

优化问题的新算法

罗页

在解决科学问题时，我们常常需要求解最大或最小值问题，即最优化问题。目前解决最优化问题的方法主要有两大类：决定性方法和随机性方法。决定性方法虽然很快，但是很容易陷入局域极值，从而无法遍历整个相空间得到全局最小。随机性方法却恰恰相反，虽然不易陷入局域极值，可是它不能在有限时间内保证找到全局最小值。在本文解决 Thomson 问题中，同时用到了随机性方法——模拟退火方法和决定性方法——共轭梯度法，从而得到了比较好的全局最值。

模拟退火算法(SA)原理

模拟退火的基本思想来源于固体退火过程中，这里面的物理图像很明晰。在固体从液体或气体开始冷却时，院子的热运动逐渐减弱，随着温度的降低，原子运动渐趋于有序。当温度降低至相变温度后，原子运动被束缚在晶格点上做微小振动，液体凝固成固体。这种从高温降温至低温的过程称为退火。整个过程中熵减小，系统能量也随着温度同步降低趋于最小值。

基于固体退火的启示，Kirkpatrick 等人首先意识到固体退火过程与优化问题之间的相似性。他们把 Metropolis 算法引入优化过程，得到迭代算法，称为“模拟退火算法”。其过程为：

- 1、初始时刻在温度 $T(t=1)$ ，选取一个初始态 x_t ，计算 $E(x_t)$
- 2、根据某种跃迁分布，随机从 x_t 跳到 x_{t+1}
- 3、计算 $E(x_{t+1})$ ，如果 $E(x_{t+1}) < E(x_t)$ ，则用 x_{t+1} 取代 x_t ，否则计算接受率 $p = \exp\{-[E(x_{t+1}) - E(x_t)]/T(t)\}$ 和一个 $(0,1)$ 之间的随机数 r ，如果 $p > r$ ，就用 x_{t+1} 取代 x_t ，否则保持不变。
- 4、根据降温公式，计算下一步的 $T(t+1)$ ，回到(2)重复以上过程，直到 E 达到某个预定的精度范围。

其中的第(3)步就是典型的 Metropolis 算法的步骤，接受率 p 即是根据 Boltzmann 因子从一个态跳到另一个态的条件概率。

不过模拟退火算法并没有，规定跃迁分布、接收方式和降温方式。经典的模拟退火算法，其跃迁分布是满足高斯形式的，接收方式是 Boltzmann 统计形式，而降温方式是指数形式的。

模拟退火算法能将时间复杂度 $O(e^{\beta N})$ 的问题简化为 $O(\alpha N^\beta)$ 。

推广的模拟退火算法(GSA)

推广的模拟退火算法基于 Tsallis 统计，相比经典的模拟退火算法，其降温方式、跃迁几率和降温方式都有所改动。

1)接收率为

$$P_{q_a}(X_t \rightarrow X_{t+1}) = \begin{cases} 1 & E(X_{t+1}) - E(X_t) < 0 \\ \frac{1}{\left[1 + (q_a - 1)(E(X_{t+1}) - E(X_t))/T^{\alpha}(q_a)\right]^{\frac{1}{q_a-1}}} & E(X_{t+1}) - E(X_t) \geq 0 \end{cases}$$

2)跃迁分布为

$$g_{q_v}(\Delta X_t) = \left(\frac{q_v - 1}{\pi}\right)^{D/2} \frac{\Gamma\left(\frac{1}{q_v - 1} + \frac{D-1}{2}\right)}{\Gamma\left(\frac{1}{q_v - 1} - \frac{1}{2}\right)} \frac{[T_{q_v}^v(t)]^{\frac{D}{3-q_v}}}{\left\{1 + (q_v - 1) \frac{(\Delta X_t)^2}{[T_{q_v}^v(t)]^{\frac{2}{3-q_v}}}\right\}^{\frac{1}{q_v - 1} + \frac{D-1}{2}}}$$

3)降温方式为

$$T_{q_v}^v(t) = T_{q_v}^v(1) \frac{2^{q_v-1} - 1}{(1+t)^{q_v-1} - 1}$$

q_a 为控制接收率的参量, q_v 为控制降温的一个参量, 为了简单起见, $T_{q_v}^v = T_{q_a}^a$ 。整个步骤和经典模拟退火无异。

Thomson 问题

本问题源于物理学家 J.J.Thomson 在本世纪初提出的原子模型。在一个球面上如何排布电子才能使得体系的能量最低, 这就是 Thomson 问题。该问题被很多科学家研究, 但是解析方面的工作很少, 绝大部分是数值方法。

Thomson 问题的能量形式为

$$E = \frac{1}{2} \sum_{i \neq j} \frac{1}{|r_i - r_j|}$$

Erber 等人指出, Thomson 问题的局域最小数量随 $\exp(cN)$ 增长, 且全局最小的窗口很小, 因此, Thomson 问题被认为测试优化方法的理想体系。

Thomson 问题的解已经被很多优化方法求解, 如最速下降、蒙特卡洛、模拟退火、遗传算法等搜索过。最速下降很容易给出局域极小, 无法跳出该极小寻找下一个极小。而 Monte Carlo 和模拟退火虽然能跳出极小去寻找全局极小, 但是花费的时间很长, 所以也不实际。

共轭梯度法(CG)

鉴于推广模拟退火算法是一种随机数算法, 主要用于寻找全局最小的位置, 其精度只告诉我们全局最小大概在什么位置, 如果要得到精确值, 时间必须无限长。说以我们可以在该位置附近进行搜索, 加速收敛。这里, 我们使用共轭梯度法。

共轭梯度法属于决定性算法, 相比于最速下降法, 其每次选取共轭的方向来搜索极小值, 能够避免在试探步长时浪费时间, 而且其只需要消耗 N 的空间, 选取适当方向经过 N 步之后即可到达极小。共轭梯度法的时间复杂度 $O(\alpha N^2)$

Thomson 问题计算结果

用推广模拟退火配合共轭梯度，计算了 N 在 156-190 的数值，对照文献中的结果

| N | E _{min} | error | N | E _{min} | error |
|-----|------------------|----------|-----|------------------|----------|
| 156 | 11092.79974883 | 1.30E-07 | 173 | 13708.63765790 | 1.76E-07 |
| 157 | 11238.90447665 | 1.28E-07 | 174 | 13871.18957134 | 1.79E-07 |
| 158 | 11385.99177456 | 1.40E-07 | 175 | 14034.78419441 | 2.06E-07 |
| 159 | 11534.02543517 | 1.28E-07 | 176 | 14199.35776813 | 2.11E-07 |
| 160 | 11683.05640956 | 1.37E-07 | 177 | 14364.84099383 | 7.48E-02 |
| 161 | 11833.08629355 | 1.31E-07 | 178 | 14531.31232695 | 1.91E-07 |
| 162 | 11984.05215346 | 1.52E-07 | 179 | 14698.75737754 | 1.89E-07 |
| 163 | 12136.01503700 | 1.63E-07 | 180 | 14867.10400368 | 2.74E-07 |
| 164 | 12288.93277898 | 2.18E-07 | 181 | 15036.47044551 | 2.13E-07 |
| 165 | 12442.80698535 | 2.04E-07 | 182 | 15206.73304816 | 1.60E-07 |
| 166 | 12597.65115198 | 1.65E-07 | 183 | 15378.16936105 | 1.81E-07 |
| 167 | 12753.47223832 | 2.20E-07 | 184 | 15550.42761609 | 3.97E-07 |
| 168 | 12910.21494786 | 1.76E-07 | 185 | 15723.72674394 | 4.24E-07 |
| 169 | 13068.00879690 | 1.80E-07 | 186 | 15897.90010426 | 1.68E-07 |
| 170 | 13226.68370480 | 1.99E-07 | 187 | 16072.97799709 | 1.75E-07 |
| 171 | 13386.36028466 | 3.25E-07 | 188 | 16249.22618307 | 2.16E-07 |
| 172 | 13547.02100758 | 2.14E-07 | 189 | 16426.37508713 | 1.92E-07 |
| | | | 190 | 16604.43108024 | 1.65E-07 |

图中 error 栏表示与参考文献中的数据的差，只要有足够多的尝试计算，可以达到文献中的数值，其中 156 和 190 的能量值已经显著低于采用遗传算法计算得到的数值。

如果只采用推广模拟退火，计算结果的 error 在 1E-5，通过共轭梯度减小到 1E-7。

求解过程的分析

我们为了尽快达到最小值，必须选择合适计算参数，从而平衡好计算精度和计算时间。

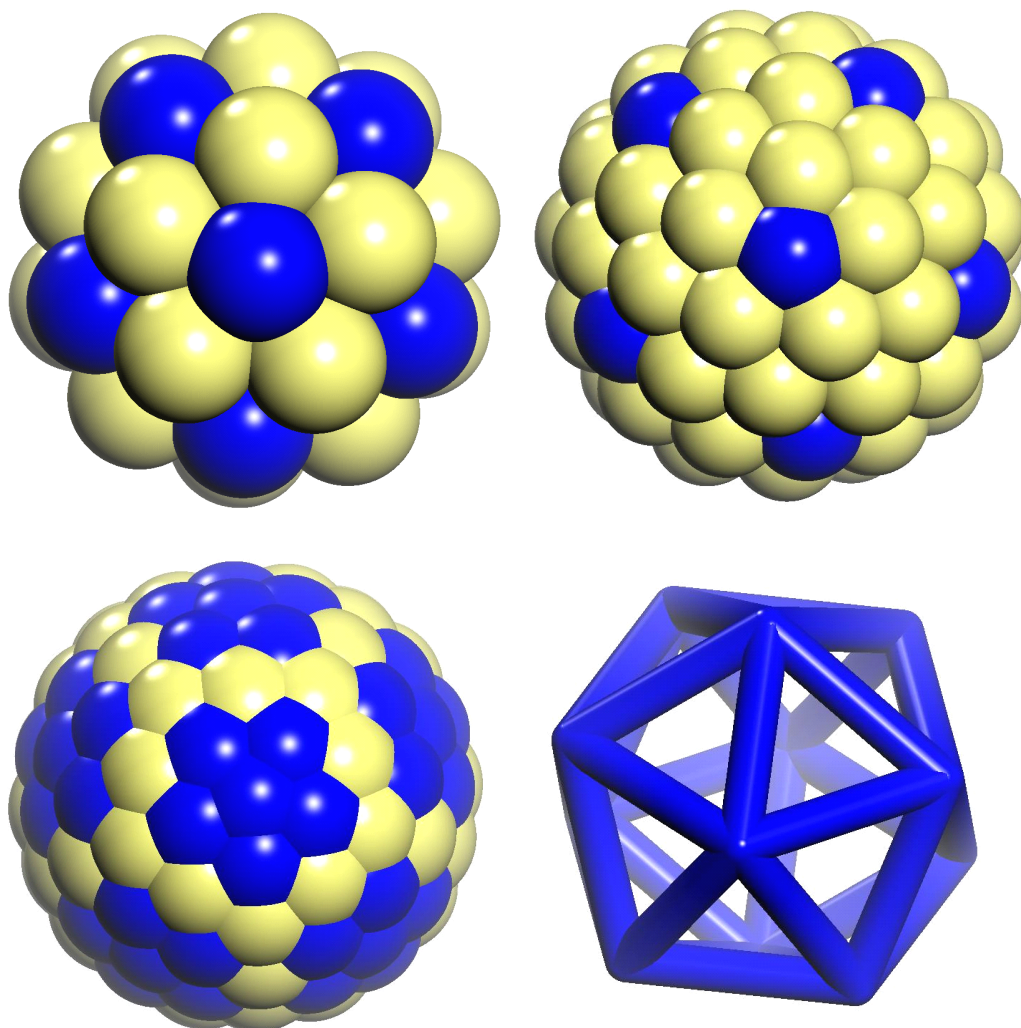
1、由于在低温下，电子构型变化的变得非常缓慢，跳出某个极小的时间变得非常长，所以最终终止的温度也无需很低，过低的温度只会浪费计算时间。

2、选择 $q_a = -3 - 0.85t$ ， $q_v = 2.62$ 。如果 q_v 很小，会导致降温很慢，计算费时。

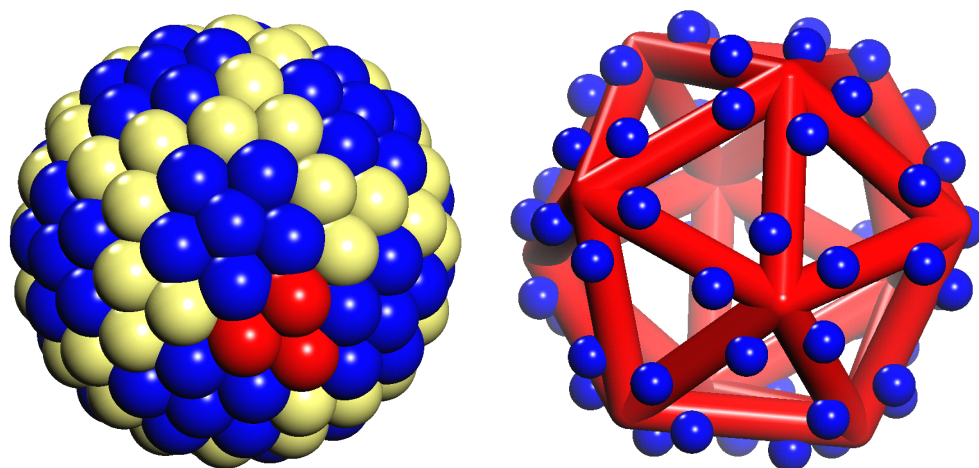
3、在低温下，采用推广模拟退火电子大部分时间只在深谷里徘徊，有时会接近该深谷的极小值，但仍然在同一个深谷里，所以，这些接近极小值的构型只需要进行一次共轭梯度的搜索，无需重复搜索。即，对于 $E(x_{t+1}) < E(x_t)$ ，如果差值不大，无需进行共轭梯度。

能量最小构型的对称性分析

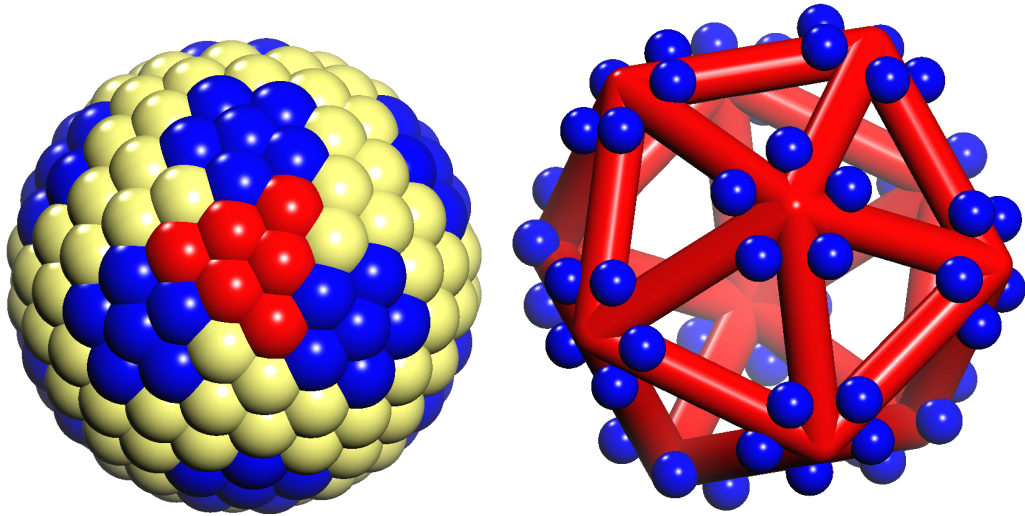
32、72、122 个电子，如图。它们都具有正 20 面体结构，蓝色球构成正 20 面体，蓝色球都是 5 近邻，黄色球都是 6 近邻。



132 个电子时，如图。蓝色电子以 5 个成团，一共 12 个团，构成正 20 面体。团之间以三角形的散点子组填充。有很高的对称性。



192 个电子时，如图。蓝色电子以 5 个成团，一共 12 个团，构成正 20 面体。团之间以 6 电子三角形填充。同样有很高对称性。



Altschuler 猜想 $N=10(m^2+n^2+mn)+2$ 的结构在具有正 20 面体构型具有最小值，这与本文的计算结果完全吻合。

结束语

推广的模拟退火算法是一种高效的优化算法，在包括计算机、生物、化学领域都用有了较大的用途。在发展方法方面，我们可以结合推广模拟退火算法和其他随机算法如遗传算法，以及决定性算法，从而得到更好的效果。