第 52 卷 第 6 期 河南师范大学学报(自然科学版) Vol.52 No.6
2024 年 11 月 Journal of Henan Normal University(Natural Science Edition) Nov. 2024

基于鞅方法的鸡群优化算法收敛性分析

周婷婷,戴家佳

(贵州大学 数学与统计学院,贵阳 550025)

摘 要:针对鸡群优化(chicken swarm optimization,CSO)算法已有的收敛性分析结果属于弱收敛,不能保证算法能在有限步内收敛到问题的全局最优这一不足,提出了运用鞅方法来研究 CSO 算法的全局收敛性.首先,基于 CSO 算法的相关定义,建立 CSO 算法的马尔可夫(Markov)链模型,分析其 Markov 性质;其次,将具有最小适应度值的鸡群状态序列转化成上鞅,利用上鞅收敛定理和 Egoroff 定理证明了 CSO 算法的几乎处处强收敛性和一致收敛性,进而得出了当鸡群状态空间有限时,CSO 算法能确保在有限步内收敛到问题的全局最优这一结论;最后,在仿真实验中成功验证了理论证明的正确性,并发现 CSO 算法比其他算法具有更强的寻优能力和更高的收敛精度.

关键词:CSO 算法;Markov 链;上鞅收敛定理;Egoroff 定理;几乎处处强收敛;一致收敛

中图分类号:TP301.6;TP18 **文献标志码**:A **文章编号**:1000-2367(2024)06-0080-08

鸡群优化算法[1](简称鸡群算法)属于群智能优化算法中的一种,主要通过模拟鸡群中的等级秩序和觅食行为,利用鸡的群体智能来解决优化问题.相较于粒子群优化[2](particle swarm optimization,PSO)算法、差分进化[3](differential evolution,DE)算法,其具有参数少、收敛速度快、简单易操作等特点[4].目前,CSO 算法已被成功地应用到了车间调度、路径规划、图像分类、资源分配等领域[5]中,并且根据实际问题的特殊性,出现了一大批 CSO 改进算法[6-8].不难发现,由于发展时间较短,CSO 算法改进与应用虽然已有了不少研究成果,但与其他算法相比,它在数学理论方面的研究还比较少,特别是关于算法的收敛性研究仍不多见,因此在一定程度上也限制了 CSO 算法的改进和应用.

作为一种常用来分析群智能算法收敛性的数学工具,随机搜索算法的收敛准则已成功运用到了蝙蝠算法(bat algorithm,BA)[9]、灰狼优化(grey wolf optimization,GWO)算法[10]和萤火虫算法(firefly algorithm,FA)[11]等算法的收敛性证明中.同样地,文献[12]通过建立 CSO 算法的 Markov 链模型,并结合该收敛准则,证明了 CSO 算法能收敛到问题的全局最优.然而,上述证明方法存在如下不足:一是通过该方法进行收敛性理论分析时过程较为复杂,且得到的收敛结果属于依概率收敛,在弱大数定律范畴中;二是无法确保 CSO 算法能在有限步内收敛到问题的全局最优.针对这两点,本文基于 CSO 算法的 Markov 链模型,提出了用鞅方法代替 Markov 链的遍历性分析来研究 CSO 算法的全局收敛性.不仅有效简化了 CSO 算法全局收敛性理论分析的过程,而且还成功证明了当鸡群状态空间有限时,CSO 算法能确保在有限步内以概率 1 收敛到问题的全局最优.

收稿日期:2023-05-04;**修回日期**:2023-07-04.

基金项目:国家自然科学基金(12361057);贵州省数据驱动建模学习与优化创新团队项目(黔科合平台人才[2020]5016).

作者简介:周婷婷(1998—),女,贵州遵义人,贵州大学硕士研究生,研究方向为智能优化算法、生物与医学统计,E-mail:3075055469@qq.com.

通信作者:戴家佳(1976—),女,贵州贵阳人,贵州大学教授,博士,研究方向为生物与医学统计、生存分析,E-mail:jjdai@gzu.edu.cn.

引用本文:周婷婷,戴家佳.基于鞅方法的鸡群优化算法收敛性分析[J].河南师范大学学报(自然科学版),2024,52(6):80-87.(Zhou Tingting,Dai Jiajia.Convergence analysis of chicken swarm optimization algorithm based on martingale method[J].Journal of Henan Normal University(Natural Science Edition),2024,52(6):80-87.DOI:10.16366/j.cnki.1000-2367.2023.05.04.0002.)

1 CSO 算法

CSO 算法综合了 PSO 算法、DE 算法的优点,是通过 4 条规则将鸡群中的等级秩序和觅食行为理想化,从而建立数学模型来求解优化问题.本文考虑最小值优化问题,即:

$$\min F(\boldsymbol{x}), F(\boldsymbol{x}) \geq 0. \tag{1}$$

CSO 算法在优化问题的求解空间中随机生成 N 只鸡形成鸡群 $\boldsymbol{x}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_N)$.鸡群中第 i 只鸡在 D 维空间中所处的位置表示为:$\boldsymbol{x}_i=(\boldsymbol{x}_{i1},\boldsymbol{x}_{i2},\cdots,\boldsymbol{x}_{iD})^{\mathrm{T}}$,$i=1,2,\cdots,N$.$\boldsymbol{x}_i$ 是优化问题在解空间中的一个潜在解,每只鸡通过适应度函数 $f(\boldsymbol{x}_i)$ 求出适应度值.由于适应度函数是定义在解空间上的任意非负实值函数,基于式(1),在解空间上定义的适应度函数就是目标函数.

鸡群中共有 N 只鸡,其中公鸡有 N_r 只、母鸡有 N_h 只、小鸡有 N_c 只、母鸡妈妈有 N_m 只,且 $N_m \subset N_h$,$N_r+N_h+N_c=N$.第 i 只鸡在第 j 维空间中第 t 次迭代的位置为 $\boldsymbol{x}_i^j(t)$.因为不同种类的鸡有不同的运动规律,故 3 种鸡的位置更新策略也各不相同.

公鸡的位置更新策略为:

$$\boldsymbol{x}_i^j(t+1)=\boldsymbol{x}_i^j(t)\times(1+R), \tag{2}$$

$$\sigma^2=\begin{cases}1, f_i \leq f_k, \\ \exp\left[\dfrac{(f_k-f_i)}{|f_i|+\varepsilon}\right], \text{其他}, \end{cases} \quad k=1,2,\cdots,N \text{ 且 } k \neq i. \tag{3}$$

式中:R 是服从 $N(0,\sigma^2)$ 的随机数;f 为对应 $\boldsymbol{x}$ 的适应度值;ε 是一个充分小的正数,用来避免式(3)中分母为 0;k 是所有公鸡中除了 i 后的任一只公鸡.

母鸡的位置更新策略为:

$$\boldsymbol{x}_i^j(t+1)=\boldsymbol{x}_i^j(t)+B_1\times V_1\times[\boldsymbol{x}_{r_1}^j(t)-\boldsymbol{x}_i^j(t)]+B_2\times V_2\times[\boldsymbol{x}_{r_2}^j(t)-\boldsymbol{x}_i^j(t)], \tag{4}$$

$$B_1=\exp\left[\frac{(f_i-f_{r1})}{|f_i|+\varepsilon}\right], \tag{5}$$

$$B_2=\exp(f_{r2}-f_i). \tag{6}$$

式中:r_1 是与第 i 只母鸡同在一个子群里的公鸡;r_2 是整个鸡群里的任意一只公鸡或母鸡,且 $r_1 \neq r_2$;B_1、B_2 为常数,取值范围为(0,1);V_1、V_2 是服从[0,1]之间均匀分布的随机数.

小鸡的位置更新策略为:

$$\boldsymbol{x}_i^j(t+1)=\boldsymbol{x}_i^j(t)+F_L\times[\boldsymbol{x}_m^j(t)-\boldsymbol{x}_i^j(t)]. \tag{7}$$

式中:m 是第 i 只小鸡的母鸡妈妈;F_L 是跟随系数,为小鸡跟着母鸡妈妈觅食时母鸡妈妈的位置对小鸡位置的影响因子,取值范围为(0,2).

2 CSO 算法的 Markov 链模型

Markov 链是对随机过程进行分析的重要手段,指的是系统在任一时刻所处状态组成的 Markov 随机序列.下面先给出一些 CSO 算法的相关定义,以便建立其 Markov 链模型.

定义 1 在 CSO 算法中,鸡群中每只鸡觅食时所处的位置构成了鸡个体状态,记为 $\boldsymbol{x}$,$\boldsymbol{x}\in\boldsymbol{C}$,$\boldsymbol{C}$ 是可行解空间,一只鸡的全部可能状态组成了鸡个体状态空间,记为:$\boldsymbol{x}=\{\boldsymbol{x} \mid \boldsymbol{x}\in\boldsymbol{C}\}$.所有鸡个体状态就组成了鸡群状态,记为:

$$\theta=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_i), 1 \leq i \leq N. \tag{8}$$

故鸡群状态空间记为:

$$\boldsymbol{\Theta}=\{\theta=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_N) \mid \boldsymbol{x}_i \in \boldsymbol{C}, 1 \leq i \leq N\}. \tag{9}$$

式中:$\boldsymbol{x}_i$ 为第 i 只鸡的状态,N 为种群大小.

定义 2 对于任意两状态 $\boldsymbol{x}_i\in\theta$,$\boldsymbol{x}_j\in\theta$,CSO 算法迭代过程中鸡个体从状态 $\boldsymbol{x}_i$ 一步转移到状态 $\boldsymbol{x}_j$,记为 $T_\theta(\boldsymbol{x}_i)=\boldsymbol{x}_j$.

定理1 CSO算法中,鸡个体从状态 $\boldsymbol{x}_i$ 转移到 $\boldsymbol{x}_j$ 的一步转移概率为:

$$P\{T_\theta(\boldsymbol{x}_i)=\boldsymbol{x}_j\}=\begin{cases}P_r\{T_\theta(\boldsymbol{x}_i)=\boldsymbol{x}_j\},\text{公鸡},\\P_h\{T_\theta(\boldsymbol{x}_i)=\boldsymbol{x}_j\},\text{母鸡},\\P_c\{T_\theta(\boldsymbol{x}_i)=\boldsymbol{x}_j\},\text{小鸡}.\end{cases}\tag{10}$$

证明 过程详见参考文献[12]中的定理1.

由于CSO算法中公鸡、母鸡和小鸡有不同的位置更新策略,故它们从状态 $\boldsymbol{x}_i$ 转移到 $\boldsymbol{x}_j$ 也有着不同的一步转移概率.公鸡的一步转移概率为:

$$P_r\{T_\theta(\boldsymbol{x}_i)=\boldsymbol{x}_j\}=\begin{cases}\dfrac{1}{|\boldsymbol{x}_i\times R|},\boldsymbol{x}_j\in[\boldsymbol{x}_i,\boldsymbol{x}_i+\boldsymbol{x}_i\times R],\\0,\text{其他}.\end{cases}\tag{11}$$

母鸡的一步转移概率为:

$$P_h\{T_\theta(\boldsymbol{x}_i)=\boldsymbol{x}_j\}=\begin{cases}\dfrac{1}{|S_1\times V_1\times(\boldsymbol{x}_{r_1}-\boldsymbol{x}_i)+S_2\times V_2\times(\boldsymbol{x}_{r_2}-\boldsymbol{x}_i)|},\\\boldsymbol{x}_j\in[\boldsymbol{x}_i,\boldsymbol{x}_i+S_1\times V_1\times(\boldsymbol{x}_{r_1}-\boldsymbol{x}_i)+S_2\times V_2\times(\boldsymbol{x}_{r_2}-\boldsymbol{x}_i)],\\0,\text{其他}.\end{cases}\tag{12}$$

小鸡的一步转移概率为:

$$P_c\{T_\theta(\boldsymbol{x}_i)=\boldsymbol{x}_j\}=\begin{cases}\dfrac{1}{|FL\times(\boldsymbol{x}_m-\boldsymbol{x}_i)|},\boldsymbol{x}_j\in[\boldsymbol{x}_i,\boldsymbol{x}_i+FL\times(\boldsymbol{x}_m-\boldsymbol{x}_i)],\\0,\text{其他}.\end{cases}\tag{13}$$

定义3 CSO算法中对于 $\forall\theta_i\in\boldsymbol{\Theta},\forall\theta_j\in\boldsymbol{\Theta}$,鸡群状态 θ_i 一步转移到 θ_j 记为 $T_{\boldsymbol{\Theta}}(\theta_i)=\theta_j$,则其一步转移概率为:

$$P\{T_{\boldsymbol{\Theta}}(\theta_i)=\theta_j\}=\prod_{k=1}^{N}P\{T_\theta(\boldsymbol{x}_{ik})=\boldsymbol{x}_{jk}\}.\tag{14}$$

定义4(Markov链)[13] 设 $\{X_n,n\in T\}$ 为一列取值离散的随机变量,参数集 T 为离散的时间序列,即 $T=\{0,1,2,\cdots\}$,全体随机变量 X_n 所对应取值组成的离散状态空间为 $I=\{i_0,i_1,i_2,\cdots\}$.若对任意的状态 $i_0,i_1,i_2,\cdots,i_{n+1}\in I$ 和整数 $n\in T$ 满足如下条件概率:

$$P\{X_{n+1}=i_{n+1}\mid X_0=i_0X_1=i_1,\cdots,X_n=i_n\}=P\{X_{n+1}=i_{n+1}\mid X_n=i_n\},\tag{15}$$

则称 $\{X_n,n\in T\}$ 为Markov链.

定理2 CSO算法中的鸡群状态序列 $\{\theta(t),t\geq0\}$ 是有限齐次Markov链.

证明 对于 $\forall\theta(t)\in\boldsymbol{\Theta},\forall\theta(t+1)\in\boldsymbol{\Theta}$,根据定义3,其一步转移概率 $P\{T_\theta(\theta(t))=\theta(t+1)\}$ 由鸡群内所有鸡的转移概率 $P\{T_\theta(\boldsymbol{x}(t))=\boldsymbol{x}(t+1)\}$ 决定.根据定理1,可知鸡群内任意鸡的状态转移概率 $P\{T_\theta(\boldsymbol{x}(t))=\boldsymbol{x}(t+1)\}$ 仅与 t 时刻的状态 $\boldsymbol{x}(t)$ 以及式(2)~(7)中随机数 R、V_1、V_2,常数 B_1、B_2,第 i 只母鸡有关的公鸡 $\boldsymbol{x}_{r_1}$,鸡群中随机选择的公鸡或母鸡个体 $\boldsymbol{x}_{r_2}$,跟随系数 F_L 和与小鸡的母鸡妈妈 $\boldsymbol{x}_m$ 有关.所以一步转移概率 $P\{T_{\boldsymbol{\Theta}}(\theta(t))=\theta(t+1)\}$ 也仅与 t 时刻的状态式(2)~(7)中的相关参数有关,而与 t 时刻无关,即:鸡群状态序列 $\{\theta(t),t\geq0\}$ 具有Markov性;对于任何优化算法来说,其搜索空间都是有限的,因此CSO算法中每只鸡的状态 $\boldsymbol{x}_i$ 也都是有限的.而鸡群状态空间 $\boldsymbol{\Theta}$ 又由 N 只鸡的状态共同组成,故 $\boldsymbol{\Theta}$ 也是有限的;再根据定理1可知,状态转移概率 $P\{T_\theta(\boldsymbol{x}(t))=\boldsymbol{x}(t+1)\}$ 仅与时刻的状态 $\boldsymbol{x}(t)$ 有关,而与 t 时刻无关.故鸡群状态序列 $\{\theta(t),t\geq0\}$ 是有限齐次Markov链.证毕.

3 CSO算法的全局收敛性分析

3.1 基础理论

定义5[14] 若 $\boldsymbol{G}=\{\boldsymbol{x}^*\mid\forall\boldsymbol{x}\in\boldsymbol{\Theta},f(\boldsymbol{x}^*)\leq f(\boldsymbol{x})\}$ 为鸡群状态序列 $\{\theta(t),t\geq0\}$ 的全局最优状态集,如果 $\lim\limits_{t\to\infty}P\{[\theta(t)\cap\boldsymbol{G}\neq\varnothing]\}=1$($\lim\limits_{t\to\infty}P\{[\theta(t)\subset\boldsymbol{G}]\}=1$),则称鸡群状态序列 $\{\theta(t),t\geq0\}$ 依概率弱(强)

收敛到全局最优状态集 $\boldsymbol{G}$；如果 $P\{\lim\limits_{t\to\infty}[\theta(t)\cap \boldsymbol{G}\neq\varnothing]\}=1$（$P\{\lim\limits_{t\to\infty}[\theta(t)\subset \boldsymbol{G}]\}=1$），则称序列$\{\theta(t),t\geq 0\}$ 几乎处处弱（强）收敛到全局最优状态集 $\boldsymbol{G}$.

定理 3　鸡群状态序列$\{\theta(t),t\geq 0\}$的 4 种收敛性有如下关系：(1)几乎处处强收敛⇒几乎处处弱收敛⇒依概率弱收敛；(2)几乎处处强收敛⇒依概率强收敛⇒依概率弱收敛.

证明　由 $[\theta(t)\subset \boldsymbol{G}]\subset[\theta(t)\cap \boldsymbol{G}\neq\varnothing]$ 可知，$P\{[\theta(t)\subset \boldsymbol{G}]\}\leq P\{[\theta(t)\cap \boldsymbol{G}\neq\varnothing]\}$，所以依概率强收敛可推出依概率弱收敛，几乎处处强收敛可推出几乎处处弱收敛. 又因 $\lim\limits_{t\to\infty}[\theta(t)\cap \boldsymbol{G}\neq\varnothing]=\bigcup\limits_{t=1}^{\infty}\bigcap\limits_{k=t}^{\infty}[\theta(t)\cap \boldsymbol{G}\neq\varnothing]$，由概率的单调性可得：$P\{\lim\limits_{t\to\infty}[\theta(t)\cap \boldsymbol{G}\neq\varnothing]\}=\lim\limits_{t\to\infty}P\{[\bigcap\limits_{k=t}^{\infty}\theta(t)\cap \boldsymbol{G}\neq\varnothing]\}\leq\lim\limits_{t\to\infty}P\{[\theta(t)\cap \boldsymbol{G}\neq\varnothing]\}$. 所以几乎处处弱收敛可推出依概率弱收敛.同理可得，几乎处处强收敛可推出依概率强收敛.证毕.

定义 6[15]　如果对 $\forall i\in \boldsymbol{E}$，有 $\sum\limits_{j\in \boldsymbol{E}}p_{ij}=1$，称状态空间 $\boldsymbol{E}$ 为闭集.

定义 7[15]　如果闭集 $\boldsymbol{E}$ 有真闭子集，则称 $\boldsymbol{E}$ 为可约的，反之称 $\boldsymbol{E}$ 不可约.

命题 1　鸡群状态空间 $\boldsymbol{\Theta}$ 是一个闭集.

证明　根据定理 2，鸡群状态序列$\{\theta(t),t\geq 0\}$是有限齐次 Markov 链，故有 $\forall i\in\boldsymbol{\Theta}$，$\sum\limits_{j\in\boldsymbol{\Theta}}P_{ij}=1$.根据定义 6，可知 $\boldsymbol{\Theta}$ 是一个闭集. 证毕.

命题 2　在 CSO 算法中，全局最优状态集 $\boldsymbol{G}$ 是状态空间 $\boldsymbol{\Theta}$ 上的一个真闭子集.

证明　首先，当 $\boldsymbol{G}=\boldsymbol{\Theta}$ 时，$\boldsymbol{\Theta}$ 中的每一个解都是最优解，优化问题没有讨论意义，故本文是基于 $\boldsymbol{G}\subset\boldsymbol{\Theta}$ 来讨论优化问题.其次，对于 $\forall\theta_i\in \boldsymbol{G}$，$\forall\theta_j\neq \boldsymbol{G}$，$\theta_j\in\boldsymbol{\Theta}$.由定义 3 可知，$T_{\boldsymbol{\Theta}}(\theta_i)=\theta_j$ 的转移概率为：$P\{T_{\boldsymbol{\Theta}}(\theta_i)=\theta_j\}=\prod\limits_{n=1}^{N}P\{T_\theta(\boldsymbol{x}_{in})=\boldsymbol{x}_{jn}\}$.若在 $\boldsymbol{G}$ 中至少有 1 只鸡的状态达到了最优，令 $\boldsymbol{x}^*\sim\boldsymbol{x}_{i_0 n}$ 为最优状态，那么有 $\boldsymbol{x}_{i_0 n}\in \boldsymbol{G}$，使得 $P\{T_\theta(\boldsymbol{x}_{i_0 n})=\boldsymbol{x}_{j_0 n}\}=0$，此时 $P\{T_{\boldsymbol{\Theta}}(\theta_i)=\theta_j\}=0$.故 $\boldsymbol{G}$ 是 $\boldsymbol{\Theta}$ 上的一个真闭子集. 证毕.

命题 3　鸡群状态空间 $\boldsymbol{\Theta}$ 是可约的.

证明　首先从命题 1 可以知道整个鸡群状态空间 $\boldsymbol{\Theta}$ 是一个最大的闭集.其次由命题 2 可知，鸡群的全局最优状态集 $\boldsymbol{G}$ 是状态空间 $\boldsymbol{\Theta}$ 上的一个真闭子集.最后由定义 7 可知鸡群状态空间 $\boldsymbol{\Theta}$ 是可约的. 证毕.

3.2　CSO 算法的几乎处处强收敛性分析

由于利用群智能算法求解优化问题的过程是一个迭代寻优的过程，所以可把它转换成一类收敛问题，利用鞅的收敛定理对其过程进行理论分析.本文考虑的是目标函数为最小值的情况，即所求鸡群的适应度值越小越好，故只有找到全局最小适应度值时，才能得到全局最优状态集（全局最优解）.假若鸡群在第 t 次迭代时得到的最佳个体适应度值达到全局最优状态集所对应的最小适应度值 $f(\theta^*)$，即：$f(\theta_t)=f(\theta^*)$，那么在第 t 次迭代之后的任何一次迭代中鸡群的最佳个体的适应度值都不会大于全局最优解所对应的最小适应度值 $f(\theta^*)$，即下一次迭代时鸡群的最小适应度值关于当前迭代鸡群的条件期望不大于当前迭代鸡群的最小适应度值.因此可把鸡群的最小适应度函数 $f(\theta_t)$ 转变为一个上鞅，将研究鸡群状态序列$\{\theta(t),t\geq 0\}$ 的收敛性转化为研究 $f(\theta_t)$ 的收敛性. 根据定理 3 可以知道，几乎处处强收敛是强于依概率收敛的，故接下来本文将通过利用上鞅的性质和上鞅收敛定理来证明 CSO 算法的几乎处处强收敛性.

定义 8[15]　如果随机过程 $\{Y_k,k\geq 0\}$ 与$\{X_k,k\geq 0\}$ 满足以下条件：(1)$E(|Y_k|)<\infty$；(2)$E(Y_{k+1}\mid X_0,X_1,\cdots,X_k)\leq Y_k$（或 $E(Y_{k+1}\mid X_0,X_1,\cdots,X_k)\geq Y_k$）；(3)$Y_k$ 是 $X_0,X_1,\cdots,X_k$ 的函数；则称随机过程$\{Y_k,k\geq 0\}$ 是关于$\{X_k,k\geq 0\}$ 的上鞅（或下鞅）.

引理 1（下鞅收敛定理）[15]　若随机过程 $\{Y_k,k\geq 0\}$ 是一个下鞅，且满足 $\sup E\mid Y_k\mid<\infty$，则一定存在一个随机变量 $Y^*\in\{Y_k,k\geq 0\}$ 使得 $E\mid Y^*\mid<\infty$，且$\{Y_k,k\geq 0\}$ 以概率 1 收敛至 Y^*，即：$P\{\lim\limits_{t\to\infty}Y_k=Y^*\}=1$.

引理 2[15]　引理 1 的结论对上鞅也成立.

定理 4　随机过程 $\{f(\theta_t),t\geq 0\}$ 关于$\{\theta_t,t\geq 0\}$ 是一个非负有界上鞅函数列，即对任意 $t\geq 0$，则有 $E[f(\theta_{t+1})\mid\theta_0,\theta_1,\cdots,\theta_t]\leq f(\theta_t)$.

证明　由于适应度函数是定义在解空间上的任意非负实值函数，故 $f(\theta_t)\geq 0$.又因 CSO 算法保证了下

一次迭代时鸡群的最小适应度值不会大于当前迭代鸡群的最小适应度值,所以可以得到:$f(\theta_{t+1}) \leq f(\theta_t) \leq f(\theta_{t-1})$,所以$\{f(\theta_t), t \geq 0\}$是非负有界的. 再根据定理 2 的马尔可夫性可得:

$$E[f(\theta_{t+1}) \mid \theta_0, \theta_1, \cdots, \theta_t] = E[f(\theta_{t+1}) \mid \theta_t]. \tag{16}$$

因此只需证明对任意 $\boldsymbol{x} \in \boldsymbol{\Theta}$ 有:

$$E[f(\theta_{t+1}) \mid \theta_t = \boldsymbol{x}] \leq f(\boldsymbol{x}), \tag{17}$$

根据式(17)和随机过程的相关理论知识,式(17)不等号左边可以转化为:

$$E[f(\theta_{t+1}) \mid \theta_t = \boldsymbol{x}] = \sum_{y \in \boldsymbol{\Theta}} f(y) P(\theta_{t+1} = y \mid \theta_t = \boldsymbol{x}), \tag{18}$$

因此需要证明的式(17)就转化成:

$$\sum_{y \in \boldsymbol{\Theta}} f(y) P(\theta_{t+1} = y \mid \theta_t = \boldsymbol{x}) \leq f(\boldsymbol{x}). \tag{19}$$

由于任意群状态 $\boldsymbol{x}$ 之后的群状态都存在:$f(\theta_{t+1} = y) \leq f(\theta_t = \boldsymbol{x})$, $\forall y \in \boldsymbol{\Theta}$, 可得:

$$\sum_{y \in \boldsymbol{\Theta}} f(y) P(\theta_{t+1} = y \mid \theta_t = \boldsymbol{x}) \leq \sum_{y \in \boldsymbol{\Theta}} f(\boldsymbol{x}) P(\theta_{t+1} = y \mid \theta_t = \boldsymbol{x}) \leq f(\boldsymbol{x}) \sum_{y \in \boldsymbol{\Theta}} P(\theta_{t+1} = y \mid \theta_t = \boldsymbol{x}) \leq f(\boldsymbol{x}). \tag{20}$$

由于 $\boldsymbol{\Theta}$ 是一个闭集,故 $\sum_{y \in \boldsymbol{\Theta}} P(\theta_{t+1} = y \mid \theta_t = \boldsymbol{x}) = 1$,所以式(20)成立.综上,可以证得随机过程$\{f(\theta_t), t \geq 0\}$关于$\{\theta_t, t \geq 0\}$是一个非负有界上鞅函数列. 证毕.

定理 5 如果非负有界上鞅函数列$\{f(\theta_t), t \geq 0\}$满足 $\sup E \mid f(\theta_t) \mid < \infty$,那么当 $t \to \infty$ 时,一定存在一个群状态随机变量 $\theta^* \in \{\theta_t, t \geq 0\}$ 使得鸡群状态序列$\{\theta_t, t \geq 0\}$收敛,即 $P\{\lim_{t \to \infty} \theta_t = \theta^*\} = 1$.

证明 由于优化问题的全局最小适应度值 $f(\theta^*)$ 存在(否则优化问题无意义),则有:$f(\theta^*) \leq f(\theta_t) < \infty$,根据数学期望的性质有:$E[f(\theta^*)] \leq E[f(\theta_t)] < \infty$,所以 $\sup E \mid f(\theta_t) \mid < \infty$.又根据引理 2,当$t \to \infty$时,一定存在 $f(\theta^*) \in \{f(\theta_t), t \geq 0\}$,使得$\{f(\theta_t), t \geq 0\}$以概率 1 收敛到 $f(\theta^*)$.相应地,当 $t \to \infty$ 时,存在 $\theta^* \in \{\theta_t, t \geq 0\}$,使得$\{\theta_t, t \geq 0\}$以概率 1 收敛到 θ^*,即 $P\{\lim_{t \to \infty} \theta_t = \theta^*\} = 1$. 证毕.

由定理 5 中可知,若 θ^* 的取值是在全局最优状态集 $\boldsymbol{G}$ 中,则 CSO 算法将几乎处处强收敛至全局最优. 因此在定理 4 和定理 5 的条件下,记 $P_t^* = \min P(\theta_{t+1} \mid \theta_t)$ 为 CSO 算法从状态 θ_t 转移到 θ_{t+1} 的最小转移概率, 从而引出要证的定理 6.

定理 6 当 $\sum_{t=1}^{\infty} P_t^* = \infty$ 时,$P\{\lim_{t \to \infty}(\theta_t \subset \boldsymbol{G})\} = 1$, 即 CSO 算法以概率 1 几乎处处强收敛到全局最优.

证明 令 $M_t = \{\theta_t \cap \boldsymbol{G} = \varnothing \mid \theta_t = a, t \geq 0\} = \{f(\theta_t) \neq f(\theta^*)\}$,$N_t = \{\theta_{t+1} \cap \boldsymbol{G} \neq \varnothing \mid \theta_{t+1} = b, t \geq 0\} = \{f(\theta_t) = f(\theta^*)\}$.再令 $\beta = \min\{\mid f(a) - f(b) \mid f(a) \neq f(b)\}$,则存在一个正实数 λ,使得 $f(a) - f(b) \geq \lambda\beta$.由条件期望的性质可得:$E[f(\theta_t)] - E[f(\theta_{t+1})] = E[f(\theta_t)] - E\{E[f(\theta_{t+1} \mid \theta_t)]\} = \sum_{a \in \boldsymbol{\Theta}} P(\theta_t = a)\{f(\theta_t) - E[f(\theta_{t+1} \mid \theta_t = a)]\} = \sum_{a \in \boldsymbol{\Theta}} P(\theta_t = a) \sum_{b \in \boldsymbol{\Theta}} P(b \mid a)[f(a) - f(b)]$.根据 $P_t^* = \min P(\theta_{t+1} \mid \theta_t)$, $t \geq 0$ 与以下两个不等式:$\sum_{a \in \boldsymbol{\Theta}} P(b \mid a) \geq P(b \mid a)$,$\sum_{a \in \boldsymbol{\Theta}} P(\theta_t = a) \geq \sum_{a \cap \boldsymbol{G} = \varnothing} P(\theta_t = a)$,代入可得:$E[f(\theta_t)] - E[f(\theta_{t+1})] \geq \sum_{a \in \boldsymbol{\Theta}} P(\theta_t = a) \sum_{b \in \boldsymbol{\Theta}} P(b \mid a)\lambda\beta \geq \sum_{a \in \boldsymbol{\Theta}} P(\theta_t = a) P(b \mid a)\lambda\beta \geq \sum_{a \in \boldsymbol{\Theta}} P(\theta_t = a)\lambda\beta P_t^* \geq \sum_{a \cap \boldsymbol{G} = \varnothing} P(\theta_t = a)\lambda\beta P_t^* \geq \lambda\beta P_t^* P(M_t)$, 对上式中的 t 从 1 到 N 求和可得:$E[f(\theta_1)] \geq E[f(\theta_1)] - E[f(\theta_{N+1})] \geq \lambda\beta \sum_{t=1}^{N} P(M_t) P_t^*$,当 $N \to \infty$ 时,可得 $\sum_{t=1}^{\infty} P(M_t) P_t^* < \infty$.当 $\sum_{t=1}^{\infty} P_t^* = \infty$ 时,$P(M_t) \to 0$,有 $P\{\lim_{t \to \infty} f(\theta_t) \neq f(\theta^*), t \geq 0\} = 0$ 或 $P\{\lim_{t \to \infty} f(\theta_t) = f(\theta^*), t \geq 0\} = 1$,即:$P\{\lim_{t \to \infty} \theta_t \cap \boldsymbol{G} = \varnothing \mid \theta_t = a, t \geq 0\} = 0$ 或 $P\{\lim_{t \to \infty} \theta_t \subset \boldsymbol{G} \mid \theta_t = a, t \geq 0\} = 1$.当 $t \to \infty$ 时,鸡群状态序列$\{\theta_t, t \geq 0\}$收敛至群状态随机变量 θ^*.此外,$\theta^* \in \{\theta_t, t \geq 0\}$且 θ^* 的取值在全局最优状态集 $\boldsymbol{G}$ 中.即:当 $\sum_{t=1}^{\infty} P_t^* = \infty$ 时, CSO 算法几乎处处强收敛到全局最优.证毕.

3.3 CSO 算法的一致收敛性分析

定义 9(一致收敛)[16]　设在同一个可测集 F 上有函数列$\{f_n(\boldsymbol{x}),n\in N\}$和函数 $f(\boldsymbol{x})$，$\forall\varepsilon>0$，$\exists N(\varepsilon)>0$，使得当 $n\geq N(\varepsilon)$ 时，对 $\forall\boldsymbol{x}\in F$，都有 $|f_n(\boldsymbol{x})-f(\boldsymbol{x})|<\varepsilon$，则称在 F 上有函数列$\{f_n(\boldsymbol{x})\}$一致收敛于 $f(\boldsymbol{x})$.

引理 3(Egoroff 定理)[16]　若 $\mu(A)<\infty$，$\{f_n(\boldsymbol{x}),n\in N\}$是可测集 A 上几乎处处有限并可测的函数列，它在 A 上几乎处处收敛于函数 $f(\boldsymbol{x})$，即：$\lim\limits_{n\to\infty}f_n(\boldsymbol{x})=f(\boldsymbol{x})$.则对每个 $\varepsilon>0$，存在可测集 $A_c\subset A$，使得在 A_c 上函数列$\{f_n(\boldsymbol{x}),n\in N\}$一致收敛至函数 $f(\boldsymbol{x})$.

定理 7　函数列$\{f(\theta_t),t\geq 0\}$一致收敛于函数 $f(\theta^*)$.

证明　因鸡群状态空间 $\boldsymbol{\Theta}$ 是有限的，故 $|\boldsymbol{\Theta}|$ 中仅包含有限个不同的点.又因$\{f(\theta_t),t\geq 0\}$是非负有界上鞅，所以其极限几乎处处存在且是有界的.由定理 5 可知，$\lim\limits_{t\to\infty}f(\theta_t)=f(\theta^*)$，因此根据引理 3 可推出函数列$\{f(\theta_t),t\geq 0\}$具有一致收敛性，即对 $\forall\varepsilon>0$，$\exists N(\varepsilon)>0$，使得当 $t\geq N(\varepsilon)$ 时，对 $\forall\theta\in\boldsymbol{G}$，$\boldsymbol{G}\subset\boldsymbol{\Theta}$，有 $|f(\theta_t)-f(\theta^*)|<\varepsilon$.所以，函数列$\{f(\theta_t),t\geq 0\}$一致收敛于函数 $f(\theta^*)$，由此可推出 $\theta_t(\varepsilon)\subset\boldsymbol{G}$.也就是说，能确保在有限步数 $N(\varepsilon)$ 内，CSO 算法以概率 1 收敛到全局最优.一致收敛性具有“有限终止”的特点，因此它是强于几乎处处强收敛性的.证毕.

4 CSO 算法的实验与分析

为了验证本文理论证明的正确性及比较各算法的收敛性能和特点，本节决定对 DE 算法、PSO 算法和 CSO 算法进行仿真实验，即：选取具有不同特征的 16 个标准测试函数，再利用 DE、PSO 和 CSO 分别对它们进行重复 30 次的寻优计算，再从平均值、标准差、平均耗时 3 个方面来进行算法间的比较.其中，算法的收敛精度和寻优能力体现在平均值上，算法抗局部极值的能力和稳定性体现在标准差上，算法的快速性体现在平均耗时上.在 16 个标准测试函数中，$F_1\sim F_4$ 是单峰函数，分别是 Sphere，Schwefel's 2.22，Rosenbrock，Quartic，且最优值都为 0.单峰函数能很好地测试算法的收敛精度和寻优能力；$F_5\sim F_{10}$是多峰函数，分别是 Rastrigin，Griewank，Beale，Schaffer，Kowalik，Branin，且 Rastrigin，Griewank，Beale，Schaffer 的最优值为 0，Kowalik 和 Branin 的最优值分别为 0.000 3 和 0.398 0.多峰函数能很好地展现算法的全局搜索能力以及避免局部最优的能力；$F_{11}\sim F_{16}$是固定维多峰函数，分别是 2 维的 Shekel's Foxholes，Drop-Wave，Three-hump Camel，Levy N.13，4 维的 Styblinski-Tang 和 6 维的 Powell，且 Shekel's Foxholes 的最优值为 1，Drop-Wave 的最优值为－1，Three-hump Camel，Levy N.13 和 Powell 的最优值为 0，Styblinski-Tang 的最优值为－156.664 0.固定维多峰函数能很好地衡量算法在低维时的全局搜索能力.

此外，为了方便对算法进行比较，决定把 3 个算法的通用参数设置一样，固定维度除外，即：最大迭代次数 $M=1\ 000$，种群大小 $N=30$，维度 $D=30$.3 个算法的其他参数设置如下，CSO 算法：$N_r=0.2N$，$N_h=0.6N$，$N_m=0.1N_h$，$N_c=N-N_r-N_h$，$G=10$，$FL\in[0.5,0.9]$；PSO 算法：$c_1=c_2=2$，$w=0.75$；DE 算法：$F=0.5$，$CR=0.3$，$X_{\min}=-30$，$X_{\max}=30$. 在仿真实验中，若实验运行结果与函数的最优值相差小于 10^{-8}，就认为寻优成功.表 1 为实验运行得到的最终结果，并用粗体来标出每个测试函数中平均值、标准差、平均耗时的最优值.

在表 1 中，CSO 算法在函数 F_1，F_2，F_7，F_8，F_{10}，F_{12}，F_{13}，F_{14} 和 F_{16} 上都找到了它们的全局最小值，这表明 CSO 算法在有限的迭代次数内可以收敛到问题的全局最优，从而验证本文定理 7 中所得到的结论.从收敛精度和寻优能力来看，与其他两种算法相比，CSO 算法在 F_1，F_2，F_3，F_4，F_5，F_7，F_8，F_{10}，F_{12}，F_{13}，F_{14} 上收敛精度最高，寻优能力最强.虽然 DE 算法的收敛精度和寻优能力在函数 F_6，F_9，F_{11}，F_{15} 和 F_{16} 上是最高的，但 CSO 算法所得到的结果仅次于它.从对抗局部极值的能力和算法的稳定性来看，相较于 PSO 算法和 DE 算法，CSO 算法在函数 F_1，F_2，F_3，F_4，F_5，F_7，F_8，F_{12}，F_{13}，F_{14} 上抗局部极值能力较强，稳定性较好.而在 F_6，F_9，F_{10}，F_{11}，F_{15} 和 F_{16} 上，CSO 算法的抗局部极值能力和稳定性稍差一点，在 3 种算法中居于第 2.对于所有的标准测试函数，从平均耗时上看，PSO 算法平均耗时最短，CSO 算法的平均耗时要稍大于 DE 算法的平均耗时.总的来说，CSO 算法的收敛性能要好于 DE 算法和 PSO 算法，同时还有着较好的抗局部极值

能力和稳定性.虽然在计算时间上要稍弱于其他算法,但是对于计算时间要求不是特别高的优化问题,CSO算法比其他两种算法明显在应用上更具有优势.

表 1 测试函数运行结果

Tab. 1 Test function running results

函数	性能指标	CSO	PSO	DE	函数	性能指标	CSO	PSO	DE
F_1	mean	**3.447 2e-41**	0.083 6	4.057 7e-13	F_9	mean	0.001 1	0.001 3	**0.001 0**
	std	**1.809 7e-40**	0.046 9	3.836 3e-13		std	2.278 3e-04	0.003 6	**7.807 8e-06**
	time/s	0.464 8	**0.318 6**	0.438 3		time/s	0.490 6	**0.352 9**	0.469 0
F_2	mean	**1.779 7e-37**	3.293 0	3.642 5e-07	F_{10}	mean	**0.398 0**	0.397 9	0.397 9
	std	**6.192 8e-37**	1.501 0	1.271 4e-07		std	2.390 4e-04	**0**	**0**
	time/s	0.488 1	**0.349 4**	0.478 3		time/s	0.454 2	**0.295 0**	0.416 8
F_3	mean	**28.032 6**	70.714 9	34.986 9	F_{11}	mean	−3.817 8	−3.862 8	**−3.219 0**
	std	**0.593 2**	32.411 2	23.450 6		std	0.029 3	**0**	1.464 2
	time/s	0.454 0	**0.312 3**	0.426 6		time/s	0.523 3	**0.368 3**	0.521 1
F_4	mean	**0.009 3**	0.580 9	0.372 7	F_{12}	mean	**−1**	**−1**	**−1**
	std	**0.013 6**	0.271 4	0.095 3		std	**0**	**0**	**0**
	time/s	0.594 8	**0.428 0**	0.554 7		time/s	0.420 3	**0.278 0**	0.355 7
F_5	mean	**0.487 3**	55.696 8	88.081 1	F_{13}	mean	**4.324 4e-315**	4.596 4e-45	4.223 7e-176
	std	**2.668 8**	10.686 0	8.610 7		std	**0**	1.648 0e-44	**0**
	time/s	0.458 4	**0.330 1**	0.454 9		time/s	0.426 6	**0.282 0**	0.365 4
F_6	mean	0.003 4	0.010 2	**2.925 5e-13**	F_{14}	mean	**1.349 8e-31**	**1.349 8e-31**	**1.349 8e-31**
	std	0.014 3	0.016 7	**9.632 9e-13**		std	**6.680 9e-47**	**6.680 9e-47**	**6.680 9e-47**
	time/s	0.486 0	**0.334 6**	0.485 9		time/s	0.416 6	**0.283 7**	0.358 5
F_7	mean	**0**	0.024 5	**0**	F_{15}	mean	−154.200 3	−148.653 9	**−156.664 7**
	std	**0**	0.133 9	**0**		std	4.767 3	8.034 2	**0**
	time/s	0.446 0	**0.298 3**	0.413 9		time/s	0.431 4	**0.293 2**	0.385 1
F_8	mean	**0**	0.003 9	1.509 1e-05	F_{16}	mean	1.132 8e-09	−211.907 0	**9.285 9e-20**
	std	**0**	0.004 8	6.260 1e-05		std	3.214 7e-09	13.632 1	**2.435 1e-19**
	time/s	0.453 3	**0.300 6**	0.423 9		time/s	0.435 5	**0.297 0**	0.376 4

5 结 论

本文在 CSO 算法 Markov 链模型的基础上,通过一系列的定义、定理把具有最小适应度值的鸡群状态序列转化成一个上鞅,利用鞅方法证明了 CSO 算法的几乎处处强收敛性,进一步也证明了 CSO 算法的一致收敛性.此外,仿真实验也成功验证了本文理论证明的正确性并体现出了 CSO 算法的特点.与现有的结论相比,得到了 CSO 算法更强的收敛性证明结果,这将对未来 CSO 算法的改进和应用提供更多的理论基础.下一步将对 CSO 算法的改进优化进行研究,并利用鞅方法对改进的 CSO 算法进行更深入的收敛性理论分析.

参 考 文 献

[1] MENG X B,LIU Y,GAO X Z,et al.A new bio-inspired algorithm:chicken swarm optimization[C]// International Conference in Swarm Intelligence.Cham:Springer,2014:86-94.

[2] 李冰晓,万睿之,朱永杰,等.基于种群分区的多策略综合粒子群优化算法[J].河南师范大学学报(自然科学版),2022,50(3):85-94.
LI B X,WAN R Z,ZHU Y J,et al.Multi-strategy comprehensive article swarm optimization algorithm based on population partition[J].Journal of Henan Normal University(Natural Science Edition),2022,50(3):85-94.

[3] 方景远,季益胜,赵新超.基于高斯分布估计的对位差分进化算法[J].河南师范大学学报(自然科学版),2021,49(3):27-32.

FANG J Y，JI Y S，ZHAO X C.Opposition-based differential evolution algorithm with Gaussian distribution estimation[J].Journal of Henan Normal University(Natural Science Edition)，2021，49(3)：27-32.
[4] 郭堃，薛太林，耿杰，等.基于改进鸡群算法的无功优化综合分析[J].电气自动化，2021，43(6)：36-38.
GUO K，XUE T L，GENG J，et al.Comprehensive analysis of reactive power OptimizationBased on improved chicken population algorithm [J].Electrical Automation，2021，43(6)：36-38.
[5] 寇德昌.改进的鸡群算法及其应用[D].北京：北京建筑大学，2021.
[6] 王英聪，刘驰，王延峰.基于刺激-响应机制的改进鸡群算法[J].控制与决策，2023，38(1)：58-66.
WANG Y C，LIU C，WANG Y F.Chicken swarm optimization algorithm based on stimulus-response mechanism[J].Control and Decision，2023，38(1)：58-66.
[7] 杨菊蜻，张达敏，张慕雪，等.一种混合改进的鸡群优化算法[J].计算机应用研究，2018，35(11)：3290-3293.
YANG J Q，ZHANG D M，ZHANG M X，et al.Hybrid improved for chicken swarm optimization algorithm[J].Application Research of Computers，2018，35(11)：3290-3293.
[8] 李春红，陆安江，邵丽萍.基于线性递减权值更新的鸡群算法[J].智能计算机与应用，2020，10(2)：43-47.
LI C H，LU A J，SHAO L P.Rooster algorithm based on linear decrement weight update[J].Intelligent Computer and Applications，2020，10(2)：43-47.
[9] 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.
[10] 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.
[11] 张大力，夏红伟，张朝兴，等.改进萤火虫算法及其收敛性分析[J].系统工程与电子技术，2022，44(4)：1291-1300.
ZHANG D L，XIA H W，ZHANG C X，et al.Improved firefly algorithm and its convergence analysis[J].Systems Engineering and Electronics，2022，44(4)：1291-1300.
[12] 吴定会，孔飞，纪志成.鸡群算法的收敛性分析[J].中南大学学报(自然科学版)，2017，48(8)：2105-2112.
WU D H，KONG F，JI Z C.Convergence analysis of chicken swarm optimization algorithm[J].Journal of Central South University(Science and Technology)，2017，48(8)：2105-2112.
[13] 钱伟民，梁汉营，杨国庆.应用随机过程[M].北京：高等教育出版社，2014.
[14] 张文修，梁怡.遗传算法的数学基础[M].西安：西安交通大学出版社，2000.
[15] LAWLER G F.Introduction to stochastic processes[M].2nd ed.London：CRC Press，2006.
[16] 程其襄，张奠宙，魏国强，等.实变函数与泛函分析基础[M].2 版.北京：高等教育出版社，2003.

Convergence analysis of chicken swarm optimization algorithm based on martingale method

Zhou Tingting，Dai Jiajia

(School of Mathematics and Statistics，Guizhou University，Guiyang 550025，China)

Abstract：The most convergence analysis on chicken swarm optimization(CSO) algorithm belonged to weak convergence，and it cannot infer in general that the CSO algorithm would be convergent to a global optimum in a finite number of evolution steps. In order to make up for this deficiency，a martingale method was proposed to study the global convergence of CSO algorithm. Firstly，based on relevant definitions of CSO algorithm，the Markov chain model of CSO algorithm was established，and its Markov properties were analyzed. Secondly，the chicken swarm state sequence with the minimum fitness value was transformed into a supermartingale. By using the supermartingale convergence theorem and the Egoroff′s theorem，it was proved that the CSO algorithm had almost surely strong convergence and uniform convergence. Furthermore，it was concluded that the CSO algorithm can surely convergence to a global optimum in a finite number of evolution steps when the chicken swarm state space was finite. Finally，the validity of the theoretical proofs were verified successfully in the simulation experiment，and it was found that the CSO algorithm had stronger ability to search for excellence and higher convergence accuracy than PSO algorithm and DE algorithm.

Keywords：chicken swarm optimization(CSO) algorithm；Markov chain；supermartingale convergence theorem；Egoroff′s theorem；almost sure strong convergence；uniform convergence

[责任编校　陈留院　赵晓华]