

浮点遗传算法中一种新的杂交算子

汪剑鸣, 许镇琳

(天津大学 自动化学院, 天津 300072)

摘要: 为了提高浮点遗传算法在优化计算时向最优解收敛的速度, 提出了一种新的遗传算子: 代间差分杂交算子. 通过应用于非线性参数估计的仿真计算, 表明了这种杂交算子的有效性及其相对于普通杂交算子的优点.

关键词: 遗传算法; 全局优化; 参数估计

中图分类号: O242.23

文献标识码: A

New crossover operator in float-point genetic algorithms

WANG Jian-ming, XU Zhen-lin

(School of Automation Tianjin University, Tianjin 300072, China)

Abstract: To improve the convergent speed of float-point genetic algorithms, a new genetic operator named intergenerational differential crossover operator was proposed. The operator was applied to simulating parameter estimation of nonlinear system, and the resulted showed its validity and its superiority over a general crossover operator.

Key words: genetic algorithms; global optimization; parameter estimation

1 引言(Introduction)

在过去的几十年中经典智能理论得到了长足的发展, 例如采用联结机制的人工神经网络理论和采用符号机制的专家系统理论已经被广泛地应用到各个领域. 虽然这样, 当人们面对越来越突出的关于处理模式信号、信息组合爆炸等困难时, 发现经典智能技术对此也显得无能为力. 幸运的是, 人们逐渐认识到了可以向已经有几十亿年进化历史的自然界学习, 也正是在这样一种背景下, 产生了以遗传算法(GA)为代表的采用进化机制的智能理论^[1]. 遗传算法属于随机优化算法, 具有内在并行性、自适应、自学习等特征^[2,3], 通过对自然界中生物进化过程的模拟, 可以在对系统了解很少或完全不了解的情况下找到系统的最优解.

遗传算法最初只采用二进制编码进行遗传操作, 这类遗传算法被称为简单遗传算法(SGA). 但是进入 90 年代后, 实数编码被引入到遗传算法中, 从而产生了采用实数编码进行遗传操作的浮点遗传算法(FGA).

在 FGA 的应用过程中, 人们发现收敛速度慢是这种算法的一个较为明显的不足. 为了提高 FGA 向

最优解的收敛速度, 本文提出了一种新的遗传算子: 代间差分杂交算子 (intergenerational differential crossover operator, IDCO), 通过仿真计算证实了这种算子的有效性及其相对于普通杂交算子的优点.

2 代间差分杂交算子(Intergenerational differential crossover operator)

在 FGA 中, 杂交算子通常是从同一代群体中随机选择两个个体, 然后进行重新组合产生新个体, 如式(1)所示.

$$\begin{cases} x_i^{n+1} = \eta x_k^n + (1 - \eta)x_l^n, \\ x_{i+1}^{n+1} = (1 - \eta)x_k^n + \eta x_l^n. \end{cases} \quad (1)$$

其中 $0 < i, k, l, \leq N; \eta \in (0, 1)$.

N 为种群的规模, 即种群中各体的总数, 一般在计算过程中保持不变. η 为 0 到 1 之间的一个随机数. x_k^n 和 x_l^n 为从第 n 代种群中随机选取的个体, x_i^{n+1} 和 x_{i+1}^{n+1} 为杂交产生的新个体. 为了便于区别, 本文中把具有式(1)形式的杂交算子称为普通杂交算子, 采用普通杂交算子的 FGA 称为普通的 FGA.

从式(1)可以看出, 普通杂交算子产生的新个体仅仅是某一代种群中现有信息的重新组合.

本文中介绍的 IDCO 则是从相邻两代群体中各

随机选择一个个体来产生新个体,如式(2)所示.

$$x_i^{n+1} = \alpha x_k^n + \beta (x_k^n - n_l^{n-1}) + \lambda (x_{\text{best}}^n - x_{\text{best}}^{n-1}). \quad (2)$$

其中 $0 < i, k, l < N; \alpha, \beta, \lambda \in (0, 1]$.

α, β 和 λ 是小于 1 的正常数. x_{best}^{n-1} 和 x_{best}^n 分别为第 $n-1$ 代和第 n 代中适应性最好的个体. 其它与式(1)相同. 又按式(3)计算个体 x :

$$x = x_{\text{best}}^n + \lambda (x_{\text{best}}^n - x_{\text{best}}^{n-1}). \quad (3)$$

即在新种群产生后,又进行如下操作:从新种群中随机选择一个个体,如该个体的适应值比 x 好,则保持不变,否则用 x 代替该个体. 式(3)中 λ 与式(2)中相同.

根据自然界中生物进化适者生存的思想,随着种群进化代数的增加,种群的平均适应值将以大概率增加. 这样在精英算法条件下,有理由假定种群的适应值随进化代数的增加而单调增加,从相邻两代种群中各自随机选择一个个体,则两个体的差以一定概率代表了种群适应值增加的方向,也就是所希望的进化方向. 所以,比较普通的杂交算子和 IDCO 可以发现,后者较前者更充分地利用了进化种群之间包含的信息,产生的新个体将以更高的概率向最优解接近,使进化过程以更快的速度向系统的最优区域进行.

需要指出的是,由于采用了差分的形式来产生新个体,新个体可能会越出定义域. 如果不严格要求个体完全限定在定义域内,对此可不作处理,此时带有 IDCO 的 FGA 可以找到位于定义域以外的最优解;如果要求个体严格限定在定义域内,则需要在实现 IDCO 时对生产的新个体加以判断,对于不在定义域内的个体丢掉不用,并重新进行一次杂交计算.

3 仿真研究(Simulation research)

人们在进行科学研究时,经常要用回归分析方法来处理实验数据. 尤其近年来,非线性回归分析更是被广泛地应用到许多学科领域. 但由于非线性模型的表达式往往比较复杂,通常的参数估计方法要依赖非线性模型的表达式而不能普遍适用. 遗传算法可适用于一般非线性系统模型的参数估计^[4]. 下面将分别用含有普通杂交算子和 IDCO 的 FGA 来实

现非线性回归分析的参数估计. 通过两种算法结果的比较,来表明 IDCO 的有效性和优点.

本文中采用的回归模型为

$$f(x) = a - bc^x. \quad (4)$$

为了验证 IDCO 的作用,令 $a = 0.5, b = 0.5, c = 0.5$, 取 5 个不含噪声的采样点,用于非线性参数估计的采样数据如表 1.

表 1 用于非线性参数估计的采样数据

Table 1 Sampled data for nonlinear parameter estimation

x	1.0	2.0	3.0	4.0	5.0
$f(x)$	0.2500	0.3750	0.4375	0.4688	0.4844

在进行仿真计算时,两种 FGA 的下列控制参数相同:种群规模 $N = 50$;繁殖概率 $p_r = 0.1$;杂交概率为 $p_c = 0.9$;变异概率 $p_m = 0.15$.

IDOC 算子的参数如下: $\alpha = 1, \beta = 0.2, \lambda = 0.8$. 定义仿真结果评价误差,如式(5)所示.

$$SE = \sum_{i=0}^2 (a_i - 0.5)^2, i = 0, 1, 2. \quad (5)$$

式(5)中 $a_0 = a, a_1 = b, a_2 = c$. 并认为当 $|a_i - 0.5| \leq 0.1, i = 0, 1, 2$ 成立时,计算结果收敛. 算法结束条件为:进化过程达到设定的最大进化代数. 在三个不同的解空间中利用两种 FGA 进行搜索,得到了三组仿真结果. 在每一组仿真结果中,随机进行 10 次计算,取在收敛条件下的最好和最差的结果(利用 SE 来评价最好与最差的结果)及不收敛次数进行比较,如表 2~4.

通过分析计算结果可以发现:在搜索较小的解空间时,两种 FGA 都可以得到比较满意的结果,收敛速度没有明显的区别,表明 IDCO 是一种有效的杂交算子;搜索的解空间较大时,在相同的最大进化代数下,带有 IDCO 的 FGA 比普通 FGA 有更小的 SE,表明前者的计算结果更接近真实值. 同时,在相同的最大进化代数下,前者的不收敛次数明显少于后者,并且随着最大进化代数的增加前者的收敛次数减小的速度明显比后者要快. 因此可以说,带有 IDCO 的 FGA 的收敛速度明显比普通 FGA 的收敛速度要快.

表 2 第一组仿真结果
Table 2 The first simulation result

进化 代数	收索的解空间 $a, b \in [0,1]$ $c \in [0,1]$	普通 FGA			带 IDCO 的 FGA		
		最好 结果	最差 结果	不收敛 次 数	最好 结果	最差 结果	不收敛 次 数
20	<i>ab</i>	0.5001	0.4813	0	0.4976	0.5034	0
	<i>c</i>	0.4935	0.5875		0.5138	0.4654	
	SE	0.4913	0.4121		0.4904	0.5313	
		0.000161	0.015725		0.000290	0.002185	
50	<i>ab</i>	0.5005	0.4795	0	0.5004	0.4912	0
	<i>c</i>	0.5036	0.5198		0.4936	0.5236	
	SE	0.4920	0.4237		0.5058	0.4648	
		0.000077	0.006623		0.000075	0.001874	
100	<i>ab</i>	0.4977	0.4848	0	0.5003	0.4972	0
	<i>c</i>	0.5116	0.5515		0.5013	0.5108	
	SE	0.4859	0.4413		0.5023	0.4900	
		0.000339	0.006330		0.000008	0.000266	

表 3 第二组仿真结果
Table 3 The second simulation result

进化 代数	收索的解空间 $a, b \in [0,5]$ $c \in [0,1]$	普通 FGA			带 IDCO 的 FGA		
		最好 结果	最差 结果	不收敛 次 数	最好 结果	最差 结果	不收敛 次 数
500	<i>ab</i>	0.5006	0.5065	4	0.4981	0.5027	0
	<i>c</i>	0.5136	0.5725		0.4997	0.4664	
	SE	0.4851	0.4647		0.4966	0.5280	
		0.000407	0.005650		0.000015	0.001918	
1000	<i>ab</i>	0.5073	0.4784	3	0.5018	0.4985	0
	<i>c</i>	0.4813	0.5992		0.5014	0.5250	
	SE	0.5164	0.4057		0.5000	0.4829	
		0.000673	0.019200		0.000005	0.000918	
2000	<i>ab</i>	0.4966	0.5194	0	0.4974	0.5039	0
	<i>c</i>	0.5078	0.4828		0.5017	0.4771	
	SE	0.4896	0.5672		0.4963	0.5244	
		0.000180	0.005163		0.000023	0.001132	

表 4 第三组仿真结果
Table 4 The third simulation result

进化 代数	收索的解空间 $a, b \in [0,20]$ $c \in [0,1]$	普通 FGA			带 IDCO 的 FGA		
		最好 结果	最差 结果	不收敛 次 数	最好 结果	最差 结果	不收敛 次 数
1000	<i>ab</i>	0.4992	0.5043	6	0.5068	0.5333	2
	<i>c</i>	0.5237	0.5689		0.5156	0.4963	
	SE	0.4665	0.4478		0.5127	0.5831	
		0.001688	0.007490		0.000454	0.008030	
2000	<i>ab</i>	0.5060	0.5368	7	0.4956	0.4951	1
	<i>c</i>	0.4775	0.5990		0.5027	0.5635	
	SE	0.5365	0.5364		0.4991	0.4447	
		0.001876	0.013048		0.000027	0.007111	
3000	<i>ab</i>	0.5095	0.4989	5	0.5086	0.4943	0
	<i>c</i>	0.4966	0.6000		0.5163	0.5698	
	SE	0.5261	0.4385		0.5023	0.4537	
		0.000781	0.014356		0.000349	0.007056	

4 结论(Conclusion)

为了提高 FGA 向最优解收敛的速度,本文首先假定认为进化种群的平均适应值随着进化代数的增加而单调增加,然后在此基础上提出了代间差分杂交算子.把带有代间差分杂交算子的 FGA 和普通的 FGA 同时用于非线性参数估计的仿真计算,计算结果表明代间差分杂交算子是有效的,并且同普通杂交算子相比它能使进化过程更快更准确地向最优解收敛.

参考文献(References)

- [1] Dong Cong, Guo Xiaohua. The research and advances on some focus questions in computation intelligence [J]. Control Theory and Appli-

cations, 2000, 17(5):691 - 697 (in Chinese)

- [2] Eiben A E, Rudolph G. Theory of evolutionary algorithms: a bird's eye view[J]. Theoretical Computer Science 1999,229(1):3 - 9
 [3] Pan Zhengjun, Kong Lishan. Evoulutionary Algorithms [M]. Beijing:Tsinghua University Press,1998 (in Chinese)
 [4] Jiang Bo, Wang Bingwen. Parameter estimation of nonlinear system based on genetic algorithms [J]. Control Theory and Applications, 2000,17(1):150 - 152 (in Chinese)

本文作者简介

汪剑鸣 1974 年生.现为天津大学自动化学院博士研究生.主要研究方向:自动控制的智能化、网络化. Email:sword4000@eyou.com

许镇琳 1939 年生.现为天津大学自动化学院教授,博士生导师.主要研究方向:自动化,电力电子与电力传动.

(上接第 976 页)

参考文献(References)

- [1] Hen J, Nett C N. The Caratheodory-Fejer problem and H_∞/L_1 identification: a time domain approach [J]. IEEE Trans. Automatic Control, 1995, 40(4): 729 - 735

本文作者简介

高 林 1973 年生.分别在 1993 年、1996 年、1999 年于东北大学

自动控制系本科、硕士、博士毕业.清华大学博士后.研究方向:复杂系统的建模,优化与仿真,智能优化方法,计算机应用. Email:lingao_neu@263.net

王书宁 1956 年生.1982 年湖南大学电气工程系本科毕业.1984 年和 1988 年于华中理工大学自动控制系硕士、博士毕业.现为清华大学自动化系教授、博士生导师.主要研究方向为系统辨识,参数估计,优化理论和算法以及决策分析.

王 伟 1975 年生.1995 年江苏理工大学本科毕业.现为清华大学自动化系硕士生.