

Procesos estocásticos

Sesión 3. Introducción a las cadenas de Markov

Enrique Miranda

Universidad of Oviedo

Máster Universitario en Análisis de Datos
para la Inteligencia de Negocios

Contenidos

1. Definición.
2. Ejemplos.
3. Matrices de transición.
4. Ecuaciones de Chapman-Kolmogorov.

Cadenas de Markov: introducción

Las **cadenas de Markov** son un tipo de procesos estocásticos a **tiempo discreto** que permiten modelizar una gran cantidad de fenómenos biológicos, sociales, médicos, físicos económicos, etc.

Son procesos relativamente sencillos ya que asumen que **el futuro sólo depende del pasado a través del presente**.

Esto ha facilitado el que se haya desarrollado para ellas una teoría bastante completa que permite describirlas con precisión y emplearlas para realizar predicciones.

Ejemplo: el problema de la ruina del jugador

Consideremos un juego de apuestas en el cual, en cada turno, podemos ganar 1 euro con probabilidad $p = 0.4$ o bien perderlo con probabilidad $1 - p = 0.6$.

Supongamos que decidimos jugar hasta que, o bien logremos reunir N euros o bien nos arruinemos, es decir, nuestra fortuna alcance los 0 euros.

Llamemos X_n a la cantidad a la que asciende nuestra fortuna tras haber jugado n turnos.

Problema de la ruina: continuación

El proceso $\{X_n\}_{n \geq 1}$ tiene la **propiedad de Markov**: si en la etapa n estamos aún jugando, es decir, si $X_n = i$ para algún $0 < i < N$, entonces, para cualquier posible historia $i_{n-1}, i_{n-2}, \dots, i_1, i_0$, se tiene que

$$P(X_{n+1} = i + 1 | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = 0.4,$$

y

$$P(X_{n+1} = i - 1 | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = 0.6.$$

Cadenas de Markov: definición

Se dice que $\{X_n\}_{n \geq 1}$ es una **cadena de Markov en tiempo discreto** si para todo n la distribución de $X_{n+1} | X_n, \dots, X_0$ es independiente de $X_{n-1}, \dots, X_0$, es decir, si

$$\begin{aligned} &P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) \\ &= P(X_{n+1} = j | X_n = i, X_{n-1} = i'_{n-1}, \dots, X_1 = i'_1, X_0 = i'_0) \end{aligned}$$

para todos $i, j, i_0, \dots, i_{n-1}, i'_0, \dots, i'_{n-1}$ en el conjunto de estados.



A. A. Markov (1886).

- ▶ Los primeros resultados fueron realizados por **Andrei Markov** en 1906 para espacios finitos de estados.
- ▶ La extensión a espacios numerables de estados fue realizada por **Kolmogorov**.
- ▶ Otros trabajos relevantes se deben a **Erlang** en teoría de colas.

Cadenas homogéneas y no homogéneas

Una cadena de Markov se dice **homogénea** cuando

$$P(X_{n+1} = j | X_n = i) = P(X_1 = j | X_0 = i)$$

para todo n y para todo par de estados i, j . En otro caso la cadena de Markov se dice **no homogénea**.

En este curso trabajaremos fundamentalmente con cadenas de Markov homogéneas, las cuales vienen caracterizadas por su **matriz de transición**

$$p(i, j) = P(X_1 = j | X_0 = i).$$

Elementos de una cadena de Markov

Una cadena de Markov queda completamente determinada a partir de:

- ▶ Su *espacio de estados* Ω .
- ▶ Su *matriz de transición* P , de dimensión $n \times n$, siendo $n = |\Omega|$.
- ▶ Si el estado de partida X_0 es desconocido o aleatorio, necesitamos también la *distribución de probabilidades inicial* sobre Ω .

Ejemplo 1: la ruina del jugador

Para el ejemplo del problema de la ruina, las probabilidades de transición son

$$p(i, i+1) = 0.4 \quad p(i, i-1) = 0.6 \quad \text{si } 0 < i < N$$

$$p(0, 0) = 1 \quad p(N, N) = 1$$

Por ejemplo, en el caso $N = 4$, la **matriz de transición** sería

	0	1	2	3	4
0	1	0	0	0	0
1	0.6	0	0.4	0	0
2	0	0.6	0	0.4	0
3	0	0	0.6	0	0.4
4	0	0	0	0	1

Ejemplo 2: la cadena del tiempo

Aunque las cadenas de Markov no son los procesos más adecuados para modelizar el tiempo meteorológico, supongamos, a modo de ejemplo sencillo, que en una determinada zona el tiempo de cada día depende sólo de cómo haya sido el del día anterior.

Supongamos que los posibles estados del tiempo son *lluvioso* (L), *nublado* (N) y *soleado* (S). El espacio de estados para este ejemplo es por tanto

$$\Omega = \{L, N, S\}.$$

Como matriz de probabilidades de transición podríamos tomar, por ejemplo,

	L	N	S
L	0.4	0.6	0
N	0.2	0.5	0.3
S	0.1	0.7	0.2

Matrices estocásticas

Se llama **matriz estocástica** a cualquier matriz cuadrada para la cual se verifique $p(i, j) \geq 0 \quad \forall i, j$, y

$$\sum_{j \in \Omega} p(i, j) = 1 \quad \forall i \in \Omega.$$

Toda matriz estocástica define una cadena de Markov.

Si en una matriz estocástica ocurre además que todas las **columnas** suman 1, se dice que la matriz es **doblemente estocástica**. Las cadenas de Markov asociadas a este tipo de matrices tienen características especiales.

Ejercicio: cadena de reparación

Una máquina cuenta con tres motores (1,2 y 3) que pueden fallar con probabilidades 0.01, 0.02 y 0.04 respectivamente. La máquina está diseñada de tal manera que no puede fallar más de un motor en el mismo día, y funciona correctamente siempre que al menos dos de los motores estén funcionando.

Cuando dos de los motores están estropeados, son reemplazados por otros dos nuevos (cada uno de su tipo) y la máquina vuelve a funcionar con todos sus motores en buenas condiciones.

- ↪ Determina la matriz de transición del proceso estocástico con espacio de estados 'motores estropeados'.

Ejercicio: políticas de control de inventario

Un establecimiento ha comprobado que la demanda diaria de cierto producto es de 0, 1, 2, 3 unidades con probabilidades 0.3, 0.4, 0.2 y 0.1 respectivamente. Su política de control de inventario es la siguiente:

- ▶ si al final de una jornada en el almacén quedan 1 o 0 unidades del producto, hace un pedido hasta llegar a un total de 5 unidades, que llegan al establecimiento antes de la hora de apertura del día posterior;
- ▶ si quedan dos o más unidades, no realiza ningún pedido.

Denotemos por X_n el número de unidades almacenadas por el establecimiento tras cierre del n -ésimo día.

↪ Determina la matriz de transición.

Probabilidades de transición en varias etapas

Las probabilidades de transición $p(i, j) = P(X_{n+1} = j | X_n = i)$ se refieren a pasar del estado i al j en una etapa (o en un salto).

A continuación vamos a analizar las probabilidades de pasar i a j en m etapas,

$$p^m(i, j) = P(X_{n+m} = j | X_n = i),$$

para $m > 1$.

Ejemplo

Consideremos la cadena de tiempo meteorológico,

1. Si hoy está nublado, cuál es la probabilidad de que mañana esté soleado y pasado mañana lluvioso?

$$\begin{aligned}P(X_2 = L, X_1 = S | X_0 = N) \\ = p(N, S)p(S, L) = 0.3 \times 0.1 = 0.03.\end{aligned}$$

2. Si hoy está nublado, cuál es la probabilidad de que pasado mañana esté lluvioso?

$$\begin{aligned}P(X_2 = L | X_0 = N) &= \sum_k p(N, k)p(k, L) \\ &= 0.2 \times 0.4 + 0.5 \times 0.2 + 0.3 \times 0.1 = 0.21.\end{aligned}$$

Matriz de transición en dos etapas: generalización

De forma similar podría comprobarse que, en general, para cualquier cadena de Markov X_n con espacio de estados Ω se verifica

$$P(X_2 = j | X_0 = i) = \sum_{k \in \Omega} p(i, k)p(k, j).$$

Esta ecuación indica que la entrada (i, j) de la matriz que indica las probabilidades de transición en dos etapas es la entrada (i, j) de la matriz de transiciones (p) multiplicada por sí misma, es decir, p^2 .

Ecuaciones de Chapman-Kolmogorov

Las probabilidades de transición en varias etapas vienen determinadas por las **ecuaciones de Chapman-Kolmogorov**:

$$p^{m+n}(i, j) = \sum_{k \in \Omega} p^m(i, k) p^n(k, j) \quad \forall i, j \in \Omega.$$

La idea intuitiva es que para pasar desde i a j en $m + n$ etapas, necesitamos ir desde i hasta algún estado k en m etapas y después llegar desde ese estado k hasta el j en n etapas.

La propiedad de Markov implica que *las dos partes del trayecto son independientes*.

Teorema (probabilidades de transición en m etapas)

La matriz de probabilidades de transición en m etapas,

$$[p^m(i, j)]_{i, j \in \Omega} = [P(X_{n+m} = j | X_n = i)]_{i, j \in \Omega},$$

es la *potencia m -ésima de la matriz de transición*

$$p = [p(i, j)]_{i, j \in \Omega} = [P(X_{n+1} = j | X_n = i)]_{i, j \in \Omega}.$$

Ejemplo 1: la ruina del jugador

Para el caso $N = 4$ tenemos por ejemplo que

$$p^2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.6 & 0.24 & 0 & 0.16 & 0 \\ 0.36 & 0 & 0.48 & 0 & 0.16 \\ 0 & 0.36 & 0 & 0.24 & 0.4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

y

$$p^{20} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.87655 & 0.00032 & 0 & 0.00022 & 0.12291 \\ 0.69186 & 0 & 0.00065 & 0 & 0.30749 \\ 0.41842 & 0.00049 & 0 & 0.00032 & 0.58437 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Vemos que no es posible pasar del estado 2 al 3 en una cantidad par de etapas, puesto que la cadena alterna entre estados pares e impares hasta que queda atrapada en 0 o en 4. Este fenómeno se conoce como **periodicidad**.

Ejemplo 2: clima

Ya hemos visto que

$$p^2 = \begin{pmatrix} 0.28 & 0.54 & 0.18 \\ 0.21 & 0.58 & 0.21 \\ 0.20 & 0.55 & 0.25 \end{pmatrix}$$

Aplicando el teorema anterior obtenemos

$$p^4 = \begin{pmatrix} 0.2278 & 0.5634 & 0.2088 \\ 0.2226 & 0.5653 & 0.2121 \\ 0.2215 & 0.5645 & 0.2140 \end{pmatrix}$$

y

$$p^{20} = \begin{pmatrix} 0.22355 & 0.56470 & 0.21175 \\ 0.22352 & 0.56471 & 0.21177 \\ 0.22352 & 0.56471 & 0.21177 \end{pmatrix}$$

Este ejemplo parece indicar que p^n *está convergiendo rápidamente a una matriz con todas sus filas iguales.*

¿Y si el estado inicial es aleatorio?

Para cualquier cadena de Markov X_n ,

$$\begin{aligned} P(X_n = y) &= \sum_{x \in \Omega} P(X_0 = x, X_n = y) \\ &= \sum_{x \in \Omega} P(X_0 = x) P(X_n = y | X_0 = x). \end{aligned}$$

Denotando $\mu(x) = P(X_0 = x)$ podemos re-escribir la ecuación anterior como

$$P(X_n = y) = \sum_{x \in \Omega} \mu(x) p^n(x, y).$$

Es decir, la probabilidad de estar en el estado Y en la etapa n -ésima puede obtenerse como el producto del vector de probabilidades iniciales por la matriz de probabilidades de transición en n etapas.

Ejemplo

En el ejemplo del clima, supongamos que comenzamos nuestra observación en un día en el que las probabilidades de lluvioso, nublado y soleado son 0.2, 0.3 y 0.5, respectivamente.

La probabilidad de tiempo soleado dos días después sería:

$$\begin{aligned}P(X_2 = S) &= 0.2p^2(L, S) + 0.3p^2(N, S) + 0.5p^2(S, S) \\&= 0.2 \cdot 0.18 + 0.3 \cdot 0.21 + 0.5 \cdot 0.25 = 0.224.\end{aligned}$$