Page 1

Lecture IV

Monte Carlo in more detail

A word of warning! ⚠️

> "Monte Carlo is an extremely bad method, it should only be used when all alternative methods are worse!"
> - Alan Sokal in his introduction to "Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms"

Curse of Dimensionality

Consider a grid with 10 evenly spaced sample points in each dimension. The total number of points grows exponentially with the dimension, d:

Dimension	# of Points
1D	10^1
2D	10^2
3D	10^3
nD	10^n

导出到 Google 表格

This shows exponential growth with dimension.

Page 2

Conventional numerical integration requires an exponentially increasing number of function evaluations as the dimension 'd' increases.

By contrast, with Monte Carlo methods, the error scales as $\sim \frac{1}{\sqrt{N}}$, where N is the total number of function evaluations. This error scaling is independent of the system's dimensionality.

This is a result of the **central limit theorem**: distributions get sharper the more events contribute to the system.

Page 3

Markov Chain Monte Carlo

A **Markov process** transfers a system from a state i to a state j with a transition probability $p(i \to j)$.

This probability has the following characteristics:

1. It is time-independent.
2. It depends only on the initial and final states.
3. The sum over all possible final states is one: $\sum_j p(i \to j) = 1$.

A **Markov chain** is a sequence of Markov processes. If a chain is "positive recurrent" (meaning it has a finite return time to every state), it will have a stationary distribution of visited states.

Page 4

An **ergodic** (or irreducible) process is one where every state can be reached from any other state. This property ensures a unique stationary distribution.

Markov Chain Monte Carlo (MCMC) methods construct a Markov chain that generates a Boltzmann distribution of microstates.

This is achieved by enforcing the **"detailed balance"** condition:
$p(i)p(i \to j) = p(j)p(j \to i)$

This is a stronger condition than simply having a stationary distribution. It implies that the ratio of transition probabilities is:
$\frac{p(i \to j)}{p(j \to i)} = e^{-\beta(E_j - E_i)}$

Page 5

If the detailed balance condition is fulfilled, the **ergodic hypothesis** holds, meaning the time average of a property is equal to its ensemble average.

Time Average:
$\langle A \rangle_t = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau dt A(r(t), p(t)) \approx \frac{1}{l} \sum_{n=1}^{l} A(\underline{r}(t_n), \underline{p}(t_n))$

Ensemble Average:
$\langle A \rangle_\rho = \mathcal{N} \int d\underline{r}^N d\underline{p}^N \rho(\underline{r}^N, \underline{p}^N) A(\underline{r}^N, \underline{p}^N)$

So, we can state: $\langle A \rangle_t = \langle A \rangle_\rho = \langle A \rangle$

Page 6

Example: Metropolis-Hastings Algorithm

This was first described in the 1953 paper "Equation of State Calculations by Fast Computing Machines" by Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller.

The basic algorithm is:

1. Start from a random initial state.
2. Introduce a small, random, ergodic perturbation to create a new state.
3. Accept or reject the new state based on an acceptance probability, p.

The acceptance probability is:

$$p = \begin{cases} 1 & \text{if } E_{new} < E_{old} \\ \exp(-\beta(E_{new} - E_{old})) & \text{if } E_{new} \geq E_{old} \end{cases}$$

Why use small perturbations?

To ensure the acceptance probability is reasonably high for efficient sampling of the phase space.

Page 7

Example I: Metropolis Algorithm for the Ising Model

The Hamiltonian for the Ising model is:

$H = -J \sum_{\langle i,j \rangle} s_i s_j - B \sum_i s_i$

where the spins $s_i = \pm 1$ and periodic boundary conditions are used. The 2D case has an exact analytical solution by Onsager (1944), which serves as a useful reference.

- **Ergodic perturbation**: A single spin flip. Any state can be transformed into any other state through a series of single spin flips. The choice of acceptance criterion ensures ergodicity.
- **Other perturbations**: Cluster flips can also be used.
- **Important considerations**: Finite size effects and phase transitions.

Page 8

Example II: Monoatomic Fluid

The system is often modeled with the **Lennard-Jones** potential:

$$H = \sum_{i>j} 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right]$$

- **Ergodic perturbation**: Small random displacements of atoms (either individually, as a group, or all at once). This ensures that all possible configurations $\{\underline{r}_i'\}$ can be reached from any initial state $\{\underline{r}_i\}$.

Page 9

MC in Different Thermodynamic Ensembles

So far, we have been sampling from the Boltzmann distribution, which corresponds to the **canonical (NVT) ensemble**.

However, real experiments are often conducted under conditions of constant pressure, not constant volume. This corresponds to the **isothermal-isobaric (NPT) ensemble**.

- Heat is exchanged with the environment (fixed T).
- No particle exchange (fixed N).
- The system can expand or contract (fixed P).

In the NPT ensemble, the energy and volume can fluctuate. The ergodic perturbation must therefore include moves that change the system's volume in addition to the particle positions.

Page 10

MC in the NPT Ensemble

Volume becomes an additional degree of freedom.

Simply changing the volume without adjusting particle positions leads to very low acceptance rates.

- **Solution**: Rescale the particle positions when the volume is changed.

We can define scaled, dimensionless coordinates s_i for each particle, such that the real position r_i is given by $r_i = Ls_i$, where L is the length of the cubic simulation box ($V = L^3$).

The ergodic perturbation now consists of two types of moves:

1. A change to the scaled positions, s_i.
2. A change to the box length, L, which rescales all particle positions.

Page 11

Homework: Write an MC program to compute π 🥧

The method is as follows:

1. Generate random points within a 1x1 square.
2. Check if a point (x, y) is inside the inscribed quarter circle (i.e., if $\sqrt{x^2 + y^2} \leq 1$).
3. The value of π can be approximated from the ratio of points inside the quarter circle to the total number of points.

$$\pi \approx 4 \times \frac{\#\text{ points inside}}{\#\text{ total points}}$$