

# 2022 年山西省“金地杯”大学生 论文

## 题    目            车辆轨迹数据的分析应用

### 摘    要

近年来,随着我国经济社会发展和人民群众生活水平提高,高速公路交通流量大幅增加,高速公路用户对高速公路服务水平的要求也不断提高,因停车收费造成交通拥堵的现象时有发生。因此为了克服传统的全人工及半自动收费中存在的种种弊端,以不断提升高速公路通行效率和服务水平,电子收费系统(ETC)应运而生。该系统最主要的组成部分是建造在高速公路主线横断面上的门架,主要包含车辆视觉识别、天线、边缘计算服务单元等关键设备。ETC 依托高速公路一定规模区域内密集、连续布设的门架系统,结合高速公路收费站等其他必要的资源,实现车辆行驶轨迹检测。研究分析车辆的轨迹有利于达到最大限度提高路网安全畅通的目的,从而更有利于提高城市的运行速率,推动城市进步。

对于问题 1,这是有关多个元素的问题,我们可以用线性回归的方法求解,利用给定的数据,借助线性代数和 Excel,得出各个时间段的维修门架及维修时间

对于问题 2 我们用 excel 和算法算出差值除以平均值就是漏失记录率

对于问题 3 我们用非线性模型求解,

对于问题 4 我们用预测模型,如灰色模型

**关键词**    多元线性回归模型    数据分析

## 目 录

一、问题重述·····	(1)
二、问题分析·····	(2)
三、模型假设·····	(3)
四、定义与符号说明·····	(4)
五、模型的建立与求解·····	(5)
5.1 问题 1 的模型·····	(6)
5.1.1 模型的建立·····	(7)
5.1.2 模型求解·····	(8)
5.1.3 结论·····	(9)
·····	
5.2 问题 2 的模型·····	(10)
5.2.1 模型的建立·····	(11)
5.2.2 模型求解·····	(12)
5.2.3 结论·····	(13)
·····	
5.3 问题 3 的模型·····	( )
5.3.1 模型的建立·····	( )
5.3.2 模型的求解·····	( )
六、模型的评价	
6.1 模型的优点·····	(16)
6.2 模型的缺点·····	(17)
参考文献·····	( )
附录·····	( )

# 一、问题重述

## 问题背景

近年来,随着我国经济社会发展和人民群众生活水平提高,高速公路交通流量大幅增加,高速公路用户对高速公路服务水平的要求也不断提高,因停车收费造成交通拥堵的现象时有发生。因此为了克服传统的全人工及半自动收费中存在的种种弊端,以不断提升高速公路通行效率和服务水平,电子收费系统(ETC)应运而生。该系统最主要的组成部分是建造在高速公路主线横断面上的门架,主要包含车辆视觉识别、天线、边缘计算服务单元等关键设备。ETC 依托高速公路一定规模区域内密集、连续布设的门架系统,结合高速公路收费站等其他必要的资源,实现车辆行驶轨迹检测。

## 问题设计

为了分析某高速路区间的交通状况,对经过该高速路区间车辆的信息进行采集。该区间以收费站 13 为起点,依次包含收费站 12、11、15、10、9、8、14、7、6、5、4、3、2,以收费站 1 为终点,共 15 个收费站。区间内同时设有用于车辆识别的 ETC 门架共 15 个,依次从门架 1 至门架 15。收费站、门架分布以及以相继两个门架为起、终点进行分段的各路段间距见附件 1“高速路段图”。附件 2 记录了由收费站和门架采集的自 2022 年 2 月 22 日至 2 月 27 日的六天时间内经过该高速路区间的全部车辆的轨迹信息。对轨迹表的解释见附件 3“轨迹表字段释义”。

建立数学模型以解决下面的问题。

1. 在该段时间内,运维部门对一些 ETC 门架进行了维修。请描述车流量的时空分布并依据分布指出维修门架的编号和维修的时间。
2. 门架是否会漏失对过路车辆的记录,如果会,请指出各门架的漏失记录率和容易漏失记录的车辆的特征。
3. 说明收费站流量与路段流量的关系,并确定各站收费车道数以及应急收费系统(便携式收费机或移动电站收费车)的数量,使得在保证通行顺畅的前提下达到费用的最省。
4. 预测各路段未来 24 小时的车流量,并利用交通流时空分布不均衡的特点,依据预测模块实时优化未来 24 小时各个路段的限速值,以达到最大限度提高路网安全畅通的目的。

## 二、问题分析

### 2.1 问题 1 的分析

问题 1 的研究有助于更好的发挥智能交通系统的作用，对设施进行更精准有效的检测，达到合理有效高效的进行运行。

问题 1 属于线性回归数学问题，对于解决此类问题一般用线性代数模型

此题已经给出数据，这些是非常庞大且复杂的，且信息量多，需要加工提取

由于以上原因，我们可以先对数据进行观察分析，运用 excel 的工具然后对数据进行筛选，进一步统计绘制图表求回归方程可以得出需要维修的门架及时间。

### 2.2 问题 2 的分析

研究问题 2 有助于推动完善系统，使系统可以更好地运行，避免一些其他问题的产生。有助于适应更好的城市节奏。

还是需要数据的处理与分析，且使用的还是庞大复杂的数据

解答可以把问题 1 中求得的结果运到第二题中，然后再反带 excel 数据表中筛选

### 2.3 问 3 题的分析

这是一道非线性规划的问题，用非线性规划模型求解

### 2.4 问题 4 的分析

这是一道预测题，可以用灰色模型解决

## 三、模型假设

1. 假设题目所给的数据真实可靠；
2. 假设所有车都安装有 ETC 可以识别的标识；
3. 假设维修部到达的时间很短，可以忽略不计；
4. 假设测试时间内天气等各种次要因素的影响可忽略不计。

## 四、定义与符号说明

符号定义	符号说明
y	车流量
x	时间段

## 五、模型的建立与求解

数据的预处理：

1. 假设次要数据全部丢失不予处理
2. 对数据测试的特点，如周期等进行分析。
3. 如果数据残缺，根据数据挖掘等理论根据数据的变化趋势进行补充。
4. 对数据特点进行提取。

用 excel 软件聚类分析和各个不同问题的需要，采得 3 组采样，每组 24 个采样值。将采样所对应的特征值进行列表或图示。

根据数据特点，对总体和个体的特点进行比较，以表格或图示方式显示。

### 5.1 问题 1 的模型建立与求解

#### 5.1.1 线性回归模型的建立

模型建立的内容要点如下：筛选数据，插入图表，建立回归方程。

模型的主要类别：线性回归模型

几种常见的建模目的：描述变量之间的变量关系，预测因变量

建模过程常见的几个要点：提出问题，分析各种和模型有关因素，建模

模型的基本要求：有明显的多元关系

模型选择要点：根据数据处理的实用性和规范性，合理性

### 5.1.2 模型的求解

通过 excel 软件将需要的各门架在各个时间段内的数据筛选出来，然后汇总，绘制散点图

### 5.1.3 结果

通过是上述过程的结果是可用 excel 图标来表示

## 5.2 问题 2 的模型建立与求解

### 5.2.1 \*\*\*模型的建立

我们需要解决的问题是门架的遗失记录率和遗失记录的车特征。我们还可以用 excel 表格来表示，用各时间段的们家属减去平均数除以平均数即为要求的值，最后反带会表格，查看结果即可

### 5.2.2 模型的求解

将预处理数据带入上述模型，通过 excel 软件得到结果模型求解结果  
门架 1、2、3、6、11、13、14 分别遗失记录

### 5.2.3 结果

依据表格可以得出各门架的遗失记录率

## 5.3 问题 3 的模型建立与求解

### 5.3.1 非线性模型的建立

我们需要解决的问题是收费站流量与路段流量的关系，，题目要求是使得在保证通行顺畅的前提下达到费用的最省，我们可以用非线性回归模型

### 5.3.2 非线性回归模型的求解

无

### 5.3.3 结果

无

## 5.4 问题 4 的模型建立与求解

### 5.4.1 灰色模型的建立

## 六、模型的评价

### 6.1 误差分析

#### 6.1.1 针对于问题 1 的误差分析

最终数值结果的正确性或合理性是第一位的；对数值结果或模拟结果进行必要的检验。结果不正确、不合理、或误差大时，分析原因，对算法、计算方法、或模型进行修正、改进；

#### 6.1.2 针对于问题 2 的误差分析

最终数值结果的正确性或合理性是第一位的；对数值结果或模拟结果进行必要的检验。结果不正确、不合理、或误差大时，分析原因，对算法、计算方法、或模型进行修正、改进；2 的误差分析

### 6.2 模型的优点

得到较合理的解

较好地解决了线性问题

使模型得到简化

使结果更合理，避免差值带来的较大误差

使问题描述比较清晰

减少大的计算量

## 6.3 模型的缺点

主观性过强

建立在什么的前提条件下

有一定的局限性

存在不确定性

有一定的偏差

## 参考文献

1 司守奎 孙玺菁等著《数学建模的算法与应用》国防工业出版社 2020 年 211-216 218-219 458-460

2 张慧 《浅谈数学建模中的方法、技巧和须注意的问题》 2011 年 第 21 期



## 附录

Excel 表格