

同济大学 2016 年数学建模竞赛论文  
本科生组  
B 题

参赛者信息:

姓名	学号	学院	专业	联系方式
----	----	----	----	------

# 游乐园客流疏导方案

## 摘要

对于目前游乐场所等公共场所越来越易出现人流大量汇集以至拥堵的现象，如何及时分流提供正确引导的重要性越发突出。针对此现象以某一具体游乐园为例，提出了两个问题。本文运用了模拟、优化、拟合等算法模型成功解决了这两个问题。并得到人流分布、排队时间、优化措施、预定数量的具体算法流程。

针对问题一，首先确定影响游客选择不同游乐项目的因素，接着根据这些因素利用选择行为模型求出游客在不同项目间的选择概率，以 **Markov** 过程的算法得到自然状态下人流的分布。然后利用遗传算法的思想确定引导方式，利用 **Matlab** 编程模拟遗传算法的迭代过程，得出优化模型，分析该模型的有效性。

针对问题二，首先利用历史数据统计分析影响因素与每月入住房间数量的关系，建立相关线性方程组，并进行误差分析。然后利用最小二乘算法拟合入住数量与预定数量的关系，得到每月预定数量的预测值，参照影响因素的权重确定每天预定数量。

最后对所建立的模型的算法进行客观评价，并提出相关改进方法。

关键词： 人流分布 行为选择模型 **Markov** 过程 统计分析 最小二乘拟合

## 一、问题重述

### 1.1 问题引言

本市建有最多过山车的游乐园 Youth 游乐园即将开业。其受到青少年的热捧，预计届时园区将迎来每天 1 万人的最大客流量。现要根据具体客流情况，及时分流人群，为顾客提供游园线路引导，保障游客的最佳游园体验。使 Youth 游乐园稳定且良好运行。

### 1.2 问题的提出

围绕园区总体建设布局、游乐项目分布、游客入住安排等方面，本文提出以下问题：

(1) 根据附件中 Youth 游乐园的规划图，游乐园中建设了 A-J 共 10 个项目点，游客可以按照附件中标注的线路往返下一个游乐项目。在保障每位游客较好体验游乐设施的前提下，要求建立适合的数学模型对每个游乐项目的等候游客进行游览提醒和及时疏导，以达到最佳游园体验。同时附表中给出了 A-J 每个游乐项目的具体安排参数。

(2) 皇冠假日酒店是 Youth 游乐园内已经开业的酒店，为有需要的游客提供住宿的便利。根据附件中该酒店的历史预定数据信息，综合考虑季节、周末、法定假日、暑假等影响房间预定量的因素建立数学模型，预测 2016 年 1 月至 3 月每天的预定房间数量。

## 二、模型假设

(1) 每位游客会综合考虑各个游乐项目对自身的效用，总会选择对自己效用最大的游览方式。

(2) 对于每个游乐项目游客进入项目场地的时间忽略不计。

(3) 大部分游客会听从游乐园工作人员及相关标识的疏导和指引。

(4) 每个游乐项目是开放性的系统，并且会在一定时间后趋于稳定状态，进入和流出游客数量均衡。

(5) 鉴于此模型主要用于高峰人流期的指引和疏导，对于进入游乐园的游客流量考虑高峰流量时期，不考虑初始状态游客数量较少的时期。

## 三、符号说明

表 1 公式符号及含义对应表

$t = (t_1, t_2, \dots, t_{10})^T$	游乐园人数分布向量， $t_i$ 为第 $i$ 个游乐项目处人数
$U_{ij}$	MNL 模型中的个人效用函数
$P_{ij}$	在项目 $i$ 处前往游乐项目 $j$ 的概率
$D$	两游乐项目间的给定距离

$N$	两游乐项目是否相邻
$V$	游乐项目被参观的可能性
$T$	游乐项目每场持续时间比值
$C$	游乐项目容纳人数比值
$T_m$	游客最长等待时间
$f$	适应度函数
$P_i$	遗传算法中个体存活概率
$P_c$	遗传算法中交叉概率
$P_m$	遗传算法中变异概率
$s$	参观可能性控制的最小值
$i$	表示月份，取值 1-16 表示 2015 年一月到 2016 年 4 月
$x_1$ 、 $x_2$ 、 $x_3$	工作日、双休日、节假日每天的入住房间数
$y_1$ 、 $y_2$ 、 $y_3$ 、 $y_4$	由于暑假、夏季、秋季、冬季导致的每月入住房间数变化
$a_{i1}$ 、 $a_{i2}$ 、 $a_{i3}$	第 $i$ 月的工作日、双休日及节假日天数
$a_{i4}$ 、 $a_{i5}$ 、 $a_{i6}$ 、 $a_{i7}$	取 0-1 表示第 $i$ 月是否为暑假、夏季、秋季、冬季
$b_i$	表示第 $i$ 月的入住房间数
$c_i$	表示第 $i$ 月预订房间数
$l_1$ 、 $l_2$	月预订房间数与当月及下一月入住房间数的线性回归系数
$s_{i1}$	第 $i$ 月预订的当天入住的房间数
$s_{i2}$	第 $i$ 月预订的非当天入住的房间数
$M_i$	第 $i$ 月总天数
$r_{ij}$	第 $i$ 月某天对应的权重
$R_{ij}$	第 $i$ 月某天对应的预定房间数

## 四、问题分析

### 4.1 问题一问题分析

问题一需要建立对每个游乐项目的等候游客进行游览提醒和疏导的模型，以达到游客游园体验最优。首先，这是一个建立优化模型的问题，但在进行优化之前，可以对游乐园人群分布先进行一个大致的模拟，然后通过新模型观察优化程度，大致模拟可以应用 Markov 过程，最优解可以用迭代得到；其次进行游览提醒和疏导的操作，应该体现在疏导模型中；最后关于游客的体验，可以用一个或几个参量来进行评价，也是作为优化过程的控制条件和目标条件。

### 4.2 问题二问题分析

对于问题二，要求建立数学模型预测 2016 年 1 月至 3 月每天预定房间数。由于双休日、节假日、寒暑假及季节等因素将会直接影响每月入住房间数。根据生活经验知道，上述因素直接影响的是每月入住的数量而非预订数量。故而不能直接以每月的影响因素来预测当月的预定房间数量。

因此首先求出 2015 年 1 至 12 月每月实际入住房间数与各月份影响因素关系的参数值。假定这些因素与最后结果呈线性关系，设定系数，并利用最小二乘法结合其中 8 个月的数据求出系数。利用得到的实际入住数量与影响因素间的关系对另外 4 个月的数据做出预测，比较预测数据与真实值的偏差。反复调整方案直至预测数据与真实数据偏差较小。然后，具体分析数据得到各月入住数量与预订数量之间的关系。最后预测 2016 年 1 月到 4 月的入住数据，从而借助其得到 2016 年 1 月至 3 月的预订数据。最后，分析当天预订当天入住的比例，结合入住数量与各因素之间的关系得到每天预订数量的预测值。

## 五、模型的建立与求解

### 5.1 问题一模型建立与求解

Youth 游乐园中设 A-J 共 10 个项目点，游客可沿着一定的线路往返于游乐项目间。在保障每位游客体验游乐设施的前提下，以下建立对等候游客进行游览提醒和疏导的模型，以达到游客游园体验最优。

首先，在实际情况下，游乐园在开园后，游客陆续进入园区。游乐园内总体游客数随时间变化基本上是先迅速增加，再维持动态平衡，最后人数迅速减少为零。而需要对游客进行疏导的时间就是在人数达到高峰段时。于是在问题中可以假设游客进出游乐园达到动态平衡，也即可以认为总人数基本不变。由于问题中游乐园只有一个出入口，在前面的假设下，可以认为往返于出入口到 A 和出入口到 B 的游客是在 A 和 B 间移动，于是，该问题可以转化为一个封闭系统下的人群流动模型。

当游客在游乐园里游玩时，会前往自己感兴趣的游乐项目，但由于不同项目容纳人数有限和持续时间不同，游客会在一些项目前等候起来，而过长等待时间对游园体验是不太好的。同时由于人数的累积，及时的人群疏导也是很有必要的。

下面将问题分为两大部分：第一部分研究自然状态下人群的分布和游客最长等待时间；第二部分讨论在加入疏导条件后，人群分布和等待时长的改变。

### 5.1.1 自然状态下的人群

根据假设，游客在游乐园内游玩时选择最短的线路，而且在路程上的时间不予考虑，于是人群的主要信息便集中在 10 个游乐项目上。记每个游乐项目处人数为  $t_i$ ，可得到整个游乐园人数分布向量  $t = (t_1, t_2, \dots, t_{10})^T$ ，其中 A-J 按 1-10 排序。根据假设可知， $\sum t_i$  为一固定值。

为研究自然状态下人数的平衡状态，可以这是认为是一个 Markov 过程。为了和实际情况更加接近，即认为游客倾向于向人数较少的地方走，在迭代过程中状态转移矩阵会发生改变，即游客前往各个项目的概率会发生变化。

#### 5.1.1.1 状态转移矩阵

为了得到状态转移矩阵，就需要得到游客前往各个游乐项目的概率。下面使用 MNL (Multinomial Logit model) 模型来得到游客在项目  $i$  处前往游乐项目  $j$  的概率。

MNL 模型是一种选择行为模型，它建立在效用最大化的基础上，即假说认为做出决断的个体或集体，在相互独立的选择项目中会选择对自己效用最大的选项。而在游乐项目  $i$  处选择项目  $j$  的概率为：

$$P_{ij} = \frac{\exp(U_{ij})}{\sum_j \exp(U_{ij})}$$

其中  $U_{ij}$  是个人效用函数，是根据各种对选项的影响因素计算出的。

在此问题中，我们设定的影响因素为：

(1) 两点间距离  $D$ （单位：米），在效用函数计算中取对数，一是平衡量纲，二是使得该点到该点本身的选择概率为 0，即游客在游玩项目后不会再次选择此项目；

(2) 两点是否相邻  $N$ ，取值为 0 和 1，这一项与游乐园的线路分布有关，同时也的确关系到游客的选择；

(3) 参观可能性  $V$ ，将游乐项目分为 1 到 10 共十个等级，等级越高，参观可能性越高，而为了便于带入计算取值为相互间的比值；

(4) 持续时间比  $T$ ，即各个项目持续时间的比值；

(5) 容纳人数比  $C$ ，即各个项目容纳人数的比值。

在计算中，定义个人效用函数为：

$$U_{ij} = \alpha \cdot \ln(D_{ij}) + \beta \cdot N_{ij} + \gamma \cdot V_{ij} + \varepsilon \cdot T_{ij} + \delta \cdot C_{ij}$$

#### 5.1.1.2 个人效用函数系数的确定

在 MNL 模型中一般是使用统计数据计算得出系数的估计值，由于数据不充

足，在这里，我们使用层次分析法计算相关系数的值。

设定目标层为个人效用函数，准则层为路程，是否相邻，参观可能性，持续时间比，容纳人数比，如下图所示：

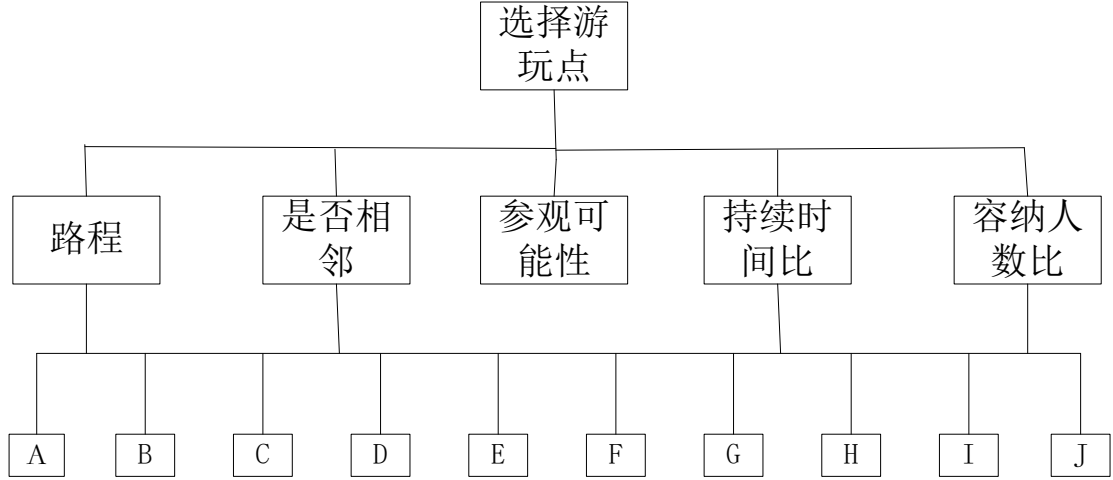


图 1 选择游乐项目的层次结构

因为各个影响因素的相对影响力不同，此处采用 1 到 9 等级的划分，等级越高，相对影响越大。若设准则  $i$  对  $j$  的影响为  $a_{ij}$ ， $j$  对  $k$  的影响为  $a_{jk}$ ，则  $j$  对  $i$  的影

响为  $\frac{1}{a_{ij}}$ ， $i$  对  $k$  的影响为  $a_{ik} = a_{ij} \cdot a_{jk}$ ，于是可以得到层次分析矩阵。由于准则数目

较多，先进行一对一的影响判断，后对一致性进行检验。

在此问题中，层次分析矩阵为：

$$A = \begin{pmatrix} 1 & 5 & \frac{1}{3} & 3 & 7 \\ \frac{1}{5} & 1 & \frac{1}{5} & 3 & 3 \\ 3 & 5 & 1 & 7 & 7 \\ \frac{1}{3} & \frac{1}{5} & \frac{1}{7} & 1 & 3 \\ \frac{1}{7} & \frac{1}{3} & \frac{1}{7} & \frac{1}{3} & 1 \end{pmatrix}$$

计算其最大特征值  $\lambda$  和从属于此特征值的特征向量  $v$ ，由一致性检验方法计算一致性指标：

$$CI = \frac{\lambda - n}{n - 1}$$

取随机一致性指标  $RI = 1.12$ ，则一致性比率  $CR = \frac{CI}{RI} = 0.0858 < 0.1$ ，通过了一致性检验。而将特征向量  $v$  归一化即为准则层对应系数。

### 5.1.1.3 自然状态下的平衡

通过层次分析法得到个人效用函数的系数后，便可由 MNL 模型得出状态转

移矩阵。为了实现状态转移矩阵的变化，可以通过改变参观可能性来改变概率。而参观可能性的改变取决于当前的人数分布。由于人数越多，参观可能性越低，于是可以设参观可能性 (%) = 1 - 人数所占比例 (%)。于是在每次迭代中，由得出的人数比改变参观可能性，从而改变概率和状态转移矩阵。

通过编程实现这一迭代过程，控制条件为当各个项目处人数变化少于 1 人，由每天总的人流量为 1 万人，这里不妨取总人数为 1 万人，因为在收敛状态下，结果与初始值关系不大，不妨取各个游乐项目的初始人数分布为：

$t = (10^4, 0, 0, 0, 0, 0, 0, 0, 0, 0)^T$ ，列向量  $t$  共 10 个元素，每个元素分别表示 A-J 游乐

项目的人数，同时定义最长等待时间  $T_m$

$$T_m = \max_i \left\{ \frac{t_i \cdot T_i}{C_i} \right\}$$

其中  $t_i$  为第  $i$  个项目前人数， $T_i$  为项目  $i$  持续时间， $C_i$  为项目  $i$  容量。

计算得到人数分布向量  $t = (3458, 611, 675, 790, 710, 898, 754, 797, 682, 625)^T$ ，最长等待时间为  $T_m = 285.35$ （单位：分钟）。

可以看到，人群主要集中分布在游乐项目 A 处，人数超过总数的三分之一，而等待的最长时间超过 4 个小时，可见在自然状态下人群分布游客体验不理想。

### 5.1.2 加入疏导控制后的人群分布

在游乐园中每个游客所掌握的信息总是有限的，而管理方对整个园区拥有更多的信息，从而可以进行人为控制而使得人群分布达到最优。从自然分布下的计算结果可以看到，人群多分布在游乐项目 A 处，于是管理方可以劝说 A 处聚集的游客前往另外的景点，而其他地方的游客尽量避免前往 A 处。显然这一过程也是动态变化的，从而进行疏导的强度也可以不同。下面将使用遗传算法的想法进行动态模拟和求解等待时间最少的最优解。

#### 5.1.2.1 遗传算法的应用

遗传算法是模拟自然界生物进化的一种随机和自适应的算法，主要过程是将实际问题的求解参数转化为生物群体，然后进行迭代遗传，最后找出满足条件的群体，得出实际问题的最优解。流程图如下：



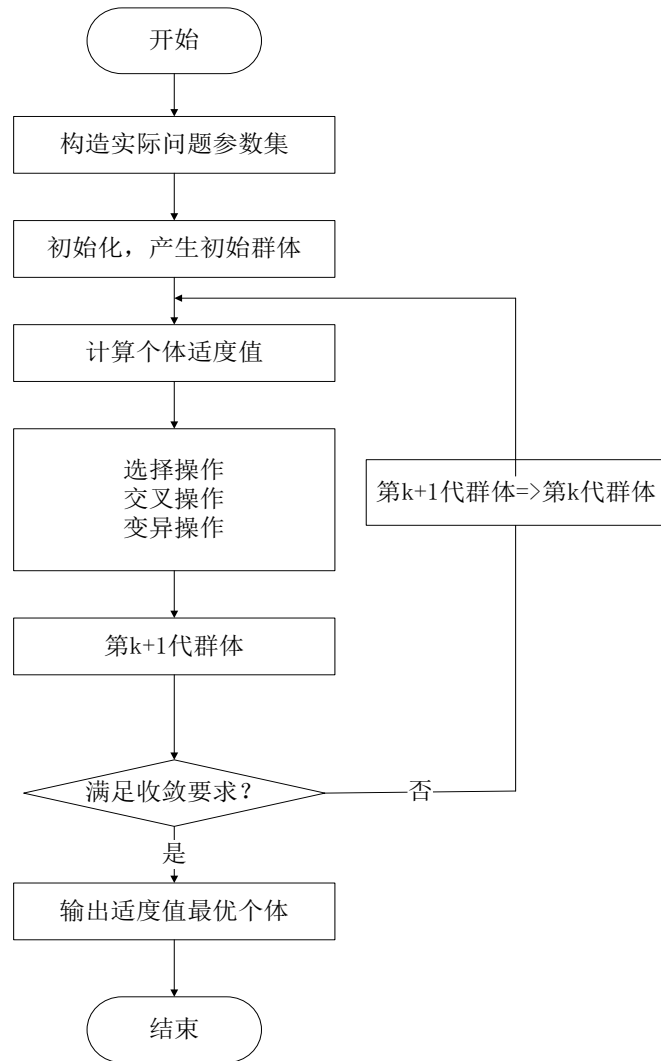


图 2 遗传算法流程图

在这个问题中，将参观可能性视为生物个体。而遗传算法一开始需要的是一个生物群体，一般初始的一组解是随机生成的。初始个体数目  $M$  越大，搜索范围越广，但效率会低，一般  $M$  取  $20 \sim 100$ ，这里计算取的是 20。

在遗传过程中，需要进行选择，交叉和变异三项操作。而进行选择过程需要一个适应度函数，其作用是判别个体的适应度，即类比生物在自然选择下是否能适者生存。此问题中适应度函数定义为最长等待时间的倒数，即

$$f = \frac{1}{T_m}$$

这样在第一代  $M=20$  个个体对应 20 个函数值，记为  $f_i$ ， $i=1,2,\dots,20$ ，从而

每个个体存活概率为  $P_i = \frac{f_i}{\sum f_i}$ ，选取存活概率最高的前 10 个个体，即等待时

间最短的 10 个参观可能性条件，作为下一代的产生者。

在选择过程中，仅仅对优秀个体进行选取，而不产生下一代。接下来在交叉过程中，先给定一个交叉概率  $P_c$ ，计算中取值为 0.8，然后任意选择两个个体进行基因交叉，此问题中选择了算数交叉法，设原个体为  $x_1$  和  $x_2$ ，新个体为  $x'_1$  和  $x'_2$ ，则算数交叉为

$$\begin{aligned}x'_1 &= \alpha \cdot x_1 + (1-\alpha)x_2 \\x'_2 &= (1-\alpha)x_1 + \alpha \cdot x_2\end{aligned}$$

其中  $\alpha$  为 (0,1) 上的随机数。

第三步操作是变异。其目的是为了增加找到最优解的可能性，维持群体多样性。而变异概率  $P_m$  一般取 0.0001~0.1，以防止把较优的个体去掉。此问题中取值为 0.1。变异操作按照下式进行

$$x' = x + 2(\alpha - 0.5)x_m$$

其中  $x'$  为变异后个体， $x$  为变异前个体， $x_m$  为变异操作中最大可能改变值， $\alpha$  为 (0,1) 上的随机数。

最后进行遗传迭代，进化代数一般取 100~500，同时设置收敛条件，一般是对适度值进行控制。此问题中设置的适度值条件为最长等待时长不超过 90 分钟。

#### 5.1.2.2 控制条件下求解

为了体现管理方对人群的疏散，在算法中体现为对参观可能性的强制性调整。由于参观可能性不能取值为 0，否则会对计算过程产生影响，所以可以取一较小值，此问题中，最小取为 0.1 来替代 0，在实际问题中反映为强烈建议游客不前往某处。而这一地点的选取即定为人分布最多的地点。

将这一控制条件加入选择中，进行遗传算法的模拟计算多次试验后，得出一个人数分布最佳结果为  $t = (2295, 643, 1461, 407, 1606, 2002, 361, 441, 354, 430)^T$ ，

等待最长时间为  $T_m = 151$ （单位：分钟）。相比自然状态下，A 点减少了一千多人，而最长等待时长约为两个半小时，相比四个小时有所减少。

根据人群分布  $T$  的变化，可知主要是将 A 处人群疏散至 C，E，F 处。

改变参观可能性控制的最小值  $s$ ，计算可得

$s = 0.05$ ， $T = (1866, 691, 855, 945, 1262, 1070, 872, 899, 778, 761)^T$ ，主要将 A 处人群疏散到 C，D，E，F，H，最长等待时间为 150 分钟；

$s=0.15$ ,  $T=(2113,332,834,665,2102,789,402,538,346,1879)^T$ , 主要将 A 处人群疏散到 C, E, J, 最长等待时间为 164 分钟。

结合游乐园地图可知, 当  $s$  取值越小, 分散地点越多且越远, 而  $s$  定义为参观可能性控制的最小值, 所以  $s$  越小, 即疏散力度越大, 这与模拟计算结果是一致的。

## 5.2 问题二模型建立与求解

### 5.2.1 预测每月入住数量与各因素的关系

首先统计 2015 年 1 到 12 月各月的预订房间数如下表:

表 2 每月预定房间数量分析

月份	预订房间数	当月预订当月入住房间数	入住房间数	入住的提前一月预订比例	预订的当月入住比例
1	46	10	10		
2	11	4	5		
3	320	167	188		
4	3065	2471	2601	5.00%	80.62%
5	2576	2188	2703	19.05%	84.94%
6	2561	2236	2733	18.19%	87.31%
7	2831	2525	2830	10.78%	89.19%
8	3033	2714	3001	9.56%	89.48%
9	3114	2679	2985	10.25%	86.03%
10	2987	2555	2984	14.38%	85.54%
11	2724	2482	2970	16.43%	91.12%
12	2657	2495	2738	8.88%	93.90%

观察到前 3 个月的预订数量显著地小于其他月份。猜测其原因为季节以及寒假因素或者是因为 Youth 游乐园开放时间。由统计数据得到 2015 年预订的于 2016 年 1 月入住的房间数为 177 间。据此与入住房间提前预订比例推知 2016 年一月入住数量与表中 2015 年 1 月预订和入住数量有显著差异。从而推断前三个月预订数量少于后几个月的原因是受皇冠假日酒店刚开业因素影响。从而可以主要分析 4-12 月的数据。

统计 2015 年 1 到 12 月的双休日、节假日、季节、是否为寒、暑假等信息如下表:

表 3 每月相关属性统计表

年份	月份	正常天数	双休日	节假日	是否暑假	季节
2015	1	21	7	3	0	4
2015	2	16	5	7	0	4
2015	3	23	8	0	0	1
2015	4	21	6	3	0	1
2015	5	20	8	3	0	1

2015	6	21	6	3	0	2
2015	7	23	8	0	1	2
2015	8	20	10	0	1	2
2015	9	22	6	3	0	3
2015	10	17	7	7	0	3
2015	11	22	8	0	0	3
2015	12	23	8	0	0	4
2016	1	20	8	3	0	4
2016	2	16	6	7	0	4
2016	3	23	8	0	0	1
2016	4	20	7	3	0	1

注：最后一列季节中 1、2、3、4 分别表示 春季、夏季、秋季、冬季。

设春季工作日、双休日、节假日每天的入住人数为  $x_1$ 、 $x_2$ 、 $x_3$ 。设暑假每月会多  $y_1$  人入住，夏季、秋季、冬季相比春季每月各多  $y_2$ 、 $y_3$ 、 $y_4$  人入住 ( $y_1$ 、 $y_2$ 、 $y_3$ 、 $y_4$  可以取负值)。记  $a_{i1}$ 、 $a_{i2}$ 、 $a_{i3}$  为第  $i$  月的工作日、双休日及节假日天数， $a_{i4}$ 、 $a_{i5}$ 、 $a_{i6}$ 、 $a_{i7}$  取 0-1 表示第  $i$  月是否属于暑假、夏季、秋季、冬季， $b_i$  ( $i$  取 1 到 12) 表示第  $i$  月的实际入住房间数。这样对于 4-12 月的数据一共可以得到 9 个线性方程。我们选取 4、5、6、7、9、11、12 月的 7 个方程试求其解，代入 8、9 月的数据加以检验后发现偏差较大。故而选取 4、5、6、7、8、9、11、12 月对应的 8 个方程。由于一共有 7 个变元，这通常是一个矛盾方程组，但可以求其最小二乘法意义下的近似解，方法如下：

$$\text{对于矛盾方程组} \begin{cases} a_{4,1}x_1 + \cdots a_{4,7}y_4 = b_4 \\ \cdots \cdots \cdots \\ a_{12,1}x_1 + \cdots a_{12,7}y_4 = b_{12} \end{cases}$$

考虑使方程：  $[(a_{11}x_1 + \cdots a_{17}y_4 - b_1)^2 + \cdots + (a_{81}x_1 + \cdots a_{87}y_4 - b_8)^2]$  实际值最小的解。由多元微分学知识知此时必有

$$\begin{cases} \frac{\partial [(a_{4,1}x_1 + \cdots a_{4,7}y_4 - b_4)^2 + \cdots + (a_{12,1}x_1 + \cdots a_{12,7}y_4 - b_{12})^2]}{\partial x_1} = 0 \\ \cdots \cdots \cdots \\ \frac{\partial [(a_{4,1}x_1 + \cdots a_{4,7}y_4 - b_4)^2 + \cdots + (a_{12,1}x_1 + \cdots a_{12,7}y_4 - b_{12})^2]}{\partial y_4} = 0 \end{cases}$$

这是一个由七个方程构成的七元线性方程组，代入数据求其解为：

$$\begin{cases} x_1 = 71.1 \\ x_2 = 147.7 \\ x_3 = 74.0 \\ y_1 = -74.5 \\ y_2 = 132 \\ y_3 = 268.6 \\ y_4 = -78.1 \end{cases}$$

代入九月的数据，求得这个模型预测的九月入住数量为 3029 次，与真实值的偏差为 1.47%。可见这个模型是可靠的。

因此可代入 2016 年 1-4 月的数据，求得预测的每月入住房间数，设为

$b_{13}$ 、 $b_{14}$ 、 $b_{15}$ 、 $b_{16}$ 。有

$$\begin{cases} b_{13} = 2747 \\ b_{14} = 2463 \\ b_{15} = 2817 \\ b_{16} = 2678 \end{cases}$$

#### 5.2.2 预测入住数量与预订数量的关系，求出 2016 年 1-3 月每天预订数量。

对于每月的预订数量，其包括当月入住的数量与今后入住的数量。由统计数据知对于 2015 全年的数据，预订时间与入住时间相隔一月（30 天）以上的占 4.1%，但是预订时间与入住时间跨月的只占 1.6%。因此我们可以假设每月预订的数量只与当月与下月的入住数量有关，并且呈线性关系。我们记 2015 年

1 月到 2016 年 3 月的每月预订房间数为  $c_i$ ， $i$  取 1 到 15。假设其呈现关系

$c_i = l_1 b_i + l_2 b_{i+1}$ 。为避免前 3 个月数据的影响，我们取 2015 年 5 到 10 月的数据，利用最小二乘的方法求参数的估计值为：

$$\begin{cases} l_1 = 0.5879 \\ l_2 = 0.3938 \end{cases}$$

借助 MATLAB 统计工具箱中的 regress 对其进行分析，发现  $l_1$ 、 $l_2$  对  $c_i$  的影响并不是很显著，可见这个模型是有一定误差的。

借助上面的公式可以得到 2016 年 1-3 月的每月预订数量为：

$$\begin{cases} c_{13} = 2585 \\ c_{14} = 2557 \\ c_{15} = 2711 \end{cases}$$

最后根据每月预订数量来估计每天预订数量。考虑到每月预定分为当天预定当天入住，当天预定非当天入住两类。因此由每月预定总数预测每天预定数量时将对上述两种类型的预定分开计算。对于当天预定当天入住类型，由 2015

年统计数据知，当天预订当天入住的房间总数为 13170（间），占总数 50.8%。对于此类型应考虑 5.2.1 中列出的入住数量与工作日、双休日、节假日等影响因素的关系按权分配。对于另一类型，由于非当天入住，该类型的每天预定房间数量不受当天的各种影响因素影响。因此可将此类型的预定数量平均分配。具体分配过程如下：

(1). 计算每月预订数量中当天预订当天入住的房间数量设为  $s_{i1}$  和非当天入住的房间数量  $s_{i2}$ ，以及该月的总天数设为  $M_i$ ， $M_i$  由该月的工作日、双休日及节假日天数三部分组成，则有以下公式：

$$M_i = a_{i,1} + a_{i,2} + a_{i,3}$$

(2). 计算每天预定房间数量占第  $i$  月总预定房间数量的权重：对于第  $i$  月的某一天，设工作日、双休日、节假日权重分别为  $r_{i1}, r_{i2}, r_{i3}$  则有下列计算公式：

$$r_{i1} = \frac{x_1}{x_1 a_{i,1} + x_2 a_{i,2} + x_3 a_{i,3}}, r_{i2} = \frac{x_2}{x_1 a_{i,1} + x_2 a_{i,2} + x_3 a_{i,3}}, r_{i3} = \frac{x_3}{x_1 a_{i,1} + x_2 a_{i,2} + x_3 a_{i,3}}$$

(3). 设工作日、双休日、节假日预定房间数量分别为  $R_{i1}, R_{i2}, R_{i3}$  则：

$$R_{ij} = s_{i,1} r_{ij} + \frac{s_{i,2}}{M} \quad (j = 1, 2, 3)$$

下面列出 1 月的每天预定数量，具体详细结果请见附件 2。

表 4 一月每天预定数量统计表

	周日	周一	周二	周三	周四	周五	周六
一月						1	2
						75	75
	3	4	5	6	7	8	9
	75	74	74	74	74	74	110
	10	11	12	13	14	15	16
	110	74	74	74	74	74	110
	17	18	19	20	21	22	23
	110	74	74	74	74	74	110
	24	25	26	27	28	29	30
	110	74	74	74	74	74	110
	31						
	110						

注：红色填充表示节假日，蓝色填充表示双休日。黑色字体表示日期，蓝色字体表示人数。

## 六、模型评价及改进

### 6.1 模型评价

#### 6.1.1 问题一模型评价

问题一中使用了多个模型，不同阶段使用不同模型算法的组合方式效果较为理想，其中较优的几点是：

- (1) 对 Markov 过程中状态转移矩阵的调整，这使得和自然状态下更加符合；
- (2) 遗传算法的使用，增加了随机性这一条件，从而使得寻找最优解可能性增加，虽然每次计算结果可能不同，但有总体趋势可以观察。
- (3) 遗传算法中选择过程的控制和实际操作吻合性很好，也验证了这种实际疏散方式的可行性。

#### 6.1.2 问题二模型评价

该数学模型考虑因素全面，分开探讨了每月入住房间数量与月份、双休日、节假日、暑假、季节等影响因素的关系及预订数量与入住数量之间的关系。并且将预定按照是否当天入住的标准分为两类，更好的预测每天的预定房间数量。

但是由于舍去了前三个月的数据，无法较为准确的描述寒假这一假日对相关月份入住房间数量的影响。关于每月预订数量与当月及下个月入住房间数量关系的研究中，线性拟合存在一定的误差。

### 6.2 模型改进

#### 6.2.1 问题一模型改进

- (1) 对于游乐园初始状态下人流分布，可以调研具体数据及周围实际情况后采用 Anylogic 仿真软件进行仿真模拟。
- (2) 层次分析法主观因素较强，所以 MNL 模型在使用层次分析法得出的系数时，可以利用统计软件 SAS 结合实际数据对系数进行计算检验。
- (3) 个人效用函数中对距离的处理可以增加一个常量，从而可以讨论游客在体验一个游乐项目后重新体验该相同游乐项目的可能性。
- (4) 遗传算法中交叉概率和变异概率可以具体细化调控，以寻找最合适的情况。

#### 6.2.2 问题二模型改进

对于以入住数量确定预定数量这一算法，可以进一步分析数据探讨工作日、双休日、节假日等因素与游客提前多少天预定的实际分布。以此来完善该模型，使每天预定数量的预测更为准确。

## 七、参考文献

- [1] 董文永等. 最优化技术与数学建模[M]. 北京：清华大学出版社，2010.
- [2] 黄平等. 最优化理论与方法[M]. 北京：清华大学出版社，2009.
- [3] 谭永基，蔡志杰. 数学模型[M]. 上海：复旦大学出版社，2011.
- [4] 王德，马力等. 2010 年上海世博会场内人流模拟分析. 城市规划学刊[J]. 2006, 3: 58-63, 2006.
- [5] 张栋，张力楠，史宇恒等. 基于 MNL 模型的有轨电车出行选择行为研究. 交通与信息安全[J]. 2011, 4: 75-79, 2011.

## 源程序索引

### 问题一源程序

MATLAB 程序（具体程序请见附件 1）：

problem1-1.m %状态转移矩阵计算  
problem1-2.m %效用函数计算  
problem1-3.m %自然状态下的平衡计算  
problem1-4.m %遗传算法迭代计算



## 附件 1: MATLAB 源程序

### 1.1 problem1-1.m

%状态转移矩阵计算

function P=eff(v,V)

```
D=[0 300 600 1050 350 1550 900 1050 600 250;
    300 0 300 750 350 1250 900 1250 800 550;
    600 300 0 450 500 950 1050 1400 950 850;
    1050 750 450 0 950 500 1150 1550 1400 1300;
    350 350 500 950 0 1200 550 900 450 600;
    1550 1250 950 500 1200 0 650 1050 1500 1800;
    900 900 1050 1150 550 650 0 400 850 1150;
    1050 1250 1400 1550 900 1050 400 0 450 800;
    600 800 950 1400 450 1500 850 450 0 350;
    250 550 850 1300 600 1800 1150 800 350 0];
N=[1 1 0 0 1 0 0 0 0 1;
    1 1 1 0 1 0 0 0 0 0;
    0 1 1 1 1 0 0 0 0 0;
    0 0 1 1 0 1 0 0 0 0;
    1 1 1 0 1 0 1 0 1 0;
    0 0 0 1 0 1 1 0 0 0;
    0 0 0 0 1 1 1 1 0 0;
    0 0 0 0 0 1 1 1 0;
    0 0 0 0 1 0 0 1 1 1;
    1 0 0 0 0 0 0 1 1];
t=[33 1.25 2.5 2.5 5 2.5 2 1.5 1.5 2];T=[];
for i=1:10
    l=t./(t(i));
    T=[T;l];
end
c=[400 30 50 30 100 50 30 30 20 50];C=[];
for i=1:10
    k=c./(c(i));
    C=[C;k];
end
U=ones(10);
for i=1:10
    for j=1:10
        U(i,j)=[log(D(i,j)),N(i,j),V(i,j),T(i,j),C(i,j)]*v;
    end
end
P=U;
```

### 1.2 problem1-2.m

%效用函数的计算

```
function [f, y]=suit(Vsc)
A=[1 5 1/3 3 7;1/5 1 1/5 3 3;3 5 1 7 7;1/3 1/3 1/7 1 3;1/7 1/3 1/7 1/3
1];
[v, d]=eig(A);
coef=-v(:, 1);
V=[];
for j=1:10
    l=Vsc./(Vsc(j));
    V=[V; l];
end
U=eff(coef, V);
S=sum(exp(U), 2);
for i=1:10
    for j=1:10
        P(i, j)=exp(U(i, j))/S(i);
    end
end
P=P';
tour=[10000, 0, 0, 0, 0, 0, 0, 0, 0, 0]';
for n=1:30
    t=P*tour;
    if max(abs(t-tour))<1
        %disp(n)
        break
    end
    tour=t;
end
y=tour;
ptime=max(tour'.*[33 1.25 2.5 2.5 5 2.5 2 1.5 1.5 2]./[400 30 50 30
100 50 30 30 20 50]);
f=1/ptime;
```

### 1.3 problem1-3.m

%层次分析法和自然状态下的平衡

```
clear, clc
A=[1 5 1/3 3 7;1/5 1 1/5 3 3;3 5 1 7 7;1/3 1/3 1/7 1 3;1/7 1/3 1/7 1/3
1];
[v, d]=eig(A);
coef=-v(:, 1);
CI=(d(1, 1)-5)/4;RI=1.12;
CR=CI/RI
```

```

V=randi(10,[10,10]);
U=eff(coef,V);
S=sum(exp(U),2);
for i=1:10
    for j=1:10
        P(i,j)=exp(U(i,j))/S(i);
    end
end
P=P'
tour=[10000,0,0,0,0,0,0,0,0,0]';
for n=1:30
    t=P*tour;
    Vsc=10*((ones(10,1)-t/sum(t))');%线性减少
    f=suit(Vsc);
    disp('time=')
    disp(1/f)
    Vb=[];
    for j=1:10
        l=Vsc./(Vsc(j));
        Vb=[Vb;l];
    end
    U=eff(coef,Vb);
    S=sum(exp(U),2);
    for i=1:10
        for j=1:10
            P(i,j)=exp(U(i,j))/S(i);
        end
    end
    P=P';
    disp(t')
    disp(sum(t))
    if max(abs(t-tour))<1
        disp(n)
        break
    end
    tour=t;
end

```

#### 1.4 problem1-4.m

%遗传算法迭代计算

```
clear,clc
```

```
nV=randi(10,[20,10]);
```

```
xx=300;
```

```
for n=1:100
```

```

        F=[];P=[];T=[];
    for i=1:20
        [f,tour]=suit(nV(i,:));
        F=[F,f];
        T=[T;tour'];
    end
    fm=max(F);
    if fm>1/90
        break
    end
    for i=1:20
        p=F(i)/sum(F);
        P=[P,p];
    end
    %selection
    S=[];TT=[];
    for i=1:10
        [z,1]=max(P);
        P(1)=0;S=[S;nV(1,:)];TT=[TT;T(1,:)];
    end
    [x,y]=max(TT');
    for j=1:10
        S(j,y(j))=0.05;
    end
    %crossover
    pc=rand;
    if pc<0.8
        nV=S;
        for j=1:5
            %disp('发生交叉')
            a=rand;
            k=randi(10,[1,2]);
            M=[a*S(k(1),:)+(1-a)*S(k(2),:);(1-a)*S(k(1),:)+a*S(k(2),:)];
            nV=[nV;M];
        end
    else
        nV=[S;S];
    end
    %mutation
    pm=rand;
    if pm<0.1
        disp('发生变异')
        l=randi(20);a=rand;
        nV(1,:)=nV(1,:)+2*(a-0.5)*10*ones(1,10);
    end
end

```

```

for i=1:10
    if nV(1,i)<0.1
        nV(1,i)=0.1;
    else
        if nV(1,i)>10
            nV(1,i)=10;
        end
    end
end
end
end
if xx>1/fm
    xx=(1/fm);
end
end
disp(xx)
disp(T(1,:))

```

附件 2：预测 1-3 月每天预定房间数量

	周日	周一	周二	周三	周四	周五	周六
一月						1	2
						75	75
	3	4	5	6	7	8	9
	75	74	74	74	74	74	110
	10	11	12	13	14	15	16
	110	74	74	74	74	74	110
	17	18	19	20	21	22	23
	110	74	74	74	74	74	110
	24	25	26	27	28	29	30
	110	74	74	74	74	74	110
	31						
	110						
二月		1	2	3	4	5	6
		80	80	80	80	80	119
	7	8	9	10	11	12	13
	81	81	81	81	81	81	81
	14	15	16	17	18	19	20
	119	81	81	81	81	81	119
	21	22	23	24	25	26	27
	119	81	81	81	81	81	119
	28	29					

	119	81					
三月			1	2	3	4	5
			80	80	80	80	118
	6	7	8	9	10	11	12
	118	80	80	80	80	80	118
	13	14	15	16	17	18	19
	118	80	80	80	80	80	118
	20	21	22	23	24	25	26
	118	80	80	80	80	80	118
	27	28	29	30	31		
	118	80	80	80	80		

注：红色填充表示节假日，蓝色填充表示双休日。黑色字体表示日期，蓝色字体表示人数。