差分进化算法马尔可夫链模型及收敛性分析

孙成富,赵建洋,陈剑洪
(淮阴工学院 计算机工程学院,江苏 淮安 223003)

摘　要:差分进化算法是一种基于种群差异的优化算法,主要应用于解决连续空间的优化问题。目前,研究人员主要在算法的改进和应用方面研究差分进化算法,很少从理论角度对其进行研究。为了分析差分进化算法的收敛性,定义优化个体、种群的状态转移,并提出种群的最优状态集合。根据差分进化算法的操作算子计算出个体的状态迁移概率,并证明种群状态序列是有限齐次马尔可夫链,进而建立差分进化算法的马尔可夫链模型;最后,证明差分进化算法无法保证全局收敛。理论研究结果表明,适当保证种群的多样性能够提高差分进化算法的性能。

关键词:差分进化;马尔可夫链;收敛性分析;全局收敛;局部收敛
中图分类号:TP311　　**文献标识码**:A　　**文章编号**:1673-629X(2013)08-0062-04
doi:10.3969/j.issn.1673-629X.2013.08.016

Analysis of Differential Evolution's Markov Chain Model and Convergence

SUN Cheng-fu,ZHAO Jian-yang,CHEN Jian-hong
(College of Computer Engineering,Huaiyin Institute of Technology,Huaian 223003,China)

Abstract:As a modern optimization algorithm,differential evolution algorithm which is based on the individual differential reconstruction idea is designed for the global continuous optimization problem. Up to now,the improvement and application of the algorithm are mainly focused by researchers but theoretical analysis of the algorithm is seldom taken into account. In order to analyze the convergence of the algorithm,the concepts of state transition for individual and population are defined and the optimal state set of population is proposed. The individual state transition probability is computed according to the operators of differential evolution algorithm. The state sequence of population has been proved to be Finite Nonhomogeneous Markov chain and the Markov chain model of differential evolution is proposed. At last,the theory analysis of the differential evolution demonstrates that it is not able to guarantee the global convergence. The result of the theory research shows that keeping the population diversity will improve the performance of the algorithm.

Key words:differential evolution;Markov chain;convergence analysis;global convergence;local convergence

0　引　言

差分进化算法(Differential Evolution,简称DE)是一种基于种群差异的新兴进化计算技术,最初由Rainer Storn和Kenneth Price于1995年提出[1]。基于种群的优化算法主要包括遗传算法、粒子群算法、差分进化算法等。目前,群智能优化技术已经广泛地应用于很多领域,并取得很好的效果。

差分进化算法通过对当前种群的变异、交叉和选择等操作产生下一代种群,并逐步使种群靠近或包含最优解。由于差分进化算法具有收敛速度快、鲁棒性能优和简单易用等优点,它已经在许多领域得到了广泛的应用,譬如聚类分析[2]、电力系统经济调度[3]、水火混合电力系统经济调度[4]、流水车间优化调度[5]、订单可分的协作计划优化[6]等。标准的差分进化算法是一种基于实数编码的具有保优思想的遗传算法,因此在低维空间上具有收敛速度快、解的质量高、稳定性好等特点,但在高维、多峰值的复杂函数优化中也会陷入局部最优,出现“早熟”现象。

为了进一步提高差分进化算法的性能,许多学者对其进行了深入的改进研究,如:对差分进化算法的变异操作算子进行改进,以保证优化过程中种群的多样性[7];利用最优个体收缩搜索空间,以进一步提高算法的搜索效率[8];将差分进化算法与其他优化算法相结

收稿日期:2012-10-14　　**修回日期**:2013-01-18　　**网络出版时间**:2013-04-08
基金项目:江苏省科技支撑计划(BE2012112);淮安市科技支撑计划(工业)项目(HAG2011044,HAG2011045)
作者简介:孙成富(1979-),男,山东苍山人,讲师,博士,研究方向为智能计算、优化算法理论。
网络出版地址:http://www.cnki.net/kcms/detail/61.1450.TP.20130408.1600.027.html

合,发挥各自优势,设计出高效的优化算法[9]。尽管针对差分进化算法的改进获得了深入广泛的研究,但是其研究成果相当分散,缺乏系统性和理论依据。主要原因就是差分进化算法的理论研究还没有重大突破。

由于差分进化算法是一种随机优化算法,因此对其收敛或不能保证收敛进行证明都存在一定难度。文中对标准差分进化算法-DE/rand/1 进行收敛性分析与研究。最终,为差分进化算法的改进提供指导思想。

1　标准差分进化算法

差分进化算法就是通过变异、交叉、选择等操作算子实现从一代种群演化到新一代种群,最终解决优化问题。为了深入分析标准差分进化算法的收敛性,下面将介绍其变异、交叉、选择这三个操作算子。

(1)变异操作。对种群中的每个目标个体 x_i^G,在当前种群中随机选择三个互不相同的个体 $x_{r_1}^G, x_{r_2}^G, x_{r_3}^G(r_1 \neq r_2 \neq r_3 \neq i)$,并按下面的公式产生变异个体 v_i^G:

$$v_i^G = x_{r_1}^G + F * (x_{r'_2}^G - x_{r_3}^G) \tag{1}$$

其中,F 为缩放因子,它控制矢量差 $x_{r'_2}^G - x_{r_3}^G$ 的幅值。根据变异个体不同的生成方法,从而形成不同的差分进化方案。

(2) 交叉操作。为了增加种群的多样性,差分进化算法引入交叉操作。将变异操作产生的变异个体与目标个体按照下面的公式进行交叉,以生成新的试验个体 u_i^G。试验个体 u_i^G 的第 j 维分量的生成规则如下:

$$u_{i,j}^G = \begin{cases} v_{i,j}^G & \text{if}(\eta_j \leqslant \text{CR})\text{or}(j = q) \\ x_{i,j}^G & \text{otherwise} \end{cases} \tag{2}$$

其中 $\eta_j \in [0,1]$,表示 0 与 1 之间的随机数,q 表示 1 到优化问题的维数之间的一个随机整数;CR 是交叉概率,控制种群的多样性;$x_{i,j}^G, v_{i,j}^G$ 和 $u_{i,j}^G$ 分别是第 i 目标个体,变异个体和试验个体的第 j 维分量。

(3) 选择操作。对试验个体 $u_{i,j}^G$ 和目标个体 $x_{i,j}^G$ 的目标函数值进行比较,对于最小化问题,则选择目标函数值小的个体进入下一代种群。

$$x_i^{G+1} = \begin{cases} u_i^G & \text{if } f(u_i^G) \leqslant f(x_i^G) \\ x_i^G & \text{otherwise} \end{cases} \tag{3}$$

2　差分进化算法的马尔可夫链模型及其收敛性分析

下面将对最常用的差分进化方案进行数学建模和收敛性分析。除非明确说明都将对最小化问题 f 进行研究,并假定种群中的个体互不相同。

2.1　相关定义与定理

定义 1 (个体状态及种群状态):个体 I 的状态由其位置向量 x 决定;而种群的状态由其构成个体的状态来决定,即若 POP $= \{I_1, I_2, \cdots, I_{N_p}\}$,则种群 POP 的状态 $S = \{x_1, x_2, \cdots, x_{N_p}\}$,其中 $x_1, x_2, \cdots, x_{N_p}$ 分别表示个体 $I_1, I_2, \cdots, I_{N_p}$ 的状态。

定义 2 (种群状态空间):种群所有可能的状态构成的集合为种群状态空间,SP $= \{S_i = (x_{i1}, x_{i2}, \cdots, x_{iN_p}) \mid i = 1,2,\cdots\}$。

定义 3 (个体状态转移):在差分进化算法迭代中,个体 I 的状态由 x_1 一步转移到 x_2,记作 $x_1 \to x_2$,其中 x_1, x_2 为个体状态空间中的任意两个状态。

定义 4 (种群的状态转移):在差分进化算法中,种群的状态由 $S_i = (x_{i1}, x_{i2}, \cdots, x_{iN_p})$ 一步跳转到状态 $S_j = (x_{j1}, x_{j2}, \cdots, x_{jN_p})$,记作 $S_i \to S_j$。

定义 5(最优种群状态集合 Ψ):假定优化问题的全局最优解为 g^*,则最优种群状态集合可定义为:$\Psi = \{S = (x_1, x_2, \cdots, x_{N_p}) \mid \exists x_i$ 使得 $f(x_i) = f(g^*), 1 \leqslant i \leqslant N_p\}$。

在最优种群状态集合 Ψ 中的每个种群状态中至少有一个个体的适应值达到全局最优解 g^* 的适应值。

定理 1:在差分进化算法中,个体 I 的状态由 x_1 一步转移到 x_2 的概率为 $P(x_1 \to x_2) = P_{x_1 \to v_1} * P_{(v_1, x_1) \to x_2} * P_{\text{select}}$。

其中,$P_{x_1 \to v_1} = \alpha, (1 \geqslant a \geqslant \frac{1}{(N_p - 1) * (N_p - 2) * (N_p - 3)})$ (N_p 表示种群的大小)。

$$P_{(v_1, x_1) \to x_2} = \prod_{i=1}^{\text{DIM}} K(i)$$

$$(K(i) = \begin{cases} x_{1i} == v_{1i} \text{ 或者 } i == q \\ (1 - \text{CR}) * \text{Sel}(\text{rand}(i) > \text{CR}) + \\ \text{CR} * \text{Sel}(\text{rand}(i) \leqslant \text{CR}) \text{otherwise} \end{cases}$$

$\text{Sel}(x) = \begin{cases} 1 & x \text{ 为真} \\ 0 & x \text{ 为假} \end{cases}$;rand($i$) 表示针对第 i 维而生成的(0,1B) 之间的随机数;DIM 表示优化问题的维数;q 为整数并且 $q \in [1, \text{DIM}]$)。

$$P_{\text{select}} = \begin{cases} 1 & f(x_1) \geqslant f(x_2) \\ 0 & f(x_1) < f(x_2) \end{cases}$$

证明:假定差分进化算法的变异、交叉和选择等操作均相互独立,并且在变异操作中三个随机个体的选择过程也是相互独立的随机过程。

根据差分进化算法中的变异操作,即 $v_1 = x_{r_1} + F * (x_{r_2} - x_{r_3})$ (F 为固定常数),由目标个体 x_1 到变异个体 v_1 的概率为:$P_{x_1 \to v_1} = \alpha, (1 \geqslant a \geqslant \frac{1}{(N_p - 1) * (N_p - 2) * (N_p - 3)})$ (N_p 表示种群的大

小)。

根据差分进化过程中交叉操作的定义,即针对优化问题的每一维进行如下操作:

$$x_2(i)=\begin{cases}v_1(i) & \text{rand}(i)\leqslant \text{CR 或者 } i==q\\ x_1(i) & \text{otherwise}\end{cases}$$

(CR 为交叉概率,rand(i) 表示针对第 i 维而生成的(0,1)之间的随机数,q 表示 1 到优化问题的维数之间的一个随机整数)。事件 rand(i) ⩽ CR 发生的概率为 $P(\text{rand}(i)\leqslant \text{CR})=1*\text{CR}=\text{CR}$。由于 rand(i) ⩽ CR 与 rand(i) > CR 是相互排斥的事件,因此 $P(\text{rand}(i)>\text{CR})=1-\text{CR}$。可以得到由 $(v_1,x_1)\rightarrow x_2$ 的概率为:

$$P_{(v_1,x_1)\rightarrow x_2}=\prod_{i=1}^{\text{DIM}}K(i)$$

$$(K(i)=\begin{cases}1 & x_{1i}==v_{1i}\text{ 或者 } i==q\\ (1-\text{CR})*\text{Sel}(\text{rand}(i)>\text{CR})+ & \\ \text{CR}*\text{Sel}(\text{rand}(i)\leqslant \text{CR}) & \text{otherwise}\end{cases}$$

$\text{Sel}(x)=\begin{cases}1 & x\text{ 为真}\\ 0 & x\text{ 为假}\end{cases}$;rand(i) 表示针对第 i 维而生成的(0,1)之间的随机数;DIM 表示优化问题的维数;q 为整数并且 $q\in[1,\text{DIM}]$);变异和交叉操作之后,选择操作被执行。状态 x_2 被确定为 I 的状态 x_1 的下一个状态,是由 x_2 的适应值决定的。如果 x_2 的适应值优于 x_1 的适应值,则 x_2 就被定为个体 I 的状态 x_1 的下一个状态;否则个体 I 保持状态 x_1 不变。因此,x_2 最终成为个体 I 的状态 x_1 的下一个状态的概率为:

$$P_{\text{select}}=\begin{cases}1 & f(x_1)\geqslant f(x_2)\\ 0 & f(x_1)<f(x_2)\end{cases}$$

综上所述,在差分进化算法中,个体 I 的状态由 x_1 一步转移到 x_2 的概率为 $P(x_1\rightarrow x_2)=P_{x_1\rightarrow v_1}*P_{(v_1,x_1)\rightarrow x_2}*P_{\text{select}}$。至此定理证毕。

定理 2:在差分进化算法的迭代过程中,种群状态由 S_i 一步跳转到状态 S_j 的概率为 $P(S_i\rightarrow S_j)=\prod_{k=1}^{N_p}P(x_{ik}\rightarrow x_{jk})$($N_p$ 表示种群所包含的个体数)。

证明:$S_i\rightarrow S_j$ 表示种群状态 S_i 中的所有个体的状态同时转移到种群状态 S_j 中相应个体的状态,即 $x_{ik}\rightarrow x_{jk}$,$k=1,2,\cdots,N_p$,并假定所有个体的状态转移是相互独立的事件,因此 $P(S_i\rightarrow S_j)=\prod_{k=1}^{N_p}P(x_{ik}\rightarrow x_{jk})$。至此定理证毕。

2.2 差分进化算法的 Markov 链模型

在差分进化算法的迭代过程中,单个个体的状态转移不但依赖于个体当前的状态,而且还取决于该个体所在的整个群体的状态,因此单个个体的状态转移过程不是马尔可夫过程。但是整个种群状态转移的过程却只与种群的当前状态有关,因此种群状态的转移是马尔可夫过程。下面将给出详细证明。

定理 3:差分进化算法在计算机上实现时,种群状态序列 $\{S(t)\mid t\geqslant 0\}$ 是有限齐次马尔可夫链。

证明:(1) 在计算机上优化任何问题,都以一定的精度表示优化问题的变量。为了不失一般,将假定优化问题只含有一个变量,其取值范围是 $[x_{\min},x_{\max}]$。若设精度为 ε,那么 $[x_{\min},x_{\max}]$ 将可以被有限的离散数来表示。因此当在计算机上实现差分进化算法时,单个个体的状态空间是有限的并且是离散的。一个种群的状态 $S=(x_1,x_2,\cdots,x_{N_p})$ 由 N_p 个体的状态组成,N_p 为有限正整数,因此种群状态空间 SP 也是有限的。

(2) 由定理 2 可知,种群状态序列 $\{S(t)\mid t\geqslant 0\}$ 中的任意两个种群状态 $S(t-1)=(x_{1,(t-1)},x_{2,(t-1)},\cdots,x_{N_p,(t-1)})$,$S(t)=(x_{1,t},x_{2,t},\cdots,x_{N_p,t})$,它们之间的转移概率可表示为:

$$P(S(t-1)\rightarrow S(t))=\prod_{i=1}^{N_p}P(x_{i,(t-1)}\rightarrow x_{i,t})$$

由定理 1 可知:

$$P(x_{i,(t-1)}\rightarrow x_{i,t})=P_{x_{i,(t-1)}\rightarrow v_1}*P_{(v_1,x_{i,(t-1)})\rightarrow x_{i,t}}*P_{\text{select}}$$

种群中个体状态的转移概率 $P(x_{i,(t-1)}\rightarrow x_{i,t})$ 是由 $P_{x_{i,(t-1)}\rightarrow v_1}$,$P_{(v_1,x_{i,(t-1)})\rightarrow x_{i,t}}$,$P_{\text{select}}$ 决定的。并且 $P_{x_{i,(t-1)}\rightarrow v_1}$,$P_{(v_1,x_{i,(t-1)})\rightarrow x_{i,t}}$,$P_{\text{select}}$ 仅仅与 $t-1$ 时刻的种群状态有关,所以 $P(x_{i,(t-1)}\rightarrow x_{i,t})$ 也仅仅与 $t-1$ 时刻的种群状态有关。

由以上分析可得到,种群状态序列 $\{S(t)\mid t\geqslant 0\}$ 中的任意两个种群状态 $S(t-1)=(x_{1,(t-1)},x_{2,(t-1)},\cdots,x_{N_p,(t-1)})$,$S(t)=(x_{1,t},x_{2,t},\cdots,x_{N_p,t})$ 之间的状态转移概率 $P(S(t-1)\rightarrow S(t))$ 也仅仅与 $t-1$ 时刻的种群状态有关。所以种群状态序列 $\{S(t)\mid t\geqslant 0\}$ 具有马尔可夫特性。又由于种群的状态空间 SP 既是离散的又是有限的,因此种群状态序列 $\{S(t)\mid t\geqslant 0\}$ 是有限马尔可夫链。

(3) 由定理 1 可知,$P_{x_{i,(t-1)}\rightarrow v_1}$,$P_{(v_1,x_{i,(t-1)})\rightarrow x_{i,t}}$,$P_{\text{select}}$ 与时刻 $t-1$ 无关,因此 $P(x_{i,(t-1)}\rightarrow x_{i,t})$ 也与时刻 $t-1$ 无关,从而可以得到 $P(S(t-1)\rightarrow S(t))$ 也与时刻 $t-1$ 无关,因此差分进化算法在计算机上实现时,种群状态序列 $\{S(t)\mid t\geqslant 0\}$ 是有限齐次马尔可夫链。至此定理证毕。

2.3 差分进化算法的收敛性分析

由于当 $\Psi=\text{SP}$ 没有任何优化问题的存在,因此下

面关于差分进化算法的收敛性分析是在 $\Psi \subset SP$ 的前提下进行的。

定理4 当 $\Psi \subset SP$ 时,即最优种群状态集合为种群状态空间的真子集,差分进化算法无法保证全局收敛。

证明:因为 $\Psi \subset SP$,因此在可行解空间内至少存在一点 x,使得 $f(x) > f(g^*)$。构造一个种群状态 $S_t = (x_{t1}, x_{t2}, \cdots, x_{tN_p})$,并且其所有的 N_p 个体的状态都为 x。并假定 $SP_t = \{S_t\}$ 为状态空间的一个子集。因为 $\exists S_i \notin SP_t$ 和 $\forall S_t \in SP_t$,则种群状态转移 $S_i \to S_t$ 的概率为 $P(S_i \to S_t) = \prod_{k=1}^{N_p} P(x_{ik} \to x_{tk}) > 0$。但对于 $\forall S_j \notin SP_t$ 和 $\forall S_t \in SP_t$,若要实现种群的状态转移 $S_t \to S_j$,则种群状态 S_t 中至少需要一个个体 I_i 发生如下状态转移 $x_{ti} \to \hat{x}_j$ 并且 $f(\hat{x}_j) \neq f(x_{ti})$。但是由差分进化算法的迭代公式和定理1与2可知,$P(x_{ti} \to x_{ti}) = 1$,并且可以得到 $P(S_t \to S_t) = 1$,因此 $P(S_t \to S_j) = 0$。由于 $f(x) > f(g^*)$,所以 SP_t 是完全不同于最优状态集合 Ψ,即 $SP_t \cap \Psi = \varnothing$。由定理 3 可知,种群状态序列 $\{S(t) \mid t \geqslant 0\}$ 为有限齐次马尔可夫链。又因为种群状态序列 $\{S(t) \mid t \geqslant 0\}$ 这一状态空间中存在真子集 SP_t,且 $SP_t \cap \Psi = \varnothing$。$SP_t$ 为这一有限齐次马尔可夫链的吸收态。系统一旦进入该吸收态,将无法跳出。综上所述,可知当 $\Psi \subset SP$ 时,差分进化算法无法保证全局收敛。至此定理证毕。

3 结束语

文中对差分进化算法作了深入的理论分析,这将为在具体工程实践中改进差分进化算法提拱指导思想。首先,对差分进化算法的基本概念作了严格的数学描述和定义;然后建立差分进化算法的 Markov 链模型,并证明差分进化算法优化过程中群体状态的转移过程是有限齐次 Markov 链;最后,分析差分进化算法的收敛性,并证明了差分进化算法无法保证全局收敛。虽然差分进化算法无法保证全局收敛,但是仿真实验和实际应用显示该优化算法具有很强的全局搜索能力。粒子群算法同样无法保证全局收敛,但是也被广泛地应用于解决各种各样的全局优化问题,并显示出搜索能力强、稳定性好等优势。

参考文献:

[1] Storn R, Price K. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces[J]. Global Optimization, 1997, 11(4): 341-359.

[2] Paterlinia S, Krinkb T. Differential evolution and particle swarm optimization in partitional clustering[J]. Computational Statistics & Data Analysis, 2006, 50(5): 1220-1247.

[3] Mandal K K, Chakraborty N. Short-term combined economic emission scheduling of hydrothermal power systems with cascaded reservoirs using differential evolution[J]. Energy Conversion and Management, 2009, 50(1): 97-104.

[4] Yuan X H, Wang L, Zhang Y C, et al. A hybrid differential evolution method for dynamic economic dispatch with valve-point effects[J]. Expert System with Application, 2009, 36(2): 4042-4048.

[5] Pan Q K, Wang L, Qian B. A novel differential evolution algorithm for bi-criteria no-wait flow shop scheduling problems[J]. Computer & Operations Research, 2009, 36(8): 2498-2511.

[6] 包　融,王伟业,顾汉杰,等. 订单可分的协作计划模型及其进化算法[J]. 计算机技术与发展, 2010, 20(10): 58-61.

[7] Fan H Y, Lampinen J. A trigonometric mutation operation to differential evolution[J]. Journal of global optimization, 2003, 27(1): 105-129.

[8] Lakshminarasimman L, Subramanian S. A modified hybrid differential evolution for short-term scheduling of hydrothermal power systems with cascaded reservoirs[J]. Energy Conversion and Management, 2009, 49(10): 2513-2521.

[9] Hu C, Yan X F. An immune self-adaptive differential evolution algorithm with application to estimate kinetic parameters for homogeneous mercury oxidation[J]. Chinese Journal of Chemical Engineering, 2009, 17(2): 232-240.

(上接第 61 页)

[3] 刘　刚,于力超. 搜索引擎中网络蜘蛛的设计与实现[J]. 电脑与信息技术, 2007, 15(4): 39-42.

[4] Roslak J. Active lighting systems for improved road safety[C]//Proc. of IEEE Intelligent Vehicles Symposium. [s. l.]: [s. n.], 2004: 682-685.

[5] 陈　勇,刘　勇. 中医药主题搜索网络机器人的设计与实现[J]. 计算机技术与发展, 2010, 20(5): 162-166.

[6] 尹西杰. 基于智能 Agent 的 Web 个性化信息检索系统[D]. 济南:山东大学, 2006.

[7] 何顺志,徐文芬,黄　敏,等. 贵州中药资源种类与分布的研究[J]. 世界科学技术:中医药现代化, 2005, 7(2): 95-102.

[8] 胡元军. 基于 Agent 的分布式专业信息采集系统[D]. 北京:北京化工大学, 2007.

[9] 王汝传,徐小龙,黄海平. 智能 AGENT 及其在信息网络中的应用[M]. 北京:北京邮电大学出版社, 2007.

[10] 刘　迁,贾惠波. 中文信息处理中自动分词技术的研究与展望[J]. 计算机工程与应用, 2006, 42(3): 175-177.

差分进化算法马尔可夫链模型及收敛性分析

作者： 孙成富， 赵建洋， 陈剑洪， SUN Cheng-fu， ZHAO Jian-yang， CHEN Jian-hong
作者单位： 淮阴工学院 计算机工程学院,江苏 淮安,223003

刊名： 计算机技术与发展 ISTIC
英文刊名： Computer Technology and Development
年，卷(期)： 2013(8)