文章编号:1672 - 058X(2012)01 - 0024 - 06

改进的灰色马尔可夫组合预测模型应用研究*

周廷慰[1], 周宗福[2]

(1. 蚌埠学院 数学与物理系,安徽 蚌埠 233030;2. 安徽大学 数学科学学院,合肥 230039)

摘　要:根据灰色预测模型和马尔可夫预测思想,将灰色预测模型结合马尔可夫预测模型,并对组合模型进行了改进,采用递进转移概率矩阵,并结合 1998-2007 年安徽入境游客人数来预测 2008、2009 年的入境游客人数;结果表明:采用该改进后灰色马尔可夫的组合模型,可以得到比传统灰色马尔可夫组合预测模型更为准确的数据。

关键词:灰色马尔可夫组合预测;入境旅游;安徽省

中图分类号:O211.62　　　　**文献标志码**:A

1　灰色马尔可夫组合预测模型

把灰色模型与马尔可夫模型结合后,就得到灰色马尔可夫组合模型,模型克服了灰色模型与马尔可夫模型各自的缺点,不仅反映了数据序列的发展趋势,而且还通过状态转移概率矩阵的变换,提取出数据中的随机响应,因此预测的适应性更强,预测精度也更高。

1.1　模型的原理和方法

灰色预测模型主要适用于预测时间短,数据资料少,波动不大的系统对象,由于其自身的缺陷,使其仅能适用于原始数据序列按指数规律变化的系统[1,2],而马尔柯夫链预测模型适用于预测随机波动大的动态过程,在这一点上可以弥补灰色预测模型的缺陷。旅游业是一个波动极大的行业,2001 年的“9.11”事件、2003 年“非典”事件都对安徽入境旅游产生了较大的影响,因此仅使用单一的灰色预测方法来进行预测结果并不十分满意,为了提高预测的精度,考虑将灰色预测模型与马尔柯夫链预测模型进行组合,建立起灰色马尔可夫组合模型,从而实现功能互补,突出优势,使得预测精度提高[3]。

一般的灰色马尔可夫组合模型首先根据灰色 GM(1, 1)模型求出相应的拟合值 $\hat{x}^{(0)}(k)$,然后求出残差 $\Delta(k)=x^{(0)}(k)-\hat{x}^{(0)}(k)$,得到残差相对值为:

$$e^{(0)}(k) = \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \times 100\ \%$$

然后将残差相对值根据问题的需要和原始数据的多少,将其进行状态的划分,尽可能的使每一个划分出的状态区间内的数据都差不多。一般来说,当原始的数据较多的时候,不妨多划分一些区间,这样可以从原始信息中获取到更多的信息,从而提高预测的精度。但当数据较少时,划分的区间则不宜过多,以便于增多各个状态之间的转移次数,从而提高预测的精度[4,5]。

收稿日期:2011 - 04 - 20;**修回日期**:2011 - 05 - 10.

* **基金项目**:国家自然科学基金(10771001).

作者简介:周廷慰(1979-),男,安徽蚌埠人,讲师,硕士研究生,从事灰色理论研究.

根据上面的状态划分，可以将残差相对值分为若干状态，将其记为 $E_1, E_2, \cdots, E_m$，把 n 步转移概率定义为残差相对值序列由状态 E_i 经 n 步转移到状态 E_j 的概率，记为 $P_{ij}(n) = \frac{M_{ij}}{M_i}, i, j = 1, 2, \cdots, m$. 其中的 M_i 为状态 E_i 出现的次数，M_{ij}为状态 E_i 经 n 步转移到状态 E_j 的的次数，在计算 n 步转移概率时，由于数据序列的最后状态转向不确定，所以在计数时要把数据序列的最后 n 个数据去掉。需要注意的是当在求 n 步状态转移矩阵时，计数 M_i 时就把最后 n 个数据去掉，但在记数 M_{ij}时，则不要把数据序列的最后 n 个数据去掉。

这样就可以得到 $m \times m$ 阶的 n 步状态转移概率矩阵：

$$P_n = \begin{pmatrix} P_{11}(n) & P_{12}(n) & \cdots & P_{1m}(n) \\ P_{21}(n) & P_{22}(n) & \cdots & P_{2m}(n) \\ \vdots & \vdots & \vdots & \vdots \\ P_{m1}(n) & P_{m2}(n) & \cdots & P_{mm}(n) \end{pmatrix}, \text{其中} \sum_{j=1}^{m} P_{ij} = 1$$

由上可知 $P(1) = P(0)P_1$，若在 $P(1)$ 的 m 个值中，$\max\limits_{j} P_{kj} = P_{kl}$，那么就认为下一个时刻序列将从状态 E_k 转向状态 E_l，即下个时刻系统最有可能处于状态 E_l。

根据 $P(2) = P(1)P_1, \cdots, P(n) = P(n-1)P_1$，这样就可以知道 n 步后系统最有可能所处的状态。

下面通过考察状态转移矩阵来编制可以预测系统未来变换发展的预测表。由于状态转移矩阵反映了系统中各种状态转移的统计规律性，所以在编制预测表时，首先选取离预测时刻较近的 n 个节点，按离预测节点的远近，确定转移步数为 $1, 2, \cdots, n$。在转移步数所对应的转移矩阵中，选取初始状态所对应的行向量，将它们组成新的概率矩阵，然后将其列向量求和，最后在加权平均求出概率的期望值 δ。最后利用概率的期望值 δ 及公式 $\mu'^{(0)}_{k+1} = \hat{\mu}^{(0)}_{k+1} \div (1-\delta)$ 对灰色预测的结果进行修正。就可以得到利用灰色马尔可夫组合模型所获得的预测值了。

1.2　灰色马尔可夫组合预测模型的改进

现有文献所见的灰色马尔可夫模型都是采用有限步 n 的转移矩阵，一般的 n 都不大（常为 1 或 2），但发现当预测较远节点的状态时，已发生的很久之前的节点其数据状态对预测数据的精度影响是不大的，但现有文献中所采用的数据都是固定不变的，这就会影响到预测的精度。所以考虑采用递进转移概率矩阵来进行预测。所谓的递进转移概率矩阵就是去掉已发生的最远节点的状态数据，然后补充进去最新预测节点的数据状态，这样就可以形成一个新的 1 步状态转移概率矩阵，从而就可以以最新的预测节点的数据状态作为初始分布，再来接着预测下一个节点的数据状态，这样就可以充分利用到各个状态之间的内在规律性，充分反映数据发展的趋势，提高预测的精度。

2　模型的建立

依据 1998-2007 年安徽入境游客人数及外汇收入的原始数据，来预测 2008、2009 年的相关数据。根据表中的 1998-2007 年的实际入境人数 $x^{(0)}$，利用相关知识，可以建立起基于 1998-2007 年的实际入境人数 $x^{(0)}$ 的灰色(1，1)模型。这样就可以依据模型拟合出 1998-2007 年的预测入境人数 $\hat{x}^{(0)}$，见表 1：

可以发现模型的入境预测人数值 $\hat{x}^{(0)}$ 与实际入境人数 $x^{(0)}$ 的相对残差 $e^{(0)}(k)$ 的范围为（$-56\%, 25\%$），考虑到相对残差 $e^{(0)}(k)$ 在 0 的左测一边较多，且为了保证划分后的各个状态中数据的数量基本保持一致，将序列根据相对残差从小到大的顺序共划分为 3 个状态，前 3 个数据划为第一状态，中间 4 个数据为第二状态，后 3 个数据为第三状态。入境游客人数的 3 个状态分别用 N_1, N_2, N_3 来表示。划分后的 3 个状态的区间分别为 $N_1 = (-56\%, 0\%]$，$N_2 = (0\%, 20\%]$，$N_3 = (20\%, 25\%]$，各年相对残差 $e^{(0)}(k)$ 所处具体状态见表 1。

表 1　1998-2007 年安徽入境旅游人数相关数据

序号	年份	实际入境人数 $x^{(0)}$	预测入境人数 $\hat{x}^{(0)}$	人数 $e^{(0)}(k)$	状态
1	1998	184 198	184 198	0	N_2
2	1999	251 183	204 370	18.63	N_2
3	2000	318 430	247 276	22.35	N_3
4	2001	380 852	299 189	21.44	N_3
5	2002	459 051	362 001	21.14	N_3
6	2003	280 819	438 000	−55.97	N_1
7	2004	500 966	529 953	−5.79	N_1
8	2005	632 895	641 213	−1.31	N_1
9	2006	803 722	775 829	3.47	N_2
10	2007	1 064 296	938 707	11.80	N_2

通过上述状态的划分可以得到入境游客人数状态转移概率矩阵：

入境游客人数 1 步状态转移概率矩阵：

$$P_1^N = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}$$

注意到 2007 年的入境旅游人数处在状态 N_2，那么 2008 年的绝对分布为：

$$P(1) = (p_0(1), p_0(2), \cdots, p_0(n))P_1 = (0,1,0)\begin{pmatrix} \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix} = \left(0, \frac{2}{3}, \frac{1}{3}\right)$$

所以 2008 年安徽入境旅游人数最有可能处于状态 N_2，即有 $e^{(0)}(k) \in \mathbf{N}_2 = (0\%, 20\%]$，这样就可以求出 2008 年的平均相对残差：

$$\bar{e}^{(0)}(k) = \frac{2}{3} \times \frac{0\% + 20\%}{2} + \frac{1}{3} \times \frac{20\% + 25\%}{2} = 0.142$$

又由 $e^{(0)}(k) = \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \times 100\%$，得 $x^{(0)}(k) = \frac{\hat{x}^{(0)}(k)}{1 - e^{(0)}(k)}$，根据 1998-2007 年安徽入境旅游人数所建立的 GM(1,1) 模型，可以得到 2008 年的入境旅游预测人数为 1 135 779，所以需要预测的 2008 年的安徽入境旅游人数可以根据 $x^{(0)}(k) = \frac{\hat{x}^{(0)}(k)}{1 - e^{(0)}(k)}$，对由 GM(1,1) 模型所得到预测人数 1 135 779 进行修正得到灰色马尔可夫模型下为 1 323 752 人次。而 2008 年的安徽入境旅游实际人数为 1 320 947 人次。可以很明显看到组合预测要比单独的灰色模型预测要精确许多。

现在来预测 2009 年安徽入境旅游人数。首先使用常规文献所见的一般的灰色马尔可夫模型来进行预测，当应用马尔可夫模型来预测 2009 年的入境旅游人数时，依然以 2007 年为初始年份，由于距离 2009 年时隔 2 年，所以采用两步概率转移矩阵，其具体为：

$$P_2^N=\begin{pmatrix}\frac{1}{3}&\frac{2}{3}&0\\0&0&1\\\frac{2}{3}&0&\frac{1}{3}\end{pmatrix}$$

注意到 2007 年的入境旅游人数处在状态 N_2,那么 2009 年的绝对分布为:

$$P(2)=(p_0(1),p_0(2),\cdots,p_0(n))P_2=(0,1,0)\begin{pmatrix}\frac{1}{3}&\frac{2}{3}&0\\0&0&1\\\frac{2}{3}&0&\frac{1}{3}\end{pmatrix}=(0,0,1)$$

从上式知道 2009 年的状态应该为 N_3,可以求出 2009 年的平均相对残差为 $\bar{e}^{(0)}(k)=\frac{20\%+25\%}{2}=0.225$,根据 1998-2007 年安徽入境旅游人数相关数据所建立的 GM(1,1)模型,可以得到 2009 年的入境旅游预测人数为 1 374 225。根据 2009 年的平均相对残差及公式 $x^{(0)}(k)=\frac{\hat{x}^{(0)}(k)}{1-e^{(0)}(k)}$,可以得到经过马尔可夫模型修正的 2009 年的入境旅游预测人数为 1 773 194。

下面用改进的马尔可夫模型来进行预测 2009 年的入境旅游人数,注意到 2008 年数据以 $\frac{2}{3}$ 的概率处于状态 N_2,$\frac{1}{3}$ 的概率处于状态 N_3,所以考虑采用递进转移概率矩阵来进行预测。去掉已发生的最远节点即 1998 年的数据状态,然后补充进最新的预测节点即 2008 年的数据状态,这样就可以形成一个新的 1 步概率转移矩阵,从而可以最新的 2008 年数据作为初始分布,再来接着预测下一个节点即 2009 年的数据状态,由于 2008 年的数据状态有两种可能,所以下面分别给出 2008 年数据状态处于状态 N_2 和 N_3 时的 1 步概率转移矩阵:

$$\tilde{P}_1^{N_2}=\begin{pmatrix}\frac{2}{3}&\frac{1}{3}&0\\0&\frac{2}{3}&\frac{1}{3}\\\frac{1}{3}&0&\frac{2}{3}\end{pmatrix},\qquad \tilde{P}_1^{N_3}=\begin{pmatrix}\frac{2}{3}&\frac{1}{3}&0\\0&\frac{1}{3}&\frac{2}{3}\\\frac{1}{3}&0&\frac{2}{3}\end{pmatrix}$$

所以 2009 年的绝对分布为:

$$P(1)=\frac{2}{3}(0,1,0)\begin{pmatrix}\frac{2}{3}&\frac{1}{3}&0\\0&\frac{2}{3}&\frac{1}{3}\\\frac{1}{3}&0&\frac{2}{3}\end{pmatrix}+\frac{1}{3}(0,0,1)\begin{pmatrix}\frac{2}{3}&\frac{1}{3}&0\\0&\frac{1}{3}&\frac{2}{3}\\\frac{1}{3}&0&\frac{2}{3}\end{pmatrix}=\left(\frac{1}{9},\frac{4}{9},\frac{4}{9}\right)$$

这样可以求出 2009 年的平均相对残差为:

$$\bar{e}^{(0)}(k)=\frac{1}{9}\times\frac{-56\%+0\%}{2}+\frac{4}{9}\times\frac{0\%+20\%}{2}+\frac{4}{9}\times\frac{20\%+25\%}{2}=0.113$$

根据 1998-2007 年安徽入境旅游人数相关数据所建立的 GM(1,1)模型,可以得到 2009 年的入境旅游预测人数为 1 374 225。根据上述 2009 年的平均相对残差及公式 $x^{(0)}(k)=\frac{\hat{x}^{(0)}(k)}{1-e^{(0)}(k)}$,可以得到经过改进的马尔可夫模型修正的 2009 年的入境旅游预测人数为 1 549 295。

这样就可以得到下列这张安徽入境旅游人数预测表,见表2:

表2 2008-2009年安徽入境旅游人数预测数据

年份	实际入境人数	GM(1,1)		一般灰色马尔可夫模型		改进的灰色马尔可夫模型	
		预测值	预测精度	预测值	预测精度	预测值	预测精度
2008	1 320 947	1 135 779	85.98%	1 323 752	99.79%	1 323 752	
2009	1 561 600	1 374 225	88.00%	1 773 194	86.45%	1 549 295	99.21%

3 模型的检验与分析

在表2中由2008年的预测数据可以得出,针对安徽入境旅游人数的原始数据,灰色马尔可夫组合预测的精度会大大高于单一的灰色预测方法。而从2009年的预测数据中又可以发现随着预测节点的渐行渐远,单一的灰色预测和一般的灰色马尔可夫预测距离2009年的实际入境人数误差都较大,而改进的灰色马尔可夫预测模型则取得了非常满意的效果,达到了较高的预测精度。

灰色GM(1.1)模型实际上是以指数型曲线去拟合原始数据,所以用预测值做出的图形是一条光滑的指数曲线,因此只有当历史数据呈现稳定单调的趋势,并且波动不大时,模型的预测才会达到较高的精度,并且一般来说处理的数据也不宜过多,以5-10个数据为宜[6]。就选用了10个数据,用灰色GM(1.1)模型预测曲线来反映数据列的发展趋势。但是这10个数据呈现出较大的波动性,利用传统的灰色GM(1.1)进行预测达不到满意的预测精度。因此利用马尔可夫模型来对灰色预测进行修正,充分利用到马尔可夫模型能够反映数据列的波动规律。注意到当预测较远节点的状态时,已发生的很久之前的节点其数据状态对预测数据的精度影响是不大的,所以考虑采用递进转移概率矩阵来进行预测。去掉已发生的最远节点的状态数据,然后补充进去最新预测节点的数据状态,这样就可以形成一个新的1步状态转移概率矩阵,从而就可以以最新的预测节点的数据状态作为初始分布,再来接着预测下一个节点的数据状态。另外在状态划分上,虽然说划分的状态数越多,预测的精度越高,但划分过多也会导致各状态内样本点过少,造成转移概率所反映出的规律不强,因此将状态划分为3种状态。这样就可以充分利用到各个状态之间的内在规律性,充分反映数据发展的趋势,这样即使历史数据发生较大的波动时,往往也能达到非常满意的预测精度。从所得到的结果来看,改进后的灰色马尔可夫模型非常适合安徽省入境旅游人数的预测。

参考文献:

[1] 邓聚龙.灰理论基础[M].武汉:华中科技大学出版社,2002

[2] 刘思峰.灰色系统理论及其应用[M].北京:科学出版社,2010

[3] 陈萍,张玉红.重庆直辖十年入境旅游市场特征分析及趋势预测[J].重庆工商大学学报,2007,7(3):49-51

[4] 黄银珠.基于灰色模型的福建省入境旅游客源预测[J].北京第二外国语学院学报 2009,21(8):90-92

[5] 黄芳,黎洁,连传鹏.西安市入境旅游市场特征及趋势预测研究[J].西北农林科技大学学报,2008,34(3):86-91

[6] SONG H, LI G. Tourism demand modeling and forecasting — A review of recent research [J]. Tourism Management,2008,29(2):203-220

Application of Improved Grey Markov Combination Predication Model

ZHOU Ting-wei[1], **ZHOU Zong-fu**[2]

(1. Department of Mathematics and Physics, Bengbu University, Anhui Bengbu 233030, China;
2. School of Mathematical Science, Anhui University, Hefei 230039, China)

Abstract: According to grey predication model and Markov predication model ideas, the grey predication model is integrated with Markov predication model, this combination model is improved. By using progressive transition probability matrix, the inbound tourists number to Anhui in 2008-2009 is predicated on the basis of Anhui inbound tourists number during 1998-2007 by the models, and the results show that application of improved grey Markov combination predication model can obtain more accurate data than application of traditional grey Markov combination predication model.

Key words: grey Markov combination predication model; inbound tourism; Anhui Province

责任编辑:田 静

(上接第 23 页)

Compensation Strategy Theory for Wholesale Price Dependence under Supply Chain Management

CHEN Xiu-su[1,2], **CHEN Rui**[3]

(1. Chongqing Municipal Key Laboratory for E-commerce and Supply Chain System, Chongqing Technology and Business University, Chongqing 400067, China;
2. School of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China;
3. School of Electrical Engineering, Chongqing University, Chongqing 400044, China)

Abstract: Compensation strategy theory for wholesale price dependence under supply chain management was studied by He Ju, Zhou Jing and He Yong [1], but their real expression of deduction and conclusion was not completely correct. This paper gives a correct expression for their conclusion. Under the condition that wholesale price is related to purchase quantity of a product and the market demand quantity for the product is stochastic variable, compensation strategy theory model for wholesale price dependence under supply chain management is obtained through discussing compensation strategy influenced by pricing mechanism of wholesale price of a product, furthermore, equilibrium analysis and sensitivity analysis are conducted on this theoretical model under the condition that market demand quantity for the product follows homogeneous distribution, finally, the intrinsic relationship between purchase quantity of a product, compensation price and the product demand as well as wholesale price is pointed out. This paper provides helpful analysis method for enterprises to make decision.

Key words: supply chain management; compensation strategy; equilibrium analysis; wholesale price

责任编辑:田 静