Lecture: Monte Carlo method for Ising model (1)

- Monte Carlo is a sampling method
  - it solves the following problem:

    given a probability distribution

    $$P(x_1, \dots x_N)$$

    how can we efficiently evaluate averages over $P(x_1, \dots, x_N)$, i.e. quantities like

    $$\langle A \rangle = \int dx_1 \dots dx_N \, P(x_1, \dots, x_N) A(x_1, \dots, x_N)$$

- conditions on $P(x_1 \dots, x_N)$ — $P(x_1 \dots x_N)$ and $A(x_1 \dots x_N)$ → known functions!

  1.) $P(x_1, \dots, x_N) \geq 0$ for any $x_1 \dots x_N$

  2.) $\int dx_1 \dots dx_N \, P(x_1, \dots, x_N) = 1$

- in principle one could solve this by straightforward integration based on discretization

  $$\to \int dx_1 \dots dx_N \, P(x_1 \dots x_N) A(x_1 \dots x_N)$$

  $$\approx \sum_{i_1=1}^{L} \dots \sum_{i_N=1}^{L} \Delta x^N P(x_{i_1}, \dots x_{i_N}) A(x_{i_1}, \dots, x_{i_N})$$

  but this sum has L^N terms

  for $N = 100$ and $L = 100$

  100^{100} → too big for modern computers!!!

- basic idea behind Monte Carlo:
  - in large dimensional averages like $\langle A \rangle$ the part of configuration space where the contributions to the integral are large are actually very localized
    - => when we evaluate $\langle A \rangle$ by summing as

$$\sum_{i_1=1}^{L} \cdots \sum_{i_N=1}^{L} dx^N P(x_{i_1} \cdots x_{i_N}) A(x_{i_1} \cdots x_{i_N})$$

we go everywhere, not only to places where P is significant => wasting computer time

- to give an example of a sharply peaked distribution
  - => Boltzmann distribution in statistical mechanics is sharply peaked around the configuration with the largest contribution when the thermodynamic limit is taken
- need to develop a method in which the configurations visited fall in places where $P(x_1 \cdots x_N)$ is large and where $P(x_1 \cdots x_N)$ is small, the algorithm goes less frequently => importance sampling

Basic Monte Carlo method: Metropolis MC

- will use a ~~one~~ one-variable distribution $P(x)$, but the idea can easily be generalized to many dimensions
- MC: uses a Markov chain to generate a sequence of configurations
  - Markov chain:
    - given a sequence

$$x_1 \to x_2 \to x_3 \to x_4 \cdots \to x_N$$

x_2 depends solely on x_1, x_3 depends solely on $x_2, \ldots, x_{i+1}$ depends solely on x_i

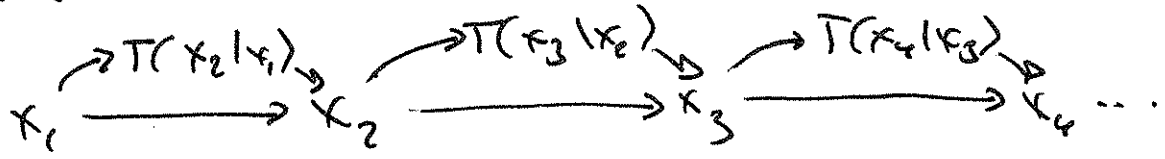
One can write the probability of arriving at x_{i+1} solely in terms of the probability of ~~arriving~~ in x_i, rather than in terms of all the previous variables

- in general $P(x_{i+1}) = \int P(x_{i+1} | x_1 \ldots x_i)\, dx_1 \ldots dx_i$

for markov process: $P(x_{i+1}) = \int dx_i\, P(x_i) T(x_{i+1} | x_i)$

where $T(x_{i+1} | x_i)$ is a transition matrix

in other words:



in the sequence (chain) position of x_{i+1} depends only x_i

$T(x_{i+1}|x_i)$ – ~~transition~~ matrix – conditional probabilities: the probability of arriving at the position x_{i+1} given that we started at x_i

~~note that~~ →

- each step in a Monte Carlo algorithm can be represented by a transition matrix $T(x'|x)$
- the actual procedure of Metropolis Monte Carlo: = propose a ~~rand~~ random move (change the position x_i by a random value, thereby generating a new potential configuration)
  - accept or reject the proposed configuration

$$\Rightarrow \quad T(x'|x) = M(x'|x)\,A(x'|x)$$

- T can be represented as a product of two matrices M and A → M represents the proposal of a new configuration, A represents the probability that the move is accepted

- how do we construct T, more specifically, H and A?
- we want our Markov chain to have a distribution $P(x)$
  - in this case $x_1 \to x_2 \to x_3 \to \dots \to x_N$

    we can evaluate averages $\frac{1}{L}\sum_{i=1}^{L} A(x_i) = \langle A \rangle$

- ~~what~~ to generate such a distribution, we require that $T(x'|x)$ satisfy detailed balance

$$P(x_i)T(x_{i+1}|x_i) = P(x_{i+1})T(x_i|x_{i+1})$$

if this condition is satisfied, then if our variable x_i is sampled from $P(x)$, then so will x_{i+1} be

$$\Rightarrow \int P(x_i)\,T(x_{i+1}|x_i)\,dx_i = \int P(x_{i+1})\,T(x_i|x_{i+1})\,dx_i$$

$$= P(x_{i+1})$$

since $\int dx_i\, T(x_i|x_{i+1}) = 1$

(the probability of arriving anywhere from x_{i+1})

- so we must construct T, or H and A so that detailed balance is satisfied

M - represents proposed random move

$$x'_{i+1} = x_i + \xi \qquad -\frac{\Delta}{2} < \xi < \frac{\Delta}{2}$$

ξ - random number between $-\frac{\Delta}{2}, \frac{\Delta}{2}$

- in this case $M(x'|x) = M(x|x')$

$\Rightarrow$ using this the detailed balance condition can be rewritten as $P(x) A(x'|x) = P(x') A(x|x')$

A - choose as $A(x'|x) = \min[1, P(x')/P(x)]$

- prove that detailed balance is satisfied
  - consider case $P(x) > P(x')$

$$P(x) A(x'|x) = P(x') A(x|x')$$

$$P(x) \min[1, P(x')/P(x)] = P(x') \min[1, P(x)/P(x')]$$

$$P(x) P(x')/P(x) = P(x')$$

$$P(x') = P(x') \quad \checkmark$$

  - consider case $P(x) < P(x')$

$$P(x) A(x'|x) = P(x') A(x|x')$$

$$P(x) \min[1, P(x')/P(x)] = P(x') \min[1, P(x)/P(x')]$$

$$P(x) = P(x') P(x)/P(x')$$

$$P(x) = P(x)$$

- Monte Carlo procedure:

given some initial configuration x_i

1.) generate a proposed move x'_{i+1}

$$x'_{i+1} = x_i + \xi \qquad -\frac{\Delta}{2} < \xi < \frac{\Delta}{2}$$

2.) accept or reject proposed configuration according to

$$A(x'_{i+1} | x_i) = \min[1, P(x'_{i+1})/P(x_i)]$$

if new configuration is accepted then

$$x_{i+1} = x'_{i+1}$$

if not $x_{i+1} = x_i$

- for the Ising model we have a many-variable system σ_i where $i = 1, \ldots, N$ (N is total number of lattice sites)

in this case: MC can be implemented based on spin-flips (individual)

- start with some configuration $\{\sigma_i\}_0$
- choose a spin randomly σ_j
  - flip it => $\sigma_j \to -\sigma_j$
  - ~~new config~~
  - accept or reject according to ratio of Boltzmann factors of the old and proposed configurations

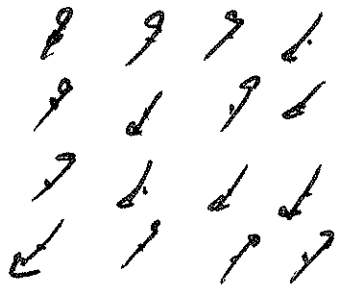
in more detail: Ising model

- Hamiltonian: $H = \sum_{\langle i,j \rangle} \sigma_{ij} \sigma_i \sigma_j - h \sum_i \sigma_i$

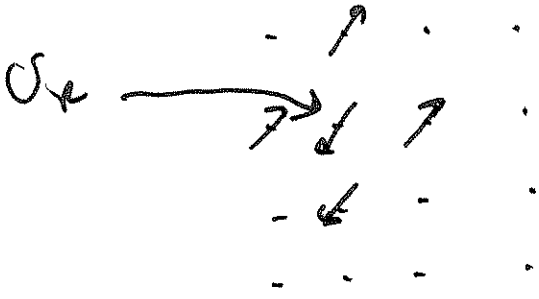
- for simplicity assume nearest neighbor couplings, and that all coupling ~~are~~ are equal

$$H = J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i$$

- MC procedure: start with some configuration (can be randomly generated)

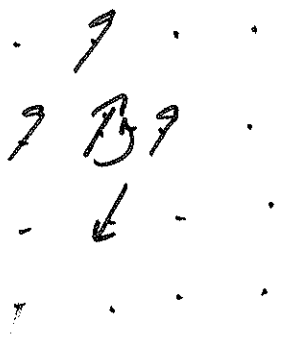


- choose one spin randomly



=> old configuration

- flip it



=> proposed configuration

$\sigma_k' \leftarrow \sigma_k$

- calculate ratio of Boltzmann factors

$$P(\{\sigma_i\}) = e^{-\beta J \sum_{\langle i,j \rangle} \sigma_i \sigma_j + \beta h \sum_i \sigma_i}$$

$$P(\{\sigma_i'\}) = e^{-\beta J \sum_{\langle i,j \rangle} \sigma_i' \sigma_j' + \beta h \sum_i \sigma_i'}$$

$\{\sigma_i\}$ - old configuration

$\{\sigma_i'\}$ - new configuration

- given that only one spin is different between the new and old configurations, in calculating the ratio of Boltzmann factors and the acceptance probability, enough to consider the bonds which are connected to the flipped spin

$$\underbrace{A(\sigma_k', \sigma_k) = \exp[-\beta J \sum_i \sigma_i(\sigma_k' - \sigma_k)]}_{\text{acceptance probability}}$$

accept with this probability

- then choose another spin randomly and repeat procedure

- call MC pass: one sequence of accepted or rejected moves in which N spins were randomly chosen for proposed moves
N times ⇒ every spin was given a chance to move

- MC procedure:
  - setup initial configuration
  - make a large number of MC passes so that the system reaches equilibrium (the first configuration may not be representative of the equilibrium fluctuations we wish to average)
  - after equilibration, keep generating new configurations, and use them to calculate averages

☺ ☺ ☺

NEXT QUESTION: how large is large?

(answer below)

- MC can calculate equilibrium averages

for Ising model:

- average energy: $\bar{E} = J \sum_{\langle i,j \rangle} \langle \sigma_i \sigma_j \rangle - h \sum_i \langle \sigma_i \rangle$
- magnetization: $\bar{M} = \frac{1}{N} \sum_i \langle \sigma_i \rangle$
- susceptibility: $\bar{\chi} = $ ~~$\sum_{i,j} \sigma_i \sigma_j$~~

$$\bar{\chi} = \frac{\partial \bar{M}}{\partial h} = \frac{1}{N} \sum_i \sum_j \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle$$

- correlation functions:

$$C_{ij} = \frac{\langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle}{\langle \sigma_i^2 \rangle - \langle \sigma_i \rangle^2}$$

- answer question: how large is large?

MC proceeds in steps/passes

- can also calculate correlation function in MC passes

for example: take one spin and watch how it changes after each MC pass

σ_k^α $\quad \alpha \rightarrow$ MC pass

$\alpha = 1, \ldots, \Gamma$

we can calculate: $C_\beta = \frac{1}{N} \sum_{\alpha=1}^{N} \frac{\langle \sigma_k^\alpha \sigma_k^{\alpha+\beta} \rangle - \langle \sigma_k^\alpha \rangle^2}{\langle \sigma_k^{\alpha 2} \rangle - \langle \sigma_k^\alpha \rangle^2}$

$C_\beta = 1$ for $\beta = 0$

as $\beta \to \infty$ $C_\beta \to 0$

approximately

- if we assume that C_β is exponential decay (which is approximately true usually)

$$C_\beta = e^{-\beta/\tau}$$ $\tau \to$ correlation time

$$\int_0^\infty d\beta\, C_\beta = \int_0^\infty d\beta\, e^{-\beta/\tau} = -\tau e^{-\beta/\tau}\Big|_0^\infty = \tau$$

$\Rightarrow$ correlation time estimated for MC passes

$$\sum_{\beta=1}^{\Gamma} C_\beta \sim \tau$$

Γ must be "large"

Γ must be such that

$$C_\beta \to 0$$

- one can also calculate correlation function in MC passes for observables which depend on the whole lattice

$$M^\alpha = \frac{1}{N}\sum_i \sigma_i^\alpha \Rightarrow C_\beta = \frac{1}{\Gamma}\sum_{\alpha=1}^{\Gamma} \frac{\langle M^\alpha M^{\alpha+\beta}\rangle - \langle M^\alpha\rangle^2}{\langle M^{\alpha 2}\rangle - \langle M^\alpha\rangle^2}$$

- to obtain averages, one must equilibrate first for a large number of correlation times
- also, if averages are taken, one does not have to take samples after every MC pass, but only after the passing of a correlation time

⇒ in other words

- equilibrate first (it possible, several hundred correlation times of MC passes)
- after equilibration, keep generating configurations

$$\{\sigma_i\}_I \rightarrow \{\sigma_i\}_{I+1} \rightarrow \{\sigma_i\}_{I+2} \rightarrow \cdots$$

but average only configurations which are separated by number of passes ~~which are~~ equal to τ for maximum efficiency