



中南大學
CENTRAL SOUTH UNIVERSITY

矿业系统工程

资源与安全工程学院
采矿工程系 杨珊 副教授



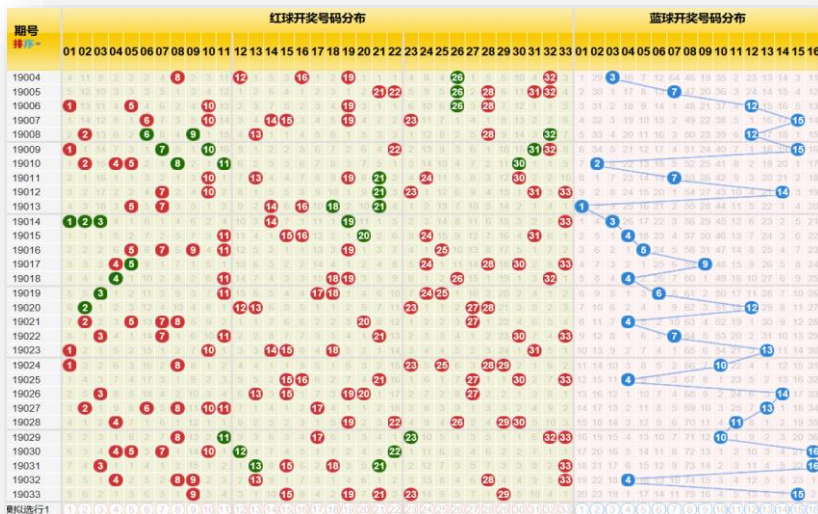
主要内容

第一节 预测概述

第二节 定性预测技术

第三节 时间序列预测

第四节 灰色预测技术



体育彩票



股市



黄金价格



煤矿百万吨死亡率



隆
中
对

亮答曰：“自董卓已来，豪杰并起，跨州连郡者不可胜数。曹操比于袁绍，则名微而众寡，然操遂能克绍，以弱为强者，非惟天时，抑亦人谋也。今操已拥百万之众，挟天子而令诸侯，此诚不可与争锋。孙权据有江东，已历三世，国险而民附，贤能为之用，此可以援而不可图也。荆州北据汉、沔，利尽南海，东连吴会，西通巴、蜀，此用武之国，而其主不能守，此殆天所以资将军，将军岂有意乎？益州险塞，沃野千里，天府之土，高祖因之以成帝业。刘璋暗弱，张鲁在北，民殷国富而不知存恤，智能之士思得明君。将军既帝室之胄，信义著于四海，总揽英雄，思贤如渴，若跨有荆、益，保其岩阻，西和诸戎，南抚夷越，外结好孙权，内修政理；天下有变，则命一上将将荆州之军以向宛、洛，将军身率益州之众出于秦川，百姓孰敢不箪食壶浆以迎将军者乎？诚如是，则霸业可成，汉室可兴矣。”



预测概述

1.1 预测的概念

预测，是通过对客观事实历史和现状进行科学的调查和分析，由过去和现在去推测未来，由已知去推测未知，从而**揭示客观事实未来发展的趋势和规律**。预测是对尚未发生或目前还不确定的事物进行预先的估计和推断，是现时对事物将要发生的结果所进行的探讨和研究。

1.1 预测的概念

系统预测是以系统为研究对象，根据以往旧系统或类似系统的历史统计资料，运用现代科学方法理论以及各种经验、判断和知识，对系统中某些因素或系统发展趋势**进行推测，做出评价，提出措施方案**，使系统沿着有利的方向发展。

系统预测的方法和手段称为**预测技术**。

1.2 预测的产生与发展

- 预测是人类自古就有的活动。古人说：“凡事预则立，不预则废。”
- 公元前5世纪，范蠡扶助越王勾践战胜吴王夫差后，弃官经商，成为巨富，号陶朱公。他有句名言，叫作“论其有余不足，则知贵贱，贵上极则反贱，贱下极则反贵”。是对商品供求与价格变化之间关系的精辟论述。



商王问卜



范蠡离越

1.2 预测的产生与发展

- 《孙子兵法》中“兵者,国之大事,生死之地,存亡之道,不可不察也”,这个“察”包含了预测的意思。这部兵书历时2000多年长盛不衰,至今仍被军事战略家、企业家奉为经典,主要原因是它提供的种种预测方法,能够帮助人们进行正确决策。



孙子兵法



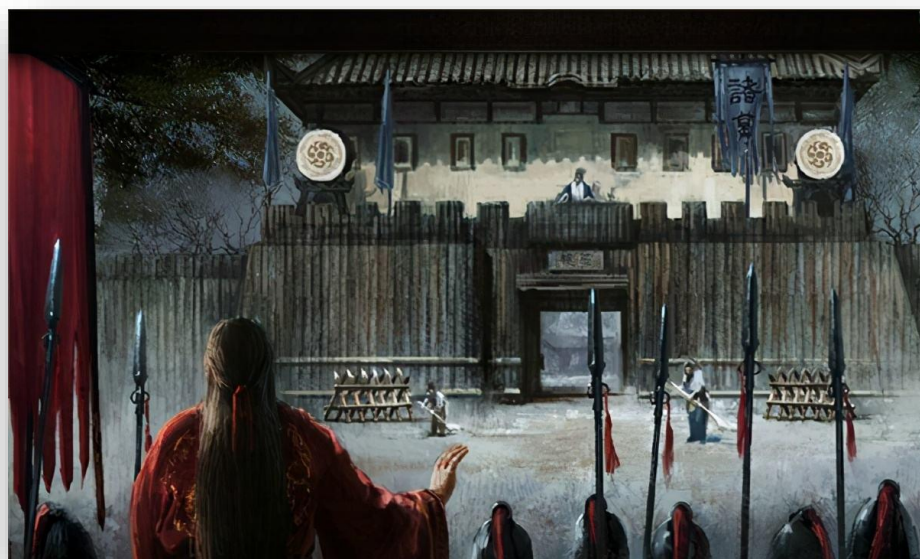
所谓兵者

1.2 预测的产生与发展

- 诸葛亮“借东风”胸有成竹，是基于他对当地气象变化的预测。
- “空城计”，是基于他对司马懿军事决策行为特点的分析和预测。



借东风



空城计

1.2 预测的产生与发展

预测作为一种社会实践活动,已有几千年的历史。它是适应社会经济的发展和管理需要而产生、发展起来的。

Jakob Bernoulli(1654-1705)创立了预测学, 目的在于减少人类生活各个方面由于不确定性导致错误决策所产生的风险。不确定性是几乎所有问题的根源,因此预测学将成为更有效的问题解决工具的关键。

1.2 预测的产生与发展

第二次世界大战以后由于科学技术和世界经济取得了前所未有的快速发展，社会经济现象的不确定因素显著增加，诸如政治危机、经济危机、能源危机等，所有这些不确定因素增加了人们从心理上[了解和掌握未来](#)的必要性和迫切性。

人们日益意识到科学预测的重要性，这也成为预测学科进一步发展的推动力。

1.3 预测的原则

□ 系统性原则

- 预测者所研究的事物和自然界的其他事物一样，都有自己的过去、现在和将来，也能是存在着一种纵向的发展关系——**因果关系**。而这种因果关系要受某种规律的支配。
- 预测者必须全面分析预测事物本身及与其本身有关联的所有因素的发展规律。将事物本身与周围的环境视为一个**系统综合体**进行研究。
- 系统性原则要求预测者只能客观、如实地反映预测对象及其相关因素的发展规律与组合方式，不能随意增减某些因素或改变其组合方式。

1.3 预测的原则

□ 关联性原则

- 预测对象与相关因素之间存在依存关系。不同的相关因素之间也可能存在某种依存关系。预测者应对这些关系进行全面分析。
- 关联性原则就是要求预测者充分考虑相关因素的联系及其作用与反作用的依存关系。如果不重视这一原则，顾此失彼，有可能导致严重的后果。

1.3 预测的原则

□ 动态性原则

- 预测对象的相关因素和环境不是一成不变的，而是处于不断发展变化的过程中。这些因素或环境的各个发展阶段对预测对象都有影响，有时甚至会改变预测对象的发展方向或性质。
- 相关因素或外部环境是预测对象内部矛盾性的外因。如果外因（或外部环境或相关因素）变化很平稳，或处于相对稳定的状态，则预测者可以利用历史数据进行外推，预测事物的发展，但是情况往往并非如此理想。自然灾害、资料缺失、意外变故，预测时都要充分考虑。

1.4 预测的分类

(1) 按预测的范围或层次分类

□ 宏观预测

是指针对国家或部门、地区的社会经济活动进行的各种预测。它以整个社会经济发展的总图景作为考察对象，研究社会经济发展中各项指标之间的联系和发展变化。宏观经济预测，是政府制定方针政策，编制和检查计划，调整经济结构的重要依据。

1.4 预测的分类

(1) 按预测的范围或层次分类

□ 微观预测

是针对基层单位的各项活动进行的预测。它以企业或个人生产经营发展的前景作为考察对象，研究微观经济中各项指标间的联系和发展变化。

微观经济预测，是企业制定生产经营决策，编制和检查计划的依据。

1.4 预测的分类

(2) 按预测的时间长短分类

- **近期预测**。一般指3个月以下的预测。它是制订月、旬计划和明确规定近期经济活动具体任务的依据。
- **短期预测**。一般指3个月以上1年以下的预测。它是制订年度计划、季度计划和明确规定短期经济发展具体任务的依据。
- **中期预测**。一般指1年以上5年以下的预测。它是制订经济发展五年计划，规划经济五年发展任务的依据。
- **长期预测**。一般指5年以上的预测。它是制订经济发展长期计划、远景规划，规定经济长期发展任务的依据。

1.4 预测的分类

(3) 按预测的内容分类

- 社会预测，即对有关社会问题的预测。
- 经济预测，即对经济领域问题的预测。
- 科技预测，即对科学技术问题的预测。
- 军事预测，即对国防战争方面问题的预测。

1.4 预测的分类

(4) 按预测时是否考虑时间因素进行分类

□ **静态预测**。是指不包含时间变动因素，根据事物在同一时期的因果关系进行预测。

□ **动态预测**。是指包含时间变动因素，根据事物发展的历史和现状，对其未来发展前景做出的预测。

1.4 预测的分类

(5) 按预测技术的性质分类

- 定性预测
- 定量预测
- 灰色预测
- 回归分析预测
- 时间序列预测

□ 定性预测

定性预测方法是一种依靠人的主观判断预测未来的方法，主要是定性地估计某一事件的发展趋势、优劣程度和发生的概率。预测是否准确完全取决于预测者的知识和经验。定性预测一般用于对缺乏历史统计资料的事件进行预测。

- 在定量分析之前首先进行定性分析，明确发展趋势，为定量分析做准备工作。
- 在缺乏定量预测的数据时，直接进行预测。
- 与定量分析方法结合使用，以提高预测的可靠程度。
- 对定量预测的结果进行评价。

□ 定性预测

定性预测的主要方法：

①**德尔菲法**：也称专家调查法，1946年由美国兰德公司创始实行，其大致流程是在对所预测的问题征得专家的意见之后，进行整理、归纳、统计，再匿名反馈给各专家，再次征求意见，再集中，再反馈，直至得到一致的意见。

②**市场调查法**：市场调查法是一种企业组织有关人员进行市场调查分析确定促销效果的方法。这种方法比较适合于评估促销活动的长期效果。它包括确定调查项目和调查法的实施方式两方面内容。

□ 定性预测

③**头脑风暴法**：由美国BBDO广告公司的亚历克斯·奥斯本首创，该方法主要由价值工程工作小组人员在正常融洽和不受任何限制的气氛中以会议形式进行讨论、座谈，打破常规，积极思考，畅所欲言，充分发表看法。

④**类比法**：类比法是按同类事物或相似事物的发展规律相一致的原则，对预测目标事物加以对比分析，来推断预测目标事物未来发展趋向与可能水平的一种预测方法。类比法应用形式很多，如由点推算面、由局部类推整体、由类似产品类推新产品、由相似国外国际市场类推国内国际市场等。

□ 定量预测

定量预测法，是指在大量掌握与预测对象有关的各种信息资料的基础上，运用数学方法对资料进行处理，**建立能够反映各种变量之间规律性联系的数学模型**的预测过程。

①趋势外推法，是指根据预测对象的发展规律，结合各种制约条件对预测对象的未来发展进行分析判断的一种预测方法。

②因果预测法，是指根据各个变量之间的因果关系建立数学模型，对预测对象未来发展趋势的预测。

1 预测概述

□ 定量预测

(1) 简单平均法

简单平均法是使用统计中的简单算术平均数的方法进行的预测法。它是以历史数据为依据,进行简单平均得出的:

$$x = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

式中, x -预测的平均值;

$x_1, x_2, \dots, x_n$ -各个历史时期的实际值;

n -时期数。

□ 定量预测

(2) 趋势平均法

趋势平均法是假设未来时期的销售量是与其接近时期的销售量的直接延伸，而与较远时期的销售量关系较小，同时为了尽可能缩小偶然因素的影响，可用最近若干时期的平均值作为预测期的预测值的基础。

趋势平均法将前后各期对预测值的影响一视同仁，认为所有使用的信息不管其是过去还是近期的资料对预测结果的影响是一样的。

□ 定量预测

(3) 加权移动平均法

使用加权移动平均法就是在计算平均数时,使用一个权数来计算。

$$x = \sum f_x / f$$

其中, f 表示权数。

权数的选择是一个比较重要的因素。一般情况下,近期的权数比较大,远期的权数比较小。

□ 定量预测

(4) 指数平滑法

指数平滑法对前后各期区别对待, 分别给予不同的权数, 考虑到近期信息对预测值的影响比远期大, 因而越是近期的信息, 权数越大, 而远期的权数小。这样计算的结果可比趋势平均法更客观。

$$Y_t = aA_{t-1} + (1 - a)Y_{t-1}$$

式中 Y_t 表示本期预测值; A_{t-1} 表示上期的实际销售额;
 a 为平滑系数, 其取值范围为 $0 < a < 1$ 。

□ 定量预测

(5) 百分比率递增法

该方法是根据历史资料,用环比发展速度的序时平均计算出其每年增长的百分比来预测该项目的未来期望值。计算公式:

$$s = p(1+x)^n$$

式中 s -预测值; p -基期实际值;

x -年平均递增率; n -期数。

如果只计算其增长部分,则其计算公式为:

$$g = p[(1+x)^n - 1]$$

□ 灰色预测

灰色系统理论是邓聚龙教授1982年首次提出并创立的,灰色系统理论是控制论的观点和方法应用于社会经济系统,也是控制理论与运筹学等相结合的产物。

传统的预测方法大多采用回归分析法,需要大量的数据,且要求分布典型、计算量大、预测结果误差大,因而不便于实施目标管理和决策。而灰色预测模型可在信息不完全的情况下,进行较为精确的预测,单数据微分模型有较好的拟合和外推特征,具有**所需样本数据少,运算简便**等特点。

□ 回归分析预测

回归分析是英国生物学家高尔登(F.Galton)在研究人类身高的遗传特性时首先提出来的。通过大量的统计资料,高尔登发现,身材高的父亲,他的儿子身材比其略矮;反之,身材矮的父亲,他的儿子身材比其略高。即人类的身高有向平均数靠近的倾向。这种现象称之为回归。

回归分析预测方法是一种因果分析预测。它是研究某一个随机变量与一个或几个变量之间的数量关系,用一个或几个非随机变量来预测某一个随机变量的方法。

□ 时间序列预测

时间序列,是指某一变量或指标的数值或观测值,按照其出现时间的先后顺序,以相同的间隔时间排列的一组数值。时间序列中不仅包含系统所有变量过去的信息,而且还包含参与系统演化的所有变量的大量信息。

预测要求根据大量的观测数据对系统进行分析、预测出系统未来的特性以便对系统的特性进行处理或控制。时间序列的预测是用被预测事物过去和现在的观测数据,构造依时间变化的序列模型,并借助一定的规则推测未来,是预测领域的重要组成部分。

1.5 预测的应用领域和发展现状

- 20世纪50年代以来,预测学渐渐地形成一门独立的学科,国内外各部门、各行业不断应用各种预测理论和方法来进行社会预测、经济预测、科学预测、技术预测和军事预测等。
- 预测学广泛应用于人口环境、资源、教育、金融、交通运输、市规划、医药卫生、材料科学和科技管理等领域。

1.5 预测的应用领域和发展现状

在矿业工程领域中，需要预测的方面及参数很多。

- 顶板抗压强度、抗压时间、抗压步距的估算；
- 支护阻力的估计；
- 围岩变形预测；
- 涌水量预测；
- 硫化矿自燃预测；
- 安全生产状况的推断；
- 矿石价格预测；
- 矿石产量及经济效益的预测等。

1.6 预测的程序

- ① 明确预测任务，制定预测计划；
- ② 搜集、审核和整理资料；
- ③ 选择预测方法和建立数学模型；
- ④ 利用模型进行预测；
- ⑤ 模型检验，分析误差，评价预测结果；
- ⑥ 向决策者提交预测报告



主要内容

第一节 预测概述

第二节 定性预测技术

第三节 时间序列预测

第四节 灰色预测技术

2.1 定性预测的定义

- **定性预测技术**是指预测者依靠熟悉业务知识、具有丰富经验和综合分析能力的人员与专家，根据已掌握的历史资料和直观材料，对事物的未来发展做出性质和程度上的判断，然后综合各方面的意见，作为预测未来的主要依据。
- **定性预测主要用于研究和探讨对象在未来所表现的性质、状态**。事物发展的总体趋势、事件发生和发展的各种可能性及其造成的影响，以及目前确定并将要执行的决策是否会达到制定决策的目的等。这类预测主要是凭借预测者的主观经验和逻辑推理能力，对事物未来表现的性质进行推测和判断。

2.2 定性预测方法的特点

- (1) 着重对事物发展的性质进行预测，主要凭借人的经验以及分析能力；
- (2) 着重对事物发展的趋势、方向和重大转折点进行预测。
 - 优点：注重于事物发展在性质方面的预测，具有较大的灵活性，易于充分发挥人的主观能动作用，且简单的迅速，省时省费用。
 - 缺点：易受主观因素的影响，比较注重于人的经验和主观判断能力，从而易受人的知识、经验和能力的多少大小的束缚和限制，尤其是缺乏对事物发展作数量上的精确描述。

2.3 定性预测方法的分类

- 状态列举型预测
- 状态判定型预测
- 数值型预测
- 状态发生概率型预测

(1) 状态列举型预测

只要能够预计到列举的状态就可认为预测正确



(2) 状态判定型预测

判定具体发生的状态，要求预测者给出是与非的回答。



(3) 数值型预测

对未来事物发展的某个量的预测，一般用预测误差来评价预测方法的好坏。



(4) 状态发生概率型预测

预测各状态发生的概率。



2.4 定性预测与定量预测的比较

(1) 预测现有趋势延续或转折的能力

定量预测不能识别预测趋势的转折。

如，经济的增长一般是波动性增长，而不是一直持续性增长，仅仅应用定量预测方法不能预测出经济发展的波动。趋势的转折是暂时的还是长期的，必须依赖于定性判断。

2.4 定性预测与定量预测的比较

(2) 信息应用的充分性

定量预测不能充分运用历史数据所包含的信息。

- ✓ 时间序列方法一般容易忽视较早的历史数据所包含的信息，而过分依赖近期数据。
- ✓ 回归分析则轻视近期数据所包含的信息，它赋予所有历史数据相同的权重。

定性预测可以充分利用各种信息，包括有关预测环境信息、过去类似案例及其失误等，使决策者可以充分利用他的经验，并结合各种信息作出判断。

2.4 定性预测与定量预测的比较

(3) 预测的客观性

定量预测的最大优点在于其具有客观性。只要选择好模型,任何人应用同样的数据都会得出相同的预测结果。

定性预测则不一样,根据同样的信息,不同的人可得出完全不同的结论。定性预测受决策者性格、情绪和阅历等方面的影响。

2.4 定性预测与定量预测的比较

表6-1 定性预测与定量预测对应比较表

方法	适用情况	应做工作
定性预测法	对缺乏历史统计资料或趋势面临转折的事件	需要做大量的调研工作
回归预测法	自变量与因变量之间存在相关关系	为所有变量收集历史资料
移动平均法	不带季节变动的预测	预测变量的历史资料
指数平滑法	不带季节变动的预测	预测变量的历史资料
灰色预测法	适用于时序发展呈指数趋势	预测变量的历史资料



主要内容

第一节 预测概述

第二节 定性预测技术

第三节 时间序列预测

第四节 灰色预测技术

3.1 时间序列预测

- 时间序列分析是一种数据处理的动态统计方法，主要用于描述和探索现象随时间时发展变化的数量规律性。时间序列分析方法由Box-Jenkins于1976年提出，适用于各种领域的时间序列分析。
- 时间序列预测是指根据事物的历史和现实数据，寻求事物随时间推移而发展变化的规律，从而推测未来状况的常用预测方法。
- 时间序列预测基于随机过程理论和数理统计学方法，研究随机数据序列所遵从的统计规律，用于解决实际问题。包括一般统计分析，统计模型的建立与推断以及关于时间序列的最优预测、控制与滤波等内容。

3 时间序列预测

3.2 移动平均法

● 简单移动平均法

设时间序列为： $y_1, y_2, \dots, y_n$ ， N 为间隔期，则简单移动平均公式为：

$$M_t = S_t = \frac{1}{N} (y_t + y_{t-1} + \dots + y_{t-N+1})$$

其中， $t \geq N$ ， y_t 表示 t 时刻的真实值或观察值； M_t 为 t 期移动平均数；上式也可以写成下式：

$$M_t = \frac{1}{N} (y_t - y_{t-N}) + M_{t-1}$$

预测公式为： $\hat{y}_{t+1} = M_t$

即以第 t 期移动平均数作为第 $t+1$ 期的预测值。

3 时间序列预测

【例】我国2000-2012年工业生产者出厂价格总指数，如表2.1所示。
试用简单移动平均法，预测2013年的工业生产者出厂价格总指数。

解：分别取 $N=3$ 和 $N=4$ ，按简单移动预测公式得

$$\hat{y}_{t+1} = \frac{1}{3} (y_t + y_{t-1} + y_{t-2})$$

和
$$\hat{y}_{t+1} = \frac{1}{4} (y_t + y_{t-1} + y_{t-2} + y_{t-3})$$

计算3-4年移动平均预测值、预测误差，如表6-2所示。

其预测曲线如下图所示。

表6-2 中国工业生产者出厂价格总指数

年份	工业生产者 出厂价格分 类指数 (%)	3年移动平 均 预测值	相对误 差 %	4年移动平均 预测值	相对误 差 %
2000	102.8				
2001	98.7				
2002	97.8				
2003	102.3	99.77	2.48		
2004	106.1	99.60	6.13	100.40	5.37
2005	104.9	102.07	2.70	101.23	3.50
2006	103	104.43	1.39	102.78	0.22
2007	103.1	104.67	1.52	104.08	0.95
2008	106.9	103.67	3.02	104.28	2.46
2009	94.6	104.33	10.29	104.48	10.44
2010	105.5	101.53	3.76	101.90	3.41
2011	106	102.33	3.46	102.53	3.28
2012	98.3	102.03	3.80	103.25	5.04
2013		103.27		101.10	

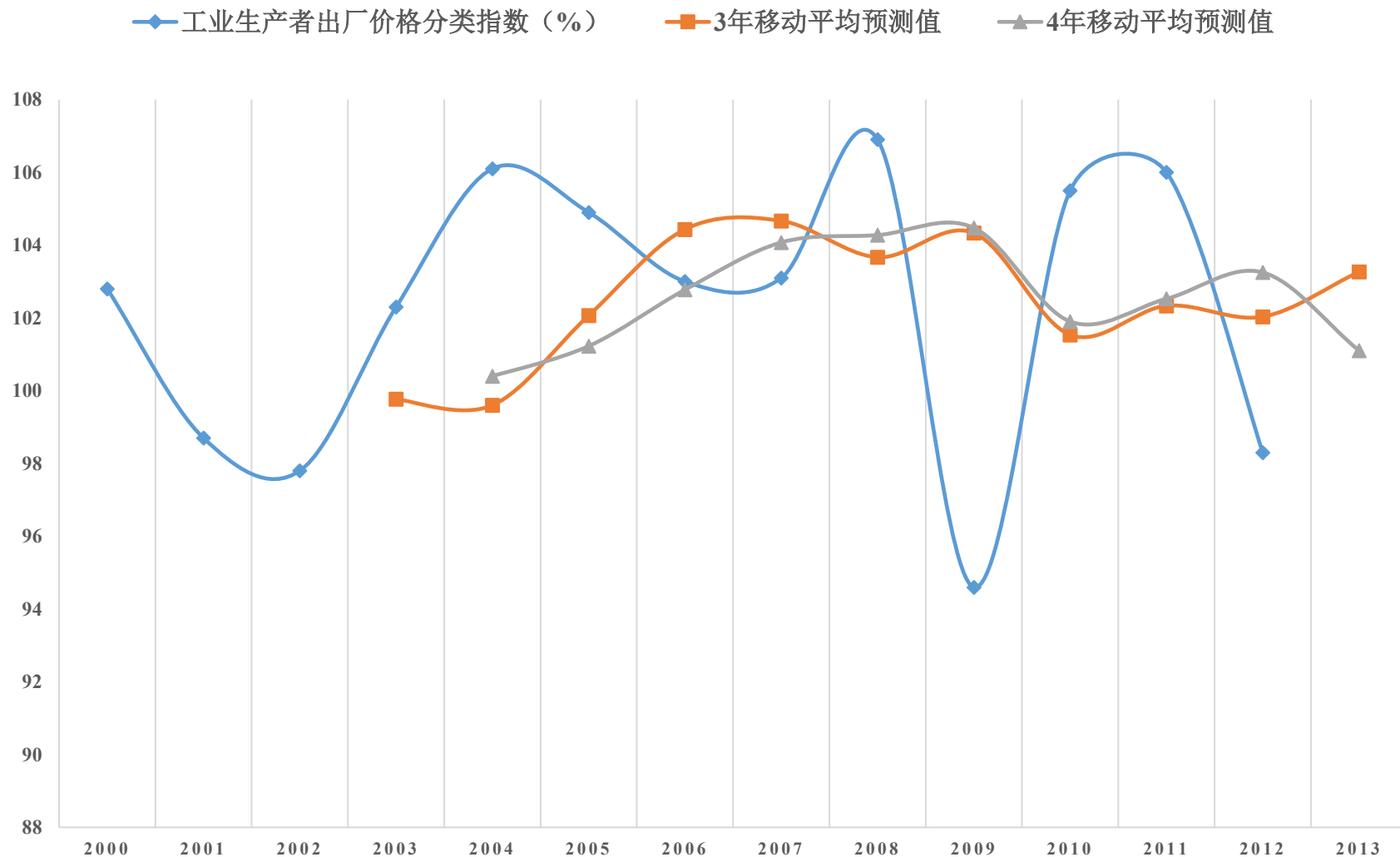


图6-5 2000-2012年中国工业生产者出厂价格总指数和

3 时间序列预测

3.2 移动平均法

- 趋势移动平均法

一次移动的平均数为： $M_t^{(1)} = \frac{1}{N} (y_t + y_{t-1} + \cdots + y_{t-N+1})$

在一次移动平均的基础上再进行一次移动平均就是二次移动平均，其计算公式为： $M_t^{(2)} = \frac{1}{N} (M_t^{(1)} + M_{t-1}^{(1)} + \cdots + M_{t-N+1}^{(1)})$

它的递推公式为： $M_t^{(2)} = M_{t-1}^{(2)} + \frac{1}{N} (M_t^{(1)} - M_{t-N}^{(1)})$

3 时间序列预测

3.2 移动平均法

- 趋势移动平均法

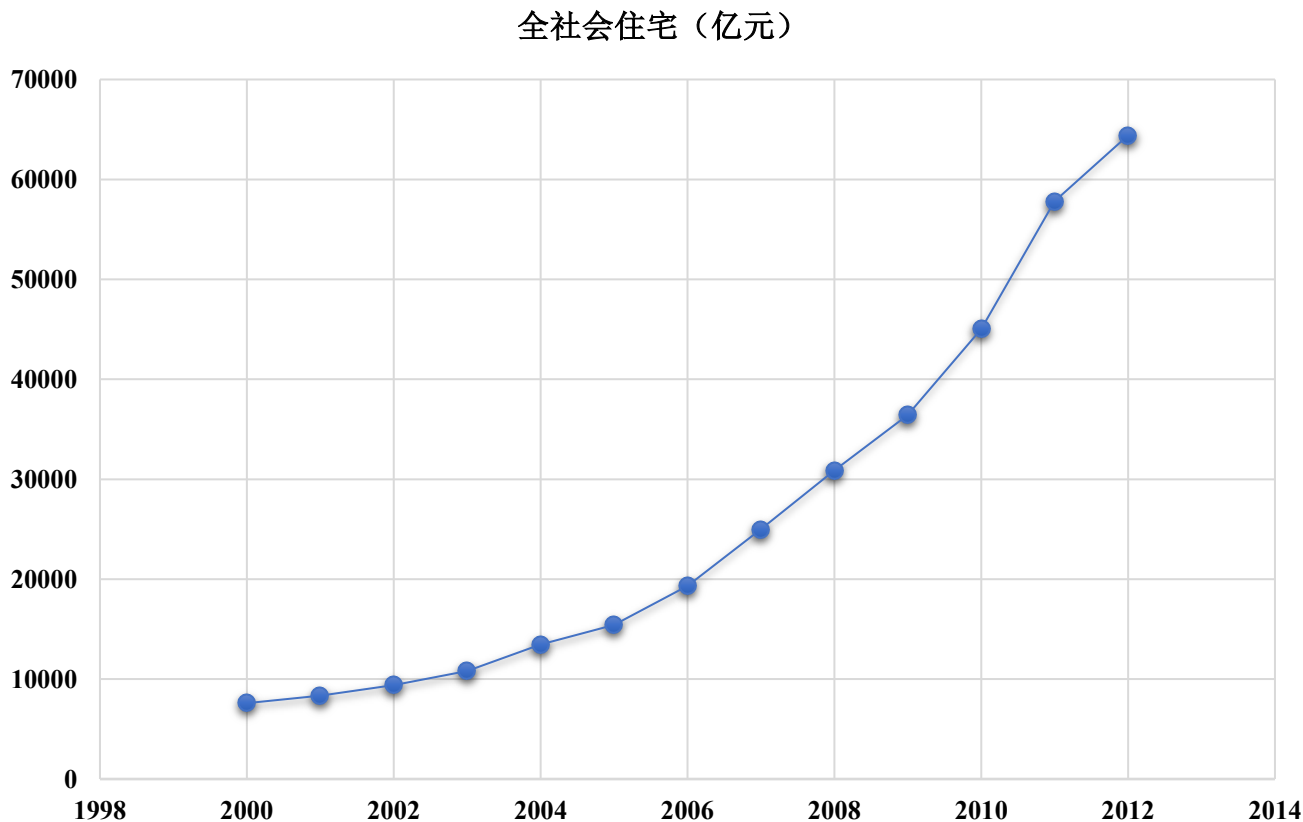
设时间序列 $\{y_t\}$ 从某时期开始具有直线趋势，且认为未来时期也按此直线趋势变化，则可用下式直线趋势模型进行预测。

$$\hat{y}_{t+T} = a_t + b_t T$$
$$\begin{cases} a_t = 2M_t^{(1)} - M_t^{(2)} \\ b_t = \frac{2}{N-1} (M_t^{(1)} - M_t^{(2)}) \end{cases}$$

其中， t 为当前时期数； T 为由 t 至预测期的时期的时期数； a 为截距； b 为斜率。两者又称为平滑系数。

【例】我国2000-2012年全社会住宅额，如表2.1所示。试用趋势移动平均法预测2013-2014年的全社会住宅额。

解：由散点图2.6可以看出，我国全社会住宅总值基本呈直线上升趋势，可用趋势移动平均法来预测。



取 $N=3$,分别计算一次移动平均值、二次移动平均值,如下表所示。

表6-3 我国全社会住宅及一、二次移动平均值计算表

年份	全社会住宅（亿元）	一次移动平均 $N=3$	二次移动平均 $N=3$
2000	7594.1		
2001	8339.1		
2002	9407.1	8446.77	
2003	10792.3	9512.83	
2004	13464.1	11221.17	9726.92
2005	15427.2	13227.87	11320.62
2006	19333.1	16074.80	13507.94
2007	25005	19921.77	16408.14
2008	30881.2	25073.10	20356.56
2009	36428.2	30771.47	25255.44
2010	45027	37445.47	31096.68
2011	57824.4	46426.53	38214.49
2012	64412.8	55754.73	46542.24

$$a_{12} = 2M_{2012}^{(1)} - M_{2012}^{(2)} = (2 \times 55754.73 - 46542.24) = 64967.2(\text{亿元})$$

$$b_{12} = \frac{2}{3-1} \left(M_{2012}^{(1)} - M_{2012}^{(2)} \right) = \frac{2}{2} (55754.73 - 46542.24) = 9212.5(\text{亿元})$$

$T=2012$ 时，直线趋势预测模型为

$$\hat{y}_{2012+T} = 64967.2 + 9212.5T$$

预测2013-2014年的我国全社会住宅总值为

$$\hat{y}_{2013} = 64967.2 + 9212.5 \times 1 = 74179.7(\text{亿元})$$

$$\hat{y}_{2014} = 64967.2 + 9212.5 \times 2 = 83392.2(\text{亿元})$$

3.3 指数平均法

- 一次指数平滑法

预测模型为：

$$\hat{y}_{t+1} = S_t^{(1)}$$

即

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha) \hat{y}_t$$

就是以第 t 期指数平滑值作为 $t+1$ 期预测值。

3

时间序列预测

试运用一次指数平滑方法预测2013年工业生产者出厂价格总指数。

解：分别取 $\alpha=0.3$ ， $\alpha=0.8$ ，

初始值 $S_0^{(1)} = \frac{y_1 + y_2}{2} = 100.75$

依据一次指数平滑预测模型

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha) \hat{y}_t$$

$$\hat{y}_1 = S_0^{(1)} = 100.75$$

计算各期预测值、预测误差，预测曲线如下表。

当 $\alpha=0.3$ ，平均相对误差为3.41%；

当 $\alpha=0.8$ ，平均相对误差为3.97%。

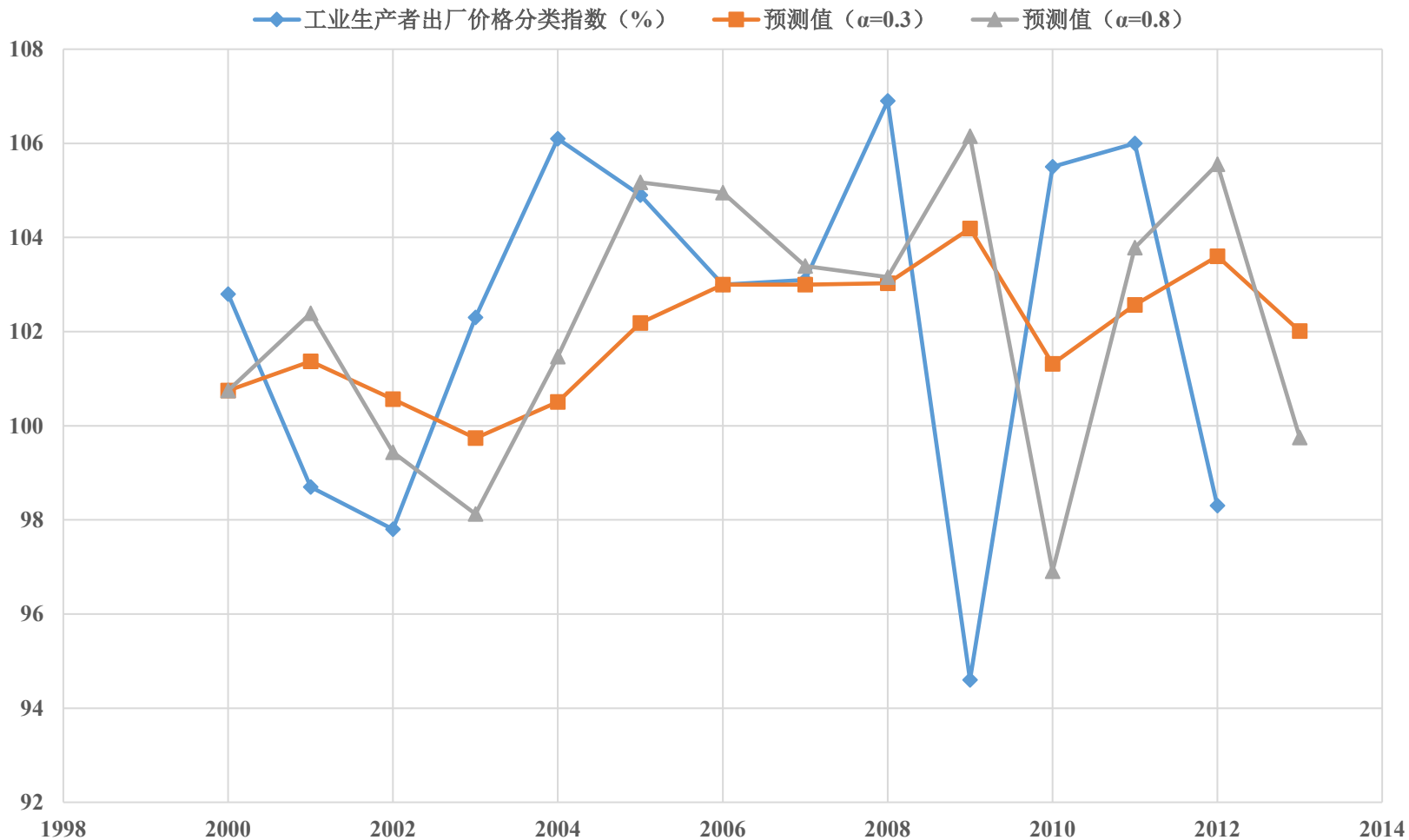
显然 $\alpha=0.3$ ，平均相对误差较小，故选取 $\alpha=0.3$ 。

预测2013年中国工业生产者出厂价格总指数为：

$$\hat{y}_{2013} = 102.01(\text{万元})$$

表6-4 中国工业生产者出厂价格总指数

年份	工业生产者出厂价格 分类指数 (%)	预测值 ($\alpha=0.3$)	相对误差 (%)	预测值 ($\alpha=0.8$)	相对误差 (%)
2000	102.8	100.75	1.99	100.75	1.99
2001	98.7	101.37	2.70	102.39	3.74
2002	97.8	100.57	2.83	99.44	1.67
2003	102.3	99.74	2.51	98.13	4.08
2004	106.1	100.51	5.27	101.47	4.37
2005	104.9	102.18	2.59	105.17	0.26
2006	103	103.00	0.00	104.95	1.90
2007	103.1	103.00	0.10	103.39	0.28
2008	106.9	103.03	3.62	103.16	3.50
2009	94.6	104.19	10.14	106.15	12.21
2010	105.5	101.31	3.97	96.91	8.14
2011	106	102.57	3.24	103.78	2.09
2012	98.3	103.60	5.39	105.56	7.38
2013		102.01		99.75	



3 时间序列预测

3.4 非线性时间序列预测

时间序列数据存在于自然科学和社会科学的各个领域，如天文观测、股市每日的交易价格等。

时间序列预测就是根据被预测事物的过去和现在的观测数据，构造特定的模型，并借助一定的规则推测未来。

时间序列预测一般分为线性时间序列预测和非线性时间序列预测两类。

3.4 非线性时间序列预测

在线性时间序列预测领域，1970年Box和Jenkins出版的专著《time series analysis: forecasting and control》提出了线性自回归滑动平均(ARMA)模型的辨别、参数估计、适用性检验等一整套理论，标志着线性时间序列预测的理论已经日臻成熟。

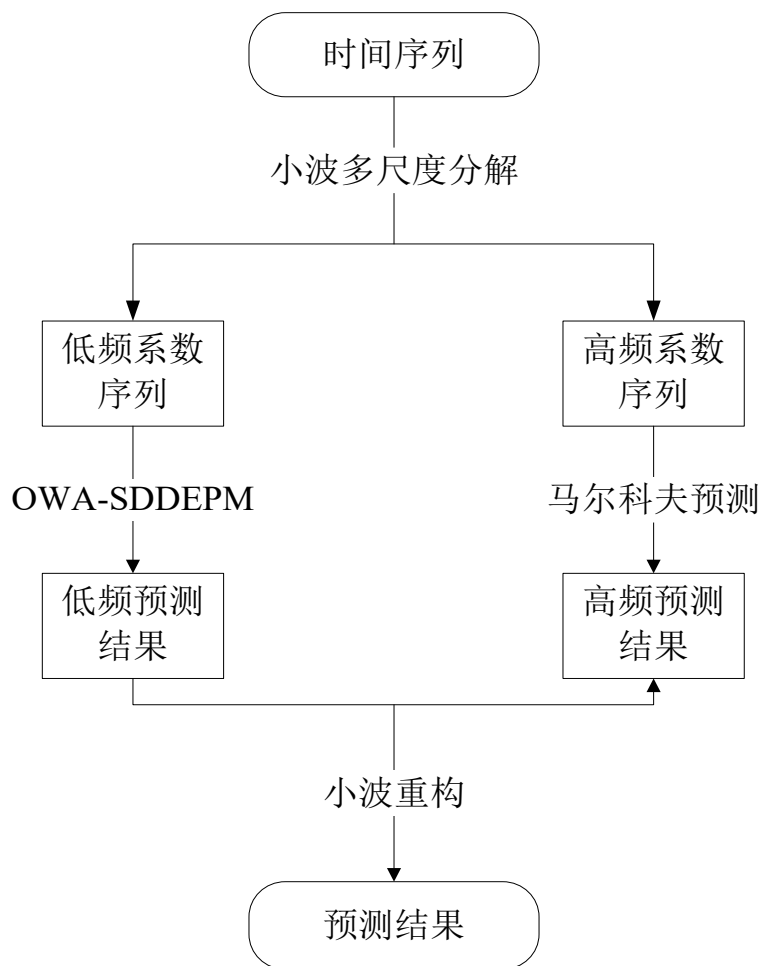
现实世界中许多时间序列数据呈现出很强的非线性特征，特别是对一些非线性复杂系统的研究中。为此，大量非线性预测模型被相继提出，如神经网络预测模型、支持向量预测模型等，利用非线性模型对时间序列数据进行预测已经成为国内外研究的热点。

3.4 非线性时间序列预测

不稳定性的非线性时间序列的预测问题产生于多个领域，在信号处理、地震波分析以及社会—经济系统中，非平稳序列较之平稳序列更为常见，因此，近年来非平稳时间序列分析问题倍受关注。

首先，对时间序列数据进行小波多尺度分解，提取出原时间序列的低频趋势序列和高频细节波动序列，然后分别利用基于OWA算子的滑动离散差分方程预测模型(OWA-SDDEPM)和马尔可夫预测模型对低频和高频序列进行预测，接着对各层的预测结果进行小波重构，便得到了原时间序列的预测值。

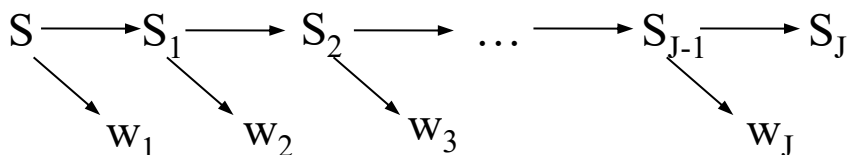
非平稳时间序列分析



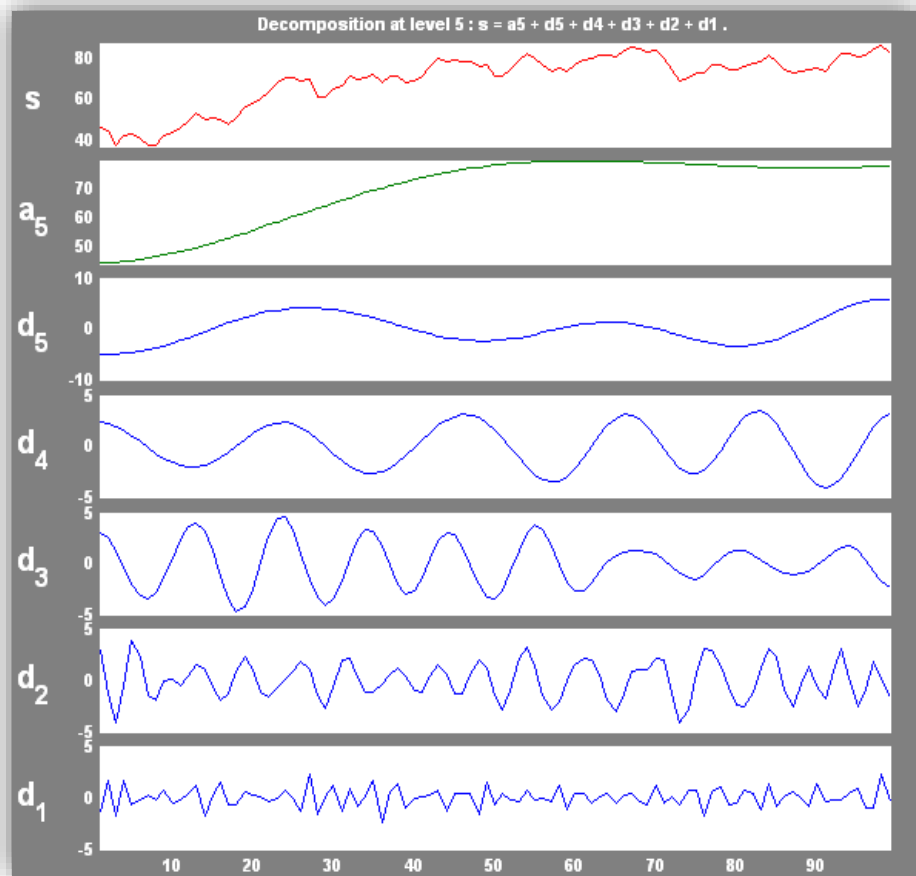
牙+鼻子+耳朵+腿+身体+尾巴=大象

非平稳时间序列分析

小波分解



小波分解，可以将一个复杂的时间序列信号分解成一系列不同频率的简单信号，分解后的各层信号具有单一的频率，从而具有更好的平稳性。



3 时间序列预测

实证分析—国际原油价格的预测

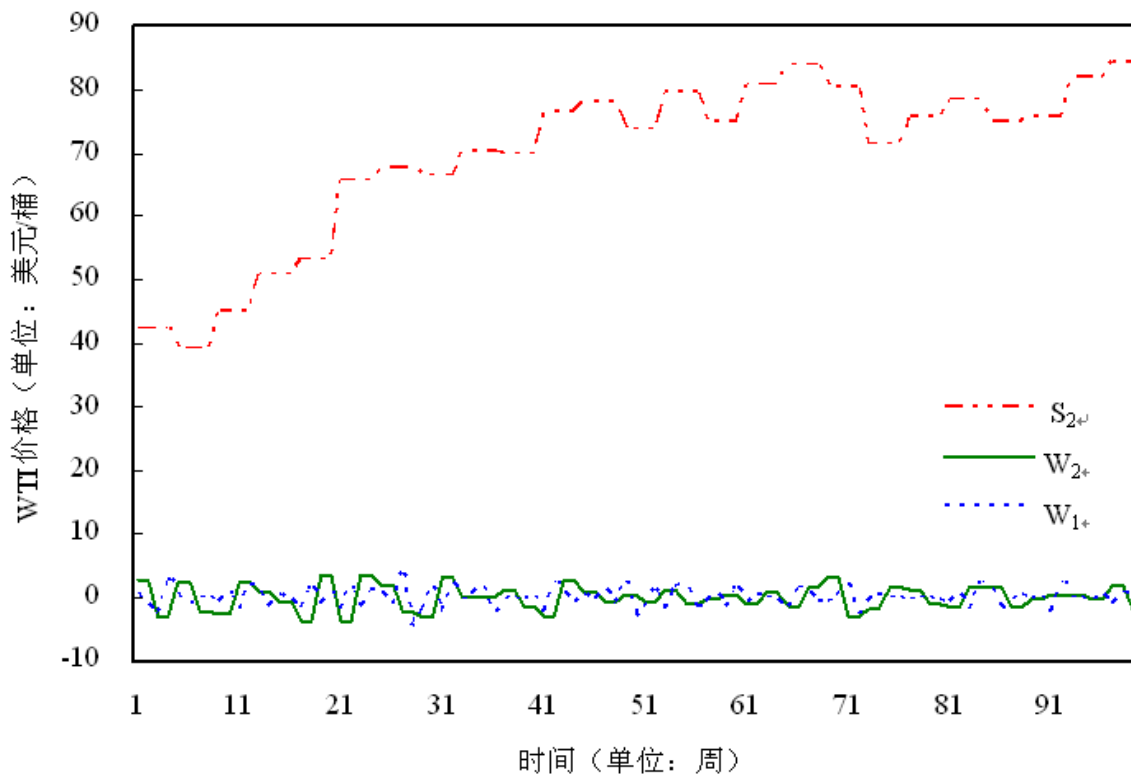
在国际原油市场上，消耗与供应是调节原油价格的主要因素，但因为能源有其本身的特殊性，原油价格还受到国际政治格局、消费模式、人口增长以及突发事件等

各个方面的影响。近几十年来国际原油价格骤升骤降，所以对国际原油价格进行及时有效的预测面临很大的困难和挑战。

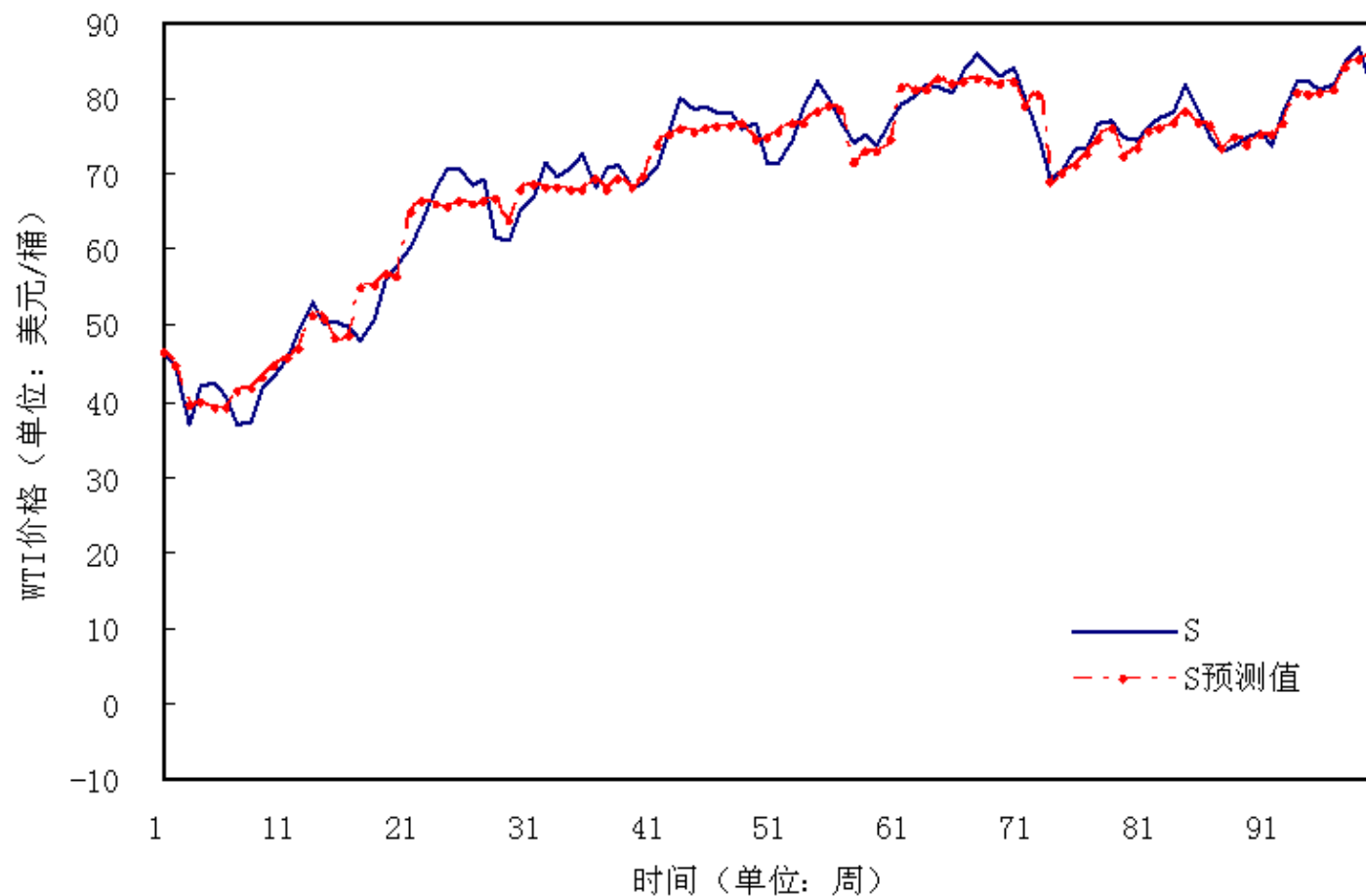


实证分析—国际原油价格的预测

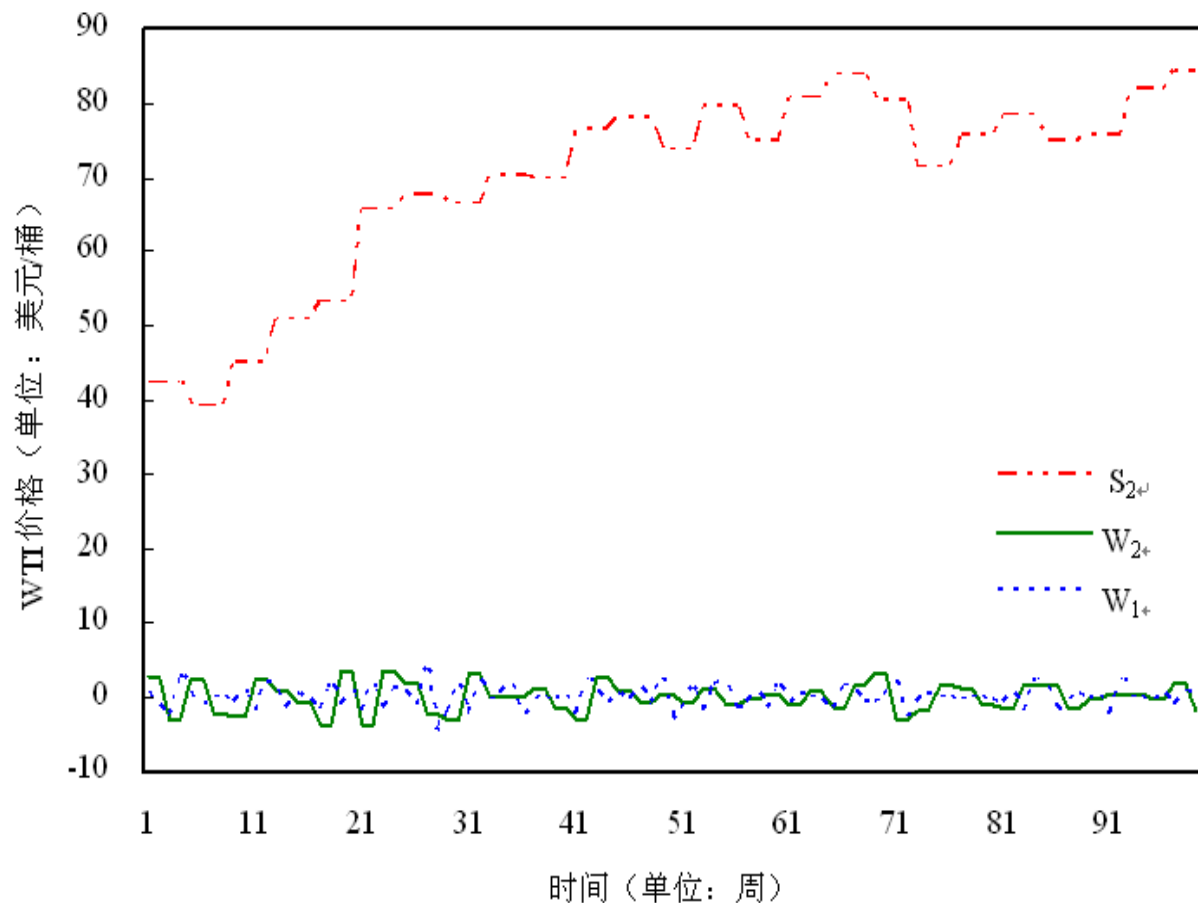
选取2009年1月2日至2010年11月19日共99周的WTI原油价格，将其小波分解，得



预测结果



预测效果对比



小波-余弦预测模型的特点是能更好的反映时间序列的周期性，小波-AR预测模型则适用于预测小波分解后各层都比较平稳的时间序列，而对于周期性不明显和不稳定性的非线性时间序列，此方法具有更好的预测效果。

分别采用平均绝对误差(MAE)、误差平方和(SSE)、均方误差(MSE)、平均绝对百分比误差(MAPE)和均方百分比误差(MSPE)五种常用的误差指标来评价五种模型的预测精度。

	MAE	SSE	MSE	MAPE	MSPE
本文预测模型	2.07	614.10	0.25	3.21×10^{-2}	4.15×10^{-3}
小波-AR	3.91	2358.37	0.49	5.90×10^{-2}	7.96×10^{-3}
小波-余弦	3.84	2276.47	0.48	5.73×10^{-2}	7.28×10^{-3}
GM(1,1)	6.82	6349.72	0.80	11.21×10^{-2}	15.04×10^{-3}
GARCH	6.70	7906.13	0.90	11.75×10^{-2}	18.11×10^{-3}



主要内容

第一节 预测概述

第二节 定性预测技术

第三节 时间序列预测

第四节 灰色预测技术

邓聚龙教授首创灰色系统理论

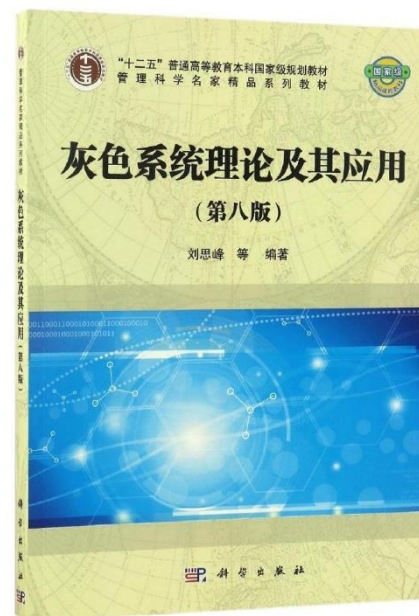
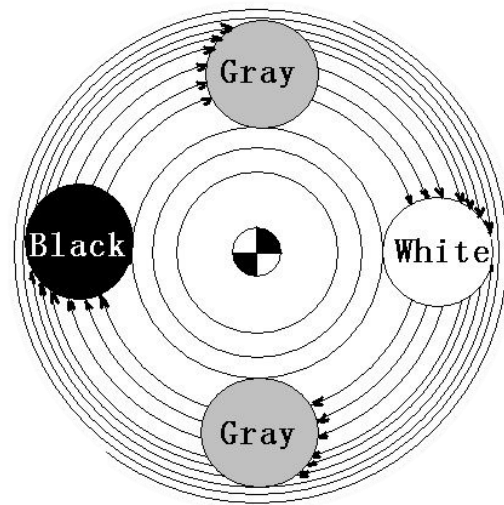
2007年，在首届IEEE灰色系统与智能服务国际会议上，邓聚龙教授荣获灰色系统理论创始人奖。



邓聚龙教授

4.1 灰色预测概述

- 灰色系统：部分信息已知，部分信息未知的系统。
- 比如人体是一个灰色系统，人体某些外形参数如身高、体重等是已知的，某些内部参数如血压、脉搏等也是已知的；但有更多的信息是未知的，如人体信息网络、人体功能机制等。
- 严格说来，灰色系统是绝对的，而白色系统和黑色系统是相对的。
- 灰色预测：基于灰色系统理论(简称灰色理论或灰理论)的GM(1, 1)模型的预测。



4.2 灰色预测的基本思想

当一时间序列无明显趋势时，采用累加的方法可生成一趋势明显的时间序列。

如时间序列 $x^{(0)} = (32, 38, 36, 35, 40, 42)$ 的趋势并不明显，但将其元素进行“累加”后所生成的时间序列

$$x^{(0)} = (32, 70, 106, 141, 181, 223),$$

则是一趋势明显的数列，按该数列的增长趋势可建立预测模型并考虑灰色因子的影响进行预测，然后采用“累减”的方法进行逆运算，恢复原时间序列，得到预测结果，这就是灰色预测的基本思想。

4.3 灰色预测的类型

灰色预测按其应用的对象不同可分为五种类型：

①**数列预测**。这类预测是对系统行为特征值大小的发展变化所进行的预测，称为系统行为数据列的变化预测，简称数列预测。如粮食产量的预测等。

②**灾变预测**。对系统行为的特征值超过某个阈值(界限值)的异常值将在何时再出现的预测称为灾变预测。所以灾变预测就是对**异常值出现时刻**的预测。

4 灰色预测技术

③**季节性突变预测**。若系统行为特征量异常值的出现，或某种事件的发生是在一年中某个特定的时区，则这种预测称为季节性灾变预测。

④**拓扑预测**。这类预测是对一段时间内系统行为特征数据波形的预测。

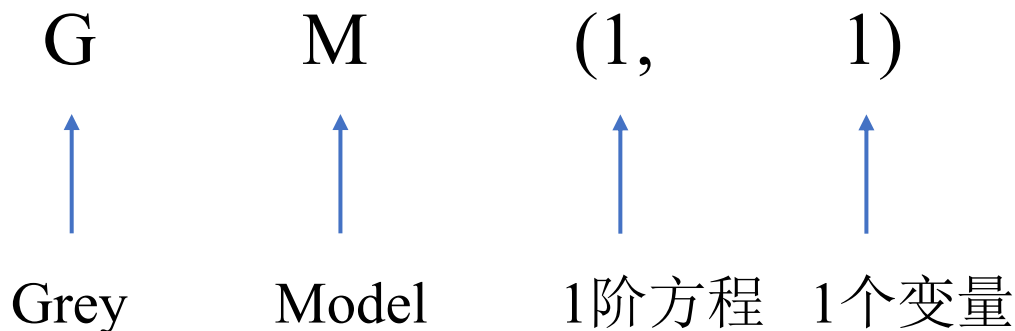
⑤**系统预测**。系统预测是指对同一系统多种行为变量的预测。如对粮食系统中粮食亩产量、粮食总产量、化肥施用量等的预测就是系统预测。系统预测模型不同于前面四种预测模型，它是基于一串相互关联的GM(1, 1)模型。

4 灰色预测技术

4.4 GM模型机理

GM模型：灰色理论的微分方程模型。

一般建模是用数据列建立差分方程，而灰色建模则是将原始数据列作生成处理后建立**微分方程**。



GM(1, 1)模型的建模

GM(1, 1)模型是生成模型**生成数**:

分为**累加生成数 (AGO)** 与**累减生成数 (IAGO)**

累加生成数 1-AGO指一次累加生成。

记原始序列为 $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$

生成序列为 $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}$

上标“0”表示原始序列，上标“1”表示一次累加生成序列。其中，

$$x^{(1)}(k) = \sum_{i=0}^k x^{(0)}(i) = x^{(1)}(k-1) + x^{(0)}(k)$$

4 灰色预测技术

GM(1,1)模型的建模步骤

令 $X^{(0)}$ 为GM(1,1)建模序列, $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$

$X^{(1)}$ 为 $X^{(0)}$ 的1-AGO序列,

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

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)$$

$$k = 1, 2, \dots, n$$

令 $Z^{(1)}$ 为 $X^{(1)}$ 的紧邻均值 (MEAN) 生成序列

$$Z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n))$$

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

则 GM(1,1) 的定义型, 即 GM(1,1) 的灰微分方程模型为

$$x^{(0)}(k) + az^{(1)}(k) = b$$

式中 a 称为发展系数, b 为灰色作用量。设 $\hat{\alpha}$ 为待估参数向量, 即 $\hat{\alpha} = (a, b)^T$, 则灰微分方程的最小二乘估计参数列满足

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

其中

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \dots & \dots \\ -z^{(1)}(n) & 1 \end{bmatrix}, \quad Y_n = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \dots \\ x^{(0)}(n) \end{bmatrix}$$

4 灰色预测技术

经过证明与推导，得到灰微分方程的解为：

$$\hat{x}^{(1)}(k+1) = [x^{(1)}(0) - \frac{b}{a}]e^{-ak} + \frac{b}{a}, \quad k = 1, 2, \dots, n$$

取 $x^{(1)}(0) = x^{(0)}(1)$ ，则

$$\hat{x}^{(1)}(k+1) = [x^{(0)}(1) - \frac{b}{a}]e^{-ak} + \frac{b}{a}, \quad k = 1, 2, \dots, n$$

还原值

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

灰色系统模型的检验方法：常见的有后验差检验，关联度检验，残差检验等。其中残差检验是一种较为客观的检验方法，GM(1,1)模型的检验通常为残差检验。

- 残差检验：指按所建模型计算出累加数列后，再按累减生成还原，还原后将其与原始数列 $x^{(0)}$ 相比较，求出两序列的差值即为残差，通过计算相对精度以确定模型精度的一种方法。
- 如果相对精度均满足要求精度，则模型通过检验；如果不满足要求精度，可通过上述残差数列建立残差GM(1, 1)模型对原模型进行修正。残差大小检验，即对模型值和实际值的残差进行逐点检验。

4 灰色预测技术

【例1】设有原始数据序列

$$X^{(0)} = \{2.874, 3.278, 3.337, 3.390, 3.679\}$$

试用GM(1,1)模型对 $X^{(0)}$ 进行模拟。

解：第一步，对 $X^{(0)}$ 作1-AGO,得：

$$\begin{aligned} X^{(1)} &= \{x^{(1)}(1), x^{(1)}(2), x^{(1)}(3), x^{(1)}(4), x^{(1)}(5)\} \\ &= \{2.874, 6.152, 9.489, 12.897, 16.558\} \end{aligned}$$

第二步，对 $X^{(1)}$ 进行紧邻均值生成。

$$\text{令： } z^{(1)}(k) = \frac{1}{2}x^{(1)}(k) + \frac{1}{2}x^{(1)}(k-1)$$

$$\begin{aligned} \text{得 } Z^{(1)} &= \{z^{(1)}(1), z^{(1)}(2), z^{(1)}(3), z^{(1)}(4), z^{(1)}(5)\} \\ &= \{4.513, 7.820, 11.184, 14.718\} \end{aligned}$$

4 灰色预测技术

于是：

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ -z^{(1)}(4) & 1 \\ -z^{(1)}(5) & 1 \end{bmatrix} = \begin{bmatrix} -4.5113 & 1 \\ -7.820 & 1 \\ -11.184 & 1 \\ -14.718 & 1 \end{bmatrix}, Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \\ x^{(0)}(5) \end{bmatrix} = \begin{bmatrix} 3.278 \\ 3.337 \\ 3.390 \\ 3.679 \end{bmatrix}$$

第三步，对参数 $\hat{a} = (a, b)^T$ 进行最小二乘估计。得

$$\hat{a} = (B^T B)^{-1} B^T Y = \begin{bmatrix} -0.037 & 20 \\ 3.065 & 36 \end{bmatrix}$$

第四步，确定模型：

$$\frac{dx^{(1)}}{dt} - 0.0372x^{(1)} = 3.06536$$

及时间响应式：

$$\begin{aligned} \hat{x}^{(1)}(k+1) &= \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-at} + \frac{b}{a} \\ &= 85.276151e^{0.0372k} - 82.402151 \end{aligned}$$

4 灰色预测技术

第五步，求 $X^{(1)}$ 的模拟值：

$$\begin{aligned}\hat{X}^{(1)} &= \{\hat{x}^{(1)}(1), \hat{x}^{(1)}(2), \hat{x}^{(1)}(3), \hat{x}^{(1)}(4), \hat{x}^{(1)}(5)\} \\ &= \{2.874\ 0, 6.106\ 0, 9.465\ 0, 12.942\ 2, 16.555\ 8\}\end{aligned}$$

第六步，还原求出 $X^{(0)}$ 的模拟值。由：

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

$$\begin{aligned}\text{得到 } \hat{X}^{(0)} &= \{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \hat{x}^{(0)}(3), \hat{x}^{(0)}(4), \hat{x}^{(0)}(5)\} \\ &= \{2.8740, 3.2320, 3.3545, 3.4817, 3.6136\}\end{aligned}$$

第七步，检验误差。由下表可算出残差平方和：

$$s = \varepsilon \varepsilon^T = [\varepsilon(2), \varepsilon(3), \varepsilon(4), \varepsilon(5)] \begin{bmatrix} \varepsilon(2) \\ \varepsilon(3) \\ \varepsilon(4) \\ \varepsilon(5) \end{bmatrix} = 0.015 \ 11$$

$$\text{平均相对误差: } \Delta = \frac{1}{4} \sum_{k=2}^5 \Delta_k = 1.602 \ 5\%$$

误差检验表

序号	实际数据 $x^{(0)}(k)$	测量数据 $\hat{x}^{(0)}(k)$	残差 $\varepsilon(k)$ $= x^{(0)}(k) - \hat{x}^{(0)}(k)$	相对误差 $\left(\Delta_k = \frac{ \varepsilon(k) }{x^{(0)}(k)} \right) / \%$
2	3.278	3.232 0	0.046 0	1.40
3	3.337	3.354 5	-0.017 5	0.52
4	3.390	3.481 7	-0.091 7	2.71
5	3.679	3.613 6	0.065 4	1.78

4 灰色预测技术

【例2】某大型企业1999年至2004年的产品销售额如下表，试建立GM(1,1)预测模型，并预测2005年的产品销售额。

年份	1999	2000	2001	2002	2003	2004
销售额 (亿元)	2.67	3.13	3.25	3.36	3.56	3.72

解：设 $X^{(0)} = \{2.67, 3.13, 3.25, 3.36, 3.56, 3.72\}$

第1步 构造累加生成序列

$$X^{(1)} = \{2.67, 5.80, 9.05, 12.41, 15.97, 19.69\}$$

第2步构造数据矩阵 B 和数据向量 Y_n

$$B = \begin{bmatrix} -Z^{(1)}(2) \\ -Z^{(1)}(3) \\ -Z^{(1)}(4) \\ -Z^{(1)}(5) \\ -Z^{(1)}(6) \end{bmatrix} = \begin{bmatrix} 4.235 \\ 7.425 \\ 10.73 \\ 14.19 \\ 17.83 \end{bmatrix}$$

$$Y_n = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \\ x^{(0)}(5) \\ x^{(0)}(6) \end{bmatrix} = \begin{bmatrix} 3.13 \\ 3.25 \\ 3.36 \\ 3.56 \\ 3.72 \end{bmatrix}$$

4 灰色预测技术

第3步 计算 $\hat{\alpha} = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y_n$

$$B^T B = \begin{bmatrix} 707.46375 & -54.41 \\ -54.41 & 5 \end{bmatrix}$$

$$(B^T B)^{-1} = \begin{bmatrix} 0.008667 & 0.094319 \\ 0.094319 & 1.226382 \end{bmatrix}$$

$$\hat{\alpha} = \begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y_n = \begin{bmatrix} -0.043879 \\ 2.925663 \end{bmatrix}$$

第4步 得出预测模型

$$\frac{dx^{(1)}}{dt} - 0.043879x^{(1)} = 2.925663$$

$$\hat{x}^{(1)}(k+1) = 69.3457e^{0.043879k} - 66.6757$$

$$x^{(0)} = 2.67 \quad \frac{b}{a} = -66.6757$$

第5步 残差检验

(1) 根据预测公式, 计算 $\hat{x}^{(1)}(k)$, 得

$$\hat{x}^{(1)}(k) = \{2.67, 5.78, 9.03, 12.43, 15.97, 19.68, 19.69\} \\ (k=0, 1, \dots, 6)$$

(2) 累减生成 $\hat{X}^{(0)}(k)$ 序列, $k=1, 2, \dots, 6$

$$\hat{X}^{(0)}(k) = \{2.67, 3.11, 3.25, 3.40, 3.54, 3.71\}$$

原始序列:

$$X^{(0)}(k) = \{2.67, 3.13, 3.25, 3.36, 3.56, 3.72\}$$

4 灰色预测技术

(3) 计算绝对残差和相对残差序列

绝对残差序列 $\Delta^{(0)} = \{0, 0.02, 0, 0.04, 0.02, 0.01\}$

相对残差序列 $\Phi = \{0, 0.64\%, 0, 1.19\%, 0.56\%, 0.01\}$

相对残差不超过1.19%，模型精度高。

第8步 预测：

$$k=7, \quad x^{(0)}(8) = x^{(1)}(8) - x^{(1)}(7) = 4.23$$

即2005年的产品销售额预测值为4.23亿元。

4 灰色预测技术

【例3】选取我国2003~2008年的煤矿百万吨死亡率采用灰色GM(1,1)模型预测。

2003~2008年的煤矿百万吨死亡率和预测结果

序号	年份	实际值 $x^{(0)}(k)$	模拟值 $\hat{x}^{(0)}(k)$	残差值 $\varepsilon^{(0)}(k)$	相对误差
1	2003	4.170	4.170	0.000 0	0.000 0
2	2004	3.100	3.248	-0.048	-0.048
3	2005	2.836	2.556	0.280	0.099
4	2006	2.041	2.011	0.030	0.015
5	2007	1.485	1.583	-0.098	0.066
6	2008	1.182	1.246	-0.064	-0.054

4 灰色预测技术

(1) 根据灰色预测理论,建立煤矿百万吨死亡率的预测模型如下:

$$\hat{x}^{(1)}(t+1) = \left(4.170 - \frac{4.6514}{0.2396}\right) e^{-0.2396t} + 19.4151$$

$$\hat{X}^{(1)} = \{4.170, 7.270, 10.106, 12.147, 13.632, 14.814\}$$

$$\hat{X}^{(0)} = \{4.170, 3.248, 2.556, 2.011, 1.583, 1.246\}$$

(2) 检验精度

原始序列平均值: $\bar{x} = \frac{1}{n} \sum_{k=1}^n x^{(0)}(k) = 2.469$

残差的平均值: $\bar{\varepsilon} = \frac{1}{n} \sum_{k=1}^n \varepsilon^{(0)}(k) = 0$

原始序列的方差: $s_1^2 = \frac{1}{n-1} \sum_{k=1}^n |x^{(0)}(k) - \bar{x}|^2 = 1.2468$

残差的方差: $s_2^2 = \frac{1}{n-1} \sum_{k=1}^n |\varepsilon^{(0)}(k) - \bar{\varepsilon}|^2 = 0.023$

4 灰色预测技术

求得:

$$c = \frac{s_2}{s_1} = 0.1359 < 0.35$$

$$p = p\{\varepsilon_1^{(0)} | (i) - \bar{\varepsilon}_1^{(0)}(i) < 0.6745s_1\} = 1 > 0.95$$

综上所述,对照表可知,计算精度为一级,证明了该方法的合理性和准确性。

(3)预测

模型的精确度达到了要求,可以对下个时间序号进行预测,在煤矿百万吨死亡率GM(1,1)的时间响应函数下,令 $t=7,8$,可以预测未来2年我国煤矿的百万吨死亡率为0.980, 0.772。

(4) 预测结果分析

由预测结果可知,全国煤矿的百万吨死亡率呈递减趋势,煤矿的安全形势不断提高。这种预测结果是在模型精度可靠的前提下进行的,同时灰色预测具有完整严密的理论系统和科学的分析计算方法,预测的相对误差 $<10\%$,因此预测结果科学可靠,对煤矿的安全管理有重大的指导意义。

4.5 灰色关联分析方法

(1) 关联分析概述

例如在社会系统中，人口是一种重要的子系统。影响人口发展变化的有社会因素，如计划生育、社会治安、社会道德风尚、社会的生活方式等。影响人口发展变化的因素还有经济的，如社会福利、社会保险；还有医疗的，如医疗条件、医疗水平等。总之，人口是多种因素互相关联、互相制约的子系统。这些因素的分析对于控制人口、发展生产是必要的。

(2) 关联系数与关联度

数据列的表示方式

做关联分析先要指定参考数据列。参考数据列常记为 x_0 ，记第1个时刻的值为 $x_0(1)$ ，第2个时刻的值为 $x_0(2)$ ，第 k 个时刻的值为 $x_0(k)$ 。因此，参考序列 x_0 可表示为 $x_0 = (x_0(1), x_0(2), \dots, x_0(n))$

关联分析中被比较数列常记为 $x_1, x_2, \dots, x_k$ ，类似参考序列 x_0 的表示方法，有 $x_1 = (x_1(1), x_1(2), \dots, x_1(n))$ ， $\dots$ ， $x_k = (x_k(1), x_k(2), \dots, x_k(n))$

关联系数计算公式

对于一个参考数据列 x_0 ，有几个比较数列 $x_1, x_2, \dots, x_n$ 的情况。可以用下述关系表示各比较曲线与参考曲线在各点（时刻）的差。

$$\xi_i(k) = \frac{\min_i(\Delta_i(\min)) + 0.5 \max_i(\Delta_i(\max))}{|x_0(k) - x_i(k)| + 0.5 \max_i(\Delta_i(\max))}$$

式中， $\xi_i(k)$ 是第 k 个时刻比较曲线 x_i 与参考曲线 x_0 的相对差值，它称为 x_i 对 x_0 在 k 时刻的关联系数。其中，0.5 是分辨系数，记为 ζ 一般在 0 与 1 之间选取；

$$\min_i(\Delta_i(\min)) = \min_i(\min_k |x_0(k) - x_i(k)|)$$

$$\max_i(\Delta_i(\max)) = \max_i(\max_k |x_0(k) - x_i(k)|)$$

关联系数计算

由于作关联度计算的数列的量纲最好是相同的，当量纲不同时要化为无量纲。为了解决这两个问题，计算关联系数之前，先将数列作初值化处理。

无量纲化

无量纲化的方法常用的有**初值化**与**均值化**。

初值化是指所有数据均用第1个数据除，然后得到一个新的数列，这个新的数列即是各个不同时刻的值相对于第一个时刻的值的百分比。经济序列中常用此法处理。

均值化处理则是用平均值去除所有数据，以得到一个占平均值百分比的数列。

[例] 关联系数的计算

给出已出初值化的序列如下：

$$x_0 = (1, 1.1, 2, 2.25, 3, 4) \quad x_1 = (1, 1.166, 1.834, 2, 2.314, 3)$$

$$x_2 = (1, 1.125, 1.075, 1.375, 1.625, 1.75) \quad x_3 = (1, 1, 0.7, 0.8, 0.9, 1.2)$$

下面分三步计算关联系数：

第一步 求差序列

各个时刻 x_i 与 x_0 的绝对差如下

4 灰色预测技术

序 号	1	2	3	4	5	6
$\Delta_1 = x_0(k) - x_1(k) $	0	0.066	0.166	0.25	0.686	1
$\Delta_2 = x_0(k) - x_2(k) $	0	0.025	0.925	0.875	1.375	2.25
$\Delta_3 = x_0(k) - x_3(k) $	0	0.1	1.3	1.45	2.1	2.8

第二步 求两级最小差与最大差

容易求出 $\min(\min_k |x_0(k) - x_i(k)|) = 0$ $\max(\max_i |x_0(k) - x_i(k)|) = 2.8$

第三步 计算关联系数

将数据代入关联系数计算公式，得

$$\xi_i(k) = \frac{0 + 0.5 \times 2.8}{|x_0(k) - x_i(k)| + 0.5 \times 2.8} = \frac{1.4}{\Delta_i(k) + 1.4}$$

4 灰色预测技术

令 $i=1$, 我们有

序 号	1	2	3	4	5	6
$\Delta_i(k)$	0	0.066	0.166	0.25	0.686	1
	$\Delta_1(1)$	$\Delta_1(2)$	$\Delta_1(3)$	$\Delta_1(4)$	$\Delta_1(5)$	$\Delta_1(6)$

因此, 我们有

$$\xi_1(1) = \frac{1.4}{\Delta_1(1) + 1.4} = \frac{1.4}{0 + 1.4} = 1$$

$$\xi_1(2) = \frac{1.4}{\Delta_1(2) + 1.4} = \frac{1.4}{0.066 + 1.4} = 0.955$$

$$\xi_1(3) = \frac{1.4}{\Delta_1(3) + 1.4} = \frac{1.4}{0.166 + 1.4} = 0.894$$

$$\xi_1(4) = \frac{1.4}{\Delta_1(4) + 1.4} = \frac{1.4}{0.25 + 1.4} = 0.848$$

$$\xi_1(5) = \frac{1.4}{\Delta_1(5) + 1.4} = \frac{1.4}{0.686 + 1.4} = 0.679$$

$$\xi_1(6) = \frac{1.4}{\Delta_1(6) + 1.4} = \frac{1.4}{1 + 1.4} = 0.583$$

作关联系数 $\xi_1(k)$ 在各个时刻的值的集合，得关联系数序 ξ_1



灰色预测技术

$$\xi_1 = (\xi_1(1), \xi_1(2), \xi_1(3), \xi_1(4), \xi_1(5), \xi_1(6)) = (1, 0.955, 0.894, 0.848, 0.679, 0.583)$$

同理有

$$\xi_2 = (\xi_2(1), \xi_2(2), \xi_2(3), \xi_2(4), \xi_2(5), \xi_2(6)) = (1, 0.982, 0.602, 0.615, 0.797, 0.383)$$

$$\xi_3 = (\xi_3(1), \xi_3(2), \xi_3(3), \xi_3(4), \xi_3(5), \xi_3(6)) = (1, 0.933, 0.52, 0.49, 0.4, 0.34)$$

关联度

关联系数的数很多，信息过于分散，不便于比较，为此有必要将各个时刻关联系数集中为一个值，求平均值便是做这种信息处理集中处理的一种方法。

关联度的一般表达式为：

$$r_i = \frac{1}{N} \sum_{k=1}^N \xi_i(k)$$

4 灰色预测技术

【例4】山西省汾河上游的**输沙量与降雨径流**的灰色关联分析 汾河是山西省的主要河流，在汾河下游距太原市100多公里的西山修建了汾河水库。该水库不但对农业灌溉、防洪蓄水、鱼类养殖等起着很大作用，并且还为太原市的用水提供了保证。建库以来，人们经常在考虑如何防止库容被泥沙淤塞，使水库能长期有效为工农业生产与人民生活服务。

影响泥沙输入水库的因素较多，比如**降雨量、径流量、植被覆盖率**等。在这些因素中哪些是主要的，哪些是次要的有待研究和量化分析。

4 灰色预测技术

以输沙量为参考数列 x_0 ，以年径流量为 x_1 ，平均年降雨量为 x_2 平均汛期降雨量为 x_3 则相应的关联系数序列如下：

$$\xi_1(k) = (1, 0.4, 0.4, 0.32, 0.86, 0.23, 0.29, 0.2, 0.53, 0.45, 0.17, 0.29, 0.73, 0.36, 0.27, 0.31, 0.35, 0.44, 0.42, 0.34, 0.61, 0.51, 0.36, 0.4, 0.23, 0.34, 0.33)$$

$$\xi_2(k) = (1, 0.24, 0.17, 0.17, 0.29, 0.22, 0.15, 0.1, 0.24, 0.14, 0.1, 0.2, 0.22, 0.14, 0.2, 0.18, 0.21, 0.1, 0.22, 0.16, 0.21, 0.13, 0.13, 0.23, 0.17, 0.19, 0.14)$$

$$\xi_3(k) = (1, 0.26, 0.17, 0.2, 0.32, 0.19, 0.16, 0.11, 0.23, 0.18, 0.1, 0.27, 0.25, 0.16, 0.24, 0.22, 0.23, 0.1, 0.25, 0.16, 0.23, 0.21, 0.13, 0.24, 0.17, 0.26, 0.19)$$

4 灰色预测技术

根据关联系数求关联度得

$r_1 = 0.41$ （年径流量与输沙量的关联程度）

$r_2 = 0.21$ （年平均降雨量与输沙量的关联程度）

$r_3 = 0.23$ （平均汛期降雨量与输沙量的关联程度）

相应的关联序为 $r_1 > r_3 > r_2$

上述关联序表明对输沙量影响最大的是年径流量，其次是汛期降雨量，再其次是平均年降雨量。

实际上，强度大的暴雨冲刷力大，难以被土壤吸收，从而在地表形成径流，造成水土流失，引起河道泥沙流量的形成



THANK YOU

谢谢大家!

