

Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas

FI3104-01 Métodos Numéricos para la Ciencia e Ingeniería

Tarea N°8: Métodos aleatorios: Cálculo del centro de masa de un sólido y obtención de una distribución aleatoria

Bruno Scheihing, RUT: 18.954.350-6

17 de Noviembre de 2015

Resumen

En la presente tarea se buscó calcular la posición del centro de masa de un sólido usando el algoritmo de Monte-Carlo 1. Adicionalmente, se buscó producir una distribución aleatoria a partir de una función normalizable $W(x)$, utilizando el algoritmo de Metropolis. Se calculó satisfactoriamente el centro de masa del sólido, obteniendo un resultado intuitivamente esperable, dadas las simetrías del problema. Se obtuvo una distribución aleatoria a partir de $W(x)$ que calza muy bien con la función original al hacer las normalizaciones adecuadas. Para el tamaño de la muestra usado ($N \approx 10^7$) se encontró que la incertidumbre asociada al histograma es, en la sección donde hay mayor densidad de datos, al menos dos órdenes de magnitud inferior que los tamaños característicos del histograma mismo, presentando así el algoritmo de Metropolis un muy buen funcionamiento.

1. Introducción

Los métodos aleatorios son una importante herramienta para hacer estimaciones de cálculos complejos mediante algoritmos relativamente simples. En la presente tarea se busca calcular la posición del centro de masa de un sólido generado por la intersección de dos regiones en el espacio:

$$\begin{aligned} z^2 + \left(\sqrt{x^2 + y^2} - 3 \right)^2 &\leq 1 \\ (x - 2)^2 + z^2 &\leq 1 \end{aligned}$$

que corresponden a un volumen toroidal y cilíndrico respectivamente. Este sólido posee una densidad dada por:

$$\rho(x, y, z) = 0.5 * (x^2 + y^2 + z^2)$$

Para hacer el cálculo se usará el método de Monte-Carlo 1. Se determinará el error asociado a este cálculo repitiendo el proceso (para semillas aleatorias diferentes); en vez de tomar el error dentro de cada simulación de Monte-Carlo, se realizará la simulación $N_{sim} = 100$ veces, y se tomará el error asociado como la desviación estándar de los resultados obtenidos sobre todas las simulaciones.

Por otra parte, en esta tarea también se buscará generar una distribución aleatoria con la función de distribución (no normalizada):

$$W(x) = 3.5 \times \exp\left(\frac{-(x-3)^2}{3}\right) + 2 \times \exp\left(\frac{-(x+1.5)^2}{0.5}\right)$$

a partir del algoritmo de Metropolis. Se usará una distribución proposición $x_P = x_n + \delta \cdot r$, con $\delta = 4$ y $r \sim U(0, 1)$, generando una muestra de puntos $\{x_n\}_{n=1, \dots, N}$ con $N = 10^7$, usando $x_0 = 0$.

Para determinar la incertidumbre asociada a la distribución generada, se repite el procedimiento anterior 100 veces (con semillas aleatorias diferentes), y se asocia un error a cada intervalo relevante del espacio (por definir en *Procedimiento*) a partir de la desviación estándar de la distribución en cada intervalo sobre las 100 repeticiones. Esto permitirá evaluar si hay garantías de que para el N elegido, el algoritmo de Metropolis genera de forma satisfactoria la distribución asociada a $W(x)$.

2. Procedimiento

2.1. Cálculo del centro de masa de un sólido

Se generarán números aleatorios dentro de un volumen que contiene la intersección de ambas regiones, y se guardarán los puntos (x, y, z) que caigan dentro de la región de interés. Desde este set de números $\{\vec{r}_i = (x_i, y_i, z_i)\}_{i=1, \dots, N_{datos}}$, se calcula la posición del centro de masa usando que:

$$\vec{r}_{CM} = \frac{\int \vec{r} dm}{\int dm} = \frac{\int_V \vec{r} \rho(x, y, z) dV}{\int_V \rho(x, y, z) dV} \approx \frac{\sum_{i=1}^{N_{datos}} \vec{r}_i \rho(x_i, y_i, z_i)}{\sum_{i=1}^{N_{datos}} \rho(x_i, y_i, z_i)}$$

donde hemos simplificado todas las constantes asociadas a cada suma, que aparecen tanto en el numerador como en el denominador desde la construcción del método de Monte-Carlo (el volumen de la región y el número de datos usados).

El volumen elegido para esto será una región rectangular, definida por:

$$(x, y, z) \in [0.9, 3.1] \times [-3.9, 3.9] \times [-1.1, 1.1]$$

que contiene completamente a la intersección de las dos regiones: vemos que ambas cumplen por construcción que $z^2 \leq 1$, por lo que $z \in [-1, 1]$. Por otra parte, el cilindro está centrado en $x = 2$, y su radio es 1, y así $x \in [1, 3]$. Finalmente, usando esto en la expresión del volumen toroidal podemos reducir el dominio de y al intervalo $[-3.9, 3.9]$, desde donde vemos que la región antes presentada contiene a todos los puntos posibles de interés.

Se reportará el error asociado a la estimación como la desviación estándar del resultado obtenido sobre 100 cálculos diferentes (con semillas aleatorias distintas entre sí). Cada cálculo será hecho a partir de $N_{datos} = 100000$ puntos (x, y, z) aleatorios en la región definida.

2.2. Obtención de distribución $W(x)$ usando el algoritmo de Metropolis

El algoritmo de Metropolis permite generar una distribución de probabilidad descrita por una función $W(x)$ (no necesariamente normalizada) a partir de

una posición x_0 . La implementación particular en esta tarea, para determinar el punto x_{n+1} a partir de x_n , consiste en:

- (a) Se genera un número aleatorio $r \sim U(-1, 1)$.
- (b) Dado un número δ (fijo), se propone $x_P = x_n + \delta \cdot r$.
- (c) Se genera otro número aleatorio $s \sim U(0, 1)$.
- (d) Se evalúa la proposición $W(x_P)/W(x_n) > s$. Si es verdadera, entonces se “acepta” y se asigna $x_{n+1} = x_P$. Si es falsa, entonces se “rechaza” y se asigna $x_{n+1} = x_n$.

El proceso recién descrito se repite $N = 10^7$ veces para generar una muestra con 10 millones de puntos más 1 (el inicial). Como se dijo en *Introducción*, se usa $\delta = 4$, pues experimentalmente se determinó que este valor permitía aceptar alrededor de la mitad de las proposiciones x_P (se dijo en clases que los buenos valores son los que permiten aceptar aproximadamente el 50 % de las proposiciones). Se imprimirá en pantalla cuántas proposiciones son aceptadas, y se presentará un gráfico con el histograma generado a partir de los $\{x_n\}_{n=0, \dots, 10^7}$ y la función de distribución deseada, ambos debidamente normalizados. No es difícil verificar que la normalización para $W(x)$ está dada por:

$$W_{norm}(x) = \frac{3.5 \times \exp\left(\frac{-(x-3)^2}{3}\right) + 2 \times \exp\left(\frac{-(x+1.5)^2}{0.5}\right)}{3.5\sqrt{3\pi} + 2\sqrt{0.5\pi}}$$

Para determinar la incertidumbre asociada al histograma generado, se repite el procedimiento anterior mismo 100 veces (con diferentes semillas aleatorias) y luego se calcula la desviación estándar para cada *bin* del histograma. Se usarán 100 *bins* del mismo “ancho” sobre el intervalo $[-5, 10]$. Se presentará un histograma, con barras de error asociadas, y la distribución deseada. El histograma será el promedio de los 100 histogramas generados (no hay motivos para elegir uno por sobre otro), pues el promedio de histogramas normalizados es equivalente a armar un solo gran histograma con todos ellos y normalizarlo.

En caso de que el gráfico anterior no sea fácil de interpretar, se agregará una tabla (o una lista de valores) con desviaciones estándar representativas de cada sección.

3. Resultados

3.1. Cálculo del centro de masa de un sólido

Después de realizar el procedimiento descrito para calcular el centro de masa del sólido, se obtuvo:

$$x_{CM} = 2.080 \pm 0.003$$

$$y_{CM} = 0.00 \pm 0.01$$

$$z_{CM} = 0.000 \pm 0.002$$

donde se muestra el resultado hasta la primera cifra significativa que presenta un error asociado. Mediante argumentos de simetría se podía determinar que la

posición del centro de masa en los ejes Y y Z es cero, lo que es consistente con el resultado obtenido.

Sin embargo, el cálculo en el eje X es analíticamente más complejo, debido no solamente a los límites de integración asociados al cálculo sobre el volumen de interés, sino además a una densidad variable con la posición, por lo que no es trivial efectuar comparaciones precisas.

3.2. Obtención de distribución $W(x)$ usando el algoritmo de Metropolis

Con el valor de δ utilizado, se aceptan 4998501 proposiciones de x_P , por lo que se concluye que es un buen valor, ya que corresponde al 49.985 % de las proposiciones.

En la Figura 1 se muestra la comparación del histograma obtenido usando el algoritmo de Metropolis contra la distribución deseada.

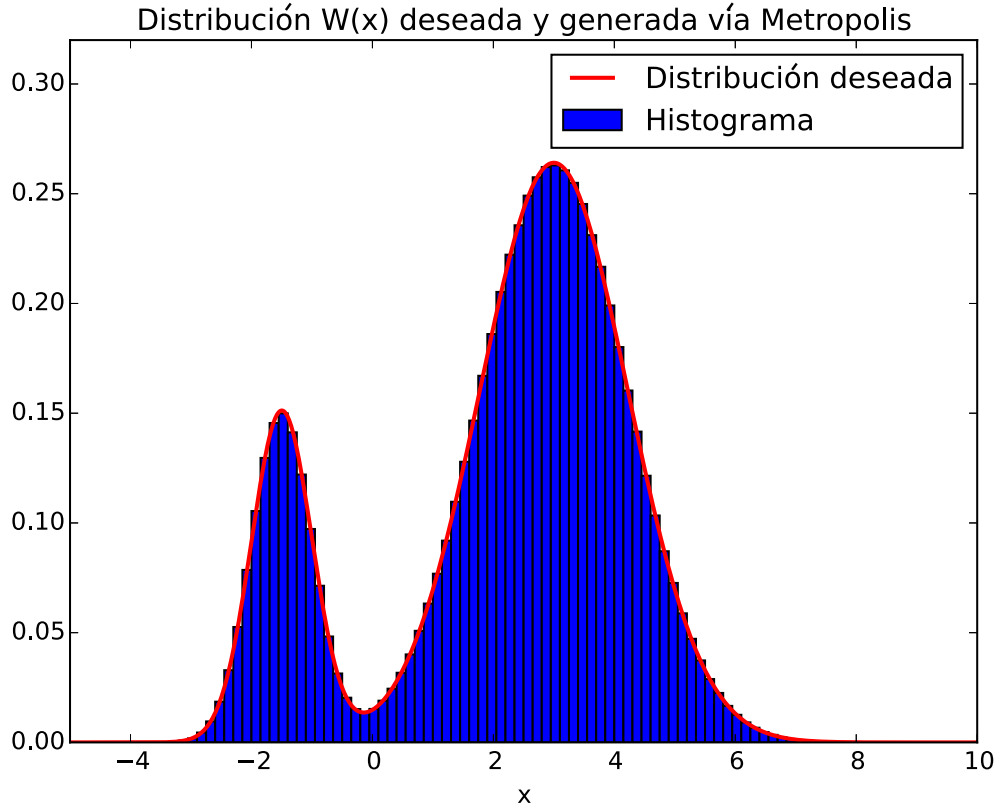


Figura 1: Histograma normalizado de la distribución generada usando el algoritmo de Metropolis (azul) y distribución deseada $W_{norm}(x)$ (rojo).

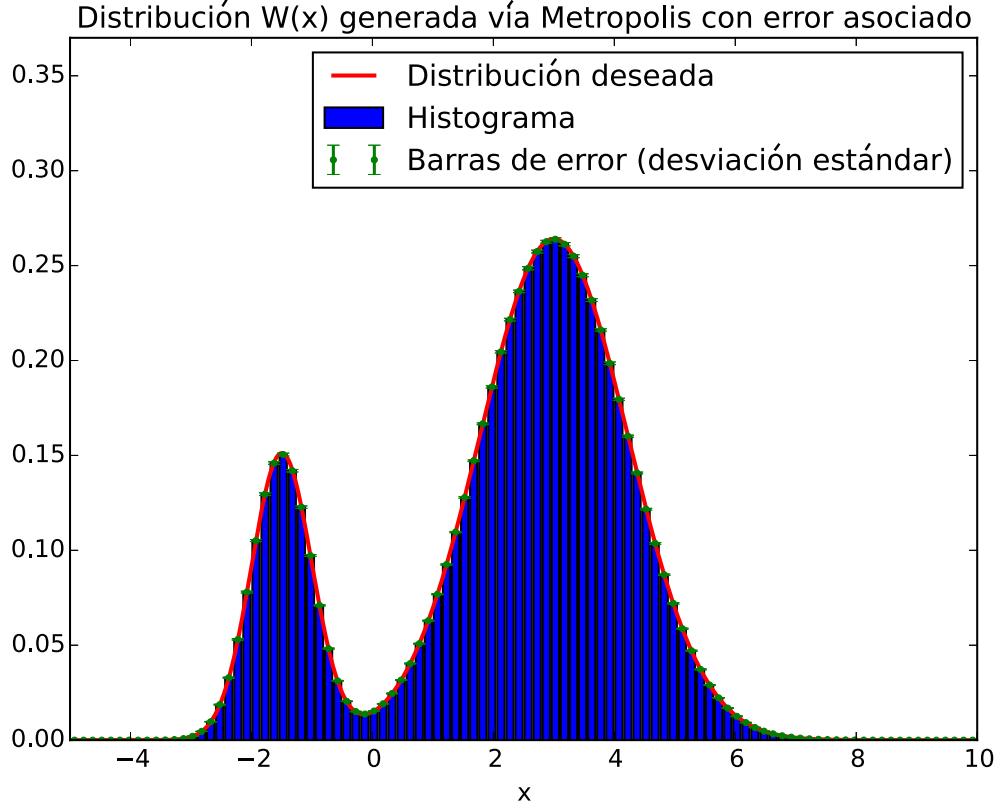


Figura 2: Histograma normalizado promedio de la distribución generada usando el algoritmo de Metropolis (azul), barras de error asociadas a la desviación estándar de los *bins* (verde), y distribución deseada $W_{norm}(x)$ (rojo).

Se observa un alto grado de coincidencia entre los bins generados y $W_{norm}(x)$, evidenciando el buen funcionamiento del algoritmo.

La Figura 2 muestra el resultado de ejecutar 100 veces el procedimiento descrito para obtener la Figura 1. El histograma presentado es el promedio de todos los histogramas generados. Se observa que la desviación estándar asociada a cada *bin* es muy pequeña relativa al tamaño característico del gráfico completo.

Al observar el Cuadro 1 junto con la Figura 2 se aprecia que mayores desviaciones estándar (i.e. incertidumbre mayor) ocurren cuando el valor mismo del histograma es más grande, es decir, donde la distribución concentra más valores. En general la incertidumbre calculada está más de dos órdenes de magnitud por debajo de los valores en el histograma, salvo por los valores de las “colas” donde ya la distribución es prácticamente nula y el error asociado resulta mayor.

En caso de querer obtener los datos directamente, en el archivo “histdesviacion.npy” se encuentran registradas las desviaciones estándar para cada uno de

Cuadro 1: Algunas desviaciones estándar de los *bins*

x	-1.925	-0.125	1.675	3.475	5.275	7.075
σ	$5.3 \cdot 10^{-4}$	$1.2 \cdot 10^{-4}$	$4.5 \cdot 10^{-4}$	$7.8 \cdot 10^{-4}$	$2.8 \cdot 10^{-4}$	$4.2 \cdot 10^{-5}$

DESVIACIONES ESTÁNDAR σ ASOCIADAS A LAS BARRAS DE ERROR DEL HISTOGRAMA EN FIGURA 2. SE MUESTRAN VALORES QUE SE ESPERA SEAN REPRESENTATIVOS DE SU VECINDAD EN EL EJE x . EL PUNTO USADO PARA REPRESENTAR CADA *bin* ES EL PUNTO MEDIO DEL INTERVALO QUE OCUPA.

los 100 *bins*, ordenadas según la posición de cada uno en orden creciente. (en el archivo “histpromedio.npy” se encuentran los valores promedio en la misma forma)

4. Conclusiones

La posición del centro de masa del sólido definido fue estimada satisfactoriamente, obteniendo una posición que es intuitivamente aceptable, dadas las simetrías del cuerpo y su distribución de densidad. Se esperaba un centro de masa en el cero de las coordenadas para (y, z) (por simetría), pero según x no quedaba claro qué efecto iba a dominar, aunque era esperable que $x \in (1.5, 3)$.

El error asociado al cálculo es mayor en la coordenada que tenía más valores para recorrer (y), lo que es esperable pues un intervalo más grande permite una mayor dispersión de los valores. Hay que notar que el método de Monte-Carlo no es particularmente preciso sino para un gran número de puntos, desde donde se explica que aún con 100000 puntos el error asociado no sea despreciable.

La distribución obtenida usando el algoritmo de Metropolis es también satisfactoria, pues se ajusta muy bien a la distribución deseada. Pareciera que se podrían haber usado más *bins* sin que apareciera una discrepancia de la distribución con $W_{norm}(x)$.

Al estudiar la incertidumbre asociada al histograma generado, se encuentra que el valor de N elegido es suficientemente grande como para que la distribución no presente variaciones significativas al cambiar la semilla, obteniendo una desviación estándar órdenes de magnitud inferior al tamaño de la sección central del histograma.

Se podría continuar este estudio variando N y comparando los distintos valores para la desviación estándar que se obtienen en cada caso, presumiblemente viendo de qué forma disminuye el error al aumentar el tamaño de la muestra.