

# 2010 高教社杯全国大学生数学建模竞赛

# 承诺书

我们仔细阅读了中国大学生数学建模竞赛的竞赛规则.

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。

我们参赛选择的题号是（从A/B/C/D中选择一项填写）： B

我们的参赛报名号为（如果赛区设置报名号的话）: \_\_\_\_\_

所属学校（请填写完整的全名）：天津农学院

参赛队员（打印并签名）：1. 李建忠

2. 田 金 歌

3. 王姣姣

指导教师或指导教师组负责人 (打印并签名): \_\_\_\_\_

日期: 2010 年 9 月 12 日

---

赛区评阅编号（由赛区组委会评阅前进行编号）：

## 2010 高教社杯全国大学生数学建模竞赛

### 编 号 专 用 页

赛区评阅编号（由赛区组委会评阅前进行编号）：

赛区评阅记录（可供赛区评阅时使用）：

评 阅 人										
评 分										
备 注										

全国统一编号（由赛区组委会送交全国前编号）：

全国评阅编号（由全国组委会评阅前进行编号）：

# 定量评估上海世博会的影响力

## 摘要

本文主要讨论 2010 年上海世博会在加强各国文化交流与促进上海和谐发展的影响力。首先通过上海世博会与布鲁塞尔、大阪等六届世博会影响力的比较建立纵向评估模型，得到文化交流影响力的模糊评价指标；然后以上海本市的经济、环境、精神面貌为指标，建立横向评估模型，可以定量地说明上海世博会对其和谐发展的影响力；最后对模型进行评价分析。

**模型一** 基于纵向评估的文化交流模型。本文通过参展国家、观展人数、历时天数和场馆面积 4 个指标，建立纵向评估模型，来评估世博会在文化交流方面的影响力。首先用层次分析法给出各项评估指标的权重，再用熵权法对权重值进行修正；然后我们建立纵向评估模型，利用模糊评价法得出布鲁塞尔、蒙特利尔、大阪、塞维利亚、汉诺威、爱知以及上海的模糊评估指标分别为 0.0936、0.1369、0.1711、0.1286、0.0986、0.0901 和 0.2810。通过以上数据比较可知，上海世博在文化交流方面的影响力明显优于以往世博所带来的影响。

**模型二** 基于横向评估的和谐发展模型。本文以经济、环境、精神面貌为指标，建立横向综合评估模型，来评估世博会在促进上海市和谐发展方面的影响力。首先在传统 GM(1,1)模型的基础上建立改进的无偏 GM(1,1)模型，预测 2010 年无世博影响下 3 项指标的影响因子值；然后利用各项指标建立 3 个多元回归模型，将 3 个回归方程作动态加权综合处理，得到综合评价模型；最后将 3 项指标的预测值带入综合评价模型，得到预测结果。受世博影响情况下综合指标值比该结果增长了 11.96 个百分点。

本文的亮点在于：纵向比较模型，选取影响力在世界上相对较大的六届世博会作为比较对象；在对选取的指标运用层次分析法赋予权重后，用熵权法对权值进行修正，使权值更为准确。横向比较模型中，我们选取经济、环境、精神三个指标，并对它们进行了加权综合，从而避免对单个指标的分析，相对提高了预测精度。这样，较全面且精确的得出世博会对上海地区的综合影响力。

**关键字：**层次分析 熵值取权法 模糊评价 多元回归 无偏 GM(1,1) 模型

## 1 问题分析

从 1851 年伦敦的“万国工业博览会”开始，世博会正日益成为各国人民交流历史文化、展示科技成果、体现合作精神、展望未来发展等的重要舞台。中国是一个有着悠久历史文明的国家，2010 年上海世博会是首次在中国举办的世界博览会，这无疑给了中国一次向世界展示文明的机会，同时也给了中国一次发展的机会。本题要求定量评估 2010 年上海世博会的影响力。我们选取历史文化交流与经济影响因子、环境因子、精神面貌表现等影响和谐的因素为两个侧重点分别进行纵向比较和横向比较，来说明上海世博会的影响力。

(1) 在对于各国人民历史文化交流方面进行横行比较。我们选取了已经成功举办的 6 届世博会包括比利时布鲁塞尔、加拿大蒙特利尔、日本大阪、西班牙塞维利亚、德国汉诺威、日本爱知。通过对选取的各届参展国家、参展天数、观展人数和世博占地面积的分析，并且通过与上海世博预计数据进行分析对比突出上海世博在各国人民历史文化交流方面的巨大影响。

(2) 在经济影响因子、环境因子、精神面貌表现等影响和谐的因素方面进行横向比较。选取这些影响因子作为评价一级指标，选取影响一级指标的影响因素作为二级评价指标。通过互联网获得这些指标的数据，分析相关数据获得指标间的关系。通过预测上海市无世博会影响情况下这些指标的预测值，与有世博会影响的指标值进行比较，定量估算出上海世博会对上海和谐因素的影响力。

## 2 符号说明

$CR$ ——一致性检验指标数；

$a_{m,n}$ ——文化交流评价因子；

$W$ ——模糊评价评价指标组合权重矩阵；

$p_{ij}$ ——评价指标模糊化后的数据矩阵；

$S$ ——模糊综合评价权重矩阵；

$G_i$ ——上海市第  $i$  年生产总值 (GDP)；

$G(x_i)$ ——上海经济影响指标 GDP；

$H(a_i)$ ——环境因子指标资金投入；

$J(b_i)$ ——精神面貌体现指标；

$A_i$ ——世博会不在上海举行情况下的综合评价指标；

$G'(x_i)$ ——为上海市第*i*年预测生产总值 GDP;

$r_i$ ——第*i*年上海 GDP 的增长率;

$B_i$ ——世博会影响下上海市和谐因素综合评价指标;

$s$ ——二者比较情况下的增长点;

### 3 问题假设

假设 1: 收集的上海世博预测结果都与实际数值相差不大;

假设 2: 上海世博会前第三产业 GDP 受世博影响忽略不计;

假设 3: 建模收集数据真实可靠;

假设 4: 建模中涉及主观分析的结论基本与事实相符;

### 4 模型建立及求解

我们建模的总体思路是: 首先利用历届世博的相关数据与上海世博数据建立纵向评估模型, 分析上海世博在文化交流方面的影响力, 然后再利用上海的经济、环境和精神面貌的各项指标来建立横向评估模型; 最后对模型的结果进行分析说明, 并作相应的模型讨论、误差分析、模型评价、改进及推广。

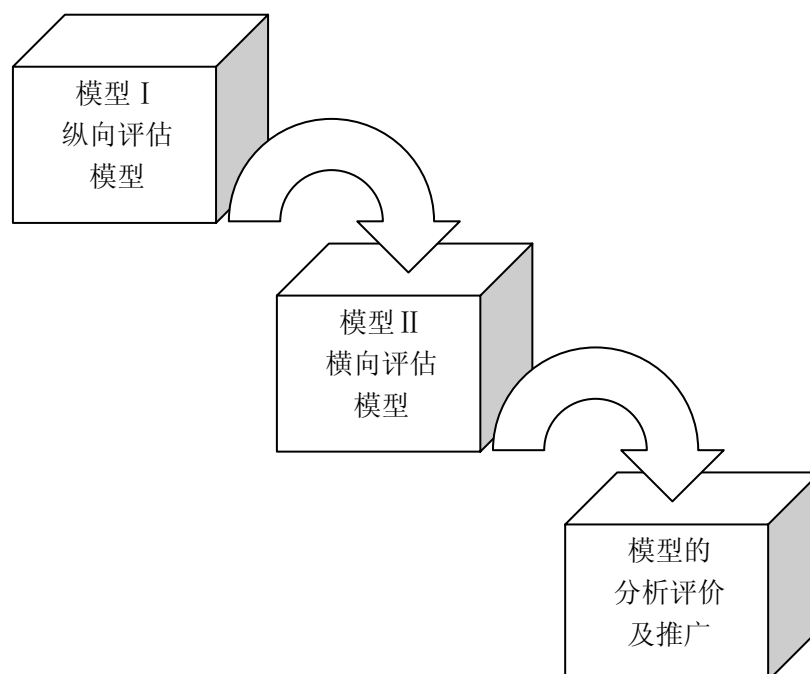


图 1 建模的总体思路图

## 4.1 纵向评估模型

### 4.1.1 层次分析法求权重

在定量评估 2010 年上海世博会的影响力过程中我们考虑到世博会涉及的直接影响因素，例如参展国家、参展天数、观展人数、基建投入、资金总投入等。此处我们考虑到分析实际问题的基本准则是对于最重要因素进行分析，我们可以根据体现文化交流的重要性差异，进一步将其简化归类为四个主要因子：参展国家、观展人数、历时天数和场馆面积。此处我们利用层次分析法对上述 4 个因子进行权重分析，基本的层次结构如下图所示：

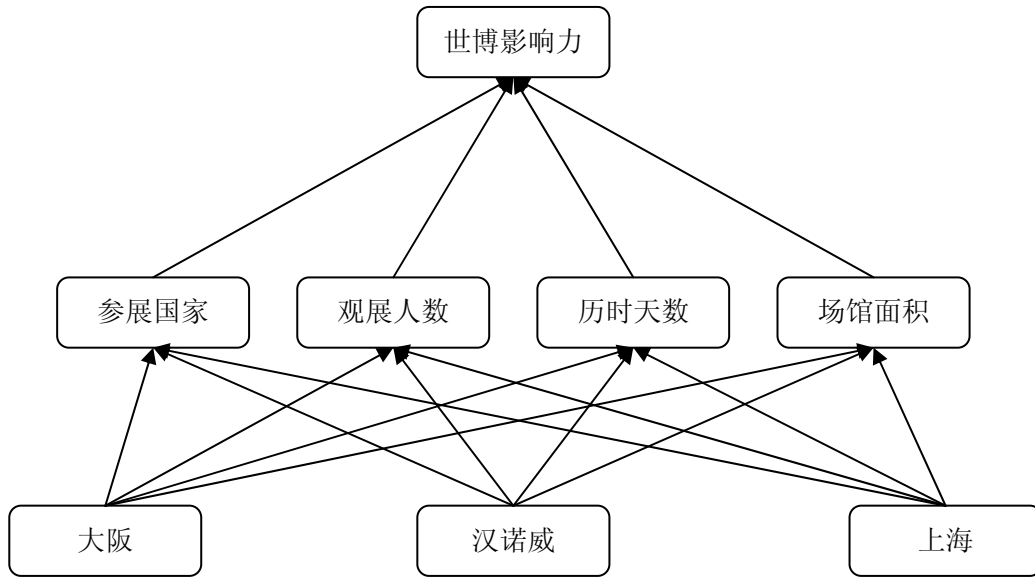


图 2 层次分析法示意图

层次分析法是一种定性分析和定量计算相结合的方法，根据相关文献资料，构造因素间的成对比较矩阵

参展国家、观展人数、历时天数和场馆面积比较矩阵

$$m_{ij} = \begin{bmatrix} 1 & \frac{1}{5} & 4 & 5 \\ 5 & 1 & 6 & 7 \\ \frac{1}{4} & \frac{1}{6} & 1 & 4 \\ \frac{1}{5} & \frac{1}{7} & \frac{1}{4} & 1 \end{bmatrix}$$

首先将矩阵  $m_{ij}$  做归一化处理，然后做按行求和最后将求和结果进行归一化处理得权重向量

$$w_1 = [0.2537 \quad 0.4623 \quad 0.1701 \quad 0.1140]$$

由  $\lambda = 4.405$  ,  $RI = 0.90$  ,  $CI = \frac{\lambda - n}{n - 1} = 0.0135$  可得一致性指标:

$$CR = \frac{CI}{RI} = 0.015 < 0.1$$

该结果满足一致性要求。

#### 4.1.2 熵值取权法对权重的校正

由于层次分析法成对比较矩阵的给出存在一定的主观因素, 于是我们想到了利用熵值取权法进行校正。

熵值取权法的优势在于, 它是通过判断各个因素的变化剧烈程度来决定该因素在最终目标中所占的权重。<sup>[2]</sup>比如参展时间对于世博会影响力的影响很重要, 但是如果在一段时间里世博会参展时间保持基本不变, 但是世博占地总面积却因某些因素导致不断的变化, 这样从熵值取权法的角度来看这对市场潜力影响就比参展时间大, 这与人们的想法也是一样, 同时它和层次分析法得出的结论是互补的且是客观的。据此,

我们利用熵值取权法客观地给出一个 4 个因素的  $1 \times 4$  的权重矩阵  $w_2$ , 对层次分析法给出的  $1 \times 4$  的权重矩阵  $w_1$ , 以 0.3:0.7 的比例进行校正, 从而给出最终 4 个因素对合理化指标的组合权重值矩阵  $W = 0.7w_1 + 0.3w_2$ , 然后对矩阵  $W$  进行归一化处理得到最终的组合权重值矩阵  $W = [w_i]_{1 \times n}$

此处我们选取大阪、汉诺威、上海世博会的基本分析数据建立  $m \times n$  的指标矩阵:

$$a_{m \times n} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

现在对每一个指标的列向量做归一化处理  $d_{ij} = \frac{a_{ij}}{\sum_{i,j=1}^{m,n} a_{ij}}$ , 其中  $0 < d_{ij} < 1$ , 于是归一化后的指标矩阵:

$$d_{m \times n} = \begin{bmatrix} d_{11} & \cdots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{m1} & \cdots & d_{mn} \end{bmatrix}$$

计算出  $d$  中每一个元素的熵值, 利用公式  $\gamma_n = -k \sum_{m=1}^m (d_{ij} \times \ln(d_{ij}))$ , 算得一个未加工

的权重矩阵  $\lambda_n$ , 现在利用公式  $w_i = \frac{1 - \lambda_i}{\sum_{i=1}^n (1 - \lambda_i)}$  <sup>[4]</sup> 求出权重矩阵  $w_2 = [w_i]_{1 \times n}$ : 我们利用大



阪、汉诺威、上海世博会的基本分析数据求解得到：

$$w_2 = [0.2563 \quad 0.2802 \quad 0.1982 \quad 0.2653]$$

则最终 4 个因素对合理化指标的组合权重值矩阵：

$$W = [0.2742 \quad 0.4392 \quad 0.1923 \quad 0.0943]$$

#### 4.1.3 建立模糊评价模型

在模糊评价模型中我们选取成功举办的 6 届世博会与上海世博会进行分析对比，具体的分析数据如表 1 所示：

表 1 世博会直接相关数据表<sup>[7]-[10]</sup>

往届世博	参展国家 (个)	观展人数 (个)	参展时间 (天)	场馆面积 (公顷)
比利时布鲁塞尔	42	41454412	186	200
加拿大蒙特利尔	62	50306648	185	400
日本大阪	81	64218770	183	330
西班牙塞维利亚	114	41814571	176	215
德国汉诺威	177	18100000	153	160
日本爱知	125	22049544	185	173
中国上海	249	70000000	184	528

由上面权重分析我们已经得到了 4 个世博会的直接影响因素的权重，此处我们利用模糊数学公式  $p = \frac{f(x_{ij}) - \min}{\max - \min}$  (max 为  $f(x_{ij})$  中的最大者，min 为  $f(x_{ij})$  中的最小值)

将 4 个评价指标的数字  $f(x_{ij})$  进行模糊处理<sup>[3]</sup>，建立模糊关系矩阵：

$$p_{ij} = \begin{bmatrix} 0.0000 & 0.0966 & 0.1884 & 0.3478 & 0.4010 & 0.6522 & 1.0000 \\ 0.4071 & 0.4039 & 0.3973 & 0.3743 & 0.4039 & 0.2989 & 0.4006 \\ 0.3499 & 0.4679 & 0.6534 & 0.3547 & 0.0911 & 0.0385 & 0.7305 \\ 0.2065 & 0.6159 & 0.4726 & 0.2372 & 0.1512 & 0.1246 & 0.8779 \end{bmatrix}$$

由权重分析可知 4 个影响指标的权重为：

$$W = [0.2742 \quad 0.4392 \quad 0.1923 \quad 0.0943]$$

把  $W$  与模糊关系矩阵  $p_{ij}$  相乘得模糊综合评价结果：

$$k = W \times p_{ij}$$

$$= [0.2742 \quad 0.4392 \quad 0.1923 \quad 0.0943]$$

$$\times \begin{bmatrix} 0.0000 & 0.0966 & 0.1884 & 0.3478 & 0.4010 & 0.6522 & 1.0000 \\ 0.3499 & 0.4679 & 0.6534 & 0.3547 & 0.0911 & 0.0385 & 0.7305 \\ 0.4071 & 0.4039 & 0.3973 & 0.3743 & 0.4039 & 0.2989 & 0.4006 \\ 0.2065 & 0.6159 & 0.4726 & 0.2372 & 0.1512 & 0.1246 & 0.8779 \end{bmatrix}$$

$$= [0.2514 \quad 0.3678 \quad 0.4596 \quad 0.3455 \quad 0.2419 \quad 0.2649 \quad 0.7548]$$

对矩阵  $k$  进行归一化处理得到最终的模糊综合评价结果矩阵  $S = [k_i]_{1 \times j}$ ，即

$$S = [0.0936 \quad 0.1369 \quad 0.1711 \quad 0.1286 \quad 0.0986 \quad 0.0901 \quad 0.2810]$$

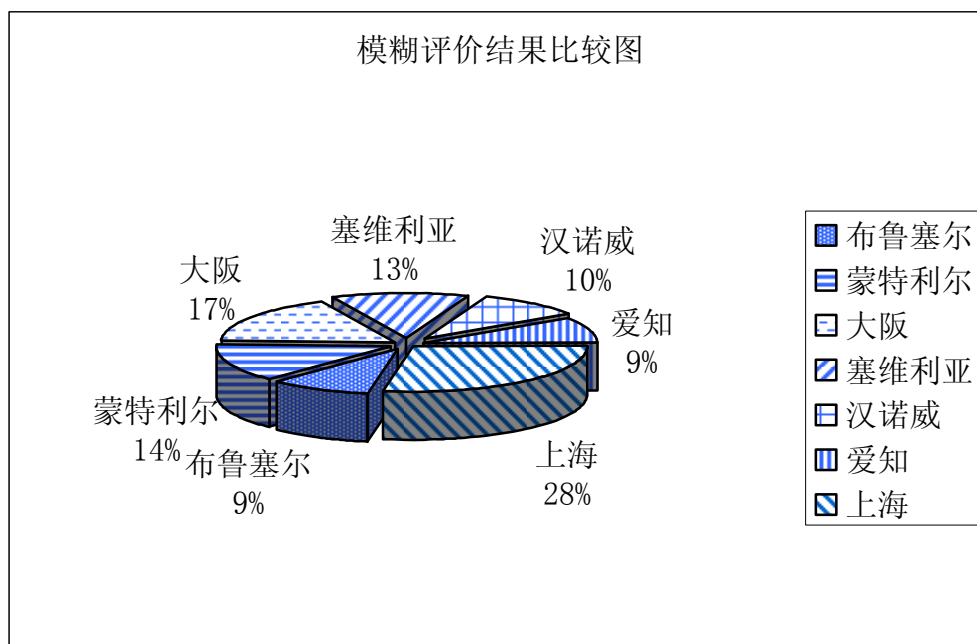
#### 4.1.4 结果分析及说明

将上述结果整理成表见表 2

**表 2 模糊综合评价结果表**

国家	布鲁塞尔	蒙特利尔	大阪	塞维利亚	汉诺威	爱知	上海
权重	0.0936	0.1369	0.1711	0.1286	0.0986	0.0901	0.2810

此处将上表模糊结果用饼形图表示：



**图 3 模糊评价结果比较图**

从上示饼形图我们可以清楚地看出上海世博会在各国文化交流方面的影响明显优

于已成功举办的其他世博会，无论是从规模、参展国家、还是从观展人数方面上海世博会都有着明显的优势，由此我们不难得出上海世博会对于世界文化交流所产生的影响是空前的、是其他世博会所未能企及的。世界各国的文化交流必定会因上海世博会而更近一成，上海世博会也必定为以后举办世博的国家留下不可忽略的见解财富。

4.2 横向比较——综合评估模型

4.2.1 指标的选取

当今社会的主题是和谐。为了突出这个主题，对上海市横向比较时，我们以影响和谐的因素为主选取经济影响因子指标 GDP、环境因子指标环境投入资金、精神面貌表现指标精神方面资金投入三个方面作为综合评价模型的一级指标。每一个影响和谐的因素又都具有各自的影响因素，把这些影响因素作为综合评价的二级指标。

二级指标的选取的原则是对一级指标影响较大。根据这个原则，我们分别对三个一级指标选取影响因素。

首先，对经济因子指标进行选取影响因素。在上海世博举行之前，政府及相关行业、企业的投资主要用作场馆建设、配套设施和整改环境，对第二产业的建筑业有较大拉动影响，而对第三产业的影响不大。世界博览会举行时以及举办后的一段时间里，必然会拉动第三产业，使其具有较大的发展，因此选取对第三产业影响较大的指标交通运输、仓储及邮电通信，教育经费，旅游业，金融保险业，科技成果，房地产业，卫生事业机构个数作为经济因子指标的影响因素。对环境因子指标进行选取时，选取对环境影响较大的指标可吸入颗粒物、二氧化硫、二氧化氮作为其影响因素。对精神面貌指标选取时，选取具有代表性的福利机构数量、图书馆、文化活动机构作为其影响因素。

指标的划分如表 3：

表 3 各级指标

综合指标	和谐因素				
一级指标	经济因子			环境因子	精神面貌
二级指标	运输行业	教育经费	旅游业	可吸入颗粒物	福利机构数量
	金融保险	科技成果	房地产	二氧化硫	图书馆
	卫生机构			二氧化碳	文化活动机构

先考虑上海在不举办世博会的情况下，预测 2010 年上海市影响和谐的综合指标。首先对选取的二级指标利用无偏 GM (1,1) 模型进行预测，预测出 2009 年、2010 年的二级指标值。然后将选取一级指标作为因变量，把影响该指标的多个二级指标作为自变量，建立反映而二者关系的多元回归模型。将测出的 2010 年二级指标值带入多元回归模型，就可以得到上海不举行世博会情况下三个一级指标的预测值，分别赋予三个指标不同权值，估算出无世博会影响下上海市的综合评价指标。

其次考虑上海在举办世博会情况下，2010 年影响和谐的综合指标。根据上海市 GDP 的增长率建立上海市生产总值增长模型，根据 2009 年上海 GDP 预测出世博会影响下 2010 年经济影响因子。再根据官方网站对环境资金投入和精神方面资金投入的预测，

得出世博会影响下 2010 年环境因子指标值和精神面貌体现指标值。最后, 我们分别赋予三个指标不同权值, 估算出世博会影响下上海市的综合评价指标。

最后，建立比较模型，对两个综合评价指标进行比较，计算出有世博会影响下上海市综合评价指标的值比无世博会影响下增长的百分点。这样就可以得到世博会对上海综合的影响。

对于上海经济影响因子指标 GDP 较其他指一级标的分析较为复杂, 需要根据上海市生产总值与第三产业 GDP 的关系, 通过最小二乘拟合估算上海市 GDP 值。从而得到经济影响因子指标的求解模型。

为了使表达清晰, 将横向建模思路建立为如下流程图:

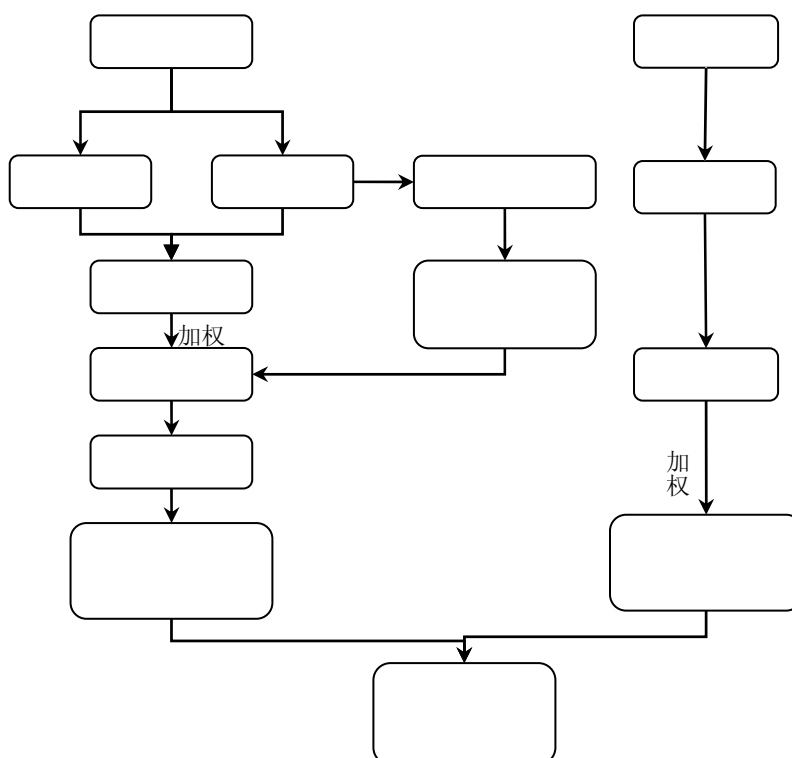


图 4 建模思路

### 4.2.2 无偏 GM(1,1)模型

### (1) 模型建立

无偏  $GM(1,1)$  模型对实际问题进行预测时发现, 当原始数据序列增长率变化较小, 即  $GM(1,1)$  中的  $|a|$  较小时, 预测精度较高; 当原始数据序列增长率变化较大, 即  $GM(1,1)$  模型预测方法中的  $|a|$  较大时, 预测精度低<sup>[4]</sup>。由于上海市经济影响因子指标GDP、环境因子指标以及精神面貌体现指标变化率较大, 我们运用无偏  $GM(1,1)$  模型对这些和谐

影响因素进行预测。

根据原始数据，建立时间序列：

$$X^{(0)} = \{X^{(0)}(1), X^{(0)}(2), \dots, X^{(0)}(n)\}$$

通过将原始数据进行一次累加生成新序列：

$$X^{(1)} = \{X^{(1)}(1), X^{(1)}(2), \dots, X^{(1)}(n)\}$$

建立 GM (1, 1) 模型相应的微分方程为：

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = \mu \quad (1)$$

其中：a 称为发展灰数；μ 称为内生控制灰数。

设  $\hat{\alpha}$  为待估参数向量， $\hat{\alpha} = \begin{bmatrix} a \\ \mu \end{bmatrix}$ ，可利用最小二乘法求解。解得：

$$\hat{\alpha} = (B^T B)^{-1} B^T Y_n \quad (2)$$

$$Y_n = [x^{(0)}(2) \quad x^{(0)}(3) \quad \dots \quad x^{(0)}(n)]^T$$

$$B = \begin{bmatrix} -\frac{1}{2}(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -\frac{1}{2}(x^{(1)}(2) + x^{(1)}(3)) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}(x^{(1)}(n-1) + x^{(1)}(n)) & 1 \end{bmatrix}$$

求解微分方程，即可得预测模型：

$$\hat{X}^{(1)}(k+1) = \left[ X^{(0)}(1) - \frac{\mu}{a} \right] e^{-ak} + \frac{\mu}{a}, \quad k = 0, 1, 2, \dots, n \quad (3)$$

设原始序列为严格的指数序列，即

$$x^{(0)}(k) = A e^{\alpha(k-1)}, k = 1, 2, \dots, n \quad (4)$$

则该序列进行一次累加生成序列为：

$$x^{(1)}(k) = A \frac{1 - e^{\alpha k}}{1 - e^{\alpha}}, k = 1, 2, \dots, n \quad (5)$$

由传统 GM (1,1) 模型建模方法可得：

$$Y_n = \begin{bmatrix} Ae^\alpha & Ae^{2\alpha} & \dots & Ae^{(n-1)\alpha} \end{bmatrix}$$

$$B = \begin{bmatrix} -\frac{1}{2}A \frac{2-e^\alpha-e^{2\alpha}}{1-e^\alpha} & 1 \\ -\frac{1}{2}A \frac{2-e^{2\alpha}-e^{4\alpha}}{1-e^\alpha} & 1 \\ \vdots & \vdots \\ -\frac{1}{2}A \frac{2-e^{(n-1)\alpha}-e^{n\alpha}}{1-e^\alpha} & 1 \end{bmatrix}$$

经过以上矩阵推导可得：

$$\begin{bmatrix} a \\ \mu \end{bmatrix} = (B^T B)^{-1} B^T Y_n = \begin{bmatrix} \frac{2(1-e^\alpha)}{1+e^\alpha} \\ \frac{2A}{1+e^\alpha} \end{bmatrix} \quad (6)$$

此处利用公式（6）可得：

$$\alpha = \ln \frac{2-a}{2+a}, A = \frac{2\mu}{2+a}$$

无偏  $GM(1,1)$  模型最终拟合预测方程为：

$$\begin{cases} \hat{x}^{(0)}(1) = A \\ \hat{x}(k) = \frac{Ae^\alpha(1-e^\alpha)}{1-e^\alpha} e^{-\alpha(k-1)}, k = 2, 3, \dots, n \end{cases} \quad (7)$$

比较（4）和（7）可知，两式存在差异较大。因为采用传统  $GM(1,1)$  模型对指数序列建模总存在偏差，而利用无偏  $GM(1,1)$  模型可以消除传统  $GM(1,1)$  模型的固有偏差<sup>[5]</sup>。

## （2）预测模型的求解

通过以上建立的无偏  $GM(1,1)$  模型，对经济因子指标、环境因子指标、精神面貌体现指标三个一级指标各自的影响因素（即二级指标）预测出 2009 年、2010 年的二级指标的值。

### 影响经济因子的二级指标

通过互联网，在上海统计局数据中我们可以得到影响第三产业 GDP 的因素交通运输行业、教育经费、旅游业、金融保险业、科技成果、房地产业、其他服务业等指标的相关数据如表 4 所示：

**表 4 影响经济因子的二级指标表（单位亿元）**

年份	运输行业	教育经费	旅游业	金融保险	科技成果	房地产	卫生机构
2000	385.61	200.89	181.40	602.95	1102	263.35	4400
2001	426.32	232.04	204.26	529.26	1338	328.59	3813
2002	479.13	273.97	272.53	485.25	1418	384.86	2422
2003	547.81	307.02	244.71	517.97	1508	487.82	2319
2004	493.6	383.27	385.45	612.45	1629	666.30	2577
2005	582.6	422.95	444.54	675.12	1701	676.12	2527
2006	669.01	426.34	464.63	825.20	1953	688.10	2519
2007	723.13	370.73	520.10	1209.08	2396	806.79	2646
2008	712.99	431.83	526.47	1414.21	1866	939.34	2809

根据无偏  $GM(1,1)$  模型的预测方程，使用 MATLAB 工具箱对此模型进行求解（代码见附录），可预测出 2009 年、2010 年各个影响经济因子的第二指标预测值，结果如表 5 所示：

**表 5 影响经济因子第二指标预测值表（单位亿元）**

年份	运输行业	教育经费	旅游业	金融保险	科技成果	房地产	卫生机构
2009	912	545	549	1561	2087	1267	2763
2010	1056	562	697	1856	2384	1623	2934

### 影响环境因子的二级指标

根据上海统计年鉴，可以得到可吸入颗粒物、二氧化硫、二氧化氮影响环境因子二级指标<sup>[13]</sup>的相关数据，如表 6 所示：

**表 6 影响环境因子的二级指标表（单位  $mg/m^3$ ）**

年份	可吸入颗粒物	二氧化硫	二氧化碳
2001	0.100	0.024	0.044
2002	0.108	0.030	0.051
2003	0.097	0.043	0.057
2004	0.099	0.055	0.062
2005	0.088	0.061	0.061
2006	0.086	0.051	0.055
2007	0.088	0.055	0.054
2008	0.084	0.051	0.056
2009	0.081	0.035	0.053

根据无偏  $GM(1,1)$  模型的预测方程，使用 MATLAB 工具箱对此模型进行求解（代码见附录），可解得 2009 年、2010 年各个影响环境因子的第二指标的预测值，结果如表 7 所示：

表 7 影响环境因子的第二指标预测值表（单位  $mg/m^3$ ）

年份	可吸入颗粒物	二氧化硫	二氧化碳
2010	0.0534	0.0603	0.0532

### 影响精神面貌的二级指标

根据上海统计年鉴，可以得到福利机构数量、图书馆、文化活动机构等影响精神面貌体现的二级指标<sup>[11]</sup>的相关数据，如表 8 所示：

表 8 影响精神面貌的二级指标表（单位个）

年份	福利机构数量	图书馆	文化活动机构
2000	426	31	340
2001	409	32	276
2002	450	32	270
2003	453	35	261
2004	449	28	251
2005	482	28	248
2006	513	28	250
2007	569	30	247
2008	591	29	245

根据无偏  $GM(1,1)$  模型的预测方程，使用 MATLAB 工具箱编程（代码见附录）对此模型进行求解，可预测出 2009 年、2010 年各个影响精神面貌的二级指标的预测值，结果如表 9 所示：

表 9 影响精神面貌的二级指标预测值表（单位个）

指标年份	福利机构数量	图书馆	文化活动机构
2009	590	30	243
2010	601	32	240

### 4.2.3 多元线性回归模型

#### （1）模型建立

如果世界博览会不在上海举行，根据 2000—2008 年 GDP、环境因子指标以及精神面貌体现指标等影响和谐的因素，利用多元线性回归的方法<sup>[6]</sup>，建立的多元回归模型，模型相同，如下：

$$\text{一般称} \begin{cases} Y = X\beta + \varepsilon \\ E(\varepsilon) = 0, COV(\varepsilon, \varepsilon) = \sigma^2 I_n \end{cases} \text{为高斯—马尔柯夫线性模型（} k \text{ 元线性回归模}$$

型），并简记为  $(Y, X\beta, \sigma^2 I_n)$



$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$  称为回归平面方程。

用最小二乘法求  $\beta_0, \dots, \beta_k$  的估计量，作离差平方和：

$$Q = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_k x_{ik})^2$$

选择  $\beta_0, \dots, \beta_k$  使  $Q$  达到最小，解得估计值：

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

得到的  $\hat{\beta}_i$  代入回归平面方程得：

$$y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k$$

此方程为经验回归平面方程。 $\hat{\beta}_i$  称为经验回归系数。

## (2) 模型求解

此处我们利用经济影响因素 GDP、环境因子指标、精神面貌体现指标的相关数据进行多元回归拟合，然后利用无偏 GM(1,1) 模型的预测值进行各项指标投入的相关预测。

### 经济因子指标——GDP

根据影响经济因子的二级指标具体数据（见表），使用 MATLAB 软件编程（代码见附录），从而得到回归模型：

$$y_i = -161.5496 + 5.0041x_1 - 2.5285x_2 + 2.8967x_3 + 2.0690x_4 - 1.2401x_5 + 3.5910x_6$$

利用以上多元回归方程式，带入无偏 GM(1,1) 模型中预测出的 2010 年第三产业 GDP 的影响因子，解得第三产业 GDP 值为：

$$y_{2010} = 7819.5 \text{ 亿元}$$

要得到经济因子指标 GDP 与第三产业 GDP 的影响因子的关系，就需要找到上海市 GDP 与第三产业 GDP 的关系。利用最小二乘拟合得到如下函数关系式：

$$G_i(x) = a_1x + a_2$$

$G_i$  表示上海市第  $i$  年生产总值 (GDP);  $a_1, a_2$  为拟合系数。

通过中国统计年鉴可以得到上海市总生产总值(GDP)<sup>[11]</sup>以及第三产业GDP值, 具体数据如表 10 所示:

**表 10 生产总值与第三产业的 GDP 的数据关系表 (单位亿元)**

年份	1993	1994	1995	1996	1997	1998	1999	2000
GDP	1519.2	1990.9	2499.4	2957.6	3438.8	3801.1	4188.7	4771.2
第三产业	579.0	794.8	1020.2	1292.1	1592.7	1855.4	2129.6	2486.9
年份	2001	2002	2003	2004	2005	2006	2007	2008
GDP	5210.1	5741	6694.2	8072.8	9154.2	10366.4	12188.8	14069.9
第三产业	2728.9	3038.9	3404.2	4097.3	4620.9	5244.2	6408.5	7872.2

利用最小二乘模型求解得到的拟合方程式, 带入 2010 年上海市第三产业 GDP 的值, 可以得到上海市经济因子指标 GDP, 解得:

$$G(x_{2010}) = 14561.0 \text{ 亿元}$$

#### 环境因子指标——环境资金投入

根据影响环境因子的二级指标的相关数据 (见表 6 影响环境因子的二级指标表), 使用MATLAB软件编程 (代码见附录), 解得回归系数, 从而得到关于环境因子的多元回归方程, 如下:

$$H(a_i) = 778 - 345x_7 + 9967x_8 - 17211x_9$$

将无偏 GM(1,1)模型中预测出的 2010 环境影响因子带入该多元回归方程, 解得 2010 年环境投入为:

$$H(a_{2010}) = 444.7842 \text{ 亿元}$$

#### 精神面貌体现指标——精神面貌资金投入

根据影响精神面貌的二级指标的相关数据 (见表 8 影响精神面貌的二级指标表), 使用MATLAB软件编程 (代码见附录), 解得回归系数, 从而得到关于精神面貌体现指标的多元回归方程, 如下:

$$J(b_i) = -582.3323 + 20.2256x_{10} - 77.7327x_{11} - 9.1928x_{12}$$

将无偏 GM(1,1)模型中预测出的 2010 精神面貌表现指标影响因子带入该多元回归方程, 解得 2010 年精神方面投入为:

$$J(b_{2010}) = 6879.5 \text{ 亿元}$$

#### 4.2.4 综合评价模型

通过多元线性回归模型就可以分别得到上海经济影响指标 GDP ( $G(x_i)$ )、环境因

子指标  $H(a_i)$  以及精神面貌体现指标  $J(b_i)$  的多元回归方程。通过这些指标各自的多元回归方程，可以预测第  $i$  年各项指标的值。我们将这三项指标分别赋予权重  $w_1, w_2, w_3$ ，建立综合评价模型：

$$A_i = w_1 G(x_i) + w_2 H(a_i) + w_3 J(b_i)$$

其中， $A_i$  为世博会不在上海举行情况下的综合评价指标。

如果 2010 年世博会不在上海举行，我们分别赋予经济影响因子 GDP、环境因子以及精神面貌表现三个不同评价指标权重为 0.7、0.1、0.2。运用以上模型求得的结果，可以计算出这种情况下上海市和谐因素的综合评价指标，可得：

$$A_{2010} = 0.7G(x_{2010}) + 0.1H(a_{2010}) + 0.2J(b_{2010})$$

解得： $A_{2010} = 11613.0$

#### 4.2.5 上海市 GDP 增长模型

根据以上建立的多元回归模型，利用无偏 GM(1,1)模型可预测出在不举行世博会情况下 2010 年第三产业 GDP 值，再利用拟合模型估算出上海市 GDP 值。在世博会影响下，上海无论是第一、二产业，还是第三产业均会有较大增长，2010 年上海市生产总值必然会有较大的增长。可以得到 GDP 增长模型如下：

$$G'(x_i) = G'(x_{i-1})(1 + r_i)$$

$G(x_i)$  为上海市第  $i$  年生产总值 GDP； $r_i$  为第  $i$  年上海 GDP 的增长率。

根据上海第十三届人民代表大会第三次会议上正市长报告可知，2010 年上海市 GDP 较 2009 增加 8% 以上，2009 年上海市 GDP 的值为 14900.93 亿元。这样，就可以预测出 2010 年上海在世博会影响下的 GDP 值，为：

$$G'(x_{2010}) = 16093.00 \text{ 亿元}$$

#### 4.2.6 比较模型

##### (1) 模型建立

2010 年世博会在上海举行，必然会对上海经济、环境、精神、文化等多方面造成影响，使上海的经济、环境、精神等影响和谐的因素发生较大的变化。这就需要在不举办世博会的情况下预测出的综合评价指标与在世博影响下的综合评价指标进行比较。首先建立世博会影响下的综合评价指标模型，如下：

$$B_i = w_1 G'(x_i) + w_2 H'(a_i) + w_3 J'(b_i)$$

然后，建立二者的比较模型，如下：

$$s = \frac{B_i - A_i}{A_i} \times 100\%$$

其中， $B_i$  为世博会影响下上海市和谐因素综合评价指标； $s$  为二者比较情况下的增长点。

通过以上模型，我们就可以比较出有世博会影响下和无世博会影响下上海生产总值的变化，由此，我们可以说明世博会对上海经济发展的影响。

## (2) 模型求解

世博会影响下，上海市和谐影响因素综合评价指标为：

$$B_{2010} = 0.7G'(x_{2010}) + 0.1H'(a_{2010}) + 0.2J'(b_{2010})$$

通过权威数值预测可得到上海市在世博会影响下环境投入指标  $H'(a_{2010}) = 482.79$  亿元，精神方面投资  $J'(b_{2010}) = 8445.1$  亿元，解得：

$$B_{2010} = 13002.0$$

根据以上数据，可以解得上海市在世博会影响下和不受世博会影响下，综合评价指标增长百分点，结果为：

$$s = 11.96\%$$

### 4.2.7 结果分析

通过对评价模型的求解，可以得到三个一级指标的在世博会影响情况下比无世博会影响情况下的增长百分比。其中经济影响因子指标 GDP、环境因子指标、精神面貌表现指标的增长百分比分别为 10.52%、8.54%、22.76%。对于第三个一级指标数据误差较大，而采用动态加权法对三个指标赋予权值之后，得到综合评价模型。使用该模型解出的上海市影响和谐的因素综合评价指标在世博会影响比无世博会影响下的值增长的百分比为 11.96%。这个数据较为合理且相对于单项预测更为准确。能相对准确对上海世博会的影响做出预测。

## 5 模型讨论与误差分析

### 5.1 模型讨论

#### 模型一

模型一的建立的基础是对已成功举办多届世博的影响力进行分析，例如参展国家、参展天数、观展人数、场馆面积都具有相对精确的数据支持，所以带来的影响力可以具有一定的定量关系，我们通过上海世博的实际数据、预测数据与历届世博作相应的对比，从侧面反映出上海世博所带来的影响力是历届世博所未能达到的。

在对数据进行分析的过程中，针对不同模型的需求，对数据进行了不同的分析处理，例如数据的归一化处理、数据的模糊化处理，不仅解决了量纲不同所带来的问题，

也一定程度上减小了数据变化较大所带来的相对误差。

## 模型二

为了更加突出上海世博所带来的巨大影响，世博所带来的和谐是很据有代表性的，所以将经济发展、环境投入、文化精神建设进行综合评价，来突出世博所带来的影响力。

在数据的处理过程中，利用无偏  $GM(1,1)$  模型来预测不同的数据，比较合理的预测各项指标值。在对不同指标的分析中，根据不同的数据建立起不同的多元回归模型，并且将不同的指标进行加权综合处理，求得基本预测增长率。

### 5.2 误差分析

我们只对模型二进行误差分析。在模型二中大量采用多元回归模型进行了数据的拟合，此处我们利用指标因子较多的经济指标进行误差的分析，利用 MATLAB 编程可得基本的误差分析指标以及残差分析图



图 5 残差分析图 1

相关系数  $r^2=0.9979$ ， $F=161.1005$ ，显著性水平  $p=0.0062$

由图像可知第四组数据的相对误差较大，所以此处我们将其剔除在进行数据的回归拟合，所得结果如下所示：



图 6 残差分析图 2

相关系数  $r^2=1$ ， $F=22259$ ，显著性水平  $p=0$

由回归系数为 1，显著性水平为 0 可以看出，回归模型所得结果误差极小，不会影响模型的基本预测分析，所以本模型具有较高的合理性。

## 6 模型评价与推广

### 6.1 模型评价

在建立模型一中，我们采用了层次分析法和熵值取权法来选取世博文化交流主要表现因子的权重，层次分析法的应用虽然将各因素的权重比较理性的解决了，但主观色彩比较浓，于是我们采用与层次分析法具有明显的互补性质的熵值取权法对权重进行了修正，这样得到的组合权重就比较客观而可信。模糊评估模型的运用避免了传统印象评价的局限性与片面性，将人类思维与数学方法相结合，进一步提高了数据的可靠性与比较性。

通过模型二，我们可以较为预测出上海市在世博会影响下的和谐发展的增长百分点，且预测出的结果较为准确、可靠。因此该模型对于本题具有较强的可行性。除此之外，我们利用了  $GM(1,1)$  模型的改进模型无偏  $GM(1,1)$  模型，与传统的  $GM(1,1)$  模型

相比，无偏模型不存在传统  $GM(1,1)$  模型所固有的偏差，因而也就消除传统  $GM(1,1)$  模型在原始数据序列增长率较大时失效的现象。这样，就可以相对减小预测数值的误差。然而，由于数据的缺乏，该模型考虑的指标较少，且只能预测短期内的影响力，实用性较差。

#### 优点

1. 层次分析法的使用将准则目标得以很好的量化，使评估模型更加简单明确。
2. 熵值取权法不具有如何主观色彩，权重向量的决定是客观的，具有评价过程的透明性，减少了单独采用层次分析法因主观因素对结果造成的影响，使结果更加精确，有说服力。
3. 无偏  $GM(1,1)$  模型的运用消除了传统  $GM(1,1)$  模型所存在的固有偏差，所以可以更精确的进行数据预测。

4. 对于世博会对上海市的影响，我们采用动态加权、综合预测的方法，避免了单指标考虑，提高模型的精确度，还能定量估算出世博会对上海市的综合影响。

#### 缺点

1. 由于所查到的数据量的局限性，上海世博会的长期影响以及其对周边省市乃至全国的影响不能够定量预测。
2. 由于上海世博会尚未结束，我们所得到的 2010 年上海市世博会影响的指标均是预测值，这必将与现实值具有一定误差。
3. 我们所采用的一级指标及二级指标，在一定程度上有局限性，这必然会导致模型误差在一定程度上增大。

## 6.2 模型改进

上海世博会无论是对本市、中国还是世界必然会产生深远影响。由于数据问题我们未能进行全面考虑，但可以利用以往各届世博对于本国及世界产生的影响建立改进模型，对上海世博深远影响进行预测与估计。除此之外，还可以考虑通过上海周边省市经济增长变化率建立改进模型，来预测上海世博会对周边地区的带动情况。

## 6.3 模型推广

本文的模糊评价模型不仅可以用于指标的对比，而且还可以推广到其他评估模型，例如对教学质量的评估、圣体素质的评估等。

本文的无偏  $GM(1,1)$  模型可以推广到其他与经济增长率较大有关的预测问题。

## 7 参考文献

- [1] 《改进熵值法问题的初探》，浙江理工大学，苏洁
- [2] 张成科，基于熵的水质模糊评价模型及应用[J]. 系统工程理论与实践，1998 (6) : 80 - 85
- [3] 《用模糊数学评判法对普通高校教师工作的评估》，罗青山，1992
- [4] 吉培荣，黄巍松等. 无偏灰色预测模型[J]. 系统工程与电子技术. 2000, 22(6)
- [5] 孙皓，基于灰色理论和神经网络的道路交通事故预测研究[J]. 检测技术与自动化装置 2007, 5, 23--27
- [6] 《数学建模与实验 MATLAB 教程》，第 11 讲，回归分析
- [7] 唐子来、付磊. 世博会的经典案例研究之一 1970 年大阪世博会[J]. 城市规划汇刊，2004 年 1 月：6-9.
- [8] 唐子来、陈琳. 世博会的经典案例研究之二 1998 年里斯本世博会[J]. 城市规划汇刊，2004 年 2 月：13-16.
- [9] 唐子来、赵渺希. 世界博览会的经典案例研究之三 1992 年塞维利亚世博会[J]. 城市规划汇刊，2004 年 3 月：39-42.
- [10] 唐子来、朱弋宇. 世界博览会的经典案例研究之四 2000 年汉诺威世博会[J]. 城市规划汇刊，2004 年 4 月：49-52.
- [11] 中华人民共和国国家统计局数据库 <http://www.stats.gov.cn/tjsj/> 2010-9-10
- [12] 上海统计局数据库 <http://www.stats-sh.gov.cn/2008shtj/index.asp> 2010-9-10
- [13] 上海环境状况公报 <http://www.sepb.gov.cn/> 2010-9-10

## 8 附录

### 熵值取权法

```
clc
clear
%熵值取权法
a=[81    64218770    183 330
249 70000000    184 528
177 18100000    153 160];
s=zeros(1,4);
```

```

for j=1:4
    for i=1:3
        s(j)=s(j)+a(i,j);
    for m=1:4
        for n=1:3
            b(n,m)=a(n,m)/s(m);
        end
    end
end
end
b% 归一化
t=zeros(1,4);
k=-1/log(4);%2 代表指标个数
for m=1:4
    for n=1:3
        t(m)=t(m)+b(n,m)*log(b(n,m));
        f=k*t;
    end
end
f% 基本权重

g=0;
for i=1:4
    g=g+f(i);
end
g
for j=1:4
    h(j)=f(j)/g;
end
h% 权重归一化
m=[0.2537 0.4623 0.1701 .01140]
f=0.7*m+0.3*h;
g=0;
for i=1:4
    g=g+f(i);
end
g
for j=1:4
    h(j)=f(j)/g;
end
h% 权重归一化
模糊评价
function y=pls(a)

```



loda a.txt%原始数据存放在纯文本文件 a.txt 中

```
s=zeros(1,4);
for j=1:4
    for i=1:7
        s(j)=s(j)+a(i,j);
    end
end
for m=1:4
    for n=1:7
        b(n,m)=a(n,m)/s(m);
    end
end
end
m=max(b);
m=max(m)
n=min(b);
n=min(n);
c=(b-n)/(m-n);
c% 归一化
t=h*c'
g=0;
for i=1:7
    g=g+t(i);
end
g
for j=1:7
    h(j)=t(j)/g;
end
h% 权重归一化
```

## 灰色预测

function y=pls(t)

loda t.txt%原始数据存放在纯文本文件 t.txt 中

```
m=0;
i=1;
while i<13
    m=m+a(i);
    s(i)=m;
    i=i+1;
end
s% 数据叠加
k=a(1,2:12);
y=k';
m=-1/2*s(1,2:12)
n=ones(1,11);
```

```

b=[m;n];
b=b';
u=inv(b'*b)*b'*y;
end
u%求得时间响应式的参数 a 与 u
m=log((2-u(1))/(2+u(1)));
n=2*u(2)/(2+u(1));
for i=1:23
    s1(i)=n*exp(m*(i-1));%预测模型 3
end
s1
for i=1:9
c(i)=(s1(i)-a(i))/a(i);
end
c
t=0;
for i=2:9
    t=t+c(i);
end
t=t/8

```

```

多元回归分析
clc
clear
function y=pls(pz)
loda pz.txt%原始数据存放在纯文本文件 pz.txt 中

[b, bint,r,rint,stats]=regress(Y,X,0.05)
rcoplot(r,rint)

```

## 多元回归分析

```

clc
clear
function y=pls(pz)
loda pz.txt%原始数据存放在纯文本文件 pz.txt 中

[b, bint,r,rint,stats]=regress(Y,X,0.05)
rcoplot(r,rint)
GDP 拟合代码:
clc
clear
y=[781.66 893.77 1114.32 1519.23 1990.86 2499.43 2957.55 3438.79 3801.09 4188.73
4771.17 5210.12 5741.03 6694.23 8072.83 9154.18 10366.37 12188.85 14069.87];
x=[241.82 309.07 402.77 579.03 794.8 1020.2 1292.11 1592.74 1855.36 2129.6 2486.86
2728.94 3038.9 3404.19 4097.26 4620.92 5244.2 6408.5 7872.23];
a=polyfit(x,y,1)
z=polyval(a,x);
plot(x,y,'k+',x,z,'r')
xlabel('第三产业 GDP');
ylabel('总 GDP');
title('上海')
a =

```

```

1.7999 486.3957

```

拟合图像：

