Cadenas de Markov

Los procesos de paseo aleatorio en realidad son un caso particular de procesos más generales que son las cadenas de Markov. En esencia, una cadena es un proceso en tiempo discreto en el que una variable aleatoria X_n va cambiando con el paso del tiempo. Las cadenas de Markov tienen la propiedad de que la probabilidad de que $X_n = j$ sólo depende del estado inmediatamente anterior del sistema: X_{n-1}. Cuando en una cadena dichas probabilidades no dependen del tiempo en que se considere, n,

$$P\left(X_n = j \mid X_{n-1} = i\right)$$

se denomina *cadena homogénea*, esto es, las probabilidades son las mismas en cada paso.

Probabilidades de Transición

En una cadena homogénea finita con m posibles estados $E_1, E_2, \ldots, E_m$ se puede introducir la notación

$$p_{ij} = P\left(X_n = j \mid X_{n-1} = i\right),$$

donde $i, j = 1, 2, \ldots, m$. Si $p_{ij} > 0$ entonces se dice que el estado E_i puede *comunicar* con E_j. La comunicación puede ser mutua si también $p_{ji} > 0$.

Para cada i fijo, la serie de valores $\{p_{ij}\}$ es una distribución de probabilidad, ya que en cualquier paso puede ocurrir alguno de los sucesos $E_1, E_2, \ldots, E_m$ y son mutuamente excluyentes. Los valores p_{ij} se denominan *probabilidades de transición* que satisfacen la condición

$$p_{ij} > 0,$$
$$\sum_{j=1}^{m} p_{ij} = 1,$$

para cada $i = 1, 2, \ldots, m$. Todos estos valores se combinan formando una *matriz de transición* T de tamaño $m \times m$, donde

$$T = [p_{ij}] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix}$$

Se puede observar que cada fila de la matriz es una distribución de probabilidad, es decir, $\sum_{j=1}^{m} p_{ij} = 1$.

Observación.

Si las matrices $A = [a_{ij}]$ y $B = [b_{ij}]$ son matrices estocásticas, entonces $C = A \cdot B$ es también estocástica.

Por la regla de multiplicación de matrices,

$$C = [c_{ij}] = [a_{ij}] \cdot [b_{ij}] = \left[\sum_{k=1}^{m} a_{ik} b_{kj} \right]$$

De este modo

$$\sum_{j=1}^{m} c_{ij} = \sum_{j=1}^{m} \sum_{k=1}^{m} a_{ik} b_{kj} = \sum_{k=1}^{m} a_{ik} \sum_{j=1}^{m} b_{kj} = \sum_{k=1}^{m} a_{ik} \cdot 1 = 1.$$

Una consecuencia es que cualquier potencia de la matriz T es también una matriz estocástica: T^n.

Probabilidad $p_j^{(n)}$

Una probabilidad de bastante interés es la probabilidad de llegar a E_j después de n pasos, dada una distribución de probabilidad $\left\{ p_i^{(0)} \right\}$.

Se observa que $\left\{p_i^{(0)}\right\}$ es la probabilidad de que el sistema ocupe inicialmente el estado E_i, de modo que

$$\sum_{i=1}^{m} p_i^{(0)} = 1.$$

Si se denomina $p_j^{(1)}$ a la probabilidad de alcanzar E_j en un solo paso, entonces, por el teorema de probabilidad total

$$p_j^{(1)} = \sum_{i=1}^{m} p_i^{(0)} p_{ij}.$$

Esto se puede expresar de forma vectorial: sean $\mathbf{p}^{(0)}$ y $\mathbf{p}^{(1)}$ los vectores fila de probabilidad dados por

$$\mathbf{p}^{(0)} = \left(p_1^{(0)}, \ldots, p_m^{(0)}\right)$$

y

$$\mathbf{p}^{(1)} = \left(p_1^{(1)}, \ldots, p_m^{(1)}\right),$$

donde $\mathbf{p}^{(0)}$ es la distribución de probabilidad inicial y $\mathbf{p}^{(1)}$ es la probabilidad de que se alcance cada uno de los estados $E_1, \ldots, E_m$ después de un paso. Con esta notación, se puede expresar

$$\mathbf{p}^{(1)} = \left[p_j^{(1)}\right] = \left[\sum_{i=1}^{m} p_i^{(0)} p_{ij}\right] = \mathbf{p}^{(0)} T$$

donde T es la matriz de transición.

Del mismo modo,

$$\mathbf{p}^{(2)} = \mathbf{p}^{(1)} T = \mathbf{p}^{(0)} T^2$$

y en n pasos,

$$\mathbf{p}^{(n)} = \mathbf{p}^{(n-1)} T = \mathbf{p}^{(0)} T^n$$

donde

$$\mathbf{p}^{(n)} = \left(p_1^{(n)}, \ldots, p_m^{(n)}\right)$$

y de manera general,

$$\mathbf{p}^{(n+r)} = \mathbf{p}^{(r)} T^n$$

NOTA: $p_j^{(n)}$ es la probabilidad incondicional de estar en el estado E_j en el n-ésimo paso, dado que la probabilidad inicial es $\mathbf{p}^{(0)}$, esto es,

$$P\left(X_n = j\right) = p_j^{(n)},$$

que es tal que

$$\sum_{j=1}^{m} p_j^{(n)} = 1.$$

Probabilidad de transición en n pasos $p_{ij}^{(n)}$

Se define $p_{ij}^{(n)}$ como la probabilidad de que la cadena esté en el estado E_j después de n pasos, dado que la cadena empezó en el estado E_i. Se tiene que

$$p_{ij}^{(n)} = P\left(X_n = j \mid X_0 = i\right)$$

por la propiedad markoviana se tiene que

$$p_{ij}^{(n)} = \sum_{k=1}^{m} P\left(X_n = j, X_{n-1} = k \mid X_0 = i\right),$$

para $n \geq 2$, ya que la cadena debe haber pasado por uno de los m posibles estados en la etapa $n-1$.

NOTA: Se tiene la siguiente igualdad, para tres posibles sucesos A, B y C :

$$P\left(A \cap B \mid C\right) = P\left(A \mid B \cap C\right) \cdot P\left(B \mid C\right)$$

Si se sustituye

$$\begin{aligned} A &\rightarrow (X_n = j) \\ B &\rightarrow (X_{n-1} = k) \\ C &\rightarrow (X_0 = i) \end{aligned}$$

entonces

$$\begin{aligned}
p_{ij}^{(n)} &= \sum_{k=1}^{m} P\left(X_n = j, X_{n-1} = k \mid X_0 = i\right) = \\
&= \sum_{k=1}^{m} P\left(X_n = j \mid X_{n-1} = k, X_0 = i\right) P\left(X_{n-1} = k \mid X_0 = i\right) = \\
&= \sum_{k=1}^{m} P\left(X_n = j \mid X_{n-1} = k\right) P\left(X_{n-1} = k \mid X_0 = i\right) = \\
&= \sum_{k=1}^{m} p_{kj}^{(1)} p_{ik}^{(n-1)} = \sum_{k=1}^{m} p_{ik}^{(n-1)} p_{kj}^{(1)}
\end{aligned}$$

usando la propiedad markoviana nuevamente. La ecuación anterior se denomina de *Chapman-Kolmogorov.*

Haciendo n igual a $2, 3, \ldots$ se obtiene que las matrices con esos elementos son

$$\begin{aligned}
\left[p_{ij}^{(2)}\right] &= \left[p_{ik}^{(1)} p_{kj}^{(1)}\right] = T^2 \\
\left[p_{ij}^{(3)}\right] &= \left[p_{ik}^{(2)} p_{kj}^{(1)}\right] = T^3
\end{aligned}$$

ya que $p_{ik}^{(2)}$ son los elementos de T^2 y así sucesivamente. Así,

$$\left[p_{ij}^{(n)}\right] = T^n.$$

Ejemplo.

En una cierta región el tiempo atmosférico sigue la siguiente secuencia: Un día se denomina *soleado* (S) si el sol luce más de la mitad del día, y se denomina *nublado* (N), si lo hace menos. Por experiencia, se sabe que si hay un día nublado, es igual de probable que el día siguiente sea también nublado. Si el día es soleado hay una probabilidad de $2/3$ de que sea también soleado.

1. Construye la matriz de transición T de este proceso.

2. Si hoy está nublado, ¿cuál es la probabilidad de que dentro de tres días esté también nublado? ¿y de que esté soleado?

3. Calcula T^5 y T^{10}. ¿Cuál es el comportamiento de T^n cuando $n \to \infty$? ¿Cómo se comporta $\mathbf{p}^{(n)}$ cuando $n \to \infty$? ¿Depende el límite de $\mathbf{p}^{(0)}$?

(1) Asumimos que el proceso es markoviamo y que el tiempo sólo depende del día anterior. Se tiene una cadena de Markov con dos estados: $E_1 \equiv$ Día nublado N, $E_2 \equiv$ Día soleado S. La matriz de transición es:

	N	S
N	$\frac{1}{2}$	$\frac{1}{2}$
S	$\frac{1}{3}$	$\frac{2}{3}$

es decir,

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}.$$

(2) Empezando los pasos a partir del día actual, se tiene que

$$\mathbf{p}^{(0)} = \begin{pmatrix} p_1^{(0)} & p_2^{(0)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

De este modo, si hoy está nublado, dentro de tres días, se tiene que

$$\mathbf{p}^{(3)} = \mathbf{p}^{(0)} T^3 = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}^3 =$$
$$= \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{29}{72} & \frac{43}{72} \\ \frac{43}{108} & \frac{65}{108} \end{pmatrix} = \begin{pmatrix} \frac{29}{72} & \frac{43}{72} \end{pmatrix} = \begin{pmatrix} 0{,}403 & 0{,}597 \end{pmatrix}$$

Así, la probabilidad de que haya un día nublado en el tercer día es $p_1^{(3)} = \frac{29}{72}$ y que sea soleado es $p_2^{(3)} = \frac{43}{72}$.

(3)

$$T^5 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}^5 = \begin{pmatrix} 0{,}40008 & 0{,}59992 \\ 0{,}39995 & 0{,}60005 \end{pmatrix}$$

$$T^{10} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}^{10} = \begin{pmatrix} 0{,}4 & 0{,}60000 \\ 0{,}40000 & 0{,}6 \end{pmatrix}.$$

Cálculos así se facilitan bastante usando programas tipo `MatLab`, de modo que

$$T^n \to \begin{pmatrix} 0{,}4 & 0{,}6 \\ 0{,}4 & 0{,}6 \end{pmatrix} = Q$$

de este modo,

$$\begin{aligned}
\mathbf{p}^{(n)} &= \mathbf{p}^{(0)} T^n = \begin{pmatrix} p_1^{(0)} & p_2^{(0)} \end{pmatrix} \begin{pmatrix} 0{,}4 & 0{,}6 \\ 0{,}4 & 0{,}6 \end{pmatrix} = \\
&= \begin{pmatrix} \left(p_1^{(0)} + p_2^{(0)}\right) 0{,}4 & \left(p_1^{(0)} + p_2^{(0)}\right) 0{,}6 \end{pmatrix} = \\
&= \begin{pmatrix} 0{,}4 & 0{,}6 \end{pmatrix},
\end{aligned}$$

ya que $p_1^{(0)} + p_2^{(0)} = 1$.

Se observa que cuando $n \to \infty$ $\mathbf{p}^{(n)}$ no depende de dónde se parte: $\mathbf{p}^{(0)}$.

Potencias de una matriz de transición

Es obvio que resulta necesario definir un método para calcular potencias de una matriz. Esto se hace en términos de diagonalización de matrices en términos de autovalores y autovectores.

Sea T una matriz de transición de orden $m \times m$, los autovalores de T se encuentran a partir de la ecuación característica

$$|T - \lambda I_m| = 0,$$

donde I_m es la matriz identidad de orden $m \times m$. Supongamos que los autovalores son distintos y los denotamos por $\lambda_1, \ldots, \lambda_m$. Los correspondientes autovectores cumplen las siguientes ecuaciones

$$(T - \lambda_i I_m)\, r_i = 0$$

para $i = 1, \ldots, m$.

Construyo la matriz

$$C = \begin{pmatrix} r_1 & r_2 & \cdots & r_m \end{pmatrix},$$

donde los autovectores están colocados por columnas. Multiplicando las matrices,

$$
\begin{aligned}
TC &= T \left(\begin{array}{cccc} r_1 & r_2 & \cdots & r_m \end{array} \right) = \\
\left(\begin{array}{cccc} Tr_1 & Tr_2 & \cdots & Tr_m \end{array} \right) &= \left(\begin{array}{cccc} \lambda_1 r_1 & \lambda_2 r_2 & \cdots & \lambda_m r_m \end{array} \right) = \\
\left(\begin{array}{cccc} r_1 & r_2 & \cdots & r_m \end{array} \right) \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{pmatrix} &= CD
\end{aligned}
$$

donde D es la matriz diagonal con los autovalores situados en la diagonal principal.

Como se deduce que

$$
\begin{aligned}
TC &= CD \Longrightarrow \\
T &= CDC^{-1}
\end{aligned}
$$

Se puede calcular

$$
T^2 = \left(CDC^{-1}\right)\left(CDC^{-1}\right) = CD^2C^{-1}
$$

y del mismo modo, en general,

$$
T^n = CD^nC^{-1},
$$

donde

$$
D^n = \begin{pmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m^n \end{pmatrix}.
$$

Aunque este método es bastante útil resulta un poco pesado a la hora de hacer los cálculos, incluso para cadenas a partir de $m = 4$ ó 5.

En `MatLab` la sentencia para calcular autovectores y autovalores es

```
[C,D]=eig(T);
```

Ejemplo.

Calcular los autovectores y autovalores de la matriz estocástica

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

y determinar T^n y el límite $\lim_{n\to\infty} T^n$.

si se calcula

$$|T - \lambda I_m| = 0,$$

se obtiene

$$\begin{vmatrix} \frac{1}{4}-\lambda & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4}-\lambda & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2}-\lambda \end{vmatrix} = 0 \Longrightarrow$$

$$(1-\lambda)\left(\frac{1}{4}-\lambda\right)\left(-\frac{1}{4}-\lambda\right) = 0$$

luego los autovalores son

$$\lambda_1 = 1, \quad \lambda_2 = \tfrac{1}{4} \quad \lambda_3 = -\tfrac{1}{4}$$

Del mismo modo, los autovectores se calculan a partir de

$$(T - \lambda_i I_m)\, r_i = 0$$

y se obtiene

$$r_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad r_1 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \quad r_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

y de este modo

$$C = \begin{pmatrix} r_1 & r_2 & r_2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 0 \end{pmatrix}.$$

$$\begin{aligned} T^n &= CD^nC^{-1} = \\ & \begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \left(\frac{1}{4}\right)^n & 0 \\ 0 & 0 & \left(-\frac{1}{4}\right)^n \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{6} & -\frac{1}{6} & \frac{1}{3} \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}. \end{aligned}$$

Cuando $n \to \infty$, se tiene que

$$\begin{aligned} T^n &\to \begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{6} & -\frac{1}{6} & \frac{1}{3} \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix} = \\ &= \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} = Q. \end{aligned}$$

Supongamos que T es la matriz de transición de una cadena de Markov con 3 estados y que la probabilidad inicial es $\mathbf{p}^{(0)}$, entonces, después de n etapas

$$\mathbf{p}^{(n)} = \mathbf{p}^{(0)} T^n$$

y la distribución límite es

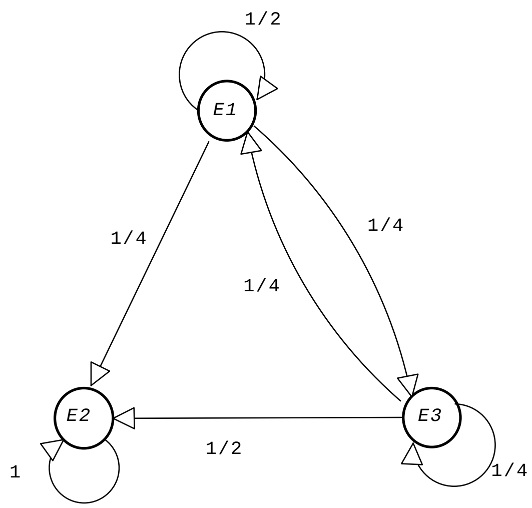
$$\mathbf{p} = \lim_{n\to\infty} \mathbf{p}^{(n)} = \lim_{n\to\infty} \mathbf{p}^{(0)} T^n = \mathbf{p}^{(0)} Q = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

El vector $\mathbf{p}$ da las probabilidades de los estados a largo plazo, y esta distribución es independiente del estado inicial $\mathbf{p}^{(0)}$.

Ejemplo.

Se pueden representar las transiciones entre los estados mediante grafos, donde los arcos representan las transiciones entre los estados. Por ejemplo, sea el siguiente diagrama de transición:

cuya matriz de transición es:



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

Se puede observar, que una vez se entra en el estado E_2 ya no se puede volver a salir de él. Es lo que se denomina un estado *absorbente*.

Los autovalores son

$$\lambda_1 = 1, \quad \lambda_2 = \tfrac{1}{4} \quad \lambda_3 = \tfrac{3}{4}$$

y la matriz de autovectores es

$$T = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

$$\begin{aligned} T^n &= CD^nC^{-1} = \\ & \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \left(\frac{1}{4}\right)^n & 0 \\ 0 & 0 & \left(\frac{3}{4}\right)^n \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{pmatrix}. \end{aligned}$$

Cuando $n \to \infty$, se tiene que

$$\begin{aligned} T^n &\to \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} = Q. \end{aligned}$$

esto implica que

$$\mathbf{p} = \lim_{n\to\infty} \mathbf{p}^{(n)} = \lim_{n\to\infty} \mathbf{p}^{(0)} T^n = \mathbf{p}^{(0)} Q = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$$

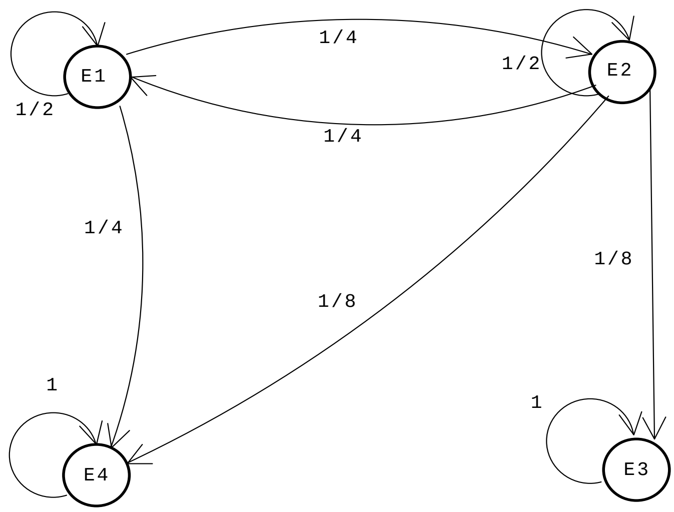
es decir, el sistema termina yendo a E_2 con probabilidad 1.

Ejemplo.

Modelo de enfermedad. Este modelo se puede representar por 4 estados: E_1 donde las personas no tienen enfermedad, E_2 donde las personas tienen la enfermedad, E_3 estado muerte como consecuencia de la enfermedad y E_4 estado muerte por otras causas Suipongamos la siguiente matriz de transición

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

el correspondiente diagrama de transición es



Se tienen dos estados absorbentes E_3 y E_4, y la interpretación que de las probabilidades individuales son las siguientes: $p_{11} = \frac{1}{2}$ significa que una persona dado que no está enferma en un periodo tiene una probabilidad de $1/2$ de no estarlo en el siguiente periodo, $p_{24} = \frac{1}{8}$ significa que una persona dado que no está enferma en un periodo tiene una probabilidad de $1/8$ de morir en el siguiente periodo por otras causas y así sucesivamente.

En este ejemplo la matriz de transición tiene la siguiente forma típica:

$$T = \begin{pmatrix} A & B \\ O_{22} & I_2 \end{pmatrix}$$

donde, en este caso,

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{pmatrix} \quad B = \begin{pmatrix} 0 & \frac{1}{4} \\ \frac{1}{8} & \frac{1}{8} \end{pmatrix} \quad O_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Se puede observar que tanto A como B no son matrices estocásticas. Se obtiene que

$$T^2 = \begin{pmatrix} A & B \\ O_{22} & I_2 \end{pmatrix} \begin{pmatrix} A & B \\ O_{22} & I_2 \end{pmatrix} = \begin{pmatrix} A^2 & (I_2 + A)\,B \\ O_{22} & I_2 \end{pmatrix}.$$

$$T^2 = \begin{pmatrix} A^3 & (I_2 + A + A^2)\,B \\ O_{22} & I_2 \end{pmatrix}$$

y, en general,

$$T^n = \begin{pmatrix} A^n & \left(I_2 + A + A^2 + \cdots + A^{n-1}\right) B \\ O_{22} & I_2 \end{pmatrix}$$

Si se denomina

$$S_n = I_2 + A + A^2 + \cdots + A^{n-1}$$

entonces,

$$AS_n = A + A^2 + \cdots + A^n$$

de modo que

$$\left(I_2 - A\right) S_n = \left(I_2 + A + A^2 + \cdots + A^{n-1}\right) - \left(A + A^2 + \cdots + A^n\right) = I_2 - A^n$$

luego

$$S_n = \left(I_2 - A\right)^{-1} \left(I_2 - A^n\right)$$

quedando

$$T^n = \begin{pmatrix} A^n & \left(I_2 - A\right)^{-1} \left(I_2 - A^n\right) B \\ O_{22} & I_2 \end{pmatrix}.$$

Los cálculos, a continuación, son semejantes al caso anterior.

Los autovalores de A son $\lambda_1 = \frac{1}{4}$, $\lambda_2 = \frac{3}{4}$, la matriz de autovectores es

$$C = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

quedando

$$A^n = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \left(\frac{1}{4}\right)^n & 0 \\ 0 & \left(\frac{3}{4}\right)^n \end{pmatrix} \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Cuando $n \to \infty$, $A^n \to O_{22}$, de modo que

$$\begin{aligned} T^n &= \begin{pmatrix} O_{22} & \left(I_2 - A\right)^{-1} B \\ O_{22} & I_2 \end{pmatrix} = \\ &= \begin{pmatrix} 0 & 0 & \frac{1}{6} & \frac{5}{6} \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

En este caso, la distribución límite es

$$\mathbf{p} = \begin{pmatrix} 0 & 0 & \frac{1}{6}p_1^{(0)} + \frac{1}{3}p_2^{(0)} + p_3^{(0)} & \frac{5}{6}p_1^{(0)} + \frac{2}{3}p_2^{(0)} + p_4^{(0)} \end{pmatrix}$$

luego la distribución límite depende de las condiciones iniciales. Por ejemplo, la probabilidad de que un individuo sano muera por enfermedad es, dado que la condición inicial es

$$\mathbf{p}^{(0)} = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathbf{p} = \begin{pmatrix} 0 & 0 & \frac{1}{6} & \frac{5}{6} \end{pmatrix}$$

es decir, $\frac{1}{6}$.

Caso de autovalores complejos.

En el caso de que los autovalores de la matriz de transición sean complejos, el procedimiento es igual:

Sea la matriz

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{8} & \frac{3}{8} \\ 1 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

los autovalores son

$$\begin{pmatrix} \lambda_1 = -\frac{1}{8} + \frac{3}{8}i & \lambda_2 = -\frac{1}{8} - \frac{3}{8}i & \lambda_3 = 1 \end{pmatrix}$$

y la matriz de autovectores se obtiene como antes (con un poco más de esfuerzo, o bien usando `MatLab`) mediante

$$(T - \lambda_i I_m)\, r_i = 0$$

y se obtiene

$$C = \begin{pmatrix} -\frac{1}{8} + \frac{3}{8}i & -\frac{1}{8} - \frac{3}{8}i & 1 \\ 1 & 1 & 1 \\ -\frac{1}{2} - \frac{3}{4}i & -\frac{1}{2} + \frac{3}{4}i & 1 \end{pmatrix}$$

y su inversa es

$$C^{-1} = \begin{pmatrix} -\frac{4}{15} - \frac{8}{15}i & \frac{2}{5} + \frac{2}{15}i & -\frac{2}{15} + \frac{2}{5}i \\ -\frac{4}{15} + \frac{8}{15}i & \frac{2}{5} - \frac{2}{15}i & -\frac{2}{15} - \frac{2}{5}i \\ \frac{8}{15} & \frac{1}{5} & \frac{4}{15} \end{pmatrix}$$

$$D^n = \begin{pmatrix} \left(-\frac{1}{8}+\frac{3}{8}i\right)^n & 0 & 0 \\ 0 & \left(-\frac{1}{8}-\frac{3}{8}i\right)^n & 0 \\ 0 & 0 & 1 \end{pmatrix} \underset{n\to\infty}{\longrightarrow} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

ya que $\left|-\frac{1}{8}+\frac{3}{8}i\right| = \sqrt{\frac{10}{8}} < 1$. Finalmente,

$$T^n \underset{n\to\infty}{\longrightarrow} \begin{pmatrix} \frac{8}{15} & \frac{1}{5} & \frac{4}{15} \\ \frac{8}{15} & \frac{1}{5} & \frac{4}{15} \\ \frac{8}{15} & \frac{1}{5} & \frac{4}{15} \end{pmatrix} = Q$$

Clasificación de los estados

En una cadena homogénea con m estados $E_1, E_2, \ldots, E_m$ y matriz de transición $T = [p_{ij}]$, $(1 \leq i, j \leq m)$ el valor de p_{ij} es la probabilidad de que haya una transición entre E_i y E_j en un momento dado. Según lo anterior se pueden clasificar los estados de una cadena.

Estado absorbente

Un estado es absorbente cuando una vez que se entra en él no se puede salir del mismo. Un estado E_i es absorbente si

$$\begin{aligned} p_{ii} &= 1 \\ p_{ij} &= 0 \qquad (i \neq j, \quad j = 1, \ldots, m) \end{aligned}$$

en la i-ésima fila de T.

Estado periódico

La probabilidad de que se regrese al estado E_i en el paso n es $p_{ii}^{(n)}$. Sea t un número entero mayor que 1. Supongamos que

$$\begin{aligned} p_{ii}^{(n)} &= 0 \qquad \text{para } n \neq t, 2t, 3t, \ldots \\ p_{ii}^{(n)} &\neq 0 \qquad \text{para } n = t, 2t, 3t, \ldots \end{aligned}$$

En este caso se dice que el estado E_i es *periódico* de periodo t. Si para un estado no existe dicho valor de t entonces se dice que el estado es *aperiódico*.

Alternativamente, se puede definir

$$d(i) = mcd\left\{n \mid p_{ii}^{(n)} > 0\right\},$$

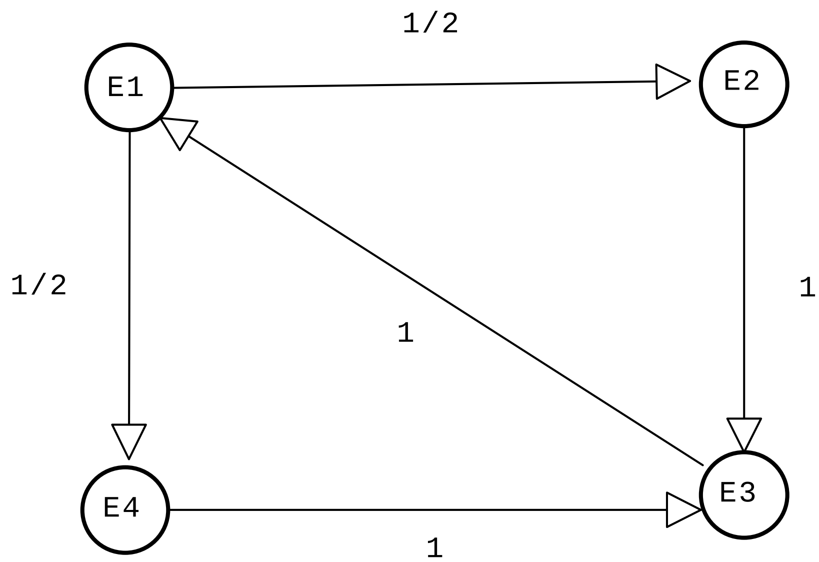
es decir, el máximo común divisor de (mcd) del conjunto de los enteros n para los que $p_{ii}^{(n)} > 0$. Entonces, el estado E_i es periódico si $d(i) > 0$ y aperiódico si $d(i) = 1$

Ejemplo.

Sea una cadena de Markov con la siguiente matriz de transición:

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

cuyo diagrama de estados es



Observando este grafo, se ve que todos los estados tienen periodo 3. Por ejemplo, si se empieza en E_1, entonces los regresos a ese estado sólo se producen en los pasos $3, 6, 9, \ldots$

Se puede comprobar la periodicidad en términos de la matriz de transición. Sea

$$S = T^3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

En este ejemplo, resulta que

$$S^2 = T^6 = SS = S,$$

de modo que

$$S^r = T^{3r} = S$$

para $r = 1, 2, \ldots$ y siempre tiene elementos distintos de 0 en la diagonal.

Por otro lado

$$T^{3r+1} = S^r T = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$T^{3r+2} = S^r T^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 & 0 \end{pmatrix},$$

y ambas matrices tienen elementos igual a 0 en la diagonal para $r = 1, 2, \ldots$

Así, para $i = 1, 2, 3, 4$

$$\begin{aligned} p_{ii}^{(n)} &= 0 \qquad n \neq 3, 6, 9, \ldots \\ p_{ii}^{(n)} &\neq 0 \qquad n = 3, 6, 9, \ldots \end{aligned}$$

lo que implica que los estados tienen periodo 3.

Estado recurrente

Denominamos como $f_j^{(n)}$ la probabilidad de que la primera visita al estado E_j ocurra en la etapa n. Esta probabilidad no es la misma que $p_{jj}^{(n)}$, que es la probabilidad de que se produzca un retorno en el n-ésimo paso y esto incluye a los posibles retornos en los pasos $1, 2, 3, \ldots, n-1$ también. Se deduce que

$$\begin{aligned} p_{jj}^{(1)} &= f_j^{(1)} \\ p_{jj}^{(2)} &= f_j^{(2)} + f_j^{(1)} p_{jj}^{(1)} \\ p_{jj}^{(3)} &= f_j^{(3)} + f_j^{(1)} p_{jj}^{(2)} + f_j^{(2)} p_{jj}^{(1)}, \end{aligned}$$

es decir, la probabilidad de un retorno en el paso 3 es igual a la probabilidad de un primer retorno en el paso 3 ó la probabilidad de un primer retorno en el primer paso y un retorno dos pasos después ó la probabilidad de un retorno en el segundo paso y un retorno un paso después.

Así, en general

$$p_{jj}^{(n)} = f_j^{(n)} + \sum_{r=1}^{n-1} f_j^{(r)} p_{jj}^{(n-r)}.$$

Se puede expresar en términos de $f_j^{(n)}$

$$\begin{aligned} f_j^{(1)} &= p_{jj}^{(1)} \\ f_j^{(n)} &= p_{jj}^{(n)} - \sum_{r=1}^{n-1} f_j^{(r)} p_{jj}^{(n-r)}. \end{aligned}$$

La probabilidad de regresar en algún paso al estado E_j es

$$f_j = \sum_{n=1}^{\infty} f_j^{(n)},$$

si $f_j = 1$, entonces seguro que se regresa a E_j y se denomina a E_j *estado recurrente.*

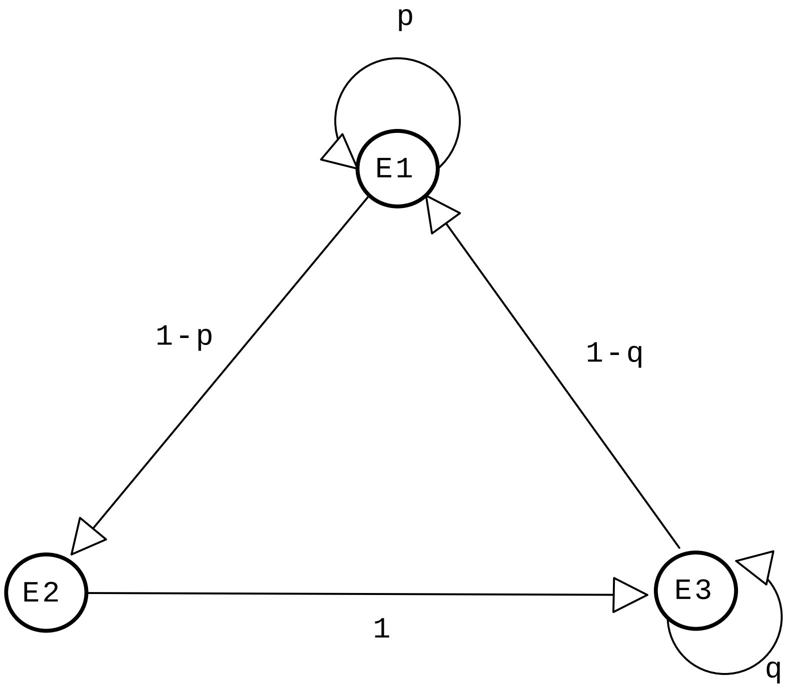
Ejemplo.

Una cadena de Markov con tres estados tiene como matriz de transición

$$T = \begin{pmatrix} p & 1-p & 0 \\ 0 & 0 & 1 \\ 1-q & 0 & q \end{pmatrix},$$

donde $0 < p < 1$ y $0 < q < 1$. Se puede comprobar que E_1 es recurrente.

Si se considera el grafo asociado



se observa que si se empieza en E_1 el primer regreso a este estado se puede hacer en todos los pasos excepto para $n = 2$, ya que después de dos pasos la cadena debe estar en E_3. A partir de la figura se puede observar que

$$\begin{aligned} f_1^{(1)} &= p \\ f_1^{(2)} &= 0 \\ f_1^{(3)} &= (1-p) \cdot 1 \cdot (1-q) \end{aligned}$$

y, en general $(n \geq 3)$,

$$f_1^{(n)} = (1-p) \cdot 1 \cdot q^{n-3} \cdot (1-q).$$

Para $n \geq 3$, la secuencia que da $f_1^{(n)}$ es

$$E_1 \quad E_2 \quad \overbrace{E_3 E_3 \cdots E_3}^{(n-3)\text{ veces}} \quad E_1$$

La probabilidad f_1 de que el sistema regrese al menos una vez al estado E_1 es

$$f_1 = \sum_{n=1}^{\infty} f_1^{(n)} = p + \sum_{n=3}^{\infty} (1-p)(1-q)q^{n-3}$$

si se hace el cambio $s = n - 3$, entonces

$$\begin{aligned} f_1 &= p + (1-p)(1-q)\sum_{s=0}^{\infty} q^s = \\ &= p + (1-p)(1-q)\frac{1}{1-q} = 1 \end{aligned}$$

donde se ha usado la fórmula de la suma de los elementos de una serie geométrica.

Así, $f_1 = 1$ y el estado E_1 es recurrente.

En los estados recurrentes se puede estudiar, también, el tiempo medio de recurrencia μ de un estado recurrente E_j, el cual, cuando

$$f_j = \sum_{n=1}^{\infty} f_j^{(n)} = 1,$$

se expresa como

$$\mu_j = \sum_{n=1}^{\infty} n f_j^{(n)}.$$

En el ejemplo, el estado E_1 es recurrente con un tiempo medio de recurrencia dado por

$$\begin{aligned} \mu_1 &= \sum_{n=1}^{\infty} n f_1^{(n)} = p + (1-p)(1-q)\sum_{n=3}^{\infty} n q^{n-3} = \\ &= \frac{3 - 2q - 2p + pq}{1-q} \end{aligned}$$

ya que

NOTA:

$\sum_{n=3}^{\infty} nq^{n-3} = \sum_{s=0}^{\infty}(s+3)q^s = \frac{3}{1-q} + \sum_{s=0}^{\infty} sq^s$ (por ser $\sum_{s=0}^{\infty} q^s = \frac{1}{1-q}$)

Derivando ambos términos de $\sum_{s=0}^{\infty} q^s = \frac{1}{1-q}$, se obtiene,

$\sum_{s=1}^{\infty} sq^{s-1} = \frac{1}{(1-q)^2}$, es decir, $\sum_{s=0}^{\infty} sq^s = \frac{q}{(1-q)^2}$, luego

$\sum_{n=3}^{\infty} nq^{n-3} = \sum_{s=0}^{\infty}(s+3)q^s = \frac{3}{1-q} + \sum_{s=0}^{\infty} sq^s = \frac{3}{1-q} + \frac{q}{(1-q)^2}$.

Sustituyendo en la expresión general,

$p + (1-p)(1-q)\sum_{n=3}^{\infty} nq^{n-3} = p + (1-p)(1-q)\left(\frac{3}{1-q} + \frac{q}{(1-q)^2}\right) =$

$=\frac{p-pq+3(1-p)(1-q)+q-pq}{1-q}$

Simplificando términos, esto es igual a $\frac{p-2pq+3-3q-3p+3pq+q}{1-q} = \frac{3-2p-2q+pq}{1-q}$.

La expresión anterior es finita en este caso, aunque no tiene por qué ser siempre así.

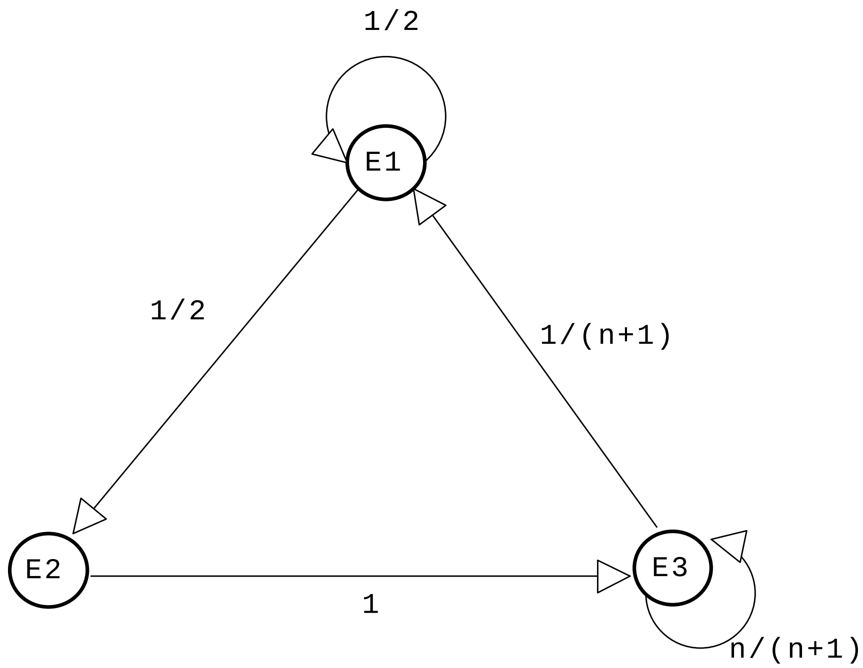
Un estado recurrente se dice que es *nulo*, si $\mu_j = \infty$ y es no nulo en caso contrario. Para obtener una cadena con estados recurrentes nulos basta hacer que las probabilidades de transición dependan del número de pasos n. En el caso de probabilidades de transición constantes no aparecen estados recurrentes nulos.

Ejemplo.

Supongamos una cadena con la siguiente matriz de transición

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ \frac{1}{n+1} & 0 & \frac{n}{n+1} \end{pmatrix},$$

cuyo diagrama es



Se tiene que

$$\begin{aligned}
f_1^{(1)} &= \frac{1}{2}\\
f_1^{(2)} &= 0\\
f_1^{(3)} &= \frac{1}{2}\cdot 1\cdot\frac{1}{4}
\end{aligned}$$

en general, para $n \geq 4$,

$$f_1^{(n)} = \frac{1}{2}\cdot 1\cdot\frac{3}{4}\cdot\frac{4}{5}\cdots\frac{n-1}{n}\cdot\frac{1}{n+1} = \frac{3}{2n(n+1)}.$$

Así,

$$f_1 = \frac{1}{2}+\frac{1}{8}+\frac{3}{2}\sum_{n=4}^{\infty}\frac{1}{n(n+1)}.$$

Se tiene que como

$$\frac{1}{n(n+1)} = \frac{1}{n}-\frac{1}{n+1}$$

$$\sum_{n=4}^{\infty}\frac{1}{n(n+1)} = \lim_{N\to\infty}\sum_{n=4}^{N}\left(\frac{1}{n}-\frac{1}{n+1}\right) = \lim_{N\to\infty}\left(\frac{1}{4}-\frac{1}{N+1}\right) = \frac{1}{4},$$

entonces

$$f_1 = \frac{1}{2}+\frac{1}{8}+\frac{3}{2}\frac{1}{4} = 1,$$

lo que significa que E_1 es un estado recurrente.

Por otro lado, el tiempo medio de recurrencia

$$\begin{aligned}
\mu_1 &= \sum_{n=1}^{\infty} n f_1^{(n)} = \frac{1}{2}\cdot 1+\frac{1}{8}\cdot 3+\frac{3}{2}\sum_{n=4}^{\infty}\frac{n}{n(n+1)} =\\
&= \frac{7}{8}+\frac{3}{2}\sum_{n=4}^{\infty}\frac{1}{(n+1)}\\
&= \frac{7}{8}+\frac{3}{2}\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\cdots\right) =\\
&= \frac{7}{8}+\frac{3}{2}\sum_{n=5}^{\infty}\frac{1}{n},
\end{aligned}$$

pero $\sum_{n=5}^{\infty}\frac{1}{n}$ es la llamada serie armónica (menos los cuatro primeros términos) que es divergente. Así $\mu_1 = \infty$, con lo cual E_1 es recurrente y nulo.

Los estados pueden ser a la vez recurrentes y periódicos.

Probabilidad de primera pasada en general

Se puede generalizar el concepto de probabilidad de primera pasada, calculando la probabilidad de pasar de un estado i a otro estado j en n etapas por primera vez: $f_{ij}^{(n)}$ sin que en las etapas intermedias se haya pasado antes por j :

$$f_{ij}^{(n)} = P\left(X_n = j, X_r \neq j \mid X_0 = i\right),$$

$\forall r = 1, 2, \ldots, n-1$ donde $n = 1, 2, 3, \ldots$

– Para $n = 1$,

$$f_{ij}^{(1)} = P\left(X_1 = j \mid X_0 = i\right) = p_{ij}.$$

– Para $n = 2$,

$$\begin{aligned} f_{ij}^{(2)} &= P\left(X_2 = j, X_1 \neq j \mid X_0 = i\right) = \\ &= \sum_{k \neq j} P\left(X_1 = k \mid X_0 = i\right) \cdot P\left(X_2 = j \mid X_1 = k\right) = \sum_{k \neq j} p_{ik} p_{kj} = \\ &= \sum_{k \neq j} p_{ik} f_{kj}^{(1)}. \end{aligned}$$

– Para $n > 2$,

$$f_{ij}^{(n)} = \sum_{k \neq j} P\left(X_1 = k \mid X_0 = i\right) \cdot f_{kj}^{(n-1)} = \sum_{k \neq j} p_{ik} f_{kj}^{(n-1)}.$$

Se puede observar que para calcular $f_{ij}^{(n)}$ se necesita calcular de modo iterativo $f_{kj}^{(s)}$ con $s < n$ para todo $k \neq j$, y se puede expresar la última igualdad como el producto de la j-ésima coordenada de la matriz T por el vector de los $f_{\cdot j}^{(n-1)}$:

$$f_{ij}^{(n)} = \begin{pmatrix} p_{i1} & \ldots & p_{i,j-1} & 0 & p_{i,j+1} & \ldots \end{pmatrix} \begin{pmatrix} f_{1j}^{(n-1)} \\ \vdots \\ f_{j-1j}^{(n-1)} \\ f_{jj}^{(n-1)} \\ f_{j+1,j}^{(n-1)} \\ \vdots \end{pmatrix}$$

En general, se tiene que

$$f_{\cdot j}^{(n)} = \begin{pmatrix} p_{11} & \dots & p_{1,j-1} & 0 & p_{1,j+1} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p_{i1} & \dots & p_{i,j-1} & 0 & p_{i,j+1} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} f_{1j}^{(n-1)} \\ \vdots \\ f_{j-1j}^{(n-1)} \\ f_{jj}^{(n-1)} \\ f_{j+1,j}^{(n-1)} \\ \vdots \end{pmatrix}$$

es decir, con la columna j-ésima de la matriz de transición sustituida por una columna de ceros.

También, se puede generalizar f_j cuando se accede a j desde un estado i cualquiera:

$$f_{ij} = \sum_{n=0}^{\infty} f_{ij}^{(n)}.$$

Estado transitorio

En un estado recurrente, la probabilidad de que se regrese por primera vez a ese estado en algún paso es 1, pero para otros estados sucede que

$$f_j = \sum_{n=1}^{\infty} f_j^{(n)} < 1$$

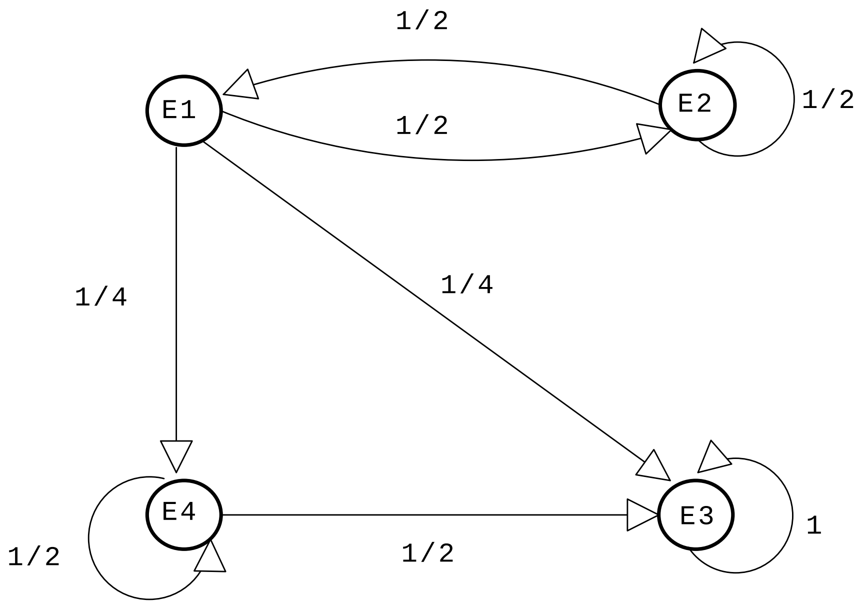
lo que significa es que no se regresa al estado E_j de modo seguro. Un estado así se denomina *transitorio.*

Ejemplo.

Supongamos una cadena con la siguiente matriz de transición

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

cuyo diagrama es



El estado E_1 es transitorio ya que,

$$
\begin{aligned}
f_1^{(1)} &= 0 \\
f_1^{(2)} &= \frac{1}{2} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^2 \\
f_1^{(3)} &= \left(\frac{1}{2}\right)^3 \\
& \dots \\
f_1^{(n)} &= \left(\frac{1}{2}\right)^n,
\end{aligned}
$$

así,

$$f_1 = \sum_{n=1}^{\infty} f_1^{(n)} = \sum_{n=2}^{\infty} \left(\frac{1}{2}\right)^n = \frac{1}{1-\frac{1}{2}} - \frac{1}{2} - 1 = \frac{1}{2} < 1$$

luego E_1 es un estado transitorio. De hecho no se puede acceder a E_1 desde E_3 ó E_4.

Estado ergódico

Un estado bastante importante es aquel que es recurrente, no nulo y aperiódico. Recibe el nombre de *ergódico*. Los estados ergódicos son importantes en la clasificación de cadenas y para probar la existencia de distribuciones de probabilidad límite.

Ejemplo.

En la cadena de Markov ya estudiada con tres estados y matriz de transición

$$T = \begin{pmatrix} p & 1-p & 0 \\ 0 & 0 & 1 \\ 1-q & 0 & q \end{pmatrix},$$

donde $0 < p < 1$ y $0 < q < 1$. Se puede comprobar que E_1 es ergódico.

Se tenía que

$$\begin{aligned} f_1^{(1)} &= p \\ f_1^{(2)} &= 0 \\ f_1^{(3)} &= (1-p) \cdot 1 \cdot (1-q) \end{aligned}$$

y, en general ($n \geq 3$),

$$f_1^{(n)} = (1-p) \cdot (1-q) \cdot q^{n-3}.$$

de modo que E_1 era recurrente.

Se tiene que el tiempo de recurrencia medio es

$$\mu_1 = \sum_{n=1}^{\infty} n f_1^{(n)} = p + (1-p) \cdot (1-q) \sum_{n=3}^{\infty} n q^{n-3}$$

Como $\sum_{n=3}^{\infty} n q^{n-3} = \frac{3}{1-q} + \frac{q}{(1-q)^2}$ como se vio antes, y sustituyendo

$$\begin{aligned} \mu_1 &= p + (1-p) \cdot (1-q) \left(\frac{3}{1-q} + \frac{q}{(1-q)^2} \right) = \\ &= p + 3(1-p) + \frac{q(1-p)}{1-q} < \infty \end{aligned}$$

La convergencia de μ_1 implica que E_1 es no nula, además se demuestra fácilmente que los elementos diagonales $p_{ii}^{(n)} > 0$ para $n \geq 3$ e $i = 1, 2, 3$, lo que significa que E_1 es aperiódico. De este modo por la definición establecida antes E_1 es ergódica.

Clasificación de cadenas

En esta sección se definen propiedades de las cadenas que, en realidad, son propiedades comunes de los estados de la cadena.

Cadenas irreducibles

Una cadena irreducible es aquella en la que todos los estados son alcanzables desde cualquier otro estado de la cadena en un número finito de pasos. Eso implica que se puede llegar a cualquier estado E_j desde otro estado E_i esto es $p_{ij}^{(n)} > 0$, para algún número entero n.

Una matriz $A = [a_{ij}]$ se dice que es positiva si $a_{ij}^{(n)} > 0$ para todos los i, j.

Una matriz de transición T se dice que es regular si existe un número entero N tal que T^N es positivo.

Una cadena regular obviamente es irreducible, sin embargo, lo contrario no tiene por qué ser necesariamente cierto.

Ejemplo.

Suponemos la siguiente matriz de transición de una cadena irreducible

$$T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Como

$$T^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$

y

$$T^{2n+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = T$$

para $n = 1, 2, 3, \ldots$ entonces ninguna potencia de T es una matriz positiva y por tanto no es una cadena regular.

Ejemplo.

Se puede ver que una cadena en la tercera etapa, con matriz de transición

$$T = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

define una cadena regular y, así, una cadena irreducible también.

$$T^2 = \begin{pmatrix} \frac{4}{9} & \frac{1}{9} & \frac{4}{9} \\ 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

y

$$T^3 = \begin{pmatrix} \frac{16}{27} & \frac{4}{27} & \frac{7}{27} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{4}{9} & \frac{1}{9} & \frac{4}{9} \end{pmatrix}$$

Así, T^3 es una matriz positiva, lo que significa que la cadena es regular.

Una propiedad muy importante de las cadenas irreducibles es que todos sus estados son del mismo tipo, esto es, o bien todos son transitorios o bien todos son recurrentes (nulos o no nulos) y todos tienen el mismo periodo. Esto significa que la clasificación de todos los estados de una cadena se puede deducir a partir de la clasificación conocida de uno de los estados.

También es obvio que todos los estados de una cadena finita irreducible no pueden ser transitorios, ya que eso significaría que el regreso a alguno de los estados no sería seguro, aunque todos los estados fueran accesibles desde cualquiera de ellos en un número finito de pasos.

Conjuntos cerrados

Una cadena de Markov puede contener algunos estados que sean recurrentes, otros que sean transitorios y otros que sean absorbentes. Los estados recurrentes pueden ser parte de subcadenas cerradas.

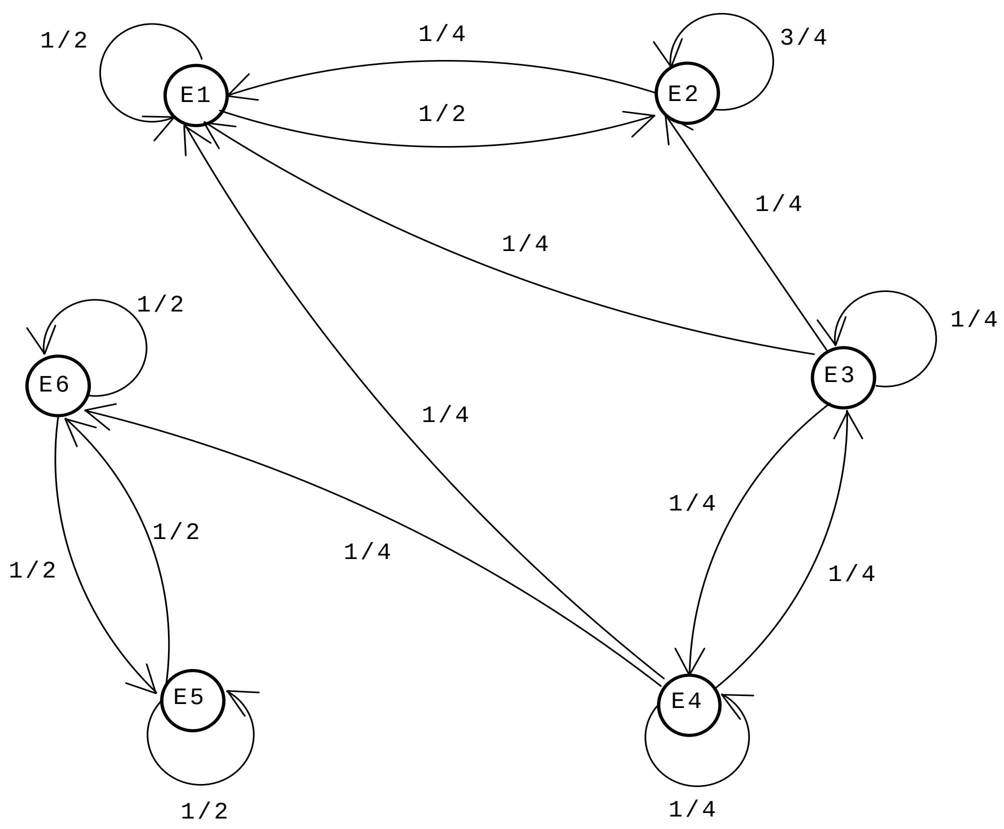
Un conjunto de estados C en una cadena de Markov se dice que es cerrado si cualquier estado dentro de C puede alcanzarse desde cualquier otro estado de C y ningún otro estado fuera de C puede ser alcanzado desde cualquier estado dentro de C. Así una condición necesaria para que esto ocurra es que

$$p_{ij} = 0 \qquad \forall E_i \in C, \ \forall E_j \notin C$$

Los estados absorbentes son cerrados con sólo un elemento. Se puede ver que un subconjunto cerrado es él mismo una subcadena irreducible de una cadena de Markov completa.

Ejemplo.

Supongamos la siguiente cadena cuyo diagrama es



y la matriz de transición asociada es

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

Observando el diagrama se tiene que el conjunto $\{E_1, E_2\}$ forman una subcadena cerrada e irreducible ya que ningún estado fuera de E_1 y E_2 se puede alcanzar desde ellos. De modo similar $\{E_5, E_6\}$ forman una subcadena cerrada e irreducible. Los estados E_3 y

E_4 son transitorios. Todos los estados son aperiódicos, lo que significa que E_1, E_2, E_5 y E_6 son ergódicos.

Cadenas ergódicas

Se tenía que todos los estados en una cadena irreducible pertenecen a la misma clase. Si todos los estados son ergódicos, esto es, recurrentes, no nulos y aperiódicos entonces se define la cadena como ergódica.

Ejemplo.

Supongamos la siguiente cadena cuya matriz de transición es

$$T = \begin{pmatrix} \frac{1}{5} & \frac{4}{5} & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Los autovalores de la matriz son

$$\lambda_1 = 1 \quad \lambda_2 = -\tfrac{2}{5} + \tfrac{4}{5}i \quad \lambda_3 = -\tfrac{2}{5} - \tfrac{4}{5}i$$

y la matriz de autovectores es

$$C = \begin{pmatrix} 1 & -\frac{12}{25} + \frac{16}{25}i & -\frac{12}{25} - \frac{16}{25}i \\ 1 & -\frac{2}{5} - \frac{4}{5}i & -\frac{2}{5} + \frac{4}{5}i \\ 1 & 1 & 1 \end{pmatrix}$$

y su inversa es

$$C^{-1} = \begin{pmatrix} \frac{5}{13} & \frac{4}{13} & \frac{4}{13} \\ -\frac{5}{26} - \frac{35}{104}i & -\frac{2}{13} + \frac{37}{104}i & \frac{9}{26} - \frac{1}{52}i \\ -\frac{5}{26} + \frac{35}{104}i & -\frac{2}{13} - \frac{37}{104}i & \frac{9}{26} + \frac{1}{52}i \end{pmatrix}$$

Para calcular la distribución límite

$$\begin{aligned} Q &= \lim_{n\to\infty} T^n = C \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} C^{-1} = \\ &= \begin{pmatrix} \frac{5}{13} & \frac{4}{13} & \frac{4}{13} \\ \frac{5}{13} & \frac{4}{13} & \frac{4}{13} \\ \frac{5}{13} & \frac{4}{13} & \frac{4}{13} \end{pmatrix}. \end{aligned}$$

entonces como

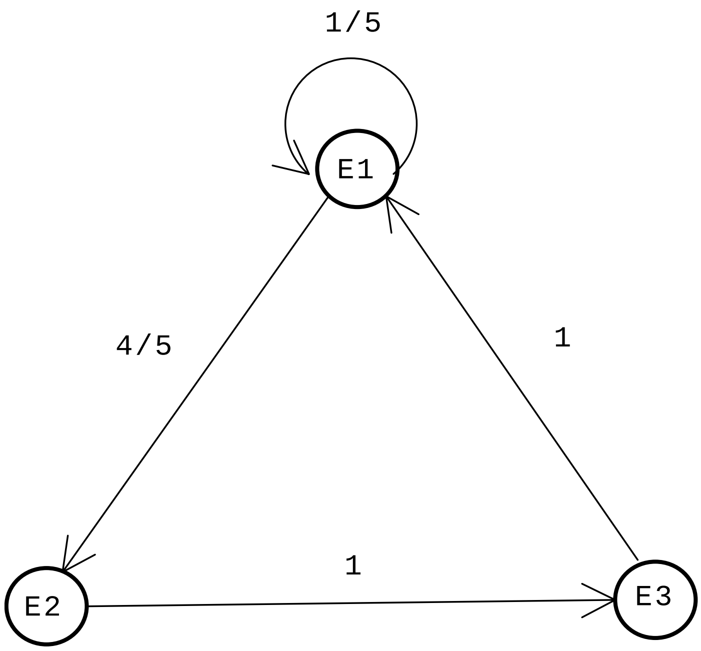
$$\mathbf{p}^{(n)} = \mathbf{p}^{(0)} T^n,$$

la distribución límite es

$$\begin{aligned}
\mathbf{p} &= \lim_{n\to\infty} \mathbf{p}^{(n)} = \lim_{n\to\infty} \mathbf{p}^{(0)} T^n = \mathbf{p}^{(0)} Q = \\
&= \begin{pmatrix} p_1^{(0)} & p_2^{(0)} & p_3^{(0)} \end{pmatrix} \begin{pmatrix} \frac{5}{13} & \frac{4}{13} & \frac{4}{13} \\ \frac{5}{13} & \frac{4}{13} & \frac{4}{13} \\ \frac{5}{13} & \frac{4}{13} & \frac{4}{13} \end{pmatrix} = \\
&= \begin{pmatrix} \frac{5}{13} & \frac{4}{13} & \frac{4}{13} \end{pmatrix}.
\end{aligned}$$

El vector $\mathbf{p}$ da las probabilidades de los estados a largo plazo, y esta distribución es independiente del estado inicial $\mathbf{p}^{(0)}$.

Si consideramos, a continuación, el diagrama de estados



La probabilidad de primer retorno para cada uno de los estados $f_i^{(n)}$ se puede calcular observando el diagrama de estados previo.

Así,

$$\begin{aligned}
f_1^{(1)} &= \frac{1}{5} \\
f_1^{(2)} &= 0 \\
f_1^{(3)} &= \frac{4}{5} \cdot 1 \cdot 1 = \frac{4}{5} \\
f_1^{(n)} &= 0, \qquad n \geq 4
\end{aligned}$$

Del mismo modo,

$$\begin{aligned}
f_2^{(1)} &= f_3^{(1)} = 0 \\
f_2^{(2)} &= f_3^{(2)} = 0 \\
f_2^{(n)} &= f_3^{(n)} = 1 \cdot 1 \cdot \left(\frac{1}{5}\right)^{n-3} \cdot \frac{4}{5} = \frac{4}{5}\left(\frac{1}{5}\right)^{n-3}, \qquad n \geq 3.
\end{aligned}$$

Para calcular los tiempos de recurrencia medios, en cada caso,

$$\mu_1 = \sum_{n=1}^{\infty} n f_1^{(n)} = \frac{1}{5} + 3\frac{4}{5} = \frac{13}{5}.$$

$$\mu_2 = \mu_3 = \sum_{n=1}^{\infty} n f_2^{(n)} = \frac{4}{5}\sum_{n=1}^{\infty} n \left(\frac{1}{5}\right)^{n-3} \overset{(*)}{=} \frac{4}{5} \cdot \frac{65}{16} = \frac{13}{4}$$

(*) donde se ha empleado la misma relación que en los ejemplos anteriores:

$$\begin{aligned}
\sum_{n=3}^{\infty} n q^{n-3} &= \frac{3}{1-q} + \frac{q}{(1-q)^2} = \\
&= \frac{3}{1-\frac{1}{5}} + \frac{\frac{1}{5}}{\left(1-\frac{1}{5}\right)^2} = \frac{65}{16}.
\end{aligned}$$

Se observa que el vector de recíprocos

$$\begin{pmatrix} \frac{1}{\mu_1} & \frac{1}{\mu_2} & \frac{1}{\mu_3} \end{pmatrix} = \begin{pmatrix} \frac{5}{13} & \frac{4}{13} & \frac{4}{13} \end{pmatrix}$$

es igual que el vector $\mathbf{p}$ calculado mediante los autovalores y autovectores.

Para cadenas ergódicas se obtiene siempre que la distribución invariante es el recíproco del vector de tiempos medios de recurrencia.

Representación Canónica de una matriz de transición

Se trata de reordenar la matriz de transición separando los estados recurrentes de los transitorios. A su vez, los estados recurrentes se tienen que reordenar juntando aquellos que se comunican entre sí.

Los estados de una cadena se puede dividir en dos subconjuntos disjuntos (alguno puede ser $\varnothing$): uno está formado por los estados transitorios y el otro por los recurrentes, de modo que los estados transitorios son inaccesibles desde los estados recurrentes.

Los estados recurrentes se pueden subdividir de manera única en conjuntos cerrados e irreducibles. Dentro de cada conjunto cerrado todos los estados se comunican y son del mismo tipo (recurrentes positivos o recurrentes nulos) y con el mismo periodo en su caso.

Si dos estados están en distintos conjuntos cerrados, no se pueden comunicar. Además, si dos estados i y j pertenecen al mismo conjunto cerrado e irreducible entonces

$$f_{ij} = f_{ji} = 1$$

y si están en conjuntos distintos entonces

$$f_{ij} = f_{ji} = 0.$$

De acuerdo con este resultado se puede hacer una reordenación de los estados, y así de las filas y columnas de la matriz de transición, colocando los estados transitorios en último lugar.

$$P = \begin{pmatrix} P_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & P_2 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & P_r & \cdots & 0 \\ Q_1 & Q_2 & \cdots & Q_r & \cdots & Q \end{pmatrix}$$

donde las submatrices P_i $(i = 1, \ldots, r)$ están asociadas a subcadenas recurrentes cerradas, Q_i $(i = 1, \ldots, r)$ dan la probabilidad de pasar a los estados recurrentes desde los transitorios y Q da la probabilidad de pasar de transitorios a transitorios.

Cadenas Infinitas

Consideramos que la cadena no es finita y que el proceso continua de manera indefinida.

Se entiende por visita a un estado j como la etapa o momento en el que la cadena se encuentra en el estado j, y se denota por N_j el número de visitas que se hacen al estado j en el proceso.

Distribución de N_j

Los posibles valores que puede tomar N_j son $0, 1, 2, 3, \ldots$

Así, se tienen dos casos:

(i)

$$P\left(N_j = m \mid X_0 = i\right) = \begin{cases} f_{ij} f_{jj}^{m-1} \left(1 - f_{jj}\right) & \text{si } m = 1, 2, 3 \ldots \\ 1 - f_{ij} & \text{si } m = 0 \end{cases}$$

donde $f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$.

Equivale a la probabilidad de ir de i a j una vez, y de ir a j $(m-1)$ veces, y la de no volver nunca a j.

(ii)

$$P\left(N_j = m \mid X_0 = j\right) = \begin{cases} f_{jj}^{m-1} \left(1 - f_{jj}\right) & \text{si } m = 1, 2, 3 \ldots \\ 0 & \text{si } m = 0 \end{cases}$$

dado que se parte ya de una visita al estado j.

Si j es un estado recurrente, entonces $f_{jj} = 1$ y

$$P\left(N_j = m \mid X_0 = i\right) = \begin{cases} 1 - f_{ij} & \text{si } m = 0 \\ 0 & \text{si } m = 1, 2, 3 \ldots \\ f_{ij} & \text{si } m = \infty \end{cases}$$

$$P\left(N_j = m \mid X_0 = j\right) = \begin{cases} 0 & \text{si } m = 0, 1, 2, 3 \ldots \\ 1 & \text{si } m = \infty \end{cases}$$

es decir, si j es recurrente la probabilidad de que partiendo de i se visite j es f_{ij} y el número de veces que se visita es infinito, pues una vez en j se vuelve a ese estado infinitas veces.

Número medio de visitas la estado j

Se demuestra que calculando la esperanza de la variable N_j se obtiene, por estar relacionada con la distribución geométrica, que

$$E[N_j \mid X_0 = i] = \frac{f_{ij}}{1 - f_{jj}}$$

y del mismo modo

$$E[N_j \mid X_0 = j] = \frac{1}{1 - f_{jj}}$$

Se denota como

$$R_{ij} = \frac{f_{ij}}{1 - f_{jj}}$$

al número medio de visitas a j partiendo de i y

$$R_{jj} = \frac{1}{1 - f_{jj}}$$

al número medio de visitas a j partiendo de j.

Así, si j es recurrente, $f_{jj} = 1$ de manera que $R_{jj} = \infty$

Si j es transitorio entonces $f_{jj} < 1$ y $R_{jj} < \infty$, es decir, en algún momento dado ya no se vuelve a j. Así, en realidad, el cálculo de R_{jj} tiene interés para un estado transitorio.

Proposición

Se cumple que

$$R_{ij} = \sum_{n=0}^{\infty} T_{ij}^{(n)}$$

Demostración

Fijado el estado j, se define un *contador* de visitas como

$$N_j^n = \begin{cases} 1 & \text{si } X_n = j \\ 0 & \text{si } X_n \neq j \end{cases}$$

de este modo,

$$N_j = \sum_{n=0}^{\infty} N_j^n$$

entonces

$$E\left[N_j^n \mid X_0 = i\right] = 1 \cdot P\left(X_n = j \mid X_0 = i\right) = T_{ij}^{(n)}$$

con lo que

$$\begin{aligned} E\left[N_j \mid X_0 = i\right] &= E\left[\sum_{n=0}^{\infty} N_j^n \mid X_0 = i\right] = \\ &= \sum_{n=0}^{\infty} E\left[N_j^n \mid X_0 = i\right] = \sum_{n=0}^{\infty} T_{ij}^{(n)}. \end{aligned}$$

Si se denomina R a la matriz $\left(R_{ij}\right)_{i,j\in E}$ entonces se tiene que

$$R = I + T + T^2 + \cdots = \sum_{n=0}^{\infty} T^n.$$

Comportamiento en el límite de una cadena de Markov

Se trata de calcular la probabilidad de visitar un estado j desde un estado i (f_{ij}) teniendo en cuenta cómo sean los estados.

(i) Si i y j son ambos recurrentes entonces $f_{ij} = 1$ cuando ambos pertenecen al mismo conjunto cerrado, ó $f_{ij} = 0$ cuando pertenecen a conjuntos cerrados diferentes.

(ii) Si i es recurrente y j es transitorio entonces $f_{ij} = 0$, ya que no se puede acceder a ningún estado transitorio desde uno recurrente.

(iii) Si i y j son ambos transitorios, entonces, como se tenía que $R_{ij} = \frac{f_{ij}}{1-f_{jj}}$ y $R_{jj} = \frac{1}{1-f_{jj}}$, despejando se obtiene que

$$\begin{aligned} f_{ij} &= \frac{R_{ij}}{R_{jj}} \\ f_{jj} &= 1 - \frac{1}{R_{jj}} \end{aligned}$$

Cálculo de R_{ij}

Se parte de la descomposición canónica de la matriz de transición

$$P = \begin{pmatrix} K & 0 \\ L & Q \end{pmatrix}$$

como $R = I + P + P^2 + \cdots = \sum_{n=0}^{\infty} P^n$, entonces

$$P^n = \begin{pmatrix} K^n & 0 \\ L_n & Q^n \end{pmatrix}$$

donde $L_n \neq L^n$, entonces

$$R = \sum_{n=0}^{\infty} P^n = \begin{pmatrix} \sum_{n=0}^{\infty} K^n & 0 \\ \sum_{n=0}^{\infty} L_n & \sum_{n=0}^{\infty} Q^n \end{pmatrix}$$

es la matriz que proporciona el número medio de visitas a cada estado desde otro estado.

Si sólo interesa considerar estados i, j transitorios, basta tomar la matriz

$$\begin{aligned} S &= \sum_{n=0}^{\infty} Q^n = \\ &= (I - Q)^{-1}. \end{aligned}$$

De este modo, se enuncia la siguiente proposición.

Proposición Dados $i, j \in E$ transitorios, la probabilidad de pasar de i a j es

$$\begin{aligned} f_{ij} &= \frac{S_{ij}}{S_{jj}} \\ f_{jj} &= 1 - \frac{1}{S_{jj}} \end{aligned}$$

donde $S = (I - Q)^{-1}$, siendo Q la submatriz correspondiente a los estados transitorios en la representación canónica

Caso en que i es transitorio y j recurrente Se tiene que todo estado recurrente j se encuentra en un conjunto cerrado e irreducible. De este modo, si i es transitorio y j pertenece a un conjunto cerrado e irreducible de estados recurrentes, entonces

$$\forall j, k \in C \Longrightarrow f_{ij} = f_{ik}$$

ya que

$$\forall j, k \in C \Longrightarrow f_{jk} = f_{kj} = 1$$

Debido a esto, se puede considerar la siguiente definición:

Definición

Se llama probabilidad de absorción de un estado transitorio i por un conjunto cerrado e irreducible C, a la probabilidad f_{ij}, $\forall j \in C$.

En realidad, sólo es necesario calcular las probabilidades de absorción, y los conjuntos cerrados e irreducibles presentes en una cadena se toman como estados absorbentes. Así de la matriz canónica P con r conjuntos cerrados irreducibles y m estados transitorios se pasa a otra matriz reducida

$$P = \begin{pmatrix} P_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & P_2 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & P_r & \cdots & 0 \\ Q_1 & Q_2 & \cdots & Q_r & \cdots & Q \end{pmatrix}$$

$$\bar{P} = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ b_1 & b_2 & \cdots & b_r & \cdots & Q \end{pmatrix} = \begin{pmatrix} \begin{matrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{matrix} & 0 \\ B & Q \end{pmatrix}$$

de modo que se considera cada conjunto cerrado irreducible como un único estado absorbente y las matrices Q_i pasan a ser vectores columna b_i que dan las probabilidades de transición de los estados transitorios al conjunto cerrado correspondiente i (y así a todos sus correspondientes estados recurrentes).

De este modo se define

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1r} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mr} \end{pmatrix}$$

siendo, en cada caso,

$$b_{ij} = \sum_{k \in C_j} P_{ik}$$

$\forall i \in E$ transitorio.

Así se suman las probabilidades de transición en una etapa, del estado transitorio i a todos los estados recurrentes que están contenidos en el conjunto cerrado correspondiente C_j.

Proposición

Sea Q la matriz obtenida a partir de la representación de la matriz canónica P, eliminando las filas y columnas pertenecientes a los estados recurrentes, y sea B la matriz construida a partir de los vectores columna b_j, donde $j = 1, \ldots, r$, entonces la matriz

$$G = (I - Q)^{-1}B = S \cdot B$$

recoge las probabilidades de absorción de los estados transitorios por los conjuntos cerrados e irreducibles.

Se demuestra calculando $\bar{P}^n$ y haciendo $n \to \infty$.

Así, $G_{ij} = f_{ik}\ \forall k \in C_j$, donde i es transitorio.

El siguiente concepto que es interesante de considerar, es el tiempo que se tarda en ser absorbido.

Definición

Se define el tiempo medio de absorción desde un estado transitorio i, como el número medio de etapas que, comenzando en i, la cadena recorrerá antes de pasar a cualquier conjunto cerrado e irreducible.

Este número de etapas es la suma de las visitas que se hacen a todos los estados transitorios, es decir

$$A_i = \sum_{j \in D} R_{ij}$$

donde D es el conjunto de los estados transitorios.

Como la matriz fundamental de la representación canónica S recoge el número medio de visitas a estados transitorios desde estados transitorios, basta fijar la fila i y sumar en j los elementos S_{ij} cuando $j \in D$. Así, A_i se calcula sumando los elementos de la fila $i-$ésima de la matriz fundamental S.

Se puede generalizar esta idea, y plantearse calcular el número medio de etapas que hay que recorrer antes de visitar un estado transitorio j por primera vez. En este caso bastará con sumar sólo aquellos R_{ik} donde $k \neq j$, es decir si se impone la visita a j por primera vez, el número medio de etapas será

$$\sum_{k \neq j} R_{ik}$$

donde $k \in D$.

Cadenas periódicas

Si un estado j es periódico con periodo δ y otro estado i comunica con él, entonces el estado i también es periódico con el mismo periodo. De este modo, el periodo es una característica común del conjunto irreducible y cerrado y se puede hablar del periodo de una subcadena irreducible.

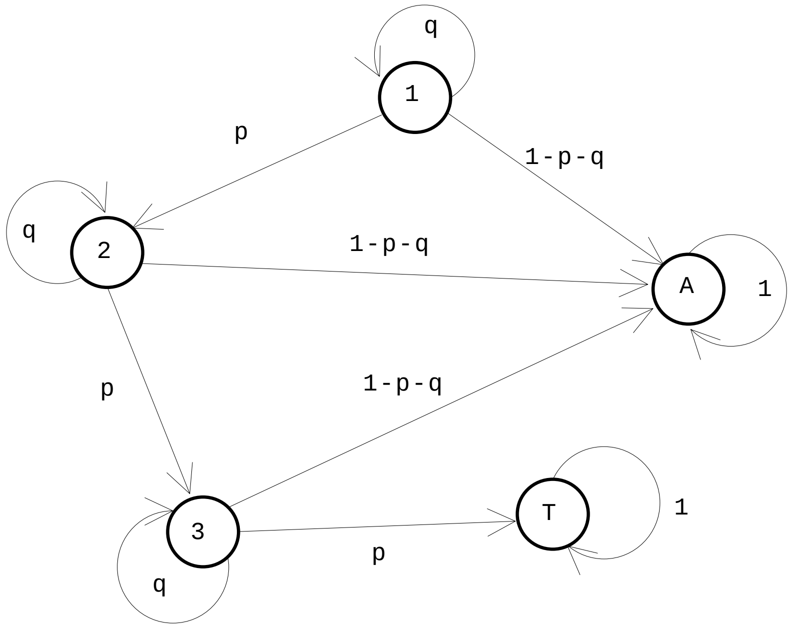
Proposición

Si una subcadena cerrada e irreducible de estados $C \in E$ tiene periodo δ, entonces C puede subdividirse de manera única en δ subconjuntos $D_0, D_1, \ldots, D_{\delta-1}$, tal que $\bigcup_{i=0}^{\delta-1} D_i = C$, de manera que la evolución de la cadena se produce cíclicamente pasando en cada etapa de D_r a D_{r+1} y de $D_{\delta-1}$ a D_0.

Ejemplo. En una carrera que dura 3 años, una persona tiene cada año una probabilidad p de pasar al siguiente curso, una probabilidad q de repetir curso y una probabilidad $1 - p - q$ de abandonar los estudios.

Consideramos, así el espacio de estados $E = \{1, 2, 3, A, T\}$ donde A es abandonar, y T es terminar los estudios.

El grafo correspondiente es



Entonces la matriz de transición es

$$T = \begin{array}{c} {\scriptstyle 1} \\ {\scriptstyle 2} \\ {\scriptstyle 3} \\ {\scriptstyle A} \\ {\scriptstyle T} \end{array} \begin{pmatrix} q & p & 0 & 1-p-q & 0 \\ 0 & q & p & 1-p-q & 0 \\ 0 & 0 & q & 1-p-q & p \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

donde los estados absorbentes son $\{A, T\}$ y los estados transitorios son $\{1, 2, 3\}$. Se reordena la matriz y se obtiene la matriz en forma canónica, poniendo primero los estados recurrentes y luego los transitorios:

$$P = \begin{array}{c} {\scriptstyle A} \\ {\scriptstyle T} \\ {\scriptstyle 1} \\ {\scriptstyle 2} \\ {\scriptstyle 3} \end{array} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1-p-q & 0 & q & p & 0 \\ 1-p-q & 0 & 0 & q & p \\ 1-p-q & p & 0 & 0 & q \end{pmatrix}$$

Se puede preguntar alguien por cosas como: cuál es la probabilidad de acabar los estudios? ¿cuál es el número medio de años que pasa una persona estudiando en el centro?

Así, nos planteamos:

(i) ¿cuál es el número medio de años que pasa una persona estudiando en el primer curso?

En este caso,

$$Q = \begin{pmatrix} q & p & 0 \\ 0 & q & p \\ 0 & 0 & q \end{pmatrix}$$

luego

$$S = (I - Q)^{-1} = \frac{1}{(1-q)^3} \begin{pmatrix} (1-q)^2 & p(1-q) & p^2 \\ 0 & (1-q)^2 & p(1-q) \\ 0 & 0 & (1-q)^2 \end{pmatrix}$$

Se pide calcular R_{11}.

Como el estado 1 es transitorio

$$R_{11} = S_{11} = \frac{(1-q)^2}{(1-q)^3} = \frac{1}{1-q}.$$

Así, si suponemos que la probabilidad de pasar de curso es $p = 0{,}4$, la de repetir curso es $q = 0{,}55$ y la probabilidad de abandonar es $1 - p - q = 0{,}05$, entonces

$$R_{11} = \frac{1}{1-q} = \frac{1}{1-0{,}55} = 2{,}22 \text{ años}$$

Si aumenta el valor de q, entonces aumenta también el número medio de años:

$$R_{11} = \frac{1}{1-q} = \frac{1}{1-0{,}75} = 4 \text{ años.}$$

(ii) ¿cuál es la probabilidad de que se llegue a tercer curso?

La probabilidad de pasar a tercer curso es f_{13}.

Como 1 y 3 son transitorios, entonces

$$f_{13} = \frac{S_{13}}{S_{33}}$$

donde $S = (I - Q)^{-1}$. Entonces

$$f_{13} = \frac{\frac{p^2}{(1-q)^3}}{\frac{(1-q)^2}{(1-q)^3}} = \frac{p^2}{(1-q)^2}.$$

Si $p = 0{,}4$, y $q = 0{,}55$

$$f_{13} = \frac{p^2}{(1-q)^2} = \frac{0{,}16}{0{,}2025} = 0{,}7901$$

Si se toma $p = 0{,}2$, y $q = 0{,}75$ esta probabilidad disminuye

$$f_{13} = \frac{p^2}{(1-q)^2} = \frac{0{,}04}{0{,}0625} = 0{,}64$$

(iii) ¿cuál es la probabilidad de que no se llegue a segundo curso?

Se pide $(1 - f_{12})$.

Como

$$f_{12} = \frac{S_{12}}{S_{22}} = \frac{\frac{p(1-q)}{(1-q)^3}}{\frac{(1-q)^2}{(1-q)^3}} = \frac{p}{1-q},$$

luego

$$(1 - f_{12}) = 1 - \frac{p}{1-q}$$

En el caso de tomar $p = 0{,}2$, y $q = 0{,}75$

$$(1 - f_{12}) = \frac{0{,}05}{0{,}45} = 0{,}111.$$

(iv) ¿Si se empieza en primero, cuál es la probabilidad de acabar la carrera?

Los estados $\{A, T\}$ son recurrentes, de modo que se busca f_{1T} donde 1 es un estado transitorio y T es recurrente. Como T es absorbente (sólo él forma un conjunto cerrado e irreducible) entonces basta calcular la probabilidad de absorción del 1 por el conjunto absorbente T.

Se calcula entonces $G = (I - Q)^{-1}B$ donde

$$B = \begin{pmatrix} 1-p-q & 0 \\ 1-p-q & 0 \\ 1-p-q & p \end{pmatrix}.$$

Así,

$$\begin{aligned} G &= (I-Q)^{-1}B = \frac{1}{(1-q)^3}\begin{pmatrix} (1-q)^2 & p(1-q) & p^2 \\ 0 & (1-q)^2 & p(1-q) \\ 0 & 0 & (1-q)^2 \end{pmatrix}\begin{pmatrix} 1-p-q & 0 \\ 1-p-q & 0 \\ 1-p-q & p \end{pmatrix} = \\ &= \begin{pmatrix} 1-\frac{p^3}{(1-q)^3} & \frac{p^3}{(1-q)^3} \\ 1-\frac{p^2}{(1-q)^2} & \frac{p^2}{(1-q)^2} \\ 1-\frac{p}{1-q} & \frac{p}{1-q} \end{pmatrix}. \end{aligned}$$

Por lo tanto, la probabilidad de acabar la carrera (pasar de 1 a T) es $\frac{p^3}{(1-q)^3}$.

Así, si $p = 0{,}4$, y $q = 0{,}55$,

$$f_{1T} = \frac{0{,}4^3}{0{,}45^3} = 0{,}7023$$

De aquí se deduce que la probabilidad de abandonar la carrera (pasar de 1 a A) es

$$f_{1A} = 1 - 0{,}7023 = 0{,}2977$$

(v) ¿Cuál es el número medio de años que se pasa en la carrera?

Se trata de calcular, en realidad, el número medio de etapas que se pasa en los estados transitorios, cuando se parte del estado transitorio 1. Esto es equivalente a calcular el tiempo medio de absorción desde $i = 1$, es decir, A_1.

Se calcula el número medio de visitas (etapas) que se hacen a 1, a 2 y a 3.

$$A_1 = R_{11} + R_{12} + R_{13}$$

Se considera $S = (I - Q)^{-1}$, entonces

$$A_1 = \frac{1}{(1-q)^3}\left[(1-q)^2 + p(1-q) + p^2\right]$$

Si $p = 0{,}4$, y $q = 0{,}55$,

$$A_1 = \frac{1}{0{,}45^3}\left(0{,}45^2 + 0{,}4 \cdot 0{,}45 + 0{,}4^2\right) = 5{,}95 \text{ años}$$

Si $p = 0{,}2$, y $q = 0{,}75$,

$$A_1 = \frac{1}{0{,}25^3}\left(0{,}25^2 + 0{,}2 \cdot 0{,}75 + 0{,}2^2\right) = 16{,}16 \text{ años}$$

(vi) ¿Cuál es el número medio de años que se tarda en llegar a tercero?

Se tiene que calcular el número medio de etapas que se tienen que hacer desde primero antes de llegar a tercero.

$$\tau_{13} = R_{11} + R_{12} = \frac{(1-q)^2}{(1-q)^3} + \frac{p(1-q)}{(1-q)^3}$$

Si $p = 0{,}4$, y $q = 0{,}55$,

$$\tau_{13} = \frac{0{,}2025 + 0{,}4 \cdot 0{,}45}{0{,}091125} = 4{,}1975 \text{ años.}$$

Distribuciones estacionarias

Se trata de determinar las distribuciones sobre el conjunto de estados E tales que la distribución del proceso X_n en cualquier instante es la misma que en el instante inicial X_0. Se le denomina, también, distribución de *equilibrio*.

El caso más interesante corresponde a cadenas irreducibles y aperiódicas.

Definición (*Distribución estacionaria*)

Sea E el conjunto de estados de una cadena de Markov con matriz de transición T. Una distribución $\pi = [\pi_i]$ donde $i \in E$ se dice que es estacionaria o de equilibrio si

$$\pi = \pi T$$

Alternativamente, se puede escribir como

$$\pi_j = \sum_{i \in E} \pi_i T_{ij}$$

Teorema

Dada la matriz de transición T de una cadena de Markov finita, aperiódica e irreducible con m estados, existe, entonces una única solución al sistema de ecuaciones

$$\pi_j = \sum_{i=1}^{m} \pi_i T_{ij}$$

para todo $j = 1, \ldots, m$, de modo que $\sum_{i=1}^{m} \pi_i = 1$.

Además, la solución viene dada por

$$\pi_j = \lim_{n \to \infty} T_{ij}^{(n)}$$

para todo $i, j = 1, \ldots, m$.

Observaciones.

– Este teorema asegura que si una cadena es finita, irreducible y aperiódica (así es ergódica) existe una única distribución estacionaria que se obtiene resolviendo el correspondiente sistema de ecuaciones.

– El hecho de que para todo $i = 1, \ldots, m$ resulta que $\pi_j = \text{lím}_{n\to\infty} T_{ij}^{(n)}$ implica que cualquiera que sea el estado de partida, la probabilidad de transición en n etapas al estado j se estabiliza y es igual a π_j.

Tiempos medios de recurrencia

Las probabilidades estacionarias de una cadena finita, irreducible y aperiódica (formada por estados recurrentes positivos) están relacionadas con los tiempos medios de recurrencia de los estados.

Proposición

Sea $\{X_n\}$ una cadena finita, irreducible y aperiódica, entonces, $\forall j \in E$ el tiempo medio de recurrencia al estado j, μ_j, se relaciona con π_j así:

$$\pi_j = \lim_{n\to\infty} T_{jj}^{(n)} = \frac{1}{\mu_j}.$$

Observaciones.

– Esto indica que cuanto mayor es el tiempo medio de recurrencia del estado j, menor es la probabilidad de que nos encontremos en ese estado j, cuando pase un tiempo suficientemente grande.

Por otro lado, μ_j se puede calcular como

$$\mu_j = \sum_{n=1}^{\infty} n f_j^{(n)}$$

– En el caso de que se tenga una cadena de Markov **no** finita, aperiódica e irreducible se cumple el teorema anterior, aunque se requiere que todos sus estados sean recurrentes positivos.

Probabilidades de transición límite

En general, dada una matriz en forma canónica P se puede estudiar qué valores tomaría

$$P^{\infty} = \lim_{n\to\infty} P^n$$

donde P es la matriz de transición de una cadena de Markov finita, no necesariamente irreducible y aperiódica.

La matriz P, en su forma canónica, tiene la siguiente expresión

$$P = \begin{pmatrix} P_1 & & & \\ & \ddots & & 0 \\ & & P_r & \\ & \mathbf{L} & & \mathbf{Q} \end{pmatrix}$$

Casos:

(i) Cada una de las submatrices P_i, $i = 1, \ldots, r$ está asociada a r cadenas irreducibles, y para este tipo de cadenas se tiene una distribución límite que es estacionaria, esto es, $\forall i$,

$$\pi_j = \lim_{n\to\infty} P_{ij}^{(n)}.$$

(ii) En los estados transitorios las distribuciones límite asignan probabilidad 0, luego

$$\lim_{n\to\infty} P_{ij}^{(n)} = 0$$

$\forall j$ transitorio.

(iii) Para i transitorio y j recurrente,

$$\lim_{n\to\infty} P_{ij}^{(n)} = \lim_{n\to\infty} f_{ij} P_{jj}^{(n)} = f_{ij} \lim_{n\to\infty} P_{jj}^{(n)} = f_{ij}\pi_j,$$

donde

$$f_{ij} = \sum_{k=0}^{\infty} f_{ij}^{(k)}.$$

Ejemplo.

Un grupo de personas elige un lugar de viaje de fin de fin curso entre Bahamas, Europa o Hawai utilizando las siguientes reglas:

Si estuvo en Bahamas el año anterior, se elige Europa con probabilidad $p = \frac{2}{3}$ y Hawai con $p = \frac{1}{3}$.

Si estuvo en Europa el año anterior, se elige Bahamas con probabilidad $p = \frac{3}{8}$, Europa con $p = \frac{1}{8}$ y Hawai con $p = \frac{1}{2}$.

Si estuvo en Hawai se elige Bahamas o Europa con probabilidad $p = \frac{1}{2}$.

En este caso, $E = \{B, E, H\}$ con matriz de transición

$$T = \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{3} \\ \frac{3}{8} & \frac{1}{8} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

Es una cadena finita, aperiódica e irreducible, por lo que se calcula su distribución estacionaria.

$$\begin{pmatrix} \pi_B & \pi_E & \pi_H \end{pmatrix} = \begin{pmatrix} \pi_B & \pi_E & \pi_H \end{pmatrix} \begin{pmatrix} 0 & \frac{2}{3} & \frac{1}{3} \\ \frac{3}{8} & \frac{1}{8} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

de modo que queda el sistema de ecuaciones

$$\begin{aligned} \pi_B &= \frac{3}{8}\pi_E + \frac{1}{2}\pi_H \\ \pi_E &= \frac{2}{3}\pi_B + \frac{1}{8}\pi_E + \frac{1}{2}\pi_H \\ \pi_H &= \frac{1}{3}\pi_B + \frac{1}{2}\pi_E \\ \pi_B + \pi_E + \pi_H &= 1 \end{aligned}$$

dado que las tres primeras ecuaciones son linealmente dependientes. Se obtiene

$$\begin{aligned} \pi_B &= 0{,}3 \\ \pi_E &= 0{,}4 \\ \pi_H &= 0{,}3 \end{aligned}$$

Se puede interpretar que después de un largo periodo de tiempo se prefiere Europa con probabilidad 0,4, y después Bahamas y Hawai con probabilidad 0,3.