Cadenas de Markov

Definición 1 *Un* proceso estocástico *es una colección de variables aleatorias* $\{X_t\}_{t\in\mathcal{T}}$ *sobre un mismo espacio de probabilidad, donde* $\mathcal{T}$ *es cualquier conjunto de índices.*

El conjunto E *de posibles valores de cada* X_t *(que asumimos que es independiente del índice* t*) se llama el* espacio de estados *del proceso.*

El conjunto de índices $\mathcal{T}$ se toma a menudo como un subconjunto numerable o continuo de $\mathbb{R}$, considerando de esa forma un proceso aleatorio como una variable aleatoria que evoluciona a lo largo del tiempo.

Definición 2 *Una* cadena de Markov *de tiempo discreto es un proceso estocástico* $\boldsymbol{X} = \{X_t\}_{t\in\mathbb{Z}_{\geq 0}}$ *que cumple la propiedad de Markov:*

$$\mathbb{P}(X_{t+1} \mid X_0, \dots, X_t) = \mathbb{P}(X_{t+1} \mid X_t), \quad \forall t \in \mathbb{Z}_{\geq 0}$$

Una cadena de Markov es homogénea *en el tiempo si las probabilidades anteriores no dependen del índice* t*:*

$$\mathbb{P}(X_{t+1} \mid X_t) = \mathbb{P}(X_1 \mid X_0), \quad \forall t \in \mathbb{Z}_{\geq 0}$$

En este tema solo consideramos cadenas de Markov de tiempo discreto homogéneas.

La distribución de una cadena de Markov viene determinada por

1. La distribución marginal de X_0, llamada *distribución inicial*.
2. La distribución condicional $X_{t+1} \mid X_t$, llamada *distribución de transición*.

Si el espacio de estados de la cadena de Markov es finito, $E = \{x_1, \dots, x_d\}$, entonces la distribución inicial se puede representar como un vector $(p_1, \dots, p_d)^{\mathrm{T}} \in \mathbb{R}^d_{\geq 0}$ tal que

$$\sum_{i=1}^{d} p_i = 1$$
$$\mathbb{P}(X_0 = x_i) = p_i, \quad \forall i = 1, \dots, d$$

y la distribución de transición como una matriz $P = (p_{ij})_{i,j=1,\dots,d} \in \mathbb{R}^{d\times d}_{\geq 0}$ tal que

$$\sum_{j=1}^{d} p_{ij} = 1, \quad \forall i = 1, \dots, d$$
$$\mathbb{P}(X_{t+1} = x_j \mid X_t = x_i) = p_{ij}, \quad \forall i, j = 1, \dots, d$$

Si el espacio de estados es infinito numerable, entonces las distribuciones inicial y de transición se pueden representar como un vector y una matriz infinitas, respectivamente.

Si el espacio de estados es continuo, entonces debemos pensar en la distribución inicial como una distribución no condicional y en la distribución de transición como una distribución condicional.

Cadenas de Markov con espacio de estados finito o infinito numerable

Proposición 3 *Denotemos por $p_{ij}^{(t)}$ la probabilidad de que una cadena de Markov $\{X_t\}_{t\in\mathbb{Z}_{\geq 0}}$ pase del estado i al estado j en t unidades de tiempo, esto es*

$$p_{ij}^{(t)} = \mathbb{P}(X_t = j \mid X_0 = i)$$

Las ecuaciones de Chapman-Kolmogorov *establecen que*

$$p_{ij}^{(t+t')} = \sum_{k\in E} p_{ik}^{(t)} p_{kj}^{(t')}, \quad \forall i, j \in E, \forall t, t' \geq 0$$

Para las cadenas de Markov con espacio de estado finito o infinito numerable, denotando por $P^{(t)}$ la matriz de probabilidades de transición en t pasos, las ecuaciones anteriores se traducen en que $P^{(t)}$ se puede obtener multiplicando P consigo misma t veces.

Definición 4 *Sea $\boldsymbol{X}$ una cadena de Markov con espacio de estados E y matriz de transición P. Sea $\boldsymbol{\pi} = (\pi_i)_{i\in E}$ una distribución de probabilidad sobre E, es decir,*

$$0 \leq \pi_i \leq 1, \ \forall i \in E \qquad \text{y} \qquad \sum_{i\in E} \pi_i = 1$$

- *$\boldsymbol{\pi}$ es* invariante *o* estacionaria *para $\boldsymbol{X}$ si*

  $$X_0 \sim \boldsymbol{\pi} \quad \Rightarrow \quad X_t \sim \boldsymbol{\pi}, \ \forall t \in \mathbb{Z}_{\geq 0}$$

  Es fácil comprobar que esto se verifica si y solo si $\boldsymbol{\pi}^{\mathrm{T}} = \boldsymbol{\pi}^{\mathrm{T}} P$.

- *$\boldsymbol{\pi}$ es una* distribución límite *de $\boldsymbol{X}$ si*

  $$\pi_j = \lim_{t\to+\infty} p_{ij}^{(t)}, \quad \forall i, j \in E$$

Ejemplo: Modelo genético

El tipo más simple de herencia de rasgos en animales ocurre cuando el rasgo está gobernado por un par de genes, cada uno de los cuales puede ser de dos tipos, digamos G y g. Un individuo se dice dominante si tiene la combinación GG, recesivo si tiene la combinación gg e híbrido si tiene la combinación Gg (que es genéticamente la misma que gG).

Consideremos un proceso de apareamiento continuado en el que siempre uno de los progenitores es híbrido. Asumimos que siempre hay al menos un descendiente, uno de los cuales es elegido al azar para participar en el siguiente apareamiento.

Podemos modelizar este proceso mediante una cadena de Markov en la que los estados son pares de genes:

$$E = \{\mathrm{GG}, \mathrm{Gg}, \mathrm{gg}\}$$

La matriz de transición del proceso de apareamiento vendría dada entonces por

$$P = \begin{array}{c} \\ \mathrm{GG} \\ \mathrm{Gg} \\ \mathrm{gg} \end{array} \begin{array}{c} \begin{array}{ccc} \mathrm{GG} & \mathrm{Gg} & \mathrm{gg} \end{array} \\ \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \end{array}$$

Para determinar la distribución invariante $\boldsymbol{\pi} = (\pi_{\mathrm{GG}}, \pi_{\mathrm{Gg}}, \pi_{\mathrm{gg}})^{\mathrm{T}}$ de esta cadena de Markov basta resolver el sistema de ecuaciones dado por

$$\begin{gathered} (\pi_{\mathrm{GG}}, \pi_{\mathrm{Gg}}, \pi_{\mathrm{gg}})^{\mathrm{T}} P = (\pi_{\mathrm{GG}}, \pi_{\mathrm{Gg}}, \pi_{\mathrm{gg}})^{\mathrm{T}} \\ \pi_{\mathrm{GG}} + \pi_{\mathrm{Gg}} + \pi_{\mathrm{gg}} = 1 \end{gathered}$$

obteniéndose $\boldsymbol{\pi} = (1/4, 1/2, 1/4)^{\mathrm{T}}$ como resultado.

Por otra parte, es fácil demostrar por inducción que

$$P^{(t)} = \begin{array}{c} \\ \mathrm{GG} \\ \mathrm{Gg} \\ \mathrm{gg} \end{array} \begin{array}{c} \begin{array}{ccc} \mathrm{GG} & \mathrm{Gg} & \mathrm{gg} \end{array} \\ \begin{pmatrix} \frac{1}{4} + \frac{1}{2^{t+1}} & \frac{1}{2} & \frac{1}{4} - \frac{1}{2^{t+1}} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} - \frac{1}{2^{t+1}} & \frac{1}{2} & \frac{1}{4} + \frac{1}{2^{t+1}} \end{pmatrix} \end{array}$$

para todo $t > 0$, por lo que $\boldsymbol{\pi}$ también es una distribución límite de la cadena de Markov.

Se puede demostrar que toda cadena de Markov con espacio de estados finito tiene al menos una distribución invariante. En cambio, la existencia de distribución límite no está asegurada en general.

El siguiente ejemplo muestra una cadena de Markov sin distribución límite porque su conjunto de estados se divide en dos subconjuntos en los que la cadena se queda «atrapada», por lo que la probabilidad límite de cada estado depende del estado inicial.

Ejemplo: Cadena de Markov sin distribución límite por no ser irreducible

Consideremos la cadena de Markov con espacio de estados $E = \{1,2,3\}$ y matriz de transición

$$P = \begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{array}$$

Es fácil comprobar que $P^{(t)} = P$, para todo $t > 0$, de lo que se deduce que esta cadena de Markov no tiene distribución límite, ya que, por ejemplo

$$\lim_{t \to +\infty} p_{i3}^{(t)} = \begin{cases} 0 & \text{si } i \neq 3, \\ 1 & \text{si } i = 3. \end{cases}$$

Definición 5

1. *Un estado j se dice que es* accesible *desde un estado i si $p_{ij}^{(t)} > 0$, para algún $t \geq 0$.*

2. *Dos estados i y j se dice que* se comunican, *$i \leftrightarrow j$, si cada uno de ellos es accesible desde el otro.*

Proposición 6 *La relación de comunicación es una relación de equivalencia sobre el espacio de estados de la cadena de Markov. Es decir,*

1. *Para todo estado $i \in E$, $i \leftrightarrow i$.*

2. *Para cualesquiera estados $i, j \in E$, si $i \leftrightarrow j$, entonces $j \leftrightarrow i$.*

3. *Para cualesquiera estados $i, j, k \in E$, si $i \leftrightarrow j$ y $j \leftrightarrow k$, entonces $i \leftrightarrow k$.*

Dos estados que se comunican se dice que están en la misma clase *(de equivalencia).*

Definición 7 *Una cadena de Markov se dice que es* irreducible *si posee una única clase comunicante. Es decir, si todos sus estados se comunican entre sí.*

Por su parte, el siguiente ejemplo muestra una cadena de Markov irreducible sin distribución límite porque el estado que toma en cada instante «rota cíclicamente», por lo que la probabilidad límite de cada estado no puede existir.

Ejemplo: Cadena de Markov sin distribución límite por no ser aperiódica

Consideremos la cadena de Markov con espacio de estados $E = \{1,2,3\}$ y matriz de transición

$$P = \begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{array}$$

Es fácil comprobar que, para todo $t > 0$,

$$P^{(t)} = \begin{cases} \left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{smallmatrix}\right) & \text{si } t \text{ es de la forma } 3t' + 1, \\ \left(\begin{smallmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{smallmatrix}\right) & \text{si } t \text{ es de la forma } 3t' + 2, \\ \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right) & \text{si } t \text{ es de la forma } 3t' \end{cases}$$

Por lo tanto, esta cadena de Markov no tiene distribución límite, ya que no existe $\lim_{t\to+\infty} p_{ij}^{(t)}$ para ningún $i, j \in E$.

Definición 8 *Un estado i se dice que tiene periodo n si*

$$n = \mathrm{mcd}\{t > 0 \mid p_{ii}^{(t)} > 0\}$$

Si $n > 1$ el estado se dice que es periódico *y si $n = 1$ se dice que es* aperiódico.

Proposición 9

1. *Todos los estados de una misma clase comunicante tienen el mismo periodo.*
2. *En una cadena de Markov irreducible todos los estados tienen el mismo periodo.*

La irreducibilidad y aperiodicidad de una cadena de Markov con espacio de estados finito son condiciones suficientes para que tenga distribución límite.

Definición 10 *La variable aleatoria $\tau(i)$ que dada una cadena de Markov $\{X_t\}_{t\in\mathbb{Z}_{\geq 0}}$ cuenta el mínimo número de transiciones necesarias para alcanzar el estado i se define como*

$$\tau(i) = \inf\{t \geq 1 \mid X_t = i\}$$

donde $\tau(i) = +\infty$ si el conjunto anterior es el conjunto vacío.

Teorema 11 *Sea $\boldsymbol{X}$ una cadena de Markov irreducible y aperiódica con espacio de estados finito $\boldsymbol{E}$. Entonces,*

1. *$\boldsymbol{X}$ tiene una única distribución invariante $\boldsymbol{\pi} = (\pi_i)_{i\in E}$.*
2. *$\boldsymbol{\pi}$ es distribución límite de $\boldsymbol{X}$.*
3. *$\pi_i = \frac{1}{m_i}$ para todo $i \in E$, donde $m_i = \mathbb{E}[\tau(i) \mid X_0 = i]$.*

En una cadena de Markov con espacio de estados infinito numerable, sin embargo, no basta con esas dos condiciones, ya que si los estados no se «revisitan con la suficiente asiduidad», entonces su probabilidad límite será nula.

Ejemplo: Cadena de Markov sin distribución límite por no ser positivo recurrente

Consideremos un paseo aleatorio en $\mathbb{Z}$, es decir, una cadena de Markov con espacio de estados $E = \mathbb{Z}$ y distribución de transición

$$\mathbb{P}(X_{t+1} = j \mid X_t = i) = \begin{cases} p & \text{si } j = i + 1, \\ 1 - p & \text{si } j = i - 1, \\ 0 & \text{en otro caso.} \end{cases}$$

con $p \in [0, 1]$.

Si $\pi = (\pi_i)_{i \in \mathbb{Z}}$ es una distribución invariante de la cadena de Markov, entonces es fácil establecer que debe verificarse

$$\pi_i = p\pi_{i-1} + (1 - p)\pi_{i+1} \qquad \text{para todo } i \in \mathbb{Z} \tag{1}$$

En el caso $p \neq 1/2$, la solución general a la ecuación (1) viene dada por

$$\pi_i = a + b\left(\frac{p}{1-p}\right)^i \qquad \text{para todo } i \in \mathbb{Z}$$

que para ningún valor de a y b cumple $\pi_i \in [0, 1]$, para todo $i \in \mathbb{Z}$, y $\sum_{i \in \mathbb{Z}} \pi_i = 1$, luego la cadena de Markov no puede tener distribución invariante. La explicación intuitiva es que el paseo está sesgado a la derecha (si $p > 1/2$) o a la izquierda (si $p < 1/2$) y, en consecuencia, la probabilidad de no volver a visitar un estado es mayor que 0.

En el caso $p = 1/2$, la solución general a la ecuación (1) viene dada por

$$\pi_i = a + bi \qquad \text{para todo } i \in \mathbb{Z}$$

que para ningún valor de a y b cumple $\pi_i \in [0, 1]$, para todo $i \in \mathbb{Z}$, y $\sum_{i \in \mathbb{Z}} \pi_i = 1$, luego la cadena de Markov no puede tener distribución invariante. La explicación intuitiva es que, aunque la probabilidad de volver a visitar un estado es 1, el promedio de pasos requerido para ello es infinito.

Cualquier distribución límite debe ser, necesariamente, distribución invariante, de lo que se deduce que ningún paseo aleatorio en $\mathbb{Z}$ tiene distribución límite.

Definición 12

1. *Un estado i se dice que es* recurrente *si* $\mathbb{P}(\tau(i) < +\infty \mid X_0 = i) = 1$, *en caso contrario se dice que es* transitorio.

2. *Un estado recurrente i se dice que es* positivo recurrente *si* $\mathbb{E}[\tau(i) \mid X_0 = i] < +\infty$, *en caso contrario se dice que es* cero recurrente.

Proposición 13

1. *La recurrencia (positiva) y la transitoriedad son propiedades de clase.*
2. *En una cadena de Markov con espacio de estados finito todo estado recurrente es positivo recurrente.*
3. *Todos los estados de una cadena de Markov irreducible con espacio de estados finito son positivo recurrentes.*

Teorema 14 *Sea $\boldsymbol{X}$ una cadena de Markov irreducible y aperiódica con espacio de estados infinito numerable $\boldsymbol{E}$. Entonces ocurre uno de los siguientes casos:*

- *Todos los estados son transitorios y*

$$\lim_{t \to +\infty} p_{ij}^{(t)} = 0, \quad \forall i, j \in E$$

- *Todos los estados son cero recurrentes y*

$$\lim_{t \to +\infty} p_{ij}^{(t)} = 0, \quad \forall i, j \in E$$

- *Todos los estados son positivo recurrentes y*
    1. *$\boldsymbol{X}$ tiene una única distribución invariante $\boldsymbol{\pi} = (\pi_i)_{i \in E}$.*
    2. *π es distribución límite de $\boldsymbol{X}$.*
    3. *$\pi_i = \frac{1}{m_i}$ para todo $i \in E$, donde $m_i = \mathbb{E}[\tau(i) \mid X_0 = i]$.*

Las cadenas de Markov reversibles proporcionan una vía para construir cadenas de Markov con distribuciones invariantes prefijadas.

Definición 15 *Una cadena de Markov $\{X_t\}_{t \in \mathbb{Z}_{\geq 0}}$ estacionaria se dice que es* reversible *si*

$$\mathbb{P}(X_{t+1} \mid X_{t+2} = i) = \mathbb{P}(X_{t+1} \mid X_t = i)$$

para todo $i \in E$.

Ejemplo: Cadena de Markov reversible

Consideremos la cadena de Markov con espacio de estados $E = \{1, 2\}$ y matriz de transición

$$P = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} 0{,}9 & 0{,}1 \\ 0{,}3 & 0{,}7 \end{pmatrix} \end{array}$$

Es fácil comprobar que $\boldsymbol{\pi} = (3/4 \;\; 1/4)^{\mathrm{T}}$ es la distribución invariante de la cadena de Markov. Entonces, asumiendo que $X_0 \sim \boldsymbol{\pi}$, se tiene que

$$\begin{aligned}
\mathbb{P}(X_{t+1} = 1 \mid X_{t+2} = 1) &= \frac{\mathbb{P}(X_{t+2} = 1 \mid X_{t+1} = 1)\mathbb{P}(X_{t+1} = 1)}{\mathbb{P}(X_{t+2} = 1)} = \frac{0{,}9 \times 3/4}{3/4} = 0{,}9 \\
&= \mathbb{P}(X_{t+1} = 1 \mid X_t = 1) \\
\mathbb{P}(X_{t+1} = 2 \mid X_{t+2} = 1) &= \frac{\mathbb{P}(X_{t+2} = 1 \mid X_{t+1} = 2)\mathbb{P}(X_{t+1} = 2)}{\mathbb{P}(X_{t+2} = 1)} = \frac{0{,}3 \times 1/4}{3/4} = 0{,}1 \\
&= \mathbb{P}(X_{t+1} = 2 \mid X_t = 1) \\
\mathbb{P}(X_{t+1} = 1 \mid X_{t+2} = 2) &= \frac{\mathbb{P}(X_{t+2} = 2 \mid X_{t+1} = 1)\mathbb{P}(X_{t+1} = 1)}{\mathbb{P}(X_{t+2} = 2)} = \frac{0{,}1 \times 3/4}{1/4} = 0{,}3 \\
&= \mathbb{P}(X_{t+1} = 1 \mid X_t = 2) \\
\mathbb{P}(X_{t+1} = 2 \mid X_{t+2} = 2) &= \frac{\mathbb{P}(X_{t+2} = 2 \mid X_{t+1} = 2)\mathbb{P}(X_{t+1} = 2)}{\mathbb{P}(X_{t+2} = 2)} = \frac{0{,}7 \times 1/4}{1/4} = 0{,}7 \\
&= \mathbb{P}(X_{t+1} = 2 \mid X_t = 2)
\end{aligned}$$

Definición 16 *La matriz de transición P de una cadena de Markov se dice que está en* balance detallado *con una distribución de probabilidad $\boldsymbol{\pi}$ si*

$$\pi_i p_{ij} = \pi_j p_{ji}, \quad \forall i, j \in E$$

Teorema 17 *Sean $\boldsymbol{\pi}$ una distribución de probabilidad sobre E y $\boldsymbol{X}$ una cadena de Markov cuya matriz de transición está en balance detallado con $\boldsymbol{\pi}$. Entonces:*

1. *$\boldsymbol{\pi}$ es la distribución invariante de $\boldsymbol{X}$.*
2. *Tomando $\boldsymbol{\pi}$ como distribución inicial, $\boldsymbol{X}$ es reversible.*

Luego, dada $\boldsymbol{\pi}$ una distribución de probabilidad sobre E, para construir una cadena de Markov con distribución invariante $\boldsymbol{\pi}$ basta construirla de tal forma que su matriz de transición esté en balance detallado con $\boldsymbol{\pi}$.

Es importante hacer notar, sin embargo, que la reversibilidad es únicamente una condición suficiente.

Ejemplo: Cadena de Markov no reversible, pero con distribución invariante

Consideremos la cadena de Markov $\boldsymbol{X}$ con espacio de estados $E = \{1, 2, 3\}$ y matriz de transición

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 0 & 0{,}8 & 0{,}2 \\ 0{,}2 & 0 & 0{,}8 \\ 0{,}8 & 0{,}2 & 0 \end{pmatrix} \end{array}$$

Es fácil comprobar que $\boldsymbol{\pi} = (1/3 \;\; 1/3 \;\; 1/3)^{\mathrm{T}}$ es distribución invariante de la cadena de Markov, pero que, sin embargo, no está en balance detallado con P.

De hecho, si consideramos la cadena de Markov «de tiempo inverso» $\widetilde{\boldsymbol{X}}$, su matriz de transición es $\tilde{P} = (\tilde{p}_{ij})$, con $\tilde{p}_{ij} = \mathbb{P}(X_{t+1} = j \mid X_{t+2} = i)$. Entonces, tomando $\boldsymbol{\pi}$ como distribución inicial, se puede verificar que $\tilde{P} = P^{\mathrm{T}}$. Esto quiere decir que $\boldsymbol{X}$ recorre cíclicamente los estados favoreciendo el orden creciente, mientras que $\widetilde{\boldsymbol{X}}$ lo hace favoreciendo el orden decreciente y, por tanto, son totalmente distinguibles.

Finalmente, para las cadenas de Markov también se tiene una ley fuerte de los grandes números, que es la base del Método de Montecarlo por Cadenas de Markov (MCMC, *Markov Chain Monte Carlo*).

Teorema 18 *Sean $\{X_t\}_{t \in \mathbb{Z}_{\geq 0}}$ una cadena de Markov irreducible con distribución estacionaria $\boldsymbol{\pi}$ y $g\colon E \to \mathbb{R}$ una función acotada. Entonces,*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{t=0}^{n} g(x_t) = \sum_{i \in E} g(i)\pi_i = \mathbb{E}_{\boldsymbol{\pi}}[g(X)]$$

con probabilidad 1, para cualquier realización $(x_t)_{t=0,\dots,n}$ de la cadena de Markov.

E continuo

Asumimos $E = \mathbb{R}^d$ y que la distribución de transición de la cadena de Markov se expresa mediante un *núcleo de transición* $K\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_{\geq 0}$, de manera que

$$\mathbb{P}(X_{t+1} \in A \mid X_t = x) = \int_A K(x, y)\, dy$$

para todo conjunto de Borel $A \subseteq \mathbb{R}^d$. Es decir, $K(x, \cdot)$ es justamente la función de densidad condicionada de X_{t+1} dado $X_t = x$.

Ejemplo: Paseo aleatorio en $\mathbb{R}^d$

Un paseo aleatorio es una cadena de Markov con espacio de estados $E = \mathbb{R}^d$ y en el que la transición se realiza como

$$\boldsymbol{X}_{t+1} = \boldsymbol{X}_t + \boldsymbol{Z}_t, \qquad \text{con } \boldsymbol{Z}_t \overset{\text{i. i. d.}}{\sim} g$$

con g una función de densidad sobre E prefijada de antemano.

La función de distribución de la transición del paseo aleatorio es, por tanto,

$$\begin{aligned} F(\boldsymbol{y} \mid \boldsymbol{x}) &= \mathbb{P}\big(\boldsymbol{X}_{t+1} \leq \boldsymbol{y} \mid \boldsymbol{X}_t = \boldsymbol{x}\big) \\ &= \mathbb{P}\big(\boldsymbol{X}_t + \boldsymbol{Z}_t \leq \boldsymbol{y} \mid \boldsymbol{X}_t = \boldsymbol{x}\big) \\ &= \mathbb{P}\big(\boldsymbol{Z}_t \leq \boldsymbol{y} - \boldsymbol{x}\big) \\ &= F_g(\boldsymbol{y} - \boldsymbol{x}) \end{aligned}$$

de lo que se deduce que el núcleo de transición del paseo aleatorio es

$$K(\boldsymbol{x}, \boldsymbol{y}) = \frac{\partial^d F(\boldsymbol{y} \mid \boldsymbol{x})}{\partial y_1 \cdots \partial y_d} = \frac{\partial^d F_g(\boldsymbol{y} - \boldsymbol{x})}{\partial y_1 \cdots \partial y_d} = g(\boldsymbol{y} - \boldsymbol{x})$$

Para las cadenas de Markov con espacio de estados continuo se obtienen resultados análogos a los obtenidos para las cadenas de Markov con espacio de estados finito, con la diferencia que los conceptos deben referirse a subconjuntos de estados, en lugar de a estados puntuales.

Proposición 19 *La transición en t unidades de tiempo de la cadena de Markov a partir del estado inicial x_0 viene dada por*

$$\mathbb{P}^{(t)}_{x_0}(A) = \mathbb{P}(X_t \in A \mid X_0 = x_0) = \int_A K^{(t)}(x_0, x_t)\, dx_t$$

para todo conjunto de Borel $A \subseteq \mathbb{R}^d$, donde

$$K^{(t)}(x_0, x_t) = \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} K(x_0, x_1) \cdots K(x_{t-1}, x_t)\, dx_{t-1} \cdots dx_1$$

Las ecuaciones de Chapman-Kolmogorov *establecen que*

$$K^{(t+t')}(x, y) = \int_{\mathbb{R}^d} K^{(t)}(x, z) K^{(t')}(z, y)\, dz$$

Definición 20 *Una distribución de probabilidad π con función de densidad f_π es* invariante *o* estacionaria *para el núcleo de transición K si*

$$f_\pi(y) = \int_{\mathbb{R}^d} f_\pi(x) K(x, y)\, dx, \quad \forall y \in \mathbb{R}^d$$

Si $\{X_t\}_{t\in\mathbb{Z}_{\geq 0}}$ es una cadena de Markov con distribución invariante π y tomamos π como su distribución inicial, entonces $X_t \sim \pi$ para todo t $\in \mathbb{Z}_{\geq 0}$.

Definición 21 *Dada una medida φ sobre $\mathbb{R}^d$, una cadena de Markov se dice que es φ-irreducible si para todo conjunto de Borel $A \subseteq \mathbb{R}^d$ tal que $\varphi(A) > 0$ y para todo $x \in \mathbb{R}^d$ existe un $t > 0$ tal que $\mathbb{P}^{(t)}_x(A) > 0$.*

Definición 22 *Una cadena de Markov con distribución invariante π se dice que es* aperiódica *si no existen $n > 1$ y subconjuntos de Borel disjuntos $A_1, \dots, A_n \subseteq \mathbb{R}^d$ tales que:*

1. *$\pi(A_i) > 0$, para todo $i = 1, \dots, n$.*
2. *$\mathbb{P}(X_{t+1} \in A_{i+1} \mid X_t \in A_i) = 1$, para todo $i = 1, \dots, n-1$ y $\mathbb{P}(X_{t+1} \in A_1 \mid X_t \in A_n) = 1$.*

Definición 23 *El* número de visitas *de una cadena de Markov a un conjunto $A \subseteq \mathbb{R}^d$ viene dado por*

$$V(A) = \sum_{t=0}^{+\infty} \mathbb{1}_A(X_t)$$

Definición 24

1. *Un conjunto de Borel* $A \subseteq \mathbb{R}^d$ *se dice que es* recurrente *si para todo* $x \in A$

$$\mathbb{E}[V(A) \mid X_0 = x] = +\infty$$

2. *Una cadena de Markov se dice que es* recurrente *si*

   a) *La cadena es* φ*-irreducible para alguna medida* φ*.*
   b) *Todo conjunto de Borel* $A \subseteq \mathbb{R}^d$ *con* $\varphi(A) > 0$ *es recurrente.*

Definición 25

1. *Un conjunto de Borel* $A \subseteq \mathbb{R}^d$ *se dice que es* Harris-recurrente *si para todo* $x \in A$

$$\mathbb{P}\big(V(A) = +\infty \mid X_0 = x\big) = 1$$

2. *Una cadena de Markov se dice que es* Harris-recurrente *si*

   a) *La cadena es* φ*-irreducible para alguna medida* φ*.*
   b) *Todo conjunto de Borel* $A \subseteq \mathbb{R}^d$ *con* $\varphi(A) > 0$ *es Harris-recurrente.*

Proposición 26 *Toda cadena de Markov* φ*-irreducible que tenga una distribución invariante es recurrente.*

La irreducibilidad y aperiodicidad de una cadena de Markov con espacio de estados continuo son condiciones suficientes para que tenga distribución límite.

Definición 27 *La* distancia de variación total *entre dos distribuciones de probabilidad* π_1 *y* π_2 *se define como*

$$\|\pi_1 - \pi_2\| = \sup_{A \subseteq \mathbb{R}^d \text{ de Borel}} |\pi_1(A) - \pi_2(A)|$$

Teorema 28 *Si una cadena de Markov es* φ*-irreducible y aperiódica y tiene distribución invariante* π*, entonces para* π*-casi todo* $x \in \mathbb{R}^d$ *se tiene que*

$$\lim_{t \to +\infty} \left\| \mathbb{P}_x^{(t)} - \pi \right\| = 0$$

En particular, $\lim_{t \to +\infty} \mathbb{P}_x^{(t)}(A) = \pi(A)$*, para todo conjunto de Borel* $A \subseteq \mathbb{R}^d$*.*
Si la cadena es Harris-recurrente, entonces el resultado se da para todo $x \in \mathbb{R}^d$*.*

Las cadenas de Markov con distribuciones invariantes prefijadas también se obtendrán a partir de cadenas de Markov reversibles.

Definición 29 *Un núcleo de transición K se dice que está en* balance detallado *con una distribución π con función de densidad f_π si*

$$f_\pi(x)K(x,y) = f_\pi(y)K(y,x)$$

para todo $x, y \in \mathbb{R}^d$.

Teorema 30 *Sea $\boldsymbol{X}$ una cadena de Markov cuyo núcleo de transición está en balance detallado con una distribución de probabilidad π. Entonces,*

1. *π es la única distribución invariante de $\boldsymbol{X}$.*
2. *Tomando π como distribución inicial, $\boldsymbol{X}$ es reversible.*

Finalmente, la ley fuerte de los grandes números también se cumple para las cadenas de Markov con espacio de estados continuo.

Teorema 31 *Sean $\{X_t\}_{t\in\mathbb{Z}_{\geq 0}}$ una cadena de Markov φ-irreducible con distribución estacionaria π y $g\colon \mathbb{R}^d \to \mathbb{R}$ una función π-integrable. Entonces, para π-casi todo $x_0 \in \mathbb{R}^d$*

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{t=0}^{n} g(x_t) = \int_{\mathbb{R}^d} g(x) f_\pi(x)\, dx = \mathbb{E}_\pi[g(X)]$$

con probabilidad 1 para cualquier realización $\{x_t\}_{t=0,\dots,n}$ de la cadena de Markov.

Si la cadena es Harris-recurrente, entonces el resultado se verifica para todo $x_0 \in \mathbb{R}^d$.