

模拟退火算法及其应用

作者：周思阳 学号：1500012628 指导老师：陈斌

【摘要】模拟退火算法的思想最早在 1953 年由 Metropolis 提出,在 1983 年被 Kirkpatrick 等人成功引入组合优化领域。由于它具有很强的实用性和极佳的性能表现,迅速引起了很多专家学者的兴趣,不断对其进行研究。本文阐述了模拟退火算法的提出背景和原理步骤,并从文献角度展示了近年来较为成功并被广泛采用与地球物理学反演过程的对模拟退火算法进行的改进,旨在凸显该算法在地球科学研究中的重要作用。

【关键词】模拟退火 反演

1 模拟退火算法的提出与原理

1.1 模拟退火算法的提出

模拟退火算法是局部搜索算法的扩展,从理论上来说,它是一个全局最优算法。

模拟退火算法的思想来源于对固体退火降温过程的模拟。即将固体加温至充分高,再让其徐徐冷却。在加热固体时,固体中原子的热运动不断增强,内能增大,随着温度的不断升高,固体的长程有序被彻底破坏,固体内部粒子随温度的升高而变为无序状。冷却时,粒子逐渐趋于有序,在每个温度下都达到平衡状态,最后在常温下达到基态,同时内能也减为最小。

在实际应用中,我们可以将内能 E 模拟为目标函数值 f ,将温度 T 模拟为控制参数,然后从一给定解开始,从其邻域中随机产生一个新解,接受准则允许目标函数在一定范围内接受使目标函数恶化的解,算法持续进行“产生新解—计算目标函数差—判断是否接受新解—接受或舍弃”的迭代过程,对应着固体在某一恒定温度下趋于热平衡的过程。经过大量的解变化后,可以求得给定控制参数 T 值的时候优化问题的相对最优解。然后减小控制参数 T 的值,重复执行上述迭代过程。当控制参数逐渐减小并趋于零时,系统也越来越趋于平衡状态,最后系统状态对应于优化问题的整体最优解。退火过程由一组称作冷却进度表(Cooling schedule)的参数控制,包括控制参数的初始值 T 以及衰减因子 Δt 、每个 T 值时的迭代次数(称为一个 Markov 链的长度) L 和终止条件 S 。

1.2 模拟退火算法的步骤

1)给定模型每一个参数变化范围,在这个范围内随机选择一个初始模型 m_0 ,并计算相应的目标函数值 $E(m_0)$;

2)对当前模型进行扰动产生一个新模型 m ,计算相应的目标函数值 $E(m)$,得到:

$$\Delta E = E(m) - E(m_0)$$

3)若 $\Delta E < 0$,则新模型 m 被接受;若 $\Delta E > 0$,则新模型 m 按概率 $P = \exp(-\Delta E/T)$ 进行接受, T 为温度。当模型被接受时,置 $m_0 = m$;

4)在温度 T 下,重复一定次数的扰动和接受过程,即重复步骤 2)、3);

5)缓慢降低温度 T ;

6)重复步骤 2)、5),直至收敛条件满足为止。

为了避免最好的解在优化过程中被忽略掉,可以稍做改进,即在整个搜索过程中随时记下最好的值,因为该法的特点决定了最后的优化值并不一定就是最优的值,而只能是较优值。

2 基于快速模拟退火算法的改进

2.1 改进目的

人类对地球内部物理性质(包括速度、密度、电导率、温度等)以及矿产资源分布的了解,大多来自地表地质和地球物理、地球化学资料的反演和解释。反演方法可以分为线性反演和非线性反演两种,线性反演已成为一套科学的反演理论,然而,绝大部分地球物理问题都是非线性的,并且实践表明,线性反演方法有容易陷入局部极值和依赖于初始值等缺点。因此,地球物理学者们不断的尝试开发非线性反演方法,比如人工神经网络、小波多尺度反演、模拟退火算法等。

然而,由于模拟退火法是建立在随机搜寻方法的基础上,要达到一定的精度要求,每一模型参数的离散点必须足够大。每次迭代必须进行多次目标函数计算,因而在处理实际资料时计算效率不高,影响着它的广泛应用。为了提高模拟退火算法的计算效率,出现了许多改进的方法,如采用依赖温度的 Cauchy 或似 Cauchy 分布代替常规模拟退火方法中的高斯分布产生新模型等。

2.2 算法改进

2.2.1 模型扰动

模拟退火算法中新模型的产生是对当前模型进行扰动得到的,这个扰动是用随机函数控制的。Ingber 于 1989 年提出了速度非常快的模拟退火算法 VFSA(Very Fast Simulated annealing),在该算法中,采用了依赖于温度的似 Cauchy 分布产生新的模型,即(2)式:

$$\begin{aligned} x'_i &= x_i + y_i (x_{i\max} - x_{i\min}) \\ y_i &= T_k \operatorname{sgn}(\xi - 0.5) \\ &\quad \times [(1 + 1/T_k)^{|2\xi - 1|} - 1] \end{aligned}$$

式中, x_i 为当前模型参数, x'_i 为扰动后的模型参数, ξ 为(0,1)区间上的随机数, sgn 为符号函数, y_i 称之为扰动因子。

在快速模拟退火算法的改进扰动策略基础上,有一种强化局部搜索能力的算

法,该算法借鉴了遗传算法中的非均匀变异思想,用非均匀变异策略对当前模型参数扰动产生新的模型参数,即(3)式:

$$x_i' = x_i + y_i (x_{imax} - x_{imin})$$

$$y_{ic} = r(1 - t/N)^\lambda \text{sgn}(r - 0.5)$$

其中, r 为(0,1)之间的随机数, t 为当前温度, N 为最大迭代次数, N 与最大温度、最低温度有关, λ 是确定非均匀性程度的常数,也称之为形状因子。(3)式表明,在高温时期,搜索的范围比较广,随着温度的降低,逐渐缩小搜索范围。

蒋龙聪,刘江平进一步分析了当 $\lambda=0.5,1,2,5,10$ 时相应的 y_{ic} 值,指导实际计算中如何取值,并且跟(2)式 y_i 做了对比分析(图 1)。

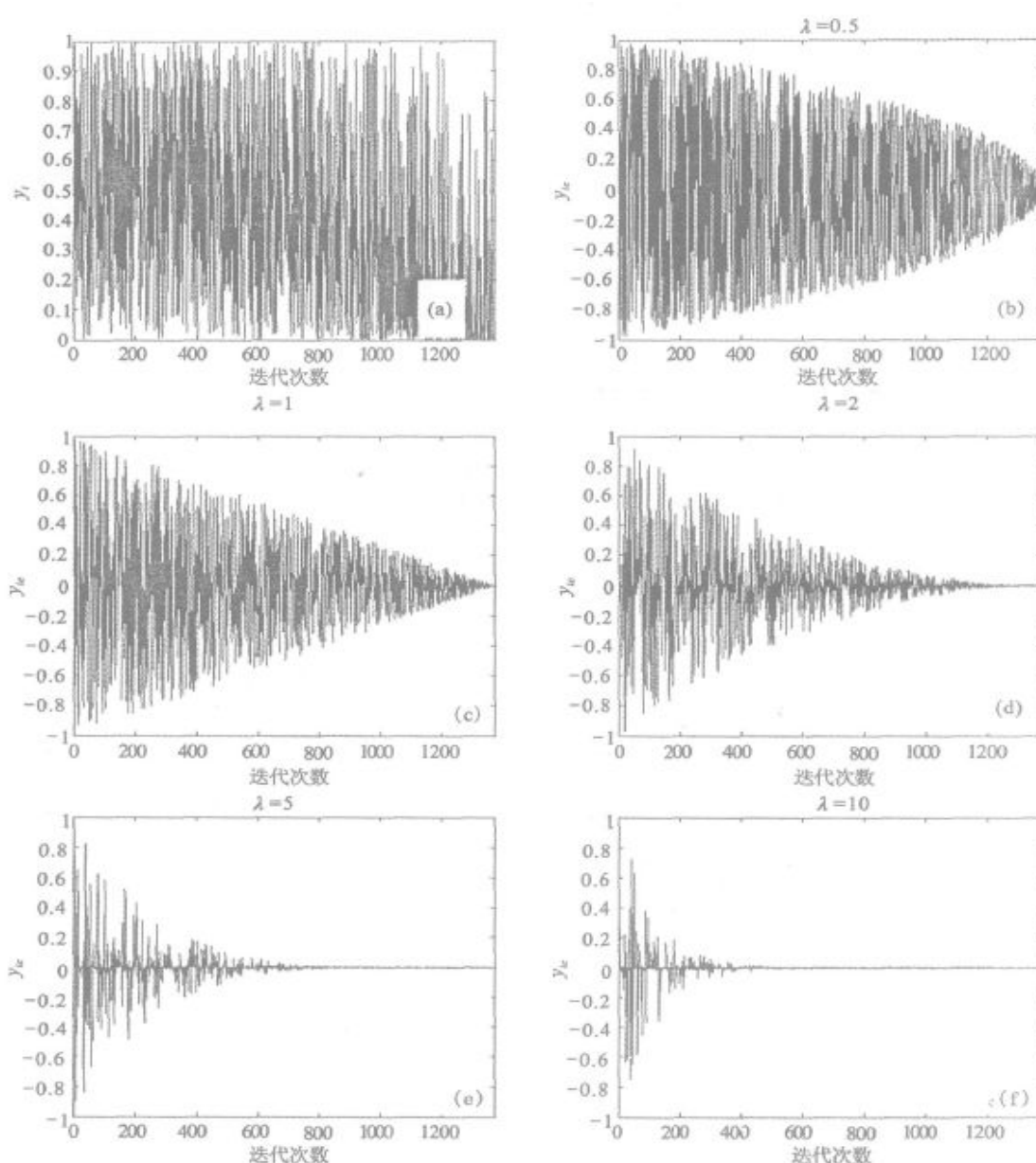


图 1 扰动因子随迭代次数变化的曲线
 (a) 为常规模拟退火的扰动; (b)、(c)、(d)、(e)、(f) 为 $\lambda=0.5, 1, 2, 5, 10$
 Fig. 1 Perturbation factor via iterative number

从图 1 可以得出,随着迭代次数的增加(即温度的降低), y_{ic} 的值逐渐减小,也即缩小了搜索范围,加强了局部搜索的能力,常规的模拟退火算法和快速模拟退火算法都是随机进行大范围的搜索,而且,当形状因子 λ 越大,越容易进行局部搜索,因此,在实际计算中,要合理地控制好 λ 值, λ 太大容易陷入局部极值中,Michalewicz 的建议表明是 λ 取[2,5]是比较好的。随着温度的降低,新接收的模型越来越接近真实的求解模型,因此,加强局部搜索更有利于提高算法的收敛性能。

2.2.2 接收概率

对于接受概率,快速模拟退火算法采用广义 Gibbs 分布产生接收概率 P ,即

$$P = [1 - (1 - h) \Delta E / T]^{1/(1-h)} \quad (4)$$

式中 ΔE 表示能量差, T 表示温度, h 为一实数。按照(4)式计算得到的概率进行判断是否接收新模型,该式只考虑了能量的绝对变化,然而,物体由熔化状态逐渐冷却至结晶状态所出现的相变规则,必须考虑能量的相对变化,为此顾汉明等[12]对此进行了修正,即:

$$P = [1 - (h \Delta E / T)(E_1^\alpha / E^\beta)]^{1/h} \quad (5)$$

式中, E 为新模型的目标函数, E_1 为当前模型的目标函数, α, β 为非负实数。

2.2.3 降温方式

对于退火方案,实践表明, T 取指数变化的模式比较切合退火的实质。Ingber 给出的快速退火算法中的退火方式为

$$T(K) = T_0 \exp(-CK^{1/N}) \quad (6)$$

式中: T_0 为初始温度, K 为迭代次数, C 为给定的常数, N 为待反演的参数的个数。该式子也可以改成

$$T(K) = T_0 \alpha^{K^{1/N}} \quad (7)$$

通常选择 $0.7 \leq \alpha \leq 1$,在本文中,用 1 代替 $1/N$, α 取为 0.99。

在改进的算法中 RSA,采用(3)式对当前模型扰动产生新的模型,新模型按照(5)进行接收,按照(7)式缓慢降温。

2.3 对该算法改进的总结

1)改进的模拟退火算法相对于常规的模拟退火算法,加强了局部搜索功能,更有利于提高算法的收敛速度,数值优化结果证实改进算法的有效性和高效性。

2)对于改进的算法, λ 的选择至关重要, λ 太大容易陷入局部最小值, λ 太小则降低了收敛速度,因此要通过测试再决定取 λ 的值。

3)从本质上讲,模拟退火算法属于蒙特卡罗方法,它用随机的方法代替了线性反演中的步长,相对于线性反演来讲,模拟退火算法的执行效率还是远远不及线性反演,因此,如何考虑将线性反演与非线性反演结合起来,相互利用对方的优势,并且用于实际反演中将是接下来重点研究的对象。

参考文献

庞峰.2006.模拟退火算法的原理及算法在优化问题上的应用

蒋龙聪,刘江平.2006.模拟退火算法及其改进