# Probabilidad

## Cadenas de Markov
28/04/2023

Lara Sánchez Peña[1]

UTDT 2023 - MEC

[1] Basado en las notas de Andrea Rotnitzky

# Cadenas de Markov discretas

Definimos un **proceso estocástico discreto** como una familia de variables aleatorias $X_t$, con $t \in \mathbb{N}_0$ que toma valores en un espacio de **estados** $\Omega = \{s_1, \cdots, s_n\}$. Supongamos que $\#(\Omega) = n$.
Un proceso estocástico **tiene la propiedad de Markov** si para cualquier $t$, $X_{t+1}$ depende de $X_t$ pero no de $X_{t-1}, X_{t-2}, \cdots, X_1, X_0$.
Es decir,

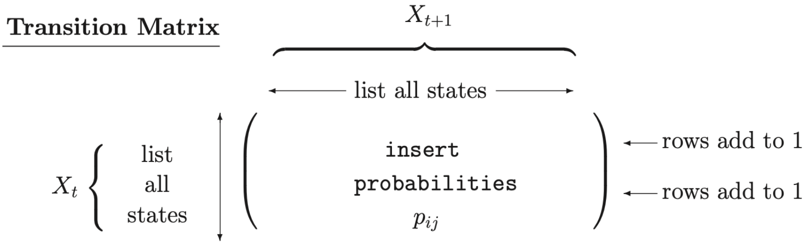
$$P(X_{t+1} = s | X_t = s_{(t)}, \cdots, X_1 = s_{(1)}, X_0 = s_{(0)}) = P(X_{t+1} = s | X_t = s_{(t)}). \tag{1}$$

Definimos una **cadena de Markov** como un proceso estocástico que cumple la propiedad de Markov (1) para todo $t = 0, 1, 2, \cdots$ y todos posibles valores $s, s_{(t)}, \cdots, s_{(1)}, s_{(0)}$.

# Cadenas de Markov: Matriz de transición

Definimos la matriz de transición $P$ como la matriz donde, en el lugar $ij$, se escriben las **probabilidades de transición** de pasar en el período $t$ del estado $s_i$ a pasar en el período $t+1$ al estado $s_j$ y se denotan

$$p_{ij} = Prob(X_{t+1} = s_j | X_t = s_i).$$




- Notemos que la **suma de los elementos de cada una de las filas** de $P$ suman 1 porque $\sum_{j=1}^{n} p_{ij} = \sum_{j=1}^{n} P(X_{t+1} = s_j | X_t = s_i) = 1$.
- Las suma de los elementos de cada una de las columnas de $P$ no tiene por qué sumar 1.

# Cadenas de Markov: prob. transición entre k períodos

¿Cuál es la **probabilidad de hacer una transición** del estado $s_i$ al $s_j$ a lo largo de **dos períodos**?

$$
\begin{aligned}
P(X_2 = s_j|X_0 = s_i) &= \sum_{k=1}^{n} P(X_2 = s_j|X_1 = s_k, X_0 = s_i)P(X_1 = s_k|S_0 = s_i) \\
&= \sum_{k=1}^{n} P(X_2 = s_j|X_1 = s_k)P(X_1 = s_k|X_0 = s_i) \\
&= \sum_{k=1}^{n} p_{kj} \cdot p_{ik} = \sum_{k=1}^{n} p_{ik} \cdot p_{kj} \\
&= (P^2)_{ij}
\end{aligned}
$$

De manera inductiva se puede probar que las probabilidades de transición de pasar del estado $s_i$ al $s_j$ a lo largo de $k$ períodos es:

$$
\begin{aligned}
P(X_3 = s_j|X_0 = s_i) &= (P^3)_{ij} \\
P(X_t = s_j|X_0 = s_i) &= P(X_{k+t} = j|X_t = i) = (P^k)_{ij}
\end{aligned}
$$

# Cadenas de Markov: Probabilidades marginales

Consideremos por otro lado las **distribuciones no condicionadas** de las variables $X_0, X_1, \cdots$.
Llamamos a la **distribución de probabilidades iniciales**:

$$\pi^{\mathsf{T}} = (\pi_1, \pi_2, \cdots, \pi_n) = (P(X_0 = s_1), P(X_0 = s_2), \cdots, P(X_0 = s_n))$$

Por lo tanto, para calcular

$$\begin{aligned} P(X_1 = s_j) &= \sum_{k=1}^{n} P(X_1 = s_j | X_0 = s_k) P(X_0 = s_k) \\ &= \sum_{k=1}^{n} p_{kj} \cdot \pi_k = \sum_{k=1}^{n} \pi_k \cdot p_{kj} \\ &= (\pi^{\mathsf{T}} P)_j \end{aligned}$$

De manera inductiva se puede probar que:

$$\begin{aligned} P(X_2 = s_j) &= (\pi^{\mathsf{T}} P^2)_j \\ P(X_k = s_j) &= (\pi^{\mathsf{T}} P^k)_j \end{aligned}$$

# Cadenas de Markov: probabilidad conjunta

Si queremos calcular la probabilidad conjunta de una trayectoria (también llamada **historia parcial**) $s_{(0)}, s_{(1)}, s_{(2)}, \cdots, s_{(t)}$ con $X_0 \sim \pi^{\mathsf{T}}$ se tiene que

$$
\begin{aligned}
&P\left(X_0 = s_{(0)}, X_1 = s_{(1)}, \ldots, X_t = s_{(t)}\right)\\
&= P\left(X_t = s_{(t)} \mid X_{t-1} = s_{(t-1)}, \ldots, X_0 = s_{(0)}\right) \cdot P\left(X_{t-1} = s_{(t-1)}, \ldots, X_0 = s_{(0)}\right)\\
&= P\left(X_t = s_{(t)} \mid X_{t-1} = s_{(t-1)}\right) \cdot P\left(X_{t-1} = s_{(t-1)}, \ldots, X_0 = s_{(0)}\right) \qquad \text{(Markov Property)}\\
&= p_{s_{(t-1)}, s_{(t)}} P\left(X_{t-1} = s_{(t-1)} \mid X_{t-2} = s_{(t-2)}, \ldots, X_0 = s_{(0)}\right) \cdot P\left(X_{t-2} = s_{(t-2)}, \ldots, X_0 = s_{(0)}\right)\\
&= p_{s_{(t-1)}, s_{(t)}} \cdot p_{s_{(t-2)}, s_{(t-1)}} \cdot \ldots \cdot p_{s_{(0)}, s_{(1)}} \cdot P\left(X_0 = s_{(0)}\right)\\
&= p_{s_{(t-1)}, s_{(t)}} \cdot p_{s_{(t-2)}, s_{(t-1)}} \cdot \ldots \cdot p_{s_{(0)}, s_{(1)}} \cdot \pi_{s_0}.
\end{aligned}
$$

# Cadenas de Markov: clasificación de los estados

- Definimos que un estado $s_i$ es **recurrente** si, empezando desde el estado $s_i$, la probabilidad de que la cadena de Markov vuelva **eventualmente** al estado $s_i$ es igual a 1.
- Definimos que un estado $s_i$ es **transitorio** si, empezando desde el estado $s_i$, la probabilidad de que la cadena de Markov vuelva **eventualmente** al estado $s_i$ es menor a 1. Es decir, puede ocurrir con probabilidad positiva que la cadena de Markov nunca vuelva al estado $s_i$.

Vale que si con probabilidad positiva no volvemos al estado $s_i$ entonces en algún momento la cadena no volverá al estado $s_i$, podemos pensarlo de la siguiente manera:

Consideremos un estado $s_i$ transitorio, podemos considerar $p$, la probabilidad de que, empezando desde el estado $s_i$, la cadena nunca vuelva al estado $s_i$. Por lo tanto, partiendo desde el estado $s_i$ la cantidad de veces que la cadena vuelve al estado $s_i$ antes de no volver nunca más tiene distribución Geométrica con parámetro $p$.

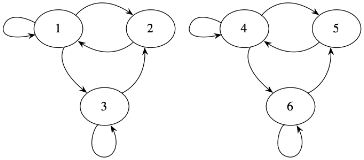
Un caso particular en el que sabremos que **todos los estados serán recurrentes** es cuando la cadena de Markov sea **irreducible**.

# Cadenas de Markov: definiciones

- Decimos que una cadena de Markov con matriz de transición $P$ es **irreducible** si para dos estados cualesquiera $s_i$ y $s_j$ es posible ir del estado $i$ al estado $j$ en una cantidad finita de períodos con probabilidad positiva. Esto es sinónimo de decir que para cualesquiera estados $s_i$ y $s_j$ existe un número $n \in \mathbb{N}$ de manera que la matriz $(P^n)_{ij} > 0$.
- **En una cadena de Markov irreducible con finitos estados, todos los estados son recurrentes**. Es decir, no puede haber eventos transitorios. Ver Blitzstein p.504. La intuición es que todos los estados no pueden ser transitorios porque sino la cadena de Markov dejaría eventualmente todos los estados y no tendría adónde ir. SPDG suponemos que $s_1$ es un estado recurrente. Consideremos el estado $s_i$. Como la cadena de Markov es irreducible existe un $n$ de manera que $(P^n)_{ij} > 0$, es decir, con probabilidad positiva luego de $n$ períodos podríamos pasar del estado $s_1$ al estado $s_i$. Pero como $s_1$ es un estado recurrente este proceso ocurrirá infinitas veces.
- Una cadena de Markov que no es irreducible se dice **reducible**.

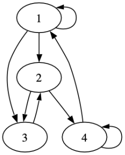
# Cadenas de Markov

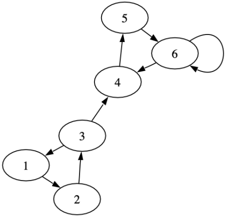
Notemos que puede suceder que una Cadena de Markov sea reducible y que sin embargo todos sus estados sean recurrentes. Por ejemplo:




Notemos que la recurrencia o transitividad es una propiedad de cada estado mientras que la reducibilidad o irreducibilidad es una propiedad de la cadena.

# Cadenas de Markov: estados transitorios y recurrentes




$$P_1 = \begin{pmatrix} 1/3 & 1/3 & 1/3 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 1 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 \end{pmatrix}$$

$$P_2 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \end{pmatrix}$$

En estas cadenas de Markov que no especifican las probabilidades es equiprobable salir del estado $s_i$ por cualquiera de las flechas que salen de ese estado.

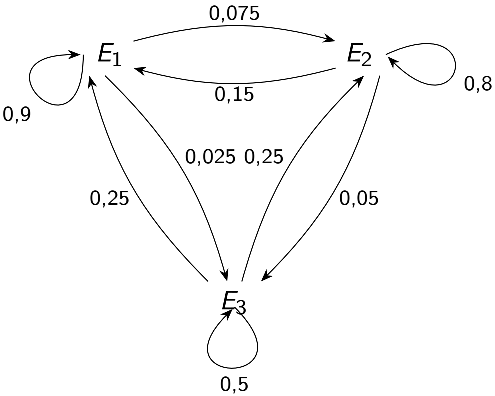
En la cadena de la izquierda todos los estados son recurrentes, en cambio en la cadena de la derecha los estados 1,2 y 3 son transitorios porque una vez que se pasa al estado 4 no se puede volver a dichos estados.

# Cadenas de Markov: ejemplo

Consideramos un modelo simple de 3 estados para un mercado financiero con 3 estados:

- $E_1$ : Mercado en crecimiento.
- $E_2$ : Mercado en decrecimiento.
- $E_3$ : Mercado estancado.

La cadena de Markov se puede representar gráficamente de la siguiente manera:



# Cadenas de Markov

Por lo tanto, podemos escribir la siguiente matriz de transición para describir la cadena de Markov antes mencionada.

$$P = \begin{bmatrix} 0,9 & 0,075 & 0,025 \\ 0,15 & 0,8 & 0,05 \\ 0,25 & 0,25 & 0,5 \end{bmatrix}$$

Supongamos que queremos conocer cómo se comporta este mercado a **largo plazo**. Para ellos, diagonalicemos la matriz $M$. Sus autovalores son:

$$\lambda_1 = 1, \lambda_2 = 0,741421356237310, \lambda_3 = 0,458578643762690$$

Además podemos encontrar una matriz de cambio de base $C$ tal que

$$CPC^{-1} = D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0,7414213562373091 & 0 \\ 0 & 0 & 0,4585786437626905 \end{pmatrix}$$

# Cadenas de Markov: comportamiento a largo plazo

donde $C = \begin{pmatrix} 0,5773502691896265 & 0,5773502691896251 & 0,5773502691896256 \\ 0,44371856511363317 & -0,8113070602072252 & -0,380650350100204 \\ -0,03400257431430442 & -0,13017637781608127 & 0,9909076322234505 \end{pmatrix}$

Entonces se tiene que

Luego $\lim_{t\to+\infty} P^t = \lim_{t\to+\infty} C^{-1} D^t C = \begin{pmatrix} 0,625 & 0,625 & 0,625 \\ 0,3125 & 0,3125 & 0,3125 \\ 0,0625 & 0,0625 & 0,0625 \end{pmatrix}$

Esto significaría que el mercado va a estar un 62,5 % del tiempo en estado alcista, un 32, 15 % en estado bajista y un 6, 25 % estancado. Este vector es un autovector de autovalor 1 de P es decir:

$$U_\infty^{\mathsf{T}} \cdot P = U_\infty^{\mathsf{T}}$$

donde $U_\infty^{\mathsf{T}} = \begin{pmatrix} 0,625 & 0,3125 & 0,0625 \end{pmatrix}$.

# Distribución estacionaria

- Definimos una distribución estacionaria como $r$ donde $r^\top P = r^\top$.
- Queremos comprender el comportamiento a largo plazo de una cadena de Markov. A largo plazo, sabemos que la cadena de Markov solamente estará en estados que sean recurrentes pero ¿qué "fracción del tiempo pasará en cada uno de ellos"? Esto lo determina la distribución estacionaria.
- Para una cadena de Markov **irreducible** existe una única **distribución estacionaria**. En esta distribución todos los estados tienen probabilidad positiva por el teorema de Perron-Frobenius.
- Para que la distribución marginal de la cadena de Markov **converja** a la distribución estacionaria tiene que ocurrir que la cadena de Markov sea **irreducible y aperiódica**.
- El período del estado $s_i$ es el máximo común divisor del conjunto $n \in \mathbb{N}$ de las probabilidades de $(P^n)_{ii} > 0$, partiendo del estado $s_i$ volver al estado $s_i$ luego de $n$ períodos de tiempo que sean positivas.
- Decimos que una cadena de Markov es **aperiódica** si cada estado tiene período 1.

# A veces no existe un comportamiento de largo plazo

Para la cadena de Markov **irreducible** que tiene asociada esta matriz

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

se puede ver probar que $r^{\mathsf{T}} = 0{,}2(1,1,1,1,1)$ es una distribución estacionaria de $P$ y sin embargo $P^n$ no converge. En esta cadena lo que sucede es que es períodica con período 5 y por lo tanto si bien hay una distribución asintótica, no sucede que la distribución converja a la distribución marginal del estado estacionario.

# ¿Puede no haber estado estacionario?

- Para cualquier cadena de Markov (bajo cualquier circunstancia) existe (al menos) una distribución de estado estacionario. En general esa(s) distribucion(es) de estado estacionario podrían potencialmente tener probabilidades iguales a 0 y podrían no ser únicas.
- Si el proceso es irreducible la distribución estacionaria es única y las probabilidades de la distribución estacionaria son todas positivas. Además de ser irreducible, si el proceso es aperiódico, sabés que la distribución de estado estacionario coincide con cualquier columna de lim n tendiendo a infinito de $P^n$ (este límite existe y en cada columna de ese límite tenés la distribución estacionaria que es el autovector de autovalor 1).

# ¿Puede no haber estado estacionario?

- Lo lindo es notar que para cualquier cadena de Markov vas a tener un autovector de autovalor 1. Acá hay dos claves:

  1. Los autovalores de $P$ y de $P$ transpuesta son los mismos (los autovectores sí son diferentes)
  2. la suma de los elementos de cada fila de $P$ suman 1.

- Buscamos $r$ la distribución de estado estacionario

$$r^{\mathsf{T}} P = r^{\mathsf{T}}.$$

- Ahora bien, notemos que si considero un vector $v = (1, 1, 1, 1, ...., 1)^{\mathsf{T}}$, vale que

$$P \cdot v = v$$

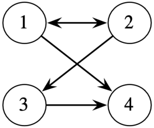
  porque justamente $P \cdot v$ suma los elementos de cada fila y cada una de esas sumas da 1.

# ¿Puede no haber estado estacionario?

- Entonces $v^{\mathsf{T}}$ es un autovector de $P^{\mathsf{T}}$ porque transponiendo la igualdad $Pv = v$ queda que
$$v^{\mathsf{T}} P^{\mathsf{T}} = v^{\mathsf{T}}.$$
- Entonces $\lambda = 1$ es autovalor de $P^{\mathsf{T}}$ y por lo tanto $\lambda = 1$ es autovalor de $P$. El tema es que el autovector de $P$ asociado a $\lambda = 1$ no tiene por qué ser $v$.
- Por ejemplo si $P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$ tanto $(1, 0, 0)^{\mathsf{T}}$ como $(0, 1, 0)^{\mathsf{T}}$ son distribuciones estacionarias de $P$ pero
$$\lim_{t \to +\infty} P^t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

# Google Page rank

Page rank busca ordenar las páginas según su relevancia: para cada página web tiene en cuenta a cuántas otras páginas dirige y también la relevancia de cada una de esas páginas. Imaginemos un ejemplos sencillo:




Hay páginas que no tienen links que dirigen hacia otras páginas como la 4. Entonces cuando esto sucede un navegador abre otra página y de manera aleatoria abre una página web. Entonces, cuando una página que no tiene links se convierte en una página que tiene links a cualquier otra incluida a sí misma.

# Google Page rank

La matriz de transición en este caso es

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

El problema es que lo que querríamos encontrar es una distribución estacionaria pero como no podemos garantizar que la cadena de Markov sea irreducible se agrega lo siguiente:
Se tira una moneda que con probabilidad $\alpha$ sale cara y en ese caso se aprieta un link aleatorio de la página en la que nos encontramos en el momento $t$. Con probabilidad $1-\alpha$ sale ceca y se abre una página al azar. Entonces, si consideramos la matriz de transición:

$$G = \alpha P + \frac{1}{N}(1-\alpha)J$$

donde $N$ es la cantidad de páginas que hay en la red, y $J$ es una matriz de $N \times N$ de unos.

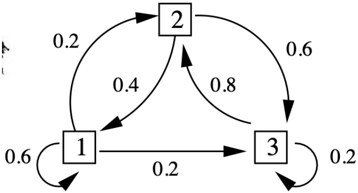
# Google Page rank

En este caso tenemos que la matriz $G$ tiene todas entradas positivas y por lo tanto buscar la distribución estacionaria es buscar $r$ de manera que $r^{\mathsf{T}} G = r^{\mathsf{T}}$.

Esto es buscar un autovector de autovalor 1, cosa que podría ser cuentosa, pero con cadenas de Markov podemos buscar $\lim_{t \to +\infty} vG^t$.

En la práctica Larry Page y Sergey Brin recomiendan $\alpha = 0{,}85$.

# Cadenas de Markov: ejercicio




Para la cadena de Markov calcule:

(a) Encuentre la matriz de transición $P$.

(b) Encuentre $P(X_2 = 3 \mid X_0 = 1)$

(c) Suponga que $\pi^{\mathsf{T}} = (1/3 \quad 1/3 \quad 1/3)$. Encuentre la distribución de $X_1$.

(d) Suponga que $\pi^{\mathsf{T}} = (1 \quad 0 \quad 0)$ Encuentre la distribución de $X_2$.

(e) Suponga que $\pi^{\mathsf{T}} = (1/3 \quad 1/3 \quad 1/3)$. Encuentre la probabilidad de que la cadena pase por los siguientes estados desde $t = 0$ hasta $t = 4$: $(3, 2, 1, 1, 3)$.

# Cadenas de Markov: ejercicio (soluciones)

(a) $P = \begin{pmatrix} 0{,}6 & 0{,}2 & 0{,}2 \\ 0{,}4 & 0 & 0{,}6 \\ 0 & 0{,}8 & 0{,}2 \end{pmatrix}$

(b)

$$\begin{aligned} P\left(X_2 = 3 \mid X_0 = 1\right) = \left(P^2\right)_{13} &= \begin{pmatrix} 0{,}6 & 0{,}2 & 0{,}2 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdot & \cdot & 0{,}2 \\ \cdot & \cdot & 0{,}6 \\ \cdot & \cdot & 0{,}2 \end{pmatrix} \\ &= 0{,}6 \cdot 0{,}2 + 0{,}2 \cdot 0{,}6 + 0{,}2 \cdot 0{,}2 \\ &= 0{,}28 \end{aligned}$$

Nota: Podríamos haber calculado toda la matriz $P^2$ pero no es necesario.

# Cadenas de Markov: ejercicio (soluciones)

(c) De esta información la distribución de $X_0$ es $\pi^{\mathsf{T}} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. Necesitamos calcular $X_1 \sim \pi^{\mathsf{T}} P$.

$$\boldsymbol{\pi}^{\mathsf{T}} P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 0{,}6 & 0{,}2 & 0{,}2 \\ 0{,}4 & 0 & 0{,}6 \\ 0 & 0{,}8 & 0{,}2 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

(d) La distribución de $X_0$ es $\boldsymbol{\pi}^{\mathsf{T}} = (1, 0, 0)$. Tenemos que calcular $X_2 \sim \boldsymbol{\pi}^{\mathsf{T}} P^2$.

$$\boldsymbol{\pi}^{\mathsf{T}} P^2 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0{,}6 & 0{,}2 & 0{,}2 \\ 0{,}4 & 0 & 0{,}6 \\ 0 & 0{,}8 & 0{,}2 \end{pmatrix} \begin{pmatrix} 0{,}6 & 0{,}2 & 0{,}2 \\ 0{,}4 & 0 & 0{,}6 \\ 0 & 0{,}8 & 0{,}2 \end{pmatrix}$$

# Cadenas de Markov: ejercicio (soluciones)

(d) Continuamos aquí:

$$\pi^{\mathsf{T}} P^2 = \begin{pmatrix} 0{,}6 & 0{,}2 & 0{,}2 \end{pmatrix} \begin{pmatrix} 0{,}6 & 0{,}2 & 0{,}2 \\ 0{,}4 & 0 & 0{,}6 \\ 0 & 0{,}8 & 0{,}2 \end{pmatrix}$$
$$= \begin{pmatrix} 0{,}44 & 0{,}28 & 0{,}28 \end{pmatrix}.$$

Thus $P(X_2 = 1) = 0{,}44$, $P(X_2 = 2) = 0{,}28$, $P(X_2 = 3) = 0{,}28$.
Notemos que conviene multiplicar $\pi^{\mathsf{T}}$ por P primero: nonecesitamos calcular toda la matriz $P^2$.

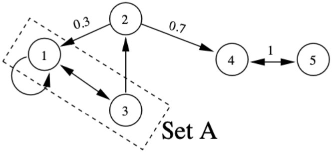
(e)

$$P(3,2,1,1,3) = P(X_0 = 3) \cdot p_{32} \cdot p_{21} \cdot p_{11} \cdot p_{13}$$
$$= \frac{1}{3} \cdot 0{,}8 \cdot 0{,}4 \cdot 0{,}6 \cdot 0{,}2$$
$$= 0{,}0128$$

# Cadenas de Markov: hitting probabilities

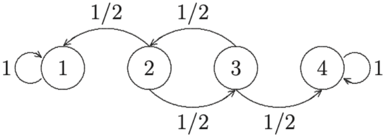
Consideremos una familia $A \subseteq \Omega$ de nodos, definimos la **probabilidad de pegarle a la familia** $A$ empezando desde el estado $s_i$ como $h_{iA} = P(X_t \in A | X_0 = s_i)$

Para $A = \{1, 3\}$




$h_{1A} = 1$, $h_{2A} = 0{,}3$, $h_{3A} = 1$, $h_{4A} = 0$, $h_{5A} = 0$

Para $A = \{4\}$




$h_{14} = 0$, $h_{24} = 0{,}5 h_{34}$, $h_{34} = 0{,}5 + 0{,}5 h_{24}$, $h_{44} = 1$. Entonces $h_{24} = \frac{1}{3}$ y $h_{34} = \frac{2}{3}$.

# Cadenas de Markov: expected hitting time

Definimos el tiempo esperado de parada al conjunto $A$ empezando desde el estado inicial $s_i$ como $m_{iA} = E(T_A | X_0 = s_i)$ donde $T_A = \text{mín}\{t \geq 0 : s_{(t)} \in A\}$
Vale que

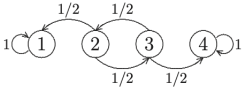
$$m_{iA} = \begin{cases} 0 & \text{si} \quad s_i \in A \\ 1 + \sum_{j \notin A} p_{ij} m_{jA} & \text{si} \quad s_i \notin A \end{cases}$$

Demostración: Si $m_{iA} = 0$ si $s_i \in A$. Supongamos que $s_i \notin A$. Entonces

$$\begin{aligned} m_{iA} =& E\left(T_A \mid X_0 = i\right) \\ =& 1 + \sum_{s_j \in S} E\left(T_A \mid X_1 = s_j\right) P\left(X_1 = s_j \mid X_0 = i\right) \\ =& 1 + \sum_{s_j \in S} m_{jA} p_{ij} \quad \text{(por definición)} \\ =& 1 + \sum_{s_j \notin A} p_{ij} m_{jA}, \quad \text{porque } m_{jA} = 0 \text{ para } s_j \in A. \end{aligned}$$

# Cadenas de Markov: expected hitting time

Empezando desde el estado $s_i = 2$, queremos calcular el tiempo esperado de llegar al conjunto $A = \{1, 4\}$ (el conjunto de los estados absorbentes).




Buscamos $m_{iA} = m_{2A}$.

$$m_{iA} = \begin{cases} 0 & \text{si} \quad i \in \{1, 4\}, \\ 1 + \sum_{j \notin A} p_{ij} m_{jA} & \text{si } i \notin \{1, 4\}. \end{cases}$$

Sabemos que

$$m_{1A} = 0 \quad (\text{ porque } 1 \in A)$$
$$m_{4A} = 0 \quad (\text{ porque } 4 \in A)$$

# Cadenas de Markov: expected hitting time

Entonces

$$m_{2A} = 1 + \frac{1}{2} m_{1A} + \frac{1}{2} m_{3A} \Rightarrow \quad m_{2A} = 1 + \frac{1}{2} m_{3A}$$

$$m_{3A} = 1 + \frac{1}{2} m_{2A} + \frac{1}{2} m_{4A} = 1 + \frac{1}{2} m_{2A}$$

$$m_{3A} = 1 + \frac{1}{2} \left( 1 + \frac{1}{2} m_{3A} \right) \Rightarrow \quad \frac{3}{4} m_{3A} = \frac{3}{2} \Rightarrow \quad m_{3A} = 2.$$

Entonces

$$m_{2A} = 1 + \frac{1}{2} m_{3A} = 2$$

El tiempo esperado para llegar a los estados absorbentes es de $E\left(T_A\right) = 2$ pasos.