Proyecto Probabilidad

Juan David Duarte Yara

December 2023

1. Cadenas de Markov.

.

Las cadenas de Markov, nombradas en honor al matemático ruso Andréi Markov, son casos particulares de procesos estocásticos que describen la evolución de un sistema a lo largo del tiempo, donde la probabilidad de que el sistema se encuentre en un estado particular depende únicamente del estado actual y no de cómo se llegó a ese estado. Este enfoque hace que las cadenas de Markov sean especialmente útiles para modelar fenómenos dinámicos en los cuales la historia pasada no tiene impacto directo en el futuro, dadas las condiciones actuales.

Una cadena de Markov se compone de un conjunto finito de estados y las probabilidades de transición entre ellos. Estas cadenas de Markov tienen la propiedad de "perdida de memoria"que se demostrara mas adelante en la ecuacion (2), lo que implica que la probabilidad de transición a un nuevo estado depende únicamente del estado actual, lo que hace aplicables las cadenas de Markov en diversas disciplinas, como estadística, teoría de la información, economía, biología, y etc...

La versatilidad de las cadenas de Markov radica en su capacidad para modelar sistemas dinámicos en los que la incertidumbre y la variabilidad desempeñan un papel crucial. Su utilidad se extiende desde la predicción de comportamientos futuros hasta la resolución de problemas en los cuales la aleatoriedad y la dependencia de estados son factores esenciales. Por lo tanto las cadenas de Markov son herramientas fundamentales en el análisis de procesos estocásticos y ofrecen una base conceptual para comprender y abordar una amplia gama de fenómenos en ciencias aplicadas y teóricas.

2. Primera parte.

1. Escriba un programa que simule una cadena de Markov a partir del vector $\vec{a}$ y la matriz de transición P.

Vease la carpeta llamada "primera parte.[en] el siguiente repositorio de git-hub, especificamente el documento llamado "main.py".
`https://github.com/JuanDDY/Proyecto_probabilidad_2023-2.git`

En el codigo se implementan las funciones f y g. Se define X_0 usando la funcion g y las X_{n+1} para $n \geq 0$ usando la funcion f.
Al final se imprimen las X_i ordenadas en una lista, para ver como evoluciona la cadena de Markov a lo largo de los tiempos.

I. Defina la función $g(u) = \sum_{i=1}^{|S|} i\mathbf{1}_{(\sum_{k=1}^{i-1} a_k, \sum_{k=1}^{i} a_k]}(u)$.

II. Defina $X_0 = g(U_0)$.

III. Defina la función $f(i,u) = \sum_{i=1}^{|S|} j\mathbf{1}_{(\sum_{k=1}^{j-1} p_{ik}, \sum_{k=1}^{j} p_{ik}]}(u)$.

IV. Para $n \geq 0$, defina $X_{n+1} = f(X_n, U_{n+1})$.

Figura 1: Funciones del codigo

Se definen las constantes "SIZE"que es el tamaño del espacio S y se define "TAMANIO_SUCESION"que es el numero de tiempos que se quiere que la cadena de Markov avance. Estas constantes se invitan a que sean cambiadas para ver cono funciona esta cadena de Markov.

Cabe aclarar que el vector de estados iniciales a se genera automaticamente con una funcion $Uniforme(0,1)$ y luego se normaliza y se hace una correccion para corregir la precision finita de la maquina. Lo mismo se hace con la Matriz P que se genera de manera automatica. Se realiza el mismo proceso con cada fila de la matriz.

Si se quiere insertar un vector para probar, se debe reemplazar el vector A que es una constante para el codigo Y lo mismo para la matriz P la que tambien se puede reemplazar si se quiere.

2.1. Propiedades de las cadenas de Markov:

2. Muestre que una cadena de Markov satisface que:

$$\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i_n) = a_{i_0} p_{i_0 i_1} \ldots p_{i_{n-1} i_n} \tag{1}$$

Demostración. Se revisaran dos casos.

Caso 1: Consideremos el caso en que para algun $k \in \{0, \ldots, n\}$ $p_{i_k i_{k+1}} = 0$. Es decir $\mathbb{P}(X_{k+1} = i_{k+1} | X_k = i_k) = 0$. En otras palabras, como nuestros estados son finitos, es imposible pasar del estado i_k al estado i_{k+1}. Y la probabilidad de que los estados i_k y i_{k+1} ocurran de manera consecutiva es 0. Por tanto $\mathbb{P}(X_j = i_j, X_{j+1} = i_{j+1}) = 0$. En particular:

$$\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i_n) = 0$$

Y se cumple la ecuacion (1) para cuando $p_{i_k i_{k+1}} = 0$.

Caso 2: Consideremos el caso $p_{i_k i_{k+1}} > 0$ para todo $k \in \{0, \ldots, n\}$. Y procederemos por induccion sobre n.

Caso base: Tomemos a $n = 1$

$$\begin{aligned} \mathbb{P}(X_1 = i_1, X_0 = i_0) &= \mathbb{P}(X_1 = i_1 | X_0 = i_0) \mathbb{P}(X_0 = i_0) \\ &= p_{i_0 i_1} a_{i_0} \end{aligned}$$

Por lo tanto la ecuacion (1) se cumple para cuando n es igual a 1

Paso inductivo: Suponga que la ecuacion (1) vale para n, es decir que:

$$\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i_n) = a_{i_0} p_{i_0 i_1} \ldots p_{i_{n-1} i_n}$$

Entonces calculemos

$$\begin{aligned}
\mathbb{P}(X_{n+1} = i_{n+1}, X_n = i_n, \ldots, X_0 = i_0) &= \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n, \ldots X_0 = i_0)\mathbb{P}(X_n = i_n, \ldots X_0 = i_0) \\
&= \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n)\mathbb{P}(X_n = i_n, \ldots X_0 = i_0) \\
&= p_{i_n i_{n+1}}(a_{i_0} p_{i_0 i_1} \ldots p_{i_{n-1} i_n}) \\
&= a_{i_0}\, p_{i_0 i_1} \ \ldots \ p_{i_{n-1} i_n}\, p_{i_n i_{n+1}}
\end{aligned}$$

Entonces se cumple en el paso inductivo, por tanto la ecuacion (1) vale para toda $n \in \mathbb{Z}^+$

□

3. Muestre que:

$$\mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m | X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) = \mathbb{P}(X_1 = j_1, \ldots, X_m = j_m | X_0 = i) \tag{2}$$

Siempre que $\mathbb{P}(X_0 = j_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) > 0$

Demostración. Veamos que cuando tenemos la expresion $\mathbb{P}(X_1 = j_1, \ldots, X_m = j_m | X_0 = i)$ se supone que con seguridad $X_0 = i$. Por lo tanto definimos $a'_i := \mathbb{P}(X_0 = i) = 1$ bajo esta suposicion. Entonces calculamos:

$$\begin{aligned}
\mathbb{P}(X_1 = j_1, \ldots, X_m = j_m | X_0 = i) &= \frac{\mathbb{P}(X_0 = i, X_1 = j_1, \ldots, X_m = j_m)}{\mathbb{P}(X_0 = i)} \\
&= \frac{a'_i\, p_{i j_1} \ldots p_{j_{m-1} j_m}}{a'_i} \\
&= p_{i j_1} \ldots p_{j_{m-1} j_m} \\
&= p_{i j_1} \ldots p_{j_{m-1} j_m} \frac{\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i)}{\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i)} \\
&= \frac{a_{i_0}\, p_{i_0 i_1} \ldots p_{i_{n-1} i}\, p_{i j_1} \ldots p_{j_{m-1} j_m}}{\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i)} \\
&= \frac{\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i, X_{n+1} = j_1, \ldots, X_{n+m} = j_m)}{\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i)} \\
&= \mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m | X_0 = i_0, \ldots, X_n = i)
\end{aligned}$$

De esta manera tenemos que la ecuacion (2) de perdida de memoria se cumple para las cadenas de Markov. □

4. Muestre que para todo $n \geq 1$

$$\mathbb{P}(X_n = j | X_0 = i) = p_{ij}^{(n)}. \tag{3}$$

donde $P^n = (p_{ij}^{(n)})$.

Demostración. Vamos a demostrar la igualdad con induccion sobre n.

Caso base: Tomemos a $n = 1$. Por tanto $P^n = P$. Por definicion de las cadenas de Markov tenemos que $\mathbb{P}(X_1 = j | X_0 = i) = p_{ij}^{(1)} = p_{ij}$. De manera que la ecuacion (3) se cumple para cuando n es igual a 1

Paso inductivo: Suponga que la ecuacion (3) vale para n, es decir que:

$$\mathbb{P}(X_n = j | X_0 = i) = p_{ij}^{(n)}. \tag{4}$$

$$P^{n+1} = P^n P = (p_{ij}^{(n)})(p_{ij}) = (\sum_{k=1}^{n} p_{ik}^{(n)}\, p_{kj})$$

Por lo tanto $p_{ij}^{(n+1)} = \sum_{k=1}^{n} p_{ik}^{(n)}\, p_{kj}$

Ahora, para la expresion estamos asumiendo que con segutidad ocurre que $X_0 = i$ por tanto podemos asumir que $\mathbb{P}(X_0 = i) = 1$ o por lo menos diferente de cero.

$$\begin{aligned}
\mathbb{P}(X_{n+1} = j | X_0 = i) &= \mathbb{P}(X_{n+1} = j, X_0 = i) \frac{1}{\mathbb{P}(X_0 = i)} \\
&= \sum_{k \in S} \mathbb{P}(X_{n+1} = j, X_0 = i, X_n = k) \frac{1}{\mathbb{P}(X_0 = i)} \\
&= \sum_{k \in S} \mathbb{P}(X_{n+1} = j | X_n = k, X_0 = i) \frac{\mathbb{P}(X_n = k, X_0 = i)}{\mathbb{P}(X_0 = i)} \\
&= \sum_{k \in S} \mathbb{P}(X_{n+1} = j | X_n = k) \mathbb{P}(X_n = k | X_0 = i) \\
&= \sum_{k \in S} p_{kj}\, p_{ik}^{(n)} \\
&= \sum_{k \in S} p_{ik}^{(n)}\, p_{kj} \\
&= p_{ij}^{(n+1)}
\end{aligned}$$

Entonces se cumple en el paso inductivo, por tanto la ecuacion (3) vale para toda $n \in \mathbb{Z}^+$

□

5. Sea $\{X_n, n \geq 0\}$ cadena de Markov con matriz de transición P y distribución inicial π, con π distribución estacionaria. Muestre que para todo $n \geq 1$ y todo estado j se tiene que

$$\mathbb{P}(X_n = j) = \pi(j). \tag{5}$$

Demostración. .

Importante:

$$\begin{bmatrix}\pi(1), & \pi(2), & \cdots & \pi(n)\end{bmatrix} \times \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} = \begin{bmatrix} \sum_{k\in S}\pi(k)p_{k1}, & \cdots & \sum_{k\in S}\pi(k)p_{kn} \end{bmatrix}$$

Como $\pi P = \pi$, tenemos que $\pi(j) = \sum_{k\in S}\pi(k)p_{kj}$
Adicionalmente tenemos que $\pi P^n = \pi P = \pi$, Por tanto $\pi(j) = \sum_{k\in S}\pi(k)p^{(n)}_{kj}$

$$\begin{aligned}\mathbb{P}(X_n = j) &= \sum_{k\in S}\mathbb{P}(X_n = j|X_0 = k)\mathbb{P}(X_0 = k) \\ &= \sum_{k\in S} p^{(n)}_{kj}\pi(k) \\ &= \sum_{k\in S} p_{kj}\pi(k) \\ &= \pi(j)\end{aligned}$$

□

6. Muestre que si P es reversible con μ que satisface las ecuaciones de balance detallado, entonces μ es distribucion estacionaria.

Demostración. Primero veamos que $\sum_{k\in S}\mu(k) = \mu(\Omega) = 1$.
Ahora queremos ver que $\mu P = \mu$ sabendo que $\mu(i)p_{ij} = \mu(j)p_{ji}$

$$\begin{aligned}\mu P &= (\sum_{k\in S}\mu(k)p_{k1}\,,\ \sum_{k\in S}\mu(k)p_{k2}\,,\ \cdots \sum_{k\in S}\mu(k)p_{kn}) \\ &= (\sum_{k\in S}\mu(1)p_{1k}\,,\ \sum_{k\in S}\mu(2)p_{2k}\,,\ \cdots \sum_{k\in S}\mu(n)p_{nk}) \\ &= (\mu(1)\sum_{k\in S}p_{1k}\,,\ \mu(2)\sum_{k\in S}p_{2k}\,,\ \cdots \mu(n)\sum_{k\in S}p_{nk}) \\ &= (\mu(1)\,,\ \mu(2)\,,\ \cdots \mu(n)) \\ &= \mu\end{aligned}$$

Y como $\mu P = \mu$, podemos decir que μ es una distribucion estacionaria. □

7. Sea Ψ una matriz de transición irreducible y π una distribución sobre S. Defina una nueva matriz P dada por

$$p_{ij} = \begin{cases} \psi_{ij}\, min(\frac{\pi(j)\psi_{ji}}{\pi(i)\psi_{ij}}, 1) & si \quad j \neq i \\ \\ 1 - \sum_{k \neq i} \psi_{ik}\, min(\frac{\pi(k)\psi_{ki}}{\pi(i)\psi_{ik}}, 1) & si \quad j = i \end{cases}$$

Demostración. Para ver que π es la única distribución estacionaria de P, vamos a ver que la cadena de Marcov que esta dada por la matriz P es irreducible, es decir para todo$i, j \in S$ existe un n tal que $\mathbb{P}(X_n = j | X_0 = i) > 0$.
Como Ψ es una matriz de transición irreducible entonces para todo $i, j \in S$, $\psi_{ij}^{(n)} > 0$ para algun $n \in \mathbb{Z}^+$

Veamos el caso donde $j \neq i$ en donde

$$p_{ij}^{(n)} = \psi_{ij}^{(n)}\, min(\frac{\pi(j)\psi_{ji}^{(n)}}{\pi(i)\psi_{ij}^{(n)}}, 1)$$

Y recordando que $\pi(j) = \sum_{k \in S} \psi_{kj}^{(n)} \pi(k)$.
Pero como algun $\psi_{kj}^{(n)} > 0$, entonces y como π es una distribucion tenemos que $\pi(j) > 0$ y asi sabemos que para algun n, $p_{ij}^{(n)} > 0$.

Veamos el caso donde $j = i$, con algun n en donde

$$p_{ii}^{(n)} = 1 - \sum_{k \neq i} \psi_{ik}^{(n)}\, min(\frac{\pi(k)\psi_{ki}^{(n)}}{\pi(i)\psi_{ik}^{(n)}}, 1)$$

Como se itera sobre todo el espacio, tenemos que para algun $k \in S$, $\pi(k)\psi_{ki}^{(n)} < \pi(i)\psi_{ik}^{(n)}$ o lo que es equivalente decir que $\frac{\pi(k)\psi_{ki}^{(n)}}{\pi(i)\psi_{ik}^{(n)}} < 1$. Esto implica que $\sum_{k \neq i} \psi_{ik}^{(n)}\, min(\frac{\pi(k)\psi_{ki}^{(n)}}{\pi(i)\psi_{ik}^{(n)}}, 1) < 1$. Por lo tanto $p_{ii}^{(n)} > 0$.

Asi la cadena de Markov sobre la matriz P es irreducible entonces existe una unica distribucion estacionaria. □

8. Dados diferentes valores de λ, simule la cadena de Markov correspondiente a la matriz P y dejela correr por los tiempos que estime necesarios para acercarse lo mas posible a la configuracion que alcanza el maximo. Para esto use el grafo dado en esta libreria: `http://bqp.cs.uni-bonn.de/library/html/g05_60.0.html`

Demostración. Vease la carpeta llamada "max-cut.en el siguiente repositorio de git-hub, especificamente el documento llamado "max-cut.py"donde se implemente el codigo.

`https://github.com/JuanDDY/Proyecto_probabilidad_2023-2.git`

Tambien se encuentra un archivo lamado "grafo60.txt"que contiene el grafo con 60 nodos 885 aristas y que se pide y el archivo "lecturaGrafo.py"lee el archivo que contiene el grafo. Este archivo es importado por el archivo principal "max-cut.py"

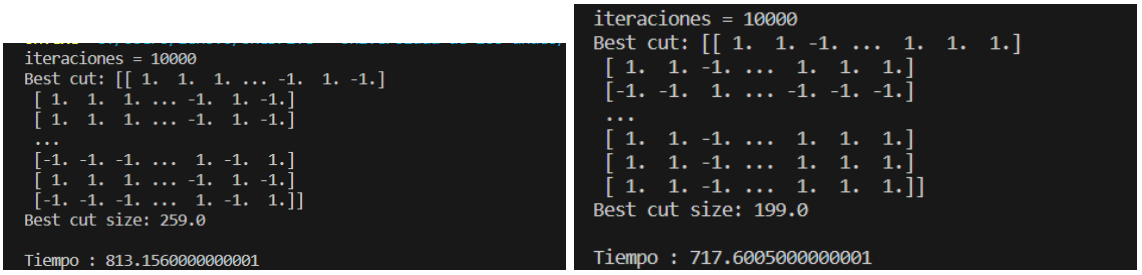
Lo primero que se hace es generar el grafo en una matriz de abyacencia y con esta matriz se ejecuta el algoritmo.

Se genera un vector aleatorio de tamaño $n \times 1$ con valores dados por una distribucion $Normal(0, 1)$ Con este vector se crea una matriz que tiene el vector en la diagonal y a esta matris se le saca el signo. Ahora se suma sobre los signos que son positivos y asi obtenemos el numero de corte que se encontro.

Ahora se compara el numero de cortes con dicho vector aleatorio con el actual mayor numero de cortes. Ahora se repite esto para hallar el corte maximo entre las iteraciones que se indiquen.

A continuacion se ven algunos ejemplos con cierto numero de iteraciones:

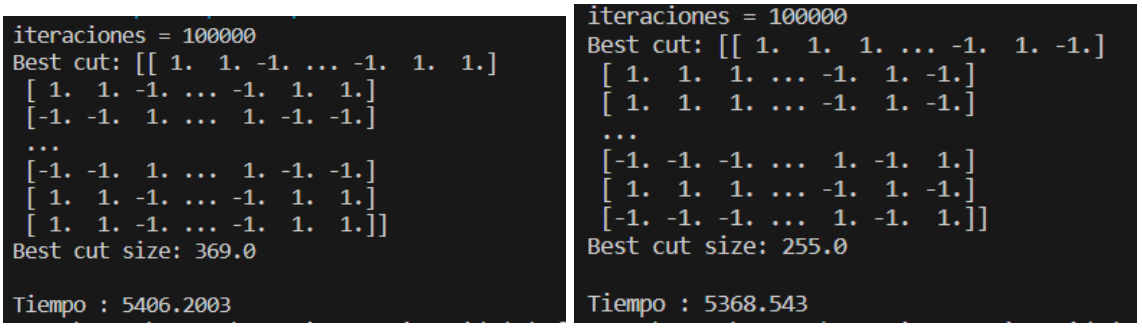
10000 iteraciones (Diez mil iteraciones)



(a) Prueba 1 (b) Prueba 2

Figura 2: Prueba con Diez mil iteraciones

100000 iteraciones (Cien mil iteraciones)



(a) Prueba 1 (b) Prueba 2

Figura 3: Prueba con cien mil iteraciones

1000000 iteraciones (Un millón de iteraciones)

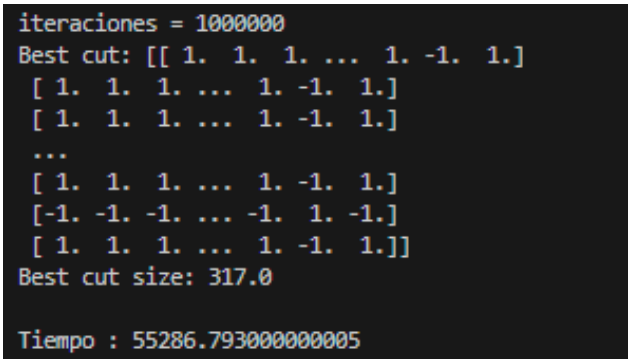


Figura 4: Prueba con un millón de iteraciones

10000000 iteraciones (Diez millones de iteraciones)

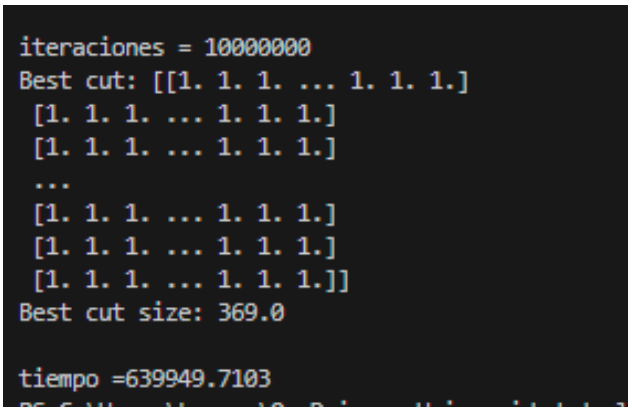


Figura 5: Prueba con diez millones de iteraciones

Dadas estas pruebas, se concluye que con 10000 (Diez mil) iteraciones se pueden obtener resultados muy buenos y los tiempos de procesamiento no son altos ya que para este numero de iteraciones se gastaron 5 segundos en cada prueba. Para un millon de iteraciones se gastaron 50 segundos.
Tambien cabe resaltar que el codigo tiene complegidad temporal de $O(n \times E)$ con n el numero de iteraciones y E el numero de arcos. Como se uso el mismo grafo para estas pruebas, el tiempo parece lineal con respecto de n el numero de iteraciones.

□