第 12 章　整数规划

第 3 章描述了几个例子,你会发现线性规划能够用来解决各种问题。然而,线性规划的可分割性假定,限制了它的应用范围。即决策变量要求取非整数值(见 3.3 节)。但在很多实际问题中,决策变量只有取整数值才有意义,如必须给一个任务分派整数的人、机器或车辆。如果一个问题与线性规划的不同之处仅在于要求变量取值为整数,那它就是一个整数规划(**IP**)问题(更复杂的名称是整数线性规划。形容词线性通常被省略掉,除非是与更高深的整数非线性规划问题对比,这超出了本书涉及的范围)。

整数规划的数学模型就是线性规划的数学模型(见 3.2 节)再加上一个变量必须取整数值的额外限制。只要求部分变量取整数值的(可分割性假定对其余变量仍然适用)称为混合整数规划(**MIP**)。为了加以区分,我们要求全部变量取整数值的问题称为纯整数规划问题。

例如,在 3.1 节的 Wyndor Glass 公司问题中,如果两个决策变量 x_1、x_2 分别代表产品 1 和产品 2 的总产量,而不是生产率,那么,该模型就是一个整数规划问题。因为两种产品(玻璃门和木框窗)都是不可分割的,所以必须考虑 x_1、x_2 的整数约束。

当变量不满足可分割性假定时,许多此类问题可以作为线性规划直接扩展的整数规划来解决。然而,另一领域的应用显得尤为重要,即涉及是或否的决策问题。在这种决策中,只有两个可能的选择是或否。例如,我们是否应该实施一项准备就绪的工程?我们是否应该进行某项投资?我们是否应该在一个特定地点安置设备?

由于只有两种选择,我们只给决策变量取两个值 0 和 1,就能够表达这种决策。x_j 代表第 j 个是或否的决策,有

$$x_j=\begin{cases}1,\text{决策 } j \text{ 为是}\\0,\text{决策 } j \text{ 为否}\end{cases}$$

这种变量称为二值变量(或 0-1 变量)。所以只包含二值变量的整数规划问题有时称为二值整数规划问题(或 0-1 整数规划问题)。

12.1 节提出一个简单的典型二值整数规划问题。12.2 节考察了各种二值整数规划的应用。其他二值变量的一些模型将在 12.3 节中讨论。12.4 节提出了一系列建模的例子。12.5 节~12.8 节讨论了解决整数规划问题的方法,包括二值整数规划和混合整数规划问题。12.9 节介绍了一个令人激动的最新发展(约束规划),这将显著提高建模和求解整数规划模型的能力。

12.1　范　　例

加州制造公司正在考虑建一个新工厂来扩大公司规模。新工厂可能建在洛杉矶或旧金山,甚至可能在两座城市都建新工厂。该公司同时正在考虑建至多一个新仓库,但是新仓库的选址需视新工厂的厂址而定。表 12.1 的第 4 列是每种选择的净现值(考虑到资金时间价值的总收益)。最右边一列给出了目前投资所需要的资金(已经包含在净现值中)。可用的总资金为 1000 万美元。目标是找到一种可行的选择组合,以使总净现值最大。

表 12.1 加州制造公司的数据

决策数	答案为是或否的问题	决策变量	总净现值	需要的资金
1	是否在洛杉矶建新工厂	x_1	900 万美元	600 万美元
2	是否在旧金山建新工厂	x_2	500 万美元	300 万美元
3	是否在洛杉矶建新仓库	x_3	600 万美元	500 万美元
4	是否在旧金山建新仓库	x_4	400 万美元	200 万美元
可用资本:1000 万美元				

12.1.1 二值整数规划模型

这个问题非常简单,稍做调查就能很快得出结论(在两座城市都建新工厂,不建新仓库)。让我们用一个整数规划模型说明。所有决策变量都是二值形式,即

$$x_j=\begin{cases}1,\text{决策 } j \text{ 为是}\\0,\text{决策 } j \text{ 为否}\end{cases}\qquad (j=1,2,3,4)$$

令

$$Z=\text{这些决策的总净现值}$$

如果投资建一个工厂或仓库(相应的决策变量取值为1),则该投资的预期净现值由表12.1的第4列给出。如果没有投资(相应的决策变量取值为0),净现值为0。因此,以百万美元为单位:

$$Z=9x_1+5x_2+6x_3+4x_4$$

表12.1最右边一列说明用于工厂和仓库的投资金额不能超过1000万美元,故仍然以百万美元为单位,该模型的一个约束条件为

$$6x_1+3x_2+5x_3+2x_4\leqslant 10$$

因为后两个决策是互斥的(公司至多新建一个仓库),我们需要以下约束条件:

$$x_3+x_4\leqslant 1$$

另外,决策3和决策4是可能决策,因为它们分别要视决策1和决策2而定(只有当公司打算在某城市新建一个工厂时,公司才会考虑是否在那里建一个仓库),因此,对于决策3来说,我们要求若 $x_1=0$,则 $x_3=0$。这个对决策3的限制(当 $x_1=0$ 时)通过以下约束条件实现:

$$x_3\leqslant x_1$$

类似地,若 $x_2=0$,则 $x_4=0$。可以由以下约束条件实现:

$$x_4\leqslant x_2$$

因此,我们重写这两个约束条件,把所有变量都移到左边,完整的0-1整数规划模型为

$$\text{Max}\quad Z=9x_1+5x_2+6x_3+4x_4$$

s. t.

$$\begin{aligned}
6x_1+3x_2+5x_3+2x_4&\leqslant 10\\
x_3+x_4&\leqslant 1\\
-x_1+x_3&\leqslant 0\\
-x_2+x_4&\leqslant 0\\
x_j&\leqslant 1\\
x_j&\geqslant 0
\end{aligned}$$

且 x_j 是整数,$j=1,2,3,4$。

该模型的最后 3 行可以等价转换为一个约束条件：

$$x_j \text{ 是二值变量}, j=1,2,3,4$$

这个例子虽然简单,但是当主要变量为 0-1 变量时,它能典型地代表许多整数规划的实际应用。正如本例的第二对决策,一组是或否的决策经常构成一组互斥的选择。也就是在一组中只能有一个决策的结果为是。每一组需要一个约束条件,相应地,0-1 变量的和必须等于 1(如果一组中刚好有一个决策的结果为是)或小于等于 1(如果一组中至多有一个决策的结果为是)。有时,是或否的决策是可能决策,也就是依赖先前决策结果的后续决策。例如,如果只有当一个决策结果为是时,另一个决策才可能为是,那么,我们就说后一个决策要视前一个决策而定。当可能决策涉及一个后续的行动时这种情况就发生了;如果前一个决策结果为否,那么,可能决策就变得无足轻重甚至不可能了。最终得到的约束条件通常采用本例中第三个与第四个约束条件的形式。

12.1.2　用于求解此类模型的软件

所有用于运筹学软件包(Excel、LINGO/LINDO 和 MPL/Solvers)都包括求解 0-1(纯或混合)整数规划模型的算法,也包括求解一般(纯或混合)整数规划模型的算法,此时,变量必须是整数而不是 0-1 变量。二值变量处理起来较一般整数变量更为容易,所以前者的算法能够解决问题的范围比后者大。

在使用 Solver 时,处理过程与线性规划基本相同。当为了增加约束条件,单击"Solver"对话框的"Add"按钮时,会出现一点不同。除了满足线性规划的约束条件之外,还需要添加整数约束。在整数变量而不是 0-1 变量的情况下,这要在"Add Constraint"对话框中完成,在左边选择整数变量的范围并且在弹出菜单中选择"int"。对于二值变量来说,在弹出菜单中选择"bin"。

本章的一个 Excel 文件给出了加州制造公司案例的完整的建模和结果的程序列表。本书网站的解题案例包括了两个整数约束的最小值案例。这个案例解释了整数规划的模型和图形解法,也包括电子表格式解法。

一个 LINGO 模型使用@ BIN( )函数说明圆括号例的变量为 0-1 变量。对于一般整数变量(取整数值而不是 0-1 变量),@ GIN( )函数的使用方法与之相同。在两种方法中,函数都可以嵌入一个@ FOR 声明,给整个变量组规定二值或整数约束。

在一个 LINDO 模型中,二值或整数约束是插在 END 声明后面的。输入 GIN X 是指定变量 X 为一般整数变量。另外。对于任何正整数 n 声明 GIN n 指定前 n 个变量是一般整数变量。除了用 INTEGER 替换 GIN 之外,对 0-1 变量的处理方法是相同的。

对于个 MPL 模型,关键词 INTEGER 用来指明一般整数变量,而 BINARY 用于 0-1 变量。在 MPL 模型的变量部分,你要做的只是在 VARIABLES 标签前加适当的形容词(整数或二值的)指明标签下列表中的变量是哪种类型的,或者你可以不在变量部分给出指定,而是在模型部分在其他的约束后面加入二值或整数约束。使用这种方法时,变量组上的标签会变成 INTEGER 或 BINARY。

学生版本的 MPL 包括经典的解决线性规划问题的方法——CPLEX、GUROBI、CoinMP 和 SULUM。这 4 种工具可以用于求解解纯或混合整数规划或二值整数规划模型。例如,运用 CPLEX,通过选择 Options 菜单下 CPLEX Parameters 对话框中的 MIP Strategy 标签。一个有经验的人能够从完成算法的众多选项中找到最适合的选项。

当你看到这些不同的软件包应用于例子时,上面的介绍就会变得更加清晰。运筹学课程软件中本章的 Excel、LINGO/LINDO 和 MPL/Solvers 文件,说明了怎样将各个软件应用于本节介绍的范例,以及后面的整数规划的例子。

本章的后半部分将主要讨论整数规划算法,这些算法与软件包中所使用的算法类似。12.6 节将用本范例演示纯二值整数规划算法的应用。

12.2 整数规划的应用

正如加州制造公司的例子,管理者经常要面对是或否的决策,因此二值整数规划广泛用于诸如此类的辅助决策。

我们现在要介绍不同类型的是或否的决策。本章节包括了两个应用用于解释这两种类型。对于其他类型,我们也提到了一些其他的关于二值整数规划解决实际应用的例子。所有这些例子都包含在了本章参看文献中,这些文献可通过本书网站进行浏览。

12.2.1 投资分析

人们有时利用线性规划完成资金预算决策,决定对不同的项目分别投资多少。然而,正如加州制造公司的例子,一些资金预算决策不是决定投资多少,而是决策是否进行一项固定金额的投资。尤其在该例中 4 个决策为是否投入固定的资金在一座城市(洛杉矶或旧金山)新建工厂或仓库。

管理层必须经常面对的决策是有关是否进行一项固定投资(需要投入的资金是已经预先确定的)。我们是否应该收购正从另一家公司中分离出来的子公司?我们是否应该买下某个原料来源?我们是否应该引进一条新生产线自己生产某种原料,而不是继续从供应商那里获得它?

总之,关于固定投资的资金预算决策是一种是或否的决策,其表达形式如下。

每个是或否的决策:

我们是否应该进行一项固定投资?

$$\text{它的决策变量}=\begin{cases}1, & \text{是}\\0, & \text{否}\end{cases}$$

案例来源于参考文献[A6]。主要的运筹学问题,是关于南非国防部用更少的支出改善国防能力。这个案例中的“投资”是获得物的成本和正在发生的支出,为了获得某种特定的军事能力,这些投入都是必需的。它们建立了一个混合二值整数规划模型,选择那些能够使国防部整体效力最大化的特定能力,同时满足了预算限制。这个模型的变量超过 16000 个(包括 256 个 0-1 变量),函数约束超过 5000 个。国防部优化后的编制每年能节约 11 亿多美元和重要的非货币收益。

参考文献[A2]给出了另外一个成功的关于投资分析的混合二值整数规划。该模型被 Grantham、Mayo、VanOtterloo 公司用来定量地组合它所控制的资金,资金额超过 80 亿美元。每次公司都组合其资产使之与目标组合相近(以部门和安全为依据)。但是该组合的股票数目要少得多而且更加容易控制。一个 0-1 变量用来表示某支股票是否应该包括在资产组合中,而一个连续变量则代表组合中包含的股票数目。给出一个需要被平衡的现有投资组合,是希望减少为了得到最终投资组合所需要的交易次数,降低交易成本。所以另一个 0-1 变量被引入模型,代表是否完成改变持股票数目所进行的交易。使用该模型以后,每年能使持有资本所带来的成本减少 400 万美元以上。

应用案例

美国中西部独立传输系统公司(简称 MISO)是一家成立于 1998 年的非盈利组织,组织负责美国中西部电力的生产与传输。

该公司服务于 4000 万用户(包括个人和公司),控制了近 6 万 mile 的高压传输电线和能够生产 14600000 万 W 电力的电力工厂 1000 多个。基础设施覆盖了美国中西部 13 个州和加拿大的马尼托巴湖省。

该公司主要任务是向所有用户提高可靠且高效率的电力。该公司将运用混合二元线性规划求解在满足用电需求的条件下,花费的最小化。模型中的每一个二元变量需要提供电厂在特定时间段提供电力的决策("是"或"否")模型结论将应用到电力输出级别和电价的线性规划模型。

混合二元线性规划模型包括 330 万个连续变量,45 万个 0-1 变量和 390 万个约束条件。运用特殊的算法(拉格朗日松弛算法)求解该复杂巨系统模型。

该改进运筹学算法,4 年(2007 至 2010 年)为该公司共节约了 25 亿美元,预计到 2020 年能为公司节约 70 亿美元。该公司因为该算法获得了 2011 年 Franz Edelman 运筹学和管理科学国际竞赛第一名。

来源:B. Carlson and 12 co-authors,"MISO Unlocks Billions in Savings Through the Application of Operations Research for Energy and Ancillary Services Markets,"*Interfaces*,42(1):58-73,Jan. -Feb. 2012.(我们的网址提供了本文链接:www. mhhe. com/hillier。)

12.2.2 选址

在如今的全球经济中,许多公司正在全世界各个地方建立新工厂,为的是获得低劳动力成本等好处。在为新工厂选址之前,需要分析和比较很多地点(在加州制造公司的例子中,有两个可供选择的厂址)。每个可供选择的地点都涉及一个是或否的决策,其表达形式如下。

每个是或否的决策:

是否应该选择某个地点来建新设施?

$$它的决策变量=\begin{cases}1, & 是\\0, & 否\end{cases}$$

在许多案例中,目标是地点的选择以使新建设施的总成本最小化,这些新设施能满足生产的需要。

正如参考文献[A10]描述的,AT&T 利用一个 BIP 模型来帮助几十个客户选择电话促销中心的地址。该模型使劳动力、交通和实际土地成本最小化,同时提供了每个中心理想的覆盖面积。仅仅在一年中,这个方法就使 AT&T 的 46 位客户迅速而放心地做出了对选址是与否的决策,一年就产生了 3.75 亿美元的网络服务费用和 3100 万美元 AT&T 设备销售额。

正如参考文献[A5]描述的,全球造纸公司 Norske Skog 也应用了类似模型。但是这次模型主要用于解决关闭工厂的问题,而不是建立新工厂。由于电子媒体逐步取代纸质媒体,所有该公司的造纸产量被迫下降。为此,该公司建立了一个巨大的混合二元线性规划模型(312 个 0-1 变量,47000 个连续变量和 2600 个约束条件),得到了关闭两个造纸工厂和一个造纸机器的决策,从而每年为公司节约 1 亿美元。

下面将要讨论这类问题对于许多公司都是很重要的,其中选址起到了关键性的作用。

12.2.3 设计生产和销售网络

如今,制造商们为了使产品更快地进入市场同时降低生产和销售成本,面临巨大的压力。因此,任何在大范围内(甚至全世界)销售产品的公司必须不断关注生产和销售网。

这种设计包括下述类型的是或否的决策:

是否应该让某工厂继续运营?

是否应该选择某地点开一家新工厂?

是否应该保持某销售中心继续营业?

是否应该选择某地点建一个新销售中心?

如果每个市场都仅由一个销售中心来提供服务,那么,对于每个市场与销售中心的组合,我们产生了另一种是或否的决策。

是否应该指定某一销售中心为某一市场服务?

对于每个此类是或否的决策:

$$它的决策变量=\begin{cases}1, & 是\\0, & 否\end{cases}$$

正如本章关于美国中西部独立传输系统公司的应用案例,运用巨型混合二元线性规划模型为公司节约了上亿美元。

12.2.4 发送运输

生产和销售网络投入运行后,关于如何送货的日常运营决策是必不可少的。某些此类决策同样属于是或否的决策。

举个例子来说,假设用卡车运送货物,每辆卡车在每一次行程中都将货物送给几个顾客,这时候,就有必要为每一辆卡车选择一个路线(顾客的次序),故每个候选路线产生一个是或否的决策。

是否应该为一辆卡车选择某一路线?

$$它的决策变量=\begin{cases}1, & 是\\0, & 否\end{cases}$$

目标是选择路线使送达所有货物的总成本最小。

我们还要考虑到具体操作中的复杂性。例如,如果存在不同的卡车型号,每个备选方案既包括路线也包括相应卡车的型号。类似地,如果时间是一个因素,出发的时间段也可以包含在是或否的决策中。考虑到这两个因素,每个是或否的决策有如下形式。

在一次送货过程中,是否应该同时考虑以下因素:

(1) 一条路线;

(2) 卡车的载重;

(3) 出发的时间段。

$$它的决策变量=\begin{cases}1, & 是\\0, & 否\end{cases}$$

正如参考文献[A7]描述的该类型的混合二元线性规划模型可以以巴西最大的全球石油企业 Petrobras 为例,该公司由近1900雇员运用40架直升机在4个陆地基地和80个海岸石油工厂运输石油。模型通过优化直升机路径和时间,每年为公司节约2000万美元。

12.2.5 安排相互联系的活动

每天我们都要安排互相关联的活动,甚至安排什么时候开始不同的家庭作业。所以管理者必须对相关活动进行安排。我们应该什么时候开始生产不同的新产品?我们应该什么时候开始为不同的新产品作市场推广?我们应该什么时候进行不同的投资扩大生产能力?

对于每一项活动,对于每一个可能的开始时间对应一个决策,关于什么时候开始决策都可以通过是或否的决策表达。

某活动是否应该在某一时间开始?

$$它的决策变量=\begin{cases}1, & 是\\0, & 否\end{cases}$$

由于特定活动只能在一个时间开始,对各种时间的选择是一组互斥选择,所以如果只有一个时间可供选择,那么,决策变量的值为1。

正如参考文献[A4]描述的，瑞典政府运用大型混合二元线性规划模型安排 4000 个员工的工作时间和路径为老年人提供家庭护理。该模型不仅节约了 3000 万~4500 万美元，而且提高了家庭护理的质量。

12.2.6 航空应用

航空业在其日常运营中广泛地使用运筹学。很多运筹学专家在该领域工作。主要的航空公司都有专门从事运筹学应用的内部机构。另外，一些著名的咨询公司仅仅关注公司涉及运输的问题，如航空公司。我们将要提及两个应用，它们都使用混合二元线性规划模型。

一个是飞机安排问题。给定几种不同类型的可用飞机，问题是为每个日程安排中的航班指定某种机型，以使完成日程安排获得利润最大化。基本问题是如果给航班分配的飞机太小，那么，将失去潜在的乘客。如果飞机太大，则因为存在空座位，公司将承担更大的开销。

对于每个机型与航班的组合，我们有如下是或否的决策。

$$\text{它的决策变量}=\begin{cases}1, & \text{是}\\0, & \text{否}\end{cases}$$

应用案例

荷兰铁路公司是荷兰主要的铁路建设和运输管理公司。该公司用于 5500 个乘客运输线路，每天运输 110 万乘客，每年的收益约 15 亿欧元(约 20 亿美元)。

过去几年荷兰铁路公司的乘客数量稳定增长，由于 2002 年国家开始扩建 3 个基础设施，所以所有的荷兰列车时刻表需要更新，每列火车的始发和到站时间都需要改变。因此，为了确保新的时刻表满足需求，以及最大化利用资源，针对荷兰铁路规划的运筹学分析就此开始。该项任务由公司的后勤部门、欧洲顶尖的运筹学专业大学以及软件公司共同完成。

新的时刻表于 2006 年 12 月开始正式运行，包括相关配套运输资源的调配以及员工工作时间的分配。混合二元线性规划模型解决了相关问题。同样地，该模型也有效解决了西南航空公司人员调度的相关问题。

运用该运筹学方法每年为公司增加了 6000 万美元的收入，今年收入将增加到 1.05 亿美元。该公司因为该算法获得了 2008 年 Franz Edelman 运筹学和管理科学国际竞赛第一名。

来源：L. Kroon, D. Huisman, E. Abbink, P. -J. Fioole, M. Fischetti, G. Maróti, A. Schrijver, A. Steenbeck, and R. Ybema, "The New Dutch Timetable: The OR Revolution," *Interfaces*, **39**(1): 6-17, Jan. -Feb. 2009.（我们的网址提供了本文链接：www.mhhe.com/hillier。）

2010 年，西北航空公司完成了合并，Delta 航空公司每天的班次超过 2500 个，拥有属于 10 种不同机型的 450 架飞机。它们运用巨大的整数规划模型(大概 40000 个约束条件、20000 个 0-1 变量和 40000 个一般整数变量)解决飞机安排的问题。每年为 Delta 航空公司节省大约 1 亿美元。

一个很相似的应用是机组人员安排问题。在这里，不是为航班分配机型，而是把飞行员和乘务人员分配给某条航线。因此。每条可行的航线都是从机组人员的基地出发再返回同一基地，对每条航线，是或否的决策是：是否应该指定某队机组人员飞行某航线？

$$\text{它的决策变量}=\begin{cases}1, & \text{是}\\0, & \text{否}\end{cases}$$

目标是使机组人员要飞行所有班次，且所花的费用最小。

此类的一个比较完整的建模案例将在 12.4 节给出。

与航空公司有关的一个问题是：有时由于天气恶劣、飞机机械故障或人员缺乏等原因，出现航班延误或取消的情况时，需要很快调整机组人员的安排。正如 2.2 节(参考文献[A12])所描述的，大陆航空公司在发生上述情况时，使用一个基于 BIP 的复杂决策支持系统进行机组人员安排，第 1 年就节省了 4000 万美元。

航空公司所面临的问题在其他运输业同样会出现。因此,运筹学研究在航空业的一些应用也扩展到其他部门,包括铁路运输。例如,本章第二个应用案例描述了荷兰铁路公司运用线性规划和约束规划(见12.9节)解决了实际问题,并获得广泛认可。

12.3 0-1变量在模型构建中的创新应用

前面我们看了一些例子,这些问题的基本决策结果都为是或否,因此引入0-1变量表示这些决策。现在我们看一些0-1变量的其他有效应用。特别地,有时这些变量有助于解决复杂模型问题,并将它们重新建模,使之成为纯或混合IP问题。

如果问题的原模型对应于一个整数规划模型或线性规划模型,只是在模型中的组合关系稍有不同,那么,就可以对它们重新建模,通过问题表达这些组合关系,而对问题的回答只能为是或否,于是,辅助0-1变量被引入模型以表达是或否的决策(与其作为原问题的决策变量考虑,不如在问题模型中引入一个辅助的二元变量帮助建立一个纯或混合的BIP模型)。引进这些变量可以将问题变成一个混合整数规划(MIP)问题(或一个纯IP问题,如果所有的原始变量都要求是整数值)。

下面介绍一些可以用这种方法解决的例子。x_j 表示问题的原始变量(可以是连续变量或者是整数变量),y_i 表示用于重新建模的辅助0-1变量。

12.3.1 "或"约束

考虑这种重要情况:在两个约束条件之间选择1个,即只有一个约束条件(两者之中任一)是必须保留的(尽管另一个也可以保留,但我们不要求这么做)。例如,为了达到某种目的,在两种资源中择其一使用,因此,在计算上,有必要把其中一种资源的可用性约束保留下来。为了说明适用于这种情况,假设整个问题的一个要求为

$3x_1+2x_2\leqslant 18$

或

$x_1+4x_2\leqslant 16$

也就是说,两个不等式中至少有一个要保留,但不必保留两个。为了使这个要求满足线性规划模型,必须改变它的形式,因为在线性规划模型中,所有约束条件都必须被保留下来。假设 M 是一个很大的正数,则这个要求可以重写为

$3x_1+2x_2\leqslant 18$

$x_1+4x_2\leqslant 16+M$

或

$3x_1+2x_2\leqslant 18+M$

$x_1+4x_2\leqslant 16$

关键是把 M 加到约束式的右边后,约束条件不再起作用,因为满足该问题其余约束条件的解都必然满足该约束(这个公式假设总体问题的可行解集是一个有限集合,而 M 足够大以至于不会漏掉任何一个可行的解)。这等价于下面这种约束条件集:

$3x_1+2x_2\leqslant 18+My$

$x_1+4x_2\leqslant 16+M(1-y)$

因为辅助变量 y 必须是0或1,这个公式保证原始约束条件中的一个被满足,而另一个实际

上是不起作用的。把这个新约束集加入总体模型的其他约束集中,使之成为一个纯或混合的 IP 问题(主要取决于 x_j 是整数变量还是连续变量)。

我们先前讨论了通过问题表达组合联系,而对问题的回答只能为是或否,这种方法与之直接相关。这里涉及的组合联系主要是模型的其他约束条件与可选约束的第一个的组合,然后是与第二个的组合。这两种约束组合哪一种更好呢(根据可以取得的目标函数值)？为了用是或否来重写这个问题,我们提出两个补充问题。

(1) 公式 $x_1+4x_2\leqslant16$ 应该作为被保留的约束吗?

(2) 公式 $3x_1+2x_2\leqslant18$ 应该作为被保留的约束吗?

因为两个问题中的一个要被肯定地回答,所以我们让 0-1 变量 y 和 $1-y$ 相应地表示是或否决策。因此,当第一个问题回答是(第二个问题回答否)时,$y=1$;当第二个问题回答是(第一个回答否)时,$1-y=1$(也就是 $y=0$)。既然 $y+1-y=1$(其中一个回答是)自动满足,就不再需要加入另一个约束迫使这两个决策是互斥的了(如果用单独的 0-1 变量 y_1 和 y_2 表示这两个是或否决策,那么,还需要附加额外的约束 $y_1+y_2=1$ 使之互斥)。

对于这种方法,以后我们将在更常见的例子中正式介绍。

12.3.2 保留 N 个约束条件中的 K 个

考虑这样一个例子:总体模型有一个大小为 N 的可能约束集,而只有其中的 K 个约束条件要保留下来(假设 $K<N$)。优化过程的部分工作是选择 K 个约束的组合使目标函数达到可能的最佳值。其余没有被选择的 $N-K$ 个约束条件,实际上,这个问题不起作用,虽然可行解可能恰巧也能使它们中的部分成立。

本例是前面假设 $K=1,N=2$ 例子的一般化。N 个可能的约束如下所示:

$$f_1(x_1,x_2,\cdots,x_n)\leqslant d_1$$
$$f_2(x_1,x_2,\cdots,x_n)\leqslant d_2$$
$$\vdots$$
$$f_N(x_1,x_2,\cdots,x_n)\leqslant d_N$$

然后,应用与前述例子相同的逻辑,可以找到必须保留 K 个约束条件所要求的一个等式:

$$f_1(x_1,x_2,\cdots,x_n)\leqslant d_1+My_1$$
$$f_2(x_1,x_2,\cdots,x_n)\leqslant d_2+My_2$$
$$\vdots$$
$$f_N(x_1,x_2,\cdots,x_n)\leqslant d_N+My_N$$
$$\sum_{i=1}^{N} y_i = N - K$$

且 y_i 是 0-1 变量,$i=1,2,\cdots,N$。

式中,M 是一个极大正数。对每一个二值变量 $y_i(i=1,2,\cdots,N)$,注意到,$y_i=0$ 时,$My_i=0$,这使得新约束 i 变成初始约束 i。若 $y_i=1$,d_i+My_i 变得相当大(仍然假定在一个范围内),以至于新约束 i 能够自动被任一满足其余新约束的解所满足,它消除了初始约束 i 的作用。因此,因为对 y_i 的约束保证了其中的 K 个变量等于 0 而剩下的等于 1,初始约束中的 K 个变量将不会改变,而另外($N-K$)个实际将被剔除。通过对整个问题应用适当的算法,能够决定哪些约束条件被保留在模型中,以得到对所有变量最佳的解。

12.3.3 有 N 个可能取值的函数

考虑下面的情况:一个给定函数要求取给定 N 个值中的一个,如下所示:

$f(x_1,x_2,\cdots,x_n)=d_1$ 或 $d_2,\cdots,d_N$

一种特殊情况为

$$f(x_1,x_2,\cdots,x_n)=\sum_{j=1}^{n}a_jx_j$$

类似一个线性规划约束条件的左边项。另一种特殊情况是 $f(x_1,x_2,\cdots,x_n)=x_j$,即对于给定的 j,必须取给定的 N 个值之一。

等价的 IP 方程如下表示:

$$f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{N}d_iy_i$$

$$\sum_{i=1}^{N}y_i=1$$

且 y_i 是 0-1 变量,$i=1,2,\cdots,N$。

因此,新的约束集将会取代整个问题的要求。该约束集提供了一个等价形式,因为必须有个 y_i 等于 1,而其他的等于 0,所以正好有一个 d_i 被选为函数值。在这个例子中,提出了 N 个是或否问题,那么,是否应该选择 $d_i(i=1,2,\cdots,N)$ 作为函数值呢?因为 y_i 表示相应是或否决策,所以第 2 个约束条件使得它们成为互斥的可选择。

为了说明为什么会出现这种情况,重新考虑 3.1 节提出的 Wyndor Glass 公司的问题。在第 3 个工厂,每周有 18h 的生产时间可以用来生产两种新产品,或者生产即将投入生产的某些产品。为了保留一些生产力来生产未来的产品,管理层现在想限制两种新产品的生产时间,每周 6h、12h 或者 18h。因此,初始模型的第 3 个约束($3x_1+2x_2\leqslant18$)现在变成

$$3x_1+2x_2=6 \text{ 或 } 12 \text{ 或 } 18$$

在之前的符号中,$N=3$,而 $d_1=6,d_2=12,d_3=18$。因此,管理层的新要求可以用如下公式表示:

$$3x_1+2x_2=6y_1+12y_2+18y_3$$

$$y_1+y_2+y_3=1$$

且 y_1、y_2、y_3 是 0-1 变量。

现在这个新问题的总体模型由两部分构成,原模型(见 3.1 节)加上替换原始第 3 个约束的新约束集。该替换产生一个易处理的 MIP 模型。

12.3.4 固定支出问题

当举办一项活动时,产生固定支出或组织成本的情况是很常见的。例如,一个工厂接下生产一批产品的订单后,必须建立相应的生产设施以启动生产,这时就产生了一笔费用。在这种情况下,项目的总开销就是与项目进度相关的可变支出以及用来启动项目的启动资金。可变支出通常大致与项目进度成比例。如果是这种情况,项目(称为项目 j)的总开销可以用如下的函数形式表示:

$$f_j(x_j)=\begin{cases}k_j+c_jx_j, & x_j>0\\ 0, & x_j=0\end{cases}$$

式中:x_j 表示项目 j 的进度($x_j\geqslant0$);k_j 表示组织成本;c_j 表示多生产一个单位所带来的成本。如

果不是因为组织成本 k_j，成本结构表明，可以用线性规划模型决定项目最佳的生产进度。幸运的是，即使有 k_j，仍然可以使用 MIP 模型。

为了建立总体模型，假设有 n 个项目，它们的成本结构如前所述（每个项目的 $k_j \geq 0$，对于某些 $j=1,2,\cdots,n$ 来说，$k_j>0$），这样，问题表示为

Min　$Z=f_1(x_1)+f_2(x_2)+\cdots+f_n(x_n)$

s. t. 给定的线性规划约束条件

为了将问题转化为 MIP 形式，我们提出 n 个问题，这些问题必须回答是或者不是。也就是说，对于每一个 $j=1,2,\cdots,n$，项目 j 应该实施（$x_j>0$）吗？每一个是或否决策用一个辅助 0-1 变量 y_j 表示，所以

$$Z = \sum_{j=1}^{n} (c_j x_j + k_j y_j)$$

其中

$$y_j = \begin{cases} 1, & x_j>0 \\ 0, & x_j=0 \end{cases}$$

因此，为可以被看成可能决策，与 12.1 节中提到的类型类似（但不等同），假设 M 是一个极大正数，它大于任意 $x_j(j=1,2,\cdots,n)$ 的最大可能值，那么约束：

$$x_j \leq My_j, \quad j=1,2,\cdots,n$$

将保证当 $x_j>0$ 时 $y_j=1$ 而不是 0。剩下的一个难题是当 $x_j=0$ 时，y_j 自由地选择 0 或者 1。幸运的是，这个难题会因目标函数自身的特点而被自动解决。$k_j=0$ 的情况将会被忽略，因为这时 y_j 将从模型中去掉。所以我们考虑另一种情况，也就是 $k_j>0$。当 $x_j=0$ 时，既然约束集允许 y_j 在 $y_j=0$ 和 $y_j=1$ 之间选择，$y_j=0$ 必然产生一个比 $y_j=1$ 更小的 Z 值。既然如此，因为目标是最小化 Z。所以产生最优解算法会在 $x_j=0$ 的时候选择令 $y_j=0$。

总之，固定支出问题的 MIP 模型如下：

Min $Z = \sum_{j=1}^{n} (c_j x_j + k_j y_j)$

s. t. 原始约束条件，加上

$x_j - My_j \leq 0$

且 y_j 是 0-1 变量 $j=1,2,\cdots,n$。

如果 x_j 也被限制为整数，那么，它将会变成一个纯 IP 问题。

为了说明这种方法，再次回顾 3.4 节所介绍的 Nori&Leets 公司空气污染问题。考虑第一种消除污染的方法，增加烟囱的高度，实际上包括为增高烟囱而做准备产生的固定支出和大致与增高量成比例的可变支出。在转化为模型中相应的年支出后，对于每一个鼓风炉和平炉来说，固定支出将会是 200 万美元，而可变支出如表 3.14 所列。因此，在之前的符号中，$k_1=2$，$k_2=2$，$c_1=8$ 以及 $c_2=10$，这里目标函数是以百万美元为单位的。因为另一种消除污染的方法不涉及任何固定支出，即对于 $j=3,4,5,6$，$k_j=0$，所以，此问题的新 MIP 模型可表示如下：

Min　　$Z=8x_1+10x_2+7x_3+6x_4+11x_5+9x_6+2y_1+2y_2$

s. t.

3.4 节中给出的约束条件，加上

$x_1 - My_1 \leq 0$

$x_2 - My_2 \leq 0$

且 y_1、y_2 是 0-1 变量。

12.3.5 一般整数变量的二值表示

假设有一个纯 IP 问题,其中大多数变量都是 0-1 变量,但是一般整数变量的存在使我们没有行之有效的 BIP 算法解决这个问题。一种绕开该难点的有效办法是对每一个一般整数变量进行二值表示。特别地,如果一个整数变量 x 的取值范围为

$0 \leqslant x \leqslant u$

而 N 被定义成如下所示的整数:

$2^N \leqslant u \leqslant 2^{N+1}$

那么,x 的二值表示为

$$x = \sum_{i=0}^{N} 2^i y_i$$

式中:变量 y_j 是(辅助)二值变量。用二值表示替换每一个一般整数变量(用不同的辅助二值变量集),这样就将整个问题简化成一个 BIP 模型。

例如,假设一个 IP 问题有两个一般整数变量 x_1 和 x_2 以及许多二值变量,再假设 x_1 和 x_2 都是非负的,函数约束包括:

$x_1 \qquad \leqslant 5$

$2x_1+3x_2 \quad \leqslant 30$

这些约束条件显示对于 x_1,$u=5$,而对于 x_2,$u=10$,所以根据 N 的定义,对应于 N,$N=2(2^2 \leqslant 5<2^3)$,而对 x_2,$N=3(2^3 \leqslant 10<2^4)$。因此,这些变量的二值表示为

$x_1 = y_0+2y_1+4y_2$

$x_2 = y_3+2y_4+4y_5+8y_6$

在用上式替换在所有函数约束和目标函数中的变量后,上述的两个函数约束变成

$y_0+2y_1+4y_2 \leqslant 5$

$2y_0+4y_1+8y_2+3y_3+6y_4+12y_5+24y_6 \leqslant 30$

注意:x_1 的可能值对应于向量(y_0,y_1,y_2)的可能值,同样,x_2 对应于(y_3,y_4,y_5,y_6)。例如,$x_1=3$对应于$(y_0,y_1,y_2)=(1,1,0)$,而当 $x_2=5$ 时,$(y_3,y_4,y_5,y_6)=(1,0,1,0)$。

对于一个 IP 问题,其所有的变量都是(有界的)一般整数变量,我们有可能利用相同的方法将其简化为一个 BIP 模型。然而,大多数情况下这么做并不可取,因为其中包含的变量将会急剧增加。对一个原始 IP 模型使用一个好的 IP 算法,通常会比在一个更大的 BIP 模型上应用一个好的 BIP 算法更为有效。

概括地说。对于本节介绍的所有用辅助二值变量表达的形式,我们需要小心谨慎。这种方法有时要求加入大量变量,这可能使模型在计算上是不可行的(12.5 节将会讨论可求解的 IP 问题的规模)。

12.4 一些建模举例

我们现在给出一些例子,阐明各种用 0-1 变量建模的方法,包括前面几节所讨论的内容。为了清楚起见,例子都很简单(包含很多 0-1 变量和约束条件的大型案例可以通过本书网站查找)。实际应用时,在一个更为巨大的模型中,这些模型是典型小的组成部分。

12.4.1　例 1　当决策变量是连续变量时的选择

GOOD PRODUCTS 公司研发部开发了 3 种可行的新产品。然而,为了使产品的生产线不至过于多样化,管理层决定实施以下限制。

限制 1:在 3 种新产品中,至多有 2 种被投入生产。

每一种产品可能由 2 个工厂中的任何一个生产。出于管理方面的考虑,管理层实行了第 2 条限制。

限制 2:2 个工厂中,仅有一个能作为新产品的唯一生产者。

对于 2 个工厂来说,每种新产品的单位生产成本都是相同的。然而,由于 2 个工厂的生产设备不同,每种产品的单位生产时间可能是不同的。数据在表 12.2 中给出。还有其他一些信息,包括在投产后每周每种新产品的预期销售数量。目标是选择新产品、工厂和生产新产品的生产率,以使利润最大化。

表 12.2　案例 1 的数据(GOOD PRODUCTS 公司问题)

	单位产品的生产时间/h			每周可用生产时间/h
	产品 1	产品 2	产品 3	
工厂 1	3h	4h	2h	30h
工厂 2	4h	6h	2h	40h
单位利润/千美元	5	7	3	
可能销售量/周	7	5	9	

在某种程度上,这个问题类似一个标准的产品混合问题,正如 3.1 节所描述的 Wyndor Glass 公司的例子。实际上,如果我们去掉两个约束条件并且满足 12.2 列出的 2 个工厂(所以这 2 个工厂生产产品的工艺是不同的)生产每种新产品所用的时间。那么,原问题就变成了此类问题。特别地,令 x_1、x_2、x_3 分别代表新产品的生产率,那么,模型变为

$$\text{Max}\quad Z=5x_1+7x_2+3x_3$$

s. t.

$$3x_1+4x_2+2x_3\leqslant 30$$

$$4x_1+6x_2+2x_3\leqslant 40$$

$$x_1\leqslant 7$$

$$x_2\leqslant 5$$

$$x_3\leqslant 9$$

且

$$x_1\geqslant 0,x_2\geqslant 0,x_3\geqslant 0$$

对于实际问题,必然会给模型加入限制条件 1。

严格大于零的决策变量(x_1,x_2,x_3)必须≤2。

这个约束条件无法用于一个线性或整数规划模型,所以关键问题是怎样把它转化成此类模型,以便使用相应的算法求解总体模型。如果决策变量是 0-1 变量,那么,约束条件就可以表达为 $x_1+x_2+x_3\leqslant 2$。然而,我们需要一个更为复杂的模型,不仅涉及辅助 0-1 变量,还包含连续变量。

第 2 个约束要求前 2 个约束条件($3x_1+4x_2+2x_3\leqslant 30$ 与 $4x_1+6x_2+2x_3\leqslant 40$)被下述约束条件代替:

$$3x_1+4x_2+2x_3\leqslant 30$$

或

$$4x_1+6x_2+2x_3\leqslant 40$$

至于哪个约束条件被保留,对应于选择哪个工厂生产新产品,我们在前面的章节已经讨论了怎样把约束转化成为一个线性或整数规划形式,再次用到一个辅助 0-1 变量。

使用辅助 0-1 变量建模:为了处理第一个限制条件,我们引入 3 个辅助 0-1 变量(y_1,y_2,y_3)它们的含义如下:

$$y_j=\begin{cases}1, & x_j>0 \text{ 被保留(生产产品 } j\text{)}\\ 0, & x_j=0 \text{ 必须保留(不能生产产品 } j\text{)}\end{cases}\qquad (j=1,2,3)$$

为了把该含义融入模型中,我们把 M(一个非常大的正数)加到约束条件当中:

$$x_1\leqslant My_1$$

$$x_2\leqslant My_2$$

$$x_3\leqslant My_3$$

$$y_1+y_2+y_3\leqslant 2$$

y_j 是 0-1 变量,$j=1,2,3$。

或约束与非负约束使决策变量的可行域是有限的(每个 $x_j\leqslant M$)。因此,在每个约束条件 $x_j\leqslant My_j$ 中,$y_j=1$ 使 x_j 能取到可行域中的任何值,而 $y_j=0$ 迫使 $x_j=0$(反过来,$x_j\geqslant 0$ 迫使 $y_j=1$,而 $x_j=0$ 时允许 $y_j=0$ 或 1)。结果,因为第 4 个约束条件令至多能有 2 个 $y_j=1$,所以它等价于至多能有 2 种新产品被投入生产。

为了处理第 2 个限制条件,我们引入第 2 个辅助 0-1 变量 y_4,其含义如下:

$$y_4=\begin{cases}1, & 4x_1+6x_2+2x_3\leqslant 40 \text{ 必须保留(关闭工厂 2)}\\ 0, & 3x_1+4x_2+2x_3\leqslant 30 \text{ 必须保留(关闭工厂 1)}\end{cases}$$

正如 12.3 节所论述的,通过加上如下约束条件表达该含义:

$$3x_1+4x_2+2x_3\leqslant 30+My_4$$

$$4x_1+6x_2+2x_3\leqslant 40+M(1-y_4)$$

y_4 是 0-1 变量。

结果,我们把所有变量移到约束条件的左边后,完整的模型为

$$\text{Max}\quad Z=5x_1+7x_2+3x_3$$

s. t.

$$x_1\leqslant 7$$

$$x_2\leqslant 5$$

$$x_3\leqslant 9$$

$$x_1-My_1\leqslant 0$$

$$x_2-My_2\leqslant 0$$

$$x_3-My_3\leqslant 0$$

$$y_1+y_2+y_3\leqslant 2$$

$$3x_1+4x_2+2x_3-My_4\leqslant 30$$

$$4x_1+6x_2+2x_3+My_4\leqslant 40+M$$

且

$x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0$

y_j 是 0-1 变量,$j=1,2,3,4$。

现在这是一个 MIP 模型,3 个变量(x_j)不要求是整数,4 个 0-1 变量,所以可以用 MIP 算法求解这个模型。求得结果是(在用一个相当大的数替换 M 之后)①:最优解为 $y_1=1, y_2=0, y_3=1, y_4=1, x_1=5\frac{1}{2}, x_2=0$ 和 $x_3=9$;也就是说,选择生产第 1 个和第 3 个新产品,选择第 2 个工厂生产新产品,并且第 1 个产品的生产率是每周 $5\frac{1}{2}$ 个,第 3 个产品的生产率是每周 9 个。总利润是每周 54500 美元。

12.4.2　例 2　违反比例性

SUPERSUDS 公司正在为明年的新产品开发制定市场计划。对于 3 个新产品,公司决定在国内电视上购买 5 个时段播出广告。问题是怎样把这 5 个时段分配给 3 个产品,对每个产品最多只能分配 3 个(最小值为零)时段。

表 12.3 显示了把时段 0、1、2、3 分配给每个产品带来的预期效果,这个效果是由在该时段播出广告所产生销售增长带来的利润(以百万美元为单位)来衡量的,也要考虑到制作广告和购买时段所花的费用。目标是把 5 个时段分配给产品使利润最大化。

这个小问题可以很容易用动态规划或推断解决(最优解是分配 2 个时段给产品 1,3 个时段给产品 3,不为产品 2 做广告)。然而,我们将给出两个不同的 BIP 模型阐明这个问题。如果这个小问题要与一个涉及为公司所有新产品的市场活动分配资源的大 IP 模型组合起来,则需要应用这类模型。

0-1 变量的模型:一个自然的建模方式是令 x_1、x_2、x_3 分别为分配给每种产品的电视广告的时间段数。每个 x_j 对目标函数的贡献由表 12.3 的相应列给出。然而,每一列都不满足 3.3 节提出的成比例的假定。因此,我们无法用这些整数决策变量写出一个线性的目标函数。

表 12.3　案例 2 数据(SUPERSUDS 公司的问题)

	利润		
	产品		
电视时段数	1	2	3
0	0	0	0
1	1	0	-1
2	3	2	2
3	3	3	4

现在看一下当我们为每个正整数 $x_i=j(j=1,2,3)$ 引入一个辅助 0-1 变量 y_{ij} 会发生什么,y_{ij} 的定义为

$$y_{ij}=\begin{cases}1, & x_i=j\\ 0, & \text{否则}\end{cases}$$

(例如,$y_{21}=0, y_{22}=0$ 以及 $y_{23}=1$,意味着 $x_2=30$)结果这个线性 BIP 模型为

① 实际上,给 M 取值时,M 应足够大,以免消除任何可行解,但也应尽可能小,这是为了避免在进行 LP 松弛(避免数字上的不稳定)时,无谓地扩大可行域的范围。经过对约束条件的检查,M 的最小可行值为 $M=9$。

$$\text{Max} \quad Z=y_{11}+3y_{12}+3y_{13}+2y_{22}+3y_{23}-y_{31}+2y_{32}+4y_{33}$$

s. t.

$$y_{11}+y_{12}+y_{13}\leqslant 1$$

$$y_{21}+y_{22}+y_{23}\leqslant 1$$

$$y_{31}+y_{32}+y_{33}\leqslant 1$$

$$y_{11}+2y_{12}+3y_{13}+y_{21}+2y_{22}+3y_{23}+y_{31}+2y_{32}+3y_{33}=5$$

且每个 y_{ij}是 0-1 变量。

注意:前 3 个约束条件确保每个 x_i 只能取一个可能值(在这里 $y_{i1}+y_{i2}+y_{i3}=0$ 对应 $x_i=0$,这对于目标函数没有贡献)。最后一个约束条件确保 $x_1+x_2+x_3=5$。根据表 12.3,这个线性目标函数给出了总利润。

解这个 BIP 模型,最优解为

$y_{11}=0,y_{12}=1,y_{13}=0$,故 $x_1=2$

$y_{21}=0,y_{22}=0,y_{23}=0$,故 $x_2=0$

$y_{31}=0,y_{32}=0,y_{33}=1$,故 $x_3=3$

用 0-1 变量建模的另一个例子:现在重新定义辅助 0-1 变量 y_{ij}如下:

$$y_{ij}=\begin{cases}1, & x_i\geqslant j\\ 0, & \text{否则}\end{cases}$$

故不同之处在于,当 $x_i\geqslant j$,而不是 $x_i=j$ 时,有 $y_{ij}=1$。因此,有

$x_i=0\Rightarrow y_{i1}=0,y_{i2}=0,y_{i3}=0$

$x_i=1\Rightarrow y_{i1}=1,y_{i2}=0,y_{i3}=0$

$x_i=2\Rightarrow y_{i1}=1,y_{i2}=1,y_{i3}=0$

$x_i=3\Rightarrow y_{i1}=1,y_{i2}=1,y_{i3}=1$

所以

$$x_i=y_{i1}+y_{i2}+y_{i3}$$

因为 $y_{i2}=1$ 时才可能有 $y_{i1}=1,y_{i2}=1$ 时才可能有 $y_{i3}=1$,这些定义通过加入下述约束条件来实现:

$y_{i2}\leqslant y_{i1}$ 并且 $y_{i3}\leqslant y_{i2},i=1,2,3$。

y_{ij}的新定义同样改变了目标函数,图 12.1 显示了目标函数中产品 1 的部分。因为 y_{11}、y_{12}、y_{13}提供了 x_1 值(从 0 开始)的连续增加(如果有),y_{11}、y_{12}、y_{13}的系数分别由表 12.3 中产品 1 相应的增长量给出($1-0=1,3-1=2,3-3=0$)。这些增长量在图 12.1 中是斜率,得到 $y_{11}+2y_{12}+0y_{13}$作为目标函数的产品 1 部分。注意:对 3 种产品应用这个方法必然会得到一个线性目标函数,我们把所有变量移到约束条件的左边,最后得到完整的 BIP 模型:

$$\text{Max} \quad Z=y_{11}+2y_{12}+2y_{22}+y_{23}-y_{31}+3y_{32}+2y_{33}$$

s. t.

$$y_{12}-y_{11}\leqslant 0$$

$$y_{13}-y_{12}\leqslant 0$$

$$y_{22}-y_{21}\leqslant 0$$

$$y_{23}-y_{22}\leqslant 0$$

$$y_{32}-y_{31}\leqslant 0$$

$$y_{33}-y_{32}\leqslant 0$$

$$y_{11}+y_{12}+y_{13}+y_{21}+y_{22}+y_{23}+y_{31}+y_{32}+y_{33}=5$$

且每个 y_{ij}是 0-1 变量。

解这个 BIP 模型,最优解为

$y_{11}=1, y_{12}=1, y_{13}=0$,故 $x_1=2$

$y_{21}=0, y_{22}=0, y_{23}=0$,故 $x_2=0$

$y_{31}=1, y_{32}=1, y_{33}=1$,故 $x_3=3$

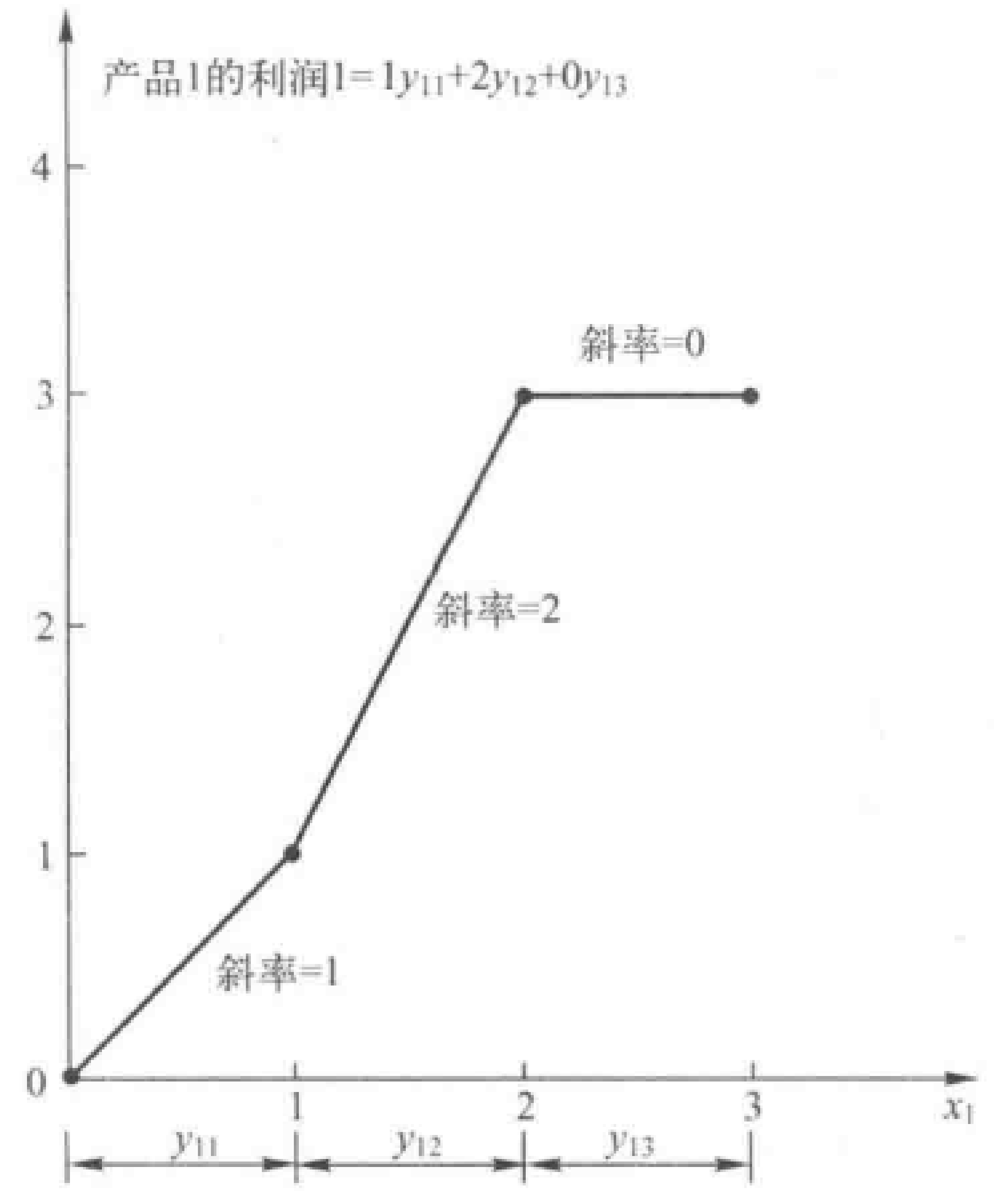

图 12.1　产品 1 播出 x_1 个时段的广告导致销售增加而带来的利润,
其中范围绘出了例 2 的第 2 个 BIP 模型中目标函数的相应系数(SUPERSUDS 公司的问题)

除了个人偏好外,在这个 BIP 模型与前一个模型之间没有优劣之分。它们有相同数目的 0-1 变量(决定 BIP 问题计算效果的主要因素)。它们都有一些特定结构(第 2 个模型中存在互斥的约束条件,第 2 个模型中存在相互关联的决策)提高模型效率。第 2 个模型较第 1 个具有更多的函数约束。

12.4.3　例 3　覆盖所有特征

西南航空公司需要分配它的机组人员,使其覆盖所有将要飞行的航班。我们研究的重点是,为驻扎在旧金山的 3 队机组人员指定如表 12.4 第 1 列所列的所有航班,另外 12 列显示的是 12 条可行的航线(每列的数字代表该航线覆盖的航班及其顺序号)。在这些航线中,需要选择 3 条(一队机组人员负责一条航线),但是要保证覆盖所有的航班(允许在一个航班上有多队机组人员,多出来的机组人员被视为乘客,但是工会合同要求,多余的机组人员被视为正在工作,应向其支付工资)。把一队机组人员分配给某条航线的成本由表中的最后一行列出(以千美元为单位)。目标是分配 3 队机组人员,使他们飞行所有航班的总成本最小。

用 0-1 变量建模:有 12 条可行的航线,相应地,我们有 12 个是或否的决策:

应该指定一队机组人员飞行航线 j 吗?($j=1,2,\cdots,12$)

$$x_j=\begin{cases}1, & \text{指定一组人员飞行航线 } j\\ 0, & \text{否则}\end{cases}$$

该模型最有趣的地方是,每个约束条件实际上是保证一个航班被覆盖。例如,考虑表 12.4 的最后一个航班(西雅图到洛杉矶)。5 条航线(也就是 6 航线、9 航线、10 航线、11 航线和 12 航

线)包括该航班,因此,公司至少会选择其中的一条航线飞行。结果约束条件为

$x_6+x_9+x_{10}+x_{11}+x_{12}\geqslant 1$

对另外10个航班使用类似的约束,完整的BIP模型为

$$\text{Min } Z=2x_1+3x_2+4x_3+6x_4+7x_5+5x_6+7x_7+8x_8+9x_9+9x_{10}+8x_{11}+9x_{12}$$

s. t.

$x_1+x_4+x_7+x_{10}\geqslant 1$(旧金山-洛杉矶)

$x_2+x_5+x_8+x_{11}\geqslant 1$(旧金山-丹佛)

$x_3+x_6+x_9+x_{12}\geqslant 1$(旧金山-西雅图)

$x_4+x_7+x_9+x_{10}+x_{12}\geqslant 1$(洛杉矶-芝加哥)

$x_1+x_6+x_{10}+x_{11}\geqslant 1$(洛杉矶-旧金山)

$x_4+x_5+x_9\geqslant 1$(芝加哥-丹佛)

$x_7+x_8+x_{10}+x_{11}+x_{12}\geqslant 1$(芝加哥-西雅图)

$x_2+x_4+x_5+x_9\geqslant 1$(丹佛-旧金山)

$x_5+x_8+x_{11}\geqslant 1$(丹佛-芝加哥)

$x_3+x_7+x_8+x_{12}\geqslant 1$(西雅图-旧金山)

$x_6+x_9+x_{10}+x_{11}+x_{12}\geqslant 1$(西雅图-洛杉矶)

$$\sum_{j=1}^{12} x_j = 3$$

且 x_j 是0-1变量,$j=1,2,\cdots,12$。

该BIP模型的一个最优解为

$x_3=1$(航线3)

$x_4=1$(航线4)

$x_{11}=1$(航线11)

其余 $x_j=0$,总成本为18000美元(另一最优解是 $x_1=1,x_5=1,x_{12}=1$ 其他 $x_j=0$)。

表12.4 例3的数据(西南航空公司的问题)

	可行的航线											
航班	1	2	3	4	5	6	7	8	9	10	11	12
1. 旧金山至洛杉矶	1			1			1			1		
2. 旧金山至丹佛		1			1			1			1	
3. 旧金山至西雅图			1			1			1			1
4. 洛杉矶至芝加哥				2			2		3	2		3
5. 洛杉矶至旧金山	2					3				5	5	
6. 芝加哥至丹佛				3	3				4			
7. 芝加哥至西雅图							3	3		3	3	4
8. 丹佛至旧金山		2		4	4				5			
9. 丹佛至芝加哥					2			2			2	
10. 西雅图至旧金山			2				4	4				5
11. 西雅图至洛杉矶						2			2	4	4	2
费用/千美元	2	3	4	6	7	5	7	8	9	9	8	9

这个例子阐明了一类更常见的问题,即集合覆盖问题①。任何集合覆盖问题都可以用涉及一些可能的活动(如航线)与特征(如航班)的一般模型描述。每个活动处理一些特征但不是全部特征。目标是决定成本最小的活动组合,所有活动必须覆盖所有特征。因此,令 S_i 是覆盖特征 i 的所有活动的集合。集合 S_i 中,至少有一个被选择,所以对于每个特征 i,一个约束条件为

$$\sum_{j \in S_i} x_j \geq 1$$

另一类相关问题是集合分离问题,把每个约束条件变成

$$\sum_{j \in S_i} x_j = 1$$

所以每个集合 S_i 中恰好有一个被包括在选择出来的活动中。用于该例,也就是对选出的航线,每个航班只能在一条航线中出现,消除了在任何航班上有多余机组人员的情况。

12.5　求解整数规划问题的若干展望

表面看来 IP 问题应该相对容易解决,毕竟我们能够采取很有效的方法解决线性规划问题,而它们的区别也仅仅是对于 IP 问题,我们需要考虑的解相对更少,事实上,具有有限可行域的纯 IP 问题一定有有限个可行解。

不幸的是,在这个推理过程中存在两个谬误。一个是有限个数的可行解保证了问题是容易解决的,然而,有限个数也可以是天文数字那么大,如考虑一个简单的 BIP 问题。如果有 n 个变量,则需要考虑 2^n 个解(随后其中一些解可能因为不符合函数约束条件而被舍去)。因此,n 的个数每增长一个,解的数量就会变成 2 倍。这种情况使得问题难度呈指数增长。如果 $n=10$,就有 1000 多个(1024)解,如果 $n=20$ 或 $n=30$,就有 10 亿多个解,依此类推。因此,对于有几十个变量的 BIP 问题,即使是最快的计算机也不可能一一列举(检查每一个解的可行性,如果可行,计算目标函数值),更不用说有同样多整数变量的一般 IP 问题了。但幸运的是,应用后面各节中的思路,目前最好的整数规划算法远远优于上述枚举的算法。这种改进在过去二三十年内十分明显。25 年前需用一年计算时间的 BIP 问题,用目前最好的商业软件只需几秒钟时间。速度的极大提高主要基于以下 3 个方面的巨大进展:BIP 算法及其他 IP 算法的改进、整数规划程序中频繁调用的线性规划算法的改进以及计算机(包括台式计算机)运算速度的提高。因此,目前规模十分巨大的 BIP 的求解,与过去 10 年相比具有了更大的可能性。今天最好的算法已能求解某些具有几十万个变量的纯 BIP 问题。然而,由于变量指数增长,即使是最好的算法,也不一定能解出每一个相对小的问题(有少于几百个 0-1 变量)。由于小问题自身的特点,有时候它们比那些规模大得多的问题要难解得多。

当要处理的变量是一般整型变量而非 0-1 变量时,能够求解的问题规模往往小得多。然而,仍然存在例外。

另一个谬误是从一个线性规划问题里去除一些可行解(非整型的)将使得问题更易于解决。相反,正是所有的这些可行解才保证了能够得到一个位于顶点的可行解(CPF 解)(也是一个相应基本可行解(BF 解)),这个 CPF 解是整个问题的最优解。这个保证是高效率实现单纯形法的关键。因此,线性规划问题通常比 IP 问题更容易求解。

因此,大多数成功的整数规划算法都通过把 IP 问题的一部分与相应的线性规划算法联系起

① 严格地讲,一个集合覆盖问题不包括任何其他函数的约束,如上例中最后一个函数约束。有时还假定目标函数的每个系数都被最小化为 1,在不满足这个假定时称其为加权的集合覆盖问题。

来,如单纯形法或对偶单纯形法。对给定的IP问题,从中去除变量的整约束,得到的相应线性规划问题通常称作它的LP松弛。在接下来两节里所描述的算法阐述了一个IP问题某些部分的一系列LP松弛是如何有效地解决整个IP问题的。

在一种特殊情况下,解决IP问题不再比用单纯形法解决它的LP松弛困难,也就是说,后者的最优解恰好满足IP问题的整数约束。当发生这种情况时,这个解也一定是IP问题的最优解,因为这个解是LP松弛所有可行解中的最优解,也包括了IP问题的所有可行解。因此,对于一个IP算法来说,用单纯形法求解LP松弛问题,首先检查这个偶然的结果是否已经产生,是很常见的事情。

通常来说,尽管发生LP松弛的最优解也是整数这种情况的概率很小,但是,事实上,存在几种特殊类型的IP问题,其结果一定是整数。在第9章和第10章,你已经看到这些特殊类型的典型代表,即最小费用流问题(带有整型参数)和它的特殊案例(包括运输问题、分配问题、最短路径问题和最大流问题)。这些类型的问题保证了这个最优解是整数,因为它们具有一种特殊的结构(如表9.6)。正如9.1节和10.6节关于整数解的性质里提到的,这种结构保证了每一个BF解都是整数。因此,这些特殊类型的IP问题可以看做线性规划问题,因为它们完全可以用改进的单纯形法求解。

尽管这种较大程度的简化并不常见,但实际上IP问题常常有某种特殊的结构,这种结构可以用来简化问题(12.4节的例2和例3就属于此类,因为它们有互斥的约束条件、可能决策约束条件或者集合覆盖约束条件)。有些时候,这类问题的很大一部分都能够被成功地解决。在整数规划问题上,用于求解一定特殊结构的、具有特殊用途的算法正变得越来越重要。

因此,有3个主要的因素决定了IP问题的计算难度,它们是:①整数变量的数量;②这些变量是0-1变量还是一般整型变量;③问题中的特殊结构。这种情况正好与线性规划相反,在线性规划中(函数)约束条件的数量要比变量的数量重要得多。在整数规划中,约束条件的数量也是比较重要的(尤其是如果正在求解LP松弛)。但是其重要性严格次于上面提到的3个因素。事实上,偶尔会出现,约束条件的数量增加了而计算时间却减少了,产生这种情况的原因是可行解的数量减少了。对MIP问题来说,整数变量的数量才是重要的,而不是所有变量的总数,因为连续变量几乎对计算量没有什么影响。

因为IP问题通常比线性规划问题难解得多,所以有时候采用一种近似程序是很吸引人的。首先用单纯形法求解LP松弛,然后,在最终结果中,把非整数值凑整成为整数值。这种方法对这些应用来说可能就足够了,尤其是变量的值非常大,以至于调整所产生的误差相对很小时。然而,应该注意这个方法有两个缺陷。

一个缺陷是线性规划的最优解,在凑整之后,并不一定是可行的。通常来说,预见到怎样调整能够保持可行性是困难的。在调整之后,甚至有必要将一些变量的值改变一个或更多单位。为了说明这种情况,考虑下面这个问题:

Max　$Z=x_2$

s. t.

$$-x_1+x_2\leqslant\frac{1}{2}$$

$$x_1+x_2\leqslant 3\frac{1}{2}$$

且

$$x_1\geqslant 0,x_2\geqslant 0$$

x_1、x_2 是 0-1 变量。

应用案例

2013 年，塔可钟公司在美国拥有近 5600 家快餐店，在 20 多个国家拥有 250 家快餐店。年均销售 20 亿份套餐。

塔可钟快餐店，每天的生意非常可观，特别是在饭点顾客非常拥挤。如何规划每个时段员工的数量及其应该完成工作是一件非常复杂和令人焦虑的问题。

为了解决该问题，塔可钟公司的管理者雇用了一个运筹学团队（包括一些咨询师）开发一套新的人力管理系统。该系统主要有 3 个功能模块：①一个预测模型，用来预测任何时刻的顾客订单；②一个仿真模型（如第 20 节所描述的），用来将顾客订单转换为人力需求；③一个整数规划模型，用来调度员工以满足人力需求同时最小化薪金。

该整数规划模型的整数决策变量是不同时间段每次轮班需要的员工数。轮班的时长也是决策变量（受最小和最大轮班时间的限制），该模型还包括连续决策变量，因此该模型称为混合整数规划（MIP）模型。主要的约束是每 15min 间隔内工作的员工数要大于或等于该间隔内需要的最少员工数（根据预测模型）。

该 MIP 模型与 3.4 节介绍的联合航空公司分配员工的线性规划模型类似。然而，主要的区别是塔可钟快餐轮班的工作员工比联合航空 100 个雇员少得多，因此必须将这些决策变量限制为整数（对于联合航空公司的非整数解可以将 100 多个雇员整数化，同时整数化后精确度的损失很小）。

通过应用 MIP 模型及人力管理系统，每年为塔可钟公司节约了大约 1300 万美元的人力成本。

资料来源：J. Hueter and W. Swart："An Integrated LaborManagement System for Taco Bell," *Interfaces*, **28**(1): 75-91, Jan.-Feb. 1998.（我们的网址提供了本文链接：www.mhhe.com/hillier。）

如图 12.2 所示，LP 松弛的最优解是 $x_1=1\frac{1}{2}$，$x_2=2$，但是不可能把非整数解 x_1 改进到 1 或 2（或其他任何整数）并且保持可行性。可行性的唯一保持方法是同时改变 x_2 的值。可以想象，当有成百上千个约束条件和变量时，这种改进会有多困难。

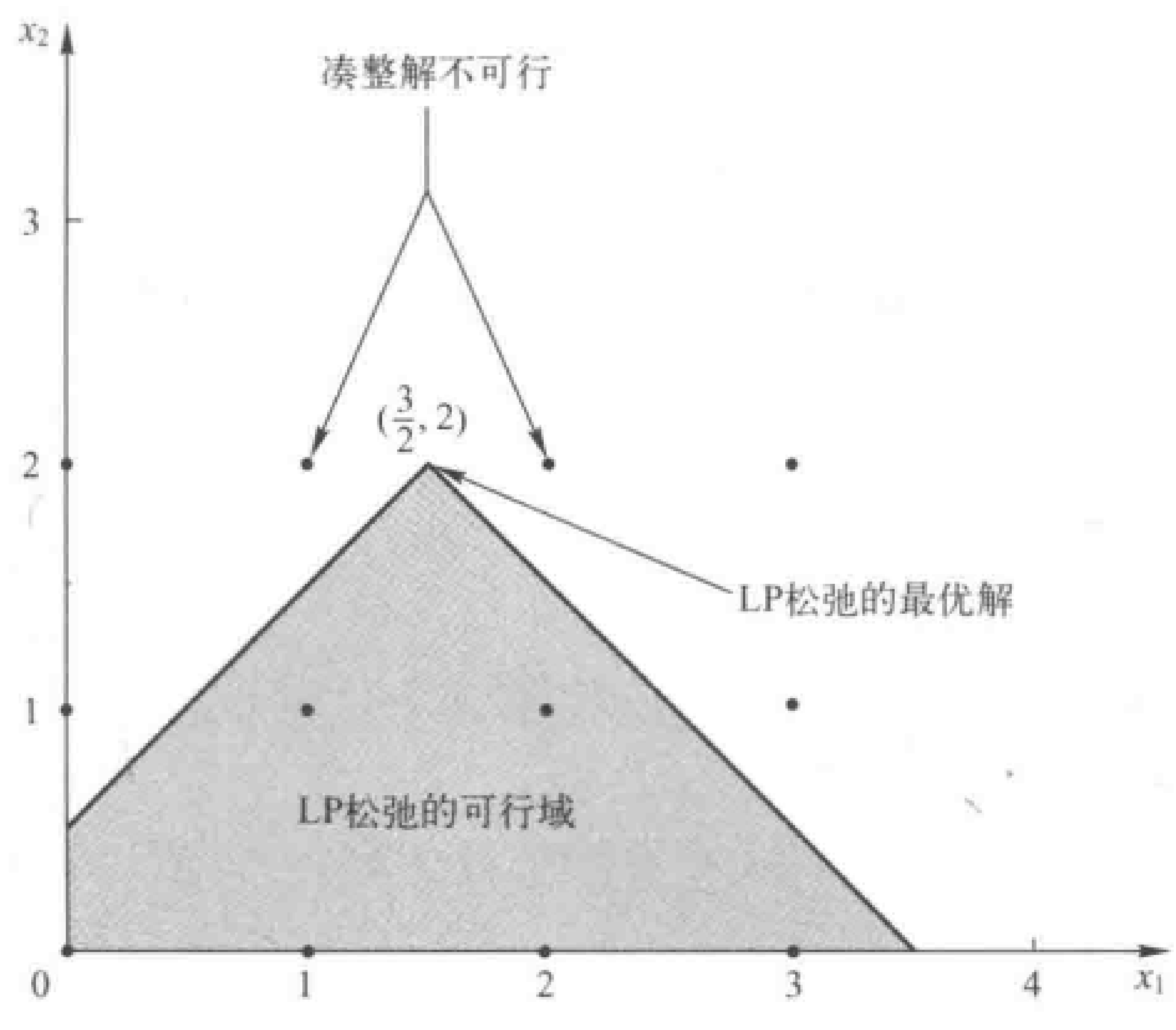

图 12.2　LP 松弛的最优无论怎么凑整，也不是可行解的 IP 问题的例子

即使 LP 松弛的最优解被成功改进，仍存在另一个缺陷。不能保证这个改进的解是最优整数解。实际上，从目标函数值的角度考虑，它可能远离了最优。这个事实可由以下问题来说明：

Max　$Z=x_1+5x_2$

s. t.

$x_1+10x_2\leqslant 20$

$x_1\leqslant 2$

且

$x_1 \geqslant 0, x_2 \geqslant 0$

x_1、x_2 是0-1变量。

因为只有两个决策变量,这个问题可以用图形描述,如图12.3所示。无论用图解法还是单纯形法都可以得到LP松弛的最优解是 $x_1 = 2, x_2 = \frac{9}{5}, Z = 11$。如果不能使用图解法(有更多变量的情况),那么,非整数变量 $x_2 = \frac{9}{5}$ 一般会朝着可行的方向调整到 $x_2 = 1$。最后的整数解是 $x_1 = 2, x_2 = 1$,得到 $Z = 7$。注意:这个解远离最优解 $(x_1, x_2) = (0,2), Z = 10$。

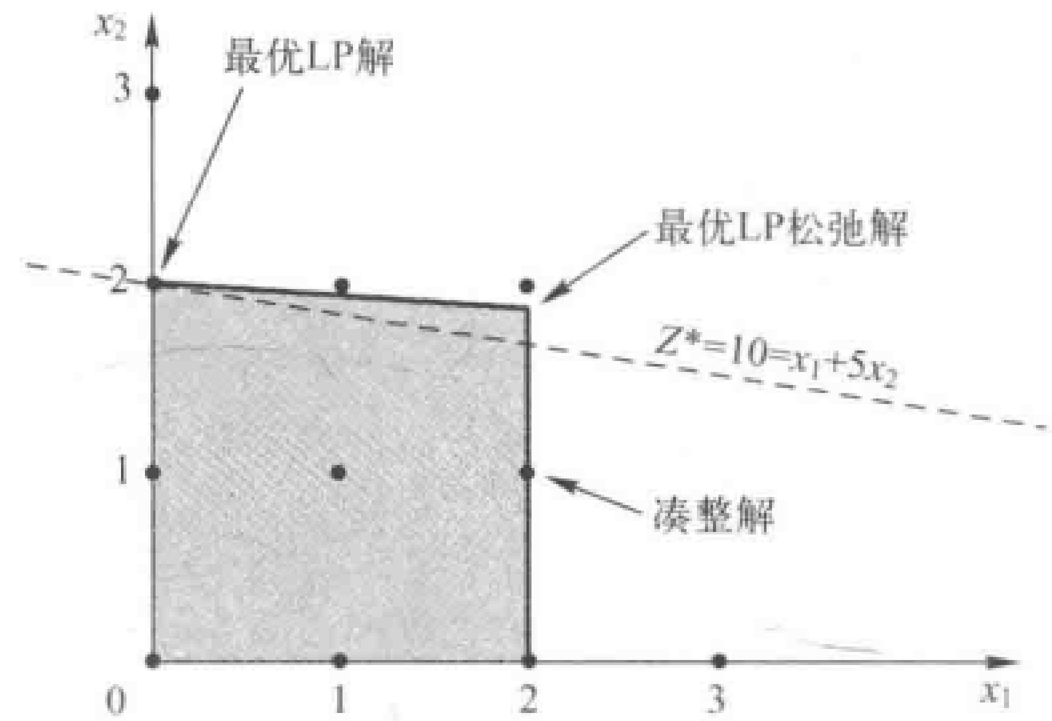

图12.3　对LP松弛最优解的凑整解远离了LP问题的最优解的例子

因为这两个缺陷,当IP问题过大而无法求解时,一种更好的方法是启发式的算法。这种算法对于大型问题是相当有效的,但是不能保证找到一个最优解。然而,在寻找非常好的可行解方面,这种算法比刚刚讨论的改进的方法要有效得多。

近年来,运筹学取得了一个令人兴奋的发展,即为处理各种组合问题如IP问题,而开发的高效启发式的算法(通常称作元启发式算法)发展快速。这些复杂的元启发式算法,甚至可以用于整数非线性规划问题,这种问题的局部最优解可能是远离全局最优解的。它们还可以用于各种联合最优问题,这种问题常常在一个有整数变量但约束条件比IP模型复杂的模型中出现。

回到整数线性规划,对一个小到可以求解的IP问题来说,有很多算法是可用的。但是,在计算效率方面,没有任何一种IP算法可以与单纯形法(除了问题的特殊性)相提并论。因此,开发IP算法仍然是一个活跃的研究领域。幸运的是,已经取得了一些令人兴奋的算法进展,并且预期在未来的几年中将获得更多成果。这些进展将在12.8节与12.9节论述。

在IP算法中,最常见、最传统的方法是使用分支定界技术和相关思想,枚举可行整数解,我们应该主要讨论这种方法。下一节将提出分支定界技术,并且用一个解BIP问题的基本分支定界算法来阐述。12.7节提出了另一个用于求解一般MIP问题的同类算法。

12.6　分支定界法及其在求解0-1整数规划中的应用

任何一个有界纯整数线性规划问题只有有限数目的可行解,所以我们很自然地考虑是否可以用类似枚举的方法找出最优解。不幸的是,正如前文所述,有限的数目通常也很大。因此,必须巧妙地设计枚举方法使我们只需检查一小部分可行解。例如,动态规划(见第11章)就为那些只有有限数目可行解的问题提供了这样的方法(虽然对大多数线性规划问题其效率并不高)。另一种方法则是分支定界法。这种方法及其变形已经成功地应用于各种运筹学问题,尤其在解决整数规划问题方面更为出色。

分支定界法的基本思想是拆分排除法。对于那些很难直接处理的大问题,我们把它拆分成越来越小的子问题,直到这些子问题能被处理。拆分(分支)的工作是通过把整个可行解的集合分成越来越小的子集完成的。排除(剪枝)的工作是通过界定子集中的最好的解"好"的程度,然后舍弃这样的子集——其边界值表明它不可能包含原问题的最优解来完成。

现在我们依次介绍这3个基本步骤——分支、定界、剪枝,并通过应用分支定界算法的一个

原型例子来解释这些步骤。例子(加州制造公司)出现在12.1节,现在复述一下(对约束条件进行编号以便下文引用)。

Max　$Z=9x_1+5x_2+6x_3+4x_4$

s. t.

(1) $6x_1+3x_2+5x_3+2x_4\leqslant10$

(2) $x_3+x_4\leqslant1$

(3) $-x_1+x_3\leqslant0$

(4) $-x_2+x_4\leqslant0$

(5) x_j 是0-1变量,$j=1,2,3,4$

12.6.1　分支

在处理0-1变量时,把可行解集合拆分成子集的最直接的办法是在某个子集中令某个变量(如 x_1)取值为 $x_1=0$,在另一个子集中令此变量为1。在本例中,原问题被拆分成如下两个小的子问题。

子问题1:

令 $x_1=0$ 相应的子问题为

Max　$Z=5x_2+6x_3+4x_4$

s. t.

(1) $3x_2+5x_3+2x_4\leqslant10$

(2) $x_3+x_4\leqslant1$

(3) $x_3\leqslant0$

(4) $-x_2+x_4\leqslant0$

(5) x_j 是0或1,$j=2,3,4$

子问题2:

令 $x_1=1$ 相应的子问题为

Max　$Z=9+5x_2+6x_3+4x_4$

s. t.

(1) $3x_2+5x_3+2x_4\leqslant4$

(2) $x_3+x_4\leqslant1$

(3) $x_3\leqslant1$

(4) $-x_2+x_4\leqslant0$

(5) x_j 是0或1,$j=2,3,4$

图12.4用树图展示了这个分支过程。树图中,从所有节点(对应于包含所有可行解的原问题)分支(弧)生成与两个子问题相对应的两个节点。经过一次次迭代,树会继续长出新枝,我们称这种树为求解树(或枚举树)。在每次迭代中被赋值用以产生分支的变量(如上文的 x_1)称为分支变量(选择分支变量方法很复杂,它是某些分支定界算法重要的研究内容,不过,为了简化问题,我们在本节中只是以变量的自然顺序——x_1、x_2、x_3 进行选择)。

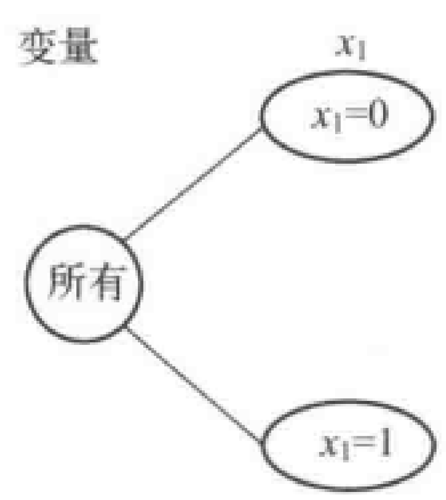

图12.4　对12.1节中的例子应用0-1整数规划的分支定界法进行第1次迭代分立后产生的求解树

在本节以下部分你将看到某些子问题被直接处理(剪

枝)。另一些子问题通过赋值 $x_1=0$ 或 $x_2=1$ 被进一步拆分成更小的子问题。

对于那些整数变量有两个以上取值的整数规划问题,可对其分支变量的每个方面都赋予单个值产生分支,从而建立两个以上的子问题。不过,较好的选择是对分支变量设置取值区间(如 $x_j \leqslant 2$ 或 $x_j \geqslant 3$)产生子问题,12.7 节中描述了这种方法。

12.6.2 定界

对每个子问题,我们需要得到一个边界值,此边界值可表示该子问题是好的可行解的“好”的程度。完成这项工作的标准方式是快速求解该子问题的松弛问题,此松弛问题比该子问题更为简单。在大多数情况下。我们通过删除(放松)某些使问题难以解决的约束条件就能很容易地得到一个问题的松弛问题。对于整数规划问题,最烦人的约束是要求变量为整数。因此,最广泛采用的松弛方法是删除此类约束的线性松弛法。

下面举个例子。考查 12.1 节中绘出并在本节一开始重复的问题。它的松弛问题是把模型的最后一行($x_j=0$ 或 $1, j=1,2,3,4$)换成如下新的约束:

(5) $0 \leqslant x_j \leqslant 1, j=1,2,3,4$

利用单纯形法,我们就能很快得出线性松弛问题的最优解为

$$(x_1,x_2,x_3,x_4)=\left(\frac{5}{6},1,0,1\right), Z=16\frac{1}{2}$$

因此,对原 0-1 整数规划问题的所有可行解都有 $Z \leqslant 16\frac{1}{2}$(因为这些解是线性松弛问题可行解的子集)。事实上,正如下文总结所述,$16\frac{1}{2}$这个上限可以下调至 16。因为目标函数的所有系数都是整数,所有的整数解必定产生整数值 Z。

原问题的边界值 $Z \leqslant 16$。

我们以同样的方式求得两个子问题的边界值。对于子问题 1,固定变量 $x_1=0$,这很容易在线性规划松弛式中通过加入约束 $x_1 \leqslant 0$ 得到。因为将其同现有约束 $0 \leqslant x_1 \leqslant 1$ 结合,得到 $x_1=0$。类似地,对于子问题 2,当固定变量 $x_1=1$,即在线性规划松弛式中加入约束 $x_1 \geqslant 1$。应用单纯形法可以求得对这些线性规划松弛问题的最优解,如下所示。

子问题 1 的松弛问题:$x_1 \leqslant 0$ 和(5)$0 \leqslant x_j \leqslant 1, j=2,3,4$

最优解:$(x_1,x_2,x_3,x_4)=(0,1,0,1), Z=9$

子问题 2 的线性松弛问题:$x_1=1$ 和(5)$0 \leqslant x_j \leqslant 1, j=2,3,4$

最优解:$(x_1,x_2,x_3,x_4)=\left(1,\frac{4}{5},0,\frac{4}{5}\right), Z=16\frac{1}{5}$

子问题 1 的边界值:$Z \leqslant 9$

子问题 2 的边界值:$Z \leqslant 16$

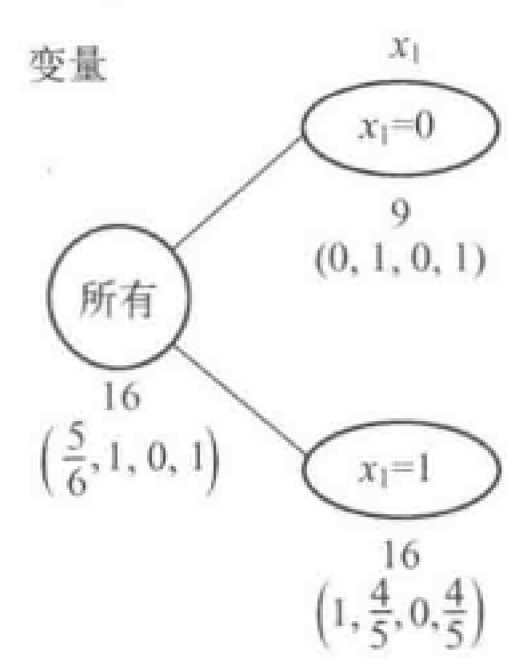

图 12.5 对 12.1 节中的例子应用 0-1 整数规划的分支定界法进行第 1 次迭代分支后产生的边界值

图 12.5 汇总了这些结果,节点下方的数字即为边界值,而边界值下方则为线性松弛问题的最优解。

12.6.3 剪枝

子问题可采用以下 3 种方式处理(剪枝),以避免进一步的考查。

第一种方式可通过图 12.5 中 $x_1=0$ 节点处子问题 1 的结果看出。注意:它的线性松弛问题的(唯一)最优解 $(x_1,x_2,x_3,x_4)=(0,1,0,1)$ 是整数解。因此该解也必定是子问题 1 的最优解。