Florentino García Santos: In Memoriam
Universidad de Granada, 2011, págs. 1-8

Jugando con matrices positivas: eficiencia de un estado inicial.

Margarita Arias • Juan Campos •

Resumen Conocedores del interés que el profesor García Santos mostraba por todos los temas relacionados con la docencia, en este trabajo presentamos algunos resultados, en parte novedosos, sobre el comportamiento asintótico de ciertos sistemas dinámicos generados por matrices positivas, que bien podrían estar incluidos en uno de los cursos de álgebra lineal que tantas veces impartió Floro.

1. Introducción

Una matriz A, cuadrada de orden k y positiva, es decir, con todas sus entradas positivas o cero, permite definir un sistema dinámico en $\mathbb{R}_+^k = [0, +\infty)^k$ por medio de la ecuación

$$\vec{v}_{n+1} = A\vec{v}_n. \tag{1}$$

Son muy diversos los modelos matemáticos que se ajustan a este tipo de ecuaciones. Por ejemplo, algunos modelos de poblaciones estructuradas por edad, como el modelo de Leslie, o los clásicos modelos de economía de Leontief (ver, por ejemplo, [1], [3]). La matriz A es la llamada *matriz de transición* y $\vec{v}_n$ es el *vector de estados* en la etapa n-ésima.

Aunque la ecuación (1), conocido el vector de estados inicial, $\vec{v}_0$, permite determinar por recurrencia el vector de estados en cualquier etapa posterior, no proporciona de forma inmediata un retrato general de la dinámica del proceso que describe. Para realizar este estudio es necesario investigar los valores propios y los vectores propios

Margarita Arias
Departamento de Matemática Aplicada, Facultad de Ciencias, Universidad de Granada, 18071-Granada
marias@ugr.es
Juan Campos
Departamento de Matemática Aplicada, Facultad de Ciencias, Universidad de Granada, 18071-Granada
campos@ugr.es

de la matriz. Como A es positiva es conocido que

$$\lambda_p = \max\{|\lambda|/\ \lambda \in \sigma(A)\} \tag{2}$$

es valor propio y que $\ker(A - \lambda_p I_k) \bigcap \mathbb{R}^k_+ \neq \emptyset$. Nos referiremos a λ_p como *valor propio principal*. Se dice que el valor propio principal es *dominante* cuando es algebráicamente simple, esto es, es simple como raíz del polinomio característico, y

$$\lambda_p > |\lambda|\ , \forall \lambda \in \sigma(A) - \{\lambda_p\}. \tag{3}$$

Existen muchos resultados que permiten asegurar la dominancia del valor propio principal. Probablemente el más conocido sea el clásico Teorema de Perron-Fröbenious, [4] (ver también [3]), que asegura que λ_p es dominante si la matriz A es tal que para algún $s \in \mathbb{N}$, A^s tiene todas sus entradas estrictamente positivas (lo que algunos autores denominan *matrices ergódicas*).

Tambien se conocen, por ejemplo, condiciones sencillas para matrices de transición de un modelo de Leslie. Recordemos que los modelos de Leslie son modelos para estudiar la evolución del número de hembras de una población estructurada por edades. En su versión más simple, se divide la población en k grupos de edad equiespaciados de duración d, de forma que la edad máxima de supervivencia de un individuo es kd y se recuenta cada d unidades de tiempo; la evolución de la población se describe mediante la ecuación (1) con

$$A = \begin{pmatrix} f_1 & f_2 & \cdots & f_{k-1} & f_k \\ s_1 & 0 & \cdots & 0 & 0 \\ 0 & s_1 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & s_{k-1} & 0 \end{pmatrix}$$

donde $f_i \geq 0$ es el número medio de crías hembras que tiene cada hembra del grupo i (la llamada tasa de fertilidad), y $0 < s_i < 1$ es la probabilidad de que un individuo del grupo i sobreviva al siguiente (la tasa de supervivencia del grupo). Se demuesta, [2], que el valor propio principal de A es dominante siempre que existan dos grupos fértiles consecutivos, es decir, cuando $f_i f_{i+1} > 0$ para algún $i \in \{1, \ldots, k-1\}$.

Bajo condiciones de dominancia es evidente que $\lambda_p > 0$. Además, si denotamos por $\vec{p}$ el único vector propio cuyas componentes suman 1, se puede demostrar (ver, por ejemplo, [2]) que dado cualquier vector de estados inicial $\vec{v}_0 \in \mathbb{R}^k_+$, existe una constante $c(\vec{v}_0) \in \mathbb{R}_+$ tal que, si $\vec{v}_n = A^n \vec{v}_0,\ n \in \mathbb{N}$, es la solución de (1) que parte de ese estado inicial, entonces

$$\frac{1}{\lambda_p^n} \vec{v}_n \to c(\vec{v}_0)\vec{p},\ n \to +\infty. \tag{4}$$

En otras palabras, la solución de (1) que parte del estado inicial $\vec{v}_0$ se comporta a largo plazo como $c(\vec{v}_0)\lambda_p^n\vec{p}$. Vamos a llamar a la constante $c(\vec{v}_0)$ *eficiencia* del estado $\vec{v}_0$. El cálculo de la eficiencia de un estado inicial requiere del conocimiento de una base de $\mathbb{R}^k$ en la que A se escriba en algún tipo de forma canónica y su obtención no es en absoluto trivial.

El propósito de este trabajo es relacionar la eficiencia con un vector propio de la matriz traspuesta de A, A^t, asociado al valor propio dominante. Hay que observar que, como es evidente a partir de (2) y (3), el valor propio principal y la condición de dominancia son invariantes por trasposición.

Aunque toda esta teoría es válida para cualquier matriz positiva, nuestro interés se centra en las matrices de transición de los modelos lineales discretos de dinámica de poblaciones, como el de Leslie, puesto que el conocimiento de la eficiencia de cualquier estado inicial permite conocer a priori una estimación precisa del número de individuos en cada estado, no sólo su comportamiento asintótico. Como consecuencia se puede determinar, por ejemplo, la estrategia más conveniente de distribución inicial de la población en procesos de repoblación.

2. Eficiencia de un estado inicial.

Dado $\vec{v} \in \mathbb{R}_+^k$, denotamos por $|\vec{v}| = \sum_{i=1}^k v_i$ donde v_i es la i-ésima componente de $\vec{v}$. Con esta notación $\vec{p}$ cumple $|\vec{p}| = 1$ y es un vector de tantos por uno que en virtud de (4) es el límite en n de

$$\frac{\vec{v}_n}{|\vec{v}_n|}$$

para cualquier solución de (1) no eventualmente nula. Podemos, por tanto, llamar a $\vec{p}$ el *vector de proporciones asintóticas* del sistema dinámico (1).

La solución de (1) dada por

$$\vec{p}_n = \lambda_p^n \vec{p} \tag{5}$$

representa la evolución de un individuo inicialmente distibuido en la proporción asintótica del modelo. Como consecuencia de la normalización $|\vec{p}_n| = \lambda_p^n$.

Llamamos *eficiencia* de un estado inicial $\vec{v}_0$ al "límite"del cociente entre la solución $\vec{v}_n$ que comienza en dicho estado y la solución (5) cuando $n \to \infty$. Este límite puede entenderse componente a componente en las componentes de $\vec{p}$ no nulas o el cociente entre la suma total de las componentes de cada vector, y coincide con el valor de $c(\vec{v}_0)$ de la fórmula (4). Se deduce de (4) que $c : \mathbb{R}_+^k \to \mathbb{R}$ es la restricción de una aplicación lineal definida en $\mathbb{R}^k$ (de hecho (4) se puede hacer para cualquier $\vec{v}_0 \in \mathbb{R}^k$).

Si consideramos $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_k\}$ la base canónica de $\mathbb{R}^k$, el valor $c_i = c(\vec{e}_i)$ se puede ver como la eficiencia de un individuo del estado i en relación con un individuo distribuido en la proporción asintótica del modelo.

Definición. Denominamos *vector de eficiencia* del modelo al vector formado por las eficiencias de los vectores de la base caónica de $\mathbb{R}^k$, es decir, $\vec{c} = (c_1, \ldots, c_k)^t$.

Con esta notación se tiene que, para cualquier $\vec{v} \in \mathbb{R}^k$,

$$c(\vec{v}) = \sum_{i=1}^{k} c_i v_i := < \vec{c}, \vec{v} >$$

donde $<,>$ denota el producto escalar usual en $\mathbb{R}^k$.

Teorema 1 *Supongamos que el vector propio principal λ_p de una matriz positiva A es dominante. Entonces, el vector de eficiencia $\vec{c}$ es el vector propio de A^t asociado a λ_p que cumple $< \vec{c}, \vec{p} >= 1$.*

Demostración. Sea w un vector propio de A^t asociado a λ_p. Dado $\vec{v} \in \mathbb{R}^k$,

$$< \vec{w}, \vec{v} >= \frac{1}{\lambda_p^n} < (A^t)^n \vec{w}, \vec{v} >= \frac{1}{\lambda_p^n} < \vec{w}, A^n \vec{v} >,$$

para todo $n \in \mathbb{N}$. Haciendo $n \to \infty$ y teniendo en cuenta (4), se tiene entonces que

$$< \vec{w}, \vec{v} >= c(\vec{v}) < \vec{w}, \vec{p} > . \tag{6}$$

Si $< \vec{w}, \vec{p} >= 0$, entonces $< \vec{w}, \vec{v} >= 0$ para cualquier $\vec{v} \in \mathbb{R}^k$ y $\vec{w} = \vec{0}$, en contradición con que es un vector propio de A^t. Por tanto, $< \vec{w}, \vec{p} > \neq 0$ y

$$c(\vec{v}) = \frac{< \vec{w}, \vec{v} >}{< \vec{w}, \vec{p} >},$$

para todo $\vec{v} \in \mathbb{R}^k$. En particular,

$$\vec{c} = \frac{1}{< \vec{w}, \vec{p} >} \vec{w}, \tag{7}$$

con lo que $\vec{c}$ es el vector propio de A^t asociado a λ_p que cumple $< \vec{c}, \vec{p} >= 1$, c.q.d.

Observación. Dado que $\text{rg}(A^t - \lambda_p I) = \text{rg}(A - \lambda_p I) = k - 1$, el vector de eficiencia $\vec{c}$ es la única solución del sistema

$$(A^t - \lambda_p I)\vec{x} = \vec{0}, \tag{8}$$

que cumple $< \vec{x}, \vec{p} >= 1$. Esta última ecuación resulta de imponer que la eficiencia del vector de proporciones asintóticas del sistema dinámico es igual a 1.

Un ejemplo. En [2], ejercicio 5, pag. 345, los autores proponen un sencillo ejercicio en el que se trata de describir la distribución por edades y el comportamiento asintótico de una población estructurada en dos grupos de edad con $f_1 = 1$, $f_2 = 3$ y $s_1 = 2/3$, partiendo de un estado inicial de 100 miembros en el primer grupo de edad y ninguno en el segundo.

Es un ejercicio elemental determinar que para la matriz

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

$\lambda_p = 2$ y $\vec{p} = (3/4,\, 1/4)^t$. Estos dos datos nos permiten asegurar que la población total se hace cada vez más grande y que a largo plazo la proporción de individuos en cada grupo de edad se ajustará al vector de proporciones $\vec{p}$, es decir, un $75\,\%$ del total pertenecerá al primer grupo de edad y el $25\,\%$ restante al segundo.

Sin embargo, si calculamos la eficiencia del estado inicial podemos ser más concretos en nuestras predicciones. Se comprueba que $\vec{c} = (8/9,\, 4/3)^t$, con lo que la eficiencia del estado inicial dado es $c((100,\, 0)^t) = 800/9$ y, por tanto, la población en el recuento n, $\vec{v}_n$, cumple

$$\vec{v}_n = A^n \vec{v}_0 = \begin{pmatrix} 1 & 3 \\ 2/3 & 0 \end{pmatrix}^n \begin{pmatrix} 100 \\ 0 \end{pmatrix} \simeq 2^n \frac{800}{9} \begin{pmatrix} 3/4 \\ 1/4 \end{pmatrix}.$$

La siguiente tabla presenta los datos reales y los obtenidos mediante esta aproximación.

n	0	5	10	15	20
$A^n \begin{pmatrix} 100 \\ 0 \end{pmatrix}$	100 0	2100 733	68300 22733	2184500 728200	69905100 23301667
$2^n \frac{800}{9} \begin{pmatrix} 3/4 \\ 1/4 \end{pmatrix}$	67 22	2133 711	68267 22756	2184533 728178	69905067 23301689

El vector de eficiencia $\vec{c}$ indica la mejor estrategia a seguir en este modelo con vistas a la repoblación: 100 individuos del primer estadio se comportan a largo plazo como $\frac{800}{9} \simeq 88{,}89$ individuos repartidos según las proporciones de $\vec{p} = (3/4,\, 1/4)^t$, mientras que si partimos de 100 individuos en el segundo estadio, como $c((0,\, 100)^t) = 400/3 \simeq 133{,}33$, su evolución corresponderá a la de 133,33 individuos repartidos según $\vec{p}$. También es posible analizar hasta qué punto es rentable pagar un precio más alto por individuos adultos.

3. El caso no dominante.

Cuando el valor propio principal λ_p no es dominante, el estudio del comportamiento de $A^n \vec{v}_0$ para una condición inicial dada, $\vec{v}_0$, se puede llegar a complicar mucho, tanto por la posible existencia de valores propios complejos con norma λ_p, como por la multiplicidad del valor propio principal. Sin embargo, cuando λ_p es geométricamente simple aún es posible obtener resultados semejantes a los de la sección anterior.

Diremos que λ_p es *geométricamente dominante* si verifica (3) y además

$$\dim(\ker\,(A - \lambda_p I)) = 1.$$

En este caso, $\vec{p}$ viene naturalmente definido como el único elemento de ker $(A - \lambda_p I)$ con $|\vec{p}| = 1$. Se demuestra (ver anexo)

Lema 1 *Para cada $\vec{v} \in \mathbb{R}^k$ existe una constante $c(\vec{v})$ tal que*

$$\frac{1}{n^{s-1}\lambda_p^n} A^n \vec{v} \to c(\vec{v})\vec{p}, \tag{9}$$

donde s es la multiplicidad algebráica de λ_p (multiplicidad como raíz del polinomio característico). Además existen vectores tales que $c(\vec{v}) \neq 0$.

Podemos por tanto llamar, de forma análoga al caso dominante, a la constante $c(\vec{v})$ eficiencia del estado inicial $\vec{v}$ y el vector de eficiencia del modelo será $\vec{c} = (c_1, \ldots, c_k)^t$, con $c_i = c(\vec{e}_i),\ i = 1, \ldots, k$. Obsérvese que $\vec{c} \neq \vec{0}$ puesto que hay vectores con $c(\vec{v}) \neq 0$. Con esta notación,

$$c(\vec{v}) =< \vec{c}, \vec{v} >,\ \vec{v} \in \mathbb{R}.$$

Como en el caso dominante, el resultado anterior permite dar una estimación del número de individuos en el estado n, conocido el estado inicial $\vec{v}_0$, siempre que $c(\vec{v}_0) \neq 0$:

$$|\vec{v}_n| = |A^n \vec{v}_0| \simeq n^{s-1}\lambda_p^n c(\vec{v}_0),$$

y la proporción en cada clase tiende a asemejarse a la determinada por el vector de proporciones asintóticas, $\vec{p}$. Como en el caso dominante, se tiene el siguiente resultado, cuya demostración incluimos también en el anexo final.

Teorema 2 *Sea A una matriz positiva con valor propio principal λ_p geométricamente dominante. Entonces, el vector de eficiencia, $\vec{c}$, es un vector propio de la matriz A^t asociado a λ_p.*

Observación 1. Dado que $A^n\vec{p} = \lambda_p^n \vec{p}$, es evidente que $c(\vec{p}) = 0$ y, por tanto, no está claro el tipo de normalización necesario para obtener el vector de eficiencia en este caso. Un problema abierto es obtener una normalización apropiada que permita determinar $\vec{c}$ sin necesidad de conocer una base de Jordan de la matriz de partida.

Tampoco es obvio que halla una solución modelo con la que comparar, como ocurre en el caso dominante con la determinada por el vector de proporciones asintóticas.

Observación 2. Como hemos comentado antes, cuando el valor propio principal no es geométricamente dominante la dinámica de la ecuación (1) puede llegar a complicarse enormemente. Si lo que ocurre es que hay valores propios complejos de norma λ_p, es posible todavía obtener alguna información puesto que las oscilaciones están bien estudiadas y se trata de tomar valor medio en un periodo. Cuando λ_p es geométricamente múltiple, cada vector propio asociado a λ_p puede tener una eficiencia diferente y parece difícil determinar a priori cual va a ser el comportamiento de un estado inicial.

4. Anexo: demostración del Lema 1 y del Teorema 2.

Por simplicidad y, principalmente, por cuestiones de espacio, vamos a demostrar únicamente el caso $s = 2$. Sea $J = \text{ diag } (J_1, \ldots, J_r)$ una forma canónica de Jordan de la matriz A. Como $\dim(\text{ker } (A - \lambda_p I)) = 1$ y la multiplicidad algebráica de λ_p es 2, hay un único bloque de Jordan asociado a este valor propio y tiene orden 2. Podemos suponer sin perdida de generalidad que es el primero, es decir, $J_1 = \lambda_p I + N_2$, siendo N_2 la matriz nilpotente de orden 2

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

Entonces, $\frac{1}{n\lambda_p^n} A^n = P \text{ diag } (\frac{1}{n\lambda_p^n} J_1^n, \ldots, \frac{1}{n\lambda_p^n} J_r^n) P^{-1}$, para cierta matriz P no singular. Por (3), es evidente que $\frac{1}{n\lambda_p^n} J_i^n \to 0,\ n \to \infty$ para todo $i = 2, \ldots, k$, luego hemos de centrarnos únicamente en el primer bloque de Jordan. Se tiene

$$\frac{1}{n\lambda_p^n} J_1^n = \frac{1}{n}(I + \frac{1}{\lambda_p} N_2)^n = \frac{1}{n}(I + n\frac{1}{\lambda_p} N_2),$$

ya que $N_2^2 = 0$ y

$$\lim_n \frac{1}{n\lambda_p^n} J_1^n = \frac{1}{\lambda_p} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Fijado $\vec{v} \in \mathbb{R}$, se tiene entonces que, poniendo $P = (\vec{w}_1, \ldots, \vec{w}_k)$,

$$\lim_n \frac{1}{n\lambda_p^n} A^n \vec{v} = \frac{1}{\lambda_p} (\vec{0}, \vec{w}_1, \vec{0}, \ldots, \vec{0}) P^{-1} \vec{v},$$

pero si $\vec{v}$ tiene coordenadas $(\alpha_1, \ldots, \alpha_k)$ respecto de la base $\{\vec{w}_1, \ldots, \vec{w}_k\}$, es claro que $P^{-1}\vec{v} = (\alpha_1, \ldots, \alpha_k)^t$ y

$$\lim_n \frac{1}{n\lambda_p^n} A^n \vec{v} = \frac{\alpha_2}{\lambda_p} \vec{w}_1.$$

Basta tener en cuenta que, por construcción, $A\vec{w}_1 = \lambda_p \vec{w}_1$ para concluir la demostración del Lema 1.

En cuanto a la demostración del Teorema 2, también por construcción, se observa que $A\vec{w}_2 = \lambda_p \vec{w}_2 + \vec{w}_1$. Por tanto, fijado $\vec{z} \in \ker\,(A^t - \lambda_p I)$,

$$< \vec{z}, \vec{w}_1 >=< \vec{z}, A\vec{w}_2 - \lambda_p \vec{w}_2 >=< (A^t - \lambda_p I)\vec{z}, \vec{w}_2 >= 0.$$

Además, para cada $i \in \{3, \ldots, k\}$, $A^n \vec{w}_i$ está en el subespacio invariante generado por $\{\vec{w}_3, \ldots, \vec{w}_k\}$ y, por tanto,

$$\frac{1}{\lambda_p^n} A^n \vec{w}_i \to \vec{0},\ n \to \infty, \tag{10}$$

con lo que $< \vec{z}, \vec{w}_i >= 0,\ i \in \{3, \ldots, k\}$ y $\vec{z}$ es ortogonal al subespacio generado por $\{\vec{w}_1, \vec{w}_3, \ldots, \vec{w}_k\}$.

Por otra parte, como $\vec{w}_1 \in \ker\,(A - \lambda_p I),\ < \vec{c}, \vec{w}_1 >= 0$ y teniendo en cuenta (10) y la definición de eficiencia, es evidente que $< \vec{c}, \vec{w}_i >= 0,\ i \in \{3, \ldots, k\}$, con lo que $\vec{c}$ es también ortogonal al subespacio generado por $\{\vec{w}_1, \vec{w}_3, \ldots, \vec{w}_k\}$ y los vectores $\vec{c}$ y $\vec{z}$ han de ser colineales, lo que permite concluir la demostración.

Agradecimientos

Los autores agradecen la financiación del proyecto D.G.I. MTM2008-02502, Ministerio de Educación y Cultura.

Referencias

[1] H. Anton, C. Rorres, *Elementary Linear Algebra with Applications*, John Wiley and Sons, Inc., New York, 1987.

[2] F. Brauer, C. Castillo-Chávez, *Mathematical Models in Population Biology and Epidemiology*, Texts in Applied Mathematics, 40, Springer-Verlag New York, Inc., 2001.

[3] C. Meyer, *Matrix analysis and applied linear algebra*, SIAM, 2000.

[4] O. Perron, Zur Theorie der Matrices, *Mathematische Annalen* **64 (2)** (1907), 248–263.