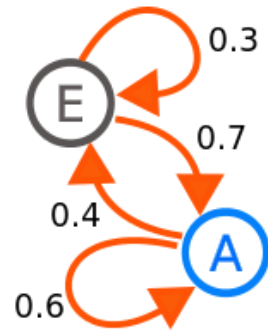


Cadena de Márkov

En la teoría de la probabilidad, se conoce como **cadena de Márkov** o **modelo de Márkov** a un tipo especial de proceso estocástico discreto en el que la probabilidad de que ocurra un evento depende *solamente* del evento inmediatamente anterior. Esta característica de incluir una memoria reciente recibe el nombre de *propiedad de Markov* en contraste con los eventos independientes que no tienen memoria de ningún evento anterior. En un primer artículo de 1906 A. A. Markov definió la "cadena simple" como "una secuencia infinita $x_1, x_2, \dots, x_k, x_{k+1}, \dots$, de variables conectadas de tal modo que x_{k+1} para cualquier k es independiente de x_1, x_2, \dots, x_{k-1} , en el caso de que x_k sea conocida". Markov llamó a la cadena "homogénea" si la distribución condicional de x_{k+1} dado x_k fuese independiente de k . También consideró cadenas "complicadas (complex en inglés)" en las que "cada número está conectado directamente no sólo con uno, sino con varios números anteriores".¹

Recibe su nombre del matemático ruso Andréi Márkov (1856-1922), que lo introdujo en 1906.¹

Estos modelos estadísticos cuentan con un gran número de aplicaciones reales.



Cadena simple biestable de Markov

Índice

Definición

Características

- Cadenas homogéneas y no homogéneas
- Probabilidades de Transición
- Matriz de Probabilidades de Transición
- Ecuación de Chapman-Kolmogorov
- Clases de comunicación
- Periodicidad
- Tiempos de Primera Visita
 - Probabilidad de Primera Visita
- Recurrencia
- Tiempo Medio de Recurrencia
- Distribuciones Estacionarias
 - Existencia y Unicidad
 - Convergencia a la distribución estacionaria
 - Convergencia para Cadenas de Markov

Tipos de Cadenas de Markov

- Cadenas irreducibles
- Cadenas Recurrentes Positivas
- Cadenas Regulares
- Cadenas Absorbentes
- Cadenas de Markov a tiempo continuo

Aplicaciones

- Meteorología
- Modelos epidemiológicos
- Internet
- Simulación
- Juegos de azar
- Economía y finanzas
- Genética
- Música
- Operaciones

Definición

En matemáticas, una **Cadena de Markov** es un proceso estocástico a tiempo discreto $\{X_n : n = 0, 1, 2, \dots\}$ con espacio de estados discreto S que para cualquier entero $n \geq 0$ y para cualesquiera $x_0, x_1, \dots, x_{n+1} \in S$ satisface

$$P[X_{n+1} = x_{n+1} | X_0 = x_0, X_1 = x_1, \dots, X_n = x_n] = P[X_{n+1} = x_{n+1} | X_n = x_n]$$

a esta propiedad se le conoce como propiedad de Markov.

Características

Cadenas homogéneas y no homogéneas

Se dice que una **Cadena de Markov** es **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, esto es:

$$P[X_{n+1} = j | X_n = i] = P[X_1 = j | X_0 = i]$$

para todo $n \geq 0$ y para cualquier $i, j \in S$.

Si para alguna pareja de estados y para algún tiempo n la propiedad antes mencionada no se cumple entonces diremos que la Cadena de Markov es **no homogénea**.

Probabilidades de Transición

Sean i y j dos estados de una Cadena de Markov. La probabilidad de ir del estado i en el tiempo n al estado j en el tiempo $n + 1$ se denota por

$$p_{ij}(n, n + 1) = P[X_{n+1} = j | X_n = i].$$

Cuando la cadena es homogénea, esta probabilidad se denota por

$$p_{ij} = P[X_{n+1} = j | X_n = i],$$

que representa la probabilidad de pasar del estado i al estado j en una unidad de tiempo.

Matriz de Probabilidades de Transición

Teniendo las probabilidades de transición en un paso p_{ij} , si variamos los índices i, j sobre el espacio de estados $S = \{0, 1, 2, \dots\}$ obtenemos la matriz P llamada **matriz de probabilidades de transición** en un paso, es decir:

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

donde la entrada (i, j) representa la probabilidad de pasar del estado i al estado j en un paso.

La matriz P es una matriz estocástica pues satisface

- $p_{ij} \geq 0$
- $\sum_{j \in S} p_{ij} = 1$

Similarmente se define la matriz de probabilidades de transición en n pasos, esta se denota por $P(n)$ y está dada por

$$P(n) = \begin{bmatrix} p_{00}(n) & p_{01}(n) & p_{02}(n) & \cdots \\ p_{10}(n) & p_{11}(n) & p_{12}(n) & \cdots \\ p_{20}(n) & p_{21}(n) & p_{22}(n) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

donde la entrada (i, j) representa la probabilidad de pasar del estado i al estado j en n pasos.

Ecuación de Chapman-Kolmogorov

Para cualesquiera $r, n \in \mathbb{Z}$ tales que $0 \leq r \leq n$ y para cualesquiera estados $i, j \in S$ se cumple

$$p_{ij}(n) = \sum_{k \in S} p_{ik}(r) p_{kj}(n-r)$$

Como consecuencia de este resultado, la probabilidad de transición en n pasos, $p_{ij}(n)$, está dada por la entrada (i, j) de la n -ésima potencia de la matriz de probabilidades de transición en un paso, es decir

$$p_{ij}(n) = (P^n)_{ij}$$

Con lo anterior, el problema de calcular las probabilidades de transición en n pasos se convierte en halla la n -ésima potencia de la matriz de probabilidades de transición en un paso, esto es

$$\begin{aligned} P(n) &= \begin{bmatrix} p_{00}(n) & p_{01}(n) & \cdots \\ p_{10}(n) & p_{11}(n) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \\ &= \begin{bmatrix} p_{00} & p_{01} & \cdots \\ p_{10} & p_{11} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}^n \\ &= P^n \end{aligned}$$

Clases de comunicación

Para dos estados i y j en el espacio de estados S , diremos que el estado j es accesible desde el estado i y escribiremos $i \rightarrow j$ si $\exists n \in \mathbb{Z}^+$ tal que

$$p_{ij}(n) > 0$$

si $i \rightarrow j$ y $j \rightarrow i$ entonces diremos que el estado i se comunica con el estado j y escribiremos $i \longleftrightarrow j$.

La propiedad " \longleftrightarrow " es una relación de equivalencia. Esta relación induce una partición del espacio de estados. A estas clases de equivalencia las llamaremos *clases de comunicación*.

Dado un estado $i \in S$, denotaremos a su clase de comunicación como $C(i)$, por lo que $i \longleftrightarrow j$ si y sólo si $C(i) = C(j)$.

Si $C(i) = S$ entonces se dice que la cadena es irreducible.

Periodicidad

El *periodo* de un estado $i \in S$ se define como:

$$d(i) = \text{mcd}\{n \geq 1 : p_{ii}(n) > 0\}$$

donde **mcd** denota el máximo común divisor.

- Si $d(i) = 1$ diremos que i es un estado **aperiódico**.
- Si $d(i) = k \geq 2$ diremos que i tiene periodo k .

Una cadena de Márkov se dice **aperiódica** si todos sus estados son *aperiódicos*, es decir, si $d(i) = 1 \quad \forall i \in S$.

Tiempos de Primera Visita

Si $C \subset S$, definimos el **tiempo de primera visita** a C como la variable aleatoria

$$\tau_C = \begin{cases} \min\{n > 0 | X_n \in C\} & \text{si } \{n > 0 | X_n \in C\} \neq \emptyset \\ 1 & \text{si } \{n > 0 | X_n \in C\} = \emptyset \end{cases}$$

esto es, τ_C denota la primera vez que la cadena entra al conjunto de estados C .

Probabilidad de Primera Visita

Se define

$$f_{ij}(n) = P[X_n = j, X_{n-1} \neq j, \dots, X_1 \neq j | X_0 = i]$$

como la probabilidad de que una cadena que inicia en el estado i llegue al estado j por primera vez en n pasos, donde $f_{ij}(0) = 0$.

En particular, cuando $i = j$, $f_{ii}(n)$ denota la probabilidad de regresar por primera vez al estado i en n pasos.

Y se definen

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n)$$

como la probabilidad de una eventual visita a partir del estado i al estado j y

$$f_{ii} = \sum_{n=1}^{\infty} f_{ii}(n)$$

como la probabilidad de partir del estado i y regresar a él mismo en un tiempo finito.

Recurrencia

En una cadena de Markov con espacio de estados S , diremos que:

- i es un estado *recurrente* si $f_{ii} = 1$.
- i es *transitorio* si $f_{ii} < 1$.

o utilizando las probabilidades de transición en n pasos:

- i es recurrente si $\sum_{n=1}^{\infty} p_{ii}(n) = \infty$
- i es transitorio si $\sum_{n=1}^{\infty} p_{ii}(n) < \infty$

La recurrencia es una propiedad de clase pues

- Si i es recurrente e $i \longleftrightarrow j$ entonces j es recurrente.
- Si i es transitorio e $i \longleftrightarrow j$ entonces j es transitorio.

AA.

Tiempo Medio de Recurrencia

Se define como el tiempo medio de recurrencia de un estado recurrente j a partir del estado i como la esperanza de

$$\tau_{ij} = \min\{n \geq 1 : X_n = j | X_0 = i\}$$

y se denota por μ_{ij}

$$\mu_{ij} = E[\tau_{ij}] = \sum_{n=1}^{\infty} n f_{ij}(n),$$

Esta esperanza representa el número de pasos promedio que a la cadena le toma regresar al estado recurrente j .

En particular, cuando $i = j$ escribimos μ_i en lugar de μ_{ij} .

Se dice que un estado recurrente i es

- recurrente nulo si $\mu_i = \infty$.
- recurrente positivo si $\mu_i < \infty$.

La recurrencia positiva es una propiedad de clase pues

- Si i es recurrente positivo e $i \longleftrightarrow j$ entonces j es recurrente positivo.
- Si i es recurrente nulo e $i \longleftrightarrow j$ entonces j es recurrente nulo.

Distribuciones Estacionarias

Se dice que el vector $\pi = (\pi_0, \pi_1, \dots)$ es una distribución de probabilidad si

- $\pi_i \geq 0$
- $\sum_i \pi_i = 1$

Se dice que una distribución de probabilidad $\pi = (\pi_0, \pi_1, \dots)$ es **estacionaria** para una Cadena de Markov con matriz de probabilidades de transición $P = (p_{ij})$ si

$$\pi_j = \sum_{i \in S} \pi_i p_{ij}$$

En forma matricial lo anterior es equivalente a $\pi = \pi P$ y significa que si una variable aleatoria inicial X_0 tiene una distribución π entonces la distribución de X_n también es π , es decir, esta distribución no cambia con el paso del tiempo.

Para encontrar una posible distribución estacionaria de una cadena con matriz P , un método consiste en resolver el sistema de ecuaciones

$$\left\{ \begin{array}{l} \pi = \pi P \\ \text{Sujeto a:} \\ \sum_{j \in S} \pi_j = 1 \\ \pi_j \geq 0 \end{array} \right.$$

La distribución estacionaria puede no ser única o incluso no existir.

Existencia y Unicidad

Si una Cadena de Markov es **irreducible** y **recurrente positiva** entonces tiene una única distribución estacionaria y esta está dada por

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

donde μ_j es el tiempo medio de recurrencia del estado j .

Convergencia a la distribución estacionaria

Si una cadena de Markov es

- Irreducible
- Aperiódica
- Con distribución estacionaria π

entonces para cualesquiera $i, j \in S$

$$\lim_{n \rightarrow \infty} p_{ij}(n) = \pi_j$$

Convergencia para Cadenas de Markov

Si una cadena de Markov es

- Irreducible
- Recurrente positiva
- Aperiódica

entonces las probabilidades límite

$$\pi_j = \lim_{n \rightarrow \infty} p_{ij}(n)$$

existen, están dadas por

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

y constituyen la única solución al sistema de ecuaciones

$$\left\{ \begin{array}{l} \pi = \pi P \\ \text{Sujeto a:} \\ \sum_{j \in S} \pi_j = 1 \\ \pi_j \geq 0 \end{array} \right.$$

Tipos de Cadenas de Markov

Cadenas irreducibles

Una cadena de Markov se dice **irreducible** si se cumple cualquiera de las siguientes condiciones (equivalentes entre sí):

1. Desde cualquier estado de S se puede acceder a cualquier otro.
2. Todos los estados se comunican entre sí.

3. $C(i) = S$ para algún $i \in S$.
4. $C(i) = S$ para todo $i \in S$.
5. El único conjunto cerrado es el total.

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

Cadenas Recurrentes Positivas

Una cadena de Markov se dice **recurrente positiva** si todos sus estados son recurrentes positivos. Si la cadena es además irreducible es posible demostrar que existe un único vector de probabilidad invariante y está dado por:

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

Cadenas Regulares

Una cadena de Márkov se dice *regular* (también *primitiva* o *ergódica*) si existe alguna potencia positiva de la matriz de transición cuyas entradas sean todas estrictamente mayores que cero.

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

$$\lim_{n \rightarrow \infty} P^n = W$$

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.

Cadenas Absorbentes

Una cadena de Márkov con espacio de estados finito se dice absorbente si se cumplen las dos condiciones siguientes:

1. La cadena tiene al menos un estado absorbente.
2. De cualquier estado no absorbente se accede a algún estado absorbente.

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 siempre se puede llevar a una de la 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.

- $P_x(T_A < \infty) = 1$, esto es, no importa en donde se encuentre la cadena, eventualmente terminará en un estado absorbente.

Cadenas de Markov a tiempo continuo

Si en lugar de considerar una secuencia discreta $X_1, X_2, \dots, X_i, \dots$ con i indexado en el conjunto \mathbb{N} de números naturales, se consideran las variables aleatorias X_t con t que varía en un intervalo continuo del conjunto \mathbb{R} de números reales, tendremos una cadena en tiempo continuo. Para este tipo de cadenas en tiempo continuo la propiedad de Márkov se expresa de la siguiente manera:

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

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

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

La cadena se denomina homogénea si $\mathbf{P}(t_1, t_2) = \mathbf{P}(t_2 - t_1)$. Para una cadena de Márkov en tiempo continuo homogénea y con un número finito de estados puede definirse el llamado generador infinitesimal como:²

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

Y puede demostrarse que 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

Meteorología

Si consideramos el tiempo atmosférico de una región a través de distintos días, es posible asumir que el estado actual solo depende del último estado y no de toda la historia en sí, de modo que se pueden usar cadenas de Markov para formular modelos climatológicos básicos. Por ejemplo, se han desarrollado modelos de recurrencia de las lluvias basados en cadenas de Markov.³

Modelos epidemiológicos

Una importante aplicación de las cadenas de Markov se encuentra en el proceso Galton-Watson. Este es un proceso de ramificación que se puede usar, entre otras cosas, para modelar el desarrollo de una epidemia (véase modelaje matemático de epidemias).

Internet

El pagerank de una página web (usado por Google en sus motores de búsqueda) se define a través de una cadena de Markov, donde la posición que tendrá una página en el buscador será determinada por su peso en la distribución estacionaria de la cadena.

Simulación

Las cadenas de Márkov son utilizadas para proveer una solución analítica a ciertos problemas de simulación, por ejemplo en teoría de colas el Modelo M/M/1⁴ es de hecho un modelo de cadenas de Markov.

Juegos de azar

Son muchos los juegos de azar que se pueden modelar a través de una cadena de Márkov. El modelo de la ruina del jugador (*Gambler's ruin*), que establece la probabilidad de que una persona que apuesta en un juego de azar finalmente termine sin dinero, es una de las aplicaciones de las cadenas de Márkov en este rubro.

Economía y finanzas

Las cadenas de Márkov se pueden utilizar en modelos simples de valuación de opciones para determinar cuándo existe oportunidad de arbitraje, así como en el modelo de colapsos de una bolsa de valores o para determinar la volatilidad de los precios. En los negocios, las cadenas de Márkov se han utilizado para analizar los patrones de compra de los deudores morosos, para planear las necesidades de personal y para analizar el reemplazo de equipo.

Genética

Se emplean cadenas de Márkov en teoría de genética de poblaciones, para describir el cambio de frecuencias génicas en una población pequeña con generaciones discretas, sometida a deriva genética. Ha sido empleada en la construcción del modelo de difusión de Motō Kimura.

Música

Diversos algoritmos de composición musical usan cadenas de Márkov, por ejemplo el software Csound o Max. Uno de los compositores que usó esta técnica en sus composiciones fue Iannis Xenakis con su obra *Analogique A et B* (1958–59).

Operaciones

Se emplean cadenas de Márkov en inventarios, mantenimiento y flujo de proceso.

Redes neuronales

Se utilizan en las máquinas de Boltzmann.

Referencias

- Basharin, Gely P.; Langville, Amy N.; Naumov, Valeriy A. (2004). «The Life and Work of A. A. Markov» (<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.4.4887>). *Linear Algebra and its Applications* (en inglés) **386**: 3-26. Consultado el 31 de marzo de 2010.
- Masaki Kijima, 1997, p. 175
- R. Gabriel & J. Neumann (2006): *A Markov chain model for daily rainfall occurrence at Tel Aviv* (<http://onlinelibrary.wiley.com/doi/10.1002/qj.49708837511/abstract;jsessionid=13F79369B87DF0E62F46E141489D1AC7.f01t02?deniedAccessCustomisedMessage=&userIsAuthenticated=falseK.>)
- Masaki Kijima, 1997, pp. 287-290.

Bibliografía

- A.A. Márkov. "Rasprostranenie zakona bol'shih chisel na velichiny, zavisyaschie drug ot druga". *Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete*, 2-ya seriya, tom 15, pp. 135–156, 1906.
- A.A. Markov. "Extension of the limit theorems of probability theory to a sum of variables connected in a chain". reprinted in Appendix B of: R. Howard. *Dynamic Probabilistic Systems, volume 1: Markov Chains*. John Wiley and Sons, 1971.
- Classical Text in Translation: A. A. Markov, An Example of Statistical Investigation of the Text Eugene Onegin Concerning the Connection of Samples in Chains, trans. David Link. *Science in Context* 19.4 (2006): 591–600. Online: <http://journals.cambridge.org/production/action/cjoGetFulltext?fulltextid=637500>
- Leo Breiman. *Probability*. Original edition published by Addison-Wesley, 1968; reprinted by Society for Industrial and Applied Mathematics, 1992. ISBN 0-89871-296-3. (See Chapter 7.)
- J.L. Doob. *Stochastic Processes*. New York: John Wiley and Sons, 1953. ISBN 0-471-52369-0.
- S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. London: Springer-Verlag, 1993. ISBN 0-387-19832-6. en línea: [1] (https://web.archive.org/web/20100619010320/https://netfiles.uiuc.edu/meyn/www/spm_files/book.html) . Second edition to appear, Cambridge University Press, 2009.
- S. P. Meyn. *Control Techniques for Complex Networks*. Cambridge University Press, 2007. ISBN 978-0-521-88441-9. Appendix contains abridged Meyn & Tweedie. en línea: https://web.archive.org/web/20100619011046/https://netfiles.uiuc.edu/meyn/www/spm_files/CTCN/CTCN.html
- Booth, Taylor L. (1967). *Sequential Machines and Automata Theory* (<https://archive.org/details/SequentialMachinesAndAutomataTheory>) (1st edición). Nueva York: John Wiley and Sons, Inc. Library of Congress Card Catalog Number 67-25924. Extensive, wide-ranging book meant for specialists, written for both theoretical computer scientists as well as electrical engineers. With detailed explanations of state minimization techniques, FSMs, Turing machines, Markov processes, and undecidability. Excellent treatment of Markov processes pp. 449ff. Discusses Z-transforms, D transforms in their context.
- Kemeny, John G.; Mirkil, Hazleton; Snell, J. Laurie; Thompson, Gerald L. (1959). *Finite Mathematical Structures* (https://archive.org/details/finitemathematic0000keme_h5g0) (1st edición). Englewood Cliffs, N.J.:

Prentice-Hall, Inc. Library of Congress Card Catalog Number 59-12841. Classical text. cf Chapter 6 *Finite Markov Chains* pp. 384ff.

- Kijima, Masaaki (1997). *Markov Processes for Stochastic Modeling* (1st edición). Cambridge: Chapman & Hall. ISBN 0 412 60660 7.
- E. Nummelin. "General irreducible Markov chains and non-negative operators". Cambridge University Press, 1984, 2004. ISBN 0-521-60494-X

Enlaces externos

- Cadenas de Markov en tiempo discreto (http://www.investigaciondeoperaciones.net/cadenas_de_markov.html)
 - Ejemplo de una Cadena de Markov en tiempo discreto (<http://www.gestiondeoperaciones.net/cadenas-de-markov/ejemplo-de-una-cadena-de-markov-en-tiempo-discreto/>)
 - Techniques to Understand Computer Simulations: Markov Chain Analysis (<http://jasss.soc.surrey.ac.uk/12/1/6.html>) (en inglés)
 - Desambiguación mediante Cadenas de Markov (<https://web.archive.org/web/20140506010450/http://sistemassquestionanswering.es/pagina-principal/desambiguacionmarkov.html>)
 - Una explicación visual por Victor Powell (<http://setosa.io/blog/2014/07/26/markov-chains/index.html>)
-

Obtenido de «https://es.wikipedia.org/w/index.php?title=Cadena_de_Márkov&oldid=144861479»

Esta página se editó por última vez el 19 jul 2022 a las 19:45.

El texto está disponible bajo la Licencia Creative Commons Atribución Compartir Igual 3.0; pueden aplicarse cláusulas adicionales. Al usar este sitio, usted acepta nuestros términos de uso y nuestra política de privacidad.
Wikipedia® es una marca registrada de la Fundación Wikimedia, Inc., una organización sin ánimo de lucro.