

IV. MATRICES QUE DEJAN CONOS INVARIANTES

EN MUCHAS de las aplicaciones de la teoría de matrices aparecen matrices reales con entradas no negativas. Por ejemplo, si consideramos problemas de economía como los estudiados por el modelo de Leontief (ver capítulo II); problemas de probabilidad como los examinados por medio de matrices estocásticas (ver capítulo III y otros ejemplos en este capítulo); problemas de gráficas estudiados por medio de matrices de adyacencia (siguiente capítulo). El hecho más relevante relativo a los valores propios de matrices no negativas está contenido en el importante *teorema de Perron-Frobenius*: el radio espectral de una matriz no negativa $A \in M_n(\mathbf{R})$ es un valor propio de A que acepta un vector propio con entradas no negativas.

Una matriz $A \in M_n(\mathbf{R})$ es no negativa si y solamente si el cono $K = \{x \in \mathbf{R}^n : x \geq 0\}$ de vectores no negativos en \mathbf{R}^n queda invariante bajo la acción de A . Esta sencilla observación da lugar a la importante generalización del teorema de Perron-Frobenius que explicamos en este capítulo: dada una matriz $A \in M_n$ que deja invariante un cono propio sólido K en \mathbf{R}^n , entonces el radio espectral es un valor propio de A que acepta un vector propio que pertenece al cono K . Este teorema tiene un rango de aplicaciones más amplio que el teorema de Perron-Frobenius, como veremos en breves consideraciones de matrices asociadas a gráficas, sistemas de ecuaciones diferenciales y otros.

1. CONOS

1.1. Sea V el espacio vectorial \mathbf{R}^n . En lo que sigue consideraremos a V como un espacio topológico con la topología inducida por la norma euclidiana.

Un cono K es un subconjunto cerrado de V que satisface:

- (K1) $K + K \subset K$ (esto es, para todos $v, w \in K$, se tiene $v + w \in K$).
- (K2) $\alpha K \subset K$ para todo $\alpha \geq 0$ (esto es, para todo $v \in K$, se tiene $\alpha v \in K$).

Decimos que el cono K es *propio* si, además, se cumple:

- (K3) $K \cap (-K) = \{0\}$, donde $-K = \{-x : x \in K\}$,
- y es *sólido* en V si se cumple:

(K4) El interior K° de K no es vacío.

Ejemplos de conos

(a) Sea $V = \mathbf{R}^n$, el cono positivo $V^+ = \{v \in V: v(i) \geq 0 \text{ para } 1 \leq i \leq n\}$.

(b) Si $A: V \rightarrow V$ es una transformación lineal y K es un cono, entonces $A^{-1}(K)$ es un cono.

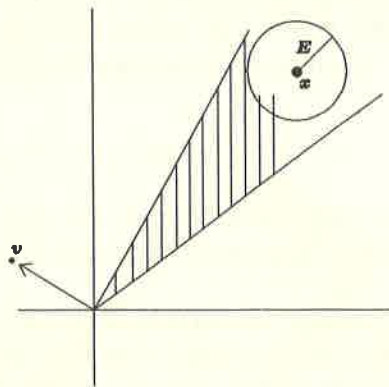
(c) Consideremos el espacio dual $V^* = \text{Hom}_{\mathbf{R}}(V, \mathbf{R})$. El cono ortogonal K^\perp está definido por

$$K^\perp = \{f \in V^*: f(v) \geq 0, \text{ para cada } v \in K\}.$$

Lema. Para un cono K son equivalentes:

- i) K es sólido.
- ii) $K - K = V$.
- iii) K contiene una base de V .

Demostración: (i) \Rightarrow (ii): Supongamos que K es sólido. Existe $x \in K^\circ$; o sea, podemos encontrar una bola con centro en x y radio ε suficientemente pequeño de manera que $B_\varepsilon(x) \subset K$.



Sea $v \in V$, entonces existe $\delta > 0$, $\omega = \delta v + x \in B_\varepsilon(x) \subset K$. Luego

$$v = \delta^{-1}\omega - \delta^{-1}x \in K - K.$$

(ii) \Rightarrow (iii): Sea $\{v_1, \dots, v_s\}$ un conjunto maximal de vectores linealmente independientes en K . Como $K - K = V$, $\{v_1, \dots, v_s\}$ genera V . Así $s = n$.

(iii) \Rightarrow (i): Sea $\{v_1, \dots, v_n\}$ una base de V con $v_i \in K$, $i = 1, \dots, n$.

El conjunto $C = \left\{ \sum_{i=1}^n \lambda_i v_i: 0 \leq \lambda_i \leq 1 \right\}$ está contenido en K (observe que C es un poliedro con n lados). Claramente, $C^\circ = \left\{ \sum_{i=1}^n \lambda_i v_i: 0 < \lambda_i < 1 \right\} \neq \emptyset$. Luego $K^\circ \neq \emptyset$ y K es sólido. \square

Ejercicios

1. Sea K un cono, $x \in K^\circ$, $y \in K$. Entonces $x + y \in K^\circ$.
2. Dé condiciones sobre un cono K en V y una transformación $A: V \rightarrow V$ para que el cono $A^{-1}(K)$ sea sólido.

1.2. Para un vector $x \in K$ también escribimos $x \geq 0$. Si K es sólido y $x \in K^\circ$, escribimos $x \succ 0$.

Observemos que si K es un cono propio \geq nos da en efecto un orden en V : $x \geq y$ si y sólo si $x - y \geq 0$. En efecto, si $x \geq y$, $y \geq x$, entonces $x - y \in K \cap (-K) = \{0\}$, $x = y$; si $x \geq y$, $y \geq z$, entonces $x - y, y - z \in K$, $x - z = (x - y) + (y - z) \in K$; si $x \geq y$, $\alpha \geq 0$, $\alpha x \geq \alpha y$.

Un vector $x \in K$ se llama *extremal* si $x = y + z$, con $y, z \in K$, implica que y, z son múltiplos escalares de x .

Un cono se dice *generado* por un conjunto $\{v_i\}_{i \in I}$ de vectores si $K = \left\{ \sum_{i \in I} \lambda_i v_i: \lambda_i \geq 0 \right\}$. Si I es finito, entonces K es *poliédrico*.

Ejercicios

1. En \mathbf{R}^3 considere el cono K dado por todos los puntos (x, y, z) que satisfacen las desigualdades $x^2 + y^2 \leq z^2$ y $z \geq 0$. Pruebe que este cono (que llamaremos *el cono de helado*) no es poliédrico.
2. En un cono poliédrico ¿cuáles son los vectores extremales?

Proposición. Un cono K en V está generado por sus vectores extremales.

Demostración: Por inducción sobre n , $V = \mathbf{R}^n$.

Si $n = 1, 2$, el resultado es obvio. Si K no es sólido, podemos suponer que $K \subset \mathbf{R}^{n-1}$ y el resultado se sigue por inducción.

Sea $x \in K^\circ$ y sea $u \in K$ linealmente independiente de x . Consideremos el plano P generado por x, u . Luego $K' = P \cap K$ es un cono con $x \in K'$. Así, $x = x_1 + x_2$ y x_1, x_2 son extremales en K' . Luego x_1, x_2 están en la frontera ∂K de K . Para probar el resultado basta con considerar vectores en ∂K . Supongamos $x \in \partial K$ y x no es extremal. Luego $x = u + v$, con $u, v \in K$ linealmente independientes. Si $u \in K^\circ$, también $x \in K^\circ$; entonces $u, v \in \partial K$. El conjunto $\{\alpha u + \beta v: \alpha, \beta \geq 0\}$ es un cono contenido en ∂K .

Sea S el cono más grande con $x \in S \subset \partial K$. Sea H el espacio vectorial de dimensión mínima con $S \subset H$. Por (1.1), $\dim H \leq n - 1$. Por hipótesis de inducción, x está generado por los vectores extremales de S . Basta probar que estos vectores son también extremales en K .

Supongamos que $y \in S$ no es extremal en K . Luego $y = u + v$ con $u, v \in K$ linealmente independientes. Si $u \notin S$ podemos construir el cono

$$S' = \{w + \alpha u; \quad w \in S, \quad \alpha \geq 0\}.$$

Como $w + \alpha u = w + \alpha(y - v) \stackrel{K}{\leq} w + \alpha y \in S \subset \partial K$, entonces $S' \subset \partial K$, lo que contradice la maximalidad de S . Así, $u, v \in S$, lo que implica que y no es extremal en S . \square

1.3. Un *subcono* de K es un cono contenido en K . Un *subcono extremal* de K es un subcono generado por vectores extremales. Un subcono extremal contenido en ∂K se llama una *cara* de K . Una cara F de K siempre está contenida en un subespacio de dimensión menor que n ; al menor de dichos subespacios lo denotamos por H_F .

Lema. Para cada $x \in \partial K$ existe una cara F_x tal que:

- i) $x \in F_x^\circ$, donde el interior se toma en el espacio H_{F_x} ;
- ii) $F_x = \partial K \cap H_{F_x}$;
- iii) $0 \stackrel{K}{\leq} y \stackrel{K}{\leq} x$, implica que $y \in F_x$.

Demostración: Por (1.2), podemos escribir

$$x = \sum_{i=1}^m \lambda_i x_i$$

con x_i extremal, $\lambda_i > 0$, $i = 1, \dots, m$. Sea F_x' el cono generado por x_1, \dots, x_m . Claramente F_x' satisface (i). Sea $F_x' \subset F_x$ la mayor cara que satisface (i). Evidentemente, $F_x \subset \partial K \cap H_{F_x}$. Pero $\partial K \cap H_{F_x}$ es una cara que satisface (i), y por tanto se da la igualdad, $F_x = \partial K \cap H_{F_x}$.

Finalmente, supongamos $0 \stackrel{K}{\leq} y \stackrel{K}{\leq} x$, entonces $y \in \partial K$. Supongamos que $y \notin H_{F_x}$. Consideremos el espacio H' generado por H_{F_x} y $\{y\}$. Entonces $F' = \partial K \cap H'$ es una cara, $x - y \in F'$, $y \in (F')^\circ$ en el espacio $H_{F'}$.

Por tanto, $x \in (F')^\circ$ en el espacio $H_{F'}$, lo que contradice la maximalidad de F_x . Luego $y \in H_{F_x}$, $y \in F_x$. \square

Ejercicios

1. Sea K un cono sólido en \mathbf{R}^3 . Sea F una cara de K de forma que el espacio H_F es de dimensión $\dim_k H_F = 1$. Demuestre que entonces K no es poliédrico. ¿Puede generalizarse esta observación a dimensión n ?

2. Sea K un cono sólido en V . Sea $x \in K$ y F_x la cara definida en el Lema anterior. Pruebe que

$$F_x = K \cap H_{F_x}.$$

[Ayuda: Si $y \in K^\circ$, entonces y no está en el espacio vectorial generado por F_x , ya que de otra manera se tendría $y = y_1 - y_2$ con $y_1, y_2 \in F_x$, lo que contradice el inciso (iii) del Lema.]

1.4. Proposición. Sea K un cono propio sólido en V . Entonces las siguientes afirmaciones son equivalentes:

- a) K es poliédrico;
- b) K tiene un número finito de caras;
- c) existe un número finito de vectores $v_1, \dots, v_s \in V$ de forma que

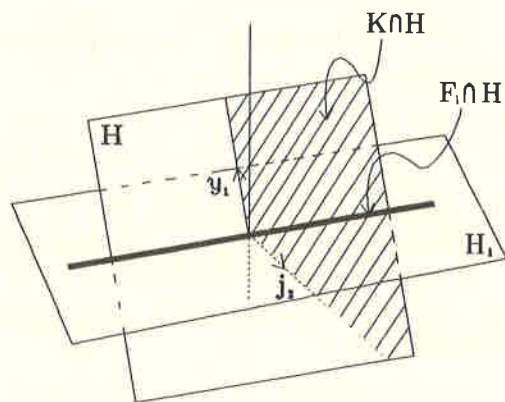
$$K = \bigcap_{i=1}^s H_{v_i}^+,$$

donde $H_{v_i}^+ = \{x \in \mathbf{R}^n: x^T v_i \geq 0\}$ es un *semiespacio* asociado al vector v_i .

Demostración: (a) \Rightarrow (b): es claro.

(b) \Rightarrow (c): Sean F_1, \dots, F_n caras de K que contienen elementos $x_i \in F_i$ tales que $F_i = F_{x_i}$ con la notación de (1.3). En particular, $H_i = H_{F_i}$ es un espacio vectorial tal que $F_i = \partial K \cap H_i$.

Sean $w_1^{(i)}, \dots, w_{s_i}^{(i)}$ vectores ortonormales tales que $H_i = \{x \in \mathbb{R}^n: x^T \cdot w_j^{(i)} = 0, j = 1, \dots, s_i\}$. Supongamos que existen $y_1, y_2 \in K$ tales que $y_1^T w_j^{(i)} > 0$ y $y_2^T w_j^{(i)} < 0$. Denotemos por un momento $w = w_j^{(i)}$. Consideremos el plano H generado por y_1, y_2 .



Dado que $H \cap K$ es un cono en H , por continuidad podemos encontrar $y \in H_i \cap H \cap K$. Multiplicando por escalares convenientes podemos suponer que $y = y_1 + y_2$. Por el ejercicio (1.3.2) tenemos que $y \in F_i$. Esto implica que y_1 está en F_i , lo que contradice que F_i está contenido en H_i . Así, podemos suponer que $K \subset H_w^+$. Haciendo $v_j^{(i)} = w_j^{(i)}$ o bien $v_j^{(i)} = -w_j^{(i)}$, de forma que $K \subset H_{v_j^{(i)}}^+$ para toda (i, j) , obtenemos que $K \subset \bigcap_{j=1}^{s_i} H_{v_j^{(i)}}^+$.

El resto de la demostración se deja como ejercicio. \square

Ejercicios

1. Demuestre que $K = \bigcap_{i=1}^n \bigcap_{j=1}^{s_i} H_{v_j^{(i)}}^+$ en la proposición anterior.

2. Demuestre la implicación restante (c) \Rightarrow (a) en la proposición anterior.

1.5. La proposición anterior tiene la siguiente consecuencia para los conos ortogonales definidos en el espacio dual como en (1.1). Usaremos este resultado más adelante.

Lema. Consideremos el isomorfismo natural

$$\varphi: V \rightarrow V^{**}, \quad \varphi(v)(g) = g(v), \quad \text{para } v \in V, g \in V^*.$$

Si K es un cono poliédrico, entonces $K^{\perp\perp} = \varphi(K)$.

Demostración: En general, $\varphi(K) \subset K^{\perp\perp}$: si $v \in K$, $f \in K^{\perp}$, entonces $\varphi(v)(f) = f(v) \geq 0$, de donde $\varphi(v) \in K^{\perp\perp}$.

Si K es poliédrico, entonces $K = \bigcap_{i=1}^s H_{v_i}^+$, donde $H_{v_i}^+$ es el semiespacio asociado a v_i (1.4). Mostremos que para dos semiespacios H_v^+ , H_w^+ se tiene:

$$(i) (H_v^+)^{\perp\perp} = \varphi(H_v^+);$$

$$(ii) (H_v^+ \cap H_w^+)^{\perp\perp} = \varphi(H_v^+) \cap \varphi(H_w^+);$$

de aquí se sigue que $K^{\perp\perp} = \varphi(K)$.

Como antes, $\varphi(H_v^+) \subset (H_v^+)^{\perp\perp}$. Si $g \in (H_v^+)^{\perp\perp}$, $g = \varphi(w)$, tenemos que probar que $w \in H_v^+$. En efecto, consideramos $f: V \rightarrow \mathbb{R}$, $x \rightarrow xv^T$, tal que $f \in (H_v^+)^{\perp}$. Luego, $0 \leq g(f) = \varphi(w)(f) = f(w) = wv^T$, o sea $w \in H_v^+$.

Exactamente de la misma manera se prueba (ii). \square

2. MATRICES QUE DEJAN CONOS INVARIANTES

2.1. Sea A una matriz real de tamaño $n \times n$ y K un cono en $V = \mathbb{R}^n$. Se dice que A deja invariante el cono K si $A(K) \subset K$.

En lo que sigue nos ocuparemos del estudio de propiedades espectrales de matrices que dejan conos invariantes. Un par de ejemplos:

(a) Si $A = (a_{ij})$ con $a_{ij} \geq 0$ para $1 \leq i, j \leq n$, entonces $A(V^+) \subset V^+$, donde V^+ es el cono positivo en V . De hecho esto caracteriza a las matrices con coeficientes no negativos.

(b) Consideremos la matriz

$$A = \begin{bmatrix} 1 & a \\ 1 & -1 \end{bmatrix}, \quad \text{con } a > 0,$$

que envía un vector (x) en $(x') = \begin{bmatrix} x + ay \\ x - y \end{bmatrix}$. Luego $0 \leq y \leq x$ implica $0 \leq y' \leq x'$. Entonces, si definimos el cono $K = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbf{R}^2 : 0 \leq y \leq x \right\}$, tenemos $A(K) \subset K$. Observemos que $\rho(A) = \sqrt{a+1}$ con vector propio $\left(1, \frac{1}{a}(\rho(A) - 1)\right) \in K$.

Recordamos del capítulo I que para una matriz A y un valor propio λ de A , el grado $d(\lambda)$ de λ es el tamaño del bloque de Jordan más grande con valor propio λ en la forma de Jordan de A .

El siguiente resultado es el teorema principal de este capítulo. Históricamente tiene su origen en los resultados de Perron que estudiaremos más tarde. En la forma del teorema que aquí presentamos, el teorema fue demostrado en trabajos de G. Birkhoff (1967) y J. S. Vandergraft (1968). En nuestro tratamiento invertiremos el orden histórico y el teorema de Perron-Frobenius se obtendrá como consecuencia del teorema más general de Birkhoff-Vandergraft.

Teorema. Sea K un cono propio sólido y A una matriz tal que $A(K) \subset K$. Entonces:

- i) $\rho(A)$ es valor propio de A ;
- ii) $d(\rho(A)) \geq d(\lambda)$, para cualquier otro valor propio $|\lambda| = \rho(A)$;
- iii) K contiene un vector propio $0 \neq v \in K$, con $Av = \rho(A)v$.

Además, si A satisface las condiciones (i) y (ii), entonces K deja un cono propio sólido invariante.

Demostración: Sea $\varepsilon_1, \dots, \varepsilon_n$ una base de Jordan de A ordenada de la siguiente manera:

$\varepsilon_{m_1}, \dots, \varepsilon_{m_\ell}$ son los vectores propios de A , con $m_1 < m_2 < \dots < m_\ell$ y $\lambda_1, \dots, \lambda_\ell$ los respectivos valores propios con $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_\ell|$.

Para cada $1 \leq j \leq \ell$, $\lambda_j = |\lambda_j|e^{i\alpha_j}$. Para cada $1 \leq h \leq n$ existe una única $1 \leq j(h) \leq \ell$ tal que $m_{j(h)-1} < h \leq m_{j(h)}$, con $m_0 := 0$ y $m_\ell = n$,

$$A(\varepsilon_h) = \lambda_{j(h)}\varepsilon_h + \varepsilon_{h+1}, \quad \text{si } h < m_{j(h)},$$

$$A(\varepsilon_{m_{j(h)}}) = \lambda_{j(h)}\varepsilon_{m_{j(h)}}.$$

Al aplicar repetidamente A tenemos:

$$A^r(\varepsilon_h) = \sum_{k=h}^{m_{j(h)}} \lambda_{j(h)}^{r-(k-h)} \binom{r}{k-h} \varepsilon_k, \quad \text{con } \binom{r}{k-h} = 0 \text{ si } r < k-h. \quad (*)$$

Denotamos por $\rho = \rho(A) = \max \{|\lambda_i| : i = 1, \dots, \ell\}$.

Sea K un cono propio y sólido tal que $A(K) \subset K$. Deseamos demostrar que (i), (ii) y (iii) se satisfacen.

Primeramente, si $\rho = 0$ entonces A es nilpotente. Elegimos $1 < r$ mínima con $A^r = 0$ en K (si no, $A(K) = 0$ y $A = 0$). Sea $x \in K$ tal que $0 \neq w = A^{r-1}(x) \in K$. Entonces $A(w) = 0 = \rho w$. Además, $d(\rho) = n$.

Podemos suponer entonces que $\rho > 0$.

Denotemos por b_j el tamaño del bloque de Jordan de A correspondiente al vector ε_{m_j} . Observemos que $b_j = m_j - m_{j-1}$.

Sea $M = \max \{b_j : |\lambda_j| = \rho, 1 \leq j \leq \ell\}$. Podemos suponer que $1, 2, \dots, t$ son los índices j tales que $|\lambda_j| = \rho$ y $b_j = M$.

En (2.2) demostraremos que existen números $d_j \in \mathbf{C}$, $1 \leq j \leq t$ con

$$0 \neq y = \sum_{j=1}^t d_j \varepsilon_{m_j} \in K.$$

Ésta es una parte esencial (pero un tanto técnica) de la prueba. Usemos esta afirmación para concluir la demostración que nos ocupa.

Supongamos que $\lambda_1 \notin \mathbf{R}$ (es decir, $\lambda_1 \neq \rho$). Por un lema elemental que mostraremos en (2.3), existen números reales $0 \leq c_0, \dots, c_q$, no todos cero, de manera que

$$\sum_{p=0}^q c_p \lambda_1^p = 0.$$

Definimos

$$\begin{aligned} y^* &= \sum_{p=0}^q c_p A^p(y) = \sum_{p=0}^q c_p A^p \left(\sum_{j=1}^t d_j \varepsilon_{m_j} \right) \\ &= \sum_{j=2}^t d_j \left(\sum_{p=0}^q c_p \lambda_j^p \right) \varepsilon_{m_j} \in K, \end{aligned}$$

que es combinación lineal de $\{\varepsilon_{m_j} : 1 < j \leq t\}$ (aquí hemos usado que K es un cono invariante bajo A). Además, $0 \neq y^*$, si no

$$\sum_{p=1}^q c_p A^p(y) = -c_0 y \in K \cap (-K) = \{0\},$$

de donde $c_0 = 0$; $\sum_{p=2}^q c_p A^p(y) = -c_1 A(y)$, lo que implica $c_1 = 0$, y continuando llegaríamos a $c_p = 0$, $0 \leq p \leq q$ (observe que $A^p(y) \neq 0$ por (*)), lo que es absurdo.

Si $\lambda_2 \notin \mathbf{R}$, como antes, podemos construir $0 \neq y^{**} \in K$, combinación lineal de $\{\varepsilon_{m_j} : 2 < j \leq t\}$. Este proceso es claramente finito, por lo que existe un vector $0 \neq \tilde{y} \in K$, combinación lineal de $\{\varepsilon_{m_j} : \lambda_j = \rho, b_j = M\}$. Luego,

$$A(\tilde{y}) = \rho \tilde{y},$$

$$d(\rho) = M \geq d(\lambda) \text{ si } |\lambda| = \rho.$$

Con lo que probamos que (i), (ii) y (iii) se satisfacen.

La última afirmación del teorema la mostraremos en (2.4), después de completar los argumentos que quedaron pendientes de la primera parte. \square

2.2. Con base en las hipótesis y la notación que teníamos en (2.1), probaremos que existe

$$0 \neq y = \sum_{j=1}^t d_j \varepsilon_{m_j} \in K.$$

Como K es un cono sólido, existe $0 \neq z = \sum_{h=1}^n c_h \varepsilon_h \in K^0$ con $0 \neq c_h \in \mathbf{R}$ (en efecto, dado $0 \neq z' \in \sum_{h=1}^n c'_h \varepsilon_h \in B_\delta(z') \subset K^0$ con $\delta > 0$, podemos modificar los coeficientes para obtener $c_h \neq 0$ con $0 \neq y = \sum_{h=1}^n c_h \varepsilon_h \in B_\delta(z')$). La expresión (*) para z toma la forma

$$A^r(z) = \sum_{h=1}^n c_h A^r(\varepsilon_h) = \sum_{h=1}^n \left(\sum_{m_{j(h)-1} < k \leq h} c_k \lambda_{j(h)}^{r-(h-k)} \binom{r}{h-k} \right) \varepsilon_h,$$

tomando $r \geq n$.

Queremos calcular $\lim_{r \rightarrow \infty} \frac{A^r(z)}{\rho^r r^{(M-1)}}$; recordamos que $M = \max\{b_j : |\lambda_j| = \rho, 1 \leq j \leq \ell\}$. Podemos establecer este límite para un bloque de Jordan $J_s(\mu)$:

$$J_s(\mu)^r = \begin{bmatrix} \mu^r & r\mu^{r-1} & \binom{r}{2}\mu^{r-2} & \dots & \binom{r}{s-1}\mu^{r-(s-1)} \\ & \mu^r & r\mu^{r-1} & \dots & \binom{r}{s-2}\mu^{r-(s-2)} \\ & & \mu^r & \ddots & \vdots \\ & & & \ddots & r\mu^{r-1} \\ 0 & & & & \mu^r \end{bmatrix}$$

Luego, $\lim_{r \rightarrow \infty} \frac{J_s(\mu)^r}{\rho^r r^{(M-1)}} = 0$ si $|\mu| < \rho$, o bien si $s < M$. Si $|\mu| = \rho$ y $s = M$, entonces

$$\frac{J_s(\mu)^r}{\rho^r r^{(M-1)}} = \begin{bmatrix} 0 & \dots & 0 & \frac{e^{ir\alpha_j}}{(M-1)! \rho^{(M-1)}} \\ 0 & \dots & 0 & 0 \\ & \vdots & & \\ 0 & \dots & \dots & 0 \end{bmatrix} + c(r)$$

tal que $\lim_{r \rightarrow \infty} \frac{c(r)}{\rho^r r^{(M-1)}} = 0$. En este caso, podemos elegir una sucesión r_1, r_2, \dots , tal que

$$\lim_{i \rightarrow \infty} \frac{J_s(\mu)^{r_i}}{\rho^{r_i} r_i^{(M-1)}} = \begin{bmatrix} 0 & \dots & 0 & \beta \\ 0 & \dots & 0 & 0 \\ & \vdots & & \\ 0 & \dots & 0 & 0 \end{bmatrix} \text{ existe,}$$

con $\beta \neq 0$. Regresando a nuestro problema original con la sucesión $(A^r(z))_r$, concluimos que podemos hallar una sucesión r_1, r_2, \dots , tal que

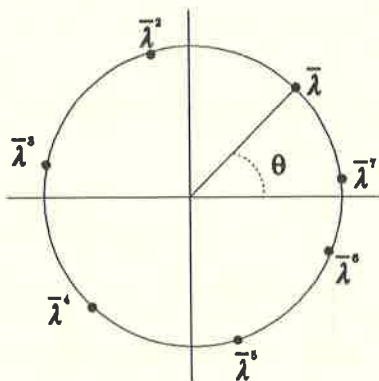
$$w = \lim_{i \rightarrow \infty} w_i = \sum_{j=1}^h d_j \varepsilon_{m_j} \in K,$$

con $w_i = \frac{A^{r_i}(z)}{\|A^{r_i}(z)\|} = \frac{\rho^{r_i} r_i^{(M-1)}}{\|A^{r_i}(z)\|} \left(\frac{A^{r_i}(z)}{\rho^{r_i} r_i^{(M-1)}} \right)$, $\|w_i\| = 1$. Esto demuestra uno de los puntos pendientes del teorema (2.1).

2.3. Lema. Sea $\lambda \in \mathbb{C}$ pero no en \mathbb{R}^+ , entonces existen números $c_0, \dots, c_q \in \mathbb{R}^+$, no todos ceros, de forma que

$$\sum_{p=0}^q c_p \lambda^p = 0.$$

Demostración: En efecto, escribimos $\lambda = re^{i\theta}$ con $0 < \theta < 2\pi$. Consideremos $\bar{\lambda} = e^{i\theta}$ en el círculo unitario.



El polígono convexo formado por los vértices $\{\bar{\lambda}^p : p \in \mathbb{N}\}$ contiene al 0. Luego existe $q \geq 1$ y constantes $\bar{c}_1, \dots, \bar{c}_q$ tales que

$$0 = \sum_{p=0}^q \bar{c}_p \bar{\lambda}^p = \sum_{p=0}^q (r^{-p} \bar{c}_p) \lambda^p. \quad \square$$

2.4. Procederemos a demostrar la última parte del teorema (2.1).

Ordenamos los valores propios $\lambda_1, \dots, \lambda_\ell$ de A como se ve a continuación:

$$\lambda_1 = \rho(A) \text{ y } |\lambda_1| = \dots = |\lambda_\nu| > |\lambda_{\nu+1}| \geq \dots \geq |\lambda_\ell|.$$

Definamos

$$c = \begin{cases} 1, & \text{si } \nu = \ell, \\ \lambda_1 - |\lambda_{\nu+1}|, & \text{en caso contrario.} \end{cases}$$

Podemos modificar la base de Jordan de manera que obtengamos una base $\varepsilon_1, \dots, \varepsilon_n$ que satisfaga

$$A(\varepsilon_h) = \lambda_j \varepsilon_h + c \varepsilon_{h+1}, \text{ si } m_{j-1} < h < m_j,$$

$$A(\varepsilon_{m_j}) = \lambda_j \varepsilon_{m_j},$$

con $0 = m_0 < m_1 < m_2 < \dots < m_\ell = n$ (ver 3.1). Además, suponemos que los vectores complejos ε_i vienen por pares conjugados.

Por hipótesis (ii), tenemos $m_1 \geq m_i - m_{i-1}$ para $i = 1, \dots, \nu$.

Definamos el conjunto

$$K = \left\{ x \in \mathbb{R}^n : x = \sum_{h=1}^n \alpha_h \varepsilon_h \text{ tal que para } m_{j-1} < h \leq m_j; \right. \\ \left. |\alpha_h| \leq \alpha_{h-m_{j-1}}, \text{ si } h - m_{j-1} \leq m_1 \text{ y } |\alpha_j| \leq \alpha_{m_1}, \text{ si no;} \right. \\ \left. \text{además, } \alpha_i = \bar{\alpha}_p \text{ si } \varepsilon_i = \bar{\varepsilon}_p \right\}.$$

Mostraremos que K es un cono sólido invariante.

En primer lugar observemos que K es un cono en \mathbb{R}^n bien definido.

Tomemos un elemento arbitrario $y \in \mathbb{R}^n$, $y = \sum_{h=1}^n \gamma_h \varepsilon_h$. Entonces

$$y = \sum_{h=1}^n \beta_h \varepsilon_h - \sum_{h=1}^{m_1} \delta_h \varepsilon_h,$$

donde

$$\beta_h = \begin{cases} \gamma_h & \text{si } m_1 < h, \\ \max \{ |\gamma_h|, |\gamma_{h+m_{j-1}}| : h + m_{j-1} \leq m_j \} & \text{si } h < m_1, \\ \max \{ |\gamma_{m_1}|, |\gamma_{i+m_{j-1}}| : m_1 \leq i \leq m_j - m_{j-1} \} & \text{si } h = m_1, \end{cases}$$

con $\delta_h = \beta_h - \gamma_h, \quad h = 1, \dots, m_1.$

Claramente, $y \in K - K$. Esto muestra que $\mathbb{R}^n = K - K$ y K es sólido.

Para mostrar la invarianza de K , tomemos

$$x = \sum_{h=1}^n \alpha_h \varepsilon_h \in K$$

y

$$A(x) = \sum_{h=1}^n \beta_h \varepsilon_h.$$

Por la definición de la base obtenemos

$$\beta_h = \begin{cases} \alpha_h \lambda_j + c \alpha_{h+1} & \text{si } m_{j-1} < h < m_j; \\ \alpha_h \lambda_j & \text{si } h = m_j. \end{cases}$$

Claramente tenemos que $\beta_h = \bar{\beta}_p$ si $\varepsilon_h = \bar{\varepsilon}_p$. Por tanto, basta con probar que

$$|\beta_h| \leq \beta_{h-m_{j-1}}, \text{ si } h - m_{j-1} \leq m_1,$$

$$|\beta_h| \leq \beta_{m_1}, \text{ en caso contrario.}$$

Consideramos varios casos:

Caso 1. $m_{j-1} < h < m_j$, $h - m_{j-1} < m_1$. Entonces,

$$|\beta_h| = |\alpha_h \lambda_j + c \alpha_{h+1}| \leq \alpha_{h-m_{j-1}} \lambda_1 + c \alpha_{h+1-m_{j-1}} = \beta_{h-m_{j-1}}.$$

Caso 2. $h = m_j$, $h - m_{j-1} < m_1$. Entonces,

$$|\beta_h| = |\alpha_h \lambda_j| \leq \alpha_{h-m_{j-1}} \lambda_1 + c \alpha_{h+1-m_{j-1}} = \beta_{h-m_{j-1}}.$$

Caso 3. $m_{j-1} < h < m_j$, $h - m_{j-1} \geq m_1$. Entonces,

$$|\beta_h| = |\alpha_h \lambda_j + \alpha_h \lambda_j + c \alpha_{h+1}| \leq \alpha_{m_1} (|\lambda_j| + c) \leq \alpha_{m_1} \lambda_1 = \beta_{m_1}.$$

Caso 4. $h = m_j$, $h - m_{j-1} \geq m_1$. Entonces,

$$|\beta_h| = |\alpha_h \lambda_j| \leq \alpha_{m_1} \lambda_1 = \beta_{m_1}. \quad \square$$

2.5. Supongamos que A deja invariante el cono $K \subset V = \mathbf{R}^n$. Decimos que K es *propriadamente invariante* bajo la acción de A si $A^m(K) \subset K^0$

para alguna potencia $m \geq 1$. En este caso podemos obtener una versión más fuerte del teorema (2.1).

Teorema. Supongamos que A deja propriadamente invariante el cono propio sólido $K \subset V = \mathbf{R}^n$. Entonces:

- i) $\rho(A)$ es valor propio simple de A y, para cualquier valor propio $\lambda \neq \rho(A)$ de A , se tiene $|\lambda| < \rho(A)$;
- ii) un vector propio de A correspondiente a $\rho(A)$ pertenece a K^0 .

Demostración: Por (2.1) sabemos que $\rho(A)$ es valor propio de A , y que un vector propio de A correspondiente a $\rho(A)$ pertenece a K .

(1) Probaremos que A tiene exactamente un vector propio en K y este vector propio está en K^0 .

Escribamos $\rho = \rho(A)$ y sea $u \in K$ un vector propio de A correspondiente a ρ . Entonces $A^m(u) = \rho^m u \in K^0$ y $u \in K^0$.

Supongamos que $u' \in K^0$ es otro vector propio de A correspondiente al valor propio λ . En primer lugar observemos que $Au' = \lambda u'$ implica que $\lambda \in \mathbf{R}$. Si $\lambda < 0$, entonces $Au' = \lambda u' \in K \cap (-K) = \{0\}$. Luego $0 \leq \lambda \leq \rho$. Definimos $t_0 = \inf \{t > 0 : tu' - u \in K\} > 0$. (En efecto, como $u' \in K^0$, existe $\varepsilon > 0$, $B_\varepsilon(u') \subset K$; para $\delta > 0$ suficientemente pequeña, $u' - \delta u \in B_\varepsilon(u')$.) También

$$y = t_0 u' - u,$$

que debe estar en ∂K . Si $\rho \neq 0$, entonces

$$A(y) = t_0 \lambda u' - \rho u = \rho \left[t_0 \frac{\lambda}{\rho} u' - u \right] \in K.$$

Por lo tanto, $\frac{\lambda}{\rho} \geq 1$. Pero como también sabíamos que $\frac{\lambda}{\rho} \leq 1$, obtenemos que $\lambda = \rho$. Entonces $Ay = \rho y$, lo que implica que A tiene un vector propio en la frontera ∂K . Esto es absurdo ya que

$$y = \frac{1}{\rho^m} A^m(y) \in K^0.$$

En caso de que $\rho = 0$, también $\lambda = 0$ y $Ay = 0$. Esto también es absurdo porque $0 = A^m(y) \in K^0$. Esto concluye la demostración de (1); en particular hemos probado (ii).

(2) El valor propio ρ de A es raíz simple del polinomio característico $\chi_A(t)$. Supongamos que ρ no es raíz simple de $\chi_A(t)$. Existen entonces

vectores u, u' linealmente independientes con $u \in K^0$, $Au = \rho u$ y

$$Au' = \rho u' \quad \text{o bien} \quad Au' = \rho u' + u.$$

Si u' fuera vector propio de A , para alguna $t > 0$, $y = tu + u' \in K$ con y vector propio de A diferente de u . Esto contradice el punto (1). Luego se debe cumplir que $Au' = \rho u' + u$.

En particular, $-u' \notin K$ (si no, para m suficientemente grande, $u' + \frac{m}{\rho}u \in K$, ya que $u \in K^0$). Luego, $A^m(u') = \rho^m u' + m\rho^{m-1}u \in K \cap (-K) = \{0\}$, contradicción). Podemos entonces definir $t_0 = \inf \{t > 0: tu - u' \in K\} > 0$. Como en (1), $\rho > 0$ y

$$A(t_0 u - u') = \rho \left[\left(t_0 - \frac{1}{\rho} \right) u - u' \right] \in K,$$

lo que contradice la elección de t_0 . Por lo tanto, ρ es valor propio simple de A .

(3) Sea λ valor propio de A , $\lambda \neq \rho$. Probaremos que $|\lambda| < \rho$. Supongamos que $|\lambda| = \rho$, entonces $\lambda = \rho e^{i\theta}$ con $0 < \theta < 2\pi$. Sea $v \in \mathbb{C}^n$ con $A(v) = \lambda v$.

Probaremos primero que para alguna φ , $\operatorname{Re}(e^{i\varphi}v) \in K$. En efecto, para cada φ , definimos $t_\varphi \in \mathbb{R}$ como

$$t_\varphi = \begin{cases} 0, & \text{si } \operatorname{Re}(e^{i\varphi}v) \in K, \\ \inf \{t > 0: tu + \operatorname{Re}(e^{i\varphi}v) \in K\} > 0, & \text{en caso contrario.} \end{cases}$$

Llamamos

$$y_\varphi = t_\varphi u + \operatorname{Re}(e^{i\varphi}v) \in K.$$

Si $t_\varphi = 0$ ya probamos lo que queríamos; si no $y_\varphi \in \partial K$, y

$$A^m(y_\varphi) = \rho^m t_\varphi u + \rho^m \operatorname{Re}(e^{i(\varphi+m\theta)}v) \in K^0.$$

Luego $t_\varphi > t_{\varphi+m\theta}$ y entonces $\inf \{t_\varphi\} = 0$. De este modo, como K es cerrado, para alguna φ_0 tenemos que

$$y_0 = \operatorname{Re}(e^{i\varphi_0}v) \in K.$$

Como $0 < \theta < 2\pi$, por (2.3) existen números $\xi_0, \dots, \xi_s \geq 0$ y $\xi_0 > 0$, tales que

$$\sum_{k=0}^s \xi_k e^{ik\theta} = 0.$$

Luego, $\sum_{k=0}^s \xi_k A^k y_0 = \sum_{k=0}^s \xi_k \operatorname{Re}(e^{i(k\theta+\varphi_0)}v) = \operatorname{Re} \left(\sum_{k=0}^s \xi_k e^{ik\theta} e^{i\varphi_0} v \right) = 0$. Por lo tanto,

$$\sum_{k=1}^s \xi_k A^k y_0 = -\xi_0 y_0 \in K \cap (-K) = \{0\},$$

entonces $y_0 = 0$. Esto implica $v' = e^{i(\varphi_0+\pi/2)}v \in \mathbb{R}^n$. Como $A(v') = \lambda v' \in \mathbb{R}^n$, se debe tener $\lambda = \rho$ o $\lambda = -\rho$. Luego, $\lambda = -\rho$. Elegimos

$$t_0 = \inf \{t > 0: tu + v' \in K\} > 0;$$

entonces

$$A^{2m}(t_0 u + v') = \rho^{2m}(t_0 u + v') \in K^0$$

y debe existir $0 < t < t_0$, con $tu + v' \in K$, que es una contradicción. Esto muestra que $|\lambda| < \rho$. Con ello concluimos la demostración del teorema. \square

Ejercicios

1. Demuestre que una matriz cuadrada con coeficientes reales estrictamente positivos deja propiamente invariante un cono propio sólido.
2. Demuestre que la matriz

$$A = \begin{pmatrix} 3 & 1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

deja propiamente invariante el cono de helado.

2.6. Ejemplo: Sea n un número natural mayor o igual a 2. Consideremos la matriz:

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

Formemos ahora la matriz

$$C = -PP^{-T} = \begin{pmatrix} -1 & n \\ -n & n^2 - 1 \end{pmatrix}.$$

La matriz C deja un cono propio sólido invariante. En efecto, consideremos en \mathbf{R}^2 los vectores $p_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ y $p_2 = \begin{pmatrix} 1 \\ n \end{pmatrix}$. Sea K el cono generado por los vectores de la forma $C^m p_1$ y $C^m p_2$ con $m \geq 1$. Se verifica fácilmente que:

- (i) K está contenido en $(\mathbf{R}^2)^+$ (por inducción, con $n \geq 2$).
- (ii) K es un cono sólido invariante de C .

El radio espectral de C es $\rho = \frac{(n^2 - 2 + n\sqrt{n^2 - 4})}{2}$ y un vector propio en el cono K está dado por $v = \begin{pmatrix} 1 \\ \frac{1+\rho}{n} \end{pmatrix}$.

Ejercicios

1. Demuestre que para la matriz C , como antes, existe siempre un cono propio K_r invariante bajo C y C^{-1} simultáneamente. Este cono es sólido si y solamente si $n \geq 3$.

2. Considere la gráfica orientada

y $P = (p_{ij})$ la matriz $n + 1 \times n + 1$ definida por $p_{0i} = 1, p_{ii} = 1$ y cero en todos los demás casos. Como antes, llamamos $C = -PP^{-T}$. Demuestre:

i) Para $n \geq 4$ hay un cono K propio sólido en \mathbf{R}^{n+1} invariante bajo C . ¿Cuál es el radio espectral de la matriz C ?
[Ayuda: Considere el cono generado por $C^m p_i$ con $m \geq 1$, donde p_i son las columnas de P .]

ii) Existe un cono propio K_r invariante bajo C y C^{-1} simultáneamente. Este cono es sólido si y solamente si $n \geq 5$.

3. Sea Δ una gráfica orientada bipartita (esto es, dada una flecha $i \rightarrow j$ en Δ , no hay flechas que comiencen en j) con n vértices. Defina

$P = (p_{ij})$ tal que $p_{ii} = 1$ y, para $i \neq j$, $p_{ij} = 1$ si hay una flecha de i a j y 0 en caso contrario. Como antes, $C = -PP^{-1}$.

i) Demuestre que si la gráfica no orientada subyacente a Δ contiene un *diagrama de Dynkin extendido* (ver la definición en el capítulo V), entonces existe un cono propio sólido K en \mathbf{R}^n invariante bajo C .

ii) En ese caso el radio espectral de C es mayor o igual a 1. ¿Cuándo es estrictamente mayor que 1?

La matriz C así definida se llama la *matriz de Coxeter* de la gráfica Δ .

2.7. Como otro ejemplo más consideraremos las *matrices exponencialmente no negativas*.

Consideremos el sistema de ecuaciones diferenciales

$$y'(t) = Ay(t),$$

donde $t \in \mathbf{R}^+$ y $A \in M_n(\mathbf{R})$. Para un conjunto K en \mathbf{R}^n dado, una pregunta importante es la siguiente: ¿en qué condiciones toda solución de la ecuación que se origina en K se conserva en K ? Formalmente, quisiéramos caracterizar la propiedad:

$$x(0) \in K, \text{ entonces } x(t) = e^{tA}x(0) \in K, \text{ para toda } t \geq 0.$$

En caso de que K sea un cono propio, esto es equivalente a pedir que K sea invariante bajo la matriz e^{tA} para toda $t \geq 0$. En este caso se dice que A es *exponencialmente K -no negativa*.

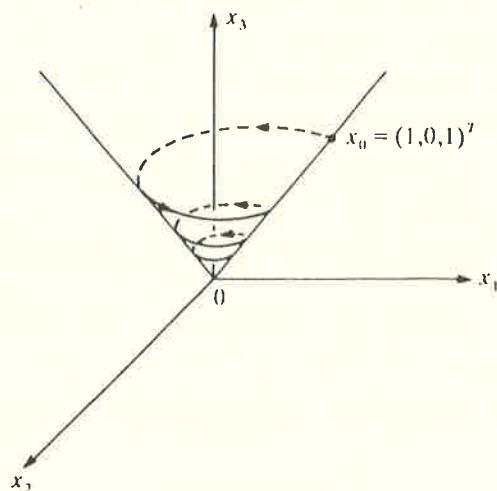
Ejercicios

1. Sea K el cono de helado en \mathbf{R}^3 . Considere la matriz

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

Demuestre que A es exponencialmente K -no negativa.

La siguiente figura muestra la trayectoria del punto $x_0 = (1, 0, 1)^T$.



2. Muestre que la siguiente matriz B es también exponencialmente K -no negativa:

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

Calcule las trayectorias de los puntos x en K con $x_3 = 1$.

3. Supongamos que A es exponencialmente K -no negativa para un cono propio K en \mathbb{R}^n . Por el Teorema (2.1), el radio espectral $\rho(e^{tA})$ es valor propio de e^{tA} . Por (1.5.1) sabemos que $\text{Spec}(e^{tA}) = e^{t \text{Spec}(A)}$. Luego, $\rho(e^{tA}) = e^{t\lambda(t)}$, donde $\lambda(t) \in \text{Spec}(A)$. Demuestre que entonces se tiene:

- i) $\lambda(t) = \lambda_0$ es constante.
- ii) Para todo valor propio $\lambda \in \text{Spec}(A)$ con $\text{Re } \lambda = \text{Re } \lambda_0$, se tiene que $d(\lambda) \leq d(\lambda_0)$.

iii) K contiene un vector propio de A correspondiente al valor propio λ_0 .

4. Enuncie y pruebe la versión exponencial del Teorema de Birkhoff-Vandergraft.

3. EL TEOREMA DE PERRON-FROBENIUS

3.1. Una matriz $A = (a_{ij})$ se llama *no negativa* (y escribimos $A \geq 0$) si $a_{ij} \geq 0$ para todos $1 \leq i, j \leq n$.

Se dice que es *positiva* (y escribimos $A > 0$) si $a_{ij} > 0$, para todos $1 \leq i, j \leq n$.

Se dice que A es *reducible* si existe una matriz permutación P (esto es, P se obtiene de I_n por medio de una permutación de columnas y la misma permutación en los renglones), tal que

$$PAP^t = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix},$$

donde B es una matriz cuadrada de tamaño $m \times m$ con $m < n$.

Ejercicios

1. Encuentre todas las matrices reducibles de tamaño 2×2 .
2. Demuestre el siguiente Lema.

Lema. Si A es una matriz reducible, existe un subespacio propio ($0 \neq W$) de $V = \mathbb{R}^n$ tal que $AW \subset W$ y W está generado por elementos de la base canónica.

Definición. Decimos que A es *irreducible* si A no es reducible.

Lema. Sea $0 \leq A$ matriz irreducible, entonces $(I_n + A)^{n-1} > 0$.

Demostración: Es suficiente probar que si $0 \leq y \in \mathbb{R}^n \setminus \{0\}$, entonces $(I_n + A)^{n-1}y > 0$.

Mostraremos que $(I_n + A)y$ tiene menos coordenadas 0 que y , de donde se sigue el resultado. Por medio de una permutación podemos suponer que

$$y = \begin{bmatrix} u \\ 0 \end{bmatrix}, \quad z = (I_n + A)y = \begin{bmatrix} v \\ 0 \end{bmatrix} \geq y$$

con $u > 0$, $v > 0$ del mismo tamaño.

Partimos adecuadamente A en bloques

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

de modo que tenemos

$$\begin{bmatrix} u \\ 0 \end{bmatrix} + \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} u \\ 0 \end{bmatrix} = \begin{bmatrix} v \\ 0 \end{bmatrix},$$

de donde $A_{21}u = 0$ y $A_{21} = 0$, lo que contradice que A es irreducible. \square

3.2. Teorema. (Perron-Frobenius). *Sea $0 \leq A$ matriz irreducible. Entonces:*

i) $\rho(A)$ es valor propio simple de A .

ii) Existe $v > 0$ tal que $Av = \rho(A)v$.

Supongamos que $\lambda_1 = \rho(A), \lambda_2, \dots, \lambda_h$ son los valores propios de A con módulo $\rho(A)$, entonces $\lambda_1, \dots, \lambda_h$ son las raíces de la ecuación

$$\lambda^h - \rho(A)^h = 0.$$

Más en general, todo el espectro $\text{Spec } A = \{\lambda_1, \dots, \lambda_n\}$ como subconjunto de \mathbb{C} es invariante bajo una rotación del plano por un ángulo $2\pi/h$. Si $h > 1$, A puede llevarse por medio de una permutación a la "forma cíclica"

$$A = \begin{bmatrix} 0 & A_{12} & 0 & \dots & 0 \\ 0 & 0 & A_{23} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_{h-1,h} \\ A_{h,1} & 0 & 0 & \dots & 0 \end{bmatrix},$$

donde los bloques de la diagonal principal son cuadrados.

Para demostrar el teorema comencemos por recopilar lo que ya sabemos según el teorema de la sección 3. Denotamos $\rho = \rho(A)$.

(1) Existe un vector $0 \neq v \geq 0$ con $Av = \rho v$.

Esto se sigue del hecho de que el cono positivo $K = V^+$ queda invariante bajo la acción de A .

(2) El vector v tiene todas sus coordenadas positivas, o sea $v > 0$.

En efecto, $(I_n + A)^{n-1}K \subset K^0$ por (3.1) y $(I_n + A)^{n-1}v = (1 + \rho)^{n-1}v$, de manera que $v \in K^0$.

(3) Para cada $p \geq 0$, definimos $A^p = (a_{ij}^{(p)})$. Existe una constante $c > 0$ tal que $a_{ij}^{(p)} \leq c\rho^p$ para toda $p \geq 0$.

En efecto, $a_{ij}^{(p)}v_j \leq \sum_{\ell} a_{i\ell}^{(p)}v_{\ell} = (A^p v)_i = \rho^p v_i$, $a_{ij}^{(p)} \leq \rho^p \max \left\{ \frac{v_i}{v_j} : i, j \right\}$.

(4) La multiplicidad geométrica de ρ como valor propio de A es 1.

En efecto, si u es otro vector propio de A con valor propio ρ , entonces $(I_n + A)u = (1 + \rho)u$. Por el teorema (3.5), $\rho(I_n + A) = 1 + \rho$ es valor propio simple de $I_n + A$. Entonces u y v son múltiplos escalares.

(5) ρ es valor propio simple de A .

En efecto, si definimos $P = \frac{1}{\rho}A$, $P^q = (p_{ij}^{(q)})$, tenemos $0 \leq p_{ij}^{(q)} \leq c$ para toda $q \geq 0$, donde c es la constante construida en (3). Luego, la multiplicidad geométrica y la algebraica de 1 como valor propio de P coinciden. Por (4), la multiplicidad geométrica de 1 como valor propio de P es 1. Entonces 1 es valor propio simple de P , o equivalentemente, ρ es valor propio simple de A .

Esto concluye la demostración de la primera parte del teorema. Antes de demostrar la segunda, probaremos algunos resultados auxiliares.

3.3. Proposición. (Principio min-max de Wielandt). *Sea $0 \leq A$ matriz irreducible. Para cada vector $0 \neq x = (x_1, \dots, x_n)$ definamos*

$$r_x = \min_{1 \leq i \leq n} \left\{ \frac{(Ax)_i}{x_i} : x_i \neq 0 \right\}, r^x = \max_{1 \leq i \leq n} \left\{ \frac{(Ax)_i}{x_i} : x_i \neq 0 \right\}.$$

Entonces,

$$\rho = \max_{0 \neq x \geq 0} r_x = \min_{0 \neq x \geq 0} r^x.$$

Demostración: Sea $r = \max_{0 \neq x \geq 0} r_x$. Demostraremos que r es un valor propio de A , lo que implica que $r \leq \rho$. Como, además, si $0 < v$ con $Av = \rho v$, se tiene $r_v = \rho$, entonces $r = \rho$.

En efecto,

$$\begin{cases} r \geq r_v = \rho > 0. & \text{Además, si } 0 \leq x \neq 0 \text{ se tiene} \\ r_x x \leq Ax \\ r_x y \leq Ay & \text{con } y = (I_n + A)^{n-1}x > 0, \end{cases}$$

de donde $r_x \leq y$, $r = \max \{r_x : x > 0, \|x\| = 1\}$.

Como la función r_x es continua en $\{x: x > 0, \|x\| = 1\}$, entonces existe $u > 0, \|m\| = 1$ con $r = r_u$. Supongamos que $ru \neq Au$. Entonces $Au - ru \geq 0$ y haciendo $x = (I_n + A)^{n-1}u > 0$ se tiene

$$Ax - rx > 0.$$

Por tanto podemos encontrar $\varepsilon > 0$, de manera que $r_x \geq r + \varepsilon > r$, lo cual es absurdo. Así, $ru = Au$, que es lo que deseábamos probar. \square

3.4. Lema. Sean $A = (a_{ij})$ y $C = (c_{ij})$ dos matrices de tamaño $n \times n$, donde suponemos que A es irreducible y $|c_{ij}| \leq a_{ij}$ para toda pareja (i, j) . Entonces todo valor propio λ de C satisface

$$|\lambda| \leq \rho(A).$$

Además, se da $|\lambda| = \rho(A)$ si y solamente si

$$C = e^{i\varphi} D A D^{-1},$$

donde $e^{i\varphi} = \lambda/\rho(A)$ y $D = (d_{ij})$ es una matriz diagonal con $|d_{ii}| = 1, i = 1, \dots, n$.

Demostración: Denotemos $C^+ = (|c_{ij}|)$. Tomemos λ como un valor propio de C con vector propio asociado y , esto es, $Cy = \lambda y$. Entonces,

$$|\lambda|y^+ \leq C^+y^+ \leq Ay^+.$$

Luego, $|\lambda| \leq r_{y^+} \leq \rho(A) =: \rho$, por (4.3).

Supongamos ahora que $|\lambda| = \rho(A)$. Por la demostración de (4.3), y^+ es vector propio de A y, por (4.2), y^+ tiene todas sus coordenadas positivas. Luego, $\rho y^+ = C^+y^+$. Como $C^+ \leq A$, obtenemos que $C^+ = A$.

Hagamos $y = (y_1, \dots, y_n)$, con $y_j = |y_j|e^{i\varphi_j}, j = 1, \dots, n$.

Definamos $D = \text{diag}(e^{i\varphi_1}, \dots, e^{i\varphi_n})$ de modo que

$$y = Dy^+.$$

Probaremos que $C = e^{i\varphi} D A D^{-1}$. Hagamos primeramente $B = e^{-i\varphi} D^{-1} C D$. Entonces, $By^+ = e^{-i\varphi} D^{-1} C y = e^{-i\varphi} \lambda y^+ = \rho y^+$. Luego

$By^+ = Ay^+$, con $B^+ \leq A$. Entonces $B^+ = A$, lo que también implica que $B^+y^+ = \rho y^+$.

Como $y^+ > 0$, esto sólo es posible si $B = B^+$. O sea,

$$A = e^{-i\varphi} D^{-1} C D,$$

que es lo que queríamos demostrar. \square

Concluimos ahora con la *prueba del teorema*. Supongamos que $A \geq 0$ tiene exactamente h valores propios con módulo ρ :

$$\lambda_0 = \rho e^{i\varphi_0}, \lambda_1 = \rho e^{i\varphi_1}, \dots, \lambda_{h-1} = \rho e^{i\varphi_{h-1}}$$

con $0 = \varphi_0 < \varphi_1 < \dots < \varphi_{h-1} < 2\pi$.

Hagamos $C = A$ y $\lambda = \lambda_j$ en el lema. Obtenemos para cada $j = 0, \dots, h-1$,

$$A = e^{i\varphi_j} D_j A D_j^{-1},$$

donde D_j es una matriz diagonal con $D_j^+ = I_n$. En particular, no sólo $\lambda_0 = \rho$, sino todos los valores propios $\lambda_1, \dots, \lambda_{h-1}$ de A son simples. Tomamos $y^{(0)} = z > 0$ como vector propio correspondiente a λ_0 . Fijamos $y^{(j)}$ como vector propio correspondiente a λ_j con $y^{(j)+} = z$ y su primera entrada positiva. De esta manera, D_j está determinada en forma única con la primera entrada diagonal igual a 1. Observemos en particular que $D_0 = I_n$.

(a) Mostraremos que los números $e^{i\varphi_0}, \dots, e^{i\varphi_{h-1}}$ y las matrices D_0, \dots, D_{h-1} forman dos grupos abelianos isomorfos.

En efecto, si $j, \ell \in \{0, \dots, h-1\}$, tenemos

$$A = e^{i(\varphi_j + \varphi_\ell)} D_j D_\ell A D_\ell^{-1} D_j^{-1}$$

y $D_j D_\ell z$ es un vector propio de A con valor propio $\rho e^{i(\varphi_j + \varphi_\ell)}$. Entonces $e^{i(\varphi_j + \varphi_\ell)} = e^{i\varphi_t}$ para alguna t y $D_j D_\ell = D_t$. Similarmente para $\varphi_j - \varphi_\ell$.

(b) $\text{Spec}(A)$ es cerrado bajo la multiplicación por $e^{i2\pi/h}$.

Como $\{e^{i\varphi_0}, \dots, e^{i\varphi_{h-1}}\}$ es un grupo con h elementos, entonces $e^{i\varphi_j}$ es raíz h -ésima de la unidad, $j = 0, \dots, h-1$. Luego $\varphi_j = \frac{2j\pi}{h}$, $j = 0, \dots, h-1$. Esto es, $\lambda_0, \dots, \lambda_{h-1}$ son un conjunto completo de soluciones de la ecuación $\lambda^h - \rho^h = 0$. En particular obtenemos

$$A = e^{i2\pi/h} D_1 A D_1^{-1},$$

que implica la afirmación (b).

(c) Hagamos $D = D_1$. Por una permutación de A (simultánea con D) podemos arreglar D en forma diagonal

$$D = \begin{bmatrix} \eta_0 E_0 & & & 0 \\ & \eta_1 E_1 & & \\ \dots & & \dots & \\ 0 & & & \eta_{s-1} E_{s-1} \end{bmatrix} \quad A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & A_{22} & \dots & A_{2s} \\ \dots & \dots & \dots & \dots \\ A_{s1} & A_{s2} & \dots & A_{ss} \end{bmatrix},$$

con E_j como matriz unidad y $\eta_j = e^{i\psi_j}$, $\psi_j = n_j \frac{2\pi}{h}$, $0 < n_0 < \dots < n_{s-1} < h$. Partimos A en bloques de tamaño correspondiente a los de D .

Haciendo $\varepsilon = e^{i2\pi/h}$, tenemos $A = \varepsilon D A D^{-1}$. Equivalentemente,

$$\varepsilon A_{pq} = \frac{\eta_{q-1}}{\eta_{p-1}} A_{pq}, \quad p, q = 1, \dots, s.$$

Luego $\varepsilon = \frac{\eta_{q-1}}{\eta_{p-1}}$ o $A_{pq} = 0$.

Tomemos $p = 1$. Como no todas las matrices A_{12}, \dots, A_{1s} pueden ser 0, entonces uno de los números $\frac{\eta_1}{\eta_0}, \dots, \frac{\eta_{s-1}}{\eta_0}$ es igual a ε . Como $\eta_0 = 1$, entonces $n_1 = 1$, $\eta_1 = \varepsilon$ y $A_{11} = A_{1s} = \dots = A_{1s} = 0$. Procediendo similarmente con otras p , obtenemos

$$A = \begin{bmatrix} 0 & A_{12} & 0 & \dots & 0 \\ 0 & 0 & A_{13} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_{s-1,s} \\ A_{s1} & A_{s2} & A_{s3} & \dots & A_{ss} \end{bmatrix}$$

con $n_1 = 1, n_2 = 2, \dots, n_{s-1} = s - 1$. Pero entonces,

$$\frac{\eta_{q-1}}{\eta_{s-1}} = e^{(q-s) \frac{2\pi}{h} i}, \quad q = 1, \dots, s,$$

y alguno de estos números debe ser ε . Esto sólo es posible si $s = h$ y $q = 1$. Luego, $A_{s2} = \dots = A_{ss} = 0$. Así se obtiene el resultado deseado. Esto concluye completamente la demostración del Teorema de Perron-Frobenius. \square

3.5. Procedemos en primer lugar a generalizar el teorema de Perron-Frobenius al caso de una matriz no negativa (no necesariamente irreducible).

Teorema. Sea $0 \leq A$ una matriz $n \times n$ no negativa. Entonces:

- El radio espectral $\rho(A)$ es un valor propio de A .
- Si λ es un valor propio de A con $|\lambda| = \rho(A)$, entonces $\lambda = \varepsilon \rho(A)$, con $\varepsilon^m = 1$ para alguna $m \leq n$.
- Existe un vector propio $0 \neq v \geq 0$ con $Av = \rho(A)v$.
- Para cada $0 < x$ se tiene

$$\rho(A) \leq \max_i \frac{(Ax)_i}{x_i}.$$

Demostración: (a) y (c) se siguen inmediatamente de la sección (2). Si A no es irreducible, podemos escribirla como

$$A = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix},$$

donde B y D son matrices cuadradas, B de tamaño $k \times k$. Se tiene entonces que

$$\text{Spec}(A) = \text{Spec}(B) \cup \text{Spec}(D).$$

Si λ es valor propio de B con $|\lambda| = \rho(A)$, por hipótesis de inducción $\lambda = \varepsilon \rho(A)$ con $\varepsilon^m = 1$.

Consideremos un vector $0 < x$. Podemos partirlo como $x = \begin{pmatrix} y \\ z \end{pmatrix}$ en forma compatible con la partición de A . Entonces

$$\begin{aligned} \rho(B) &\leq \max_{1 \leq i \leq k} \frac{(By)_i}{y_i} = \max_{1 \leq i \leq k} \frac{(Ay)_i}{y_i} \leq \max_{1 \leq i \leq n} \frac{(Ax)_i}{x_i} \\ \rho(D) &\leq \max_{k+1 \leq i \leq n} \frac{(Dz)_i}{z_i} \leq \max_{k+1 \leq i \leq n} \frac{(Cy)_i + (Dz)_i}{y_i} \\ &= \max_{k+1 \leq i \leq n} \frac{(Ax)_i}{x_i} \leq \max_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}. \end{aligned}$$

Luego $\rho(A) = \max\{\rho(B), \rho(D)\} \leq \max_{1 \leq i \leq n} \frac{(Ax)_i}{x_i}$. \square

3.6. Decimos que la matriz no negativa A es primitiva si

- A es irreducible, y
- para todo valor propio $\rho(A) \neq \lambda \in \text{Spec}(A)$, se tiene $|\lambda| < \rho(A)$ (esto es, $h = 1$ en el teorema de Perron-Frobenius).

Observe que si A es primitiva, entonces A^T es también primitiva.

Teorema. Sea $0 \leq A$, matriz $n \times n$ primitiva. Sea $\rho = \rho(A)$. Considere mos vectores $z > 0$, $y > 0$ tales que

$$Az = \rho z, \quad A^T y = \rho y, \quad y^T z = 1.$$

Entonces el límite

$$P^\infty = \lim_{k \rightarrow \infty} \rho^{-k} A^k = \begin{pmatrix} z_1 y_1 & \cdots & z_1 y_n \\ & & \vdots \\ z_n y_1 & \cdots & z_n y_n \end{pmatrix}$$

existe y tiene rango 1.

Demostración: Como A es positiva, entonces la matriz

$$P = \frac{1}{\rho} A$$

tiene $\rho(P) = 1$ valor propio simple. Existe entonces T invertible tal que

$$T^{-1} P T = \begin{bmatrix} 1 & 0 \\ 0 & B \end{bmatrix},$$

de forma que $\rho(B) < 1$. Hemos demostrado que, en ese caso,

$$\lim_{k \rightarrow \infty} B^k = 0.$$

Por tanto $T^{-1} P^\infty T = \lim_{k \rightarrow \infty} T^{-1} P^k T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ existe.

Consideremos $T = \begin{bmatrix} s_1 & & \\ \vdots & * & \\ s_n & & \end{bmatrix}$ y $T^{-1} = \begin{bmatrix} t_1 & \cdots & t_n \\ & * & \end{bmatrix}$, de forma que

$$P^\infty = T \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} T^{-1} = \begin{pmatrix} s_1 t_1 & \cdots & s_1 t_n \\ & & \vdots \\ s_n t_1 & \cdots & s_n t_n \end{pmatrix},$$

que tiene $rk P^\infty \leq 1$.

Como antes, $Az = \rho z$, entonces $Pz = z$. Luego, $P^\infty z = z$. Como también $P^\infty \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix}$ y el rango de P^∞ es 1, entonces z y $\begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix}$ son múltiplos escalares. Podemos suponer $z_j = s_j$, $1 \leq j \leq n$. De la misma manera se sigue que $y_j = t_j$, $1 \leq j \leq n$. \square

Ejercicio: Caracterice las matrices de tamaño 2×2 que son primitivas.

3.7. Un sencillo criterio de primitividad es el siguiente.

Proposición. Sea $0 \leq A$ una matriz $n \times n$. Son equivalentes:

- Existe k_0 tal que para toda $k \geq k_0$, $A^k > 0$.
- Existe k tal que $A^k > 0$.
- A es primitiva.

Demostración: (a) \Rightarrow (b) es clara.

(b) \Rightarrow (c): Probemos primero que A es irreducible. Si no existe una matriz permutación tal que

$$P A P^{-1} = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix}$$

con B y D matrices cuadradas, entonces

$$P A^k P^{-1} = \begin{bmatrix} B^k & 0 \\ * & D^k \end{bmatrix},$$

que no es positiva, entonces A^k tampoco es positiva. Luego, A es irreducible.

Sean $\lambda_1, \dots, \lambda_k$ los valores propios de A con $|\lambda_i| = \rho(A)$. Deseamos probar que $k = 1$. Sean v_1, \dots, v_h vectores propios de A con

$$A v_i = \lambda_i v_i.$$

Entonces, $A^{kh} v_i = \lambda_i^{kh} v_i = \rho(A)^{kh} v_i$ y la matriz $A^{kh} > 0$ es irreducible. Por el teorema de Perron-Frobenius, $\rho(A^{kh}) = \rho(A)^{kh}$ es valor propio simple de A^{kh} , entonces $h = 1$, como deseábamos probar.

(c) \Rightarrow (a): Con la notación de (3.6),

$$\lim_{k \rightarrow \infty} \rho^{-k} A^k = \begin{pmatrix} z_1 y_1 & \cdots & z_1 y_n \\ & \vdots & \\ z_n y_1 & \cdots & z_n y_n \end{pmatrix} > 0;$$

por tanto, $\rho^{-k} A^k > 0$ para k suficientemente grande. \square

4. MATRICES ESTOCÁSTICAS

4.1. Analicemos con detalle el caso 2×2 de una matriz

$$A = \begin{bmatrix} 1 - \alpha & \beta \\ \alpha & 1 - \beta \end{bmatrix},$$

con $0 \leq \alpha, \beta \leq 1$ (en particular A es no negativa). Calculamos primero los valores propios de A . Si $\alpha = 0 = \beta$, $A = I_2$ y no tenemos nada interesante que decir. Supongamos que $\alpha \neq 0$. Los valores propios son:

$$\lambda_1 = 1 \quad \text{y} \quad \lambda_2 = 1 - \alpha - \beta,$$

con vectores propios

$$v_1 = \begin{bmatrix} \beta \\ \alpha \end{bmatrix} \quad \text{y} \quad v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

respectivamente. Observe que $\rho(A) = 1$ es valor propio simple y A es diagonalizable. Haciendo

$$S = \begin{bmatrix} \beta & 1 \\ \alpha & -1 \end{bmatrix}, \quad S^{-1} = \frac{1}{\alpha + \beta} \begin{bmatrix} 1 & 1 \\ \alpha & -\beta \end{bmatrix}$$

se tiene $S^{-1} A S = \text{diag}(\lambda_1, \lambda_2)$. Como $\lambda_2 < 1$ obtenemos

$$\begin{aligned} A^\infty &= \lim_{k \rightarrow \infty} A^k = S \left(\lim_{k \rightarrow \infty} \text{diag}(\lambda_1, \lambda_2)^k \right) S^{-1} \\ &= S \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} S^{-1} = \frac{1}{\alpha + \beta} \begin{bmatrix} \beta & \beta \\ \alpha & \alpha \end{bmatrix}. \end{aligned}$$

Observe que la suma de las entradas de una columna de A^∞ también es 1.

4.2. Recordamos que una matriz $0 \leq A$ de tamaño $n \times n$ se llama *estocástica* si

$$\sum_{j=1}^n a_{ij} = 1, \quad \text{para } i = 1, \dots, n.$$

Claramente, $0 \leq A$ es estocástica si y solamente si 1 es valor propio de A con vector propio $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

Lema. a) Si A y B son estocásticas, también AB lo es.

b) Si A es estocástica, $\rho(A) = 1$.

c) Si $A^\infty = \lim A^k$ existe, entonces A^∞ es estocástica.

Demostración: (a) Sea $A = (a_{ij})$, $B = (b_{ij})$, $AB = (c_{ij})$. Entonces

$$\sum_{j=1}^n c_{ij} = \sum_{j=1}^n \sum_{\ell=1}^n a_{i\ell} b_{\ell j} = \sum_{\ell=1}^n a_{i\ell} \left(\sum_{j=1}^n b_{\ell j} \right) = \sum_{\ell=1}^n a_{i\ell} = 1.$$

(b) Como $A^k = (a_{ij}^{(k)})$ es estocástica, entonces $a_{ij}^{(k)} \leq 1$. Luego, por (I.4.4), $\rho(A) \leq 1$. Como además 1 es valor propio de A , entonces $\rho(A) = 1$. \square

4.3. Veamos el tipo general de aplicaciones de las matrices estocásticas. Sea S un sistema que puede encontrarse en n estados diferentes, a los que denotamos por $1, \dots, n$. Podemos pensar en muchos ejemplos diferentes para ilustrar el sistema S , por ejemplo: el sistema S puede ser una colección de m focos que pueden estar encendidos o apagados, y así un estado del sistema es una función $f: \{1, \dots, m\} \rightarrow \{0, 1\}$ que indica qué focos están encendidos y cuáles están apagados; hay por tanto 2^m estados del sistema S . Este sistema está sujeto a un *proceso* (esto es, cambio de estados) \mathcal{A} , cuya acción sobre S no está perfectamente determinada. Sólo se sabe que bajo el proceso \mathcal{A} , la probabilidad de que el sistema S pase del estado i al estado j es $a_{ij} \geq 0$. Señalamos que esta probabilidad sólo depende del estado i y no de toda la historia

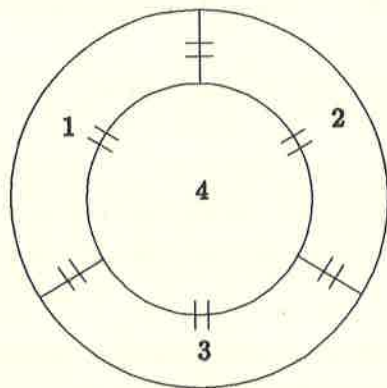
previa del sistema. A un proceso \mathcal{A} como éste lo llamamos un *proceso estocástico*.

Describiremos al proceso estocástico \mathcal{A} por medio de su *matriz de transición* $A = (a_{ij})$. En particular, la probabilidad de que el sistema S pase bajo el proceso \mathcal{A} del estado i a cualquier otro estado es $\sum_{j=1}^n a_{ij}$; por lo tanto tenemos que $\sum_{j=1}^n a_{ij} = 1, i = 1, \dots, n$.

En general, nos interesa el comportamiento del sistema S cuando se aplica el proceso \mathcal{A} en repetidas ocasiones. Con frecuencia, el problema fundamental será el de la existencia del límite $\lim_{k \rightarrow \infty} A^k$ y cómo puede calcularse este límite.

Consideraremos algunos *ejemplos*:

(a) Tenemos un laberinto con cuatro cámaras, cada una de las cuales está unida a todas las demás por medio de una puerta en la forma en que se indica en la figura.



Un experimentador instala un ratón en este laberinto. Podemos considerar al laberinto y al ratón como un sistema con cuatro estados: el sistema se encuentra en el estado i ($1 \leq i \leq 4$) cuando el ratón se ubica en la cámara i precisamente. El proceso estocástico al que está sujeto el sistema es el siguiente: supongamos que el ratón está en la cámara i ; con probabilidad a se quedará en esa cámara; con probabilidad $b = \frac{1-a}{3} \geq 0$ se moverá a la cámara j ($j \neq i$). Obtenemos entonces la matriz de transición

$$A = \begin{pmatrix} a & b & b & b \\ b & a & b & b \\ b & b & a & b \\ b & b & b & a \end{pmatrix}.$$

Nuestro problema es el siguiente: después de un largo tiempo, ¿cuál es la probabilidad de que el ratón se encuentre en la cámara i ?

En caso de que el ratón estuviera en la cámara 1 al inicio del proceso, en el momento k la probabilidad de que se encuentre en la cámara j es de

$$\left(A^k \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right)_j.$$

Por tanto, nuestro problema consiste en calcular el límite $\lim A^k$. Definimos la matriz auxiliar

$$F = \begin{bmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{bmatrix},$$

de forma que tenemos $A = (a-b)E + bF$, con $E = I_4$ y $F^2 = 4F$. En general, $F^j = 4^{j-1}F$ si $j \geq 1$. De aquí obtenemos,

$$\begin{aligned} A^k &= ((a-b)E + bF)^k = \sum_{j=0}^k \binom{k}{j} (a-b)^{k-j} (bF)^j \\ &= (a-b)^k E + \sum_{j=1}^k \binom{k}{j} (a-b)^{k-j} b^j 4^{j-1} F \\ &= (a-b)^k E - \frac{(a-b)^k}{4} F + \frac{1}{4} \sum_{j=0}^k \binom{k}{j} (a-b)^{k-j} (4b)^j F \\ &= (a-b)^k E - \frac{(a-b)^k}{4} F + \frac{1}{4} (a-b+4b)^k F \\ &= (a-b)^k \left(E - \frac{1}{4} F \right) + \frac{1}{4} F. \end{aligned}$$

En caso de que $|a - b| < 1$, tenemos que

$$\lim A^k = \frac{1}{4} F = \begin{pmatrix} \frac{1}{4} & \cdots & \frac{1}{4} \\ \vdots & & \vdots \\ \frac{1}{4} & \cdots & \frac{1}{4} \end{pmatrix},$$

lo que nos indica que después de mucho tiempo todos los estados son igualmente probables.

El caso $1 \leq a - b$ sólo puede darse si $a = 1$ y $b = 0$, lo que implica que $A = E$ y $\lim A^k = E$. En esta situación (como es obvio), el sistema permanece siempre en su estado inicial. En el caso restante, $-1 \geq a - b$ implica que $a = 0$, $b = 1$. Es claro que el límite $\lim A^k$ no existe.

(b) La instrucción primaria en México tiene una duración de 6 años. Al final de cada año el escolar se enfrenta a las siguientes opciones: dejar la escuela, repetir el curso en el que estuvo registrado ese año o bien ingresar al siguiente curso del ciclo.

Al final del año j -ésimo ($1 \leq j \leq 6$), el escolar enfrenta las diferentes opciones de acuerdo con las siguientes expectativas:

(i) con una probabilidad de $p_j < 1$ no continuará los estudios primarios;

(ii) con una probabilidad de $q_j < 1$ repetirá curso.

(iii) si $j < 6$, con una probabilidad de r_j , el estudiante será promovido al siguiente curso; o bien si $j = 6$, con una probabilidad de r_6 , el estudiante recibirá su certificado de estudios primarios.

En particular, tenemos que $p_j + q_j + r_j = 1$. Nuestro problema es calcular la probabilidad que tiene un estudiante que se encuentra en el j -ésimo año del ciclo primario para obtener su certificado después de algún tiempo. Para resolver el problema supondremos que no hay límite de tiempo en que el estudiante puede estar inscrito en el ciclo primario.

La situación de un escolar en el ciclo de educación primaria puede entenderse como un sistema con ocho estados, que numeraremos de la siguiente manera:

1: El escolar recibe su certificado de educación primaria.

2: El escolar suspendió sus estudios.

$2 + j$ ($1 \leq j \leq 6$): El escolar está en el j -ésimo año del ciclo de instrucción.

La matriz de transición del sistema es, por lo tanto,

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & p_1 & q_1 & r_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & p_2 & 0 & q_2 & r_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & p_5 & 0 & 0 & 0 & 0 & \cdots & q_5 & r_5 \\ r_6 & p_6 & 0 & 0 & 0 & 0 & \cdots & 0 & q_6 \end{bmatrix}.$$

Nuestro problema consiste en demostrar que $\lim A^k$ existe y en calcularlo. En efecto, suponiendo que la matriz $\lim A^k$ está bien definida, la entrada $(1, j + 2)$ de ella es la probabilidad de que un escolar en el año j del ciclo primario reciba alguna vez el certificado de primaria.

Para la solución de este problema requerimos demostrar el resultado técnico (5.4). Aplicándolo directamente obtenemos que el límite $\lim A^k$ existe y que es igual a

$$\lim A^k = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ b_1 & 1 - b_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_6 & 1 - b_6 & 0 & \cdots & 0 \end{bmatrix},$$

donde $b_j = q_j b_j + r_j b_{j+1}$ para $1 \leq j \leq 5$ y $b_6 = r_6 + q_6 b_6$.

De aquí se sigue que las probabilidades que conciernen a la solución de nuestro problema son las siguientes:

$$b_6 = \frac{r_6}{1 - q_6} = \frac{r_6}{p_6 + r_6},$$

$$b_j = \frac{r_j b_{j+1}}{1 - q_j} = \prod_{k=j}^6 \frac{r_k}{p_k + r_k}.$$

Ejercicios

1. Justifique que b_j es la probabilidad de que un estudiante en el año j -ésimo del ciclo escolar obtenga alguna vez el certificado de primaria.

2. Suponiendo que la probabilidad de deserción es una constante p a lo largo de todo el ciclo primario y que el índice de reprobación es

también una constante r , calcule cuál es el máximo valor de r que permite que al menos tres cuartas partes de la población del país obtengan el certificado de primaria.

4.4. Lema. (a) Sea C la matriz

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ 0 & c_{22} & \cdots & c_{2n} \\ & \vdots & & \\ 0 & 0 & \cdots & c_{nn} \end{bmatrix},$$

con entradas complejas c_{ij} , $|c_{jj}| < 1$, $j = 1, \dots, n$. Entonces $\lim_{k \rightarrow \infty} C^k = 0$.

(b) Consideremos ahora la matriz

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ & & \vdots & & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} E & 0 \\ B & C \end{bmatrix}$$

como una matriz estocástica con $0 \leq a_{jj} < 1$ para $j = 3, \dots, n$. Entonces $P = \lim_{k \rightarrow \infty} A^k$ existe y es de la forma

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ b_3 & 1 - b_3 & 0 & \cdots & 0 \\ \vdots & & & & \\ b_n & 1 - b_n & 0 & \cdots & 0 \end{bmatrix},$$

con $b_j = a_{j1} + \sum_{k=j}^n a_{jk} b_k$ ($j = 3, \dots, n$).

Demostración: (a) Se sigue de los siguientes pasos:

(1) Consideramos primero el caso más simple de una matriz

$$M = \begin{pmatrix} q & m & m & \cdots & m \\ 0 & q & m & \cdots & m \\ 0 & 0 & q & \cdots & m \\ \vdots & & & & \\ 0 & \cdots & \cdots & 0 & q \end{pmatrix},$$

con $m \geq 0$, $0 \leq q < 1$. Mostraremos que $\lim_{k \rightarrow \infty} M^k = 0$.

En efecto, sea $M = qE + N$ con $N^n = 0$. Luego

$$M^k = (qE + N)^k = \sum_{j=0}^k \binom{k}{j} q^{k-j} N^j = \sum_{j=0}^{n-1} \binom{k}{j} q^{k-j} N^j$$

$$\binom{k}{j} = \frac{k \cdots (k-j+1)}{j!} \leq \frac{k^j}{j!} \leq k^{n-1} \text{ para } j \leq n-1.$$

Como $0 \leq q < 1$, entonces $\lim_{k \rightarrow \infty} k^{n-1} q^k = 0$ y $\lim_{k \rightarrow \infty} \binom{k}{j} q^{k-j} = 0$.

(2) Si $B = (ij)$ y $|c_{ij}| \leq b_{ij}$, hacemos $C^k = (c_{ij}^{(k)})$, $B^k = (b_{ij}^{(k)})$ y entonces $|c_{ij}^{(k)}| \leq b_{ij}^{(k)}$.

(3) En el caso general, hacemos $q = \max_j |c_j|$, $m = \max_{i < j} |c_{ij}|$ y aplicamos (1) y (2).

(b) Claramente

$$A^k = \begin{pmatrix} E & 0 \\ a_{31}^{(b)} & a_{32}^{(k)} \\ \vdots & \\ a_{n1}^{(k)} & a_{n2}^{(k)} \\ & c^k \end{pmatrix},$$

donde

$$1 \geq a_{j1}^{(k+1)} = a_{j1}^{(k)} + \sum_{t=3}^n a_{jt}^{(k)} a_{t1} \geq a_{j1}^{(b)}.$$

Luego obtenemos que $b_j = \lim_{k \rightarrow \infty} a_{j1}^{(k)}$ existe y

$$P = \lim A^k = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ b_3 & 1 - b_3 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_n & 1 - b_n & 0 & \cdots & 0 \end{pmatrix}.$$

Como $P = AP$, entonces

$$b_j = a_{j1} + \sum_{k=j}^n a_{jk} b_k \quad (3 \leq j \leq n);$$

luego

$$\begin{aligned} b_n(1 - a_{nn}) &= a_{n1} \\ b_{n-1}(1 - a_{n-1,n-1}) &= a_{n-1,1} + a_{n-1,n} b_n \\ &\vdots \\ b_3(1 - a_{33}(1 - a_{33})) &= a_{31} + a_{34} + \cdots + a_{3n} b_n, \end{aligned}$$

y como $1 - a_{jj} \neq 0$ ($3 \leq j \leq n$) se pueden resolver estas ecuaciones para b_j , obteniéndose el resultado deseado. \square

Ejercicios

1. Un examen consta de 50 preguntas. Cuando un estudiante promedio contesta correctamente a una pregunta tiene una probabilidad de $\frac{3}{4}$ de responder correctamente la siguiente pregunta. Suponga que la primera pregunta fue contestada correctamente por el 90% de los estudiantes. Entonces:

- i) ¿Cuál fue la calificación promedio del examen?
- ii) ¿Cuántos estudiantes contestaron correctamente la última pregunta?

2. Una fábrica tiene n máquinas del mismo tipo y desea mantener trabajando la mayor cantidad posible. Por ello, la política de la fábrica es que toda máquina que se encuentre descompuesta al principio de una semana sea reparada para el inicio de la semana siguiente. La probabilidad de que una máquina se descomponga a lo largo de una semana es de p con $0 < p < 1$. Calcule:

i) La probabilidad de que en una semana fallen exactamente j de las máquinas que al inicio de la semana estaban en buen estado.

ii) Encuentre la matriz de transición de este proceso (definiendo el estado i como la situación en la que al principio de una semana hay i máquinas que funcionan correctamente y $n - i$ que están descompuestas).

iii) ¿Cuál es la probabilidad de que una máquina dada se encuentre funcionando correctamente después de mucho tiempo? ¿De qué se encuentre descompuesta?