

Física Computacional

Modelo de Ising: Una Introducción al Método Monte Carlo

1. Introducción.

- En el contexto de la **Mecánica Estadística de equilibrio** se deriva que un sistema en equilibrio termodinámico con una temperatura T , número de partículas N y volumen V , la probabilidad de encontrar a las N partículas en un estado $X = (x(1), \dots, x(N))$ está dada por

$$P_N(X) = Z^{-1} e^{-\beta H_N(X)} \quad Z = \int dX e^{-\beta H_N(X)} \quad (1)$$

donde $\beta = 1/(k_B T)$ y k_B es la constante de Boltzmann. Z es la **función de partición** y $H_N(X)$ es el hamiltoniano que define la dinámica microscópica correspondiente a las N partículas. Así, conocida $H_N(X)$, podemos calcular el **valor esperado** de cualquier variable (que coincidirá con el valor **observado** en un experimento):

$$\langle A \rangle = Z^{-1} \int dX e^{-\beta H_N(X)} A_N(X) \quad (2)$$

- La Mecánica Estadística conecta las ecuaciones del movimiento microscópicas de un sistema de partículas con sus propiedades termodinámicas macroscópicas. Sin embargo, en la práctica, sólo en unos pocos casos de interés, podemos realizar de forma analítica alguna de esas integrales N dimensionales. De esta forma debemos acudir a métodos numéricos para conocer el comportamiento del sistema.
- La estrategia más natural que se aplica es usar la ley de los grandes números. Ésta nos dice que **independientemente** de la distribución de probabilidad $P_N(X)$, el valor esperado de cualquier observable puede ser hallado mediante sumas parciales de M *realizaciones concretas* de los mismos, esto es

$$\langle A \rangle = \frac{1}{M} \sum_{i=1}^M A(X_i) + O(M^{-1/2}) \quad (3)$$

donde X_i deben de seguir la distribución $P_N(X)$ y han de ser **estadísticamente independientes**.

- Para generar esas configuraciones podemos utilizar la técnica del **rechazo**. Ésta, se basa en el hecho obvio de que si $f(x)$ es una cierta densidad de probabilidad (de una sola variable por sencillez), entonces si una y cumple que $f(y) = 2f(x)$ entonces y es dos veces más probable que x . Osea, lo que tenemos que conseguir es que la **frecuencia relativa** entre eventos sea la correcta. Así el esquema para generar números aleatorios que sigan una cierta distribución $f(x)$ hemos de seguir el siguiente algoritmo:
 - (0) Generar un valor x aleatoriamente y uniformemente en su dominio de definición.
 - (1) Generar un número aleatorio uniforme $\xi \in [0, 1]$.
 - (2) Si $\xi < f(x) / \max_y f(y)$ entonces x se acepta como representante de la distribución $f(x)$.
 - (3) Ir a (0).

Si intentamos utilizar este sencillo algoritmo para generar configuraciones que sigan la distribución $P_N(X)$, nos pasaríamos **todo el tiempo rechazando configuraciones**. Esto es debido a que cuando N es grande (como es el caso en los sistemas que se quieren simular), el hamiltoniano es una variable extensiva en N y la distribución de probabilidad $P_N(X) \simeq \exp -N(e(X) - \langle e \rangle)$ lo que la hace extremadamente picuda alrededor de sus valores de equilibrio termodinámico. De esta forma si aplicamos el algoritmo de **rechazo** tendríamos dos efectos adversos: (1) las configuraciones aleatorias son, normalmente, mucho más numerosas que las demás y por lo tanto, la probabilidad de escoger al azar una configuración aleatoria es prácticamente uno y (2) las configuraciones aleatorias son las que tienen más energía (las correspondientes en un gas a temperatura infinita) por lo que $P_N(X) \simeq 0$ es cero y ningún número aleatorio uniforme generado por nosotros va a ser menor o igual a cero.

- Para solucionar este problema [Metropolis](#), [Rosenbluth](#), [Rosenbluth](#), [Teller](#) y [Teller](#) desarrollaron en 1953 un algoritmo que abandona la idea original de generar configuraciones estadísticamente independientes. En este caso, las configuraciones son construidas a través de **cadenas de Markov**. Una cadena de Markov de eventos

sucesivos X_1, X_2, \dots, X_N tiene como probabilidad:

$$P_N(X_1, X_2, \dots, X_N) = P_1(X_1)T(X_1 \rightarrow X_2)T(X_2 \rightarrow X_3)\dots T(X_{N-1} \rightarrow X_N) \quad (4)$$

donde $T(X \rightarrow Y)$ es la probabilidad de transición de estando en el estado X ir al estado Y . Obviamente

$$\sum_Y T(X \rightarrow Y) = 1 \quad (5)$$

Si $g(X, t)$ es la probabilidad de que en el paso t de la cadena de Markov nos encontremos en el estado X podemos escribir fácilmente su ecuación de evolución

$$g(X, t+1) = \sum_{X'} g(X', t)T(X' \rightarrow X) \quad (6)$$

$$= \sum_{X' \neq X} g(X', t)T(X' \rightarrow X) + g(X, t)T(X \rightarrow X) \quad (7)$$

$$= g(X, t) + \sum_{X' \neq X} [g(X', t)T(X' \rightarrow X) - g(X, t)T(X \rightarrow X')] \quad (8)$$

que no es más que la llamada **Ecuación Maestra**. Una solución estacionaria de este proceso de Markov la encontramos pidiendo que

$$g(X')T(X' \rightarrow X) = g(X)T(X \rightarrow X') \quad (9)$$

donde $g(X)$ es la probabilidad estacionaria de que el sistema se encuentre en el estado X . Esta condición suficiente se denomina **condición de balance detallado**.

- En este contexto, la estrategia a seguir para generar estados X cuya distribución estacionaria sea $P_N(X)$ supuestamente conocida es la siguiente:

- (0) Construimos una $T(X' \rightarrow X)$ a partir de la condición de balance detallado. En nuestro caso $P_N(X) \simeq \exp[-\beta H_N(X)]$ así la condición de balance detallado debe de cumplir

$$T(X' \rightarrow X) = e^{-\beta[H_N(X) - H_N(X')]} T(X \rightarrow X') \quad (10)$$

Obviamente funciones $T(X \rightarrow X')$ que cumplan la anterior ecuación hay infinitas pero vamos a elegir la utilizada por Metropolis et al.

$$T(X \rightarrow X') = \min(1, e^{-\beta[E(X') - E(X)]}) \quad (11)$$

- (1) Dar un estado inicial $X = X_0$.
 - (2) Elegir un estado X' . Generar un número aleatorio uniforme $\xi \in [0, 1]$. Si $\xi < T(X \rightarrow X')$ entonces mover el sistema al estado X' .
 - (3) Ir a (2).
- Esta dinámica de estados que nos hemos inventado **recorre todos los estados accesibles por el sistema** y, para tiempos largos, **tiende a muestrear los estados con probabilidad $P_N(X)$** . Obviamente, si elegimos al azar X' tendríamos el mismo problema que nos encontrábamos con el algoritmo de rechazo. El truco aquí es que *podemos elegir un X' que no sea más que una pequeña modificación de X* . De esta forma si en algún momento de la cadena de Markov las X son las más probables según $P_N(X)$ los siguientes estados generados seguirán estando concentrados alrededor de ese máximo de probabilidad.

2. Modelo Básico

- Lenz en 1927 propuso a su estudiante Ising modelar de forma sencilla unos materiales que sufrían un cambio brusco en sus propiedades magnéticas con la temperatura. A temperaturas altas eran paramagnéticos (magnetización espontánea cero) y a temperaturas suficientemente bajas eran ferromagnéticos (magnetización espontánea no nula). El modelo (hoy llamado de Ising) se define en una red cuadrada bi-dimensional N^2 donde en cada nudo de la red existe una variable llamada de espín $s(i, j)$ ($i, j = 1, 2, \dots, N$) que puede tener dos valores 1 y -1 . La energía de interacción del sistema dada una configuración de espines S es:

$$E(S) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N s(i, j) [s(i, j+1) + s(i, j-1) + s(i+1, j) + s(i-1, j)] \quad (12)$$

donde suponemos que el sistema tiene **condiciones de contorno periódicas** *i.e.* $s(0, j) = s(N, j)$, $s(N+1, j) = s(1, j)$, $s(i, 0) = s(i, N)$ y $s(i, N+1) = s(i, 1)$. Dada una temperatura T , cuando el sistema está en equilibrio termodinámico la probabilidad de encontrar una configuración S viene dada por

$$P(S) = \frac{1}{Z} e^{-\beta E(S)} \quad (13)$$

Así, notar que cuando $T = 0$, $P(S) = \frac{1}{2} \Pi_{i,j} \delta(s(i, j) - 1) + \frac{1}{2} \Pi_{i,j} \delta(s(i, j) + 1)$. Esto es, todos los espines tienen valor 1 o todos tienen valor -1 . Por otra parte, si $T = \infty$ entonces $P(S) = \text{cte}$ o sea, todas las configuraciones son igual de probables. Para generar configuraciones típicas con probabilidad de equilibrio $P(S)$ utilizaremos el algoritmo de Metropolis:

- (0) Dar una $T \in [0, 5]$. Generar una configuración inicial de espines. Por ejemplo una configuración ordenada *i.e.* $s(i, j) = 1 \quad \forall \quad i, j$, o desordenada (elegir aleatoriamente 1 o -1 en cada nudo de la red con probabilidad $1/2$).

- (1) Elegir un punto al azar, (n, m) , de la red.
- (2) Evaluar $p = \min(1, \exp -[\Delta E/T])$ donde $\Delta E = 2s(n, m)[s(n + 1, m) + s(n - 1, m) + s(n, m + 1) + s(n, m - 1)]$. Usar las condiciones de contorno periódicas.
- (3) Generar un número aleatorio uniforme $\xi \in [0, 1]$. Si $\xi < p$ entonces cambiar el signo del espin (n, m) *i.e.* $s(n, m) = -s(n, m)$.
- (4) Ir a (1).

Notar que este algoritmo cumple la premisa anteriormente contemplada, la diferencia entre una configuración S y la siguiente después de un paso de Markov S' es de un solo espín, $s(m, n)$, todos los demás son iguales. De esta forma, la unidad de tiempo básica es el *paso Monte Carlo* que equivale a realizar N^2 intentos de cambio de un espin. Así, en un paso Monte Carlo, en promedio todos los espines han intentado cambiar una vez.

3. Problemas

Obligatorio:

Realizar el programa que simular el Modelo de Ising con el método Monte Carlo. Mostrar su evolución por pantalla para diversas temperaturas.

Voluntario 1:

Dado el tamaño N , elegir una configuración inicial ordenada *i.e.* $s(i, j) = 1$ y dejar evolucionar el sistema 10^6 pasos Monte Carlo (pMC).

Se pide obtener valores medios y sus correspondientes errores para las siguientes magnitudes:

- La magnetización promedio:

$$m_N = \left\langle \frac{1}{N^2} \left| \sum_{i=1}^N \sum_{j=1}^N s(i, j) \right| \right\rangle \quad (14)$$

- La energía media:

$$e_N = \frac{\langle E(S) \rangle}{2N} \quad (15)$$

- El calor específico:

$$c_N = \frac{1}{N^2 T} [\langle E(S)^2 \rangle - \langle E(S) \rangle^2] \quad (16)$$

- La función de correlación:

$$f(i) = \frac{1}{N^2} \sum_{(n,m)} \langle s(n, m) s(n + i, m) \rangle \quad (17)$$

donde $\langle \cdot \rangle$ es un promedio sobre distintas medidas realizadas cada 100 pMC, esto es promedios sobre 10^4 medidas. Realizar el experimento para 10 temperaturas en un intervalo $T \in [1,5, 3,5]$ y para $N = 16, 32, 64$ y 128. Hacer como mínimo el siguiente análisis:

- Describir el comportamiento de las anteriores magnitudes para el rango de temperaturas y tamaños. Comparar con el resultado exácto de Onsager. Describir el efecto del tamaño en cada una de las variables.
- Obtener de la literatura información sobre exponentes críticos y teoría de tamaño finito.
- Estimar el valor del punto crítico: Para cada valor de N obtener una estimación del máximo del calor específico, $T_c(N)$, y estudiar su comportamiento con N extrapolando para $N \rightarrow \infty$.
- Obtener numéricamente el exponente crítico β de la magnetización y comparar con el resultado exacto.

- Estudiar de la función $f(i)$ con la temperatura y el tamaño del sistema. Extraer la longitud de correlación y su exponente crítico característico.

Voluntario 2: Modelo de Hopfield de red neuronal

Usar el modelo de Ising para simular y estudiar el comportamiento emergente de una red neuronal de Hopfield para cualquier temperatura T . Para ello se asumen las siguientes hipótesis del modelo:

1. Cada variable de espín representa el estado de una neurona que está disparando o está sin disparar. Es útil entonces transformar el hamiltoniano de Ising del código $s(i) = -1, 1$ (espín más o menos) al código $s(i) = 1, 0$ (neurona disparando o no) obteniendo:

$$H(\mathbf{s}) = -\frac{1}{2} \sum_{i,j} J_{ij} s_i s_j + \sum_i \theta_i s_i$$

2. Las interacciones entre espines $J_{ij} = w_{ij}$ representan ahora el estado de las sinapsis o conexiones entre neuronas y se denominan pesos sinápticos. Estos vienen dados en término de P configuraciones particulares de la red $\xi^\mu = \{\xi_i^\mu = 1, 0 \mid i = 1, \dots, N\}$, llamados patrones o memorias mediante la regla de aprendizaje Hebbiano:

$$w_{ij} = \frac{1}{a(1-a)N} \sum_{\mu=1}^P (\xi_i^\mu - a)(\xi_j^\mu - a)$$

con $a = \langle \xi_i^\mu \rangle$.

3. En vez de conexiones a vecinos próximos la red de Hopfield asume interacciones de largo alcance y en particular una red totalmente conectada.

4. No se permiten autoconexiones $w_{ii} = 0$, y los umbrales de disparo θ_i vienen dados en términos de los pesos en la forma:

$$\theta_i = \frac{1}{2} \sum_j w_{ij}$$

Se pide:

1. Ver como la red es capaz de recordar los patrones almacenados en función de la temperatura y el número de patrones almacenados en los pesos sinápticos, partiendo de una condición inicial aleatoria o un patrón deformado, y calculando el llamado solapamiento con cada patrón definido como:

$$m^\mu(s) = \frac{1}{a(1-a)N} \sum_i (\xi_i^\mu - a)(s_i - \frac{1}{2})$$

Utilizar diferentes tipo de patrones, incluyendo el caso de patrones aleatorios, y prefijados, como por ejemplo los dígitos del 0 al 9, un conjunto de caras o un conjunto de símbolos.

2. Considerando una red de $N = 400$ neuronas y haciendo $T = 0$, calcular como decae la recuperación de la memoria en función del número de patrones almacenados, y asumiendo que un patrón dado se recuerda sin apenas error cuando el solapamiento es mayor que 0.75, calcular la fracción máxima $\alpha_c = P_c/N$ de patrones que la red podrá almacenar de forma que todos los patrones se pueden recordar. En este caso es mejor trabajar con patrones aleatorios.