计算智能——模拟退火算法

作者 柯良军

西安交通大学 电信学院

October 21, 2015

目录

热力学中的退火过程

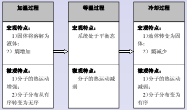

图: 退火的三个过程

退火过程是一个典型的物理过程。它包括加温过程、等温过程和冷却过程。金属物体在退火后会变得柔韧。图1给出了在每个过程中物体的宏观和微观变化特点。一般地，退火过程中的温度是逐渐下降的。如果高温金属物体的温度急剧下降，它没有达到平衡态，而处于非均匀的亚稳态，这就是淬火过程。经过淬火过程，物体能量并没有达到最小值，它能提高金属的强度和硬度，但会减弱韧度。

退火的启示

鉴于物理系统倾向于能量较低的状态，而热运动又妨碍它准确落到最低态的物理形态，采样时只需着重取那些有贡献作用的状态则可较快达到较好的效果。1953年，Metropolis等人受到蒙特卡罗模拟方法的启示提出了一种重要性采样法，即以概率接受新状态。在温度t，从当前状态i 的邻域随机产生新状态j，两者的能量分别为E_i 和E_j，若$E_i > E_j$，则接受新状态j 为当前状态；否则，依据以下概率

$$p_t = \exp\left(\frac{-(E_j - E_i)}{kt}\right) \tag{1}$$

来接受状态j，其中k 为Boltzmann常数，$\exp(\cdot)$ 为指数函数。当这个过程多次重复，即经过大量迁移后，系统将趋于能量较低的平衡态。各状态的概率分布将趋于一定的正则分布。这种接受新状态的方法被称为Metropolis 准则，它能够大大减少采样的计算量。

退火的启示

由Metropolis准则可知：
1. 高温下可接受与当前状态能量差较大的新状态
2. 在低温下只能接受比当前能量低的新状态或与当前能量差较小的新状态。这与不同温度下的热运动影响金属原子重新排列的过程一致。
3. 在温度趋于零时，就不能接收任一个$E_j > E_i$的新状态j了

退火的启示

表：模拟退火算法求解过程与物理退火过程的对应关系

模拟退火算法	物理退火
解	状态
目标函数	能量函数
初始状态	高温状态
温度降低	降温过程

基本思想

将待优化问题的可行解看作是物体的原子状态，将目标函数看作是物体原子的能量函数，模拟物体退火过程中粒子的热运动。降温过程中结合概率突跳特性的Metropolis抽样准则在解空间中随机搜索全局最优解，在陷入局部最优时，能以一定概率跳出最终趋于全局最优解。

要素

- 状态表示
- 邻域定义
- 热平衡过程
- 降温控制

状态表达

在模拟退火算法求解优化问题时，一个状态对应于一个解。为了有效求解优化问题，要采用合适的编码来表示解或状态。

状态表达

在模拟退火算法求解优化问题时，一个状态对应于一个解。为了有效求解优化问题，要采用合适的编码来表示解或状态。

例如，在求解背包问题时，可采用0-1编码；而求解旅行商问题时，可采用顺序编码；而对于连续函数优化，可以采用实数编码。

邻域定义

模拟退火算法从一个解出发，探索其邻域寻找改进解。邻域决定搜索范围。模拟退火算法的邻域定义与局部搜索中的定义一样。例如在车辆路径问题中，一种邻域定义是采用一系列的k-边交换操作作为邻域，k表示交换边的个数，k 越大，其邻域解与当前解差别越大。

邻域定义

模拟退火算法采用一种特殊的Metropolis准则的邻域移动方法。如果一个新解优于当前解，当前解被新解替换，那么就称从当前解移动到这个新解，否则，依据一定的概率来决定当前解是否移向新解。Metropolis准则状态转移概率P_{ij}定义如下：

$$p_{ij} = \begin{cases} 1 & \text{如果} f(j) \leqslant f(i) \\ exp(-\frac{f(j)-f(i)}{T}) & \text{其他} \end{cases} \qquad (2)$$

其中i 为当前状态，j 为其邻域中的一个状态，它们的目标函数值分别为$f(i)$和$f(j)$，当前温度为T。可见，当抽样得到的邻域新解优于当前解时，无条件移动；当新解劣于当前解时，以一定概率移动。

邻域定义

从Metropolis准则中可以看出，当T 很大时，状态转移概率趋于1，当前邻域中的任一状态都可能被接受，此时算法正在进行全局搜索；而当T 很小时，状态转移概率趋于0，它只会接受当前邻域中的更好的状态，此时算法进行局部搜索。因此模拟退火算法既具备跳出局部优解的能力又具备探索全局最优的能力。

热平衡过程

工业退火要求温度缓慢下降，以保证金属原子在每个温度下达到稳定的能量态。但是在实际应用中达到理论的平衡状态是不可能的。模拟退火算法采用在每一个温度，在邻域中寻找一定数量的解的，这一循环称为算法的内循环。内循环次数应该尽可能地大，以达到近似热平衡过程。内循环次数的选择与问题的解空间大小有关。当解空间较大时，可以将内循环次数设为较大的数。另外，也可以依据其他条件来动态调节内循环次数。例如高温时迭代次数减少，低温时迭代次数增加。

降温函数

模拟退火算法的搜索能力与退火速度（温度的降低速度）密切相关。在高温状态下，当前邻域中几乎所有的解都会被接受，算法进行全局搜索；当温度变低时，当前邻域中越来越多的解将会被拒绝，算法进行邻域搜索。在同一温度下，需要对邻域进行充分地搜索以达到热平衡状态。如果温度下降速度太快，则可能错过最优解，过早地陷入局部最优；如果温度下降速度太慢，又可能会降低算法的收敛速度。因此，降温函数的选择对于模拟退火算法的性能有重要影响。

降温函数

在邻域搜索过程中，如果解的质量变差的概率呈Boltzmann分布，理论上可以证明对数降温方式可使模拟退火算法收敛于全局最优解：

$$T(k) = \frac{t_0}{\lg(1+k)} \tag{3}$$

上式中，k 为降温的次数，$T(k)$ 代表t 次降温后的温度，t_0 为初始温度，它为正的常数，$\lg(\cdot)$ 为以十为底的对数函数。在经典模拟退火算法中采用这一降温函数。

降温函数

在邻域搜索过程中，如果解的质量变差的概率呈Cahchy 分布时，理论可以证明按如下降温方式能保证模拟退火算法收敛于全局最优解：

$$T(t) = \frac{k}{1+t} \tag{4}$$

上式中，k 为降温的次数，$T(k)$ 代表k 次降温后的温度，t_0 为正的常数。在快速模拟退火算法中采用这一降温函数。

降温函数

另外常用到的降温函数还有指数降温函数$T(k+1) = \tau \cdot T(k)$，其中τ 为指数下降比率和直线降温函数$T(k+1) = T(k) - \Delta T$，其中ΔT 温度下降步长。

降温函数

除了上述各要素之外，初始温度的选择对模拟退火算法的性能也会有很大的影响。一般来说，初始温度T_0选择充分的大，保证算法在开始时能够处在一个平衡状态，同时获得高质量解的概率就大。但花费的计算时间将会增加。因此初始温度应该选择足够的高，同时兼顾优化质量和优化效率。在实际应用中，初始温度的选择选择方法主要有以下几种：

（1）均匀地抽样一组状态，选各状态目标值的方差作为初始温度。

降温函数

（2）随机产生一组状态，计算最大的目标函数值与最小的目标函数值的差值Δf，然后依据一定的规则计算初始温度。例如，可依据下式得到：

$$T_0 = -\frac{\Delta f}{\ln(p)} \tag{5}$$

上式中，$\ln(\cdot)$ 为自然对数，p为初始接受概率。也可采用下式得到：

$$T_0 = k\Delta f \tag{6}$$

上式中，k为常数充分大。

降温函数

（3）根据经验公式给出初始温度。

算法步骤

给定优化问题$\min_{i \in S} f(i)$，模拟退火算法主要步骤如下：

$step1$：设定初值，随机选一个可行解作为初始解$i = i_0$，按照上一节初始温度选择方法设定初温$T = T_0$ ，设定最大迭代次数k_{max} 作为降温过程的终止条件，初始化迭代计数器$k = 0$。

$step2$：从i 的邻域$N(i)$中随机产生一个新解j ，根据目标函数分别计算当前解i 和新解j 对应的$f(i)$ 和$f(j)$，并计算差量$\Delta f = f(j) - f(i)$。

算法步骤

$step3$：如果$\Delta f < 0$ 或$\Delta f > 0$ 且$\exp\left(\frac{-\Delta f}{T}\right) > random(0,1)$，那么接受新解$j$，令$i = j$。

否则不接受新解，当前解保持不变。

$step4$：更新内循环次数$n = n + 1$，若$n > n_{max}$，热平衡过程，其中n_{max} 为内循环最大次数，转$step5$；否则转$step2$。

$step5$：令$k = k + 1$，若$k > k_{max}$,则算法停止，输出最优解i 和其目标函数$f(i)$；否则根据降温函数降低更新当前温度T，转$step2$。