9

VNIVERSIDAD NACIONAL
AVFNOMA DE
MEXICO

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

ALGUNOS ASPECTOS DE LA TEORIA DE MATRICES POSITIVAS Y CADENAS DE MARKOV.

TESIS

QUE PARA OBTENER EL TITULO DE

MATEMATICO

PRESENTA :

SERGIO ESPINOSA FLORES

MEXICO, D.F. 1991.

FALLA DE ORIGEN

TESIS CON FALLA DE ORIGEN

ALGUNOS ASPECTOS DE LA TEORIA DE MATRICES POSITIVAS Y CADENAS DE MARKOV

CONTENIDO

INTRODUCCION

El presente trabajo consta de dos partes, la primera que esta basada en un ejercicio de la referencia 4 donde se expone un desarrollo de la teoría sobre formas canónicas de Jordan y Racional a partir de una demostración del Teorema de Descomposición en subgrupos cíclicos de un grupo abeliano donde los órdenes de los elementos están acotados. El tratamiento seguido muestra que para un operador $T : V \longrightarrow V$ que tenga polinomio mínimio, V se descompone como suma directa de Subespacios T-cíclicos. Nótese que en la demostración dada de la afirmación anterior no se usa que $\dim(V) < \infty$, sino únicamente la existencia del polinomio mínimo para T. Cuando $\dim(V) < \infty$, obtenemos como consecuencia del Teorema anterior, la existencia de las formas canónicas de Jordan y Racional, mostrando también como calcularlas.

Usamos la teoría sobre la forma canónica de Jordan para demostrar la existencia de límites de potencias de matrices positivas, en particular de matrices estocásticas, caracterizando las matrices A tales que $\lim A^n$ existe.

La segunda parte está basada en la referencia 6 y presenta algunos aspectos sobre la teoría de matrices positivas tales como el lema de Wielandt y algunos criterios para encontrar el límite de matrices en función de sus vectores característicos, enfocando toda esta teoría a las matrices estocásticas citando algunos tipos de ellas como las regulares o las absorbentes.

Posterior a ésto se define a las matrices escindibles, así también se dan algunos criterios para reconocerlas, finalmente se enuncia y se prueba el Teorema de Frobenius, el cual nos clasifica los estados de un cierto experimento o fenómeno que esté representado por una matriz estocástica.

En el presente capitulo hablaré acerca de la forma Canónica de Jordan ya que la necesidad de encontrar límites de matrices a lo largo del trabajo, nos obliga en cierta manera a hacerlo. Empezaremos recordando algunas definiciones básicas y daremos también unos ejemplos que nos serán de gran utilidad , posteriormente daremos una descomposición cíclica de grupos abelianos, haciendo lo análogo para espacios vectoriales. Finalmente daremos la forma para encontrar la matriz de Jordan asociada a una matriz A y demostraremos un teorema acerca de límite de matrices.

DEFINICION-1.0)

Sea $A \in M_{n\times n}(F)$ (F un campo),

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & A_{22} & \dots & A_{2s} \\ \vdots & & \ddots & \vdots \\ A_{s1} & \dots & & A_{ss} \end{pmatrix}$$

es una $(s, (n_1, \dots , n_s), n)$ partición de A si :

i) A_{ii} son matrices cuadradas tal que (suponiendo que A_{ii} es de $n_i \times n_i$) $n_1 + n_2 + \dots + n_s = n$

ii) $A_{kl} \in M_{n_k x n_l}$ $k,l \in \{1,2,\dots,s\}$

DEFINICION-2.0)

Sean A y B matrices cuadradas (n x n) sobre un cierto campo F y sean $P = (i, (n_1, n_2, \dots , n_i), n)$ y $Q = (i, (n_1, n_2, \dots , n_i), n)$ particiones de A y B respectivamente, se define el "*producto partición*" de P y Q (denotado por $\odot$) como $(P \odot Q)_{kl} = (P_k \cdot Q_l)$, donde $\cdot$ es el producto usual de matrices.

PROPOSICION-1.0) $(P \odot Q)$ es una $(i,\ (n_1,\ n_2,\ \ldots\ ,n_i),\ n)$ partición de $A \cdot B$.

Demostración:
Sea $(P \odot Q)_{k,l}$ el elemento k,l de la partición *producto* y sea $(P \odot Q)^{ij}_{kl}$ la entrada i,j del elemento k,l de la partición *producto*, entonces lo que hay que demostrar es que se cumpla la igualdad siguiente :

e.d. P.D. $(P \odot Q)^{ij}_{kl} = (A \cdot B)_{(n_1+\ldots+n_{i-1})+k\ ,\ (n_1+\ldots+n_{j-1})+l}$

$$(P \odot Q)^{ij}_{kl} = (\sum_{s=1}^{m} P^{is}\, Q^{sj})_{kl} = (P^{i1}Q^{1j} + P^{i2}Q^{2j} + \ldots + P^{im}Q^{mj}) \ \ =^{*}_{kl}$$

pero $(P^{i1}\ Q^{1j})_{kl} = (\sum_{k=1}^{n_1} P_{kt}\ Q_{kl}\ )$

$$\therefore\ \ * = \sum A_{(n_1+n_2+\ldots+n_{i-1})+k,\,t}\ \ B_{t,\,(n_1+\ldots+n_{j-1})+l} \qquad + \ldots +$$

$$+ \sum A_{(n_1+\ldots+n_{i-1})+k,\,n_1+t}\ \ B_{n_1+t,\,(n_1+\ldots+n_{j-1})+l}$$

$$= \sum_{t=1}^{n_1+\ldots+n_k} A_{(n_1+\ldots+n_{i-1})+k,\,p}\ \ B_{p,\,(n_1+\ldots+n_{j-1})+l}$$

$$= (A \cdot B)_{(n_1+\ldots+n_2)+k\ ,\ (n_1+\ldots+n_{j-1})+l}$$

■

OBSERVACION-1.0)

Sea C una matriz de kxk y sea

$$C = \begin{pmatrix} X & 0 \\ Y & I \end{pmatrix} \text{ una } (2,\ (n_1,\ n_2), k) \text{ partición de } C$$

entonces

$$C^n = \begin{pmatrix} X^n & 0 \\ Y \sum_{j=0}^{n-1} X^j & I \end{pmatrix}$$

dem. por inducción.
base de inducción
n = 1 es claro

hipótesis de inducción
supongamos ahora que vale para (n - 1) $\quad n \geq 2$

es decir $\quad C^{n-1} = \begin{pmatrix} X^{n-1} & 0 \\ Y \sum_{j=0}^{n-2} X^j & I \end{pmatrix}$

p.d. que vale para (n)

$$C^n = C^{n-1}\, C = \begin{pmatrix} X^{n-1} & 0 \\ Y \sum_{j=0}^{n-2} X^j & I \end{pmatrix} \cdot \begin{pmatrix} X & 0 \\ Y & I \end{pmatrix} = *$$

$$* = \begin{pmatrix} X^{n-1} \cdot X + 0 \cdot Y & X^{n-1} \cdot 0 + 0 \cdot I \\ (Y \sum_{j=0}^{n-2} X^j)\, X + I \cdot Y & (Y \sum_{j=0}^{n-2} X^j) \cdot 0 + I \cdot I \end{pmatrix}$$

$$* = \begin{pmatrix} X^n & 0 \\ Y \sum_{j=0}^{n-1} X^j & I \end{pmatrix}$$

DEFINICION-3.0) Un conjunto $R \neq \emptyset$ se dice que es un grupo si está definida una operación binaria $* : R \times R \longrightarrow R$ denotada por $*$ tal que:

1) $a, b \in R \Rightarrow a * b \in R$ es cerrado bajo $*$

2) $a, b, c \in R \Rightarrow a * (b * c) = (a * b) * c$ (cumple asociatividad).

3) $\exists\ e \in R$ tal que $a * e = e * a = a$ para toda $a \in R$ (elemento identidad)

4) Para toda $a \in R\ \exists\ a^{-1} \in R$ tal que $a * a^{-1} = a^{-1} * a = e$ (existencia de inversos)

DEFINICION-4.0)Decimos que un grupo R que es abeliano (conmutativo)
si además para todas $a, b \in R$ se tiene $a * b = b * a$

DEFINICION-5.0) $(R, +, \cdot)$ es anillo si $(R, +)$ es grupo abeliano y $\cdot : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ cumple:

1) Para toda $a, b, c \in R$ $a \circ (b \circ c) = (a \circ b) \circ c$

2) $a \circ (b + c) = a \circ b + a \circ c$ y $(b + c)a = b \circ a + c \circ a$

DEFINICION-6.0) $(R, +, \circ)$ es un anillo con elemento unitario si $\exists$ $1 \neq 0 \in R$ tal que $a \circ 1 = 1 \circ a = a$

Ejemplos:

$(\mathbb{Z}, +, \circ)$, $(\mathbb{Z}n, +, \circ)$, $Hom_F(V, V)$ con V_F espacio vectorial. $M_{nxn}(F)$, $F[x]$, $\mathbb{Z}[x]$ son anillos.

DEFINICION-7.0) Sea R un anillo y sea M un conjunto distinto del vacio, decimos que M es un R-módulo izquierdo si:

1) $(M, +)$ es un grupo abeliano

2) $\exists \circ : R \times M \longrightarrow M$ tal que:

a) $1_R \circ m = m$ para toda $m \in M$

b) $r_1 (r_2 \circ m) = (r_1 \circ r_2) \circ m$ para toda $r_1, r_2 \in R$ y para toda $m \in M$

c) $(r_1 + r_2)m = r_1m + r_2m$ para toda $r_1, r_2 \in R$ y para toda $m \in M$

d) $r(m_1 + m_2) = rm_1 + rm_2$ para toda $r \in R$ y para toda m_1, $m_2 \in M$

Hemos dicho R - módulo izquierdo pues, pues sólo dejamos multiplicar por el lado izquierdo, de manera análoga podemos definir R - módulo derecho o simplemente R - módulo donde se vale multiplicar por ambos lados, notemos además que si R es un campo entonces los R módulos serán los espacios vectoriales.

Demos algunos ejemplos de R - módulos izquierdos:

Sea G un grupo abeliano y tomemos el anillo de los enteros definiendo $\cdot : \mathbb{Z} \times G \longrightarrow G$ de la siguiente manera:

i) $0 \cdot g = 0_G$ con $0 \in \mathbb{Z}$ y $g, 0_G \in G$

ii) Para toda $n \in \mathbb{Z}^+$ supongamos que ng está definido y definamos $(n + 1)\cdot g = n\cdot g + g$ y si $n \notin \mathbb{N} \Rightarrow (-n)g = -(ng)$

Con nuestra definición anterior de la operación producto, G es un $\mathbb{Z}$-módulo.

Demostración:

a) P.D. Para toda $g \in G$ $1 \circ g = g$

$1 \circ g = (0 + 1)g = 0 \circ g + g = 0_G + g = g$

b) P.D. $n_1(n_2g) = (n_1n_2)g$ Para toda $n_1, n_2 \in \mathbb{Z}$ y para toda $g \in G$

La demostración se hará por inducción sobre n_1

i) $n_1 = 0$

$0 \circ (n_2g) = 0_G$ y $(0 \circ n_2)g = 0 \circ g = 0_G$

ii) Supongamos cumple para n_1 e.d. $n_1(n_2g) = (n_1n_2)g$

iii) P.D. cumple para $n_1 + 1$ e.d. $(n_1 + 1)(n_2g) = ((n_1 + 1)\circ n_2)g$

$(n_1 + 1)(n_2g) = n_1(n_2g) + n_2g \overset{(2)}{=}$

$$(n_1n_2)g + n_2g = (n_1 \underset{\mathbb{Z}}{\circ} n_2 + n_2)g = [(n_1+1)n_2]g$$

∎

con lo cual queda demostrado para todo entero positivo. Falta ver que cumple para $\mathbb{Z}^-$.

Si $n \in \mathbb{N} \Rightarrow -n \in \mathbb{Z}^-$ P.D. $(-n_1)(n_2g) = (-n_1n_2)g$

$(-n_1)(n_2g) = -(n_1)(n_2)g = -(n_1n_2)g = (-n_1n_2)g$

c) P.D. $(n_1+n_2)g = n_1g + n_2g$ para toda $n_1, n_2 \in \mathbb{Z}$ y para toda $g \in G$. Nuevamente hagamos la demostración por inducción sobre n_1.

i) $n_1 = 0$ P.D. $(0 + n_2)g = 0g + n_2g$

$(0 + n_2)g = (n_2)g$ y $0g + n_2g =$

$0_G + n_2g = n_2g \therefore (0 + n_2)g =$

$0g + n_2g$

ii) Suponemos vale para n_1 e.d. $(n_1+n_2)g = n_1g + n_2g$

iii) P.D. vale para n_1+1 e.d. $[(n_1 + 1) \underset{\mathbb{Z}}{+} n_2]g =$

$(n_1+1)g + n_2g$ pero $((n_1+1) + n_2)g =$

$((n_1+n_2) + 1)g = (n_1+n_2)g + g = n_1g + n_2g + g =$

$n_1g + g + n_2g = (n_1+1)g + n_2g$

∎

Con lo cual queda demostrado para $\mathbb{N}$

Veamos que cumple con $\mathbb{Z}^-$, si $n_1 \in \mathbb{N} \Rightarrow -n_1 \in \mathbb{Z}^-$

P.D. $(-n_1+n_2)g = -n_1g + n_2g$.

$(-n_1+n_2)g = -(n_1-n_2)g = -(n_1+(-n_2))g =$

$-[(n_1+(-n_2))g] = -[n_1g + (-n_2)g = -n_1g + n_2g$

∎

EJEMPLO (2)

Definamos el producto de un vector por un polinomio de la siguiente manera :

Sea $T : V_F \longrightarrow V_F$ un operador lineal ($\dim V = n$) y sea $p \in F[t]$ y $v \in V_F$ entonces definimos el producto de p y v como $p\cdot v = p(T)(v)$.

Entonces V es un F[t]-módulo bajo el producto que antes definimos pues es claro que, V cumple las condiciones de grupo abeliano bajo la suma y ademas :

a) $1_p\cdot v = 1_p(T)(v) = I(v) = v \quad \forall\, v \in V$, con 1_p el polinomio constante 1.

b) $(P_1P_2)\cdot(v) = P_1((P_2)\cdot(v))$ con $P_i \in F[t]$ y $v \in V$.

Antes de probar esto mencionemos el siguiente teorema

Teorema-0.0)

Sean P_1 y $P_2 \in F[t]$ y sea $T : V_F \longrightarrow V_F$ un operador lineal ($\dim V = n$) entonces $(P_1P_2)(T) = P_1(T)\circ P_2(T)$

Prueba:

Lema $\forall\, m \in \mathbb{N}$ y $\forall\, g \in F[t]$ se tiene que $(t^m g)(T) = T^m \circ g(T)$

Probemos esto por inducción sobre m

i)base inducción

$m = 0$ tenemos que $t^0 = 1$

por lo que $(1g)(T) = (g)(T) = I\circ(g)(T) = 1(T)\circ g(T)$

Supongamos $g(t) = \sum\limits_{i=0}^{n} b_i t^i$

$$\text{entonces} \quad (t\, g(t))(T) = \left(\sum_{i=0}^{n} b_i t^{i+1}\right)(T) = \left(\sum_{i=1}^{n} b_i T^{i+1}\right) =$$

$$= T\left(\sum_{i=0}^{n} b_i T^i\right) = T\circ g(T) = t(T)\circ g(T)$$

ii) paso inductivo

Supongamos $(t^m g)(T) = T^m \circ g(T) \quad \forall\ g(t) \in F[t]$

Demostremos que $(t^{m+1}g)(T) = t^{m+1}(T)\circ g(T)$

$$(t^{m+1}g)(T) = (tt^m g)(T)$$
$$T\circ[(t^m g)(T)] = T\circ(T^m\circ g(T))$$
$$T^{m+1}\circ g(T) = t^{m+1}(T)\circ g(T) \qquad \blacksquare$$

Demostración del teorema:

Por inducción sobre el grado de p_1

i) base inducción

Denotemos por $\delta(p_1)$ el grado de p_1

si $\delta(p_1) = 0$ entonces se tiene que $p_1 = a \in F$ de donde por el lema anterior

$$(p_1p_2)(T) = (ap_2)(T) = ap_2(T) = aI \circ p_2(T) = a(I) \circ p_2(T)$$
$$= p_1(T) \circ p_2(T).$$

ii) hipótesis de inducción

Supongamos que $\delta(p_1) > 0$, p_1 puede ser expresado de la siguiente forma :

$$p_1 = tq_1 + r \text{ donde } r = 0 \text{ ó } \delta(r) < 1$$

e.d. $p_1 = tq_1 + r$ con $r \in F$ de hecho por el teorema de residuo $r = p_1(0)$.

$$\begin{aligned}(p_1p_2)(T) &= ((tq_1 + r)(p_2))(T)\\
&= (tq_1p_2 + rp_2)(T)\\
&= (tq_1p_2)(T) + (rp_2)(T)\\
&= T \circ ((q_1p_2)(T)) + rI \circ p_2(T) \quad \text{hipótesis de inducción}\\
&= T \circ (q_1(T) \circ p_2(T)) + rT \circ p_2(T)\\
&= (T \circ q_1(T) + r(T)) \circ p_2(T)\\
&= ((tq_1)(T) + r(T)) \circ p_2(T)\\
&= (tq_1 + r)(T) \circ p_2(T)\\
&= p_1(T) \circ p_2(T)\end{aligned}$$

$\therefore$ por el teorema anterior $(p_1p_2)v = (p_1p_2)(T)(v)$

$$= p_1(T) \circ p_2(T)v = p_1 \circ (p_2(T))v) = p_1 \circ (p_2v) = p_1(p_2v)$$

■

c) $(P_1 + P_2) \cdot v = P_1 \cdot v + P_2 \cdot v$:

$$\begin{aligned}(P_1 + P_2) \cdot v &= [(P_1 + P_2)(T)](v) = (P_1(T) + P_2(T))(v)\\
&= P_1(T)(v) + P_2(T)(v)\\
&= = P_1 \cdot v + P_2 \cdot v\end{aligned}$$

d) $P \cdot (v + w) = (P \cdot v) + (P \cdot w)$

$$\begin{aligned}P \cdot (v + w) &= P(T)(v + w)\\
&= P(T)(v) + P(T)(w)\\
&= (P \cdot v) + (P \cdot w)\end{aligned}$$

■

Definamos ahora lo que es un submódulo de algún R-módulo M.

DEFINICION-8.0) Sea M un R-módulo y sea $N \subseteq M$, $N \neq \emptyset$
N es un submódulo de M $\Longleftrightarrow$

i) $a + b \in N$ para toda $a, b \in N$ (cerrado)

ii) $r\, a \in N$ para toda $r \in R$ y $a \in N$

Y ahora definamos lo que es un ideal

DEFINICION-9.0) Un ideal izquierdo de R es un submódulo de ${}_R R$

EJEMPLO. Si $A = \{0\}$, entonces A es un ideal izquierdo de ${}_{(0)}\mathbb{Z}$.

Demostración:
i) $A \neq \emptyset$
ii) Para toda $a \in A$ $a + a \in A$ pues el único elemento de A es 0 y $0 + 0 = 0 \in A$

iii) $r \circ a = 0 \in A$ para toda $r \in \mathbb{Z}$ y para toda $a \in A$

■

Sea R un anillo conmutativo e $I \leq R$

DEFINICION-10.0) $r_1 \equiv r_2 \bmod I$ si $r_1 - r_2 \in I$ esta relación es de equivalencia
Demostración:
i) Reflexividad
$0 = r - r \in I \therefore r \equiv r \bmod I,\ \forall\, r \in R.$

ii) Simetría

$r_1 \equiv r_2 \bmod I \Rightarrow r_2 \equiv r_1 \bmod I$

$r_1 \equiv r_2 \bmod I \Rightarrow r_1 - r_2 \in I \Rightarrow -(r_1 - r_2) \in I$

$\Rightarrow r_2 - r_1 \in I \Rightarrow r_2 \equiv r_1 \bmod I$

{* Notación $N \leq M$ nos denotará un submódulo de M y se lee N es submódulo de M *}

iii) Transitividad

si $r_1 \equiv r_2 \bmod I$ y $r_2 \equiv r_3 \bmod I$ entonces $r_1 \equiv r_3 \bmod I$

$r_1 - r_2 \in I$ y $r_2 - r_3 \in I$ entonces $(r_1 - r_2) + (r_2 - r_3) \in I$

$\therefore\ r_1 - r_3 \in I\ \therefore\ r_1 \equiv r_2 \bmod I$

$\therefore$ esta relación es de equivalencia.

∎

$$
\begin{aligned}
[r] &= \{s \in R \mid r \equiv s \bmod I\} \\
&= \{s \in R \mid r - s \in I\} \\
&= \{s \in R \mid s - r \in I\} \\
&= \{s \in R \mid s - r = i,\ \text{p.a. } i \in I\} \\
&= \{s \in R \mid s = r + i,\ i \in I\} = r + I.
\end{aligned}
$$

$R/I = \{\ [r] \mid r \in R\ \}$

Definanse dos operaciones binarias (una llamada suma y otra producto) como sigue:

$+ : R/I \times R/I \longrightarrow R/I$

tal que $(r + I) + (s + I) = (r + s) + I$

$\odot : R/I \times R/I \longrightarrow R/I$

tal que $(r + I) \odot (s + I) = (r * s) + I$

Hay que ver que estas dos operaciones estan bien definidas.

Probemos primero que la + esta bien definida, esto equivale a probar que si $[r] = [r']$ y $[s] = [s']$ entonces

$[r] + [s] = [r'] + [s']$

Supongamos que $[r] + [s] \neq [r'] + [s']$

por definición $[r] + [s] = [r + s]$

y $[r'] + [s'] = [r' + s']$

luego si $[r + s] \neq [r' + s']$

entonces $(r + s) - (r' + s') \notin I$

$\therefore$ $(r - r') + (s - s') \notin I$ lo es una contradicción pues $[r] = [r']$ y $[s] = [s']$ e I es un ideal de R.

$\therefore [r] + [s] = [r'] + [s']$.

Probemos ahora que $\odot$ esta bien definida, tomemos de nuevo las siguientes clases $[r] = [r']$ y $[s] = [s']$ y probemos que

$[r] \odot [s] = [r'] \odot [s']$

sabemos que $(r-r') \in I$ y $(s-s') \in I$

$\therefore$ $(r-r')s \in I$ y $(s-s')r' \in I$

$\therefore$ $(rs-r's) \in I$ y $(sr'-s'r') \in I$

tenemos que $(rs-r's) + (sr'-s'r') =$

$(rs - s'r') \in I$

$\therefore rs \equiv s'r' \bmod I$

de donde $[rs] = [r's']$

$\therefore$ $[r]\odot[s] = [r']\odot[s']$

<u>AFIRMACION-1.0</u>)

$(R/I, +, \odot)$ es un anillo

Demostración

1)P.D. $(R/I,+)$ forma un grupo abeliano

a) Sea $[r_1] \in R/I$ y $[r_2] \in R/I$

p.d. $[r_1] + [r_2] \in R/I$

sea $s_1 \in [r_1]$ y $s_2 \in [r_2]$

entonces $s_1 = r_1 + \imath$ y $s_2 = r_2 + \imath$

por lo que $s_1 + s_2 = (r_1 + \imath) + (r_2 + \imath)$

$= (r_1 + r_2) + \imath \in R/I$

entonces $s_3 = r_3 + \imath$

$s_3 - r_3 = \imath$

$s_3 - r_3 \in I$

$r_3 - s_3 \in I$

$\therefore$ $r_3 \equiv s_3 \bmod I$

$\therefore$ $[r_3] \in R/I$

b) P.D. $s_1 + (s_2 + s_3) = (s_1 + s_2) + s_3$

con $s_1 \in [r_1]$ $s_2 \in [r_2]$ $s_3 \in [r_3]$

$s_1 = r_1 + \imath$ $s_2 = r_2 + \imath$ $s_3 = r_3 + \imath$

$$
\begin{aligned}
[s_1 + (s_2 + s_3)] &= (r_1 + \imath) + [(r_2 + \imath) + (r_3 + \imath)] \\
&= (r_1 + \imath) + [(r_2 + r_3) + \imath)] \\
&= [(r_1 + (r_2 + r_3)) + \imath] \\
&= [((r_1 + r_2) + r_3) + \imath]
\end{aligned}
$$

$= [(r_1 + r_2) + \imath] + (r_3 + \imath)$
$= [(r_1 + \imath) + (r_2 + \imath)] + (r_3 + \imath)$
$= (s_1 + s_2) + s_3$

∎

c) Existencia del elemento neutro aditivo:

Como R es anillo $\exists\ 0 \in R$ tal que $0 + r = r,\ \forall\ r \in R$

p.d. $\exists\ [r_0]$ tal que $[r_0] + [r_\imath] = [r_\imath] \quad \forall\ \imath$

sea $s_0 \in [r_0] \qquad s_\imath \in [r_\imath]$

entonces $s_0 = 0 + \imath \qquad s_\imath = r_\imath + \imath$

entonces $s_0 + s_\imath = (0 + \imath) + (r_\imath + \imath)$
$= (0 + r_\imath) + \imath$
$= (r_\imath) + \imath$
$= r_\imath + \imath$
$= S_\imath \qquad \forall\ \imath \quad \therefore [0] = [r_0]$

∎

d) $\forall\ [r] \in R/I \qquad \exists\ [r'] \in R/I$

tal que $[r] + [r'] = [0]$

Sea $s \in [r]$ entonces $s = r + \imath$

sea $s' \in [r']$ tal que $s + s' = 0$ con $0 \in [0]$

$s = r + \imath$ y $s' = r' + \imath$

entonces $(r + \imath) + (r' + \imath) = [r + r'] + \imath$

$[r + r'] + \imath = [0] + \imath$ entonces

$r + r' = 0$

$r = -r'$

$\therefore s' = -r + \imath$

e) P.D. $[r] + [r'] = [r'] + [r] \quad \forall\ [r],\ [r']$ en R/I

Sean s y s' en $[r]$ y $[r']$ respectivamente

entonces $s = r + \imath$ y $s' = r' + \imath$

$s + s' = (r + \imath) + (r' + \imath)$
$= (r + r') + \imath$
$= (r' + r) + \imath$

$= (r' + \imath) + (r + \imath)$

$= s' + s$

$\therefore\ [r] + [r'] = [r'] + [r] \quad \forall\ [r],\ [r']$ en R/I

■

Con todo lo anterior hemos demostrado que (R/I, +) es un grupo conmutativo, falta ahora demostrar las propiedades del producto para demostrar que R/I es un anillo.

f) P.D. que $\forall$ s , s' en [r] y [r'] respectivamente se tiene que $s \odot s' \in R/I$

Sean s y s' en [r] y [r'] respectivamente

entonces se tiene que $s = r + \imath$ y $s' = r' + \imath$ po lo que

$s \odot s' = (r + \imath) \odot (r' + \imath)$

$= (r * r') + \imath \in R/I$

$\therefore\ s \odot s' \in R/I$

g) P.D. que $(s \odot s') \odot s'' = s \odot (s' \odot s'')$

$\forall$ s, s', s'' en [r], [r'], [r''] respectivamente

entonces se tiene que

$s = r + \imath,\ s' = r' + \imath,\ s'' = r'' + \imath$

por lo que

$$
\begin{aligned}
(s \odot s') \odot s'' &= [(r + \imath) \odot (r' + \imath)] \odot (r'' + \imath)\\
&= ((r * r') + \imath) \odot (r'' + \imath)\\
&= ((r * r') * r'') + \imath\\
&= (r * (r' * r'')) + \imath\\
&= (r + \imath) \odot ((r' * r'') + \imath)\\
&= (r + \imath) \odot [(r' + \imath) \odot (r'' + \imath)]\\
&= s \odot (s' \odot s'')
\end{aligned}
$$

con lo que concluimos que

$(R/I\ ,\ \odot)$ es asociativo.

■

h) P.D. $s \odot (s' + s'') = s \odot s' + s \odot s''$

$\forall$ s, s', s'' en [r], [r'], [r''] respectivamente

sean $s = r + \imath,\ s' = r' + \imath,\ s'' = r'' + \imath$

entonces $s \odot (s' + s'') = (r + \imath) \odot [(r' + \imath)+(r'' + \imath)]$

$$= (r + i) \odot [(r' + r'') + i)]$$
$$= (r^*(r' + r'')) + i$$
$$= ((r^*r') + (r^*r'')) + i$$
$$= ((r^*r') + i) \pm ((r^*r'') + i)$$
$$= ((r + i) \odot (r' + i)) + ((r + i) \odot (r'' + i))$$
$$= s \odot s' + s \odot s''$$

Con lo cual queda demostrado que $(R/I, +, \odot)$ forma una anillo, cuando R es conmutativo.

Se puede demostrar que si I es un ideal bilateral de R entonces R/I es un anillo.

■

Sean R y R' anillos y sea $\Psi : R \longrightarrow R'$ entonces decimos que Ψ es un morfismo de anillos si:

i) $\Psi(ab) = \Psi(a)\Psi(b)$

ii) $\Psi(a + b) = \Psi(a) + \Psi(b)$

<u>OBSERVACION-2.0</u>)

El núcleo de Ψ es un ideal bilateral de R, donde Núcleo de $\Psi = N = \{ x \in R \mid \Psi(x) = o \}$

Demostración

i) P.D. que $\forall\ x, y \in N \quad (x + y) \in N$

$$\Psi(x + y) = \Psi(x) + \Psi(y) = 0 + 0 = 0$$
$$\therefore (x + y) \in N$$

ii) P.D. $(ra) \in N,\ \forall\ r \in R$ y $a \in N$

$$\Psi(ra) = \Psi(r)\Psi(a) = \Psi(r)0 = 0$$

$\therefore$ $(ra) \in N$ con r, a en R y N respectivamente, analogamente podemos hacer ésto por el lado derecho de donde concluimos que N es ideal bilateral de R.

■

Sea G un grupo y $g \in G$, denotaremos por $O(G)$ el orden de G, i.e. el número de elementos de G y $O(g)$ al orden de g.

<u>TEOREMA-1.0</u>)Sea G un p-grupo abeliano tal que $\{O(x) \mid x \in G\}$ esta acotado y sea m el menor natural tal que $p^m G = 0$, entonces G es suma directa de subgrupos cíclicos .

Demostración:
La demostración la haremos por inducción sobre m.

i) base de la inducción ($m=1$).

Si $pG = 0$ entonces G es un Espacio Vectorial sobre el campo $\mathbb{Z}_p$ y entonces existe una base β de G tal que $G = \bigoplus_{\alpha \in \beta} \mathbb{Z}_p x_\alpha$ y como $\mathbb{Z}_p x_\alpha = \mathbb{Z} x_\alpha$ entonces $G = \bigoplus_{\alpha \in \beta} \mathbb{Z} x_\alpha$.

ii) supongamos que $m > 1$.

sabemos que $pG \leq G$, ahora bien $pG \neq 0$ puesto que $m > 1$ y p^m se escogió como mínimo con la propiedad de que $p^m G = 0$. Ahora,
$p^m G = p^{m-1}(pG) = 0$ y por hipótesis de inducción podemos tomar
$pG = \oplus \mathbb{Z} x_\alpha$

veamos el siguiente diagrama:

$$\begin{tikzcd}
\text{p-zoc}(G) \arrow[r, hook] & G & \sum \mathbb{Z} y_\alpha \arrow[l, hook'] \\
\text{p-zoc}(pG) \arrow[r, hook] \arrow[u, hook] & pG = \oplus\, \mathbb{Z} x_\alpha \arrow[u, hook] &
\end{tikzcd} \qquad (*)$$

Tenemos que:

$$\text{p-zoc}(G) = \text{p-zoc}(pG) \oplus W \hookrightarrow G \hookleftarrow \sum \mathbb{Z} y_\alpha \qquad (✱)$$

hacemos las siguientes afirmaciones:

(✱) $\text{p-zoc}(G) = \{ x \in G \mid px = 0 \}$
e.d. p-zoc(G) es un espacio vectorial sobre $\mathbb{Z}_p$, ∴ cualquier subgrupo de p-zoc(G) es un sumando directo de p-zoc(G) puesto que todo subespacio de un espacio vectorial es un sumando directo ∴ $\text{p-zoc}(G) = \text{p-zoc}(pG) \oplus W$.
(*) Aquí $y_\alpha \in G$ se escoge de manera que $py_\alpha = x_\alpha \in pG$.

Afirmación 1.- $\sum \mathbb{Z}y_\alpha = \oplus\mathbb{Z}y_\alpha$

Afirmación 2.- $W \cap (\oplus \mathbb{Z}y_\alpha) = \{0\}$

Afirmación 3.- ${}_{\mathbb{Z}p}W \oplus (\oplus\mathbb{Z}y_\alpha) = G$

LEMA-1.0)Sea $\mathbb{Z}y_\alpha$ tal que $p^k y_\alpha = 0$ para alguna k (tomemos k mínima con esta propiedad) entonces el p-zoc$(\mathbb{Z}y_\alpha) = \mathbb{Z}z \ \forall\ z \neq 0$ con $z \in$ p-zoc$(\mathbb{Z}y_\alpha)$.

Demostración

Sea $z \neq 0$ con $z \in$ p-zoc$(\mathbb{Z}y_\alpha)$, entonces $\mathbb{Z}z \subset$ p-zoc$(\mathbb{Z}y_\alpha)$, pues si $x \in \mathbb{Z}z$ entonces $x = az$ con $a \in \mathbb{Z}$, pero $z \in$ p-zoc$(\mathbb{Z}y_\alpha)$

$\Rightarrow p(az) = a(pz) = a0 = 0$

$\Rightarrow x \in$ p-zoc$(\mathbb{Z}y_\alpha)$.

Demostremos ahora la otra inclusión.

Sea $0 \neq z \in$ p-zoc$(\mathbb{Z}y_\alpha)$ entonces $z = ry_\alpha$ pero $pz = 0$ y $pz = pry_\alpha$, sea $I = \{ s \in \mathbb{Z} \mid sy_\alpha = 0 \}$ entonces I es un subgrupo de $\mathbb{Z}$, de donde $I = a\mathbb{Z}$ con $a \in \mathbb{N}$. Note $p^k \in a\mathbb{Z}$. Como p^k es mínimo tal que $p^k y_\alpha = 0$ entonces $p^k \leq a$, pero si $p^k \in a\mathbb{Z}$ entonces $p^k = am$ de donde se ve que $a \leq p^k$, de donde $p^k = a$ y $\therefore a\mathbb{Z} = I = p^k\mathbb{Z}$. Pero $pr \in I = p^k\mathbb{Z}$ $\therefore pr = p^k b$, p.a. $b \in \mathbb{Z}$ y $\therefore r = p^{k-1}b$. Si multiplicamos la igualdad anterior por y_α tenemos que $ry_\alpha = p^{k-1}by_\alpha$ y por lo tanto $z = p^{k-1}by_\alpha$, de donde se ve que b no es dividido por p , pues si b fuera dividido por p entonces $b = pb_1$ y $z = p^{k-1}pb_1y_\alpha$ y así tenemos que $z = b_1p^ky_\alpha = 0$, lo cual es una contradicción.

Ahora bien si $z = p^{k-1}by_\alpha$ y p no divide a b entonces tenemos que $(p,b) = 1$

En resumen, si $z \in$ p-zoc$(\mathbb{Z}y_\alpha)$ entonces $z = b(p^{k-1}y_\alpha)$, con $(b,p)=1$ es decir p-zoc$(\mathbb{Z}y_\alpha) \subseteq \mathbb{Z}(p^{k-1}y_\alpha)$.

Como $0 \neq p^{k-1}y_\alpha \in$ p-zoc$(\mathbb{Z}y_\alpha)$ entonces se tiene que p-zoc$(\mathbb{Z}y_\alpha) = \mathbb{Z}(p^{k-1}y_\alpha)$.

Ahora hay que ver que $\mathbb{Z}(p^{k-1}y_\alpha) \subseteq \mathbb{Z}z$ tal que $0 \neq z \in$ p-zoc$(\mathbb{Z}y_\alpha)$.

Como hemos visto $z = bp^{k-1}y_\alpha$ con $(p,b) = 1$ entonces $(p^k, b) = 1$

$\therefore 1 = \gamma p + \beta b$ con $\gamma, \beta \in \mathbb{Z}$

$$\therefore\ p^{k-1}y_\alpha = p^{k-1}y_\alpha 1 = p^{k-1}y_\alpha\gamma p + p^{k-1}y_\alpha\beta b = p^ky_\alpha\gamma + p^{k-1}y_\alpha\beta b$$
$$= \beta bp^{k-1}y_\alpha = \beta z.$$

$\therefore\ \mathbb{Z}(p^{k-1}y_\alpha) \subseteq \mathbb{Z}z$

$\therefore$ p-zoc$(\mathbb{Z}y_\alpha) = \mathbb{Z}p^{k-1}y_\alpha = \mathbb{Z}z \quad \forall\ 0 \neq z \in$ p-zoc$(\mathbb{Z}y_\alpha)$.

■

Demostración de la afirmación 1.-

P.D. $\sum \mathbb{Z}y_\alpha = \oplus\mathbb{Z}y_\alpha$

Sea $0 = r_1y_{\alpha 1} + \dots + r_ny_{\alpha n}$ una combinación lineal de elementos de $\sum \mathbb{Z}y_\alpha$ con $y_{\alpha i} \in \mathbb{Z}y_{\alpha i}$, hay que ver $r_iy_{\alpha i} = 0\ \forall\ i \in \{1,2,\dots n\}$

Multiplicando por p tenemos que:

$0 = p0 = r_1py_{\alpha 1} + \dots + r_npy_{\alpha n}$ pero $py_{\alpha i} = x_{\alpha i}\ \ \forall\ i$ luego entonces

$$p0 = r_1x_{\alpha 1} + \dots + r_nx_{\alpha n} \in \oplus\ \mathbb{Z}x_{\alpha n}$$

lo cual implica que $r_ix_{\alpha i} = 0\ \ \forall\ i \in \{1,2,\dots n\}$

pero $\qquad r_ix_{\alpha i} = pr_iy_{\alpha i} = 0$

por lo que $r_iy_{\alpha i} \in$ p-zoc$(\mathbb{Z}y_{\alpha i}) =$ p-zoc$(\mathbb{Z}x_{\alpha i}) \subseteq \mathbb{Z}x_{\alpha i}$

por la forma como se escogió $py_{\alpha i} = x_{\alpha i}$ se tiene que $\mathbb{Z}x_{\alpha i} \subseteq \mathbb{Z}y_{\alpha i}$

pues por el lema p-zoc$(\mathbb{Z}y_{\alpha i}) =$ p-zoc$(\mathbb{Z}x_{\alpha i})$

$$\|$$

$$\mathbb{Z}z \quad \forall\ z \in \text{p-zoc}(\mathbb{Z}x_{\alpha i})$$

$$\therefore \qquad 0 = r_1y_{\alpha 1} + \dots + r_ny_{\alpha n} \in \oplus\mathbb{Z}x_\alpha$$

con $r_iy_{\alpha i} \in \mathbb{Z}x_{\alpha i}$

y $\therefore\ r_1y_{\alpha 1} = \dots = r_ny_{\alpha n} = 0$

■

Demostración de la afirmación 2.

Sea $w = r_1y_{\alpha 1} + \dots + r_ny_{\alpha n} \in W \cap (\mathbb{Z}y_\alpha)$,

por lo tanto $w \in$ p-zoc(G) así que $pw = 0$ y recordemos que

$py_{\alpha i} = x_{\alpha i}$ por lo que $0 = r_1x_{\alpha 1} + \dots + r_nx_{\alpha n} \in \oplus\mathbb{Z}x_{\alpha i}$

en consecuencia $r_1x_{\alpha 1} = \dots = r_nx_{\alpha n} = 0$ de donde tenemos que

$r_iy_{\alpha i} \in$ p-zoc$(\mathbb{Z}y_{\alpha i}) =$ p-zoc$(\mathbb{Z}x_{\alpha i}) = \mathbb{Z}x_{\alpha i} \subset pG$

$\therefore\ w \in$ p-zoc$(pG) \cap W = \{0\}$

■

Demostración de la afirmación 3.

Sea $g \in G$ entonces $pg \in pG = \oplus \mathbb{Z}x_\alpha$

$\therefore pg = r_1 x_{\alpha 1} + \ldots + r_n x_{\alpha n}$ y consideremos

$z = r_1 y_{\alpha 1} + \ldots + r_n y_{\alpha n}$ entonces tenemos que $pz = pg$ y esto es por ser $py_{\alpha i} = x_{\alpha i}$ de donde vemos que $p(z-g) = 0$ y $\therefore$ $(z-g) \in p\text{-zoc}(G)$ y $(g-z) \in p\text{-zoc}(G) = p\text{-zoc}(pG) + W$

sea $g-z = z' + w$ con $z' \in p\text{-zoc}(pG)$.

Ahora $p\text{-zoc}(pG) = p\text{-zoc}(\oplus \mathbb{Z}x_\alpha) = \oplus p\text{-zoc}(\mathbb{Z}x_\alpha) = \oplus p\text{-zoc}(\mathbb{Z}y_\alpha)$ (*)

$= p\text{-zoc}(\oplus \mathbb{Z}y_\alpha) \subseteq \oplus \mathbb{Z}y_\alpha$

$\therefore z' \in \oplus \mathbb{Z}y_\alpha$

$\therefore z' + z \in \oplus \mathbb{Z}y_\alpha$

$\therefore g = (z + z') + w \in (\oplus \mathbb{Z}y_\alpha + W)$

■

(*) Observación $p\text{-zoc}(\oplus \mathbb{Z}x_\alpha) = \oplus p\text{-zoc}(\mathbb{Z}x_\alpha)$

Prueba: sea $z \in p\text{-zoc}(\oplus \mathbb{Z}x_\alpha)$, $z = m_1 x_{\alpha 1} + \ldots + m_n x_{\alpha n}$ $m_i \in \mathbb{Z}$

entonces $pz = pm_1 x_{\alpha 1} + \ldots + pm_n x_{\alpha n} = 0$ pero

$0 = pm_1 x_{\alpha 1} + \ldots + pm_n x_{\alpha n} \in \oplus \mathbb{Z}x_\alpha$

$$\Leftrightarrow \quad p(m_i x_{\alpha i}) = 0$$

$$\Leftrightarrow \quad m_1 x_{\alpha 1} + \ldots + m_n x_{\alpha n} \in \oplus p\text{-zoc}(\mathbb{Z}x_\alpha)$$

■

TEOREMA DE DESCOMPOSICION PARA ESPACIOS VECTORIALES

Hemos visto que si G es un grupo abeliano entonces podemos descomponerlo como una suma directa de subgrupos ciclicos. Lo que ahora haremos será una analogía para espacios vectoriales, veremos que si V es un espacio vectorial entonces lo podemos descomponer como una suma directa de subespacios T-ciclicos, pero primero veamos el siguiente teorema.

<u>TEOREMA-2.0</u>) Sea $T : V_F \longrightarrow V_F$ un operador lineal tal que la $\dim(V_F) < \infty$ entonces $\exists\ \mu(T,t) \in F[t]$ tal que
$\mu(T,t)F[t] = \{ f \in F[t] \mid f(T) = T_0 \}$

Demostración

Como la dimensión de V_F es finita podemos entonces encontrarle una base, sea $\{v_1, v_2, \ldots, v_n\}$ una base ordenada de V_F podemos tomar el siguiente conjunto $\{v_1, T(v_1), \ldots T^{k_1}(v_1)\}$ de manera que sea linealmente dependiente mínimo p.a. k_1

$\therefore\ T^{k_1}(v_1) = -a_{k-1}T^{k_1-1}(v_1) - \ldots - a_0 v_1$

$\therefore\ p_1(t) = t^{k_1}(v_1) + a_{k-1}t^{k_1-1}(v_1) + a_0 \quad \in F[t]$ es tal que me anula a T en v_1 e.d $p_1(T)(v_1) = 0$

análogamente $\exists\ p_i \in F[t]$ tal que $p_i(T)(v_i) = 0$ por lo tanto

$p_1(T)\circ p_2(T)\circ \ldots \circ p_n(T)(v) = 0 \quad \forall\ v \in V_F$

e.d. $p_1(T)\circ p_2(T)\circ \ldots \circ p_n(T) = p_1\circ p_2\circ \ldots \circ p_n(T) = T_0$

con $T_0 : V_F \longrightarrow V_F$ $\quad \therefore\ \{\ell \in F[t] \mid \ell(T) = T_0\} \neq \emptyset$

es claro que $\{\ell \in F[t] \mid \ell(T) = T_0\}$ es un ideal bilateral de $F[t]$. Por lo que $\{\ell \in F[t] \mid \ell(T) = T_0\} = \mu(T,t)F[t]$, e.d $\mu(T,t) \mid \ell(t) \quad \forall\ \ell \in F[t]$ con $\ell(T) = T_0$, sin pérdida de generalidad se puede suponer que $\mu(T,t)$ es mónico.

A $\mu(T,t)$ se le conoce como polinomio mínimo para T.

■

$\underline{\text{TEOREMA}}$ $\underline{\text{DE}}$ $\underline{\text{DESCOMPOSICION-3.0}}$) Sea $T : V_F \longrightarrow V_F$ un operador lineal tal que la dimensión de $V = n$ (con $n < \infty$) y sea $\mu(T,t) = p_1^{m_1}(t) \cdot \ldots \cdot p_k^{m_k}(t)$ el polinomio mínimo de T factorizado como producto de polinomios irreducibles, entonces

$$V = K_{p_1} \oplus \ldots \oplus K_{p_k} \quad \text{con } K_{p_i} = \{ v \in V \mid p_i^{m_i} \cdot v = 0 \}$$
$$= \{ v \in V \mid p_i^{m_i}(T)v = 0 \}$$
$$= \mathrm{Ker}\{ p_i^{m_i}(T) \}$$

con K_{p_i} suma directa de subespacios T-cíclicos.

Antes de empezar la demostración recordemos que V es un F[t]-módulo bajo el producto $\cdot$ que definimos en el ejemplo (2) de esta sección como $p \cdot v = p(T)(v)$

donde $T : V_F \longrightarrow V_F$ un operador lineal (dim $V = n$) y $p \in F[t]$ y $v \in V_F$. (*)

AFIRMACION-1.0 .- K_{p_i} es un subespacio T-invariante de V .

Sea $x \in K_{p_i}$ entonces
$$p_i^{m_i}(T(x)) = p_i^{m_i}(T)(T(x))$$
$$= (p_i^{m_i}(T) \circ T)(x)$$
$$= [(([p_i^{m_i}(t)t](T)](x)$$
$$= [(([tp_i^{m_i}(t))](T)](x)$$
$$= (T \circ p_i^{m_i}(T))(x)$$
$$= T(p_i^{m_i}(T)(x))$$
$$= T(0)$$
$$= 0$$

por lo que tambien $T(x) \in K_{p_i}$, así que K_{p_i} es un subespacio T-invariante de V.

∎

(*) Nótese que H es un subespacio de V $\Leftrightarrow$ H es un F[t]-submódulo de V.

Notación: $T_0 : V_F \longrightarrow V_F$ es la función tal que $T_0(v) = 0, \ \forall \ v \in V$.

Pasemos ahora a demostrar que $\sum_{i=1}^{k} K_{p_i}$ es directa y que $V = \bigoplus_{i=1}^{k} Kp_i$

Por hipótesis tenemos que el polinomio mínimo de T se factoriza como $\mu(T,t) = p_1^{m_1}(t) \cdot \ldots \cdot p_k^{m_k}(t)$
hagamos la demostración por inducción sobre k

i) Si k = 1

entonces $\mu(T,t) = p_1^{m_1}(t)$

por lo que $K_{p_1} = V$

ii) Supongamos ahora que para espacios vectoriales con polinomio mínimo de T factorizable en la forma

$$\mu(T,t) = p_1^{m_1}(t) \cdot \ldots \cdot p_k^{m_k}(t) \qquad (*)$$

se tiene que

$$V = K_{p_1} \oplus \ldots \oplus K_{p_k}$$

P.D. que se cumple para espacios tales que el polinomio mínimo de de T se factoriza como

$$\mu(T,t) = p_1^{m_1}(t) \cdot \ldots \cdot p_k^{m_k}(t) \cdot p_{k+1}^{m_{k+1}}(t)$$

Sea V espacio vectorial tal que su polinomio mínimo es

$$\mu(T,t) = p_1^{m_1}(t) \cdot \ldots \cdot p_k^{m_k}(t) \cdot p_{k+1}^{m_{k+1}}(t)$$

y hagamos $g(t) = p_1^{m_1}(t) \cdot \ldots \cdot p_k^{m_k}(t)$

$h(t) = p_{k+1}^{m_{k+1}}(t)$

y sean $W_1 = \{v \in V \ / \ g(t) \cdot v = 0\}$

$W_2 = \{v \in V \ / \ h(t) \cdot v = 0\}$

Tenemos que $g(t)$ y $h(t)$ son primos relativos por lo que existen $q(t)$ y $f(t) \in F[t]$ tales que $g(t)q(t) + h(t)f(t) = 1$

de donde $g(T)q(T) + h(T)f(T) = I$

Ahora bien $[h(t)f(t) \cdot v] \in W_1$

pues $g(t)[h(t)f(t) \cdot v] =$

$f(t)[g(t)h(t) \cdot v] =$

$f(t) \cdot 0 = 0$

(*) Tomemos cada P con coeficiente principal 1, e.d. mónico.

de igual forma $\quad [g(t)q(t)\cdot v] \in W_2$

pues $\quad h(t)[g(t)q(t)\cdot v] =$

$q(t)[h(t)g(t)\cdot v] =$

$q(t)[0] = 0$.

Puesto que $\quad g(T)q(T) + h(T)f(T) = I$

tenemos que $\ g(T)q(T)\cdot v + h(T)f(T)\cdot v = Iv = v \quad \forall\, v \in V$

y así tenemos que $V = W_1 + W_2$

Supongamos que $v \in W_1 \cap W_2$

entonces $\ v = g(T)q(T)\cdot v + h(T)f(T)\cdot v = 0 + 0 = 0$

$\therefore\ v = 0 \quad \therefore\ W_1 \cap W_2 = \{0\}$

$\therefore\ V = W_1 \oplus W_2$

Por hipótesis de inducción tenemos que (*)

$$W_1 = K_{p_1} \oplus \ldots \oplus K_{p_k}$$

de donde $\quad V = K_{p_1} \oplus \ldots \oplus K_{p_k} \oplus K_{p_{k+1}}$

que es lo que necesitabamos probar.

■

(*) Nótese que el polinomio mínimo para $T|_{W_1} = g$

pues $g(t)\cdot w = g(T)(w) = 0\ \forall\ w \in W_1$, por definición de W_1

$\therefore\ \mu(T|_{W_1},\ t) \mid g(t).$ (1)

de igual forma $\ \mu(T|_{W_2},\ t) \mid h(t).$ (2)

y tenemos que $\forall\ v \in V$ se tiene que

$\mu(T|_{W_1}, t)\ \mu(T|_{W_2},\ t)\cdot v = 0 \in V$

por lo que $\ \mu(T,\ t) \mid \ \mu(T|_{W_1},\ t)\ \mu(T|_{W_2},\ t) \mid g(t)h(t) = \mu(T, t)$

entonces $\mu(T|_{W_1},\ t)\ \mu(T|_{W_2},\ t) = s(t)\ \mu(T,\ t)$, con $s(t) \in F[t]$

con $gr(s(t)) = 0$

pero $\mu(T,\ t)$, $\mu(T|_{W_1},\ t)$, $\mu(T|_{W_2},\ t)$ son todos mónicos

$\therefore\ s(t) = 1 \in F[t]$

$\mu(T,\ t) = \ \mu(T|_{W_1},\ t)\ \mu(T|_{W_2},\ t)$

de (1) y (2) tenemos que

$g(t) = q_1(t)\mu(T|_{W_1}, t)$ y $h(t) = q_2(t)\mu(T|_{W_2}, t)$

por hipótesis

$\mu(T,t) = p_1^{m_1}(t) \cdot \ldots \cdot p_k^{m_k}(t) \cdot p_{k+1}^{m_{k+1}}(t) = g(t)h(t)$

$= \mu(T|_{W_1}, t)\ \mu(T|_{W_2}, t)$

Pero $g(t)$, $h(t)$, $(T|_{W}, t)$ y $\mu(T|_{W_2}, t)$

son mónicos por lo que $q(t) = q(t) = 1 \in F[t]$

$\therefore g(t) = \mu(T|_{W_1}, t)$ y $h(t) = \mu(T|_{W_2}, t)$.

$\square$

Para $T : V \longrightarrow V$ tenemos que:

$$T|_{Kp_1}: Kp_1 \longrightarrow Kp_1$$
$$T|_{Kp_2}: Kp_2 \longrightarrow Kp_2$$
$$\vdots \qquad \vdots \qquad \vdots$$
$$T|_{Kp_k}: Kp_k \longrightarrow Kp_k$$

$\therefore$ si $x = x_1 + x_2 + \ldots + x_k$ con $x_i \in Kp_i$

entonces $T(x) = T|_{Kp_1}(x_1) + T|_{Kp_2}(x_2) + \ldots + T|_{Kp_k}(x_k)$

Observemos que Kp_i es un módulo sobre el anillo $F[t]/\langle p_i^{m_i}\rangle$. Esto es análogo al caso en que en un grupo abeliano $p^n g = 0 \in G$ $\forall\, g \in G$ y p.a. $n \in \mathbb{N}$.

Definición.- Tenemos como en el teorema para grupos abelianos que $p\text{-zoc}(Kp) = \{x \in K_{p_i} \mid px = 0\} = \mathrm{Ker}\,(p(T))$ así que $p\text{-zoc}(Kp)$ es un subespacio de Kp.

TEOREMA-4.0) $\quad K_{p_i} = \underset{\alpha\in X}{\oplus} x_\alpha F[t]$, $\quad x_{\alpha_j} \in K_p$.

(Suma directa de F[t]-submódulos y tambien es suma directa de espacios vectoriales)

La demostración del teorema la haremos por inducción sobre m con $m = \min\{l \;/\; p^l(T) = T_0\}$

i) base de inducción m = 1

Si $m = 1$ entonces $V = K_p$ y $pK_p = \{0\}$

e.d. $p(T)(x) = 0 \;\forall\; x \in K_p$ $\therefore$ V es un $F[t]/_{<p>}$ - módulo,

e.d. es un espacio vectorial sobre el campo $F[t]/_{<p>}$ y como los espacios vectoriales tienen bases, $V = \underset{\alpha\in X}{\oplus}(F[t]/_{<p>})x_\alpha = \underset{\alpha\in X}{\oplus}F[t]x_\alpha$.

ii) Hipótesis de inducción

supongamos que $pK_p = \oplus\, x_\alpha F[t]$ $\quad (\; p^{m-1}(pK_p) = p^m K_p = \{0\}\;)$.

entonces en el siguiente diagrama podemos observar que:

$$\begin{array}{ccc} p\text{-zoc}(K_p) & \hookrightarrow & K_p \\ \uparrow & & \uparrow \\ p\text{-zoc}(pK_p) & \hookrightarrow & pK_p = \oplus\, x_\alpha F[t] \end{array}$$

(por hipótesis)

Si $x_\alpha \in pK_p$ entonces $x_\alpha = py_\alpha = p(T)y_\alpha$, p.a. $y_\alpha \in Kp$.

Puesto que $p\text{-zoc}(pK_p) \hookrightarrow p\text{-zoc}(K_p)$ y $p\text{-zoc}(pK_p)$ es un subespacio vectorial de $p\text{-zoc}(K_p)$, como espacio sobre el campo $F[t]_{/<p>}$, entonces $p\text{-zoc}(K_p) = p\text{-zoc}(pK_p) \oplus W$ suma directa de $F[t]_{/<p>}$ espacios vectoriales, $\therefore$ suma directa de F[t]-módulos, $\therefore$ es suma directa de F espacios vectoriales.

{*notación $\hookrightarrow$ nos indicará la función inclusión ...*}

Afirmación 1.- $\sum y_\alpha F[t] = \oplus\, y_\alpha F[t]$. e.d. la suma es directa.

Afirmación 2.- $W \cap (\oplus\, y_\alpha F[t]) = \{0\}$.

Afirmación 3.- $K_p = (\oplus\, \mathbb{Z}_{y_\alpha}) \oplus W$.

Observemos que el módulo $z_\alpha F[t]$ como espacio vectorial sobre F es el subespacio generado por $\{z_\alpha, T(z_\alpha), T^2(z_\alpha), \ . \ . \ . \}$ que se llama el espacio vectorial T-cíclico generado por z_α. Puesto que para $z_\alpha \in Kp$ se tiene que:

$$\begin{aligned} z_\alpha F[t] &= \{ f(t)z_\alpha \mid f(t) \in F[t] \} \\ &= \{ f(T)z_\alpha \mid f(t) \in F[t] \} \end{aligned}$$

$$\begin{aligned} z_\alpha F[t] &= \mathcal{L}(\{ f(T)z_\alpha \mid f(t) \in F[t] \}) \\ &= \mathcal{L}\{ Iz_\alpha , T(z_\alpha) , T^2(z_\alpha) , T^3(z_\alpha) , \ . \ . \ . \ \} \end{aligned}$$

$$(f(T) = a_nT^n + a_{n-1}T^{n-1} + \ . \ . \ . + a_1T + a_0I$$
$$\Rightarrow f(t)z_\alpha = a_nT^n(z_\alpha) + \ . \ . \ . + a_0z_\alpha).$$

E.d. $z_\alpha F[t]$ es cíclico como $F[t]$-módulo y es T-cíclico como espacio vectorial sobre F.

Demostración de la afirmación 1.

Sea $0 = r_1y_{\alpha_1} + \ldots + r_ny_{\alpha_n}$ una combinación lineal de elementos de $y_\alpha F[t]$, con $r_iy_{\alpha_i} \in y_{\alpha_i}F[t]$

P.D. $r_iy_{\alpha_i} = 0 \quad \forall i \in \{1, \ldots ,n\}$

$0 = p(0)$ entonces

$$\begin{aligned} 0 = p(0) &= p(r_1y_{\alpha_1} + \ldots + r_ny_{\alpha_n}) \\ &= r_1p(y_{\alpha_1}) + \ldots + r_np(y_{\alpha_n}) \\ &= r_1x_{\alpha_1} + \ldots + r_nx_{\alpha_n} \in \bigoplus_{i=1}^{n} x_\alpha F[t] \end{aligned}$$ (por hipótesis de inducción) y por lo tanto $r_ix_\alpha = 0_i$ $\forall i$, y $r_ix_\alpha = r_ip(y_\alpha)_i$ $\forall\, i \in \{1, \ldots n\}$.

Se sigue que: $r_i y_{\alpha_i} \in$ p-zoc$(y_{\alpha_i} F[t])$ = p-zoc$(x_{\alpha_i} F[t]) \subset \mathbb{Z}x_{\alpha_i}$,

$\therefore 0 = r_1 y_{\alpha_1} + \ldots + r_n y_{\alpha_n} \in \oplus\, x_\alpha F[t]$ y $r_i y_{\alpha_i} \in x_{\alpha_i} F[t]$

$\forall\, i \in \{1, \ldots, n\}$

por lo tanto $r_i y_{\alpha_i} = 0 \;\forall\, i \in \{1, \ldots n\}$ y por lo tanto se tiene que $\sum y_\alpha F[t] = \oplus\, y_\alpha F[t]$.

■

Demostración de la afirmación 2.-

Sea $w \in W \cap (\oplus\, y_\alpha F[t])$ entonces

$w = r_1 y_{\alpha_1} + \ldots + r_n y_{\alpha_n}$ puesto que w está en la intersección.

Como $w \in W \leq$ p-zoc(Kp)

entonces $w \in$ p-zoc$(\oplus y_\alpha F[t])$ = $\oplus$p-zoc$(y_\alpha F[t])$ = $\oplus$p-zoc$(x_\alpha F[t])$

$\leq \oplus\, x_\alpha F[t] = pK_p$

$\therefore w \in pK_p \cap W = \{0\}$

$\therefore w = 0$.

■

Demostración de la afirmación 3.-

Sea $v \in K_p$ entonces $pv \in pK_p = \oplus\, x_\alpha F[t]$ (por hipótesis de inducción) $\therefore$ $p \cdot v$ puede escribirse como una combinación lineal de elementos de $x_\alpha F[t]$, sea $pv = r_1 x_{\alpha_1} + \ldots + r_n x_{\alpha_n}$ con $r_i x_{\alpha_i} \in x_{\alpha_i} F[t]$, consideremos:

$$z = r_1 y_{\alpha_1} + \ldots + r_n y_{\alpha_n}$$

con $r_i y_{\alpha_i} \in y_{\alpha_i} F[t]$. $pz = pv$

$$\begin{aligned} \text{pues } pz &= p(r_1 y_{\alpha_1} + \ldots + r_n y_{\alpha_n}) \\ &= pr_1 y_{\alpha_1} + \ldots + pr_n y_{\alpha_n} \\ &= r_1 p y_{\alpha_1} + \ldots + r_n p y_{\alpha_n} \end{aligned}$$

$= r\, x_{1\,\alpha_1} + \dots + r\, x_{n\,\alpha_n}$

$= pv.$

Por lo tanto $p(z - v) = 0$,

e.d. $z-v \in$ p-zoc(Kp)

tanto $(z-v)$ como $(-z+v)$ están en $\text{p-zoc}(K_p) = \text{p-zoc}(pK_p) \oplus W$ entonces $(-z+v) = (z'+v)$ con $z' \in \text{p-zoc}(pK_p)$ y $w \in W$. Entonces $v = z' + w + z$ con $z \in \oplus\, y_\alpha F[t]$

como $z' \in \text{p-zoc}(pK_p) = \text{p-zoc}(\oplus\, x_\alpha F[t])$

$= \oplus(\text{p-zoc}(\, x_\alpha F[t]))$

$= \oplus(\text{p-zoc}(\, y_\alpha F[t]))$

$= \text{p-zoc}(\oplus\, y_\alpha F[t]) \leq \oplus y_\alpha F[t]$,

entonces $z' \in \oplus\, y_\alpha F[t]$

$\therefore\ z + z' \in \oplus\, y_\alpha F[t]$

$\therefore\ v = z' + z + w$ pertenece a $\oplus\, y_\alpha F[t] \oplus W$.

$\therefore\ K_p \leq \oplus\, y_\alpha F[t] \oplus W \leq K_p$.

$\therefore\ K_p = \oplus\, y_\alpha F[t] \oplus W$.

Lo que demuestra que K_p es suma directa de submódulos cíclicos. W es una suma directa de $F[t]$-submódulos cíclicos pues $W \leq \text{p-zoc}(K_p)$ así que se puede aplicar el argumento de la base de la inducción. Con lo cual queda demostrado el teorema de descomposición cíclica.

∎

FORMA CANONICA RACIONAL Y DE JORDAN

Sea V un espacio vectorial de dimensión finita sobre un cierto campo F y $T : V_F \longrightarrow V_F$ un operador lineal tal que su polinomio mínimo es $\mu(T,t) = p_1^{k_1}(t) \cdot \ldots \cdot p_\ell^{k_\ell}(t)$ entonces sabemos que

$$V = K_{p_1} \oplus \ldots \oplus K_{p_\ell} \quad \text{donde} \quad K_{p_i} = \{ v \in V \mid p_i^{k_i} v = 0 \}$$
$$= \{ v \in V \mid p_i^{k_i}(T)v = 0 \}$$
$$= \mathrm{Ker}\{ p_i^{k_i}(T) \} \leq V$$

Nótese que : $T|_{K_{p_i}} : K_{p_i} \longrightarrow K_{p_i}$

Ahora $K_{p_i} = \oplus_{\alpha=1} F[t]z_\alpha$

$F[t]z_\alpha = \{f(t)(z_\alpha) \mid f(t) \in F[t] \}$

$= \mathcal{L}\{ z_\alpha, T(z_\alpha), T^2(z_\alpha), \ldots \}.$

Notemos que $F[t]z_\alpha \leq K_{p_i}$ y es T-invariante pues

$T(\mathcal{L}\{ z_\alpha, T(z_\alpha), T^2(z_\alpha), \ldots \}) = \mathcal{L}(T(z_\alpha), T^2(z_\alpha), \ldots \}.$

<u>OBSERVACION-3.0</u>) Sea $T : V_F \longrightarrow V_F$ un operador lineal con V_F espacio vectorial de dimensión finita y

$\mu(T,t) = p_1^{k_1}(t) \cdot \ldots \cdot p_\ell^{k_\ell}(t)$ su polinomio mínimo

$K_{p_i} \leq V_F$, tenemos $T|_{K_i} : K_{p_i} \longrightarrow K_{p_i}$

además su polinomio mínimo es $\mu(T|_{K_{p_i}}, t) = p_i^{k_i}(t)$.(ver la prueba del teorema de descomposición para espacios vectoriales)

<u>OBSERVACION-4.0</u>)

$K_{p_i} = \oplus_{\alpha=1}^{s} F[t]z_\alpha$ y tenemos que $T|_{F[t]z_\alpha} : F[t]z_\alpha \longrightarrow F[t]z_\alpha$

Demostraremos que $\dim_F F[t]z_\alpha = k_\alpha m_i$, donde m_i es el grad(p_i) y

$\mu(T|_{F[t]z_\alpha}, t) = p_i^{k_\alpha}(t)$

Demostración:

Es claro que si $\mu(T|_{F[t]z_\alpha}, t) \mid p_1^{k_1}(t)$

entonces $\mu(T|_{F[t]z_\alpha}, t) = p_1^{k_\alpha}(t) \quad 1 \le k_\alpha \le k_1$

$F[t]z_\alpha = \mathcal{L}\{ T^i(z_\alpha) \mid i \in \mathbb{N} \cup \{0\}\}$

como $F[t]z_\alpha \le K_{p_1} \le V_F$ y V es de dimensión finita entonces

$F[t]z_\alpha = \mathcal{L}\{ z_\alpha, \ldots, T^{m-1}(z_\alpha)\}$ $m \in \mathbb{N}$ y podemos escoger m mínima con esta propiedad.

<u>AFIRMACION-2.0</u>) $\{ z_\alpha, \ldots, T^{m-1}(z_\alpha)\}$ es linealmente independiente pues en caso contrario existiría $n \le m-1$ tal que

$T^n(z_\alpha) \in \mathcal{L}\{ z_\alpha, \ldots, T^{n-1}(z_\alpha)\}$ pero entonces

$T^{n+1}(z_\alpha) \in \mathcal{L}\{T(z_\alpha), \ldots, T^n(z_\alpha)\} \le \mathcal{L}\{ z_\alpha, \ldots, T^{n-1}(z_\alpha)\}$

así que por inducción $T^k(z_\alpha) \in \mathcal{L}\{ z_\alpha, \ldots, T^{n-1}(z_\alpha)\}$ $\forall\ k \in \mathbb{N}$ con $n-1 < m-1$,
contradiciendo la elección de m.

□

Como el polinomio mínimo de $T|_{F[t]z_\alpha}$ es $\mu(T|_{F[t]z_\alpha}, t) = p_1^{k_\alpha}(t)$

entonces $p_1^{k_\alpha}(T)(z_\alpha) = 0.$

Si $grad(p_1) = m_1$ entonces

$T^{m_1 k_\alpha}(z_\alpha) \in \mathcal{L}\{ z_\alpha, \ldots, T^{m_1 k_\alpha - 1}(z_\alpha)\}$

$\therefore\ m-1 \le m_1 k_\alpha - 1 \quad \therefore\ m \le m_1 k_\alpha$.

Como $T^m(z_\alpha) \in \mathcal{L}\{ z_\alpha, \ldots, T^{m-1}(z_\alpha)\}$

entonces $T^m(z_\alpha) = \sum^{m-1} (-a_i)T^i(z_\alpha)$

$\therefore\ T^m(z_\alpha) + \sum^{m-1} (a_i)T^i(z_\alpha) = 0$

$\therefore\ (t^m + \sum^{m-1} a_i t^i)(T) = T_0 |_{F[t]z_\alpha}$

$\therefore\ \mu(T \mid_{F[t]z_\alpha}, t) \mid (t^m(z_\alpha) + \sum^{m-1} a_i t^i)$

$\therefore \operatorname{grad}(\mu(T \mid_{F[t]z_\alpha}, t) = k_\alpha m_1 \mid m$

$\therefore \quad k_\alpha m_1 \leq m$

$\therefore m = k_\alpha m_1$

$\therefore \dim_F (F[t]z_\alpha) = k_\alpha m_1$

donde $m_1 = \operatorname{grad}(p_1)$

y $\mu(T \mid_{F[t]z_\alpha}) = p_1^{k_\alpha}(t)$

$\therefore \dim(F[t]z_\alpha) = \operatorname{grad}(p_1^{k_\alpha}(t)) = \operatorname{grad}(\mu(T \mid_{F[t]z_\alpha}, t))$

$\therefore \beta_\alpha = \{ z_\alpha, \ . \ . \ . \ . \ , T^{m_1 k_\alpha - 1}(z_\alpha)\} \xhookrightarrow[\text{base}]{} F[t]z_\alpha$

$\therefore \dim F[t]z_\alpha = m_1 k_\alpha$ ∎

<u>OBSERVACION-5.0</u>)

$$[T|_{F[T]z_\alpha}]_{\beta_\alpha}^{\beta_\alpha} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ . & 0 & 1 & \cdots & . \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ . & & & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \left(\left[\; T^{k_\alpha m_1}(z_\alpha) \; \right]^{\beta_\alpha}_{\beta_\alpha} \right) =$$

$$= \begin{pmatrix} 0 & 0 & 0 & \cdots & & -a_0 \\ 1 & 0 & 0 & \cdots & & -a_1 \\ 0 & 1 & 0 & \cdots & & \\ . & 0 & 1 & \cdots & & . \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ . & \vdots & \vdots & & 0 & . \\ 0 & 0 & 0 & \cdots & 1 & -a_{k_\alpha m_1 - 1} \end{pmatrix} \qquad (*)$$

Pues $T^{k_\alpha m_1}(z_\alpha) = \sum^{k_\alpha m_1 - 1} (-a_i) T^i(z_\alpha)$

y entonces $T^{k_\alpha m_1}(z_\alpha) + \sum^{k_\alpha m_1 - 1} (a_i)(T^i)(z_\alpha) = 0$

donde $T^{k_\alpha m_1} + \sum (a_i) T^i$ es $\mu(T|_{F[t]z_\alpha}, t)$

$\therefore$ la última columna de $[T_{|F[t]z_\alpha}]_{\beta_\alpha}^{\beta_\alpha}$ está formada por los inversos aditivos de los coeficientes $a_1, a_2, \ . \ . \ . \ , a_{k_\alpha m_i - 1}$ del del polinomio mínimo de de $T|_{K_{\alpha_i}}$.

<u>DEFINICION-11.0</u>) La matriz (*)se llama la matriz compañera del polinomio $t^{k_\alpha m_i} + \sum_{i=0}^{k_\alpha m_i} a_i t^i = \mu(T_{|F[t]z_\alpha}, t)$

Como $Kp_i = \oplus F[t]z_\alpha$ es una suma finita

y $\beta_{\alpha_i} \xhookrightarrow[\text{base}]{} F[t]z_{\alpha_i}$

entonces $\beta = \beta_{\alpha_1} \cup \beta_{\alpha_2} \cup \ . \ . \ . \ \cup \beta_{\alpha_s} \xhookrightarrow[\text{base}]{} Kp_i$

digamos que $\mu(T\ |_{F[t]z_{\alpha_i}}) = p_i^{k_{\alpha_i}}(t)$ y $\text{grad}(p_i) = m_i$

entonces

$$
[\ T|_{Kp_i}]_\beta^\beta = \left(\begin{array}{c|c|c}
\begin{matrix} 0 & 0 & \cdots & * \\ 1 & 0 & \cdots & * \\ 0 & 1 & 0 \cdots & * \\ \vdots & & & * \\ & & 0 & * \\ & & 1 & * \end{matrix} & & 0 \\
\hline
& \begin{matrix} 0 & 0 & \cdots & * \\ 1 & 0 & \cdots & * \\ 0 & 1 & 0 \cdots & * \\ \vdots & & & * \\ & & 0 & * \\ & & 1 & * \end{matrix} & \\
\hline
0 & & \ddots \quad \boxed{}
\end{array}\right)
$$

Donde el ℓ-ésimo bloque es $[\, T|_{F[t]z_\ell} \,]_{\beta_\ell}^{\beta_\ell}$.

DEFINICION-12.0) $\beta_i = \beta_{\alpha_1} \cup \beta_{\alpha_2} \cup \; . \; . \; . \cup \beta_{\alpha_s}$ donde $\beta_{\alpha_j} = \{x_{\alpha_j} \, , \; T(x_{\alpha_j}), \; T^2(x_{\alpha_j}) \, , \; . \; . \; . \; \}$ se llama base canónica racional para Kp_i y escribiremos $\beta_i \xhookrightarrow[\text{b. c. rac.}]{} Kp_i$.

Como $V = Kp_1 \oplus Kp_2 \oplus \ldots \oplus Kp_\ell$ y tenemos que si $\beta_i \xhookrightarrow[\text{b. c. rac.}]{} Kp_i$ entonces $\beta_1 \cup \beta_2 \cup \; . \; . \; . \cup \beta_\ell$ se llama una base canónica racional para V_F respecto a T.

Pasemos a construir la forma de Jordan.

Sabemos que si $T : V_F \longrightarrow V_F$ un operador lineal tal que su polinomio mínimo es $\mu(T,t) = p_1^{k_1}(t) \cdot \ldots \cdot p_\ell^{k_\ell}(t)$ (con p_i irreducible) y $F = \mathbb{C}$ entonces $p_i = (t-\lambda_i)$.

Sea V_F como en el teorema, e.d. $V_F = K_{p_1} \oplus \ldots \oplus K_{p_\ell}$ (con $F = \mathbb{C}$) y fijemos una i , escribiendo $K_p = K_{p_i}$ con polinomio mínimo $\mu(T|_{Kp}, t) = p^k(t)$ y $p = t-\lambda$

entonces $Kp = \bigoplus_{j=1}^{m} x_j F[t]$ y $\mu(T|_{x_i F[t]}, t) = p^{k_j}(t)$

con $k_1 \geq k_2 \geq \; . \; . \; . \geq k_m$.

Sea $\beta_i = \{x_i, \; (T-\lambda)x_i \; , \; . \; . \; . \; . \; , \; (T-\lambda)^{k_i - 1}x_i\} \xhookrightarrow[\text{base ord.}]{} x_i F[t]$

y hagamos $U = T - \lambda I$ es claro que si $\mu(T|_{x_i F[t]}, t) = (t-\lambda)^{k_i}$

entonces $U^{k_i}|_{x_i F[t]} = U_0 = 0$ y

$\beta_i = \{x_i, \; Ux_i \; , \; . \; . \; . \; . \; , \; U^{k_i - 1}x_i\} \xhookrightarrow[\text{base ord.}]{} x_i F[t]$

$$[U]_{\beta_1}^{\beta_1} = \begin{pmatrix} 0 & 0 & . & . & . & 0 & 0 \\ 1 & 0 & . & . & . & 0 & 0 \\ 0 & 1 & . & . & . & & 0 \\ . & & . & & & & . \\ . & & & & . & & . \\ . & & & & & . & . \\ 0 & & . & . & . & 1 & 0 \end{pmatrix} = [T - \lambda I]_{\beta_1}^{\beta_1} =$$

$$= [T]_{\beta_1}^{\beta_1} - \lambda [I]_{\beta_1}^{\beta}$$

por lo que $\quad [T]_{\beta_1}^{\beta_1} = \lambda [I]_{\beta_1}^{\beta} + [U]_{\beta_1}^{\beta}$

$$\therefore \quad [T]_{\beta_1}^{\beta_1} = \begin{pmatrix} \lambda & 0 & . & . & . & 0 & 0 \\ 1 & \lambda & . & . & . & 0 & 0 \\ 0 & 1 & . & . & . & & 0 \\ . & & . & & & & . \\ . & & & & . & & . \\ . & & & & & \lambda & . \\ 0 & & . & . & . & 1 & \lambda \end{pmatrix}$$

si tomamos a $\gamma_1 = \{(T-\lambda)^{k_1-1}x_1, \; . \; . \; . \; . \;, (T-\lambda)x_1 \;, \; x_1\}$ que es β_1 ordenada de manera inversa

$$[T]_{\gamma_1}^{\gamma_1} = \left([I]_{\gamma_1}^{\beta_1} \right)^{-1} \begin{pmatrix} \lambda & 0 & . & . & . & 0 & 0 \\ 1 & \lambda & . & . & . & 0 & 0 \\ 0 & 1 & . & . & . & & 0 \\ . & & . & & & & . \\ . & & & & . & & . \\ . & & & & & \lambda & . \\ 0 & & . & . & . & 1 & \lambda \end{pmatrix} [I]_{\gamma_1}^{\beta_1}$$

notemos que $[I]_{\gamma_1}^{\beta_1}$ es una matriz autoinvertibles puesto que invertir el orden de los elementos de una base es una operación autoinvertible.

$$\therefore \quad [T]_{\gamma_1}^{\gamma_1} = P \, [T]_{\beta_1}^{\beta_1} \, P$$

$$
= \begin{pmatrix}
0 & 0 & \cdot & \cdot & & 1 & \lambda \\
0 & \cdot & \cdot & \cdot & 1 & \lambda & 0 \\
 & \cdot & \cdot & \cdot & \lambda & & 0 \\
\cdot & \cdot & \cdot & & & & \cdot \\
\cdot & & \cdot & & & & \cdot \\
1 & \lambda & & & & & \cdot \\
\lambda & & \cdot & \cdot & \cdot & & 0
\end{pmatrix} P
$$

$$
= \begin{pmatrix}
\lambda & 1 & \cdot & \cdot & \cdot & 0 & 0 \\
0 & \lambda & \cdot & \cdot & \cdot & 0 & 0 \\
0 & 0 & \cdot & \cdot & \cdot & & 0 \\
\cdot & & \cdot & & & & \cdot \\
\cdot & & & & \cdot & 1 & \cdot \\
\cdot & & & & & \lambda & 1 \\
0 & & \cdot & \cdot & \cdot & 0 & \lambda
\end{pmatrix}
$$

que es un bloque Jordan de orden $k \times k$ y donde k es el grado de $\mu(T|_{K_p}, t) = p^k(t)$ por lo que si tomamos $\gamma = \gamma_1 \cup \gamma_2 \cup \ldots \cup \gamma_m$ $(\gamma_1 \xhookrightarrow[\text{base ord.}]{} x_1 F[t])$, tenemos que γ es base de V y que

$$
[\,T|_{K_p}]_\gamma^\gamma = \left(\begin{array}{c|c|c}
\begin{matrix}
\lambda & 1 & \cdot & \cdot & 0 & 0 \\
0 & \lambda & 1 & \cdot & \cdot & 0 \\
0 & 0 & \lambda & \cdot & \cdot & \cdot \\
\cdot & & & \cdot & & \\
\cdot & & & & \cdot & 1 \\
0 & \cdot & & \cdot & \cdot & \lambda
\end{matrix} & & 0 \\
\hline
 & \begin{matrix}
\lambda & 1 & \cdot & \cdot & 0 & 0 \\
0 & \lambda & 1 & \cdot & \cdot & 0 \\
0 & 0 & \lambda & \cdot & \cdot & \cdot \\
\cdot & & & \cdot & & \\
\cdot & & & & \cdot & 1 \\
0 & \cdot & & \cdot & \cdot & \lambda
\end{matrix} & \\
\hline
0 & & \ddots \quad \begin{matrix}
\lambda & 1 & & 0 \\
\vdots & & \ddots & \vdots \\
0 & \cdot & \cdot & \lambda
\end{matrix}
\end{array}\right)
$$

donde la dimensión del ℓ bloque J_ℓ es el $\text{grad}((t-\lambda)^{k_\ell}) = k_\ell$

■

Que es la forma de Jordan que queríamos construir.

Pasemos a ver una forma para calcular la forma racional de $T : V_F \longrightarrow V_F$, donde $V_F = \oplus F[t]x_i$, suponiendo que $\mu(T,t) = p^{\ell}$, con $p \in F[t]$ irreducible y $\mu(T_{|F[t]z_\alpha}, t) = p^{k_i}$ y $k_1 \geq k_2 \geq . . . \geq k_m$

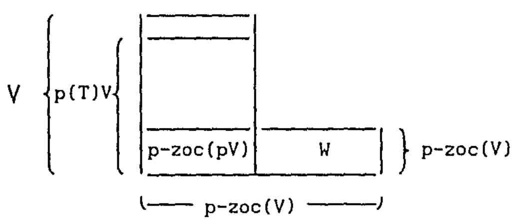
Denotemos con $S_{p,k}(V)$ el número de sumandos directos en $V = (\oplus_{i=1}^{m} x_i F[t])$ que tengan polinomio mínimo p^k.

<u>TEOREMA-5.0</u>) Sea T un operador lineal ($T: V \longrightarrow V$) y V espacio vectorial tal que $V = (\oplus_{i=1}^{m} x_i F[t])$ y $\mu(T,t) = p^{\ell}$ con p irreducible donde δ = grado de p entonces:

$$m = \frac{1}{\delta} (\dim_F(N(p(T))) = \frac{1}{\delta}(\dim_F(V)-\dim_F(p(V)))$$

y $S_{p,k}(V) = \frac{1}{\delta} (\dim(p^{k-1}V) - 2\dim(p^k V)) + \dim(p^{k+1}V) \quad \forall\ k \geq 1.$

Demostración:



$$\boxed{\begin{array}{l} \text{p-zoc}(V) = \text{p-zoc}(pV) \oplus W \\ p(T)V = \oplus y_i F[t] \\ V = \oplus x_i F[t] \oplus W \\ px_i = y_i \end{array}}$$

Recordemos que $\text{p-zoc}(V) = \text{p-zoc}(pV) \oplus W$ con $p(V) = \oplus y_i F[t]$ y $y_i = p(T)x_i$, $V = (\oplus x_i F[t]) \oplus W$.

Notemos que $\text{p-zoc}(V) = \overset{m}{\underset{i=1}{\oplus}} \text{p-zoc}(x_i F[t])$

$\text{p-zoc}(x_i F[t])$ es un espacio vectorial sobre $F[t]_{/<p>}$, de dimensión uno $\therefore$ $\text{p-zoc}(x_i F[t])$ es isomorfo a $F[t]_{/<p>}$

$\therefore \quad m = \dim_{F[t]_{/<p>}} (\text{p-zoc}(V))$.

(Observemos que si $x \in \text{p-zoc}(V)$ entonces $\dim_{F[t]_{/<p>}} xF[t] = 1$ pues $\{x\} \xhookrightarrow[\text{base}]{} F[t]_{/<p>} \; xF[t])$

Y que $\dim_F(xF[t]) = \text{grad}(p)$

pues $_F(xF[t]) = \mathcal{L}\{x, T(x), \ldots, T^{\delta-1}(x)\}$ y donde $\delta = \text{grad}(p)$).

$\therefore$ Si $m = \dim_{F[t]_{/<p>}} \text{p-zoc}(V)$ entonces $\text{p-zoc}(V) = \overset{m}{\underset{i=1}{\oplus}} (x_i F[t])$

y $\dim_F \text{p-zoc}(V) = m(\text{grad}(p)) = m\delta$

e.d. $m = \frac{1}{\delta}(\dim_F \text{p-zoc}(V))$

$\blacksquare$

$S_{p,k}(V)$ denota el número de sumandos directos de $V = (\overset{m}{\underset{i=1}{\oplus}} x_i F[t])$ con polinomio mínimo p^k y ya vimos que

$m = \frac{1}{\delta}(\dim_F \text{p-zoc}(V)) = \frac{1}{\delta}(\dim_F N(p(T))) = (\dim_{F[t]_{/<p>}} \text{p-zoc}(V))$.

Ahora $\dim_{F[t]_{/<p>}}(W) = S_{p,1}(V) = (\dim_{F[t]_{/<p>}} \text{p-zoc}(V))$

$\qquad - (\dim_{F[t]_{/<p>}} \text{p-zoc}(pV))$

$S_{p,2}(V) = S_{p,1}(pV) = \dim_{F[T]_{/<P>'}} (\text{p-zoc}(pV)) - \dim_{F[T]_{/<P>'}} (\text{p-zoc}(ppV))$

$\qquad = \dim_{F[T]_{/<P>'}} (\text{p-zoc}(pV)) - \dim_{F[T]_{/<P>'}} (\text{p-zoc}(p^2V))$

en general:

$S_{p,k}(V) = S_{p,k-1}(pV) = S_{p,k-2}(p^2V) = \ldots = S_{p,1}(p^{k-1}V)$

$$= \dim_{F[T]/\langle P\rangle'}(p\text{-zoc}(p^{k-1}V)) - \dim_{F[T]/\langle P\rangle'}(p\text{-zoc}(p^{k}V))$$

$$\therefore S_{p,k}(V) = \dim_{F[T]/\langle P\rangle'}(p\text{-zoc}(p^{k-1}V)) - \dim_{F[T]/\langle P\rangle'}(p\text{-zoc}(p^{k}V))$$

$$\therefore S_{p,k}(V) = \dim_{F[t]/\langle p\rangle}(p\text{-zoc}(p^{k-1}V)) - \dim_{F[t]/\langle p\rangle}(p\text{-zoc}(p^{k}V))$$

entonces $S_{p,k}(V) = \frac{1}{\delta}(\dim_F(p\text{-zoc}(p^{k-1}V)) - \dim_F(p\text{-zoc}(p^{k} V)))$

(por la observación en la página anterior).

Ahora $p\text{-zoc}(p^{k-1}V) = \{x \in (p^{k-1}V) \mid p\circ x = 0\}$

$$= p^{k-1}(V) \cap N(p(T))$$

$$= N(p(T)_{|p^{k-1}(V)})$$

con $(p(T)_{|p^{k-1}(V)}) : p^{k-1}V \longrightarrow p^{k-1}V$

puesto que $\dim_F(V) = \dim_F(N(T)) + \dim_F(\text{rango}(T))$. Para $T: V \longrightarrow V$

se tiene que $\dim_F(p^{k-1}V) = \dim_F(p^{k-1}(V) \cap N(p(T))) + \dim_F(p^{k}(V))$.

$\therefore \dim_F(p\text{-zoc}(p^{k-1}V)) = \dim_F(p^{k-1}V) - \dim_F(p^{k}V)$.

$\therefore S_{p,k}(V) = \frac{1}{\delta}(\dim_F(p^{k-1}V) - \dim_F(p^{k}V)) - \dim_F(p^{k}V) + \dim_F(p^{k+1}V)$

$\therefore \qquad\qquad = \frac{1}{\delta}(\dim_F(p^{k-1}V) - 2\dim_F(p^{k}V)) + \dim_F(p^{k+1}V)$.

∎

<u>COROLARIO-1.0</u>) Si $T: V_F \longrightarrow V_F$ tiene polinomio mínimo $p^{k}(t)$, $p \in F[t]$ irreducible y $V = \bigoplus_{i=1}^{m} F[t]x_i$ con $\mu(T_{|F[t]x_i}, t) = p^{k_i}$, $k_1 \geq k_2 \geq \ldots \geq k_m$.

Si $\beta \xhookrightarrow[\text{b.c.rac.}]{} V_F$ entonces el número de bloques racionales de tamaño $k_i\text{grad}(p)$ es $S_{p,k_i}(T)$ y el número de bloques es m, calculados como en el teorema anterior.

∎

TEOREMA-6.0)Sean T , $U : V_F \longrightarrow V_F$ operadores lineales entonces T, U son similares si y solo si T, U admiten la misma forma canónica racional.

T, U son similares si existe un isomorfismo $\psi : V_F \longrightarrow V_F$ tal que el siguiente diagrama conmuta.

$$\begin{array}{ccc} V_F & \xrightarrow{T} & V_F \\ \downarrow\psi & & \downarrow\psi \\ V_F & \xrightarrow[U]{} & V_F \end{array}$$

COROLARIO-2.0) Sean $T, U : V_F \longrightarrow V_F$ tales que $\mu(T,t) = \mu(U,t)$ se factoriza como un producto de factores de grado 1, entonces T, U son similares si y solo si T, U admiten la misma forma canónica de Jordan.

Demostración del Teorema.

$\Leftarrow$) Supongamos existe $\beta \xhookrightarrow[\text{b.c.rac.}]{} V_F$ y $\gamma \xhookrightarrow[\text{b.c.rac.}]{} V_F$ tal que:

$[\,T\,]_\beta^\beta = [\,U\,]_\gamma^\gamma$

es claro que $f(T) = T_0 \Leftrightarrow f([\,T\,]_\beta^\beta) = 0 \in \mathcal{M}_{n\times n}(F)$

$\therefore \mu(T,t) = \mu([\,T\,]_\beta^\beta, t) = \mu([\,U\,]_\gamma^\gamma, t) = \mu(U,t).$

$\mu(T,t) = p_1^{k_1}(t) \cdot \ \dots \ \cdot p_r^{k_r}$

$V = K_{p_1} \oplus \dots \oplus K_{p_r}$ con $K_{p_i} = N(p_i^{k_i}(T))$ y también

$V = K'_{p_1} \oplus \dots \oplus K'_{p_r}$ con $K'_{p_i} = N(p_i^{k_i}(U))$.

Ahora $\beta = (\beta \cap K_{p_1}) \cup \dots \cup (\beta \cap K_{p_r})$

definamos $\beta_i = \beta \cap K_{p_i}$ y

$\gamma = (\gamma \cap K'_{p_1}) \cup \dots \cup (\gamma \cap K'_{p_r})$

definamos $\gamma_i = \gamma \cap K_{p_i}$, por definición de base canónica racional.

Como

$$[T]_{\beta}^{\beta} = \begin{pmatrix} [T|_{K_{p_1}}] & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ & & & & [T|_{K_{p_r}}] \end{pmatrix} = [U]_{\gamma}^{\gamma}$$

entonces:

$$[T|_{K_{p_i}}]_{\beta_i}^{\beta_i} = [U|_{K_{p_i}}]_{\gamma_i}^{\gamma_i} \qquad (*)$$

$$\therefore\ K_{p_i} = \bigoplus_{j=1}^{m_i} z_{ij}F[t] \qquad \mu(T|_{z_{ij}F[t]}, t) = p^{k_j},\ k_1 \geq \ldots \geq k_{m_i}$$

$$K'_{p_i} = \bigoplus_{j=1}^{n_i} y_{ij}F[t] \qquad \mu(U|_{y_{ij}F[t]}, t) = p^{k'_j},\ k'_1 \geq \ldots \geq k'_{n_i}$$

Y el corolario 1.0 nos asegura que $m_i = n_i$ ya que ambos son el número de bloques de la matriz (*) y demás
$S_{p_i k_j}(Kp_i) = S_{p_i k_{j'}}(K'p_i)$

$$\therefore \mu(T|_{z_{ij}F[t]}, t) = \mu(U|_{y_{ij}F[t]}, t) \quad \forall\ j \in \{1, \ldots, m_i = n_i\}.$$

Definamos $\psi_i : Kp_i \longrightarrow Kp'_i$

$$z_{i,j} \longrightarrow y_{i,j}$$

$$f(t)\circ z_{i,j} \longrightarrow f(t)\circ y_{i,j}$$

e.d. $\quad f(T)(z_{i,j}) \longrightarrow f(U)(y_{i,j})$

e.d. que

$$\{z_{i,j}, T(z_{i,j}), T^2(z_{i,j}) \ldots\} \longrightarrow \{y_{i,j}, U(y_{i,j}), U(y^2_{i,j}) \ldots\}$$

Entonces es claro que ψ_i es un isomorfismo lineal entre Kp_i y Kp'_i ya que mapea biyectivamente una base de Kp_i en una base de Kp'_i .

Además es claro que el siguiente diagrama conmuta.

$$\begin{tikzcd} Kp_i \arrow[r, "\psi_i"] & Kp'_i \\ Kp_i \arrow[u, "T"] \arrow[r, "\psi_i"'] & Kp'_i \arrow[u, "U'"'] \end{tikzcd}$$

Pues se tienen dos casos:

Caso a)

$$\begin{tikzcd} T^{\ell+1}(z_{i,j}) \arrow[rrr] & & & U^{\ell+1}(z_{i,j}) \\ & Kp_i \arrow[r, "\bar\psi_i"] & Kp'_i & \\ & Kp_i \arrow[u, "T"] \arrow[r, "\bar\psi_i"'] & Kp'_i \arrow[u, "U'"'] & \\ T^{\ell}(z_{i,j}) \arrow[uuu] \arrow[rrr] & & & U^{\ell}(z_{i,j}) \arrow[uuu] \end{tikzcd}$$

Caso b) $T^{\ell}(z_{i,j})$ es el último elemento de

$\{z_{i,j}, T(z_{i,j}), T^2(z_{i,j}) \ldots T^{\ell}(z_{i,j})\}$

e.d. $T^{\ell+1}(z_{i,j}) + a_{\ell}T^{\ell}(z_{i,j}) + \ldots + a_o z_{i,j} = 0$

e.d. $T^{\ell+1} + a_{\ell}T^{\ell} + \ldots + a_o$

es el polinomio mínimo para $T|_{z_{i,j}} F[t]$

entonces

$$\begin{tikzcd} -\sum\limits_{r=0}^{\ell} a_r T^r(z_{i,j}) \arrow[rrr] & & & -\sum\limits_{r=0}^{\ell} a_r U^r(y_{i,j}) \\ & Kp_i \arrow[r, "\bar\psi_i"] & Kp'_i & \\ & Kp_i \arrow[u, "T"] \arrow[r, "\bar\psi_i"'] & Kp'_i \arrow[u, "U'"'] & \\ T^{\ell}(z_{i,j}) \arrow[uuu] \arrow[rrr] & & & U^{\ell}(z_{i,j}) \arrow[uuu] \end{tikzcd}$$

Que conmuta, pues ya observamos que

$\mu(T|_{z_{ij}}F[t], t) = \mu(U|_{y_{ij}}F[t], t)$.

Definimos $\psi \ : \ V_F \longrightarrow V_F$ con $V_F = K_{p_1} \oplus \ldots \oplus K_{p_r}$ y

$V_F = K'_{p_1} \oplus \ldots \oplus K'_{p_r}$

e.d. $\psi : K_{p_1} \oplus \ldots \oplus K_{p_r} \longrightarrow K'_{p_1} \oplus \ldots \oplus K'_{p_r}$

tal que $v_1 + \ldots + v_r \longrightarrow \psi_1(v_1) + \ldots + \psi_r(v_r)$
donde $v_i \in Kp_i$.
Es claro que ψ es un isomorfismo y que el siguiente diagrama conmuta

$$\begin{tikzcd} V_F \arrow[r, "T"] \arrow[d, "\psi"'] & V_F \arrow[d, "\psi"] \\ V_F \arrow[r, "U"] & V_F \end{tikzcd}$$

pues

$$\begin{aligned} U\psi(v_1 + \ldots + v_r) &= U(\psi_1(v_1) + \ldots + \psi_r(v_r)) \\ &= U\psi_1(v_1) + \ldots + U\psi_r(v_r) \\ &= \psi_1 T(v_1) + \ldots + \psi_r T(v_r) \\ &= \psi_1 T(v_1 + \ldots + v_r) \end{aligned}$$

■

$\Rightarrow$) Si

$$\begin{tikzcd} V_F \arrow[r, "T"] \arrow[d, "\psi"'] & V_F \arrow[d, "\psi"] \\ V_F \arrow[r, "U"] & V_F \end{tikzcd}$$

conmuta , con ψ isomorfismo
e.d. $\psi \ T \ \psi^{-1} = U$. Es claro que $f(T) = T_0 \Leftrightarrow f(U) = U_0$ pues como

$$\begin{tikzcd} V_F \arrow[r, "T"] \arrow[d, "\psi"'] & V_F \arrow[r, "T"] \arrow[d, "\psi"'] & V_F \arrow[d, "\psi"] \\ V_F \arrow[r, "U"] & V_F \arrow[r, "U"] & V_F \end{tikzcd}$$

conmuta, tenemos que $\psi \ T^2 = U^2\psi$

e.d. $\psi\ (T^2)\ \psi^{-1} = U^2$ de manera inductiva podemos demostrar que $\psi\ (T^k)\ \psi^{-1} = U^k \quad \forall\ k \in \mathbb{N}$, Así que en general :
$\psi\ (f(T))\ \psi^{-1} = f(U)$, e.d.

$\psi\ f(T) = f(U)\psi$.

$\therefore$ T, U tienen el mismo polinomio mínimo

$\mu(T,t) = p_1^{k_1}(t) \cdot \ldots \cdot p_r^{k_r}(t)$.

$\therefore\ V = K_{p_1} \oplus \ldots \oplus K_{p_r}$ donde $K_{p_i} = N(p_i^{k_i}(T))$

$= K'_{p_1} \oplus \ldots \oplus K'_{p_r}$ con $K'_{p_i} = N(p_i^{k_i}(U))$

Observemos que $K_{p_1} \xrightarrow[\psi]{-} K'_{p_1}$ es un isomorfismo

$$
\begin{tikzcd}
K_{p_1} \arrow[rrr, "T|_{K_{p_1}}"] \arrow[dddd, "\psi|_{K_{p_1}}"'] \arrow[dr] & & & K_{p_1} \arrow[dddd, "\psi"] \arrow[dl] \\
& V_F \arrow[r, "T"] \arrow[d, "\psi"'] & V_F \arrow[d, "\psi"] & \\
& V_F \arrow[r, "U"] & V_F & \\
& & & \\
K'_{p_1} \arrow[rrr, "U"'] \arrow[uur] & & & K'_{p_1} \arrow[uul]
\end{tikzcd}
$$

$x \in K_{p_1}$ entonces $p_1^{k_1}(T)(x) = 0 \ \Rightarrow \psi(p_1^{k_1}(T))(x) = 0$

$= p_1^{k_1}(U)\psi(x)$

$\Rightarrow \psi(x) \in N(p_1^{k_1}(U)) = K'_{p_1}$

$\therefore\ K_{p_1} \xrightarrow[\psi|_{K_{p_1}}]{-} K'_{p_1} \quad \therefore\ \dim(K_{p_1}) \le \dim(K'_{p_1})$

por simetría $\dim(K'_{p_1}) \le \dim(K_{p_1}) \quad \therefore\ \dim(K_{p_1}) = \dim(K'_{p_1})$

$\therefore\ K_{p_1} \xrightarrow[\psi|_{K_{p_1}}]{-} K'_{p_1}$ es isomorfismo y además

$$\begin{array}{ccc} Kp_1 & \xrightarrow{T_{|Kp_1}} & Kp_1 \\ {\scriptstyle \psi_{|Kp_1}}\downarrow & & \downarrow{\scriptstyle \psi} \\ Kp_1' & \xrightarrow[U]{} & K'_{p_1} \end{array}$$ conmuta.

Es decir hemos reducido el problema a demostrar que si

$$\begin{array}{ccc} V_F & \xrightarrow{T} & V_F \\ {\scriptstyle \psi}\downarrow & & \downarrow{\scriptstyle \psi} \\ V_F & \xrightarrow{U} & V_F \end{array}$$ conmuta con ψ isomorfismo

y $\mu(T,t) = p^{\ell}(t)$ con $p \in F[t]$ irreducible entonces T y U tienen la misma forma canónica racional.

T induce como descomposición $V = \bigoplus_{i=1}^{m} z_i F[t]$ con

$\mu(T_{|z_i F[t]}, t) = p^{k_i}$ $\quad k_1 \geq . . . \geq k_m$

U induce como descomposición $V = \bigoplus_{i=1}^{m} y_i F[t]$

$\mu(U_{|y_i F[t]}, t) = p^{\ell_i}$.

pero $m = \frac{1}{\text{grad}(p)} \dim(p\text{-zoc}(V)) = \frac{1}{\text{grad}(p)} \dim(N(p(T)))$

además $N(p(T)) \cong N(p(U))$ pues como $\psi f(T) = f(U)\psi \quad \forall f \in F[t]$

entonces $x \in N(p(T)) \Rightarrow p(T)(x) = 0$

$\Rightarrow \psi p(T)(x) = 0$

$\Rightarrow p(U)(\psi(x)) = 0$

$\Rightarrow \psi(x) \in N(p(U))$

$\therefore \psi_{|N(p(T))} : N(p(T)) \rightarrowtail N(p(U))$ es inyectiva

Simétricamente $\psi^{-1}_{|N(p(T))}: N(p(U)) \rightarrowtail N(p(T))$ es inyectiva

$\therefore \dim N(p(T)) = \dim N(p(U))$

$\therefore \psi_{|N(p(T))} : N(p(T)) \rightarrowtail N(p(U))$ es un isomorfismo

$\therefore\ m = \frac{1}{\text{grad}(p)}\dim(N(p(T))) = \frac{1}{\text{grad}(p)}\dim(N(p(U)) = n$.

Sea $S_{p,k}(V)$ el número de sumandos con polinomio mínimo correspondiente p^k en $V = \overset{m}{\underset{i=1}{\oplus}} z_i F[t]$

y $S'_{p,k}(V)$ el número análogamente definido para la descomposición $V = \overset{m}{\underset{i=1}{\oplus}} y_i F[t]$

Como $S_{p,k}(V) = \frac{1}{\delta}\ (\dim(p^{k-1}(T)(v) - 2\dim(p^k(T)(v)) + \dim(p^{k+1}(T)(v)$

y $S_{p,k}(U) = \frac{1}{\delta}\ (\dim(p^{k-1}(U)(v) - 2\dim(p^k(U)(v)) + \dim(p^{k+1}(U)(v)$, para demostrar que son iguales basta demostrar que $\dim(p^k(U)(v)) = \dim(p^k(T)(v))\ \ \forall\ k \in \{o, \ldots , \ell\}$.

Pero ésto se sigue de que $\psi f(T) = f(U)\psi\ \ \forall\ f \in F[t]$:

Sea $x \in p^k(T)(v)$ entonces $x = p^k(T)(w)$

$= p^k(U)\psi(w) \in p^k(U)(v)$

$\therefore\ p^k(T)(v) \xrightarrow{\psi} p^k(U)(v)$ es lineal e inyectiva.

Simétricamente $p^k(U)(v) \xrightarrow{\psi^{-1}} p^k(T)(v)$ es lineal e inyectiva

$\therefore\ \dim_F(p^k(T)(v)) = \dim_F(p^k(U)(v))$

$\therefore\ S_{p,k}(V) = S'_{p,k}(V)$ y por el corolario 2.0 T y U tienen la misma forma canónica racional.

∎

TEOREMA-7.0) Sea $T : V_F \longrightarrow V_F$ un operador lineal (con $F = \mathbb{C}$) y sea $[T]_\beta^\beta$ con $\beta \underset{b.\,ord.}{\hookrightarrow} V_F$, entonces: $\lim_{n\to\infty} [[T]_\beta^\beta]^n$ existe si y sólo si se satisfacen las condiciones siguientes.

1) Si λ es un eigenvalor de $[T]_\beta^\beta$ entonces $|\lambda| \leq 1$.

2) Si λ es un eigenvalor de $[T]_\beta^\beta$ tal que $|\lambda| = 1$, entonces λ es el número real 1.

3) Si 1 es un eigenvalor de $[T]_\beta^\beta$, entonces la dimensión del eigenespacio correspondiente a 1 es igual a la multiplicidad de 1 como eigenvalor de A. (igual a la multiplicidad de 1 como raíz de $X(T,t)$).

Antes de proceder la demostración haremos algunas observaciones.

Sea $\mu(T,t) = p_1^{k_1}(t) \cdot \ldots \cdot p_\ell^{k_\ell}(t)$ el polinomio mínimo de T (con p_i irreducible), puesto que $F = \mathbb{C}$ entonces sabemos que $p_i = (t-\lambda_i)$. y por el teorema de descomposición para espacios vectoriales sabemos que existe $\gamma \underset{b.\,ord.}{\hookrightarrow} V_F$ tal que $[T]_\gamma^\gamma = J_{[T]}$ (forma canónica de Jordan) además $[T]_\beta^\beta = Q\,[T]_\gamma^\gamma Q^{-1}$, ($Q = [I]_\gamma^\beta$) y $\lim_{n\to\infty} [[T]_\beta^\beta]^n = Q\,[[T]_\gamma^\gamma]^n Q^{-1}$ por lo tanto si probamos que $\lim_{n\to\infty}[[T]_\gamma^\gamma]^n$ existe habremos probado que $\lim_{n\to\infty} [[T]_\beta^\beta]^n$ también existe.

Supongamos que $|\lambda| < 1$ y Sea J_λ un bloque de Jordan asociado al eigenvalor λ , entonces

$$J_\lambda = \begin{pmatrix} \lambda & 1 & 0 & . & . & . & 0 \\ 0 & \lambda & 1 & 0. & . & . & . \\ 0 & 0 & \lambda & 1 & . & . & \vdots \\ \vdots & & \cdot & & . & & 1 \\ 0 & & . & . & \vdots & & \lambda \end{pmatrix} \in \mathcal{M}_{m\times m}$$

hagamos $N = J_\lambda - \lambda I_m$

Observemos que $N^m = 0 \in \mathcal{M}_{m\times m}$

Demostración.

por inducción sobre m

i) $m = 2$

$J_\lambda = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ por lo que $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ y claramente $N^2 = 0 \in \mathcal{M}_{2\times 2}$

ii) Supongamos que $N^m = 0$, con $N, 0 \in \mathcal{M}_{m\times m}$

Sea $N \in \mathcal{M}_{(m+1)\times(m+1)}$ queremos probar que $N^{m+1} = 0 \in \mathcal{M}_{(m+1)\times(m+1)}$ y hagamos una $(2,(m,1),m+1)$ partición de J_λ (Ver definición 1.0 del capítulo 2) e.d.

$$N = \begin{pmatrix} \begin{pmatrix} 0 & 1 & 0 & . & . & . & 0 \\ 0 & 0 & 1 & 0. & . & . & . \\ 0 & 0 & 0 & 1 & . & . & \vdots \\ \vdots & & . & & . & & 1 \\ 0 & & . & & . & \vdots & 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 1 \end{pmatrix} \\ (0 \;\; 0 \;\; . \;\; . \;\; . \;\; 0) & (0) \end{pmatrix} \in \mathcal{M}_{(m+1)\times(m+1)}$$

e.d. $N = \begin{pmatrix} N & A \\ 0 & 0 \end{pmatrix}$ y $N^{m+1} = \begin{pmatrix} N^{m+1} & N^m A \\ 0 & 0 \end{pmatrix}$

pero $N \in \mathcal{M}_{m\times m}$ y por hipótesis de inducción $N^m = 0$

así que $N^{m+1} = N\ N^m = N\ 0 = 0$ y $N^m A = 0A = 0$

$\therefore N^{m+1} = 0 \quad \therefore \quad (J_\lambda - \lambda I_m)^m = 0$ con $J_\lambda \in \mathcal{M}_{m\times m}$ $\forall\ m \in \mathbb{N}$.

∎

Ahora $J_\lambda = \lambda I + N$ y entonces por el teorema del binomio (que se puede aplicar pues $\lambda IN = N\lambda I = \lambda N$) así tenemos que
$J_\lambda^r = \binom{r}{0}(\lambda I)^r + \binom{r}{1}(\lambda)^{r-1}N + \ldots + \binom{r}{j}(\lambda I)^{r-j} + \binom{r}{r}N^r$

Como $N = \begin{pmatrix} 0 & 1 & 0 & . & . & . & 0 \\ 0 & 0 & 1 & 0 & . & . & 0 \\ \vdots & & & . & . & & \vdots \\ \vdots & & & & & . & 1 \\ 0 & 0 & 0 & . & . & . & 0 \end{pmatrix}$

tenemos que $Ne_1 = o$, $Ne_2 = e_1$, $Ne_3 = e_2$, ... , $Ne_m = e_{m-1}$
$\therefore N^2 e_1 = o$, $N^2 e_2 = Ne_1 = o$, $N^3 e_3 = Ne_2 = e_1$, . . . , $N^m e_m = Ne_{m-1} = e_{m-2}$

$$\therefore\quad N^2 = \begin{pmatrix} 0 & 0 & 1 & . & . & . & 0 \\ 0 & 0 & 0 & 1 & . & . & \vdots \\ \vdots & & & \ddots & & . & 1 \\ \vdots & & & & & . & 0 \\ 0 & 0 & 0 & . & . & . & 0 \end{pmatrix}$$

En general $\quad N^k = \begin{array}{c} \leftarrow \ k+1 \ \rightarrow \\ \begin{pmatrix} 0 & 0 & 0 & \ldots 1 \ldots & 0 \\ 0 & 0 & 0 & 0 \ldots \ddots & \vdots \\ \vdots & & \ddots & & 1 \\ \vdots & & & \ddots & 0 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix} \end{array} \quad$ y $\quad N^m = 0.$

Así que $J_\lambda^r = \sum_{i=0}^{m-1} \binom{r}{i} \lambda^{r-i} N^i \quad$ si $r \geq m-1$

$$\therefore\quad J_\lambda^r = \begin{pmatrix} \binom{r}{0}\lambda^r & \binom{r}{1}\lambda^{r-1} & \binom{r}{2}\lambda^{r-2} & \ldots & \binom{r}{m-1}\lambda^{r-m+1} \\ & \ddots & & & \vdots \\ & & \ddots & & \binom{r}{2}\lambda^{r-2} \\ & 0 & & \ddots & \binom{r}{1}\lambda^{r-1} \\ & & & & \binom{r}{0}\lambda^r \end{pmatrix}$$

Demostración del Teorema.

Demostraremos que $\lim J^n$ existe $\Leftrightarrow$ valen i) ii) y iii) para λ

$\Rightarrow$) Como los coeficientes de la diagonal principal de J^r son λ^r tenemos que $\lim\limits_{r\to\infty} J^r$ existe $\Rightarrow \lim\limits_{r\to\infty} \lambda^r$ existe $\Rightarrow |\lambda| \leq 1$. Además $\lim\limits_{r\to\infty} \lambda^r$ existe con $|\lambda| = 1$ sólo si $\lambda = 1$. (Pues si $\lambda = e^{i\theta}$ $0 < \theta < 2\pi$, si $\theta \neq 0$ entonces $\lambda^r = e^{ir\theta}$ y $\therefore$ $|\lambda^{r+1} - \lambda^r| = |e^{i(r+1)\theta} - e^{ir\theta}| = |e^{ir\theta+i\theta} - e^{ir\theta}| = |e^{ir\theta}(e^{i\theta} - 1)| = |e^{i\theta} - 1| > 0$ $\therefore$ $\{\lambda^r\}_{r\in\mathbb{N}}$ no es Cauchy $\therefore$ $\{\lambda^r\}_{r\in\mathbb{N}}$ no converge). Con lo cual queda demostrado i) y ii) .

Si $\lambda = 1$ entonces

$$J_\lambda^r = \begin{pmatrix} \binom{r}{0} & \binom{r}{1} & \binom{r}{2} & \cdots & \binom{r}{m-1} \\ & \ddots & & & \vdots \\ & & \ddots & & \binom{r}{2} \\ & & & \ddots & \binom{r}{1} \\ & & & & \binom{r}{0} \end{pmatrix}$$

y como $\binom{r}{1} = r$ que tiende a infinito cuando r tiende a infinito, entonces $\lim_{r \to \infty} J^r$ existe $\Rightarrow$ $J = (\binom{r}{0}) = (1) \in M_{1\times 1}(\mathbb{C})$

$\therefore$ cada bloque de Jordan correspondiente a 1 es (1).

Si $\mu(T,t) = (t-1)^{k_1} \cdot \ldots \cdot (t-\lambda_\ell)^{k_\ell}$ entonces $K_1 = N(T - I)^{k_1}$ tiene una base canónica de Jordan β_1 tal que

$$[T|_{K_1}]_{\beta_1}^{\beta_1} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

$\therefore$ cada elemento de β_1 es un vector propio de T correspondiente a 1 $\therefore$ $\beta_1 \hookrightarrow E_1 = \{x \in V_F \mid T(x) = x\} \leq K_1$ $\therefore$ $K_1 = \mathcal{L}(\beta_1) \leq E_1 \leq K_1$

$\therefore$ $K_1 = E_1$ $(m_1 = 1)$.

Ahora si β es una base canónica de Jordan entonces

$$[T]_\beta^\beta = \begin{pmatrix} \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} & & & 0 \\ & \begin{pmatrix} \begin{pmatrix} \lambda & 1 & 0 \\ & \ddots & 1 \\ 0 & & \lambda \end{pmatrix} & & 0 \\ & \ddots & \\ 0 & & \begin{pmatrix} \lambda & 1 & 0 \\ & \ddots & 1 \\ & & \lambda \end{pmatrix} \end{pmatrix} & & \\ & & \ddots & \\ 0 & & & \begin{pmatrix} \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \ddots & 1 \\ & & \lambda \end{pmatrix} & & 0 \\ & \ddots & \\ 0 & & \begin{pmatrix} \lambda & 1 & 0 \\ 0 & & 1 \\ & & \lambda \end{pmatrix} \end{pmatrix} \end{pmatrix}$$

Es una matriz triangular superior de donde se vé que:

$\chi(T,t) = (t-1)^{m_1} \cdot \ldots \cdot (t-\lambda_k)^{m_k}$ así que m_1 = multiplicidad de 1 como raiz de $\chi(T,t) = \dim(K_1) = \dim(E_1)$.

con lo cual tenemos iii)

$\blacksquare$

$\Leftarrow$) Supongamos que λ es un eigenvalor de T tal que satisface i), ii) y iii). Demostraremos que $\lim_{r\to\infty} J^r$ existe.

caso a) $\lambda = 1$

en este caso tenemos que dim (E_1) = multiplicidad de 1 como raíz de $\chi(T,t)$.

Consideremos la forma canónica de Jordan de $T|_{K_1}$ donde $K_1 = \{x \in V_F \mid (T - I)^s x = 0$ p.a. $s \in \mathbb{N}\}$, que en principio es de la forma

$$\begin{pmatrix} \begin{pmatrix} 1 & 1 & & 0 \\ & 1 & 1 & \\ & & \ddots & 1 \\ 0 & & & 1 \end{pmatrix} & & & 0 \\ & \begin{pmatrix} 1 & 1 & & & \\ & 1 & \ddots & & 0 \\ & & \ddots & 1 & \\ & & & 1 & 1 \\ 0 & & & & 1 \end{pmatrix} & & \\ & & \ddots & \\ 0 & & & \begin{pmatrix} 1 & 1 & & 0 \\ & 1 & 1 & \\ & & \ddots & 1 \\ 0 & & & 1 \end{pmatrix} \end{pmatrix} \qquad [*]$$

ahora cada 1 en la diagonal contribuye en un factor $(1 - t)$ a $X(T,t)$, $\therefore$ el número de 1 en la diagonal de la matriz (*) es la multiplicidad de 1 como raiz de $X(T,t)$. Es claro que el número de 1 en la diagonal de (*) también es la dimensión de K_1 $\therefore$ $\dim(K_1)$ = número de 1 en la diagonal de (*) que es igual a la multiplicidad de 1 como eigenvalor de T que es igual (por hipótesis) a $\dim(E_1)$, $\therefore \dim(K_1) = \dim(E_1)$ y como $E_1 \leq K_1$ tenemos que $E_1 = K_1$, por lo que cada elemento de K_1 es un eigenvector correspondiente a 1 $\therefore$ la forma canónica de Jordan de $T|_{K_1}$ es: $\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ & & \ddots & 0 \\ 0 & & & 1 \end{pmatrix}$

es decir cada bloque de Jordàn es (1) $\in M_{1\times 1}(\mathbb{C})$, y es claro que $\lim_{r \to \infty} J_1^r = \lim_{r \to \infty} (1)^r = (1)$.

caso b) Si $|\lambda| < 1$ demostraremos que $\lim_{r \to \infty} J^r = 0$, demostrando que $\lim_{r \to \infty} \binom{r}{i}\lambda^{r-i} = 0$.

Para demostrar que $\lim_{r \to \infty} \binom{r}{i}\lambda^{r-i} = 0$ basta demostrar que $\lim_{r \to \infty} |\binom{r}{i}\lambda^{r-i}| = \lim_{r \to \infty} \binom{r}{i}|\lambda|^{r-i} = 0$. En otras palabras podemos suponer $\lambda \in \mathbb{R}$ con $0 \le \lambda < 1$.

$\binom{r}{i}\lambda^{r-i} = \dfrac{r!}{i!\,(r-i)!}\ \lambda^{r-i}$ que tiende a cero si y solo si

$\dfrac{r!}{(r-i)!}\lambda^{r-i}$ tiende también a cero.

$$\frac{r!}{(r-i)!}\ \lambda^{r-i} = r(r-1)\ \ldots\ (r-i+1)\ \lambda\ \ldots\ \lambda = S_r$$

$$S_{r+1} = (r+1)r(r-1)\ \ldots\ (r-i+1)\ \lambda\ \ldots\ \lambda\ \lambda = \frac{(r+1)}{(r-i+1)}\ \lambda\ S_r$$

fijemos $r \in \mathbb{N}$. entonces

$$S_{N+1} = \frac{(N+1)}{(N-i+1)}\ \lambda\ S_N$$

$$S_{N+2} = \frac{(N+2)\qquad(N+1)}{(N-i+2)\ (N-i+1)}\ \lambda^2\ S_N$$

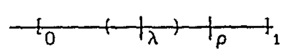
$$\vdots$$

$$S_{N+j} = \frac{(N+j)}{(N-i+j)}\ \cdot\ \cdot\ \cdot\ \frac{(N+2)\qquad(N+1)}{(N-i+2)\ (N-i+1)}\ \lambda^j\ S_N$$

Ahora $\lim_{j \to \infty} \dfrac{(N+j)}{(N-i+j)} = \lim_{j \to \infty} \dfrac{(N/j+1)}{(N/j - i/j + 1\ \ )} = 1$

$\therefore \lim_{j \to \infty} \dfrac{(N+j)}{(N-i+j)}\ \lambda = \lambda \qquad 0 \le \lambda < 1$

escojamos una ρ tal que $\lambda < \rho < 1$



Existe K tal que si $j > K$ entonces

$\dfrac{(N+j)}{(N-i+j)}\ \lambda$ cae dentro de la vecindad de λ marcada en la figura anterior, es decir que $0 \le \dfrac{(N+j)}{(N-i+j)}\ \lambda < \rho$

$\therefore$ si $\quad j > K$

$$S_{n+j} = \underset{\underset{\rho}{\wedge}}{\left(\frac{(N+j)}{(N-i+j)}\lambda\right)} \cdots \underset{\underset{\rho}{\wedge}}{\left(\frac{(N+K+1)}{(N-i+K+1)}\lambda\right)} \underbrace{\left(\frac{(N+K)}{(N-i+K)}\lambda\right) \cdots \frac{(N+1)}{(N-i+1)}\lambda \quad S_N}_{\substack{\text{constante } =:\ C \\ C\, >\, 0}}$$

$$< C\,\rho^{j-k} = \frac{C}{\rho^{k}}\,\rho^{j} \quad \text{donde} \quad \frac{C}{\rho^{k}} =:\ C_1 \in \mathbb{R}^{+}.$$

$\therefore\ 0 \le \lim_{r\to\infty} S_r = \lim_{j\to\infty} S_{N+j} \le \lim_{j\to\infty} C_1\,\rho^{j} = 0$ pues $0 \le \rho < 1$

$\therefore\ \lim_{r\to\infty} S_r = 0.$

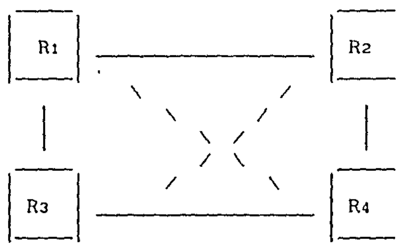
Con lo que se termina la prueba del teorema.

■

MATRICES NO NEGATIVAS

Introduzcámonos a la teoría de matríces no negativas por medio de dos ejemplos que a continuación presentaremos, tomados de la referencia 6 dada en la bibliografía.

Supongamos que una sociedad que consta de cuatro regiones, R_1, R_2, R_3 y R_4 comercian entre sí con cierto articulo no renovable. R_i le vende uno de sus productos a la industria R_j, la industria R_j no lo podrá regresar, pero sí vender sus demas productos.



Los porcentajes de embarque de una región a otra fueron los que a continuación se muestran en la siguiente matriz :

$$C = \begin{pmatrix} 0.1 & 0.3 & 0.0 & 0.0 \\ 0.7 & 0.6 & 0.0 & 0.0 \\ 0.1 & 0.0 & 1.0 & 0.0 \\ 0.1 & 0.1 & 0.0 & 1.0 \end{pmatrix}$$

en donde la entrada c_{ij} es la proporción del articulo presente en R_j que será embarcado cada dia a R_i . Supongamos que las coordenadas de (v_1, v_2, v_3, v_4) son las cantidades del articulo presente hoy en R_1, R_2, R_3 y R_4 respectivamente, a nosotros nos interesaría preguntarnos como será la distribución del producto dos semanas o un mes después a partir de hoy. Para mañana la respuesta está dada por el producto de Cv, así que en dos semanas la respuesta es $C^{14}v$ y en general $C^n v$ nos daría la distribución para n días despues a partir de hoy.

Haciendo uso de los primeros resultados del capítulo anterior podemos pensar a C de la siguiente manera:

$$C = \begin{pmatrix} X & 0 \\ Y & I \end{pmatrix}$$ que es una (2, (2 , 2), 4) partición de C

en donde $X = \begin{pmatrix} 0.1 & 0.3 \\ 0.7 & 0.6 \end{pmatrix}$ $\qquad Y = \begin{pmatrix} 0.1 & 0.0 \\ 0.1 & 0.1 \end{pmatrix}$

$$0 = \begin{pmatrix} 0.0 & 0.0 \\ 0.0 & 0.0 \end{pmatrix} \qquad I = \begin{pmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \end{pmatrix}$$

entonces

$$\underbrace{C\cdot C\cdot C\cdot \;.\;.\;.\; C\cdot C}_{n \text{ veces}} = C^n = \begin{pmatrix} X^n & 0 \\ Y \sum_{j=0}^{n-1} X^j & I \end{pmatrix}$$

Y de aqui si evaluáramos en alguna n deseada nos daríamos cuenta de como sería el embarque de una industria a otra.

<u>DEFINICION-13.0</u>)
Sea A una matriz cuadrada, entonces se define el *Radio Espectral* de A, como máx $\{ \, |\lambda| \;|\; \lambda$ es eigenvalor de $A \, \}$ denotado por $|A|$.

Para propósitos de nuestro problema supondremos que el radio espectral es un eigenvalor lo que demostraremos más adelante.

<u>LEMA-2.0</u>)

Sea A una matriz cuadrada (k x k) sobre un cierto campo y sea $|A|$ su radio espectral entonces $|A| \leq \max\limits_{1<j<k} \sum_{j=1}^{k} |A_{ij}| \qquad A_{ij} \in \mathbb{R}$

Demostración:

Sea $\bar{x} = (x_1, \ldots, x_k)$ un eigenvecor asociado a $|A|$ y sea $x^* = |x_r|$ donde $|x_r| = \max\{|x_i| \text{ con } 1 \leq i \leq k\}$
entonces:

$$|A|x^* = |A||x_r| = |(A\,x)_r| = \left|\sum_{j=1}^{k} A_{rj}\,x_j\right| \leq \sum_{j=1}^{k} |A_{rj}\,x_j|$$

$$\leq \sum_{j=1}^{k} |A_{rj}||x_r| = \sum_{j=1}^{k} |A_{rj}|\,x^*$$

$$\therefore\ |A| \leq \sum_{j=1}^{k} |A_{rj}| \qquad \blacksquare$$

Calculando $|X|$ tenemos que $|X| = .9 < 1$ entonces tenemos que $\lim_{n\ \infty} X^n = 0$, por el teorema 7.0 y aseguramos que $I - X$ es invertible, pues si no lo fuera entonces:

$$\text{Nuc}(\ I - X\ ) \neq \{\ 0\ \}$$

e.d. $\exists\ \bar{x} \neq \bar{0}$ tal que $(\ I - X\ )\bar{x} = \bar{0}$

$$(\ I - X\ )\bar{x} = I\bar{x} - X\bar{x}$$

$$I\bar{x} = X\bar{x} = \bar{x}$$

$\therefore$ 1 es un eigenvalor de X asociado a $\bar{x}$ lo cual es una contradicción, pues $|X| < 1$; y de aquí que $(\ I - X\ )$ sea invertible.

<u>OBSERVACION-7.0</u>)

Sea X una matriz cuadrada $(k \times k)$ sobre un cierto campo F

tal que $|X| < 1$ entonces $\quad I = (I - X)\sum_{n=0}^{\infty} X^n$

Prueba:

$$(I - X)\sum_{n=0}^{\infty} X^n$$

$$= \sum_{n=0}^{\infty} X^n - \sum_{n=0}^{\infty} X^{n+1}$$

$$=\lim_{n\ \infty} (X^0 + X^1 + \ldots + X^n - X^1 - X^2 - \ldots - X^n - X^{n+1})$$

$$=\lim_{n\ \infty} X^0 - \lim_{n\ \infty} X^{n+1} \quad (\lim_{n\ \infty} X = \theta \text{ pues } |X| < 1, \text{ teorema 7.0})$$

$$= I - 0$$

$$= I$$

■

Y como (I - X) es invertible entonces tenemos que:

$$(I - X)^{-1} = \sum_{n=1}^{\infty} X^n$$

Por consiguiente haciendo los cálculos necesarios encontramos que:

$$\sum_{n=0}^{\infty} X^n = (I - X)^{-1} = \begin{pmatrix} 0.9 & -0.3 \\ -0.7 & 0.4 \end{pmatrix}^{-1} = \begin{pmatrix} 8/3 & 2 \\ 14/3 & 6 \end{pmatrix}$$

concluyendo que $\lim_{n\ \infty} C^n = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 4/15 & 1/15 & 1 & 0 \\ 11/15 & 4/5 & 0 & 1 \end{pmatrix}$

y haciendo el producto del limite por el vector de distribución inicial tenemos :

$$\lim_{n\ \infty} C^n\ v = \begin{pmatrix} 0 \\ 0 \\ ((4v_1)/5) + ((v_2)/5) + v_3 \\ ((11v_1)/15) + ((4v_2)/5) + v_4 \end{pmatrix}$$

Así que con el tiempo R_3 y R_4 tendrán acumulado todo el producto.

Tratemos ahora el ejemplo conocido como el problema del subsidio. Sucede que en cierta ciudad industrial existen m industrias, en donde la producción de cada una de ellas depende del producto de las otras m-1 industrias y de ella misma. El gobierno ha decidido destinar una cantidad de dinero para subsidiar a cada una de ellas. Por otra parte a las industrias les gustaría que el subsidio fuera repartido equitativamente, e.d. que ninguna de ellas gaste más de lo que recibe de subsidio para costo de su producción. Por otra parte el gobierno quiere que la producción de cada una de ellas sea tan grande como sea posible, por lo que ellas están de acuerdo en cumplir con cuotas para llevar a cabo una distribución equitativa de costos.

Dado que la i-ésima industria le paga a la j-ésima industria c_{ij} pesos por unidad de su propia producción suponemos que $c_{ij} > 0$ la pregunta sería ¿puede ser repartido equitativamente el subsidio? ¿y en que medida esta posibilidad depende de los costos relativos c_{ij} ?

Veamos más de cerca el problema, supongamos que la i-ésima industria produce u_i unidades anuales. La j-ésima industria paga un total de $\sum_{i=1}^{k} c_{ij}u_j$ pesos al año a las demás industrias y por otra parte esta recibe un total de :

$\sum_{i=1}^{k} c_{ji}u_i$ pesos del subsidio anual en pago de las demás industrias. Así que su costo de producción subsidiado anual ν_j está dado por

$\nu_j = \gamma_j u_j - \sum_{i=1}^{k} c_{ji}u_i$ donde $\gamma_j = \sum_{i=1}^{k} c_{ij}$. Aparte para que ésto sea equitativo debe suceder que para toda j $\nu_j \leq 0$, pero esto implica que $\nu_j = 0$. Así que nuestro problema se reduce a encontrar un sistema de producción de cuotas u_1 , u_2 , ... , u_k tal que (puesto que queremos que que $\nu_j = 0$)

$$0 = \gamma_j u_j - \sum_{i=1}^{k} c_{ji} u_i$$

$$\Rightarrow \quad \gamma_j u_j = \sum_{i=1}^{k} c_{ji} u_i$$

$$\Rightarrow \quad u_j = \sum_{i=1}^{k} \frac{c_{ji}\ u_i}{\gamma_j} \qquad \text{con } j = 1, 2, \ldots, k \text{ ------------(1)}$$

y sujeto a que cada $u_j > 0$ con $i = 1, 2, \ldots, k$ ---------(2)

y (3) que $\sum_{j=1}^{k} u_j \geq \sum_{j=1}^{k} x_j$ para cualquier x, tal que $x_j > 0$ y

$$x_j = \sum_{i=1}^{k} \frac{c_{ji}\ x_i}{\gamma_j} \qquad j = 1, 2, \ldots, k.$$

La ecuación (1) nos asegura que el sistema sea equitativo, la (2) que el sistema sea factible y la (3) que sea óptimo.

De nuevo la pregunta sigue siendo si existirá solución y de ser así, cuál solución será factible. (1) puede ser escrito como un sistema homogeneo de k ecuaciones lineales con k incógnitas y desde luego esperaríamos soluciones distintas de cero si el rango del sistema fuera menor que k, ahora bien, si hubieran soluciones distintas de cero ¿qué nos aseguraria que hay una que cumpla (2)? Pareciera que la solución está en función de los costos c_{ij}.

En 1905 el matemático alemán *O. Perron* probó un teorema básico acerca de matrices positivas y una parte fuerte que el teorema garantiza es que si A es una matriz real positiva entonces $Ax = |A|x$ tiene una solución x con cada $x_i > 0$.

Así que si $a_{ji} = c_{ji} / \gamma_j$ entonces

$$\sum_{i=1}^{k} \frac{c_{ji}}{\gamma_j} x_i = |A| x_j \quad \text{para cualquier } j = 1, 2, \dots, k.$$

Por otra parte haciendo $g = (\gamma_1, \gamma_2, \dots, \gamma_k)^{tr}$ tenemos que:

$g^{tr}A = g^{tr}$; $\therefore g^{tr}x = g^{tr}Ax = g^{tr}|A|x = |A|g^{tr}x$ y como el escalar

$g^{tr}x \neq 0$ entonces se tiene que $|A| = 1$. También cualquier vector de la forma τx con $\tau > 0$ es una solución a la condición (1) además de que es factible. El teorema de *O. Perron* también nos asegura que el espacio de soluciones de $Ax = |A|x$ es unidimensional y porlo tanto las únicas soluciones que cumplan (1) y sean factibles son de la forma rx con $r > 0$.

$\sum_{i=1}^{k} \tau x_i$ será la producción total bajo el sistema de cuotas dado

por τx y el costo total de producción bajo este sistema es

$\sum_{j=1}^{k} \gamma_j(\tau_j)$ que no puede exceder a σ, que es el subsidio dado por el

gobierno, así que si tomamos al sistema de cuotas $u_i = \dfrac{\sigma}{\sum_{j=1}^{k} \gamma_j x_j} x_i$

con $i = 1, 2, \dots, k$. obtenemos una solución que cumple (1), es factible y además es óptima.

En este sistema la i-ésima industria recibe $\gamma_i u_i$ como su parte del subsidio total.

Es por eso que el teorema nos da la respuesta: el problema del subsidio siempre puede ser resuelto y además de manera única, no importando cuales sean los costos c_{ij}.

Notemos que en los ejemplos anteriores se ha tratado con matrices que no tienen ninguna entrada negativa, así que las definiremos de la siguiente manera.

DEFINICION-14.0)
Sea $A \in M_{k\times k}(\mathbb{R})$, A es no-negativa si $A_{ij} \geq 0 \quad \forall\, i,\ j$. con $i, j \in \{1,\dots,k\}$.

Tambien podemos comparar matrices en el sentido de decir quien es mas "grande" y ésto aceptando la siguiente convención. Si M , N son matrices cuadradas (kxk) con entradas en un cierto campo F entonces decimos que $M \geq N$ si $M_{ij} \geq N_{ij} \ \forall\, i,j$ con $1 \leq i,j \leq k$

y $\qquad M > N$ si $M_{ij} > N_{ij} \ \forall\, i,j$ con $1 \leq i,j \leq k$

Con ésto podemos decir ahora que una matriz A es no negativa si $A \geq 0$ y A es positiva si $A > 0$. Donde 0 es la matriz con entradas de puros ceros.

Por ejemplo $A = \begin{pmatrix} 3 & 9 & 4 \\ 6 & 5 & 1 \\ 4 & 8 & 2 \end{pmatrix}$ es una matriz positiva mientras

que $B = \begin{pmatrix} 0 & 4 & 3 \\ 5 & 9 & 1 \\ 4 & 0 & 1 \end{pmatrix}$ es una matriz no negativa.

OBSERVACION-8.0) Sea $A > 0$ y $x > 0$ entonces $Ax > 0$.

Demostración:

$Ax = y$ hay que demostrar que $y > 0$

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{pmatrix} \quad \text{y} \quad y_i = \sum_{j=1}^{k} A_{ij} x_j$$

$A_{ij} > 0 \ \forall\, i,j$ y puesto que $x > 0$ entonces existe al menos una j tal que $x_j > 0 \quad \therefore \ A_{ij}x_j > 0$

y de aquí que $y_i = \sum_{j=1}^{k} A_{ij}x_j \geq A_{ij}x_j > 0$ para toda i

$\therefore \quad y > 0$.

∎

N.B. Tomemos la convención de que salvo que se digamos otra cosa, unicamente trabajaremos con matrices cuadradas de orden k, con coeficientes en $\mathbb{C}$.

OBSERVACION-9.0)
Sea A una matriz positiva y sean z, w vectores tales que $z \geq w$, entonces :

i) $Az \geq Aw$

ii) $Az = Aw \Leftrightarrow z = w$

Demostración

Sea $z' = Az$ y $w' = Aw$

p.d. que $z' \geq w'$

$z' = (z'_1 , z'_2 , \ldots z'_k)$ $w' = (w'_1 , w'_2 , \ldots , w'_k)$

$$z'_i = \sum_{j=1}^{k} A_{ij} z_j \qquad w' = \sum_{i=1}^{k} A_{ij} w_j$$

pero $z_j \geq w_j$ $\forall\ j \in \{1,2, \ldots , k\}$
entonces $A_{ij} z_j \geq A_{ij} w_j$ $\forall\ j \in \{1,2, \ldots , k\}$ pues A es positiva

$$\therefore \sum_{j=1}^{k} A_{ij} z_j \geq \sum_{i=1}^{k} A_{ij} w_j \qquad \therefore z' \geq w'$$

y $\therefore$ $Az \geq Aw$.

también es claro $Az = Aw \Leftrightarrow z = w$.

■

LEMA-3.0)

(*Wielandt*) Sea A una matriz positiva entonces:

i) $|A|$ es un eigenvalor positivo de A, teniendo asociado un eigenvector positivo.

ii) El módulo de cualquier otro eigenvalor de A es mas chico que el radio espectral .

Observemos que $|\lambda| = \max\{|\lambda| \ / \ \lambda$ es eigenvalor de $A\}$ entonces $\lambda \in \mathbb{R}$.
Pues si λ es tal que $|\lambda| = |A|$ entonces $\lambda = |A|$
y por lo tanto $\lambda = |\lambda| \Leftrightarrow \lambda \in (0, \infty)$ y aún algo más
si $A > 0$ entonces A admite un eigenvalor $\lambda \in \mathbb{R}$ con $\lambda > 0$.

Demostración del lema de Wielandt:

Supongamos que μ es un eigenvalor de módulo máximo asociado a un eigenvector $x \neq 0$, entonces tenemos que $|\mu| = |A|$ y tambien que $Ax = \mu x$

definamos $p = (|x_1|, |x_2|, \dots, |x_k|)$

sabemos que $\mu x_i = \sum_{j=1}^{k} A_{ij} x_j$ para cada i

$$|A| p_j = |A| |x_j| = |(Ax)_j| = |\sum_{j=1}^{k} A_{ij} x_j| \leq \sum_{k=1}^{k} |A_{ij} x_j| =$$

$$= \sum_{j=1}^{k} A_{ij} |x_j| = \sum_{j=1}^{k} A_{ij} p_j$$

$$\therefore \quad |A| p \leq Ap$$

De aquí se sigue que $(Ap - |A|p) \geq 0$ por lo que para demostrar que $|A|$ es un eigenvalor tenemos que probar que se da la igualdad.

Razonemos por contradicción, es decir, supongamos que $(A - |A|)p > 0$ y que $z = (A - |A|I)p$

entonces tenemos que $Az > 0$ y $Ap > 0$ por la observación 2.0 dado ésto podemos encontrar un $\varepsilon > 0$ tan pequeña como se desée de tal forma que:

$Az \geq \varepsilon Ap$, sabemos que $Az = [A\,(A - |A|I)]p$

entonces:

$A^2p - |A|Ap \geq \varepsilon\, Ap$

$A^2p \geq \varepsilon\, Ap + |A|Ap$

$A^2p \geq (\varepsilon + |A|)Ap$ como $\varepsilon > 0$ y $|A| > 0 \Rightarrow \dfrac{1}{\varepsilon + |A|} > 0$

$(\varepsilon + |A|)^{-1}A^2p \geq Ap$ y ésto es que :

$[(\varepsilon + |A|)^{-1}A\,]\,Ap \geq Ap$

Sea $B = [(\varepsilon + |A|)^{-1}A\,]$ es claro que $B > 0$

entonces $B\,Ap \geq Ap$ de donde se sigue que

$B^{n+1}Ap \geq Ap \quad \forall\, n \geq 1$

pero $|B| = |(\varepsilon + |A|)^{-1}A| = |\,(\varepsilon + |A|)^{-1}|\,|A| = \dfrac{|A|}{|\,\varepsilon + |A||} < 1$

$\therefore \lim\limits_{n\,\infty} B^n = 0$ entonces $\lim\limits_{n\,\infty} B^nAp \geq Ap$

de donde se ve que $0 \geq Ap$ lo cual es una contradicción pues $Ap > 0$.

y de aquí se infiere que $(A - |A|I)p = 0$ y por lo tanto queda demostrado que $|A|$ es un eigenvalor de A asociado a p que es un eigenvector positivo.

∎

Probemos ahora la parte dos que equivale a decir que $|A|$ es el único eigenvalor de tamaño $|A|$.

Supongamos que no es único, es decir, que existe λ eigenvalor asociado a un eigenvector positivo y.

entonces tenemos que $|\lambda| = |A|$

hagamos de nuevo $q = (|y_1|\,, |y_2|\,, \ldots, |y_k|)$

Aplicando el mismo razonamiento a q y a λ tenemos que

$$|A|q = Aq$$

de donde $\left| \sum_{j=1}^{k} A_{ij} y_j \right| = \sum_{j=1}^{k} A_{ij} |y_j|$

y esta desigualdad se da sólo si los argumentos de los y_i's son iguales, es decir que si escribimos a cada y_i en coordenadas polares tendríamos que: $y_j = |y_j| e^{i\theta}$ con θ real e independiente de j, sucediendo ésto para todo valor de j, pero $|y_j| = q_j$ $\therefore$ $y = qe^{i\theta}$ de donde vemos que y es un múltiplo del eigenvector asociado a $|A|$, así que y tambien está asociado a $|A|$ y como y está asociado al eigenvalor λ concluímos que $\lambda = |A|$.

∎

En nuestros cursos de Algebra Lineal aprendimos que para encontrar el límite M^n para alguna matriz M nos es de gran ayuda encontrar sus eigenvalores y diagonalizarla si es posible, ya que analizando el tamaño de sus eigenvalores podemos afirmar si el límite de M^n existe o no. Utilizando estos hechos probemos en el siguiente teorema que $(|A|^{-1}A)^n$ converge.
(ver tambien el capítulo de la forma racional y de Jordan).

TEOREMA-8.0)
Sea A una matriz cuadrada y $|A| \neq 0$ su radio espectral entonces $(|A|^{-1}A)^n$ es convergente.

Demostración:
Sea $B = (|A|^{-1}A)$, entonces $|B| = \left| \, |A|^{-1}A \, \right| = \left| |A|^{-1} \right| \, |A|$

$$= |A|^{-1}|A| = 1$$

$$\therefore \; |B| = 1$$

y por el lema anterior sabemos que para todo eigenvalor λ de B se cumple que $|\lambda| < |B| = 1$.
Sea p el eigenvector asociado a $|B| = 1$, entonces $Bp = |B|p = p$ y de aquí que $B^n p = p$
sea $p^* = \max p_i$ y $p_* = \min p_i$, con $i \in \{1, 2, \dots, k\}$
y sea $(B_{ij}^{(n)})$ la entrada ij de B^n

entonces se tiene que :

$$p^* \ge p_i = \sum_{j=1}^{k} B_{ij}^{(n)} p_j \ge \sum_{j=1}^{k} B_{ij}^{(n)} p_* \ge B_{ij}^{(n)} p_* \qquad \text{para toda } i,\ j$$

$\therefore\ B_{ij}^{(n)} \le (p^* \,/\, p_*)$ y de aquí que B^n sea acotada.

Antes de continuar veamos la siguiente observación.

OBSERVACION-10.0)

Si $B \in M_{k\times k}(F)$ y $J_B \in M_{k\times k}(F)$ es la forma de Jordan asociada a la matriz B. Entonces se tiene que si B^n está acotada entonces tambien (J_B) está acotada.

Demostración:

Sabemos que existe $Q \in M_{k\times k}(F)$ invertible tal que $B = Q\, J_B\, Q^{-1}$ y queremos ver que $(J_B)^n$ está acotada.

Pero $J_B = Q^{-1} B\, Q$ de donde

$$(J_B)^n = \underbrace{(Q^{-1}B\,Q)(Q^{-1}B\,Q)\ \dots\ (Q^{-1}B\,Q)}_{\text{n-veces}} =$$

$$= Q^{-1}B\,(Q\,Q^{-1})\,B\,(Q\,Q^{-1})\,B\ \ \dots\ \ (Q\,Q^{-1})\,B\,Q$$

$$= Q^{-1}(\underbrace{B\,B\,B\,\dots\,B\,B}_{\text{n-veces}})\,Q$$

$$= Q^{-1}\,B^n\,Q$$

$\therefore\ (J_B)^n = Q^{-1}\,B^n\,Q$

Tomemos una entrada de $(Q^{-1}\,B^n\,Q)$

entonces:

$$(Q^{-1}\,B^n\,Q)_{ij} = [\,(Q^{-1}\,B^{(n)})\,(Q)\,]_{ij} = \sum_{s=1}^{k} (\sum_{j=1}^{k} Q_{ij}^{-1}\,B_{js}^{(n)})\,Q_{sl} =$$

$$= \sum_{s=1}^{k}\ \sum_{j=1}^{k} Q_{ij} B_{js}^{(n)} Q_{sl}$$

pero B^n esta acotada $\therefore$ para toda $i, j \in \{1,\ 2,\ \dots\ ,\ k\}$ y para toda $n \in \mathbb{N}$ sabemos que existe una $M \in \mathbb{N}$ tal que $B_{ij}^{(n)} \le M$

entonces $\sum\sum Q_{ij}^{-1}\ B_{js}^{(n)} Q_{sl} \ \le\ \sum\ \sum Q_{ij}^{-1}\ M\ Q_{sl} = (*)$

pero $Q_{ij}^{-1} < \infty$, $Q_{ij} < \infty$ $\in F$ por lo tanto Q^{-1} y Q son finitas y de aquí que $(*) < \infty$.

$\therefore\ (J_B)_{ij} = (\,Q^{-1}\ B^{(n)}\ Q\,)_{ij} \le (*)$ para toda $i,j \in \{1,\ 2,\ \dots\ ,k\}$

por lo tanto J_B está acotada.

∎

Sabemos que $|B| = 1$ y por lo tanto para cualquier eigenvalor λ de B se tiene que $|\lambda| < |B| = 1$, y como J_B está acotada, ésto implica que J_B no puede tener bloques asociados al eigenvalor 1 de B que sean de la forma:

$$J_{B_1} = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 \\ 0 & 1 & 1 & \dots & 0 \\ 0 & . & . & \dots & 0 \\ \vdots & & \vdots & \cdot & \vdots \\ 0 & \dots & \dots & 1 & 1 \end{pmatrix}_{m\times m}$$

como enseguida demostramos:

En la proposición 1.0 vimos que podemos partir una matriz y mas aún que podemos hacer el producto por particiones, supongamos que tuviéramos un bloque de Jordan asociado a $|B| = 1$ como el de arriba y hagámosle una partición.

$$J_{B_1} = \begin{pmatrix} J^{11}_{B1} & J^{12}_{B1} & \dots & J^{1s}_{B1} \\ J^{21}_{B1} & J^{22}_{B1} & \dots & J^{2S}_{B1} \\ \vdots & \vdots & \ddots & \vdots \\ J^{S1}_{B1} & J^{S2}_{B1} & \dots & J^{SS}_{B1} \end{pmatrix}$$

Tal que $J^{ij}_B = \begin{cases} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} & \text{si} \quad i = j \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \text{si} \quad i \neq j \end{cases}$

es claro que para $i = j$ $\quad (J^{ij}_B)^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$ y de aquí que divergiría cuando $n \to \infty$, $\therefore (J^{ij}_B)^n$ no esté acotada y en general $(J_{B1})^n$ estaría acotada. $\therefore$ los bloques de Jordan asociados a $|B| = 1$ son todos de 1×1 y como $|\lambda| < 1 = |B|$ entonces se tiene que cada bloque de Jordan de la forma (J_{B_λ}) converge, por lo tanto la forma de Jordan J asociada a B es convergente y como $J_B \simeq B$ se tiene entonces que B^n es convergente que es lo que queríamos probar. ∎

COROLARIO-3.0)

$|A|$ es un eigenvalor simple de A

Demostración:

Lo que se quiere es equivalente a que $|A|$ es una raíz de multiplicidad 1 del polinomio característico de A. Sea $B = |A|^{-1}A$

ya vimos que $J_B = I_{n\times n} \oplus J_{\lambda 1} \oplus J_{\lambda 2} \oplus \dots \oplus J_{\lambda k}$ y sabemos que

$\mathrm{mult}|A| = n = \mathrm{mult}|B|$ como eigenvalor de A y de B respectivamente. Por demostrar n=1:

Tenemos que $J_B - I = 0_{n\times n} \oplus J_{\lambda_1 - 1} \oplus J_{\lambda_2 - 1} \oplus \dots \oplus J_{\lambda_k - 1}$

$\therefore$ nul $(J_B - I) = n = \mathrm{nul}(B - I)$ pues $J \simeq B$.

También, para toda $n \geq 1$ existe $x = (x_1, x_2, \dots, x_k) \in \mathbb{R}^k$ tal que $B^n x = x$ ya que si la n fuera estrictamente positiva $(n > 1)$ existiría entonces $y \in \mathbb{R}^k$ tal que $By = y$ con $x \neq \alpha y$ y $\alpha \in F$

como $x = (x_1, x_2, \dots, x_k)$ y $y = (y_1, y_2, \dots, y_k)$

hagamos $\tau = \max(x_i / y_i) = (x_j / y_j)$ con j fija

entonces $\tau y \geq x$ pues $(x_j / y_j) \geq (x_i / y_i)$ para toda $i \in \{1, 2 \dots k\}$

entonces $\tau \geq (x_i / y_i)$, $\tau y_i \geq x_i$ para toda i, $\therefore \tau x \geq y$

y $\tau y - x \geq 0$.

Como $B^n > 0$ se tiene que:

$$B^n(\tau y - x) > 0$$

$$\tau B^n y - B^n x > 0$$

$$\tau y - x > 0$$

$$\tau y > x$$

y para toda $i \in \{1, \dots, k\}$ $\qquad \tau y_i > x_i$

en particular para $y_j \wedge x_j$ $\qquad \tau y_j > x_j$

pero $\qquad \tau y_j = (x_j / y_j) y_j > x_j$

lo cual es una contradicción y $\therefore n = 1$

∎

La herramienta anterior nos servirá para demostrar el siguiente teorema que es el Teorema de Perron, pero antes hablaremos un poco acerca de cadenas de Markov, relacionándolas con el teorema y resaltando ciertos tipos de ellas más adelante.

CADENAS DE MARKOV Y MATRICES ESTOCASTICAS

Supongamos que estamos tratando un cierto experimento que puede estar en un número finito de formas diferentes o puede transformarse conforme pasa el tiempo, a estas formas les llamaremos estados, (supongamos que son k) entonces estos k estados difieren segun las condiciones de su evolución, es decir conforme evoluciona puede o no recorrer los k estados y puede o no regresar a su estado inicial.

Para generalizar, hablemos no únicamente de experimentos sino de cualquier tipo de fenómeno. Denotemos cada estado por e_i, y denotemos la probabilidad de pasar de un estado i a un estado j por e_{ij}. Supongamos que el cambio del estado i al estado j depende unicamente del estado i, es decir de nuestro estado inmediato anterior, entonces a este proceso se le conoce como proceso de *Markov* y si este proceso es infinito entonces se llamará *Cadena de Markov*. Es claro que en un principio (al tiempo t_0) nuestro fenómeno puede estar en cualquiera de los k estados, entonces podemos construir un vector de k entradas en donde la primer entrada nos diga la probabilidad de que el fenómeno se encuentre en el primer estado, la segunda entrada nos diga la probabilidad de que nuestro fenómeno se encuentre en el segundo estado, hasta llegar a la k-ésima entrada que nos dirá la probabilidad de que se encuentre en el k-ésimo estado. Por darse estas probabilidades en un principio (tiempo t_0) a este vector se le denomina vector inicial de probabilidades y se denota a menudo de la siguiente manera $p_0 = (e_1, e_2, \ldots, e_k)$, cabe hacer notar que en todo vector de probabilidades la suma de sus entradas debe ser igual a uno, ya que en cada entrada se está asignando una probabilidad para cada elemento de nuestro espacio muestral y en este vector se encuentran las probabilidades de todos los estados que forman nuestro espacio muestral.

También podemos formar una matriz de tamaño k (k = número de estados) con las probabilidades de pasar de un estado i a uno j, a dicha matriz se le denomina matriz de probabilidades, estocástica o de transición y tiene la siguiente forma:

$$P = \begin{array}{c c} & \begin{array}{ccc} e_1 & \cdots & e_k \end{array} \\ \begin{array}{c} e_1 \\ \vdots \\ e_k \end{array} & \begin{pmatrix} e_{11} & \cdots & e_{1k} \\ \vdots & & \vdots \\ e_{k1} & \cdots & e_{kk} \end{pmatrix} \end{array}$$

donde cada renglón (o columna) es un vector de probabilidades. definamos bien lo anterior.

DEFINICION-15.0)

Decimos que un vector v en $\mathbb{R}^k$ es un vector de probabilidades si $\sum_{i=1}^{k} v_i = 1$ y $v_i \geq 0$

DEFINICION-16.0)

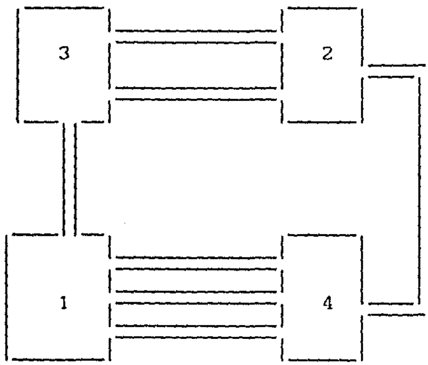
Decimos que una matriz A en $\mathbb{R}^k \times \mathbb{R}^k$ es de probabilidades si cada uno de sus renglones (columnas) es un vector de probabilidades.

Veamos el siguiente ejemplo para aclarar las dudas, supongamos que la siguiente figura nos representa a cuatro jaulas (guaridas) de una rata, las cuales están unidas por siete túneles arreglados como en el dibujo de la página siguiente.

Si suponemos en un principio que la rata puede estar a un tiempo inicial t_0 en cualquiera de las cuatro guaridas entonces podemos formar un vector de probabilidad inicial $v=(p_1, p_2, p_3, p_4)$ donde cada p_i nos dice la probabilidad de que la rata esté en la jaula

i al tiempo inicial t_0.

Sin embargo si suponemos que la rata la soltamos al tiempo inicial t_0 en la jaula 3, entonces nuestro vector de probabilidad inicial tendrá la forma $v = (0,\ 0,\ 1,\ 0)$



Construyamos ahora una matriz que nos indique las probabilidades de que la rata pase, digamos a la jaula j , dado que en el instante de tiempo anterior estaba en la jaula i, supongamos primero el caso en que la rata está en permanente movimiento, es decir que a cada intervalo de tiempo t la rata cambia de lugar.
La matriz que formemos será de 4×4 puesto que tenemos cuatro estados y tiene la forma:

$$P = \begin{array}{c} \\ e_1 \\ e_2 \\ e_3 \\ e_4 \end{array} \begin{array}{c} \begin{array}{cccc} e_1 & e_2 & e_3 & e_4 \end{array} \\ \begin{pmatrix} 0 & 0 & 1/4 & 3/4 \\ 0 & 0 & 2/3 & 1/3 \\ 1/3 & 2/3 & 0 & 0 \\ 3/4 & 1/4 & 0 & 0 \end{pmatrix} \end{array} \text{---------------(1)}$$

resulta fácil ver que cada renglón es un vector de probabilidades.
También podemos construir la siguiente matriz de probabilidades si suponemos que la rata puede permanecer en la misma guarida aún

pasado un intervalo de tiempo t, es decir la rata puede elegir entre quedarse en la misma guarida o mudarse a otra.

$$P = \begin{array}{c|cccc} & e_1 & e_2 & e_3 & e_4 \\ \hline e_1 & 1/5 & 0 & 1/5 & 3/5 \\ e_2 & 0 & 1/4 & 2/4 & 1/4 \\ e_3 & 1/4 & 2/4 & 1/4 & 0 \\ e_4 & 3/5 & 1/5 & 0 & 1/5 \end{array} \text{---------------} (2)$$

En cada una de las matrices anteriores la entrada P_{ij} nos indica la probabilidad de pasar del estado (en este caso jaula) i al estado j, por ejemplo la entrada P_{14} de la primer matriz es el número 3/4 , es decir este número nos indica que la rata tiene 3/4 de probabilidad de pasar a la jaula 4 dado que estuviera en la jaula 1. Mientras que en la segunda matriz la entrada P_{44} es el número 1/5, es decir este número nos indica la probabilidad de que la rata se quede en la jaula 4 dado que estuviera en esa misma jaula o sea que tiene 1/5 de probabilidad de permanecer en la misma jaula dado que tiene las opciones de las jaulas 1 y 2.

Veamos el siguiente ejemplo el cual es un juego llamado el juego de *Craps*, este juego consiste en tirar un par de dados donde si la suma de los puntos es igual a 7 u 11 el jugador gana el juego, si la suma de los puntos es un 2, 3, o un 12 entonces el jugador pierde el juego, sin embargo si el jugador tira los dados y la suma es igual a cualquier otro número diferente de los anteriores entonces se dice que este número es el punto del jugador y este repite su tirada una y otra vez hasta que salga su punto (ganando el juego) o le salga un siete (perdiendo el juego) a este juego lo podemos ver como una cadena de *Markov* de la siguiente manera.

El siguiente arreglo es el espacio muestral de la suma de los puntos de dos dados comunes no cargados:

D_1+D_2	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

y los estados en los que podemos estar son :

s1: ganar el juego con (7, 11)

s2: perder el juego con (2, 3, 12)

s3: tener como punto al 4

s4: tener como punto al 5

s5: tener como punto al 6

s6: tener como punto al 8

s7: tener como punto al 9

s8: tener como punto al 10

entonces el espacio muestral y los estados anteriores nos establece el siguiente vector de probabilidad inicial:

v_o = (p(s1), P(s2), P(s3), P(s4), P(s5), P(s6), P(s7), P(s8))
donde cada P(si) es igual a la probabilidad de estar en el estado i con i = 1 , 2 , ... , k.

entonces es claro que v_o = (1/36)(8, 4, 3, 4, 5, 5, 4, 3)
y nuestra matriz de probabilidad es la siguiente:

$$
C = \begin{array}{c} \\ s1 \\ s2 \\ s3 \\ s4 \\ s5 \\ s6 \\ s7 \\ s8 \end{array}
\begin{array}{c}
\begin{array}{cccccccc} s1 & s2 & s3 & s4 & s5 & s6 & s7 & s8 \end{array} \\
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
3/36 & 6/36 & 27/36 & 0 & 0 & 0 & 0 & 0 \\
4/36 & 6/36 & 0 & 26/36 & 0 & 0 & 0 & 0 \\
5/36 & 6/36 & 0 & 0 & 25/36 & 0 & 0 & 0 \\
5/36 & 6/36 & 0 & 0 & 0 & 25/36 & 0 & 0 \\
4/36 & 6/36 & 0 & 0 & 0 & 0 & 26/36 & 0 \\
3/36 & 6/36 & 0 & 0 & 0 & 0 & 0 & 27/36
\end{pmatrix}
\end{array}
$$

analicemos el renglón cinco (estado cinco), viendo cada una de sus entradas; la entrada $(s5,s1)$ nos dice que si hemos sacado como punto al No 6 , ¿Entonces que probabilidad tenemos de ganar el juego? dado que las condiciones del problema son repetir el 6 para ganar, y vemos que se repite el 6 cinco veces de treinta y seis posibilidades en el espacio muestral se infiere entonces que la probabilidad de ganar es 5/36. De la misma manera la probabilidad de perder es 6/36, la probabilidad de permanecer en ese estado es 27/36 puesto que no importa el número que salga (distinto de seis o siete) nuestro estado no cambiará y de aquí mismo que la probabilidad de cambiar a algún otro estado distinto del 1, 2 o 5 sea cero.

En cada uno de los ejemplos anteriores la información que nos proporciona cada una de las matrices estocásticas es la probabilidad que tenemos de pasar de un estado a otro, incluyendo el mismo en el que estamos. Pero eso no es todo, pues si tomamos n potencias de una matriz estocástica, la matriz resultante nos dirá la probabilidad de pasar de un estado a otro después de n veces y si aplicamos nuestra matriz estocástica a nuestro vector de probabilidad inicial, la entrada j del vector resultante nos dirá la probabilidad de estar en el estado j dado que ya transcurrió un periodo de tiempo de longitud uno, o de longitud n si estamos aplicando la potencia n de nuestra matriz estocástica a nuestro vector de probabilidad inicial.

Es fácil notar que si aplicamos nuestra matriz estocástica C o P al vector $e = (1, 1, 1, \dots ,1)^{tr} \in \mathbb{R}^k$ entonces el vector resultante será de nuevo e y ésto es por ser C y P estocásticas, entonces tendremos que se cumple la igualdad $Ce = 1e$ de donde se infiere que 1 es un eigenvalor de cualquier matriz estocástica y por el lema 1 tambien se infiere que el radio espectral de cualquier matriz estocástica es igual a 1.

Ya con el material anterior enunciemos el siguiente teorema

TEOREMA-9.0) *(PERRON)*

Sea A una matriz de transición de $k\times k$ tal que para todo estado i y todo estado j, se tiene que la probabilidad de pasar del estado i al j (denotado por p_{ij}) es positiva, e.d. $p_{ij} > 0 \;\forall\; i,j$ con $i,j \in \{1,2,\dots,k\}$

entonces :

i) $|A|$ *es un eigenvalor simple de* A *teniendo asociado un vector* de probabilidades estrictamente *positivo* y $|A| = 1$

ii) $|A| > |\lambda|$ *para todo eigenvalor* λ *de* A *que sea distinto a* $|A|$.

Demostración.- Aplicando el lema de Wielandt queda la demostración casi completa, solo falta demostrar que $|A| = 1$ pero por el lema 1 tenemos que $|A| \leq \sum^{k} p_{ij} = 1$, pero para toda matriz estocástica A se tiene que $Az = 1z$, con $z \in \mathbb{R}^k$ tal que $z = (1,1, \dots, 1)$ entonces tenemos que 1 es un eigenvalor de A y sabemos que $|A| \geq \lambda$ para todo eigenvalor λ de A, entonces se tiene que $|A| \geq 1$, entonces $|A| = 1$.

∎

TEOREMA-10.0)

Si $A > 0$, p es cualquier eigenvector asociado a $|A|$ en A y q es cualquier eigenvector asociado a $|A|$ en A^{tr} entonces *:*

$$\lim_{n\to\infty}(|A|^{-1}A) = (q^{tr}p)^{-1}(pq\quad)^{tr}$$

Demostración:

Por el teorema 8.0 sabemos que el $\lim\limits_{n\to\infty} B^n = (|A|^{-1}A)^n$ existe sea $L = \lim(|A|^{-1}A)^n$ entonces cada columna de L está en el eigenespacio asociado a $|B| = 1$, es decir $BL_j = L_j$ para cualquier $j \in \{1, 2, \dots, k\}$ y donde L_j nos representa la columna j de la matriz L, pues $L = \lim_{n \infty} B^{n+1} = \lim_{n \infty} B\, B^n = B\, L = L$

como $|B|$ es de multiplicidad uno, entonces se tiene que el eigenespacio generado por el vector asociado a $|B|$ es de dimensión uno, por lo tanto $L_j = r_j\ p$ (donde p cumple $(|A|^{-1}A)p = p$ o $Ap = |A|p$)

por lo tanto:

$$L = \begin{pmatrix} r_1 p_1 & r_2 p_1 & \dots & r_k p_1 \\ r_1 p_2 & r_2 p_2 & \dots & r_k p_2 \\ \vdots & \vdots & & \vdots \\ r_1 p_k & r_2 p_k & & r_k p_k \end{pmatrix}$$

Sea $r = (r_1, r_2, \dots, r_k)^{tr}$ y $p = (p_1, p_2, \dots, p_k)$,

entonces:

$$L = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_k \end{pmatrix} (r_1, r_2, \dots, r_k) \qquad \text{e.d.} \quad L = pr^{tr}.$$

Ahora bien, sabemos que existe un vector q tal $A^{tr}q = |A|q$ que es lo mismo que decir que $(|A|^{-1}A^{tr})\, q = q$, e.d. $B^{tr}q = q$ y se cumple que:

$$(B^{tr})^n q = q$$

$\Leftrightarrow$ $\quad q^{tr}B^n = q^{tr}$

$\Rightarrow$ $\quad q^{tr}\, L = q^{tr}$. Pero sustituyendo $L = p\, r^{tr}$, tenemos que

$q^{tr}p\ r^{tr} = q^{tr}$, $\quad (q^{tr}p$ es un escalar no cero)

$\therefore \quad r^{tr} = (q^{tr}p)^{-1}q^{tr}$ multiplicando ambas partes por p

$pr^{tr} = p\ (q^{tr}p)^{-1}q^{tr}$

$\therefore \quad L = p\ (q^{tr}p)^{-1}q^{tr}$

■

Es decir, dada una matriz B positiva tal que $|B| = 1$ podemos encontrar el límite de sus potencias en términos de únicamente

sus vectores propios que se pueden calcular facilmente.
Los teoremas más conocidos acerca de limites tambien se cumplen para matrices cuadradas. Veamos los siguientes resultados que utilizaremos a lo largo de este trabajo.

DEFINICION-16.0)
Sea A una matriz cuadrada de kxk entonces decimos que $\lim_{n\,\infty} A^n = L_A$ $\Leftrightarrow$ la sucesión $A_{ij}^{(n)}$ converge a $(L_A)_{ij}$ para cada i y para cada j. L_A es la matriz limite de las potencias de A.

Haciendo uso de la definición anterior probemos los siguientes resultados.

TEOREMA-11.0) *De limite* 1)
Sean A y B matrices cuadradas de kxk , tales que los límites de sus potencias existen e.d. $\lim_{n\,\infty} A^n = L_A$ y $\mathrm{Lim}_{n\,\infty} B^n = L_B$

$$\text{entonces } \mathrm{Lim}_{n\,\infty} (A^n + B^n) = L_A + L_B$$

Demostración
$\mathrm{Lim}_{n\,\infty} (A^n + B^n) = L_{A+B} \Leftrightarrow (A^{(n)} + B^{(n)})_{ij}$ converge a $(L_{A+B})_{ij}$ (para cualquier i y para cualquier j)
$\Leftrightarrow (A^{(n)})_{ij} + (B^{(n)})_{ij}$ converge a $(L_A)_{ij} + (L_B)_{ij}$ (pues sabemos que esto ocurre para sucesiones)
$\Leftrightarrow \lim_{n\,\infty} A^n + \lim_{n\,\infty} B^n = L_A + L_B$

$\therefore \mathrm{Lim}_{n\,\infty} (A^n + B^n) = L_A + L_B$

■

TEOREMA-11.0) *de límite* 2)
Sean A y B matrices cuadradas (kxk) tales que

$$\lim_{n\to\infty} A^n = L_A \quad y \quad \lim_{n\to\infty} B^n = L_B$$

$$\text{entonces} \quad \lim_{n\to\infty} (A^n \circ B^n) = L_A \circ L_B$$

Demostración.

$\lim_{n\to\infty} (A^n \circ B^n) = (L_{AB})$

$\Leftrightarrow \ (A^n \circ B^n)_{ij}$ converge a $(L_{AB})_{ij}$

pero $(A^n \circ B^n)_{ij} = \sum_{s=1}^{k} A_{is}^{(n)} B_{sj}^{(n)}$

$$\sum_{s=1}^{k} A_{is}^{(n)} B_{sj}^{(n)} = \sum_{s=1}^{k} (L_A)_{is} (L_B)_{sj}$$

$$y \quad \sum_{s=1}^{k} (L_A)_{is} (L_B)_{sj} = L_A \cdot L_B$$

$$\therefore \lim_{n\to\infty} (A^n B^n) = L_A L_B$$

■

TEOREMA-12.0) *de límite* 3)

Sean A, B matrices cuadradas (kxk) tales que $\lim_{n\to\infty} A^n = L_A$

$$\text{entonces} \quad \lim_{n\to\infty} (BA^n) = B \lim_{n\to\infty} A^n = B\, L_A$$

Demostración)

$$\lim_{n\to\infty} A^n = L_A \quad \Leftrightarrow \quad \lim_{n\to\infty} (A^{(n)})_{ij} = (L_A)_{ij} \qquad (*)$$

$$\lim_{n\to\infty} (BA^n)_{ij} = \lim_{n\to\infty} \left(\sum_{s=1}^{k} B_{is} A_{sj}^{(n)}\right) = \sum_{s=1}^{k} \lim_{n\to\infty} (B_{is} A_{sj}^{(n)}) =$$

$= \sum_{s=1}^{k} B_{is} \left(\lim_{n\to\infty} A_{sj}^{(n)}\right)$ y para cada $s \in \{1,2,\ldots,k\}$ se cumple (*)

$$\therefore \quad \sum_{s=1}^{k} B_{is} \left(\lim_{n \infty} A_{sj}^{(n)}\right) = \sum_{s=1}^{k} B_{is} \, (L_A)_{sj} = (B\,(L_A))_{ij}$$

de donde se sigue que $\lim_{n \infty}(BA^n) = B \lim_{n \infty} A^n = B\, L_A$ ∎

MATRICES PRIMITIVAS

Dentro del conjunto de matrices no negativas existe un tipo de matriz que se caracteriza por transformarse en una matriz estrictamente positiva al tomar las potencias de la misma, este tipo de matriz es la que definiremos como primitiva.

DEFINICION-17.0)

Sea A una matriz no negativa entonces diremos que A es *Primitiva* si para alguna $0 < n \in \mathbb{N}$ se tiene que $A^n > 0$.

Cabe mencionar que si $A \geq 0$ y $A \in M_{2x2}(\mathbb{R})$ entonces A^n siempre admite una entrada igual a cero y tambien si $A > 0$ entonces $A^n > 0$ para toda $n \in \mathbb{N} \setminus \{0\}$, con esta observación se excluyen a las matrices de 2x2 de las matrices primitivas.
Las matrices primitivas tambien cumple con las conclusiones del teorema de *Perron* como veremos en el siguiente teorema.

TEOREMA-13.0)

Sea A una matriz primitiva entonces $|A|$ es un eigenvalor de multiplicidad uno, asociado a un eigenvector positivo y nigún otro eigenvalor tiene su módulo igual a $|A|$.

Demostración:
Veamos primero la siguiente

OBSERVACION 5.0)

$|A|^n = |A^n|$ para toda $n \in \mathbb{N}$.

Demostración:

demostración por inducción / n.

Para n = 1 no hay problema

Supongámosla válida para n-1,

e.d. $|A^{n-1}| = |A|^{n-1}$

y demostremos para n

$|A^n| = |A^{n-1}A| = |A^{n-1}||A| = |A|^{n-1}|A| = |A|^n$

$\blacksquare$

Como A es primitiva entonces existe $n \in \mathbb{N}$ tal que $A^n > 0$ entonces aplicando el teorema de Perron a esta nueva matriz sabemos que $(A^n)p = |A^n|p$ para algún eigenvector positivo asociado a $|A^n|$

hagamos $\quad p = \sum_{i=0}^{n-1} |A|^{-i}A^i p$

entonces aplicando A a p tenemos que

$$Ap = A\Big(\sum_{i=0}^{n-1} |A|^{-i}A^i p\Big) =$$

$$= \sum_{i=0}^{n-1} |A|^{i} A^{i+1}p =$$

$$= \sum_{i=0}^{n-2} |A|^{-i}A^{i+1}p + |A|^{1-n} A^n p =$$

$$= \sum_{i=1}^{n-1} |A|^{-i+1}A^i p + |A|^{1-n} |A^n| p$$

$$= |A|\Big(\sum_{i=1}^{n-1} |A|^{-i}A^i p + p\Big) \qquad \text{y como } |A^n| = |A|^n$$

$$= |A|p.$$

Nótese que p es una suma de puros vectores positivos, pues tanto $|A|$ como p son distintos de cero y de aquí que p sea estrictamente positivo.

Notemos también que p es también un eigenvector para $|A|$.

Mostremos ahora que $|A|$ es un eigenvalor simple de A, la multiplicidad de $|A|$ como eigenvalor de A no excede a la multiplicidad de $|A^n|$ como eigenvalor de A^n

Supongamos ahora que existe μ eigenvalor de A tal que $|\mu| = |A|$ como $Aq = \mu q$ para alguna $q \neq 0$ entonces $A^n q = \mu^n q$
pero $|\mu^n| = |\mu|^n = |A|^n = |A^n|$
como $|A^n| \geq \nu$, con $|A^n| \neq \nu$ y $\mu^n \geq \nu$ para todo eigenvalor ν de A^n entonces tenemos que $|A^n| = \mu^n$.
$\therefore\ A^n q = \mu^n q = |A^n| q$,
pero $A^n p = |A^n| p$ y de aqui que $q = \lambda p$ pues el eigenespacio asociado a $|A|$ es unidimensional.
$\therefore\ |A|^n = |A^n| = \mu^n$
y como $|A|^{n-1}|A| = \mu^{n-1}\mu$ entonces tenemos que $|A| = \mu$.

Probemos que en este tipo de matrices tambien se puede calcular el limite de sus potencias por medio de sus vectores propios y su radio espectral.

TEOREMA-14.0)
Sea A una matriz primitiva, p cualquier eigenvector positivo de A asociado a $|A|$ y q cualquier eigenvector positivo de A^{tr} asociado a $|A|$, entonces :

$$\lim_{n \to \infty} (|A|^{-1}A)^n = (q^{tr}p)^{-1} p q^{tr}.$$

Demostración:
Nuestro teorema anterior nos asegura que podemos encontrar los eigenvectores de A y A^{tr} asociados a $|A|$, y por nuestro primer teorema sabemos que el limite buscado existe pues aplicando de nuevo el teorema anterior sabemos que $(|A|^{1}A)^n > 0$ y que $|(|A|^{-1}A)^n| = |(|A|^{-1}A)| = 1$ que es a la vez de multiplicidad uno, con $1 > \nu$ para toda ν eigenvalor de $(|A|^{-1}A)$ las cuales son las condiciones necesarias para que este limite exista.
la demostración de el limite es la misma que la del teorema 1.0

∎

A nosotros nos gustaría tener algún tipo de prueba que nos facilitara el reconocimiento de cuando una matriz no negativa fuera primitiva, ya que nos podemos pasar bastante tiempo calculando las potencias de alguna matriz no negativa y tal vez ni siquiera sea primitiva.
El siguiente criterio es una prueba para saber si alguna matriz M no negativa es primitiva y se demostrará mas adelante.

TEOREMA-15.0)
Sea A una matriz no negativa de $k \times k$ entonces A es primitiva si $A^{(k-1)^2+1} > 0$.

Una generalización del teorema de Perron sería saber que pasa si A es una matriz que no es primitiva, es natural pensar que no se cumplan todas las afirmaciones del teorema de Perron como veremos en el siguiente ejemplo donde la matriz dada no es primitiva.

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

haciendo ciertos calculos nos damos cuenta que $|A| = 2$ y que $|A|$ tiene asociado un eigenvector de la forma $(a, a, t, 0)$ el cual no es positivo.

Sin embargo lo que sí se puede asegurar es que $|A|$ es un eigenvalor de A que está asociado a un eigenvector que nunca es negativo, como lo veremos en el siguiente teorema:

TEOREMA-16.0)
Si $A \geq 0$ entonces $|A|$ es un eigenvalor de A asociado a un eigenvector no negativo.

Demostración:
Sea $B_n = A + (1/n) * E$ donde E es una matriz de $k \times k$ y $E_{ij} = 1$ para toda i y j en el siguiente conjunto $\{1, \ldots, k\}$
es claro que $0 \leq A < B_{n+1}$ para toda $n \in \mathbb{N}$

Veamos antes la siguiente observación

OBSERVACION-11.0) Sean A, C matrices del mismo tamaño tal que cumplen que $0 \le A < C$ entonces $|A| < |C|$

Demostración:

Multipliquemos a nuestra primer desigualdad por $|C|^{-1} > 0$ que no cambia el sentido de la desigualdad e.d.

$$|C|^{-1}0 \le |C|^{-1}A < |C|^{-1}C \qquad \text{---------------------}(1)$$

$$0 \le |C|^{-1}A < (|C|^{-1}C) \qquad \text{---------------------}(2)$$

$$0^n \le (|C|^{-1}A)^n < (|C|^{-1}C)^n \text{------------------}(3)$$

puesto que C es una matriz estrictamente positiva sabemos que el lado derecho de la desigualdad es convergente y entonces $(|C|^{-1}A)^n$ está acotada, tomemos pues el radio espectral en cada parte de la desigualdad (2) teniendo asi:

$$|0| \le ||C|^{-1}A| < ||C|^{-1}C| \qquad \text{---------------------}(2)$$

$$0 \le |C|^{-1}|A| < 1$$

y de aquí que $0 \le |A| < |C|$

∎

De la observación anterior tenemos entonces que

$$|A| \le |B_{n+1}| < |B_n| \qquad \text{para toda } n \in \mathbb{N}$$

Sea $\lambda = \min\{|Bn| : n \ge 1\}$

entonces $|A| \le \lambda = \lim_{n\ \infty} |Bn|$

cada Bn es una matriz positiva, entonces aplicando el teorema de Perron a cada una de ellas encontraremos $y^{(n)}$ vectores positivos tales que

$$Bn\, y^{(n)} = |Bn| y^{(n)} \qquad \forall\ n > 1$$

haciendo $z^{(n)} = \left(\sum_{i=1}^{k} y_i^{(n)} \right)^{-1} y^{(n)} \qquad \forall\ n > 1$

teniendo que para cada $n > 1$

$$Bn\, z^{(n)} = |Bn| z^{(n)} \quad , \quad z^{(n)} > 0 \quad \text{y} \quad \sum_{i=1}^{k} z_i^{(n)} = 1$$

$S = \{ z \geq 0 : \sum_{l=1}^{k} z_l = 1 \}$ es un subconjunto compacto de $\mathbb{R}^{(k)}$

y entonces existe una subsucesión de elementos de S que converge a un elemento z de S es decir $\lim_{m\, \infty} z^{(n_m)} = z$ con $z \in S$

se sigue que $\lim_{m\, \infty} (Bn_m\, z^{(n_m)}) = \lim_{m\, \infty} (Bn_m) * \lim_{m\, \infty} (z^{(n_m)}) = A * z$

tambien $\lambda\, z = \lim_{m\, \infty} |Bn_m| \lim_{m\, \infty} z^{(n_m)} = \lim_{m\, \infty} [\, |Bn_m| z^{(n_m)}] =$

$$= \lim_{m\, \infty} [Bn_m\; z^{(n_m)}] = A\, z$$

$\therefore\ Az = \lambda z$

pero $z \neq 0$ pues $0 \notin S$ asi que λ tiene asociado un eigenvector z no negativo

$|\lambda| \leq |A|$ por ser λ un eigenvalor de A y

$|A| \leq \lambda$ por lo anterior $\therefore\ \lambda = |A|$

∎

CADENAS REGULARES

Según lo visto para las matrices positivas se tiene que una matriz es primitiva si a partir de cierto momento se tiene que esta matriz se vuelve positiva, esto tambien lo podemos "traducir" para matrices estocásticas diciendo que una de este tipo es *regular* si a partir de cierto tiempo no importando el estado en el que estuvimos inicialmente, siempre existirá una probabilidad positiva para estar en cualquier estado, en otras palabras a nuestra cadena la podemos llamar de manera análoga a las primitivas como *Regular*.

Ya vimos antes que el radio espectral de cualquier matriz estocástica es igual a 1, entonces si aplicamos el teorema (10.0) tenemos que $\lim_{n\, \infty} P^n = \bar{1}\, q^{tr}$ donde q es un vector estocástico columna de P^{tr} asociado a $|P| = 1$

pues si $Pp = |P|p$ y $P^{tr}q = |P|q$

entonces $\lim(|P|^{-1}P)^n = (q^{tr}p)^{-1}pq^{tr}$

pero $|P| = 1$ y $(q^{tr}p)^{-1} = 1$ pues q es estocástico y $p^{tr} = (1,1,\ldots,1)$, finalmente $\lim_{n\ \infty} P^n = 1\ q^{tr}$

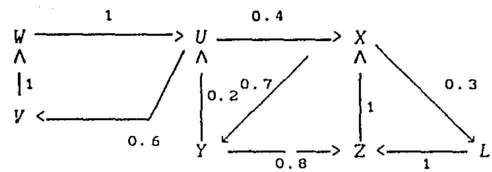
y se aquí claramente se ve que $1\ q^{tr} > \bar{0}$ no importando como se halla elegido $q^{(0)}$, de aquí vemos que $\lim_{n\ \infty} P^n$ es convergente.

CADENAS DE MARKOV ERGODICAS

Decimos que una cadena es *ergódica* si sabemos que con el paso del tiempo existirá una probabilidad positiva para estar en cualquier estado no importando en que estado estuvimos al principio. Una forma equivalente sería pedir $\sum_{n=1}^{m} A^n > 0$ para alguna m donde A es la matriz que nos determina nuestra cadena de Markov. Es fácil darnos cuenta que toda matriz regular es ergódica, el juego de Craps es un ejemplo de una cadena que no es ergódica pues una vez que hemos perdido el juego o lo hemos ganado ya no podemos cambiar de estado. Sin embargo el experimento de la rata en el laberinto si es una cadena ergódica pues siempre existe una probabilidad positiva de que la rata esté en cualquier jaula. Más adelante veremos como los estados de una cadena ergódica pueden ser clasificados de tal forma que A pueda ser expresada como una suma directa de matrices primitivas.

Veamos el siguiente ejemplo:

Supongamos que tenemos de nuevo una rata que puede estar en cualquiera de las siete jaulas siguientes denotadas por *W*, *U*, *X*, *V*, *Y*, *Z* y *L*, las cuales están intercomunicadas de la siguiente manera y con las siguientes probabilidades:



Entonces la rata tiene probabilidad 1 de pasar a la jaula U si estuviera en la jaula W y tiene probabilidad cero para regresarse puesto que así lo denota el sentido de la gráfica. En este ejemplo tambien suponemos que la rata está en constante movimiento por lo que la rata no puede permanecer en el mismo lugar cuando ya ha pasado un intervalo de tiempo, por lo que la probabilidad de que permanezca en un mismo lugar es cero.

La matriz que nos determina esta cadena de Markov es la siguiente:

$$
A = \begin{array}{c} \\ x \\ v \\ l \\ y \\ w \\ u \\ z \end{array}
\begin{array}{c}
\begin{array}{ccccccc} x & v & l & y & w & u & z \end{array} \\
\left(\begin{array}{cc|ccc|cc}
0 & 0 & 0.3 & 0.7 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0.2 & 0.8 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\hline
0.4 & 0.6 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right)
\end{array}
$$

Notemos en el dibujo que independientemente del estado en que estemos al principio tenemos una probabilidad positiva para llegar a cualquier estado y este hecho nos hace pensar que nuestra cadena es ergódica.

Nos damos cuenta de que A pudo ser partida en nueve bloques que aunque son de distinto tamaño nos clasifican a nuestros estados. Sin embargo A no parece cumplir con la propiedad de ser ergódica. Recordemos que la probabilidad positiva de estar en cualquier estado esta dada por el tiempo. Veamos entonces qué sucede conforme pasa el tiempo, es decir tomemos sus potencias y démonos cuenta de lo siguiente:

haciendo el producto por bloques como vimos en la proposición 1.0 tenemos que

$$A^2 = \left(\begin{array}{cc|ccc|cc} 0 & 0 & 0 & 0 & 0 & 0.14 & 0.86 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.88 & 0.12 & 0 & 0 & 0 & 0 & 0 \\ 0.4 & 0.6 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0.12 & 0.28 & 0.6 & 0 & 0 \\ 0 & 0 & 0.3 & 0.7 & 0 & 0 & 0 \end{array}\right),$$

$$A^3 = \left(\begin{array}{cc|ccc|cc} 0.96 & 0.084 & 0 & 0 & 0 & 0 & 0 \\ 0.4 & 0.60 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0.3 & 0.7 & 0 & 0 & 0 \\ 0 & 0 & 0.264 & 0.616 & 0.12 & 0 & 0 \\ 0 & 0 & 0.12 & 0.28 & 0.6 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0.656 & 0.344 \\ 0 & 0 & 0 & 0 & 0 & 0.14 & 0.86 \end{array}\right).$$

Dándonos cuenta del recorrido transversal que van haciendo los bloques, recorriendo cada entrada de A. Para este caso nos damos cuenta que de A^3 es una suma directa de matrices primitivas que además son estocásticas y de ahí el hecho de que A sea ergódica. Una vez que tenemos a A^3 como una suma directa de matrices positivas es fácil calcular los siguientes límites.

Supongamos que $\lim_{n\ \infty} A^{3n} = L$

entonces

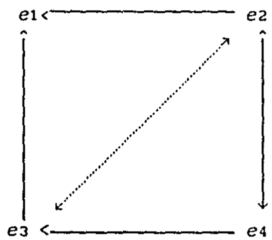
$$\lim_{n\ \infty} A^{3n+1} = L \cdot A$$

y $$\lim_{n\ \infty} A^{3n+2} = L \cdot A \cdot A = L \cdot A^2.$$

Lo cual nos sirve para calcular la convergencia de nuestro vector de probabilidad inicial, es decir supongamos que tenemos $p^{(0)}$ y queremos estimar $p^{(756)}$. Como $756 = 3 * 251 + 2$ por lo que tenemos que calcular $\lim_{n\ \infty} p^{(0)} A^{3n+2}$ y este nos dará la estimación.

CADENAS ABSORBENTES

Supongamos de nuevo que tenemos una rata corriendo a través del siguiente laberinto:



De donde al construir su matriz de transición tenemos que:

$$A = \begin{array}{c} \\ e1 \\ e2 \\ e3 \\ e4 \end{array} \begin{array}{c} \begin{array}{cccc} e1 & e2 & e3 & e4 \end{array} \\ \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \end{array} \right) \end{array}$$

Por el sentido de las flechas y tomando en cuenta que la rata tiene la opción de permanecer en el mismo lugar o de cambiar de lugar a cada instante de tiempo, nos damos cuenta que una vez que la rata ha caído en el casillero uno de ahí ya no puede salir . Cuando sucede algo similar en cualquier tipo de cadena markoviana decimos que caímos en un estado absorbente. Demos las siguientes definiciones para formalizar.

DEFINICION-18.0)
Un estado e_i se dice absorbente si y solamente si una vez entrado en dicho estado, éste no pude ser dejado.

DEFINICION-19.0)
Una cadena se llama absorbente si y solo si empezando en cualquier estado es posible que eventualmente el sistema estará en un estado absorbente.
Esto equivale a pedir que si s_i es arbitrario entonces existe s_j absorbente (que depende de i) y un tiempo t (que depende de i y de j) tal que $p_{ij}^{(m)} > 0$.

es decir una vez que entramos en un estado absorbente siempre existirá una probabilidad cero de salir de él.

DEFINICION-20.0)

Una cadena de Markov se dice absorbente si esta tiene uno o más estados absorbentes.

Supongamos que tenemos una cadena de Markov con k estados de los cuales s son absorbentes ($s \leq k$) entonces nosotros podemos clasificar los estados de la cadena agrupando primero los s estados absorbentes como en la siguiente matriz:

$$P = \begin{pmatrix} I_s & 0 \\ U & T \end{pmatrix}$$

en donde el bloque I_s nos representa los s estados absorbentes pues al caer en el la probabilidad de permanecer en el será 1 y obviamente la probabilidad de dejarlo es 0, es por eso que el bloque I_s no es más que la matriz identidad de $s\times s$. Si tomamos las potencias de P nos daremos cuenta de lo que pasa con los demás estados.

Por la observación 1.0 tenemos que:

$$P^n = \begin{pmatrix} I_s & 0 \\ \sum_{m=0}^{n-1} T^m U & T^n \end{pmatrix} \text{ para toda } n \geq 1$$

Supongamos que s_j es un estado absorbente y que para cierta i y m tenemos que $p_{ij}^{(m)} > 0$. Entonces $p_{ij}^{(m+r)} \geq p_{ij}^{(m)}\, p_{jj}^{(r)} = p_{ij}^{(m)} > 0$ para toda $r \geq 0$ pues $p_{jj}^{(r)} = 1$ por ser s_j absorbente. Como la cadena es absorbente entonces existe un estado absorbente s_{j_i} y un entero m_i tal

que $p_{ij_i}^{(m_i+r)} > 0$ para toda $r \geq 0$ y de aquí que $p_{ij_i}^{(n)} > 0$, $\forall\, i > s$

para toda $n \geq \max_{k \geq i > s} m_i = \mu$ por lo que la suma de los renglones de T^{μ} es menor que uno y de aquí que $|T^{\mu}| < 1$ y tambien $|T| < 1$ por lo que T^n converge a cero, entonces tenemos que:

$$\lim_{n \to \infty} P^n = \begin{pmatrix} I_s & 0 \\ & \\ (I - T)^{-1}U & 0 \end{pmatrix} \text{ para toda } n \geq 1$$

Nótese que esta matriz hace que los vectores de probabilidad $p^{(m)}$ converjan pero el límite puede depender de la elección del vector inicial $p^{(0)}$.

MATRICES ESCINDIBLES

Hasta ahora hemos conocido bastantes propiedades acerca de matrices no negativas y en particular hemos manejado las matrices estocásticas, ahora toca estudiar las matrices *escindibles* que son aquellas que por medio de una matriz permutante se pueden poner como una matriz de cuatro bloques, tal que los bloques de la diagonal son matrices cuadradas y no negativas y el bloque inferior izquierdo es una matriz de ceros no necesariamente cuadrada.

DEFINICION-21.0)

Decimos que una matriz cuadrada P es permutante si se obtuvo por medio de las permutaciones en las columnas de I (matriz identidad del mismo tamaño de P).

Ejemplos:

$$P_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$P_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

son matrices permutantes de tamaño 2 y 3 respectivamente y notemos que sus transpuestas tambien lo son.

OBSERVACION-12.0)

Si P es una matriz permutante entonces $P^{-1} = P^{tr}$

Ya que $(PP^{tr})_{ij} = \sum P_{ir}P_{rj} = \begin{cases} = 1 \text{ si } i = j \\ = 0 \text{ si } i <> j \end{cases}$, es decir por ser P una matriz unitaria.

Si M es una matriz cuadrada del mismo tamaño que P entonces MP es la matriz que se obtiene de intercambiar las columnas de M en el mismo sentido que se intercambiaron las de P, como P^{tr} es tambien una matriz permutante entonces la matriz $P^{tr}M$ es la matriz que se obtiene apartir de M permutando sus renglones en el mismo sentido que se permutaron las columnas de P y por lo tanto $P^{tr}MP$ es una matriz que se obtiene de permutar los renglones y columnas de M en el mismo sentido que se intercambiaron las columnas de P.

Ejemplo:

$$\text{Si} \quad M = \begin{pmatrix} 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad y \quad P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

entonces tenemos que $P^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ y

$$P^{-1}M\,P = \begin{pmatrix} 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{pmatrix}$$

que no es mas que nuestra matriz M permutada de columnas y renglones como lo fuerón P y P^{-1}.

DEFINICION-22.0)

Una matriz $M \in \mathcal{M}_{m\times m}(F)$ es escindible $\Leftrightarrow$ $k \geq 2$ y existe una matriz permutante P tal que:

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

donde C y E son bloques cuadrados posiblemente de diferentes tamaños y el bloque debajo de la diagonal tiene entradas idénticamente cero.

DEFINICION-23.0)

Una matriz es *Inescindible* si no es escindible.

Si se nos diera una matriz cuadrada ($k \times k$) A y nos preguntaramos si es escindible o no, no sería tan fácil contestar usando tan solo la definición, pues existen k! matrices permutantes de ese tamaño. Sin embargo, el siguiente teorema nos dá un camino más fácil para contestar lo anterior.

TEOREMA-17.0)

$$A \text{ es } \textit{inescindible} \Leftrightarrow A + A^2 + \ldots + A^K > 0$$

Antes de dar la demostración veremos unos resultados acerca de matrices no negativas

LEMA-4.0)

Sea A una matriz no negativa de $k \times k$ entonces :

$A + A^2 + \ldots + A^n > 0$ para alguna $n \geq 1 \Leftrightarrow$ para toda i,j $\exists$ $m(i,j)$ tal que $a_{ij}^{(m)} > 0$.

Demostración

$\Rightarrow$) Sea $B = \sum A^s$ y $b_{ij} = \sum a_j^{(s)}$ donde $a_{ij}^{(s)}$ denota la entrada ij de la matriz A^s

entonces $\forall$ ij y $\forall$ s $a_{ij}^{(s)} \geq 0$ y $b_{ij} > 0$

$\therefore$ para cada a_{ij} existe almenos una s tal que $a_{ij}^{(s)} > 0$

□

$\Leftarrow$) Supongamos ahora que $\forall$ ij $\exists$ $m(ij)$ tal que $a_{ij}^{(m)} > 0$

sea $b_{ij} = \sum a_{ij}^{(s)}$ claramente $b_{ij} > 0$

y sea $M = \{m(ij) \mid a_{ij}^{(m)} > 0\}$

y hagamos $r = \min M$, es claro que tenemos un número finito de r's tantas como $k \times k$.

entonces definamos $R = \{r_1, r_2, \ldots, r_{kxk}\}$

sea $n = \max R$ entonces si $b_{ij} = \sum a_{ij}^{(m)}$ se tiene que $b_{ij} > 0$ $\forall i,j$, finalmente sea $B = A + A^2 + \ldots + A^n$

de donde vemos que $B > 0$

■

TEOREMA-18.0)

A es inescindible $\Leftrightarrow$ i) $\forall i,j\ \exists\ m(ij)$ tal que $a_{ij}^{(m)}>0$

ii) $A + A^2 + \ldots + A^k > 0$

Demostración:

$\Rightarrow$)

Ya vimos que se tiene i) $\Leftrightarrow$ ii) entonces ahora supongamos que i) no se cumple, para así llegar a una contradicción, Supongamos que $\exists\ i,j$ tal que $\forall\ m\ a_{ij}^{(m)} = 0$ y sea $S = \bigcup_{m>0} \{\ i \mid a_{ij}^{(m)}>0\ \}$ y $T = \bigcup_{m>0} \{\ i \mid a_{ii}^{(m)}>0\ \}$

P.D. que $S \cap T = \phi$

supongamos que $S \cap T \neq \phi$

sea $i \in S \cap T$ entonces $a_{ii}^{(m)} > 0$ para alguna m y $a_{ii}^{(n)}>0$ para alguna n pero por hipótesis $a_{ij}^{(m+n)} = 0$

por otra parte $a_{ij}^{(m+n)} = \sum a_{ii}^{m} + a_{ij}^{n} \geq a_{ii}^{m}\ a_{ij}^{n} > 0$ lo cual es una contradicción y $\therefore$ $S \cap T = \phi$

S es distinto del ϕ, ya que si $S = \phi$ entonces para toda i $a_{ij} = 0$ y de aqui que la columna j de A es de puros ceros, es decir

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots 0 \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots 0 \ldots & a_{2k} \\ \cdot & \cdot & 0 \ldots & a_{3k} \\ \vdots & & \vdots & \vdots \\ a_{k1} & a_{k2} & \ldots 0 \ldots & a_{kk} \end{pmatrix}$$

$\uparrow$ columna j

la cual por medio de una matriz permutante P la podemos poner de la forma siguiente

$$A = \left(\begin{array}{c|ccc} 0 & a_{12} \ldots & a_{11} & \ldots a_{1k} \\ \hline \cdot & \cdot & & \cdot \\ \cdot & & \ddots & \cdot \\ \cdot & & & \cdot\ \cdot \\ 0 & a_{k2} \ldots & a_{k1} & \ldots a_{kk} \end{array}\right)$$

que nos diría que A es inescindible tomando C = 0 y E igual al otro bloque de la diagonal, de la misma manera con T.

Observemos que si $p \in S$ y $i \notin S$ entonces $a_{ip} = 0$ pues si $p \in S$ entonces $a_{pj}^{(m)} > 0$ para alguna m

y como $i \notin S$ entonces $a_{ij}^{(m)} = 0$ para toda m

$\therefore\ 0 = a_{ij}^{(m+1)} \geq a_{ip}\, a_{pj}^{(m)} = 0 \quad \therefore a_{ij} = 0$

$\Leftarrow$)Supongamos ahora que A es escindible, entonces existe una matriz permutante P tal que

$$P^{tr} A\, P = \begin{pmatrix} C & D \\ 0 & E \end{pmatrix}$$

pero $$P^{tr} A^n P = \begin{pmatrix} A^n & (D \sum A^{n-i} E^{i-1} \\ 0 & (E^n) \end{pmatrix}$$

por el lema anterior sabemos que $\forall\ ij\ \exists\ m(ij)$ tal que $a_{ij}^{(m)} > 0$ y $\exists$ n tal que $\sum_{m=0}^{n} A^m > 0$ y $\therefore P^{tr}[\sum_{m=0}^{n} A^m]P > 0$ lo cual es una contradicción.

□

<u>TEOREMA-19.0</u>)

Si $A \geq 0$ y $n \geq 1$ entonces $a_{ij}^{(n)} > 0 \Leftrightarrow \exists\ j_1, j_2, \ldots, j_{n-1}$

tal que $a_{ij_1} a_{j_1 j_2} \ldots a_{j_{n-1} j} > 0$

$\Rightarrow$)Demostración (por inducción sobre n)

i)base inducción

$n = 2$

Si $a_{ij}^{2} > 0$ tenemos que $a_{ij}^{2} = \sum^{k} a_{il} a_{lj} > 0$ por lo tanto $\exists\ l = j_1$ tal que $a_{ij_1} a_{j_1 j} > 0$

ii)hipótesis de inducción

Supongamos que si $a_{ij}^{(n-1)} > 0$ entonces $\exists\ j_1, j_2, \ldots, j_{n-2}$ tal que $a_{ij_1} a_{j_1 j_2} \ldots a_{j_{n-2} j} > 0$

iii)P.D. cumple tambien para n

es decir que si $a_{ij}^{(n)} > 0 \overset{PD.}{\Longrightarrow} \exists\ j_1, \ldots, j_{n-1}$ tal que

$a_{ij_1} \ldots a_{j_{n-1}j} > 0$

$a_{ij}^{(n)} = \sum^{k} a_{il}^{(n-1)} a_{lj} > 0 \;\therefore\; \exists\; l = j_{n-1}$ tal que $a_{ij_{n-1}}^{(n-1)} a_{j_{n-1}j} > 0$

$\therefore\; a_{j_{n-1}j} > 0$ y por hipótesis de inducción existen l's iguales

a $j_1, j_2, \ldots j_{n-2}$

$\therefore\; \exists\; j_1, j_2, \ldots j_{n-2}, j_{n-1}$ tal es que

$a_{ij_1} a_{j_1 j_2} \ldots a_{j_{n-2} j_{n-1}} a_{j_{j-1} j} > 0$

∎

$\Leftarrow$)Supongamos que existen $j_1, j_2, \ldots, j_{n-1}$ tal que

si $a_{ij_1} a_{j_1 j_2} \ldots a_{j_{n-1} j} > 0 \Longrightarrow a_{ij}^{(n)} > 0$

Demostración (inducción / n)

i) n = 2

$a_{ij}^{(2)} = \sum a_{il} a_{lj}$ pero por hipótesis existe $l = j_1$ tal que

$a_{ij_1} a_{j_1 j} > 0$ y por lo tanto $a_{ij}^{(2)} > 0$

ii)hipótesis de inducción

supongamos que existen n-2 j's tal que

$a_{ij_1} a_{j_2 j_3} \ldots a_{j_{n-2} j} > 0 \Longrightarrow a_{ij}^{(n-1)} > 0$

iii) P.D. para n

P.D. $a_{ij}^{(n)} > 0$ suponiendo que existen n-1 j's tales que

$a_{ij_1} a_{j_1 j_2} \ldots a_{j_{n-1} j} > 0$

$a_{ij}^{(n)} = \sum a_{il}^{(n-1)} a_{lj}$, por hipótesis existe $l = j_{n-1}$ tal que

$a_{j_{n-1} j} > 0$ y por hipótesis de inducción existen n-2 j's tales

que $a_{ij_{n-1}}^{(n-1)} > 0 \;\therefore\; a_{ij}^{(n)} > 0$ lo cual completa la demostración ∎

Con el lema y los teoremas anteriores podemos demostrar el teorema deseado

TEOREMA-20.0)

Sea A una matriz no negativa de tamaño k $\times$ k entonces:

$$A \text{ es } \textit{inescindible} \Longleftrightarrow A + A^2 + \ldots + A^k > 0$$

Demostración.

La implicación hacia la izquierda esta dada por el teorema anterior. Probemos la implicación hacia la derecha.

==>) Por el teorema anterior sabemos que si A es inescindible entonces existe n tal que $A + A^2 + \ldots + A^n > 0$ por lo que si $n \leq k$ no hay problema, supongamos que $n > k$ y probemos que existe

una $n' < n$ tal que $A + A^2 + \ldots + A^{n'} > 0$

Sea $a_{ij}^{(n)} > 0$ entonces sabemos que existen $j_1, j_2, \ldots, j_{n-1}$ tal que $a_{ij_1} a_{j_1 j_2} \ldots a_{j_{n-1}j} > 0$

fijémonos en los segundos indices e.d. en el siguiente conjunto

$S = \{j_1, j_2, \ldots, j_{n-1}, j\}$ este conjunto tiene n elementos y como $n > k$ entonces existen indices que se repiten e.d. existe l y s tal que $j_l = j_s$ (digamos que $l < s$)

por la existencia de las j's sabemos que pasa lo siguiente

$$a_{ij_1} a_{j_1 j_2} \ldots a_{j_{l-1} j_l} \underbrace{a_{j_l j_{l+1}} \ldots a_{j_{s-1} j_s}}_{R} a_{j_s j_{s+1}} \ldots a_{j_{n-1} j} > 0$$

como cada $a_{j_r j_{r+1}} > 0$ entonces podemos cancelar lo subrayado e.d.

$$R^{-1}(a_{ij_1} a_{j_1 j_2} \ldots a_{j_{n-1} j}) = a_{ij_1} a_{j_1 j_2} \ldots a_{j_{l-1} j_l} a_{j_s j_{s+1}} \ldots a_{j_{n-1} j} > 0$$

$\therefore$ existe $n' < n$ con $n' = n - (s - l)$ y en consecuencia n' j's

tales que $a_{ij_1} a_{j_1 j_2} \ldots a_{j_{n'-1} j} > 0$

y por lo tanto $a_{ij}^{(n')} > 0$ para alguna $n' < n$ si la n' fuera todavía mayor que k repetiríamos el procedimiento

∎

TEOREMA DE FROBENIUS (1912) (21.0)

Sea A una matriz de transición, tal que A es inescindible y de tamaño $k\times k$ con $k > 1$, entonces:

i) existe una matriz de permutación P y un número natural d llamado el periodo de A tal que:

$$P^{-1}A\,P = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 & A_d \\ A_1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & A_2 & 0 & \dots & 0 & 0 & 0 \\ . & 0 & A_3 & \dots & 0 & 0 & 0 \\ . & . & . & \dots & . & . & . \\ . & . & . & \dots & . & . & . \\ 0 & 0 & 0 & \dots & . & 0 & 0 \\ 0 & 0 & 0 & \dots & . & A_{d-1} & 0 \end{bmatrix}$$

Donde los bloques de la diagonal son matrices cuadradas de ceros y

$P^{-1}A^d\ P = \left[\bigoplus_{i=1}^{d} B_i\right]$ (suma directa de matrices B_i)

ii) Todas las matrices B son matrices primitivas y cada una de ellas tiene el mismo radio espectral de A^d

Antes de la demostración del teorema de Frobenius veamos los siguientes lemas.

LEMA-5.0)

Sea A una matriz cuadrada $(k \times k)$ inescindible y sea $[A^r]_1$ y $[A]_1$ la primer columna de A^r y de A respectivamente, entonces

$$[A^r]_1 = 0 \Rightarrow [A]_1 = 0$$

Demostración.
Por reducción al absurdo.

Supongamos que $[A\ ]_1 \neq 0$

Sea $R = \{ r \ /\ [A^r]_1 = 0 \}$

y sea r = minimo de R.

Entonces tenemos que $[A^r]_1 = 0$ pero $[A^{r-1}]_1 \neq 0$ por lo que se debe cumplir que $[A^{r-1}]_1 \neq 0$ y $[A]\ [A^{r-1}]_1 = 0$

supongamos que el renglón j de $[A^{r-1}]_1$ es distinto de cero y denotemoslo por $[A^{r-1}]_1^j$ entonces como $[A^r] = 0$ se tiene que $(A_{ij})([A^{r-1}]_1^j) = 0$ para toda i

$$0 = [A^r]_1 = \underset{A}{\begin{array}{c} \text{columna } j \text{ de } A \\ \downarrow \\ \begin{pmatrix} * & \cdots & 0 & \cdots & * \\ * & \cdots & 0 & \cdots & * \\ \vdots & & \vdots & & \vdots \\ * & \cdots & 0 & \cdots & * \end{pmatrix} \end{array}} \underset{[A^{r-1}]_1}{\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}} \longleftarrow \text{renglón } j \text{ de } [A^{r-1}]_1$$

lo cual es una contradicción pues A es una matriz inescindible.

∎

LEMA-6.0)

Sea $S \subset \mathbb{N}$ tal que S sea cerrado bajo la suma y sea $d = \text{m.c.d}(S)$ entonces a partir de una cierta n_0 suficientemente grande se tiene que $nd \in S$, con $n_0 \leq n$ y $n_0,\ n \in \mathbb{N}$.

Demostración.

Primero probemos que $\text{m.c.d.}(S) = \text{m.c.d.}(\{a_1,\ a_2,\ \dots\ ,\ a_s\})$ para los s primeros elementos de S.

e.d. $\text{m.c.d.}(S) = \text{m.c.d.}(\mathcal{S})$, con $\mathcal{S} = \{a_1,\ a_2,\ \dots\ ,\ a_s\}$

Construyamos la siguiente sucesión de números naturales, sea $\{d_i\}_{i \in \mathbb{N}}$ tal que

$$\begin{aligned} d_1 &= \text{m.c.d.}(a_1) = a_1 \\ d_2 &= \text{m.c.d.}(a_1, a_2) \\ d_3 &= \text{m.c.d.}(a_1, a_2, a_3) \\ d_4 &= \text{m.c.d.}(a_1, a_2, a_3, a_4) \\ &\ \vdots \\ d_s &= \text{m.c.d.}(a_1,\ a_2,\ \dots\ ,\ a_s) \end{aligned}$$

es claro que $d_1 \geq d_2 \geq \dots \geq d_s \geq \dots \geq 1$

y supongamos que $\{d_i\}_{i \in \mathbb{N}}$ converge a d_s , con $d_s \in \mathbb{N}$

p.d. que d_s divide a a_{s+1}

si $d_s = 1$ entonces no hay nada que probar, entonces podemos suponer que $d_s \neq 1$

supongamos que d_s no es un divisor de

a_{s+1}

entonces existe d' tal que

$d' = m.c.d.(a_1, a_2, \ldots, a_s, a_{s+1})$

entonces $d_s \geq d'$ por la convergencia que existe, pero por otra $\{d_i\}_{i \in \mathbb{N}}$ converge a d_s entonces $d' \geq d_s$

$\therefore d' = d_s$

aplicando el mismo razonamiento para $d_{s+2}, d_{s+3}, \ldots$

tenemos que $d_s = m.c.d.(S)$.

Sea $S = \{a_1, a_2, \ldots, a_s\}$ y sea $d = m.c.d.(S)$ donde d lo podemos escribir como combinación lineal de los elementos de S e.d. $d = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_s a_s$ o de manera equivalente

$$1 = \alpha_1(a_1/d) + \alpha_2(a_2/d) + \ldots + \alpha_s(a_s/d)$$

y hagamos $S' = \{ (a_1/d), (a_2/d), \ldots, (a_s/d) \}$ de donde es claro que $m.c.d.(S') = 1$ y es equivalente probar el lema para S' sea $(a_i/d) = b_i$ para toda i entonces $S' = \{b_1, b_2, \ldots, b_s\}$ con $b_1 < b_2 < \ldots < b_s$

entonces es claro que $1 = \alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_s b_s$

Sea $N = \beta_1 b_1 + \beta_2 b_2 + \ldots + \beta_s b_s$ una combinación lineal positiva de elementos de S' entonces hay que ver que $N + 1$, $N + 2, \ldots$ estan en S'.

$$N = \beta_1 b_1 + \beta_2 b_2 + \ldots + \beta_s b_s$$

$$N + 1 = (\beta_1 b_1 + \beta_2 b_2 + \ldots + \beta_s b_s) + (\alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_s b_s)$$
$$= [(\beta_1 + \alpha_1)b_1 + (\beta_2 + \alpha_2)b_2 + \ldots + (\beta_s + \alpha_s)b_s]$$

$N + 2 = \dots$

$\vdots$

$$N + (b_1 - 1) = [\beta_1 b_1 + \beta_2 b_2 + \dots + \beta_s b_s] + [(b_1 - 1)b_1 + (b_1 - 1)b_2 + \dots + (b_1 - 1)b_s]$$

$$= [\{\beta_1 + (b_1 -1)\alpha_1\}b_1 + \{\beta_2 + (b_1 -1)\alpha_2\}b_2 + \dots + \{\beta_s + (b_1 -1)\alpha_s\}b_s]$$

de donde todos los coeficientes deben ser positivos por lo que basta tomar

$$\beta_1 > (1 - b_1)\alpha_1$$
$$\beta_2 > (1 - b_1)\alpha_2$$
$$\vdots$$
$$\beta_s > (1 - b_1)\alpha_s$$

que siempre tiene solución y $\therefore$ para toda $n \geq N$ se tiene que $n + 1, n + 2, \dots \in S'$ y por lo tanto $nd \in S$

$\blacksquare$

Demostración del teorema de *Frobenius*

Consideremos el siguiente conjunto $S = \{ n \in \mathbb{N} \; / \; a^{n}_{11} > 0 \}$
nótese que $S \neq \emptyset$ pues A es inescindible
y sea d como en el lema 6.0 , es claro que S es cerrado bajo suma pues si:

n, m estan en S entonces $a^{(n+m)}_{11} \geq a^{(n)}_{11}\, a^{(m)}_{11} > 0$ por lo tanto

$n + m$ esta en S y en consecuencia $a^{nd}_{11} > 0$.

i) $a^{nd}_{11} > 0$ para toda n suficientemente grande.

Notemos que este conjunto esta determinado por los tiempos n en los cuales la probabilidad de permanecer en nuestro primer estado es positiva.

Así como se ha construido el conjunto de tiempos a los cuales tenemos probailidad positiva de permanecer en nuestro estado 1, también podemos hacer un análisis de los tiempos a los cuales los demás estados se "conectan"* o tienen probabilidad positiva de pasar al estado i.

Veamos el siguiente inciso:

ii) Para toda j $\exists$ r_j con $0 \leq r_j < d$ tal que $a^m_{j1} > 0$ con $m = dn_o + r_j$ para alguna n_o y haciendo:

$S = \{\ k \in \mathbb{N} \mid a^k_{j1} > 0\ \}$ con j fija y $j \in \{1,2,\ldots,k\}$

entonces los elementos de S dejan el mismo residuo r_j módulo d es decir $k = (nd + r_j) \in S$ para toda $n \geq n_0$.

Demostración:

Fijemos al estado j (arbitrario). Como A es inescindible tenemos que existen números naturales t y ℓ tales que

$$a^{(t)}_{j1} > 0 \qquad \wedge \qquad a^{(\ell)}_{1j} > 0$$

Sea r_j el residuo cuando t es dividido por d es decir $t = dp + r_j$. Ahora, para toda $n \in \mathbb{N}$, $a^{(n+\ell)}_{11} \geq a^{(\ell)}_{1j}\ a^{(n)}_{j1}$. De donde si $a^{(n)}_{j1} > 0$ entonces se tiene que $n + \ell$ es un múltiplo de d. Así que si $a^{(m)}_{j1} > 0$ entonces concluimos que $t + \ell$ y $m + \ell$ son ambos múltiplos de d, es decir $t + \ell = dn_1$ y $m + \ell = dn_2$.

Resolvamos el siguiente sistema:

$$\begin{cases} \ell = dn_1 - t \\ \ell = dn_2 - m \end{cases} \Rightarrow \quad dn_1 - t = dn_2 - m \quad \Rightarrow \quad dn_1 + m = dn_2 + t$$

$$\Rightarrow \begin{cases} m = d(n_2 - n_1) + t \\ \text{haciendo } q = n_2 - n_1 \end{cases} \Rightarrow \begin{cases} m = dq + t \\ \text{y sabemos que } t = dp + r_j \end{cases}$$

$$\therefore\ \{\ m = dq + dp + r_j \quad \Rightarrow \quad \{\ m = d(p + q) + r_j$$

haciendo $n_0 = p + q$ tenemos finalmente que $m = dn_0 + r_j$

y sustituyendo tenemos que $a^{(dn+r_j)}_{j1} > 0$

* Diremos que el estado i esta conectado con el estado j si la probabilidad de pasar del estado i al j es estrictamente positiva, e.d. $p_{ij}>0$ o $a_{ij}>0$.

Recíprocamente si $m = dn + r_j$ y $n \geq p$ como $t = dp + r_j$ tenemos entonces que $m - t = dn - dp$ por lo tanto $m = d(n - p) + t$

pero $a_{j1}^{(m)} \geq a_{j1}^{(t)}\, a_{11}^{d(n-p)}$ y por hipótesis sabemos que

$a_{j1}^{(t)} > 0$ y $a_{11}^{d(n-p)} > 0$ para n suficientemente grande, de aqui

que $a_{j1}^{(m)} = a_{j1}^{dn+r_j} > 0$ para toda m suficientemente grande. ∎

Hagamos una clasificación de todos los estados que están conectados con en el estado $\imath$ a tiempos que dejen residuo r módulo d por medio de los siguientes conjuntos

Sea $C_r = \bigcup_{n\geq 0} \{ j \mid a_{j1}^{(dn+r)} > 0 \}$ $r = 0, 1, 2, \ldots, d-1$

es decir C_r es el conjunto de todos los estados j que tienen una probabilidad positiva de pasar al estado uno a tiempos que dejan un residuo r módulo d.

AFIRMACION 1)

$$C_r \text{ no es vacío}$$

pues si $C_r = \emptyset$ entonces la primera columna de A^n sería de ceros es decir que ningún estado estaría conectado con el estado $\imath$ a un cierto tiempo n y de aquí se infiere que al tiempo inicial tampoco, es decir que la primer columna de A tambien sería de ceros, (pues haciendo uso del lema(5) tenemos que A es escindible), lo cual sería una contradicción.

AFIRMACION 2)

$$C_r \cap C_s = \emptyset$$

los conjuntos C_r son ajenos dos a dos, pues supongamos que un estado $j \in C_r \cap C_s$

entonces $a_{j1}^{(dq+r)} > 0$ y $a_{j1}^{(dq'+s)} > 0$ pero por ii)

$dq + r = dn_0 + r_j$ y $dq' + s = dn_1 + r_j$

entonces $d(q - n_0) = (r_j - r)$ y $d(q' - n_1) = (r_j - s)$

de donde $d \mid (r_j - r)$ y $d \mid (r_j - s)$

lo cual es una contradicción pues $r_i = 0, 1, 2, \ldots, d-1$

y por lo tanto $|r_i - r_j| < d$ lo que quiere decir que $d \nmid (r_j - r)$ y $d \nmid (r_j - s)$ a menos que $r_j = r = s$

y de aquí que $C_r \cap C_s = \emptyset$ ∎

iii)Supongamos que los estados i, j estan conectados es decir $a_{ij} > 0$ y que el estado $j \in C_r$ entonces $i \in C_{r+1}$ para toda $0 \le r \le d-1$ pues :

si $j \in C_r \Rightarrow a_{j1}^{(dn+r)} > 0$ y como $a_{ij} > 0$ entonces

$a_{ij}a_{j1}^{(dn+r)} = a_{i1}^{(dn+(r+1))} > 0$ y por lo tanto $i \in C_{r+1}$

Observemos que si $a_{ij} > 0$ y $j \in C_r$ entonces $i \in C_{r+1}$ para cualquier $0 \le r < d$, pues que $j \in C_r$ quiere decir que $a_{j1}^{dn+r} > 0$ con $0 \le r < d$ por lo que $0 < a_{ij}a_{j1}^{dn+r} \le a_{i1}^{dn+(r+1)}$ por lo que concluimos que $i \in C_{r+1}$ hagamos la aclaración que si $r+1 = d$ entonces $i \in C_0$.

Para construir la matriz permutante veamos el siguiente ejemplo que que nos ilustrará de como hacerlo.
Sea

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

Anteriormente construimos los siguientes conjuntos

$$C_r = \bigcup_{n \ge 0} \{ j : a_{ij}^{dn+r} > 0 \} \quad \text{con} \quad r = 0,1,2,\ldots,d-1$$

calculemos entonces a d de la siguiente manera, denotemos por A_n a la primer columna de A^n, colocando un "+" cuando alguna de sus entradas sea distinta de cero., entonces es facil calcular A_n para toda $n \in \mathbb{N}$ de la siguiente manera:
$A_1 = A_1$, $AA_1 = A_2$, $AA_2 = A_3$, etc. y viendo nuestro ejemplo podemos darnos cuenta que:

$$A_1 = \begin{pmatrix} 0 \\ + \\ + \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad AA_1 = A_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ + \\ 0 \\ + \\ 0 \end{pmatrix}, \quad AA_2 = A_3 = \begin{pmatrix} + \\ 0 \\ 0 \\ 0 \\ + \\ 0 \\ 0 \end{pmatrix}, \quad AA_3 = A_4 = \begin{pmatrix} 0 \\ + \\ + \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = A_1$$

entonces nos damos cuenta que formamos un ciclo de longitud tres es decir que a tiempos múltiplos de tres tendremos una probabilidad positiva de estar en un estado j, para toda $1 \leq j \leq k$ vease el siguiente dibujo.

$$A_1 \quad A_2 \quad A_3 \quad A_4 \quad A_5 \quad A_6 \quad A_7 \quad A_8 \quad A_9 \quad A_{10} \quad A_{11} \quad A_{12} \quad A_{13} \quad A_{14} \quad \text{etc.}$$

lo que quiere decir que :

$a_{j1}^{(4)} > 0 \Leftrightarrow a_{j1}^{(1)} > 0$

$a_{j1}^{(2)} > 0 \Leftrightarrow a_{j1}^{(5)} > 0$

$a_{j1}^{(3)} > 0 \Leftrightarrow a_{j1}^{(6)} > 0$

$a_{j1}^{(4)} > 0 \Leftrightarrow a_{j1}^{(7)} > 0$

$a_{j1}^{(5)} > 0 \Leftrightarrow a_{j1}^{(8)} > 0$

para toda j entre 1 y k

o de manera equivalente $a_{j1}^{3n+r} > 0 \Leftrightarrow a_{j1}^{3(n+1)+r} > 0$

con $1 \leq r < d$, $1 \leq j \leq k$ y para toda $n \in \mathbb{N}$ es decir que a intervalos de tiempo de longitud tres regresamos a nuestro estado original o mas bien tenemos cierta probabilidad positiva de estar en nuestro estado original, también podemos pensar que nuestra cadena de Markov sea una función periódica, y de aquí que denotaremos por d a la longitud del periódo, llamandola el periódo de A.

Recordando que si $a_{ij} > 0$ y $j \in C_r$ entonces $i \in C_{r+1}$ y que $C_r = \bigcup \{j \ : \ a_{ij}^{dn+r} > 0 \}$ con $r = 0,1,2,\dots,d-1$ tenemos lo siguiente:

$C_1 = \{2, 3\}$, $C_2 = \{4, 6, 7\}$, $C_3 = C_0 = \{1, 5\}$

Nuestra matriz permutante es aquella que tiene como primeras columnas a las columnas mencionadas en C_0, como siguientes columnas a los elementos de C_1 y como últimas columnas a los elementos de C_2, es decir si numeramos de izquierda a derecha a las columnas de la matriz identidad (I) y asociando cada elemento de cada conjunto con la numeración en I tendrermos el orden en que permutaremos las columnas de I, por lo que la matriz permutante es:

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

Ya vimos que $P^{-1} = P^{tr}$ por lo que haciendo el producto obtenemos:

$$P^{-1} A P = \left(\begin{array}{cc:cc:ccc} 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{array}\right)$$

O haciendo una (3, (2 , 2 , 3), 7) partición de A tenemos de A

obtenemos de igual forma $P^{-1} A P = \begin{pmatrix} 0 & 0 & A_3 \\ A_1 & 0 & 0 \\ 0 & A_2 & 0 \end{pmatrix}$

Habiendo visto el ejemplo anterior procedamos a construir la matriz P para la demostración supongamos que $C_r = \{i_{r_1}, \dots, i_{r_{c_r}}\}$ con $0 \le r < d$ con "c_r" denotandonos la cardinalidad de C_r.
Sea P la matriz permutante cuyas primeras c_0 son las columnas $i_{0_1}, \dots, i_{0_{c_0}}$ de I seguidas de las $i_{1_1}, \dots, i_{1_{c_1}}$ columas de I hasta las últimas $i_{d-1_1}, \dots, i_{d-1_{c_r}}$ columnas de I.

Aplicando la observación anterior tenemos:

$$P^{-1} A P = \begin{pmatrix} 0 & 0 & \dots & 0 & A_d \\ A_1 & 0 & \dots & 0 & 0^d \\ . & A_2 & \dots & . & 0 \\ . & . & . & . & . \\ 0 & . & & . & . \\ 0 & 0 & \dots & .A_{d-1} & 0 \end{pmatrix}$$

Es facil notar que los bloques de la diagonal son bloques cuadrados de tamaño $c_r \times c_r$ y consecuentemente A_{r+1} es de $c_{r+1} \times c_r$ y las únicas entradas positivas estan en las A_i posiblemente con algunas entradas de ceros, visto lo anterior podemos afirmar que $P^{-1}A^dP = \bigoplus_{i=1}^{d} B_i$ de donde haciendo el producto y recordando un ejercicio anterior tenemos que:

$B_1 = A_d\ A_{d-1}A \dots A_2\ A_1$
$B_2 = A_1\ A_d\ A_{d-1} \dots A_3\ A_2$
.
.
.
$B_d = A_{d-1}A_{d-2} \dots A_1\ A_2$

Solo nos resta probar que cada matriz B_i es primitiva es decir que para cada i existe una $n \in \mathbb{N}$ tal que $B_i^n > 0$

Empecemos demostrando que B_1 es una matriz primitiva.
Supongamos que $j \in C_0$ y por hipótesis tenemos que A es inescindible entonces $a_{1j}^{dn}j > 0$ como A es inescindible podemos

afirmar que existe una $l \in \mathbb{N}$ tal que $a^{l}_{1j} > 0$ además puesto que $j \in Co$ tenemos que $a^{dn}_{j1} > 0$ con la n dependiendo de la j pero entonces:

$0 < a^{l}_{1j}\ a^{dn}_{j1} \leq a^{(l+dn)}_{11}$

$\therefore (l + dn) \in S$ entonces $l + dn = md$

$\therefore l = (m - n)d$

$\therefore l = n_j d$

$\therefore\ a^{dn_j}_{1j} > 0$

bien sabemos entonces que $a^{dn_j}_{1j} > 0$ y además que $a^{dn}_{11} > 0$ para toda n suficientemente grande así que:

$a^{d(n_j+n)}_{1j} > 0$ $\therefore$ existe n_o que no depende de j tal que $a^{dn_o}_{1j} > 0$ para toda $j \in C_o$. (con $n_o \geq (n_j + n)$)

Por otra parte si $i \in C_o$ entonces es bastante claro que $a^{dn_i}_{i1} > 0$ para alguna n_i pero de nuevo $a^{dn}_{11} > 0$ para toda n suficientemente grande, así que $a^{d(n_i+n)}_{i1} > 0$, entonces si tomamos cualquier $m_o \geq (n_i + n)$ podemos encontrar m_o que no dependa de i tal que $a^{dn_i}_{i1} > 0$ para toda $i \in C_o$.

$\therefore a^{d(m_o+n_o)}_{ij} \geq a^{dm_o}_{i1}\ a^{dn_o}_{1j} > 0$ y por lo tanto B_1 es primitiva.

∎

Queremos ver tambien que $B_2, \ldots, B_d$ son primitivas es decir $B_2^{n_2} > 0$ para alguna n_2 , $B_3^{n_3} > 0$ para alguna n_3 , hasta $B_d^{n_d} > 0$ para alguna n_d.

Puesto que hemos probado que B_1 es primitiva sabemos que existe una $n_1 \in \mathbb{N}$ tal que $B_1^{n_1} > 0$ bien pero $B_2^{n_1+1} = A_1 B_1^{n_1} A_d \ldots A_2$ y por la observación 0.0 tenemos que $B_2^{n_2} > 0$ para toda $n_2 > n_1$ de manera análoga podemos afirmar que:

$B_3^{n_3} > 0$ para toda $n_3 > n_2$

$B_4^{n_4} > 0$ para toda $n_4 > n_3$

$\vdots$

$B_d^{n_d} > 0$ para toda $n_d > n_{d-1}$

por lo que hemos podido encontrar una n_i para cada B_i tal que cada potencia de las B_i es estrictamente positiva y por lo tanto cada B_i es una matriz primitiva.

Por último demostraremos que el radio espectral de cada matriz B_i es el mismo para toda i e igual al radio espectral de A^d, demostrando en primera instancia la observación siguiente:

OBSERVACION-13.0)

Sea A una matriz primitiva y sea λ cualquier eigenvalor tal que tiene asociado un eigenvector positivo entonces $|A| = \lambda$.

Demostración:

sup $|A| \neq \lambda$

$Ax = \lambda x$ y en general $A^n x = \lambda^n x$

Sea $L = \lim\limits_{n\ \infty} (|A|^{-1}A)$

entonces $(|A|^{-1}A)x = Lx = \lim (\lambda/|A|)^n x$

pero $\lambda < |A|$ $\therefore$ $(\lambda/|A|) < 1$

$\therefore$ $(\lambda/|A|)^n \xrightarrow[n\ \infty]{} 0$

$\therefore$ $\lim\limits_{n\ \infty} (\lambda/|A|)^n x = 0$

pero $L > 0$ y $x > 0$ entonces $Lx > 0$

$\therefore$ $0 < Lx = \lim\limits_{n\ \infty} (\lambda/|A|)^n x = 0$ lo cual es una contradicción

por lo tanto $\lambda = |A|$

■

Puesto que B_1 y B_2 son matrices primitivas tenemos que
$B_1 p = |B_1| p$ p.a. $p > 0$ y $B_2 q = |B_2| q$ p.a. $q > 0$
tambien tenemos que :
$A_1 B_1 = B_2 A_1$ y de aquí que:
$A_1 B_1 p = |B_1| A_1 p$
$\therefore B_2(A_1 p) = |B_1|(A_1 p)$
por lo que $|B_1|$ es un eigenvalor positivo para B_2 teniendo asociado el eigenvector positivo $A_1 p$ y haciendo uso de la observación anterior tenemos que $|B_2| = |B_1|$, análogamente podemos probar que $|B_i| = |B_1|$ para toda $1 \le i \le d-1$.

finalmente la observación 5.0 nos dá la pauta para concluir que $|B_1| = |A|^d$. Con lo cual damos por terminada la demostración del teorema de *Frobenius*.

■

CONCLUSIONES

Con el presente trabajo tenemos los conocimientos necesarios para poder distinguir cuando un Grupo Abeliano se descompone en una suma directa de subgrupos cíclicos y cuando un Espacio Vectorial en una suma directa de subgrupos T-cíclicos o de F[t]-submódulos si nos referimos al anillo de polinomios, pues pudimos notar que W es un subespacio de V si y sólo si W es un F[t]-submódulo de V, con V espacio vectorial.

Así también, tenemos una manera sencilla de calcular la forma canónica Racional y de Jordan con la cual podemos darnos cuenta de la convergencia de algunos tipos de matrices positivas, en particular las estocásticas.

Existen infinidad de sistemas y situaciones reales que cumplen con las características de un proceso de Markov y de este modo ser representadas por una matriz estocástica. Por lo que hay diversas aplicaciones del Teorema de Frobenius, además que nos proporciona una clasificación de los estados y también los periódos de tiempo a los cuales existe una probabilidad positiva de pasar de un estado i a un estado j.

R E F E R E N C I A S.

1 - ALGEBRA LINEAL
STEPHEN H. FRIEDBERG, ARNOLD J. INSEI, LAWRENCE E. SPENCE.
ED. PUBLICACIONES CULTURAL.

2 - ALGEBRA ABSTRACTA
I.N. HERSTEIN.
ED. TRILLAS.

3 - ALGEBRA MODERNA
I.N. HERSTEIN.
ED. TRILLAS.

4 - CURSO DE ALGEBRA MODERNA
HILTON & YEL-CHIANG.

5 - APPLIED LINEAR ALGEBRA
BEN NOBLE & JAMES W. DANIEL
ED. PRENTICE HALL.

6 - MATRIX THEORY AND IT'S APPLICATIONS (*TOPICOS SELECTOS*)
N.J. PULLMAN.
MARCEL DEKKER, INC. NEW YORK AND BASEL.