**灰色-马尔可夫链算法**

**1.灰色系统**

人类社会进入一个新的历史时代，有人称之为原子时代，有人称之为信息时代等，我们认为这个新时代称之为系统时代更为恰当。系统是无处不在、无所不包的。人体是一个系统，它由循环系统、呼吸系统、消化系统、骨豁系统和肌肉系统等组成；天上的飞禽地上的走兽，百花齐放，万物争，，自然界为我们安排了奇妙的生物系统；庞大的地球只不过是太阳的一颗行星，太阳又置身于浩瀚的银河系之中，而银河系也只是茫茫字宙的一片星云，这就是天体系统；铁路、公路、水路和航空构成交通运输系统；发电厂、输电线、变电站和用电设备构成电力系统；还有城市系统、商业系统、水利系统、电讯系统等等。因而，可以断言，我们无时无刻不与一定的系统相处。

系统的三要素就是物质、能量和信息，这三者缺一不可。物质由联系而沟成系统，而联系必须依靠能量和信息的作用一切物质都在运动，一切物质都具有能量。信息就是信息，不是物质也不是能量，但发送信息的信源，接受信息的信宿，都是物质。信息必须通过物质才能储存和传播。

对于系统，许多学者都留下过定义：贝塔朗菲认为：系统是“相互作用诸要素的综合体”；韦氏大辞典把系统定义为：“有组织或被组合化的整体，结合着的整体所形成的概念和原理的综合；由有规则的相互作用、相互依存的形式组成的要素集合等等”；日本工业标准，定义系统为：“许多组成要素保持有机的秩序，向同一目标行动的事物”；凯斯特等人认为：系统是一项有组织曲整体，由两个或两个以上的相关联的“个体”或“构成体”，或“次系体”所构成，存在于其外在的高级系统之内，具有明确的边界者。”

我们认为：系统是一个有机的整体，并且不同的系统都有其不同的特定功能，系统内部要素之间总是有相互联系、相互影响和相互作用的关系。概念是普遍适用的，即一切物质都构成系统，宇宙间的一切都是系统。

人们常用颜色来简要地描述信息的完备程度。1953年艾什比首先使用黑盒（Black Box）—词，这个世界中总有一些事物，我们对其不了解，所谓不了解就是对事物内部结构不确定，以及事物的组成的不确定，要想认识该事物，只能从事物的外部关系来认识，这种类型的事物我们通常叫做黑盒或者黑箱。

黑箱原理、黑箱问题至今在系统辩识、建摸、仿真，以及大系统研究中仍占有一席之地。此后，又有人提出“灰箱”（Grey Box）—词，既指客观事物中以掌握其部分信息一类问题。由此可以推论，若信息完备的事可称之谓白箱（White Box）。

我们认为用“箱”一词不如用“系统”一词来得确切、全面、恰到好处。因为箱必有箱壁，有边界使得人们不得不从箱外去研究。而系统则可以从内部去研究，从其结构功能及相应的参数去研究。由于每个系统的信息完整度不同，即按照信息完整度对系统进行划分，分别为：灰色系统、白色系统、黑色系统。

白色系统是指信息完整的系统如一台电脑，它的、显卡、硬盘、装载的操作系统、键盘、鼠标的型号和出厂期都是可以知道的；一架彩色电视机包括什么零部件、电子元件、显像管都是清楚的；一个家庭有多少人口、各成员的性别、年龄。及其相互关系部是清楚的，家庭收入、日常消费支出都可以计算。

黑色系统是指那些毫无信息的系统，例如浩激的宇宙，现有的科学技术只知道地球属银河系，但是银河系之外的所存在的星球，对其一无所知，所以这些星球属于黑色系统；我国神农架的野人，有什么群体关系？有何种生活习惯？使用何种语言？今后发展趋势如何？可以说是一无所知，上述系统均为黑色系统。

但是白色、黑色系统都与人的认识水平有关，随着人类的进步，科技的发展，过去认为不可知的事物，现在已变为可知的了。灰色系统是指一些系统的一部分信息可知，而另一部分的信息未知且该类系统内可知的信息为白色参数，而不可知道的信息称之为黑色参数。例如我们人体这个灰色系统，可知的信息有：升高、年龄、性别、体温、脉搏这些都称为白色参数。但尚有许多参数至今尚不清楚，如人体究竟有多少穴位，穴位究竞有哪些物理的、化学的、生理的机能，穴位与人体的哪些部位有关，能起什么作用等等尚说不清楚。至于人体的特异功能，更是奥妙无穷。地理系统、生态系统、社会系统等一些大型并且复杂的系统也都是灰色系统。

**2.灰色系统的概念**

“凡事预则立，不预则废”，这是我国古代学者对预测的认识和评价，这句话精辟地概括了预测的深刻意义。可见，预测之事，古己有之，诸葛亮草船借箭，借东风，火烧赤壁的故事传为佳话、春秋末期，越国范龜弃政从商，曾制定“水则资车，旱则资舟”之策，显见其颇有远见卓识。古人虽已需要预测，也有成功之处，但毕竟限于历史条件，未能形成预测科学。在当时，人们需要预测，但又在缺乏科学方法的情况下，不得不求助于星相占卜之举，例如，吉普赛人用水晶球；印度人用茶叶；国人用拆字看相等等方法预知未来凶吉。但这种预测常带有迷信色彩。

决策的失误必将导致严重后果，当今时代处于高速发展节奏，对决策者的要求即越来越高，如何做出正确的决策是每一位决策者必然思考的问题，只有运用科学预测方法作为决策的主要参考依据，才会做出正确的决策，如果不重视预测而盲目决策，必然导致决策失误。例如：在旧的经济体制下，产销脱节，生产企业不关心市场需求情息，一味盲目生产，工业报喜，商业报忧，供过于求，造成产品滞销，资金积压，企业亏损，甚至瀕临倒闭者有之；自由建设，投资大，效益低，投资回收遥遥无期者有之，相反决策者若能对企业的未来作充分的预测，能掌控今后若干年内本企业产品的市场需求情况，必将带领一个企业走向辉煌。在企业中预测是为决策服务的，同样的在工程实践中正确科学的预测也同样会给施工带来较高的效率，以及安全。科学准确的掌握边坡位移，对矿山生产和工作人员的安全，有着重要的意义。只有科学准确的预测边坡位移，才能为矿山企业做出正确的决策提供保障。

所谓预测就是去发现和控制与人类有联系事物的发展趋势，为了达到目的现代预测通常将科学的理论方法与经验相结合，对历史的数据进行分析和预估，对事物的未来做出预测，以帮助决策者指定正确的方向，降低事物未来的不确定性。简而言之，预测的本质就是将过去和现在作为基础，在此基础上对未来进行估计，减少未来的不确定性。预测仅仅是手段，并不是目的。

科学技术水平的不断发展，越来越多的领域需要预测，致使预测的方法层出不穷，近年来有关预测著作近多种，介绍的预测方法不下种，从不同的角度，不同的方法提供多种预测方法。本文所介绍无偏灰色模糊马尔可夫预测方法属于应用数学工具预测方法。按照应用数学工具的及其性质可分为：

1. 直观法，是定性预测的一种方法，主要凭个人的经验和分析综合能力。属于这一类的有个人判断法，专家会议法、智力激励法、情景描述法、台尔菲法扣相互影响矩阵法等。

2. 外推法，是定量预测的一种方法，它主要是仿效物理学惯性原理而产生的概念。在一个力学系统中，当在结构不变、外力不变的情况下，运动物体将按惯性原理维持其原来的运动方式继续运动。对于一个经济系统或是一个社会系统亦可用惯性原理来进行研究，这就是说经济系统的未来必然是过去和现在的延续，过去的发展趋势可以作为未来的预测。系统规模愈大，惯性愈大，系统形成的时间愈早、愈长；系统结构愈复杂，惯性愈大。惯性原理是外控法的理论依据。

属于外推法的预测方法有：时间序列法、移动平均法、指数平滑法、季节系数法、生长曲线法、替代曲线法、学习曲线法等。在灰色预测模型中，*GM(1，1)*和*GM(2，1)* 等方法亦外推法的一种新的开拓。

外推法适用于系统结构复杂，作用因素众多，预测期内结构上无重大变化，无突发事件的情况，一般预测期不宜过长，否则，惯性逐渐衰减，导致如预测精度明显下降。

相关法，是定量预测的重要方法，它主要是寻找预测变量与其他影响因素之间的函数关系。辩证法的重要原则之一就是认为任何事物都不是孤立的，而是联系，相互制约的，事物之间存在着明显的因果关系，追根觅源是一种预测的好方法。

属于这一类的有回归分析法，弹性系数法等。灰色预测模型，是一种新的幵拓。因果法或相关法适用于因果关系明显、统计数据齐全的场合，对于系统结构变化有一定的适应性，预测精度较高，但原始资料的收集有一定的难度，计算方法比较复杂。

4. 系统模型法，它是定量预测的高级形式。从系统观点来看，一个系统是有层次的，每个层次、每个子系统都有自己的目标体系，因而，对多层次、多目标的系统显然不能用单方程的模型进行预测，必须建立多方程的模型体系，用以从数学上描述系统行为，或对系统进行仿真，达到预测的目的，为系统的控制、管理、规划提供依据。

属于这类的有：经济计量模型，经济控制论模型，投入产出模型，系统动力学模型等。灰色系统模型的建立是用一些具体的参数描述经济、社会等系统用以协调、优化系统结构，具有良好的效果。

传统的预测模型是以概率统计为基础，并结合回归分析得出的，且该模型本身就是离散的，因而产生了一系列的问题：（1）由该模型是以概率统计为基础，即对数据的需求量相当大，只有得到大量的数据才可以统计得到系统的预测规律；（2）对变量严格要求，其必须正态分布并且具有平稳的过程；（3）传统的预测结果是由概率构成，从结果上很难获取较多的信息量；（4）预测结果精度较低，在一些情况下系统出现不稳定；（预测结果过于宏观，不足以分析变化过程；（5）大量的数据需要计算导致工作量繁重。

然而灰色系统的出现运用一种全新的角度和方法，解决了概率统计自身所存在的问题，通过对毫无规律的、离散的数据中寻找出其变化规律，并且建立灰色模型，通过灰色模型来重新认识系统。

灰色系统理论具有很强的包容性，并且具有立体的思维模式。且灰色系统理论的决策是可以多目标的，预测方法也可以是多样化的。

随机变量的概念只限于概率统计中，在灰色系统内随机变量等同变化的灰色量，与概率统计不同的是灰色理论更注重对灰色量经行有效分析处理，而不是一味的去寻找灰色量的统计规律。

灰色系统理论的特点如下所述：

1.在对杂乱无章的原始数据建模前，灰色系统会对数据进行处理，经处理后的数据具有一定的规律性，这样的方法定义为生成，生成后旳结果也就是新产生的数列称之为“模块”。生成的作用就是消除原始数据随机性的特点。

2. 生成的基本思想是：客观存在的系统内，各因素看似毫无关系，但实际必然存在某种联系，以帮助系统完成某项功能，因此系统内必将存在某种规律，生成模块的目的就是寻找这种客观存在的规律的方法。生成方法有许多种，但最重要也是最常用的是累加生成，通过累加生成可以将毫无规律波动的时间序列，转变为单调增长的指数规律序列。

3. 灰色系统的关联分析法，主要对系统的变化机理和发展做特征分析，该类型分析属于拓扑分析，根据不同的发展趋势建立关联度。通过运用此种分析方法可以对灰色理论所构成的函数检验。

**3. 灰色预测模型*GM(1，1)***

灰色预测模型*GM(1，1)*由于所需数据量少，不需要考虑数据分布规律，适合数据波动不大的模型。预测数据的随机性很大程度上影响着预测效果，所以需要对原始数据进行处理以弱化随机性，常用的数据处理方法有累加、累减、累积等方法[3]。通过对比，发现累加生成的数据序列具有单调凹升与近似指数变化规律，适合用*GM(1，1)*模型进行预测。所以，*GM(1，1)*采用累加生成的方法对原始数据序列进行处理，再对累加序列建立灰色微分方程，通过求解该微分方程对销售量进行预测。

给定某一商品历史销售量数据序列：

** (1)

对销售量序列进行累加生成得到累加序列

 (2)

式中：。

考虑到偶然性，收集到的数据可能存在空缺或者异常值，为了减小这些数据对预测结果的影响，用式(2)得到的序列进行均值处理[4]，得到紧邻均值序列：

 (3)

式中:。

然后对建立灰色模型的微分方程：

 (4)

式中： 为待定系数。用最小二乘法求解，有

 (5)

式中，

求得方程的解为：

 (6)

初始状态，从而得到数据还原模型：

 (7)

本文在预测过程中将实时对预测模型进行修正，每次预测结束，将当前预测点的真实值加入历史数据序列，同时舍弃最老的数据，将获得的最新的序列对下一时间点进行预测。

通过式（1）对**进行预测，预测结束新的历史数据序列为

|  |  |
| --- | --- |
|  | (8) |

通过实时更新历史数据序列，能够修正灰色预测的偏差，以获得最佳的预测效果。

**4. 马尔可夫链模型**

马尔可夫链指对于随机过程*X(t)*，当*t0*时刻所处的状态已知时，此后的状态只与时刻*t0*的状态有关，而与*t0*之前的状态无关。

设马尔可夫过程的参数集T是离散时间序列，即，则*xn*的取值空间是离散的状态集。

定义随机过程对于任意的及任意的，有



 (9)

称为马尔可夫链，并记

 (10)

式中，*pab(k)*表示当*m*时刻系统状态为*b*，时刻*m+k*系统状态处于*a*的概率。

由系统中各种可能状态的转移概率*pab(k)*构成矩阵：

 (11)

式中：*k*为转移步数。

*P*即为马尔可夫链状态转移矩阵，且具有以下性质：。