

报名号 # 1384

第三届“ScienceWord 杯”数学中国

数学建模网络挑战赛
承 诺 书

我们仔细阅读了第三届“ScienceWord 杯”数学中国数学建模网络挑战赛的竞赛规则。

我们完全明白，在竞赛开始后参赛队员不能以任何方式（包括电话、电子邮件、网上咨询等）与队外的任何人（包括指导教师）研究、讨论与赛题有关的问题。

我们知道，抄袭别人的成果是违反竞赛规则的，如果引用别人的成果或其他公开的资料（包括网上查到的资料），必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺，严格遵守竞赛规则，以保证竞赛的公正、公平性。如有违反竞赛规则的行为，我们将受到严肃处理。

我们允许数学中国网站(www.madio.net)公布论文，以供网友之间学习交流，数学中国网站以非商业目的的论文交流不需要提前取得我们的同意。

我们的参赛报名号为：1384

参赛队员（签名）：

队员 1：

队员 2：

队员 3：

参赛队教练员（签名）：

参赛队伍组别：大学组

报名号 # 1384

第三届“ScienceWord 杯”数学中国

数学建模网络挑战赛

编 号 专 用 页

参赛队伍的参赛号码：（请各个参赛队提前填写好）：

竞赛统一编号（由竞赛组委会送至评委团前编号）：

竞赛评阅编号（由竞赛评委团评阅前进行编号）：

报名号 # 1384

2010 年第三届 “ScienceWord 杯” 数学中国 数学建模网络挑战赛

题 目 图书馆的智能服务问题关 键 词 数据处理、聚类分析、综合评价、博弈、回归分析

摘 要：

通过对题目的深入分析，本文是数据分析和综合评价类问题。

对于第一问，1) 本文用模糊聚类分析算法对超期书籍进行分类，建立模糊聚类模型，以简化数据。2) 将超期天数分成三个阶段，对每一阶段导致借阅超期的原因按主次分别进行量化；统计出已分类的每类书籍中处于每一超期阶段的书籍数目，并以此为输入，分别用主观赋权法和客观分析法求解每种原因引起超期的权重；为了消除上述两种方法过于主观或过于客观的缺点，结合其优点，我们用基于离差最大化的综合赋权算法将上面两结果综合，得到最终的各种原因引起借阅超期的权重。3) 本文对每类超期书籍建立灰色关联——模糊综合评价模型及基于 ANFIS 的各类书籍超期程度的综合评价模型进行综合评价，通过变成求解得到各类超期书籍在总超期书籍中所占权重，即每类书籍容易超期的程度。

对于第二问，我们首先根据博弈论对图书馆智能服务与借阅人员的借书问题进行定性分析，得到两者之间存在一均衡点；然后对题目中所给出的策略根据第一问中所确定的因素进行简单评价，并利用回归模型综合考虑图书种类、借书者造成超期原因以及超期天数与图书馆进行罚款处罚之间的关系提出了新的方案。

最后，本文对建立的模型进行评价并分析进一步的改进方法。

参赛队号 1384所选题目 C

参赛密码 _____

(由组委会填写)

报名号 # 1384

英文摘要（选填）

Library Intelligent service issues

Through in-depth analysis of the topic, this is the data analysis and evaluation issues.

For the first question, 1) the use of the extended fuzzy clustering algorithm to classify books, to build a fuzzy clustering model, to simplify the data. 2) the number of days will be extended into three stages, each stage led by primary and secondary reasons for borrowing were extended to quantify; the statistics have been classified in each category of books for each extended period in the number of books and use it as input, respectively, with the subjective weighting method and objective analysis of the causes of each method of extended weight; in order to eliminate the above-mentioned two methods are too subjective or too objective shortcomings, combined with its advantages, we use based on maximizing deviations above the integrated enabling algorithm Comprehensive two results, a variety of causes, the final weight of overdue loan. 3) For each category of the establishment of extended book Grey - Fuzzy comprehensive evaluation model and the types of books based on ANFIS evaluation of extended degree evaluation model, extended through into solving all kinds of books are books in the share of the total extended weight of each type of book that is easily extended level.

For the second question, we first game of the library according to intelligence services and lending library personnel issues for qualitative analysis, there is an equilibrium between the two; then given the title of the strategy in accordance with the first question the factors identified in a simple evaluation, and use regression model took into account books Zhong Lei, borrowers has caused the number of days of extended reasons and extended the library to fine the relationship between punishment put forward new proposals.

Finally, the paper evaluate the model we have established and analysis the further improvement.of the model.

报名号 # 1384

目录

高校图书的馆智能服务问题	3
1 引言	3
2 基本假设	3
3 相关符号说明	3
4 问题的分析与建模思路	4
4.1 问题一的分析	4
4.2 问题二的分析	4
5 问题一模型的建立及求解	4
5.1 模糊聚类分析模型的建立	5
5.1.1 模型的准备	5
5.1.2 模糊聚类分析算法原理及步骤	5
5.1.3 编程求解	6
5.2 借阅超期影响因素权重确定	6
5.2.1 对各影响因素进行量化	6
5.2.2 层次分析法的具体步骤:	7
5.2.3 编程求解	8
5.3 客观赋权（改进的熵权法）模型的建立	8
5.3.1 熵权法的原理	8
5.3.2 熵权法的步骤	8
5.3.3 改进的熵权法	9
5.3.4 用 MATLAB 软件编程求解	10
5.4 离差平方组合赋权模型的建立	10
5.4.1 确定最优化模型	10
5.4.2 构造 Lagrange 函数	11
5.4.3 编程求解	12
5.5 灰色关联——模糊综合评价及排序	12
5.5.1 灰色关联法求解步骤	12
5.5.2 编程求解	13
5.6 基于 ANFIS（自适应神经模糊推理系统）的书籍易超期程度评价模型	13
6 问题二模型的建立与求解	18
6.1 博弈模型	18
6.1.1 博弈论方法与纳什均衡的概述	18
6.1.2 高校收费的纳什均衡理论分析	18
6.2 回归分析模型的建立	20

报名号 # 1384

6.2.1 多元线性回归原理.....	20
6.2.2 模型的建立.....	20
6.2.3 参数估计.....	21
6.2.4 统计分析.....	21
6.2.5 回归模型的假设检验.....	22
6.2.6 回归系数的假设检验和区间估计.....	22
6.2.7 对本文的具体模型的建立.....	23
6.2.8 模型的应用.....	23
7 模型的评价与改进	23
7.1 模型的评价	23
7.1.1 优点.....	23
7.1.2 缺点.....	24
7.2 模型的改进	24
8 参考文献	24
9 附录	25

附表索引

表 3-1 符号说明一览表.....	4
表 5-1 聚类分析及统计结果.....	6
表 5-2 超期天数各原因量化表.....	6
表 5-3 尺度 a_{ij} 的含义.....	7
表 5-4 准则层对目标层判断表.....	7
表 5-5 Saaty 的一致性指标 RI	7
表 5-6 层次分析法所求各影响因素权重.....	8
表 5-7 熵权法所求各影响因素权重.....	10
表 5-8 各影响因素综合权重.....	12
表 5-9 各类书籍容易超期程度排序.....	13
表 5-10 书籍容易超期程度需求结果.....	16
表 5-11 各类书籍容易超期程度加权平均结果.....	17
表 5-12 各类书籍各种超期原因所占权重.....	17

报名号 # 1384

高校图书馆的智能服务问题

1 引言

目前高校图书馆存在着许多不良的现象，尤其是借阅超期现象。超期现象是读者借阅图书资料时超过规定节约期限后不归还的一种普遍现象。而处理超期罚款是流通服务中常遇到的事，特殊情况则需酌情处理。这一现象虽然是小事，但处理不当会影响图书馆与读者之间的关系，从而影响图书馆的声誉，以及馆藏资源的利用率，最终影响图书馆功能的正常发挥。对于借阅图书超期，原因众多。考虑到高校的特点，大体上可以把原因分为三类：1：无意错过还书期限；2、由于教学科研需要，长期使用；3、由于罚款数目较少，故意拖延。

关于图书超期及图书馆的借还管理等方面问题，目前已有许多专家学者进行过相关方面的研究。如胡福文先生对于高校图书馆读者借阅超期有关问题的思考^[1]，陈京先生对于高校图书馆超期借阅罚款中存在的矛盾以及相应策略的研究等等。这些研究对于高校图书馆超期现象做出了很好的分析，提出了很好的对策。本文所研究的是分析高校图书馆中哪类图书容易超期及其超期的原因，从而可以从针对不同的图书制定相应的惩罚策略，达到更好的效果。

本文所研究的问题可叙述如下：

1.通过给出的数据或自行查找相关数据，建立合理的模型，分析哪些图书容易发生超期现象，以及发生超期时其可能的原因。以便图书馆采取相应的措施改进服务。

2.假设某高校的借书规则为：借书时限为一个月，可以续借一次，续借时限为 15 天，超期一天罚款 0.1 元。试根据问题一中的模型评价这种规则的优劣。请设计一种更加合理的借阅规则。

2 基本假设

结合本题实际，为了确保模型求解的准确性和合理性，我们排除了一些未知因素的干扰，提出了以下几点假设：

- [1]假设附件中提供的数据都是真实可靠的或所给的数据都在误差允许范围之内；
- [2]假设可以忽略某些个别同学对其他的专业书籍的兴趣；
- [3]假设图书馆借阅超期与书籍本身、借书人罚款承受能力等因素有关；
- [4]假设第二文中可以看做图书馆管理方与借书人之间的博弈；
- [5]假设第二文中考虑罚款金额为常数的情况。

3 相关符号说明

为了便于描述问题，我们用一些符号来代替问题中涉及的一些基本变量，如表 3-1 符号说明一览表所示，其它一些变量将在文中陆续说明。

报名号 # 1384

表 3-1 符号说明一览表

主要符号	符号说明
x_i	六类书籍
s_i	各类书籍的数目
a_{ij}	层次分析法中的尺度
c_i	三种影响因素的权重
u_i	影响惩罚金额的三种因素
Y_i	惩罚金额

4 问题的分析与建模思路

4.1 问题一的分析

第一问要求分析哪些书容易超期及其超期的最可能的原因。

由于所给数据繁多，直接利用每个数据进行分析不可行，于是本文首先用模糊聚类分析算法对超期书籍进行分类，建立模糊聚类模型，以简化数据。

然后根据分析，将超期天数分成三阶段，对每一阶段导致借阅超期的原因按主次分别进行量化；然后统计出已分类的每类书籍中处于每一超期阶段的书籍数目（换算成比例），并以此为输入，分别用主观赋权法（层次分析法）和客观分析法（熵权法）求解每种原因引起超期的权重；为了消除上述两种方法过于主观或过于客观的缺点，结合其优点，于是我们用基于离差最大化的综合赋权算法将上面两结果综合，得到最终的各种原因引起借阅超期的权重。

之后，本文对每类超期书籍进行综合评价，以得到各类超期书籍在总超期书籍中所占权重，即每类书籍容易超期的程度。再结合前面步骤可进一步得到每类书籍中每种原因引起借阅超期权重系数。进行综合评价时，我们建立灰色关联——模糊综合评价模型及基于 ANFIS 的各类书籍超期程度的综合评价模型分别求解，并将二者所得结果进行比较，从而得到更精确的结果。

最后，我们根据模型所得结果进行分析，并提出相应的措施及改进方案。

4.2 问题二的分析

对于问题二，我们首先根据博弈论对图书馆智能服务与借阅人员的借书问题进行定性分析，得到两者之间存在一均衡点；然后对题目中所给出的策略根据第一问中所确定的因素进行简单评价，并利用回归模型综合考虑图书种类、借书者造成超期原因以及超期天数与图书馆进行罚款处罚之间的关系提出了新的方案。

5 问题一模型的建立及求解

根据分析，我们首先对超期书籍进行分类。

报名号 # 1384

5.1 模糊聚类分析模型的建立

5.1.1 模型的准备

在许多工程技术和经济管理中，常常需要对某些指标按一定的标准(相似程度、亲疏关系等)进行分类处理。

为了简化数据，对所给数据进行分析，再结合高校的借阅读者特点，可以发现读者证件可以反映学生或教师所在的学院及专业，而同一专业读者所借阅的书籍必定在一定程度上与其专业相关(除了高校基础课，而且这里我们假设可以忽略某些个别同学对其他专业书籍的兴趣)，所以可分为一个类别。根据这一点，我们现根据读者证件号将书籍分为 26 个类别，如证件号前几位为 J000 的超期书籍可以分为一个类别，前几位为 J390 的超期书籍可以分为一个类别，另外注意基础课(大学数学、大学物理等)分为一类。

我们将 26 个类别的书籍数据进行统计。但是 26 个类别对于问题的解决还是比较繁琐，于是我们在已有数据的基础上再进行分区，这里我们用模糊聚类分析方法^[2]。

5.1.2 模糊聚类分析算法原理及步骤

题目中这种对客观事物按一定标准进行分类的数学方法主要就是聚类分析法，而模糊聚类分析法就是根据事物的某些模糊性质进行分类的一种数学方法，下面给出模糊聚类分析方法的一般步骤：

1) 数据标准化

(1) 获取数据：论域 $U = \{x_1, x_2, \dots, x_n\}$ 为所需分类研究的对象，每个对象又由 m 个指标表示其性态，即 $x_i = \{x_{i1}, x_{i2}, \dots, x_{im}\} (i=1, 2, \dots, n)$ ，于是，可以得到原始数据矩阵为

$$A = (x_{ij})_{n \times m}$$

(2) 数据的标准化处理：在实际问题中的数据可能有不同的性质和不同的量纲，为了使原始数据能够适合模糊聚类的要求，需要将原始数据矩阵 A 做标准化的处理，即通过适当的数据交换和压缩，将其转化为模糊矩阵。常用的方法有以下两种：

(i) 平移-标准差变换

如果原始数据之间有不同的量纲，则可以采用这种变换后使每个变量的均值为 0，标准差为 1，即可以消除量纲的差异的影响。即令

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (i=1, 2, \dots, n; j=1, 2, \dots, m)$$

其中

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad s_j = \left[\frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 \right]^{\frac{1}{2}} \quad (j=1, 2, \dots, m)$$

(ii) 平移-极差变换

如果经过平移-标准差变换后还有某些，则还需对其进行平移-极差变换，即令

$$x''_{ij} = \frac{x'_{ij} - \min_{1 \leq i \leq n} \{x'_{ij}\}}{\max_{1 \leq i \leq n} \{x'_{ij}\} - \min_{1 \leq i \leq n} \{x'_{ij}\}} \quad (j=1, 2, \dots, m)$$

显然所有的 $x''_{ij} \in [0, 1]$ ，且也不存在量纲因素的影响，从而可以得到模糊矩阵

报名号 # 1384

$$R = (x_{ij}^*)_{n \times m}。$$

2) 建立模糊相似矩阵

设论域 $U = \{x_1, x_2, \dots, x_n\}$, $x_i = \{x_{i1}, x_{i2}, \dots, x_{im}\}$ ($i = 1, 2, \dots, n$), 即数据矩阵为 $A = (x_{ij})_{n \times m}$, 称之为相似系数, 确定相似系数有多种不同的方法。包括数量积法、夹角余弦法、相关系数法、指数相似系数法、最大最小值法、算术平均法、几何平均值法、绝对值倒数法、绝对值指数法、欧氏距离法等。

其中相关系数法：令

$$r_{ij} = \frac{\left| \sum_{k=1}^m (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) \right|}{\sqrt{\sum_{k=1}^m (x_{ik} - \bar{x}_i)^2 \cdot \sum_{k=1}^m (x_{jk} - \bar{x}_j)^2}} \quad (i, j = 1, 2, \dots, n)$$

3) 聚类

所谓的聚类方法就是依据模糊矩阵将所研究的对象进行分类的方法, 对于不同的置信水平, 可得不同的分类结果, 从而可以形成动态聚类图。常用的方法有传递闭包法、布尔矩阵法、直接聚类法等。

5.1.3 编程求解

根据以上原理及步骤, 我们利用 MATLAB 编程^[3] (代码见附件 1) 进行求解, 得到六个超期书籍分组情况, 根据每个类别的性质我们对其命名, 然后进行相关统计, 得到下表 5-1 聚类分析及统计结果:

表 5-1 聚类分析及统计结果

书籍类别	每类书籍数目
哲学理学	395
社会政治	548
文化艺术	1175
生物化学	417
专业科学技术	1360
基础科学	759

5.2 借阅超期影响因素权重确定

5.2.1 对各影响因素进行量化

导致书籍借阅超期的因素主要可以分为以下三类: 1. 无意错过还书期限; 2. 由于教学科研需要, 长期使用; 3. 由于罚款数目较少, 故意拖延。

根据题目所给数据及高校书籍借阅的特点, 可以知道, 对于超期时间较少的书籍, 其超期原因中“由于罚款数目较少故意拖延”所占的比率较大, 而对于超期时间很长, 甚至超过了一年、两年的书籍, 其超期原因中“无意错过还书期限”的比率较大, 而时间中等的则偏于由于教学科研需要, 长期使用而不归还。根据此分析, 我们根据超期时间将三种原因所占比例进行量化。得到表 5-2 超期天数各原因量化表如下表所示:

表 5-2 超期天数各原因量化表

超期天数	第一种原因所占比例	第二周原因比例	第三种原因所占比例
30 天—150 天	2	1	4
151 天—365 天	1	4	2
一年以上	4	2	1

报名号 # 1384

下面用主观赋权法——层次分析法来确定各影响因素在所有超期书籍中所占的比重。

5.2.2 层次分析法的具体步骤：

将各个因素两两比较确定其相对重要性构造判断矩阵^[4]。常用的是 Saaty 比较尺度，取值可以是 1, 2, ..., 9 及其 1, 1/2, ..., 1/9，这是因为人们在进行定性成对比较时常见有 5 种明显的等级（根据心理学的研究成果）（见下表 5-3 尺度 a_{ij} 的含义）

表 5-3 尺度 a_{ij} 的含义

Y_i/Y_j	同等重要	略为重要	明显重要	尤为重	异常重要	等级之间
r_{ij}	1	3	5	7	9	2, 4, 6, 8

我们得到如下判断表及判断矩阵。

表 5-4 准则层对目标层判断表

	第一种因素	第二种因素	第三种因素
第一重因素	1	4	2
第二种因素	1/4	1	1/2
第三种因素	1/2	2	1

$$A = \begin{bmatrix} 1 & 1.2214 & 1.2214 & 1 & 2.7183 & 1.4918 \\ 0.8187 & 1 & 1 & 1.8221 & 4.0552 & 2.7183 \\ 0.8187 & 1 & 1 & 1.8221 & 4.0552 & 2.7183 \\ 1 & 0.5488 & 0.5488 & 1 & 1.8221 & 1.4918 \\ 0.3679 & 0.2466 & 0.2466 & 0.5488 & 1 & 1 \\ 0.6703 & 0.3679 & 0.3679 & 0.6703 & 1 & 1 \end{bmatrix}$$

这样构造出来的判断矩阵，其中的数据是依据实际情况自行设计的，衡量判断矩阵的合格标准时它是否具有 consistency。由于客观事物的复杂性和人的认识的多样性，完全一致的判断往往是不现实的，一般只需近似地满足 consistency 即可。在给出判断矩阵后应进行 consistency 检验，其方法如下。

计算一致性指标 CI 。公式为

$$CI = \frac{\lambda_{\max} - n}{n - 1}$$

式中， n 为判断矩阵 A 的阶数， λ_{\max} 为 A 的最大特征值。

查询平均随机一致性指标 RI (见表 5-5 Saaty 的一致性指标)

表 5-5 Saaty 的一致性指标 RI

n	1	2	3	4	5	6	7	8	9	10	11
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51

计算一致性比例 CR 。公式为

$$CR = \frac{CI}{RI}$$

理论说明：当 $CR < 0.1$ 时，一般认为判断矩阵 A 的一致性可以接受；否则就需要重

报名号 # 1384

新进行比较，对 A 加以调整，使之具有满意的一致性。

5.2.3 编程求解

根据上述原理和求解过程，用 yaahp 软件求得 $\lambda_{\max} = 4$ ， $CI = 0$ ， $RI = 0$ ，所以认为满足一致性指标，判断矩阵是可行的。

我们还可以用 MATLAB 编程（代码见附件 2）求解，最终得到最大特征值对应的权向量如下表 5-6 层次分析法所求各影响因素权重所示：

表 5-6 层次分析法所求各影响因素权重

影响因素	第一种影响因素	第二种影响因素	第三种影响因素
权重	0.4688	0.2453	0.2859

由上表可得出三种超期影响因素所占的权重。

上面我们用主观赋权法进行了竞争力评价指标的权重确定，但是这种方法会由于我们的经验的缺乏和个人的偏好，使决策带有主观随意性，因此，下面我们着手建立客观赋权模型——改进的熵权法，重新对竞争力评价指标进行权重确定。

5.3 客观赋权（改进的熵权法）模型的建立

5.3.1 熵权法的原理

设有 m 个待评方案， n 项评价指标，形成原始指标数据矩阵 $X = (x_{ij})_{m \times n}$ ，对于某项指标 X_j ，指标值 x_{ij} 的差距越大，则该指标在综合评价中所起的作用越大；如果某项指标的指标值全部相等，则该指标在综合评价中不起作用。在信息论中，信息熵

$H(x) = -\sum_{i=1}^n p(x_i) \ln p(x_i)$ 是系统无序程度的度量，信息是系统有序程度的度量，二者

绝对值相等，符号相反。某项指标的指标值变异程度越大，信息熵越小，该指标提供的信息量越大，该指标的权重也应越大；反之，某项指标的指标值变异程度越小，信息熵越大，该指标提供的信息量越小，该指标的权重也越小。所以，可以根据各项指标值的变异程度，利用信息熵这个工具，计算出各指标的权重，为多指标综合评价提供依据。

5.3.2 熵权法的步骤

用熵值法^[5]进行综合评价的步骤是：

1) 将各指标同度量化，计算第 j 项指标下第 i 方案指标值的比重 p_{ij}

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}}$$

2) 计算第 j 项指标的熵值 e_j

$$e_j = -k \sum_{i=1}^m p_{ij} \ln p_{ij}$$

其中 $k > 0$ ， \ln 为自然对数， $e_j \geq 0$ 。如果 x_{ij} 对于给定的 j 全部相等，那

报名号 # 1384

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}} = \frac{1}{m}$$

此时 e_j 取极大值, 即

$$e_j = -k \sum_{i=1}^m \frac{1}{m} \ln \frac{1}{m} = k \ln m$$

若设 $k = \frac{1}{\ln m}$, 于是有 $0 \leq e_j \leq 1$

3) 计算第 j 项指标的差异性系数 g_j

对于给定的 j , x_{ij} 的差异性越小, 则 e_j 越大; 当 x_{ij} 全部相等时, $e_j = e_{\max} = 1$,

此时对于方案的比较, 指标 x_j 毫无作用; 当各方案的指标值相差越大时, e_j 越小, 该项指标对于方案比较所起的作用越大。定义差异性系数

$$g_j = 1 - e_j$$

则当 g_j 越大时, 指标越重要。

4) 定义权数

$$a_j = \frac{g_j}{\sum_{j=1}^n g_j}$$

5) 计算综合经济效益系数 v_i

$$v_i = \sum_{j=1}^n a_j p_{ij}$$

v_i 为第 i 个方案的综合评价价值。

5.3.3 改进的熵权法

在用熵值进行企业经济效益评价时, 时常会遇到一些极端值。指标值为负时不能直接计算比重, 也不能取对数, 而为保证数据的完整性这两个值又不能删去, 因此需要对该项指标数据进行变换。

1) 用功效系数法进行变换

取第 j 项指标值中最好值为 $x_j^{(h)}$, 最差值为 $x_j^{(s)}$, 用下列公式进行变换:

$$x_{ij} = \frac{x_{ij} - x_j^{(s)}}{x_j^{(h)} - x_j^{(s)}} \times \alpha + (1 - \alpha)$$

为避免变换后的数据出现零, α 的范围应取 $(0, 1)$ 。

在用公式(1) 进行变换时, 实际上加入了评价者的主观因素, 因为 A 的选取是由评

报名号 # 1384

价者决定的。如果评价者要加大该指标的权重, 可将 A 取大一些, 这时数据范围大, 用熵值法计算的权重就大; 同理, 如果要减小该指标的权重, 可系数法对数据变换后的熵值法不是严格的客观赋权法, 而是一种主、客观结合赋权法。

用功效系数法变换后, 对极端值作一定的处理, 消除了指标值中负值的问题, 然后可按公式(1)~ (5)的步骤进行评价, 但取不同的 A 可能会出现不同的评价结果。

2) 用标准化法进行变换

$$x_{ij}' = (x_{ij} - \bar{x}_j) / s_j$$

其中 \bar{x}_j 第 j 项指标值的均值, s_j 为第 j 项指标值的标准差。

一般地, x_{ij} 的范围在-5 到 5 之间, 为消除负值, 可将坐标平移, 令

$$z_{ij} = 5 + x_{ij}$$

然后用 z_{ij} 代替 x_{ij} 按公式(1)~ (5) 的步骤进行评价, 评价结果见表 1 中熵值法 1 的数据。

(1) 用标准化法变换不需要加入任何主观信息, 是一种完全意义的客观赋权法。

(2) 用功效系数法变换因 α 的选取不同使得评价结果可能是不唯一的, 而用标准化法进行变换评价结果是唯一的。

(3) 标准化法有利于缩小极端值对综合评价的影响。

5.3.4 用 MATLAB 软件编程求解

根据上述改进的熵权法模型的原理, 我们用 MATLAB 编程 (代码见附件 3) 求解, 得到各食品安全影响因素所占权重如表 5-7 熵权法所求各影响因素权重:

表 5-7 熵权法所求各影响因素权重

影响因素	第一种影响因素	第二种影响因素	第三种影响因素
权重	0.4320	0.2387	0.3293

5.4 离差平方组合赋权模型的建立

上面, 我们分别用主观赋权法和客观赋权法进行了权重求解, 这两种方法一种过于依赖决策者的主观随意性, 一种只建立于客观的数学理论知识, 忽略了决策者的主观信息, 因此这两种方法都有其长处和短处。

于是, 我们着手建立一种综合权重确定模型, 通过将主客观赋权法所得的各指标权重通过集成的方法来形成指标的最终权重的思想^[3], 使之既能客观的反应各指标的重要性, 又能反应决策者的主观愿望。

5.4.1 确定最优化模型

经过以上模型的建立与求解, 我们已经得到各安全影响因素的主观权重和客观权重向量分别为 $\bar{w}_1 = (w_{11}, \dots, w_{110})^T$ 和 $\bar{w}_2 = (w_{21}, \dots, w_{210})^T$, 其中 w_{1j}, w_{2j} 分别表示影响因素 P_j 的主观权重和客观权重 ($j=1, \dots, 10$)。记集成的最终权向量为 $\bar{w} = (w_1, \dots, w_{10})^T$, 其中 w_j 为属性 $p_j (j=1, 2, \dots, 10)$ 的集成权重。为了较好地反映主客观的信息, 我们先考虑所确定的集成

权重与已有的主、客观权重之间的离差平方和最小, 即 $\sum_{j=1}^{10} (w_i - w_{1j})^2$ 和 $\sum_{j=1}^{10} (w_i - w_{2j})^2$ 最

报名号 # 1384

小。从而，我们构造如下的最优化模型：

$$\begin{aligned} \min z &= \sum_{j=1}^{10} \left[\alpha (X_i - \bar{X})^2 + \beta (w_j - w_{2j})^2 \right] \\ \text{s.t. } &\sum_{j=1}^{10} w_j = 1, w_j \geq 0 (j=1, \dots, 10) \end{aligned}$$

其中满足 $\alpha, \beta \geq 0$ 均为给定的常数。 $\frac{\alpha}{\alpha + \beta}$ 反应了决策者偏向主观权重的程度，相

应 $\frac{\alpha}{\alpha + \beta}$ 反应决策者偏向客观权重的程度。另一方面，考虑到方案的综合价值越大。其对应的方案就越优，于是权重的确定应使方案的综合评价之和尽可能大。从而构造如下的最优化模型：

$$\begin{aligned} \max \sum_{i=1}^m u_i &= \sum_{i=1}^m \sum_{j=1}^{10} b_{ij} w_j \\ \text{s.t. } &\sum_{j=1}^{10} w_j = 1, w_j \geq 0 (j=1, 2, \dots, 10) \end{aligned}$$

为了综合求解以上两个最优化模型，我们构造如下模型：

$$\begin{aligned} \min \sum_{j=1}^{10} \left[\alpha (X_i - \bar{X})^2 + \beta (w_j - w_{2j})^2 \right] &= \sum_{i=1}^m \sum_{j=1}^{10} b_{ij} w_j \\ \text{s.t. } &\sum_{j=1}^{10} w_j = 1, w_j \geq 0 (j=1, 2, \dots, 10) \end{aligned}$$

5.4.2 构造 Lagrange 函数

我们依据 (3a) 和 (3b) 构造 Lagrange 函数

$$F(w_1, w_2, \dots, w_{10}, \lambda) = \sum_{j=1}^{10} \left[\alpha (X_i - \bar{X})^2 + \beta (w_j - w_{2j})^2 \right] - \sum_{i=1}^m \sum_{j=1}^{10} b_{ij} w_j + 2\lambda \left(\sum_{j=1}^{10} w_j - 1 \right),$$

式中 λ 是 Lagrange 因子。令 $\frac{\partial F}{\partial w_j} = 0 (j=1, \dots, 10)$ 和 $\frac{\partial F}{\partial \lambda} = 0$ 并化简得方程组：

$$\begin{cases} (\alpha + \beta) w_j + \lambda = \alpha w_{1j} + \beta w_{2j} + \frac{1}{2} \sum_{i=1}^m b_{ij}, (j=1, 2, \dots, 10) \\ \sum_{j=1}^{10} w_j = 1 \end{cases}$$

解此方程组，得：

$$w_j = \frac{\alpha}{\alpha + \beta} w_{1j} + \frac{\beta}{\alpha + \beta} w_{2j} + \frac{1}{2(\alpha + \beta)} \left(\sum_{i=1}^m b_{ij} - \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^{10} b_{ij} \right)$$

报名号 # 1384

$$= \frac{1}{\alpha + \beta} \left[\alpha w_{1j} + \beta w_{2j} + \frac{1}{2} \left(\sum_{i=1}^m b_{ij} - \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^{10} b_{ij} \right) \right]$$

记

$$b_j = \sum_{i=1}^m b_{ij} - \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^{10} b_{ij} \quad (j=1, 2, \dots, 10)$$

则

$$w_j = \frac{1}{\alpha + \beta} \left(\alpha w_{1j} + \beta w_{2j} + \frac{1}{2} b_j \right)$$

定理 1 设 $w_{1j} \geq 0, w_{2j} \geq 0 (j=1, 2, \dots, 10)$ ，且分别满足 $\sum_{j=1}^{10} w_{2j} = 1, \sum_{j=1}^{10} w_{2j} = 1$ 则存在正数

$c (= \alpha + \beta) \geq 1$ ，使得由 (7) 式确定的 w_i 满足：

$$w_j \geq 0 (j=1, 2, \dots, 10) \text{ 且 } \sum_{j=1}^{10} w_j = 1$$

5.4.3 编程求解

根据以上模型的建立及求解过程，我们用 MATLAB 编程，求解得到三种影响因素的最终权重如表 5-8 各影响因素综合权重所示：

表 5-8 各影响因素综合权重

影响因素	第一种原因	第二种原因	第三种原因
权重	0.3786	0.3551	0.2663

由上表可看出，对总体超期图书而言，最可能的原因是无意错过还书期限，然后是由于罚款数目较少而故意拖欠，最后是因为科研需要而延长超期。

接下来，我们利用上面求出的各影响因素的权重，对聚类得到的六类超期书籍进行排序，即求各类超期书籍容易超期的程度。

首先我们用灰色关联——模糊综合评价算法来求解。

5.5 灰色关联——模糊综合评价及排序

灰色关联^[6]是灰色系统理论中的一种系统分析方法，其基本思想是通过分析比较数列指标变化对参考数列指标的影响来判别其关联度。

5.5.1 灰色关联法求解步骤

(1) 根据评价目的确定评价指标体系，建立原始数据矩阵（原始数据矩阵同 TOPSIS 法）。

(2) 对指标数据用内插法进行无量纲化在对实测值进行无量纲化中，为体现各纯净水厂间的差异程度，以评价指标中实测数据以及自来水标准值设立区间，在 $[0, 1]$ 内进行内插，低优指标以低值归一为 0，大值归一为 1。高优指标则相反。参考序列取最优值，参考序列 (00000000)

低优指标无量纲化：

$$x_i(k) = \frac{x'_i(k) - x'_{\min}(k)}{x'_{\max}(k) - x'_{\min}(k)}$$

报名号 # 1384

高优指标无量纲化:

$$x_i(k) = \frac{x'_{\min}(k) - x'_i(k)}{x'_{\min}(k) - x'_{\max}(k)}$$

由上式得无量纲化后矩阵。

(3) 逐个计算每个被评价对象指标序列（比较序列）与参考序列对应元素的绝对差值。

即 $|x_0(k) - x_i(k)| (k=1, \dots, m \quad i=1, \dots, n)$ n 为评价对象的个数。由于参考序列为(0 0 0 0 0 0 0)。故差值表结果与表无量纲化矩阵相同。

$$(4) \text{ 确定 } \min_{i=1}^n \min_{k=1}^n |x_0(k) - x_i(k)| \text{ 与 } \max_{i=1}^n \max_{k=1}^n |x_0(k) - x_i(k)|$$

$$\min_{i=1}^n \min_{k=1}^n |x_0(k) - x_i(k)| = 0 \quad \max_{i=1}^n \max_{k=1}^n |x_0(k) - x_i(k)| = 1$$

(5) 确定分辨系数 ρ

$\rho \in [0, 1]$, 其取值不同, 分辨能力不同, 其值愈大, 分辨能力愈强, 但对整个顺序趋势无影响, 在水质质量评价中, 只是求得其关联系数的大小顺序即可。为了简便计算, 文中不对取值作进一步讨论, 取 $\rho = 0.5$

(6) 计算关联系数

$$\zeta(k) = \frac{\min_{i=1}^n \min_{k=1}^n |x_0(k) - x_i(k)| + \rho \cdot \max_{i=1}^n \max_{k=1}^n |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \cdot \max_{i=1}^n \max_{k=1}^n |x_0(k) - x_i(k)|} \quad (k=1, \dots, m)$$

(7) 计算关联序

对各评价对象（比较序列）分别计算其 m 个指标与参考序列对应元素的关联系数的均值, 以反映各评价对象与参考序列的关联关系, 并称其为关联序。

5.5.2 编程求解

根据原理及步骤, 我们用 MATLAB 编程（程序代码见附件 4）求解, 得到各类书籍容易超期的程度, 即权重系数如表 5-9 各类书籍容易超期程度排序所示:

表 5-9 各类书籍容易超期程度排序

书籍类别	权重系数
哲学理学	0.2235
社会政治	0.0699
文化艺术	0.0989
生物化学	0.1686
专业科学技术	0.2818
基础科学	0.1573

5.6 基于 ANFIS（自适应神经模糊推理系统）的书籍易超期程度评价模型

为了有对比性, 使结果更加精确, 我们又建立了基于 ANFIS（自适应神经模糊推理系统^[7]）的援助程度评价模型, 得到新的排序结果。下面介绍这一模型的建立及求解。

报名号 # 1384

1. 自适应神经模糊推理系统

ANFIS 是 Sugeno 型模糊系统。对于 2 个输入的 (x、y) 1 个输出 (z) 的一阶 Sugeno 型模糊系统，有以下两条规则：

$$\text{If } (x \text{ is } A_1) \text{ and } (y \text{ is } B_1) \text{ then } f_1 = p_1x + q_1y + r_1$$

$$\text{If } (x \text{ is } A_2) \text{ and } (y \text{ is } B_2) \text{ then } f_2 = p_2x + q_2y + r_2$$

与该一阶 Sugeno 型模糊系统等效的 ANFIS 结构图如图 5-1ANFIS 结构图：

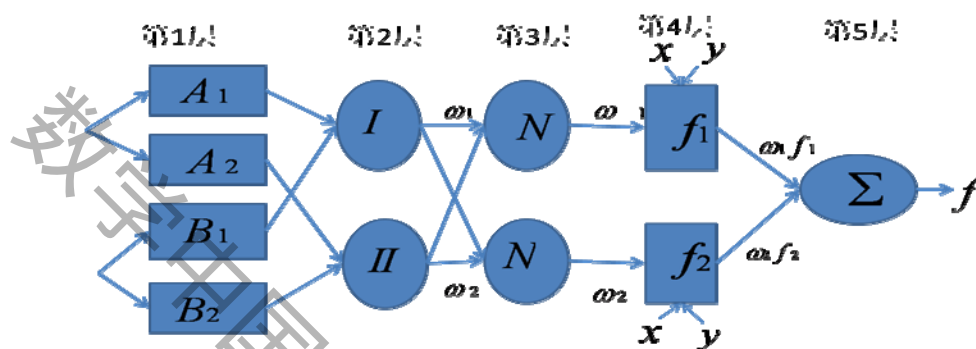


图 5-1ANFIS 结构图

图 1 中，节点间的连线仅表示信号的流向，没有权值与之联系；方形节点表示带有可调参数的节点，圆形节点表示不带有可调参数的节点。ANFI 的结构可分为 5 层。

第1层：将输入变量模糊化，并输出对应模糊集的隶属度：

$$O_i^1 = \mu_{A_i}(x) \quad i = 1, 2$$

$$O_j^1 = \mu_{B_j}(y) \quad j = 1, 2$$

式中：x、y 为节点输入； O_i^1 、 O_j^1 为第 1 层对应节点的输出； A_i 、 B_j 为与该层对

应节点的隶属度函数有关的语言表示（如“大”或“小”）； $\mu_{A_i}(x)$ 、 $\mu_{B_j}(y)$ 分别为输入变量 x、y 的隶属度函数。常用的隶属度函数有三角隶属函数、梯形隶属函数、高斯隶属函数和钟形隶属函数。三角形和梯形隶属函数形式简单，计算效率高，然而由于这两种隶属函数都由直线段构成，在由参数指定的拐角点处不够光滑和简洁的表示，是定义模糊集合最流行的形式。钟形隶属函数：

$$\mu_{A_i}(x) = \left[1 + \left[\frac{x - c_i}{a_i} \right]^{2b_i} \right]^{-1}$$

等式中参数集 $\{a_i, b_i, c_i\}$ 称为前件参数集。

第 2 层：实现条件部分的模糊集运算，输出每条规则的适用度：

$$O_i^2 = \omega_i = \mu_{A_i}(x) \bullet \mu_{B_i}(y)$$

第 3 层：将各条规则的适用度归一化，如第 i 条规则的归一化适用度：

$$O_i^3 = \bar{\omega}_i = \omega_i / (\omega_1 + \omega_2)$$

报名号 # 1384

第 4 层：去模糊运算，计算每条规则的输出：

$$O_i^4 = \bar{\omega}_i f_i = \bar{\omega}_i (p_i x + q_i y + r_i)$$

第 5 层：计算模糊系统的总输出，即所有规则的输出之和：

$$O_i^5 = f = \sum \bar{\omega}_i f_i = \sum \bar{\omega}_i f_i / \sum \bar{\omega}_i$$

前件参数和后件参数都是未知参数，通过混合学习算法训练 ANFIS，可以得到这些参数值，进而达到模糊建模的目的。

2. ANFIS混合学习方法

混合学习算法是将最小二乘法法和梯度下降法相结合的一种学习算法，它可以降低梯度法中搜索空间的维数，提高收敛速度。对于每一次样本训练，混合学习算法都存在前向传递和反向传递两个过程。首先，固定前件参数，将输入量前向传递到 ANFIS，的第 4 层，此时，系统总输出可表示为后件参数的线性组合，即：

$$z = (\bar{\omega}_1 x) p_1 + (\bar{\omega}_1 y) q_1 + \bar{\omega}_1 r_1 + (\bar{\omega}_2 x) p_2 + (\bar{\omega}_2 y) q_2 + \bar{\omega}_2 r_2 = A \bullet X$$

公式中：\$\{p_1, q_1, r_1, p_2, q_2, r_2\}\$ 构成向量 \$X\$；\$A\$、\$X\$、\$z\$ 为矩阵，其维数为 \$P \times 6\$、\$6 \times 1\$、

\$P \times 1\$；\$P\$ 为训练数据的组数。应用最小二乘法求得后件参数 \$X^* = (A^T A)^{-1} A^T z\$

进行误差计算，将获得的误差信号反向传递，用梯度下降法更新前件参数，改变隶属度函数的形状。

3. 基于 MATLAB 软件的 ANFIS

MATLAB 软件将模糊推理技巧与神经网络的自适应建模结合起来，提供了基于 Sugeno 模型的模糊神经网络的实现算法。具体步骤：①产生训练数据和检验数据；②确定输入变量的隶属度函数类型和个数；③由 genfis1 函数使用训练数据产生初始的 FIS 结构；④设定 ANFIS 训练参数；⑤利用 anfis 函数训练 ANFIS，若满足精度要求，则认为该 ANFIS 系统是有效合理的基于模糊推理的系统模型。

4. 基于 ANFIS 援助需求程度评价模型

设有 \$n\$ 类超期书籍集合，其中 \$m\$ 项影响因素指标，则有实测指标矩阵

$$X_{m \times n} = (x_{ij})_{m \times n}, \quad i = 1, 2, \dots, m; \quad j = 1, 2, \dots, n$$

式中 \$x_{ij}\$ 为样本 \$j\$ 指标 \$i\$ 的实测值。

若 \$m\$ 项指标按照 \$c\$ 类级别的标准进行评价，对于各类超期书籍的数值区间型类别取其标准值，则有标准指标矩阵

$$Y_{m \times c} = (y_{ih})_{m \times c}$$

式中：\$h\$ 为标准指标矩阵的类别号，\$h = 1, 2, \dots, c\$；\$c\$ 为标准类别数；\$y_{ih}\$ 为 \$h\$ 级标准类别指标 \$i\$ 的标准值。

以 \$m\$ 项评价指标为输入，各类书籍容易超期程度为输出，标准指标矩阵和已分的国家类别矩阵为学习样本，采用 MATLAB 软件对模糊神经识别网络进行训练；将待识别的实测指标矩阵代入训练好的 ANFIS 评价模型，根据系统输出值确定各类书籍容易超期程度。

报名号 # 1384

各类书籍容易超期程度综合评价流程见图 5-2 基于 ANFIS 的综合评价流程。

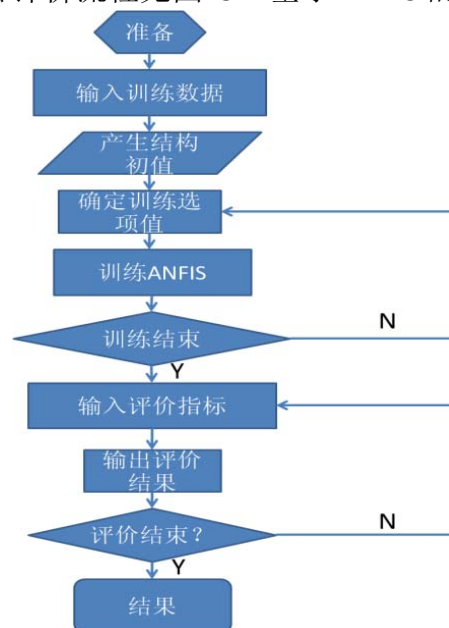


图 5-2 基于 ANFIS 的综合评价流程

现采用 ANFIS 方法对个样本点的容易超期程度进行综合评价，取各影响因素权重评价指标作为网络的输入，共 18 个输入节点；容易超期程度级别作为网络的输出，共 1 个输出节点。按隶属度不同，将输入节点划分为 $[1,0.2]$ 之间等分的区间作为输出区分。

按照图 2 流程对处理后的数据样本进行学习，经过 20 次训练即可得出理想的样本识别效果。将数据带入训练好的 ANFIS 结构，得出相应的书籍容易超期程度需求结果。如表 5-10 所示：

表 5-10 书籍容易超期程度需求结果

书籍类别	满意度值
哲学理学	0.887
社会政治	0.798
文化艺术	0.799
生物化学	0.835
专业科学技术	0.919
基础科学	0.887

6.3.3.两种综合评价模型的比较及结合

灰色关联——模糊综合评价模型具有计算简便、应用灵活、结果较合理等特点。但是，利用欧式距离进行计算存在一个容易忽视的不足之处，即与理想解欧式距离小的方案可能与负理想解的欧式距离也小，这也是模型的一个缺点。

基于 ANFIS 的各类书籍容易超期程度评价模型能够比较全面了各评价因素的影响，其计算速度快、结果更准确合理，克服了灰色关联——模糊综合评价体系模型的缺点。

根据两者优缺点的总结，我们取二者结果的加权平均值作为最终结果。对基于 ANFIS 的综合评价方法结果进行归一化，然后与灰色关联综合评价模型的结果相加取平均值得到最终各类书籍容易超期程度如下表 5-11 各类书籍容易超期程度加权平均结果所示：

报名号 # 1384

表 5-11 各类书籍容易超期程度加权平均结果

书籍类别	权重
哲学理学	0.2125
社会政治	0.0812
文化艺术	0.0976
生物化学	0.1425
专业科学技术	0.2856
基础科学	0.1806

下面求引起每种超期书籍中每种影响因素所占的比重。

前面求解时曾将超期天数分成三阶段，并将每一阶段所引起的因素所占比例进行量化（表 5-2），根据该量化表，我们再统计出每一类书籍中每一超期阶段的书籍数目（所占比例），与量化表中的数据进行加权求和，然后归一化，可以得到使每一类书籍超期的原因的权重。

模型建立如下：

$$a'_{ik} = \sum_{j=1}^3 b_{ij} c_{jk} (i=1,2,\dots,6; k=1,2,3)$$

$$a_{ik} = \frac{a'_{ik}}{\sum_{i=1}^6 \sum_{k=1}^3 a'_{ik}} (i=1,2,\dots,6; k=1,2,3)$$

其中， b_{ij} 表示第 i 类书籍中在 j 阶段超期的书籍所占比例； c_{jk} 表示第 j 阶段第 k 个因素的量化比例；而得到的 a'_{ik} 则代表第 i 类书籍中第 k 个影响因素所占比例； a_{ik} 是对 a'_{ik} 归一化后的结果，表示第 i 类书籍中第 k 个影响因素所占权重，且可知：

$$\sum_{i=1}^6 \sum_{k=1}^3 a_{ik} = 1$$

根据上述模型建立及统计计算我们得到每类书籍每种因素引起超期的程度权重，如表 5-12 各类书籍各种超期原因所占权重所示。

表 5-12 各类书籍各种超期原因所占权重

	总权重	第一种原因权重	第二种原因权重	第三种原因权重
哲学理学	0.2125	0.0897	0.0981	0.0247
社会政治	0.0812	0.0379	0.0132	0.0301
文化艺术	0.0976	0.0542	0.0322	0.0112
生物化学	0.1425	0.0520	0.0498	0.0407
专业科学技术	0.2856	0.0672	0.1289	0.0895
基础科学	0.1806	0.0776	0.0329	0.0701
总权重	1.0000	0.3876	0.3551	0.2663

从上表可以很直观的看出，在六类所分书籍中，哲学理学书籍以及专业科学技术书籍两类书最容易超期，而二者超期原因有个共同的特点，即“由于科学研究而不能归还”这一原因而导致超期所占比重较大；其次容易超期的是基础科学书籍与生物化学书籍，而导致其超期的主要原因是“无意错过还书期限”。而就每种原因所占总权重分析，“无意错过还书期限”引起借阅超期比例最大。

综上分析可知，大多数借阅超期是由于无意错过还书期限或者由于科学研究而不能

报名号 # 1384

归还书籍，而不是故意不还。就此我们对图书馆惩罚管理提出以下建议：

1、罚款应该设定一个金额上限。因为超期时间越长，其因为“无意错过还书期限”而超期的可能性越大，既然非故意超期，而图书馆罚款是为了治理超期现象而不是为了赚钱，则若罚款太多不合情理，也会影响图书馆形象。

2、短期内提高每天罚款金额。现在大多数图书馆都是罚款每超期一天 0.1 元，可以适当提高价钱，比如每天 0.2 元。这样的话可以催促那些因为觉得罚款金额少而故意超期的同学尽早还书，也给经常忘记还书的同学提个醒，记得还书。

3、图书馆应加强教育。要充分利用读者入学教育的机会扎实开展读者入馆教育；而且做好日常宣传教育，可以在图书馆贴日常标语，宣传广泛。

4、图书馆加强自身管理。从所给数据中可以看出，有些书籍已经超期好几年甚至十几年，这些学生已经毕业，明显书籍已经还不了了，但是图书馆只是将这些信息记录下来，并不实施措施解决。有些书籍在超期半年或者一年时就应该根据学生证件号通知其所在学院、班级及个人，让其作出解释，并酌情令其还书或者采取其他手段。

6 问题二模型的建立与求解

6.1 博弈模型

6.1.1 博弈论方法与纳什均衡的概述

高校图书馆结束超期应对策略的确定受诸多因素的影响，是多种要素相互作用的结果。只看到其中一个方面很难得出令人信服的结论。在这里，我们尝试用一种新的视角来分析、确定高校图书馆的智能服务问题。这就是高校图书馆智能服务确定的博弈分析[8]方法。

在现实生活中，人们的许多决策之间是相互依赖的。你的最优决策依赖于别人的决策，别人的最优决策依赖于你的决策；你的决策还依赖于别人认为你怎么决策，你怎么认为别人认为你怎么决策，如此等等。研究这种互动环境下的理性选择的理论叫博弈论(theory of rational choices under interactions)。

在博弈理论中包含一个重要的概念和思想，这就是“纳什均衡”。纳什均衡(Nash equilibrium)是由美国数学家纳什首先提出并证明的。所谓纳什均衡，指的是所有参与人最优选择的一种组合。在这种组合下给定其他人的选择，没有任何人有积极性做出新的选择。纳什均衡也可以从另一个角度来理解：假设博弈中的所有参与人事先达成一项协议，等于说这个协议构成一个纳什均衡：在给定遵守协议的情况下，没有人有积极性偏离协议规定的自己的行为规则。这就是纳什均衡的理论思想。

6.1.2 高校收费的纳什均衡理论分析

用纳什均衡理论来分析高校图书馆的智能服务问题，我们就会发现，在一定条件下，高校图书馆的智能服务是图书馆限制超期与借书者超期之间的一种纳什均衡。这时，图书馆不可能大幅度罚款的数额，而借书者也不能做到全都按时归还图书。图书馆与借阅人员自觉履行满足纳什均衡要求的“收费协议”。在这种情况下，无论对图书馆还是借书者个人都是有利的。反之，如果图书馆大幅度增加罚款的数额，可能造成图书交流收阻碍，而且在一定程度上会影响图书馆声誉，更重要的是超过部分借阅者的金钱支撑能力。这样，一方面借书的人会减少，不能满足社会对高素质人才的需求，阻碍社会经济的发展；另一方面也会使图书馆资源大量闲置浪费。如果图书馆大幅度增加罚款数额，可能会出现借书人数锐减的现象。这对图书馆而言是得不偿失的。从借书者角度来看，因为借阅书刊必定会带来很多个人收益，如果减少买书支出，丧失了借书机会是不划算

报名号 # 1384

的。同时，我们也应看到，这种均衡只是一种动态的平衡，随着人们收入水平的提高，图书馆规模的扩大，这种平衡必然会被打破，并在更高的水平上达到新的平衡。当然，在这个过程中，各方都会努力维护自己的利益，都希望自己的愿望能实现。因而，其间充满纷争和妥协。图书馆罚款的博弈过程具体体现为以下形式：

1) 图书馆罚款的承诺博弈

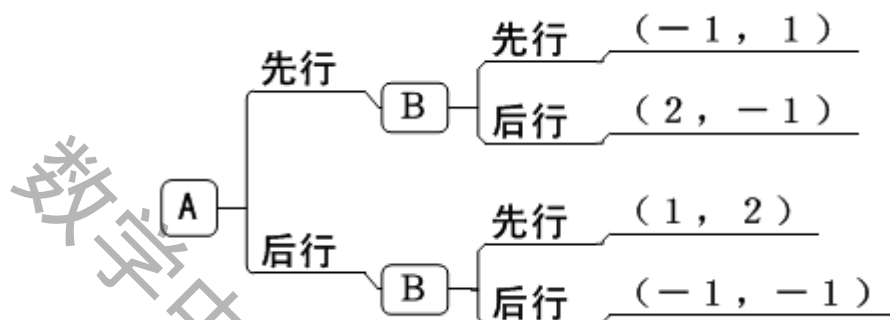


图 6-1 承诺博弈

借书人员处于有利的位置，因而借书人员的按时还书可以看成是第一个承诺。而图书馆处于劣势的位置。

因此，政府提高罚款金额、借书者增加超期是最有可能出现的情况。因而，图书馆收费的出现及提高是一种不可避免的现象。

2) 图书馆罚款形成博弈

用博弈理论来分析图书馆智能服务问题，我们就会发现，图书馆智能服务的最终形成是图书馆和借阅人员之间相互博弈的结果。从图书馆的角度来看，图书馆希望增加罚款使得借阅人员能按时归还所借的图书。从借阅人员的角度来看，他们不愿看到图书馆罚款上涨，图书馆罚款越低，对他们越有利。

通过以上分析，我们不难发现，现有的博弈关系使借书者处于不利的地位。因而，要改变这种状况须从改变图书馆与借阅人员两者之间的博弈关系入手。我们认为具体的做法有：

(1) 在图书馆罚款决策博弈过程中，要尽可能听取公众和社会舆论的呼声，倾听借书者的意见，多做社会调查。

(2) 在借书人员承诺博弈过程中，应将图书馆的要求作为纳什均衡的第一承诺。为了保证图书借阅的顺利发展，应将图书馆的罚款承诺作为第二承诺，应充分发挥图书馆服务策略作为高等教育投资主渠道的作用。当罚款不足以弥补成本时，由增加罚款以补足，即由借阅人员→图书馆→借阅人员。

(3) 在图书馆服务形成博弈过程中，应要求图书馆确定科学的借书策略，注意节约资源，不断图书的利用效率。

由上述分析，在图书馆的智能服务与借阅人员的还书问题之间存在一平衡点，在这点处，可以达到图书馆罚款数目最少，而且借书人员又能够接受因超期所造成的经济压力。

题目中所给的图书馆罚款策略，根据第一问的结果，虽然此策略在一定程度上解决了超期现象。它一方面增加了借书人员的还书意识，使得更多的人员能按时归还所借图书，同时为图书馆损失的图书提供了资金来源；另一方面，由于罚款对于不同的人采用了相同罚款数目而且没有充分考虑到各种图书超期比重和借书人员超期原因，使得罚款

报名号 # 1384

出现一定程度的盲目性，不能完全起到智能服务的功能。

以下我们对上述方案中出现的问题进行相应的改进，充分考虑到不同类别图书对超期天数的影响，综合借阅人员的经济承受能力与图书超期现象因素利用回归分析方法建立相关模型。

6.2 回归分析模型的建立

6.2.1 多元线性回归原理

回归分析中最简单的形式是 $y = \beta_0 + \beta_1 x$ ， x, y 均为标量， β_0, β_1 为回归系数，称一元线性回归。它的一个自然推广是 x 为多元变量，形如

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m \quad (6-1)$$

$m \geq 2$ ，或者更一般地

$$y = \beta_0 + \beta_1 f_1(x) + \cdots + \beta_m f_m(x) \quad (6-2)$$

其中 $x = (x_1, \cdots, x_m)$ ， $f_j (j=1, \cdots, m)$ 是已知函数。这里 y 对回归系数 $\beta = (\beta_0, \beta_1, \cdots, \beta_m)$ 是线性的，称为多元线性回归。不难看出，对自变量 x 作变量代换，就可将 (6-2) 化为 (6-1) 的形式，所以下面以 (1) 为多元线性回归的标准型。

6.2.2 模型的建立

在回归分析中自变量 $x = (x_1, x_2, \cdots, x_m)$ 是影响因变量 y 的主要因素，是人们能控制或能观察的，而 y 还受到随机因素的干扰，可以合理地假设这种干扰服从零均值的正态分布，于是模型记作

$$\begin{cases} y = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m + \varepsilon \\ \varepsilon \sim N(0, \sigma^2) \end{cases} \quad (6-3)$$

其中 σ 未知。现得到 n 个独立观测数据 $(y_i, x_{i1}, \cdots, x_{im})$ ， $i=1, \cdots, n, n > m$ ，由 (3) 得

$$\begin{cases} y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_m x_{im} + \varepsilon_i \\ \varepsilon_i \sim N(0, \sigma^2), \quad i=1, \cdots, n \end{cases} \quad (6-4)$$

记

$$X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1m} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & \cdots & x_{nm} \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad (6-5)$$

$$\varepsilon = [\varepsilon_1 \quad \cdots \quad \varepsilon_n]^T, \quad \beta = [\beta_0 \quad \beta_1 \quad \cdots \quad \beta_m]^T$$

(6-4) 表为

$$\begin{cases} Y = X\beta + \varepsilon \\ \varepsilon \sim N(0, \sigma^2) \end{cases} \quad (6-6)$$

报名号 # 1384

6.2.3 参数估计

用最小二乘法估计模型 (6-3) 中的参数 β 。

由 (4) 式这组数据的误差平方和为

$$Q(\beta) = \sum_{i=1}^n \varepsilon_i^2 = (Y - X\beta)^T (Y - X\beta) \quad (6-7)$$

求 β 使 $Q(\beta)$ 最小，得到 β 的最小二乘估计，记作 $\hat{\beta}$ ，可以推出

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (6-8)$$

将 $\hat{\beta}$ 代回原模型得到 y 的估计值

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_m x_m \quad (6-9)$$

而这组数据的拟合值为 $\hat{Y} = X\hat{\beta}$ ，拟合误差 $e = Y - \hat{Y}$ 称为残差，可作为随机误差 ε 的估计，而

$$Q = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6-10)$$

为残差平方和（或剩余平方和），即 $Q(\hat{\beta})$ 。

6.2.4 统计分析

不加证明地给出以下结果：

(i) $\hat{\beta}$ 是 β 的线性无偏最小方差估计。指的是 $\hat{\beta}$ 是 Y 的线性函数； $\hat{\beta}$ 的期望等于 β ；在 β 的线性无偏估计中， $\hat{\beta}$ 的方差最小。

(ii) $\hat{\beta}$ 服从正态分布

$$\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1}) \quad (6-11)$$

(iii) 对残差平方和 Q ， $EQ = (n - m - 1)\sigma^2$ ，且

$$\frac{Q}{\sigma^2} \sim \chi^2(n - m - 1) \quad (6-12)$$

由此得到 σ^2 的无偏估计

$$s^2 = \frac{Q}{n - m - 1} = \hat{\sigma}^2 \quad (6-13)$$

s^2 是剩余方差（残差的方差）， s 称为剩余标准差。

报名号 # 1384

(iv) 对总平方和 $S = \sum_{i=1}^n (y_i - \bar{y})^2$ 进行分解, 有

$$S = Q + U, U = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (6-14)$$

其中 Q 是由 (6-10) 定义的残差平方和, 反映随机误差对 y 的影响, U 称为回归平方和, 反映自变量对 y 的影响。

6.2.5 回归模型的假设检验

因变量 y 与自变量 x_1, \dots, x_m 之间是否存在如模型 (1) 所示的线性关系是需要检验的, 显然, 如果所有的 $|\hat{\beta}_j|$ ($j=1, \dots, m$) 都很小, y 与 x_1, \dots, x_m 的线性关系就不明显, 所以可令原假设为

$$H_0: \beta_j = 0 (j=1, \dots, m)$$

当 H_0 成立时由分解式 (6-14) 定义的 U, Q 满足

$$F = \frac{U/m}{Q/(n-m-1)} \sim F(m, n-m-1) \quad (6-15)$$

在显著性水平 α 下有 $1-\alpha$ 分位数 $F_{1-\alpha}(m, n-m-1)$, 若 $F < F_{1-\alpha}(m, n-m-1)$, 接受 H_0 ; 否则, 拒绝。

注意 拒绝 H_0 只说明 y 与 x_1, \dots, x_m 的线性关系不明显, 可能存在非线性关系, 如平方关系。

还有一些衡量 y 与 x_1, \dots, x_m 相关程度的指标, 如用回归平方和在总平方和中的比值定义

$$R^2 = \frac{U}{S} \quad (6-16)$$

$R \in [0, 1]$ 称为相关系数, R 越大, y 与 x_1, \dots, x_m 相关关系越密切, 通常, R 大于 0.8 (或 0.9) 才认为相关关系成立。

6.2.6 回归系数的假设检验和区间估计

当上面的 H_0 被拒绝时, β_j 不全为零, 但是不排除其中若干个等于零。所以应进一步作如下 m 个检验 ($j=1, \dots, m$):

$$H_0^{(j)}: \beta_j = 0$$

报名号 # 1384

由(6-11)式, $\hat{\beta}_j \sim N(\beta_j, \sigma^2 c_{jj})$, c_{jj} 是 $(X^T X)^{-1}$ 对角线上的元素, 用 s^2 代替 σ^2 ,

由(6-11) - (6-13)式, 当 $H_0^{(j)}$ 成立时

$$t_j = \frac{\hat{\beta}_j / \sqrt{c_{jj}}}{\sqrt{Q/(n-m-1)}} \sim t(n-m-1) \quad (6-17)$$

对给定的 α , 若 $|t_j| < t_{1-\frac{\alpha}{2}}(n-m-1)$, 接受 $H_0^{(j)}$; 否则, 拒绝。

(6-17)式也可用于对 β_j 作区间估计 ($j=0,1,\dots,m$), 在置信水平 $1-\alpha$ 下, β_j 的置信区间为

$$[\hat{\beta}_j - t_{1-\frac{\alpha}{2}}(n-m-1)s\sqrt{c_{jj}}, \hat{\beta}_j + t_{1-\frac{\alpha}{2}}(n-m-1)s\sqrt{c_{jj}}] \quad (6-18)$$

其中 $s = \sqrt{\frac{Q}{n-m-1}}$ 。

6.2.7 对本文的具体模型的建立

在此问题中我们综合考虑图书种类、借书者造成超期原因以及超期天数与图书馆进行罚款处罚之间的关系, 根据第一问所得到的不同种类图书超期天数不同而得到的超期难易权重, 我们建立回归模型

$$Y = C_1 u_1 + C_2 u_2 + C_3 u_3$$

其中 Y 表示罚款金额, u_1 表示图书种类, u_2 表示借书者造成超期原因, u_3 表示超期天数, C_1 、 C_2 、 C_3 分别为要确定的系数。

6.2.8 模型的应用

利用该模型以及上述原理可以通过搜集更多高校的数据, 对上述具体模型的参数进行估计, 进而得到确定的经验公式, 及根据图书馆管理方期望超期情况即可确定罚款金额。

7 模型的评价与改进

7.1 模型的评价

7.1.1 优点

- 1 问题一中综合考虑到了不同种类书会出现不同的超期现象, 而且不同的借阅者对超期有不同的原因, 并对此作了权重处理. 因此本文在建立模型时考虑因素较为全面, 使模型及体系具有很强的使用性与代表性
- 2 模型中加入合理的主观判断, 使结果更加符合实际情况.
- 3 模型二本文运用博弈论中的纳什平衡建立评价体系, 对于原有学费制定体系进行评价分析并提出新体系的准则, 文章具有很强的逻辑性与合理性;
- 4 本文将原模糊问题具体划分为若干问题, 文章的思路清晰且具有很好的连贯性;

报名号 # 1384

7.1.2 缺点

- 1 对于模型的建立没有完全考虑到所有的因素；
- 2 对部分因素的影响效果的叙述与讨论有些抽象，只做了一些定性的分析；

7.2 模型的改进

- 1 对于问题二中确定的模型，确定的单一的罚款方案，仍然会导致一些不公平现象，例如对于由于进行科研而导致超期过长的借书人员，或者由于忘记归还图书的借书者可能罚款数会因此偏高，而对于故意不归还图书的人员则偏低。因此我们可以建立因人而异的罚款方式，将 y 看做一列向量，对不同种类的书分别制定相应罚款。
- 2 仍然对于问题二中确定的模型，由于不同时间段借还书的情况会发生较大变化，因此，我们可以建立一动态的模型，一段时间变化一次其中一种或者几种类型书的罚款策略。
- 3 模型的建立中应该增加借阅人员的经济承受能力，即罚款上限。
- 4 可以建立单目标最优化问题，其中最终目标是超期天数最少，约束条件由罚款上限、罚款金额最小值、图书种类等决定。

8 参考文献

- [1] 胡福文，高校图书馆读者借阅超期有关问题的思考，学理论，158-159，2009
- [2] 姜启源，谢金星. 数学模型（第三版）. 北京：高等教育出版社. 2003. 8.
- [3] 龙脉工作室，刘会灯，朱 飞，MATLAB 编程基础与典型应用，北京：北京邮电出版社，2008. 7。
- [4] 陈建中，徐玖平，群决策理论与方法及实现，北京：清华大学出版社，2009. 4
- [5] 韩中庚，数学建模方法及其应用，北京：高等教育出版社，2005 年 6 月
- [6] 王庚，王敏生，现代数学建模方法，北京：科学出版社，2008
- [7] 鲁斌，何华灿，自适应神经模糊推理系统建模研究，计算机科学，第十期，2003
- [8] 范丽娟，博弈论在图书馆管理中的应用，黑河学刊，第六卷，2003
- [9] 盛骤，谢式千，潘承毅，概率论与数理统计，北京：高等教育出版社，2008. 6

报名号 # 1384

9 附录

附件一：模糊聚类分析代码

```
X=input('please input X:');
[N,M]=size(X);
%
P=zeros(M,1);
for j=1:M
    for i=1:N
        P(j)=P(j)+X(i,j);
    end
    P(j)=P(j)/N;
end
P1=zeros(M,1);
for j=1:M
    for i=1:N
        P1(j)=P1(j)+(X(i,j)-P(j))*(X(i,j)-P(j));
    end
    P1(j)=P1(j)/N;
    P1(j)=sqrt(P1(j));
end
for i=1:N
    for j=1:M
        X1(i,j)=(X(i,j)-P(j))/P1(j);
    end
end
for j=1:M
    A(j)=min(X1(:,j));
    B(j)=max(X1(:,j));
end
for i=1:N
    for j=1:M
        X2(i,j)=(X1(i,j)-A(j))/(B(j)-A(j));
    end
end
X2
R1=zeros(N,N);
for i=1:N
    for k=1:N
        if i==k ;
        else

            for j=1:M
```

报名号 # 1384

```
        R1(i,k)= R1(i,k)+X2(i,j)*X2(k,j);
    end
end
end
m=max(R1);
m=max(m);
R=zeros(N,N);
for i=1:N
    for j=1:N
        if i==j
            R(i,j)=1;
        else
            R(i,j)=R1(i,j)/m;
        end
    end
end
end
R
flag=0;
C=R;
T=zeros(N,N);
while flag==0
    for i=1:N
        for j=1:N
            for k=1:N
                T(i,j)=max(min(C(i,k),C(k,j)),T(i,j));
            end
        end
    end
    if T==C
        flag=1;
    else
        C=T;
    end
end
end
T
q=0.25;
for i=1:N
    for j=1:N
        if T(i,j)>=q
            T1(i,j)=1;
        else
            T1(i,j)=0;
        end
    end
end
```

报名号 # 1384

```

end
end
T1
F=zeros(N,N);
k=1;
for i=1:N
    e=1;
    F(k,e)=i;
    for j=1:N
        if i==j;
        else
            if isequal(T1(i,:),T1(j,:))
                e=e+1;
                F(k,e)=j;
            end
        end
    end
    k=k+1;
end
F

```

附件二

```

function [T,ind]=huiseguanlianfa(A,W)
[ma,na]=size(A);
Dopt=zeros(1,na);
d=zeros(ma,na);
E=zeros(ma,na);
L=zeros(ma,1);
p=0.5;
MAX=max(A);
MIN=min(A);
for i=1:ma
    for j=1:na
        A(i,j)=(A(i,j)-MIN(j))/(MAX(j)-MIN(j));
    end
end
Dopt=min(A);
for i=1:ma
    for j=1:na
        d(i,j)=abs(A(i,j)-Dopt(j));
    end
end
dmax=max(max(d));
dmin=min(min(d));

```

报名号 # 1384

```

for i=1:ma
    for j=1:na
        E(i,j)=(dmin+p*dmax)/(d(i,j)+p*dmax);
    end
end
for i=1:ma
    for j=1:na
        L(i)=L(i)+W(j)*E(i,j);
    end
end
[T,ind]=sort(L);

```

附件三 熵权法代码

```

function b=shangquanfa(A)
[ma,na]=size(A);
B=zeros(1,na);
P=zeros(1,na);
G=zeros(1,na);
g=0;
a=0.9;
MAX=max(A);
MIN=min(A);
for i=1:ma
    for j=1:na
        A(i,j)=(A(i,j)-MIN(j))/(MAX(j)-MIN(j))*a+(1-a);
    end
end
for j=1:na
    for i=1:ma
        B(j)=B(j)+A(i,j);
    end
end
for i=1:ma
    for j=1:na
        A(i,j)=A(i,j)/B(j);
    end
end
for j=1:na
    for i=1:ma
        P(j)=P(j)+A(i,j)*log(A(i,j));
    end
end
k=1/log(na);
for j=1:na
    G(j)=1-(-k)*P(j);

```


报名号 # 1384

```
g=g+G(j);  
end  
for j=1:na  
    b(j)=G(j)/g;  
end
```

附件 4：灰色关联分析代码

```
function [T,ind]=huiseguanlianfa(A,W)  
[ma,na]=size(A);  
Dopt=zeros(1,na);  
d=zeros(ma,na);  
E=zeros(ma,na);  
L=zeros(ma,1);  
p=0.5;  
MAX=max(A);  
MIN=min(A);  
for i=1:ma  
    for j=1:na  
        A(i,j)=(A(i,j)-MIN(j))/(MAX(j)-MIN(j));  
    end  
end  
Dopt=min(A);  
for i=1:ma  
    for j=1:na  
        d(i,j)=abs(A(i,j)-Dopt(j));  
    end  
end  
dmax=max(max(d));  
dmin=min(min(d));  
for i=1:ma  
    for j=1:na  
        E(i,j)=(dmin+p*dmax)/(d(i,j)+p*dmax);  
    end
```