

基于 GM (1, 1) - M arkov模型的我国人口城市化水平预测

石留杰, 李艳军, 臧雨亭, 戴晓明, 张妩娜

(安阳工学院数理学院, 河南 安阳 455000)

摘要: 结合灰色 GM (1, 1)模型和马尔可夫链理论的优点, 建立耦合的 GM (1, 1) - M arkov预测模型。实例预测 2009年—2015年我国人口城市化水平, 其结果证明 GM (1, 1) - M arkov模型预测精度较高, 具有较强的科学性和实用性。

关键词: GM (1, 1)模型; M arkov链; 城市化; 预测; M atlab

中图分类号: O241

文献标识码: A

引 言

改革开放以来, 我国城市化得到了健康有序地发展, 如何确保我国城市化的可持续发展, 制定正确的城市发展方针政策显得极为重要。对城市化发展趋势的准确把握无疑会对城市发展方针政策的制定提供重要的理论依据。

针对灰色模型在长期预测时, 数据序列拟合较差, 预测精度偏低; 而马尔可夫链适用于长期、数据序列随机波动大的模型。本文依据我国城市化发展特征, 建立了 GM (1, 1) - M arkov预测模型, 分析我国人口城市化水平的发展趋势。所建模型通过了精度检验, 具有较高的预测精度, 其预测结果基本上反映了我国人口城市化水平发展的实际情况。

1 模型建立

1.1 灰色 GM (1, 1)模型

设灰色信息原始数据的时间序列为

$$x^{(0)} = \{x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}\} \tag{1}$$

为了使原始数据的时间序列内部规律性强化, 随机波动性减弱, 可对 (1)进行一次累加 (AGO)处理

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

得到序列

$$x^{(1)} = \{x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}\} \tag{2}$$

对 $x^{(1)} = \{x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}\}$ 建立白化形式方程:

$$\frac{dx^{(1)}}{dt} + \alpha x^{(1)} = u \tag{3}$$

这是一个灰色一阶微分方程, 记为 GM (1, 1)。式中:

α, u 为由已知序列确定的常数, 可按最小二乘法求解。

设 \hat{a} 为参数向量, B 为灰色序列矩阵, $\hat{a} = [au]^T$, $Y_n = [x_2^{(0)}, x_3^{(0)}, \dots, x_n^{(0)}]^T$ 及

$$B = \begin{bmatrix} -\frac{1}{2}[x_1(1) + x_1(2)] & 1 \\ -\frac{1}{2}[x_1(2) + x_1(3)] & 1 \\ \dots & \dots \\ -\frac{1}{2}[x_1(n-1) + x_1(n)] & 1 \end{bmatrix}$$

利用最小二乘法, 通过 m atlab拟合求解参数

$$\alpha, u。即 \hat{a} = (B^T B)^{-1} B^T Y_n = \begin{bmatrix} \hat{\alpha} \\ \hat{u} \end{bmatrix}$$

该模型的时间响应函数为:

$$\hat{x}_{kt+1}^{(1)} = \left[x_1^{(1)} - \frac{u}{a} \right] e^{-ak} + \frac{u}{a} \tag{4}$$

经过一次累减即可得到原数列的预测值:

$$\begin{aligned} \hat{x}^{(0)}(k+1) &= x^{(1)}(k+1) - x^{(1)}(k) \\ &= \left[x_1^{(1)} - \frac{u}{a} \right] (1 - e^{-a}) e^{-ak} \end{aligned} \tag{5}$$

$$k = 0, 1, 2, 3, \dots$$

1.2 M arkov链模型

设 $\{X_n, n \geq 0\}$ 是一个随机变量序列, $X_n \in E$ 。如果对于任意的 $n \geq 1$, 任意非负整数 $i_0, i_1, i_2, \dots, i_{n-1}, i_j \in E$, 恒有 $P(X_{n+1} = j | X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i) = P(X_{n+1} = j | X_n = i)$, 则称 $\{X_n, n \geq 0\}$ 为 Markov 链。

若 P_{ij} 表示从状态 i 转移到状态 j 的概率, 则转移概率矩阵可表示为,

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix} \quad (6)$$

$P_{ij} \geq 0$ 且 $\sum_{j \in E} P_{ij} = 1$ 。

1.3 GM (1, 1) - Markov模型

下面给出 GM (1, 1) - Markov 模型的具体建模步骤。

(1) 建立 GM (1, 1) 模型并求解, 得到 $\hat{x}^{(0)}(k+1)$ 的预测序列, 该序列给出了原始序列的变化趋势。

$$\begin{aligned} \hat{x}^{(0)}(k+1) &= x^{(1)}(k+1) - x^{(1)}(k) \\ &= [x_1^{(1)} - \frac{u}{a}] (1 - e^{-a}) e^{-ak} \end{aligned}$$

(2) 状态划分和状态转移矩阵的计算。将 $\hat{x}^{(0)}(k+1)$ 划分为 n 个状态, 任一状态可表示为, $\Delta_i = [\Delta'_{1i}, \Delta'_{2i}]$; $\Delta'_i \in \Delta, i = 1, 2, 3, \dots, n$ 。

$$\left. \begin{aligned} \Delta'_{1i} &= Y(k) + A_i \\ \Delta'_{2i} &= Y(k) + B_i \end{aligned} \right\}$$

式中, 灰元 $\Delta'_{1i}, \Delta'_{2i}$ 随时间 k 变化而变化, 状态划分的数目 n 以及 A_i, B_i 可根据预测数据 $\hat{x}^{(0)}(k+1)$ 来确定状态转移概率:

$$P_{ij}(m) = \frac{M_{ij}(m)}{M_j}, \quad (i = 1, 2, \dots, n)$$

状态转移概率矩阵:

$$P(m) = \begin{pmatrix} P_{11}(m) & P_{12}(m) & \dots & P_{1n}(m) \\ P_{21}(m) & P_{22}(m) & \dots & P_{2n}(m) \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1}(m) & P_{n2}(m) & \dots & P_{nn}(m) \end{pmatrix} \quad (7)$$

式中, $M_{ij}(m)$ ——由状态 Δ_i 经过 m 步转移到状态 Δ_j 的原始数据样本数;

M_i ——处于状态 Δ_i 的原始数据样本数;

$P_{ij}(m)$ ——由状态 Δ_i 经过 m 步转移到状态 Δ_j 的概率。

为了保证一步转移的确定性, 每一行的转移概率 P_{ij} 都要确保有唯一的最大值。

(3) 计算预测值。由 $P(m)$ 可知系统未来时刻最有可能的转移状态, 且认为预测值 $\varphi(k+1)$ 的取值最有可能为灰区间的中点处, 即

$$\varphi(k+1) = \frac{1}{2} (\Delta'_{1i} + \Delta'_{2i}) = Y(k) + \frac{1}{2} (A_i + B_i) \quad (8)$$

2 我国城市化水平的趋势分析与预测

本文选取我国 1997-2008 年间的人口城市化水平原始数据进行分析。其原始数据序列如表 1 所示。

表 1 1997-2008 年全国人口城市化水平变动状况表

年份	城市化水平 (%)	年份	城市化水平 (%)
1997	31.91	2003	40.53
1998	33.35	2004	41.76
1999	34.78	2005	42.99
2000	36.22	2006	43.90
2001	37.66	2007	44.94
2002	39.09	2008	45.68

采用 GM (1, 1) - Markov 模型进行预测, 并与 GM (1, 1) 模型进行对比。

(1) 根据 GM (1, 1) 得到数列的预测模型:

$$\hat{x}^{(1)}(k+1) = 215.5327e^{-0.125k} + 246.3327$$

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

$$Y(k) = \hat{x}^{(0)}(k+1), \quad k = 0, 1, 2, \dots$$

根据灰色预测模型精度检验等级参照表, 上述建立的飞行事故灰色新陈代谢 GM (1, 1) 模型预测精度为一级, 模型能很好地反映我国城市化水平的总体变化趋势, 可以用于城市化水平预测。

(2) 状态划分: 根据 1997-2008 年我国人口城市化水平的实际值和拟合值的百分比, 将城市化水平状态划分为表 2。

表 2 城市化水平状态划分

编号	状态	实际值 / 预测值
Δ_1	微上升年度	80% - 90%
Δ_2	一般上升年度	90% - 100%
Δ_3	显著上升年度	100% - 115%
Δ_4	大幅上升年度	115% - 135%

由表 2 可得到城市化水平一步状态转移概率矩阵如下:

$$P(1) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{4} & 0 & \frac{2}{4} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

根据以上状态转移概率矩阵可以预测城市化水平。由表 1 知 1998 年我国人口城市化水平处于 Δ_3 的状态, 则根据状态转移概率确定方法, 考察状态转移矩阵 P 的第 3 行, 有 $\max P_{3j} = P_{33} (j = 1, 2, 3, 4)$, 经过 1 年转移,

1999年人口城市化水平最有可能处于 Δ_3 的状态, 根据 (3)、(5)和 (8)式, 得 1999年人口城市化水平最有可能

$$\varphi(1) = \frac{1}{2}(1 + 1.1) \times 32.86 = 34.50$$

(3)给出预测值。

我们得到了 1998-2008年的预测值, 与 GM(1, 1)模型的预测值比较(见表 3):

表 3 GM(1, 1) - Markov模型与 GM(1, 1)模型检验性预测结果比较

年份	实际值	GM(1, 1)模型		GM(1, 1) - Markov模型	
		预测值	相对误差 (%)	预测值	相对误差 (%)
1998	33.35	31.36	- 5.97	33.05	- 0.90
1999	34.78	32.79	- 5.72	34.50	- 0.81
2000	36.22	35.46	- 2.10	36.68	1.27
2001	37.66	38.69	2.76	37.42	- 0.64
2002	39.09	38.92	4.34	39.60	1.30
2003	40.53	41.66	2.79	40.21	- 0.79
2004	41.76	42.69	2.23	41.96	0.48
2005	42.99	43.66	1.56	42.35	1.49
2006	43.90	44.90	2.28	43.93	0.07
2007	44.94	46.52	3.52	44.98	0.09
2008	45.68	48.69	6.59	45.73	0.11

由表 3可知 GM(1, 1) - Markov模型预测值与实际值拟合较好, 较传统 GM(1, 1)模型预测精度要高。因此给出了 2009-2015年我国人口城市化水平的预测值(见图 1)。

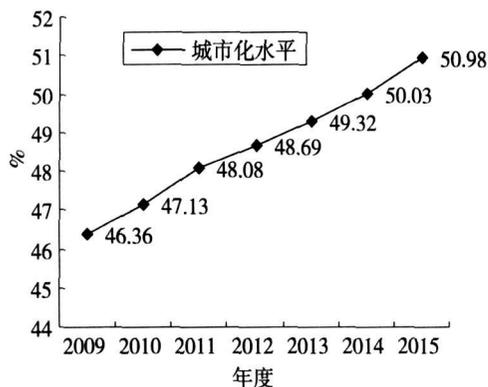


图 1 2009-2015年我国人口城市化水平预测

3 结论

(1) 实验证明 GM(1, 1) - Markov预测模型去掉已失去参考价值的历史老信息, 补充新信息, 克服了随机波动性数据的影响, 提高了灰色预测的应用水平, 较以往文献的预测精度要高。

(2) 根据预测, 至 2014年我国人口城市化水平为 50.03%, 达到发达国家 1950年城市化的平均水平, 我国开始步入城市化的快速发展阶段。为确保我国城市化的可持续发展, 急需制定正确的城市发展方针, 本文的预测数据为我国城市发展方针、政策的制定提供了重要的理论依据。

参考文献

- [1] 燕列雅, 毛联霞. GM(1, 1)模型应用[J]. 系统工程理论与实践, 1998, 18(10): 104-106
- [2] 李艳军, 赵玉亮. 基于灰色 GM(1, 1)模型的河南省能源需求量预测[J]. 河南科学, 2009, 27(12): 1508-1511
- [3] 刘晓叙. 灰色预测与一元线性回归预测的比较[J]. 四川理工学院学报:自然科学版, 2009, 22(1): 67-71
- [4] 叶宗文, 罗珊, 曾波, 李彦. 股票价格的马氏链预测法[J]. 重庆师范大学学报:自然科学版, 2006, 23(1): 75-80
- [5] 唐晓彬. Markov机制转换的状态空间模型及其在我国经济周期中的应用研究[J]. 统计研究, 2010, 27(2): 94-99
- [6] 邓聚龙. 灰色预测与决策[M]. 武汉: 华中工业出版社, 1986
- [7] 中华人民共和国国家统计局. 中国统计年鉴(2009)[M]. 北京: 中国统计出版社, 2009
- [8] 刘思峰, 党耀国, 方志耕. 灰色系统理论及其应用[M]. 北京: 科学出版社, 2004
- [9] 白先春, 李炳俊. 基于新陈代谢 GM(1, 1)模型的我国人口城市化水平分析[J]. 统计与决策, 2006, 22(5): 62-65

Forecast of the Level of China's Urbanization Based on GM(1, 1)-Markov Model

SHI Liu-jie, LI Yan-jun, ZANG Yu-ting, DAI Xiaoming, ZHANG Wu-na

(College of Mathematics and Physics Anyang Institute of Technology, Anyang 455000, China)

Abstract By combining the advantages of both Gray prediction and Markov prediction chain theory, a new Gray Markov metabolism GM(1, 1) model is proposed. The forecasting result of the level of China's urbanization from 2009 to 2015 by this method shows that this method is practical and scientific.

Keywords GM(1, 1) model; Markov chains; urbanization; forecast; Matlab