Monte Carlo via Cadeias de Markov - Aplicações Físicas

Gabriel Moreira da Silva Campos



Instituto de Física Armando Dia Tavares Universidade do Estado do Rio de Janeiro

gabrielmscampos@gmail.com

10 de junho de 2021

- Objetivo
- 2 Resumo
- Cadeias de Markov
- Monte Carlo
- 5 Monte Carlo via Cadeias de Markov
- 6 Oscilador Quântico Anarmônico
- Integrais de caminho
- Simulação via MCMC
- Resultados

- Objetivo
- Resumo
- Cadeias de Markov
- Monte Carlo
- 5 Monte Carlo via Cadeias de Markov
- 6 Oscilador Quântico Anarmônico
- Integrais de caminho
- Simulação via MCMC
- Resultados



Objetivo

 Apresentar Cadeias de Markov, método de Monte Carlo, a simulação de Monte Carlo via Cadeias de Markov (Markov Chain Monte Carlo, MCMC) bem como suas aplicações em problemas físicos.

- Objetivo
- 2 Resumo
- Cadeias de Markov
- Monte Carlo
- 5 Monte Carlo via Cadeias de Markov
- 6 Oscilador Quântico Anarmônico
- Integrais de caminho
- Simulação via MCMC
- Resultados

Resumo

- Casos em que soluções analíticas são inviáveis para equações matemáticas, integrações numéricas de aproximações de ordem N podem ser utilizadas para encontrar a solução com grande custeio computacional. Como alternativa, a inferência através de simulações estatística constituem um ferramental interessante para solução desses problemas.
- Simulações de Monte Carlo via Cadeias de Markov são métodos que podem ser utilizados para gerar de uma distribuição que aproxima uma distribuição alvo f.

6/43

- Objetivo
- Resumo
- Cadeias de Markov
- Monte Carlo
- 5 Monte Carlo via Cadeias de Markov
- 6 Oscilador Quântico Anarmônico
- Integrais de caminho
- Simulação via MCMC
- Resultados

Cadeias de Markov: Definição

Uma Cadeia de Markov é um processo estocástico (coleção de quantidades randômicas) onde o estado atual, passado e futuro da cadeia são independentes.

$$P(X_{n+1} \in A | X_n = x, X_{n-1} \in A_{n-1}, ..., X_0 \in A_0) = P(X_{n+1} \in A | X_n = x)$$
 (1)

para todos os conjuntos $A_0,...,A_{n-1},A\subset S$ e $x\in S$, onde S representa o espaço de estados do processo. Equivalentemente, em um caso de espaço de estados discreto:

$$P(X_{n+1} = y | X_n = x, X_{n-1} = x_{n-1}, ..., X_0 = x_0) = P(X_{n+1} = y | X_n = x)$$
 (2)

para todo $x_0, ..., x_{n-1}, x, y \in S$.

◄□▶◀圖▶◀불▶◀불▶ 불 ∽Q҈

Cadeias de Markov: Transições

As mudanças de estado de um sistema são ditas transições. Esse processo é caracterizado por um espaço de estados, uma matriz de transição descrevendo as probabilidades particulares de transição e um estado inicial.

Em geral, as probabilidades em 1 dependem de x, A e n. Quando essas não dependem de n, a cadeia é dita homogênea. Nesse caso, podemos definiar uma função de transição P(x,A).

- **1** Para todo $x \in S$, $P(x, \cdot)$ é uma distribuição de probabilidade sobre S;
- ② Para todo $A \subset S$, a função $x \mapsto P(x, A)$ pode ser computada.

Quando tratamos do caso discreto, é importante identificar $P(x, \{y\}) = P(x, y)$. A função de transição de probabilidade deve satisfazer:

- $P(x,y) \ge 0, \ \forall x,y \in S;$



Cadeias de Markov: Passeio Aleatório

Considere uma partícula se movendo independentemente para esquerda e direita em uma linha com deslocamentos sucessivos de sua posição atual governados por uma função de probabilidade f sobre inteiros e X_n representando sua posição no instante $n, n \in \mathbb{N}$. Inicialmente, X_0 é distribuído de acordo com uma distribuição π_0 . As posições podem ser relacionadas como:

$$X_n = X_{n-1} + w_n = w_1 + w_2 + ... + w_n$$
 (3)

onde w_i são variáveis randômicas independentes com função de probabilidade f. Assim, contituímos uma Cadeia de Markov $\{X_n:n\in\mathbb{N}\}$ sobre \mathbb{Z} .

Cadeias de Markov: Passeio Aleatório

A posição da cadeia no instante t = n é descrito probabilisticamente pela distribuição $w_1 + ... + w_n$.

Se f(1) = p, f(-1) = q e f(0) = r com p + q + r = 1 então as probabilidades de transição são dadas por:

$$P(x,y) = \begin{cases} p, & \text{if } y = x+1\\ q, & \text{if } y = x-1\\ r, & \text{if } y = x\\ 0, & \text{if } y \neq x-1, x, x+1 \end{cases}$$

- Objetivo
- Resumo
- Cadeias de Markov
- Monte Carlo
- Monte Carlo via Cadeias de Markov
- 6 Oscilador Quântico Anarmônico
- Integrais de caminho
- Simulação via MCMC
- Resultados

O que são os métodos de Monte Carlo?

O objetivo dos métodos de Monte Carlo é resolver um ou os dois dos seguintes problemas:

- Gerar amostras* $\{x^{(r)}\}_{r=1}^R$ de uma distribuição de probabilidade P(x).
- Estimar os valores esperadores das funções sobre essa distribuição.

$$\Phi = \langle \phi(x) \rangle = \int d^N x P(x) \phi(x) \tag{4}$$

A distribuição de probabilidade P(x) é o que chamados **densidade alvo** ou **função alvo**, essa pode ser uma distribuição de física estatística ou uma distribuição condicional de uma problema de modelagem.

* A palavra "amostra" é usada no sentido: uma amostra de uma distribuição P(x) é uma única realização x no qual sua distribuição de probabiliadade é P(x). Essa definição constrasta com a usual utilizada em estatística, onde "amostra" se refere a uma colução de realizações $\{x\}$.

O que são os métodos de Monte Carlo?

Se solucionarmos o primeiro problema (amostragem), acabamos solucionando o segundo problema, pois, utilizamos as amostras randômicas $\{x^{(r)}\}_{r=1}^R$ no estimador

$$\hat{\Phi} = \frac{1}{R} \sum_{r} \phi(x^{(r)}) \tag{5}$$

Claramente, se os vetores $\{x^{(r)}\}_{r=1}^R$ são gerados de P(x) então o valor esperado $\hat{\Phi}$ é Φ . É importante notar que a medida que o número de amostras R aumenta, a variância de $\hat{\Phi}$ irá diminuir com $\frac{\sigma^2}{R}$, onde σ^2 é a variância de ϕ ,

$$\sigma^2 = \int d^N x P(x) (\phi(x) - \Phi)^2$$
 (6)

• A acurácia da estimativa partir de Monte Carlo (5) é independente da dimensionalidade do espaço amostral.

Primeiro, devemos assumir que a densidade alvo que desejamos gerar amostras, P(x), possa ser computado, pelo menos com uma constante multiplicativa. Isto é, podemos computador uma função $P^*(x)$, tal que

$$P(x) = \frac{P^*(x)}{Z} \tag{7}$$

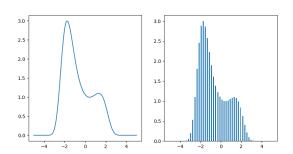
Naturalmente, se podemos computar $P^*(x)$, poderíamos diretamente computar P(x), contudo, tipicamente não conhecemos a constante de normalização

$$Z = \int d^N x P^*(x) \tag{8}$$

Além disso, ainda que conhecessemos Z, gerar amostras de P(x) ainda é uma ação custosa, especialmente em espaços N-dimensionais.

Para demonstrar a dificuldade, seja o caso unidimensional tal que queiramos gerar amostras de P(x) que satisfaça (7), onde

$$P^*(x) = \exp[0.4(x - 0.4)^2 - 0.08x^4], x \in (-\inf, \inf)$$
 (9)



Mesmo que possamos traçar a função (figura 16, esquerda), isso não significa que conseguimos gerar amostras dela. Contudo, se podemos discretizar a variável x e pedir amostras discretas da distribuição de probabilidade sobre um conjunto uniformemente separado de pontos $\{x_i\}$ (figura 16, direita).

Se podemos computar $p_i^* = P^*(x_i)$ para cada ponto x_i , então podemos calcular

$$Z = \sum_{i} p_i^* \tag{10}$$

е

$$p_i = \frac{p_i^*}{Z} \tag{11}$$

de tal sorte que podemos amostrar a distribuição de probabilidade $\{p_i\}$ através da geração de números randômicos. Contudo, qual é o custo desse procedimento e o quanto ele escala com a dimensionalidade do espaço, N?

Para computar Z (10), deveriamos visitar cada ponto do espaço. Na figura (16, direita) existem 50 pontos uniformemente espaçados em uma dimensão. Se nosso sistema tivesse N dimensões, diagmos N=1000, o número de pontos correspondentes seria 50^{1000} . Ainda que cada componente x_n tivesse apenas dois valores discretos, o números de cálculos para avaliar P^* seria 2^{1000} .

Um sistema com 2^{1000} estados é uma coleção de 1000 spins, por exemplo, um fragmento 30×30 de um modelo de Ising onde a distribuição de probabilidade é proporcional a

$$P^*(x) = \exp[-\beta E(x)] \tag{12}$$

e $x_n \in \{\pm 1\}$ e

$$E(x) = -\frac{1}{2} \sum_{m,n} J_{m,n} x_m x_n - \sum_n H_n x_n$$
 (13)

A função energia E(x) é facilmente avaliada para qualquer x. No entando, se quisersemos avaliar a função para todos os estados x, o tempo computacional requirido seria 2^{1000} avaliações da função.

Estratégia: Amostragem Uniforme

Como não podemos visitar cada posição x no estado de espaço, consideramos resolver o segundo problema (estimar os valores esperados da função $\phi(x)$) gerando amostras randômicas $\{x^{(r)}\}_{r=1}^R$ uniformemente do espaço de estado e computado $P^*(x)$ nesses pontos. Assim, poderíamos introduzir Z_R , definido por

$$Z_R = \sum_{r=1}^R P^*(x^{(r)}),$$
 (14)

e estimar $\Phi = \int d^N x \phi(x) P(x)$ por

$$\hat{\Phi} = \sum_{r=1}^{R} \phi(x^{(r)}) \frac{P^*(x^{(r)})}{Z_R}$$
 (15)

Estratégia: Amostragem Uniforme

A eficácia dessa estratégia depende fortemente das funções $\phi(x)$ e $P^*(x)$. Vamos assumar que $\phi(x)$ é bem comportada variando suavemente e concentrada na natureza de $P^*(x)$. Uma distribuição altamente dimensional é normalmente concentrada em uma pequena região do espaço de estado conhecida como conjunto típico T, onde o volume é dado por $|T| \simeq 2^{H(X)}$, onde H(x) é a entropia de Shannon-Gibbs da distribuição de probabilidade P(x),

$$H(X) = \sum_{x} P(x) \log_2 \frac{1}{P(x)}$$
 (16)

Se toda massa probabílistica está localizada no conjunto típico e phi(x) é uma função bem comportada, o valor de $\Phi = \int d^N x \phi(x) P(x)$ será determinado principalmente pelos valores de $\phi(x)$ tomados no conjunto típico. Ou seja, a amostragem uniform só dará estimativas razoáveis para Φ se o número de amostrar R for suficientemente grande para encontrarmos o conjunto típico.

* Modelo de Ising para temperaturas intermediárias irá requerir $R_{min} \simeq 2^{N/2}$.

Métodos de Monte de Carlo

- Amostragem por importância
- Amostragem por rejeição
- Algoritmo de Metropolis-Hastings
- Amostragem de Gibbs

- Objetivo
- Resumo
- Cadeias de Markov
- Monte Carlo
- 5 Monte Carlo via Cadeias de Markov
- Oscilador Quântico Anarmônico
- Integrais de caminho
- Simulação via MCMC
- Resultados

Monte de Carlo via Cadeias de Markov

A abordagem MCMC para amostrar de uma distribuição de probabilidade Q(x) é a construção de uma cadeia de Markov irredutível e aperiódica com distribuição estacionária Q(x), e rodar a cadeia por um tempo suficientemente longo até que a cadeia convirja para distribuição estacionária.

A densidade proposta (ou distribuição de referência) Q(x) para amostragem P(x) não precisa ser completamente similiar à P(x), a densidade proposta apenas precisa depender do estado atual x da cadeia.

- Algoritmo de Metropolis-Hastings
- Amostragem de Gibbs

Algoritmo de Metropolis-Hastings

Dado $X^t = x^t$, o algoritmo para gerar X^{t+1} é dado por:

- Gerar um candidato X' de uma **densidade proposta** Q(x)
- ② Determinar se aceitamos um novo estado a partir da razão $a = \frac{P^*(x')}{P^*(x^{(t)})} \frac{Q(x^t;x')}{Q(x';x^t)}$
- $oldsymbol{\circ}$ Se $a \geq 1$ o novo estado é aceitado, do contrário ele é aceitado com probabilidade a.

Se o passo é aceitado, fazemos $x^{(t+1)}=x'$. Se o passo é rejeitado, fazemos $x^{(t+1)}=x^{(t)}$. Note que, os casos rejeitados não são descartados imeditamente e não influenciam na amostras $\{x^{(r)}\}$ coletadas. A rejeição causa o estado atual a ser escrito em uma lista de pontos em outro ponto no tempo.

- Objetivo
- Resumo
- Cadeias de Markov
- Monte Carlo
- 5 Monte Carlo via Cadeias de Markov
- 6 Oscilador Quântico Anarmônico
- Integrais de caminho
- Simulação via MCMC
- Resultados

Oscilador Quântico Anarmônico

- Um dos poucos problemas em sistema mecânicos quânticos solúveis
- Modelo apresenta problemas em alguns regimes
- Exemplo: Expansão termal de sólidos
- Termos de ordem superior para aumentar a interação potencial
- $\lambda = 0 \rightarrow$ oscilador harmônico!

$$\hat{H}\psi = \left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{m\omega^2 x^2}{2} + \lambda x^4\right)\psi = E\psi \tag{17}$$

onde m é a massa da partícula, $\omega=(k/m)^{1/2}$ é a frequência de oscilação, k é a rigidez do potencial e λ é a constante de acoplamento do termo quártico do potencial.

Soluções para o OQA

- ullet Teoria de pertubação (2) ightarrow diverge para $\lambda>0$
- Integração numérica da equação de Schrödinger
- Avaliar a integral de caminho para esse sistema com o método MCMC
- Métodos de Monte Carlo Variacionais

$$E_0(\lambda) = \frac{1}{2}\hbar\omega + \hbar\omega \sum_{n=1}^{\infty} A_n \left(\frac{\lambda\hbar}{m\omega^2}\right)^n$$
 (18)

- Objetivo
- Resumo
- Cadeias de Markov
- Monte Carlo
- 5 Monte Carlo via Cadeias de Markov
- 6 Oscilador Quântico Anarmônico
- Integrais de caminho
- Simulação via MCMC
- Resultados

Integrais de caminho de Feynman

A equação de Schrödinger dependente do tempo,

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi,\tag{19}$$

para o operador Halmitoniano

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}),\tag{20}$$

tem a solução formal

$$\psi(x,t) = e^{-i\hat{H}t/\hbar}\psi(x,0) \tag{21}$$

onde o fator exponential é o operador de evolução.

29 / 43

Integrais de caminho de Feynman

A conexão com as integrais de caminho advém dos elementos de matriz do operador de evolução entre quaisquer dois autoestados da posição inicial e final.

$$\langle x_f, t_f | x_i, t_i \rangle = \langle x_f | e^{-i\hat{H}(t_f - t_i)/\hbar} | x_i \rangle = \int [Dx(t)] \exp \left[-\frac{1}{\hbar} \int_{t_i}^{t_f} L(x(t)) \right] dt \quad (22)$$

onde L é a Lagrangeana clássica correspondente ao operador Hamiltoniano:

$$L(x(t)) = \frac{m}{2} \left(\frac{dx}{dt}\right)^2 - V(x(t))$$
 (23)

- [Dx(t)] significa que a integral inclui todos os caminhos (x,t).
- A fase de cada caminho é determinada pela ação clássica (8) sobre o caminho.

$$S = \int_{t_i}^{t_f} L(x(t)) dt$$
 (24)

Integrais de caminho com tempo imaginário

• Solução alternativa \rightarrow onde t é substituído por $-i\tau$.

$$\langle x_f, \tau_f | x_i, \tau_i \rangle = \langle x_f | e^{-i\hat{H}(\tau_f - \tau_i)/\hbar} | x_i \rangle = \int [Dx(\tau)] \exp \left[-\frac{1}{\hbar} \int_{\tau_i}^{\tau_f} L_E(x(\tau)) \right] d\tau$$
(25)

$$L_{E}(x(\tau)) = \frac{m}{2} \left(\frac{dx}{d\tau}\right)^{2} - V(x(\tau))$$
 (26)

A energia E_0 do estado fundamental pode ser obtida através do valor esperado do Hamiltoniano (1) onde, em conjunto com o Teorema do Virial, temos:

$$E_0 = m\omega^2 \langle x^2 \rangle + 3\lambda \langle x^4 \rangle \tag{27}$$

- Objetivo
- Resumo
- Cadeias de Markov
- Monte Carlo
- Monte Carlo via Cadeias de Markov
- 6 Oscilador Quântico Anarmônico
- Integrais de caminho
- 8 Simulação via MCMC
- Resultados

Monte Carlo

• Simulações de Monte Carlo (MC) são feitas em intervalos de tempo com N_{τ} incrementos δ_{τ} com pontos $x_n = n\delta_{\tau}$ para $n = 0, 1, 2, 3..., N_{\tau}$

$$\langle x_f, \tau_f | x_i, \tau_i \rangle = \langle x_f | e^{-i\hat{H}(\tau_f - \tau_i)/\hbar} | x_i \rangle = \lim_{N_\tau \to \infty} \int \prod_{k=1}^{N_\tau} dx_k \left(\frac{m}{2\pi\hbar\delta_\tau} \right) e^{-S(x_1, x_2, \dots, x_n)/\hbar},$$
(28)

onde

$$S(\lbrace x_k \rbrace) = \delta_\tau \sum_{i=1}^{N_\tau} \left[\frac{m}{2} \left(\frac{x_{i+1} - x_i}{\delta_\tau} \right) + V(x_i) \right]$$
 (29)

33 / 43

Mudança de variável

Fazendo a mudança de variáveis, de tal sorte que cada variável seja expressa em função do espaçamento $\delta_{ au}$,

$$\tilde{m} = m\delta_{\tau} \tag{30}$$

$$\tilde{\omega} = \omega \delta_{\tau} \tag{31}$$

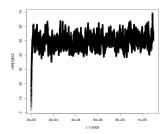
$$\tilde{x_i} = \frac{x_i}{\delta_{\tau}} \tag{32}$$

$$S(\{x_k\}) = \sum_{i=1}^{N_{\tau}} \left[\frac{\tilde{m}}{2} (x_{i+1} - x_i)^2 + \frac{\tilde{m}\tilde{\omega}^2 \tilde{x}_i^2}{2} + \tilde{\lambda}\tilde{m}^2 \omega^3 \tilde{x}_i^4 \right]$$
(33)

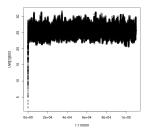
Aplicação: Monte Carlo via Cadeias de Markov

- Método baseado em determinar a estatística dos caminhos observáveis.
- Começamos em um caminho inicial com um array de números aleatórios (hot start) ou zeros (cold start). O caminho é atualizado aplicando o algoritmo de Metropolis-Hastings para cada elemento xi do caminho em ordem aleatória, chamado de sweep.
- O elemento central do método MCMC é a seleção de caminhos para os cálculos. Os caminhos devem representar o equilíbrio da distribuição, então os caminhos devem primeiro chegar ao equilíbrio em uma configuração inicial.
- O número N_{therm} de *sweeps* requeridos para obter o equilíbrio é determinado quando a quantidade mensurada flutua sobre um estado estável.

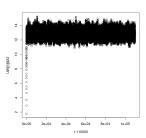
Equilíbrio MCMC



(a)
$$\tilde{\lambda}=0$$
, $\delta_{ au}=0$,1



(b)
$$\tilde{\lambda}=1$$
, $\delta_{ au}=0.1$

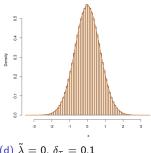


(c)
$$ilde{\lambda}=$$
 10, $\delta_{ au}=$ 0,1

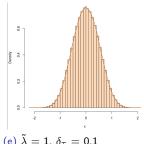
- Objetivo
- Resumo
- Cadeias de Markov
- Monte Carlo
- 5 Monte Carlo via Cadeias de Markov
- 6 Oscilador Quântico Anarmônico
- Integrais de caminho
- Simulação via MCMC
- Resultados

Energia no estado fundamental: Densidade de Probabilidade

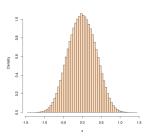
Número de sweeps: 110000







(e) $\tilde{\lambda}=1,\,\delta_{ au}=0.1$



(f)
$$\tilde{\lambda}=10,\,\delta_{ au}=0.1$$

Energia no estado fundamental: $ilde{\lambda}=0$

 \bullet $\delta_{ au}=0.1$

N. Sweeps	Energia Média	Erro Estatístico	
500	0.309	0.043	
1000	0.37	0.045	
1500	0.432	0.057	
2000	0.434	0.044	
2500	0.464	0.035	
3000	0.479	0.033	
3500	0.467	0.041	
4000	0.471	0.025	
8000	0.494	0.019	
8500	0.494	0.022	
9000	0.483	0.024	
9500	0.466	0.012	
10000	0.5	0.021	
110000	0.497	0.003	

Energia no estado fundamental: $ilde{\lambda}=1$

 $\delta_{ au} = 0.1$

N. Sweeps	Energia Média	Erro Estatístico	
500	0.664	0.076	
1000	0.742	0.061	
1500	0.771	0.056	
2000	0.744	0.028	
2500	0.757	0.023	
3000	0.788	0.026	
3500	0.788	0.023	
4000	0.795	0.025	
8000	0.775	0.011	
8500	0.781	0.011	
9000	0.789	0.011	
9500	0.793	0.009	
10000	0.789	0.011	
110000	0.793	0.003	

Energia no estado fundamental: $ilde{\lambda}=10$

$$\bullet$$
 $\delta_{ au}=$ 0.1

N. Sweeps	Energia Média	Erro Estatístico
110000	1.451	0.002

Resultados

• Número de sweeps: 110000

$\tilde{\lambda}$	Teoria de	Per-	MCMC	Exato
	turbação			
0	0.50		0.497 ± 0.003	0.500000
1	1.25		0.793 ± 0.003	0.803771
10	8.00		1.451 ± 0.002	1.504972

Referências I



D. J. C. Mackay.

Introduction to monte carlo methods.



R. Rosenfelder.

Path integrals in quantum physics, 2017.



Shikhar Mittal, Marise J E Westbroek, Peter R King, and Dimitri D Vvedensky.

Path integral monte carlo method for the quantum anharmonic oscillator. *European Journal of Physics*, 41(5):055401, Aug 2020.