

文章编号: 1001-0920(2013)06-0950-03

# GM(1,2)模型改进方法研究及应用

雷鸣雳, 冯祖仁

(西安交通大学 a. 系统工程研究所, b. 机械制造系统工程国家重点实验室, 西安 710049)

**摘要:** 针对GM(1,2)建模难点和模型缺陷提出两种改进方法: 一是运用相关匹配算法, 在历史数据库中搜索与主序列具有强关联特性的数据序列, 确定为模型参考序列; 二是引入粒子群算法, 以模型预测性能评价指标为目标函数对模型参数进行辨识, 改善模型预测性能. 算例结果表明了改进方法的适用性和有效性.

**关键词:** 灰色预测; GM(1,2)模型; 粒子群算法; 相关系数

中图分类号: N94

文献标志码: A

## Study on the improved method for GM(1,2) model and its application

LEI Ming-li, FENG Zu-ren

(a. Systems Engineering Institute, b. State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China. Correspondent: LEI Ming-li, E-mail: mllei768@126.com)

**Abstract:** Two improved method are proposed for the difficulty and defects of the traditional GM(1,2) model. Firstly, the correlation matching algorithm is adopted to search the reference sequence which has the maximum grey relation grade in the history database. Secondly, the particle swarm optimization algorithm is used to identify the model parameters by taking the evaluation function of the forecasting performance as the objective function. The simulation results show the effectiveness and the applicability of the presented method.

**Key words:** grey forecasting; GM(1,2) model; particle swarm optimization; correlation coefficient

### 0 引言

与GM(1,1)相比, GM(1,2)在模型输入端引入与主序列具有强关联特性的参考序列, 能够有效解决波动性变化序列预测问题. 已有学者对GM(1,2)进行研究并取得成效<sup>[1-3]</sup>, 但GM(1,2)中参考序列的确定是建模面临的难题; 另外, 最小二乘法(LSM)参数辨识目标函数与模型预测性能评价指标不同, 影响模型预测精度. 为了解决以上问题, 本文提出一种基于相关匹配运算的参考序列搜索方法, 并引入粒子群算法对模型参数进行辨识, 能够有效改善模型的预测性能.

### 1 GM(1,2)建模原理及缺陷分析

#### 1.1 GM(1,2)建模原理

记原始主序列和参考序列分别为 $y^{(0)}$ 和 $r^{(0)}$ , 其一次累加序列分别为 $y^{(1)}$ 和 $r^{(1)}$ , 建立灰色微分方程

$$y^{(0)}(k) + az^{(1)}(k) = br^{(1)}(k). \quad (1)$$

其中:  $a$ 为发展系数;  $b$ 为灰作用量;  $Z^{(1)}(k)$ 为模型背景值. 式(1)对应的白化微分方程为

$$\dot{y}^{(1)} + ay^{(1)} = br^{(1)}. \quad (2)$$

求解方程(2)即可得原始主序列预测值

$$\hat{y}^{(1)}(k+1) = \left( y^{(0)}(1) - \frac{b}{a}r^{(1)}(k+1) \right) e^{(-ak)} + \frac{b}{a}r^{(1)}(k+1), \quad (3)$$

其中参数 $a$ 和 $b$ 由LSM依据式(1)辨识确定.

#### 1.2 GM(1,2)建模难点和缺陷分析

首先, 由模型原理分析可知, 建模用参考序列必须同时满足两个要求: 与主序列具有强关联特性; 预测时刻数值已知. 可见, 合理确定参考序列是GM(1,2)建模需解决的难题. 其次, 利用LSM对模型参数辨识的目标为误差平方和最小, 而模型预测性能评价常以平均绝对百分比误差为核心指标. 显然, 目标函数与模型预测性能评价指标不同, 影响模型的预测性能.

### 2 GM(1,2)模型改进方法

#### 2.1 采用相关匹配算法进行参考序列搜索

设由预测变量历史数据构成的时间序列为

收稿日期: 2011-12-26; 修回日期: 2012-05-12.

基金项目: 国家自然科学基金项目(60875043); 教育部博士点基金项目(2010020111003).

作者简介: 雷鸣雳(1971-), 男, 博士生, 从事电力市场、灰色预测理论的研究; 冯祖仁(1953-), 男, 教授, 博士生导师, 从事多智能体、机器人等研究.

$$H = \{H(k)|k = 1, 2, \dots, N, \dots, M\}. \quad (4)$$

其中:  $M$  为历史时间序列长度,  $N$  为模型训练数据长度,  $M \gg N$ .

以  $H$  为母板, 将主序列  $y^{(0)}$  与母板逐位滑动进行相关匹配运算, 匹配运算次数为  $(M - N)$ , 待匹配完成后, 确定匹配相关系数最大值, 若最大值出现在第  $j$  次匹配, 则模型参考序列可确定为

$$r^0 = \{r^0(i)|r^0(i) = H(i + j - 1)\}. \quad (5)$$

第  $k$  次滑动匹配相关系数定义如下<sup>[4]</sup>:

$$R(k) = \frac{N \sum_{i=1}^N y^0(i)H(i+k) - \sum_{j=1}^N y^0(j) \sum_{i=1}^N H(i+k)}{\sqrt{N \sum_{i=1}^N (y^0(i))^2 - \sum_{i=1}^N (y^0(i))^2}} \times \frac{1}{\sqrt{N \sum_{i=1}^N (H(i+k))^2 - \sum_{i=1}^N (H(i+k))^2}}. \quad (6)$$

### 2.2 采用粒子群算法进行模型参数辨识

应用粒子群算法的非线性参数优化能力<sup>[5]</sup>, 对 GM(1,2) 模型参数  $a, b$  进行辨识, 具体描述如下.

**Step 1:** 初始化粒子群. 在  $D = 2$  维空间中随机产生  $m$  个粒子, 初始化粒子位置和速度. 第  $k$  次迭代中第  $i$  个粒子位置向量  $x_{(i)}^k = \{x_{i1}^k, x_{i2}^k\}$ , 速度向量  $v_{(i)}^k = \{v_{i1}^k, v_{i2}^k\}$ , 粒子在二维空间中的坐标对应待辨识参数  $a$  和  $b$  的某一组潜在解.

**Step 2:** 计算粒子第  $k$  次迭代中第  $i$  个粒适应度, 有

$$f_i^k(a, b) = \frac{1}{1 + e_{MAPE}}, \quad (7)$$

其中  $e_{MAPE}$  为平均绝对百分比误差, 有

$$e_{MAPE} = \frac{1}{N} \sum_{j=1}^N e_{APE}(j),$$

$$e_{APE}(j) = \frac{|y^{(0)}(j) - \hat{y}^{(0)}(j)|}{y^{(0)}(j)} \times 100\%. \quad (8)$$

**Step 3:** 更新粒子速度和位置

$$v_{id}^{k+1} = \omega v_{id}^k + c_1 r_1 (p_i - x_{id}^k) + c_2 r_2 (p_g - x_{id}^k),$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1}. \quad (9)$$

其中:  $k$  为当前更新迭代次数;  $v_{id}^k$  和  $x_{id}^k$  分别为粒子当前速度和位置;  $v_{id}^{k+1}$  和  $x_{id}^{k+1}$  分别为粒子更新后速度和位置;  $c_1$  和  $c_2$  为加速度常数, 设为常数 2;  $r_1$  和  $r_2$  为范围在  $[0,1]$  内变化的随机数;  $\omega$  为惯性权重系数.

**Step 4:** 更新个体最优位置  $p_i$  和种群最优位置  $p_g$ , 若  $f_i^k(a, b) > p_i$ , 则  $p_i = f_i^k(a, b)$ ; 若  $f_i^k(a, b) > p_g$ , 则  $p_g = f_i^k(a, b)$ .

**Step 5:** 判定算法是否停止. 若迭代次数达到最大化代数, 则转至 Step 6; 否则, 返回 Step 2.

**Step 6:** 输出最优辨识参数  $a$  和  $b$ , 算法结束.

### 3 算例分析

给出北欧 Nordpool 电力市场 3 个典型时段电价预测算例, 各算例中匹配母板由市场 2006 年 1 月 1 日 ~ 2007 年 4 月 30 日典型时段历史电价数据构成, 序列长度  $M = 485$ ; 电价主序列由 2007 年 4 月 23 日 ~ 2007 年 4 月 30 日典型时段历史电价数据构成, 序列长度  $N = 8$ ; 相关匹配运算次数为 477 次.

**例 1**(谷负荷段电价预测) 2007 年 4 月 23 日 ~ 4 月 30 日谷负荷段电价主序列

$$y^{(0)} = \{22.29, 22.53, 21.74, 21.52, 19.53, 19.87, 17.90, 17.72, 19.09\}.$$

应用相关匹配算法搜索得到电价参考序列

$$r^{(0)} = \{52.06, 51.10, 50.95, 50.08, 45.96, 45.33, 42.29, 40.57, 40.78\}.$$

477 次匹配运算中, 相关系数最大值为 0.9866, 出现在第 114 次匹配, 已在图 1 中用圆圈标出. 以此建立普通 GM(1,2) 和改进 GM(1,2) 模型, 预测结果如表 1 所示.

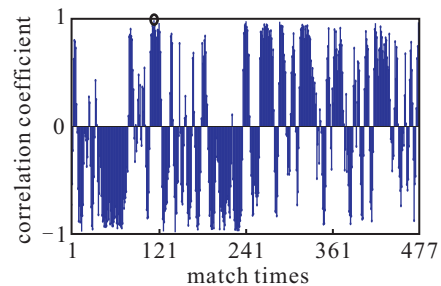


图 1 相关匹配运算相关系数曲线图(谷负荷段)

表 1 数据序列预测结果比较(谷负荷段)

预测日期	原始数据	普通 GM(1,2) 模型		改进 GM(1,2) 模型	
		拟合值	$e_{APE}/\%$	拟合值	$e_{APE}/\%$
4月23日	22.29	22.29	0.00	22.29	0.00
4月24日	22.53	19.26	14.53	22.04	2.18
4月25日	21.74	24.01	10.43	21.89	0.71
4月26日	21.52	22.11	2.76	21.52	0.00
4月27日	19.53	19.88	1.77	19.75	1.12
4月28日	19.87	19.51	1.81	19.48	1.97
4月29日	17.90	18.19	1.61	18.17	1.52
4月30日	17.72	17.45	1.55	17.43	1.62
$e_{APE}/\%$		4.31		1.14	
辨识参数		$a = 2.06$ $b = 0.89$		$a = 10.00$ $b = 4.30$	

**例 2**(峰负荷段电价预测) 2007 年 4 月 23 日 ~ 4 月 30 日峰负荷段电价主序列

$$y^{(0)} = \{25.68, 25.12, 24.15, 23.47, 22.40, 19.60, 10.99, 21.40, 17.62\}.$$

应用相关匹配算法搜索得到电价参考序列

$$r^{(0)} = \{82.27, 80.25, 78.52, 75.35, 72.06, 62.31, 42.88, 67.22, 65.95\}.$$

477次匹配运算中, 相关系数最大值为0.9925, 出现在第240次匹配, 已在图2中用圆圈标出. 以此建立普通GM(1,2)和改进GM(1,2)模型, 预测结果如表2所示.

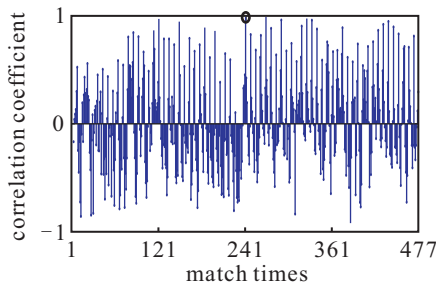


图 2 相关匹配运算相关系数曲线图(峰负荷段)

表 2 数据序列预测结果比较(峰负荷段)

预测日期	原始数据	普通 GM(1,2) 模型		改进 GM(1,2) 模型	
		拟合值	$e_{APE}/\%$	拟合值	$e_{APE}/\%$
4月23日	25.68	25.68	0.00	25.68	0.00
4月24日	25.12	21.57	14.15	24.94	0.72
4月25日	24.15	26.20	8.49	24.46	1.28
4月26日	23.47	23.63	0.70	23.47	0.00
4月27日	22.40	22.21	0.83	22.45	0.20
4月28日	19.60	19.15	2.30	19.41	0.98
4月29日	10.99	13.17	19.85	13.36	21.53
4月30日	21.40	20.65	3.52	20.94	2.16
$e_{APE}(\%)$		6.23		3.36	
辨识参数		$a = 2.21$ $b = 0.68$		$a = 9.50$ $b = 2.96$	

例3(平负荷段电价预测) 2007年4月23日~4月30日平负荷段电价主序列

$$y^{(0)} = \{23.97, 24.37, 23.19, 22.87, 21.47, 17.92, 17.35, 21.28, 18.00\}.$$

应用相关匹配算法搜索得到电价参考序列

$$r^{(0)} = \{24.96, 25.38, 24.63, 24.74, 22.85, 19.84, 19.27, 23.52, 23.70\}.$$

477次相关匹配运算中, 相关系数最大值为0.9895, 出现在第443次匹配, 已在图3中用圆圈标出. 以此建立普通GM(1,2)和改进GM(1,2)模型, 预测结果如表3所示.

以上算例中, 进行相关匹配运算获得的参考序列与主序列相关系数均大于0.9, 具有强关联性, 依据

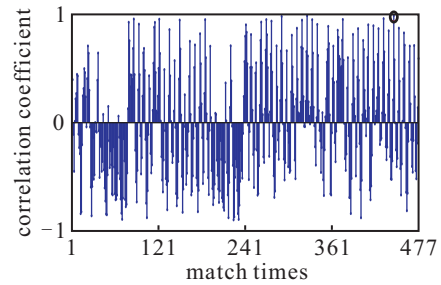


图 3 相关匹配运算相关系数曲线图(平负荷段)

表 3 数据序列预测结果比较(平负荷段)

预测日期	原始数据	普通 GM(1,2) 模型		改进 GM(1,2) 模型	
		拟合值	$e_{APE}/\%$	拟合值	$e_{APE}/\%$
4月23日	23.97	23.97	0.00	23.97	0.00
4月24日	24.37	20.59	15.50	22.46	7.85
4月25日	23.19	24.47	5.52	22.87	1.37
4月26日	22.87	23.22	1.53	22.87	0.00
4月27日	21.47	21.20	1.27	21.12	1.62
4月28日	17.92	18.37	2.52	18.34	2.34
4月29日	17.35	17.84	2.82	17.81	2.67
4月30日	21.28	21.77	2.32	21.74	2.17
$e_{APE}(\%)$		3.93		2.25	
辨识参数		$a = 2.41$ $b = 2.23$		$a = 5.36$ $b = 4.96$	

搜索到的参考序列建立的GM(1,2)模型均表现出较好的预测性能, 表明基于相关匹配运算的参考序列搜索方法能有效应用于GM(1,2)模型的建立. 另外, 基于粒子群算法参数辨识的GM(1,2)改进模型较普通GM(1,2)模型拟合误差平均降低1.7~3个百分点, 表现出更好的预测能力, 尤其对波动性较大数据序列改善效果更加明显.

### 4 结 论

通过理论分析和算例验证, 获得结论如下:

1) 本文提出的基于主序列和预测变量历史数据序列相关匹配运算的参考序列搜索方法, 能够有效解决GM(1,2)模型中参考序列确定的难题, 为模型推广应用提供了新思路;

2) 本文采用粒子群算法, 以模型预测性能评价指标为目标函数对模型参数a和b进行辨识, 明显降低了模型预测误差, 改善了模型的预测性能.

### 参考文献(References)

[1] 谢乃明, 刘思峰. GM(n, h)模型建模序列数据乘变化特性研究[J]. 控制与决策, 2009, 24(9): 1294-1299. (Xie N M, Liu S F. Research on property of GM(n, h) model under data multiple transformation[J]. Control and Decision, 2009, 24(9): 1294-1299.)