

一个多目标优化演化算法的收敛性分析框架*

覃 俊^{1,2}, 康立山¹

(1. 武汉大学 软件工程国家重点实验室, 湖北 武汉 430072; 2. 中南民族大学 计算机科学学院, 湖北 武汉 430074)

摘 要: 由于演化算法求解多目标优化问题所得结果是一个优化解集——Pareto 最优集, 而现有的演化算法收敛性分析只适合针对单目标优化问题的单个最优解。利用有限马尔科夫链给出了演化算法求解多目标优化问题的收敛性分析框架, 并给出了一个分析实例。

关键词: 多目标优化; 演化算法; Pareto 前沿; 收敛性; 有限马尔科夫链

中图法分类号: TP301.6 文献标识码: A 文章编号: 1001-3695(2005)02-0068-03

A Convergence Analysis Framework for Multi-objective Optimization Evolutionary Algorithm

QIN Jun^{1,2}, KANG Li-shan¹

(1. State Key Laboratory of Software Engineering, Wuhan University, Wuhan Hubei 430072, China; 2. College of Computer Science, South Center University for Nationalities, Wuhan Hubei 430074, China)

Abstract: Based on Finite Markov Chain theory, a new convergence analysis framework for Multi-objective Optimization Evolutionary Algorithm (MOEA), which aims to converge to Pareto optimal set rather than single optimal point, is presented. And a case study is given too.

Key words: Multi-objective Optimization; Evolutionary Algorithm; Pareto Dominance; Convergence; Finite Markov Chain

1 引言

多目标优化作为运筹学的一个重要分支, 一直引起学术界和工程界的重视。但由于经典方法的局限性, 它在过去一段时间内发展缓慢, 近几年随着计算机技术的高速发展, 国际学术界兴起了一种求解多目标优化问题的新方法——多目标优化演化算法(Multi-objective Optimization Evolutionary Algorithms, MOEA)。该类算法的特点是不需要人为定义各目标的权重, 而是由算法找出所有的非劣解(Pareto 前沿), 一般来说, 一个多目标优化问题的非劣解是一个集合, 再由决策者进行权衡选择。

由于该方案比传统的方法有明显的优势, 已吸引了很多学者进行研究, 并取得了不少研究成果, 大多数成果集中在设计高效的演化算法^[1~3], 如为了保持群体多样性而采用的适应值分享策略、精英策略等, 著名的算法有 Schaffer 的 VEGA, Fonseca 的 FFGA, Zitzler 的 SPEA 等。

尽管这些算法在数值实验中得到了较好的结果, 但大多数都没有理论分析结果, 特别是算法的收敛性分析、收敛速度等很少涉及^[4]。MOEA 的收敛性分析对于设计一个高效的算法提供了理论保障。现有的演化算法收敛性分析框架只适合收敛到单个解的单目标优化演化算法, 而 MOEA 收敛到一个集合(Pareto 最优集)。本文试图利用有限马尔科夫链理论, 给出

一个新的收敛于一个解集(Pareto 最优集)的定义。

2 相关定义和定理

2.1 MOEA 基本定义

定义 1 Pareto 占优(Pareto Dominance): 向量 $u = (u, u_1, u_2, \dots, u_k)$ 比向量 $v = (v_1, v_2, \dots, v_k)$ 占优($u < v$) 是指, $\forall i \in \{1, 2, \dots, k\} u_i \leq v_i$, 并且 $\exists i \in \{1, 2, \dots, k\}$, 使得 $u_i < v_i$ 。

定义 2 Pareto 最优解(Pareto Optimality): 一个候选解 $X \in \Omega$ 是 Pareto 最优解是指, $\neg \exists X' \in \Omega$ 使得 $F(X') < F(X)$ 。

定义 3 Pareto 最优集(Pareto Optimal Set):

$P = \{X \in \Omega | \neg \exists X' \in \Omega: F(X') < F(X)\}$ 。

定义 4 Pareto 前沿(Pareto Front): $PF = \{F(X) | X \in P\}$ 。

从定义可以看出, Pareto 前沿是 Pareto 最优集在目标函数空间中的像。另外, 还要引入概念非劣解(Nondominated), 非劣解是指 Pareto 最优解通过函数 F 映像到目标函数空间中的向量。

2.2 预备定义

定义 5 设 Γ 是严格偏序集($S, <$)的子集, $a \in \Gamma$ 叫做 Γ 的极小元, 是指不存在 $x \in \Gamma$, 使 $x < a$; 定义集合 $M(\Gamma, <) = \{a \in \Gamma | a \text{ 是 } \Gamma \text{ 的极小元}\}$, 称 $M(\Gamma, <)$ 为关于 Γ 的极小元集。

定义 6 称极小元集 $M(\Gamma, <)$ 为完全的, 如果对于每一个 $x \in \Gamma$, 都至少有一个 $x^* \in M(\Gamma, <)$, 满足 $x^* < x$ 。

定义 7 设 $f: \Omega \rightarrow \Gamma$ 是一个从集合 Ω 到偏序集($\Gamma, <$)的一个映射, 那么对于 $A \in \Omega$, 定义集合 $M_f(A, <) = \{a \in A | f(a) \in M(f(A), <)\}$ 包含了 A 中某些元素, 这些元素对应的

收稿日期: 2004-02-07; 修返日期: 2004-05-31

基金项目: 国家自然科学基金资助项目(69635030, 60073043, 70071042)

像在像空间 $f(A) = \{ f(a) \mid a \in A \}$ 中是极小元。

2.3 有限马尔科夫链基本定义和定理

定义 8 (有限马尔科夫链) 如果 S 是一个有限集并且 $\{X_t: t \in N_0\}$ 是一个从 S 中取值的随机序列, 对于所有的 $t \geq 0$ 和所有的 $(i, j) \in S \times S$ 有以下性质:

$P\{X_{t+1} = j \mid X_t = i, X_{t-1} = i_{t-1}, \dots, X_0 = i_0\} = P\{X_{t+1} = j \mid X_t = i\} = P_{i,j}$, 则称序列 $\{X_t: t \in N_0\}$ 是一个在状态空间 S 上的齐次有限马尔科夫链。

由于 S 是有限的, 转移概率可用一个状态转移矩阵 $P = (P_{i,j})_{i,j \in S}$ 表示。向量 $\pi(t)$ 表示在第 $t \geq 0$ 步马尔科夫链的分布情况, 其中 $\pi_i(t) = P\{X_t = i\}$ 。由于 $\pi(t) = \pi(t-1)P = \pi(0)P^t$

对于所有 $t \geq 1$, 一个齐次有限马尔科夫链可以用它的初始分布 $\pi(0)$ 和它的转移矩阵 P 描述。通过 k 步从状态 i 转换到状态 j 的概率是:

$P_{i,j}(k) = P\{X_k = j \mid X_0 = i\} = e_i P^k e_j$, 其中 e_i 是第 i 项为 1 的单位向量, 因此有:

$$\pi_j(t) = \sum_{i \in S} \pi_i(0) \cdot P_{ij}(t)$$

一个 $n \times m$ 矩阵 A , 如果对于所有的 $i = 1, \dots, n$ 和所有的 $j = 1, \dots, m$, 有 $a_{ij} \geq 0$, 则称 A 是非负的, 若有 $a_{ij} > 0$, 则称 A 是正的。如果一个非负矩阵的每一行之和等于 1, 则称它是随机的。所以有限齐次马尔科夫链的转移矩阵是随机矩阵。称随机矩阵 P 是不可约的, 如果有:

- $\forall i, j \in S: \exists k \in N: P_{ij}(k) > 0$ 。
- $\forall$ 称 P 是原始的或规则的, 如果有:
 $\exists k \in N: \forall i, j \in S: P_{ij}(k) > 0$ 。

因此, 所有的正矩阵 P 都是规则的, 所有的规则矩阵 P 都是不可约的。

定义 9 一个随机矩阵被称为正对角的, 如果对角线上每个元素都是正数, 被称为列可容的, 如果每一列至少存在一个元素是正数。因此, 所有的正对角矩阵都是列可容的。

定理 1 一个具有有限状态空间与不可约转移矩阵的齐次马尔科夫链, 以概率 1 无限次地访问每一状态而与它的初始分布无关(证明略)。

定理 2^[5] 令 I, D, C, P, A 为随机矩阵, 其中 I 是不可约的, D 是正对角矩阵, C 是列可容的, P 是正的, A 是任意的, 那么

- (1) AP 和 PC 是正的。
 - (2) ID 和 DI 是不可约的。
- 证明略。

2.4 MOEA 收敛性定义

定义 10 令 A_t 是某演化算法在第 t 代的群体, $F_t = f(A_t)$ 是其相应的适应值(目标空间)集合, F^* 是所求问题目标空间集合的极小元集。

如果 $\lim_{t \rightarrow \infty} d(F_t, F^*) \rightarrow 0$ 以概率 1 成立, 则称该演化算法以概率 1 收敛于极小元整集。

如果 $\lim_{t \rightarrow \infty} \delta_b^*(F_t) \rightarrow 0$ 以概率 1 成立, 则称该演化算法以概率 1 收敛于极小元集。

其中, $d(A, B) = |A \cup B| - |A \cap B|$, $\delta_b(A) = |A| - |A \cap B|$

说明: 对于第一种情况, 很显然, 种群数要增长到 $|F^*|$ 。

在单目标问题中, $|F^*| = 1$; 但是在多目标问题中, $|F^*|$ 可能是有限的, 也可能是无限的。很显然对于无限的 $|F^*|$, 有限种群的 MOEA 不可能以概率 1 收敛于极小元整集。对于种群规模 $N < |F^*|$ 的算法, 很显然也不能以概率 1 收敛于极小元整集。

3 MOEA 收敛性分析框架

假定这里所指的 MOEA 算法都是求解多目标优化问题的 Pareto 前沿。利用有限马尔科夫链进行 MOEA 收敛性分析的基本思想是:

- (1) 证明算法在执行过程中, 它的种群序列是一个具有有限状态空间和不可约转移矩阵的齐次马尔科夫链。这样, 定理 1 可以保证种群可以概率 1 遍历每一个状态, 当然也包括了包含 Pareto 最优解的状态。
- (2) 要保证, 一旦获得了 Pareto 最优解, 算法应该具有某种机制把最优解保存下来, 即找到的最优解一旦进入某存储机构中就不再被淘汰, 如精英策略。
- (3) 还要证明, 作为结果输出的集合中的所有非最优解能以概率 1 被淘汰。

当然这只是基本思路, 在实际算法中要根据算法的具体特点及所用的算子进行针对性的讨论。为简单起见, 本文不考虑某些问题的 Pareto 最优解的数量无穷或相当大以至于超出存储空间时的情况。

4 算法实例分析

4.1 一个求解多目标优化的算法 MOEA

该 MOEA 算法框架最初由文献[6]提出, 利用本文的工作对该算法收敛性分析得知, 该算法由于变换算子缺陷, 使得该算法在本文定义的收敛性意义下不收敛。这也可以从该算法在求解高维问题 and 多目标优化等这些目标空间巨大的问题时效果不好的事实推测出来。针对该算法的变换算子的不足, 我们加入了一个变异算子和显式精英策略。可以证明改进后的算法在本文的定义下收敛, 我们的很多实验结果也说明了这一点^[7,8]。

算法总框架描述如下:

- (1) 随机产生 N 个个体(群体规模 $= N$, 个体在问题空间实数编码) 形成初始化种群 A_t ;
- (2) 按 better 函数要求把 N 个个体排序, 产生初始 Pareto 最优集 S_{best} ;

While 不满足终止条件{ 从 N 个个体中随机取 M 个点, 根据 M 个点进行多父体杂交产生后代 V^* , 随机挑选某个体进行变异, 对此 $N+1$ 个个体按 better 函数重新排序, 丢掉最差个体, 形成当前种群 A_t ; 更新 S_{best} }。

算法结束时, S_{best} 中的个体即为所求问题的 Pareto 最优解集。

算法有关说明如下:

由于该算法为了实现排序, 定义个体间的 better 关系为

$$\text{better}(X, Y) = \begin{cases} \text{True} & \text{if } X \prec Y \\ \text{False} & \text{if } X \succ Y \\ \text{Niche}(X, Y) & \text{否则} \end{cases}$$

Niche 机制^[2] 用来使群体分布均匀, 是一个布尔函数。具

体定义由于与收敛性无关,这里从略。

多父体杂交算子具体设计:从当前群体中任意选取 M 个个体,通过它们的非线性组合形成新个体 V^* :

$$V^* = \sum_{i=1}^M a_i x_i \quad \text{其中 } a_i \text{ 满足 } \sum_{i=1}^M a_i = 1 \quad -0.5 \leq a_i \leq 1.5$$
 S_{best} 的保留是一种精华策略的体现,更新 S_{best} 的方式如下:淘汰 $S_{best} \cup A_t$ 中的所有劣个体,如果剩下的非劣解个数小于 $|S_{best}|$ 的最大值 m ,那么就把 $S_{best} \cup A_t$ 中的非劣解拷贝到下一代 S_{best} 中,如果超过了限定数目,则从 A_t 中的非劣解中随机挑选 $(m - |S_{best}|)$ 中的非劣解个体个数) 个个体 + S_{best} 中所有个体拷贝到下一代 S_{best} 中,从而使 S_{best} 的大小不超过限定值。

4.2 算法收敛性分析

本节利用前面的收敛性分析基本思想以及相关的定义和定理对该算法进行收敛性分析。在这里,我们不考虑问题的 Pareto 集合是无穷的情况,且假定 $|S_{best}| > |F^*|$ 。

结论 1 该算法的种群序列 $(A_t)_{t \geq 0}$ 是具有正转移矩阵的齐次有限马尔科夫链。

证明:若算法求解的是一个连续问题,但在确定的精度意义下,仍可以把状态空间 S 看成是有限集。由于算法中的参数不随时间 t 改变,因此有,

$$P\{X_{t+1} = j | X_t = i, X_{t-1} = i_{t-1}, \dots, X_0 = i_0\} = P\{X_{t+1} = j | X_t = i\} = P_{i,j}$$

根据定义 8,序列 $(A_t)_{t \geq 0}$ 的转移矩阵 P 与时间 t 无关。

现在要证明转移矩阵是正矩阵。转移矩阵通常是一些变换算子(变异、杂交)和选择算子的转移矩阵的积。本算法中的变换算子有两个:①变异算子,该算子可以产生任意的一个新个体;②杂交算子是对最好的 M 个个体进行凸组合。通过这两种操作交替执行,一个个体通过一代的演化,存在变换成为任意一个个体的可能性,因而它的转移矩阵是正的。

我们的选择算子是一种确定性的选择算子,每次总是淘汰本种群中最差的那个个体,因此,选择算子的转移矩阵是列可容的。根据定理 2,变换算子和选择算子的转移矩阵的积是正矩阵。

综上所述可以得出,该算法中的种群序列 $(A_t)_{t \geq 0}$ 是具有正转移矩阵的齐次有限马尔科夫链。

结论 2 如果算法中的种群序列 $(A_t)_{t \geq 0}$ 是具有正转移矩阵的齐次有限马尔科夫链,那么该演化算法以概率 1 收敛于极小元集。

证明:设 $f(A_t)$ 是 A_t 的像集,以下的偏序关系符号“ $\leq$ ”在算法的执行过程中,将实例化为 better 函数。 $\Gamma^* = M(\Gamma, \leq)$ 是所求某一个问题的极小元集且是完全的(Pareto 最优解集)。

首先要证明最优元素一旦被找到进入 $f(S_{best})$ 就永远不会被淘汰。 S_{best} 的更新方式说明:一个元素 $a \in S_{best}$ 从 S_{best} 被删除,当且仅当存在一个元素 $b \in A_t$: $f(b) \leq f(a)$ 。因此,根据 $\Gamma^* = M(\Gamma, \leq)$ 的定义, Γ^* 中的元素一旦进入 $f(S_{best})$ 将永远不会被删除。

下面证明种群中的非最优个体在有限步内必将被取代。令元素 $a \in S_{best}$ 且 $f(a) \notin \Gamma^*$ 。由于 $M(\Gamma, \leq)$ 是完全的,根据定义 6 一定存在一个元素 $b \in \Omega$ 满足 $f(a) < f(b)$ 。根据 S_{best} 的更新方法,只要这个更好的元素经过一定的迭代次数 t 在 $(A_t)_{t \geq 0}$ 出现,它一定会取代 a 。这一过程将不断发生,直到 $f(S_{best})$ 中没有非最优元素。由于 $(A_t)_{t \geq 0}$ 是具有正转移矩阵的

有限马尔科夫链,定理 1 保证了状态空间中的所有元素都会以概率 1 被无限次访问。因此,非最优元素经过有限次迭代后会以概率 1 被淘汰。

综上所述,所有进入 $f(S_{best})$ 的最优解都不会被淘汰,并且所有 $f(S_{best})$ 中的非最优解都会以概率 1 在有限步内被淘汰。因此,当 $T \rightarrow \infty$ 时, $f(S_{best}) \subset \Gamma^*$, 因而有 $\delta_1^*(f(S_{best})) \rightarrow 0$ 以概率 1 成立,根据定义 10,该算法以概率 1 收敛于极小元集。

5 结论

本文利用有限马尔科夫链的相关定义和结论给出了一个用来证明 MOEA 的收敛性的分析框架,并对于一个著名的 MOEA 算法,运用本分析框架给出了收敛性证明。通过分析我们可以发现,搜索算子和选择算子不仅对提高 MOEA 的效率有重要作用,而且对算法的收敛性也有很大的影响。另外,精英策略(S_{best})对算法的收敛性也有重要的作用,它可以保证精英个体在演化的过程中不被丢失,为算法收敛到 Pareto 最优集提供了保障。这些结论对设计 MOEA 的研究者来说都是很有指导意义的。本文对 MOEA 的理论工作还只是一个开端,有些细节问题还有待深入研究,如该收敛定义是否是一个 MOEA 收敛的必要条件或者只是一个充分条件。在所求问题的 Pareto 最优集是一个无穷集合,怎样定义算法收敛到极小元整集以及收敛速度等问题还有很多工作要做,这也是我们进一步的研究计划。

参考文献:

[1] alyanmoy Deb. Multi-objective Optimization Using Evolutionary Algorithms[M] .Chichester: John Wiley & Sons Ltd, 2001.34-76.

[2] Eckart Zitzler, Kalyanmoy Deb, Lothar Thiele. Comparison of Multi-objective Evolutionary Algorithms: Empirical Result[J] . Evolutionary Computation, 2000, 8(2) : 173-195.

[3] Daid A, Van Veldhuizen, Gary B Lamont. Multi-objective Evolutionary Algorithms: Analyzing the State-of-the-Art [A] . Kalyanmoy Deb, Samir Agrawal, Amrit Pratap, T Meyarivan. A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II[C] . Paris, France: Proceedings of the Parallel Problem Solving from Nature VI Conference, 2000. 849-858.

[4] 覃俊,康立山,等.演化算法的收敛性分析及算法改进[J] . 计算机工程与应用, 2003, 39(19) : 91-94.

[5] G Rudolph, A Agapie. Convergence Properties of Some Multi-objective Evolutionary Algorithm[C] . Seoul, Korea: Proceedings of the 2001 IEEE Congress on Evolutionary Computation, 2001. 27-30.

[6] 郭涛,康立山,李艳.一种求解不等式约束下函数优化问题的新算法[J] . 武汉大学学报(自然科学版), 1999, 45(5 B) : 771-775.

[7] Yan Zhenyu, Kang Lishan. Do Search and Selection Operator Play Important Roles in Multi-objective Optimization: A Case Study[J] . Wuhan University Journal of Natural Sciences, 2003, 54(8) .

[8] 覃俊,康立山,等.基于遗传算法求解多目标优化问题 Pareto 前沿[J] . 计算机工程与应用, 2003, 39(23) : 42-44.

作者简介:

覃俊(1968-),女,副教授,研究领域为遗传算法、电子商务与信息安全;康立山(1932-),男,教授,博士生导师,研究领域为演化计算、并行计算等。