

## 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](http://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：1812

参赛队员（签名）：

队员 1：林伟照

队员 2：洪希曦

队员 3：方纾婕

参赛队教练员（签名）：

参赛队伍组别：大学组

## 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）：

1812

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）：

# 2010年第三届“ScienceWord杯”数学中国

## 数学建模网络挑战赛

题 目 高校图书馆的智能服务

关 键 词 数据挖掘 矩阵模型 样条拟合 合理度 图书馆 超期规则

### 摘 要:

本文针对高校图书馆图书超期的现象,探讨借书制度的问题。在对相关数据进行收集并用数据挖掘的知识对数据进行分析的基础上,首先通过建立模型 I 探讨超期原因的影响因素,然后建立模型 II 分析超期原因对罚款、期限的影响,接着建立一个模型 III 制定理想规则,并以此为参照规则,建立评价模型 IV 评价原本借书规则的方法的合理性,最后再制定出科学合理的借书规则。

**模型 I --超期数的分布规律矩阵模型** 用“超期数”表示同一种类别的书在某一个特定时间段内超期的总本数。为了分析超期数的规律,运用数据挖掘的统计方法建立超期数的分布规律矩阵模型。

**模型 II --超期原因矩阵模型** 根据各种图书的种类和超期天数对超期原因的影响,运用了加权的思想,利用经验数据分别对不同类别和超期天数的三种原因进行加权赋值,从而得到一个超期原因矩阵模型。其中,在超期天数的加权中,我们采用了先给定某些离散的超期天数对应的原因权值进行样条拟合方法拟合出三条随超期天数变化的原因权值函数,引入不同类别和不同超期天数位置上原因确定算法,最终得到原因矩阵模型。

**模型 III --规则制定模型** 通过对模型 I 和 II 的联合利用,分析了规则与超期原因和超期天数的关系,给出了规则制定的法则之后,建立了规则制定模型。并利用该模型得出理想规则。

**模型 IV --规则评价模型** 在此模型中设定了期限指标和罚款指标组成的来评价规则的合理度。利用该评价模型,计算得出本题目所给规则的合理度为 0.32,超过给定的合理区间;而本文的设计规则的合理度为 0.19,合理度较高。

**本文的亮点在于:** 创新性的建立了不同于一般表达式的超期数分布规律矩阵模型和超期原因矩阵;采用了样条拟合方法拟合出以超期天数为自变量的权值函数,能更加准确的确定权值;制定了规则制定算法和理想规则;建立了规则评价指标和罚款评价指标,在数据极具缺少的情况下,解决了规则的评价问题。

参赛队号 1812

参赛密码 \_\_\_\_\_  
(由组委会填写)

所选题目 C

## 英文摘要

To deal with the problem of overdue books in the university libraries, the present study is designed to discuss the rationality of the rules of book borrowing in the basis of relative data collection and the use of data mining.

There are four models set to analyze and solve the problem. Firstly, model I is to discuss the influence factors of extended period of book borrowing. Secondly, model II is to find out the impacts of extended period on the fine and deadline. Thirdly, model III, a method of evaluating the quantity of money and the model II, is set up to evaluate the rationality of the original fine rules. Finally, a scientific and rational borrowing and penalty rules is developed to better the library system.

### **Model I : the Distribution Rule Matrix Model of overdue books**

“overdue quantity” is used to define the total quantity of a certain kind of overdue books in a certain period of time. To analyze the rule of overdue quantity, the statistic method of data mining is used to set up the Distribution Rule Matrix model of overdue books.

### **Model II : the Overdue Reason Matrix Model**

Based on the influence of book kinds and overdue days on overdue reasons, weighed method is used to evaluate the data so that an overdue reason matrix can be developed.

### **Model III: the Rule Formulating Model**

With the combination of Model I and II, this model analyze the correlation of rule and the overdue reason and days, and provide a series of new rules to establish the Rule Formulating Model, so that come up with an ideal rule.

### **Model IV: the Evaluating Rationality of Rule Model**

To justify the flaw of the original rule and the rationality of the newly-build rule by Model III, two terms are set up: The deadline term and the penalty term. Based on the evaluating model, the degree of rationality is 0.32, which is excessive to the given rational interval, while the degree of rationality of the rule made by the model is 0.19, which is in accordance with the rationality

**The light spot of this paper** is the creativity that we use overdue reason matrix to illustrate the distribution rule. The Spline fitting curving method can fit the weighted function of the overdue days, which can set up a more accurate weighed value. And a evaluation of fine and rule can help solve the rule evaluating problem.

数学中国提供 (www.madio.net)

## 1 问题重述

高校图书馆存在着许多不良的现象，其中，所借的图书超期是每一个图书馆都面临的一个最基本问题。超期是读者借阅图书资料时超过规定借阅期限后不归还的一种普遍现象。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。

(1)对于借阅图书超期，原因众多。考虑到高校的特点，大体上可以把超期原因分为三类：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，所以故意拖延。请通过给出的数据或自行查找相关数据，建立合理的模型，分析哪些书容易发生超期现象，以及当发生超期时，分析其可能的原因。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 1 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 (1) 中的模型评价这种借书规则的优劣。请设计一种更加合理的借阅规则。

## 2 问题的分析

### 2.1 分析造成读者借阅超期的原因分析

#### 2.1.1 造成读者借阅超期的原因

##### (1) 故意超期

所谓故意超期，就是读者明知图书借阅到期并且能够及时办理还书手续，而不予办理相关手续的行为，即明知故犯。故意超期的情形主要有：

①图书借阅到期且不能续借时，读者因某种急需，如考研、考级等，还需继续借阅，于是通过违章继续占用到期图书。

②有的读者认为，每天 0.1 元的超期罚款算不了什么，因而未把按时主动还书当回事，导致借阅超期；

③有的图书很紧俏，市场上不易买到，有的读者出于种种原因，为达到长期独占目的，在故意超期后谎报丢失，即使交超期罚款和赔书款也自认为值得。

##### (2) 过失超期

与故意超期不同，过失超期是读者不知道图书借阅到期，因而未能按时还书所造成的，即主观上不希望超期。过失超期的情形主要有：

①读者实际不知道或不记得借阅到期的准确日期而导致超期；

②疏忽大意的过失，即读者因忙于学习或工作等种种原因完全忘记借阅到期，甚至忘记自己曾借过图书而导致超期。

##### (3) 其他原因

其他原因是指读者既非故意，又非过失而导致超期的情形，主要指出现意外事件（如所借的图书被盗、遗失等）后读者未及时到图书馆进行处理而导致超期，造成赔书又超期的双重损失<sup>[1]</sup>。

#### 2.1.2 各原因造成借阅超期所占的比例

在福建工程学院的一份问卷调查<sup>[2]</sup>中，可以总结出导致图书超期的原因主要包括：

- (1) 没有留意或不能明确还书期限，约 50%；
- (2) 读者的读书习惯和学习兴趣，约 15%；
- (3) 图书内容和教学安排，约 10%；
- (4) 图书的借阅周期制度，约 15%；
- (5) 图书的私自互借，约 10%。

结合借阅超期的不同原因的人以及其比例，本文在分析超期原因时将分为以下三类：

- 1、无意错过还书期限；
- 2、非恶意的故意超期（由于教学科研使用需要）；
- 3、恶意的故意超期（由于罚款数目较少而故意拖延）。

## 2.2 对问题所需要的数据的处理与分析

### 2.2.1 超期的天数统计

根据题目所提供的数据，运用数据挖掘中的随机采样，可做出以下的统计，将超期的天数分成 5 段，分别为：40 及 40 天以下、40-100 天、100-300 天、300-1000 天、1000 天以上。并从中根据等距抽样的方式抽取出两份样本，各有 526 个数据。可以发现，抽样后的数据的超期天数分布与总数据非常接近，可以认为样本 1 和样本 2 都具备随机性，所反映的情况可以代表总数据。样本 1 将作为模型建立的数据，而样本 2 则作为检验模型的数据。

表 1 总数据的超期天数情况

超期天数分段	汇总	百分比
1 ( $\leq 40$ )	940	21.43%
2 (40-100)	534	12.18%
3 (100-300)	833	18.99%
4 (300-1000)	930	21.20%
5 ( $> 1000$ )	1149	26.20%
总计	4386	100%

表 2 样本 2 的超期天数情况

超期天数分段	汇总	百分比
1 ( $\leq 40$ )	105	19.96%
2 (40-100)	65	12.36%
3 (100-300)	91	17.30%
4 (300-1000)	113	21.48%
5 ( $> 1000$ )	152	28.90%
总计	526	100%

表 3 样本 1 的超期天数情况

天数分段	汇总	百分比
1 ( $\leq 40$ )	118	22.43%
2 (40-100)	67	12.74%
3 (100-300)	92	17.49%
4 (300-1000)	110	20.91%
5 ( $> 1000$ )	139	26.43%
总计	526	100%

### 2.2.2 超期天数的统计分类

利用 excel 的关键字筛选工具对抽取的样本进行分类, 分类标准采取对世界图书馆分类学有相当大的影响, 被大部分美国的公共图书馆和学校图书馆都采用的杜威十进制图书分类法<sup>[3]</sup>。

000 计算机科学、资讯与总类

100 哲学与心理学

200 - 宗教

300 - 社会科学

400 - 语言

500 - 科学 (指自然科学)

600 - 技术应用科学

700 - 艺术与休闲

800 - 文学

900 - 历史、地理与传记

为方便起见, 只对书籍做 10 个大分类, 不再细分为分类法中的小分类。即分别用 0~9 10 个数字来分别代表 000~900 这 10 个类别。

## 3 基本假设

- (1) 假设罚款的目的纯粹是为了促使借阅者按期归还书本;
- (2) 假设本文所分的三类超期原因占借阅超期原因的极大部分;
- (3) 假设所使用的数据都真实准确;
- (4) 对于超期数大于 1000 天的书本, 视为丢失(不在考虑的超期原因之内)。

## 4 符号说明

$R_i$ : 超期原因; ( $i=1,2,3$ ) $A_{ij}$ : 第 $i$ 类书、第 $j$ 类天数的超期数; $\alpha_{ij}$ : 矩阵模型的元素; $m$ : 行数; $n$ : 列数; $F^{zl}_i$ 种类的权值; $F^{ts}_j$ 天数的权值;	$qx_i^{gd}$ : 原规则的期限; $\Delta qx_i$ : 合理的增加期限; $md_j$ : 第 $j$ 类天数区间的中心值; $FK$ : 罚款矩阵; $jfk$ : 基本罚款; $bfk$ : 补偿罚款; $gfk$ : 故意罚款;
---	---



$R$ ：原因矩阵模型； $y_i$ ： $R_i$ 原因； $f_i(t)$ ：连续变化规律； $qx_i$ ：合理期限； $fk_{ij}^{yd}$ ：罚款规则。；	$dGZ$ ：表示规则的合理指标； $dqx$ ：期限指标； $dFK$ ：罚款指标； $qx_i^{yd}$ ：现定的期限；
---	--

## 5 模型的建立与求解

从总体出发，分成四个步骤来对图书的超期原因进行分析，并制定出合理的还书体系。

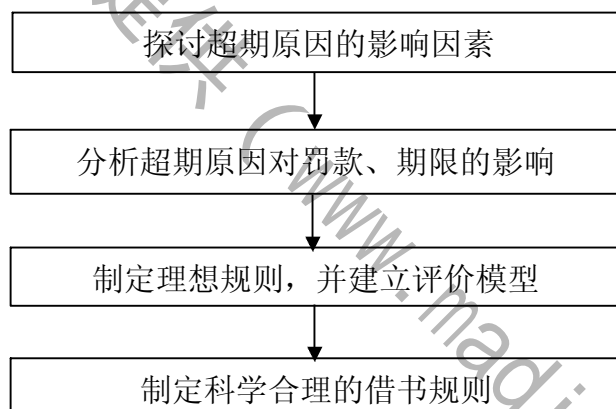


图 1 总体建模思路

第一步是探讨超期原因的影响因素，即在确定超期原因为无意，科研，有意三种的情况下，各种学科和超期的天数对超期原因的影响。然后收集所需的相关数据算出各个超期数的分布规律。

第二步是分析超期原因的影响因素。运用了加权的思想，根据各种书籍的性质，利用经验数据分别对不同类别的三种原因进行加权赋值，从而得出各种图书的种类和超期天数对超期原因的影响。

第三步是在前两步的基础上，分析了规则与超期原因和超期天数的关系，给出了规则制定的法则之后，建立了规则制定模型。并利用该模型得出理想规则。

第四步是根据超期原因的影响因素，设定了期限指标和罚款指标组成的来评价规则的合理度，制定出科学合理的借书规则，即确定最优的罚款数额和借期。

以上四个步骤是对图书的超期原因进行总体分析。

### 5.1 模型 I ——超期数的分布规律矩阵模型建立与应用

用“超期数”表示同一种类别的书在某一个特定时间段内超期的总本数。为了分析超期数的规律，将运用数据挖掘的统计方法建立超期数的分布规律矩阵模型。

### 5.1.1 超期数影响因素的分析

超期数的分布规律与图书类别和超期的天数有关。不同类别的书，由于其性质的不同，会导致不同的超期现象。例如：一般情况下，应用科学类书籍阅读时间较长，文学类阅读时间较短；超期天数的长短也影响到超期数的大小，一般情况下，超期天数越大，超期数会变小。为了形象地显示超期数分布规律与两者的关系，可将类别和超期天数分别表示为矩阵的行和列，建立一个矩阵模型。

### 5.1.2、模型的建立

如上述分析，建立超期数的分布矩阵模型如下：

$$A = (\alpha_{ij})_{mn} \quad i = 1, 2, \dots, m \quad j = 1, 2, \dots, n \quad (1.1)$$

其中： $i$  表示第  $i$  类书， $j$  表示第  $j$  类天数， $A_{ij}$  表示第  $i$  类书、第  $j$  类天数的超期数， $\alpha_{ij}$  为矩阵模型的元素， $m$ 、 $n$  分别表示行、列数。

运用矩阵模型对本题的数据进行分析，得到结果如下：

表 4 样本 1 的不同种类的超期天数统计表

天数 类别	1 ( $\leq 40$ )	2 ( $40-100$ )	3 ( $100-300$ )	4 ( $300-1000$ )	5 ( $>1000$ )	总计 (本)
0	30	13	10	20	28	101
1	0	1	1	1	1	4
2	1	0	0	2	0	3
3	9	11	15	17	19	71
4	6	2	12	11	11	42
5	9	3	4	13	11	40
6	30	20	22	21	54	147
7	10	7	13	8	10	48
8	7	6	10	16	12	51
9	3	2	4	4	6	19
总计	105	65	91	113	152	526

由此可以分析得，第 6 类（技术应用科学类）的书籍的超期数最多，第 0 类（计算机科学、资讯与总类）的超期数次之，两者的超期数与其他类别的相差较大，因此，可以得出结论：技术应用科学类与计算机科学、资讯与总类的书比较容易超期。

## 5.2 模型 II ——超期原因矩阵模型的建立与应用

根据各种图书的种类和超期天数对超期原因的影响，运用了加权的思想，利用经验数据分别对不同类别的三种原因进行加权赋值，从而得到一个超期原因矩阵模型。

### 5.2.1 对超期原因的分析

对于高校图书馆的借阅超期的原因，在上文的分析中已分为 1、无意错过还书期限；2、非恶意的故意超期（由于教学科研需要使用）；3、恶意的故意超期（由于罚款数目较少而故意拖延）。分别用  $R_1$ ,  $R_2$ ,  $R_3$  来表示。

对于每种原因，其实与借阅者本身、借阅的类别和超期天数都有关系。而借阅者的属性一般在其借阅书本的类别可以体现出来。因此，可以根据不同的借阅书籍类别和超期的天数得出该超期现象的最有可能原因。

### 5.2.2 指标选取

超期原因与借阅书籍类别的关系

由上述的分析得知，借阅者的属性已经被糅合在借阅书籍的类别里面。不同的书籍，对于超期原因是有一定的关联的。比如，借阅科学类的书籍时，极有可能是思想上比较热爱学习的人在需要的情况下借阅的。因此，针对该类书籍，可能性相对较大就是  $R_2$ ，而由于现实情况的复杂性，其他两个方面的原因也是有一定的发生概率，因此可运用加权的思想，根据各种书籍的性质，利用经验数据分别对不同类别的三种原因进行加权赋值。

$$F_i^{zl} = (\eta_{i1}, \eta_{i2}, \eta_{i3}) \quad i = 1, 2, \dots, m \quad (2.2)$$

其中， $F_i^{zl}$  表示第  $i$  类别书本的原因权属向量； $\eta_{i1}$ 、 $\eta_{i2}$ 、 $\eta_{i3}$  分别表示第  $i$  类别书本的原因  $R_1$ ,  $R_2$ ,  $R_3$  各自的权值。

超期原因与超期天数的关系

基于三大类超期原因的性质，可以得知，随着超期天数的变化，可能出现的原因也是不断变化着的。针对  $R_1$ ，可以得知，随着超期天数的增加， $R_1$  的出现可能是一个比较平缓的减少的规律；针对  $R_2$ ，随着超期天数的增加，其出现可能度是一个先增后减的规律；针对  $R_3$ ，随着超期天数的增加，其出现的概率可能是先减后增的。此外，在同一时间三种原因都可能同时作用于一个超期现象。为了能够区分说明，同样可采用加权赋值的方法。根据实验数据，可简单的得知离散点（不同超期天数上）的权值，因此，可采用插值拟合的方法，分别拟合出三类原因以超期天数为自变量的连续变化函数。

$$y_k^{ts} = f_k(t) \quad k = 1, 2, 3 \quad (2.3)$$

其中， $y_k^{ts}$  表示  $R_k$  原因， $f_k(t)$  第  $k$  种原因的连续变化规律。根据以上给定的插值点，运用保样样条插值法拟合出  $f_k(t)$  的曲线如下（其中褐： $R_1$ ；蓝： $R_2$ ；红： $R_3$ ）

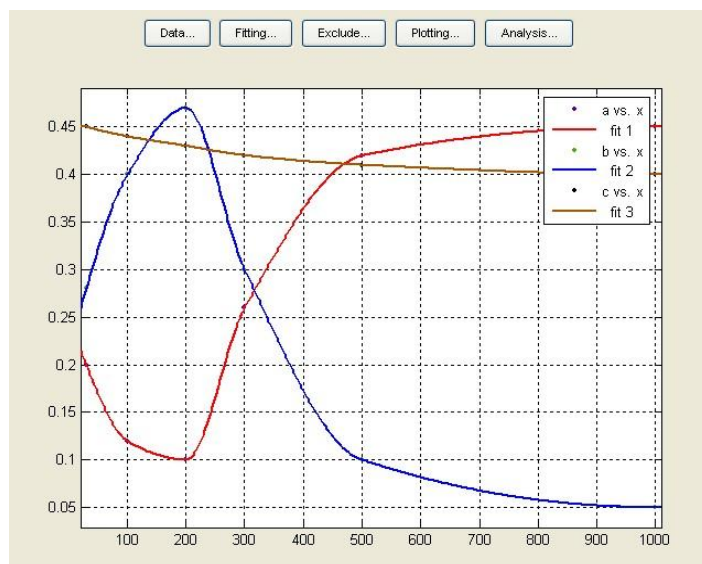


图 2 保样样条插值法拟合曲线

根据不同的超期天数分类，利用上述的连续变化函数，可以确定不同的离散加权赋值

$$F^{ts}_j = (y^{ts}_{j1}, y^{ts}_{j2}, y^{ts}_{j3}) \quad j=1, 2, \dots, n \quad (2.4)$$

说明： $F^{ts}_j$  表示第  $j$  类超期天数的权值向量； $y^{ts}_{j1}, y^{ts}_{j2}, y^{ts}_{j3}$  分别表示第  $j$  类超期天数的原因 R1, R2, R3 各自的权值。

### 5.2.3 超期原因权值的制定

由以上模型可以得出，不同类别书籍和不同超期天数共同决定着超期原因；而不同的原因，依据不同类别和超期天数都给定了相应的权值。因此，为了确定在某类书籍某个超期时间下的超期原因，可采用以下方法处理：

对于类别和超期天数的权值，采用加权平均的方法，求出在第  $i$  类书籍、第  $j$  类超期天数时，每种原因的权值：

$$F_k^{ij} = \frac{1}{2} F_i^{zl} + \frac{1}{2} F_j^{ts} \quad i=1, 2, \dots, m \quad j=1, 2, \dots, n \quad k=1, 2, 3 \quad (2.5)$$

其中  $F_k^{ij}$  表示第  $i$  类书籍、第  $j$  类超期天数时原因的权值向量， $k$  表示原因的类别。原因矩阵模型的建立

由上述分析可以得出，在第  $i$  类书籍、第  $j$  类超期天数时原因的权值向量  $F_k^{ij}$ ，同时判断第  $i$  类书籍、第  $j$  类的超期天数的三种原因的权值大小。取权值最大的原因为第  $i$  类书籍、第  $j$  超期天数的超期原因，表示为如下模型：

$$r_{ij} = \begin{cases} 1, & F_1^j = \max_{k=1,2,3} \{F_k^j\} \\ 2, & F_2^j = \max_{k=1,2,3} \{F_k^j\} \\ 3, & F_3^j = \max_{k=1,2,3} \{F_k^j\} \end{cases} \quad (2.6)$$

基于上述分析，建立原因矩阵模型如下：

$$R = (r_{ij})_{mn} \quad i = 1, 2, \dots, m \quad j = 1, 2, \dots, n \quad (2.7)$$

其中， $r_{ij}$  表示第  $i$  类书籍、第  $j$  类超期天数的原因。

#### 5.2.4 模型的运用

根据该模型 II 的原理，分析该题目的数据。

根据不同的种类，赋予的权值阵为：

$$F^{zl}_i = \begin{pmatrix} 0.2 & 0.5 & 0.2 \\ 0.4 & 0.3 & 0.2 \\ 0.4 & 0.3 & 0.2 \\ 0.3 & 0.3 & 0.3 \\ 0.25 & 0.45 & 0.2; \\ 0.2 & 0.5 & 0.2 \\ 0.2 & 0.55 & 0.15 \\ 0.4 & 0.3 & 0.2 \\ 0.5 & 0.2 & 0.2 \\ 0.35 & 0.35 & 0.2 \end{pmatrix}$$

根据不同的超期天数类，可依据上文对超期天数的分类，代入  $y_k^{ts}$  得到：

$$F^{ts}_j = \begin{pmatrix} 0.445 & 0.28 & 0.2 \\ 0.44 & 0.4 & 0.12 \\ 0.42 & 0.3 & 0.26 \\ 0.4 & 0.05 & 0.45 \end{pmatrix}$$

再根据式 (2.5) 和 (2.6)，利用 Matlab 编程得到的超期原因矩阵如下：

$$R = \begin{pmatrix} 2 & 2 & 2 & 3 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 3 \\ 2 & 2 & 2 & 3 \\ 2 & 2 & 2 & 3 \\ 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

由该超期原因矩阵可以得出，每一类书本在不同类的超期天数的原因。经过观察该矩阵发现，第 0 类（计算机科学、资讯与总类）、第 4 类（语言）、第 5 类（科学（指自然科学））和第 6 类（技术应用科学）的原因是 R2（教学科研使用需要），根据类别的分析，得出该结果合理性较高；其他类别的原因是 R1（无意错过还书期限）占很大的比例；这与 2.1.2 的各原因造成借阅超期所占的比例的分析结果是相对吻合的。

### 5.3 模型Ⅲ——规则制定模型的建立与应用

#### 5.3.1 对制定规则的分析

在借阅书籍的规则方面，制定方向主要是根据借阅期限和过期罚款额。期限是与实际的阅读时间直接相关的。一般规律是，书籍的阅读难度和阅读量越大，阅读时间越长，因此，其期限相应必须越长。基于我们所建立的模型 I 与模型 II 可以发现，借阅难度和阅读量的大小可以体现在超期原因上。而期限的长度设定也应该与超期数有关。比如，

在第  $i$  类书籍、第  $j$  类超期天数的超期原因显示，需要延长期限时，延长的长度，还必须与相应的超期数与该类的超期总数的比例有关。因此，比较合理的算法思想定义如下。

对于第  $i$  类书籍：

1) 在不同类超期天数时，根据模型 II 判断其超期原因，来确定需要修改位置  $(i, j)$ ，

2) 再对需要修改位置  $(i, j)$  上的超期数作分析，利用模型 I，以该需要修改位置上超期数占该类书籍总的超期数的比例作为该位置  $(i, j)$  的权值；以该位置超期天数区间的中心值为基准。

说明：假设区间为  $[a, b]$  时，其中心值为  $\frac{a+b}{2}$ 。

3) 需要修改的位置  $(i, j)$  的基准和其权值的乘积的和就是合理增加期限，该合理增加期限加上本来规定的期限就可以得出合理期限。

下面，通过举例子进一步形象地说明这个算法。假设要对第  $i$  类书籍进行分析。通

过模型 II 可以得出，该类书籍在 ( $<40$ ) 和 ( $40-100$ ) 这两类的超期天数上需要修改。若这两类的超期数分别为 60、40，则可以求出它们各自超期数占该类书籍总的超期数的比例。假设分别为 0.3 和 0.2。因此，合理增加期限为：

$0.3 * \frac{40+0}{2} + 0.2 * \frac{100+40}{2} = 20$ 。即若本来的规定期限为 30 天，所以可以得到合理期限应该为 50 天。

罚款额的确定则是与人的心理有关，额数的大小一般与借阅者还书的重视度成正比。即：罚款额越大，其重视程度会越高，但是考虑到有些是无意超期的情况和借阅者对图书馆的看法，罚款额又不能太大。基于模型 I 与模型 II，通过对不同原因可以确定出不同的罚款额。这样分类别的罚款规定相比统一单调的规则来说是比较合理的。因此，对于不同原因，根据数据经验性设定其罚款形式：

原因 1：基本罚款

原因 2：基本罚款-需要补偿款

原因 3：基本罚款+故意罚款

根据模型 II 的原因矩阵和本文规定罚款规则，可以算出罚款额矩阵。

### 5.3.2 模型的建立与应用

根据上述分析建立期限和罚款模型。

为了分析的方便，将模型 I 与模型 II 这两个矩阵模型结合成一个矩阵模型

$$C = (c_{ij})_{mn} \quad i=1,2,\dots,m \quad j=1,2,\dots,n \quad (3.1)$$

其中， $c_{ij} = (\alpha_{ij}, r_{ij})$ 。

因此，建立的期限模型为：

$$\begin{aligned} qx_i &= qx_i^{gd} + \Delta qx_i \\ \Delta qx_i &= \sum_j^n \alpha_{ij} * md_j \quad \text{if } r_{ij} = 2 \end{aligned} \quad (3.2)$$

其中， $qx_i$  表示合理期限； $qx_i^{gd}$  表示原规则的期限； $\Delta qx_i$  合理的增加期限； $\alpha_{ij}$  为 i 类书籍 j 类天数的超期数； $md_j$  表示第 j 类天数区间的中心值。

注意：假设  $md_j$  区间为  $[a, b]$ ，则  $md_j = \frac{a+b}{2}$ 。

建立的罚款矩阵模型为：

$$FK = (fk_{ij})_{mn}$$

$$fk_{ij} = \begin{cases} jfk & \text{if } r_{ij} = 1 \\ jfk - bfk & \text{if } r_{ij} = 2 \\ jfk + gfk & \text{if } r_{ij} = 3 \end{cases}$$

其中， $FK$  表示罚款矩阵； $jfk$  表示基本罚款； $bfk$  表示补偿罚款； $gfk$  表示故意罚款。

针对本题数据，我们假设该题数据的期限是 1.5 个月。我们根据利用 Matlab 进行编程得可以得到设定的合理期限如表 5 所示。

表 5 不同类别书籍的合理借阅期限 单位：天

类别	0	1	2	3	4	5	6	7	8	9
期限	87	45	45	45	75	59	121	45	45	45

即：第 0 类（计算机科学、资讯与总类）、第 4 类（语言）、第 5 类（科学（指自然科学））和第 6 类（技术应用科学）都是需要调整期限的，他们调整后的合理期限分别为 87, 75, 59、121 天。

根据罚款模型可以得到罚款矩阵，在此可假定，基本罚款额  $jfk = 0.1$ ，补偿罚款额  $bfk = 0.05$ ，故意罚款额  $gfk = 0.2$ ，另外，根据本文的超期天数的分类，决定 1000 以上的不去考虑。然后再根据该罚款矩阵模型，编程得出 1000 之前的罚款矩阵如下，

$$FK = \begin{pmatrix} 0.0500 & 0.0500 & 0.0500 & 0.3000 \\ 0.1000 & 0.1000 & 0.1000 & 0.1000 \\ 0.1000 & 0.1000 & 0.1000 & 0.1000 \\ 0.1000 & 0.1000 & 0.1000 & 0.3000 \\ 0.0500 & 0.0500 & 0.0500 & 0.3000 \\ 0.0500 & 0.0500 & 0.0500 & 0.3000 \\ 0.0500 & 0.0500 & 0.0500 & 0.0500 \\ 0.1000 & 0.1000 & 0.1000 & 0.1000 \\ 0.1000 & 0.1000 & 0.1000 & 0.1000 \\ 0.1000 & 0.1000 & 0.1000 & 0.1000 \end{pmatrix}$$

根据种类的性质，可以得出不同种类的罚款规则。

第 0 类（计算机科学、资讯与总类）：罚款额改为 300 天之前罚款 0.05 元，300-1000 的罚款 0.3 元/天。



第4类（语言）和第5类（科学（指自然科学））：罚款规则同第0类。

第6类（技术应用科学）：罚款额改为1000天之前罚款0.05元/天。

其他种类不变，还是0.1元。

对于1000天之后的，由于时间太久，我们可以把这部书当做丢失的书。

## 5.4 模型IV——评价规则的合理性模型

### 5.4.1 评价规则的合理性的分析

根据模型III，可以得知算出的合理期限 $qx_i$ 和合理的罚款的矩阵 $FK$ 。并以此为标准。为判断已定规则的合理性，可设定了两个指标：期限指标和罚款指标。

由于已经有了对比标准，因此可建立的指标模型如下：

$$dqx = \frac{\sum_i^m |qx_i - qx_i^{yd}|}{\sum_i^m qx_i} \quad (4.1)$$

$$dFK = \frac{\sum_i^m \sum_j^n |fk_{ij} - fk_{ij}^{yd}|}{\sum_i^m \sum_j^n fk_{ij}} \quad (4.2)$$

其中， $dqx$ ， $dFK$ 分别表示期限指标和罚款指标； $qx_i^{yd}$ ， $fk_{ij}^{yd}$ 分别为现定的期限和罚款规则。

为了统一衡量规则是否合理，对两个指标进行加权平均，得到：

$$dGZ = \frac{2}{3}dqx + \frac{1}{3}dFK \quad (4.3)$$

其中， $dGZ$ 表示规则的合理指标。

通过经验数据，可以假设为以下的评判区间：

表6 合理度评判区间

类别	合理	不合理
区间	(0,0.3]	(0.3,1]

### 5.4.2 评价规则的合理性的运用

在模型III的应用中，已经得到了合理的期限，当期限为45天（1.5个月），可计算得出 $dqx=0.265$ ，

在模型III的应用中，已经得到了合理的罚款矩阵，当罚款额为 0.1 元是，算出：  
 $dFK = 0.427$

因此，由（4.3）得到  $dGZ = \frac{2}{3}dqx + \frac{1}{3}dFK = 0.32$ 。

根据确定的评判区间发现，该规则是不合理的。

### 5.4.3 合理规则

考虑到本文中讨论的种类比较多，不能根据文中所设模型算出来的合理规则就直接作为某一个图书馆的规则。因为这种情况下制定的规则比较复杂，一方面学生难以遵守，一方面会增加管理难度。考虑到III模型得出的结果，综合以上情况可归纳出以下规则：理科类（第 0、5、6 类）：期限为 90 天（3 个月），罚款前 300 天为 0.05, 300-1000 为 0.3 元；文科类（第 1、2、3、4、7、8、9 类）：期限为 45 天，罚款为 0.1 元。运用模型IV对本文定义的规则进行评价：

计算得到  $dGZ = \frac{2}{3} * 0.12 + \frac{1}{3} * 0.31$   $dGZ = 0.19$ 。

此时根据判断区间发现，该规则是合理且可行的。

## 6 模型评价与改进

### 6.1 模型的优点

①模型建立的合理性，模型是建立在对收集的数据进行充分的挖掘的基础之上的，通过数据之间的关系提炼出各个变量之间的关系，建立起模型；

②对一些未量化的指标建立模型，进行合理的量化，例如对超期原因的量化，建立超期数的分布规律矩阵模型；

③模型的建立是按照问题的解决的思路进行的，首先分析和发现现有规律得出超期数的分布规律矩阵模型，然后对现有的规律进行简化得到可能性矩阵模型，再建立规则制定模型，最后对准则进行评价，层次渐进易于理解；

④模型的有良好的可推广性，模型只需根据不同高校图书馆的特点修改必要的变量，就可以适用于不同学校的图书超期问题。

### 6.2 模型缺点

①指标的定义存在一定的主观性，由于问题的复杂因素较多，不能对所有因素进行全面的考虑。

②模型涉及到很多方面的相关数据，对数据的收集有一定的依赖性。

### 6.3 模型的改进

①针对缺点一，可以通过参照更多的数据，同时设定一个相关的调查来对学生借阅超期的原因进行分析，根据统计出来的结果来制定权值。

②针对缺点二，可再通过一组样本数据来检验模型的适用性与合理性。从而根据

数据可以发现更加有效的准确评价因素。

## 7 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 页, 2009 年第 17 期  
 [2]夏燕燕, 高校图书馆图书超期因素及其对策探讨, 华章, 142~143 页, 2008 年 7 月  
 [3]杜威十进图书分类法 <http://zh.wikipedia.org/zh-cn/>

## 8 附录

```
Rm.m/*得出超期原因矩阵
Fz=[0.2 0.5 0.2;0.4 0.3 0.2;0.40.3 0.2;0.30.3 0.3;0.25 0.45 0.2;
    0.2 0.5 0.2;0.2 0.55 0.15;0.4 0.3 0.2;0.50.2 0.2;0.35 0.35 0.2];
Ft=[0.445 0.28 0.2;0.44 0.4 0.12;0.42 0.3 0.26;0.4 0.05 0.45];
[m1,l1]=size(Fz);[m2,l2]=size(Ft);
R=zeros(m1,m2);B=zeros(1,3);
for i=1:m1
    for j=1:m2
        B=l(i,:)+t(j,:);
        if (B(1)<=B(2))
            if (B(2)<=B(3)) R(i,j)=3;
            else R(i,j)=2;
            end
        else if (B(1)<=B(3)) R(i,j)=3;
            else R(i,j)=1;
            end
        end
        B=zeros(1,3);
    end
end
R,

cij.m/*得出增加合理期限
function cij(A)
R m;
[m,l]=size(A);
s=zeros(m,1);sum=0;
for i=1:m
```

```

        for j=1:l
            sum=sum+A(i,j);
        end
        s(i,1)=sum;
        sum=0;
    end
    k=zeros(m,l);
    for i=1:m
        k(i,:)=A(i,:)/s(i,1);
    end
    q=zeros(10,1);f=0;
    b=k';
    for i=1:10
        if (R(i,1)==2)
            f=b(i,1)*20;
        end
        if (R(i,2)==2)
            f=f+b(i,2)*70;
        end
        if (R(i,3)==2)
            f=f+b(i,3)*200;
        end
        q(i,1)=f;
        f=0;
    end
    q',

```

Fk.m

R m; /\*得出罚款矩阵

F=zeros(10,4);

for i=1:10

for j=1:4

if (R(i,j)==1)

F(i,j)=0.1;

end

if (R(i,j)==2)

F(i,j)=0.05;

end

if (R(i,j)==3)

F(i,j)=0.3;

end

end

end

F,

数学中国提供 (www.madio.net)