Lección n° 3: Invariancia, recurrencia y ergodicidad 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 [3], pero también en un recopilatorio mejor condesando en [2] y otros mencionados en la primera lección.

Las cadenas de Markov describen sistemas cuya evolución en el tiempo es aleatoria, pero cuya incertidumbre está gobernada por una regla de dependencia simple: el futuro depende únicamente del presente. Esta propiedad de "memoria nula" o propiedad markoviana confiere a las cadenas de Markov una estructura matemática rica y una sorprendente capacidad para modelar fenómenos diversos, desde procesos biológicos y físicos hasta sistemas de aprendizaje o dinámicas económicas.

En esta lección, se introducen las *medidas invariantes* y la noción de *equilibrio estacionario*, así como la distinción entre *recurrencia positiva* y *recurrencia nula*. Finalmente, se desarrolla el *teorema ergódico*, que formaliza la convergencia de las medias temporales hacia los valores esperados en el régimen estacionario.

1. Medidas invariantes

Definición 1.1 (Medida invariante)**.** Una medida $\mu \not\equiv 0$ sobre E es **invariante** para el núcleo de transición $\mathbf{P}$ cuando para todo $y \in E$:

$$\mu(y) = \sum_{x \in E} \mu(x)\mathbf{P}(x, y)$$

Esto se escribe $\mu = \mu\mathbf{P}$, o equivalentemente $\mu\mathbf{L} = 0$ donde $\mathbf{L} = \mathbf{P} - \mathbf{I}$ es el generador.

Observación 1.2. Para una medida de probabilidad μ, su invarianza para $\mathbf{P}$ significa que si $(X_n)_{n\in\mathbb{N}}$ es una $\mathrm{CM}(E, \mathbf{P})$ con $X_0 \sim \mu$, entonces $X_n \sim \mu$ para todo tiempo n. Representa un equilibrio en ley invariante por la dinámica, y la ecuación $\mu = \mu\mathbf{P}$ es una ecuación de punto fijo.

Desde el punto de vista del grafo de transiciones, si multiplicamos el peso de cada flecha que llega a x por el peso $\mu(y)$ donde y es el estado de partida de la flecha, la suma total de estos productos debe ser igual a $\mu(x)$ (ecuación de balance).

Ejemplo 1.3 (Medidas invariantes)**.** Algunos ejemplos:

- **Cadena con dos estados**: Para la matriz:

$$\mathbf{P} = \begin{pmatrix} 1 - p_0 & p_0 \\ p_1 & 1 - p_1 \end{pmatrix}$$

  una medida μ es invariante sii $p_0\mu(0) = p_1\mu(1)$. Cuando $p_0 + p_1 > 0$, la única ley invariante es:

$$\mu(0) = \frac{p_1}{p_0 + p_1}, \quad \mu(1) = \frac{p_0}{p_0 + p_1}$$

  Si $p_0 = p_1 = 0$, todas las medidas son invariantes.

- **Proceso de Bernoulli**: Para el proceso de Bernoulli del ejemplo 5.2.2, una medida μ sobre $\mathbb{N}$ es invariante sii:

$$\mu(0) = (1 - p)\mu(0) \quad \text{y} \quad \mu(x) = (1 - p)\mu(x) + p\mu(x - 1) \quad \text{para todo } x > 0$$

  Así, si $0 < p < 1$, no existe medida invariante (la clase transitoria $\{0\}$ pierde masa). Si $p = 0$, todas las medidas son invariantes, mientras que si $p = 1$, toda medida invariante μ es constante.

- **Marcha aleatoria sobre** $\mathbb{Z}$: Para la marcha aleatoria con incrementos de ley $p\delta_1 + (1 - p)\delta_0$, $0 < p < 1$, una medida μ es invariante sii:

$$\mu(x) = \alpha + \beta \left(\frac{p}{1 - p}\right)^x \quad \text{para todo } x \in \mathbb{Z}, \quad \text{con } \alpha, \beta \in \mathbb{R}$$

  No hay unicidad salvo normalización.

- **Proceso de Derman**: Es la cadena de Markov irreducible sobre $E = \mathbb{N}$ con núcleo $\mathbf{P}(x, x+1) = p_x$, $\mathbf{P}(x, 0) = 1 - p_x$ para $x \geq 1$, y $\mathbf{P}(0, 1) = 1$, donde $0 < p_x < 1$ y $\prod_{x=1}^{\infty} p_x > 0$. Si μ es una medida invariante, entonces $\mu(x+1) = \mu(x)p_x$, luego $\mu(x+1) = \prod_{y=1}^{x} p_y$, mientras que:

$$\mu(0) = \sum_{x=1}^{\infty} \mu(x)(1 - p_x) = \sum_{x=1}^{\infty} (\mu(x) - \mu(x+1)) = \mu(1) - \lim_{x\to\infty} \mu(x) = \mu(1) - \prod_{x=1}^{\infty} p_x$$

  Pero $\mu(1) = \mu(0)$ pues $\mathbf{P}(0, 1) = 1$ es la única transición que lleva a 1, de donde una contradicción pues $\prod_{x=1}^{\infty} p_x \neq 0$.

La linealidad de la ecuación $\mu\mathbf{P} = \mu$ implica que el conjunto de medidas invariantes es un cono convexo: si μ_1 y μ_2 son invariantes y $\alpha_1, \alpha_2 \geq 0$ con $\alpha_1 + \alpha_2 > 0$, [2] entonces la medida $\alpha_1\mu_1 + \alpha_2\mu_2$ es también invariante. El conjunto de medidas invariantes es también convexo (tomar $\alpha_1 + \alpha_2 = 1$). Si μ es una medida invariante, entonces todo múltiplo de μ es también invariante. En particular, si μ es invariante y verifica $\mu(E) < \infty$ (siempre cierto si E es finito), entonces $\mu(E)^{-1}\mu$ es una ley invariante.

Recurrencia positiva y recurrencia nula

Definición 1.4 (Número medio de pasajes antes del retorno)**.** Si $(X_n)_{n\in\mathbb{N}}$ es una CM(E), para todos $x, y \in E$, el número medio de pasajes en y antes del retorno a x es:

$$\mu_x(y) := \mathbb{E}^x \left(\sum_{n=0}^{T_x^* - 1} \mathbb{1}_{\{X_n = y\}} \right) \in [0, \infty]$$

Notar que $\mu_x(x) = 1$, y $\mu_x(y) = \delta_x(y)$ si x es absorbante. Además, $\mu_x(y) > 0$ si y sólo si $x \to y$.

Lema 1.5 (Propiedades de μ_x)**.** *Si $\mathbf{P}$ es un núcleo de transición irreducible sobre E, entonces para todo $x \in E$:*

(i) $\mu_x \geq \mu_x\mathbf{P}$, es decir $\mu_x(y) \geq \sum_{z\in E} \mu_x(z)\mathbf{P}(z, y)$ para todo $y \in E$

(ii) μ_x es una medida que carga todos los estados: $0 < \mu_x(y) < \infty$ para todo $y \in E$

(iii) μ_x tiene masa total $\mathbb{E}^x(T_x^*) \in [0, \infty]$, es decir $\mu_x(E) = \mathbb{E}^x(T_x^*)$

(iv) Si $\mathbf{P}$ es recurrente entonces hay igualdad en (i) y μ_x es una medida invariante

Demostración. Para (i), tenemos $\mu_x(y) \geq \mathbb{E}^x \left(\sum_{n=1}^{T_x^*} \mathbb{1}_{\{X_n = y\}} \right)$ con igualdad si $\mathbf{P}$ es recurrente pues en ese caso $T_x^* < \infty$ y $X_{T_x^*} = x$. Este desplazamiento del tiempo permite explotar una transición al final de la trayectoria. Más precisamente:

$$\mathbb{E}^x \left(\sum_{n=1}^{T_x^*} \mathbb{1}_{\{X_n = y\}} \right) = \mathbb{E}^x \left(\sum_{n=1}^{\infty} \mathbb{1}_{\{X_n = y, T_x^* \geq n\}} \right) = \sum_{n=1}^{\infty} \mathbb{P}^x(X_n = y, T_x^* \geq n)$$

lo que permite un desglose según X_{n-1} y el uso de la propiedad de Markov débil, dando (i):

$$\begin{aligned}
\sum_{n=1}^{\infty} \mathbb{P}^x(X_n = y, T_x^* \geq n) &= \sum_{n=1}^{\infty} \sum_{z\in E} \mathbb{P}^x(X_n = y; X_{n-1} = z; T_x^* \geq n) \\
&= \sum_{z\in E} \sum_{n=1}^{\infty} \mathbb{P}(X_n = y \mid X_{n-1} = z, T_x^* \geq n)\mathbb{P}^x(X_{n-1} = z, T_x^* \geq n) \\
&= \sum_{z\in E} \sum_{n=1}^{\infty} \mathbb{P}(X_n = y \mid X_{n-1} = z)\mathbb{P}^x(X_{n-1} = z, T_x^* > n - 1) \\
&= \sum_{z\in E} \mathbf{P}(z, y) \sum_{n=0}^{\infty} \mathbb{P}^x(X_n = z, T_x^* > n) \\
&= \sum_{z\in E} \mathbf{P}(z, y)\mu_x(z)
\end{aligned}$$

Para (*), $\{T_x^* \geq n\} = \{X_1 \neq x, \ldots, X_{n-1} \neq x\} \in \mathscr{F}_{n-1}$, y si $z \neq x$ entonces $\{X_{n-1} = z, X_{n-1} \neq x\} = \{X_{n-1} = z\}$, mientras que si $z = x$ entonces $\mathbb{P}^x(X_{n-1} = z, T_x^* \geq n) = 0$.

Para (ii), por (i), $\mu_x(x)\mathbf{P}^{m_1}(x,y) \leq \mu_x(y)$ y $\mu_x(y)\mathbf{P}^{m_2}(y,x) \leq \mu_x(x)$. Basta elegir m_1, m_2 tales que $\mathbf{P}^{m_1}(x,y) > 0$ y $\mathbf{P}^{m_2}(y,x) > 0$ por irreductibilidad de $\mathbf{P}$, y usar $\mu_x(x) = 1$.

Finalmente (iii) viene de $T_x^* = \sum_{y \in E} \sum_{n=0}^{T_x^*-1} \mathbb{1}_{\{X_n = y\}}$. Hemos usado varias veces el teorema de Fubini-Tonelli en $[0, \infty]$ para conmutar sumas o esperanzas. □

Teorema 1.6 (Medidas invariantes para núcleos irreducibles)**.** *Sea* $\mathbf{P}$ *un núcleo de transición irreducible sobre* E*. Entonces:*

1. *Toda medida invariante* μ *carga todos los estados y verifica* $\mu(y) \geq \mu(x)\mu_x(y)$ *para todos* x, y
2. *Se presentan dos casos distintos:*
   - ***Caso transitorio****:* $\mu(E) = \infty$ *para toda medida invariante* μ*,* E *es necesariamente infinito y no hay distribución invariante*
   - ***Caso recurrente****:* μ_x *es invariante para todo* x*, las medidas invariantes son proporcionales, y si* μ *es invariante entonces* $\mu(y) = \mu(x)\mu_x(y)$ *para todos* x, y*, en particular* $\mu_x(y)\mu_y(x) = 1$

Demostración. Sea (X_n) una $\mathrm{CM}(E, \mathbf{P})$. Sea μ una medida invariante para $\mathbf{P}$. Existe al menos un estado z tal que $\mu(z) > 0$. Para todo y, existe un entero n tal que $\mathbf{P}^n(z,y) > 0$ pues $\mathbf{P}$ es irreducible. Se deduce que:

$$\mu(y) = \sum_x \mu(x)\mathbf{P}^n(x,y) \geq \mu(z)\mathbf{P}^n(z,y) > 0$$

Así, μ carga todos los estados. Sea μ invariante con $\mu(x) = 1$. La fórmula $\mu(y) = \mu(x)\mathbf{P}(x,y) + \sum_{z \neq x} \mu(z)\mathbf{P}(z,y)$ usada recursivamente da, si $y \neq x$:

$$\mu(y) \geq \mathbb{P}^x(X_1 = y, T_x^* > 1) + \cdots + \mathbb{P}^x(X_n = y, T_x^* > n)$$

De donde $\mu(y) \geq \mu_x(y)$ cuando $n \to \infty$.

Caso donde $\mathbf{P}$ es recurrente. Si μ es invariante y $\mu(x) = 1$, entonces sabemos que $\mu \geq \mu_x$, luego $\mu - \mu_x$ es una medida sobre E. Como μ_x es invariante (Lema 5.7.2) y $\mu(x) = \mu_x(x) = 1$, se tiene que para todo n y todo y:

$$0 = (\mu - \mu_x)(x) = ((\mu - \mu_x)\mathbf{P}^n)(x) \geq (\mu - \mu_x)(y)\mathbf{P}^n(y,x)$$

de donde $(\mu - \mu_x)(y) = 0$ eligiendo, gracias a la irreductibilidad de $\mathbf{P}$, un entero n tal que $\mathbf{P}^n(y,x) > 0$.

Caso donde $\mathbf{P}$ es transitorio. En este caso $\mathbf{G}(x,y) < \infty$ para todos x, y, de donde $\lim_{n\to\infty} \mathbf{P}^n(x,y) = 0$. Si $\mu(E) < \infty$ entonces por convergencia dominada, $\mu(y) = \lim_n \sum_x \mu(x)\mathbf{P}^n(x,y) = 0$ para todo y, lo que es imposible, luego $\mu(E) = \infty$. □

Definición 1.7 (Recurrencia positiva y nula)**.** Un estado recurrente x es:

- **Recurrente positivo** si $\mathbb{E}^x(T_x^*) = \mu_x(E) < \infty$ (siempre el caso cuando x es absorbante)
- **Recurrente nulo** si $\mathbb{E}^x(T_x^*) = \mu_x(E) = \infty$ (en este caso T_x^* tiene cola pesada)

Corolario 1.8 (Caracterización de la recurrencia positiva)**.** *Para todo núcleo de transición* $\mathbf{P}$ *sobre* E*, irreducible recurrente, uno de los dos casos siguientes se realiza:*

1. $\mathbf{P}$ *es recurrente positivo: todos los estados son recurrentes positivos. Entonces:*
   - $\mu(E) < \infty$ *para toda medida invariante y existe una única distribución invariante:* $\mu(x) = \frac{1}{\mathbb{E}^x(T_x^*)}$*,* $x \in E$
   - *¡Siempre el caso si* E *es finito!*
2. $\mathbf{P}$ *es recurrente nulo: todos los estados son recurrentes nulos. Entonces:*
   - $\mu(E) = \infty$ *para toda medida invariante* μ
   - E *es necesariamente infinito, y no existe distribución invariante*

Demostración. Por el Teorema 1.6, para todos $x, y \in E$, las medidas invariantes μ_x y μ_y son proporcionales, luego $\mathbb{E}^x(T_x^*) = \mu_x(E)$ y $\mathbb{E}^y(T_y^*) = \mu_y(E)$ también. Por tanto x e y son simultáneamente recurrentes nulos o recurrentes positivos. Si μ es una distribución invariante entonces sumando sobre y en $\mu(y)/\mu(x) = \mu_x(y)$ se obtiene $\mu(x) = 1/\mu_x(E) = 1/\mathbb{E}^x(T_x^*)$. □

Ejemplo 1.9 (Medidas invariantes)**.** Ejemplos de medidas invariantes:

- Toda cadena irreducible sobre un espacio de estados finito es recurrente positiva: es recurrente pues toda trayectoria pasa infinitas veces por al menos un estado, y recurrente positiva pues toda medida es finita.
- La marcha aleatoria simple sobre $\mathbb{Z}$ es irreducible, y la medida de conteo es invariante. Como esta medida tiene masa infinita, la cadena es recurrente nula o transitoria. Sabemos por otra parte que es recurrente pues $\mathbf{G}(x, x) = \mathbf{G}(0, 0) = \infty$ para todo x. Es por tanto recurrente nula.
- Para todo $0 < p < 1$, la marcha aleatoria sobre $E = \mathbb{N}$ reflejada en cero con núcleo dado por $\mathbf{P}(x, x+1) = p$, $\mathbf{P}(x+1, x) = 1-p$, $x \in \mathbb{N}$, $\mathbf{P}(0, 0) = 1-p$, y $\mathbf{P}(0, 1) = p$, es irreducible. La medida $\mu(x) = \rho^x$, $x \in E$, donde $\rho = p/(1-p)$ es invariante. Tiene masa finita sii $p < 1-p$ es decir $p < 1/2$. Se deduce que si $p < 1/2$ entonces la cadena es recurrente positiva con distribución invariante $\mathcal{G}_{\mathbb{N}}(p/(1-p))$. Para $p = 1/2$, la cadena es recurrente nula, mientras que para $p > 1/2$, es transitoria.

5.8 Criterio de Foster y función de Lyapunov

Lema 1.10 (Criterio de recurrencia positiva)**.** *Si $(X_n)_{n\in\mathbb{N}}$ es una* CM(E) *irreducible, y existe $F \subset E$ finito no vacío tal que $\mathbb{E}^x(T_F^*) < \infty$ para todo $x \in F$, donde $T_F^* := \inf\{n \geq 1 : X_n \in F\}$, entonces X es recurrente positiva.*

Demostración. Notemos $(T^n)_{n\in\mathbb{N}}$ los tiempos de retorno sucesivos de X en F, con $T^0 := 0$, y $T^1 = T_F^*$. Sea $x \in F$. Por la propiedad de Markov fuerte, $(Y_n)_{n\in\mathbb{N}} := (X_{T^n})_{n\in\mathbb{N}}$ es una CM(F) irreducible pues X es irreducible. Además, como F es finito, Y es recurrente positiva, y en particular $\mathbb{E}^x(S_x) < \infty$ donde S_x es el tiempo de retorno de Y en $x \in F$. Si T_x^* es el tiempo de retorno en x de X, el objetivo ahora es controlar $\mathbb{E}^x(T_x^*)$ con $\mathbb{E}^x(S_x) < \infty$. Para ello, observamos que $T_x^* = \sum_{k=0}^{\infty} U_k \mathbb{1}_{\{k<S_x\}}$ donde $U_k := T^{k+1} - T^k$, de donde:

$$\mathbb{E}^x(T_x^*) = \sum_{k=0}^{\infty} \mathbb{E}^x(U_k \mathbb{1}_{\{k<S_x\}}), \quad \text{y} \quad \mathbb{E}^x(U_k \mathbb{1}_{\{k<S_x\}}) = \sum_{y\in F} \mathbb{E}^x(U_k \mathbb{1}_{\{k<S_x\}} \mathbb{1}_{\{X_{T^k}=y\}})$$

Por la propiedad de Markov fuerte para T^k, observando que $\{k < S_x\}$ pertenece al pasado de X en el tiempo T^k:

$$\begin{aligned}\mathbb{E}^x(U_k \mathbb{1}_{\{k<S_x\}} \mathbb{1}_{\{X_{T^k}=y\}}) &= \mathbb{E}^x(U_k \mid k < S_x, X_{T^k} = y)\mathbb{P}^x(k < S_x, X_{T^k} = y) \\ &= \mathbb{E}^x(U_k \mid X_{T^k} = y)\mathbb{P}^x(k < S_x, X_{T^k} = y)\end{aligned}$$

Ahora $\mathbb{E}^x(U_k \mid X_{T^k} = y) = \mathbb{E}^y(T_F^*)$, de donde:

$$\mathbb{E}^x(T_x^*) \leq \left(\max_{y\in F} \mathbb{E}^y(T_F^*)\right) \sum_{k=0}^{\infty} \mathbb{P}^x(S_x > k) = \left(\max_{y\in F} \mathbb{E}^y(T_F^*)\right) \mathbb{E}^x(S_x) < \infty$$

□

Teorema 1.11 (Criterio de recurrencia positiva de Foster)**.** *Sea $\mathbf{P}$ un núcleo de transición irreducible sobre E. Entonces $\mathbf{P}$ es recurrente positivo si admite una función de Lyapunov, es decir $V : E \to \mathbb{R}$, $F \subset E$ finito no vacío, y $\varepsilon > 0$, tales que:*

(i) $\inf_E V > -\infty$

(ii) $\mathbf{P}V < \infty$ *sobre* F

(iii) $\mathbf{P}V \leq V - \varepsilon$ *sobre* F^c

Demostración. Gracias a (i) podemos suponer sin pérdida de generalidad que $V \geq 0$, pues (ii) y (iii) son invariantes por adición de una constante a V. Consideremos $T_F^* = \inf\{n \geq 1 : X_n \in F\}$. Como F es finito no vacío, según el Lema 5.8.1, basta establecer que $\mathbb{E}^x(T_F^*) < \infty$ para todo $x \in F$.

Sea (X_n) una $\mathrm{CM}(E, \mathbf{P})$ y $Y_n := V(X_n)\mathbb{1}_{\{T_F^* \geq n\}}$. La propiedad (iii) se escribe $\mathbb{E}(V(X_{n+1}) \mid X_n = x) \leq V(x) - \varepsilon$ para todo $x \notin F$. Si $x \notin F$, entonces:

$$\begin{aligned}\mathbb{E}^x(Y_{n+1} \mid X_0, \ldots, X_n) &= \mathbb{E}^x(V_{n+1}\mathbb{1}_{\{T_F^* > n\}} \mid X_0, \ldots, X_n) = \mathbb{1}_{\{T_F^* > n\}}\mathbb{E}^x(V(X_{n+1}) \mid X_n) \\ &\leq \mathbb{1}_{\{T_F^* > n\}}(V(X_n) - \varepsilon) \leq Y_n - \varepsilon\mathbb{1}_{\{T_F^* > n\}}\end{aligned}$$

donde la primera desigualdad usa que $X_n \notin F$ sobre $\{T_F^* > n\}$, y la segunda que $V \geq 0$ sobre F. Tomando esperanza y usando $V \geq 0$, esto da:

$$0 \leq \mathbb{E}^x(Y_{n+1}) \leq \mathbb{E}^x(Y_n) - \varepsilon\mathbb{P}^x(T_F^* > n)$$

de donde, iterando:

$$0 \leq \mathbb{E}^x(Y_0) - \varepsilon \sum_{k=0}^{n} \mathbb{P}^x(T_F^* > k)$$

Ahora $Y_0 = V(X_0)\mathbb{1}_{\{T_F^* > 0\}} = V(X_0)$, y como $\sum_{k=0}^{\infty} \mathbb{P}^x(T_F^* > k) = \mathbb{E}^x(T_F^*)$, se tiene para todo $x \notin F$:

$$\mathbb{E}^x(T_F^*) \leq \varepsilon^{-1}V(x) < \infty$$

Por otra parte, para todo $x \in F$, el desglose según X_1, la propiedad de Markov débil, y (ii) dan:

$$\mathbb{E}^x(T_F^*) \leq 1 + \sum_{y \notin F} \mathbf{P}(x, y)\mathbb{E}^y(T_F^*) \leq 1 + \varepsilon^{-1} \sum_{y \notin F} \mathbf{P}(x, y)V(y) \leq 1 + \varepsilon^{-1} \sum_{y \in E} \mathbf{P}(x, y)V(y) < \infty$$

□

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] 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.

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