

## 5. 逢山开路(CUMCM 1994 年 A 题)

要在一山区修建公路,首先测得一地点的高程,数据见表 1(平面区域  $0 \leq x \leq 5600$ ,  $0 \leq y \leq 4800$ ,表中数据为坐标点的高程,单位:m).数据显示:在  $y = 3200$  处有一东西走向的山峰;从坐标(2400,2400)到(4800,0)有一西北—东南走向的山谷;在(2000,2800)附近有一山口湖,其最高水位略高于 1350 m,雨季在山谷中形成一溪流.经调查知,雨量最大时溪流水面宽度  $W$  与(溪流最深处的) $x$  坐标的关系可近似表示为

$$W(x) = \left( \frac{x - 2400}{2} \right)^{3/4} + 5, \quad 2400 \leq x \leq 4000.$$

公路从山脚(0,800)处开始,经居民点(4000,2000)至矿区(2000,4000).已知路段工程成本及对路段坡度  $\alpha$ (上升高程与水平距离之比)的限制如表 2.

1) 试给出一种线路设计方案,包括原理、方法及比较精确的线路位置(含桥梁、隧道),并估算该方案的总成本.

2) 如果居民点改为  $3600 \leq x \leq 4000$ ,  $2000 \leq y \leq 2400$  的居民区,公路只须经过居民区即可,那么你的方案有什么改变.

表一 ↑北

4800	1350	1370	1390	1400	1410	960	940	880	800	690	570	430	290	210	150
4400	1370	1390	1410	1430	1440	1140	1110	1050	950	820	690	540	380	300	210
4000	1380	1410	1430	1450	1470	1320	1280	1200	1080	940	780	620	460	370	350
3600	1420	1430	1450	1480	1500	1550	1510	1430	1300	1200	980	850	750	550	500
3200	1430	1450	1460	1500	1550	1600	1550	1600	1600	1600	1550	1500	1500	1550	1550
2800	950	1190	1370	1500	1200	1100	1550	1600	1550	1380	1070	900	1050	1150	1200
2400	910	1090	1270	1500	1200	1100	1350	1450	1200	1150	1010	880	1000	1050	1100
2000	880	1060	1230	1390	1500	1500	1400	900	1100	1060	950	870	900	930	950
1600	830	980	1180	1320	1450	1420	1400	1300	700	900	850	840	380	780	750
1200	740	880	1080	1130	1250	1280	1230	1040	900	500	700	780	750	650	550
800	650	760	880	970	1020	1050	1020	830	800	700	300	500	550	480	350
400	510	620	730	800	850	870	850	780	720	650	500	200	300	350	320
0	370	470	550	600	670	690	670	620	580	450	400	300	100	150	250
Y/X	0	400	800	1200	1600	2000	2400	2800	3200	3600	4000	4400	4800	5200	5600

表二

工程种类	一般路段	桥梁	隧道
工程成本(元/m)	300	2000	1500(长度 $\leq 300$ m);3000(长度 $> 300$ m)
对坡度 $\alpha$ 的限制	$\alpha < 0.125$	$\alpha = 0$	$\alpha < 0.100$

(西安电子科技大学 何大可提供)

# 逢山开路模型

清华大学 陈连飞 林 涛 兰 嵩

指导教师 高策理

**编者按** 本答案对于地貌、路线、裕量和环境等因素所作的假设基本合理,在确定了几个控制点(桥头,隧道口)后,对一般路线提出了两个模型:利用等高线决定最大坡度路线和在小方格内进行局部优化.本答案特别在居民点东面设一分叉点,分别通往桥西头和隧道南口,这在工程设计上有所创新.对一般路线虽然是局部寻优,但总成本是 370 万元,较合理.

**摘 要** 本文讨论的是在山区修建公路的路线选择问题.

1. 针对问题,本文阐述了局部优化的原理,并根据这个原理提出了对山区的具体情形设置控制点的方法.这种设置控制点局部优化的方法对一般修路问题都适用.
2. 结合具体实例,详细说明具体的设置控制点局部优化方法.
3. 利用计算机进行计算并配以图形  $\leq \alpha$  处理显示,说明这种方法.
4. 对其他方法进行一些比较和讨论,考虑各种方法对本类问题的适用性.

## 问题的提出

在某山区修建公路.已知该山区的地形高度分布.该山区里有一东西走向的山峰,以及一山口湖.雨季时山口湖溢出在山谷形成溪流,最大水面宽度与溪流最深处的  $x$  坐标近似满足

$$W(x) = \left( \frac{x - 2400}{2} \right)^{3/4} + 5 \quad (2400 \leq x \leq 4000).$$

现在设计一条线路,从  $(0, 800)$  出发,经居民点  $(4000, 2000)$  到矿区  $(2000, 4000)$ .路段工程的单位成本和路段坡度的限制已知.

(1) 提出一个较经济的方案,包括原理、方法、线路位置和工程成本.

(2) 居民点改为  $3600 \leq x \leq 4000, 2000 \leq y \leq 2400$  居民区,而公路只需经过居民点即可,那么设计方案做何变动?

工程种类	一般路段	桥梁	隧道
	300	2000	1500(长度 $\leq 300$ m)
工程成本(元 / m)			3000(长度 $> 300$ m)
对坡度 $\alpha$ 的限制	$\alpha < 0.125$	$\alpha = 0$	$\alpha < 0.100$

## 符号系统

$A \rightarrow B$ :表示从  $A$  到  $B$  的公路

$R(A \rightarrow B)$ :所有从  $A$  到  $B$  的可行路线的集合

$\text{cost}(A \rightarrow B)$ :建造公路  $A \rightarrow B$  的成本

$\text{length}(A \rightarrow B)$ : 公路  $A \rightarrow B$  的长度

$J = \{T_i\}_{i=1}^4$ :  $T_i$  为公路类型

$T_1$  = 一般公路,  $T_2$  = 桥梁,  $T_3$  = 长度小于等于 300 m 的隧道

$T_4$  = 长度大于 300 m 的隧道

$a(T_i)$ : 对  $T_i$  的倾斜度限制

$P(T_i)$ : 对  $T_i$  的单位距离工程成本

### 问题分析

设  $\{A_i\}_{i=0}^N$  是公路  $A_0 \rightarrow A_N$  上  $N+1$  个点的集合, 定义  $T(A_k \rightarrow A_{k+1})$  为公路  $A_k \rightarrow A_{k+1}$  的类型.

上式成立的条件是  $A_{k_j} \rightarrow A_{k_j+1}$  和  $A_{k_i} \rightarrow A_{k_i+1}$  没有重叠部分, 其中:  $k_j \neq k_i$ . 在此条件下:

$$\begin{aligned} & \min_{A_0 \rightarrow A_N \in R(A_0 \rightarrow A_N)} \{ \text{cost}(A_0 \rightarrow A_N) \} \\ &= \min_{A_0 \rightarrow A_N \in R(A_0 \rightarrow A_N)} \left\{ \sum_{k=0}^{N-1} \text{cost}(A_k \rightarrow A_{k+1}) \right\} \\ &= \sum_{k=0}^{N-1} \min_{A_0 \rightarrow A_N \in R(A_0 \rightarrow A_N)} \{ \text{cost}(A_k \rightarrow A_{k+1}) \} \\ &= \sum_{k=0}^{N-1} \min_{A_k \rightarrow A_{k+1} \in R(A_k \rightarrow A_{k+1})} \{ \text{length}(A_k \rightarrow A_{k+1}) \}. \end{aligned}$$

由上式, 我们可以进行分段讨论, 我们可在整条路上取几个控制点  $\{A_i\}_{i=0}^N$ . 设定  $T(k) = T(A_k \rightarrow A_{k+1})$ , 与  $A_k \rightarrow A_{k+1}$  的具体线路无关. 控制点  $A_i$  的选择是根据具体情况分析得到的.

对本题, 考虑到桥梁和隧道成本远高于一般公路, 因此在线路设计时往往为节省成本而让一般公路多绕几个弯.

利用 MATLAB 软件, 我们将数据画成等势图和三维视图, 可以看出:

1. 从起点到居民点须经过一条小谷, 雨季形成溪流.
2. 从居民点到矿区, 有一山峰阻挡, 且高度很高, 坡度很陡.

因此, 我们设置控制点:

$A_0$ : 起始点 (0, 800)

$A_1$ : 狭谷西岸 (待定)

$A_2$ : 狭谷东岸 (待定)

$A_3$ : 居民点 (4000, 2000)

$A_4$ : 山峰南侧某处 (待定)

$A_5$ : 山峰北侧某处 (待定)

$A_6$ : 矿区 (2000, 4000)

$A_1 \rightarrow A_2$ :  $T(1)$  = 桥梁

$A_4 \rightarrow A_5$ :  $T(4)$  = 隧道

其余道路为一般公路.

对第二个问题,  $A_3$  改成待定, 在范围  $3600 \leq r \leq 4000$  内, 且  $2000 \leq y \leq 2400$ .

### 模型假设

1. 地貌假设: 假设题目提供的数据是精确和充分的. 在四个相邻数据点间的单位矩形内没有太大起伏. 如果两条对角线两端数据的均值的差远小于矩形边长, 则可以近似认为该矩形为平面.

2. 路线假设:

- a. 只考虑公路为一几何线而不计其宽度, 忽略横向坡度对宽度的限制.
- b. 设计路线时, 暂不考虑路线急转弯角度、缓急的限制.
- c. 不考虑地质情况及气候条件等的影响.

3. 裕量假设:

- a. 不妨假设溪流宽度已经留有桥梁高于水面的余地.
- b. 坡度的上限也包含裕量.

4. 环境假设:

- a. 假设该地区内无原公路可利用.
- b. 新修公路应限于所给区域内.

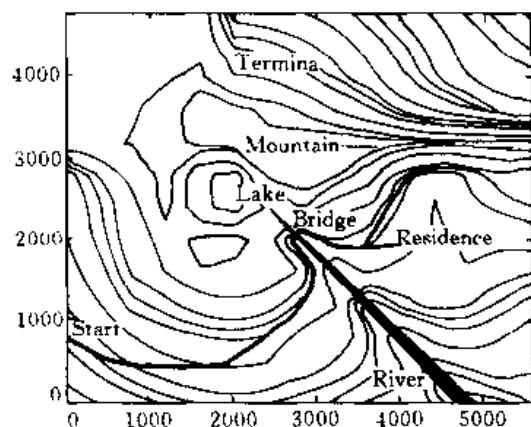


图1 模型一

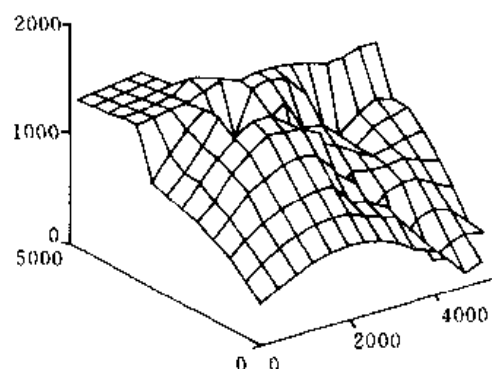


图2 地貌图

### 模型设计与结果

#### 一、模型一

这一模型比较简单, 只用于确定两个控制点间的最短路径.

从等势图上可以看出, 两点间的坡度为高度差除以两点间的水平距离. 设起始点高  $h_0$ , 等高线间高度差为  $\Delta h$ . 若以起始点为中心,  $\frac{\Delta h}{\alpha}$  为半径作圆弧, 交  $h_0 + \Delta h$  和  $h_0 - \Delta h$  对应的等高线于  $A, B$  两点, 则这两点间圆弧上各点的高度  $h'$  满足  $h_0 + \Delta h \geq h' \geq h_0 - \Delta h$ , 则从起始点到这些点的坡度小于  $\alpha_0$ . 可从这些点中选取某点为第一步的终点. 例如想尽量爬高时选  $h_0 + \Delta h$  的点.

如图1所示, 为本问题的初步路线.

这一方案直观、简便. 缺点首先是不精确, 只能得到大概的路线. 其次在坡度变化大的地

方需要等势线足够密(即  $\Delta h$  足够小)才能实现.

## 二、模型二

### 1. 道路的选择(一般公路)

两个控制点之间,如何选择合适道路使长度短而坡度满足要求,是决定成本的重要条件,也就是怎样实现区域优化,我们选择了逐步定线的方法.原理与模型一有相似之处.

先将  $A_i \rightarrow A_j$  分  $a_{i_0} \rightarrow a_{i_1} \rightarrow a_{i_2} \rightarrow \cdots$  成若干段,  $a_{i_n}$  是位于某单元矩形的边上的点,  $a_{i_n} \rightarrow a_{i_{n+1}}$  同是直线段,坡度小于或等于  $a$  的最大值且方向尽可能地靠向  $a_{i_n} \rightarrow A_j$  的直线方向.

如图 3 所示,设  $a_{i_n}$  位于矩形 1 和矩形 2 的棱上,除  $a_{i_n}$  所在棱外还有 6 条棱,记为  $l_1 \sim l_6$ .不妨设  $a_{i_n}$  与  $A_j$  的连线交  $l_1 \sim l_6$  中某棱于  $P_0$  点.可由假设 1 认为  $P_3, P_4$  间是直线.由此求出  $P_0$  点的高度,从而求出  $a_{i_n} \rightarrow P_0$  的坡度,设为  $a_0$ .

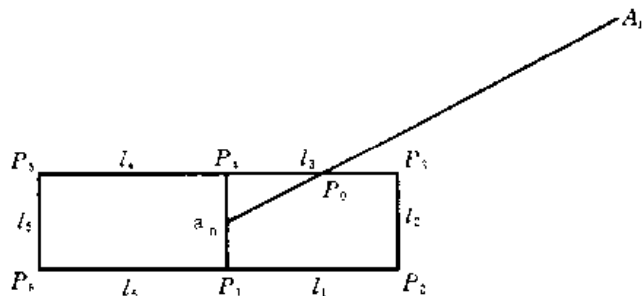


图 3

当  $|a_0| < 0.125$  时,取  $a_{i_{n+1}}$  为  $P_0$ .

当  $|a_0| > 0.125$  时,再在  $P_0$  附近棱上找  $P'$ ,使  $a_{i_n}$  对  $P'$  的坡度正好为 0.125,且  $a_{i_n} \rightarrow P'$  与  $a_{i_n} \rightarrow A_j$  的角度偏差最小.设  $(x_1, y_1), (x_2, y_2)$  为靠近  $a_{i_n} \rightarrow A_j$  的一条棱,且这两点中对应  $a_{i_n}$  的坡度一个大于 0.125,一个小于 0.125,那么可以知道该棱上必可找到坡度等于 0.125 的点.设该点为  $(x_0, y_0)$ .

不妨设  $x_0 = x_1 = x_2$ ,

$y_0$  由方程: 
$$\frac{(y_0 - y_2)z_1 + (y_1 - y_0)z_2 - z}{[(x_0 - x_1)^2 + (y_0 - y_1)^2]} = \gamma \times 0.125$$
 决定,

$z_1, z_2$  是  $(x_1, y_1), (x_2, y_2)$  点的高度,  $z$  是点  $a_{i_n}$  的高度.

$$\gamma = \begin{cases} 1 & a > 0.125 \\ -1 & a < -0.125 \end{cases}, (\text{主要考虑上下坡问题})$$

易知这样的解是存在的,利用这个解求出  $P'$  即是  $a_{i_{n+1}}$ ,依次进行下去逐步接近  $A_j$ .一般情况下,这样的路线是较优的.

### 2. 山谷溪流的处理和桥梁

从图中和表中可以看出,谷底是直线的,谷的两侧基本也是对称的.可以由此算出雨量最大的液面界限:

谷底方程:  $x + y = 4800 (2400 \leq x \leq 4800)$

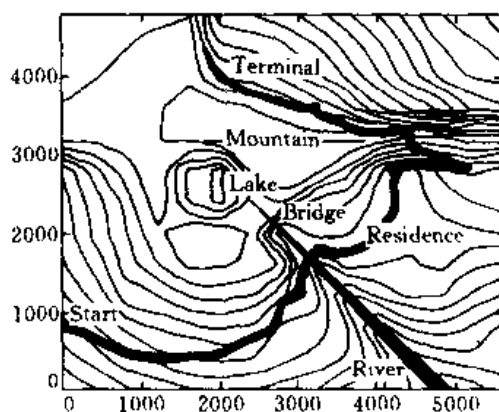


图4 模型二之方案一

$$\text{西岸方程: } 4800 - x - y = \frac{\sqrt{2}}{2} w\left(\frac{x - y + 4800}{2}\right) \quad (2400 \leq x \leq 4800)$$

$$\text{东岸方程: } x + y - 4800 = \frac{\sqrt{2}}{2} w\left(\frac{x - y + 4800}{2}\right) \quad (2400 \leq x \leq 4800)$$

$$\text{其中: } w(x) = \left(\frac{4800 - x}{2}\right)^{3/4} + 5$$

由于桥的成本与一般公路相比较, 所以建桥最好使桥正好跨在两岸的界限上, 同时要顾及桥的两端高度相等.

### 3. 隧道位置的选择:

由于隧道的成本很高, 且长度超过 300 m 后成本增长一倍, 因此最好能选择长度小于 300 米的隧道. 从等势图上可以看出,  $x = 4400$  处的山峰特别尖, 在这里修隧道可能实现这一点. 因此以下计算中固定隧道位置的横坐标  $x = 4400$ .

想让隧道的长度小于 300 m, 需将路修到一定高度. 如图 5 所示, 对水平隧道

$$Z_c = 1500 - \frac{300 \times 650}{400 \times \left(1 + \frac{650}{600}\right)} = 1266$$

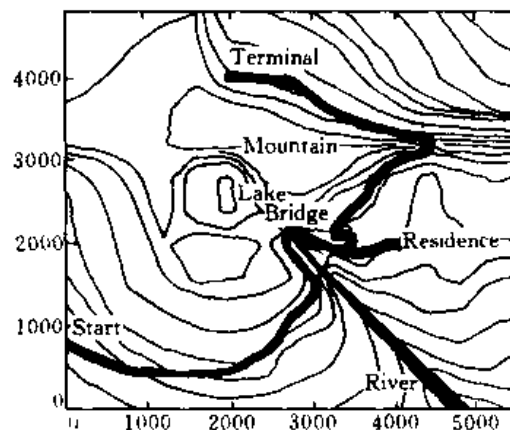


图5 模型三

即使有坡度,隧道高度也不能低于 1240 m.对于矿区,高度 1320 m 可以实现这一点.对于居民区,高度 950 m,上升到这个高度相当困难.本模型采取了迂回绕道的方法达到爬高的目的.显然这样做省下的隧道成本远大于多修公路的成本.

#### 4. 模型的实现

对本模型我们使用了 FORTRAN 语言编程实现.通过上机算出两种方案.

方案一:路径总长度 10.685 km,总费用 3704.6 千元.

方案二:路径总长度 10.802 km,总费用 4266.0 千元.

(除模型三中举出一例,其它模型结果具体路径见附录)

#### 5. 模型的不足

1. 对地貌假设的范围.由地貌假设,每个单元矩形近似平面,这在山两侧不成立,会影响桥的位置.

2. 采用迂回绕道的方法虽然最后隧道符合要求,但拐弯过于尖锐,实际需改进.

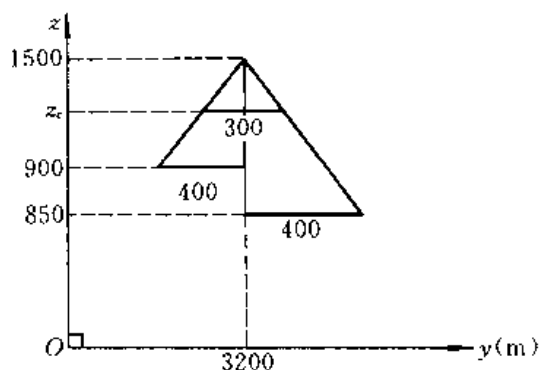


图 6

### 三、模型三

模型三是在模型二的基础上改进而成.

1. 模型三仍采用局部优化,逐步定线的方法.

2. 在河谷和山脊附近的地貌假设不合理,如  $x \in [2800, 3200]$ ,  $y \in [1200, 1600]$  这个区域四个顶点高度分别为 1300, 700, 1040, 900, 则  $(1300 + 900)/2 - (1040 + 700)/2 = 230$ , 对这么大的差值仍将该矩形视为平面,显然不太合理.

考虑到实际情况,可以将这类矩形近似成两个三角平面,其公共边是地表的一条凸出或者下陷的线.这要通过三维视图判断哪条对角线是公共边.

根据这一假设,公路到了岸边的高度后,还要经过下降的坡才能到桥边,因此公路过桥前还要沿河边走一段路才行.这与模型二是不一样的.我们在处理山脊附近的点的情况时,也采用了这种假设.这样得到的数据点更多,更精确.

3. 从居民点到隧道的爬高过程仍是一个难题.考虑到按上一假设计算,桥东头到居民点的路上的有些点也很高,我们可以这样处理:假如公路允许分叉的话,我们可以在桥东侧

到居民点某处再取一个控制点  $A'$ ,将原来的  $A_2 \rightarrow A_3 \rightarrow A_4$  变成了  $A_2 \rightarrow A' \rightarrow A_3 \rightarrow A_4$ ,这样一

方面减低公路总长,另一方面可以降低迂回绕道的次数和角度.

4. 为了得到精确的结果,我们采用 Mathematica 软件逐点计算,应用该模型解决问题一,结果如下:

```

x = {0 400 800 1200 1600 2000 2400 2479 2788 2800 3017.7
2925.4 2800 2677.6 2677.6 2711 2800 3100 3200 3325 3600 4000
3600 3325 3527.4 3200 3457.2 3600 3786.8 4000 4400 4400 4000
3600 3200 3138.5 2800 2400 2000)
y = {800 639.6 474.6 419.8 419.8 486.9 666 800 1200 1212 1382
1600 1819.7 2000 2089 2122.5 2109.1 2000 1950 1875 1857.7 2000
1875.7 1875 2000 2033.6 2400 2465.2 2800 2966.6 3080.8 3296.7 3345.4
3457.2 3561.2 3600 3930.4 3965.2 4000)

```

其中  $x$  向量是所得路径的横坐标,相应的  $y$  向量是所得路径的纵坐标.

路径的总长度:10.72273 km

总费用:3557.4941 千元

桥梁位置(2677.56,2089.033)——(2710.96,2122.45)

桥长:47.24 m

桥费用:94489 元

隧道位置:(4400,3080)——(4400,3296.74)

隧道长度:217.12 m

隧道造价:325525 元

模型三的结果分析:

由于三角平面的近似,使公路在桥边必须缓慢下坡,这样使得桥的位置上移,长度减小.可以看出模型三的结果比模型二的更合理更实用.

四、用模型三解决问题二

对问题二,相当于把控制点  $A_3$  改为可变,其位置由(4000,2000)改为  $3600 \leq x \leq 4000$ ,  $2000 \leq y \leq 2400$ .我们在区域里取了几个点进行计算.结果为:

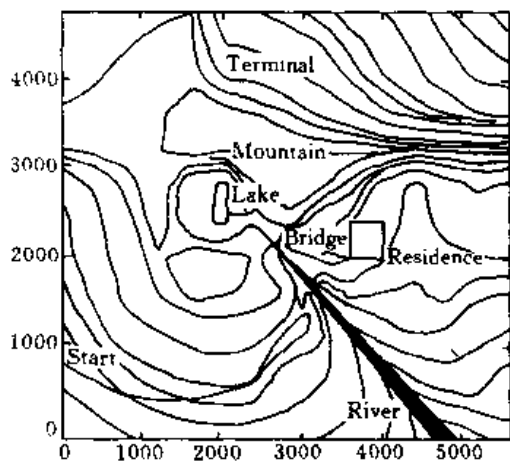


图7 问题二之方案一

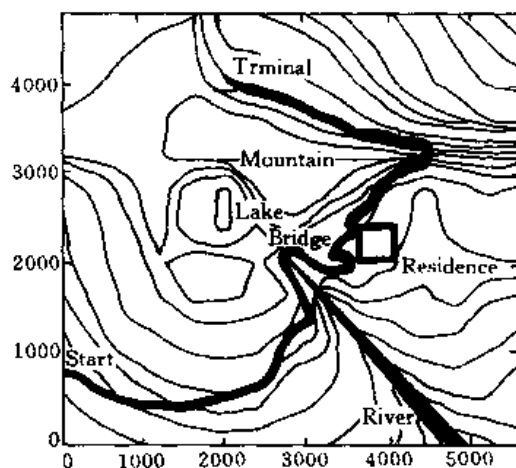


图8 问题二之方案二



方案一 总路径长 10.1315 km,总费用 338.024 万元.

方案二 总路径长 10.0995 km,总费用 337.054 万元.

从结果上看,这两个点的选取确实降低了造价.而就这两点比较,造价很接近.

## 模型的设计与分析

### 一、桥址的选择对总成本的影响

桥址的选择即  $A_1, A_2$  的选择.用模型二试计算:

第一种情况:  $A_1 = (3080.79, 1651.40), A_2 = (3133.72, 1782.60)$

$$\text{length}(A_1 \rightarrow A_2) = 95.93 \text{ m}$$

$$\text{cost} = 3704.6 \text{ 千元}$$

第二种情况:  $A_1 = (3092.09, 1653.59), A_2 = (3138.58, 1444.09)$

$$\text{length}(A_1 \rightarrow A_2) = 405.4 \text{ m}$$

$$\text{cost} = 4266 \text{ 千元}$$

可见桥址的选择与造价有很大关系.应尽量使桥短些.

### 二、居民点的变化对总成本的影响

对模型三,令居民点(4000,2000),  $\text{cost} = 3557.49$  万元

居民点(3600,2000),  $\text{cost} = 3380.24$  万元

居民点(3600,2400),  $\text{cost} = 3370.54$  万元

对后面两种情况,  $\Delta y = 400, \Delta y/y = 20\%$ , 而  $\frac{\Delta \text{cost}}{\text{cost}} = 0.3\%$ , 可知居民点的变化对总成本的影响不大.

## 讨论及优缺点

### 1. 与连续模型的比较

对地面起伏不大,数据充足情况下,可以分区拟合地面的函数.设为  $f(x, y, z) = 0$ . 那么可用变分法求最佳路径:

$$\begin{cases} \min \int_0^1 |x'(t)| dt \\ x(0) = A \\ x(1) = B \\ f(x, y, z) = 0 \end{cases}$$

不过,对地面起伏不大的地区,我们的方法也适用,且结果与变分法的相差不会大.而对于地面不规则的地区,如本问题,则用变分法是求不到解的.

2. 三个模型的精确程度为:模型一较差,模型二较好,模型三最好.而模型二和三的计算量大.

3. 本模型的优点在于灵活性好,得到的结果接近实际.通过灵敏度分析可以知道哪些点的选择要求灵敏度高,需要进一步测量.

4. 我们无法证明我们得到的方案最佳,另外在选择路线时只考虑了成本,没有考虑使用是否方便及长期经济效益.

5. 改进方向:可以加入其他条件的限制,使得到的方案更加实用.

# 关于“逢山开路”赛题的分析和答卷评阅

中国科学院应用数学研究所

韩继业

本题的构思保留了工程实际背景的一些基本特征,涉及到地貌、路线、环境等自然条件以及费用系数,这些在实际的工程设计上必须注意的重要因素我们在解决本题时也应考虑到.在用数学模型方法解题时,除了从数学角度上思考问题的求解之外,适当地考虑有关实际因素,对于我们建立合理的数学模型提供了重要的依据条件,也会使我们设计的解题方法比较可行和有效.本题在这一方面表现得更加明显.

本题的重要步骤是求两点间的最短路.由于要在河流上架桥及开挖隧道,直接求从山脚  $S$  到居民点  $R$  再到矿区  $M$  的最小费用的路线是困难的.一种简便方法是先根据对地形和不同路段费用系数的分析,确定桥头和隧道口的若干候选地点,然后寻求从  $S$  到桥西头  $B_1$ ,从桥东头  $B_2$  到  $R$ ,从  $R$  到隧道南口  $D_1$  和从隧道北口  $D_2$  到  $M$  的最短路,也是最小费用路径,其中  $B_1, B_2$  和  $D_1, D_2$  均有若干候选点.最后再综合考虑修桥和隧道的费用,从候选路径得到全局最小费用路线.

下面介绍求两点(如  $S$  和  $B_1$ )间最短路的一种方法.

图 1 表示一个图  $G=(V, E)$ ,  $V$  是顶点  $V_1, \dots, V_9$  的集合,  $E$  是连接顶点的弧的集合,弧旁的数字不妨看作二顶点间的距离.求图中任意两点间最短路的一种有效算法是 Dijkstra 提出的所谓标号法,用它可以同时求出从图中某一点(如  $V_1$ )到其他各点的最短路,因而适合于本题中求  $S$  到有若干候选点的  $B_1$  的最短路问题.

这个方法的基本思想是从起点  $V_1$  开始,逐步地寻找到达各点的最短路,在每一步都对每一顶点记录一个数,称为该点的标号.它表示  $V_1$  到该点的最短距离的上界(称  $T$  标号),或者就是  $V_1$  到该点的最短距离(称  $P$  标号).实际上每一步都通过比较把至少一个具有  $T$  标号的点变为  $P$  标号,这样最多经过  $9-1=8$  步就可以找到  $V_1$  到各点的最短路.

将 Dijkstra 方法用于求本题的最短路时,必须确定图的顶点集  $V$  和弧集  $E$ .这可以在原来网格的基础上构造,现给出两种形式.

1. 对原网格加密.如将  $400\text{ m} \times 400\text{ m}$  加密为  $100\text{ m} \times 100\text{ m}$ ,新格点构成图的顶点集  $V$ ,如图 2 所示(新格用虚线表示),其高度( $z$  坐标)由原格点插值得到.图的弧集  $E$  可以只包括  $x$  方向和  $y$  方向的路径,这时相邻两顶点  $V_i, V_j$  间的距离,即弧长  $d_{ij}$  定义为(且考虑到坡度限制)

$$d_{ij} = \begin{cases} [(\Delta x)^2 + (\Delta z)^2]^{1/2}, & \left| \frac{\Delta z}{\Delta x} \right| \leq 0.125, \\ \infty, & \left| \frac{\Delta z}{\Delta x} \right| > 0.125, \end{cases} \quad (1)$$

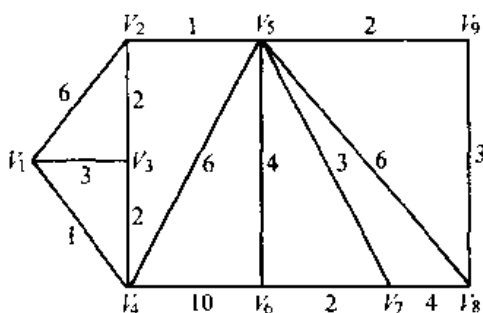


图 1

其中  $\Delta z$  是  $V_i$  与  $V_j$  的  $z$  坐标之差,  $\Delta x = 100$  m, 且可视为  $\Delta y = 100$  m. 弧集  $E$  也可以包括对角线方向的连线, 如图中左上角的格子所示. 这时  $d_{ij}$  应定义为

$$d_{ij} = \begin{cases} [(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2]^{1/2}, & \frac{|\Delta z|}{((\Delta x)^2 + (\Delta y)^2)^{1/2}} \leq 0.125, \\ \infty, & \frac{|\Delta z|}{((\Delta x)^2 + (\Delta y)^2)^{1/2}} > 0.125 \end{cases} \quad (2)$$

2. 对原网格只作  $200 \text{ m} \times 200 \text{ m}$  的加密, 然后在新网格的边上细分. 如分为  $50 \text{ m}$  一个点, 构成顶点集  $V$ . 而弧集  $E$  由某顶点出发到它相邻的 2 个 (或 4 个) 新方格内任一点的连线构成, 如图 3 所示, 这时  $d_{ij}$  的定义同 (2) 式, 图中折线表示一条从  $V_1$  到  $V_2$  的路径.

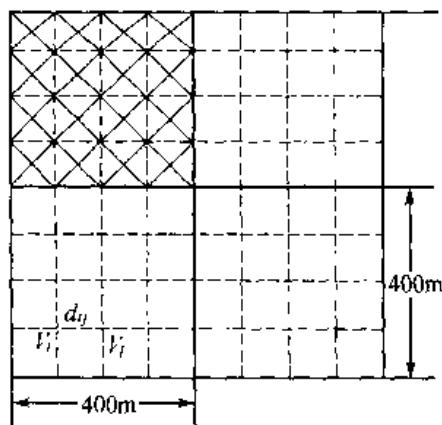


图 2

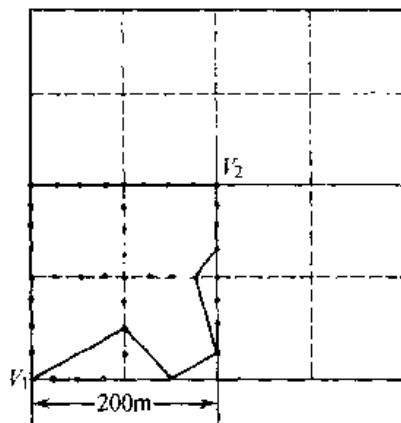


图 3

顶点集  $V$  和弧集  $E$  一经确定, 就可以直接用 Dijkstra 方法求图上任一点到其他各点的最短路了.

这次数模竞赛中本题的绝大多数答案都利用了平面上格点的加密, 以及地面高度的插值. 插值的方法多为线性插值, 少数答案为了使结果更精确, 还利用了局部二次或三次插值. 大多数答案也都先根据地形和不同路段的费用系数的分析和计算, 提出了桥和隧道的几个可供选择的地点, 然后再对整个路线进行优化考虑, 最后确定出各自的最优路线. 但大多数答案中确定最优路线的数学方法不够先进, 只有极少数答案提出了一般路线的优化模型, 或叙述了动态规划方法, 但又缺乏动态规划方法计算过程的说明与分析. 多数答案对整个路线的优化方法主要是通过对几种不同方案的比较来得到; 部分答案在确定局部路线时基本上是用图上作业方法, 即在等高线图上沿着允许坡度的路线来确定, 这样得到的路线的总费用往往偏高, 但是多数答案的总费用仍属于可以接受的范围. 某些答案确定的路线的总费用过高, 其原因多为隧道和桥的位置选择得不合理, 致使过长. 例如桥长超过了  $100 \text{ m}$ , 隧道长超过了  $500 \text{ m}$ , 这使得整个路线的总费用超过了  $400$  万元. 多数答案在建立数学模型时提出的假设条件是比较合理的, 分析问题也比较细致, 但在确定出最优路线后, 进一步对结果进行稳定性分析的答案很少, 这是一个不足之处.

总起来看, 通过各个参赛队的认真细致的分析和计算, 大多数答案都取得了相当好的成绩, 彼此水平差距并不大. 一、二等奖是从横向比较选择出来的.

# 公路工程中的选线问题

易智华 朱定国 许礼林

(武汉测绘科技大学 430070)

指导教师 胡元明

**编者按** 本答案的假设合理,特别是使用了等高线跟踪插值方法与等高线搜索前进方法,图上作业法,并且比较了几种方案,考虑得比较细致,文字图形清楚,计算结果也令人满意.

**摘 要** 本文采用线性插值手工描绘的方法,对已知部分高程点的山区作出等高线图,并以此图为基础,进行线路设计.本文重点讲叙了等高线“跟踪”插值过程和利用等高线图进行公路设计,提出了几种不同的方案,并且对这些方案进行了比较和讨论,提供一种认为最满意的结果.本文还尝试运用层次分析法对方案进行评估.最后,本文对所涉及到的公路选线作了推广,并总结出线路选线需考虑的一般问题.

## 一、问题的提出

要在一山区修建公路,首先测得一些地点的高程,数据见表 1(平面区域  $0 \leq x \leq 5600$ ,  $0 \leq y \leq 4800$ ,表中数据为坐标点的高程,单位:米).数据显示:在  $y = 3200$  处有一东西走向的山峰;从坐标  $(2400, 2400)$  到  $(4800, 0)$  有一西北——东南走向的山谷;在  $(2000, 2800)$  附近有一山口湖,其最高水位略高于 1350 米,雨季在山谷中形成一溪流.经调查知,雨量最大时溪流水面宽度  $w$  与(溪流最深处的) $x$  坐标关系可近似表示为  $w(x) = \left(\frac{x - 2400}{2}\right)^{\frac{3}{4}} + 5(2400 \leq x \leq 4000)$ .

公路从山脚  $(0, 800)$  处开始,经居民点  $(4000, 2000)$  至矿区  $(2000, 4000)$ .已知路段工程成本及对路段坡度  $\alpha$  (上升高程与水平距离之比)的限制如表 2.

表 1:略.

表 2 工程限制条件及单位费用表

工程种类	一般路段	桥 梁	隧 道
工程成本(元/米)	300	2000	1500(长度 $\leq 300$ 米);3000(长度 $> 300$ 米)
对坡度 $\alpha$ 的限制	$\alpha < 0.125$	$\alpha = 0$	$\alpha < 0.100$

1) 试给出一种路线设计方案,包括原理、方法及比较精确的线路位置(含桥梁,隧道),并估算该方案的总成本.

2) 如果居民点改为  $3600 \leq x \leq 4000, 2000 \leq y \leq 2400$  的居民区,公路只须经过居民区即可,那么你的方案有什么改变.

## 二、建模的假设

1. 所给地区的地形是连续的,不存在断层、悬崖,可用线性插值方法确定直线的走向.

2. 所给地区的工程地质条件都满足工程建设的要求,不存在对工程建设有害的因素,如地震带、溶岩地区等.
3. 由于题目没有给出公路的等级及规格,可以忽略公路的宽度,而只考虑其线形.
4. 桥梁的长度必须大于溪流水面最宽时的宽度.
5. 隧道的建设不考虑通风等次要因素.
6. 尽管填方、挖方在线路上的工程费用中占很大比例,但题目题设没有给出填方、挖方的具体要求,建模时暂不予考虑.

### 三、与问题有关的某些概念及原理

#### 等高线的概念、特征和应用:

1. 概念:等高线是地面上高程相等的相邻各点连线的闭合曲线,也就是水平面与地面相交的曲线.

2. 特征:

- (1) 在同一条等高线上的各点的高程相等;
- (2) 等高线是闭合的曲线;
- (3) 不同高程的等高线不能相交;
- (4) 等高线与分水线(山脊线)、合水线(山谷线)正交;
- (5) 两等高线间的垂直距离称为平距,等高线平距的大小与地面坡度的大小成反比.

3. 等高线在地形图中的应用:

(1) 可以根据等高线确定地面点的高程;

(2) 根据等高线间的平距确定其坡度.其原理为:若已知地形图上等高距为  $h$ ,要确定图上两等高线的倾角或坡度,则量出该两等高线间的平距宽度  $a$ ,按下式计算:

$$\text{坡度 } \alpha = \operatorname{tg} \tau = \frac{h}{a}.$$

如图 I:  $\alpha = \operatorname{tg} \tau \frac{20}{300} = \frac{1}{15}$  (比例尺, 1:10000 单位:  $m$ )

### 四、建模的初期工作

由前述可知,要进行公路选线,必须有一定精度的地形图,但在问题中只提供了一些规则格网点上的高程.因此,我们必须根据这些高程尽可能准确地反映出该地区的地势,以便确定路线的走向和成本的预算.为此,我们建立等高线“跟踪”插值模型,以便“生成”一幅地形图.

#### 等高线“跟踪”插值模型<sup>[2]</sup>

• 数据特点 所给高程数据点位间隔过大,不利于用二次曲面或者其它类似方法拟合,但由于它是以规则格网数据形式给出的,因此有利于构造“跟踪”信息,实现数据插值.

• 操作特点 采用简单的线性插值,简便易行,同时其数据场很有规则,插值步骤具有很强规律性,适于计算机编程计算.

• 具体过程

- (1) 依次计算各等高线和网格边交点的坐标值.

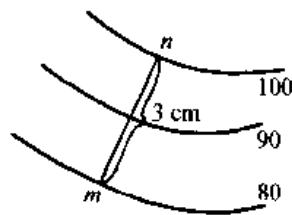


图 1

(2) 找出一条等高线起始等高点并确定判断和识别条件,以追踪一条等值线的全部等高点.

(3) 根据前面提到等高线的特征及地势的大致走向,联结各等高点绘制光滑曲线.

#### • 算法分析

(1) 内插等值点的位置:

假设制图地区由  $m \times n$  个网格点数据组成,并设沿  $x$  方向的分割记为  $j = 1, 2, \dots, n$ ; 沿  $y$  方向的分割记为  $i = 1, 2, \dots, m$ . 则对于任一网格点的高程数据可以表示为  $B_{i,j} = j \cdot d$ ;  $y_{i,j} = i \cdot d$ . 显然,对于  $m \times n$  个网格点区域,只有  $(m-1) \cdot n$  条纵边和  $(n-1) \cdot m$  条横边. 对于位于任一条边上的等高点位置,我们可以表示为:

纵边上的等高点  $H_{i,j} \quad i = 1, 2, \dots, (m-1);$

$j = 1, 2, \dots, n.$

横边上的等高点  $S_{i,j} \quad i = 1, 2, \dots, m;$

$j = 1, 2, \dots, (n-1).$

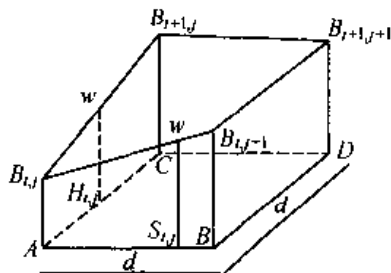


图 2

为了计算等高点在网格边上的位置,首先要确定等值线与网格边相交的条件. 设等高线高程值  $w$ , 显然, 只有  $w$  值处于相邻网格点数值之间. 该边才能有等高点. 因此, 我们可用下式来判断:  $(B_{i,j} - w) \cdot (B_{i,j+1} - w) < 0$  时, 在横边上有等高点;

$(B_{i,j} - w) \cdot (B_{i+1,j} - w) < 0$  时, 在纵边上有等高点.

若上式成立, 我们即可采用线性内插方法计算出等值点位置. 我们有 ABCD 网格 (图 2), 其高程依次为  $B_{i,j}$ 、 $B_{i,j+1}$ 、 $B_{i+1,j}$  和  $B_{i+1,j+1}$ , 若在横边 AB 边内插等高点  $A'$ ,  $A'$  离 A 点的距离记为  $S_{i,j}$ , 则:

$$\frac{w - B_{i,j}}{B_{i,j+1} - B_{i,j}} = \frac{S_{i,j}}{d}, \text{ 令 } d = 1$$

$$S_{i,j} = \frac{w - B_{i,j}}{B_{i,j+1} - B_{i,j}}$$

显然这里计算的  $S_{i,j}$  的一个相对比值, 是横边的一个单位的几分之几, 即  $0 \leq S_{i,j} \leq 1$ . 同理, 对于纵边 AC 之间插等高点  $B'$  时, 该点到 A 的距离记为  $H_{i,j}$ , 当纵边  $d = 1$  时, 有:

$$H_{i,j} = \frac{w - B_{i,j}}{B_{i+1,j} - B_{i,j}}$$

同样  $0 \leq H_{i,j} \leq 1$ .

(2) 追踪等高点

为了确定追踪方案, 我们要确定某一等值线在矩形网格内走向的几种可能, 并通过确定等值线走向与等高点坐标之间的关系来建立跟踪条件. 由于等高点位于网格边上, 所以等高线通过相邻网格的走向只有四种可能: 自下而上、自左向右、自上而下、自右向左. 因此, 若找到某一值线头位于某一网格边上, 该网格边往往是相邻网格的公共边, 即是前一网格的出口边又是后一网格的进入边, 则进入边的方向对于每一个网格都有上、下、左、右四种情形, 即追踪等高点有四种可能. 下面分别讨论这四种可能.

先谈自下而上追踪, 在图 3 中, 我们可以看到, 在方格 I 上有等高点  $a_1$ , 它的位置有三种情况, 即  $H(i, j)$ 、 $S(i, j)$  和  $H(i, j+1)$ . II 号方格上  $a_2$ ; 等高点为  $S(i+1, j)$ , 显然, 我

们比较  $a_1$  和  $a_2$  的坐标位置,可以得到  $a_1$  点取整的纵坐标,一定小于  $a_2$  点取整的纵坐标.因此,只要满足  $i_{a_1} < i_{a_2}$  的条件( $i_{a_1}$  表示  $a_1$  点横坐标),即可自下而上的追踪.如果有  $a_3$  点,它一定位于方格 II 的另外三边上.

同理,可以讨论得到自左而右(图 4),自上而下(图 5)和自右向左(图 6)的追踪情况.

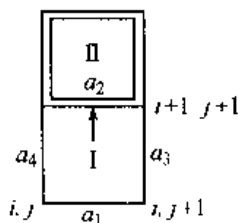


图 3

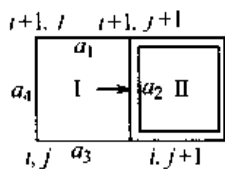


图 4

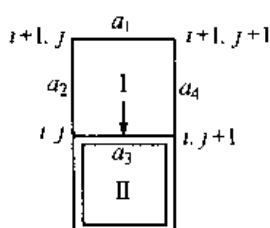


图 5

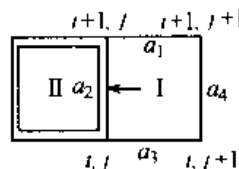


图 6

已知  $a_3$  点是位于 II 号方格的其余三边上,如何确定是其中的哪一边呢?这是一个十分重要的问题,必须合理地选择其余三边上的一个等高点.不然,将会出现同一等高线的交叉和分支走向不确定的多义性.例如某一网格上的四点联结的状况可有三种情况(图 7).a、b 即为多义性 c 是不允许的,必须排除.对于等高线点可以有多种方式联结,手工勾绘等高线时经常出现,这情况的处理往往根据实践经验和对制图对象物理背景的理解作出的,即参考周围等高线的定向和趋势,为强调等高线之间的协调一致,突出表现区域特征而作出各种选择.根据一般的规律,首先是考虑等值线原来前进的方向,即顺着原来等高线走向延伸下去,其次是根据距离远近的选择  $a_3$  点.

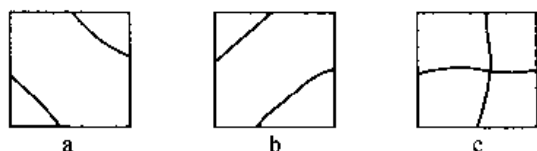


图 7

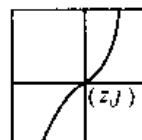


图 8

在内插等高点时,当遇到网格高程值和待高线相等的情况,此时,等高线必须通过网格点.而该网格点同时又是四个相邻网格的公共交点(图 8).于四个相邻横边和纵边上得到不是 0 就是 1 的四个值(即  $S(i,j) = 0, H(i,j) = 0, S(i,j-1) = 1, H(i-1,j) = 1$ ).追踪点时,一定会发生重复和追踪混乱的问题.对此情况,必须预先给予处理,其方法是对该网格点加上一个足够小的数组给予纠正.

根据以上步骤,我们绘制了一幅 1:20000 等高线地形图,作为下一步工作的依据.(如图 15,其中注记了高程格网数据)

此方法采用的是线性插值法,以便实现手工绘图.如果采用二元分段等距抛——抛插值方法或者双三次样条插值方法加密数据场.这样精度要高些.由于时间的关系,我们只编制了二元分段插值的程序,附在后面,以供参考.

## 五、建立公路定线的操作模型

### 1. 建造桥梁、隧道的一般原则:

(1) 中小型桥梁、隧道走向服从道路的走向;

(2) 道路的走向服从大桥梁、大隧道的走向.

2. 对坡度值的要求, 本题要求公路的坡度小于 0.125, 桥梁的坡度为 0, 隧道的坡度小于 0.1.

3. 不考虑“1.”所说的情况, 现在将地形图上进行公路定线的原理[3,4] 及方法叙述如下:

### (1) 展线法

展线的方法是先确定该路段容许的设计平均坡度, 然后, 据此计算相邻两等高线间所应有的长度(宜等于等高线间距与平均坡度之比值), 并按地形图比例尺将其换算成图纸上应表示的已缩小的长度, 以此缩小的长度为准, 自路线起、终点开始分别向上、向下的两等高线之间依次作截距交会于各等高线上成若干平面转折点. 最后将各转点相连成平面上的转折线, 并在有关转角点间布置相应必要的平面线和回头曲线.

### (2) 回头曲线法

当路线起、终点位于同一很陡的山坡面, 为了克服高差过大, 一方面要顺山坡逐步展线; 另一方面又需一次或多次地将路线折回到原来的方向, 形成“之”字线”路线. 这种顺地势反复盘旋面上的展线, 往往会遇到路线平面转折角大于  $90^\circ$ , 按通常设置平面线方法, 曲线长会过短, 纵坡会过大. 为了克服这种困难, 常采用在转角顶点的外侧设置回头曲线的方法来布置路线.

### (1) 展线法和回头曲线法的例子

#### ① 展线法

以图 9 为例, 已知路段起、终点 A、D 之间高差为 18 米, 距离  $L_1$  为 225 米, 其纵坡  $i_1$  为  $18/225$ ; 由于  $i_1$  较大, 决定采取平均坡度为 5% 进行展线. 该处比例尺为 1:5000, 先计算相邻两等高线间为符合坡度 5% 要求的必要长度  $\Delta L = \frac{5 \text{ 米}}{0.05} = 100 \text{ 米}$ , 则在  $\Delta L$  在地形图上经按上述比例尺缩小后的实际长度应为  $100/500$  米, 即 0.02 米. 今自 A、D 点分别以 0.02 米为截距作弧依次交各等高线于  $\Delta_1, \Delta_2, \Delta_3$ . 然后加以连接, 并且在有关平面转折点处布置必要的适合曲线, 即得出图 9 所示的展线方案.

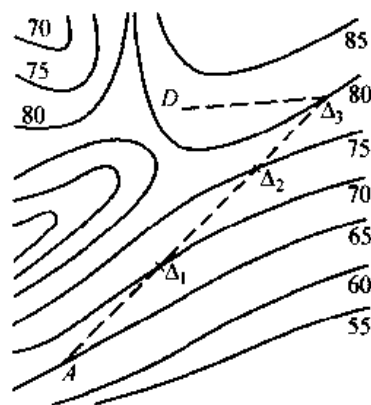
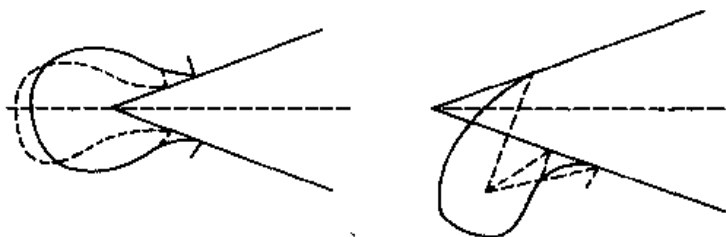


图 9

#### ② 回头曲线法 这里列举几种回头曲线形式:



### 4. 问题的求解

在 A、B 区域(见图 14、15), 如果不进行回头曲线的测设的话, 公路的坡度将超过 0.125, 这就不符合问题的要求; 但如果我们测设回头曲线, 将形成“之”字形的盘山道, 大大



增加公路的里程和营造费用. 所以, 我们在选取路线时, 在满足坡度限制的条件下, 尽可能不测设回头曲线, 也就是说尽量避免在 A 区和 B 区测设公路.

(1) 为了选线的方便, 我们找到了一种比较简便的方法, 来确定公路取向范围. 方法具体如下:

① 由公路必过的定点为圆心, 以  $R$  为搜索半径, 画圆弧, 交相邻的等高线于两点.

② 取其中一点为圆心, 再以搜索半径  $R$  作圆弧, 返回上一步.

$R$ ——搜索半径  $\theta$ ——路线在该点取向范围  $R$  的选取由坡度  $\alpha$  决定(如图 10)

$\Delta D$ ——AB 两点高差

$$\alpha = \frac{\Delta D}{R} \Rightarrow R = \frac{\Delta D}{\alpha}$$

由上式可知, 如果  $\alpha$  越小, 则搜索半径  $R$  越大; 同样, 我们将  $R$  取比较大时,  $\alpha$  将比较小. (如图 11).

由于等高线是通过线性插值的方法获得的, 这就存在着截断误差, 下面分析这个误差对  $R$  的影响.

$\Delta D = h_S - h_A$   $h_S, h_A$  的为通过线性插值得到的高程.

假设截断误差中的误差为  $m_h$ , 则有:

$$dR = \frac{d(\Delta D)}{\alpha},$$

$$dR = \frac{dh_B - dh_A}{\alpha}$$

$$m_R^2 = \frac{1}{\alpha^2} (m_{h_B}^2 + m_{h_A}^2) - 2m_h^2/\alpha^2$$

$$\therefore m_R = \frac{\sqrt{2}}{\alpha} \cdot m_h$$

由此可知, 如果截断误差  $m_h$  很大,  $R$  的中误差  $m_R$  将很大; 若搜索半径仅取  $R = \Delta D$  的话, 仍将可能造成线路的坡度过大. 我们可以通过在实践操作中, 选取较大的  $R' = R + m_R$  来妥善解决, 这里我们取搜索半径  $R = 800 + 10m$ .

由于我们所绘制的等高线是由格网高程插值点得到的, 因此我们对搜索半径  $R$  的传递误差可以不考虑.

## (2) 矿区部分道路拟定原则

根据展线法和回头曲线以高差 100 米、最大坡度 0.125 进行展线, 可大致表示为如图 12 所示.

从绘制的等高线图上我们可以看出, 若公路在 I 区走线, 偏离山峰越来越远, 导致道路走线拉长, 隧道长度增加, 故不可取; 若在 III 区走线, 虽使得隧道长度有所减少, 但会使对面山坡公路走线带来困难. 因为对面山坡坡面较陡, 要满足  $\frac{|H_N - H_S|}{D_{AB}} < 0.1$  (其中:  $H_N$ 、 $H_S$  分别表示隧道北南洞口选址点的高程,  $D_{AB}$  是两点间的水平距离), 必须修建大量的盘山

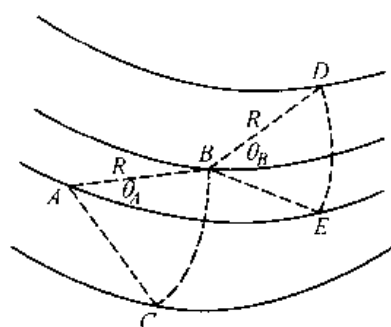


图 10

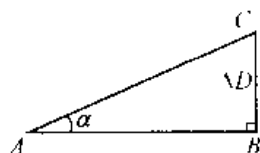


图 11

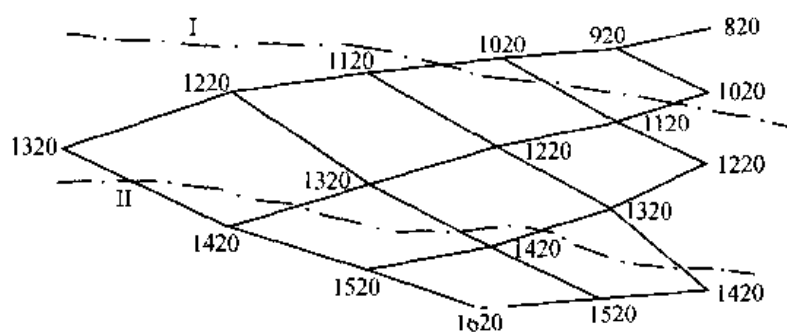


图 12

道,不仅走线拉长,也会给运输带来麻烦.因此我们考虑在 II 区( $h \geq 1120, h \leq 1400$ )设计公路走线.另外,从图上可看出在山峰南坡  $x > 4400$  米,等高线间的平距越来越大,而在其对应的北面坡等高线的变化并不十分明显;在山峰南部  $x < 4000$  米,南北两坡等高线平都有明显增大趋势.考虑到隧道的开凿,再满足其他条件的情况下,使最小长度尽可能小,因此我们在  $x \geq 4000$  米和  $x \leq 4400$  米的范围内进行选址.

### (3) 对于居民点的情况分析

根据以上的方法,我们可以分段进行求值.(见图 13,图 14,图 15).

#### ① 从公路起始点到山谷,可得出三种走线:(见图 20)

$H-I-J-L-P-Q$ —沿 1000 米等高线到桥 1;

$H-I-J-M$ —沿 850m 的等高线到桥 3;

$H-I-J-N$ —沿 900m 的等高线到桥 2.

#### ② 考虑到地形、引桥等多种因素,采用 $S = 1.4w(x)$ 米计算桥长.

#### ③ 在考虑到桥梁处的展图过程中,我们有以下几种走线:

$R-III$ —沿着 1000m 的等高线到桥 1;

$R$ —桥 2;

$R-II-II_2$ —沿 850 的等高线到桥 3.

注:在 800m 等高线上建桥,不仅桥身长,而且走线盘曲多折,拉线长,没有以上三种效果好,故舍去.

#### ④ 从居民点 $R$ 到隧道的几种走线

$R-IV$ —隧道  $B_1A_2$ ;

$R-III-III_1-III_2-III_{21}-III_{22}$ —沿 1300m 的等高线到隧道  $B_2A_2$ ;

$R-III-III_3-III_{31}-III_{32}$ —沿 1300m 的等高线到隧道  $B_2A_2$ ;

$R-III-III_3-III_{34}-III_{35}$ —沿 1300m 的等高线到隧道  $B_2A_2$ .

#### ⑤ 量得隧道长: $A_1B_1 = 720m, A_2B_2 = 210m$ .

#### ⑥ 从矿区 $G$ 到隧道展线所形成的几种走线

$G-A_{21}-B_{21}$ —沿 1120m 等高线到隧道  $A_1B_1$ ;

$G-A_{21}-B_{11}$ —沿 1320m 等高线到隧道  $A_2B_2$ ;

$G$ —沿着 1320m 等高线到隧道  $A_2B_2$ ;

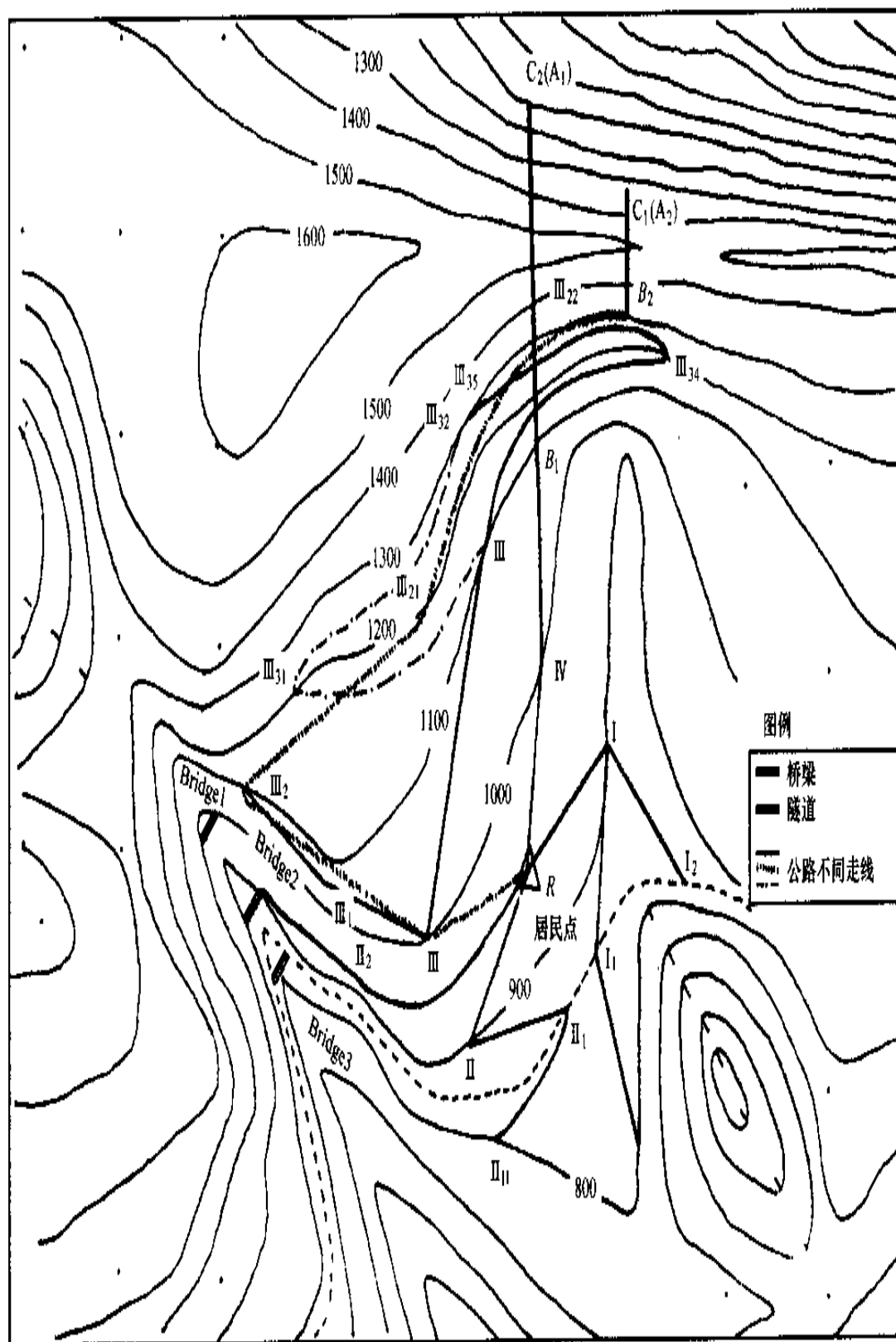


图13

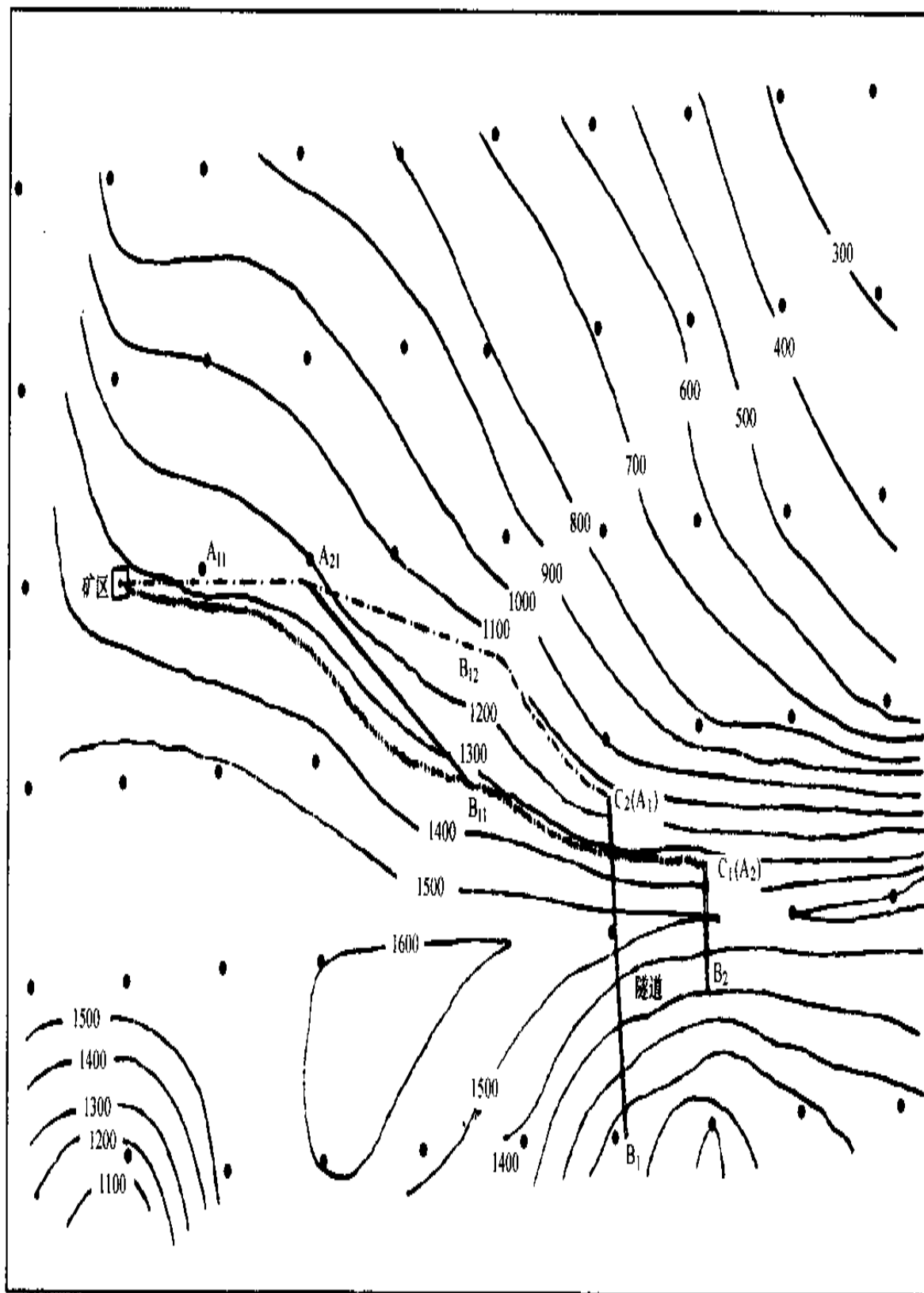
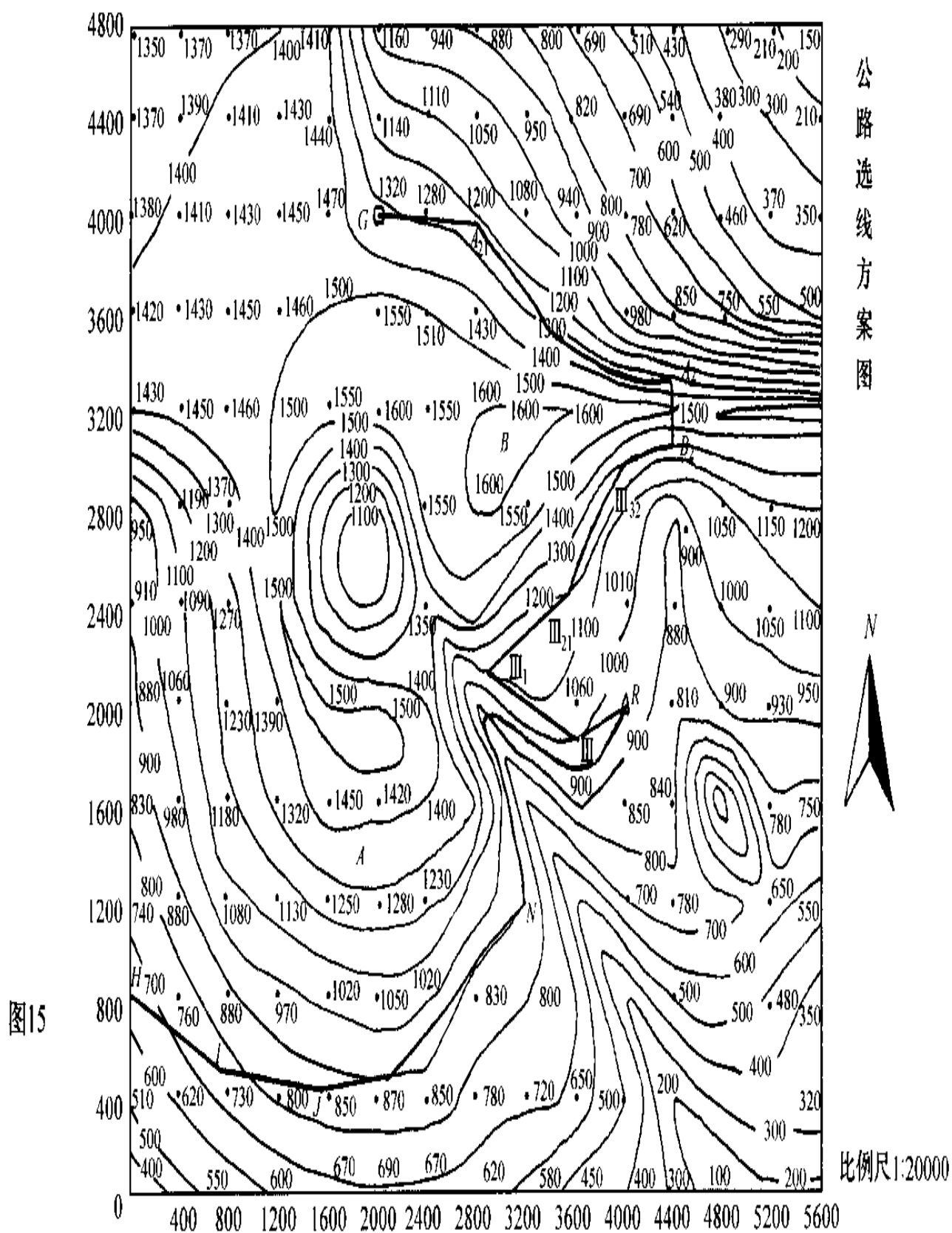


图 14



公路选线方案图



方案I: ---  
方案II: ---

比例尺: 1:20000

图16

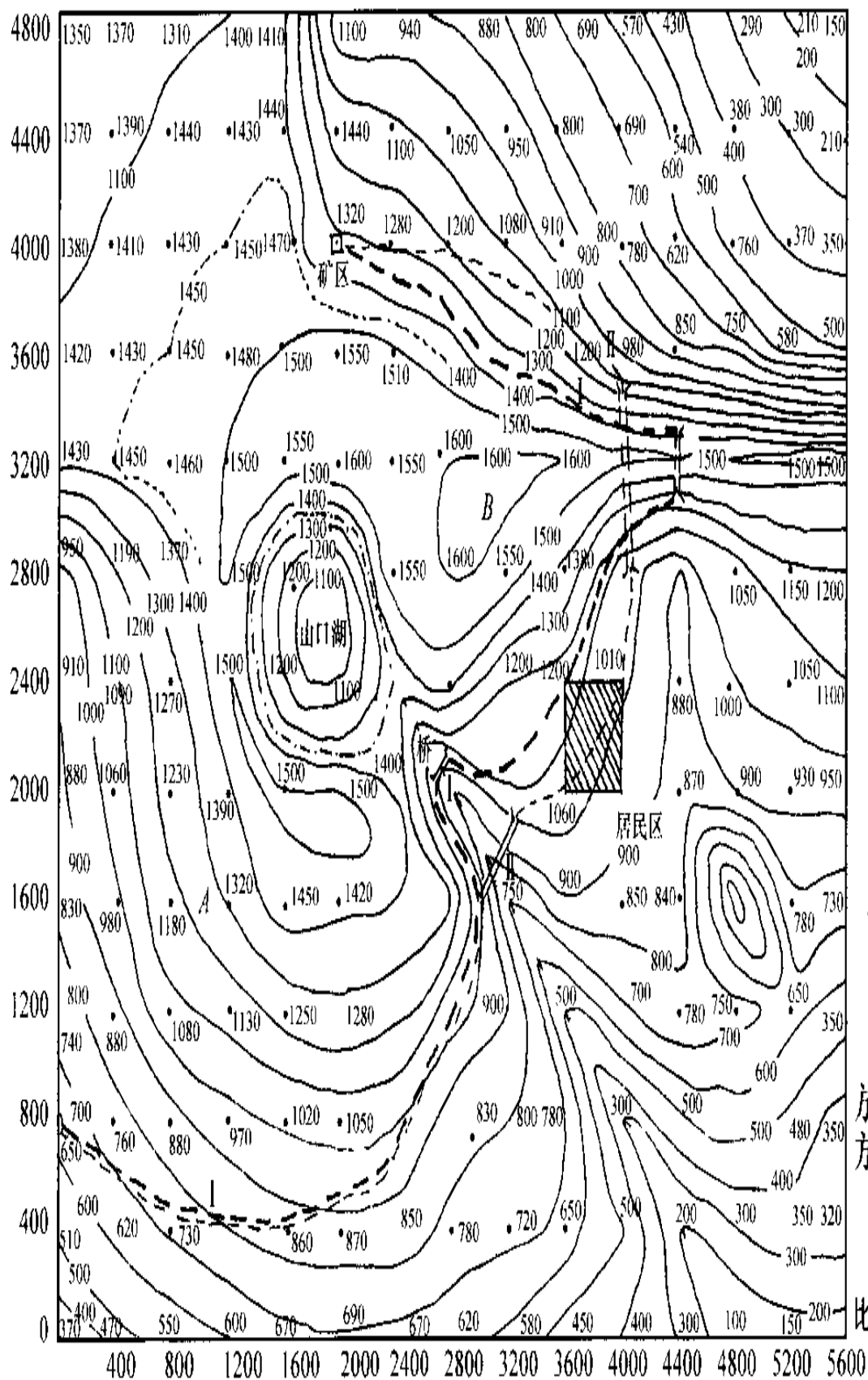


表 3 工程费用计算表

起点至桥 (m)	桥长(m)	桥至居 民点(m)	居民点至 隧道(m)	隧道 长度(m)	隧道至 矿区(m)	最小 总费用
4141.18	→ 68.46	→ 1293.33	800	→ 720	→ 2093.33	367 万元
3202.88	→ 81.48	→ 1786.66	3106.67	↘ 210	↗ 2560	
4091.43	→ 88.06	→ 1133.33	3666.67	↗	↘ 2453.33	

说明:1. 由于量取路线长度时进行了比例尺变换,因此出现了小数点后面的数值,这并不意味着量线的精度有这么高,实际量线的精度为  $\pm 20\text{m}$ .

2. 考虑量线的误差,总费用的误差取 5%,选取路线设计方案为:

$H-I-J-N$ —沿 900m 等高线到桥 2— $R-III-III_1-III_{21}-III_{22}$ —沿 1300m 等高线到隧道  $B_2A_2$ —沿 1320m 等高线到  $B_{11}-A_{21}-G$

注:线路的坐标可以通过图纸量取.(如图 15 所示)

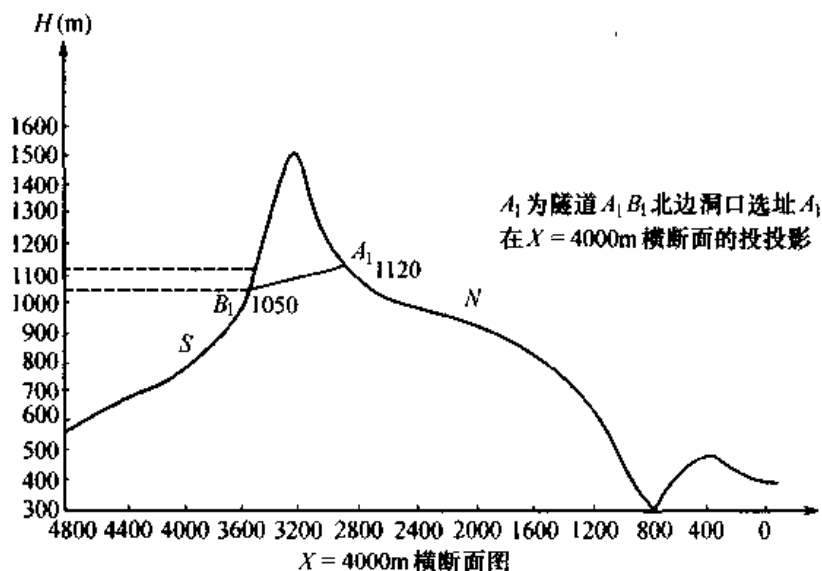


图 17

(隧道的断面图见图 17,图 18,平面图见图 19.)

(4) 居民点改为居民区的情况

根据条件,我们确定了二个方案,作比较(见图 16),下面给出两个方案的工程费用表:

表 4 工程费用表

方 案 项 目	单价(元)	方 案 1		方 案 2	
		长度(m)	费用(万元)	长度(m)	费用(万元)
道路	300	8840	265.2	6960	208.8
桥梁	2000	54	10.8	480	96
隧道	1500 或 3000	210	39	720	216
总计		9594	315	8160	520.8

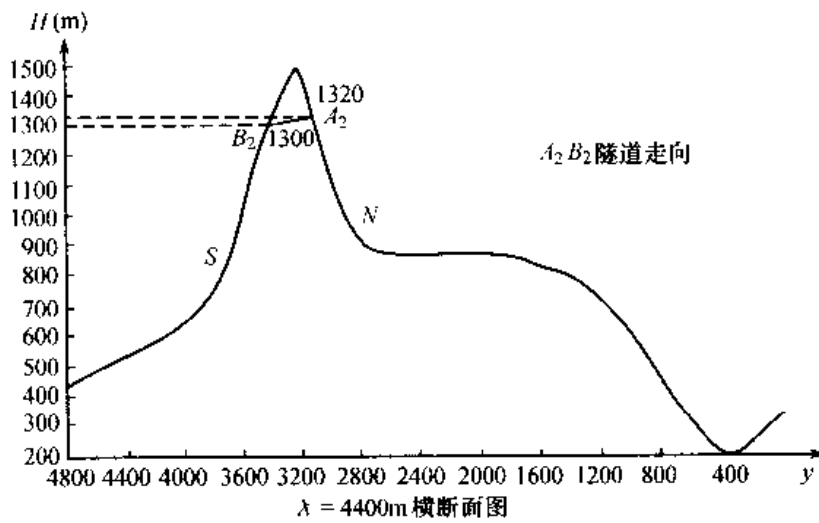


图 18

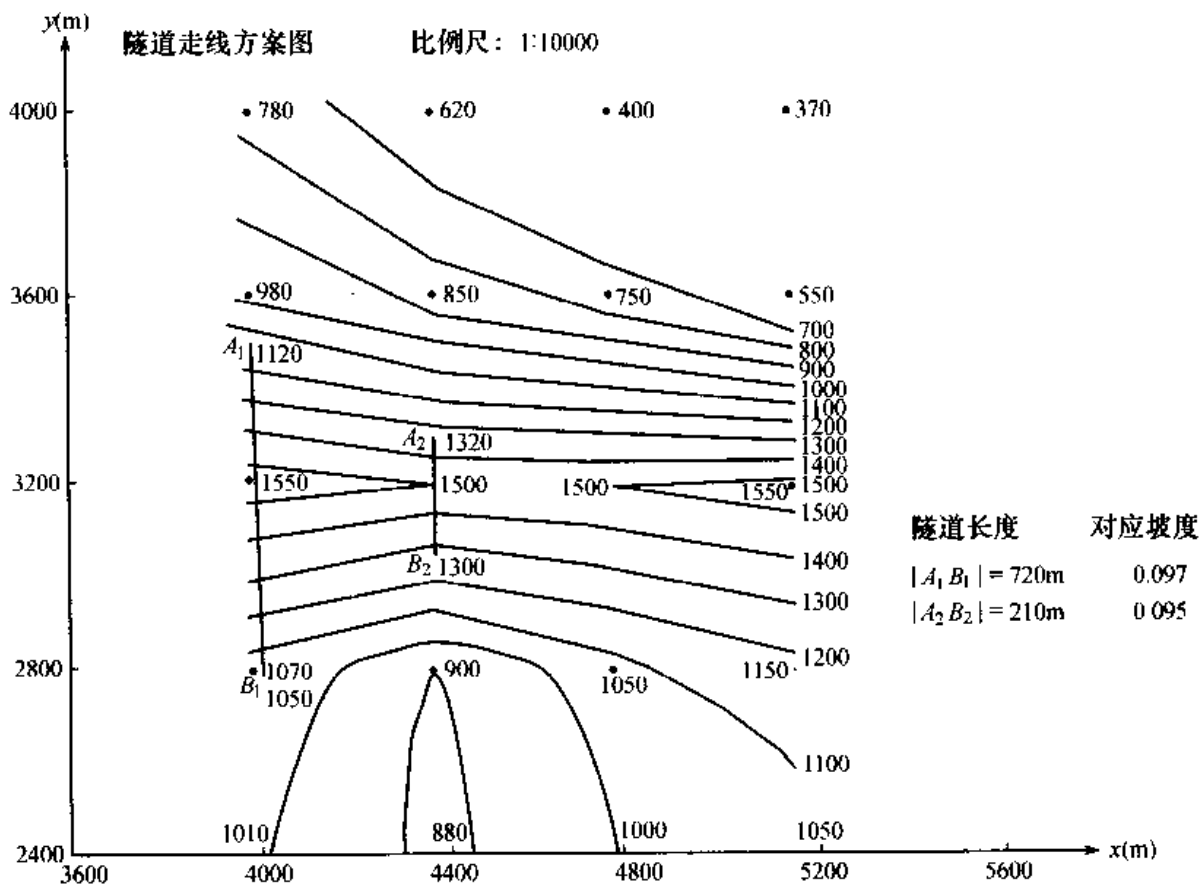


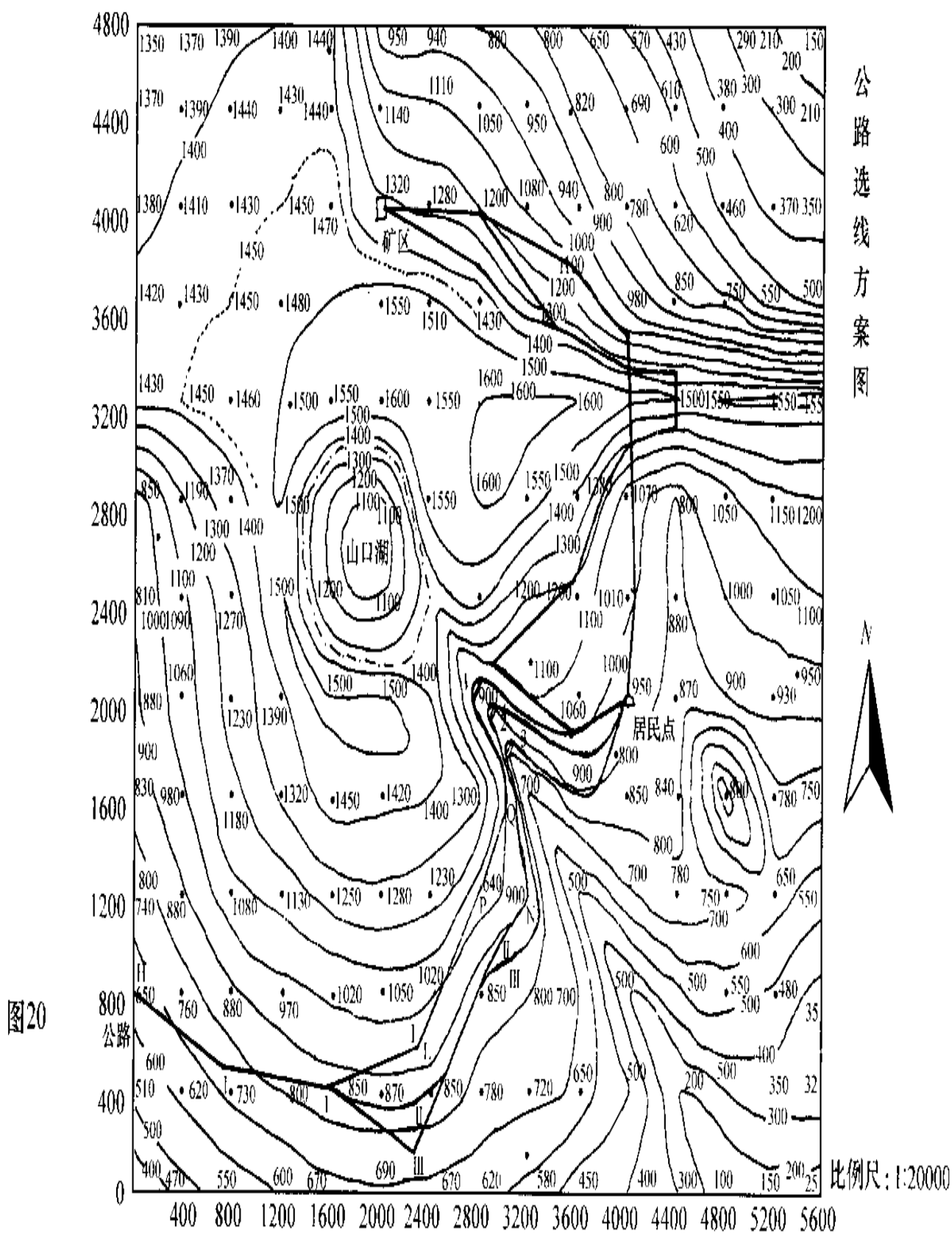
图 19

方案比较：

方案 1 造价比较低,但路程比较长,考虑到矿山公路对居民区的影响,故把公路选在居民区的外围。

方案 2 营造费用高,但路程较短,通过了居民区,对居民区产生了不利影响。



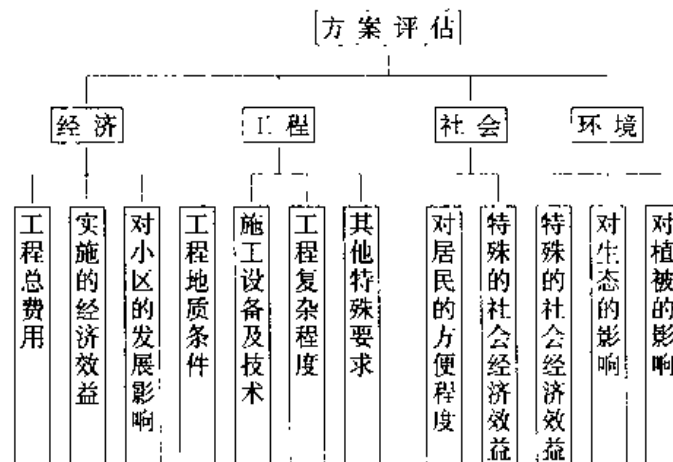


故我们选方案 1。

如果需要减缓其转弯程度,可通过适当增加桥梁和隧道的长度加以解决(这样就增加了费用)。

## 六、动用层次分析法对方案进行评估

我们采取的层次如图所示,在准则层中考虑了经济、工程、社会和环境等因素。



考虑实际工程中的多因素情况,可以采用多准则层的层次分析法对方案进行评估,由于本问题给予的有关信息很少,因此我们仅仅考虑费用与方便程度这两个方面:

表 5 评估权重表

方案评估	经济	社会	优先权重
经济	1	1/5	1/6
社会	5	1	5/6

费用转化评估标度

$I_0 = 100$ , 费用每增加 10 万元, 标度减 1。

方便程度评估标度:

以增加的路线为标度,  $I_0 = 100$ , 长度每增加 200 米, 标度减 1。

现在我们对居民区的两种路线方案进行评估:

表 6 居民区路线评估表

因 素 \ 方 案	I	II
费用(万元)	310	520
路线长度(米)	9100	8100

计算权重:

$$P_I = 5/6 \times (100 - 310/10) + 1/6 \times (100 - 8100/200) = 67.4$$

$$P_{II} = 5/6 \times (100 - 520/10) + 1/6 \times (100 - 8100/200) = 49.9$$

$P_I > P_{II}$ , 故选方案 I.

这里仅举此例以作说明,原因是我们的定权及确定标度都有很强的主观因素.要想通过层次分析法得到更客观的评估标准,需要建立专家评估系统,使评估尺度科学化、规则化、标准化,这样才更具说服力.

## 七、关于建模的几点推广及对线路测设的认识

1. 本问题所考虑的只是公路选线的问题,我们可以把它推广到铁路、管道、给排水系统等线形工程.

2. 如果在更大范围内考虑有线形工程构成的网络系统,运用此方法将会创造巨大的经济效益、社会效益和生态效益.

3. 通过本问题的研究,我们认识到在进行某项重大决策时,主管部门进行专家咨询和多个方案的比较,是至关重要的.

## 八、本模型的优点及改进方向

本模型比较系统地介绍了由已知高程点格网绘制等高线进行公路定线及方案比较的全过程,具有条理性 and 实用性,能够推广到类似的工程中去.

本题所讨论的地形比较简单,对工程具体要求比较粗糙,这主要是由于所给问题没有提供更多的人文和地理信息如地质、水文、气候和地形条件,以及居民点、矿区的规模和性质.否则,我们将考虑得更仔细,甚至可以做出居民区的规划图.

本题所提供的高程数据比较稀疏,因此不便进行土方量的计算.在实际工程中,土方量的工程费用是很大的,如果能够提供更详细的地形资料,如 1:1000 的地形图,那我们将进行这方面的考虑.

## 参考文献

- [1] 测量学,中国建筑工业出版社,1990.
- [2] 计算机地图制图原理,武汉测绘科技大学出版社,1993.
- [3] 秦纭,桥梁工程测量,测绘出版社,1991.
- [4] 城市道路与交通,中国建筑工业出版社,1981.

# 山区公路线路选优方法

汪 昕 石 斌 严红辉

(东南大学 210018)

指导教师 孙志忠

**编者按** 答案从工程和数学两个方面进行了较为清楚的分析,确定了用计算机计算及人工纸上作业相结合的寻优原则。

在道路设计中,作者在确定“桥”和“隧道”位置之后,利用类似的启发式算法寻找各段的较优路线,这些工作有特点,现将这部分内容发表如下。

## 支态规划法

动态规划法是从终止点开始逆推的多阶段最优决策过程。多段最优决策的过程中每一段的结束就是下一段的初始状态。动态规划得到的是全局最优。它在工程上运用很广。如最优路线(选定路线,运输路线,各种管道等)的规划,资源,物资,设备的最优分配,设备的更新等。

如果不采用尝试法,就需要事先由地形图确定一般桥梁、隧道的所有可能的位置赋以相应的权值,用狄克斯特累方法或动态规划等方法,求出整体的最小成本费用。程序见附录(略)。

以上是计算机搜寻路径的方法,理论上只要网格无限细分,并且对地形的了解足够的话,就一定可以找到最优的线路设计方案,但由于数据无法细分到合理的情况(估计应至少分到两点之间的间隔为10米左右才行,此时才能在地形图上表示出任意方向的直线。而视觉上不产生失真。但受机器限制,数据规模不能太大),存在很大误差,因此不能作为定量分析的依据。但由于误差的性质和大小每次计算都差不多,该方法完全可以定性地分析出最优化的线路所在的大概位置,并为进一步的定量分析——纸上定线法作出十分有意义的指导。

在纸上定线之前,我们先规定了几条原则,这些原则有些并未经过严格证明,只是为了符合工程要求或简化问题。

原则1 按允许的最大角度爬坡,即 $\alpha$ 取0.125,此时路最短。

原则2 在坡度较陡的地方,首先考虑自然展线,不得已时方采用回头展线,应尽量利用支脉和平缓山坡等有利地形,并尽量避免在一个山坡上布设较多和相距很近的回头曲线。

原则3 回头曲线的转折处应满足最小平面曲线半径的要求,(可通过插入一定曲率的过渡曲线解决)

原则4 考虑桥的长度时,可取水面宽度加一个常数 $C$ 作为桥长,如取 $C=10$ 米。

原则5 在进行越岭路线选择时,可以考虑通过垭口;对垭口很高,但山体薄窄和分水岭,应考虑在山体的最薄窄处采用过岭隧道。

在纸上实践时,可以利用计算机定性算出的最优路线作为参考,并根据线路的起止点和中间控制点,作好整体布局。

纸上定线最后阶段可以得到一条折线段,必须通过修正和调整,才能最终得出符合标准的设计曲线.这条曲线一般比较接近最优曲线.根据该,可以直接从图上计算出总成本.

在第二问中,居民点改为一个矩形的居民区,因此活动的范围加大了,但根据路的走向,可以分三种情况考虑;即道路经过居民区西北角,中部和东南角.通过对这三种形式分别计算发现西北角最佳.纸上分析也可得相同结论.该结论可通过地形图直观解释:居民区西北角的海拔为 1150 米,最接近隧道的开口位置(海拔约 1300 米),因此爬坡的路线会短很多.当然,当线路改变时,桥的位置也有相应的改变,但变化不大,隧道的位置几乎不变.

数学科学有两个主要方面.第一个也是较为抽象的方面可以叙述为:对结构、模式以及模式的结构和谐性的研究.探求抽象模式结构中的对称性和规则性是纯粹数学的核心.这些探求的目的通常在于了解抽象的概念,但是也常常对其他领域产生实践的和理论的影响.例如,积分几何就构成了 X 射线断层分析照相术(CAT Scan 计算机辅助层析扫描)的理论基础,素数的算术导致了互联网上数据安全传输的完备码的生成,而群的无穷维表示使人们有可能设计通讯中大规模、经济有效的高连通性网络.

数学科学的第二个方面是由要对生活中,通常是由物理学、生物学和商业中,碰到的事件或系统的(数学)建模所激发的.这个方包括以下三个步骤:

对常常并非是清楚明白的实际情况创建一个有明确的(数学)模型.这种(数学)建模需要在忠实于实际情形的模型要求和数学上易于处理的要求之间进行妥协.恰当的妥协通常需要该领域的专家和数学专家的合作.

一通过解析或计算的方法或两者混合的方法来求解该(数学)模型.

一开发在求解特殊数学模型时大概可以重复使用的一般工具.

数学建模的例子包括量子计算机计划,以 DNA 为基础的分子设计,生物学中的模式(图象)形成,工程师日常所用的数值计算的快速 Fourier 变换和多路算法.……数学具有双重的性质:她既是为其精确性和内在的优美而受到敬重的独立的学科,她也是应用领域的丰富的工具资源.可以把数学描述为具有内部抽象性和外部有效性的学科.

(下转第 57 页)

## 6. 锁具装箱(CUMCM 1994 年 B 题)

某厂生产一种弹子锁具,每个锁具的钥匙有 5 个槽,每个槽的高度从 1,2,3,4,5,6 个数(单位略)中任取一数;由于工艺及其他原因,制造锁具时对 5 个槽的高度还有两个限制:至少有 3 个不同的数;相邻两槽的高度之差不能为 5;满足以上条件制造出来的所有互不相同的锁具称为一批.

从顾客的利益出发,自然希望在每批锁具中“一把钥匙开一把锁”.但是在当前工艺条件下,对于同一批中两个锁是否能够互开,有以下试验结果:若二者相对应的 5 个槽的高度中有 4 个相同,另一个槽的高度差为 1,则可能互开;在其他情形下,不可能互开.

原来,销售部门在一批锁具中随意地取 60 个装一箱出售.团体顾客往往购买几箱到几十箱,他们抱怨购得的锁具会出现互开的情形.现聘你为顾问,回答并解决以下的问题:

- 1) 每一批锁具有多少个,装多少箱.
- 2) 为销售部门提出一种方案,包括如何装箱(仍是 60 个锁具一箱),如何给箱子以标志,出售时如何利用这些标志,使团体顾客不再或减少抱怨.
- 3) 采取你提出的方案,团体顾客的购买量不超过多少箱,就可以保证一定不会出现互开的情形.
- 4) 按照原来的装箱办法,如何定量地衡量团体顾客抱怨互开的程度(试对购买一二箱者给出具体结果).

(俞文虬 谭永基提供)

(上接第 56 页)

这种双重性的两部分是深刻地联系在一起的。探求模式中的次序、对称性和规则性是纯数学研究的核心。这种研究的结果是非常持久的,有时候会在发现这些结果的几十年后以一种意想不到的方式找到重要的应用。一个重要的理由就是数学的结果一经证明,决不会被否定——即使它们可能会被更强的结果所取代。相比之下,其他科学是经由一个逐次逼近的过程来达到真理的。

《资深评估小组对美国的数学科学的国际评估报告》,美国科学基金会 (NSF), 1998.

# 关于锁具装箱的数学模型

黄宗虎 李 波 李春福

(电子科技大学 610054)

指导教师 徐全智

**编者按** 本文首先利用组合计数的方法求得一批锁具的总数并按槽高和的奇偶性分类装箱,说明团体顾客一次购买量不超过 49 箱时一定不会出现互开现象.作者试图利用图论的 *König* 定理证明方案的最优性,但在证明过程中,从奇类锁与偶类锁具的对称性并不能得到图中所含有饱和  $V_1$  中每个顶点的匹配结论.最后,引进平均互开对数概念,用概率论方法简洁地得到随机装箱时顾客的抱怨程度的量化结果.

整篇文章建模假设合理,综合运用多种学科工具,较圆满完成题目的要求,并有一定的创见.叙述简洁清楚,结构严谨是一篇较优秀的参赛论文.

**摘 要** 本文建立了一个关于如何对一批弹子锁具进行装箱和标志的模型.

本文首先用组合数学的方法求得了一批锁具的总数为 5880 件,接着分析了能够互开的锁具之间的特性,从而得到以下装箱方案:根据钥匙槽高度之和的不同奇偶性将锁具分类装箱.按照这个方案,当购买量不超过 49 箱时,就可以保证一定不会出现互开的情形.文中用图论知识证明了这个方案是最优的.本文从概率论的角度,引进平均互开对数  $E(m_k)$ ,衡量了按原方案装箱时顾客的抱怨程度.并将本文的方案与原方案进行比较,得出新方案明显减少了顾客抱怨程度的结论.

## 一、问题的重述与分析

某锁厂生产一种弹子锁具,每个锁具的钥匙有 5 个槽,令  $h_i (i = 1, 2, 3, 4, 5)$  为钥匙第  $i$  个槽的高度,则一批锁具应满足如下三个条件.

条件 1 对任意一种槽高排列  $h_1 h_2 h_3 h_4 h_5$  有

$$h_i \in \{1, 2, 3, 4, 5, 6\} (i = 1, 2, 3, 4, 5)$$

条件 2 对任意一种槽高排列  $h_1 h_2 h_3 h_4 h_5$ , 至少有三个槽高互不相同.

条件 3 对任意一种槽高排列  $h_1 h_2 h_3 h_4 h_5$ , 有

$$|h_i - h_{i-1}| \neq 5 \quad (i = 2, 3, 4, 5).$$

而同时满足下面两个条件的两个锁具可以互开,并把这两个锁具称为一个互开对:

\* 1. 两个锁的钥匙有四个槽高度相同;

\* 2. 其余一个槽高度相差 1.

锁厂销售部门原先在一批锁具中随机地取 60 个装为一箱出售,这样一来,成箱购买锁具的顾客总抱怨购得的锁具有互开现象.

我们所关心的是:每一批锁具共有多少个,如何衡量随机装箱造成的团体顾客的抱怨程度以及采取何种方案装箱来尽量避免团体顾客的抱怨.

由一批锁具中,互开对总数是确定的,但在随机装箱之后,对于每一箱而言,锁具互开现象就不可避免地有了随机性,因而可以用统计平均值定量地衡量随机装箱造成的团体顾

客的抱怨程度.

为了能够提出一种方案来装箱和标志,以尽量避免团体顾客的抱怨,就需要找出能够互开的锁具之间的特性,从而使能够互开的锁具分开装箱和标记.

文中用到的符号及其说明

符号	说明	符号	说明
$h_i (i = 1, 2, 3, 4, 5)$	锁具钥匙的第 $i$ 个槽的高度	$\alpha_n(G)$	图 $G$ 的最小覆盖中的顶点个数
$H$	$H = \sum_{i=1}^5 h_i$	$\beta_1(G)$	图 $G$ 的最大匹配中的边数
$D$	满足条件 1 但不满足条件 2 或条件 3 的钥匙槽高度的排列方式的集合,称为除去集	$\beta_0(G)$	图 $G$ 的最大独立顶点集中的顶点个数
$D_i (i = 1, 2, 3, 4, 5)$	$D$ 的一个完全划分	$N_G(V_0)$	图 $G$ 中 $V_0$ 的邻集
$E$	一批锁具中平均每把锁具与其它锁具所能组成的互开对数	$M$	图的最大匹配
$E_K$	$K$ 箱锁具中平均每把锁具与其它锁具所能组成的互开对数	$ V $	集合 $V$ 的元素个数
$m$	总互开对数	$sd$	互开对数
$E(m_k)$	$K$ 箱锁具中平均含有的互开对数		
$S$	某锁具		
$V$	锁具集合 / 顶点集合		
$X$	互开锁具对的关系构成的边的集合		
$G(V, X)$	由 $V$ 和 $X$ 组成的图		
$P(G)$	图 $G$ 的顶点个数		

## 二、模型假设

1. 锁具厂在生产锁具过程中能够准确地知道钥匙的每个槽的高度.
2. 对于同一批中两个锁具,若二者相对应的 5 个槽的高度中有 4 个相同,且另一个槽的高度差为 1,则一定能互开.

## 三、模型的建立及求解

### 1. 确定一批锁具的总数

我们根据生产一批锁具的三个条件,用排列组合的知识对一批锁具的总数目进行求解,其主要过程如下:

- (1) 根据条件 1, 钥匙槽高度的可能排列有  $6^5 = 7776$  种
- (2) 受条件 2 和条件 3 的约束,实际制锁时还要除去一部分钥匙槽高度的排列方式,我们称这些排列方式形成的集合为除去集  $D$ .
- (3) 条件 3 可等价于钥匙槽高度排列方式中不能出现 1 和 6 相邻的情况.
- (4) 对除去集  $D$  可进行如下划分:



令  $D_1 = \{h_1 = h_2 = h_3 = h_4 = h_5 \text{ 的排列}\}$ ;

$D_2 = \{h_i (i = 1, 2, 3, 4, 5) \text{ 中只有两个不同数的排列}\}$ ;

$D_3 = \{h_i (i = 1, 2, 3, 4, 5) \text{ 中有三个不同数且 1 和 6 相邻的排列}\}$ ;

$D_4 = \{h_i (i = 1, 2, 3, 4, 5) \text{ 中有四个不同数且 1 和 6 相邻的排列}\}$ ;

$D_5 = \{h_i (i = 1, 2, 3, 4, 5) \text{ 各不相同且 1 和 6 相邻的排列}\}$ ;

显然,  $D_i (i = 1, 2, 3, 4, 5)$  是  $D$  的一个完全划分, 即  $D_1 \cup D_2 \cup D_3 \cup D_4 \cup D_5 = D$  且  $D_i \cap D_j = \emptyset (i, j = 1, 2, 3, 4, 5 \text{ 且 } i \neq j)$  我们分别求出了  $D_i (i = 1, 2, 3, 4, 5)$  中的元素个数(详细计算见附录 1) 如下:  $|D_1| = 6$ ;  $|D_2| = 450$ ;  $|D_3| = 456$ ;  $|D_4| = 792$ ;  $|D_5| = 192$  ( $|D|$  表示集合  $D$  中元素的个数)

综上, 一批锁具的总数为

$$7776 - (6 + 456 + 792 + 192) = 5880 \text{ 件}$$

可装的箱数为

$$\frac{5880}{60} = 98 \text{ 箱}$$

## 2. 装箱方案

### (1) 对锁具进行分类

当两个锁具相对应的 5 个槽的高度中有 4 个相同, 另一个槽的高度差为 1 时, 它们可以互开. 我们惊奇地发现了下面的规律:

**定理** 在一批锁具中, 能够互开的两锁具的槽高排列  $h_1 h_2 h_3 h_4 h_5$  和  $h'_1 h'_2 h'_3 h'_4 h'_5$ , 其各槽高度之和  $H = \sum_{i=1}^5 h_i$  和  $H' = \sum_{i=1}^5 h'_i$  必然具有不同的奇偶性.

因为互开的两个锁具有四个槽高度相同, 仅有一个槽高度差 1, 那么高度之和  $H = \sum_{i=1}^5 h_i$  和  $H' = \sum_{i=1}^5 h'_i$  必为两个相邻的自然数, 而两个相邻的自然数中, 必有一个为奇数, 另一个为偶数.

根据定理, 我们把锁具按钥匙槽高度之和  $H$  的奇偶分为两类,  $H$  为奇数的属于奇类,  $H$  为偶数的属于偶类. 这样, 能够互开的锁具一定分属于奇类和偶类.

### (2) 求解奇类和偶类中锁具的个数

对一个锁具的钥匙槽高度的排列  $h_1 h_2 h_3 h_4 h_5$ , 我们用 7 分别减去每个高度  $h_i (i = 1, 2, 3, 4, 5)$ , 形成另一个与其对偶的排列:  $(7 - h_1)(7 - h_2)(7 - h_3)(7 - h_4)(7 - h_5)$  记原排列的高度和  $H_0 = \sum_{i=1}^5 h_i$ , 新排列的高度和  $H_1 = \sum_{i=1}^5 (7 - h_i)$

则有

$$H_0 + H_1 = 35$$

为使上式成立,  $H_0, H_1$  中必有一个偶数和一个奇数. 而每一个奇(偶)类中的排列必能从偶(奇)类中找出与其对应的对偶排列, 这种对偶关系是一一对应的. 所以:

$$\text{奇类中锁具数} = \text{偶数中锁具数} = \frac{5880}{2} = 2940$$

故奇类和偶类中的锁具数都为 2940 件.

### (3) 装箱和标记

基于以下讨论,我们可得到如下方案:

① 生产锁具过程中记录每个钥匙槽的高度,从而确定  $H$  的奇偶性,将生产出来的锁具分为奇类和偶类.

② 将奇类和偶类分别装为 49 箱并做“奇”和“偶”的标志.

这样,销售部门可以根据所做标记只选同类的箱子售给团体顾客.只要他们一次购买的箱数不超过 49 箱(即 2940 个锁具),我们就可以保证他们的锁具不会有互开现象,他们也不会抱怨了.

### 3. 对方案最优性的证明

我们用计算机对互开对总数进行穷举法计算(见附录 2),得到在一批锁具中互开对总数为 22778 对.

我们将锁具的互开关系用图表示出来.锁具集合用顶点集合  $V$  表示,如果两个锁具  $V_i, V_j (i, j = 1, 2, \dots, 5880 \text{ 且 } i \neq j)$  能够互开,就用边  $X_{ij}$  连接起来,所有  $X_{ij}$  组成边集合  $X$ .用这种方法得到的图  $G_0(V, X)$  有 5880 个顶点和 22778 条边.

我们引用图论中的一些概念:

最大匹配

最小覆盖

最大独立顶点集

邻集  $N_G(V_0)$

二分图  $G(V_1, V_2, X)$

最大匹配中的边数  $\beta_1(G)$

最大覆盖中的顶点个数  $\alpha_0(G)$

最大独立顶点集中顶点个数  $\beta_0(G)$

这些概念在一般图论理论书中可见到(可参见文献[1],[2]).

我们的目的是寻找图  $G_0$  中的最大独立顶点集,即不包含互开锁对的最大锁具集合.

**引理 1** 二分图  $G(V_1, V_2, X)$  含有饱和  $V_1$  的每个顶点的匹配的充要条件是对所有  $V_0 \subseteq V_1, |N_G(V_0)| \geq |V_0|$  ( $|V_0|$  表示集合  $V_0$  中元素的个数). (1) (见[1])

**引理 2** 在二分图  $G$  中,最大匹配  $M$  的边数等于最小覆盖的顶点数  $\alpha_0(G)$ .

**引理 3**  $\alpha_0(G) + \beta_0(G) = P(G)$ . (见[1])

其中  $P(G)$  为图  $G$  的顶点总数.

对二分图  $G_0(V, X)$ ,我们将奇(偶)类顶点集合定义为  $V_1(V_2)$  得到  $G_0(V_1, V_2, X)$ .

**定理** 二分图  $G_0(V_1, V_2, X)$  的  $V_1, V_2$  是它的两个最大独立顶点集.

证明:我们的图  $G_0(V, X)$ ,对表示奇类锁具的顶点集合  $V_1$ ,由奇类锁具与偶类锁具的对称性可知,  $G_0(V_1, V_2, X)$  满足(1),即图  $G_0(V_1, V_2, X)$  含有饱和  $V_1$  中每个顶点的匹配  $M$ .

又因为  $V_1$  中各顶点互不邻接,则

$$\frac{P(G_0)}{2} \leq |M| \leq \beta_1(G_0) = \alpha_0(G_0) \quad (\text{由引理 2 得})$$

$$\beta_0(G_0) = P(G_0) - \alpha_0(G_0) \quad (\text{由引理 3 得})$$

$$\leq P(G_0) - \frac{P(G_0)}{2}$$

$$= \frac{P(G_0)}{2} = \frac{5880}{2} = 2940$$

即图  $G_0$  的最大独立顶点集中的顶点数不大于 2940,而我们划分的独立集(奇类集和偶

类集)的顶点个数为 2940,因此奇类集  $V_1$  和偶类集  $V_2$  是  $G_0(V_1, V_2, X)$  的两个最大独立顶点集.定理得证.

所以,我们的方案可以使锁具不出现互开情况时购买量达到最大.这就证明了我们的方案是最优的.

#### 4. 定量分析顾客抱怨互开的程度

##### (1) 对于随机装箱的方案

在—批锁具中,互开对总数为  $m = 22778$  对,则平均每个锁具与其它所有锁具能组成的互开对数为  $E = \frac{m}{2940} = 7.75$ (对)

对于任一个指定锁具  $S$ ,从其它 5879 把钥匙中任取一把,能打开该锁的概率是均等的.这样,如果从另外 5879 个锁具中随意取出 59 个锁具与  $S$  装成一箱,则在这一箱中能与  $S$  组成互开对的锁具个数平均有

$$E_1 = \frac{E \times 59}{5879} = 0.078(\text{个})$$

平均含有的互开对数为

$$E(m_1) = \frac{60E_1}{2} = 2.33(\text{对})$$

同理可以得知: $k$  箱锁具中,能与某一个锁具互开的锁具个数平均为  $E_k = E \times \frac{60k-1}{5879}$ ,平均含有的互开对数为

$$E(m_k) = \frac{60k \times E_k}{2} = \frac{60k(60k-1)}{2 \times 5879} \times E = \frac{22778}{5879 \times 98} \times k \times (60k-1)$$

显然, $E_k$  和  $E(m_k)$  的大小反映了购买  $k$  箱的顾客的抱怨程度,即  $E_k$  或  $E(m_k)$  越大,顾客抱怨程度越大.

表 1  $k = 1, k = 2$  和  $k = 49$  的具体结果

箱数 $k$	1	2	49
含有的互开对数 $E(m_k)$	2.33	9.41	5693.5
能与某一锁具 $S$ 互开的锁具数 $E_k$	0.078	0.157	3.87

##### (2) 对于奇偶分类装箱的方案

由前面的分析知,给箱子进行奇偶标志后再出售,只要顾客购买量  $k$  不超过 49 箱,就可以保证不出现互开的情况,顾客自然不会抱怨.当购买量  $k$  超过 49 箱时,可以先从奇(偶)类锁具中取出 49 箱,再从偶(奇)类锁具中任取  $k-49$  箱出售给顾客,互开对显然只产生在 49 箱奇(偶)类锁具与  $k-49$  箱偶(奇)类锁具之间.此时, $k$  箱锁具中的平均互开对数为

$$E'(m_k) = 22778 \times \frac{60 \times (k-49)}{2940} = 22778 \times \frac{k-49}{49}(\text{对}),$$

对比随意装箱的互开对数为

$$E(m_k) = \frac{22778}{5879 \times 98} \times k(60k-1),$$

可知:

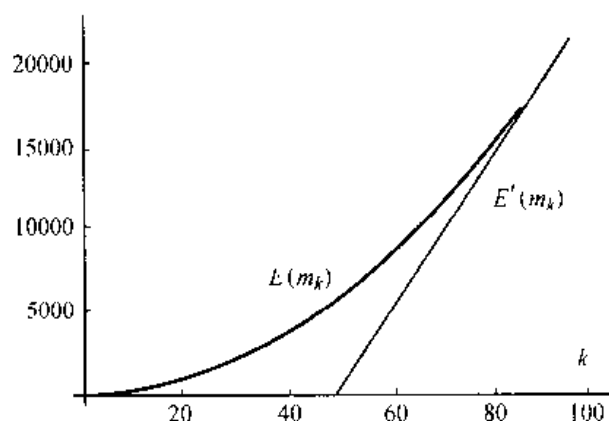
当  $k \leq 49$  时,  $E'(m_k) = 0, E(m_k) > 0$

此时顾客的抱怨被避免了;

当  $49 < k < 98$  时,  $E'(m_k) < E(m_k)$

此时顾客的抱怨程度减少了.

我们在右面画出了  $E'(m_k)$  和  $E(m_k)$  的比较图.



### (3) 在奇偶分类基础上再分类

对于某个锁具  $S$ , 一批中能与它互开的锁具个数  $sd$  是一定的, 那么  $sd$  的范围是多大呢, 我们用穷举法利用计算机得到  $4 \leq sd \leq 10$ , 并且得到了对应于每个  $sd$  的属于奇类的锁具个数 (见下面程序结果), 由于奇偶的对称性, 偶类也有同样的结果.

程序运行结果如下

Total: = 5880

odd: = 2940 Even: 2940

互开总对数: 22778

互开对为 3, 0; 互开对为 4, 45; 互开对为 5, 105;

互开对为 6, 296; 互开对为 7, 699; 互开对为 8, 901;

互开对为 9, 744; 互开对为 10, 150; 互开对为 11, 0;

我们设想在奇偶分类后, 再对奇(偶)类按每个锁具的  $sd$  从小到大对锁具进行装箱, 并标记以序号  $1, 2, \dots, 49$ . 这样当团体顾客购买箱数  $k > 49$  时, 取奇(偶)类 49 箱后剩下的  $k - 49$  箱在偶(奇)类中按序号从小到大取箱, 使得在  $k$  箱中互开总数尽量减少, 这样必然更能减小顾客的抱怨程度.

## 四、模型评价

1. 本模型利用数的奇妙的奇偶性为制锁厂提供了一项很好的锁具装箱方案. 此方案在最大范围内消除了锁具的互开现象, 而且方案简单明了, 施行方便.

2. 本模型综合运用了多种方法对问题进行求解. 其中穷举法简单易懂, 用计算机编程求解的运行时间仅几秒钟; 而概率、图论、组合数学等方理论性较强, 逻辑严密, 而且易于为大家所接受.

3. 当锁具的种类变化时, 即锁具的槽数和每个槽的可能高度的集合发生变化时, 我们的方案仍然适用, 将我们的计算机程序稍作修改, 就能很快得出一批锁具的总数和奇类、偶

类锁具的个数.

4. 我们的模型还能应用到其它类似性质的事物上, 如对条形码的排列组合和识别分类问题, 只需将可能高度改为可能码数, 槽数改为条形码的条数即可.

### 参考文献

- [1] 王朝瑞,《图论》(修订本),北京工业学院出版社出版,1987年.  
[2] 李慰萱,《图论》,湖南科学技术出版社,1980年.

### 六、附录清单

附录 1 除去集  $D$  的元素个数的详细求解

附录 2 程序框图、源程序及运行结果(略)

1. 程序说明

2. PASCAL 源程序

#### 附 录 1

除去集  $D_i (i = 1, 2, 3, 4, 5)$  的元素个数的详细求解

1. 根据  $D_1$  的性质,  $h_i (i = 1, 2, 3, 4, 5)$  都取同一个数, 明显有  $D_1$  个元素个灵敏为

$$C_6^1 = 6 \text{ 个}$$

2. 根据  $D_2$  的性质,  $h_i (i = 1, 2, 3, 4, 5)$  中只有两个不同数  $\omega_1$  和  $\omega_2$ , 我们进行如下考虑, 让每个  $h_i (i = 1, 2, 3, 4, 5)$  都可以任取  $\omega_1$  和  $\omega_2$ , 有  $2^5$  种方式, 但这样做又重复了  $D_1$  中五个槽高均为  $\omega_1$  和五个槽高均为  $\omega_2$  的两种组合, 因而须从  $2^5$  中减去 2, 最后考虑到  $\omega_1, \omega_2 \in \{1, 2, 3, 4, 5, 6\}$ , 可有  $C_6^2$  种取法, 故而  $D_2$  的元素个数为

$$C_6^2(2^5 - 2) = 450 \text{ 个}$$

3. 根据  $D_3$  的性质,  $h_i (i = 1, 2, 3, 4, 5)$  中必有 1 和 6, 记另外一个与 1 和 6 不同的高度数为  $\omega$ , 我们考虑如下完全划分:

(a)  $h_i (i = 1, 2, 3, 4, 5)$  中三个 1 和一个 6.

第一步, 从  $h_1 h_2 h_3 h_4 h_5$  中选出了三个来安排 1, 有  $C_5^3$  种方法, 第二步, 从剩余的两个槽位中选一个来安排 6, 剩余最后一个位置自然要排  $\omega$ , 因而有  $P_2^1 C_5^3$  种安排方法. 但这样做纳入了  $\bar{D}$  集中的两个元素 111 $\omega$ 6 和 6 $\omega$ 111, 须从安排方法中减去, 又  $\omega$  可以  $\{2, 3, 4, 5\}$  中任取一数, 故而共有排列方式

$$C_4^1(2C_5^3 - 2) = 72 \text{ 种}$$

(b)  $h_i (i = 1, 2, 3, 4, 5)$  中三个 6 和一个 1.

与(a)同理可知共有排列方式

$$C_4^1 2(C_5^3 - 2) = 72 \text{ 种}$$

(c)  $h_i (i = 1, 2, 3, 4, 5)$  中有两个 1 和两个 6.

第一步从  $h_1 h_2 h_3 h_4 h_5$  中任选两上来排 1, 有  $C_5^2$  种选法; 第二步, 从剩余三个位置中任选两个安排 6, 剩下最后一个位置自然要排  $\omega$ , 因而有  $C_3^2 C_5^2$  种排法. 但这样做纳入  $\bar{D}$  集中的两元素 11 $\omega$ 66 和 66 $\omega$ 11, 须从排法中减去 2. 又  $\omega$  可以  $\{2, 3, 4, 5\}$  中任取一数, 故而共有排列

方式.

$$C_4^1(C_3^3C_3^2 + 2) = 112 \text{ 种}$$

(d)  $h_i (i = 1, 2, 3, 4, 5)$  中有两个 1 和一个 6.

第一步, 从  $h_1h_2h_3h_4h_5$  中选出二个安排 1, 有  $C_5^2$  种方法, 第二步, 从剩余 3 个位置中选一个排 6, 剩余的一个位置都排  $\omega$ , 因而有  $C_3^3C_3^1$  种排法. 但这样做纳入了  $\bar{D}$  中的 9 个元素  $11\omega\omega6, 11\omega6\omega, 6\omega\omega11, \omega6\omega11, 1\omega1\omega6, 1\omega6\omega1, 6\omega1\omega1, \omega11\omega6$  和  $6\omega11\omega$ , 须从排法中减去. 又  $\omega$  可以  $\{2, 3, 4, 5\}$  中任取一个数, 故而共有排列方式

$$C_4^1(C_3^3C_3^1 - 9) = 84 \text{ 种}$$

(e)  $h_i (i = 1, 2, 3, 4, 5)$  中有两个 6 和一个 1.

与(d) 同理可知共有排列方式

$$C_4^1(C_3^3C_3^1 - 9) = 84 \text{ 种}$$

(f)  $h_i (i = 1, 2, 3, 4, 5)$  中只有一个 1 和一个 6.

这样考虑, 三个  $\omega$  可形成四个间隔位, 将 1, 6 做为一个数插入四个间隔位有  $C_4^1$  种方法, 这样得到的组合方式显然属于  $D_3$ , 同样可将 6, 1 做为一个数插入四个间隔位, 而  $\omega$  可从  $\{2, 3, 4, 5\}$  中任取一数, 故而共有排列

$$C_4^1(P_4^1P_2^2) = 32 \text{ 种}$$

综上所述,  $D_3$  集中元素的个数为

$$72 + 72 + 112 + 84 + 84 + 32 = 456 \text{ 个}$$

4. 根据  $D_4$  的性质,  $h_i (i = 1, 2, 3, 4, 5)$  中必有 1 和 6, 记另外两个与 1 和 6 不同而且相互之间也不同的高度数为  $\omega_1, \omega_2$ , 我们考虑如下完全划分:

(a)  $h_i (i = 1, 2, 3, 4, 5)$  中有两个 1 和一个 6.

与 3 - (d) 的情况相类比, 现在的情况是  $\omega_1 \neq \omega_2$  且  $\omega_1, \omega_2 \in \{2, 3, 4, 5\}$ , 而前面的情况是有两个相同的  $\omega$  且  $\omega \in \{2, 3, 4, 5\}$ . 因而同理容易推知现在共有排列方式

$$C_4^2[P_2^2(C_3^3C_3^1 - 9)] = 252 \text{ 种}$$

(b)  $h_i (i = 1, 2, 3, 4, 5)$  中有一个 1 和两个 6.

同理, 与 3 - (e) 的情况类比可知共有排列方式

$$C_4^2[P_2^2(C_3^3C_3^1 - 9)] = 252 \text{ 种}$$

(c)  $h_i (i = 1, 2, 3, 4, 5)$  中只有一个 1 和一个 6.

与 3 - (f) 的情况类比, 现在的情况相当于原来的构成间隔位的方式增多了, 可以是两个  $\omega_1$  和一个  $\omega_2$  或两个  $\omega_2$  和一个  $\omega_1$  构成间隔位, 它们都有  $C_3^2C_1^1$  种方式, 又  $\omega_1, \omega_2 \in \{2, 3, 4, 5\}$ , 因而现在的情形下共有排列方式

$$C_4^2[C_3^2C_1^1(P_4^1P_2^2) + C_3^2C_1^1(P_4^1P_2^2)] = 288 \text{ 种}$$

综上所述,  $D_4$  集中元素个数为

$$252 + 252 + 288 = 792 \text{ 个}$$

5. 根据  $D_5$  的性质,  $h_i (i = 1, 2, 3, 4, 5)$  中只有一个 1 和一个 6, 记其余三个与 1 和 6 不同的数为  $\omega_1, \omega_2, \omega_3 \in \{2, 3, 4, 5\}$ , 故有  $P_4^3$  种构成间隔位的形成. 由此可知  $D_5$  集中元素的个数为

$$P_4^3(P_4^1P_2^2) = 192 \text{ 个}.$$

# 装箱与销售模型

兰州铁道学院 刘 振 杨文青 何新宇

指导教师 俞建宁

**编者按** 本文较详细地讨论了装箱和连续销售问题,证明了从任一箱起顺次销售能保证至少有 35 箱不会发生互开现象,进而用字典装箱可改进到 42 箱.这点是很有创见和特色的.

## 假 设

1. 两把锁具对应的 5 个槽的高度中有 4 个对应相同,另一个槽的高度差为 1 时,两锁必能互开.

## 模型 I (1)

在考虑使团体顾客不再或减少抱怨的前提下、兼顾销售方便,我们拟采用连续数字作为箱的编号,销售时按箱的编号连续销售.

设一批锁具中第  $i$  把钥匙从一端开始顺次各槽的高度值组成一个向量记为  $B_i$

$$B_i = (b_1^{(i)}, b_2^{(i)}, \dots, b_m^{(i)}),$$

其中  $m$  为一批锁具中各钥匙槽的个数(本模型  $m = 5$ ).

从而可有如下结论:

结论 1:若  $|B_i| - |B_j| \neq 1$ ,则第  $i$  把锁和第  $j$  把锁不能互开(证略).

由结论 1 得如下推论.

推论 1:要使团体顾客不再或减少抱怨,必须使  $||B_i| - |B_j|| = 1$  的两锁具装入编号相距尽可能大的箱中.

推论 2:模同为偶数的锁具间不能互开.

模同为奇数的锁具间不能互开.

模相等的锁具间不能互开.

由以上分析得到下面的装箱方案.

将模为偶数的锁装入前  $i$  箱,模为奇数的锁装入后  $j$  箱( $i + j = Y$ ,  $Y$  为总箱数,本题  $Y = 98$ ),且把模为偶数和奇数的锁具分别按模从小到大的顺序排列后,依次装箱.

记模为  $|B_i|$  的所有锁具数为  $\alpha(|B_i|)$ .

本文用计算机程序对不同  $|B_i|$  的  $\alpha$  进行计算,结果如下:

$$2(8) = 20, 2(9) = 50, 2(10) = 120, 2(11) = 162, 2(12) = 251,$$

$$2(13) = 322, 2(14) = 405, 2(15) = 508, 2(16) = 539,$$

$$2(17) = 563,$$

$$2(18) = 563, 2(19) = 539, 2(20) = 508, 2(21) = 405, 2(22) = 322,$$

$$2(23) = 251, 2(24) = 162, 2(25) = 120, 2(26) = 50, 2(27) = 20.$$

由统计数据及装箱方案可得表 1:

表 1 模型 I (1) 的装箱方案

$n \setminus  B_i $ (编号)	8	10	12	14	16	18	20	22	24	26	9	11	13	15	17	19	21	23	25	27
$a \setminus  B_i $	20	120	251	405	539	563	508	322	162	50	50	162	322	508	563	539	405	251	120	20
混合箱	1	3	7	14	23	32	41	46	49	50	53	58	67	76	85	92	96	98		
顺次装箱 (自然排序)		2	4	8	15	24	33	42	47	51	54	59	68	77	86	93	97			
			5	9	16	25	34	43	48	52	55	60	69	78	87	94				
			6	10	17	26	35	44					56	61	70	79	88	95		
				11	18	27	36	45					57	62	71	80	89			
				12	19	28	37							63	72	81	90			
				13	20	29	38							64	73	82	91			
					21	30	39								65	74	83			
					22	31	40								66	75	84			

对表 1 说明如下:

(1) 表中第一行为上述装箱方案对应的模  $|B_i|$  的排列.

(2) 表中第二行为模为  $|B_i|$  的锁具的个数.

(3) 第三行以下为模为  $|B_i|$  的锁装入箱的对应编号.

若锁具  $|B_i|$  与锁具  $|B_j|$  能互开, 记  $|B_i|$  所在箱与  $|B_j|$  所在箱在自然下的距离为  $|S_{ij}|$ .

记  $M(S_{ij})$  为从任意箱处顺次连续拿取一定不会出现互开锁的最大箱数, 则

$$M(S_{ij}) = \min\{|S_{ij}|\} - 1.$$

据表 1 有:

$$S_{ij} = \{50, 49, 48, 47, 46, 45, 43, 42, 41, 40, 39, 38, 37, 36\},$$

$$M(S_{ij}) = 36 - 1 = 35.$$

据以上结果作购买量和出售方案对照表(表 2), 表 2 中曲线上各标值表示由该点对应列箱号开始顺次连续购买, 一定不会出现互开锁的最大箱数. 显然, 保证一定不会出现互开锁的最大购买量为 49 箱.

表 2 使用说明:

(1) 若购买箱数  $N \leq 35$  箱, 可以从任一标号开始, 顺次连续取.

(2) 若购买箱数  $N = 49$  箱, 只能从第 1 箱或第 50 箱或第 98 箱连续取.

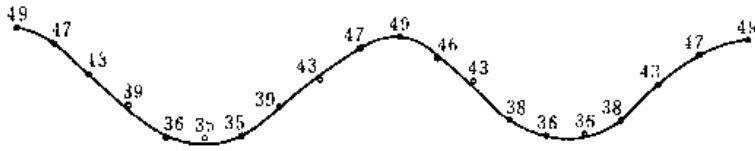
(3) 若购买  $35 < a \leq N \leq b < 49$  箱, (其中  $a, b$  为表 2 上可取的标值范围的外限)

(i) 当  $b$  位于  $a$  点左边时,  $b$  点取对应箱号的上限,  $a$  点取对应箱号的下限.

(ii) 当  $b$  位于  $a$  的右边时,  $b$  取对应箱的下限,  $a$  取对应箱的上限.



表 2

购买箱数																								
$ B_i $	8	10	12	14	16	18	20	22	24	26	9	11	13	15	17	19	21	23	25	27				
$\alpha  B_i $	20	120	251	405	539	563	508	322	162	50	50	162	322	508	563	539	405	251	120	20				
出售 标 号	1	3	7	14	23	32	41	46	49	50	53	58	67	76	85	92	96	98						
	2	4	8	15	24	33	42	47	51	54	59	68	77	86	93	97								
	5	9	16	25	34	43	48	52	55	60	69	78	87	94										
	6	10	17	26	35	44					56	61	70	79	88	95								
	11	18	27	36	45					57	62	71	80	89										
	12	19	28	37									63	72	81	90								
	13	20	29	38									64	73	82	91								
	21	30	39													65	74	83						
	22	31	40													66	75	84						

## 模型 I (2)

模型 I (1) 只考虑了不同  $|B_i|$  之间的排序, 没有考虑相同  $|B_i|$  的锁具之间的排序. 因而装箱是方便的, 但购买量并非最大. 为了进一步增大从任一箱起的最大购买量, 我们对相同  $|B_i|$  的锁具之间进行字典排序后, 并依次按 1, 2, 3,  $\dots$ , 5880 进行标记.

记  $|B_i| = (b_1^{(i)}, b_2^{(i)}, \dots, b_m^{(i)})$  中的元素组成一个  $m$  位数值为

$$M_i = b_1^{(i)} b_2^{(i)} \cdots b_m^{(i)},$$

则  $B_j = (b_1^{(j)}, b_2^{(j)}, \dots, b_m^{(j)})$ .

$m$  位数值为:

$$M_j = b_1^{(j)} b_2^{(j)} \cdots b_m^{(j)}.$$

结论 2:  $M(S_0)$  必然出现在  $|b_1^{(i)} - b_1^{(j)}| = 1$  的情形中.

证明: 将  $|B_i|$  组成的有序向量记为  $T$ , 则

$$T = (8, 10, 12, 14, 16, 18, 20, 22, 24, 26, \\ 9, 11, 13, 15, 17, 19, 21, 23, 25, 27).$$

将其对应的有序向量集空间定义为  $ZK$ , 则

$$ZK = (\{20\}, \{120\}, \{251\}, \{405\}, \{539\}, \{563\}, \{208\}, \{322\}, \\ \{162\}, \{50\}, \{50\}, \{162\}, \{322\}, \{508\}, \{563\}, \{539\}, \\ \{405\}, \{251\}, \{120\}, \{20\}).$$

在向量集空间  $ZK$  中,对于  $\|B_i\| - \|B_j\| = 1$  的两个子向量集  $\{a \in B_i\}$  和  $\{a \in B_j\}$ ,在按字典排序的条件下,为了保证不互开,只有当首位数  $|b_1^{(i)} - b_1^{(j)}| = 1$  时向量所对应的锁具编号才相距“最近”。

下面对向量集空间  $ZK$  中的每一个子向量集  $\{a \in B_i\}$  按首位数  $b_1^{(i)} = 1, 2, \dots, 6$  的大小进行字典排序,由表 2 图象可知,只需在  $\|B_i\|$  为  $14 \leftrightarrow 13, 16 \leftrightarrow 15, 18 \leftrightarrow 17, 20 \leftrightarrow 19, 22 \leftrightarrow 21$  这五种情况间进行分析,  $M(S_{ij})$  必在其间,具体分析见表 3。

分析表 3,可得  $M(S_{ij}) = 42.7$ ,取下整为 42。

也可用计算机进行全范围搜索求解。

将有序向量集空间中的各向量按  $1, 2, 3, \dots, 5880$  标记以后,运用计算机依据前述的方案进行搜索,求出各种可能情况下的最小距离  $M(S_{ij})$ ,以该最小距离对应的向量数,标记号差值。

运行结果为:

最小距离对应向量标号之差为 2562。

相应向量 128 个,即互开锁具数为 128 把。

由  $M(S_{ij}) = 2562/60 = 42.7$ ,取下整为 42 箱。

装箱方案:将字典排序后的有序向量集空间中的 5880 个向量按顺序提出来,排成一列,然后将其对应锁具每 60 把一箱,依次装箱,并把各箱的  $1 \sim 98$  依次编号。

销售方案:出售时按箱编号顺次取出。显然,保证不互开的最小“最大购买量”为 42 ~ 49。

若购买 42 箱,可以从任一箱开始顺次连续拿取,两模型都是在连续生产、循环销售的条件下建立的。但模型 I(1) 和模型 I(2) 相比,模型 I(1) 生产简单,装箱方便。模型 I(2) 在保证不会出现互开锁的前提下,购买箱数由 35 箱增到 42 箱,可是在一定程度上增加了生产、装箱的复杂度。

表 3

箱 距 $ B_i $	14						16						18						20						22						总 和
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	
13	44.38																														90
		44.35																													80
			45.95																												60
				44.28																											52
					44.28																										30
																															10
15							43.05																								102
							43.05																								112
								43.05																							102
									43.15																						85
										43.22																					70
																															34
																															78
17																															112
																															106
																															104
																															96
																															67
																															50
																															84
19																															96
																															107
																															106
																															95
																															19
21																															52
																															60
																															84
																															44.72
总 和	98	92	84	60	52	19	95	106	108	96	84	50	67	96	104	106	112	78	34	70	88	102	112	102	12	28	52	60	80	90	98

# 关于不互开箱数上限的一个讨论

魏明骏 于清娟 王振宇

(中国科学技术大学 230026)

指导教师 程继新

**编者按** 本文讨论了锁具分箱问题中,顾客最多购买多少箱时不会出现互开现象的问题,且给出了箱数的一个上限.虽然所得上限过大,但证明方法巧妙、简洁,具有启发性,特将有关部分予以发表.

我们按槽高之和的奇偶性将锁具划分为  $A$ 、 $B$  两个集合.容易证明,每个集合的锁具数为  $5880/2 = 2940$  个, $A$  和  $B$  中的锁具可分别装 49 箱.以  $A \cup B = V$  为顶集,当且仅当  $a \in A, b \in B, a$  与  $b$  可互开时,在顶  $a$  与  $b$  间连一条边,得一图  $G(V, E)$ .可以证明  $G(V, E)$  为二分图,  $A$  与  $B$  为独立集.于是,不互开锁具箱数的下限为 49.下面,我们证明不互开箱数的上限为 54.证明如下:

假设团体顾客购买了  $n$  箱锁具而不互开,且  $n > 54$ ,  $n$  箱锁具必分属于  $A$ 、 $B$  两个集合.不妨设其中有  $n_1$  箱锁具,这些锁具的集合  $N_1 \subset A$ , (可以证明  $N_1 \neq A$ ) 和有  $n_2$  箱锁具,这些锁具的集合  $N_2 \subset B$ .令  $A - N_1 = N_3, |N_3| = n_3$  (集合  $N_3$  中的锁具装在  $n_3$  箱中),  $B - N_2 = N_4, |N_4| = n_4$ , 且  $n_1, n_2, n_3, n_4$  满足:

$$n_1 + n_2 = n, n_1 + n_3 = 49, n_2 + n_4 = 49 \quad (1)$$

故

$$n_3 + n_4 = 98 - (n_1 + n_2)$$

又

$$n_1 + n_2 = n > 54$$

于是

$$n_3 + n_4 < 44 \quad (2)$$

由于  $n$  箱锁具不互开,所以在  $G$  中  $N_1 \cup N_2$  为独立集;又  $N_1 \cup N_3 = A$  为独立集,所以  $N_1$  中的顶只能与  $N_4$  中的顶相邻,而  $N_4$  中的顶则与  $N_1 \cup N_3$  中的顶可能相邻.所以  $N_1$  中顶的总度数  $D_1$  与  $N_4$  中顶的总度数  $D_4$  满足:

$$D_1 \leq D_4 \quad (3)$$

同理

$$D_2 \leq D_3,$$

于是

$$D_1 + D_2 \leq D_3 + D_4 \quad (4)$$

借助计算机可统计出  $G$  中顶的总度数  $D = 45556$ ,

又

$$\sum_{i=1}^4 D_i = D, \text{再由(4)得}$$

$$D_3 + D_4 \geq D/2 = 22778 \quad (5)$$

借助计算机统计,我们还可以知道度数的精确分布(即不同度数各有多少锁具)如下:

$$R(1) = R(2) = R(3) = 0, R(4) = 90, R(5) = 210,$$

$$R(6) = 592, R(7) = 1398, R(8) = 1802,$$

$$R(9) = 1488, R(10) = 300$$

其中  $R(i)$  表示  $i$  度顶的个数.

由此,可求出43箱锁具能容纳的最大度数为 22728;

44箱锁具能容纳的最大度数为 23208.

于是

$$\text{当 } n_3 + n_4 < 44 \text{ 时, } D_3 + D_4 < 22778 \quad (6)$$

与(5)矛盾,由此证出  $n > 54$  不成立,即不互开锁具箱数的上限为 54,证毕.

### 参考文献

- [1] 王树禾编著,《图论及其算法》,中国科学技术大学出版社,1990.

确实,数学是决定性的相互交叉的学科。数学是取得全面进步的出发点。数学既是强有力的洞察工具,又是科学的共同语言。我把她称为科学的“世界语”。(原文:Indeed, Mathematics is the ultimate cross-cutting discipline. It is the springboard for advances across the board. Mathematics is both a powerful tool for insight, and it's a common language for science, I refer it as the "Esperanto" of science. Rita Colwell (Director of NSF)(美国科学基金会主任 Rita Colwell)

NSF Launches Major Initiative in Mathematics, Allyn Jackson, Notices of AMS, v. 48(2001), no. 2, 190 - 192.

# 独立集最大性的讨论

杨 琮 尹雪钰 李 程

(北京师范大学 100875)

指导教师 曾文艺

**编者按** 本部分将最优装箱问题转化为图论的二分图问题,用图论方法说明了  $S, T$  分别是两个极大无关集且  $|S| = |T| = 2940$ ,并借助计算机找到了一个匹配,其边数为 2928,说明了最大匹配边数  $\geq 2928$ ,再利用 Konig 定理得出最大独立集个数  $\leq 2952 < 3000$ ,从而得到文中不能互开的最大箱数为 49 箱.

本部分思路清晰,行文流畅;但未能从理论上证明 2940 就是最大独立集数,只是用计算机给出了 2952 这个上界.

三 采取我们的方案,团体顾客购买量不超过 49 箱,可以保证一定不会出现互开情形,

我们分析如下:(注: $S = \{\text{各位数字和为奇数的合格锁的集合}\}$

$T = \{\text{各位数字和为偶数的合格锁的集合}\}$ )

将锁看成点,两个锁具互开则可看成两点之间存在一条边.则一批锁具构成顶点数为  $P = 5880$  的简单图.由前所证,在同  $S$  或  $T$  集合中不会有互开关系,互开关系只会发生在  $S$  中点与  $T$  中点之间.由图论中二部图定义,这  $P$  个锁具构成二部图  $G = (S, T, E)$

我们引入以下几个概念:(见[1] 田丰等,《图与网络流理论》,科学出版社)

(1) 边无关集:给定图  $G = (V, E)$ ,若边集  $M \subseteq E$  中任两条边都没有公共顶点,称  $M$  是  $G$  的边无关集;记最大边无关集的边数为  $\beta_1(G)$ ,显然  $\beta_1(G) \leq P/2$  ([1]p.60)

(2) 点无关集:对图  $G = (V, E)$ ,若点集  $I \subseteq V$  中任意两个点间都不存在边,称  $I$  是最大点无关集;包含点数最多的点无关集称为最大点无关集,其中点的个数记为  $\beta_0(G)$  ([1]p.77)

再不加证明地引入以下几个定理:(其中  $\alpha_0(G)$  为图的点覆盖数)

① 设  $G = (S, T, E)$  是二部图,则  $\beta_1(G) = \alpha_0(G)$  ([1]p.64)

②  $\alpha_1(G) + \beta_0(G) = P(G)$

易由 ①② 得,对于二部图,  $\beta_1(G) + \beta_0(G) = P(G)$  (\*)

显然对于本题的二部图  $G = (S, T, E)$ ,  $S = T = 2940$ ,而  $S, T$  都是极大点无关集,即若再往  $S$  或  $T$  中增加集合外一点,则集合不再是无关集,于是  $\beta_0(G) \geq 2940$  (\*\*)

另外,我们利用计算机程序(见算法)找出了一个边无关集,其中边数为 2928,从而最大边无关数  $\beta_1(G) \geq 2928$ .由(\*)得

$$\beta_0(G) = P(G) - \beta_1(G) \leq 2952$$

综合(\*)得  $2940 \leq \beta_0(G) \leq 2952$ ,这就是说明我们如从一批锁具中任取 50 箱,其锁数为  $3000 > 2952 \geq \beta_0(G)$ ,其中必存在相关(可互开)的锁.

由前面的装箱办法,我们将  $S$  类中的 2940 把锁装成 49 箱,  $T$  类中的 2940 把锁装成 49

箱;从上面的分析可知我们提出的两个内部不互开的 49 箱的分类是最优的. 团体顾客购买不超过 49 箱,可以保证一定不会出现互开.

#### 附

利用程序找尽量大的边无关集的计算步骤:

算法主要目的是找出尽可能多的匹配,从而提高  $\beta_1(G)$  的下限,也即缩小了最大点无关集个数  $\beta_0(G)$  的上限.

① 产生 5880 把合格锁,按数字特征(锁各位数字之和分组)

② 对每把锁求出与之互开的锁

③ 标记所有的锁为没有匹配(NoMate)

④ 从数字特征的 8 的组开始,在标记为 NoMate 的锁中查找互开锁(与之互开的锁数字特征必为 9),找到后将两锁都标记为已匹配(HasMate)

⑤ 依次对特征为 9 - 27 的所有锁进行步骤 ④,若发现与某未匹配锁互开的锁都已匹配,则产生一不可匹配的锁,直到所有锁都处理完.

程序运行后,给出 2928 对匹配点,即 2928 条无关的边,因此  $\beta_1(G) \geq 2928$ .

那么数学的“应用”是由什么组成的呢?在我看来它能描述如下,它的任务在于基于物质世界某些对象的性态所遵从的一般规律,预言这些对象在给定条件下的行为。通过把物质对象对应到认定能“表示”这些物质对象的数学对象以及把控制前者的规律对应到数学对象之间的数学关系,就能构造所研究的情形的数学模型;这样,把原来的问题翻译为数学问题,如果能以精确或近似方式求解此数学问题,就可以再把所得到的解翻译回去,从而“解”出原先提出的问题。

《当代数学——为了人类心智的荣耀》——让·迪厄多内著,沈永欢译,上海教育出版社, pp. 21 - 22, 1999.

# 抱怨程度模型

顾宗良 党 京 卢敦陆

(合肥工业大学 230009)

指导教师 周永务

**编者按** 在本次竞赛中,很大部分试卷中都用平均互开次数的数学期望值定义为抱怨程度的度量.这自然就得出购买箱数越多,抱怨程度越大,这不太符合顾客的心理状态.本文认为顾客抱怨程度与购买箱数成反比,与检验结果成指数关系,在定义抱怨程度函数上是很有特色的.

## 1. 总体互开概率

将一批 5880 个锁具两两比较,总的比较依次数为  $C_{5880}^2$ ,其中能互开的次数可通过编程让计算机统计出

两两比较次数总和为  $TOTAL = C_{5880}^2$

两两比较中能互开的次数和为  $HK = 22778$

互开率为  $P_{\text{总}} = \frac{HK}{TOTAL} = \frac{22778}{C_{5880}^2} = 0.132\%$

通过  $P_{\text{总}}$  我们可看出这批锁具的互开程度,并可衡量一批锁具的技术指标.若  $P_{\text{总}}$  越小,互开现象就越小,锁具的安全性越好.

## 2. 团体顾客购买若干箱时互开概率

### (1) 根据总体概率分析估算

若抽取 1 箱,里面 60 个锁具两两比较次数为  $C_{60}^2$ .若认为每箱的互开率与一批总体互开率  $P_{\text{总}}$  相同,则互开次数的,数学期望  $E_1 = C_{60}^2 \cdot P = 2.34$ .

设在  $C_{60}^2$  次独立重复抽取事件中,互开次数  $\xi \sim B(C_{60}^2, P)$ ,  $P(i)$  代表在  $C_{60}^2 = 1770$  次比较中出现  $i$  对互开现象的概率.

由二项分布的基本原理得:

$$P(K) = P(\xi = K) = C_{1770}^K P^K (1 - P)^{1770 - K}$$

$$(K = 0, 1, 2, \dots, 1770, P = 0.132\%)$$

上述分布律亦可用下表表示:

K	0	1	2	3	4	5
P(K)	0.096	0.226	0.264	0.205	0.120	0.056
K	6	7	8	9	$\geq 10$	
P(K)	0.022	0.007	0.002	0.001	$\approx 0$	

从上表我们可以看出,一箱内不存在互开的可能性是很小的,仅 0.096,面出现两次互



开的概率最大.可见互开现象较严重,所以采用新的装箱、标志方法非常重要.

同理抽取二箱时,互开次数  $\xi \sim B(C_{120}^2, P)$

$$E_2 = C_{120}^2 \cdot P = 9.42$$

$$P = (\xi = K) = P(K) = C_{7140}^K P^K (1 - P)^K$$

作出的表格为:

K	0	1	2	3	4	5	6
P(K)	8E-5	0.00076	0.0036	0.011	0.026	0.050	0.078
K	7	8	9	10	11	12	13
P(K)	0.106	0.125	0.131	0.123	0.105	0.083	0.060
K	14	15	16	17	18	19	$\geq 20$
P(K)	0.041	0.025	0.015	0.0083	0.0043	0.0029	0.0021

对一般情况抽取  $n$  箱时,互开次数  $\xi \sim B(C_{60 \cdot n}^2, P)$

## (2) 随机抽样模拟

i) 在实际生活中,一般顾客在大多数情况下,往往采取从已买的若干箱锁具中抽取若干个锁具,检验统计后进行估计的方法,由此行出互开程度的大小.我们认为此种方法虽然在理论上是合理的,但实际中的检验方法不可能满足估计量的评选标准,因为,即使取一箱,在其中任取一对检验,也有  $C_{60}^2 = 1770$  种可能,若取两箱,  $C_{120}^2$  就更大,在实际中不可能.做到与之可类比的检验次数,而取到任意一箱的种数也有  $C_{5880}^{60}$  种,况且不同的人在一批锁具中的不同箱子中,可能按照不同抽取方法,不同的抽取次数来抽样检验,故一般顾客采取的抽样方法在估计中的置信程度、偏差程度、有效性、相合性等方面存在很大欠缺,为此我们采取计算机随机抽样方法来模拟互开的概率情况.

ii) 利用计算机模拟随机抽样.对此我们有多种方法:蒙特卡洛模拟中的迭代取中法移位法、同余法.但考虑问题的实际情况,我们采用了线性同余法来模拟(见程序(略)).

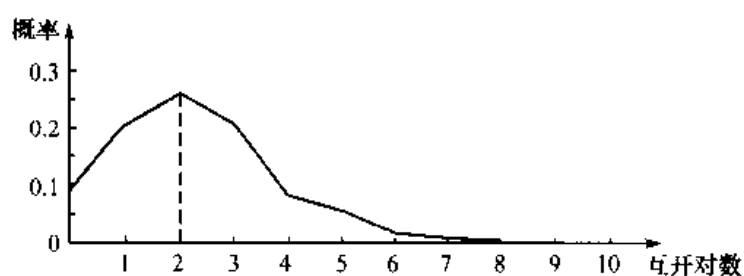
我们基本方法如下:

从一批总体 5880 个中随机抽取 60 个组成一箱,箱内 60 个锁具两两比较,互开次数记为  $N$ ,  $N$  的取值  $0, 1, 2, 3, \dots, C_{60}^2$ .

通过计算大量产生随机数,模拟抽取,并进行比较,对  $N$  值取值为  $0, 1, 2, 3, \dots$  的次数加以分类统计,记为  $M_i (i = 1, 2, 3, \dots)$ ,以  $P(i) = \frac{M_i}{M}$  代表一箱内出现  $i$  对的概率,则根据强大数定律得:当随机抽样次数  $M$  充分大时,  $P(i)$  趋向一批总体概率分布.

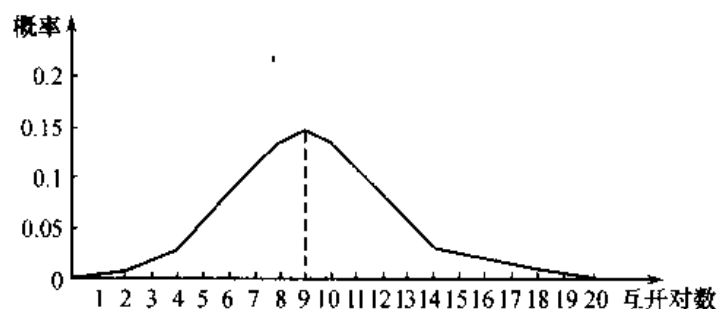
我们利用计算机模拟顾客购买一箱和两箱锁具的互开情况,计算结果如下表和图.

一箱	$K_0$	$K_1$	$K_2$	$K_3$	$K_4$	$K_5$
次数	991	2300	2689	2173	1100	527
互开率	0.0991	0.2300	0.2689	0.2173	0.1100	0.0527
二箱	$K_6$	$K_7$	$K_8$	$K_9$	$K \geq 10$	
次数	82	27	11	1	0	
互开率	0.0082	0.0027	0.0110	0.0001	0	



从图表可以看出,抽取一箱时互开时数为 2 的概率最大

二箱	$K_0$	$K_1$	$K_2$	$K_3$	$K_4$	$K_5$	$K_6$
次数		8	33	112	257	508	771
互开率	$8E-5$	0.0008	0.0033	0.0112	0.0257	0.0508	0.0771
二箱	$K_7$	$K_8$	$K_9$	$K_{10}$	$K_{11}$	$K_{12}$	$K_{13}$
次数	1053	1254	1325	1212	1053	826	587
互开率	0.1053	0.1254	0.1325	0.1212	0.1053	0.0826	0.0587
二箱	$K_{14}$	$K_{15}$	$K_{16}$	$K_{17}$	$K_{18}$	$K_{19}$	$K \geq 20$
次数	424	236	157	87	45	29	23
互开率	0.0424	0.0236	0.0157	0.0087	0.0045	0.0029	0.0005



从图可以看出,抽取二箱时,互开对数为 9 的概率最大

抱怨函数的构造:

抱怨函数的构造建立在以下假设的基础上.

1. 假设团体顾客在购买时,至多一次购买  $G$  箱 ( $G \leq 98$ ),且顾客从所购的  $G$  箱中随机抽取检验的箱数占其购买箱数的百分率  $g$  与他所购买的箱数成二次关系.

即有:  $g = -aG^2 + bG + c$  (其中  $a, b, c$  为常数)

由于箱数必须为整数,故我们认为顾客抽取的检验指数符合  $J = \text{INT}(g \times G + 0.5)$  的规律.

这是因为考虑到团体顾客购买后不可能象单个顾客购买几把或十几把可以完全检验所购锁的互开次数.

例如:若想检验一箱锁中有几次互开,需检验  $C_{60}^2 = 1770$  次.

若检验二箱,则需  $C_{120}^2 = 7140$  次

故此团体顾客购买后只能从中抽取一定数量的箱数来检验,且考虑到顾客的心理,当他购买箱数较少时,自然检验率较高,能发现几乎所有的互开情况,当购买数量大时,相对检验

率会减少,有的互开情况发现不了.

2. 由于当团体顾客购买量较大(通常为几十箱)时,他也不可能对抽出箱内的锁作完全互开试验,因此,我们又假设顾客在对所抽取的箱中的锁作检验时,从每箱当中抽取相同数量的锁具  $n$  把,然后把  $n$  把锁具混合起来作完全互开试验,这样检验的工作量不大,而又达到了检验目的.例如,从 40 箱里抽出 3 箱检验,再从 3 箱的每箱中又抽出 10 把锁具混合起来,对这 30 把锁具作完全检验.

显然,顾客的抱怨程度最主要一方面取决于他所购买的总数量,另一方面取决于他作检验的结果(即互开次数).并且从心理学的解度考虑,顾客更偏重于检验的结果.由此我们认为顾客的抱怨程度与购买的总箱数成反比,与检验结果成  $e$  的指数关系

由此,就可以定义出抱怨函数的表示为

$$C(T, \delta_n) = \frac{K_1}{T} e^{K_2 \delta_n},$$

其中  $K_1, K_2$  为常数,其实际意义为:

$K_1$ : 表示顾客购买箱数在整个抱怨程度中所占的比重

$K_2$ : 表示检验到互开次数在整个抱怨程序中所点的比重.

其确定可分别依实际的用户而确定.

$T$ : 表示顾客购买的总箱数.

$\delta_n$ : 表示顾客检验到有  $n$  次互开的互开率.

即  $\delta_n$  等于  $J \times n$  把锁作完全互开试验后得到的  $n$  次互开比上  $J \times n$  把锁的所有组合

$$\text{即 } \delta_n = \frac{n \text{ 次互开}}{C_{J \times n}^2} \times 100\%$$

下面我们针对在一批锁具中购一、二箱的情况作具体分析.

我们用计算机进行 1000 次的模拟数据检验(程序见附录(略)),结果见下表.

互开次数 $n$	0	1	2	3	4	5	6	$\geq 7$
概率 $P_n(\%)$	13.7	26.9	28.6	17.9	8.7	2.9	0.9	0

$n$ : 一箱中有  $n$  次互开;  $P_n$ : 1000 次检验下一箱中有  $n$  次互开的概率.

我们认为:

① 如果甲顾客购买箱数是乙顾客的两倍,而甲、乙检验到相同的  $\delta_n$ ,则甲的抱怨程度是乙的两倍;

② 如果仅购买一箱,由互开数大于 7 的为 0 次,故检验到的互开率为  $\frac{6}{C_{10}^2} = \frac{2}{15}$  时抱怨达到极值设为 100.

由 ①, 得到  $K_1 = 1$ ,

由 ②, 得到  $K_2 = \frac{15}{2} \ln 100$

则得到抱怨函数的具体表达式为  $C(T, \delta) = \frac{1}{T} e^{\delta \cdot \frac{15}{2} \ln 100}$ .

对购买一箱者：
$$\delta_n = \frac{n}{C_{10}^2}$$

为简便起见,我们假设顾客多次重复做了这样的试验,并对各种互开情况顾客均能检验到,则平均互开率:

$$\begin{aligned}\delta_1 &= \sum_{n=1}^6 \delta_n \cdot P_n \\ &= \frac{1}{C_{10}^2} (1 \times 0.269 + 2 \times 0.286 + 3 \times 0.179 + 4 \times 0.087 \\ &\quad + 5 \times 0.029 + 6 \times 0.009) = 0.04\end{aligned}$$

对于二箱的情况:

$$\begin{aligned}\delta'_n &= \frac{n}{C_{20}^2} \\ \delta_2 &= \sum_{n=1}^6 \delta'_n P_n \\ &= \frac{1}{C_{10}^2} (1 \times 0.269 + 2 \times 0.286 + 3 \times 0.179 + 4 \times 0.087 \\ &\quad + 5 \times 0.029 + 6 \times 0.009) = 0.01\end{aligned}$$

则购买一箱的顾客的平均抱怨程度为:

$$C(1, \bar{\delta}_1) = e^{\bar{\delta}_1 \cdot \frac{15}{2} \ln 100} = e^{0.04 \times \frac{15}{2} \ln 100} = 3.98$$

则买二箱的顾客的平均抱怨程度为:

$$\bar{C}(2, \bar{\delta}_2) = \frac{1}{2} e^{\bar{\delta}_2 \cdot \frac{15}{2} \ln 100} = 0.5 \times e^{0.01 \times \frac{15}{2} \ln 100} = 0.71$$

对结果的解释:

由上述结果可以看出,购买一箱锁具的顾客的平均抱怨率明显高于购买两箱锁具的顾客的平均抱怨率,这是与我们的一贯看法所吻合的.因为虽然在个别情况下购买两箱锁具的顾客的抱怨率会高于购买一箱锁具的顾客的抱怨率,但顾客购买的数量越少,就越容易检验出其中的互开情况,而如果购买的数量较大,相对就不易检验出互开的情况,并且由于购买的总箱数在一定程度上起了平衡顾客心理的作用,故上述的结果在数值上差别就不难理解了.

# 关于顾客抱怨程度的定量描述

喻甫祥 薛剑耿 林 东

(中国科学技术大学 230026)

指导教师 王树禾 徐俊明

**编者按** 本文在讨论锁具随机装箱情况(即原来的装箱方法)之下,顾客不满意程度的描述时,首先引入了三个特征量:互开的“锁具对”的数目  $x$ ,能被某把其他锁具打开的锁具的个数  $y$ ,必须报废的锁具的最小数目  $a$ .通过计算机模拟求出了  $x$  和  $y$  的期望值.作者们在讨论一把锁被  $n$  把其他锁具打开的问题中,证明了  $n \geq 3$  的情况出现的概率很小,可以忽略.在此基础上给出了抱怨程度的一个描述,并给出了改进的描述方法.虽然其中的一些问题还有待于进一步探讨,仍不失为一种富有特色的考虑方法.

## 1. 分析和假设

当锁具随机装箱后出售给团体顾客时,可能由于某些锁具互开而引起抱怨.我们设计概率模型来定量地衡量团体顾客抱怨互开的程度.模型应该较好地反映出大多数时候团体顾客抱怨互开的程度,而又能兼顾到极端情况,尽量使团体顾客抱怨程度极大而抱怨函数却极小的情况很少出现.

假设团体顾客发现购买的锁具中有互开的情况时,会更换一些锁具以避免互开,我们称这些被更换的锁具被报废.我们引入三个特征量:互开锁的“锁具对”的数目  $x$ ,能被某把其它锁具打开的锁具的个数  $y$ ,必须报废的锁具的最小数目  $a$ .如果以每个锁具为顶点,能互开的锁具之间用一条边相连,构造出一个图  $G$ ,那么  $x$  就是  $G$  的边数, $y$  就是  $G$  的非孤立顶点数, $a$  就是  $G$  的点覆盖数.

另外,我们用五位数来代表锁具,各位数字分别与相应槽的高度相对应,按五位数的数字和的奇偶性,把锁具分为奇锁具和偶锁具两大类,容易证明,一批具有 5880 个,其中奇、偶锁具各有 2940 个;不同的奇锁具之间不互开,不同的偶锁具之间也不互开,因此  $G$  是一个二分图.

## 2. 模型的设计

抱怨程度与互开的锁具对数  $x$ ,能被其它某把锁具打开的锁具的个数  $y$  以及必须报废的锁具的最小数目  $a$  有关.

理论推导可以得到  $x$  的期望值,在一箱中,由于边的总数为 22778,可以得到  $x$  的期望值为  $E_x = \frac{22778}{C_{5880}^2} \times C_{60}^2 = 2.33$ .通过计算机模拟的方法可以得到  $x, y$  的分布和数学期望.如当团体顾客的购买量为一箱时, $x$  和  $y$  的数学期望分别为 2.33 和 4.43;两箱时,则分别为 9.40 和 17.39.

称出现一个锁具能与其它两个锁具互开的情况为一个二分叉,能与其它三个锁具互开的情况为三分叉,依次类推.在一箱锁具中,出现一个锁具被(至少三个)其它锁具打开的概

率  $p_0$  是相当小的,证明如下:(设一箱中一个给定锁具被至少三个其它锁具打开的概率为  $p_1$ )

$$P_1 \leq \max \left\{ \sum_{l_1=3}^l \frac{\binom{l}{l_1} \binom{5879-l}{59-l_1}}{\binom{5879}{59}}; l \leq 10 \right\} \leq 0.0001$$

$$\therefore p_0 \leq 60 p_1 \leq 0.006$$

所以这种情况可以忽略不计.

在剩余的情况中,我们用函数  $2x + y$  来衡量顾客抱怨的程度.若四个锁具能分为两对,每对中两个锁具可以互开,不同对中的两个锁具不能互开,我们称之为情况 A;同时我们用情况 B 来表示三个锁具形成一个二分叉.显然,情况 A 和情况 B 都恰有两对互开的锁具,但是情况 A 可能需要报废两个锁具,情况 B 一般只需要报废一个锁具.所以我们可以说情况 A 比情况 B 更坏一些,在排除了三分叉和更多分叉的情况后,我们可以看出如果有一个锁具能被另两个不同的锁具打开(即出现情况 B),则  $2x + y$  小于将这二把锁及另外某把锁 A 的情况且不影响其它锁具的互开关系时它的值,而且可以得出有  $n_1 = 2x - y$  个锁具能被两个不同的锁具打开,有  $n_2 = 2(y - x)$  个锁具能被恰好一个其它的锁具打开, ( $n_1$  和  $n_2$  满足  $2n_1 + n_2 = 2x$ ,  $n_1 + n_2 = y$ ) 其它的锁具都不会被箱中任意别的锁具打开.这样当每增加一对锁具能互相打开而不影响其它锁具的互开关系时  $2x + y$  将增大.可见,  $2x + y$  对 A 和 B 这两种情况给予了恰当的区分.由此可知  $2x + y$  在绝大多数情况下比较好地描述了顾客的不满意程度.根据计算机模拟的结果可得到  $2x + y$  对 A 和 B 这两种情况给予了恰当的区分.由此可知  $2x + y$  在绝大多数情况下比较好地描述了顾客的不满意程度.根据计算机模拟的结果可得到  $2x + y$  的数学期望值是

$$E(2x + y) = 2Ex + Ey = 2 \times 2.33 + 4.43 = 9.09.$$

另外当箱数为二时,计算可得存在一个锁具被另外四个锁具打开的概率小于百分之二,所以这种情况也可以忽略不计.这时  $2x + y$  仍能比较好地反映顾客的不满意程度.此时的数学期望值为

$$E(2x + y) = 2Ex + Ey = 2 \times 9.40 + 17.39 = 36.19.$$

为了更精确地描述顾客的不满意程度,我们引进新的参数  $y_1$  和  $y_2$ .  $y_1$  为能被箱中某个另外的锁具的个数,  $y_2$  为能被箱中某个另外的锁具打开的偶锁具的个数.我们用  $u = y_1 y_2$  表征顾客不满意程度.由一条变时,顾客的不满意程度也变大.容易看出  $u$  很好的反映了这种要求.另外我们可以考虑另一个参数  $a$  为最少要报废的锁具数目(即顾客使用其它的锁具可以做到任意两把锁具不可以互开).从经济的角度来看,  $a$  十分精确地反映了顾客的不满意程度.而且可以看出  $a$  大致为  $\min(y_1, y_2)$ , 由于  $y$  的分布已由计算机模拟得出,而且  $x$  的分布用 poisson 分布( $\lambda = 2.33$ ) 较好地逼近,所以  $u$  中包含了相当多的  $a$  信息量.由于  $a$  的分布很难用数学方法或计算机模拟得到,而  $u$  的分布在箱数不多时用计算机可以很快得到,所以我们采用  $u$  作为表征顾客不满意程度的量.

下面估算  $u$  中包含  $a$  的信息量  $I(u, a)$  与  $a$  的熵值  $H(a)$  的比值(即  $u$  反映  $a$  的程度):当箱数为二时,由计算机模拟得到的  $c = \min(y_1, y_2)$  的分布( $l = 1000$ ) 知:

$$P\{a = 0\} = P(u = 0) = 0.102$$

$$0.224 = P\{u = 1\} \leq P\{a = 1\} \leq P\{u = 1\} + P\{u = 2\} + 0.01 = 0.274$$

$$0.614 \leq P\{a \geq 2\} \leq 0.664$$

所以

$$H(a) \geq 0.84$$

$$\begin{aligned} \frac{H_u(a)}{H(a)} &= \sum_{k=0}^{900} P\{u = k\} \frac{H_{u-k}(a)}{H(a)} \\ &\leq -0.0061 \ln 0.006 - 0.0491 \ln 0.049 \leq 0.15 \end{aligned}$$

$$(1) \quad \frac{I(u, a)}{H(a)} \geq 0.85, I(u, a) \geq 0.7$$

注:分三种情况

A) 存在一把锁可以被与它不同的三把锁打开.

B)  $0 \leq u \leq 20$ , 且不在 A) 中.

C)  $21 \leq u \leq 900$ , 且不在 A) 中.

由此很容易看出熵值估计式的正确性.

由(1)式知道如果给定  $u$ , 则在大约百分之八十五以上的程度上决定了  $a$ , 所以我们取  $u$  作为表征量是十分合理的.

### 3. 模型的评价

我们的模型对下述情况没有给出很好的描述: 抱急函数较大但团体顾客因为实际需要报废的锁具数并不多而不十分抱怨. 这是一种“好”的情况(销售部门担心的是对团体顾客抱怨互开的程度估计不足), 这类情况里较多的出现二分叉、三分叉等, 前面已指出这类情况发生的概率很小. 发生概率较大的是接近极坏的情况, 二分叉、三分叉等出现较少, 需要报废的锁具数目与  $x$  十分接近, 此时  $y \approx 2x$ ,  $2x + y \approx 4x$ ,  $\sqrt{y_1 y_2} \approx x$ , 可见两种抱怨函数在多数情况下是一致的.

在以上的讨论中, 我们认为两个锁具只要满足互开时锁具对应槽的高度关系就能互开的概率  $p = 1$ , 事实上并不能保证它们一定能互开, 即应有  $0 < p < 1$ , 若某个团体顾客购买的锁具中存在  $n$  对满足互开时锁具对应槽的高度关系, 如果认为每一对都有概率  $p$  的可能互开, 那么实际能互开的锁具的对数(数学期望) 大约为  $pn$ . 因此对生产厂家来说, 通过提高工艺水平来降低  $p$ , 可以降低团体顾客抱怨互开的程度.

# 关于“锁具装箱”问题的评阅工作

中国科学院系统科学研究所

项可凤

“锁具装箱”问题(B题)送交评阅的试卷共 58 份,由于经过赛区筛选,卷子质量普遍较好,绝大多数卷子都达到或超过组委会所提供的参考答案的要求.根据这样情况,我们评阅小组认为凡是未达到参考答案要求,或文字表达极差的卷子,经评卷小组成员半数以上认定,评卷组长确认,按未达到标准而被淘汰.这样,在阅卷第一阶段,就淘汰了 17 份卷子.评卷的第二阶段要在余下的 41 份卷子中,评选出全国一等奖.对于一等奖,我们认为其内容上必须超过参考答案要求,特别在下述几个方面,至少要在二、三个问题上有所创见或特色:

## 1)不能互开锁具的最大数的论证

这是一个较难的理论问题.参考答案只要求按槽高之和的奇偶分成二类,各分装 49 箱, 2940 把锁不能互开.但 2940 是否为不能互开的最大数呢?有少量卷子用图论方法给出讨论或给出算法.我们认为,能在这个问题上给出论证本身就反映学生思考问题的严密性和解决理论问题的能力.

## 2)关于装箱销售问题

在这个问题上,大多数卷子仅仅讨论了按槽高之和的奇偶标记分箱,可以保证团体顾客购买不超过 49 箱时,不会出现互开问题.我们认为仅讨论到这一步还不够.在门市销售中,实际的顾客流是序贯的.有少量卷子,讨论了序贯销售时,装箱的标记方法,按此方法,在连续销售中,可以保证任意购买不超过 42 箱均不会出现互开问题,这样的讨论就更深入和符合实际了.

## 3)关于顾客抱怨程度

这个问题灵活性较大,没有固定解答,学生有较大发挥余地,答案也是形形色色的.但大部分卷子都用平均互开总对数,或它的单调增函数去定义抱怨程度.因此,购买箱数越多,抱怨程度越大.我们认为这样定义抱怨函数有一定不合理性.因为购买越多,自然互开概率越大,这是顾客所意料中的,不应有太多的抱怨.顾客所不能容忍的是在购买少量锁具而出现互开现象.因此,我们更欣赏把购买箱数作为一个因素考虑到抱怨函数中.理想的抱怨函数应该是,开始随购买量增加而增加,到一定量后下降,这才更合情理.

4)在参赛的卷子中,普遍存在不重视文字表述问题.我们认为,入选一等奖的卷子,必须文字叙述清楚、简练、流畅.

根据上述 4 条原则,经过评卷小组成员反复地横向比较,一致同意确定 15 份优秀卷子,作为全国一等奖提名.

最后,余下 26 份卷子都是至少有评卷小组成员半数以上同意入选的,自然就是全国二等奖.在二等奖中,成绩跨度较大,这是因为数学建模竞赛和一般竞赛不同,它没有标准答案,只能从横向比较中确定优劣.