Una pregunta pendiente

¿Cómo podemos construir un generador (casi) uniforme para una relación?

Recuerde el problema KS definido en la sección anterior y la relación:

$$R_{\text{KS}} \quad = \quad \{((\vec{a}, b), \vec{x}) \mid \vec{a} \in \mathbb{N}^n \text{ y } \vec{x} \in \{0,1\}^n \text{ para } n \geq 1, b \in \mathbb{Z} \text{ y } \vec{a} \cdot \vec{x} \leq b\}$$

Vamos a responder primero una pregunta más específica: ¿Cómo podemos construir un generador (casi) uniforme para R_{KS}?

- La respuesta a esta pregunta va a tener los ingredientes necesarios para responder la pregunta más general

Una secuencia de variables aleatorias para generar R_{KS}

Fije $n \geq 1$, $\vec{a} \in \mathbb{N}^n$ y $b \in \mathbb{N}$

- Y suponga que $\Omega = \{\vec{x} \in \{0,1\}^n \mid \vec{a} \cdot \vec{x} \leq b\}$

Nótese que $\Omega \neq \emptyset$

Considere una secuencia $\{X_t\}_{t \in \mathbb{N}}$ de variables aleatorias con recorrido Ω

- El dominio de cada variable X_t es D_t, el cual no necesitamos definir

Decimos que Ω es el conjunto de estados de la secuencia $\{X_t\}_{t \in \mathbb{N}}$

Una secuencia de variables aleatorias para generar R_{KS}

Dado $t \in \mathbb{N}$ y $\vec{x} \in \Omega$, nos interesa calcular $\mathbf{Pr}(X_t = \vec{x})$

Para calcular esta probabilidad necesitamos definir la dinámica de la secuencia

- Vale decir, necesitamos definir cómo se cambio de estado al pasar de tiempo t a tiempo $t + 1$

De manera formal, dado $t \in \mathbb{N}$ y $\vec{x}, \vec{y} \in \Omega$, necesitamos entregar:

$$\mathbf{Pr}(X_{t+1} = \vec{y} \mid X_t = \vec{x})$$

Una secuencia de variables aleatorias para generar R_{KS}

Suponga que en tiempo t estamos en el estado $\vec{x} = (x_1, \ldots, x_n)$

El estado $\vec{y}$ en el tiempo $t + 1$ se obtiene utilizando el siguiente procedimiento:

GenerarSiguiente($\vec{x}$)
    Escoja $c \in \{0, 1\}$ con distribución uniforme
    if $c = 0$ **then return** $\vec{x}$
    else
        Escoja $i \in \{1, \ldots, n\}$ con distribución uniforme
        $\vec{u} := (x_1, \ldots, x_{i-1}, 1 - x_i, x_{i+1}, \ldots, x_n)$
        if $\vec{a} \cdot \vec{u} \leq b$ **then return** $\vec{u}$
        else return $\vec{x}$

Tenemos entonces que $\vec{y} =$ **GenerarSiguiente**($\vec{x}$)

Algunas propiedades de la secuencia

Ejercicios

Sea $t \in \mathbb{N}$

1. Dados $\vec{x}, \vec{y} \in \Omega$, calcule $\mathbf{Pr}(X_{t+1} = \vec{y} \mid X_t = \vec{x})$
   - Considere de manera separada los casos $\vec{x} \neq \vec{y}$ y $\vec{x} = \vec{y}$

2. Demuestre que para todo $\vec{x}, \vec{y} \in \Omega$, existe $t' > t$ tal que:

$$\mathbf{Pr}(X_{t'} = \vec{y} \mid X_t = \vec{x}) \quad > \quad 0$$

3. Dados $\vec{x}_0, \vec{x}_1, \ldots, \vec{x}_t, \vec{x}_{t+1} \in \Omega$, demuestre que:

$$\mathbf{Pr}(X_{t+1} = \vec{x}_{t+1} \mid X_0 = \vec{x}_0 \wedge X_1 = \vec{x}_1 \wedge \cdots \wedge X_t = \vec{x}_t) \; = \; \mathbf{Pr}(X_{t+1} = \vec{x}_{t+1} \mid X_t = \vec{x}_t)$$

¿A qué converge la secuencia?

Sea $\vec{x}$ un vector arbitrario en Ω

Suponga que en tiempo 0 estamos en el estado $\vec{x}$, y que en cada instante cambiamos de estado utilizando el procedimiento **GenerarSiguiente**

- ¿Cuáles son los estados a los que podríamos llegar en un tiempo $t \gg 0$? ¿Cuál es la probabilidad de estar en un estado específico en este tiempo t?
- ¿Es posible llegar a una distribución estacionaria, vale decir, un tiempo t' tal que para todo $\vec{y} \in \Omega$ se tiene que $\mathbf{Pr}(X_{t'+1} = \vec{y}) = \mathbf{Pr}(X_{t'} = \vec{y})$?
- ¿Existe una única distribución estacionaria?

¿Cuáles son las respuestas a las preguntas anteriores si cambiamos $\vec{x}$ por otro vector inicial?

La secuencia converge a la distribución uniforme

Ejercicios

1. Dado $t \in \mathbb{N}$ y $\vec{x} \in \Omega$, demuestre que:

$$\begin{aligned}\sum_{\vec{y} \in \Omega} \mathbf{Pr}(X_{t+1} = \vec{y} \mid X_t = \vec{x}) &= 1 \\ \sum_{\vec{y} \in \Omega} \mathbf{Pr}(X_{t+1} = \vec{x} \mid X_t = \vec{y}) &= 1\end{aligned}$$

2. Demuestre que la distribución uniforme es una distribución estacionaria.
   - Vale decir, demuestre que si para $t \in \mathbb{N}$ se tiene que $\mathbf{Pr}(X_t = \vec{x}) = \frac{1}{|\Omega|}$ para cada $\vec{x} \in \Omega$, entonces:

$$\mathbf{Pr}(X_{t+1} = \vec{y}) = \mathbf{Pr}(X_t = \vec{y}) \text{ para cada } \vec{y} \in \Omega$$

Generando R_{KS} con distribución (casi) uniforme

Para generar los elementos de Ω con distribución (casi) uniforme utilizamos el siguiente procedimiento:

> Sea $\vec{x}$ un elemento arbitrario de Ω
> $t = f(|x|)$
> **for** $i := 1$ **to** t **do**
> $\vec{x} :=$ **GenerarSiguiente**$(\vec{x})$
> **return** $\vec{x}$

¿Qué condiciones deben cumplirse para que este procedimiento genere Ω con distribución (casi) uniforme y en tiempo polinomial?

Las condiciones para la generación (casi) uniforme

- La secuencia debe converger a la distribución uniforme desde cualquier punto de partida $\vec{x} \in \Omega$
- La distribución uniforme debe ser la única distribución a la que converge la secuencia
- $f(n)$ debe estar acotada superiormente por un polinomio

Las condiciones para la generación (casi) uniforme

- Debe ser posible calcular **GenerarSiguiente**$(\vec{x})$ en tiempo polinomial

- Después de ejecutar t pasos se debe tener una garantía de que estamos cerca de la distribución uniforme
  - Obtenemos entonces un generador casi uniforme para los elementos de Ω

Todas estas condiciones han sido estudiadas para las cadenas de Markov.

- Vamos a introducir y estudiar este herramienta esencial para el muestreo de variables aleatorias

Las cadenas de Markov

Considere una sucesión $\{X_t\}_{t \in \mathbb{N}}$ de variables aleatorias.

Definición

$\{X_t\}_{t \in \mathbb{N}}$ es una cadena de Markov con conjunto de estados Ω si:

1. *Ω es un conjunto finito o infinito enumerable, y $X_t : D_t \to \Omega$ para cada $t \in \mathbb{N}$*

2. *Existe $p : \Omega \times \Omega \to [0, 1]$ tal que para cada $t \in \mathbb{N}$ y cada secuencia $a_0, \ldots, a_t, a_{t+1}$ de elementos de Ω:*

$$\mathbf{Pr}(X_{t+1} = a_{t+1} \mid X_t = a_t \wedge \cdots \wedge X_0 = a_0) \;=\; \mathbf{Pr}(X_{t+1} = a_{t+1} \mid X_t = a_t) \;=\; p(a_{t+1}, a_t)$$

Las cadenas de Markov

En una cadena de Markov la distribución de probabilidades en el tiempo $t + 1$ sólo depende de la distribución de probabilidades en el tiempo t

- Consideramos cadenas discretas: el tiempo es un conjunto infinito enumerable, y el conjunto de estados Ω es finito o infinito enumerable

Además, consideramos cadenas de Markov donde las probabilidades de transición no dependen del tiempo.

- Para cada $t_1, t_2 \in \mathbb{N}$ y $a, b \in \Omega$:

$$\mathbf{Pr}(X_{t_1+1} = a \mid X_{t_1} = b) \quad = \quad \mathbf{Pr}(X_{t_2+1} = a \mid X_{t_2} = b)$$

Estos son llamadas cadenas de Markov homogéneas.

En general, consideramos cadenas de Markov con conjunto de estados Ω finito.

Representando una cadena de Markov

Sea $\{X_t\}_{t \in \mathbb{N}}$ una cadena de Markov con conjunto de estados Ω

- Suponemos además que $p : \Omega \times \Omega \to [0, 1]$ es la función que define las probabilidades de transición.

Podemos representar cada variable aleatoria X_t como un vector $\vec{x}_t$ tal que:

$$\text{para cada } a \in \Omega \text{ se tiene que } \vec{x}_t[a] = \mathbf{Pr}(X_t = a)$$

Además, podemos representar la cadena de Markov como una matriz P de $|\Omega| \times |\Omega|$:

$$\text{para cada } a, b \in \Omega, \text{ se tiene que } P[a, b] = p(a, b)$$

P es llamada la matriz de transición de la cadena de Markov.

Representando una cadena de Markov

Dado $a \in \Omega$ y $t \in \mathbb{N}$, tenemos que:

$$\begin{aligned}\mathbf{Pr}(X_{t+1} = a) \ &= \ \sum_{b \in \Omega} \mathbf{Pr}(X_{t+1} = a \mid X_t = b) \cdot \mathbf{Pr}(X_t = b) \ = \\ &\sum_{b \in \Omega} p(a, b) \cdot \mathbf{Pr}(X_t = b) \ = \ \sum_{b \in \Omega} P[a, b] \cdot \mathbf{Pr}(X_t = b)\end{aligned}$$

Dado que $\vec{x}_{t+1}[a] = \mathbf{Pr}(X_{t+1} = a)$ y $\vec{x}_t[b] = \mathbf{Pr}(X_t = b)$ para cada $a, b \in \Omega$ y $t \in \mathbb{N}$, concluimos que:

$$P\vec{x}_t \ = \ \vec{x}_{t+1}$$

Algunas propiedades de la matriz de transición

Ejercicios

Sea P la matriz de transición de una cadena de Markov.

1. Demuestre que para cada columna de P, se tiene que la suma de sus valores es 1
2. Para cada fila de P, ¿se debe tener que la suma de sus valores es 1?
   - ¿Era cierta esta propiedad para la matriz P de la cadena de Markov para R_{KS}?

Una cadena de Markov como un grafo

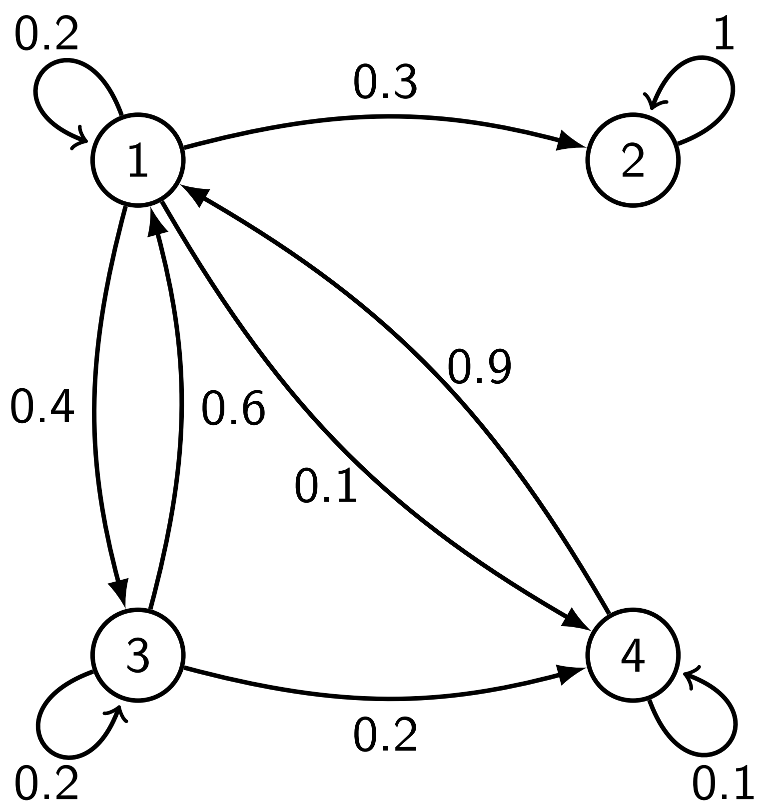
La matriz de transición P de una cadena de Markov con conjunto de estados Ω puede ser vista como un grafo G_P con pesos:

- Ω es el conjunto de nodos de G_P
- Dados $a, b \in \Omega$, el peso del arco (a, b) en G es $P[b, a]$
  - Vale decir, el peso de (a, b) representa la probabilidad de pasar al estado b dado que estábamos en el estado a

Una cadena de Markov como un grafo: un ejemplo

$$\Omega = \{1, 2, 3, 4\} \qquad\qquad P = \begin{pmatrix} 0.2 & 0 & 0.6 & 0.9 \\ 0.3 & 1 & 0 & 0 \\ 0.4 & 0 & 0.2 & 0 \\ 0.1 & 0 & 0.2 & 0.1 \end{pmatrix}$$

$G_P =$



Algunas propiedades de las cadenas de Markov

Sea $\{X_t\}_{t \in \mathbb{N}}$ una cadena de Markov con conjunto de estados Ω y matriz de transición P

- Además, defina P^0 como la matriz identidad y $P^{t+1} = PP^t$ para todo $t \in \mathbb{N}$

Ejercicios

1. Demuestre que para cada $t \in \mathbb{N}$ y $a, b \in \Omega$ se tiene que:

$$\mathbf{Pr}(X_t = a \mid X_0 = b) \quad = \quad P^t[a, b]$$

2. Demuestre para cada $t_1, t_2 \in \mathbb{N}$ y $a, b \in \Omega$ se tiene que:

$$\mathbf{Pr}(X_{t_1+t_2} = a \mid X_{t_1} = b) \quad = \quad \mathbf{Pr}(X_{t_2} = a \mid X_0 = b)$$

La distribución estacionaria de una cadena de Markov

Sea $\{X_t\}_{t \in \mathbb{N}}$ una de cadena de Markov con conjunto de estados Ω y matriz de transición P

- Y sea $\vec{\pi} \in [0,1]^{|\Omega|}$ tal que $\displaystyle\sum_{a \in \Omega} \vec{\pi}[a] = 1$

Definición

$\vec{\pi}$ es una distribución estacionaria para $\{X_t\}_{t \in \mathbb{N}}$ si $P\vec{\pi} = \vec{\pi}$

Vale decir, $\vec{\pi}$ es una distribución estacionaria para $\{X_t\}_{t \in \mathbb{N}}$ si esta distribución no cambia al realizar una transición de la cadena de Markov

Algunos ejemplos de distribuciones estacionarias

Ejercicios

1. Considere una cadena de Markov con la siguiente matriz de transición:

$$P = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$$

   Muestre que esta cadena de Markov tiene una única distribución estacionaria, y construya esta distribución.

2. Construya una cadena de Markov que tenga al menos dos distribuciones estacionarias.
   - El dominio de esta cadena debe ser finito

3. Construya una cadena de Markov que no tenga distribución estacionaria.
   - El dominio de esta cadena debe ser infinito enumerable