Cadenas de Markov

Sumario

- Introducción
- Procesos estocásticos
- Concepto de cadena de Markov
- Ecuaciones de Chapman-Kolmogorov
- Clasificación de estados
- Cadenas absorbentes
- Distribución estacionaria

Introducción.
Causas de la incertidumbre

Existen varias causas de incertidumbre que tienen que ver con la información, el conocimiento y la representación.

- Información incompleta, poco confinable, ruido, distorsión.
- Conocimientro impreciso o contradictorio.
- Representación no adecuada o con falta de poder descriptivo.

Introducción.
Manejo de Incertidumbre

- Para tratar la incertidumbre, hay que considerarla de forma explícita en la representación e inferencia

- Para ello se han desarrollado diversa formas de representar y manejar la incertidumbre

Introducción.
Técnicas numéricas

Probabilista
- Cadenas de Markov
- Procesos de decisión de Markov
- Redes bayesianas

Alternativas
- Empíricas (MYCIN, Prospector)
- Lógica difusa
- Teoría de Dempster-Shafer

Procesos estocásticos

Procesos estocásticos

- Un sistema informático complejo se caracteriza por demandas de carácter aleatorio y por ser dinámico. A veces interesa saber cómo cambia una variable aleatoria a lo largo del tiempo
- El estudio de cómo evoluciona una variable aleatoria a lo largo del tiempo incluye los procesos aleatorios o estocásticos
- Es necesario una herramienta que modele procesos aleatorios en el tiempo

Un proceso estocástico es una familia de variables aleatorias parametrizadas por el tiempo

Procesos estocásticos

- Es un concepto matemático que sirve para representar magnitudes aleatorias que varían con el tiempo o para caracterizar una sucesión de variables aleatorias (estocásticas) que evolucionan en función de otra variable, que generalmente es el tiempo.

- Cada una de las variables aleatorias del proceso tiene su propia función de distribución de probabilidad.

Procesos estocásticos

- Cada variable o conjunto de variables sometidas a influencias o efectos aleatorios constituye un proceso estocástico.

Matemáticamente, un proceso estocástico X_t es un conjunto a uniparamétrico de variables aleatorias indexadas mediante el tiempo t.

- Los procesos estocásticos permiten tratar procesos dinámicos en los que hay cierta aleatoriedad

Procesos estocásticos

Un proceso estocástico es un conjunto de variables aleatorias definida sobre un espacio de probabilidad

$$\left\{ X_t : \Omega \rightarrow \Re, \quad t \in T \right\}$$

$$\omega \rightarrow X_t(\omega) = X(\omega, t)$$

Procesos estocásticos

X es una función de dos argumentos. Fijado $\omega=\omega_0$, obtenemos una función determinista

$$X(\cdot,\omega_0):T \to \Re$$

$$t \to X(t,\omega_0)$$

Procesos estocásticos

Igualmente, fijado $t = t_0$, obtenemos una de las variables aleatorias de la familia

$$X(t_0,\cdot): \Omega \to \Re$$

$$\omega \to X(t_0,\omega)$$

Procesos estocásticos

El espacio de estados S de un proceso estocástico es el conjunto de todos los posibles valores que puede tomar dicho proceso

$$S = \left\{ X_t\left(\omega\right) \middle| \, t \in T \wedge \omega \in \Omega \right\}$$

Ejemplo de proceso estocástico

- Lanzamos una moneda al aire 6 veces. El jugador gana 1 € cada vez que sale cara, y pierde 1 € cada vez que sale cruz

Sean las variables aleatorias

X_i = cuentas del jugador después de la *i*-ésima jugada

La familia de variables aleatorias $\{X_1, X_2, \ldots, X_6\}$ constituye un proceso estocástico

Ejemplo de proceso estocástico

- Ω={CCCCCC,CCCCCF,…}
- card(Ω) = 2^6 = 64
- $P(\omega)=1/64 \quad \forall\ \omega \in \Omega$
- T={1, 2, 3, 4, 5, 6}
- S={–6, –5, …, –1, 0, 1, 2, …, 5, 6}
- $X_1(\Omega)=\{-1, 1\}$
- $X_2(\Omega)=\{-2, 0, 2\}$

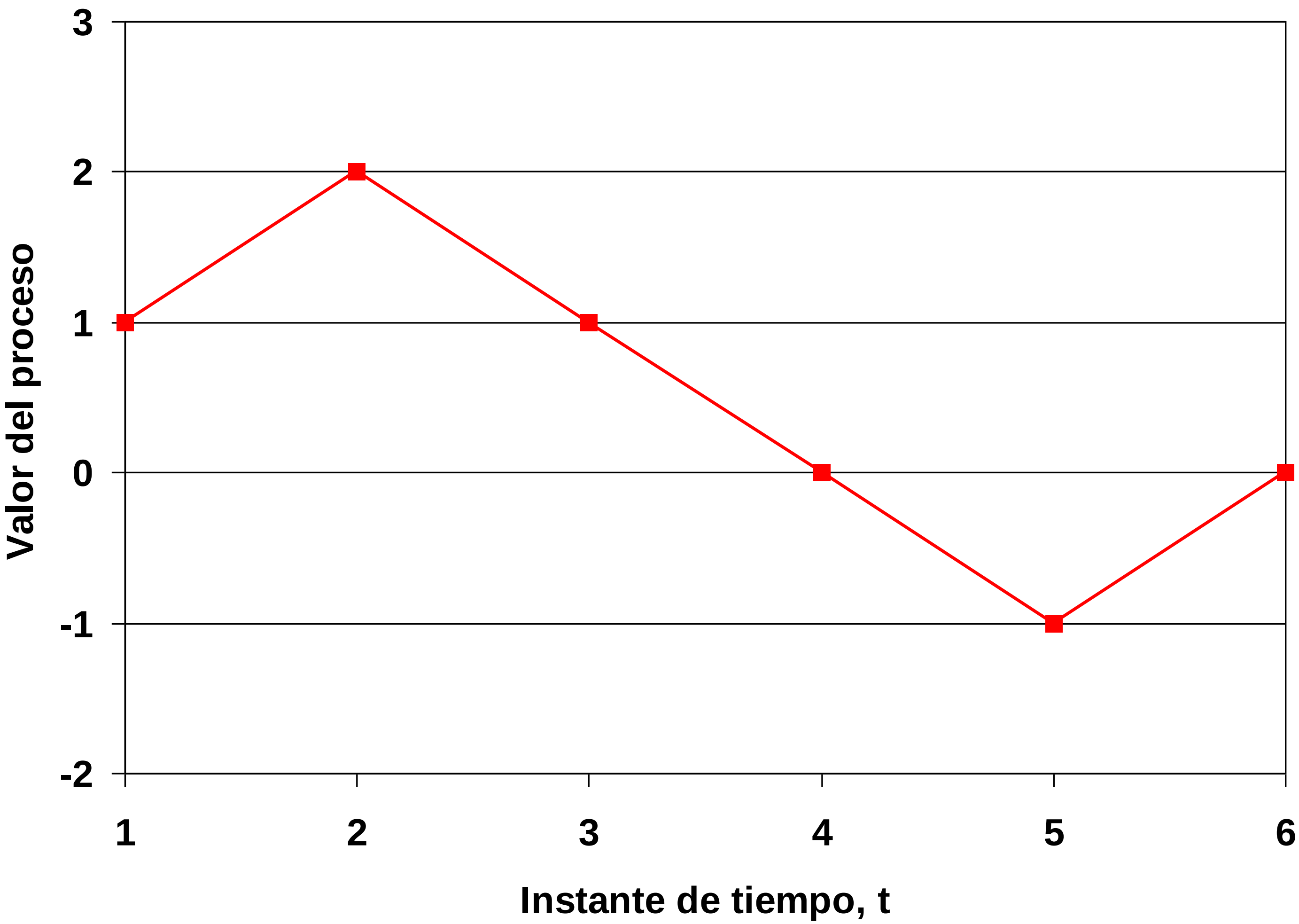
Ejemplo de proceso estocástico

Si se fija ω, por ejemplo ω_0=CCFFFC, se obtiene una secuencia de valores completamente determinista:

$X_1(\omega_0)=1$, $X_2(\omega_0)=2$, $X_3(\omega_0)=1$, $X_4(\omega_0)=0$, $X_5(\omega_0)= -1$, $X_6(\omega_0)=0$

Se puede graficar, con estos valores, la *trayectoria del proceso*

Ejemplo de proceso estocástico



Ejemplo de proceso estocástico

Si fijo t, por ejemplo, $t_0=3$, obtengo una de las variables aleatorias del proceso

$$X_3 : \Omega \to \Re$$

$$\omega \to X_3(\omega)$$

Los posibles valores que puede tomar el proceso en $t_0=3$ son: $X_3(\Omega)=\{-3, -1, 1, 3\}$

Ejemplo de proceso estocástico

Se puede hallar la probabilidad de que el proceso tome uno de estos valores:

$$P\left[X_3(\omega)=1\right]=P\left[CFC\right]+P\left[CCF\right]+P\left[FCC\right]=3\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{3}{8}$$

$$P\left[X_3(\omega)=3\right]=P\left[CCC\right]=\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{8}$$

$$P\left[X_3(\omega)=-1\right]=P\left[FCF\right]+P\left[FFC\right]+P\left[CFF\right]=3\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{3}{8}$$

$$P\left[X_3(\omega)=-3\right]=P\left[FFF\right]=\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{8}$$

Clasificación de los procesos estocásticos

	S discreto	**S continuo**
T discreto	Cadena	Sucesión de variables aleatorias continuas
T continuo	Proceso puntual	Proceso continuo

Ejemplos de los tipos de procesos estocásticos

- Cadena: Ejemplo anterior
- Sucesión de variables aleatorias continuas: cantidad de lluvia caída cada mes
- Proceso puntual: Número de clientes esperando en la cola de un supermercado
- Proceso continuo: velocidad del viento

Funciones asociadas a los procesos estocásticos

- Función de distribución de primer orden:

$$F : \Re^2 \to [0,1]$$

$$(x,t) \to F(x,t) = P[X(t,\omega) \leq x]$$

- Función de densidad de primer orden:

$$f(x,t) = \frac{\partial F(x,t)}{\partial x}$$

Funciones asociadas a los procesos estocásticos

- Función de distribución de 2º orden:

$$F : \Re^4 \rightarrow [0,1]$$

$$(x_1, x_2, t_1, t_2) \rightarrow F(x_1, x_2, t_1, t_2) = P[(X(t_1, \omega) \leq x_1) \wedge (X(t_2, \omega) \leq x_2)]$$

- Función de densidad de 2º orden:

$$f(x_1, x_2, t_1, t_2) = \frac{\partial^2 F(x_1, x_2, t_1, t_2)}{\partial x_1 \partial x_2}$$

Funciones asociadas a los procesos estocásticos

- Función valor esperado o valor medio (determinista)

$$m : T \to \Re$$
$$t \to E\left[X_t\right]$$

- Función varianza (es determinista)

$$\sigma^2 : T \to \Re$$
$$t \to \mathrm{var}\left(X_t\right)$$

Concepto de cadena de Markov

Cadenas de Markov

Las cadenas de Markov y los procesos de Markov son un tipo especial de procesos estocásticos que poseen la siguiente propiedad, denominada **Propiedad de Markov**

Conocido el estado del proceso en un momento dado, su comportamiento futuro no depende del pasado

Dicho de otro modo, “dado el presente, el futuro es independiente del pasado”

Cadenas de Markov

- Se van a estudiar las cadenas de Markov, es decir, procesos estocásticos con espacio de estados S e instantes de tiempo T discretos, T={t_0, t_1, t_2,…}

Matemáticamente, una cadena de Markov (CM) es una sucesión de variables aleatorias X_i, $i \in$ N, tal que

$$P\left[X_{t+1} = j \middle/ X_0, X_1, \ldots, X_t \right] = P\left[X_{t+1} = j \middle/ X_t \right]$$

es la expresión algebraica de la propiedad de Markov para T discreto.

Probabilidades de transición

- Las Cadenas de Markov están completamente caracterizadas por las probabilidades de transición en una etapa,

$$P\left[X_{t+1}=j \middle/ X_t = i\right], \quad i, j \in S, t \in T$$

Sólo trabajaremos con CM homogéneas en el tiempo, que son aquellas en las que

$$\forall i, j \in S \ \forall\, t \in T, P\left[X_{t+1}=j \middle/ X_t = i\right] = q_{ij}$$

donde $\mathbf{q_{ij}}$ se llama probabilidad de transición en una etapa desde el estado i hasta el estado j

Matriz de transición

Las probabilidades q_{ij} se agrupan en la denominada **matriz de transición de la cadena de Markov**

$$Q = \begin{pmatrix} q_{00} & q_{01} & q_{02} & \cdots \\ q_{10} & q_{11} & q_{12} & \cdots \\ q_{20} & q_{21} & q_{22} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} = \left(q_{ij}\right)_{i,\,j\in S}$$

Propiedades de la matriz de transición

Por ser los q_{ij} probabilidades,

$$\forall i, j \in S, \quad q_{ij} \in [0,1]$$

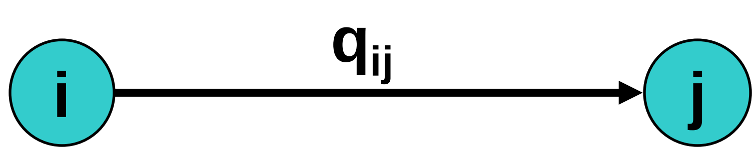
Por ser 1 la probabilidad del suceso seguro, cada fila ha de sumar 1, es decir,

$$\forall i \in S, \quad \sum_{j \in S} q_{ij} = 1$$

Una matriz que cumpla estas dos propiedades se denomina **matriz estocástica por filas**

Diagrama de transición de estados

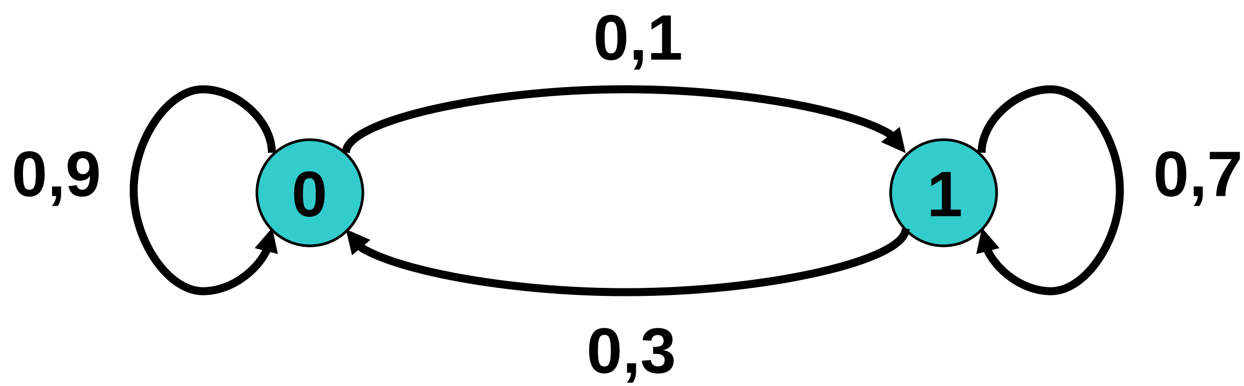
El diagrama de transición de estados (DTE) de una cadena de Markov es un grafo dirigido cuyos nodos son los estados de la cadena y cuyos arcos se etiquetan con la probabilidad de transición entre los estados que unen. Si dicha probabilidad es nula, no se pone arco.



Ejemplo: línea telefónica

Sea una línea telefónica con dos estados: ocupado = 1 y desocupado = 0. Si en el instante t está ocupada, en el instante t+1 estará ocupada con probabilidad 0,7 y desocupada con probabilidad 0,3. Si en el instante t está desocupada, en el t+1 estará ocupada con probabilidad 0,1 y desocupada con probabilidad 0,9.

$$Q = \begin{pmatrix} 0{,}9 & 0{,}1 \\ 0{,}3 & 0{,}7 \end{pmatrix}$$

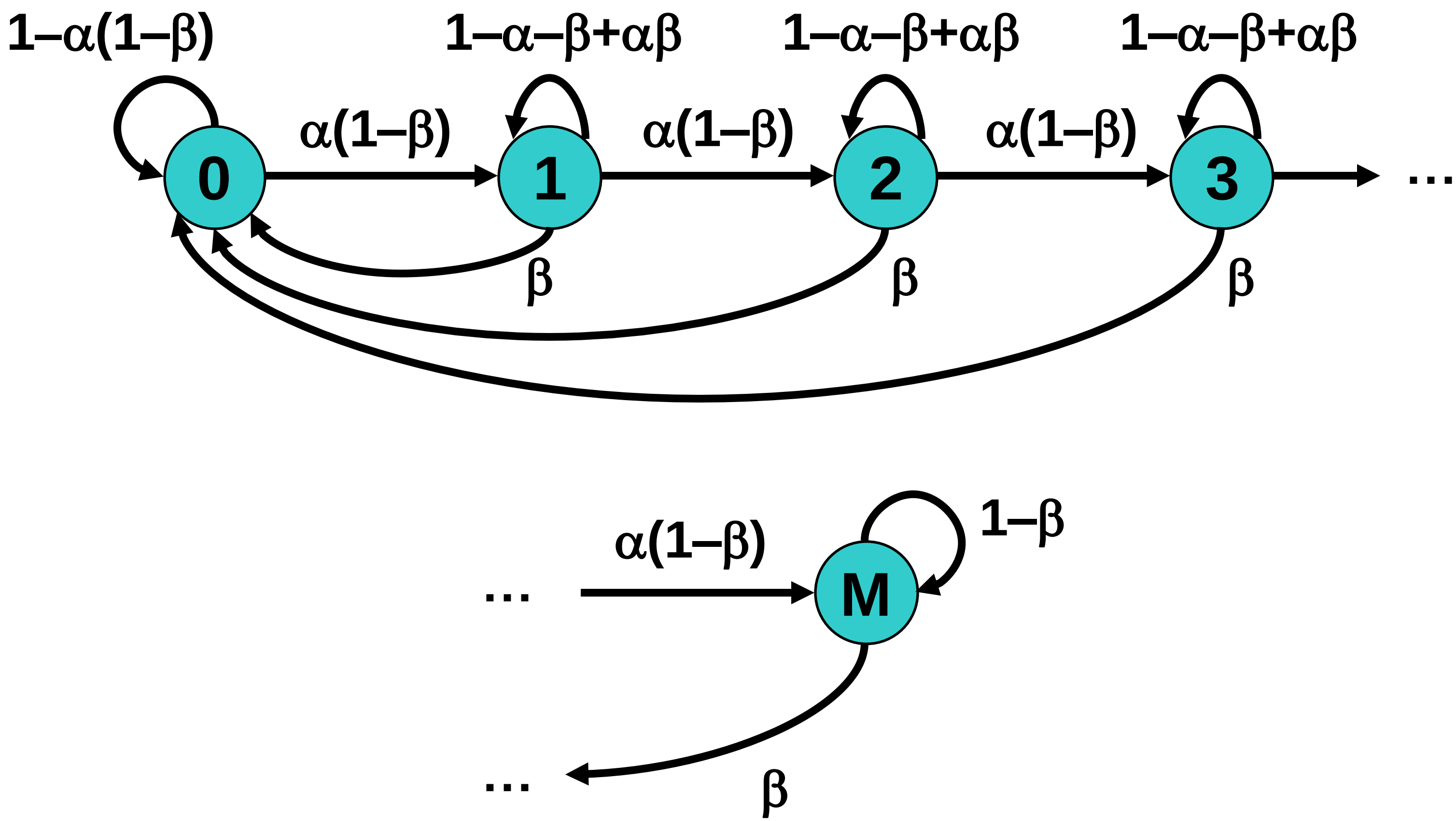


Ejemplo: *buffer* de E/S

- Supongamos que un *buffer* de E/S tiene espacio para M paquetes. En cualquier instante de tiempo se puede insertar un paquete en el *buffer* con probabilidad α o bien el *buffer* puede vaciarse con probabilidad β. Si ambos casos se dan en el mismo instante, primero se inserta y luego se vacía.

Sea X_t = nº de paquetes en el *buffer* en el instante t. Suponiendo que las inserciones y vaciados son independientes entre sí e independientes de la historia pasada, $\{ X_t \}$ es una cadena de Markov, donde S={0, 1, 2, …, M}

Ejemplo: *buffer* de E/S

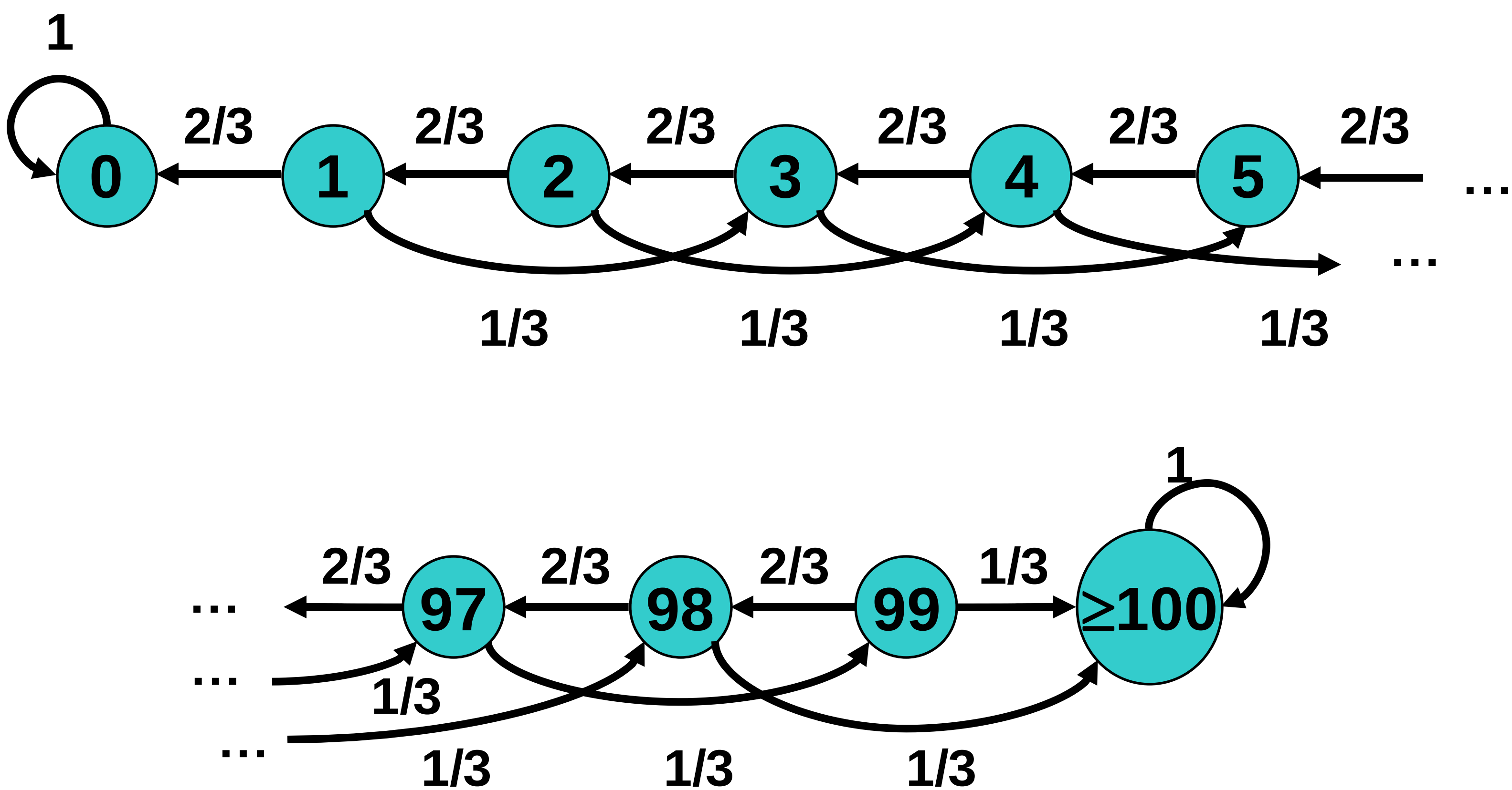


Ejemplo: Lanzamiento de un dado

- Se lanza un dado repetidas veces. Cada vez que sale menor que 5 se pierde 1€, y cada vez que sale 5 ó 6 se gana 2 €. El juego acaba cuando se tienen 0 ó 100€.

Si X_t = estado de cuentas en el instante t entonces, $\{ X_t \}$ es una cadena de Markov con S={0, 1, 2, …, 100}

Ejemplo: Lanzamiento de un dado



Ejemplo: organismos unicelulares

- Se tiene una población de organismos unicelulares que evoluciona así: cada organismo se duplica con probabilidad 1–p o muere con probabilidad p.

Sea X_n el nº de organismos en el instante n. La cadena de Markov $\{ X_n \}$ tendrá S = { 0, 1, 2, 3, … } = **N**

Si hay i organismos en el instante n, en el instante n+1 tendremos k organismos que se dupliquen e i–k que mueran, con lo que habrá 2k organismos.

Ejemplo: organismos unicelulares

Mediante la distribución binomial podemos hallar las probabilidades de transición $q_{i,2k}$, ya que el resto de probabilidades son nulas

$$\forall k \in \{0,1,2,\ldots,i\}, \quad q_{i,2k} = \binom{i}{k} (1-p)^k \, p^{i-k}$$

Ecuaciones de Chapman-Kolmogorov

Ecuaciones de Chapman-Kolmogorov

Teorema: Las probabilidades de transición en n etapas vienen dadas por la potencia n de la matriz de transición de estados Q^n

$$\forall i, j \in S, P\left[X_{t+n} = j \big/ X_t = i \right] = q_{ij}^{(n)}$$

Demostración: Por inducción sobre n

Caso base (n=1). Se sigue de la definición de q_{ij}

Ecuaciones de Chapman-Kolmogorov

Hipótesis de inducción. Suponemos cierta la conclusión del teorema para un cierto n.

Paso inductivo, se demuestra para (n+1), para cualesquier i, j $\in$ S

$$P\left[X_{t+n+1}=j \middle/ X_t=i\right]=\sum_{k\in S}P\left[\left(X_{t+n}=k\right)\wedge\left(X_{t+n+1}=j\right) \middle/ X_t=i\right]=$$

$$=\sum_{k\in S}P\left[X_{t+n}=k \middle/ X_t=i\right]P\left[X_{t+n+1}=j \middle/ X_{t+n}=k\right]=\{H.I.\}=$$

$$=\sum_{k\in S}q_{ik}^{(n)}P\left[X_{t+n+1}=j \middle/ X_{t+n}=k\right]=\sum_{k\in S}q_{ik}^{(n)}q_{kj}=q_{ij}^{(n+1)}$$

Ecuaciones de Chapman-Kolmogorov

Por este teorema sabemos que **la probabilidad de transitar desde el estado i hasta el estado j en n pasos es el elemento (i, j) de la matriz** Q^n

Para evitar calculos de potencias elevadas de matrices, se intenta averiguar el comportamiento del sistema en el límite cuando $n \rightarrow \infty$, llamado también comportamiento a largo plazo

Distribución de probabilidad

Sea $p_i(n) = P(X_n = i)$

$$\mathbf{Q = P(n) = [p_0(n) \quad p_1(n) \quad p_2(n) \quad \ldots]}$$

Sea p(0) la distribución de probabilidad inicial con $p_0(n) = P(X_0 = i)$

$$p(0) = [p_0(0) \quad p_1(0) \quad p_2(0) \quad \ldots]$$

La distribución de probabilidad después de la primera transición es

$$p(1) = p(0)Q$$

La distribución de probabilidad después de dos transiciones es

$$p(2) = p(1)Q = (p(0)Q)Q = p(0)Q^2$$

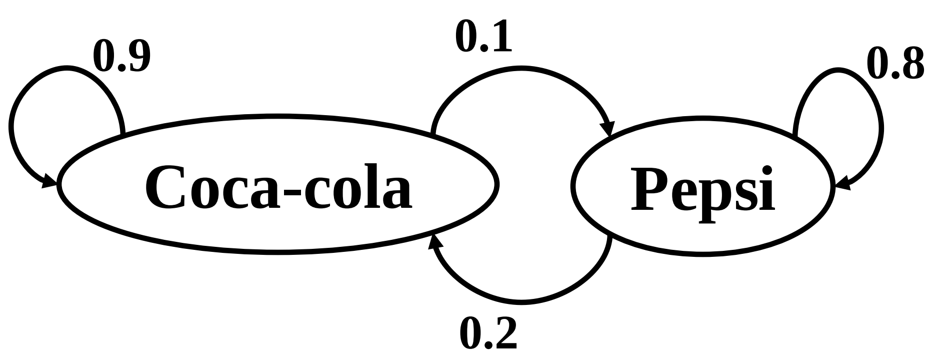
La distribución de probabilidad después de n transiciones es

$$p(n) = p(0)Q^n$$

Ejemplo

Dado que la última compra de cola de una persona fue Coca-Cola, hay un 90% de posibilidades de que su próxima compra de cola también sea Coca-Cola. Por contra, si la última compra de cola de una persona fue Pepsi, hay un 80% de probabilidades de que su próxima compra de cola también sea Pepsi.

La matriz de transición es $Q = \begin{bmatrix} 0{,}9 & 0{,}1 \\ 0{,}2 & 0{,}8 \end{bmatrix}$



Dado que una persona compra actualmente Pepsi, ¿cuál es la probabilidad de que compre Coca-Cola dentro de dos compras?

Ejemplo

P[Pepsi→?→CC] = P[Pepsi→CC→CC] + P[Pepsi→ Pepsi →CC] =

= 0.2 * 0.9 + 0.8 * 0.2 = 0.34

$$Q^2 = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}$$

Ejemplo

Dado que una persona compra actualmente Coca-Cola, ¿cuál es la probabilidad de que compre Pepsi dentro de tres compras?

$$Q^3 = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix} = \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix}$$

Ejemplo

Supongamos que cada persona hace una compra de cola a la semana. Ademá, supongamos que el 60% de la gente bebe Coca-Cola y el 40% Pepsi.

¿Qué porcentaje de personas beberá Coca-Cola dentro de tres semanas?

$$Q = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \qquad Q^3 = \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix}$$

$P[X_3 = CC] = 0.6 * 0.781 + 0.4 * 0.438 = 0.6438$

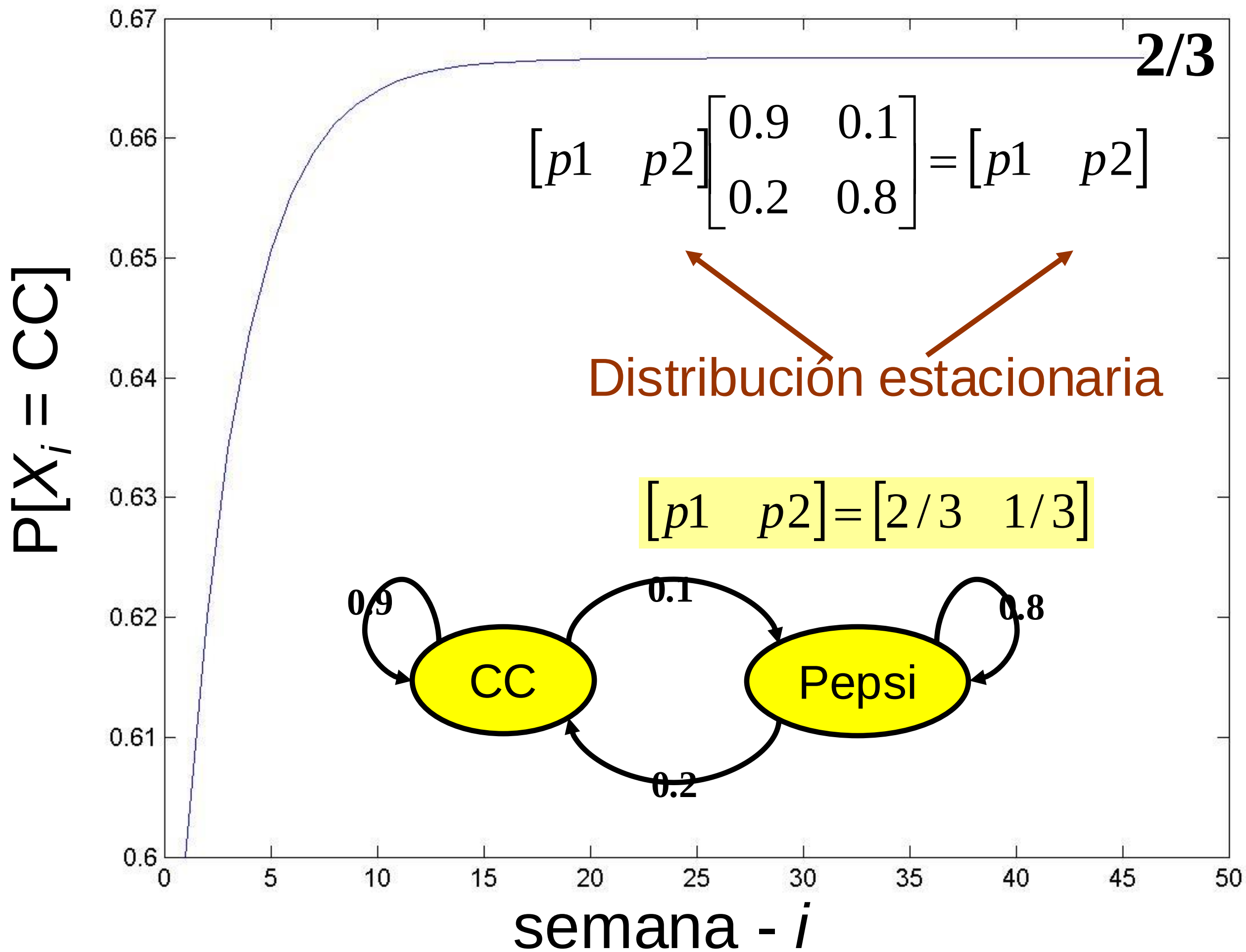
p_i = distribución en la semana i

$p_0 = (0.6, 0.4)$ – distribución inicial

$p_3 = p_0 * Q^3 = (0.6438, 0.3562)$

Ejemplo

Distribución estacionaria



Ejemplo

Para beber agua un animal puede ir a un lago o a un río. Se sabe que no va al lago dos días seguidos y que si toma agua en el río la probabilidad de que el día siguiente beba agua en cada uno de los sitios es la misma. ¿Calcula la matriz de transición?

$$\mathrm{Q} = \begin{bmatrix} 0 & 1/2 \\ 1 & 1/2 \end{bmatrix}$$

Cálculo de la potencia de una matriz

Un escalar λ se llama valor propio de A si existe una solución no trivial $\bar{x} = \bar{0}$ de la ecuación $A\bar{x} = \lambda\bar{x}$

Una de esas soluciones no triviales $\bar{x}$ se denomina vector propio de A asociado al valor propio λ.

Cálculo de la potencia de una matriz

$$\mathrm{A}\bar{x} = \lambda\bar{x} \qquad \Longrightarrow \qquad (\mathrm{A} - \lambda I)\bar{x} = \bar{0}$$

Para que λ sea valor propio de la matriz cuadrada A, el sistema homogéneo ha de tener soluciones no triviales, luego el determinante de la matriz cuadrada (A– λ I) ha de ser cero.

Ecuación característica de A

$$\mathrm{Det}\,(\mathrm{A} - \lambda I) = 0$$

Polinomio característico de A

$$p(\lambda) = \mathrm{Det}\,(\mathrm{A} - \lambda I) = 0$$

Los valores propios de una matriz cuadrada son las raíces de su polinomio característico

El **orden** del valor propio λ es la **multiplicidad** k de λ como raíz del polinomio característico.

Cálculo de la potencia de una matriz

Una matriz cuadrada A se dice diagonalizable si existe una matriz regular P que cumple que:

$$\mathbf{A} = \mathbf{P} \cdot \mathbf{D} \cdot \mathbf{P}^{-1} \quad , \quad \mathbf{D} \ \text{diagonal}$$

Esto nos va a permitir calcular rápidamente A^k para valores grandes de k.

¿Cómo se diagonaliza una matriz cuadrada A diagonalizable?

1.- Calcular los valores propios de A indicando sus órdenes.

$$\underbrace{\lambda_1 \cdots \lambda_1}_{k_1} \underbrace{\lambda_2 \cdots \lambda_2}_{k_2} \cdots \underbrace{\lambda_r \cdots \lambda_r}_{k_r}$$

2.- Calcular los subespacios propios $V(\lambda_i)$ y bases B_i de cada subespacio.

$$\mathbf{B_1 \cup B_2 \cup \cdots \cup B_r = B}$$

3.- Escribir las matrices D y P tales que: $\mathbf{D = P^{-1} \cdot A \cdot P}$

<u>columnas de P</u>: vectores de la base B formada por los vectores propios de A encontrada en 2.

En orden

<u>elementos de la diagonal principal de D</u>: valores propios de A calculados en 1.

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & -2 \end{pmatrix}$$

$$\begin{cases} \lambda_1 = 1 & , \quad k_1 = 2 \\ \lambda_2 = -2 & , \quad k_2 = 1 \quad (\text{v.p. simple}) \end{cases}$$

$$\mathbf{V}(\lambda_1) = \mathbf{V}(1) = \{(a, b, 0)/a, b \in \mathbb{R}\} \qquad \mathbf{B}_1 = \{\overline{u}_1 = (1, 0, 0), \overline{u}_2 = (0, 1, 0)\}$$

$$\mathbf{V}(\lambda_2) = \mathbf{V}(-2) = \{(a, -2a, 3a)/a \in \mathbb{R}\} \qquad \mathbf{B}_2 = \{\overline{u}_3 = (1, -2, 3)\}$$

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} = \mathbf{P}^{-1} \cdot \mathbf{A} \cdot \mathbf{P} \quad ; \quad \mathbf{P} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 3 \end{pmatrix}$$

Como ya sabemos, el cálculo de las potencias A^k puede ser bastante tedioso. Sin embargo, si A es diagonalizable y hemos calculado P y D, entonces sabemos que

$$\mathbf{A} = \mathbf{P} \cdot \mathbf{D} \cdot \mathbf{P}^{-1}$$

$$\mathbf{A}^2 = \left(\mathbf{P} \cdot \mathbf{D} \cdot \mathbf{P}^{-1}\right) \cdot \left(\mathbf{P} \cdot \mathbf{D} \cdot \mathbf{P}^{-1}\right) = \mathbf{P} \cdot \mathbf{D}^2 \cdot \mathbf{P}^{-1}$$

Con lo cual, iterando el proceso llegamos a:

$$\boxed{\mathbf{A}^k = \mathbf{P} \cdot \mathbf{D}^k \cdot \mathbf{P}^{-1}}$$

Como el cálculo de D^k equivale a elevar sólo los elementos diagonales de D a la k-ésima potencia, vemos que A^k es fácil de obtener.

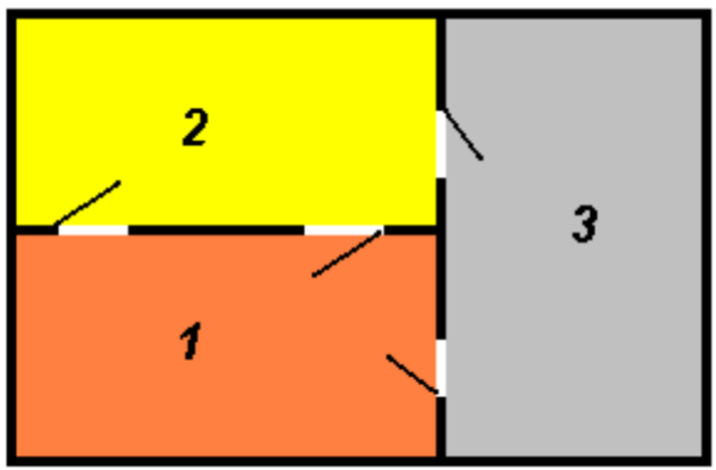
Ejemplo

Dada una cadena de Markov con matriz de transición A, calcula la tendencia a largo plazo para una distribución inicial $p_0 = (0, 0, 0, 1, 0, 0)$.

$$A = \begin{pmatrix} 1 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1 \end{pmatrix}.$$

Ejemplo

Supongamos que en un laboratorio se coloca un conjunto de ratones en una caja dividida en tres compartimentos comunicados y todos con la misma facilidad de acceso, tal y como se indica en la figura. Los compartimentos permanecen cerrados y se abren cada lunes. Sabiendo que semana tras semana todos los ratones cambian de ubicación y que los ratones cuando salen eligen un compartimento al azar, calcula la distribución de los ratones al cabo de "infinitas" semanas.



Ejemplo

Los trabajadores de un parque natural se clasifican en 3 categorías profesionales: científicos X1, personal auxiliar X2 y colaboradores X3. En cada generación t, representaremos a la fuerza de trabajo del parque por el número de personas incluidas en las tres categorías

anteriores, es decir X(t) = (X1(t), X2(t), X3(t)). Supongamos que cada trabajador activo sólo tiene un hijo que sigue trabajando en el parque. Además, la mitad de los hijos de los científicos lo son también, la cuarta parte pasa a ser personal auxiliar especializado y el resto es personal colaborador no especializado. Por otro lado, los hijos del personal auxiliar se reparten entre las 3 categorías según los porcentajes 30%, 40%, 30%. Para los hijos de los colaboradores las proporciones de reparto entre las categorías son 50%, 25% y 25%.mento al azar. Calcula la distribución al cabo de “infinitas” Semanas suponiendo X(0) = (1, 1, 1)

Clasificación de estados

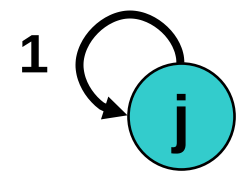
Clasificación de estados

Probabilidad de alcanzar un estado:

$$\forall i, j \in S, \quad \nu_{ij} = P\left[X_n = j \text{ para algún } n > 0 \Big/ X_0 = i \right]$$

Diremos que un estado $j \in S$ es **alcanzable** desde el estado $i \in S$ sii $\nu_{ij} \neq 0$. Esto significa que existe una sucesión de arcos (camino) que van desde i hasta j.

Un estado $j \in S$ es **absorbente** sii $q_{jj}=1$



Subconjuntos cerrados

- Sea $C \subseteq S$, con $C \neq \emptyset$. Diremos que C es **cerrado** sii $\forall i \in C\ \forall j \notin C$, j no es alcanzable desde i, o lo que es lo mismo, $\nu_{ij}=0$. En particular, si $C=\{i\}$, entonces i es absorbente. S siempre es cerrado.

- Un subconjunto cerrado $C \subseteq S$ se dice que es **irreducible** sii no contiene ningún subconjunto propio cerrado

Estados recurrentes y transitorios

- Si S es irreducible, se dice que la cadena de Markov es **irreducible**. Esto ocurre sii dados i,j cualesquiera, j es alcanzable desde i

- Diremos que un estado $j \in S$ es **recurrente** sii $\nu_{jj}=1$. En otro caso diremos que j es **transitorio**. Se demuestra que una cadena de Markov sólo puede pasar por un estado transitorio como máximo una cantidad finita de veces. En cambio, si visitamos un estado recurrente, entonces lo visitaremos infinitas veces.

Estados recurrentes y transitorios

- **Proposición:** Sea $C \subseteq S$ cerrado, irreducible y finito. Entonces $\forall i \in C$, i es recurrente

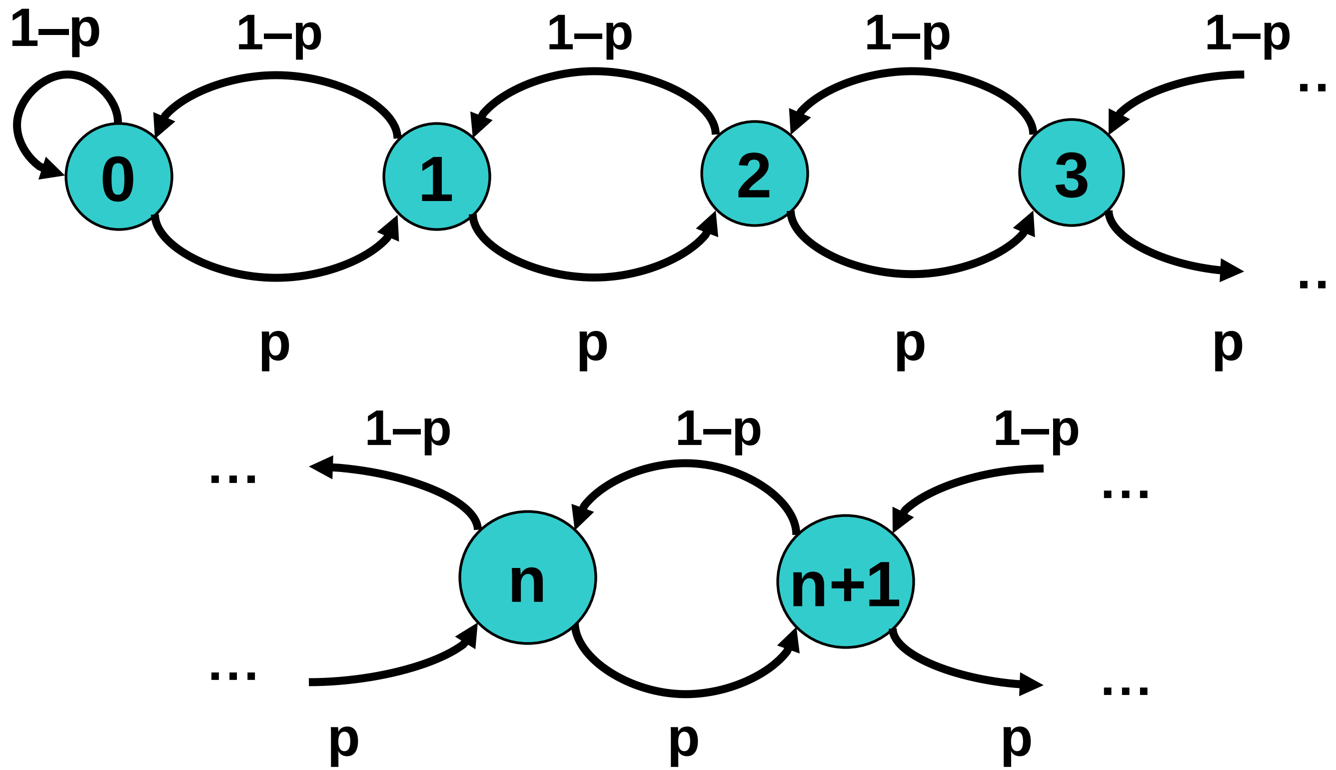
- **Ejemplos:** La CM de la línea telefónica es irreducible. Como además es finita, todos los estados serán recurrentes. Lo mismo ocurre con el ejemplo del *buffer*

- **Ejemplo:** En el lanzamiento del dado, tenemos los subconjuntos cerrados $\{0\}$, $\{\geq 100\}$, con lo que la CM no es irreducible. Los estados 0 y ≥ 100 son absorbentes, y el resto son transitorios

Estados recurrentes y transitorios

Proposición: Sea *X* una CM irreducible. Entonces, o bien todos sus estados son recurrentes (y decimos que *X* es recurrente), o bien todos sus estados son transitorios (y decimos que *X* es transitoria).

Ejemplo: Estado de cuentas con una persona rica. Probabilidad p de ganar 1 € y 1–p de perder 1€. Cuando otra persona se arruina, el rico le presta dinero para la siguiente tirada:



Cadenas recurrentes y transitorias

- Esta cadena es irreducible e infinita. Se demuestra que es transitoria sii $p>0,5$ y recurrente en otro caso ($p\leq 0,5$)

- La cadena es transitoria cuando la “tendencia global” es ir ganando dinero. Esto implica que una vez visitado un estado, al final dejaremos de visitarlo porque tendremos más dinero.

Periodicidad

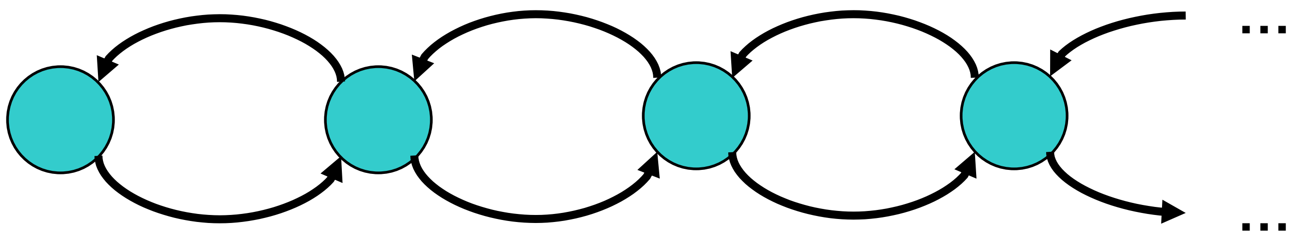
- Sea j$\in$S tal que ν_{jj}>0. Sea

$$k = mcd\ \{n \in \mathbf{N} - \{0\} \mid q_{jj}^{(n)} > 0\}$$

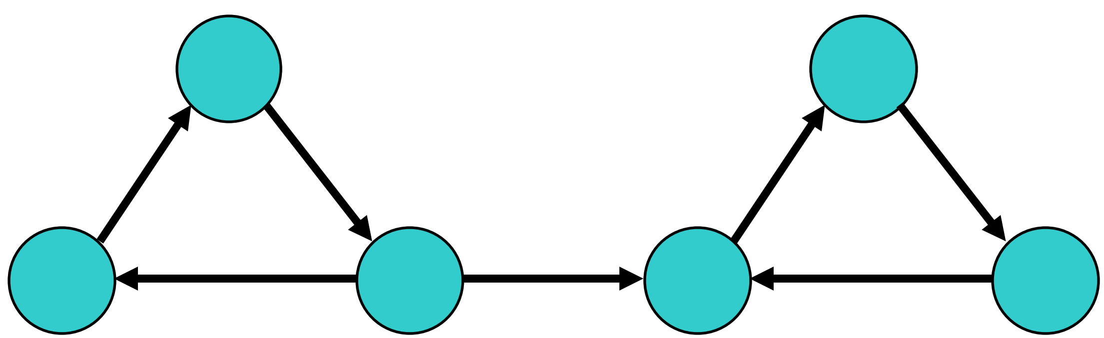
Si k>1, entonces diremos que j es periódico de periodo k. El estado j será periódico de periodo k>1 sii existen caminos que llevan desde j hasta j pero todos tienen longitud mk, con m>0

Periodicidad

Ejemplo: En la siguiente CM todos los estados son periódicos de periodo k=2:



Ejemplo: En la siguiente CM todos los estados son periódicos de periodo k=3:

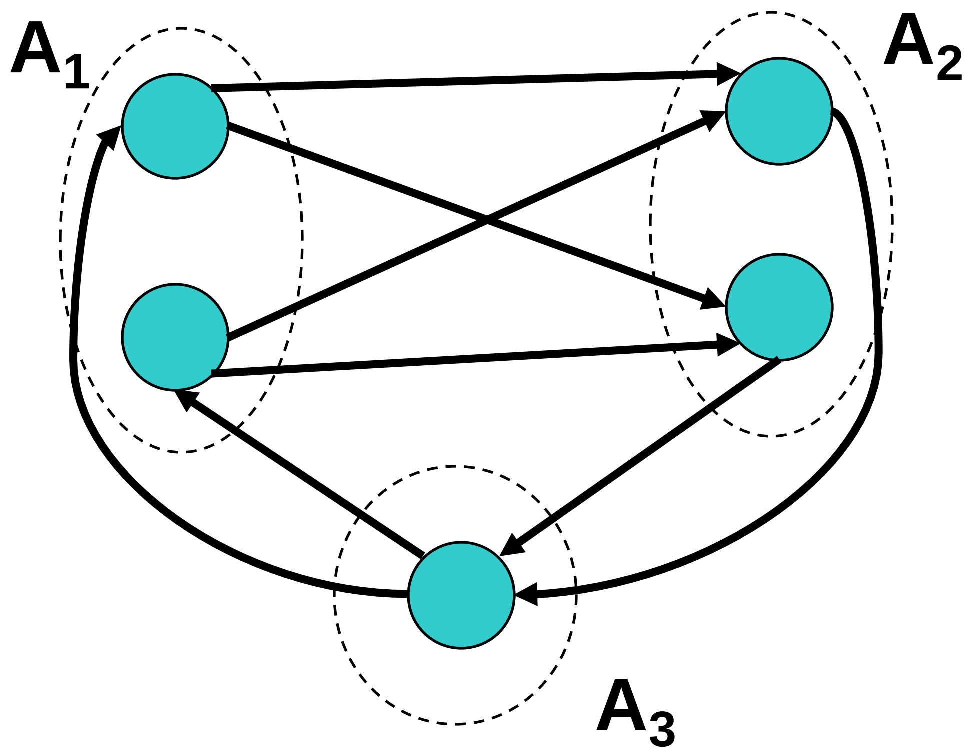


Periodicidad

Proposición: Sea X una CM irreducible. Entonces, o bien todos los estados son periódicos de periodo k (y decimos que X es periódica de periodo k), o bien ningún estado es periódico (y decimos que X es aperiódica)

Periodicidad

Ejemplo de cadena de Markov periódica de periodo k=3



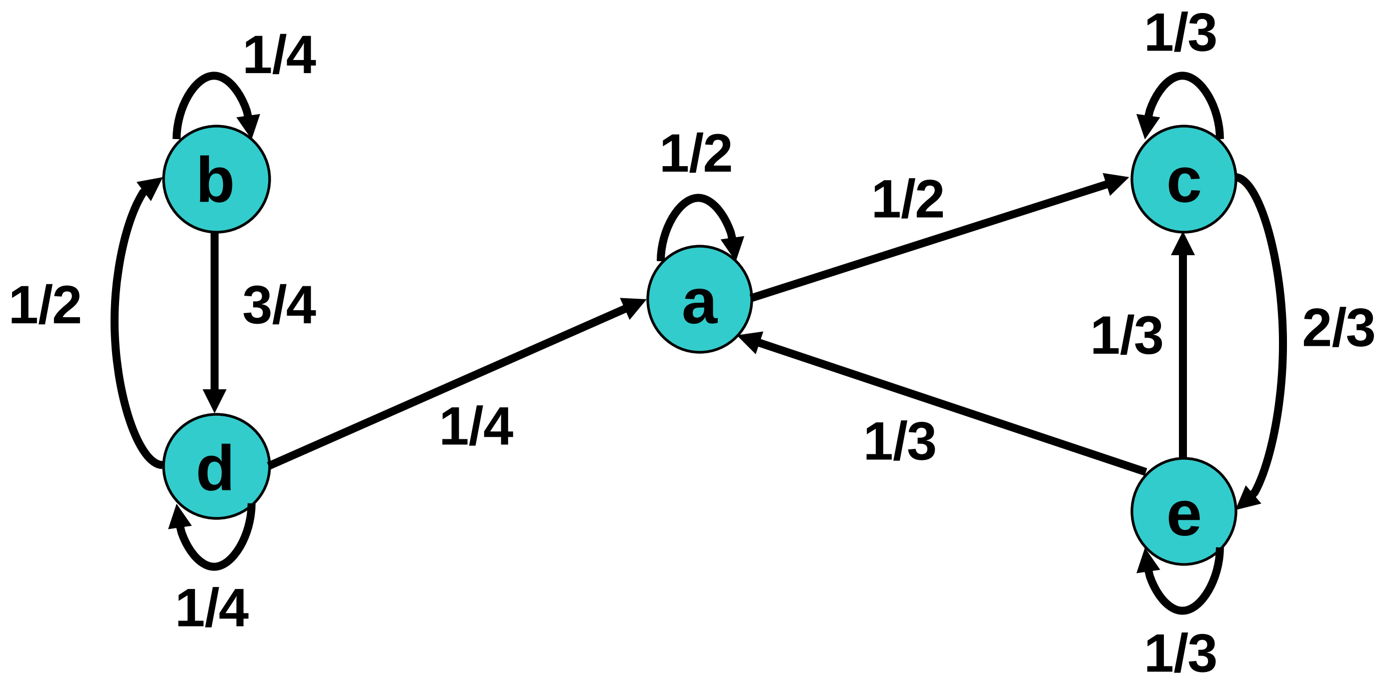
Cadenas ergódicas

Sea X una CM finita. Diremos que X es ergódica sii es irreducible, recurrente y no periódica

Ejemplo: Analizar la siguiente CM, con S={a, b, c, d, e}

$$Q = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \end{pmatrix}$$

Ejemplo



Ejemplo

Hallar los conjuntos cerrados

Tomado un estado i, construimos un conjunto cerrado C_i con todos los alcanzables desde él en una o más etapas (el propio i también se pone):

$C_a=\{a, c, e\}=C_c=C_e$

$C_b=\{b, d, a, c, e\}=C_d=S$

La cadena de Markov no será irreducible, ya que C_a es un subconjunto propio cerrado de S

Ejemplos

Clasificamos los estados

- Recurrentes: a, c, e
- Transitorios: b, d
- Periódicos: ninguno
- Absorbentes: ninguno

Reorganizamos Q. Dada una CM finita, siempre podemos agrupar los estados recurrentes por un lado y los transitorios por otro, y hacer:

$$Q = \left(\begin{array}{c|c} \begin{array}{c} \text{Movimientos entre} \\ \text{recurrentes} \end{array} & 0 \\ \hline \begin{array}{c} \text{Paso de transitorios} \\ \text{a recurrentes} \end{array} & \begin{array}{c} \text{Movimientos entre} \\ \text{transitorios} \end{array} \end{array} \right)$$

Ejemplos

En nuestro caso, la nueva ordenación de S es S={a, c, e, b, d}, con lo que obtenemos:

$$Q = \left(\begin{array}{ccc|cc} 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/3 & 2/3 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \\ \hline 0 & 0 & 0 & 1/4 & 3/4 \\ 1/4 & 0 & 0 & 1/2 & 1/4 \end{array}\right)$$

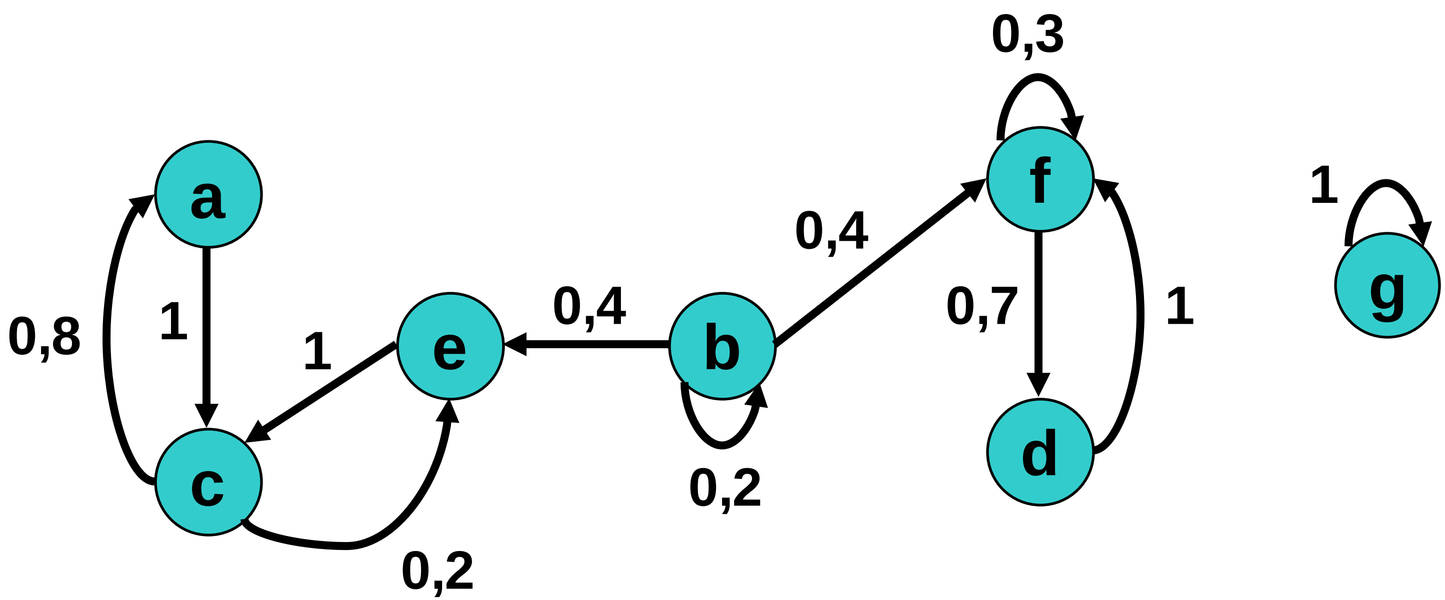
Clasificamos la cadena. No es irreducible, con lo cual no será periódica, ni aperiódica, ni recurrente, ni transitoria ni ergódica.

Ejemplos

Ejemplo: Analizar la siguiente CM, con S={a, b, c, d, e, f, g}:

$$Q = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0{,}2 & 0 & 0 & 0{,}4 & 0{,}4 & 0 \\ 0{,}8 & 0 & 0 & 0 & 0{,}2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0{,}7 & 0 & 0{,}3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Ejemplos



Ejemplos

Los conjuntos cerrados son

- $C_a=\{a, c, e\}=C_c=C_e$
- $C_f=\{f, d\}=C_d$
- $C_g=\{g\}$
- S

Clasificamos los estados

- Recurrentes: a, c, d, e, f, g
- Transitorios: b
- Periódicos: a, c, e (todos de periodo 2)
- Absorbentes: g

Ejemplos

Reorganizamos Q. Cuando hay varios conjuntos cerrados e irreducibles de estados recurrentes (por ejemplo, n conjuntos), ponemos juntos los estados del mismo conjunto:

$$Q = \left(\begin{array}{c:c:c:c:c:c}
P_1 & 0 & 0 & \dots & 0 & 0 \\
\hdashline
0 & P_2 & 0 & \dots & 0 & 0 \\
\hdashline
0 & 0 & P_3 & \dots & 0 & 0 \\
\hdashline
\dots & \dots & \dots & \dots & \dots & \dots \\
\hdashline
0 & 0 & 0 & \dots & P_n & 0 \\
\hdashline
Z_1 & Z_2 & Z_3 & & Z_n & Z
\end{array}\right)$$

Ejemplos

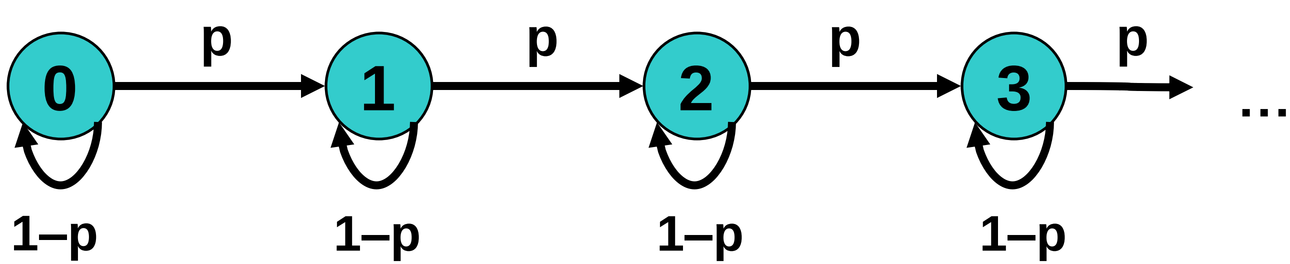
En nuestro caso, reordenamos S={a, c, e, d, f, g, b} y obtenemos:

$$Q = \left(\begin{array}{ccc:cc:c:c} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0,8 & 0 & 0,2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0,7 & 0,3 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \hdashline 0 & 0 & 0,4 & 0 & 0,4 & 0 & 0,2 \end{array} \right)$$

Ejemplo

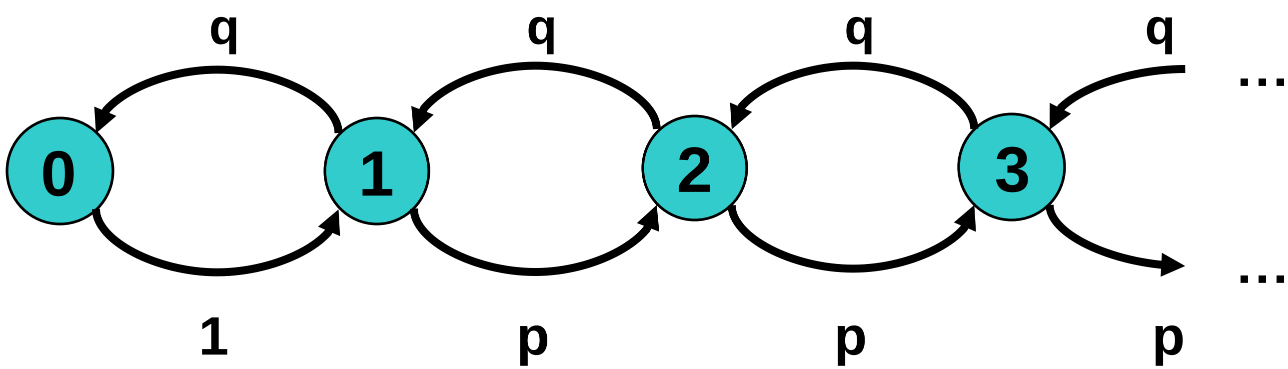
Clasificar la cadena. No es irreducible, con lo cual no será periódica, ni aperiódica, ni recurrente, ni transitoria ni ergódica.

Ejemplo: Número de éxitos al repetir indefinidamente una prueba de Bernouilli (probabilidad p de éxito). No es una cadena de Markov irreducible, porque por ejemplo $C_1=\{1, 2, 3, \ldots\}$ es cerrado. Todos los estados son transitorios.



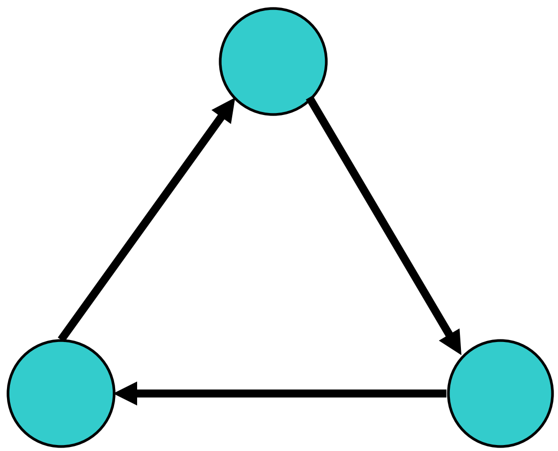
Ejemplo

- *Recorrido aleatorio*. Es una CM irreducible y periódica de periodo 2. Se demuestra que si $p \leq q$, todos los estados son recurrentes, y que si $p > q$, todos son transitorios.



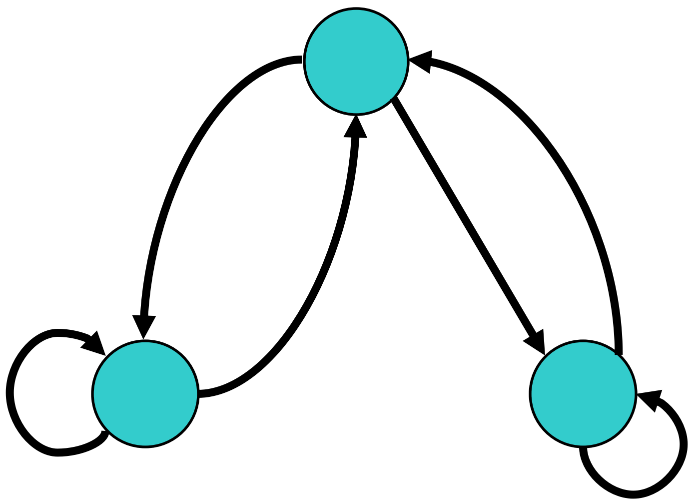
Ejemplo

La siguiente CM es irreducible, recurrente y periódica de periodo 3.
No es ergódica.



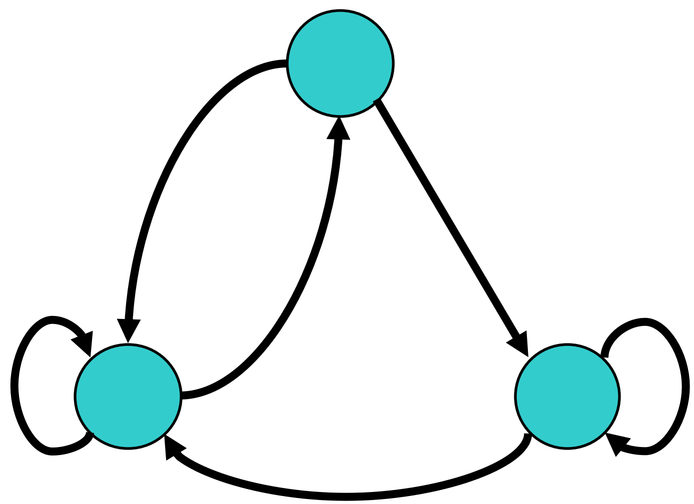
Ejemplo

La siguiente CM es irreducible, aperiódica, recurrente y ergódica.



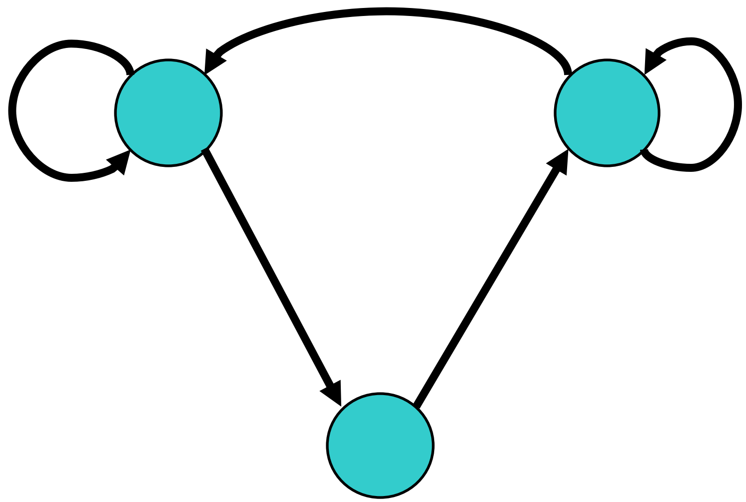
Ejemplo

La siguiente CM es irreducible, aperiódica, recurrente y ergódica



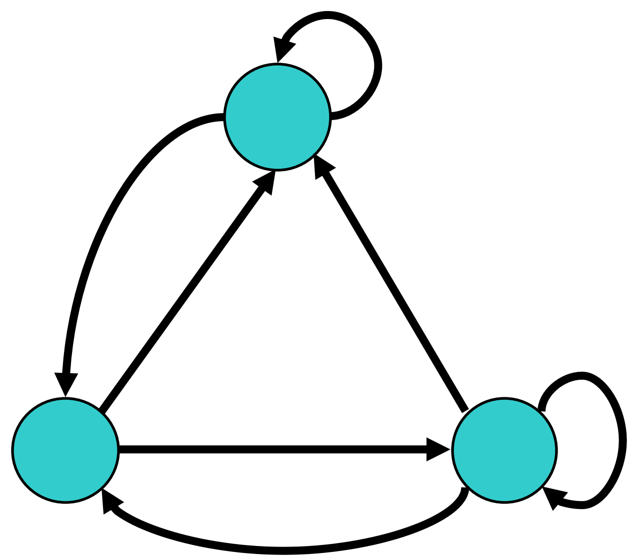
Ejemplo

La siguiente CM es irreducible, aperiódica, recurrente y ergódica



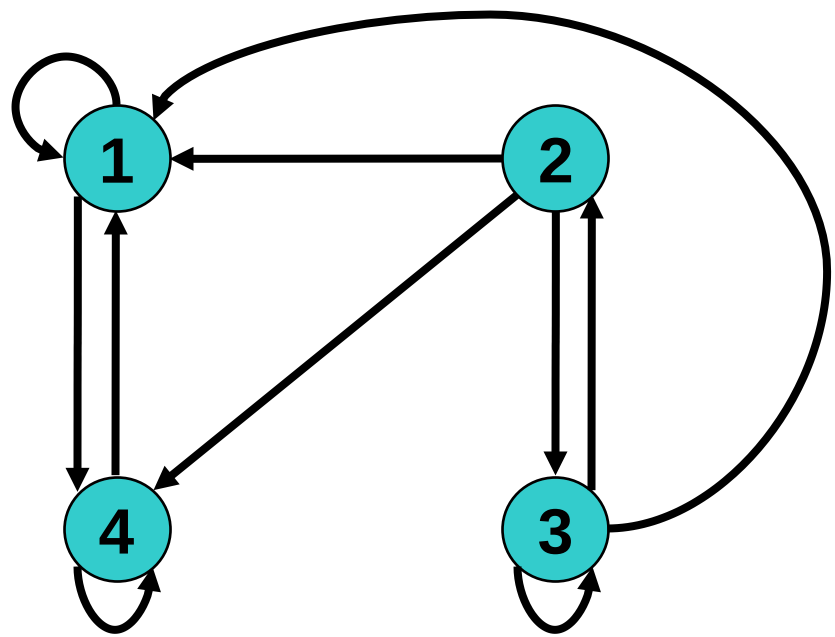
Ejemplo

La siguiente CM es irreducible, aperiódica, recurrente y ergódica



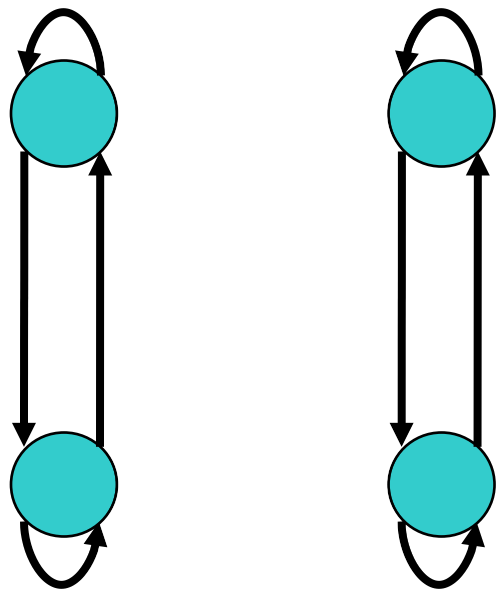
Ejemplo

La siguiente CM no es irreducible, y por tanto no es de ninguno de los demás tipos. Los estados 1 y 4 son recurrentes; 2 y 3 son transitorios.



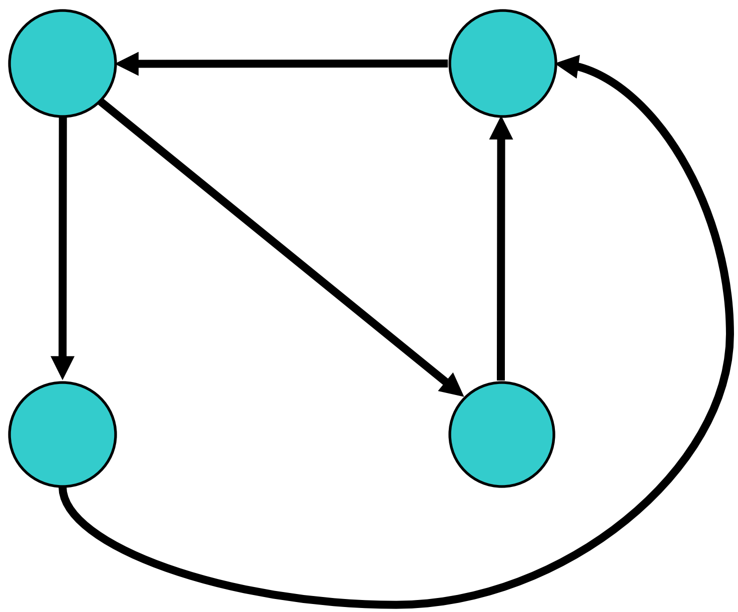
Ejemplo

La siguiente CM no es irreducible, y por tanto no es de ninguno de los demás tipos. Todos los estados son recurrentes y ninguno es periódico.



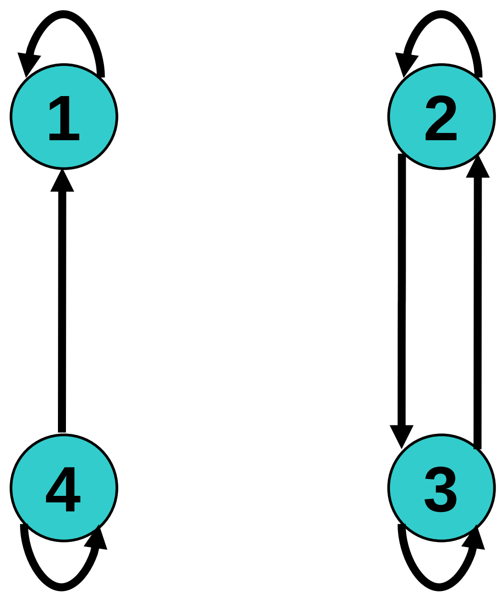
Ejemplo

La siguiente CM es irreducible, recurrente y periódica de periodo 3. No es ergódica.



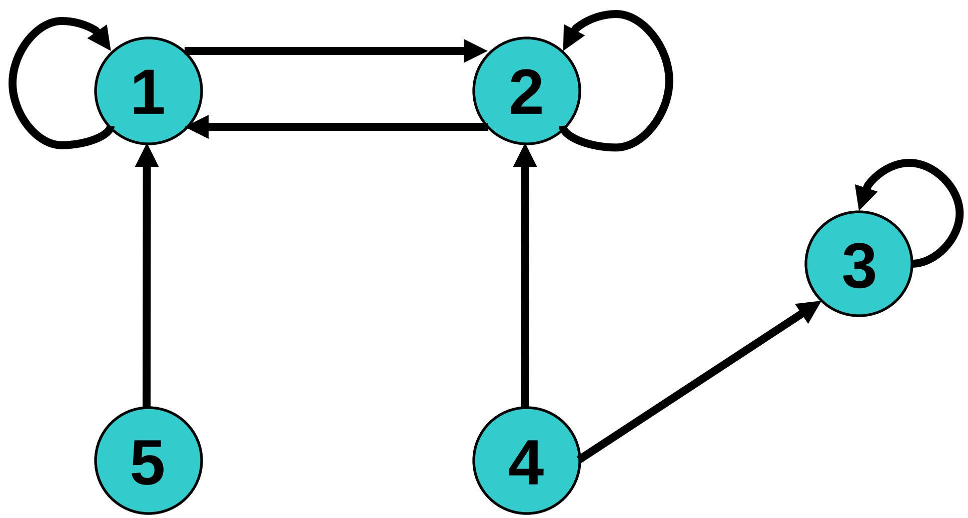
Ejemplo

La siguiente CM no es irreducible, y por tanto no es de ninguno de los demás tipos. Ningún estado es periódico. 4 es transitorio, y el resto recurrentes. 1 es absorbente.



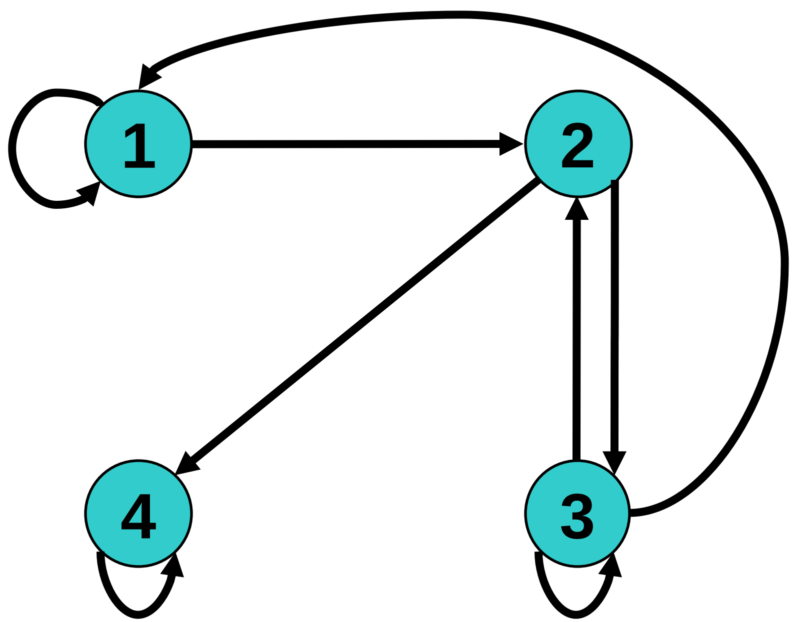
Ejemplo

La siguiente CM no es irreducible, y por tanto no es de ninguno de los demás tipos. Ningún estado es periódico. 4 y 5 son transitorios, y el resto recurrentes. 3 es absorbente.



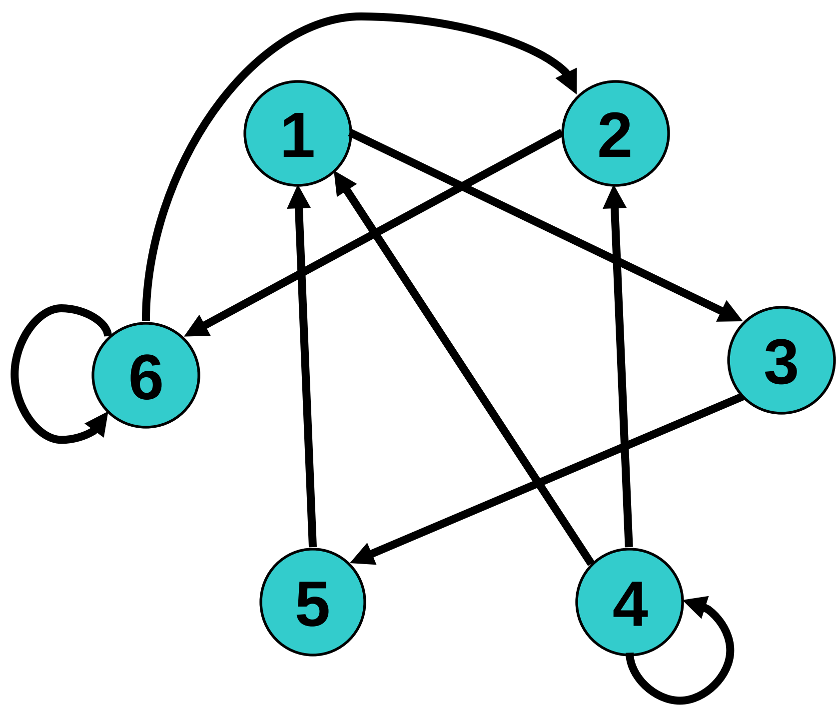
Ejemplo

La siguiente CM es no es irreducible, y por tanto tampoco de ninguno de los demás tipos. 4 es absorbente, y el resto son transitorios.



Ejemplo

La siguiente CM no es irreducible y por tanto no es de ninguno de los demás tipos. 1,3 y 5 son recurrentes de periodo 3. 2 y 6 son recurrentes, pero no periódicos. 4 es transitorio.



Cadenas absorbentes

Concepto de cadena absorbente

Sea *X* una CM cuyos estados son todos transitorios o absorbentes. En tal caso diremos que *X* es absorbente

Si *X* es finita y absorbente, reordenamos S poniendo primero los estados transitorios y obtenemos

$$Q = \left(\begin{array}{c:c} Q' & R \\ \hdashline 0 & I \end{array} \right)$$

Resultados sobre cadenas absorbentes

Proposición: El número medio de etapas que se estará en el estado transitorio $j \in S$ antes de la absorción, suponiendo que empezamos en el estado transitorio $i \in S$, viene dado por el elemento (i,j) de $(I - Q')^{-1}$

Nota: La etapa inicial también se cuenta, es decir, en la diagonal de $(I - Q')^{-1}$ todos los elementos son siempre mayores o iguales que 1

Ejemplo de CM absorbente

En un juego participan dos jugadores, A y B. En cada turno, se lanza una moneda al aire. Si sale cara, A le da 1 € a B. Si sale cruz, B le da 1 € a A. Al principio, A tiene 3 € y B tiene 2 €. El juego continúa hasta que alguno de los dos se arruine. Calcular:

a) La probabilidad de que A termine arruinándose.

b) La probabilidad de que B termine arruinándose.

c) El número medio de tiradas que tarda en acabar el juego.

Ejemplo de CM absorbente

Tendremos una CM con un estado por cada posible estado de cuentas de A: S={1, 2, 3, 4, 5, 0}. Descomponemos Q:

$$Q = \left(\begin{array}{cccc:cc} 0 & 0,5 & 0 & 0 & 0 & 0,5 \\ 0,5 & 0 & 0,5 & 0 & 0 & 0 \\ 0 & 0,5 & 0 & 0,5 & 0 & 0 \\ 0 & 0 & 0,5 & 0 & 0,5 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$Q' = \begin{pmatrix} 0 & 0,5 & 0 & 0 \\ 0,5 & 0 & 0,5 & 0 \\ 0 & 0,5 & 0 & 0,5 \\ 0 & 0 & 0,5 & 0 \end{pmatrix}$$

$$R = \begin{pmatrix} 0 & 0,5 \\ 0 & 0 \\ 0 & 0 \\ 0,5 & 0 \end{pmatrix}$$

Ejemplo de CM absorbente

Realizamos los cálculos necesarios:

$$(I-Q')^{-1} = \begin{pmatrix} 1 & -0,5 & 0 & 0 \\ -0,5 & 1 & -0,5 & 0 \\ 0 & -0,5 & 1 & -0,5 \\ 0 & 0 & -0,5 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1,6 & 1,2 & 0,8 & 0,4 \\ 1,2 & 2,4 & 1,6 & 0,8 \\ 0,8 & 1,6 & 2,4 & 1,2 \\ 0,4 & 0,8 & 1,2 & 1,6 \end{pmatrix}$$

$$(I-Q')^{-1}R = \begin{pmatrix} 0,2 & 0,8 \\ 0,4 & 0,6 \\ 0,6 & 0,4 \\ 0,8 & 0,2 \end{pmatrix}$$

Ejemplo de CM absorbente

Probabilidad de que A termine arruinándose.

La ruina de A está representada por el estado 0, que es el 2º estado absorbente. Como empezamos en el 3er estado transitorio (A empieza con 3 €), debemos consultar la 3ª fila, 2ª columna de $(I-Q')^{-1}R$, que nos da una probabilidad de 0,4 de que A empiece con 3 € y termine en la ruina.

Probabilidad de que B termine arruinándose

Como es el suceso contrario del apartado a), su probabilidad será 1–0,4=0,6. También podríamos haber consultado la 3ª fila, 1ª columna de $(I-Q')^{-1}R$.

Ejemplo de CM absorbente

- Número medio de tiradas que tarda en acabar el juego

  Sumamos los números medios de etapas que se estará en cualquier estado transitorio antes de la absorción, suponiendo que empezamos en el 3er estado transitorio. Dichos números medios son los que forman la 3ª fila de la matriz $(I-Q')^{-1}$. El promedio es: 0,8+1,6+2,4+1,2=6 tiradas.

- Nota: si observamos la 1ª columna de $(I-Q')^{-1}R$, vemos que los valores van creciendo. Esto se debe a que, cuanto más dinero tenga al principio A, más probabilidad tiene de ganar el juego.