第 26 卷第 6 期　　四川理工学院学报(自然科学版)　　Vol. 26　No. 6
2013 年 12 月　　Journal of Sichuan University of Science & Engineering(Natural Science Edition)　　Dec. 2013
文章编号:1673-1549(2013)06-0093-04　　DOI:10.3969/j.issn.1673-1549.2013.06.022

入境旅游发展的灰色马尔可夫预测模型研究

陈杏莉, 廖文杰

(苏州工业职业技术学院, 江苏　苏州　215104)

摘　要:旅游业是城市服务业的支柱产业,对国民经济和社会发展的贡献显著,入境旅游发展预测具有重要的经济价值和社会意义。将 GM(1,1)灰色系统理论和马尔可夫理论相结合,建立灰色马尔可夫数学模型。根据苏州市入境旅游外汇收入历史数据进行预测,结果显示,灰色马尔可夫预测模型精度高,预测效果好,具备良好的参考价值。

关键词:灰色预测;入境旅游;GM(1,1)模型;马尔可夫链

中图分类号:O21　　**文献标志码**:A

引　言

根据世界旅游组织预测,自 2010 年到 2020 年,世界旅游经济总量预计将保持 4.4% 的年增长率。到 2020 年,国际旅游人数将达到 16 亿人次,中国将成为世界第一大旅游目的地[1]。苏州,作为具有一定国际知名度的旅游目的地将直接面临机遇和挑战,预测未来 3 ~ 5 年的入境旅游发展情况,对于客观分析苏州在国际旅游市场的战略部署、合理调整阶段性发展策略、保障苏州旅游的国际化发展目标具有重要的实践意义。文献资料显示,旅游预测的方法主要有灰色系统模型[2]、ARIMA 模型[3]、IOWHA 算子组合模型[4]、SSVR 模型[5]及灰色神经网络组合模型[6]等,各种预测模型在实际中都有适用范围,选择合适的预测模型可以提高预测精度。

灰色系统理论对原始数据进行处理,变成有规律的时间序列数据,采用以数据寻找数据的方法建立动态模型。在灰色预测模型中,单序列一阶线性微分方程模型 GM(1,1)是常用的模型,但如果数据有较大波动性,预测精度降低,所以要对模型进行改进。本文建立灰色马尔可夫预测模型,通过采集苏州市近年来旅游外汇收入数据,在灰色预测 GM(1,1)预测值的基础上进行精度检验,采用马尔可夫预测理论优化模型并进行二次预测,然后与实际值作精确度比较,检验预测效果。

1　灰色马尔可夫预测模型理论

1.1　GM(1,1)预测模型

设初始时间序列 $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots x^{(0)}(n))$, n 为数据个数,将原始数据累加弱化随机序列的波动性,得到符合灰指数率的新数据序列

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \cdots x^{(1)}(n))$$

其中,

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

对 $x^{(1)}(t)$ 建立一阶线性微分方程:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \tag{1}$$

其中 a, b 为待定系数,在模型中分别称为发展系数和灰色作用量,求出参数 a, b, 就能求出 $x^{(1)}(t)$, 再求出 $x^{(0)}$ 的预测值。令灰参数 $\hat{a} = \begin{pmatrix} a \\ b \end{pmatrix}$, 由最小二乘法得到

收稿日期:2013-09-12

基金项目:江苏省社科联基金项目(13SWC-052)

作者简介:陈杏莉(1978-),女,江苏南通人,讲师,硕士,主要从事概率论与数理统计方面的研究,(E-mail)chenxingl@siit.edu.cn

$$\hat{a} = (B^T B)^{-1} B^T Y_n \quad (2)$$

其中,

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

$$Y_n = (x^{(0)}(2), x^{(0)}(3), \cdots, x^{(0)}(n))^T$$

将 $\hat{a} = \begin{pmatrix} a \\ b \end{pmatrix}$ 代入(1)式,得出微分方程解的近似表达式

$$\hat{x}^{(1)}(t+1) = (x^{(0)}(1) - \frac{b}{a})e^{-at} + \frac{b}{a} \quad (3)$$

原序列的近似表达式通过 $\hat{x}^{(1)}(t+1)$ 和 $\hat{x}^{(1)}(t)$ 的离散值还原,即

$$\hat{x}^{(0)}(t+1) = \hat{x}^{(1)}(t+1) - \hat{x}^{(1)}(t) \quad (4)$$

计算相对误差

$$m(t) = \frac{e^{(0)}(t)}{x^{(0)}(t)} = \frac{x^{(0)}(t) - \hat{x}^{(0)}(t)}{x^{(0)}(t)}$$

进行残差检验、关联度检验及后验差检验,测试模型精度。

1.2 马尔可夫链

GM(1,1)预测模型适用于短期性、随机性和波动性较弱的数据序列预测,若数据序列表现出强随机波动性时,可以在 GM(1,1)预测模型的基础上建立马尔可夫预测模型。

设 $\{\xi_n, n = 1,2,\cdots\}$ 是一个随机序列,状态空间 E 为有限集或可列集,对于任意的正整数 m, n, 若 i, j, $i_k \in E(k = 1,2,\cdots,n-1)$, 有 $P(\xi_{n+m} = j | \xi_n = i, \xi_{n-1} = i_{n-1}, \cdots, \xi_1 = i_1) = P(\xi_{n+m} = j | \xi_n = i)$, 则称 $\{\xi_n, n = 1, 2, \cdots\}$ 为一个马尔可夫链。它的含义是系统在已知现在情况的条件之下,系统未来时刻的情况只与现在有关,而与过去的历史无直接关系[7]。

如果条件概率 $P(\xi_{n+m} = j | \xi_n = i) = p_{ij}(m)$ 与 n 无关,则称 $\{\xi_n, n = 1,2,\cdots\}$ 为齐次的马氏链,称 $p_{ij}(m)$ 为系统由状态 i 经过 m 个时间间隔(或 m 步)转移到状态 j 的转移概率。对于一个马氏链,称以 m 步转移概率 $p_{ij}(m)$ 为元素的矩阵 $P(m) = (p_{ij}(m))$ 为马氏链的 m 步转移矩阵。当 $m = 1$ 时,记 $P(1) = P$ 称为马尔可夫链的一步转移矩阵。

根据柯尔莫哥洛夫—开普曼定理,m 步转移概率矩阵 $P(m) = P^m$ 。

1.3 灰色马尔可夫预测模型

在 GM(1,1)预测模型的基础上,根据原始数据和预测值求出残差和相对残差序列,将相对残差序列划分成 n 种独立的状态,任意一种状态可以表示成: $Q_i = [Q_{1i}, Q_{2i}]$, $i = 1,2,3,\cdots,n$, 其中 Q_{1i}、Q_{2i} 称为灰元,且 $Q_{1i} = \hat{x}(t) + A_i$, $Q_{2i} = \hat{x}(t) + B_i$, A_i, B_i 是根据具体问题确定的常数, Q_{1i}, Q_{2i} 随时间变化,故状态 Q_i 具有动态特征。

设 $M_{ij}(k)$ 表示原始数据序列由状态 Q_i 经过 k 步转移到状态 Q_j 的次数, M_i 表示原始数据序列中数据落入状态 Q_i 的次数。则从状态 Q_i 经过 k 步转移到状态 Q_j 的状态转移概率 $p_{ij}(k)$ 可以表示为:

$$p_{ij}(k) = \frac{M_{ij}(k)}{M_i} i = 1,2,\cdots,n \quad (5)$$

状态转移矩阵

$$P(k) = \begin{bmatrix} p_{11}(k) \cdots p_{1n}(k) \\ \vdots \qquad\qquad \vdots \\ p_{n1}(k) \cdots p_{nn}(k) \end{bmatrix} \quad (6)$$

计算状态转移概率矩阵,一般只考虑一步转移矩阵 $P(1)$ 。若预测对象目前在状态 Q_i, 考虑一步转移矩阵第 i 行,若第 j 列是此行中最大值,则下一时刻最有可能转向状态 Q_j; 若第 i 行中有两个或两个以上相同概率值,需继续考虑 m 步转移概率矩阵 $P(m) = P^m$ $(m \geqslant 2)$ 确定。当转移状态确定后,预测值范围确定,二次预测值可以取:

$$\tilde{x}(t) = \frac{1}{2}(Q_{1i} + Q_{2i}) = \hat{x}(t) + \frac{1}{2}(A_i + B_i) \quad (7)$$

2 苏州入境旅游发展预测模型

2.1 入境旅游发展 GM(1,1)预测模型建立

以苏州市 2004 ~ 2012 年入境旅游外汇收入为基础数据(表 1),建立 GM(1,1)预测模型:由(2)式得

$$\hat{a} = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -0.132853602475478 \\ 5.388688714263504 \end{pmatrix}$$

由(3)式求

$$\hat{x}^{(1)}(t+1) = 45.4211 e^{0.132853602475478t} - 40.5611$$

$$t = 0,1,2,\cdots,9$$

由(4)式求

$$\hat{x}^{(0)}(t+1) = 5.650690638e^{0.132853602475478t}$$

计算结果见表 2。

表 1　苏州市入境旅游外汇收入历史数据(人民币:亿元)

年份	2004	2005	2006	2007	2008	2009	2010	2011	2012
$x^{(0)}$ 实际收入	4.86	6.39	7.48	8.89	9.95	9.97	12.51	14.70	16.47

表 2　苏州市入境旅游外汇收入 GM(1,1)模拟值(人民币:亿元)

年份	GM(1,1)模拟值 $\hat{x}^{(0)}$	年份	GM(1,1)模拟值 $\hat{x}^{(0)}$
2004	4.86	2009	10.98
2005	6.45	2010	12.53
2006	7.37	2011	14.33
2007	8.42	2012	16.36
2008	9.62	2013	18.68

对模型精度进行分析,计算平均相对误差、关联度、均方差比值和小误差概率,结果见表 3。

表 3　苏州市入境旅游外汇收入 GM(1,1)精度检验

检验项目	平均相对误差 $\bar{\varphi}$(%)	关联度 r	均方差比值 C	小误差概率 P
检验数据	2.73	0.724	0.084106652	1

检验数据显示,残差检验、关联度检验、后验差检验均通过,可以继续建立马尔可夫预测模型。

2.2　入境旅游发展灰色马尔可夫预测模型建立

根据 GM(1,1)模型预测值与真实值的相对误差划分状态,以 -2%,0,2%,4%,6% 为分界点,可以划分五种状态,分别是状态 1 为 -2% 以下,状态 2 为(-2%,0],状态 3 为(0,2%],状态 4 为(2%,4%],状态 5 为(4%,6%],各年份所处状态见表 4。并据此得到状态转移情况(表 5)。

表 4　苏州市入境旅游外汇收入灰色马尔可夫模拟分析

年份	实际收入(亿元)	GM(1,1)预测值(亿元)	相对误差 %	状态	马尔可夫二次预测值(亿元)	相对误差 %
2004	4.86	4.86	0	2	4.86	0
2005	6.39	6.45	-0.939	2	6.39	0
2006	7.48	7.37	1.471	3	7.45	0.40
2007	8.89	8.42	5.287	5	8.86	0.34
2008	9.95	9.62	3.317	4	9.92	0.30
2009	9.97	10.98	-10.130	1	10.33	-3.61
2010	12.51	12.53	-0.160	2	12.41	0.80
2011	14.70	14.33	2.517	4	14.78	-0.54
2012	16.47	16.35	0.729	3	16.52	-0.30

考察状态转移概率矩阵,由表 5 的结论和(5)式、(6)式,可以得出苏州近年外汇收入量的 1 步状态转移概率矩阵

表 5　各年份状态转移情况

	状态 1	状态 2	状态 3	状态 4	状态 5
状态 1	0	1	0	0	0
状态 2	0	1	1	1	0
状态 3	0	0	0	0	1
状态 4	1	0	1	0	0
状态 5	0	0	0	1	0

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

2012 年处于状态 3,考察一步概率转移矩阵第三行,第 5 列出现此行最大概率,所以未来 2013 年状态转向最可能是状态 5。根据 GM(1,1)模型预测数据显示,2013 年预测值为 18.68,通过灰色马尔可夫预测模型修正,2013 年预测区间为[19.46,19.87),最终二次预测值 $\hat{x}(t)$ 可由(7)式取上述区间中点,即

$$\hat{x}(t) = \frac{1}{2}(19.46 + 19.87) =$$

$$18.68 + \frac{1}{2}(0.78 + 1.19) = 19.67$$

按照相同的方法可以计算 2004 至 2012 各个年份的马尔可夫二次模拟情况,结果见表 4 第 6、7 列。除了受到 2008 年金融危机影响,2009 年的预测数据相对误差较大外,二次模拟平均相对误差为 0.70%,对比 GM(1,1)模型平均相对误差 2.73%,预测精度显著提高。

3　结束语

本文在灰色理论和马尔可夫理论的基础上,建立灰色马尔可夫预测模型,对苏州入境旅游外汇收入问题进行分析。实证研究结果显示,灰色马尔可夫预测模型结合了灰色预测模型和马尔可夫预测模型的优点,对随机性和波动性较大的数据序列具有较高的预测精度。该模型还可预测入境旅游人次等旅游发展预测指标,因此在旅游发展预测中具有较高的应用价值。

参 考 文 献:

[1] 苏州市旅游业发展"十二五"规划[DB/OL]. http://www.visitsz.com/gov/info/201110241319440402.html, 2011-08-16/2011-10-24.

[2] 唐晓云,赵黎明,秦 彬.灰色系统理论及其在旅游预测中的应用[J].西安电子科技大学学报,2007,17(2):1-5.

[3] 陈 鹏,吴 玲,宋 徽.基于 ARIMA 模型的安徽省入境旅游人数预测[J].安徽农业大学学报,2012,21(1):32-35.

[4] 杨 洋,李东和,靳 非.基于 IOWHA 算子组合模型的旅游总收入预测—以安徽省旅游总收入为例[J].乐山师范学院学报,2011,26(12):42-45.

[5] 陈 荣,梁昌勇.基于季节 SVR 的节假日旅游客流量预测[J].统计与决策,2013(9):82-84.

[6] 杨 盼.基于灰色系统和神经网络的旅游需求预测[D].江西:东华理工大学,2012.

[7] 袁荫棠.经济应用数学基础三:概率论与数理统计[M].2 版.北京:中国人民大学出版社,1990.

Study on Grey Markov Prediction Model of the Inbound Tourism Development

CHEN Xing-li, *LIAO Wen-jie*

(Suzhou Industrial Institute of Vocational Technology, Suzhou 215104, China)

Abstract: Tourism is a pillar industry of city services industry. It is significant to the national economy and social development. The prediction of inbound tourism development has important economic value and social significance. GM (1, 1) grey system theory is combined with Markov theory, and the gray Markov mathematical model is established. According to the history data of Suzhou municipal tourism foreign exchange earnings, the grey Markov prediction model is of high precision, good predictive ability, and has a good reference value.

Key words: grey prediction; inbound tourism; GM(1,1) model; Markov chain