Lección n° 2: Fundamentos de las cadenas de Markov AMARUN 2025

El estudio de las cadenas de Markov es central en la teoría de los procesos estocásticos. Los resultados fundamentales de este capítulo pueden encontrarse en textos clásicos como los de [1] y [4], pero también en un recopilatorio mejor condesando en [3] y otros mencionados en la primera lección.

En esta segunda lección, se establecen los conceptos básicos: el espacio de estados, el núcleo de transición, las propiedades de homogeneidad, la clasificación de los estados y la noción de irreductibilidad. Se presentan ejemplos representativos, como la caminata aleatoria y el proceso de Bernoulli, que ilustran la variedad de comportamientos posibles.

1. Cadenas de Markov a tiempo y espacio discreto

1.1. Sucesiones recurrentes aleatorias y núcleos de transición

Sea $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias con valores en un conjunto a lo más numerable E. La ley de la sucesión es una medida de probabilidad sobre $E^{\mathbb{N}}$ dotado de la tribu generada por los cilindros. El teorema de Carathéodory-Kolmogorov indica que está caracterizada por sus marginales de dimensión finita, es decir, por las leyes de los vectores aleatorios $(X_0, \ldots, X_m)$ para todo $m \in \mathbb{N}$ [3]. Ahora bien, E es a lo más numerable, y por tanto la ley de $(X_0, \ldots, X_m)$ está caracterizada por el conocimiento de $\mathbb{P}(X_m = x_m, \ldots, X_0 = x_0)$ para todos $x_0, \ldots, x_m$ en E. Por condicionamientos sucesivos, podemos escribir

$$\mathbb{P}(X_m = x_m, \ldots, X_0 = x_0) = \prod_{n=0}^{m-1} \mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n \ldots, X_0 = x_0)\, \mathbb{P}(X_0 = x_0).$$

Si un evento de la forma $\{X_n = x_n \ldots, X_0 = x_0\}$ es negligeable, lo mismo ocurre con

$$\{X_m = x_m, \ldots, X_0 = x_0\}.$$

Así, en todo lo que sigue, estos casos particulares se omiten sistemáticamente por brevedad.

El índice n de la sucesión $(X_n)_{n\in\mathbb{N}}$ se interpreta como un tiempo. La variable X_n representa la posición espacial en el instante n, la tribu $\sigma(X_0, \ldots, X_{n-1})$ representa su pasado mientras que la tribu $\sigma(X_{n+1}, X_{n+2}, \ldots)$ representa su futuro.

Definición 1.1 (Cadena de Markov)**.** Se dice que una sucesión $(X_n)_{n\in\mathbb{N}}$ de v.a. con valores en un conjunto a lo más numerable E es una **cadena de Markov** de espacio de estados E cuando para todo $n \in \mathbb{N}$ y $x_0, \ldots, x_{n+1} \in E$,

$$\mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n, \ldots, X_0 = x_0) = \mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n).$$

Esto también se escribe

$$\mathcal{L}(X_{n+1} \mid X_n, \ldots, X_0) = \mathcal{L}(X_{n+1} \mid X_n) \quad \text{para todo } n \in \mathbb{N}.$$

Se dice que la cadena es **homogénea** (en el tiempo) cuando para todo $n \in \mathbb{N}$ y todos $x, y \in E$,

$$\mathbb{P}(X_{n+1} = y \mid X_n = x) = \mathbb{P}(X_1 = y \mid X_0 = x).$$

Esto también se escribe

$$\mathcal{L}(X_{n+1} \mid X_n = x) = \mathcal{L}(X_1 \mid X_0 = x) \quad \text{para todos } n \in \mathbb{N} \text{ y } x \in E.$$

Las cadenas de Markov modelizan procesos estocásticos sin memoria. Son indispensables, presentes prácticamente en todas partes: informática, física, estadística, finanzas, ingeniería, biología, ciencias humanas y sociales, etc. Desde el punto de vista matemático, su aparente simplicidad conceptual es engañosa. Este ámbito plantea, aún hoy en día, arduas cuestiones abiertas y una investigación internacional muy viva y de alto nivel [2, 3].

Ejemplo 1.2 (Sucesiones recurrentes aleatorias)**.** Sea X_0 una variable aleatoria con valores en un conjunto a lo más numerable E.
Sea $(\varepsilon_{n+1})_{n\in\mathbb{N}}$ una sucesión de v.a. independientes con valores en un espacio medible F, independiente de X_0. Para todo $n \in \mathbb{N}$, sea $g_n : E \times F \to E$ medible.
La sucesión recurrente aleatoria $(X_n)_{n\in\mathbb{N}}$ definida por $X_{n+1} = g_n(X_n, \varepsilon_{n+1})$ es una cadena de Markov sobre E. Es homogénea en el tiempo si $\mathcal{L}(\varepsilon_{n+1})$ y g_n no dependen del tiempo n: $(\varepsilon_{n+1})_{n\in\mathbb{N}}$ es i.i.d. y $g_n = g$. En efecto, como $(X_0, \ldots, X_n)$ es $\sigma(X_0, \varepsilon_1, \ldots, \varepsilon_n)$-medible, entonces es independiente de ε_{n+1},

$$\mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n, \ldots, X_0 = x_0) = \mathbb{P}(g(x_n, \varepsilon_{n+1}) = x_{n+1}).$$

Ejemplo 1.3 (Caminata aleatoria sobre $\mathbb{Z}$)**.** Sea $\mathrm{g}(x, \varepsilon) = x + \varepsilon$ de donde $X_{n+1} = X_0 + \varepsilon_1 + \cdots + \varepsilon_{n+1}$.

Ejemplo 1.4 (Caminata aleatoria sobre el grupo simétrico $\mathbb{S}_n$)**.** Sea $g(x, \varepsilon) = x\varepsilon$ de donde $X_{n+1} = X_0\varepsilon_1 \cdots \varepsilon_{n+1}$.

Recíprocamente, toda cadena de Markov es una sucesión recurrente aleatoria de este tipo (teorema 5.5.1). Para un entero $r \geq 1$, consideremos ahora una sucesión recurrente aleatoria $(X_n)_{n\in\mathbb{N}}$ de orden r en E, dada por $X_{n+1} = f(X_n, \ldots, X_{n-r}, \varepsilon_{n+1})$, donde $(\varepsilon_{n+1})_{n\geq r}$ es una sucesión de v.a. i.i.d. independiente de $X_0, \ldots, X_{r-1}$. Se dice que se trata de una cadena de Markov de orden r, es decir, que $((X_{n+r-1}, \ldots, X_n))_{n\in\mathbb{N}}$ es una CM sobre E^r.

Núcleos de transición

Definición 1.5. Sea (X_n) una cadena de Markov de espacio de estados E. Su **núcleo de transición** $\mathbf{P} : E \times E \to [0, 1]$ se define por

$$\mathbf{P}(x, y) = \mathbb{P}(X_1 = y \mid X_0 = x), \quad x, y \in E.$$

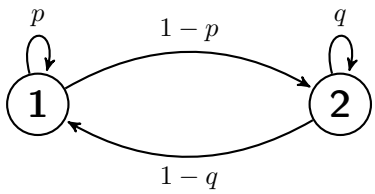

Figura 1: Cadena de Markov de dos estados: transiciones con probabilidades p, q, $1-p$, $1-q$.

Por homogeneidad se tiene que $\mathbb{P}(X_{n+1} = y \mid X_n = x) = \mathbf{P}(x, y)$ para todo $n \in \mathbb{N}$ y todos $x, y \in E$. Si $v = \mathcal{L}(X_0)$, la definición 1.1 da para todo $n \in \mathbb{N}$ y todos $x_0, \ldots, x_n$ en E,

$$\mathbb{P}(X_n = x_n, \ldots, X_0 = x_0) = v(x_0) \prod_{k=0}^{n-1} \mathbf{P}(x_k, x_{k+1}).$$

Así, la ley de la sucesión $(X_n)_{n\in\mathbb{N}}$ está completamente caracterizada por la ley inicial $v = \mathcal{L}(X_0)$ y el núcleo de transición $\mathbf{P}$. La notación $\mathrm{CM}(E, v, \mathbf{P})$ significa «cadena de Markov de espacio de estados E, de ley inicial v, y de núcleo de transición $\mathbf{P}$».

Ejemplo 1.6. El núcleo de transición de la cadena de Markov $(X_n)_{n\in\mathbb{N}}$ del ejemplo 1.2 es $\mathbf{P}(x, y) = \mathbb{P}(g(x, \varepsilon_1) = y)$.

La noción de núcleo puede aislarse del concepto de cadena de Markov. Una aplicación $\mathbf{P} : E \times E \to \mathbb{R}$ es un **núcleo de transición** sobre E si y sólo si para todo x, $\mathbf{P}(x, \cdot)$, vista como $\sum_{y\in E} \mathbf{P}(x, y)\delta_y$, es una ley sobre E:

$$0 \leq \mathbf{P}(x, y) \leq 1 \quad \text{y} \quad \sum_{z\in E} \mathbf{P}(x, z) = 1.$$

A fuerza de numerar los elementos de E, un núcleo de transición $\mathbf{P}$ puede verse como una matriz cuadrada, infinita si E es infinito, llamada **matriz de transición**, cuyas filas son leyes sobre E.

$$E = \{x_0, x_1, x_2, \ldots\} \quad \text{y} \quad \mathbf{P} = \begin{pmatrix} \mathbf{P}(x_0, x_0) & \mathbf{P}(x_0, x_1) & \mathbf{P}(x_0, x_2) & \cdots \\ \mathbf{P}(x_1, x_0) & \mathbf{P}(x_1, x_1) & \mathbf{P}(x_1, x_2) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Sean p_0 y p_1 en $[0,1]$. El núcleo de transición $\mathbf{P}$ con dos estados es el núcleo sobre $E = \{0,1\}$ que se escribe

$$\mathbf{P}(0,1) = 1 - \mathbf{P}(0,0) = p_0 \quad \text{y} \quad \mathbf{P}(1,0) = 1 - \mathbf{P}(1,1) = p_1.$$

Vista como una matriz 2×2, el núcleo $\mathbf{P}$ se escribe

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}(0,0) & \mathbf{P}(0,1) \\ \mathbf{P}(1,0) & \mathbf{P}(1,1) \end{pmatrix} = \begin{pmatrix} 1-p_0 & p_0 \\ p_1 & 1-p_1 \end{pmatrix}.$$

Una cadena de Markov con este núcleo es una cadena con dos estados. Es útil representar un núcleo de transición con un **grafo de transiciones**, como en la figura 1.1.

Estructura del conjunto de los núcleos

Si $\mathbf{P}$ y $\mathbf{Q}$ son dos núcleos de transición sobre E, y si α y β son dos reales positivos tales que $\alpha+\beta = 1$, entonces $\alpha\mathbf{P}(x,y)+\beta\mathbf{Q}(x,y)$ define un núcleo de transición sobre E denotado $\alpha\mathbf{P}+\beta\mathbf{Q}$. Por otra parte, $\sum_{z\in E}\mathbf{P}(x,z)\mathbf{Q}(z,y)$ define también un núcleo de transición sobre E denotado $\mathbf{PQ}$. Así, el conjunto de los núcleos de transición sobre E es convexo, y constituye un semigrupo no abeliano para el producto, de elemento neutro $\mathbf{I}$ definido por $\mathbf{I}(x,y) = \mathbb{1}_{x=y}$. Por otra parte, las matrices de permutación de E forman un grupo, incluido en este semigrupo, cuya acción a derecha, a izquierda, o por conjugación sobre el semigrupo corresponde a la invarianza por permutación de las filas y/o columnas.

Núcleo iterado y escritura matricial

Sea $(X_n)_{n\in\mathbb{N}}$ una $\mathrm{CM}(E,v,\mathbf{P})$. La aplicación $\mathbf{P}^n : E\times E \to [0,1]$ definida por

$$\mathbf{P}^n(x,y) := \mathbb{P}\left(X_n = y \mid X_0 = x\right)$$

es un núcleo de transición sobre E. En particular $\mathbf{P}^1 = \mathbf{P}$, y $\mathbf{P}^0 = \mathbf{I}$. Para todos $n \geq 1$ y x_0, x_n en E,

$$\mathbf{P}^n\left(x_0,x_n\right) = \mathbb{P}\left(X_n = x_n \mid X_0 = x_0\right) = \sum_{x_1,\ldots,x_{n-1}\in E} \mathbb{P}\left(X_n = x_n,\ldots,X_1 = x_1 \mid X_0 = x_0\right).$$

La sucesión $x_0,\ldots,x_n$ constituye un camino en E. Por recurrencia sobre n, se tiene

$$\mathbf{P}^n\left(x_0,x_n\right) = \sum_{x_1,\ldots,x_{n-1}\in E} \mathbf{P}\left(x_0,x_1\right)\cdots\mathbf{P}\left(x_{n-1},x_n\right) = \sum_{x\in E}\mathbf{P}^{n-1}\left(x_0,x\right)\mathbf{P}\left(x,x_n\right).$$

Estas fórmulas son las del producto matricial: $\mathbf{P}^n = \mathbf{P}\mathbf{P}^{n-1} = \mathbf{P}^{n-1}\mathbf{P}$.
Para el proceso de Bernoulli del ejemplo 1.8, el núcleo iterado es $\mathbf{P}^n(x,\cdot) = \mathcal{L}\left(B_n \mid B_0 = x\right) = \delta_x * \mathrm{Bin}(n,p)$.
Identificando toda ley v sobre E con un vector fila, y toda función $f : E \to \mathbb{R}$ con un vector columna, la media de f para v se escribe, cuando tiene sentido, $vf = \sum_{x\in E} f(x)v(x)$ (serie si E es infinito). Identificando todo núcleo de transición con una matriz cuadrada infinita, si $(X_n)_{n\in\mathbb{N}}$ es una $\mathrm{CM}(E,v,\mathbf{P})$, entonces $\mathcal{L}\left(X_n\right) = v\mathbf{P}^n$ y $\mathbb{E}\left(f\left(X_n\right)\right) = v\mathbf{P}^n f$, pues

$$\mathbb{E}\left(f\left(X_n\right)\right) = \sum_{y\in E}\mathbb{P}\left(X_n = y\right)f(y) = \sum_{x,y\in E} v(x)\mathbf{P}^n(x,y)f(y) = v\mathbf{P}^n f.$$

La familia $\{\mathbf{P}^n : n\in\mathbb{N}\}$ es un **semigrupo (de Markov)** de operadores lineales que actúan sobre las medidas de probabilidad sobre E. Denotando $v_n := \mathrm{Law}\left(X_n\right) = v\mathbf{P}^n$, la sucesión $(v_n)_{n\in\mathbb{N}}$ se calcula por recurrencia lineal

$$v_0 = v \quad \text{y} \quad v_{n+1} = v_n\mathbf{P}.$$

Las trayectorias $(X_n)_{n\in\mathbb{N}}$ no son ni deterministas ni lineales, pero la evolución de las leyes instantáneas es determinista y lineal. Así, la estructura markoviana de los fenómenos de evolución aleatoria sin memoria sobre E hace aparecer una estructura lineal, no sobre E que no está necesariamente estructurado, sino sobre el conjunto de las leyes sobre E.

Notación condicional

Si $(X_n)_{n\in\mathbb{N}}$ es una cadena de Markov sobre E, entonces para todo $x \in E$, todo $A \subset E$, y toda v.a. Z, se nota

$$\mathbb{P}^x(A) = \mathbb{P}\left(A \mid X_0 = x\right) \quad \text{y} \quad \mathbb{E}^x(Z) = \mathbb{E}\left(Z \mid X_0 = x\right).$$

1.2. Ejemplos de cadenas de Markov

Ejemplo 1.7 (Caminatas aleatorias sobre grupos y sobre grafos)**.** La **caminata aleatoria simple** (CAS) sobre $\mathbb{Z}^d, d \geq 1$, es la sucesión recurrente aleatoria $(X_n)_{n\in\mathbb{N}}$ de $\mathbb{Z}^d$ definida por

$$X_{n+1} = X_n + \varepsilon_{n+1} = X_0 + \varepsilon_1 + \cdots + \varepsilon_n$$

donde $(\varepsilon_{n+1})_{n\in\mathbb{N}}$ es una sucesión de v.a. i.i.d. de ley uniforme sobre $\{\pm e_1, \ldots, \pm e_d\}$ donde $\{e_1, \ldots, e_d\}$ es la base canónica de $\mathbb{R}^d$, y donde X_0 es una v.a. sobre $\mathbb{Z}^d$ independiente de $(\varepsilon_n)_{n\geq 1}$. Es una CM $\left(\mathbb{Z}^d, \mathrm{Law}\left(X_0\right), \mathbf{P}\right)$ de núcleo

$$\mathbf{P}(x,y) = \begin{cases} \frac{1}{2d} & \text{si } x \text{ e } y \text{ son vecinos en } \mathbb{Z}^d, \text{ es decir } |x-y|_1 = 1 \\ 0 & \text{en otro caso.} \end{cases}$$

Esto modela una difusión, como el movimiento browniano (ver Lección 1), que constituye su límite de escala.

Ejemplo 1.8 (Proceso de Bernoulli del juego de cara o cruz)**.** Los resultados sucesivos en un juego de cara o cruz se modelizan por una sucesión $(\varepsilon_{n+1})_{n\in\mathbb{N}}$ de v.a. i.i.d. de ley de Bernoulli $p\delta_1 + (1-p)\delta_0, p \in [0,1]$. El número de ganancias en n partidas es $B_n = \varepsilon_1 + \cdots + \varepsilon_n$. Posando $B_0 := 0$, la sucesión $(B_n)_{n\in\mathbb{N}}$ es una CM $(\mathbb{N}, \delta_0, \mathbf{P})$ llamada **proceso de Bernoulli**, de núcleo de transición

$$\mathbf{P}(x,y) = \begin{cases} p & \text{si } y = x+1 \\ 1-p & \text{si } y = x \\ 0 & \text{en otro caso.} \end{cases}$$

Es una caminata aleatoria sobre el semigrupo conmutativo $(\mathbb{N}, +)$.

1.3. Propiedad de Markov fuerte

Definición 1.9 (Propiedad de Markov débil)**.** Sea $(X_n)_{n\in\mathbb{N}}$ una CM$(E, v, \mathbf{P})$. Para todos $m, n \geq 0$, y todos $A \in E^n$ y $C \in E^m$,

$$\begin{aligned} &\mathbb{P}\left((X_{m+n}, \ldots, X_m) \in C, (X_m, \ldots, X_0) \in A \mid X_m = x\right) \\ &= \mathbb{P}\left((X_{m+n}, \ldots, X_n) \in C \mid X_m = x\right) \mathbb{P}\left((X_n, \ldots, X_0) \in A \mid X_m = x\right) \\ &= \mathbb{P}\left((X_n, \ldots, X_0) \in C \mid X_0 = x\right) \mathbb{P}\left((X_m, \ldots, X_0) \in A \mid X_m = x\right). \end{aligned}$$

Es decir, para todo $m \in \mathbb{N}$, y condicionalmente a $\{X_m = x\}$, la sucesión $(X_{m+n})_{n\in\mathbb{N}}$ es una CM $(E, \delta_x, \mathbf{P})$ independiente de $\sigma\left(X_0, \ldots, X_m\right)$.

Teorema 1.10 (Propiedad de Markov fuerte)**.** *Sea $(X_n)_{n\in\mathbb{N}}$ una* CM(E)*, y $T : \Omega \to \mathbb{N} \cup \{\infty\}$ un tiempo de parada para la filtración natural $(\mathscr{F}_n)_{n\in\mathbb{N}}$. Entonces para todo $x \in E$, sabiendo $\{T < \infty$ y $X_T = x\}$, $(X_{T+n})_{n\in\mathbb{N}}$ es una* CM $(E, \delta_x, \mathbf{P})$*, independiente de $\mathscr{F}_T$.*

Equivalentemente, para todo $A \in \mathscr{F}_T$, todo evento cilíndrico $C \subset E^{\mathbb{N}}$, y todo $x \in E$,

$$\mathbb{P}\left(\left\{(X_{T+n})_{n\in\mathbb{N}} \in C\right\} \cap A \mid T < \infty, X_T = x\right) = \mathbb{P}^x\left((X_n)_{n\in\mathbb{N}} \in C\right) \mathbb{P}\left(A \mid T < \infty, X_T = x\right).$$

Demostración. Para todos $m \in \mathbb{N}, A \in \mathscr{F}_T$, y $x \in E$, $A \cap \{T = m\} \cap \{X_T = x\} \in \mathscr{F}_m$, y por la propiedad de Markov débil,

$$\mathbb{P}\left(\{X_{T+n} = x_n, \ldots, X_T = x_0\} \cap A \cap \{T = m\} \cap \{X_T = x\}\right) = \mathbb{P}\left(X_n = x_n, \ldots, X_0 = x_0 \mid X_0 = x_0\right) \mathbb{P}\left(A \cap \{T = m\} \cap \{X_T = x\}\right)$$

para todos $x_1, \ldots, x_n \in E$ y con $x_0 := x$. Sumando sobre $m \in \mathbb{N}$,

$$\mathbb{P}\left(\{X_{T+n} = x_n, \ldots, X_T = x_0\} \cap A \cap \{T < \infty\} \cap \{X_T = x\}\right) = \mathbb{P}\left(X_n = x_n, \ldots, X_0 = x_0 \mid X_0 = x_0\right) \mathbb{P}\left(A \cap \{T < \infty\} \cap \{X_T = x\}\right).$$

La propiedad deseada se obtiene dividiendo por $\mathbb{P}\left(T < \infty \text{ y } X_T = x\right)$. $\square$

1.4. Tiempos de primer paso, tiempos de alcance, número de visitas, potencial

Definición 1.11 (Tiempos de alcance y número de visitas)**.** Sea $(X_n)_{n\in\mathbb{N}}$ una CM(E) y $F \subset E$. Definimos:

- **Tiempo de primer paso** en F: $T_F^* := \inf\{n \geq 1 : X_n \in F\}$
- **Tiempo de alcance** de F: $T_F := \inf\{n \geq 0 : X_n \in F\}$
- **Número de visitas**: $N_F^* := \sum_{n=1}^{\infty} \mathbb{1}_{\{X_n \in F\}}$, $N_F := \sum_{n=0}^{\infty} \mathbb{1}_{\{X_n \in F\}}$

Cuando $F = \{x\}$, denotamos T_x^*, T_x, N_x^*, N_x.

Definición 1.12 (Potencial de Green)**.** Los números medios de visitas en y partiendo de x están dados por:

$$\mathbf{G}^*(x,y) := \mathbb{E}^x\left(N_y^*\right) = \sum_{n=1}^{\infty} \mathbf{P}^n(x,y), \quad \mathbf{G}(x,y) := \mathbb{E}^x\left(N_y\right) = \sum_{n=0}^{\infty} \mathbf{P}^n(x,y)$$

La probabilidad de alcance: $\rho(x,y) := \mathbb{P}^x\left(T_y^* < \infty\right)$.

Teorema 1.13 (Tiempos de retorno y potencial)**.** *Para una* CM($E, \mathbf{P}$)*:*

(i) Para todos $x, y \in E$, $\mathbb{P}^x\left(N_y = \infty\right) = \mathbb{P}^y\left(N_y = \infty\right)\rho(x,y)$

(ii) Si $\mathbf{G}^*(y,y) < \infty$, entonces para todo $x \in E$, $\mathbf{G}^*(x,y) < \infty$ y $\rho(x,y) = \frac{\mathbf{G}^*(x,y)}{\mathbf{G}(y,y)}$

(iii) Para todos $x, y \in E$ y $\lambda > 0$, $\mathbb{E}^x\left(\mathrm{e}^{-\lambda T_y^*}\right) = \frac{\mathbf{G}_\lambda^*(x,y)}{1+\mathbf{G}_\lambda^*(y,y)}$

Demostración. Para la parte (i) correspondiente a la probabilidad de visitas infinitas, queremos demostrar que:

$$\mathbb{P}^x(N_y = \infty) = \mathbb{P}^y(N_y = \infty)\mathbb{P}^x(T_y^* < \infty).$$

Recordemos que $T_y^* = \inf\{n \geq 1 : X_n = y\}$ es el tiempo de primer paso a y (excluyendo el tiempo 0) y que $N_y = \sum_{n=0}^{\infty} \mathbb{1}_{\{X_n = y\}}$ es el número total de visitas a y. La clave es usar la **propiedad de Markov fuerte** en el tiempo de parada T_y^*. Observemos que:

Si $T_y^* = \infty$, entonces nunca se visita y después del tiempo 0, por lo que N_y es finito. Por tanto, el evento $\{N_y = \infty\}$ sólo puede ocurrir cuando $T_y^* < \infty$:

$$\mathbb{P}^x(N_y = \infty) = \mathbb{P}^x(N_y = \infty, T_y^* < \infty)$$

Ahora, descomponemos las visitas usando T_y^*:

$$N_y = \text{visitas antes de } T_y^* + \text{visitas después de } T_y^*$$

Pero antes de T_y^*, sólo podría haber una visita en el tiempo 0 si $x = y$. Para $n \geq 1$, no hay visitas a y antes de T_y^* por definición. Así, en el evento $\{T_y^* < \infty\}$, tenemos:

$$N_y = \mathbb{1}_{\{X_0 = y\}} + \sum_{n=T_y^*}^{\infty} \mathbb{1}_{\{X_n = y\}}$$

Pero si $T_y^* < \infty$, entonces $X_{T_y^*} = y$, y el número de visitas después de T_y^* determina si N_y es infinito. Por la propiedad de Markov fuerte, condicional a $\{T_y^* < \infty, X_{T_y^*} = y\}$, el proceso $(X_{T_y^*+n})_{n\geq 0}$ es una cadena de Markov con la misma distribución que empezar en y, independiente del pasado. Por lo tanto:

$$\mathbb{P}^x\left(\sum_{n=T_y^*}^{\infty} \mathbb{1}_{\{X_n = y\}} = \infty \mid T_y^* < \infty, X_{T_y^*} = y\right) = \mathbb{P}^y(N_y = \infty)$$

Finalmente,

$$\begin{aligned}
\mathbb{P}^x(N_y = \infty) &= \mathbb{P}^x(N_y = \infty, T_y^* < \infty) \\
&= \mathbb{E}^x\left[\mathbb{1}_{\{T_y^* < \infty\}}\mathbb{1}_{\{N_y = \infty\}}\right] \\
&= \mathbb{E}^x\left[\mathbb{1}_{\{T_y^* < \infty\}}\mathbb{P}^y(N_y = \infty)\right] \\
&= \mathbb{P}^y(N_y = \infty)\mathbb{P}^x(T_y^* < \infty)
\end{aligned}$$

Queremos demostrar (ii), observemos que si $T_y^* = \infty$, entonces $N_y^* = 0$. Si $T_y^* < \infty$, entonces para $n \geq 1$:

- Para $n < T_y^*$: $\mathbb{1}_{\{X_n=y\}} = 0$ (por definición de T_y^*)
- Para $n \geq T_y^*$: pueden haber visitas a y

Por tanto, podemos escribir:

$$N_y^* = \mathbb{1}_{\{T_y^*<\infty\}} \sum_{n=T_y^*}^{\infty} \mathbb{1}_{\{X_n=y\}}.$$

Tomando esperanza:

$$\mathbf{G}^*(x,y) = \mathbb{E}^x[N_y^*] = \mathbb{E}^x\left[\mathbb{1}_{\{T_y^*<\infty\}} \sum_{n=T_y^*}^{\infty} \mathbb{1}_{\{X_n=y\}}\right].$$

Por la propiedad de Markov fuerte, condicional a $\{T_y^* < \infty, X_{T_y^*} = y\}$:

$$\mathbb{E}^x\left[\sum_{n=T_y^*}^{\infty} \mathbb{1}_{\{X_n=y\}} \mid T_y^* < \infty, X_{T_y^*} = y\right] = \mathbb{E}^y[N_y] = \mathbf{G}(y,y).$$

Por lo tanto:

$$\begin{aligned}\mathbf{G}^*(x,y) &= \mathbb{E}^x\left[\mathbb{1}_{\{T_y^*<\infty\}}\mathbf{G}(y,y)\right] \\ &= \mathbf{G}(y,y)\mathbb{P}^x(T_y^* < \infty).\end{aligned}$$

En la parte (iii), queremos demostrar que para $\lambda > 0$:

$$\mathbb{E}^x\left[e^{-\lambda T_y^*}\right] = \frac{\mathbf{G}_\lambda^*(x,y)}{1+\mathbf{G}_\lambda^*(y,y)}.$$

donde:

$$\mathbf{G}_\lambda^*(x,y) = \sum_{n=1}^{\infty} e^{-\lambda n}\mathbf{P}^n(x,y) = \mathbb{E}^x\left[\sum_{n=1}^{\infty} e^{-\lambda n}\mathbb{1}_{\{X_n=y\}}\right].$$

Nuevamente, si $T_y^* = \infty$, entonces no hay visitas para $n \geq 1$. Si $T_y^* < \infty$, descomponemos la suma:

$$\sum_{n=1}^{\infty} e^{-\lambda n}\mathbb{1}_{\{X_n=y\}} = \mathbb{1}_{\{T_y^*<\infty\}} \sum_{n=T_y^*}^{\infty} e^{-\lambda n}\mathbb{1}_{\{X_n=y\}}$$

Haciendo el cambio de índice $n = T_y^* + m$ con $m \geq 0$:

$$\sum_{n=T_y^*}^{\infty} e^{-\lambda n}\mathbb{1}_{\{X_n=y\}} = e^{-\lambda T_y^*} \sum_{m=0}^{\infty} e^{-\lambda m}\mathbb{1}_{\{X_{T_y^*+m}=y\}}$$

Por lo tanto:

$$\mathbf{G}_\lambda^*(x,y) = \mathbb{E}^x\left[\mathbb{1}_{\{T_y^*<\infty\}}e^{-\lambda T_y^*} \sum_{m=0}^{\infty} e^{-\lambda m}\mathbb{1}_{\{X_{T_y^*+m}=y\}}\right]$$

Por la propiedad de Markov fuerte, condicional a $\{T_y^* < \infty, X_{T_y^*} = y\}$:

$$\sum_{m=0}^{\infty} e^{-\lambda m}\mathbb{1}_{\{X_{T_y^*+m}=y\}} \overset{d}{=} \sum_{m=0}^{\infty} e^{-\lambda m}\mathbb{1}_{\{X_m=y\}} \quad \text{bajo } \mathbb{P}^y$$

Bajo $\mathbb{P}^y$, esta suma es:

$$\sum_{m=0}^{\infty} e^{-\lambda m}\mathbb{1}_{\{X_m=y\}} = 1 + \sum_{m=1}^{\infty} e^{-\lambda m}\mathbb{1}_{\{X_m=y\}} = 1 + \mathbf{G}_\lambda^*(y,y)$$

(el término 1 viene de $m = 0$ cuando $X_0 = y$). Por tanto:

$$\begin{aligned}\mathbf{G}^*_\lambda(x,y) &= \mathbb{E}^x\left[\mathbb{1}_{\{T^*_y<\infty\}}e^{-\lambda T^*_y}(1+\mathbf{G}^*_\lambda(y,y))\right]\\ &= (1+\mathbf{G}^*_\lambda(y,y))\mathbb{E}^x\left[\mathbb{1}_{\{T^*_y<\infty\}}e^{-\lambda T^*_y}\right]\\ &= (1+\mathbf{G}^*_\lambda(y,y))\mathbb{E}^x\left[e^{-\lambda T^*_y}\right]\end{aligned}$$

(pues $e^{-\lambda T^*_y} = 0$ cuando $T^*_y = \infty$)

Despejando, obtenemos que

$$\mathbb{E}^x\left[e^{-\lambda T^*_y}\right] = \frac{\mathbf{G}^*_\lambda(x,y)}{1+\mathbf{G}^*_\lambda(y,y)}$$

□

Observación 1.14. Observemos que (ii) se puede obtener como caso límite de (iii) cuando $\lambda \to 0$:

Demostración. Por convergencia monótona:

- $\lim_{\lambda\to 0} \mathbf{G}^*_\lambda(x,y) = \mathbf{G}^*(x,y)$
- $\lim_{\lambda\to 0} \mathbb{E}^x\left[e^{-\lambda T^*_y}\right] = \mathbb{P}^x(T^*_y < \infty)$

Tomando límite en (iii):

$$\mathbb{P}^x(T^*_y < \infty) = \frac{\mathbf{G}^*(x,y)}{1+\mathbf{G}^*(y,y)} = \frac{\mathbf{G}^*(x,y)}{\mathbf{G}(y,y)}$$

que es exactamente la relación (ii). □

1.5. Clasificación de estados, recurrencia y transitoriedad

Definición 1.15 (Estados recurrentes y transitorios)**.** Sea $(X_n)_{n\in\mathbb{N}}$ una $\mathrm{CM}(E, v, \mathbf{P})$. Un estado $x \in E$ es:

- **Recurrente** si $\rho(x,x) = 1$ (equivalentemente $\mathbf{G}(x,x) = \infty$)
- **Transitorio** si $\rho(x,x) < 1$ (equivalentemente $\mathbf{G}(x,x) < \infty$)

Estado x recurrente	Estado x transitorio
$\rho(x,x) = 1, \mathbf{G}(x,x) = \infty$	$\rho(x,x) < 1, \mathbf{G}(x,x) < \infty$
$\rho(x,y) > 0 \Rightarrow \mathbb{P}^x(N_y = \infty) = 1$	$\mathbb{P}^x(N_x = \infty) = 0$
$\rho(y,x) = \mathbb{P}^y(N_x = \infty)$	$\mathrm{Ley}(N_x \mid X_0 = x) = \mathrm{Geo}_{\mathbb{N}^*}(1-\rho(x,x))$
$\rho(y,x) > 0 \Rightarrow \mathbf{G}(y,x) = \infty$	$\mathbf{G}^*(y,x) = \rho(y,x)/(1-\rho(x,x)) < \infty$

Cuadro 1: Propiedades de estados recurrentes y transitorios

Definición 1.16 (Clases irreductibles)**.** La relación de comunicación $x \leftrightarrow y$ es una relación de equivalencia. Las clases de equivalencia se llaman **clases irreductibles**. En cada clase, los estados son todos recurrentes o todos transitorios.

Teorema 1.17 (Descomposición del espacio de estados)**.** *Para todo núcleo de transición* $\mathbf{P}$ *sobre* E*:*

1. *La relación* $\leftrightarrow$ *es de equivalencia y partitiona* E *en clases irreductibles*
2. *En cada clase, los estados son todos recurrentes o todos transitorios*
3. *Las clases de estados recurrentes son cerradas (**clases de recurrencia**)*

Ejemplo 1.18 (Cadena con tres clases irreducibles)**.** Consideremos una cadena con espacio de estados $E = \{1, 2, 3, 4, 5\}$ y matriz de transición:

$$\mathbf{P} = \begin{pmatrix} 0.5 & 0.5 & 0 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0.1 & 0 & 0 & 0.4 & 0.5 \\ 0 & 0 & 0 & 0.2 & 0.8 \end{pmatrix}$$

Las clases irreducibles son:

- $C_1 = \{1, 2\}$: Los estados 1 y 2 se comunican entre sí
- $C_2 = \{3\}$: El estado 3 es absorbante (forma su propia clase)
- $C_3 = \{4, 5\}$: Los estados 4 y 5 se comunican entre sí

Ejemplo 1.19 (Determinando recurrencia/transitoriedad)**.** En el ejemplo anterior:

- $C_1 = \{1, 2\}$: Estados recurrentes (clase finita y cerrada)
- $C_2 = \{3\}$: Estado recurrente (absorbante)
- $C_3 = \{4, 5\}$: Estados transitorios (desde 4 y 5 se puede ir a 1, pero no se puede volver)

¿Por qué C_3 es transitoria?

- Desde el estado 4: $\mathbf{P}(4, 1) = 0.1 > 0$, pero desde 1 no se puede volver a 4
- Desde el estado 5: $\mathbf{P}(5, 4) = 0.2 > 0$, pero desde 4 eventualmente se puede ir a 1 y no volver

Definición 1.20 (Clase cerrada)**.** Una clase $C \subset E$ es **cerrada** si:

$$\mathbb{P}^x(T_{C^c} < \infty) = 0 \quad \text{para todo } x \in C$$

Es decir, una vez que la cadena entra en C, nunca puede salir.

Ejemplo 1.21 (Clases cerradas vs no cerradas)**.** En nuestro ejemplo:

- $C_1 = \{1, 2\}$ es **cerrada**: Desde 1 y 2 sólo se puede ir a 1 o 2
- $C_2 = \{3\}$ es **cerrada**: Estado absorbante
- $C_3 = \{4, 5\}$ es **no cerrada**: Desde 4 se puede ir a 1

Proposición 1.22 (Propiedad de clase)**.** *Dentro de una misma clase irreducible:*

- *Todos los estados son del mismo tipo: todos recurrentes o todos transitorios*
- *Si un estado es recurrente positivo/nulo, todos lo son*
- *El período es el mismo para todos los estados de la clase*

Ideas de la demostración. Si $x \leftrightarrow y$, entonces existen m, n tales que:

$$\mathbf{P}^{m+n}(x, x) \geq \mathbf{P}^m(x, y)\mathbf{P}^n(y, x) > 0$$

Esto implica que la recurrencia/transitoriedad se transmite entre estados que se comunican. Más precisamente:

- Si x es recurrente: $\sum_{k=0}^{\infty} \mathbf{P}^k(x, x) = \infty$
- Como $\mathbf{P}^{m+n}(x, x) > 0$, la serie para y también diverge
- Por tanto y también es recurrente

$\square$

Ejemplo 1.23 (Cadena con múltiples clases recurrentes)**.** Consideremos una cadena con matriz:

$$\mathbf{P} = \begin{pmatrix} 0.8 & 0.2 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 \\ 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0.4 & 0.6 \end{pmatrix}$$

- **Clases irreducibles**: $C_1 = \{1, 2\}$, $C_2 = \{3, 4\}$
- **Ambas son cerradas**: No hay transiciones entre clases
- **Ambas son recurrentes**: Clases finitas y cerradas
- **Descomposición**: $E = C_1 \cup C_2$

Ejemplo 1.24 (Cadena con estados transitorios y recurrentes)**.** Consideremos:

$$\mathbf{P} = \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 \\ 0.2 & 0.3 & 0.4 & 0.1 \\ 0.1 & 0.2 & 0.3 & 0.4 \end{pmatrix}$$

- **Clase recurrente**: $C_1 = \{1, 2\}$ (cerrada)
- **Estados transitorios**: $\{3, 4\}$ (no forman clase cerrada)
- **Comportamiento**: Desde 3 o 4, eventualmente la cadena es absorbida por $\{1, 2\}$

1.6. Simulación de trayectorias

Teorema 1.25 (Simulación por sucesiones recurrentes aleatorias)**.** *Para todo núcleo de transición* $\mathbf{P}$ *sobre* E *y toda ley inicial* v*, existen funciones* $f : [0, 1] \to E$ *y* $g : E \times [0, 1] \to E$ *tales que la sucesión:*

$$X_0 = f(U_0), \quad X_{n+1} = g(X_n, U_{n+1})$$

es una $\mathrm{CM}(E, v, \mathbf{P})$ *para cualquier sucesión* $(U_n)_{n\in\mathbb{N}}$ *de v.a. i.i.d. uniformes en* $[0, 1]$.

Demostración. Identificando E con $\mathbb{N}$, definimos:

$$f(u) := \text{mín}\{m \in \mathbb{N} : v(0) + \cdots + v(m) \geq u\}$$

$$g(n, u) := \text{mín}\{m \in \mathbb{N} : \mathbf{P}(n, 0) + \cdots + \mathbf{P}(n, m) \geq u\}$$

Estas son las funciones de distribución inversas de v y $\mathbf{P}(n, \cdot)$. □

1.7. Absorción y Cadenas de Markov como martingalas

Definición 1.26 (Probabilidad y tiempo medio de absorción)**.** Para $F \subset E$ cerrado:

$$a_F(x) := \mathbb{P}^x(T_F < \infty), \quad m_F(x) := \mathbb{E}^x(T_F)$$

Teorema 1.27 (Ecuaciones de absorción)**.** *Sea* $(X_n)_{n\in\mathbb{N}}$ *una* $\mathrm{CM}(E, v, \mathbf{P})$ *y* $F \subset E$*:*

1. *El vector* a_F *es la solución mínima no negativa de:*

$$a_F(x) = \begin{cases} 1 & \text{si } x \in F \\ \sum_{y\in E} \mathbf{P}(x, y) a_F(y) & \text{si } x \notin F \end{cases}$$

2. *El vector* m_F *es la solución mínima no negativa de:*

$$m_F(x) = \begin{cases} 0 & \text{si } x \in F \\ 1 + \sum_{y\in E} \mathbf{P}(x, y) m_F(y) & \text{si } x \notin F \end{cases}$$

Ejemplo 1.28 (Marcha aleatoria sobre $\mathbb{N}$ y ruina del jugador)**.** Sea X_n una marcha aleatoria sobre $\mathbb{N}$ con $\mathbf{P}(0,0) = 1$, $\mathbf{P}(x, x+1) = p$, $\mathbf{P}(x, x-1) = 1-p$ para $x > 0$. Para $p = 1/2$, la probabilidad de absorción en 0 satisface $a(x) = 1$ para todo x, es decir, la ruina es casi segura.

Teorema 1.29 (Martingala de cadena de Markov)**.** *Sea* $(X_n)_{n\in\mathbb{N}}$ *una* $\mathrm{CM}(E, v, \mathbf{P})$, $\mathbf{L} = \mathbf{P} - \mathbf{I}$ *su generador, y* $f : E \to \mathbb{R}$ *integrable. Entonces:*

$$M_n := f(X_n) - f(X_0) - \sum_{k=0}^{n-1} (\mathbf{L}f)(X_k)$$

es una martingala. Además:

- *Si f es armónica ($\mathbf{L}f = 0$), entonces $f(X_n)$ es martingala*
- *Si f es subarmónica ($\mathbf{L}f \geq 0$), entonces $f(X_n)$ es submartingala*
- *Si f es superarmónica ($\mathbf{L}f \leq 0$), entonces $f(X_n)$ es supermartingala*

Demostración. Definiendo $Z_k := f(X_k) - \mathbb{E}(f(X_k) \mid \mathscr{F}_{k-1})$, tenemos:

$$M_n = \sum_{k=1}^{n} Z_k, \quad \text{con } \mathbb{E}(Z_k \mid \mathscr{F}_{k-1}) = 0.$$

□

Ejercicio 1. Sea $(X_n)_{n\geq 1}$ la realización de un dado no sesgado de 6 caras en el n-ésimo lanzamiento (supondremos que los lanzamientos son independientes). Denotamos por Y_n el valor máximo entre las primeras n realizaciones:

$$Y_n = \max_{1\leq i\leq n} X_i$$

y por Z_n el número de veces que salió 6 entre las primeras n realizaciones:

$$Z_n = \sum_{i=1}^{n} \mathbb{1}_{\{6\}}(X_i).$$

1. Mostrar que las cadenas (Y_n) y (Z_n) son cadenas de Markov, y especificar sus matrices de transición.
2. Representar gráficamente estas cadenas y determinar las clases de estados, así como los estados que son recurrentes o transitorios.
3. Mostrar que existe una única medida invariante para (Y_n), y especificarla.

Ejercicio 2. Sea $(X_n)_{n\in\mathbb{N}}$ una cadena de Markov homogénea con matriz de transición P.

1. En cada uno de los casos siguientes, se pide determinar si existen estados absorbentes, cuáles son las clases de comunicación de la cadena, y estudiar la periodicidad y la recurrencia (o transitoriedad) de los estados.

   a) $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ b) $P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$ c) $P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$

   d) $P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$

   e) $P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$

2. Para cada matriz P, estudiar la existencia y unicidad de la(s) posible(s) probabilidad(es) invariante(s).

Referencias

[1] P. Brémaud. *Markov Chains*. Vol. 31. Texts in Applied Mathematics. Gibbs fields, Monte Carlo simulation, and queues. New York: Springer-Verlag, 1999.

[2] Pierre Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. Vol. 31. Texts in Applied Mathematics. Covers stochastic stability, ergodicity, and applications in queueing and simulation. New York: Springer, 1999. ISBN: 978-0-387-98793-9. DOI: 10.1007/978-1-4757-3124-8.

[3] Djalil Chafaï et al. *Processus stochastiques: Ébauche de notes de cours de M1*. Département de mathématiques et applications (DMA), École normale supérieure – PSL, Paris. Équipe pédagogique: Djalil Chafaï (cours) y Alexis Metz–Donnadieu (TD), revisado por Laëtitia Comminges y Arnaud Guyad. 2024.

[4] G. F. Lawler. *Introduction to Stochastic Processes*. 2nd. Boca Raton, FL: Chapman & Hall/CRC, 2006.