

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	5
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	Clasificación de los estados	12
6	Teoremas límite	12
7	Ejemplos	12
7.1	Ejemplo 1	12
7.2	Ejemplo 2	12

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: Cadena de Markov

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 $0 \leq u \leq s$, se cumple que:

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

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: Cadena de Markov

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: 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:

$$P(\mathbb{X}_n = j \mid \mathbb{X}_{n-1} = i) = P(\mathbb{X}_1 = j \mid \mathbb{X}_0 = i)$$

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 $h + 1$:

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

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

Obtenemos que:

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

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

2.2.2 Segunda propiedad

Demostración 2.2

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:

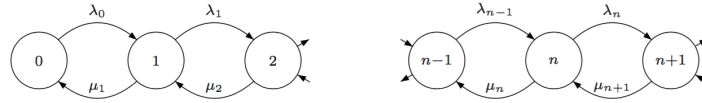


Figure 1: Ejemplo de grafo de una CMTC-h

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.

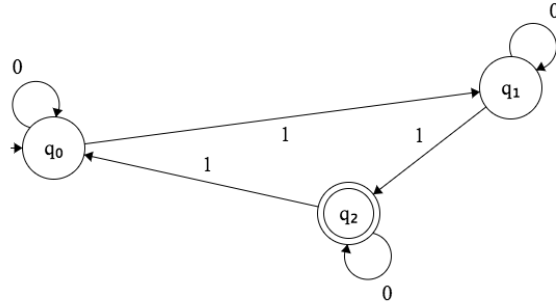
Para concluir presentaremos la definición de autómata finito determinista, por si el lector está interesado en el estudio del mismo:

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 $\delta(q, a) = q_2 \Rightarrow q_1 = q_2$ 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 $\delta t.q$:

$\delta(q_0, 0) = q_0$	$\delta(q_1, 0) = q_1$	$\delta(q_2, 0) = q_2$
$\delta(q_0, 1) = q_1$	$\delta(q_1, 1) = q_2$	$\delta(q_2, 1) = q_0$
- $F = q_2$

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: $Q = E$, $\Sigma = W$ 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:

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

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

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

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

■

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:

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

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

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

$$= P[T_i > t].$$

Puesto que T_i es una variable continua verificando la propiedad de falta de memoria, tenemos que $T_i \rightsquigarrow \exp(\lambda_i)$, 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 T_i , deducimos:

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

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:

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

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

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

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:

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

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

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

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

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

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

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

o equivalentemente:

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

4 Ecuación de Kolmogorov

5 Clasificación de los estados

6 Teoremas límite

A continuación presentamos el Teorema Ergódico.

Teorema 6.1 *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,*

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

7 Ejemplos

7.1 Ejemplo 1

7.2 Ejemplo 2