

打车软件对出租车资源配置的影响

摘要

打车难问题是当今社会热点问题,而近年来推出的打车软件正致力于缓解这一问题。本文主要研究出租车资源在不同时空的供求匹配程度,探讨现有出租车补贴方案对缓解打车难问题的作用,并设计了新的出租车补贴方案。

针对问题一,首先从空间角度进行分析,定义空间差异性为一二三线城市并分别选择北京、成都和合肥作代表,为控制时间变量选择 2015 年数据作为研究对象。再根据 Delphi 法确定影响供求匹配程度指标,通过 ISM 解释结果模型分析各指标间的关系来建立评估指标体系。然后结用拉开档次法确定主观权重,用 D-S 理论确定客观权重,用主客观组合赋权法结合两者确定综合权重。最后在修正指标归一化后,求得北京、成都和合肥的出租车资源供求匹配程度得分为:

北京	成都	合肥
0.856504	0.779608	0.583116

对比得出城市综合发展水平越高,出租车资源供求匹配程度越好的结论。

然后从时间角度进行分析,首先定义时间差异性为 24 小时,选定北京市进行分析以控制空间变量。通过网络爬虫数据挖掘软件在滴滴苍穹大数据平台上挖掘一天中每小时的打车难易度,用 9 月 5 日采集到的数据代表休息日,用 9 月 7 日采集到的数据代表工作日。最后通过系统聚类对各时间段的供求匹配程度进行分级,分别将工作日分为 4 级,将休息日分为 3 级。

针对问题二,通过出租车满载率上升率体现对缓解打车难问题的帮助。根据出租车满载率函数及其变量间的相互关系,推导出出租车满载率上升率函数

$$\alpha = \frac{(1+\beta)(1+\gamma)(1+\lambda)}{1+\mu} - 1$$

补贴金额 B_1B_2 通过影响函数中的出租车工作效率 λ 来改变出租车满载率上升率。剔除其他因素的影响,仅考虑补贴金额的作用,得到出租车工作效率提升率函数:

$f(B_1, B_2) = -2.4141 + 0.0003B_1^2B_2^2 - 0.0189B_1^2 - 0.082B_2^2 - 0.0440B_1B_2 + 0.3073B_1 + 1.0766B_2$
并根据此函数计算得到主流打车软件公司的补贴方案都能够缓解打车难问题。

针对问题三,首先以打车软件公司盈利最高和出租车满载率上升率最大建立双目标规划模型,然后求得最优解为对乘客每单补贴 3 元但不对司机进行补贴的补贴方案。最后分别从打车软件的持续发展、对缓解打车难问题的作用和公司的盈利模式三个方面论证该补贴方案的合理性。

关键词: Delphi 法 ISM 解释结构模型 D-S 理论 主客观组合赋权法
大数据挖掘 系统聚类 多目标规划

一、问题重述

出租车是市民出行的重要交通工具之一，“打车难”是人们关注的一个社会热点问题。随着“互联网+”时代的到来，有多家公司依托移动互联网建立了打车软件服务平台，实现了乘客与出租车司机之间的信息互通，同时推出了多种出租车的补贴方案。

请你们搜集相关数据，建立数学模型研究如下问题：

- (1) 试建立合理的指标，并分析不同时空出租车资源的“供求匹配”程度。
- (2) 分析各公司的出租车补贴方案是否对“缓解打车难”有帮助？
- (3) 如果要创建一个新的打车软件服务平台，你们将设计什么样的补贴方案，并论证其合理性。

二、模型假设

- 1、本文从网站搜集的数据真实可靠，能够真实的反应当今的实际情况。
- 2、出租车的运营在长时间内没有太大的波动，例如某地发生自然灾害、地震等。
- 3、对于某些缺失数据，虽然为人为增添，但是对模型应用无影响，并且能更好地说明模型的实用性。

三、符号说明

符号	说明	符号	说明
w	权重系数向量	M	概率分配函数
Bel	信任函数	Pl	似然函数
μ_{ij}	评估数据	S	决策方案集
W	权重	N	出租车满载率
E	出租车工作效率	β	人口数量增加率
γ	收入水平提升率	B	出租车补贴金额

四、问题一模型的建立与求解

4.1 问题分析

本问要求我们建立指标，分别从时间和空间两个角度分析出租车的供求匹配程度。从空间角度考虑时，首先需要保持时间方面的变量一致，可以选择某一段时间内不同地区的数据，例如某一年内不同地区的数据。其次需要定义空间差异性，用不同的分类指标和分类方法来体现空间上的区别，再根据空间差异性定义结果搜集相关数据并分析其相互关系，建立合理评价指标集并给出指标权重，得到供求匹配程度并对比分析得出空间上的差异。从时间角度考虑时，同样需要保持空间方面的变量一致，然后定义时间差异性，并对不同时间的供求匹配程度进行对比分析，得到供求匹配程度的时间分布规律。

4.2 出租车资源供求匹配程度空间角度分析

为保证数据的时效性，我们选择 2015 年的数据进行分析。为反映地域差异性，我们将差异性定义为一线城市、二线城市和三线城市，并分别选择北京代表一线城市，成都代表二线城市，合肥代表三线城市，通过对这三个城市采集数据并分析，来体现一二线城市各自的出租车资源供求匹配水平。

4.2.1 Delphi 法确定影响供求匹配程度指标^[1]

Delphi 法是一种采用专家小组形式，可以充分利用专家的经验 and 学识；每个成员独立发言，使每一位专家都能独立自由地做出自己的判断，不为对方所干扰；预测过程经过数轮反馈，使专家的意见逐渐趋同。

人们对不同因素下的打车软件的使用影响积极性反应程度均不相同。对于本文的研究来说，这些因素的影响作用之间存在一定的差异，于此同时，这些因素还会产生交互影响，层级逻辑关系复杂。通过 Delphi 法进行分析，可以迅速理清错综复杂的影响因素，并在此基础上确定初始外生因素以及深层次的关键影响因素。

本文使用 Delphi 法针对出租车行业的专业以及从业者建立专家小组，分别从出租车公司、出租车司机与乘客三方面的角度出发，最终确定影响供求匹配程度的 14 个因素，如下表所示：

表 1 供需匹配因素

P1: 供需程度	P8: 万人拥有量
P2: 乘客消费水平	P9: 车辆空载率
P3: 城市人口数量	P10: 司机车收入水平
P4: 人均收入水平	P11: 城市交通水平
P5: 城市人口密度	P12: 城市道路面积
P6: 出租车服务能力	P13: 公共交通水平
P7: 里程利用率	P14: 城市环保水平

4.2.2 基于 ISM 解释结构模型的指标关系分析

通过 Delphi 法我们确立了影响供求匹配程度的 14 个因素，但对于其内在关系我们无法给出准确判断和归类，进而无法建立系统的评估指标体系，所以在此我们研究各评估指标间的关系和相互间的联系。

ISM 解释结构模型^[1]

1973 年, J.N.Warfield 为了分析复杂的社会经济系统结构问题开发了 ISM(Interpretive Structural Model)解释结构模型, ISM 模型是静态的定性模型。ISM 解释结构模型的特点是通过将复杂的系统分解为多个子系统, 凭借人们的实践经验, 利用矩阵、有向图, 以及计算机软件的辅助, 把经过处理的信息构建成多级解释结构模型, 并对各个元素进行文字说明。最终将系统构造成一个多层递阶解释结构模型, 并通过文字加以解释说明, 明确问题的层次和整体结构。提高对问题的认识和理解程度。由于 ISM 模型清晰地表示了问题的层次以及结构, 所以提高了人们对问题的认识和理解程度。

模型建立步骤

对提出的问题, 在充分了解系统边界, 提出各因素的基础上, 就可以进行 ISM 方法建模。建模的过程简要来说主要包括以下几个步骤:

第一步, 使用 Delphi 法, 组织专业人员对以文字形式设定问题, 相关成员提出构成问题的要素, 最后归纳总结为 10-30 个相关要素。

第二步, 小组成员对问题和要素进行确切的文字描述。

第三步, 明确各要素之间的相互关系, 并将两项关系的有无总结成邻接矩阵 A。邻接矩阵的各元素取值原则: 两项之间有关系则为 1, 反之为 0。

第四步, 生成可达矩阵, 方法为: 将邻接矩阵 A 与单位矩阵 I 相加, 之后对矩阵(A+I)做 n 次幂运算, 其中幂运算基于布尔运算

$$(0 + 0 = 0, 0 + 1 = 1, 1 + 1 = 1, 1 \times 0 = 0, 0 \times 1 = 0, 1 \times 1 = 1), \text{直到} \\ M = (A + I)^{n+1} = (A + I)^n = \dots = (A + I)^2 = (A + I)$$

其中:

$$(A + I)^n = I + A + A^2 + \dots + A^n$$

为止, 其中 $M = (A + I)^n$ 即为所要求的可达矩阵。

第五步, 对可达矩阵 M 进行区域分解、级间分解和骨架矩阵提取, 得到骨架矩阵。

第六步, 基于骨架矩阵, 建立层级关系有向图, 根据有向图得到解释结构模型。

模型的求解与结果分析

1.建立邻接矩阵

通过小组成员的讨论, 提出各个影响因素之间的逻辑关系, 得到影响因素之间的相互关系示意图如下所示:

			A					A				A	P1: 供需程度
									A	A	A		P2: 乘客消费水平
									V	V			P3: 城市人口数量
V	V	V	V					V	X				P4: 人均收入水平
						V		V					P5: 城市人口密度
				A	A	A	A						P6: 出租车服务能力
					X								P7: 里程利用率
					A								P8: 万人拥有量
			A										P9: 车辆空载率
			V										P10: 司机车收入水平
A	A	A											P11: 城市交通水平
													P12: 城市道路面积
													P13: 公共交通水平
													P14: 城市环保水平

图 1 影响供求匹配程度指标关系图

其中要素间的逻辑关系用 V、A、X 表示。V 表示方格图中的行要素直接影响到列要素，A 表示列要素对行要素有直接影响，X 表示行列两要素相互影响。

随后建立邻接矩阵 A:

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14
P1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P2	1	0	0	0	0	0	0	0	0	0	0	0	0	0
P3	0	1	0	1	1	0	0	0	0	0	0	0	0	0
P4	0	1	0	0	1	1	0	0	0	0	1	1	1	1
P5	0	1	0	1	0	1	0	1	0	0	0	0	0	0
P6	1	0	0	0	0	0	0	0	0	0	0	0	0	0
P7	0	0	0	0	0	1	0	0	1	0	0	0	0	0
P8	0	0	0	0	0	1	0	0	0	0	0	0	0	0
P9	0	0	0	0	0	1	1	0	0	0	0	0	0	0
P10	0	0	0	0	0	1	0	0	1	0	1	0	0	0
P11	1	0	0	0	0	0	0	0	0	0	0	0	0	0
P12	0	0	0	0	0	0	0	0	0	0	1	0	0	0
P13	0	0	0	0	0	0	0	0	0	0	1	0	0	0
P14	0	0	0	0	0	0	0	0	0	0	1	0	0	0

图 2 邻接矩阵图

本文使用 14*14 的方形矩阵来表示，14 即为影响因素的个数，并规定当影响因素 S_i 对 S_j 有影响时，矩阵元素 a_{ij} 为 1，反之矩阵元素 a_{ij} 则为 0。即

$$a_{ij} = \begin{cases} 1 & \text{当 } S_i \text{ 对 } S_j \text{ 有影响时} \\ 0 & \text{当 } S_i \text{ 对 } S_j \text{ 无影响时} \end{cases}$$

2.可达矩阵的求解

由于矩阵内因素较多，运算量庞大，所以通过计算机软件 Matlab 求解可达矩阵。可达矩阵表示了两个影响因素之间是否存在连接路径，它是用矩阵的形式来说明有向图各个定点之间通过一定路径可以到达的程度，程序见附录 1。经计算得出当 $K=4$ 时，求出可达矩阵 M ：

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14
P1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
P2	1	1	0	0	0	0	0	0	0	0	0	0	0	0
P3	0	1	1	0	0	0	0	0	0	0	0	0	0	0
P4	0	1	0	1	0	0	0	0	0	0	0	0	0	0
P5	0	1	0	0	1	0	0	0	0	0	0	0	0	0
P6	1	0	0	0	0	1	0	0	0	0	0	0	0	0
P7	0	0	0	0	0	1	1	0	0	0	0	0	0	0
P8	0	0	0	0	0	1	0	1	0	0	0	0	0	0
P9	0	0	0	0	0	1	0	0	1	0	0	0	0	0
P10	0	0	0	0	0	1	0	0	0	1	0	0	0	0
P11	1	0	0	0	0	0	0	0	0	0	1	0	0	0
P12	0	0	0	0	0	0	0	0	0	0	1	1	0	0
P13	0	0	0	0	0	0	0	0	0	0	1	0	1	0
P14	0	0	0	0	0	0	0	0	0	0	1	0	0	1

图 3 可达矩阵图

确定各层级包含因素后，建立表格。根据可达矩阵列举出每个因素 P_i 对应行以及列与其他因素之间的对应关系，定义可达集合 $R(S_i)$ 是可达矩阵中要素 S_i 这一行中对应值 1 的矩阵元素所在的列要素的集合。可达集合 $Q(S_i)$ 是可达矩阵中要素 S_i 这一列中对应值 1 的矩阵元素所在的所有行要素的集合。当 $R(S_i) = R(S_i) \cap Q(S_i)$ 时， S_i 即为该层的元素，剔除元素集合中的 S_i ，继续进行下一层的元素计算。

3.建立递阶有向图

经过计算推导，可以建立如下骨架矩阵 R^* 构成系统的结构模型。

	P1	P2	P6	P11	P3	P4	P5	P7	P8	P9	P10	P12	P13	P14
P1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
P2	1	1	0	0	0	0	0	0	0	0	0	0	0	0
P6	1	0	1	0	0	0	0	0	0	0	0	0	0	0
P11	1	0	0	1	0	0	0	0	0	0	0	0	0	0
P3	0	1	0	0	1	0	0	0	0	0	0	0	0	0
P4	0	1	0	0	0	1	0	0	0	0	0	0	0	0
P5	0	1	0	0	0	0	1	0	0	0	0	0	0	0
P7	0	0	1	0	0	0	0	1	0	0	0	0	0	0
P8	0	0	1	0	0	0	0	0	1	0	0	0	0	0
P9	0	0	1	0	0	0	0	0	0	1	0	0	0	0
P10	0	0	1	0	0	0	0	0	0	0	1	0	0	0
P12	0	0	0	1	0	0	0	0	0	0	0	1	0	0
P13	0	0	0	1	0	0	0	0	0	0	0	0	1	0
P14	0	0	0	1	0	0	0	0	0	0	0	0	0	1

图 4 骨架矩阵图

骨架矩阵中所对应的全部行因素为一个递阶结构层次，从骨架矩阵中可以看出，影响打车软件使用积极性的因素可大致分为三层，并且它们之间的层次关系形成了有一定逻辑关系的影响因素链，可绘出影响因素的层次结构图。根据已求得骨架矩阵中元素之间的关系，建立出最终的 ISM 结构模型图如下所示：

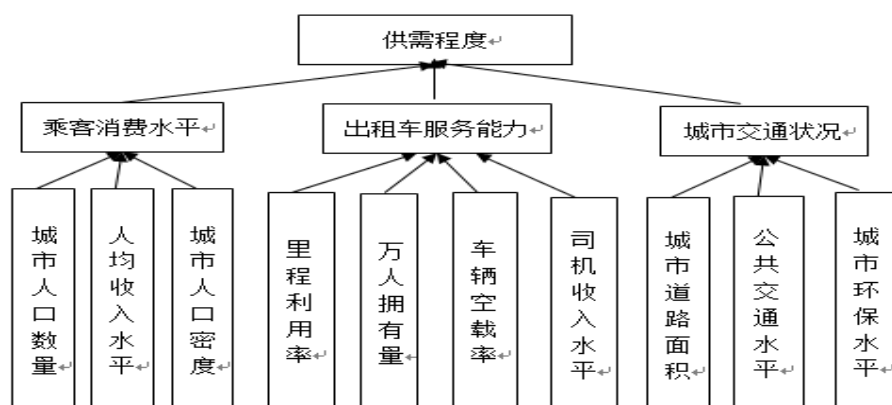


图 5 ISM 解释结构模型

4.2.3 基于主客观组合赋权法的指标权重确定^[2]

完成评价指标集的建立后我们需要确定评价指标的权重，在此我们突破传统上只单一考虑主观指标或者客观指标的方法，综合考虑主观信息和客观信息来确定指标权重。

拉开档次法确定主观指标权重

设 x_1, x_2, \dots, x_m 为极大型指标（即指标取值越大越好），取其线性函数

$$y = w_1 x_1 + w_2 x_2 + \dots + w_m x_m = w^T x$$

为被评价对象的综合函数。式中 $w = (w_1, w_2, \dots, w_m)^T$ 是 m 维持定正向量（其

作用相当于权系数向量), $x = (x_1, x_2, \dots, x_m)^T$ 为评价对象的状态的状态向量。如将第 i 个被评对象 o_i 的标准观测值 $x_{i1}, x_{i2}, \dots, x_{im}$, 带入上式中, 既得

$$y = w_1 x_{i1} + w_2 x_{i2} + \dots + w_m x_{im}$$

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad A = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{pmatrix}, \quad \text{则公式可写成 } y = Aw$$

确定权重系数向量 w 的准则是求指标向量 x 的线性函数 $w^T x$, 使此函数对 n 个被评价对象取值的分散程度或方差尽可能地大。

而变量 $y = w^T x$ 按 n 个被评价对象取值构成样本的方差为

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{y^T y}{n-1} - \frac{n}{n-1} \bar{y}^2$$

将 $y = Aw$ 代入上式, 并注意到原始数据的标准化处理, 可知 $\bar{y} = 0$, 于是有

$$(n-1)s^2 = w^T A^T A w = w^T H w$$

式中, $H = A^T A$ 为实对称矩阵。

显然, 对 w 不加限制是, 式 $(n-1)s^2 = w^T A^T A w = w^T H w$ 可取任意大的值。这里限定 $w^T w = 1$, 选择 w 使得

$$\begin{aligned} & \max w^T H w \\ & s.t. w^T w = 1 \\ & w > 0 \end{aligned}$$

通过拉开档次法进行计算, 求得各影响因素的主观权重如下:

表 2 各因素的主观权重

影响因素	P1	P2	P6	P1	P3	P4	P5
主观权重	1	0.	0.	0.	0.	0.	0.
影响因素	P7	P8	P9	P1	P1	P1	P1
主观权重	0.	0.	0.	0.	0.	0.	0.

基于 D-S 理论的客观指标权重确定^[3]

证据理论 (D-S) 由德普斯特首先提出, 后由沙佛发展起来。该理论主要用于处理一些有不确定性的信息, 既属于不确定性的推方法, 也是有效的数据融合方法之一。

识别框架 D 表示一个不变的两两互斥又可穷举的元素的集合, 即

$D = \{A_1, A_2, \dots, A_n\}$, 其中一个元素能被解释可能的答案, 且只有一个正确答案。

概率分配函数、信任函数和似然函数分别设为 M 、 Bel 和 Pl , 对于 $2^D \rightarrow [0, 1]$, 若 $A \subseteq D$, 称

$$\begin{cases} M(\phi) = 0, \\ \sum_{A \subseteq D} M(A) = 1, \end{cases}$$

$$Bel(A) = \sum_{B \subseteq A} M(B)$$

$$Pl(A) = 1 - Bel(-A)$$

$M(A)$ 、 $Bel(A)$ 、 $Pl(A)$ 分别为 A 的概率分配函数、信任函数和似然函数。

$Bel(A)$ 只能反映出对 A 为真的信任度，而不能反映出相信 A 为非假的程度，从而用似函数来表示怀疑命题 A 为非假的程度。

权重系数关系到整个评价的效果，文中假设各专家是同等重要的，即决策者不能明确确定各专家的主观权重，认为他们的权重均相同。由于每位专家的实际经验不尽相同，所以在给出各指标权重时也会产生一些冲突，为此，文中采用证据距离的方法确定权重，这样能够大大减弱由于各数据间的明显冲突而导致的融合数据不精确程度。

定义 E_1 和 E_2 是识别框架 D 下的两个证据， m_1 和 m_2 分别为其概率分配函数，焦点分别为 A_i 和 B_j ，则 m_1 与 m_2 的距离可表示为

$$d(m_1, m_2) = \sqrt{\frac{1}{2} (m_1 - m_2)^T D (m_1 - m_2)}$$

式中， D 为一个 $2^N \times 2^N$ 的矩阵，则 m_1 与 m_2 的距离也可表示为

$$d(m_1, m_2) = \sqrt{\frac{1}{2} (\|m_1\|^2 - \|m_2\|^2 - 2\langle m_1, m_2 \rangle)}$$

其中

$$\langle m_1, m_2 \rangle = \sum_{i=1}^{2N} \sum_{j=1}^{2N} m_1(A_i) m_2(B_j) \frac{|A_i \cap B_j|}{|A_i \cup B_j|}$$

$$\|m\|^2 = \langle m, m \rangle$$

由公式可知 $0 \leq d(m_i, m_j) \leq 1$ ，且两证据间距离越大，则分歧就越大，反之距离越小，则分歧越小。

定义 m_i 与其他证据的总距离为

$$d_i = \sum_{j=1, j \neq i}^n d(m_i, m_j)$$

令 $\tilde{d} = \min \{d_i\} (i=1, 2, \dots, n)$ ，可以看出 \tilde{d} 与其他证据的距离是最近的，即可认为 \tilde{d} 为关键证据，则权重

$$\lambda_i = \frac{\tilde{d}}{d_i}, i=1, 2, \dots, n$$

由此便可确定权重系数矩阵

$$\lambda_n = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \Rightarrow \begin{pmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1n} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{m1} & \mu_{m2} & \cdots & \mu_{mn} \end{pmatrix}$$

μ_{ij} 表示第 j 个专家对第 i 个指标的评估数据, λ_i 由 μ_{ij} 通过证据 D-S 距离融合而成。

通过计算, 求得各影响因素的客观权重如下:

表 3 各因素的客观权重

影响因素	P1	P2	P6	P1	P3	P4	P5
客观权重	1	0.	0.	0.	0.	0.	0.
影响因素	P7	P8	P9	P1	P1	P1	P1
客观权重	0.	0.	0.	0.	0.	0.	0.

主客观组合赋权确定综合权重^[4]

关于权重的确定, 目前来讲主要有主观赋权法和客观赋权法两大类。主观赋权法是根据决策者主观信息进行赋权的一类方法; 而客观赋权法是无决策者任何信息, 各个指标根据一定的规则进行自动赋权的一类方法。运用主观赋权法确定各指标间的权重系数, 反映了决策者的意向, 决策或评价结果具有很大的主观性。而运用客观赋权法确定各指标间的权重系数, 决策或评价结果虽然具有较强的数学理论依据, 但没有考虑决策者的意向。因此, 主、客观赋权法各具有一定的局限性。为使多决策分析结果既含主观信息又含客观信息, 本文使用主客观综合赋权方法, 使其确定的属性权重同时反映主观程度和客观程度。

设多属性决策问题的决策方案集为 $S = \{S_1, S_2, \dots, S_m\}$, 属性 (或指标) 集为 $P = \{P_1, P_2, \dots, P_n\}$, 指标加权向量 $W = (W_1, W_2, \dots, W_n)^T$, 且满足 $W_j \geq 0 (j=1, 2, \dots, n)$, $\sum_{j=1}^n W_j = 1$, 方案 S_i 对属性 P_i 的属性值记为 $a_{ij} (i=1, 2, \dots, m; j=1, 2, \dots, n)$, 矩阵 $\bar{A} = (a_{ij})_{m \times n}$ 称为决策矩阵, 记已规范化的决策矩阵为 $B = (b_{ij})_{m \times n}$ 。

记由主观赋权得出 (或由决策者直接给出) 的属性权重向量为 $W' = (W'_1, W'_2, \dots, W'_n)^T$, 且满足 $0 \leq W'_j \leq 1$, $\sum_{j=1}^n W'_j = 1$ 。由客观赋权得出的属性权重向量为 $W'' = (W''_1, W''_2, \dots, W''_n)^T$, 且满足 $0 \leq W''_j \leq 1$, $\sum_{j=1}^n W''_j = 1$ 。记 α, β 分别表现 W' 和 W'' 的重要程度, 且 $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$, $\alpha + \beta = 1$ 。

在此我们取主客观权重分别为:

主观赋权	客观赋权
0.4	0.6

考虑到将主观权重向量与客观权重向量进行综合, 则令 $W = \alpha W' + \beta W''$ 得到主客观综合赋权法确定的权重为:

表 4 各因素的综合权重

影响因素	P1	P2	P6	P1	P3	P4	P5
综合权重	1	0.	0.	0.	0.	0.	0.
影响因素	P7	P8	P9	P1	P1	P1	P1
综合权重	0.	0.	0.	0.	0.	0.	0.

4.2.4 不同城市出租车资源供求匹配程度分析

数据归一化处理^[5]

归一化是一种无量纲处理手段，数据标准化（归一化）处理是数据挖掘的一项基础工作，不同评价指标往往具有不同的量纲和量纲单位，这样的情况会影响到数据分析的结果，为了消除指标之间的量纲影响，需要进行数据标准化处理，以解决数据指标之间的可比性。原始数据经过数据标准化处理后，各指标处于同一数量级，适合进行综合对比评价。本文将数据归一化：

$$x_i^* = \frac{x_i}{\max x_i}$$

归一化处理后三个城市的各项指标值如下所示：

表 5 三大城市归一化数据

城市	城市人口数量	人均收入	人口密度	车辆空载率	出租车万人拥有量	司机收入	里程利用率	道路面积	公共交通水平	城市环保状况
北京	1	0.58	1	0.8	1	0.87	0.94	0.38	1	1
成都	0.27	1	0.70	0.8	0.69	1	0.93	0.80	0.8	0.9
合肥	0.38	0.36	0.34	1	0.35	0.65	1	1	0.7	0.8

出租车资源供求匹配程度评分计算与结果分析

结合上文通过主客观组合赋权法确定的权重和归一化处理后的各项指标值，得到三个城市的出租车资源供求匹配程度评分为：

北京	成都	合肥
0.856504	0.779608	0.583116

结果分析

根据上文中求得的一二三线城市的出租车资源供求匹配程度评分我们可以得出结论，就出租车供求匹配程度而言：

一线城市 > 二线城市 > 三线城市

城市的综合发展水平越高，出租车资源的供求匹配程度越好，市民打车难度越小。所以发展城市综合水平，将能够一定程度上缓解打车难的问题。

4.3 出租车资源供求匹配程度时间角度分析

只从时间角度考虑时就要剔除空间方面的影响，所以在此我们只选定一个城市进行分析。由于中国各地区的上班规律和作息时无明显差异，因此我们选择出租车系统更稳定的北京市作为研究对象，对比分析其不同时间的出租车资源供求匹配程度来研究时间方面的影响。

4.3.1 基于打车难易度的供求匹配程度分时段模型

在此我们分别考虑工作日和休息日两种情况，通过统计一天中 24 个时段的打车难易度来反映该时段出租车的供求匹配程度。

大数据平台打车难易度全时段挖掘

滴滴打车是由大数据底层构架支撑的智能出行平台，可以实时呈现出租车出行的相关数据。通过此网站可以在打车难易度、打车需求量和被抢单时间这三个可视化维度上观察全国和各大重要城市的出租车使用情况。

网络爬虫是一个自动提取网页的程序，通过搜索引擎从万维网上下载网页，并根据一定的网页分析算法过滤与主题无关的链接，保留有用的链接并将其放入等待抓取的 URL 队列，然后根据一定的搜索策略从队列中选择下一步要抓取的网页 URL，并重复上述过程，直到达到系统的某一条件时停止，所有被爬虫抓取的网页将会被系统存贮，进行一定的分析、过滤，并建立索引，以便之后的查询和检索，通过滴滴打车大数据平台，我们可以用网络爬虫软件挖掘出租车相关数据，在此我们主要从工作日和休息日两方面对打车难易度进行全时段挖掘，难易度分为 0-10 共十一个等级，时间段分为 0-24 小时共二十四个时间段。

在此我们选用 2015 年 9 月 5 日的各时段打车难易度来代表休息日的打车状况，其数据如下：

表 6 休息日各时段打车难易度

打车 难度 时间	0	1	2	3	4	5	6	7	8	9	10
1	267	1	1	0	2	1	3	0	24	2	0
2	275	1	0	2	0	2	0	21	0	0	0
3	278	4	0	0	0	0	0	1	14	4	0
4	268	24	3	0	0	0	0	1	4	1	0
5	277	6	1	0	3	1	0	10	2	1	0
6	263	8	1	0	2	1	2	0	22	2	0
7	212	5	1	10	4	12	31	22	3	1	0
8	203	4	6	15	7	3	13	33	12	4	1
9	195	3	6	11	14	4	9	33	23	3	0
10	188	1	3	2	8	6	9	4	66	12	2
11	221	0	0	1	6	3	5	39	23	2	1
12	206	0	1	0	6	3	8	60	15	1	1
13	221	0	0	2	4	3	4	47	18	1	1
14	216	1	1	3	3	2	8	56	10	0	1
15	216	1	0	1	3	1	4	52	16	3	4
16	202	2	2	1	5	4	6	54	24	1	0
17	212	0	1	3	7	2	5	48	16	2	5
18	186	0	4	7	7	8	11	51	22	5	0
19	214	1	6	0	9	4	5	44	16	1	1
20	215	1	2	4	6	3	6	45	18	1	0
21	223	0	3	4	6	4	5	41	13	1	1
22	216	1	4	11	7	6	4	2	46	3	1
23	238	4	2	3	9	6	5	0	30	2	2
24	257	0	1	3	5	1	4	26	2	0	1

同时选用 2015 年 9 月 7 日的各时段打车难易度来代表工作日的打车状况，其数据如下：

表 7 工作日各时段打车难易度

打车 难度 时间	0	1	2	3	4	5	6	7	8	9	10
1	267	1	1	0	2	1	3	0	24	2	0
2	275	1	0	2	0	2	0	21	0	0	0
3	278	4	0	0	0	0	0	1	14	4	0
4	268	24	3	0	0	0	0	1	4	1	0
5	277	6	1	0	3	1	0	10	2	1	0
6	263	8	1	0	2	1	2	0	22	2	0
7	212	5	1	10	4	12	31	22	3	1	0
8	203	4	6	15	7	3	13	33	12	4	1
9	195	3	6	11	14	4	9	33	23	3	0
10	188	1	3	2	8	6	9	4	66	12	2
11	221	0	0	1	6	3	5	39	23	2	1
12	206	0	1	0	6	3	8	60	15	1	1
13	221	0	0	2	4	3	4	47	18	1	1
14	216	1	1	3	3	2	8	56	10	0	1
15	216	1	0	1	3	1	4	52	16	3	4
16	202	2	2	1	5	4	6	54	24	1	0
17	212	0	1	3	7	2	5	48	16	2	5
18	186	0	4	7	7	8	11	51	22	5	0
19	214	1	6	0	9	4	5	44	16	1	1
20	215	1	2	4	6	3	6	45	18	1	0
21	223	0	3	4	6	4	5	41	13	1	1
22	216	1	4	11	7	6	4	2	46	3	1
23	238	4	2	3	9	6	5	0	30	2	2
24	257	0	1	3	5	1	4	26	2	0	1

基于系统聚类法的各时段供求匹配程度分级^[6]

1. 系统聚类的概述与步骤

系统聚类是聚类方法中方法最多、理论最完整的聚类方法，其聚类效果也较好。系统聚类不仅可以用于样本聚类，也可以用于变量聚类，它的优点是可以清楚地了解聚类的整个过程。

在进行系统聚类的过程中首先对数据进行初始化，根据样品的特征，规定样品之间的距离，将所有的距离制作成一张表格。其次进行并类选取表格中做小的非零数，然后计算新类距离，利用递推公式计算新类和其他类之间的距离。之后进行迭代，在表中选取最小的非零数，其对应的两类又构成新类，再利用递推公式计算新类与其他类之间的距离，分别删除原表中的相应的行与列，并新增一行和一系列添上的新类和旧类之间的距离，结果产生新的表格，依次类推直至所有的

样本点归为一类为止，最后将样本进行聚类。

2. 系统聚类的求解及结果分析

(1) 工作日打车难易时段聚类模型的求解与结果分析

将所给不同时段打车难易度指标进行系统聚类，根据这些一级指标的数据进行聚类，可以得到各时段聚类分析冰挂图和打车难易时段分析聚类图(见附录)。利用打车难易时段分析聚类图可以得到不同时段分类数，利用这些冰挂图可以对出租车资源“供求匹配”进行分类

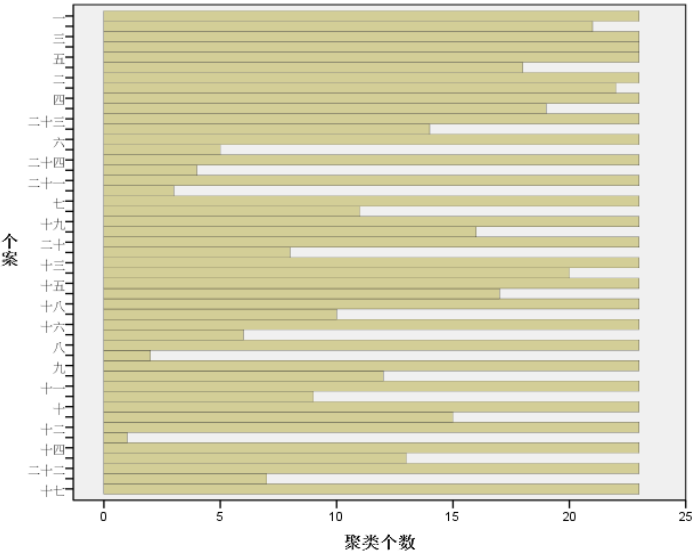


图 6 出租车资源“供求匹配”冰挂图

在进行分类的过程中，需要遵循将组类距离控制在 5 左右，而组间距离大概在 15~20 之间的原理，如果组类距离远小于组建距离的话，那么我们可以认为分组时合理的。

根据图中的数据进行分组，结果如下表所示：

表 8 出租车工作日“供求匹配”冰挂图分组

组号	时段	组内距	组间距
1	二十一时至次日六时	4	9
2	七、八、十三、十五、十八至二十时	6	19
3	九至十二时	8	17
4	十四、二十二、十七时	7	13

通过上表的组内距和组间距结果可以分析出分 4 组是可行的。第 1 组的组内距有波动，与标准不符，但是该组的组内距在分组 4 组以上是不可避免的，所以可以忽略。组间距都在 7 左右。第一组和第二组的组间距是确定的。分组对二者的组间距是不影响的，从而可以得到该种分组方法比较可靠，所以对于平时打车难易度聚类分析我们将所有的时段聚类为 4 组。

(2) 节假日打车难易时段聚类模型的求解与结果分析

对于节假日打车难易时段的聚类模型，根据上文解算方法，得到节假日时段分析聚类图(见附录)和节假日打车难易聚类分析冰挂图如下：

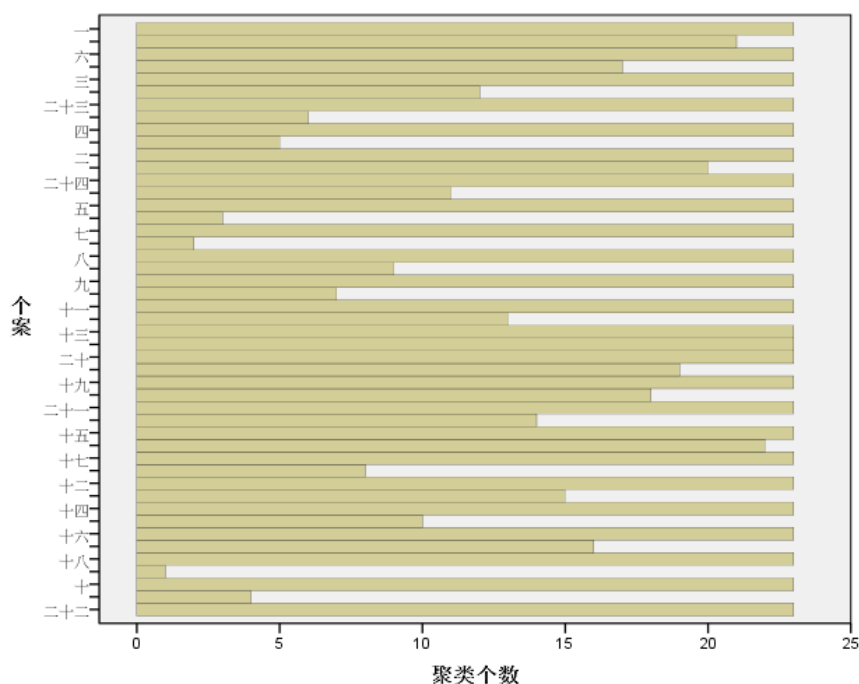


图 7 节假日打车难易聚类分析冰挂图

在进行分组的过程中，首先利用平时打车的聚类方法进行聚类。在进行聚类之后，根据它的分组标准进行分组，得到最接近分组标准的相似分组，如下表所示：

表 9 出租车节假日日“供求匹配”冰挂图分组（1）

组号	时段	组内距	组间距
1	二十三时至次日六时	5	17
2	七时	0	0
3	八九时、十一时至二十一 时	7	9
4	十时、二十二时	4	12

表 10 出租车节假日日“供求匹配”冰挂图分组（2）

组号	时段	组内距	组间距
1	二十三时至次日七时	5	17
2	八时、九时、十一时至二 十一时	7	9
3	十时、二十二时	4	12

通过比较这两组分组，可以看出两组的组内距大小基本相同，但是分成 3 组比分成 4 组的组间距大，更加符合标准，所以经过筛选我们选取将节假日聚类分组成 3 组的方案。

3. 供求匹配程度分级

(1) 工作日

对工作日聚类结果进行分析，我们可以得出：

出租车资源供求匹配程度好的时间段为二十一时至次日六时，

出租车资源供求匹配程度较好的时间段为十四、二十二、十七时，

出租车资源供求匹配程度较差的时间段为九至十二时，

出租车资源供求匹配程度差的时间段为七、八、十三、十五、十八至二十时。

(2) 休息日

对休息日聚类结果进行分析，我们可以得出：

出租车资源供求匹配程度好的时间段为二十三时至次日七时，

出租车资源供求匹配程度中的时间段为十时、二十二时，

出租车资源供求匹配程度差的时间段为八时、九时、十一时至二十一时。

至此我们从时间角度完成了对出租车资源供求匹配程度的分析。

五、问题二模型的建立与求解

5.1 问题分析

该问要求我们分析打车软件公司的补贴方案能否缓解打车难问题，由于出租车的满载率能够反映出出租车资源的供求匹配程度，满载率提高将优化供求匹配程度，从而缓解打车难问题，据此我们将缓解打车难问题定义为出租车的满载率提高，即用满载率上升率反映打车难问题的缓解度。

就补贴方案而言，其作用具体表现在对司机和乘客的补贴能够吸引更多的司机和乘客使用打车软件，从而提高信息共享率和出租车载客效率，缓解打车难的问题，即提高满载率。说明补贴方案能够对缓解打车难问题产生作用，从数学范畴上就是指补贴金额和满载率上升率存在一定的函数关系，下面我们将具体求其函数关系，并用此函数关系评估现有主流打车软件公司的补贴方案对缓解打车难的作用和影响。

5.2 仅受补贴金额影响的出租车满载率上升率模型

由于我们主要为了求解补贴金额和满载率上升率的函数关系，所以我们需要剔除其他因素的影响，建立仅受补贴金额影响下的出租车满载率上升率模型。

5.2.1 出租车满载率函数

通过在客流集散较为集中的地点选取几个长期观测点，统计单位时间通过道路观测点的载有乘客的出租车数量占总通过的出租车数量的比值，该比值就叫做出租车满载率，即：

$$\text{出租车满载率} = \frac{\text{载客车辆数}}{\text{总通过车辆数}} \times 100\%$$

根据其定义，出租车满载率由乘客和出租车司机两个群体共同决定。从乘客方面考虑，其对出租车满载率的影响主要体现在人口数量和人均收入水平两方面；从出租车司机方面考虑，其对出租车满载率的影响主要体现在出租车数量和出租车工作效率上，所谓出租车工作效率，就是指出租车司机能遇到乘客的机会大小，在现实生活中，出租车司机有相当一部分时间仍然花费在寻找乘客和等待乘客上，并不是一直都有机会能够遇到乘客，该机会大小，将能够决定其载客数量，从而影响其车辆满载率。具体关系图如下所示：

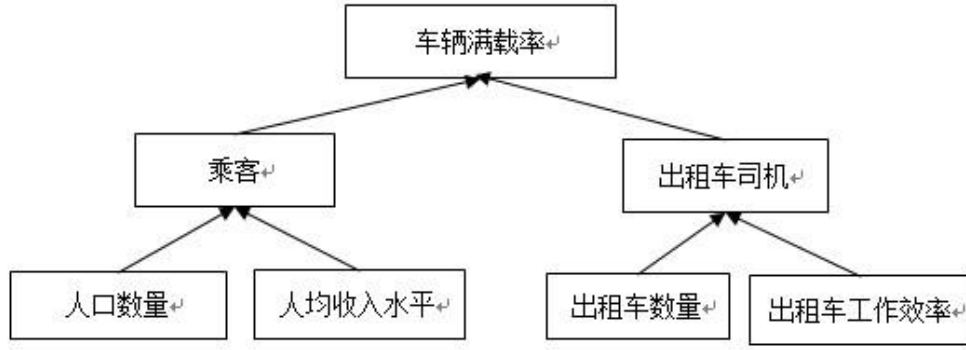


图 8 出租车满载率关系图

根据上文所分析的出租车满载率影响因素，设出租车满载率函数为：

$$N = f(P, M, T, E)$$

其中 N 表示出租车满载率， P 表示人口数量， M 表示人均收入水平， T 表示出租车数量， E 表示出租车工作效率。

5.2.2 出租车满载率函数自变量和因变量的关系

通过分析出租车满载率自变量和因变量在实际情况中的联系和相互间的作用，得出其在函数中的关系如下。

人口数量

城市人口越多，出行人数越多，对出租车的需求量更大，出租车载客数量增加，导致满载率提高，所以人口数量与出租车满载率成正相关，即：

$$N = k_1 P$$

其中 k_1 为常数，且 $k_1 > 1$ 。

人均收入水平

居民人均收入水平越高，选择出行方式时将更倾向于选择成本较高但相对便捷的交通方式，相比于公共交通方式，此时出租车被选择作为出行方式的机会将会更大，出租车载客频率也将提高，从而提升出租车的满载率，所以人均收入水平和出租车满载率成正相关，即：

$$N = k_2 M$$

其中 k_2 为常数，且 $k_2 > 1$ 。

出租车数量

出租车数量越多，其相互之间的竞争越大，载客机会将会减小，相应的出租车的车辆满载率也会降低，所以出租车数量和出租车满载率成负相关，即：

$$N = k_3 T$$

其中 k_3 为常数，且 $0 < k_3 < 1$ 。

出租车工作效率

出租车的工作效率越高，司机在行驶过程中遇到乘客的机会越大，载客机会越高，所以出租车满载率将提高，所以出租车工作效率和出租车满载率成正相关，即：

$$N = k_4 E$$

其中 k_4 为常数，且 $k_4 > 1$ 。

5.2.3 出租车满载率上升率函数

根据出租车满载率函数：

$$N = f(P, M, T, E)$$

可知出租车满载率和人口数量、人均收入水平、出租车数量和出租车工作效率这四个因素有关。

再根据其自变量和因变量的关系：

$$\begin{cases} N = k_1 P \\ N = k_2 M \\ N = k_3 T \\ N = k_4 E \end{cases}, \text{ 其中 } k_1, k_2, k_4 > 1, 0 < k_3 < 1$$

得出该四个因素与出租车满载率成正相关或负相关关系。

结合以上两方面分析，可推导出如下出租车满载率上升率函数：

$$\alpha = \frac{(1+\beta)(1+\gamma)(1+\lambda)}{1+\mu} - 1$$

其中 α 表示出租车满载率上升率， β 表示人口数量增加率， γ 表示人均收入水平提升率， μ 表示出租车数量增加率， λ 表示出租车工作效率提高率。

5.2.4 以补贴金额为自变量的出租车工作效率提高率函数

出租车的工作效率其本质为出租车司机在行驶过程中遇到乘客的机会，传统的打车模式中，乘客一般采用招手打车的方式，这种方式随机性较大且主动性不高，由于乘客和司机间信息不对称造成出租车的空驶。而打车软件能够解决这种信息不对称的问题，优化资源配置，提高司机的载客效率，减少空驶时间，所以打车软件使用的普及，将对出租车的工作效率产生积极影响，即使用打车软件的乘客群体和司机群体数量越多，出租车的工作效率越高。

各家公司的补贴方案具体表现在对司机的补贴金额和对乘客的补贴金额上，补贴金额的数量将影响乘客和司机使用打车软件的积极性，补贴金额越高，其积极性越强，使用打车软件的乘客和司机的数量增长越快，出租车的工作效率提高得也越快。由此可见，出租车的工作效率提高率与补贴金额成正比关系，其具体关系可用下图表示：

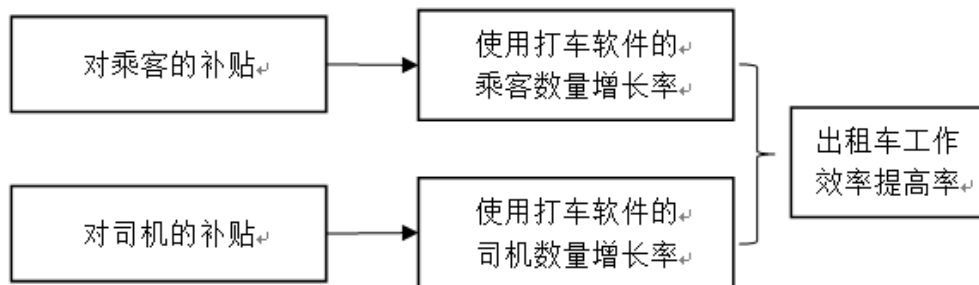


图 9 出租车工作效率关系图

推导求得出以补贴金额为自变量的出租车工作效率提高率函数为：

$$\lambda = f(B_1, B_2)$$

其中 λ 表示出租车工作效率， B_1 表示对乘客的补贴金额， B_2 表示对司机的补贴金额。

5.2.5 以补贴金额为自变量的出租车满载率上升率函数

结合上文推导得出的满载率上升率函数和以补贴金额为自变量的租车工作效率提高率函数：

$$\begin{cases} 1 + \alpha = \frac{(1 + \beta)(1 + \gamma)(1 + \lambda)}{1 + \mu} \\ \lambda = f(B_1, B_2) \end{cases}$$

联解求得以补贴金额为自变量的出租车满载率上升率函数：

$$\alpha = \frac{(1 + \beta)(1 + \gamma)[1 + f(B_1, B_2)]}{1 + \mu} - 1$$

5.2.6 函数表达式 $f(B_1, B_2)$ 的求解

由于以补贴金额为自变量的出租车满载率上升率函数中出租车工作效率提高率 λ 未知，所以我们需要对 λ 即：

$$\lambda = f(B_1, B_2) = \frac{\alpha + \mu + \alpha\mu - \beta - \gamma - \beta\gamma}{1 + \beta + \gamma + \beta\gamma}$$

进行求解，进而得出以补贴金额为自变量的出租车满载率上升率函数。

通过多方渠道信息，打车软件最早始于 2012 年 UBER 推出叫车服务，所以在数据的搜集方面我们只需要搜集 2012 年到 2015 年这四年来历年的出租车满载率上升率、人口数量增加率、人均收入水平提升率和出租车数量增加率，同时统计归纳这四年来主流打车软件公司历年的补贴方案即对乘客和对司机的补贴金额，再通过对这些数据进行拟合得出函数表达式。

表 11 2012-2015 年出租车满载率上升率

年份	2012	2013	2014	2015
出租车满载率	60%	74%	80%	81%
增长率	——	23.3%	13.5%	1.25%

表 12 2012-2015 年人口数量增加率

年份	2012	2013	2014	2015
全国城镇人口（万人）	69079	71182	73111	74916
增长率	3.14%	3.04%	2.71%	2.47%

表 13 2012-2015 年人均收入水平提升率

年份	2012	2013	2014	2015
城镇人均可支配收入	24565	26955	28844	31151
增长率	12.6%	8.1%	8%	8% (估)

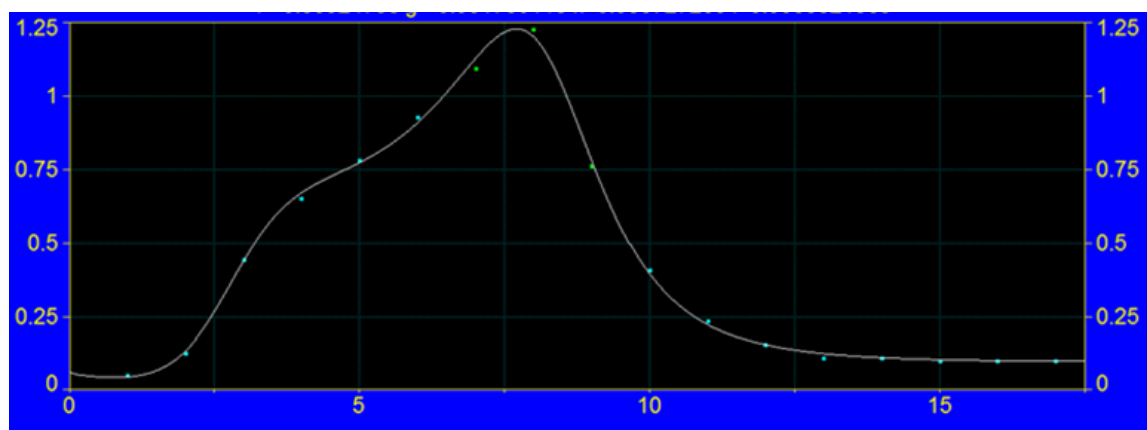
表 14 2012-2015 年出租车数量增加率

年份	2012	2013	2014	2015
全国出租车数量	1026678	1053580	1083080	1112323
增长率	2.43%	2.6%	2.8%	2.7% (估)

2012-2015 年主流打车软件公司补贴方案

拟合函数表达式 $f(B_1, B_2)$

通过以上搜集得到的 2012 年到 2015 年这四年来历年的出租车满载率上升率、人口数量增加率、人均收入水平提升率和出租车数量增加率，以及四年来的主流打车软件公司历年的补贴方案，对这些数据用 Table curve2D 软件进行拟合如下图所示：



得到出租车工作效率提升率

$$f(B_1, B_2) = -2.4141 + 0.0003B_1^2B_2^2 - 0.0189B_1^2 - 0.082B_2^2 - 0.0440B_1B_2 + 0.3073B_1 + 1.0766B_2$$

5.3 当今主流打车软件公司对打车难问题的缓解度分析

滴滴打车其现今的补贴方案为对乘客补贴 5 元，对司机补贴 2 元，求得其出租车工作效率提升率为 6.51%

快的打车其现今的补贴方案为对乘客补贴 4 元，对司机补贴 3 元，求得其出租车工作效率提升率为 5.19%

对比可知，两家主流打车软件公司都对缓解打车难问题有帮助，其中滴滴打车的补贴方案对缓解打车难问题的效果更明显。

六、问题三模型的建立与求解

6.1 问题分析

该问需要我们提出补贴方案并论证其合理性，对于如何提出补贴方案，可以转化为优化问题，即如何提出最优补贴方案，最优性由其对打车难缓解程度和公司自身盈利共同决定，所以这将是一个多目标规划问题。在此我们需要定义目标和约束，构建多目标规划，对模型求解得到最佳补贴方案，并分别从其对打车难的缓解能力和公司的盈利论证其合理性。

6.2 基于多目标规划的最佳补贴方案模型

在此我们通过建立多目标规划模型并求解，根据所得结果得出软件公司的最佳补贴方案。

6.2.1 模型准备

多目标规划^[7]

多目标规划就是存在多个目标的规划问题。只要目标函数为两个或者两个以上，都可以归为多目标规划的范畴。

一般的多目标规划问题具有如下的形式：

$$\text{Min}F(x) = (f_1(x_1, x_2, \dots, x_m), \dots, f_k(x_1, x_2, \dots, x_m))$$

$$\text{Max}G(x) = (g_1(x_1, x_2, \dots, x_m), \dots, g_k(x_1, x_2, \dots, x_m))$$

$$\text{s.t.} \begin{cases} h_i(x_1, x_2, \dots, x_m) \geq 0, i = 1, \dots, s \\ h_j(x_1, x_2, \dots, x_m) \leq 0, j = s+1, \dots, s+t \end{cases}$$

其中， $k+n \geq 2$ ； $F(\cdot)$ 和 $G(\cdot)$ 是函数向量，元素 f_i 和 g_i 是目标函数； $x = (x_1, x_2, \dots, x_m)^T$ 是决策变量， $h_j(\cdot)$ 是约束条件。

上述是混合模型，通过最大值与最小值的转化可以将上述模型变为标准的多目标最大化或者多目标最小化模型。不同于单目标模型，多目标决策中经常不存在使所有目标达到最优化的解，而只能使其中的某些目标达到最优。若使另一些目标也达到最优，则会降低已经达到最优的目标的优化程度。所以多目标规划中经常不存在绝对最优解，其最优解是以有效解或者非劣解的形式出现。若 x 是有效解，则不存在另一可行解 y ，使得 y 对应的所有目标函数的值都不劣于 x 对应的目标值，而且至少有一个 y 对应的目标要优于 x 相应的目标。由于多目标规划的有效解集中有效解的数目众多，因此，多目标决策的一个核心问题就是如何根据决策者的主观价值对有效解进行比较分析。

决策者在选择的时候需要在有效解之间进行权衡，进而得到最终的满意解。由于决策者自身的决策原则不同，因此，它们根据自身的偏好从众多有效解中得到的满意解也往往是不同的。如果能够将决策者的偏好信息、选择原则通过关于各个目标的效用函数来表示，那么就能够将多目标规划问题转化为单目标规划问题，进而借助单目标理论进行求解，然而事实上，很难得到能够确切表示决策者偏好的函数。所以长期以来，理论界对多目标的求解技术进行了深入的探讨，得到了许多求解多目标规划模型的一般方法。

多目标规划求解方法

在约束条件限制下，使多目标规划中的目标函数尽可能的接近预定的目标值。不同于前面的理想点法，此处预定的目标值只是反映了决策者的期望，不存在任何限制，是满足既定条件的逼近解，基本形式如下：

$$\begin{aligned} & \text{Min}[P_s \sum_{j=1}^n (\lambda_{s,j}^+ \delta_j^+ + \lambda_{s,j}^- \delta_j^-)]_{s=1}^L \\ & \text{s.t.} \begin{cases} f_j(x) - \delta_j^- + \delta_j^+ = f_j^0 \\ \delta_j^- \geq 0, \delta_j^+ \geq 0, j=1, \dots, n \\ h_j(x_1, x_2, \dots, x_m) \geq 0, i=1, \dots, k \\ h_j(x_1, x_2, \dots, x_m) \leq 0, j=k+1, \dots, k+t \end{cases} \end{aligned}$$

其中 P_s 表示优先层级，共有 L 个层级。具体过程如下：

首先计算第一层级 P_1 ，得到对应的最优解，然后求解 P_2 ，但是要以 P_1 的计算结果为限制条件。依次继续，每一步都是一个单目标规划，最后一个层级得到的结果是多目标规划的最优解。

6.2.2 建立最佳补贴方案多目标规划模型

1.模型目标

(1) 打车软件公司盈利目标

打车软件公司盈利 S 必须最大化，所以有

$$\max S$$

(2) 出租车满载率上升率目标

出租车满载率上升率越大，对缓解打车难问题的帮助越大，所以出租车满载率上升率必须最大化，所以有

$$\max \alpha$$

2.模型约束

(1) 打车软件公司盈利非负约束

打车软件公司是以盈利为目标的机构，所以盈利必须非负，所以有

$$S \geq 0$$

(2) 出租车满载率上升率非负约束

出租车满载率上升率不能下降，否则将会加剧打车难现象，所以有：

$$\alpha \geq 0$$

(3) 使用打车软件乘客数量约束

使用打车软件的出租车司机数量 A_1 不能超过出租车司机的总数 m ，所以有：

$$A_1 \leq m$$

(4) 使用打车软件出租车司机数量约束

使用打车软件的乘客数量 A_2 不能超过出租车司机的总数 n ，所以有：

$$A_2 \leq n$$

(5) 补贴非负约束

打车软件公司不可能在市场尚未稳固的情况下，主动向司机和乘客索要费用，所以有补贴非负约束：

$$B_1 \geq 0, B_2 \geq 0$$

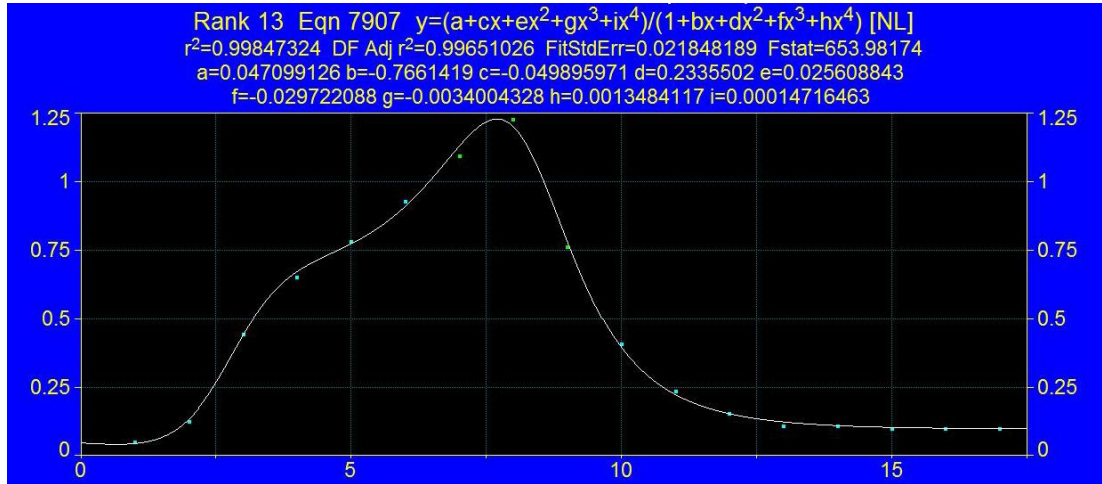
3.对乘客的补贴与使用打车软件乘客数量的关系

统计得到历年对乘客的补贴和乘客注册数如下：

表 15 历年对乘客的补贴和乘客注册数

年份	补贴（元）	乘客注册数（万）
2012	10	340
2013	17	1920
2014	9	2600
2015	3	3960

通过 Table curve2D 软件进行如下拟合



得到对乘客的补贴与使用打车软件乘客数量的关系

$$B_1 = \frac{0.0126 - 0.0197A_1 + 0.000015A_1^2 - 0.000000554A_1^3 + 0.00000000142A_1^4}{1 - 0.0221A_1 + 0.000219A_1^2 - 0.000000966A_1^3 + 0.00000000154A_1^4}$$

4.对出租车的补贴与使用打车软件出租车司机数量的关系

统计得到历年对司机的补贴和司机注册数如下：

表 15 历年对司机的补贴和司机注册数

年份	补贴（元）	司机注册数（万）
2012	10	60
2013	15	240
2014	5	300
2015	0	340

同理得到对出租车司机的补贴与使用打车软件的出租车司机数量的关系

$$B_2 = \frac{-0.00415 + 0.005A_2 + 0.0000509A_2^2 - 0.000000208A_2^3 + 0.00000000046A_2^4}{1 - 0.0208A_2 + 0.000169A_2^2 - 0.000000613A_2^3 + 0.000000000811A_2^4}$$

5.打车软件公司盈利与使用软件的司机和乘客数量的关系

根据相关渠道统计，得到每天每名司机使用打车软件可为打车软件公司带来 4 元收入，而乘客可以带来 6 元收入，所以打车软件公司盈利与使用软件的司机和乘客数量的关系为

$$S = 6A_1 + 4A_2 - A_1B_1 - A_2B_2$$

6.模型的建立

根据上文确立的目标、约束和关系，建立双目标规划模型如下：

$$\begin{aligned} & \max S, \max \alpha \\ s.t. & \begin{cases} S = 6A_1 + 4A_2 - A_1B_1 - A_2B_2 \\ B_2 = \frac{-0.00415 + 0.005A_2 + 0.0000509A_2^2 - 0.000000208A_2^3 + 0.00000000046A_2^4}{1 - 0.0208A_2 + 0.000169A_2^2 - 0.000000613A_2^3 + 0.000000000811A_2^4} \\ B_1 = \frac{0.0126 - 0.0197A_1 + 0.000015A_1^2 - 0.000000554A_1^3 + 0.00000000142A_1^4}{1 - 0.0221A_1 + 0.000219A_1^2 - 0.000000966A_1^3 + 0.00000000154A_1^4} \\ S, \alpha, B_1, B_2 \geq 0 \\ A_1 \leq m \\ A_2 \leq n \end{cases} \end{aligned}$$

6.2.3 模型的求解

对上文中的双目标规划模型进行求解，得到最优解为：

$$\begin{cases} B_1 = 3 \\ B_2 = 0 \end{cases}$$

即对乘客补贴每单 3 元，对司机不实行补贴。

6.3 模型的合理性论证

通过双目标规划，我们提出对乘客补贴每单 3 元，对司机不实行补贴的政策，下面分别就打车软件的持续发展、对缓解打车难问题的帮助和公司的盈利模式进行论证。

1. 打车软件的持续发展

在该补贴模式下，能够保持使用软件的乘客数量和出租车司机数量持续增长，从而保证打车软件的生命力，促进其持续发展。

2. 对缓解打车难问题的帮助

由于能够保证出租车满载率上升率非负，所以该补贴方案下，出租车满载率将持续上升，打车难的问题将会被逐渐解决。

3. 公司的盈利模式

通过双目标规划模型得出的打车补贴方案，能够保证公司盈利，满足公司存在的基本条件，使其能够长久经营下去。

综合以上分析论证，证明了我们提出的补贴方案的合理性。

七、模型的评价

7.1 模型的优点

1、本文模型是应用综合模型，很好地克服了主观因素对模型结果的影响，使模型的实用性更好，更加贴近实际。

2、本文的在数据处理方面，消除了量纲的缺陷，使数据得到了充分的利用。

7.2 模型的有待改进之处

本文的数据由网络收集而来,真伪性不好分辨,可能与实际的结果有所偏离。

参考文献

- [1]. 王一帆. 基于打车软件的出租车服务模式优化研究[D]. 上海交通大学, 2014
- [2]. 郭亚军, 李玉玲, 易平涛, 李伟伟. 拉开档次法稳定性分析及改进[J]. 运筹与管理, 2015
- [3]. 韩立岩, 周芳. 基于 D-S 证据理论的知识融合及其应用[J]. 北京航空航天大学学报, 2006
- [4]. 赵萱, 张权, 樊治平. 多属性决策中权重确定的主客观赋权法[J]. 沈阳工业大学学报, 1997
- [5]. ChaoSimple, 数据归一化和两种常用的归一化方法,
<http://www.cnblogs.com/chaosimple/archive/2013/07/31/3227271.html>, 2015/9/13
- [6]. 段明秀. 层次聚类算法的研究及应用[D]. 中南大学, 2009
- [7]. 王丽珍. 基于多目标规划的产险公司最优资本规模研究[D]. 南开大学, 2012

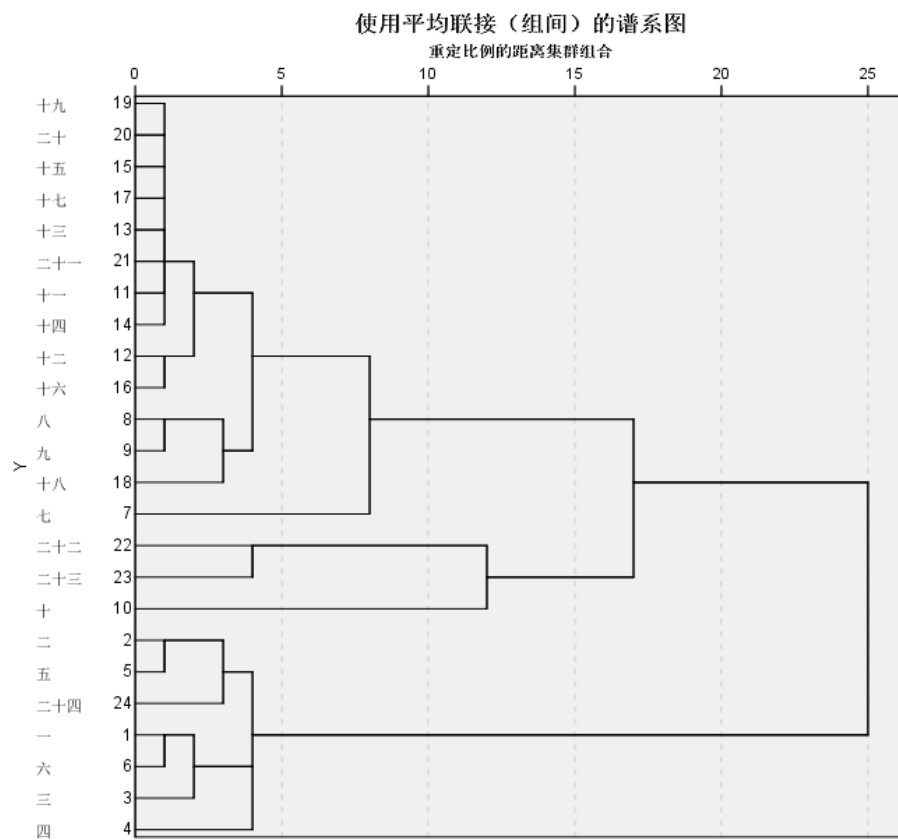
附录

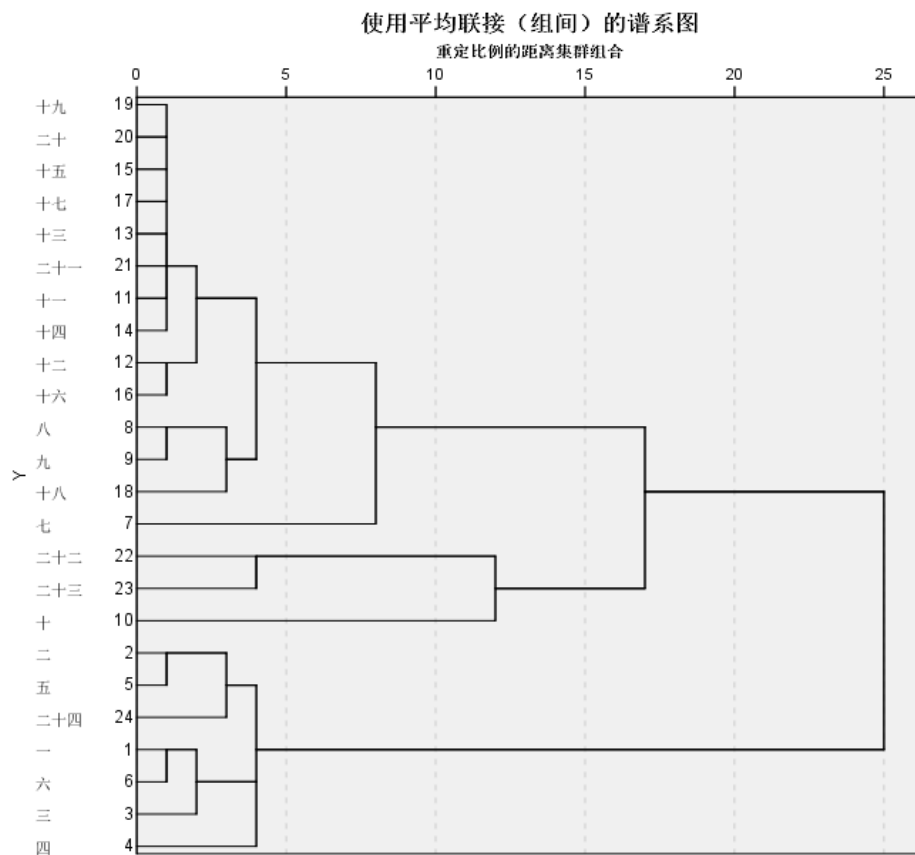
附录 1 邻接矩阵 MATLAB 程序

邻接矩阵程序

```
A=input(' 输入矩阵:');  
I=eye(14);  
B=A+I;  
i=0;  
while 1  
    Bnew=B*(A+I)>0;  
    if isequal(B, Bnew)  
        Bnew  
        i=i+1  
        break  
    end  
    B=Bnew;  
    i=i+1;  
end  
answer=Bnew
```

附录 2





不同时段打车难易谱系图

附录 3 不同时期工作效率影响表

月份	F (B1, B2)	司机 B2	乘客 B1
1	0.0518	10	10
2	0.1243	13	12
3	0.4459	17	8
4	0.6538	15	9
5	0.7842	8	7
6	0.928	5	7
7	1.097	3.2	6.3
8	1.288	3.32	4.6
9	0.7638	2.95	3.4
10	0.4092	2.58	3.5
11	0.2376	1.34	3.2
12	0.1537	0	3
13	0.108	0.237	3.21
14	0.109	0.245	2.96
15	0.0999	0.012	3.01
16	0.0992	0	2.987
17	0.1003	0	3.026
18	0.0987	0	2.9765

附录 4 不同年份乘客注册数

年份	补贴（元）	乘客注册数（万）
2012	10	340
2013	17	1920
2014	9	2600
2015	3	3960

附录 5 不同年份司机注册数

年份	补贴（元）	司机注册数（万）
2012	10	60
2013	15	240
2014	5	300
2015	0	340

附录 6 不同补贴下的司机与乘客数

补贴	乘客数（十万）	补贴	司机数（万）
0	0	0	3
0.683696	20	0.603261	40
1.640606	40	1.447593	70
5.885326	70	5.192935	100
8.629348	100	7.61413	130
10.35047	130	9.132764	160
12.24845	160	10.80745	190
14.47904	190	12.77562	220
17.0000	230	15.0000	260
10.08121	250	8.895186	270
5.400932	270	4.765528	280
3.136025	290	2.767081	290
2.028649	310	1.789984	300
1.425466	330	1.257764	310
1.438665	350	1.26941	320
1.318556	370	1.163432	330
1.309317	385	1.15528	340
1.323835	395	1.16809	350
1.302717	410	1.149457	360