

出版社资源配置的问题研究

摘 要

本文针对出版社的资源配置问题进行了模型研究，以出版社的经济效益为目标函数，以总社为分社的各个学科分配的书号数量为决策变量建立了非线性规划模型，并运用多种求解方法对模型进行求解，并得到了最优的出版社资源配置方案。

我们首先用各课程的年销售量表示出总经济效益，并分析出影响销售量的因素为课程的需求量，市场竞争力和分配到的书号数；然后，通过对附件中相关数据的统计，我们对 2001 年至 2005 年的需求量和市场竞争力进行了量化分析，得到了每一年每门课程的需求量系数和市场竞争力系数，并应用灰色系统分析方法预测了 2006 年每门课程的需求量系数和市场竞争力系数。

为了得到销售量和课程需求量，市场竞争力及分配到的书号数的具体函数关系，在已知 2001 年至 2005 年各课程需求量系数和市场竞争力系数的条件下，我们应用统计回归的方法借助 MATLAB 的统计工具箱较为准确地拟合出了这一函数关系。

在模型建立中，我们将灰度预测所得 2006 年的需求量系数和市场竞争力系数代入到拟合出的销售量的函数式中，求出 2006 年经济效益的表达式，并以此作为目标函数，以总社为分社的各个学科分配的书号数量为决策变量，并考虑相关约束条件建立了非线性规划模型。

在模型求解中，我们首先应用 LINGO 软件得到了全局最优解，而后又引入强势度的概念，提出了快速贪心算法，并得到了相同的最优分配方案。

最后，我们根据模型求解结果为 2006 年的强势产品定位，并比较了各分社之间的经济效益，同时对求解出的 2006 年出版社总销售量和总销售额的准确度进行了分析。

关键字：需求量系数，市场竞争力系数，灰度预测，统计回归，强势度，快速贪心算法

目 录

问题重述.....	3
模型假设.....	3
符号说明.....	3
问题分析.....	4
模型建立和求解.....	4
1. 模型分析：	4
1.1 经济效益的表示：	4
1.2 年需求量系数：	5
1.3 市场竞争力系数：	7
1.4 各课程书目年销售量的统计回归分析：	11
2. 模型建立：	14
3. 模型求解：	14
3.1 LINGO 求解：	14
3.2 贪心算法求解：	15
4. 结果分析：	17
4.1 2006 年强势产品定位：	17
4.2 2006 年各分社经济效益比较：	17
4.3 2006 年总销售量的准确度分析：	18
4.4 2006 年总销售额的准确度分析：	18
模型评价.....	19
参考文献.....	19

问题重述

出版社的资源主要包括人力资源、生产资源、资金和管理资源等，它们都捆绑在书号上，经过各个部门的运作，形成成本（策划成本、编辑成本、生产成本、库存成本、销售成本、财务与管理成本等）和利润。

某个以教材类出版物为主的出版社，总社领导每年需要针对分社提交的生产计划申请书、人力资源情况以及市场信息分析，将总量一定的书号数合理地分配给各个分社，使出版的教材产生最好的经济效益。事实上，由于各个分社提交的需求书号总量远大于总社的书号总量，因此总社一般以增加强势产品支持力度的原则优化资源配置。资源配置完成后，各个分社（分社以学科划分）根据分配到的书号数量，再重新对学科所属每个课程作出出版计划，付诸实施。

资源配置是总社每年进行的重要决策，直接关系到出版社的当年经济效益和长远发展战略。由于市场信息（主要是需求与竞争力）通常是不完全的，企业自身的数据收集和积累也不足，这种情况下的决策问题在我国企业中是普遍存在的。

对于这种情况，我们需要根据该出版社所掌握的一些数据资料，利用数学建模的方法，在信息不足的条件下，提出以量化分析为基础的资源（书号）配置方法，给出一个明确的分配方案，向出版社提供有益的建议。

模型假设

1. 总社直接为各个分社的各个学科分配书号数，各个分社自己不能分配书号数。
2. 出版社所有课程书目的年销售总量主要受市场的年需求量的影响，因此出版社年销售总量的变化率与市场年需求量的变化率近似相等。
3. 当分社分配到的书号数对应的工作量大于该分社工作人员的工作能力时，分社仍有可能通过其他方式完成任务工作量。通过观察附件 4 和附件 5，我们发现，数学类分社在 2001 年至 2005 年间每年分得的书号数都大于该社工作人员能完成的书号个数，这证明该社通过其他方式（与其他分社合作或加班等）完成了任务工作量。故在模型建立中不考虑分社工作人员的工作能力限制。

符号说明

1. x_i ：总社为分设的每门课程实际分配的书号个数（ $i=1,2,\dots,72$ ）。
2. p_i ：每门课程对应书目的平均价格（ $i=1,2,\dots,72$ ）。
3. s_i ：每门课程对应书目的年销售量（ $i=1,2,\dots,72$ ）。
4. r_i ：每门课程对应书目的年需求量系数（ $i=1,2,\dots,72$ ）。

5. c_i : 每门课程对应书目的市场竞争力系数 ($i=1,2,\dots,72$)。
6. a_i : 分社为每门课程申请的书号数 ($i=1,2,\dots,72$)。
7. b : 出版社所有教材的统一利润率。
8. Q : 出版社一年的总经济效益。

问题分析

本问题的最终目的就是为了求得总社为分社的各门课程实际分配的书号数,以使出版的教材产生最好的经济效益。这里首先就是一个如何衡量经济效益的问题,根据题目的已知信息,我们选择用每门课程对应书目的平均价格,年销售量和利润率来表示出版社的经济效益。这里问题的关键就是如何表示每门课程对应书目的年销售量。我们分析认为,每门课程的年销售量应该与该门课程书目的年需求量,该出版社的市场竞争力以及总社分配的该门课程的书号个数有关。为此,我们就要借助于 2001 年至 2005 年的已知相关信息建立起年销售量与这三个因素的具体函数关系,并把这一关系应用到 2006 年。由于 2006 年的年需求量和出版社竞争力未知,我们就需要应用前 5 年的年需求量和出版社竞争力情况通过预测和统计的方式得到。在已知年需求量和市场竞争力的情况下,我们就可以建立起年销售量和每门课程分配书号数的函数关系,进而表示出总的经济效益。这样,我们就可以以总社为分社每门课程分配的书号数为决策变量建立规划模型,通过模型求解得到最优的书号分配方案。

模型建立和求解

1. 模型分析:

1.1 经济效益的表示:

如题目要求所说,我们建立模型的目的就是给出版社提供一个明确的书号数分配方案,使出版的教材产生最好的经济效益。这里首先就是一个如何衡量经济效益的问题,按题目所说,经济效益就是出版的所有教材的利润总和。这里的利润有两种表示方法,首先是出版教材的收入减去出版成本(包括策划成本,编辑成本,生产成本,库存成本,销售成本,财务与管理成本等),再者就是用每本教材的价格乘以该本教材的利润率。在本题目中,由于没有给出出版成本的具体信息,而是在附件 4 中给出了每门课程不同书目价格均值,并规定该出版社在定价时保持对所有教材利润率统一,所以我们在模型中采用每本教材的价格乘以该本教材的利润率的方式来表示该本教材的利润,并求出所有教材的利润总和来表示总的经济效益。

下面讨论如何定量表示总的经济效益。已知同一课程不同书目价格差别不大,同时销售量相近,故假设每门课程对应书目的平均价格为 p_i , 年销售量平均值为 s_i , 并取定

所有教材的统一利润率为 b , 则总的经济效益 Q 可以表示为:
$$Q = \sum_{i=1}^{72} p_i \cdot s_i \cdot b。$$

在上面的经济效益表达式中, p_i 为题目已知, b 为一确定常数, 故我们需要确定

的是每门课程对应书目的年销售量平均值 s_i 。经过分析,我们认为,每门课程对应书目的年销售量应该与该书目的年需求量,该出版社的市场竞争力以及总社分配的该门课程的书号个数有关。由于总社分配的每门课程的书号个数 x_i 为最终的求解对象,因此我们将对每门课程对应书目的年需求量和出版社市场竞争力对年销售量的影响进行量化分析,为此,我们分别引入了年需求量系数和出版社市场竞争力系数的概念,具体分析如下:

1.2 年需求量系数:

1.2.1 01-05 年年需求量的量化分析:

为了定量的分析大学生对教材的需求量,我们需要得知近几年来大学生对每门课程书目的使用情况,进而预测出下一年大学生对每门课程书目的需求情况。附件 2 中的问卷调查数据即反映了在过去几年(2001 年至 2005 年)大学生对每门课程书目的使用情况。在此问卷调查数据表中,对于每一年,我们首先筛选出包含出版社所关注的 72 门课程的调查问卷,然后分别统计包含每门课程的调查问卷的数量与调查问卷总数量的比值,并把这一比值当作当年该门课程的相对需求量系数。表 1 给出了我们得出的 2001 年至 2005 年各门课程的相对年需求量系数:

课程代码	2001	2002	2003	2004	2005
1	0.00547	0.005477	0.005489	0.005476	0.005475
2	0.014585	0.014606	0.014637	0.014602	0.0146
3	0.001443	0.001445	0.001449	0.001445	0.001445
4	0.001975	0.001978	0.001982	0.001977	0.001977
5	0.003039	0.003043	0.00305	0.003042	0.003042
6	0.01261	0.012628	0.012655	0.012625	0.012623
7	0.007824	0.007836	0.007852	0.007605	0.007756
...
70	0.000836	0.000837	0.000839	0.000837	0.000836
71	0.000912	0.000913	0.000915	0.000913	0.000912
72	0.001291	0.001293	0.001296	0.001293	0.001293

表 1 2001 年至 2005 年各课程的相对需求量系数(部分)

我们从表 1 中可以看出,各课程的相对年需求量系数只能反映出某一年各课程之间的需求量比较,而不能反映出某课程需求量随每一年的变化情况。也就是说,表 1 中的相对需求量系数不能反映出市场总需求量随年份变化对每门课程需求量的影响。为了体现市场总需求量的影响,我们定义 2001 年的总需求量系数为 1,并计算出 2001 年至 2005 年每一年总需求量的增长率,进而得到每一年的总需求量系数;然后,用每年的总需求量系数乘以当年的各个课程的相对需求量系数,这样就得到了各课程各年份的绝对需求量值。

为了计算总需求量的增长率,我们首先假设出版社所有课程书目的年销售总量主要受市场的年需求量的影响,因此出版社年销售总量的变化率与市场年需求量的变化率近似相等。我们在附件 3 中累加各课程的实际销售量得到出版社每年的总销售量,进而计算出出版社年销售量的变化率,我们用这一变化率来表示总需求量的变化率,从而得到每一年的总需求量系数。表 2 列出了我们求出的总需求量系数值:

2001	2002	2003	2004	2005
1	1.0800	1.2970	1.6829	2.0835

表 2 2001 年至 2005 年总需求量系数值

最后,我们用表 2 中每一年的总需求量系数乘以表 1 中对应年的相对年需求量系数,就得到了 2001 年至 2005 年各课程的绝对需求量系数,我们把这一系数称为年需求量系数。表 3 给出了 2001 年至 2005 年各课程的年需求量系数值:

课程代码	2001	2002	2003	2004	2005
1	0.00547	0.005916	0.007119	0.009216	0.011407
2	0.014585	0.015774	0.018984	0.024575	0.030419
3	0.001443	0.001561	0.001879	0.002432	0.00301
4	0.001975	0.002136	0.002571	0.003328	0.004119
5	0.003039	0.003286	0.003955	0.00512	0.006337
6	0.01261	0.013638	0.016414	0.021248	0.0263
7	0.007824	0.008463	0.010185	0.012799	0.01616
...
70	0.000836	0.000904	0.001088	0.001408	0.001743
71	0.000912	0.000986	0.001187	0.001536	0.001901
72	0.001291	0.001397	0.001681	0.002176	0.002693

表 3 2001 年至 2005 年各课程的需求量系数 (部分)

1.2.2 06 年年需求量的预测:

在已知每门课程前 5 年的需求量系数后,我们需要预测出每门课程在 2006 年的需求量系数,为了预测 2006 年的需求量系数,我们只有前 5 年的需求量系数作为参考,作为一个预测问题,显然供预测的数据量太少了,如果采用一般的预测方法,则预测的结果很可能出现很大的误差,为此,我们采用灰色系统分析方法对需求量系数进行灰度预测。所谓灰色系统分析方法主要是根据具体灰色系统的行为特征数据,充分利用数量不多的数据和信息寻求相关因素自身与各因素之间的数学关系,即建立相应的数学模型;而灰色预测则是依据目前已有的数据,利用灰色模型对系统行为特征的发展变化规律进行估计预测,同时对在特定时区内发生事件的未来时间分布情况做出研究。

在我们的预测问题中,把前 5 年的需求量系数看作灰色系统的行为特征数据并建立相应的灰色模型,进而利用灰色预测得到 2006 年的需求量系数。

在图 1 中,我们列举了几个利用灰度预测得到的预测函数图像与原始数据的位置关系:

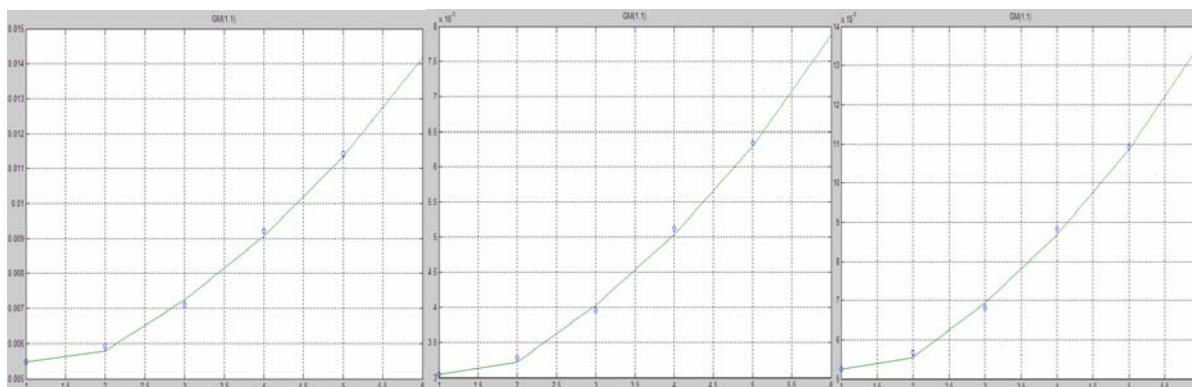


图 1 灰度预测得到的预测函数图像与原始数据的位置关系

从图 1 我们可以看出，灰度预测得到的预测函数图像能够很好的反映出数据的变化趋势，因此可以认为由灰度预测得到的 2006 年的预测值也是较为准确的。在表 4 中我们给出了利用灰度预测得到的 2006 年的 72 门课程书目的需求量系数值。

课程代码	1	2	3	4	5	6	7	8	9
需求量系数	0.0142	0.0378	0.0037	0.0051	0.0079	0.0327	0.0199	0.0112	0.0200
课程代码	10	11	12	13	14	15	16	17	18
需求量系数	0.0071	0.0116	0.0136	0.0134	0.0352	0.0165	0.0102	0.0079	0.0218
课程代码	19	20	21	22	23	24	25	26	27
需求量系数	0.0331	0.0065	0.0199	0.0264	0.3752	0.0083	0.0278	0.0774	0.0053
课程代码	28	29	30	31	32	33	34	35	36
需求量系数	0.2319	0.0250	0.0538	0.2767	0.0085	0.2211	0.0098	0.0061	0.0100
课程代码	37	38	39	40	41	42	43	44	45
需求量系数	0.0051	0.0142	0.0098	0.0059	0.0459	0.1274	0.0720	0.1260	0.1266
课程代码	46	47	48	49	50	51	52	53	54
需求量系数	0.1181	0.0887	0.0356	0.0118	0.0024	0.0069	0.0112	0.0045	0.0071
课程代码	55	56	57	58	59	60	61	62	63
需求量系数	0.0026	0.0037	0.0031	0.0033	0.0020	0.0037	0.0020	0.0039	0.0033
课程代码	64	65	66	67	68	69	70	71	72
需求量系数	0.0043	0.0026	0.0049	0.0033	0.0033	0.0039	0.0022	0.0024	0.0033

表 4 预测法得到的 2006 年的 72 门课程书目的需求量系数值

1.3 市场竞争力系数:

1.3.1 01-05 年市场竞争力的量化分析:

出版社在市场上的竞争力主要表现在两个方面：首先是消费者对出版社的偏爱程度，再者就是出版社产品在市场上的占有率，下面进行分别讨论：

1) 偏爱程度:

消费者对出版社的偏爱程度又由两个因素决定：消费者对出版社产品的满意度和消费者对出版社的忠诚度。这其中忠诚度是主要因素，若消费者对出版社的忠诚度高，则消费者在购买任何书时都会首选甚至只选该出版社的书，所以忠诚度在很大程度上决定了消费者对出版社的偏爱程度。再者就是消费者对出版社产品的满意度，消费者的满意度越高，其再次选择该出版社产品的可能性就越大，对出版社的偏爱程度也就越高。

在附件 2 中，消费者对出版社的忠诚度可以反映为该出版社在消费者心中的位置，位置越高，则忠诚度越高。而消费者对出版社出版教材的满意度评价打分则直接反映了消费者对出版社产品的满意度。为了量化的表示消费者对出版社的偏爱程度，我们用出版社在消费者心中的位置值的倒数作为忠诚度，用消费者对出版社出版教材的满意度评价打分（四类）之和作为满意度，并以两者的乘积作为消费者对出版社的偏爱程度值。即：偏爱程度=满意度×忠诚度。

在附件 2 中，对于每一年，我们首先筛选出出版社所关注的 72 门课程的调查问卷，然后分别按照上述量化方法计算每张问卷对应的偏爱程度值。最后对同一出版社同一门课程对应的调查问卷（可能有 0 张，一张或多张）的偏爱程度值取均值作为该出版社该门课程的偏爱程度值。这样我们就得到了每一年的 24 个出版社的 72 门课程的各自的偏爱程度值，表 5 给出了 2001 年的偏爱程度值：

出版社 课程	P030	P044	P063	P091	P102	P106	P110	P115	...	P522	P534
1	0	0	0	8.55	0	0	0	5.52	...	0	0
2	0	0	0	7.10	0	0	0	6.85	...	0	0
3	0	0	0	3.56	0	0	0	0	...	0	0
4	0	0	0	0	0	0	0	5.33	...	0	0
5	0	0	2.66	3.75	0	0	0	2.83	...	0	0
6	0	2.83	0	4.4	0	0	0	5.98	...	0	0
...
71	0	0	0	0	0	0	0	5.29	...	0	4.27
72	0	0	0	0	0	0	0	5.27	...	0	0

表 5 2001 年各出版社各课程的偏爱程度值（部分）

在上表中，某出版社某课程的偏爱程度为 0 表示没有针对该出版社该课程的调查问卷，也反映出消费者对这一课程数目的偏爱程度很低。

2) 市场占有率：

影响出版社市场竞争力的另一个重要因素就是出版社产品的市场占有率。对本问题来说，市场占有率还是应该通过附件 2 的课程信息进行计算。对于每一年，我们首先筛选出出版社所关注的 72 门课程的调查问卷，然后分别统计各个出版社各个课程所对应的调查问卷的数量。最后用各个出版社各个课程的调查问卷数量除以该课程对应的所有出版社的问卷数量，我们把这个商就作为该出版社该课程的市场占有率。这样我们就得到了每一年的 24 个出版社的 72 门课程的各自的市场占有率，2001 年的数据如表 6 所示：

出版社 课程	P030	P044	P063	P091	P102	P106	P110	P115	...	P522	P534
1	0	0	0	0.069	0	0	0	0.194	...	0	0
2	0	0	0	0.016	0	0	0	0.114	...	0	0
3	0	0	0	0.263	0	0	0	0	...	0	0
4	0	0	0	0	0	0	0	0.038	...	0	0

5	0	0	0.075	0.050	0	0	0	0.050	...	0	0
6	0	0.012	0	0.012	0	0	0	0.084	...	0	0
...
71	0	0	0	0	0	0	0	0.417	...	0	0.167
72	0	0	0	0	0	0	0	0.765	...	0	0

表 6 2001 年各出版社各课程的市场占有率（部分）

在上表中，某出版社某课程的市场占有率为 0 表示没有针对该出版社该课程的调查问卷，也反映出消费者对这一课程数目的市场占有率很低。

通过前面的讨论，我们得到了偏爱程度和市场占有率的量化表示，这样我们便可以得到市场竞争力的表示方法。我们定义：市场竞争力=偏爱程度×市场占有率。这样，对于每一年，我们将前面得到的偏爱程度和市场占有率的两个 72×24 的矩阵的对应位置相乘，便得到了每一年的 24 个出版社的 72 门课程的各自的市场竞争力矩阵，表 7 给出了 2001 年的数据：

出版社 课程	P030	P044	P063	P091	P102	P106	P110	P115	...	P522	P534
1	0	0	0	0.594	0	0	0	1.073	...	0	0
2	0	0	0	0.111	0	0	0	0.785	...	0	0
3	0	0	0	0.937	0	0	0	0	...	0	0
4	0	0	0	0	0	0	0	0.205	...	0	0
5	0	0	0.199	0.188	0	0	0	0.142	...	0	0
6	0	0.034	0	0.053	0	0	0	0.504	...	0	0
...
71	0	0	0	0	0	0	0	2.204	...	0	0.711
72	0	0	0	0	0	0	0	4.030	...	0	0

表 7 2001 年各出版社各课程的市场竞争力（部分）

由于我们的目的是要讨论 A 出版社的各门课程的市场竞争力情况，而表 6 中 A 出版社所在列的市场竞争力值只反映了 A 出版社各门课程市场竞争力的绝对大小，如果单独讨论 A 出版社的市场竞争力，则这些绝对值就失去意义。因此我们需要得到 A 出版社各门课程市场竞争力的相对大小。为此，我们将 A 出版社各门课程的市场竞争力绝对值除以所有出版社在该门课程的市场竞争力绝对值之和，就得到了 A 出版社各门课程市场竞争力的相对值，我们把这些市场竞争力的相对值作为 A 出版社各门课程的市场竞争力系数。表 8 给出了 2001 年的市场竞争力系数：

课程代码	2001	2002	2003	2004	2005
1	0.18015	0.22315	0.23864	0.31289	0.25105
2	0.12782	0.088638	0.13521	0.12609	0.12566
3	0	0.10714	0.0625	0.28037	0.085308
4	0.036521	0.1334	0.076696	0.092445	0.1619
5	0.03166	0.055479	0.04829	0.071571	0.066445

6	0.084142	0.054424	0.069917	0.077067	0.094705
7	0.24875	0.28219	0.30473	0.23909	0.27575
...
70	0.59076	0.67062	0.59585	0.65534	1
71	0.403	0.39802	0.31288	0.45946	0.69296
72	0.82236	0.72527	0.75281	0.94954	1

表 8 2001 年 A 出版社各课程的市场竞争力系数（部分）

1.3.2 06 年市场竞争力的预测：

在已知每门课程前 5 年的市场竞争力系数后，我们需要预测出每门课程在 2006 年的市场竞争力系数，我们这里仍通过灰度预测法来预测 2006 年的市场竞争力系数。在图 2 中，我们列举了几个利用灰度预测得到的预测函数图像与原始数据的位置关系：

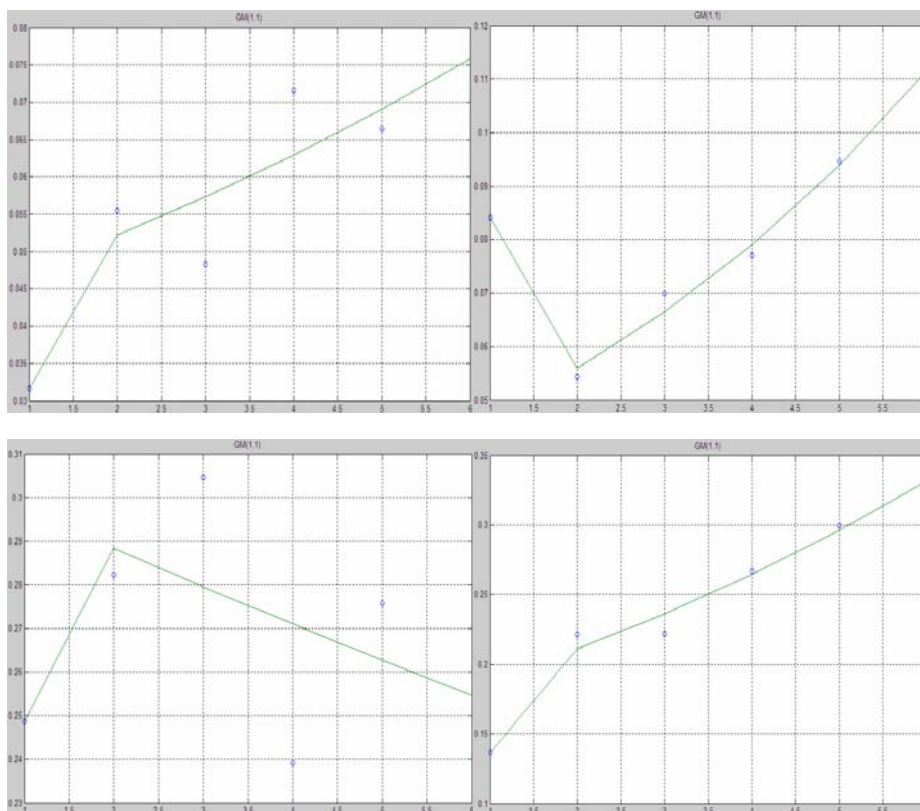


图 2 灰度预测得到的预测函数图像与原始数据的位置关系

从图 2 中我们可以看到，利用灰度预测得到的预测函数图像与原始数据的变化趋势是大致相同的。但有些数据由于变化趋势不明显或杂乱无章，使得灰度预测的效果不佳。对于这些变化无规律的数据，我们回避预测的方法而转向分析其统计特性，求出这些数据的均值，并把均值作为其对应课程 2006 年的市场竞争力系数。

经过预测和均值处理后的 2006 年课程的市场竞争力系数如表 9 所示：

课程代码	1	2	3	4	5	6	7	8	9
竞争力系数	0.2958	0.1444	0.1671	0.1507	0.0758	0.1116	0.2546	0.0347	0.1974
课程代码	10	11	12	13	14	15	16	17	18
竞争力系数	0.3316	0.8185	0.6132	0.0903	0.0618	0.1296	0.2781	0.7380	0.2347

课程代码	19	20	21	22	23	24	25	26	27
竞争力系数	0.5315	0.8123	0.4320	1.0198	0.9997	1.0980	1.0236	0.8818	1.0347
课程代码	28	29	30	31	32	33	34	35	36
竞争力系数	0.0140	0.3676	0.5440	0.0859	0.2245	0.0070	0.6474	0.6123	0.5160
课程代码	37	38	39	40	41	42	43	44	45
竞争力系数	0.1561	0.1775	0.3315	0.2538	0.2800	0.4047	0.3717	0.4061	0.4941
课程代码	46	47	48	49	50	51	52	53	54
竞争力系数	0.4072	0.4391	0.4857	0.9070	0.4891	0.6620	0.8410	1.0850	1.0021
课程代码	55	56	57	58	59	60	61	62	63
竞争力系数	0.1600	0.4904	0.5886	0.1530	1.0000	1.0538	1.0000	0.8739	1.0000
课程代码	64	65	66	67	68	69	70	71	72
竞争力系数	1.0000	0.5040	1.0000	1.0371	1.0539	1.0178	1.0660	0.8207	1.1429

表 9 灰度预测和均值处理后的 2006 年 72 门课程书目的市场竞争力系数值

1.4 各课程书目年销售量的统计回归分析：

如前所述，我们认为，每门课程对应书目的年销售量应该与该书目的年需求量，该出版社的市场竞争力以及总社分配的该门课程的书号个数有关。在前面的讨论中，我们已经可以定量的表示 2001 至 2005 各年出版社每门课程书目的年需求量以及出版社各个课程的市场竞争力，而且，附件 3 和附件 4 又给出了各年总社为各个课程分配的书号数以及各个课程书目每一年的实际销售量，这样，对于每一年，我们都可以得到 A 出版社对应于 72 门课程的 72 组数据，每组数据都由年需求量，市场竞争力，书号分配个数和年销售量组成。综合考虑 5 年的情况，则可以得到 360 组数据。表 10 给出了这 360 组数据的数据格式：

年份	课程代码	年需求量 r_i	市场竞争力 c_i	分配书号数 x_i	年销售量 s_i
2001	1	0.0055	0.1802	10	1240
	2	0.0146	0.1278	10	1809

	71	0.0009	0.4030	3	444
	72	0.0013	0.8224	2	949
2002	1	0.0059	0.2232	11	1243
	2	0.0158	0.0886	11	1706

	71	0.0010	0.3980	1	443
	72	0.0014	0.7253	1	1182
2003	1	0.0071	0.2386	12	1850
	2	0.0190	0.1352	12	2681

	71	0.0012	0.3129	3	568
	72	0.0017	0.7528	1	1250
2004	1	0.0092	0.3129	11	2641
	2	0.0246	0.1261	12	3298

	71	0.0015	0.4595	2	759
	72	0.0022	0.9495	2	2038
2005	1	0.0114	0.2511	12	2692
	2	0.0304	0.1257	12	3927

	71	0.0019	0.6930	2	1209
	72	0.0027	1.0000	2	2527

表 10 A 出版社 5 年的年需求量, 市场竞争力, 书号分配个数和年销售量数据

我们这里对这 360 组数据应用统计回归的分析方法, 估计出年销售量 s 与年需求量 r , 市场竞争力 c 和书号分配个数 x 的函数关系, 即 $s = f(r, c, x)$ 。具体讨论如下:

首先大致地分析 s 与 r , c , x 的关系, 我们认为, s 的值随着 r , c 的增加是呈线性增长趋势的, 因此我们分别用两个线性方程来表示 s 和 r 以及 s 和 c 的线性关系:

$$s = \beta_0 + \beta_1 \cdot r + \varepsilon$$

$$s = \beta_0 + \beta_1 \cdot c + \varepsilon$$

对于 s 与 x 的关系, 我们认为, 在初始阶段, s 随 x 的增大有较大的增长趋势, 而当 x 增大到一定程度时, s 的增长趋势就逐渐减缓了, 因此我们用 $\frac{1}{2}$ 次幂函数方程来表示 s 和 x 的关系:

$$s = \beta_0 + \beta_1 \cdot x^{\frac{1}{2}} + \varepsilon$$

综合以上的分析, 将上面的 s 与 r , c , x 的关系方程结合起来, 就可以建立如下的统计回归方程:

$$s = \beta_0 + \beta_1 \cdot r + \beta_2 \cdot c + \beta_3 \cdot x^{\frac{1}{2}} + \varepsilon$$

式中, r , c , x 为回归变量, $\beta_0, \beta_1, \beta_2, \beta_3$ 为回归系数, ε 为随机误差。回归方程表示给定年需求量 r , 市场竞争力 c 和书号分配个数 x 时, 年销售量 s 的平均值。影响 s 的其他因素的作用都包含在随机误差 ε 中, 并且 ε 应大致服从均值为零的正态分布。由于在本题中, 当 r , c , x 均为 0 时, s 也为 0, 所以可知常数项 β_0 为 0, 故统计回归方程应为:

$$s = \beta_1 \cdot r + \beta_2 \cdot c + \beta_3 \cdot x^{\frac{1}{2}} + \varepsilon$$

我们这里直接利用 MATLAB 统计工具箱的命令 regress 对上面的统计回归方程进行求解, 使用格式为:

$$[b, bint, r, rint, stats] = \text{regress}(y, x, \alpha)$$

其中输入 y 为统计回归方程中 s 的数据 (360 维的向量), x 为对应于回归系数

$\beta=(\beta_1,\beta_2,\beta_3)$ 的数据矩阵 $[r \ c \ x^{\frac{1}{2}}]$ (360×3 矩阵), α 为置信水平 (缺省时 $\alpha=0.05$); 输出 b 为 β 的估计值, 常记作 $\hat{\beta}$, bint 为 b 的置信区间, r 为残差向量 $y-x \cdot \hat{\beta}$, rint 为 r 的置信区间, stats 为回归方程的检验统计量, 有 3 个值, 第 1 个是回归方程的决定系数 R^2 (R 是相关系数), 第 2 个是 F 统计量值, 第 3 个是与 F 统计量相对应的概率值 p 。

经过求解得到的统计回归方程的回归系数估计值及其置信区间 (置信水平 $\alpha=0.05$), 检验统计量 R^2 , F , p 的结果如表 11 所示:

参数	参数估计值	参数置信区间
β_1	430240	[413040, 447430]
β_2	11832	[9986.2, 13678]
β_3	-1697.6	[-2150.4, -1244.9]
$R^2=0.4558 \quad F=167.95 \quad p=0.0000$		

表 11 统计回归方程的参数求解结果

在表 11 中, $R^2=0.4558$ 表示因变量 y (年销售量 s) 的 45.58% 可由统计回归方程来确定。这说明由统计回归方程确定的年销售量 s 有 54.42% 的可能是错误的。这样的统计回归方程显然是不能使用的。为此, 我们在原有基础上改变统计回归方程的形式, 加入了 r 与 c 和 r 与 x 的交叉项, 修改后的统计回归方程为:

$$s = \beta_1 \cdot r + \beta_2 \cdot c + \beta_3 \cdot x^{\frac{1}{2}} + \beta_4 \cdot r \cdot c + \beta_5 \cdot r \cdot x + \varepsilon$$

同样应用 MATLAB 的 regress 命令对方程进行求解, 得到的结果如表 12 所示:

参数	参数估计值	参数置信区间
β_1	-2884.1	[-5466.4, -301.88]
β_2	-237.26	[-434.87, -39.653]
β_3	161.22	[114.47, 207.96]
β_4	961510	[956450, 966570]
β_5	331.27	[194.98, 467.57]
$R^2=0.99407 \quad F=17962 \quad p=0.0000$		

表 12 带交叉项的统计回归方程的参数求解结果

表 12 显示, $R^2=0.99407$ 表示因变量 y (年销售量 s) 的 99.47% 可由统计回归方程来确定, F 值远远超过 F 检验的临界值, p 远小于 α 。并且, 回归系数 ($\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$) 的置信区间都远离零点, 这说明回归变量 r, c, x 对年销售量 s 的影响是显著的。从上述统计特征我们可以看出, 统计回归方程很好的表示出了年销售量 s 与 r, c, x 的关系。因此, 我们将回归系数的估计值代入到方程中 (令 $\varepsilon=0$), 得到如下方程:

$$s = -2884.1 \cdot r - 237.26 \cdot c + 161.22 \cdot x^{\frac{1}{2}} + 961510 \cdot r \cdot c + 331.27 \cdot r \cdot x$$

我们把这一方程作为出版社各门学科年销售量 s 与该学科年需求量 r , 市场竞争力 c 和分配书号数 x 的函数关系。即:

$$s = f(r, c, x) = -2884.1 \cdot r - 237.26 \cdot c + 161.22 \cdot x^{\frac{1}{2}} + 961510 \cdot r \cdot c + 331.27 \cdot r \cdot x$$

2. 模型建立:

通过以上分析, 我们已经能够完整用出版社各课程书目的年需求量 r , 市场竞争力 c 和书号分配个数 x 表示该课程书目的年销售量 s , 进而应用总经济效益的表达式

$$Q = \sum_{i=1}^{72} p_i \cdot s_i \cdot b$$

得到出版社当年的经济效益。我们把一年经济效益作为模型的目标函数,

把总社为分设的每门课程 (共 72 门) 实际分配到的书号个数 x_i ($i=1,2,\dots,72$) 作为模型的决策变量。对于约束条件, 由附件 4 可知, 总社为分社的每门课程实际分配的书号数 x_i 要不大于分社为每门课程申请的书号数 a_i , 同时也要不小于分社申请书号数的一半; 又由题目可知, 总社每年供分配书号数的总量是一定的, 为此我们统计了 2001 年至 2005 年的书号数总量, 得到每年的书号数总量都是 500 个, 所以 2006 年的总社为分社分配的书号数总量也应为 500 个。至此, 我们便针对出版社的资源配置问题建立了如下的数学模型:

目标函数: $Q = \sum_{i=1}^{72} p_i \cdot s_i \cdot b$

决策变量: x_i ($i=1,2,\dots,72$)

约束条件: $s_i = -2884.1 \cdot r_i - 237.26 \cdot c_i + 161.22 \cdot x_i^{\frac{1}{2}} + 961510 \cdot r_i \cdot c_i + 331.27 \cdot r_i \cdot x_i$
($i=1,2,\dots,72$)

$$\frac{1}{2} a_i \leq x_i \leq a_i \quad (i=1,2,\dots,72) \quad (1)$$

$$\sum_{i=1}^{72} x_i = 500 \quad (2)$$

3. 模型求解:

3.1 LINGO 求解:

由于我们前面所建立的模型是一个非线性规划模型，其目标函数包含 $\frac{1}{2}$ 次项，因此我们应用 LINGO8 就对模型进行求解时，求解的速度较慢。最终求解出的总社为各课程分配书号数的最优方案如表 13 所示：

课程代码	1	2	3	4	5	6	7	8	9
分配书号数	9	9	4	4	4	8	6	3	8
课程代码	10	11	12	13	14	15	16	17	18
分配书号数	4	5	4	4	4	4	4	4	8
课程代码	19	20	21	22	23	24	25	26	27
分配书号数	6	4	6	19	52	4	12	17	6
课程代码	28	29	30	31	32	33	34	35	36
分配书号数	6	12	11	40	2	2	11	5	8
课程代码	37	38	39	40	41	42	43	44	45
分配书号数	3	6	5	6	2	10	8	10	5
课程代码	46	47	48	49	50	51	52	53	54
分配书号数	6	7	9	4	5	8	3	8	6
课程代码	55	56	57	58	59	60	61	62	63
分配书号数	2	3	4	4	2	5	4	6	4
课程代码	64	65	66	67	68	69	70	71	72
分配书号数	4	2	3	8	5	4	3	3	4

表 13 用 LINGO 求解出的总社为各课程分配书号数的最优方案

我们根据求解得到的分配方案计算出的出版社 2006 年的总销售额为 24596538 元。

3.2 贪心算法求解：

在前面的讨论中，我们是以出版社的总利润为标准，一次性的为各课程分配的书号数。这种分配方法原理简单，较为直接，但模型的求解难度较大，且求解速度慢。为此，我们提出了一种快速的贪心算法，以每个课程增加一个书号所多出的利润（称作该课程的强势度）为标准，逐个地为各课程分配书号。具体的算法如下：

Step 1: 由于题目中要求每门课程都至少得到分社为其申请书号数的一半，因此我们首先为 72 门课程的每一门都分配其申请书号数一半的书号数；

Step 2: 假设为每一门课程分别再分配一个书号数，如果此时某门课程分得的书号数已经大于其申请的书号数，则令该门课程的强势度为 0；否则就把增加这个书号后每门课程增加的利润当作该课程当前的强势度；

Step 3: 比较所有课程当前的强势度，找出强势度最大的课程，并为该课程分配一个书号。

Step 4: 若总共分配的书号数未达到 500 个，则转到 Step 2; 若总共分配的书号数已经达到 500 个，则终止分配。

由上述算法得到的算法流程图如图 3 所示：

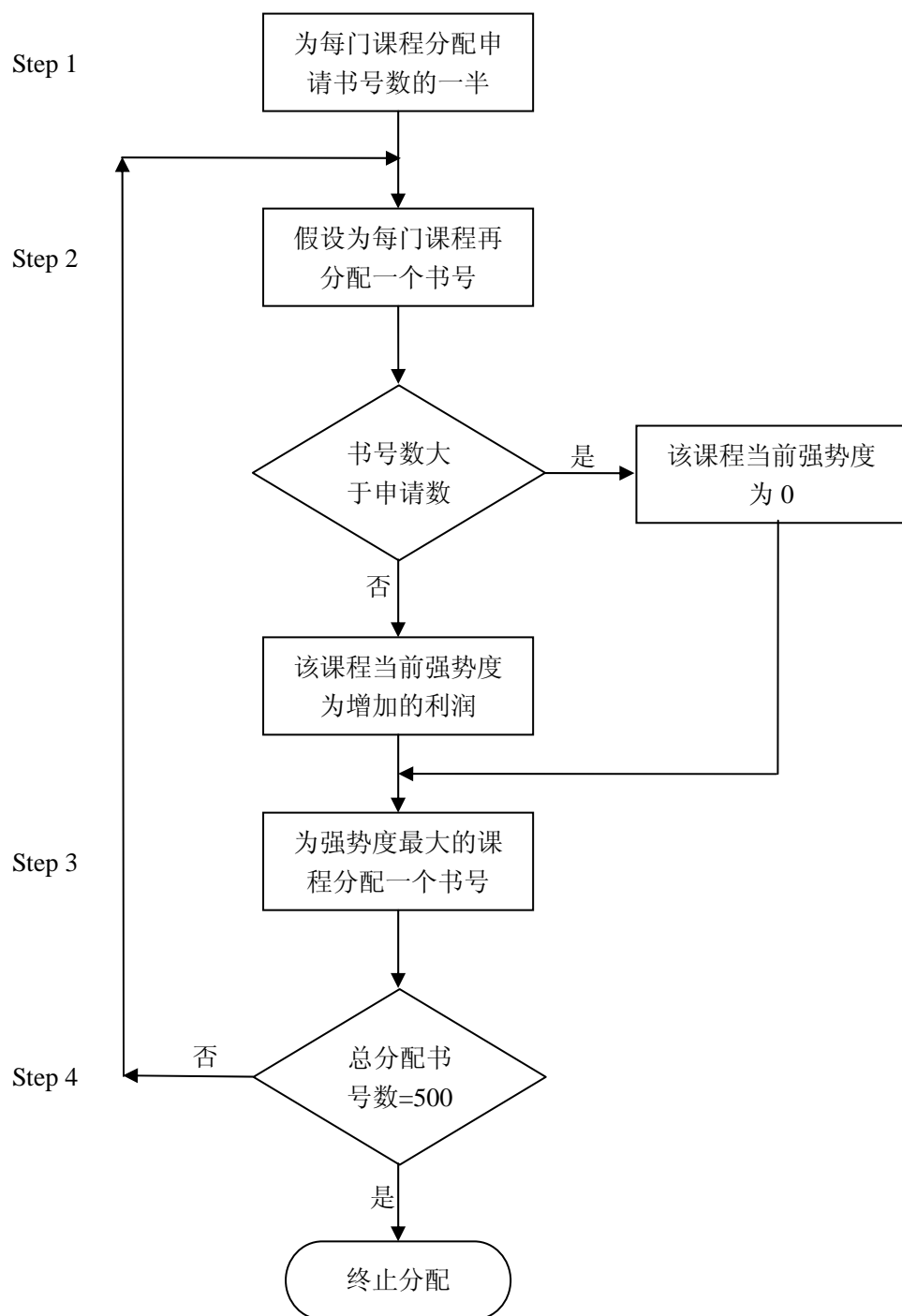


图3 贪心算法流程图

表14 列出了由贪心算法求解出的总社为各课程分配书号数的最优方案:

课程代码	1	2	3	4	5	6	7	8	9
分配书号数	9	9	4	4	4	8	6	3	8
课程代码	10	11	12	13	14	15	16	17	18
分配书号数	4	5	4	4	4	4	4	4	8
课程代码	19	20	21	22	23	24	25	26	27
分配书号数	6	4	6	19	52	4	12	17	6
课程代码	28	29	30	31	32	33	34	35	36

分配书号数	6	12	11	40	2	2	11	5	8
课程代码	37	38	39	40	41	42	43	44	45
分配书号数	3	6	5	6	2	10	8	10	5
课程代码	46	47	48	49	50	51	52	53	54
分配书号数	6	7	9	4	5	8	3	8	6
课程代码	55	56	57	58	59	60	61	62	63
分配书号数	2	3	4	4	2	5	4	6	4
课程代码	64	65	66	67	68	69	70	71	72
分配书号数	4	2	3	8	5	4	3	3	4

表 14 用贪心算法求解出的总社为各课程分配书号数的最优方案

我们根据求解得到的分配方案计算出的出版社 2006 年的总销售额为 24596538 元。

比较两种算法得到的结果，我们发现，用 LINGO 求解和用贪心算法求解得到的总社为各课程分配书号数的最优方案是完全一样的。由于用 LINGO 求解得到的结果是模型的全局最优解，这说明，我们提出的贪心算法不仅可以极大地缩短运算速度，并且其得到的结果也是准确的，最优的。同时，两种算法得到相同的结果也说明我们求得的书号分配方案是最优的。

4. 结果分析：

4.1 2006 年强势产品定位：

如题目所述，总社一般以增加强势产品支持力度的原则优化资源配置。这里，我们用每个课程所产生的年销售额来衡量该课程的强势程度。在前面的讨论中，我们已经得到了总社最优的资源配置方案，因此我们可以由该分配方案求出每个课程所产生的年经济效益，并以此比较各课程的强势程度。

在表 15 中，我们列举出了强势程度最高的前 10 门课程，并给出了相应的书号申请数，实际分配数，年需求量系数，市场竞争力系数，销售量和销售额数据：

课程代码	申请书号数	分配书号数	需求量系数	竞争力系数	销售量	销售额
23	52	52	0.37517	0.99974	366942	9107433
26	34	17	0.07732	0.88181	66259	1542510
42	10	10	0.12736	0.40473	50031	939364.8
31	40	40	0.27674	0.085882	26720	920376.8
44	10	10	0.12598	0.40614	49664	831493.9
45	6	5	0.12663	0.49431	60249	809783
30	12	11	0.053837	0.54402	28608	733983.4
43	8	8	0.072	0.37167	26098	692940.3
46	8	6	0.11814	0.40715	46442	685059.5
47	12	6	0.088673	0.439	37645	656945.3

表 15 2006 年强势程度最高的前 10 门课程的相关数据

4.2 2006 年各分社经济效益比较：

在前面的分析中，我们都是以各个课程为单位进行讨论，这里我们要以各分社为单位，比较它们在 2006 年的经济效益。在得到了 2006 年书号分配方案后，我们把各分社包含课程的经济效益求和，就得到了各分社的经济效益。图 4 比较了各分社的经济效益在总经济效益中所占的份额：

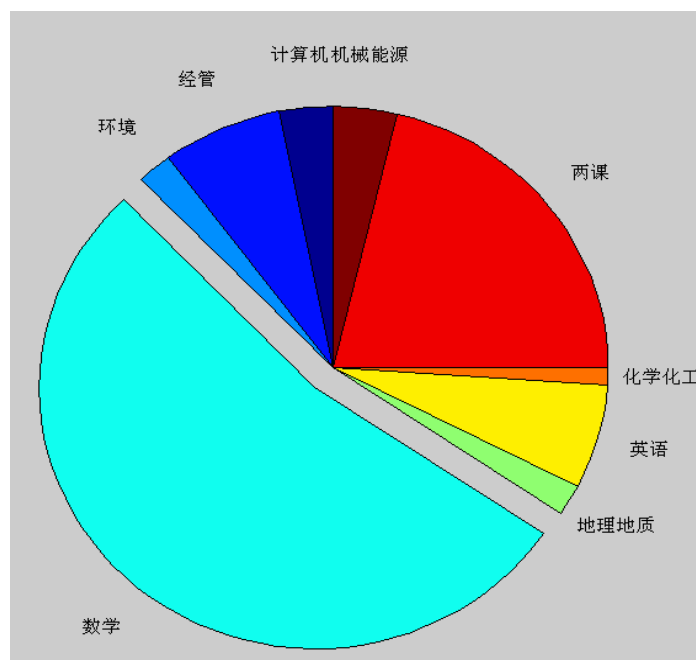


图 4 各分社经济效益比较

由上图可以看出，2006 年数学类分社的经济效益在总经济效益中占了绝大比重。

4.3 2006 年总销售量的准确度分析：

在前面的讨论中，我们已经求得 2006 年各门学科分到的书号数，并可以由此计算出各学科的年销售量，进而得到出版社在 2006 年的总销售量。而在附件三中，我们已知了 2001 年至 2005 年的总销售量，可以预测出 2006 年的总销售量。在图 5 中，我们把这两种方法得到的 2006 年的总销售量进行了比较：

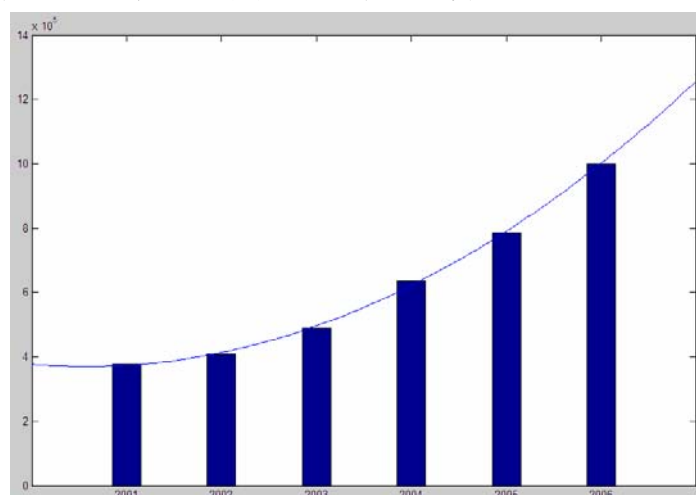


图 5 2006 年总销售量的准确度比较

由图 5 可以看出，通过模型计算出 2006 年的总销售量与通过预测得到的总销售量十分接近，这也证明我们模型的建立和求解是较为准确的。

4.4 2006 年总销售额的准确度分析：

与总销售量的准确度分析相同，我们对用计算方法和预测方法得到的总销售额进行了比较，如图 6 所示：

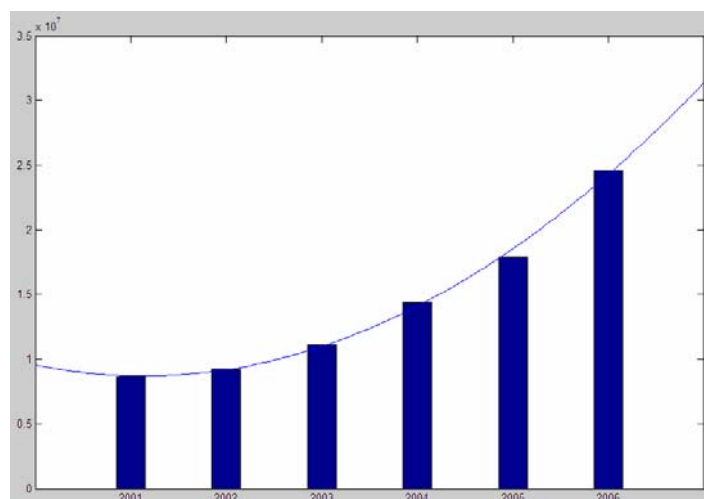


图 6 2006 年总销售额的准确度比较

与图 5 相似，在图 6 中，通过模型计算出 2006 年的总销售额与通过预测得到的总销售额也十分接近，这也证明我们模型的建立和求解是较为准确的。

模型评价

在模型的建立过程中，我们对题目已知数据进行了有效地统计分析，合理地提出了需求量系数和市场竞争力系数的概念，在已知 2001 年至 2005 年的这两类系数值的情况下，我们应用灰度预测的方法准确地预测出了 2006 年的系数值；同时，我们合理地应用统计回归的方法，较为准确地拟合出了销售量与需求量，市场竞争力和分配书号数的函数关系。在模型求解过程中，我们不仅用 LINGO 给出了全局最优解，还引入了强势度的概念设计了快速的贪心算法对模型进行了高效的求解，并得到了满意的结果。最后，我们对求解的结果进行了全面准确地分析。

同时，由于我们在模型建立过程中引入的需求量系数，市场竞争力系数以及强势度等主观定义的参量，这使得模型的求解结果受主观因素的影响较大，我们也未对这一影响做出准确的灵敏度分析。

参考文献

- [1] 姜启源，谢金星，叶俊，《数学模型》，北京：高等教育出版社，2003.1。
- [2] 《现代数学手册》编纂委员会，《现代数学手册•经济数学卷》，湖北武汉，华中科技大学出版社，2001.1
- [3] 宋兆基，徐流美等，《MATLAB 6.5 在科学计算中的应用》，北京：清华大学出版社，2005.1。
- [4] 韩中庚. 数学建模方法及其应用. 北京高等教育出版社. 2005