Paseos aleatorios sobre grafos

Antonio Sevilla Sastre

26 de abril de 2023

Epítome

Esta es la memoria de un trabajo final para la asignatura Taller de Tecnomatemática, impartida por David Gómez Castro en la Universidad Complutense de Madrid. Se proponen algoritmos para la simulación de caminatas aleatorias sobre distintos grafos infinitos. Posteriormente se procede al estudio de la recurrencia.

Abstract

This is the report of a final project for the subject Technomathematics Workshop taught by David Gómez Castro at the Complutense University of Madrid. Algorithms for the simulation of random walks on different infinite graphs are proposed. After that, it is proposed the study of its recurrence.

Motivación

El estudio de los paseos aleatorios sobre mallados tiene sus raíces en la teoría de grafos y la física estadística. En la década de 1950, los matemáticos comenzaron a interesarse por el análisis de la dinámica de partículas en redes regulares y aleatorias, lo que dio lugar al desarrollo de la teoría de los procesos de Markov y los modelos de caminatas aleatorias.

En la década de los 60, el estudio de los paseos aleatorios se expandió a una variedad de aplicaciones en física, química, biología y ciencias sociales, y los mallados comenzaron a ser utilizados como modelos para la descripción de sistemas complejos. En particular, los paseos aleatorios sobre mallados se convirtieron en una herramienta importante para el estudio de fenómenos como la difusión, la propagación de epidemias y la percolación.

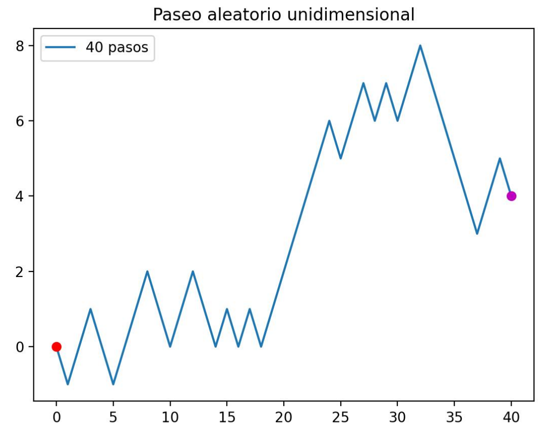
En las últimas décadas, el estudio de los paseos aleatorios sobre mallados ha sido objeto de intensa investigación, y se han desarrollado técnicas cada vez más sofisticadas para analizar su comportamiento y propiedades estadísticas. Hoy en día, esta área de investigación continúa siendo muy importante para matemáticos, físicos, biólogos y otros científicos que buscan entender y modelar sistemas complejos en la naturaleza y la sociedad.

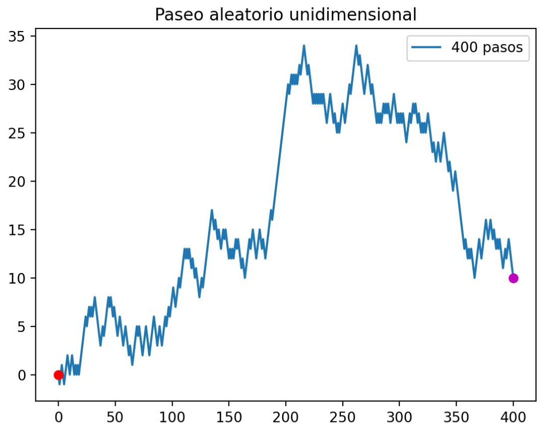
Este trabajo sigue la línea de lo estudiado durante la asignatura, la cual introduce ciertas herramientas para la simulación y el estudio de procesos estocásticos.

1. Algoritmos para la simulación y representación de paseos aleatorios

1.1. Caminata unidimensional

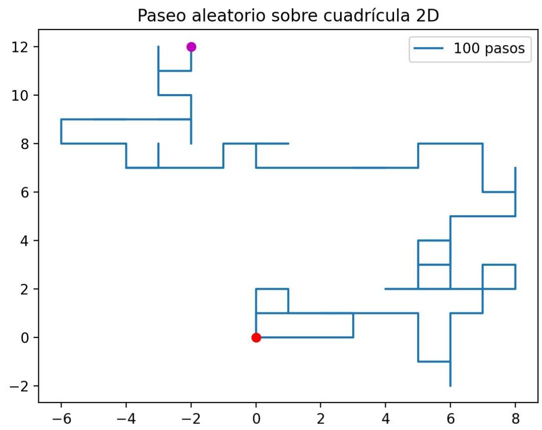
Supongamos que un hombre borracho sale de un bar y empieza a dar tumbos equiprovablemente hacia su izquierda y hacia su derecha. La caminata que describe sobre un eje es equivalente a la que siguen las bolas de la máquina de Galton, la cual hemos estudiado en clase. Como vimos, al aumentar el número de pasos el recorrido se asemeja al movimiento browniano unidimensional.

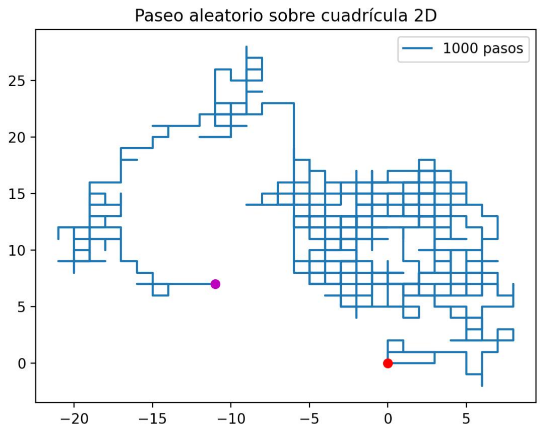


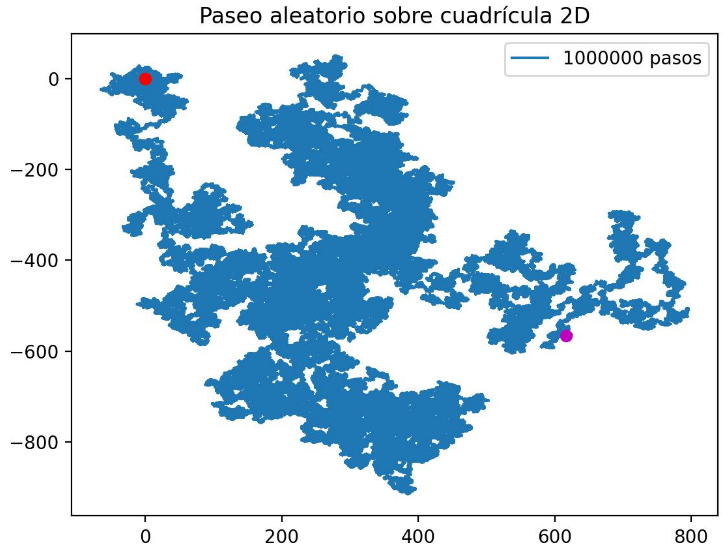


1.2. Caminata sobre cuadrícula bidimensional

Supongamos ahora que nuestro hombre borracho intenta llegar a su casa andando. Está en Barcelona (o en Manhattan si se quiere), las calles conforman una cuadrícula. Cuando llega al final de una manzana decide aleatoria y equiprobablemente si girar a la derecha, a la izquierda, seguir de frente o regresar por donde ha venido.Se adjunta una representación de una simulación de la caminata.



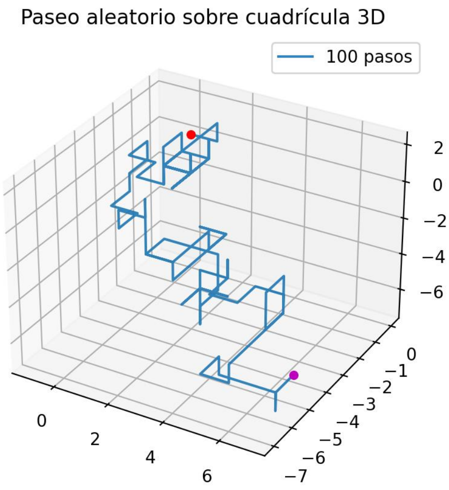


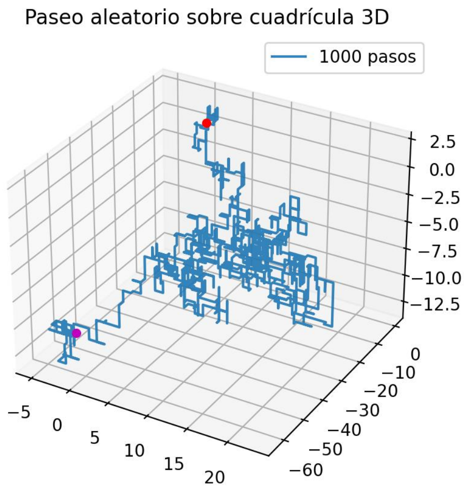


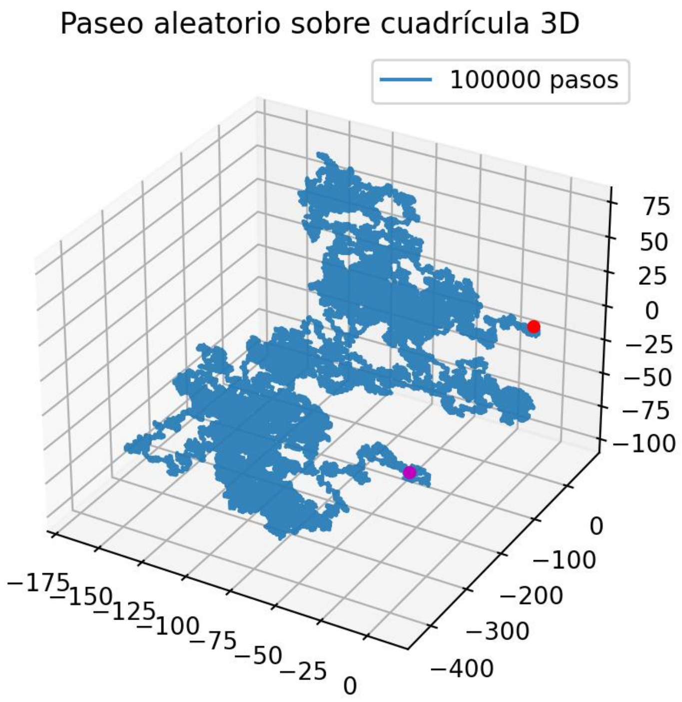
Las representaciones de caminatas con muchos pasos, al ocupar tan poco cada paso, recuerdan al movimiento browniano bidimensional. La reproducción de ciertos hongos, por ejemplo, sigue distribuciones similares.

1.3. Caminata sobre cuadrícula tridimensional

Consideramos ahora una abeja borracha (de néctar quizá) que emprende el vuelo. Cada cierto tiempo decide aleatoriamente y de forma equiprobable entre girar en cada una de las 4 direcciones cardinales, ascender o descender. Se adjunta una representación de la simulación de dicho vuelo.







De nuevo, si se realizan muchos pasos, los resultados recuerdan al movimiento browniano de 3 dimensiones en una escala grande. La última gráfica se parece a la forma de un hormiguero, por ejemplo.

1.4. Caminata sobre cuadrícula m-dimensional

Supongamos ahora un ser que se puede desplazar entre m dimensiones y, por supuesto, está borracho. Toma sus decisiones de movimiento similarmente al hombre o a la abeja pero sobre una cuadrícula m-dimensional. Su caminata se puede simular para cualquier m. Se adjunta el código.

```
def gen_direccion(m):
    """
    este metodo genera una direccion y un sentido para
    la caminata sobre cuadricula m-dimensional
    """
    direccion = random.randint(0, m-1)
    sentido = random.choice([1, -1])
    return (direccion, sentido)
```

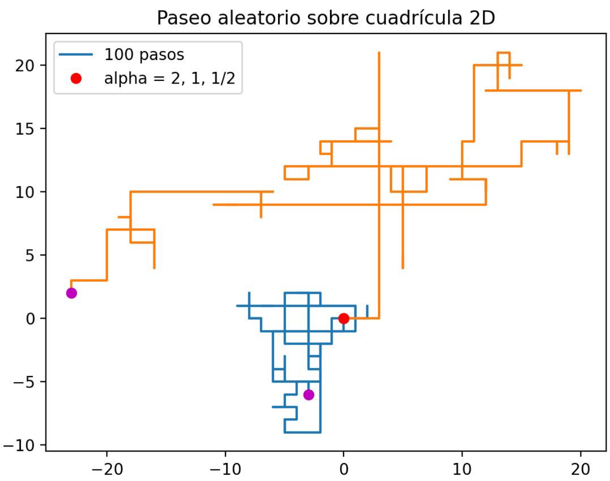
```
def paseo_md(n, m):
    """
    este metodo realiza una simulacion de caminata aleatoria
    sobre una cuadricula m-dimensional de n pasos
    la salida es una matriz de m vectores de las posiciones
    (incluyendo la inicial) en cada coordenada
    """
    matrix = []
    for p in range(m):
        matrix.append(np.zeros(n+1))
    for paso in range(1, n+1):
        (direccion, sentido) = gen_direccion(m)
        for h in range(m):
            matrix[h][paso] = matrix[h][paso-1]
        matrix[direccion][paso] += sentido
    return matrix
```

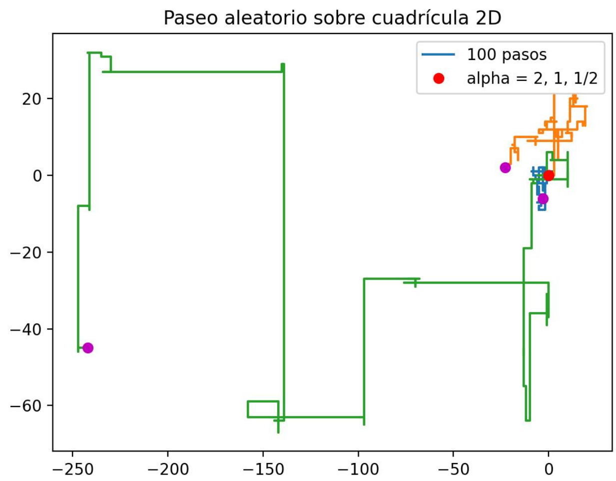
Para representar la caminata sobre una cuadrícula m-dimensional se puede emplear cualquier conjunto de r representaciones del estilo de las anteriores, cada una de ellas con dimensión $1 \leq dim_i \leq 3$ de manera que $\sum_1^r dim_i = m$.

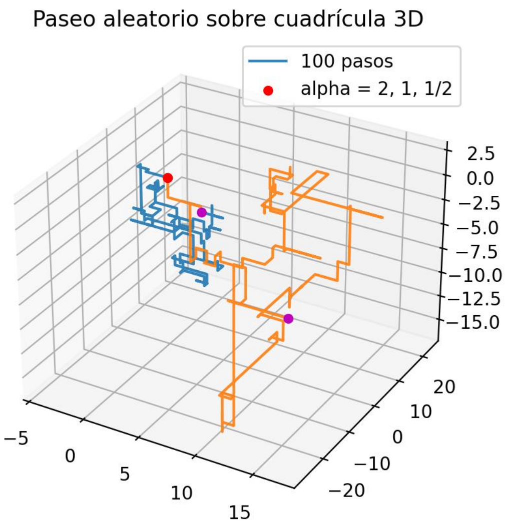
Por ejemplo, la caminata 4-dimensional sobre una cuadrícula puede representarse con un gráfico tridimensional `paseo_3d_plot` y uno de la posición frente al tiempo `paseo_1d_plot;` o también como dos gráficos bidimensionales como `paseo_2d_plot`. Se adjunta en src el código pertinente para dicha representación, `PaseomD.py`. No genera imágenes ya que no se incluyen en el documento, pero las plotea.

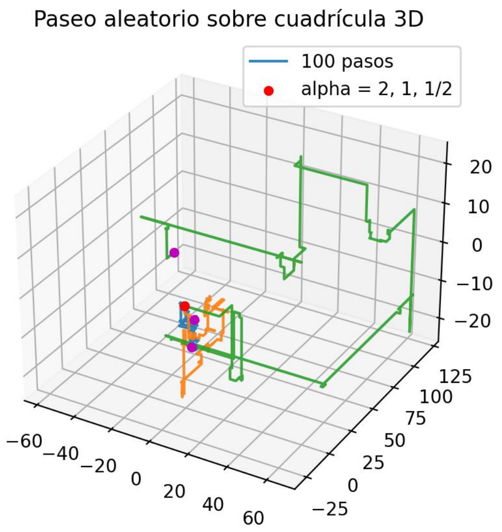
1.5. Caminata sobre cuadrícula m-dimensional con saltos

Utilizamos los anteriores algoritmos y lo estudiado en la práctica 3 para simular los salto siguiendo la probabilida respecto a su longitud una ley de potencias.





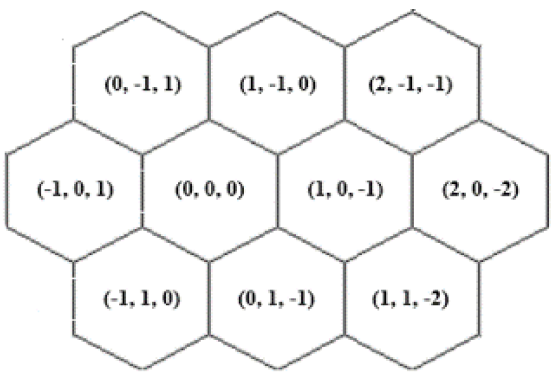


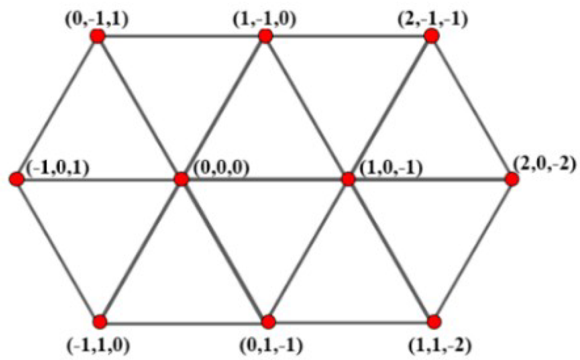


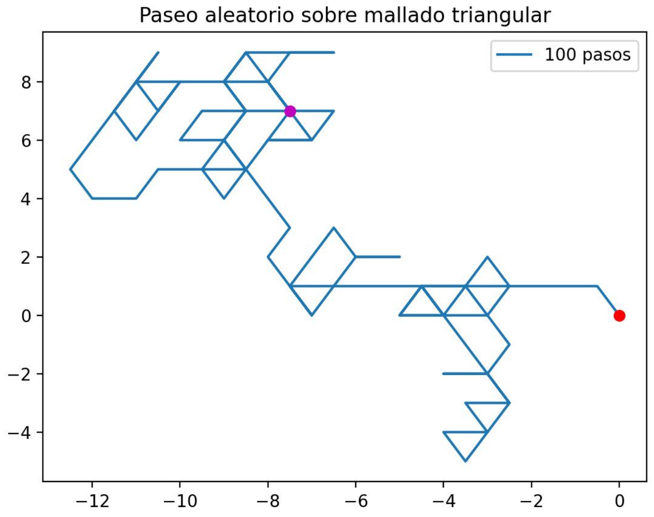
1.6. Caminata sobre mallado triangular

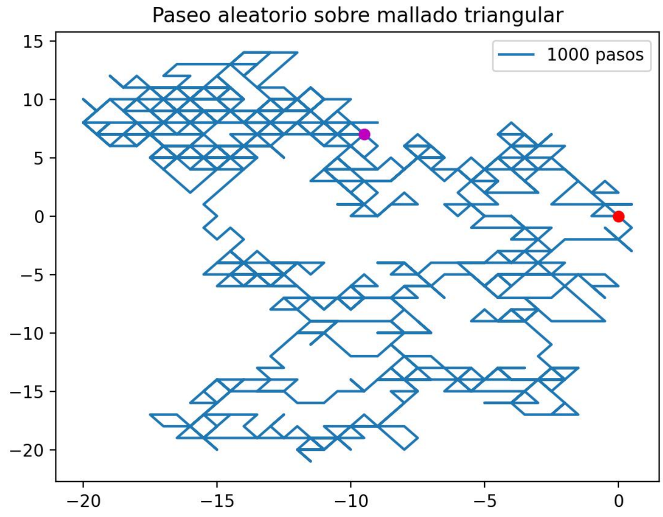
Nuestra abeja borracha se ha posado sobre su panal. Empieza a moverse aleatoriamente entre las celdas del mismo, eligiendo cada cierto tiempo una de las 6 celdas contiguas a la que está, todas con igual probabilidad.

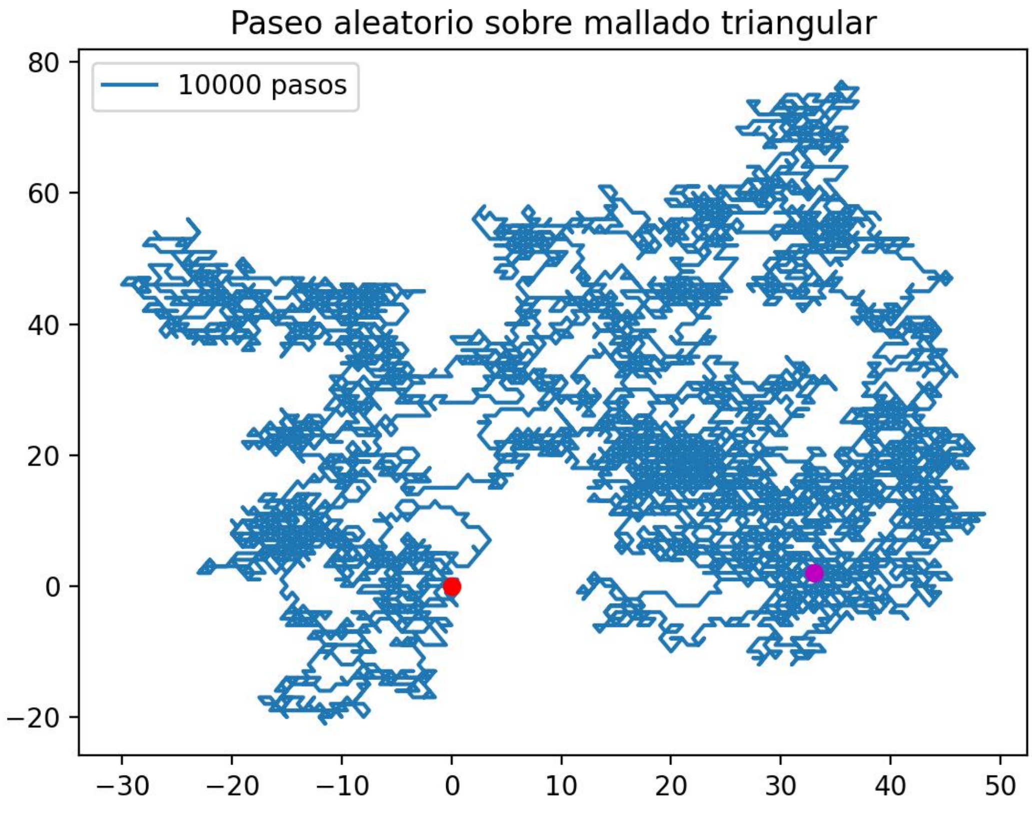
Al desplazarse entre hexágonos sobre un plano recubierto por ellos, el grafo por el que se mueve es un mallado triangular. Asumiremos los hexágonos regulares y consideraremos el sistema de coordenadas hexagonales propuesto en [1] para realizar la simulación.






Paseo aleatorio sobre mallado triangular
100 pasos

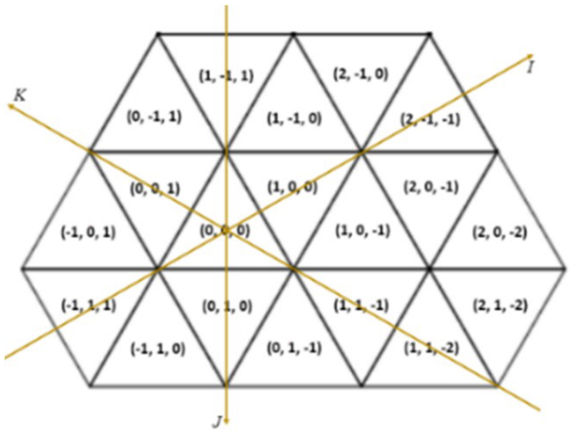

Paseo aleatorio sobre mallado triangular
1000 pasos

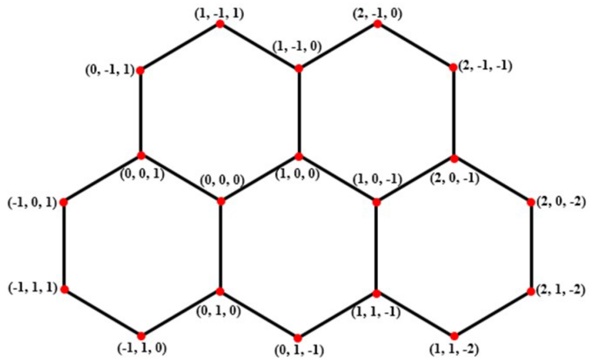

Paseo aleatorio sobre mallado triangular
10000 pasos

1.7. Caminata sobre mallado hexagonal

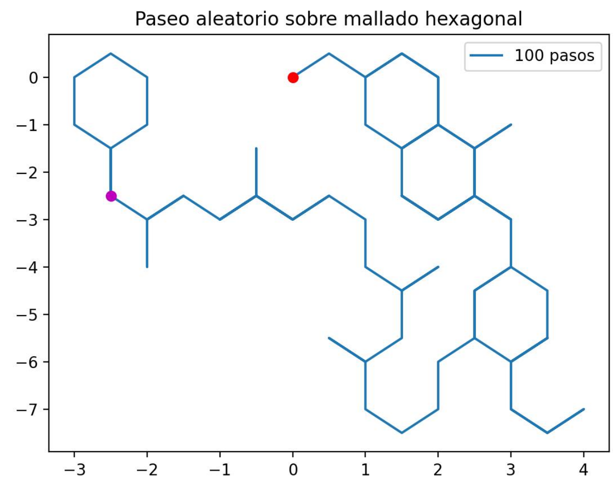
La abeja borracha ha decidido desplazarse ahora sobre los bordes de las celdas hexagonales. Lo hace también aleatoriamente, eligiendo cuando llega a una esquina uno de los 3 vértices contiguos, todos con igual probabilidad.

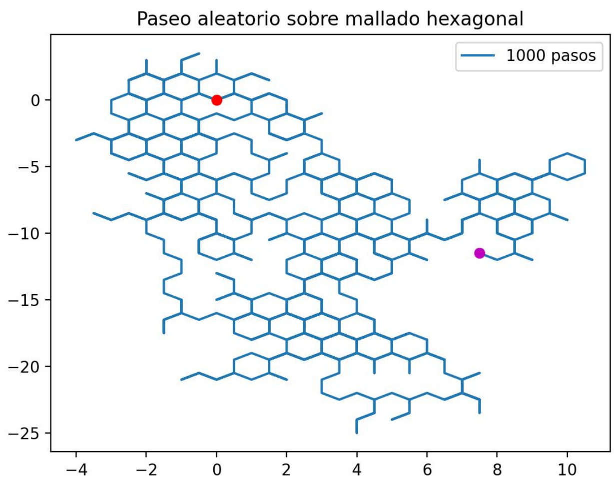
Desplazarse sobre los vértices de un hexágono, es decir sobre un mallado hexagonal, es equivalente a hacerlo entre celdas triangulares. En este sentido, se dice que el triángulo y el haxágono son *duales*, tal y como se expone en [1]. El cuadrado, de hecho, es *autodual*. Por ello, el hombre borracho hubiera descrito un recorrido de aspecto similar si en lugar de recorrer las calles andando hubiera ido saltando de azotea en azotea ortogonalmente.

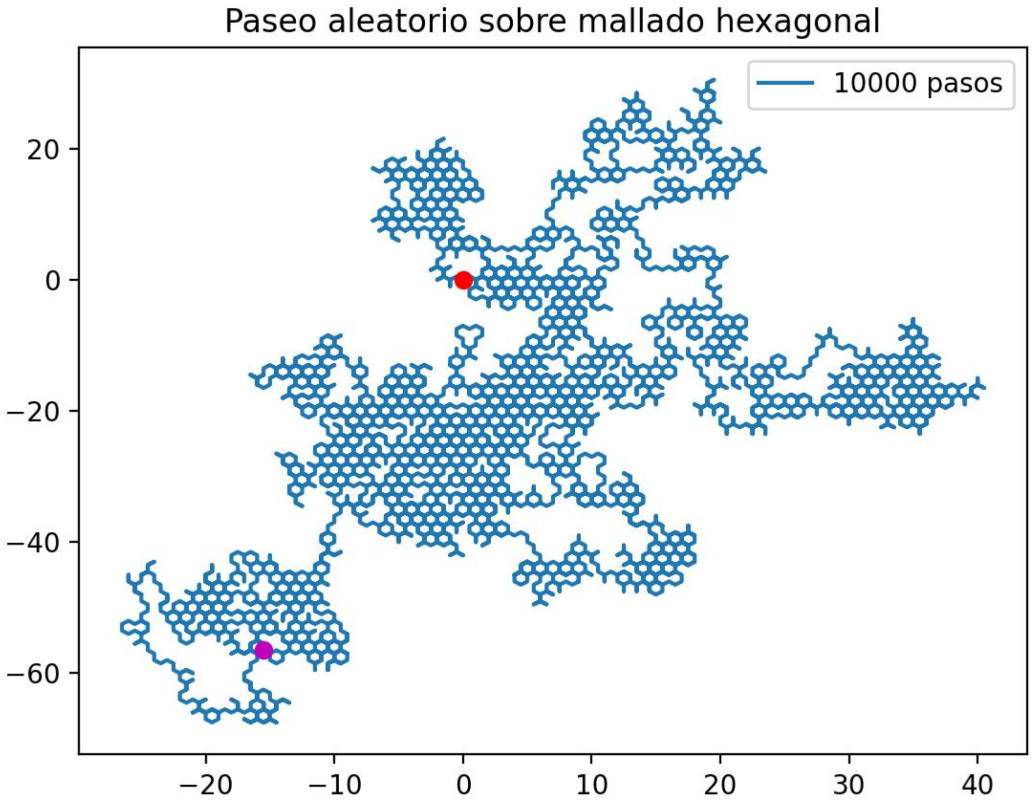




Asumiremos los hexágonos regulares y consideraremos el sistema de coordenadas triangulares que se sugiere en dicho artículo para realizar la simulación.







2. Estudio de la recurrencia

Una vez se han desarrollado procedimientos para simular y representar todos estos paseos aleatorios, una pregunta que parece bastante lógico realizarse es si se producirá eventualmente un regreso al origen.

Sin embargo, resulta inviable abordar esta cuestión computacionalmente. Cuando decimos *eventualmente* significa *transcurrida una cantidad infinita de tiempo*. Mediante un ordenador podríamos simular, a lo sumo, una cantidad lo suficientemente grande de tiempo como para hacernos una idea de la frecuencia con la que dicho regreso sucede. Por supuesto, esto no responde formalmente a nuestra pregunta.

A continuación se procede al estudio formal de los paseos aleatorios vistos como cadenas de Márkov.

2.1. Las caminatas aleatorias como cadenas de Márkov

Las cadenas de Márkov son un tipo especial de proceso estocástico discreto en el que la probabilidad de que ocurra un evento depende solamente del evento inmediatamente anterior.

Las caminatas aleatorias sobre grafos son por definición una cadena de Márkov. Se dice que estas cadenas constan de estado discretos. En nuestro caso, los estados son los puntos sobre el mallado elegido.

2.2. Caracterización de la recurrencia de una cadena de Márkov

La recurrencia es una propiedad sobre un estado de la cadena, es decir sobre un punto del mallado. El estudio de la recurrencia en el origen es equivalente a la pregunta que nos realizamos al principio de la sección. De hecho, se puede demostrar que la pérdida de memoria de las cadenas de Márkov garantiza que todos los estados (puntos) son recurrentes si y solo si lo es el estado inicial (origen).

Un punto de la cadena es recurrente si y solo si la probabilidad de volver es 1. Si la probabilidad es menor, entonces se dice que es transitorio. Para calcular dicha probabilidad vamos a estudiar la esperanza de la variable *número de regresos*, a la que llamaremos V.

Si un punto es recurrente, se puede aplicar iterativamente dicha propiedad junto con la pérdida de memoria de las cadenas de Márkov, obteniendo que eventualmente se regresa a él infinitas veces. Por tanto, $P(V \geq 1) = 1 \Rightarrow P(V = \infty) = 1$.

Por otro lado, $E[V]$ se puede descomponer como

$$\sum_{i=1}^{\infty} i \cdot P(V = i) = \sum_{i=1}^{\infty} P(V \geq i)$$

Si llamamos r a $P(V \geq 1)$, al no tener memoria el proceso $P(V \geq 2) = r^2$ y en general $P(V \geq k) = r^k$. Con esto, $E[V] = \sum_i^{\infty} r^i$. Esta es la serie geométrica, que converge a $\frac{r}{1-r}$ si $r < 1$. Un punto es transitorio si $r < 1$ y recurrente si $r = 1$.

Así pues, se tiene que

$$\text{Punto recurrente} \Rightarrow P(V = \infty) = 1 \Rightarrow E[V] = \infty$$

$$\text{Punto transitorio} \Rightarrow P(V = \infty) < 1 \Rightarrow E[V] < \infty$$

Al ser la recurrencia y la transitoriedad dicotómicas, se tiene que

$$\text{Punto recurrente} \Leftrightarrow P(V = \infty) = 1 \Leftrightarrow E[V] = \infty$$

$$\text{Punto transitorio} \Leftrightarrow P(V = \infty) < 1 \Leftrightarrow E[V] < \infty$$

para cualquier cadena de Márkov [2], [3].

2.3. Recurrencia de la caminata sobre cuadrícula unidimensional

Para calcular explícitamente $E[V]$, vamos a expresar primeramente V como la suma $\sum_{n=1}^{\infty} A_n$, siendo $A_n = 1$ si se ha producido un regreso en el paso n y $A_n = 0$ en caso contrario. Como la esperanza es un operador lineal, $E[V] = \sum_{n=1}^{\infty} E[A_n]$.

La probabilidad de que un paseo unidimensional de longitud n concreto suceda es trivialmente $\frac{1}{2}^n$, pues en cada paso solo se puede elegir entre 2 direcciones. Para que se produzca un regreso en un paseo aleatorio unidimensional, se tienen que haber producido n desplazamientos a la derecha y n a la izquierda. Esto garantiza que cualquier paseo con regreso tiene longitud par. Además, fijada cierta longitud $2n$, hay $\frac{2n!}{n!n!}$ de ellos, pues son las permutaciones de $2n$ elementos con 2 grupos de n elementos idénticos entre ellos.

Aplicando el toerema de la probabilidad total, podemos asegurar que la probabilidad de regresar tras $2n$ pasos es exactamente

$$\frac{1}{2}^{2n} \cdot \frac{2n!}{n!n!}$$

y que por tanto

$$E[V] = \sum_{n=1}^{\infty} \frac{1}{2}^{2n} \cdot \frac{2n!}{n!n!}$$

Se puede comprobar que dicha serie se comporta como $\frac{1}{\sqrt{n}}$ [4], con lo cual $E[V]$ diverge.

Esto significa que el hombre borracho que da tumbos hacia los lados eventualmente volverá a su ubicación original.

2.4. Recurrencia de la caminata sobre cuadrícula bidimensional

Procederemos de forma similar al caso de la caminata unidimensional. La probabilidad de que un camino concreto suceda es de $\frac{1}{4}^{n}$, al considerar las 4 direcciones de la cuadrícula. Para regresar se debe haber girado tantas veces hacia el norte como hacia el sur y tantas veces hacia el este como al oeste. Eso garantiza nuevamente que todo camino de vuelta tiene longitud par. Notando como i el número de giros hacia el norte, para un camino de longitud $2n$ el número de giros hacia el sur es también i y el de giros hacia el este y el oeste es $(n-i)$ y $(n-i)$. Fijado i, el número de caminos de regreso de longitud $2n$ es

$$\frac{2n!}{n! \cdot n! \cdot (n-i)! \cdot (n-i)!}$$

Así, el número de paseos posibles es

$$\sum_{i=0}^{n} \frac{2n!}{n! \cdot n! \cdot (n-i)! \cdot (n-i)!}$$

De nuevo, aplicando el teorema de la probabilidad total se tiene que la probabilidad de regresar tras $2n$ pasos es

$$\frac{1}{4}^{2n} \cdot \sum_{i=0}^{n} \frac{2n!}{n! \cdot n! \cdot (n-i)! \cdot (n-i)!}$$

y que por tanto

$$E[V] = \sum_{n=1}^{\infty} \frac{1}{4}^{2n} \cdot \sum_{i=0}^{n} \frac{2n!}{n! \cdot n! \cdot (n-i)! \cdot (n-i)!}$$

Tras una pesada manipulación descrita detalladamente en [4], se puede comprobar que dicha serie crece como $\frac{1}{n}$, serie conocidamente divergente.

Por suerte, podemos estar seguros de que en algún nuestro hombre borracho regresará a su casa. Quizá no sea tan alentador saber que eventualmente regresará también al bar desde el que emprendió el paseo.

2.5. Recurrencia de la caminata sobre cuadrícula tridimensional

Se puede aplicar un razonamiento muy similar para este caso. Considerábamos que solo se puede realizar un movimiento en 6 direcciones. De nuevo, el número de movimientos en cada dirección de los ejes x, y, z debe ser igual al número de movimientos en la opuesta para que el camino sea cerrado sobre el origen. Asumiendo que realiza, por ejemplo, i movimientos hacia arriba y j hacia el norte esto resulta en un número esperado de retornos

$$E[V] = \sum_{n=1}^{\infty} \frac{1}{6}^{2n} \cdot \sum_{i,j=0}^{n} \frac{2n!}{i! \cdot i! \cdot j! \cdot j! \cdot (n-i-j)! \cdot (n-i-j)!}$$

Tras una ingeniosa manipulación descrita detalladamente en [4], empleando métodos de [5] podemos comprobar que dicha serie crece como $\frac{1}{n^{\frac{2}{3}}}$; conocidamente convergente.

Esto seignifica que es posible que la abeja borracha nunca regrese a casa. En [6] se calcula la probabilidad exacta de que lo haga, que es de entorno a $\frac{1}{3}$.

2.6. Recurrencia de la caminata sobre cuadrícula m-dimensional

Mediante el mismo proceso se puede estudiar el comportamiento del ser que se desplaza en 4 dimensiones, solo añadiendo que realiza k desplazamientos en *kata* y en *ana*. Se obtiene

$$E[V] = \sum_{n=1}^{\infty} \frac{1}{8}^{2n} \cdot \sum_{i,j,k=0}^{n} \frac{2n!}{i! \cdot i! \cdot j! \cdot j! \cdot k! \cdot k! \cdot (n-i-j-k)! \cdot (n-i-j-k)!}$$

cuyo comportamiento resulta convergente.

George Pólya demostró [2] que, en general, todos los puntos del paseo sobre una cuadrícula m-dimensional son recurrentes si y solo si $m \leq 2$.

Una aproximación intuitiva a por qué sucede esto se da en [7]. Cuanto más grande es m más decisiones de movimiento se pueden tomar. De alguna manera hay más espacio donde perderse y eso reduce las probabilidades de regresar al origen.

2.7. Recurrencia de las caminatas planas

Como hemos visto, si conseguimos obtener la esperanza de V, podemos conocer la recurrencia de un mallado. Sin embargo, para las caminatas sobre grafos triangulares o hexagonales no resulta tan fácil establecer el número de caminos posibles de regreso mediante una razón sencilla como pasaba en las cuadrículas. El camino de regreso ni siquiera tiene por qué constar de un número par de pasos. Solo para hacernos una idea, en este artículo [8] Yurii Lahodiuk se pregunta cuántos caminos con retorno posibles de longitud N se pueden realizar sobre el contorno de un solo hexágono.

Sin embargo, el teorema 4.11 de [3] asegura que todo grafo 2-dimensional es recurrente. Mientras que si emprende el vuelo puede que se pierda para siempre, nuestra abeja borracha regresará a su celda del panal infinito si camina sobre él.

2.8. Recurrencia de las caminatas con saltos

Por otro lado, los paseos aleatorios con saltos no son paseos aleatorios simples, con lo que ninguno de los resultados anteriores se puede aplicar para determinar la recurrencia de los mismos. Intuitivamente, un paseo aleatorio con saltos más cortos (distribuidos por ejemplo como una serie de potencias de parámetro menor o una función con orden de crecimiento menor) parece más propenso a ser recurrente.

Referencias

[1] Benedek Nagy and Khaled Abuhmaidan. A continuous coordinate system for the plane by triangular symmetry. *Symmetry*, 11(2), 2019.

[2] Jonathan Novak. Polya's random walk theorem. *The American Mathematical Monthly*, 121(8):711–716, 2014.

[3] Gregory F Lawler and Vlada Limic. *Random walk: a modern introduction*, volume 123. Cambridge University Press, 2010.

[4] Michael Kozdron. An introduction to random walks from polya to self-avoidance. Technical report, Technical report, Duke University, 1998.

[5] Albert R Meyer and Ronitt Rubinfeld. Generating functions. *Mathematics for Computer Science, Massachusetts Institute of Technology*, pages 1–13, 2005.

[6] Steven R Finch. *Mathematical constants.* Cambridge university press, 2003.

[7] Mathemaniac. *Random walks in 2D and 3D are fundamentally different (Markov chains approach).* 2022. URL: https://youtu.be/iH2kATv49rc.

[8] Yurii Lahodiuk. *http://lagodiuk.github.io/computer_science/2017/04/17/random_walks_in_a_hexagon.html*, 2017.