Mixba'al
Revista Metropolitana de Matemáticas
No. 1, Vol. II, junio 2011, 41–58

PRODUCCIÓN DE ENTROPÍA EN CADENAS DE MARKOV

JORGE BOLAÑOS SERVÍN

RESUMEN. La descomposición en ciclos de cadenas de Markov finitas y de la tasa de producción de entropía debida a Qian-Kalpazidou es expuesta brevemente. Además se obtenien condiciones equivalentes a la reversibilidad de la cadena en términos de los ciclos.

1. INTRODUCCIÓN

Las Cadenas de Markov (C.M.), llamadas así en honor al matemático Andrey Markov –quien sentó la base teórica de las mismas–, son ampliamente usadas para modelar fenómenos tanto físicos como sociales, desde la actividad enzimática, para la cual se emplea la cinemática de Michaelis - Menten - Henri, ver [4], hasta más recientemente la manera en que Google cataloga una página electrónica mediante su sistema *PageRank* desarrollado en la universidad de Standford, ver [2] y [3].

Siguiendo el trabajo de Kalpazidou en [5], la escuela china de los Qian da una descripción de la tasa de producción entropía de una C.M. en términos de los ciclos formados, ver [6]. El presente trabajo es una breve exposición de la construcción de esta teoría para una C.M. finita, rellenando detalles y presentando los teoremas de mayor importancia junto con un ejemplo nuevo. (Una exposición autocontenida con todas las demostraciones y una aplicación, que no dan los Qian, a una clase de sistemas cu ánticos abiertos puede encontrarse en [1].)

La noción de circuito dirigido y ciclo es el punto de partida de la teoría. Estos conceptos son introducidos en la Sección 2. En la Sección 3, se estudia la cadena derivada y su distribución invariante con la cual se obtiene una representación probabilística y una descripción de la tasa de producción entropía en términos de los ciclos. Finalmente, en la Sección 3 se extienden los resultados al caso de tiempo continuo. También se presenta un ejemplo discreto y un ejemplo continuo en su respectiva Sección.

2. CIRCUITOS

Un circuito o ciclo es un concepto topológico que puede definirse geométricamente o algebraicamente. Desde el punto de vista geométrico, un circuito de puntos distintos es la imagen, bajo cierta función, de un círculo o de cualquier curva homeomorfa a un círculo, mientras que la definición algebraica requiere la noción de orientación, es decir, distinguir un punto inicial y un punto final en cada arco. Cuando los arcos de un circuito tienen la misma orientación lo llamaremos circuito dirigido.

Una propiedad que poseen los circuitos dirigidos es que regresan periódica- mente a sus puntos, esto motiva una definición que exprese dicha periodicidad.

Definicion 2.1. Una *función de circuito dirigido* en un conjunto numerable S es una función periódica c, $c : \mathbb{Z} \to S$.

A las parejas $(c(n), c(n+1))$, $n \in \mathbb{Z}$, les llamamos aristas, mientras que al menor entero $p = p_c \geq 1$ que satisface la ecuación $c(n+p) = c(n)$ para todo $n \in \mathbb{Z}$ le llamaremos periodo de c.

Con cada función de circuito dirigido c podemos asociar una clase de funciones de circuito dirigido c', construida a partir de c mediante el grupo de translaciones en $\mathbb{Z}$ de

2010 *Mathematics Subject Classification.* 60J99.
Palabras clave. Cadenas de Markov, entropía, ciclos, Qian, Kalpazidou, reversibilidad.

la siguiente manera: para cualquier $i \in \mathbb{Z}$ fijo, definimos la función $t_i(n) := n+i$, $n \in \mathbb{Z}$. Con ello obtenemos una nueva función de circuito dirigido c' mediante la composición $c' = c \circ t_i$, es decir, $c'(n) = c(n+i)$, $n \in \mathbb{Z}$.

El hecho de que c y c' no difieren escencialmente motiva definir la siguiente relación.

Definicion 2.2. Decimos que *dos funciones de circuito dirigido c y c' están relacionadas*, denotándolo por $c \sim c'$, si y sólo si existe $i \in \mathbb{Z}$ tal que $c' = c \circ t_i$.

Esta relación es de equivalencia sobre la clase de todos los circuitos dirigidos:

I) Reflexiva Una función de circuito dirigido c satisface $c \sim c$, pues $c = c \circ t_0$.
II) Simétrica Supongamos que $c \sim c'$, es decir existe $i \in \mathbb{Z}$ tal que $c'(n) = c(n+i)$. A partir de ello vemos que $-i$ satisface $c(n) = c'(n + (-i))$. Por lo tanto $c' \sim c$.
III) Transitiva Si $c \sim c'$ y $c' \sim c''$, entonces existen $i, j \in \mathbb{Z}$ tales que $c'(n) = c(n+i)$ y $c''(n) = c(n+j)$. De donde $c''(n) = c'(n+j) = c(n+i+j)$. Así que $c \sim c''$.

Definicion 2.3. Le llamamos *circuito dirigido* a cada una de las clases de equivalencia inducidas por $\sim$.

Un circuito dirigido c está completamente determinado por

- El periodo p_c
- Cualquier (p_c+1)-tupla $(i_1, i_2, \ldots, i_{p_c}, i_{p_c+1})$ con $i_{p_c+1} = i_1$ ó cualesquiera p_c parejas ordenadas $(i_1, i_2), (i_2, i_3), \ldots, (i_{p_c}, i_{p_c+1})$ con $i_{p_c+1} = i_1$ donde $i_l = c(n+l-1)$, $1 \le l \le p_c$ para algún $n \in \mathbb{Z}$.

Definicion 2.4. El *ciclo dirigido* asociado con un circuito dirigido dado $c = (i_1, i_2, \ldots, i_p, i_1)$, $p \ge 1$, con puntos distintos $i_1, i_2, \ldots, i_p$, es la sucesión ordenada $\hat{c} = (i_1, \ldots, i_p)$.

Como consecuencia de las definiciones anteriores, un ciclo dirigido es invariante respecto a permutaciones cíclicas.

Definicion 2.5. Dado un circuito dirigido $c = (i_1, i_2, \ldots, i_p, i_1)$ definimos el *circuito en reversa* como $c_- = (i_1, i_p, i_{p-1}, \ldots, i_2, i_1)$.

2.1. Funciones de pasaje. Dado un conjunto finito S y un circuito dirigido c en S, para el estudio de sus propiedades es útil saber cuándo un circuito pasa a través de un punto. La manera más sencilla de hacerlo es usar una idea similar a la función indicadora de dicho evento.

Definicion 2.6. Dado un circuito dirigido $c = (i_1, \ldots, i_p, i_1)$, definimos la *función de pasaje de c* denotada por J_c como

$$J_c(k) = \text{card}\{l \in \mathbb{Z} : i_{l+1} = k, 0 \le l \le p_c - 1\}$$

Decimos que *c pasa por k* si y sólo si $J_c(k) \neq 0$. $J_c(k)$ es el número de veces que c pasa por k.

Proposición 2.1. *Si c es un circuito dirigido entonces $J_{c \circ t_j}(k) = J_c(k)$ para todo $j \in \mathbb{Z}$.*

Demostración. Como $c \circ t_j$ es una permutación cíclica de c es claro que $J_{c\circ t_j}(k) = J_c(k)$. □

Definicion 2.7. Dados $k_1, \ldots, k_r \in S$ con $r > 1$ y un circuito dirigido c en S con periodo p, definimos $J_c(k_1, \ldots, k_r)$ como

$$J_c(k_1, \ldots, k_r) = \text{card}\{l \in \mathbb{Z} : c \circ t_l(m) = k_m, m = 1, 2, \ldots, r, 0 \le l \le p_c - 1\}$$

Decimos que *c pasa através de* $(k_1, \ldots, k_r)$ si y sólo si $J_c(k_1, \ldots, k_r) \neq 0$. $J_c(k_1, \ldots, k_r)$ es el número de veces que c pasa por $(k_1, \ldots, k_r)$.

En particular para $r = 2$ en la definición anterior se tiene que

$$J_c(i,j) = \begin{cases} 1 & \text{si } (i,j) \text{ es arista de } c; \\ 0 & \text{otro.} \end{cases}$$

3. Cadenas de Markov tiempo discreto

Sea $\xi = \{\xi_n(\omega)\}_{n\in\mathbb{Z}}$ es una cadena de Markov a tiempo discreto irreducible, positiva recurrente y estacionaria en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, con espacio de estados S finito, matriz de transición $P = (p_{i,j})_{i,j\in S}$ y distribución invariante $\Pi = \{\pi_i\}_{i\in S}$. Pensaremos a $(\Omega, \mathcal{F}, \mathbb{P})$ como el espacio canónico dado por el Teorema de Extensión de Kolmogorov para ξ, es decir,

$$\Omega = S^{\mathbb{Z}} = \{\omega = (\omega_k)_{k\in\mathbb{Z}} : \omega_k \in S,\ \forall k \in \mathbb{Z}\}$$

con $\xi_n(\omega) = \omega_n$.

Supongamos que la órbita ω de ξ es

$$\{3, 4, 7, 3, 1, 2, 5, 2, 4, 1, \dots\}$$

En cada tiempo n, sea $\eta_n(\omega)$ la órbita de la cadena original hasta el tiempo n, respetando el orden en que aparecen los estados, pero descartando los ciclos formados hasta dicho tiempo. Es decir,

n	0	1	2	3	4	5	6
$\xi_n(\omega)$	3	4	7	3	1	2	1
$\eta_n(\omega)$	$[3]$	$[3,4]$	$[3,4,7]$	$[3]$	$[3,1]$	$[3,1,2]$	$[3,1]$
Ciclos descartados				$(3,4,7)$			$(1,2)$

En la siguiente sección se definiŕa $\{\eta_n(\omega)\}_n$ de manera rigurosa.

3.1. La cadena derivada.

Definicion 3.1. Definimos a $[S]$ como el conjunto de sucesiones finitas ordenadas de **elementos distintos** de S. En símbolos

$$[S] = \{[i_1, i_2, \dots, i_r] : i_1, i_2, \dots, i_r \in S \text{ y son distintos}, r \geq 1\}.$$

La concatenación de dos elementos de $[S]$ está bien definida siempre y cuando no tengan puntos en común, i.e.

$$[[i_1, \dots, i_r], [i_{r+1}, \dots, i_{r+n}]] = [i_1, \dots, i_r, i_{r+1}, \dots, i_{r+n}].$$

Claramente $[S]$ es finito. Tomaremos la σ-álgebra $[\Sigma] = 2^{[S]}$.

Definicion 3.2. Definimos una *operación binaria*

$$\uplus : [S] \times S \to [S]$$

mediante

$$[i_1, i_2, \dots, i_r] \uplus i = \begin{cases} [i_1, i_2, \dots, i_r, i] & \text{si } i \notin \{i_1, \dots, i_r\} \\ [i_1, i_2, \dots, i_k], & \text{si } i = i_k, \text{para algún } 1 \leq k \leq r \end{cases}$$

Definicion 3.3. Definimos el *proceso estocástico* $\{\eta_n(\omega)\}_{n\in\mathbb{Z}}$, $\eta_n : \Omega \to [S]$ recursivamente como sigue:

$$\eta_0(\omega) = [\xi_0(\omega)] \qquad \eta_n(\omega) = \eta_{n-1}(\omega) \uplus \xi_n(\omega) \text{ para } n \geq 1.$$

Proposición 3.1. *El proceso estocástico $\{\eta_n(\omega)\}_{n\in\mathbb{Z}}$ está adaptado a la filtración $\{\mathcal{F}_n\}_{n\geq 0}$ donde $\mathcal{F}_n = \sigma(\xi_k : 0 \leq k \leq n)$.*

Demostración. Tenemos que ver que η_n es $\mathcal{F}_n$-medible para cada n. Usaremos inducción. Sea A cualquier medible de $[S]$.

Para $n = 0$,

$$\eta_0^{-1}(A) = \{\omega \in \Omega : [\xi_0(\omega)] = \eta_0(\omega) \in A\} = \bigcup_{[n_1]\in A} \xi_0^{-1}(\{n_1\}) \in \mathcal{F}_0.$$

Supongamos que el resultado es válido para n. La función $\uplus : [S] \times S \longrightarrow [S]$ es trivialmente medible al ser S finito. Por otro lado para $(\eta_n, \xi_{n+1}) : \Omega \longrightarrow [S] \times S$ tomamos un conjunto rectangular $A \times B$ en $[S] \times S$.

$$(\eta_n,\xi_{n+1})^{-1}(A\times B)=\eta_n^{-1}(A)\cap\xi_{n+1}^{-1}(B)\in\mathcal{F}_{n+1} \text{ pues } \mathcal{F}_n\subset\mathcal{F}_{n+1}$$

La familia de conjuntos rectangulares medibles es una álgebra, por tanto es π-sistema contenido en la familia $\mathscr{L}$

$$\mathscr{L}=\{C\in[S]\times S:(\eta_n,\xi_{n+1})^{-1}(C)\in\mathcal{F}_{n+1}\}$$

Es facil ver que $\mathscr{L}$ es un λ-sistema. Por el lema de Dynkin (η_n,ξ_{n+1}) es medible. Usando la relación de recurrencia $\eta_{n+1}=\biguplus\circ(\eta_n,\xi_{n+1})$ concluímos que η_{n+1} es $\mathcal{F}_{n+1}$ medible. □

Definicion 3.4. Definimos $[S]_i$ como el conjunto de elementos $[i_1,\ldots,i_r]$, $r\geq 1$ de [S] tales que $i_1=i$ y $p_{i_k,i_{k+1}}>0\ \forall\ 1\leq k\leq r$.

De acuerdo a la manera en que está definida η, si $\eta_0(\omega)=[i]$ entonces $\eta_n(\omega)\in[S]_i$ para todo n.

Lema 3.2. *$\eta=\{\eta_n\}_{n\geq 0}$ es una cadena de Markov con espacio de estados $[S]$ y distribución inicial*

$$\mathbb{P}(\eta_0=[i])=\begin{cases}\pi_i & i\in S\\ 0 & otro\end{cases}$$

Cada $[S]_i$ es una clase irreducible y positiva recurrente de η. Además para cualquier par de estados $y_1=[i_1,\ldots,i_s], y_2=[j_1,\ldots,j_r]\in[S]_i$ la probabilidad de transición en un solo paso viene dada por

$$\tilde{p}_{y_1,y_2}=\begin{cases}p_{i_s,j_r} & si\ r\leq s\ y\ i_1=j_1,\ i_2=j_2,\ldots,i_r=j_r\\ & ó\ r=s+1\ y\ i_1=j_1,\ i_2=j_2,\ \ldots,i_s=j_s,\\ 0 & otro\ caso\end{cases}$$

Y la distribución invariante única $\tilde{\Pi}^i$ de η en cada clase irreducible $[S]_i$ satisface

$$\tilde{\Pi}^i([i])=\pi_i$$

La propiedad de Markov y la descripción de las probabilidades de transición es clara pues para transiciones permitidas $y_2=[y_1,j_r]$ ó $y_1=[y_2,[i_{r+1},\ldots,i_s]]$ y cualesquiera $z_1,\ldots z_{n-1}\in[S]_i$ adecuados, tenemos que

$$\begin{aligned}&\mathbb{P}(\eta_{n+1}(\omega)=y_2|\eta_n(\omega)=y_1,\eta_{n-1}(\omega)=z_{n-1},\ldots,\eta_0(\omega)=[i])\\ =&\mathbb{P}(\eta_{n+1}(\omega)=y_2,\xi_{n+1}(\omega)=j_r|\xi_n(\omega)=i_s,\eta_n(\omega)=y_1,\ldots,\eta_0(\omega)=[i])\\ =&\mathbb{P}(\xi_{n+1}(\omega)=j_r|\xi_n(\omega)=i_s,\eta_n(\omega)=y_1,\eta_{n-1}(\omega)=z_{n-1},\ldots,\eta_0(\omega)=[i])\\ =&\mathbb{P}(\xi_{n+1}(\omega)=j_r|\xi_n(\omega)=i_s)\\ =&p_{i_s,j_r},\end{aligned}$$

donde hemos usado la propiedad de Markov de ξ.

La irreducibilidad no es dificil de verificar mientras que al ser $[S]_i$ finito la cadena es recurrente positiva. Como consecuencia de un hecho bien conocido de las cadenas de Markov irreducibles y positivas reccurrentes se tiene que

$$\begin{aligned}\tilde{\Pi}^i([i]) &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mathbb{P}(\eta_k=[i]|\eta_0=[i])\\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mathbb{P}(\xi_k=[i]|\xi_0=[i])=\pi_i.\end{aligned}$$

3.2. Distribución estacionaria de η. En esta sección se probarán algunos resultados necesarios para dar la forma explícita la distribución estacionaria, $\tilde{\Pi}^i$, de η, en términos de las entradas de la matriz de transición P de ξ.

Definicion 3.5. Dado un conjunto índice $H = \{h_1, \dots, h_k\}$, definimos $D(H)$ *como el subdeterminante de la matriz* $\mathbf{D} = I - P$ tomando como renglones y columnas a los elementos de H, es decir,

$$D(H) = \begin{vmatrix} d_{h_1,h_1} & d_{h_1,h_2} & \cdots & d_{h_1,h_k} \\ d_{h_2,h_1} & d_{h_2,h_2} & \cdots & d_{h_2,h_k} \\ \vdots & \vdots & \ddots & \vdots \\ d_{h_k,h_1} & d_{h_k,h_2} & \cdots & d_{h_k,h_k} \end{vmatrix}.$$

Si $H = \emptyset$ definimos a $D(H) = 1$.

Lema 3.3. *La distribución estacionaria única* $\Pi = (\pi_i)_{i\in S}$ *de la cadena de Markov* ξ *puede expresarse como*

(1)
$$\pi_i = \frac{D(\{i\}^c)}{\sum_{j\in S} D(\{j\}^c)}.$$

Demostración. Sabemos que Π existe, es única y satisface

$$\Pi P = P, \quad \Pi \mathbf{1} = \mathbf{1},$$

equivalentemente,

$$\Pi \mathbf{D} = \mathbf{0}, \quad \Pi \mathbf{1} = \mathbf{1}.$$

Como la suma de cada renglón de $\mathbf{D}$ es cero, el anterior sistema de ecuaciones es equivalente al sistema

$$(\pi_1, \dots, \pi_N) \begin{pmatrix} 1 & d_{1,1} & d_{1,2} & \cdots & d_{1,j-1} & d_{1,j+1} & \cdots & d_{1,N} \\ 1 & d_{2,1} & d_{2,2} & \cdots & d_{2,j-1} & d_{2,j+1} & \cdots & d_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & d_{N,1} & d_{N,2} & \cdots & d_{N,j-1} & d_{N,j+1} & \cdots & d_{N,N} \end{pmatrix} = (1, 0, \dots, 0),$$

para cada $j \in S$.

Llamaremos $\mathbf{D}_j$ a la matriz del sistema anterior. Como consecuencia del teorema de Perron-Frobenius se tiene que $\det \mathbf{D}_j \neq 0$. Resolviendo el sistema obtenemos,

$$\begin{aligned} \Pi &= (1, 0, \dots, 0)\mathbf{D}_j^{-1} \\ &= (\text{vector de cofactores de la primera columna de } \mathbf{D}_j). \end{aligned}$$

La j-ésima entrada de Π es

(2)
$$\pi_j = \frac{D(\{j\}^c)}{(-1)^{j+1} \det \mathbf{D}_j}.$$

Recordando que $\sum_l d_{i,l} = 0 \Rightarrow -d_{i,k} = \sum_{l\neq k} d_{i,l}$ para toda k, sumamos todas las columnas, excepto la primera, a la segunda columna. Como el intercambio de columas dentro de un determinante sólo le cambian el signo, obtenemos que si $2 \le j \le N$

$$\begin{aligned} \det \mathbf{D}_j &= \\ &= \det \begin{pmatrix} 1 & -d_{1,j} & d_{1,2} & \cdots & d_{1,j-1} & d_{1,j+1} & \cdots & d_{1,N} \\ 1 & -d_{2,j} & d_{2,2} & \cdots & d_{2,j-1} & d_{2,j+1} & \cdots & d_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & -d_{N,j} & d_{N,2} & \cdots & d_{N,j-1} & d_{N,j+1} & \cdots & d_{N,N} \end{pmatrix} \\ &= (-1)^{j+1} \det \mathbf{D}_1 \end{aligned}$$

Por otro lado dado que $\sum_{j\in S} \pi_j = 1$ y que $\pi_j(-1)^{j+1} \det \mathbf{D}_j = D(\{j\}^c)$

$$\det \mathbf{D}_1 = \det \mathbf{D}_1 \sum_{j\in S} \pi_j = \sum_{j\in S} \pi_j (-1)^{j+1} \det \mathbf{D}_j = \sum_{j\in S} D(\{j\}^c)$$

Finalmente combinando lo anterior

$$\pi_i = \frac{D(\{i\}^c)}{(-1)^{i+1}\det \mathbf{D}_i} = \frac{D(\{i\}^c)}{\det \mathbf{D}_1} = \frac{D(\{i\}^c)}{\sum_{j\in S} D(\{j\}^c)}.$$

□

Lema 3.4.

$$\begin{aligned}&D(\{i_1, i_2, \dots, i_{s-1}\}^c)\\ &= d_{i_s,i_s} D(\{i_1,\dots,i_s\}^c)\\ &- \sum_{r>0, j_1,\dots,j_r} p_{i_s,j_1}p_{j_1,j_2}\cdots p_{j_{r-1},j_r}p_{j_r,i_s} D(\{j_1,\dots,j_r,i_1,\dots,i_s\}^c),\end{aligned}$$

en donde la suma se toma sobre todas las elecciones distintas de $j_1,\dots,j_r \in \{i_1,\dots,i_s\}^c$.

Lema 3.5. *Para todo* $j \in \{i_1,\dots,i_s\}^c$ *fijo tenemos que*

$$\begin{aligned}&D(\{i_1,\dots,i_s\}^c)\\ \\ &= \sum_{k=1}^{s} \sum_{\substack{r\geq 0\\ j_1,\dots,j_r}} p_{j,j_1}p_{j_1,j_2}\cdots p_{j_r,i_k} D(\{j,j_1,\dots,j_r,i_1,\dots,i_s\}^c),\end{aligned}$$

donde la suma está tomada sobre las distintas elecciones $j_1,\dots,j_r \in \{j,i_1,\cdots,i_s\}^c$.

Lema 3.6. *Para cualquier* $i \in$

$$\sum_{j\in S} D(\{j\}^c) = \sum_{[i_1,\dots,i_s]\in[S]_i} p_{i_1,i_2}\cdots p_{i_{s-1},i_s} D(\{i_1,\dots,i_s\}^c).$$

Para el caso $s=1$*, entendemos a* $p_{i_1,i_2}\cdots p_{i_{s-1},i_s}$ *como 1.*

Demostración. Usando el lema anterior y sumando

$$\begin{aligned}\sum_{j\in S} D(\{j\}^c) &= \sum_{\substack{j\in S\\ j\neq i}} \sum_{\substack{r\geq 0\\ j_1,\dots,j_r}} p_{i,j_1}p_{j_1,j_2}\cdots p_{j_r,j} D(\{i,j_1,\dots,j_r,l\}^c) + D(\{i\}^c)\\ &= \sum_{\substack{s\geq 2, i_2,\dots,i_s\\ i_1=1}} p_{i_1,i_2}\cdots p_{i_{s-1},i_s} D(\{i_1,\dots,i_s\}^c) + D(\{i\}^c)\\ &= \sum_{[i_1,\dots,i_s]\in[S]_i} p_{i_1,i_2}\cdots p_{i_{s-1},i_s} D(\{i_1,\dots,i_s\}^c).\end{aligned}$$

□

Finalmente con la ayuda de los lemas anteriores podemos dar una descripción de la distribución invanriante de la cadena derivada η,

Teorema 3.7. *Si* ξ *es una cadena de Markov estacionaria con matriz de transición* $P = (p_{i,j})_{i,j\in S}$*, irreducible, positiva recurrente con espacio de estados* S *finito y distribución estacionaria* $\Pi = (\pi_i)_{i\in S}$*, entonces su cadena derivada* η *es recurrente positiva en cada* $[S]_i$ *con distribución invariante* $\tilde{\Pi}^i$ *dada por*

$$\tilde{\Pi}^i([i_1,\dots,i_s]) = p_{i_1,i_2}p_{i_2,i_3}\cdots p_{i_{s-1},i_s}\frac{D(\{i_1,\dots,i_s\}^c)}{\sum_{j\in S} D(\{j\}^c)}. \tag{3}$$

Además se tiene que

$$\tilde{\Pi}^i([i_1,\dots,i_s])p_{i_s,i_1} = \sum_{k=1}^{s} \sum_{\substack{r\geq 1\\ j_2,\dots,j_r}} \tilde{\Pi}^{j_1}([j_1,\dots,j_r,i_k,i_{k+1},\dots,i_{k+s-1}])p_{i_{k-1},i_s}, \tag{4}$$

donde j_1 *es fijo en* $\{i_1,\dots,i_s\}^c$*, la suma es sobre las distintas elecciones* $j_2,\dots,j_r \in \{j_1,i_1,\dots,i_s\}^c$ *y las sumas donde aparece* k *se entienden modulo* s*.*

Demostración. Sólo resta demostrar las ecuaciones (3) y (4). Como η es recurrente positiva en la clase irreducible $[S]_i$ entonces $\tilde{\Pi}^i$ es la única solución del sistema de ecuaciones

$$\begin{aligned} \tilde{\Pi}^i\tilde{P}_i &= \tilde{\Pi}^i \\ \sum_{[i_1,\ldots,i_s]\in[S]_i} \tilde{\Pi}^i([i_1,\ldots,i_s]) &= 1. \end{aligned}$$

Por el lema anterior se sigue que $\tilde{\Pi}^i$ dado por (3) satisface la segunda ecuación. Recordando que $d_{j,j} = 1 - p_{j,j}$ y usando el lema (3.4) obtenemos que

$$\tilde{\Pi}^i([i_1,\ldots,i_s]) = p_{i_1,i_2}p_{i_2,i_3}\cdots p_{i_{s-1},i_s}\frac{D(\{i_1,\ldots,i_s\}^c)}{\sum_{j\in S} D(\{j\}^c)}$$

$$=p_{i_1,i_2}p_{i_2,i_3}\cdots p_{i_{s-1},i_s}\frac{D(\{i_1,\ldots,i_{s-1}\}^c)}{\sum_{j\in S} D(\{j\}^c)}$$

$$+p_{i_1,i_2}p_{i_2,i_3}\cdots p_{i_{s-1},i_s}p_{i_s,i_s}\frac{D(\{i_1,\ldots,i_s\}^c)}{\sum_{j\in S} D(\{j\}^c)}$$

$$+\sum_{\substack{r\geq 1\\ j_1,\ldots,j_r}} p_{i_1,i_2}p_{i_2,i_3}\cdots p_{i_{s-1},i_s}p_{i_s,j_1}p_{j_1,j_2}\cdots p_{j_{r-1},j_r}p_{j_r,i_s}\frac{D(\{j_1,\ldots,j_r,i_1,\ldots,i_s\}^c)}{\sum_{j\in S} D(\{j\}^c)}$$

$$\begin{aligned} &= \underbrace{p_{i_{s-1},i_s}}_{c)}\tilde{\Pi}^i([i_1,\ldots,i_{s-1}]) + \underbrace{p_{i_s,i_s}}_{b)}\tilde{\Pi}^i([i_1,\ldots,i_s]) \\ &\quad+\sum_{\substack{r\geq 1\\ j_1,\ldots,j_r}} \underbrace{p_{j_r,i_s}}_{a)}\tilde{\Pi}^i([i_1,\ldots,i_s,j_1,\ldots,j_r]) \end{aligned}$$

Veamos que esto es exactamente la primera ecuación. Recordemos que si $x = [x_1,\ldots,x_n], y = [y_1,\ldots,y_m] \in [S]_i$ entonces $\tilde{p}_{x,y} \neq 0$ sólo cuando

a) $m < n$
b) $m = n$
c) $m = n+1$

y en los tres casos $\tilde{p}_{x,y} = p_{x_n,y_m}$.

$$\begin{aligned} &= \underbrace{\tilde{p}_{x',y}}_{\substack{x'=[i_1,\ldots,i_{s-1}]\\ y=[x',i_s]\\ c)}}\tilde{\Pi}^i([i_1,\ldots,i_{s-1}]) + \underbrace{\tilde{p}_{x'',y}}_{\substack{x''=y=[i_1,\ldots,i_s]\\ b)}}\tilde{\Pi}^i([i_1,\ldots,i_s]) \\ &\quad+\sum_{\substack{r\geq 1\\ j_1,\ldots,j_r}} \underbrace{\tilde{p}_{x,y}}_{\substack{x=[i_1,\ldots,i_s,j_1,\ldots,j_r]\\ y=[i_1,\ldots,i_s]\\ a)}}\tilde{\Pi}^i([i_1,\ldots,i_s,j_1,\ldots,j_r]) \\ &= \sum_{x\in[S]_i}\tilde{\Pi}^i(x)\tilde{p}_{x,[i_1,\ldots,i_s]}. \end{aligned}$$

Esto demuestra (3). Usando la descripción de $\tilde{\Pi}^i$ junto con el lema (3.5), fijando un $j_1 \in \{i_1,\ldots,i_s\}^c$ se concluye la validez de (4).

□

3.3. Representación probabilística en ciclos.

Definicion 3.6. Le llamamos $\mathcal{C}_n(\omega)$ a la clase de todos los ciclos que ocurren hasta el tiempo n a lo largo de la órbita $\{\xi_l(\omega)\}_{l\geq 0}$

Definicion 3.7. Para un ciclo c, definimos

$$w_{c,n}(\omega) = \sum_{l=1}^{n} 1_{\bigcup_{k=1}^{s}\{\tilde{\omega}\,:\,\eta_{l-1}(\tilde{\omega}) = [\eta_l(\tilde{\omega}),[i_k,i_{k+1},\ldots,i_{k+s-1}]]\}}(\omega).$$

Es claro que $\omega_{c,n}(\omega)$ cuenta el número de veces que aparece el ciclo c ó alguna permutación cíclica en la órbita $\{\xi_l(\omega)\}_{l\geq 0}$ hasta el tiempo n.

Buscamos obtener una familia de ciclos, a la que llamaremos $\mathcal{C}_\infty$, que no dependa de ω y que contenga a todos los ciclos posibles de la cadena ξ. Esto se hace mediante un proceso estocástico auxiliar γ

$$\begin{aligned}\gamma_n &: \Omega \longrightarrow S \times S \\ \gamma_n(\omega) &= (\xi_{n-1}(\omega), \xi_n(\omega))\end{aligned}$$

Proposición 3.8. *El proceso estocástico γ es una cadena de Markov positiva recurrente e irreducible con distribución estacionaria única dada por*

$$\Pi_\gamma(i,j) = \pi_i p_{i,j}$$

Definicion 3.8. Denotamos por $\sigma_n(\omega; i, j)$ al *número promedio de transiciones de i a j en la órbita de $\{\xi_l(\omega)\}_l \geq 0$ hasta el tiempo n.*

En símbolos

$$\sigma_n(\omega; i, j) = \frac{1}{n} \operatorname{card}\{m : 0 \leq m < n, \xi_m(\omega) = i, \xi_{m+1}(\omega) = j\}$$

Proposición 3.9. *Para todo $i, j \in S$*

$$\lim_{n\to\infty} \sigma_n(\omega; i, j) = \pi_i p_{i,j} \quad c.s.$$

Demostración. Es consecuencia de la ley de los grandes números para cadenas de Markov aplicada al proceso γ. □

Proposición 3.10. *Para cada ciclo $c = (i_1, \dots, i_s)$ el siguiente límite existe*

$$\lim_{n\to\infty} \frac{w_{c,n}(\omega)}{n} = w_c \quad c.s.,$$

donde

$$w_c = p_{i_1,i_2} p_{i_2,i_3} \cdots p_{i_{s-1},i_s} p_{i_s,i_1} \frac{D(\{i_1, \dots, i_s\}^c)}{\sum_{j\in S} D(\{j\}^c)}$$

Es importante notar que el límite w_c **no** depende de ω.

Por otro parte la sucesión $(\mathcal{C}_n(\omega))_n$ converge casi seguramente a una clase de ciclos que denotamos por $\mathcal{C}_\infty$. Esto es claro observando que los $\mathcal{C}_n(\omega)$ forman una sucesión es creciente. El límite es la unión de todos ellos.

Para ver que no depende de la trayectoria, sean ω_1, ω_2 fuera del conjunto de probabilidad cero dado por la proposición anterior. Entonces para cada $c \in \mathcal{C}_\infty(\omega_1)$

$$\lim_{n\to\infty} \frac{w_{c,n}(\omega_1)}{n} = w_c = \lim_{n\to\infty} \frac{w_{c,n}(\omega_2)}{n},$$

y así $c \in \mathcal{C}_\infty(\omega_2)$.

Definicion 3.9. La *clase de ciclos* $\mathcal{C}_\infty$ dada por la proposición anterior es la colección de ciclos que ocurren en casi todas las órbitas $\{\xi_n(\omega)\}_n$. Al conjunto de pesos $\{w_c : c \in \mathcal{C}_\infty\}$ de las proposiciones anteriores se les llama *distribución circulatoria* de ξ.

Finalmente podemos obtener una representación probabilística en ciclos, siendo esto la mitad del trabajo para obtener una descripción en ciclos de la entropía.

Teorema 3.11. ***(Representación Probabilística en Ciclos)*** *Si ξ es una cadena de Markov estacionaria con matriz de transición $P = (p_{i,j})_{i,j\in S}$, irreducible, positiva recurrente con espacio de estados S finito y distribución estacionaria $\Pi = (\pi_i)_{i\in S}$, entonces*

$$\begin{aligned}\pi_i p_{i,j} &= \lim_{n\to\infty} \sum_{c\in\mathcal{C}_\infty} \frac{w_{c,n}(\omega)}{n} J_n(i,j) \quad c.s \\ &= \sum_{c\in\mathcal{C}_\infty} w_c J_c(i,j) \quad \forall i, j \in S.\end{aligned}$$

Demostración. Para esta demostración definimos la función $\epsilon_n(\omega; i, j,)$ como

$$\epsilon_n(\omega; i, j) = \begin{cases} 1 & \text{si la última transición de } i \text{ a } j \text{ pertenece} \\ & \text{ a algún ciclo de } \mathcal{C}_n(\omega) \\ 0 & \text{otro.} \end{cases}$$

Con ello vemos que

$$\sigma_n(\omega; i, j) = \sum_{c \in \mathcal{C}_n(\omega)} \frac{w_{c,n}(\omega)}{n} J_c(i, j) + \frac{\epsilon_n(\omega; i, j)}{n}.$$

Tomando límite cuando $n \to \infty$ en ambos lados y usando la Proposición 3.9, obtenemos

$$\pi_i p_{i,j} = \sum_{c \in \mathcal{C}_\infty} w_c J_c(i, j).$$

□

3.4. Producción de entropía. Sea r la transformación medible que invierte el tiempo,

$$r : (\Omega, \mathcal{F}) \to (\Omega, \mathcal{F})$$
$$rw(n) = w(-n) \quad n \in \mathbb{Z}.$$

Definicion 3.10. Definimos *el proceso estocástico* ξ^- como

$$\xi_n^-(\omega) = \xi_n(r\omega) = \xi_{-n}(\omega)$$

y la medida $\mathbb{P}^- = r\mathbb{P}$ sobre $(\Omega, \mathcal{F})$, en donde

$$\mathbb{P}^-(A) = r\mathbb{P}(A) = \mathbb{P}(r^{-1}(A)) \quad \forall A \in \mathcal{F}.$$

ξ^- es el proceso ξ con el tiempo en reversa, y su distribución viene dada por $\mathbb{P}^-$.

El proceso ξ^- es una cadena de Markov estacionaria en $(\Omega, \mathcal{F}, \mathbb{P})$ con matriz de transición

$$P^- = (p_{i,j}^-)_{i,j \in S} = \left(\frac{\pi_j p_{j,i}}{\pi_i} \right)_{i,j \in S},$$

y distribución invariante $\Pi^- = \Pi$. Es bien sabdico que ξ es reversible si y sólo si $\mathbb{P} = \mathbb{P}^-$

A continuación definimos entropía relativa de medidas de probabilidad.

Definicion 3.11. Supongamos que μ y λ son dos medidas de probabilidad sobre un espacio medible $(M, \mathcal{A})$, *la entropía relativa de* μ *respecto a* λ se define como

$$H(\mu, \lambda) = \begin{cases} \displaystyle\int_M \log \frac{d\mu}{d\lambda}(x) \mu(dx) & \text{si } \mu \ll \lambda \text{ y } \log \frac{d\mu}{d\lambda} \in L_1(\mu) \\ +\infty & \text{otro} \end{cases}$$

En adelante usaremos la siguiente notación para la sub σ-álgebra generada por un número finito de variables aleatorias sucesivas y para las medidas restringidas a dichas sub σ-álgebras

$$\mathcal{F}_m^n = \sigma(\xi_k : m \le k \le n)$$
$$\mathbb{P}_{[m,n]} = \mathbb{P}|_{\mathcal{F}_m^n}$$
$$\mathbb{P}_{[m,n]}^- = \mathbb{P}^-|_{\mathcal{F}_m^n}.$$

Definicion 3.12. La *tasa de producción de entropía de la cadena de Markov estacionaria* ξ se define por

$$e_p = \lim_{n \to \infty} \frac{1}{n} H(\mathbb{P}_{[0,n]}, \mathbb{P}_{[0,n]}^-)$$

donde $H(\mathbb{P}_{[0,n]}, \mathbb{P}_{[0,n]}^-)$ es la entropía relativa de $\mathbb{P}$ con respecto a $\mathbb{P}^-$ restringidas a la σ-álgebra $\mathcal{F}_0^n$.

Lema 3.12. *Si la matriz de transición* P *de* ξ *satisface*

$$p_{i,j} > 0 \Longleftrightarrow p_{j,i} > 0 \quad \forall i, j \in S$$

entonces para todo $m, n \in \mathbb{Z}$ *las medidas* $\mathbb{P}_{[m,m+n]}$ *y* $\mathbb{P}_{[m,m+n]}^-$ *son equivalentes, es decir,* $\mathbb{P}_{[m,m+n]} \ll \mathbb{P}_{[m,m+n]}^-$ *y* $\mathbb{P}_{[m,m+n]}^- \ll \mathbb{P}_{[m,m+n]}$.

Además la derivada de Radon-Nikodym viene dada por

$$\frac{d\mathbb{P}_{[m,m+n]}}{d\mathbb{P}^-_{[m,m+n]}}(\omega) = \frac{\pi_{\xi_m(\omega)}p_{\xi_m(\omega),\xi_{m+1}(\omega)}\cdots p_{\xi_{m+n-1}(\omega),\xi_{m+n}(\omega)}}{\pi_{\xi_{m+n}(\omega)}p_{\xi_{m+n}(\omega),\xi_{m+n-1}(\omega)}\cdots p_{\xi_{m+1}(\omega),\xi_m(\omega)}} \quad \mathbb{P}-c.s \tag{5}$$

Demostración. La medida de un conjunto generador $B \times S^{\mathbb{Z}-N}$ de la σ-álgebra $\mathcal{F}_m^{m+n}$, es decir, de un conjunto de la forma $\{i_m\} \times \cdots \{i_{m+n}\} \times S^{\mathbb{Z}-N}$ es

$$\begin{aligned}\mathbb{P}_{[m,m+n]}(B \times S^{\mathbb{Z}-N}) &= \mathbb{P}(\xi_m(\omega), \ldots, \xi_{m+n}(\omega) \in B)\\ &= \pi_{i_m}p_{i_m,i_{m+1}}\cdots p_{i_{m+n-1},i_{m+n}}.\end{aligned}$$

Similarmente

$$\begin{aligned}\mathbb{P}^-_{[m,m+n]}(B \times S^{\mathbb{Z}-N}) &= \pi^-_{i_m}p^-_{i_m,i_{m+1}}\cdots p^-_{i_{m+n-1},i_{m+n}}\\ &= \pi_{i_{m+n}}p_{i_{m+n},i_{m+n-1}}\cdots p_{i_{m+1}(\omega),i_m}.\end{aligned}$$

Y usando la relación entre P y P^- obtenemos que

$$\mathbb{P}_{[m,m+n]}(B \times S^{\mathbb{Z}-N}) = \frac{\pi_{i_m}p_{i_m,i_{m+1}}\cdots p_{i_{m+n-1},i_{m+n}}}{\pi_{i_{m+n}}p_{i_{m+n},i_{m+n-1}}\cdots p_{i_{m+1}}}\mathbb{P}^-_{[m,m+n]}(B \times S^{\mathbb{Z}-N}).$$

Así bajo la hipótesis de la proposición las medidas son equivalentes.

Recordemos que los conjuntos cilíndricos forman un álgebra, por lo que forman un π-sistema contenido en el siguiente conjunto

$$\mathscr{L} = \{A \in \mathcal{F}_m^{m+n} : \mathbb{P}_{[m,m+n]}(A) = \int_A f d\mathbb{P}^-_{[m,m+n]}\}$$

donde f es la función dada por (5). Es facil verificar que $\mathscr{L}$ es un λ-sistema y por el lema de Dynkin $\mathscr{L} = \mathcal{F}_m^{m+n}$. Por la unicidad de la derivada de Radon-Nikodym se concluye (5). □

Teorema 3.13. *La tasa de produción de entropía e_p de la cadena de Markov estacionaria ξ puede expresarse como*

$$e_p = \frac{1}{2}\sum_{i,j\in S}(\pi_i p_{i,j} - \pi_j p_{j,i})\log\frac{\pi_i p_{i,j}}{\pi_j p_{j,i}} \tag{6}$$

$$= \frac{1}{2}\sum_{c\in\mathcal{C}_\infty}(w_c - w_{c_-})\log\frac{w_c}{w_{c_-}} \tag{7}$$

Demostración. Basta considerar el caso en que

$$p_{i,j} > 0 \Longleftrightarrow p_{j,i} > 0 \quad \forall i,j \in S,$$

de lo contrario $e_p = +\infty$ mientras que en el lado derecho de (6) ningún término puede ser $-\infty$ y por lo menos uno de ellos es $+\infty$. En dicho caso (6) se cumple.

Así pues, bajo la hipótesis del lema anterior,

$$\begin{aligned} e_p &= \lim_{n\to\infty}\frac{1}{n}H(\mathbb{P}_{[0,n]}, \mathbb{P}^-_{[0,n]})\\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{i_0,\ldots,i_n\in S}\pi_{i_0}p_{i_0,i_1}\cdots p_{i_{n-1},i_n}\log\frac{\pi_{i_0}p_{i_0,i_1}\cdots p_{i_{n-1},i_n}}{\pi_{i_n}p_{i_n,i_{n-1}}\cdots p_{i_1,i_0}}\end{aligned}$$

y utilizando propiedades del logaritmo para separar productos en sumas,

$$\begin{aligned} &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\sum_{i_k,i_{k+1}\in S}\pi_k p_{i_k,i_{k+1}}\log\frac{\pi_k p_{i_k,i_{k+1}}}{\pi_{i_{k+1}}p_{i_{k+1},i_k}}\\ &= \sum_{i,j\in S}\pi_i p_{i,j}\log\frac{\pi_i p_{i,j}}{\pi_j p_{j,i}}, \quad \text{pues la suma interna no depende de } n\\ &= \tfrac{1}{2}\sum_{i,j\in S}(\pi_i p_{i,j} - \pi_j p_{j,i})\log\frac{\pi_i p_{i,j}}{\pi_j p_{j,i}}\end{aligned}$$

Esto demuestra el primer renglón. Para el segundo renglón se tiene que

$$\begin{aligned}
e_p &= \tfrac{1}{2}\sum_{i,j\in S}\sum_{c\in\mathcal{C}_\infty}(w_c - w_{c_-})J_c(i,j)\log\frac{\pi_i p_{i,j}}{\pi_j p_{j,i}} \\
&= \tfrac{1}{2}\sum_{c\in\mathcal{C}_\infty}\sum_{i,j\in S}(w_c - w_{c_-})J_c(i,j)\log\frac{\pi_i p_{i,j}}{\pi_j p_{j,i}} \\
&= \tfrac{1}{2}\sum_{\substack{c\in\mathcal{C}_\infty \\ c=(\ldots,i_k,i_{k+1},\ldots)}}\sum_{k=1}^{s}(w_c - w_{c_-})\log\frac{\pi_{i_k}p_{i_k,i_{k+1}}}{\pi_{i_{k+1}}p_{i_{k+1},i_k}} \\
&= \tfrac{1}{2}\sum_{c\in\mathcal{C}_\infty}(w_c - w_{c_-})\log\prod_{k=1}^{s}\frac{\pi_{i_k}p_{i_k,i_{k+1}}}{\pi_{i_{k+1}}p_{i_{k+1},i_k}} \\
&= \tfrac{1}{2}\sum_{c\in\mathcal{C}_\infty}(w_c - w_{c_-})\log\frac{w_c}{w_{c_-}}.
\end{aligned}$$

□

Hemos encontrado condiciones equivalentes para la reversibilidad de una cadena de Markov en términos de los ciclos ó distribución circulatoria:

1) La cadena deMarkov ξ es reversible.
2) La cadena de Markov está en balance detallado, esto es,
$$\pi_i p_{i,j} = \pi_j p_{j,i} \quad \forall i,j \in S.$$
3) La cadena de Markov satisface el criterio de Kolmogorov,
$$p_{i_1,i_2}p_{i_2,i_3}\cdots p_{i_{s-1},i_s} = p_{i_s,i_{s-1}}\cdots p_{i_3,i_2}p_{i_2,i_1},$$
para cualquier colección finita $\{i_1,\ldots,i_s\}$ de elementos distintos de S.
4) Los elementos de la distribución circulatoria $\{w_c : c \in \mathcal{C}_\infty\}$ satisfacen la condición de simetría
$$w_c = w_{c_-} \quad \forall c \in \mathcal{C}_\infty.$$
5) La tasa de producción de entropía es cero
$$e_p = 0$$

Ejemplo 3.14.

El caso mas simple no trivial es cuando el espacio de estados es $S = \{1,2,3\}$ y matriz de transición

$$P = \begin{pmatrix} 0 & p & q \\ q & 0 & p \\ p & q & 0 \end{pmatrix},$$

donde $p,q > 0$ y $p+q = 1$. La distribución invariante es $\Pi = (\frac{1}{3},\frac{1}{3},\frac{1}{3})$.

Los ciclos que aparecen en casi todas las trayectorias son

$$\mathcal{C}_\infty = \{(1,2,3),(3,2,1),(1,2),(2,3),(3,1)\}$$

Recordemos que un ciclo es equivalente a sus permutaciones cíclicas. Los pesos asociados a cada ciclo en virtud de la proposición (3.10) son

$$w_{(1,2,3)} = \frac{p_{1,2}p_{2,3}p_{3,1}}{\begin{vmatrix} 1 & -p \\ -q & 1 \end{vmatrix} + \begin{vmatrix} 1 & -p \\ -q & 1 \end{vmatrix} + \begin{vmatrix} 1 & -p \\ -q & 1 \end{vmatrix}} = \frac{p^3}{3(1-pq)},$$

$$w_{(3,2,1)} = \frac{q^3}{3(1-pq)},$$

$$w_{(1,2)} = w_{(2,3)} = w_{(3,1)} = \frac{pq}{3(1-pq)},$$

Es importante notar que para cada ciclo c de longitud 2, se tiene que c_- es la única permutación cíclica de c, y por la equivalencia conocida entonces $w_c = w_{c_-}$. Esto y la ecuación (7) muestra que los ciclos que contribuyen a la tasa de producción de entropía son aquellos de longitud 3 o más.

En nuestro ejemplo la tasa de producción de entropía es

$$e_p = (w_{(1,2,3)} - w_{(3,2,1)}) \log \frac{w_{(1,2,3)}}{w_{(3,2,1)}} = \frac{p^3 - q^3}{3(1-pq)} \log \left(\frac{p}{q}\right)^3 = \frac{p^3 - q^3}{1-pq} \log \frac{p}{q},$$

$$= \frac{(p-q)(p^2+pq+q^2)}{1-pq} \log \frac{p}{q} = (p-q) \log \frac{p}{q}.$$

Pues si $p + q = 1$, entonces $(p + q)^2 = 1$ y con ello $p^2 + pq + q^2 = 1 - pq$.

De lo anterior vemos que la cadena de Markov ξ es reversible si y sólo si $p = q = \frac{1}{2}$.

4. Cadenas de Markov tiempo continuo

Para el caso cuando $\xi = \{\xi_t\}_{t \in \mathbb{R}}$ es una cadena de Markov a tiempo continuo, irreducible, recurrente y estacionaria con espacio de estados finito $S = \{1, \ldots, N\}$ y distribución invariante $\Pi = \{\pi_i\}_{i \in S}$ sobre el espacio de probabilidad canónico $(\Omega, \mathcal{F}, \mathbb{P})$ de ξ. Denotaremos a su matriz de intensidades de transición, o generador infinitesimal, por $Q = (q_{i,j})_{i,j \in S}$ donde

$$q_i := \sum_{j \in S, \neq i} q_{i,j} = -q_{i,i} < \infty \quad \forall i \in S.$$

Denotaremos por $\tilde{\xi}$ a la cadena de Markov encajada, y $\tilde{\Pi}$ su distribución invariante.

Proposición 4.1. *Las distribuciones invariantes de ξ y $\tilde{\xi}$ satisfacen las siguientes relaciones*

$$\tag{8} \begin{aligned} \tilde{\pi}_i &= \frac{D(\{i\}^c)}{\sum_{j \in S} D(\{j\}^c)} = \frac{\tilde{D}(\{i\}^c) q_i}{\sum_{j \in S} \tilde{D}(\{j\}^c) q_j} \quad \forall i \in S. \\ \pi_i &= \frac{\tilde{D}(\{i\}^c)}{\sum_{j \in S} \tilde{D}(\{j\}^c)} = \frac{\tilde{\pi}_i / q_i}{\sum_{j \in S} \tilde{\pi}_j / q_j} \quad \forall i \in S. \end{aligned}$$

Demostración. Las primeras igualdades en (8) se siguen del Lema 3.3 de la sección anterior y del hecho de que $\Pi Q = 0$.

Para obtener la última igualdad para $\tilde{\pi}_i$ bastará ver que para $i \neq j$ se tiene que $D(\{j\}^c)\tilde{D}(\{i\}^c)q_i = D(\{i\}^c)\tilde{D}(\{j\}^c)q_j$. En efecto, utilizando la relación entre los elementos de matriz de Q y D y propiedades de los determinantes

$$\begin{aligned} D(\{j\}^c)\tilde{D}(\{i\}^c)q_i &= D(\{j\}^c)\tilde{D}(\{i\}^c)q_i \prod_l \frac{-q_l}{-q_l} \\ &= \tilde{D}(\{j\}^c)\tilde{D}(\{i\}^c)q_j \prod_{l \neq i} \frac{1}{-q_l} \\ &= \tilde{D}(\{j\}^c)D(\{i\}^c)q_j. \end{aligned}$$

A partir de esta relación se ve que la última igualdad para π_i es inmediata pues,

$$\begin{aligned} \tilde{D}(\{j\}^c)q_j\tilde{\pi}_i &= \left(\frac{\tilde{D}(\{j\}^c)q_j}{\sum_{l \in S} \tilde{D}(\{l\}^c)q_l} \right) \tilde{D}(\{i\}^c)q_i \\ &= \tilde{\pi}_l \tilde{D}(\{i\}^c)q_i. \end{aligned}$$

□

Aplicando los resultados de la sección anterior a la cadena encajada obtenemos la descomposición en ciclos de la cadena continua. Sin embargo, hay algunas consideraciones técnicas que hay que tratar con cuidado, las cuales abordaremos a continuación.

4.1. Representación Probabilística.

Definicion 4.1. Definimos $n_t(\omega)$ como

$$n_t(\omega) = \sup\{n \geq 0 : T_n(\omega) \leq t\}$$

De la definición se ve que $n_t(\omega)$ es el *número de transiciones en la órbita ω hasta el tiempo t de la cadena ξ.*

Definicion 4.2. Para la cadena ξ definimos

$$\tilde{w}_{c,t}(\omega) = w_{c,n_t(\omega)}$$

Es claro que $\tilde{w}_{c,t}(\omega)$ es el *número de veces que se completa el ciclo c en la órbita ω hasta el tiempo t.*

La sección anterior establece la existencia de $\lim_{n\to\infty} \frac{w_{c,n}(\omega)}{n}$, mientras que para el caso continuo buscamos la existencia de $\lim_{t\to\infty} \frac{\tilde{w}_{c,t}(\omega)}{t}$. El siguiente lema nos proporciona el paso intermedio para lograrlo.

Lema 4.2. *Para $\mathbb{P}$-casi todo $\omega \in \Omega$*

$$\lim_{t\to+\infty} \frac{n_t(\omega)}{t} = \frac{\sum_{i\in S} \tilde{D}(\{i\}^c) q_i}{\sum_{i\in S} \tilde{D}(\{i\}^c)}$$

Proposición 4.3. *Para cada ciclo $c = (i_1, \ldots, i_s)$ el siguiente límite existe*

$$\lim_{t\to\infty} \frac{\tilde{w}_{c,t}(\omega)}{t} = w_c \quad c.s.,$$

donde

$$w_c = (-1)^{s-1} q_{i_1,i_2} q_{i_2,i_3} \cdots q_{i_{s-1},i_s} q_{i_s,i_1} \frac{\tilde{D}(\{i_1,\ldots,i_s\}^c)}{\sum_{j\in S} \tilde{D}(\{j\}^c)}.$$

Demostración. Para cada ciclo c por la Proposición 3.10, aplicada a la cadena encajada $\tilde{\xi}$, obtenemos que casi seguramente

$$\begin{aligned}
\lim_{n\to\infty} \frac{w_{c,n}(\omega)}{n} &= \tilde{p}_{i_1,i_2}\tilde{p}_{i_2,i_3} \cdots \tilde{p}_{i_{s-1},i_s}\tilde{p}_{i_s,i_1} \frac{D(\{i_1,\ldots,i_s\}^c)}{\sum_{j\in S} D(\{j\}^c)} \\
&= \frac{q_{i_1,i_2}}{q_{i_1}} \cdots \frac{q_{i_s,i_1}}{q_{i_s}} \frac{\prod_{l\in\{i_1,\ldots,i_s\}^c} \frac{1}{-q_l} \tilde{D}(\{i_1,\ldots,i_s\}^c)}{\prod_{l\in S} \frac{1}{-q_l} \sum_{j\in S} -q_j \tilde{D}(\{j\}^c)} \\
&= \prod_{l\in\{i_1,\ldots,i_s\}} -q_l \frac{q_{i_1,i_2}}{q_{i_1}} \cdots \frac{q_{i_s,i_1}}{q_{i_s}} \frac{\tilde{D}(\{i_1,\ldots,i_s\}^c)}{\sum_{j\in S} -q_j \tilde{D}(\{j\}^c)} \\
&= (-1)^{s-1} q_{i_1,i_2} q_{i_2,i_3} \cdots q_{i_{s-1},i_s} q_{i_s,i_1} \frac{\tilde{D}(\{i_1,\ldots,i_s\}^c)}{\sum_{j\in S} q_j \tilde{D}(\{j\}^c)}.
\end{aligned}$$

Y usando el lema anterior

$$\begin{aligned}
w_c &= \lim_{t\to\infty} \frac{\tilde{w}_{c,t}(\omega)}{t} \\
&= \lim_{t\to\infty} \frac{n_t(\omega)}{t} \frac{w_{c,n_t(\omega)}(\omega)}{n_t(\omega)} \\
&= (-1)^{s-1} q_{i_1,i_2} q_{i_2,i_3} \cdots q_{i_{s-1},i_s} q_{i_s,i_1} \frac{\tilde{D}(\{i_1,\ldots,i_s\}^c)}{\sum_{i\in S} \tilde{D}(\{i\}^c)}.
\end{aligned}$$

$\square$

Al igual que en caso discreto, la sucesión $\{\mathcal{C}_t(\omega)\}_{t>0}$ converge casi seguramente a una clase de ciclos que denotaremos por $\mathcal{C}_\infty$.

Teorema 4.4. ***(Representación Probabilística en Ciclos)*** *Si ξ es una cadena de Markov estacionaria con generador Q, irreducible, recurrente positiva con espacio de estados S finito y distribución estacionaria Π, entonces*

$$\pi_i q_{i,j} = \sum_{c\in\mathcal{C}_\infty} w_c J_c(i,j) \quad \forall i,j\in S \quad i\neq j.$$

Demostración. Para cada $c\in\mathcal{C}_\infty$ denotamos $\tilde{w}_c = \lim_{n\to\infty} \frac{w_{c,n}(\omega)}{n}$ $c.s$, entonces claramente $\{\tilde{w}_c : c\in\mathcal{C}_\infty\}$ es la distribución circulatoria de la cadena encajada $\tilde{\xi}$, aún más por la Proposición 4.3 sabemos que

$$\tilde{w}_c = w_c \frac{\sum_{i\in S} \tilde{D}(\{i\}^c)}{\sum_{j\in S} q_j \tilde{D}(\{j\}^c)}.$$

Utilizando la representación probabilística en ciclos de la cadena encajada $\tilde{\xi}$ dada por el Teorema 3.11,

$$\tilde{\pi}_i\tilde{p}_{i,j} = \sum_{c\in\mathcal{C}_\infty} \tilde{w}_c J_c(i,j),$$

las relaciones entre las distribuciones invariantes Π y $\tilde{\Pi}$, de la Proposición 4.1 obtenemos que

$$\pi_i q_{i,j} = \sum_{c\in\mathcal{C}_\infty} w_c J_c(i,j).$$

□

4.2. Producción de entropía.

Mediante la transformación que invierte el tiempo,

$$r : (\Omega,\mathcal{F}) \to (\Omega,\mathcal{F}),$$
$$rw(t) = \lim_{s\uparrow -t} w(s), \quad \forall t\in\mathbb{R},$$

definimos el proceso estocástico ξ^- como sigue.

Definicion 4.3. Sea

$$\xi_t^-(\omega) = \xi_t(r\omega),$$

y $\mathbb{P}^- = r\mathbb{P}$ una medida sobre $(\Omega,\mathcal{F})$, en donde

$$\mathbb{P}^-(A) = r\mathbb{P}(A) = \mathbb{P}(r^{-1}(A)), \quad \forall A\in\mathcal{F}.$$

ξ^- *es la cadena en reversa de* ξ y su distribución viene dada por $\mathbb{P}^-$.

El proceso ξ^- es una cadena de Markov estacionaria en $(\Omega,\mathcal{F},\mathbb{P})$ con generador

$$Q^- = (q_{i,j}^-)_{i,j\in S} = \left(\frac{\pi_j q_{j,i}}{\pi_i}\right)_{i,j\in S},$$

y distribución invariante $\Pi^- = \Pi$.

En adelante usamos la siguiente notación, para cualquier $s<t$ denotamos por $\mathcal{F}_s^t$ a la sub σ-álgebra generada por $\{\xi_u : s\le u\le y\}$, es decir,

$$\mathcal{F}_s^t = \sigma(\xi_u : s\le u\le t)$$
$$\mathbb{P}_{[s,t]} = \mathbb{P}|_{\mathcal{F}_s^t}$$
$$\mathbb{P}_{[s,t]}^- = \mathbb{P}^-|_{\mathcal{F}_s^t}.$$

Definicion 4.4. La *tasa de producción de entropía de una cadena de Markov estacionaria* ξ se define por

$$e_p = \lim_{t\to\infty} \frac{1}{t} H(\mathbb{P}_{[0,t]}, \mathbb{P}^-_{[0,t]}),$$

donde $H(\mathbb{P}_{[0,t]}, \mathbb{P}^-_{[0,t]})$ es la entropía relativa de $\mathbb{P}$ con respecto a $\mathbb{P}^-$ restringidas a la σ-álgebra $\mathcal{F}_0^t$.

Definicion 4.5. Para cualquier $t > 0$, $n \geq 0$ y $[i_0, i_1, \ldots, i_n] \in [S]$, denotamos por

$$A_{i_0,i_1,\ldots,i_n}(t) = \{\omega \in \Omega : n_t(\omega) = n, \xi_{T_k(\omega)}(\omega) = i_k, k = 0, 1, \ldots, n\},$$

al conjunto de trayectorias ω que hasta el tiempo t contienen n transiciones sobre el conjunto $i_0, i_1, \ldots, i_n$.

Lema 4.5. *Si la Q-matriz de ξ satisface*

$$q_{i,j} > 0 \Longleftrightarrow q_{j,i} > 0 \quad \forall i, j \in S,$$

entonces para todo $s, t \in \mathbb{R}$, $t > 0$ las medidas $\mathbb{P}_{[s,s+t]}$ y $\mathbb{P}^-_{[s,s+t]}$ son equivalentes, es decir, $\mathbb{P}_{[s,s+t]} \ll \mathbb{P}^-_{[s,s+t]}$ y $\mathbb{P}^-_{[s,s+t]} \ll \mathbb{P}_{[s,s+t]}$.

Demostración. Para $t > 0$, $n \geq 0$, $[i_0, i_1, \ldots, i_n] \in [S]$, $0 < t_1 < \ldots < t_n < t$ y δ suficientemente pequeño para que no haya traslapes escribimos

$$A = \{\omega \in A_{i_0,i_1,\ldots,i_n}(t) : t_1 < T_1(\omega) \leq t_1 + \delta t_1, \cdots, t_n < T_n(\omega) \leq t_n + \delta t_n\}.$$

Con esto,

$$\mathbb{P}(A) = \pi_{i_0}\tilde{p}_{i_0,i_1} \int_{t_1}^{t_1+\delta t_1} qi_0 e^{-q_{i_0}s_1} \tilde{p}_{i_1,i_2} \int_{t_2-s_1}^{t_2+\delta t_2 - s_1} q_{i_1} q^{-q_{i_1}s_2} \cdots \tilde{p}_{i_{n-1},i_n}$$
$$\int_{t_n-\sum_{k=1}^{n-1}s_k}^{t_n+\delta t_n-\sum_{k=1}^{n-1}s_k} q_{i_{n-1}} e^{-q_{i_{n-1}}s_n} \int_{t-\sum_{k=1}^{n}s_k}^{\infty} q_{i_n} e^{-q_{i_n}s_{n+1}} ds_{n+1} ds_n \cdots ds_2 ds_1$$
$$= \int_{t_1}^{t_1+\delta t_1} \cdots \int_{t_n-\sum_{k=1}^{n-1}s_k}^{t_n+\delta t_n-\sum_{k=1}^{n-1}s_k} \int_{t-\sum_{k=1}^{n}s_k}^{\infty} \pi_{i_0} q_{i_0,i_1} \cdots q_{i_{n-1}i_n} q_{i_n}$$
$$\prod_{k=0}^{n} e^{-q_{i_k}s_{k+1}} ds_{n+1} \cdots ds_1.$$

De manera similar calculando $\mathbb{P}^-(A)$ tenemos que

$$\mathbb{P}(A) = 0 \Longleftrightarrow q_{i_0,i_1} \cdots q_{i_{n-1}i_n} q_{i_{n-1},i_n} = 0$$
$$\Longleftrightarrow$$
$$q^-_{i_0,i_1} \cdots q^-_{i_{n-1}i_n} q^-_{i_{n-1},i_n} = 0 \Longleftrightarrow \mathbb{P}^-(A) = 0.$$

Aún más, para $\mathbb{P}(A) > 0$ usando la relación entre $q_{i,j}$ y $q^-_{i,j}$

$$\mathbb{P}(A) = \frac{\pi_{i_0} q_{i_0,i_1} \cdots q_{i_{n-1}i_n} q_{i_{n-1},i_n}}{\pi_{i_n} q_{i_n,i_{n-1}} \cdots q_{i_2,i_1} q_{i_1,i_0}} \mathbb{P}^-(A).$$

Como $\mathcal{F}_0^t \subset \sigma(n_t, \xi_0, T_i, \xi_{T_1}, \ldots, T_k, \xi_{T_k}, \ldots)$, esta última siendo generada por conjuntos del tipo de A, entonces tenemos que $\mathbb{P} \ll \mathbb{P}^-$ y $\mathbb{P}^- \ll \mathbb{P}$ en $\mathcal{F}_0^t$, en donde si $\omega \in A_{i_0,\ldots,i_1}(t)$

$$\frac{d\mathbb{P}_{[0,t]}}{d\mathbb{P}^-_{[0,t]}}(\omega) = \frac{\pi_{\xi_0(\omega)} q_{\xi_0(\omega),\xi_{T_1}(\omega)} \cdots q_{\xi_{T_{n-1}(\omega)},\xi_{T_n}(\omega)} q_{\xi_{T_{n-1}(\omega)},\xi_{T_n}(\omega)}}{\pi_{\xi_n(\omega)} q_{\xi_n(\omega),\xi_{T_{n-1}(\omega)}} \cdots q_{\xi_{T_2}(\omega),\xi_{T_1}(\omega)} q_{\xi_{T_1}(\omega),\xi_{T_0}(\omega)}} \quad \mathbb{P} - c.s.$$

□

Teorema 4.6. *Si ξ es una cadena de Markov estacionaria con generador $Q = (q_{i,j})_{i,j\in S}$, irreducible, positiva recurrente con espacio de estados S finito, y con distribución estacionaria*

$\Pi = (\pi_i)_{i \in S}$, *entonces su tasa de producción de entropía puede ser expresada como*

$$e_p = \frac{1}{2} \sum_{i,j \in S} (\pi_i q_{i,j} - \pi_j q_{j,i}) \log \frac{\pi_i q_{i,j}}{\pi_j q_{j,i}} \tag{9}$$

$$= \frac{1}{2} \sum_{c \in \mathcal{C}_\infty} (w_c - w_{c_-}) \log \frac{w_c}{w_{c_-}}. \tag{10}$$

Al igual que en el caso discreto tenemos condiciones equivalentes a la reversibilidad de a cadena de Markov en términos de su distribución circulatoria:

1) La cadena de Markov estacionaria ξ es reversible.
2) La cadena de Markov estacionaria ξ está en balance detallado, esto es,

$$\pi_i q_{i,j} = \pi_j q_{j,i} \quad \forall i, j \in S$$

3) La cadena de Markov estacionaria satisface el criterio de Kolmogorov

$$q_{i_1,i_2} q_{i_2,i_3} \cdots q_{i_{s-1},i_s} = q_{i_s,i_{s-1}} \cdots q_{i_3,i_2} q_{i_2,i_1}$$

   para cualquier colección finita $\{i_1, \dots, i_s\}$ de elementos distintos de S.
4) Los elementos de la distribución circulatoria $\{w_c : c \in \mathcal{C}_\infty\}$ satisfacen la condición de simetría

$$w_c = w_{c_-} \quad \forall c \in \mathcal{C}_\infty.$$

5) La tasa de producción de entropía de ξ es cero

$$e_p = 0.$$

Ejemplo 4.7.

Consideremos la cadena de Markov estacionaria, irreducible y recurrente ξ sobre el espacio de estados $S = \{0, 1, 2, 3\}$ con generador

$$Q = \begin{pmatrix} -(q_{0,1} + q_{0,2}) & q_{01} & q_{0,2} & 0 \\ q_{1,0} & -(q_{1,0} + q_{1,3}) & 0 & q_{1,3} \\ q_{2,0} & 0 & -(q_{2,0} + q_{2,3}) & q_{2,3} \\ 0 & q_{3,1} & q_{3,2} & -(q_{3,1} + q_{3,2}) \end{pmatrix},$$

donde las entradas indicadas son no nulas, y distribución invariante Π.

Los ciclos que aparecen en casi todas las trayectorias son

$$\begin{aligned} \mathcal{C}_\infty = \{&(0,1,2,3), (0,1,3,2), (0,2,1,3), (0,2,3,1), (0,3,1,2), (0,3,2,1), \\ &(0,1,2), (0,2,1), (0,1,3), (0,3,1), (1,2,3), (1,3,2), (0,2,3), \\ &(0,3,2), (0,1), (0,2), (0,3), (1,2), (1,3), (2,3)\}. \end{aligned}$$

Sabemos que los ciclos que contribuyen son aquellos de longitud 3 ó más, pues si $c = (i_1, i_2)$ entonces $c_- = (i_2, i_1)$, es decir, c y c_- son equivalentes en $\mathcal{C}_\infty$ al ser c_- una permutación cíclica de c y con ello $w_c = w_{c_-}$.

Calculando los pesos vemos que todos ellos se anulan excepto los correspondientes a los ciclos $\{(0, 1, 3, 2), (0, 2, 3, 1)\}$, de modo que en este ejemplo la tasa de producción de entropía es

$$\begin{aligned} e_p &= (w_{(0,2,3,1)} - w_{(0,1,3,2)}) \log \frac{w_{(0,2,3,1)}}{w_{(0,1,3,2)}} \\ &= -\frac{q_{0,2}q_{2,3}q_{3,1}q_{1,0} - q_{0,1}q_{1,3}q_{3,2}q_{2,0}}{\sum_{j=0}^{3} \tilde{D}(\{j\}^c)} \log \frac{q_{0,2}q_{2,3}q_{3,1}q_{1,0}}{q_{0,1}q_{1,3}q_{3,2}q_{2,0}}. \end{aligned}$$

Y la cadena ξ es reversible si y sólo si $q_{0,2}q_{2,3}q_{3,1}q_{1,0} = q_{0,1}q_{1,3}q_{3,2}q_{2,0}$.

Referencias

[1] Bolaños J., Quezada R., Producción de entropía en Cadenas de Markov, Tesis de Maestría, Posgrado en Matemáticas, UAM-I, México 2010.
[2] Richardson M., Domingos P., The intelligent surfer: Probabilistic combination of link and content information in PageRank, Advances in N.I.P.S. 14 (pp. 1441–1448), 2002.
[3] Langville A.N.,Meyer C.D., Deeper inside page rank, Internet Mathematics Vol 1 No. 23, 2003.
[4] Merrill S.J., Sathananthan S., Approximate Michaelis-Menten kinetics displayed in a stochastic model of cell-mediated cytotoxicity, Mathematical Biosciences Volume 80 Issue 2 ,1986.

[5] Kalpazidou S.L., Cycle Representations of Markov Processes, Springer, 2006.
[6] Qian M-P., Qian M., Jiang D-J., Mathematical Theory of Nonequilibrium Steady States, Springer, 2003.
[7] Norris J.R., Markov Chains, Cambridge Series in Statistical and Probabilistic Mathematics (No. 2), Cambridge University Press, 1997.
[8] Chung K.L., Markov Chains with Stationary Transition Probabilities, Grund. Math. Wiss 104, Springer-Verlag, Primera edición 1960.

Dirección del autor:
Jorge Bolaños Servín
Universidad Autónoma Metropolitana,
Unidad Iztapalapa,
División de Ciencias Básicas e Ingeniería,
Departamento de Matemáticas.
Av. San Rafael Atlixco 186, Col. Vicentina
Del. Iztapalapa, C.P. 09340 México, D.F.
e-mail: `kajit@gmail.com`