
Física Estatística Computacional

Tereza Mendes

IFSC – USP

<http://lattice.ifsc.usp.br/cbpf.html>

Aula 5: MONTE CARLO



Método de Monte Carlo

Sistemas estocásticos são simulados no computador usando um **gerador de números aleatórios**



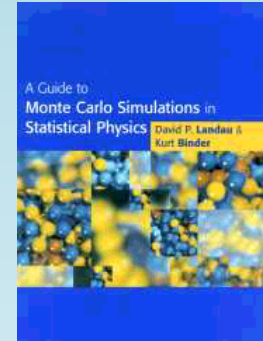
⇒ tratamento teórico, com aspectos experimentais:

- dados, erros
- “medidas” no tempo

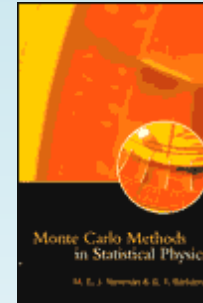


Referências

- A Guide to Monte Carlo Simulations in Statistical Physics, [Landau & Binder](#) (Cambridge, 2000)



- Monte Carlo Methods in Statistical Physics, [Newman & Barkema](#) (Oxford, 1999)



- Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms, [Sokal](#) (1996),

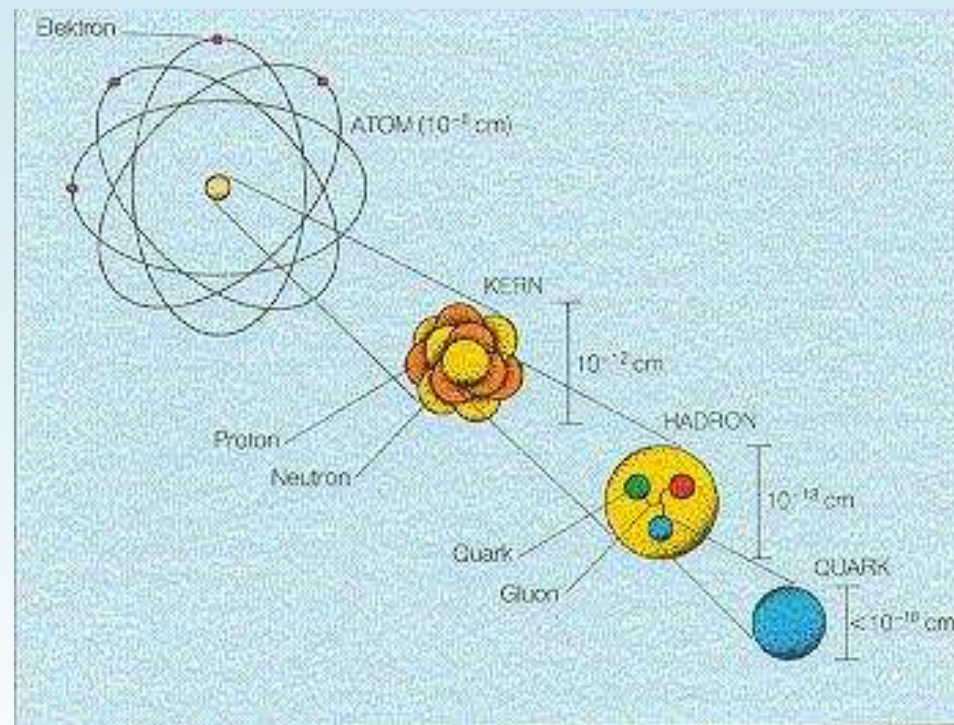
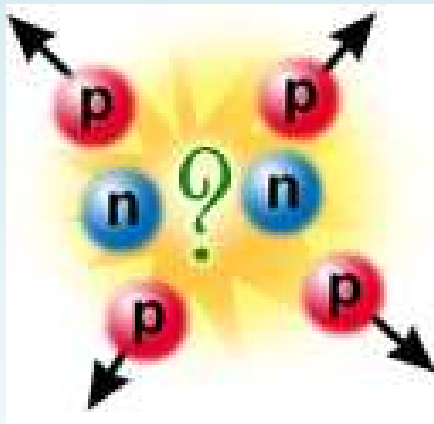
<http://citeseer.nj.nec.com/sokal96monte.html>

- Notas de aula - Transição de fase e fenômenos críticos (2002), <http://lattice.if.sc.usp.br/>

Aplicações

- **Mecânica Estatística**: descrição de **sistemas de muitos corpos** ($\approx 10^{23}$ corpos...) utilizando grandezas médias \Rightarrow comportamento **macroscópico** (termodinâmica) a partir da descrição **microscópica** de sistemas como fluidos/gases, modelos de **materiais magnéticos**, sistemas biológicos; tratamento de **fenômenos críticos**, sistemas complexos.
- **Matéria Condensada**: descrição aproximada de sistemas quânticos, polímeros, fluidos complexos, propriedades condutoras/magnéticas.

- **Cromodinâmica Quântica (QCD):** teoria quântica de campos que descreve a **força nuclear** como interação forte entre **quarks** e **glúons**; Formulação de **Rede** \leftrightarrow Mecânica Estatística.



O Problema

Mecânica Estatística: probabilidade de uma configuração S para um sistema em equilíbrio à temperatura T é dada (no **ensemble canônico**) em termos de sua **hamiltoniana** $\mathcal{H}(S)$ pela **distribuição de Boltzmann**

$$P(S) = \frac{e^{-\beta\mathcal{H}(S)}}{Z}; \quad Z = \int dS e^{-\beta\mathcal{H}(S)}; \quad \beta = 1/KT$$

Média termodinâmica do observável A dada por

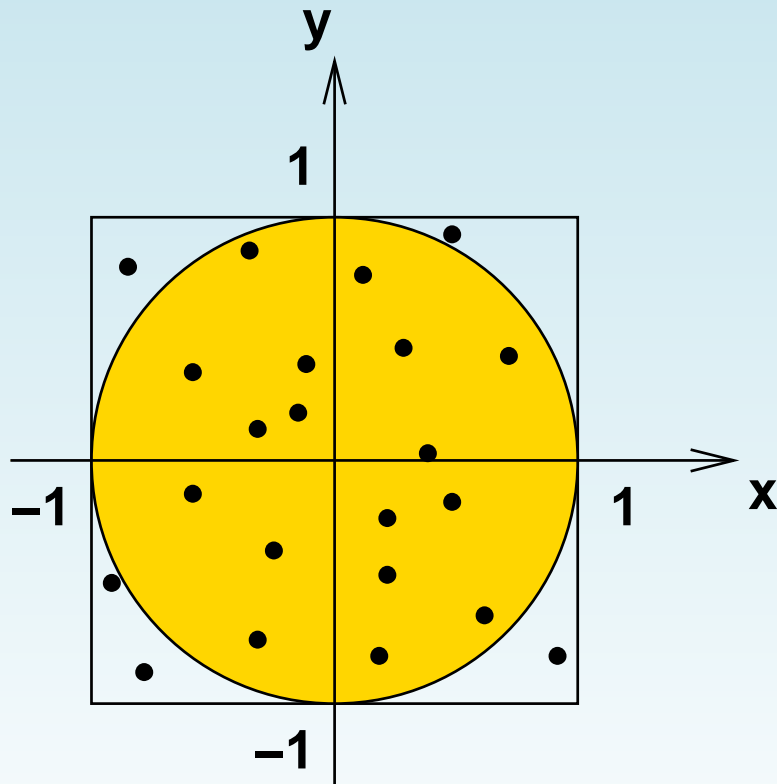
$$\langle A \rangle = \int dS A(S) P(S)$$

e.g. energia: $E = \langle \mathcal{H}(S) \rangle$

Integral (multi-dimensional) muito complicada!

Área do Círculo de Raio 1

Lançando N pontos aleatórios uniformemente no quadrado: $x \in [-1, 1]$, $y \in [-1, 1]$



razão entre as áreas

$$\frac{A_{\circ}}{A_{\square}} = \frac{\pi}{4} = \frac{n}{N}$$

$n < N$ é o número de pontos no círculo

Integral Unidimensional

Integral como soma de variáveis aleatórias igualmente distribuídas

$$I = \int_0^1 f(x) dx \rightarrow \frac{\sum_i f(x_i)}{N}$$

com x_i uniformemente distribuídos em $[0,1]$.

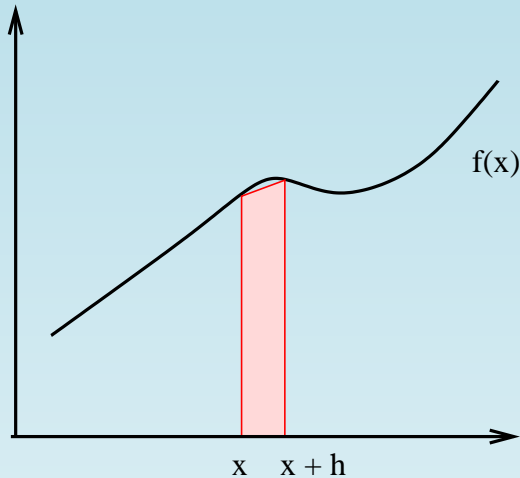
Na verdade, para N finito $\bar{I} \equiv \frac{\sum_i f(x_i)}{N}$

é uma **variável aleatória**, que converge para seu valor médio I com erro proporcional a $1/\sqrt{N}$ (teorema central do limite).

$$\sigma_{\bar{I}}^2 = \frac{\sigma_f^2}{N} = \frac{\langle f^2 \rangle - \langle f \rangle^2}{N}$$

Exercício: derive a expressão acima

Métodos Determinísticos



Regra do trapezóide: estima a área compreendida entre x e $x+h$ como a área do trapezóide definido pela aproximação linear da função entre estes dois pontos.

$$\int_x^{x+h} f(x') dx' = \frac{h}{2} [f(x) + f(x+h)] + \mathcal{O}(h^3 f'')$$

erro para integral em $[a, b]$ é $\mathcal{O}(h^2) \sim N^{-2}$

Regra de Simpson: aproximação de 3 pontos para $f(x) \Rightarrow$
erro é $\sim N^{-4}$

Comparação

- $d = 1$:

métodos determinísticos tipicamente têm erros $\mathcal{O}(N^{-2})$ (regra do trapézio) ou $\sim \mathcal{O}(N^{-4})$ (regra de Simpson); Monte Carlo tem $\mathcal{O}(N^{-1/2})$: com $2N$ pontos o erro diminui por um fator 4 (trapezóide), 16 (Simpson) ou $\sqrt{2}$ (Monte Carlo)

- $d > 1$:

para integral d -dimensional $N \sim 1/h^d \Rightarrow$ erro $N^{-2/d}$ (trapezóide) ou $N^{-4/d}$ (Simpson) \Rightarrow Monte Carlo começa a ser vantagem a partir de $d = 8...$

Comparação

Tipicamente, em mecânica estatística o número de dimensões (i.e. número de graus de liberdade) é $d \sim 10^3$ (e.g. modelo de Ising em 3d com 10 pontos por direção)

⇒ tempo para somar 2^{1000} em computador de 1 Tflops:

$$t = 10^{288} \text{ s} = 10^{270} \times \text{idade do universo}$$

Comparação

Tipicamente, em mecânica estatística o **número de dimensões** (i.e. número de **graus de liberdade**) é $d \sim 10^3$ (e.g. modelo de Ising em 3d com 10 pontos por direção)

⇒ tempo para somar 2^{1000} em computador de 1 Tflops:

$$t = 10^{288} \text{ s} = 10^{270} \times \text{idade do universo}$$

⇒ **Monte Carlo não é a melhor escolha,
é a única escolha!**

Amostragem por Importância

Note: calcular

$$I = \int_0^1 f(x) w(x) dx = \frac{\sum_i f(x_i) w(x_i)}{N}$$

[onde $\int_0^1 w(x) dx = 1$] com x_i uniformemente distribuídos em $[0,1]$ é **muito ineficiente** se $w(x)$ é concentrada.

Portanto toma-se

$$I = \int_0^1 f(x) w(x) dx = \frac{\sum_i f(x_i)}{N}$$

onde os x_i são gerados com a distribuição $w(x_i)$.

Amostragem de distribuições

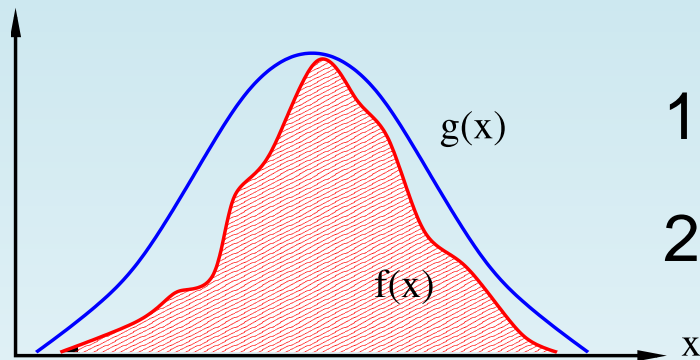
Precisamos de métodos para produzir x_i com distribuição $w(x_i)$ (**amostrar w**) a partir da distribuição uniforme.

- Distribuição discreta, e.g. $x = 1$ com probabilidade p , $x = -1$ com $1 - p$ obtida gerando r uniforme em $[0,1]$ e tomando $x = 1$ se $r \leq p$, $x = -1$ se $r > p$
- distribuição exponencial $w(x) = e^{-x}$ em $[0, \infty]$ é amostrada gerando r uniforme em $[0,1]$ e tomando $x = -\log(1 - r)$, já que isso corresponde a

$$P(x) dx = P(r) dr \Rightarrow P(x) = \frac{d}{dx}(1 - e^{-x}) = w(x)$$

Método da Rejeição

Podemos amostrar uma distribuição $f(x)$ se soubermos amostrar $g(x)$ (não-normalizada) tal que $g(x) \geq f(x)$ para todo x .



1. **gera** x com dist. prop. a $g(x)$
2. **aceita** com prob. $f(x)/g(x)$

Os x aceitos terão distribuição $f(x)$. Aceitação média:

$$A = \frac{\int_{-\infty}^{\infty} f(x) dx}{\int_{-\infty}^{\infty} g(x) dx}$$

Distribuição de Boltzmann

Mesmo com bons métodos e amostragem por importância, para uma distribuição conjunta de muitos graus de liberdade como a distribuição de Boltzmann

$$\langle A \rangle = \int A(x) w(x) dx, \quad w(x) = \frac{e^{-\beta \mathcal{H}(x)}}{Z}$$

não há esperança de amostragem direta!

Solução: Monte Carlo dinâmico. **Inventamos** uma evolução temporal de modo que as configurações geradas sejam distribuídas de acordo com $w(x)$. Isto pode ser feito para uma **dinâmica markoviana** escolhida de forma conveniente.

Cadeia de Markov

Processo estocástico X_0, X_1, \dots, X_t tal que

$$P(X_{t+1}|x_0 \dots x_t) = P(X_{t+1}|x_t)$$

(futuro depende do passado apenas através do presente)

⇒ história da cadeia determinada pela distribuição inicial $P(X_0)$ e pela **matriz de transição** p_{xy}

p_{xy} = probabilidade de ir de x para y em 1 passo

Note: $p_{xy}^{(2)}$ é a probabilidade de $x \rightarrow y$ em 2 passos

Claramente, $\sum_y p_{xy} = 1$ para todo x .

Distribuição Estacionária

Se existir $w(x)$ tal que

$$\sum_x w(x) p_{xy} = w(y) \quad \text{para todo } y$$

e a cadeia for **aperiódica** então o processo converge para a **distribuição estacionária** $w(x)$ independentemente da distribuição inicial $P(X_0)$: $\lim_{t \rightarrow \infty} p_{xy}^{(t)} = w(y)$

Nosso problema (inverso): será possível encontrar p_{xy} que tenha a distribuição desejada $w(x)$ (e.g. a distribuição de Boltzmann) como distribuição estacionária? Daí médias temporais na cadeia convergem (quando $t \rightarrow \infty$) para médias na distribuição de equilíbrio $w(x)$ do sistema considerado. **Note**: desprezo o transiente inicial.

Distribuição Estacionária

Condições sobre a **dinâmica** p_{xy} da **cadeia de Markov** para que haja convergência à distribuição $w(x)$ para tempos longos

(A) Irreducibilidade (**ergodicidade**): para todo x, y há n tal que $p_{xy}^{(n)} \neq 0$

(B) $w(x)$ é estacionária: $\sum_x w(x) p_{xy} = w(y)$

Nota: é possível também impor a **condição suficiente**

(B') balanço detalhado: $w(x) p_{xy} = w(y) p_{yx}$

Exercício: mostre que **(B') \rightarrow (B)**

Método de Metropolis

based on proposing and accepting/rejecting a step $x \rightarrow y$

- **accept** if $w(y)/w(x) \geq 1$
- otherwise **accept with probability** $w(y)/w(x)$

the probability of acceptance is $A_{xy} = \min \{1, w(y)/w(x)\}$. Then consider the transition matrix $p_{xy} = T_{xy} A_{xy}$ (with general $T_{xy} = T_{yx}$)

Exercise: show that the above choice satisfies **detailed balance**

For the Boltzmann distribution this means

$$w(x) = \frac{e^{-\beta E(x)}}{Z} \Rightarrow \frac{w(y)}{w(x)} = e^{-\beta \Delta E}; \quad \Delta E \equiv E(y) - E(x)$$

\Rightarrow accept if $\Delta E \leq 0$; otherwise accept with probability $e^{-\beta \Delta E}$

Note: if proposed step is rejected, keep old value and move to a new site; when possible, choose T_{xy} such that acceptance is 50%

Método do Banho Térmico

Geralmente a evolução das configurações (e.g. x or S) do sistema é feita **congelando** as variáveis de campo em todos os pontos menos um. Este ponto é então amostrado por um **método local** (pode ser Metropolis). Uma iteração do algoritmo, i.e. um passo da cadeia de Markov, é obtido percorrendo-se assim todos os sítios do sistema

Algoritmo de banho térmico: amostragem exata da distribuição (condicional) local; claramente uma maneira válida de amostrar a distribuição conjunta (reamostragem parcial)

mais difícil de implementar do que Metropolis...

Receita

follow the dynamics $X(t) = x_i$ and compute time averages

$$\langle A \rangle = \int A(x) w(x) dx = \frac{\sum_i A(x_i)}{N}$$

which are expectation values in the desired distribution, i.e. the Boltzmann distribution. The resulting averages + errors are the output of our Monte Carlo simulation. **Note: initial transient must be discarded.**

But... we have a problem: samples are **not** independent.



The program above is subject to systematic effects.

The time correlation between different steps of the Markov chain is

$$C(k) = \frac{\langle A_i A_{i+k} \rangle - \langle A_i \rangle^2}{\langle A_i^2 \rangle - \langle A_i \rangle^2}$$

\Rightarrow independent samples only after $C(k) \approx 0$; $k =$ decorrelation time

Correlações

(Monte Carlo) average of A :
$$\bar{A} = \frac{1}{N} \sum_{i=1}^N A_i$$

Variance:
$$\sigma_A^2 = \frac{\sigma_A^2}{N} \left[1 + 2 \sum_{k=1}^{N-1} C(k) \right] = \frac{\sigma_A^2}{N} (2\tau)$$

where the **temporal correlation** $C(k)$ was given above and τ is the **auto-correlation time** for observable A .

Consider $C(k) = e^{-k/\tau}$, τ large (but $\tau \ll N$)

$$\begin{aligned} 1 + 2 \sum_{k=1}^{N-1} C(k) &\approx 2 \sum_{k=0}^{\infty} e^{-k/\tau} - 1 \\ &\approx 2\tau \int_0^{\infty} e^{-u} du - 1 \approx 2\tau \end{aligned}$$

We therefore define $\tau \equiv \frac{1}{2} + \sum_{k=1}^{N-1} C(k)$

Monte Carlo Method: Summary

Integral becomes sum of random variables

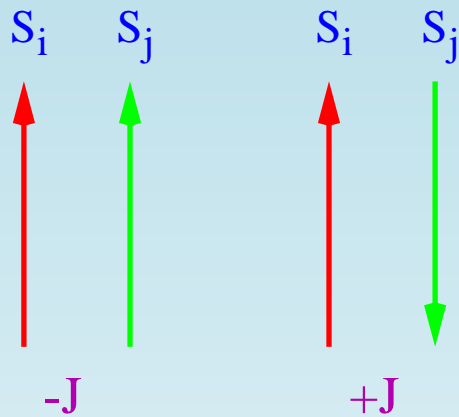
$$\int f(x) d\mu, \quad d\mu = \frac{e^{-\beta\mathcal{H}(x)}}{Z} dx \quad \Rightarrow \quad \frac{1}{N} \sum_{i=1}^N f(x_i)$$

where x_i have statistical distribution μ

- **Static** Monte Carlo: independent sampling (error $\sim 1/\sqrt{N}$)
- **Dynamic** Monte Carlo: Simulation of a Markov chain with equilibrium distribution μ (error $\sim \sqrt{\tau/N}$). Autocorrelation time τ related to **critical slowing-down**. **Note**: similar to **experimental methods**, but temporal dynamics was artificially introduced

Errors: either consider only effectively independent samples (via temporal correlation analysis) and error is given by standard deviation, **jack-knife**, **bootstrap** or consider all samples and error is estimated taking correlations into account: **binning** method, self-consistent windowing method

Aplicação ao Modelo de Ising (I)

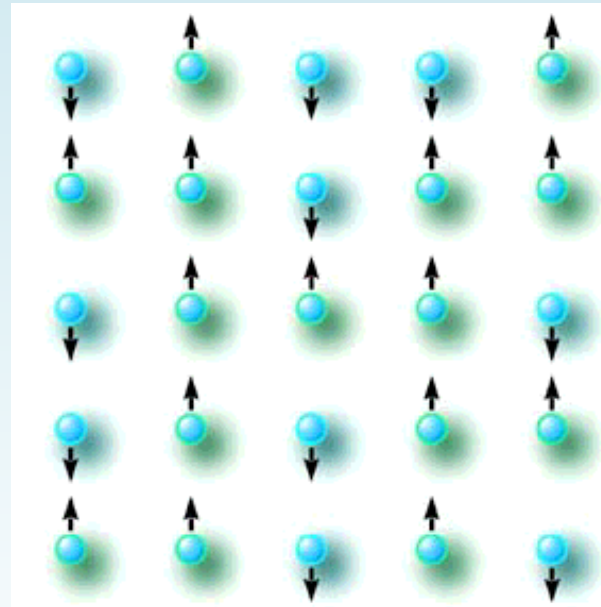


spins de dois estados, que preferem ficar alinhados

$$\mathcal{H}(S) = -J \sum_{\langle i,j \rangle} S_i S_j - H \sum_i S_i$$

Observáveis de interesse

- Energy: $E = \langle \mathcal{H}(S) \rangle$
- Specific Heat: $C_V = \partial E / \partial T$
- Magnetization: $M = \langle \sum_i S_i \rangle$
- Suscetibility: $\chi = \partial M / \partial H$



Aplicação ao Modelo de Ising (II)

Metropolis method for the Ising model: sweep over the lattice, at each site **propose to flip** the spin, i.e. $S_i \rightarrow -S_i$. Acceptance probability

$$\frac{e^{-\beta E(y)}}{e^{-\beta E(x)}} = \frac{e^{+\beta J S_i \sum_{j \text{ n.n. } i} S_j}}{e^{-\beta J S_i \sum_{j \text{ n.n. } i} S_j}} = e^{2\beta J S_i h_i}$$

an **iteration** consists of a complete “sweep” over the lattice. At the end of the iteration compute $A(S)$ for the generated configuration, and restart the process of generating configurations

Heat-bath method for the Ising model: exact sampling of the conditional probability at site i . Sweep over the lattice, at each site **pick a new value** for S_i independently of the old one, keeping all other spins fixed. Unnormalized probability $P(S_i) = e^{\beta J S_i \sum_{j \text{ n.n. } i} S_j} \times \text{const.}$

Thus

$$\blacksquare P(S_i = +1) = e^{\beta J h_i} / (e^{\beta J h_i} + e^{-\beta J h_i}) \equiv p$$

$$\blacksquare P(S_i = -1) = 1 - p$$

Exercício

Escreva programas simples para simular o modelo de Ising bidimensional utilizando os métodos de Metropolis e do banho térmico.

Compare a eficiência dos dois métodos. Faça gráficos da dependência com o campo magnético e com a temperatura.

Note: maiores detalhes e exemplos, assim como programas parcialmente prontos, estão disponíveis no site <http://lattice.if.sc.usp.br/oldlattice/CADSC03/escola.html>