Universidad del País Vasco
Matemática Aplicada y Estadística

Matrices no negativas, paseos aleatorios y cadenas de Markov

Juan-Miguel Gracia

Extracto: Damos algunas definiciones de digrafo, cadena de Markov, paseo aleatorio, así como su relación con la teoría de Perron-Frobenius de matrices no negativas. Contiene programas en MATLAB.

mepgrmej@vc.ehu.es
11 de noviembre de 2002 Versión final

Índice General

1. Digrafos

La terminología sobre grafos no está estandarizada y varía más de lo que sería de desear de unos autores a otros. Hasta donde llega nuestra experiencia sobre este tema, nos induce a leer con cuidado las definiciones de cada autor. En la literatura, muchos conceptos no son siquiera definidos y se debe apelar a su significado intuitivo. Nuestras definiciones persiguen la claridad formal, conscientes del sacrificio intuitivo del tema; éste aspecto lo fiamos a las figuras que insertamos.

Un **digrafo (grafo dirigido u orientado)** D es un par (V, E) donde V es un conjunto finito de elementos llamados *vértices (puntos o nodos)* y E es un subconjunto de $V \times V$. Todo par ordenado $(a, b) \in E$ se llama *arco* del digrafo D. Nótese que un digrafo puede contener tanto los arcos (a, b) y (b, a) con a y b vértices distintos como arcos de la forma (a, a); éstos últimos arcos se denominan *lazos* . Los vértices a y b de un arco $\alpha = (a, b)$ se llaman extremos del arco α. Además, a es el *vértice inicial u origen* y b es el *vértice final* de α. Se dice que el arco (a, b) es *incidente* con los vértices a y b. El número de arcos que salen de un vértice es el **grado saliente** del vértice. El número de arcos que entran en un vértice es el *grado entrante* del vértice. Convenimos que un lazo en un vértice aporta una unidad al grado saliente y también una unidad al grado entrante. Se dice que el arco (a, b) es *incidente* con los vértices a y b.

1.1. Subdigrafo

Un *subdigrafo* del digrafo $D = (V, E)$ es un digrafo $D_1 = (V_1, E_1)$ tal que $V_1 \subset V$ y $E_1 \subset E$. Se dice que el subdigrafo D_1 es un *subdigrafo generador* de D si D_1 tiene el mismo conjunto de vértices que D, es decir, si $V_1 = V$.

1.2. Paseo

Un *paseo* de D es una secuencia finita de vértices $w = (v_0, v_1, \ldots, v_p)$ tal que $(v_{i-1}, v_i) \in E$ para $i = 1, 2, \ldots, p$ y $p > 0$. El número p es la *longitud* del paseo w. Este número coincide con el número de arcos

$$(v_0, v_1), (v_1, v_2), \ldots, (v_{p-1}, v_p)$$

del paseo w; contando los arcos cuantas veces se pase por ellos para ir desde el vértice v_0 al vértice v_p.

Ejemplo 1.1 Consideremos el digrafo D representado por la Figura 1.

Aquí $D = (V, E)$ con $V = \{1, 2, 3, 4, 5, 6\}$ y

$$E = \{(1, 2), (2, 1), (2, 2), (2, 3), (3, 5), (3, 6), (4, 2), (4, 4), (5, 2), (6, 3)\}.$$

El paseo $w_1 = (1, 2, 3, 6)$ tiene longitud 3 y $w_2 = (1, 2, 2, 1, 2)$ es un paseo de longitud 4.

Un *camino* de D es un paseo $w = (v_0, v_1, \ldots, v_p)$ que tiene todos los vértices distintos. Un *ciclo* de D es un paseo $w = (v_0, v_1, \ldots, v_p)$ con los vértices

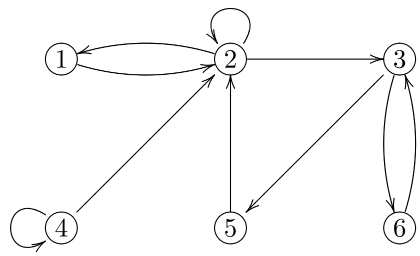


Figura 1: Digrafo D.

$v_0, v_1, \ldots, v_{p-1}$ distintos y $v_0 = v_p$. En particular, los lazos (a, a) son ciclos. Tenemos la clasificación siguiente de los paseos:

$$\text{paseos} \begin{cases} \text{caminos,} \\ \text{ciclos,} \\ \text{paseos} \quad w = (v_0, v_1, \ldots, v_p) \\ \quad \text{con repeticiones en } v_0, v_1, \ldots, v_{p-1} \text{ y/o en } v_1, \ldots, v_{p-1}, v_p. \end{cases}$$

Un digrafo se llama *acíclico* si no tiene ciclos. En el digrafo anterior el paseo $C = (2, 3, 5, 2)$ es un ciclo (Figura 2) y el paseo $P = (1, 2, 3, 5)$ es un camino (Figura 3) . El número de vértices distintos de un camino P o de un ciclo C será denotado por $|P|$ o $|C|$, respectivamente. Así, en el ejemplo último, $|P| = 4$ y $|C| = 3$. También $|P| =$ longitud de $P\ +1$, $|C| =$ longitud de C. La notación $|X|$ indica el cardinal de un conjunto finito X. Aquí se comete un ligero abuso de notación pues P y C no son conjuntos.

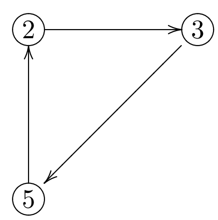


Figura 2: Ciclo $C = (2, 3, 5, 2)$.

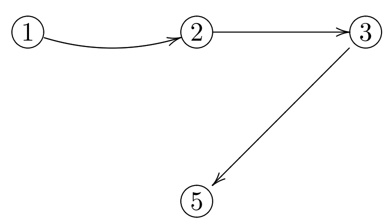


Figura 3: Camino $P = (1, 2, 3, 5)$.

1.3. Digrafo fuertemente conexo

Un digrafo $D = (V, E)$ se llama **fuertemente conexo** si para *cualquier par* $(i, j) \in V \times V$, con $i \neq j$, existe un paseo $(i = v_0, v_1, \ldots, v_p = j)$ que conecta i con j.

Obsérvese que los vértices i y j pueden estar conectados por un paseo mientras que j e i no estarlo.

Ejercicio 1.1 Comprobar que el digrafo de la Figura 4 es fuertemente conexo.

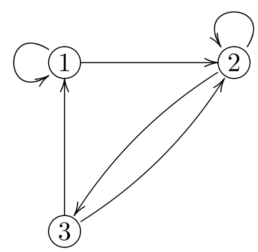


Figura 4: Digrafo fuertemente conexo.

1.4. Digrafo ponderado

Un **digrafo ponderado** es una terna (V, E, p) en la que $D = (V, E)$ es un digrafo y $p : E \to \mathscr{R}$ es una aplicación de E en un anillo conmutativo $\mathscr{R}$, que asocia a cada arco $(\alpha, \beta) \in E$ su *peso* $p((\alpha, \beta))$.

1.5. Isomorfismo de digrafos

Sean los digrafos $D_i = (V_i, E_i),\ i = 1, 2$. Se llama **isomorfismo de los digrafos** D_1 y D_2 a toda aplicación biyectiva $\varphi\colon V_1 \to V_2$ que satisface la condición

$$(i, j) \in E_1 \Longleftrightarrow (\varphi(i), \varphi(j)) \in E_2.$$

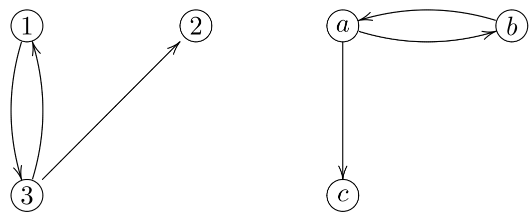
Ejemplo 1.2 Sean $D_1 = (V_1, E_1)$ y $D_2 = (V_2, E_2)$ los digrafos dados por la Figura 5. Si definimos la aplicación $\varphi\colon \{1, 2, 3\} \to \{a, b, c\}$ dada por

$$\varphi(1) = b, \quad \varphi(2) = c, \quad \varphi(3) = a,$$

se tiene que

$$\begin{aligned}
(1, 3) \in E_1 &\Longleftrightarrow (\varphi(1), \varphi(3)) = (b, a) \in E_2, \\
(3, 1) \in E_1 &\Longleftrightarrow (\varphi(3), \varphi(1)) = (a, b) \in E_2, \\
(3, 2) \in E_1 &\Longleftrightarrow (\varphi(3), \varphi(2)) = (a, c) \in E_2.
\end{aligned}$$

Por tanto, φ es un isomorfismo.



(a) D_1.   (b) D_2.

Figura 5: Digrafos isomorfos.

Si los digrafos son *ponderados* en el mismo anillo $\mathscr{R}$, digamos (V_1, E_1, p) y (V_2, E_2, q), un **isomorfismo** entre ellos es una aplicación biyectiva φ entre los conjuntos de vértices que conserva no solo los arcos, sino también su peso respectivo; i.e.

$$\text{para todo } (i, j) \in E_1, \quad p((i, j)) = q(\varphi(i), \varphi(j)).$$

1.6. Paseo aleatorio

Sea ahora $G = (V, E, p)$ un digrafo ponderado donde $p\colon E \to \mathbb{R}$ asocia al arco (i, j) un peso $p\big((i, j)\big) = p_{ij} > 0$, y tal que para todo vértice i se tenga que la suma de los pesos de los arcos que salen de i sea igual a 1. Si denotamos por $(i, j_1), (i, j_2), \ldots, (i, j_{k_i})$ a todos estos arcos, esta condición se rescribe así

$$\sum_{s=1}^{k_i} p_{ij_s} = 1.$$

Para cada $i \in V$, supongamos que el intervalo $(0, 1]$ es partido en los subintervalos disjuntos dos a dos

$$(a_{i0}, a_{i1}], (a_{i1}, a_{i2}], \ldots, (a_{ik_i-1}, a_{ik_i}],$$

donde $a_{is} := p_{ij_1} + \cdots + p_{ij_s}$, $s = 1, 2, \ldots, k_i$, y $a_{i0} := 0$.

Un **paseo aleatorio** por un digrafo ponderado $G = (V, E, p)$ es un proceso aleatorio indefinido en cuyas etapas un caminante situado en el vértice i elige un número al azar x con distribución uniforme en el intervalo (0,1] y salta al vértice j_s si

$$x \in (a_{is-1}, a_{is}], \quad s \in \{1, 2, \ldots, k_i\}.$$

Allí vuelve a repetir el procedimiento de elegir un nuevo vértice al que saltar. Puede ocurrir que $p_{ii} = 1$, en cuyo caso suponemos que el caminante salta todo el rato de i a i. Si al comienzo del paseo o en un momento dado, el caminante está situado en un vértice sin arcos salientes, o en vértice i tal que $p_{ii} = 1$, se supone que se queda allí absorbido.

2. Potencias de una matriz

Dada $A \in \mathbb{C}^{n\times n}$ consideremos algunos resultados importantes sobre el comportamiento asintótico de la sucesión de potencias de A: $A^k, \quad k = 0, 1, 2, \ldots$. Un valor propio λ de A se llama **semisimple** si la multiplicidad algebraica de λ coincide con su multiplicidad geométrica: $\mathrm{m}(\lambda, A) = \mathrm{mg}(\lambda, A)$. Por tanto, un valor propio simple es semisimple, pero no al revés. También, el decir que un valor propio λ es semisimple es equivalente a decir que su índice, $\nu(\lambda)$, es igual a 1. Se llama **radio espectral** de una matriz $A \in \mathbb{C}^{n\times n}$ al máximo de los valores absolutos de sus valores propios; se denota por $\varrho(A)$.

El límite $\lim_{k\to\infty} A^k$ existe si y solo si (i) $\varrho(A) < 1$, o bien (ii) $\varrho(A) = 1$, 1 es el único valor propio de A de módulo uno, y 1 es semisimple.

Cuando existe, $\lim_{k\to\infty} A^k$ es el proyector sobre $\mathrm{Ker}(I - A)$ a lo largo de $\mathrm{Im}(I - A)$. Además, $\lim_{k\to\infty} A^k = 0$ si y solo si $\varrho(A) < 1$. Cuando existe $\lim_{k\to\infty} A^k$, se dice que A es una **matriz convergente**.

Se dice que A es **sumable Cesàro** (o meramente *sumable*) a G si existe el límite

$$\lim_{k\to\infty} \frac{I + A + A^2 + \cdots + A^{k-1}}{k} = G.$$

La matriz $A \in \mathbb{C}^{n\times n}$ es sumable Cesàro si y solo si $\varrho(A) < 1$ o bien $\varrho(A) = 1$ con todos los valores de módulo uno semisimples.

Cuando existe, el límite Cesàro

$$\lim_{k\to\infty} \frac{I + A + A^2 + \cdots + A^{k-1}}{k} = G$$

es el proyector sobre $\mathrm{Ker}(I - A)$ a lo largo de $\mathrm{Im}(I - A)$. Si A es convergente a G, entonces A sumable a G, pero no recíprocamente. Las demostraciones de estos resultados son consecuencia de la forma de las potencias de una matriz en forma canónica de Jordan; el detalle puede consultarse en [5, p. 630, 633].

3. Matrices no negativas

Una matriz $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ se dice *no negativa* si todos sus elementos son ≥ 0. Si todos los elementos de A son positivos (i.e > 0), entonces A se llama *positiva*. Si m o n son iguales a 1, hablaremos de vectores no negativos, resp. positivos. Denotamos por $A \geq 0$ (resp. $A > 0$) cuando *para todo* par (i, j) se tiene que $a_{ij} \geq 0$ (resp. $a_{ij} > 0$). Nótese que $A \geq 0$ y $A \neq 0$ no implican $A > 0$.

La duplicación de terminología con las matrices (hermíticas) definidas no negativas y definidas positivas es desafortunada. Sin embargo, no hay peligro de confusión pues raramente se mezclan ambos conceptos con los de matriz no negativa y matriz positiva. Además, es de esperar que el lector esté suficientemente atento al contexto para evitarlo.

Los notables teoremas de Perron y Frobenius constituyen el resultado central de la teoría de las matrices no negativas, y serán enunciados en los Teoremas 3.3, 3.4 y 3.5.

Sea $A = (a_{ij}) \in \mathbb{C}^{n\times n}$, se llama digrafo asociado a la matriz A al digrafo $D(A) = (V, E, a)$ cuyo conjunto de vértices es $V = \{1, 2, \ldots, n\}$ y con arcos $(i, j) \in E$ siempre que $a_{ij} \neq 0$ y de peso a_{ij}. Por ejemplo, si

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ -3 & 5 & 2 & 0 \\ 0 & 0 & -8 & 1 \\ 0 & 0 & 5 & -4 \end{pmatrix},$$

entonces su digrafo $D(A)$ viene dado por la Figura 6.

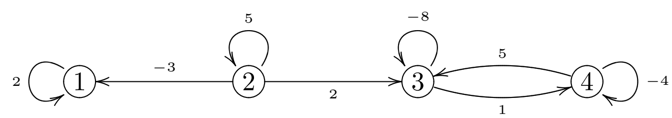


Figura 6: Digrafo D.

Sea σ una permutación de $\{1, 2, \ldots, n\}$. Sea $M(\sigma) := (\delta_{i,\sigma(j)})$ la **matriz asociada a la permutación** σ, donde δ_{ik} son las deltas de Kronecker; $M(\sigma)$ es una matriz de ceros y unos con un solo 1 en cada fila y columna; es decir, $M(\sigma)$ es una matriz de permutación. Si permutamos las columnas de $A = (a_{ij})$ de acuerdo con la permutación σ y después permutamos las filas de la matriz resultante de la misma manera, obtenemos una matriz $B = (b_{ij})$, cuyos elementos b_{ij} vienen dados en función de los elementos de A por la expresión

$$b_{ij} = a_{\sigma(i),\sigma(j)}.$$

Además,

$$B = M(\sigma)^{\mathrm{T}} A M(\sigma).$$

También nos interesar señalar que los digrafos asociados a las matrices A y B son isomorfos; siendo el isomorfismo precisamente la permutación

$$\sigma\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}.$$

Ejemplo 3.1 Sea la matriz

$$A = \begin{pmatrix} 0 & 0 & a_{13} \\ 0 & 0 & 0 \\ a_{31} & a_{32} & 0 \end{pmatrix},$$

su digrafo asociado aparece en la parte (a) de la Figura 7. Si consideramos la permutación

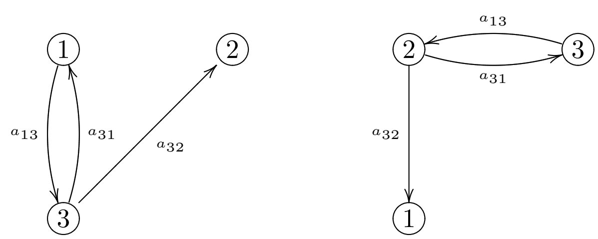
$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

de $\{1,2,3\}$, se tiene que

$$B = M(\sigma)^{\mathrm{T}} AM(\sigma) =$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & a_{13} \\ 0 & 0 & 0 \\ a_{31} & a_{32} & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ a_{32} & 0 & a_{31} \\ 0 & a_{13} & 0 \end{pmatrix}.$$

El digrafo de B aparece en la parte (b) de la Figura 7.



(a) $D(A)$). (b) $D(B)$.

Figura 7: Digrafos σ-permutados.

La matriz $A \in \mathbb{C}^{n\times n}$ se llama **reducible** si existe una matriz de permutación P de orden n tal que

$$P^{\mathrm{T}} AP = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

donde A_{11} y A_{22} de órdenes menores que n y ≥ 1, cuando $n \geq 2$; o si $n = 1$ y $A = (0)$. Si no existe P entonces A es **irreducible**. Otra definición equivalente de matriz reducible es la siguiente. Si el conjunto $\{1, 2, \ldots, n\}$ puede ser partido en dos subconjuntos complementarios no vacíos

$$\{i_1, i_2, \ldots, i_\mu\} \text{ y } \{j_1, j_2, \ldots, j_\nu\}, \qquad (\mu + \nu = n),$$

tales que

$$a_{i_\alpha j_\beta} = 0 \quad (\alpha = 1, \ldots, \mu; \beta = 1, \ldots, \nu),$$

se dice que la matriz $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ es reducible [3, p. 46].

Teorema 3.1 *Una matriz cuadrada es irreducible si y solo si su digrafo es fuertemente conexo.*

Demostración. Véase [4, Theorem 1, p. 529].

Teorema 3.2 *Sea $M \in \mathbb{R}^{n\times n}$ una matriz no negativa. Sea k un entero positivo. Entonces el elemento $m_{ij}^{(k)}$ situado en el lugar (i,j) de la potencia M^k es distinto de cero si y solo si hay un paseo del vértice i al j de longitud k en el digrafo $D(M)$, asociado a M.*

Demostración. Véase [2, p. 89, Theorem 4.4].

3.1. Forma normal de una matriz reducible

Si una matriz $A \in \mathbb{C}^{n\times n}$ es reducible, entonces, por definición, existe una matriz de permutación P tal que

$$P^{\mathrm{T}}AP = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix},$$

donde X y Z de órdenes menores que n y ≥ 1. Por conveniencia, denotamos esto escribiendo

$$A \sim \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}.$$

Si X o Z es reducible, entonces puede ser realizada otra permutación para producir

$$\begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix} \sim \begin{pmatrix} R & S & T \\ 0 & U & V \\ 0 & 0 & W \end{pmatrix},$$

donde R, U y W son cuadradas. Repitiendo este proceso se llega eventualmente a

$$A \sim \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1k} \\ 0 & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X_{kk} \end{pmatrix},$$

donde cada X_{ii} es irreducible o $X_{ii} = (0)_{1\times 1}$. Un bloque diagonal $X_{ii}, 1 \leq i \leq k$, se dice *aislado* si

$$X_{ij} = 0 \text{ para } j = i+1, i+2, \ldots, k.$$

Mediante una permutación de filas y columnas de bloques para llevar los bloques diagonales aislados a la parte de abajo, se llega a

$$A \sim N := \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1r} & A_{1,r+1} & A_{1,r+2} & \cdots & A_{1m} \\ 0 & A_{22} & \cdots & A_{2r} & A_{2,r+1} & A_{2,r+2} & \cdots & A_{2m} \\ \vdots & & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_{rr} & A_{r,r+1} & A_{r,r+2} & \cdots & A_{rm} \\ 0 & 0 & \cdots & 0 & A_{r+1,r+1} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & A_{r+2,r+2} & \cdots & 0 \\ \vdots & & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & A_{mm} \end{pmatrix}, \quad (1)$$

donde cada bloque diagonal $A_{11}, \ldots, A_{mm}$ es, o bien irreducible, o bien $(0)_{1\times 1}$, y cada fila de bloques

$$A_{i,i+1}, A_{i,i+2}, \ldots, A_{i,r+1}, A_{i,r+2}, \ldots, A_{im} \quad (i = 1, \ldots, r)$$

tiene al menos un bloque A_{ij} diferente de cero. Además, como dijimos al permutar las líneas de una matriz los digrafos ponderados de A y de la matriz N son isomorfos. A la matriz N se le llama **forma normal para la matriz reducible** A.

3.2. Teoremas de Perron y Frobenius

Teorema 3.3 (Perron) *Una matriz positiva cuadrada A tiene un valor propio real y positivo r que tiene multiplicidad algebraica 1 y es mayor que el valor absoluto de cualquier otro valor propio de A. El valor propio r tiene asociado un vector propio por la derecha positivo.*

DEMOSTRACIÓN. Véase [4, Corollary 2, p. 542].

Teorema 3.4 (Frobenius, Parte 1ª) *Sea $A \in \mathbb{R}^{n\times n}$ una matriz no negativa e irreducible, entonces*

(a) *La matriz A tiene un valor propio positivo r, igual al radio espectral de A;*

(b) *Existe un vector propio positivo (por la derecha) asociado al valor propio r;*

(c) *El valor propio r tiene multiplicidad algebraica 1.*

El único vector $\boldsymbol{p}$ definido por

$$A\boldsymbol{p} = r\boldsymbol{p}, \quad \boldsymbol{p} > 0, \quad \|\boldsymbol{p}\|_1 = 1,$$

se llama el **vector de Perron**.

Teorema 3.5 (Frobenius, Parte 2ª) *Sea la matriz $A \in \mathbb{R}^{n\times n}$ no negativa e irreducible y llamemos $\lambda_1, \lambda_2, \ldots, \lambda_n$ a sus valores propios, donde puede haber repetidos. Si hay exactamente d valores propios $\lambda_1 = r, \lambda_2, \ldots, \lambda_d$ de valor absoluto r y si $\omega_1, \ldots, \omega_d$ son las raíces d-ésimas distintas de la unidad, entonces $\lambda_j = \omega_j r, j = 1, 2, \ldots, d$.*

Además, los puntos del plano complejo que representan a $\lambda_1, \lambda_2, \ldots, \lambda_n$ forman un conjunto invariante bajo rotaciones alrededor del origen de ángulo $2\pi/d$. Si $d > 1$, entonces existe una matriz de permutación P tal que $P^{\mathrm{T}}AP$ tiene la forma partida ***(Forma de Frobenius)***

$$\begin{bmatrix} 0 & A_{12} & 0 & \cdots & 0 \\ 0 & 0 & A_{23} & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & A_{d-1,d} \\ A_{d1} & 0 & \cdots & 0 & 0 \end{bmatrix}, \tag{2}$$

en la que los bloques diagonales son cuadrados.

DEMOSTRACIÓN. Véanse las páginas 536 y 540 de[4].

3.3. Matrices reducibles

Teorema 3.6 *Si la matriz $A \in \mathbb{R}^{n\times n}$ es no negativa, entonces*

(a) *La matriz A tiene un valor propio real, r, igual al radio espectral de A;*

(b) *Hay un vector propio no negativo (por la derecha) asociado al valor propio r.*

3.4. Matrices primitivas

Sea A una matriz cuadrada no negativa e irreducible. Sea d el número de valores propios distintos de A de valor absoluto igual al radio espectral $\varrho(A)$. Se dice que A es una **matriz primitiva o imprimitiva** según que sea $d = 1$ o $d > 1$. El número d se llama el **índice de imprimitividad** de A. Veremos que este concepto esta relacionado con el posible periodo de repetición de cierto patrón en una cadena de Markov. Por el Teorema de Perron (Teorema 3.3) se tiene que toda matriz cuadrada positiva es primitiva.

Ejemplo 3.2 Sea la matriz

$$A = \begin{pmatrix} 0 & 0{,}7 & 0{,}3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0{,}8 & 0 & 0 & 0 & 0 & 0{,}2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0{,}6 & 0 & 0 & 0{,}4 & 0 \end{pmatrix};$$

el polinomio característico de A es $p(\lambda) = \lambda^6 - \lambda^3 = \lambda^3(\lambda^3 - 1)$; de aquí que los valores propios de A son 0 (triple) y las raíces cúbicas de la unidad: $-0{,}5 \pm 0{,}866\mathrm{i}, 1$. Por tanto, A es una matriz irreducible con índice de imprimitividad 3 (no primitiva). Sea

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 6 & 2 & 3 & 5 & 4 \end{pmatrix}$$

la permutación de 1,2,3,4,5,6 que lleva las líneas de A a su forma de Frobenius

$$A_{\mathrm{F}} = \left(\begin{array}{cc|ccc|c} 0 & 0 & 0{,}7 & 0{,}3 & 0 & 0 \\ 0 & 0 & 0{,}6 & 0 & 0{,}4 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0{,}8 & 0{,}2 & 0 & 0 & 0 & 0 \end{array}\right), \tag{3}$$

entonces $A_{\mathrm{F}} = M(\sigma)^{\mathrm{T}} A M(\sigma)$, siendo $M(\sigma)$ la matriz de permutación asociada a σ.

Una función de MATLAB que construye la forma de Frobenius A_{F} de una matriz irreducible no primitiva A y la matriz de permutación de paso M, es frobform.m.

Teorema 3.7 *Una matriz cuadrada no negativa A de orden n es primitiva si y solo si existe un entero positivo p tal que $A^p > 0$.*

DEMOSTRACIÓN. Véase [4, Theorem 1, p. 546].

Puede demostrarse [4, p. 547] que p puede tomarse $\leq n^2 - 2n + 2$. Es decir, que si A^q tiene algún elemento igual a 0 para todo $q = 1, \ldots, n^2 - 2n + 2$, entonces A no es primitiva.

El índice de imprimitividad de una matriz está relacionado con los exponentes correspondientes a los coeficientes no nulos del polinomio característico.

Teorema 3.8 *Sea $\lambda^n + a_1\lambda^{n_1} + \cdots + a_q\lambda^{n_q}$ el polinomio característico de una matriz irreducible $A \geq 0$, , donde $a_1, \ldots, a_q$ son diferentes de cero y $n > n_1 > \cdots > n_q \geq 0$. Entonces, el índice de imprimitividad d de A es igual al máximo común divisor de las diferencias $n - n_1, n_1 - n_2, \ldots, n_{q-1} - n_q$.*

DEMOSTRACIÓN. Véase [1, p. 34, Theorem 2.2.27].

Una matriz irreducible A con $r = \varrho(A)$ es primitiva si y solo si $\lim_{k\to\infty}(A/r)^k$ existe, en cuyo caso

$$\lim_{k\to\infty}\left(\frac{A}{r}\right)^k = G = \frac{\boldsymbol{p}\boldsymbol{q}^{\mathrm{T}}}{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{p}} > 0,$$

donde $\boldsymbol{p}$ y $\boldsymbol{q}$ son los vectores de Perron respectivos de A y A^{T}. G es el proyector (espectral) sobre $\mathrm{Ker}(A - rI)$ a lo largo de $\mathrm{Im}(A - rI)$.

Este hecho está demostrado en [5, p. 675].

3.5. Matrices estocásticas

Una matriz real $P = (p_{ij}) \in \mathbb{R}^{n\times n}$ se llama **estocástica** si $P \geq 0$ y la suma de todas sus filas es igual a 1:

$$\sum_{j=1}^{n} p_{ij} = 1, \quad i = 1, \ldots, n. \tag{4}$$

Teorema 3.9 *Una matriz no negativa P es estocástica si y solo si tiene el valor propio 1 con vector propio (por la derecha) dado por $\boldsymbol{e} = [1, 1, \ldots, 1]^{\mathrm{T}}$. Además, el radio espectral de una matriz estocástica es 1.*

DEMOSTRACIÓN. Véase[4, Theorem 1, p. 547].

4. Cadenas de Markov discretas

Empezaremos hablando de un ejemplo típico. Sea una rata que se mueve entre los cuatro compartimentos 1, 2, 3 y 4 del laberinto descrito en la Figura 8. Supongamos que en determinados instantes de tiempo $t = 0, 1, 2, \ldots$ la rata elige una de las salidas o elige permanecer en el compartimento (en un intervalo de tiempo dado) con la misma probabilidad; así pues, si la rata está en 1 tiene

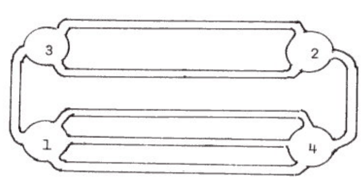


Figura 8: Laberinto 1º.

ante sí cinco sucesos igualmente probables (elegir la salida a 3, elegir la primera salida a 4, elegir la segunda salida a 4, elegir la tercera salida a 4, quedarse en 1). Tiene unas probabilidades de 1/5 de permanecer en 1, 1/5 de ir a 3 y 3/5 de ir a 4, en el movimiento siguiente, respectivamente.

Admitimos la hipótesis de que la rata "no usa su memoria"; es decir, que el movimiento que va a hacer en cada instante es independiente de la sucesión de lugares que ha ido visitando hasta entonces. Se trata de predecir el comportamiento de la sucesión aleatoria de sitios que va visitando; por ejemplo 1 4 1 1 3 1 4 2 3 3 3 1 3 2 3 3 3 1 4 1 4 2 4 4 1 4 1 3 1 4 1 1 1 4 2 2 4 1 1 4 4 . . . ; observamos que no hay transiciones del 1 al 2, ni del 2 al 1.

¿Cuál será la evolución a largo plazo? A corto plazo, si empieza en 1, ¿cuál será el número medio de etapas hasta que visite 2 por vez primera? ¿. . . , por vez tercera? etc.

En general, consideremos un proceso aleatorio en el que se produce un cambio de estado en ciertos instantes de tiempo discretos $t = 0, 1, 2, \ldots$. Supongamos que hay un número finito de estados posibles $S_1, S_2, \ldots, S_n$. Surge así una sucesión (o cadena) de situaciones $X_0, X_1, \ldots, X_t, \ldots$, el la que cada X_t es igual a uno de los n estados. Este proceso se llama **cadena de Markov** si las probabilidades condicionadas que expresan este cambio de situaciones satisfacen la *propiedad de Markov:*

$$P(X_{t+1} = S_j | X_t = S_i, X_{t-1} = S_{i_{t-1}}, \ldots, X_0 = S_{i_0}) = P(X_{t+1} = S_j | X_t = S_i)$$

para todos los instantes t, todos los estados S_i, S_j, y todas las posibles sucesiones $S_{i_0}, \ldots, S_{i_{t-1}}$ de estados previos. Esta propiedad traduce la condición de que las cadenas de Markov no guardan su historia pasada en la memoria; en cada momento, la evolución al estado siguiente hace tabla rasa de lo anterior y empieza de nuevo.

El valor

$$p_{ij}(t) := P(X_t = S_j | X_{t-1} = S_i)$$

es la probabilidad de que la cadena esté en el estado S_j en el instante t dado que haya pasado por el estado S_i en el instante $t-1$; por esto, a $p_{ij}(t)$ se le llama la **probabilidad de transición** de moverse del estado S_i al S_j en el instante t. La matriz de las probabilidades de transición

$$P(t) := \big(p_{ij}(t)\big)$$

es no negativa para todo t, y sus filas suman 1. Así pues, $P(t)$ es una matriz estocástica. Cuando las probabilidades de transición no dependen del tiempo (digamos que $p_{ij}(t) = p_{ij}$ para todo t), la cadena de Markov se dice *estacionaria* (u *homogénea*), y la **matriz de transición** es la matriz constante $P = (p_{ij})$. Supondremos siempre esta condición. Recíprocamente, toda matriz estocástica $P = (p_{ij}) \in \mathbb{R}^{n\times n}$ define una cadena de Markov de n estados porque los elementos p_{ij} definen un conjunto de probabilidades de transición, que puede ser interpretado como una cadena de Markov estacionaria con n estados.

Una cadena de Markov también puede ser interpretada como un paseo aleatorio por el digrafo ponderado $D(P)$ de su matriz estocástica $P = (p_{ij})$. Es claro que ahora al arco (i, j) de $D(P)$ se le atribuye el peso $p_{ij} > 0$.

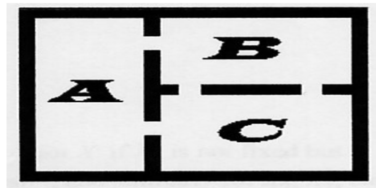


Figura 9: Laberinto 2º.

Un **vector de distribución de probabilidad** se define como un vector no negativo $\boldsymbol{p} = (p_1, p_2, \ldots, p_n)$ tal que $\sum_k p_k = 1$. (Toda fila de una matriz estocástica es uno de estos vectores.) Para una cadena de markov de n-estados, el **vector de distribución de probabilidad de la etapa** k-ésima está definido por

$$\boldsymbol{p}^{\mathrm{T}}(k) = (p_1(k), p_2(k), \ldots, p_n(k)), \quad k = 1, 2, \ldots, \text{ donde } p_j(k) = P(X_k = S_j).$$

En otras palabras, $p_j(k)$ es la probabilidad de estar en el estado j-ésimo después de la k-ésima etapa, pero antes de la etapa $(k+1)$-ésima. El **vector de distribución inicial** es

$$\boldsymbol{p}^{\mathrm{T}}(0) = (p_1(0), p_2(0), \ldots, p_n(0)), \text{ donde } p_j(0) = P(X_0 = X_j)$$

es la probabilidad de que la cadena empieze en S_j.

En el ejemplo de la Figura 8, la matriz de las probabilidades de transición es

$$P = \begin{pmatrix} 1/5 & 0 & 1/5 & 3/5 \\ 0 & 1/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & 1/4 & 0 \\ 3/5 & 1/5 & 0 & 1/5 \end{pmatrix}.$$

Un cálculo estándar con Matlab nos da que el espectro de P es aproximadamente $\Lambda(P) = \{1,\ 0{,}55,\ -0{,}5608,\ -0{,}0892\}$. Así pues, se ve que una matriz estocástica P es una matriz no negativa que tiene radio espectral $\varrho(P) = 1$. Además $\boldsymbol{e} = (1, 1, 1, 1)^{\mathrm{T}}$ es un vector propio positivo asociado a $\varrho(P) = 1$. Con Matlab hemos realizado tres simulaciones de 30 etapas, tomando el primer compartimento 1, 2, 3, ó 4, con probabilidad 1/4 cada uno. He aquí los resultados:

- 3 2 3 2 3 2 3 1 4 1 3 2 3 2 3 2 2 3 3 1 4 1 4 1 4 4 4 1 3 2 3
- 4 1 3 2 4 2 2 2 3 3 2 4 2 3 3 2 3 2 4 1 4 1 4 1 1 1 4 2 4 1 1
- 1 4 1 4 1 3 3 2 2 4 2 4 1 4 1 4 4 2 3 3 2 3 1 4 2 3 2 3 1 4 1

Consideremos ahora un segundo ejemplo de rata en otro laberinto (Figura 9). Aquí hay tres cámaras $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$. La rata coge una puerta al azar en cada instante $t = 0, 1, 2, \ldots$. Supongamos que la rata es forzada a salir de su cámara electrificando el suelo. Si la rata está inicialmente en la cámara 2, entonces el vector de distribución inicial es $\boldsymbol{p}^{\mathrm{T}}(0) = (0, 1, 0) = \boldsymbol{e}_2^{\mathrm{T}}$. Pero si el proceso comenzase arrojando la rata al aire de manera que aterrizase al azar en una de las cámaras, entonces una distribución inicial razonable sería $\boldsymbol{p}^{\mathrm{T}}(0) = (0{,}37,\ 0{,}30,\ 0{,}33)$ porque en estas proporciones parecen estar las áreas de las cámaras. La matriz de transición de esta cadena de Markov es la matriz estocástica

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/3 & 0 & 2/3 \\ 1/3 & 2/3 & 0 \end{pmatrix}. \tag{5}$$

El digrafo ponderado de esta matriz, o de esta cadena de Markov, viene dado en la Figura 10. Cuatro simulaciones de esta cadena, hechas con MATLAB, en

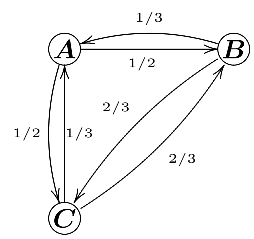


Figura 10: Digrafo ponderado del laberinto 2º.

las que el estado inicial se elige de acuerdo con el vector de distribución inicial $\boldsymbol{p}^{\mathrm{T}}(0) = (0{,}37,\ 0{,}30,\ 0{,}33)$ son:

- ***ABCBCBABCACACBCBCBCABABCBABCBACB***
- ***CABCBCBCABCACBCBACBCACBCBABACBAC***
- ***BACBABCBABACACACBCACBACABACABCAB***
- ***CBACBCACACBACBCACACBCBCABCACBABA***

Ejercicio 4.1 *Estimación de una matriz de las probabilidades de transición.* Una cadena de Markov de dos estados, 0 y 1, fue observada durante 50 transiciones. Los estados sucesivos ocupados por la cadena fueron los siguientes

$$\begin{array}{l} 0\,1\,0\,1\,1\,1\,0\,1\,0\,0\,1 \\ \quad 0\,1\,0\,1\,0\,1\,1\,1\,1\,0 \\ \quad 1\,0\,0\,1\,0\,1\,1\,0\,0\,1 \\ \quad 0\,1\,1\,0\,1\,0\,0\,0\,1\,0 \\ \quad 1\,1\,0\,0\,1\,1\,0\,1\,1\,0. \end{array}$$

Hagáse una estimación de la matriz de transición. *Solución.*

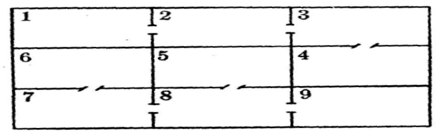
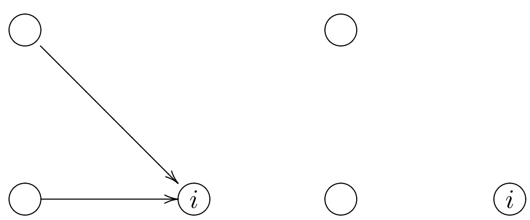


Figura 11: Laberinto 3º.

Ejercicio 4.2 Se coloca una rata blanca en el laberinto de la Figura 11. La rata se mueve al azar de un compartimento al otro; si hay k puertas elige cada una de ellas con la misma probabilidad $1/k$. En cada instante hace un cambio de compartimento. El estado del sistema es el número del compartimento en el que está la rata. Dibujar el digrafo de esta cadena de Markov y predecir la evolución a largo plazo. Solución.

Observación 4.1 Algunos digrafos nunca pueden ser digrafos de una cadena de Markov. Si un digrafo tiene un vértice i del que no salen flechas, no puede corresponder a una matriz estocástica (o cadena de Markov), pues tendría la fila i toda de ceros. La Figura 12 ilustra esta situación.



(a) Vértice i "colgado". (b) Grados salientes = 0.

Figura 12: Vértices "colgados".

Dos asuntos importantes surgen en el análisis markoviano respecto del comportamiento asintótico de la cadena. En otras palabras, reclaman nuestra atención las cuestiones siguientes.

- Describir del vector de distribución de probabilidad $\boldsymbol{p}^{\mathrm{T}}(k)$ de la etapa k-ésima para cualquier vector de distribución inicial $\boldsymbol{p}^{\mathrm{T}}(0)$.
- Determinar si existe $\lim_{k\to\infty} \boldsymbol{p}^{\mathrm{T}}(k)$, y si existe, determinar su valor.

- Si no existe una distribución límite, determinar la posibilidad de que exista el límite según Cesàro
$$\lim_{k\to\infty} \frac{\boldsymbol{p}^{\mathrm{T}}(0) + \boldsymbol{p}^{\mathrm{T}}(1) + \cdots + \boldsymbol{p}^{\mathrm{T}}(k-1)}{k}.$$
Si existe tal límite, interpretar su significado, y determinar su valor.

El vector $\boldsymbol{p}^{\mathrm{T}}(k)$ se puede describir fácilmente usando los teoremas elementales de las probabilidades. En particular, recordemos que $P(A \cup B) = P(A) + P(B)$ cuando A y B son sucesos disjuntos, y la probabilidad condicionada de que ocurra A dado que B ha sucedido se define por $P(A|B) = P(A \cap B)/P(B)$. Para determinar la componente j-ésima $p_j(1)$ del vector $\boldsymbol{p}^{\mathrm{T}}(1)$ para un vector dado $\boldsymbol{p}^{\mathrm{T}}(0)$, escribamos

$$\begin{aligned}
p_j(1) &= P(X_1 = S_j) = P\Big[X_1 = S_j \cap (X_0 = S_1 \cup \cdots \cup X_0 = S_n)\Big] \\
&= P\Big[(X_1 = S_j \cap X_0 = S_1) \cup \cdots \cup (X_1 = S_j \cap X_0 = S_n)\Big] \\
&= \sum_{i=1}^{n} P\Big[X_1 = S_j \cap X_0 = S_i\Big] = \sum_{i=1}^{n} P\Big[X_0 = S_i\Big] P\Big[X_1 = S_j | X_0 = S_i\Big] \\
&= \sum_{i=1}^{n} p_i(0) p_{ij} \text{ para } j = 1, \ldots, n.
\end{aligned}$$

En consecuencia, $\boldsymbol{p}^{\mathrm{T}}(1) = \boldsymbol{p}^{\mathrm{T}}(0)P$. Esto nos dice lo que podemos esperar en la etapa siguiente al comienzo con $\boldsymbol{p}^{\mathrm{T}}(0)$. Pero la propiedad de Markov de "falta de memoria" nos dice que el estado de cosas al final de la segunda etapa está determinado por el estado al final de la primera etapa— que es lo mismo que empezar otra vez pero con $\boldsymbol{p}^{\mathrm{T}}(1)$ como la distribución inicial. En otras palabras, que $\boldsymbol{p}^{\mathrm{T}}(2) = \boldsymbol{p}^{\mathrm{T}}(1)P$, y $\boldsymbol{p}^{\mathrm{T}}(3) = \boldsymbol{p}^{\mathrm{T}}(2)P$, etc. Por tanto,
$$\boldsymbol{p}^{\mathrm{T}}(k) = \boldsymbol{p}^{\mathrm{T}}(k-1)P = \boldsymbol{p}^{\mathrm{T}}(k-2)P^2 = \cdots = \boldsymbol{p}^{\mathrm{T}}(0)P^k,$$
y, de este modo, la distribución en la etapa k-ésima está determinada por el producto de la matriz fila $\boldsymbol{p}^{\mathrm{T}}(0)$ por la potencia k-ésima de la matriz P
$$\boldsymbol{p}^{\mathrm{T}}(k) = \boldsymbol{p}^{\mathrm{T}}(0)P^k. \tag{6}$$
Nótese que si adoptamos la notación $P^k =: [p_{ij}^{(k)}]$, y si tomamos $\boldsymbol{p}^{\mathrm{T}}(0) = \boldsymbol{e}_i^{\mathrm{T}}$ en (6), entonces obtenemos que $p_j(k) = p_{ij}^{(k)}$ para cada $i = 1, \ldots, n$, llegando así a la conclusión siguiente.

El elemento (i, j) de P^k representa la probabilidad de moverse de S_i a S_j en k pasos exactamente. Por esta razón, a P^k se le llama a menudo la **matriz de transición del paso k-ésimo.**

Para analizar las propiedades asintóticas de las cadenas de Markov, dividimos la clase de las matrices estocásticas P en cuatro categorías mutuamente exclusivas.

1. Irreducible con convergencia de la sucesión $P^k, k = 0, 1, \ldots$ (i.e. P es primitiva).

2. Irreducible con no convergencia de la sucesión $P^k, k = 0, 1, \ldots$ (i.e. P es imprimitiva).
3. Reducible con convergencia de la sucesión $P^k, k = 0, 1, \ldots$.
4. Reducible con no convergencia de la sucesión $P^k, k = 0, 1, \ldots$.

Cadenas de Markov irreducibles

Sea P la matriz de transición de una cadena de Markov irreducible de n estados $\{S_1, S_2, \ldots, S_n\}$ (i.e. P es una matriz $n \times n$ estocástica irreducible), y sea $\boldsymbol{\pi}^{\mathrm{T}}$ el vector de Perron por la izquierda de P. Los enunciados siguientes son ciertos para cualquier distribución inicial $\boldsymbol{p}^{\mathrm{T}}(0)$.

- La matriz de transición del paso k-ésimo es P^k porque el elemento (i, j) de P^k es la probabilidad de pasar de S_i a S_j en k pasos exactamente.
- El vector de distribución del paso k-ésimo viene dado por $\boldsymbol{p}^{\mathrm{T}}(k) = \boldsymbol{p}^{\mathrm{T}}(0)P^k$.
- Si P es primitiva y $\boldsymbol{e}$ el vector de todo unos, entonces
$$\lim_{k\to\infty} P^k = \boldsymbol{e}\boldsymbol{\pi}^{\mathrm{T}}, \qquad \lim_{k\to\infty} \boldsymbol{p}^{\mathrm{T}}(k) = \boldsymbol{\pi}^{\mathrm{T}}.$$

- Si P es imprimitiva, entonces
$$\lim_{k\to\infty} \frac{I + P + P^2 + \cdots + P^{k-1}}{k} = \boldsymbol{e}\boldsymbol{\pi}^{\mathrm{T}}$$
y
$$\lim_{k\to\infty} \frac{\boldsymbol{p}^{\mathrm{T}}(0) + \boldsymbol{p}^{\mathrm{T}}(1) + \cdots + \boldsymbol{p}^{\mathrm{T}}(k-1)}{k} = \boldsymbol{\pi}^{\mathrm{T}}.$$
- Tanto si P es primitiva como si es imprimitiva, la componente j-ésima π_j de $\boldsymbol{\pi}^{\mathrm{T}}$ representa la fracción de tiempo que está la cadena en S_j a largo plazo.
- $\boldsymbol{\pi}^{\mathrm{T}}$ se llama a menudo el **vector de distribución estacionario** de la cadena pues es el único vector de distribución que satisface $\boldsymbol{\pi}^{\mathrm{T}}P = \boldsymbol{\pi}^{\mathrm{T}}$.

Estos resultados están probados en [5, p.692–693].

4.1. Estados recurrentes, transitorios y absorbentes

En los ejemplos de cadenas de Markov considerados hasta aquí, hemos visto que en ocasiones la cadena pasa indefinidamente por todos los estados, o nunca pasa por algunos estados. Algunas otras situaciones son posibles al analizar el comportamiento a largo plazo de la cadena. En el ejemplo de la Figura 13 una mosca se pasea por el digrafo con las probabilidades de transición indicadas. En los estados 1 y 4 hay sendas arañas. Si la mosca cae en ellos, es atrapada por la araña y queda allí absorbida el resto del tiempo. Si la mosca empieza en 2 ó 3, estará algunos pasos en estos estados, pero en un momento determinado caerá en 1 ó 4 para no retornar. Los estados 2 y 3 son transitorios.

Estas observaciones van a ser substanciadas con las definiciones siguientes. Llamaremos a los estados $1, 2, \ldots, n$ en vez de $S_1, S_2, \ldots, S_n$, por comodidad.

Definición 4.1 Un estado j se llama **accesible** desde el estado i si es posible moverse en un número finito de etapas desde el estado i al estado j. Indicamos esta situación con la notación $i \to j$.

Una definición equivalente es que para algún momento $k \geq 1$ la probabilidad de paso $p_{ij}^{(k)}$ es positiva. Por el Teorema 3.2, esto ocurre si y solo si hay un paseo

$$i, i_1, i_2, \ldots, i_{k-2}, i_{k-1}, j$$

en el digrafo formado por arcos de probabilidad positiva.

Proposición 4.1 *Si $i \to j$ y $j \to k$, entonces $i \to k$.*

Por tanto, la relación $\to$ es transitiva.

Definición 4.2 Se dice que el estado i **comunica** con el estado j si $i \to j$ y $j \to i$. En cuyo caso escribimos $i \leftrightarrow j$ para denotar esta relación.

Proposición 4.2 *La relación de comunicación $\leftrightarrow$ es simétrica y transitiva. Es decir, para estados cualesquiera i, j y k,*

$$i \to j \quad \textit{implica} \quad j \to k,$$
$$i \to j \quad y \quad j \to k \quad \textit{implica} \quad i \to k.$$

Pero la relación $\leftrightarrow$ no es reflexiva necesariamente. Para que así lo fuera tendríamos que tener que $\forall i,\ i \leftrightarrow i$. Cosa imposible a veces. Véanse, por ejemplo, los estados 1 y 2 en el digrafo de la Figura 14.

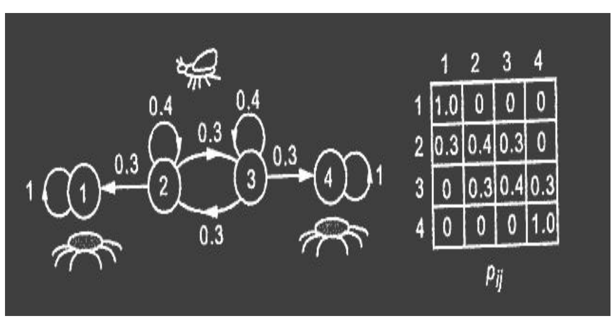


Figura 13: Mosca y arañas

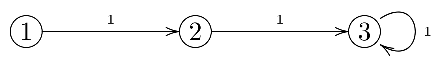


Figura 14: No reflexiva.

Dado un estado i de una cadena de Markov, se define su **clase comunicante** $C(i)$ como el subconjunto de los estados j en la cadena que comunican con i; con símbolos,

$$C(i) := \{j \in \{1, 2, \ldots, n\} : i \leftrightarrow j\}.$$

Puede suceder que la clase $C(i)$ sea vacía (es decir, i no se comunica con ningún estado, ni siquiera consigo mismo). En este caso se llama a i estado *sin retorno* . Si $C(i)$ es no vacía, entonces i pertenece a $C(i)$; para ver esto, sea k un elemento cualquiera de $C(i)$, entonces $i \leftrightarrow k$, pero como $\leftrightarrow$ es simétrica, $k \leftrightarrow i$; y dado que $\leftrightarrow$ es también transitiva deducimos que $i \leftrightarrow i$. Por tanto, si $C(i)$ es no vacía $i \in C(i)$. Un estado que se comunica consigo mismo se denomina *con retorno.*

Definición 4.3 Se dice que un subconjunto $\mathcal{C}$ de estados no vacío es una **clase comunicante** si, para cierto estado i, $\mathcal{C}$ es igual a $C(i)$.

Proposición 4.3 *Si $\mathcal{C}_1$ y $\mathcal{C}_2$ son clases comunicantes, entonces, o bien $\mathcal{C}_1 = \mathcal{C}_2$, o bien $\mathcal{C}_1 \cap \mathcal{C}_2 = \varnothing$.*

En la Figura 15 tenemos dos cadenas de Markov: (a) y (b). En la cadena (a) las clases comunicantes son $\{1\}, \{2, 3\}$ y $\{4, 5\}$. En la cadena (b) las clases comunicantes son $\{1\}, \{2, 3\}$ y $\{6\}$. En la cadena de la Figura 14 la única clase comunicante es $\{3\}$.

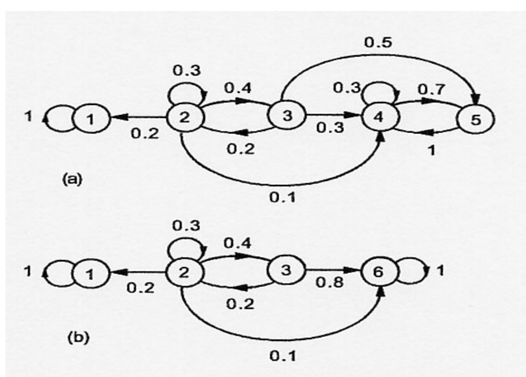


Figura 15: Clases comunicantes

De ahora en adelante por **clase** querremos significar **clase comunicante**, salvo que se especifique más. Se dice que el estado $i \notin \beta$ **tiene acceso a la clase** β si $i \to j$ para algún $j \in \beta$. Una clase α tiene **acceso** a una clase β (se escribe $\alpha \to \beta$) si $i \to j$ para algunos $i \in \alpha$ y $j \in \beta$. También se dice en este caso que la clase β **es accesible desde** la clase α. Una clase se llama **final** si no tiene acceso a ninguna otra clase. Si una clase final contiene un único estado entonces el estado se llama **absorbente**. Un estado $j \notin \alpha$ se dice **accesible desde la clase** α si $i \to j$ para algún $i \in \alpha$. Si no existe tal $i \in \alpha$ se dice que el estado j **no es accesible desde la clase** α.

En la la cadena (a) de la Figura 15, la clase $\{2,3\}$ tiene acceso a las clases $\{1\}$ y $\{4,5\}$. Estas dos últimas clases son clases finales. El estado 1 es absorbente.

Definición 4.4 Un estado i se llama **recurrente (persistente, ergódico)** si para todo j tal que $i \to j$, también es $j \to i$. Esto es, si para todo j accesible desde i, i es accesible desde j.

Así pues, a partir de cualquier momento futuro hay siempre alguna probabilidad de regresar a i y, mediando suficiente tiempo, esto ocurrirá ciertamente.

Definición 4.5 Un estado i se llama **transitorio** si no es recurrente.

Pasado bastante tiempo, el estado transitorio i ya no podrá ser visitado. Por lo que un estado transitorio solo será visitado un número finito de veces.

Se dice que una clase $\mathcal{C}$ es **recurrente (resp. transitoria)** si algún estado de $\mathcal{C}$ es recurrente (resp. transitorio).

Proposición 4.4 *Sea $\mathcal{C}$ una clase. Entonces si $\mathcal{C}$ es recurrente, todos los estados de $\mathcal{C}$ son recurrentes. Si $\mathcal{C}$ es transitoria, entonces todos los estados de $\mathcal{C}$ son transitorios.*

En la cadena (a) de la Figura 15 la clase $\{2,3\}$ es transitoria, las clases $\{1\}$ y $\{4,5\}$ son recurrentes.

Descomposición de una cadena de Markov

El conjunto de los estados puede ser descompuesto en la unión disjunta de una o más clases recurrentes y un subconjunto de estados transitorios.

- **Periodicidad**

Una situación importante en una clase de recurrencia es la presencia o ausencia de un cierto patrón periódico en la manera en que son visitados los estados de la clase. Una clase de recurrencia α se llama **periódica** si existe una partición $E_1, E_2, \ldots, E_d$ de α en $d(>1)$ subconjuntos no vacíos disjuntos tal que todas las transiciones desde un subconjunto conducen al subconjunto siguiente; véase la Figura 16. Con mayor precisión,

$$\text{si} \quad i \in E_k \quad \text{y} \quad p_{ij} > 0, \text{entonces} \quad \begin{cases} j \in E_{k+1}, & \text{si } k = 1, \ldots, d-1, \\ j \in E_1, & \text{si } k = d. \end{cases}$$

Una clase de recurrencia que no sea periódica se llama **aperiódica**..

Así pues, en una clase recurrente periódica, nos movemos a través de la sucesión de subconjuntos $E_1, E_2, \ldots, E_d$ ordenadamente, y tras d pasos volvemos al mismo subconjunto. Consideremos el ejemplo de la Figura 16, donde hemos ponderado los arcos del digrafo de acuerdo con la matriz de transición P dada en (7)

$$P = \left(\begin{array}{cc|ccc|c} 0 & 0 & 0{,}7 & 0{,}3 & 0 & 0 \\ 0 & 0 & 0{,}6 & 0 & 0{,}4 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0{,}8 & 0{,}2 & 0 & 0 & 0 & 0 \end{array} \right). \tag{7}$$

La matriz P es irreducible, no primitiva con índice de imprimitividad 3 y está en la forma de Frobenius.

Tomamos el primer estado eligiendo con igual probabilidad alguno de los 6 estados: 1,2,3,4,5 ó 6.

6 1 3 6 2 5 6 1 3 6 1 3 6 1 4 6 1 4 6 1 3 6 1 4 6 2 3 6 1 4 6 1 3 ...

E_3 E_1 E_2 E_3 E_1 E_2 E_3 E_1 E_2 E_3 E_1 E_2 E_3 E_1 E_2 E_3 E_1 E_2 E_3 E_1 E_2 E_3 E_1 E_2 E_3 E_1 E_2 E_3 E_1 E_2 E_3 E_1 E_2 ...

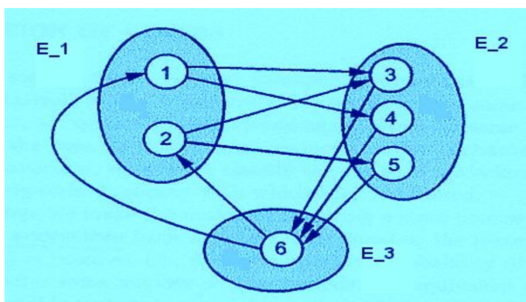


Figura 16: Una clase de recurrencia de periodo 3.

Vemos arriba la cadena de estados obtenida mediante simulación con MATLAB y la sucesión de los subconjuntos E_1, E_2, E_3 a los que pertenecen. Se observará la repetición periódica E_3, E_1, E_2 de los subconjuntos.

Nótese que dados una clase recurrente periódica, un entero positivo m y un estado j de la clase, debe existir algún estado i tal que $p_{ij}^{(m)} = 0$. La razón es que, por la definición de periodicidad, se sigue que los estados están agrupados en subconjuntos $E_1, E_2, \ldots, E_d$, y el subconjunto al que pertenece j puede ser alcanzado en m pasos desde los estados situados en uno solo de los subconjuntos. Así pues, una manera de verificar la aperiodicidad de una clase recurrente α, es comprobar si hay algún instante especial m_0 y un estado especial k_0 que pueda ser alcanzado en el paso m_0-ésimo desde todos los estados (iniciales) de α, i.e., $p_{ik_0}^{(m_0)} > 0$ para todo $i \in \alpha$. Como ejemplo, consideremos la cadena de la Figura 17 que solo tiene una clase de recurrencia. El estado $k_0 = 2$ puede ser alcanzado en el paso $m_0 = 2$ empezando desde cualquier estado; por tanto, esta clase recurrente es aperiódica.

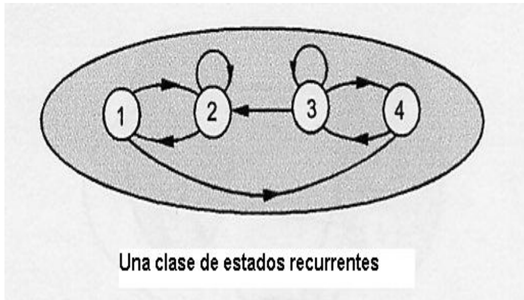


Figura 17: Una clase recurrente aperiódica.

Un enunciado recíproco es verdadero también: si una clase recurrente no es periódica, entonces un instante m_0 y un estado especial k_0 con las propiedades antes mencionadas pueden ser encontrados siempre.

Periodicidad

Consideremos una clase recurrente α.

- La clase se llama **periódica** si sus estados pueden ser agrupados en $d > 1$ subconjuntos no vacíos disjuntos a pares $E_1, E_2, \ldots, E_d$, de manera tal que todas las transiciones desde E_k conducen a E_{k+1} (o a E_1 si $k = d$).

Sea i un estado con retorno (i.e. tal que $i \leftrightarrow i$). El **periodo** $\mathrm{d}(i)$ de i se define como el máximo común divisor del conjunto de enteros k tales que $p_{ii}^{(k)} > 0$; con símbolos

$$\mathrm{d}(i) := \mathrm{mcd}\{k \in \mathbb{N} : p_{ii}^{(k)} > 0\}.$$

Se dice que un estado es aperiódico si tiene periodo 1.

El periodo $\mathrm{d}(i)$ se caracteriza por la propiedad siguiente: es posible volver a i en m pasos si y solo si $\mathrm{d}(i)$ divide a m. De hecho, el conjunto $\{k \in \mathbb{N} : p_{ii}^{(k)} > 0\}$ es igual al conjunto $\overset{\circ}{\widehat{\mathrm{d}(i)}}$ de números naturales múltiplos de $\mathrm{d}(i)$. También puede demostrarse que dos estados que se comunican tienen el mismo periodo; y que si una clase recurrente tiene un estado de periodo $d > 1$, entonces la clase es periódica de periodo d. Una clase recurrente es aperiódica si y solo si tiene un estado i tal que $p_{ii} > 0$; pues $p_{ii}^{(1)} = p_{ii}$, lo que implica que $\{k \in \mathbb{N} : p_{ii}^{(k)} > 0\} = \mathbb{N}$. Sea α un subconjunto de $\{1, 2, \ldots, n\}$, y sea $M \in \mathbb{R}^{n \times n}$. Por $M[\alpha]$ denotamos la submatriz principal de M formada por las filas y columnas correspondientes a α.

Aperiodicidad

Consideremos α una clase recurrente.

- La clase α es **aperiódica** (no periódica) si y solo si existen un instante m_0 y un estado especial k_0 de la clase, tales que para todo $i \in \alpha$, $p_{ik_0}^{(m_0)} > 0$.
- La clase α es **aperiódica** si y solo si existe un entero $m_0 \geq 1$ tal que la submatriz $(P^{m_0})[\alpha]$, de P^{m_0}, tiene una columna con todos sus elementos > 0. Este criterio es fácil de aplicar analizando las potencias sucesivas de la matriz de transición $P, P^2, P^3, \ldots$

Ejemplo 4.1 En al Figura 18 damos el digrafo de la matriz de transición P de una cadena de Markov de 11 estados, que constituyen una sola clase recurrente α.

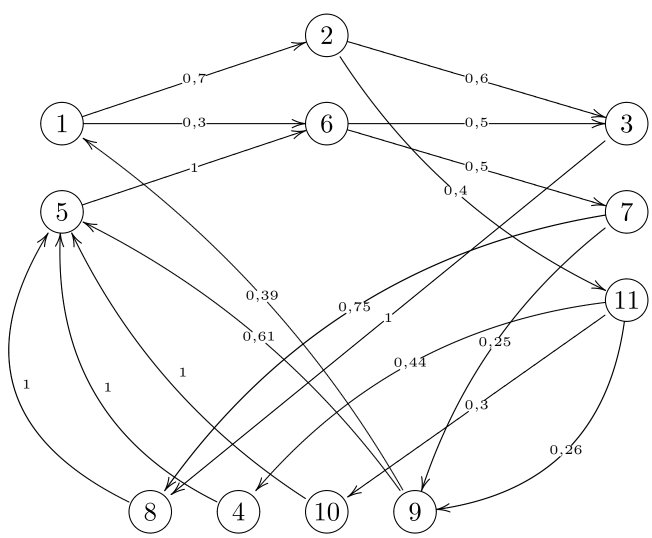


Figura 18: Clase recurrente α, periódica de periodo 4.

La matriz de transición asociada a esta cadena es obviamente

$$P = \begin{pmatrix} 0 & 0{,}7 & 0 & 0 & 0 & 0{,}3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0{,}6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0{,}4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0{,}5 & 0 & 0 & 0 & 0{,}5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0{,}75 & 0{,}25 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0{,}39 & 0 & 0 & 0 & 0{,}61 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0{,}44 & 0 & 0 & 0 & 0 & 0{,}26 & 0{,}3 & 0 \end{pmatrix}. \quad (8)$$

Como el digrafo de esta matriz es fuertemente conexo, la matriz es irreducible. Además, el polinomio característico de P es

$$\lambda^{11} - 0{,}9943\lambda^7 - 0{,}0057\lambda^3.$$

Por el Teorema 3.8 se sigue que el índice de imprimitividad d de P viene dado por $d = \text{mcd}\{11-7, 7-3\} = 4$. Por tanto, la matriz P no es primitiva, y, en consecuencia, la clase recurrente α es periódica de periodo 4. Los subconjuntos E_1, E_2, E_3, E_4 en que se descompone α y por los que la cadena transita periódicamente en el orden

$$\cdots E_1 \to E_2 \to E_3 \to E_4 \to E_1 \to \cdots$$

son

$$E_1 = \{1,5\},\ E_2 = \{2,6\},\ E_3 = \{3,7,11\},\ E_4 = \{8,4,10,9\}.$$

La permutación σ de 1,2,3,4,5,6,7,8,9,10,11 que cambia las líneas de P en su forma de Frobenius, P_{F}, es

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 1 & 5 & 2 & 6 & 3 & 7 & 11 & 8 & 4 & 10 & 9 \end{pmatrix},$$

puesto que si reescribimos los vértices de la Figura 18 de acuerdo con esta permutación, obtenemos la Figura 19, donde utilizamos la convención $\sigma_i := \sigma(i)$ para $i = 1, \ldots, 11$.

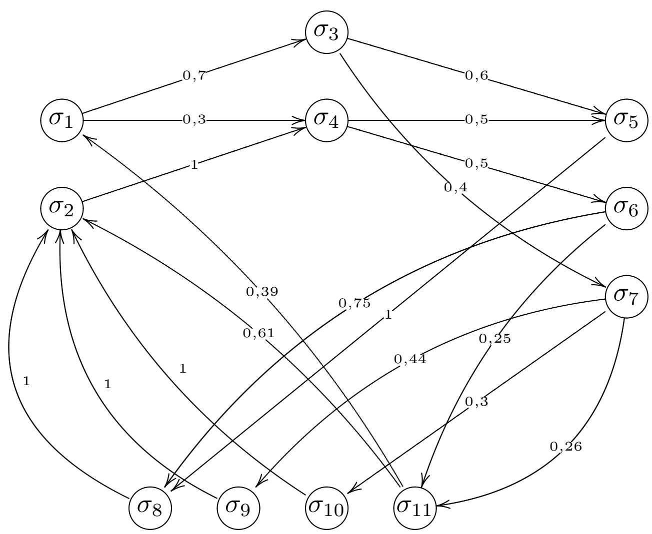


Figura 19: Vértices de la clase α renombrados .

La matriz $M(\sigma)$, calculada con Matlab, es

$$M(\sigma) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix};$$

por consiguiente, la matriz en forma de Frobenius $P_{\mathrm{F}} = M(\sigma)^{\mathrm{T}} P M(\sigma)$ viene dada por

$$P_{\mathrm{F}} = \left(\begin{array}{cc|cc|ccc|cccc} 0 & 0 & 0{,}70 & 0{,}30 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0{,}6 & 0 & 0{,}4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0{,}5 & 0{,}5 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0{,}75 & 0 & 0 & 0{,}25 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0{,}44 & 0{,}3 & 0{,}26 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0{,}39 & 0{,}61 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right). \tag{9}$$

4.2. Cadenas de Markov reducibles

Una cadena de Markov se llama **reducible** si su matriz de transición es reducible. Estamos interesados en predecir el comportamiento a largo plazo de estas cadenas más generales que las irreducibles. Empezamos considerando un ejemplo; se trata de la cadena cuyo digrafo viene dado en la Figura 20. Omitimos el peso p_{ij} de cada arco (i, j), por comodidad.

Esta cadena tiene quince estados a los que llamaremos $1, 2, \ldots, 15$. Los estados que llevan al lado una t son transitorios, los que llevan una r son recurrentes, y los que llevan una a son absorbentes. Hay un solo estado absorbente que es el 8. Hay 5 clases comunicantes; de ellas, hay 3 recurrentes, 2 transitorias. También hay un estado sin retorno: el 3; que por tanto no pertenece a ninguna clase. Mirando la figura es fácil averiguar el comportamiento a largo plazo de la cadena. Por ejemplo, si el estado inicial es 1 ó 9, como hay una probabilidad de saltar del 9 al 3, en determinado instante, la cadena salta a 3 y ya no retorna a la clase transitoria $\{1, 9\}$. Desde 3 pasa *seguramente* a 14, i.e. con probabilidad 1. De 14 con igual probabilidad salta a 5. En 5 tiene dos posibilidades, saltar a 11 o a 13; si saltara a 11, la cadena sería atrapada en la clase recurrente $\{11, 10, 4, 2\}$ y ya

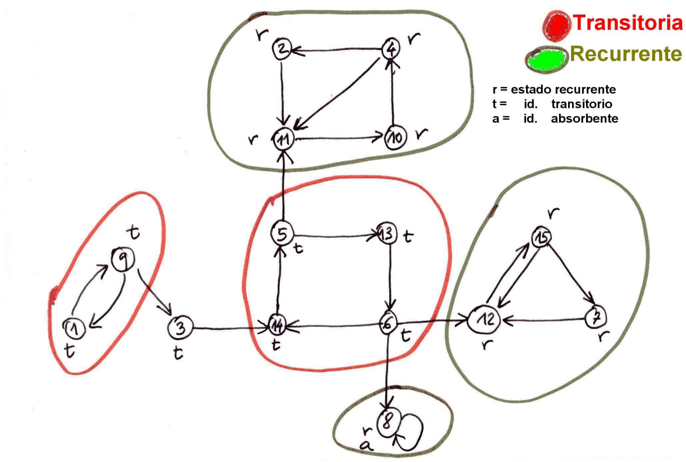


Figura 20: Cadena reducible.

no saldría de allí; si saltara a 13, iría *seguramente* a 6 y desde este estado podría, o bien, pasar a la clase recurrente $\{12, 15, 7\}$, o bien, al estado absorbente 8, o bien regresar a 14 y vuelta a empezar. Es claro que entrará ciertamente en alguna de las tres clases de recurrencia, pues el suceso tiene probabilidad positiva, por lo que tras un número *suficientemente grande* de pasos entrará con probabilidad 1. También es evidente que una vez que entre la cadena en una de las tres clases recurrentes no saldrá de ella. Hay otras posibilidades, como arrancar de 3; en este caso, es *seguro* que no visitará ni 3, ni 9, ni 1. Si comienza en 7, seguro que no saldrá de la clase de 7, etc. Ninguna de las tres clases recurrentes es periódica.

Cuestiones básicas

Llamando P a la matriz de transición de la cadena, las cuestiones básicas son las siguientes: ¿existe el $\lim_{k\to\infty} \boldsymbol{p}^{\mathrm{T}}(k)$? Si existe, ¿depende de $\boldsymbol{p}^{\mathrm{T}}(0)$? ¿Existe el $\lim_{k\to\infty} P^k$? Si no existen estos dos límites, ¿cuál es el comportamiento asintótico de las dos sucesiones (de vectores y matrices)? ¿Oscilan alrededor de un número finito de valores de adherencia (o límites subsecuenciales)? Si oscilan, ¿lo hacen con alguna periodicidad, o con algún otro patrón?

Sea la matriz Q

$$P \sim Q := \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1r} & P_{1,r+1} & P_{1,r+2} & \cdots & P_{1m} \\ 0 & P_{22} & \cdots & P_{2r} & P_{2,r+1} & P_{2,r+2} & \cdots & P_{2m} \\ \vdots & & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & P_{rr} & P_{r,r+1} & P_{r,r+2} & \cdots & P_{rm} \\ 0 & 0 & \cdots & 0 & P_{r+1,r+1} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & P_{r+2,r+2} & \cdots & 0 \\ \vdots & & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & P_{mm} \end{pmatrix}, \tag{10}$$

la forma normal de la matriz reducible P, donde cada bloque diagonal $P_{11}, \ldots, P_{rr}$ es, o bien irreducible, o bien $(0)_{1\times 1}$, $P_{r+1,r+1}, \ldots, P_{mm}$ son irreducibles (estos bloques no pueden ser cero porque la suma de las filas de cada uno es igual a 1), y cada fila de bloques

$$P_{i,i+1}, P_{i,i+2}, \ldots, P_{i,r+1}, P_{i,r+2}, \ldots, P_{im} \quad (i = 1, \ldots, r)$$

tiene al menos un bloque P_{ij} diferente de cero. Además, como dijimos al permutar las líneas de una matriz los digrafos ponderados de P y de la matriz Q son isomorfos. Si llamamos σ a la permutación de $1, 2, \ldots, n$ que transforma P en Q se tiene que las líneas (i.e. filas y columnas) de Q se disponen según la ordenación

$$\sigma(1), \sigma(2), \ldots, \sigma(n).$$

Si el bloque $P_{11} \neq 0$ es de orden n_1, entonces $\{\sigma(1), \sigma(2), \ldots, \sigma(n_1)\}$ constituye la clase a él asociada. Si $P_{22} \neq 0$ tiene orden n_2, entonces $\{\sigma(n_1+1), \ldots, \sigma(n_1+n_2)\}$ el la clase a él asociada, etc. Si P_{mm} es de orden n_m, su clase asociada es $\{\sigma(n_1 + \cdots + n_{m-1} + 1), \ldots, \sigma(n)\}$. Si para algún $j \in \{1, \ldots, r\}, P_{jj} = (0)_{1\times 1}$, entonces el estado $\sigma(n_1 + \cdots + n_j + 1)$ es *sin retorno*, y **no** constituye una clase comunicante. Se puede demostrar que $\varrho(P_{kk}) < 1$ para cada $k = 1, \ldots, r$. Los estados transitorios son reflejados en las r primeras filas de bloques. Cada bloque no nulo $P_{ii}, i = 1, \ldots, r$ está asociado a una clase transitoria:$\{\sigma(n_1+\cdots+n_{i-1}+1), \ldots, \sigma(n_1+\cdots+n_{i-1}+n_i)\}$ Cada una de las matrices $P_{r+1,r+1}, \ldots, P_{mm}$ corresponde a una clase recurrente.

Llamemos

$$T_{11} = \begin{pmatrix} P_{11} & \cdots & P_{1r} \\ & \ddots & \vdots \\ & & P_{rr} \end{pmatrix}, \quad T_{12} = \begin{pmatrix} P_{1,r+1} & \cdots & P_{1m} \\ \vdots & & \vdots \\ P_{r,r+1} & \cdots & P_{rm} \end{pmatrix},$$

y

$$T_{22} = \begin{pmatrix} P_{r+1,r+1} & & \\ & \ddots & \\ & & P_{mm} \end{pmatrix}.$$

Entonces

$$P \sim Q := \begin{pmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{pmatrix}.$$

Como existe una matriz de permutación M tal que $Q = M^{\mathrm{T}}PM$, las cuestiones planteadas más arriba equivalen a ser contestadas para la cadena de Markov cuya matriz de transición es Q.

Cadenas de Markov reducibles

Si $\boldsymbol{\pi}_j^{\mathrm{T}}$ es el vector de Perron por la izquierda para $P_{jj}, r+1 \leq j \leq m$, entonces

$$\lim_{k\to\infty} \frac{I + Q + \cdots + Q^{k-1}}{k} = \begin{pmatrix} 0 & (I - T_{11})^{-1}T_{12}E \\ 0 & E \end{pmatrix},$$

donde

$$E = \begin{pmatrix} \boldsymbol{e\pi}_{r+1}^{\mathrm{T}} & & \\ & \ddots & \\ & & \boldsymbol{e\pi}_m^{\mathrm{T}} \end{pmatrix}.$$

Además, el $\lim_{k\to\infty} Q^k$ existe si y solo si las matrices estocásticas $P_{r+1,r+1}, \ldots, P_{mm}$ de (10) son primitivas, en cuyo caso

$$\lim_{k\to\infty} Q^k = \begin{pmatrix} 0 & (I - T_{11})^{-1}T_{12}E \\ 0 & E \end{pmatrix}.$$

4.3. Cadenas de Markov cualesquiera

Los resultados citados a continuación son válidos tanto si la matriz de transición P es reducible o irreducible. Sobrentenderemos que la forma normal de una matriz irreducible consta de un solo bloque diagonal $P_{r+1,r+1}$ igual a ella misma. Así pues, todo el análisis se basa en la forma normal (10) de una matriz racional. Llamemos $\psi(\lambda)$ al polinomio mínimo de P, y

$$C(\lambda) := \psi(\lambda)(\lambda I - P)^{-1}$$

a la **matriz adjunta reducida** de P. Sean $\Delta(\lambda)$ el polinomio característico de P y

$$B(\lambda) := \Delta(\lambda)(\lambda I - P)^{-1}$$

la **matriz adjunta** de P.

Una matriz estocástica P se llama **propia** si no tiene valores propios λ_0 de módulo 1, salvo $\lambda_0 = 1$, y se llama **regular** si es propia y si 1 es un valor propio simple de P. Se puede demostrar que una matriz estocástica P es propia si y solo si las matrices las matrices $P_{r+1,r+1}, \ldots, P_{mm}$ de (10) son primitivas. Para que P sea una matriz regular es preciso además que $r + 1 = m$.

Sobre $\lim_{k\to\infty} P^k$

(1) Existe el límite $L = (l_{ij}) = \lim_{k\to\infty} P^k$ si y solo si las matrices $P_{r+1,r+1}, \ldots, P_{mm}$ de (10) son primitivas (matriz P **propia**). En cuyo caso, $L = \frac{C(1)}{\psi'(1)}$.

La probabilidad $p_{ij}^{(k)}$ de ir del estado i al j en k pasos tiende al valor l_{ij} cuando $k \to \infty$. Por ello, si j es un estado transitorio, entonces $\lim_{k\to\infty} p_{ij}^{(k)} = 0$ para todo estado i; si j es un estado recurrente, entonces $\lim_{k\to\infty} p_{ij}^{(k)} > 0$ para todo estado i de su misma clase recurrente, y es $= 0$ para los demás estados i.

(2) Existe el $L := \lim_{k\to\infty} P^k$ y cada columna de L es proporcional al vector columna $[1, 1, \ldots, 1]^{\mathrm{T}}$ si y solo si $P_{r+1,r+1}$ es primitiva y $r+1 = m$ en (10), i.e. en T_{22} solo hay un bloque diagonal primitivo (matriz P **regular**).

En este caso, $L = \frac{C(1)}{\psi'(1)} = \frac{B(1)}{\Delta'(1)}$.

(2') El resultado (2) nos dice que la probabilidad $p_{ij}^{(k)}$ de ir del estado i al j en k pasos tiende a ser independiente del estado inicial i cuando $k \to \infty$.

(3) $L > 0$ si y solo si la matriz P es irreducible primitiva (i.e. la única clase de recurrencia es aperiódica).

Sobre $\lim_{k\to\infty} \boldsymbol{p}^{\mathrm{T}}(k)$

- Existe el límite $\ell := \lim_{k\to\infty} \boldsymbol{p}^{\mathrm{T}}(k)$ si y solo si las matrices $P_{r+1,r+1}, \ldots, P_{mm}$ de (10) son primitivas. Además, en este caso, para todo estado transitorio $\sigma(j), j = 1, \ldots, n_1 + \cdots + n_r$, se tiene que $\lim_{k\to\infty} p_{\sigma(j)}(k) = 0$.
- Existe el $\lim_{k\to\infty} \boldsymbol{p}^{\mathrm{T}}(k)$ y es independiente del vector inicial $\boldsymbol{p}^{\mathrm{T}}(0)$ si y solo si $P_{r+1,r+1}$ es primitiva y $r + 1 = m$ en (10), i.e. en T_{22} solo hay un bloque diagonal primitivo.
- $\ell > 0$ y es independiente de $\boldsymbol{p}^{\mathrm{T}}(0)$ si y solo si P es irreducible primitiva.

Consideraremos ahora el caso más general de cadena. Sean $h_{r+1}, \ldots, h_m$ los índices de imprimitividad de las matrices $P_{r+1,r+1}, \ldots, P_{mm}$. Sea h el mínimo común múltiplo de $h_{r+1}, \ldots, h_m$. Entonces la matriz P^h es propia; además las matrices $P, P^2, \ldots, P^{h-1}$ no son propias. Llamaremos al número h el **periodo de la cadena de Markov.**

Ya que P^h es propia, existe el

$$\lim_{k\to\infty} P^{hk} =: L_h.$$

Por lo que para todo $s = 0, 1, \ldots, h-1$, existen los límites

$$\lim_{k\to\infty} P^{s+hk} =: L_s = P^s L_h.$$

Comportamiento asintótico general

Sea una cadena de Markov de periodo h. Entonces las sucesiones

$$I, P, P^2, P^3, \ldots \quad \text{y} \quad \boldsymbol{p}^{\mathrm{T}}(0), \boldsymbol{p}^{\mathrm{T}}(1), \boldsymbol{p}^{\mathrm{T}}(2), \ldots$$

se pueden partir en h subsucesiones mutuamente disjuntas, cada una de las cuales tiene un límite $L_s = P^s L_h$, $\boldsymbol{\ell}_s = \boldsymbol{p}^{\mathrm{T}}(s)\boldsymbol{\ell}_h$, para $s = 0, 1, \ldots, h-1$; siendo $\boldsymbol{\ell}_h := \lim_{k\to\infty} \boldsymbol{p}^{\mathrm{T}}(hk)$. Los términos de estas dos sucesiones oscilan periódicamente alrededor de estos h valores de adherencia.

- Las sucesiones $P, P^2, P^3, \ldots$ y $\boldsymbol{p}^{\mathrm{T}}(0), \boldsymbol{p}^{\mathrm{T}}(1), \boldsymbol{p}^{\mathrm{T}}(2), \ldots$ son sumables Cesàro a $\tilde{P}$ y $\tilde{\boldsymbol{p}}^{\mathrm{T}}$, respectivamente. Las componentes $\sigma(1), \ldots, \sigma(n_1 + \cdots + n_r)$ de $\tilde{\boldsymbol{p}}^{\mathrm{T}}$, correspondientes a los estados transitorios, son cero.
- La suma Cesàro $\tilde{\boldsymbol{p}}^{\mathrm{T}}$ es independiente de $\boldsymbol{p}^{\mathrm{T}}(0)$ si y solo si la matriz P es regular. El vector $\tilde{\boldsymbol{p}}$ está determinado por la ecuación $P^{\mathrm{T}}\tilde{\boldsymbol{p}} = \tilde{\boldsymbol{p}}$.

5. Clases de una matriz no negativa

Las definiciones y resultados de la subsección 4.1 sobre las clases comunicantes (transitorias, recurrentes, periódicas, etc.) son trasladables sin más a una matriz cuadrada **cualquiera** no negativa $n \times n$ A. Así lo entenderemos en adelante. Una clase α tal que $\varrho(A[\alpha]) = \varrho(A)$ se llama *clase básica* . Los bloques A_{ii} no nulos de la forma normal (1) de la matriz A se corresponden biunívocamente con las clases comunicantes y si son $(0)_{1\times 1}$ con los estados sin retorno. Los bloques aislados, i.e. los bloques $A_{r+1,r+1}, \ldots, A_{mm}$ se corresponden con las clases finales. La clase correspondiente al bloque A_{ii} es básica si y solo si $\varrho(A_{ii}) = \varrho(A)$. Como el conjunto valores propios de una matriz triangular superior por bloques, es igual a la unión de los espectros de los bloques diagonales, es obvio que debe haber una clase básica al menos.

Las clases comunicantes α de A corresponden a las componentes fuertemente conexas del digrafo cuya matriz de adyacencia es A. El subconjunto β de $\{1, 2, \ldots, n\}$ es una clase comunicante si y solo si la submatriz principal $A[\beta]$ es irreducible. Una matriz cuadrada M es irreducible si y solo si las **potencias booleanas** sucesivas de la matriz Boole(M), obtenida al sustituir los elementos no nulos de M por 1, se estacionan en la matriz con **todos** sus elementos iguales a 1. Esto será utilizado para escribir una función de Matlab, `normform.m`, que nos dará estas clases.

El segundo de los dos teoremas siguientes relaciona ciertas matrices no negativas y las matrices estocásticas. De hecho, lo que viene a decir es que una matriz no negativa A que tiene un vector propio $\boldsymbol{x} > \mathbf{0}$ asociado a su radio espectral r, es diagonalmente semejante a una matriz estocástica P, salvo el factor r.

Teorema 5.1 *El radio espectral r de la matriz $n \times n$, $A \geq 0$ tiene asociado un vector propio positivo si y solo si las clases finales de A son precisamente sus clases básicas.*

Demostración. Véase [1, Theorem 3.10, p. 40]

Teorema 5.2 *Sea A una matriz no negativa $n \times n$ con radio espectral r. Si existe un vector propio (por la derecha) positivo $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ asociado a r y escribimos $X = \operatorname{diag}(x_1, x_2, \ldots, x_n)$, entonces*

$$A = rXPX^{-1},$$

donde P es una matriz estocástica.

Demostración. [4, Theorem 2, p. 548]

6. Matlab

Programa cadena.m

```
Cadena=[ ];
x=rand(1);
% Elección del estado inicial
if x < 0.25,
    i=1;
elseif x < 0.50,
    i=2;
elseif x < 0.75,
    i=3;
else
    i=4;
end
Cadena=[Cadena,i]';
% Elección de cada uno de los estados siguientes
for k=1:30
   j=salto(i);
   Cadena=[Cadena;j];
   i=j;
end
% Cadena al final de 30 saltos
Cadena
% Fin del fichero cadena.m
```

Función salto.m

```
function j=salto(i)
% j=salto(i) devuelve el estado j al que se pasa
%  aleatoriamente desde el estado i.
%
% Matriz de transición
P=[
     1/5  0  1/5  3/5
      0 1/4  1/2  1/4
     1/4  1/2  1/4  0
     3/5  1/5  0  1/5];
% Elección del estado al que se saltará
switch i
  case 1
   a=rand(1);
    if a<P(1,1)
       j=1;
    elseif P(1,1)<=a & a<P(1,1)+P(1,2)
       j=2;
    elseif P(1,1)+P(1,2)<=a & a<P(1,1)+P(1,2)+P(1,3)
       j=3;
    else
       j=4;
    end

  case 2
   a=rand(1);
    if  a<P(2,1)
       j=1;
    elseif P(2,1)<=a & a<P(2,1)+P(2,2)
       j=2;
    elseif P(2,1)+P(2,2)<=a & a<P(2,1)+P(2,2)+P(2,3)
       j=3;
    else
       j=4;
    end

  case 3
   a=rand(1);
    if a<P(3,1)
       j=1;
    elseif P(3,1)<=a & a<P(3,1)+P(3,2)
       j=2;
    elseif P(3,1)+P(3,2)<=a & a<P(3,1)+P(3,2)+P(3,3)
       j=3;
```

```
    else
       j=4;
    end

  case 4
   a=rand(1);
    if a<P(4,1)
       j=1;
    elseif P(4,1)<=a & a<P(4,1)+P(4,2)
       j=2;
    elseif P(4,1)+P(4,2)<=a & a<P(4,1)+P(4,2)+P(4,3)
       j=3;
    else
       j=4;
    end
end
% Fin del fichero salto.m
```

Matriz de la permutación σ

```
% Fichero sigma.m
sigma=[1 5 2 6 3 7 11 8 4 10 9];
for i=1:11
    for j=1:11
        if i==sigma(j),
                    M(i,j)=1;
        else
                    M(i,j)=0;
        end

    end
end
M
% Fin del fichero sigma.m
```

- **Forma normal de una matriz reducible en Matlab**

Un algoritmo para la búsqueda automática de la forma normal (1) para matrices reducibles, ha sido implementado en la siguiente función de Matlab **normform.m**. Esta función de A produce los valores siguientes.

[N,M,K,r,m]=**normform**(A)

- N es la forma normal de la matriz reducible $n \times n$ no negativa A. Los estados de las clases transitorias van al principio.
- M es la matriz de permutación tal que $N = M^{\mathrm{T}} A M$.

- K es el vector ("cell array") cuyas componentes K{1}, K{2},...,K{m} son vectores fila de las clases comunicantes o estados sin retorno, ordenadas según conviene a la forma normal.
- r es el número total de las clases transitorias + los estados sin retorno.
- m es el número total de bloques en la diagonal de N.

Esta función requiere la función **irredble.m**.

Ejemplo 6.1 Consideremos la cadena de Markov dada por la Figura 20 donde asignamos un peso positivo p_{ij} a cada arco (i, j), indicado por $p_{ij} := *$, pues no importa mucho su valor. Nosotros hemos hecho los cálculos en Matlab con $* := 2$, pero podríamos haber puesto cualquier otra cantidad en cada p_{ij} no nulo.

Entonces la matriz de transición 15×15 de la cadena está representada en

$$P = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\
0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & * & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & * & 0 & * & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \\
0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0
\end{pmatrix} \tag{11}$$

Las clases de P están dadas en K: K{1}={1, 9}, K{2}={3}, K{3}={ 5, 6, 13, 14}, K{4}={ 2, 4, 10, 11}, K{5}={7, 12, 15} y K{6}={8}.

El estado 3 es sin retorno (i.e no se comunica con sí mismo) y el estado 8 es absorbente. Propiamente, la cadena tiene 5 clases de comunicación más el estado 3. Las clases K{1} y K{3} son transitorias. Las clases K{4}, K{5} y K{6} son recurrentes (o finales). En puridad, K{2}={3} no es una clase comunicante.

La permutación σ de 1,...,15 que transforma (las filas y columnas de) P en su forma normal N, se obtiene poniendo en la segunda línea de σ los elementos de las clases:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 1 & 9 & 3 & 5 & 6 & 13 & 14 & 2 & 4 & 10 & 11 & 7 & 12 & 15 & 8 \end{pmatrix}.$$

Por tanto, la matriz $M = (\delta_{i,\sigma(j)})$ es igual a

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

y la forma normal de P es la matriz N que damos en la ecuación (12)

$$N = \left|\begin{array}{cc|c|cccc|cccc|ccc|c}
\hline
0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
 & & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\cline{3-15}
 & & & 0 & 0 & * & 0 & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 \\
 & & & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & * & 0 & * \\
 & & & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\cline{4-15}
 & & & & & & & 0 & 0 & 0 & * & & & & \\
 & & & & & & & * & 0 & 0 & * & & & & \\
 & & & & & & & 0 & * & 0 & 0 & & & & \\
 & & & & & & & 0 & 0 & * & 0 & & & & \\
\cline{8-15}
 & & & & & & & & & & & 0 & * & 0 & \\
 & & & & & & & & & & & 0 & 0 & * & \\
 & & & & & & & & & & & * & * & 0 & \\
\cline{12-15}
 & & & & & & & & & & & & & & * \\
\hline
\end{array}\right| \tag{12}$$

Es claro que $r = 3$ y que $m = 6$. Hay $m - r = 3$ bloques diagonales aislados.

- **Forma de Frobenius en Matlab**
  - [F,P,E] =**frobform**(A)

    da la forma de Frobenius F de una matriz no negativa, irreducible y no primitiva A, de índice de imprimitividad d; así como la matriz de permutación P tal que $F = P^{\mathrm{T}}AP$ y el vector E con los subconjuntos $E_1, E_2, \ldots, E_d$ por los que pasa periódicamente un paseante aleatorio sobre el digrafo de A. Requiere **invierte** y **mcd**.

  - u=**invierte**(v) devuelve los elementos del vector $v = (v_1, v_2, \ldots, v_n)$ ordenados en sentido contrario $u = (v_n, \ldots, v_2, v_1)$.

- d=**mcd**(a) da el mcd (máximo común divisor) de los numeros enteros contenidos en el vector a. Utiliza la función **gcd**.

Los ficheros **frobform.m**, **invierte.m** y **mcd.m** pueden bajarse pinchando en ellos.

• Relaciones binarias en Matlab

Una relación binaria R en el conjunto finito S es un subconjunto del producto cartesiano S×S. Las funciones siguientes comprueban si una relación binaria R, dada por su matriz booleana T, es reflexiva, simétrica, antisimétrica, transitiva, de orden, o de equivalencia. Los ficheros.m correspondientes son: esreflexi.m, essimetri.m, esantisim.m, estransi.m, esorden.m, esequiva.m y corte.m. Asimismo, damos ficheros reflexiva.m, simetrica.m, transitiva.m, equivalencia.m, que contienen funciones para hallar las clausuras reflexiva, simétrica y transitiva de una relación binaria R, dada por T; así como una función que construye la mínima relación de equivalencia $\sim$ que contiene a R.

- s=**esreflexi**(T)

  da 1 si T es la matriz booleana de una relación reflexiva. Da 0 en caso contrario.

- s=**essimetri**(T)

  da 1 si T es la matriz booleana de una relación simétrica. Da 0 en caso contrario.

- s=**esantisim**(T)

  da 1 si T es la matriz booleana de una relación antisimétrica. Da 0 en caso contrario. Requiere la función **corte**.

- s=**estransi**(T)

  da 1 si T es la matriz booleana de una relación transitiva. Da 0 en caso contrario. Requiere la función **potboole**.

- s=**esorden**(T)

  da 1 si T es la matriz booleana de una relación de orden. Da 0 en caso contrario. Requiere las funciones **esreflexi, esantisim** y **estransi**.

- s=**esequiva**(T)

  da 1 si T es la matriz booleana de una relación de equivalencia. Da 0 en caso contrario. Requiere las funciones **esreflexi, essimetri** y **estransi**.

- Y=**corte**(A,B)

  da la matriz $m \times n$ Y a partir de las matrices booleanas $m \times n$ A y B mediante Y(i,j)=A(i,j) & B(i,j). El operador & es AND. Esto es, **corte** produce el producto booleano de A y B componente a componente.

- Y=**reflexiva**(T) toma la matriz booleana T (una relación binaria) y devuelve su clausura reflexiva.
- Y=**simetrica**(T) toma la matriz booleana T (una relación binaria) y devuelve su clausura simétrica.
- Y=**transitiva**(T) toma la matriz booleana T (una relación binaria) y devuelve su clausura transitiva.
- Y=**equivalencia**(T) toma la matriz booleana T (de una relación binaria R) y devuelve la mínima relación de equivalencia que contiene a R. Requiere las funciones **reflexiva, simetrica** y **transitiva**.

Ejercicio 6.1 Dada una relación binaria R en el conjunto finito S, no igual a S×S, ¿cuándo es igual a S×S la mínima relación de equivalencia que contiene a R?

7. Soluciones

Solución del Ejercicio 4.1:

$$p_{00} \approx 6/23, \quad p_{01} \approx 17/23, \quad p_{10} \approx 17/27, \quad p_{11} \approx 10/27.$$

Solución del Ejercicio 4.2:

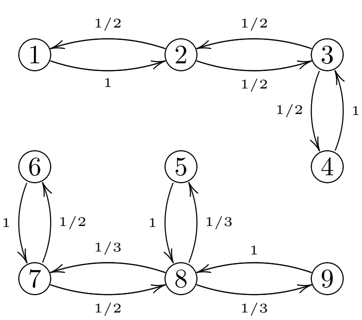


Figura 21: Digrafo ponderado del laberinto 3º.

Es obvio que hay dos posibilidades. Si la rata inicia su andadura en una de las cámaras 1,2,3 ó 4, permanece en algunas de ellas todo el tiempo y cada una de estas cámaras es visitada *indefinidamente* todo el tiempo. Lo mismo cabe decir si arranca de alguno de los compartimentos 5,6,7,8 ó 9. Así pues, estas dos situaciones alternativas describen la evolución de la cadena. Es claro también que la distribución de probabilidades del paso k-ésimo $\boldsymbol{p}^{\mathrm{T}}(k)$ depende de la posición inicial de la rata en una de estas dos "viviendas" estancas.

8. Acreditaciones

Las Figuras 13, 15, 16 y 17 han sido tomadas de los Apuntes de D.P. Bertsekas y J.N. Tsitsiklis: Introduction to Probability. La Figura 8 ha sido tomada del libro [7]. La clasificación de los estados de una cadena de Markov está inspirada en el libro [6, p. 244, Sec. 6.3, p. 258, Sec. 6.5]. Los resultados de las Subsecciones 4.2 y 4.3 pueden verse demostrados en [5, p. 694–698] y [3, p. 83–93].

8.1. Agradecimientos

El autor agradece a José María González de Durana la lectura de este documento y por avisarle de algunas erratas.

Referencias

[1] A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, New York, 1979. 13, 34

[2] M. Fiedler. *Special Matrices and their Applications in Numerical Mathematics*. Martinus Nijhoff, Dordrecht, 1986. 9

[3] F. R. Gantmacher. *Théorie des Matrices*, volume 2. Dunod, Paris, 1966. Traduit par Ch. Sarthou. 9, 41

[4] P. Lancaster and M. Tismenetsky. *The Theory of Matrices with Applications*. Academic Press, second edition, 1985. 9, 11, 12, 13, 34

[5] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, Philadelphia, 2000. 7, 13, 19, 41

[6] E. Parzen. *Procesos Estocásticos*. Paraninfo, Madrid, 1972. 41

[7] N. Pullman. *Matrix Theory and Its Applications*. Marcel Dekker, Inc., New York, 1976. 41

Índice alfabético