

报名序号：2572

论文题目：基于优化调度的共享单车现状及发展研究



论文题目：基于优化调度的共享单车现状及发展研究

摘要

随着共享单车的出现和普及，共享单车系统在城市公共交通中的地位越来越重要，由于其快捷、方便、环保的特点，共享单车已成为居民解决出行“最后一公里”问题的重要选择。

首先，我们利用 **VBA 编程** 对数据进行处理，整理成为包含以下五项指标的数据表：**出发时间、到达时间、出发区域、到达区域、骑行时长**，对骑行时长进行方差分析，从而假设从 A 地出发到达 B 地的骑行时长以平均值代替，并假设骑行速度都统一为正常人的平均速度 15km/h，从而解出各个区域相对距离矩阵。

针对问题一，我们分为时间分布和空间分布讨论。从时间分布上看，各个区域单车在 420—900 分钟流动频繁、分布密集，其他时段逐渐下降，1380 分钟以后分布趋于静止。从空间分布来说，我们通过 **VBA 编程** 统计分析得到从出发区域至到达区域车次的 **OD 矩阵**，可以明显看出，共享单车在 **1 到 2、2 到 3、3 到 4、……、9 到 10** 之间的分布较为密集、其他地区基本呈现均匀分布。

针对问题二，建立了**单调度中心调度模型**和**动态调度优化模型**。根据附件 1、附件 2 的数据计算出现有单车情况下的各个地区的单车需求量，确定不同地区单车的需求时间和可接受时间。根据问题 1 结果估计不同地区间的最短距离，建立单调度中心调度模型并利用 **MATLAB 软件** 和 **遗传算法** 求解得出**初始调度方案**，**最优调度路线为 2-5-6-3-7-10-9-2**。以单调度中心软时间窗调度模型为基础建立动态需求调度优化模型，向初始静态优化解中不断插入新的调度需求，应用“**初始静态优化+实时动态优化**”的方法对多个连续静态调度问题进行求解不断优化调度路线，**初次优化结果为 5-2-3-4-8-10-9**，其中 **5、2、3 已完成调度**。

针对问题三，要判断各区域所需共享单车的满足程度。首先我们定义了从出发区域 i 到到达区域 j 的**需求量、满足量以及需求满足比**作为衡量各区域所需共享单车满足程度的指标。根据问题一得到的 **OD 矩阵**以及附件 2 所给的需求矩阵，即可计算出从出发区域 i 到到达区域 j 的需求满足比，进而判断各区域所需共享单车的满足程度。由结果可知，**各区域的满足程度从 78%到 99%不等**，总体来看，满足程度较好。按照问题二的需求比例即可求得 100 辆单车的投放方案。

针对问题四，要分析研究共享单车的投入对该地区打车市场的影响。根据附件 3 的数据，我们将单车投放量作为自变量，打车人次作为因变量，作出散点图。观察得到在单车投放量小于 4000 辆时呈现线性关系，大于 4000 辆时呈现非线性关系。因此我们以 $x=4000$ 为分段点建立**一元线性回归模型、一元非线性回归模型**，利用 **MATLAB 求解**，得出方程。通过观察方程，我们得出共享单车投放量与打车人次呈负相关，即共享单车投放量越大，打车人次越少，从而对打车市场产生负面影响。并且搜集到上海市 4 月份单车使用次数及打车人次的数据，从时间维度分析我们可以得出，固定日期分时段单车使用次数与打车人次数的图像呈现近似的趋势，而固定时段分日期单车使用次数与打车人次呈现**线性负相关的关系**。

关键词：单调度中心调度模型 动态调度优化模型 一元线性回归模型
一元非线性回归模型 共享单车 SPSS MATLAB



目录

一、问题重述.....	1
二、模型假设.....	1
三、符号说明.....	2
四、问题分析.....	2
4.1 数据分析.....	2
4.2 问题分析.....	4
五、模型的建立与求解.....	5
5.1 问题一.....	5
5.1.1 共享单车的时间分布.....	5
5.1.2 共享单车的空间分布.....	7
5.2 问题二模型的建立及求解.....	8
5.2.1 模型的建立.....	8
5.2.1.1 单调度中心调度模型.....	8
5.2.1.2 动态需求调度优化模型.....	9
5.2.2 模型的求解.....	10
5.2.2.1 数据的预处理.....	10
5.2.2.2 优化调度模型的求解.....	11
5.3 问题三模型的建立及求解.....	13
5.3.1 不同时空下共享单车的满足程度评价模型的建立.....	13
5.3.2 不同时空下共享单车的满足程度评价模型的求解.....	14
5.3.3 100 辆单车的分配求解.....	15
5.4 问题四模型的建立和求解.....	16
5.4.1 一元线性回归模型.....	17
5.4.1.1 模型的建立.....	17
5.4.1.2 模型的求解.....	18
5.4.1.3 模型的检验.....	18
5.4.1.4 模型的检验结果.....	20
5.4.2 一元非线性回归模型.....	21
5.4.3 量化研究.....	22
六、模型的评价及灵敏度、敏感性分析.....	27
6.1 模型评价.....	27
6.1.1 动态优化模型的评价.....	27
6.1.2 一元回归模型的评价.....	28
6.2 灵敏度分析.....	28
七、参考文献.....	29

一、问题重述

共享单车是指企业在校园、地铁站点、公交站点、居民区、商业区、公共服务区等提供自行车单车共享服务，是一种分时租赁模式。共享单车是一种新型共享经济。共享单车已经越来越多地引起人们的注意，由于其符合低碳出行理念，政府对这一新鲜事物也处于善意的观察期。

很多共享单车公司的单车都有 GPS 定位，能够实现动态化地监测车辆数据、骑行分布数据，进而对单车做出全天候供需预测，为车辆投放、调度和运维提供指引。

请根据下面附件给出数据及结合根据需要自己收集的数据，完成以下问题：

(1) 根据附件 1 中共享单车的骑行数据，估计共享单车的时空分布情况。如从某地点 A 出发，到达不同地点的分布情况。可分时间段讨论。

(2) 假如根据调查，得到人们的骑行需求估计数据，见附件 2。

根据问题 1 的估计结果，建立数学模型解决如何优化共享单车的调度问题。

(3) 根据附件 1 的骑行数据和附件 2 的需求数据，判断各区域所需共享单车的满足程度，给出你的度量指标。若增加 100 辆单车，如何进行投放更优。

(4) 附件 3 是某地区投入不同数量共享单车后打车人次的数据。据此分析研究共享单车的投入对该地区打车市场的影响。同时请你收集实际数据进行量化研究。

二、模型假设

- ❖ 假设 0 分钟为凌晨 0 点；
- ❖ 假设各单车的行驶速度均匀固定；
- ❖ 假设出发区域 i 到到达区域 j 的相对距离=骑行时长 \times 骑行速度。
- ❖ 假设区域 1 到 10 只有一个调度中心且位于区域 1，配备足够多的同型号调度车辆，能为地区 1 到 10 内所有地点提供调度服务；
- ❖ 假设调度车辆必须从调度中心出发，完成调度任务后需返回调度中心；
- ❖ 假设各地区的调度需求不能大于调度车辆容量；
- ❖ 假设各地区间的最短路网距离已知，调度车辆在任意两地区间行驶时必须选择两地区间的路网最短路径；
- ❖ 假设调度车辆在路网中的平均行驶速度已知且固定，不受其他因素影响；
- ❖ 假设前一 Δt 内完成调度的最后一个地点是下一 Δt 内新调度路径的起点；
- ❖ 假设所收集到的上海市单车使用人次数和出租车使用人次真实可靠。

三、符号说明

这里只列出部分通用符号，个别模型单独使用符号会在下文中具体说明。

符号	说明
N	地中心管理的所有地区数量
G	片区调度中心配备的调度车辆总数
D	调度车装载共享单车的最大容量
V	调度车辆在路网中的的平均行驶速度
φ	权重系数
D_{ij}	各地区间的路网最短距离
Q_i	地区 i 的调度需求
d_m	调度车辆在城市路网条件下的最大行驶里程
D_{g0}	从片区调度中心出发时调度车辆 g 上的初始装载 自行车数
L_{gi}	第 g 台调度车辆的调度路径集合， $L_{gi} = \{l_{gi} \mid g = 1, 2, \dots, N\}$
L_{g0}	表示地点 i 需要调度服务，且由调度车辆 g 完成该 地点的调度任务
$[T_{1i}, T_{2i}]$	地区 i 的调度最佳时间窗
$[T'_{1i}, T'_{2i}]$	地区 i 的调度可接受时间窗

四、问题分析

4.1 数据分析

1. 数据解读

结合部分附件 1 中共享单车的骑行数据，我们对数据进行解读。如下表所示，表中数据代表了一号自行车，在 468.28 分钟从区域 8 出发，于 471.67 分钟到达区域 1，又在 569.13 分钟从区域 1 出发到达区域 4，……，具体过程见图 4.1.1。

表 4.1.1 自行车 1 的行驶过程

自行车序号：1		出发区域：8			
出发时间	468.28	到达时间	471.67	到达区域	1
出发时间	569.13	到达时间	580.61	到达区域	4
出发时间	655.88	到达时间	669.49	到达区域	6
出发时间	724.07	到达时间	728.71	到达区域	10
出发时间	778.06	到达时间	782.69	到达区域	5
出发时间	877.31	到达时间	891.67	到达区域	4

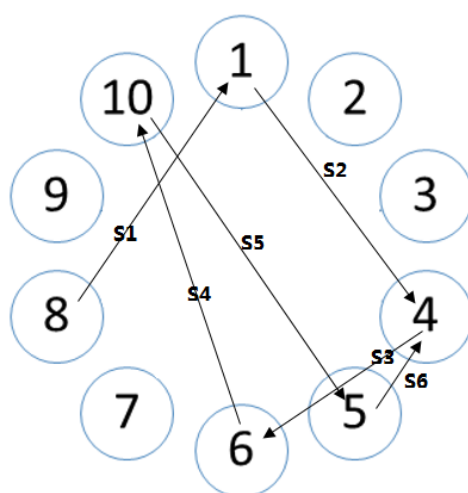


图 4.1.1 一号自行车骑行轨迹

2. 数据预处理

在对数据进行统计分析时，我们做出假设：从区域 A 到区域 B 的骑行时间可用一个平均值代替，并且假设人们的骑行速度都符合正常人平均骑行速度 15km/h。最后用平均骑行时间*骑行速度计算得出各区域相对距离矩阵，为之后的问题求解提供理论基础。

首先我们对各区域间骑行时间做方差检验，如下表，发现方差值极小，假设成立。

表 4.1.2 骑行时间方差检验

	1	2	3	4	5	6	7	8	9	10
1		0.338	0.769	0.410	0.210	0.277	0.366	0.436	0.542	0.207
2	0.333		0.068	0.142	0.596	0.508	0.322	0.226	0.041	0.399
3	0.649	0.081		0.594	0.262	0.110	0.710	0.059	0.604	0.154
4	0.389	0.163	0.671		0.052	0.180	0.429	0.063	0.154	0.434
5	0.215	0.597	0.270	0.065		0.373	0.541	0.425	0.747	0.701
6	0.270	0.535	0.106	0.156	0.408		0.128	0.062	0.099	0.648
7	0.382	0.344	0.665	0.497	0.496	0.101		0.057	0.699	0.679
8	0.465	0.254	0.061	0.069	0.396	0.065	0.072		0.514	0.049
9	0.575	0.034	0.683	0.147	0.736	0.097	0.686	0.549		0.065
10	0.235	0.398	0.142	0.354	0.553	0.511	0.559	0.043	0.067	

注：表格中 i 行 j 列代表从 i 区域到 j 区域骑行时间样本方差值

然后利用公式求得区域间距离矩阵

$$l_{ij} = \bar{v}_{\text{自行车}} \times t_{ij}$$

其中 l_{ij} 代表从 i 区域到 j 区域的距离， t_{ij} 代表从 i 区域到 j 区域的骑行时间。

表 4.1.3 区域间相对距离表（单位：m）

	1	2	3	4	5	6	7	8	9	10
1		2600.7	3569.5	2680.3	2026.3	2259.8	2724.8	2840.6	3087.6	2054.7
2	2600.7		1219.7	1705.5	3415.7	3139.8	2535.6	2143.3	876.5	2731.8
3	3569.5	1219.7		3343.5	2227.4	1356.4	3469.8	1006.3	3503.9	1777.9
4	2680.3	1705.5	3343.5		1012.0	1685.1	3032.6	1064.4	1715.1	2752.7
5	2026.3	3415.7	2227.4	1012.0		2579.4	3049.8	2795.5	3700.0	3409.4
6	2259.8	3139.8	1356.4	1685.1	2579.4		1515.1	1077.3	1312.8	3393.5
7	2724.8	2535.6	3469.8	3032.6	3049.8	1515.1		1108.3	3484.4	3396.0
8	2840.6	2143.3	1006.3	1064.4	2795.5	1077.3	1108.3		3156.2	920.6
9	3087.6	876.5	3503.9	1715.1	3700.0	1312.8	3484.4	3156.2		1112.3
10	2054.7	2731.8	1777.9	2752.7	3409.4	3393.5	3396.0	920.6	1112.3	

4.2 问题分析

针对问题一，在分析共享单车的时空分布情况时，考虑到时空既可以指不同时间下的空间分布，又可以指不同区域单车的时间分布。我们将其综合考虑，对于时间分布，直接将各个点做出散点图并不直观，所以我们将其划分成每 60 分钟一个时间段，统计时间段内的单车出发、到达车次，绘制图表观察。对于空间分布，我们通过建立 OD 矩阵，描述区域间交通中一定时间范围内所有出行起点和终点之间的出行交通量，间接反映空间分布情况。

针对问题二，题目要求根据问题 1 的共享单车的时空分布情况的估计结果，建立数学模型解决如何优化共享单车的调度问题。首先，假设骑行人的骑行距离一定，由平均骑行时间可计算出各地区间的相对距离。其次，根据客户需求是否实时变动，调度问题可分为静态和动态调度问题。对动态调度问题求解时可将动态调度问题转化为多个连续静态调度问题进行求解，即“初始静态优化+实时动态优化”。先根据系统内各地点在某一时刻的调度需求，形成初始的调度方案，然后不断更新各地点的调度需求，同时更新调度方案，实现调度需求与调度路径的不断反馈，最终达到动态优化的目的。

针对问题三，要判断各区域所需共享单车的满足程度。首先我们应该衡量各区域所需共享单车满足程度的指标，建立指标评价体系，但是由于题目所给数据不全面，无法得到指标体系的具体数据，所以我们只用需求满足比来作为衡量各区域所需共享单车满足程度的指标。

针对问题四，要分析研究共享单车的投入对该地区打车市场的影响。根据附件 3 的

数据，我们可以将单车投放量作为自变量，打车人次作为因变量，作出散点图，观察单车投放量与打车人次的关系，探究是线性还是非线性，是正相关还是负相关。再用 MATLAB 软件得到具体的函数关系。进而得到单车投放量对打车市场的影响关系。

五、模型的建立与求解

5.1 问题一

共享单车的时空分布情况是指：通过共享单车系统运行数据，对共享单车在一定时间内的空间分布，以及在一定空间内共享单车到达、出发数量随时间变化情况的分析。

5.1.1 共享单车的时间分布

为了使结果更加直观，我们将出发、到达时间分为每 60 分钟一段，分别统计每段时间内的单车数量变化情况，并分区域绘制出单车数量—时间变化曲线。

表 5.1.1 分区域到达车次时间分布表

到达时间 (分) \ 到达区域	1	2	3	4	5	6	7	8	9	10
360-420	40	37	52	47	50	54	45	63	49	58
420-480	71	99	95	92	84	99	88	114	106	84
480-540	89	107	86	108	108	79	107	87	89	98
540-600	94	83	112	88	106	108	94	125	85	94
600-660	95	74	110	116	110	86	86	109	92	84
660-720	78	94	92	98	93	100	100	97	94	95
720-780	84	86	83	75	72	111	89	96	92	79
780-840	74	78	77	82	92	85	90	73	71	98
840-900	71	86	63	88	62	85	68	84	84	69
900-960	68	71	68	66	69	65	61	76	70	81
960-1020	55	60	62	63	61	72	65	61	53	57
1020-1080	54	60	50	56	59	43	55	60	59	50
1080-1140	35	61	48	52	49	65	58	56	44	64
1140-1200	50	50	52	26	47	41	47	51	40	50
1200-1260	33	49	40	48	29	32	36	47	39	46
1260-1320	25	34	26	34	37	39	33	24	30	17
1320-1380	20	36	34	31	27	24	25	20	25	23
1380-1440	18	7	7	16	21	13	13	16	8	7

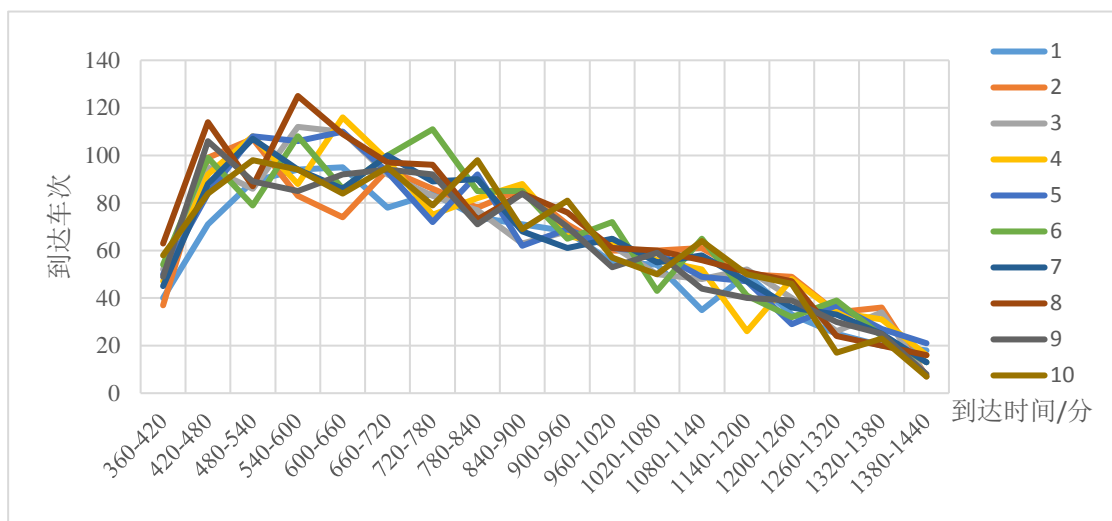


图 5.1.1 分区域到达车次时间分布曲线

表 5.1.2 分区域出发车次时间分布表

出发 时间 (分)	出发 区域	1	2	3	4	5	6	7	8	9	10
360-420		60	71	55	56	56	63	56	85	52	60
420-480		84	91	91	107	115	115	92	111	101	101
480-540		80	87	89	88	82	84	88	113	108	85
540-600		93	106	104	111	102	86	104	91	85	98
600-660		91	92	106	93	104	102	97	122	79	86
660-720		80	73	100	104	98	84	85	114	87	95
720-780		74	77	99	87	97	102	82	88	97	85
780-840		78	82	60	78	70	98	92	79	89	81
840-900		70	75	80	88	78	75	76	65	62	67
900-960		60	79	67	73	69	70	62	69	68	62
960-1020		53	61	48	54	53	67	54	63	70	77
1020-1080		49	47	60	49	58	61	66	61	46	46
1080-1140		55	69	48	54	53	47	38	57	51	49
1140-1200		38	41	41	43	42	45	51	49	40	59
1200-1260		37	47	43	30	39	31	48	41	31	41
1260-1320		23	32	31	32	20	33	21	42	25	30
1320-1380		19	34	25	39	21	28	29	16	24	20
1380-1440		8	17	11	8	7	7	11	9	9	9

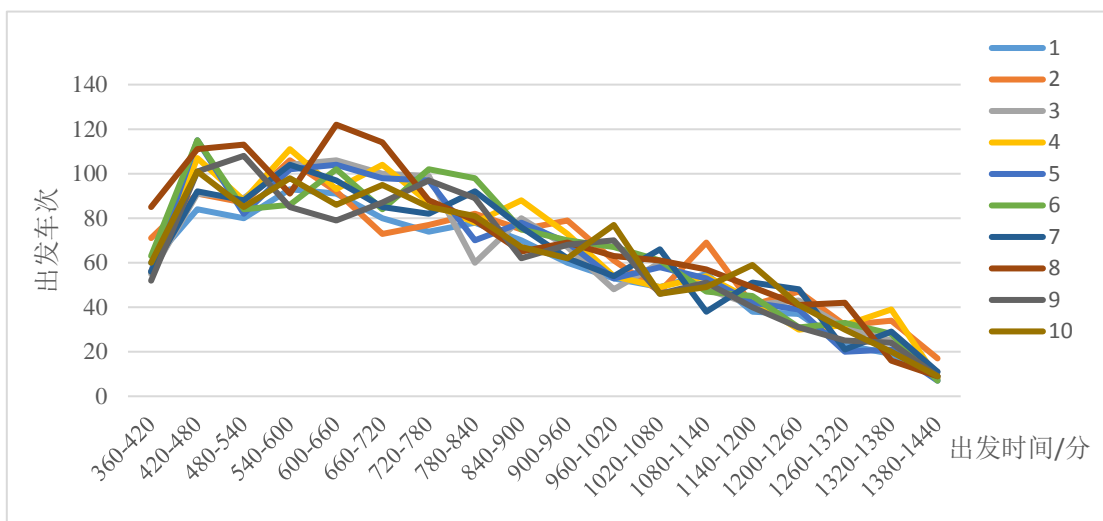


图 5.1.2 分区域出发车次时间分布曲线

从折线图中，我们可以明显的看出，各个区域单车数量从早晨 6:00 开始在平均水平波动，在 7:00—15:00 有一个小高峰，出发车次最多高达 122 次，到达车次 114 次，其他时间呈现逐渐下降趋势，23:00 以后达到最小值，车次下降至个位数。单车的时间分布与人们的活动规律紧密联系，我们得出的单车时间分布规律与人们白天出行、夜间休息的出行规律相符合，结果合理。

5.1.2 共享单车的空间分布

OD 矩阵描述了区域间交通中一定时间范围内所有出行起点和终点之间的出行交通量。OD 矩阵中“OD”取自英文单词 Origin(起点)和 Destination(终点)的第一个字母,从英文的字面意思可知道,OD 矩阵即“起讫点矩阵”。OD 矩阵主要反映的是单车使用者从一个区域到另外一个区域的动态迁移状况,在一定程度上反映了单车的空间分布情况。

通过对原始数据的统计分析，我们得到了各区域单车实际使用情况数据，i 行 j 列数据代表从区域 i 到区域 j 使用共享单车的人次，整理出 OD 矩阵如下。

表 5.1.3 共享单车出行人次的 OD 矩阵

到达 出发	1	2	3	4	5	6	7	8	9	10
1		229	95	101	134	105	92	113	87	96
2	122		222	115	116	120	115	121	113	137
3	116	127		230	97	107	112	113	140	116
4	110	98	141		247	134	118	117	113	116
5	117	122	110	107		248	132	115	100	113
6	133	119	131	126	120		221	114	122	112
7	111	125	116	130	112	120		231	97	110
8	124	146	121	130	129	132	134		237	122
9	99	112	109	121	103	134	104	110		232
10	122	94	112	126	118	101	132	225	121	

我们可以直接看出，共享单车在 1 到 2、2 到 3、3 到 4、……、9 到 10 之间的分布较为密集、其他地区基本呈现均匀分布。

5.2 问题二模型的建立及求解

我们以单调度中心软时间窗调度模型为基础建立动态需求调度优化模型，求解时先求得初始静态优化解，再不断插入新的调度需求，不断优化调度路线，直至调度任务结束。

5.2.1 模型的建立

5.2.1.1 单调度中心调度模型

单调度中心调度优化模型以调度路径长度最短和乘客等待时间最短为优化目标，建立多目标优化模型如下：

$$\min Z_1 = \sum_{g=1}^G \sum_{i=0}^N \sum_{j=0}^N F_{gi} Y_{gij} d_{l_{gi} l_{gj}}$$

$$s.t. \sum_{i=0}^N \sum_{j=0 \& j \neq i}^N F_{gi} Y_{gij} d_{l_{gi} l_{gj}} \leq d_m$$

$$\sum_{g=1}^G \sum_{j=1}^N Y_{g0j} \leq G$$

$$\forall g, L_{g1} \cap L_{g2} = \emptyset$$

$$\sum_{j=1}^N Y_{g0j} = \sum_{j=1}^N Y_{gj0} = 1, g = 1, 2, \dots, G$$

$$F_{gi} = \begin{cases} 1 & \text{地点 } i \text{ 由第 } g \text{ 台车辆进行调度} \\ 0 & \text{否则} \end{cases}$$

$$Y_{gij} = \begin{cases} 1 & \text{第 } g \text{ 台车辆由地点 } i \text{ 驶向 } j \\ 0 & \text{否则} \end{cases}$$

上述模型中，目标函数要求所有调度车辆的路径长度之和最短，各约束条件意义如下：

- 约束 1：各条调度路径的总长度不能大于调度车辆的路网最大行驶里程；
- 约束 2：调度车辆总数不能超过该片区调度中心配备的调度车辆总数 G ；
- 约束 3：任意时刻，调度车上的自行车数量不能大于调度车装载共享单车的最大容量 D ；

约束 4：同一地点不能被多个调度车辆服务；

所有调度车辆均必须从调度中心出发，完成自身调度任务后还必须返回调度中心。以乘客等待时间最短为目标函数建立模型如以下公式所示：

$$\min Z_2 = \theta \sum_{i=1}^N \max \left\{ (T_{li} - T_{l_{gi}}), (T_{l_{gi}} - T_{li}), 0 \right\}$$

$$s.t. T_{l_{gi}} + t_0 \cdot |Q_i| + t_{l_{gi}l_{gij}} = T_{l_{gi}} \quad (i=1,2,\dots,N)$$

上述模型中，目标函数要求各地点的人等待时间总和最小， θ 为时间延误惩罚系数，用来对超出最佳时间窗的解进行惩罚，并协调与目标函数间的量级关系，约束条件意义如下：

调度车辆 g 到达地点 j 的时刻=到前一地点 i 的时刻+车均服务时间 \times |地点 i 的调度需求|+地点 i 到地点 j 的行驶时间。

$$\min \{\varphi_1 Z_1 + \varphi_2 Z_2\}$$

如上式，其中 φ_1 、 φ_2 为目标函数权重系数，此即为单调度中心软时间窗调度模型。

5.2.1.2 动态需求调度优化模型

在动态需求调度优化模型中，以调度路径长度最短为目标函数建立模型时要将参与本次调度任务的调度车辆分两类分析。

1. 已完成调度任务的调度车辆

对于已完成调度任务的调度车辆，以调度路径长度最短为目标函数建立模型如式

$$Z_1 = \sum_{g=1}^G \sum_{i=0}^N \sum_{j=0}^N F_{gi} Y_{gij} d_{i_g j_{gi}}$$

2. 仍在进行调度任务的调度车辆

假设在 $n\Delta t$ 时刻系统对调度需求进行变更，此时调度车辆 g 正在或已经完成地点 h 的调度任务，以调度路径长度最短为目标函数建立模型如式。

$$Z_2 = \sum_{g=1}^g \sum_{i=0}^{l_{gi}} \sum_{j=0}^{l_{gij}} F_{gi} Y_{gij} d_{l_{gi} l_{gij}} + \sum_{g=1}^g \left\{ \left[\sum_{i=l_{gi}}^N \sum_{j=l_{gij}}^N F_{gij} Y_{gij} d_{l_{gi} l_{gij}} \right] + d_{l_{gi} l_{gij}} \right\}$$

其中，第一项表示调度车辆 g 在时刻 Δt 已经行驶的里程，第二项表示调度路径优化后调度车辆 g 对剩下地点进行服务并返回调度中心的行驶里程。

$$Z_3 = Z_1 + Z_2$$

以上两式相加即为动态需求调度优化模型中，调度路径长度最短的目标函数模型，模型约束条件与单调度中心软时间窗调度模型相同。

$$s.t. \sum_{i=0}^N \sum_{g=1}^N F_{gi} Y_{gij} d_{l_{gi} l_{gij}} \leq d_m$$

$$\sum_{g=1}^G \sum_{j=1}^N Y_{g0j} \leq G$$

$$\forall i, 0 \leq D_{g0} + \sum_{j=1}^N Q_{gi} \leq D$$

$$\forall g, L_{g1} \cap L_{g2} = \emptyset$$

$$\sum_{j=1}^N Y_{g0j} = \sum_{j=1}^N Y_{gj0} = 1, g = 1, 2, \dots, G$$

以乘客等待时间最短为目标函数建立模型所示，模型约束条件与单调度中心软时间窗调度模型相同。

$$\min Z_4 = \theta \sum_{i=1}^N \max \left\{ (T_{li} - T_{l_{gi}}), (T_{l_{gi}} - T_{li}), 0 \right\}$$

$$s.t. T_{l_{gi}} + t_0 \cdot |Q_i| + t_{l_{gi} l_{gij}} = T_{l_{gi}} \quad (i = 1, 2, \dots, N)$$

对两个目标函数给予不同的权重，将多目标转问题转化为单目标问题如下式，其中 φ_1 和 φ_2 为目标函数权重系数，此即为动态需求调度优化模型。

$$\min \{ \varphi_1 Z_3 + \varphi_2 Z_4 \}$$

5.2.2 模型的求解

5.2.2.1 数据的预处理

1. 计算各地区需求差

首先，我们将附件 1 中数据整理成一天中 1-10 各个地区中间行驶人次表 5.1.3，将其与附件 2 骑行需求估计表进行作差，得出各个路段间实际行驶人次与需求人次的差。接着，我们以到达目的地为标准，统计 1-10 各地区的需求差，把需求差按到达地区不同分成 1 到 10 地区的需求差。

将一天中的骑行时间每小时分段，把实际出发人次按出发地点不同划分出不同时间段的骑行出发次数，同时计算每个时间段的出行人次占一天中出行总人次的百分比。把需求差按不停时间段的实际出行比例划分得出数据。

2. 计算不同地区间的最短距离 d_{ij}

统计各地区之间的骑行时间，去除无效值和误差较大的值，对它们取平均值，取平均骑行速度为 15km/h，得到各地区的最短相对距离如下：

表 5.2.1 各地最短相对距离（单位：m）

到达 出发	1	2	3	4	5	6	7	8	9	10
1	2601	0	0	0	0	0	0	0	0	0
2	3569	1220	0	0	0	0	0	0	0	0
3	2680	1706	3344	0	0	0	0	0	0	0
4	2026	3416	2227	1012	0	0	0	0	0	0
5	2260	3140	1356	1685	2579	0	0	0	0	0
6	2725	2536	3470	3033	3050	1515	0	0	0	0
7	2841	2143	1006	1064	2795	1077	1108	0	0	0
8	3088	877	3504	1715	3700	1313	3484	3156	0	0
9	2055	2732	1778	2753	3409	3393	3396	921	1112	0
10	2601	1220	1220	1220	1220	1220	1220	1220	1220	0

共享单车系统目前依靠专门车辆进行调度作业，在此假设调度中心的调度车辆足够多，每辆调度车辆的容量为 30 辆，最大行驶距离为 40 公里，高峰时段行驶速度为 30km/h，配有 2 名调度作业人员。

5.2.2.2 优化调度模型的求解

1. 算法参数

初始种群为 $N=50$ ，遗传代数 $T=100$ ，调度车辆从调度中心出发时的自行车数在 $[0,30]$ 之间随机生成。

2. 求解过程

(1) 单调度中心模型调度方案

通过数据统计得到第 840 分钟到第 900 分钟（下午 2:00-3:00）需要调度的地区有 2、3、5、6、7、9、10。得到下表：

表 5.2.2 各地区需求及时间窗

地区	调度需求	最佳时间窗	可接受时间窗
1	-	-	-
2	19	2:30-2:40	2:30-2:50
3	34	2:35-2:45	2:20-2:50
4	-	-	-
5	-10	2:20-2:30	2:15-2:40
6	12	2:35-2:45	2:35-3:00
7	17	2:40-2:55	2:30-2:50
8	-	-	-
9	-8	2:50-3:00	2:35-3:10
10	12	2:35-2:50	2:20-2:55

注：调度需求为正表示从地点调出车辆，调度需求为负表示向地点调入车辆。

确定各地区是否需要调度的方法是：由数据统计得出各地点在第 840 分钟到第 900 分钟的实需调度数量。对各地区的实需调度数量取平均，得到平均数为 7 人次，则实需调度数量的绝对值大于平均需求数的地区为需要调度的地区。

利用本文建立的单调度中心软时间窗调度模型及模型求解方法对以上案例进行分析，利用 MATLAB 软件得到多组局部最优解如下表所示。

表 5.2.3 单调度中心软时间窗模型求解结果

序号	调度车辆数	调度路径	目标函数
1	4	2-5-6-3-7-10-9-2 2-6-7-2 2-7-10-5-6-2	9.24
2	5	2-5-6-2 2-6-7-9-5-10-3-2	8.19
3	4	2-9-7-3-6-5-2 2-5-10-7-2	8.53
4	4	2-10-5-6-7-9-3-2 2-6-7-2	8.69
5	4	2-3-5-2 2-9-10-5-7-6-2 2-10-5-6-7-9-3-2	8.93

观察表 5.2.3 列出的 5 组局部最优解，此时，系统普遍需要 4-5 辆调度车才能完成调度任务，目标函数的值基本稳定在 9 附近，5 组解中的最优解 1 的目标函数值为 9.24 对局部最优解 2 的时间约束进行分析，根据初始调度需求求解得出初始调度方案如表 5.2.4 所示。

表 5.2.4 初始调度方案

地区	到达时间	最佳时间窗	可接受时间窗
1	2:30	—	—
2	2:31'13"	2:30-2:40	2:30-2:50
5	2:35'28"	2:20-2:30	2:15-2:40
6	2:41'19"	2:35-2:45	2:35-3:00
3	2:45'14"	2:35-2:45	2:20-2:50
10	2:52'19"	2:35-2:50	2:20-2:55
7	2:53'21"	2:40-2:55	2:40-3:00
9	2:55'44"	2:50-3:00	2:35-3:10
1	3:04	—	—

(2) 动态优化调度方案

3:00 时动态优化调度方案采用确定初始调度需求的方法更新调度需求,更新结果如表 5.2.5。此时,地点 2,5,6,已经完成调度,对还未调度的地点进行动态优化,如下表所示。并综合已完成的调度路径,得到动态优化调度方案如表 5.2.6 所示。

表 5.2.5 需求更新

地区	调度需求	最佳时间窗	可接受时间窗	备注
4	16	2:40-2:55	2:35-3:00	新增
6	-	-	-	取消
8	18	2:45-2:55	2:35-3:05	新增
7	-	-	-	取消
10	-20	2:35-2:50	2:35-3:00	调度量变化

表 5.2.6 动态优化调度方案

地区	到达时间	最佳时间窗	可接受时间窗	备注
1	2:20	-	-	
5	2:26'28"	2:20-2:30	2:15-2:40	已完成调度
2	2:31'13"	2:30-2:40	2:30-2:50	已完成调度
3	2:45'14"	2:35-2:45	2:20-2:50	已完成调度
4	2:49'58"	2:40-2:55	2:35-3:00	
8	2:52'58"	2:45-2:55	2:35-3:05	
10	2:58'44"	2:35-2:50	2:35-3:00	
9	3:05'44"	2:50-3:00	2:35-3:10	
1	3:14'	-	-	

表 5.2.6 的动态优化调度方案中,大部分地点都能在最佳时间窗内完成调度,只 3、8 和 9 没有在最佳时间窗内完成调度,但调度时间也在可接受时间窗内,此方案乘客等待时间较短。系统将继续以 60 分钟为间隔更新调度需求,调度方案也将不断优化,直至调度任务结束。

5.3 问题三模型的建立及求解

5.3.1 不同时空下共享单车的满足程度评价模型的建立

从该地区各区域及一天内各时间段的共享单车需求、满足情况的角度出发,可以建立共享单车的满足程度评价模型。

为了得到合适的指标,来评价该地区不同时空的满足程度,可以从满足量、需求量、需求满足比这三个角度进行分析。

①需求量

将位于第*i*区域内，需要骑行共享单车到达第*j*区域的人次定义为从第*i*区域到第*j*区域的需求量 Q_{ij} 。

②满足量

将从第*i*区域内出发，已经骑行共享单车到第*j*区域的人次定义为第*i*区域到第*j*区域的满足量 M_{ij} 。

③需求满足比

依据以上获得的第*i*区域到第*j*区域的需求量和满足量，我们可以得到第*i*区域到第*j*区域的需求满足比作为衡量第*i*区域内共享单车满足程度的指标。所以，我们得出需求满足比的函数表达式为：

$$\eta_{ij} = \frac{M_{ij}}{Q_{ij}} \quad i, j = 1, 2, \dots, 10$$

5.3.2 不同时空下共享单车的满足程度评价模型的求解

由表 5.1.3 和附件 2，我们可以得到各区域的需求满足比，其中第*i*行第*j*列数据代表从区域*i*到区域*j*的需求满足比。

表 5.3.1 各区域的需求满足比

区域	1	2	3	4	5	6	7	8	9	10
1		0.95	0.80	0.82	0.92	0.83	0.80	0.89	0.78	0.94
2	0.90		0.99	0.89	0.97	0.95	0.83	0.95	0.93	0.98
3	0.93	0.96		0.92	0.88	0.91	0.89	0.82	0.89	0.92
4	0.95	0.85	0.95		0.95	0.93	0.89	0.83	0.91	0.97
5	0.91	0.92	0.89	0.92		0.91	0.96	0.96	0.97	0.86
6	0.84	0.93	0.92	0.90	0.83		0.91	0.86	0.87	0.97
7	0.86	0.89	0.93	0.89	0.83	0.87		0.97	0.84	0.81
8	0.93	0.87	0.83	0.97	0.91	0.95	0.92		0.97	0.95
9	0.94	0.92	0.84	0.88	0.84	0.94	0.91	0.92		0.98
10	0.88	0.87	0.85	0.93	0.98	0.89	0.92	0.93	0.90	

从表 5.3.1 可以看出，所有区域的需求满足程度都在 78%到 99%之间，并且相邻区域之间的满足程度都高达 90%以上，说明该地区各区域的需求满足程度很好。

5.3.3 100 辆单车的分配求解

对于新投放的 100 辆共享单车，需要按照第二问中各地区需求比例分别分配到 1 到 10 地区。下表为根据附件 1 统计出的各地实际单车骑行人次。

表 5.3.2 各地实际单车骑行人次

到达 出发	1	2	3	4	5	6	7	8	9	10
1	0	229	95	101	134	105	92	113	87	96
2	122	0	222	115	116	120	115	121	113	137
3	116	127	0	230	97	107	112	113	140	116
4	110	98	141	0	247	134	118	117	113	116
5	117	122	110	107	0	248	132	115	100	113
6	133	119	131	126	120	0	221	114	122	112
7	111	125	116	130	112	120	0	231	97	110
8	124	146	121	130	129	132	134	0	237	122
9	99	112	109	121	103	134	104	110	0	232
10	122	94	112	126	118	101	132	225	121	0
总计	1054	1172	1157	1186	1176	1201	1160	1259	1130	1154

把上表中数据分别与附件 2 中数据一一作差，得到各地区在现有骑行基础上的的需求量。

表 5.3.3 现有骑行基础上的的需求量

到达 出发	1	2	3	4	5	6	7	8	9	10
1	0	11	24	22	11	21	23	14	25	6
2	13	0	3	14	4	6	24	6	8	3
3	9	5	0	21	13	10	14	24	18	10
4	6	17	7	0	14	10	14	24	11	3
5	11	11	14	9	0	25	6	5	3	19
6	25	9	12	14	24	0	23	18	18	4
7	18	15	9	16	23	18	0	6	19	25
8	10	22	24	4	13	7	11	0	7	6
9	6	10	20	17	20	9	10	9	0	5
10	16	14	19	10	3	12	12	18	13	0
总计	114	114	132	127	125	118	137	124	122	81

将数据按到达地区分类，得到 1 到 10 每个地区的需求量，计算每个地区占总需求量的比例，把 100 辆车按照比例分配，可以得出分配方案如下表：

表 5.3.4 分配方案

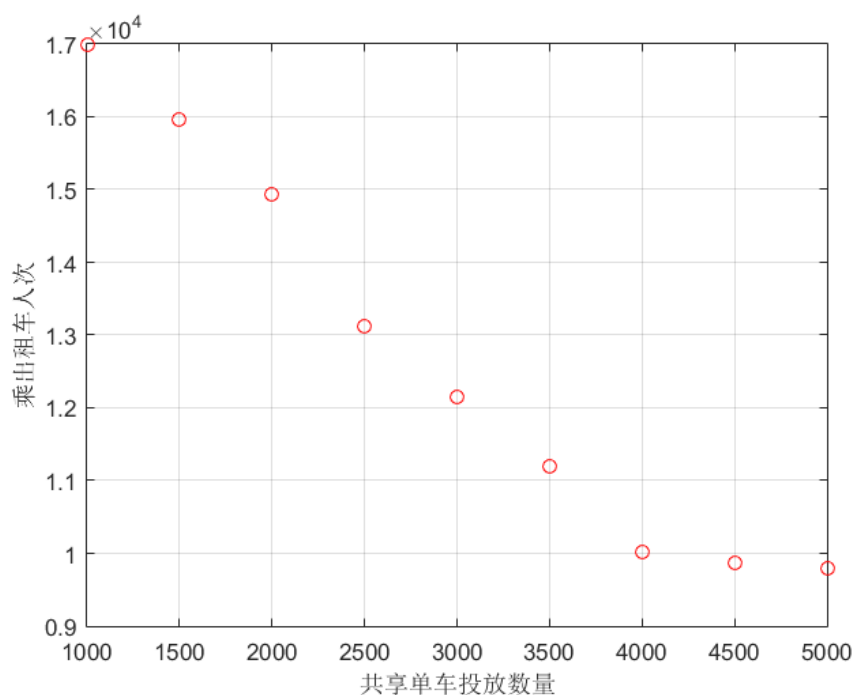
出发\到达	1	2	3	4	5	6	7	8	9	10
需求车量	114	114	132	127	125	118	137	124	122	81
比例	9.5%	9.5%	11.1%	10.6%	10.5%	9.9%	11.5%	10.4%	10.2%	6.8%
配比车辆	9	10	11	11	9	10	12	10	10	7

5.4 问题四模型的建立和求解

设该地区投入共享单车的数量为 x_i ，打车人次为 y_i 。为了大致分析 x_i 和 y_i 的关系，首先利用表 5.4.1 的数据做出 y_i 对 x_i 关系的散点图，如图 5.4.1 所示。

表 5.4.1 投入共享单车乘出租车人次数据

单车投放量	1000	1500	2000	2500	3000	3500	4000	4500	5000
打车人次	17000	15960	14940	13120	12150	11200	10017	9875	9796

图 5.4.1 y_i 对 x_i 关系的散点图

可以看出，随着共享单车投入量的增加，乘出租车人次减少，而且两者在 $x = 4000$ 之前有着很强的线性关系，而在 $x = 4000$ 之后呈现非线性关系。因此，我们认为需要分段建立模型，分别建立一元线性回归模型和一元非线性回归模型来对共享单车投入量对乘出租车人次的影响进行分析。

5.4.1 一元线性回归模型

5.4.1.1 模型的建立

由题意建立一元线性回归模型

$$y = \beta_0 + \beta_1 x + \varepsilon$$

其中, β_0, β_1 为回归系数, ε 是随机误差项, 总是假设 $\varepsilon \sim N(0, \sigma^2)$, 则随机变量 $y \sim N(\beta_0 + \beta_1 x, \sigma^2)$ 。

若对 y 和 x 分别进行了 n 次独立观测, 得到以下 n 对观测值

$$(y_i, x_i), i = 1, 2, \dots, n$$

这 n 对观测值之间的关系符合模型

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, i = 1, 2, \dots, n$$

这里, x_i 是自变量在第 i 次观测时的取值, 它是一个非随机变量, 并且没有测量误差。对应于 x_i , y_i 是一个随机变量, 它的随机性是由 ε_i 造成的。 $\varepsilon_i \sim N(0, \sigma^2)$, 对于不同的观测, 当 $i \neq j$ 时, ε_i 与 ε_j 是相互独立的。

用最小二乘法估计 β_0, β_1 的值, 即取 β_0, β_1 的一组估计值 $\hat{\beta}_0, \hat{\beta}_1$, 使 y_i 与 $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ 的误差平方和达到最小。若记

$$Q(\beta_0, \beta_1) = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$$

则

$$Q(\hat{\beta}_0, \hat{\beta}_1) = \min_{\beta_0, \beta_1} Q(\beta_0, \beta_1) = \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

显然 $Q(\beta_0, \beta_1) \geq 0$, 且关于可微, 则由多元函数存在极值的必要条件得

$$\frac{\partial Q}{\partial \beta_0} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\frac{\partial Q}{\partial \beta_1} = -2 \sum_{i=1}^n x_i (y_i - \beta_0 - \beta_1 x_i) = 0$$

求解可以得到

$$\begin{cases} \hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \\ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \end{cases}$$

其中， \bar{x} ， \bar{y} 分别是 x_i 与 y_i 的样本均值，即

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

5.4.1.2 模型的求解

根据表 5.4.1 的数据，对本模型直接利用 MATLAB 统计工具箱求解，得到参数估计值及其置信区间的结果见表 5.4.2。

表 5.4.2 回归模型参数估计值	
参数	参数估计值
β_0	1.9423×10^4
β_1	-2.376

将参数估计值代入得到：

$$y_i = 1.9423 \times 10^4 - 2.376 x_i$$

5.4.1.3 模型的检验

下面我们从以下几方面来对拟合效果进行分析。

1. 残差的样本方差

记残差

$$e_i = y_i - \hat{y}_i, i = 1, 2, \dots, n$$

残差的样本均值为

$$\bar{e} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) = 0$$

残差的样本方差为

$$MSE = \frac{1}{n-2} \sum_{i=1}^n (e_i - \bar{e})^2 = \frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

一个好的拟合方程，其残差总和应越小越好。残差越小，拟合值与观测值越接近，各观测点在拟合直线周围聚集的紧密程度越高，也就是说，拟合方程解释 y 的能力越强。

2. 判定系数（拟合优度）

对应于不同的值，观测值的取值是不同的。建立一元线性回归模型的目的，就是试图以 x 的线性函数来解释 y 的变异。接下来计算回归模型究竟能以多大的精度来解释 y 的变异。

记

$SST = \sum_{i=1}^n (y_i - \bar{y})^2$ ，这是原始数据的总变异平方和。

$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ ，这是用拟合直线可解释的变异平方和。

$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ ，这是残差平方和。

所以有

$$SST = SSR + SSE$$

这个公式可以同时从两个方面说明拟合方程的优良程度：

（1） SSR 越大，用回归方程来解释变异的部分越大，回归方程对元数据解释得越好；

（2） SSE 越小，观测值绕回归直线越紧密，回归方程对原数据的拟合效果越好。

因此，可以定义一个测量标准来说明回归方程对原始数据的拟合程度，这就是所谓的判定系数，有些文献上也称之为拟合优度。

判定系数是指可解释的变异占总变异的百分比，用 R^2 表示，有

$$R^2 = \frac{SSR}{SST} = (1 - \frac{SSE}{SST})$$

从判定系数的定义看，我们可以知道：

（1）当 $R^2 = 1$ ，有 $SSR = SST$ ，也就是说，此时原数据的总变异完全可以由拟合值的变异来解释，并且残差为零，即拟合点与原数据完全吻合；

（2）当 $R^2 = 0$ ，回归方程完全不能解释原数据的总变异， y 的变异完全由与 x 无关的因素引起，这时 $SSE = SST$ 。

3. 回归模型的线性关系检验

为了检验是否可以用 x 的线性方程式来解释 y ，可以进行下面的统计检验。记 y_i 关于 x_i 的总体回归系数为 β_1 ，则 F 检验的原假设 H_0 与备择假设 H_1 分别是

$$H_0: \beta_1 = 0, H_1: \beta_1 \neq 0$$

检验的统计量为

$$F = \frac{MSR}{MSE} \sim F(1, n-2)$$

若 $F > F_\alpha(1, n-2)$ ，则否定 H_0 ，接受 H_1 。这时认为 β_1 显著不为零，可以用 x 的线性关系来解释 y 。

4. 回归系数的显著性检验

回归系数的检验是考察每一个自变量对因变量的影响是否显著。换句话说，就是要检验每一个总体参数是否显著不为零。

t 检验的原假设 H_0 与备择假设 H_1 分别是 $H_0: \beta_1 = 0, H_1: \beta_1 \neq 0$

检验统计量为

$$t_1 = \frac{\hat{\beta}_1}{S(\hat{\beta}_1)}$$

若 $|t_1| > t_{\frac{\alpha}{2}}(n-2)$ ，则拒绝 H_0 ，认为 β_1 显著不为零。还可以确定 β_1 的置信度为 $1-\alpha$ 的置信区间为

$$\hat{\beta}_1 - t_{\frac{\alpha}{2}}(n-2)S(\hat{\beta}_1) \leq \beta_1 \leq \hat{\beta}_1 + t_{\frac{\alpha}{2}}(n-2)S(\hat{\beta}_1)$$

同样地，也可以对总体参数 β_0 进行显著性检验，并且求出它的置信区间。 β_0 的置信度为 $1-\alpha$ 的置信区间为 $\hat{\beta}_0 - t_{\frac{\alpha}{2}}(n-2)S(\hat{\beta}_0) \leq \beta_0 \leq \hat{\beta}_0 + t_{\frac{\alpha}{2}}(n-2)S(\hat{\beta}_0)$

5.4.1.4 模型的检验结果

表 5.4.3 R^2 检验

模型汇总

模型	R	R 方	调整 R 方	标准估计的 误差	更改统计量				
					R 方更改	F 更改	df1	df2	Sig. F 更改
1	.997a	.994	.992	226.191	.994	772.164	1	5	.000

a. 预测变量: (常量), 共享单车投放量。

由 R^2 检验可知， $R^2 = 0.994$ 说明回归方程能在 99.4% 的程度解释原数据的变异，即回归方程对原始数据的拟合程度较好。

表 5.4.4 F 检验

Anova^b

模型	平方和	df	均方	F	Sig.
1 回归	39505752.893	1	39505752.893	772.164	.000a
残差	255811.964	5	51162.393		
总计	39761564.857	6			

由 F 检验可知, $F = 1.0e + 004 * 0.0772$, $F > F_{\alpha}(1, n-2)$, 则否定 H_0 , 接受 H_1 。这时认为 β_1 显著不为零, 可以用 x 的线性关系来解释 y。

表 5.4.5 T 检验

系数 a

模型	非标准化系数		标准系数	t	Sig.	B 的 95.0% 置信区间	
	B	标准误差	试用版			下限	上限
1 (常量)	19422.964	230.195		84.376	.000	18831.230	20014.699
共享单车投放量	-2.376	.085	-.997	-27.788	.000	-2.595	-2.156

由 T 检验可知, $t_1 = 1.0e + 004 * 0.0084$, $t_0 = 1.0e + 004 * (-0.0028)$, 可以看出自变量对因变量的影响显著。且 β_0 和 β_1 的置信区间分别为 $[1.8831 \times 10^4, 2.0015 \times 10^4]$, $[-2.595, -2.156]$ 。

由以上的模型可知, 在共享单车投放量小于 4000 辆时, 乘出租车人次与共享单车投放量呈现显著的线性关系, 并且是负相关。即共享单车投放量越大, 乘出租车人次越少, 从而使打车市场的收益减小。

5.4.2 一元非线性回归模型

由图 5.4.1 我们可以观察得到, 在共享单车投放量大于 4000 辆时, 乘出租车人次与共享单车投放量没有明显的显著关系, 所以我们利用单车投放量大于 3000 辆的数据建立一元非线性回归模型来分析二者之间的关系。

用 MATLAB 分别对数据进行幂函数拟合、多项式函数拟合以及对数函数拟合。拟合好的函数分别为:

$$\text{幂函数: } y = 4.855 \times 10^5 x^{-0.4623}$$

$$\text{多项式函数: } y = -1.501 \times 10^{-9} x^4 + 2.422 \times 10^{-5} x^3 - 0.144 x^2 + 372.8 x - 3.421 \times 10^5$$

$$\text{对数函数: } y = -1.116 \times 10^4 \log_{10}(x) + 5.075 \times 10^4$$

三个拟合函数的图像如图 5.4.2 所示

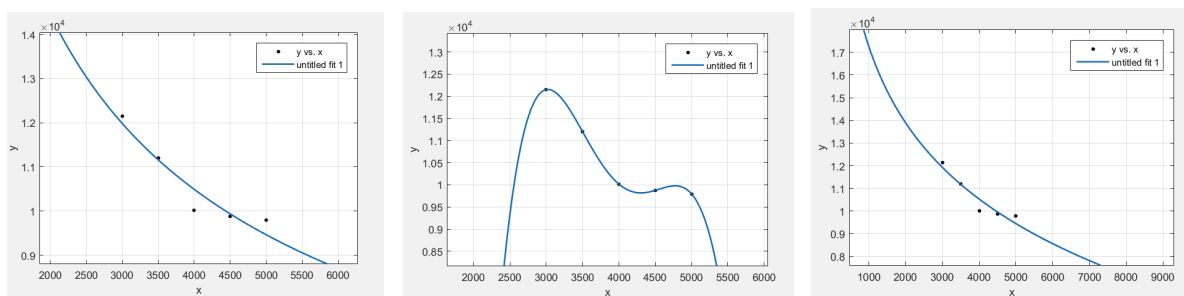


图 5.4.2 拟合函数

为了方便更好地识别哪一个函数能更好地拟合，我们将用 SPSS 输出的拟合优度制成表格来进行对比。从表 5.4.6 中可以看出，拟合优度从优到差排序为多项式函数>幂函数>对数函数。所以我们选择多项式函数来表示乘出租车人数与共享单车投入量之间的关系。

表 5.4.6 各函数拟合优度表

参数	幂函数	多项式函数	对数函数
SSE	3.688e+05	4.472e-19	4.425e+05
R-square	0.9137	1	0.8965
Adjusted R-square	0.8849	NaN	0.862
RMSE	350.6	NaN	384

通过分析多项式函数的图像，我们可以知道，该地区在共享单车投放量大于 4000 辆时，将单车投放量再增加 1000 辆对乘出租车人次无明显影响。继续增加共享单车投放量，则会使乘出租车人次大幅减小，从而使打车市场的收益大幅减小。

5.4.3 量化研究

对于第二小问收集实际数据进行量化研究，我们搜集了 2017 年 4 月份上海地区分时段共享单车使用人次以及乘出租车人次。我们将从时间维度分两个方面来研究上海地区共享单车的使用对打车市场的影响。

1. 固定日期分时段单车使用次数对成出租车人次数的影响

(1) 工作日——以 4 月 10 日和 4 月 17 日为例

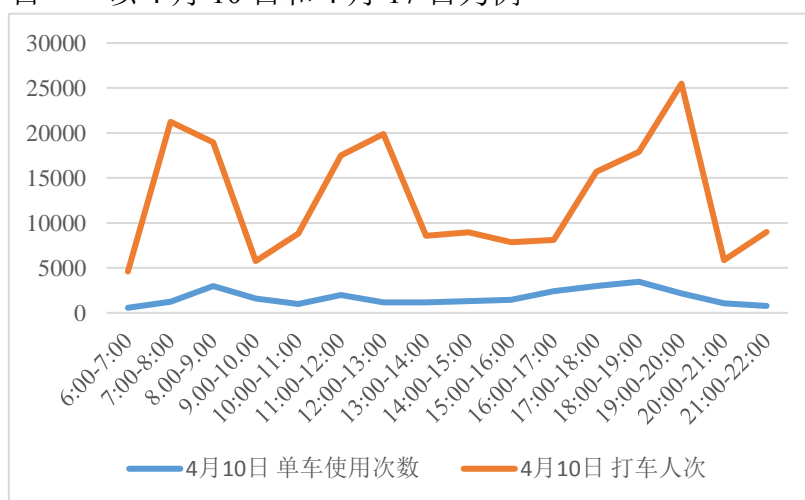


图 5.4.3 4 月 10 日单车使用次数及打车人次

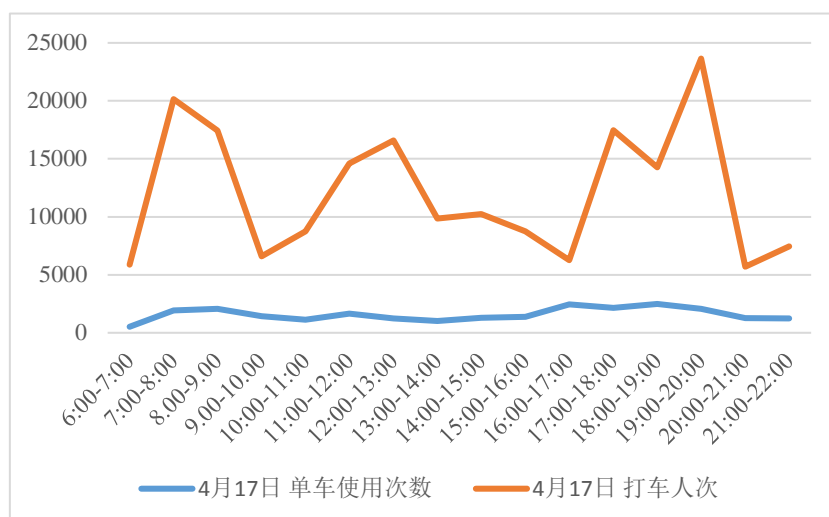


图 5.4.4 4 月 17 日单车使用次数及打车人次

观察图 5.4.3 和图 5.4.4，我们不难看出，两图的走势呈现近似的趋势，共享单车使用次数的高峰与打车人次出现的高峰几乎都在同一时间段内。例如图 5.4.4，打车人次第一个高峰出现在早上 7 点到 8 点，而共享单车使用人次也在 7 点到 8 点出现了第一个高峰。又如图 5.4.3，打车人次最后一个低谷出现在晚上 8 点到 9 点，而共享单车使用人次也在晚上 8 点到 9 点出现最后一个低谷。除此之外，我们还能观察出其具有很显著的潮汐现象，即早高峰和晚高峰单车使用次数和打车人次很多，其他时段较平稳。

(2) 节假日——以 4 月 2 日和 4 月 9 日为例

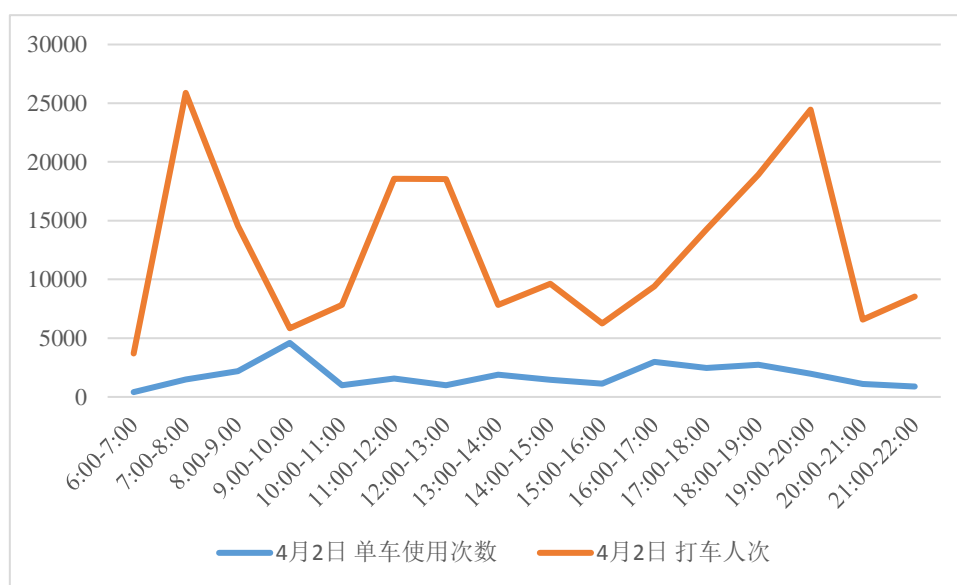


图 5.4.5 4 月 2 日单车使用次数及打车人次

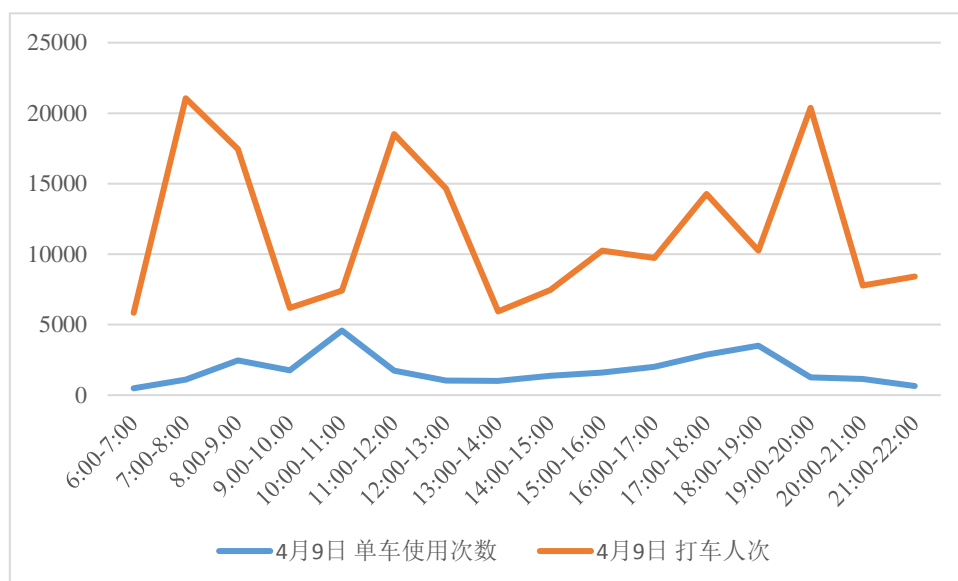


图 5.4.6 4 月 9 日单车使用次数及打车人次

观察图 5.4.5 和图 5.4.6，我们不难看出，两图的走势大体上依旧呈现近似的趋势，但潮汐现象不是很显著。例如图 5.4.6，在早上 7 点到 8 点，中午 12 点到 13 点，晚上 8 点到 9 点共享单车使用次数和打车人次出现了明显的三个高峰。

(3) 二者对比

通过对比图 5.4.3 和图 5.4.5，我们可以看出，打车人次的高峰在两天出现的时间段大致相同，都为早上 7 点到 8 点。而共享单车使用次数则有显著的差别，即工作日时高峰出现在早上 8 点到 9 点，而节假日时高峰则出现在上午 9 点到 10 点。

2. 固定时段分日期单车使用次数对出租车人次的影响

(1) 高峰期——以早上 8 点到 9 点为例

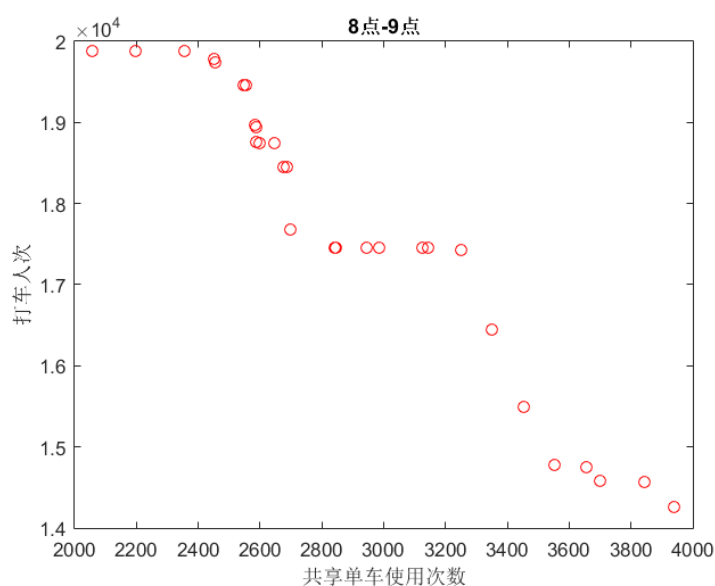


图 5.4.7 8 点-9 点共享单车使用次数及打车人次散点图

将共享单车使用次数作为自变量，打车人次数作为因变量，用 MATLAB 做出散点图。通过观察图 5.4.7，我们可以看到二者呈现线性关系，利用问题四第一小问建立的一元线性回归模型，我们可以得到衡量共享单车使用次数和打车人次数关系的函数表达式，函数表达式为

$$y = 27940.401 - 3.515x$$

拟合函数图像如图 5.4.8 所示

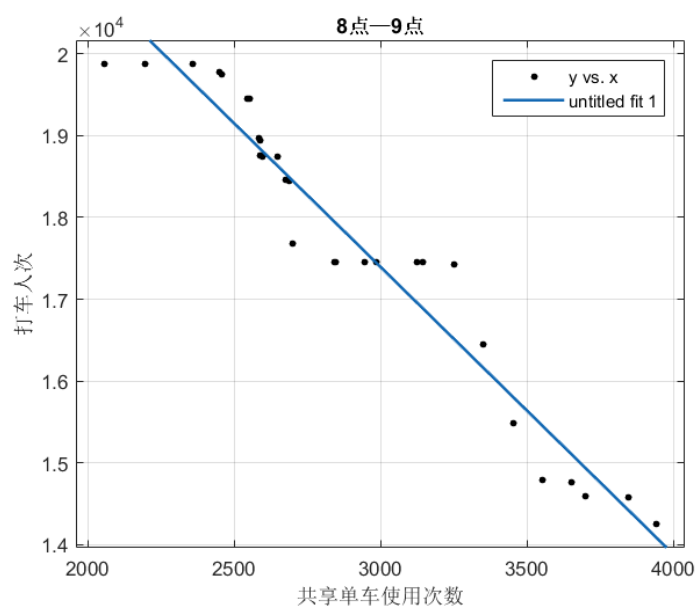


图 5.4.8 8 点-9 点拟合函数图

相应的检验值如表 5.4.7 所示。根据 R^2 我们可以得到拟合优度为 94.1%，即这个模型可以从 94.1%的程度上来解释共享单车使用次数和打车人数的关系，拟合效果较好。

表 5.4.7 检验参数表

检验参数	参数值
R^2 检验	0.941
F 检验	$1.0e+004*0.0458$
t_1 检验	$1.0e+004*0.0058$
t_0 检验	$1.0e+004*(-0.0021)$

(2) 低谷期——以上午 10 点到 11 点为例

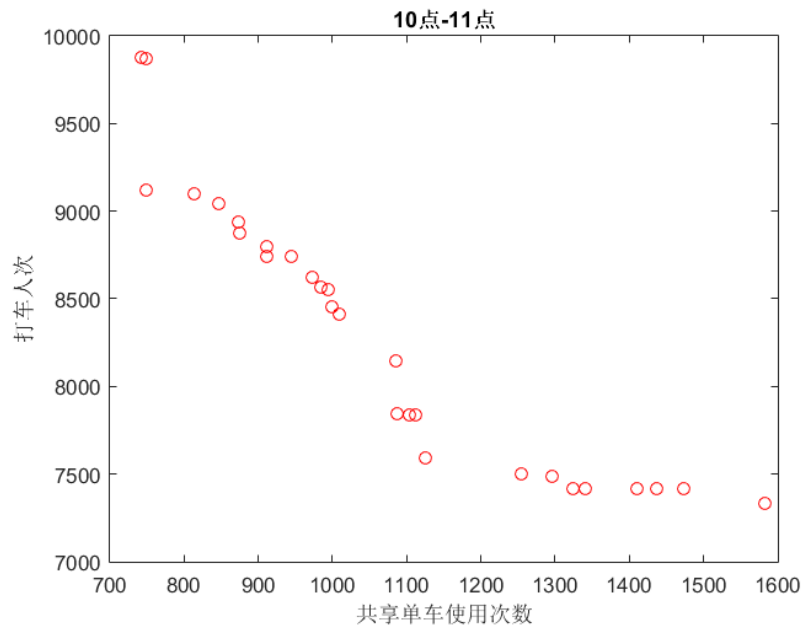


图 5.4.9 10 点-11 点共享单车使用次数及打车人次散点图

将共享单车使用次数作为自变量，打车人次作为因变量，用 MATLAB 做出散点图。通过观察图 5.4.9，我们可以看到二者呈现线性关系，利用问题四第一问建立的一元线性回归模型，我们可以得到衡量共享单车使用次数和打车人次关系的函数表达式，函数表达式为

$$y = 11536.193 - 3.029x$$

拟合函数图像如 5.4.10 所示。

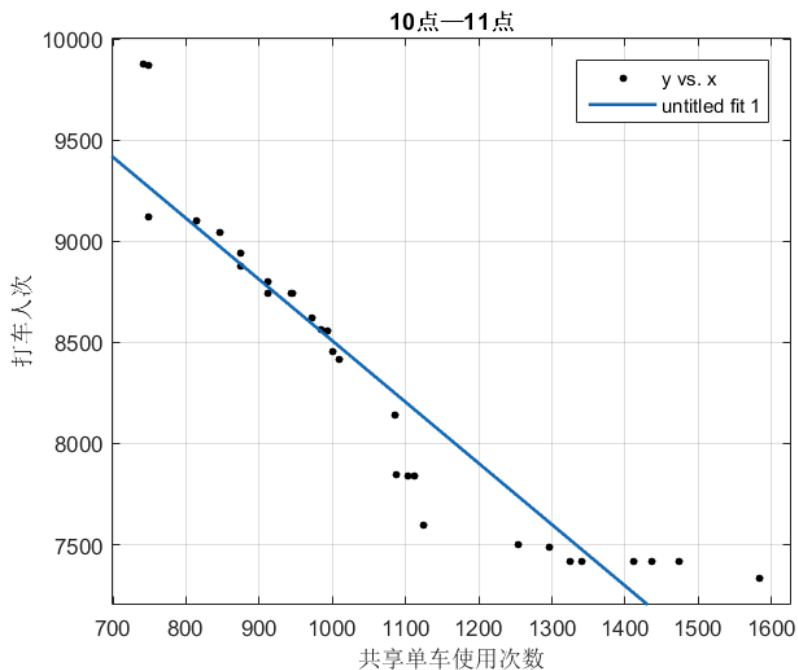


图 5.4.10 10 点-11 点拟合函数图像

相应的检验值如表 5.4.8 所示。根据 R^2 我们可以得到拟合优度为 86.3%，即这个模型可以从 86.3% 的程度上来解释共享单车使用次数和打车人数的关系，拟合程度较好。

表 5.4.8 检验参数表

检验参数	参数值
R^2 检验	0.863
F 检验	$1.0e+004*0.0182$
t_1 检验	$1.0e+004*0.0047$
t_0 检验	$1.0e+004*(-0.0012)$

六、模型的评价及灵敏度、敏感性分析

6.1 模型评价

6.1.1 动态优化模型的评价

1. 优点

- 本模型及其算法能够快速的找出局部最优解，在调度路径长度最短的基础上使乘客等待时间大大降低，为公共自行车调度问题的研究提供了可靠的依据。
- 本模型以遗传算法为基础，以 MATLAB 软件为工具，结合公共自行车调度优化问题特点，应用“初始静态优化+实时动态优化”的方法对多个连续静态调度问题进行求解，最终实现对动态需求调度模型的准确求解。
- 本模型给予长度最短和乘客等待时间最短两目标函数不同的权重，将多目标问题转化为单目标问题，简化了模型，也为模型求解提供了便利。

2. 缺点

- 本文 θ 的取值是经过多次试验得到的， θ 取值过大易造成模型无可行解， θ 取值过小则时间窗的约束力度不够。但是在实际中，乘客等待时间与调度车辆延误时间之间的关系是复杂的，仅仅用时间延误惩罚系数 θ 不足以完全表述两者之间的关系。
- 求解动态需求调度优化模型时，首先要对调度需求进行较高精度的预测，但是共享单车出行行为的随机性很大，影响因素也很多，因此调度需求的确定不是十分准确。
- 在动态需求调度优化模型中，每隔一段时间系统需求将发生变化，调度车辆有可能已经驶离关键点，处于关键点和原调度计划中下一租赁点之间，因此存在一定误差。

- 在动态需求调度优化模型中，每隔一段时间系统需求将发生变化，调度车辆有可能已经驶离关键点，处于关键点和原调度计划中下一租赁点之间，因此存在一定误差。

6.1.2 一元回归模型的评价

1. 优点

- 回归分析可以准确地计量共享单车与出租车出行次数之间的相关程度与回归拟合程度的高度，提高预测方程式的效果。
- 本模型中的线性回归反应的是准确的函数关系，可以将共享单车与出租车出行次数两变量量化，得出相对稳定的关系。

2. 缺点

- 本回归分析模型的拟合曲线具有人为特征性，因此会存在一定的误差。
- 在回归分析法时，由于实际一个变量仅受单个因素的影响的情况较少，因此本模型具有特定的覆盖范围，即共享单车对出租车的影响显著的时期。

6.2 灵敏度分析

问题二中，动态优化调度方案时采用确定初始调度需求的方法更新调度需求，更新结果显示地点 2, 5, 3 已经完成调度，对还未调度的地点进行动态优化，并综合已完成的调度路径，得到动态优化调度方案。

我们改变动态调度模型中的需求，更新调度结果如下表，得出新的调度优化方案，根据新方案的合理程度来检测模型的灵敏度。

表 6.2.1 新的调度需求

地区	调度需求	最佳时间窗	可接受时间窗	备注
3	—	—	—	取消
7	27	2:40–2:50	2:35–3:00	新增
6	—	—	—	取消
8	–12	2:25–2:40	2:20–2:45	新增
9	18	2:45–2:55	2:35–3:10	调度量变化

表 6.2.2 新需求下的优化调度方案

地区	到达时间	最佳时间窗	可接受时间窗	备注
1	2:15	—	—	
5	2:22'28"	2:20–2:30	2:15–2:40	已完成调度
8	2:27'46"	2:25–2:40	2:20–2:45	已完成调度
2	2:32'13"	2:30–2:40	2:30–2:50	已完成调度
7	2:41'56"	2:40–2:50	2:35–3:00	
10	2:52'44"	2:35–2:50	2:35–3:00	
9	3:02'44"	2:45–2:55	2:35–3:10	
1	3:14	—	—	

上表的动态优化调度方案中，大部分地点都能在最佳时间窗内完成调度，其中 9 和 10 地点的调度时间也在可接受时间窗内，此方案乘客等待时间较短。由此可知，在需求改变的情况下，优化调度模型依然能得出良好的优化方案，条件的改变对模型基本无影响。

七、参考文献

- [1]刘亚楠. 共享单车发展研究分析[J]. 时代金融, 2017, (08):251+254.
- [2]李敏莲. 共享单车市场调研与分析[J]. 财经界(学术版), 2017, (05):121-123.
- [3]徐鑫垚. 共享单车 App 后台管理系统的优化[J]. 电子技术与软件工程, 2017, (04):80-81.
- [4]冯天军. 公共自行车交通系统效果分析[D]. 吉林大学, 2016.
- [5]曹雪柠. 基于 IC 卡数据的公共自行车使用特性与动态调度优化研究[D]. 东南大学, 2016.
- [6]王恺. 公共自行车调度优化研究[D]. 西南交通大学, 2016.
- [7]江迎. 基于大数据的长沙市出租车时空分布研究[A]. 中国城市规划学会城市交通规划学术委员会. 2016 年中国城市交通规划年会论文集[C]. 中国城市规划学会城市交通规划学术委员会:, 2016:9.
- [8]李锦霞. 公共自行车调度优化研究[D]. 长沙理工大学, 2013.
- [9]史彩霞. 公共自行车系统运行数据时空分析及智能调度系统的研究[D]. 浙江工业大学, 2012.