

大规模数值优化中变量相关性学习的研究

自然科学类学术论文

陈文祥

自然计算与应用实验室(NICAL)
计算机科学与技术学院
中国科学技术大学

2007级本科生



概要

1 引言

- 黑盒数值优化
- 大规模数值优化

2 协同演化

- 分而治之:降低问题复杂度
- 相关性难题
- 策略选择困境

3 新颖的算法框架

- 变量相关性学习
- 算法框架的性能

4 结论

- 实验总结
- 作品意义

5 参考文献

数值优化

Definition

搜索给定函数的最大/最小值点

Example

$$\min F_{\text{sphere}}(\mathbf{x}) = \sum_{i=1}^D x_i^2 \quad (1)$$

最优解为 $F(0, \dots, 0) = 0$.

黑盒问题

黑盒优化问题

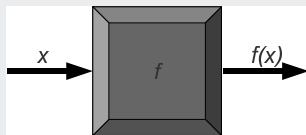


Figure: Illustration of BlackBox

大规模问题

大规模问题

- 问题本身有成千上万的决策变量的数目
- 本作品以维数1000作为例子进行研究

现实生活中有海量的工业生产和工程项目的问题都可以抽象成大规模黑盒数值优化问题

生活中的大规模问题

Example

- 金融衍生产品（期货、掉期等）的定价
- 民航航路的规划



生活中的大规模问题

Example

- 金融衍生产品（期货、掉期等）的定价
- 民航航路的规划



大规模问题：维度的诅咒(Curse of Dimensionality)

Definition

问题维度增加时，解空间体积**指数级**增加的难题。

Example

维度为1 10^1 个均匀分布的点能够保证在单位区间上进行间距不超过0.1的采样

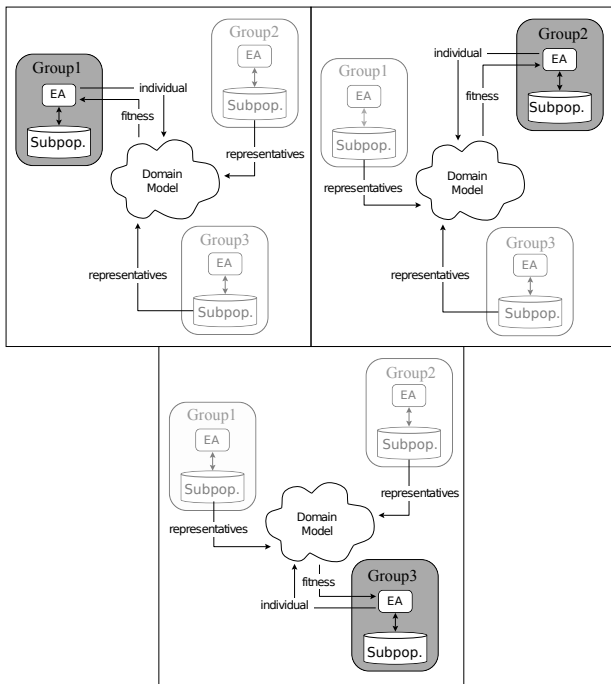
维度为3 10^3 个均匀分布的点能够保证在单位立方体上进行间距不超过0.1小方格的采样

问题规模 $1 \rightarrow 3$ ，需采样点数目增加了 10^2 倍

人工智能中的协同演化 (Cooperative Coevolution)

Definition

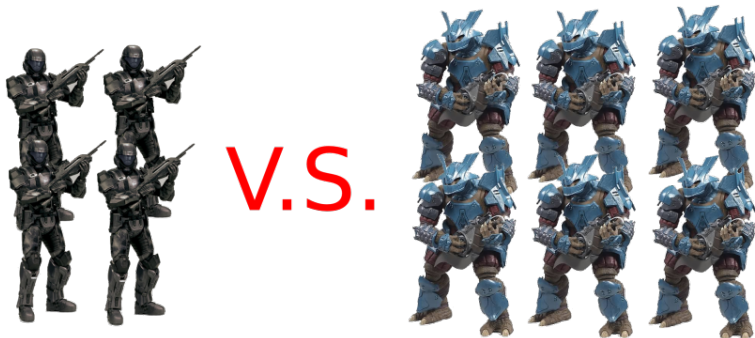
基于**分而治之**，将问题分解规模较小的若干个子问题，逐个解决子问题以期达到最终解决原大规模问题的目的



一个形象的例子

背景 2012年有一群外星人入侵地球…

目标 尽可能快消灭所有入侵的外星人



情景1:分而治之的成功应用

原问题各组分之间相互独立

- 将敌军**分解**个若干个小集团
- **逐个击破**
- 最终消灭所有侵略者



什么都没看见...



NICAL

协同演化：数值优化中的分而治之

优化球型(Sphere)函数

$$F_{\text{sphere}}(\mathbf{x}) = \sum_{i=1}^D x_i^2 \quad (2)$$

特征:

$$\arg \min_{(x_1, \dots, x_N)} F_{\text{sphere}}(x_1, \dots, x_N) = \left(\arg \min_{(x_1)} F_{\text{sphere}}(x_1, \dots), \dots, \arg \min_{(x_N)} F_{\text{sphere}}(\dots, x_N) \right) \quad (3)$$

全局最优点可以通过逐维优化而达到

协同演化：数值优化中的分而治之

优化球型(Sphere)函数

$$F_{\text{sphere}}(\mathbf{x}) = \sum_{i=1}^D x_i^2 \quad (2)$$

特征:

$$\arg \min_{(x_1, \dots, x_N)} F_{\text{sphere}}(x_1, \dots, x_N) = \left(\arg \min_{(x_1)} F_{\text{sphere}}(x_1, \dots), \dots, \arg \min_{(x_N)} F_{\text{sphere}}(\dots, x_N) \right) \quad (3)$$

全局最优点可以通过逐维优化而达到

使用分而治之方法的前提

分解后的子问题之间应该不存在相关性

情景2:分而治之可能存在的问题

子问题之间具有相关性

- 逐个解决子问题不能够保证最终解决原问题
- 必须将相关的子问题合并



协同演化：数值优化中的相关性问题的

Definition

- 所有变量之间都不存在相关性的函数称为可分函数，如方程[2]所示的球型函数。
- 否则，若有某对变量之间有相关性，该函数为不可分函数，如方程4的Rosenbrock函数。

不可分函数：Rosenbrock函数

$$F_{\text{rosenbrock}}(\mathbf{x}) = \sum_{i=1}^{D-1} [100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2] \quad (4)$$

注意：不可分函数必须将相关的变量放在一组优化，否则无法最终达到全局最优点。

黑盒问题：两难的分组选择

如果孤注一掷…

- **选整体优化策略:** 对于可分解或者部分可分解的一类问题, 优化效率低下
- **选单维优化策略:** 对于不可分解的一类问题, 算法将很快收敛与局部最优解, 失去搜索全局最优解的能力

黑盒问题：两难的分组选择

如果孤注一掷...

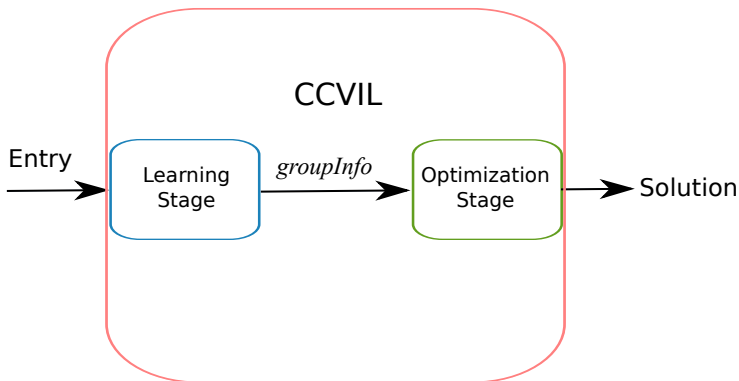
- **选整体优化策略**: 对于可分解或者部分可分解的一类问题, 优化效率低下
- **选单维优化策略**: 对于不可分解的一类问题, 算法将很快收敛与局部最优解, 失去搜索全局最优解的能力

我的解决方案

- 花费一定代价学习变量的相关性信息
- 再根据所获取的信息合理分组, 高效进行优化

预期目标: 在特征各异的多类问题上, 算法总体性能都比较好, 显著增强大规模优化算法的健壮性

算法流程示意



变量相关性理论基础

可分解函数定义

一个函数是可分解的，当且仅当它满足 5. [2]

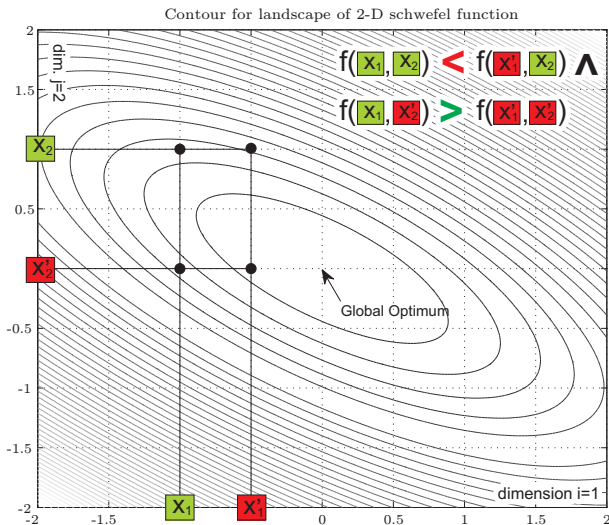
$$\arg \min_{(x_1, \dots, x_N)} f(x_1, \dots, x_N) = \left(\arg \min_{(x_1)} f(x_1, \dots), \dots, \arg \min_{(x_N)} f(\dots, x_N) \right) \quad (5)$$

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 \bar{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[1].

$$\exists \bar{x}, x'_i, x'_j: f(\bar{x}_1, \dots, \bar{x}_i, \bar{x}_j, \dots, \bar{x}_n) \text{ (green)} < f(\bar{x}_1, \dots, \bar{x}_i, x'_j, \dots, \bar{x}_n) \text{ (red)} \wedge \\ f(\bar{x}_1, \dots, \bar{x}_i, \bar{x}_j, \dots, \bar{x}_n) \text{ (green)} > f(\bar{x}_1, \dots, x'_i, \bar{x}_j, \dots, \bar{x}_n) \text{ (red)}$$

相关性学习简单实例

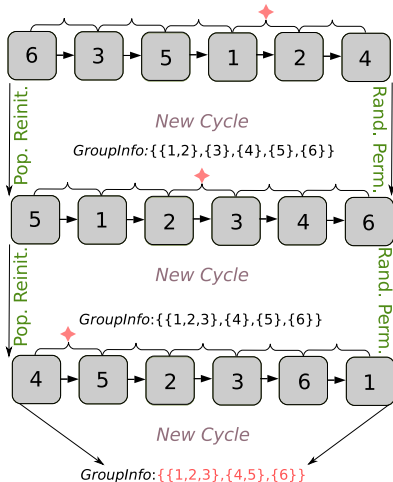


变量相关性学习阶段演示

Target: $\{\{1,2,3\},\{4,5\},\{6\}\}$

Initialize GroupInfo:

GroupInfo: $\{\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}\}$



实验模拟结果比较

	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	9.62e+12	3.43e+12	5.00e+12	3.38e+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

实验结果总结

- 与国际上最先进的大规模协同演化算法DECC-G[3]和MLCC[4]相比较，在绝大部分函数上，本作品的算法都有统计显著的优势
- 与嵌入在本作品算法框架的算法JADE[5]相比，新的算法框架同样使得原算法的性能在多数测试问题上有统计显著的提升
- 在有些测试问题上，本作品算法所得解的标准差较大

作品意义(1)

学术影响

- 在协同演化算法中首次系统的引入了具有切实理论基础变量相关性学习机制
- 通过实验模拟，表明了在大规模优化相关性学习的重要性，为将来的研究提供有益思路
- 不同组分之间的相互依赖的关系在人工神经网络、多智能体系统和分布式系统等研究领域广泛存在，本文所提出的简单而高效的方法提供了有益的参考

本作品的一部分内容已经发表于国际会议

PPSN' 2010 (Parallel Problem Solving From Nature)，会议录由Springer的LNCS系列出版。论文被

SCI(000286453000031)、EI(20104513369063)和ISTP检索。

作品意义(2)

社会影响

本作品所研究的课题在工业生产和工程项目中有着广泛的应用:

- 发动机叶片的自动设计
- 无线传感器网络部署(Wireless Sensor Network Deploymen)
- ...



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.



Ke Tang, Xiaodong Li, Ponnuthurai Nagaratnam Suganthan, Zhenyu Yang, and Thomas Weise.

Benchmark functions for the CEC’ 2010 special session and competition on large scale global optimization.

Technical report, NICAL, USTC, Hefei, Anhui, China, 2009.

<http://nical.ustc.edu.cn/cec10ss.php>.



Zhenyu Yang, Ke Tang, and Xin Yao.

Large scale evolutionary optimization using cooperative coevolution.

Information Sciences, 178(15):2985 – 2999, 2008.



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.



Jingqiao Zhang and Arthur C. Sanderson.

JADE: adaptive differential evolution with optional external archive.
IEEE Transactions on Evolutionary Computation, 13(5):945 – 958,
2009.

感谢关注！

Email: chenwx.ustc@gmail.com

WWW : <http://cs-chen.net>