

# 标准粒子群优化算法的马尔科夫链分析

潘峰<sup>1</sup> 周倩<sup>1</sup> 李位星<sup>1</sup> 高琪<sup>1</sup>

**摘要** 根据粒子群优化 (Particle swarm optimization, PSO) 算法的差分模型定义粒子状态序列和群体状态序列, 并分析其马尔科夫性质, 证明了粒子及种群的最优状态集的封闭性, 以及计算粒子一步转移概率; 进一步基于全概率公式和马氏链的性质, 推导了群体状态转到最优状态集的转移概率; 根据该转移概率, 对 PSO 算法的惯性权重  $\omega$  和加速度因子  $c$  进行了讨论和解释, 研究了算法早熟收敛和发散等问题, 最后分析表明标准 PSO 算法以一定概率收敛到全局最优。

**关键词** 粒子群优化算法, 马尔科夫链, 全概率公式, 全局收敛

**引用格式** 潘峰, 周倩, 李位星, 高琪. 标准粒子群优化算法的马尔科夫链分析. 自动化学报, 2013, 39(4): 381–389

**DOI** 10.3724/SP.J.1004.2013.00381

## Analysis of Standard Particle Swarm Optimization Algorithm Based on Markov Chain

PAN Feng<sup>1</sup> ZHOU Qian<sup>1</sup> LI Wei-Xing<sup>1</sup> GAO Qi<sup>1</sup>

**Abstract** According to the proposed particle swarm optimization (PSO) difference model in this paper, the state sequence of a single particle and swarm state sequence are defined first, and their Markov property are analyzed, after that, it is demonstrated that the set of optimal states are closed set. Moreover, the one-step transition probability of a particle is calculated. Considering the complete probability formula and the Markov properties, the transition probability to the optimal set is deduced. According to the derived conclusion, the inertia weight  $\omega$  and accelerate factor  $c$  of PSO are discussed. Finally, the premature convergence and divergent problem are explained, furthermore, it is proved that the standard PSO algorithm reaches the global optimum in probability.

**Key words** Particle swarm optimization (PSO), Markov chain, complete probability formula, global convergence

**Citation** Feng Pan, Qian Zhou, Wei-Xing Li, Qi Gao. Analysis of standard particle swarm optimization algorithm based on Markov chain. *Acta Automatica Sinica*, 2013, 39(4): 381–389

近年来, 越来越多的研究关注于标准的和各种改进的粒子群优化<sup>[1]</sup> (Particle swarm optimization, PSO) 算法在参数选择、稳定性、收敛性和统计特性等方面的研究. Shi 等<sup>[2]</sup> 的研究是早期通过统计实验结果, 对比不同参数取值对标准 PSO 优化性能影响的代表。

PSO 算法参数选择和稳定性的研究主要基于动态系统的各种相关稳定性理论. Trelea<sup>[3]</sup> 运用常系数线性动态系统理论进行了 PSO 的参数选取条件和稳定分析; Clerc 等<sup>[4]</sup> 建立了由 5 个参数描述的 PSO 约束模型, 分析了其收敛性和在相平面的粒子运动轨迹特性; 曾建潮等<sup>[5]</sup> 提出一种 PSO 的连续模型, 并给出了在不同参数取值条件下, PSO 进化

方程渐近收敛的条件; 文献 [6] 进一步建立了由积分环节与震荡环节组成的系统所对应的改进微粒群算法, 并讨论了参数的选择方法; 但是, PSO 算法中粒子运动的动态方程并不是一个线性时不变系统, 因此基于时不变动态系统的分析结果具有一定的局限性; Kadiramanathan 等<sup>[7]</sup> 将陷入停滞后的群体最优粒子描述为一个非线性反馈系统, 基于 Lyapunov 稳定性原理分析了 PSO 稳定的充分条件, 但由于其保守性, 对参数的选取非常严格, 粒子只能维持很小的局部振动; Fernández-Martínez 等<sup>[8]</sup> 证明 PSO 在物理上可以描述成一个离散的随机阻尼质量弹簧系统, 并且推断出不同 PSO 模型在保持探索和开发平衡上具有不同的性能, 其收敛性与其一阶和二阶系统的稳定性有关. 潘峰等<sup>[9]</sup> 将 PSO 作为动态时变系统, 分析其在无 Lipschitz 条件约束下稳定性的充分条件, 将惯性因子  $\omega$  值域扩展到  $(-1, 1)$ ; 并进一步分析说明了当群体停滞于解空间中某一局部极值点时算法的搜索性能<sup>[9–13]</sup> 以及同 BB-PSO 的相似性<sup>[14]</sup>.

PSO 算法收敛性的一类研究是以一般性随机搜索算法框架为出发点. van den Bergh 等<sup>[15–16]</sup> 证

收稿日期 2012-05-08 录用日期 2012-08-06  
Manuscript received May 8, 2012; accepted August 6, 2012  
国家自然科学基金 (60903005) 资助  
Supported by National Natural Science Foundation of China (60903005)  
本文责任编辑 方海涛  
Recommended by Associate Editor FANG Hai-Tao  
1. 北京理工大学自动化学院 北京 100081  
1. School of Automation, Beijing Institute of Technology, Beijing 100081  
该文的英文版同时发表在 *Acta Automatica Sinica*, vol. 39, no. 4, pp. 381–389, 2013.

明了尽管标准 PSO 非常高效,但并不是一种全局优化的方法,也不能保证一定收敛到某个局部最优点,并提出保证局部收敛性的改进; Jiang 等<sup>[17]</sup>对标准 PSO 在不同参数选择条件下对算法随机收敛性的影响进行了研究; Zhang 等<sup>[18]</sup>对随机优化算法的全局收敛性准则作了详细解释,并应用此全局收敛性准则对 PSO 算法的全局收敛性进行了理论分析,指出了 PSO 算法并不满足随机优化算法的全局收敛性准则应具备的两个条件。

基于随机过程理论的研究是 PSO 算法收敛性研究的另一个主要思路. Nakama<sup>[19]</sup>在遗传算法 (Genetic algorithm, GA) 中加入加法和乘法噪声,建立马尔科夫模型,分析了 GA 算法的转移特性和收敛性;论文最后提出了确保 GA 算法最终以概率 1 收敛到全局最优解的充要条件. Zhang 等<sup>[20]</sup>以成功概率来衡量 GA 算法的性能,并应用马尔科夫链理论来计算成功概率,表明成功概率随着字符串长度  $L$  的增加呈线性下降的趋势. Sahin 等<sup>[21]</sup>通过使用齐次马尔科夫链分析了 FGA 算法的收敛速度,运用 Dobrushin 收缩系数来衡量 FGA 马尔科夫链模型的收敛速度,并且运用两种不同的测试平台说明了所给仿生算法的有效性. 金欣磊等<sup>[22]</sup>从随机过程的角度出发,把带随机因子的 PSO 动力学模型在概率意义下转换成线性定常系统加以分析,给出了系统依概率收敛的一个充分条件;袁代林等<sup>[23]</sup>基于马氏链理论启发,使 PSO 算法中粒子具有遗忘特性,建立了算法的马氏链过程的数学模型;李宁等<sup>[24]</sup>证明了粒子状态链为马尔科夫链,以及种群状态序列是有限齐次马尔科夫链,进而也从差分方程的角度分析研究了粒子的运动轨迹,但是论文中没有给出利用马尔科夫链研究 PSO 收敛性等问题的进一步结论;任子晖等<sup>[25]</sup>从粒子状态所构成的马尔科夫链着手,证明了粒子状态空间的可约性和非齐次性,并验证了粒子状态空间是非常返态的,最后表明马尔科夫链不存在平稳过程的条件,也不是全局收敛的;蔡昭权等<sup>[26]</sup>建立了 PSO 的吸收态马尔科夫过程模型,并提出了可达状态集作为收敛性分析的关键指标和 PSO 全局收敛性改进的方法,但是有待给出对于标准 PSO 的收敛性分析的完整证明;Poli 等<sup>[27]</sup>利用有限元网格,建造了 BB-PSO (Bare bones particle swarm) 的离散的马氏链模型,分析说明了此模型可以逼近任意精度的连续问题,迭代链的转移矩阵可以获得每一代 BB-PSO 中粒子行为的精确信息,包括找到全局最优,或者寻到错误解的概率,同时对标准 PSO 算法的加速度因子分别采用柯西、高斯和其他概率分布的特点,进行了解释说明。

本文根据标准 PSO 算法的差分模型,在第 1 节定义粒子状态序列,分析 PSO 粒子状态序列的性质,并计算粒子状态的转移概率;第 2 节给出种群

状态序列的定义,利用全概率公式计算了群体状态转移到非最优状态的概率;根据该结论,第 3 节对 PSO 算法的早熟收敛和发散以及标准 PSO 以一定概率收敛到全局最优等问题进行分析. 论文分析证明过程的整体结构如图 1 所示。

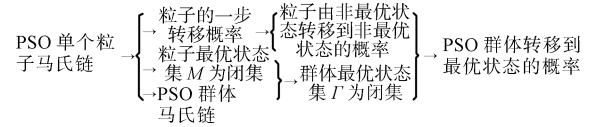


图 1 本文分析证明过程的整体结构

Fig. 1 The overall structure of certification process in this paper

## 1 标准 PSO 算法单个粒子马氏链分析

在任意  $k$  时刻,第  $i$  个粒子的速度、位置更新公式定义为

$$\begin{cases} v_i^{(k+1)} = \omega \cdot v_i^{(k)} + c_1 \cdot r_{i1}^{(k)} \cdot (p_i^{(k)} - x_i^{(k)}) + \\ c_2 \cdot r_{i2}^{(k)} \cdot (g^{(k)} - x_i^{(k)}) \\ x_i^{(k+1)} = x_i^{(k)} + v_i^{(k+1)} \end{cases} \quad (1)$$

其中,  $i = 1, 2, \dots, m$ ,  $\omega$  是惯量因子,  $x_i^{(k)}$  是粒子的当前的位置向量;  $v_i^{(k)}$  是粒子运动速度向量,  $p_i^{(k)}$  为粒子个体位置最优值;  $g^{(k)}$  为群体位置最优值;  $c_1, c_2$  是加速因子,  $r_{i1}^{(k)}, r_{i2}^{(k)}$  是  $[0, 1]$  之间的随机数. 根据式 (1) 推导得到 PSO 差分模型, 如式 (2) 所示:

$$\begin{aligned} x_i^{(k+1)} &= x_i^{(k)} + \omega \cdot (x_i^{(k)} - x_i^{(k-1)}) + \\ &\varphi_{i1}^k \cdot (p_i^{(k)} - x_i^{(k)}) + \varphi_{i2}^k \cdot (g^{(k)} - x_i^{(k)}) \end{aligned} \quad (2)$$

其中,  $\varphi_{i1}^k = c_1 r_{i1}^{(k)}, \varphi_{i2}^k = c_2 r_{i2}^{(k)}$ .

**定义 1.** 设  $\{y^{(k)}, k \geq 0\}$  为一列取值离散的随机变量, 离散值的集合为  $S = \{j\}$ , 称  $S$  为状态空间. 若对任意  $k \geq 1, i^{(l)} \in S (l \leq k+1)$  有:

$$\begin{aligned} p(y^{(k+1)} = i^{(k+1)} | y^{(k)} = i^{(k)}, \dots, y^{(0)} = i^{(0)}) = \\ p(y^{(k+1)} = i^{(k+1)} | y^{(k)} = i^{(k)}) \end{aligned} \quad (3)$$

称  $\{y^{(k)}, k \geq 0\}$  为马尔科夫链。

**定义 2.** 根据式 (2), 定义 PSO 群体中粒子  $i$  在  $k$  时刻的状态  $\xi_i^{(k)}$  为  $(x_i^{(k)}, x_i^{(k-1)}, p_i^{(k)}, g^{(k)})$ , 即  $\xi_i^{(k)} = (x_i^{(k)}, x_i^{(k-1)}, p_i^{(k)}, g^{(k)})$ .

**引理 1.** PSO 中粒子  $i$  在  $k$  时刻的状态  $\xi_i^{(k)}$  所构成的序列  $\{\xi_i^{(k)}, k \geq 1\}$  为马尔科夫链。

**证明.** 由式 (2) 定义有  $\{\xi_i^{(k)}, k \geq 1\}$  为一列取值离散的随机变量, 且  $\xi_i^{(k+1)} = (x_i^{(k+1)}, x_i^{(k)}, p_i^{(k+1)}, g^{(k+1)})$  的取值只和前一时刻的状态  $\xi_i^{(k)} = (x_i^{(k)}, x_i^{(k-1)}, p_i^{(k)}, g^{(k)})$  有关. 设  $\xi_i^{(k)}$  的状态空间为  $S$ , 对于任意的  $k \geq 1, \xi_i^{(l)} \in S (l \leq$

$k + 1$ ) 有:

$$p(\xi_i^{(k+1)} | \xi_i^{(k)}, \dots, \xi_i^{(0)}) = p(\xi_i^{(k+1)} | \xi_i^{(k)}) \quad (4)$$

因此 PSO 群体中粒子  $i$  在  $k$  时刻的状态  $\xi_i^{(k)} = (x_i^{(k)}, x_i^{(k-1)}, p_i^{(k)}, g^{(k)})$  所构成的序列  $\{\xi_i^{(k)}, k \geq 1\}$  为马尔科夫链.  $\square$

**引理 2.** 当 PSO 中加速度因子  $\varphi_{i1}^{(k)}, \varphi_{i1}^{(k)} \sim U(0, c)$ , 粒子  $i$  由  $x_i^{(k)}$  转移到以  $x_i^{(k+1)}$  为中心, 任意  $\varepsilon$  为半径的球区中的一步转移概率为

$$p(\xi_i^{(k+1)} | \xi_i^{(k)}) = \frac{\varepsilon}{\omega \|x_i^{(k)} - x_i^{(k-1)}\|} \frac{\varepsilon}{c \|p_i^{(k)} - x_i^{(k)}\|} \frac{\varepsilon}{c \|g^{(k)} - x_i^{(k)}\|}$$

其中当  $|f(p_i^{(k)}) - f(x_i^{(k)})| \leq \delta$  时,  $\frac{\varepsilon}{c \|p_i^{(k)} - x_i^{(k)}\|} = 1$ ,  $|f(g^{(k)}) - f(x_i^{(k)})| \leq \delta$  时,  $\frac{\varepsilon}{c \|g^{(k)} - x_i^{(k)}\|} = 1$  ( $\delta$  为一个趋于 0 的值).

**证明.** 考虑单个粒子模型:  $x_i^{(k+1)} = x_i^{(k-1)} + \omega(x_i^{(k)} - x_i^{(k-1)}) + \varphi_{i1}^{(k)}(p_i^{(k)} - x_i^{(k)}) + \varphi_{i2}^{(k)}(g^{(k)} - x_i^{(k)})$ , 因为  $\varphi_{i1}^{(k)}, \varphi_{i2}^{(k)}$  是随机数, 故  $x_i^{(k+1)}$  的取值由  $\omega, \varphi_{i1}^{(k)}, \varphi_{i2}^{(k)}$  确定. 考虑下面三种情况:

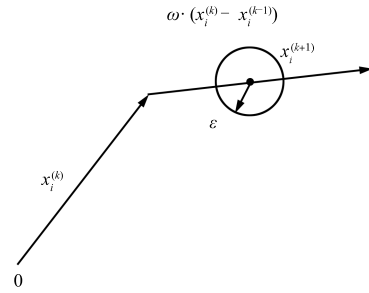
1) 当  $|f(p_i^{(k)}) - f(x_i^{(k)})| \leq \delta$  且  $|f(g^{(k)}) - f(x_i^{(k)})| \leq \delta$  时,  $x_i^{(k+1)} = x_i^{(k-1)} + \omega(x_i^{(k)} - x_i^{(k-1)})$ , 图形解释如图 2(a) 所示. 此时  $x_i^{(k+1)}$  的终点落在  $\omega(x_i^{(k)} - x_i^{(k-1)})$  向量上, 因此  $x_i^{(k+1)}$  可由  $\omega$  唯一确定; 当  $\omega$  为未知常数时, 有:

$$p(x_i^{(k+1)} | \xi_i^{(k)}) = \frac{\int_{x_i^{(k+1)} - 0.5\varepsilon}^{x_i^{(k+1)} + 0.5\varepsilon} dy}{\int_{x_i^{(k)}}^{x_i^{(k)} + \omega(x_i^{(k)} - x_i^{(k-1)})} dy} = \frac{\varepsilon}{\omega \|x_i^{(k)} - x_i^{(k-1)}\|}$$

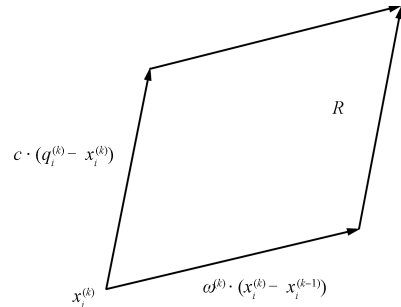
2) 当  $|f(p_i^{(k)}) - f(x_i^{(k)})| \leq \delta$  或  $|f(g^{(k)}) - f(x_i^{(k)})| \leq \delta$  时, 令  $q_i^{(k)} \in \{p_i^{(k)}, g^{(k)}\}$ ,  $x_i^{(k+1)} = x_i^{(k-1)} + \omega(x_i^{(k)} - x_i^{(k-1)}) + \varphi_i^{(k)}(q_i^{(k)} - x_i^{(k)})$ , 图形解释如图 2(b) 所示.  $\varphi_i^{(k)}$  为随机数, 故  $x_i^{(k+1)}$  位于由  $\omega(x_i^{(k)} - x_i^{(k-1)})$  和  $\varphi_i^{(k)}(q_i^{(k)} - x_i^{(k)})$  所确定的区域内. 因此,  $x_i^{(k+1)}$  的取值就由  $\omega, \varphi_i^{(k)}$  同时确定, 因此:

$$p(x_i^{(k+1)} | \xi_i^{(k)}) = \frac{\int_{x_i^{(k+1)} - 0.5\varepsilon}^{x_i^{(k+1)} + 0.5\varepsilon} dy}{\int_{x_i^{(k)}}^{x_i^{(k)} + \omega(x_i^{(k)} - x_i^{(k-1)})} dy} \cdot \frac{\int_{\varphi_i^{(k)} - 0.5\varepsilon}^{\varphi_i^{(k)} + 0.5\varepsilon} dy}{\int_{x_i^{(k)}}^{x_i^{(k)} + c(q_i^{(k)} - x_i^{(k)})} dy} = \frac{\varepsilon}{\omega \|x_i^{(k)} - x_i^{(k-1)}\|} \cdot \frac{\varepsilon}{c \|q_i^{(k)} - x_i^{(k)}\|}$$

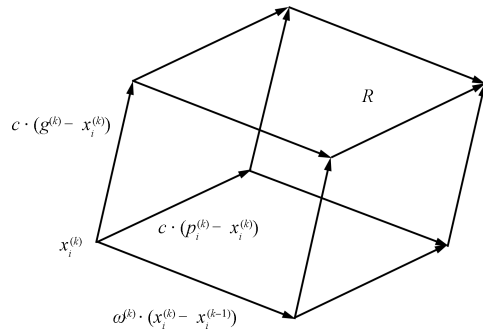
3) 当  $|f(p_i^{(k)}) - f(x_i^{(k)})| > \delta$  且  $|f(g^{(k)}) - f(x_i^{(k)})| > \delta$  时,  $x_i^{(k+1)} = x_i^{(k-1)} + \omega(x_i^{(k)} - x_i^{(k-1)}) + \varphi_{i1}^{(k)}(p_i^{(k)} - x_i^{(k)}) + \varphi_{i2}^{(k)}(g^{(k)} - x_i^{(k)})$ , 其图形解释如图 2(c) 所示. 当  $x_i^{(k+1)} \in R$  时, 只有在  $\omega, \varphi_{i1}^{(k)}, \varphi_{i2}^{(k)}$  同时确定时,  $x_i^{(k+1)}$  才能确定, 因此:



(a) 情况 1  
(a) Situation 1



(b) 情况 2  
(b) Situation 2



(c) 情况 3  
(c) Situation 3

图 2 单个粒子一步移动范围示例  
Fig. 2 Examples of single particle one-step moving range

$$p(x_i^{(k+1)} | \xi_i^{(k)}) = \frac{\int_{x_i^{(k+1)} - 0.5\varepsilon}^{x_i^{(k+1)} + 0.5\varepsilon} dy}{\int_{x_i^{(k)}}^{x_i^{(k)} + \omega(x_i^{(k)} - x_i^{(k-1)})} dy} \cdot \frac{\int_{\varphi_i^{(k)} - 0.5\varepsilon}^{\varphi_i^{(k)} + 0.5\varepsilon} dy}{\int_{x_i^{(k)}}^{x_i^{(k)} + c(q_i^{(k)} - x_i^{(k)})} dy}$$

$$\frac{\int_{\varphi_{i1}^{(k)}-0.5\varepsilon}^{\varphi_{i1}^{(k)}+0.5\varepsilon} dy}{\int_{x_i^{(k)}}^{x_i^{(k)}+c(p_i^{(k)}-x_i^{(k-1)})} dy} \cdot \frac{\int_{\varphi_{i2}^{(k)}-0.5\varepsilon}^{\varphi_{i2}^{(k)}+0.5\varepsilon} dy}{\int_{x_i^{(k)}}^{x_i^{(k)}+c(g^{(k)}-x_i^{(k-1)})} dy} = \frac{\varepsilon}{\omega\|x_i^{(k)}-x_i^{(k-1)}\|} \cdot \frac{\varepsilon}{c\|p_i^{(k)}-x_i^{(k)}\|} \cdot \frac{\varepsilon}{c\|g^{(k)}-x_i^{(k)}\|}$$

综合 1)~3) 有:

$$p(x_i^{(k+1)}|\xi_i^{(k)}) = \frac{\varepsilon}{\omega\|x_i^{(k)}-x_i^{(k-1)}\|} \cdot \frac{\varepsilon}{c\|p_i^{(k)}-x_i^{(k)}\|} \cdot \frac{\varepsilon}{c\|g^{(k)}-x_i^{(k)}\|}$$

其中,  $\lim_{p_i^{(k)} \rightarrow x_i^{(k)}} \frac{\varepsilon}{c\|p_i^{(k)}-x_i^{(k)}\|} = 1$ ,  $\lim_{g^{(k)} \rightarrow x_i^{(k)}} \frac{\varepsilon}{c\|g^{(k)}-x_i^{(k)}\|} = 1$ .

而  $p_i^{(k+1)}$ ,  $g^{(k+1)}$  由  $x_i^{(k+1)}$  确定, 因此粒子  $i$  从状态  $\xi_i^{(k)}$  转移到  $\xi_i^{(k+1)}$  的概率为

$$p(\xi_i^{(k+1)}|\xi_i^{(k)}) = \frac{\varepsilon}{\omega\|x_i^{(k)}-x_i^{(k-1)}\|} \cdot \frac{\varepsilon}{c\|p_i^{(k)}-x_i^{(k)}\|} \cdot \frac{\varepsilon}{c\|g^{(k)}-x_i^{(k)}\|} \quad (5)$$

其中,  $\varepsilon$  为  $x_i^{(k+1)}$  的取值半径, 当  $|f(p_i^{(k)}) - f(x_i^{(k)})| \leq \delta$  时,  $\frac{\varepsilon}{c\|p_i^{(k)}-x_i^{(k)}\|} = 1$ ,  $|f(g^{(k)}) - f(x_i^{(k)})| \leq \delta$  时,  $\frac{\varepsilon}{c\|g^{(k)}-x_i^{(k)}\|} = 1$  ( $\delta$  为一个趋于 0 的值). □

**定义 3.** 粒子最优状态集. 设优化问题的全局最优为  $g^*$ , 那么粒子最优状态集为:  $M = \{\xi^* = (x^{(k)}, x^{(k-1)}, g^*, g^*), k \geq 1\}$ , 其中,  $\xi^* \in S$  为粒子最优状态. 若  $M = S$ , 则在可行域中, 每个解是最优解, 优化问题就被解决了, 因此考虑  $M \subset S$  的情况.

**引理 3.** 粒子最优状态集  $M$  是闭集.

**证明.** 根据标准 PSO 的选择策略, 粒子状态  $\xi_i^{(k)} = (x_i^{(k)}, x_i^{(k-1)}, p_i^{(k)}, g^{(k)})$  为最优状态时, 即  $\xi_i^{(k)} = \xi^* = (x^{(k)}, x^{(k-1)}, g^*, g^*)$ , 其下一个时刻的状态  $\xi_i^{(k+1)} = (x^{(k+1)}, x^{(k)}, g^*, g^*)$  也为最优状态, 即如果  $\xi_i^{(k)} \in M, k \in [0, T]$ , 那么  $\forall l \in [k+1, T], \xi_i^{(l)} \in M$ . 因此有:

$$\begin{cases} P(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \in M) = 1 \\ P(\xi_i^{(k+1)} \notin M | \xi_i^{(k)} \in M) = 0 \end{cases} \quad (6)$$

所以粒子最优状态集  $M = \{\xi^* = (x^{(k)}, x^{(k-1)}, g^*, g^*), k \geq 1\}$ , 其中,  $\xi^* \in S$  是闭集. □

**引理 4.** 当粒子  $k$  时刻的状态  $\xi_i^{(k)} \notin M$  时, 其

下一时刻的状态  $\xi_i^{(k+1)} \notin M$  的概率为

$$p(\xi_i^{(k+1)} \notin M | \xi_i^{(k)} \notin M) = 1 - \frac{\varepsilon_k}{\omega\|x_i^{(k)}-x_i^{(k-1)}\|} \cdot \frac{\varepsilon_k}{c\|p_i^{(k)}-x_i^{(k)}\|} \cdot \frac{\varepsilon_k}{c\|g^{(k)}-x_i^{(k)}\|}$$

其中, 当  $g^* \in R$  时,  $\varepsilon_k = \varepsilon$ ; 当  $g^* \notin R$  时,  $\varepsilon_k = 0$ .

**证明.** 分析标准 PSO 算法中粒子的运动规律, 粒子在  $k$  时刻的状态  $\xi_i^{(k)} \notin M$  已知, 由于加速度因子  $\varphi_{i1}^{(k)}, \varphi_{i2}^{(k)}$  的不确定性, 粒子在  $k+1$  时刻的位置存在于一个可行域  $R$  内, 因此全局最优  $g^*$  的位置未知, 即可能位于  $R$  域中, 也可能位于  $R$  域之外, 根据引理 2, 在  $k$  时刻  $\forall \xi_i^{(k+1)} \in R (k \geq 1)$ , 有:

$$p(\xi_i^{(k+1)}|\xi_i^{(k)}) = \frac{\varepsilon}{\omega \cdot \|x_i^{(k)}-x_i^{(k-1)}\|} \cdot \frac{\varepsilon}{c \cdot \|p_i^{(k)}-x_i^{(k)}\|} \cdot \frac{\varepsilon}{c \cdot \|g^{(k)}-x_i^{(k)}\|}$$

而:

$$p(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M) = \begin{cases} 0, & g^* \notin R \\ \frac{\varepsilon}{\omega \cdot \|x_i^{(k)}-x_i^{(k-1)}\|} \cdot \frac{\varepsilon}{c \cdot \|p_i^{(k)}-x_i^{(k)}\|} \cdot \frac{\varepsilon}{c \cdot \|g^{(k)}-x_i^{(k)}\|}, & g^* \in R \end{cases} \quad (7)$$

又有:

$$P(\xi_i^{(k+1)} \notin M | \xi_i^{(k)} \notin M) = 1 - P(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M)$$

因此:

$$p(\xi_i^{(k+1)} \notin M | \xi_i^{(k)} \notin M) = \begin{cases} 1, & g^* \notin R \\ 1 - \frac{\varepsilon}{\omega \cdot \|x_i^{(k)}-x_i^{(k-1)}\|} \cdot \frac{\varepsilon}{c \cdot \|p_i^{(k)}-x_i^{(k)}\|} \cdot \frac{\varepsilon}{c \cdot \|g^{(k)}-x_i^{(k)}\|}, & g^* \in R \end{cases} \quad (8)$$

综合式 (8), 可得:

$$p(\xi_i^{(k+1)} \notin M | \xi_i^{(k)} \notin M) = 1 - \frac{\varepsilon_k}{\omega \cdot \|x_i^{(k)}-x_i^{(k-1)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|p_i^{(k)}-x_i^{(k)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|g^{(k)}-x_i^{(k)}\|} \quad (9)$$

当  $g^* \in R$  时,  $\varepsilon_k = \varepsilon$ ; 当  $g^* \notin R$  时,  $\varepsilon_k = 0$ . □

## 2 PSO 群体马氏链分析

**定义 4.** PSO 群体状态, 定义为:  $\zeta^{(k)} = (\xi_1^{(k)}, \xi_2^{(k)}, \dots, \xi_m^{(k)})$ ,  $m$  为群体粒子个数,  $k$  为当前迭代次数.

**引理 5.** 由  $m$  个粒子构成的 PSO 群体, 其状态序列  $\{\zeta^{(k)} = (\xi_1^{(k)}, \xi_2^{(k)}, \dots, \xi_m^{(k)}), k \geq 1\}$  构成马尔科夫链.

**证明.** 设  $\{\zeta^{(k)}, k \geq 1\}$  的状态空间为  $\Lambda$ , 对于任意的  $k \geq 1, \zeta^{(l)} \in \Lambda (l \leq k + 1)$  有:

$$p(\zeta^{(k+1)}|\zeta^{(k)}, \dots, \zeta^{(k)}) = p(\xi_1^{(k+1)}, \dots, \xi_m^{(k+1)}|\xi_1^{(k)}, \dots, \xi_m^{(k)}, \dots, \xi_1^{(1)}, \dots, \xi_m^{(1)}) \quad (10)$$

由引理 1 有  $p(\xi^{(k+1)}|\xi^{(k)}, \dots, \xi^{(k)}) = p(\zeta^{(k+1)}|\zeta^{(k)})$ , 因此:

$$p(\zeta^{(k+1)}|\zeta^{(k)}, \dots, \zeta^{(k)}) = p(\xi_1^{(k+1)}, \dots, \xi_m^{(k+1)}|\xi_1^{(k)}, \dots, \xi_m^{(k)}) = p(\zeta^{(k+1)}|\zeta^{(k)}) \quad (11)$$

因此, 由  $m$  个粒子构成的 PSO 群体, 其状态序列  $\{\zeta^{(k)} = (\xi_1^{(k)}, \xi_2^{(k)}, \dots, \xi_m^{(k)}), k \geq 1\}$  构成一个马尔科夫链.  $\square$

**定义 5 (群体最优状态集  $\Gamma$ ).** PSO 群体状态为  $\zeta^{(k)} = (\xi_1^{(k)}, \xi_2^{(k)}, \dots, \xi_m^{(k)})$ , 定义群体最优状态为  $\zeta_k^*$ , 对于  $\zeta_k^* = (\xi_1^{(k)}, \dots, \xi_i^{(k)}, \dots, \xi_m^{(k)})$ ,  $\exists i \in [1, m]$ , 使得  $\xi_i^{(k)} \in M$  群体最优状态集  $\Gamma = \{\zeta_k^*, k \geq 1 | \zeta_k^* = (\xi_1^{(k)}, \dots, \xi_i^{(k)}, \dots, \xi_m^{(k)}), \exists i \in [1, m], \text{使 } \xi_i^{(k)} \in M\}$ .

若群体最优状态集  $\Gamma = \Lambda$  (群体状态空间), 表示 PSO 群体状态都为最优状态, 即表示算法已完成优化; 因此考虑  $\Gamma \subset \Lambda$  的情况.

**引理 6.** 群体最优状态集  $\Gamma$  为闭集.

**证明.** 由引理 3 知, 粒子最优状态集  $M$  为闭集, 因此有:

$$P(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \in M) = 1$$

而  $\Gamma = \{\zeta_k^*, k \geq 1 | \zeta_k^* = (\xi_1^{(k)}, \dots, \xi_i^{(k)}, \dots, \xi_m^{(k)}), \exists i \in [1, m], \text{有 } \xi_i^{(k)} \in M\}$ , 如果在  $k$  时刻群体状态  $\zeta^{(k)} = (\xi_1^{(k)}, \xi_2^{(k)}, \dots, \xi_m^{(k)}) \in \Gamma$ , 即对于群体状态  $\zeta^{(k)}, \exists i \in [1, m], \text{有 } \xi_i^{(k)} \in M$ , 那么在  $k+1$  时刻必有  $\xi_i^{(k+1)} \in M$ , 因此在  $k+1$  时刻群体状态:  $\zeta^{(k+1)} = (\xi_1^{(k+1)}, \dots, \xi_m^{(k+1)}), \exists i \in [1, m], \text{有 } \xi_i^{(k+1)} \in M$ , 因此  $\zeta^{(k+1)} \in \Gamma$ , 所以有:

$$\begin{cases} p(\zeta^{(k+1)} \in \Gamma | \zeta^{(k)} \in \Gamma) = 1 \\ p(\zeta^{(k+1)} \notin \Gamma | \zeta^{(k)} \in \Gamma) = 0 \end{cases} \quad (12)$$

因此群体最优状态集为  $\Gamma$  闭集.  $\square$

**引理 7.** 群体在  $k+1$  时刻的状态  $\zeta^{(k+1)} \in \Gamma$  的

概率为

$$P(\zeta^{(k+1)} \in \Gamma) = 1 - p(\zeta^{(1)} \notin \Gamma) \cdot \prod_{l=1}^k \prod_{i=1}^m \left(1 - \frac{\varepsilon_l}{\omega \cdot \|x_i^{(l)} - x_i^{(l-1)}\|} \cdot \frac{\varepsilon_l}{c \cdot \|p_i^{(l)} - x_i^{(l)}\|} \cdot \frac{\varepsilon_l}{c \cdot \|g^{(l)} - x_i^{(l)}\|}\right)$$

**证明.** 根据全概率公式, 群体在  $k+1$  时刻的状态  $\zeta^{(k+1)} \notin \Gamma$  的概率为

$$p(\zeta^{(k+1)} \notin \Gamma) = p(\zeta^{(k+1)} \notin \Gamma | \zeta^{(k)} \notin \Gamma) \cdot p(\zeta^{(k)} \notin \Gamma) + p(\zeta^{(k+1)} \notin \Gamma | \zeta^{(k)} \in \Gamma) \cdot p(\zeta^{(k)} \in \Gamma)$$

而  $p(\zeta^{(k+1)} \notin \Gamma | \zeta^{(k)} \in \Gamma) = 0$ . 因此:

$$\begin{aligned} p(\zeta^{(k+1)} \notin \Gamma) &= p(\zeta^{(k+1)} \notin \Gamma | \zeta^{(k)} \notin \Gamma) \cdot p(\zeta^{(k)} \notin \Gamma) = \\ &= p(\zeta^{(k+1)} \notin \Gamma | \zeta^{(k)} \notin \Gamma) \cdot p(\zeta^{(k)} \notin \Gamma | \zeta^{(k-1)} \notin \Gamma) \cdots p(\zeta^{(1)} \notin \Gamma) = \\ &= p(\zeta^{(1)} \notin \Gamma) \cdot \prod_{l=1}^k p(\zeta^{(l+1)} \notin \Gamma | \zeta^{(l)} \notin \Gamma) \end{aligned}$$

而群体状态  $\zeta^{(k)} = (\xi_1^{(k)}, \xi_2^{(k)}, \dots, \xi_m^{(k)}) \notin \Gamma$  为:  $\forall i \in [1, m], \exists \xi_i^{(k)} \in M$ , 其中  $\exists$  为不存在. 因此:

$$p(\zeta^{(k+1)} \notin \Gamma | \zeta^{(k)} \notin \Gamma) = \prod_{i=1}^m p(\xi_i^{(k+1)} \notin M | \xi_i^{(k)} \notin M) = \prod_{i=1}^m \left(1 - \frac{\varepsilon_k}{\omega \cdot \|x_i^{(k)} - x_i^{(k-1)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|p_i^{(k)} - x_i^{(k)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|g^{(k)} - x_i^{(k)}\|}\right) \quad (13)$$

所以:

$$\begin{aligned} p(\zeta^{(k+1)} \notin \Gamma) &= p(\zeta^{(1)} \notin \Gamma) \cdot \prod_{l=1}^k p(\zeta^{(l+1)} \notin \Gamma | \zeta^{(l)} \notin \Gamma) = \\ &= p(\zeta^{(1)} \notin \Gamma) \cdot \prod_{l=1}^k \prod_{i=1}^m \left(1 - \frac{\varepsilon_l}{\omega \cdot \|x_i^{(l)} - x_i^{(l-1)}\|} \cdot \frac{\varepsilon_l}{c \cdot \|p_i^{(l)} - x_i^{(l)}\|} \cdot \frac{\varepsilon_l}{c \cdot \|g^{(l)} - x_i^{(l)}\|}\right) \end{aligned}$$

故群体在  $k+1$  时刻的状态  $\zeta^{(k+1)} \in \Gamma$  的概率为

$$\begin{aligned} p(\zeta^{(k+1)} \in \Gamma) &= 1 - p(\zeta^{(k+1)} \notin \Gamma) = \\ &= 1 - p(\zeta^{(1)} \notin \Gamma) \cdot \prod_{l=1}^k \prod_{i=1}^m \left(1 - \frac{\varepsilon_l}{\omega \cdot \|x_i^{(l)} - x_i^{(l-1)}\|} \cdot \frac{\varepsilon_l}{c \cdot \|p_i^{(l)} - x_i^{(l)}\|} \cdot \frac{\varepsilon_l}{c \cdot \|g^{(l)} - x_i^{(l)}\|}\right) \end{aligned} \quad (14)$$

$\square$

### 3 分析与讨论

#### 3.1 群体规模的影响

考虑式 (13), 当  $m \rightarrow \infty$  时, 即群体中粒子个数为无穷时, 由于粒子随机分配状态 (即粒子之间相互独立), 且  $1 - \frac{\varepsilon_k}{\omega \cdot \|x_i^{(k)} - x_i^{(k-1)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|p_i^{(k)} - x_i^{(k)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|g^{(k)} - x_i^{(k)}\|} < 1$ , 因此:  $\prod_{i=1}^{\infty} (1 - \frac{\varepsilon_k}{\omega \cdot \|x_i^{(k)} - x_i^{(k-1)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|p_i^{(k)} - x_i^{(k)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|g^{(k)} - x_i^{(k)}\|}) \rightarrow 0$  即  $p(\zeta^{(k+1)} \notin \Gamma | \zeta^k \notin \Gamma) \rightarrow 0$ , 则  $p(\zeta^{(k+1)} \in \Gamma) \rightarrow 1$ . 说明当群体中粒子个数趋于无穷时, 算法以概率 1 收敛至全局最优. 其特点类似于在解空间中进行无数次随机抽样.

#### 3.2 PSO 惯性权重 $\omega$ 的分析

考虑  $x_i^{(k+1)} = x_i^{(k)} + \omega \cdot (x_i^{(k)} - x_i^{(k-1)}) + \varphi_{i1}^{(k)} \cdot (p_i^{(k)} - x_i^{(k)}) + \varphi_{i2}^{(k)} \cdot (g^{(k)} - x_i^{(k)})$ , 粒子  $i$  的可达域由  $\omega, \varphi_{i1}^{(k)}, \varphi_{i2}^{(k)}$  确定, 当  $x_i^{(k)}$  为多维时,  $\varphi_{i1}^{(k)}, \varphi_{i2}^{(k)} \sim U(0, c), \varphi_{i1}^{(k)} \cdot (p_i^{(k)} - x_i^{(k)}) + \varphi_{i2}^{(k)} \cdot (g^{(k)} - x_i^{(k)})$  所覆盖的区域可形成超球体, 而  $\omega \cdot (x_i^{(k)} - x_i^{(k-1)})$  则是将此超球体转移了一个步长, 如图 3 所描述.

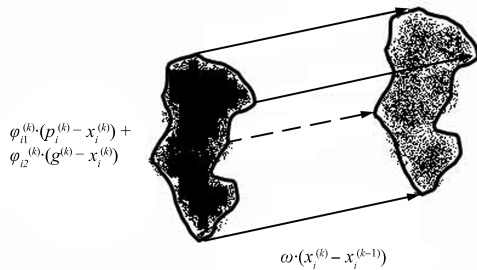


图 3 粒子可达域  $R$  图示

Fig. 3 The graphical representation of particle reachable area  $R$

所以,  $\varphi_{i1}^{(k)}, \varphi_{i2}^{(k)}$  确定可达域  $R$  的大小,  $|\omega|$  确定可达域  $R$  的步长, 也就是说  $\varphi_{i1}^{(k)}, \varphi_{i2}^{(k)}$  确定的超立方体按照  $|\omega|$  的大小, 在搜索空间中移动, 从而寻找全局最优. 因此,  $|\omega|$  就会影响可达域  $R$  所能达到的范围, 同时也控制着可达域  $R$  的搜索精度. 较大的  $|\omega|$  可以增大算法对搜索空间的搜索, 却使得搜索精度减小, 可能错过全局最优  $g^*$ , 甚者会使算法发散, 从而无法搜索至全局最优; 而较小的  $|\omega|$  会提高搜索精度, 但是搜索时间增加, 甚者早熟收敛至局部最优.

考虑式 (7) 和式 (14):

$$p(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M) = \frac{\varepsilon_k}{\omega \cdot \|x_i^{(k)} - x_i^{(k-1)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|p_i^{(k)} - x_i^{(k)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|g^{(k)} - x_i^{(k)}\|}$$

其中, 当  $g^* \in R$  时,  $\varepsilon_k = \varepsilon$ ; 当  $g^* \notin R$  时,  $\varepsilon_k = 0$ .

若  $k+1$  时刻  $g^* \notin R$  时,  $\varepsilon_k = 0$ , 故  $p(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M) = 0$ , 若  $k+1$  时刻  $g^* \in R$  时,  $\varepsilon_k = \varepsilon$ , 此时  $p(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M)$  与  $|\omega|$  成反比, 也就是说  $|\omega|$  确定的步长越大时,  $p(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M)$  的值越小, 算法容易错过全局最优; 而在  $|\omega|$  较小时,  $p(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M)$  较大, 说明保证全局最优  $g^*$  位于可达域  $R$  中时, 较小的  $|\omega|$  可使算法具有较小的步长, 使得粒子在可达域内搜索精度增大, 从而搜索到全局最优. 但是  $|\omega|$  控制算法的步长, 即控制可达域  $R$  的大小, 因此,  $|\omega|$  也会影响全局最优  $g^*$  是否位于可达域  $R$  中, 较大的  $|\omega|$  使  $g^* \in R$  的概率增大, 反之较小的  $|\omega|$  使  $g^* \in R$  的概率减小. 因此  $|\omega|$  在取值比较合适的情况下, 才能使  $p(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M)$  具有较好的值.

#### 3.3 PSO 加速度因子讨论

考虑式 (7) 和式 (14):

$$p(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M) = \frac{\varepsilon_k}{\omega \cdot \|x_i^{(k)} - x_i^{(k-1)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|p_i^{(k)} - x_i^{(k)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|g^{(k)} - x_i^{(k)}\|}$$

$$p(\zeta^{(k+1)} \in \Gamma) = 1 - p(\zeta^{(k+1)} \notin \Gamma) = 1 - p(\zeta^{(1)} \notin \Gamma) \cdot \prod_{l=1}^k \prod_{i=1}^m (1 - \frac{\varepsilon_l}{\omega \cdot \|x_i^{(l)} - x_i^{(l-1)}\|} \cdot \frac{\varepsilon_l}{c \cdot \|p_i^{(l)} - x_i^{(l)}\|} \cdot \frac{\varepsilon_l}{c \cdot \|g^{(l)} - x_i^{(l)}\|})$$

其中, 当  $g^* \in R$  时,  $\varepsilon_l = \varepsilon$ ; 当  $g^* \notin R$  时,  $\varepsilon_l = 0$ ;  $\varphi_{i1}^{(k)}, \varphi_{i2}^{(k)} \sim U(0, c)$ .

##### 3.3.1 $c$ 取值过小

当  $c$  取值过小时, 即便算法初期全局最优  $g^{(k)}$  和粒子最优位置  $p_i^{(k)}$  与当前位置  $x_i^{(k)}$  相聚较远, 但  $c$  取值过小, 使  $c \cdot \|p_i^{(k)} - x_i^{(k)}\| \cdot c \cdot \|g^{(k)} - x_i^{(k)}\|$  的值较小, 从而使可行域  $R$  的区域过小,  $g^{(k)}, p_i^{(k)}$  随着迭代的进行, 越来越接近于  $x_i^{(k)}$ , 从而使  $c \cdot \|p_i^{(k)} - x_i^{(k)}\| \cdot c \cdot \|g^{(k)} - x_i^{(k)}\|$  随着迭代减小; 若  $g^* \in R, \varepsilon_k = \varepsilon > 0$ , 则  $\lim_{k \rightarrow \infty} P(\zeta^{(k+1)} \in \Gamma) = 1$ , 算法能快速收敛到全局最优  $g^*$ ; 若  $g^* \notin R, \varepsilon_k = 0$ , 因此  $\lim_{k \rightarrow \infty} P(\zeta^{(k+1)} \in \Gamma) = 0$ , 算法就快速收敛到一点 (非全局最优). 因此, 若  $c$  取值过小, 可达域  $R$  越来越小, 使群体快速收敛到一点, 算法出现早熟收敛.

##### 3.3.2 $c$ 取值过大

观察可达域  $R$ , 在惯性权重  $\omega$  取值确定, 即便当前时刻的  $g^{(k)}, p_i^{(k)}$  与  $x_i^{(k)}$  接近, 但是由于  $c$  取值过大, 使  $c \cdot \|p_i^{(k)} - x_i^{(k)}\| \cdot c \cdot \|g^{(k)} - x_i^{(k)}\|$  过大, 使  $k+1$  时刻的  $g^{(k+1)}, p_i^{(k+1)}$  与  $x_i^{(k)}$  距离很远, 从而导致  $c \cdot \|p_i^{(k)} - x_i^{(k)}\| \cdot c \cdot \|g^{(k)} - x_i^{(k)}\|$  越来越大, 故  $c \cdot \|p_i^{(k)} - x_i^{(k)}\| \cdot c \cdot \|g^{(k)} - x_i^{(k)}\|$  为单调增函数, 即

随着迭代进行, 可达域  $R$  不断增大, 从而导致算法发散.

当算法发散时, 可达域  $R$  的范围随着迭代增大, 最优位置  $g^* \in R$  的概率增大, 可以认为算法有可能搜索到  $g^*$ , 但是通过对  $k \rightarrow \infty$  时,  $P(\zeta^{(k+1)} \in \Gamma)$  的取值, 就会得出否定回答.

PSO 发散时,  $c \cdot \|p_i^{(k)} - x_i^{(k)}\| \cdot c \cdot \|g^{(k)} - x_i^{(k)}\|$  为单调增函数, 随迭代进行不断增大, 因此可以认为  $p(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M) = \frac{\varepsilon_k}{c \cdot \|p_i^{(k)} - x_i^{(k)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|g^{(k)} - x_i^{(k)}\|}$  是随时间单调递减的函数, 即  $P(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M)$  越来越小且  $\lim_{k \rightarrow \infty} p(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M) \rightarrow 0$ , 因此:

$$\lim_{k \rightarrow \infty} \prod_{l=1}^k \prod_{i=1}^m \frac{\varepsilon_l^2}{c^2 \cdot \|p_i^l - x_i^l\| \cdot \|g^l - x_i^l\|} \rightarrow 0$$

所以:

$$\lim_{k \rightarrow \infty} p(\zeta^k \in \Gamma) = 1 - p(\zeta^1 \notin \Gamma) \cdot \prod_{l=1}^{\infty} \prod_{i=1}^m (1 - \frac{\varepsilon_l^2}{c^2 \cdot \|p_i^l - x_i^l\| \cdot \|g^l - x_i^l\|}) = 0$$

说明当 PSO 算法发散时, 即便可达域  $R$  很大, 使全局最优  $g^* \in R$ , 也不能使算法搜索到最优位置  $g^*$ . 因此不是保证全局最优位置  $g^* \in R$ , 就能保证算法搜索到  $g^*$ , 而是在  $g^* \in R$  的同时,  $R$  的范围的取值保持在一个合适的区域内, 即  $p(\xi_i^{(k+1)} \notin M | \xi_i^{(k)} \notin M) = 1 - \frac{\varepsilon_k}{c \cdot \|p_i^{(k)} - x_i^{(k)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|g^{(k)} - x_i^{(k)}\|} < 1$ , 同时  $p(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M) = \frac{\varepsilon_k}{c \cdot \|p_i^{(k)} - x_i^{(k)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|g^{(k)} - x_i^{(k)}\|}$  不是一个随时间单调递减的函数, 才会在迭代时间趋于无穷时有:  $\lim_{k \rightarrow \infty} p(\zeta^{(k)} \in M) = 1$ .

### 3.4 标准 PSO 算法以一定的概率搜索到全局最优

考虑式子  $p(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M) = \frac{\varepsilon_k}{|\omega| \cdot \|x_i^{(k)} - x_i^{(k-1)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|p_i^{(k)} - x_i^{(k)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|g^{(k)} - x_i^{(k)}\|}$ , 由于标准 PSO 算法的选择机制存在随机性, 因此不能确保在每次迭代过程中, 最优位置  $g^*$  总位于可达域  $R$  内, 即  $\varepsilon_k \geq 0$ , 因此:

$$p(\xi_i^{(k+1)} \in M | \xi_i^{(k)} \notin M) = \frac{\varepsilon_k}{|\omega| \cdot \|x_i^{(k)} - x_i^{(k-1)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|p_i^{(k)} - x_i^{(k)}\|} \cdot \frac{\varepsilon_k}{c \cdot \|g^{(k)} - x_i^{(k)}\|} \leq 1$$

从而导致  $0 \leq p(\zeta^\infty \in M) \leq 1$ , 也就是说, 标准 PSO 算法既不能保证搜索到全局最优, 也不能保证搜索不到全局最优.

若在 PSO 算法优化过程中, 存在某些时刻, 使得最优位置  $g^* \in R$ , 并且算法参数  $\omega, c$  选择保证算

法不发散, 则有  $j \in \{1, 2, \dots, n\}$ , 使:

$$p(\xi_i^{(kj+1)} \in M | \xi_i^{(kj)} \notin M) = \frac{\varepsilon}{|\omega| \cdot \|x_i^{(kj)} - x_i^{(kj-1)}\|} \cdot \frac{\varepsilon}{c \cdot \|p_i^{(kj)} - x_i^{(kj)}\|} \cdot \frac{\varepsilon}{c \cdot \|g^{(kj)} - x_i^{(kj)}\|} > \delta$$

因此有:

$$1 - \frac{\varepsilon}{|\omega| \cdot \|x_i^{(kj)} - x_i^{(kj-1)}\|} \cdot \frac{\varepsilon}{c \cdot \|p_i^{(kj)} - x_i^{(kj)}\|} \cdot \frac{\varepsilon}{c \cdot \|g^{(kj)} - x_i^{(kj)}\|} < 1 - \delta$$

所以:

$$\prod_{j=1}^{\infty} (1 - \frac{\varepsilon}{|\omega| \cdot \|x_i^{(kj)} - x_i^{(kj-1)}\|} \cdot \frac{\varepsilon}{c \cdot \|p_i^{(kj)} - x_i^{(kj)}\|} \cdot \frac{\varepsilon}{c \cdot \|g^{(kj)} - x_i^{(kj)}\|}) = 0$$

故, 若  $k \rightarrow \infty$  时, 有  $n$  次迭代使得  $g^* \in R$ , 则:

$$p(\zeta^{(k+1)} \in \Gamma) = 1 - p(\zeta^{(1)} \notin \Gamma) \cdot \prod_{l=1}^k \prod_{i=1}^m (1 - \frac{\varepsilon_l}{|\omega| \cdot \|x_i^{(l)} - x_i^{(l-1)}\|} \cdot \frac{\varepsilon_l}{c \cdot \|p_i^{(l)} - x_i^{(l)}\|} \cdot \frac{\varepsilon_l}{c \cdot \|g^{(l)} - x_i^{(l)}\|}) = 1 - p(\zeta^{(1)} \notin \Gamma) \prod_{j=1}^n \prod_{i=1}^m (1 - \frac{\varepsilon}{|\omega| \cdot \|x_i^{(kj)} - x_i^{(kj-1)}\|} \cdot \frac{\varepsilon}{c \cdot \|p_i^{(kj)} - x_i^{(kj)}\|} \cdot \frac{\varepsilon}{c \cdot \|g^{(kj)} - x_i^{(kj)}\|})$$

当  $n \rightarrow \infty$  时,  $\lim_{n \rightarrow \infty} p(\zeta^{(k+1)} \in \Gamma) = 1$ , 此时算法收敛到全局最优. 但是实际 PSO 算法无法保证  $n \rightarrow \infty$ , 因此  $0 < p(\zeta^{(k+1)} \in \Gamma) < 1$ , 即标准 PSO 算法以一定的概率搜索到全局最优.

## 4 结论

与文献 [20] 相比, 本文基于 PSO 的差分模型定义粒子和群体状态序列, 根据 PSO 加速度因子的随机特性计算得到粒子转移概率, 而非单纯的将粒子可达域看作一个超球体: 然后, 利用全概率公式计算得到了群体状态转移到非最优状态的概率, 从而得到群体转移到最优状态的概率, 而非只是计算群体状态的转移概率. 依据文中得到的概率, 分析了惯性权重, 加速度因子对算法早熟收敛和发散等问题的影响, 得到标准 PSO 算法中参数的选取无法使得全局最优位置总位于种群的可达域  $R$  内, 因此标准 PSO 算法既不能保证搜索到全局最优, 也不能保证搜索不到全局最优, 而是以一定概率搜索到全局最优. 而较好的参数选择会使最优位置位于可行区域内的概率增大, 并且使  $R$  域的范围适中, 从而增大算法搜索到全局最优位置的概率.

## References

- 1 Kennedy J, Eberhart R C. Particle swarm optimization. In: Proceedings of IEEE International Conference on Neural Networks. Perth, WA: IEEE, 1995. 1942–1948
- 2 Shi Y H, Eberhart R C. Parameter selection in particle swarm optimization. In: Proceedings of the 7th International Conference on Evolutionary Programming. New York, USA, 1998. 591–600
- 3 Trelea I C. The particle swarm optimization algorithm: convergence analysis and parameter selection. *Information Processing Letters*, 2003, **85**(6): 317–325
- 4 Clerc M, Kennedy J. The particle swarm: explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation*, 2002, **6**(1): 58–73
- 5 Zeng Jian-Chao, Cui Zhi-Hua. A new unified model of particle swarm optimization and its theoretical analysis. *Journal of Computer Research and Development*, 2006, **43**(1): 96–100  
(曾建潮, 崔志华. 微粒群算法的统一模型及分析. 计算机研究与发展, 2006, **43**(1): 96–100)
- 6 Zeng Zhi-Hua, Cui Jian-Chao. Analysis and improvement about particle swarm optimization based on linear control theory. *Journal of Chinese Computer Systems*, 2006, **27**(5): 849–853  
(崔志华, 曾建潮. 基于控制理论的微粒群算法的分析与改进. 小型微型计算机系统, 2006, **27**(5): 849–853)
- 7 Kadiramanathan V, Selvarajah K, Fleming P J. Stability analysis of the particle dynamics in particle swarm optimizer. *IEEE Transactions on Evolutionary Computation*, 2006, **10**(3): 245–255
- 8 Fernández-Martínez J L, García-Gonzalo E, Saraswathi S, Jernigan R, Kloczkowski A. Particle swarm optimization: a powerful family of stochastic optimizers: analysis, design and application to inverse modelling. In: Proceedings of the 2nd International Conference on Advances in Swarm Intelligence. Berlin, Heidelberg: Springer-Verlag, 2011. 1–8
- 9 Pan Feng, Chen Jie, Gan Ming-Gang, Cai Tao, Tu Xu-Yan. Model analysis of particle swarm optimizer. *Acta Automatica Sinica*, 2006, **32**(3): 368–377  
(潘峰, 陈杰, 甘明刚, 蔡涛, 涂序彦. 粒子群优化算法模型分析. 自动化学报, 2006, **32**(3): 368–377)
- 10 Clerc M. Stagnation analysis in particle swarm optimization or what happens when nothing happens [Online], available: <http://clerc.maurice.free.fr/pso/stagnation>, May 1, 2006
- 11 Jiang M, Luo Y P, Yang S Y. Stagnation analysis in particle swarm optimization. In: Proceedings of Swarm Intelligence Symposium. Honolulu, HI: IEEE, 2007. 92–99
- 12 Poli R, Broomhead D. Exact analysis of the sampling distribution for the canonical particle swarm optimiser and its convergence during stagnation. In: Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation. London, England: ACM, 2007. 134–141
- 13 Pan F, Chen J, Cai T, Gan M G, Wang G H. Stability, convergence of harmonious particle swarm optimizer and its application. *Journal of Beijing Institute of Technology*, 2009, **17**(1): 35–40
- 14 Pan Feng, Chen Jie, Xin Bin, Zhang Juan. Several characteristics analysis of particle swarm optimizer. *Acta Automatica Sinica*, 2009, **35**(7): 1010–1016  
(潘峰, 陈杰, 辛斌, 张娟. 粒子群优化方法若干特性分析. 自动化学报, 2009, **35**(7): 1010–1016)
- 15 van den Bergh F. An Analysis of Particle Swarm Optimizers [Ph. D. dissertation], University of Pretoria, South Africa, 2001
- 16 van den Bergh F, Engelbrecht A P. A study of particle swarm optimization particle trajectories. *Information Sciences*, 2006, **176**(8): 937–971
- 17 Jiang M, Luo Y P, Yang S Y. Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm. *Information Processing Letters*, 2007, **102**(1): 8–16
- 18 Zhang H B, Wang H B, Hu Z J. Analysis of particle swarm optimization algorithm global convergence method. *Computer Engineering and Applications*, 2011, **47**(34): 61–63
- 19 Nakama T. Markov chain analysis of genetic algorithms applied to fitness functions perturbed concurrently by additive and multiplicative noise. *Computational Optimization and Applications*, 2012, **51**(2): 601–622
- 20 Zhang Y A, Ma Q L, Sakamoto M, Furutani H. Effect of mutation to distribution of optimum solution in genetic algorithm. *Natural Computing*, 2010, **2**(3): 380–387
- 21 Sahin C S, Gundry S, Uyar M Ü. Markov chain analysis of self-organizing mobile nodes self-organizing mobile nodes. *Journal of Intelligent and Robotic Systems*, 2012, **67**(12): 133–153
- 22 Jin Xin-Lei, Ma Long-Hua, Wu Tie-Jun, Qian Ji-Xin. Convergence analysis of the particle swarm optimization based on stochastic processes. *Acta Automatica Sinica*, 2007, **33**(12): 1263–1268  
(金欣磊, 马龙华, 吴铁军, 钱积新. 基于随机过程的 PSO 收敛性分析. 自动化学报, 2007, **33**(12): 1263–1268)
- 23 Yuan Dai-Lin, Chen Qiu. Particle swarm optimization algorithm based on Markov model and its stochastic process analysis. *Computer Engineering and Applications*, 2009,



45(31): 49–52

(袁代林, 陈虬. 马氏模型 PSO 及其随机过程分析. 计算机工程与应用, 2009, 45(31): 49–52)

- 24 Li Ning. Analysis and Application of Particle Swarm Optimization [Ph. D. dissertation], Huazhong University of Science and Technology, China, 2006

(李宁. 粒子群优化算法的理论分析与应用研究 [博士学位论文], 华中科技大学, 中国, 2006)

- 25 Ren Zi-Hui, Wang Jian, Gao Yue-Lin. The global convergence analysis of particle swarm optimization algorithm based on Markov chain. *Control Theory and Applications*, 2011, 28(4): 462–466

(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. 控制理论与应用, 2011, 28(4): 462–466)

- 26 Cai Zhao-Quan, Huang Han, Zheng Zong-Hui, Luo Wei. Convergence improvement of particle swarm optimization based on the expanding attaining-state set. *Journal of Huazhong University of Science and Technology (Nature Science Edition)*, 2009, 37(6): 44–47

(蔡昭权, 黄翰, 郑宗晖, 罗伟. 基于可达状态集扩张的粒子群算法收敛性改进. 华中科技大学学报(自然科学版), 2009, 37(6): 44–47)

- 27 Poli R, Langdon W B. Markov chain models of bare-bones particle swarm optimizers. In: Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation. London, England: ACM, 2007. 142–149



**潘 峰** 北京理工大学自动化学院副教授. 2005 年获北京理工大学自动化学院模式识别与智能系统专业博士学位. 主要研究方向为群智能理论和计算智能. 本文通信作者.

E-mail: andropan@gmail.com

(**PAN Feng** Associate professor at the School of Automation, Beijing Institute of Technology. He received his Ph.D. degree in

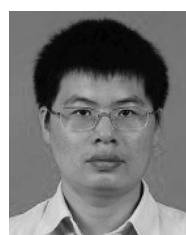
pattern recognition and intellectual system in 2005. His research interest covers computational intelligence, swarm intelligence, and unconventional computation. Corresponding author of this paper.)



**周 倩** 北京理工大学自动化学院硕士研究生. 2011 年获北京理工大学自动化学院学士学位. 主要研究方向为计算智能与群智能算法.

E-mail: szhouqian@gmail.com

(**ZHOU Qian** Master student at the School of Automation, Beijing Institute of Technology. She received her bachelor degree from Beijing Institute of Technology in 2011. Her research interest covers computational intelligence and swarm intelligence.)



**李位星** 北京理工大学自动化学院讲师. 2004 年获北京理工大学自动化学院模式识别与智能系统专业硕士学位. 主要研究方向为智能视频监控, 运动目标跟踪, 计算智能.

E-mail: liweixing@bit.edu.cn

(**LI Wei-Xing** Lecturer at the School of Automation, Beijing Institute of Technology. He received his master degree in pattern recognition and intellectual system in 2004. His research interest covers intelligent video surveillance, moving object tracking, and computational intelligence.)



**高 琪** 北京理工大学自动化学院讲师. 主要研究方向为模式识别应用, 复杂网络和智能监控技术.

E-mail: gaoqi@bit.edu.cn

(**GAO Qi** Lecturer at the School of Automation, Beijing Institute of Technology. His research interest covers pattern recognition and its applications, complex network, and intelligent monitoring technology.)