Lecture 4 - Metropolis algorithm. Classical Monte Carlo method. Simulated annealing

We have learned that importance sampling permits to decrease the variance in estimation of the value of the integral of the type:

$$\underbrace{\int f(R)\,P(R)\,dR}_{\text{integral of interest}} \simeq \frac{1}{M}\sum_{i=1}^{M} f(R_i)$$

real

some function
positive/negative/complex

non-negative

where
- $R = \{x_1, \dots, x_N\}$ is a point in a N-dimensional space
- $P(R)$ is the probability distribution function
- $f(R)$ quantity which is averaged
- $R_1, \dots, R_M$ points in N-dimensional space distributed according to $P(R)$

- For some simple PDFs it is easy to generate random values with appropriate distribution (i.e. uniform, Gaussia, etc)
- what about the possibility to generate an arbitrary non-uniform distribution $P(R)$?
- This can be achieved by generating a Random Walk with a specific transition probability $T(R_i \rightarrow R_j)$
- It can be shown that it is sufficient

- It can be shown that it is sufficient (not necessary) to satisfy the detailed balance condition

$$p(R)\,T(R \to R') = p(R')\,T(R' \to R)$$

- The detailed balance condition does not specify $T(R \to R')$ in a unique way

- A simple choice is

$$T(R \to R') = \min\left[1, \frac{p(R')}{p(R)}\right]$$

Metropolis algorithm:

- generate a trial position R_i' from the present position R_i
- calculate relative weight $\omega = \dfrac{p(R_i')}{p(R_i)}$
- generate a random number $0 < u < 1$ with a uniform PDF
- evaluate $\mathrm{int}(\omega + u)$

* if $\mathrm{int}(\omega + u) > \emptyset$ then $\boxed{\text{accept}}$ the

$\mathrm{int}(\omega) = 1; 2; 3; \ldots$

move $R_{i+1} = R_i'$

* if $\mathrm{int}(\omega + u) = \emptyset$ then $\boxed{\text{reject}}$ the move $R_{i+1} = R_i$

- repeat

- repeat

NB 1 trial position can be generated by displacing R by

a) $\xi = (2u-1)\cdot \Delta t$ with $0 < u < 1$

a uniformly distributed random variable

in interval $(-\Delta t; \Delta t)$

b) ξ drawn from a normal probability distribution

$$p(\xi) = e^{-\frac{\xi^2}{2(\Delta t)^2}}$$

where the amplitude of displacement Δt is a free parameter

$x_i' = x_i + \xi$ for each degree of freedom

NB 2 It is convenient to calculate the acceptance ratio P_{acc}

$$P_{acc} = \frac{N_{acc}}{N_{total}} \; ; \quad 0 \le P_{acc} \le 1$$

N_{acc} – number of accepted moves

N_{total} – number of total moves

- if $P_{acc} \lesssim 0.1$ then acceptance is too low $\Rightarrow$ decrease Δt (inefficient sampling)

- if $P_{acc} \geq 0.9$ the acceptance is too high $\Rightarrow$ increase Δt (inefficient sampling)

NB 3 Advantages of Metropolis algorithm

- can be used to sample an arbitrary probability distribution
- it is not necessary to know its normalization
- works in arbitrary number of dimensions

Disadvantages

- the obtained Random Process is strongly correlated
- estimation of the statistical error as $\frac{\sigma}{\sqrt{M}}$ underestimates the real error. Here σ is the variance of the observable and M is the number of measurements. This happens as the number of independent measurements $M^* < M$ which increases the error $\frac{\sigma}{\sqrt{M^*}} > \frac{\sigma}{\sqrt{M}}$

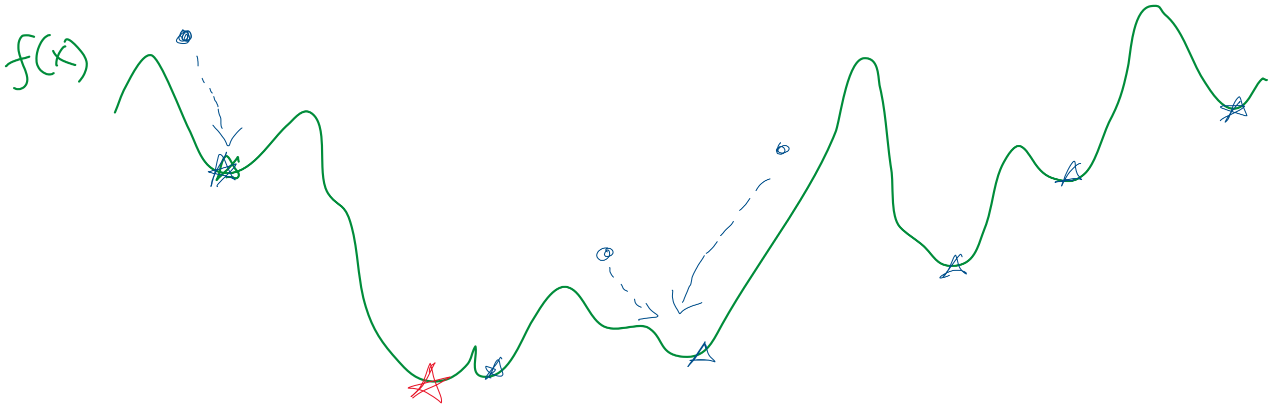
NB 4 If the initial configuration R

NB 4 If the initial configuration R has a low probability $P(R)$, the method will asymptotically converge to the desired probability distribution, but the first iterations have to be thrown away

Simulated annealing method

Simulated annealing method

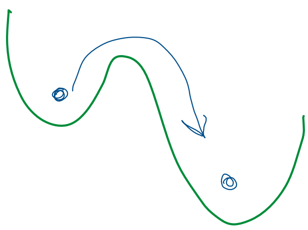
- Problem of optimization → finding the global optimum of a given function
- Generally, the function can have a large number of variables / degrees of freedom
- Newton steepest descent method can be inefficient



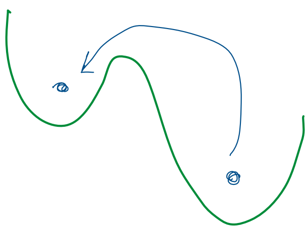
Global minimum

Local minima

- number of local minuma can be large
- a better efficiency can be obtained if
  - jumps over barriers are allowed
  - deepest wells have larger probability of being considered



more interesting



less interesting

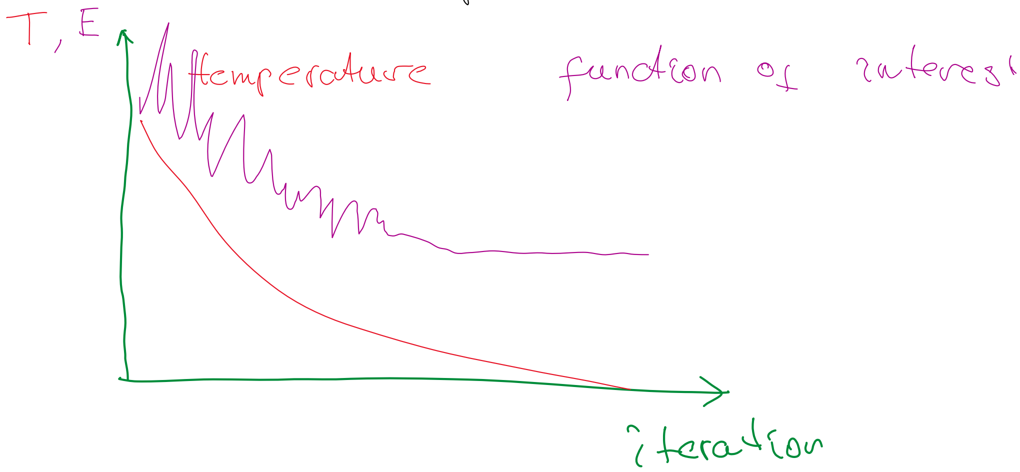
- Annealing method: uses fictitious temperature
- Name comes from metallurgy

$E(R) = E(x_1, x_2, x_3 \ldots)$

- is a function of interest to be minimized
- sample $p = \exp\left(-\frac{E(R)}{T}\right)$ distribution wher T is artifitial temperature
- use Metropolis algorithm to do that
- lower the temperature

T, E

- lower the temperature



- initial temperature larger than the height the barrier
- gradualy decrease the temperature

$$T := T \cdot 0{,}999$$

Thompson atom

$$E_{pot} = \underbrace{\sum_{i=1}^{N} \frac{1}{2} m \omega^2 r_i^2}_{\text{harmonic trap}} + \underbrace{\sum_{i<j}^{N} \frac{q^2}{|r_i - r_j|}}_{\text{Coulomb interaction}}$$

- potential energy

Dimensionless units:

$$E_{pot} = \sum_{i=1}^{N} r_i^2 + \sum_{i<j} \frac{1}{|r_i - r_j|}$$

$$= \sum_{i=1}^{N} (x_i^2 + y_i^2) + \sum_{i<j}^{N} \frac{1}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}$$

Metropolis algorithm implement

$R \to R'$ move generation:

$\frac{p(R')}{p(R)}$ where

$$p(R) = \exp\left\{-\frac{E_{pot}(R)}{T}\right\}$$

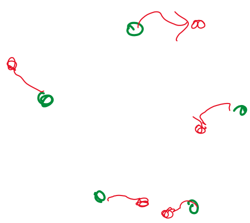
Maxwell - Boltzmann distribution

1) Moving all particles (Global move)

$$x_i' = x_i + (2u_i - 1)\cdot\Delta t$$
$$y_i' = y_i + (2u_i' - 1)\,\Delta t$$

uniform random distribution

$$0 < u_i < 1$$
$$0 < u_i' < 1$$



→ calculate $\omega = p(R')/p(R)$

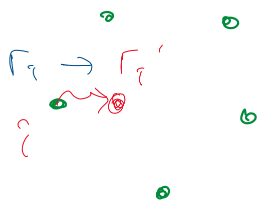
→ accept or reject with ω

→ adjust timestep Δt in order to have $\approx 50\%$ acceptance rate

2) Move single particle (local move)

i ; $x_i' = x_i + (2u - 1)\Delta t$

$\quad y_i' = y_i + (2u' - 1)\Delta t$



$j \neq i$ $\quad x_j' = x_j$

$\quad y_j' = y_j$

Calculate weight

$$\omega = \frac{p(R')}{p(R)} = \frac{\exp\{-E_{pot}(R')/T\}}{\exp\{-E_{pot}(R)/T\}}$$

$$= \exp\left\{-\frac{E_{pot}(R') - E_{pot}(R)}{T}\right\}$$

$$= \exp\left\{-\underbrace{\sum_{i=1}^{N}\frac{(r_i'^2 - r_i^2)}{T}}_{N \text{ terms}} - \underbrace{\frac{1}{T}\sum_{i<j}^{N}\left(\frac{1}{|r_i' - r_j'|} - \frac{1}{|r_i - r_j|}\right)}_{\frac{N(N-1)}{2} \text{ terms}}\right\}$$

applies both the global and local moves

move i

$$= \exp\left\{-\underbrace{\frac{(r_i')^2 - r_i^2}{T}}_{1 \text{ term}} - \underbrace{\frac{1}{T}\sum_{j\neq i}^{N}\left(\frac{1}{|r_i' - r_j|} - \frac{1}{|r_i - r_j|}\right)}_{(N-1) \text{ terms}}\right\}$$

- allowed timestep Δt is much larger
- does not depend on N

(instead it decreases with N in the global move)