

# 2021 第六届“数维杯”大学生数学建模竞赛论文

题 目                      中小城市地铁运营与建设优化设计

---

## 摘 要

轨道交通影响一个城市的运转效率。呼和浩特市在 2019 年建成地铁并投入使用，然而地铁的建设和维护需要消耗大量的财力和物力，需要考虑车厢节数、发车间隔、新站点选址这些指标，使其达到成本尽可能小，而乘客体验更优。所以我们就从这几个变量入手，进行优化模型的建立。

在第一个问题中，我们首先使用混合高斯模型对地铁客流量数据进行仿真，通过 Fisher 聚类方法对地铁运营的高峰、低峰时段进行划分。通过构建目标函数，加以条件限制得到 4-6 节车厢编组的建议，并给出了详细的发车时间间隔表。

第二问，依据该市热点图和拥堵路段，确定需要改进的路段。采用最小二乘估计进行建模，之后对找到重点节点，对重点区域进行加权最小二乘估计的建模，拟合出所需的路线。对于地铁盈利的确定则是考虑影响成本的主要条件，加以限制，用优化模型解决。

第三问基于高峰情况发生的背后原因以及错峰出行的实质意义，建立优化模型，利用算法求得结果，输出出行时间与目的节点对应表格。为疫情防控提出切实可行的方案。

第四问在综合考虑公交地铁互补出行以及现有快速路的分布下，给出了公交线路规划方案以及一些注意事项。

最后对于模型的部分闲置进行总结和反思，总结了局限性和展望。

**关键词:** 交通规划、仿真模拟、高斯混合模型、最优化方法、最小二乘法

# 目 录

<b>一、问题重述</b>	<b>1</b>
1.1 引言 . . . . .	1
1.2 要解决的具体问题 . . . . .	1
<b>二、名词解释与符号说明</b>	<b>2</b>
2.1 名词解释与说明 . . . . .	2
2.2 主要符号与说明 . . . . .	2
<b>三、问题分析</b>	<b>3</b>
3.1 数据分析 . . . . .	3
3.1.1 地铁客流随时间的分布研究 . . . . .	4
3.1.2 客流量数据的特征分析 . . . . .	5
3.2 乘客体验感指标 . . . . .	6
3.3 问题一的分析 . . . . .	7
3.4 问题二的分析 . . . . .	7
<b>四、模型假设</b>	<b>8</b>
<b>五、模型的建立与求解</b>	<b>9</b>
5.1 问题一的客流量模拟模型 . . . . .	9
5.1.1 时间度量的映射 . . . . .	9
5.1.2 高斯混合模型的建立 . . . . .	10
5.1.3 客流量数据的生成模型 . . . . .	10
5.1.4 客流量数据的生成过程 . . . . .	11
5.1.5 最优车厢数量的求解 . . . . .	12
5.1.6 最优发车间隔的求解 . . . . .	13
5.2 问题二的分析和求解 . . . . .	13

5.2.1	限制条件的推导 . . . . .	14
5.2.2	模型建立 . . . . .	16
5.2.3	考虑地铁运营收益 . . . . .	18
5.3	问题三的求解和分析 . . . . .	19
5.4	问题四的求解和分析 . . . . .	19
 <b>六、模型的评价与推广模型的评价与推广</b>		<b>20</b>
 <b>七、模型的改进</b>		<b>21</b>

# 一、问题重述

## 1.1 引言

呼和浩特市地处中国的华北地区、北部地区，建成区面积约为 260 平方千米，常住人口在 344 万左右。截止 2020 年 10 月，呼和浩特市已经开通运营路线 2 条，地铁里程总长约 49 千米，车站总计 43 站。

在工作日，地铁通常用于满足人们通勤的需求，这样符合地铁建设的初衷，缓解交通高峰时段的交通拥堵。呼和浩特市的城市总面积较大，但常住人口相对较少，人口基数小，会直接应该影响地铁的当日客流量，从而直接影响运营的成本。

本文的目的就是想从车厢数量和发车间隔的确定入手，继而进行新增地铁线路的选址，并尽可能保证地铁可以营利。最后要基于地铁和公交现状，提出新的地铁和公交互补的公交线路，普惠于呼和浩特人民。

## 1.2 要解决的具体问题

### 1. 当前的地铁运营是否合理？

呼和浩特市地铁的基本信息如表 1-1，错峰发车的时间间隔如 1-1 所示，考察网络数据和所给数据不难发现，尽管在官方报道的客流量数据中，地铁客流量在逐渐增加，2020 年 12 月 6 日，地铁客运量达到了 17 万人次，但是每天单向会发出 100 次，粗略的平均每列地铁的人次在 400 左右，载客率过低，也导致亏损十分严重。

综合运营成本和乘客的乘坐体验，我们需要给出合理的车厢节数和发

运营时间	地铁型号	承载量上限	高峰间隔	平峰间隔	低峰间隔
6: 00-22:00	6B	400	6min	10min	12min

**表 1-1:** 基础信息表

	早高峰	晚高峰
工作日	7:00-9:00	17:00-19:00
节假日	9:00-11:00	16:00-18:00

**表 1-2:** 早晚高峰时间表

车的时间间隔。

## 2. 建设新的线路选址在哪里？

当前的 2 号地铁线难以覆盖整个呼和浩特市，仅在一定程度上缓解了小部分的交通拥堵，并且人们并没有养成地铁出行的习惯，所以地铁处于亏损的状态。需要选择新的地点用于建设地铁，以及要求出可以让地铁盈利的乘坐人的数量。

## 3. 实现学生党和工作党的错峰出行，实现平峰目标该怎么做？

考虑到疫情的特殊背景，我们需要对不同站点的学生和工作者进行出行时间设定，以满足疫情防控的需要以及平峰的目标，为乘客提供更好的乘车环境。[6]

## 4. 如何对公交新增线路，实现地铁公交更好的互补？

对于一个城市来说，交通系统下各交通工具应该是相辅相成为该市的居民服务的。所以在公交的基础上建造地铁，之后又要在地铁的基础上完善公交。

# 二、名词解释与符号说明

## 2.1 名词解释与说明

1. **断面客流量**：在单位时间内，沿同一方向通过对到交通线路某断面你的乘客数量，即通过该断面你所在区间的客流量，分为上行断面客流量。
2. **最大断面客流量**：在单位时间内，通过轨道交通线路各断面的客流量的最大值。
3. **满载率**：反应车辆乘客满载程度的相对值，衡量车量利用程度的指标，可通过实际载客量与额定载客量之比求得。
4. **乘客到达率**：乘客在某单位时刻在某一站点下车人数，单位通常为人/min。
5. **地铁客运周转量**：乘客乘坐里程的总和。

## 2.2 主要符号与说明

序号	符号	符号说明
1	$M$	地铁车站的总数量
2	$N$	当日的发车次数
3	$t_i^j$	第 $i$ 辆地铁到达第 $j$ 个站点的时间
4	$\Delta t_i^j$	$t_{i-1}^j$ 到 $t_i^j$ 的时间间隔
5	$F(\Delta t_i^j)$	在 $\Delta t_i^j$ 时间间隔内的进站人数
6	$f_j(t)$	站点 $j$ 的乘客到达率函数
7	$f(t)$	地铁线路的乘客到达率
8	$\Delta(t)$	发车间隔
9	$\bar{F}$	地铁的最大荷载人数
10	$\eta_i$	第 $i$ 量车辆的载客率
11	$W_i$	乘客要乘第 $i$ 量车的等待时间
12	$K$	$k$ 取值为 1-43, 一号线二号线依次编号, 1a 为 1, 1t 为 20, 2a 为 1...
13	$S$	所有站点指标的集合, 如 1a、2u。
14	$t$	时间变量, 从 $t=0$ 为一天的开始
15	$Q_{in}(s, t)$	一名乘客在时刻 $t$ 从 $s$ 进站的概率
16	$Q_{out}(s, t)$	一名乘客在时刻 $t$ 从 $s$ 出站的概率
17	$L_i$	$i=1,2$ 地铁线路的总长度, $L_1, L_2$
18	$Z_i$	$i=1,2$ 地铁 $i$ 号线客运周转量
...	...	

### 三、问题分析

本次比赛的四个问题中, 本质上也都是在围绕一个核心的意义-如何让地铁运营方的收益和乘客的体验感同时可以到达一个更好的水平, 也就是如何更好的“以乘客为本, 兼顾效益”。

从乘客的角度, 影响体验感的变量会有: 平均等待时长、票价、地铁空间体验感等等。从运营者角度需要考虑的问题就是: 发车间隔、站点选址、票价定价、运营成本。

#### 3.1 数据分析

附件一给出了每一个地铁站的经纬度坐标, 容易发现, 一号线共有 20 个站点, 二号线共有 24 个站点, 二号线的交点为新华广场站, 在一号线中标号为 1h, 在二号线中标记为 2k, 地铁一号线的总运行时间在 45 分钟, 二号线的总运行时间在 47 分钟。同时根据当前的发车时间间隔数据, 可以求

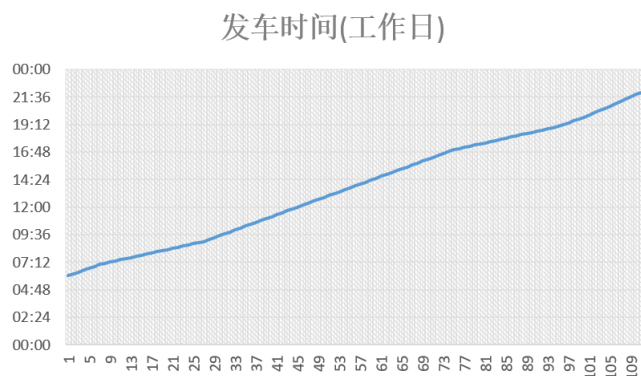


图 3-1: 发车时间图

出发车时间表，如下图3-1

附件二中给出了在 9 月 1 日到 9 月 14 日共计 14 天的，自早上 6 点到晚上 22 点 45 分的各站点进出站的数据。数据间隔为 15 分钟，以 9 月 1 日 6 点的 1A 为例，进站数据为 62 人，即在 6 点到 6 点 15 期间有 62 人刷卡进站。在本文中我们假设乘客的到达服从间隔内的均匀分布。在上例中，就是每分钟约有 4 人进站。同时我们也认为乘客的等待时间服从均匀分布，如果是在非高峰期，那么就是服从均值为 5，取值在  $[0,10]$  的均匀分布，如果在高峰期就是均值为 3，取值为  $[0,6]$  的均匀分布。

此处我们引入一个转移矩阵  $X_{ij}$ ，是由  $x_{ij}$  组成的，代表一名乘客从  $i$  站点上车，从  $j$  站点下车的概率。从  $i$  站出发到达  $i$  站的可能性我们记为 0，矩阵每行每列各概率值相加均为零，代表着从某一站上车一定会从某一站下车。

$$\begin{bmatrix} 0 & x_{12} & x_{13} & \dots & x_{1n} \\ x_{21} & 0 & x_{23} & \dots & x_{2n} \\ x_{31} & x_{32} & 0 & \dots & x_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & x_{n3} & \dots & 0 \end{bmatrix} \quad (1)$$

转移矩阵的数据可以由实际的运行的闸机数据获得，此处只能根据到站数据占总数据的比，来推测每个站点的热度最为权重，视作转移概率。对于进站的数据(非端点站)，认为其有 0.5 的概率向两个相反的方向。

### 3.1.1 地铁客流随时间的分布研究

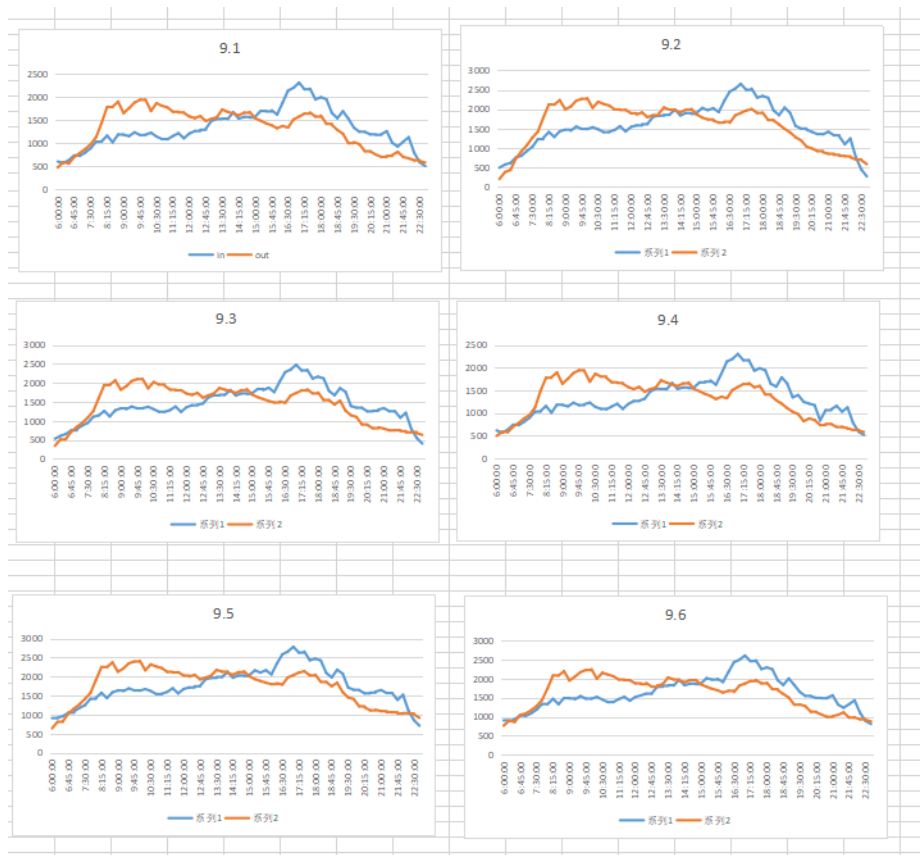


图 3-2: 客流量图

### 3.1.2 客流量数据的特征分析

将附件 2 中的数据进行一些处理得到每个取样时刻点的地铁站内客流总和，取样 9 月 1 日到 9 月 6 日，将数据绘图3-2, 在图中，橙色的线表示的是出站的数据，蓝色的是进站的数据。通过分析图片发现附件 2 中的客流量数据存在诸多明显不符合实际情况的地方，其中两处最明显的错误如下：

- 每天早上出站人数均远多于进站人数。例如对于 9 月 1 日的数据，截至早上 9 时，累计进站 11517 人次，累计出站 14591 人次，这在实际生活中是明显不可能出现的情况。
- 末班车在晚上 22 时由起点站发车，但是在末班车开走之后，仍有旅客进站。例如对于 9 月 1 日的数据，1 号线于 22 时由 1a 站和 1t 站分别发出当天的最后一班列车，但是在 22:15 分还有 8 人进入 1a 站，在 22:30 分还有 9 人进入 1a 站，在 22:15 分还有 10 人进入 1a 站，这些旅客不可能再坐上任何一趟地铁。



- 同时 6 天的数据边缘重合度较高，每一时点和下一个时点的变化量之间没有什么明显的变化，无论是平日还是节假日，都是有一个进站的高峰和一个出站的高峰。在这种现实情况下，乌鲁木齐地铁并不需要将工作日与节假日的高峰期与平峰期等分开。

忽视这些问题，我们我们在分析了附件 2 中数据的基础上，概括提炼出了如下的数据特征：

- 一天之内会形成两个较为平缓的客流高峰，分别位于 8:15 左右和 16:00 左右。并且早高峰的峰值会更高一些，说明早高峰的客流比较集中。晚高峰的客流相对缓和。工作日与非工作日的波动性差距不大，非工作日的数据会工作日的数据更上移一些。
- 每日客流约为 10 万人次，其中 9 月 10 日客流量最少为 77984 人次，9 月 7 日客流量最多，为 122495 人次。
- 出站高峰期在早上 9 时左右，出站客流量大约为 2000 人次每 15 分钟。
- 进站高峰期在晚上 18 时左右，进站客流量大约为 2500 人次每 15 分钟。
- 早上 9 时，1a-1e、2u 站的进站客流量较大，大约在 50-100 人次每 15 分钟每站。
- 晚上 18 时，1e、1f、1m、1p、1s、1t、2b-2i、2k、2q、2u 站的进站客流量较大，大约在 50-200 人次每 15 分钟每站。
- 早上 9 时，1m-1p、1s-1t、2b、2e-2k、2u 站的出站客流量较大，大约在 50-200 人次每 15 分钟每站。
- 晚上 18 时，1a-1b、1e-1f、1m、1p、2u 站的出站客流量较大，大约在 50-150 人次每 15 分钟每站。
- 从早上 10 时到晚上 17 时，1f、1m、1p、2q、2u 站的出站客流量持续较大，大约在 50-150 人次每 15 分钟每站。

### 3.2 乘客体验感指标

平均等待时间: 在前面我们已经陈述，在此处我们认为等待时间服从均匀分布，所以此时的平均等待时间的期望为 3 分钟，5 分钟，和 6 分钟。平均等待时间会影响乘客对地铁的选择，对确定发车间隔有着重要作用，参

考 [7]。

乘客舒适度体验：地铁车厢的拥挤程度会直接影响乘客的乘车舒适度，我们会沿车辆内乘客站立人员。在此处因为呼和浩特的地铁人员偏稀少，所以我们不对这部分加以展开，就选择一个合适的密度来进行计算。在此处，我们在求车厢节数的时候，极限站席密度会用  $7/m^{-2}$ 。这种情况下的拥挤程度为感到有些拥挤，站席范围有些突破。

### 3.3 问题一的分析

通过前面对于数据的初步观察，可以看出目前的运营模式下有几点不太贴合现实情况，所以我们初步可以认定目前发车方案的不合理。

1. 数据不合理需要重新进行仿真。2. 当前的高峰平峰时间的划分不合理，需要用 Fisher 聚类重新划分，参考 [8]。

在结局这两个问题的基础上进行车厢节数的划分和发车间隔的确定。

问题一首先需要对数据进行简单地处理，对于通过数据可以得出的结论进行总结和讨论。然后通过数据来推测最优车厢节数，以及最优发车间隔。最优车厢节数的推算，需要用到最大断面流量。最大断面流量的获得先需要对客流特征进行提取，建立客流到达率函数。通过到到达率函数用 Fisher 有序样本聚类分析，得到客流量高峰、低峰和平峰，用于错峰发车，划分调动时段。最后考虑乘客的等车时间成本和车辆的利用率用遗传算法确定发车间隔。再通过混合高斯模型进行数据的仿真，对模型进行检测。

### 3.4 问题二的分析

问题二需要根据呼和浩特的城市特点，提出合理的选址方案，让民众更加倾向于选择地铁这种出行方式，这样才能更大程度上缓解交通的拥堵情况。在这一部分，我们是以事实为依据，对于现实生活中实时路况信息进行调查，对于拥堵严重的地区，考虑用地铁缓解交通的压力。确定热点位置之后，通过最小二乘法确定路线，。

中国的城市轨道发展至今，地铁运营行业一直需要面对的课题就是：如何提高地铁运营企业的经济效益，如何降低运营成本，如何增加客票收入，实现“减亏”。

通常意义下定义的地铁运营企业的收支平衡，是指剔除了还本付息、折扣和大修基金之后的相对平衡，也就是在考虑收支平衡的时候不考虑前期的成本。地铁的前期建造成本以亿为单位进行计算，如果考虑完全成本，那么地铁盈利是基本不可能的，所以我们也同样不考虑建设成本，只计算电费、维修、人工、管理和相关税费，与票价冲抵之后的平衡。

首先对未来的客流量进行一定的预测。由于我们可收集到的数据有限，只有 6 组单日客流量数据以及题目中给出的仿真数据，官方数据如下所示，所以我们此处只引入集中计算未来客流量的简单模型。由于呼和浩特市地铁正在蓬勃发展，那么我们可以假设增速为一定常数来进行建模。在实际应用时，考虑  $Y$  时需要将工作日数据与节假日数据分开进行预测。具体公式如下所示

$$Y = Y \times (1 + \beta)^t \quad (2)$$

$Y$  为计划年度运量， $Y$  为基期年度运量， $\beta$  为年平均增长率， $t$  为年数，然后会根据这个数据进行预测到达多少人才会实现地铁运营的目标。

## 四、模型假设

1. 假设乘客的等待时间服从均匀分布，期望为发车间隔的 1/2。
2. 假设在手机的数据中，乘客的到达服从 15 分钟内的均匀分布。
3. 假设所给数据能推出的结论是合理的。
4. 假设在运营过程中，地铁按时发车和到达，排除故障等因素。
5. 网站的路面信息可以反映真实的交通拥堵状况。
6. 假设各区域人口服从均匀分布，在本题不考虑居民的分布情况。

日期	单日客流量
2019/12/29	58171
2019/12/30	57714
2019/12/31	82211
2020/1/1	119613
2020/10/31	162629
2020/12/5	166234
2020/12/6	171687

7. 在地铁选址的过程中，不考虑实际的地点是否可行。
8. 在考虑地铁的运营成本时，不考虑地铁的前期成本。

## 五、模型的建立与求解

### 5.1 问题一的客流量模拟模型

首先，我们分析并提炼附件 2 中客流量的数据特征，使用高斯混合模型 (GMM, Gaussian mixture model) 结合随机噪声 (random noise)，生成了 5 组仿真模拟数据。随后，根据地铁的发车间隔和运行速度，我们编写代码模拟 (simulate) 了整个地铁的运行过程。最后，我们统计了每趟列车的车载人数的变化情况，分析得出最优的车厢数量方案，并使用聚类算法 (cluster algorithm) 分段求解最优的发车间隔。

#### 5.1.1 时间度量的映射

我们设  $t$  是一个表示时间的变量。 $t$  的值加 1 表示过了一分钟。我们设定当  $t = 0$  时表示一天的开始时刻，则  $t = 720$  时表示一天的正午时刻， $t = 1440$  时表示一天的结束时刻。地铁的首班列车发车时间是早上 6 时，此时  $t = 360$ ，末班列车发车时间是晚上 22 时，此时  $t = 1320$ 。

我们考虑一个连续型的概率密度函数 (Probability Density Function)，不妨记为  $P(x)$ 。由于  $P(x)$  是一个定义在  $(-\infty, \infty)$  上的函数，而地铁一天的运行时间只有 16 个小时，因此我们需要对时间建立一个从  $(-\infty, \infty)$  到  $[360, 1320]$  的映射。设映射后的概率密度函数为  $Q(t)$ ，映射算子为  $T(\cdot)$ ，映射过程如公式 3 所示。映射过程的含义是，仅考虑  $P(x)$  在  $[360, 1320]$  上的取值，再做一个归一化操作。

$$Q(t) = T(P(t)) = \frac{1}{\int_{360}^{1320} P(x) dx} P(t) \quad (3)$$

其中， $\int_{360}^{1320} P(x) dx$  是归一化因子，确保概率密度函数  $Q(t)$  在区间  $[360, 1320]$  上的积分值为 1。

同理，对于离散型的概率密度函数  $Q(s)$ ，设其定义域为集合  $S$ 。映射算子  $T(\cdot)$  的定义如公式 4 所示。映射过程的含义是，仅考虑  $P(x)$  在集合  $S$

上的取值，再做一个归一化操作。

$$Q(s) = T(P(s)) = \frac{1}{\sum_{x \in S} x} P(s) \quad (4)$$

### 5.1.2 高斯混合模型的建立

高斯混合模型 (GMM, Gaussian mixture model) 是多个高斯分布 (Normal Distribution) 的叠加。形式化地，一个均值为  $\mu$  方差为  $\sigma^2$  的高斯分布，其概率密度函数  $P(x; \mu, \sigma^2)$  如公式5所示。

$$P(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (5)$$

现有  $N$  个高斯分布，其均值分别为  $\mu_1, \mu_2, \dots, \mu_N$ ，其方差分别为  $\sigma_1^2, \sigma_2^2, \dots, \sigma_N^2$ ，则其概率密度函数分别为  $P_1(x; \mu_1, \sigma_1^2), P_2(x; \mu_2, \sigma_2^2), \dots, P_N(x; \mu_N, \sigma_N^2)$ 。设每个高斯分布发生的概率为  $\alpha_1, \alpha_2, \dots, \alpha_N$ ，则高斯混合模型  $P_0(x)$  可按照公式6构建。

$$P_0(x) = \sum_{i=1}^N \alpha_i P_i(x; \mu_i, \sigma_i^2) \quad (6)$$

其中， $\sum_{i=1}^N \alpha_i = 1$ 。

### 5.1.3 客流量数据的生成模型

我们定义二维离散型的概率密度函数  $Q_{in}(t, s)$  表示一名旅客在时刻  $t$  从  $s$  站进站的概率，同理用  $Q_{out}(t, s)$  表示一名旅客在时刻  $t$  从  $s$  站进站的概率。其中  $t$  的取值范围是区间  $[360, 1320]$  中的整数， $s$  的取值范围为集合

$$S = \{1a, 1b, \dots, 1t, 2a, 2b, \dots, 2x\}$$

我们考虑  $Q_{in}(t, s)$  和  $Q_{out}(t, s)$  的边缘概率密度函数 (Marginal Probability Density Function)

$$R_{in}(t) = \sum_{s \in S} Q_{in}(t, s)$$

$$R_{out}(t) = \sum_{s \in S} Q_{out}(t, s)$$

$R_{in}(t)$  和  $R_{out}(t)$  反映了在  $t$  时刻所有站点的进站概率和出站概率，是服从混合高斯分布的概率密度函数，其形式化定义如公式7所示。

$$\begin{aligned} R_{in}(t) &= T \left( \max \left( \sum_i \alpha_i P_i(t; \mu_i, \sigma_i^2), c \right) \right) \\ R_{out}(t) &= T \left( \max \left( \sum_i \alpha_i P_i(t; \mu_i, \sigma_i^2), c \right) \right) \end{aligned} \quad (7)$$

其中， $c$  是一个介于 0 到 1 之间的小数，定义了每个时刻的最小概率。这样做的目的是，防止高斯分布的“ $3 - \sigma$  原则”导致某些时刻的概率近乎于 0。

在  $t_0$  时刻， $Q_{in}(t_0, s)$  和  $Q_{out}(t_0, s)$  反映了在  $t_0$  时刻站点  $s$  的进站概率和出站概率，也是服从混合高斯分布的概率密度函数，其形式化定义如公式8所示。

$$\begin{aligned} Q_{in}(t_0, s) &= T \left( \max \left( \sum_i \alpha_i P_i(s; \mu_i, \sigma_i^2), c \right) \right) \\ Q_{out}(t_0, s) &= T \left( \max \left( \sum_i \alpha_i P_i(s; \mu_i, \sigma_i^2), c \right) \right) \end{aligned} \quad (8)$$

#### 5.1.4 客流量数据的生成过程

我们构建每时刻进出站概率的高斯混合模型，如表5-3所示，每个站点进出站概率的高斯混合模型，如表5-4所示。

**表 5-3:** 每时刻进出站概率的高斯混合模型

时刻	$t$	进站 GMM		出站 GMM	
		权重	高斯分布	权重	高斯分布
6 时 - 22 时	[360, 1320]	0.4	$P(t; 8, 0.5)$	0.4	$P(t; 8, 0.5)$
		0.4	$P(t; 18, 0.5)$	0.4	$P(t; 18, 0.5)$
		0.2	$P(t; 21, 0.2)$	0.2	$P(t; 21, 0.2)$

根据表5-3和表5-4的高斯混合模型，我们计算出每时刻每站点进出站的概率。

我们按照如下步骤生成模拟数据：

表 5-4: 每个站点进出站概率的高斯混合模型

时刻	$t$	特别热点站 GMM		热点站 GMM		普通站 GMM	
		权重	高斯分布	权重	高斯分布	权重	高斯分布
6 时 - 14 时	[360, 840]	0.7	$P(t; 8, 0.5)$	0.6	$P(t; 8, 1)$	0.5	$P(t; 8, 2)$
		0.3	$P(t; 10, 0.5)$	0.4	$P(t; 10, 1)$	0.5	$P(t; 10, 2)$
14 时 - 22 时	[840, 1320]	0.7	$P(t; 18, 0.5)$	0.6	$P(t; 18, 1)$	0.5	$P(t; 18, 2)$
		0.3	$P(t; 21, 0.5)$	0.4	$P(t; 21, 1)$	0.5	$P(t; 21, 2)$

1. 根据图??中的概率选择一个单元格，记为  $(t, s_f)$ ，表示旅客在时刻  $t$  从站点  $s_f$  进站；
2. 根据时刻  $t$  的概率（图??中时刻  $t$  的那一行）选择一个单元格，记为  $s_t$ ，表示旅客想要前往站点  $s_t$ ；

于是，我们得到了一个形如  $(t, s_f, s_t)$  的元组，表示一名旅客在时刻  $t$  从站点  $s_f$  进站，前往站点  $s_t$ 。由于每日的客流量约为 10 万人次，因此我们生成 10 万条左右这样的元组，以模拟地铁一天的运行情况。

### 5.1.5 最优车厢数量的求解

我们根据模拟数据，使用 Python 程序实时模拟地铁一天的运行情况，具体的模拟过程如下：

1. 确定一天所有列车的发车时刻；
2. 在时刻  $t$ ，将所有站点上车的旅客加入每个站点的等待队列  $L(s, up)$  或  $L(s, down)$  中，两个等待队列分别代表上行和下行两个方向；
3. 在时刻  $t$ ，更新所有列车状态；
4. 在时刻  $t$ ，若有列车  $r$  进站  $s$ ，将该站点等待队列  $L(s, d)$  中的所有旅客，移入该列车的车载旅客列表  $R(r, t)$  中，并将  $R(r, t)$  中所有在该站下车的旅客移除；
5. 时刻  $t$  加一，跳回步骤 2，直到所有列车结束；

我们使用最大断面客流量确定最大车厢数，最大断面客流量的定义如公式9所示。最大断面客流量的含义是，所有列车在所有时刻的运行过程中，某列车某时刻的车载人数的最大值。[3]

$$W = \max_{r,t} R(r, t) \quad (9)$$

根据最大断面客流量，我们可以计算出最多需要的车厢数量，这也是最优的车厢数量，其定义如公式10所示。

$$C = \left\lfloor \frac{W}{400} \right\rfloor \quad (10)$$

根据模拟数据，我们分析得出，在早上 7:36 从 1a 站发出的列车，从 1g 站运行到 1h 站的途中，达到最大断面客流量 852 人，最优车厢节数为 3 节。

但考虑到实际情况和乘坐舒适性，我们并不建议使用 3 节车厢的编组方式。在我们的认知范围内，没有城市采用 3 节车厢的编组方式，我们建议采用 4-6 节的编组方式最为合适。

#### 5.1.6 最优发车间隔的求解

目前地铁采用高峰期发车间隔 6 分钟、平峰期发车间隔 10 分钟、晚上 20 点之后发车间隔 12 分钟的发车模式。参考 [5] 结合实际情况，我们对发车间隔提出了如下规定：

1. 相邻两趟列车的发车间隔不得少于 3 分钟，以防发生安全事故；
2. 相邻两趟列车的发车间隔不得多于 15 分钟，地铁间隔时间过长影响乘客的出行体验，特别是对于外地来呼和浩特旅游的旅客；
3. 发车间隔采用动态调整的方式，在从高峰期向平峰期过渡的阶段，发车间隔可依次变大，例如从 3 分钟一趟变为 6 分钟一趟，再变为 10 分钟一趟；
4. 以 30 分钟为最小考虑跨度，即 6:00-6:30、6:30-7:00 每 30 分钟区间采用相同的发车间隔；

我们根据每时刻的进站客流量，结合实际情况，得出了表5-5的发车间隔时刻表。

## 5.2 问题二的分析和求解

由于题目中只要求我们选择地址，但是因为同时存在多个热点区域，所以为了缓解地面交通压力，降低碳排放，现拟建地铁三号线。优化地铁规划



表 5-5: 发车间隔时刻表

时刻	发车间隔
6 时 - 7 时	10 分钟/趟
7 时 30 分 - 8 时	5 分钟/趟
8 时 - 9 时	3 分钟/趟
9 时 - 9 时 30 分	5 分钟/趟
9 时 30 分 - 11 时	10 分钟/趟
11 时 - 13 时	8 分钟/趟
13 时 - 16 时 30 分	10 分钟/趟
16 时 30 分 - 17 时	6 分钟/趟
17 时 - 18 时 30 分	3 分钟/趟
18 时 30 分 - 19 时	5 分钟/趟
19 时 - 20 时	8 分钟/趟
20 时 - 21 时	10 分钟/趟
21 时 - 22 时	15 分钟/趟

建设是建设新地铁线路前必须解决的问题。因为科学的规划建造方案，才能形成有层次性、稳定性的地铁线网结构。

5.2.1 限制条件的推导

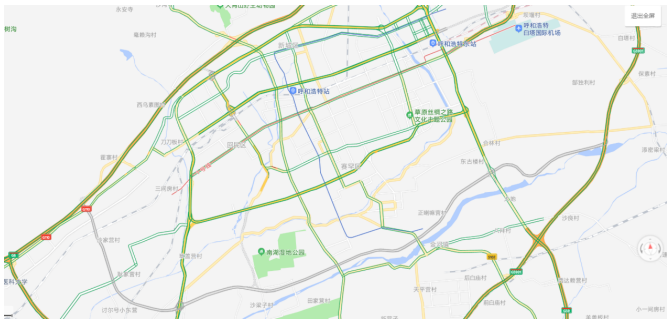


图 5-3: 呼市地图

首先对该地区进行总体分析，下图 5-3为该地区地图，通过对地图上主要交通路线分布情况呼和浩特主要交通路线聚集在城市中部以及西北方向。

如下 5-4为城市热力图，可以在一定程度上影响该市的交通状况，颜色越深的地方越有可能人员聚集，距离中心较远区域地广人稀，且自然山脉地形限制，暂不考虑将地铁延拓到这类相关区域。市区人员出行人流、车流

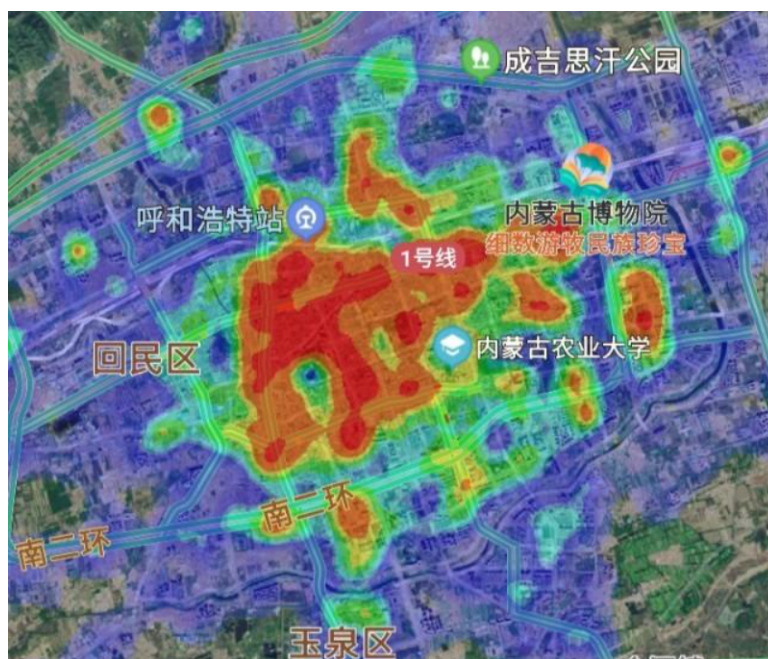


图 5-4: 城市热力图

量最大一般在市中心，副中心，居民区，工业区等。



图 5-5: 拥堵图

通过实时监测路况拥堵程度，提取地面车辆密度信息。如下图5-5，是以兴茂家园附近路线为例，其中绿色到黄色再到红色分别表明畅通到拥挤不同程度的路况。由于本文旨在缓解交通压力，减少碳排放，因此地面车辆密度的高低可以很好的反应相关信息。通过对路况信息的监控，拥堵区域集中在小区、学校、大型市场附近。

综上所述，我们可以大致给出关于地铁三号线的建设的初步考虑：1. 穿过城市中心。2. 为缓解地面交通压力，要在尽可能多的途径拥堵路段或其附近区域。3. 由于建造维护成本高，暂不考虑覆盖偏远地区。

### 5.2.2 模型建立

1. 结构基础在三线地铁的基本结构中，三角形结构最为稳定，根据《成都地铁规划》中，换乘次数、换乘压力、线网吸引力覆盖强度，达到最优。在此基础上我们寻找需要覆盖的重要节点。

2. 引入重要节点在确定了线路整体趋势的基础上，地铁规划又一重要因素在于停车站点的设置。设有  $k$  个节点，用集合  $S$  表示。为有效控制建设成本，我们设置  $k < 20$ ，超过时将次要点淘汰。为此我们按照重要性优先级顺序先后考虑如下三类节点：

id	经度	纬度
1	111.6455	40.854321
2	111.713097	40.847481
3	111.691971	40.844742
4	111.680588	40.840221
5	111.668419	40.839783
6	111.671967	40.835911
7	111.663864	40.837093
8	111.697327	40.837112
9	111.707542	40.834732
10	111.704674	40.83046
11	111.700783	40.830231
12	111.693687	40.828114
13	111.647552	40.793724
13	111.644638	40.804459
14	111.647697	40.798195
15	111.670311	40.798656
16	111.695641	40.798409
17	111.710930	40.807781

图 5-6: 经纬度图

I：地面拥堵点监测地面实时路况时，仅保留主要日常交通要道监测对象排除了拥堵的立交桥，快速内环，国道，省道的情况。在此基础上排除附近 1km 内有地铁的路段（假定通过一些措施可以对地面交通进行分流，提高地铁出行率）。通过观察 24 小时拥堵情况变化，利用 Python 提取出城市中心区域 17 个高拥堵点的经纬度。如 5-6所示：对于给定的关键大交通流量节点，以及学校工作聚集地，利用 requests 函数获取关键节点的经纬度信息。基于现有节点信息

II：换乘站点用于和 1、2 号线衔接，建造更高覆盖率、更方便快捷的地铁网。站点选择参考进出站热度，以及利用编程软件计算出的待分流区域到各个站点三维坐标轴下距离平方和，为参考依据，最后选择 1j 和 2i 作为 3 号线与 1/2 号线的换乘站点。

III：过渡/其他站点若两个站点之间距离较远，可以选择中间新增一个站点，满足更多乘客的出行需求。若地铁线附近存在有标志性建筑（商圈、公园等），可以考虑增加相关站点。此站点的选择可以根据初步得到的线路进行进一步优化，目前暂时无法确定。

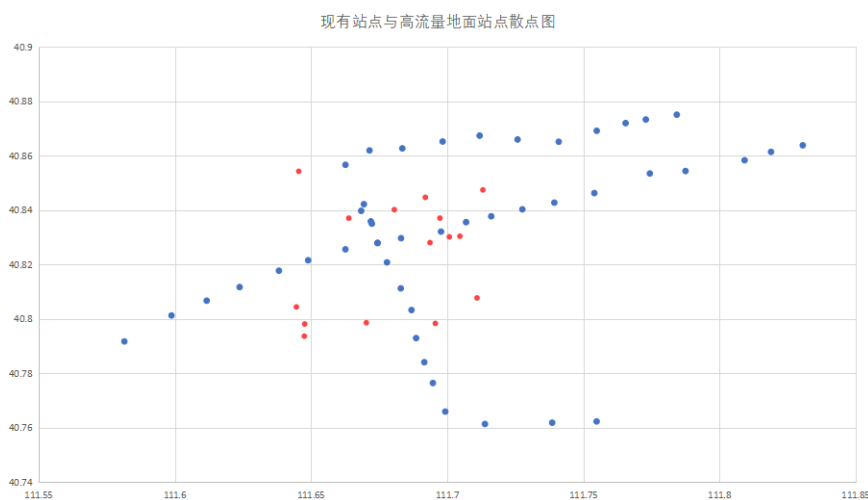


图 5-7: 选址散点图

3. 模糊规划为简化计算，文中通过经纬度信息近似距离信息，假设在平面直角坐标系下，该信息依旧有较高的参考价值。其中蓝色代表现有 1/2 号线对应站点红色代表待分流拥堵点 5-7。

(1) 直接拟合从图中可以看到标出的点（出去左上角的点）呈现明显的圆环趋势，因此考虑直接拟合环形地铁，但这类地铁制造、维护成本较高，在接下来尝试是否有更加接近直线的地铁建造方案。

(2) 线性规划在考虑近似直线地铁建造规划时，首先想到利用最小二乘法进行线性回归。这种方法可以拟合出一条直线到各散点的距离差平方和最小。但是这种方法的弊端在于使得各个点所含的信息量相同，与实际情况出入较大。我们应该考虑尽可能地接近人流量大的拟合点，为了优化现有的线性规划模型，我们考虑在进行最小二乘法距离求和时，给每个点的

距离平方引入权重系数。权重的确定采用打分系统，80% 由交通拥堵程度  $M$  决定，20% 由周边建筑信息  $N$  决定。 $M$  的确定：在监控 24 小时路况时，针对每小时拥堵情况进行累加。

$$W_i = 0.8 * M_i / (\text{sum}(M_i)) + 0.2 * N_i / (\text{sum}(N_i)) \quad (11)$$

4. 实际规划上部分线路的确定仅是提供模糊规划方案，实际线路确定应当围绕所给直线结合实际地形、地质等多方面因素综合考虑。由于信息有限，这里仅提供建议。(1) 互补原则：和现有机动车快速干道没有长距离重复 (2) 低成本性：避免大量拆迁建筑等等

### 5.2.3 考虑地铁运营收益

我们把整跳线路上的点转化为 1-43 的数字，根据轨道交通的特性， $i$  到  $i+1$  站点间的断面总客流量为  $Q_{i,i+1}$ ，是每一个站点断面数据之和。此处我们只考虑一号线，二号线同理，即取值点为 1-20. [2]

$$Q_{i,i+1} = \sum_{m=1}^i \sum_{n=i+1}^s q_{m,n} (1 \leq i \leq s-1) \quad (12)$$

引入呼和浩特地铁票价关于距离的函数，记为  $p(L_{m,n})$ ,  $L_{m,n}$  表示  $m$  站和  $n$  站之间的距离. 具体分段函数的取值如下：单向的轨道运营总收入可求。

$$P_{total} = \sum_{m=1}^{s-1} \sum_{n=2}^s p(L_{m,n}) q_{m,n} (1 \leq m < n \leq s) \quad (13)$$

路况	畅通（绿色）	较拥挤（黄色）	拥挤（红色）
计数	0	1	1.5

乘坐里程（千米）	票价（元）
0-5	2
5-10	3
10-15	4
15-21	5
21-28	6
>28	每 10 千米增加一元

双向的只要在此基础上再将  $m$  与  $n$  调换一下顺序。

$$P_{total} = \sum_{m=1}^{s-1} \sum_{n=2}^s p(L_{m,n})(q_{m,n} + q_{n,m})(1 \leq m < n \leq s) \quad (14)$$

参考 [4] 运营成本我们考虑人工成本与电力成本两个方面，这两个变量均当日的班次数量有关。当日发车次数为  $N$ ，当日的行驶总里程是  $L * M$ ，电力成本系数为  $k$  元/公里。人员成本数据应为一个确定的已知的数据，记为  $s_0$ ，不同数量的乘客对应车辆有不同程度的损耗，那么在这种情况下我们引入一个列车周转量的变量  $\sum_{m=1}^{s-1} \sum_{n=2}^s (q_{m,n}L_{m,n} + q_{n,m}L_{m,n})$  上下行两个方向的运营成本记为  $F_{cost} = f_2(L * M, \sum_{m=1}^{s-1} \sum_{n=2}^s (q_{m,n}L_{m,n} + q_{n,m}L_{m,n}), s_0)$

### 5.3 问题三的求解和分析

可以用 (起始节点, 目的节点) 表示一条出行轨迹，放入新集合中。为简化计算可以将终节点到达时间视作上班时间，中间经过此时间问题变为在已知起点、终点节点坐标以及路程中间经过的所有路段条件下，调节到达时间，使得各个节点流量每个时间控制在相对稳定的情况下。

$x_a^{pq}$ : 乘客  $x$  的起始乘车点为  $p$  点，终止点是  $q$ , 到达时间为  $a$

$$I_{tu}(x_a^{pq}) = \begin{cases} 0 & t & x & u \\ 1 & t & x & u \end{cases} \quad (15)$$

可通过列车的进出站点以及进出站时间，以及乘客刷卡乘车进出站节点与时间求出每个  $x$  对应的不同他  $t, u$  下示性函数的值。

此时目标函数变为

$$\min_{t,u} \max_i \sum_x I_t(x_a^{pq}) \quad (16)$$

该模型的限制条件在于对  $a$  的调整， $a$  每次调整的幅度  $\Delta_t$  为发车间隔，由于此时为高峰时间段，所以  $\Delta_t$  均为 6 分钟。本文建议采用贪心算法把所有合理可调整的变化大小进行遍历，最终可求得各个点最优解。

### 5.4 问题四的求解和分析

在充分考虑呼和浩特市城市快速路的基础之上，你能否提出一个地铁和公交互补的若干新增公交线路，以满足更多人员高峰出行时期的出行效

表 5-6: 列车发车时刻与到站点对应表

发车时间	6:30	6:40	6:50	7:00	7:06	7:12	7:18	7:24	7:30
到站节点	1j	2a	1c	1g	1i	1l	2o	2e	2b
	2p		1m	1r	2c	2t		2i	
			2w						
发车时间	7:36	7:42	7:48	7:54	8:00	8:06	8:12	8:18	8:24
到站节点	1e	2j	1b	1o	1k	1d	2d	2f	1h
	2u		1f				2l	2v	
发车时间	8:30	8:36	8:42	8:48	8:54	9:00	9:10	9:20	9:30
到站节点	1a	1s	1n	1p	2h	2k	2q	2g	1t
	1q	2r		2n			2s	2m	
				2x					

率

## 六、模型的评价与推广模型的评价与推广

### 1. 问题分析:

将模型进行数值计算，并与附件中的真实采样值（进行列表或图示）比较。对误差进行数据分析，给出误差分析的理论估计。在地面路况监控的过程中，我们发现在距离城区较远的地方也会存在拥堵情况，此时建造地铁成本较大，因此考虑利用灵活的公交线路对地铁未抵达区域进行互补。由于城市快速路主要位于呼和浩特市的中部偏东北部，因此考虑在西南部地区安排公交线路。

### 2. 建造原则

在确定公交线路是需要遵循如下步骤：

STEP1：确定带规划区域大致范围，标出主要出行线路和道路等级。

STEP2：计算周围居民每天出行量与公交乘车出行意愿，在此基础上确定公交为环线或双向单线。

STEP3：找到最优最大覆盖公交路径，极大化公交居民出行意愿。

此外还要注意如下细节：

(1) 线路长度适中，停靠站点数量合理



(2) 做好与主干线的衔接，提供更多的换乘可能性

(3) 结合周边环境、自然进行规划设计

### 3. 总结反思

在实际情况中，公交选线一般将提案与实际情况相结合，按照经验与试运营情况权衡出合适的线路。关注城市的发展，与客流数据的动态变化，找到合适的方案。

## 七、模型的改进

1. 实际情况反思。在现有地铁经过的站点附近 1km 为半径的环形区域内，路面交通情况依旧不容乐观，存在着较多时段的拥堵，人们在选择出行方式时，更偏好于地面交通。因此在可以考虑进行低碳出行宣传，以及制定相关政策，对地铁出行方式的偏斜。以便更好地缓解交通压力和碳排放。

2. 在相关数据的采集方面可以更加完善，在站点设置上可以通过调查问卷方式了解当地居民的看法，合理化地铁规划。同时可以对不同区域进行细分，基于更多的信息考虑各个区域的权重。



## 参考文献

- [1] 陈瑜. 地铁运营成本特性分析. 现代城市轨道交通, 4, 91–93.
- [2] 朱顺应, 刘滨, 刘斌, 杨友珍, AND 王红. 考虑运营收益与车厢空间感知的轨道列车发车间隔优化.
- [3] 刘岩, 张宁, AND 邵星杰. 城市轨道交通断面客流短时预测. 都市快轨交通 28, 001 (2015), 77–81.
- [4] 宋瑞, 邹彦雯, AND 李婷婷. 城市轨道交通线路运营计划优化方法. 北京交通大学学报: 自然科学版 (2015).
- [5] 林震, AND 杨浩. 城市轨道交通发车间距优化模型探讨. 土木工程学报 36, 001 (2003), 1–5.
- [6] 杨晓光, AND 褚浩然. 错峰出行对城市交通的影响分析. 同济大学学报 (自然科学版) 34, 007 (2006), 899–903.
- [7] 刘兰芬, AND 杨信丰. 地铁客流分析及列车发车间隔优化研究. 武汉理工大学学报 (交通科学与工程版) 39, 006 (2015), 1119–1124.
- [8] FAZHI, L. I., 李发智, LEI, X., 肖蕾, DONGYUAN, Y., AND 杨东援. 公交发车间隔模型研究. In 2011 年全国博士生学术论坛——交通运输工程 (2011).