

基于改进的 GM_Markon 模型对北京创意产业发展的预测

刘鸿雁, 齐小芸

(华北电力大学 经济管理系, 河北 保定 071003)

摘要:创意产业产值受周边环境因素影响较大,在某些年份甚至出现局部较大畸变,通过相关文献学习,首先通过对灰色模型预测,然后经过残差修正增加模型规律性(采用两次残差累加),最后在残差修正灰色模型基础上应用改进的马尔科夫链提高预测精度,对2005—2014年北京创意产业产值数据验证算法有效性进行研究。此算法优点在于弥补灰色模型数据少,随机性大的缺陷,很好修正畸变数据和提高拟合度。

关键词:北京创意产业;灰色模型;残差修正;改进灰色马尔科夫

中图分类号:G315 **文献标志码:**A **文章编号:**1671-1807(2016)03-0028-04

创意产业以文化为基础,强调文化产业与第一、第二以及第三产业的融合和渗透,推动经济、社会和文化三个维度的全面转型。是推动经济发展的原动力和塑造国际形象的主要途径。本文研究北京创意产业产值时间序列短、信息不完全,通常预测此类数据采用灰色系统。灰色系统模型 GM(1,1)通过变量的一阶微分方程模型揭示内在发展规律,对系统未来发展趋势进行预测。灰色理论模型中累加只能增强规律性不能弱化随机性。因此有必要进行残差修正提高预测精度。然而残差修正仍然不能满足精度要求,文献^[4]用灰色马尔科夫预测负荷,根据灰色模型的残差通过马尔科夫链进行修正,可以小范围修正预测精度。本文采用修正后的 GM(1,1)模型与改进马尔科夫链相结合的方法,首先用灰色模型对北京创意产业发展进行预测,然后利用残差进行修正,最后通过马尔科夫链对残差再次修正,来达到提高精度目的。

1 改进的灰色马尔科夫理论介绍

1.1 建立原始数据的 GM(1,1)模型^[1]

1)原始数据序列 $x^{(0)}$

$$x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)\}$$

(1)

$$B = \begin{bmatrix} \frac{1}{2}(x^{(1)}(2) + x^{(1)}(1)) & \frac{1}{2}(x^{(1)}(3) + x^{(1)}(2)) & \dots & \frac{1}{2}(x^{(1)}(n) + x^{(1)}(n-1)) \\ 1 & 1 & \dots & 1 \end{bmatrix}^T \quad (6)$$

根据 GM(1,1)模型可以求出 $x^{(1)}(k)$ 的预测值,

式中: n 为采集的样本数。

2)原始数据一次累加序列 $(1-AG0) x^{(1)}$

$$x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), x^{(1)}(3), \dots, x^{(1)}(n)\} \quad (2)$$

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

$x^{(1)}$ 的临近均值生成序列 $z^{(1)}$

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

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

3)建立白化(影子)方程

$$\frac{\partial x^{(1)}}{\partial t} + ax^{(1)} = u \quad (4)$$

其中: a 为发展系数, u 称为内生控制系数。

4)利用最小二乘法求解 a 和 u

$$\begin{bmatrix} a \\ u \end{bmatrix} = (B^T B)^{-1} B^T Y \quad (5)$$

其中: $Y = (x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n))^T$ 。

计算公式如下:

收稿日期:2015-10-30

作者简介:刘鸿雁(1971—),女,河北保定人,华北电力大学经济管理系,副教授,硕士生导师,研究方向:公司金融、企业绩效评价等;齐小芸(1989—),女,河北保定人,华北电力大学经济管理系,工业工程硕士研究生,研究方向:现代企业信息统计。

$$\dot{x}^{(1)}(i+1) = (x^{(0)}(1) - \frac{u}{a})e^{-ai} + \frac{u}{a} \quad (7)$$

其中: $i = 1, 2, \dots, n$ 。

原始数据模型值 $\dot{x}^{(0)}(i+1)$ 计算公式如下:

$$\dot{x}^{(0)}(i+1) = (1 - e^a)(x^{(0)}(1) - \frac{u}{a})e^{-ai} \quad (8)$$

1.2 建立 GM(1,1)残差模型

原始数据序列建立 GM(1,1)模型残差检验精度没有达到要求精度时,可以 GM(1,1)残差模型提高预测精度。

1)定义残差序列 $\epsilon^{(0)}$

$$\epsilon^{(0)} = \{\epsilon^{(0)}(1), \epsilon^{(0)}(2), \epsilon^{(0)}(3), \dots, \epsilon^{(0)}(n)\} \quad (9)$$

其中: $\epsilon^{(0)}(j) = x^{(1)}(j) - \hat{x}^{(1)}(j)$, $\hat{x}^{(1)}(j)$ 根据公式(6)计算 $x^{(1)}(j)$ 预测值。

2)计算生成 $\epsilon^{(1)}$, 建立相应 GM(1,1)模型

$$\dot{\epsilon}^{(1)}(j+1) = (\epsilon^{(0)}(1) - \frac{u}{a})e^{-aj} + \frac{u}{a} \quad (10)$$

3)修正模型

$$\dot{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{u}{a}\right)e^{-ak} + \frac{u}{a} + \delta(k-j)(-a_e)\left(\epsilon^{(0)}(1) - \frac{u}{a}\right)e^{-ak} \quad (11)$$

其中: $\delta(k-j) = \begin{cases} 1 & k \geq j \\ 0 & k < j \end{cases}$ 为修正参数, i 为修正起点。

本文因残差数据一次累加后指数规律不明显,因此采用再次累加。

1.3 改进马尔科夫链修正模型^[3]

1)状态 E_i 。状态指事物在某个时刻出现的一种结果。这里定义创意产业产值的相对误差范围 E_i 为 n 种状态 $(E_1, E_2, \dots, E_n)$ 中一种。

2)状态概率向量 S 。状态概率向量指在某阶段所有状态可能出现的概率向量,用 S 表示:

$$S = (S_1, S_2, \dots, S_n)$$

满足以下条件:

$$\begin{cases} 0 < S_i < 1 \\ \sum_{i=1}^n S_i = 1 \end{cases} \quad (12)$$

其中: S_i 为状态概率,表示某阶段系统处于状态 E_i 的可能性大小。

3)状态转移概率 $p_{ij}(k)$ 。 $p_{ij}(k)$ 表示 k 步转移概率,即:

$$p_{ij}(k) = p(E_i \xrightarrow{k} E_j) \quad (13)$$

其中: $i, j = 1, 2, \dots, n$ 。

建模步骤如下:

1)划分创意农业产值相对误差阶段和状态。

a. 确定取值范围为 $[x_{\min}, x_{\max}]$

b. 确定相对误差的状态数 n

c. 确定状态步长 $L = (x_{\max} - x_{\min})/n$

2)确定状态转移概率和状态转移概率矩阵。一般用频率近似概率的思想计算,当样本空间足够大时,可以近似用状态相互转移频率来统计状态量和状态转移。设 M_i 为处于状态 E_i 的样本数, $M_{ij}(k)$ 为状态 E_i 经过 k 步转移到状态 E_j 的样本数, $p_{ij}(k)$ 计算如下:

$$p_{ij}(k) = p(E_i \xrightarrow{k} E_j) = \frac{M_{ij}(k)}{M_i} \quad (14)$$

相应 k 步状态转移概率矩阵 $p(k)$ 如下式:

$$p(k) = \begin{bmatrix} p_{11}(k) & p_{12}(k) & \dots & p_{1n}(k) \\ p_{21}(k) & p_{22}(k) & \dots & p_{2n}(k) \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1}(k) & p_{n2}(k) & \dots & p_{nn}(k) \end{bmatrix} \quad (15)$$

3)建立马尔科夫修正模型。由 $t-k$ 阶段的状态和状态转移概率矩阵,可以求出 $t-k$ 阶段以后任一阶段的状态向量,计算公式如式(16):

$$S(t) = S(t-k) \times p(k) \quad (16)$$

4)对马尔科夫预测模型进行改进。当灰色模型与马尔科夫链相结合,考虑到灰色模型中数据有限,马尔科夫模型进行状态转移时,转移概率矩阵有一定不稳定性和概率相同性,为减少由于一步状态转移所产生误差和不确定性,故本文采用两部平均法进行改进。

$$S(t) = \frac{S(t-1) \times p(1) + S(t-2) \times p(2)}{2} \quad (17)$$

5)灰色马尔科夫链修正。由相对误差公式,推导出灰色马尔科夫链修正后预测值 $\hat{x}^{(0)}(k)$:

$$\hat{x}^{(0)}(k) = \frac{\dot{x}^{(0)}(k)}{1 - \dot{\epsilon}^{(0)}(k)} \quad (18)$$

其中: $k = 1, 2, \dots, n$ 。

2 算例应用

2.1 建立 GM(1,1)模型并残差修正

样本数据来源北京统计年鉴其中 2005—2012 为训练样本,2013—2014 年做误差分析样本。灰色模型和残差修正后灰色模型预测结果如图 1 所示。残差修正后灰色模型预测总体更好地拟合创意产业产值真实值变化情况,2009—2011 年原始数据称凸型,与修正前灰色模型比较,修正后的灰色模型修正了偏移,提高拟合度。样本结构图如图 1。

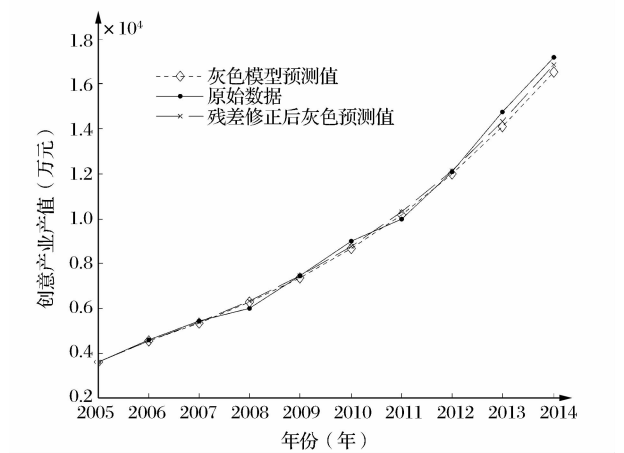


图 1 灰色模型和残差修正后灰色模型预测结果

表 1 灰色预测值和残差修正后灰色预测值相对误差

年份(年)	原始值(万元)	灰色预测		残差修正后灰色预测	
		预测值(万元)	相对误差(%)	预测值(万元)	相对误差(%)
2013	15 570.9	13 764	9.28%	14 005	7.69%
2014	17 498.0	16 126	7.31%	16 457	5.41%

表 2 状态划分

状态编号	相对误差范围
1	-0.05~-0.027
2	-0.027~-0.004
3	-0.004~0.019
4	0.019~0.042

表 3 残差修正和相对误差及所属状态

年份	原始值	残差修正	相对误差	所属状态
2005	3 614.8	3 615	0	3
2006	4 601.6	4 565	0.008 0	3
2007	5 439.6	5 371	0.012 6	3
2008	5 985.7	6 320	-0.055 9	1
2009	7 442.3	7 439	0.000 5	3
2010	9 012.2	8 757	-0.028 3	4
2011	9 983.0	10 312	0.033 0	1
2012	12 082.3	12 147	-0.005 3	2

表 5 残差修正后预测值和改进灰色马尔科夫预测值及相对误差

年份(年)	原始值(万元)	残差修正后灰色预测		改进灰色马尔科夫预测	
		预测值(万元)	相对误差(%)	预测值(万元)	相对误差(%)
2013	15 170.9	14 005	7.69%	14 446	4.78%
2014	17 398.0	16 457	5.41%	16 581	4.70%

2005—2012 年波动比较平缓,但是 2013—2014 年样本值波动较大,残差修正后使得误差有一定程度减小,但是不满足高精度要求,如表 1 所示。

2.2 改进的马尔科夫误差修正

表 2 中对进行状态表的编制,先求得 2005 到 2012 年相对误差,划分四个状态:状态 1 相对范围为 -0.05~-0.027;状态 2 相对范围为 -0.027~-0.004;状态 3 相对范围为 0.009~0.038 6;状态 4 相对范围为 0.019~0.042。相应所属状态如表 3 所示。

根据表 3 中的所属状态情况,结合平均多步转移式(16),可得到 2013 与 2014 年状态概率向量,具体数值如下表 4。

表 4 改进马尔科夫预测状态向量

年份	状态向量			
	状态 1	状态 2	状态 3	状态 4
2013	0	0	0	1
2014	0.25	0	0.5	0.25

由表 3、表 4 可看到 2013—2014 年状态向量均处于状态 3,误差范围 0.009~0.038 6。根据式(2)一式(18)对 2013—2014 年创意产业产值修正。具体计算结果如表 5。

由表 5 可以看出改进灰色马尔科夫预测 2013 年和 2014 年创意产业产值相对误差均在 5%以内,相对于残差修正后灰色预测预测精度有明显提高。

3 结论

创意产业产值受诸多因素影响,随机性大,容易

出现畸变。本实例应用中,创意产业产值预测有灰色相对误差 7%~9%,误差较大。经残差修正后相对误差范围缩小到 5%~7%,误差范围明显减小,但仍

不能满足精度要求。故残差修正基础上然后经过马尔科夫再次修正,使得精度满足要求,同时证明了本算法对畸变较大的样本具有有效性。

灰色预测具有在小样本数据下,可以达到较高精度要求,缺点是容易受外界干扰,随机波动性大,经累加后指数规律不明显。在算法中初始值选取,紧邻均值生成有待改进。马尔科夫链使用随机过程的状态转移行为,刚好可以弥补灰色预测的局限。残差灰色模型具有平稳过程,等均值等特点,满足马尔科夫要求,两者结合,实现优势互补。但是马尔科夫模型状态出现的频率近似有频率来计算的,所以在小样本情况下容易出现不稳定性和不确定性。因此采用多步平均法来弥补这一缺陷,实例中证明此算法可以很好提高预测精度。对状态数量确定和相对误差范围计算预测结果有待深入研究。

参考文献

[1] 杨华龙,刘金霞,郑斌. 灰色预测 GM(1,1)模型的改进及应用[J]. 数学的实践与认识,2011(23):39—46.

[2] 刘宗明,贾志绚,李兴莉. 基于灰色马尔科夫链模型的交通量预测[J]. 华东交通大学学报,2012(1):30—34.

[3] 夏耀杰,程浩忠. 基于改进灰色马尔科夫预测法的中长期负荷预测[J]. 电力需求侧管理,2015(1):21—25.

[4] 刘笑冰,陈建成,何忠伟. 基于 MGM(1,N)模型的北京创意农业发展灰色预测[J]. 中国人口. 资源与环境,2013(4):62—66.

[5] 夏秀峰,董彦军. 基于改进马尔科夫模型的航空备件需求预测[J]. 兵工自动化,2013(11):39—41,48.

Predict the Development of Creative Industries of Beijing Based on Improved GM Markov Model

LIU Hong-yan, QI Xiao-yun

(department of Economic Management, North China Electric Power University, Baoding Hebei 071003, China)

Abstract: Creative industry output is influenced by lots of surroundings factors Even the output appears a partial large distortion in some years. Through literature study, First through the gray model prediction output, then increased regularity by residual correction model(two residual accumulation), Finally, improving the forecast accuracy meet to requirements using Improved Markov chain based on residuals error correction of the gray mode, research on validation algorithms effectiveness data Of the creative industry output of Beijing for 2005—2014. The advantage of this algorithm is that s making up the small data of gray mode and randomness large, correcting distortion and improving goodness of fit.

Key words: creative Industry of Beijing; Grey model; residual error correction; improved gray Markov

(上接第 27 页)

Research on Capacity Adaptability of Steel Industry in Hebei Province Based on DEA

LI Li, XING Ting-ting, ZHANG Jing-min

(School of Economy and Management, Hebei University of Technology, Tianjin 300400, China)

Abstract: This paper uses C^2R model and BC^2 model of DEA about the input oriented model to study the capacity adaptability of steel industry in Hebei Province. It sees the number of employees, energy consumption per ton, total assets, main business costs as input indicators, and the unit of industrial output, total assets contribution rate as output indicators. The results show that the steel industry is not always over capacity in Hebei Province, and also does not always meet the development of national economy. The capacity adaptability of steel industry may be influenced by the policy that local government and companies make.

Key words: data envelopment analysis; steel industry; capacity adaptability