

基于非参数统计的生物启发式优化算法性能评估

赵 辉, 李牧东*, 翁兴伟, 周 欢

(空军工程大学航空航天工程学院, 陕西西安, 710038)

摘要 由于生物启发式优化算法性能评估方法存在单一性、片面性且无法进行统一地检验分析问题, 从而严重影响了对优化算法性能的深入研究而导致各类优化算法无法准确地应用于解决实际问题。针对这一问题, 利用非参数统计检验中典型的 Wilcoxon 符号秩检验和 Quade 检验方法, 对 5 种生物启发式优化算法在 36 种测试函数条件下的仿真结果进行检验分析。测试结果表明: 上述检验方法能够有效地对不同类型的优化算法性能进行分析比较, JADE 算法相比于其他 4 种算法, 在收敛速度及搜索精度方面表现最优, 而 GWO 算法在精度稳定性方面相比于其他 4 种算法表现出较优的性能, 对各类生物启发式优化算法优化性能的评估与比较提供了新的思路。

关键词 生物启发式优化算法; 非参数统计; Wilcoxon 符号秩检验; Quade 检验

DOI 10.3969/j.issn.1009-3516.2015.01.020

中图分类号 TP301.6 **文献标志码** A **文章编号** 1009-3516(2015)01-0089-06

Performance Evaluation for Biology-inspired Optimization Algorithms Based on Nonparametric Statistics

ZHAO Hui, LI Mu-dong, WENG Xing-wei, ZHOU Huan

(Aeronautics and Astronautics College, Air Force Engineering University, Xi'an 710038, China)

Abstract: Aimed at the problems that the biology-inspired optimization algorithms are of oneness, one-sidedness and fail to check and analyze uniformly the performance evaluation, thus having a strong impact on the intensive study for performance of the optimization algorithms and failing to solve practical problems accurately, two classical nonparametric statistics methods named Wilcoxon Sign Rank test and Quade test are utilized for testing and analyzing the simulation results of five different BOAs under the conditions of thirty-six different test functions. The experimental results show that the two test methods can be used effectively to compare and analyze the optimization performances of different optimization algorithms. JADE algorithm is most superior in convergence speed and search accuracy compared with the other four algorithms, whereas, GWO has comparatively superior performance in the aspect of stability compared with other four algorithms. And this provides a new idea for evaluating and comparing the performances of different BOAs.

Key words: biology-inspired optimization algorithms; nonparametric statistics; Wilcoxon sign rank test;

收稿日期: 2014-09-22

基金项目: 航空科学基金资助项目(20105169016); 中国博士后基金资助项目(2012M5211807)

作者简介: 赵 辉(1974—), 男, 河北沧州人, 教授, 博士, 主要从事武器系统与运用工程、最优化方法等研究。

* 通信作者: 李牧东(1987—), 男, 博士生, 从事智能优化算法、武器系统与运用工程等研究。E-mail: modern_lee@163.com

引用格式: 赵辉, 李牧东, 翁兴伟, 等. 基于非参数统计的生物启发式优化算法性能评估[J]. 空军工程大学学报: 自然科学版, 2015, 16(1): 89-94. ZHAO Hui, LI Mudong, WENG Xingwei, et al. Performance Evaluation for Biology-inspired Optimization Algorithms Based on Nonparametric Statistics[J]. Journal of air force engineering university: natural science edition, 2015, 16(1): 89-94.

Quade test

生物启发式优化算法(Biology-inspired Optimization Algorithms, BOA)是基于生物学理论与原理设计并发展的自然启发式算法,主要包括进化算法(Evolutionary Algorithms, EA)和群集智能算法(Swarm Intelligence, SI)两大类^[1]。Holland于1992年提出的受到达尔文自然选择理论以及遗传学启发的遗传算法(Genetic Algorithms, GA)^[2]开启了生物启发式优化算法的先河。随后,相继出现了一系列典型的优化算法,如较为经典的粒子群算法(Particle Swarm Optimization, PSO)^[3]、蚁群算法(Ant Colony Optimization, ACO)^[4]及其改进算法^[5],以及近两年提出的人工蜂群算法(Artificial Bee Colony, ABC)^[6]、灰狼算法(Grey Wolf Optimizer, GWO)^[7]和布谷鸟搜索算法(Cuckoo-Search Algorithm, CK)^[8]等等。

随着各类生物启发式优化算法的种类与性能以及各类测试函数的不断发展,如何准确地、方便地评价各类算法的优化性能成为了制约算法应用发展的一个关键问题。目前,较为广泛的方法是利用不同类型的测试函数,对各类优化算法独立运行后的优化结果计算其平均值或方差等统计参数的方法进行逐一比较并加以评价。然而通过这种方法分析出的结果较为片面性且无法进行统一地定性分析,特别是对于多种优化算法在多个测试函数条件下,由于得出的实验数据较多,准确地对其性能进行比较评价显得更加困难。为此,本文利用非参数统计学中Wilcoxon符号秩检验^[9]和Quade检验^[10]2种方法来解决不同类型生物启发式优化算法的性能评估问题,分别通过上述2种方法所对应的成对符号检验和多符号检验的方法对5种典型生物启发式优化算法的仿真结果进行检验分析,得出不同算法优化性能的比较结果。

1 2种非参数统计检验方法

针对各类生物启发式优化算法在比较其优化性能时对于不同的测试函数存在2种算法之间的比较以及多种算法之间的比较存在繁琐、缺乏直观表现的问题,本文将非参数统计中2种与之相对应的典型检验方法引入到对优化算法的优化性能评估中,即Wilcoxon符号秩检验和Quade检验,以期对各类不同算法给出较准确且可靠的性能评估。

1.1 Wilcoxon符号秩检验

Wilcoxon符号秩检验是用来回答2组独立样

本之间是否存在明显差异问题^[9]。在Wilcoxon符号秩检验中,它将观测值和零假设的中心位置之差的绝对值的秩分别按照不同的符号相加作为其检验统计量。不同于T检验,Wilcoxon符号秩检验不求成对的数据之差服从正态分布,只要求其对称即可,因此能够用于优化算法之间对于不同测试函数优化结果的比较它的检验过程具体如下。令 d_i 为2种算法在优化 n 个问题时所对应的第 i 个问题之差,将 n 个 d_i 进行排序,并计算它们的秩。

1)计算秩和。令 R^+ 为第一个算法优于第二个算法的秩和, R^- 为第一个算法劣于第二个算法的秩和,计算公式如下:

$$R^+ = \sum_{d_i > 0} \text{rank}(d_i) + \frac{1}{2} \sum_{d_i = 0} \text{rank}(d_i) \quad (1)$$

$$R^- = \sum_{d_i < 0} \text{rank}(d_i) + \frac{1}{2} \sum_{d_i = 0} \text{rank}(d_i) \quad (2)$$

式中: R^+ 与 R^- 之和为 $0.5n(n+1)$ 。

2)双边检验。在零假设条件下,对于双边检验 $H_0: M = M_0 \leftrightarrow H_1: M \neq M_0$, R^+ 与 R^- 应差异不大,即2种算法性能并无明显差异;若当中之一很小,即 $T = \min(R^+, R^-)$ 很小,则怀疑假设,即2种算法之间优化性能存在明显差异。

3)查表求 p 值。根据得到的 T 值,查Wilcoxon符号秩检验的分布表以得到在零假设下的 p 值,也可以通过软件计算出 p 值,如目前流行的SPSS、SAS和R等软件均可以有效地计算 p 值。如果 p 值较小(即小于或等于给定的显著水平 $\alpha = 0.05$),则可以拒绝零假设。

Wilcoxon符号秩检验是在样本数据对称的假定条件之下进行的,相比于符号检验具有更多的信息,因而得到的结果也更加可靠。因此,这种检验方法可以用来对2种不同优化算法对于各类测试函数所得出的平均值或者方差等统计结果进行成对的比较,从而可以对2种优化算法的优化性能进行统一的定性分析。

1.2 Quade检验

在不同算法性能的比较中,往往需要比较2种以上不同算法对于各类测试函数的优化性能,这时就需要进行多组数据之间的比较。上述Wilcoxon符号秩检验仅适用于2种算法之间相互比较,对于多种算法优化性能的同时比较却束手无策,而Quade检验作为用来处理完全随机区组设计数据的方法,为解决多算法之间性能相互比较提供了新的思路。

Quade 检验要求各组数据之间相互独立,且适用于数据类型为非正态分布的计量资料。检验假设 H_0 为各类算法总体分布相同,即算法性能几乎没有差异; H_1 为各算法总体分布不全相同,即各类算法之间存在差异。对于每个待优化的问题 i ,不同算法的优化结果按照最优值到最差值的顺序从 1 到 k 进行排序,将这些排序后的结果记为 $r_{i,j} (1 \leq j \leq k)$ 。

1) 计算组内极差。组内极差为对于同一个优化问题 i ,不同优化算法优化结果的最优值与最差值之间的差,即:

$$Y_{i,j} = \max_j x_{i,j} - \min_j x_{i,j} \quad (3)$$

2) 计算组内各观察值得相对大小 $S_{i,j}$ 以及合计值 S_j 。令 Q_i 为各组极差值由小到大编排的秩次,则:

$$S_{i,j} = Q_i \left[r_{i,j} - \frac{k+1}{2} \right] \quad (4)$$

$$S_j = \sum_{i=1}^n S_{i,j} \quad (5)$$

式中: $(k+1)/2$ 为组内的平均秩次, $j=1,2,\dots,k$ 。为了方便起见,我们也可以省略平均秩次,即:

$$W_{i,j} = Q_i [r_{i,j}] \quad (6)$$

而与之相对应的第 j 种算法的合计值 T_j 则按下式计算:

$$T_j = \frac{W_j}{n(n+1)/2} \quad (7)$$

式中: $W_j = \sum_{i=1}^n W_{i,j} (j=1,2,\dots,k)$ 。

3) 计算检验统计值 F_Q 。

$$F_Q = \frac{(n-1)B}{A-B} \quad (8)$$

式中: A 为总平方和; B 为处理平方和。

$$A = n(n+1)(2n+1)k(k+1)(k-1)/72 \quad (9)$$

$$B = \frac{1}{n} \sum_{j=1}^k kS_j^2 \quad (10)$$

4) 假设判断。根据自由度 $(k-1)$ 和 $(n-1)(k-1)$ 及显著水平 α ,通过查询 F 分布表得出 F 界值。若计算出的 F_Q 大于 F 界值,则在水平上拒绝 H_0 ,接收 H_1 。

本文利用 Quade 检验中计算各类算法的 T_j 与 S_j 值实现评价比较各类优化算法优化性能,合计值越小,则算法对应的性能越好。

2 仿真实验及结果分析

2.1 实验准备

对于测试不同的 BOA 算法性能,目前较为认可的方法是利用不同类型的标准测试函数,通过多

次独立运行计算其优化后最优值的平均值和方差等统计值来进行性能评估。本文在此基础上,利用非参数统计中 Wilcoxon 符号秩检验和 Quade 检验两种方法对实验结果进行分析,得出更为可靠的性能分析结果。首先,本文从文献[11]中所提到的测试函数中选取 36 个典型的标准测试函数见表 1。

表 1 36 个典型测试函数

Tab.1 Thirty-six benchmark test functions

函数	名称	类型	维数	搜索区间	最优值
F1	Ackley	MN	30	[-32,32]	0
F2	Alpine	MS	30	[-10,10]	0
F3	Bird	MN	2	[-2 π ,2 π]	-106.764 537
F4	Camel6	MN	2	[-5,5]	-1.031 6
F5	Colville	MN	4	[-10,10]	0
F6	Cross-in-tray	MN	2	[-10,10]	-2.062 612 18
F7	Cube	UN	2	[-10,10]	0
F8	Dixon&Price	UN	30	[-10,10]	0
F9	Easom	MS	2	[-100,100]	-1
F10	De Jong's 4	MN	30	[-1.28,1.28]	0
F11	Trigonometric2	MN	30	[-500,500]	1
F12	Foxholes	MS	2	[-65.536,65.536]	0.998 004
F13	Goldstein Price	MN	2	[-2,2]	3
F14	Giunta	MS	2	[-1,1]	0.060 447 04
F15	Griewank	MN	30	[-600,600]	0
F16	Hartman6	MN	6	[0,1]	-3.322 36
F17	Himmelblau	MN	2	[-5,5]	0
F18	Langerman	MN	5	[0,10]	-1.5
F19	Leon	UN	2	[-1.2,1.2]	0
F20	Matyas	UN	2	[-10,10]	0
F21	Michalewics	MS	2	[0, π]	-1.801 3
F22	Michalewics	MS	5	[0, π]	-4.687 658
F23	Michalewics	MS	10	[0, π]	-9.660 15
F24	Pen Holder	MN	2	[-11,11]	-0.963 54
F25	Powell	UN	24	[-4,5]	0
F26	Powersum	MN	4	[0,4]	0
F27	Quartic	US	30	[-1.28,1.28]	0
F28	Rosenbrock	UN	30	[-30,30]	0
F29	Rotated Hyper	MN	30	[-65.536,65.536]	0
F30	Schaffer4	UN	2	[-100,100]	0.292 579
F31	Schaffer6	MN	30	[-100,100]	0
F32	Schwefel2,21	US	30	[-100,100]	0
F33	Shekel5	MN	4	[0,10]	-10.153 2
F34	Sphere	US	30	[-100,100]	0
F35	Sum Squares	US	30	[-5.12,5.12]	0
F36	Sum Powers	US	30	[-1,1]	0

考虑到文章篇幅限制,表 1 中仅给出了这些测试函数的基本描述,具体描述见文献[11]。其中 M 为多峰函数, U 为单峰函数, S 为可分函数, N 为不可分函数。为了充分说明本文采用的优化算法性能评估方法的有效性,本文对 5 种目前较为流行的生物启发式优化算法进行比较,分别为广泛学习粒子群算法 (Comprehensive Learning PSO, CLP-SO)^[12]、ABC、CK、GWO 和 JADE^[13] 算法,其中 JADE 算法作为标准算法与其它 4 种算法进行比较,从而验证其优化性能。上述算法的简要描述和参数设置如下:

1)CLPSO:CLPSO 算法是在粒子群算法基础上利用广泛学习策略提出的改进算法,该算法的参数设置为 $C=1.494\ 45, m=0, \omega_0=0.9, \omega_1=0.4$;

2)ABC: ABC 算法是模拟蜜蜂采蜜过程而提出的新型算法,参数设置为 $\text{limit}=1\ 000$,观察蜂个数=雇佣蜂个数=种群个数/2;

3)GWO:GWO 算法是通过模拟灰狼群在猎食时的行为特点设计并提出的新型算法,参数设置为 $\alpha = 2 - 1 \times (2/\text{Max Cycle})$ 。

4)CK:CK 算法是模拟布谷鸟产蛋活动而提出的新型生物启发式优化算法,该算法的参数设置为 $b = 1.50, p_0 = 0.25$;

5)JADE: JADE 算法是通过利用自适应原理对 DE 算法的控制参数加以改进来提高算法的性能,参数设置为 $CR_m = 0.5, F_m = 0.5, \text{Adactor} = 1, c = 1/10, p = 0.05$;

为了更好地验证算法的优化性能,所有试验均在硬件配置为 Intel(R) Core(TM),CPU:i5-3470,主频 3.20 GHz,内存为 3.46GB 的计算机上进行,程序采用 MATLAB 2013a 编写。各算法种群个数设置为 40,最大迭代次数(Max Cycle)为 10 000,算法评价次数为 $40 \times 10\ 000$ 次,算法终止条件为优化结果满足最小误差精度 $\text{ErrGoal}=10 - 50$ 或者达到最大迭代次数。

2.2 仿真结果及分析

通过仿真实验,表 2 列出了 5 种算法在上述参数设置条件下独立运行 30 次的仿真结果,其中 M 为算法求得最优值的平均值, S 为标准差。从表 2 中可以看出,面对多个测试函数直接对 JADE 算法的优化性能进行评估显得十分繁琐,且不能对其性能在整体上进行比较评价。

表 2 5 种算法对于 36 种测试函数的优化结果

Tab.2 Optimization results of five algorithms for optimizing thirty-six test functions

Sta.	F	CLPSO	ABC	CK	GWO	JADE	F	CLPSO	ABC	CK	GWO	JADE
M	1	7.46E-15	3.22E-14	3.85E-02	7.11E-15	6.21E-02	19	1.48E-32	6.73E-03	1.87E-20	6.43E-11	0.00E-00
S	1	1.08E-15	3.84E-15	2.11E-01	0.00E-00	2.36E-01	19	8.10E-32	8.22E-03	4.64E-20	2.25E-10	0.00E-00
M	2	2.81E-05	7.54E-16	1.64E-00	3.08E-52	3.48E-16	20	3.09E-52	1.29E-16	1.23E-25	8.56E-53	1.94E-54
S	2	1.52E-05	1.99E-16	5.81E-01	1.69E-51	4.41E-16	20	1.69E-51	1.47E-16	2.84E-25	4.69E-52	1.06E-53
M	3	-106.76	-106.76	-106.76	-106.76	-106.76	21	-1.82E-00	-1.82E-00	-1.82E-00	-1.82E-00	-1.82E-00
S	3	3.08E-14	3.39E-14	5.05E-14	1.26E-08	4.34E-14	21	9.034E-16	7.12E-16	9.03E-16	1.39E-10	9.03E-16
M	4	-1.03E-00	-1.03E-00	-1.03E-00	-1.03E-00	-1.03E-00	22	-4.69E-00	-4.69E-00	-4.69E-00	-4.65E-00	-4.69E-00
S	4	4.51E-16	4.46E-16	4.52E-16	7.05E-13	4.52E-16	22	7.86E-16	7.56E-16	6.37E-14	2.94E-08	7.53E-03
M	5	0.98E-03	0.38E-01	1.17E-07	1.56E-00	0.00E-00	23	-9.66E-00	-9.66E-00	-9.32E-00	-8.29E-00	-9.65E-00
S	5	1.38E-03	5.78E-02	4.06E-07	1.84E-04	0.00E-00	23	3.30E-16	8.73E-16	1.53E-01	4.01E-06	1.85E-02
M	6	-2.06E-00	-2.06E-00	-2.06E-00	-2.06E-00	-2.06E-00	24	-0.96E-00	-0.96E-00	-0.96E-00	-0.96E-00	-0.96E-00
S	6	9.03E-16	9.03E-16	1.08E-15	4.99E-12	9.03E-16	24	0.00E-00	0.00E-00	2.62E-16	1.43E-11	0.00E-00
M	7	1.95E-30	3.55E-03	6.52E-09	1.23E-10	0.00E-00	25	2.12E-03	3.71E-03	1.91E-05	5.15E-12	8.31E-08
S	7	1.07E-29	6.50E-03	2.59E-08	3.48E-10	0.00E-00	25	6.92E-04	4.458E-05	6.71E-05	2.09E-12	1.47E-07
M	8	0.40E-00	1.28E-14	6.67E-01	6.67E-01	6.67E-01	26	3.21E-04	4.61E-03	4.38E-04	1.30E-05	6.74E-15
S	8	0.28E-00	3.45E-14	9.53E-15	9.75E-10	0.00E-00	26	2.49E-04	3.47E-03	4.79E-04	5.89E-06	1.46E-14
M	9	-0.03E-00	-0.03E-00	-1.00E-00	-1.00E-00	-6.67E-02	27	3.12E-03	0.11E-00	6.03E-04	1.41E-05	7.02E-04
S	9	0.18E-00	0.18E-00	4.32E-13	8.33E-11	2.54E-01	27	8.27E-04	1.98E-02	2.66E-03	2.01E-05	2.35E-04
M	10	2.34E-52	1.68E-16	3.31E-52	2.78E-52	2.42E-52	28	3.41E-00	0.27E-00	3.69E-01	2.71E+01	0.93E-00
S	10	1.28E-51	5.26E-17	1.81E-51	1.52E-51	1.33E-51	28	5.24E-00	0.52E-00	8.52E-01	3.87E-03	1.71E-00
M	11	0.90E-00	1.00E-00	4.79E+01	4.80E+01	4.19E-01	29	2.51E-40	4.75E-16	4.75E-46	2.39E-52	3.10E-52
S	11	0.30E-00	4.91E-16	1.18E+01	5.08E-02	7.17E-01	29	2.13E-40	6.23E-17	1.89E-45	1.314E-51	1.70E-51
M	12	0.99E-00	0.99E-00	0.99E-00	1.07E+01	0.99E-00	30	0.29E-00	0.29E-00	0.29E-00	0.29E-00	0.29E-00
S	12	0.00E-00	0.00E-00	3.72E-16	3.64E-14	0.00E-00	30	5.55E-17	7.13E-10	8.13E-09	6.47E-11	5.65E-17
M	13	3.00E-00	3.00E-00	3.00E-00	3.00E-00	3.00E-00	31	3.86E-02	0.24E-00	7.73E-02	9.72E-03	4.27E-03
S	13	1.24E-15	1.34E-04	1.77E-15	4.428E-10	1.77E-15	31	7.48E-03	4.78E-02	1.84E-02	8.26E-15	1.42E-02
M	14	6.44E-02	6.45E-02	6.45E-02	6.45E-02	6.45E-02	32	0.49E-00	6.61E-03	1.39E-01	2.74E-52	1.71E-02
S	14	5.65E-17	6.67E-17	5.65E-17	5.71E-13	5.54E-17	32	8.42E-02	1.89E-02	2.61E-01	1.50E-51	2.67E-02
M	15	0.00E-00	2.47E-04	0.00E-00	2.81E-00	4.26E-03	33	-7.89E-03	-2.58E-03	-1.02E+01	-1.02E+01	-3.01E-04
S	15	0.00E-00	1.35E-03	0.00E-00	1.12E-04	8.63E-03	33	1.38E-02	1.76E-18	2.06E-13	3.14E-07	2.76E-19
M	16	-3.04E-00	-3.04E-00	-3.04E-00	-2.98E-00	-3.02E-00	34	5.95E-41	4.56E-16	1.57E-46	1.67E-52	2.53E-52
S	16	1.37E-15	1.36E-15	1.36E-15	1.13E-06	3.06E-02	34	4.42E-41	7.84E-17	1.64E-46	9.24E-52	1.39E-51
M	17	2.63E-32	3.64E-18	1.72E-13	2.53E-10	2.10E-31	35	1.46E-42	4.21E-16	7.134E-48	2.99E-52	3.22E-52
S	17	1.44E-31	3.18E-18	2.26E-13	6.46E-10	3.55E-31	35	1.08E-42	9.26E-17	9.46E-48	1.64E-51	1.76E-51
M	18	-1.49E-00	-1.31E-00	-1.49E-00	-1.49E-00	-1.37E-00	36	1.29E-52	3.08E-17	2.83E-52	5.23E-53	2.46E-52
S	18	6.78E-16	2.68E-01	1.74E-05	4.21E-09	0.24E-00	36	7.10E-52	1.03E-17	1.55E-51	2.87E-52	1.35E-51

因此首先利用 Wilcoxon 符号秩检验对其性能进行分析,设置显著水平为 0.05。表 3 给出了 30 次独立运行后算法对 36 个测试函数优化后的最优值

的 Wilcoxon 符号秩检验结果。其中 win 表示两种算法比较中性能较优的算法,“+”表示 JADE 算法优于相比较的算法,“-”表示 JADE 算法差于相比

较的算法,“=”表示两种算法性能并无明显差异。

从表 3 中通过统计“+/=/-”值可以看出,对于 36 个不同的测试函数,JADE 算法整体在算法收敛速度及搜索精度方面明显优于与之比较的 CLPSO 算法、ABC 算法、CK 算法以及 GWO 算法。

然而 Wilcoxon 符号秩检验仅仅适用于两两之间的比较,而对于同时评价并按照优化性能的好坏对这 5 种算法的性能进行比较排序,则需要利用 Quade 检验这一方法完成。

表 3 Wilcoxon 符号秩检验结果
Tab.3 The results of Wilcoxon sign rank test

F	CLPSO vs JADE				ABC vs JADE				CK vs JADE				GWO vs JADE			
	p-value	R ⁺	R ⁻	win	p-value	R ⁺	R ⁻	win	p-value	R ⁺	R ⁻	win	p-value	R ⁺	R ⁻	win
1	8.6695E-03	171	39	+	3.1731E-04	406	59	+	1.5247E-01	35	85	=	1.0623E-02	153	37	+
2	1.7344E-06	465	0	+	4.8969E-04	402	63	+	1.7344E-06	465	0	+	3.1468E-04	1	152	-
3	5.3373E-07	351	0	+	4.5928E-06	231	0	+	1.5165E-06	465	0	+	9.0256E-07	465	0	+
4	1.0000E-00	0	0	=	1.0000E-00	1	0	=	1.0000E-00	1	0	=	3.3221E-07	465	0	+
5	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.7279E-06	465	0	+
6	1.0000E-00	0	0	=	1.0000E-00	0	0	=	3.9063E-03	45	0	+	1.2350E-06	465	0	+
7	1.0000E-00	1	0	=	1.7344E-06	465	0	+	1.7344E-06	465	0	+	9.0382E-07	465	0	+
8	3.8542E-03	92	373	-	1.7235E-06	0	465	-	4.8828E-04	78	0	+	1.7344E-06	465	0	+
9	1.0000E-00	0	0	=	1.0000E-00	1	0	=	1.7279E-06	465	0	+	7.8866E-07	465	0	+
10	1.0000E-00	1	2	=	1.7344E-06	465	0	+	1.0000E-00	2	1	=	1.0000E-00	2	1	=
11	6.9564E-04	243	33	+	2.5149E-04	409	56	+	1.7344E-06	465	0	+	1.7344E-06	465	0	+
12	1.0000E-00	0	0	=	1.0000E-00	1	0	=	1.0615E-06	465	0	+	4.1427E-07	465	0	+
13	9.1553E-04	114	7	+	1.1279E-04	285	15	+	1.7389E-05	300	0	+	1.6976E-06	465	0	+
14	2.5135E-01	70	120	=	3.8770E-01	52	26	=	4.8828E-04	78	0	+	1.6065E-06	465	0	+
15	7.8125E-03	0	36	-	1.3113E-01	39	97	=	7.8125E-03	0	36	-	1.7344E-06	465	0	+
16	1.0000E-00	0	0	=	4.8828E-04	0	78	+	4.8828E-04	0	78	-	1.7344E-06	465	0	+
17	3.9063E-02	5	40	-	1.7344E-06	465	0	+	1.7344E-06	465	0	+	2.6136E-07	465	0	+
18	1.5625E-02	0	28	-	1.1138E-03	391	74	+	3.7094E-01	276	189	=	3.6920E-01	276	189	=
19	1.0000E-00	1	0	=	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.0144E-07	465	0	+
20	1.0000E-00	2	1	=	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.0000E-00	2	1	=
21	1.0000E-00	1	0	=	3.4142E-07	351	0	+	1.0000E-00	0	0	=	1.9773E-07	465	0	+
22	2.9602E-02	98	22	+	3.3667E-01	70	35	=	9.5156E-05	351	27	+	1.6106E-06	465	0	+
23	3.0092E-01	78	132	=	1.4063E-01	84	170	=	1.7344E-06	465	0	+	1.7105E-06	465	0	+
24	1.0000E-00	1	0	=	1.0000E-00	1	0	=	1.9531E-03	55	0	+	9.9323E-07	465	0	+
25	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.7344E-06	0	465	-
26	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.7344E-06	465	0	+
27	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.7344E-06	0	465	-
28	8.2167E-03	361	104	+	3.7094E-01	276	189	=	1.6503E-01	300	165	=	1.7344E-06	465	0	+
29	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.0000E-00	1	2	=
30	1.0000E-00	1	0	=	8.0914E-07	465	0	+	1.7344E-06	465	0	+	1.7084E-06	465	0	+
31	1.1748E-02	355	110	+	1.7344E-06	465	0	+	6.9838E-06	451	14	+	1.7344E-06	0	465	-
32	1.7344E-06	465	0	+	3.0861E-01	183	282	=	3.8723E-02	333	132	+	1.7344E-06	0	465	-
33	1.7344E-06	0	465	-	4.3205E-08	0	465	+	1.7344E-06	465	0	+	1.2826E-06	0	465	-
34	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.0000E-00	1	2	=
35	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.7344E-06	465	0	+	1.0000E-00	1	2	=
36	1.0000E-00	1	2	=	1.7344E-06	465	0	+	1.0000E-00	2	1	=	1.0000E-00	2	0	=
+/=/-	16/15/5				24/11/1				27/7/2				23/7/6			

表 4 给出了 30 次独立运行后各算法对 36 个测试函数优化后的平均最优值及平均方差的 Quade 检验结果。从表 4 中可以看出,在收敛速度及搜索

精度方面,JADE 算法性能最优,其次分别为 CLPSO 算法、CK 算法、GWO 算法,ABC 算法相对性能最差;而在搜索精度稳定性方面,GWO 算法性能最

优,其次分别为 JADE 算法、CLPSO 算法、ABC 算法,CK 算法在这一方面表现最差。

表 4 Quade 检验结果

Tab.4 Results of Quade test

Ranks		CLPSO	ABC	CK	GWO	JADE
R_{mean}	S_j 排序	2.7245E+00	3.3904E+00	3.0818E+00	3.0901E+00	2.7132E+00
		2	5	3	4	1
	T_j 排序	-1.8350E+02	2.6000E+02	5.4500E+01	6.0000E+01	-1.9100E+02
		2	5	3	4	1
R_{Std}	S_j 排序	2.9324E+00	3.1224E+00	3.4227E+00	2.6517E+00	2.8709E+00
		3	4	5	1	2
	T_j 排序	-4.5000E+01	8.1500E+01	2.8150E+02	-2.3200E+02	-8.6000E+01
		3	4	5	1	2

3 结语

随着进化计算的迅猛发展,作为其重要分支的生物启发式优化算法,其种类、性能也随之得到了前所未有的突破,而与之对应的测试函数的种类也逐渐增多。在此情况下,如何准确地、方便地、有效地评价各类优化算法的优化性能成为一个新的问题,对于此,本文提出了利用统计学中非参数统计的 2 类典型方法:Wilcoxon 符号秩检验和 Quade 检验,分别从算法成对比较和多组比较的角度出发,对 5 种不同的生物启发式优化算法在 36 种不同类型的测试函数下的性能进行了比较分析,实验结果表明, JADE 算法相比于其它四种算法,在收敛速度及搜索精度方面表现最优,而 GWO 算法在精度稳定性方面相比于其它 4 种算法表现出较优的性能。

本文为评价生物启发式优化算法的性能提供了新的思路,后续在开展对其算法性能研究的过程中也可利用其它非参数统计方法对不同优化算法的各类优化性能进行评价分析,例如符号检验、Friedman 检验等方法。

参考文献(References):

- [1] Chiong R. Nature-Inspired Algorithms for Optimization[M]. Heidelberg: Springer-Verlag Berlin Heidelberg, 2009.
- [2] Holland J H. Genetic Algorithms[J]. Sci Am 1992, 267:66-72.
- [3] Kennedy J, Eberhart R. Particle Swarm Optimization [C]//IEEE International Conference in Neural Networks, Perth, WA: IEEE, 1995: 1942-1948.
- [4] Dorigo M, Birattari M, Stutzle T. Ant colony Optimization [J]. Computation Intelligence Magazine, 2006, 1:28-39.
- [5] 陈义雄,梁昔明,黄亚飞. 基于佳点集构造的改进量子粒子群优化算法[J].中南大学学报:自然科学版, 2013, 44(4): 1409-1415.

CHEN Yixiong, LIANG Ximing, HUANG Yafei, Improved Quantum Particle Swarm Optimization Based on Good-point set [J]. Journal of Central South University: Science and Technology, 2013, 44 (4): 1409-1415.

- [6] Basturk B, Karaboga D. An Artificial Bee Colony (ABC) Algorithm for Numeric Function Optimization[C]//IEEE Swarm Intelligence Symposium, Indiana, USA: IEEE, 2006: 12-16.
- [7] Mirjalili S, Mirjalili S M, Lewis A. Grey Wolf Optimizer[J]. Advances in Engineering Software, 2014, 69:46-61.
- [8] Yang X S, Deb S. Cuckoo Search: Recent Advances and Applications [J]. Neural Computation & Application, 2014, 24:169-174.
- [9] Sheskin D J. Handbook of Parametric and Nonparametric Statistical Procedures[M].4th ed.[s.l.]:Chapman & Hall/CRC, 2006.
- [10] Quade D. Using Weighted Rankings in the Analysis of Complete Blocks with Additive Block Effects[J], Journal of the American Statistical Association, 1979, 74:680-683.
- [11] Jamil M, Yang X S. A literature Survey of Benchmark Functions for Global Optimization Problems [J].International Journal of Mathematical Modelling and Numerical Optimisation,2013, 4(2): 150-194.
- [12] Liang J J, Qin A K, Suganthan P N, et al. Comprehensive learning Particle Swarm Optimizer for Global Optimization of Multimodal Functions [J]. IEEE Transaction Evolution Computation, 2006, 10 (3), 281-295.
- [13] Zhang J, Sanderson A C. JADE: Adaptive Differential Evolution with Optional External Archive[J].IEEE Transaction Evolution Computation, 2009, 13 (5): 945-958.

(编辑:姚树峰)