

近似非齐次无偏GM(1,1)模型的递推解法及应用

江艺羨, 张岐山

(福州大学 经济与管理学院, 福州 350108)

摘要: 针对传统近似非齐次灰建模可能出现参数复数解的问题, 提出无偏灰色GM(1,1)模型的递推解法, 从而减少由差分方程向微分方程跳跃而导致误差的问题. 给出不同初始条件下非齐次无偏GM(1,1)模型的递推预测公式, 并在此基础上, 将递推公式运用于时间序列分段, 提出基于近似非齐次无偏GM(1,1)模型的时间序列分段表示方法. 实例结果表明, 所提出的递推模型能够获得较高的拟合精度, 分析结果验证了基于灰色预测模型在时间序列分段表示中的有效性和实用性.

关键词: 灰色GM(1,1)模型; 非齐次指数; 递推解法; 时间序列; 分段表示

中图分类号: TP273

文献标志码: A

Recursive solution to approximate non-homogeneous unbiased GM(1,1) model and its application

JIANG Yi-xian, ZHANG Qi-shan

(School of Economics and Management, Fuzhou University, Fuzhou 350108, China. Correspondent: JIANG Yi-xian, E-mail: 371466977@qq.com)

Abstract: As there may be complex solutions in the traditional approximate non-homogeneous GM(1,1) model, this paper proposes the recursive solution to unbiased GM(1,1) model, which can reduce the errors from the differential equation to differential equation. And its predictive formulas are given under different initial conditions with the recursive method. On this basis, by applying the recursive formulas to time series piecewise representation, the method of time series piecewise representation based on the new model is proposed. The results show that the recursive model has higher fitting precision, and also verify the effectiveness and the practicability of the representation method of time series based on the grey forecasting model.

Keywords: grey GM(1,1) model; non-homogeneous exponential; recursive solution; time series; piecewise representation

0 引言

灰色系统理论是我国著名学者邓聚龙教授于1982年所创立的, 是一种研究少数据、贫信息不确定性问题的新方法^[1]. 为了提高预测精度, 许多学者和实际工作者都从理论和应用上进行了广泛深入的研究, 并成功地解决了生产、生活和科学研究中的大量实际问题. 现实生活中, 具有非齐次指数特征的数据明显多于齐次指数特征的数据, 因此对于非齐次指数特征数据的研究更具实际意义.

目前已有大量研究针对近似非齐次指数律序列的模拟与预测, 并取得了一定的成果^[2-7]. 在实际应用中, 由最小二乘估计方法求得的近似非齐次GM(1,1)模型的参数 a 可能会存在复数解, 从而使得

实验计算复杂, 并可能影响模拟精度. 为了避免这种误差, 本文基于近似非齐次无偏GM(1,1)模型的统一表达式, 提出近似非齐次无偏GM(1,1)模型的递推解法, 并在此基础上, 研究在不同初始条件下模型的递推预测公式.

时间序列一般是海量、高维数据, 在原始数据上进行数据挖掘及分析的效率较低, 因此在对数据进行分析之前, 需要对时间序列进行降维^[8]. 数据压缩及噪声过滤大多使用的方法是数据序列的线性分段表示方法^[9-11]. 文献[12]通过传统GM(1,1)模型验证了灰色预测模型在时间序列分段表示方法上的优越性, 本文在此基础上, 提出基于近似无偏GM(1,1)模型的时间序列分段表示方法.

收稿日期: 2014-12-09; **修回日期:** 2015-02-28.

基金项目: 福建省软科学项目(2013R0057).

作者简介: 江艺羨(1983-), 女, 博士生, 从事灰色系统、数据挖掘的研究; 张岐山(1962-), 男, 教授, 博士生导师, 从事灰色系统、物流管理等研究.

1 无偏GM(1,1)模型的递推解法

定义1 设 $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ 为非负序列, 称 $X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n))$ 为 $X^{(0)}$ 的一阶累加生成算子(1-AGO)序列, 其中

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

定理1 若 $x^{(1)}(t)$ 满足微分方程 $\frac{dx^{(1)}}{dt} + ax^{(1)} = vt + w$, 即

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

的充分必要条件是

$$x^{(1)}(k) = e^{-a}x^{(1)}(k-1) + \frac{e^a - 1}{ae^a}vk + w\frac{e^a - 1}{ae^a}, \quad k = 2, 3, \dots, n. \quad (1)$$

证明 1) 必要性. 将 $x^{(1)}(k-1)$ 代入 $e^{-a}x^{(1)}(k-1) + \frac{e^a - 1}{ae^a}vk + w\frac{e^a - 1}{ae^a}$, 可得

$$\begin{aligned} & e^{-a}x^{(1)}(k-1) + \frac{e^a - 1}{ae^a}vk + w\frac{e^a - 1}{ae^a} = \\ & e^{-a} \left\{ \left[x^{(1)}(1) + \frac{v}{a(e^a - 1)} - \frac{w}{a} - \frac{v}{a} \right] e^{-a(k-1-1)} + \frac{v}{a}(k-1) + \left[\frac{w}{a} - \frac{1}{a(e^a - 1)}v \right] \right\} + \frac{e^a - 1}{ae^a}vk + w\frac{e^a - 1}{ae^a} = \\ & \left[x^{(1)}(1) + \frac{v}{a(e^a - 1)} - \frac{w}{a} - \frac{v}{a} \right] e^{-a(k-1)} + \frac{v}{a}k + \left[\frac{w}{a} - \frac{1}{a(e^a - 1)}v \right] = x^{(1)}(k). \end{aligned}$$

可证等式

$$x^{(1)}(k) = e^{-a}x^{(1)}(k-1) + \frac{e^a - 1}{ae^a}vk + w\frac{e^a - 1}{ae^a}$$

成立. 必要性得证.

2) 充分性. 因为

$$x^{(1)}(k) = e^{-a}x^{(1)}(k-1) + \frac{e^a - 1}{ae^a}vk + w\frac{e^a - 1}{ae^a},$$

即, ① 当 $k = 2$ 时

$$\begin{aligned} & x^{(1)}(2) = \\ & e^{-a}x^{(1)}(1) + 2\frac{e^a - 1}{ae^a}v + w\frac{e^a - 1}{ae^a} = \\ & x^{(1)}(1)e^{-a} + \left[\frac{v}{a(e^a - 1)} - \frac{w}{a} - \frac{v}{a} \right] e^{-a} - \\ & \frac{v}{a(e^a - 1)} + \frac{2v}{a} + \frac{w}{a} - \frac{v}{a}e^{-a} = \\ & \left[x^{(1)}(1) + \frac{v}{a(e^a - 1)} - \frac{w}{a} - \frac{v}{a} \right] e^{-a} + \\ & \frac{2v}{a} + \left[\frac{w}{a} - \frac{v}{a(e^a - 1)} \right]. \end{aligned}$$

显然等式成立.

② 当 $k = t$ 时, 等式成立, 即

$$\begin{aligned} & x^{(1)}(t) = \\ & \left[x^{(1)}(1) + \frac{v}{a(e^a - 1)} - \frac{w}{a} - \frac{v}{a} \right] e^{-a(t-1)} + \\ & \frac{vt}{a} + \left[\frac{w}{a} - \frac{v}{a(e^a - 1)} \right], \end{aligned}$$

则, 当 $k = t + 1$ 时

$$\begin{aligned} & x^{(1)}(t+1) = \\ & e^{-a}x^{(1)}(t) + \frac{e^a - 1}{ae^a}v(t+1) + \frac{e^a - 1}{ae^a}w = \\ & e^{-a} \left[x^{(1)}(1) + \frac{v}{a(e^a - 1)} - \frac{w}{a} - \frac{v}{a} \right] e^{-a(t-1)} + \\ & \frac{vt}{a}e^{-a} + \left[\frac{w}{a} - \frac{v}{a(e^a - 1)} \right] e^{-a} + \\ & \frac{e^a - 1}{ae^a}v(t+1) + \frac{e^a - 1}{ae^a}w = \\ & \left[x^{(1)}(1) + \frac{v}{a(e^a - 1)} - \frac{w}{a} - \frac{v}{a} \right] e^{-a(t+1-1)} + \frac{vt}{ae^a} + \\ & \frac{e^a - 1}{ae^a}v(t+1) - \frac{v}{ae^a(e^a - 1)} + \frac{w}{ae^a} + \frac{e^a - 1}{ae^a}w = \\ & \left[x^{(1)}(1) + \frac{v}{a(e^a - 1)} - \frac{w}{a} - \frac{v}{a} \right] e^{-a(t+1-1)} + \\ & \frac{v}{a}(t+1) + \left[\frac{w}{a} - \frac{v}{a(e^a - 1)} \right] \end{aligned}$$

等式成立.

由①和②可知, 对于任意 $k \geq 2$, 等式

$$\begin{aligned} & x^{(1)}(k) = \\ & \left[x^{(1)}(1) + \frac{v}{a(e^a - 1)} - \frac{w}{a} - \frac{v}{a} \right] e^{-a(k-1)} + \frac{v}{a}k + \left[\frac{w}{a} - \frac{v}{a(e^a - 1)} \right] \end{aligned}$$

均成立. 充分性得证. \square

定义2 $X^{(0)}$ 和 $X^{(1)}$ 如定义1所述, 若参数列 $\hat{u} = (u_1, u_2, u_3)^T$ 满足式(2), 则称无偏灰色模型的最小二乘估计参数列满足

$$\hat{u} = (B^T B)^{-1} B^T Y.$$

其中

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

定理2 设 Y 、 B 和 \hat{u} 如定义2所述, $\hat{u} = (B^T B)^{-1} B^T Y$, 取初始值 $x^{(1)}(1) = x^{(0)}(1)$, 则无偏估计模型的解为

$$\begin{aligned} & x^{(1)}(k) = \\ & \begin{cases} \left[x^{(1)}(1) + \frac{u_1 u_2}{(1-u_1)^2} + \frac{u_2}{u_1-1} + \frac{u_3}{u_1-1} \right] u_1^{k-1} - \\ \frac{u_2}{u_1-1}k - \frac{u_1 u_2}{(1-u_1)^2} + \frac{u_3}{1-u_1}, \quad u_1 \neq 1; \\ x^{(1)}(1) + \frac{(k-1)(k+2)}{2}u_2 + u_3k, \quad u_1 = 1; \end{cases} \quad (3) \end{aligned}$$

$$x^{(0)}(k) = \begin{cases} \left[(u_1 - 1)x^{(1)}(1) + \frac{u_1 u_2}{u_1 - 1} + \right. \\ \left. u_2 + u_3 \right] u_1^{k-2} - \frac{u_2}{u_1 - 1}, u_1 \neq 1; \\ u_2 k + u_3, u_1 = 1. \end{cases} \quad (4)$$

其中 $k = 2, 3, \dots, n$.

证明 由 $x^{(1)}(k) = u_1 x^{(1)}(k-1) + u_2 k + u_3$ 可知

$$\begin{aligned} x^{(1)}(2) &= u_1 x^{(1)}(1) + 2u_2 + u_3, \\ x^{(1)}(3) &= u_1 x^{(1)}(2) + 3u_2 + u_3, \\ &\vdots \\ x^{(1)}(n) &= u_1 x^{(1)}(n-1) + nu_2 + u_3. \end{aligned}$$

将 $x^{(1)}(t-1)$ 依次代入 $x^{(1)}(t), t = 2, 3, \dots, k$, 可得

$$\begin{aligned} x^{(1)}(k) &= \\ u_1 x^{(1)}(k-1) + u_2 k + u_3 &= \\ u_1 [u_1 x^{(1)}(k-2) + u_2(k-1) + u_3] + u_2 k + u_3 &= \\ u_1^2 x^{(1)}(k-2) + [u_1(k-1) + k]u_2 + (u_1 + 1)u_3 &= \\ u_1^2 [u_1 x^{(1)}(k-3) + u_2(k-2) + u_3] + & \\ [u_1(k-1) + k]u_2 + (u_1 + 1)u_3 &= \\ u_1^2 x^{(1)}(k-3) + [u_1^2(k-2) + u_1(k-1) + & \\ k]u_2 + (u_1^2 + u_1 + 1)u_3 &= \\ &\vdots \\ u_1^{k-1} x^{(1)}(1) + \frac{u_1^{k-1} - k}{u_1 - 1} u_2 - (1 - u_1^{k-1}) \frac{u_1 u_2}{(1 - u_1)^2} + & \\ (1 - u_1^{k-1}) \frac{u_3}{1 - u_1}. & \end{aligned}$$

当 $u_1 \neq 1$ 时

$$\begin{aligned} x^{(1)}(k) &= \\ \left[x^{(1)}(1) + \frac{u_1 u_2}{(1 - u_1)^2} + \frac{u_2}{u_1 - 1} + \frac{u_3}{u_1 - 1} \right] u_1^{k-1} - & \\ \frac{u_2}{u_1 - 1} k - \frac{u_1 u_2}{(1 - u_1)^2} + \frac{u_3}{1 - u_1}; & \end{aligned}$$

当 $u_1 = 1$ 时

$$x^{(1)}(k) = x^{(1)}(1) + \frac{(k-1)(k+2)}{2} u_2 + u_3 k.$$

取初始值 $x^{(1)}(1) = x^{(0)}(1)$, 可得 $x^{(1)}(k)$ 的预测公式为

$$\begin{aligned} x^{(1)}(k) &= \\ \begin{cases} \left[x^{(1)}(1) + \frac{u_1 u_2}{(1 - u_1)^2} + \frac{u_2}{u_1 - 1} + \frac{u_3}{u_1 - 1} \right] u_1^{k-1} - \\ \frac{u_2}{u_1 - 1} k - \frac{u_1 u_2}{(1 - u_1)^2} + \frac{u_3}{1 - u_1}, u_1 \neq 1; \\ x^{(1)}(1) + \frac{(k-1)(k+2)}{2} u_2 + u_3 k, u_1 = 1; \end{cases} \\ k = 2, 3, \dots, n. \end{aligned}$$

利用 $x^{(1)}(k) - x^{(1)}(k-1)$ 可求得 $x^{(0)}(k)$ 的预测公式如下:

当 $u_1 \neq 1$ 时

$$\begin{aligned} x^{(0)}(k) &= x^{(1)}(k) - x^{(1)}(k-1) = \\ \left[x^{(1)}(1) + \frac{u_1 u_2}{(1 - u_1)^2} + \frac{u_2}{u_1 - 1} + \frac{u_3}{u_1 - 1} \right] u_1^{k-1} - & \\ \frac{u_2}{u_1 - 1} k - \frac{u_1 u_2}{(1 - u_1)^2} + \frac{u_3}{1 - u_1} - \left[x^{(1)}(1) + \right. & \\ \left. \frac{u_1 u_2}{(1 - u_1)^2} + \frac{u_2}{u_1 - 1} + \frac{u_3}{u_1 - 1} \right] u_1^{k-1-1} + & \\ \frac{u_2}{u_1 - 1} (k-1) + \frac{u_1 u_2}{(1 - u_1)^2} - \frac{u_3}{1 - u_1} = & \\ \left[(u_1 - 1)x^{(1)}(1) + \frac{u_1 u_2}{u_1 - 1} + \right. & \\ \left. u_2 + u_3 \right] u_1^{k-2} - \frac{u_2}{u_1 - 1}; & \end{aligned}$$

当 $u_1 = 1$ 时

$$\begin{aligned} x^{(0)}(k) &= \\ x^{(1)}(1) + \frac{(k-1)(k+2)}{2} u_2 + u_3 k - & \\ x^{(1)}(1) - \frac{(k-2)(k+1)}{2} u_2 - u_3 k + u_3 &= \\ u_2 k + u_3. & \end{aligned}$$

取初始值 $x^{(1)}(1) = x^{(0)}(1)$, 可得 $x^{(0)}(k)$ 的预测公式为

$$\begin{aligned} x^{(0)}(k) &= \\ \begin{cases} \left[(u_1 - 1)x^{(1)}(1) + \frac{u_1 u_2}{u_1 - 1} + u_2 + \right. \\ \left. u_3 \right] u_1^{k-2} - \frac{u_2}{u_1 - 1}, u_1 \neq 1; \\ u_2 k + u_3, u_1 = 1; \end{cases} \\ k = 2, 3, \dots, n. \quad \square \end{aligned}$$

根据经典 GM(1,1) 模型的研究成果, 模拟预测函数并不一定经过第 1 个值, 因此为完善近似非齐次无偏 GM(1,1) 模型, 根据新息原理, 提出以累加序列终点作为模型的初始条件, 构建另一种无偏 GM(1,1) 模型的递推求解方法.

定理 3 设 Y, B 和 \hat{u} 如定义 2 所述, $\hat{u} = (B^T B)^{-1} B^T Y$, 取初始值 $x^{(1)}(1) = x^{(1)}(n)$, 则无偏估计模型的解为

$$\begin{aligned} x^{(1)}(k) &= \\ \begin{cases} x^{(1)}(n) - \frac{(n-k)(n+k+1)}{2} u_2 - \\ u_3(n-k), u_1 = 1; \\ \left[x^{(1)}(n) + \frac{u_2 n}{u_1 - 1} + \frac{u_1 u_2}{(u_1 - 1)^2} + \right. \\ \left. \frac{u_3}{u_1 - 1} \right] (u_1)^{k-n} - \frac{u_2}{u_1 - 1} k - \frac{u_1 u_2}{(u_1 - 1)^2} - \\ \frac{u_3}{u_1 - 1}, u_1 \neq 1; \end{cases} \quad (5) \end{aligned}$$

$$x^{(0)}(k) = \begin{cases} \left[(u_1 - 1)x^{(1)}(n) + u_2n + \frac{u_1u_2}{u_1 - 1} + \right. \\ \left. u_3 \right] (u_1)^{k-1-n} - \frac{u_2}{u_1 - 1}, u_1 \neq 1; \\ u_2k + u_3, u_1 = 1. \end{cases} \quad (6)$$

其中 $k = 2, 3, \dots, n$.

2 基于近似无偏 GM(1,1) 模型的时间序列分段表示方法

文献 [12] 已经验证了灰色预测模型在时间序列分段方法上的优越性, 本文在此基础上, 采用非齐次无偏 GM(1,1) 模型代替文献中的传统 GM(1,1) 模型, 以期得到更高的拟合精度.

定义 3 基于无偏 GM(1,1) 模型分段表示 UGMPR_rs. 设有时间序列 $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$, 在拟合误差阈值 ε 下, 它的分段函数表示为

$$\hat{x}^{(0)}(k) = \begin{cases} \left[(u_{11} - 1)x^{(1)}(1) + \frac{u_{11}u_{12}}{u_{11} - 1} + u_{12} + \right. \\ \left. u_{13} \right] u_{11}^{k-2} - \frac{u_{12}}{u_{11} - 1}, k = 2, 3, \dots, t_1; \\ \left[(u_{21} - 1)x^{(1)}(t_1) + \frac{u_{21}u_{22}}{u_{21} - 1} + u_{22} + \right. \\ \left. u_{23} \right] u_{21}^{k-1-t_1} - \frac{u_{22}}{u_{21} - 1}, k = t_1 + 1, t_1 + 2, \dots, t_2; \\ \vdots \\ \left[(u_{m1} - 1)x^{(1)}(t_{m-1}) + \frac{u_{m1}u_{m2}}{u_{m1} - 1} + \right. \\ \left. u_{m2} + u_{m3} \right] u_{m1}^{k-1-t_{m-1}} - \frac{u_{m2}}{u_{m1} - 1}, \\ k = t_{m-1} + 1, t_{m-1} + 2, \dots, t_m. \end{cases} \quad (7)$$

其中: 初始值为 $x^{(1)}(1) = x^{(0)}(1)$, $[u_{i1}, u_{i2}, u_{i3}]^T$ 为第 i 个子序列通过无偏 GM(1,1) 模型所得到的参数列, $t_{i-1} + 1$ 和 t_i 分别为第 i 个子序列第 1 个和最后 1 个数据在原序列的位置. 式 (7) 中, 当 $u_{i1} = 1$ 时, 将对应的分段表示函数改为 $\hat{x}^{(0)}(k) = u_{i2}(k - t_{i-1}) + u_{i3}$, $k = t_{i-1} + 1, t_{i-1} + 2, \dots, t_i$. 若初始值取 $x^{(1)}(1) = x^{(1)}(n)$, 则将式 (7) 中的等式改成对应的式 (5).

2.1 时间序列的 UGMPR 表示算法

基于非齐次无偏 GM(1,1) 模型的时间序列分段表示方法中, 拟合函数初始值取 $x^{(1)}(1) = x^{(0)}(1)$. 算法步骤如下.

算法输入: 时间序列 $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$, 拟合误差阈值 ε_0 ;

算法输出: 时间序列的 UGMPR 表示.

Step 1: 令分段队列 $queue = [1, 1, 0, 1, 0, 0, 0]$ 分别表示 [子序列起始点, 子序列终点, 子序列拟合误差, 模型输出参数 u (用于判断 e^a 与 0 的关系), 模型参

数 a , 模型参数 b , 模型参数 c (若采用 GM(1,1) 模型, 则 $c = 0$); $num = 1$ (原始序列位置).

Step 2: 如果 num 不大于原始序列长度 n , 则令当前 $queue$ 最后一行中子序列终点为新子序列的起点. $\varepsilon_{new} = 0$, 执行 **Step 3**; 否则执行 **Step 5**.

Step 3: while $\varepsilon_{new} \leq \varepsilon_0$, $num = num + 1$, 判断 num , 若不大于序列长度, 则构造新子序列, 采用拟合函数 (式 (7)) 判断新子序列的拟合误差 ε_{new} ; 否则, $\varepsilon_{new} = \varepsilon_0 + 1$. 执行直至 $\varepsilon_{new} > \varepsilon_0$.

Step 4: 若 $\varepsilon_{new} > \varepsilon_0$, 则删除新子序列的最后一个数据构成新子序列, 采用拟合函数, 在 $queue$ 中添加新子序列对应的信息. 令 $num = num - 1$, 如果此时 $num = n$, 则 $num = n + 1$, 执行 **Step 2**.

Step 5: 当原始序列扫描一遍后, 删除分段队列 $queue$ 中的第 1 行, 则 $queue$ 已保存原始序列的分段方法以及各子序列拟合函数对应的参数, 从而可计算出分段表示方法的拟合误差.

2.2 时间与空间复杂度分析

输入数据 $Data_{1 \times n}$ 的时间复杂度和空间复杂度均为 $O(n)$; 算法开始执行时, 设各个分段中序列的长度为 $L_i = |t_i - t_{i-1}|$, 其中 $L_{\max} = \max\{L_i\}$ 为各分段长度中的最大值, 则分段运行非齐次无偏 GM(1,1) 模型所需的时间复杂度和空间复杂度均为 $O(L_{\max})$. 因此, 整个算法最不理想的情况下, 时间复杂度和空间复杂度均为 $O(n)$.

3 实验分析

本文利用 Matlab 7.0 软件, 分别应用相关的计算公式, 建模计算模型的模拟数据, 并选用残差检验法对数据模拟精度进行比较. 为了方便比较, 以序列首项为初始条件的模型记为 Model- $x(1)$, 以累加序列最后一项为初始条件的模型记为 Model- $x(n)$;

实验 1 本文采用文献 [5] 中的 6 组不同类型的数据序列建立本文优化模型, 分析模拟结果, 并与文献 [5] 和文献 [6] 的模拟精度进行比较.

表 1 文献 [5] 中的 6 组数据

序号	$x(1)$	$x(2)$	$x(3)$	$x(4)$	$x(5)$
1	1.2	2.9	4.2	5.1	5.8
2	8.5	16.4	32.3	64.2	128.1
3	5.8	5.1	4.2	2.9	1.2
4	128.1	64.2	32.3	16.4	8.5
5	5	11	29	83	245
6	1.4	2.0	2.8	3.9	5.4

计算新背景值模型对应的平均相对误差 (MRE), 数据拟合结果见表 2. 从表 2 可以看出, 本文无偏优化模型的模拟精度均较高, 优于文献 [5] 和文献 [6] 的优化模型.

表 2 模拟数据平均相对误差 (MRE) %

序号	1	2	3	4	5	6
文献 [5]	0.2958	0.1754	0.3321	0.0756	0.0000	0.1099
文献 [6]	0.2200	0.0143	0.4000	0.0300	0.0000	0.0300
Model-x(1)	0.1831	0.0118	0.3095	0.0217	0.0000	0.0235
Model-x(n)	0.1881	0.0106	0.3157	0.0238	0.0000	0.0233

实验 2 采用 3 组不同类型的数据序列建立优化模型, 分析模拟结果, 并与文献 [7] 的模拟精度比较。

表 3 3 组不同类型的实验数据

序号	x(1)	x(2)	x(3)	x(4)	x(5)	x(6)	x(7)	x(8)
1	451	453	446	455	452	457	449	450
2	376	379	386	387	386	389	394	393
3	344	339	350	351	350	345		

计算新背景值模型对应的平均相对误差 (MRE), 数据拟合结果见表 4. 由表 4 可看出, 本节 Model-x(1) 和 Model-x(n) 两者的拟合精度均优于文献 [7] 的优化模型。

表 4 模拟数据平均相对误差 (MRE/%)

序号	Model-x(1)			MRE
	u_1	u_2	u_3	
1	-0.0929	494.2242	-45.3335	0.6211
2	1.2560	-96.1417	477.4086	0.5263
3	-0.2090	422.0667	-89.1853	0.4058

序号	文献 [7]			MRE
	a	b	c	
1	-0.0248 - 3.1416i	3.2935 + 0.0000i	452.2323	1.1422
2	-0.5953	1.6445	385.2880	17.5118
3	1.0093 - 3.1416i	29.2359 + 0.0000i	348.9795 - 0.0000i	0.5400

序号	Model-x(n)			MRE
	α	β	γ	
1	-0.0904	493.1328	-455.2420	0.5519
2	0.7301	-102.5343	-383.0776	0.5417
3	-0.9024	663.3220	-369.9756	0.4924

实验 3 采用文献 [12] 中的数据 (IBM 公司普通股从 1961-05-17 ~ 1962-11-02 每天的收盘价格 (数据长度 369)).

图 1 为传统 GM(1,1) 模型 (gm) 与非齐次无偏 GM(1,1) 模型 (ugm) 在实验中误差阈值的设置与分段数间的关系, 虚线为传统 GM(1,1) 模型的关系图, 实线为非齐次无偏 GM(1,1) 模型的关系图. 由图 1 可知, 对应相同的分段数, 非齐次无偏模型的误差阈值小于传统模型。

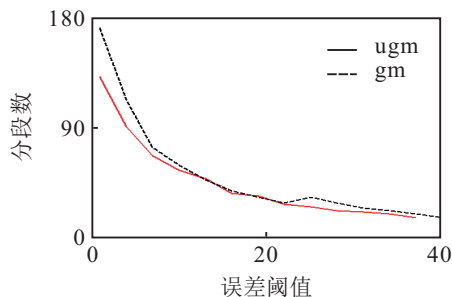


图 1 误差阈值与分段数

图 2 和图 3 分别表示分段数为 63 和 50 时拟合数据与原始数据的对比图形. 从图中可以看出, gm 模型与 ugm 模型均能够较好地拟合原始数据. gm 模型在 63 分段和 50 分段时的拟合误差分别为 53.2183 和 68.2416, 而 ugm 模型对应的拟合误差分别为 41.6955 和 62.1553. 可以看出 ugm 模型在时间序列拟合上更优于 gm 模型。

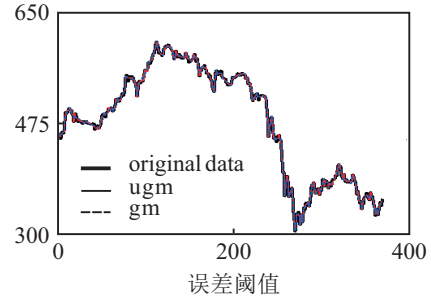


图 2 分段数为 63 时的拟合

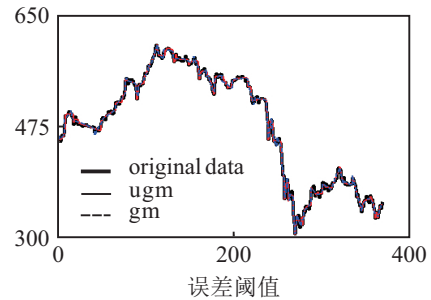


图 3 分段数为 50 时的拟合

图 4 为拟合误差随压缩率变化图, 随压缩率的增加, 拟合误差逐渐变大. 对应相同的压缩率, ugm 模型的拟合误差均小于 gm 模型的拟合误差. 随着压缩率的增加, 分段数变少. 因此, 拟合误差随分段数的增加逐渐减少 (如图 5).

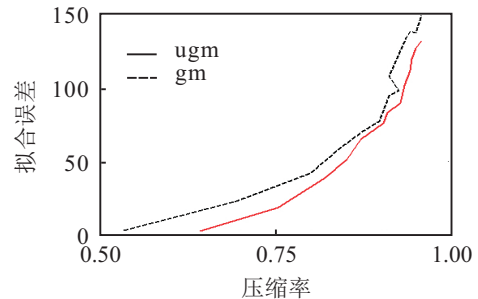


图 4 拟合误差随压缩率变化

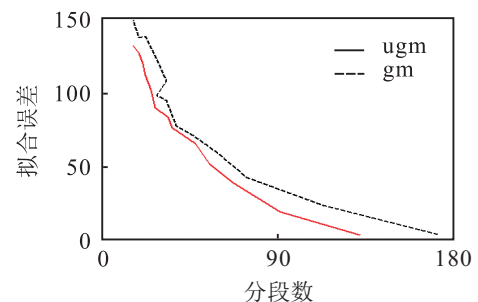


图 5 拟合误差随分段数变化

4 结 论

本文推导得到了非齐次指数序列的无偏模型的递推解法,给出了无偏GM(1,1)模型在不同初始值下的建模方法和表达式.实例分析表明,所提方法的模拟和预测精度得到了明显提高,从而扩大了模型的适用范围.在递推模型的基础上,将其运用于时间序列的分段表示算法中,提出了基于非齐次无偏GM(1,1)模型的时间序列分段表示算法.通过实例验证,本文所提方法在相同的压缩率下,能以较少的误差拟合原始数据.

参考文献(References)

- [1] Liu S F, Lin Y. Grey information: Theory and practice applications[M]. London: Springer, 2006: 4-7.
- [2] Wang Z X, Dang Y G, Liu S F. The optimization of background value in GM(1,1) model[J]. J of Grey System, 2007, 10(2): 69-74.
- [3] 谢乃明, 刘思峰. 近似非齐次指数序列的离散灰色模型特性研究[J]. 系统工程与电子技术, 2008, 30(5): 863-867.
(Xie N M, Liu S F. Research on the non-homogenous discrete grey model and its parameter's properties[J]. Systems Engineering and Electronics, 2008, 30(5): 863-867.)
- [4] 崔杰, 党耀国, 刘思峰. 一种新的灰色预测模型及其建模机理[J]. 控制与决策, 2009, 24(11): 1702-1706.
(Cui J, Dang Y G, Liu S F. Novel grey forecasting model and its modeling mechanism[J]. Control and Decision, 2009, 24(11): 1702-1706.)
- [5] 曾波, 刘思峰. 近似非齐次指数序列的DGM(1,1)直接建模方法[J]. 系统工程理论与实践, 2011, 31(2): 297-301.
(Zeng B, Liu S F. Direct modeling approach of DGM(1,1) with approximate non-homogeneous exponential sequence[J]. Systems Engineering – Theory & Practice, 2011, 31(2): 297-301.)
- [6] 陈芳, 魏勇. 近非齐次指数序列GM(1,1)模型灰导数的优化[J]. 系统工程理论与实践, 2013, 33(11): 2874-2878.
(Chen F, Wei Y. Approximate non-homogeneous index sequece GM(1,1) model of grey derivative optimization[J]. Systems Engineering – Theory & Practice, 2013, 33(11): 2874-2878.)
- [7] 战立青, 施化吉. 近似非齐次指数数据的灰色建模方法与模型[J]. 系统工程理论与实践, 2013, 33(3): 689-694.
(Zhan L Q, Shi H J. Methods and model of grey modeling for approximation non-homogenous exponential data[J]. Systems Engineering – Theory & Practice, 2013, 33(3): 689-694.)
- [8] 宋辞, 裴韬. 基于特征的时间序列聚类方法研究进展[J]. 地理科学进展, 2012, 31(10): 1307-1317.
(Song C, Pei T. Research progress in time series clustering methods based on characteristics[J]. Progress in Geography, 2012, 31(10): 1307-1317.)
- [9] Keogh E, Chakrabarti K, Mehrotra S, et al. Locally adaptive dimensionality reduction for indexing large time series databases[J]. ACM Transactions on Database Systems, 2001, 27(2): 188-228.
- [10] Keogh E. The UCR time series data mining archive[DB/OL]. (2014-12-01)[2002-07-01]. <http://www.cs.ucr.edu/~eamonn/TSDMA/index.html>.
- [11] Prat K B, Fink E. Search for patterns in compressed time series[J]. Int J of Image and Graphics, 2002, 2(1): 89-106.
- [12] 江艺羨. 基于GM(1,1)模型的时间序列分段表示方法[J]. 系统工程, 2014, 32(7): 137-142.
(Jiang Y X. Method of time series piecewise representation based on model GM(1,1)[J]. Systems Engineering, 2014, 32(7): 137-142.)

(责任编辑: 齐 霖)