



UNIVERSIDAD SIMÓN BOLÍVAR

LA-CoNGA PHYSICS

---

Métodos de Monte Carlo.  
b) Muestreo de importancia

---

*Profesora*  
GLORIA BUENDÍA

3 de abril de 2023

# Contenido General

<b>1</b>	<b>Técnicas de muestreo de importancia. Métodos de Monte Carlo.</b>	<b>1</b>
<b>2</b>	<b>Cadenas de Márkov</b>	<b>2</b>
<b>3</b>	<b>Cadena de Márkov Homogénea</b>	<b>3</b>
3.1	Propiedades . . . . .	3
<b>4</b>	<b>Algoritmo de Metropolis (1953)</b>	<b>6</b>
<b>5</b>	<b>Ejemplo: algoritmo de Metropolis. Modelo de Ising en una red cuadrada</b>	<b>7</b>
<b>6</b>	<b>Agunas ventajas y desventajas de este tipo de algoritmo</b>	<b>9</b>
<b>7</b>	<b>Algunos puntos a tomar en cuenta cuando se hacen simulaciones numéricas</b>	<b>10</b>

# 1. Técnicas de muestreo de importancia. Métodos de Monte Carlo.

Queremos calcular promedios sobre configuraciones  $[\varphi]$  de la forma

$$\langle O \rangle = \int [d\varphi] O[\varphi] P[\varphi] \quad (1)$$

**Idea:** seleccionar un número de configuraciones del espacio de fase de aquella región que sea la que más contribuya a la integral (o suma en el caso de variables discretas).

Se sustituye el promedio  $\langle O \rangle$  por un promedio sobre  $M$  configuraciones tal que

$$\langle O \rangle \rightarrow \bar{O} = \frac{1}{M} \int [d\varphi] O(\varphi) \quad (2)$$

dónde las  $M$  configuraciones son seleccionadas con la probabilidad  $P[\varphi]$  que en el caso de un ensemble canónico es el factor de *Boltzmann*

$$P(\varphi) = \frac{e^{-\beta\mathcal{H}(\varphi)}}{\int [d\varphi] e^{-\beta\mathcal{H}(\varphi)}} \quad (3)$$

Para que  $\bar{O}$  sea una buena aproximación a  $\langle O \rangle$  la secuencia de configuraciones se genera realizando una caminata aleatoria en el espacio de configuraciones. Esta secuencia debe ser una cadena de *Márkov*. Andrei Márkov (1906).

La secuencia de configuraciones  $\varphi_1, \varphi_2, \dots, \varphi_M$  es una realización de la cadena de *Márkov*, definida por una configuración original y una matriz estocástica  $W_{ij}$  que da la probabilidad de que en un paso de la transición se vaya de la configuración  $\varphi_i$  a la  $\varphi_j$ .

## 2. Cadenas de Márkov

: Vienen dadas por  $\begin{cases} \varphi_0 \\ W_{ij} \end{cases}$

Queremos construir la matriz de transiciones  $W_{ij}$  de forma que la distribución límite de la cadena sea independiente de la configuración inicial,  $\varphi_0$ , y tenga la probabilidad  $P[\varphi]$ , en el caso de promedios térmicos  $P[\varphi]$  debe ser la probabilidad de *Boltzmann* (3).

Si la probabilidad de obtener  $\varphi_j$  después de  $M$  pasos es (comenzando por  $\varphi_i$ )

$$W_{i \rightarrow j}^M = \sum_{\ell} W_{i \rightarrow \ell}^{M-1} W_{\ell \rightarrow j} \quad \forall i, j \quad (4)$$

$[W_{i \rightarrow j}^M$  probabilidad de que empezando en  $\varphi_i$  después de  $M$  pasos lleguemos a  $\varphi_j$ ]. Queremos que

$$\lim_{M \rightarrow \infty} W_{i \rightarrow j}^M = P_j \quad \forall i, j \quad (5)$$

es decir que después de un número muy grande de pasos, las nuevas configuraciones sean generadas con la probabilidad deseada, independientemente de la configuración inicial. ¿Que condiciones debe cumplir  $W$  para que esto ocurra?  $W_{ij}$  la probabilidad condicional de que en un paso se vaya de la configuración  $i$  a la  $j$ .

$W_{ij}$  debe verificar las condiciones que definen una cadena de *Márkov* homogénea, para que se cumpla (5)

$$W_{ij} > 0 \quad \forall i, j$$

### 3. Cadena de Márkov Homogénea

$$\varphi_0, \varphi_1, \varphi_2 \dots \varphi_M$$

#### 3.1. Propiedades

i)

$$\sum_j W_{ij} = 1 \quad \forall i \quad \text{Normalizacin} \quad (6)$$

El sistema siempre debe de llegar a alguna configuración de la cadena

ii) Si  $P(\varphi_k) > 0$  y  $P(\varphi_j) > 0 \implies W_{kj} > 0$ 

**Ergodicidad** Todos los estados posibles deben ser accesibles desde cualquier configuración.

iii) Balance detallado - Reversibilidad microscópica

$$\frac{W_{kj}}{W_{jk}} = \frac{P_j}{P_k} \implies P_k W_{kj} = P_j W_{jk} \quad \forall j, k$$

(No hay suma sobre índices repetidos)

Debemos probar que

1. Que la cadena converge.
2. Que converge con la probabilidad deseada.
3. Que después que converge se mantiene estable.

$$W_{ij}^M = P_j \implies W_{ik}^{M+1} = P_k \quad \forall i$$

→ Una vez que la cadena converge (dá la probabilidad deseada) las configuraciones siguientes tienen la probabilidad deseada  $\implies$  Estabilidad

*Prueba:*

Si la cadena converge después de  $M$  pasos:  $W_{ij}^M = P_j$ , que pasa con  $W_{ik}^{M+1}$

$$W_{ik}^{M+1} = \sum_{\ell} W_{i\ell}^M W_{\ell k} = \sum_{\ell} P_{\ell} W_{\ell k}$$

La condición iii) balance detallado.

$$P_\ell W_{\ell k} = P_k W_{k\ell} \quad (\text{No hay suma sobre índices repetidos})$$

$$\implies W_{ik}^{M+1} = \sum_\ell P_\ell W_{\ell k} = \sum_\ell P_k W_{k\ell} = P_k \sum_\ell W_{k\ell} = P_k$$

ya que de i)  $\sum_\ell W_{k\ell} = 1$

Otra forma de ver que una vez alcanzado el equilibrio éste es estable es escribir

$$\frac{dP_i}{dt} = - \sum_k P_i W_{ik} + \sum_k P_k W_{ki}$$

La probabilidad de tener una configuración  $i$  en un momento dado depende de si otras configuraciones cambian a  $i$  (+) ó si las configuraciones  $i$  cambian a otras (-)

$$\frac{dP_i}{dt} = \sum_k [P_k W_{ki} - P_i W_{ik}] = 0 \quad (\text{Balance detallado})$$

En el equilibrio la probabilidad no depende del tiempo.

Ahora debemos probar que la desviación entre las probabilidades que se obtienen y la probabilidad deseada disminuye con cada paso de la cadena de *Márkov*.

Definimos  $d_M$  = distancia después del paso  $M$  entre las distribución obtenida y la distribución deseada.

$$d_M = \sum_K |W_{iK}^M - P_K| \quad ; \quad d_{M+1} = \sum_K |W_{iK}^{M+1} - P_K|$$

$$W_{iK}^{M+1} = \sum_j W_{ij}^M W_{jK}$$

$$\implies d^{M+1} = \sum_K \left| \sum_j W_{ij}^M W_{jK} - P_K \right| = \sum_K \left| \sum_j W_{ij}^M W_{jK} - P_K \sum_j W_{Kj} \right|$$

Como  $\sum_j W_{Kj} = 1$

$$d^{M+1} = \sum_K \left| \sum_j (W_{ij}^M W_{jK} - P_K W_{Kj}) \right|$$

↳ Ojo aquí no hay suma sobre  $K$

$$= \sum_K \left| \sum_j (W_{ij}^M W_{jK} - P_j W_{jK}) \right| \quad (\text{Balance detallado})$$

Como  $|A + B + C + \dots| \leq |A| + |B| + |C| + \dots$

$$\begin{aligned}
 d^{M+1} &= \sum_K \left| \sum_j (W_{ij}^M - P_j) W_{jK} \right| && \text{Recordemos que } W_{jK} > 0 \\
 &= \sum_K \left| (W_{i1}^M - P_1) W_{1K} + (W_{i2}^M - P_2) W_{2K} + \dots \right| \\
 &\leq \sum_j \left| W_{ij}^M - P_j \right| \sum_K W_{jK} && ; \quad \sum_K W_{jK} = 1 \\
 d^{M+1} &\leq \sum_j \left| W_{ij}^M - P_j \right| = d^M
 \end{aligned}$$

La distancia a la probabilidad deseada se acorta con cada paso de la cadena.

En 1953, Metropolis, Rosenbluth, Rosenbluth, Teller, Teller, investigadores que trabajaban en Los Alamos, en el laboratorio creado durante la segunda guerra mundial para desarrollar el llamado "Proyecto Manhattan" publicaron un artículo que sentó las bases para lo que sería una nueva era de simulaciones basadas en algoritmos basados en cadenas de Márkov.

## 4. Algoritmo de Metropolis (1953)

Este algoritmo propone una matriz de transición muy sencilla que satisface las condiciones de la cadena de Márkov.

$$W_{ij} = \begin{cases} C & E_i > E_j \\ C e^{-\beta(E_j - E_i)} & E_j > E_i \end{cases}$$

$C$  es una constante de normalización que asegura  $\sum W_{ij} = 1$

Se escoge una configuración inicial arbitraria  $\varphi_i$ , se calcula su energía

Se genera otra configuración al azar  $\varphi_j$ ,  $W_0(i \rightarrow j) = W_0(j \rightarrow i)$  (por ejemplo volteando un espín)

- Calcule  $(E_j - E_i) = \Delta E$ ,  $e^{-\beta\Delta E}$
- Escoja un número al azar  $r$  entre 0 y 1, con probabilidad uniforme
- Si  $r \leq e^{-\beta\Delta E}$  acepte la nueva configuración  
 $r > e^{-\beta\Delta E}$  rechazela

Note que si  $\Delta E < 0$ , la nueva configuración siempre es aceptada, no se necesita generar números al azar.

Este algoritmo acepta todas las configuraciones que bajan la energía, pero también acepta con probabilidad  $e^{-\beta\Delta E}$  algunas que la aumentan.

Probar que el algoritmo de Metropolis satisface balance detallado

a)  $E_j > E_i$

$$W_{ij} = P_0(i \rightarrow j) e^{-\beta(E_j - E_i)}$$

$P_0(i \rightarrow j)$  es la probabilidad de generar una configuración  $j$  al azar a partir de otra configuración  $i$

$$\begin{aligned} W_{ji} &= P_0(j \rightarrow i) \rightarrow \text{Bajar la energía siempre se acepta} \\ \implies \frac{W_{ij}}{W_{ji}} &= \frac{P_0(i \rightarrow j)}{P_0(j \rightarrow i)} e^{-\beta(E_j - E_i)} & P_0(i \rightarrow j) &= P_0(j \rightarrow i) \\ \implies e^{-\beta E_i} W_{ij} &= e^{-\beta E_j} W_{ji} & & \text{(Balance detallado)} \end{aligned}$$

b)  $E_j < E_i$

$$\begin{aligned}
 W_{ij} &= P_0(i \rightarrow j) \\
 W_{ji} &= P_0(j \rightarrow i)e^{-\beta(E_i - E_j)} \\
 \frac{W_{ij}}{W_{ji}} &= \frac{P_0(i \rightarrow j)}{P_0(j \rightarrow i)} e^{\beta(E_i - E_j)} \\
 e^{-\beta E_i} W_{ij} &= e^{-\beta E_j} W_{ji} \rightarrow \text{Balance detallado}
 \end{aligned}$$

## 5. Ejemplo: algoritmo de Metropolis. Modelo de Ising en una red cuadrada

Tenemos un sistema de spines que toman valores  $\pm 1$ , supongamos un Hamiltoniano muy simple en que solo interaccionan con sus primeros vecinos y con un campo externo constante. Este es un modelo muy sencillo para explicar algunos aspectos del comportamiento magnético de materiales, pero sobre todo es muy importante para explorar transiciones de fase.

Tenemos un sistema de  $N=L \times L$  espines, con un Hamiltoniano de la forma

$$H = -J \sum_{nn} S_i S_j - h \sum_i S_i \quad (7)$$

escojo un espin al azar  $S_k$  y trato de generar una nueva configuración en que este espin está invertido,  $S_k \rightarrow -S_k$ .

Energía inicial  $E_0$  y energía final  $E_f$

$$\begin{aligned}
 E_0 &= \dots - JS_k(S_{k+1} + S_{k+L}) - hS_k \dots \\
 E_f &= \dots - J(-S_k)(S_{k+1} + S_{k+L}) - h(-S_k) \dots
 \end{aligned}$$

los otros términos de la energía no cambian. Aquí hemos considerado como primeros vecinos solo el de la derecha y el de abajo del espín para evitar contar doble las interacciones.

$$\Delta E = E_f - E_i = 2JS_k(S_{k+1} + S_{k+L}) + 2hS_k \quad (8)$$

Si  $\Delta E \leq 0$  se acepta el cambio, la nueva configuración tiene el espin  $-S_k$  y su nueva energía es  $E \rightarrow E + \Delta E$ . Si  $\Delta E > 0$  se escoge un número al azar  $r$ ,  $0 \leq r \leq 1$ . Si  $r \leq e^{-\beta \Delta E}$  se acepta la nueva configuración, si no se rechaza.

A partir de una configuración de espines, se considera que se ha generado una configuración nueva luego que se han hecho  $N=L \times L$  intentos de voltear espines al azar. Esto es lo que constituye un paso de Monte Carlo, los promedios son generados sobre las configuraciones obtenidas después de cada paso de Monte Carlo.

---

$$\langle O \rangle = \frac{1}{MCS} \sum_{c=1}^{MCS} O_c \quad (9)$$

donde  $O_c$  es el valor de  $O$  en la configuración  $c$ , note que ahora no tenemos que multiplicar por la probabilidad de Boltzmann ya que el algoritmo me asegura que las configuraciones están generadas con la configuración deseada.

## 6. Algunas ventajas y desventajas de este tipo de algoritmo

### a) Hamiltonianos con interacciones locales

Algoritmos tipo Metropolis funcionan muy bien si las interacciones son locales, cada spin interactúa con un número reducido de vecinos. En este caso el cálculo de la variación de la energía luego de voltear un espín es muy rápido (y como veremos posteriormente podemos hacer una tabla con los posibles cambios lo que hace que este proceso tenga muy poco costo computacional), pero si el Hamiltoniano del sistema a estudiar es no lineal o tiene interacciones de muy largo alcance el proceso se hace extremadamente lento. Por ejemplo imaginen que el Hamiltoniano tiene un término no local, como un determinante (no es un caso traído por los pelos, es exactamente el caso para sistemas de tienen grados de libertad fermiónicos), un mínimo cambio que se haga en el sistema requiere que se recalcule la energía, en este caso el determinante. A pesar de que hay muchos métodos eficientes para calcular determinantes, sigue siendo un proceso que consume mucho tiempo de computación, generar una nueva configuración requiere del cálculo de  $N$  determinantes (que depende de la dimensión del sistema) y en muchos casos y dependiendo de las condiciones la mayoría de cambios son rechazados con lo que es muy lento generar nuevas configuraciones.

### b) Generación de configuraciones independientes

Con este algoritmo en muchos casos hay una fuerte correlación entre una configuración y la siguiente. Hay que tener mucho cuidado al estimar los errores de los promedios. En algunos casos, por ejemplo para algunas temperaturas, casi ninguna configuración nueva es aceptada, por lo que podemos estar haciendo promedios sobre la misma configuración una y otra vez. En el próximo tema trataremos este problema, llamado critical slowing down.

## **7. Algunos puntos a tomar en cuenta cuando se hacen simulaciones numéricas**

### a) Pasos de calentamiento

Es inevitable comenzar nuestras simulaciones en alguna configuración inicial, las medidas que tomemos en esa configuración inicial y las que le siguen no están distribuidas con la probabilidad deseada, debemos esperar a que la cadena de Márkov converga antes de empezar a tomar medidas. En la práctica, como sabemos que las configuraciones que estamos generando son independientes de la configuración inicial? es decir que hemos generado un número de pasos suficientes tal que la cadena converge. Lo que se hace es medir algunas cantidades físicas y ver el número de pasos de Monte Carlo que se necesitan para ver que las cantidades medidas se estabilizan. Estos pasos que se necesitan para que el sistema alcance el equilibrio son llamados pasos de calentamiento. Para los promedios sólo se deben de tomar en cuenta configuraciones obtenidas despues de estos pasos.

### b) Efectos de tamaño finito

La mecánica estadística funciona para estudiar sistemas macroscópicos donde el número de grados de libertad, partículas, etc es enorme (del orden del número de Avogadro,  $10^{23}$  pero en la práctica sólo podemos estudiar sistemas relativamente reducidos en el computador. Como lidiamos con este problema? Hablaremos de ello luego.

### c) Cálculo de errores

Ya que las cantidades medidas con estos algoritmos de Monte Carlo son promedios estadísticos sobre configuraciones, es fundamental que tengamos un estimado del margen de error cometido. Como veremos mas adelante esto no es tan sencillo como calcular la desviación cuadrática media.