

特征统计算法及其在 NP 组合优化问题上的应用

Characteristic Statistic Algorithm (CSA) and Its Application to NP Combined Optimization Problems

刘志宏/LIU Zhi-hong¹, 胡永明/HU Yong-ming¹, 施工/SHI Gong²

1. 清华大学核能与新能源技术研究院, 北京 100084
2. 清华大学工程物理系, 北京 100084

1. Institute of Nuclear and New Energy Technology, Tsinghua University, Beijing 100084, China
2. Department of Engineering Physics, Tsinghua University, Beijing 100084, China

[摘要] 特征统计算法是为了解决复杂多极值优化问题而开发的一种新的全局优化算法。为了检验该算法的性能,应用它在一类具有代表性的 NP 组合优化问题-旅行商问题(TSP)上作了计算。结果发现,该算法虽不是专为 TSP 问题而开发,却在该问题上取得了很好的结果。所得到的结果表明,特征统计算法可以作为解决这类 NP 组合优化问题的一个新的途径。

[关键词] 特征统计算法(CSA); NP 问题; 组合优化

[中图分类号] O 224

[文献标识码] A

[文章编号] 1000-7857(2006)11-0028-03

收稿日期: 2006-06-12

作者简介: 刘志宏,男,清华大学核能与新能源技术研究院,助理研究员,研究方向为反应堆物理计算及优化算法等方面的研究;
E-mail: Lzh@tsinghua.org.cn

表 5 F 检验相关数据
Tbl. 5 F relative inspection data

组分	平均值 /%	自由度 n	标准偏差 S	标准偏差 S ²	F=S ² /S ² ₁	F	有无显著 性差异
丙酮 (面积归一化法)	23.1	5	0.11	0.01	2.51	6.26	无
丙酮(外标法)	22.11	6	0.17	0.03			
甲丙酮 (面积归一化法)	7.72	5	0.09	0.008	2.48	6.26	无
甲丙酮(外标法)	8.2	6	0.06	0.003			
双基酮 (面积归一化法)	51.05	5	0.25	0.06	3.24	6.26	无
双基酮(外标法)	53.55	6	0.45	0.2			

误差都小于 6.0%,对于准确度要求不太高(如中控分析)的场合是可以的。

3 结论

3.1 定性方法

运用色谱-质谱联用仪的计算机检索和标准样品叠加法对二异丙基酮冷凝液定性准确而方便,避免了单纯运用标样叠加法的繁琐性和局限性,是一较好的定性方法。

3.2 定量方法

表 6 t 检验相关数据
Tbl. 6 t relative inspection data

组分	平均值 /%	自由度 n	标准偏差 S	t 值	t _{0.05}	有无显著 性差异	相对误差 /%
丙酮 (面积归一化法)	23.11	5	0.15	11.37	2.26	有	4.5
丙酮(外标法)	22.11	6					
甲丙酮 (面积归一化法)	7.72	5	0.07	17.74	2.26	有	5.8
甲丙酮(外标法)	8.2	6					
双基酮 (面积归一化法)	51.05	5	0.37	11.06	2.26	有	4.7
双基酮(外标法)	53.55	6					

采用外标法对二异丙基酮定量,结果的精密度和准确度都非常高,是较好的定量方法之一。采用面积归一化法对二异丙基酮定量,精密度很好,准确度稍差,存在一定的系统误差,只适合于分析频次多、对定量结果准确度要求不太高的场合使用。若需准确定量时,必须采用外标法,以减少误差。

参考文献 (References)

- [1] 武汉大学. 分析化学[M]. 北京: 高等教育出版社, 1978: 125-130.
- [2] 刘虎威. 气相色谱方法及应用[M]. 北京: 化学工业出版社, 2000.

(责任编辑 赵佳)

Abstract: Characteristic statistic algorithm (CSA) is a newly developed global optimization algorithm for complex multi-extreme optimization problems. To test the optimization efficiency of CSA, it is used on a benchmark NP problem - TSP problem. Calculated results show that CSA has excellent ability on TSP problems, though it is not specially developed for TSP. It is proved that CSA can be a new method for solving these NP combined optimization problems.

Key Words: characteristic statistic algorithm (CSA); NP problem; combined optimization

CLC Number: O 224

Document Code: A

Article ID: 1000-7857(2006)11-0028-03

1 引言

全局优化问题是运筹学中一个重要的研究方向。对于单极值优化问题,已经发展了很多成熟的方法来获得它的最优点;而对于多极值问题和离散变量的组合优化问题,目前还没有一种方法能够确保在可接受的时间内获取其全局最优点。许多实际的工程设计优化问题都是复杂的多极值问题或者组合优化问题,对于这类问题,如何有效地找到它的全局最优(或接近于全局最优)解是一件迫切需要而又富有挑战性的工作。因此,全局优化算法研究也是当前优化算法研究中的一个热点。大家都希望找到一种既不需要耗费太多计算时间而又能够使所得解更加接近于全局最优解的算法。

目前国际上关于这方面的研究并不非常成熟,已有的算法有模拟退火算法(SA)、遗传算法(GA)、禁忌搜索(TS)、人工神经网络(ANN)、蚁群搜索算法(ACO),以及在这些算法基础上改进的其他算法^[1-4]。这些算法都在许多领域内取得了较好的效果,但是在实际应用中,当优化问题非常复杂或者比较特殊时,它们都存在搜索效率不高的问题。

通过分析发现,这些算法是利用某一种通用性质来进行搜索的,而不管具体的优化问题是否符合这种性质。因此,当实际的优化问题不具有这些性质或者很难转化为具有该性质的问题时,这些算法的搜索效率就会大打折扣^[5]。笔者开发了一种新的全局优化算法——特征统计算法(注:在以前的文献中曾经称之为统计归纳算法^[5])。特征统计算法(CSA)能够针对不同优化问题自身具体的特点,灵活地选取一些不同的特征量,然后利用最优点在每个特征量上的一维分布概率来构造整个自变量空间上的最优点分布概率情况,从而引导搜索方向。因此,CSA能够获得很高的搜索效率和很好的全局性,非常适合于用来解决情况比较复杂的多极值问题和组合优化问题。

2 CSA 算法

CSA 算法根据具体问题的特点,给出一些特征量,从而确定出应该如何从已计算的样本点中提取有用的信息,然后用它来指导下一步的搜索。图1给出了CSA算法的基本流程。该算法的基本步骤如下。

1) 在定义域内随机地给出 N_0 个解,并对这 N_0 个解按照目标函数值大小进行排序,取出其中函数值最好的 N_1 个解作为统计样本群体。

2) 根据优化问题自身的性质,给出一些特征量作为统计项目,然后对这 N_1 个解进行统计。对于每一个特征量,可以由这 N_1 个解所对应的 N_1 个值得到一条统计分布曲线。

3) 根据这些统计分布曲线,构造出 N_1 个新解,使得这 N_1 个新解在每一个特征量上的统计分布与上一步所得到的统计分布相近。

4) 将这 N_1 个新解加入到所有已计算的解中,重新排序,选出 N_1 个当前最优解作为统计样本,然后返回步骤2。

关于在步骤2中应该如何来选取特征量,需根据问题自身的特点来确定。这一点将在第3节中进一步讨论。

采用 CSA 算法在一些连续函数的标准测试问题和组合优化的旅行商问题(TSP)上分别作了计算,均得到了很好的结果。同时,CSA 算法还在一些实际的全局优化问

题(如压水堆堆芯装料优化问题、辐射屏蔽设计优化问题,以及一个图像重建问题等)上得到了很好的验证。

3 CSA 算法在 NP 组合优化问题上的应用

这里选择组合优化领域中比较富有挑战性的一个 NP 难度问题——“旅行商问题”(TSP)来检验本算法的性能。所谓 TSP 问题,可以简单地这样来描述:给定 N 个城市,任意 2 个城市之间的距离已知,希望找到一条经过每一个城市的闭合路径,使得该路径的长度最短。

对于 TSP 问题,事先我们可以知道该问题的特点是:目标函数值的大小只与每个城市前后连接的是什么城市有关。因此,本算法第2步中所选择的特征量就是每个城市前后的城市分布情况,即所选取的 N 个统计项目 $\{G_i\}$ 为:

G_i = 样本中与城市 i 相连接的城市号

这样,得到的统计分布就是其他城市与该城市相连接的概率分别为多少。总共得到 N 个这样的统计分布 (N 为城市数

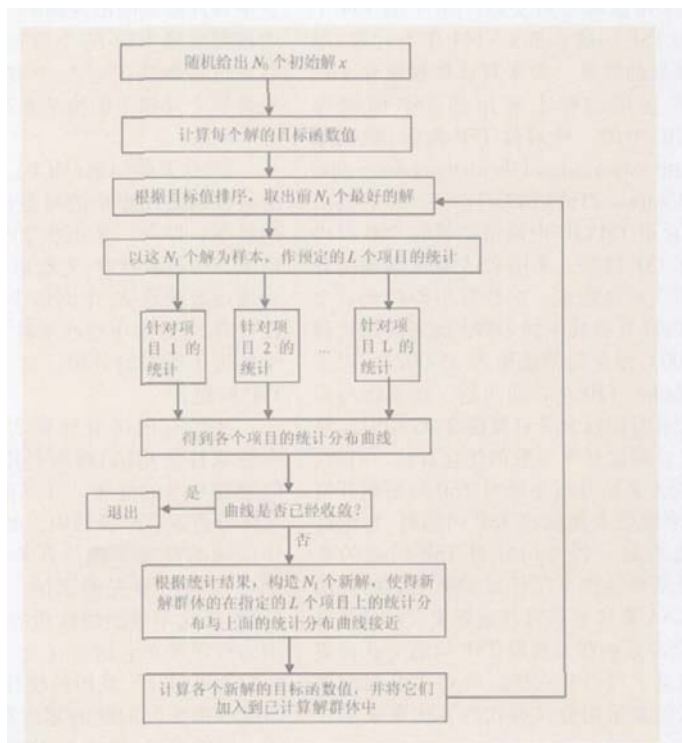


图1 CSA 算法流程图

Fig. 1 Flowchart of CSA

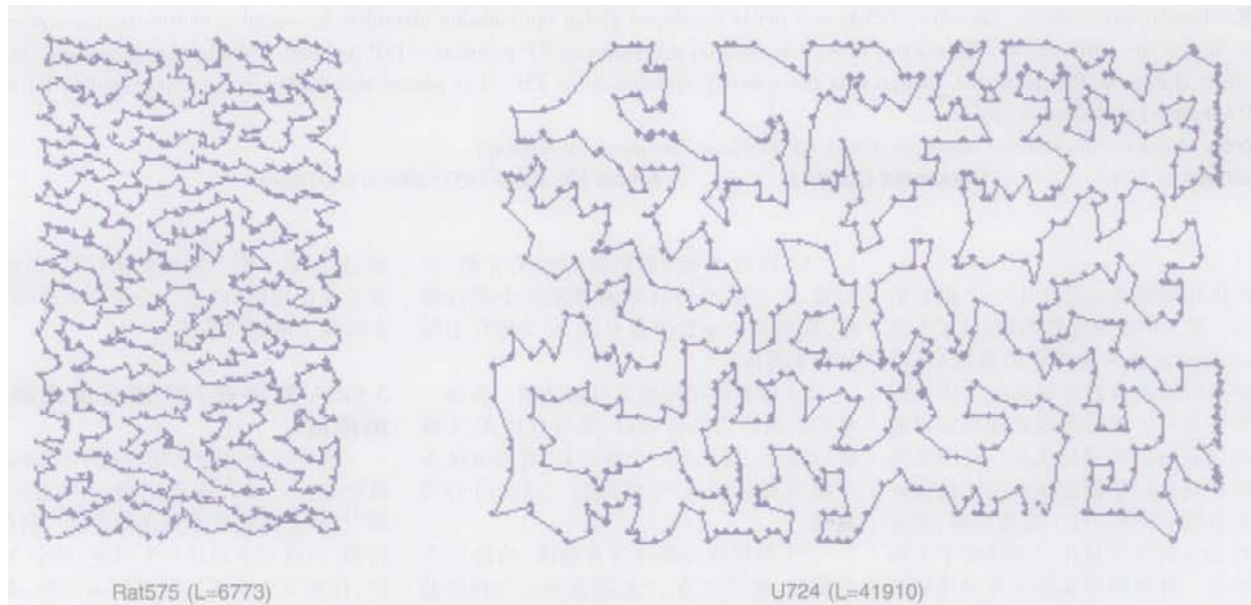


图2 TSPLIB 中两个 TSP 问题的计算结果

Fig. 2 Results of two TSP problems in TSPLIB

目)。而第 3 步中构造新点时,由于涉及到新点的可行性问题,不像连续问题那样简单,需要一定的技巧。如何得到 N_1 个既满足所需的统计分布,又是可行点的新点呢?程序中采用了这样一种办法:首先随机给出 N_1 个新的可行解,然后根据所需的统计分布,对这 N_1 个序列中的元素进行移动操作,使之逐渐接近于所需的统计分布。

为了检验本算法对 TSP 问题的优化性能,曾采用该算法对文献[1]中中国 144 个城市的 TSP 问题实例 CN144 作了计算,取得了很好的效果。后来对计算程序作了改进,并选用国际上通用的 TSP 问题库“TSPLIB”中的一些对称 TSP 例题(数据来源为 <http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/>)作了计算。结果发现,对于 TSPLIB 中城市规模在 700 以内的对称 TSP 问题,采用 CSA 算法编制的计算程序几乎都能在一次计算中即找到其全局最优解(只有其中的 D493 找到的最优解为 35 004,而全局最优解为 35 002)。对于城市规模在 150 左右的问题,该算法所需要的计算时间仅为 3 s(赛扬 2.4G CPU)。该结果已经远远好于一般的优化算法。当然,由于 CSA 算法并不是针对 TSP 问题而开发的,当求解更大规模的 TSP 问题时,它的性能可能不如一些专门针对 TSP 问题的算法。但是和这些专门针对 TSP 问题的算法相比,CSA 算法更具有普遍意义。而且如果希望 CSA 算法在大规模 TSP 问题上获得更大的效果,可以对它进行针对 TSP 问题的修改(例如采用分区叠代的方法来减小计算量)。

图 2 给出了 CSA 算法对 TSPLIB 中 2 个对称 TSP 问题的计算结果。图中算法所

得到的解均为该问题的全局最优解。

上面针对 TSP 问题分别给出了本算法的实现过程。事实上,本算法几乎适用于所有的全局优化问题。例如文献[1]中的其他几个 NP 组合优化问题,都可以轻易地给出采用 CSA 算法的实现方案。以前 2 个为例作如下阐述。

1) 0-1 背包问题(ZKP)。根据该问题的性质,所应该选取的特征量是在 N_1 个样本点中各件物品的出现概率。在构造新点时,就按照此概率向 N_1 个背包中给出物品,然后通过将物品在这 N_1 个背包中进行倒换,使得每个背包中的物品总重量尽量趋于相等。

2) 最大截问题(MCP)。它所应该选取的特征量是图形中的每条边落在记数集中的概率,即每一条边的 2 个端点不同色的概率。构造新点时,先给出 N_1 个可行解,然后通过改换这 N_1 个图形中各个点的颜色,使得 N_1 个新解中这些项目上的统计分布尽量接近于所需的分布。这一步过程类似于 TSP 问题。

从已有的所有计算实例中发现,CSA 在选取特征量 $\{G_i\}$ 时所利用的性质都是该问题最基本的性质,并不需要很多复杂的具体分析或者经验知识。而且,当我们对具体问题的性质了解得并不是很清楚时,完全可以利用事先假定的一些共有的性质(例如变化不剧烈的性质等),采用一个通用的程序来对它进行优化,同样会取得较好的效果。另外,其中的统计样本数 N_1 的大小应根据实际问题的复杂程度来选取。对于一般的问题,如果 N_1 的取值在几百以上,一般都能够较好地表示出每个项目的统计性质。

4 结论

特征统计算法(CSA)是一种全新的算法,它在智能性和全局性方面均取得了明显效果。和其他的全局优化算法相比,CSA 算法的特点主要体现在以下 2 个方面。

1) CSA 算法可以根据优化问题自身的特点灵活地选取不同的特征量作为统计项目,这样可以最充分地利用问题自身的性质。这一点 SA、GA 等其他算法都无法做到。因此 CSA 算法能够解决很多其他算法难以解决的复杂优化问题。

2) CSA 算法采用最优解在每个特征量上的一维分布概率来构造最优解在整个自变量的多维空间上的概率分布,这样能够精练而又准确地提炼出由已计算点获得的信息,获得更高的搜索效率。

同时,CSA 算法具有固有的并行性,因此它在全局优化方面具有很大的发展潜力,非常适合于用来解决各种实际工程中复杂的多极值问题和组合优化问题。

参考文献(References)

- [1] 康立山,谢云,尤矢勇,等.非数值并行算法-模拟退火算法[M].北京:科学出版社,1998.
- [2] 刘勇,康立山,陈旒屏.非数值并行算法-遗传算法[M].北京:科学出版社,1998.
- [3] 邢文训,谢金星.现代优化计算方法[M].北京:清华大学出版社,1999.
- [4] STUTZLE T, HOOS H H. Max-min ant system [J]. Future Generation Computer System, 2000, 16: 889-914.
- [5] 刘志宏,施工,胡永明.一种新的全局优化算法-统计归纳算法[J].清华大学学报,2002(5):580-583.

(责任编辑 王 芷)