

采用协同演化算法解决 大规模数值优化问题的研究

Large-Scale Numerical Optimization using Cooperative
Coevolutionary Algorithms

陈文祥

导师：唐珂 教授

中国科学技术大学



纲要

1 研究背景

2 本论文主要工作

- 提出变量相关性定义
- 分组策略对协同演化影响的研究
- 将变量相关性学习引入协同演化
- 提升变量相关性学习的效率

3 总结

协同演化算法(Cooperative Coevolutionary Algorithms)

协同演化算法

- 相关性信息 **已知** → 解决大规模问题有潜力的方法
- 实际应用中相关性信息未知(黑盒优化)
- 极大限制协同演化的应用

相关性学习的发展情况

遗传算法中的Linkage Learning

- 已经有近20年的研究基础: 如 David E. Goldberg, Georges Harik, Dirk Thierens...
- **局限**: 基于二进制串表示

相关性学习的发展情况

遗传算法中的Linkage Learning

- 已经有近20年的研究基础: 如 David E. Goldberg, Georges Harik, Dirk Thierens...
- **局限**: 基于二进制串表示

实数表示的变量相关性学习

- 主流的数值优化算法都以实数表示, 如PSO, DE, CMA-ES...
- 鲜有工作研究基于实数的变量相关性

提出变量相关性的数学定义

变量相关性定义[1]

任意两个维度*i*和*j*的决策变量是**相关**的, 当且仅当存在一个候选解 \vec{x} , 其第*i*th和第*j*th变量在替换为 x'_i 和 x'_j 后, 满足方程:

$$\exists \vec{x}, x'_i, x'_j: f(\boxed{x_1, \dots}, \boxed{x_i}, \boxed{}, \boxed{x_j}, \boxed{\dots, x_n}) < f(\boxed{x_1, \dots}, \boxed{x'_i}, \boxed{}, \boxed{x_j}, \boxed{\dots, x_n}) \wedge \\ f(\boxed{x_1, \dots}, \boxed{x_i}, \boxed{}, \boxed{x'_j}, \boxed{\dots, x_n}) > f(\boxed{x_1, \dots}, \boxed{x'_i}, \boxed{}, \boxed{x'_j}, \boxed{\dots, x_n})$$

本部分研究动机

背景：有大量研究致力于在协同演化算法中设计合理的分组策略，尽量使得子问题之间相互独立，子问题内的变量紧密相关

现有研究的不足

没有工作研究分组策略的对协同演化的影响究竟有多大：

- 理想分组策略下，算法能够达到什么样的性能？
- 是不是在理想分组策略下的协同演化算法性能就一定最好？
- 是不是越接近理想分组，算法性能就呈现正比例的提高？

本部分工作将通过全面的实验研究回答上述问题

实验结果

	Prob.	先验分组信息所占比例						
		0%	1%	3%	5%	7%	9%	100%
f ₁₄	mean	5.97e+08	4.41e+08	2.19e+08	4.97e+07	1.75e+07	1.14e+07	5.78e+06
	std	5.98e+07	2.68e+07	2.18e+07	7.66e+06	2.85e+06	2.77e+06	7.16e+05
	group	1000	745	345	120	60	30	20
f ₁₅	mean	1.42e+04	1.18e+04	7.14e+03	3.03e+03	1.84e+03	1.66e+03	1.60e+03
	std	4.40e+02	4.42e+02	3.52e+02	2.03e+02	1.22e+02	7.66e+01	9.98e+01
	group	1000	764	346	122	46	30	20
f ₁₆	mean	3.95e+02	4.01e+02	1.94e+02	3.90e+01	1.72e+01	5.36e+00	4.06e-01
	std	1.36e+00	2.61e+00	2.27e+01	4.64e+00	2.84e+00	1.99e+00	2.54e-01
	group	1000	773	316	122	53	28	20
f ₁₇	mean	1.03e+06	8.26e+05	4.42e+05	1.17e+05	2.98e+04	8.13e+03	6.15e+01
	std	1.01e+05	8.07e+04	5.53e+04	1.97e+04	7.70e+03	2.82e+03	7.63e+01
	group	1000	756	305	108	51	26	20

结果分析

随着先验分组信息的增加，分组策略趋于合理，协同演化算法性能的确有显著的提升

本部分研究动机

目前不存在有效的分组策略

几种现有典型分组策略的弊端：

- 所有变量一齐优化 → 复杂度高，不适合求解大规模问题
- 对每维变量都单独考虑 → 未考虑相关性，局部最优[3]
- 随机分组 → 将相关变量都分于一组的概率低[5, 6]

本部分思路

- 提出一种二阶段协同演化算法
- 主动学习变量相关性
- 尽量接近理想分组，从而显著提高算法性能

测试函数的特征

	Function	Sep	Multi modal	Groups	
				real	found
f ₁	Shifted Elliptic Function	Yes	No	1000	1000
f ₂	Shifted Rastrigin' s Function	Yes	Yes	1000	1000
f ₃	Shifted Ackley' s Function	Yes	Yes	1000	969
f ₄	Single-group Shifted 50-rotated Elliptic Function	No	No	951	963
f ₅	Single-group Shifted 50-rotated Rastrigin' s Function	No	Yes	951	952
f ₆	Single-group Shifted 50-rotated Ackley' s Function	No	Yes	951	921
f ₇	Single-group Shifted 50-dimensional Schwefel' s	No	No	951	952
f ₈	Single-group Shifted 50-dimensional Rosenbrock' s	No	Yes	951	1000
f ₉	10-group Shifted 50-rotated Elliptic Function	No	No	510	627
f ₁₀	10-group Shifted 50-rotated Rastrigin Function	No	Yes	510	516
f ₁₁	10-group Shifted 50-rotated Ackley Function	No	Yes	510	501
f ₁₂	10-group Shifted 50-dimensional Schwefel' s	No	No	510	522
f ₁₃	10-group Shifted 50-dimensional Rosenbrock' s	No	Yes	510	1000
f ₁₄	20-group Shifted 50-rotated Elliptic Function	No	No	20	232
f ₁₅	20-group Shifted 50-rotated Rastrigin' s Function	No	Yes	20	37
f ₁₆	20-group Shifted 50-rotated Ackley Function	No	Yes	20	39
f ₁₇	20-group Shifted 50-dimensional Schwefel' s Function	No	No	20	42
f ₁₈	20-group Shifted 50-dimensional Rosenbrock	No	Yes	20	1000
f ₁₉	Shifted Schwefel' s Function 1.2	No	No	1	1
f ₂₀	Shifted Rosenbrock' s Function	No	Yes	1	1000

- 多数测试函数上，CCVIL通过学习变量相关性，得到接近理想的分组策略

实验研究结果

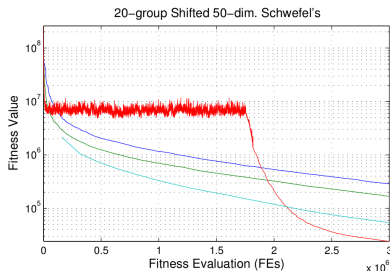
	CCVIL		DECC-G		MLCC		R ₁	Naive JADE		R ₂
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev		Mean	Std Dev	
f ₁	1.55e-17	7.75e-17	2.93e-07	8.62e-08	1.53e-27	7.66e-27	-	1.57e+04	1.38e+04	W
f ₂	6.71e-09	2.31e-08	1.31e+03	3.24e+01	5.55e-01	2.20e+00	W	7.66e+03	9.67e+01	W
f ₃	7.52e-11	6.58e-11	1.39e+00	9.59e-02	9.86e-13	3.69e-12	L	4.52e+00	2.41e-01	W
f ₄	5.00e+12	3.38e+12	9.62e+12	3.43e+12	1.70e+13	5.38e+12	W	6.14e+09	3.81e+09	L
f ₅	1.76e+08	6.47e+07	2.63e+08	8.44e+07	3.84e+08	6.93e+07	W	1.35e+08	1.21e+07	L
f ₆	2.94e+05	6.09e+05	4.96e+06	8.02e+05	1.62e+07	4.97e+06	W	1.94e+01	1.79e-02	-
f ₇	8.00e+08	2.48e+09	1.63e+08	1.38e+08	6.89e+05	7.36e+05	-	2.99e+01	3.30e+01	-
f ₈	6.50e+07	3.07e+07	6.44e+07	2.89e+07	4.38e+07	3.45e+07	-	1.19e+04	4.92e+03	L
f ₉	6.66e+07	1.60e+07	3.21e+08	3.39e+07	1.23e+08	1.33e+07	W	2.70e+07	2.08e+06	L
f ₁₀	1.28e+03	7.95e+01	1.06e+04	2.93e+02	3.43e+03	8.72e+02	W	8.50e+03	2.30e+02	W
f ₁₁	3.48e+00	1.91e+00	2.34e+01	1.79e+00	1.98e+02	6.45e-01	W	9.29e+01	9.66e+00	W
f ₁₂	8.95e+03	5.39e+03	8.93e+04	6.90e+03	3.48e+04	4.91e+03	W	6.21e+03	1.34e+03	-
f ₁₃	5.72e+02	2.55e+02	5.12e+03	3.95e+03	2.08e+03	7.26e+02	W	1.87e+03	1.11e+03	W
f ₁₄	1.74e+08	2.68e+07	8.08e+08	6.06e+07	3.16e+08	2.78e+07	W	1.00e+08	8.84e+06	L
f ₁₅	2.65e+03	9.34e+01	1.22e+04	9.10e+02	7.10e+03	1.34e+03	W	3.65e+03	1.09e+03	W
f ₁₆	7.18e+00	2.23e+00	7.66e+01	8.14e+00	3.77e+02	4.71e+01	W	2.09e+02	2.01e+01	W
f ₁₇	2.13e+04	9.16e+03	2.87e+05	1.97e+04	1.59e+05	1.43e+04	W	7.78e+04	5.87e+03	W
f ₁₈	1.33e+04	1.00e+04	2.46e+04	1.05e+04	7.09e+03	4.77e+03	-	3.71e+03	9.58e+02	L
f ₁₉	3.52e+05	2.04e+04	1.11e+06	5.00e+04	1.36e+06	7.31e+04	W	3.48e+05	1.67e+04	-
f ₂₀	1.11e+03	3.04e+02	4.06e+03	3.66e+02	2.05e+03	1.79e+02	W	2.06e+03	2.01e+02	W

● “W”:统计显著优; “L”:统计显著劣; “-”:无统计显著差别

● 每个算法 3×10^6 次采样机会, 独立运行25次

实验研究结果总结

- CCVIL是目前世界上已发表的**性能最好**的协同演化算法



典型函数收敛曲线

- 红色为**CCVIL**
- 虽然花费较多计算时间用于学习，然而一旦进入优化阶段，算法求解速度**显著快于其他算法**

研究动机

CCVIL变量相关性学习机制的不足

- 花费大量计算代价(多达60%)
- 仍然无法学习到所有的变量相关性

本部分研究目的

提高学习效率，使用尽可能少的计算代价达到理想分组

奥卡姆剃刀原理的应用

奥卡姆剃刀[4]

切勿浪费较多东西，去做‘用较少的东西，同样可以做好的事情’

本部分思路

在学习阶段采用最简洁的搜索算法“随机采样”，代替复杂的JADE，新算法记为（RS-LP）

实验研究结果-分组情况

RS-LP与原相关性学习方法(Ori-LS)分组情况的比较

Benchmark	f ₁	f ₂	f ₃	f ₄	f ₅	f ₆	f ₇	f ₈	f ₉	f ₁₀
Ori-LS found	1000	1000	969	963	952	921	952	1000	627	516
RS-LP found	1000	1000	397	951	951	394	951	1000	510	510
Real group num.	1000	1000	951	951	951	951	951	951	510	510

Benchmark	f ₁₁	f ₁₂	f ₁₃	f ₁₄	f ₁₅	f ₁₆	f ₁₇	f ₁₈	f ₁₉	f ₂₀
Ori-LS found	501	522	1000	232	37	39	42	1000	1	1000
RS-LP found	286	510	1000	21	20	20	20	1000	1	1000
Real group num.	510	510	510	20	20	20	20	20	1	1

结果分析

- 简洁的方法却获得更高的相关性学习效率
- 大部分函数上已经能够达到理想分组

实验研究结果-算法整体性能

	Ori-LS-CCVIL		R ₁	RS-LP-CCVIL		R ₂	MA-SW-Chains	
f ₁	1.55e-17	7.75e-17	-	0.00e+00	0.00e+00	W	2.10e-14	1.99e-14
f ₂	6.71e-09	2.31e-08	L	2.97e-08	7.29e-08	W	8.10e+02	5.88e+01
f ₃	7.52e-11	6.58e-11	L	7.07e-01	2.69e-01	L	7.28e-13	3.40e-13
f ₄	5.00e+12	3.38e+12	W	3.66e+12	7.55e+12	-	3.53e+11	3.12e+10
f ₅	1.76e+08	6.47e+07	-	2.10e+08	1.39e+08	-	1.68e+08	1.04e+08
f ₆	2.94e+05	6.09e+05	L	6.81e+05	5.87e+05	L	8.14e+04	2.84e+05
f ₇	8.00e+08	2.48e+09	L	9.22e+08	4.61e+09	L	1.03e+02	8.70e+01
f ₈	6.50e+07	3.07e+07	W	3.99e+07	3.48e+07	L	1.41e+07	3.68e+07
f ₉	6.66e+07	1.60e+07	W	5.88e+06	9.87e+05	W	1.41e+07	1.15e+06
f ₁₀	1.28e+03	7.95e+01	W	9.67e+02	4.15e+01	W	2.06e+03	1.40e+02
f ₁₁	3.48e+00	1.91e+00	-	2.44e+00	1.66e+00	W	3.77e+01	6.85e+00
f ₁₂	8.95e+03	5.39e+03	W	6.62e+01	1.73e+02	L	3.62e-06	5.92e-07
f ₁₃	5.72e+02	2.55e+02	-	5.57e+02	8.05e+01	W	1.25e+03	5.72e+02
f ₁₄	1.74e+08	2.68e+07	W	1.80e+07	2.21e+06	W	3.10e+07	2.19e+06
f ₁₅	2.65e+03	9.34e+01	W	2.20e+03	7.22e+01	W	2.72e+03	1.22e+02
f ₁₆	7.18e+00	2.23e+00	W	3.28e+00	1.89e+00	W	1.01e+02	1.45e+01
f ₁₇	2.13e+04	9.16e+03	W	1.49e+02	1.24e+02	L	1.24e+00	1.25e-01
f ₁₈	1.33e+04	1.00e+04	W	2.43e+03	3.54e+03	L	1.30e+03	4.36e+02
f ₁₉	3.52e+05	2.04e+04	-	3.47e+05	1.77e+04	L	2.85e+05	1.78e+04
f ₂₀	1.11e+03	3.04e+02	-	1.18e+03	2.59e+02	-	1.07e+03	7.29e+01

- R₁: 和原CCVIL算法 (Ori-LS-CCVIL) 的统计测试比较结果; R₂和测试集的世界冠军MA-SW-Chains[2]的统计测试比较结果;
- 比原CCVIL算法 (Ori-LS-CCVIL) 有全面的、统计显著的性能提升
- 性能甚至超过测试集上的世界冠军MA-SW-Chains

从连系学习的方法中寻找灵感

研究动机

- 连系学习具备丰富的研究基础
- 存在可以借鉴的思想

进一步提升学习效率的三种策略

- ① 使用随机游走代替随机采样
- ② 推广变量相关性的定义
- ③ 利用折半查找将相关性学习的复杂度 $O(N^2)$ 降至 $O(N \times \log(N))$

实验研究结果

	RS-LP	RS-GenDef-LP	RS-GenDef-BinSearch
f ₁₄	2.98e+06 (60)	2.99e+05 (20)	4.41e+04 (20)
f ₁₅	5.16e+05 (20)	2.99e+05 (20)	4.41e+04 (20)
f ₁₆	4.50e+05 (20)	2.99e+05 (20)	4.41e+04 (20)
f ₁₇	2.98e+06 (73)	2.99e+05 (20)	4.41e+04 (20)
f ₁₈	2.94e+06 (976)	2.99e+06 (241)	5.55e+04 (20)
f ₁₉	2.84e+05 (1)	4.00e+03 (1)	4.06e+04 (1)
f ₂₀	2.94e+06 (974)	2.99e+06 (217)	5.67e+04 (1)
	RW-LP	RW-GenDef-LP	RW-GenDef-BinSearch
f ₁₄	1.55e+06 (20)	2.67e+05 (20)	4.41e+04 (20)
f ₁₅	7.94e+05 (20)	2.67e+05 (20)	4.41e+04 (20)
f ₁₆	4.69e+05 (20)	2.67e+05 (20)	4.41e+04 (20)
f ₁₇	7.75e+05 (20)	2.67e+05 (20)	4.41e+04 (20)
f ₁₈	2.99e+06 (873)	1.67e+06 (1)	5.47e+04 (20)
f ₁₉	1.13e+04 (1)	4.00e+03 (1)	4.06e+04 (1)
f ₂₀	2.98e+06 (871)	1.39e+06 (1)	5.57e+04 (1)

注：左边数字为学习阶段收敛时消耗的适应度评估次数; 右边括号内数字是收敛时分组的数目

结果分析

- RW-GenDef-BinSearch只花费RS-LP1%~10%的计算代价, 就能检测到所有相关变量, 达到理想分组
- 新的相关性学习策略(如RW-GenDef-BinSearch)比测试集世界冠军MA-SW-Chains会有更大幅性能提升

论文发表

Results Topic=(Large-Scale Global Optimization using Cooperative Coevolution with Variable Interaction Learning) ScientificWebPlus View Web Results >>
Timespan=All Years.

Results: 2 Page 1 of 1 Go Sort by: Publication Date

Refine Results
Search within results for Search

General Categories Refine
SCIENCE & TECHNOLOGY (2)

1. Title: Large-Scale Global Optimization Using Cooperative Coevolution with Variable Interaction Learning
Author(s): Chen, WX, Weise, T, Yang, ZY, et al.
Conference Information: 11th International Conference on Parallel Problem Solving from Nature, Date: SEP 11-15, 2010 AGU Univ Krakow POLAND
Source: PARALLEL PROBLEM SOLVING FROM NATURE-PPSN XI, PT II Volume: 6239 Pages: 300-309 Published: 2010

- 相关论文已经发表于演化计算领域最好会议 PPSN' 2010(Parallel Problem Solving From Nature XI)
- 论文被SCI (000286453000031), EI (20104513369063)和ISTP 权威检索
- 目前已经被它引一次，引用者正是协同演化发明者所在的美国研究小组

部分工作的获奖情况

“挑战杯”安徽省大学生课外学术科技作品竞赛特等奖



论文贡献

- ① 提出变量相关性定义(已发表)
 - 奠定变量相关性学习的理论基础
 - 填补实数表示相关性分析的空白
- ② 研究分组策略对协同演化影响(待发表)
 - 回答悬而未决的关键问题
- ③ 将变量相关性学习引入协同演化(已发表)
 - 提出一种性能国际领先的协同演化算法
- ④ 提升变量相关性学习的效率(待发表)
 - 提出性能显著优于测试集世界冠军的若干协同演化算法

结束

欢迎老师和同学提供指导和建议！

参考文献 I



Wenxiang Chen, Thomas Weise, Zhenyu Yang, and Ke Tang.

Large-Scale Global Optimization Using Cooperative Coevolution with Variable Interaction Learning.
In Robert Schaefer, Carlos Cotta, Joanna Kolodziej, and Günter Rudolph, editors, Parallel Problem Solving from Nature – PPSN XI, volume 6239 of Lecture Notes in Computer Science, pages 300 – 309. Springer Berlin / Heidelberg, 2010.
10.1007/978-3-642-15871-1_31.



Daniel Molina, Manuel Lozano, and Francisco Herrera.

Ma-sw-chains: Memetic algorithm based on local search chains for large scale continuous global optimization.
In IEEE Congress on Evolutionary Computation, pages 1 – 8. IEEE, 2010.



Mitchell A. Potter.

The Design and Analysis of a Computational Model of Cooperative Coevolution.
PhD thesis, George Mason University, 1997.



C. E. Rasmussen and Z. Ghahramani.

Occam’ s razor.

In T. K. Leen, T. G. Dietterich, and V. Tresp, editors, Advances in Neural Information Processing Systems 13, pages 294 – 300. MIT Press, Cambridge, MA, 2001.



Zhenyu Yang, Ke Tang, and Xin Yao.

Large Scale Evolutionary Optimization Using Cooperative Coevolution.
Information Sciences, 178(15):2985 – 2999, 2008.

参考文献 II



Zhenyu Yang, Ke Tang, and Xin Yao.

Multilevel cooperative coevolution for large scale optimization.

In IEEE Congress on Evolutionary Computation, pages 1663 – 1670. IEEE Press, 2008.