

分子建模与模拟导论

王延颀

2012年12月29日

13. 优化技术

13.1. 寻找局部最小值

13.1.1. 最陡下降法 (Steepest Decent Method)

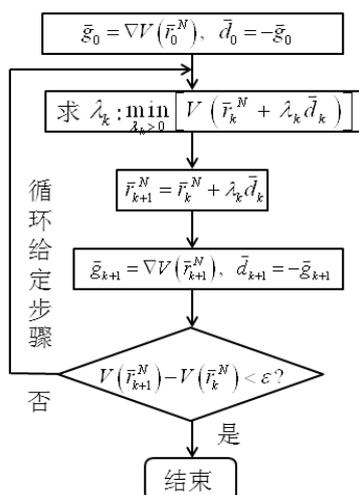
优化时每次都沿着下降最快的方向（即梯度方向）行走。设待优化函数为 $V(\vec{r}^N)$ ，梯度方向为 $\vec{g}(\vec{r}^N) = -\vec{F}(\vec{r}^N) = -\vec{\nabla}_{\vec{r}^N} V(\vec{r}^N)$ 。一维情况下做一次梯度计算就可以找到局部最小值。对于多维情形，每一步沿着梯度下降方向走 λ_k ：

$$\vec{r}_{k+1}^N = \vec{r}_k^N - \lambda_k \vec{\nabla} V(\vec{r}_k^N) = \vec{r}_k^N - \lambda_k \vec{g}(\vec{r}_k^N) \quad (13.1)$$

选择 λ_k 使得 $k+1$ 点的梯度与 k 点垂直：

$$\frac{d}{d\lambda_k} V(\vec{r}_{k+1}^N) = \nabla V(\vec{r}_{k+1}^N)^T \cdot \frac{d}{d\lambda_k} \vec{r}_{k+1} = -\vec{g}(\vec{r}_{k+1}^N) \cdot \vec{g}(\vec{r}_k^N) = 0 \quad (13.2)$$

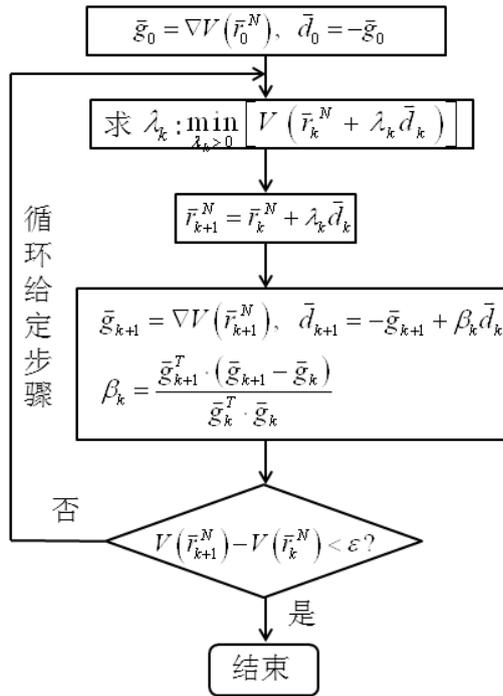
算法流程：



最陡下降法的优点是远离极值点时收敛很快，缺点是到极值点附近收敛很困难。

13.1.2. 共轭梯度法 (Conjugate Gradient Method)

每一次调整方向为两步梯度的矢量叠加，从而使得在极值点附近能较快且稳定地收敛。缺点在于在远离极值点处相对于最陡下降法收敛较慢。算法流程：



13.1.3. Newton-Raphson 方法

每一步用一个二次方程对给定函数 V 作近似，然后移动至此二次方程的极小值。待优化函数近似为

$$V(\bar{r}^N) \approx V(\bar{r}_k^N) + (\bar{r}^N - \bar{r}_k^N)^T \cdot \bar{g}_k + \frac{1}{2} (\bar{r}^N - \bar{r}_k^N)^T \cdot \bar{H}_k \cdot (\bar{r}^N - \bar{r}_k^N) \quad (13.3)$$

其中 $\bar{g}_k = \bar{\nabla} V(\bar{r}_k^N)$ ，Hessian 矩阵

$$\vec{H}_k = \nabla^2 V(\vec{r}_k^N) = \begin{bmatrix} \frac{\partial^2 V}{\partial x_1^2} & \frac{\partial^2 V}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 V}{\partial x_1 \partial x_n} \\ \frac{\partial^2 V}{\partial x_2 \partial x_1} & \frac{\partial^2 V}{\partial x_2^2} & \cdots & \frac{\partial^2 V}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 V}{\partial x_n \partial x_1} & \frac{\partial^2 V}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 V}{\partial x_n^2} \end{bmatrix} \quad (13.4)$$

这里用 (x_1, x_2, \dots, x_n) 代表 $(r_{1x}, r_{1y}, r_{1z}, \dots)$ 。每一步移动方向为 $\vec{d}_k = -\vec{H}_k^{-1} \cdot \vec{g}_k$ ，移动大小为

$\vec{r}_{k+1}^N = \vec{r}_k^N + \lambda_k \vec{d}_k$ ，其中 λ_k 用 backtracking 方法确定大小。

13.2. 寻找全局最小值

13.2.1. 模拟退火 (Simulated Annealing)

恒温的 MC 或 MD 模拟原则上总有几率找到全局最小，但是随着势垒的增加，几率指数减小。所以大多数情况下用基本的 MC 或 MD 无法使系统自发找到全局最小态。模拟退火是先在高温模拟，然后串行降温，直至给定温度。MC 或 MD 皆可。模拟退火的优点是实现非常简单，缺点是不能保证找到全局最小。

13.2.2. 副本交换 (Replica Exchange or Parallel Tempering)

副本交换方法同时在多个温度下运行 MC 或 MD，隔一段时间比较一下系统势能，根据 Metropolis Algorithm 决定是否交换构型（或者等价地，交换温度）。算法推导如下。

整个系统的广义配分函数为

$$Z_{\text{extended}} = \prod_{i=1}^m Z_i = \prod_{i=1}^m \frac{1}{h^{3N} N!} \int \exp(-\beta_i E_p(\vec{r}_i^N)) d\vec{r}_i^N \quad (13.5)$$

参照基本 MC 的算法，根据 Metropolis Algorithm 有

$$\begin{aligned} & P(i, \beta_i) \cdot P(j, \beta_j) \cdot \alpha \left[(i, \beta_i), (j, \beta_j) \rightarrow (j, \beta_i), (i, \beta_j) \right] \\ & \cdot \text{acc} \left[(i, \beta_i), (j, \beta_j) \rightarrow (j, \beta_i), (i, \beta_j) \right] \\ & = P(i, \beta_j) \cdot P(j, \beta_i) \cdot \alpha \left[(i, \beta_j), (j, \beta_i) \rightarrow (i, \beta_i), (j, \beta_j) \right] \\ & \cdot \text{acc} \left[(i, \beta_j), (j, \beta_i) \rightarrow (i, \beta_i), (j, \beta_j) \right] \end{aligned} \quad (13.6)$$

令 Markov 链的转换矩阵 α 相等，有

$$\frac{\text{acc}\left[(i, \beta_i), (j, \beta_j) \rightarrow (j, \beta_i), (i, \beta_j)\right]}{\text{acc}\left[(i, \beta_j), (j, \beta_i) \rightarrow (i, \beta_i), (j, \beta_j)\right]} = \frac{\exp\left[-\beta_i E_p(j) - \beta_j E_p(i)\right]}{\exp\left[-\beta_i E_p(i) - \beta_j E_p(j)\right]} \quad (13.7)$$

$$= \exp\left[(\beta_j - \beta_i) \cdot (E_p(j) - E_p(i))\right]$$

所以舍选条件为

$$\text{acc}\left[(i, \beta_i), (j, \beta_j) \rightarrow (j, \beta_i), (i, \beta_j)\right] = \min\left(1, \exp\left((\beta_j - \beta_i) \cdot (E_p(j) - E_p(i))\right)\right) \quad (13.8)$$

副本交换法的优点在于 1) 一般情况下可以保证找到全局最小态；2) 每个副本的 Boltzmann 分布依然成立。缺点在于当系统很大时势能会很大，低温下需要非常多的副本温度以使 E_p 的涨落有交叠，所以一般只能模拟比较小的体系。

13.2.3. 基因算法 (Genetic Algorithm)

给定一个序列，按一定规则随机地改变序列元素的值，并按评判函数决定如何舍选新产生的序列。

举例：对于一个有 6 个二面角的短肽分子寻找其全局最小构型。选取评判函数为构型所对应的势能函数的高低，越低的认为越好。每次对二面角的角度做随机组合并对角度产生小的扰动变化，每次根据评判函数留下一组好的构型。

基因算法的优点是 1) 在接近最优解处比较稳定；2) 易于实现。缺点在于 1) 理论框架不够严谨；2) 最后步骤需要其它算法帮助最终收敛到最优解处。

13.3. 寻找反应路径 (Locating Reaction Path)

13.3.1. Nudged Elastic Band

针对有反应势垒的情况，给定初态和终态以及猜想的中间态，对连接各态的“弹簧力”进行优化，直到力的垂直分量为 0，平行分量使构型间保持一定距离。平衡后的构型给出的就是局部最优的反应路径。等效于在路径空间寻找反应路径的局部最小值。

13.3.2. Transition Path Sampling

用 **Metropolis Algorithm** 在路径空间对路径进行舍选，从而原则上可以在路径空间采样到不同路径的几率分布。