研究报告

灰色预测模型GM（1，1）的适用性分析及在火灾风险预测中的应用

陈子锦[1]，王福亮[2]，陆守香[2]
（1.中国科学技术大学统计与金融系，合肥 230026；
2.中国科学技术大学火灾科学国家重点实验室，合肥 230027）

［摘要］ 通过对灰色预测模型——GM（1，1）的理论分析，证明了该模型的预测值及其变化趋势均具有单调性，进而提出了GM（1，1）模型的适用性判据，并给出了该判据在火灾风险灰色预测中的应用实例。

［关键词］ 火灾预测；GM（1，1）模型；火灾伤人率
［中图分类号］ O29 ［文献标识码］ A ［文章编号］ 1009－1742（2007）05－0091－04

灰色系统理论建立以来，相关模型也逐渐走进了火灾科学研究者的视线，文献［1］用灰关联，文献［2］用灰色GM（1，1）模型预测我国的火灾损失，文献［3～5］预测我国在今后的几年内的火灾形势，使用灰色GM（1，1）模型对火灾次数，人员伤亡和对灾变发生年做出了预测。该模型通过原始数据的累加可以消除一定的随机性，但它并不是一种万能方法，对于有些数据模型的解并不合理，于是用通过预测值和实际值的残差占实际值百分比来判断模型的好坏，后验差检验，或者用灰色关联度检验[3,5,6]。下面将介绍一种方法，即通过观察分析数据就可以大概判断该模型适用与否。

1 方法理论

考虑原始离散数据序列 $x^{(0)}=\left[x_1^{(0)},x_2^{(0)},\cdots,x_n^{(0)}\right]$，其中0表示是原始数据，$n$ 为序列长度，对其进行一次累加生成处理得：$x_k^{(1)}=\sum_{i=1}^{k}x_i^{(0)},k=1,2,\cdots,n$，则生成序列 $x^{(1)}=\left[x_1^{(1)},x_2^{(1)},\cdots,x_n^{(1)}\right]$ 为基础建立灰色GM（1，1）模型：

$$\frac{\mathrm{d}x^{(1)}}{\mathrm{d}t}+ax^{(1)}=u \tag{1}$$

其中 a，u 是待估计的参数。

取 θ 表示参数向量（a，u），

$$X_n=(x_2^{(0)},x_3^{(0)},\cdots,x_n^{(0)})$$

$$B=\begin{bmatrix}-\left[x_2^{(1)}+x_1^{(1)}\right]/2 & 1\\ \cdots & \cdots\\ -\left[x_n^{(1)}+x_{n-1}^{(1)}\right]/2 & 1\end{bmatrix} \tag{2}$$

θ 的估计为

$$\theta=\left[B^TB\right]^{-1}B^TX_n \tag{3}$$

求解微分方程得到

$$x^{(1)}(k+1)=\left[x^{(0)}(1)-u/a\right]e^{-ak}+u/a,\quad k=(1,2,\cdots,n) \tag{4}$$

由 $x^{(1)}$ 得生成过程可以知

$$x^{(0)}(k+1)=x^{(1)}(k+1)-x^{(1)}(k)=(1-e^{a})\left[x^{(0)}(1)-u/a\right]e^{-ak} \tag{5}$$

再做一次差分，记为序列 $x^{(-1)}$，

$$x^{(-1)}(k+1)=x^{(0)}(k+1)-x^{(0)}(k)=(1-e^{a})^2\left[x^{(0)}(1)-u/a\right]e^{-ak},\quad k=1,\cdots,n-1 \tag{6}$$

不论 u，a 的估计如何，式（4），式（5）都是关于 k 的单调函数，式（4）表明估计出来的值要么单调增，要么单调减，式（5）表明估计出来的值增加或者减少的幅度也是单调的。比如说估计

［收稿日期］ 2006－04－20
［基金项目］ 国家重点基础研究发展规划资助项目（2001CB409600）
［作者简介］ 陈子锦（1979－），男，安徽省铜陵县人，中国科学技术大学统计与金融系博士生

的值是单调增的，那么它增加的越来越快，或者增加的越来越慢，如果数据适合用 GM（1，1）模型来拟合，这数据应该有上述描述的特点，这样对于有波动的数据则估计的效果就会很差，尤其是在波动程度很大的时候。其中参数 a，u 决定了模型，为了进一步研究参数的一些性质，作者给出以下的定义和定理：

定义 1：设有 2 正数列
$X=\{x_i, i=1,2\cdots\}$，$Y=\{y_i, i=1,2,\cdots\}$，称数列 X 相对于数列 Y 是 TP_2 的，如果对于 $\forall_{i,k}\in\{1,2,\cdots\}$，有 $x_iy_i+x_ky_k\geqslant x_iy_k+x_ky_i$。

定义 2：设有 2 正数列
$X=\{x_i, i=1,2,\cdots\}$，$Y=\{y_i, i=1,2,\cdots\}$，称数列 X 相对于数列 Y 是 RR_2 的，如果对于
$\forall_{i,k}\in\{1,2,\cdots\}$，有 $x_iy_i+x_ky_k\leqslant x_iy_k+x_ky_i$。

引理 1：设数列 X 相对于数列 Y 是 TP_2 的，则

$$n\sum_{i=1}^{n}x_iy_i-\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i\geqslant 0。$$

证明：

$$\begin{aligned}&n\sum_{i=1}^{n}x_iy_i-\sum_{i=1}^{n}x_i\sum_{j=1}^{n}y_j=\\&(n-1)\sum_{i=1}^{n}x_iy_i-\sum_{i\neq j}\sum x_iy_j=\\&(n-1)\sum_{i=1}^{n}x_iy_i-\sum_{i<j}\sum(x_iy_j+x_jy_i)=\\&\sum_{i<j}\sum(x_iy_i+x_jy_j-x_iy_j-x_jy_i)\geqslant 0\end{aligned}\tag{7}$$

最后不等式是由于数列 X 相对于数列 Y 是 TP_2，直接由定义得。证毕。

引理 2：设有正数列 $x^{(0)}=(x_1^{(0)},x_2^{(0)}\cdots,x_n^{(0)})$，$x_k^{(1)}=\sum_{i=1}^{k}x_i^{(0)}$，$k=1,\cdots,n$，$b_i=x_{i+1}^{(0)}$，

$-a_i=-\dfrac{x_i^{(1)}+x_{i+1}^{(1)}}{2}=-x_i^{(1)}-\dfrac{x_{i+1}^{(0)}}{2}, i=1,\cdots,n-1$，

如果 $x_i^{(0)}$ 是关于 i 单调增，则数列 $a=\{a_i\}$ 相对于数列 $b=\{b_i\}$ 是 TP_2，如果 $x_I^{(0)}$ 是关于 i 单调减，则数列 $a=\{a_i\}$ 相对于数列 $b=\{b_i\}$ 是 RR_2。

证明：

$$\begin{aligned}&a_ib_i+a_kb_k-a_ib_k-a_kb_i=\\&x_i^{(1)}x_{i+1}^{(0)}+x_{i+1}^{2(0)}/2+x_k^{(1)}x_{k+1}^{(0)}+x_{k+1}^{2(0)}/2-\\&x_i^{(1)}x_{k+1}^{(0)}-x_k^{(1)}x_{i+1}^{(0)}-x_{k+1}^{(0)}x_{i+1}^{(0)}=x_{i+1}^{(0)}(x_i^{(1)}-x_k^{(1)})+\\&x_{k+1}^{(0)}(x_k^{(1)}-x_i^{(1)})+(x_{i+1}^{(0)}-x_{k+1}^{(0)})^2/2=\\&(x_{i+1}^{(0)}-x_{k+1}^{(0)})(x_i^{(1)}-x_k^{(1)})+(x_{i+1}^{(0)}-x_{k+1}^{(0)})^2/2\end{aligned}\tag{8}$$

或

$$(x_{i+1}^{(0)}-x_{k+1}^{(0)})(x_i^{(1)}-x_k^{(1)}+x_{i+1}^{(0)}/2-x_{k+1}^{(0)}/2)\tag{9}$$

当 $x_i^{(0)}$ 关于 i 单调增时，考虑式(8)中第一个部分，不论 $i>k$，或是 $i<k$，都大于等于 0，显然第 2 部分大于等于 0，故式(8)大于等于 0，所以 a 相对于数列 b 是 TP_2。

当 $x_i^{(0)}$ 关于 i 单调减时，考虑式(9)，不论 $i>k$，或是 $i<K$，都大于等于 0，所以 A 相对于数列 b 是 RR_2。证毕

引理 3：设有正数列 $x^{(0)}=(x_1^{(0)},x_2^{(0)}\cdots x_n^{(0)})$，$x_k^{(1)}=\sum_{i=1}^{k}x_i^{(0)}, k=1,\cdots,n, b_i=x_{i+1}^{(0)}$，

$-a_i=-\dfrac{x_i^{(1)}+x_{i+1}^{(1)}}{2}=-x_i^{(1)}-\dfrac{x_{i+1}^{(0)}}{2}, i=1,\cdots,n-1$，

如果 $x_i^{(0)}$ 是关于 i 单调减，则

$$\sum_{i=1}^{n-1}a_i^2\sum_{i=1}^{n-1}b_i-\sum_{i=1}^{n-1}a_i\sum_{i=1}^{n-1}a_ib_i>0。$$

证明：

$$\begin{aligned}&\sum_{i=1}^{n-1}a_i^2\sum_{i=1}^{n-1}b_i-\sum_{i=1}^{n-1}a_i\sum_{i=1}^{n-1}a_ib_i=\\&\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}a_i^2b_j-\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}a_ia_jb_j=\\&\sum_{i\neq j}\sum a_i^2b_j-a_ia_jb_j=\\&\sum_{i<j}\sum a_i^2b_j-a_ia_jb_j+a_j^2b_i-a_ja_ib_i=\\&\sum_{i<j}\sum(a_ib_j-a_jb_i)(a_i-a_j)\end{aligned}\tag{10}$$

显然 a_i 严格单调增，由题设 b_i 单调减，且都是正数，所以式（10）大于 0。证毕。

定理 1：有原始离散数据序列
$x^{(0)}=(x_1^{(0)},x_2^{(0)}\cdots x_n^{(0)})$，当 $x_i^{(0)}$ 是关于 i 单调增的正的数列，则按灰色 GM(1,1) 模型估计出来的参数 $\hat{a}\leqslant 0$。当 $x_i^{(0)}$ 关于 i 单调减时，$\hat{a}\geqslant 0$，$\hat{u}>0$。

证明：设 $x_k^{(1)}=\sum_{i=1}^{k}x_i^{(1)}$，取 θ 表示参数向量 (a,u)，$X_n=(x_2^{(0)},x_3^{(0)},\cdots\cdots,x_n^{(0)})$。由式（3）知 $\theta=(B^TB)^{-1}B^TX_n$，其中 B 见式(2)。

$$设 -a_i=-\frac{x_i^{(1)}+x_{i+1}^{(1)}}{2}=-x_i^{(1)}-\frac{x_{i+1}^{(0)}}{2},$$

$$b_i=x_{i+1}^{(0)}, i=1,\cdots,n-1。$$

$$\Delta=(B^TB)=\begin{vmatrix}\sum_{i=1}^{n-1}a_i^2 & -\sum_{i=1}^{n-1}a_i\\ -\sum_{i=1}^{n-1}a_i & n-1\end{vmatrix}=$$

$$(n-1)\sum_{i=1}^{n-1}a_i^2-\left(\sum_{i=1}^{n-1}a_i\right)^2\neq 0 \tag{11}$$

$\dfrac{\Delta}{(n-1)^2}$ 可以看成是一个以概率 $1/(n-1)$ 取值于 $a_i(i=1,\cdots,n-1)$ 随机变量的方差，所以 $\Delta>0$。

$$(B^TB)^{-1}=\begin{pmatrix}\sum_{i=1}^{n-1}a_i^2 & -\sum_{i=1}^{n-1}a_i\\ -\sum_{i=1}^{n-1}a_i & n-1\end{pmatrix}^{-1}= \tag{12}$$

$$\frac{1}{\Delta}\begin{pmatrix}n-1 & \sum_{i=1}^{n-1}a_i\\ \sum_{i=1}^{n-1}a_i & \sum_{i=1}^{n-1}a_i^2\end{pmatrix}$$

$$\theta=(B^TB)^{-1}B^TX_n=$$

$$\frac{1}{\Delta}\begin{pmatrix}n-1 & \sum_{i=1}^{n-1}a_i\\ \sum_{i=1}^{n-1}a_i & \sum_{i=1}^{n-1}a_i^2\end{pmatrix}\begin{pmatrix}-\sum_{i=1}^{n-1}a_ib_i\\ \sum_{i=1}^{n-1}b_i\end{pmatrix}=$$

$$\frac{1}{\Delta}\begin{pmatrix}-(n-1)\sum_{i=1}^{n-1}a_ib_i+\sum_{i=1}^{n-1}a_i\sum_{i=1}^{n-1}b_i\\ \sum_{i=1}^{n-1}a_i^2\sum_{i=1}^{n-1}b_i-\sum_{i=1}^{n-1}a_i\sum_{i=1}^{n-1}a_ib_i\end{pmatrix} \tag{13}$$

由引理 1，2，3 可知，$x_i^{(0)}$ 关于 i 单调增时，$\hat{a}\leqslant 0$，当 $x_i^{(0)}$ 关于 i 单调减时 $\hat{a}\geqslant 0$，$\hat{u}>0$。证毕。

以上定理告诉我们，如果 $x_i^{(0)}$ 关于 i 单调增时，$\hat{a}\leqslant 0$，由式(5) 可知 $[x^{(0)}(1)-u/a]>0$，否则与 $x_i^{(0)}$ 关于 i 单调增矛盾，由式(6) 可知，增加的幅度也越来越大，如果 $x_i^{(0)}$ 关于 i 单调减，同理知 $\hat{a}\geqslant 0$，$[x^{(0)}(1)-u/a]<0$，减小的幅度也越来越小。这就可根据数据的特点基本判断该数据是否适合使用灰色 GM（1，1）模型。适合使用灰色 GM（1，1）模型的数据要么单调增，要么单调减，单调增的时候，增幅越来越大，单调减的时候，减幅越来越小，这种判定避免了盲目的使用该模型。

2　应用实例

从正反两个例子来说明使用上面的方法。

例 1：考虑文献［2］中我国公众聚集场所发生火灾起数数据列于表 1。

表 **1**　我国公众聚集场所发生火灾起数数据

Table 1　Public fire data of China

年份	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
原始起数/次	50	33	24	23	20	24	27	9	29	6
估计起数/次	55.8	35.6	32.6	29.9	27.5	25.2	23.1	21.2	19.4	17.8
原数据波动/次	0	17	9	1	3	−4	−3	18	−20	23
估计波动/次	0	20.2	3	2.7	2.4	2.3	2.1	1.9	1.8	1.6

首先数据是波动的，而且波动幅度也是完全没有规律，后面几年的波动非常大，由上面的讨论知道这批数据是不适合 GM（1，1）模型条件的，如果非要用 GM（1，1）模型，可以估计出模型：

$$\hat{x}_{k+1}^{(1)}=486.0127-469.0127e^{-0.0863k} \tag{14}$$

计算出来的参数 $a>0$，由定理 1 知道在数据单调减的时候 $a>0$，且下降的幅度越来越小，再观察原数据，虽然不是单调减的，但是整体是下降趋势的。正如定理 1 所述 $a>0$，估计出来的数据单调减，减幅越来越小，见表 1。但是和原始数据差别就很大，所以用这批数据建立 GM（1，1）模型是不合理的。其实一开始通过观察原始数据的特点已经可以排除使用 GM（1，1）模型了。需要说明的是，作者证明的都是最理想的情况，如果严格比照上述的数据特点来建 GM（1，1）模型的话，估计极少有数据会满足，只要与定理中数据特点差的不是很大，还是可以应用的，模型估计出来后再检验下模型的合理性，务必不要盲目的使用模型，像上述举的例子和模型条件相差太大，故不合适。

例 2：考虑 1997 年到 2003 年火灾伤人率，由于未知的原因 1997 年以前的数据和 97 年的有很大的差异，所以作者选择了从 1997 年开始，数据见表 2。首先观察到数据单调减，符合 GM（1，1）模型要求，由定理 1 知，如果适合 GM（1，1）模型，数据的减小的幅度还应该越来越小，观察减小幅度，除了 1999 年和 2001 年减小幅度增加，其它都是减小的，尤其后面几年单调减小，基本符合定理 1 要求，去除了 2003 年的数据再建模，得到模

型如下：

$$x^{(1)}(k+1)=$$
$$\left[x^{(0)}(1)-4.571\ 884\ 3/0.094\ 423\ 6\right]e^{-0.094\ 423\ 6k}+$$
$$4.571\ 884\ 3/0.094\ 423\ 6 \tag{15}$$

并用模型预测2003年的数据为2.496，与实际数据比较接近，估计数据与原始数据最大相对误差小于5％（见表2）。注意估计波动中开始有个－0.002，那是由于式（5）中 k 不能取0，后面的波动就正常了，符合定理1，数据单调减，$a>0$，且减小的幅度越来越小。基于开始对数据的观察，使用作者提出的方法发现了GM（1，1）模型是合理的，最后建立模型估计出结果，都表明了模型和数据较好的拟合。

表2　我国1997—2003年火灾伤人率数据分析及预报

Table 2　Analysis and prediction of fire injury rate of China

10^{-6}

年份	1997	1998	1999	2000	2001	2002	2003
原始伤人率	4	3.9	3.7	3.5	2.96	2.66	2.38
估计伤人率	4	4.002	3.642	3.314	3.015	2.743	2.496
相对误差	0	0.026 2	－0.016	－0.053	0.018 6	0.031 3	0.048 8
原数据波动	0	0.1	0.2	0.2	0.54	0.3	0.28
估计波动	0	－0.002	0.36	0.328	0.299	0.272	0.247

3　结论

研究了满足GM（1，1）模型的数据应该具有的特点：该模型的数据具有单调性，并且增加或是减小的幅度也具有某种单调性；分析了数据单调增或是减的时候，参数的正负，和数据变化幅度的关系，为判断数据是否适合GM（1，1）模型建立了理论依据。GM（1，1）模型有其很大的局限性，但是它也有其优点就是所需数据不是很多即可，所以在现在我国火灾数据很匮乏的情况下还是有一定的优势的，但是注意不要滥用该模型。

参考文献

［1］ 卢兆明，胡宝清，陆君安．高层建筑火灾风险灰关联评估[J]．武汉大学学报（工学版），2004，37(5)：62～66

［2］ 王艺栋．公众聚集场所火灾及危险源的数学分析[J]．上海第二工业大学报，2005，22(22)：48～54

［3］ 徐志胜，白国强，冯　凯．灰色预测理论在火灾损失预测中的应用[J]．湘潭矿业学院学报，2004，19(3)：24～26

［4］ 姜学鹏，徐志胜．我国火灾损失的时间趋势分析及动态预测[J]．消防科学与技术，2005，24(4)：412～414

［5］ 徐志胜，白国强．灾变预测在火灾预测中的作用[J]．安全与环境学报，2003，3(6)：71～73

［6］ 邓聚龙．灰色预测与决策[M]．武汉：华中理工大学出版社，1985

［7］ Karlin S．Total Positivity [M]．Stanford California：Stanford University Press，1968

［8］ 公安部消防局编．中国消防年鉴（2000）[M]．北京：中国人事出版社，2004．304

Analysis of GM(1,1) Model and Its Application in Fire Risk Prediction

Chen Zijin[1]，Wang Fuliang[2]，Lu Shouxiang[2]

（1．*Department of Statistics and Finance*，*University of Science and Technology of China*．*Hefei*　230027，*China*；2．*State Key Laboratory of Fire Science*，*University of Science and Technology of China*．*Hefei*　230027，*China*）

［**Abstract**］　Theoretical analysis of grey prediction model GM(1,1) is present in this paper. Monotonicity of predicted value and its variation tendency predicted by GM(1,1)model is proved. Based on the monotonicity of predicted value and its variation tendency, applicability criterion of GM(1,1) is brought forward. Example applications of the criterion in fire risk grey prediction are discussed.

［**Key words**］　fire forecast；GM(1,1)；rate of the fire injured