Cadenas de Markov

Definición.

Cadena de Markov o modelo de Markov es un proceso estocástico que consiste en un número finito de estados, el cual describe una secuencia de posibles eventos, en los que la probabilidad de cada evento depende solamente del estado alcanzado en el evento anterior con unas probabilidades que están fijas.

En definición formal, una cadena de Markov es una secuencia X1, X2, X3,… de variables aleatorias. El valor de Xn es el estado del proceso en el tiempo n. De esta manera se obtiene la propiedad markoviana:

$$P\,(X_{n+1} = X_{n+1} \mid X_n = x_n, X_{n-1} = x_{n-1,\dots,} X_2 = x_{2,} X_1 = x_1) = P\,(X_{n+1} = X_{n+1} \mid X_n)$$

Donde x_i es el estado del proceso en el instante *i*.

Propiedades de las cadenas de Markov.

Las propiedades de esta cadena son variadas y cuentan con funcionalidades que permiten procesar ciertos casos específicos:

Cadena irreducible:

Es irreducible o ergódica cuando existe una total comunicación entre todos sus estados, lo que significa que la probabilidad de que un estado pase a otro, es positiva.

Presentando las siguientes condiciones:

- En una cadena de Markov un estado e_j se dice que es **accesible** desde otro estado e_i si la probabilidad de ir desde el estado e_i al e_j en algún momento futuro es distinta de cero.
- Una cadena de Markov se dice que es **irreducible** si todos los estados están comunicados entre sí.
- Un estado ei está **comunicado** con otro e_j si e_j es accesible desde e_i y e_i lo es desde e_j.

La cadena de Ehrenfest o la caminata aleatoria sin barreras absorbentes son ejemplos de cadenas de Márkov irreducibles.

Cadena regulares:

Una cadena es regular en caso de que se pueda pasar de cualquier estado a otro al seguir una cantidad de pasos específicos, mientras que el estado inicial siga siendo independiente. Presenta las siguientes condiciones:

- Se dice que es **regular** si alguna potencia de la matriz de transición tiene todos sus elementos positivos (no hay ceros)
- Si una cadena es regular entonces es **irreducible** y **no absorbente**.

Matemáticamente, cuando el espacio de estados **S** es finito, si **P** denota la matriz de transición de la cadena se tiene que:

$$\lim_{n\to 1} P^n = W$$

donde **W** es una matriz con todos sus renglones iguales a un mismo vector de probabilidad **w**, que resulta ser el vector de probabilidad invariante de la cadena. En el caso de cadenas regulares, este vector invariante es único.

Cadena recurrentes positivas:

Se dice que es **recurrente positiva** si todos sus estados son recurrentes positivos. Además, si la cadena es irreducible es posible demostrar que existe un único vector de probabilidad invariante y está dado por:

$$\pi_j = \frac{1}{\mu_j}$$

Cadena absorbente:

1. Un estado e_j se dice absorbente si es imposible abandonar dicho estado, es decir:
$$P(X_k = e_j \mid X_{k-1} = e_j) = 1$$
2. Es absorbente cuando alguno de sus estados ya es absorbente.
3. De cualquier estado no absorbente se accede a algún estado absorbente. Es decir, se puede llevar al estado absorbente, ya que suelen ser transitorios después de un tiempo infinito medio y se convierte en un promedio positivo recurrente.

Si denotamos como **A** al conjunto de todos los estados absorbentes y a su complemento como **D**, tenemos los siguientes resultados:

- Su matriz de transición es de la siguiente forma:

$$P = \begin{pmatrix} Q & R \\ 0 & I \end{pmatrix}$$

  donde la submatriz Q corresponde a los estados del conjunto **D**, **I** es la matriz identidad, **0** es la matriz nula y **R** alguna submatriz.

- La siguiente representación significa que no importa en donde se encuentre la cadena, eventualmente terminará en un estado absorbente.

$$P_x(T_A < 1) = 1$$

Cadena homogénea: Es una cadena homogénea si la probabilidad de ir del estado i al estado j en un paso no depende del tiempo en el que se encuentra la cadena, es decir, es homogénea cuando la probabilidad de transición es plenamente independiente en relación al tiempo.

$$P\ (X_n = j \mid X_{n-1} = i\ ) = P\ (X_1 = j \mid X_o = i\ )$$

Para todo n y para cualquier i, j.

Si para alguna pareja de estados y para determinado el tiempo n, la propiedad antes mencionada no se cumple se puede decir que la cadena de Markov es no homogénea.

Cadena de Markov a tiempo continuo: Si se consideran las variables aleatorias X_t con t que varía en un intervalo continuo del conjunto R de los números reales, tendremos una cadena en tiempo continuo. Para este tipo de cadenas la propiedad de Márkov se expresa de la siguiente manera:

$P(X(t_{n+1}) = x_{n+1}|X(t_n) = x_n, \ldots, X(t_1) = x_1) = P(X(t_{n+1}) = x_{n+1}|X(t_n) = x_n)$ tal que $t_{n+1} > t_n > t_{n-1} > \cdots > t_1$

Para una cadena Markov continua con número finito de estados puede definirse una matriz estocástica dada por:

$$\mathbf{P}(t_1, t_2) = [p_{ij}(t_1, t_2)]_{i,j=1,\ldots,N}, \qquad p_{ij}(t_1, t_2) = P[X(t_2) = j|X(t_1) = i],\ 0 \geq t_1 < t_2$$

La cadena se denomina homogénea si $\mathbf{P}(t_1, t_2) = \mathbf{P}(t_2 - t_1)$. Para esta cadena en tiempo continuo homogénea y con número finito de estados se define el llamado generador infinitesimal como:

$$\mathbf{Q} = \lim_{h \to 0^+} \frac{\mathbf{P}(h) - \mathbf{I}}{h}$$

Y la matriz estocástica viene dada por:

$$\mathbf{P}(t) = e^{\mathbf{Q}t} = \sum_{n=0}^{\infty} \frac{\mathbf{Q}^n t^n}{n!}$$

Aplicaciones.

La aplicación y utilidad de las cadenas de Markov está presente en una amplia gama de temas.

Física: Aplicado en termodinámica y mecánica estadística, usados para representar detalles desconocidos o no modelados del sistema. Por ejemplo, el método **Markov Chain Monte Carlo** para extraer muestras al azar de una caja negra y aproximar la distribución de probabilidad de los atributos en un rango de objetos. O usada para la formulación integral de **caminos** en mecánica cuántica.

Operaciones: Empleada en inventarios, mantenimiento y flujo de procesos.

Meteorología: Considerando el tiempo atmosférico de una región en distintos días, es posible asumir que su estado actual solo depende del último estado y no de todo la historia en sí, por ende con estas cadenas podemos formular modelos climatológicos básicos, como los **modelos de recurrencias de lluvias**.

Redes neuronales: Utilizada en máquinas de Boltzmann.

Biología: En **Filogenética y bioinformática**, donde los modelos de evolución del ADN utilizan cadenas para describir el nucleótido presente en un sitio determinado del genoma.
Dinámica de poblaciones, es una herramienta central para el estudio teórico de modelos de población matriciales
Neurobiología, para simular la neocorteza de mamíferos.
Biología de sistemas, modelo de infección viral de células individuales.
Modelado de brotes de enfermedades y epidemias.

Reconocimiento de voz: Son la base de la mayoría de sistemas automáticos de reconocimiento de voz modernos.

Química: Para analizar el crecimiento y la composición de los copolímeros. Enfocar el crecimiento basado en fragmentos de productos químicos en silicio hacia una clase deseada de compuestos. El modelo clásico de actividad enzimática, la cinética de **Michaelis-Menten**, donde cada reacción es una transición de estado en una cadena de Markov.

$$\mathrm{E} + \mathrm{S} \underset{\substack{\text{Substrate}\\ \text{binding}}}{\rightleftharpoons} \mathrm{ES} \xrightarrow{\substack{\text{Catalytic}\\ \text{step}}} \mathrm{E} + \mathrm{P}$$

Internet: Utilizado para analizar el comportamiento de navegación web de los usuarios. **Pagerank** usado por Google en sus motores de búsqueda, donde la posición que tendrá la página al buscarse será determinada por su peso en la distribución estacionaria de la cadena.

Estadísticas: Generar secuencia de números aleatorios para reflejar con precisión distribuciones de probabilidad deseadas muy complejas, utilizando un proceso llamado cadena de Markov Monte Carlo**(MCMC)**, esto ha revolucionado la viabilidad de los métodos de inferencia bayesianos.

Economía y finanzas: Usado para modelar una **variedad de fenómenos**, incluida la distribución del ingreso, los precios de los activos, los colapsos de una bolsa de valores o la volatilidad de los precios.
En la **macroeconómica** se usa para modelar exógenamente los precios de las acciones en un entorno de equilibrio general.
En los **negocios**, analizar patrones de compra de los deudores morosos, planear las necesidad de personal y analizar reemplazo de equipo.
Las agencias de **calificación crediticia** elaboran tablas de probabilidad de transición para determinar los bonos de acuerdo a las diferentes calificaciones crediticias.

Pronósticos probabilísticos: Empleado para pronósticos en diversas áreas, tendencias de precios, energía eólica e irradiación solar, la cadena de Markov presenta una variedad de configuraciones que permiten llegar a estas implementaciones.

Juegos: Utilizado para modelar muchos juegos de azar

Simulación: Empleado para proveer una solución analítica a ciertos problemas de simulación, como la **teoría de colas**. Las cadenas de Markov son la base para el tratamiento **analítico de las colas**. Esto les permite optimizar el rendimiento de las redes de telecomunicaciones, donde los mensajes a menudo deben competir por recursos limitados (como el ancho de banda). Numerosos modelos de colas utilizan cadenas de Markov de tiempo continuo.

Ciencias sociales: Es común usarlo para **modelar** como una vez un país alcanza un nivel de desarrollo económico, la configuración de factores estructurales, el tamaño de la clase media, la proporción de residencia urbana a rural, etc. Como el avance en estos factores genera una mayor probabilidad de **transitar** de **situaciones autoritarias** al **régimen democratico**.

Música: Se emplean en la **composición algorítmica** de música, particularmente en software como Csound, Max y SuperCollider. En una cadena de primer orden, los estados del sistema se convierten en valores de nota o tono, y se construye un vector de probabilidad para cada nota, completando una matriz de probabilidad de transición. Uno de los compositores que usó esta técnica fue **Lannis Xenakis** considerado el pionero de la música electrónica y creador de la música estocástica, compuso la obra *Analoguique A et B* (1958–59).

Teoría de información: Estas cadenas se utilizan en todo el procesamiento de la información. Diversos modelos pueden capturar muchas de las **regularidades estadísticas** de los sistemas, incluso sin describir perfectamente la estructura completa, causando una compresión de datos muy efectiva a través de técnicas de codificación de entropía como la **codificación aritmética**. También permiten una **estimación** efectiva del **estado** y el **reconocimiento de patrones**.

El modelo de Markov también juega un papel importante en el **aprendizaje por refuerzo(RL)**. Además, son la base de los modelos ocultos de Markov, las cuales, son herramientas importantes en campos como las **redes telefónicas, reconocimiento de voz y bioinformática**. Asimismo, el algoritmo de compresión de datos sin pérdida **LZMA** combina cadenas de Markov con compresión **Lempel-Ziv** para lograr relaciones de entendimiento muy altas.

Estas son algunas de las aplicaciones para las cadenas de Markov, como se puede denotar sus aplicaciones son notables y abarcan diversos campos, considerando la utilidad e importancia de este modelo.

Ejemplos

Ejercicio Camiones
https://drive.google.com/file/d/1LpTj385mUwh615J06jp3-DauYJtohlyO/view?usp=sharing

Generador de Frases
https://github.com/aarroyoc/blog-ejemplos/blob/master/markov-hector-del-mar/main.py