

模拟退火算法在优化中的研究进展

项宝卫¹, 余雪芬², 骆兆文³

(1 台州学院 信息与工程学院, 浙江 临海 317009)

2 黄岩区域关中学, 浙江 台州 318024 3 椒江区三甲中学, 浙江 台州 318000)

摘 要: 模拟退火算法在处理全局优化、离散变量优化等困难问题中, 具有传统优化算法无可比拟的优势。在计算机计算速度不断提高的条件下, 具有不可低估的发展潜力和重要的研究价值。本文研究分析了基本的模拟退火算法原理, 主要介绍了模拟退火算法的发展历史和特点, 综合了近期关于模拟退火算法研究进展。本文在比较了模拟退火算法与传统的优化方法的基础上, 最后指出了它的发展趋势和研究热点, 对模拟退火算法在优化中的推广应用具有重要意义。

关键词: 模拟退火; 全局优化; 传统优化方法; 隐含并行性

中图分类号: O 224 **文献标识码:** A **文章编号:** 1672-3708(2005)06-0006-04

1 引言

优化算法种类繁多, 广义上可分为基于微分法和直接搜索法。如果优化函数及其导数的表达式容易求得, 且计算导数值并没有庞大的运算量, 则基于微分法常常较为有效。但在实际应用中, 优化函数往往具有很复杂的结构, 如遇到非凸和不连续函数时只能采用直接搜索法。从数学的观点, 检查优化问题的凸性是很重要的。如果优化问题是凸规划, 只有一个全局极值, 传统的数值优化算法比较有效; 如果是非凸问题, 则存在若干个局部极值, 最小的一个是全局最优解, 即问题的最优解。^[1]

随着问题种类的不同以及问题规模的扩大, 要寻求到一种能以有限的代价来求解最优化问题的通用方法仍是一个难题, 建立以最大可能(概率)求解全局解仍是一个重要问题。而且传统的优化方法对以下问题的处理比较困难: (1)非凸可行域以及具有多个局部极值问题; (2)不连通的可行域; (3)设计变量全部或部分为离散的、整型的; (4)目标函数具有多个极值; (5)难于求解目标函数和约束函数的梯度。

模拟退火算法(Simulated Annealing 简称 SA)为求解上述传统方法难处理的问题提供了一个有效的途径和通用框架, 并逐渐发展成一种迭代自适应启发式概率性搜索算法。用以求解不同的非线性问题; 对不可微甚至不连续的函数优化, SA能以较大概率求得全局优化解; 具有较强的鲁棒性、全局收敛性、隐含并行性及广泛的适应性; 并且能处理不同类型的优化设计变量(离散的、连续的和混合型的); 不需要任何的辅助信息, 对目标函数和约束函数没有任何要求。利用 Metropolis算法并适当地控制温度下降过程, 在优化问题中具有很强的竞争力, 因此研究 SA 算法在优化中的应用显得尤为重要。^[2, 3] 本文研究分析了基本的模拟退火算法原理, 主要介绍了模拟退火算法的发展历史和特点, 综合了近期关于模拟退火算法研究进展, 以推广 SA 算法在优化中的应用。

2 模拟退火算法的基本原理

SA 全局优化算法的基本原理如下: (1)给定初始温度 T_0 和初始点 x_0 , 计算该点的函数值 $f(x_0)$; (2)随机产生扰动 Δx 得到新点 $x' = x + \Delta x$ 计算新点的函数值 $f(x')$ 和差 $\Delta f = f(x') - f(x)$; (3)若 $\Delta f \leq 0$ 则接受新点, 作为下一次模拟退火的初始点; (4)若 $\Delta f > 0$ 则计算新点的接受概率: $p(\Delta f) =$

$\exp(-\Delta f(K, T))$, 产生 $[0, 1]$ 区间上均匀分布的伪随机数 $r, r \in [0, 1]$ 。若 $p \leq r$ 则接受新点作为下一次模拟的初始点; 否则仍取原来的点作为下一次模拟退火的初始点。以上步骤称为 *Metropolis* 过程。逐渐降低控制温度, 重复 *Metropolis* 过程, 直至达到结束准则, 就构成了 *SA* 算法。*SA* 算法能收敛到全局最优点或近似全局最优点。

3 模拟退火算法的历史回顾

Metropolis 准则是 1953 年 *Metropolis N.* 等人在研究二维相变时提出的。*SA* 算法最早分别由 *Kirkpatrick S.* 等人 (1980) 和 *Cerny V.* (1981) 独立地提出, 但收敛速度很慢。1983 年 *Kirkpatrick S.* 等用这种方法设计大规模集成电路 (*VLSI*)。^[4] 1986 年 *Szu H.* 提出了一种退火率与时间成反比的快速模拟退火算法 (*FSA*)。^[5] 1987 年 *Laarhoven P.* 和 *Aarts E.* 出版了《模拟退火的理论 and 应用》一书,^[6] 对 *SA* 算法作了比较系统的总结, 促进了 *SA* 算法的理论研究和实际应用的发展, 这是 *SA* 算法发展史上的一块里程碑。

1990 年 *Gun t e D.* 和 *Tobias S.* 研究了 *SA* 算法中初始温度的临界值的确定方法。^[7] 1993 年 *Kirkpatrick S.* 等人将 *SA* 算法用于优化问题, 取得了相当不错的效果。^[3] 1995 年 *Tarek M.* 等人将 *SA* 算法进行了并行化计算的研究, 以提高 *SA* 算法的计算效率, 用来解决比较复杂的科学和工程计算。^[8] 1997 年胡山鹰等人在无约束非线性规划问题全局优化的 *SA* 算法基础上, 进行有约束问题求解的进一步探讨, 对不等式约束条件提出了检验法和罚函数法的处理方法, 对等式约束条件开发了罚函数法和解方程法的求解步骤, 并进行了分析比较, 形成了完整的求取非线性规划问题全局优化的模拟退火算法。^[9] 1999 年康立山等出版了《非数值并行算法》(第一册), 对并行的 *SA* 算法作了概括和较系统的总结。^[10]

2000 年向阳等人介绍了推广模拟退火方法的基本思想及其统计基础, 通过一系列标准函数测试了推广 *SA* 算法的性能, 探讨了推广模拟退火方法的效率随体系复杂性的变化。^[11] 2001 年都志辉等人提出一种混合 *SPMD-SA* 算法, 在克服经典 *SA* 算法内在串行性的同时, 进一步和下山法结合起来, 并综合多种优化方法, 在一定的处理机规模内取得了可扩展的并行效果, 显著提高了算法的收敛速度, 克服了算法性能对初始值和参数选择的过分依赖, 在提高算法性能的同时, 方便了算法的使用。^[12] 2002 年耿平等采用人工神经网络方法建立多变量与多目标函数之间的关系, 并将模拟退火算法与人工神经网络 *BP* 算法相结合, 解决了这类复杂系统中多函数变量与多目标函数之间没有确定的解析关系因而无法进行直接优化的难题, 并为解决多变元非线性复杂系统的优化问题提供了一种新的有效的方法。^[13] 目前, *SA* 算法迎来了兴盛时期, 无论是理论研究还是应用研究都成了十分热门的课题。尤其是 *SA* 算法的应用研究显得格外活跃。

4 模拟退火算法与其它优化算法的比较及其特点

SA 算法不仅在理论上能突破传统算法难以解决的难题, 而且具有很强的科学和实际的工程应用价值, 因而被誉为解决许多高难度优化问题的救星。下面分析 *SA* 算法和其它传统搜索方法的对比。

解析法是常用的搜索方法之一。它通常是通过求解使目标函数梯度为零的一组非线性方程来进行搜索的。一般而言, 若目标函数连续可微, 解的空间方程比较简单, 解析法还是可以用的。但是若方程的变量有几十或几百时, 它就无能为力了。爬山法也是常用的搜索方法, 它和解析法一样都是属于寻找局部最优解的方法。对于爬山法, 只有在更好的解位于当前解附近的前提下, 才能继续向优解搜索。显然这种方法对于具有单峰分布性质的解空间才能进行行之有效的搜索, 并得到最优解。

另一种典型的搜索方法是穷举法。该方法简单易行, 即在一个连续有限搜索空间或离散无限搜索空间中, 计算空间中每个点的目标函数, 且每次计算一次。显然, 这种方法效率太低而鲁棒性不强。许多实际问题所对应的搜索空间都很大, 不允许一点一点地慢慢求解。

SA 方法也是利用随机化技术来指导对于最小能量状态的搜索。当面临更为复杂的问题时, 必须

采用像 SA 算法这样采用随机化技术的优化方法。SA 算法的特点主要有以下几个方面:

(1) 以一定的概率接受恶化解。SA 算法在搜索策略上与传统的随机搜索方法不同, 它不仅引入了适当的随机因素, 而且还引入了物理系统退火过程的自然机理。这种自然机理的引入使 SA 算法在迭代过程中不仅接受使目标函数变“好”的试探点, 而且还能以一定的概率接受使目标函数值变“差”的试探点, 迭代中出现的状态是随机产生的, 并且不强求后一个状态一定由于前一个状态, 即以一定的可能容忍的退化状态的出现。接受概率随着温度的下降而逐渐减小。传统的方法往往是从解空间一个初始点开始最优解的迭代搜索过程。如登山法, 若一个细微变动能改善质量, 则沿该方向前进, 否则取相反方向。然而复杂问题会使解空间中出现若干局部最优解, 传统的方法很容易限于局部最优解而停滞不前。很多传统的优化算法往往是确定性的。从一个搜索点到另一个搜索点的转移有确定的转移方法和转移关系, 这种确定性往往可能使得搜索永远达不到最优点, 因而限制了算法的应用范围。而 SA 算法是以一种概率的方式来进行的, 从而增加了其搜索过程的灵活性。

(2) 引进算法控制参数 T 。引进类似于退火温度的算法控制参数 T , 它将优化过程分成各个阶段, 并决定各个阶段下随机状态的取舍标准, 接受函数由 Metropolis 算法给出一个简单的数学模型。SA 算法的两个重要步骤是: 一是在每个控制参数 T 下, 由前迭代点 $x(i)$ 出发, 产生邻近的随机状态 $x(i+1)$, 由 T 确定的接受准则决定此新状态的取舍, 并由此形成一定长度的随机 Markov 链; 二是缓慢降低控制参数 T 提高接收准则, 直至 $T \rightarrow 0$ 状态链稳定于优化问题的最优状态, 提高 SA 算法获得全局最优解的可靠性。

(3) 使用对象函数值(即适应值)进行搜索。传统搜索算法不仅需要利用目标函数值, 而且往往需要目标函数的导数值等其它一些辅助信息才能确定搜索方向。当这些信息不存在时, 算法就无效了。而 SA 算法仅使用由目标函数变换来的适应度函数值, 就可确定进一步的搜索方向和搜索范围, 无需其它一些辅助信息。需要着重提出的是, SA 算法的适应度函数不仅不受连续可微的约束, 而且其定义域可以任意设定。对适应度函数唯一要求是, 对于输入可计算出加以比较的正的输出。这个特性对很多无法或很难求导数的函数, 或导数不存在的函数的优化问题, 以及组合优化问题等, 应用 SA 算法就显得比较方便。

(4) 隐含并行性。并行算法是 60 年代发展起来的, 发展迅速。从目前情况看, 并行算法的设计主要采用两种方法: 一是对现有的串行算法加工改造, 使之成为好的并行算法; 二是结合所用并行计算机的结构特点, 直接设计新的并行算法。对模拟退火算法改造为并行算法还是比较容易的。目前常见的有以下几种并行策略: 操作并行策略, 试演并行策略, 区域分裂策略, 混乱松弛策略。这几种并行算法在不同程度上对解的质量、收敛速度方面较模拟退火算法优。由此可以预见, 大规模的并行计算模式将成为研究全局优化问题的主流。即 SA 算法隐含并行性 (Implicit Parallelism), 它是优于其它求解过程的关键所在。另外 SA 算法的隐含并行性还有助于处理非线性问题。

(5) 搜索复杂区域。SA 算法最善于搜索复杂地区, 从中找出期望值高的区域。但在求解简单问题时效率并不高。正如遗传算法创始人 Holland H 所指出的“如果只对几个变量作微小的改动就能进一步改进解, 则最好使用一些更普通的方法, 来为遗传算法助一臂之力”。SA 算法在这一点上与遗传算法类似, 但比遗传算法更加适合搜索复杂区域。

上述具有特色的技术和方法使得 SA 算法使用简单、鲁棒性强、易于并行化, 从而应用范围甚广。

5 模拟退火算法的研究发展趋势

目前, 关于 SA 算法的研究通常分为两类。第一类, 基于有限状态奇异马尔可夫链的有关理论, 给出 SA 的某些关于理想收敛模型的充分条件或充要条件, 这些条件在理论上证明了当退火三原则(初始温度足够高、降温速度足够慢、终止温度足够低)满足时, SA 以概率 1 达到全局最优解。第二类, 针

对某些具体问题, 给出了 SA 的若干成功应用。前者在指导应用方面作用有限, 在定参过程中, 往往很难给出有益的定量关系。后者各自的领域中有应用价值, 但过分依赖于问题, 不具有普遍意义。事实上, 在现有情况下给出关于 SA 的具有普遍意义的定量关系式是不现实的。因此, 对 SA 进行的有意义研究应集中在引入新思想在此基础上提出在应用中实现新思想的可能途径, 并通过典型实验对其效果进行验证。SA 的未来发展方向应着重解决以下几个问题:

(1) 如何把传统的启发式搜索方法与 SA 随机搜索算法结合起来;

(2) 如何把 SA 算法与 GA (遗传) 算法有机结合起来, 开发出一种更具有理论意义和应用价值的随机搜索算法;

(3) 由于 SA 及其变种所固有的特点, 它们在解决某些特殊领域的问题时具有很好的性能。寻找更多的使用领域, 并在这些领域中给出成功的应用系统, 也是一个颇具理论意义和应用价值的研究方向。

参考文献:

- [1] 王凌. 智能优化算法及其应用 [M]. 北京: 清华大学出版社, 2001
- [2] 杨若黎, 顾基发. 一种高效的模拟退火全局优化算法 [J]. 系统工程理论与实践, 1997, 17(5): 30-33
- [3] Ingber L, Rosen B. Genetic algorithms and very fast simulated annealing: A comparison. *Mathematical Computer Modeling* [J]. 1992, 16(11): 87-100
- [4] Kirkpatrick S, Gelatt C, and Vecchi M. Optimization by Simulated Annealing. *Science* [J]. 1983, (220): 671-680
- [5] Szűcs H, and Hartley P. Fast Simulated Annealing [J]. *Physics Letters A* 1987, (122): 157-162
- [6] Laarhoven P and Aarts E. *Simulated Annealing: Theory and Applications* [M]. D. Reidel Publishing Company, Dordrecht, 1987.
- [7] Deck G, Scheuer T. Threshold accepting: A general purpose optimization algorithm appearing superior to simulated annealing. *Journal of Computational Physics* [J]. 1990, 90(1): 161-175
- [8] Tarek M, Nabhan A, Bert Y, Zouaya A. parallel simulated annealing algorithm with low communication overhead [J]. *IEEE Transactions on Parallel and Distributed Systems* 1995, 6(12): 1226-1233
- [9] 胡山鹰, 陈丙珍, 何小荣, 等. 非线性规划问题全局优化的模拟退火法 [J]. 清华大学学报 (自然科学版), 1997 (6): 5-9
- [10] 康立山, 谢云. 非数值并行算法 (第一册) [M]. 北京: 科学出版社, 1994
- [11] 向阳, 龚新高. 推广模拟退火方法及其应用 [J]. 物理学进展, 2000, 20(3): 319-334
- [12] 都志辉, 李三立, 吴梦月, 等. 混合 SPMD 模拟退火算法及其应用 [J]. 计算机学报, 2001, 24(1).
- [13] 耿平, 刘静, 曾梅光. 多变元非线性复杂系统的优化与模拟退火算法 [J]. 东北大学学报 (自然科学版), 2002, 23(3): 270-272

Research and Development of Simulated Annealing Algorithm in Optimization

XIANG Bao-wei¹, YU Xue-fen², LUO Zhao-wen³

(1. Dept. of Computer Science, Taizhou University, Linhai 317000; 2. The City Middle School of Huangyan, Taizhou 318029;

3. The Sanjia Middle School of Jiaojiang, Taizhou 318000, China)

Abstract Can be used to solve the difficult problems like global optimization or discrete variables. Simulated annealing algorithm, superior to traditional deterministic optimization methods with the speed of computer rising, it has a magnificent prospects and an invaluable effect on research work. In this paper, the research developments of SA are surveyed, including the basic principles of SA summarized in the first section, and then the history of development and characteristics for SA, a survey of internal and external research and the trend of development. On basis of comparison between SA and traditional optimization methods, it has been pointed out that SA can be applied to problems which are difficult to deal with by the traditional deterministic optimization methods.

Key words simulated annealing; global optimization; traditional optimization methods; implicit parallelism

(责任编辑: 耿继祥)