

试卷综合评价分析的数学模型

摘要

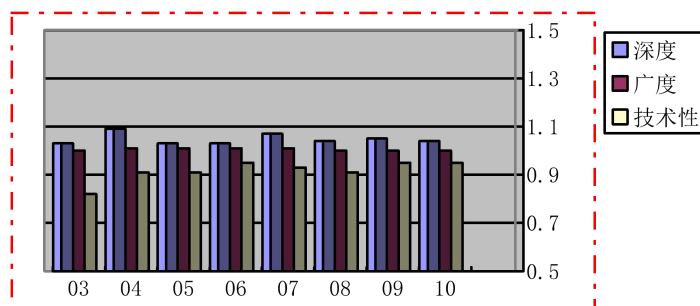
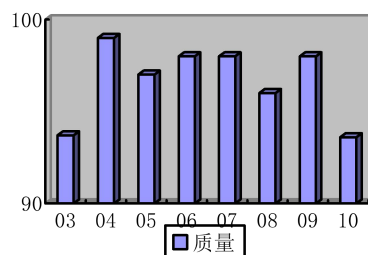
本文解决的是试卷综合评价问题，基于对试题质量标准的理解，结合综合评优、图形分析等方法，建立分层模型、灰色系统模型，对 2006 年至 2009 年数学一考研试题进行评价，对 2003 年至 2009 年数学一考研试题的命题规律进行分析。

对于问题一：首先结合大纲和试题确定考研数学一的深度值、广度值和技术性指标，然后根据专家评定、数据统计、小组预测给出标准试卷的各项标准值，确立历年试卷与标准试卷的数据对比关系，后建立分层模型，分别计算出试卷的深度、广度、技术性以及综合质量。结果如下表：

评价指标 年份	相对于标准试卷的权向量			综合评分
	深度	广度	技术性	
2006	1.03	1.01	0.95	98
2007	1.07	1.00	0.93	98
2008	1.04	1.00	0.91	96
2009	1.05	1.00	0.95	98

对各项指标分析：深度均偏高但幅度很小；广度基本吻合标准；技术性均偏低但波动不大；综合评分均高，保持在一恒定范围。

对于问题二：基于问题一方法、结果的进一步应用，求得 2003 年到 2009 年的各项指标，后用灰色系统模型计算出 2010 的深度值、广值和技术性指标以及综合质量，从而运用 MATLAB 软件分别绘制 03 年到 10 年的质量变化图和深度权向量、广度权向量、技术性权向量的变化图。其变化如图：



权向量图

关键字： 考研数学一 分层模型 灰色系统模型 图形分析

1、问题重述

各种考试对于试卷质量的要求非常严格。如何保证试卷的质量，确保考试的公正、高效、合理，是命题人和考生都特别关注的事情。不同的考试有不同的要求，考试的方式和内容也有所不同，而考试的目的如水平考试、选拔考试、竞赛考试等对试题的要求也不尽相同，因此，评价试卷质量的标准也各不相同。考研是一项选拔性考试，科目数学一又有别于其他科目。结合考试大纲和试题，运用定性和定量相结合的方法，确立试卷的质量标准，对历年考研数学一试卷进行评价，并分析命题的规律性，是总结以往试卷质量，改善今后试卷质量的重要措施。基于以上分析和探讨，给出历年考研数学一的试题，在参考考试大纲的基础上，提出两个具体问题：1. 结合大纲要求，对2006年到2009年的考研数学一试题进行定量分析评价；2. 总结以往的考试规律，对2003年到2009年考研数学命题的规律性进行分析。

2、问题分析

试卷质量的综合评价和命题的规律性分析分别属于评价问题和预测问题。鉴于试卷评价的层次性和复杂性，我们选取层次模型加以分析，而对于试卷命题的规律性分析，我们运用灰色系统模型和数学软件加以解决。就考研数学一试卷本身来看，知识面有高等数学、线性代数、概率统计三个部分组成，难易程度有难、中、易三个层次，题型有选择题、填空题、综合题三种，而综合性和知识覆盖面也是重要的隐含信息。从而，我们对所要解决的两个问题具体分析如下：

针对问题一：试卷质量的评价标准受到多方面因素的影响，具体总结为深度、广度和技术性三个层面，而在试卷的命题过程中，三者的权重同等重要。其中深度与每套试卷中难、中、易三类题的分值大小相关；广度与高等数学、线性代数、概率统计所占分值相关，技术性则包括覆盖面、综合性、题型比例，其中题型比例由选择题、填空题、综合题所占分值决定。在此基础上，通过试卷与大纲的综合分析，对各影响因素给出决定比值，从而依据各因素的相互影响因素建立分层模型，综合得出2006年到2009年试卷的综合评定值。通过所得值与标准值的比较，确定试卷质量。

针对问题二：首先进行定量分析：在问题一中，计算出2006年到2009年考研数学一试卷的深度值、广度值、技术性指标以及综合质量值，通过进一步计算，我们得到2003年到2009年的深度值、广度值、技术性指标以及综合质量值。运用灰色模型对2010年深度值、广度值、技术性指标分别加以计算、预测，对质量结合问题一进行计算。然后进行定性分析：根据2003年到2010年各项值，运用数理统计方法找出其变化规律并用MATLAB软件作图，分析规律。

3、模型假设

- 1) 假定自2003年起考试大纲没有大的变动；
- 2) 标准试卷的各项指标为最优；
- 3) 学生能力大致符合正态分布；
- 4) 专家对问题的难度估计无误；
- 5) 各个指标的定权无误；
- 6) 深度、广度和技术性指标能被文中要素全面反映。

4、符号说明

序号	符号	符号说明
1	$X^{(0)}$	原始序列
2	$\hat{X}^{(0)}$	为相应的模拟误差序列
3	$X^{(1)}$	将原始数据累加生成德序列
4	$X^{(1)}(K)$	将原始数据作累加生成, 第 K 个累加生成成为元素
5	$X^{(0)}(K)$	表示深度权向量中第 K 套试卷与标准试卷的比值
6	$\hat{X}^{(1)}$	表示相应的模拟序列
7	$\hat{x}^{(1)}(k)$	为相应的模拟值
8	$Z^{(1)}$	对 $X^{(1)}$ 作紧邻均值生成的序列
9	\mathcal{E}	X 与 \hat{X} 的灰色关联度
10	z_1, z_2, z_3	分别表示试卷的深度、广度和技术性
11	y_{11}, y_{12}, y_{13}	分别表示试卷中的基本题、中等题、难题
12	$W_{z_i}(y_i)$	y_i 对 z_i 的权向量
13	y_{2i} ($i=1, 2, 3$)	依次表示高数、线性代数、概率
14	$x_0, x_1 \dots x_7$	分别表示标准试卷, 03 年试卷... 09 年试卷
15	$W_{y_{1i}}(X)$	X 对 y_{1i} ($i=1, 2, 3$) 的权向量
16	$W_{y_{2i}}(X)$	X 对 y_{2i} ($i=1, 2, 3$) 的权向量
17	$W_{y_{3i}}(X)$	X 对 y_{3i} ($i=1, 2, 3$) 的权向量
18	$W_{z_i}(X)$	X 对 Z_i 的权向量
19	$W(X)$	试卷质量的权向量

5、模型的建立与求解

5.1 模型准备

参照大纲、结合试题，对 2003 年到 2009 年考研数学一试卷的深度值、广度值和技术性指标分别统计如下：

表 5.1：零三年到零九年考研数学的深度值

深度 年份	难 (分值)	中 (分值)	易 (分值)	总分
2003	32	69	49	150
2004	36	73	41	150
2005	31	70	49	150
2006	35	72	43	150
2007	34	75	41	150
2008	33	74	43	150
2009	34	76	40	150

分析： 1. 难、中、易三类题在历年试卷中所占分值基本稳定
2. 难、中、易三类题的比例基本符合 34：72：44

表 5.2：零三年到零九年考研数学的广度值

广度 年份	高等数学 (分值)	线性代数 (分值)	概率统计 (分值)	总分
2003	90	30	30	150
2004	90	30	30	150
2005	90	30	30	150
2006	90	30	30	150
2007	82	34	34	150
2008	82	34	34	150
2009	82	34	34	150

分析： 1. 高等数学、线性代数、概率统计在历年中占分值基本稳定
2. 高等数学所占比重最大
3. 三者所占分值比例基本为 68：32：32

表 5.3: 零三年到零九年考研数学的技术性指标

年份 \ 技术性	各项技术评分					
	覆盖面 (评分)	综合性 (评分)	题型比例			
			选择题 (分值)	填空题 (分值)	综合题 (分值)	评分
2003	8	9	24	24	102	8
2004	9	5	32	24	94	10
2005	8	10	32	24	94	10
2006	9	9	32	24	94	10
2007	9	9	40	24	86	9
2008	9	8	32	24	94	10
2009	9	9	32	24	94	10

说明: 1. 覆盖面的满分值为 10
 2. 综合性的满分值为 10
 3. 题型比例的综合评分满分为 10

分析: 技术性指标中覆盖面、综合性、题型比例各项值均稳定

5.2 模型建立的建立、求解

5.2.1 问题一模型:

5.2.1.1 建立分层模型结构简图如下:

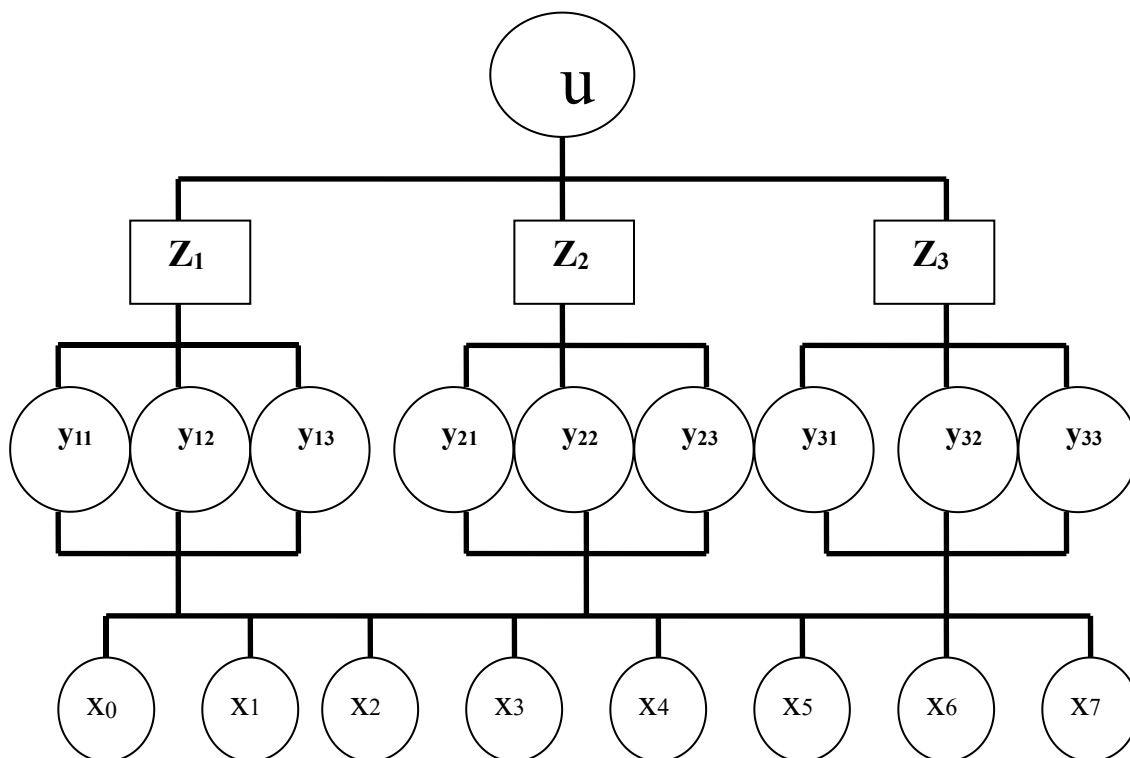


图 1 模型简图

5.2.1.2 建立模型

5.2.1.2.1 确立标准深度

设试卷的深度、广度、技术性分别为 z_1 、 z_2 、 z_3 ，它们对评价试卷的质量是同等重要的。因此对试卷评定的权向量为 $w(z) = (0.33, 0.33, 0.33)^T$ 。

试卷中的基本题、中等题、难题分别表示为 y_{11} 、 y_{12} 、 y_{13} ，则令 $y_1 = \{y_{11}, y_{12}, y_{13}\}$ ，对于一套标准试卷来说，比较适宜的难度比例应该是 30: 50: 20。 y_1 对 z_1 的成对比较矩阵为 A， y_1 对 z_1 的权向量为 $W_{z_1}(y_1)$ ，则：

$$A = \begin{bmatrix} 1.00 & 0.6 & 1.5 \\ 1.67 & 1 & 2.5 \\ 0.67 & 0.4 & 1 \end{bmatrix}$$

将 A 按行求和取算术平均值，再归一化得 $W_{z_1}(y_1) = (0.2, 0.5, 0.3)^T$

5.2.1.2.2 确立标准广度

在对试题的广度进行分析时，我们分别从高数、线性代数和概率三个方面的出题情况进行考虑。

y_{2i} ($i=1,2,3$) 1、2、3 依次表示高数、线性代数、概率

根据最近大纲的规定，此三部分的比例应为 56: 22: 22 比较恰当。

$$W_{z_2}(y_2) = (0.56, 0.22, 0.22)^T$$

5.2.1.2.3 确立标准技术性指标

在技术性方面的三个指标中，依据它们对试卷质量的的重要性对它们进行定权为覆盖面：题类：综合试题指标=5: 3: 4。

$$W_{\bar{y}_3}(y_3) = (0.41, 0.26, 0.33)^T$$

5.2.1.2.4 确立历年试卷与标准试卷的数据对比关系

从 x_1 到 x_7 表示从 2003 年到 2009 年, x_0 表示标准试卷。它们在各项评定标准中的得分或者数据应该如下表所示:

表 5.4 2003 年到 2009 年考研数学一在各项评定标准中的数据或得分

评定标准 试卷(年)		深度			广度			技术性		
		易	中	难	高等数学	线性代数	概率统计	覆盖面	综合性	题型比例
		y_{11}	y_{12}	y_{13}	y_{21}	y_{22}	y_{23}	y_{31}	y_{32}	y_{33}
标准	x_0	30	75	45	84	33	33	10	10	10
2003	x_1	32	69	49	90	30	30	8	8	9
2004	x_2	36	73	41	90	30	30	9	10	8
2005	x_3	31	70	49	90	30	30	8	10	10
2006	x_4	35	72	43	90	30	30	9	10	9
2007	x_5	34	75	41	82	34	34	9	9	9
2008	x_6	33	74	43	82	34	34	9	10	8
2009	x_7	34	76	40	82	34	34	9	10	9

- 说明:**
1. 标准深度(易、中、难比例)为 30: 75: 45
 2. 标准广度(高数、现代、概率的比例)为 84: 33: 33
 3. 标准技术性中覆盖面、综合性、题型比例均为 10

5.2.1.3 模型求解

令 $X = \{ x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7 \}$

先求出 X 对 y_{1i} ($i=1, 2, 3$) 的权向量

$$W_{y_{1i}}(X) = \begin{bmatrix} 0.11 & 0.13 & 0.13 \\ 0.12 & 0.12 & 0.14 \\ 0.14 & 0.13 & 0.12 \\ 0.12 & 0.12 & 0.14 \\ 0.13 & 0.12 & 0.12 \\ 0.13 & 0.13 & 0.12 \\ 0.12 & 0.13 & 0.12 \\ 0.13 & 0.13 & 0.11 \end{bmatrix}$$

再求出 X 对 y_{2i} ($i=1, 2, 3$) 的权向量

$$= W_{y_{2i}}(X) \begin{bmatrix} 0.12 & 0.13 & 0.13 \\ 0.13 & 0.12 & 0.12 \\ 0.13 & 0.12 & 0.12 \\ 0.13 & 0.12 & 0.12 \\ 0.13 & 0.12 & 0.12 \\ 0.12 & 0.13 & 0.13 \\ 0.12 & 0.13 & 0.13 \\ 0.12 & 0.13 & 0.13 \end{bmatrix}$$

最后求出 X 对 y_{3i} ($i=1, 2, 3$) 的权向量

$$W_{y_{3i}}(X) = \begin{bmatrix} 0.14 & 0.13 & 0.14 \\ 0.11 & 0.10 & 0.13 \\ 0.13 & 0.13 & 0.11 \\ 0.11 & 0.13 & 0.14 \\ 0.13 & 0.13 & 0.13 \\ 0.13 & 0.12 & 0.13 \\ 0.13 & 0.13 & 0.11 \\ 0.13 & 0.13 & 0.13 \end{bmatrix}$$

设 X 对 z_1 的权向量为 $W_{z_1}(X)$ ，则 $W_{z_1}(X) = (W_{y_{1i}}(X), W_{y_{2i}}(X), W_{y_{3i}}(X)) W_{z_1}(y_i)$ 。
经计算得：

$$W_{z_1}(X) = (0.34, 0.33, 0.33)^T;$$

$$W_{z_2}(X) = (0.34, 0.33, 0.33)^T;$$

$$W_{z_3}(X) = (0.5, 0.3, 0.4)^T。$$

X 对试卷质量的权向量 $W(X) = (W_{y_{1i}}(X), W_{y_{2i}}(X), W_{y_{3i}}(X)) W(Z)$

$$W(X) = \begin{pmatrix} 0.1240 & 0.1244 & 0.1367 \\ 0.1240 & 0.1256 & 0.1117 \\ 0.1310 & 0.1256 & 0.1250 \\ 0.1240 & 0.1256 & 0.1241 \\ 0.1230 & 0.1256 & 0.1300 \\ 0.1280 & 0.1244 & 0.1267 \\ 0.1250 & 0.1244 & 0.1250 \\ 0.1260 & 0.1244 & 0.1300 \end{pmatrix} \begin{pmatrix} 0.33 \\ 0.33 \\ 0.33 \end{pmatrix}$$

$$= (0.1284 \quad 0.1204 \quad 0.1272 \quad 0.1246 \quad 0.1262 \quad 0.1264 \quad 0.1248 \quad 0.1268)^T$$

得出此向量后即得出了模型评价的综合指标。根据这个指标数据的比例关系可以得出从 2006 年到 2009 年这四年间考研数学题目的质量得分：

$$W(x_i) \text{ 分别为 } 98、98、96、98 \quad (i = 4, 5, 6, 7)$$

对于一份试卷的质量来说，我们不能只看综合指标，由于三个因素间可能会互补，因此我们分别看三个指标：

对于深度来说所得的深度的权向量为：

$$(1 \quad 1.03 \quad 1.09 \quad 1.03 \quad 1.03 \quad 1.07 \quad 1.04 \quad 1.05)^T ;$$

对于广度来说所得的权向量为：

$$(1 \quad 1.01 \quad 1.01 \quad 1.01 \quad 1.01 \quad 1 \quad 1 \quad 1)^T ;$$

对于技术性来说所得的权向量为：

$$(1 \quad 0.82 \quad 0.91 \quad 0.91 \quad 0.95 \quad 0.93 \quad 0.91 \quad 0.95)^T .$$

5.2.2 问题二模型

定量分析：

5.2.2.1 运用灰色系统模型对 2010 年考研数学一的深度权向量进行预测

设 $X^{(0)}(k)$ 表示深度权向量中第 k 套试卷与标准试卷的比值。则可得原始序列：

$$\begin{aligned} X^{(0)} &= \{X^{(0)}(1), X^{(0)}(2), X^{(0)}(3), X^{(0)}(4), X^{(0)}(5), X^{(0)}(6), X^{(0)}(7)\} \\ &= \{1.033, 1.092, 1.033, 1.025, 1.067, 1.042, 1.050\} \end{aligned}$$

$$\begin{aligned} X^{(1)} &= \{X^{(1)}(1), X^{(1)}(2), X^{(1)}(3), X^{(1)}(4), X^{(1)}(5), X^{(1)}(6), X^{(1)}(7)\} \\ &= \{1.033, 2.125, 3.158, 4.183, 5.250, 6.292, 7.342\} \end{aligned}$$

对 $X^{(1)}$ 作紧邻均值生成，令

$$Z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1)$$

$$\begin{aligned} Z^{(1)} &= \{Z^{(1)}(1), Z^{(1)}(2), Z^{(1)}(3), Z^{(1)}(4), Z^{(1)}(5), Z^{(1)}(6), Z^{(1)}(7)\} \\ &= \{1.033, 1.579, 2.6415, 3.6705, 4.7165, 5.771, 6.817\} \end{aligned}$$

于是,

$$B = \begin{bmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(1)}(3) & 1 \\ -Z^{(1)}(4) & 1 \\ -Z^{(1)}(5) & 1 \\ -Z^{(1)}(6) & 1 \\ -Z^{(1)}(7) & 1 \end{bmatrix} = \begin{bmatrix} -1.579 & 1 \\ -2.6415 & 1 \\ -3.6705 & 1 \\ -4.7165 & 1 \\ -5.771 & 1 \\ -6.817 & 1 \end{bmatrix}, Y = \begin{bmatrix} X^{(0)}(2) \\ X^{(0)}(3) \\ X^{(0)}(4) \\ X^{(0)}(5) \\ X^{(0)}(6) \\ X^{(0)}(7) \end{bmatrix} = \begin{bmatrix} 1.092 \\ 1.033 \\ 1.025 \\ 1.067 \\ 1.042 \\ 1.050 \end{bmatrix}$$

$$B^T B = \begin{bmatrix} 124.9646 & -25.1955 \\ -25.1955 & 6.0000 \end{bmatrix}$$

$$(B^T B)^{-1} = \begin{bmatrix} 0.0522 & 0.2191 \\ 0.2191 & 1.0869 \end{bmatrix}$$

$$B^T Y = \begin{pmatrix} -26.4189 \\ 6.3090 \end{pmatrix}$$

$$\hat{a} = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y = \begin{pmatrix} 0.0039 \\ 1.0677 \end{pmatrix}$$

白化方程为: $\frac{dx^{(1)}}{dt} - ax^{(1)} = 1.0677$

及时间响应公式:

$$\begin{aligned} \hat{x}^{(1)}(t) &= (x^{(1)}(0) - \frac{b}{a})e^{-at} + \frac{b}{a} \\ &= -272.736e^{-0.0039k} + 273.769 \end{aligned}$$

求 $x^{(1)}$ 的模拟值

$$\begin{aligned} \hat{X}^{(1)} &= \{\hat{x}^{(1)}(1), \hat{x}^{(1)}(2), \dots, \hat{x}^{(1)}(7)\} \\ &= \{1.033, 2.0946, 3.1521, 4.2054, 5.2547, 6.2998, 7.3409\} \end{aligned}$$

还原出 $x^{(0)}$ 的模拟值由

$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$ 得

$$\begin{aligned} \hat{X}^{(0)} &= \{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \hat{x}^{(0)}(3), \hat{x}^{(0)}(4), \hat{x}^{(0)}(5), \hat{x}^{(0)}(6), \hat{x}^{(0)}(7)\} \\ &= \{1.033, 1.0616, 1.0578, 1.0533, 1.0493, 1.0451, 1.0411\} \end{aligned}$$

计算 x 与 \hat{x} 的灰色关联度：

$$|S| = \left| \sum_{k=2}^6 (x(k) - x(1)) + \frac{1}{2}(x(7) - x(1)) \right|$$

$$= 0.1025$$

$$|\hat{S}| = \left| \sum_{k=2}^6 (\hat{x}(k) - \hat{x}(1)) + \frac{1}{2}(\hat{x}(7) - \hat{x}(1)) \right|$$

$$= 0.10615$$

$$|\hat{S} - S| = \left| \sum_{k=2}^6 [(x(k) - x(1)) - (\hat{x}(k) - \hat{x}(1))] + \frac{1}{2}[(x(7) - x(1)) - (\hat{x}(7) - \hat{x}(1))] \right|$$

$$= 0.00365$$

$$\varepsilon = \frac{1 + |S| + |\hat{S}|}{1 + |S| + |\hat{S}| + |\hat{S} - S|} = \frac{1 + 0.1025 + 0.10615}{1 + 0.1025 + 0.10615 + 0.00365} = \frac{1.20865}{1.2123}$$

$$= 0.997 > 0.90$$

精度为一级，关联度为一级，可用

$$\begin{cases} \hat{x}^{(1)}(k+1) = -272.736e^{-0.0039k} + 273.769 \\ \hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \end{cases}$$

进行预测

$$\hat{X}^{(1)} = \{\hat{x}^{(1)}(8)\} = 8.378$$

$$\hat{X}^{(0)} = \{\hat{x}^{(0)}(8)\} = \hat{x}^{(1)}(8) - \hat{x}^{(1)}(7)$$

$$= 1.0371$$

所以可知 2010 年考研数学一的深度权向量为 1.0371。

5.2.2.1 广度权向量、技术性权向量以及质量的计算

运用问题二中灰色模型同理可推算出 2010 年考研数学一的广度权向量、技术性权向量分别为：1.04、1.00。

运用问题一分层模型可以计算出 2010 年考研数学一的质量值为 93.6

则 2003 年到 2009 年深度权向量、广度权向量、技术性权向量以及质量的指标或分值如下表：

表 5.5 :2003 年到 2010 年考研数学一各项评价结果

评价指标 年份	相对于标准试卷的权向量			综合评分
	深度	广度	技术性	
2003	1.03	1.01	0.82	93.7
2004	1.09	1.01	0.91	99
2005	1.03	1.01	0.91	97
2006	1.03	1.01	0.95	98
2007	1.07	1.00	0.93	98
2008	1.04	1.00	0.91	96
2009	1.05	1.00	0.95	98
2010	1.04	1.00	0.94	93.6

定性分析:

运用 MATLAB 软件对 2003 年到 2010 年考研数学一的深度权向量、广度权向量、技术性权向量进行整体分析, 从而得出权向量变化规律图:

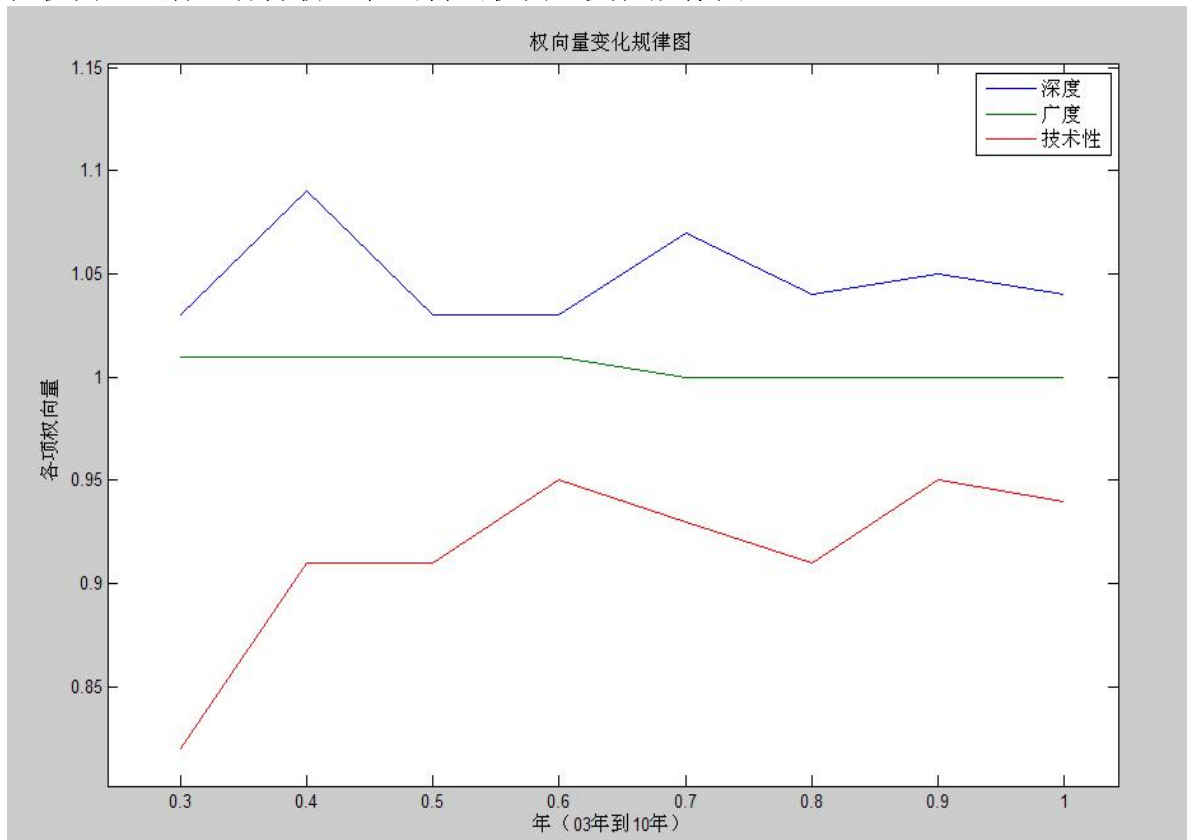


图 2 权向量变化规律图

质量变化图:

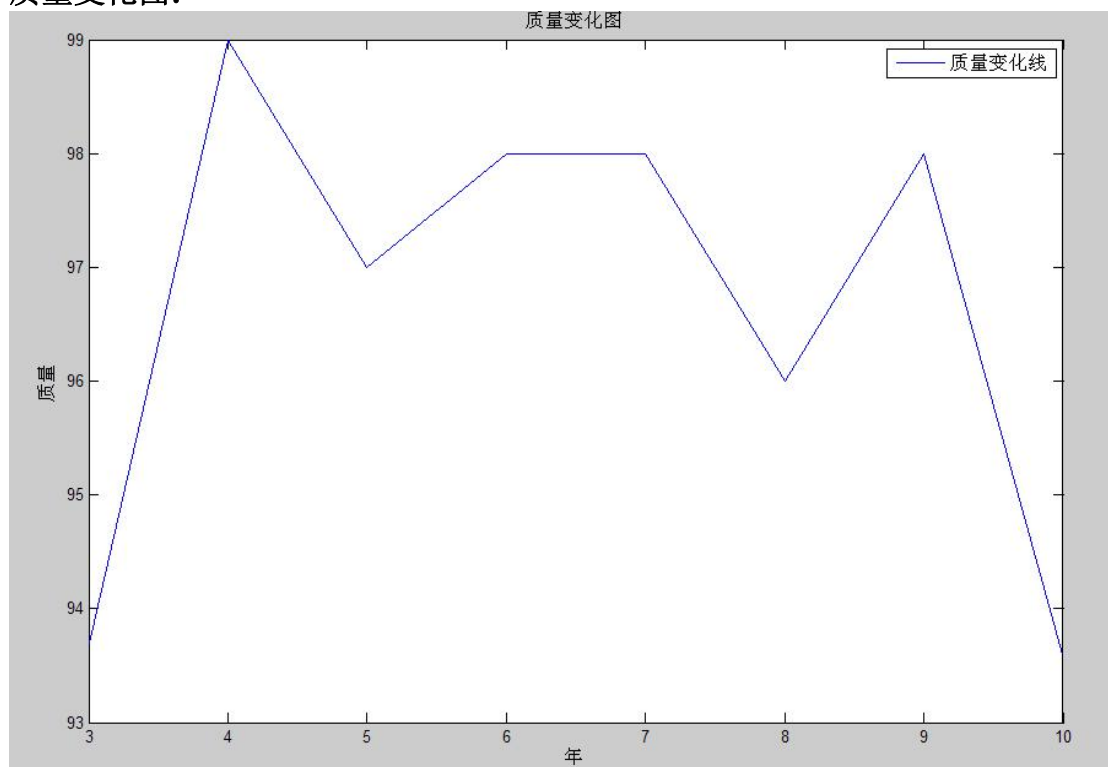


图 3 质量变化图

6、模型结果的分析

6.1 模型一

对 2006 年到 2009 年考研数学一质量分析互发现，几乎保持在同一水平，2008 年稍略于其他三年。说明考研试卷的整体质量高。

对于三个指标，我们不需要通过打分来具体分析，只需要通过数值来定量分析即可。

我们统计了七年的数据，再次我们只需要分析最后一年的数据，各个权向量的第一个元素是标准试卷的相应指标值，将此值归一化以便能更明晰地分析各个指标。从而从各个指标的变化来分析试卷。

首先，深度指标是大于一的，说明历年出题的难度略微偏难，不过鉴于考研属于选拔性考试而指标数值也只是略微偏大，所以在可以接受的范围内。在这四年内，2006 年的试题最接近于标准值，在难度的控制上比其它三年略胜一筹，而 2007 年的难度较大。

其次，广度的契合度都很好，基本上没有什么偏差，说明出题人对于大纲出题范围和各个学科出题比例的把握都相当到位。

最后对于技术性来说，虽然各个偏差都在一定的范围内可以接受，但是各年还是有明显差别 2006 年和 2009 年的技术性较好，而其他两年的技术性略差。

6.2 模型二

6.2.1 定量计算结果分析

本题在对 10 年试卷预测时，采用了深度、广度和技术指标权向量中 03 年到 09 年与标准试卷的比值作为各项指标的度量标准，将标准试卷中各项指标标准取 1。利用灰

度预测，得到 10 年的深度、广度和技术指标值分别为 1.0371、0.9959、0.9366，各项指标灰色关联度分别为 0.997、0.9998 和 0.994，均大于 0.9，说明精度为一级，关联度为一级，估计值可以用。根据估计可知，10 年试卷的各项指标和标准试卷的差别很小，说明 10 年的试卷和标准试卷很接近。

6.2.2 对定性分析的结果图 2（权向量变化规律图）和图 3（质量变化图）进行分析 图 2

根据历年试题的各项指标的示意图可以看出，广度的分布基本上没有变化，这说明出题人对于高考各部分的出题分量的把握是相当恰当的。

对于试题的深度也即试题的难度来说，把握起来还是有一定的困难的，所以会出现有一定规律的波动现象，这是因为命题专家在参考了上一年的试题答题情况和难度情况后，都会适当的调整试题的难度以便于试题更符合大纲的要求，更能反映学生的真实水平。

对于试题的技术含量来说，整体上呈现上升的趋势，而在近几年则维持在比较高的水平上起伏不大了，这是因为随着教育的改革，我国教育体制不断完善，而因为多年考研经验的积累，也使命题专家更能提高试题的技术含量，使研究生入学考试的试题得到不断完善，是试题的可信度和稳定性越来越强。

图 3

从试题的综合质量指标来看每一年都有波动，尤其以 2003 年和 2010 年为甚。出现此情况的原因是 2003 年试题的大纲和以后的大纲有一些变动，而我们在模型假设的时候是以大纲不变动作为基础来评价的。而 2010 年的指标是预测出来的，虽然在误差范围内，但是主观因素还是比较强，所以偏差较大。而中间几年的试题综合质量指标是令人满意的，都能维持在一个较高的水平上而又有略微的浮动。

7、模型的评价与推广

结合问题 1、2 对所建模型加以评价和推广：

7.1 模型评价

优点：1.化繁为简，化模糊为清晰。通过试题来从定量和定性两个方面来评价试卷的质量不是一个简单的过程，因为它并没有提供可供利用的直接数据，我们不能直接通过成绩分布来确定试题的质量所以评价的方法就相当模糊。而我们通过运用层次分析法，引入三个跟试题质量有关的指标深度、广度和技术性来定量评价试题的质量。试题的深度、广度和技术性可以通过试题和考试大纲来确定题目质量的好坏，虽然在一定程度上说是有主观色彩在里面的，但是却大大降低了定性和定量分析的难度，并在一定的范围内可以分析和预测试题。

2.统计样本少。在第二问中，我们运用的是灰度分析预测模型，运用灰度模型的最大优点就是所需样本比较少，因为灰度模型就是对小样本分布的预测分析，而对于考研试题来说，我们本身不可能收集到太多数据，因此用函数拟合的方法是不太现实的。而运用灰度模型既在一定程度上保证了预测的准确性又大大减少了所需样本的数量。

缺点：由于我们实现并不了解试题的各项权威数据，所以在做定量分析的时候有主观因素在里面，就针对我们这次所建立起来的模型来说，虽然经过三个人的判断，但最终结果还是不很理想。很大一部分因素是我们的主观性较强。而我们还没有做考研数学的复习，虽然以前学过这些知识，可是由于时间流逝而淡忘和考研和平时做题深度不一样的原因，偏差还是比较可观。

7.2 模型的推广

本次我们运用的是层次分析法和灰度预测模型，这两种模型的结合运用主要体现在小样本多因素多层次的样本问题的分析和预测的解决中。比如各种教育系统试题分析和产品上市前对市场的预分析等等。这种模型能较好的解决此类评价和预测问题。虽然主观性比较强，但是如果召集对本行业比较有经验的专家来进行分析确定指标的话，评价和预测效果还是相当不错的。

8、参考文献

- [1] 于桂珍，评价试卷质量的层次分析模型，天津大学。
- [2] 王守本，层次分析在竞赛评比中的应用，工科数学，1990.4。
- [3] 买焕章，郑垣模，试卷质量的模糊综合评价。
- [4] 陈衍泰，陈国宏，李美娟，综合评价方法分类及研究进展，管理科学学报，第7卷第2期：69-78，2004.4。
- [5] 李中复，吕秀芳，王大雷，正态分布密度及学生考试成绩统计，辽宁工学院学报，第6卷第5期：109-110，2004.10。
- [6] 郝永红，王学萌，灰色动态模型及其在人口预测中的应用，数学的实践与认知，第32卷第5期：2002.9。
- [7] 姜启源，谢金心，叶俊，数学建模，北京：高等教育出版社，2003。