文章编号:1005-0523(2013)04-0019-05

基于多次修正残差灰色模型的铁路客流预测

骆 晨,刘 澜

(西南交通大学交通运输与物流学院,四川 成都 610031)

摘要: 为了提高全国铁路客流预测精度,针对铁路客流变化的影响因素及特点,对$GM(1,1)$模型进行改进,提出了多次修正残差灰色模型的铁路客流预测方法。克服了传统的$GM(1,1)$模型的指数函数预测精度较差、单次修正残差精度不够的缺点,通过多次残差修正,减少预测误差。结合2003—2010年铁路客运数据实例分析。结果证明该方法预测误差小、精度高、计算简便、可操作性强,为铁路客流预测提供了一种更为可行的途径。

关键词: 灰色理论;修正残差灰色;客流预测;精度

中图分类号: U491.1+4 **文献标志码:** A

客流预测是指利用一定的方法和技术对未来一定时期内客流的需求、性质进行预先推测和判断[1]。客流预测一般分为短期预测和长期预测两种。长、短期预测方法主要采用基于时间序列的、基于现代数学算法的和结合客流特点的复合模型等方法研究客流数据随时间变化的规律[2]。目前,国内外学者针对客流预测展开了不同的探索与实践。霍保世等[3]运用Fuzzy集值统计聚类分析法对多因素的客流预测方案进行方案的归类,得出客流预测方案的区间。李振宇[4]等对指数平滑法进行改进,提出了基于交通流季节性调整的预测方法。焦永兰等[5]建立了基于改进灰色理论的影子响应$GM(1,1)$模型、Dynamic $GM(1,1)$模型、修饰型残差$GM(1,1)$预测模型用于预测全国铁路年度客货运量。马彦祥等[6]构建出一种涵盖平日客流量和特殊时段客流量的客流激发能级模型。Gaudry和Dagenias提出Dogit模型以及经过Box-Cox变换得到的BCL模型和BCD进行轨道交通客流客流预测。Stouffer提出了介入机会模型,摒弃传统模型中同时需要客运生成量和吸引量的方法,其减少预测条件。现有的研究结果显示,采用灰色模型预测客流量可以取得较好的预测效果。传统的$GM(1,1)$ 模型归类为指数函数模型,因其数据要求高而使得其远期预测精度较差,在实际预测应用中受到较大限制。针对单次修正$GM(1,1)$的残差而言,虽相对误差达到要求,但由于客运量基数较大,客运量预测误差达到几千万。为此,本文在使用灰色预测模型预测铁路客流量时,对$GM(1,1)$进行模型改进,通过多次修正$GM(1,1)$的残差,减少预测差值量,降低预测的绝对误差,提高预测精度。并通过实例证明了该模型具有较高的准确性。

1 灰色预测模型

灰色预测法主要针对在一系统中含有不确定因素而进行的预测的方法。根据已知的少量含有不确定信息的数据信息系统进行推测,该方法的特点是根据已有的数据建立动态微分方程,再合理预测自身发展。

1.1 模型的建立

灰色系统理论基础是通过模型的原始数据一次累加处理,使得其数据呈现明显的指数变化特点,因此,$GM(1,1)$模型的微分方程可表达为

收稿日期: 2013-05-25

作者简介: 骆晨(1989-),男,硕士研究生,研究方向为交通运输。

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

式中：a 表示 $GM(1,1)$ 模型的发展灰数；u 为 $GM(1,1)$ 模型的灰作用量。

1.2 模型的系数估计

在 $GM(1,1)$ 模型的原始数据中，为方便计算，首先假设模型中的数据序列为非负序列，$x^{(0)}=\left\{x_{(1)}^{(0)},x_{(2)}^{(0)},\cdots,x_{(n)}^{(0)}\right\}$，其中：$x_{(j)}^{(0)}\geqslant 0$，$i=1,2,\cdots,n$；原始序列 $x^{(0)}$，其中，初始数据的原始序列记为 $x^{(0)}$，经一次累加(1-AGO)得到生成序列 $x^{(1)}$，即

$x^{(1)}=\left\{x_{(1)}^{(1)},x_{(2)}^{(1)},\cdots,x_{(n)}^{(1)}\right\}$，其中：$x_{(k)}^{(1)}=\sum_{j=1}^{k}x_{(1)}^{(0)}$，$z^{(1)}$ 为 $x^{(1)}$ 的紧邻均值生成序列，即 $z^{(1)}=\{z_{(2)}^{(1)},z_{(3)}^{(1)},\cdots,z_{(n)}^{(1)}\}$，其中：$z_{(k)}^{(1)}=\frac{1}{2}(x_{(k)}^{(1)}+x_{(k-1)}^{(1)})$，11)，$k=1,2,\cdots,n$，则 $GM(1,1)$ 模的微分方程 $\frac{dx^{(1)}}{dt}+ax^{(1)}=u$；待估参数向量为 $\hat{\boldsymbol{A}}$，即 $\hat{\boldsymbol{A}}=[a,u]^{T}$，由最小二乘法估计可知[7]，即 $\hat{\boldsymbol{A}}$ 为

$$\hat{\boldsymbol{A}}=[a,u]^{\mathrm{T}}=\left[B^{\mathrm{T}}B\right]^{-1}B^{\mathrm{T}}Y \tag{1}$$

式中：B 为紧邻均值生成序列构成的参数向量；Y 为初始数据的原始序列构成的参数向量，即可表述为

$$\boldsymbol{B}=\begin{bmatrix}-Z_{(2)}^{(1)}\\-Z_{(3)}^{(1)}\\\vdots\\-Z_{(N)}^{(1)}\end{bmatrix}=\begin{bmatrix}-\frac{1}{2}\left[X_{(2)}^{(1)}+X_{(1)}^{(1)}\right] & 1\\-\frac{1}{2}\left[X_{(3)}^{(1)}+X_{(2)}^{(1)}\right] & 1\\\vdots & \vdots\\-\frac{1}{2}\left[X_{(N)}^{(1)}+X_{(N-1)}^{(1)}\right] & 1\end{bmatrix},\boldsymbol{Y}=\begin{bmatrix}X_{(2)}^{(0)}\\X_{(3)}^{(0)}\\\vdots\\X_{(N)}^{(0)}\end{bmatrix}$$

1.3 预测未来量

通过求解 $GM(1,1)$ 模型的微分方程，可得预测模型方程为

$$\hat{X}_{(k+1)}^{(1)}=(X_{(1)}^{(0)}-\frac{u}{a})e^{-ak}+\frac{u}{a},k=1,2,\cdots,n \tag{2}$$

由 $x_{(k+1)}^{(0)}=x_{(k+1)}^{(1)}-x_{(k)}^{(1)}$，通过一次数据累减，可得到初始数据的数据还原值[8]，即原始数据序列的灰色预测公式可以表达为

$$\begin{cases}\hat{x}_{(1)}^{(1)}=\hat{x}_{(1)}^{(0)}\\\hat{x}_{(k+1)}^{(0)}=\hat{x}_{(k+1)}^{(1)}-\hat{x}_{(k)}^{(0)}=(1-e^{a})(x_{(1)}^{(0)}-\frac{u}{a})e^{-ak},k=1,2,\cdots,n\end{cases} \tag{3}$$

2 修正的 $GM(1,1)$ 残差灰色预测模型

根据公式(3)可以得到一组预测数据 $\hat{x}_{(k)}^{(0)}=\left[\hat{x}_{(1)}^{(0)},\hat{x}_{(2)}^{(0)},\cdots,\hat{x}_{(n)}^{(0)}\right]$，预测模型具体运行步骤如下：

Step 1： 记 $e_{(k)}^{(0)}=x_{(k)}^{(0)}-\hat{x}_{(k)}^{(0)},k=1,2,\cdots,n$ 为残差序列；

Step 2： 对 $e_{(k)}^{(0)}$ 取部分子数列，其中分子数列在选择中，应该剔数据波动幅度较大的分子，排除数据噪音的存在，则重新排序有：$e_{(k')}^{(0)}=\left[e_{(1')}^{(0)},e_{(2')}^{(0)}\cdots,e_{(n')}^{(0)}\right]$；$e_{(k')}^{(0)}$ 建模型的数据具体要求如下：

1）若数列中所有数值均为非负数，则直接建立 $GM(1,1)$ 模型；

2）若数列中的所有数值为负数，则首先在不考虑符号前提下，建立 $GM(1,1)$ 模型，然后在最终结果的数据前再加上负号；

3）若数列中的数值有正数亦有负数，则按下述步骤进行：第一步，做非负处理：即对数列中所有数值都

加上最小负数的绝对值;第二步,建立 $GM(1,1)$ 模型;第三部,在求出最终反馈值后再减去所添加的最小负数的绝对值;

Step 3: 对 $e^{(0)}_{(k')}$ 建立GM(1,1)模型,其时间响应函数的离散形式为

$$\hat{e}^{(0)}_{(k+1')}=(1-e^{a'})\left[e^{(0)}_{(1')}-\frac{u'}{a'}\right]e^{-a'k'} \tag{4}$$

Step 4: 以 $\hat{e}^{(0)}_{(k+1')}$ 作为 $x^{(0)}_{(k+1)}$ 的修正模型可得:

$$\hat{x}^{(0)}_{(k+1)}=(1-e^{\hat{a}})\left[x^{(0)}_{(1)}-\frac{u'}{a'}\right]e^{-\hat{a}k}+\delta(k-i)(1-e^{a'})\left[e^{(0)}_{(1')}-\frac{u'}{a'}\right]e^{-a'k'}$$

式中:

$$\begin{cases} i=n-n' \\ \delta(k-1)=\begin{cases}1, & k\geqslant i\\ 0, & k<i\end{cases}\end{cases} \tag{5}$$

Step 5: (4)(5)所求的一次残差值如不满足精度要求,则带入Step 1中重复上述步骤进行二次残差计算,得到二次残差预测表格。若需进行多次修正残差,则需多次迭代[5]。

3 实例验证

根据第2部分给出的修正 $GM(1,1)$ 残差灰色预测模型的建模原理与步骤,以2003—2010年为原始数据预测2011年的客运量,如表1所示。

表1　2000—2010年我国铁路客运总量

Tab.1　Passenger flow volume during 2000—2010

亿人

年份	2003	2004	2005	2006	2007	2008	2009	2010
客流量	9.73	11.17	11.55	12.58	13.56	14.45	15.25	16.81

建模时,依据铁路客运量按年份顺序依次取1~8作为其序号,记其为

$X^{(0)}=\{9.73,11.17,11.55,12.58,13.56,14.45,15.25,16.81\}$,1-AGO序列为

$X^{(1)}=\{9.73,20.90,32.45,45.03,58.59,73.04,88.29,150.10\}$

由(1)式得:$\hat{\boldsymbol{A}}=[a,u]^{\mathrm{T}}=[-0.069\,3,9.906\,2]^{\mathrm{T}}$,由公式(3) $\hat{X}^{(0)}_{(k+1)}=10.62e^{-\hat{a}k}=10.62e^{0.069\,3k}$ 则

$\hat{X}^{(0)}=\{9.73,11.38,12.19,13.07,14.01,15.02,16.08,17.25\}$,其残差序列为

$e^{(0)}=\{0,-0.21,-0.64,-0.49,-0.45,-0.57,-0.83,-0.44\}$,如表2总结所示。

表2　灰色预测值

Tab.2　The gray predictions

序号	年份	初始客流量/亿人	预测客流量/亿人	相对误差	相对误差度/%
1	2003	9.73	9.73	0	0
2	2004	11.17	11.38	−0.21	−1.88
3	2005	11.55	12.19	−0.64	−5.54
4	2006	12.58	13.07	−0.49	−3.89
5	2007	13.56	14.01	−0.45	−3.31
6	2008	14.45	15.02	−0.57	−3.94
7	2009	15.25	16.08	−0.83	−5.44
8	2010	16.81	17.25	−0.44	−2.61

如表2所示，对所有年份数据进行一次残差计算，并对残差值进行排序得

$\hat{e}^{(0)}_{(k)}=\{-0.83,-0.64,-0.57,-0.49,-0.45,-0.44,-0.21,0\}$，先对该数列作正化处理得：

$\hat{e}1^{(0)}_{(k')}=\{0,0.21,0.44,0.45,0.49,0.57,0.64,0.83\}$，其1-AGO序列为：

$\hat{e}^{(1)}_{(1)}=\{0,0.21,0.65,1.10,1.59,2.16,2.80,3.63\}$对$e^{(1)}_{1}$进行$GM(1,1)$建模，来修正上面的原序列。

由(4)式得：$\hat{e}^{(0)}_{(k')}=\{0,0.365,0.477,0.681,0.494,0.569,0.813,0.343\}$将所得预测数代入(5)式得第一次残差预测值，如表3所示。

表3 一次残差灰色预测值

Tab.3 Predictions of the first residual error gray

序号	年份	初始客流量/亿人	预测客流量/亿人	相对误差	相对误差度/%
1	2003	9.73	9.73	0	0
2	2004	11.17	11.015	−0.155	−0.14
3	2005	11.55	11.710	−0.160	−1.38
4	2006	12.58	12.389	−0.191	1.52
5	2007	13.56	13.506	−0.054	−0.40
6	2008	14.45	14.731	−0.281	−1.94
7	2009	15.25	15.267	−0.017	0.11
8	2010	16.81	16.907	−0.097	0.58

对2005，2006两年的数据进行二次残差运算，将其残差进行排序得

$\hat{e}^{(0)}_{(k)}=\{-0.281,-0.16,-0.155,-0.054,-0.017,0,0.097,0.191,\}$先对该数列作正化处理得：$\hat{e}^{(0)}_{2(k)}=\{0,0.121,0.126,0.227,0.264,0.281,0.378,0.472\}$

由公式(4)可知，$\hat{e}^{(0)}_{(k')}=\{0.094,0.308\}$将所得预测数代入(5)式得，其二次残差预测值如表4

表4 二次残差灰色预测值

Tab.4 Predictions of the second residual error gray

序号	年份	初始客流量/亿人	预测客流量/亿人	相对误差	相对误差度/%
1	2003	9.73	9.73	0	0
2	2004	11.17	11.015	−0.055	−0.490
3	2005	11.55	11.616	0.066	0.570
4	2006	12.58	12.697	0.111	0.882
5	2007	13.56	13.506	−0.034	−0.250
6	2008	14.45	14.731	−0.081	−0.556
7	2009	15.25	15.267	−0.013	0.083
8	2010	16.81	16.907	0.037	0.220

通过全国客流量预测结果分析，可知在经过传统$GM(1,1)$模型的预测后，其客流预测的相对误差在−5.54%~−1.88%之间；而经过改进的修正残差灰色预测模型的计算预测后，其客流预测的相对误差

在-1.94%~1.52%之间;在进行二次残差修正的灰色模型计算后,其客流预测的相对误差在-0.556%~-0.882%之间。通过预测数据比较发现,经过修正后的多次残差灰色预测模型,其在预测结果的精度上明显高于传统的灰色预测模型和单次残差修正模型。若进行多次残差,则预测误差将会进一步降低,使预测精度进一步提高。

4　总结

旅客发送量是铁路客运站的重要指标,客流经济效益是给铁路部门带来经济效益的最直接来源,同时客流量也是一个主要影响客运稳定安全的因素。准确的客流量预测能给铁路客票部门、乘务等工作提供重要数据依据。本文提出的基于多次修正残差灰色预测模型的客流预测方法,克服传统灰色理论的指数缺点,强化修正残差灰色预测计算精度。该方法计算简便、可操作性强,且预测结果可以达到较高精度要求。若进行深层次的残差计算更能提高预测精度。然而由于客流量影响因素复杂多样,只有结合现有数学理论模型和实际客流影响因素才能使预测值更符合实际值,为今后的铁路客运组织决策提供合理可靠的依据。

参考文献:

[1] 潘亮. 基于EEMD-GSVM的高速铁路短期客流预测[D]. 北京. 北京交通大学,2012:20-25.

[2] 徐薇,黄厚宽,秦勇. 基于时空数据挖掘的铁路客流预测方法[J]. 北方交通大学学报,2004,28(05):16-19.

[3] 霍保世,张秀媛. Fuzzy聚类分析原理在客流预测中的应用[J]. 数量经济技术经济研究,2001(12):91-93.

[4] 刘廷新,李振宇. 指数平滑法在交通参数短期预测中的应用[J]. 山东交通学院学报,2002,10(3):29-36.

[5] 焦永兰,孙秉珍. 基于灰色理论的铁路客货运量预测研究[J]. 兰州交通大学学报,2008,27(3):74-77.

[6] 马彦祥,刘军马,敏书,等. 基于原子激发能级的铁路客流短期预测的研究[J]. 铁道运输与经济,2009,31(9):72-78.

[7] 陈怡乾. 铁路客票预售期制定及动态优化计算系统研究[D]. 成都:西南交通大学,2011.

[8] 邓聚龙. 灰色系统理论及其应用[M]. 北京:科学技术出版社,1992.

[9] 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,2012,29(1):30-34.

Prediction of Railway Passenger Flow Based on Residual Error Corrections of the Gray Model

Luo Chen,Liu Lan

(School of Traffic and Logistics,Southwest Jiaotong University,Chengdu 610031,China)

Abstract:In order to increase the forecast accuracy of the national railway passenger flow,this study improves GM (1,1) model and proposes the railway passenger flow prediction method based on residual error gray model in light of factors and characteristics of railway passenger flow. By conquering some weaknesses,such as the inaccurate prediction of the exponential function in traditional *GM* (1,1) model and the inaccurate single residual correction,the forecast error is reduced with many times of residual modification. Based on the analysis of railway passenger flow data during 2003—2010,the prediction method based on residual error gray model has such advantages as small error,high precision,easy calculation and operability,which provides a feasible approach for the railway passenger flow forecast.

Key words:gray theory;residual error gray;passenger flow prediction;accuracy