

Cadenas de Markov

IIC3810

Marcelo Arenas, Luis Alberto Croquevielle y Thomas
Reisenegger

Motivación: una pregunta pendiente del capítulo anterior

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

Motivación: una pregunta pendiente del capítulo anterior

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

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

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

Motivación: una pregunta pendiente del capítulo anterior

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

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

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

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 $\Omega \neq \emptyset$

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

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

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 $\mathbf{Pr}(X_t = \vec{x})$

Una secuencia de variables aleatorias para generar R_{KS}

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

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 $t + 1$

Una secuencia de variables aleatorias para generar R_{KS}

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

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 definir:

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

Una secuencia de variables aleatorias para generar R_{KS}

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

Una secuencia de variables aleatorias para generar R_{KS}

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

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 $i \in \{1, \dots, n\}$ con distribución uniforme

$\vec{u} := (x_1, \dots, x_{i-1}, 1 - x_i, x_{i+1}, \dots, x_n)$

if $\vec{a} \cdot \vec{u} \leq b$ **then return** $\vec{u}$

else return $\vec{x}$

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

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:

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

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

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

¿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**

¿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 ?

¿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 $\mathbf{Pr}(X_{t'+1} = \vec{y}) = \mathbf{Pr}(X_{t'} = \vec{y})$?

¿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?

¿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:

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

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

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  $\Omega$   
 $t = f(|\vec{a}| + |b|)$   
for  $i := 1$  to  $t$  do  
     $\vec{x} := \mathbf{GenerarSiguiente}(\vec{x})$   
return  $\vec{x}$ 
```

Donde $|\vec{a}| + |b|$ es el tamaño de la entrada $(\vec{a}, b)$

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  $\Omega$   
 $t = f(|\vec{a}| + |b|)$   
for  $i := 1$  to  $t$  do  
     $\vec{x} := \mathbf{GenerarSiguiente}(\vec{x})$   
return  $\vec{x}$ 
```

Donde $|\vec{a}| + |b|$ es el tamaño de la entrada $(\vec{a}, b)$

¿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 $\vec{x} \in \Omega$

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

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

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

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 esta herramienta esencial para el muestreo de variables aleatorias

Las cadenas de Markov

Considere una sucesión $\{X_t\}_{t \in \mathbb{N}}$ 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 $X_t : D_t \rightarrow \Omega$ para cada $t \in \mathbb{N}$
2. Existe $p : \Omega \times \Omega \rightarrow [0, 1]$ tal que para cada $t \in \mathbb{N}$ y cada secuencia $a_0, \dots, a_t, a_{t+1}$ de elementos de Ω :

$$\Pr(X_{t+1} = a_{t+1} \mid X_t = a_t \wedge \dots \wedge X_0 = a_0) =$$

$$\Pr(X_{t+1} = a_{t+1} \mid X_t = a_t) = p(a_{t+1}, a_t)$$

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

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 $a, b \in \Omega$:

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

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 $a, b \in \Omega$:

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

Estos son llamadas cadenas de Markov homogéneas.

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 $a, b \in \Omega$:

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

Estos son llamadas cadenas de Markov homogéneas.

En general, consideramos cadenas de Markov con conjuntos finitos de estados.

Representando una cadena de Markov

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

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

Representando una cadena de Markov

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

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

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

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

Representando una cadena de Markov

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

- Suponemos además que $p : \Omega \times \Omega \rightarrow [0, 1]$ 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:

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

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} ?

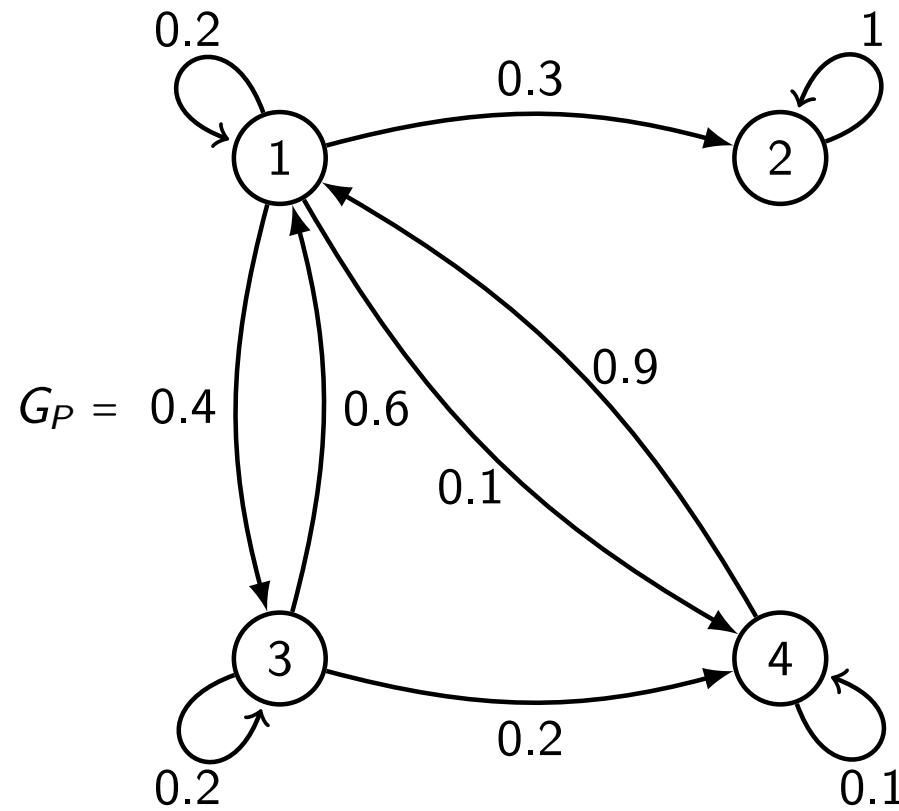
Una cadena de Markov como un grafo

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

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



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 $t \in \mathbb{N}$

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:

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

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

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

La distribución estacionaria de una cadena de Markov

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

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

La distribución estacionaria de una cadena de Markov

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

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

Definición

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

La distribución estacionaria de una cadena de Markov

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

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

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

Una primera propiedad fundamental: irreducibilidad

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

Una primera propiedad fundamental: irreducibilidad

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

Definición

$\{X_t\}_{t \in \mathbb{N}}$ es irreducible si G_P es un grafo fuertemente conexo.

Una primera propiedad fundamental: irreducibilidad

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

Definición

$\{X_t\}_{t \in \mathbb{N}}$ es irreducible si G_P es un grafo fuertemente conexo.

Tenemos que $\{X_t\}_{t \in \mathbb{N}}$ es irreducible si y sólo si para cada $a, b \in \Omega$, existe $t > 0$ tal que:

$$\Pr(X_t = a \mid X_0 = b) > 0$$

Una segunda propiedad fundamental: aperiodicidad

Sea $a \in \Omega$. Si $\{n > 0 \mid \mathbf{Pr}(X_n = a \mid X_0 = a) > 0\}$ no es vacío, entonces el periodo de a es definido como:

$$\text{MCD} \{n > 0 \mid \mathbf{Pr}(X_n = a \mid X_0 = a) > 0\}$$

En caso contrario, el periodo de a no está definido.

Una segunda propiedad fundamental: aperiodicidad

Sea $a \in \Omega$. Si $\{n > 0 \mid \mathbf{Pr}(X_n = a \mid X_0 = a) > 0\}$ no es vacío, entonces el periodo de a es definido como:

$$\text{MCD} \{n > 0 \mid \mathbf{Pr}(X_n = a \mid X_0 = a) > 0\}$$

En caso contrario, el periodo de a no está definido.

Definición

Un estado $a \in \Omega$ es aperiódico si su periodo está definido y es igual a 1. Además, $\{X_t\}_{t \in \mathbb{N}}$ es aperiódica si cada estado $b \in \Omega$ es aperiódico.

Irreducibilidad y aperiodicidad: algunos ejemplos

Ejercicios

1. Muestre que la cadena de Markov definida para la relación R_{KS} es irreducible y aperiódica.
2. Demuestre que en una cadena de Markov irreducible todos los estados tienen el mismo periodo.
 - ▶ Concluimos entonces que si una cadena de Markov irreducible tiene un estado aperiódico, entonces la cadena es aperiódica.

Una solución para el segundo ejercicio

Sea $\{X_t\}_{t \in \mathbb{N}}$ una cadena de Markov irreducible con conjunto de estados Ω y matriz de transición P , y sean $b, c \in \Omega$

Suponemos que los periodos de b y c son ℓ_b y ℓ_c , respectivamente.

- ▶ ¿Por qué sabemos que estos periodos existen?

Dado que $\{X_t\}_{t \in \mathbb{N}}$ es una cadena de Markov irreducible:

- ▶ Existe un camino en G_P desde b a c de largo $k_{b,c}$, y existe un camino en G_P desde c a b de largo $k_{c,b}$

Una solución para el segundo ejercicio

Entonces existe un camino de b a b de largo $k_{b,c} + k_{c,b}$, de lo cual concluimos que $\Pr(X_{k_{b,c}+k_{c,b}} = b \mid X_0 = b) > 0$

- ▶ Dado que ℓ_b es el periodo de b , concluimos que $\ell_b \mid (k_{b,c} + k_{c,b})$

Una solución para el segundo ejercicio

Entonces existe un camino de b a b de largo $k_{b,c} + k_{c,b}$, de lo cual concluimos que $\Pr(X_{k_{b,c}+k_{c,b}} = b \mid X_0 = b) > 0$

- ▶ Dado que ℓ_b es el periodo de b , concluimos que $\ell_b \mid (k_{b,c} + k_{c,b})$

Sea $n > 0$ tal que $\Pr(X_n = c \mid X_0 = c) > 0$

- ▶ Existe un camino en G_P de c a c de largo n

Una solución para el segundo ejercicio

Entonces existe un camino de b a b de largo $k_{b,c} + k_{c,b}$, de lo cual concluimos que $\Pr(X_{k_{b,c}+k_{c,b}} = b \mid X_0 = b) > 0$

- ▶ Dado que ℓ_b es el periodo de b , concluimos que $\ell_b \mid (k_{b,c} + k_{c,b})$

Sea $n > 0$ tal que $\Pr(X_n = c \mid X_0 = c) > 0$

- ▶ Existe un camino en G_P de c a c de largo n

Tenemos entonces que existe un camino en G_P de b a b de largo $k_{b,c} + n + k_{c,b}$, por lo que $\Pr(X_{k_{b,c}+n+k_{c,b}} = b \mid X_0 = b) > 0$

- ▶ Dado que ℓ_b es el periodo de b , concluimos que $\ell_b \mid (k_{b,c} + n + k_{c,b})$

Una solución para el segundo ejercicio

Entonces existe un camino de b a b de largo $k_{b,c} + k_{c,b}$, de lo cual concluimos que $\Pr(X_{k_{b,c}+k_{c,b}} = b \mid X_0 = b) > 0$

- ▶ Dado que ℓ_b es el periodo de b , concluimos que $\ell_b \mid (k_{b,c} + k_{c,b})$

Sea $n > 0$ tal que $\Pr(X_n = c \mid X_0 = c) > 0$

- ▶ Existe un camino en G_P de c a c de largo n

Tenemos entonces que existe un camino en G_P de b a b de largo $k_{b,c} + n + k_{c,b}$, por lo que $\Pr(X_{k_{b,c}+n+k_{c,b}} = b \mid X_0 = b) > 0$

- ▶ Dado que ℓ_b es el periodo de b , concluimos que $\ell_b \mid (k_{b,c} + n + k_{c,b})$
- ▶ Así, dado que $\ell_b \mid (k_{b,c} + k_{c,b})$, deducimos que $\ell_b \mid n$

Una solución para el segundo ejercicio

Por lo tanto, $\ell_b \mid n$ para cada $n > 0$ tal que $\mathbf{Pr}(X_n = c \mid X_0 = c) > 0$

- ▶ De lo cual deducimos que $\ell_b \leq \ell_c$, puesto que
 $\ell_c = \text{MCD} \{n > 0 \mid \mathbf{Pr}(X_n = c \mid X_0 = c) > 0\}$

Una solución para el segundo ejercicio

Por lo tanto, $\ell_b \mid n$ para cada $n > 0$ tal que $\Pr(X_n = c \mid X_0 = c) > 0$

- ▶ De lo cual deducimos que $\ell_b \leq \ell_c$, puesto que
 $\ell_c = \text{MCD} \{n > 0 \mid \Pr(X_n = c \mid X_0 = c) > 0\}$

De la misma forma se puede demostrar que $\ell_c \leq \ell_b$, de lo cual concluimos que $\ell_b = \ell_c$ □

Existencia del límite en el caso finito: irreducibilidad y aperiodicidad

Vamos a demostrar que irreducibilidad y aperiodicidad son condiciones suficientes para la convergencia de una cadena Markov con un conjunto finito de estados.

- ▶ La convergencia no depende del punto de partida

Existencia del límite en el caso finito: irreducibilidad y aperiodicidad

Vamos a demostrar que irreducibilidad y aperiodicidad son condiciones suficientes para la convergencia de una cadena Markov con un conjunto finito de estados.

- ▶ La convergencia no depende del punto de partida

Esto nos va a permitir demostrar que una cadena de Markov irreducible, aperiódica y con un conjunto finito de estados tiene una única distribución estacionaria.

Existencia del límite

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

Teorema

Si $a \in \Omega$, se tiene que:

1. $\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b)$ existe para cada $b \in \Omega$
2. $\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b) = \lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = c)$ para cada $b, c \in \Omega$

Existencia del límite

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

Teorema

Si $a \in \Omega$, se tiene que:

1. $\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b)$ existe para cada $b \in \Omega$
2. $\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b) = \lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = c)$ para cada $b, c \in \Omega$

Antes de demostrar este teorema, vamos a estudiar en detalle sus consecuencias.

Irreducibilidad, aperiodicidad y el caso finito

Sea $\{X_t\}_{t \in \mathbb{N}}$ una cadena de Markov irreducible, aperiódica y con conjunto finito de estados Ω

Dado $a \in \Omega$, definimos:

$$\lambda_a = \lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b),$$

donde b es un elemento arbitrario en Ω

Irreducibilidad, aperiodicidad y el caso finito

Sea $\{X_t\}_{t \in \mathbb{N}}$ una cadena de Markov irreducible, aperiódica y con conjunto finito de estados Ω

Dado $a \in \Omega$, definimos:

$$\lambda_a = \lim_{n \rightarrow \infty} \Pr(X_n = a \mid X_0 = b),$$

donde b es un elemento arbitrario en Ω

Lema

$$\sum_{a \in \Omega} \lambda_a = 1$$

La demostración del lema

Fije un elemento $b \in \Omega$

La demostración del lema

Fije un elemento $b \in \Omega$

Para cada $n \in \mathbb{N}$ tenemos que:

$$\sum_{a \in \Omega} \mathbf{Pr}(X_n = a \mid X_0 = b) = 1$$

La demostración del lema

Fije un elemento $b \in \Omega$

Para cada $n \in \mathbb{N}$ tenemos que:

$$\sum_{a \in \Omega} \mathbf{Pr}(X_n = a \mid X_0 = b) = 1$$

Por lo tanto tenemos que:

$$\lim_{n \rightarrow \infty} \left(\sum_{a \in \Omega} \mathbf{Pr}(X_n = a \mid X_0 = b) \right) = 1$$

La demostración del lema

Además, sabemos que:

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(\sum_{a \in \Omega} \mathbf{Pr}(X_n = a \mid X_0 = b) \right) &= \sum_{a \in \Omega} \lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b) \\ &= \sum_{a \in \Omega} \lambda_a \end{aligned}$$

La demostración del lema

Además, sabemos que:

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(\sum_{a \in \Omega} \mathbf{Pr}(X_n = a \mid X_0 = b) \right) &= \sum_{a \in \Omega} \lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b) \\ &= \sum_{a \in \Omega} \lambda_a \end{aligned}$$

Concluimos que $\sum_{a \in \Omega} \lambda_a = 1$



Irreducibilidad, aperiodicidad y el caso finito: una segunda propiedad fundamental

Lema

Si $\{X_t\}_{t \in \mathbb{N}}$ es una cadena de Markov irreducible, aperiódica y con un conjunto finito de estados Ω , entonces $\lambda_a > 0$ para cada $a \in \Omega$

Irreducibilidad, aperiodicidad y el caso finito: una segunda propiedad fundamental

Lema

Si $\{X_t\}_{t \in \mathbb{N}}$ es una cadena de Markov irreducible, aperiódica y con un conjunto finito de estados Ω , entonces $\lambda_a > 0$ para cada $a \in \Omega$

Ejercicio

Demuestre el lema.

Existencia y unicidad de la distribución estacionaria en el caso finito

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

- ▶ Y sea $\vec{\pi} \in [0, 1]^{|\Omega|}$ un vector tal que $\vec{\pi}[a] = \lambda_a$ para cada $a \in \Omega$

Existencia y unicidad de la distribución estacionaria en el caso finito

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

- Y sea $\vec{\pi} \in [0, 1]^{|\Omega|}$ un vector tal que $\vec{\pi}[a] = \lambda_a$ para cada $a \in \Omega$

Teorema

$\vec{\pi}$ es la única distribución estacionaria para $\{X_t\}_{t \in \mathbb{N}}$

Existencia y unicidad de la distribución estacionaria en el caso finito

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

- ▶ Y sea $\vec{\pi} \in [0, 1]^{|\Omega|}$ un vector tal que $\vec{\pi}[a] = \lambda_a$ para cada $a \in \Omega$

Teorema

$\vec{\pi}$ es la única distribución estacionaria para $\{X_t\}_{t \in \mathbb{N}}$

Demostración: Primero tenemos que demostrar que $P\vec{\pi} = \vec{\pi}$

- ▶ Vale decir, tenemos que demostrar que $(P\vec{\pi})[a] = \vec{\pi}[a]$ para cada $a \in \Omega$

Existencia y unicidad en el caso finito: demostración

Dado $a \in \Omega$ y un estado arbitrario $c \in \Omega$, tenemos que:

$$\begin{aligned}(P\vec{\pi})[a] &= \sum_{b \in \Omega} P[a, b] \cdot \vec{\pi}[b] \\&= \sum_{b \in \Omega} \left(\Pr(X_1 = a \mid X_0 = b) \cdot \lim_{n \rightarrow \infty} \Pr(X_n = b \mid X_0 = c) \right) \\&= \lim_{n \rightarrow \infty} \sum_{b \in \Omega} \left(\Pr(X_1 = a \mid X_0 = b) \cdot \Pr(X_n = b \mid X_0 = c) \right) \\&= \lim_{n \rightarrow \infty} \sum_{b \in \Omega} \left(\Pr(X_{n+1} = a \mid X_n = b) \cdot \Pr(X_n = b \mid X_0 = c) \right) \\&= \lim_{n \rightarrow \infty} \Pr(X_{n+1} = a \mid X_0 = c) \\&= \lim_{m \rightarrow \infty} \Pr(X_m = a \mid X_0 = c) \\&= \lambda_a = \vec{\pi}[a]\end{aligned}$$

Existencia y unicidad en el caso finito: demostración

En segundo lugar, tenemos que demostrar que para toda distribución estacionaria $\vec{\alpha}$ para $\{X_t\}_{t \in \mathbb{N}}$, se tiene que $\vec{\alpha} = \vec{\pi}$

Existencia y unicidad en el caso finito: demostración

En segundo lugar, tenemos que demostrar que para toda distribución estacionaria $\vec{\alpha}$ para $\{X_t\}_{t \in \mathbb{N}}$, se tiene que $\vec{\alpha} = \vec{\pi}$

Suponga que $P\vec{\alpha} = \vec{\alpha}$. Dado $a \in \Omega$ tenemos que:

$$\begin{aligned}\vec{\alpha}[a] &= (P\vec{\alpha})[a] \\ &= \sum_{b \in \Omega} P[a, b] \cdot \vec{\alpha}[b] \\ &= \sum_{b \in \Omega} P[a, b] \cdot (P\vec{\alpha})[b] \\ &= \sum_{b \in \Omega} P[a, b] \cdot \left(\sum_{c \in \Omega} P[b, c] \cdot \vec{\alpha}[c] \right) \\ &= \sum_{b \in \Omega} \sum_{c \in \Omega} P[a, b] \cdot P[b, c] \cdot \vec{\alpha}[c] \\ &= \sum_{c \in \Omega} \left(\sum_{b \in \Omega} P[a, b] \cdot P[b, c] \right) \cdot \vec{\alpha}[c]\end{aligned}$$

Existencia y unicidad en el caso finito: demostración

$$\begin{aligned} &= \sum_{c \in \Omega} P^2[a, c] \cdot \vec{\alpha}[c] \\ &= \sum_{c \in \Omega} P^2[a, c] \cdot (P\vec{\alpha})[c] \\ &= \sum_{c \in \Omega} P^2[a, c] \cdot \left(\sum_{d \in \Omega} P[c, d] \cdot \vec{\alpha}[d] \right) \\ &= \sum_{c \in \Omega} \sum_{d \in \Omega} P^2[a, c] \cdot P[c, d] \cdot \vec{\alpha}[d] \\ &= \sum_{d \in \Omega} \left(\sum_{c \in \Omega} P^2[a, c] \cdot P[c, d] \right) \cdot \vec{\alpha}[d] \\ &= \sum_{d \in \Omega} P^3[a, d] \cdot \vec{\alpha}[d] \end{aligned}$$

Existencia y unicidad en el caso finito: demostración

Siguiendo con este proceso, concluimos que para todo $n \geq 1$:

$$\vec{\alpha}[a] = \sum_{b \in \Omega} P^n[a, b] \cdot \vec{\alpha}[b]$$

Existencia y unicidad en el caso finito: demostración

Siguiendo con este proceso, concluimos que para todo $n \geq 1$:

$$\vec{\alpha}[a] = \sum_{b \in \Omega} P^n[a, b] \cdot \vec{\alpha}[b]$$

Por lo tanto:

$$\vec{\alpha}[a] = \sum_{b \in \Omega} \mathbf{Pr}(X_n = a \mid X_0 = b) \cdot \vec{\alpha}[b]$$

Existencia y unicidad en el caso finito: demostración

Tomando el límite cuando n tiende a infinito obtenemos:

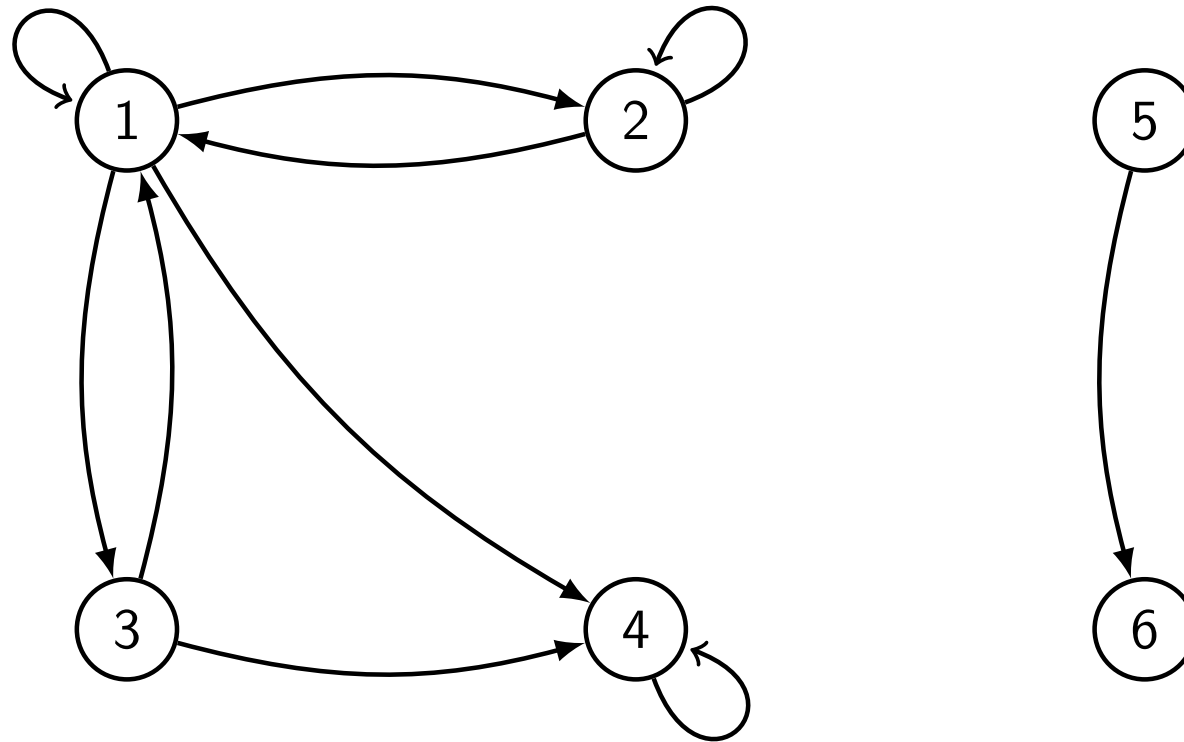
$$\begin{aligned}\vec{\alpha}[a] &= \lim_{n \rightarrow \infty} \left(\sum_{b \in \Omega} \mathbf{Pr}(X_n = a \mid X_0 = b) \cdot \vec{\alpha}[b] \right) \\ &= \sum_{b \in \Omega} \left(\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b) \right) \cdot \vec{\alpha}[b] \\ &= \sum_{b \in \Omega} \lambda_a \cdot \vec{\alpha}[b] \\ &= \lambda_a \cdot \sum_{b \in \Omega} \vec{\alpha}[b] \\ &= \lambda_a \cdot 1 = \lambda_a = \vec{\pi}[a]\end{aligned}$$

Concluimos que $\vec{\alpha} = \vec{\pi}$, puesto que $\vec{\alpha}[a] = \vec{\pi}[a]$ para todo $a \in \Omega$



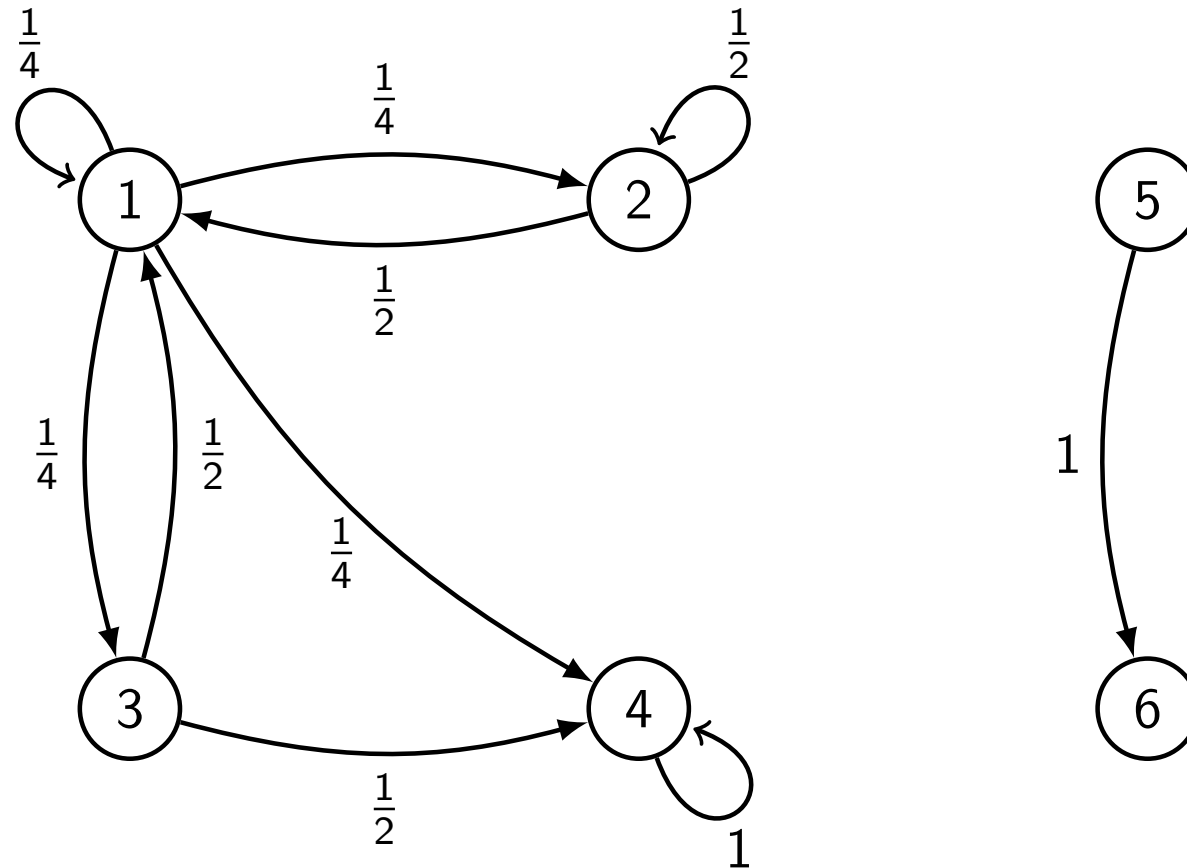
Un ejemplo importante: PageRank

La estructura de links en una pequeña Web:



Un ejemplo importante: PageRank

La estructura de links vista como una caminata aleatoria:



Un ejemplo importante: PageRank

Podemos representar la caminata aleatoria como una matriz P

- ▶ $P[i,j]$ es la probabilidad de ir de la página j a la página i

Un ejemplo importante: PageRank

Podemos representar la caminata aleatoria como una matriz P

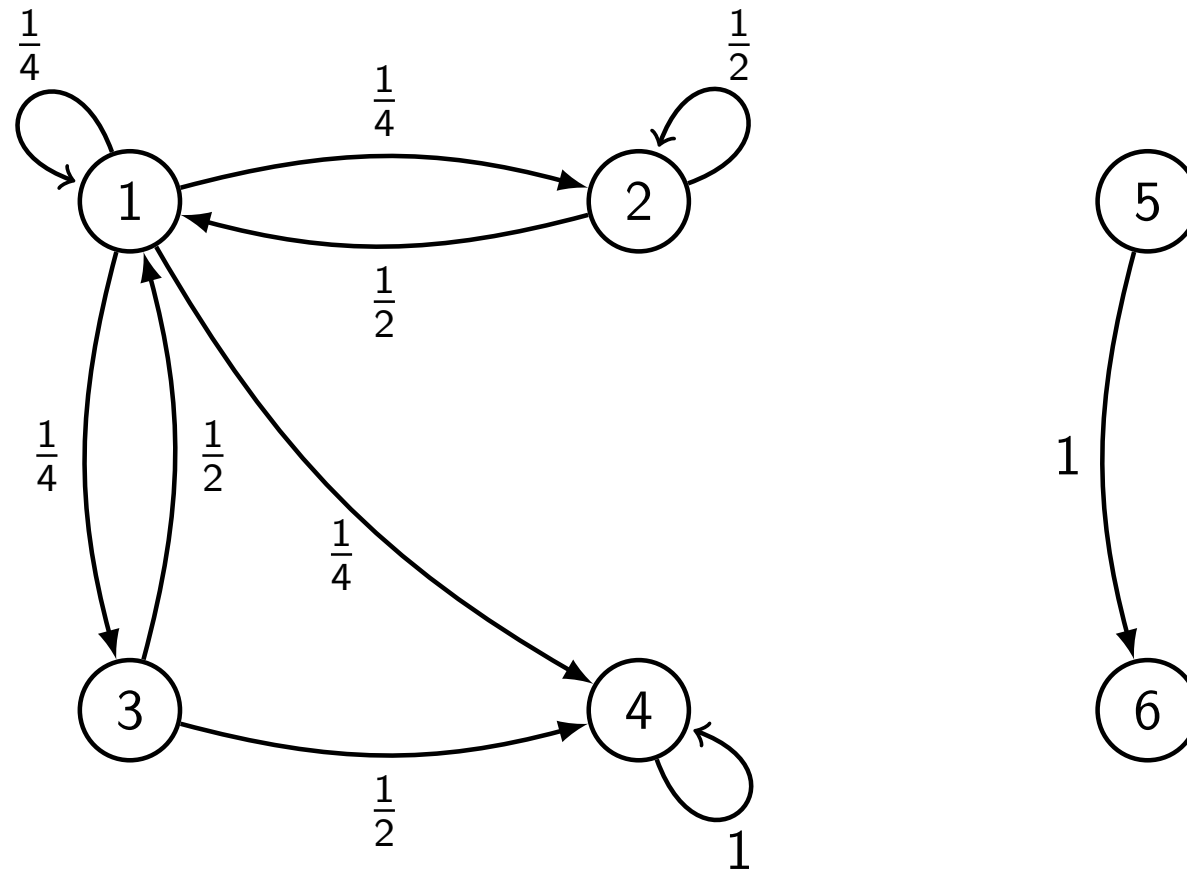
- ▶ $P[i,j]$ es la probabilidad de ir de la página j a la página i

La matriz P de nuestro ejemplo:

$$P = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{2} & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

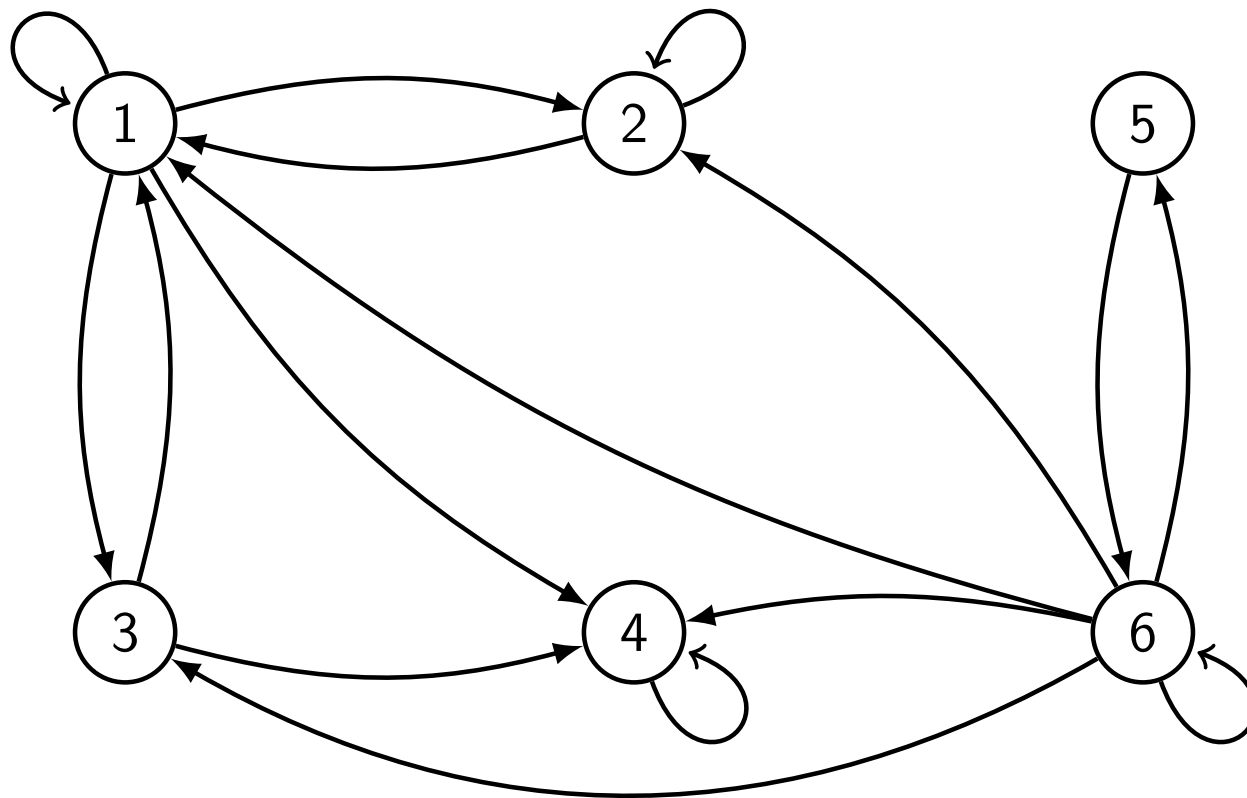
Un ejemplo importante: PageRank

Esperamos que la caminata aleatoria no se detenga, lo cual no sucede en nuestra pequeña Web porque la página 6 no tiene links de salida:



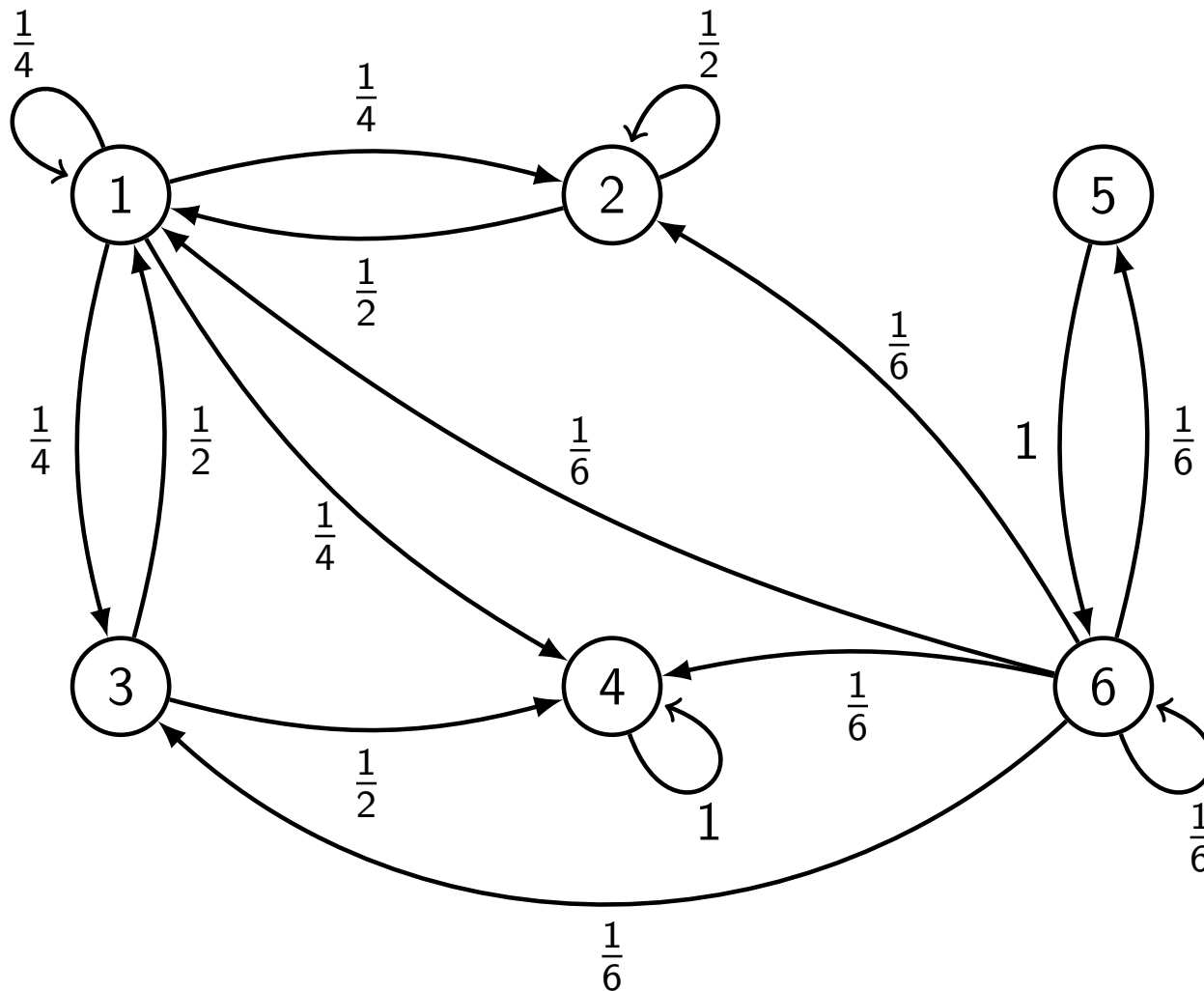
Un ejemplo importante: PageRank

Para solucionar el problema, suponemos que la página 6 no tiene preferencias al navegar y la conectamos con todas las páginas:



Un ejemplo importante: PageRank

Obtenemos entonces el siguiente grafo para la caminata aleatoria:



Un ejemplo importante: PageRank

Así, consideramos la siguiente matriz en nuestro ejemplo:

$$P = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{6} \\ \frac{1}{4} & 0 & 0 & 0 & 0 & \frac{1}{6} \\ \frac{1}{4} & 0 & \frac{1}{2} & 1 & 0 & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 1 & \frac{1}{6} \end{pmatrix}$$

Un ejemplo importante: PageRank

El ranking de una página i se define como la probabilidad x_i de que el caminante aleatorio llegue a i

Un ejemplo importante: PageRank

El ranking de una página i se define como la probabilidad x_i de que el caminante aleatorio llegue a i

- ▶ ¿Qué características de la Web influyen en el valor x_i ?

Un ejemplo importante: PageRank

El ranking de una página i se define como la probabilidad x_i de que el caminante aleatorio llegue a i

- ▶ ¿Qué características de la Web influyen en el valor x_i ?

En el ejemplo tenemos que:

$$x_i = \sum_{j=1}^6 P[i,j] \cdot x_j$$

Un ejemplo importante: PageRank

El ranking de una página i se define como la probabilidad x_i de que el caminante aleatorio llegue a i

- ▶ ¿Qué características de la Web influyen en el valor x_i ?

En el ejemplo tenemos que:

$$x_i = \sum_{j=1}^6 P[i, j] \cdot x_j$$

Así, si $\vec{x}$ es un vector tal que $\vec{x}[i] = x_i$ para cada $i \in \{1, \dots, 6\}$, entonces:

$$P\vec{x} = \vec{x}$$

PageRank como una cadena de Markov

El proceso anterior se puede generalizar para la Web real.

- ▶ Suponemos que la Web tiene N páginas
- ▶ Cada página $i \in N$ que no tiene links de salida es conectada con cada página $j \in N$
- ▶ P es una matriz de $N \times N$ construida a partir de la estructura de links generada, como fue mostrado en las transparencias anteriores

Además, definimos $\vec{\pi}$ como un vector tal que $\vec{\pi}[i]$ es la probabilidad de que el caminante aleatorio llegue a la página i , para cada $i \in \{1, \dots, N\}$

PageRank como una cadena de Markov

P puede ser considerada como la matriz de transición de una cadena de Markov $\{X_t\}_{t \in \mathbb{N}}$ con conjunto de estados $\Omega = \{1, \dots, N\}$

PageRank como una cadena de Markov

P puede ser considerada como la matriz de transición de una cadena de Markov $\{X_t\}_{t \in \mathbb{N}}$ con conjunto de estados $\Omega = \{1, \dots, N\}$

- ▶ El vector $\vec{\pi}$ es entonces una distribución estacionaria de $\{X_t\}_{t \in \mathbb{N}}$

PageRank como una cadena de Markov

P puede ser considerada como la matriz de transición de una cadena de Markov $\{X_t\}_{t \in \mathbb{N}}$ con conjunto de estados $\Omega = \{1, \dots, N\}$

- ▶ El vector $\vec{\pi}$ es entonces una distribución estacionaria de $\{X_t\}_{t \in \mathbb{N}}$

¿Cómo podemos asegurar que $\vec{\pi}$ existe y es único? ¿Cómo podemos calcular $\vec{\pi}$?

PageRank como una cadena de Markov

P puede ser considerada como la matriz de transición de una cadena de Markov $\{X_t\}_{t \in \mathbb{N}}$ con conjunto de estados $\Omega = \{1, \dots, N\}$

- ▶ El vector $\vec{\pi}$ es entonces una distribución estacionaria de $\{X_t\}_{t \in \mathbb{N}}$

¿Cómo podemos asegurar que $\vec{\pi}$ existe y es único? ¿Cómo podemos calcular $\vec{\pi}$?

¡Irreducibilidad y aperiodicidad son las propiedades que necesitamos!

- ▶ Dado que el conjunto de estados de la cadena de Markov es finito

Irreducibilidad, aperiodicidad y PageRank

Dos preguntas por responder:

- ▶ ¿Podemos asegurar que la cadena de Markov dada por P es irreducible?
- ▶ ¿Podemos asegurar que la cadena de Markov dada por P es aperiódica?

Irreducibilidad, aperiodicidad y PageRank

Dos preguntas por responder:

- ▶ ¿Podemos asegurar que la cadena de Markov dada por P es irreducible? **No**
- ▶ ¿Podemos asegurar que la cadena de Markov dada por P es aperiódica?

Irreducibilidad, aperiodicidad y PageRank

Dos preguntas por responder:

- ▶ ¿Podemos asegurar que la cadena de Markov dada por P es irreducible? **No**
- ▶ ¿Podemos asegurar que la cadena de Markov dada por P es aperiódica? **No**

Irreducibilidad, aperiodicidad y PageRank

Dos preguntas por responder:

- ▶ ¿Podemos asegurar que la cadena de Markov dada por P es irreducible? **No**
- ▶ ¿Podemos asegurar que la cadena de Markov dada por P es aperiódica? **No**

¿Cómo solucionamos estos problemas?

Irreducibilidad, aperiodicidad y PageRank

Para solucionar los problemas suponemos que el caminante aleatorio también puede decidir moverse a cualquier página sin considerar la estructura de la Web.

- ▶ Esto es controlado por un parámetro α

Irreducibilidad, aperiodicidad y PageRank

Para solucionar los problemas suponemos que el caminante aleatorio también puede decidir moverse a cualquier página sin considerar la estructura de la Web.

- ▶ Esto es controlado por un parámetro α

Sea U una matriz de $N \times N$ tal que $U[i, j] = \frac{1}{N}$, y sea $\alpha \in (0, 1)$

Irreducibilidad, aperiodicidad y PageRank

Para solucionar los problemas suponemos que el caminante aleatorio también puede decidir moverse a cualquier página sin considerar la estructura de la Web.

- ▶ Esto es controlado por un parámetro α

Sea U una matriz de $N \times N$ tal que $U[i, j] = \frac{1}{N}$, y sea $\alpha \in (0, 1)$

Definimos una matriz W_α de $N \times N$ como:

$$W_\alpha = \alpha \cdot P + (1 - \alpha) \cdot U$$

Irreducibilidad, aperiodicidad y PageRank

Con probabilidad α el caminante utiliza los links para moverse, y con probabilidad $(1 - \alpha)$ se mueve a cualquier página (sin considerar la estructura de la Web).

- ▶ Cuanto más grande es α , más consideramos la estructura de la Web

Irreducibilidad, aperiodicidad y PageRank

Con probabilidad α el caminante utiliza los links para moverse, y con probabilidad $(1 - \alpha)$ se mueve a cualquier página (sin considerar la estructura de la Web).

- ▶ Cuanto más grande es α , más consideramos la estructura de la Web

Tenemos que W_α define una cadena de Markov irreducible y aperiódica.

- ▶ ¿Por qué?

La definición de PageRank

Consideramos un valor fijo de α

- ▶ En el artículo original sobre PageRank se consideraba $\alpha = 0.85$

La definición de PageRank

Consideramos un valor fijo de α

- ▶ En el artículo original sobre PageRank se consideraba $\alpha = 0.85$

Definimos entonces $\vec{\pi}$ como la distribución estacionaria de la cadena de Markov $\{X_t\}_{t \in \mathbb{N}}$ cuya matriz de transición es W_α

- ▶ Esta distribución existe y es única porque la cadena de Markov es irreducible, aperiódica y tiene un conjunto finito de estados
- ▶ Además se tiene que $\vec{\pi}[i] > 0$ para cada $i \in \{1, \dots, N\}$, por lo que ninguna página Web es considerada absolutamente irrelevante 😊

El cálculo de PageRank

Para cada página i , calculamos $\vec{\pi}[i]$ considerando la igualdad:

$$\vec{\pi}[i] = \lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = i \mid X_0 = 1)$$

El cálculo de PageRank

Para cada página i , calculamos $\vec{\pi}[i]$ considerando la igualdad:

$$\vec{\pi}[i] = \lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = i \mid X_0 = 1)$$

Vale decir, la caminata comienza en la página 1 (una página arbitraria), y el valor $\vec{\pi}[i]$ se calcula utilizando un valor de n suficientemente grande.

El cálculo de PageRank

Para cada página i , calculamos $\vec{\pi}[i]$ considerando la igualdad:

$$\vec{\pi}[i] = \lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = i \mid X_0 = 1)$$

Vale decir, la caminata comienza en la página 1 (una página arbitraria), y el valor $\vec{\pi}[i]$ se calcula utilizando un valor de n suficientemente grande.

- Obtenemos una aproximación del valor de $\vec{\pi}[i]$

El cálculo de PageRank

Las ideas anteriores se traducen en la práctica en el siguiente algoritmo

AproxPageRank para aproximar $\vec{\pi}$

En **AproxPageRank** utilizamos un parámetro $\varepsilon \in (0, 1)$ que indica cuándo debe detenerse el algoritmo.

- ▶ Si la distancia entre los últimos vectores generados es menor que ε el algoritmo se detiene

En **AproxPageRank** la similitud entre dos vectores $\vec{a}$ y $\vec{b}$ es medida a través de la distancia Manhattan:

$$\|\vec{a} - \vec{b}\|_1 = \sum_{i=1}^N |\vec{a}[i] - \vec{b}[i]|$$

El algoritmo **AproxPageRank**

AproxPageRank(ε)

Sea $\vec{\pi}_0$ tal que $\vec{\pi}_0[1] = 1$ y $\vec{\pi}_0[i] = 0$ para cada $i \in \{2, \dots, N\}$

Sea $\vec{\pi}_1 = W_\alpha \vec{\pi}_0$

if $\|\vec{\pi}_0 - \vec{\pi}_1\|_1 < \varepsilon$ **then return** $\vec{\pi}_1$

else

for $n := 2$ **to** ∞ **do**

$\vec{\pi}_n = W_\alpha \vec{\pi}_{n-1}$

if $\|\vec{\pi}_{n-1} - \vec{\pi}_n\|_1 < \varepsilon$ **then return** $\vec{\pi}_n$

La demostración de la existencia del límite

Sea $\{X_t\}_{t \in \mathbb{N}}$ una cadena de Markov irreducible y aperiódica con un conjunto finito de estados Ω

En las transparencias anteriores utilizamos la siguiente propiedad para cada $a \in \Omega$:

1. $\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b)$ existe para cada $b \in \Omega$
2. $\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b) = \lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = c)$ para cada $b, c \in \Omega$

La demostración de la existencia del límite

Sea $\{X_t\}_{t \in \mathbb{N}}$ una cadena de Markov irreducible y aperiódica con un conjunto finito de estados Ω

En las transparencias anteriores utilizamos la siguiente propiedad para cada $a \in \Omega$:

1. $\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b)$ existe para cada $b \in \Omega$
2. $\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b) = \lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = c)$ para cada $b, c \in \Omega$

En las siguientes transparencias vamos a demostrar esta propiedad.

La existencia del límite: una noción fundamental

Definición

Dada una matriz A de $\ell \times \ell$ de números reales.

1. *A es no negativa si para todo $a, b \in \{1, \dots, \ell\}$, se tiene que $A[a, b] \geq 0$*

La existencia del límite: una noción fundamental

Definición

Dada una matriz A de $\ell \times \ell$ de números reales.

- 1. A es no negativa si para todo $a, b \in \{1, \dots, \ell\}$, se tiene que $A[a, b] \geq 0$*
- 2. A es cuasipositiva si A es no negativa y existe un número natural $n \geq 1$, tal que para todo $a, b \in \{1, \dots, \ell\}$, se tiene que $A^n[a, b] > 0$*

Irreducibilidad, aperiodicidad y cuasipositividad

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

Proposición

P es una matriz cuasipositiva

Irreducibilidad, aperiodicidad y cuasipositividad

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

Proposición

P es una matriz cuasipositiva

Demostración: Dado $a \in \Omega$, defina $J(a) = \{n \geq 1 \mid P^n[a, a] > 0\}$

- ▶ Por aperiodicidad, se tiene que $\text{MCD}(J(a)) = 1$

La demostración de la proposición

Nótese que $P^{n+m}[a, a] \geq P^n[a, a] \cdot P^m[a, a]$

- ▶ Por lo tanto, tenemos la siguiente propiedad de clausura: si $n, m \in J(a)$, entonces $n + m \in J(a)$

La demostración de la proposición

Nótese que $P^{n+m}[a, a] \geq P^n[a, a] \cdot P^m[a, a]$

- ▶ Por lo tanto, tenemos la siguiente propiedad de clausura: si $n, m \in J(a)$, entonces $n + m \in J(a)$

Lema

Para todo $a \in \Omega$, existe una par de números consecutivos en $J(a)$

La demostración de la proposición

Nótese que $P^{n+m}[a, a] \geq P^n[a, a] \cdot P^m[a, a]$

- ▶ Por lo tanto, tenemos la siguiente propiedad de clausura: si $n, m \in J(a)$, entonces $n + m \in J(a)$

Lema

Para todo $a \in \Omega$, existe una par de números consecutivos en $J(a)$

Demostración: suponga que existe $a \in \Omega$ tal que $J(a)$ no contiene algún par de números consecutivos

- ▶ Hay una distancia mínima $k \geq 2$ entre los elementos de $J(a)$

La demostración de la proposición

Tenemos entonces que existe $n_0 \in J(a)$ tal que $n_0 + k \in J(a)$

La demostración de la proposición

Tenemos entonces que existe $n_0 \in J(a)$ tal que $n_0 + k \in J(a)$

Además, existe $n_1 \in J(a)$ tal que $n_1 = m \cdot k + d$, con $m \in \mathbb{N}$ y $1 \leq d \leq k - 1$

- ▶ De otra forma, se tendría que $k|n$ para cada $n \in J(a)$, lo cual implicaría que $\text{MCD}(J(a)) \geq k > 1$

La demostración de la proposición

Tenemos entonces que existe $n_0 \in J(a)$ tal que $n_0 + k \in J(a)$

Además, existe $n_1 \in J(a)$ tal que $n_1 = m \cdot k + d$, con $m \in \mathbb{N}$ y $1 \leq d \leq k - 1$

- ▶ De otra forma, se tendría que $k|n$ para cada $n \in J(a)$, lo cual implicaría que $\text{MCD}(J(a)) \geq k > 1$

Por lo tanto, por la propiedad de clausura: $(m + 1) \cdot (n_0 + k) \in J(a)$ y $n_1 + (m + 1) \cdot n_0 \in J(a)$

La demostración de la proposición

La distancia entre $(m + 1) \cdot (n_0 + k)$ y $n_1 + (m + 1) \cdot n_0$ es:

$$\begin{aligned}(m + 1) \cdot (n_0 + k) - (n_1 + (m + 1) \cdot n_0) &= (m + 1) \cdot k - n_1 \\ &= k - (n_1 - m \cdot k) \\ &= k - d \\ &< k\end{aligned}$$

La demostración de la proposición

La distancia entre $(m + 1) \cdot (n_0 + k)$ y $n_1 + (m + 1) \cdot n_0$ es:

$$\begin{aligned}(m + 1) \cdot (n_0 + k) - (n_1 + (m + 1) \cdot n_0) &= (m + 1) \cdot k - n_1 \\ &= k - (n_1 - m \cdot k) \\ &= k - d \\ &< k\end{aligned}$$

Encontramos entonces dos números en $J(a)$ cuya distancia es menor que k , lo cual contradice el supuesto inicial.

- Concluimos que existe un par de números consecutivos en $J(a)$ □

La demostración de la proposición

Lema

Dado $k \in \mathbb{N}$ con $k \geq 1$, existe $n_0 \in \mathbb{N}$ tal que $n_0 \geq 1$ y:

$$\{n \in \mathbb{N} \mid n \geq n_0\} \subseteq \{n_1 \cdot k + n_2 \cdot (k + 1) \mid n_1, n_2 \in \mathbb{N} \text{ y } n_1 \geq 1\}$$

La demostración de la proposición

Lema

Dado $k \in \mathbb{N}$ con $k \geq 1$, existe $n_0 \in \mathbb{N}$ tal que $n_0 \geq 1$ y:

$$\{n \in \mathbb{N} \mid n \geq n_0\} \subseteq \{n_1 \cdot k + n_2 \cdot (k + 1) \mid n_1, n_2 \in \mathbb{N} \text{ y } n_1 \geq 1\}$$

Demostración: Dado $n \geq k^2$, existen $m \in \mathbb{N}$ y $d \in \{0, \dots, k-1\}$ tales que

$$n - k^2 = m \cdot k + d$$

La demostración de la proposición

Podemos entonces expresar n como:

$$\begin{aligned} n &= k^2 + m \cdot k + d \\ &= (k - d + m) \cdot k + d \cdot (k + 1) \end{aligned}$$

Dado que $k - d > 0$, concluimos que para $n_0 = k^2$ se tiene que $n_0 \geq 1$ y:

$$\{n \in \mathbb{N} \mid n \geq n_0\} \subseteq \{n_1 \cdot k + n_2 \cdot (k + 1) \mid n_1, n_2 \in \mathbb{N} \text{ y } n_1 \geq 1\}$$



La demostración de la proposición

Fije $c \in \Omega$

Por el primer lema, existe $k \in J(c)$ tal que $k + 1$ también es un elemento de $J(c)$

- ▶ Por la definición de $J(c)$ sabemos que $k \geq 1$

Por lo tanto, por el segundo lema existe $n_c \geq 1$ tal que:

$$\begin{aligned} \{n \in \mathbb{N} \mid n \geq n_c\} &\subseteq \{n_1 \cdot k + n_2 \cdot (k + 1) \mid n_1, n_2 \in \mathbb{N} \text{ y } n_1 \geq 1\} \\ &\subseteq J(c) \end{aligned}$$

Nótese que la última inclusión es consecuencia de que $n + m \in J(c)$ si $n, m \in J(c)$

La demostración de la proposición

Sean $a, b \in \Omega$

- ▶ Como la cadena de Markov considerada es irreducible, existen $r_{a,c}, s_{c,b} \geq 1$ tales que $P^{r_{a,c}}[a, c] > 0$ y $P^{s_{c,b}}[c, b] > 0$

La demostración de la proposición

Sean $a, b \in \Omega$

- ▶ Como la cadena de Markov considerada es irreducible, existen $r_{a,c}, s_{c,b} \geq 1$ tales que $P^{r_{a,c}}[a, c] > 0$ y $P^{s_{c,b}}[c, b] > 0$

Sea $n_{a,b} = r_{a,c} + n_c + s_{c,b}$

La demostración de la proposición

Sean $a, b \in \Omega$

- ▶ Como la cadena de Markov considerada es irreducible, existen $r_{a,c}, s_{c,b} \geq 1$ tales que $P^{r_{a,c}}[a, c] > 0$ y $P^{s_{c,b}}[c, b] > 0$

Sea $n_{a,b} = r_{a,c} + n_c + s_{c,b}$

Dado $n \geq n_{a,b}$ con $\ell = n - n_{a,b}$, tenemos que:

$$\begin{aligned} P^n[a, b] &= P^{n_{a,b}+\ell}[a, b] \\ &= P^{r_{a,c}+n_c+\ell+s_{c,b}}[a, b] \\ &\geq P^{r_{a,c}}[a, c] \cdot P^{n_c+\ell}[c, c] \cdot P^{s_{c,b}}[c, b] > 0 \end{aligned}$$

Nótese que $P^{n_c+\ell}[c, c] > 0$ dado que $\{n \in \mathbb{N} \mid n \geq n_c\} \subseteq J(c)$

La demostración de la proposición

Para todo $a, b \in \Omega$ y para todo $n \geq n_{a,b}$, tenemos que $P^n[a, b] > 0$

La demostración de la proposición

Para todo $a, b \in \Omega$ y para todo $n \geq n_{a,b}$, tenemos que $P^n[a, b] > 0$

Sea $n^* = \max_{a,b \in \Omega} n_{a,b}$

La demostración de la proposición

Para todo $a, b \in \Omega$ y para todo $n \geq n_{a,b}$, tenemos que $P^n[a, b] > 0$

Sea $n^* = \max_{a,b \in \Omega} n_{a,b}$

Tenemos que $P^{n^*}[a, b] > 0$ para todo $a, b \in \Omega$

- Concluimos que P es cuasipositiva (dado que P es no negativa) $\square$

Existencia del límite: cuasipositividad

Sea $\{X_t\}_{t \in \mathbb{N}}$ una cadena de Markov con conjunto finito de estados Ω y matriz de transición P , donde P es cuasipositiva.

Teorema

Si $a \in \Omega$, se tiene que:

1. $\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b)$ existe para cada $b \in \Omega$
2. $\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b) = \lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = c)$ para cada $b, c \in \Omega$

Existencia del límite: cuasipositividad

Sea $\{X_t\}_{t \in \mathbb{N}}$ una cadena de Markov con conjunto finito de estados Ω y matriz de transición P , donde P es cuasipositiva.

Teorema

Si $a \in \Omega$, se tiene que:

1. $\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b)$ existe para cada $b \in \Omega$
2. $\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b) = \lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = c)$ para cada $b, c \in \Omega$

Obtenemos entonces como corolario el teorema que queríamos demostrar, dada la proposición anterior.

La demostración de la existencia del límite

Fije $a \in \Omega$

Dado $n \in \mathbb{N}$, definimos:

$$\begin{aligned} m_a^{(n)} &= \min_{b \in \Omega} P^n[a, b] \\ M_a^{(n)} &= \max_{b \in \Omega} P^n[a, b] \end{aligned}$$

La demostración de la existencia del límite

Tenemos que $m_a^{(n)} \leq m_a^{(n+1)}$, puesto que:

$$\begin{aligned} m_a^{(n+1)} &= \min_{b \in \Omega} P^{n+1}[a, b] \\ &= \min_{b \in \Omega} \sum_{c \in \Omega} P^n[a, c] \cdot P[c, b] \\ &\geq \min_{b \in \Omega} \sum_{c \in \Omega} \left(\min_{d \in \Omega} P^n[a, d] \right) P[c, b] \\ &= \min_{b \in \Omega} \left(\min_{d \in \Omega} P^n[a, d] \right) \sum_{c \in \Omega} P[c, b] \\ &= \min_{b \in \Omega} \left(\min_{d \in \Omega} P^n[a, d] \right) \\ &= \min_{d \in \Omega} P^n[a, d] = m_a^{(n)} \end{aligned}$$

La demostración de la existencia del límite

Tenemos que $m_a^{(n)} \leq m_a^{(n+1)}$, puesto que:

$$\begin{aligned} m_a^{(n+1)} &= \min_{b \in \Omega} P^{n+1}[a, b] \\ &= \min_{b \in \Omega} \sum_{c \in \Omega} P^n[a, c] \cdot P[c, b] \\ &\geq \min_{b \in \Omega} \sum_{c \in \Omega} \left(\min_{d \in \Omega} P^n[a, d] \right) P[c, b] \\ &= \min_{b \in \Omega} \left(\min_{d \in \Omega} P^n[a, d] \right) \sum_{c \in \Omega} P[c, b] \\ &= \min_{b \in \Omega} \left(\min_{d \in \Omega} P^n[a, d] \right) \\ &= \min_{d \in \Omega} P^n[a, d] = m_a^{(n)} \end{aligned}$$

De la misma forma se puede demostrar que $M_a^{(n)} \geq M_a^{(n+1)}$

La demostración de la existencia del límite

Lema

Existen $r_a, s_a \in \mathbb{R}$ tales que:

$$\lim_{n \rightarrow \infty} m_a^{(n)} = r_a$$

$$\lim_{n \rightarrow \infty} M_a^{(n)} = s_a$$

La demostración de la existencia del límite

Lema

Existen $r_a, s_a \in \mathbb{R}$ tales que:

$$\begin{aligned}\lim_{n \rightarrow \infty} m_a^{(n)} &= r_a \\ \lim_{n \rightarrow \infty} M_a^{(n)} &= s_a\end{aligned}$$

Ejercicio

Demuestre el lema considerando que $(m_a^{(n)})_{n \in \mathbb{N}}$ es una sucesión de números reales monótona creciente y acotada superiormente.

- ¿Qué propiedades debe usar el caso de la sucesión $(M_a^{(n)})_{n \in \mathbb{N}}$?

La demostración de la existencia del límite

Para demostrar el teorema basta entonces demostrar que:

$$\lim_{n \rightarrow \infty} (M_a^{(n)} - m_a^{(n)}) = 0$$

De hecho, si esto se cumple, para cada $b \in \Omega$ se tiene que:

$$\lim_{n \rightarrow \infty} \mathbf{Pr}(X_n = a \mid X_0 = b) = r_a$$

Dado que $\mathbf{Pr}(X_n = a \mid X_0 = b) = P^n[a, b]$, $m_a^{(n)} \leq P^n[a, b] \leq M_a^{(n)}$
y $r_a = s_a$

La demostración de la existencia del límite

Sea $\ell = |\Omega|$, y sea $n_0 \in \mathbb{N}$ tal que $n_0 \geq 1$ y $P^{n_0}[b, c] > 0$ para cada $b, c \in \Omega$

- ▶ Además, sea $t = \min_{b, c \in \Omega} P^{n_0}[b, c]$

La demostración de la existencia del límite

Sea $\ell = |\Omega|$, y sea $n_0 \in \mathbb{N}$ tal que $n_0 \geq 1$ y $P^{n_0}[b, c] > 0$ para cada $b, c \in \Omega$

- ▶ Además, sea $t = \min_{b, c \in \Omega} P^{n_0}[b, c]$

Dados estados $b_0, c_0 \in \Omega$, defina:

$$\begin{aligned}\Omega_1 &= \{d \in \Omega \mid P^{n_0}[d, b_0] \geq P^{n_0}[d, c_0]\} \\ \Omega_2 &= \Omega \setminus \Omega_1\end{aligned}$$

La demostración de la existencia del límite

Nótese que:

$$\begin{aligned}\sum_{d \in \Omega_1} P^{n_0}[d, b_0] + \sum_{d \in \Omega_2} P^{n_0}[d, b_0] &= \sum_{d \in \Omega} P^{n_0}[d, b_0] \\ &= 1\end{aligned}$$

La demostración de la existencia del límite

Nótese que:

$$\begin{aligned}\sum_{d \in \Omega_1} P^{n_0}[d, b_0] + \sum_{d \in \Omega_2} P^{n_0}[d, b_0] &= \sum_{d \in \Omega} P^{n_0}[d, b_0] \\ &= 1\end{aligned}$$

A partir de lo cual obtenemos:

$$\begin{aligned}\sum_{d \in \Omega_1} (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) &= 1 - \sum_{d \in \Omega_2} P^{n_0}[d, b_0] - \sum_{c \in \Omega_1} P^{n_0}[d, c_0] \\ &\leq 1 - \sum_{d \in \Omega_2} t - \sum_{d \in \Omega_1} t \\ &= 1 - \ell \cdot t\end{aligned}$$

La demostración de la existencia del límite

Dado $n \geq 0$, tenemos que:

$$\begin{aligned} P^{n_0+n}[a, b_0] - P^{n_0+n}[a, c_0] &= \sum_{d \in \Omega} P^n[a, d] \cdot P^{n_0}[d, b_0] - \sum_{d \in \Omega} P^n[a, d] \cdot P^{n_0}[d, c_0] \\ &= \sum_{d \in \Omega} P^n[a, d] \cdot (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) \\ &= \sum_{d \in \Omega_1} P^n[a, d] \cdot (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) + \\ &\quad \sum_{d \in \Omega_2} P^n[a, d] \cdot (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) \\ &\leq \sum_{d \in \Omega_1} M_a^{(n)} \cdot (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) + \\ &\quad \sum_{d \in \Omega_2} m_a^{(n)} \cdot (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) \end{aligned}$$

La demostración de la existencia del límite

Dado que:

$$\sum_{d \in \Omega_1} P^{n_0}[d, b_0] + \sum_{d \in \Omega_2} P^{n_0}[d, b_0] = 1 = \sum_{d \in \Omega_1} P^{n_0}[d, c_0] + \sum_{d \in \Omega_2} P^{n_0}[d, c_0]$$

Tenemos que:

$$\sum_{d \in \Omega_2} (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) = - \sum_{d \in \Omega_1} (P^{n_0}[d, b_0] - P^{n_0}[d, c_0])$$

La demostración de la existencia del límite

Dado que:

$$\sum_{d \in \Omega_1} P^{n_0}[d, b_0] + \sum_{d \in \Omega_2} P^{n_0}[d, b_0] = 1 = \sum_{d \in \Omega_1} P^{n_0}[d, c_0] + \sum_{d \in \Omega_2} P^{n_0}[d, c_0]$$

Tenemos que:

$$\sum_{d \in \Omega_2} (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) = - \sum_{d \in \Omega_1} (P^{n_0}[d, b_0] - P^{n_0}[d, c_0])$$

Concluimos entonces que:

$$\begin{aligned} \sum_{d \in \Omega_1} M_a^{(n)} \cdot (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) + \sum_{d \in \Omega_2} m_a^{(n)} \cdot (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) = \\ \sum_{d \in \Omega_1} M_a^{(n)} \cdot (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) - \sum_{d \in \Omega_1} m_a^{(n)} \cdot (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) \end{aligned}$$

La demostración de la existencia del límite

Así, dado que $\sum_{d \in \Omega_1} (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) \leq 1 - \ell \cdot t$, obtenemos:

$$\begin{aligned} P^{n_0+n}[a, b_0] - P^{n_0+n}[a, c_0] &\leq \sum_{d \in \Omega_1} M_a^{(n)} \cdot (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) + \\ &\quad \sum_{d \in \Omega_2} m_a^{(n)} \cdot (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) \\ &= \sum_{d \in \Omega_1} M_a^{(n)} \cdot (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) - \\ &\quad \sum_{d \in \Omega_1} m_a^{(n)} \cdot (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) \\ &= (M_a^{(n)} - m_a^{(n)}) \sum_{d \in \Omega_1} (P^{n_0}[d, b_0] - P^{n_0}[d, c_0]) \\ &\leq (1 - \ell \cdot t) \cdot (M_a^{(n)} - m_a^{(n)}) \end{aligned}$$

La demostración de la existencia del límite

En particular, escogemos b_0, c_0 tales que

$$\begin{aligned}M_a^{(n_0+n)} &= P^{n_0+n}[a, b_0] \\m_a^{(n_0+n)} &= P^{n_0+n}[a, c_0]\end{aligned}$$

La demostración de la existencia del límite

En particular, escogemos b_0, c_0 tales que

$$\begin{aligned}M_a^{(n_0+n)} &= P^{n_0+n}[a, b_0] \\m_a^{(n_0+n)} &= P^{n_0+n}[a, c_0]\end{aligned}$$

De esta forma, obtenemos:

$$M_a^{(n_0+n)} - m_a^{(n_0+n)} \leq (1 - \ell \cdot t) \cdot (M_a^{(n)} - m_a^{(n)})$$

La demostración de la existencia del límite

Para cada $k \geq 1$, tenemos que:

$$\begin{aligned} M_a^{(k \cdot n_0 + n)} - m_a^{(k \cdot n_0 + n)} &= M_a^{(n_0 + (k-1) \cdot n_0 + n)} - m_a^{(n_0 + (k-1) \cdot n_0 + n)} \\ &\leq (1 - \ell \cdot t) \cdot (M_a^{((k-1) \cdot n_0 + n)} - m_a^{((k-1) \cdot n_0 + n)}) \\ &= (1 - \ell \cdot t) \cdot (M_a^{(n_0 + (k-2) \cdot n_0 + n)} - m_a^{(n_0 + (k-2) \cdot n_0 + n)}) \\ &\leq (1 - \ell \cdot t)^2 \cdot (M_a^{((k-2) \cdot n_0 + n)} - m_a^{((k-2) \cdot n_0 + n)}) \\ &\leq \dots \\ &\leq (1 - \ell \cdot t)^k \cdot (M_a^{(n)} - m_a^{(n)}) \end{aligned}$$

Notése que la desigualdad también es válida para $k = 0$

La demostración de la existencia del límite

Considerando $n = 0$, obtenemos la sub-sucesión $(M_a^{(k \cdot n_0)} - m_a^{(k \cdot n_0)})_{k \in \mathbb{N}}$ de la sucesión $(M_a^{(n)} - m_a^{(n)})_{n \in \mathbb{N}}$

Dado que $0 \leq (1 - \ell \cdot t) < 1$ (ya que $t, \ell > 0$), para esta sub-sucesión se tiene que:

$$0 \leq \lim_{k \rightarrow \infty} (M_a^{(k \cdot n_0)} - m_a^{(k \cdot n_0)}) \leq \lim_{k \rightarrow \infty} (1 - \ell \cdot t)^k \cdot (M_c^{(n)} - m_c^{(n)}) = 0$$

Vale decir,

$$\lim_{k \rightarrow \infty} (M_a^{(k \cdot n_0)} - m_a^{(k \cdot n_0)}) = 0$$

La demostración de la existencia del límite

Para todo $n \in \mathbb{N}$ se tiene que:

- ▶ $m_a^{(n)} \leq m_a^{(n+1)}$
- ▶ $M_a^{(n)} \geq M_a^{(n+1)}$
- ▶ $m_a^{(n)} \leq M_a^{(n)}$

Por lo tanto, para todo $n \in \mathbb{N}$ se tiene que:

- ▶ $M_a^{(n)} - m_a^{(n)} \geq 0$
- ▶ $M_a^{(n+1)} - m_a^{(n+1)} \leq M_a^{(n)} - m_a^{(n)}$

La demostración de la existencia del límite

Dado que $(M_a^{(n)} - m_a^{(n)})_{n \in \mathbb{N}}$ es una sucesión de números reales monótona decreciente y acotada inferiormente, concluimos que existe $u \in \mathbb{R}$ tal que:

$$\lim_{n \rightarrow \infty} (M_a^{(n)} - m_a^{(n)}) = u$$

La demostración de la existencia del límite

Dado que $(M_a^{(n)} - m_a^{(n)})_{n \in \mathbb{N}}$ es una sucesión de números reales monótona decreciente y acotada inferiormente, concluimos que existe $u \in \mathbb{R}$ tal que:

$$\lim_{n \rightarrow \infty} (M_a^{(n)} - m_a^{(n)}) = u$$

Además, para la sub-sucesión $(M_a^{(k \cdot n_0)} - m_a^{(k \cdot n_0)})_{k \in \mathbb{N}}$ de $(M_a^{(n)} - m_a^{(n)})_{n \in \mathbb{N}}$ se cumple que:

$$\lim_{k \rightarrow \infty} (M_a^{(k \cdot n_0)} - m_a^{(k \cdot n_0)}) = 0$$

La demostración de la existencia del límite

Dado que $(M_a^{(n)} - m_a^{(n)})_{n \in \mathbb{N}}$ es una sucesión de números reales monótona decreciente y acotada inferiormente, concluimos que existe $u \in \mathbb{R}$ tal que:

$$\lim_{n \rightarrow \infty} (M_a^{(n)} - m_a^{(n)}) = u$$

Además, para la sub-sucesión $(M_a^{(k \cdot n_0)} - m_a^{(k \cdot n_0)})_{k \in \mathbb{N}}$ de $(M_a^{(n)} - m_a^{(n)})_{n \in \mathbb{N}}$ se cumple que:

$$\lim_{k \rightarrow \infty} (M_a^{(k \cdot n_0)} - m_a^{(k \cdot n_0)}) = 0$$

¡Concluimos que $u = 0$!

La demostración de la existencia del límite

Dado que $(M_a^{(n)} - m_a^{(n)})_{n \in \mathbb{N}}$ es una sucesión de números reales monótona decreciente y acotada inferiormente, concluimos que existe $u \in \mathbb{R}$ tal que:

$$\lim_{n \rightarrow \infty} (M_a^{(n)} - m_a^{(n)}) = u$$

Además, para la sub-sucesión $(M_a^{(k \cdot n_0)} - m_a^{(k \cdot n_0)})_{k \in \mathbb{N}}$ de $(M_a^{(n)} - m_a^{(n)})_{n \in \mathbb{N}}$ se cumple que:

$$\lim_{k \rightarrow \infty} (M_a^{(k \cdot n_0)} - m_a^{(k \cdot n_0)}) = 0$$

¡Concluimos que $u = 0$!

- ¿Cómo se demuestra que las dos sucesiones tienen el mismo límite? ¿Por qué es importante aquí que $n_0 \geq 1$? □