

# 基于GWO优化的L1正则化分数阶灰色时间幂预测模型

董帮强, 张延飞\*, 丁木华, 陈萍

东华理工大学理学院, 江西 南昌

收稿日期: 2021年9月13日; 录用日期: 2021年10月6日; 发布日期: 2021年10月15日

## 摘要

灰色预测模型是一类处理小样本数据预测的有效方法, 其中单变量灰色预测模型  $GM(1,1,t^\alpha)$  是一个重要的研究对象。  $FAGM(1,1,t^\alpha)$  模型是在基于一阶累加的  $GM(1,1,t^\alpha)$  模型的基础上引入分数阶累加形式的预测模型。然而, 该模型的精度不够高, 且容易存在过拟合现象。本文结合Lasso回归中的L1正则化思想, 对分数阶累加的灰色时间幂模型  $FAGM(1,1,t^\alpha)$  进行正则化, 提出正则化分数阶灰色时间幂预测模型  $LFAGM(1,1,t^\alpha)$ , 使用坐标下降算法替代最小二乘估计来求解模型的参数。同时, 使用灰狼优化算法(GWO)搜索  $LFAGM(1,1,t^\alpha)$  模型的最优非线性参数。并基于中国农业耕地灌溉面积情况(2008~2019年)进行算例分析, 结果表明,  $LFAGM(1,1,t^\alpha)$  模型具有更高的预测精度。

## 关键词

分数阶灰色预测模型, L1正则化, 灰狼优化算法, 坐标下降

# L1 Regularized Fractional Grey Time Power Forecasting Model Based on GWO Optimization

Bangqiang Dong, Yanfei Zhang\*, Muhua Ding, Ping Chen

School of Science, East China University of Technology, Nanchang Jiangxi

Received: Sep. 13<sup>th</sup>, 2021; accepted: Oct. 6<sup>th</sup>, 2021; published: Oct. 15<sup>th</sup>, 2021

\*通讯作者。

文章引用: 董帮强, 张延飞, 丁木华, 陈萍. 基于GWO优化的L1正则化分数阶灰色时间幂预测模型[J]. 应用数学进展, 2021, 10(10): 3277-3287. DOI: 10.12677/aam.2021.1010344

## Abstract

Grey prediction model is an effective method for dealing with small sample data prediction, in which univariate grey prediction model  $GM(1,1,t^\alpha)$  is an important research object. The  $FAGM(1,1,t^\alpha)$  model is a prediction model based on the first-order cumulative  $GM(1,1,t^\alpha)$  model by introducing the fractional-order cumulative form. However, the accuracy of the model is not high enough and it is prone to over-fitting. This paper combines the L1 regularization idea in Lasso regression to regularize the fractional cumulative gray time power model  $FAGM(1,1,t^\alpha)$ , and proposes a regularized fractional gray time power prediction model  $LFAGM(1,1,t^\alpha)$ , using coordinate descent algorithm instead of least squares estimation to solve the model parameter. At the same time, Grey Wolf Optimization (GWO) is used to search the optimal nonlinear parameters of  $LFAGM(1,1,t^\alpha)$  model. Based on the agricultural farmland irrigation area in China (2008~2019), the results show that the  $LFAGM(1,1,t^\alpha)$  model has higher prediction accuracy.

## Keywords

Fractional Gray Prediction Model, L1 Regularized, Grey Wolf Optimization Algorithm, Coordinate Descent

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

灰色预测是对既含有已知信息又含有不确定信息的系统进行预测,对在一定范围内变化的、与时间有关的灰色过程进行预测控制的理论。灰色预测模型由邓聚龙教授[1]首次提出,在专家学者的不懈努力和完善下,弥补了由于数据少所带来的预测问题,尤其在在实际应用中如何提高预测模型的精度一直是广大学者研究的热点问题。在初始值优化角度方面,Ding [2]等人对初始值进行了非线性优化,使得所构建的模型的初始值各分量加权系数都具备了可调节性的特点;王义闹[3]等人同样以误差平方和最小值为目标函数,优化了指数模型的一个参数,并且得到了最小的平方误差和;Wang [4]等人提出了一种基于矩阵的算法,建立了初始条件和发展系数间的关系。在模型参数优化的角度方面,通过怎么样的方式去优化灰色预测模型是当下灰色预测模型较为重要的内容,例如Liu [5]等人利用Weibull累积分布函数构建了一个双形参数来替代原来较为传统的常数参数值,并且通过拟合实例验证了所构建的模型具有较高的拟合度;Zhao [6]等人使用蚁狮优化算法优化了灰色预测模型参数,很大程度上提升了模型的预测精度,同时证明了该算法可行性;Yi-Shian-Lee [7]设置误差的最小绝对值为目标函数,利用遗传算法来提高预测模型的精度。从灰色预测模型的累加方式上看,其中Wen Kunli [8]提出了局部的序列灰色累加方式;刘解放[9]等人提出并使用了分数阶反向累加的离散灰色模型,通过实例证明了其可行性和优越性;谢波[10]等人通过在 $GM(1,1,t)$ 预测模型基础上选取合适的的数据维度来增加数据长度,然后取多期的预测值的平均值作为最终的预测值,进一步提高了模型的精度,减少了数据波动;陈英超[11]等人则是通过将

累加生成变为卷积变换,从而建立带有时间项的  $GM(1,1,t)$  灰色预测模型。但随着灰色系统越发复杂,传统的灰色预测模型已经不能单单只应用指数特征去进行描述,所以应该在这些较为复杂的系统中进行内在的演化与分析,钱吴永[12]等人建立了带有时间幂次项的  $GM(1,1,t^\alpha)$  模型并研究讨论了参数  $\alpha$  在不同取值下模型的性质,并对沿海高速路软土地地基沉降进行了模拟分析;而 Wu Lifeng [13]等人将灰色预测模型累加次数从整数过度到分数,构建了分数阶灰色  $FGM(1,1)$  预测模型;Wu [14]等人引入了分数阶的累加方式,建立了分数阶含有时间幂次项的灰色  $FAGM(1,1,t^\alpha)$  预测模型,并对中国卫生支出进行了研究分析。

以上文献都是基于传统的灰色预测模型进行优化建模,并于不同优化角度都保留了最小二乘法,对于现实数据处理往往存在过拟合问题。本研究则通过在分数阶累加灰色时间幂预测模型  $FAGM(1,1,t^\alpha)$  建模基础上使用 L1 正则化,即在传统的最小二乘法中的损失函数中引入惩罚项,构建正则化预测模型  $LFAGM(1,1,t^\alpha)$ ,使用坐标下降算法求解模型的参数。并对于超参数问题,通过灰狼优化算法(GWO)优化调整 L1 正则化参数  $\lambda$ ,  $LFAGM(1,1,t^\alpha)$  模型的非线性参数  $\alpha$  以及分数阶累加阶数  $r$ ,从而避免由最小二乘法带来的过拟合问题。最后,以中国农业耕地灌溉面积情况(2008~2019 年)进行了仿真模拟,验证了方法的可行性。

## 2. 分数阶灰色预测模型

### 2.1. $GM(1,1,t^\alpha)$ 模型

$GM(1,1,t^\alpha)$  模型是由钱吴永等人首次提出。该模型将灰色作用量由常数  $b$  变为了一个关于时间呈现非线性变化的幂函数形式,扩展了单变量灰色预测模型的适用范围,很大程度上提高了模型的拟合和预测精度。设  $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))^T, n \geq 4$  为非负原始数据列。对  $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))^T$  进行一次累加得到新的生成数据列  $x^{(1)}$  ( $x^{(0)}$  的 1-AGO 序列)为

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n))^T \quad (1.1)$$

其中  $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), k = 1, 2, \dots, n$ 。令  $z^{(1)}$  为一阶累加序列  $x^{(1)}$  的紧邻均值生成序列,即

$$z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n))^T \quad (1.2)$$

其中  $z^{(1)}(m) = \delta x^{(1)}(m) + (1-\delta)x^{(1)}(m-1), m = 2, 3, \dots, n$  且  $\delta = 0.5$ 。

设灰色  $GM(1,1,t^\alpha)$  模型的白化方程为

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = bt^\alpha + c, \alpha > 0 \quad (1.3)$$

公式(1.3)中的  $a$  为  $GM(1,1,t^\alpha)$  模型的发展系数,  $bt^\alpha + c$  为灰色作用量。因此,  $GM(1,1,t^\alpha)$  模型的基本形式为

$$x^{(0)}(k) + az^{(1)}(k) = bk^\alpha + c, k = 2, 3, \dots, n \quad (1.4)$$

方程(1.4)同时也被称为灰色微分方程。注意到,线性方程(1.4)为一个超定方程组,根据最小二乘法,我们可得参数向量  $\beta = (a, b, c)^T$  的估计式为

$$(a, b, c)^T = (\Delta^T \Delta)^{-1} (\Delta^T \Psi) \quad (1.5)$$

其中

$$\Delta = \begin{pmatrix} -z^{(1)}(2) & 2^\alpha & 1 \\ -z^{(1)}(3) & 3^\alpha & 1 \\ \vdots & \vdots & \vdots \\ -z^{(1)}(n) & n^\alpha & 1 \end{pmatrix}, \quad \Psi = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{pmatrix} \quad (1.6)$$

其中  $n$  表示用于建立  $GM(1,1,t^\alpha)$  模型的样本个数。

## 2.2. 分数阶 $FAGM(1,1,t^\alpha)$ 预测模型

一般地, 一阶累加方式为分数阶累加方式的一个特例, 并且分数阶累加能够更加体现数据序列内在的规律, 更不易破坏数据间潜在的关联。另外, 分数阶累加相比一阶累加更加灵活, 能适应更多的数据。因此, 在  $GM(1,1,t^\alpha)$  模型的基础上, 将分数阶累加考虑进来, 能够显著提高  $GM(1,1,t^\alpha)$  模型的预测精度和扩大使用范围。由于一阶累加的  $GM(1,1,t^\alpha)$  模型与分数阶累加的  $FAGM(1,1,t^\alpha)$  模型在建模过程中只是累加方式不同, 由此我们先引入分数阶累加和累减的定义。

**定义 1:** 设  $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))^T$  为  $n$  个观测到的非负原始数据序列,  $T$  为转置符号。 $X^{(r)}$  是  $X^{(0)}$  的  $r$  阶累加生成序列( $r$ -AGO)序列。其中  $x^{(r)}(k) = \sum_{i=1}^k x^{(r-1)}(i)$ ,  $k=1, 2, \dots, n$ 。注意到, 为了形式更加紧凑, 该累加可以改写为更加方便的矩阵格式

$$X^{(r)} = A^r X^{(0)} \quad (1.7)$$

其中

$$A^r = \begin{pmatrix} \begin{bmatrix} r \\ 0 \end{bmatrix} & 0 & 0 & \cdots & 0 \\ \begin{bmatrix} r \\ 1 \end{bmatrix} & \begin{bmatrix} r \\ 0 \end{bmatrix} & 0 & \cdots & 0 \\ \begin{bmatrix} r \\ 2 \end{bmatrix} & \begin{bmatrix} r \\ 1 \end{bmatrix} & \begin{bmatrix} r \\ 0 \end{bmatrix} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \begin{bmatrix} r \\ n-1 \end{bmatrix} & \begin{bmatrix} r \\ n-2 \end{bmatrix} & \begin{bmatrix} r \\ n-3 \end{bmatrix} & \cdots & \begin{bmatrix} r \\ 0 \end{bmatrix} \end{pmatrix}_{n \times n} \quad (1.8)$$

$$\text{并且 } \begin{bmatrix} r \\ i \end{bmatrix} = \frac{r(r+1)\cdots(r+i-1)}{i!} = \binom{r+i-1}{i} = \frac{(r+i-1)!}{i!(r-1)!}, \begin{bmatrix} 0 \\ i \end{bmatrix} = 0, \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \binom{0}{0} = 1。$$

**定义 2:** 设第  $r$  阶逆累加生成(累减)定义为  $x^{(r-1)}(k) = x^{(r)}(k) - x^{(r)}(k-1)$ ,  $k=2, 3, \dots, n$  且  $x^{(r)}(1) = x^{(0)}(1)$ ,  $r > 0$ 。用矩阵  $D^r$  表示满足  $X^{(0)} = D^r X^{(r)}$  的  $r$ -IAGO 矩阵, 其中

$$D^r = \begin{pmatrix} \begin{bmatrix} -r \\ 0 \end{bmatrix} & 0 & 0 & \cdots & 0 \\ \begin{bmatrix} -r \\ 1 \end{bmatrix} & \begin{bmatrix} -r \\ 0 \end{bmatrix} & 0 & \cdots & 0 \\ \begin{bmatrix} -r \\ 2 \end{bmatrix} & \begin{bmatrix} -r \\ 1 \end{bmatrix} & \begin{bmatrix} -r \\ 0 \end{bmatrix} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \begin{bmatrix} -r \\ n-1 \end{bmatrix} & \begin{bmatrix} -r \\ n-2 \end{bmatrix} & \begin{bmatrix} -r \\ n-3 \end{bmatrix} & \cdots & \begin{bmatrix} -r \\ 0 \end{bmatrix} \end{pmatrix}_{n \times n} \quad (1.9)$$

$$\text{并且 } \begin{bmatrix} -r \\ i \end{bmatrix} = \frac{-r(-r+1)\cdots(-r+i-1)}{i!} = (-1)^{r(i-1)\cdots(r-i+1)} i!, \begin{bmatrix} -r \\ i \end{bmatrix} = 0, i > r.$$

由公式(1.8)和(1.9)可以明显看出一阶累加是分数阶累加的一个特殊例子, 即当  $r=1$  时为一阶累加。接下来介绍一下分数阶累加和累减算子的一个基本定理。

**定理 1:** 矩阵  $A^r$  和矩阵  $D^r$  满足  $A^r D^r = I_n$

**证明 1:** 由定义 1 可知  $\det(A^r) = 1$ , 这意味着  $A^r$  为可逆矩阵。从  $X^{(r)} = A^r X^{(0)}$  可知  $(A^r)^{-1} X^{(r)} = (A^r)^{-1} A^r X^{(0)} = X^{(0)}$ 。根据定义 2, 有  $X^{(0)} = D^r X^{(r)}$ 。因此, 显然有  $A^r D^r = I_n$ , 证毕。

由前所述,  $GM(1,1,t^\alpha)$  模型与  $FAGM(1,1,t^\alpha)$  模型的建模过程只是累加方式发生了些许改变, 因此  $FAGM(1,1,t^\alpha)$  模型的白化微分方程为

$$\frac{dx^{(r)}(t)}{dt} + ax^{(r)}(t) = bt^\alpha + c, r > 0, \alpha > 0 \quad (1.10)$$

将等式(1.10)在区间  $[k-1, k]$  进行积分可得

$$\int_{k-1}^k dx^{(r)}(t) + a \int_{k-1}^k x^{(r)}(t) dt = b \int_{k-1}^k t^\alpha dt + c \int_{k-1}^k dt \quad (1.11)$$

根据梯形公式以及  $z^{(r)}(k) = 0.5x^{(r)}(k) + 0.5x^{(r)}(k-1) + 0.5x^{(r)}(k-1)$ ,  $k = 2, 3, \dots, n$ , 上述公式(1.11)可进一步简化为

$$x^{(r)}(k) - x^{(r)}(k-1) + \alpha z^{(r)}(k) = b \frac{k^{1+\alpha} - (k-1)^{1+\alpha}}{1+\alpha} + c \quad (1.12)$$

同样地, 根据最小二乘法可得  $FAGM(1,1,t^\alpha)$  模型的参数向量  $\beta = (a, b, c)^T$  的估计式为

$$(a, b, c)^T = (B^T B)^{-1} (B^T Y) \quad (1.13)$$

其中

$$B = \begin{pmatrix} -z^{(r)}(2) & \frac{2^{1+\alpha} - 1}{1+\alpha} & 1 \\ -z^{(r)}(3) & \frac{3^{1+\alpha} - 2^{1+\alpha}}{1+\alpha} & 1 \\ \vdots & \vdots & \vdots \\ -z^{(r)}(v) & \frac{v^{1+\alpha} - (v-1)^{1+\alpha}}{1+\alpha} & 1 \end{pmatrix}, Y = \begin{pmatrix} x^{(r)}(2) - x^{(r)}(1) \\ x^{(r)}(3) - x^{(r)}(2) \\ \vdots \\ x^{(r)}(v) - x^{(r)}(v-1) \end{pmatrix} \quad (1.14)$$

**定理 2:** 假设观测序列、累加和累减方式满足前面的定义 1 和定义 2, 则  $FAGM(1,1,t^\alpha)$  模型的时间响应序列为

$$\hat{x}^{(r)}(k) = \left( x^{(0)}(1) - \frac{c}{a} \right) e^{-a(k-1)} + \frac{c}{a} + \frac{b}{2} e^{-a(k-1)} \sum_{i=1}^{k-1} \left( i^\alpha e^{a(i-1)} + (i+1)^\alpha e^{ai} \right), k = 2, 3, \dots, n \quad (1.15)$$

且  $FAGM(1,1,t^\alpha)$  模型的预测值  $\hat{x}^{(0)}(k)$  为

$$\hat{x}^{(0)} = D^r \hat{x}^{(r)} \quad (1.16)$$

**证明 2:** 已知公式(1.10)的解为齐次方程的通解及其特解的和。因此, 由常数变异法可得

$$\frac{dx^{(r)}(t)}{dt} + ax^{(r)}(t) = 0 \quad (1.17)$$

令  $x^{(r)}(t) = Ue^{-at}$ ，其中  $U$  为常数，并将  $x^{(r)}(t) = Ue^{-at}$  代入公式(1.10)得到

$$\frac{dU(t)}{dt} = e^{at}(bt^\alpha + c) \quad (1.18)$$

对等式(1.18)在区间  $[1, t]$  上进行积分可以得到

$$\int_1^t dU(s) = b \int_1^t s^\alpha e^{as} ds + c \int_1^t e^{as} ds \quad (1.19)$$

进一步可以有

$$U(t) = U(1) + b \int_1^t s^\alpha e^{as} ds + \frac{c}{a}(e^{at} - e^a)$$

又因为  $U(1) = x^{(r)}(1)e^a = x^{(0)}(1)e^a$ ，根据数值积分公式中的两点梯形近似公式可以得到

$$\begin{aligned} x^{(r)}(t) &= e^{-at} \left\{ x^{(0)}(1)e^a + b \int_1^t s^\alpha e^{as} ds + \frac{c}{a}(e^{at} - e^a) \right\} \\ &= \left( x^{(0)}(1) - \frac{c}{a} \right) e^{-a(t-1)} + \frac{c}{a} + b e^{-a(t-1)} \int_1^t s^\alpha e^{a(s-1)} ds \\ &= \left( x^{(0)}(1) - \frac{c}{a} \right) e^{-a(t-1)} + \frac{c}{a} + \frac{b}{2} e^{-a(t-1)} \sum_{i=1}^{t-1} \left( i^\alpha e^{a(i-1)} + (i+1)^\alpha e^{ai} \right) \end{aligned} \quad (1.20)$$

所以  $FAGM(1,1,t^\alpha)$  的时间响应函数为：

$$\hat{x}^{(r)}(k) = \left( x^{(0)}(1) - \frac{c}{a} \right) e^{-a(k-1)} + \frac{c}{a} + \frac{b}{2} e^{-a(k-1)} \sum_{i=1}^{k-1} \left( i^\alpha e^{a(i-1)} + (i+1)^\alpha e^{ai} \right) \quad (1.21)$$

### 3. $FAGM(1,1,t^\alpha)$ 模型的正则化

#### 3.1. 线性回归及 L1 正则化

在灰色预测模型中，通过求解(2.1)来得到模型的估计参数

$$\hat{\beta}_{linear} = \operatorname{argmin}_{\beta} \frac{1}{2} \sum_{i=1}^n (B_i \beta - y_i)^2 = \frac{1}{2} \|B\beta - y\|^2 \quad (2.1)$$

其中， $y$  为真实值， $B\beta$  为拟合值。由前所述，为了得到  $GM(1,1,t^\alpha)$  模型和  $FAGM(1,1,t^\alpha)$  模型中的参数，必须求解一个超定线性方程组(1.5)或者(1.13)。显然(1.5)和(1.13)是一个标准的线性回归。然而，线性回归可能存在过拟合现象，且不能进行变量选择，从而导致模型也较为复杂。其次，矩阵  $\Delta$  和  $B$  还可能会存在多重共线性，以及当变量间的相关系数为-1 或者 1 时，模型将会存在完全多重共线性的问题。

正则化方法是解决过拟合的一个有效方法。常见的正则化方法有 Ridge 回归、Elastic Net 回归以及 Lasso 回归等。其中，Ridge 回归也称作 L2 正则化，Elastic Net 回归也被称作弹性网络，Lasso 回归也称作 L1 正则化。而本文引入 L1 正则化法来优化改进  $FAGM(1,1,t^\alpha)$ ，构建新的  $LFAGM(1,1,t^\alpha)$  模型。该模型与模型  $FAGM(1,1,t^\alpha)$  相比，是在原来的损失函数基础上添加了一个参数的 L1 正则化项，即增加了一个惩罚函数以弥补线性回归所带来的不足。新的损失函数如(2.2)所示。

$$\hat{\beta}_{lasso} = \frac{1}{2} \sum_{i=1}^n (B_i \beta - y_i)^2 + \lambda \|\beta\|_1 = \frac{1}{2} \|B\beta - y\|^2 + \lambda \|\beta\|_1 \quad (2.2)$$

其中  $p(\lambda) = \lambda \|\beta\|_1$  ( $\lambda > 0$ ) 为惩罚函数， $\lambda$  为正则参数。

### 3.2. LFAGM(1,1,t<sup>α</sup>)模型的线性参数估计

求解最优化问题(2.2)已经有很多有效的方法,其中坐标下降法为一个简单有效的求解方法。该方法作为一种非梯度优化的方法,目前已被广泛应用于机器学习中求解大规模数据优化问题。在坐标下降法中,每一步的迭代都是沿一个坐标的方向进行搜索,通过循环使用不同的坐标方法来达到目标函数的极小值。对于一个凸函数  $f(x), x \in R^n$ , 该算法的计算步骤总结如下

$$\begin{aligned} x_1^{(k)} &= \arg \min_{x_1} f(x_1, x_2^{(k-1)}, x_3^{(k-1)}, \dots, x_n^{(k-1)}) \\ x_2^{(k)} &= \arg \min_{x_2} f(x_1^{(k)}, x_2, x_3^{(k-1)}, \dots, x_n^{(k-1)}) \\ x_3^{(k)} &= \arg \min_{x_3} f(x_1^{(k)}, x_2^{(k)}, x_3, \dots, x_n^{(k-1)}) \\ x_n^{(k)} &= \arg \min_{x_n} f(x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, \dots, x_n) \end{aligned} \quad (2.3)$$

该算法每一次迭代都会更新参数  $x$  的一个维度。其原理是通过不停的迭代来构造序列  $x^0, x^1, x^2, \dots$  来求解问题。最终收敛到期望的极值点。因此,问题(2.2)的坐标下降迭代格式为

$$\hat{\beta}_i^{lasso} = S_{\lambda/\|B_i\|^2} \left( \frac{B_i^T (Y - B_{-i} \beta_{-i})}{B_i^T B_i} \right) \quad (2.4)$$

其中  $S$  为软阈值函数。

#### 3.2.1 LFAGM(1,1,t<sup>α</sup>)模型的超参数优化

注意到,上述所有的计算都是在分数阶  $r$ , 幂指数  $\alpha$  和正则参数  $\lambda$  已知的情况下进行计算的。但是在实际计算时,模型 LFAGM(1,1,t<sup>α</sup>) 对分数阶  $r$ , 幂指数  $\alpha$  和正则参数  $\lambda$  这三个超参数非常敏感。因此,本文设计了一个优化模型(2.5)来搜索这三个参数的最优值。

$$\begin{aligned} \min_{r, \lambda, \alpha} & \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{x}^{(r)}(k) - x^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\% \\ \text{s.t.} & r > 0, \lambda > 0, \alpha > 0 \end{aligned} \quad (2.5)$$

由于模型(2.5)的具有高度非线性,求解复杂,传统的数值优化方法求解困难甚至难以求解。因此本文使用启发式群智能优化算法—灰狼优化算法(GWO)来求解优化模型(2.5)中的分数阶  $r$ , 幂指数  $\alpha$  和正则参数  $\lambda$ 。

灰狼算法[15],模拟了灰狼在也野外捕食场景。首先将灰狼分为四个阶级  $\theta, \xi, \eta, w$ 。其关系从上到下依次递减如图1。

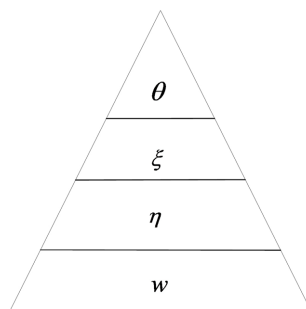


Figure 1. Grey wolf grade system  
图1. 灰狼等级制度



并且灰狼捕食过程可分为三个部分:

- 1) 跟踪, 接近猎物;
- 2) 追逐, 包围猎物逼停猎物;
- 3) 攻击猎物。

### 3.2.2. GWO 算法模型

将灰狼种群中最好的三匹狼(最优解)分别设为  $\theta$ ,  $\xi$  和  $\eta$ , 它们可以引导全狼群对目标进行搜索。而其他的狼群定义为  $w$ , 并且跟随  $\theta$ ,  $\xi$  和  $\eta$  来更新位置。

#### 第 1 步: 包围猎物

在狩猎过程中, 将灰狼的狩猎过程定义为:

$$\bar{D} = |\bar{C} \cdot \bar{X}_p(t) - \bar{X}(t)| \quad (2.6)$$

$$\bar{X}(t+1) = \bar{X}_p(t) - \bar{A} \cdot \bar{D} \quad (2.7)$$

其中(2.6)表示为狼群个体与猎物之间的距离, (2.7)为狼群更行位置。  $t$  为当前迭代数。  $\bar{A}$  和  $\bar{C}$  为系数向量,  $\bar{X}_p$  和  $\bar{X}$  为猎物的位置以及灰狼的位置。  $\bar{A}$  和  $\bar{C}$  计算公式为:

$$\bar{A} = 2\bar{\varepsilon} \cdot \bar{r}_1 - \bar{\varepsilon} \quad (2.8)$$

$$\bar{C} = 2 \cdot \bar{r}_2 \quad (2.9)$$

其中,  $\bar{\varepsilon}$  为收敛因子, 并且会随着迭代次数从 2 减为 0,  $\bar{r}_1$  和  $\bar{r}_2$  为  $[0, 1]$  上的随机数。

#### 第 2 步: 狩猎

当灰狼识别猎物位置之后  $\xi$  和  $\eta$  会在  $\theta$  的领导下包围猎物, 灰狼跟踪猎物的位置可以表示为:

$$\begin{aligned} \bar{D}_\theta &= |\bar{C}_1 \cdot \bar{X}_\theta - \bar{X}| \\ \bar{D}_\xi &= |\bar{C}_2 \cdot \bar{X}_\xi - \bar{X}| \\ \bar{D}_\eta &= |\bar{C}_3 \cdot \bar{X}_\eta - \bar{X}| \end{aligned} \quad (2.10)$$

其中,  $\bar{D}_\theta$ ,  $\bar{D}_\xi$  和  $\bar{D}_\eta$  分别为  $\theta$ ,  $\xi$  和  $\eta$  与其他个体间的距离。  $\bar{X}_\theta$ ,  $\bar{X}_\xi$  和  $\bar{X}_\eta$  为  $\theta$ ,  $\xi$  和  $\eta$  的当前位置, 且  $\bar{C}_1$ ,  $\bar{C}_2$  和  $\bar{C}_3$  为随机向量,  $\bar{X}$  为灰狼当前位置。

$$\begin{aligned} \bar{X}_1 &= \bar{X}_\theta - A_1 \cdot (\bar{D}_\theta) \\ \bar{X}_2 &= \bar{X}_\xi - A_2 \cdot (\bar{D}_\xi) \\ \bar{X}_3 &= \bar{X}_\eta - A_3 \cdot (\bar{D}_\eta) \end{aligned} \quad (2.11)$$

$$\bar{X}(t+1) = \frac{\bar{X}_1 + \bar{X}_2 + \bar{X}_3}{3} \quad (2.12)$$

其中, (2.8)分别设定了狼群中  $w$  的朝向, 以及  $\theta$ ,  $\xi$  和  $\eta$  前进步长和方向。公式(2.12)为  $w$  的最终位置。

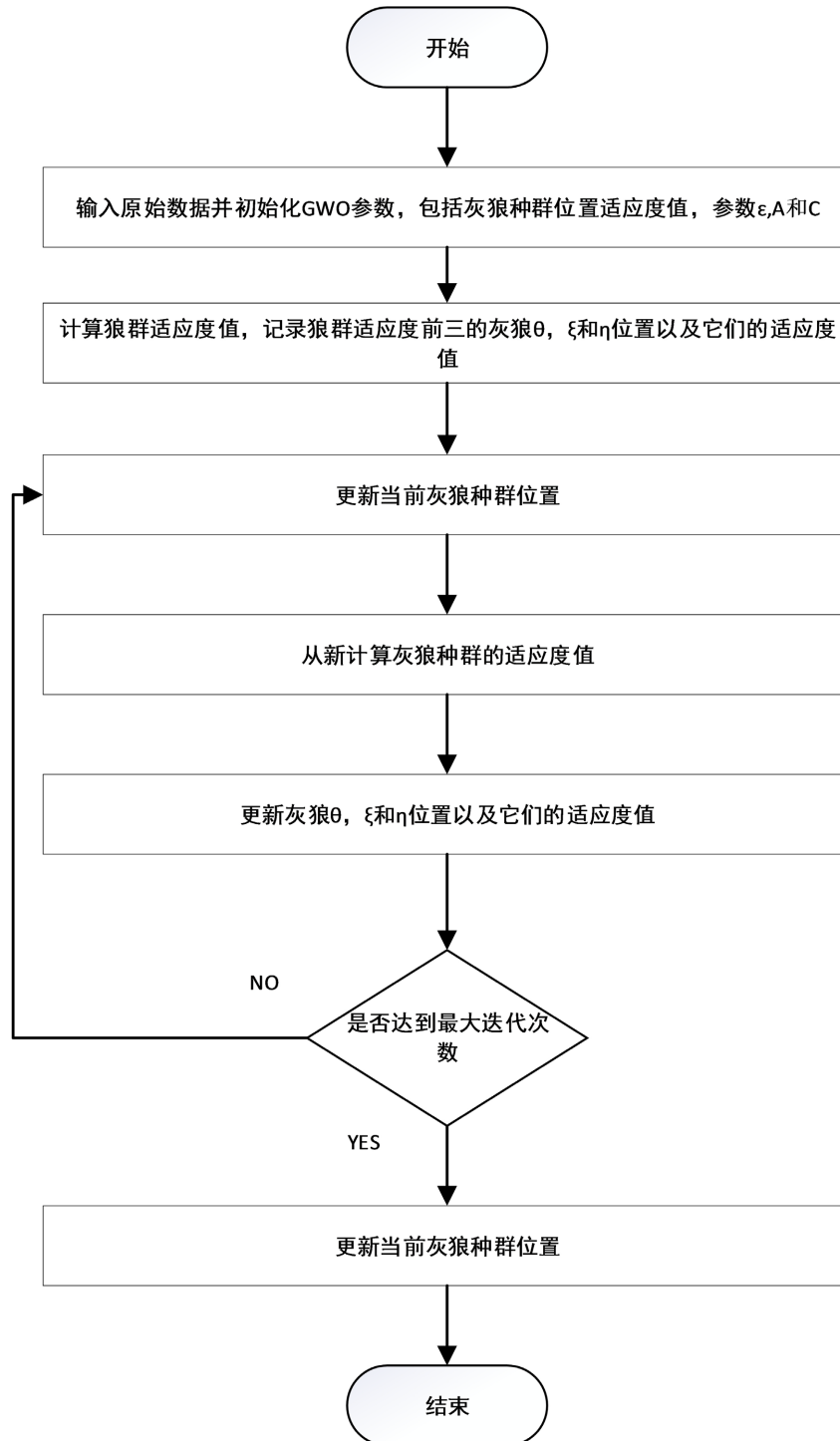
#### 第 3 步: 攻击猎物

猎物, 停止移动之后, 灰狼便会开始攻击,  $\bar{\varepsilon}$  收敛因子的值会逐渐减小, 所以  $A$  的波动范围也会随之减小。在迭代的过程中当  $\bar{\varepsilon}$  的值从 2 下降到 0 时, 而  $\bar{A}$  的值会在  $[-\bar{\varepsilon}, \bar{\varepsilon}]$  之间变化, 并且当  $|\bar{A}| < 1$  时狼群会发起攻击(局部最优), 当  $|\bar{A}| > 1$  时, 狼群会脱离猎物, 寻找更合适的猎物(全局最优)。



其中由公式(2.9)可知,  $\bar{C}$  为[0, 2]之间的随机数, 表示为狼群所在位置对猎物的影响的随机权重, 当  $C > 1$  时表示权重重大,  $C < 1$  时表示权重小。这样可以避免算法更加随机探索, 同时避免陷入局部最优。

GWO 算法流程图如图 2 所示。



**Figure 2.** Flow chart of grey wolf optimization algorithm  
**图 2.** 灰狼优化算法流程图

#### 4. 实例分析

为验证  $LFAGM(1,1,t^\alpha)$  模型的有效性和实用性, 本文分别使用  $FAGM(1,1,t^\alpha)$  模型和  $LFAGM(1,1,t^\alpha)$  模型对中国农业耕地灌溉面积情况进行模拟预测, 数据来源于《中国统计年鉴》(2009~2020年)。其中, 前7组为训练集, 后5组为测试集, 相关计算结果见表1。对于  $FAGM(1,1,t^\alpha)$  模型, 灰狼优化算法得到的最优非线性参数  $\alpha = 6.364$ , 累加阶数  $r = 0.026$ , 参数向量  $\beta = (0.167, -0.03, 1.229 \times 10^4)^T$ , 对于  $LFAGM(1,1,t^\alpha)$  模型, 灰狼优化算法得到的最优非线性参数  $\alpha = 0.728$ , 累加阶数  $r = 0.010$ , 正则化参数  $\lambda = 74.575$ , 参数向量  $\beta = (-0.028, -186.322, 0)^T$ 。

**Table 1.** Accuracy comparison of  $FAGM(1,1,t^\alpha)$  model and  $LFAGM(1,1,t^\alpha)$  model

**表 1.**  $FAGM(1,1,t^\alpha)$  模型以及  $LFAGM(1,1,t^\alpha)$  模型精度比较

时间	耕地灌溉面积 (千公顷)		$FAGM(1,1,t^\alpha)$		$LFAGM(1,1,t^\alpha)$	
	真实值	模拟值	相对误差/%	模拟值	相对误差/%	
2008年	58,471.7	58,471.7	0.0000	58,471.7	0.0000	
2009年	59,261.4	59,261.4	0.0000	59,300.7	-0.0007	
2010年	60,347.7	60,422.7	-0.0012	60,346.9	0.0000	
2011年	61,681.6	61,526.2	0.0025	61,424.2	0.0042	
2012年	62,490.5	62,528.8	-0.0006	62,490.5	0.0000	
2013年	63,473.3	63,496.0	-0.0004	63,531.1	-0.0009	
2014年	64,539.5	64,597.1	-0.0009	64,539.6	0.0000	
平均相对误差%			0.0008		0.0008	
2015年	65,872.6	66,156.6	-0.0043	65,512.8	0.0055	
2016年	67,140.6	68,740.7	-0.0238	66,449.0	0.0103	
2017年	67,815.6	73,275.9	-0.0805	67,347.3	0.0069	
2018年	68,271.6	81,203.0	-0.1894	68,207.1	0.0009	
2019年	68,678.6	94,669.2	-0.3784	69,027.9	-0.0051	
平均相对误差%			0.1353		0.0057	

由表1可见, 经过L1正则化后的  $LFAGM(1,1,t^\alpha)$  模型在测试组预测精度上会更高一些, 可见正则化可使灰色预测模型具有更高的预测精度, 能更好的避免过拟合问题, 提高模型的泛化性能。同时, 该数值实例也验证了  $LFAGM(1,1,t^\alpha)$  模型有更好的实用性。

#### 5. 小结

本文将借助 Lasso 的思想, 将 L1 正则化方法考虑到灰色时间幂预测模型。然而, 基于一阶累加的  $GM(1,1,t^\alpha)$  模型的拟合精度和预测精度都没有分数阶累加高。因此, 在  $GM(1,1,t^\alpha)$  模型的基础上, 本文将分数阶累加添加到  $GM(1,1,t^\alpha)$  模型的建模过程中, 并基于 L1 正则化方法, 构建新的灰色  $LFAGM(1,1,t^\alpha)$  模型。该模型在实际预测中具有更高的预测精度。可见, L1 正则化和分数阶累加进一

步提高了基于一阶累加的  $GM(1,1,t^\alpha)$  模型的预测精度。然而, 该方法是基于短期数据的预测处理, 对于大量规模的数据处理还有待于进一步验证有效性和实用性。

## 基金项目

国家自然科学基金项目“数据驱动的高可解释迁移模糊系统预测建模新方法及应用研究”(71961001)。

## 参考文献

- [1] 邓聚龙. 灰预测与灰决策[M]. 武汉: 华中科技大学出版社, 2002.
- [2] Ding, S. (2018) A Novel Self-Adapting Intelligent Grey Model for Forecasting China's Natural-Gas Demand. *Energy*, **162**, 393-407. <https://doi.org/10.1016/j.energy.2018.08.040>
- [3] 王义闹. GM(1,1)逐步优化直接建模方法的推广[J]. 系统工程理论与实践, 2003, 23(2): 120-124.
- [4] Wang, J., Dang, Y., Ye, J., Xu, N. and Wang, J. (2018) An Improved Grey Prediction Model Based on Matrix Representations of the Optimized Initial Value. *The Journal of Grey System*, **30**, 143-156.
- [5] Liu, X.M. and Xie, N.M. (2019) A Nonlinear Grey Forecasting Model with Double Shape Parameters and Its Application. *Applied Mathematics and Computation*, **360**, 203-212. <https://doi.org/10.1016/j.amc.2019.05.012>
- [6] Zhao, H. and Guo, S. (2016) An Optimized Grey Model for Annual Power Load Forecasting. *Energy*, **107**, 272-286. <https://doi.org/10.1016/j.energy.2016.04.009>
- [7] Lee, Y.S. and Tong, L.I. (2011) Forecasting Energy Consumption Using a Grey Model Improved by Incorporating Genetic Programming. *Energy Conversion and Management*, **52**, 147-152. <https://doi.org/10.1016/j.enconman.2010.06.053>
- [8] Wen, K.L. (2002) Grey Prediction Theory and Application. OpenTech Company, Taipei, 5-7.
- [9] 刘解放, 刘思峰, 吴利丰, 方志耕. 分数阶反向累加离散灰色模型及其应用研究[J]. 系统工程与电子技术, 2016, 38(3): 719-724.
- [10] 谢波, 肖东升. 利用多期均值方法提高新陈代谢 GM(1,1)模型的预测精度[J]. 南阳师范学院报, 2020, 19(4): 16-20.
- [11] 陈超英. 累加生成的改进和 GM(1,1,t)灰色模型[J]. 数学的实践与认识, 2007, 37(2): 105-109.
- [12] 钱吴永, 党耀国, 刘思峰. 含时间幂次项的灰色 GM(1,1,t^\alpha)模型及其应用[J]. 系统工程理论与实践, 2012, 32(10): 2247-2252.
- [13] Wu, L., Liu, S., Yao, L., Yan, S. and Liu, D. (2013) Grey System Model with the Fractional Order Accumulation. *Communications in Nonlinear Science and Numerical Simulation*, **18**, 1775-1785. <https://doi.org/10.1016/j.cnsns.2012.11.017>
- [14] Wu, W., Ma, X., Zhang, Y., Wang, Y. and Wu, X. (2019) Analysis of Novel FAGM(1,1,t^\alpha) Model to Forecast Health Expenditure of China. *Grey Systems: Theory and Application*, **9**, 232-250. <https://doi.org/10.1108/GS-11-2018-0053>
- [15] 芦方旭, 米志超, 李艾静, 王海, 田雨露. 基于灰狼算法的无人机基站三维空间优化部署[J]. 兵器装备工程学报, 2021, 42(7): 185-189.