Cadenas de Markov

Temas Selectos I

Dr. Rubén Blancas Rivera

Universidad de las Américas Puebla

Otoño 2025

Contenido

Introducción

- Ha llegado la última parte del curso. En este módulo 4 vamos a estudiar los métodos Monte Carlo vía cadenas de Markov.
- Estos son referidos como métodos MCMC (Markov Chains Monte Carlo).
- Son procedimientos para obtener muestras aleatorias aproximadas de distribuciones arbitrarias y hacen uso de la teoría de las cadenas de Markov.

Introducción

- El algoritmo original fue propuesto por Nicholas Metropolis y otros autores en (1939).
- Para referirise a los procedimientos asociados a este algoritmo se puede usar también el nombre **muestreo por cadenas de Markov**
- Los dos algoritmos MCMC más utilizados son:
  - Algoritmo Metropolis-Hastings.
  - Muestreo de Gibbs.

Repaso de Cadenas de Marov

Una cadena de Markov es un proceso estocástico a tiempo discreto

$$\{X_t : t = 0, 1, \dots\}$$

con espacio de estados también discreto y que satisface la propiedad de Markov, esto es, para cualquier tiempo $t \geq 1$ y para cualesquiera estados $x_0, \dots, x_t, x_{t+1}$,

$$P(X_{t+1} = x_{t+1} \mid X_0 = x_0, \dots, X_t = x_t) = P(X_{t+1} = x_{t+1} \mid X_t = x_t).$$

- La identidad anterior es conocida como propiedad de Markov.

Repaso de Cadenas de Markov

- Esta condición establece que la probabilidad del evento futuro $(X_{t+1} = x_{t+1})$ depende del pasado o historia $(X_0 = x_0, \ldots, X_t = x_t)$ sólo a través del presente $(X_t = x_t)$ o del pasado más próximo.
- Sin pérdida de generalidad, tomaremos como espacio de estados el conjunto discreto $S = \{0, 1, 2, \ldots\}$.
- En particular, una cadena de Markov es finita cuando el espacio de estados es finito; por ejemplo, el conjunto $S = \{0, 1, \ldots, k\}$ es un espacio de estados finito.

Repaso de Cadenas de Markov

- Sea $t \geq 0$ y sean i y j dos estados cualesquiera de una cadena de Markov.
- La probabilidad condicional

$$P(X_{t+1} = j \mid X_t = i)$$

  es la probabilidad de transición del estado i al tiempo t, al estado j al tiempo siguiente $t+1$.
- Se le denota también por $p_{ij}(t, t+1)$ y se le llama probabilidad de transición en un paso.

Repaso de Cadenas de Markov

- Cuando las probabilidades $p_{ij}(t, t+1)$ no dependen de t, se dice que la cadena es estacionaria o homogénea en el tiempo.
- Regularmente se estudian las cadenas de Markov bajo esta hipótesis, y las probabilidades de transición en un paso se escriben como p_{ij}.
- Haciendo variar los subíndices i y j en el espacio de estados $S = \{0, 1, 2, \ldots\}$ se obtiene la matriz de probabilidades de transición en un paso.

Matriz de Transiciones

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} & \cdots \\ p_{10} & p_{11} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix},$$

en donde los renglones y columnas se etiquetan con los valores de los estados $0, 1, \ldots$ de tal forma que p_{ij} es la entrada (i, j) de esta matriz, la cual se escribe $\mathbf{P} = (p_{ij})$ y cumple las siguientes dos propiedades:

- $p_{ij} \geq 0.$
- $\displaystyle\sum_{j=0}^{\infty} p_{ij} = 1.$

Distribución Inicial

- Una distribución inicial es cualquier distribución para la variable aleatoria inicial X_0 de una cadena de Markov. Esta es una distribución sobre el espacio de estados $S = \{0, 1, 2, \ldots\}$.
- En ocasiones se establece que la cadena inicia en el estado $x_0 \in S$, es decir, $X_0 = x_0$. Esto significa que la distribución inicial está completamente concentrada en el punto x_0.

Definición

- Una cadena de Markov $\mathbf{P} = (p_{ij})$ determina una dinámica estocástica en el espacio de estados $S = \{0, 1, 2, \ldots\}$ de la siguiente manera:
- Si i es un estado inicial, a partir de él se genera el siguiente estado de acuerdo a la distribución dada por el vector de probabilidad $(p_{i0}, p_{i1}, \ldots)$. Este es el i-ésimo renglón de la matriz $\mathbf{P}$.

Definición

- Supongamos que el valor obtenido es j. A partir de la distribución dada por el j-ésimo renglón de $\mathbf{P}$, esto es $(p_{j0}, p_{j1}, \ldots)$, se genera el siguiente valor.
- Supongamos que el valor es k. Se procede de manera sucesiva y se obtiene así una secuencia de puntos en el espacio de estados S.
- A una sucesión de estos puntos en el espacio de estados se le llama una trayectoria o realización de la cadena de Markov.

Algoritmo

Entrada.

1. Conjunto de estados finito $S = \{0, 1, \ldots, k\}$.
2. Matriz de transición $\mathbf{P} = [p_{ij}]_{i,j \in S}$ con $p_{ij} \geq 0$ y $\sum_j p_{ij} = 1$.
3. Horizonte n (número de pasos a simular).
4. Estado inicial fijo $X_0 = x_0 \in S$ *o* distribución inicial $\boldsymbol{\pi} = (\pi_0, \ldots, \pi_k)$.

Simulación

Algoritmo

Inicialización del estado.

1. Si el estado inicial es aleatorio: genera $U_0 \sim U(0,1)$ y toma
$$X_0 = \min\left\{j \in S : \sum_{m=0}^{j} \pi_m \geq U_0\right\}.$$
2. Si el estado inicial es fijo, simplemente define $X_0 = x_0$.

Simulación

Algoritmo

Simulación paso a paso. Para $t = 0, 1, \ldots, n-1$:

1. Genera un uniforme $U_{t+1} \sim U(0,1)$ independiente.
2. Sea $i = X_t$. Calcula la suma acumulada de la fila i:

$$c_j = \sum_{m=0}^{j} p_{im}, \qquad j = 0, 1, \ldots, k.$$

3. Define el siguiente estado como

$$X_{t+1} = \min\left\{ j \in S : \; c_j \geq U_{t+1} \right\}.$$

Salida. La trayectoria simulada $(X_0, X_1, \ldots, X_n)$.

Probabilidades de Transición en n pasos

Sean $n \geq 1$ y $m \geq 0$ números enteros, a las cantidades denotadas por

$$P(X_{n+m} = j \mid X_m = i),$$

se les llama probabilidades de transición del estado i al estado j en n pasos.

Probabilidades de Transición en n pasos

Dado que hemos supuesto la condición de homogeneidad en el tiempo, esta probabilidad no depende del tiempo m y coincide con $P(X_n = j \mid X_0 = i)$.
Haciendo variar los valores de i y de j en el espacio de estados se obtiene la matriz de probabilidades de transición en n pasos,

$$\mathbf{P}(n) = \begin{pmatrix} p_{00}(n) & p_{01}(n) & \cdots \\ p_{10}(n) & p_{11}(n) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

Probabilidades de Transición en n pasos

Cuando $n = 1$ se obtiene la matriz $\mathbf{P}$, y para $n = 0$ se define $\mathbf{P}(0)$ como la matriz identidad, es decir,

$$\mathbf{P}(0) = (p_{ij}(0)) = \delta_{ij} = \begin{cases} 0, & \text{si } i \neq j, \\ 1, & \text{si } i = j. \end{cases}$$

Ecuación de Chapman-Kolmogorov

Sean $0 \leq r \leq n$ dos números enteros. Las probabilidades de transición $p_{ij}(n)$ de una cadena de Markov satisfacen

$$p_{ij}(n) = \sum_{k=0}^{\infty} p_{ik}(r)\, p_{kj}(n-r).$$

Ecuación de Chapman-Kolmogorov

Aplicando este resultado en $n-1$ ocasiones usando $r=1$, se encuentra que $p_{ij}(n)$ resulta ser la entrada (i,j) del producto $\mathbf{P}\cdot\mathbf{P}\cdots\mathbf{P}=\mathbf{P}^n$, es decir,

$$p_{ij}(n)=(\mathbf{P}\cdots\mathbf{P})_{ij}=(\mathbf{P}^n)_{ij}.$$

Ecuación de Chapman-Kolmogorov

De forma explícita,

$$\begin{pmatrix} p_{00}(n) & p_{01}(n) & \cdots \\ p_{10}(n) & p_{11}(n) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} p_{00} & p_{01} & \cdots \\ p_{10} & p_{11} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}^n .$$

Distribución de X_n

Cuando se conoce la matriz potencia $\mathbf{P}^n$ y se tiene una distribución inicial denotada por $p_i := P(X_0 = i)$, entonces la distribución de la variable X_n se puede calcular mediante la siguiente expresión:

$$P(X_n = j) = \sum_{i=0}^{\infty} p_i \, p_{ij}(n), \qquad j = 0, 1, \ldots$$