

AQUINO VAZQUEZ, EULOGIO.

Las matrices positivas y no negativas
en el modelo de Insumo-Producto de Leon-
tief y en las cadenas de Markov.

MATEMATICO.

1975

A la bendita memoria de mi madre

T-1980



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A mi padre, hermanos, a Chacha.

A todos los asesores que cedieron parte de su tiempo para la conclusión de éste trabajo.

INDICE

0.- INTRODUCCION	II
1.- MATRICES POSITIVAS Y NO NEGATIVAS ; APLICACIONES	17
2.- MATRICES NO NEGATIVAS IRREDUCIBLES	31
3.- MATRICES NO NEGATIVAS REDUCIBLES	43
4.- APLICACIONES EN LAS CADENAS DE MARKOV.	60
APELLDICE A.1	68
APELLDICE A.2	71
CONCLUSIONES	72
BIBLIOGRAFIA	73

INTRODUCCION

El propósito de éste trabajo es proporcionar un estudio relativo a la teoría de las matrices positivas y no negativas de orden $n \times n$ y alguna de sus aplicaciones, esencialmente en los modelos Leontief y en las cadenas de Markov.

Se inicia dicho trabajo en un primer capítulo con un estudio de las matrices positivas, a partir de los radios espectrales hasta llegar al teorema fundamental para las matrices positivas, también conocido como el TEOREMA DE PERRON que es aplicado en las matrices no negativas que dan como resultado nuevos teoremas que tienen su aplicación en los modelos económicos.

Las aplicaciones en el modelo de insumo producto son modelos simples de cambio para una economía, a partir de ellos se definen los modelos Leontief cerrados y modelos Leontief abiertos en la cual el problema a ser resuelto al tratar con tales ecuaciones es determinar las condiciones que aseguren una solución para dichos modelos.

El segundo capítulo está dedicado a las matrices no negativas irreducibles que son definidas a partir de una matriz permutación. Acerca de la irreducibilidad de tales matrices la PROPOSICION 2.1 es importante para los siguientes desarrollos que concluyen con el teorema fundamental para matrices irreducibles no negativas también llamado el TEOREMA DE FROBENIUS y sus respectivos corolarios todos importantes en los desarrollos posteriores.

El tercer capítulo trata sobre las matrices no negativas reducibles, primitivas e imprimitivas, se destaca en éste capítulo la forma normal para las matrices reducibles y haciendo uso de ella se demuestra que al valor característico de cada matriz reducible le corresponde un vector característico positivo. En cuanto a las matrices primitivas e imprimitivas éstas se definen a partir de los valores característicos, especialmente y a manera de un teorema las matrices primitivas también se demuestran haciendo

uso del teorema de Perron y de la fórmula de interpolación de Lagrange-Silvester, dicha fórmula para no quitar continuidad al tema es justificada en el apéndice A.1 .

El cuarto capítulo está dedicado a las aplicaciones de las matrices irreducibles y reducibles en las cadenas de Markov, en la cual la matriz estocástica es una forma especial de una matriz no negativa por lo tanto todos los conceptos y proposiciones del segundo y tercer capítulos son aplicables. En este mismo capítulo es necesario hacer uso de los divisores elementales en las matrices estocásticas, dichos divisores elementales son posteriormente justificados en el apéndice A.2 . La aplicación de los divisores se efectúa en la existencia de la matriz de probabilidades de transición final, básica para las cadenas de Markov . Se definen cuando una cadena de Markov es regular y totalmente regular, irreducible, reducible, acíclica ó cíclica, importantes para la demostración de los teoremas Ergódico y casi-Ergódico en las cadenas de Markov. Posteriormente se hace uso de los índices de imprimitividad de una matriz para la demostración de la media de probabilidad de transición final y la media de las probabilidades absolutas que se resumen en un teorema .

Se concluye este trabajo con los apéndice A1 y A2 que son necesarios para evitar quitar continuidad a los capítulos.

1

LAS MATRICES POSITIVAS Y NO NEGATIVAS

Definición 1.1: Sea $A = \|a_{ij}\|$ una matriz cuadrada, si todos sus elementos a_{ij} son no negativos se denota como $A \geq 0$, si $a_{ij} \geq 0$ y algún $a_{ij} > 0$ se denotará $A > 0$ y si toda $a_{ij} > 0$ se denota $A \gg 0$ y se denominarán estrictamente positivas.

Así $A \gg B$ si $A - B \gg 0$, entonces se tiene que una matriz como una transformación también preserva el orden parcial definido por " \gg ".

Definición 1.2: Un vector \bar{x} será llamado positivo si todas sus componentes son positivas y no negativo si todas sus componentes son no negativas, entonces, si \bar{x} es positivo se denotará $\bar{x} > 0$ y $\bar{x} \geq 0$ en el segundo caso.

La relación $\bar{x} \geq \bar{y}$ es equivalente a decir que $\bar{x} - \bar{y} \geq 0$ y $\bar{x} > \bar{y}$ si $\bar{x} - \bar{y} > 0$

Definición 1.3: Un escalar λ es un valor característico de A si, para algún vector $\bar{x} \neq 0$, se tiene que $\bar{x}A = \lambda\bar{x}$ donde el vector \bar{x} es llamado el vector característico de A correspondiente al valor característico λ .

Sea $\Delta = \{ \lambda \mid \bar{x}A \geq \lambda\bar{x} \text{ para alguna } \bar{x} \geq 0 \}$

Siendo homogénea la relación $\bar{x}A \geq \lambda\bar{x}$ se puede suponer sin pérdida de generalidad que el vector no negativo \bar{x} correspondiente al valor λ está normalizado de tal manera que

$$\|\bar{x}\| = \sum_{i=1}^n x_i^2 = 1$$

y sea la norma de la matriz A definida como

$$\|A\| = \sum_{i,j=1}^n a_{ij}$$

PROPOSICION 1.1. - Sea $A \gg 0$ y

$$\Delta = \{ \lambda \mid \bar{x}A \geq \lambda \bar{x}, \bar{x} \geq 0 \}, \text{ entonces } \lambda_0 = \max \lambda \text{ con } \lambda \in \Delta$$

Demostración

Si $\lambda \bar{x} \leq A \bar{x}$ se tiene que

$$|\lambda| \|\bar{x}\| \leq \|\bar{x}\| \|A\| \quad \text{ó} \quad 0 \leq \lambda \leq \|A\|$$

esto demuestra que Δ es un conjunto acotado que es no vacío si A es positiva. Por otro lado supóngase que $\lambda_0 = \sup \lambda$, $\lambda \in \Delta$, sea $\{\lambda_i\}$ una sucesión de λ 's que pertenecen a Δ y que convergen a λ_0 también sea $\{\bar{x}^i\}$ el conjunto de vectores asociados de tal manera que

$$\lambda_i \bar{x}^i \leq \bar{x}^i A \quad i = 1, 2, \dots$$

ya que $\|\bar{x}^i\| = 1$ se puede elegir una subsucesión de las \bar{x}^i , sea $\{\bar{x}^j\}$ que converge a \bar{x}^0 un vector no negativo y no trivial. Ya que $\lambda_0 \bar{x}^0 \leq \bar{x}^0 A$, se sigue que $\lambda_0 \in \Delta$, lo que significa que el supremo es también un máximo.

TEOREMA 1.1. (PERRON) Si A es hermitiana y $A \gg 0$ y si λ_0 es el supremo de Δ , entonces (a) existe $\bar{x}^0 > 0$ tal que $\bar{x}^0 A = \lambda_0 \bar{x}^0$; (b) si $\lambda \neq \lambda_0$ es cualquier otro valor característico de A se sigue que $|\lambda| < \lambda_0$; (c) λ_0 es un valor característico de multiplicidad geométrica y multiplicidad algebraica de valor 1.

Demostración

a) sea $\bar{x}^0 \geq 0$, $\bar{x}^0 \neq 0$ tal que $\bar{x}^0 A \geq \lambda_0 \bar{x}^0$, ya que $A \gg 0$ se tiene que $\bar{x} A > 0$ para $\bar{x} \geq 0$, supóngase que $\bar{x}^0 = \bar{x} A > 0$ de aquí que $\bar{x}^0 A = \bar{y}^0 > 0$ y nuevamente $\bar{y}^0 A \geq \lambda_0 \bar{y}^0$, supóngase que $\bar{y}^0 A > \lambda_0 \bar{y}^0$, pero si esto sucede se puede encontrar una $\varepsilon > 0$ tal que $\bar{y}^0 A \geq (\lambda_0 + \varepsilon) \bar{y}^0$ pero al aceptar que $\lambda_0 = \sup \lambda$, $\lambda \in \Delta$ se está contradiciendo el significado de λ_0 , entonces $\bar{x}^0 A = \lambda_0 \bar{x}^0$

y necesariamente $\bar{x}^0 > 0$

b) sea $\lambda \neq \lambda_0$ un valor característico de A y sea \bar{z} su correspondiente vector característico no trivial tal que $\bar{z}A = \lambda\bar{z}$, sea $|\bar{z}|$ el vector cuyas componentes son $(|z_i|)$. Entonces, de la positividad de A se tiene que $|\bar{z}|A \geq |\lambda||\bar{z}|$ pero $|\bar{z}|A > 0$ y por (a) se tiene que $|\bar{z}|A = \lambda_0|\bar{z}|$ lo cual implica que $\lambda_0|\bar{z}| \geq |\lambda||\bar{z}|$ de esto último se tiene que $\lambda_0 \geq |\lambda|$ pero $\lambda \neq \lambda_0$ de aquí que $|\lambda| < \lambda_0$.

c) Sea S una transformación lineal tal que $S: E^n \rightarrow E^m$ se define al Ker de S como $\mathcal{N}(S) = \{\bar{x} \in E^n \mid S\bar{x} = 0\}$ el cual es un subespacio de E^n , simiilarmente se define al rango de S como $R(S) = \{S\bar{x} \mid \bar{x} \in E^n\}$, además se sabe que

$$\dim \mathcal{N}(S) + \dim R(S) = n$$

Sea $L(E^n)$ el conjunto de todas las transformaciones lineales de E^n en E^n y sea S una transformación singular de $L(E^n)$ y considérese a los subespacios $\mathcal{N}(S), \mathcal{N}(S^2), \dots$

Como $\mathcal{N}(S) = \{\bar{x} \mid S\bar{x} = 0\}$ entonces se tiene que

$\mathcal{N}(S^2) = \{\bar{x} \mid S^2\bar{x} = S(S\bar{x}) = S(0) = 0\}$ de aquí se concluye que

$\mathcal{N}(S) \subset \mathcal{N}(S^2) \subset \mathcal{N}(S^3) \subset \dots$; pero por otro lado, subespacios propios tienen cada vez dimensiones más pequeñas, entonces existe $k \in \mathbb{Z}$, $k \leq n$ para lo cual

$$\mathcal{N}(S) \subsetneq \mathcal{N}(S^2) \subsetneq \mathcal{N}(S^3) \subsetneq \dots \subsetneq \mathcal{N}(S^k) = \mathcal{N}(S^{k+1}) = \dots$$

pero como $\dim \mathcal{N}(S) + \dim R(S) = n$, relacionando esto último con lo anterior se tiene que

$$R(S) \supsetneq R(S^2) \supsetneq R(S^3) \supsetneq \dots \supsetneq R(S^k) = R(S^{k+1}) = \dots$$

Sea $(A - \lambda I)$ una transformación singular definida como

$$N_\lambda^r = \{\bar{x} \mid \bar{x}(A - \lambda I)^r = 0\} \quad \text{y} \quad M_\lambda^r = \{\bar{x} = \bar{y}(A - \lambda I)^r \mid \bar{y} \text{ arbitrario}\}$$

aplicando todo lo anterior se tiene que

$$N_\lambda^0 \subsetneq N_\lambda^1 \subsetneq N_\lambda^2 \subsetneq \dots \subsetneq N_\lambda^k = N_\lambda^{k+1} = \dots$$

para alguna matriz A y $\lambda \in \Lambda$.

El entero κ así determinado es llamado el índice de λ , la dimensión de N'_λ es llamada la multiplicidad geométrica de λ y la dimensión de N_λ^κ es llamada la multiplicidad algebraica de λ .

Pero si $N_\lambda^0 \subsetneq N'_\lambda \subsetneq N_\lambda^2 \subsetneq \dots \subsetneq N_\lambda^\kappa = N_\lambda^{\kappa+1} = \dots$

entonces $M_\lambda^0 \supsetneq M'_\lambda \supsetneq M_\lambda^2 \supsetneq \dots \supsetneq M_\lambda^\kappa = M_\lambda^{\kappa+1} = \dots$ para alguna $\kappa \in \mathbb{Z}$ y si A es hermitiana entonces la multiplicidad algebraica es igual a la multiplicidad geométrica y cuando esto suceda para $\kappa=1$ se dice que los divisores elementales son simples.

Supóngase que el conjunto de vectores característicos con valor característico λ_0 es al menos de dos dimensiones, ya que A es una matriz real existe un vector característico \bar{x}^0 que es real y además linealmente independiente de \bar{x}^0 ya que $\bar{x}^0 > 0$ existe una combinación lineal $\bar{w} = t\bar{x}^0 + \bar{z}^0$ donde $t = \min_i (-\frac{\bar{z}^0_i}{\bar{x}^0_i})$ con $\bar{w} \geq 0$ pero no $\bar{w} > 0$, ya que $\bar{w}A > 0$ y $\lambda_0\bar{w} \geq 0$ se sigue que $\bar{w}A > \lambda_0\bar{w}$ y no idénticamente cero, entonces esto último contradice la definición de λ_0 y como la multiplicidad geométrica es igual al número de vectores característicos linealmente independientes de lo anterior y de esto último se sigue que solamente existe un vector correspondiente a λ_0 y que es \bar{x}^0 de aquí que la multiplicidad geométrica de λ_0 es 1.

Por otro lado supóngase que $\bar{y}(A - \lambda_0 I) = \bar{z}$ y que $\bar{z}(A - \lambda_0 I) = 0$ por la primera parte de este argumento, \bar{z} es un múltiplo de \bar{x}^0 , es decir $\bar{z} = a\bar{x}^0$ con $a \geq 0$

Aplicando el inciso (a) existe un vector característico positivo \bar{f}^0 para la matriz estrictamente positiva A' , donde \bar{f}^0 corresponde al mismo valor característico λ_0 , tal que

$$(a\bar{x}^0, \bar{f}^0) = (\bar{y}(A - \lambda_0 I), \bar{f}^0) = (\bar{y}, (A' - \lambda_0 I)\bar{f}^0) = 0$$

ya que $(\bar{x}^0, \bar{f}^0) > 0$ debe tenerse que $a=0$, entonces si $\bar{y}(A - \lambda_0 I)^2 = 0$ necesariamente $\bar{y}(A - \lambda_0 I) = 0$ y así se tiene que $N_{\lambda_0}^1 = N_{\lambda_0}^2$.

Esto significa que la multiplicidad algebraica de λ_0 es igual a su multiplicidad geométrica.

Con esto se concluye la demostración.

Para estudiar con mayor detalle la estructura de la variedad propia asignada a λ_0 , se introduce una operación proyección, que proyecta vectores en el espacio lineal de vectores característicos asociados a λ_0 .

Para este propósito sea \bar{x}^0 el vector característico tal que $\bar{x}^0 A = \lambda_0 \bar{x}^0$ y el vector característico $\bar{f}^0 > 0$ donde $A \bar{f}^0 = \lambda_0 \bar{f}^0$, normalizándose ambos en la forma $\sum x_i^0 f_i^0 = 1$.

PROPOSICION 1.2 sea $P = \|x_j^0 f_i^0\|$ la matriz con las siguientes propiedades i) Para cualesquiera vectores \bar{x} y \bar{f} se tiene que $\bar{x} P = (\bar{x}, \bar{f}^0) \bar{x}^0$, $P \bar{f} = (\bar{f}, \bar{x}^0) \bar{f}^0$. ii) $P^2 = P$, iii) $AP = PA = \lambda_0 P$.

Demostración

$$i) \bar{x} P = \bar{x} \|x_j^0 f_i^0\| = (x_1, x_2, \dots, x_n) \begin{bmatrix} x_1^0 f_1^0 & x_2^0 f_1^0 & \dots & x_n^0 f_1^0 \\ x_1^0 f_2^0 & x_2^0 f_2^0 & \dots & x_n^0 f_2^0 \\ \dots & \dots & \dots & \dots \\ x_1^0 f_n^0 & x_2^0 f_n^0 & \dots & x_n^0 f_n^0 \end{bmatrix} =$$

$$= (x_1 \sum x_i^0 f_i^0; x_2 \sum x_i^0 f_i^0; \dots; x_n \sum x_i^0 f_i^0) = \sum_{i=1}^n x_i f_i^0 (x_1^0, x_2^0, x_3^0, \dots, x_n^0) = (\bar{x}, \bar{f}^0) \bar{x}^0$$

análogamente se tiene que

$$P \bar{f} = \|x_j^0 f_i^0\| \begin{bmatrix} f_1^0 \\ f_2^0 \\ \dots \\ f_n^0 \end{bmatrix} = \begin{bmatrix} x_1^0 f_1^0 & x_2^0 f_1^0 & \dots & x_n^0 f_1^0 \\ x_1^0 f_2^0 & x_2^0 f_2^0 & \dots & x_n^0 f_2^0 \\ \dots & \dots & \dots & \dots \\ x_1^0 f_n^0 & x_2^0 f_n^0 & \dots & x_n^0 f_n^0 \end{bmatrix} \begin{bmatrix} f_1^0 \\ f_2^0 \\ \dots \\ f_n^0 \end{bmatrix} =$$

$$= \sum_{i=1}^n f_i x_i^0 \begin{bmatrix} f_1^0 \\ f_2^0 \\ \dots \\ f_n^0 \end{bmatrix} = (\bar{f}, \bar{x}^0) \bar{f}^0$$

$$ii) P^2 = PP = \|x_j^{\circ} f_i^{\circ}\| \|x_i^{\circ} f_k^{\circ}\| = \sum x_j^{\circ} f_i^{\circ} x_i^{\circ} f_k^{\circ} = x_j^{\circ} (\sum f_i^{\circ} x_i^{\circ}) f_k^{\circ} = x_j^{\circ} f_k^{\circ} = P$$

$$iii) AP = A(\bar{x}^{\circ}, \bar{f}^{\circ}) = (A\bar{x}^{\circ}, \bar{f}^{\circ}) = (\lambda_0 \bar{x}^{\circ}, \bar{f}^{\circ}) = \lambda_0 (\bar{x}^{\circ}, \bar{f}^{\circ}) = \lambda_0 P$$

$$PA = (\bar{x}^{\circ}, \bar{f}^{\circ})A = (\bar{x}^{\circ}, A\bar{f}^{\circ}) = (\bar{x}^{\circ}, \lambda_0 \bar{f}^{\circ}) = \lambda_0 (\bar{x}^{\circ}, \bar{f}^{\circ}) = \lambda_0 P$$

Para cualquier matriz B se tiene que $(B - \lambda I)^{-1}$ representa una función analítica de λ , es decir cada componente es una función analítica de λ siempre que la inversa tenga sentido. Cada valor característico de B da lugar a un polo de $(B - \lambda I)^{-1}$ es decir un polo para alguna componente, cuyo orden está relacionado a la multiplicidad algebraica del valor característico, además:

$$(B - \lambda I)^{-1} = \lambda^{-1} \left(\frac{1}{\lambda} B - I \right)^{-1} = - \sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} B^n$$

que converge para $|\lambda| > r$ donde r denota el radio del círculo más pequeño que contiene a los valores característicos de B (también escrito como $\lambda_0(A) = r_0(A) = r$); pero por lo anterior $(B - \lambda I)^{-1}$ tiene un polo en algún valor característico λ^* donde $|\lambda^*| = \lambda_0$ y aplicando el teorema de Cauchy-Hadamard se tiene que $-\sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} B^n$ converge, si $|\lambda| > r$ donde:

$$r = \lim_{n \rightarrow \infty} \sqrt[n]{\max_{ij} |b_{ij}^{(n)}|} \quad \text{con } B^{(n)} = \|b_{ij}^{(n)}\|$$

TEOREMA 1.2: Para $A \gg 0$, $\lambda_0 = \sup \lambda$, $P = \|x_j^{\circ} f_i^{\circ}\|$ se tiene que:

$$\frac{1}{\lambda_0^n} A^n \rightarrow P$$

Demostración:

Los valores característicos de la matriz $B = A - \lambda_0 P$ deben estar estrictamente dentro del círculo de radio λ_0 ya que:

i) Cualquier valor característico distinto de cero de B es un valor característico de A .

ii) λ_0 no es un valor característico de B

Supóngase que lo anterior no fuese cierto, entonces

$$\bar{z}B = \lambda \bar{z} \text{ para algún } \bar{z} \neq 0, \text{ se sigue que: } \bar{z}BP = \lambda \bar{z}P \text{ y}$$

$$\bar{z}BP = \bar{z}(AP - \lambda_0 P^2) = \bar{z}(\lambda_0 P - \lambda_0 P) = \bar{z}\lambda_0(P - P) = 0$$

por lo tanto, si $\lambda \neq 0$, se sigue $\bar{z}BP = \lambda \bar{z}P = 0$, y se tiene que $\bar{z}P = 0$, por lo tanto si $\bar{z}B = \bar{z}(A - \lambda_0 P) = \lambda \bar{z}$ se sigue que $\bar{z}A = \lambda \bar{z}$ ya que $\bar{z}P = 0$.

En particular si $\bar{z}B = \lambda_0 \bar{z}$, entonces $\bar{z}P = 0$ y $\bar{z}A = \lambda_0 \bar{z}$; por otro lado ya que $\bar{x}^0 A = \lambda_0 \bar{x}^0$, se tiene que $\bar{z}A = \lambda_0 \bar{z}$ solo sucederá si \bar{z} es múltiplo de \bar{x}^0 , pero entonces $\bar{x}^0 P$ no puede ser cero, ésta contradicción demuestra que las afirmaciones son ciertas.

Así el radio espectral de B debe ser menor que λ_0 , y se sigue que la máxima componente de B^m satisface la cota $|b_{ij}^{(m)}| < K e^{\rho m}$ para toda m con una adecuada $\rho < \lambda_0$.

Pero $B^m = A^m - \lambda_0^m P$, de aquí se infiere que A^m / λ_0^m converge a P , con la razón de convergencia ρ / λ_0 .

Varios de los resultados enunciados anteriormente son válidos cuando $A > 0$, para ello considérese una matriz cuadrada U con componentes \perp , entonces se tiene que $A + \delta U \gg 0$ para $\delta > 0$ y un argumento de límite sobre δ establece la validez de las partes (a) y (b) del siguiente teorema:

TEOREMA 1.3. - Sea $A > 0$, $\lambda_0 = \sup \lambda$, entonces, (a) λ_0 es un valor característico y existe $\bar{x}^0 \geq 0$ tal que $\bar{x}^0 A = \lambda_0 \bar{x}^0$; (b) si λ es cualquier valor característico se tiene que $|\lambda| < \lambda_0$; (c)

$\frac{1}{m} \sum_{i=1}^m \frac{A_i}{\lambda_0}$ converge siempre que $\bar{x}^0 > 0$ y (d) si λ es un valor característico de A con $|\lambda| = \lambda_0$ se sigue que λ/λ_0 es una raíz de la unidad y $\eta^m \lambda_0$ es un valor característico de A para $m = 0, 1, 2, \dots$

Demostración:

a) - Ya que $A + \delta U \gg 0$, se tiene que $\bar{x}^0 (A + \delta U) > 0$ para $\bar{x} \geq 0$, de aquí que $\bar{x}^0 (A + \delta U) = \lambda_0 \bar{x}^0$ para $\delta > 0$ y tan pequeño como se quiera, entonces se sigue que:

$$\bar{x}^0 A = \lambda_0 \bar{x}^0 \text{ para } A > 0 \text{ y } \bar{x}^0 \geq 0$$

b) es válido por inciso (a) y teorema 1.1

c) - Por conveniencia supóngase que $\lambda_0 = 1$; obsérvese que los elementos de A^m están uniformemente acotados, en efecto; $\bar{x}^0 A^m = \bar{x}^0$ implica que:

$$\left[\sum_i^{(m)} a_{ij} x_i^0 \right] = [x_i^0] \text{ entonces } a_{ij}^{(m)} \leq x_i^0 \text{ y en particular}$$

$$\text{se tiene que: } 0 \leq a_{ij}^{(m)} \leq \frac{\max_i x_i^0}{\min_i x_i^0}$$

Si \mathcal{L} denota el espacio de elementos para los cuales $\bar{z} A = \bar{z}$, \mathcal{L} es un espacio lineal cerrado tal que para $\bar{z} \in \mathcal{L}$ se tiene que:

$$\bar{z} \frac{A + A^2 + \dots + A^m}{m} = \frac{\bar{z} A + \bar{z} A^2 + \dots + \bar{z} A^m}{m} = \frac{m \bar{z}}{m} = \bar{z}$$

y converge cuando $m \rightarrow \infty$.

Se define a $K = \{ \bar{y} \mid \bar{y} = \bar{z}(I-A) \}$, es decir K es el rango de $(I-A)$; sea:

$$S_m = \frac{A+A^2+\dots+A^m}{m} \quad \text{y ya que } \bar{y} = \bar{z}(I-A) \text{ se sigue que}$$

$$\begin{aligned} \bar{y} S_m &= \bar{y} \frac{A+A^2+\dots+A^m}{m} = \bar{z}(I-A) \frac{A+A^2+\dots+A^m}{m} = \\ &= \frac{\bar{z}A - \bar{z}A^2 + \bar{z}A^2 - \dots - \bar{z}A^m + \bar{z}A^m}{m} = \frac{\bar{z}A - \bar{z}A^{m+1}}{m} \end{aligned}$$

en consecuencia $\bar{y} S_m = \frac{\bar{z}A - \bar{z}A^{m+1}}{m}$ que converge a cero

cuando m tiende a infinito en virtud del hecho de que los coeficientes de A^{m+1} están uniformemente acotados.

Así se establece que S_m converge para todos los vectores contenidos en la suma de $L + K$.

Queda únicamente por demostrar que $L + K$ es todo el espacio n -euclideo, para esto obsérvese que cualquier vector \bar{x} puede ser escrito como:

$$\bar{x} = (\bar{x} - \bar{x} S_m) + \bar{x} S_m = \bar{y}_m + \bar{z}_m$$

Ya que S_m consiste de una sucesión de matrices uniformemente acotadas, se tiene que cuando m tiende a ∞ , \bar{z}_m converge a \bar{z}_0 , donde $\bar{z}_0 \in L$, por ésto mismo $\bar{y}_m \in K$ y si \bar{y}_m converge a \bar{y}_0 , entonces $\bar{y}_0 \in K$.

d).- Supóngase sin pérdida de generalidad que existe algún vector $\bar{f}^0 > 0$ para la cual $A\bar{f}^0 = \bar{f}^0$, (\bar{f}^0 existe por inciso (a) de éste

mismo teorema aplicado a la matriz A' , con ayuda de esto se demuestra(d), porque, supóngase que λ es un valor característico de A de módulo 1 y que $\bar{x}A = \lambda\bar{x}$ para alguna \bar{x} , es decir:

$$\sum_{i=1}^n x_i a_{ij} = \lambda x_j, \text{ entonces se tiene que}$$

$$\sum_{i=1}^n |x_i| a_{ij} \geq |x_j| \text{ ó } |\bar{x}|A \geq |\bar{x}| \text{ con } |\bar{x}| = (|x_1|, |x_2|, \dots, |x_n|)$$

pero $(\bar{f}^0, |\bar{x}|) \leq (\bar{f}^0, |\bar{x}|A) = (A\bar{f}^0, |\bar{x}|) = (\bar{f}^0, |\bar{x}|)$ y esto es una contradicción a menos que $|\bar{x}|A = |\bar{x}|$, de aquí se tiene que

$$\left| \sum_{i=1}^n x_i a_{ij} \right| = |x_j| = \sum_{i=1}^n |x_i| a_{ij}$$

esto implica la existencia de números u_j con $|u_j| = 1$ tal que

$$x_i a_{ij} = |x_i| a_{ij} u_j \text{ para cada } i \text{ y } j \quad (1.1)$$

multiplicando la relación precedente por u_j y sumándola se tiene que

$$(\bar{x}u)A = (u \cdot |\bar{x}|)Au$$

y de (1.1) tomando los sumandos sobre i se tiene que

$$|\bar{x}|u = (|\bar{x}|A) \cdot u = \bar{x}A = \lambda\bar{x}$$

usando las últimas relaciones se prueba que

$$(\bar{x} \cdot u)A = (u \cdot |\bar{x}|)A \cdot u = \lambda\bar{x}Au = \lambda^2\bar{x} \cdot u$$

entonces $\bar{x} \cdot u$ es un vector característico que corresponde a λ^2 , además:

$$u_i^2 |x_i| a_{ij} u_j = u_i^2 x_i a_{ij}$$

se sigue que $|\bar{x}| \cdot u^2 = |\bar{x}| u \cdot u = \lambda \bar{x} \cdot u$

pero $(|\bar{x}| \cdot u^2) A = \lambda (\bar{x} \cdot u) A = \lambda (\lambda^2 \bar{x} \cdot u) = \lambda^3 u \cdot \bar{x}$

y así $(|\bar{x}| u^2) A u = \lambda^3 u^2 \cdot \bar{x}$

esto demuestra que $\bar{x} \cdot u^2$ es un vector característico de λ^3 .

Procediendo de esta manera, se encuentra que λ es un valor característico de módulo 1, entonces λ es una raíz de la unidad y, $\lambda, \lambda^2, \lambda^3, \dots, \lambda^n, \dots$ son todos valores característicos como se afirmó de esto se sigue que un vector característico de λ^{r+1} es $\bar{x} \cdot u^r$, en la cual \bar{x} y u tienen el mismo significado que anteriormente se indicó, y esto completa la demostración.

El radio espectral $\lambda_0 = \lambda_0(A)$ para la matriz A que fué usado en los teoremas precedentes, es real y no negativo donde

$$\lambda_0 = \max_{\lambda} \lambda \quad \text{y} \quad \Delta = \{ \lambda \mid \bar{x} A \geq \lambda \bar{x}, \bar{x} \geq 0 \}$$

Como consecuencia se tiene:

COROLARIO 1.1 .- Si existe $\bar{x}^0 \gg 0$ tal que $\bar{x}^0 A \leq \mu \bar{x}^0$ entonces μ es una cota superior para el radio espectral de A .

Demostración

$$|a_{ij}^{(n)}| \leq \mu^n \frac{\max_i x_i^0}{\min_i x_i^0} \quad \text{donde } A^n = \left\| a_{ij}^{(n)} \right\|$$

de aquí se tiene que:

$$\sqrt[n]{\max_{i,j} |a_{ij}^{(n)}|} \leq \mu \quad \text{es decir } \lambda_0 \leq \mu$$

TEOREMA 1.4: Si $\rho > \lambda_0(A)$ y $A \geq 0$, entonces $(\rho I - A)^{-1}$ transforma vectores no negativos en vectores no negativos.

Demostración

Para $\lambda > \lambda_0(A)$ donde $\lambda_0(A)$ es el radio espectral se tiene que

$$(\lambda I - A)^{-1} = \lambda^{-1} (I - \frac{1}{\lambda} A)^{-1} = \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}}$$

y ya que cada A^k transforma un ortante no negativo en sí mismo la suma infinita de transformaciones también lo realiza.

El inverso es también válido, es decir, si $(\rho I - A)^{-1}$ transforma el ortante no negativo en sí mismo, entonces ρ es una cota superior para el radio espectral de A ; es decir, si $\bar{x} \geq 0$, $\bar{x} \neq 0$ y con más componentes cero y si $\bar{x} A \geq \lambda \bar{x}$, entonces:

$\bar{z} = \bar{x} (A - \rho I) + (\rho - \lambda) \bar{x} \geq 0$, entonces $(\rho I - A)^{-1}$ aplicado a \bar{z} produce

$$\begin{aligned} (\rho I - A)^{-1} \bar{z} &= (\rho I - A)^{-1} \bar{x} (A - \rho I) + (\rho - \lambda) \bar{x} (\rho I - A)^{-1} \\ &= -\bar{x} + (\rho - \lambda) \bar{x} (\rho I - A)^{-1} \geq 0 \end{aligned}$$

que es un vector no negativo porque

$$(\rho I - A)^{-1} = \sum_{k=1}^{\infty} \frac{A^k}{\rho^{k+1}} \quad \text{donde } A^k \geq 0 \text{ y como } \bar{z} \geq 0$$

entonces se tiene que $(\rho I - A)^{-1} \geq 0$

Pero como $\bar{x} \geq 0$ y $\bar{x} (\rho I - A)^{-1} \geq 0$ necesariamente se tiene que $\rho > \lambda$, y en particular $\rho > \lambda_0(A)$.

APLICACIONES EN EL MODELO DE INSUMO PRODUCTO DE LEONTIEF

Uno de los modelos más simples de cambio para una economía puede ser descrito como sigue

Sean n actividades (o industrias) cada una produciendo un producto. Supóngase que χ_i represente el producto total de la actividad i y que x_{ij} represente la cantidad total del producto de la actividad i usado por la j -ésima actividad, si

$$y_i = \chi_i - \sum_{j=1}^n x_{ij}$$

entonces y_i expresa la diferencia entre la cantidad total de actividad i y la cantidad de su producto consumida por las n actividades, y_i es comunmente llamada la demanda final de comodidad i .

Si $y_i = 0$ para toda i , no existen sobrantes, si esto sucede entonces $y_i = 0$ es llamado un modelo Leontief cerrado. Si el trabajo es un tipo de actividad, el trabajo total es empleado, si el acero es otro tipo de actividad, todo el acero producido es consumido por el sistema. Un modelo cerrado incluye actividades como servicios, administración etc.

Si $y_i > 0$ para al menos una i , este modelo es conocido como un modelo Leontief abierto, desde un punto de vista económico un modelo abierto implica la presencia de material exógeno, alguna cantidad de capital, trabajo y material en bruto etc. que es suministrado desde el exterior. La característica esencial de los modelos abiertos es la existencia de demanda exterior.

Se define

$$a_{ij} = \frac{x_{ij}}{\chi_j} \quad (1.2)$$

que puede ser interpretada como la cantidad de actividades i 's, como-
didad necesaria para producir una unidad de comodidad j . Las can-
tidades a_{ij} son llamados los "coeficientes de producción" y son su-
puestos a ser constantes independientes de χ_i .

∴ De (1.2) se sigue que

$$a_{ij} = \frac{x_{ij}}{x_j}, \quad x_{ij} = a_{ij} x_j, \quad \text{entonces } y = (I-A)x \quad (1.3)$$

donde $A = \|a_{ij}\|$, $x = (x_1, x_2, \dots, x_n)$ y $y = (y_1, y_2, \dots, y_n)$

En la práctica A debe ser estimada o conocida.

El problema más frecuente que concierne al modelo Leontief abierto es como sigue, dado que las y_i son conocidas, se debe encontrar x_i que satisfaga (1.3). Por supuesto la contestación es significativa si la solución $x = (I-A)^{-1}y$ existe y es un vector no negativo.

Supóngase que

$$(I-A)^{-1} = \|A_{ij}\|$$

entonces

$$x_i = \sum_j A_{ij} y_j \quad i=1, 2, \dots, n$$

de tal manera que puede ser interpretado como la cantidad por la cual el producto de la actividad i debe ser incrementada para producir una unidad adicional de comodidad j .

Otro problema que envuelven los modelos abiertos es el siguiente: consideréense n países, supóngase que a_{ij} denoten los países propensos a importar desde el país i . Sean x_i los países con ingreso nacional y sea C_i los países i 's con gasto nacional, que es independiente del ingreso, entonces

$$x_i = \sum_{j=1}^n a_{ij} x_j + C_i \quad \text{ó} \quad C = (I-A)x$$

donde C y x deben ser vectores positivos.

Este ejemplo puede ser interpretado como un modelo cerrado de tratado entre países. Supóngase que todos los ingresos para j países venga de la venta a otros países o así mismo; sea α_{ij} la fracción de países cuyo ingreso es gastado en mercancías importadas desde el país i , sea x_j el ingreso anual del país j , entonces el valor total de la exportación del país i es

$$\sum \alpha_{ij} x_j$$

Y esto, por definición de un modelo cerrado es χ_i , es decir

$$\chi_i = \sum_{j=1}^n \alpha_{ij} \chi_j \quad i = 1, 2, \dots, n$$

donde

$$\sum_{j=1}^n \alpha_{ij} = 1 \quad , \quad \alpha_{ij} \geq 0$$

El problema a ser resuelto al tratar con tales ecuaciones de este tipo, es determinar las condiciones que aseguran una solución positiva.

En el caso de un modelo cerrado se busca una solución positiva χ del sistema de ecuaciones $A\chi = \chi$. Soluciones no triviales existen si y solo si el valor 1 es un valor característico de la matriz A cuando esto sucede se asegura la existencia de un vector característico positivo.

Asegurar la existencia de soluciones positivas para $A\chi = \chi$ se requiere de algunas condiciones adicionales. Una consideración razonable sugerida por consideraciones económicas es que

$$\sum_i \chi_{ij} = \chi_j$$

es decir, que el producto total de la j -ésima actividad sea completamente usada en adquirir totalmente bienes de insumo. Esto puede ser pensado como parte de un sistema económico cerrado, por ejemplo, en una economía que consista solamente granjeros, tejedores, y carpinteros, la suposición dice que el granjero consume todo su ingreso en ropa, casa y alimentos.

El problema es determinar la cantidad a ser producida, consistente con los coeficientes de producción. En términos de los coeficientes, la suposición deja a la relación

$$\sum_{i=1}^n \alpha_{ij} = 1$$

para cada j .

Como consecuencia la matriz A debe tener radio espectral 1, pero si $\bar{u} = (1, 1, \dots, 1)$ con n componentes, entonces $\bar{u}A = \bar{u}$

y si $A\bar{z} = \lambda\bar{z}$ con $\sum |z_i| = 1$, entonces

$$(\bar{0}, A|\bar{z}|) \geq |\lambda| \sum |z_i|$$

nuevamente $|\bar{z}|$ representa al vector con componentes $|z_i|$, pero

$$(\bar{0}, |\bar{z}|) = (\bar{u}, |\bar{z}|) = \sum |z_i|$$

se sigue que $|\lambda| \leq 1$ y ya que los valores característicos de A y A' coinciden $\lambda_0 = 1$, se sigue por Teorema 1.3 que existe un vector $\bar{x} \geq 0$, $\bar{x} \neq 0$ que resuelve a $\bar{x}A = \bar{x}$, con esto se ha probado el siguiente

TEOREMA 1.5: En un sistema Leontief cerrado con

$$\sum_{i=1}^n x_{ij} = \chi_j$$

existe una solución no trivial positiva para las ecuaciones $A\bar{x} = \bar{x}$, si A tiene un vector estrictamente positivo, entonces la solución es única excepto para un factor de multiplicación; y posee solamente componentes estrictamente positivas.

La última parte del teorema se infiere del Teorema 1.1

En el caso de un sistema abierto el problema es determinar condiciones tales que $(I-A)\bar{x} = \bar{y}$ puedan ser resueltas para cualquier vector demanda final \bar{y} tal que la solución \bar{x} es un vector no negativo.

TEOREMA 1.6: Si existe algún vector producto \bar{x}^0 tal que $(I-A)\bar{x}^0$ es un vector estrictamente positivo, entonces $(I-A)^{-1}$ existe y transforma vectores no negativos en vectores no negativos.

(La expresión $(I-A)^{-1}$ puede ser vista como una transformación lineal que actúa ya sea sobre vectores columna o vectores renglón)

Demostración

Por Teorema 1.4, es suficiente demostrar que todos los valores característicos de A son de magnitud menor que 1. Sea λ tal que $\bar{z}A = \lambda\bar{z}$, $\bar{z} \neq 0$, la positividad de A implica que $|\bar{z}|A \geq |\lambda||\bar{z}|$.

La hipótesis requiere que

$$((I-A)\bar{x}_0, |\bar{z}|) = (\bar{x}_0, |\bar{z}|) - (\bar{x}_0, |\bar{z}|A) = (\bar{x}_0, |\bar{z}|) - (\bar{x}_0, |\lambda||\bar{z}|) > 0$$

Y ya que $\bar{x}_0 \geq 0$ se sigue que

$$(\bar{x}_0, |\bar{z}|)(1-|\lambda|) > 0$$

Se sigue que

$$|\lambda| < 1$$

Nota. - La hipótesis de este teorema significa que existe alguna combinación de productos que garantizan un excedente para cada comodidad.

Un corolario de esta clase, cuya demostración es un paráfrasis del teorema previo, es el siguiente

COROLARIO 1.2: Si existe algún vector precio \bar{p}^0 tal que $\bar{p}^0(I-A)$ es estrictamente positivo, entonces $(I-A)^{-1}$ transforma vectores no negativos en vectores no negativos.

Definición 2.1: Se dice que P es una matriz permutación si P es una matriz cuadrada tal que en cada fila y cada columna tiene un elemento singular de valor 1 y todos los otros elementos son cero.

Definición 2.2: Una matriz cuadrada A es reducible si existe una matriz permutación P tal que

$$PAP' = \begin{bmatrix} B & O \\ C & D \end{bmatrix}. \quad (2.1)$$

donde B y D son matrices cuadradas de ordenes $r \times r$ y $(n-r) \times (n-r)$ sino existe tal matriz permutación P que haga posible (2.1), entonces la matriz A será llamada irreducible.

PROPOSICION 2.1: Si $A \geq 0$ es una matriz irreducible de orden $n \times n$, entonces

$$(I+A)^{n-1} \gg 0$$

Demostración

Es suficiente demostrar que para cualquier vector $\bar{x} \geq 0$ $\bar{x} \neq 0$, $(I+A)^{n-1} \bar{x} > 0$, definiendo la sucesión de vectores no negativos $\bar{x}_{k+1} = (I+A)\bar{x}_k$ $0 \leq k \leq n-2$, donde $\bar{x}^0 = \bar{x}$, la prueba consiste en demostrar que \bar{x}_{k+1} tiene menos componentes cero que \bar{x}_k .

Ya que $\bar{x}_{k+1} = \bar{x}_k + A\bar{x}_k$, se puede suponer que \bar{x}_{k+1} y \bar{x}_k tienen el mismo número de componentes cero, entonces para una matriz permutación aceptable P , se puede escribir

$$P\bar{x}_{k+1} = \begin{pmatrix} \bar{\alpha} \\ 0 \end{pmatrix}, \quad P\bar{x}_k = \begin{pmatrix} \bar{\beta} \\ 0 \end{pmatrix}, \quad \bar{\alpha} > 0, \quad \bar{\beta} > 0$$

donde ambos $\bar{\alpha}$, $\bar{\beta}$ tienen m componentes para $1 \leq m < n$, haciendo

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \text{se tiene que}$$

$$\begin{bmatrix} \bar{x} \\ 0 \end{bmatrix} = \begin{bmatrix} \bar{\beta} \\ 0 \end{bmatrix} + \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \bar{\beta} \\ 0 \end{bmatrix}$$

se sigue que $A_{21} \bar{\beta} = 0$ pero $A_{21} \geq 0$ y $\bar{\beta} > 0$, esto solo puede ocurrir si $A_{21} = 0$ que contradice la hipótesis de que A sea irreducible. Así \bar{x}_{k+1} tiene menos componentes cero que \bar{x}_k y ya que \bar{x}^0 tiene a lo más $(n-1)$ componentes cero, entonces \bar{x}_k tiene a lo más $(n-k-1)$ componentes cero, así

$$\bar{x}_{n-1} = \bar{x}^0 (I + A)^{n-1} > 0$$

Supóngase que $\bar{x} = (x_1, x_2, \dots, x_n)$, $\bar{x} \neq 0$ sea un vector real fijo, y sea

$$\Gamma_{\bar{x}} = \min_{1 \leq i \leq n} \frac{(\bar{x}A)_i}{x_i}, \quad x_i \neq 0; \quad (\bar{x}A)_i = \sum_{k=1}^n x_k a_{ik}$$

de esto se tiene que $\Gamma_{\bar{x}} \geq 0$ y $\Gamma_{\bar{x}}$ es un número real que es el supremo de todos los números ρ para lo cual $\rho \bar{x} \leq \bar{x}A$

Sea

$$r = \bar{r} = \max_{\bar{x} \geq 0} \Gamma_{\bar{x}} = \max_{\bar{x} \geq 0} \min_{0 < x_i} \frac{(\bar{x}A)_i}{x_i}; \quad \bar{r} \geq 0 \quad (2.2)$$

de la definición de $\Gamma_{\bar{x}}$ se sigue que $\Gamma_{\bar{x}}$ y $\Gamma_{\lambda \bar{x}}$ tienen el mismo valor para algún escalar $\lambda > 0$; sea M el conjunto cerrado de vectores $\bar{x} \geq 0$ tales que

$$(\bar{x}, \bar{x}) = \sum_{i=1}^n x_i^2 = 1$$

si la función $\Gamma_{\bar{x}}$ fuese continua sobre M , la existencia del máximo podría ser garantizada, sin embargo, aún cuando fuese continua en cada $\bar{x} > 0$, $\Gamma_{\bar{x}}$ podría tener discontinuidades en los puntos frontera de M , en la que una de sus coordenadas desaparece, por lo tanto

se introduce en lugar de M un conjunto N de vectores \bar{y} de la forma

$$\bar{y} = \bar{x}(I+A)^{n-1}, \quad \bar{x} \in M$$

este conjunto N es cerrado y acotado debido a que M es cerrado y por la proposición 2.1, entonces N consiste de vectores positivos, además si se multiplica a ambos lados de la desigualdad $r_{\bar{x}} \bar{x} \leq \bar{x}A$ por $(I+A)^{n-1} > 0$ se obtiene

$$r_{\bar{x}} \bar{y} \leq \bar{y}A \quad \text{con} \quad \bar{y} = \bar{x}(I+A)^{n-1}$$

Sea $r_{\bar{y}} = \min \frac{(\bar{y}A)_i}{y_i}$ de esto último se tiene que

$$r_{\bar{x}} \leq r_{\bar{y}}$$

sobre el conjunto compacto N la función $r_{\bar{x}}$ es continua y toma su máximo en algún vector $\bar{z} > 0$, si r toma su máximo en algún vector $\bar{z} > 0$, es decir $r = r_{\bar{z}}$, entonces \bar{z} será llamado vector extremo.

PROPOSICION 2.2: El número $r = r_{\bar{z}}$ es positivo y es un valor característico de A y cada vector extremo \bar{z} es positivo y es un vector característico de A para el valor característico r , es decir

$$\bar{z}A = r\bar{z}, \quad r > 0, \quad \bar{z} > 0 \quad (2.3)$$

Demostración

Si $\bar{u} = (1, 1, \dots, 1)$, entonces $r_{\bar{u}} = \min \sum_{k=1}^n a_{ik}$, de aquí se tiene que $r_{\bar{u}} > 0$ porque ninguna fila de una matriz irreducible puede consistir solamente de ceros, entonces $r > 0$, ya que $r \geq r_{\bar{u}}$.

Sea $\bar{x} = \bar{z}(I+A)^{n-1} > 0$ y supóngase que $A\bar{z} - r\bar{z} \neq 0$ por Proposición 2.1 se tiene que

$$(I+A)^{n-1}(A\bar{z} - r\bar{z}) > 0$$

entonces $\bar{z}A(I+A)^{n-1} - \bar{z}(I+A)^{n-1}r = \bar{x}A - r\bar{x} > 0$

esta última desigualdad contradice la definición de r , porque para $\varepsilon > 0$, se tiene que $\bar{x}A - (r+\varepsilon)\bar{x} > 0$, es decir $r < r+\varepsilon \leq r\bar{x}$ por lo tanto $\bar{x}A = r\bar{x}$

Sea $\bar{x} = \bar{z}(I+A)^{n-1} = (1+r)^{n-1}\bar{z} > 0$
 como $(1+r)^{n-1}\bar{z} = \bar{x} > 0$ y $r > 0$ se sigue $\bar{z} > 0$
 con esto último se concluye la demostración.

Si $\bar{y}A = \alpha\bar{y}$ con $\bar{y} \neq 0$, entonces se sigue que $|\alpha|\|\bar{y}\| \leq \|\bar{y}\|A$ y en consecuencia $|\alpha| \leq r_{\|\bar{y}\|} \leq r$, con esto se demuestra que el módulo de todos los valores característicos no exceden a "r".

De las matrices irreducibles y no negativas considere la siguiente

PROPOSICION 2.3: Sean A y C dos matrices del mismo orden $n \times n$, donde A es irreducible y no negativa y C una matriz compleja, si ; $|C| \leq A$ (2.4)

donde $|C| = \begin{pmatrix} |C_{11}| & |C_{12}| & \dots & |C_{1n}| \\ \vdots & \vdots & \ddots & \vdots \\ |C_{n1}| & |C_{n2}| & \dots & |C_{nn}| \end{pmatrix}$ (2.5)

entonces $|\lambda| \leq r$, siendo λ un valor característico de la matriz C y r el máximo valor característico de la matriz A , la igualdad en (2.5) es válida cuando $\|C\| = A$, entonces C toma la forma:

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

donde $e^{i\varphi} = \frac{\lambda}{r}$ y D es una matriz diagonal en la que sus elementos son de módulo 1.

Demostración

Sea \bar{y} un vector característico de la matriz C que

corresponde al valor característico λ entonces

$$\bar{y} C = \lambda \bar{y} \quad \lambda \neq 0 \quad (2.6)$$

de (2.4) y (2.5) se tiene que

$$|\lambda| |\bar{y}| \leq |\bar{y}| |C| \leq |\bar{y}| A \quad ; \quad |\bar{y}| = (|y_1|, |y_2|, \dots, |y_n|) \quad (2.7)$$

por lo tanto $|\lambda| \leq r_{|y_j|} \leq r$

cuando $|\lambda| = r$, se sigue de (2.7) que $|\bar{y}|$ es un vector extremo para la matriz A y consecuentemente $|\bar{y}| > 0$ y es un vector característico de la matriz A correspondiente al valor característico r , entonces la relación (2.7) toma la forma

$$|\bar{y}| A = |\bar{y}| |C| = r |\bar{y}| \quad |\bar{y}| > 0 \quad (2.8)$$

y por (2.4)

$$|C| = A \quad (2.9)$$

sea $\bar{y} = (y_1, y_2, y_3, \dots, y_n)$ donde

$$y_j = |y_j| \exp(i\varphi_j) \quad ; \quad j=1, 2, \dots, n$$

y defínase a la matriz diagonal D por la igualdad

$$D = \{e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_n}\}$$

entonces

$$\bar{y} = D |\bar{y}|$$

sustituyendo esta expresión en (2.6) y haciendo $\lambda = r e^{i\theta}$ se encuentra que

$$e^{-i\theta} |\bar{y}| D^{-1} C D = r |\bar{y}| \quad (2.10)$$

comparando (2.8) con (2.10) se obtiene

$$e^{-i\theta} |\bar{y}| D^{-1} C D = |\bar{y}| |C| = |\bar{y}| A \quad (2.11)$$

de esto último se concluye que

$$e^{-i\theta} |D^{-1} C D| = |C| = A$$

y de (2.11) se sigue que

$$e^{-i\theta} |\bar{y}| D^{-1} C D = e^{-i\theta} |\bar{y}| |D^{-1} C D|$$

ya que $|S| > 0$, esta igualdad vale cuando

$$\bar{e}^{i\theta} D^{-1} C D = \bar{e}^{i\theta} |D^{-1} C D|$$

es decir

$$\bar{e}^{i\theta} D^{-1} C D = A$$

entonces

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

que es lo que se deseaba.

Aplicando la proposición anterior a una matriz irreducible $A \geq 0$ que tiene exactamente "h" valores característicos todos de módulo r , donde r es el máximo valor característico de A , es decir

$$\lambda_0 = r e^{i\theta_0}, \lambda_1 = r e^{i\theta_1}, \dots, \lambda_{h-1} = r e^{i\theta_{h-1}}$$

donde $0 = \theta_0 < \theta_1 < \dots < \theta_{h-1} < 2\pi$

Si en la proposición anterior se tiene que $C = A$ y $\lambda = \lambda_k$, $k = 0, 1, 2, \dots, h-1$, entonces

$$A = e^{i\theta_k} D_k A D_k^{-1} \quad (2.12)$$

donde D_k es una matriz diagonal con $|D_k| = I$, (I , matriz unitaria)

Nuevamente sea \bar{z} un vector característico positivo de la matriz A correspondiente al valor característico r , es decir

$$\bar{z} A = r \bar{z} \quad (2.13)$$

Entonces, si se define $\bar{y}^k = D_k |\bar{y}^k|$ (2.14)
con $|\bar{y}^k| > 0$, de lo anterior se deduce que

$$\bar{y}^k A = \lambda_k \bar{y}^k, \lambda_k = r e^{i\theta_k}, k = 0, 1, \dots, h-1 \quad (2.15)$$

las últimas desigualdades demuestran que los vectores $\bar{y}^0, \bar{y}^1, \dots, \bar{y}^{(h-1)}$ son vectores característicos de la matriz A para los valores característicos $\lambda_0, \lambda_1, \dots, \lambda_{h-1}$.

De (2.12) se tiene que

$$\begin{aligned} A\bar{y}^0 &= e^{i\theta_0} D_0 A D_0^{-1} \bar{y}^0 = e^{i\theta_0} D_0 A D_0^{-1} D_0 |\bar{y}^0| = e^{i\theta_0} D_0 A |\bar{y}^0| = \\ &= r e^{i\theta_0} D_0 |\bar{y}^0| = r e^{\theta_0} \bar{y}^0 = r \bar{y}^0 \end{aligned}$$

de esto último se sigue que $\lambda_0 = r$, donde, $D_0 = I$ y $\bar{y}^0 > 0$
de (2.12) también se concluye que $\lambda_1, \lambda_2, \dots, \lambda_{h-1}$ son simples y
que

$$A = \exp\{i(\theta_j \pm \theta_k)\} (D_j D_k^{\pm 1}) A (D_j D_k^{\pm 1})^{-1}$$

$$= \exp\{i(\theta_j \pm \theta_k)\} D_j D_k^{\pm 1} A D_k^{\pm 1} D_j^{-1}, \quad j, k = 0, 1, 2, \dots, h-1$$

y como en el caso anterior, también se concluye que $D_j D_k^{\pm 1} \bar{z}$ es un
vector característico de la matriz A que pertenece al valor característi-
co $r \exp\{i(\theta_j \pm \theta_k)\}$, por lo tanto $\exp\{i(\theta_j \pm \theta_k)\}$ coincide con
uno de los números $\exp(i\theta_\ell)$ y la matriz $D_j D_k^{\pm 1}$ coincide con la
correspondiente matriz D_ℓ es decir, existe $0 \leq \ell_1, \ell_2 \leq h-1$ tal que

$$\exp\{i(\theta_j + \theta_k)\} = \exp(i\theta_{\ell_1}); \quad \exp\{i(\theta_j - \theta_k)\} = \exp(i\theta_{\ell_2}); \quad \text{y}$$

$$D_j D_k = D_{\ell_1}; \quad D_j D_k^{-1} = D_{\ell_2}.$$

Así los números $(e^{i\theta_0}, e^{i\theta_1}, \dots, e^{i\theta_{h-1}})$ por un
lado y las correspondientes matrices diagonales D_0, D_1, \dots, D_{h-1} por
el otro forman dos grupos multiplicativos abelianos é isomorfos.

En cualquier grupo finito que consista de h elementos
distintos, la h -ésima potencia de cualquier elemento, es igual a la iden-
tidad del grupo. (El orden de cada elemento en un grupo finito divide al
orden del grupo).

Por lo tanto $e^{i\theta_0}, e^{i\theta_1}, \dots, e^{i\theta_{h-1}}$ son las h -ésimas
raíces de la unidad y $0 = \theta_0 < \theta_1 < \theta_2 < \dots < \theta_{h-1} < 2\pi$, entonces

$$\theta_k = \frac{2k\pi}{h}, \quad k = 0, 1, 2, \dots, h-1 \quad y$$

$$e^{i\theta_k} = \varepsilon^k, \quad (\varepsilon = e^{i\theta_1} = e^{2\pi i/h}) \quad (2.16)$$

necesariamente se sigue que

$$\lambda_k = r \varepsilon^k, \quad k = 0, 1, 2, \dots, h-1 \quad (2.17)$$

por (2.16) y usando el isomorfismo entre los grupos multiplicativos

$$\{e^{i\theta_0}, e^{i\theta_1}, \dots, e^{i\theta_{h-1}}\} \text{ y } \{D_0, D_1, \dots, D_{h-1}\} \text{ se sigue que}$$

$$D_k = D^k, \quad (D = D_1; k = 0, 1, 2, \dots, h-1) \quad (2.18)$$

si $A = e^{i\theta_k} D_k A D_k^{-1}$ para $k=1$ se tiene

$$A = e^{2\pi i/h} D A D^{-1} \quad (2.19)$$

es decir el segundo miembro de esta igualdad es una matriz similar de A , así el sistema entero de h valores característicos de la matriz A cuando son multiplicados por $e^{2\pi i/h}$ son llevados en sí mismo.

Además $D^h = I$, por lo tanto todos los elementos diagonales en D son h -ésimas raíces de la unidad, donde los elementos diagonales de $D^h = I$ son h a lo más y pueden repetirse, permutando los índices de la matriz A y también los de D , esta última matriz puede ser puesta en la forma casi-diagonal

$$D = \{\eta_0 I_0, \eta_1 I_1, \dots, \eta_{s-1} I_{s-1}\} \quad (2.20)$$

donde I_0, I_1, \dots, I_{s-1} son matrices idénticas, con $s \leq h$ y

$$\eta_p = e^{i\varphi_p}; \quad \varphi_p = \eta_p \frac{2\pi}{h}$$

con $\eta_p \in \mathbb{Z}$, $p = 0, 1, \dots, s-1$, $0 = \eta_0 < \eta_1 < \eta_2 < \dots < \eta_{s-1} < h$

De (2.20) se sigue que la matriz A toma la forma de bloques

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & A_{22} & \dots & A_{2s} \\ \dots & \dots & \dots & \dots \\ A_{s1} & A_{s2} & \dots & A_{ss} \end{bmatrix} \quad (2.21)$$

Si se reemplaza la igualdad (2.19) por el sistema

$$\varepsilon A_{pq} = \frac{\eta_{q-1}}{\eta_{p-1}} A_{pq} \quad (2.22)$$

$$p, q = 1, 2, \dots, s, \quad \varepsilon = e^{2\pi i/h}$$

entonces, para cualquier p y q se tiene que $\frac{\eta_{q-1}}{\eta_{p-1}} = \varepsilon$ ó $A_{pq} = 0$

cuando $p=1$ se tiene $\frac{\eta_1}{\eta_0}, \frac{\eta_2}{\eta_0}, \dots, \frac{\eta_{s-1}}{\eta_0}$ donde alguno de ellos es igual a ε y ello solo sucede cuando $n_1=1$, entonces $\varphi_1 = 2\pi/h$ y $\eta_1 = e^{2\pi i/h}$ y se tiene que $q=2$, así $\frac{\eta_1}{\eta_0} = \varepsilon$ y $A_{11} = A_{13} = \dots = A_{1s} = 0$.

Análogamente cuando $p=2$ se tiene $\frac{\eta_2}{\eta_1}, \frac{\eta_3}{\eta_1}, \dots, \frac{\eta_{s-1}}{\eta_1}$ donde alguno de ellos es igual a ε y ello sucede cuando $n_2=2$ entonces $\varphi_2 = 4\pi/h$ y $\eta_2 = e^{4\pi i/h} = \varepsilon^2$ en consecuencia $q=3$ por consiguiente $\frac{\eta_2}{\eta_1} = \frac{\varepsilon^2}{\varepsilon} = \varepsilon$ y $A_{21} = A_{22} = A_{24} = \dots = A_{2s}$; de esta mane-

ra A debe tener la forma

$$A = \begin{bmatrix} 0 & A_{12} & 0 & \dots & 0 \\ 0 & 0 & A_{23} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_{s-1,s} \\ A_{s1} & A_{s2} & A_{s3} & \dots & A_{ss} \end{bmatrix}$$

donde $n_1=1, n_2=2, \dots, n_{s-1}=s-1$

Pero cuando $p=s$ el lado derecho de la igualdad (2.22) tiene el factor

$$\frac{\eta_{q-1}}{\eta_{s-1}} = \exp\{(q-s)2\pi i/h\}; \quad q=1, 2, \dots, s \quad (2.23)$$

y uno de esos números debe ser igual a $\varepsilon = e^{2\pi i/h}$, esto es posible cuando $s=h$ y $q=1$, de esta manera se tiene que $\exp(i-h)2\pi i/h = e^{(1-h)2\pi i/h} = e^{2\pi i/h - 2\pi i} = e^{2\pi i/h} \cdot e^{-2\pi i} = e^{2\pi i/h} = \varepsilon$; en consecuencia se tiene que $A_{s2} = \dots = A_{ss} = 0$ esto demuestra que

$$D = \{ I_0, \varepsilon I_1, \varepsilon^2 I_2, \dots, \varepsilon^{h-1} I_{h-1} \}$$

y la matriz A toma la forma

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

y mediante una permutación de renglones y columnas A puede ser puesta en la forma

$$\begin{bmatrix} 0 & 0 & \dots & A_{1h} \\ A_{21} & 0 & \dots & 0 \\ 0 & A_{32} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_{h-1,h} & 0 \end{bmatrix} \quad (2.24)$$

donde los ceros de la diagonal son matrices cuadradas.

La representación (2.24) para una matriz irreducible $A \geq 0$ es conocida como la forma normal de dicha matriz.

Todas estas propiedades pueden ser resumidas en un teorema conocido como el Teorema de Frobenius y que es una generalización del Teorema de Perron para matrices no negativas irreducibles, por lo tanto se tiene:

TEOREMA 2.1 (FROBENIUS) : Una matriz no negativa A de orden $n \times n$ y valor característico positivo r , que es una raíz simple de la ecuación característica y excede el módulo de todos los otros valores característicos. A este máximo valor característico r corresponde un vector característico $\bar{z} = (z_1, z_2, \dots, z_n)$ de A con coordenadas positivas $z_i > 0$.

Además, si A tiene " h " valores característicos, $\lambda_0 = r, \lambda_1, \lambda_2, \dots, \lambda_{h-1}$ de módulo r , entonces esos números son distintos y son raíces de la ecuación $\lambda^h - r^h = 0$

Más generalmente, todo el espectro $\lambda_0, \lambda_1, \dots, \lambda_{h-1}$ de A , considerado como un sistema de puntos en el λ -plano complejo va sobre sí mismo bajo una rotación de el plano por el ángulo $2\pi/h$. Si $h > 1$, entonces A puede ser puesta por medio de una permutación en la siguiente forma

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

donde los bloques a lo largo de la diagonal principal son cuadrados.

Considérense algunas observaciones muy importantes relacionadas con el teorema que se demostró

COROLARIO 2.1 La matriz adjunta $B(\lambda)$ de la matriz característica $\lambda I - A$ definida como $B(\lambda) = \|B_{ik}(\lambda)\|$ de orden $n \times n$, donde $B_{ik}(\lambda)$ es el complemento algebraico (cofactor) del elemento $\lambda \delta_{ik} - a_{ik}$ en el determinante $\Delta(\lambda) = |\lambda I - A|$.

Del hecho de que un vector característico \bar{z} corresponde

al valor característico r se sigue que $B(r) \neq 0$ y que en cada columna distinta de cero de $B(r)$ todos los elementos son del mismo signo, esto mismo también es verdad para los renglones de $B(r)$ ya que A puede ser reemplazada por su traspuesta, de estas propiedades se sigue que todos los $B_{ik}(r)$, $(i, k = 1, 2, \dots, n)$ son distintos de cero y del mismo signo σ , por lo tanto

$$r\Delta'(r) = r \sum_{i=1}^n B_{ii}(r) > 0$$

es decir $\Delta'(r) \neq 0$ y r es una raíz simple de la ecuación característica $\Delta(\lambda) = 0$; por otro lado ya que r es la máxima raíz de $\Delta(\lambda) = \lambda^n + \dots$ $\Delta(\lambda)$ crece para $\lambda \geq r$ de esto se tiene que $\Delta'(r) > 0$ y $\sigma = 1$, es decir

$$B_{ik}(r) > 0 \quad i, k = 1, 2, \dots, n \quad (2.25)$$

de aquí que

$$B(r) > 0 \quad (2.26)$$

Nuevamente considérese la matriz adjunta reducida

definida como

$$C(\lambda) = \frac{B(\lambda)}{D_{n-1}(\lambda)}$$

donde $D_{n-1}(\lambda)$ es el máximo común divisor de todos los polinomios $B_{ik}(\lambda)$. Como $B_{ik}(r) > 0$ se sigue que $D_{n-1}(r) \neq 0$ y todas las raíces de $D_{n-1}(\lambda)$ son valores característicos distintas de r , por lo tanto dichas raíces son complejas o son reales menores que r , en consecuencia se tiene que $D_{n-1}(r) > 0$ y como $B(r) > 0$ se sigue que

$$C(r) = \frac{B(r)}{D_{n-1}(r)} > 0$$

COROLARIO 2.2 Si r es el máximo valor característico está acotado por s y S , es decir $s \leq r \leq S$, donde

$$s = \min_{1 \leq i \leq n} s_i \quad \text{con} \quad s_i = \sum_{k=1}^n a_{ik}; \quad (i = 1, 2, \dots, n)$$

$$S = \max_{1 \leq i \leq n} s_i$$

COROLARIO 2.3 : Si $r = \max_{\bar{x} \geq 0} r_{\bar{x}}$ donde $r_{\bar{x}}$ es el mayor número ρ que satisface $\rho \bar{x} \leq \bar{x}A$, similarmente se puede definir para cada vector $\bar{x} \geq 0$, ($\bar{x} \neq 0$) un número $r_{\bar{x}}$ como el menor número ρ para lo cual $\rho \bar{x} \geq \bar{x}A$, donde

$$r_{\bar{x}} = \max_{1 \leq i \leq n} \frac{(\bar{x}A)_i}{x_i}$$

si alguna $x_i = 0$ y $(\bar{x}A)_i \neq 0$ entonces $r_{\bar{x}} = +\infty$.

Si la función $r_{\bar{x}}$ toma su mínimo valor r' en algún vector $\bar{v} > 0$ entonces r' que se define como

$$r' = \min_{\bar{x} \geq 0} r_{\bar{x}} = \min_{\bar{x} \geq 0} \max_{1 \leq i \leq n} \frac{(\bar{x}A)_i}{x_i}$$

coincide con r y que el vector $\bar{v} \geq 0$ ($\bar{v} \neq 0$) para lo cual este mínimo es tomado es un vector característico de A para $\lambda = r'$, porque si

$$r' \bar{v} - \bar{v}A \geq 0$$

Supóngase que la igualdad no puede ser cierta, entonces por proposición 2.1

$$(I + A)^{n-1} (r' \bar{v} - \bar{v}A) > 0$$

y haciendo

$$\bar{u} = \bar{v}(I + A)^{n-1}$$

se tiene que

$$r' \bar{u} > \bar{u}A$$

y para $\varepsilon > 0$

$$(r' - \varepsilon) \bar{u} > \bar{u}A, \quad \bar{u} > 0$$

contradice la definición de r' , así

$$\bar{v}A = r' \bar{v}$$

así se tiene que

$$r' = r$$

en consecuencia r tiene una doble caracterización

$$r = \max_{\bar{x} \geq 0} \min_{1 \leq i \leq n} \frac{(\bar{x}A)_i}{x_i} = \min_{\bar{x} \geq 0} \max_{1 \leq i \leq n} \frac{(\bar{x}A)_i}{x_i}$$

y como el $\max_{\bar{x} > 0}$ ó el $\min_{\bar{x} > 0}$ es tomado en un vector característico positiva para $\lambda = r$, de esto también se concluye que

$$\min_{1 \leq i \leq n} \frac{(\bar{x}A)_i}{\bar{x}_i} \leq r \leq \max_{1 \leq i \leq n} \frac{(\bar{x}A)_i}{\bar{x}_i}, \quad \bar{x} \geq 0, \quad \bar{x} \neq 0 \quad (2.27)$$

por esto mismo se sigue que

$$\begin{aligned} & r\bar{z} \leq \bar{z}A, \quad \bar{z} \geq 0, \quad \bar{z} \neq 0 \\ \text{ó} & r\bar{z} \geq \bar{z}A, \quad \bar{z} \geq 0, \quad \bar{z} \neq 0 \end{aligned}$$

$$\text{entonces} \quad \bar{z}A = r\bar{z}, \quad \bar{z} > 0$$

3 MATRICES NO NEGATIVAS REDUCIBLES

En el capítulo anterior se investigó la teoría de las matrices no negativas irreducibles, considérese ahora, extensiones de esos resultados en la teoría de las matrices reducibles, para ello se tiene la siguiente

Definición 3.1: Una matriz cuadrada A y no negativa se dice que es reducible si existe una matriz permutación P tal que

$$PAP^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{13} \end{bmatrix}$$

donde A_{11}, A_{12}, A_{13} son matrices cuadradas.

Las propiedades espectrales para las matrices no negativas irreducibles que anteriormente fueron establecidas no son preservadas cuando se estudian las matrices no negativas reducibles. Sin embargo ya que cualquier matriz cuadrada no negativa puede ser puesta en una sucesión de matrices cuadradas irreducibles y positivas, simplemente reemplazando cada cero de la matriz por una $\epsilon > 0$ arbitrariamente pequeña, de aquí que A puede ser presentada como

$$A = \lim_{m \rightarrow \infty} A_m \quad \text{con } A_m > 0 \text{ e irreducibles} \\ m = 1, 2, \dots$$

Para una matriz no negativa y arbitraria se tiene el siguiente

TEOREMA 3.1: Una matriz no negativa A de orden $n \times n$ siempre tiene un valor característico r tal que el módulo de todos los valores característicos de A no exceden a r . A éste máximo valor característico corresponde un vector característico \bar{y} tal que

$$\bar{y}A = r\bar{y} \quad (\bar{y} \geq 0, \bar{y} \neq 0)$$

La matriz adjunta $B(\lambda)$ definida como $B(\lambda) = (\lambda I - A)^{-1} A$

satisface las desigualdades

$$B(\lambda) \geq 0, \quad \frac{d}{d\lambda} B(\lambda) \geq 0 \quad \text{para } \lambda \geq r$$

Demostración

Sea

$$A = \lim_{m \rightarrow \infty} A_m \quad \text{con } A_m > 0 \text{ é irreducibles, } m=1,2,\dots$$

denótese por $r^{(m)}$ y $\bar{y}^{(m)}$ al máximo valor característico de la matriz positiva A_m y su correspondiente vector característico positivo normalizado entonces

$$\bar{y}^{(m)} A_m = r^{(m)} \bar{y}^{(m)} \quad \text{tal que } (\bar{y}^{(m)}, \bar{y}^{(m)}) = 1 \quad (3.1)$$

$$\bar{y}^{(m)} \geq 0, \quad m=1,2,\dots$$

como $A = \lim_{m \rightarrow \infty} A_m$ se sigue que

$$\lim_{m \rightarrow \infty} r^{(m)} = r \quad \text{existe}$$

donde r es el valor característico para la matriz A , y como $r^{(m)} > 0$ y $r^{(m)} > |\lambda_0^{(m)}|$ donde $\lambda_0^{(m)}$ es un valor característico arbitrario de A_m , de lo anterior se tiene que

$r \geq 0$ con $|\lambda_0| \leq r$ para λ_0 un valor característico arbitrario de A y en lugar de (2.25) se obtiene que

$$B(r) \geq 0 \quad (3.2)$$

Además de la sucesión de vectores característicos $\bar{y}^{(m)}$ se selecciona una subsucesión $\bar{y}^{(m_p)}$, $p=1,2,\dots$ que converge a un vector normalizado \bar{y} entonces

$$\lim_{m \rightarrow \infty} \bar{y}^{(m_p)} A_m = \bar{y} A \quad ; \quad \lim_{m \rightarrow \infty} r^{(m_p)} \bar{y}^{(m_p)} = r \bar{y}$$

y por (3.1) se sigue que

$$\bar{y} A = r \bar{y} \quad ; \quad \bar{y} \geq 0, \bar{y} \neq 0$$

Para demostrar la última parte, se establece una inducción en n .

bre el orden de la matriz, para $n=1$

se sabe que $B(\lambda) = (\lambda I - A)^{-1} \Delta(\lambda)$ entonces $B(\lambda) \equiv I$ y $\frac{d}{d\lambda} B(\lambda) \equiv 0$
 Sea A una matriz de orden $n \times n$ y supóngase que es válido para matrices de orden $(n-1) \times (n-1)$

Extendiendo el determinante característico $\Delta(\lambda) = |\lambda I - A|$ con respecto a los elementos de la última fila y en seguida con respecto a los elementos de la última columna se obtiene

$$\Delta(\lambda) = (\lambda I - A) B_{nn}(\lambda) - \sum_{i,k=1}^{n-1} B_{ki}^{(n)}(\lambda) a_{in} a_{nk} \quad (3.3)$$

donde $B_{nn}(\lambda) = |\lambda \delta_{ik} - a_{ik}|$ es el determinante característico de una matriz de orden $n-1$ y $B_{ki}^{(n)}(\lambda)$ es el cofactor del elemento $\lambda \delta_{ik} - a_{ik}$ en $B_{nn}(\lambda)$;

$i, k = 1, 2, \dots, n-1$, es decir

$$-a_{ni} B_{ni}^{(n)} = (-a_{n-1,n} B_{n-1,i} - a_{n-2,n} B_{n-2,i} - \dots - a_{in} B_{ii}) (-a_{ni})$$

$$-a_{ni} B_{ni}^{(n)} = \left(-\sum_{k=1}^{n-1} a_{kn} B_{ki}^{(n)} \right) (-a_{ni})$$

$$\sum_{i=1}^{n-1} a_{ni} B_{ni}^{(n)} = \sum_{i=1}^{n-1} \sum_{k=1}^{n-1} B_{ki}^{(n)}(\lambda) a_{kn} a_{ni}$$

la máxima raíz no negativa de $B_{nn}(\lambda)$ será denotada por r_n , haciendo $\lambda = r_n$ en (3.3) se tiene que $B_{ki}^{(n)}(r_n) \geq 0$ $i, k = 1, 2, \dots, n-1$, de aquí que $\Delta(r_n) \leq 0$ por otro lado $\Delta(\lambda) = \lambda^n + \dots$ de tal manera que $\Delta(+\infty) = +\infty$, por lo tanto r_n es una raíz de $\Delta(\lambda)$ ó es menor que alguna raíz real de $\Delta(\lambda)$, en ambos casos $r_n \leq r$ y para cualquier menor principal B_{ij} con valor característico r_j se tiene que $r_j \leq r$ (3.4)

Además $B_{ik}(\lambda)$ puede ser representado como un menor de orden $(n-1) \times (n-1)$ de la matriz característica $\lambda I - A$ multiplicado por $(-1)^{i+k}$ y derivando se obtiene

$$\frac{d}{d\lambda} B_{ik}(\lambda) = \sum_j B_{ik}^{(j)}(\lambda) \quad i, k = 1, 2, \dots, n-1 \quad (3.5)$$

donde $B^{(j)}(\lambda) = [B_{ik}^{(j)}(\lambda)]$ $i \neq j, k \neq j$, $j = 1, 2, \dots, n$ es la matriz adjunta.

de la matriz $\|a_{ik}\|$ $i, k = 1, 2, \dots, j-1, j+1, \dots, n$ de orden $n-1$.

Por hipótesis de inducción

$$B^{(j)}(\lambda) \geq 0 \text{ para } \lambda \geq r_j \quad j = 1, 2, \dots, n$$

y por tanto por (3.4) y (3.5)

$$\frac{d}{d\lambda} B(\lambda) \geq 0 \text{ para } \lambda \geq r \quad (3.6)$$

de (2.27) y (3.4) se sigue que

$$B(\lambda) \geq 0 \text{ para } \lambda \geq r \quad (3.6')$$

COROLARIO 3.1: si A es una matriz no negativa de orden $n \times n$ con máximo valor característico r y $C(\lambda)$ es su matriz adjunta reducida, entonces

$$C(\lambda) \geq 0 \text{ para } \lambda \geq r \quad (3.7)$$

Demostración.

$$C(\lambda) = \frac{B(\lambda)}{D_{n-1}(\lambda)} \quad (3.8)$$

como $D_{n-1}(\lambda)$ es el máximo común divisor de los elementos de $B(\lambda)$ y divide al polinomio característico $\Delta(\lambda)$ y $D_{n-1} = \lambda^{n-1} + \dots$, entonces

$$D_{n-1}(\lambda) > 0 \text{ para } \lambda > r \quad (3.9)$$

pero $B(\lambda) \geq 0$ para $\lambda \geq r$ se sigue que $C(\lambda) \geq 0$ para $\lambda \geq r$

COROLARIO 3.2: si $A \geq 0$ es una matriz irreducible con máximo valor característico r , entonces

$$B(\lambda) > 0, \quad C(\lambda) > 0 \text{ para } \lambda \geq r \quad (3.10)$$

por (2.26) $B(r) > 0$, pero también $\frac{d}{d\lambda} B(\lambda) \geq 0$ para $\lambda \geq r$

por lo tanto $B(\lambda) > 0$ para $\lambda \geq r$ (3.11)

y de (3.8), (3.9) y (3.11) se tiene que $C(\lambda) > 0$ para $\lambda \geq r$

COROLARIO 3.3: si $A \geq 0$ es una matriz irreducible con máximo valor característico r , entonces

$$(\lambda I - A)^{-1} > 0 \text{ para } \lambda \geq r \quad (3.12)$$

Demostración

$$(\lambda I - A)^{-1} = \frac{B(\lambda)}{\Delta(\lambda)}$$

ya que $B(\lambda) > 0$ y $\Delta(\lambda) > 0$ para $\lambda > r$, de aquí se sigue (3.12)

COROLARIO 3.4: El máximo valor característico r' de cualquier menor principal de una matriz no negativa A no excede al máximo valor característico r de A ; es decir

$$r' \leq r \quad (3.13)$$

- a) si A es irreducible, la igualdad en (3.13) no puede ocurrir
 b) si A es reducible, la igualdad en (3.13) vale para al menos un menor principal

Demostración

a) la desigualdad (3.13) es verdadera para cualquier menor principal de orden $n-1$ por (3.4). Si A es irreducible, entonces por (2.25) $B_{jj}(r) > 0$ $j=1, 2, \dots, n$ y por lo tanto $r' \neq r$.

b) si A es reducible, entonces existe una matriz permutación P tal que

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

de aquí que r debe ser un valor característico de uno de los dos menores principales B y D , esto prueba (b).

Considérese una matriz reducible arbitraria A de orden $n \times n$ por medio de una matriz permutación P , la matriz A puede ser puesta en la forma

$$A = \begin{bmatrix} B & O \\ C & D \end{bmatrix} \quad (3.14)$$

donde B y D son matrices cuadradas. Si una de las matrices B ó D es reducible, pueden ser representadas en forma similar a (3.14) de tal manera que A toma la forma

$$A = \begin{bmatrix} K & O & O \\ H & L & O \\ F & G & M \end{bmatrix}$$

si una de las matrices K, L, M es reducible, entonces el proceso continúa, finalmente y mediante una matriz permutación aceptable la matriz A toma la forma triangular en bloques, es decir:

$$A = \begin{bmatrix} A_{11} & O & \dots & O \\ A_{21} & A_{22} & \dots & O \\ \dots & \dots & \dots & \dots \\ A_{s1} & A_{s2} & \dots & A_{ss} \end{bmatrix} \quad (3.15)$$

donde las matrices de la diagonal son cuadradas e irreducibles.

Los bloques de la diagonal A_{ii} ; ($1 \leq i \leq s$) son llamados aislados si

$$A_{ik} = 0 \quad \text{para } k = 1, 2, \dots, i-1, i+1, \dots, s$$

por lo tanto A toma la forma

$$A = \begin{bmatrix} A_1 & O & \dots & O & O & \dots & O \\ O & A_2 & \dots & O & O & \dots & O \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ O & O & \dots & A_g & O & \dots & O \\ A_{g+1,1} & A_{g+1,2} & \dots & A_{g+1,g} & A_{g+1} & \dots & O \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ A_{s1} & A_{s2} & \dots & A_{sg} & A_{s,g+1} & \dots & A_s \end{bmatrix} \quad (3.16)$$

donde A_1, A_2, \dots, A_s son matrices irreducibles, y en cada fila de $A_{g+1,1}, A_{g+1,2}, \dots, A_{g+1,g}$ se tiene que al menos una matriz es distinta de cero; y a la matriz A en (3.16) es llamada la forma normal de la matriz reducible A .

Haciendo uso de la forma normal de una matriz reducible se prueba el siguiente

TEOREMA 3.2 : Al valor característico (máximo) r de la matriz $A \geq 0$ pertenece un vector característico positivo si y solo si en la forma normal de A se tiene que a) cada una de las matrices A_1, A_2, \dots, A_g tienen a r como valor característico; b) ninguna de las matrices A_{g+1}, \dots, A_s tienen esta propiedad.

Demostración

Supóngase que $\bar{z} > 0$ sea un vector característico positivo que pertenece a r ; y es necesario dividir a \bar{z} en partes de tal manera que para

$$\bar{z}A = r\bar{z} \quad (3.17)$$

es reemplazada por dos sistemas de ecuaciones

$$\bar{z}^i A_i = r\bar{z}^i \quad i=1, 2, \dots, g \quad (3.17')$$

$$\sum_{h=1}^{j-1} \bar{z}^h A_{jh} + \bar{z}^j A_j \quad \text{para } j=g+1, \dots, s \quad (3.17'')$$

de (3.17') se sigue que r es un valor característico de cada una de las matrices A_1, A_2, \dots, A_g ; de (3.17'') se observa que

$$\bar{z}^j A_j \leq r\bar{z}^j \quad \text{y que } \bar{z}^j A_j \neq r\bar{z}^j \quad (3.18)$$

para $j = g+1, \dots, s$

Sea r_j el máximo valor característico de A_j y por (2.27) se encuentra que

$$r_j \leq \max_i \frac{(\bar{z}^i A_j)}{\bar{z}_i^j} \leq r \quad j=g+1, \dots, s$$

por otro lado $r_j \neq r$, por lo tanto $r_j < r$

Inversamente, supóngase que los máximos valores característicos de las matrices A_i para $i=1, 2, \dots, q$ son iguales a r y r_j son los máximos valores característicos para $A_{j,1}, \dots, A_{j,q-1}$; entonces se pueden definir columnas características positivas \bar{z}^i para las matrices A_i $i=1, 2, \dots, q$ por medio de (3.17') y de (3.17'') se encuentran columnas \bar{z}^j $j=q+1, \dots, s$, es decir

$$\sum_{h=1}^{j-1} \bar{z}^h A_{jh} = z^j (r_j I_j - A_j)$$

se sigue que $\bar{z}^j = (r_j I_j - A_j)^{-1} \sum_{h=1}^{j-1} \bar{z}^h A_{jh}$, $j=q+1, \dots, s$ (3.19)

donde I_j es una matriz unitaria del mismo orden que A_j y ya que $r_j < r$ y por (3.12) se tiene que

$$(r I_j - A_j)^{-1} > 0 \quad j=q+1, \dots, s \quad (3.20)$$

y como $\sum_{h=1}^{j-1} \bar{z}^h A_{jh} \geq 0$ para $\sum_{h=1}^{j-1} \bar{z}^h A_{jh} \neq 0$

que junto con (3.20) se tiene que $\bar{z}^j > 0$

Por lo tanto \bar{z} es un vector característico de A para el valor característico r .

El siguiente teorema da una caracterización para una matriz $A \geq 0$ que junto con su transpuesta A' tiene la propiedad que un vector característico positivo pertenece al máximo valor característico r

TEOREMA 3.3: Al máximo valor característico r de una matriz $A \geq 0$ pertenece un vector característico positivo para ambas A y A' si y solo si A puede ser representada en forma casi-diagonal

$$A = \{A_1, A_2, \dots, A_s\} \quad (3.21)$$

donde A_1, A_2, \dots, A_s son matrices irreducibles, cada una de las cuales tiene a r como máximo valor característico.

Demostración

Supóngase que A y A' tienen vectores característicos positivos para $\lambda = r$; por teorema 3.2 A puede ser representada en su forma normal donde A_1, \dots, A_g tienen a r como su máximo valor característico y los máximos valores característicos de A_{g+1}, \dots, A_s son menores que r , se tiene que

$$A' = \begin{bmatrix} A'_1 & \cdot & \cdot & \cdot & 0 & A'_{g+1,1} & \cdot & \cdot & \cdot & A'_{s1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & A'_g & A'_{g+1,g} & \cdot & \cdot & \cdot & A'_{sg} \\ 0 & \cdot & \cdot & \cdot & 0 & A'_{g+1} & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & 0 & 0 & \cdot & \cdot & \cdot & A'_s \end{bmatrix}$$

invirtiendo el orden de los bloques en esta matriz se tiene

$$\begin{bmatrix} A'_s & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ A'_s & A'_{s-1} & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ A'_{s1} & A'_{s-1,1} & \cdot & \cdot & \cdot & \cdot & A'_1 \end{bmatrix} \quad (3.22)$$

Ya que $A'_s, A'_{s-1}, \dots, A'_1$ son irreducibles, de aquí se obtiene una forma normal para (3.22) por una permutación de los bloques, colocando los bloques aislados a lo largo de la diagonal principal, uno de esos bloques es A'_s , por otro lado ya que la forma normal de A' debe satisfacer las condiciones del teorema precedente, el máximo valor característico de A'_s debe ser igual a r , esto es posible cuando $g=s$, pero entonces la forma normal es todo A_1, A_2, \dots, A_s .

Inversamente, si una representación de A es dada como

$$A = \{A_1, A_2, \dots, A_s\} \text{ entonces}$$

$$A' = \{A'_1, A'_2, \dots, A'_s\} \quad (3.23)$$

se deduce de (3.21) y (3.23) y por el teorema anterior que A y A' tienen vectores característicos positivos para el mismo valor característico.

COROLARIO 3.5: Si el máximo valor característico r de una matriz $A \geq 0$ es simple y si vectores característicos positivos pertenecen a A y A' entonces A es irreducible.

MATRICES PRIMITIVAS E IMPRIMITIVAS

Definición 3.2: Sea $A \geq 0$ una matriz irreducible y sea r su máximo valor característico. Supóngase que existen exactamente h valores característicos de módulo r ($\lambda_1 = |\lambda_2| = \dots = |\lambda_h| = r$). Si $h=1$ la matriz es llamada primitiva; si $h > 1$ la matriz es llamada imprimitiva y el número h es llamado el índice de imprimitividad.

El siguiente teorema da una propiedad importante para matrices primitivas.

TEOREMA 3.4: (Una matriz $A \geq 0$ es primitiva si y solo si existe una potencia de A que es positiva, es decir $A^p \gg 0$ para $p \geq 1$)

Demostración

Si $A^p > 0$ la matriz A es irreducible ya que cada potencia de una matriz reducible es reducible. Además el índice de imprimitividad h para la matriz A debe ser 1, de otro modo la matriz positiva A^p podría tener h raíces características $\lambda_1^p, \lambda_2^p, \dots, \lambda_h^p$ (que pueden ser iguales) de módulo máximo r^p lo que contradice el teorema 1.1.

Inversamente, supóngase que A es una matriz primitiva, usando la fórmula de interpolación Lagrange - Sylvester para A^p * se tiene: * para fórmula de interpolación ver apéndice A1, págs 65-70

$$A^p = \sum_{k=1}^s \frac{1}{(m_k-1)!} \left[\frac{C(\lambda) \lambda^p}{\psi^{[k]}(\lambda)} \right]_{\lambda=\lambda_k}^{(m_k-1)} \quad (3.24)$$

donde $\psi(\lambda) = (\lambda-\lambda_1)^{m_1} (\lambda-\lambda_2)^{m_2} \dots (\lambda-\lambda_s)^{m_s}$ $\lambda_j \neq \lambda_f$ para $j \neq f$
 es el polinomio mínimo de la matriz A .

$$\psi^{[k]}(\lambda) = \frac{\psi(\lambda)}{(\lambda-\lambda_k)^{m_k}} \quad k = 1, 2, \dots, s.$$

y $C(\lambda)$ es la matriz adjunta reducida, definida como

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

si se toma $\lambda_1 = r > |\lambda_2| \geq \dots \geq |\lambda_s|$ con $m_1 = 1$ (3.25)

la fórmula (3.24) toma la forma

$$A^p = \frac{C(r) \cdot r^p}{\psi'(r)} + \sum_{k=2}^s \frac{1}{(m_k-1)!} \left[\frac{C(\lambda) \lambda^p}{\psi^{[k]}(\lambda)} \right]_{\lambda=\lambda_k}^{(m_k-1)}$$

de (3.25) se sigue que

$$\lim_{p \rightarrow \infty} \frac{A^p}{r^p} = \frac{C(r)}{\psi'(r)} \quad (3.26)$$

pero $C(r) > 0$ y por (3.25) $\psi'(r) > 0$; se sigue que

$$\lim_{p \rightarrow \infty} r^{-p} A^p \gg 0$$

TEOREMA 3.5: Si $A \geq 0$ es una matriz irreducible y alguna potencia A^q de A es reducible, entonces A^q es completamente reducible, es decir, A^q puede ser representada por medio de una permutación en la forma

$$A^q = \{A_1, A_2, \dots, A_s\} \quad (3.27)$$

donde A_1, A_2, \dots, A_d son matrices irreducibles que tienen uno y el mismo valor característico, y d es el máximo común divisor de q y h , donde h es el índice de imprimitividad de A .

Demostración

Ya que A es irreducible, por el TEOREMA DE FROBENIUS los vectores característicos positivos asociados a r pertenecen tanto a A como a A^q . Pero entonces esos vectores positivos son también vectores característicos de las matrices no negativas A^q y $(A^q)'$ para el valor característico $\lambda = r^q$; aplicando el teorema 3.3 a A^q , se representa esta matriz (eligiendo una matriz-permutación aceptable) en la forma (3.27) donde A_1, \dots, A_d son matrices irreducibles con máximo valor característico r^q .

Pero A tiene h valores característicos de módulo r

$$r, r\varepsilon, \dots, r\varepsilon^{h-1} \quad \varepsilon = e^{2\pi i/h}$$

esto significa que A^q también tiene h valores característicos de módulo máximo

$$r^q, r^q\varepsilon^q, \dots, r^q\varepsilon^{q(h-1)}$$

entre las cuales d son iguales a r^q . Esto solamente es posible cuando d es máximo común divisor de q y h .

COROLARIO 3.6 Una potencia de una matriz primitiva es irreducible y primitiva.

Si $q = h$ en el teorema precedente se obtiene

COROLARIO 3.7 Si A es una matriz primitiva con índice de imprimitividad h , entonces A^h es representable en " h " matrices primitivas con el mismo máximo valor característico.

4. APLICACIONES DE LAS MATRICES NONEGATIVAS EN LAS CADENAS DE MARKOV.

Considérese un sistema físico capaz de existir en "n" estados posibles.

$$S_1, S_2, S_3, \dots, S_n \quad (4.1)$$

y una sucesión ordenada de instantes de tiempo

$$t_0, t_1, t_2, \dots, t_n, \dots$$

supóngase que en cada instante de tiempo, el sistema está en uno y solo uno de los estados (4.1). Sea p_{ij} la probabilidad condicional de que el sistema esté en el estado S_j en el instante t_k , suponiendo que el sistema estuvo en el estado S_i en el instante precedente t_{k-1} , ($i, j = 1, 2, \dots, n$; $k = 1, 2, \dots$). Supóngase que p_{ij} - también llamada probabilidad de transición, es independiente del tiempo t_k , es decir, p_{ij} no varía con el tiempo.

Si la matriz de probabilidades de transición está dada por

$$P = \|p_{ij}\|$$

de orden $n \times n$, entonces se dice que se tiene una cadena homogénea de Markov con un número finito de estados, además P satisface que

$$p_{ij} \geq 0, \quad \sum_{j=1}^n p_{ij} = 1 \quad i, j = 1, 2, \dots, n \quad (4.2)$$

Definición 4.1: Una matriz $P = \|p_{ij}\|$ de orden $n \times n$ es llamada estocástica si P es no negativa y si la suma de los elementos de cada renglón de P es igual a 1, es decir si las relaciones (4.2) son válidas.

Así, para cadenas homogéneas de Markov la matriz P es estocástica é inversamente cada matriz estocástica puede ser considerada como la matriz P de alguna cadena homogénea de Markov, y ésta es la base para investigar las cadenas homogéneas de Markov.

Como una matriz estocástica es una forma especial de una matriz no negativa, todos los conceptos y proposiciones anterior-

mente estudiadas son aplicables.

PROPOSICION 4.1 : $P \geq 0$ es estocástica si y solo si tiene el vector característico $(1, 1, \dots, 1)'$ para el valor característico 1, y en una matriz estocástica el máximo valor característico es 1.

Demostración

De la definición de una matriz estocástica se sigue que tiene valor característico 1, con vector característico positivo \bar{x} tal que $\bar{x} = (1, 1, \dots, 1)'$, inversamente, cada matriz $P \geq 0$ que tenga vector característico $(1, 1, \dots, 1)'$ para el valor característico 1 es estocástica, por otro lado el máximo valor característico siempre está incluido entre el máximo y el mínimo de las sumas de los renglones, y como en una matriz estocástica todas las sumas de los renglones son iguales a 1, entonces 1 es el máximo valor característico para dicha matriz.

TEOREMA 4.1 : Cada divisor elemental de una matriz estocástica que corresponde al valor característico 1, es de primer grado.*

Demostración

Por (3.16), la matriz estocástica P puede ser descompuesta en su forma normal.

$$\begin{bmatrix} A_1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & A_2 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & A_3 & \cdot & \cdot & 0 \\ A_{g+1,1} & \cdot & \cdot & A_{g+1,g} & A_{g+1} & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ A_{s1} & \cdot & \cdot & A_{sg} & \cdot & \cdot & \cdot & A_s \end{bmatrix}$$

donde A_1, A_2, \dots, A_s son matrices irreducibles y A_1, A_2, \dots, A_g son matrices estocásticas irreducibles con valor característico simple 1, para cada una de ellas, por otro lado, los bloques diagonales A_{g+1}, \dots, A_s son irreducibles con valores característicos menores que 1.

* En el apéndice A.2 pág 71-73 se exponen los divisores elementales

Así, la matriz P se representa en la forma

$$\begin{bmatrix} Q_1 & 0 \\ S & Q_2 \end{bmatrix}$$

donde en Q_1 al valor característico λ le corresponden divisores elementales de primer grado, la matriz Q_2 no tiene valor característico λ , y el teorema se sigue inmediatamente del siguiente

LEMA 4.1: Si una matriz tiene la forma

$$\begin{bmatrix} Q_1 & 0 \\ S & Q_2 \end{bmatrix}$$

donde Q_1 y Q_2 son matrices cuadradas y el valor característico λ_0 de A es también un valor característico de Q_1 pero no de Q_2 , y

$$|Q_1 - \lambda_0 I| = 0 \quad , \quad |Q_2 - \lambda_0 I| \neq 0$$

entonces los divisores elementales de A y Q_1 correspondientes a λ_0 son los mismos.

Demostración

i) Considérese el caso donde Q_1 y Q_2 no tengan valores característicos en común, en éste caso los divisores elementales de Q_1 y Q_2 forman el sistema de divisores elementales de A , para eso sea T una matriz tal que

$$TAT^{-1} = \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} \quad (4.3)$$

Si T es de la forma

$$T = \begin{bmatrix} I_1 & 0 \\ U & I_2 \end{bmatrix}$$

donde I_1 y I_2 son matrices unitarias, entonces

$$\begin{bmatrix} I_1 & 0 \\ U & I_2 \end{bmatrix} \begin{bmatrix} Q_1 & 0 \\ S & Q_2 \end{bmatrix} \begin{bmatrix} I_1 & 0 \\ -U & I_2 \end{bmatrix} = \begin{bmatrix} Q_1 & 0 \\ UQ_1 - Q_2U + S & Q_2 \end{bmatrix} \quad (4.3)$$

la ecuación (4.3) se reduce a la forma (4.3) si se elige a la matriz rect-

angular U de tal manera que satisfaga a la ecuación

$$Q_2 U - U Q_1 = S$$

si esto sucede entonces los divisores elementales de Q_1 y Q_2 forman el sistema de divisores elementales para la matriz A ; ya que Q_1 y Q_2 no tienen valores característicos en común, entonces su ecuación siempre tiene solución única para cada S .

ii) En el caso donde Q_1 y Q_2 tengan valores característicos en común, se reemplaza a Q_1 por su forma de Jordan J (como un resultado A es reemplazado por una matriz similar). Sea $J = \{J_1, J_2\}$ donde todos los bloques de Jordan con valores característicos λ_0 son combinados en J_1 , entonces

$$A = \begin{bmatrix} J_1 & 0 & 0 & 0 \\ 0 & J_2 & 0 & 0 \\ s_{11} & s_{12} & \vdots & \dots \\ s_{21} & s_{22} & \vdots & Q_2 \end{bmatrix} = \begin{bmatrix} J_1 & 0 & 0 & 0 \\ 0 & \vdots & \dots & \dots \\ s_{11} & \vdots & \bar{Q}_2 & \vdots \\ s_{21} & \vdots & \vdots & \vdots \end{bmatrix}$$

Esta matriz cae bajo el caso precedente ya que las matrices J_1 y \bar{Q}_2 no tienen valores en común, de aquí que los divisores elementales de la forma $(\lambda - \lambda_0)^p$ son los mismos para A y J_1 y por lo tanto para A y Q_1 .

COROLARIO 4.1: Si un vector característico positivo pertenece al máximo valor característico r de A , entonces todos los divisores elementales de A que pertenecen a λ_0 , con $|\lambda_0| = r$, son del primer grado.

Volviendo a las cadenas de Markov, supóngase que s_1, s_2, \dots, s_n sean todos los posibles estados para un sistema en una cadena de Markov y sea P la matriz estocástica determinada por esa cadena, que es formada de las probabilidades de transición p_{ij} , $i, j = 1, 2, \dots, n$

Sea $p_{ij}^{(q)}$ la probabilidad de encontrar el sistema en el estado s_j en el instante t_k sabiendo que en el instante t_{k-q} el sistema estuvo en el estado s_i para $i, j = 1, 2, \dots, n$, $q = 1, 2, \dots$ donde $p_{ij}^{(1)} = p_{ij}$ y haciendo uso de la suma y multiplicación de probabilidades ó simplemente del hecho de que $P^{m+n} = P^m P^n$ se tiene que

$$p_{ij}^{(q+1)} = \sum_{h=1}^n p_{ih}^{(q)} p_{hj} \quad i, j = 1, 2, \dots, n$$

(Esta relación es conocida como la relación Chapman-Kolmogorov para las cadenas homogéneas de Markov)

En notación matricial se tiene

$$\|P_{ij}^{(q+1)}\| = \|P_{ij}^{(q)}\| \|P_{ij}\|$$

donde ambas matrices son de orden $n \times n$ y dando valores sucesivos a q , $q=1, 2, \dots$, se obtiene que

$$\|P_{ij}^{(q)}\| = P^q \quad q=1, 2, \dots$$

Si el límite

$$\lim_{q \rightarrow \infty} P_{ij}^{(q)} = P_{ij}^{\infty}$$

o en notación matricial

$$\lim_{q \rightarrow \infty} P^{(q)} = P^{\infty} = \|P_{ij}^{\infty}\|$$

existe, entonces los números P_{ij}^{∞} son llamados las probabilidades de transición final, y P^{∞} es estocástica, también se tiene que P_{ij}^{∞} no dependen del índice k del tiempo original. Para investigar bajo que condiciones P_{ij}^{∞} existe, se introduce la siguiente terminología:

Una matriz estocástica P y su correspondiente cadena homogénea de Markov será llamada regular si P no tiene valores característicos de módulo 1, más que 1 mismo, y totalmente regular, si en adición, 1 es una raíz simple de la ecuación característica de P .

Una matriz regular se caracteriza por el hecho de que en su forma normal las matrices A_1, A_2, \dots, A_g son primitivas. Para una matriz totalmente regular debe tenerse que $g=1$.

Definición 4.2: Una cadena homogénea de Markov será irreducible, reducible, acíclica ó cíclica, si la matriz estocástica P de la cadena es respectivamente irreducible, reducible, primitiva ó imprimitiva.

Como una matriz estocástica primitiva es una forma especial de una matriz regular, por tanto una cadena de Markov acíclica es una forma especial de una cadena regular.

TEOREMA 4.2: a) En una cadena homogénea de Markov todas las probabilidades de transición final existen si y solo si la cadena es regular, en ese caso la matriz P^∞ formada por las probabilidades de transición final es determinada por

$$P^\infty = \frac{C(1)}{\Psi'(1)}$$

b) En una cadena regular homogénea de Markov las probabilidades de transición final son independientes del estado inicial si y solo si la cadena es totalmente regular, en ese caso la matriz P^∞ es determinada por

$$P^\infty = \frac{B(1)}{\Delta'(1)}$$

c) En una cadena regular homogénea de Markov todas las probabilidades de transición final son distintas de cero si y solo si la cadena es acíclica.

Demostración

a) Sea $\Psi(\lambda)$ el polinomio mínimo de la matriz regular P , entonces

$$\Psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_u)^{m_u} \quad (4.4)$$

y por teorema 4.1 se puede suponer que $\lambda_1 = 1$, $m_1 = 1$, entonces

$$P^q = \frac{C(1)}{\Psi'(1)} + \sum_{k=2}^u \frac{1}{(m_k - 1)!} \left[\frac{C(\lambda)}{\Psi^k(\lambda)} \lambda^q \right]_{\lambda=\lambda_k}^{(m_k-1)} \quad (4.5)$$

donde $C(\lambda) = (\lambda I - P)^{-1} \Psi(\lambda)$ es la matriz adjunta reducida y

$$\Psi^k(\lambda) = \frac{\Psi(\lambda)}{(\lambda - \lambda_k)^{m_k}}, \quad k=1, 2, \dots, u$$

además

$$\Psi^1(\lambda) = \frac{\Psi(\lambda)}{\lambda - 1} \quad \text{para } \Psi^1 = \Psi'(1).$$

Si P es una matriz regular, entonces $|\lambda_k| < 1$, $k=2, 3, \dots, u$, y por lo tanto todos los términos del lado derecho de (4.5) excepto el primero tienden a cero cuando "q" tiende a "∞"; de aquí que para una matriz regular P , la matriz P^∞ formada por las probabilidades de transición final existe y esta dada por

$$P^\infty = \frac{C(1)}{\Psi'(1)} \quad (4.6)$$

inversamente si

$$\lim P^q = P^\infty \quad (4.7)$$

existe, entonces la matriz P no puede tener algún valor característico λ_k para lo cual $\lambda_k \neq 1$ y $|\lambda_k| = 1$, ya que entonces el $\lim \lambda_k^q$ cuando $q \rightarrow \infty$ podría no existir, y éste límite debe existir por (4.7).

Resumiendo, si P es la matriz de transición de una cadena homogénea de Markov, entonces la matriz P^∞ (de las probabilidades de transición final) existe si y solo si la cadena es regular, y la matriz P^∞ es la matriz definida por (4.6).

Sea

$$\Delta(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdot \cdot \cdot (\lambda - \lambda_u)^{n_u}$$

el polinomio característico de P

Sea

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

la matriz adjunta de P , entonces:

$$\frac{B(\lambda)}{\Delta(\lambda)} = \frac{C(\lambda)}{\Psi(\lambda)}$$

y para $n_i = 1$, $\lambda_i = 1$ se tiene

$$\frac{n_i B^{n_i-1}(1)}{\Delta^{n_i}(1)} = \frac{C(1)}{\Psi'(1)}$$

por lo tanto (4.6) puede ser reemplazado por

$$P^\infty = \frac{n_i B^{n_i-1}(1)}{\Delta^{n_i}(1)} \quad (4.8)$$

para una cadena totalmente regular, en tanto como es una forma especial de una cadena regular, la matriz P^∞ existe y está determinada por (4.6) ó (4.8); como $n_i = 1$ se sigue que (4.8) toma la forma

$$P^\infty = \frac{B(1)}{\Delta'(1)} \quad (4.9)$$

b) Considérese una cadena regular (no totalmente regular), y sea P su correspondiente matriz estocástica, y que tenga la forma normal

$$P = \begin{bmatrix} Q_1 & \cdot & \cdot & \cdot & 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & Q_g & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ U_{g+1,1} & \cdot & \cdot & \cdot & U_{g+1,g} & Q_{g+1} & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ U_{s1} & \cdot & \cdot & \cdot & U_{sg} & \cdot & \cdot & \cdot & \cdot & U_{s-1,s} & Q_s \end{bmatrix} \quad (4.10)$$

donde Q_1, \dots, Q_g son matrices primitivas y estocásticas y Q_{g+1}, \dots, Q_s son matrices irreducibles cuyo máximo valor característico es menor que 1. Haciendo

$$U = \begin{bmatrix} U_{g+1,1} & \cdot & \cdot & \cdot & U_{g+1,g} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ U_{s1} & \cdot & \cdot & \cdot & U_{sg} \end{bmatrix} \quad W = \begin{bmatrix} Q_{g+1} & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ U_{s,g+1} & \cdot & \cdot & \cdot & Q_s \end{bmatrix}$$

entonces

$$P = \begin{bmatrix} Q_1 & \cdot & \cdot & \cdot & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & Q_g & \cdot & \cdot & \cdot & 0 \\ & & & U & & & & & \\ & & & & & & & & W \end{bmatrix}$$

En consecuencia se tiene que P^g puede ser escrita como

$$P^q = \begin{bmatrix} Q_1^q & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & Q_g^q & \dots & 0 \\ U^q & & W^q & & \end{bmatrix} \quad (4.11)$$

$$y \quad \lim_{q \rightarrow \infty} P^q = \begin{bmatrix} Q_1^\infty & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & Q_g^\infty & \dots & 0 \\ U^\infty & & W^\infty & & \end{bmatrix} = P^\infty$$

Pero $\lim_{q \rightarrow \infty} W^q = W^\infty = 0$ porque el módulo de los valores característicos de W es menor que 1, por lo tanto

$$P^\infty = \begin{bmatrix} Q_1^\infty & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & Q_g^\infty & \dots & 0 \\ U^\infty & & 0 & & \end{bmatrix} \quad (4.12)$$

Ya que Q_1, \dots, Q_g son matrices estocásticas primitivas, las matrices $Q_1^\infty, \dots, Q_g^\infty$ son positivas por (4.9) y (2.26), es decir, $Q_1^\infty > 0, \dots, Q_g^\infty > 0$:

Los estados S_1, S_2, \dots, S_n del sistema en consideración pueden partirse en grupos de la forma

$$\Sigma_1, \Sigma_2, \dots, \Sigma_g, \Sigma_{g+1}, \dots, \Sigma_s \quad (4.13)$$

donde el número de estados en el primer grupo es igual al número de renglones en el bloque Q_1 en (4.10) etc.

Los estados que pertenecen a los grupos $\Sigma_1, \dots, \Sigma_g$ son llamados esenciales, los estados de los grupos restantes son llamados no esenciales.

Obsérvese que de las formas (4.10) y (4.11) de la matriz P^q solamente los siguientes q -pasos transición desde el instante t_{k-q} al instante t_k son posibles (a) de un estado esencial a otro estado esencial, (b) de un estado no esencial a otro estado esencial, (c) de un estado no esencial a otro estado no esencial del mismo grupo o de un grupo anterior.

De (4.12) se sigue que las probabilidades de transición en un estado no esencial después de q -pasos tiende a cero cuando q tiende a infinito, por esta razón los estados esenciales son algunas veces llamados estados límites

Por otro lado

$$(I - P)P^\infty = 0$$

esto demuestra que cada una de las columnas de la matriz P^∞ es un vector característico de la matriz estocástica P correspondiendo al valor característico $\lambda = 1$.

Si P es una matriz totalmente regular, el número 1 es una raíz simple de la ecuación característica y como $g=1$ se tiene que puede ser solamente un vector característico $(1, 1, \dots, 1)'$; así, todos los elementos en una columna fija, digamos la j -ésima de la matriz P^∞ tiene el mismo valor P_{*j}^∞

$$P_{ii}^\infty = P_{*j}^\infty \geq 0 \quad , \quad j=1, 2, \dots, n; \quad \sum_{j=1}^n P_{*j}^\infty = 1 \quad (4.14)$$

lo que en términos físicos significa que para una cadena totalmente regular las probabilidades de transición final no dependen del estado inicial S_i

Supóngase inversamente que, el límite de las probabilidades de transición para una cierta cadena regular homogénea de Markov sea independiente del estado inicial, es decir, supóngase que (4.14) vale, entonces en (4.12) debe tenerse que $g=1$ y por lo tanto $h=1$ y la cadena es regular.

c) Para una cadena acíclica, que es un caso especial de una cadena totalmente regular P es una matriz primitiva y por teorema 3.4 $P^q > 0$ para alguna $q > 0$, pero entonces

$$P^\infty = \lim_{m \rightarrow \infty} P^m P^q > 0 \quad \text{para } m > q$$

inversamente si $P^\infty > 0$ se sigue que $P^q > 0$ para alguna $q > 0$, y por teorema 3.4 se tiene que P es primitiva, y entonces la cadena homogénea es acíclica.

Además si $P^\infty > 0$ implica que la cadena es acíclica y por lo tanto regular de aquí que las probabilidades de transición final no dependen del estado inicial.

Definición 4.3: Las columnas $p_1^k, p_2^k, \dots, p_n^k$ son llamadas las probabilidades absolutas del proceso, donde p_i^k es la probabilidad de encontrar el sistema en el estado S_i para $i=1, 2, \dots, n$; $k=0, 1, 2, \dots$ en el instante t_k .

$$\text{Sea } p^k = (p_1^k, p_2^k, \dots, p_n^k) \quad k=0, 1, 2, \dots \quad (4.15)$$

haciendo uso de los teoremas sobre adición y multiplicación de probabilidades se encuentra que

$$p_i^k = \sum_{h=1}^n p_h^0 p_{hi}^{(k)}, \quad i=1, 2, \dots, n, \quad k=1, 2, \dots$$

o en notación matricial

$$p^k = (P')^k p^0 \quad k=1, 2, \dots \quad (4.16)$$

donde P' es la transpuesta de P .

Todas las probabilidades absolutas pueden ser determinadas por (4.16) si las probabilidades iniciales $p_1^0, p_2^0, \dots, p_n^0$ y la matriz P son conocidas.

Es necesario introducir el límite de las probabilidades absolutas como

$$p_i^\infty = \lim_{k \rightarrow \infty} p_i^k \quad i=1, 2, \dots, n$$

ó

$$p^\infty = (p_1^\infty, p_2^\infty, \dots, p_n^\infty) = \lim_{k \rightarrow \infty} p^k$$

si en ambos lados de (4.16) $k \rightarrow \infty$, entonces de (4.16) se obtiene

$$p^\infty = (P^\infty)' p^0 \quad (4.17)$$

así, si la matriz P^∞ que es el resultado de las probabilidades de transición final, existe, entonces el límite de las probabilidades absolutas p^∞ tal que $p^\infty = (p_1^\infty, p_2^\infty, \dots, p_n^\infty)$ debe existir para probabilidades iniciales arbitrarias $p^0 = (p_1^0, p_2^0, \dots, p_n^0)$, y viceversa.

De (4.17) y de (4.12) para P^∞ se sigue que, el límite de las probabilidades absolutas correspondientes a estados no esenciales son cero.

De la igualdad $P' \cdot (P^\infty)' = (P^\infty)'$ si se multiplican ambos miembros por p^0 y usando (4.17) se obtiene que

$$P' \cdot (P^\infty)' p^0 = (P^\infty)' p^0$$

$$P' p^\infty = p^\infty \quad (4.18)$$

es decir el vector columna p^∞ es un vector característico de la matriz P' para el valor característico $\lambda = 1$.

Si la cadena es totalmente regular, entonces $\lambda = 1$ es una raíz simple de la ecuación característica de la matriz P' , en este caso el vector columna p^∞ está determinado únicamente por (4.18) ya que $p_j^\infty \geq 0$ y su suma es 1, es decir $\sum_{j=1}^n p_j^\infty = 1$.

Una propiedad más de una cadena totalmente regular se deriva al combinar (4.14) con (4.17)

$$p_j^\infty = \sum_{h=1}^n p_h^0 p_{hj}^\infty = p_{*j}^\infty \sum_{h=1}^n p_h^0 = p_{*j}^\infty, \quad j=1,2,\dots,n \quad (4.19)$$

así el límite de las probabilidades absolutas $p_1^\infty, p_2^\infty, \dots, p_n^\infty$ no dependen de las probabilidades iniciales $p_1^0, p_2^0, \dots, p_n^0$.

Inversamente, p^∞ es independiente de p^0 si y solo si todos los renglones de P^∞ son iguales, es decir

$$p_{hj}^\infty = p_{*j}^\infty \quad h, j = 1, 2, \dots, n$$

y por teorema 4.2 P es totalmente regular.

Si P es primitiva, entonces $P^\infty \gg 0$ y por (4.19) $p_j^\infty > 0$ $j=1,2,\dots,n$.

Inversamente si toda p_j^∞ es positiva y no dependen de las probabilidades iniciales, entonces todos los elementos en cada columna de P^∞ son iguales y por (4.19) $P^\infty \gg 0$ y por teorema 4.2 P es primitiva es decir la cadena es acíclica.

de estas propiedades se sigue que el teorema 4.2 puede ser escrito como sigue

- TEOREMA 4.2': a) En una cadena homogénea de Markov todos los límites de probabilidades absolutas existen para probabilidades iniciales arbitrarias si y solo si la cadena es regular.
- b) En una cadena homogénea de Markov el límite de probabilidades absolutas existe para probabilidades iniciales arbitrarias y es independiente de dichas probabilidades si y solo si la cadena es totalmente regular.
- c) En una cadena homogénea de Markov el límite de probabilidades absolutas existe y es positivo para probabilidades iniciales é independiente de ellas si y solo si la cadena es acíclica.

Este teorema es algunas veces llamado el teorema Ergódico, y el teorema precedente es conocido como el teorema general Casi-Ergódico para cadenas homogéneas de Markov.

Nuevamente, supóngase que (3.16) sea la forma normal de la matriz P y denótese por h_1, h_2, \dots, h_g los índices de imprimitividad de las matrices A_1, A_2, \dots, A_g ; sea h el mínimo común múltiplo de los enteros h_1, h_2, \dots, h_g , de esto último se sigue que no tiene valores característicos de módulo 1 mas que uno mismo, es decir, P^h es una matriz regular, además P, P^2, \dots, P^{h-1} no son regulares; llámese a h el período de la cadena homogénea de Markov.

Ya que P^h es una matriz regular el límite

$$\lim_{q \rightarrow \infty} P^{hq} = (P^h)^\infty$$

existe, y así los límites

$$P_r = \lim_{q \rightarrow \infty} P^{r+qh} = P^r (P^h)^\infty$$

$$r = 0, 1, 2, \dots, h-1$$

también existen

Así en general, la sucesión de matrices

$$P, P^2, P^3, \dots$$

se dividen en h subsucesiones con los límites $P_r = P^r (P^h)^\infty$

$$r = 0, 1, \dots, h-1.$$

Al pasar por medio de (4.16) de las probabilidades de transición a las probabilidades absolutas se encuentra que la sucesión

$$p^1, p^2, \dots, p^n, \dots$$

se divide en h subsucesiones con los límites

$$\lim_{q \rightarrow \infty} p^{r+qh} = (P^h)^\infty p^r \quad r = 0, 1, \dots, h-1$$

Para una cadena homogénea de Markov arbitraria con un número finito de estados los límites de la media aritmética siempre existen

$$\tilde{P} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N p^k = \lim_{q \rightarrow \infty} P^{qh+r} = (P^h)^\infty P^r = (I + P + \dots + P^{h-1}) (P^h)^\infty \quad (4.20)$$

y

$$\tilde{P} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N p^k = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N (P^h)^k p^0 = \tilde{P}^1 p^0 \quad (4.20')$$

donde $\tilde{P} = \|\tilde{P}_{ij}\|$ es una matriz de orden $n \times n$ y $\tilde{P}^1 = (\tilde{P}_1, \tilde{P}_2, \dots, \tilde{P}_n)$. Los valores $\tilde{P}_{ij}, (i, j = 1, 2, \dots, n)$ y $\tilde{P}_j, (j = 1, 2, \dots, n)$ son la media de las probabilidades de

transición final y la media del límite de las probabilidades absolutas, respectivamente.

Ya que

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^{N+1} P^k = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N P^k$$

se tiene que

$$\tilde{P}P = \tilde{P}$$

y por (4.20')

$$P'\tilde{P} = \tilde{P}$$

(4.21)

es decir \tilde{P} es un vector característico de P' para $\lambda = 1$.

Por (3.16) y (4.20) \tilde{P} puede ser representada en la forma

$$\tilde{P} = \begin{bmatrix} \tilde{A}_1 & 0 & \dots & 0 \\ 0 & \tilde{A}_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \tilde{A}_g \\ \tilde{U} & & & \tilde{W} \end{bmatrix}$$

donde

$$A_i = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N A_i^k \quad i = 1, 2, \dots, g$$

$$W = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N W^k$$

$$W = \begin{bmatrix} A_{g+1} & 0 & \dots & 0 \\ * & A_{g+2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ * & * & \dots & A_g \end{bmatrix}$$

Ya que todos los valores característicos de W son menores que 1 se tiene que

$$\lim_{k \rightarrow \infty} W^k = 0$$

y por lo tanto $\tilde{W} = 0$

Entonces

$$\tilde{P} = \begin{bmatrix} \tilde{A}_1 & 0 & \dots & 0 & 0 \\ 0 & \tilde{A}_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \tilde{A}_g & 0 \\ \dots & \tilde{U} & \dots & \dots & 0 \end{bmatrix} \quad (4.22)$$

ya que \tilde{P} es una matriz estocástica, las matrices $\tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_g$ también son estocásticas.

De esta representación de \tilde{P} y de (4.17) se sigue que: La media de las probabilidades correspondientes a estados no-escenciales son siempre cero.

Si $g=1$ en la forma normal de P , entonces $\lambda=1$ es una raíz simple de P' .

En este caso \tilde{p} está únicamente determinado por (4.21) y $\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_n$ no dependen de las probabilidades iniciales $p_1^0, p_2^0, \dots, p_n^0$.

Inversamente, si \tilde{p} no depende de p^0 , entonces P por (4.20') es de rango 1, pero el rango de (4.22) puede ser 1 solamente cuando $g=1$.

Todos los resultados anteriores pueden resumirse en el siguiente

TEOREMA 4.3: Para una cadena homogénea de Markov arbitraria con período h las matrices probabilidad P^k y p^k tienden a un período de repetición con período h para cuando $k \rightarrow \infty$, además, la media de las probabilidades de transición final y las probabilidades absolutas \tilde{P} y $\tilde{p} = (\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_n)$ definidas por (4.20) y (4.20') siempre existen.

La media de las probabilidades absolutas correspondientes a estados no-escenciales son siempre cero.

Si $g=1$ en la forma normal de P (y únicamente en este caso) la media del límite de las probabilidades absolutas $\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_n$ son independientes de las probabilidades iniciales $p_1^0, p_2^0, \dots, p_n^0$ y por (4.21) son determinadas en forma única.

A1

FUNCIONES DE MATRICES

Definición de una función de una matriz

Sea A una matriz cuadrada y $f(\lambda)$ una función de λ . Se pretende definir el significado de $f(A)$, es decir extender la función $f(\lambda)$ a un valor matricial.

Se sabe que en el caso simple donde $f(\lambda) = \nu_0 \lambda^p + \nu_1 \lambda^{p-1} + \dots + \nu_r$ es un polinomio en λ , y en éste caso se tiene que $f(A) = \nu_0 A^p + \nu_1 A^{p-1} + \dots + \nu_r E$, partiendo de éste caso especial se obtiene una definición para $f(A)$ en el caso general.

Sea $\Psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_s)^{m_s}$ el polinomio mínimo de A para $\lambda_i \neq \lambda_j$ si $i \neq j$

Sean $g(\lambda)$ y $h(\lambda)$ dos polinomios tales que

$$g(A) = h(A)$$

entonces la diferencia $d(\lambda) = g(\lambda) - h(\lambda)$ es un polinomio divisible por $\Psi(\lambda)$ sin residuo, es decir

$$g(\lambda) \equiv h(\lambda) \pmod{\Psi(\lambda)}$$

y como $\Psi(\lambda)$ es el polinomio mínimo se tiene que

$$d(\lambda_k) = 0, d'(\lambda_k) = 0, \dots, d^{(m_k-1)}(\lambda_k) = 0, \quad k=1, 2, \dots, s$$

$$\text{es decir } g(\lambda_k) = h(\lambda_k), g'(\lambda_k) = h'(\lambda_k), \dots, g^{(m_k-1)}(\lambda_k) = h^{(m_k-1)}(\lambda_k); \quad (A1.1)$$

para $k=1, 2, \dots, s$

Los m números:

$$f(\lambda_k), f'(\lambda_k), \dots, f^{(m_k-1)}(\lambda_k); \quad k=1, 2, \dots, s \quad (A1.2)$$

serán llamados los valores de la función $f(\lambda)$ sobre el espectro de la matriz A y el conjunto de todos esos valores será denotado por $f(\Delta_A)$.

Si para una función $f(\lambda)$ los valores (A1.1) existen, entonces se dice que la función $f(\lambda)$ está definida sobre el espectro de la matriz A , y (A1.1) demuestra que $g(\lambda)$ y $h(\lambda)$ tienen los mismos valores sobre el espectro de A , es decir:

$$g(\Delta_A) = h(\Delta_A)$$

Así dada una matriz A , los valores del polinomio $g(\lambda)$ sobre el espectro de A determinan a la matriz $g(A)$, es decir todos los polinomios $g(\lambda)$ que toman los mismos valores sobre el espectro de A tienen uno y el mismo valor matricial $g(A)$; de aquí que

Los valores de la función $f(\lambda)$ sobre el espectro de A deben determinar completamente a $f(A)$, es decir, todas las funciones $f(\lambda)$ que toman los mismos valores sobre el espectro de A deben tener el mismo valor matricial, entonces, para la definición general, es suficiente buscar un polinomio $g(\lambda)$ que tome los mismos valores sobre el espectro de A , y se tendrá que

$$f(A) = g(A)$$

Definición A1.1: si la función $f(\lambda)$ está definida sobre el espectro de A , entonces $f(\Lambda_A) = g(\Lambda_A)$, donde $g(\lambda)$ es un polinomio arbitrario que toma sobre el espectro de A los mismos valores que $f(\lambda)$, es decir $f(A) = g(A)$.

Entre todos los polinomios con coeficientes complejos que toman sobre el espectro de A los mismos valores que $f(\lambda)$ existe uno y solo un polinomio $r(\lambda)$ que es de grado menor que m , siendo m el grado del polinomio mínimo, y éste polinomio es obtenido de cualquier otro polinomio que tenga los mismos valores espectrales tomando el residuo en la división por $\psi(\lambda)$ de ese polinomio. Este polinomio $r(\lambda)$ está determinado únicamente por las condiciones de interpolación

$$r(\lambda_k) = f(\lambda_k), r'(\lambda_k) = f'(\lambda_k), \dots, r^{(m_k-1)}(\lambda_k) = f^{(m_k-1)}(\lambda_k) \quad (A1.3)$$

$k=1, 2, \dots, s$

el polinomio $r(\lambda)$ es llamado el polinomio interpolación de Lagrange-Silvester para $f(\lambda)$ sobre el espectro de A ; de aquí se sigue

Definición A1.1': sea $f(\lambda)$ una función definida sobre el espectro de A y $r(\lambda)$ el correspondiente polinomio de interpolación, entonces

$$f(A) = r(A)$$

A manera de ejemplo considérese

$$H = \begin{bmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \dots & 1 \\ 0 & 0 & 0 & \dots & \dots & 0 \end{bmatrix}$$

su polinomio mínimo es λ^n , por lo tanto los valores de $f(\lambda)$ sobre el espectro de H son los números $f(0), f'(0), \dots, f^{(n-1)}(0)$ y el polinomio $r(\lambda)$ es de la forma

$$r(\lambda) = f(0) + \frac{f'(0)}{1!} \lambda + \dots + \frac{f^{(n-1)}(0)}{(n-1)!} \lambda^{n-1}$$

por lo tanto

$$f(H) = f(0)I + \frac{f'(0)}{1!} H + \dots + \frac{f^{(n-1)}(0)}{(n-1)!} H^{n-1} = \begin{bmatrix} f(0) & \frac{f'(0)}{1!} & \dots & \dots & \frac{f^{(n-1)}(0)}{(n-1)!} \\ 0 & f(0) & & & \vdots \\ \vdots & & & & \vdots \\ \vdots & & & & \frac{f'(0)}{1!} \\ 0 & 0 & \dots & \dots & f(0) \end{bmatrix}$$

EL POLINOMIO INTERPOLACION DE LAGRANGE-SILVESTER.

Considérese el caso en la cual la ecuación característica $|\lambda I - A| = 0$ no tiene raíces múltiples, las raíces que son los valores característicos de la matriz A son denotados por $\lambda_1, \lambda_2, \dots, \lambda_n$, entonces

$$\Psi(\lambda) = |\lambda I - A| = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n)$$

y (A1.3) puede ser escrito como $r(\lambda_k) = f(\lambda_k)$ para $k=1, 2, \dots, n$.

Por otro lado sea ξ un polinomio de grado n , entonces ξ tiene a lo más n raíces; sean c_1, c_2, \dots, c_m las raíces definidas, entonces aplicando m veces se tiene que

$$\xi(t) = (t - c_1)(t - c_2) \dots (t - c_m) \delta, \quad \text{entonces}$$

$$\text{grad } \xi(t) = \text{grad}(t - c_1) + \dots + \text{grad}(t - c_m) + \text{grad } \delta \geq m$$

así $\text{grad } \xi = n$ se sigue que $m \leq n$

Constrúyase una colección de polinomios de grado $(n-1)$ que forme una base para el espacio de todos los polinomios de grado $m \leq n$:

no mayor o igual que $n-1$.

Sea c_1, \dots, c_n tales que $c_i \neq c_j$ para $i \neq j$, se define

$$R_i = \frac{(t-c_1)(t-c_2)\dots(t-c_{i-1})(t-c_{i+1})\dots(t-c_n)}{(c_i-c_1)(c_i-c_2)\dots(c_i-c_{i-1})(c_i-c_{i+1})\dots(c_i-c_n)} \quad (A1.4)$$

$$i=1, 2, \dots, n$$

R_i se obtiene al tomar el producto de todos los factores del numerador excepto el factor $(t-c_i)$ sobre el producto de todos los factores en el denominador excepto el factor (c_i-c_i) , entonces R_i es un polinomio de grado $(n-1)$ y $R_i(c_j)=0$ para $i \neq j$ y $R_i(c_i)=1, i=1, \dots, n$

Sea ξ un polinomio de grado menor o igual que $(n-1)$ y sea $p = \xi - \xi(c_1)R_1 - \xi(c_2)R_2 - \dots - \xi(c_n)R_n$, si $p \neq 0$, entonces $\text{grad } p \leq (n-1)$ de aquí que $p(c_i) = \xi(c_i) - \xi(c_1)R_1(c_i) - \xi(c_2)R_2(c_i) - \dots - \xi(c_n)R_n(c_i)$, se sigue que $R_1(c_i)=1$ y $R_i(c_i)=0$ si $i > 1$, en consecuencia $p(c_i)=0$, similarmente $p(c_2)=p(c_3)=\dots=p(c_n)=0$, esto quiere decir que p tiene " n " raíces y si $p \neq 0$, $\text{grad } p \leq (n-1)$ que es una contradicción, entonces $p=0$ estas propiedades se resumen en el siguiente

TEOREMA A1.1: Si c_1, c_2, \dots, c_n son n raíces distintas, sea R_i definido por (A1.4) para $i=1, 2, \dots, n$, supóngase que ξ sea un polinomio real tal que $\xi=0$ ó $\text{grad } \xi \leq (n-1)$, entonces

$$\xi = \xi(c_1)R_1 + \dots + \xi(c_n)R_n \quad (A1.5)$$

$c_1, c_2, \dots, c_n \in \mathbb{R}$, $c_i \neq c_j$, R_i son los polinomios de interpolación de Lagrange y (A1.5) es llamada la fórmula de interpolación de Lagrange.

Para el caso que interesa, $r(\lambda)$ es el polinomio interpolación de Lagrange para la función $f(\lambda)$ en los puntos $\lambda_1, \lambda_2, \dots, \lambda_n$, en consecuencia

$$r(\lambda) = \sum_{k=1}^n \frac{(\lambda-\lambda_1)\dots(\lambda-\lambda_{k-1})(\lambda-\lambda_{k+1})\dots(\lambda-\lambda_n)}{(\lambda_k-\lambda_1)\dots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\dots(\lambda_k-\lambda_n)} f(\lambda_k)$$

Por definición

$$f(A) = r(A) = \sum_{k=1}^n \frac{(A-\lambda_1 I)\dots(A-\lambda_{k-1} I)(A-\lambda_{k+1} I)\dots(A-\lambda_n I)}{(\lambda_k-\lambda_1)\dots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\dots(\lambda_k-\lambda_n)} f(\lambda_k)$$

Si el polinomio característico tiene raíces múltiples, pero el polinomio mínimo que es un divisor del polinomio característico, tiene raíces simples, en este caso como el precedente todos los exponentes $m_k = 1$ y $r(\lambda)$ tiene la misma representación que la anterior

Considérese el caso general, es decir

$$\Psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_s)^{m_s} \quad m_1 + m_2 + \dots + m_s = m$$

se representa a la función racional $\frac{r(\lambda)}{\Psi(\lambda)}$ (donde el grado de $r(\lambda)$

es menor que el grado de $\Psi(\lambda)$) como una suma de fracciones parciales

$$\frac{r(\lambda)}{\Psi(\lambda)} = \sum_{k=1}^s \left[\frac{\alpha_{k1}}{(\lambda - \lambda_k)^{m_k}} + \frac{\alpha_{k2}}{(\lambda - \lambda_k)^{m_k-1}} + \dots + \frac{\alpha_{km_k}}{\lambda - \lambda_k} \right] \quad (A1.6)$$

donde α_{kj} son constantes para $j=1, 2, \dots, m_k$; $k=1, 2, \dots, s$

Para determinar los numeradores α_{kj} , se multiplican ambos lados de (A1.6) por $(\lambda - \lambda_k)^{m_k}$ y se denota por $\Psi_k(\lambda)$ el polinomio $\frac{\Psi(\lambda)}{(\lambda - \lambda_k)^{m_k}}$ entonces

$$\frac{r(\lambda)}{\Psi_k(\lambda)} = \alpha_{k1} + \alpha_{k2}(\lambda - \lambda_k) + \dots + \alpha_{km_k}(\lambda - \lambda_k)^{m_k-1} + (\lambda - \lambda_k)^{m_k} p_k(\lambda) \quad (A1.7)$$

$k=1, 2, \dots, s$

donde $p_k(\lambda)$ es una función racional, regular para $\lambda = \lambda_k$.

entonces $\alpha_{k1} = \left[\frac{r(\lambda)}{\Psi_k(\lambda)} \right]_{\lambda = \lambda_k}$

$$\alpha_{k2} = \left[\frac{r(\lambda)}{\Psi_k(\lambda)} \right]'_{\lambda = \lambda_k} = r(\lambda_k) \left[\frac{1}{\Psi_k(\lambda)} \right]'_{\lambda = \lambda_k} + r'(\lambda_k) \frac{1}{\Psi_k(\lambda_k)}$$

$k=1, 2, \dots, s$

Las fórmulas (A1.8) demuestran que α_{kj} se expresan en términos de los valores de $r(\lambda)$ sobre el espectro de A , y esos valores son iguales a los valores correspondientes de la función $f(\lambda)$, por lo tanto:

$$\alpha_{k1} = \frac{f(\lambda_k)}{\Psi_k(\lambda_k)}, \quad \alpha_{k2} = f(\lambda_k) \left[\frac{1}{\Psi_k(\lambda)} \right]'_{\lambda = \lambda_k} + f'(\lambda_k) \frac{1}{\Psi_k(\lambda_k)}; \dots \quad (A1.9)$$

$k=1, 2, \dots, s.$

Las fórmulas en (A1.9) pueden ser abreviadas como sigue

$$\alpha_{kj} = \frac{1}{(j-1)!} \left[\frac{f(\lambda)}{\Psi_k(\lambda)} \right]_{\lambda=\lambda_k}^{(j-1)} \quad \begin{matrix} j=1, 2, \dots, m_k \\ k=1, 2, \dots, s \end{matrix} \quad (A1.10)$$

cuando todos los α_{kj} han sido encontrados se puede determinar $r(\lambda)$ de la siguiente fórmula obtenida de (A1.6) al multiplicar ambos lados por $\Psi(\lambda)$, es decir:

$$r(\lambda) = \sum_{k=1}^s \left[\alpha_{k1} + \alpha_{k2}(\lambda - \lambda_k) + \dots + \alpha_{km_k}(\lambda - \lambda_k)^{m_k-1} \right] \Psi_k(\lambda) \quad (A1.11)$$

En esta fórmula la expresión entre los paréntesis que multiplica a $\Psi_k(\lambda)$ es, por (A1.10) igual a la suma de los primeros m_k términos de $f(\lambda)$ en la expansión de Taylor en potencias de $(\lambda - \lambda_k)$.

Cuando en (A1.11) se sustituyen las expresiones de (A1.9) por los coeficientes de α y se combinan todos los términos que contienen uno y el mismo valor de la función $f(\lambda)$ ó por una de sus derivadas, se representa a $f(\lambda)$ en la forma:

$$r(\lambda) = \sum_{k=1}^s \left[f(\lambda_k) \varphi_{k1}(\lambda) + f'(\lambda_k) \varphi_{k2}(\lambda) + \dots + f^{(m_k-1)}(\lambda_k) \varphi_{km_k}(\lambda) \right] \quad (A1.12)$$

donde $\varphi_{kj}(\lambda)$ son polinomios en λ de grado menor que m para $k=1, 2, \dots, m_k$ $j=1, 2, \dots, s$. Esos polinomios se determinan cuando es dado $\Psi(\lambda)$ que no dependa de $f(\lambda)$; el número de esos polinomios es igual a m que es el grado del polinomio mínimo. Las funciones $\varphi_{kj}(\lambda)$ representan el polinomio interpolación de Lagrange-Silvester para la función cuyos valores sobre el espectro de A son iguales a cero excepto para $f^{(j-1)}(\lambda) = 1$.

Además todos los polinomios $\varphi_{kj}(\lambda)$ son linealmente independientes porque, supóngase que

$$\sum_{k=1}^s \sum_{j=1}^{m_k} C_{kj} \varphi_{kj}(\lambda) = 0$$

de las m condiciones supóngase que $r^{(j-1)}(\lambda_k) = C_{kj}$ (A1.13)

se sigue que

$$r(\lambda) = \sum_{k=1}^s \sum_{j=1}^{m_k} c_{kj} \varphi_{kj}(\lambda) = 0$$

por lo tanto por (A1.13) ; $c_{kj} = 0$

De (A1.12) se deduce la fórmula fundamental para $f(A)$

$$f(A) = \sum_k \left[f(\lambda_k) Z_{k1} + f'(\lambda_k) Z_{k2} + \dots + f^{(m_k-1)}(\lambda_k) Z_{km_k} \right] \quad (A1.14)$$

donde $Z_{kj} = \varphi_{kj}(A)$ $j=1, 2, \dots, m_k$; $k=1, 2, \dots, s$ (A1.15)

Las matrices Z_{kj} son determinadas cuando es dada A y no depende de la elección de $f(\lambda)$, y son llamadas las matrices componentes de A , también son linealmente independientes, porque supóngase que

$$\sum_{k=1}^s \sum_{j=1}^{m_k} c_{kj} Z_{kj} = 0$$

para
$$\chi(\lambda) = \sum_{k=1}^s \sum_{j=1}^{m_k} c_{kj} \varphi_{kj}(\lambda)$$

por (A1.15) $\chi(A) = 0$

pero el grado de $\chi(\lambda)$ es menor que m (grado del polinomio mínimo) se sigue que

$$\chi(\lambda) = 0$$

Por otro lado las m funciones $\varphi_{kj}(\lambda)$ son linealmente independientes de aquí que $c_{kj} = 0$ $j=1, 2, \dots, m_k$, $k=1, 2, \dots, s$

De la independencia lineal de Z_{kj} se sigue que ninguna de esas matrices puede ser cero. Cuando la matriz A es dada y se requiere encontrar sus componentes, hágase $f(\mu) = \frac{1}{\lambda - \mu}$ donde λ es un parámetro y se obtiene

$$f(A) = (\lambda I - A)^{-1} \quad \text{para } f(A) = r(A) \text{ donde } r(\mu) \text{ es el polinomio}$$

... mio interpolación de Lagrange - Sylvester. Del hecho de que $f(\mu)$ y $r(\mu)$ coinciden sobre el espectro de A , se sigue que $(\lambda - \mu)r(\mu) = (\lambda - \mu)f(\mu) = 1$ coinciden sobre su espectro, de aquí se sigue que $(\lambda E - A)r(A) = (\lambda E - A)f(A) = E$ y por lo anterior se obtiene

$$(\lambda I - A)^{-1} = \frac{C(\lambda)}{\Psi(\lambda)} = \sum_{k=1}^s \left[\frac{Z_{k1}}{\lambda - \lambda_k} + \frac{1! Z_{k2}}{(\lambda - \lambda_k)^2} + \dots + \frac{(m_k - 1)! Z_{km_k}}{(\lambda - \lambda_k)^{m_k}} \right], \quad (A1.16)$$

donde $C(\lambda)$ es la matriz adjunta reducida de $(\lambda E - A)$.

Las matrices $(j-1)! Z_{kj}$ son los numeradores de las fracciones parciales en la descomposición (A1.16) y por analogía con (A1.6) son expresados por los valores de $C(\lambda)$ sobre el espectro de A por fórmulas similares a (A1.8), es decir.

$$(m_k - 1)! Z_{km_k} = \frac{C(\lambda)}{\Psi_k(\lambda)} \Big|_{\lambda = \lambda_k} ; \quad (m_k - 2)! Z_{km_{k-1}} = \left[\frac{C(\lambda)}{\Psi_k(\lambda)} \right]' \Big|_{\lambda = \lambda_k} ; \dots$$

entonces se tiene que

$$Z_{kj} = \frac{1}{(j-1)!(m_k-j)!} \left[\frac{C(\lambda)}{\Psi_k(\lambda)} \right]^{(m_k-j)} \Big|_{\lambda = \lambda_k} \quad \begin{matrix} j = 1, 2, \dots, m_k \\ k = 1, 2, \dots, s \end{matrix} \quad (A1.17)$$

Quando se reemplazan en (A1.14) las componentes de la matriz A por sus expresiones dadas en (A1.17), se representa la fórmula fundamental para $f(A)$ como

$$f(A) = \sum_{k=1}^s \frac{1}{(m_k - 1)!} \left[\frac{C(\lambda)}{\Psi_k(\lambda)} f(\lambda) \right]^{(m_k - 1)} \Big|_{\lambda = \lambda_k}. \quad (A1.18)$$

En conclusión, potencias mayores de una matriz A^n pueden ser computadas por fórmulas (A1.14) haciendo $f(\lambda) = \lambda^n$.

A2. Definición A2.1: Una matriz polinomio o λ -matriz es una matriz rectangular $A(\lambda)$ cuyos elementos son polinomios en λ , es decir

$$A(\lambda) = \|a_{ik}(\lambda)\| = \|a_{ik}^0 \lambda^0 + a_{ik}^{(1)} \lambda^{(1)} + \dots + a_{ik}^{(l)} \lambda^{(l)}\| \quad \begin{matrix} i = 1, 2, \dots, m \\ k = 1, 2, \dots, n \end{matrix}$$

supóngase que $A(\lambda)$ tenga rango r , es decir la matriz tiene menores de orden r no idénticamente igual a cero. Denótese por $D_j(\lambda)$ el máximo común divisor de todos los menores de orden j en $A(\lambda)$ para $j = 1, 2, \dots, r$, entonces en las series

$$D_r(\lambda), D_{r-1}(\lambda), \dots, D_1(\lambda), D_0(\lambda) \equiv 1$$

cada polinomio es divisible por el que le precede y sean

$$i_1(\lambda) = \frac{D_r(\lambda)}{D_{r-1}(\lambda)}, \quad i_2(\lambda) = \frac{D_{r-1}(\lambda)}{D_{r-2}(\lambda)}, \quad \dots, \quad i_r(\lambda) = \frac{D_1(\lambda)}{D_0(\lambda)} \quad (A2.1)$$

Definición A2.2: Los polinomios $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$ definidos por (e.1) son llamados polinomios invariantes de la matriz rectangular $A(\lambda)$

La matriz polinomio rectangular $A(\lambda)$ es siempre equivalente a una matriz diagonal canónica

$$\begin{bmatrix} i_r(\lambda) & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & i_{r-1}(\lambda) & \dots & 0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & i_1(\lambda) & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & & 0 & 0 & & 0 \end{bmatrix}$$

donde r es el rango de $A(\lambda)$ y $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$ son los polinomios invariantes definidos por (A2.1)

Si los polinomios invariantes son descompuestos en factores irreducibles sobre el campo dado \mathbb{F} , es decir

$$i_1(\lambda) = [\varphi_1(\lambda)]^{c_1} [\varphi_2(\lambda)]^{c_2} \dots [\varphi_s(\lambda)]^{c_s}$$

\vdots

$$i_r(\lambda) = [\varphi_1(\lambda)]^{l_1} [\varphi_2(\lambda)]^{l_2} \dots [\varphi_s(\lambda)]^{l_s}$$

$$c_k \geq \dots \geq l_s \geq 0$$

$$k = 1, 2, \dots, s \quad (A2.2)$$

donde $\varphi_1(\lambda), \dots, \varphi_s(\lambda)$ son factores irreducibles sobre F

Definición A2.3: Todas las potencias entre $[\varphi_1(\lambda)]^{e_1}, \dots, [\varphi_s(\lambda)]^{e_s}$ en (A2.2) hasta donde sean distintos de 1, son llamados los divisores elementales de la matriz $A(\lambda)$ en el campo F .

Si A es una matriz de orden n , entonces su matriz característica es una λ -matriz de rango n . Sus polinomios invariantes son llamados polinomios invariantes de la matriz A y los correspondientes divisores elementales son llamados los divisores elementales de la matriz A .

Supóngase que el campo F contiene no solamente elementos de A , sino también valores característicos de la matriz, entonces los divisores elementales de A tienen la forma

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_u)^{p_u} \quad (A2.3)$$

$$p_1 + p_2 + \dots + p_u = n.$$

Considérese uno de esos divisores elementales

$$(\lambda - \lambda_0)^p$$

y asóciase con la siguiente matriz de orden p

$$\begin{bmatrix} \lambda_0 & 1 & 0 & \dots & 0 \\ 0 & \lambda_0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_0 \end{bmatrix} = \lambda_0 E^{(p)} + H^{(p)} \quad (A2.4)$$

ésta matriz tiene como único divisor elemental a $(\lambda - \lambda_0)^p$. La matriz definida por (A2.4) es llamada el bloque de Jordan correspondiente a el divisor elemental $(\lambda - \lambda_0)^p$.

Los bloques de Jordan correspondientes a los divisores elementales en (A2.3) son denotados por

$$J_1, J_2, \dots, J_u$$

El conjunto de bloques de Jordan J_1, \dots, J_u forman la matriz casi-diagonal

$$J = \{J_1, J_2, \dots, J_u\}$$

esta matriz J puede ser escrita en la forma

$$J = \{ \lambda_1 I_1 + H_1, \lambda_2 I_2 + H_2, \dots, \lambda_u I_u + H_u \}$$

donde $I_k = I^{(p_k)}$, $H_k = H^{(p_k)}$ $k=1, 2, \dots, u$

Ya que las matrices A y J tienen los mismos divisores elementales, ellas son similares, es decir, existe una matriz no singular T tal que $A = T J T^{-1}$.

La matriz J es llamada la forma de Jordan de A , y es caracterizada por ser casi-diagonal y por su estructura especial (A2:4) de los bloques diagonales.

El siguiente sistema es un ejemplo que describe la matriz de Jordan para los divisores elementales $(\lambda - \lambda_1)^2, (\lambda - \lambda_2)^3, (\lambda - \lambda_3), (\lambda - \lambda_4)^2$

$$\begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \lambda_3 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & \lambda_4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_4 \end{bmatrix}$$

Si todos los divisores elementales de una matriz son de primer grado, la matriz de Jordan es una matriz diagonal y en este caso se tiene

$$A = T \{ \lambda_1, \lambda_2, \dots, \lambda_n \} T^{-1}$$

Así, una matriz A se dice que es de estructura simple, si y solo si, todos los divisores elementales son de primer grado.

CONCLUSIONES

Históricamente PERRON (1907) encontró propiedades notables de los valores característicos y vectores característicos para las matrices positivas, más tarde FROBENIUS (1912) generalizó el TEOREMA DE PERRON al investigar las propiedades espectrales de las matrices no negativas irreducibles y considerar a las matrices positivas como un caso especial de las segundas, generalmente a ambos teoremas se les conoce como uno solo denominado el: "TEOREMA DE PERRON-FROBENIUS". Respecto a las demostraciones de dicho teorema expuestas en este trabajo se basan en las demostraciones dadas por SAMUEL KARLIN y GANTMACHER, además de éstas, existen otras pruebas similares dadas por WIELANDT (1950) y DEBREU-HERSTEIN (1953) que junto con las matrices no negativas reducibles y las matrices primitivas e imprimitivas estas últimas como un caso especial de las matrices irreducibles forman la base de las aplicaciones y todos los desarrollos expuestos en éste trabajo.

BIBLIOGRAFIA

- 1.- KARLIN SAMUEL. "Mathematical Methods and Theory in Games, Programming, and Economics"
- 2.- GANTMACHER F. "The Theory of Matrix", vol. I, II
- 3.- VARGA RICHARD. "Matrix Iterative Analysis"
- 4.- BELLMAN RICHARD. "Introduction to Matrix Analysis"
- 5.- SHIELDS PAUL "Linear Algebra"