

# 控制与决策

Control and Decision

## 基于鞅论的灰狼优化算法全局收敛性分析

孙丽君, 冯斌斌, 陈天飞

引用本文:

孙丽君,冯斌斌,陈天飞. 基于鞅论的灰狼优化算法全局收敛性分析[J]. *控制与决策*, 2022, 37(11): 2839–2848.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.0645>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 多策略融合的改进麻雀搜索算法及其应用

Improved sparrow search algorithm with multi-strategy integration and its application

*控制与决策*. 2022, 37(1): 87–96 <https://doi.org/10.13195/j.kzyjc.2021.0582>

#### 基于小波精英解学习和多角度搜索的新型阴阳平衡优化算法

A novel Yin–Yang pair optimization algorithm based on wavelet elite solutions learning and multi-angle search

*控制与决策*. 2022, 37(8): 1962–1970 <https://doi.org/10.13195/j.kzyjc.2021.0251>

#### 基于自适应正态云模型的灰狼优化算法

Grey wolf optimization algorithm based on adaptive normal cloud model

*控制与决策*. 2021, 36(10): 2562–2568 <https://doi.org/10.13195/j.kzyjc.2020.0233>

#### 一种自适应拟牛顿-状态转移混合智能优化算法及应用

A hybrid state transition optimization algorithm based on adaptive quasi-newton method and its application

*控制与决策*. 2021, 36(10): 2451–2458 <https://doi.org/10.13195/j.kzyjc.2020.0214>

#### 求解约束优化问题的改进果蝇优化算法及其工程应用

Improved fruit fly optimization algorithm for solving constrained optimization problems and engineering applications

*控制与决策*. 2021, 36(2): 314–324 <https://doi.org/10.13195/j.kzyjc.2019.0557>

# 基于鞅论的灰狼优化算法全局收敛性分析

孙丽君<sup>†</sup>, 冯斌斌, 陈天飞

(1. 河南工业大学 粮食信息处理与控制教育部重点实验室, 郑州 450001; 2. 河南工业大学 郑州市机器感知与智能系统重点实验室, 郑州 450001; 3. 河南工业大学 信息科学与工程学院, 郑州 450001)

**摘要:** 灰狼优化 (grey wolf optimization, GWO) 算法是一种基于群体智能的随机优化算法, 已成功地应用于许多复杂的优化问题的求解. 尽管 GWO 算法有很多改进形式, 但缺少严谨的收敛性分析, 导致改进后的算法不具备理论支撑. 对此, 运用鞅论分析其收敛性. 首先, 根据 GWO 算法原理建立其基本的数学模型, 通过定义灰狼状态空间及灰狼群状态空间, 建立 GWO 算法的 Markov 链模型, 并分析该算法的 Markov 性质; 其次, 介绍鞅论, 推导出一个上鞅作为最优适应度值的群进化序列; 然后, 运用上鞅收敛定理, 并结合其 Markov 性质对 GWO 算法进行收敛性分析, 证明 GWO 算法能以 1 的可能性达到全局收敛; 最后, 通过数值实验验证其收敛性能. 实验结果表明, GWO 算法具有全局收敛性强、计算耗时较低、寻优精度高等特点.

**关键词:** 灰狼优化算法; Markov 链; 鞅论; 状态转移; 数值实验; 全局收敛

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2021.0645

开放科学(资源服务)标识码(OSID):



**引用格式:** 孙丽君, 冯斌斌, 陈天飞. 基于鞅论的灰狼优化算法全局收敛性分析[J]. 控制与决策, 2022, 37(11): 2839-2848.

## Global convergence analysis of grey wolf optimization algorithm based on martingale theory

SUN Li-jun<sup>†</sup>, FENG Bin-bin, CHEN Tian-fei

(1. Key Laboratory of Grain Information Processing and Control of Ministry of Education, Henan University of Technology, Zhengzhou 450001, China; 2. Zhengzhou Key Laboratory of Machine Perception and Intelligent System, Henan University of Technology, Zhengzhou 450001, China; 3. College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450001, China)

**Abstract:** The grey wolf optimization (GWO) algorithm is a stochastic optimization algorithm based on swarm intelligence which has been successfully used to solve many complex optimization problems. At present, there are many improved forms of the GWO algorithm, but the lack of rigorous convergence analysis leads to no theoretical support for the improved algorithm. In order to make up for this deficiency, the martingale theory is used to analyze its convergence for the first time. Firstly, the basic mathematical model is established according to the principle of the GWO algorithm. By defining the gray wolf state space and the gray wolf group state space, the Markov chain model of the GWO algorithm is established, and the Markov properties of the algorithm are analyzed. Secondly, the martingale theory is introduced, and a swarm evolution sequence with the supermartingale as the optimal fitness value is derived. Thirdly, the convergence of the GWO algorithm is analyzed using the supermartingale convergence theorem and its Markov properties. It is proved that the GWO algorithm can achieve global convergence with the possibility of 1. Finally, the convergence performance is verified by numerical experiments. The experimental results show that the GWO algorithm has strong global convergence, low computation time and high optimization accuracy.

**Keywords:** grey wolf optimization algorithm; Markov chain; martingale theory; state transition; numerical experiments; global convergence

收稿日期: 2021-04-15; 录用日期: 2021-07-30.

**基金项目:** 国家自然科学基金项目(61803146, 61973104); 河南省优秀青年科学基金项目(212300410036); 河南省高校科技创新人才支持计划项目(21HASTIT029); 河南省高等学校青年骨干教师培养计划项目(2019GGJS089); 河南省青年人才托举工程项目(2019HYTP005); 河南省科技攻关项目(212102210169, 212102210086); 河南省高等教育教学改革研究与实践项目(2019SJGLX270); 河南工业大学自科创新基金支持计划项目(2020ZKCJ06); 河南工业大学青年骨干教师培育计划项目(21420080).

责任编辑: 陈家伟.

<sup>†</sup>通讯作者. E-mail: ljsun@haut.edu.cn.

## 0 引言

灰狼优化(grey wolf optimization, GWO)<sup>[1]</sup>算法作为一种仿生智能算法,其灵感来源于自然界中灰狼的社会等级和狩猎行为. GWO算法在灰狼群体合作行为的基础上,模拟了搜索追踪、包围和攻击猎物等行为以达到优化求解的目的. GWO算法与人工蜂群算法(artificial bee colony algorithm, ABC)<sup>[2]</sup>、粒子群优化算法(particle swarm optimization, PSO)<sup>[3]</sup>、蝙蝠算法(bat algorithm, BA)<sup>[4]</sup>等群智能算法相比较,具有控制参数少、鲁棒性好等优点<sup>[5]</sup>.

近年来, GWO算法的相关研究主要集中在改进和应用两个方面<sup>[6]</sup>. 根据不同的应用场景,一些学者对标准的 GWO算法提出了不同的改进形式<sup>[7-9]</sup>. 在文献[7-9]的研究中,采用仿真实验验证了 GWO算法的收敛性能,却未进行相关的理论分析. 了解 GWO算法的理论基础是非常重要的,有助于进一步掌握该算法的工作原理和性能,因此,算法的收敛性理论分析是值得研究的关键问题. 随机搜索算法的收敛准则是分析群体智能优化算法收敛性的常用数学工具,已成功地证明了 ABC算法<sup>[10]</sup>、布谷鸟搜索(cuckoo search, CS)算法<sup>[11]</sup>、BA算法<sup>[12]</sup>等群智能算法的全局收敛性. 同理,基于该收敛准则,文献[13]通过建立 Markov链的数学模型,分析了 GWO算法的收敛性. 然而,利用传统的 Markov链理论进行收敛性理论分析较为复杂. 本文在 GWO算法的收敛性理论分析中创新性地引入鞅理论,有效降低了理论分析过程的复杂性.

鞅是现代概率论中的一个重要概念,同时也是研究随机过程的重要工具. 鞅在算法收敛性及收敛概率方面,具有较强的分析能力. 群智能算法寻优的过程可以视为一种随机过程,因此,利用鞅方法可以有效地对其过程进行理论分析. 鞅作为一类具有普适意义的随机过程,其应用已经十分广泛,目前,鞅论已经成功地应用于思维进化算法(mind evolutionary algorithm, MEA)<sup>[14]</sup>、蚁群算法(ant colony optimization, ACO)<sup>[15]</sup>等群智能算法的收敛性分析. 文献[16]对保留精英遗传算法的收敛性和收敛速度进行了分析;文献[17]运用鞅方法分析了噪声环境下的精英克隆选择算法的全局收敛性. 虽然本文同样采用鞅方法进行收敛性理论分析,但本文的证明过程与文献[14-17]存在一定的差异. 首先,在相关定义及定理证明方式上存在显著不同;其次,本文将灰狼群进化序列转化为有界上鞅,并对 GWO算法收敛的两种情形进行分析证明;最后,通过数值实验对 GWO算法的收

敛性能进行验证,并通过对比直观地衡量算法的性能.

本文所使用的鞅方法避免了传统 Markov链的遍历性分析,并且将具有最佳适应度值的种群进化序列对应于上鞅,有效地简化了 GWO算法的全局收敛性理论分析. 鞅方法为 GWO算法的收敛性理论分析提供了一种新的数学工具,从而为 GWO算法的进一步发展和应用奠定了良好的理论基础.

## 1 GWO算法原理

灰狼优化算法(GWO)是一种典型的仿生智能优化算法,该算法的模型源自于灰狼的领导阶层和狩猎行为. 灰狼大多成群生活,平均种群个体数为5~12个. 该群体具有极为严格的社会统治等级制度,处于不同层级的灰狼有着不同的作用:第1层是 $\alpha$ 狼,  $\alpha$ 狼在群体中起主要作用,主要负责决定狩猎、休息的地点等;第2层是 $\beta$ 狼,作为 $\alpha$ 狼的最佳候选者,  $\beta$ 狼可以帮助 $\alpha$ 狼做决策或进行其他活动,当群体中 $\alpha$ 出现空缺时,  $\beta$ 将接替 $\alpha$ 的位置;第3层的灰狼被称为 $\delta$ 狼,  $\delta$ 狼必须服从 $\alpha$ 狼和 $\beta$ 狼,但它统治着 $\omega$ 狼,同时,适应度差的 $\alpha$ 狼和 $\beta$ 狼会降为 $\delta$ 狼;处于社会最底层的是 $\omega$ 狼,  $\omega$ 狼必须要服从以上3个层级的灰狼.

灰狼的狩猎行为包括以下3个关键步骤:

step 1: 追踪并靠近猎物;

step 2: 包围猎物;

step 3: 攻击猎物.

GWO算法的基本数学模型可根据上述灰狼的社会等级制度和狩猎行为进行建立.  $\alpha$ 狼代表最优解,  $\beta$ 狼和 $\delta$ 狼分别代表次优和第三优解,剩下的候选解记为 $\omega$ . 在 GWO算法中,狩猎过程主要由 $\alpha$ 、 $\beta$ 和 $\delta$ 引导,  $\omega$ 根据 $\alpha$ 、 $\beta$ 和 $\delta$ 的当前位置进行迭代更新.

在捕猎过程中灰狼会追踪并包围猎物,这一行为可用公式表示为

$$D = |C \cdot X_p(t) - X(t)|, \quad (1)$$

$$X(t+1) = X_p(t) - A \cdot D. \quad (2)$$

其中:  $D$ 表示灰狼个体与目标猎物之间的距离;  $t$ 为当前迭代代数;  $X_p$ 为猎物的位置向量;  $X$ 为灰狼当前的位置向量;系数向量 $A$ 、 $C$ 的计算公式如下:

$$A = 2a \cdot r_1 - a, \quad (3)$$

$$C = 2 \cdot r_2. \quad (4)$$

$a$ 为收敛因子,随着迭代次数从2到0线性减小;  $r_1$ 、 $r_2$ 为[0,1]中的随机向量. 根据式(2)对灰狼的位置进行更新.

灰狼追踪目标猎物的公式如下:

$$\begin{cases} D_\alpha = |C_1 \cdot X_\alpha - X|, \\ D_\beta = |C_2 \cdot X_\beta - X|, \\ D_\delta = |C_3 \cdot X_\delta - X|. \end{cases} \quad (5)$$

其中:  $D_\alpha$ 、 $D_\beta$ 、 $D_\delta$  分别表示  $\alpha$ 、 $\beta$  和  $\delta$  与其他个体之间的距离;  $X_\alpha$ 、 $X_\beta$ 、 $X_\delta$  分别表示  $\alpha$ 、 $\beta$  和  $\delta$  各自的位置;  $C_1$ 、 $C_2$ 、 $C_3$  为随机向量;  $X$  是当前灰狼的位置.

$$\begin{cases} X_1 = X_\alpha - A_1 \cdot (D_\alpha), \\ X_2 = X_\beta - A_2 \cdot (D_\beta), \\ X_3 = X_\delta - A_3 \cdot (D_\delta), \end{cases} \quad (6)$$

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3}. \quad (7)$$

其中:  $X_1$  表示  $\omega$  狼向  $\alpha$  靠近的步长和方向;  $X_2$  表示  $\omega$  狼向  $\beta$  靠近的步长和方向;  $X_3$  表示  $\omega$  狼向  $\delta$  靠近的步长和方向; 最后, 根据式 (7) 定义  $\omega$  狼的最终位置.

通过上述 GWO 算法的基本数学模型, 结合灰狼种群的社会等级制度和狩猎行为, 得出 GWO 算法的流程, 如图 1 所示.

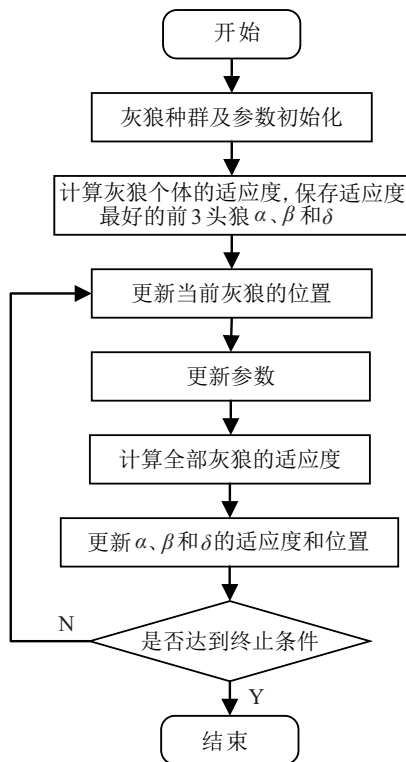


图 1 灰狼优化算法流程

## 2 GWO 算法的 Markov 链模型

GWO 算法是一种基于群体智能的随机优化算法. 灰狼群在捕猎过程中的一系列行为, 可以利用 Markov 链对这一随机过程进行表示. 下面给出一些数学描述和定义, 对 GWO 算法的 Markov 链模型进行描述.

**定义 1** 依据 GWO 算法, 设灰狼个体状态为  $X$ , 其中  $X \in Y$ ,  $Y$  表示可行解空间, 则灰狼个体状态空间可以表示为  $X = \{X | X \in Y\}$ . 因灰狼群状态是由所有灰狼个体状态组成的, 故灰狼群状态表示为

$$\theta = (X_1, X_2, \dots, X_i), 1 \leq i \leq N_\theta.$$

其中:  $X_i$  表示第  $i$  个灰狼的状态,  $N_\theta$  为种群大小. 相应地, 灰狼的群状态空间表示为

$$\Theta = \{\theta = (X_1, X_2, \dots, X_i) | X_i \in Y, 1 \leq i \leq N_\theta\}.$$

**定义 2** GWO 算法迭代过程中, 对于任意两个个体状态  $X_i, X_j \in \theta$ , 灰狼个体状态  $X_i$  一步转移到另一种状态  $X_j$ , 记为  $T_\theta(X_i) = X_j$ .

**定理 1** GWO 算法中, 灰狼个体状态从  $X_i$  转移至  $X_j$  的一步转移概率  $P(T_\theta(X_i) = X_j)$  可由下式计算:

$$P(T_\theta(X_i) = X_j) = \begin{cases} P_\alpha(T_\theta(X_i) = X_j), & \text{由 } \alpha \text{ 狼实现;} \\ P_\beta(T_\theta(X_i) = X_j), & \text{由 } \beta \text{ 狼实现;} \\ P_\delta(T_\theta(X_i) = X_j), & \text{由 } \delta \text{ 狼实现;} \\ P_\omega(T_\theta(X_i) = X_j), & \text{由 } \omega \text{ 狼实现.} \end{cases} \quad (8)$$

证明过程详见文献 [13].

GWO 算法是灰狼群通过  $\alpha$ 、 $\beta$ 、 $\delta$  和  $\omega$  四种灰狼的角色转换实现的, 所以 GWO 算法的一步转移概率  $P(T_\theta(X_i) = X_j)$  由 4 种灰狼共同决定.

**定义 3** 在 GWO 算法寻优过程中, 对于灰狼的任意两个群状态  $\theta_i, \theta_j \in \Theta$ , 从  $\theta_i$  一步转移至  $\theta_j$ , 记作  $T_\Theta(\theta_i) = \theta_j$ .

**定理 2** 在 GWO 算法中, 灰狼群状态从  $\theta_i$  一步转移至  $\theta_j$  的转移概率为

$$P(T_\Theta(\theta_i) = \theta_j) = \prod_{m=1}^{N_\theta} P(T_\theta(X_{im}) = X_{jm}). \quad (9)$$

**证明** 灰狼群状态包含所有灰狼的位置状态. 当灰狼群状态从  $\theta_i$  转移到  $\theta_j$  时, 表示群状态中所有个体的位置状态都会同时转移, 即  $T_\theta(X_{i1}) = X_{j1}$ ,  $T_\theta(X_{i2}) = X_{j2}, \dots, T_\theta(X_{im}) = X_{jm}$  同时成立, 则灰狼群状态从  $\theta_i$  一步转移至  $\theta_j$  的转移概率为

$$P(T_\Theta(\theta_i) = \theta_j) =$$

$$P(T_\theta(X_{i1}) = X_{j1}) \cdot P(T_\theta(X_{i2}) = X_{j2}) \cdot \dots \cdot$$

$$P(T_\theta(X_{im}) = X_{jm}) =$$

$$\prod_{m=1}^{N_\theta} P(T_\theta(X_{im}) = X_{jm}). \quad \square$$

**定理 3** GWO 算法中, 灰狼群状态序列  $\{\theta(t) : t > 0\}$  是有限齐次 Markov 链.

**证明** 1) 对于任何优化算法, 种群个体数与维度

均为有限正整数,个体位置的转移受相应上下限约束,且其状态空间为可列集,则任何优化算法的搜索空间都是有限的<sup>[10,13]</sup>. 因此,在GWO算法中任一灰狼状态  $X_i$  也是有限的,而且单个个体的状态空间有限且离散. 群状态空间  $\theta = (X_1, X_2, \dots, X_{N_\theta})$  是由  $N_\theta$  个灰狼组成,  $N_\theta$  为有限正整数,所以灰狼群状态空间  $\Theta$  也是有限的.

2) 由定理2得到群状态序列  $\{\theta(t) : t > 0\}$ , 其中, 对于  $\forall \theta(t-1) \in \Theta, \theta(t) \in \Theta$ , 它们之间的转移概率为  $P(T_\theta(\theta(t-1)) = \theta(t))$ . 由定理1可知, 灰狼群内任一灰狼的转移概率  $P(T_\theta(X(t-1)) = X(t))$  仅与  $t-1$  时刻的状态  $X(t-1)$  有关, 所以一步转移概率  $P(T_\theta(\theta(t-1)) = \theta(t))$  也仅与  $t-1$  时刻的状态有关, 即灰狼群状态序列  $\{\theta(t) : t > 0\}$  具有Markov性. 又因为状态空间为可列集, 故灰狼群状态序列构成一个有限Markov链.

3) 由定理1可知, 一步状态转移概率  $P(T_\theta(X(t-1)) = X(t))$  仅与  $t-1$  时刻的状态  $X(t-1)$  有关, 而与  $t-1$  无关, 所以灰狼群状态序列  $\{\theta(t) : t > 0\}$  是有限齐次Markov链.  $\square$

**定义4** 文献[18]把  $X_1^t, X_2^t, \dots, X_i^t (i \leq N_\theta)$  产生的  $\sigma$  域记作  $S_1^t$ , 且定义

$$R(S_1^t, X_{i+1}^t) = \sup |P(B|A) - P(B)|, \quad A \in S_1^t, B \in \sigma(X_{i+1}^t); \quad (10)$$

$$R = \sup_{1 \leq i \leq N_\theta - 1} R(S_1^t, X_{i+1}^t). \quad (11)$$

**定理4** 在Markov链  $\{\theta^t, t > 0\}$  中, 灰狼群状态由  $\theta^t \rightarrow \theta^{t+1}$  的转移概率满足

$$\prod_{i=1}^{N_\theta} P(X_i^{t+1}|X_i^t) - \mu \leq P(\theta^{t+1}|\theta^t) \leq \prod_{i=1}^{N_\theta} P(X_i^{t+1}|X_i^t) + \mu, \quad (12)$$

其中  $\mu = 2^{N_\theta-1}R$ . 因为灰狼个体之间的相互依赖关系很弱, 所以式(11)中  $R$  是足够小的, 故  $\mu$  也是足够小的.

**证明** 根据定义4, 可得

$$|P(B|A) - P(B)| \leq R \Rightarrow P(B) - R \leq P(B|A) \leq P(B) + R. \quad (13)$$

事件  $(X_i^{t+1}|X_i^t)$  表示灰狼  $i$  第  $t$  次迭代时的状态转移至第  $t-1$  次迭代时的状态, 把事件  $(X_i^{t+1}|X_i^t)$  记作  $A_i (i = 1, 2, \dots, N_\theta)$ , 则  $P(A_1, A_2, \dots, A_{N_\theta}) = P(\theta^{t+1}|\theta^t)$  表示从  $\theta^t \rightarrow \theta^{t+1}$  的转移概率. 由于事件  $A_1, A_2, \dots, A_{N_\theta}$  不是相互独立的, 结合式(13)和条件概率可得

$$P(A_1 A_2) = P(A_1)P(A_2|A_1) \leq P(A_1)[P(A_2) + R] \leq$$

$$P(A_1)P(A_2) + R, \\ P(A_1 A_2) = P(A_1)P(A_2|A_1) \geq P(A_1)[P(A_2) - R] \geq P(A_1)P(A_2) - R.$$

即  $P(A_1)P(A_2) - R \leq P(A_1 A_2) \leq P(A_1)P(A_2) + R$ . 因为

$$P(A_1 A_2 A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 A_2) \leq P(A_1)[P(A_2) + R][P(A_3) + R] \leq P(A_1)P(A_2)P(A_3) + 3R, \\ P(A_1 A_2 A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 A_2) \geq P(A_1)[P(A_2) - R][P(A_3) - R] \geq P(A_1)P(A_2)P(A_3) - 3R,$$

所以可得

$$P(A_1)P(A_2)P(A_3) - 3R \leq P(A_1 A_2 A_3) \leq P(A_1)P(A_2)P(A_3) + 3R.$$

进而归纳为

$$\prod_{i=1}^{N_\theta} P(A_i) - 2^{N_\theta-1}R \leq P(A_1 A_2 \dots A_{N_\theta}) \leq \prod_{i=1}^{N_\theta} P(A_i) + 2^{N_\theta-1}R.$$

令  $\mu = 2^{N_\theta-1}R$ , 得到如下不等式:

$$\prod_{i=1}^{N_\theta} P(X_i^{t+1}|X_i^t) - \mu \leq P(\theta^{t+1}|\theta^t) \leq \prod_{i=1}^{N_\theta} P(X_i^{t+1}|X_i^t) + \mu.$$

由式(1)、(2)可知, 灰狼个体与个体之间的相互依赖关系很弱, 而且  $R$  是足够小的,  $N_\theta$  是有限实数, 所以  $\mu$  是一个足够小的正数.  $\square$

### 3 鞅论

应用群体智能算法求解优化问题的过程是一个迭代寻优的过程. 在理想情况下, 随着迭代次数的增加, 所得解的质量会越来越好, 当迭代次数趋于无穷时, 将收敛到全局最优. 上述情况可以转化为一类极限问题. 对于此类问题, 运用鞅的极限定理可以很好地对其进行分析. 鞅的极限定理是将独立随机变量的中心极限定理推广到鞅, 进而得到相应的结果<sup>[19]</sup>. 极限定理是有关随机变量序列或随机过程序列的收敛性理论. 鞅论是在概率测度的基础上, 用概率方法将群体进化序列转换成鞅.

下面给出一些相关的数学描述和定义<sup>[20]</sup>.

**定义 5** 若随机过程  $\{Y^k, k \geq 0\}$  满足下列条件, 则称随机过程  $\{Y^k, k \geq 0\}$  为鞅:

- 1)  $E(|Y^k|) < \infty$ ;
- 2)  $E(Y^{k+1}|Y^0, Y^1, \dots, Y^k) = Y^k$ .

**定义 6** 若随机过程  $\{Y^k, k \geq 0\}$  和另一随机过程  $\{Z^k, k \geq 0\}$  满足下列条件, 则称随机过程  $\{Y^k, k \geq 0\}$  具有关于  $\{Z^k, k \geq 0\}$  的鞅:

- 1)  $E(|Y^k|) < \infty$ ;
- 2)  $E(Y^{k+1}|Z^0, Z^1, \dots, Z^k) = Y^k$ .

**定义 7** 若随机过程  $\{Y^k, k \geq 0\}$  和另一随机过程  $\{Z^k, k \geq 0\}$  满足下列条件, 则称随机过程  $\{Y^k, k \geq 0\}$  是关于  $\{Z^k, k \geq 0\}$  的上鞅(或下鞅):

- 1)  $E(|Y^k|) < \infty$ ;
- 2)  $E(Y^{k+1}|Z^0, Z^1, \dots, Z^k) \leq Y^k$ , 或  $(E(Y^{k+1}|Z^0, Z^1, \dots, Z^k) \geq Y^k)$ ;
- 3)  $Y^k$  是  $Z^0, Z^1, \dots, Z^k$  的函数.

文献 [20-21] 已经证明了下鞅收敛定理. 为了便于分析, 给出其结论.

**引理 1** 若随机过程  $\{Y^k, k \geq 0\}$  是一个下鞅,  $\sup_{k \geq 0} E(|Y^k|) < \infty$ , 则必定存在一个随机变量  $Y^\infty \in \{Y^k, k \geq 0\}$ , 使得  $E(|Y^\infty|) < \infty$ , 且  $\{Y^k, k \geq 0\}$  以概率 1 的可能性收敛. 即当  $E(|Y^\infty|) < \infty$  时,  $P\{\lim_{k \rightarrow \infty} Y^k = Y^\infty\} = 1$ .

**引理 2** 同引理 1, 其结论对于上鞅依然成立.

## 4 GWO 的收敛性分析

### 4.1 基础理论

**定义 8** 对于状态空间  $X$ , 当且仅当  $P_{ij} = 0$  成立时 ( $\forall i \in Q, \forall j \notin Q$ ),  $X$  的非空子集  $Q$  是一个闭集. 即从集合  $Q$  中任意一点出发, 不会到达  $Q$  以外的任何状态<sup>[20]</sup>.

**定义 9** 如果闭集  $Q$  不包含任何真闭子集, 则称为  $Q$  不可约; 否则, 称  $Q$  是可约的<sup>[20]</sup>.

**命题 1** 灰狼群位置状态空间  $\Theta$  是一个闭集.

**证明** 根据定理 3, 群状态序列  $\{\theta(t) : t > 0\}$  是有限齐次马尔可夫链, 且  $\forall i \in \Theta, \sum_{j \in \Theta} P_{ij} = 1$ . 结合定义 8, 易知  $\Theta$  是一个闭集.  $\square$

**定义 10** 设优化问题  $\langle Y, f \rangle$  的全局最优解是  $g^*$ , 灰狼群最优状态集定义如下:

$$M = \{\theta^* = (X_1, X_2, \dots, X_{N_\theta}) | f(X) = f(g^*), \theta \in \Theta\}. \quad (14)$$

如果  $M = \Theta$ , 则在可行解的空间中的任何一个解都是最优解, 此时该优化问题没有意义, 所以下文都是在  $M \in \Theta$  的情形下讨论.

**命题 2** 灰狼最优位置状态集  $M$  是群状态空间  $\Theta$  上的一个闭集.

**证明** 设  $\forall \theta_i, \theta_j \in M$ , 对于任意转移步长  $l (l \geq 1)$ , 由 Chapman-Kolmogorov 方程得

$$P_{\theta_i, \theta_j}^l = \sum_{\theta_{r1} \in \Theta} \dots \sum_{\theta_{rl} \in \Theta} P(T_\theta(\theta_i) = \theta_{r1}) \dots P(T_\theta(\theta_{r1}) = \theta_{r2}) \dots P(T_\theta(\theta_{rl-1}) = \theta_j), \quad (15)$$

其中  $P_{\theta_i, \theta_j}^l$  为灰狼群位置状态  $\theta_i$  经过  $l$  步转移至群状态  $\theta_j$  的概率. 式 (15) 展开式中的每一项表达式都存在  $P(T_\theta(\theta_{rs-1}) = \theta_{rs})$  项 ( $s \in [1, l]$ ), 满足  $\theta_{rs-1} \in M$  且  $\theta_{rs} \notin M$ , 由定理 2 可得转移概率为

$$P(T_\theta(\theta_{rs-1}) = \theta_{rs}) = \prod_{k=1}^{N_\theta} P(T_\theta(X_{ik}) = X_{jk}). \quad (16)$$

由  $\theta_{rs-1} \in M$  且  $\theta_{rs} \notin M$ , 可得  $f(X_s) > f(X_{s-1}) = f(g^*) = \inf(f(y)), y \in Y$ , 则  $\exists P(T_\theta(\theta_{rs-1}) = \theta_{rs}) = 0$ , 此时  $P_{\theta_i, \theta_j}^l = 0$ , 所以灰狼最优位置状态集  $M$  是群状态空间  $\Theta$  上的一个闭集.  $\square$

**命题 3** 灰狼群状态空间  $\Theta$  是可约的.

**证明** 由定义出发, 若群状态空间  $\Theta$  是可约的, 则  $\Theta$  应至少含有一个除整个状态空间外的闭集. 整个群状态空间  $\Theta$  构成了一个最大的闭集. 由定义 10 可知, 由于灰狼最优群状态集  $M$  是一个闭集且  $M \in \Theta$ , 再根据定义 9, 可得群状态空间  $\Theta$  是可约的. 当群体状态进入最优状态集  $M$  时, 将会收敛到全局最优.  $\square$

**命题 4** 当且仅当  $P\{\lim_{t \rightarrow \infty} (\theta^t \in M)\} = 1$  时, 灰狼群状态序列  $\{\theta^t, t \geq 1\}$  以概率 1 的可能性收敛至灰狼最优位置状态集  $M$ <sup>[20]</sup>.

### 4.2 鞅方法分析 GWO 算法收敛性

GWO 算法的目标是寻找与最优灰狼群状态相对应的全局最优解. 为此, 首先对灰狼的位置进行更新; 其次, 群状态随机变量在群状态空间  $\Theta$  中取值, 根据目标函数计算群适应度值; 然后, 根据 GWO 算法迭代策略, 在寻优过程中获得最佳的适应度值; 最后, 由最优适应度值确定全局最优解. 通过以上分析, 最佳适应度值可以记作  $\theta^t$ . 由函数  $F(\theta^t)$  的值来表示群状态的最佳适应度值, 其定义为

$$F(\theta^t) = f(g^t).$$

不失一般性, 本文考虑最小化问题, 利用鞅论对 GWO 算法的收敛性进行分析. 对于该类问题, 当且仅当找到全局最优群状态时, 才能得到全局最小适应度值. 通过研究函数  $F(\theta^t)$  的收敛性来反映 GWO 算法的收敛性. 如果  $F(\theta^t)$  的条件期望小于或等于当前群状态的值, 则  $F(\theta^t)$  可以转化为上鞅, 并进一步用于分析 GWO 算法的收敛性. 因为  $F(\theta^t)$  是单调非递增的函数, 所以可将随机过程  $\{F(\theta^t), t \geq 1\}$  转化为上鞅

来检验  $\{\theta^t, t \geq 1\}$  的收敛性.

**命题5** 随机过程  $\{F(\theta^t), t \geq 1\}$  是一个关于  $\{\theta^t, t \geq 1\}$  的有界上鞅, 记为  $E(F(\theta^{t+1})|\theta^1, \theta^2, \dots, \theta^t, t \geq 1) \leq F(\theta^t)$ . 当  $t \rightarrow \infty$  时, 必定存在一个群随机变量  $\theta^\infty \in \theta^t, t \geq 1$ , 其群状态序列  $\{\theta^t, t \geq 1\}$  收敛, 即  $P\{\lim_{t \rightarrow \infty} (\theta^t = \theta^\infty)\} = 1$ .

**证明** 序列  $\{F(\theta^t), t \geq 1\}$  的有界性是显然的. 由定理3中的马尔可夫性可得

$$E(F(\theta^{t+1})|\theta^1, \theta^2, \dots, \theta^t) = E(F(\theta^{t+1})|\theta^t). \quad (17)$$

仅需要证明

$$\forall u \in \Theta, E(F(\theta^{t+1})|\theta^t = u) \leq F(u). \quad (18)$$

式(18)中, 不等号左边可以转化为

$$E(F(\theta^{t+1})|\theta^t = u) = \sum_{v \in \Theta} F(v)P(\theta^{t+1} = v|\theta^t = u), \quad (19)$$

于是需要证明的不等式(18)转化为

$$\sum_{v \in \Theta} F(v)P(\theta^{t+1} = v|\theta^t = u) \leq F(u). \quad (20)$$

需要注意的是, 对于任何群状态  $u$  之后的群状态  $v$  均存在以下不等式:

$$F(\theta^{t+1} = v) \leq F(\theta^t = u), \forall v \in \Theta.$$

所以不等式(20)又可以转化为

$$\begin{aligned} \sum_{v \in \Theta} F(v)P(\theta^{t+1} = v|\theta^t = u) &\leq \\ \sum_{v \in \Theta} F(u)P(\theta^{t+1} = v|\theta^t = u) &= \\ F(u) \sum_{v \in \Theta} P(\theta^{t+1} = v|\theta^t = u) &= F(u). \end{aligned} \quad (21)$$

因为  $\Theta$  是一个闭集, 所以式(21)成立.

综上所述, 证明随机过程  $\{F(\theta^t), t \geq 1\}$  是有界上鞅.

根据引理2, 当  $t \rightarrow \infty, \theta^\infty \in \{\theta^t, t \geq 1\}$  时, 随机过程  $\{F(\theta^t), t \geq 1\}$  将以1的可能性收敛至  $F(\theta^\infty)$ . 相应地, 当  $t \rightarrow \infty$  时, 群状态序列  $\{\theta^t, t \geq 1\}$  将以1的可能性收敛至  $\theta^\infty$ , 即  $P\{\lim_{t \rightarrow \infty} (\theta^t = \theta^\infty)\} = 1$ .  $\square$

根据命题5, 如果  $\theta^\infty$  的取值在集合  $M$  中, 则GWO将以1的概率收敛到全局最优值. 但如果  $\theta^\infty$  的取值不在  $M$  集合中, 则GWO可能不会收敛到全局最优.

**命题6** 在上述命题5的基础上, 令  $P_t^* = \min P(\theta^{t+1}|\theta^t), t \geq 1$ , 可以得到:

- 1) 当  $\sum_{t=1}^{\infty} P_t^* = \infty$  时,  $P\{\lim_{t \rightarrow \infty} (\theta^t \in M)\} = 1$ , 即GWO算法以1的可能性达到全局收敛;
- 2) 当  $\sum_{t=1}^{\infty} P_t^* < \infty$  时,  $\theta^t$  有可能收敛到  $M$ , 即不能

保证GWO算法达到全局收敛.

**证明** 令  $G_t = \{\theta^t = u, t \geq 1 | \theta^t \notin M\}$ . 不妨假设  $\theta^t = v$  并且  $\{\theta^t = v, t \geq 1\} \cap M \neq \emptyset$ . 此时, 如果令  $\rho = \min\{|F(u) - F(v)|, F(u) \neq F(v)\}$ , 则存在正实数  $r$ , 使得  $F(u) - F(v) \geq r\rho$ .

根据条件期望的性质可以得到

$$\begin{aligned} E(F(\theta^t)) - E(F(\theta^{t+1})) &= \\ E(F(\theta^t)) - E(E(F(\theta^{t+1})|\theta^t)) &= \\ \sum_{u \in \Theta} P(\theta^t = u)(F(\theta^t) - E(F(\theta^{t+1})|\theta^t = u)) &= \\ \sum_{u \in \Theta} P(\theta^t = u) \sum_{v \in \Theta} P(v|u)(F(u) - F(v)). \end{aligned}$$

根据定理1、定理2和命题6中的题设  $P_t^* = \min(P(\theta^{t+1}|\theta^t))$ , 同时结合下面两个不等式:

$$\begin{aligned} \sum_{v \in \Theta} P(v|u) &\geq P(v|u), \\ \sum_{u \in \Theta} P(\theta^t = u) &\geq \sum_{u \cap M = \emptyset} P(\theta^t = u), \end{aligned}$$

可以得到

$$\begin{aligned} E(F(\theta^t)) - E(F(\theta^{t+1})) &\geq \\ \sum_{u \in \Theta} P(\theta^t = u) \sum_{v \in \Theta} P(v|u)r\rho &\geq \\ \sum_{u \in \Theta} P(\theta^t = u)P(v|u)r\rho &\geq \sum_{u \in \Theta} P(\theta^t = u)r\rho P_t^* \geq \\ \sum_{u \cap M = \emptyset} P(\theta^t = u)r\rho P_t^* &\geq r\rho P_t^* P(G_t). \end{aligned}$$

将上式中的  $t$  从1到  $N$  求和, 得到

$$\begin{aligned} E(F(\theta^1)) &\geq E(F(\theta^1)) - E(F(\theta^{N+1})) \geq \\ r\rho \sum_{t=1}^N P(G_t)P_t^*. \end{aligned}$$

当  $N \rightarrow \infty$  时,  $F(\theta^t)$  和  $r\rho$  都是有限的, 可以得到不等式

$$\sum_{t=1}^{\infty} P(G_t)P_t^* < \infty.$$

下面讨论GWO全局收敛的两种情形.

- 1) 当  $\sum_{t=1}^{\infty} P_t^* = \infty$  时,  $P(G_t) \rightarrow 0$ . 将题设中的  $G_t$  代入可得

$$P\{\lim_{t \rightarrow \infty} (\theta^t = u, t \geq 1 \cap M = \emptyset)\} = 0,$$

或者表示为

$$P\{\lim_{t \rightarrow \infty} (\theta^t \in M, t \geq 1)\} = 1.$$

当  $t \rightarrow \infty$  时, 因为  $\{\theta^t, t \geq 1\}$  以1的可能性收敛至群状态随机变量  $\theta^\infty \in \{\theta^t, t \geq 1\}$ , 所以  $\theta^\infty$  的取

值在集合  $M$  中. 根据命题 4, 当  $\sum_{t=1}^{\infty} P_t^* = \infty$  时, 说明 GWO 算法以 1 的可能性达到全局收敛.

2) 当  $\sum_{t=1}^{\infty} P_t^* < \infty$  时,  $P(G_t) \rightarrow 0$ , 或者  $P(G_t) \rightarrow \varepsilon, \varepsilon \in (0, 1)$ .

当  $P(G_t) \rightarrow 0$  时, 与情形 1) 类似, 同理, 可证得 GWO 算法以 1 的可能性达到全局收敛. 当  $P(G_t) \rightarrow \varepsilon$  时,  $\theta^\infty$  的取值可能不在最优群状态集  $M$  中. 虽然当  $t \rightarrow \infty$  时,  $\{\theta^t, t \geq 1\}$  以 1 的可能性收敛至群状态随机变量  $\theta^\infty \in \{\theta^t, t \geq 1\}$ , 但是  $\varepsilon \neq 0$ , 这意味着存在另一个  $\theta^\infty$  可能取值的灰狼群状态集  $\gamma$ , 满足  $\gamma \in \Theta$  且  $\gamma \cap M = \emptyset$ . 由  $\gamma$  可以产生命题 7 (见下面). 根据命题 7, 当群位置状态序列  $\{\theta^t, t \geq 1\}$  属于群状态闭集  $\gamma$  时,  $\theta^\infty$  在  $\gamma$  内取值. 当  $\gamma \cap M = \emptyset$ , 且灰狼群状态不能达到最优位置状态集  $M$  时, 所有灰狼个体都将停滞在  $g^\gamma$  而不会达到全局最优. 在这种情况下, 当  $\sum_{t=1}^{\infty} P_t^* < \infty$  时, 所有灰狼个体将聚集在一个或一些特殊的位置上. GWO 算法将会早熟收敛, 而非全局收敛. □

**命题 7** 若  $M \in \Theta$ , 则至少存在一个群状态闭集  $\gamma$  不属于  $M$  但属于  $\Theta$ . 即  $\exists \gamma \in \Theta$  且  $\gamma \cap M = \emptyset$ .

**证明** 在  $t$  次迭代之后, 如果没有达到全局最优, 并且没有找到一个更好的位置状态, 即  $\lim_{t \rightarrow \infty} X_i^t = g^\gamma$ , 而且  $g^\gamma \neq g^*$ <sup>[22]</sup>, 则灰狼个体的位置状态构造为  $X^\gamma = (g^\gamma, g^\gamma, g^\gamma, g^\gamma) \notin M \subset X$ , 该个体  $i$  的状态为吸收态. 当所有灰狼个体收敛至  $g^\gamma$  时, 灰狼群状态为  $\theta^\gamma = (X^\gamma, X^\gamma, \dots, X^\gamma)$ , 则称群状态  $\theta^\gamma$  为吸收态. 由于  $g^\gamma \neq g^*$ , 显然  $\theta^\gamma \notin M$ . 由此可以构造一个只包括  $\theta^\gamma$  的群状态闭集  $\gamma, \gamma \in \Theta$  且  $\gamma \cap M = \emptyset$ . □

上述理论分析过程证明, GWO 算法满足收敛的基本条件. 因为 GWO 算法每次迭代都保留群体中的最优解, 且存在自适应收敛因子  $a$ , 所以灰狼群位置状态最终不会到达吸收态, 确保 GWO 算法实现全局收敛. GWO 算法经连续无穷次迭代, 灰狼群体序列必定达到最优状态. 再结合式 (12) 可得  $P_t^* > 0$  并且  $\sum_{t=1}^{\infty} P_t^* = \infty$ . 所以 GWO 算法只符合命题 6 中的收敛情况 1), 即  $\sum_{t=1}^{\infty} P_t^* = \infty$ . 说明 GWO 算法以 1 的可能性达到全局收敛.

## 5 GWO 算法的数值实验与分析

本节通过数值实验来验证本文算法的有效性. 本实验的运行环境为: Windows 10 操作系统、Intel(R) Core(TM) i5-10400F CPU @2.90 GHz 处理器、16 G

内存和 GTX1650. 采用 Matlab2018a 运行算法. 由于已有文献对 ABC 算法、CS 算法、PSO 算法、BA 算法<sup>[3,10-12]</sup>进行了收敛性理论分析, 为了更加有效地测试 GWO 算法的收敛性, 将 GWO 算法与上述算法共同进行数值实验. 为使算法能够正常运行, 选取相同的通用参数. 种群大小设置为  $N_\theta = 30$ , 维度  $D = 30$ . 分别对 5 种不同的算法进行参数设置, 如表 1 所示. 可以看出 GWO 算法的控制参数是最少的.

表 1 算法参数设置

算法	参数设置
GWO	$a$ 从 2~0 随迭代次数增加, 线性递减
ABC	蜜源数目为 $N$ , 控制参数 $\text{limit} = 0.6 \times N \times D$
CS	发现概率 $P = 0.25$ , 步长调节量 $\alpha = 0.01$ , 参数 $\beta = 1.5$
PSO	惯性权重 $w = 0.75, c_1 = c_2 = 2$ , 最大速度 $V_{\max} = 0.1(X_{\max} - X_{\min})$
BA	响度 $e = 0.9$ , 脉冲速率 $b = 0.9$ , 频率 $f_{\min} = 0, f_{\max} = 2$

### 5.1 测试函数

本实验选取 10 个测试函数. 其中: Sphere ( $F_1$ )、Schwefel's 2.22 ( $F_2$ )、Schwefel's 1.2 ( $F_3$ ) 和 Schwefel's 2.21 ( $F_4$ ) 为单峰函数, 单峰函数可用于测试算法的寻优精度和寻优速度; Rastrigin ( $F_5$ ) 为多峰函数, Shifted Rastrigin ( $F_6$ ) 为多峰偏移函数, Shifted and Rotated Weierstrass ( $F_7$ ) 和 Shifted and Rotated Schwefel ( $F_8$ ) 为多峰偏移旋转函数, 多峰函数在解空间中存在大量局部极值, 能够很好地考察算法的全局搜索和避免局部最优的能力; Hybrid Function 6 ( $F_9$ ) 和 Composition Function 2 ( $F_{10}$ ) 分别为混合函数与复合函数<sup>[23]</sup>, 混合函数和复合函数用于评价算法跳出局部最优停滞的能力, 以及算法平衡探索与开发的能力.

### 5.2 仿真实验与结果分析

采用表 1 中的 5 种智能算法分别对 10 个测试函数独立运行 30 次, 最大迭代次数分别为 1000 和 2000. 对均值、标准差和每次独立运行的平均耗时 (以 s 为单位) 进行统计. 其中, 均值代表算法的搜索精度, 标准差表示算法的稳定性. 如果运行结果与真实最优值相差小于  $10^{-8}$ , 则记为寻优成功.

表 2 为 10 个测试函数迭代 1000 次并独立运行 30 次得到的结果, 每个测试函数相关性能指标的最佳结果用粗体表示. 由表 2 中的数据可以看出, 在单峰函数  $F_1 \sim F_4$  上, GWO 算法都成功地找到了全局最优值, 说明 GWO 算法具有收敛到全局最优的能力. 虽然 CS 算法在  $F_9$  函数上的收敛精度最高, 但 GWO

表2 函数  $F_1 \sim F_{10}$  运行结果(1000次迭代)

函数	性能指标	GWO	ABC	CS	PSO	BA
$F_1$	mean	<b>4.54e-59</b>	7.39e-01	4.09e-03	5.24e+01	5.45e+04
	std	<b>7.30e-59</b>	1.40e+00	2.92e-03	1.89e+01	1.05e+04
	time/s	0.177	0.823	0.405	0.573	<b>0.134</b>
$F_2$	mean	<b>1.46e-34</b>	3.42e+01	1.16e-01	4.14e-01	1.24e+02
	std	<b>2.61e-34</b>	3.37e+01	4.43e-02	8.36e-02	1.35e+01
	time/s	0.178	0.816	0.412	0.592	<b>0.135</b>
$F_3$	mean	<b>1.77e-14</b>	6.49e+04	2.83e+02	3.81e+01	9.32e+04
	std	<b>6.08e-14</b>	1.07e+04	7.49e+01	1.23e+01	3.54e+04
	time/s	0.392	1.279	0.852	0.802	<b>0.363</b>
$F_4$	mean	<b>7.33e-15</b>	6.01e+01	5.88e+00	1.33e+00	7.58e+01
	std	<b>9.92e-15</b>	5.71e+00	1.95e+00	1.50e-01	6.12e+00
	time/s	0.180	0.807	0.403	0.567	<b>0.134</b>
$F_5$	mean	<b>1.81e-01</b>	2.34e+02	7.81e+01	7.36e+00	2.01e+02
	std	<b>7.09e-01</b>	1.64e+01	1.10e+01	4.06e+00	2.92e+01
	time/s	0.180	0.842	0.450	0.573	<b>0.146</b>
$F_6$	mean	<b>8.95e+02</b>	1.04e+03	9.01e+02	8.56e+03	9.95e+02
	std	2.31e+01	1.94e+01	<b>1.44e+01</b>	1.19e+03	4.57e+01
	time/s	0.228	0.957	0.511	0.211	<b>0.155</b>
$F_7$	mean	<b>6.16e+02</b>	6.41e+02	6.29e+02	6.38e+02	6.41e+02
	std	3.36e+00	<b>8.94e-01</b>	1.55e+00	3.09e+00	2.59e+00
	time/s	<b>1.600</b>	4.086	3.326	1.744	1.770
$F_8$	mean	<b>4.57e+03</b>	9.12e+03	5.16e+03	9.82e+03	5.88e+03
	std	6.99e+02	3.39e+02	<b>1.66e+02</b>	7.80e+02	5.72e+02
	time/s	0.308	1.146	0.699	0.313	<b>0.244</b>
$F_9$	mean	2.64e+03	3.17e+03	<b>2.57e+03</b>	3.49e+07	3.51e+03
	std	1.77e+02	1.54e+02	<b>1.28e+02</b>	1.86e+08	5.32e+02
	time/s	0.288	1.136	0.655	0.285	<b>0.232</b>
$F_{10}$	mean	<b>2.60e+03</b>	2.65e+03	2.63e+03	7.85e+03	2.70e+03
	std	<b>1.29e-02</b>	5.47e+00	3.38e+00	7.07e+02	2.97e+01
	time/s	0.446	1.490	0.969	0.459	<b>0.411</b>

算法在该函数上所得结果排在了第2位,仅次于CS算法. 总体而言,GWO算法的搜索精度高于其他4种算法. 在算法的稳定性方面,除了函数  $F_6$ 、 $F_7$ 、 $F_8$ 、 $F_9$ ,GWO算法在其余函数上的稳定性优于其他算法. 由式(7)可以看出,GWO算法在进行位置更新的过程中,充分利用了前3头灰狼的位置信息,这种位置更新机制是GWO算法稳定性较好的原因之一. 从每次运行的平均时间可以看出:在函数  $F_7$  上,GWO算法的耗时最短;而在其他函数上,BA算法耗时最短. 但是,BA算法在搜索精度和稳定性方面的性能明显较差. 整体上,ABC算法的耗时最长,而ABC算法在函数  $F_7$  上的稳定性是最好的. 总体而言,GWO算法的收敛性能和稳定性优于其他4种算法,而且GWO算

法的运行耗时仅次于BA算法,CS算法的收敛性能处于中等,ABC、PSO和BA算法的收敛性能相对较弱.

在表2中,函数  $F_5 \sim F_{10}$  经过1000次迭代,暂时没有达到全局最优. 由于函数  $F_5 \sim F_{10}$  均是较为复杂的测试函数,在对应的解空间中存在大量局部极值,很难找到全局最优解. 针对函数  $F_5 \sim F_{10}$ ,在其他实验环境及条件不变的情况下,最大迭代次数增加至2000得到表3中的结果. 由表3中数据可以看出,函数  $F_5 \sim F_{10}$  在2000次迭代后,各个算法所得结果均优于表2,更加接近全局最优. 尤其在函数  $F_5$  上,GWO算法收敛到了全局最优. 同时,表3的数据显示,GWO算法在收敛精度方面的性能是较好的. 说明即使对于一些复杂函数,随着迭代次数的增

加, GWO 算法将会逐渐向全局最优靠近, 而且具有较好的收敛特性. 在函数  $F_7$  上, GWO 算法的耗时最短; 而在其余函数上, BA 算法的耗时最短.

综合表 2 和表 3 可以看出, GWO 算法的收敛性能优于其余 4 种算法, 表现出了良好的跳出局部最优的能力和较强的稳定性.

表 3 函数  $F_5 \sim F_{10}$  运行结果(2 000 次迭代)

函数	性能指标	GWO	ABC	CS	PSO	BA
$F_5$	mean	<b>0.00e+00</b>	2.29e+02	5.73e+01	4.76e+00	2.00e+02
	std	<b>0.00e+00</b>	1.04e+01	9.64e+00	3.14e+00	3.48e+01
	time/s	0.352	1.680	0.881	1.140	<b>0.286</b>
$F_6$	mean	8.90e+02	1.03e+03	<b>8.70e+02</b>	8.52e+03	9.81e+02
	std	2.19e+01	1.28e+01	<b>1.20e+01</b>	1.21e+03	2.80e+01
	time/s	0.458	1.883	1.031	0.433	<b>0.320</b>
$F_7$	mean	<b>6.15e+02</b>	6.40e+02	6.27e+02	6.37e+02	6.40e+02
	std	2.72e+00	<b>7.70e-01</b>	1.25e+00	2.87e+00	2.09e+00
	time/s	<b>3.171</b>	8.396	6.640	3.460	3.689
$F_8$	mean	<b>4.17e+03</b>	9.03e+03	4.95e+03	9.66e+03	5.47e+03
	std	8.98e+02	3.53e+02	<b>2.46e+02</b>	7.56e+02	6.81e+02
	time/s	0.610	2.231	1.375	0.613	<b>0.483</b>
$F_9$	mean	2.57e+03	3.13e+03	<b>2.55e+03</b>	3.48e+07	3.48e+03
	std	1.59e+02	1.05e+02	<b>9.67e+01</b>	1.73e+08	3.37e+02
	time/s	0.579	2.301	1.313	0.566	<b>0.463</b>
$F_{10}$	mean	<b>2.60e+03</b>	2.64e+03	2.63e+03	7.83e+03	2.68e+03
	std	<b>3.74e-03</b>	5.47e+00	1.50e+00	4.94e+02	2.51e+01
	time/s	0.889	2.961	1.909	0.908	<b>0.816</b>

## 6 结 论

本文在分析 GWO 算法原理的基础上, 运用鞅理论研究了 GWO 算法的收敛性. 将灰狼群进化序列转化为上鞅, 利用上鞅收敛定理证明了 GWO 的收敛性. 经推导证明, GWO 算法满足收敛条件  $P_t^* > 0$ ,  $\sum_{t=1}^{\infty} P_t^* = \infty$ , 证明了 GWO 算法以 1 的可能性达到全局收敛. 证明中所使用的鞅方法, 为 GWO 算法的收敛性分析提供了新的数学分析工具, 进而为 GWO 算法的发展和应用奠定了良好的理论基础. 通过进行相关数值实验, 验证了本文理论分析的正确性. 实验结果表明, GWO 算法相比于其他算法, 其搜索精度较高、稳定性较好、时间复杂较低. 关于 GWO 算法收敛性分析的其他相关方面, 如算法的运行时间<sup>[24]</sup>等, 在未来还需要进行更多的研究.

### 参考文献(References)

[1] Mirjalili S, Mirjalili S M, Lewis A. Grey wolf optimizer[J]. Advances in Engineering Software, 2014, 69: 46-61.  
 [2] Karaboga D, Gorkemli B, Ozturk C, et al. A comprehensive survey: Artificial bee colony (ABC)

algorithm and applications[J]. Artificial Intelligence Review, 2014, 42(1): 21-57.  
 [3] 任子晖, 王坚, 高岳林. 马尔科夫链的粒子群优化算法全局收敛性分析[J]. 控制理论与应用, 2011, 28(4): 462-466.  
 (Ren Z H, Wang J, Gao Y L. The global convergence analysis of particle swarm optimization algorithm based on Markov chain[J]. Control Theory & Applications, 2011, 28(4): 462-466.)  
 [4] Mirjalili S, Mirjalili S M, Yang X S. Binary bat algorithm[J]. Neural Computing and Applications, 2014, 25(3/4): 663-681.  
 [5] 张晓凤, 王秀英. 灰狼优化算法研究综述[J]. 计算机科学, 2019, 46(3): 30-38.  
 (Zhang X F, Wang X Y. Comprehensive review of grey wolf optimization algorithm[J]. Computer Science, 2019, 46(3): 30-38.)  
 [6] Faris H, Aljarah I, Al-Betar M A, et al. Grey wolf optimizer: A review of recent variants and applications[J]. Neural Computing and Applications, 2018, 30(2): 413-435.  
 [7] Heidari A A, Pahlavani P. An efficient modified grey wolf optimizer with Lévy flight for optimization tasks[J].

- Applied Soft Computing, 2017, 60: 115-134.
- [8] Gupta S, Deep K. A novel random walk grey wolf optimizer[J]. Swarm and Evolutionary Computation, 2019, 44: 101-112.
- [9] Bansal J C, Singh S. A better exploration strategy in grey wolf optimizer[J]. Journal of Ambient Intelligence and Humanized Computing, 2021, 12(1): 1099-1118.
- [10] 宁爱平, 张雪英. 人工蜂群算法的收敛性分析[J]. 控制与决策, 2013, 28(10): 1554-1558.  
(Ning A P, Zhang X Y. Convergence analysis of artificial bee colony algorithm[J]. Control and Decision, 2013, 28(10): 1554-1558.)
- [11] 刘晓东, 孙丽君, 陈天飞. 布谷鸟算法的收敛性分析及性能比较[J]. 计算机科学与探索, 2020, 14(10): 1644-1655.  
(Liu X D, Sun L J, Chen T F. Convergence analysis and performance comparison of cuckoo search algorithm[J]. Journal of Frontiers of Computer Science and Technology, 2020, 14(10): 1644-1655.)
- [12] 尚俊娜, 程涛, 岳克强, 等. 蝙蝠算法的Markov链模型分析[J]. 计算机工程, 2017, 43(7): 198-202.  
(Shang J N, Cheng T, Yue K Q, et al. Markov chain model analysis of bat algorithm[J]. Computer Engineering, 2017, 43(7): 198-202.)
- [13] 张孟健, 龙道银, 王霄, 等. 基于马尔科夫链的灰狼优化算法收敛性研究[J]. 电子学报, 2020, 48(8): 1587-1595.  
(Zhang M J, Long D Y, Wang X, et al. Research on convergence of grey wolf optimization algorithm based on Markov chain[J]. Acta Electronica Sinica, 2020, 48(8): 1587-1595.)
- [14] 郭红戈. 思维进化算法的转移概率分析及几乎处处收敛性证明[J]. 控制与决策, 2014, 29(12): 2201-2206.  
(Guo H G. Migration probabilities analysis and almost sure convergence proof of mind evolutionary algorithm[J]. Control and Decision, 2014, 29(12): 2201-2206.)
- [15] 苏兆品, 蒋建国, 梁昌勇, 等. 蚁群算法的几乎处处强收敛性分析[J]. 电子学报, 2009, 37(8): 1646-1650.  
(Su Z P, Jiang J G, Liang C Y, et al. An almost everywhere strong convergence proof for a class of ant colony algorithms[J]. Acta Electronica Sinica, 2009, 37(8): 1646-1650.)
- [16] 喻寿益, 邝溯琼. 保留精英遗传算法收敛性和收敛速度的鞅方法分析[J]. 控制理论与应用, 2010, 27(7): 843-848.  
(Yu S Y, Kuang S Q. Convergence and convergence rate analysis of elitist genetic algorithm based on martingale approach[J]. Control Theory & Applications, 2010, 27(7): 843-848.)
- [17] 洪露, 龚成龙, 王经卓, 等. 噪声环境下精英克隆选择算法的收敛性分析[J]. 控制理论与应用, 2013, 30(11): 1457-1461.  
(Hong L, Gong C L, Wang J Z, et al. Convergence analysis of elitist clonal selection algorithm in noisy environment[J]. Control Theory & Applications, 2013, 30(11): 1457-1461.)
- [18] Xiao X Y, Yin H W. Moment convergence rates in the law of logarithm for moving average process under dependence[J]. Stochastics, 2014, 86(1): 1-15.
- [19] 万成高. 鞅的极限理论[M]. 北京: 科学出版社, 2002: 25-70.  
(Wan C G. Limit theory of martingales[M]. Beijing: Science Press, 2002: 25-70.)
- [20] Lawler G F. Introduction to stochastic processes[M]. The 2nd edition. Chapman and Hall: CRC Press, 2006: 163-341.
- [21] Leake C, Kao E P C. An introduction to stochastic processes[J]. The Journal of the Operational Research Society, 1999, 50(3): 284.
- [22] Van Den Bergh F, Engelbrecht A P. A study of particle swarm optimization particle trajectories[J]. Information Sciences, 2006, 176(8): 937-971.
- [23] Liang J J, Qu B Y, Suganthan P N. Problem definitions and evaluation criteria for the CEC 2014 special session and competition on single objective real-parameter numerical optimization[R]. Zhengzhou: Computational Intelligence Laboratory, Zhengzhou University, 2014.
- [24] Yu Y, Zhou Z H. A new approach to estimating the expected first hitting time of evolutionary algorithms[J]. Artificial Intelligence, 2008, 172(15): 1809-1832.

### 作者简介

孙丽君(1968—), 女, 教授, 博士生导师, 从事智能信息处理、通信信号处理等研究, E-mail: ljsun@haut.edu.cn;

冯斌斌(1994—), 男, 硕士生, 从事群体智能算法、无线传感网络的研究, E-mail: 1563095871@qq.com;

陈天飞(1986—), 男, 副教授, 博士, 从事无线传感网络、计算机视觉等研究, E-mail: chen\_tianfei@163.com.

(责任编辑: 李君玲)