



信阳师范学院
数学与统计学院
SCHOOL OF MATHEMATICS AND STATISTICS

第13章 MATLAB预测分析



讲授人：牛言涛



日期：2020年5月9日

目录

CONTENTS



灰色预测

GM(1,1), GM(1,N),
GM(2,1), DGM(2,1),
Verhulst模型



移动平均与指数平滑法



时间序列法



微分方程与差分方程



马尔可夫预测



BP神经网络



- 灰色系统理论是我国学者**邓聚龙**教授于20世纪80年代初创立并发展的理论，它把一般系统论，信息论和控制论的观点和方法延伸到社会，经济，生态等抽象系统，结合运用数学方法发展的一套解决灰色系统的理论和方法，20多年来，灰色系统理论引起了国内外学者的广泛关注。
- 灰色系统理论已成功应用到工业，农业，社会，经济等众多领域，解决了生产，生活和科学研究中的大量实际问题。它可以对系统做出分析，建模，预测，决策，控制等。其应用案例：

例 1 1983年，邓聚龙接受了我国粮食发展决策中的一个课题，即粮食发展预测与粮食发展的长期规划。建立GM(1,1)模型，预测1983-2000年粮食产量，1983年预测值与统计值误差仅为4‰。

例 2 对湖北省2000年宏观经济的发展趋势进行预测，提出自己的结论和建议，受到湖北省科协的奖励。

例 3 河南省人民胜利渠利用灰色系统理论制定最佳灌溉方案，使全流区农田得到大丰收。

- 灰色系统、白色系统和黑色系统
 - 白色系统是指一个系统的内部特征是完全已知的，既系统信息是完全充分的。
 - 黑色系统是一个系统的内部信息对外界来说是一无所知的，只能通过它与外界的联系来加以观测研究。
 - 灰色系统介于白色和黑色之间，灰色系统内的一部分信息是已知的，另一部分信息是未知的，系统内各因素间有不确定的关系。
- 灰色系统的定义：灰色系统指既含有已知信息又含有未知信息的系统，对在一定范围内变化的，与时间有关的灰色过程进行预测。
- 灰色预测模型的定义：对灰色系统进行预测的模型。灰色模型（Grey Model，简称GM）一般表达方式为GM(n, x)模型，其含义是：用n阶微分方程对x个变量建立模型。

- 灰色系统解决了这个一向认为不能解决的连续微分方程的建模问题。为什么灰色系统理论能解决这个问题呢？其重点是灰色系统有一种新观点，即
 1. 任何随机过程都是在一定幅值范围，一定时区内变化的灰色量，称随机过程为灰色过程。
 2. 在处理手法上，灰色过程是通过原始数据的整理来寻找数的规律的，叫数的生成，这就是一种就数找数的现实规律的途径。
- 灰色系统认为：尽管客观系统表象复杂、数据离散，但它们总是有整体功能的，总是有序的。因此，它必然潜藏着某种内在规律。关键在于要用适当方式去挖掘它，然后利用它。
- 由于生成数据列有了较强的规律，有可能对变化过程做较长时间的描述。因此，有可能建立微分方程。建立微分方程模型，还要利用到灰色理论的其他成果，如：关联空间的知识，离散函数的收敛，离散函数的光滑度，灰导数，灰微分方程，平蛇等概念。

- 灰色预测模型的目的：

通过把分散在时间轴上的离散数据看成一组连续变化的序列，采用累加和累减的方式，将灰色系统中的未知因素弱化，强化已知因素的影响程度，最后构建一个以时间为变量的连续微分方程，通过数学方法确定方程中的参数，从而实现预测目的。

- 灰色系统预测模型的特点：无需大量数据样本，短期预测效果好，运算过程简单。
- 灰色系统预测模型的不足：对非线性数据样本预测效果差。
- 常用的灰色系统预测模型主要有GM(1,1)、GM(1,n)、GM(2,1)、DGM(2,1)等
- 灰色生成数列：数据生成的常用方式有累加生成、累减生成和加权累加生成。
- 数据累加生成的特点：一般经济数列都是非负数列，累加生成能使任意非负数列、摆动的与非摆动的，转化为非减的、递增的。

- 数据累加生成的特点：一般经济数列都是非负数列，累加生成能使任意非负数列、摆动的与非摆动的，转化为非减的、递增的。

```
>> X = [174, 179, 183, 189, 207, 234, 220.5, 256, 270, 285 300 320 344 365];
```

```
>> subplot(1,2,1)
```

```
>> plot(X,'b*-', 'LineWidth',1.5)
```

```
>> title('原数据趋势')
```

```
>> xlabel('序号'); ylabel('数据值')
```

```
>> grid on
```

```
>> Ago = cumsum(X); %原始数据一次累加
```

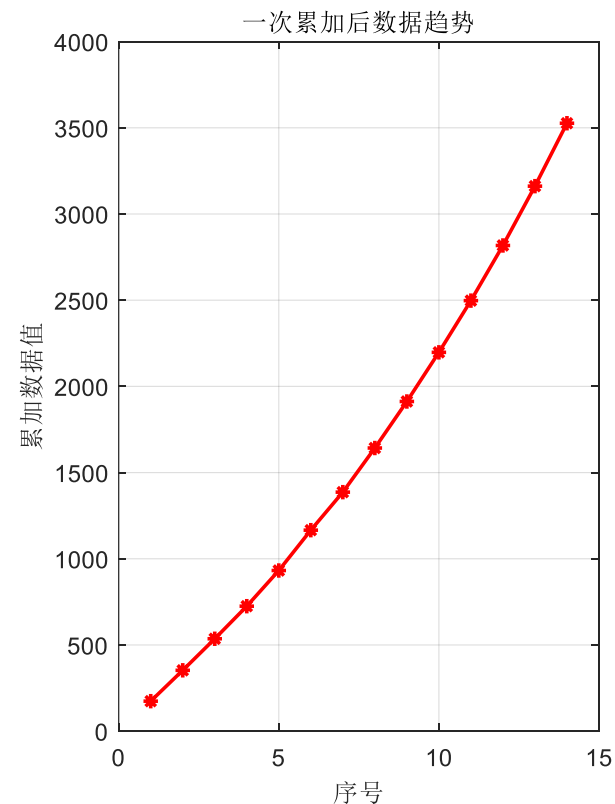
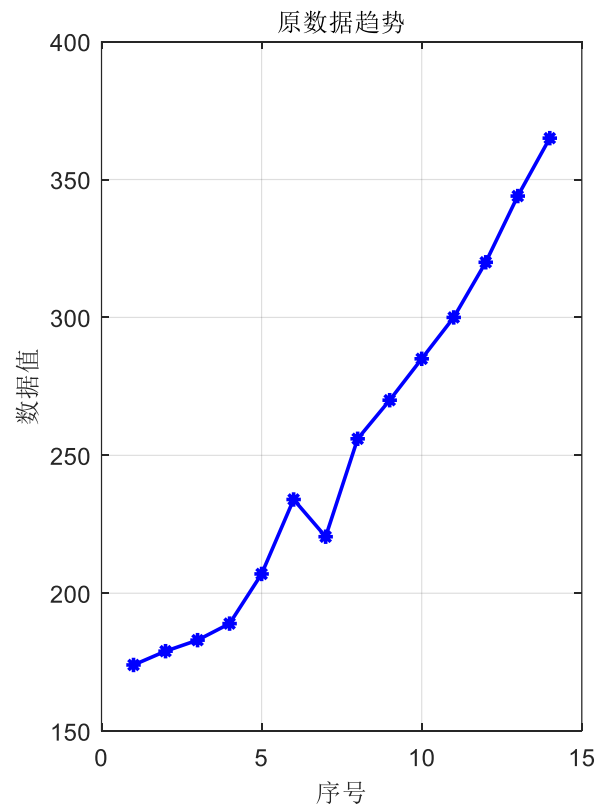
```
>> subplot(1,2,2)
```

```
>> plot(Ago,'r*-', 'LineWidth',1.5)
```

```
>> title('一次累加后数据趋势')
```

```
>> xlabel('序号'); ylabel('累加数据值')
```

```
>> grid on
```



一. 灰色预测模型GM(1,1)原理分析

GM(1,1)模型的预测原理是：对某一数据序列用累加的方式生成一组趋势明显的新数据序列，按照新的数据序列的增长趋势建立模型进行预测，然后再用累减的方法进行逆向计算，恢复原始数据序列，进而得到预测结果。

GM(1,1)建模过程：

1. 设一组原始数据为 $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ ， n 为数据个数。对 $x^{(0)}$ 累加以便弱化随机序列的波动性和随机性，得到新的数列为： $x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n))$ ，其中

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), k = 1, 2, \dots, n$$

2. 生成 $x^{(1)}$ 的邻均值等权数列

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

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

1. 灰色预测模型GM(1,1)原理分析

3. 根据灰色理论对 $x^{(1)}$ 建立关于 t 的白化形式的一阶一元微分方程, $GM(1, 1): \frac{dx^{(1)}}{dt} + ax^{(1)} = u$

其中, a, u 为待解系数, 分别称为发展系数和灰色作用量, a 的有效区间是 $(-2, 2)$, 并记 a, u 构成的矩阵为灰参数 $\hat{a} = (a, u)^T$, 只要求出参数 a, u , 就能求出 $x^{(1)}(t)$, 进而求出 $x^{(0)}$ 的预测值。

由导数定义知: $\frac{dx}{dt} = \lim_{\Delta t \rightarrow 0} \frac{x(t + \Delta t) - x(t)}{\Delta t}$, 当 Δt 很小时并且取很小的1单位时, 则近似地有

$$x(t + 1) - x(t) = \frac{\Delta x}{\Delta t}, \text{ 写成离散形式为: } \frac{\Delta x}{\Delta t} = x(k + 1) - x(k) = \Delta^{(1)}(x(k + 1))$$

这表示 $\Delta x / \Delta t$ 是 $x(k + 1)$ 的一次累减生成, 因此 $\Delta x / \Delta t$ 是 $x(k + 1)$ 和 $x(k)$ 二元组合等效值, 则称 $x(k + 1)$ 与 $x(k)$ 的二元组合为偶对, 即为 $[x(k + 1), x(k)]$, 于是定义一个从 $[x(k + 1), x(k)]$ 到 $\Delta x / \Delta t$ 的映射

$$F : [x(k + 1), x(k)] \rightarrow \frac{dx}{dt}$$

1. 灰色预测模型GM(1,1)原理分析

若定义 $R(t)$ 是 t 时刻 $\Delta x/\Delta t$ 的背景值（就是对应的 x 值），那么每一个 $\Delta x/\Delta t$ 都有一个偶对背景值 $R(t)$ 与之对应，现在考虑一阶微分方程 $dx/dt + ax = u$ ，它是 a 与 u 的线性组合。那么，作这种线性组合时， $\Delta x/\Delta t$ 所对应的背景值究竟取偶对的哪一个呢？如果认为在 $\Delta t = 1$ 的很短时间内，变量 $x(t) \rightarrow x(t + \Delta t)$ 之间不会出现突变量，那么可取偶对的平均值作为背景值。即：

$$z(t) = \frac{1}{2}[x(k) + x(k+1)]$$

于是白化形式的微分方程 $\frac{dx^{(1)}}{dt} + ax^{(1)} = u$ 离散化，即得 $\Delta^{(1)}(x^{(1)}(k+1)) + az^{(1)}(x(k+1)) = u$

其中， $\Delta^{(1)}(x^{(1)}(k+1))$ 为 $x^{(1)}$ 在 $k+1$ 时刻的累减生成序列， $z^{(1)}(k+1)$ 为 $\frac{dx^{(1)}}{dt}$ 在 $k+1$ 时刻的背景值。故，将 $\Delta^{(1)}(x^{(1)}(k+1)) = x^{(1)}(k+1) - x^{(1)}(k) = x^{(0)}(k+1)$ 与 $z^{(1)}(k+1)$ 代入得：

$$x^{(0)}(k+1) = a[-\frac{1}{2}(x^{(1)}(k) + x^{(1)}(k+1))] + u$$

1. 灰色预测模型GM(1,1)原理分析

4. 对累加生成数据做均值生成 B 与常数项向量 Y_n :

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

5. 用最小二乘法求解灰参数 \hat{a} , 则 $\hat{a} = (B^T B)^{-1} B^T Y_n$

6. 将灰参数 \hat{a} 带入 $\frac{dx^{(1)}}{dt} + ax^{(1)} = u$ 并求解得 $\hat{x}^{(1)}(t+1) = \left(x^{(1)}(1) - \frac{u}{a}\right)e^{-at} + \frac{u}{a}$

7. 将上述结果累减还原, 即可得到预测值: $\hat{x}^{(0)}(t+1) = \hat{x}^{(1)}(t+1) - \hat{x}^{(1)}(t)$

1. 灰色预测模型GM(1,1)原理分析

8. 利用模型进行预测: $\hat{x}^{(0)} = \left(\underbrace{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(n)}_{\text{原始数据}}, \underbrace{\hat{x}^{(0)}(n+1), \dots, \hat{x}^{(0)}(n+m)}_{\text{预测数据}} \right)$

9. 对建立的灰色模型进行精度检验:

(1) 残差检验: 残差 $E(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$, 相对误差 $Q(k) = \frac{E^{(0)}(k)}{x^{(0)}(k)}, k = 2, 3, \dots, n$

(2) 后验差检验:

1) 均值与标准差 $\bar{X} = \frac{1}{n} \sum_{k=1}^n x^{(0)}(k), S_1 = \sqrt{\frac{1}{n} \sum_{k=1}^n [x^{(0)}(k) - \bar{X}]^2}$

2) 残差的均值与标准差 $\bar{E} = \frac{1}{n-1} \sum_{k=2}^n E(k), S_2 = \sqrt{\frac{1}{n-1} \sum_{k=2}^n [E(k) - \bar{E}]^2}$

3) 后验差比值 $C = \frac{S_2}{S_1}$

1. 灰色预测模型GM(1,1)原理分析

4) 小概率误差 $P = p(|E(k) - \bar{E}| < 0.6745S_1)$

5) 关联度检验 $R = \frac{E_{\min} + \rho E_{\max}}{E + \rho E_{\max}}, \rho = 0.5$

(3) 预测精度等级对照如表：

等级	方差比	小概率误差	关联度	平均相对残差		
好（一级）	$C \leq 0.35$	$P \geq 0.95$	$R > 0.9$	预测等级精度	$\bar{\varphi}$	φ_n
合格（二级）	$0.35 < C \leq 0.5$	$0.95 > P \geq 0.8$	$R: 0.8 \sim 0.9$	优	< 0.01	< 0.01
勉强合格（三级）	$0.5 < C \leq 0.65$	$0.8 > P \geq 0.7$	$R: 0.7 \sim 0.8$	合格	$0.01 \sim 0.05$	$0.01 \sim 0.05$
不合格（四级）	$C > 0.65$	$P < 0.7$	$R: 0.6 \sim 0.7$ （满意）	勉强合格	$0.05 \sim 0.1$	$0.05 \sim 0.1$

1. 灰色预测模型GM(1,1)原理分析

由灰色预测方法原理， $-a$ 主要控制系统发展态势的大小，即反映预测的发展态势，被称为发展系数； u 的大小反映了数据变化的关系，被称为灰色作用量。

其中：

- ①当 $-a < 0.3$ 时, GM(1,1) 模型可用于中长期预测;
- ②当 $0.3 < -a < 0.5$ 时, GM(1,1) 模型可用于短期预测,中长期预测慎用;
- ③当 $0.5 < -a < 1$ 时, 应采用GM(1,1) 改进模型,包括GM(1,1) 残差修正模型;
- ④当 $-a > 1$ 时,不宜采用GM(1,1) 模型,可考虑其他预测方法。

2. GM(1,1)模型MATLAB编程实现



```
GM11_model.m
1 function gm11 = GM11_model(X, td)
2 % GM11_model用于灰色模型G(1, 1)的建立和预测
3 % 输入参数X为原始数据, td为未来预测期数
4 % 输出参数gm11为一个结构体, 包括:
5 % Coeff_a为发展系数, Coeff_u为灰作用量,
6 % Predict_Value为预测值, 包括当前值和未来td期预测值
7 % AbsoluteError为绝对误差, RelativeErrorMean为相对误差均值
8 % C为方差比, P误差为小概率, R为关联度
9
10 %% 输入参数的控制和默认值
11 if nargin < 2
12     warning('输入参数为2个, td将使用默认值5...')
13     td = 5;
14 elseif td < 0
15     warning('未来预测期数td不能为负值, 将使用默认值0...')
16     td = 0;
17 end
18
19 %% 数据预处理: 累加, 平均
20 n = length(X); %原始数据个数
21 Ago = cumsum(X); %原始数据一次累加, 得到1-AGO序列xi(1)。
22 %Z(i)为xi(1)的紧邻均值生成序列
23 Z = (Ago(1:n-1) + Ago(2:end))/2;
24
25 %% 构造B和Yn
26 Yn = X(2:end)'; %Yn为常数项向量, 即x(2), x(3)...
27 B = [-Z; ones(1, n-1)]'; %累加生成数据做均值
```

```
29 %% 最小二乘求解发展系数a和灰色作用量u
30 LS_solution = (B'*B)\(B'*Yn); % 利用公式求出a, u
31 a = LS_solution(1); % 发展系数a的值
32 u = LS_solution(2); % 灰作用量u的值
33
34 %% 建立灰度GM(1, 1)模型, 白化一阶一元微分方程
35 F = [X(1), (X(1) - u/a)./exp(a*(1:n+td-1))+u/a];
36 %还原序列, 得到预测数据
37 PreData = [F(1), F(2:end) - F(1:end-1)];
38
39 %% 可视化
40 t = 1:n;
41 plot(t, X, 'ko--', 'MarkerFaceColor', 'k') % 原数据图像
42 hold on;
43 grid on
44 %预测当前数据图像
45 plot(t, PreData(1:n), 'b*-', 'LineWidth', 1.5)
46 %未来td期数据图像
47 plot(n:n+td, PreData(n:n+td), 'r*-', 'LineWidth', 1.5)
48 title('GM(1, 1) model— Original VS Current And Future Predict')
49 legend('OriginalData', 'ForecastCurrentData', ...
50         'ForecastFutureData', 'Location', 'best')
51 legend('boxoff')
52 xlabel('Time')
53 ylabel('Value')
```

2. GM(1,1)模型MATLAB编程实现



```
55 %% 后验差检验
56 Err = abs(X - PreData(1:n)); %真实值与预测值误差
57 q = mean(Err./X); %平均相对误差
58 XVar = std(X,1); %原数据的标准方差, 前置因子1/n
59 ErrVar = std(Err(2:end),1); %残差(2:end)的标准方差, 前置因子1/n
60 C = ErrVar/XVar; %后验方差比
61 P = sum(abs(Err - mean(Err)) < 0.6745*XVar)/n; %小误差概率
62 R_k = (min(Err)+0.5*max(Err))./(Err+0.5*max(Err)); % rho = 0.5
63 R = sum(R_k)/length(R_k); %关联度
64
65 %% 计算过程变量组合输出结构体变量
66 gm11.Coeff_a = a;
67 gm11.Coeff_u = u;
68 gm11.Predict_Value = PreData;
69 gm11.AbsoluteError = Err;
70 gm11.RelativeErrorMean = q;
71 gm11.R = R;
72 gm11.C = C;
73 gm11.P = P;
74 end
```

例1: 1995-2008年间污水排放量, 检测结果如表, 使用灰色GM(1,1)预测未来五年的污水排放量。

年份	1995	1996	1997	1998
污水量/亿吨	174	179	183	189
年份	1999	2000	2001	2002
污水量/亿吨	207	234	220.5	256
年份	2003	2004	2005	2006
污水量/亿吨	270	285	300	320
年份	2007	2008		
污水量/亿吨	344	365		

3. GM(1,1)案例求解与分析

```
>> X = [174, 179, 183, 189, 207, 234, 220.5, 256, 270, 285 300 320 344 365];
```

```
>> gm11 = GM11_model(X,5)
```

```
gm11 =
```

包含以下字段的 struct:

Coeff_a: -0.0621

Coeff_u: 156.7876

Predict_Value: [1×19 double]

AbsoluteError: [1×14 double]

RelativeErrorMean: 0.0185

R: 0.7182

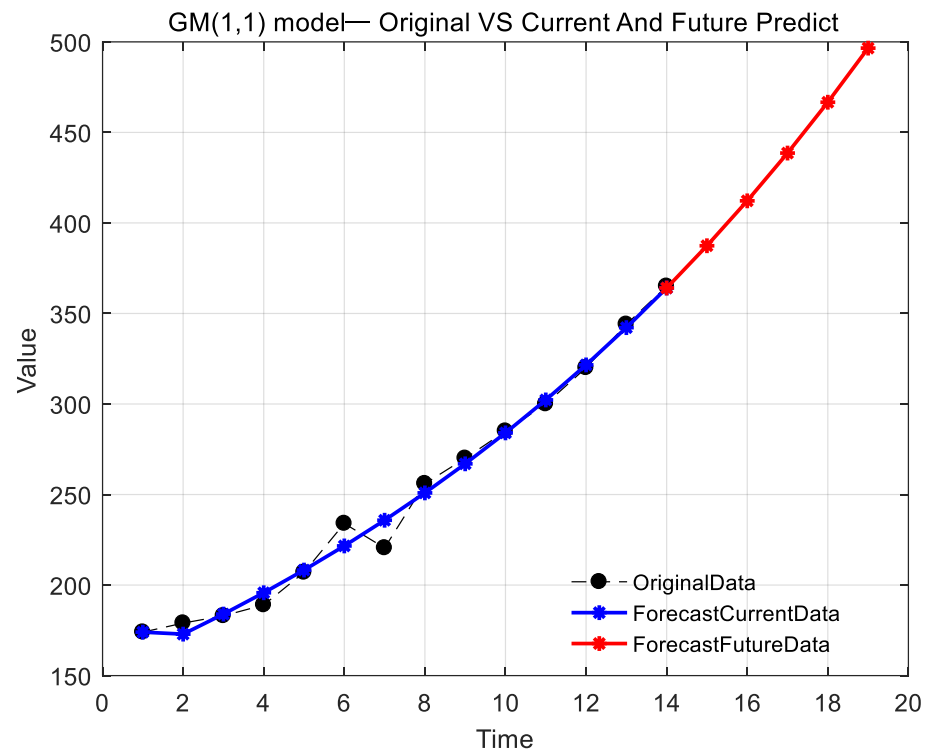
C: 0.0724

P: 1

根据模型评价标准:

- $p=1$, $C=0.0724$, 预测等级为: 好;
- 相对误差均值0.0185, 合格
- 关联度0.7182, 勉强合格
- 由于 $-a$ 系数小于0.3, 适合中长期预测。

从运行结果看, 对于线性的数据使用GM(1,1)预测, 其拟合效果还是不错。



% 未来五期的预测值

```
>> pd = gm11.Predict_Value(end-5+1:end)
```

```
pd =
```

387.3958 412.1987 438.5896 466.6702 496.5486

3. GM(1,1)案例求解与分析

例2：历年硕士招生人数，数据如程序所示，数据来源于国家数据网。

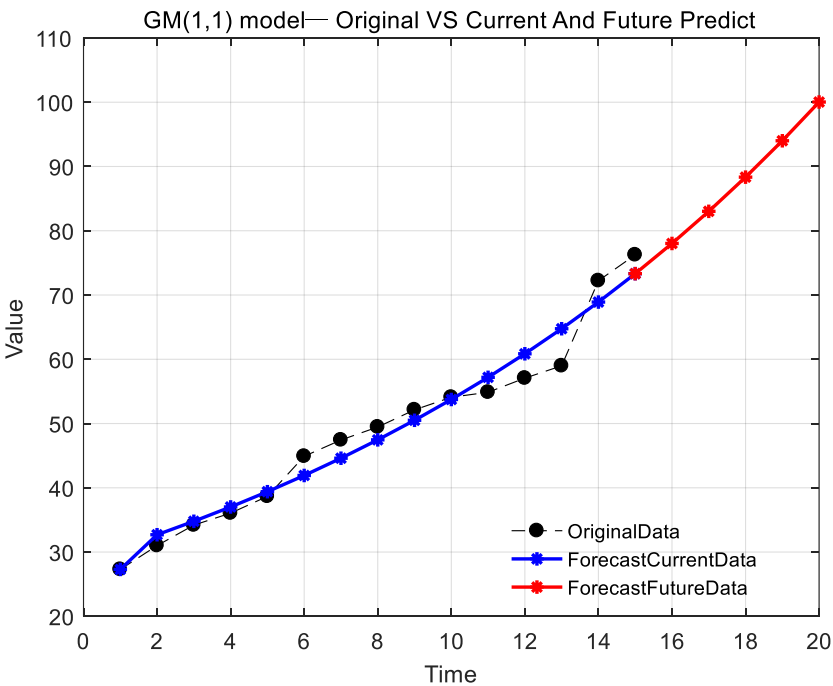
2018年	2017年	2016年	2015年	2014年	2013年	2012年	2011年
76.2464	72.2225	58.9812	57.0639	54.8689	54.0919	52.1303	49.4609
2010年	2009年	2008年	2007年	2006年	2005年	2004年	
47.4415	44.9042	38.6658	36.0590	34.1970	31.0037	27.3002	

```
>> X = [76.2464 72.2225 58.9812 57.0639 54.8689 54.0919 52.1303  
49.4609 47.4415 44.9042 38.6658 36.0590 34.1970 31.0037 27.3002];
```

```
>> X = fliplr(X);  
>> gm11 = GM11_model(X,5)
```

```
>> num = gm11.Predict_Value(end-  
4:end)  
num =  
78.0195 83.0202 88.3413  
94.0035 100.0286
```

```
gm11 =  
包含以下字段的 struct:  
  
Coeff_a: -0.0621  
Coeff_u: 29.9939  
Predict_Value: [1×20 double]  
AbsoluteError: [1×15 double]  
RelativeErrorMean: 0.0408  
R: 0.6273  
C: 0.1037  
P: 1
```



3. GM(1,1)案例求解与分析

有关建模的问题说明

- (1) 给定原始序列 x_0 中的数据不一定要全部用来建立模型，对原始序列的取舍不同，可得模型不同，即 a 、 u 的值不同。
- (2) 建模的数据取舍应保证建模序列等时距、相连，不得有跳跃出现。
- (3) 一般建模数据序列应当由最新数据及其相邻数据构成，当再出现新数据时，可采取以下两种处理方法：
 - 一是将新信息加入原始序列中，重估参数；
 - 二是去掉原始序列中最老的一个数据，再加上最新数据，所形成序列和原序列维数相等，再重估参数。

3. GM(1,1)案例求解与分析

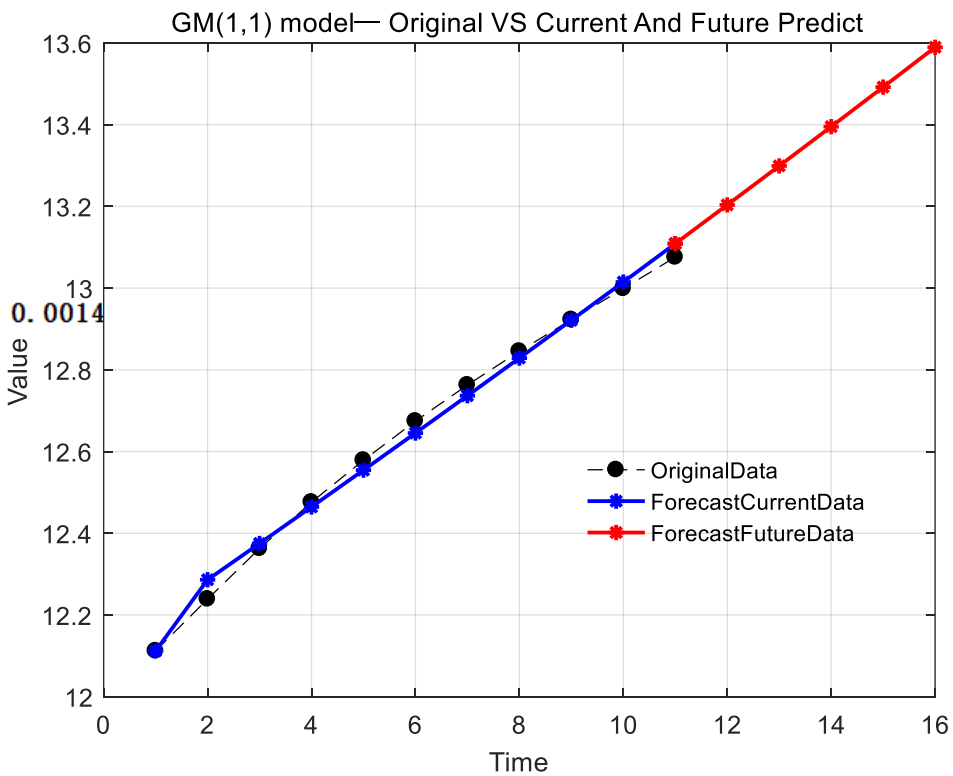
例3：人口预测

年份	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
人口(亿)	12.2389	12.3626	12.4761	12.5786	12.6743	12.7627	12.8453	12.9227	12.9988	13.0756

```
>> gm11 = GM11_model(X(21:end)', 5)
gm11 =
    包含以下字段的 struct:

    Coeff_a: -0.0072
    Coeff_u: 12.1548
    Predict_Value: [1×16 double]
    AbsoluteError: [0 -0.0473 -0.0124 0.0117 0.0241 0.0291 0.0261 0.0167 0.0014
    RelativeErrorMean: 0.0016
    R: 0.6020
    C: 0.0409
    P: 1
```

模型基本满足各种检验，可用于中长期预测。



4. GM(1,1)残差模型

若用原始时间序列 $x^{(0)}$ 建立的GM(1,1)模型检验不合格或精度不理想时, 要对建立的GM(1,1)模型进行残差修正或提高模型的预测精度。修正的方法是建立GM(1,1)的残差模型。

若用原始序列建立的GM(1,1)模型

$$\hat{x}^{(1)}(t+1) = \left(x^{(1)}(1) - \frac{u}{a} \right) e^{-at} + \frac{u}{a}$$

对于残差序列 $\varepsilon^{(0)}(t) = X^{(1)}(t) - \hat{X}^{(1)}(t)$, $(j=1,2,\dots,n)$, 若存在 k_0 , 使得当 $k > k_0$ 时, $\varepsilon^{(0)}(k)$ 的符号一致, 且 $n - k_0 \geq 4$, 则称残差序列

$$E^{(0)} = \{\varepsilon^{(0)}(k_0), \varepsilon^{(0)}(k_0+1), \dots, \varepsilon^{(0)}(n)\}$$

为可建模残差尾部。

4. GM(1,1)残差模型

计算残差序列 $E^{(0)} = \{\varepsilon^{(0)}(k_0), \varepsilon^{(0)}(k_0 + 1), \dots, \varepsilon^{(0)}(n)\}$ 的累加生成序列 $E^{(1)}$, 并建立相应的GM(1,1)模型:

$$\hat{\varepsilon}^{(1)}(k+1) = \left[\varepsilon^{(0)}(k_0) - \frac{\mu_\varepsilon}{a_\varepsilon} \right] e^{-a_\varepsilon(k-k_0)} + \frac{\mu_\varepsilon}{a_\varepsilon}$$

得修正模型:

$$\hat{X}^{(1)}(k+1) = \left[X^{(0)}(1) - \frac{\mu}{a} \right] e^{-ak} + \frac{\mu}{a} \pm a_\varepsilon \delta(k-k_0) \left[\varepsilon^{(0)}(k_0) - \frac{\mu_\varepsilon}{a_\varepsilon} \right] e^{-a_\varepsilon(k-k_0)} + \frac{\mu_\varepsilon}{a_\varepsilon}$$

其中 $\delta(k-k_0) = \begin{cases} 1 & k \geq k_0 \\ 0 & k < k_0 \end{cases}$, 正负号取值与残差尾部符号一致。

5. 基于背景值优化的改进GM(1,1)模型

设原始非负数据序列为 $X^{(0)}$ ，做一次累加(1-AGO)得 $X^{(1)}$ ，对生成的序列 $X^{(1)}(k)$ 建立灰色白化微分方程为：

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

将上式在区间 $[k, k+1]$ 上积分，得： $x^{(1)}(k+1) - x^{(1)}(k) + a \int_k^{k+1} x^{(1)}(t) dt = u$, $k = 1, 2, \dots, n-1$

即： $x^{(0)}(k+1) + a \int_k^{k+1} x^{(1)}(t) dt = u$ ，设 $z^{(1)}(k+1) = \int_k^{k+1} x^{(1)}(t) dt$ 为 $x^{(1)}(t)$ 在匀间 $[k, k+1]$ 上的背

景值，则有 $x^{(0)}(k+1) + az^{(1)}(k+1) = u$ ，然而函数 $x^{(1)}(t)$ 并不知道，但根据一阶微分方程的解为指数函数，故可令 $x^{(1)}(t) = ce^{bt}$ ，并假设该曲线过点 $(k, x^{(0)}(k))$ 及 $(k+1, x^{(0)}(k+1))$ ，则有

$$x^{(1)}(k) = ce^{bk}, \quad x^{(1)}(k+1) = ce^{b(k+1)}$$

由以上两式可解得： $b = \ln x^{(1)}(k+1) - \ln x^{(1)}(k)$, $c = \frac{x^{(1)}(k)}{e^{bk}} = \frac{[x^{(1)}(k)]^{k+1}}{[x^{(1)}(k+1)]^k}$

5. 基于背景值优化的改进GM(1,1)模型

因此背景值为：
$$z^{(1)}(k+1) = \int_k^{k+1} x^{(1)}(t) dt = \int_k^{k+1} ce^{bt} dt = \frac{x^{(1)}(k+1) - x^{(1)}(k)}{\ln x^{(1)}(k+1) - \ln x^{(1)}(k)}$$

待辨参数 a, u 可用最小二乘法求出 $\hat{\Phi} = [\hat{a} \ \hat{u}]^T = (B^T B)^{-1} B^T Y$

其中：
$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \quad B = \begin{bmatrix} -z^{(0)}(2) & 1 \\ -z^{(0)}(3) & 1 \\ \vdots & \vdots \\ -z^{(0)}(n) & 1 \end{bmatrix}, \quad z^{(1)}(k+1) \text{ 为由背景值确定。}$$

求出参数 a, u 后，可得白化方程的离散解为：
$$\hat{x}^{(1)}(k+1) = \left[x^{(1)}(1) - \frac{\hat{u}}{\hat{a}} \right] e^{-\hat{a}k} + \frac{\hat{u}}{\hat{a}}$$

进行一次累减还原，可得原始数据序列的模拟值为：

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) = (1 - e^a) \left[x^{(1)}(1) - \frac{\hat{u}}{\hat{a}} \right] e^{-\hat{a}k}$$

5. 基于背景值优化的改进GM(1,1)模型

修改程序对应代码如下，并修改函数名为：GM11_Improve

```
21 - Ago = cumsum(X); %原始数据一次累加,得到1-AGO序列xi(1)。  
22 - %Z(i)为由背景值确定  
23 - Z = (Ago(1:n-1) - Ago(2:end))./(log(Ago(1:n-1)) - log(Ago(2:end)));  
24 -  
25 - %% 构造B和Yn  
26 - Yn = X(2:end)'; %Yn为常数项向量,即x(2),x(3)...  
27 - B = [-Z; ones(1,n-1)]'; %累加生成数据做均值  
28 -  
29 - %% 最小二乘求解发展系数a和灰色作用量u  
30 - LS_solution = (B'*B)\(B'*Yn); % 利用公式求出a, u  
31 - a = LS_solution(1); % 发展系数a的值  
32 - u = LS_solution(2); % 灰作用量u的值  
33 -  
34 - %% 建立灰度GM(1,1)模型,白化一阶一元微分方程  
35 - %F = [X(1), (X(1) - u/a)./exp(a*(1:n+td-1))+u/a];  
36 - %模拟序列,得到预测数据  
37 - PreData = [Ago(1), (1-exp(a))*(Ago(1) - u/a)*exp(-a*(1:n+td-1))];  
38 -
```

```
>> gm11 = GM11_Improve(X, 5)
```

```
gm11 =
```

包含以下字段的 [struct](#):

```
        Coeff_a: -0.0620  
        Coeff_u: 157.2030  
        Predict_Value: [1×19 double]  
        AbsoluteError: [1×14 double]  
        RelativeErrorMean: 0.0184  
                R: 0.7262  
                C: 0.0731  
                P: 1
```

```
>> gm11 = GM11_model(X, 5)
```

```
gm11 =
```

包含以下字段的 [struct](#):

```
        Coeff_a: -0.0621  
        Coeff_u: 156.7876  
        Predict_Value: [1×19 double]  
        AbsoluteError: [1×14 double]  
        RelativeErrorMean: 0.0185  
                R: 0.7182  
                C: 0.0724  
                P: 1
```

二. GM(1,n)原理和求解步骤

- GM (1,n) 模型的预测原理：与GM(1,1)类似，不同在于输入数据变量是n个。
- GM(1,n)模型的建模过程：

设系统有特征数据序列： $X_1^{(0)} = (x_1^{(0)}(1), x_1^{(0)}(2), \dots, x_1^{(0)}(n))$

相关因素序列：

$$X_2^{(0)} = (x_2^{(0)}(1), x_2^{(0)}(2), \dots, x_2^{(0)}(n)), \dots, X_N^{(0)} = (x_N^{(0)}(1), x_N^{(0)}(2), \dots, x_N^{(0)}(n))$$

(1)令 $X_i^{(0)} (i=1,2,\dots,N)$ 的1-AGO序列为 $X_i^{(1)}$, 其中 $X_i^{(1)}(k) = \sum_{k=1}^n x_i^{(0)}(k)$, $(i=1,2,\dots,n)$

(2)生成 $X_1^{(1)}$ 的紧邻均值序列 $Z_1^{(1)}$, 其中 $Z_1^{(1)}(k) = \frac{1}{2}[X_1^{(1)}(k) + X_1^{(1)}(k-1)]$, $k=2,3,\dots,n$

称 $x_1^{(0)}(k) + aZ_1^{(1)}(k) = \sum_{i=2}^N b_i x_i^{(1)}(k)$ 为GM(1,n)模型。

在GM(1,n)模型中， a 被称为发展系数， b_i 称为驱动系数， $b_i x_i^{(1)}(k)$ 称为驱动项。

1. GM(1,n)原理和求解步骤

令

$$B = \begin{bmatrix} -Z^{(1)}(2) & x_2^{(1)}(2) & \cdots & x_N^{(1)}(2) \\ -Z^{(1)}(3) & x_2^{(1)}(3) & \cdots & x_N^{(1)}(3) \\ \vdots & \vdots & \vdots & \vdots \\ -Z^{(1)}(n) & x_2^{(1)}(n) & \cdots & x_N^{(1)}(n) \end{bmatrix}, Y = \begin{bmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \\ \vdots \\ x_1^{(0)}(n) \end{bmatrix}$$

再令 $\beta = (a, b_2, b_3, \dots, b_N)^T$, 由最小二乘参数估计可得 $\beta = (B^T B)^{-1} B^T Y$

当 $X_i^{(1)} (i = 1, 2, \dots, N)$ 变化幅度较小时, GM(1,n)近似时间相应式为:

$$\hat{x}_1^{(1)}(k+1) = \left[x_1^{(0)}(1) - \frac{1}{a} \sum_{i=2}^N b_i x_i^{(1)}(k+1) \right] e^{-ak} + \frac{1}{a} \sum_{i=2}^N b_i x_i^{(1)}(k+1)$$

累减还原式为 $\hat{x}_1^{(0)}(k+1) = \hat{x}_1^{(1)}(k+1) - \hat{x}_1^{(1)}(k)$

差分模拟式为 $\hat{x}_1^{(0)}(k) = -aZ_1^{(1)}(k) + \sum_{i=2}^N b_i \hat{x}_i^{(1)}(k)$

2. GM(1,n)模型的MATLAB实现



```
GM1N_model.m  +
1  function gmlN = GM1N_model(Y,X,select,X0)
2  % GM1N_model用于灰色模型G(1,1)的建立和预测,并进行可视化;
3  % 输入参数Y为因变量,X是自变量,数据格式:一变量对应一列,一行对应一个观测值;
4  % X0为未来预测数据,数据格式同X; select值为1或0,对应累减和差分模拟;
5  % 输出参数gmlN为一个结构体,包括: Coeff_a为发展系数, Coeff_b为灰作用量,
6  % PredictCurr_Value为当前预测值, AbsoluteError为绝对误差,
7  % C为方差比, P误差为小概率, RelativeErrorMean为相对误差均值,
8  % PredictFuture_Value为未来预测值。
9
10 n = size(X,1);
11 Ago = cumsum(Y); %一次累加
12 if select == 1
13     Z = (Ago(1:n-1) + Ago(2:end))/2; %Z(i)为x1(1)的紧邻均值生成序列
14 else
15     %Z(i)为由背景值确定
16     Z = (Ago(1:n-1) - Ago(2:end))./(log(Ago(1:n-1)) - log(Ago(2:end)));
17 end
18 X1 = cumsum(X); %对应自变量x的每一列进行累加一次
19
20 %% 构造Yn和B, 并采用最小二乘求解a和b, b的个数对应自变量个数
21 Yn = Y(2:end); %Yn为常数项向量
22 B = [-Z, X1(2:end,:)];
23 beta = (B'*B)\(B'*Yn); %最小二乘求解
24 a = beta(1);
25 b = beta(2:end);
26
27 %% 建立白化方程并进行数据还原预测
28 F = zeros(n,1);
```

```
29 F(1) = Y(1);
30 for k = 1:n-1 %当前值的预测
31     F(k+1) = (Y(1) - 1/a*(dot(b, X1(k+1,:))))*exp(-a*k) + ...
32         1/a*(dot(b, X1(k+1,:)));
33 end
34 if select == 1 %累减还原序列, 当前预测数据
35     PreData = [F(1);F(2:end) - F(1:end-1)];
36 else %差分模拟还原序列
37     diff_PreData = zeros(n,1);
38     diff_PreData(1) = Y(1);
39     for k = 2:n
40         diff_PreData(k) = -a*Z(k-1) + dot(b, X1(k,:));
41     end
42 end
43
44 %% 可视化
45 t = 1:n;
46 plot(t,Y,'ko--','MarkerFaceColor','k','LineWidth',1.5);
47 hold on;grid on;
48 if select == 1
49     plot(t,PreData,'r*-','LineWidth',1.5); %累减
50 else
51     plot(t,diff_PreData,'r*-','LineWidth',1.5); %差分模拟
52 end
53 title('GM(1,1) model— Original VS Future Predict')
54 legend('OriginalData','ForecastCurrentData','Location','best')
55 legend('boxoff')
56 xlabel('Time')
57 ylabel('Value')
58
```

2. GM(1,n)模型的MATLAB实现



%% 可视化

```
t = 1:n;
plot(t, Y, 'ko--', 'MarkerFaceColor', 'k', 'LineWidth', 1.5);
hold on; grid on;
if select == 1
    plot(t, PreData, 'r*- ', 'LineWidth', 1.5); %累减
else
    plot(t, diff_PreData, 'r*- ', 'LineWidth', 1.5); %差分模拟
end
title('GM(1,1) model— Original VS Future Predict')
legend('OriginalData', 'ForecastCurrentData', 'Location', 'best')
legend('boxoff')
xlabel('Time')
ylabel('Value')
```

%% 未来数据的预测

```
if nargin == 4
    m = size(X0, 1); %预测数据量
    preFuture = zeros(m, 1);
    for k = 1:m
        preFuture(k) = (Y(1) - 1/a*(dot(b, X0(k, :))))*exp(-a*k) + ...
            1/a*(dot(b, X0(k, :)));
    end
    plot(n+1:n+m, preFuture, 'b*- ', 'LineWidth', 1.5)
    legend('OriginalData', 'ForecastCurrentData', ...
        'ForecastFutureData', 'Location', 'best')
    legend('boxoff')
    hold off
end
```

```
if nargin == 4
    gmlN.PredictFuture_Value = preFuture;
end
gmlN.AbsoluteError = e;
gmlN.RelativeError = e./Y;
gmlN.RelativeErrorMean = q;
gmlN.R = R;
gmlN.C = C;
gmlN.P = P;
end
```

%% 后验差检验

```
if select == 1
    e = abs(Y - PreData); %真实值与预测值误差
else
    e = abs(Y - diff_PreData); %真实值与预测值误差
end
q = mean(e./Y); %平均相对误差
s1 = std(Y, 1); %因变量标准方差
s2 = std(e(2:end), 1); %误差的标准方差
C = s2/s1; %方差比
P = 0; %小误差概率
for i = 1:n
    if(abs(e(i) - mean(e(2:end))))<0.6745*s1
        P = P + 1;
    end
end
P = P/n; %小误差概率
R_k = (min(e(2:end))+0.5*max(e))./(e(2:end)+0.5*max(e)); % rho=0.5
R = sum(R_k)/length(R_k); %关联度
```

%% 计算过程变量组合输出结构体变量

```
gmlN.Coeff_a = a;
gmlN.Coeff_b = b;
if select == 1
    gmlN.select = '累减还原形似';
    gmlN.PredictCurr_Value = PreData;
else
    gmlN.PredictCurr_Value = diff_PreData;
    gmlN.select = '差分模拟形式';
end
```

3. GM(1,n)案例分析

例4：城市居民消费支出预测

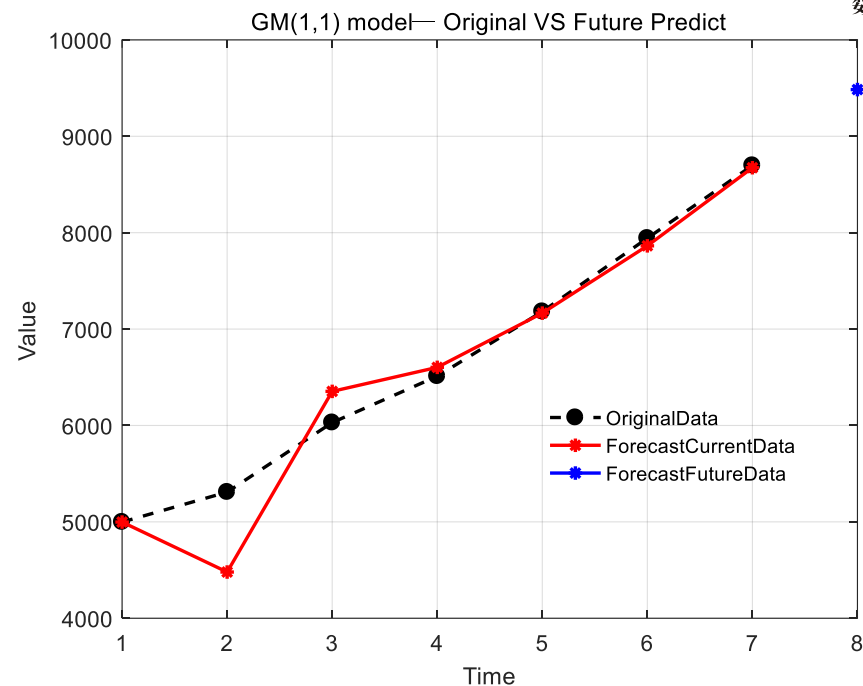
已知我国在2000~2006年的城市居民人均消费支出与收入，以及城市人口规模的调查统计数据（见下表），试建立GM(1,n)预测模型，并预测2007年城市居民人均收入和人口数量分别为15000元和6亿人时，城市居民人均消费总支出。

年份	2000	2001	2002	2003	2004	2005	2006	2007
总支出/元	4998	5309	6029.9	6510.9	7182.1	7942.9	8697.3	
总收入/元	6316.8	6868.9	8177.4	9061.2	10128.8	11320.8	12719.1	15000
人口/万人	45906	48064	50212	52376	54283	56212	57706	60000

3. GM(1,n)案例分析

```
>> Y = [4998 5309 6029.9 6510.9 7182.1 7942.9 8697.3]';  
>> X = [6316.8 6868.9 8177.4 9061.2 10128.8 11320.8 12719.1;  
        45906 48064 50212 52376 54283 56212 57706]';  
>> X0 = [15000 60000];  
>> gm1N = GM1N_model(Y,X,1, X0) %按照累减还原序列
```

```
gm1N =  
    包含以下字段的 struct:  
  
        Coeff_a: 2.2824  
        Coeff_b: [2×1 double]  
        select: '累减还原形似'  
        PredictCurr_Value: [7×1 double]  
        PredictFuture_Value: 9.4846e+03  
        AbsoluteError: [7×1 double]  
        RelativeError: [7×1 double]  
        RelativeErrorMean: 0.0341  
        R: 0.7730  
        C: 0.2291  
        P: 1
```



根据检验结果

- (1) 关联度 $r=0.7730>0.7$, 满足检验准则, 说明拟合程度好, 可以选择预测模型去进行数据预测。
- (2) 后验差比值 $C=0.2291<0.35$, 并且小残差概率 $P=1>0.95$, 说明该模型为一个精度优的预测模型, 可以选择模型去进行数据预测。
- (3) 因为平均相对残差 $m\Phi=0.0341<0.05$ 并且 $\Phi(7)=0.0027<0.05$, 一说明这是一个精度优的模型, 可以选择这个模型去预测。

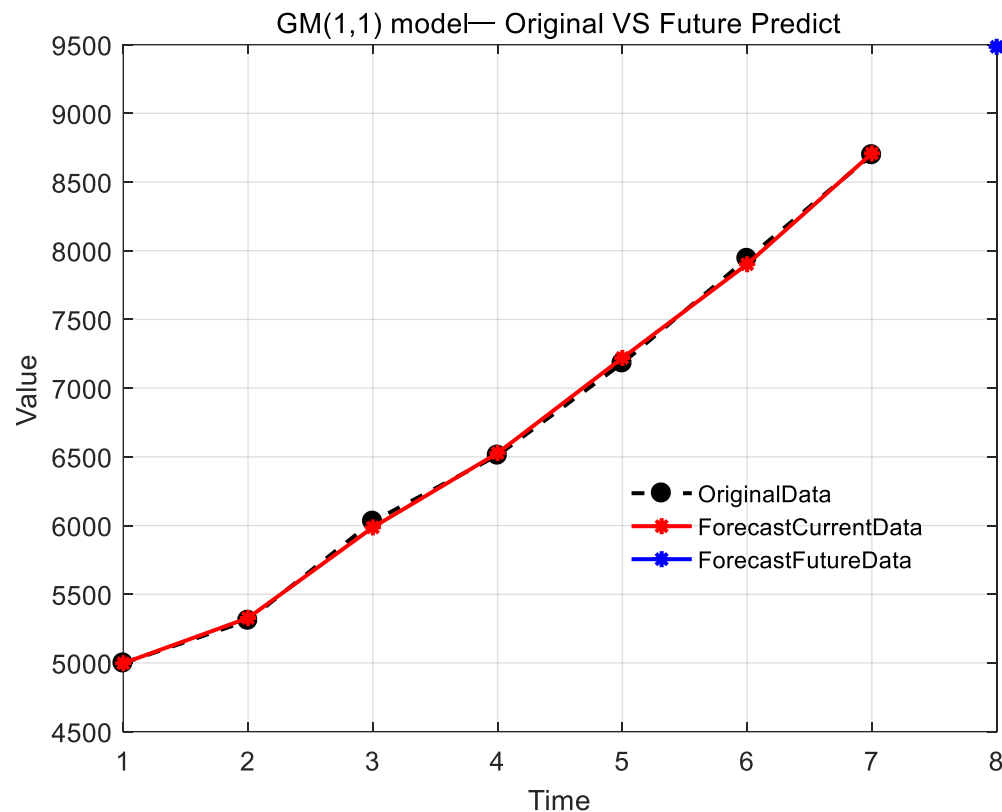
3. GM(1,n)案例分析

```
>> Y = [4998 5309 6029.9 6510.9 7182.1 7942.9 8697.3]';  
>> X = [6316.8 6868.9 8177.4 9061.2 10128.8 11320.8 12719.1;  
        45906 48064 50212 52376 54283 56212 57706]';  
>> X0 = [15000 60000];  
>> gm1N = GM1N_model(Y,X,0,X0) %差分模拟形式
```

```
gm1N =
```

包含以下字段的 [struct](#):

```
    Coeff_a: 1.8217  
    Coeff_b: [2×1 double]  
    PredictCurr_Value: [7×1 double]  
    select: '差分模拟形式'  
    PredictFuture_Value: 9.2137e+03  
    AbsoluteError: [7×1 double]  
    RelativeError: [7×1 double]  
    RelativeErrorMean: 0.0044  
    R: 0.6922  
    C: 0.0214  
    P: 1
```



$$\hat{x}_1^{(1)}(k+1) = \left[x_1^{(0)}(1) - \frac{1}{2.2822} \left(1.2547 x_2^{(1)}(k+1) + 0.0665 x_3^{(1)}(k+1) \right) \right] e^{-2.2822k} + \frac{1}{2.2822} \left[1.2547 x_2^{(1)}(k+1) + 0.0665 x_3^{(1)}(k+1) \right]$$

3. GM(1,n)案例分析



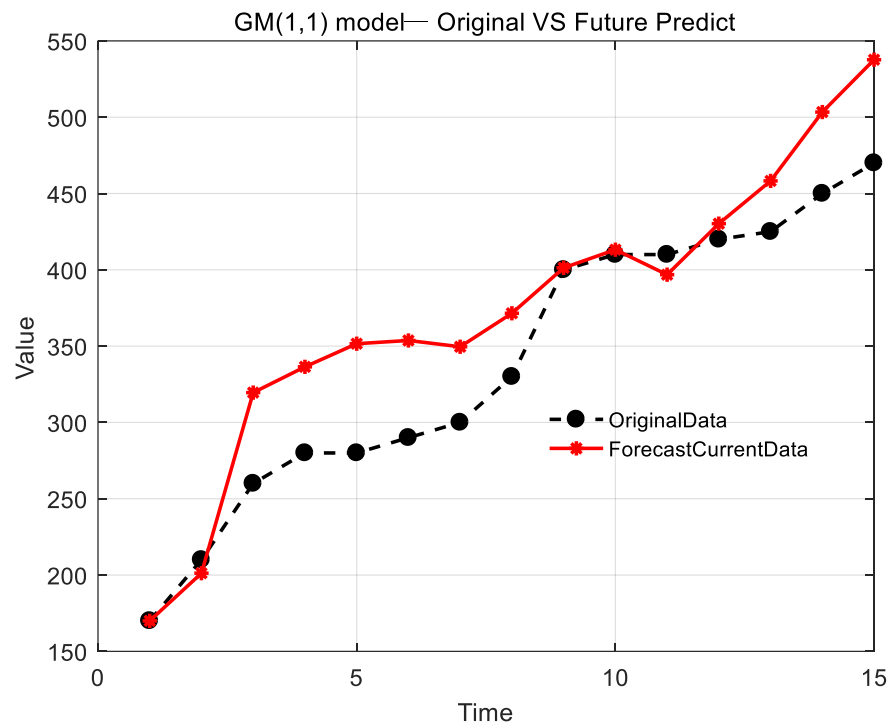
例5：考察15名不同程度的烟民的每日抽烟量（支）、饮酒量（啤酒/升）与其心电图指标(zb)的对应数据，试建立心电图指标关于日抽烟量和日饮酒量的灰度模型。

日抽烟	日饮酒	心电图	日抽烟	日饮酒	心电图	日抽烟	日饮酒	心电图	日抽烟	日饮酒	心电图
30	10	280	45	14	410	25	13	300	48	16	425
25	11	260	20	12	170	23	13	290	50	18	450
35	13	330	18	11	210	40	14	410	55	19	470
40	14	400	25	12	280	45	15	420			

```
>> data=[30 10 280;25 11 260;35 13 330;40 14 400;45 14  
410;20 12 170;18 11 210;25 12 280;25 13 300;23 13 290;  
40 14 410;45 15 420;48 16 425;50 18 450;55 19 470];  
>> data = sortrows(data,3);  
>> Y = data(:,3);  
>> X = data(:,1:2);  
>> gm1N = GM1N_model(Y,X,1) %按照累减还原序列
```

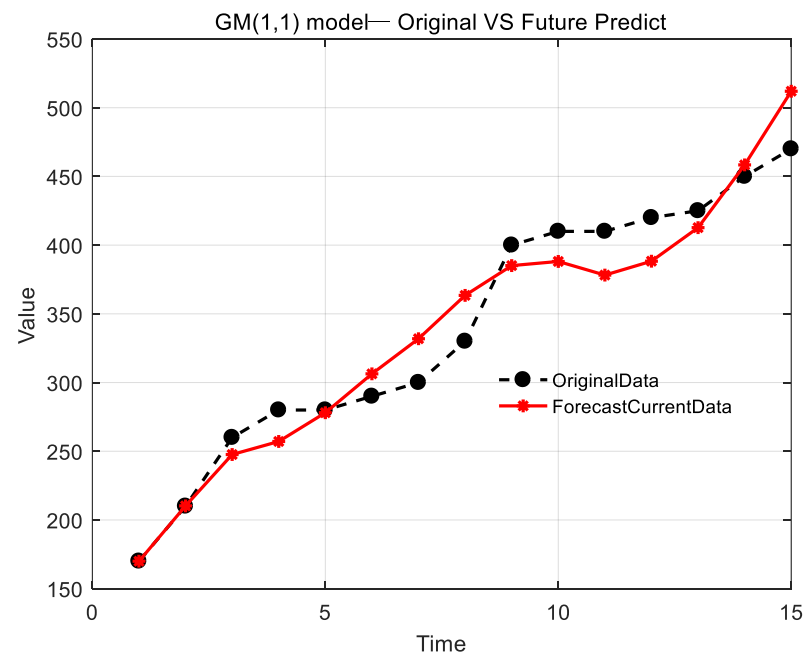
```
gm1N =  
包含以下字段的 struct:  
  
Coeff_a: 0.7079  
Coeff_b: [2×1 double]  
select: '累减还原形似'  
PredictCurr_Value: [15×1 double]  
AbsoluteError: [15×1 double]  
RelativeError: [15×1 double]  
RelativeErrorMean: 0.1098  
R: 0.5775  
C: 0.2788  
P: 1
```

3. GM(1,n)案例分析



可以看出，差分形式比累减形式精度高。
上图为累减形式图像，右图为差分形式图像。

```
gm1N =  
包含以下字段的 struct:  
  
Coeff_a: 0.6752  
Coeff_b: [2×1 double]  
PredictCurr_Value: [15×1 double]  
select: '差分模拟形式'  
AbsoluteError: [15×1 double]  
RelativeError: [15×1 double]  
RelativeErrorMean: 0.0521  
R: 0.5714  
C: 0.1371  
P: 1
```



三. GM(2,1)模型

GM(1,1)模型适用于具有较强指数规律的序列，只能描述单调的变化过程，对于非单调的摆动发展序列或有饱和的S形序列，可以考虑建立GM(2,1)，DGM和Verhulst模型。

GM(1,1)与GM(2,1)不论是建模方法和精度检验都差别不大，差别在于GM(2,1)是二阶常微分方程，GM(1,1)为一阶常微分方程。

原始序列 $x^{(0)}$ ，其1-AGO序列： $x^{(1)}$ ，1次累减生成序列1-IAGO序列：

$$\alpha^{(1)}x^{(0)} = (\alpha^{(1)}x^{(0)}(2), \dots, \alpha^{(1)}x^{(0)}(n)), \text{ 其中 } \alpha^{(1)}x^{(0)}(k) = x^{(0)}(k) - x^{(0)}(k-1), k = 2, 3, \dots, n$$

$x^{(1)}$ 的均值生成序列 $z^{(1)}$ ，则称： $\alpha^{(1)}x^{(0)}(k) + a_1x^{(0)}(k) + a_2z^{(1)}(k) = b$ 为GM(2,1)模型。

$$\text{模型的白化微分方程: } \frac{d^2x^{(1)}}{dt^2} + a_1 \frac{dx^{(1)}}{dt} + a_2x^{(1)} = b$$

1. GM(2,1)模型原理分析

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

用最小二乘法估计求解 $\hat{u} = (B^T B)^{-1} B^T Y$, 其中 $a_1 = \hat{u}(1), a_2 = \hat{u}(2), b = \hat{u}(3)$

利用边界条件, 求解GM(2,1)白化方程即可。

还原值: $\hat{x}^{(0)}(k+1) = \alpha^{(1)} \hat{x}^{(1)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$

1. GM(2,1)模型原理分析

定理: 关于GM(2,1)白化方程的解有以下结论:

1. 若 $X^{(1)}$ 是 $\frac{d^2 x^{(1)}}{dt^2} + a_1 \frac{dx^{(1)}}{dt} + a_2 x^{(1)} = b$ 的特解, $\bar{X}^{(1)}$ 是对应齐次方程 $\frac{d^2 x^{(1)}}{dt^2} + a_1 \frac{dx^{(1)}}{dt} + a_2 x^{(1)} = 0$

的特解, 则 $X^{(1)} + \bar{X}^{(1)}$ 是GM(2,1)白化方程的通解。

2. 齐次方程的通解有以下三种情况:

当特征方程 $r^2 + a_1 r + a_2 = 0$ 有两个不相等的实根时: $\bar{X}^{(1)} = C_1 e^{r_1 t} + C_2 e^{r_2 t}$

当特征方程有重根时: $\bar{X}^{(1)} = e^{rt} (C_1 + C_2 t)$

当特征方程有一对共轭复根 $r_1 = \alpha + i\beta$, $r_2 = \alpha - i\beta$ 时: $\bar{X}^{(1)} = e^{\alpha t} (C_1 \cos \beta t + C_2 \sin \beta t)$

3. 白化方程的特解有三种情况: 当零不是特征方程的根时, $X^{(1)*} = c$; 当零是特征方程的单根时, $X^{(1)*} = cx$; 当零是特征方程的重根时, $X^{(1)*} = cx^2$ 。

2. GM(2,1)模型MATLAB求解

3. GM(2,1)模型案例分析

例6：地基沉降随时间变化数据预测。

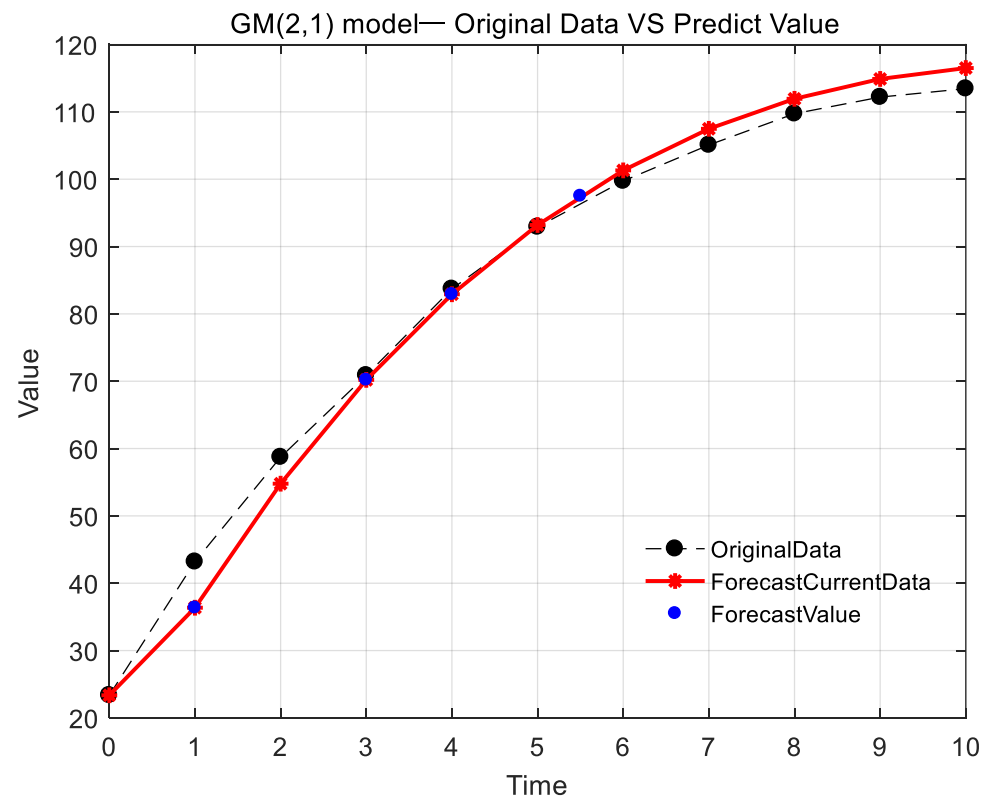
```
>> X = [23.3600 43.1900 58.7300 70.8700 83.7100 92.9100 99.7300 105.0800 109.7300 112.1900 113.4500];
```

```
>> gm21 = GM21_model(X,[2 4 5,6.5])
```

```
gm21 =
```

包含以下字段的 [struct](#):

```
    Coeff_a: [2×1 double]  
    Coeff_b: 25.8122  
Sol_Expression: [1×1 sym]  
    Predict_Data: [23.3600 36.3633 54.7816 70.2004 82.9142 93.1983 101.  
    Predict_Value: [36.3633 70.2004 82.9142 97.5099]  
    Abs_Error: [1.3038e-10 6.8267 3.9484 0.6696 0.7958 -0.2883 -1.57  
    Rel_Error: [5.5815e-12 0.1581 0.0672 0.0094 0.0095 0.0031 0.0158  
    RelErrorMean: 0.0325  
        R: 0.6864  
        C: 0.0647  
        P: 1
```



3. GM(2,1)模型案例分析

例6：森林火灾数据，来源于国家数据网，数据采集从2004年到2019年。

```
>> X = [2345 2478 3223 2034 2936 3703 3929 3966 5550 7723 8859 14144 9260 8170 11542 13466];
```

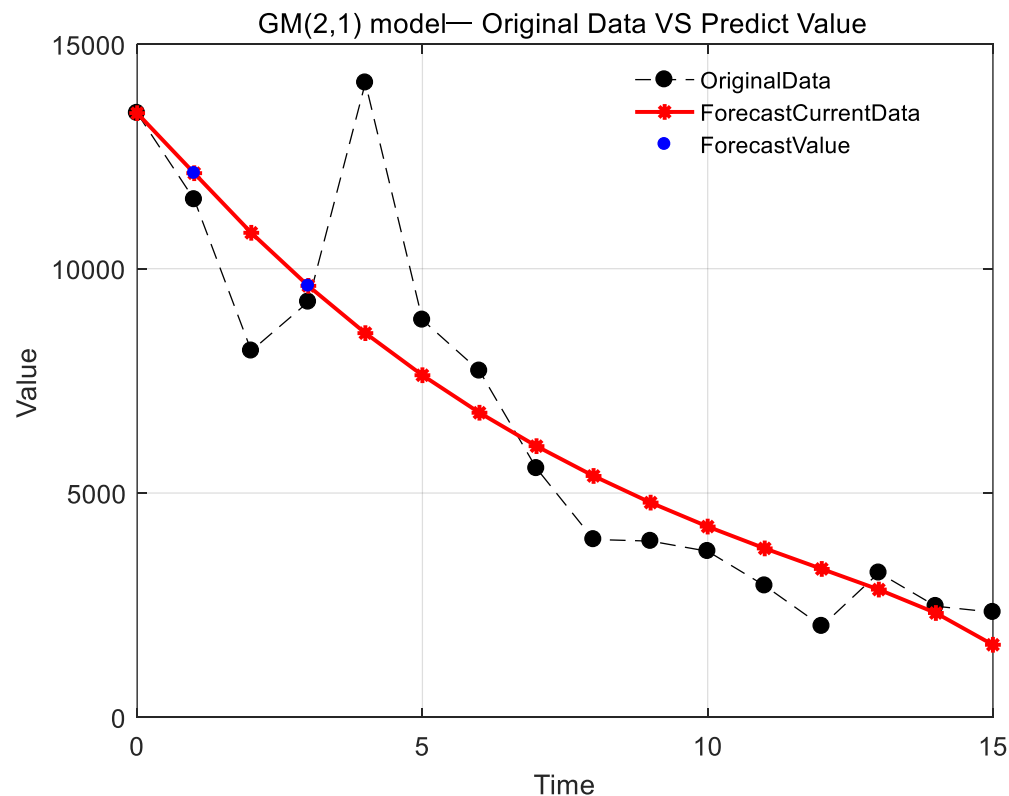
```
>> X = fliplr(X);
```

```
>> gm21 = GM21_model(X,[2 4])
```

```
gm21 =
```

包含以下字段的 [struct](#):

```
    Coeff_a: [2×1 double]  
    Coeff_b: -1.1092e+04  
Sol_Expression: [1×1 sym]  
    Predict_Data: [1×16 double]  
    Predict_Value: [1.2128e+04 9.6170e+03]  
    Abs_Error: [1×16 double]  
    Rel_Error: [1×16 double]  
    RelErrorMean: 0.2047  
         R: 0.7868  
         C: 0.3295  
         P: 0.9333
```



3. GM(2,1)模型案例分析

例7：居民消费水平数据，来源于国家数据网，数据采集从2000年到2018年。

```
>> X = [25378 23070 20877 18929 17271 15615 14075 12646 10550 9226 8483 7434 6302 5671 5056 4542 4256  
3954 3698]
```

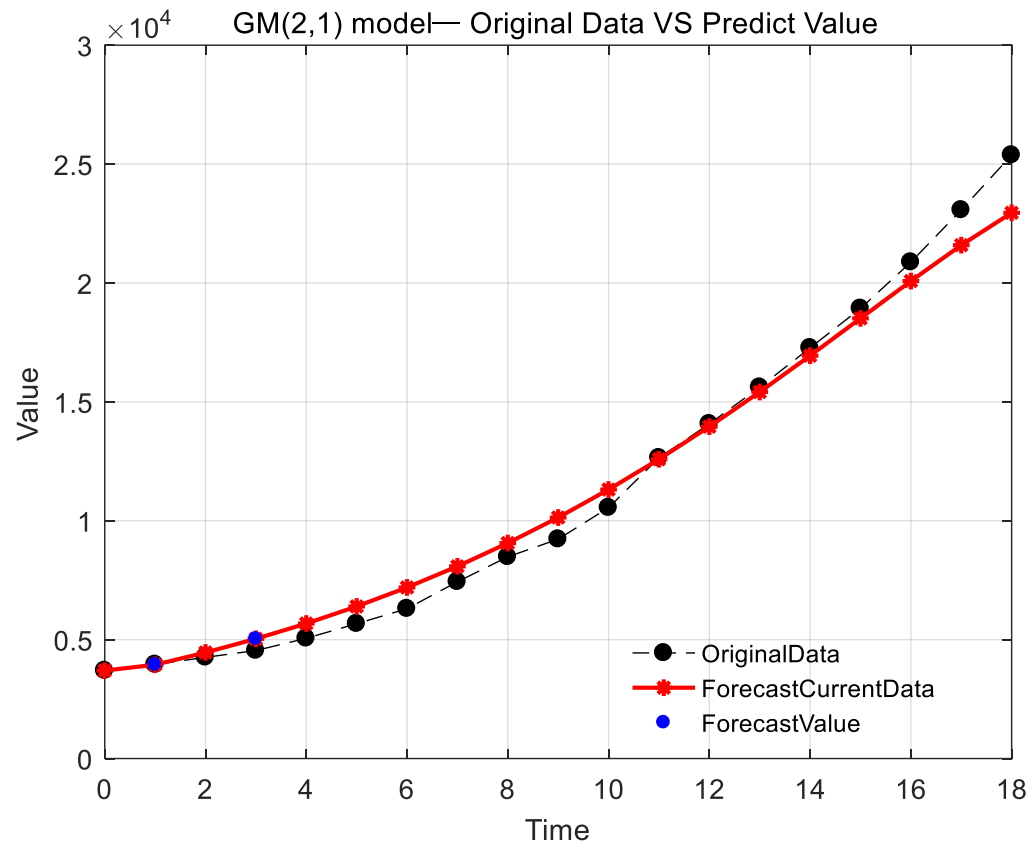
```
>> X = fliplr(X);
```

```
>> gm21 = GM21_model(X,[2 4])
```

```
gm21 =
```

包含以下字段的 [struct](#):

```
    Coeff_a: [2×1 double]  
    Coeff_b: -825.5245  
Sol_Expression: [1×1 sym]  
    Predict_Data: [1×19 double]  
    Predict_Value: [3.9455e+03 5.0340e+03]  
    Abs_Error: [1×19 double]  
    Rel_Error: [1×19 double]  
    RelErrorMean: 0.0601  
           R: 0.7039  
           C: 0.0829  
           P: 1
```



3. GM(2,1)模型案例分析



例7：农村居民消费水平指数(1978=100)数据，来源于国家数据网，数据采集从2000年到2018年。

```
>> X = [1554.5 1386.4 1258.4 1127.9 1024.0 923.4 841.3 782.1 701.7  
663.9 601.1 576.0 528.7 491.8 460.4 443.6 423.9 397.6 379.7]  
>> X = fliplr(X);  
>> gm21 = GM21_model(X,[2 4])
```

gm21 =

包含以下字段的 [struct](#):

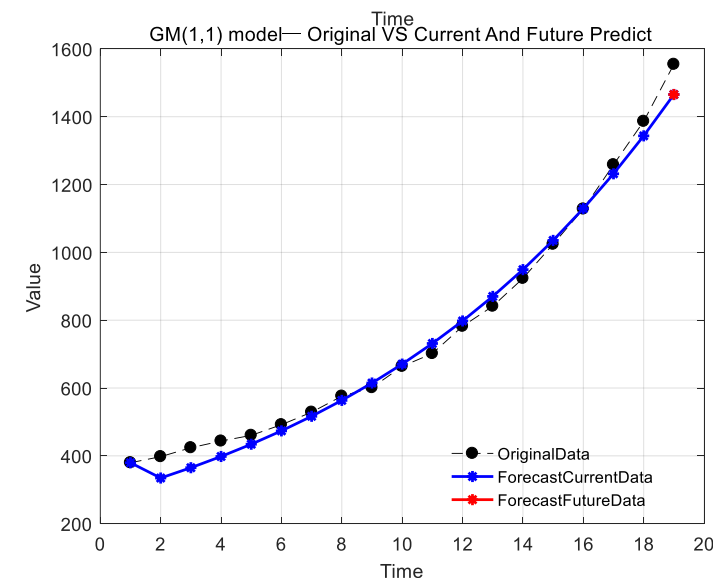
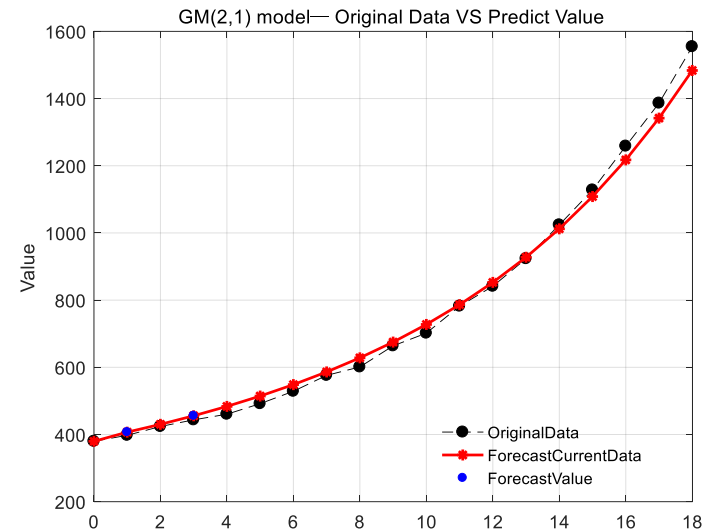
```
    Coeff_a: [2×1 double]  
    Coeff_b: -50.8798  
Sol_Expression: [1×1 sym]  
    Predict_Data: [1×19 double]  
    Predict_Value: [406.9874 455.6819]  
    Abs_Error: [1×19 double]  
    Rel_Error: [1×19 double]  
RelErrorMean: 0.0252  
        R: 0.7517  
        C: 0.0482  
        P: 1
```

```
>> gm11 = GM11_model(X,0)
```

gm11 =

包含以下字段的 [struct](#):

```
    Coeff_a: -0.0869  
    Coeff_u: 287.2902  
    Predict_Value: [1×19 double]  
    AbsoluteError: [1×19 double]  
    RelativeError: [1×19 double]  
RelativeErrorMean: 0.0430  
        R: 0.6684  
        C: 0.0638  
        P: 1
```



四. DGM(2,1)模型

原始序列 $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$, 其1-AGO序列: $x^{(1)}$, 1次累减生成序列1-IAGO序列 $\alpha^{(1)}x^{(0)}$ 。模型的灰化方程: $\alpha^{(1)}x^{(0)}(k) + ax^{(0)}(k) = b$ 为DGM(2,1)模型。

模型的白化微分方程: $\frac{d^2 x^{(1)}}{dt^2} + a \frac{dx^{(1)}}{dt} = b$

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

用最小二乘法估计求解 $\hat{u} = (B^T B)^{-1} B^T Y$

白化方程的解 (时间响应函数) 为 $\hat{x}^{(1)}(t) = \left(\frac{\hat{b}}{\hat{a}^2} - \frac{x^{(0)}(1)}{\hat{a}} \right) e^{-\hat{a}t} + \frac{\hat{b}}{\hat{a}} t + \frac{1+\hat{a}}{\hat{a}} x^{(0)}(1) - \frac{\hat{b}}{\hat{a}^2}$

DGM(2,1)模型的时间相应序列为 $\hat{x}^{(1)}(k+1) = \left(\frac{\hat{b}}{\hat{a}^2} - \frac{x^{(0)}(1)}{\hat{a}} \right) e^{-\hat{a}t} + \frac{\hat{b}}{\hat{a}} k + \frac{1+\hat{a}}{\hat{a}} x^{(0)}(1) - \frac{\hat{b}}{\hat{a}^2}$

还原值: $\hat{x}^{(0)}(k+1) = \alpha^{(1)}\hat{x}^{(1)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k)$

DGM(2,1)模型MATLAB求解

DGM(2,1)模型案例分析

例7：居民消费水平数据，来源于国家数据网，数据采集从2000年到2018年。

```
>> X = [25378 23070 20877 18929 17271 15615 14075 12646 10550 9226 8483 7434 6302 5671 5056 4542 4256  
3954 3698];
```

```
>> X = fliplr(X);
```

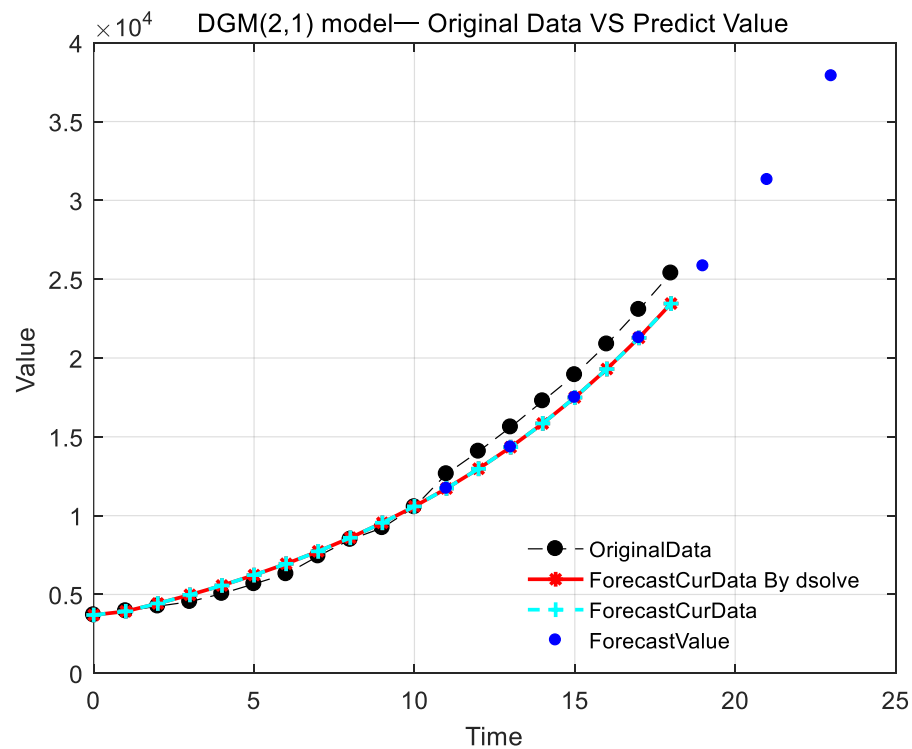
```
>> dgm21 = DGM21_model(X,[11:2:23])
```

```
dgm21 = DGM21_model(X, [11:2:23])
```

```
dgm21 =
```

包含以下字段的 [struct](#):

```
    Coeff_a: -0.0921  
    Coeff_b: 112.4363  
    Sol_Expression: [1×1 sym]  
    Predict_DataBy_dsolve: [1×19 double]  
    Predict_F: [1×19 double]  
    Predict_Value: [1.1725e+04 1.4345e+04 1.7494e+04 2.1281e+04  
    Abs_Error: [1×19 double]  
    Rel_Error: [1×19 double]  
    RelErrorMean: 0.0608  
    R: 0.6277  
    C: 0.0917  
    P: 1
```



DGM(2,1)模型案例分析

```
>> X = [2.874,3.278,3.39,3.679,3.77,3.8]; %原始数据序列
```

```
>> dgm21 = DGM21_model(X,[1:2:11])
```

```
dgm21 =
```

包含以下字段的 [struct](#):

```
    Coeff_a: 0.4240
```

```
    Coeff_b: 1.7046
```

```
    Sol_Expression: [1×1 sym]
```

```
    Predict_DataBy_dsolve: [2.8740 3.0860 3.4088 3.6201 3.7584 3.8488]
```

```
    Predict_F: [2.8740 3.0860 3.4088 3.6201 3.7584 3.8488]
```

```
    Predict_Value: [3.0860 3.6201 3.8488 3.9468 3.9888 4.0067]
```

```
    Abs_Error: [-2.7089e-14 0.1920 -0.0188 0.0589 0.0116 -0.0488]
```

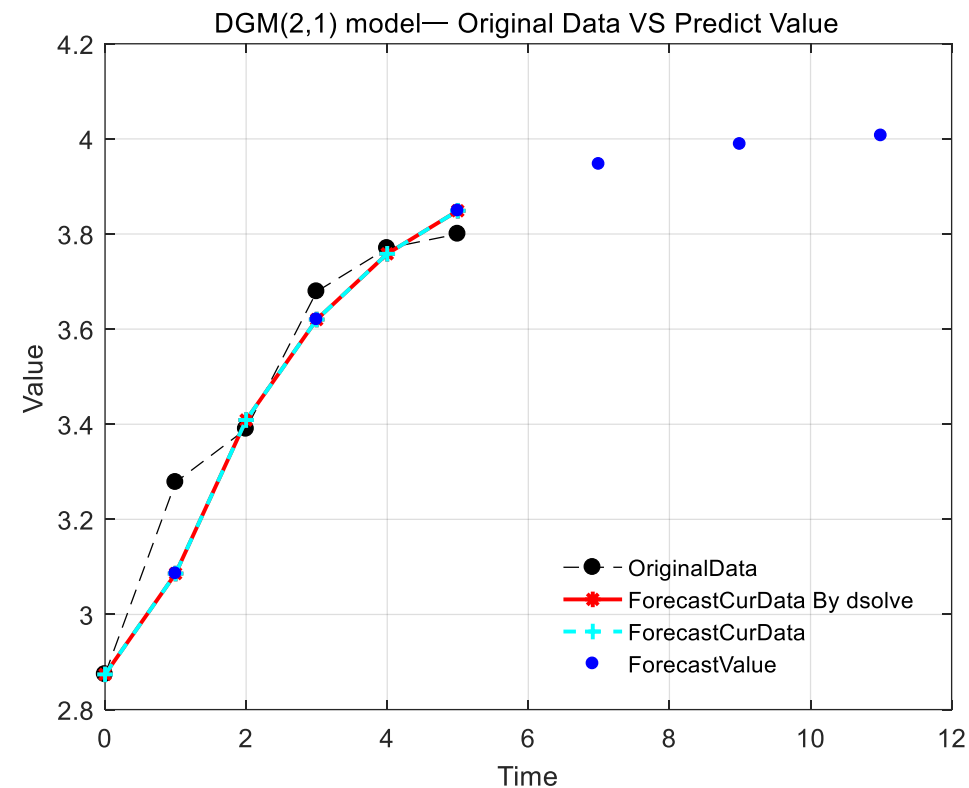
```
    Rel_Error: [9.4257e-15 0.0586 0.0056 0.0160 0.0031 0.0129]
```

```
    RelErrorMean: 0.0160
```

```
        R: 0.7498
```

```
        C: 0.1976
```

```
        P: 1
```



五. 灰色Verhulst模型

灰色Verhulst模型(即Logistic模型)

- GM(1,1)模型适用具有较强指数规律的序列，只能描述单调的变化过程。
- 但是道路交通系统是一个动态的时变系统，道路交通事故作为道路系统的行为特征量，具有一定的随机波动性，它的发展呈现某种变化趋势的非平稳随机过程，因此可建立交通事故灰色马尔可夫预测模型，以提高预测精度。但灰色马尔可夫预测模型的应用难点是如何进行状态划分，故对于非单调的摆动发展序列或具有饱和状态的S形序列，Verhulst 模型，GM(2,1)模型等更适用。
- Verhulst 模型主要用来描述具有饱和状态的过程，即S形过程，常用于人口预测、生物生长、繁殖预测及产品经济寿命预测等。近年来中国道路交通事故表现为具有饱和状态的S形过程，故可采用 Verhulst 模型。

1. 灰色Verhulst模型原理

原始序列 $x^{(0)}$ ，其1-AGO序列： $x^{(1)}$ ， $x^{(1)}$ 的均值生成序列 $z^{(1)}$

则灰色Verhulst模型为： $x^{(0)} + az^{(1)} = b\left(z^{(1)}\right)^2$ ，灰色Verhulst模型的白化方程为： $\frac{dx^{(1)}}{dt} + ax^{(1)} = b\left(x^{(1)}\right)^2$

$$\text{若 } u = [a, b]^T \text{ 为参数列, 且 } B = \begin{bmatrix} -z^{(1)}(2) & \left(z^{(1)}(2)\right)^2 \\ -z^{(1)}(3) & \left(z^{(1)}(3)\right)^2 \\ \vdots & \vdots \\ -z^{(1)}(n) & \left(z^{(1)}(n)\right)^2 \end{bmatrix}, Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$$

则参数列 u 的最小二乘法估计求解 $\hat{u} = (B^T B)^{-1} B^T Y$

灰色Verhulst白化方程的解（时间响应函数）为 $x^{(1)}(t) = \frac{\hat{a}x^{(0)}(1)}{\hat{b}x^{(0)}(1) + [\hat{a} - \hat{b}x^{(0)}(1)]e^{\hat{a}t}}$

灰色Verhulst模型的时间相应序列为 $x^{(1)}(k+1) = \frac{\hat{a}x^{(0)}(1)}{\hat{b}x^{(0)}(1) + [\hat{a} - \hat{b}x^{(0)}(1)]e^{\hat{a}k}}$

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

2. 灰色Verhulst模型MATALB求解



```
GM_Verhulst.m  x  +
1  function gmvh = GM_Verhulst(X,pre_num)
2  % GM_Verhulst用于求解灰色Verhulst模型，并用于预测
3  % 输入参数X是原数据，pre_num是未来预期
4  % 输出参数gmvh为一个结构体，包括求解过程结果和模型精度检验结果
5  %% 数据的预处理及求解
6  n = length(X);
7  H = diff(X); %对原始数列X做累减
8  Z = (X(2:end) + X(1:end-1))/2; %对原始数列X做紧邻均值生成
9  B = [-Z; Z.^2]; %构造数据矩阵
10 Yn = H';
11 u = (B*B')\B*Yn; %使用最小二乘法计算
12 a = u(1); %发展系数
13 b = u(2); %灰作用量
14
15 %% 得到预测出的数据
16 F = [X(1), (a*X(1))./(b*X(1)+(a - b*X(1))*exp(a*(1:n+pre_num-1)))];
17 Fval = F(1:n);
18 preal = F(n+1:end);
19
20 %% 可视化
21 t = 0:n-1;
22 plot(t, X, 'c-o', 'MarkerFaceColor','c');
23 hold on;grid on;
24 plot(t, Fval,'r*-','LineWidth',1.5);
25 plot(n:n+pre_num-1,preal,'b*-','LineWidth',1.5)
26 title('GM_Verhulst model— Original Data VS Predict Value')
27 legend('OriginalData','ForecastCurData','ForecastValue',...
28 'Location','best')
29 legend('boxoff')
30 xlabel('Time');ylabel('Value'); hold off
31
32 %% 后验差检验
33 e = abs(X - Fval); %真实值与预测值误差
34 q = e./X; %相对误差
35 s1 = std(X); %原数据标准方差
36 s2 = std(e); %误差标准方差
37 c = s2/s1; %方差比
38 p = 0; %小误差概率
39 for i = 2:n
40     if(abs(e(i) - mean(e))<0.6745*s1)
41         p = p + 1;
42     end
43 end
44 p = p/(n-1); %小误差概率
45 R_k = (min(e(2:end))+0.5*max(e))./(e(2:end)+0.5*max(e)); % rho = 0.5
46 R = sum(R_k)/length(R_k); %关联度
47
48 %% 计算过程变量组合输出结构体变量
49 gmvh.Coeff_a = a;
50 gmvh.Coeff_b = b;
51 gmvh.Predict_Data = Fval;
52 gmvh.Predict_Value = preal;
53 gmvh.Abs_Error = X - Fval(1:n);
54 gmvh.Rel_Error = q;
55 gmvh.RelErrorMean = mean(q);
56 gmvh.R = R;
57 gmvh.C = c;
58 gmvh.P = p;
59 end
```

3. 灰色Verhulst模型案例分析

例8: 将一定量的大肠杆菌菌种接种在液体培养基中，在一定条件下进行培养，观察其生长繁殖规律。细菌悬液的浓度与混浊度成正比，故可用分光亮度计测定细菌悬液的光密度来推知菌液的浓度。每隔5h记录OD600的值，得到下表。请你预测大肠杆菌的数量。

时间点均匀采样/5h	0	1	2	3	4	5	6	7	8	9	10
细菌培养液吸光度/OD600	0.025	0.023	0.029	0.044	0.084	0.164	0.332	0.521	0.97	1.6	2.45
时间点均匀采样/5h	11	12	13	14	15	16	17	18	19	20	
细菌培养液吸光度/OD600	3.11	3.57	3.76	3.96	4	4.46	4.4	4.49	4.76	5.01	

```
>> X = [0.025, 0.023, 0.029, 0.044, 0.084, 0.164, 0.332, 0.521, 0.97, 1.6, 2.45, 3.11, 3.57, 3.76, 3.96, 4, 4.46, 4.4, 4.49, 4.76, 5.01];
>> gmvh = GM_Verhulst(X,5)
```

3. 灰色Verhulst模型案例分析

```
>> gmvh = GM_Verhulst(X, 5)
```

```
gmvh =
```

包含以下字段的 [struct](#):

```
    Coeff_a: -0.5131
```

```
    Coeff_b: -0.1074
```

```
    Predict_Data: [1×21 double]
```

```
    Predict_Value: [4.7571 4.7647 4.7692 4.7720 4.7736]
```

```
    Abs_Error: [1×21 double]
```

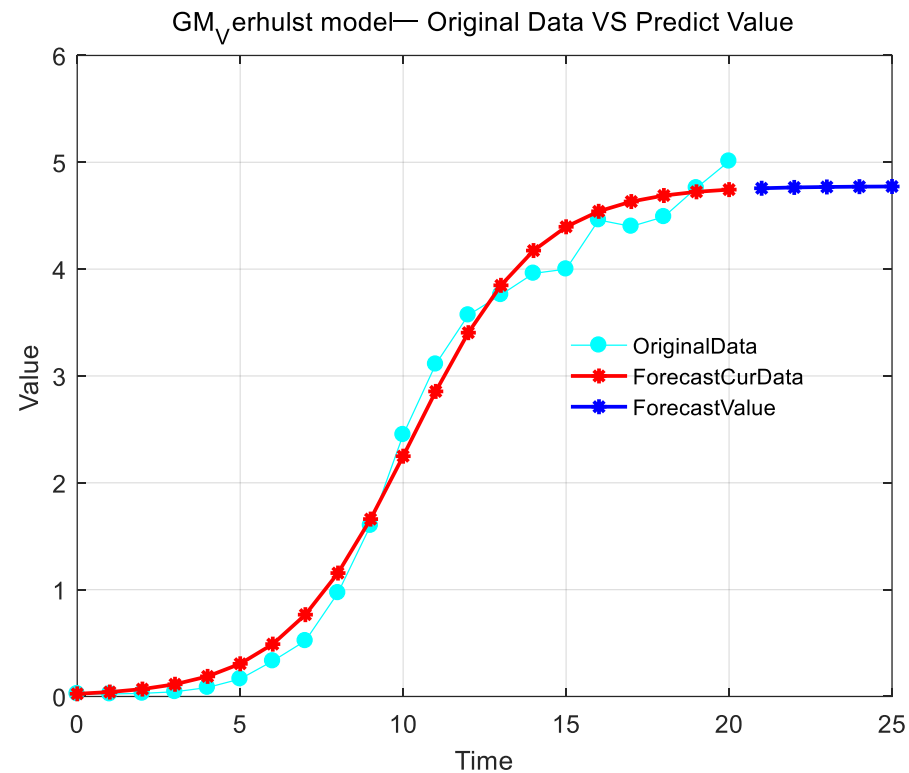
```
    Rel_Error: [1×21 double]
```

```
    RelErrorMean: 0.3634
```

```
        R: 0.6529
```

```
        C: 0.0512
```

```
        P: 1
```





感谢聆听
