简单的灰色预测----GM(1,1)预测

1982年，华中理工大学邓聚龙教授首先提出了灰色系统的概念，并建立了灰色系统理论，引起了国内外很多学者、科技人员的重视。之后，灰色系统理论得到了较深入的研究，并在众多方面获得了成功的应用。

【灰色系统】：既含有已知信息又含有未知的非确知的信息的系统。例如：人口问题、历史系统、中医系统等。

【灰色系统的描述】：灰色系统用灰色参数（灰元、灰数）、灰色方程、灰色矩阵、灰色度等综合描述，其中灰数是灰数系统的基本“单元”或“细胞”。

- 灰色系统是通过对原始数据的整理来寻找其变化规律的，这是一种**就数据寻找数据的现实规律**的途径，称为灰色序列生成。
- （灰色系统理论认为，尽管客观表象复杂，数理离乱，但总是有整体功能的，因此必然蕴含某种内在规律。关键在于如何选择适当的方式去挖掘和利用它。一切灰色序列都能***通过某种生成***弱化其随机性，显现其规律性。）
- 灰色理论中常用的生成方法有：

  累加生成（AGO），即累加生成算子；

  累减生成(IAGO)或逆累加生成以及均值生成Z。

- 生成法如下：
- 设原始数据列为：$x^{(0)}=\{x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n)\}$， 则
- 1次累加（1-AGO）：$x^{(1)}=\{x^{(1)}(1),x^{(1)}(2),\cdots,x^{(1)}(n)\}$，
  其中

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

……………………

- R次累加（r-AGO）：$x^{(r)}=\{x^{(r)}(1),x^{(r)}(2),\cdots,x^{(r)}(n)\}$，
  其中

$$x^{(r)}(k)=\sum_{m=1}^{k}x^{(r-1)}(m)=x^{(r)}(k-1)+x^{(r-1)}(k)$$ ；

- 均值生成Z：$z^{(1)}(k)=\frac{1}{2}(x^{(1)}(k)+x^{(1)}(k-1))$ ；
- 累减生成IAGO：$\alpha^{(1)}(x^{(1)}(k))=x^{(1)}(k)-x^{(1)}(k-1)=x^{(0)}(k)$。

定义1 称 $x^{(0)}(k)+az^{(1)}(k)=b$ 为**GM(1,1)模型**.

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

符号GM(1,1)的含义如下:

G	M	(1,	1)
↑	↑	↑	↑
Grey	Model	1阶方程	1个变量

定理1 设$X^{(0)}$为非负序列: $X^{(0)}=(x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n))$

其中$x^{(0)}(k)>=0,k=1,2,\ldots,n$; $X^{(1)}$为$X^{(0)}$的1-AGO序列:

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

其中$x^{(1)}(k)=\sum_{i=1}^{k}x^{(0)}(i),k=1,2,\cdots,n$; $Z^{(1)}$为$X^{(1)}$的紧邻均值生成序列:

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

其中$z^{(1)}(k)=0.5x^{(1)}(k)+0.5x^{(1)}(k-1)$;$k=2,3,\ldots,n$

若 $\hat{a}=(a,b)^T$ 为参数列,且

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

则灰色微分方程$x^{(0)}(k)+az^{(1)}(k)=b$ 的最小二乘估计参数列满足 $\hat{a}=(B^TB)^{-1}B^TY$

定义2 设$X^{(0)}$为非负序列,$X^{(1)}$为$X^{(0)}$的1-AGO序列, $Z^{(1)}$为$X^{(1)}$的紧邻均值生成序列, $\hat{a}=(B^T B)^{-1} B^T Y$, 则称

$$\frac{dx^{(1)}}{dt}+ax^{(1)}=b$$

为灰色微分方程 $x^{(0)}(k)+az^{(1)}(k)=b$

的白化方程,也叫影子方程.

定义3 称GM(1,1)模型中的参数-a为发展系数,b为灰色作用量.

定理2 设B,Y, $\hat{a}$ 如定理1所述,则

1 白化方程 $\frac{dx^{(1)}}{dt}+ax^{(1)}=b$ 的解也称时间响应函数为

$$x^{(1)}(t)=(x^{(1)}(0)-\frac{b}{a})e^{-a(t-1)}+\frac{b}{a}$$

2 GM(1,1)灰色微分方程 $x^{(0)}(k)+az^{(1)}(k)=b$ 的时间响应序列为

$$\hat{x}^{(1)}(k+1)=(x^{(1)}(0)-\frac{b}{a})e^{-ak}+\frac{b}{a};k=1,2,\cdots,n$$

3 取 $x^{(1)}(0)=x^{(0)}(1)$,则

$$\hat{x}^{(1)}(k+1)=(x^{(0)}(1)-\frac{b}{a})e^{-ak}+\frac{b}{a};k=1,2,\cdots,n$$

4 还原值

$$\hat{x}^{(0)}(k+1)=a^{(1)}\hat{x}^{(1)}(k+1)=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k);k=1,2,\cdots,n$$

定理3 GM(1,1)模型

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

可以转化为

$$x^{(0)}(k)=\beta-\alpha x^{(1)}(k-1)$$

其中

$$\beta=\frac{b}{1+0.5a} \qquad \alpha=\frac{a}{1+0.5a}$$

定理4 设 $\beta=\dfrac{b}{1+0.5a}$， $\alpha=\dfrac{a}{1+0.5a}$，且

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

为GM(1,1)模型时间响应序列,其中

$$\hat{x}^{(1)}(k)=(x^{(0)}(1)-\frac{b}{a})e^{-a(k-1)}+\frac{b}{a}$$

则 $$x^{(0)}(k)=(\beta-\alpha x^{(0)}(1))e^{-a(k-2)}$$

定义4 设序列

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

将$x^{(0)}(n)$取为时间轴的原点,则称t<n为过去,t=n为现在,t>n为未来.

定义5 设序列 $X^{(0)}=(x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n))$

$$\hat{x}^{(0)}(k+1)=(1-e^{a})(x^{(0)}(1)-\frac{b}{a})e^{-ak}$$

为其GM(1,1)时间响应式的累减还原值,则

1 当t<=n时,称 $\hat{x}^{(0)}(t)$ 为模型模拟值;

2 当t>n时,称 $\hat{x}^{(0)}(t)$ 为模型预测值.

建模的主要目的是预测,为提高预测精度,首先要保证有充分高的模拟精度,尤其是t=n时的模拟精度.因此建模数据一般应取为包括$x^{(0)}(n)$在内的一个等时距序列.

定义6 设原始数据序列

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

1 用 $X^{(0)}=(x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n))$ 建立的GM(1,1)模型称为全数据GM(1,1)

2 用 $X^{(0)}=(x^{(0)}(k_0),x^{(0)}(k_0+1),\cdots,x^{(0)}(n))$ 建立的GM(1,1)模型称为部分数据GM(1,1)

3 设$x^{(0)}(n+1)$为最新信息,将$x^{(0)}(n+1)$置入$X^{(0)}$,称用

$$X^{(0)}=(x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n),x^{(0)}(n+1))$$

建立的模型为新信息GM(1,1)

4 置入最新信息$x^{(0)}(n+1)$,去掉最老信息$x^{(0)}(1)$,称用

$$X^{(0)}=(x^{(0)}(2),\cdots,x^{(0)}(n),x^{(0)}(n+1))$$

建立的模型为新陈代谢GM(1,1).

GM(1,1)模型的适用范围

当GM(1,1)发展系数|a|>=2时,GM(1,1)模型无意义.
通过分析,可得下述结论:
(1)当-a<=0.3时,GM(1,1)可用于中长期预测
(2)当0.3<-a<=0.5时,GM(1,1)可用于短期预测,中长期预测慎用
(3)当0.5<-a<=0.8时,GM(1,1)作短期预测应十分谨慎
(4)当0.8<-a<=1时,应采用残差修正GM(1,1)
(5)当-a>1时,不宜采用GM(1,1)

后验差检验

- （1）计算原始数列的均值 $\bar{x}^{(0)}=\frac{1}{n}\sum_{i=1}^{n}x^{(0)}(i)$

- （2）计算原始数列的均方差

$$S_0=\sqrt{\frac{S_0^2}{n-1}} \qquad S_0^2=\sum_{i=1}^{n}[x^{(0)}(i)-\bar{x}^{(0)}]^2$$

- （3）计算残差的均值 $\bar{\varepsilon}^{(0)}=\frac{1}{n}\sum_{i=1}^{n}\varepsilon^{(0)}(i)$

- （4）求残差的均方差

$$S_1=\sqrt{\frac{S_1^2}{n-1}} \qquad S_1^2=\sum_{i=1}^{n}[\varepsilon^{(0)}(i)-\bar{\varepsilon}^{(0)}]^2$$

- （5）计算方差比 $c=\dfrac{S_1}{S_0}$

- （6）计算小误差概率 $p=\{|\varepsilon^{(0)}(i)-\overline{\varepsilon}^{(0)}|<0.6745\cdot S_0\}$

- （7）检验

  根据经验，一般精度等级的划分如下表

预测精度等级划分

- 在允许范围之内时，则可用所建模型进行预测，否φ则应进行**残差修正**。

值	值	预测精度等级
>0.95	<0.35	好
>0.80	<0.5	合　格
>0.70	<0.65	勉强合格
≤0.70	≥0.65	不 合 格

生成数列$\hat{x}^{(1)}$与其模拟值之差为$\varepsilon^{(0)}(j)=x^{(1)}(j)-\hat{x}^{(1)}(j)$

对 $\varepsilon^{(1)}$ 建立GM(1,1)模型， 与原始数列$x^{(0)}$ 的GM(1,1)模型

相加，则得 $x^{(0)}$ 的**残差修正GM(1,1)**模型。

第二次大作业

- 对2005年长江水质附件中的数据进行分析
- 要求：

  1、用尽可能多的分析方法和角度，以采用分析角度、方法多少及该方法完成质量评分。
- 2、每个分析要有程序或调用界面及分析数据结果和结果解释。如无程序或调用过程，直接copy已有论文结果，以0分记。
- 3、不用建模解决原题中问题
- 4、不用写建模分析
- 5、程序用附录形式，数据不用附。