

Cadenas de Markov en tiempo continuo

Santiago de Diego, Jesús Bueno, Fernando de la Cruz, Javier Ruiz

Contents

1	Introducción	3
1.1	Cadenas de Markov	3
1.2	Diferencia entre tiempo discreto y tiempo continuo	3
2	Definición y propiedades	3
2.1	Definición	3
2.2	Propiedades	4
2.2.1	Primera propiedad	4
2.2.2	Segunda propiedad	5
2.3	Relación con la teoría de autómatas	6
3	Probabilidades de transición	8
3.1	Probabilidades de transición y matriz de transición	8
3.2	Q-matriz o generador infinitesimal	10
4	Ecuación de Kolmogorov	12
5	Probabilidades absolutas y distribución estacionaria	14
6	Clasificación de los estados	15
6.1	Clasificación según ξ_i	15
6.2	Clasificación según α_{ij}	16
6.3	Clasificación según la posibilidad de ir de un estado a otro . . .	18
7	Teoremas límite	20
8	Aplicaciones	20
9	Ejemplo	21

1 Introducción

1.1 Cadenas de Markov

Primero de todo, presentamos una introducción a las Cadenas de Markov de forma genérica. Como ya sabemos, las cadenas de Markov son procesos de corta memoria en el sentido de que solo recuerdan el último estado visitado para decidir cual sería el próximo. En procesos con larga memoria el valor que toma el proceso en cada paso depende de todo el pasado.

Formalizando el concepto, el proceso $\{\mathbb{X}_n\}_{n \in \mathbb{N}}$ con espacio de estados E es una cadena de Markov si:

$$P(\mathbb{X}_{n+1} = y \mid \mathbb{X}_n = x_n, \dots, \mathbb{X}_0 = x_0) = P(\mathbb{X}_{n+1} = y \mid \mathbb{X}_n = x_n)$$

Este tipo de procesos estocásticos tienen mucho interés a la hora de modelar determinados fenómenos, como por ejemplo el tiempo de espera a un servidor en función de la tasa de llegada de los clientes.

1.2 Diferencia entre tiempo discreto y tiempo continuo

La principal diferencia entre cadenas de Markov en tiempo discreto y tiempo continuo es, como dice el propio nombre, el tiempo. En las cadenas de Markov en tiempo continuo, consideramos un $t \in T \subset \mathbb{R}$ mientras que en las cadenas de Markov en tiempo discreto, trabajamos con instantes de tiempo de la forma $t \in \mathbb{N}$.

2 Definición y propiedades

2.1 Definición

Primero de todo, definiremos una Cadena de Markov como:

Definición 2.1 *Cadena de Markov en tiempo continuo*

Sea $\{\mathbb{X}_t\}_{t \geq 0}$ un proceso estocástico en tiempo continuo, esto es, $t \in [0, T]$ con $T \in \mathbb{R}$ que toma valores en un conjunto numerable E . Decimos que $\{\mathbb{X}_t\}_{t \geq 0}$ es una Cadena de Markov en tiempo continuo si $\forall t, s \geq 0$ y $\forall i, j, x_u \in E$ con $0 \leq u < s$, se cumple que:

$$P(\mathbb{X}_{t+s} = j \mid \mathbb{X}_s = i, \mathbb{X}_u = x_u \forall 0 \leq u < s) = P(\mathbb{X}_{t+s} = j \mid \mathbb{X}_s = i)$$

Es decir, podemos definir una Cadena de Markov como un proceso estocástico en el que el futuro sólo depende del presente, independientemente de

sus estados pasados. Además también podemos considerar la definición alternativa:

Definición 2.2 Cadena de Markov en tiempo continuo

El proceso estocástico $\{\mathbb{X}_t, t \in [0, \infty]\}$ es una Cadena de Markov en tiempo continuo si para cualquier entero $n \geq 0$, cualesquiera $0 \leq t_0 < t_1 < \dots < t_{n+1}$ y $i_0, \dots, i_n, i_{n+1} \in S$ se verifica:

$$P(\mathbb{X}_{t_{n+1}} = i_{n+1} \mid \mathbb{X}_{t_0} = i_0, \dots, \mathbb{X}_{t_n} = i_n) = P(\mathbb{X}_{t_{n+1}} = i_{n+1} \mid \mathbb{X}_{t_n} = i_n)$$

Además, introducimos también el concepto de Cadena de Markov en tiempo continuo homogénea:

Definición 2.3 Cadena de Markov homogénea en tiempo continuo

Una CMTC se dice homogénea si la probabilidad de ir del estado i al estado j no depende del instante de tiempo en el que se encuentra la cadena, formalmente esto es:

$$P(\mathbb{X}_{t_n} = j \mid \mathbb{X}_{t_{n-1}} = i) = P(\mathbb{X}_{t_1} = j \mid \mathbb{X}_{t_0} = i)$$

Una vez vistas estas definiciones podemos ver las propiedades de una Cadena de Markov.

2.2 Propiedades

Podemos notar dos propiedades fundamentales en cuanto a CMTC, que son:

2.2.1 Primera propiedad

$$P(\mathbb{X}_{t_{n+h}} = i_{n+h}, h = 1, \dots, m \mid \mathbb{X}_{t_k} = i_k, k = 0, \dots, n) = P(\mathbb{X}_{t_{n+h}} = i_{n+h}, h = 1, \dots, m \mid \mathbb{X}_{t_n} = i_n) \\ \forall 0 \leq t_1 < t_2, \dots, < t_n + m, \forall i_k \in S, k = 0, \dots, n + m$$

Demostración 2.1 *Vamos a demostrarlo por inducción:*

Primero hacemos el caso $h = 1$:

$$P(\mathbb{X}_{t_{n+1}} = i_{n+1}, h = 1 \mid \mathbb{X}_{t_k} = i_k, k = 0, \dots, n) = P(\mathbb{X}_{t_{n+1}} = i_{n+1} \mid \mathbb{X}_{t_n} = i_n) \\ \forall 0 \leq t_1 < t_2 < \dots < t_{n+m} \forall m \geq 1 \text{ y } \forall i_k \in S, k = 0, \dots, n + m$$

Lo suponemos cierto para h y lo probamos para $h + 1$:

$$P(\mathbb{X}_{t_{n+h}} = i_{n+h}, h = 1, \dots, m \mid \mathbb{X}_{t_k} = i_k, k = 0, \dots, n) P(\mathbb{X}_{t_k} = i_k, k = 0, \dots, n)$$

$$\begin{aligned}
&= P(\mathbb{X}_{t_k} = i_k, k = 0, \dots, n+m+1) = P(\mathbb{X}_{t_{n+m-1}} = i_{n+m-1} | \mathbb{X}_{t_k} = i_k, k = 0, \dots, n+m) \\
&= P(\mathbb{X}_{t_k} = i_k, k = 0, \dots, n+m) = P(\mathbb{X}_{t_{n+m+1}} = i_{n+m+1} | \mathbb{X}_{t_k} = i_k, k = 0, \dots, n+m) \\
&= P(\mathbb{X}_{t_{n+h}} = i_{n+h}, h = 1, \dots, m | \mathbb{X}_{t_k} = i_k, k = 0, \dots, n) P(\mathbb{X}_{t_k} = i_k, k = 0, \dots, n) \\
&= P(\mathbb{X}_{t_{n+m+1}} = i_{n+m+1} | \mathbb{X}_{t_k} = i_k, k = 0, \dots, n+m) P(\mathbb{X}_{t_{n+h}} = i_{n+h}, h = 2, \dots, m | \mathbb{X}_{t_k} = i_k, k = 0, \dots, n) P(\mathbb{X}_{t_k} = i_k, k = 0, \dots, n)
\end{aligned}$$

Obtenemos que:

$$P(\mathbb{X}_{t_{n+h}} = i_{n+h}, h = 1, \dots, m+1 | \mathbb{X}_{t_k} = i_k, k = 0, \dots, n) = P(\mathbb{X}_{t_{n+m+1}} = i_{n+m+1}) P(\mathbb{X}_{t_{n+h}} = i_{n+h}, h = 1, \dots, m | \mathbb{X}_{t_n} = i_n)$$

Sin más que aplicar la definición tenemos que lo anterior es igual a $P(\mathbb{X}_{t_{n+m+1}} | \mathbb{X}_{t_n} = i_n)$ ■

2.2.2 Segunda propiedad

Sea \mathbb{X}_t con $t \in [0, +\infty)$ una cadena de Markov en tiempo continuo, se tiene:

$$P(\mathbb{X}_{t_{n+h}} = i_{t_{n+h}} | \mathbb{X}_{t_k} = i_k, k = 0, \dots, n) = P(\mathbb{X}_{t_{n+h}} = i_{t_{n+h}} | \mathbb{X}_{t_n} = i_n) \quad \forall h \geq 1$$

Demostración 2.2 Lo haremos por inducción sobre el número de instantes de tiempo, tal y como en la primera propiedad.

Demostremoslo para $h = 1$

$$P(\mathbb{X}_{t_{n+1}} = i_{t_{n+1}} | \mathbb{X}_{t_k} = i_k, k = 0, \dots, n) \stackrel{*}{=} P(\mathbb{X}_{t_{n+h}} = i_{t_{n+h}} | \mathbb{X}_{t_n} = i_n)$$

Supongámoslo demostrado para $h - 1$ y demostrémoslo para h .

$$\begin{aligned}
P(\mathbb{X}_{t_{n+h}} = i_{t_{n+h}} | \mathbb{X}_{t_k} = i_k, k = 0, \dots, n) &= P(\mathbb{X}_{t_{n+h}} = i_{t_{n+h}}, \bigcup_{i_{n+h-1} \in S} \mathbb{X}_{t_{n+h-1}} = i_{n+h-1}, \dots, \bigcup_{i_{n+1} \in S} \mathbb{X}_{t_{n+1}} = i_{n+1} | \mathbb{X}_{t_k} = i_k, k = 0, \dots, n) \\
&= \sum_{i_{n+h-1} \in S} \sum_{i_{n+1} \in S} P(\mathbb{X}_{t_{n+h}} = i_{t_{n+h}}, \mathbb{X}_{t_{n+h-1}} = i_{n+h-1}, \dots, \mathbb{X}_{t_{n+1}} = i_{n+1} | \mathbb{X}_{t_k} = i_k, k = 0, \dots, n) \stackrel{*}{=} \sum_{i_{n+h-1} \in S} \sum_{i_{n+1} \in S} P(\mathbb{X}_{t_{n+h}} = i_{t_{n+h}}, \mathbb{X}_{t_{n+h-1}} = i_{n+h-1}, \dots, \mathbb{X}_{t_{n+1}} = i_{n+1} | \mathbb{X}_{t_n} = i_n) \\
&= P(\mathbb{X}_{t_{n+h}} = i_{t_{n+h}}, \bigcup_{i_{n+h-1} \in S} \mathbb{X}_{t_{n+h-1}} = i_{n+h-1}, \dots, \bigcup_{i_{n+1} \in S} \mathbb{X}_{t_{n+1}} = i_{n+1} | \mathbb{X}_{t_n} = i_n) = P(\mathbb{X}_{t_{n+h}} = i_{t_{n+h}} | \mathbb{X}_{t_n} = i_n)
\end{aligned}$$

Nota: "*" significa "por definición de cadena de Markov". ■

2.3 Relación con la teoría de autómatas

En concreto, podemos asociar a una cadena de Markov un grafo, como veremos más abajo. De esta forma, el grafo resultante se comporta de igual manera que un **autómata finito determinista**, relacionándose de esta forma las cadenas de Markov con la teoría de autómatas.

Primero de todo, vamos a definir el grafo de una CMTC ya que resulta fundamental para entender lo siguiente:

Definición: Grafo asociado a una CMTC

$G = (E, U, W)$ es el grafo asociado a una CMTC sii:

- E es el conjunto de estados de la cadena
- U es el conjunto de aristas, que en este caso sería el conjunto de transiciones posibles
- W es el conjunto de ponderaciones. Podemos verlo como el conjunto de valores de cada arista

Además G es un grafo orientado, es decir, que tenemos en cuenta cual es el nodo origen y cual es el nodo destino.

Una vez presentada la definición de grafo asociado a una CMTC, podemos establecer una biyección entre el conjunto de grafos orientados asociados a una CMTC y el conjunto de los autómatas finitos deterministas, aunque estos resultados se escapan del alcance del trabajo en cuestión. Simplemente mostraremos un ejemplo de uno de estos grafos:

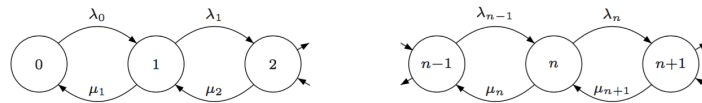


Figure 1: Ejemplo de grafo de una CMTC

En el grafo de la izquierda aparecen 3 estados, que corresponden a las variables aleatorias \mathbb{X}_{t_1} , \mathbb{X}_{t_2} y \mathbb{X}_{t_3} y representadas como λ_0 , λ_1 , μ_1 , μ_2 aparecen las 4 probabilidades de transición. Por último, las flechas indican la dirección en la que se produce la transición.

El grafo de la derecha es idéntico pero corresponde a las variables aleatorias $\mathbb{X}_{t_{n-1}}$, \mathbb{X}_{t_n} y $\mathbb{X}_{t_{n+1}}$. Además, según lo visto anteriormente tenemos que el grafo corresponde a la manera usual de representar un autómata finito determinista.

Para concluir presentaremos la definición de autómata finito determinista, por si el lector está interesado en el estudio del mismo:

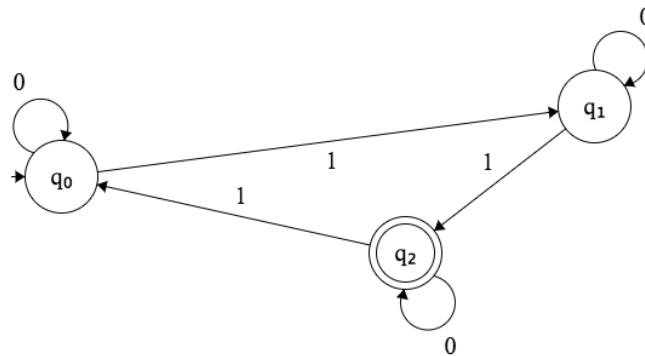
Definición de autómata finito determinista:

Un AFD es una 5-tupla $(Q, \Sigma, q_0, \delta, F)$ donde:

- Q es un conjunto de estados
- Σ es un alfabeto
- q_0 es el estado inicial
- $\delta : Q \times \Sigma \rightarrow Q$ es la función de transición
- $F \subseteq Q$ es el conjunto de estados finales

Además verifica que $\delta(q, a) = q_1$ y $\delta(q, a) = q_2 \Rightarrow q_1 \neq q_2$ y que no existen transiciones de la forma $\delta(q, \epsilon)$ donde ϵ es la cadena vacía.

Veamos un ejemplo de un autómata finito determinista muy simple:



Observando la definición y aplicándola a este caso concreto podemos diferenciar:

- $Q = \{q_0, q_1, q_2\}$
- $\Sigma = \{0, 1\}$

- $q_0 = q_0$
- Función de transición $\delta t.q$:

$$\begin{array}{lll} \delta(q_0, 0) = q_0 & \delta(q_1, 0) = q_1 & \delta(q_2, 0) = q_2 \\ \delta(q_0, 1) = q_1 & \delta(q_1, 1) = q_2 & \delta(q_2, 1) = q_0 \end{array}$$
- $F = q_2$

De esta forma, queda clara la correspondencia entre AFD y CMTC, sin más que comparar la definición de grafo de una CMTC con la definición de AFD.

Podemos ver las correspondencias: $Q = E$, $\sum = W$ y $|\delta| = U$ y además el conjunto de estados finales es el vacío, es decir, $F = \emptyset$.

3 Probabilidades de transición

Comenzamos definiendo las llamadas probabilidades de transición. Posteriormente, a partir de estas probabilidades, y restringiéndonos únicamente al caso en que la CMTC sea homogénea, construiremos la matriz de transición $P(t)$ y demostraremos sus propiedades más importantes.

Por último, partiendo de $P(t)$, generaremos y estudiaremos una nueva matriz, denominada Q -matriz o generador infinitesimal.

3.1 Probabilidades de transición y matriz de transición

Definición: Probabilidades de transición

Dada una CMTC $\{\mathbb{X}_t, t \in [0, \infty]\}$, definimos las probabilidades de transición como:

$$p_{ij}(s, t) := P(\mathbb{X}_t = j \mid \mathbb{X}_s = i), \quad 0 \leq s < t, \quad i, j \in S$$

A partir de este momento, supondremos que la CMTC es homogénea. Teniendo en cuenta la expresión anterior y siguiendo la definición de Cadena de Markov homogénea en tiempo continuo, citada en la sección (2.1), tenemos que la probabilidad de ir del estado i al estado j no depende del instante de tiempo en el que se encuentra la cadena, es decir, $p_{ij}(s, t)$ no depende de s y t , sólo depende de la diferencia $t - s$:

$$p_{ij}(s, t) = P(\mathbb{X}_t = j \mid \mathbb{X}_s = i) = P(\mathbb{X}_{t-s} = j \mid \mathbb{X}_0 = i) = p_{ij}(0, t-s) = p_{ij}(t-s)$$

Por tanto, podemos expresar las probabilidades de transición como:

$$p_{ij}(t) = P(\mathbb{X}_t = j \mid \mathbb{X}_0 = i) = P(\mathbb{X}_{t+s} = j \mid \mathbb{X}_s = i)$$

siendo t la diferencia entre los dos instantes de tiempo. En particular:

$$p_{ij}(0) = P(\mathbb{X}_0 = j \mid \mathbb{X}_0 = i) = \delta_{ij} := \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Estamos en condiciones de construir la matriz de transición, cuyos elementos son todas las posibles probabilidades de transición:

$$P(t) := (p_{ij}(t))_{i,j \in S}$$

Propiedades:

1. Para cada $t \geq 0$, $p_{ij}(t) \geq 0$, $\forall i, j \in S$,
2. Para cada $t \geq 0$, $\sum_{j \in S} p_{ij}(t) = 1$, $\forall i \in S$,
3. Ecuación de *Chapman-Kolmogorov*: $p_{ij}(s+t) = \sum_{k \in S} p_{ik}(s)p_{kj}(t)$, $\forall i, j \in S$, $\forall s, t \in [0, \infty)$.

Demostración 3.1 1. *Evidente, pues P es probabilidad.*

2. *Partiendo del hecho de que $P[\Omega] = P[\{\omega \in \Omega / \mathbb{X}_t(\omega) \in S\}] = P\left[\bigcup_{j \in S} \{\omega \in \Omega / \mathbb{X}_t(\omega) = j\}\right] = P\left[\bigcup_{j \in S} [\mathbb{X}_t = j]\right] = 1$, tenemos:*

$$1 = P\left[\bigcup_{j \in S} [\mathbb{X}_t = j] \mid \mathbb{X}_0 = i\right] \stackrel{*}{=} \sum_{j \in S} P[\mathbb{X}_t = j \mid \mathbb{X}_0 = i] - \underbrace{P\left[\bigcap_{j \in S} [\mathbb{X}_t = j] \mid \mathbb{X}_0 = i\right]}_0 = \sum_{j \in S} P[\mathbb{X}_t = j \mid \mathbb{X}_0 = i] = \sum_{j \in S} p_{ij}(t).$$

$$(*) P(A_1 \cup A_2 \cup \dots \cup A_n \mid B) = P(A_1 \mid B) + P(A_2 \mid B) + \dots + P(A_n \mid B) - P(A_1 \cap A_2 \cap \dots \cap A_n \mid B).$$

$$\begin{aligned} 3. p_{ij}(s+t) &= P[\mathbb{X}_{s+t} = j \mid \mathbb{X}_0 = i] = P\left[\mathbb{X}_{s+t} = j, \bigcup_{k \in S} [\mathbb{X}_s = k] \mid \mathbb{X}_0 = i\right] = \\ &= \sum_{k \in S} P[\mathbb{X}_{s+t} = j, \mathbb{X}_s = k \mid \mathbb{X}_0 = i] \stackrel{**}{=} \\ &= \frac{P[\mathbb{X}_{s+t} = j, \mathbb{X}_s = k, \mathbb{X}_0 = i]}{P[\mathbb{X}_0 = i]} \stackrel{**}{=} \end{aligned}$$

$$\begin{aligned}
&= \sum_{k \in S} P[\mathbb{X}_{s+t} = j \mid \mathbb{X}_s = k, \mathbb{X}_0 = i] P[\mathbb{X}_s = k \mid \mathbb{X}_0 = i] \stackrel{CMTC}{=} \\
&= \sum_{k \in S} P[\mathbb{X}_{s+t} = j \mid \mathbb{X}_s = k] P[\mathbb{X}_s = k \mid \mathbb{X}_0 = i] \stackrel{Homogénea}{=} \\
&= \sum_{k \in S} P[\mathbb{X}_t = j \mid \mathbb{X}_0 = k] P[\mathbb{X}_s = k \mid \mathbb{X}_0 = i] = \sum_{k \in S} p_{kj}(t) p_{ik}(s).
\end{aligned}$$

$$(**) P(A \cap B \mid C) = \frac{P(A \cap B \cap C)}{P(C)} = P(A \mid B \cap C) P(B \mid C)$$

■

3.2 Q-matriz o generador infinitesimal

Definición: Q-matriz o generador infinitesimal

Dada una CMTC $\{\mathbb{X}_t, t \in [0, \infty]\}$, con matriz de transición $P(t)$. Se define la Q-matriz o generador infinitesimal como:

$$Q := \lim_{t \rightarrow 0} \frac{P(t) - I}{t},$$

donde I es la matriz identidad del tamaño correspondiente y, por tanto, los elementos de la diagonal principal se expresan como:

$$q_{ii} = \lim_{t \rightarrow 0} \frac{p_{ii}(t) - 1}{t} := -q_i,$$

donde q_i es la *razón de escape* del estado i , mientras que el resto de elementos:

$$q_{ij} = \lim_{t \rightarrow 0} \frac{p_{ij}(t)}{t}, \quad i \neq j$$

donde q_{ij} es llamado *razón de transición* del estado i al j .

En lo que resta de sección, procederemos a construir esta matriz a partir de la definición de una nueva variable aleatoria. A partir de este momento, asumimos que la CMTC $\{\mathbb{X}_t, t \in [0, \infty]\}$ es continua por la derecha, es decir, si la transición de un estado $i \in S$ a un estado $j \in S$ ocurre el instante t , entonces se tiene que $\mathbb{X}_t = j$ y $\mathbb{X}_{t-} = i$.

La cuestión es: ¿Dado $\mathbb{X}_0 = i$, cuánto tiempo permanecerá el proceso en el estado i ? Para responder a esta pregunta, definimos la siguiente variable aleatoria:

$\xi_i \equiv$ Tiempo que el proceso permanece en el estado i .

Vamos a deducir la distribución que sigue ξ_i . Para ello será suficiente con probar que la variable ξ_i verifica la propiedad de falta de memoria, en efecto, dados $t, s \geq 0$:

$$\begin{aligned}
P[\xi_i > s + t \mid \xi_i > s] &= \\
&= P[\mathbb{X}_r = i, \text{ para } r \in [0, s + t] \mid \mathbb{X}_r = i, \text{ para } r \in [0, s]] = \\
&= P[\mathbb{X}_r = i, \text{ para } r \in [s, s + t] \mid \mathbb{X}_r = i, \text{ para } r \in [0, s]] \stackrel{CMTC}{=} \\
&= P[\mathbb{X}_r = i, \text{ para } r \in [s, s + t] \mid \mathbb{X}_s = i] \stackrel{\text{Homogénea}}{=} \\
&= P[\mathbb{X}_r = i, \text{ para } r \in [0, t] \mid \mathbb{X}_0 = i] = \\
&= P[\xi_i > t].
\end{aligned}$$

Puesto que ξ_i es una variable continua verificando la propiedad de falta de memoria, tenemos que $\xi_i \rightsquigarrow \exp(\lambda_i)$, con $\lambda_i > 0$, puesto que las únicas variables continuas que verifican esta propiedad son aquellas que siguen una distribución exponencial. Conocida la distribución de ξ_i , deducimos:

$$\mathbb{E}[\xi_i] = \frac{1}{\lambda_i},$$

y, por tanto, cuánto más grande sea el parámetro λ_i , más pequeño será el intervalo de tiempo en el que el proceso se encuentra en el estado i .

Estamos preparados para construir la Q -matriz, $Q = (q_{ij})_{i,j \in S}$. Sea:

$$p_{ij}(\xi_i) = P(\mathbb{X}_{\xi_i} = j \mid \mathbb{X}_0 = i), \quad j \neq i$$

la probabilidad de transición del estado i al estado j , definimos:

$$q_{ij} := \lambda_i p_{ij}(\xi_i), \quad i \neq j$$

la *razón de transición* del estado i al j . Para obtener los elementos de la diagonal, es decir, para el caso en que $i = j$, imponemos que la suma de los elementos de cada una de las filas de Q sea 0:

$$q_{ii} + \sum_{j \neq i} q_{ij} = 0$$

$$q_{ii} = - \sum_{j \neq i} q_{ij}$$

$$q_{ii} = - \sum_{j \neq i} \lambda_i p_{ij}(\xi_i)$$

$$q_{ii} = -\lambda_i \underbrace{\sum_{j \neq i} p_{ij}(\xi_i)}_1$$

$$q_{ii} = -\lambda_i := -q_i$$

y hemos obtenido q_i , la *razón de escape* del estado i . Por tanto, la Q -matriz se puede expresar de la siguiente manera:

$$Q = \begin{cases} -\sum_{j \neq i} q_{ij} & \text{si } i = j \\ q_{ij} & \text{si } i \neq j \end{cases}$$

o equivalentemente:

$$Q = \begin{cases} -\lambda_i & \text{si } i = j \\ \lambda_i p_{ij}(\xi_i) & \text{si } i \neq j \end{cases}$$

4 Ecuación de Kolmogorov

Ya que una cadena de Markov está determinada por las probabilidades de transición, que ya hemos definido, vamos a calcularlas a partir de la razón de transición. Vamos a formalizarlo con el siguiente teorema.

Teorema 4.1 (Teorema de las Ecuaciones Diferenciales de Kolmogorov)

Sea \mathbb{X}_t con $t \in [0, \infty)$ es una cadena de Markov que no tiene ningún estado instantáneo. Entonces las probabilidades de transición son diferenciables en t y $\forall i, j$, par de estados, se tiene:

$$P'(t) = P(t)Q \text{ y } P(0) = I \iff \frac{dp_{ij}(t)}{dt} = -p_{ij}q_j + \sum_{k \neq j} p_{ik}(t)q_{kj} \quad p_{ij}(0) = \delta_{ij}$$

(Ecuación adelantada)

$$P'(t) = QP(t) \text{ y } P(0) = I \iff \frac{dp_{ij}(t)}{dt} = -q_i p_{ij} + \sum_{k \neq j} q_{ik} p_{kj}(t) \quad p_{ij}(0) = \delta_{ij}$$

(Ecuación atrasada)

Demostración 4.1 En esta demostración vamos a partir de la definición de derivada de $p_{ij}(t)$ hasta llegar al resultado final aplicando la Ecuación de Chapman-Kolmogorov y las razones de transición del apartado anterior.

Conocemos la ecuación de Chapman-Kolmogorov:

$$p_{ij}(t+h) = \sum_{k \in S} p_{ik}(h)p_{kj}(t) = \begin{cases} \sum_{k \neq j} p_{ik}(t)p_{kj}(h) + p_{jj}(h)p_{ij}(t) & (1) \\ \sum_{k \neq i} p_{ik}(h)p_{kj}(t) + p_{ii}(h)p_{ij}(t) & (2) \end{cases}$$

Lo que hemos hecho, básicamente, es expresar la ecuación de Chapman-Kolmogorov de dos formas distintas que nos servirán para llegar a las ecuaciones adelantadas y atrasadas.

Veamos ahora que las ecuaciones de Kolmogorov se cumplen:

$$p'_{ij}(t) = \lim_{h \rightarrow 0} \frac{p_{ij}(t+h) - p_{ij}(t)}{h} = \lim_{h \rightarrow 0} \frac{\sum_{k \in S} p_{ik}(h)p_{kj}(t) - p_{ij}(t)}{h}$$

Si ahora aplicamos (1):

$$\lim_{h \rightarrow 0} \sum_{k \neq j} \frac{p_{ik}(t)p_{kj}(h)}{h} - \lim_{h \rightarrow 0} \frac{p_{jj}(h) - 1}{h} p_{ij}(t) \stackrel{*}{=} -p_{ij}q_j + \sum_{k \neq j} p_{ik}(t)q_{kj} \quad (\text{Ecuación adelantada})$$

Si en cambio aplicamos (2):

$$\lim_{h \rightarrow 0} \sum_{k \neq i} \frac{p_{ik}(h)p_{kj}(t)}{h} - \lim_{h \rightarrow 0} \frac{p_{ii}(h) - 1}{h} p_{ij}(t) \stackrel{*}{=} -q_i p_{ij} + \sum_{k \neq j} q_{ik} p_{kj}(t) \quad (\text{Ecuación atrasada})$$

O en notación matricial:

$$\begin{cases} P'(t) = P(t)Q : P(0) = I & (\text{Ecuación adelantada}) \\ P'(t) = QP(t) : P(0) = I & (\text{Ecuación atrasada}) \end{cases}$$

Nota: "*" significa que hemos aplicados las razones de transición dadas en el apartado anterior.

Formalmente vemos que la solución a las ecuaciones de Kolmogorov viene dada por:

$$P(t) = e^{tQ} = \sum_{i=0}^{\infty} \frac{t^i Q^i}{i!}$$

Cuando Q es de dimensión finita, la serie anterior es convergente y es la única solución para los dos sistemas de ecuaciones. Si Q es de dimensión infinita, no podemos afirmar nada.

Supongamos que Q es de dimensión finita y diagonalizable con valores propios β_0, \dots, β_n . Entonces $\exists A$ matriz invertible que verifica $Q = ADA^{-1}$ donde D

es la matriz diagonal formada por sus valores propios ordenados. Y como consecuencia:

$$Q = A \begin{pmatrix} \beta_0 & & 0 \\ \dots & \dots & \dots \\ 0 & & \beta_n \end{pmatrix} A^{-1} \Rightarrow P(t) = A \begin{pmatrix} e^{\beta_0 t} & & 0 \\ \dots & \dots & \dots \\ 0 & & e^{\beta_n t} \end{pmatrix} A^{-1}$$

Ahora vamos a reformular la matriz P, definida en el apartado de la ecuación de Kolmogorov, en función de Q de forma que nos sea útil a la hora de demostrar un resultado futuro. Las expresiones son análogas, solo cambiamos la notación.

Teorema 4.2 Sea $p_{ij}(t)$ una matriz de transición y $\alpha = \sup_{j \in S} q_j < 0$:

$$p_{ij} = e^{-\beta t} \sum_{n=0}^{\infty} \frac{\beta^n t^n}{n!} p_{ij}^{(n)}, t \geq 0, i, j \in S$$

Donde $P=(p_{ij})$ es una matriz estocástica definida como:

$$P = I + \frac{1}{\beta} Q, \quad I = P^{(0)}, \quad \beta \geq \alpha$$

5 Probabilidades absolutas y distribución estacionaria

Sea \mathbb{X}_t con $t \in [0, +\infty)$ una cadena de Markov en tiempo continuo. Definiremos las probabilidades absolutas de una cadena de Markov (en el caso unidimensional) como la probabilidad de que una cadena se encuentre en el estado j en un tiempo t, es decir:

$$a_j = P[\mathbb{X}_t = j] \quad \text{con } j \in S, \quad t \in [0, +\infty)$$

La distribución absoluta de \mathbb{X}_t se define de la siguiente forma:

$$a(t) = (a_j; j \in S)$$

Una distribución inicial $a_j(0)$ se dice estacionaria si:

$$a_j(0) = a_j(t) \quad \forall t \geq 0 \quad j \in S$$

Una cadena de Markov con probabilidades de transición estacionarias se denomina estacionaria si las distribuciones absolutas $a(t)$ son independientes de t . En dicho caso, la distribución estacionaria $a=(a_j; j \in S)$, verifica que:

$$a_j = a_1(0)p_{1j}(t) + a_2(0)p_{2j}(t) + \cdots = \sum_{i \in S} a_i(0)p_{ij}(t), \quad j \in S$$

6 Clasificación de los estados

Sea \mathbb{X}_t con $t \in [0, +\infty)$, cadena de Markov en tiempo continuo con probabilidades de transición estacionarias y sea E_i un estado del espacio de estados.

A continuación clasificaremos los estados de una cadena de Markov según:

1. ξ_i : tiempo que tarda un proceso en abandonar el estado E_i o el tiempo de permanencia en dicho estado.
2. α_{ij} : tiempo de primera llegada al estado E_j partiendo de E_i
3. Según la posibilidad de ir de un estado a otro

6.1 Clasificación según ξ_i

Como hemos notado anteriormente, ξ_i es el tiempo de permanencia en un estado E_i , concretamente, suponemos que el proceso en el instante s está en el estado E_i , definiremos entonces:

$$\xi_i = \inf\{t : t > 0, x_{s+t} \neq i\}$$

También hemos visto que el tiempo de estancia en un estado E_i sigue una distribución exponencial de parámetro q_i . El tiempo medio esperado se calcula entonces como:

$$E[\xi_i] = \frac{1}{q_i}$$

Información que usaremos para la clasificación de los estados.

- Diremos que el estado E_i es estable si:

$$0 < q_i < +\infty, \quad (0 < E[\xi_i] < +\infty)$$

- El estado E_i será absorbente si:

$$q_i = 0, \quad (E[\xi_i] = +\infty)$$

- Por último, el estado E_i será instantáneo si:

$$q_i = +\infty, \quad (E[\xi_i] = 0)$$

6.2 Clasificación según α_{ij}

Denotaremos α_{ij} tiempo de primera llegada al estado E_j partiendo de E_i , más concretamente, si suponemos que el proceso en el instante s , se encuentra en el estado E_i , definiremos entonces:

$$\alpha_{ij} = \inf\{t : t > \xi_i, x_{s+t} = j\}$$

Si $i=j$, α_{ii} es el menor tiempo de retorno del proceso al estado E_i desde que lo abandona.

Definición 6.1 Denotaremos como $F_{ij}(t)$ a la función de distribución de α_{ij} :

$$F_{ij}(t) = \begin{cases} P[\alpha_{ij} < t] & \text{si } t > 0 \\ 0 & \text{si } t \leq 0 \end{cases}$$

Definición 6.2 Si el E_i es absorbente:

$$F_{ii}(t) = \begin{cases} 1 & \text{si } t > 0 \\ 0 & \text{si } t \leq 0 \end{cases}$$

Y $F_{ij}(t)$ será nula $\forall j \in S$ y $\forall t \in (-\infty, +\infty)$

Clasificaremos los estados E_i a través de la función de distribución $F_{ii}(t)$

- Un estado E_i es transitorio si:

$$\lim_{t \rightarrow \infty} F_{ii}(t) < 1 \quad (\Leftrightarrow \int_0^\infty p_{ii}(t) dt < \infty)$$

- Un estado E_i es recurrente si:

$$\lim_{t \rightarrow \infty} F_{ii}(t) = 1 \quad (\Leftrightarrow \int_0^\infty p_{ii}(t) dt = \infty)$$

- Sea E_i un estado recurrente y $\mu_i = E[\alpha_{ii}]$

– E_i es un estado recurrente nulo si:

$$\mu_i = \infty$$

– E_i es un estado recurrente positivo o ergódico si:

$$\mu_i < \infty$$

Teorema 6.1 Sea $p_{ij}(t)$ una matriz de transición y $\alpha = \sup_{j \in S} q_j < 0$:

1. $\exists \lim_{t \rightarrow \infty} p_{ii}(t), \forall i \in S$
2. E_i es transitorio $\iff \int_0^\infty p_{ii}(t) dt < \infty$
3. Si E_i es recurrente positivo, entonces $\lim_{t \rightarrow \infty} p_{ii}(t) = \frac{1}{q_i} \frac{1}{\mu_i} > 0, \mu_i = E[\alpha_{ii}]$
4. Si E_i es recurrente nulo entonces $\lim_{t \rightarrow \infty} p_{ii}(t) = 0$

Demostración 6.1 1. Sea M una cadena de Markov en tiempo discreto con espacio de estados S , con $P=(p_{ij})$ matriz de transición en un paso y p_{ij} definido como en el teorema anterior. Si $p_{ii}(t) > 0, i \in S$, todos los estados de S son no periódicos para M . Por tanto, $p_{ii}^{(n)}$, tiene límite cuando $n \rightarrow \infty$. Notaremos ν_i a dicho límite. Del teorema anterior obtenemos que para $i \in S, t \geq 0$ y un N natural arbitrario

$$p_{ii}(t) - \nu_i = e^{-\beta t} \sum_{n=0}^N \frac{\beta^n t^n}{n!} (p_{ii}^{(n)} - \nu_i) + e^{-\beta t} \sum_{n=N+1}^{\infty} \frac{\beta^n t^n}{n!} (p_{ii}^{(n)} - \nu_i).$$

Dado $\epsilon > 0$, elegimos N de forma que $|p_{ii}^{(n)} - \nu_i| < \epsilon$ para $n > N$. Como el primer término de la parte derecha tiende a cero cuando $t \rightarrow \infty$ y el segundo término, en valor absoluto, es siempre menor que $\epsilon \implies \exists \lim_{t \rightarrow \infty} p_{ii}(t)$ y además

$$\lim_{t \rightarrow \infty} p_{ii}(t) = \lim_{n \rightarrow \infty} p_{ii}^{(n)}$$

2. Para $s \in \mathbb{R}$ definimos:

$$\Pi_{ij}(s) = \int_0^\infty e^{-st} p_{ij}(t) dt, \quad s > 0$$

$$\varphi_{ij}(s) = \int_{0-}^{\infty} e^{-st} dF_{ij}(t), \quad s \geq 0$$

Llegamos a que:

$$\Pi_{ii}(s) = \frac{1}{s + q_i} \frac{1}{1 - \varphi_{ii}(s)}, \quad s > 0$$

Si E_i es transitorio se llegamos a que:

$$\lim_{s \downarrow 0} \varphi_{ii}(s) = \lim_{t \rightarrow \infty} F_{ii}(t) < 1$$

Por lo tanto, $\Pi_{ii}(s)$ tiene un límite finito para $s \downarrow 0$, de modo que $\int_0^{\infty} p_{ii}(t) dt < \infty$. Por otra parte, si $\int_0^{\infty} p_{ii}(t) dt$ converge, entonces, $\Pi_{ii}(s)$ tiene un límite finito para $s \downarrow 0$ de forma que $F_{ii}(+\infty) < 1$, es decir, E_i es transitorio.

3. Tanto el tercer punto como el cuarto se van a demostrar a partir de las siguientes relaciones.

Supongamos que E_i es un estado recurrente, entonces $\varphi_{ii}(s) \rightarrow 1$ para $s \downarrow 0$, la existencia del límite de $p_{ii}(t)$ para $t \rightarrow \infty$ implica que:

$$\lim_{t \rightarrow \infty} p_{ii}(t) = \lim_{s \downarrow 0} s \Pi_{ii}(s) = \frac{1}{q_i} \lim_{s \downarrow 0} \frac{s}{1 - \varphi_{ii}(s)} = -\frac{1}{q_i} \lim_{s \downarrow 0} \left[\frac{d}{ds} \varphi_{ii}(s) \right]^{-1}$$

Como:

$$-\frac{d}{ds} \varphi_{ii}(s) = \int_0^{\infty} e^{-st} t dF_{ii}(t), \quad s > 0$$

Llegamos a que:

$$F_{ij}(t) = \begin{cases} -\mu_i & \text{si } \int_0^{\infty} t dF_{ii}(t) < \infty \\ -\infty & \text{si } \int_0^{\infty} t dF_{ii}(t) = \infty \end{cases}$$

De donde se deducen los dos puntos que quedaban por demostrar ■

6.3 Clasificación según la posibilidad de ir de un estado a otro

Por último, vamos a dar una clasificación de estados según sea posible ir de un estado a otro

Definición 6.3 Un estado E_i alcanza al estado E_j si siempre es posible llegar del estado i al j en un determinado tiempo:

$$\exists t > 0 \text{ tal que } p_{ij}(t) > 0$$

Definición 6.4 Se dice que dos estados E_i y E_j son comunicantes si cada uno de ellos alcanza al otro:

$$\exists t_1 > 0 \text{ tal que } p_{ij}(t_1) > 0 \text{ y } \exists t_2 > 0 \text{ tal que } p_{ji}(t_2) > 0$$

Definición 6.5 Una clase de estados S es un conjunto formado por todos los estados que se comunican entre sí. Diremos que una clase de estados es cerrada si los estados de dicha clase solo se comunican entre ellos:

$$p_{ij} = 0, \quad \forall t > 0, \quad \forall E_i \in S, \quad \forall E_j \notin S$$

Definición 6.6 Una clase de estados será minimal si es cerrada y no contiene subclases propias cerradas

Definición 6.7 Si la clase de estados de una cadena de Markov es minimal, entonces la cadena de Markov es llamada irreducible. En una cadena de Markov irreducible existe una única clase de estados y todos sus estados se comunican entre sí.

Proposición 6.1 Una cadena irreducible no puede tener estados absorbentes.

Antes de enunciar el teorema ergódico vamos a dar unas cuantas propiedades más sobre la clasificación de estados en cadenas de Markov. Vamos a definir para ello tres cadenas de Markov: M_1 cadena de Markov en tiempo continuo y M_2 y M_3 cadenas de Markov en tiempo discreto con probabilidades de transición estacionarias y espacio de estados S .

- Si $q_j = 0$ para algún $j \in S$ entonces E_j es un estado absorbente en las tres cadenas M_1 , M_2 y M_3 y la clase de estados de E_j como único elemento es minimal para todas ellas
- Si un estado E_j es no recurrente en una de las cadenas, no lo será tampoco en las otras. Análogamente si el estado E_j es recurrente en una cadena lo será en las otras dos.
- Una clase de estados cerrada (minimal) en una de las cadenas será también una clase de estados cerrada (minimal) en las otras cadenas y un estado E_i será del mismo tipo en las tres cadenas

7 Teoremas límite

A continuación presentamos el Teorema Ergódico.

Teorema 7.1 (Teorema ergódico) *Para una cadena de Markov irreducible se verifica:*

1. *Si todos los estados son recurrentes positivos:*

$$\lim_{t \rightarrow \infty} p_{i,j}(t) = u_j > 0, \forall i, j \in S$$

Además $u = (u_j, j \in S)$ es la única distribución estacionaria de la cadena y $\lim_{t \rightarrow \infty} a_j(t) = u_j, \forall j \in S$. La distribución de probabilidad u es la solución del sistema $uQ = 0$, siendo Q la Q -matriz de la cadena.

2. *Si todos los estados son recurrentes nulos o transitorios,*

$$\lim_{t \rightarrow \infty} p_{ij}(t) = 0, \forall i, j \in S; \quad \lim_{t \rightarrow \infty} a_j(t) = 0, \forall j \in S$$

8 Aplicaciones

Sin entrar en detalles, presentamos una lista con las aplicaciones más usuales de las cadenas de Markov en tiempo continuo:

- Modelos epidemiológicos
- Internet
- Simulación
- Juegos de azar
- Economía y finanzas
- Genética
- Música
- Redes neuronales

9 Ejemplo

Una chica ha decidido salir con dos chicos a la vez, a los que llamaremos A y B por respeto a la intimidad: Para ser justa en sus visitas, ha decidido gastar un tiempo aleatorio distribuido de forma exponencial de parámetro λ , de modo que si en un instante de tiempo está con A, en un tiempo exponencialmente distribuido irá a visitar a B. Los viajes los consideraremos instantáneos.

Por otra parte, sabe que mientras está con B el novio A puede llamarla y tiene que ir inmediatamente a visitarlo, lo que ocurre según un tiempo exponencialmente distribuido de tasa μ_A . Cuando esto ocurre esto, interrumpe la visita al novio B y va con el novio A.

La chica sabe que B no soportará más de 3 interrupciones, consecutivas o no. Por esto, cuando contabiliza 3, la chica apaga el móvil y pasa más tiempo con B, un tiempo exponencial de parámetro λ_B ($\lambda_B < \lambda$) lo que hace olvidar a B las tres interrupciones.

Podemos ver el grafo asociado a la cadena de Markov justo debajo:

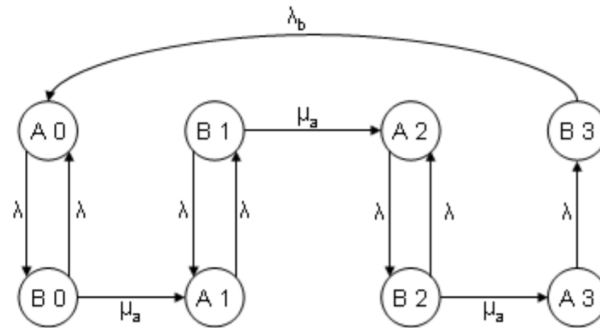


Figure 2: CMTD asociada al ejemplo