

避免近亲繁殖的遗传算法

闻骥骏 张学明 沈成武

(武汉理工大学交通学院结构工程系 武汉 430063)

摘 要 针对基本遗传算法在优化应用中遇到的诸如局部搜索能力差、计算量大、对较大搜索空间适应能力差等问题,文中提出了一种改进的遗传算法。在计算过程中,引入个体差异判别运算,有效地维持了群体的多样性,使遗传算法顺利地收敛到全局最优值。通过典型测试函数数值算例验证表明,不但可以有效地克服基本遗传算法的上述缺陷,而且计算速度和效率都有显著提高。

关键词 遗传算法 群体多样性 动态标准 全局优化

1 前言

遗传算法(Genetic Algorithm,简称 GA)是新近发展起来的一种模拟生命进化机制的搜索和优化方法^[1,2]。遗传算法的主要优点有:(1)易写一个通用程序,应用广泛;(2)非线性,不需要函数的梯度信息,也不需要函数的连续性;(3)可并行性,GA 是从一个点群开始寻优,而不是从一个初始点开始寻优,因而获得的是全局最优解。

但遗传算法也有缺陷和不足,如需要较大的群体规模,重复分析次数过多,导致搜索速度较慢;局部搜索能力不足,导致寻优迭代后期,收敛速度减慢;易出现早熟收敛和随机漫游现象等^[3]。近年来遗传算法取得了蓬勃的发展,其中一个热点就是其改进问题。

2 遗传算法的改进概述

2.1 编码

为了克服二进制编码所带来的精度与效率不高,过早收敛等问题,有人提出动态参数编码,即动态地改变参变量的定义域,当群体已经收敛至一定区域,则将参变量限制在该区域,从而可在全局最优点附近进行更精确的搜索^[3]。此外变二进制码为十进制码也可以在一定程度上弥补上述不足,且更便于实际应用。

2.2 适应值函数的建立

对于有约束参数的优化设计,可采用罚函数

的方法将目标函数和约束条件建立一个无约束的优化目标函数,然后再将目标函数作适当处理,建立适合 GA 的适应值函数。遗传算法用适应值作为复制的选择压力,如果群体的适应值变化不大或过大,会引起选择压力不足或波动,导致迭代过程过早收敛或发生振荡,若将目标函数的序号作为适应值,可有效地避免上述现象的发生。

2.3 遗传操作的改进

2.3.1 复制

采用最优个体保护策略,即父代群体中的最优个体直接进入子代群体中,这样可防止最优个体由于遗传算子的偶然性而被破坏掉,且能以概率接近 1 找到全局最优解。或者将父代与子代一起排序,按群体规模 N 截取前 N 个个体进入匹配集,作为下一步交叉操作的对象,这样可使每一代中的优良个体得到保护,而淘汰那些不良个体,以期通过交叉等操作产生更优的个体,从而提高寻优速度且不一定增加很多工作量。

2.3.2 交叉

在基本交叉操作的基础上,对算法加入人工工作成分,如限定适应值高的个体按某种规则互相交叉,以期“优”“优”交叉,快速地得到“更优”或者在种群中只对一部分适应值较差的个体进行交叉和变异操作,产生更好的新个体代替这部分个体,这种做法既能减少好的基因模式和种群结构的破坏,以使遗传算法稳定地收敛,同时每代只计算替代适应值差的个体的适应值,从而可减少计算次数,加快寻优速度。此外,采用变交叉率在一定程度上可避免陷入局部最优,增强局部搜索能力。

2.3.3 变异

第一种方式为改变通常的变异率不变的做法,如自适应性变异,即根据个体间的 Hamming 距离决定变异率 P_m 。Hamming 距离越大, P_m 越小,反之越大;或者在开始阶段取较小变异,在后期迭代到群体中已经难以产生适应值更高的个体时,将变异率增大,以期产生新的更好的个体,避免陷入局部最优。第二种方式是改变变异算子的作用范围。为了改变变异算子在整个寻优过程中的发散性,引入起过程控制作用的时变因子,该因子使变异算子在开始一段时间可以作用于整个可行域,但到一定时候则随搜索的次数增加而减小其作用范围。

2.3.4 收敛判定的改进

目前,GA 终止条件的确定有多种判定规则,如已找到可接受的最优个体或已达到预定的繁殖代数等,为了避免 GA 提前收敛或收敛于局部最优,可采用局部最优徘徊策略作为辅助终止规则,即达到上述终止条件时,先不退出循环,而重复寻优过程,重复步 $< (5 \sim 10)$,并加大交叉率和变异率,使遗传算法充分发挥多点、并行的搜索寻优,以求获得全局和近于全局的最优解,当两个规则均满足要求时可停止操作,输出最优解。

3 避免近亲繁殖的遗传算法

3.1 个体差异的衡量

初始种群的建立采用 Hamming 距离控制个体差异,防止近亲繁殖。在产生种群的过程中,若新个体与前面产生的某一个体的 Hamming 距离小于某一设定值,则重新产生新个体,直到得到所有个体之间均有一定差异的种群,由于种群的个体解遍及整个解空间,因而能很好地反映搜索空间,不仅能加快搜索进度,也可以避免局部最优。

3.2 种群多样性的控制

将子代种群的 $size$ 个个体和 c 步记忆的 NP 个个体合并在一起,得到一个含有 $M = size + NP$ 个个体的新群体;对这 M 个个体按下式计算每两个个体 X_i, X_j 的海明距离:

$$\|X_i - X_j\| = \sqrt{\sum_{k=1}^{size} (X_{ik} - X_{jk})^2}$$

$$i = 1, 2, \dots, M-1; j = i+1, \dots, M$$

当 $\|X_i - X_j\| < L$ 时,比较个体的适应度大小,并淘汰其中适应度较低的个体。

3.3 个体间距离的动态判别

先将个体间的距离判别标准设为动态函数。文中,将种群的繁殖代数 n 引入距离判别函数:

$$l = L/n$$

这样保证了在进化初期,个体间都具有较大的差异,种群拥有非常大的多样性。因此,能以较小的种群规模,在广大领域内搜索。除以上操作方法有所不同外,其余算法如适应度确定、杂交操作、变异操作规则等,二者是一致的,不再赘述。

3.4 算法运行参数

种群中个体总数 $size$ 、交叉概率 P_c 、变异概率 P_m 等参数对算法的运行性能也有影响,需结合具体问题认真选择。文中依据一般建议及计算经验,取交叉概率 P_c 为 0.7,取变异概率 P_m 为 0.1。加速循环的最优个体数目 N_A 取为 $size$ 的 10%。

4 算例分析

基于以上算法,笔者编制了相应的程序(见图1),并利用评价遗传算法性能的典型测试函数: Schaffer 函数 F_6 , 进行改进遗传算法(HGA)和基本遗传算法(GA)性能的对比研究。

Schaffer 函数 F_6 的具体形式为^[5]:

$$f_6(x_1, x_2) = 0.5 + \frac{\sin^2 \sqrt{x_1^2 + x_2^2} - 0.5}{[1.0 + 0.001(x_1^2 + x_2^2)]^2}$$

$$-100 \leq x_i \leq 100 \quad i = 1, 2$$

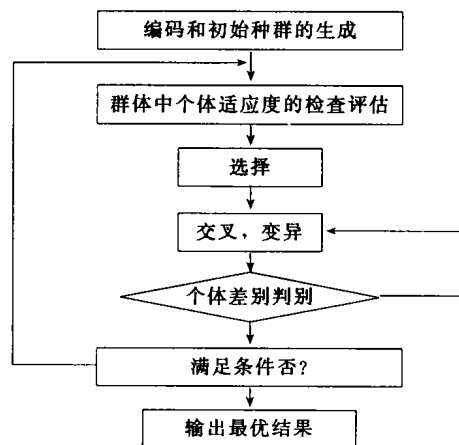


图1 改进遗传算法流程图

虽然该函数在其定义域内只有一个全局极小点 $f_6(0.0, 0.0) = 0.0$, 但由于变量的取值范围大,采用传统的直接搜索法求解时,因搜索空间太大而无法求得全局最优解,采用 GA 搜索时,由于

算法局部搜索能力差,因而需要设置相当大的种群规模,需耗费巨大的计算量以得到全局最优解。如何有效地求解这类搜索空间巨大的全局优化问题一直是人们关注的一个焦点。文中采用加强了局部搜索能力的混合遗传算法有效地解决了这一问题。表1所列为分别用GA和HGA求解该问题的收敛过程。由表1可以看出,当size取100时,与DeJohn函数 F_2 相似,GA在第三代就停留在某一局部最优值上不再进化,有明显的早熟收敛现象,不同的是,当GA种群规模增大到1500时,GA的收敛结果仍远离全局最优解。而采用

HGA搜索,当size取100时,不仅消除了早熟收敛现象,且收敛精度高于种群规模为1500时GA的精度;当种群规模仅增大到500时,HGA就用较快的速度,以 10^{-5} 的精度收敛于解析解。可见,相对于GA而言,对于搜索空间巨大的优化问题,HGA在节省计算量、提高算法效率、提高计算精度等方面的优越性更加显著。表1的结果同时表明,由于搜索空间极大,增加种群规模对GA运算性能的影响不如前一算例明显,而HGA的运算性能随种群的增大仍然有大幅度提高。可见HGA对大搜索空间的适应性明显优于GA。

表1 Schaffer函数 F_6 的GA和HGA收敛过程

种群规模	GA size = 100			GA size = 500			HGA size = 100			HGA size = 500		
	$F(x)$	x_1	x_2	$F(x)$	x_1	x_2	$F(x)$	x_1	x_2	$F(x)$	x_1	x_2
代数												
1	0.271 275 17	0.569 09	4.413 7	0.012 069	0.484 7	3.051 7	0.218 729	3.439 4	5.729 1	0.186 991 7	3.059 6	-4.494 4
2	0.178 326 15	0.061 7	6.945 1	0.012 063	0.484 8	3.051 4	0.098 999	7.6795	-5.429 7	0.087 225	-6.814 6	-4.364 7
3	0.178 222 14	0.051 3	6.993 9	0.009 972	3.004 8	0.848 4	0.013 793	-0.893 0	-3.612 1	0.035 469	-5.351 9	-2.295 0
4	0.178 222 14	0.051 3	6.993 9	0.009 731	3.011 8	0.868 7	0.013 759	3.420 9	-0.932 7	0.014 792	-2.315 7	-2.172 9
5	0.178 222 14	0.051 3	6.993 9	0.009 716	0.589 1	3.083 0	0.012 737	-1.518 3	-2.996 9	0.009 826	-2.801 9	-1.135 5
6	0.178 222 14	0.051 3	6.993 9	0.009 716	0.560 2	3.088 1	0.011 716	2.938 8	1.542 4	0.009 726	-2.976 4	-0.999 5
7	0.178 222 14	0.051 3	6.993 9	0.009 716	0.883 1	3.010 8	0.007 837	-0.011 7	0.796 1	0.009 67	-2.904 2	-0.992 9
8	0.178 222 14	0.051 3	6.993 9	0.009 716	0.565 1	3.087 2	0.006 896	0.003 6	0.069 7	0.008 621	0.004 2	-0.093 6
9	0.178 222 14	0.051 3	6.993 9	0.009 716	0.549 8	3.090 0	0.006 791	0.007 3	-0.066 4	0.000 000	0.000 0	0.000 0

5 结语

(1) 函数优化算例表明,文中方法具有很好的可行性和稳定性,能以很快的速度收敛到全局最优解或近似全局最优解,并行计算又克服了效率低的缺点。此方法有待于在更多的领域内应用,以检验其性能。

(2) 引入个体差异判别运算,有效地维持了群体的多样性,较好地改善了个体的分布特性,从而可减小遗传算法的种群规模,降低运算量,为遗传算法用于大规模工程优化实践提供了可能。

参考文献

- 1 刘杰,王媛.一种高效混合遗传算法.河海大学学报,2002(4)
- 2 沈成武,闻骥骏.大跨度桥梁预应力损失的遗传算法识别.武汉理工大学学报,2002(1)
- 3 Thomas Strassner, Markus Busold, Wolfgang A. Herrmann. MM3 parametrization of four- and five-coordinated rhenium complexes by a genetic algorithm. Journal of Computational chemistry 2002, 23
- 4 Steven Chien, Zhaowei Yang, and Edwin Hou. Genetic Algorithm Approach for Transit Route Planning and Design. Journal of Transportation Engineering 2001, 9
- 5 Byeong-Keun Choi and Bo-Suk Yang. Optimal Design of Rotor-Bearing Systems Using Immune-Genetic Algorithm. Journal of Vibration and Acoustics, 2001, 64(1-4)

Improved Genetic Algorithm and Its Application

Wen Jijun; Zhang Xueming; Sheng Chengwu

Abstract: Aimed at the limitation of genetic algorithm applied in practice such as poor local search ability, premature convergence, excessive computational cost and bad adapt ability to large search space, an improved is proposed, which can keep the diversity of population, and thereby converge conveniently on global solutions. Typical test function case studies show that the improved genetic algorithm is more efficient and reliable.

Key words: genetic algorithm; population diversity; dynamic criteria; global optimization