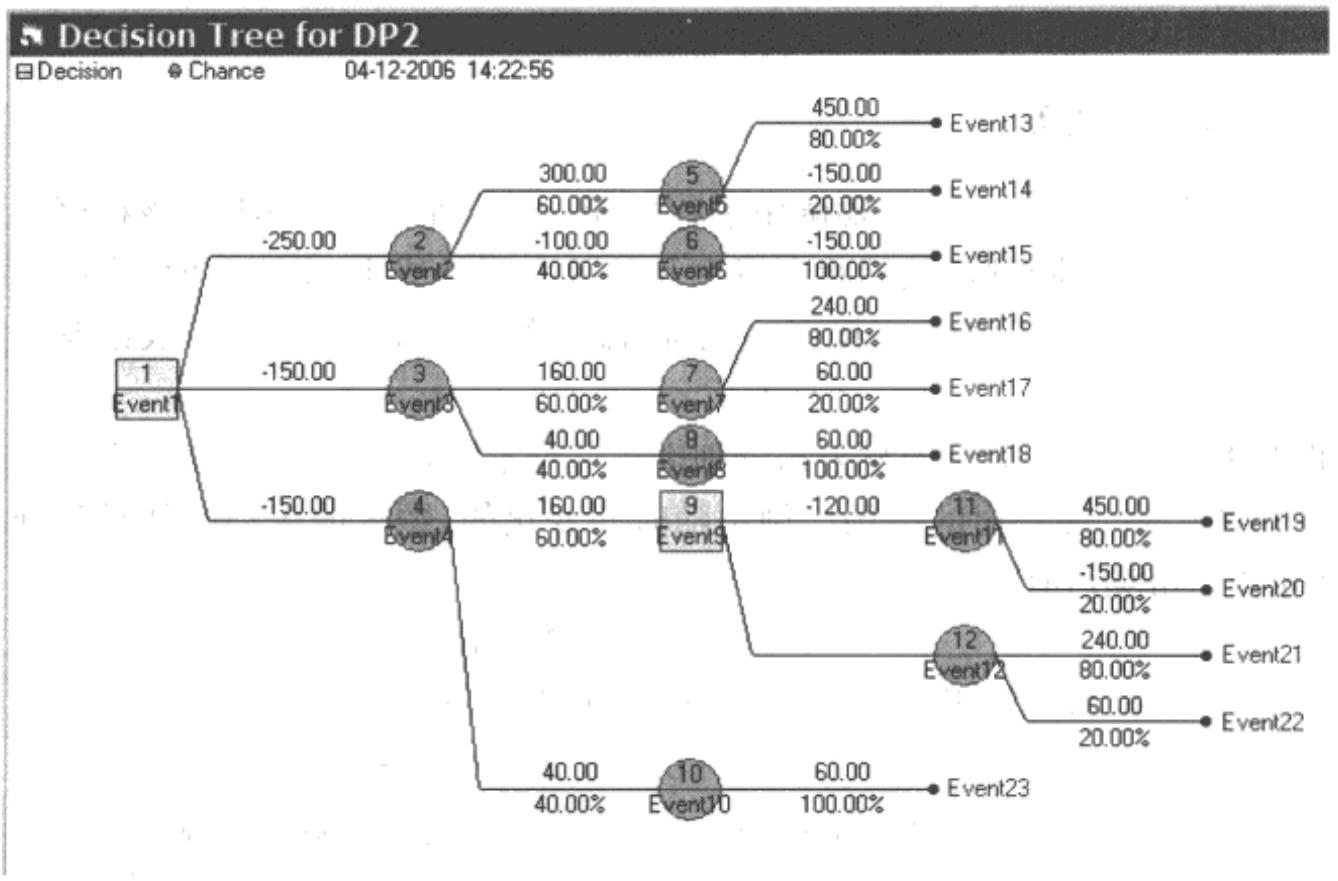


图 8.13

三、后验概率决策

WinQSB 软件作贝叶斯分析只能计算后验概率，收益期望值需手工计算。

例 8.15 某民营企业对一台机器的换代问题作决策，有三种方案：A_1 为买一台新机器；A_2 为对老机器进行改装；A_3 是维护老机器。输入不同质量的原料，三种方案的收益如表 8.25 所示。约有 30% 的原料是质量好的，还可以花 600 元对原料的质量进行测试，这种测试可靠性如表 8.26 所示，试用 WinQSB 软件求后验概率。

表 8.25 （单位：万元）

原料质量 N_i	购新机器 A_1	改建老机器 A_2	维护老机器 A_3
N_1 好(0.3)	3	1.0	0.8
N_2 差(0.7)	-1.5	0.5	0.6

表 8 - 26

$P(Z_i \mid N_i)$		N_1 好	N_2 差
测试结果	Z_1 好	0.8	0.3
	Z_2 差	0.2	0.7

解：首先建立新问题，在图 8.7 中选择 Bayesian Analysis，输入标题“DP3”（读者也可以自己取名）和状态数 2 及试验指标 2。然后输数据，在图 8.14 中，第一行输入先验概率，第二行、第三行输入条件概率，对状态指标重命名。最后，计算后验概率，点击“Solve and Analyze” → “Solve the Problem”，得到图 8.15 所示的后验概率表。在 Results 下，点击“Show Marginal Probability” 显示边际概率，点击“Show Joint Probability” 显示联合概率，点击“Show

Outcome \ State	状态1：优质	状态2：劣质
Prior Probability	0.30	0.70
指标1：结果好	0.80	0.30
指标2：结果差	0.20	0.70

图 8.14

Indicator\State	状态1：优质	状态2：劣质
指标1：结果好	0.5333	0.4667
指标2：结果差	0.1091	0.8909

图 8.15

Decision Tree Graph”显示决策树。

四、马尔柯夫决策

马尔柯夫决策分析程序需要调用子程序 MarKov Process(MKP)。该程序在给定状态的一步转移概率矩阵、初始状态概率和各状态下的成本(收益)时,系统可以计算 n(n = 1,2,…)步后的状态概率、稳定状态概率、期望成本(收益)及绘制状态概率变化图。

例 8.16 有三家通信器材公司分别生产三种不同牌子的手机。各自开展广告宣传本公司的产品。各公司的市场比例是随时间变化的。顾客的流动服从马尔柯夫过程。已知初始市场占有率及转移概率矩阵为:

$$G(0) = (0,3,0.35,0.35) \quad P = \begin{bmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0.5 & 0 \\ 0 & 0.25 & 0.75 \end{bmatrix}$$

求:三家公司最终的市场占有率;第二期后三家公司的市场占有率。

解:首先建立新问题,输入标题“MKP1”(读者也可以自己取名)和状态数 3,如图 8.16 所示。然后依次输入一步转移概率矩阵和状态的初始概率,如表 8.17 所示。如果有状态成本指标输入成本数据。然后,点击“Solve and Analyze”后,下拉菜单有三个选项:Solve Steady State(求稳态时的状态概率)、MarKov Process Step(指定转移步数求状态概率)及 Time Parametric Analysis(参数分析)。选定参数分析时,指定某个状态、给定开始期数、最后期数、步长,系统给出各期的状态概率,还可以显示状态各期的概率变化图。

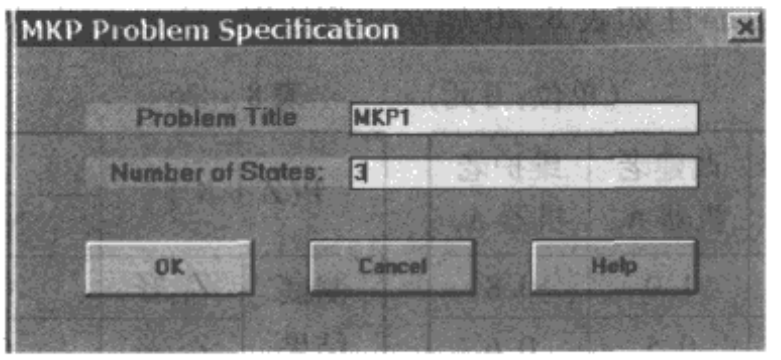


图 8.16

From \ To	State1	State2	State3
State1	0.5	0.25	0.25
State2	0.5	0.5	
State3		0.25	0.75
Initial Prob.	0.3	0.35	0.35
State Cost			

图 8.17

点击“Solve Steady State”,得到如图 8.18 所示的稳态时的状态概率表。结果表明,长期以后,三家公司的市场占有率各占 1/3。

点击“MarKov Process Step”,在图 8.19 的期数中输入 2,显示右边的结果,即二

期后三家公司的市场占有率分别为0.331 250,0.334 375,0.334 375。

04-13-2006	State Name	State Probability	Recurrence Time
1	State1	0.3333	3
2	State2	0.3333	3
3	State3	0.3333	3
	Expected	Cost/Return =	0

图 8.18

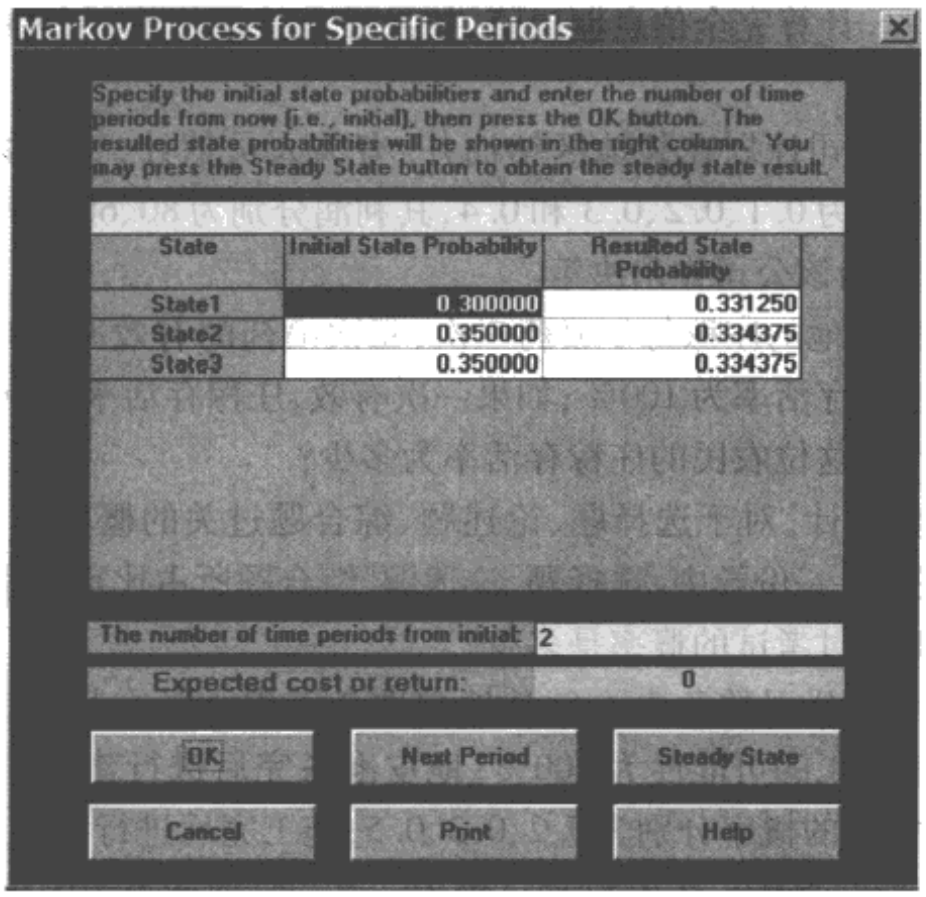


图 8.19

习题

1. 风险决策应具备什么条件?

2. 风险决策问题选择方案的标准有哪些?为什么期望值标准可以作为风险决策选择方案的标准?适合与不适合哪些情况?

3. 边际标准依据的原理是什么?计算出的概率的含义是什么?

4. 什么是先验概率与后验概率?离散变量后验概率的公式是什么?后验正态分布的均值与标准差是如何修正的?为什么后验均值用方差的倒数作权数?

5. 什么是完全信息期望值?有什么作用?如何计算?

6. 什么是决策树?什么情况下适用决策树进行决策?

7. 什么是完全风险决策?选择方案的标准有哪些?

8. 衡量效用的尺度有几种?效用曲线的基本形状有几种?

9. 什么是一阶马尔柯夫链?一阶马尔柯夫链有什么特点?

10. 齐次马尔柯夫链的合含义是什么?马尔柯夫链的遍历性是什么意思?

11. 用一阶马尔柯夫链预测市场占有率的基本假定有哪些?

12. 什么是市场占有率?市场占有率与投资利润率有什么关系?

13. 转移概率在市场占有率预测中用什么形式来反映?

14. 增加市场占有率的措施有哪些?

15. 福西斯公司在5号街上有一苹果摊,每蒲式耳进价为6美元,正常售价为12美元,当天卖不出的苹果能以每蒲式耳2元的价格处理。过去每天的需求量为5、6、7、8蒲式耳的天数分别为40、60、60、40天。试用期望利润值、期望后悔值、边际标准确定每天的最优进货量。计算完全信息期望值。如果没有处理价值,每天的最优进货量又是多少?

16. 李文斯顿公司计划在某地区钻油井。有关人员估计,钻到综合井、油井、汽井、废井的概率分别为0.1、0.2、0.3和0.4,其利润分别为80、60、40和 -50万美元。用期望后悔值标准为该公司作出决策。

17. 一位农民对他的庄稼每年喷药三次,每次喷药的有效率为0.5。如果喷药两次及以上有效,庄稼存活率为100%;如果一次有效,庄稼存活率为60%;如果全部无效庄稼存活率为零。这位农民的庄稼存活率为多少?

18. 一位学生估计,对于选择题、论述题、综合题过关的概率分别为0.9、0.6和0.75。过去经验表明,一份卷中,选择题、论述题、综合题所占比重分别为0.25、0.5和0.25。这位学生能通过考试的概率是多少?

19. 福斯特建筑公司的负责人估计地皮准备工作六月一日前完工的可能性为60%,七月一日前完工的可能性为100%;地皮准备完后进行建筑施工,建筑施工可能用1、2、3个月完工的概率分别为0.2、0.3、0.5;施工完后进行油漆,油漆可能用2、3个月完工的概率分别为0.7、0.3。则该工程十一月一日前能完工的概率是多少?

20. 德公司需要对500台机器上的轴承进行更新。统计表明,轴承在第1、2、3月的损坏率分别为0.4、0.3、0.3。大批更新费用为每台2美元,个别更新费用为10美元。为德公司推荐一更新策略,使月平均费用最少。

21. 凯斯公司有一条生产线,有效运转时间为80%,这时生产90%的合格品;无效运转时间为20%,这时生产40%的合格品。现抽两个产品均为合格品,这时生产线为有效运转的概率是多少?现抽两个产品均为不合格品,这时生产线为有效运转的概率是多少?

22. 考尔霍恩公司经营鲜花,某种鲜花进价为每枝3美元,正常售价为8美元,当天售不出的花没有处理价值。统计表明,需求量近似正态分布,平均每天售出3 600枝,标准差为800枝。公司每天的最优存货量应为多少?

23. 罗宾斯公司是一家专为大制造业研制新产品公司,现得知一企业准备招标研制一种"夏季产品",研制成功将获得60万美元的研制费。该公司面临着一种是否进行招标决策:若准备招标,需花4万美元的准备费,而中标的可能性只有0.4;若中标后,又面临用哪种方法研制,若用"标准方法"研制,需花费26万美元,研制成功的概率为80%;若用"修正方法"研制,需花费16万美元,但研制成功的概率为50%。用决策树分析该公司是否应准备进行招标?

24. 一位农民是蔬菜专业户，正准备在一块地上种夏季蔬菜。他估计天气可能出现旱、涝和正常三种情况，但不知道概率。现有三个品种供他选择：种甲品种，遇到旱、涝和正常的利润分别为500元、-50元、600元；种乙品种，遇到旱、涝和正常的利润分别为-100元、-150元、1 000元；种丙品种，遇到旱、涝和正常的利润分别为-80元、600元、400元。用完全风险决策的几种标准帮助这位农民进行决策。

25. A、B、C三个公司共同支配一个地区的通信市场，三公司的市场占有率分别为0.3、0.3、0.4。其转移矩阵如下：

$$\begin{array}{c} \\ A \\ B \\ C \end{array}\begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} 0.4 & 0.4 & 0.2 \\ 0.1 & 0.7 & 0.2 \\ 0.3 & 0.2 & 0.5 \end{bmatrix} \end{array}$$

若A公司能将丧失给B、C两公司的顾客保留一半的话，平衡状态时，A公司的市场占有率将增加多少？

26. 两自助食堂为一学校提供午餐。星期一，A食堂的营业额为60%，B食堂的营业额为40%。以往经验表明，某日在A食堂就餐的学生，次日有50%仍在A食堂就餐，其余的到B食堂就餐；某日在B食堂就餐的学生，次日有70%仍在B食堂就餐，其余的到A食堂就餐。预测星期二两食堂的销售份额各是多少？平衡状态时两食堂的销售份额各是多少？

第九章　管理决策模拟

引例:比尔洗车场决策问题

比尔洗车场是一个小型企业,它位于城市商业区的一条交通繁忙的单行道上。如图9.1所示,其设备的具体布局给作业上带来了一些限制条件,即在同一时间内只能洗刷一辆汽车,并且停车场只能停放两辆等待洗刷的汽车,沿街设立有“不准停车”、“禁止停留”等指示牌,禁止停放汽车。

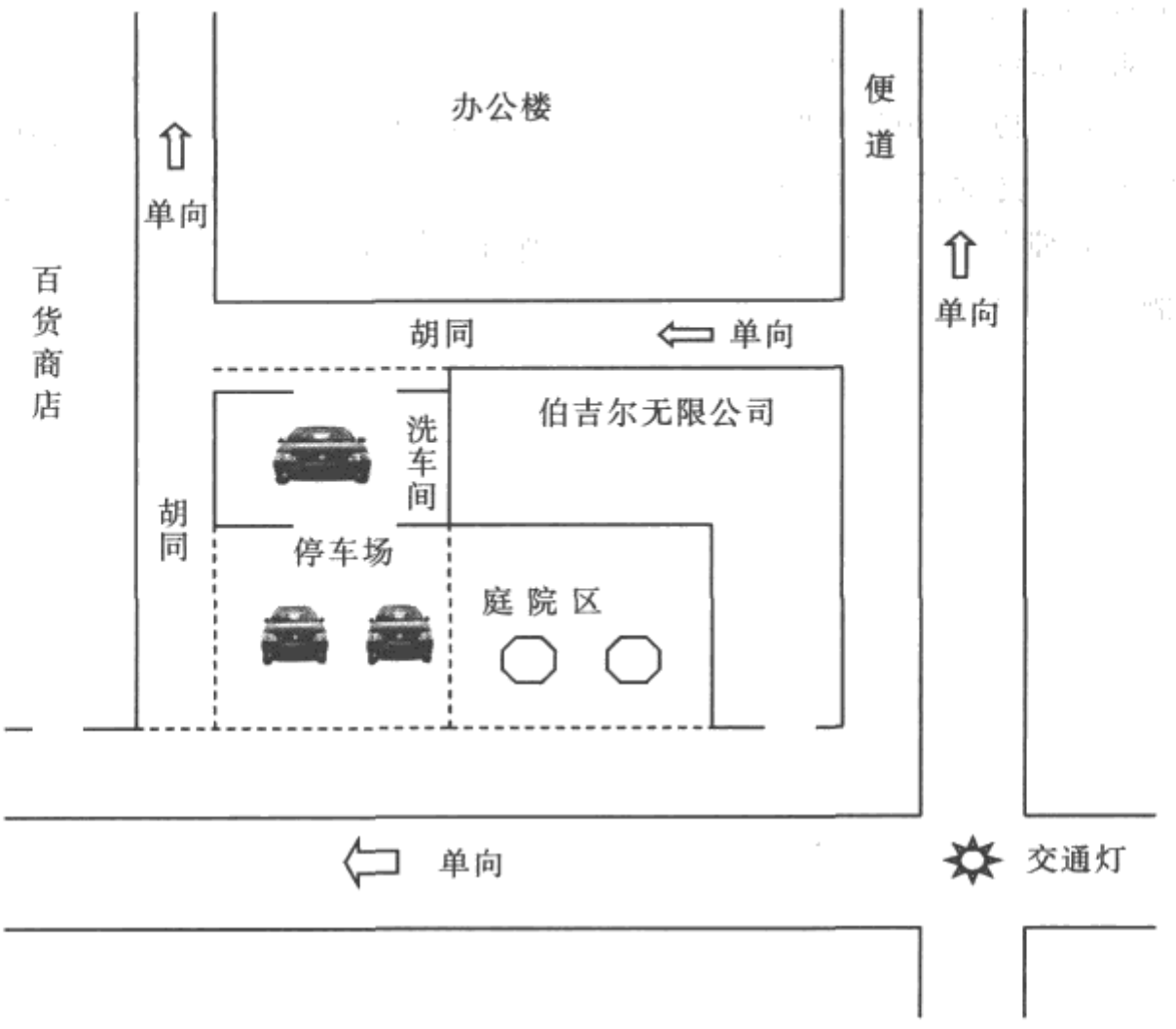


图9.1

比尔洗车场有一个原则,就是保证每个顾客在离开时,他的汽车是清洁的。由于采取这一原则,加上洗刷过程是由自动和人工作业综合而成的,这就使得洗刷一辆汽车所需的时间发生巨大的改变。决定洗刷时间的两个因素是:汽车的大小和汽车上堆积的污垢和灰尘多少。

企业主比尔·詹姆斯先生了解到,需要服务的顾客到达后,当他们发现停车场占满时,总是继续沿街而下,到另一个备有宽敞停车场的较大洗车场去。所以詹姆斯先生一段时间以来一直在考虑,试图对汽车洗刷作业作一些改革,以便把丢失给竞争者的一些业务夺回来。

詹姆斯先生考虑了几个方案，诸如安装新的高速洗刷设备以及租用邻近的庭院场地等。这块场地能存放两辆等待服务的汽车。伯吉尔无限公司的业主表示愿意出租这块空地给詹姆斯先生，因为他的顾客很少有人愿意冒着噪音和烟气的干扰去使用放在那里的野餐长凳。

但是詹姆斯先生在他确信哪些办法可以增加利润以前，是不愿做出什么改变的。他决定在许多可供选择的方案中，选出几个来评价其效果。詹姆斯先生在大学的课程中，学会了模拟的初步知识，他认为可以通过汽车洗刷业务的模拟来评价这些方案。

在前面几章中，我们讨论了用以帮助管理人员在制定决策过程方面的一些工具和技术。几乎所有这些都属于分析技术的范畴，即用它们可直接算出数字模型决策变量的最优值。然而，在一个分析模型中的情况往往是十分复杂的，以至无法列出其表达式，或者即使将表达式列出来了，也会因其复杂性而使求解十分困难。因此，用实际系统或该系统的模型进行实验也许是唯一可行的分析方法。

像在比尔洗车场出现的排队或等待问题，代表了一类常常借助于模拟模型进行分析的问题。这是因为如第七章中所指出的事实：许多复杂的排队问题不可能通过任何可用的排队模型来恰当地处理。

第一节　模拟过程

模拟是试图重现某一系统或活动的作业情况，而不必实地去建造并运转一个系统，那样会造成巨大的花费。它是一种试验的方法。分析人员首先对关心的系统或问题构造出一个模型，然后不是按照传统的观念去求解模型，而是在各种不同的作业条件的组合下，去运行这一模型。从这些模拟运行所得出的结果，可用以对所提出方案的优缺点进行评价。实质上，模拟给管理人员提供了一个进行试验的实验室，在这里一个问题可以在周密控制的条件下进行研究。

模拟可用来分析复杂的问题，同时其基本原理可为那些仅具备有限的数学基础的人员所理解和应用。模拟还有一些其他的优点，我们将在后面加以讨论。综合这些优点，就使模拟成为应用最广的管理科学技术之一。

本章将集中讨论具有随机的或概率性质的系统或问题，这些性质可以用蒙特卡罗模拟的形式进行分析。对大多数这类问题进行模拟，只需要纸和笔，但实际的应用则往往需要使用计算机。

模拟过程基本上可分为五个步骤：

(1)对所研究的问题或系统建立模型。

(2)为模型规定一组作业特性。

(3)设计与进行一系列试验，以供了解模型在指定条件下的运行情况。

(4)分析模型的运行结果，并对实际系统在类似作业条件下将如何运行进行推断。

(5)通常，还要规定一组不同的作业特性，对模型重新进行试验，即返回到第二

个步骤。这一计算过程,能对不同作业条件下的结果进行比较,以便帮助我们选择实际系统的作业条件。

第二节　建立模型

模拟过程中第一步是为所研究的过程或系统建立模型。通常这一模型是由实际中抽象出来的,模型应该足够的复杂以包含实际系统的各项重要特征。一个模拟模型通常比分析方法求解的模型所要求包含的内容更为详细,但模型应尽可能地简单。不必要的详细只能增加进行模拟所需要的时间与费用,并使输出数据的分析复杂化。

在模型构造方面,分析人员必须清晰地指出系统的各个部分,以及这些部分之间的关系。这样就可以使分析人员能把展开的模拟模型细分为若干段,以简化其结构。这样构造的模拟模型就可以预测实际系统各部分之间相互作用的结果。

一、流程图

把所研究的系统画成简图表示或流程图总是有帮助的。比尔洗车场的流程图如图 9.2 所示。第一个重要事件是汽车到达洗刷设施,然后,提出一个问题:“洗车场是否空闲?”如果回答“是”,则汽车直接进入洗车场。如果回答“否”,则提另一个问题:“停车场是否可用?”如果对第二个问题的回答“否”,则汽车的到达发生“障碍”或拒绝加入排队。假如回答“是”,则汽车加入等待线,并且等待进入洗刷设备。一旦汽车进入洗刷设备,即进行洗刷,然后退出系统。

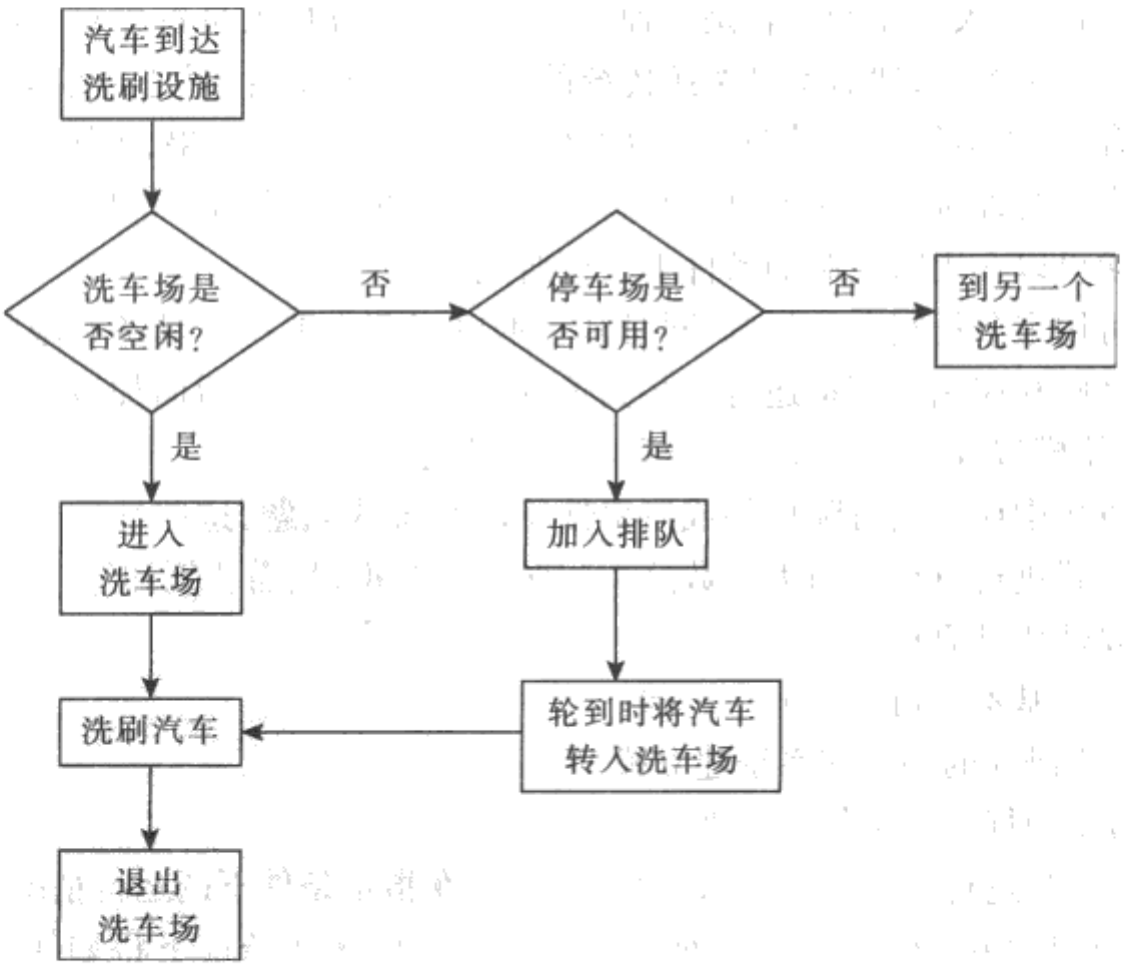


图 9.2

图 9.2 所示的模型十分简单，它包括了詹姆斯先生所面临的问题的主要特征，因此，它可以作为以后几个步骤的基础。

二、确定作业的特征

图 9.2 所示的比尔洗车场的流程图指出了几个必须确定的作业特征。我们所关心的事情主要包括：需要服务的汽车的到达方式、可供等待服务的汽车停留的车位数目以及洗刷服务过程本身的特性和持续时间等。

描述模拟到达过程通常所用的方法是确定各种到达间隔时间（连续到达之间的时间）的概率分布。詹姆斯先生让他的管理人员观察并记录了大量的连续到达之间的时间，确定了出现不同到达间隔时间的相对频率，并用它建立了概率分布表，如表 9.1 所示。

表 9.1

间隔时间（T_i 分）	概率（P_i）	期望值（T_iP_i）	间隔时间（T_i 分）	概率（P_i）	期望值（T_iP_i）
0	0.06	0.00	6	0.10	0.60
1	0.10	0.10	7	0.08	0.56
2	0.13	0.26	8	0.05	0.40
3	0.17	0.51	9	0.03	0.27
4	0.15	0.60	10	0.01	0.10
5	0.12	0.60		1.00	4.00

这一分布是经验分布，因为它是由观测数据直接决定的。到达间隔时间在 0 ~ 10 分钟之间变化，其平均间隔时间为 4 分钟。按这一平均数计算，每小时将到达 15 辆汽车（60 分钟除以平均间隔时间 4 分钟）。

注意，这一离散分布是由我们将每一观测数据化整到最接近的分钟数而得到的。因此，间隔时间 0 代表了连续到达的间隔时间在 30 秒及以下。必须注意所采用的这种化整方法，应保证不使模拟的结果产生偏差。我们可以采用较小的时间增量或者用连续分布来表示到达间隔时间，然而，一般的研究方法不论在哪一种情况下都是相同的，而且用离散分布更容易说明问题。

建立为顾客服务所消耗的时间的概率分布可用以描述服务时间。假定詹姆斯先生的管理人员也记录了汽车从开始洗刷直到它离开这一系统之间的消耗时间总数，这一工作必须对大量汽车反复的进行，从测得的洗刷时间（化整为最接近的分钟数）的相对频率，用来建立由表 9.2 给出的概率分布。

表 9.2

服务时间（T_i 分）	Ⅰ 型汽车(45%) 概率(P_i)	Ⅰ 型汽车(45%) 期望值(T_iP_i)	服务时间（T_i 分）	Ⅱ 型汽车(55%) 概率(P_i)	Ⅱ 型汽车(55%) 期望值(T_iP_i)
2	0.60	1.2	4	0.20	0.8
3	0.40	1.2	5	0.40	2.0
		期望服务时间 = 2.4	6	0.25	1.5
			7	0.10	0.7
			8	0.05	0.4
					期望服务时间 = 5.4

服务时间分两种不同型号的汽车给出。Ⅰ 型汽车是小型汽车，Ⅱ 型汽车是大型汽车。当然，两种分布可以合并成为一个单一的分布，但它们的差距应该保留下来，因为詹姆斯先生可能需要确定洗刷方式和定价标准，因而区别进入洗刷设施的汽车型号是有帮助的。

取样的汽车包括45% 的小型汽车（Ⅰ 型）和55% 的大型汽车（Ⅱ 型）。Ⅰ 型汽车的平均服务时间为2.4 分钟，而 Ⅱ 型汽车的平均洗刷时间为5.4 分钟，洗刷每一辆汽车所需总平均服务时间为4.05 分钟，即0.45 ×2.4 +0.55 ×5.4。因此，需要超过了能力，汽车洗车场必然会损失一些业务。在业务持续繁忙的期间，每小时平均洗刷汽车14.8 辆(60 分钟除以每车 4.05 分钟)，这一数字略低于每小时 15 辆汽车的平均到达率。

初始模拟是按等待线的容量为两辆汽车进行试验的，并根据先来先服务的原则洗刷汽车。

三、模拟模型的验证

从模拟得出实际系统在一定作业条件的范围内所具有的特性的结论以前，我们必须尽一切努力以验证模型的精确性，这在模拟过程中是最困难和具有挑战性的阶段之一。假如模型代表一个现在已有的系统，方法之一是在现实环境的条件下进行模拟，这样模型的输出可与实际作业系统中所收集起来的数据相比较，假如两者合理的一致，则可建立模型的可信度。

第三节　随机数的产生

为了试验一个模拟模型，需要想出一种方法产生输入信息以使模型得以运行。在许多情况下，输入信息是一些随机变量，也就是说，在某一给定的时刻，它们可选取值域内的任一数值，就要使用到随机数。所谓随机数(Random Number) 是指在两个极端值之间均匀分布(Uniform Distribution) 的数字。有关随机数的产生方法主要有三种，我们结合下面报童问题的例子来进行介绍。

例 9.1　假设有一个报童希望能够找到一个方法改善他订报的法则，以获得最多的利润。但是现在他不知道未来每天究竟会有多少顾客会买他的报纸，现在只根据

以往经验可知顾客每天的需求量及其对应的相对次数如表 9.3 所示。

表 9.3

需求量	150	160	170	180	190	200	
相对次数	1/12	2/12	4/12	3/12	1/12	1/12	12/12

现在可以用模拟的方法来产生随机数。

一、手工法(Manual Method)

此法采用抽签或抽球的方式,根据所要模拟的概率签条或者不同颜色的球分别代表不同的需求量来进行模拟,假设我们要两位数的模拟,可以在圆筒中放置 120 个签,其中 10 个写上 150,20 个写上 160,40 个写上 170,30 个写上 180,10 个写上 190 和 10 个写上 200。将所有签条搅和后,从中抽取一个,其所标明的数字即为当天的需求量,然后将此签放回,再抽第二次,即得第二天的需求量,以此类推。假设连续抽取十天,可能的次序如表 9.4 所示。

表 9.4

天数	1	2	3	4	5	6	7	8	9	10
需求量	180	170	200	180	170	160	170	150	170	180

如果我们把这十天的需求量做成次数分布表,求出次数比的概率分布,与表 9.3 所要模拟的概率分布比较,它们之间很可能有差异,但是如果模拟的日数增加,则模拟的结果必然会趋近于表 9.3 的概率分布,因为签条的安排配置方式就是按照所要模拟的对象而设计的,如果每个签条有同等被抽出的可能,则其结果应该相同。

二、随机数表(Random Number Table)

目前统计上已编有多种随机数表,其中普遍使用的是兰德公司(Rand Corporation)所发展出来的随机数表,我们在这里给出一个两位数的随机数表(表 9.5)。现在以表9.3 说明如何以随机数表产生取自该概率分布的随机观测值。设从 00 ~ 99,100 个数字中,若出现 00 ~ 07,8 个数字则表示需求量为 150 份,若出现 08 ~ 23,16 个数字则表示需求量为 160 份,若出现 88 ~ 95,8 个数字则表示需求量为 200 份。倘若出现 96 ~ 99 中任一数字,则弃而不用。依次可知,出现 150 份需求的概率为 8/96 = 1/12,出现 160 份需求的概率为 16/96 = 2/12,……,出现 200 份需求的概率为 8/96 = 1/12,一一对应与其相对次数,于是可得表 9.6。

表 9.5

33	24	52	87	13	31	14	53	65	35	02	76	07	62	93	67	23	93	42	16
50	72	85	56	18	51	49	20	94	94	06	43	09	07	51	70	88	54	35	75
13	19	79	96	61	23	74	91	76	76	17	84	97	48	48	80	77	34	90	29
82	20	86	44	47	63	04	98	43	43	32	33	63	46	79	66	60	33	70	97
59	91	72	29	60	07	04	83	73	73	70	95	41	55	44	20	07	28	93	97
30	88	20	80	29	98	80	68	52	52	55	91	46	92	56	92	57	78	33	63
24	95	12	56	03	08	83	06	15	15	62	57	59	41	90	31	90	56	73	29
02	38	21	96	23	78	87	31	54	54	30	14	18	10	08	79	38	98	35	86
15	41	99	86	67	63	04	76	94	94	06	97	79	66	68	03	21	95	38	21
38	51	58	80	61	85	21	26	52	52	45	33	21	56	21	88	83	65	29	48
12	08	04	33	62	78	49	42	61	61	15	22	03	98	17	69	41	82	45	92
85	23	36	43	13	37	21	50	09	09	96	91	17	31	62	90	86	94	58	31
92	55	01	88	68	65	65	08	96	96	94	86	06	29	28	30	66	52	76	36
79	27	84	90	59	40	21	83	87	87	88	93	63	34	10	54	73	56	35	67
59	80	13	96	77	11	15	89	47	47	28	61	22	45	41	66	88	09	96	62
11	26	06	05	73	01	49	45	69	69	61	47	39	71	66	12	28	67	52	16
97	54	15	28	63	60	82	89	02	02	10	73	20	47	08	55	52	99	90	86
39	47	73	96	60	18	56	00	98	98	17	68	62	06	10	66	60	11	90	25
71	14	64	64	28	82	47	36	87	87	48	60	37	90	15	80	54	43	61	93
16	59	96	95	85	66	88	02	37	37	88	98	85	02	90	02	13	60	60	96
77	87	27	72	76	79	15	68	23	23	46	80	88	57	41	10	33	15	47	86
78	15	18	02	65	23	14	08	54	54	62	61	13	23	43	01	23	29	77	62
40	37	69	32	79	84	37	50	76	76	23	54	60	23	72	17	56	86	67	11
83	43	17	25	54	99	29	15	13	13	64	94	50	87	51	76	82	18	73	00
28	71	96	61	27	84	81	27	99	99	34	53	34	21	74	53	95	01	11	35

表 9.6

需求量	相对次数	两位整数	概率
150	1/12	00 ~ 07	8/96
160	2/12	08 ~ 23	16/96
170	4/12	24 ~ 55	32/96
180	3/12	56 ~ 79	24/96
190	1/12	80 ~ 87	8/96
200	1/12	88 ~ 95	8/96
	12/12		96/96

三、计算机产生随机数

由于前面的产生方法颇为费时，在大规模的模拟时无法采用，近年来计算机的供应商们给出了很多套应用程序模块，只需要读进一个随机数的种子(Seed)，就可以由电脑事先设定的程序连续产生 0 ~ 1 之间的随机数。产生随机数的程序本身有多种计算方法，而最为广泛使用的是同余法(Congruential Method)，有加性、乘性和混合性三种。其中以混合性同余法(Mixed Congruential Method) 最为常见，特予说明如下：

设 x_0 为初始随机数(称为种子)，可为任何正整数，则后一随机数的产生根据前

一个随机数计算而来，其递推关系为 $x_{n+1} = (ax_n + c)(\mathrm{mod}m)$。

式中，a，c 与 m 皆为正整数，且 $a < m$，$c < m$。此式表示 x_{n+1} 为 $(ax_n + c)$ 除以 m 的余数，x_{n+1} 的可能值为 $0,1,\cdots,m-1$，这样最多可产生 m 个不同的随机数。往往对于 a 与 c 必须慎重选择，使数列 $x_1,x_2,\cdots,x_m$ 包含 m 个可能的随机数且每一数恰好出现一次。就二进制电子计算机而言，若其一字位(Word) 含有 b 个二元字节(Bit)，则通常令 $m = 2^b$，表示其在字位范围内非负整数的总个数。

一般而言，由 $m = 2^b$ 所产生的随机数，其位数较大，而应用模拟上所需要的随机数的位数可能较少。例如在某随机数表中每一随机数为五位数，而在模拟时，若只需要三位数，则每一随机数只取其中连续三位，其余的不用。惯例之一是用最前三位或最后三位。简单起见，我们使用两位数的随机数表，见表 9.5。

第四节　蒙特卡罗法(Monte Carlo Method)

蒙特卡罗法是一种随机取样的技巧，也就是一种依预定的概率分布而产生随机数的程序。蒙特卡罗模拟法是由纽曼命名的。在第二次世界大战期间，在洛沙拉莫斯科学实验室(Los Alamos Scientific Laboratory) 的物理学家被中子(Neutrons) 的动态所困惑，于是数学家纽曼和阿拉穆提议一种求解方法，逐步将个别事项依某种不规则的出现概率而求得近似的答案。因为在洛沙拉莫斯的工作要保持高度机密，纽曼将此种依人为的不规则概率分布而随机取样的求解程序予以暗号(Code Name) 蒙特卡罗(Monte Carlo)。因此，以后将此类依照某种预定概率分布而随机取样的模拟技巧称为蒙特卡罗(Monte Carlo)。

由以上的说明可知，蒙特卡罗法常用于某种视概率而定的问题的求解，对此种问题难以从事实际试验，也无法运用准确的数学公式加以分析，而且也不能从真实的母体中抽取样本。因而，运用蒙特卡罗法时，首先需要确定该变量的概率分布，然后应用随机数从此概率分布中获取有关变量特征的资料。所以事实上蒙特卡罗模拟是运用随机数产生与实际经验具有相同分布特征的一些变量，从而计算模拟问题。

一、随机数的利用

蒙特卡罗方法在概念上是简单的，且可以用一种简单的解释引出。假定抛一枚硬币，以确定两名选手中谁先发球。给定硬币的正面以及一个公正的抛掷机构，则每位选手取得先发球的机会是相等的。假定硬币的正面认为是选手1 的一次“成功”，硬币的背面认为是选手 2 的一次“成功”。首先取得五次成功的选手，判为竞赛中先发球者。

实际上，与其抛掷一枚硬币，不如用表 9.5 中给出的随机数去模拟硬币的抛数，以选择先发球的选手。取得 01 到 50 的数，表示硬币抛成正面，而取得 51 到 100 的数，表示抛成反面。我们随机的在任一点进入表 9.5，并按一预先确定的模式去选取随机数。假定第一个被选取的数是表中右上角的 16，而且我们预先确定的模式是逐个选取该列从上到下的数，那么，所选取得随机数列为 16、75、29、97 等。将这些数列入表

9.7 中。

表9.7 还给出了与随机数相关联的事件。第一个数为16,所关联的事件为硬币的正面,它意味着选手1 的一次成功。同理,76 关联着硬币的背面,意味着选手2 的一次成功。继续这一过程,直到一选手取得五次成功为止。

表 9.7

随机数	关联事件	获胜者
16	正面	1
75	背面	2
29	正面	1
97	背面	2
97	背面	2
63	背面	2
29	正面	1
86	背面	2

方法的基本思想是制定随机数序列,使得在给定的序列中取得一个数的概率正好等于该序列所关联的事件的概率。取得数 01 ~ 50 的概率为 0.5,它正好等于抛掷硬币得到正面的概率。一些补充例子将有助于弄清楚这一过程。

研究从 52 张洗匀的扑克牌中抽出一张的模拟。在每次抽出之后,将牌放回,并重新洗匀。抽得某一种花色的概率为 0.25。因此,可用随机数 01 ~ 25 表示黑桃,随机数列 26 ~ 50、51 ~ 75、76 ~ 100 则分别表示抽得红心、方块及梅花。

在有些情况中,需要将某些我们制定的随机数跳过去不用。假如我们模拟一粒骰子的重复抛掷,序列 01 ~ 10 可与抛成 1 点相关联,同样,序列11 ~ 20、21 ~ 30、31 ~ 40、41 ~ 50、51 ~ 60 可以分别与抛成2、3、4、5、6 点相关联。每当我们取得一个大于60 的数时,可以简单地将它忽略不计,接着再取下一个随机数。

最后,应该看到,可能需要大于两位数的随机数。假定我们是在模拟人出生的月份,这至少需要3 位随机数,序列 01 ~ 30 与一月相关联,32 ~ 59 与二月相关联,以此类推,直到 335 ~ 365 与十二月份相关联。更大的随机数可以将表 9.5 的两列或多列相组合而获得。

二、到达时间与服务时间的产生

上面所描述的过程,可以用来产生模拟洗车场所需要的输入数。在表9.8 中,对需要洗刷汽车的各种可能到达间隔时间,指定了一系列随机数。分配给某一给定事件的随机数个数,与随机数总数(它代表各种结果的总分布)的比值,等于该给定事件出现的概率。在表 9.8 中,到达间隔时间为 3 分钟的概率为 0.17,故分配随机数 30 ~ 46 代表 3 分钟的到达间隔时间。因为 30 ~ 46 有 17 个数,占总分配数的 100 个数的 17%,因此,取得该范围内的数的概率为 0.17。

表 9.8

到达间隔时间	概率	随机数
0	0.06	
1	0.10	
2	0.13	
3	0.17	
4	0.15	
5	0.12	
6	0.10	
7	0.08	
8	0.05	
9	0.03	
10	0.01	
	1.00	

按照表 9.8，随机数序列 36、47、82 及 13 将依次代表 3、4、6 及 1 分钟的到达间隔时间。抽取随机数的过程可以无限地继续下去，需要多少输入量就抽多少，直至完成模拟。

在产生到达汽车的服务时间之前，先要确定到达的是小型汽车还是大型汽车。我们也可以用随机数来确定，因为 Ⅰ 型车到达的概率是 0.45，故用随机数序列 01 ~ 45 代表这种车型，其余的数 46 ~ 00 代表 Ⅱ 型车，这一信息在表 9.9 中给出。

服务时间长度也能通过蒙特卡罗法产生。表 9.10 表示的是两种服务时间分布的随机数分配。Ⅰ 型车需要 2 分钟洗刷时间的概率是 0.6，需要 3 分钟服务时间的概率是 0.4。因此，对 Ⅰ 型车，随机数列 01 ~ 60 分配给代表 2 分钟的洗刷时间，其余 40 个数 61 ~ 00，分配给代表 3 分钟的洗刷时间。

例如，假定一辆汽车按到达要求洗刷。从随机数表中相邻的两个位置上抽读的数是 96 及 23，其中第一个数用来代表车型。因为 96 是在 46 ~ 00 的范围内，故它是 Ⅱ 型车；再参照表 9.10 中 Ⅱ 型车随机数的分配，由第二个数可以确定出汽车的洗刷时间，随机数 23 确定的洗刷时间为 5 分钟。

表 9.9

车型	概率	随机数
小型(Ⅰ 型)	0.45	01 ~ 45
大型(Ⅱ 型)	0.55	46 ~ 00
	1.00	

表 9.10

服务时间	Ⅰ 型汽车(小型)		服务时间	Ⅱ 型汽车(大型)	
	概率	随机数		概率	随机数
2	0.60	01 ~ 60	4	0.20	01 ~ 20
3	0.40	61 ~ 00	5	0.40	21 ~ 60
			6	0.25	61 ~ 85
			7	0.10	86 ~ 95
			8	0.05	96 ~ 00

第五节　模拟模型的完成与分析

至此,对于开始模拟运行的初始基础工作,大部分是完成了。现在,我们准备产生输入量,并按照模型中所规定的关系去处理它们。首先要运行的是模拟现行的汽车洗刷系统。

必须研究出一种方法来处理输入量,记录运行结果,追踪系统运行的状态等。在大型模拟研究中,这是通过计算机模拟模型来完成的。对于汽车洗车场的例子,这种模拟用一支笔几张模拟工作记录表就可以完成。

一、输入数据

对于 20 个不同顾客,为确定其到达间隔时间、车型及服务时间所用的随机数,列于表 9.11 中。这 20 个数就作为比尔洗车场的初始模拟到达者。表 9.11 中第 2 列的随机数是从表 9.5 的第 1 列中取得的。它们与表 9.9 中随机数的分配结合使用,产生出第 3 列中的到达间隔时间。表 9.11 中第 5 列及第 7 列中的随机数,由表 9.5 的第 2 列及第 3 列取得,它们与表 9.10 一起使用,以分别确定第 6 列和第 8 列中所表示的车型及服务时间。

表 9.11 第 4 列中给出的到达时刻,可由到达间隔时间计算出来。假定在时刻 0 汽车洗车场开始营业,第一个顾客到达间隔时间或者到达时刻,可以理解为从开始营业的时间到第一个顾客出现的时刻所经过的一段时间。对第一个顾客所产生的到达间隔时间是 3 分钟,因此第一个“到达”出现在时刻 3。

第二次到达所确定的到达间隔时间为 4 分钟,加上第一个顾客的到达时刻,就得到第二次的时刻为 7。对于顾客 3 和 4,到达间隔时间为 1 分钟和 6 分钟,由此可以得到这些顾客到达的时刻为 8 和 14。其余顾客的到达时刻,同样可以用每一顾客的到达间隔时间加上前一顾客的到达时刻得到。因为对于顾客 8 所产生的到达间隔时间为 0,所以顾客 7 和 8 的到达时刻是相同的。应该注意的是所有到达者均假定在该分钟末出现。

表 9.11

(1) 到达 序号	(2) 随机数	(3) 到达时 间间隔	(4) 到达 时刻*	(5) 随机数	(6) 到达 车型	(7) 随机数	(8) 服务 时间
1	33	3	3	24	Ⅰ	52	2
2	50	4	7	72	Ⅱ	85	6
3	13	1	8	19	Ⅰ	79	3
4	82	6	14	20	Ⅰ	86	3
5	59	4	18	91	Ⅱ	72	6
6	30	3	21	88	Ⅱ	20	4
7	24	2	23	95	Ⅱ	12	4
8	02	0	23	38	Ⅰ	21	3
9	15	1	24	41	Ⅰ	99	3
10	38	3	27	51	Ⅱ	58	5
11	12	1	28	08	Ⅰ	04	2
12	85	7	35	23	Ⅰ	36	2
13	92	8	43	55	Ⅱ	01	4
14	79	6	49	27	Ⅰ	84	3
15	59	4	53	80	Ⅱ	13	4
16	11	1	54	26	Ⅰ	06	2
17	97	9	63	54	Ⅱ	15	4
18	39	3	66	47	Ⅱ	73	6
19	71	5	71	14	Ⅰ	64	3
20	16	1	72	59	Ⅱ	96	8

* 到达者假定出现在该分钟末。

例 9.2 （罐子游戏）该游戏是对两个没有标记的罐子作随机选择，然后从所选的罐中取球。根据取得之球的颜色，给予奖赏。第一个罐子中有 5 个红球，2 个百球，3 个蓝球。第二个罐子中有 4 个红球，4 个白球及 2 个蓝球。试建立必需的随机数分配，并利用它们来模拟四次游戏。

解：

表 9.12

罐子	概率	随机数
1	0.5	01 ~ 50
2	0.5	51 ~ 00

表 9.13

罐子 1			罐子 2		
颜色	概率	随机数	颜色	概率	随机数
红	0.5	01 ~ 50	红	0.4	01 ~ 40
白	0.2	51 ~ 70	白	0.4	41 ~ 80
蓝	0.3	71 ~ 00	蓝	0.2	81 ~ 00

从表 9.5 中的第一列取出前 8 个数，并将它们分成 4 对两位数字的随机数。每对数中的第一个数用来决定选择哪一个罐子；第二个数用来决定取出的球的颜色。

33,50:罐子1,红球
13,82:罐子1,蓝球
59,30:罐子2,红球
24,02:罐子1,红球

二、模拟工作记录表

表9.11详细列出了洗车场首先到达的20辆车中每一辆的确切到达时刻和所用的服务时间。该表提供了模拟所需的全部输入数据。如表9.14所示的工作记录表称为模拟工作记录表,它用来记录系统的运行进程。我们将讨论模拟工作记录表中每一列的内容,并解释每项是如何计算的。但是,必须记住,这个工作记录表并没有标准的格式。表中每一列的安排,必须根据所模拟的问题特地加以设计。

表9.14

(1) 到达号	(2) 达到 时刻	(3) 到达 类型	(4) 服务 时间	(5) 到达 状态	(6) 是否加 入系统	(7)* 开始服 务时刻	(8)* 结束服 务时刻	(9) 排队 时间	(10) 空闲 时间
1	3	Ⅰ	2	0	是	4	5	0	3
2	7	Ⅱ	6	0	是	8	13	0	2
3	8	Ⅰ	3	1	是	14	16	5	0
4	14	Ⅰ	3	1	是	17	19	2	0
5	18	Ⅱ	6	1	是	20	25	1	0
6	21	Ⅱ	4	1	是	26	29	4	0
7	23	Ⅱ	4	2	是	30	33	6	0
8	23	Ⅰ	2	3	否	—	—	—	—
9	24	Ⅰ	3	3	否	—	—	—	—
10	27	Ⅱ	5	2	是	34	38	6	0
11	28	Ⅰ	2	3	否	—	—	—	—
12	35	Ⅰ	2	1	是	39	40	3	0
13	43	Ⅱ	4	0	是	44	47	0	3
14	49	Ⅰ	3	0	是	50	52	0	2
15	53	Ⅱ	4	0	是	54	57	0	1
16	54	Ⅰ	2	1	是	58	59	3	0
17	63	Ⅱ	4	0	是	64	67	0	4
18	66	Ⅱ	6	1	是	68	73	1	0
19	71	Ⅰ	3	1	是	74	76	2	0
20	72	Ⅱ	8	2	是	77	84	4	0

* 假定服务作业在该分钟初开始,服务结束在该分钟末。

表9.14的第2、3和4列包括了每辆汽车的到达时刻、汽车类型和服务时间的有关信息。这些信息是直接从表9.11中取得的。表9.14的第5列提供了新的到达者到达时系统的状态。系统状态就是该系统中的汽车数量,包括正在洗刷的汽车和在等待线上的汽车。这个汽车数目,可以通过确定有几个服务结束时间大于或等于该到达时刻来计算。例如,我们研究一下第6个到达者在时刻21到达时系统的情况。在表的第8列中,有一个服务结束时间大于21(25,对第5个到达者),因此,第6辆汽车的到达

状态就是1。同样,第8辆汽车的到达状态是3,因为有三辆先到达的汽车(5、6和7),其服务结束时间都大于23,即大于第8辆汽车的到达时刻。

第6列仅表示到达的汽车是加入本系统还是转到另一汽车洗刷场。如果到达状态是2或小于2,则在第6列内填上“是”,表示到达者决定在比尔汽车洗车场等待服务。如果到达状态是3,表示两个停车车位正被占用,则第6列内填入“否”,表示该到达者不加入本系统。

第7列表示何时开始服务。如果到达状态是0,则服务在汽车到达后立即开始。例如,第1辆汽车在第3分钟末到达,第4分钟初就开始服务。如果系统状态大于0,则在以前到达者中最晚的服务结束时刻之后立即开始服务。第6辆汽车在第21分钟到达,其到达状态是1,在以前到达者中,最晚的服务结束时刻是第25分钟,因此,服务开始时刻是第26分钟。

第8列表示服务结束时刻。这个时刻是服务开始时刻加上服务时间得到的。应该记住,服务开始时刻是由该分钟初开始计算,而服务结束时刻按该分钟末计算。因此,第1辆到达汽车的服务开始时刻是第4分钟初,服务时间是2分钟,表示服务结束时刻是在第5分钟末。

第9列给出了排队时间。它表示某一汽车在等待线上的时间。它可以从该汽车的“服务开始时刻”和“到达时刻”中求得。例如,第3辆汽车在第8分钟末到达(第9分钟之初),其服务开始时刻是第14分钟初,所以等待时间是5分钟。

第10列所指的是该系统从前一次到达至这一次到达之间所经过的总空闲时间。空闲时间只是在到达状态为0时才出现。从本次到达时刻减去前一辆车的出发时刻即可求得空闲时间。第14辆汽车是在第49分钟末到达的,前一辆车结束服务时间是在第47分钟末,所以该系统空闲2分钟。

三、工作记录表内容的说明

我们选择几辆汽车的到达项目加以讨论,以说明模拟工作记录表的应用。第1辆到达汽车出现在第3分钟末,由于系统此时是空闲的,所以这辆车可以加入系统,并立即进入洗刷设施。洗刷作业在第4分钟初开始,该车经过2分钟的服务时间,因此,服务结束时刻是在第5分钟末。该车是在第4和第5分钟内进行洗刷的。第1辆到达的汽车不必排队等待,所以在等待线上的排队时间为0,并列入第9列。在第一个顾客到达之前,该服务系统空闲了3分钟,因此3填入第10列中。

第4辆车到达是在第14分钟末,它所遇到的情况与第1辆车有所不同。当该顾客到达汽车洗车场时,第3辆车还在系统中,并将持续到第16分钟末。因此到达状态的项目内应记为1,这表明当第4辆车到达时,前一顾客还在系统内。但该顾客仍然加入系统,因为两个停车车位都可使用,须等到第17分钟初,才能开始服务。第四个顾客的等待时间是2分钟(即第15分钟和16分钟),将其填入第9列中。该顾客的服务时间是3分钟,所以结束服务时刻是在第19分钟末。在这辆汽车到达前,系统没有空闲时间,记入第10列中。

第8辆车到达的情况是非常有趣的。该系统已经饱和,其状态为3,第8辆车将不可能加入该系统,在第6列就表明了这一点。该系统处于饱和状态,其直接原因是由

于第5辆车延长了洗刷时间和后续车到达间隔较小。这一情况就使得系统丧失了第8辆、第9辆和第11辆车。

四、模型参数的变化

每一个有希望采用的系统或有价值的解决方案都应进行模拟，并且应建立一套评价系统有缺点的方法。这种处理方法的细节，在反复运用中是要发生变化的，以使他能够适合某一具体的情况。此外，在分析系统时，通常还应包括有形费用、无形费用以及利润等。

一旦初始的模拟运行完成以后，接着应改变模型的作业特点，并重复进行模拟。在比尔洗车场的例子中，我们将进一步研究将停车车位数目由2个增加到4个的效果。

表9.15是一个模拟工作的记录表，它表示在修改后的作业条件下，有关前20名到达顾客的计算结果。表9.11中的顾客到达时刻、汽车型号和服务时间等仍作为输入数据再次使用。因此，第二次模拟是在与初始模拟有相同的输入或需求条件下进行的，这是蒙特卡罗模拟的主要优点。由于可以利用精确的随机数序列或输入的数值，我们就能够在相同的需求条件下，对两个或多个可供选择系统的方案作出评价。这样做可以消除其他因素所引起系统特性差异的可能性，而这些差异并不是分析人员规定要改变的。从而就可以利用较小规模的抽样，在满足给定的统计标准的条件下，确定两个方案的特殊差异。

表9.15

(1) 到达号	(2) 达到 时刻	(3) 到达 类型	(4) 服务 时间	(5) 到达 状态	(6) 是否加 入系统	(7)* 开始服 务时刻	(8)* 结束服 务时刻	(9) 排队 时间	(10) 空闲 时间
1	3	Ⅰ	2	0	是	4	5	0	3
2	7	Ⅱ	6	0	是	8	13	0	2
3	8	Ⅰ	3	1	是	14	16	5	0
4	14	Ⅰ	3	1	是	17	19	2	0
5	18	Ⅱ	6	1	是	20	25	1	0
6	21	Ⅱ	4	1	是	26	29	4	0
7	23	Ⅱ	4	2	是	30	33	6	0
8	23	Ⅰ	2	4	是	34	35	10	0
9	24	Ⅰ	3	4	是	36	38	11	0
10	27	Ⅱ	5	5	否	39	43	11	0
11	28	Ⅰ	2	2	是	—	—	—	—
12	35	Ⅰ	2	2	是	44	45	8	0
13	43	Ⅱ	4	4	是	46	49	2	0
14	49	Ⅰ	3	3	是	50	52	0	0
15	53	Ⅱ	4	4	是	54	57	0	1
16	54	Ⅰ	2	2	是	58	59	3	0
17	63	Ⅱ	4	4	是	64	67	0	4
18	66	Ⅱ	6	6	是	68	73	1	0
19	71	Ⅰ	3	3	是	74	76	2	0
20	72	Ⅱ	8	8	是	77	84	4	0

* 假定开始服务时刻是由该分钟初算起，结束服务在该分钟末。

在新的作业特点情况下，特别重要的事实是：前 20 名到达的顾客中仅仅失去一名顾客。在第 8 个和第 9 个顾客在初始模拟中都是要失去的，而现在可以进入该系统了。第 11 个顾客在第 28 分钟到达系统，此时第 6 个顾客正在接受服务，并且四个停车场车位都被另外四名顾客所占用，因此它仍不能进入该系统。同时，顾客的等待时间增加，前 20 个到达的顾客花费在排队等待上的总时间为 70 分钟，这将近比第一次模拟的排队总时间 37 分钟多一倍。尽管有这些差异，在第 20 名顾客到达时，该系统所处的状态是相同的。

五、方案比较

在比尔洗车场的例子中，方案的评价是十分简单的，它只要确定由于增加了洗刷汽车的数目所得到的收益，能否补偿改变服务系统所支出的费用。

假定每辆汽车的洗刷费用为 3 美元，租用两个车位的停车场，其年租金换算为每营业小时的费用是 0.10 美元。通过减少服务设施的空闲时间可以增加洗刷汽车的数目，而洗刷这些车辆只需要直接开支，即包括水、洗浆剂和其他消耗品的费用。詹姆斯先生估计这些费用平均每辆汽车为 0.25 美元，因此，洗刷的汽车每增加一辆，就可提供 2.75 美元来补偿停车场的租金。只要收入超过他的支出就是有利的。增加停车车位的措施，可以增加每小时洗刷的汽车数近似为 0.8 辆，故每小时的期望的利润约为 2.10($0.8 \times 2.75 - 0.10 = 2.10$) 美元。因此，停车车位应该由两个增加到四个。尽管利润增加 2.10 美元的估计会有统计误差，但是这种增加幅度对于确保詹姆斯先生利润的增长是足够的。

排队时间的急剧增加，可能影响到顾客是否愿意参加排队的决定，这是一个潜在的困难，它可能使得上述结论无效。某些顾客在他们到达汽车洗刷场时，该系统已有三辆汽车正在等待，尽管这时还有一个停车车位可以使用，他们也可能决定到另一个洗车场去。顾客的这种行动会使增加停车车位的有利性降低。如果詹姆斯先生认为这种情况是可能的话，他就应该力图对这种影响进行定量的分析。但这是一项困难的工作。

第六节　模拟分析的例子

模拟方法可以而且已经应用于很多不同的场合。模拟对于排队问题的应用已经被证明是有效的，正如前面的例子所展示的一样。模拟还可以运用于库存控制、生产进度安排、仓库配置、装配线的平衡、有价证券的分析、经理人员训练以及其他情况。此外，模拟还可能有其他的用途。例如描述一个现有的系统或者设计一个新的系统。在本节中我们将就模拟的另外三个领域，列举一些简单的例子。

一、库存问题的模拟

某公司希望制定一个库存策略，即确定对某一特定产品的订货数量和订货点。这个公司面临两个方面的变化：需求量的变化以及订货与收到订货之间所需时间（交

货期限）的变化。日需求量的概率分布如表 9.16 所示。交货期限的概率分布如表 9.17 所示，其变化范围为 1 ~ 6 天。每天每件商品的存储费用为 5 美元，脱销费用或丢失销售的费用为每件 30 美元，安排一次订货的费用为 80 美元。该公司希望制定一个库存策略，使得库存的存储费、脱销费用和订货费用的总和最小。

表 9.16

需求量件数	概率	随机数
8	0.13	01 ~ 13
9	0.19	14 ~ 32
10	0.34	33 ~ 66
11	0.23	67 ~ 89
12	0.11	89 ~ 001
	1.00	

表 9.17

交货天数	概率	随机数
1	0.14	01 ~ 14
2	0.34	15 ~ 48
3	0.20	49 ~ 68
4	0.12	69 ~ 80
5	0.10	81 ~ 90
6	0.10	91 ~ 00
	1.00	

解决这个问题的一个方法，是利用我们在第六章中讨论过的概念 —— 经济订货量和期望值的分析，来建立一个策略。建立策略的方法之一是采用模拟，通过对不同的订货量和订货点进行试验可以得到模拟的结果。为了产生我们所需要的模拟输入量，其随机数的分布列入表 9.16 和 9.17 中。实际采用蒙特卡罗方法产生的随机变量的输入数值如表 9.18 的模拟工作记录表所示。

对于库存问题的分析，其模拟时间的推移是采用固定增量的方法。我们将时间每次前进一天而不是前进到下一个有关事件发生的时刻，后者是我们前面已经采用过的可变增量的方法。选择一天作为时间增量，必须假定订货总是安排在一天的终了，而新运到的货物的供应则在一天的开始。

公司已决定首先模拟这样一个策略，即每当库存下降到 30 件时就安排订货 60 件。在模拟开始时，库存是 60 件。在第三天末库存已经下降到订货点 30 件以下，因此要安排订货 60 件。随机数 54 用以确定这次订货的交货时间，由表 9.17 查得为三天。这次订货看成是三天以后收到，即在第六天末收到订货，它可以保证第七天的供应。因此，收到的 60 件订货是在第七天加入库存。在收到订货以前，库存已经用尽，并出现脱销，在第六天销售损失为 2 件。

这个模拟过程延续了 40 天。在这个经营周期内存储费是 480.50 美元，脱销费用是 1 410 美元，六次订货总共的订货费用为 480 美元，由此得出日平均费用为 59.26 美元。对于制定决策的目的来说，模拟持续时间有必要做相当的延长。

对于若干个不同的库存策略都应该重复进行模拟，并对各个方案的费用进行比较。根据第一次试验的有限结果，公司可能得出结论：脱销费用过高。于是将改变其初始策略，以减少脱销费用。为此可将订货点提高为 30 件以上。原来的订货点取为等于交货期限内的期望需求量。由表 9.17 和 9.18 中给出的概率分布可算出每日期望需求量为 10 件，期望交货期限是 3 天，故在这段时间内的期望需求量是 30 件。提高订货点将增加库存的存储费，但这部分增加的费用将会由减少脱销费用所补偿。

表 9 – 18

日期	收货	开始库存	随机数	需求量	终结库存	销售损失	随机数	交货日期
1		60	83	11	49			
2		49	70	11	38			
3		38	59	10	28*		54	3
4		28	12	8	20			
5		20	92	12	8			
6		8	46	10	0	2		
7	60	60	54	10	50			
8		50	04	8	42			
9		42	51	10	32			
10		32	99	12	20*		90	5
11		20	84	10	10			
12		10	81	11	0	1		
13		0	15	9	0	9		
14		0	36	0	0	10		
15		0	12	8	0	8		
16	60	60	54	10	50			
17		50	97	12	38			
18		38	00	12	26*		71	4
19		26	49	10	16			
20		16	44	10	6			
21		6	13	8	0	2		
22		0	23	9	0	9		
23	60	60	45	10	50			
24		50	54	10	40			
25		40	24	9	31			
26		31	50	10	21*		08	1
27		21	29	9	12			
28	60	72	61	10	62			
29		62	22	9	53			
30		53	47	10	43			
31		43	73	11	32			
32		32	18	9	23*		51	3
33		23	22	9	14			
34		14	38	10	4			
35		4	34	10	0	6		
36	60	60	54	10	50			
37		50	91	12	38			
38		38	42	10	28*		73	4
39		28	46	10	18			
40		18	30	9	9			
				961	47			

注：* 表示已到达订货点

二、失效时间的模拟

哈特门公司研制了一种新的电子系统，并希望对新建立系统的无修理运行的平均持续时间作一个估计。此系统由三个标准装配而成。当这些部件中有一个失效时，该系统就失效并进行修理。

根据系统实际运行情况来做出这个估计，则费用太高，所需的时间太长。该公司已经收集了每一部件的各种运行记录，这些记录已用于确定部件失效时间的概率分布，如表 9.19 所示。此分布可用来直接计算新系统失效时间的概率分布。另一种方案是大量地模拟系统的故障，作为估计所研究的分布的一种方法。

表 9.19

A 部件			*B* 部件			*C* 部件		
失效时间	概率	随机数	失效时间	概率	随机数	失效时间	概率	随机数
4 个月	0.10	01 ~ 10	2 个月	0.05	01 ~ 5	6 个月	0.20	01 ~ 20
5	0.20	11 ~ 30	3	0.10	06 ~ 15	7	0.30	21 ~ 50
6	0.30	31 ~ 60	4	0.20	16 ~ 35	8	0.25	51 ~ 75
7	0.20	61 ~ 80	5	0.30	36 ~ 65	9	0.15	76 ~ 90
8	0.15	81 ~ 95	6	0.25	66 ~ 90	10	0.10	91 ~ 00
9	0.05	96 ~ 00	7	0.10	91 ~ 00			

随机数的分布如表 9.19 所示，利用它可以产生出三种部件的失效时间。表 9.20 是模拟工作记录表，表示了 20 台该种电子装置运行间隔或失效时间的模拟。利用表 9.5 中的随机数，参照表 9.19 中给出的随机数分布，可以确定出三种部件（*A*、*B*、*C*）的失效时间。在第一种情况中，*A*、*B*、*C* 部件的失效时间分别求得为 6、4 和 8 个月。*B* 部件的最早失效时间（4 个月）决定了整个系统的失效时间。根据对 20 个系统的模拟，得出系统平均失效时间是 4.6 个月。根据模拟结果的相对频率，模拟应进一步扩展使能提供一个更精确的运行时间估计。

表 9.20

系统	随机数	*A* 部件失效月数	随机数	*B* 部件失效月数	随机数	*C* 部件失效月数	系统失效月数
1	33	6	24	4	52	8	4
2	50	6	72	6	85	9	6
3	13	5	19	4	79	9	4
4	82	8	20	4	86	9	4
5	59	6	91	7	72	8	6
6	30	5	88	6	20	6	5
7	24	5	95	7	12	6	5
8	02	4	38	5	21	7	4
9	15	5	41	5	99	10	5
10	38	6	51	5	58	8	5
11	12	5	08	3	04	6	3
12	85	8	23	4	36	7	4
13	92	8	55	5	01	6	5

表 9.20(续)

系统	随机数	A 部件失效月数	随机数	B 部件失效月数	随机数	C 部件失效月数	系统失效月数
14	79	7	27	4	84	9	4
15	59	6	80	6	13	6	6
16	11	5	26	4	06	6	4
17	97	9	54	5	15	6	5
18	39	6	47	5	73	8	5
19	71	7	14	3	64	8	3
20	16	5	59	5	96	10	5

三、风险分析的模拟

蒙特卡罗模拟方法也常用于从事投资风险的分析,以下将举例说明如何用于分析推出新产品的投资问题。此项投资的盈利能力需视下列因素而定,即该项产品市场的预估、本公司可能获得的市场占有率、市场的成长率、生产成本、销售价格、产品的寿命以及所需要的设备成本等,但是以上这些因素都只能做不确定的估计。

通常所用的方法是对上述各项不确定因素作一个较好的单一估计,然后再计算获利能力,例如投资净现值或者报酬率等。用此种单一估计的方法分析投资效益有下面两个缺点:

(1) 用单一估计所获得的预期利润,可能与实际情形不相符。

(2) 无法测度投资的风险,即不能确定投资计划的获利或发生亏损的概率。

风险分析既是用于克服以上两个缺点的一种运算技巧。通常是先对每个不确定因素设定一个主观的概率分布,然后再运用蒙特卡罗法,将这个概率分布结合成以投资计划获利性为整体的一项概率分布。现在用一个简单的例子说明如下:

假定推出某种新产品需要投资固定设备 5 000 美元,有三种不确定的因素需要考虑,即售价、变动成本及销售量。此项产品的寿命只有一年。表 9.21 中含有上述不确定因素的各种情况及其各自可能发生的概率。为了便于预算,设定表中各因素在统计上是独立的,否则,将因概率的相依而需要对模拟过程加以修正。

表 9.21

售价(美元)	概率	变动成本(美元)	概率	销售量(单位:千)	概率
4 000	0.3	2 000	0.1	3 000	0.2
5 000	0.5	3 000	0.6	4 000	0.4
6 000	0.2	4 000	0.3	5 000	0.4

在这个简单的例子中,只有 $3 \times 3 \times 3 = 3^3 = 27$ 种可能的不同结果,运用决策树(Decision Tree)的分析方法也可以计算出各种结果和概率。但是在较复杂的实际问题中,可能含有较多的不确定因素,而每种因素又可能有 10 种或 20 种不同情况,整个问题即可能有上百万的可能结果。在此种情形下,将模拟的技巧应用于投资计划的平均获利能力及其风险的估计则非常有效,可以用实现各种不同利润水平的概率来计算。

为了能够依照表9.21中不确定因素的各种情况所列出的概率而产生随机数，像表9.19一样，可按各种结果可能出现的概率而设定若干相对应的随机数，如下表9.22中所列。

表9.22

售价(美元)	随机数	变动成本(美元)	随机数	销售量	随机数
4 000	0 ~ 2	2 000	0	3 000	0 ~ 1
5 000	3 ~ 7	3 000	1 ~ 6	4 000	2 ~ 5
6 000	8 ~ 9	4 000	7 ~ 9	5 000	6 ~ 9

现在可以依照表9.22中所设定的与各不确定因素相对应的随机数，从随机数表9.5中选出一位的随机数，再依次决定售价、成本及销售量。这几个因素一旦决定，便可运用下式计算利润，即

利润 =（售价 - 成本）× 销售量 - 固定费用

重复运用此式多次，即可产生许多种不同的利润情况，表9.23是重复模拟25次的结果。25次的小型取样当然不足以完全正确的估计出平均盈利能力，如果将此种模拟过程设计成程序，而交由计算机去做，可以很容易的模拟几百次，甚至上千次，即可获得较正确的结果。表9.23是这25次的模拟结果。为方便说明，下面的讨论都基于表9.23。

表9.23

模拟次数	随机数	售价（美元）	随机数	成本（美元）	随机数	销售量（千）	利润（美元）
1	8	6 000	0	2 000	6	5	15 000
2	0	4 000	4	3 000	3	4	- 1 000
3	6	5 000	2	3 000	2	4	3 000
4	1	4 000	4	3 000	0	3	- 2 000
5	3	5 000	6	3 000	0	3	1 000
6	5	5 000	6	3 000	9	5	5 000
7	1	4 000	6	3 000	7	5	0
8	3	5 000	8	4 000	8	5	0
9	2	4 000	8	4 000	6	5	- 5 000
10	1	4 000	6	3 000	1	3	- 2 000
11	5	5 000	7	4 000	3	4	- 1 000
12	9	6 000	9	4 000	6	5	5 000
13	4	5 000	9	4 000	7	5	0
14	7	5 000	2	3 000	6	5	5 000
15	9	6 000	5	3 000	3	4	7 000
16	0	4 000	5	3 000	0	3	- 2 000
17	1	4 000	1	3 000	8	5	0
18	0	4 000	6	3 000	4	4	- 1 000
19	8	6 000	8	4 000	6	5	5 000
20	9	6 000	2	3 000	4	4	7 000

表 9.23(续)

模拟次数	随机数	售价(美元)	随机数	成本(美元)	随机数	销售量(千)	利润(美元)
21	0	4 000	7	4 000	7	5	-5 000
22	0	4 000	0	2 000	8	5	5 000
23	4	5 000	0	2 000	1	3	4 000
24	6	5 000	5	3 000	8	5	5 000
25	4	5 000	0	2 000	1	3	4 000
	平均						2 080

由表中的资料可以计算出,平均利润是 2 080 美元,将此项结果和简单的分析方法所获得的结果可以做一个比较。若用各不确定因素的最可能值为单一估计值,则估计利润为:

$$\text{最可能利润} = (5\,000 - 3\,000) \times 4\,000 - 5\,000 = 3\,000$$

由此可知,单数估计的结果比模拟所得的预期利润差很多。

在这个简单的例子里,可以将 27 种结果各自依概率加权而计算出预期利润,此项利润为 2 140 美元。由此可知,模拟 25 次所得的平均利润与计算得到的预期利润并不完全相同,但是预期利润的计算无法得到投资计划所附带的风险,然而模拟的 25 次结果却清楚地显示出在何种情况下会发生亏损,例如在表 9.23 中,第二、第四、第九和第十等各次所模拟的结果。为了便于显示在风险分析中所获知的风险程度,可依表 9.23 中获利能力而排列层次,用图解来表示样本的累积概率密度函数,如图 9.3 所示。

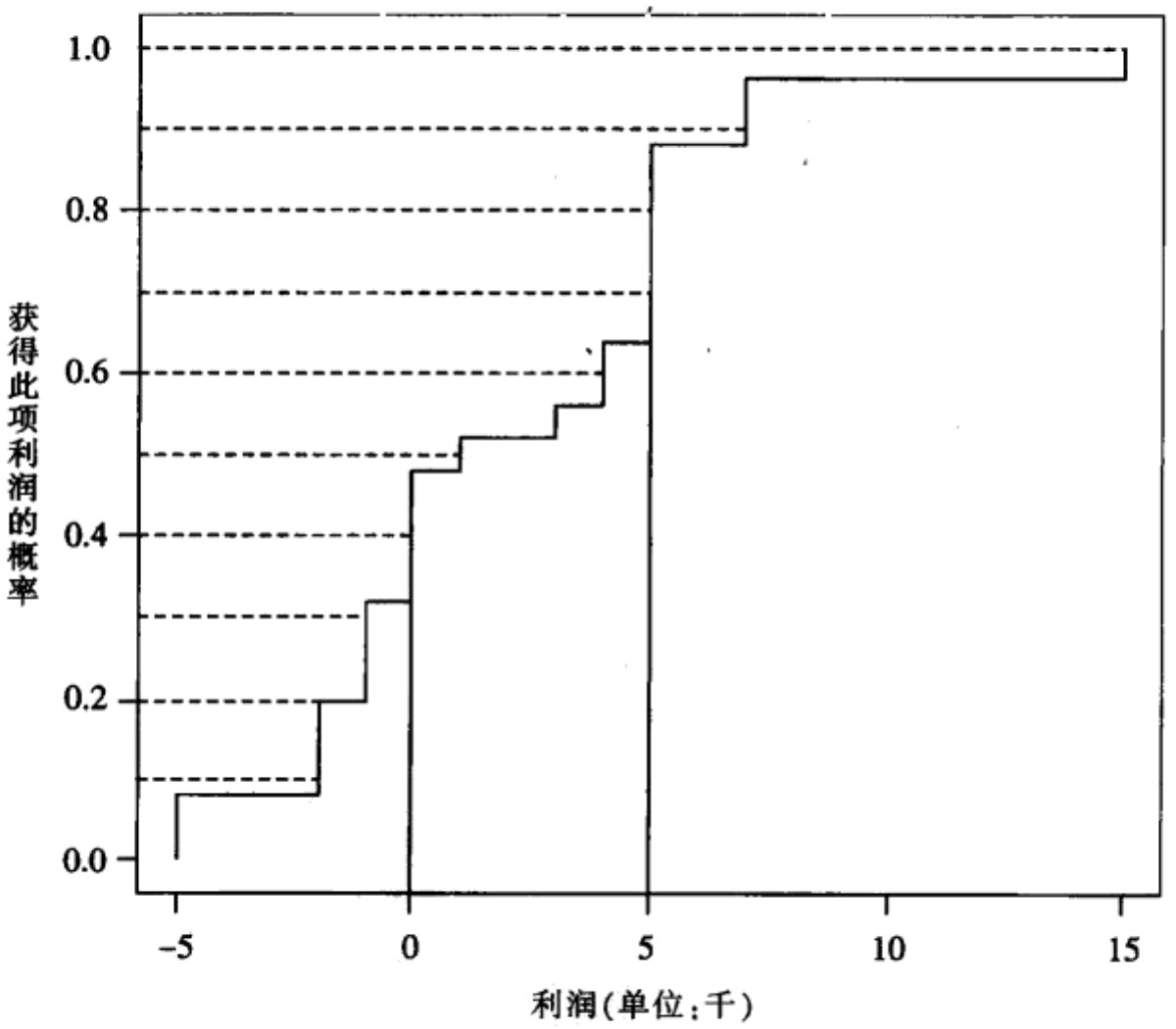


图 9.3

从图9.3中可以看出,有68% 的机会可以获得利润(有32% 的机会将蒙受损失),有36% 的机会可以获得利润5 000或5 000美元以上,获利最多为15 000美元,其可能性为4%。若模拟的次数非常多,图9.3中的阶梯线段将逐渐趋近光滑,但并不能成为连续的曲线,因为各种不同的结果是间断的。

第七节　模拟的优点和局限性

在问题分析中广泛地使用模拟,反映了模拟方法具有一系列的优点,其中一些优点前面已提到过,在本节内容中我们作进一步的补充。

模拟可以用来研究大型复杂的项目,这些项目包含着随机性的多种因素,这是使用模拟的主要原因。此外,往往由于选择不出有效的分析方法,而存在这样一个问题:是采用模拟还是不采用模拟去研究它。唯一的另一种方法,就是用实际的系统进行试验,不过这往往是不切合实际的或者说是不可能的。模拟具有强大的功能,其主要原因之一,就是我们可以把更详细、更接近实际的内容纳入模拟模型,这一点通常是可求解的分析模型办不到的。

模拟为分析人员和管理人员提供了一个"实验室",它使得分析人员和管理人员在相当短的时间内,能够预见到所研究的系统在一定运行条件下的特性。它还使得他们能够研究单个变量或参数的变化对现有系统的影响。在比尔汽车洗车场模拟的第二个方案中,我们可以了解到停车车位从两个变到四个时该系统特性的任何差异,而其他部分仍保持不变。这一点在实际系统中是不可能的,因为在同一时间内会发生各种变化,因此要把一个给定的改变措施的效果分离出来是困难的,或者说是不可能的。此外,模拟允许分析人员用某一特定变量和参数的相同数值,去充分试验,这在实际系统中也是不可能的。

模拟的一个更大的优点就是其基本方法简单,易于领会。模拟得到的结果易于为管理人员理解,而是否采纳和应用这些结果是由他们决定的。

遗憾的是,模拟除了具有通常有关模型构造的缺点外,还有某些其他的缺点,所以当我们考虑使用模拟时,必须权衡这些缺点。切记模拟模型是运行的,而不是被求解的。根据精心设计的模拟研究所选择的系统方案,只是已经考虑过的各种方案中的最佳方案。然而事实上不可能对所有可能的系统都考虑一遍,所以无法知道所选择的方案是否为最有效。与数学求解相反,模拟模型的一个根本的缺点,就是不能保证得到模型的最优解。因此,分析人员必须对那些选择方案的人的能力具有相当的信心。他们的经验和洞察力,应使得所选择的方案即使不是最优的,也应该是好的系统。然而,系统复杂性增加时,这个任务就变得特别困难。

模拟的第二个主要缺点是对系统特性估计的不精确性。这些估计存在有统计误差。即使在设计得最好的试验中,要确定哪一个研究方案为最佳方案,也是困难的或是不可能的。其中的部分困难是由于还没有完全研究出一种恰当和易于应用的统计方法,以分析和比较模拟得出的结果。

模拟处理方法的另一个重要缺点,是研制模型和进行模拟的费用高。为了分析复

杂的系统需要研究计算机程序，其费用可能是非常大的。进行模拟所需要的计算机时间的费用支出也会限制所使用的模型的复杂程度。考虑到费用问题，就可能迫使构造模拟模型的人员采取与分析模型中同样的方式，对模拟模型的假设条件进行简化。

模拟所具有的这些缺点，意味着分析人员应该把模拟看成分析方法的备用，而不能用模拟代替分析方法。如果有一个分析方法适合于采用，就应该采用它，而不是去做模拟。分析人员必须防止出现把模拟当作分析工具，而过分依赖于模拟的倾向。只有经过仔细研究，确定不能采用分析方法时，才应使用模拟。

附录 A　WinQSB 软件操作指南

A.1　WinQSB 软件简介

QSB 是 Quantitative Systems for Business 的缩写，早期的版本是在 DOS 操作系统下运行的，后来发展成为在 Windows 操作系统下运行的 WinQSB 软件，目前已经有 2.0 版。该软件是由美籍华人 Yih - Long Chang 和 Kiran Desai 共同开发的，可广泛应用于解决管理科学、决策科学、运筹学及生产管理等领域的问题。该软件界面设计友好、使用简单，使用者很容易学会并用它来解决管理和商务问题，表格形式的数据录入以及表格与图形的输出结果都给使用者带来极大的方便，同时使用者只需要借助于软件中的帮助文件就可以学会每一步的操作。

A.2　WinQSB 的一般操作

1. 安装与启动

点击 WinQSB 安装程序“Setup”，指定安装目录后，软件自动完成安装。读者在使用该软件时，只需要根据不同的问题，调用程序当中的不同模块，操作简单方便。进入某个模块以后，第一项工作就是建立新问题或者打开已经存盘的数据文件。在 WinQSB 软件安装完成后，每一个模块都提供了一些典型的例题数据文件，使用者可以先打开已有的数据文件，了解数据的输入格式，系统能够解决什么问题，结果的输出格式等内容。例如，打开线性规划文件“LP. LPP”，系统显示如图 1 的界面。

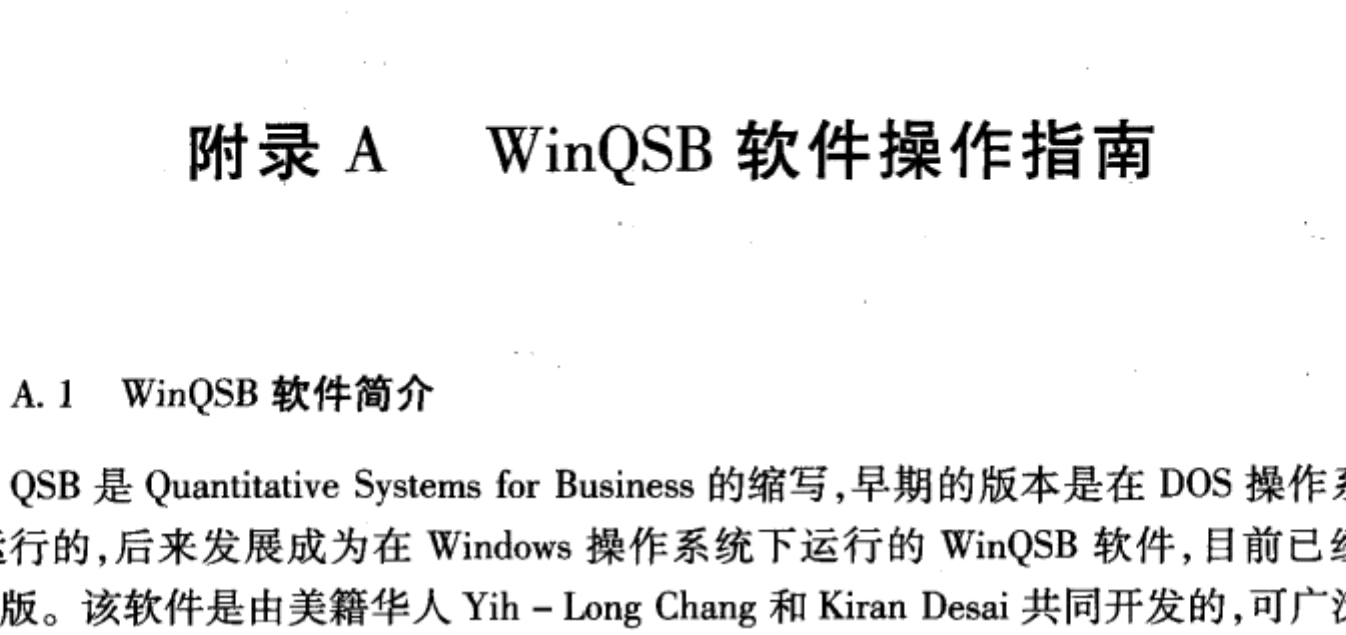


图 1

2. 数据的录入与保存

数据的录入可以直接录入，同时也可以从 Excel 或 Word 文档中复制数据到

WinQSB。首先选中要复制的电子表格中单元格的数据,点击复制,然后在 WinQSB 的电子表格编辑状态下选择要粘贴的单元格,点击粘贴即可。

如果要把 WinQSB 中的数据复制到“Office”文档中,选中 WinQSB 表格中要复制的单元格,点击“Edit”→“Copy to clipboard”即可。

数据的保存,只需要点击“File”→“Save as”即可,计算结果的保存亦相同,只是注意系统以文本格式(*. txt)保存结果,使用者可以编辑该文本文件。

A.3 WinQSB 的求解模块

关于 WinQSB 的各种模块及其功能,我们在表 1 中给出详细的说明。

表 1

序号	模块	文件名	含义	应用范畴
1	Acceptance Sampling Analysis	ASA	抽样分析	各种抽样分析,抽样方案设计、假设分析
2	Aggregate Planning	AP	综合计划编制	具有多时期正常、加班、分时、转包生产量,需求量,存储费用,生产费用等复杂的整体综合生产计划的编制方法。将问题归结为线性规划模型或运输问题。
3	Decision Analysis	DA	决策分析	确定型与风险型决策、贝叶斯决策、决策树、二人零和对策、蒙特卡罗模拟
4	Dynamic Programming	DP	动态规划	最短路问题、背包问题、生产与存储问题
5	Facility Location and Layout	FLL	设备场地布局	设备产地设计、功能布局、线路均衡布局
6	Forecasting	FC	预测	简单平均、移动平均、加权移动平均、线性趋势移动平均、指数平滑、多元线性回归、Holt – Winters 季节叠加与乘积算法
7	Linear Goal Programming and Integer Linear Goal Programming	GP – IGP	目标规划与整数规划	多目标线性规划、线性目标规划
8	Inventory Theory and Systems	ITS	存储论与存储控制系统	经济批量订货、批量折扣、单时期随机模型、多时期动态存储模型、存储控制系统
9	Job Scheduling	JOB	作业调度,编制工作进度	机器加工排序、流水线车间加工排序
10	Linear Programming and Integer Linear Programming	LP – ILP	线性规划与整数规划	线性规划、整数规划、对偶、灵敏度分析
11	Markov Process	MKPA	马尔柯夫过程	转移概率、稳态概率
12	Material Requirements Planning	MRP	物料需求计划	物料需求计划的编制、成本核算

表1(续)

序号	模块	文件名	含义	应用范畴
13	Network Modeling	Net	网络模型	运输、指派、最大流、最短路、最小树、旅行推销商等问题
14	Nonlinear Programming	NLP	非线性规划	有(无)约束条件、目标函数或约束条件非线性、目标函数与约束条件都非线性等规划问题的求解与分析
15	Project Scheduling	PERT - CPM	网络计划	关键路线法、计划评审技术、网络的优化、工程完工时间模拟、绘制甘特图与网络图
16	Quadratic Programming and Intergret Quadratic Programming	QP - IQP	二次规划	求解线性约束、目标函数是二次型的一种非线性规划问题,变量可以取整数
17	Queuing Analysis	QA	排队分析	各种排队模型的求解与性能分析、15 种分布模型求解、灵敏度分析、服务能力分析、成本分析
18	Queuing System Simulation	QSS	排队系统模拟	未知到达和服务之间分布、一般排队系统模拟计算
19	Quality Control Charts	QCC	质量控制图	建立各种质量控制图和质量分析

附录 B　参考答案

第二章

1. 可能出现唯一最优解、无穷多最优解、无界解或无可行解。出现无界解时，模型可能缺少了必要的约束条件。出现无可行解时，模型可能有相互矛盾的约束条件。

2. (1)、(3)正确，其余错误。

3. (1) 唯一最优解，$z^* = 3, x_1 = 0.5, x_2 = 0$；(2) 无可行解；(3) 有可行解，但是模型具无界解；(4) 无穷多最优解，$z^* = 66$；

4. 最优决策为：A 处采购 27.27 万吨，B 处采购 30 万吨，总费用 11 700 万元。

5. (1) 令 $x'_1 = -x_1, x_3 = x'_3 - x''_3$，化为标准型为：

$$\text{Max } z = 2x'_1 + x_2 - 2x'_3 + 2x''_3 - Mx_4 + 0x_5$$

$$s.t. \begin{cases} x'_1 + x_2 + {}'_3 - x''_3 + x_4 = 4 \\ x'_1 + x_2 - x'_3 + x''_3 + x_5 = 6 \\ x'_1, x_2, x'_3, x''_3, x_4, x_5 \geqslant 0 \end{cases}$$

(2) 令 $x'_2 = -x_2, x_4 = x'_4 - x''_4$，化为标准型为：

$$\text{Max } z = 2x_1 - x'_2 + 3x_3 + x'_4 - x''_4 + 0x_5 - Mx_6 + 0x_7 - Mx_8$$

$$s.t. \begin{cases} x_1 - x_2 + x_3 + x'_4 - x''_4 + x_5 = 7 \\ -2x_1 - 3x_2 - 5x_3 + x_6 = 8 \\ x_1 - 2x'_2 + 2x'_4 - 2x''_4 - x_7 + x_8 = 1 \\ x_1, x'_2, x_3, x'_4, x''_4, x_5, x_6, x_7, x_8 \geqslant 0 \end{cases}$$

6. (1) 最终单纯形表为：

$c_j \longrightarrow$			3	5	0	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
0	x_3	2	0	0	1	1/3	-1/3	
5	x_2	6	0	1	0	1/2	0	
3	x_1	2	1	0	0	-1/3	1/3	
	$c_j - Z_j$		0	0	0	-3/2	-1	

(2) 最终单纯形表为：

$c_j \longrightarrow$			2	-1	1	0	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	x_6	
0	x_4	10	0	0	1	1	-1	-2	
2	x_1	15	1	0	1/2	0	1/2	1/2	
-1	x_2	5	0	1	-3/2	0	-1/2	1/2	
$c_j - Z_j$			0	0	-3/2	0	-3/2	-1/2	

(3) 最终单纯形表为：

$c_j \longrightarrow$			6	2	10	8	0	0	0	θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
0	x_6	70	11	0	0	12	2	1	0	
2	x_2	5	-5	1	0	2	1	0	-2	
-1	x_3	20	-6	0	1	7	2	0	-3	
$c_j - Z_j$			76	0	-3/2	-66	-22	0	34	

为无界解。

(4) 令 $x'_1 = x_1 - 1, x'_2 = x_2 - 2, x'_3 = x_3 - 3$，代入原模型化简得：

$$\text{Max } z = x'_1 + 6x'_2 + 4x'_3 + 25$$

$$s.t. \begin{cases} -x'_1 + 2x'_2 + 2x'_3 \leqslant 4 \\ 4x'_1 - 4x'_2 + x'_3 \leqslant 21 \\ x'_1 + 2x_2 + x'_3 \leqslant 9 \\ x_1, x_2, x'_3 \geqslant 0 \end{cases}$$

最终单纯形表为：

$c_j \longrightarrow$			1	6	4	0	0	0	θ_i
C_B	X_B	b	x_1'	x_2'	x_3'	x_4	x_5	x_6	
6	x_2'	1/4	0	1	0	-1/8	-1/8	-3/8	
4	x_3'	4	0	0	1	1/2	1/6	-1/6	
1	x_1'	9/2	1	0	0	-1/4	1/12	5/12	
$c_j - Z_j$			0	0	0	-1	0	-2	

有无穷多组最优解，其中一个为 $x_1 = 11/2, x_2 = 9/4, x_3 = 7$。

本题求解使用了一点技巧，大大简化了计算量。当然不用此技巧，直接解答也可以。

7. 设 $x_1, x_2, \cdots, x_6$ 分别代表早上 2:00，6:00，直至晚上 22:00 开始上班的各时段护士数，则可建立如下线性规划模型。

$$\text{Min } z = x_1 + x_2x_3 + x_4 + x_5 + x_6$$

$$s.t.\begin{cases} x_6 + x_1 \geqslant 10 \\ x_1 + x_2 \geqslant 15 \\ x_2 + x_3 \geqslant 25 \\ x_3 + x_4 \geqslant 20 \\ x_4 + x_5 \geqslant 18 \\ x_5 + x_6 \geqslant 12 \\ x_1, x_2, x_3, x_4, x_5, x_6 \geqslant 0 \end{cases}$$

解得 $x_1 = 0, x_2 = 15, x_3 = 10, x_4 = 16, x_5 = 2, x_6 = 10$，共需要 35 名护士。

8. 设为 x_{ij} 第 i 年初投放到 j 项目的资金数，建立线性规划模型：

$$\text{Max } z = 1.2x_{31} + 1.6x_{23} + 1.4x_{34}$$

$$s.t.\begin{cases} x_{11} + x_{12} = 300\,000 \\ x_{21} + x_{23} = 1.2x_{11} \\ x_{31} + x_{34} = 1.2x_{21} + 1.5x_{12} \\ x_{12} \leqslant 150\,000 \\ x_{23} \leqslant 200\,000 \\ x_{34} \leqslant 100\,000 \\ x_{ij} \geqslant 0, i = 1,2,3, j = 1,2,3,4 \end{cases}$$

解得 $x_{11} = 166\,666.7, x_{12} = 133\,333.3, x_{21} = 0, x_{23} = 200\,000, x_{31} = 100\,000, x_{34} = 100\,000$。三年的本利和为 580 000 元。

第三章

1. (1)、(2)、(3) 正确，(4) 错误。

2. 空白处的数字为(原有的数据被删除)：

c_j									θ_i
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	x_6	
5	x_2								
0	x_5								
0	x_6								
$c_j - Z_j$									
.....									
5	x_2	80/41	0	1	0				
4	x_3	50/41	0	0	1				
3	x_1	44/41	1	0	0				
$c_j - Z_j$			0	0	0	-45/41	-24/41	-11/41	

3. (1) $\text{Max } w = 3y_1 - 5y_2 + 2y_3$

$$s.t.\begin{cases} y_1 + 2y_3 \leqslant 3 \\ -2y_1 + y_2 - 3y_3 = 2 \\ 3y_1 + 3y_2 - 7y_3 = -3 \\ 4y_1 + 4y_2 - 4y_3 \geqslant 4 \\ y_1 \leqslant 0, y_2 \geqslant 0 \end{cases}$$

(2) $\text{Max } w = 15y_1 + 20y_2 - 5y_3$

$$s.t.\begin{cases} -y_1 - 5y_2 + y_3 \geqslant -5 \\ 5y_1 - 6y_2 - y_3 \leqslant -6 \\ -3y_1 + 10y_2 - y_2 = -7 \\ y_1 \geqslant 0, y_2 \leqslant 0 \end{cases}$$

4. 原线性规划问题为:

$\text{Max } z = 6x_1 - 2x_2 + 10x_3$

$$\begin{cases} x_1 + 2x_3 \leqslant 5 \\ 3x_1 - x_2 + x_3 \leqslant 10 \\ x_1, x_2, x_3 \geqslant 0 \end{cases}$$

对偶问题的最优解为 $Y^* = (4,2)$。

5. (1) 原问题的最优解为 $X^* = (3/5,0,3,00)^T$,对偶问题最优解为 $Y^* = (1/5, 3/5)$。

(2) x_2 系数变化后,对偶问题的第二个约束相应变成 $2y_1 + 3y_2 \geqslant 3$,将原来的最优解 $Y^* = (1/5,3/5)$ 代入,约束不满足,故原问题的最优解发生了变化。

(3) 相对于新变量 x_6,因为有 $2 - (3,2)\begin{bmatrix} 1/5 \\ 3/5 \end{bmatrix} = 2 - 9/5 > 0$,故原问题最优解发生了变化。

6. (1) $15/4 \leqslant c_1 \leqslant 50, 4/5 \leqslant c_1 \leqslant 40/3$

(2) $24/4 \leqslant b_1 \leqslant 16, 9/2 \leqslant c_1 \leqslant 15$

7. (1) 以 x_1, x_2, x_3 分别代表甲、乙、丙产品的产量,最优解为 $X^* = (5,0,3)^T$,最大盈利为 35。

(2) 购进原料 B 有利,购进 B 原料 15 单位为宜。

(3) 新计划为 $X^* = (0,0,6)^T$,最大盈利为 30。

第四章

1. 最优解为 $x_1 = 4, d_1^* = 1, x_2 = x_1 - d_1 = 3, d_2^* = 0, x_3 = x_2 - d_2^* = 3, d_3 = 3, x_4 = x_3 - d_3 = 0$。

即项目 A 投资 1 万元,项目 B 投资 0 万元,项目 C 投资 3 万元,最大效益为 60 万吨。

2. $d_1^* = 5, x_2 = x_1 - 2d_1 = 0, d_2^* = 0, x_3 = x_2 - 4d_2 = 0, d_3^* = 0, x_4 = x_3 - 2d_3 = 0$。

即应取第一种物品 5 件,第二种物品 0 件,第三种物品 0 件,最高价值为 60 元,背

包没有余量。

3. 求解结果为:$x_1 = 2$, $x_2 = 1$,$f(x) = 0$。

4. $x^* = 3$, $f(x^*) = 0$

5. $x^* = (0,1)^T$

6. $x^* = (2,1)^T$

第五章

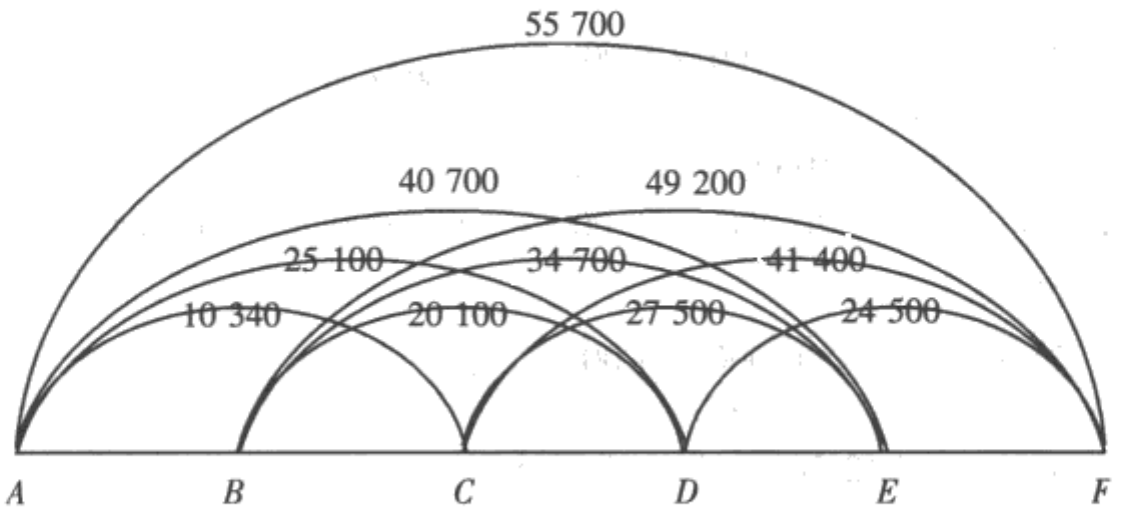
1. (2)、(5)、(6) 正确,(1)、(3)、(4) 错误。

2. (1) 最小支撑树树枝总长 12;(2) 最小支撑树树枝总长 15;(3) 最小支撑树树枝总长 12;(4) 最小支撑树树枝总长 18。

3. D。

4. 输油管总长为 10.2 海里。

5. 将线段分成 5 部分,AB,BC,CD,DE,EF,分别表示生产 5 种包装箱所需费用,AC 表示用 2 号箱代替 1 号箱时,生产 2 100 个 2 号箱费用,AD 表示用 3 号箱代替 1,2 号箱时,生产 3 500 个 3 号箱费用,以此类推,得下图:



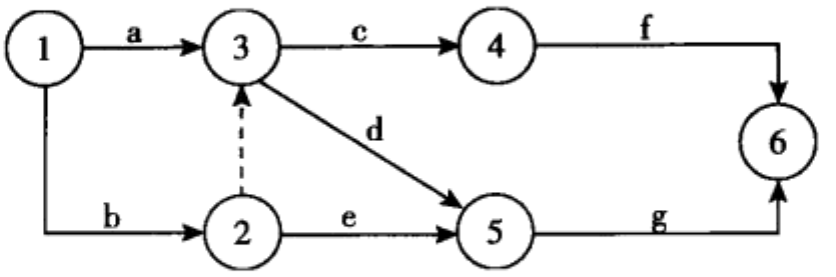
6. v_s 到 v_t 的最大流分别为:(1)5,(2)13,(3)20,(4)25。

7. 流水线生产能力达到最大的分配方案为:

工序 / 工人	Ⅰ	Ⅱ	Ⅲ	Ⅳ	Ⅴ
甲	×	×	×	×	7
乙	×	×	×	5	6
丙	×	5	×	×	×
丁	5	×	×	×	5
戊	×	7	6	×	×

8. 分别找出各自图中的奇点,然后将图中奇点按最短连线两两相连即可。

9.



10. (1)80 天;(2) 无影响;(3) 缩短 4 天;(4) 工程开工后的第 56 天;(5) 需要采取措施,应设法缩短关键路线上的工序 a,c,e,f,g,j,k,n 的工序时间,共需缩短 5 天的时间。

第六章

1. $Q^* = \sqrt{\frac{2HS}{PI}} = \sqrt{\frac{2 \times 6\,400 \times 100}{10 \times 0.2}} = 800$(个),$N = H/Q^* = 6\,400/800 = 8$(次)

订货间隔时间:360/8 = 45(天),$A = B + F = 800 + 800 = 1\,600$(美元)

2. 1) $P_1 = 40$ 时

$$Q^* = \sqrt{\frac{2HS}{PI}} = \sqrt{\frac{2 \times 40\,000 \times 5\,000}{40 \times 0.1}} = 10\,000\text{(台)},$$

$N = H/Q^* = 40\,000/10\,00 = 4$(次),

$A_1 = B_1 + F_1 + D_1 = 20\,000 + 20\,000 + 1\,600\,000 = 1\,640\,000$(美元)

2) $P_2 = 36$ 时

$$Q^* = \sqrt{\frac{2HS}{PI}} = \sqrt{\frac{2 \times 40\,000 \times 5\,000}{36 \times 0.1}} \approx 10\,541 \text{ 台}$$

(计算费用时,应以 12 500 台代入进行计算)

则 $N = H/Q^* = 40\,000/12\,500 = 3.2$(次)

$A_2 = B_2 + F_2 + D_2 = 16\,000 + 22\,500 + 1\,440\,000 = 1\,478\,500$(美元)

3) $P_3 = 32$ 时

$$Q^* = \sqrt{\frac{2HS}{PI}} = \sqrt{\frac{2 \times 40\,000 \times 5\,000}{32 \times 0.1}} \approx 11\,180\text{(台)}$$

(计算费用时,应以 20 000 台代入进行计算)

则 $N = H/Q^* = 40\,000/20\,000 = 2$(次)

$A_3 = B_3 + F_3 + D_3 = 10\,000 + 32\,000 + 1\,280\,000 = 1\,322\,000$(美元)

即每次进货 20 000 台,每年进货两次的总费用最低。

3.

$$Q^* = \sqrt{\frac{2HS}{PI(1 - U/R)}} = \sqrt{\frac{2 \times 20\,000 \times 500}{10 \times 0.1(1 - 60/300)}} = 5\,000\text{(台)}$$

4. $Q^* = \sqrt{\frac{2HS}{PI(1 - U/R)}} = \sqrt{\frac{2 \times 40\,000 \times 640}{10 \times 0.1(1 - 80/400)}} = 4\,000$(件),$N = 5$(次),

$B = 3\,200$ 元,$F = 3\,200$ 元,$D = 200\,000$ 元,$A = 206\,400$ 元。

当有 5% 的价格折扣时:

$B_1 = 3\ 200$(元),$F_1 = 4\ 000$(元),$D_1 = 190\ 000$(元),$A_1 = 197\ 200$(元)。

应接受供货方价格折扣的建议。

5. $Q^* = \sqrt{\dfrac{2HS}{K(1-U/R)}} = \sqrt{\dfrac{2\times 6\ 000\times 150}{2\times(1-30/50)}} = 1\ 500$(个),$N = 4$(次),

$A = B + F = 600 + 600 = 1\ 200$(元),最大库存为 $Q^*(1-U/R) = 600$(件)

6. (1)$Q^* = \sqrt{\dfrac{2HS}{K}} = \sqrt{\dfrac{2\times 50\times 100}{0.2\times 7}} \approx 84.52 \approx 85$(箱),$F = 59.5$(元),$B = 58.82$(元),$A = 118.32$(元)。

(2)$Q^* = \sqrt{\dfrac{2HS(K+C)}{KC}} = \sqrt{\dfrac{2\times 50\times 100(1.4+5)}{0.2\times 7\times 5}} \approx 95.62$(箱)$\approx 96$(箱),$F = 67.2$(元),$B = 52.08$(元),$A = 104.58$(元),最大缺货量 $V = KQ/(K+C) = 1.4\times 96/(1.4+5) = 21$(件)。

(实际工作中,订货次数和订货量一般应是整数,因此,当订货次数和订货量是小数时,应调整为整数再进行费用的计算,一种简便的方法是用四舍五入法。)。

7. $Q^* = \sqrt{\dfrac{2HS}{K(1-U/R)}} = \sqrt{\dfrac{2\times 30\ 000\times 1\ 000}{28\times(1-300\ 000/50\ 000)}} \approx 2\ 314.55$(件)$\approx 2\ 315$(件),

$N = H/Q = 30\ 000/2\ 315 = 12.96$(次)$\approx 13$(次),$Q = H/N = 30\ 000/13 = 2\ 307.69$(件)$\approx 2\ 308$(件)

(计算费用时将 $N = 13$,$Q = 2\ 308$ 代入计算)

$B = 13\ 000$(元),$F = 12\ 925$(元),$A = 25\ 925$(元),最高存储量为932件,订货周期间隔约为28天。

8. 合理的安全库存量为5件。

9. 90% 的订货点为251件,95% 的订货点为266件,增加的安全库存量为15(66 - 51)件,增加的存储费用为150元。

10. 只允许5% 缺货时,订货点为502本,安全库存量102本,存储费用1 020元;只允许10% 缺货时,订货点为532本,安全库存量132本,存储费用1 320元。

第七章

1. (1)、(2)、(4)、(8)、(9) 正确;(3)、(5)、(6)、(7)、(10) 错误。

2. 略。

3. (1)7天;(2)79天。

4. (1)$0.614\ 4\times 10^{-5}$;(2)$0.961\ 2\times 10^{-2}$;(3)0.036 1;(4)0.052 65。

5. (1)0.135;(2)0.270;(3)0.052 7

6. (1)0.632,0.233,0.135;(2)0.632;(3)0.368,0.368,0.264;(4)0.135,0.717,0.967

7. (1)$L = \sum_{n=0}^{4} nP_n = 2$(人);(2)$L_q = \sum_{n=3}^{4}(n-2)P_n = 0.375$(人);(3)$P_1 + 2(P_2 + P_3 + P_4) = 1.625$(人);(4)$W = \dfrac{L}{\lambda}$,$\lambda$ 应为实际进入系统人数 $\lambda = 2(1-P_4) = \dfrac{30}{16}$,

故 $W = 2 \times \frac{16}{30} = \frac{16}{15}$(小时);(5) 由$\frac{1}{\mu} - \frac{1}{\lambda}(L - L_q)$,得$\frac{1}{\mu} = \frac{26}{30}$(小时)。

8. (1)$P_0 = 1/3, L_q = 1/3, L = 4/3$;(2) 平均忙期为 1/2。

9. 因有 $P_0 = \frac{1-\rho}{1-\rho^{N+1}}, N = 4$,第一个理发店$\rho = \frac{10}{4} = 2.5$,第二个理发店$\rho = \frac{10}{6} = \frac{5}{3}$,计算得到第一个理发店 $P_0 = 0.026$,理发人员收入为 $12 \times 4 \times 11(1 - 0.026) = 514.272$(元),第二个理发店 $P_0 = 0.056$,理发人员收入为 $12 \times 6 \times 7.5(1 - 0.056) = 509.76$(元)。

10. (1)$P_0 = 0.278\,6, P_1 = 0.232\,2, P_2 = 0.193\,5, P_3 = 0.161\,2, P_4 = 0.134\,4$;(2)$L = 4.328$;(3)0.721 4;(4)0.134 4。

11. 全部车辆运转概率为 0.966 8,有一台不能运转概率为 0.032 2,两台不能运转概率为 0.000 859,三台不能运转概率为 0.000 017 2。

12. 3 名工人。

第八章

1 ~ 14 题(略)

15. 期望利润值、期望后悔值、边际标准的最优进货量均是 7 浦式耳;$EVPI = 4$(美元);有处理价格的最优进货量是 6 或 7 浦式耳。

16. 期望后悔值标准应选择钻油井(钻的 $EV = 20$,不钻的 $EV = 32$);$EVPI = 20$。

17. 庄稼存活率为 0.725。

18. 通过考试的概率是 71.25%。

19. 该工程十一月一日前能完工的概率是 0.674。

20. 更新方案有随坏随换、每一月、每二月、每三月 4 种,随坏随换的月平均费用最少,应采用随坏随换(月平均费用:随坏随换 = 2 640 美元、每一月 = 3 000 美元、每二月 = 2 650 美元、每三月 = 2 773 美元)。

21. 两个产品均为合格品时,生产线为有效运转的概率是 0.953;两个产品均为不合格品时,生产线为有效运转的概率是 0.1。

22. 公司每天的最优存货量应为 3 856 枝。

23. 决策树略。该公司应准备进行投标,投标的期望利润为 6.4 万美元。

24. 乐观标准选乙、悲观标准选甲、等概率标准选甲、后悔值标准选丙、现实主义标准取决于乐观系数。

25. 转移矩阵不变时,A、B、C 三个公司平衡状态的市场占有率分别为 0.225、0.489、0.286;

若 A 公司能将丧失给 B、C 两公司的顾客保留一半的话,平衡状态 A、B、C 三个公司的市场占有率分别为 0.367、0.4、0.233,即 A 公司的市场占有率可增加 0.142。

26. 星期二两食堂的销售份额分别是:$A = 0.42$、$B = 0.58$;平衡状态时两食堂的销售份额分别是:$A = 0.375$、$B = 0.625$。

参考文献

1. 钱颂迪,等. 运筹学. 北京:清华大学出版社,1990.
2. 汤代焱,等. 管理运筹学. 长沙:湖南大学出版社,1997.
3. 胡运权. 运筹学基础及应用. 3 版. 哈尔滨:哈尔滨工业大学出版社,1998.
4. 胡运权,等. 运筹学教程. 2 版. 北京:清华大学出版社,2003.
5. 胡运权,等. 运筹学习题集. 3 版. 北京:清华大学出版社,2002.
6. 韩伯棠,等. 管理运筹学. 2 版. 北京:高等教育出版社,2005.
7. 熊伟. 运筹学. 北京:机械工业出版社,2005.
8. 许玖平,等. 运筹学(Ⅰ类). 2 版. 北京:科学出版社,2004.
9. 邱菀华,等. 运筹学教程. 北京:机械工业出版社, 2004.
10. 杨超. 运筹学. 北京:科学出版社,2004.
11. 吴祈宗. 运筹学与最优化方法. 北京:机械工业出版社,2003.
12. 李军,徐玖平. 运筹学——非线性系统优化. 北京:科学出版社,2003.
13. 程理民,等. 运筹学模型与方法教程. 北京:清华大学出版社,2000.
14. 何坚勇. 运筹学基础. 北京:清华大学出版社,2000.
15. F. S. 麦克·劳林,R. C. 皮克·哈德. 管理决策的定量技术. 北方交通大学运输系,译. 北京:中国铁道出版社,1982.
16. 姜圣阶,等. 决策学基础(下册). 北京:中国社科出版社,1986.
17. 王艳. 创新思维. 延吉:延边大学出版社,2001.
18. 王艳. 魔鬼创新. 延吉:延边大学出版社,2001.
19. 韩大卫. 管理运筹学. 大连:大连理工大学出版社,1998.
20. 龙子泉,陆菊春. 管理运筹学. 武汉:武汉大学出版社,2002.

[General Information]
书名=管理决策模型、方法与应用　　第3版
作者=张照贵主编
页数=293
出版社=成都市：西南财经大学出版社
出版日期=2012.06
SS号=13075990
DX号=000008341140
URL=http://book.szdnet.org.cn/bookDetail.jsp?dxNumber=000008341140&d=81FAA7180D0E4CE38DB30AE87D2CD153