

基于残差思想的区间灰数预测优化模型

党耀国[†], 叶 璟

(南京航空航天大学 经济与管理学院, 南京 210016)

摘要: 针对区间灰数的预测问题, 通过构建区间灰数核的预测模型并结合残差修正思想对区间灰数的信息域进行扩展, 运用改进的区间灰数函数处理方法强化对区间灰数上下界序列的趋势拟合, 优化对未来发展状态的预测精度, 建模过程体现了区间灰数预测中“信息充分利用”的原则. 通过长江三角洲地区人均工业废水排放量的实例, 对比所提出方法和传统灰色预测方法的结果, 验证了所提出方法的有效性和实用性. 所提出的方法能够为区间灰数预测提供一种有效途径, 其原理清晰, 可广泛应用于存在区间灰数的各领域.

关键词: 区间灰数; 预测; 信息域; 残差修正; 函数变换

中图分类号: N945.1

文献标志码: A

Optimized grey prediction model of interval grey numbers based on residual corrections

DANG Yao-guo[†], YE Jing

(College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: For the prediction of interval grey numbers, the prediction model based on the kernel sequence of interval grey number is constructed, and the idea of information domain is expanded based on residual corrections in this paper. To be specific, the information domain is divided into two parts and processed by the improved function transformation to strengthen the fitting effects of the trends of the upper and lower bounds in interval grey numbers before establishing prediction models respectively. By combination of the forecasting models of the kernel sequence and the processed information domains, the prediction results for the interval grey numbers are optimized and the principle of “full usage of information” is reflected during the modeling process of the interval grey numbers. Through discussing the case of the per capita industrial wastewater discharge in the Yangtze River Delta, the results of this method is verified by compared with traditional grey prediction methods of interval grey numbers, which shows its effectiveness and practicability. The proposed method provides another feasible forecasting method for the interval grey number prediction. The clear principle and modeling mechanism of this method make it possible to be applied in every field where interval grey numbers exists.

Keywords: interval grey number; prediction; information domain; residual correction; function transformation

0 引言

现实世界中, 由于人类认知所限或受制于系统本身的特点, 多数系统行为序列往往运用区间数据表征其发展变化趋势. 自上世纪八十年代邓聚龙教授提出灰色系统理论^[1]以来, 区间灰数作为典型的研究对象, 被学者们广泛关注并研究, 在区间灰数排序处理^[2-4]、区间灰数的关联聚类分析^[5-6]、区间灰数的评价决策^[7-10]等方面取得了丰硕的成果.

在现有的区间灰数预测研究中, 学者们大多运用

有限的区间灰数信息将其转化为实数序列进行预测. 吴利丰等^[11]计算了灰数层的面积和区间灰数的认知程度; 孟伟等^[12]运用灰数定义把区间灰数分解为“白部”和“灰部”; 王大鹏等^[13]基于论域提出了合成灰数灰度的定义及性质, 并对改进的核序列和灰度序列分别建立了预测模型; 袁潮清等^[14]运用发展趋势和认知程度两个维度描述了区间灰数序列; Zeng等^[15]将区间灰色数序列转换为核序列和区域序列, 用来建立灰色预测模型; 杨德岭等^[16]建立了区间灰数核序

收稿日期: 2017-03-09; 修回日期: 2017-08-22.

基金项目: 国家自然科学基金项目(71771119, 71371098); 江苏省社会科学基金重点项目(16GLA001); 南京航空航天大学基本科研业务费重大培育项目(NP2016301).

责任编委: 赵珺.

作者简介: 党耀国(1964—), 男, 教授, 博士生导师, 从事灰色系统理论、数量经济等研究; 叶璟(1988—), 女, 博士生, 从事灰色系统理论及其应用的研究.

[†]通讯作者. E-mail: iamdangy@163.com

列的 Verhulst 模型, 然后通过信息域不减推论对核进行拓展, 得到了区间灰数上界、下界的预测表达式; Tsaur^[17] 通过对具有某种特定隶属度的样本区间数据的拟预测区间进行必要性分析, 构建了具有最小回归区间的区间灰色回归模型; Wang 等^[18] 提出了一项灰色预测区间参数综合整数规划方法 (GFIPMIP) 用来支持北京综合电气环境管理项目. 这些文献对区间灰数的预测问题从不同角度进行了探讨, 丰富了灰色预测模型的理论体系, 但鲜有从误差视角进行区间灰数信息转化的研究.

本文基于区间灰数的传统核和信息域的理论基础, 引入残差修正思想, 对区间灰数序列进行优化处理, 提高模型的预测精度; 同时, 在信息域处理上, 运用改进的函数变换模型加强信息域数据序列的光滑性、压缩级比, 在保证序列总趋势的情形下, 提高还原误差, 更进一步保证了模型的可靠性; 最后, 构建出基于信息域的区间灰数核预测优化模型. 理论应用于实际, 选取长江三角洲地区人均工业废水排放量作为实例, 比较本文新模型与原始 GM(1,1) 模型以及基于区间灰数核和信息域的建模方法^[19] 的拟合与预测结果, 验证本文方法的实用性和可靠性. 本文方法在建模过程中体现了灰色系统理论“信息充分利用”的思想, 一并结合残差修正和函数变换等数据处理方法, 进一步拓展了区间灰数预测的理论体系, 为在经济、社会、农业、地质、医学等领域日益增多的区间数据预测提供了一个新的可行方法.

1 基本概念、理论和方法

1.1 基本概念

定义 1 设 $\otimes \in [a, b]$, $a \leq b$, $a, b \in R$, 称 \otimes 为实数域 R 上的区间灰数. 若 $a = b$, 则 \otimes 退化为实数, $\otimes = a$. 全体区间灰数集合记为 $[R]$. 区间灰数 \otimes 所构成的序列 $X(\otimes)$, 记作 $X(\otimes) = (\otimes(1), \otimes(2), \dots, \otimes(n))$ ^[1].

定义 2 设 $\otimes \in [a, \tilde{\otimes}, b]$ 为连续区间灰数, 其中 $a \leq \tilde{\otimes} \leq b$, a 和 b 分别为区间灰数的下边界和上边界, $\tilde{\otimes}$ 为在此区间中取值可能性最大的数, 称为区间灰数的核. 缺乏区间分布信息时, 区间灰数的核为 $\tilde{\otimes} = \frac{a+b}{2}$. 区间灰数序列 $X(\otimes)$ 中所有元素的核所构成的序列称为区间灰数序列 $X(\otimes)$ 的核序列, 记为 $X(\tilde{\otimes}) = (\tilde{\otimes}(1), \tilde{\otimes}(2), \dots, \tilde{\otimes}(n))$ ^[1].

定义 3 设 $\otimes \in [a, b]$ 为连续区间灰数, 其中 $a \leq b$, a 和 b 分别为区间灰数的下边界和上边界, 上界和下界的差值称为区间灰数的信息域, 记作 $l_{\otimes} = b - a$. 区间灰数序列 $X(\otimes)$ 中所有元素的信息域数值所构成的序列称为区间灰数序列 $X(\otimes)$ 的信息域值

序列, 记为 $X(L_{\otimes}) = (l_{\otimes}(1), l_{\otimes}(2), \dots, l_{\otimes}(n))$ ^[16].

定义 4 设 $\otimes \in [a, \tilde{\otimes}, b]$ 为连续区间灰数, 其中 $a \leq \tilde{\otimes} \leq b$, 上界与核的差值称为上部信息域, 记作 $\bar{l}_{\otimes} = b - \tilde{\otimes}$. 区间灰数序列 $X(\otimes)$ 中所有元素的上部信息域数值所构成的序列称为区间灰数序列 $X(\otimes)$ 的上部信息域值序列, 记为 $X(\bar{L}_{\otimes}) = (\bar{l}_{\otimes}(1), \bar{l}_{\otimes}(2), \dots, \bar{l}_{\otimes}(n))$. 类似地, 核与下界的差值称为下部信息域, 记作 $l_{\otimes} = \tilde{\otimes} - a$. 区间灰数序列 $X(\otimes)$ 中所有元素的下部信息域数值所构成的序列称为区间灰数序列 $X(\otimes)$ 的下部信息域值序列, 记为 $X(\underline{L}_{\otimes}) = (\underline{l}_{\otimes}(1), \underline{l}_{\otimes}(2), \dots, \underline{l}_{\otimes}(n))$. 显然, 信息域是上部信息域和下部信息域的和.

定义 5 设序列 $X = (x(1), x(2), \dots, x(n))$, 称 $\rho(k) = x(k) / \sum_{i=1}^{k-1} x(i)$ ($k = 2, 3, \dots, n$) 为序列 X 的光滑比. 若序列 X 满足: 1) $\rho(k+1)/\rho(k) < 1$, $k = 2, 3, \dots, n-1$; 2) $\rho(k) \in [0, \varepsilon]$, $k = 3, 4, \dots, n$, $\varepsilon < 0.5$. 则称 X 为准光滑序列^[1].

1.2 基本模型

首先介绍 GM(1,1) 模型. 设原始数据序列 $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$, $x^{(0)}(k) \geq 0$, $k = 1, 2, \dots, n$. 为了验证模型的预测精度, 选取前 $n-m$ 项作为建模数据, 最后 m 项 $(x^{(0)}(n-m+1), x^{(0)}(n-m+2), \dots, x^{(0)}(n-1), x^{(0)}(n))$ 作为预测数据 (具体建模过程略^[1]), 得到时间响应序列为

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

$$k = 1, 2, \dots, n-m-1; \tag{1}$$

还原值为

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

$$k = 1, 2, \dots, n-1. \tag{2}$$

运用式 (2) 可以得到 $X^{(0)}$ 序列的拟合值和预测值, 表示为 $\hat{X}^{(0)} = (\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(n))$.

当 GM(1,1) 模型精度不符合要求时, 可用残差序列建立残差 GM(1,1) 模型, 对原模型进行修正^[1]. 设 $\varepsilon^{(0)} = (\varepsilon^{(0)}(1), \varepsilon^{(0)}(2), \dots, \varepsilon^{(0)}(n-m))$, 其中

$$\varepsilon^{(0)}(k) = x^{(1)}(k) - \hat{x}^{(1)}(k), \quad k = 1, 2, \dots, n-m, \tag{3}$$

称为 $X^{(1)}$ 的残差序列. 若存在 k_0 , 满足

$$\forall k \geq k_0, \varepsilon^{(0)}(k) \geq 0, \quad n - k_0 \geq 4,$$

或者

$$\forall k \geq k_0, \varepsilon^{(0)}(k) \leq 0, \quad n - k_0 \geq 4,$$

则称 $(|\varepsilon^{(0)}(k_0)|, |\varepsilon^{(0)}(k_0+1)|, \dots, |\varepsilon^{(0)}(n-m)|)$ 为可建残差尾段, 仍记为 $\varepsilon^{(0)} = (\varepsilon^{(0)}(k_0), \varepsilon^{(0)}(k_0 +$

1), ..., ε⁽⁰⁾(n - m)).

对序列 ε⁽⁰⁾ = (ε⁽⁰⁾(k₀), ε⁽⁰⁾(k₀ + 1), ..., ε⁽⁰⁾(n - m)) 建立 GM(1,1) 模型, 得到还原时间响应式为

$$\hat{\varepsilon}^{(0)}(k + 1) = (-a_\varepsilon) \left(\varepsilon^{(0)}(k_0) - \frac{b_\varepsilon}{a_\varepsilon} \right) e^{-a_\varepsilon(k - k_0)}, \quad k \geq k_0. \quad (4)$$

若用 ε⁽⁰⁾ 修正 $\hat{X}^{(1)}$, 则称修正后的时间响应式为

$$\hat{x}^{(1)}(k + 1) = \begin{cases} \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a}, & k < k_0; \\ \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a} \pm a_\varepsilon \left(\varepsilon^{(0)}(k_0) - \frac{b_\varepsilon}{a_\varepsilon} \right) e^{-a_\varepsilon(k - k_0)}, & k \geq k_0 \end{cases} \quad (5)$$

为残差修正 GM(1,1) 模型, 简称为残差 GM(1,1) 模型, 其中残差修正值 (4) 的符号应与残差尾段 ε⁽⁰⁾ 的符号保持一致.

2 本文模型

设连续区间灰数序列 $X(\otimes) = (\otimes(1), \otimes(2), \dots, \otimes(n))$, 将 $X(\otimes)$ 中的所有元素在二维直角平面坐标系中进行映射, 分别顺次连接相邻区间灰数的上边界、核序列、下边界, 形成上边界线、核序列线、下边界线, 如图 1 所示. 其中: 上边界线数据形成的序列称为区间灰数上边界序列, 记为 $X(\bar{\otimes}) = (b(1), b(2), \dots, b(n))$; 核序列非数据形成的序列称为区间灰数序列核序列, 记为 $X(\tilde{\otimes}) = (\tilde{\otimes}(1), \tilde{\otimes}(2), \dots, \tilde{\otimes}(n))$; 下边界线数据形成的序列称为区间灰数下边界序列, 记为 $X(\underline{\otimes}) = (a(1), a(2), \dots, a(n))$.

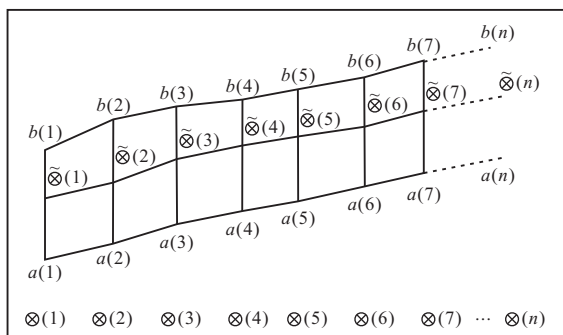


图 1 连续区间灰数上、下边界序列以及核序列连线示意图

2.1 基于核序列的 GM(1,1) 模型

对于区间灰数核序列 $X(\tilde{\otimes}) = (\tilde{\otimes}(1), \tilde{\otimes}(2), \dots, \tilde{\otimes}(n))$, 为了验证模型预测精度, 选取前 $n - m$ 项作为建模数据, 后 m 项作为预测数据. 参考 1.2 节建立的 GM(1,1) 模型, 得到区间灰数核序列的预测模型为

$$\hat{\tilde{\otimes}}^{(0)}(k + 1) = (1 - e^{a_{\tilde{\otimes}}}) \left(\tilde{\otimes}^{(0)}(1) - \frac{b_{\tilde{\otimes}}}{a_{\tilde{\otimes}}} \right) e^{-a_{\tilde{\otimes}}k}, \quad (6)$$

其中 $k = 1, 2, \dots, n - m - 1$.

运用式 (6) 可以得到核序列的拟合值和预测值, 表示为 $\hat{X}^{(0)}(\tilde{\otimes}) = (\hat{\tilde{\otimes}}^{(0)}(1), \hat{\tilde{\otimes}}^{(0)}(2), \dots, \hat{\tilde{\otimes}}^{(0)}(n))$.

2.2 基于残差修正思想的上下信息域序列构建

由定义 4 可知, 核序列的拟合值与原始区间灰数上、下界序列可以构成上、下界信息域序列. 由上信息域 $\bar{l}_\otimes = b - \tilde{\otimes}$ 和下信息域 $\underline{l}_\otimes = \tilde{\otimes} - a$ 可知, 上信息域序列 $X(\bar{L}_\otimes) = (\bar{l}_\otimes(1), \bar{l}_\otimes(2), \dots, \bar{l}_\otimes(n))$ 以及下信息域序列 $X(\underline{L}_\otimes) = (\underline{l}_\otimes(1), \underline{l}_\otimes(2), \dots, \underline{l}_\otimes(n))$ 中的各元素均非负, 这完全符合残差 GM(1,1) 模型的建模条件: 存在 k_0 , 满足 $\forall k \geq k_0, \varepsilon^{(0)}(k) \geq 0, n - k_0 \geq 4$. 由此, 可建立基于上、下信息域序列的 GM(1,1) 模型. 但是, 由于上、下信息域序列受上、下界元素以及区间灰数的核的影响, 可能呈现非平滑趋势, 首先运用定义 5 中的准光滑条件进行判别, 并根据判别结果分为以下两种情形.

1) 若满足准光滑条件, 则直接对上、下信息域分别建立 GM(1,1) 模型. 对于区间灰数上信息域序列 $X(\bar{L}_\otimes) = (\bar{l}_\otimes(1), \bar{l}_\otimes(2), \dots, \bar{l}_\otimes(n))$, 为了验证模型预测精度, 选取前 $n - m$ 项作为建模数据, 后 m 项作为预测数据. 参考 1.2 节建立的 GM(1,1) 模型, 得到区间灰数上信息域序列的预测模型为

$$\hat{\bar{l}}_\otimes^{(0)}(k + 1) = (1 - e^{a_{\bar{l}}}) \left(\bar{l}_\otimes^{(0)}(1) - \frac{b_{\bar{l}}}{a_{\bar{l}}} \right) e^{-a_{\bar{l}}k}, \quad k = 1, 2, \dots, n - m - 1. \quad (7)$$

运用式 (7) 可以得到上信息域序列的拟合值和预测值, 表示为

$$\hat{X}^{(0)}(\bar{L}_\otimes) = (\hat{\bar{l}}_\otimes^{(0)}(1), \hat{\bar{l}}_\otimes^{(0)}(2), \dots, \hat{\bar{l}}_\otimes^{(0)}(n)).$$

同理, 得到区间灰数下信息域序列 $X(\underline{L}_\otimes) = (\underline{l}_\otimes(1), \underline{l}_\otimes(2), \dots, \underline{l}_\otimes(n))$ 的预测模型为

$$\hat{\underline{l}}_\otimes^{(0)}(k + 1) = (1 - e^{a_{\underline{l}}}) \left(\underline{l}_\otimes^{(0)}(1) - \frac{b_{\underline{l}}}{a_{\underline{l}}} \right) e^{-a_{\underline{l}}k}, \quad k = 1, 2, \dots, n - m - 1. \quad (8)$$

运用式 (8) 可以得到下信息域序列的拟合值和预测值, 表示为 $\hat{X}^{(0)}(\underline{L}_\otimes) = (\hat{\underline{l}}_\otimes^{(0)}(1), \hat{\underline{l}}_\otimes^{(0)}(2), \dots, \hat{\underline{l}}_\otimes^{(0)}(n))$.

2) 若不满足准光滑条件, 则需对上、下信息域进行数据处理. 由于在区间 $[0, \pi/2]$ 上正切和余切函数的调整范围较大, 选取正切函数处理增序列, 余切函数处理减序列. 具体步骤如下.

Step 1: 数据预处理. 首先, 观察上、下信息域序列是否为单调序列, 若不是, 则将序列作累加处理 $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i)$, 保证序列的单调性. 然后, 考虑

到数据拟合的精度以及计算的简便,数据序列经幂处理、均值化等过程步骤后分别被处理到 $[0, \pi/2]$ 区间上的4个数量等级上(0.1, 0.5, 1, $\pi/2$), 得到经函数变换后的区间灰数上、下信息域序列 $X(\bar{L}_{\otimes i}) = (\bar{l}_{\otimes i}(1), \bar{l}_{\otimes i}(2), \dots, \bar{l}_{\otimes i}(n))$ 和 $X(\underline{L}_{\otimes i}) = (\underline{l}_{\otimes i}(1), \underline{l}_{\otimes i}(2), \dots, \underline{l}_{\otimes i}(n))$, 数量等级记为 $i = 1, 2, 3, 4$.

Step 2: 函数变换. 经数据标准化后的新区间灰数上、下信息域序列, 对递增序列进行正切函数变换, 即 $y = \tan x (0 \leq x \leq \pi/2)$; 对递减序列进行余切函数变换, 即 $y = \cot x (0 \leq x \leq \pi/2)$, 则有(仅以上界数据为例, 下界数据同理, 不再赘述)

$$\bar{l}'_{\otimes i}(k) = \tan \bar{l}_{\otimes i}(k), k = 1, 2, \dots, n - m, \quad (9)$$

或

$$\bar{l}'_{\otimes i}(k) = \cot \bar{l}_{\otimes i}(k), k = 1, 2, \dots, n - m. \quad (10)$$

最终, 得到函数变换数据处理后的上、下信息域序列 $Y(\bar{L}_{\otimes i}) = (\bar{l}'_{\otimes i}(1), \bar{l}'_{\otimes i}(2), \dots, \bar{l}'_{\otimes i}(n))$ 和 $Y(\underline{L}_{\otimes i}) = (\underline{l}'_{\otimes i}(1), \underline{l}'_{\otimes i}(2), \dots, \underline{l}'_{\otimes i}(n))$, 数量等级记为 $i = 1, 2, 3, 4$.

Step 3: 建立 GM(1,1) 模型. 分别运用4个数量等级的函数变换数据处理后的上、下信息域序列 $Y(\bar{L}_{\otimes i}) = (\bar{l}'_{\otimes i}(1), \bar{l}'_{\otimes i}(2), \dots, \bar{l}'_{\otimes i}(n))$ 和 $Y(\underline{L}_{\otimes i}) = (\underline{l}'_{\otimes i}(1), \underline{l}'_{\otimes i}(2), \dots, \underline{l}'_{\otimes i}(n))$, 建立 GM(1,1) 模型进行拟合计算, 得到序列 $Y(\bar{L}_{\otimes i}) = (\bar{l}'_{\otimes i}(1), \bar{l}'_{\otimes i}(2), \dots, \bar{l}'_{\otimes i}(n))$ 和 $Y(\underline{L}_{\otimes i}) = (\underline{l}'_{\otimes i}(1), \underline{l}'_{\otimes i}(2), \dots, \underline{l}'_{\otimes i}(n))$ 的拟合模型, 其中

$$\hat{l}'_{\otimes i}(0)(k+1) = (1 - e^{a'_{\otimes i}}) \left(\bar{l}'_{\otimes i}(0)(1) - \frac{b_{\bar{l}'_{\otimes i}}}{a_{\bar{l}'_{\otimes i}}} \right) e^{-a'_{\otimes i} k},$$

$$k = 1, 2, \dots, n - m, i = 1, 2, 3, 4. \quad (11)$$

Step 4: 进行逆函数变换运算还原、逆幂处理、逆均值化处理以及累减还原(如果建模前进行过累加处理), 得到4个数量等级的拟合值与预测值序列 $\hat{X}''(\bar{L}_{\otimes i}) = (\hat{l}''_{\otimes i}(0)(1), \hat{l}''_{\otimes i}(0)(2), \dots, \hat{l}''_{\otimes i}(0)(n))$ 和 $\hat{X}''(\underline{L}_{\otimes i}) = (\hat{l}''_{\otimes i}(0)(1), \hat{l}''_{\otimes i}(0)(2), \dots, \hat{l}''_{\otimes i}(0)(n))$, $i = 1, 2, 3, 4$.

Step 5: 选择合适的数量等级. 计算4个数量等级的拟合平均相对误差(MAPE), 并根据误差最小的原则, 确定上、下信息域序列的数量等级, 有

$$\text{MAPE}(\bar{l}_{\otimes i}) = \frac{1}{n - m - 1} \sum_{k=2}^{n-m} \frac{\hat{l}''_{\otimes i}(0)(k) - \bar{l}_{\otimes i}(k)}{\bar{l}_{\otimes i}(k)},$$

$$i = 1, 2, 3, 4. \quad (12)$$

得到上信息域序列数量等级(*)和下信息域序列数量等级(#)为

$$\bar{l}_{\otimes *}$$

$$\min[\text{MAPE}(\underline{l}_{\otimes 1}), \text{MAPE}(\underline{l}_{\otimes 2}), \dots, \text{MAPE}(\underline{l}_{\otimes 4})],$$

$$\underline{l}_{\otimes \#}$$

$$\min[\text{MAPE}(\underline{l}_{\otimes 1}), \text{MAPE}(\underline{l}_{\otimes 2}), \dots, \text{MAPE}(\underline{l}_{\otimes 4})].$$

Step 6: 用数量等级(*)和(#)所处的上、下信息域序列数据建立的预测模型为

$$\hat{l}'_{\otimes *}(0)(k+1) = (1 - e^{a_{\bar{l}'_{\otimes *}}}) \left(\bar{l}'_{\otimes *}(0)(1) - \frac{b_{\bar{l}'_{\otimes *}}}{a_{\bar{l}'_{\otimes *}}} \right) e^{-a_{\bar{l}'_{\otimes *}} k}, \quad (13)$$

$$\hat{l}'_{\otimes \#}(0)(k+1) = (1 - e^{a_{\underline{l}'_{\otimes \#}}}) \left(\underline{l}'_{\otimes \#}(0)(1) - \frac{b_{\underline{l}'_{\otimes \#}}}{a_{\underline{l}'_{\otimes \#}}} \right) e^{-a_{\underline{l}'_{\otimes \#}} k}, \quad (14)$$

$$k = n - m + 1, n - m + 2, \dots, n.$$

得到预测值, 并按照 Step 4 中同样的处理方法还原预测值, 最终得到该情形下上、下信息域序列的拟合值和预测值, 表示为 $\hat{X}'''(0)(\bar{L}_{\otimes *}) = (\hat{l}'''_{\otimes *}(0)(1), \hat{l}'''_{\otimes *}(0)(2), \dots, \hat{l}'''_{\otimes *}(0)(n))$ 和 $\hat{X}'''(0)(\underline{L}_{\otimes \#}) = (\hat{l}'''_{\otimes \#}(0)(1), \hat{l}'''_{\otimes \#}(0)(2), \dots, \hat{l}'''_{\otimes \#}(0)(n))$.

2.3 区间灰数上下界预测模型的构建

通过 2.1 节和 2.2 节对于核序列和上、下信息域序列的构造, 由上信息域 $\bar{l}_{\otimes} = b - \tilde{\otimes}$ 和下信息域 $\underline{l}_{\otimes} = \tilde{\otimes} - a$ 可知, 区间灰数的上、下界可以表示为

$$\begin{cases} a = \tilde{\otimes} - \underline{l}_{\otimes}, \\ b = \bar{l}_{\otimes} - \tilde{\otimes}. \end{cases} \quad (15)$$

且可以得到

$$\begin{cases} \tilde{\otimes} = (a + b)/2, \\ \underline{l}_{\otimes} = \tilde{\otimes} - a, \\ \bar{l}_{\otimes} = b - \tilde{\otimes}. \end{cases} \Leftrightarrow \begin{cases} a = (2\tilde{\otimes} - \underline{l}_{\otimes} - \bar{l}_{\otimes})/2, \\ b = (2\tilde{\otimes} + \underline{l}_{\otimes} + \bar{l}_{\otimes})/2. \end{cases} \quad (16)$$

可见, 本模型构造的区间灰数上、下界信息转化是基于有限的全信息提取的. 由式(6)~(8)或式(6)、式(13)~(16), 得到区间灰数上、下界预测模型. 对于上、下信息域满足准光滑条件的情形, 有

$$\hat{a}^{(0)}(k+1) = \hat{\tilde{\otimes}}^{(0)}(k+1) - \hat{\underline{l}}_{\otimes}^{(0)}(k+1) =$$

$$(1 - e^{a_{\tilde{\otimes}}}) \left(\hat{\tilde{\otimes}}^{(0)}(1) - \frac{b_{\tilde{\otimes}}}{a_{\tilde{\otimes}}} \right) e^{-a_{\tilde{\otimes}} k} -$$

$$(1 - e^{a_{\underline{l}}}) \left(\hat{\underline{l}}_{\otimes}^{(0)}(1) - \frac{b_{\underline{l}}}{a_{\underline{l}}} \right) e^{-a_{\underline{l}} k}, \quad (17)$$

$$\hat{b}^{(0)}(k+1) = \hat{\bar{l}}_{\otimes}^{(0)}(k+1) - \hat{\tilde{\otimes}}^{(0)}(k+1) =$$

$$(1 - e^{a_{\bar{l}}}) \left(\hat{\bar{l}}_{\otimes}^{(0)}(1) - \frac{b_{\bar{l}}}{a_{\bar{l}}} \right) e^{-a_{\bar{l}} k} -$$

$$(1 - e^{a_{\tilde{\otimes}}}) \left(\hat{\tilde{\otimes}}^{(0)}(1) - \frac{b_{\tilde{\otimes}}}{a_{\tilde{\otimes}}} \right) e^{-a_{\tilde{\otimes}} k}. \quad (18)$$

对于上、下信息域不满足准光滑条件的情形,有

$$\hat{a}^{(0)}(k+1) = \hat{\otimes}^{(0)}(k+1) - \hat{l}_{\otimes\#}^{(0)}(k+1) = (1 - e^{a_{\otimes}}) \left(\hat{\otimes}^{(0)}(1) - \frac{b_{\otimes}}{a_{\otimes}} \right) e^{-a_{\otimes}k} - (1 - e^{a_{l\#}}) \left(\hat{l}_{\otimes}^{(0)}(1) - \frac{b_{l\#}}{a_{l\#}} \right) e^{-a_{l\#}k}, \tag{19}$$

$$\hat{b}^{(0)}(k+1) = \hat{l}_{\otimes*}^{(k+1)}(k+1) - \hat{\otimes}^{(0)}(k+1) = (1 - e^{a_{l*}}) \left(\hat{l}_{\otimes}^{(0)}(1) - \frac{b_{l*}}{a_{l*}} \right) e^{-a_{l*}k} - (1 - e^{a_{\otimes}}) \left(\hat{\otimes}^{(0)}(1) - \frac{b_{\otimes}}{a_{\otimes}} \right) e^{-a_{\otimes}k}, \tag{20}$$

其中 $k = n - m + 1, n - m + 2, \dots, n - 1$.

3 实例

长江三角洲地区(包括上海市、江苏省和浙江省)是中国第一大经济区,已是国际公认的六大世界级城市群之一. 作为中国最大的经济圈,长江三角洲经济圈用仅占中国土地面积1%的区域创造了相当于中国GDP 20%的经济总量. 由此可见,长三角地区在中国社会经济发展进程中扮演着举足轻重的角色,对中国经济社会的发展有一定的指向性作用. 本文选取文献[20]中的实例进行研究,可从产业结构升级的视角,窥探长三角地区第二产业的污染排放的发展和变化,从而为全国相关产业优化改造提供参考. 作为区域性研究,单一的区域平均值难以全面反映区域内不同组成地区的发展状况,而区间灰数能够更好地表现区域内发展水平的范围,因此本实例中所获得的数据被表示为区间灰数的形式. 其中,该地区内最低水平被选为区间灰数的下边界,最高水平被选为区间灰数的上边界. 下面以2005~2012年的数据为拟合数据建立预测模型,用2013年和2014年的数据来验证预测模型的精度.

从数据上看,总体呈现下降趋势,并且在下界序列中有一定的数值波动,这给预测造成了一定的难度. 参照本文提出的预测模型,首先得到该数据的核序列,并对其进行预测;然后,得到上、下信息域序列,由于不满足准光滑条件,需要按照2.2节中的第2种方法进行处理,而后进行预测;最后,经过数据还原得到表1的拟合结果以及表2的预测结果. 为了验证本文模型,结果对比还选取了直接建模法(方法1)以及基于区间灰数核和信息域的建模方法^[19](方法2). 其中,直接建模法分别对原始区间灰数序列上、下界建立模型;基于区间灰数核和信息域的建模方法在分别将原始区间灰数序列转化成核和信息域序列之后,分别建模,再还原成上、下界数据. 最终,各模型拟合误差和预测结果分别见表1和表2.

表1 3种方法对长江三角洲地区人均工业废水排放量的拟合误差 %

	方法1		方法2		本文方法	
	下界	上界	下界	上界	下界	上界
区间灰数误差	4.66	2.16	0.62	11.08	2.90	2.43
平均误差	3.41		5.85		2.67	

表2 3种方法对长江三角洲地区人均工业废水排放量的预测误差 %

方法	2013			2014		
	下界	上界	平均误差	下界	上界	平均误差
方法1	-7.72	7.29	7.50	-8.18	13.82	11.0
方法2	-7.46	7.20	7.33	-7.89	13.72	10.80
本文方法	-4.64	4.38	4.51	2.27	6.99	4.63

从平均误差上看: 本文模型得到的拟合误差最小,说明本模型对拟合数据序列走势模拟得更准确;其次,直接建模法的拟合误差较小,为3.41%;基于区间灰数核和信息域的建模方法的拟合误差为5.85%. 3种方法的拟合误差均小于10%,处在可接受的范围内,不过本文方法的拟合程度最优. 从上、下界拟合误差看: 本文方法上、下界拟合误差均在3%以下,较为平均;基于区间灰数核和信息域的建模方法的上界拟合误差大于10%,拟合效果不十分理想;直接建模法的下界拟合误差大于本方法,上界拟合误差稍优于本方法,优势并不明显. 总的而言,相比于其他两个模型,本文方法体现出了良好的拟合效果.

由表2中的平均误差可知,本文方法对长江三角洲地区人均工业废水排放量的预测结果是3种方法中最优的. 从2013年的上、下界预测值来看,本文方法上、下界预测值误差绝对值均在5%以内,其他两种方法预测误差绝对值均处于7%~8%之间. 对于二步预测,2014年的预测值进一步显示了本文方法的有效性,误差均值仍在5%以内,下界误差为2.27%,上界误差为6.99%. 另外两种方法对于2014年数据的上界误差均大于10%,处于13%~14%之间;下界误差绝对值均大于7%,在7%~9%之间. 从2013年和2014年的预测值来看,本文方法显示出了稳定的预测效果和明显优于其他两种方法的准确性.

4 结论

随着系统数据越来越趋向于复杂化,区间灰数的应用范围越来越广泛,提高对区间灰数序列预测研究的精度有着十分重要的现实意义. 针对区间灰数预测模型的研究,本文结合区间灰数的核序列和残差修正思想,将区间灰数中点连线形成的核序列进行拟合建模,对其进行预测;然后,根据核序列拟合值与区间灰数上、下界分别形成非负的上、下信息域序列,对上、下信息域序列进行数据处理、建模预测以及数

据还原;最后,根据残差修正思想,将核序列与上、下信息域预测模型结合,得到区间灰数上、下界序列的预测值. 在本模型的建模过程中,有限的区间灰数信息得到了全信息转化,避免了信息丢失,保证了信息的有效利用. 另外,在上、下信息域序列数据处理中,运用累加和函数变换方法提高了模型精度. 在残差修正思想的推广下,将核序列与上、下信息域的预测模型结合,将区间灰数上、下界序列预测模型的准确性进一步提高,强化了本文模型的有效性. 通过长江三角洲地区人均工业废水排放量这一实例,对比本文方法与其他两种区间灰数预测模型的拟合和预测结果,表明了本文方法良好的预测效果和实用性. 在残差修正思想下,进一步探寻区间灰数序列的信息转化方法和转化序列的数据处理方法将是一个值得研究的方向.

参考文献(References)

- [1] 刘思峰, 党耀国, 方志耕, 等. 灰色系统理论及其应用[M]. 第5版. 北京: 科学出版社, 2010: 146-160.
(Liu S F, Dang Y G, Fang Z G, et al. Grey system theory and application[M]. 5th ed. Beijing: Science Press, 2010: 146-160.)
- [2] Liu S F, Lin Y. On measures of information content of grey numbers[J]. *Kybernetes*, 2006, 35(5): 899-904.
- [3] 高峰记. 可能性及区间数综合排序[J]. *系统工程理论与实践*, 2013, 33(8): 2033-2040.
(Gao F J. Possibility degree and comprehensive priority of interval numbers[J]. *Systems Engineering — Theory & Practice*, 2013, 33(8): 2033-2040.)
- [4] Xie N M, Liu S F. Novel methods on comparing grey numbers[J]. *Applied Mathematical Modelling*, 2010, 34(2): 415-423.
- [5] 张志勇, 吴声. 基于白化权函数的区间灰数关联度模型[J]. *中国管理科学*, 2015, 23(1): 154-162.
(Zhang Z Y, Wu S. Incidence degree model of interval grey number based on whitenization weight function[J]. *Chinese J of Management Science*, 2015, 23(1): 154-162.)
- [6] Wang Z X. Correlation analysis of sequences with interval grey numbers based on the kernel and greyness degree[J]. *Kybernetes*, 2013, 42(2): 309-317.
- [7] Luo D, Wang X. The multi-attribute grey target decision method for attribute value within three-parameter interval grey number[J]. *Applied Mathematical Modelling*, 2012, 36(5): 1957-1963.
- [8] Song J, Dang Y G, Wang Z X, et al. The decision-making model of harden grey target based on interval number with preference information on alternatives[J]. *J of Grey System*, 2009, 21(3): 291-300.
- [9] 王鹏飞, 李畅. 不确定多属性决策双目标组合赋权模型研究[J]. *中国管理科学*, 2012, 20(4): 104-108.
(Wang P F, Li C. The study of multiple attribute decision making based on bi-objective combined weights model[J]. *Chinese J of Management Science*, 2012, 20(4): 104-108.)
- [10] 杨保华, 方志耕, 周伟, 等. 基于信息还原算子的多指标区间灰数关联决策模型[J]. *控制与决策*, 2012, 27(2): 182-186.
(Yang B H, Fang Z G, Zhou W, et al. Incidence decision model of multi-attribute interval grey number based on information reduction operator[J]. *Control and Decision*, 2012, 27(2): 182-186.)
- [11] 吴利丰, 刘思峰, 闫书丽. 区间灰数序列的灰色预测模型构建方法[J]. *控制与决策*, 2013, 28(12): 1912-1914.
(Wu L F, Liu S F, Yan S L. Grey prediction model for hybrid sequence[J]. *Control and Decision*, 2013, 28(12): 1912-1914.)
- [12] 孟伟, 刘思峰, 曾波. 区间灰数的标准化及其预测模型的构建与应用研究[J]. *控制与决策*, 2012, 27(5): 773-776.
(Meng W, Liu S F, Zeng B. Standardization of interval grey number and research on its prediction modeling and application[J]. *Control and Decision*, 2012, 27(5): 773-776.)
- [13] 王大鹏, 汪秉文, 李睿凡. 考虑合成灰数灰度性质的改进区间灰数预测模型[J]. *系统工程与电子技术*, 2013, 35(5): 1013-1017.
(Wang D P, Wang B W, Li R F. Improved prediction model of interval grey number based on the characteristics of grey degree of compound grey number[J]. *Systems Engineering and Electronics*, 2013, 35(5): 1013-1017.)
- [14] 袁潮清, 刘思峰, 张可. 基于发展趋势和认知程度的区间灰数预测[J]. *控制与决策*, 2011, 26(2): 313-315.
(Yuan C Q, Liu S F, Zhang K. Prediction model for interval grey number based on trend and cognition[J]. *Control and Decision*, 2011, 26(2): 313-315.)
- [15] Zeng B, Li C, Long X J, et al. A novel interval grey number prediction model given kernel and grey number band[J]. *J of Grey System*, 2014, 26(3): 69-84.
- [16] 杨德岭, 刘思峰, 曾波. 基于核和信息域的区间灰数 Verhulst 模型[J]. *控制与决策*, 2013, 28(2): 264-268.
(Yang D L, Liu S F, Zeng B. Verhulst model of interval grey number based on kernel and information field[J]. *Control and Decision*, 2013, 28(2): 264-268.)
- [17] Tsaur R C. The development of an interval grey regression model for limited time series forecasting[J]. *Expert Systems with Applications*, 2010, 37: 1200-1206.
- [18] Wang X W, Cai Y P, Chen J J, et al. A grey-forecasting interval-parameter mixed-integer programming approach for integrated electric-environmental management — A case study of Beijing[J]. *Energy*, 2013, 63: 334-344.
- [19] 罗党, 李琳. 基于核和测度的区间灰数预测模型[J]. *数学的实践与认识*, 2014, 44(8): 96-100.
(Luo D, Li L. Prediction model of interval grey number based on kernels and measures[J]. *Mathematics in Practice and Theory*, 2014, 44(8): 96-100.)
- [20] 叶璟, 党耀国, 刘震. 基于余切函数变换的区间灰数预测模型[J]. *控制与决策*, 2017, 32(4): 688-694.
(Ye J, Dang Y G, Liu Z. Grey prediction model of interval grey numbers based on cotangent function transformation[J]. *Control and Decision*, 2017, 32(4): 688-694.)