Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Matemática
Profesor Joaquín Fontbona T.
Auxiliar Héctor Olivero Q.
Semestre Primavera, 2009

PROCESOS DE MARKOV A TIEMPO Y ESTADOS DISCRETOS

Resumen. Las cadenas de Markov fueron introducidas en 1907 por el matemático ruso Andrei Andreevitch Markov (1856 – 1922), mientras buscaba condiciones más débiles para la demostración del Teorema Central del Límite. A pesar de su origen instrumental la teoría de cadenas de Markov se ha vuelto un tema de estudio en si misma por su amplio espectro de aplicación.
En este capítulo del curso introduciremos las cadenas de Markov y estudiaremos sus propiedades más básicas.

A. A. Марков (1886).

Índice

1. Introducción

En lo sucesivo consideraremos un conjunto numerable E y un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Supongamos que además tenemos una suceción de variables aleatorias definidas desde $(\Omega, \mathcal{F}, \mathbb{P})$ a E, es decir:

$$X_n : (\Omega, \mathcal{F}, \mathbb{P}) \to E, \;\; \forall\, n \in \mathbb{N}$$

Diremos que $(X_n)_{n\in\mathbb{N}}$ es un proceso estocástico a tiempo discreto a valores en E.

Definición 1. [Cadena de Markov]
Un proceso $(X_n)_{n\in\mathbb{N}}$ como el anterior, se dice *Cadena de Markov* si para todo $n \in \mathbb{N}$ y para todo $(i_0, \ldots, i_n) \in E^n$ tal que $\mathbb{P}(X_0 = i_0, \ldots, X_n = i_n) > 0$, se tiene que:

$$\mathbb{P}(X_{n+1} = i_{n+1}/X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1}/X_n = i_n) \tag{1}$$

Si además se tiene que $\forall\, i, j \in E$, $\forall n \in \mathbb{N}$:

$$\mathbb{P}(X_{n+1} = i/X_n = j) = \mathbb{P}(X_1 = i/X_0 = j) \tag{2}$$

Se dice que el proceso es una *Cadena de Markov Homogenea*. En este caso $(X_n)_{n\in\mathbb{N}}$ se puede caracterizar por el par $((\lambda_n)_{n\in E}, P = (p_{ij})_{i,j\in E})$ donde:

$$\begin{aligned} \lambda_i &= \mathbb{P}(X_0 = i) \\ p_{ij} &= \mathbb{P}(X_1 = i/X_0 = j) \end{aligned}$$

Diremos que λ es la distribución inicial de la cadena y que P es su matriz de transición.

Ejemplo 1. Supongamos que $(X_n)_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias independientes. Entonces $(X_n)_{n\in\mathbb{N}}$ es cadena de Markov.

$\square$

Antes de ver más ejemplos veamos un teorema que caracteriza a las cadenas de Markov homogeneas.

Teorema 1.
Sea λ un vector de probabilidad sobre E y sea P una matriz estocástica. Entonces $(X_n)_{n\in\mathbb{N}}$ es una cadena de Markov homogenea con distribución inicial λ y matriz de transición P si y sólo si:

$$\forall\, n \in \mathbb{N},\; \forall\, i_0, \ldots, i_n \in E : \mathbb{P}(X_0 = i_0, \ldots, X_n = i_n) = \lambda_{i_0} p_{i_0 i_1} p_{i_1 i_2} \ldots p_{i_{n-1} i_n} \tag{3}$$

Demostración.

$$\begin{aligned} \Rightarrow \mathbb{P}(X_0 = i_0, \ldots, X_n = i_n) &= \mathbb{P}(X_n = i_n/X_0 = i_0, \ldots, X_{n-1} = i_{n-1})\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) \\ &= \mathbb{P}(X_{n+1} = i_{n+1}/X_n = i_n)\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) \\ &= p_{i_n i_{n-1}}\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) \end{aligned}$$

Iterando este razonamiento encontramos que:

$$\begin{aligned} \mathbb{P}(X_0 = i_0, \ldots, X_n = i_n) &= p_{i_n i_{n-1}} p_{i_{n-1} i_{n-2}} \ldots p_{i_1 i_2} \mathbb{P}(X_1 = i_1/X_0 = i_0)\mathbb{P}(X_0 = i_0) \\ &= p_{i_n i_{n-1}} p_{i_{n-1} i_{n-2}} \ldots p_{i_1 i_2} p_{i_0 i_1} \lambda_{i_0} \end{aligned}$$

Demostremos ahora la otra implicancia:

$$\begin{aligned}
\mathbb{P}(X_{n+1}=i_{n+1}/X_n=i_n) &= \frac{\mathbb{P}(X_{n+1}=i_{n+1},X_n=i_n)}{\mathbb{P}(X_n=i_n)}\\
&= \frac{\sum_{k_0,\ldots,k_{n-1}\in E}\mathbb{P}(X_{n+1}=i_{n+1},X_n=i_n,X_{n-1}=k_{n-1},\ldots,X_0=k_0)}{\sum_{k_0,\ldots,k_{n-1}\in E}\mathbb{P}(X_n=i_n,X_{n-1}=k_{n-1},\ldots,X_0=k_0)}\\
&= \frac{\sum_{k_0,\ldots,k_{n-1}\in E}\lambda_{k_0}p_{k_0k_1}p_{k_1k_2}\ldots p_{k_{n-1}i_n}p_{i_ni_{n+1}}}{\sum_{k_0,\ldots,k_{n-1}\in E}\lambda_{k_0}p_{k_0k_1}p_{k_1k_2}\ldots p_{k_{n-1}i_n}}\\
&= p_{i_ni_{n+1}}\frac{\sum_{k_0,\ldots,k_{n-1}\in E}\lambda_{k_0}p_{k_0k_1}p_{k_1k_2}\ldots p_{k_{n-1}i_n}}{\sum_{k_0,\ldots,k_{n-1}\in E}\lambda_{k_0}p_{k_0k_1}p_{k_1k_2}\ldots p_{k_{n-1}i_n}}\\
&= p_{i_ni_{n+1}}
\end{aligned}$$

Por otro lado:

$$\begin{aligned}
\mathbb{P}(X_{n+1}=i_{n+1}/X_0=i_0,\ldots,X_n=i_n) &= \frac{\mathbb{P}(X_0=i_0,\ldots,X_n=i_n,X_{n+1}=i_{n+1})}{\mathbb{P}(X_0=i_0,\ldots,X_n=i_n)}\\
&= \frac{\lambda_{i_0}p_{i_0i_1}p_{i_1i_2}\ldots p_{i_{n-1}i_n}p_{i_ni_{n+1}}}{\lambda_{i_0}p_{i_0i_1}p_{i_1i_2}\ldots p_{i_{n-1}i_n}}\\
&= p_{i_ni_{n+1}}
\end{aligned}$$

Así tenemos que:

$$\mathbb{P}(X_{n+1}=i_{n+1}/X_0=i_0,\ldots,X_n=i_n)=\mathbb{P}(X_{n+1}=i_{n+1}/X_n=i_n)$$

$\square$

Ejemplo 2. Supongamos que $E=\mathbb{Z}$ y consideremos $(Y_n)_{n\in\mathbb{N}}$ sucesión de variables aleatorias independientes e identicamente distribuidas. Entonces $(X_n)_{n\in\mathbb{N}}$ definido como:

$$X_n=\begin{cases}0 & n=0\\ Y_1+\ldots+Y_n & n\geq 1\end{cases}$$

es proceso de markov. En efecto, definamos $\mu(k)=\mathbb{P}(Y_1=k)$ y $p_{ij}=\mu(j-i)$. Además notemos que $X_0=0$ es determinista, luego:

$$\begin{aligned}
\mathbb{P}(X_0=0,X_1=k_1,\ldots,X_n=k_n) &= \mathbb{P}(Y_1=k_1,Y_2=k_2-k_1,\ldots,Y_n=k_n-k_{n-1})\\
&= \mu(k_1-0)\mu(k_2-k_1)\ldots\mu(k_n-k_{n-1})\\
&= p_{0k_1}p_{k_1k_2}\ldots p_{k_{n-1}k_n}
\end{aligned}$$

Y como una trayectoria que no parte de 0 tiene probabilidad 0 tenemos que:

$$\mathbb{P}(X_0=0,X_1=k_1,\ldots,X_n=k_n)=\delta_0(k_0)p_{k_0k_1}p_{k_1k_2}\ldots p_{k_{n-1}k_n}$$

Luego por el teorema 1 $(X_n)_{n\in\mathbb{N}}$ es cadena de Markov homogenea con matriz de transición $P=(p_{ij})_{i,j\in\mathbb{Z}}$ y distribución inicial $\delta_0(\cdot)$.

$\square$

1.1. Representación Gráfica de las Cadenas de Markov Homogeneas.

Una cadena de Markov homogenea se puede representar naturalmente como un grafo dirigido, donde los nodos del grafo representan los estados del conjunto E y los arcos representan las posibles transiciones de la cadena, es decir existe un arco desde el nodo i al nodo j si $p_{ij} > 0$. Veamos un ejemplo. Consideremos $E = \{1, 2, 3\}$ y la matriz estocástica:

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 1 & 0 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

En este caso el grafo dirigido que representa la cadena es el siguiente:

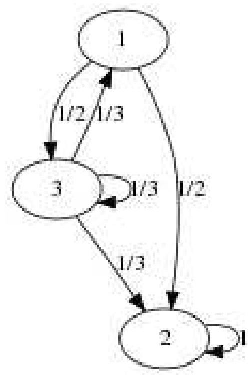

Figura 1.

Teorema 2. [Propiedad de Markov Simple:]
Un proceso $(X_n)_{n\in\mathbb{N}}$ es cadena de Markov si y sólo si $\forall\, n \in \mathbb{N}$ condicional en $\{X_n = i\}$ el proceso $(X_{n+m})_{m\in\mathbb{N}}$ es independiente de $\{X_0, \dots, X_{n-1}\}$.

Además si $(X_n)_{n\in\mathbb{N}}$ es cadena de Markov homogenea con matriz de transición P entonces condicional a $\{X_n = i\}$, $(X_{n+m})_{m\in\mathbb{N}}$ es cadena de Markov homogenea con distribución inicial δ_i y matriz de transición P.

Demostración. Probamos la doble implicancia:
$\Leftarrow$ Queremos probar que la independencia entre el pasado y el futuro condicional al presente implica que se satisfaga la definición de cadena de Markov:

$$\begin{aligned}
\mathbb{P}(X_{n+1} = i_{n+1}/X_0 = i_0, \dots, X_n = i_n) &= \frac{\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n, X_{n+1} = i_{n+1})}{\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n)} \\
&= \frac{\mathbb{P}(X_0 = i_0, \dots, X_n = i_n, X_{n+1} = i_{n+1}/X_n = i_n)\mathbb{P}(X_n = i_n)}{\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n)} \\
&= \frac{\mathbb{P}(X_0 = i_0, \dots, X_{n-1} = i_{n-1}/X_n = i_n)\mathbb{P}(X_n = i_n)}{\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n)} \times \\
&\quad \mathbb{P}(X_{n+1} = i_{n+1}/X_n = i_n) \\
&= \mathbb{P}(X_{n+1} = i_{n+1}/X_n = i_n)
\end{aligned}$$

$\Rightarrow$ Lo que queremos probar ahora es que para todo evento B que dependa sólo de $X_0, \dots, X_{n-1}$ se tiene que:

$$\begin{aligned}\mathbb{P}(B, X_{n+1} = i_{n+1}, \dots, X_{n+m} = i_{n+m}/X_n = i) &= \mathbb{P}(B/X_n = i) \times \\ &\quad \mathbb{P}(X_{n+1} = i_{n+1}, \dots, X_{n+m} = i_{n+m}/X_n = i)\end{aligned}$$

Se puede probar de que basta probar la igualdad anterior para evento B elementales, del tipo $B = X_0 = i_0, \dots, X_{n-1} = i_{n-1}$, luego nos restringiremos a este tipo de eventos. Luego, si llamamos $B_m = \{X_{n+1} = i_{n+1}, \dots, X_{n+m} = i_{n+m}\}$:

$$\begin{aligned}\mathbb{P}(B, B_m/X_n = i) &= \frac{\mathbb{P}(B, X_n = i, B_m)}{\mathbb{P}(X_n = i)} \\ &= \frac{\mathbb{P}(B_m/X_n = i, B)\mathbb{P}(X_n = i, B)}{\mathbb{P}(X_n = i)} \\ &= \mathbb{P}(B_m/X_n = i)\mathbb{P}(B/X_n = i)\end{aligned}$$

Donde la última igualdad se obtiene por la pérdida de memoria de las cadenas de Markov y la definición de probabilidad condicional.
Por último notemos que si la cadena es homogenea con matriz de transición P, entonces:

$$\mathbb{P}(B_m/X_n = i) = p_{i,i_{n+1}} p_{i_{n+1},i_{n+2}} \dots p_{i_{n+m-1},i_{n+m}}$$

Lo que corresponde a una cadena de Markov homogenea que parte "fresca" en i con matriz de transición P.

$\square$

2. Algunos Resultados sobre Cadenas de Markov Homogeneas

Definición 2. [Probabilidad de Transición en n pasos]
Dada una cadena de Markov homogenea $(X_n)_{n\in\mathbb{N}}$ la probabilidad de transición de i a j en n pasos la denotaremos por $(P^{(n)})_{ij}$ y corresponde a:

$$(P^{(n)})_{ij} = \mathbb{P}(X_n = j/X_0 = i)$$

Notemos que por ser la cadena homogenea se tiene que:

$$(P^{(n)})_{ij} = \mathbb{P}(X_{n+m} = j/X_m = i)$$

Proposición 1.
Sea $(X_n)_{n\in\mathbb{N}}$ una cadena de Markov homogenea $\sim (\lambda, P)$ entonces se tiene que:

i. $P^{(n)} = P^n$. En particular $P^{(n+m)} = P^{(n)}P^{(m)}$. Luego se satisface la ecuación de Chapman-Kolmogorov:

$$p_{ij}^{(n+m)} = \sum_{k\in E} p_{ik}^{(n)} p_{kj}^{(m)} \tag{4}$$

ii. $\mathbb{P}(X_n = i) = (\lambda^t P^n)_i$

iii. Si $f : E \to \mathbb{R}$ es acotada, entonces:

$$\mathbb{E}_i(f(X_n)) = (P^{(n)} f)_i$$

donde $\mathbb{E}_i(\cdot)$ es la esperanza bajo la ley $\mathbb{P}_i(\cdot) = \mathbb{P}(\cdot / X_0 = i)$.

iv. $\mathbb{E}(f(X_n)) = \lambda^t P^{(n)} f$

Demostración. i. Primero calculemos: $\mathbb{P}(X_n = j / X_0 = i)$.

$$\begin{aligned}
\mathbb{P}_i(X_n = j/X_0 = i) &= \sum_{k_1,\ldots,k_{n-1}\in E} \frac{\mathbb{P}(X_0 = i, X_1 = k_1, \ldots, X_{n-1} = k_{n-1}, X_n = j)}{\mathbb{P}(X_0 = i)} \\
&= \sum_{k_1,\ldots,k_{n-1}\in E} \frac{\lambda_i p_{i,k_1} p_{k_1,k_2} \ldots p_{k_{n-1},j}}{\lambda_i} \\
&= \sum_{k_1,\ldots,k_{n-1}\in E} p_{i,k_1} p_{k_1,k_2} \ldots p_{k_{n-1},j} \\
&= P^n_{ij}
\end{aligned}$$

Ejemplo 3. [Mutación de Virus:]
Supongamos que un virus puede existir en N cepas diferentes y en cada generación se mantiene como la misma cepa o muta con probabilidad α a una cepa diferente que se escoge al azar. ¿Cuál es la probabilidad de que la cepa a la cual pertenece el virus en el tiempo n sea la misma que en el tiempo 0?
COMPLETAR

3. Estructura de Clases

A veces es posible dividir una cadena de Markov en partes más pequeñas, en teoría más simples que la cadena completa. Esto se hace identificando las clases de estados comunicados presentes en la cadena. Diremos que i *lleva* a j lo que escribimos como $i \to j$ si

$$\mathbb{P}_i(\exists n \in \mathbb{N} : X_n = j) > 0$$

Diremos que i se *comunica* con j si $i \to j \wedge j \to i$.

Teorema 3.
Consideremos dos estados distintos $i, j \in E$. Entonces son equivalentes:

i. $i \to j$.
ii. $\exists\, n \in \mathbb{N}$, $\exists\, i_1, \ldots, i_{n-1} \in E : p_{ii_1} p_{i_1 i_2} \ldots p_{i_{n-1} j} > 0$.
iii. $p^{(n)}_{ij} > 0$ para algún $n \geq 0$.

Demostración. Basta con notar lo siguiente:

$$p^{(n)}_{ij} \leq \mathbb{P}_i(\exists n \in \mathbb{N} : X_n = j) \leq \sum_{n \geq 0} p^{(n)}_{ij}$$

De donde es clara la equivalencia entre (i) y (ii). Para probar que (ii) y (iii) son equivalentes basta notar que:

$$p_{ij}^{(n)} = \sum_{i_1, i_2, \ldots, i_{n-1}} p_{ii_1} p_{i_1 i_2} \ldots p_{i_{n-1} j}$$

□

Recordemos que $\forall i \in E : p_{ii}^{(0)} = 1 \Rightarrow i \to i$. Además de (ii) es claro que $i \to j \wedge j \to k \Rightarrow i \to k$. Luego $\leftrightarrow$ es una relación de equivalencia en E y entonces particiona E en clases.
Diremos que una clase C es *cerrada* si:

$$i \in C,\, i \to j \;\Rightarrow j \in C$$

Es decir una clase cerrada es una clase desde la que no hay escape. Un estado i es absorbente si $\{i\}$ es una clase cerrada. Una cadena en que E es la única clase se dice *irreducible*.

Ejemplo 4. Encuentre las clases de estados comunicados de la cadena de Markov homogenea asociada a la siguiente matriz de transición:

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

Es fácil encontrar las clases comunicadas observando el diagrama de estados de la cadena:

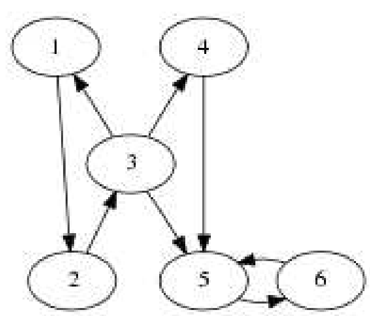

Figura 2.

4. Tiempos de Llegada y Probabilidad de Absorción

Sea $(X_n)_{n\in\mathbb{N}}$ una cadena de Markov homogenea con matriz de transición P. El tiempo de llegada al conjunto $A \subset E$ corresponde a la variable aleatoria:

$$\begin{aligned} H^A(\omega) : \Omega &\to \mathbb{N} \cup \{+\infty\} \\ H^A(\omega) &= \inf\{n \geq 0 : X_n(\omega) \in A\} \end{aligned}$$

Con la convención de que $\inf \emptyset = +\infty$. La probabilidad partiendo de $i \in E$ de que $(X_n)_{n\in\mathbb{N}}$ llegue alguna vez a A es entonces:

$$h_i^A = \mathbb{P}_i(H^A < +\infty)$$

Cuando A es una clase cerrada, h_i^A es llamada probabilidad de absorción. El tiempo medio que le toma a $(X_n)_{n\in\mathbb{N}}$ llegar a A está dado por:

$$k_i^A = \mathbb{E}_i(H^A) = \sum_{n<\infty} n\mathbb{P}_i(H^A = n) + \infty\mathbb{P}_i(H^A = \infty)$$

Con la convención de que $0 \cdot \infty = 0$.
Lo notable de los vectores $(h_i^A)_{i\in E}$ y $(k_i^A)_{i\in E}$ es que pueden ser calculados como la solución a ciertos problemas lineales. Antes de ver el resultado general, veamos un ejemplo.

Ejemplo 5. Consideremos la cadena de Markov homogenea dada por el siguiente diagrama de estados:
Partiendo de 2. ¿Cuál es la probabilidad de absorción en 4?¿Cuánto tiempo es necesario en

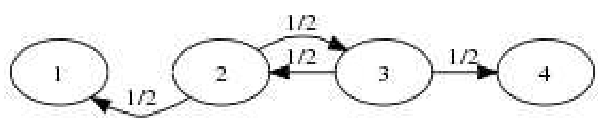

FIGURA 3.

promedio para que la cadena sea absorvida por $\{1,4\}$?. Sean:

$$h_i = \mathbb{P}_i(\text{llegar a } 4), \quad k_i = \mathbb{E}_i(\text{tiempo para llegar a } \{1,4\})$$

Claramente: $h_1 = 0$, $h_4 = 1$. Supongamos ahora que la cadena parte de 2, entonces:

$$\begin{aligned} h_2 &= \mathbb{P}_2(\text{llegar a } 4) \\ &= \mathbb{P}_2(\text{llegar a } 4/X_1 = 1)\mathbb{P}_2(X_1 = 1) + \mathbb{P}_2(\text{llegar a } 4/X_1 = 3)\mathbb{P}_2(X_1 = 3) \\ &= \frac{1}{2}\mathbb{P}_2(\text{llegar a } 4/X_1 = 1) + \frac{1}{2}\mathbb{P}_2(\text{llegar a } 4/X_1 = 3) \end{aligned}$$

Pero por la propiedad de Markov simple:

$$\mathbb{P}_2(\text{llegar a } 4/X_1 = 1) = h_1 \ \wedge \ \mathbb{P}_2(\text{llegar a } 4/X_1 = \ 1) \ = \ h_3$$

Luego:

$$h_2 = \frac{1}{2}h_1 + \frac{1}{2}h_3$$

Análogamente, encontramos que:

$$h_3 = \frac{1}{2}h_2 + \frac{1}{2}h_4$$

En resumen:

$$\begin{aligned} h_1 &= 0 \\ h_2 &= \frac{1}{2}h_1 + \frac{1}{2}h_3 \\ h_3 &= \frac{1}{2}h_2 + \frac{1}{2}h_4 \\ h_4 &= 1 \end{aligned}$$

Luego la probabilidad de llegar a 4 partiendo de 2 es 1/3.Estudiemos ahora el tiempo de absorción del conjunto $\{1,4\}$. Notemos que $k_1 = k_4 = 0$. Supongamos que empezamos en 2. Entonces:

$$\begin{aligned} k_2 &= \mathbb{E}_2(\text{tiempo para llegar a } \{1,4\}) \\ &= 1 + \mathbb{E}_2(\text{tiempo para llegar a } \{1,4\}/X_1 = 1)\mathbb{P}_2(X_1 = 1) \\ &\quad + \mathbb{E}_2(\text{tiempo para llegar a } \{1,4\}/X_1 = 3)\mathbb{P}_2(X_1 = 3) \\ &= 1 + \frac{1}{2}\mathbb{E}_2(\text{tiempo para llegar a } \{1,4\}/X_1 = 1) \\ &\quad + \frac{1}{2}\mathbb{E}_2(\text{tiempo para llegar a } \{1,4\}/X_1 = 3) \end{aligned}$$

Pero por la propiedad de Markov tenemos $\mathbb{E}_2(\text{tiempo para llegar a } \{1,4\}/X_1 = 1) = k_1$ y $\mathbb{E}_2(\text{tiempo para llegar a } \{1,4\}/X_1 = 3) = k_3$. Luego tenemos que:

$$k_2 = 1 + \frac{1}{2}(k_1 + k_3)$$

Análogamente:

$$k_3 = 1 + \frac{1}{2}(k_2 + k_4)$$

Con lo que podemos encontrar que $k_2 = 2$.

Ahora pasemos al resultado general:

Teorema 4.
El vector de probabilidades de llegada a A : $h^A = (h_i^A)_{i\in E}$ es la solución no negativa minimal del sistema:

$$(P_A)\left\{\begin{array}{lll} h_i^A & = & 1 \qquad\qquad i \in A \\ h_i^A & = & \sum_{j\in E} p_{ij}h_j^A \quad i \notin A \end{array}\right.$$

Donde la minimalidad se entiende como sigue: si $(x_i)_{i\in E}$ es otra solucion no negativa de (P_A) entonces $x_i \geq h_i^A \; \forall i \in E$.

Demostración. Primero probaremos que h^A satisface (P_A). Si $X_0 = i \in A$ entonces $H^A = 0$, luego $h_i^A = 1$. Si $X_0 = i \notin A$, entonces por la propiedad de Markov:

$$\mathbb{P}_i(H^A < \infty/X_1 = j) = \mathbb{P}_j(H^A < \infty) = h_j^A$$

Luego como:

$$h_i^A = \mathbb{P}_i(H^A < \infty) = \sum_{j\in E} \mathbb{P}_i(H^A < \infty / X_1 = j)\mathbb{P}_i(X_1 = j) = \sum_{j\in E} h_j^A p_{ij}$$

Veamos ahora que h^A es la solución minimal de (P_A). Para esto supongamos que $(x_i)_{i\in E}$ es otra solucion no negativa de (P_A) entonces:

- Para $i \in A,\ h_i^A = 1 = x_i$.
- Ahora supongamos que $i \notin A$, entonces:

  $$x_i = \sum_{j\in E} p_{ij}x_j = \sum_{j\in A} p_{ij} + \sum_{j\notin A} p_{ij}x_j$$

  Sustituyendo x_j obtenemos:

  $$\begin{aligned} x_i &= \sum_{j\in A} p_{ij} + \sum_{j\notin A} p_{ij}\left(\sum_{k\in A} p_{jk} + \sum_{k\notin A} p_{jk}x_k\right) \\ &= \sum_{j\in A} p_{ij} + \sum_{j\notin A} p_{ij}\sum_{k\in A} p_{jk} + \sum_{j\notin A} p_{ij}\sum_{k\notin A} p_{jk}x_k \\ &= \sum_{j\in A} p_{ij} + \sum_{j\notin A}\sum_{k\in A} p_{ij}p_{jk} + \sum_{j\notin A}\sum_{k\notin A} p_{ij}p_{jk}x_k \\ &= \mathbb{P}_i(X_1 \in A) + \mathbb{P}_i(X_1 \notin A, X_2 \in A) + \sum_{j\notin A}\sum_{k\notin A} p_{ij}p_{jk}x_k \\ &= \mathbb{P}_i(H^A = 1) + \mathbb{P}_i(H^A = 2) + \sum_{j\notin A}\sum_{k\notin A} p_{ij}p_{jk}x_k \\ &= \mathbb{P}_i(H^A \leq 2) + \sum_{j\notin A}\sum_{k\notin A} p_{ij}p_{jk}x_k \end{aligned}$$

  Si seguimos sustituyendo el x_k de la derecha $n-2$ veces más obtendríamos:

  $$x_i = \mathbb{P}_i(H^A \leq n) + \sum_{j_1\notin A} \dots \sum_{j_n\notin A} p_{ij_1}\dots p_{j_{n-1}j_n}x_{j_n}$$

  Como $(x_i)_{i\in E}$ es no negativo, el segundo término de la derecha también es no negativo, luego tenemos que:

  $$\forall\, n \in \mathbb{N}:\ x_i \geq \mathbb{P}_i(H^A \leq n)$$

  Entonces:

  $$x_i \geq \lim_{n\to\infty} \mathbb{P}_i(H^A \leq n) = \mathbb{P}_i(H^A < \infty) = h_i^A$$

□

Ejemplo 6. [La ruina del jugador]
Considere la cadena de Markov dada por el siguiente diagrama de estados:
donde $p = 1 - q < 1$. Las probabilidades de transición para $i = 1, 2, \dots$ son:

FIGURA 4.

$$\begin{aligned} p_{00} &= 1 \\ p_{i,i-1} &= q \\ p_{i,i+1} &= p \end{aligned}$$

Consideremos $A = \{0\}$ y notemos $h = h^A$. Del teorema sabemos que h es la solución minimal no negativa de:

$$\begin{aligned} h_0 &= 1 \\ h_i &= ph_{i+1} + qh_{i-1}\,\forall i = 1, 2, \ldots \end{aligned}$$

Esta ecuación por recurrencia la podemos resolver utilizandoe en lado izquierdo que $p+q = 1$:

$$\begin{aligned} h_i &= ph_{i+1} + qh_{i-1} \\ \Leftrightarrow (p+q)h_i &= ph_{i+1} + qh_{i-1} \\ \Leftrightarrow p(h_{i+1} - h_i) &= q(h_i - h_{i-1}) \\ \text{Sea: } a_i = h_i - h_{i-1} \Rightarrow a_{i+1} &= \frac{q}{p} a_i \;\; \forall\, i = 1, 2, \ldots \end{aligned}$$

Teorema 5.
El vector de tiempos esperados de llegada a A : $k^A = (k_i^A)_{i \in E}$ es la solución no negativa minimal del sistema:

$$(P'_A) \left\{ \begin{array}{rcll} k_i^A &=& 0 & i \in A \\ k_i^A &=& 1 + \sum_{j \notin A} p_{ij} k_j^A & i \notin A \end{array} \right.$$

5. Propiedad de Markov Fuerte

5.1. Tiempos de Parada.

Definición 3. [Tiempo de Parada]
Una variable aleatoria $T : \Omega \to \{0, 1, 2, \ldots\} \cup \{\infty\}$ es un tiempo de parada del proceso $(X_n)_{n \in \mathbb{N}}$ si el evento $\{T = n\}$ depende sólo de las variables $X_1, X_2, \ldots, X_n$ para $n = 0, 1, \ldots$ Intuitivamente, observando el proceso es posible determinar si T ha ocurrido.

Ejemplo 7. Son tiempos de parada:

i. El primer tiempo de pasada por un estado:

$$T_j = \text{ínf}\{n \geq 1 : X_n = j\}$$

Pues: $\{T_j = n\} = \{X_1 \neq j, \ldots, X_{n-1} \neq j, X_n = j\}$

ii. H^A tiempo de la primera llegada a $A \subset E$, definido en la sección anterior, pues:

$$\{H^A = n\} = \{X_0 \notin A, X_1 \notin A, \ldots, X_{n-1} \notin A, X_n \in A\}$$

5.2. Propiedad de Markov Fuerte.

Teorema 6. [Propiedad de Markov Fuerte]
Sea $(X_n)_{n\in\mathbb{N}}$ una cadena de Markov homogenea (λ, P) y sea T un tiempo de parada de $(X_n)_{n\in\mathbb{N}}$. Entonces, condicional en $\{T < \infty\}$ y $\{X_T = i\}$, $(X_{T+n})_{n\geq 0}$ es una cadena de Markov homogenea (δ_i, P) independiente de $X_0, X_1, \ldots, X_T$.

Demostración.

6. Introducción al Comportamiento en Tiempo Largo de Cadenas de Markov

6.1. Recurrencia y Transiencia.

Definición 4. [Transiencia y Recurrencia]
Sea $(X_n)_{n\in\mathbb{N}}$ una cadena de Markov homogenea con matriz de transición P decimos que un estado $i \in E$ es:

1. *Transiente* si:
$$\mathbb{P}(X_n = i, \text{ para infinitos } n) = 0$$
2. *Recurrente* si
$$\mathbb{P}(X_n = i, \text{ para infinitos } n) = 1$$

Luego un estado es recurrente si la cadena vuelve a él infinitas veces con probabilidad 1, mientras que un estado es transiente si la cadena eventualmente deja ese estado para siempre.

Nuestro primer objetivo es probar el siguiente teorema:

Teorema 7.
Se tiene la siguiente dicotomía:

i. Si $\mathbb{P}_i(T_i < \infty) = 1$ entonces i es recurrente y $\sum_{i=0}^{\infty} p_{ii}^{(n)} = \infty$.
ii. Si $\mathbb{P}_i(T_i < \infty) < 1$ entonces i es transiente y $\sum_{i=0}^{\infty} p_{ii}^{(n)} < \infty$.

Donde $T_i = \text{ínf}\{n \geq 1 : X_n = i\}$, es el primer tiempo de retorno a i.
En particular todo estado es transiente o recurrente.

Para demostrar este teorema necesitamos algunas definiciones previas y un par de lemas.

Definición 5. [Tiempos de retorno y largos de excursión]
Se define el r-ésimo tiempo de retorno por i, $T_i^{(r)}$ inductivamente de la siguiente forma:

$$T_i^{(r)} = \begin{cases} 0 & r = 0 \\ T_i & r = 1 \\ \text{ínf}\{n \geq T_i^{(r-1)} + 1 : X_n = i\} & r \geq 2 \end{cases}$$

Se define el largo de la r-ésima excursión a i, $S_i^{(r)}$ como:

$$S_i^{(r)} = \begin{cases} T_i^{(r)} - T_i^{(r-1)} & T_i^{(r-1)} < \infty \\ 0 & T_i^{(r-1)} = \infty \end{cases}$$

Lema 1.
Para $r = 2, 3, \ldots$, condicional en el evento $T_i^{(r-1)} < \infty$, $S_i^{(r)}$ es independiente de $\{X_m \,:\, m \leq T_i^{(r-1)}\}$ y

$$\mathbb{P}(S_i^{(r)} = n / T_i^{(r-1)} < \infty) = \mathbb{P}_i(T_i = n)$$

Demostración. Consideremos el tiempo de parada $T = T_i^{(r-1)}$. Luego por la definición de T, en el evento $T < \infty$ tenemos que $X_T = i$, además por la propiedad de Markov fuerte:

- $(X_{T+n})_{n\in\mathbb{N}}$ es cadena de Markov homogenea (δ_i, P).
- $(X_{T+n})_{n\in\mathbb{N}}$ es independiente de $X_0, \ldots, X_T$

Si ahora consideramos el proceso $(Y_n)_{n\in\mathbb{N}}$ definido como $Y_n = X_{T+n}$, condicional a $T < \infty$ tenemos que:

$$\begin{aligned} S_i^{(r)} &= T_i^{(r)} - T_i^{(r-1)} \\ &= \inf\{n \geq T_i^{(r-1)} + 1 : X_n = i\} - T_i^{(r-1)} \\ &= \inf\{n \geq 1 : X_{T+n} = i\} \\ &= \inf\{n \geq 1 : Y_n = i\} \end{aligned}$$

Luego $S_i^{(r)}$ es el primer tiempo de retorno a i de la cadena $(Y_n)_{n\in\mathbb{N}}$. Pero, condicional a $T < \infty$, $(Y_n)_{n\in\mathbb{N}}$ tiene la misma ley que $(X_n)_{n\in\mathbb{N}}$ condicional a $X_0 = i$. Luego:

$$\mathbb{P}(S_i^{(r)} = n / T_i^{(r-1)} < \infty) = \mathbb{P}_i(T_i = n)$$

$\square$

Introduzcamos ahora la variable V_i que representa el número de visitas a i que realiza el proceso y notemos que:

$$V_i = \sum_{n\geq 0} \mathbb{1}_{\{X_n = i\}}$$

Y además:

$$\mathbb{E}_i(V_i) = \mathbb{E}_i\left(\sum_{n\geq 0} \mathbb{1}_{\{X_n=i\}}\right) = \sum_{n\geq 0} \mathbb{E}_i(\mathbb{1}_{\{X_n=i\}}) = \sum_{n\geq 0} \mathbb{P}(X_n = i) = \sum_{n\geq 0} p_{ii}^{(n)}$$

Por último notemos, que un estado es recurrente si $\mathbb{P}_i(V_i = \infty) = 1$ y transiente si $\mathbb{P}_i(V_i = \infty) = 0$.

Lema 2.
Para $r = 0, 1, \ldots$, se tiene que $\mathbb{P}_i(V_i > r) = f_i^r$, donde $f_i = \mathbb{P}_i(T_i < \infty)$.

Demostración. Primero notemos que:

$$\mathbb{P}_i(V_i > 0) \geq \mathbb{P}_i(V_i = 1) = 1 = f_i^0$$

Ahora procedamos por inducción. Para esto supongamos que $\forall\, k = 0, \ldots, r : \ \mathbb{P}_i(V_i > k) = f_i^k$. Entonces:

$$\begin{aligned}
\mathbb{P}_i(V_i > r+1) &= \mathbb{P}_i(T_i^{(r)} < \infty, S_i^{(r+1)} < \infty) \\
&= \mathbb{P}_i(S_i^{(r+1)} < \infty / T_i^{(r)} < \infty)\mathbb{P}_i(T_i^{(r)} < \infty) \\
&= \mathbb{P}_i(T_i < \infty)\mathbb{P}_i(V_i > r) \text{ por el lema anterior.} \\
&= \mathbb{P}_i(V_i > 1)\mathbb{P}_i(V_i > r) \\
&= f_i \cdot f_i^r \text{ por H.I.}
\end{aligned}$$

□

Demostración. (del Teorema 7)
Antes de pasar a la demostración del teorema, recordemos que si X es una variable aleatoria con valores enteros no negativos se tiene que:

$$\sum_{r\geq 0} \mathbb{P}(X > r) = \sum_{r\geq 0} \sum_{v \geq r+1} \mathbb{P}(X = v) = \sum_{v\geq 1} \sum_{r=0}^{v-1} \mathbb{P}(X = v) = \sum_{v\geq 1} v\mathbb{P}(X = v) = \mathbb{E}(X)$$

Ahora pasemos al teorema:

i. Si $\mathbb{P}_i(T_i < \infty) = 1$ entonces por el lema anterior tenemos que $\mathbb{P}_i(V_i > r) = 1, \forall\, r \geq 1$. Entonces:

$$\lim_{r\to\infty} \mathbb{P}_i(V_i > r) = 1$$

Pero como los eventos $\{V_i > r\}$ son decrecientes tenemos que:

$$\lim_{r\to\infty} \{V_i > r\} = \{V_i = \infty\}$$

Lo que sumado a lo anterior nos dice que: $\mathbb{P}_i(V_i = \infty) = 1$. Por lo tanto, i es recurrente. Además:

$$\mathbb{E}_i(V_i) = \infty = \sum_{n\geq 0} p_{ii}^{(n)}$$

ii. Si $\mathbb{P}_i(T_i < \infty) < 1$, entonces $f_i < 1$, luego:

$$\sum_{n\geq 0} p_{ii}^{(n)} = \mathbb{E}_i(V_i) = \sum_{r\geq 0} \mathbb{P}_i(V_i > r) = \sum_{r\geq 0} f_i^r = \frac{1}{1 - f_i} < \infty$$

De lo que concluimos que $\mathbb{P}_i(V_i = \infty) = 0$. Es decir i es transiente.

□

En lo que sigue consideramos $(X_n)_{n\in\mathbb{N}}$ una cadena de Markov homogenea de matriz de transición P.

Teorema 8.
Sea C una clase comunicada de $(X_n)_{n\in\mathbb{N}}$ entonces, todos los estados de C son transientes o recurrentes.

Demostración. Sean $i,j \in C$ entonces existen $n,m \in \mathbb{N}$ tales que: $p_{ij}^{(n)} > 0$ y $p_{ji}^{(m)} > 0$. Luego por Chapman-Kolmogorov, $\forall\, r \geq 0$:

$$p_{ii}^{(n+r+m)} \geq p_{ij}^{(n)} p_{jj}^{(r)} p_{ji}^{(m)}$$

De donde tenemos:

$$\sum_{r\geq 0} p_{jj}^{(r)} \leq \frac{1}{p_{ij}^{(n)} p_{ji}^{(m)}} \sum_{r\geq 0} p_{ii}^{(n+r+m)}$$

Haciendo lo mismo con j en lugar de i podemos encontrar:

$$A_1 + B_1 \sum_{r\geq 0} p_{ii}^{(r)} \leq \sum_{r\geq 0} p_{jj}^{(r)} \leq A_2 + B_2 \sum_{r\geq 0} p_{ii}^{(r)}$$

De donde es directo que la serie para i converge si y sólo si la serie para j converge. Luego el estado i es transiente *(resp. recurrente)* si y sólo si el estado j es transiente *(resp. recurrente)*.

□

Teorema 9.
Toda clase recurrente es cerrada.

Demostración. COMPLETAR

□

Teorema 10.
Toda clase finita cerrada es recurrente.

Demostración. COMPLETAR

□

Teorema 11.
Supongamos que P es irreductible y recurrente. Entonces para todo $j \in E$ se tiene que $\mathbb{P}(T_j < \infty) = 1$.

Demostración. COMPLETAR

□

6.2. Recurrencia y Transiencia de Paseos Aleatorios.

6.2.1. Paseo Aleatorio en $\mathbb{Z}$.

6.2.2. Paseo Aleatorio en $\mathbb{Z}^2$.

7. Referencias

- Markov Chains - James R. Norris - Cambrige Series in Statistical and Probabilistic Mathematics