



INFORME TECNICO INTERNO

Nº. 35

**INSTITUTO DE MATEMATICA DE BAHIA BLANCA
INMABB (UNS - CONICET)**



UNIVERSIDAD NACIONAL DEL SUR

Avda. ALEM 1253 - 8000 BAHIA BLANCA

República Argentina



INFORME TÉCNICO N° 35

Matrices no negativas

por

A. Benedek

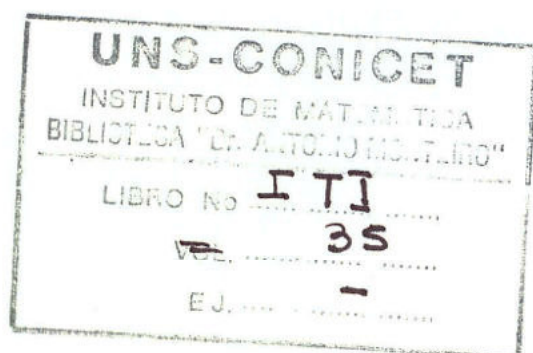
E. Fernández Stacco

R. Panzone

INMABB

UNS - CONICET

AÑO 1994



PROLOGO

Estas lecciones corresponden a un curso optativo que dicté en el primer semestre de 1994. El manuscrito original se formó con los prolijos apuntes tomados en clase por el Lic. Edgardo Fernández. Partiendo de ese material la Dra. Agnes Benedek redactó una afinada versión final y compuso las notas presentes. Salvo por dos teoremas, todos los temas fueron expuestos a los oyentes. El material del curso, con la sola excepción de un ejercicio, se encuentra aquí o allá en el excelente libro de

R. A. Horn y Ch. R. Johnson, **MATRIX ANALYSIS**, Cambridge University Press, 1985, 561 pgs..

Es presentado también en

F. R. Gantmacher, **MATRIZENRECHNUNG**, VEB Deutscher Verlag der Wissenschaften, Berlin, 1958,

y por cierto en la última versión de este clásico de la teoría de matrices, editado por Springer (1986) y titulado **MATRIZENTHEORIE**.

El tema es estudiado especialmente en

E. Seneta, **NONNEGATIVE MATRICES**, Wiley, New York, 1973.

El propósito del curso es demostrar el teorema de Perron. probado originalmente para matrices positivas, es decir, aquellas cuyos elementos son todos positivos. Y a partir de allí presentar su extensión, esencialmente debida a Frobenius, a matrices primitivas, que son las no negativas e irreducibles que poseen un único autovalor de módulo máximo.

R. P.

INDICE GENERAL



0. Definiciones básicas y notación.....	
1. Un modelo de migración.....	2
2. Otras notaciones y definiciones.....	4
3. Matrices elementales y operaciones con ellas.....	6
4. Cambio de base en espacios vectoriales.....	7
5. Irreducibilidad y conexión.....	9
6. Normas vectoriales.....	12
7. Normas matriciales.....	13
8. Matrices unitarias, hermitianas y normales.....	16
9. Problemas.....	19
10. Matrices positivas. 1ª parte del teorema de Perron.....	21
11. Problemas.....	24
12. Autovalores y autovectores.....	25
13. Consecuencias del teorema de Schur.....	31
14. Series de matrices.....	34
15. Teorema de Perron, continuación.....	35
16. Localización de los autovalores.....	39
17. Matrices no negativas.....	41
18. Generalizaciones de algunos resultados a matrices no negativas.....	44

INDICE ALFABETICO

autoespacio.....	30
autovalor.....	1
autovector.....	1
autovector a izquierda.....	31
autovector de Perron.....	23
autovalor simple.....	1
base.....	7
base ortonormal.....	16
bola unitaria, $\overline{B_1(0)}$	12
cambio de base.....	7
camino.....	10
cápsula convexa.....	37
ciclo.....	10
ciclo trivial.....	9
combinación convexa.....	37
conmutante (familia).....	28
convexo.....	37
cota de Cauchy.....	24
cota de Carmichael y Mason.....	25
cota de Montel.....	24
defectiva.....	31
disco de Geršgorin.....	40
espacio nulo.....	30
espectro.....	1
estrictamente diagonalmente dominante.....	34
grafo fuertemente conexo.....	10
grafo $\Gamma(A)$	9
idempotente.....	26
identidades de Newton.....	20
interior.....	13
invariante.....	31
longitud de un camino.....	10

M(A).....	9
matriz acompañante.....	24
matriz adjunta.....	7
matriz convergente.....	32
matriz diagonalizable.....	3
matriz doblemente estocástica.....	37
matriz estocástica.....	37
matriz hermitiana.....	18
matriz irreducible.....	1
matriz módulo.....	4
matriz normal.....	18
matriz ortogonal.....	16
matriz ortogonalmente diagonalizable.....	18
matriz permutación.....	1
matriz positiva.....	4
matriz primitiva.....	42
matriz reducible.....	1
matriz semejante.....	3
matriz traspuesta.....	1
matriz unitaria.....	7
matriz unitariamente diagonalizable.....	18
menores principales.....	19
minimal.....	31
modelo de migración.....	2
momento.....	20
multiplicidad algebraica.....	22
multiplicidad geométrica.....	22
nilpotente.....	26
norma.....	12
norma absoluta.....	13
norma espectral.....	15
norma inducida.....	15
norma matricial.....	14
norma monótona.....	13

norma vectorial.....	12
operador lineal.....	7
ortogonalmente semejante.....	3
ortonormal.....	8
polinomio característico.....	17
prenorma.....	33
producto escalar.....	8
propiedad FC.....	9
radio espectral.....	1
raíz de Perron.....	24
región conexa maximal.....	39
región de Geršgorin.....	40
representación B_2-B_1	7
rulo = ciclo trivial.....	9
semejante.....	3
submatriz principal.....	4
teorema espectral.....	18
traza.....	17
unitariamente semejante.....	3
vector de Perron a izquierda.....	24

ANALISIS MATRICIAL

Y

MATRICES NO NEGATIVAS

por A. Benedek, E. Fernández y R. Panzone.

0. DEFINICIONES BASICAS Y NOTACION

Denotamos con $M_{n,m}(C)$ al conjunto de matrices de n filas y m columnas. Los elementos de las matrices serán números complejos. Si $n=m$ escribiremos $M_n(C)$ en lugar de $M_{n,n}(C)$.

Definición 1. 1) λ es un autovalor de $A \in M_n$ si existe un vector columna $x \neq 0$, un autovector correspondiente a λ , tal que

$$(1) \quad Ax = \lambda x.$$

2) Diremos que el autovalor es *simple* si dos autovectores verificando la ecuación (1) son linealmente dependientes.

3) Con $\sigma(A)$ indicamos el conjunto de autovalores de A y lo denominaremos *espectro* de A .

4) El *radio espectral* de A , que denotamos con $\rho(A)$, es el número

$$\rho(A) = \sup \{ |\lambda|; \lambda \in \sigma(A) \}.$$

5) Diremos que $P \in M_n$ es una *matriz de permutación* si sus elementos son ceros o unos y tiene en cada fila y en cada columna exactamente un 1. Sea $a_{i\sigma(i)}$ el elemento = 1 en la fila

i . Entonces $\sigma(i)$ es una permutación (¿porqué?) y

$$\det(P) = \pm a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \pm 1. \text{ Vale además que } P^{-1} = P^T.$$

(Con A^T indicamos la *traspuesta* de A). O sea P es ortogonal.

6) Diremos que $A \in M_n$ es *reducible* si:

a) $n=1$, $A=0$, o bien

b) $n>1$ y existen una matriz de permutación P y un r ,

$1 \leq r \leq n-1$, tal que

$$P^T A P = \begin{vmatrix} B & C \\ 0 & D \end{vmatrix}, \text{ con } B \in M_r, D \in M_{n-r}, C \in M_{r,n-r}.$$

Ejercicio. En una matriz reducible aparecen por lo menos $n-1$ elementos nulos.

7) $A \in M_n$ se dice *irreducible* si no es reducible.

Consideremos el sistema lineal $Ax=y$, $A \in M_n$ y reducible, $n > 1$. Sea $\tilde{A}$ la matriz que aparece en 6)

$$\tilde{A} = P^T A P = \begin{vmatrix} B & C \\ 0 & D \end{vmatrix}$$

De aquí, si definimos $\tilde{x} = P^T x$ e $\tilde{y} = P^T y$, resulta $\tilde{A}\tilde{x} = \tilde{y}$.

Escribiendo: $\tilde{x} = \begin{vmatrix} z \\ u \end{vmatrix}$, $\tilde{y} = \begin{vmatrix} w \\ v \end{vmatrix}$, el último sistema se reduce a $\begin{cases} Bz + Cu = w \\ Du = v \end{cases}$, que son dos sistemas de orden $(n-r)$ y r respectivamente.

1. UN MODELO DE MIGRACION.

Comenzamos tratando un modelo donde aparecen naturalmente las matrices no negativas. Es un problema de migraciones entre n estados vecinos C_i , $i=1, \dots, n$. Suponemos que entre los habitantes no hay ni muertes ni nacimientos y que cada mes una proporción fija a_{ij} de los habitantes de C_j se traslada a C_i :

$$C_j \xrightarrow{a_{ij}} C_i$$

Esto es, denotando con $p_j^{(m)}$ la población de C_j en el mes m , $a_{ij}p_j^{(m)}$ = parte de la población de C_j que pasa a C_i en ese mes. Como la suma en i de estas partes debe dar toda la población $p_j^{(m)}$, debe cumplirse $\sum_{i=1}^n a_{ij} = 1$, $0 \leq a_{ij} \leq 1$.

Además, $p_i^{(m+1)} = a_{i1}p_1^{(m)} + a_{i2}p_2^{(m)} + \dots + a_{in}p_n^{(m)}$.

Escribiremos esto en forma matricial. Sea A la matriz $A = [a_{ij}]$,

y $p^{(m)}$ la matriz columna: $p^{(m)} = \begin{vmatrix} p_1^{(m)} \\ \vdots \\ p_n^{(m)} \end{vmatrix}$, $m = 0, 1, \dots$.

Se tiene entonces, $p^{(m+1)} = A p^{(m)} = A(A p^{(m-1)}) = \dots = A^{m+1} p^{(0)}$. Vamos a tomar el caso mas simple, $n = 2$. En este caso,

$$A = \begin{vmatrix} 1 - \alpha & \beta \\ \alpha & 1 - \beta \end{vmatrix} \quad \text{con } 0 \leq \alpha, \beta \leq 1.$$

Esta matriz tiene como autovalores $\lambda_1 = 1 - \alpha - \beta$ y $\lambda_2 = 1$. Observemos que $1 = |\lambda_2| \geq |\lambda_1|$, luego el radio espectral de A , $\rho(A)$, verifica $\rho(A) = 1 =$ autovalor de A . Si además $\alpha, \beta > 0$ entonces A es irreducible y el autovalor $\rho(A)$ es simple.

Supongamos simplemente que $\alpha + \beta > 0$. Entonces los autovalores son distintos y se verifica: $x = \begin{vmatrix} \beta \\ \alpha \end{vmatrix}$ es autovector para $\lambda_2 = 1$,

$z = \begin{vmatrix} 1 \\ -1 \end{vmatrix}$ es autovector para $\lambda_1 = (1 - \alpha - \beta)$. En estas condiciones A es diagonalizable, es decir existe una matriz no singular

S tal que $S^{-1} A S$ es una matriz diagonal: $\Lambda = \begin{vmatrix} 1 & 0 \\ 0 & 1 - \alpha - \beta \end{vmatrix}$.

Dicha matriz es $S = \begin{vmatrix} \beta & 1 \\ \alpha & -1 \end{vmatrix}$ y $S^{-1} = \frac{1}{\alpha + \beta} \begin{vmatrix} 1 & 1 \\ \alpha & -\beta \end{vmatrix}$.

Definición 2. $B \in M_n$ se dice *semejante* a $A \in M_n$, ($A \approx B$), si existe una matriz no singular $S \in M_n$ tal que

$$B = S^{-1} A S.$$

$\approx$ es una relación de equivalencia. A es *diagonalizable* si es semejante a una matriz diagonal. Si S es unitaria (ortogonal) diremos que B es *unitariamente (ortogonalmente) semejante* a A . Vamos a completar el ejemplo. Supongamos que $0 < \alpha + \beta < 2$. Esto implica que $|\lambda_1| < 1$. Por tanto,

$$A = S \Lambda S^{-1}, \quad A^m = S \Lambda^m S^{-1}.$$

Estamos interesados en el comportamiento asintótico de la población. Se calcula inmediatamente

$$\Lambda^m = \begin{vmatrix} 1 & 0 \\ 0 & (1 - \alpha - \beta)^m \end{vmatrix} \quad \text{de donde}$$

$$\lim A^m = S \cdot \lim \Lambda^m \cdot S^{-1} = S \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} S^{-1} = \frac{1}{\alpha + \beta} \begin{vmatrix} \beta & \beta \\ \alpha & \alpha \end{vmatrix}.$$

Para $p^{(m)}$ se tiene:

$$\lim p^{(m)} = \lim (A^m p^{(0)}) = (\lim A^m) p^{(0)} = [(p_1^{(0)} + p_2^{(0)})/(\alpha + \beta)] \begin{vmatrix} \beta \\ \alpha \end{vmatrix}.$$

Hemos verificado entonces en este caso los puntos del siguiente:

Teorema 1. Sea $A \geq 0$.

1) $\rho(A)$ es un autovalor, i.e. $\exists x \neq 0$ tal que $Ax = \rho(A)x$.

2) Si A es irreducible entonces $x > 0$.

3) Si $A > 0$ entonces $\rho(A) > 0$ es simple y $\left[\frac{A}{\rho(A)} \right]^m$

converge a una matriz de rango 1 cuyas columnas son proporcionales a x .

4) En todo caso vale: $\frac{1}{m} \sum_{j=1}^m A^j$ converge (media Cèsaro).

Ejemplo. Si $\alpha + \beta = 2$, o sea si $\alpha = 1$, $\beta = 1$,

$$A = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \quad A^{2j} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \quad A^{2j+1} = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \quad \text{luego}$$

$$\lim ((1/m) \sum_{j=1}^m A^j) = \begin{vmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{vmatrix}.$$

2. OTRAS NOTACIONES Y DEFINICIONES.

1) Sean $A, B \in M_n$. Diremos que $A > B$ si $a_{ij} > b_{ij}$ para todo par ij .

En particular, llamaremos matriz *positiva* a A si $a_{ij} > 0 \forall i, j$.

Análogamente diremos que $A \geq 0$, A no negativa, si $a_{ij} \geq 0 \forall i, j$. $A \geq B$ sii $A - B \geq 0$.

2) Diremos que $S \in M_r$, $r < n$, es una *submatriz principal* de $A \in M_n$ si existen $1 \leq k_1 < \dots < k_r \leq n$ tales que $s_{ij} = a_{k_i k_j}$.

3) Sea $A \in M_n$; $\|A\|_2 := \left(\sum_{i,j=1}^n |a_{ij}|^2 \right)^{1/2}$.

4) Si $A \in M_{n,m}$, $|A|$ es la matriz cuyo elemento ij es $|a_{ij}|$.

Ejercicios.

i) $|A| \geq 0$. $|A| = 0$ sii $A = 0$.

ii) $|aA| = |a| \cdot |A|$, si $a \in \mathbb{C}$.

iii) $|A+B| \leq |A| + |B|$.

iv) $A \geq 0$, $A \neq 0$ no implica $A > 0$.

v) $A \geq 0$, $B \geq 0$, $a, b \geq 0 \implies aA + bB \geq 0$.

- vi) $A \geq B, C \geq 0 \implies A + C \geq B + C.$
- vii) $A \geq B, B \geq C \implies A \geq C.$
- viii) Si x es un vector columna, $|Ax| \leq |A||x|.$
- ix) $|AB| \leq |A||B|.$
- x) $|A^m| \leq |A|^m.$
- xi) $0 \leq A \leq B, 0 \leq C \leq D \implies 0 \leq AC \leq BD.$
- xii) $A \geq 0 \implies A^m \geq 0; A > 0 \implies A^m > 0.$
- xiii) $A > 0, x \geq 0, x \neq 0 \implies Ax > 0.$
- xiv) $A \geq 0, x > 0, Ax = 0 \implies A = 0.$
- xv) $|A| \leq |B| \implies \|A\|_2 \leq \|B\|_2.$
- xvi) $\| |A| \|_2 = \|A\|_2.$
- xvii) La inversa de una matriz positiva no-singular tiene elementos negativos.

Teorema 2. Sean $A, B \in M_n$ tales que $|A| \leq B$. Entonces

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

Demostración. Mas adelante se demostrará que para cualquier matriz $A \in M_n$, $\rho(A) = \lim_{m \rightarrow \infty} (\|A^m\|_2)^{1/m}$. Usando este resultado, el teorema sigue si se observa que (x),xi),xv),xvi), ejercicio),

$$|A^m| \leq |A|^m \leq B^m \text{ y por lo tanto } \|A^m\|_2 \leq \||A|^m\|_2 \leq \|B^m\|_2.$$

Corolario 1. Si $0 \leq A \leq B$ entonces $\rho(A) = \rho(|A|) \leq \rho(B)$.

Corolario 2. Sea $A \geq 0, A \in M_n$. Sea $\tilde{A}$ una submatriz principal. Entonces $\rho(\tilde{A}) \leq \rho(A)$. (En particular $a_{ii} \leq \rho(A), \forall i$.)

Demostración. Sea $\tilde{A} \in M_r$. Indicamos con $\hat{A}$ a la matriz que se obtiene a partir de A si en ella se cambian los elementos que no figuran en $\tilde{A}$ por 0. Entonces $0 \leq \hat{A} \leq A$, y por lo tanto $\rho(\hat{A}) \leq \rho(A)$. Observando que los autovalores no nulos de $\tilde{A}$ y $\hat{A}$ son los mismos (¿porqué?) tenemos $\rho(\tilde{A}) = \rho(\hat{A})$, y el corolario sigue.

Ejemplo. $A = \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix}; \sigma(A) = \{0\}. \rho(A) = 0 = \max \{a_{ii}; i=1,2\}.$

Corolario 3. $0 \leq A < B \implies \rho(A) < \rho(B)$.

Demostración. Por el corolario 2, $\rho(B) > 0$. Entonces el corolario 3 vale en el caso en que $\rho(A)=0$. Supongamos $\rho(A) > 0$.

Sea $\alpha > 1$ tal que $0 \leq A \leq \alpha A < B$. Usando el teorema 2 tenemos $\rho(\alpha A) \leq \rho(B)$. Pero $\rho(\alpha A) = \alpha \cdot \rho(A)$ (¿porqué?), luego $0 < \rho(A) < \alpha \cdot \rho(A) = \rho(\alpha A) \leq \rho(B)$, QED.

3. MATRICES ELEMENTALES Y OPERACIONES CON ELLAS.

Consideremos las siguientes operaciones sobre matrices de M_n

- 1) Intercambio de la fila i con la fila j ;
- 2) Multiplicación de la fila i por el escalar c ;
- 3) Reemplazar la fila j por ella mas c veces la fila i .

Llamemos $E_{i,j}$ al resultado de aplicar a la matriz identidad I la operación 1), $M_i(c)$ la operación 2) y $S_{j,i(c)}$ la 3). O sea

$$E_{i,j} = \begin{vmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & \dots & 1 & \dots \\ & & & 1 & & \\ & & & & & \ddots \\ \dots & 1 & \dots & 0 & \dots & \\ & & & & & 1 \\ & & & & & & \ddots \\ & & & & & & & i \\ & & & & & & & j \\ & & & & & & & & 1 \end{vmatrix} \quad M_i(c) = \begin{vmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & c & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & 1 \end{vmatrix} \quad S_{i,j(c)} = \begin{vmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & \dots & & \\ & & & \ddots & & \\ \dots & c & \dots & 1 & \dots & \\ & & & & & \ddots \\ & & & & & & 1 \end{vmatrix}$$

Dada una matriz $A \in M_n$, la matriz producto $E_{i,j}A$ es el resultado de aplicar 1) a A , $M_i(c)A$ es el resultado de aplicar 2), $S_{j,i(c)}A$ es aplicar 3) a la matriz A . Por otra parte $AE_{i,j}$ es el resultado de intercambiar las columnas i y j de A , $AM_i(c)$ multiplicar la columna i de A por c y finalmente $AS_{j,i(c)}$ sumar a la columna i la columna j multiplicada por c . Como toda matriz de permutación P se puede obtener efectuando un número finito de trasposiciones de filas de la matriz identidad, resulta que P se puede escribir como producto de matrices

$P = E(1)E(2)\dots E(k)$, donde $E(i) = E_{j_i k_i}$. Como $E(j)^T = E(j)^{-1} = E(j)$ resulta $P^T = P^{-1} = E(k)\dots E(2)E(1)$.

Definición 3. 1) Denotamos con A^* la matriz adjunta de A que se define como la matriz traspuesta conjugada de A : $a_{ij}^* = \overline{a_{ji}}$.

2) $U \in M_n$ se dice unitaria si $U^* = U^{-1}$.

4. CAMBIO DE BASE EN ESPACIOS VECTORIALES.

Sea V un espacio vectorial n -dimensional y $B_1 = \{v_1, \dots, v_n\}$ una base en V . Esto es, todo elemento $x \in V$ se puede escribir, y de manera única, en la forma $x = \sum_{j=1}^n \alpha_j v_j$, $\alpha_j \in C$.

Establecemos una correspondencia uno a uno entre los elementos de V y los vectores columna $M_{n,1}$ por la aplicación

$$x \longrightarrow [x]_{B_1} = [\alpha_1 \dots \alpha_n]^T. \text{ Sea } T \text{ una transformación}$$

lineal de V en sí mismo y $B_2 = \{w_1, \dots, w_n\}$ otra base en V .

Si tenemos un vector $x \in V$ expresado en la base B_1 y expresamos Tx en la base B_2 tenemos por la linealidad:

$$[Tx]_{B_2} = \left[\sum_j \alpha_j T v_j \right]_{B_2} = \sum_j \alpha_j [T v_j]_{B_2}.$$

$$\text{Sea } [T v_j]_{B_2} = \begin{bmatrix} t_{1j} \\ \vdots \\ t_{nj} \end{bmatrix} \quad \text{y} \quad {}_{B_2}[T]_{B_1} := \begin{bmatrix} t_{11} & \dots & t_{1n} \\ \vdots & & \vdots \\ t_{n1} & \dots & t_{nn} \end{bmatrix}$$

$$\text{Entonces, } [Tx]_{B_2} = {}_{B_2}[T]_{B_1} [x]_{B_1}.$$

La matriz ${}_{B_2}[T]_{B_1}$ se denomina la *representación B_2 - B_1 de T* .

En particular si $B_1 = B_2$, habiendo identificado los elementos de V con vectores columna, un *operador lineal* T corresponde a la multiplicación del vector $[x]_{B_1}$ por la matriz ${}_{B_1}[T]_{B_1}$.

Veamos ahora el *cambio de base*. Para ello basta tomar T la transformación identidad, $Ix = x$. En efecto,

$$[x]_{B_2} = [Ix]_{B_2} = {}_{B_2}[I]_{B_1} [x]_{B_1}, \text{ donde}$$

$$B_2 [I]_{B_1} = [[v_1]_{B_2} \dots [v_n]_{B_2}].$$

Reemplazando $[x]_{B_1} = B_1 [I]_{B_2} [x]_{B_2}$ en la identidad anterior,

$$\text{resulta } [x]_{B_2} = B_2 [I]_{B_1} B_1 [I]_{B_2} [x]_{B_2} \quad \forall x \in V.$$

En consecuencia, $B_2 [I]_{B_1} B_1 [I]_{B_2} = \text{matriz identidad},$

$$\text{y } B_2 [I]_{B_1} = B_1 [I]_{B_2}^{-1} \text{ (luego, estas matrices no son singulares).}$$

Ahora es de verificación inmediata que:

$$[Tx]_{B_2} = B_2 [T]_{B_2} [x]_{B_2} = B_2 [T]_{B_2} B_2 [I]_{B_1} [x]_{B_1}$$

de donde $B_2 [T]_{B_1} = B_2 [T]_{B_2} B_2 [I]_{B_1}$. Análogamente,

$$B_1 [T]_{B_1} = B_1 [I]_{B_2} B_2 [T]_{B_2} B_2 [I]_{B_1}. \text{ O sea, } B_1 [T]_{B_1} \approx B_2 [T]_{B_2}.$$

Dada una base B_2 y $S \in M_n$ no singular, existe una base B_1 tal que $B_2 [I]_{B_1} = S$. En efecto, escribiendo $S = [s_1 \dots s_n]$ donde los s_j son vectores columna, definimos $v_j \in V$ mediante $[v_j]_{B_2} = s_j$. Como S es no singular, la familia $\{v_1, \dots, v_n\}$ es linealmente independiente (probarlo!) y es la base B_1 requerida.

Sea $V = M_{n,1} (= C^n)$. En V se define el *producto escalar* $\langle x, y \rangle$:

$$\langle x, y \rangle = y^* \cdot x = \sum x_j \overline{y_j}. \text{ Una base } B_2 = \{w_1, \dots, w_n\} \text{ se dice}$$

ortonormal si $\langle w_i, w_j \rangle = \delta_{ij}$. Vale

Teorema 3. Sea $B_2 = \{w_1, \dots, w_n\}$ una base ortonormal. Entonces $B_1 = \{v_1, \dots, v_n\}$ es una base ortonormal si y sólo si $S := B_2 [I]_{B_1}$ es una matriz unitaria.

Demostración.

$$\text{Sea } S = \begin{vmatrix} s_{11} & \dots & s_{1n} \\ s_{21} & \dots & s_{2n} \\ \vdots & & \vdots \\ s_{n1} & \dots & s_{nn} \end{vmatrix} = |s_1 \dots s_n|$$

$$\text{Entonces } s_j = [v_j]_{B_2}, \text{ o sea, } v_j = \sum_i s_{ij} w_i.$$

Luego $\langle v_k, v_r \rangle = \langle \sum_i s_{ik} w_i, \sum_j s_{jr} w_j \rangle = \sum_i \sum_j s_{ik} \overline{s_{jr}} \langle w_i, w_j \rangle =$
 $= \sum_i s_{ik} \overline{s_{ir}} =$ elemento rk de la matriz $S^* S$. Entonces B_1 es
 ortonormal si y sólo si $S^* S = I$, QED.

Vamos a enunciar aquí un teorema que probaremos más adelante:

Teorema 4 (Schur). Sea $A \in M_n(\mathbb{C})$ con autovalores $\lambda_1, \dots, \lambda_n$.

Existe una matriz unitaria $U \in M_n$, tal que

$$U^* A U = T = [t_{ij}]$$

es una matriz triangular superior, (esto es, $t_{ij} = 0$ si $j < i$).

Además $t_{ii} = \lambda_i$, y si A y los autovalores son reales, U es real.

5. IRREDUCIBILIDAD Y CONEXION.

Definición 4. Diremos que $A \in M_n$ tiene la propiedad *FC* (fuerte conexión) si para $p \neq q$, $1 \leq p, q \leq n$ existen $k_1, \dots, k_m$ ($1 \leq m \leq n$) tales que $k_1 = p$, $k_m = q$ y $a_{k_j, k_{j+1}} \neq 0$ para $j = 1, \dots, m-1$.

Definición 5. Dada $A \in M_n$, indicaremos con $M(A)$ la matriz $[m_{ij}]$ definida por: $m_{ij} = 1$ si $a_{ij} \neq 0$, $m_{ij} = 0$ si $a_{ij} = 0$.

Ejercicio. Son equivalentes:

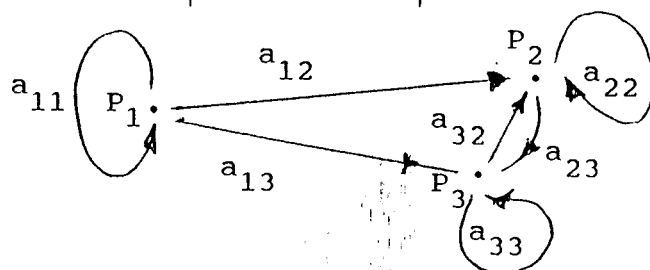
- 1) $A \in FC$
- 2) $|A| \in FC$
- 3) $M(A) \in FC$.

Definición 6. Asociamos a cada matriz $A \in M_n(\mathbb{C})$ un grafo $\Gamma(A)$ formado por n nodos $P_1, \dots, P_n$, unidos por flechas (= arco orientado) de la siguiente manera:

Si a_{ij} no es cero se tiende un flecha de P_i a P_j .

Ejemplo.

Para la matriz $A = \begin{vmatrix} -1 & -2 & 1 \\ 0 & 3 & 4 \\ 0 & 5 & 6 \end{vmatrix}$, $\Gamma(A)$ será el de la figura



La flecha que une un nodo a sí mismo se la denomina *rulo* o *ciclo trivial*.

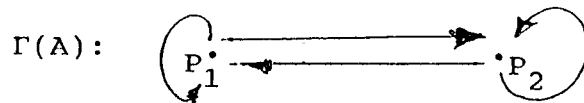
Definición 7. Por un *camino* se entiende una sucesión de flechas $P_i P_j, P_j P_k, \dots, P_r P_h$. Por *longitud* de un camino entendemos el número de arcos que lo componen. Un camino que comienza y termina en un mismo nodo y no llega dos veces a ningún nodo se denominará *ciclo*.

Diremos que el grafo $\Gamma(A)$ es *fuertemente conexo* si para cada par de nodos P_i, P_j , distintos hay un camino que comienza en P_i y termina en P_j . Dejamos al lector verificar que vale el

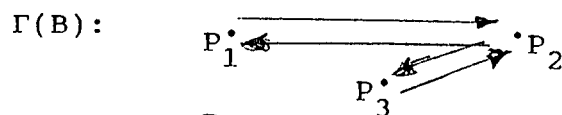
Teorema 5. $A \in FC$ si y sólo si $\Gamma(A)$ es fuertemente conexo.

Ejemplo. A y B son irreducibles, $A, B \in FC$. $C \notin FC$ y no es irreducible.

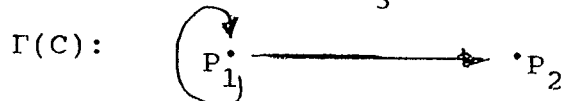
$$A = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}$$



$$B = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix}$$



$$C = \begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix}$$



Teorema 6. Sea $A \in M_n(C)$. Las siguientes proposiciones son equivalentes: a) A es irreducible;

b) $(I + M(A))^{n-1} > 0$;

c) $(I + |A|)^{n-1} > 0$

d) $\Gamma(A)$ es fuertemente conexo

e) $A \in FC$

Para la demostración necesitamos el siguiente

Lema 1. Sea $m \leq n-1$. $(|A|^m)_{ij} > 0$ si y sólo si existe un camino de longitud m que parte de P_i y llega a P_j .

Demostración. $(|A|^m)_{ij}$ es suma de términos no negativos de la forma $|a_{ik_1}| |a_{k_1 k_2}| \dots |a_{k_{m-1} j}|$. El lema sigue.

Demostración del teorema. Observemos que si hay un camino de P_i a P_j entonces hay uno de longitud $< n$. En efecto, si un camino tiene longitud $\geq n$ entonces hay un vértice P_k donde llega dos

veces. Si se suprime el tramo que parte de P_k por primera vez y llega a P_k por última vez, resulta un nuevo camino de P_i a P_j en el cual no se repite el vértice P_k . Después de suprimir en esta forma todos los vértices repetidos, el camino tendrá longitud $< n$.

c) $\iff$ d). $(I + |A|)^{n-1} = I + (n-1)|A| + \dots + |A|^{n-1} > 0$ si y sólo si $\forall i \neq j$ hay un h , $1 \leq h \leq n-1$, con $(|A|^h)_{ij} > 0$.

En vista del lema, esto equivale a que hay un camino de longitud menor que n desde P_i hasta P_j . La observación precedente concluye la prueba.

b) $\iff$ c), se deja al lector.

d) $\iff$ e), es el teorema 5.

Veamos que c) $\implies$ a) probando la contrarrecíproca.

Si A es reducible existe una matriz de permutación P tal que $P^T \tilde{A} P = A$ con $\tilde{A}$ tal que $\tilde{a}_{ij} = 0$ si $i > r$, $j \leq r$, para

cierto r , $1 \leq r < n$. La matriz $|\tilde{A}|^h$ tiene nulos estos mismos elementos (¿porqué?). Entonces $(I + |\tilde{A}|)^{n-1}$ no es positiva,

ni lo es $P^T(I + |\tilde{A}|)^{n-1}P = (P^T(I + |\tilde{A}|)P)^{n-1} = (I + |A|)^{n-1}$.

Veamos que a) $\implies$ c) también probando la contrarrecíproca.

Sea $(I + |A|)^{n-1}$ no positiva. Existen entonces $p \neq q$

tales que $(|A|^h)_{pq} = 0$ para $h = 1, \dots, n-1$. Entonces no hay

camino de P_p a P_q en el grafo $\Gamma(A)$.

Separamos los nodos en dos clases:

$S_1 = \{P_j : P_j = P_q \text{ ó existe un camino de } P_j \text{ a } P_q\}$

$S_2 = \{P_j : P_j \notin S_1\}$. Como $P_q \in S_1$ y $P_p \in S_2$, ambos son

no vacíos. Si $P_h \in S_2$, entonces no hay un camino desde P_h a ningún $P_j \in S_1$ (¿porqué?). O sea, $a_{hj} = 0$ si $P_h \in S_2$ y $P_j \in S_1$.

Renumeremos los nodos de manera que

$$S_1 = \{\tilde{P}_1, \dots, \tilde{P}_r\} \quad S_2 = \{\tilde{P}_{r+1}, \dots, \tilde{P}_n\}$$

donde $\tilde{P}_j = P_{\sigma(j)}$, σ una permutación de $\{1, \dots, n\}$.

Entonces el nuevo grafo corresponderá a la matriz $\tilde{A} = [\tilde{a}_{ij}]$ con $\tilde{a}_{ij} = a_{\sigma(i)\sigma(j)}$, o sea, $\tilde{A} = P^T A P$, P matriz de permutación. Como en ese grafo no hay camino de ningún nodo de S_2 a uno de S_1 , vale: $\tilde{a}_{ij} = 0$ para $i > r$, $j \leq r$. Luego, A es reducible, QED.

6. NORMAS VECTORIALES.

Sea C^n el espacio vectorial de las n -uplas de números complejos pensados como vectores columna. Sabemos que todo espacio vectorial de dimensión n es isomorfo a ese espacio.

Representamos con x los elementos de C^n : $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Definición 8. Llamamos *norma* en C^n a una función

$\| \cdot \| : C^n \longrightarrow [0, \infty)$ con las propiedades:

- 1) $\|x\| \geq 0$
- 2) $\|x\| = 0$ sii $x=0$
- 3) $\|c \cdot x\| = |c| \cdot \|x\|$ (homogeneidad)
- 4) $\|x + y\| \leq \|x\| + \|y\|$ (desigualdad triangular)

Denotamos con $B_r(0) := \{x; \|x\| < r\}$ y llamaremos *bola unitaria*

al conjunto $\overline{B_1(0)} = \{x; \|x\| \leq 1\}$.

Teorema 7. Dos normas $\| \cdot \|$, $[\cdot]$, sobre C^n son siempre equivalentes. Esto es, existen constantes positivas c_1 y c_2 tales que para todo $x \in C^n$: $c_1 \cdot [x] \leq \|x\| \leq c_2 \cdot [x]$.

Demostraremos este resultado luego y en un contexto más general.

Sin demostración enunciamos dos teoremas :

Teorema 8. Sea $B := \overline{B_1(0)}$ la bola unitaria en C^n respecto a una norma. Entonces el conjunto B verifica:

- a) B es acotado y cerrado (compacto).
- b) B es equilibrado : $|c|=1, c \in C, x \in B \implies cx \in B$.
- c) B es convexo : si $x, y \in B$ entonces $rx + (1-r)y \in B$ para todo número real $r \in [0,1]$.

Recíprocamente,

Teorema 9. Si $B \subset C^n$ es un conjunto compacto, equilibrado,

convexo tal que 0 es un punto interior de B, entonces hay una norma tal que B es la bola unitaria respecto a ella.

(Un punto x se dice *interior* de un conjunto H si existe un número $D > 0$ tal que para todo y de longitud 1 y todo d , $0 \leq d < D$, vale que $x + dy \in H$).

Definición 9. Una norma se dice *monótona* si

$$|x| \leq |y| \implies \|x\| \leq \|y\|.$$

Una norma se dice *absoluta* si $\|x\| = \||x|\|$.

Teorema 10. Una norma es monótona si y sólo si es absoluta.

Demostración. Sea $\|\cdot\|$ monótona, $z \in \mathbb{C}^n$. Tomando $x = z$, $y = |z|$ resulta $|x| = |y|$. La monotonía implica entonces $\|x\| \leq \|y\|$ y $\|y\| \leq \|x\|$, o sea, $\|z\| = \||z|\|$.

Supongamos $\|\cdot\|$ absoluta. Sea $\alpha \in [0, 1]$.

Probaremos primero que: si $x = [x_1 \dots x_n]^T \in \mathbb{C}^n$ e $y = [x_1 \dots x_{k-1} \quad \alpha x_k \quad x_{k+1} \dots x_n]^T$ entonces $\|y\| \leq \|x\|$.

En efecto, llamando $z := [x_1 \dots x_{k-1} \quad -x_k \quad x_{k+1} \dots x_n]^T$ tenemos $|x| = |z|$ y por lo tanto $\|x\| = \|z\|$.

Pero $y = \frac{(1+\alpha)}{2} x + \frac{(1-\alpha)}{2} z$. De la desigualdad triangular :

$\|y\| \leq ((1-\alpha)/2 + (1+\alpha)/2) \cdot \|x\| = \|x\|$. Si reemplazamos α por un número complejo α' con el mismo módulo obtenemos un vector y' tal que $\|y'\| = \|y\| \leq \|x\|$. Ahora la monotonía de la norma sigue fácilmente, QED.

Definición 10. Definiremos aquí algunas normas usuales en $\mathbb{C}^n$.

$$\|x\|_1 := \sum_i |x_i| \qquad \|x\|_2 := \left(\sum_i |x_i|^2 \right)^{1/2}$$

Estos son casos particulares de

$$\|x\|_p := \left(\sum_i |x_i|^p \right)^{1/p} \quad \text{donde el número real } p \in [1, \infty).$$

Si $p = \infty$, $\|x\|_\infty := \max \{|x_i|; i=1, \dots, n\}$.

7. NORMAS MATRICIALES.

Sea $A \in M_n(\mathbb{C})$. Se la puede identificar con un elemento de $\mathbb{C}^{n^2}$ y esa identificación respeta las operaciones de suma y producto por un escalar. Ya definimos qué se entiende por una norma

(vectorial) en $M_n(\mathbb{C})$. Repetimos los axiomas 1) a 4) :

- 1) $\|A\| \geq 0$
- 2) $\|A\| = 0 \iff A = 0$
- 3) $\|cA\| = |c| \|A\|$
- 4) $\|A+B\| \leq \|A\| + \|B\|$

Definición 11. Si una norma en $M_n(\mathbb{C})$ verifica, además de 1)-4),

- 5) $\|AB\| \leq \|A\| \cdot \|B\|$ (submultiplicatividad)

diremos que $\|\cdot\|$ es una *norma matricial*.

Nota. Sea $\|\cdot\|$ una norma matricial. Entonces valen:

- i) $A^2 = A \neq 0 \implies \|A\| = \|A^2\| \leq \|A\|^2 \implies 1 \leq \|A\|$.
- ii) en particular la matriz identidad I verifica $\|I\| \geq 1$.
- iii) $1 \leq \|I\| = \|A A^{-1}\| \leq \|A\| \cdot \|A^{-1}\|$. Luego $\|A^{-1}\| \geq 1/\|A\|$.
- iv) $\|A^k\| \leq \|A\|^k$, o equivalentemente, $\|A^k\|^{1/k} \leq \|A\|$.

Proposición 1. Las normas $\|A\|_1 = \sum_{i,j} |a_{ij}|$, $\|A\|_2 = \left(\sum_{i,j} |a_{ij}|^2 \right)^{1/2}$

son matriciales mientras que $\|A\|_\infty = \max \{|a_{ij}|; i,j=1,\dots,n\}$ no lo es.

Demostración. $\|AB\|_1 = \sum_{i,j} \left| \sum_k a_{ik} b_{kj} \right| \leq \sum_{i,j} \sum_{k,m} |a_{ik}| |b_{mj}| =$
 $= \sum_{i,k} \sum_{m,j} |a_{ik}| |b_{mj}| = \|A\|_1 \cdot \|B\|_1$.

$\|AB\|_2^2 = \sum_{i,j} \left| \sum_k a_{ik} b_{kj} \right|^2$. Aplicando la desigualdad de Schwarz,

$$\|AB\|_2^2 \leq \sum_{i,j} \left(\sum_k |a_{ik}|^2 \right) \left(\sum_m |b_{mj}|^2 \right) = \|A\|_2^2 \|B\|_2^2$$

Veamos que $\|\cdot\|_\infty$ no es matricial. Sea $J = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$; entonces

$$J^2 = 2J, \quad \|J^2\|_\infty = 2 \not\leq \|J\|_\infty^2 = 1.$$

Ejercicio. Probar que $\|A\| := n \cdot \|A\|_\infty$ es matricial.

Ejercicio. Sea U una matriz unitaria. Probar que si $A \in M_n$ entonces $\|UA\|_2 = \|A\|_2 = \|AU\|_2$, ($\|\cdot\|_2$ es unitariamente invariante).

Ejercicio. Probar que $\|A\|_p$ es matricial si y sólo si $1 \leq p \leq 2$.

Dada en $\mathbb{C}^n$ una norma vectorial $\|\cdot\|$, se le puede asociar una

norma matricial $\|A\|$ en $M_n(\mathbb{C})$ definida, para $A \in M_n(\mathbb{C})$, por

$$\|A\| := \sup \{ \|Ax\| ; x \in \mathbb{C}^n, \|x\|=1 \}.$$

La norma matricial así definida se dice *inducida* por la vectorial dada.

Ejercicio. Probar que $\|A\| = \max \{ \|Ax\| ; \|x\|=1 \} = \sup_{x \neq 0} \|Ax\|/\|x\| = \sup \{ \|Ax\| ; \|x\| \leq 1 \}$. Verificar que $\|A\|$ es una norma vectorial.

De aquí, $\|Ax\| \leq \|A\|\|x\|$ y $\|A\|$ es la menor constante k tal que, para todo $x \in \mathbb{C}^n$, $\|Ax\| \leq k \cdot \|x\|$. Veamos que es una norma matricial. Vale $\|(AB)x\| = \|A(Bx)\| \leq \|A\|\|Bx\| \leq \|A\|\|B\|\|x\|$, luego $\|AB\| \leq \|A\|\|B\|$, QED.

Observación. Si $\| \cdot \|$ es una norma inducida e I es la matriz identidad, entonces $\|I\| = 1$. Aplíquese el ejercicio precedente. ¿Cuáles son las normas inducidas por $\| \cdot \|_1$, $\| \cdot \|_2$, $\| \cdot \|_\infty$?

Denotaremos con $\| \cdot \|_p$ a la norma inducida por $\| \cdot \|_p$.

Teorema 11. Vale:

$$i) \quad \|A\|_1 = \max_j \sum_i |a_{ij}|$$

$$ii) \quad \|A\|_\infty = \max_i \sum_j |a_{ij}|$$

$$iii) \quad \|A\|_2 = \max \{ \sqrt{\lambda} ; \lambda \in \sigma(A^*A) \}, \text{ (ésta se denomina también la norma espectral de } A \text{)}.$$

Demostración. i) Designemos con $C(A)$ el miembro derecho de i), y escribamos $A = [a_1 \dots a_n]$ donde cada a_j es una matriz columna.

Tenemos, $\|Ax\|_1 = \|x_1 a_1 + \dots + x_n a_n\|_1 \leq \sum |x_j| \|a_j\|_1 \leq C(A) \|x\|_1$.

Luego, $C(A) \geq \|A\|_1$. Para probar la desigualdad contraria sea $x = e_k$ = el vector con $x_k=1$ y $x_j=0$ para $j \neq k$. Entonces, $\|x\|_1=1$,

$Ax = a_k$, implican $\|Ax\|_1/\|x\|_1 = \|a_k\|_1 \leq \|A\|_1$, o sea, $C(A) \leq \|A\|_1$.

ii) Designemos con $C(A)$ ahora el miembro derecho de ii).

Como antes, $\|Ax\|_\infty = \max_i \left| \sum_j x_j a_{ij} \right| \leq \|x\|_\infty \max_i \sum_j |a_{ij}|$.

Luego, $\|A\|_\infty \leq C(A)$. Para probar la desigualdad contraria, sea k la fila tal que $C(A) = \sum_j |a_{kj}|$, y sea x el vector tal que

$x_j = \overline{a_{kj}} / |a_{kj}|$ si $a_{kj} \neq 0$, $x_j = 1$ en caso contrario.

Entonces $\|x\|_\infty=1$, mientras que $\|Ax\|_\infty \geq |\sum_j a_{kj} x_j| = \sum_j |a_{kj}| = C(A)$.

Luego $\|A\|_\infty \geq C(A)$, lo que prueba ii).

Para ver iii), obsérvese en primer lugar que si λ es autovalor de A^*A entonces, para un cierto vector no nulo x , $A^*Ax = \lambda x$.

De aquí, $x^*A^*Ax = \lambda x^*x$, o sea, $\lambda = \|Ax\|^2/\|x\|^2 \geq 0$. Luego el miembro derecho de iii) está bien definido. Lo denotaremos con $C(A)$. Valen las igualdades

$\|A\|_2^2 = \max (\|Ax\|_2^2 ; \|x\|_2 = 1) = \max (x^*A^*Ax ; \|x\|_2 = 1) =$
 $= \max (x^*U^*A^*AUx ; \|x\|_2 = 1)$, para U una matriz unitaria fija.
 (recuérdese que U mantiene la norma y es 1:1). Usaremos una matriz U de manera que $U^*A^*AU = D = \text{diagonal}$ (tal U existe en virtud del teorema 4, ya que una matriz triangular y autoadjunta es necesariamente diagonal), siendo los elementos diagonales de D los autovalores de A^*A : $\lambda_1, \dots, \lambda_n$.

Entonces $\|A\|_2^2 = \max (x^*Dx ; \|x\|_2=1) = \max (\sum_j \lambda_j |x_j|^2 ; \|x\|_2^2=1) =$
 $= C(A)^2$, QED.

Ejercicio. Sean U y V unitarias. Entonces $\|UAV\|_2 = \|A\|_2$.

Teorema 12. Si $\|\cdot\|$ es una norma matricial entonces $\rho(A) \leq \|A\|$.

Demostración. Sea $\lambda \in \sigma(A)$ tal que $|\lambda| = \rho(A)$ y $x \neq 0$ tal que $Ax = \lambda x$. Sea X la matriz $[x, x, \dots, x]$. Entonces $AX = \lambda X$ y se verifica que $\|AX\| = |\lambda| \cdot \|X\| \leq \|A\| \cdot \|X\|$. Como $X \neq 0$, $|\lambda| \leq \|A\|$, QED.

Corolario. Para cualquier norma matricial $\rho(A) \leq \|A^k\|^{1/k}$.

Demostración. El teorema de Schur implica $\sigma(A^k) = \{\lambda^k; \lambda \in \sigma(A)\}$, de donde sigue que $\rho(A^k) = \rho^k(A) \leq \|A^k\|$, QED.

8. MATRICES UNITARIAS, HERMITIANAS Y NORMALES.

Definición 12. Dados n vectores $\{x_i\}_{i=1}^n$ en C^n se dice que

forman una *base ortonormal* si $x_i^* x_j = \delta_{ij}$.

Recordemos que una matriz se dice *unitaria* si $U^*U = I$.

Definición 13. $U \in M_n(R)$ se dice *ortogonal* si $U^T U = I$.

Teorema 13. Las siguientes proposiciones son equivalentes.

- 1) U es unitaria.
- 2) $\det U \neq 0$ y $U^* = U^{-1}$.
- 3) $UU^* = I$.
- 4) U^* es unitaria.
- 5) Las columnas de U forman una base ortonormal de C^n .
- 6) Las filas de U forman una base ortonormal de C^n .
- 7) $\forall x \in C^n$ vale : $(Ux)^*(Ux) = x^*x$. (O sea U mantiene la longitud euclídea).

Ejemplo. Si $n = 2$, las matrices ortogonales se pueden escribir como

$$T(\theta) = \begin{vmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{vmatrix} \quad \text{ó} \quad \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} T(\theta), \quad \theta \in [0, 2\pi).$$

Dejaremos al lector probar las equivalencias de este teorema ya que la mayoría sigue aplicando la definición, y sólo observamos un hecho general que servirá para ver que 1), 2) y 3) son equivalentes: Si $A, B \in M_n(C)$ y $BA = I$ entonces $\det A \neq 0$ y $AB = I$ (o sea $B = A^{-1}$). En efecto, $Ax = 0 \implies BAx = 0 = x$. Luego, $\det A \neq 0$ y el sistema lineal $Ax = y$ tiene solución (única!) para cada $y \in C^n$. Existe entonces una matriz B_d tal que

$AB_d = I$ y se verifica $B = B(AB_d) = (BA)B_d = B_d$. Luego $B = A^{-1}$.

Ejercicio. Las matrices unitarias forman un subgrupo de $GL(n, C)$ y las matrices ortogonales un subgrupo de $GL(n, R)$.

Teorema 14. 1) Si $\{U_j\}$ es una sucesión de matrices unitarias, entonces existe una subsucesión $\{U_{k_j}\}$ tal que elemento a elemento $U_{k_j} \longrightarrow U$, que también es unitaria.

2) Si A es unitariamente equivalente a B (esto es, $\exists U$ tal que $A = U^*BU$) entonces $\sum_{i,j} |a_{ij}|^2 = \sum_{i,j} |b_{ij}|^2$.

1) es una consecuencia del teorema de Heine-Borel y es dejado al lector. Para demostrar 2) presentamos algunas herramientas.

Definición 14. Se denomina el *polinomio característico* de A al polinomio mónico de grado n definido por

$$p_A(t) := \det \begin{vmatrix} t - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & t - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & t - a_{nn} \end{vmatrix} = \det (tI - A).$$

Vale, $\det (tI - A) = (-1)^n \det (A - tI)$. Observando el desarrollo del determinante se pueden precisar algunos coeficientes de este polinomio: $p_A(t) = t^n - \text{tr}(A) t^{n-1} + \dots + (-1)^n \det A$, ($\text{tr}(A) = \text{traza de } A := \sum a_{jj}$). Si $\lambda_1, \dots, \lambda_n$ son los ceros de este polinomio, i.e., los autovalores de A , entonces: $p_A(t) = (t - \lambda_1) \dots (t - \lambda_n)$.

Si la matriz B es semejante a A , i.e., existe S no singular tal que $B = S^{-1}AS$, entonces $p_A(t) = p_B(t)$. En efecto, esto sigue de $(tI - B) = S^{-1}(tI - A)S$, pues el determinante del producto es el producto de los determinantes, (T. de Cauchy). En particular, las trazas de dos matrices semejantes son iguales (!).

Demostración de 2). $\sum_{i,j} |a_{ij}|^2 = \sum_j \sum_i a_{ji}^* a_{ij} = \text{tr} (A^* A) =$

$\text{tr} (U^* B^* U U^* B U) = \text{tr} (U^* B^* B U) = \text{tr} (B^* B) = \sum_{i,j} |b_{ij}|^2$, QED.

Definición 15. 1) Se dice que $A \in M_n(\mathbb{C})$ es *normal* si $A^* A = A A^*$.

2) A se dice *hermitiana* si $A^* = A$.

3) A se dice *unitariamente diagonalizable* si $\exists U$ unitaria tal que $U^* A U = D = \text{matriz diagonal}$. En símbolos, $A \cong U D$.

4) A se dice *ortogonalmente diagonalizable*; en símbolos, $A \cong O D$, si $\exists O \in \text{ortogonal}$ tal que $O^T A O = D = \text{matriz diagonal}$.

Volvamos a las normas matriciales: recordemos que el teorema 11, iii) dice que la norma inducida por $\|\cdot\|_2$, $\|\cdot\|_2$, verifica

$\|A\|_2^2 = \rho(A^* A)$ y $A^* A$ es hermitiana.

Vale el siguiente *teorema espectral para matrices normales*:

Teorema 15. Las siguientes afirmaciones son equivalentes:

a) A es normal

b) $A \cong D$

c) $\sum_{i,j} |a_{ij}|^2 = \sum_{j=1}^n |\lambda_j|^2$, donde λ_j son las raíces de $p_A(t)$.

d) $\exists$ n autovectores ortonormales de A .

Para matrices hermitianas se puede decir más:

Teorema 16. Si A es hermitiana entonces además de a), b), c) y d), valen:

e) todos los λ_j son reales;

f) Si $A \in M_n(\mathbb{R})$ entonces $A \cong D$.

Ejercicio. Sea A hermitiana tal que $x^*Ax \geq 0 \quad \forall x \in \mathbb{C}^n$.

Entonces valen: 1) $\sigma(A) \subseteq [0, \infty)$; 2) $\text{tr } A = 0 \implies A = 0$.

9. PROBLEMAS.

Los problemas formulan, en su mayor parte, resultados teóricos que complementan la teoría que vamos presentando.

1) Ver que, si $\sum_j a_{ij} = 1$ para $i = 1, \dots, n$, $[a_{ij}] = A \in M_n(\mathbb{R})$, entonces $1 \in \sigma(A)$ y $e = [1, \dots, 1]^T$ verifica $Ae = e$. Si $\exists A^{-1}$ entonces $\sum_i a_{ij} = 1$ para $j = 1, \dots, n$. Dado un polinomio $f(t)$,

las sumas de los elementos de una fila de $f(A)$ son iguales.

2) $A \in M_n(\mathbb{R})$. Si $\lambda \in \sigma(A) \cap \mathbb{R}$, $0 \neq x \in \mathbb{C}^n$ tal que $Ax = \lambda x$, $x = u + iv$, $u, v \in \mathbb{R}^n$, entonces $Au = \lambda u$, $Av = \lambda v$. ¿Es u o v un autovector de A ? ¿Son u y v autovectores de A ?

3) Si $A = \begin{vmatrix} A_{11} & 0 \\ 0 & A_{22} \end{vmatrix}$ entonces $\sigma(A) = \sigma(A_{11}) \cup \sigma(A_{22})$.

4) La k -ésima función elemental en los n variables $x_1, \dots, x_n$, se define como $S_k(x_1, \dots, x_n) = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \prod_{j=1}^k x_{i_j}$.

Por ejemplo $S_1(x_1, \dots, x_n) = \sum_j x_j$, $S_n(x_1, \dots, x_n) = x_1 \dots x_n$.

Sean $\lambda_1, \dots, \lambda_n$, los autovalores de $A \in M_n(\mathbb{C})$ y $p_A(t)$ el polinomio característico de A . Vale: $p_A(t) = (t - \lambda_1) \dots (t - \lambda_n) = t^n - S_1(\lambda_1, \dots, \lambda_n) t^{n-1} + S_2(\lambda_1, \dots, \lambda_n) t^{n-2} - \dots \pm S_n(\lambda_1, \dots, \lambda_n)$

5) Hay $\begin{bmatrix} n \\ k \end{bmatrix}$ menores principales de orden k de A , (esto es, determinantes de submatrices principales de orden k). La suma de

ellos se denota con $E_k(A)$. En particular, $E_1 = \sum_j a_{jj}$, $E_n = \det A$.

Vale: $p_A(t) = t^n - E_1(A) t^{n-1} + E_2(A) t^{n-2} - \dots \pm E_n(A)$.

Luego, $S_k(\lambda_1, \dots, \lambda_n) = E_k(A)$.

6) Sea $p(t)$ un polinomio $p(t) = a_n t^n + \dots + a_0$ con ceros $\lambda_1, \lambda_2, \dots, \lambda_n$. El momento k -ésimo de $\{\lambda_1, \dots, \lambda_n\}$ es

$\mu_k = \lambda_1^k + \dots + \lambda_n^k$. Si R es bastante grande y $|t| > R$ entonces

$(t - \lambda_i)^{-1} = t^{-1} + \lambda_i t^{-2} + \lambda_i^2 t^{-3} + \dots$. Por lo tanto

$f(t) := \sum_{i=1}^n (t - \lambda_i)^{-1} = n \cdot t^{-1} + \mu_1 \cdot t^{-2} + \mu_2 \cdot t^{-3} + \dots$, $|t| > R$

Vale: $p'(t) = p(t) \cdot f(t)$. Luego, de esta igualdad sigue que:

$n \cdot a_n \cdot t^{n-1} + (n-1) \cdot a_{n-1} \cdot t^{n-2} + \dots + a_1 =$

$= (a_n t^n + \dots + a_0) \cdot (n \cdot t^{-1} + \mu_1 \cdot t^{-2} + \mu_2 \cdot t^{-3} + \dots)$.

Luego valen las identidades de Newton:

$k a_{n-k} + \mu_1 a_{n-k+1} + \mu_2 a_{n-k+2} + \dots + \mu_k a_n = 0$ si $1 \leq k \leq n$

y también valen:

$\mu_k a_0 + \mu_{k-1} a_1 + \dots + \mu_{n+k-1} a_{n-1} + \mu_{n+k} a_n = 0$ si $k = 1, 2, \dots$

7) Asumiendo que $\text{tr } A^k = \mu_k(A)$, probar que son equivalentes las afirmaciones: i) $\text{tr } A^k = \text{tr } B^k$ para todo $k \leq n$

ii) A y B tienen los mismos autovalores.

(Sugerencia: Como los polinomios característicos son mónicos, de las identidades de Newton sigue que los momentos determinan los coeficientes y recíprocamente. Luego, $p_A(t) = p_B(t)$ sii $\mu_k(A) = \mu_k(B)$, $k = 1, 2, \dots$).

8) Si $A \in M_n(\mathbb{C})$ es unitariamente equivalente a una matriz triangular superior T , los elementos t_{ij} no están unívocamente

determinados pero $\sum_{i < j} |t_{ij}|^2$ sí lo está.

9) Sea $j \geq n^2$. Demostrar que A^j puede escribirse como un polinomio en potencias $< n^2$ de A .

10. MATRICES POSITIVAS. 1ª parte del Teorema de Perron.

Lema 2. 1) $A \in M_n(\mathbb{R})$, $A \geq 0$. Si $\sum_j a_{ij} = c$, independiente de i , entonces $\rho(A) = \|A\|_\infty$, (esta es la norma inducida por $\|\cdot\|_\infty$).
 2) Si la suma por columnas es independiente de j , $\sum_i a_{ij} = c$, entonces $\rho(A) = \|A\|_1$.

Demostración. 1) Vale, $A \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = c \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$, luego c es autovalor,

$c = \|A\|_\infty$. Entonces $\rho(A) \geq c = \|A\|_\infty$. Como $\|\cdot\|_\infty$ es matricial,

1) sigue del teorema 12.

2) sigue por dualidad ya que $\rho(A) = \rho(A^T)$ y $\|A^T\|_\infty = \|A\|_1$.

Teorema 17. Sea $A \in M_n(\mathbb{R})$, $A \geq 0$. Entonces:

$$i) \alpha := \min_i \sum_j a_{ij} \leq \rho(A) \leq \max_i \sum_j a_{ij}$$

$$ii) \min_j \sum_i a_{ij} \leq \rho(A) \leq \max_j \sum_i a_{ij}$$

Demostración. Las cotas superiores resultan aplicando el T. 16.

La cota inferior en i) es trivial si $\alpha = 0$. Sea entonces $\alpha > 0$.

Definamos $b_{ij} := \alpha a_{ij} / \sum_k a_{ik}$. Entonces $b_{ij} \leq a_{ij}$, o sea,

$0 \leq B \leq A$, B verifica hipótesis del lema anterior con $c = \alpha$.

Luego $\alpha = \rho(B) \leq \rho(A)$ y se probó i). La cota inferior en

ii) sigue de i) usando nuevamente que $\rho(A) = \rho(A^T)$.

Corolario. 1) $A \geq 0$, $\sum_j a_{ij} > 0 \forall i \implies \rho(A) > 0$.

2) Si $A > 0$ entonces $\rho(A) > 0$.

3) Si A es irreducible, $A \geq 0$, entonces $\rho(A) > 0$ (pues si A es irreducible entonces no puede tener una fila de ceros).

Teorema 18. Sea $A \geq 0$, $x > 0$. Entonces:

$$i) \min_i \left(\sum_j a_{ij} x_j \right) / x_i \leq \rho(A) \leq \max_i \left(\sum_j a_{ij} x_j \right) / x_i$$

$$ii) \min_j \left(\sum_i a_{ij} / x_i \right) \cdot x_j \leq \rho(A) \leq \max_j \left(\sum_i a_{ij} / x_i \right) \cdot x_j$$

Demostración. Sea S la matriz diagonal $s_{ii} = x_i$. Entonces S^{-1} es la matriz diagonal $(S^{-1})_{ii} = 1/x_i$. Tenemos

$(S^{-1}AS)_{ij} = a_{ij} \cdot x_j / x_i$. Como $\rho(S^{-1}AS) = \rho(A)$, obtenemos el teorema aplicando el teorema precedente, QED.

Corolario. Sean $A \geq 0$, $x > 0$, $\alpha, \beta \geq 0$ tales que $\alpha x \leq Ax \leq \beta x$. Entonces $\alpha \leq \rho(A) \leq \beta$. Además, si $\alpha x < Ax$ entonces $\alpha < \rho(A)$ y si $Ax < \beta x$ entonces $\rho(A) < \beta$.

Demostración. Las hipótesis del corolario implican que $\alpha \leq \min_i (\sum_j a_{ij} x_j) / x_i$; $\max_i (\sum_j a_{ij} x_j) / x_i \leq \beta$ y se aplica el teorema 18. Si las desigualdades fueran estrictas entonces también se verificarían con $\alpha + \epsilon$ y $\beta - \epsilon$ y tendríamos $\alpha < \alpha + \epsilon \leq \rho(A) \leq \beta - \epsilon < \beta$, QED.

Definición 16. Dada una matriz $A \in M_n(\mathbb{C})$, diremos que λ_1 es un autovalor de *multiplicidad algebraica* L , si

$$p_A(t) = (t - \lambda_1)^L \cdot (t - \lambda_2) \dots \text{ con } \lambda_1 \neq \lambda_2, \dots$$

Diremos que λ_1 es de *multiplicidad geométrica* K , si

$K =$ la dimensión del subespacio $\{ x; Ax = \lambda_1 x \}$.

En general no coinciden y vale: $\text{mult geom}(\lambda_1) \leq \text{mult alg}(\lambda_1)$.

En efecto, sea $\Lambda = \{x; Ax = \lambda_1 x\}$. Sean $\alpha_1, \dots, \alpha_K$, vectores linealmente independientes que generan Λ . Sean $\beta_{K+1}, \dots, \beta_n$, elementos de $\mathbb{C}^n$ tales que $\{\alpha_1, \dots, \alpha_K, \beta_{K+1}, \dots, \beta_n\}$ es una base en $\mathbb{C}^n$. Si definimos $S = [\alpha_1 \dots \alpha_K \beta_{K+1} \dots \beta_n]$ (= la matriz cuyas columnas son los vectores α_i, β_j), resulta S no singular y

$AS = [\lambda_1 \alpha_1 \dots \lambda_1 \alpha_K \quad A\beta_{K+1} \dots A\beta_n]$. Como $S^{-1}S = I$, tenemos

$$B := S^{-1}AS = \begin{bmatrix} \lambda_1 I & \vdots & * \\ \dots & \ddots & \dots \\ 0 & \vdots & C \end{bmatrix} \quad \text{donde la matriz identidad } I \in M_K.$$

Entonces $p_A(t) = p_B(t) = (t - \lambda_1)^K \cdot \det [tI - C]$. Luego $K \leq L$.

Teorema 19. Sea $A > 0$. Entonces $\rho(A) > 0$, $\rho(A) \in \sigma(A)$ y $\exists z > 0$ tal que $Az = \rho(A) \cdot z$.

Demostración. En el corolario al teorema 17 se vió que $\rho(A) > 0$. Sea $\lambda \in \sigma(A)$ tal que $|\lambda| = \rho(A)$. Entonces $\exists x \neq 0$ tal que $Ax = \lambda x$,

y vale: $\rho(A) \cdot |x| = |\lambda| \cdot |x| = |\lambda x| = |Ax| \leq A|x|$.

Definamos $0 \leq y := A|x| - \rho(A) \cdot |x|$. Si probamos que $y = 0$ el teorema quedará probado pues $z = |x|$ satisface la tesis.

Supongamos que $0 < y$. Entonces $0 < Ay = A(A|x|) - \rho(A) \cdot A|x|$, $0 < A|x| =: z$, verifica $\rho(A) \cdot z < Az$. Luego, por el corolario al teorema 18, $\rho(A) < \rho(A)$, absurdo, QED.

Teorema 20. Sea $A > 0$, $\lambda \in \sigma(A)$, $\lambda \neq \rho(A)$. Entonces $|\lambda| < \rho(A)$.

Demostración. Sea $\lambda \in \sigma(A)$, $\lambda \neq \rho(A)$, $|\lambda| = \rho(A)$, con autovector x . En la demostración del teorema 19 se probó que $|x|$ es autovector correspondiente a $\rho(A)$. Entonces vale:

$|\lambda x_j| = \left| \sum_i a_{ji} x_i \right| = \rho(A) |x_j| = \sum_i a_{ji} |x_i|$. Pero esto sólo ocurre si todos los x_i están sobre un mismo rayo, esto es

$\exists \theta$ tal que $x_i = |x_i| \cdot e^{i\theta}$. Esto es lo mismo que $x = e^{i\theta} |x|$.

Entonces $|x|$ es también autovector para λ , lo cual implica que λ es real positivo, absurdo, QED.

Teorema 21. $\rho := \rho(A)$ tiene multiplicidad geométrica igual a uno.

Demostración. Supongamos que $w, z \in$ autoespacio de ρ .

Por la demostración anterior sabemos que existen θ_1, θ_2 tales que: $e^{-i\theta_1} z =: p > 0$, $e^{-i\theta_2} w =: q > 0$. Sea $0 < \beta := \min_j q_j/p_j$ y sea $r \in \mathbb{R}^n$ definido por $r_j := q_j - \beta \cdot p_j$. Entonces $r \geq 0$ verifica $Ar = \rho r$, de donde se deduce que, o bien $r = 0$, o bien $r > 0$. Como, por la definición de β , r tiene una componente nula, debe ser $r = 0$. Entonces $q = \beta p$, o sea $w = e^{i(\theta_2 - \theta_1)} z$, QED.

Hemos entonces demostrado el

Teorema de Perron (1ª parte). Sea $A \in M_n(\mathbb{C})$, $A > 0$. Vale:

- a) $\rho(A) > 0$, $\rho(A) \in \sigma(A)$
- b) $\exists x > 0$: $Ax = \rho(A)x$
- c) si $\lambda \in \sigma(A) \setminus \{\rho(A)\}$ entonces $|\lambda| < \rho(A)$
- d) $\rho(A)$ es un autovalor de multiplicidad geométrica uno.

Corolario. Sea $A > 0$. $\exists! x_0 > 0$ tal que $Ax_0 = \rho(A)x_0$ y $\sum_j x_{0,j} = 1$.

A este x_0 se la denomina *autovector de Perron*.

Dado $A > 0$, A^T es también positiva, teniendo un autovector de

Perron y_0 , $A^T y_0 = \rho(A^T) y_0 = \rho(A) y_0$. Trasponiendo resulta $y_0^T A = \rho(A) \cdot y_0^T$. y_0 es el vector de Perron a izquierda de A , $\rho(A)$ se denomina la raíz de Perron.

Vamos a dar un criterio para encontrar la raíz de Perron de A .

Teorema 22. Sea A positiva. Si x es un autovector correspondiente a λ , $x \geq 0$, entonces $\lambda = \rho(A)$ y $x > 0$. O sea, λ es la raíz de Perron y $x/\|x\|_1$ es el vector de Perron.

Demostración. Las hipótesis implican que $Ax > 0$, de donde sigue que $\lambda > 0$, $x > 0$. Además se verifica que $\lambda x \leq Ax \leq \lambda x$. Aplicando teorema 18, i) obtenemos $\lambda \leq \rho(A) \leq \lambda$, QED.

11. PROBLEMAS.

1) Sabemos que si $\|\cdot\|$ es una norma matricial y

$\lambda \in \sigma(A)$, $A \in M_n(\mathbb{C})$, entonces $|\lambda| \leq \|A\|$. Sea $f(z)$ un polinomio no constante, entonces existe un polinomio mónico $p(z)$ tal que $f(z) = K \cdot z^k \cdot p(z)$, $K \neq 0$, $p(z) = z^n + a_{n-1}z^{n-1} + \dots + a_0$, $a_0 \neq 0$. Para acotar las raíces de $f(z)$ basta acotar las de $p(z)$. La matriz acompañante de p se define como:

$$(1) \quad C(p) = \begin{vmatrix} -a_{n-1} & -a_{n-2} & \dots & -a_1 & -a_0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{vmatrix}$$

Ella verifica: $p_{C(p)}(z) = p(z)$ (Calcular $\det(zI - C(p))$ por cofactores de la 1ª columna). Luego, $p(z') = 0 \Rightarrow |z'| \leq \|C(p)\|$ cualquiera sea la norma matricial $\|\cdot\|$ sobre $M_n(\mathbb{C})$. En particular si $\|\cdot\| = \|\cdot\|_1$, vale (cota de Cauchy)

$$(2) \quad |z'| \leq \max \{ |a_0|, 1+|a_1|, 1+|a_2|, \dots, 1+|a_{n-1}| \}.$$

Si $\|\cdot\| = \|\cdot\|_\infty$, vale (cota de Montel)

$$(3) \quad |z'| \leq \max \{ 1, |a_0| + |a_1| + \dots + |a_{n-1}| \}.$$

2) Sean

$$S = \begin{vmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{vmatrix}, \quad R = \begin{vmatrix} -a_{n-1} & -a_{n-2} & \dots & -a_0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{vmatrix}$$

Entonces $S^*R = R^*S = 0$ y $\|S^*S\|_2 = 1$, $\|R^*R\|_2 = |a_0|^2 + \dots + |a_{n-1}|^2$.

(Para esto último mostrar primero que $\|A\|_2^2 = \|A^*A\|_2 = \|AA^*\|_2$ usando que $\|B\|_2 = \max \{ |y^*Bx| ; \|y\|_2 = \|x\|_2 = 1 \}$).

Entonces $\|C(p)\|_2^2 = \|C(p)^*C(p)\|_2 = \|(S+R)^*(S+R)\|_2 = \|S^*S + R^*R\|_2 \leq \|S^*S\|_2 + \|R^*R\|_2$. Luego vale (cota de Carmichael and Mason)

$$(4) \quad |z'| \leq (1 + |a_0|^2 + \dots + |a_{n-1}|^2)^{\frac{1}{2}}$$

3) Sea $q(z) = (z-1).p(z) = z^{n+1} + (a_{n-1} - 1).z^n + (a_{n-2} - a_{n-1}).z^{n-1} + \dots + (a_0 - a_1).z - a_0$. Luego $|z'| \leq \max \{1, |a_0| + |a_0 - a_1| + \dots + |a_{n-1} - 1|\}$. Como $1 \leq |a_0| + |a_1 - a_0| + \dots + |a_{n-2} - a_{n-1}| + |a_{n-1} - 1|$,

$$(5) \quad |z'| \leq |a_0| + |a_1 - a_0| + \dots + |a_{n-2} - a_{n-1}| + |a_{n-1} - 1|.$$

4) **Teorema de Kakeya.** Sea $f(z) = a_n z^n + \dots + a_1 z + a_0$,

$a_n \geq a_{n-1} \geq \dots \geq a_0 \geq 0$; $f(z') = 0 \implies |z'| \leq 1$.

(Sugerencia: usar (5)).

NB. Valen cotas por abajo, v. g.:

$|z'| \geq |a_0| / (1 + |a_0|^2 + \dots + |a_{n-1}|^2)^{\frac{1}{2}}$, (Carmichael and Mason).

12. AUTOVALORES Y AUTOVECTORES.

Supongamos $A^T = A$, $A \in M_n(\mathbb{R})$. Consideremos $x^T A x = \sum_{i,j} a_{ij} x_i x_j$ pensado como función numérica sobre $\{x \in \mathbb{R}^n; x^T x = 1\}$. Como este conjunto es cerrado y acotado, la función continua $x^T A x$ alcanza su máximo ($:= \lambda$) allí, o sea $\exists x_0$ tal que $x_0^T x_0 = 1$, $x_0^T A x_0 = \lambda$, mientras $x^T A x \leq \lambda$ si $x^T x = 1$. Entonces $f(x) := x^T A x - \lambda x^T x \leq 0$ para todo $x \in \mathbb{R}^n$ y $f(x_0) = 0$. Luego, las derivadas parciales de $f(x)$ respecto de x_j se anulan en el punto $x_0 \forall j$:

$$\frac{\partial f}{\partial x_j}(x_0) = 0 = \frac{\partial}{\partial x_j} \left(\sum (a_{ik} x_i x_k - \lambda \sum x_k^2) \right) \Big|_{x=x_0} = 2(\sum a_{jk} x_{0,k} - \lambda x_{0,j})$$

Vale entonces $Ax_0 = \lambda x_0$, o sea λ es autovalor y x_0 autovector de A . (El mayor autovalor !)

Ejercicios. 1) Si $q(t)$ es un polinomio, entonces:

$Ax = \lambda x, x \neq 0 \implies q(A)x = q(\lambda)x$. Luego $\sigma(q(A)) \supseteq \{q(\lambda); \lambda \in \sigma(A)\}$

2) Sea $A \in M_n(\mathbb{C})$. A es singular si y sólo si $0 \in \sigma(A)$.

(Sugerencia: $\det A = 0$ sii $\exists x \neq 0$ tal que $Ax = 0$).

3) Si $\exists A^{-1}$ (es decir A es no singular) y $\lambda \in \sigma(A)$ entonces $\lambda \neq 0$ y $\lambda^{-1} \in \sigma(A^{-1})$.

4) A se dice *idempotente* si $A^2 = A$. Ver que en este caso $\lambda \in \sigma(A)$ implica $\lambda = 0$ o 1 .

5) A se dice *nilpotente* si $A^q = 0$ para algún q . Ver que en este caso $\sigma(A) = \{0\}$.

6) Si A es hermitiana entonces $\sigma(A) \subset \mathbb{R}$, (sugerencia: conjugar la identidad $x^*Ax = \lambda x^*x$, calculando su adjunta).

7) Si $p_A(t) := \det(tI - A)$, ver que $\det(A - tI) = (-1)^n p_A(t)$.

8) Si $A \in M_n(\mathbb{R})$ entonces $p_A(t)$ tiene coeficientes reales.

Ver que si λ es un autovalor no real, $\bar{\lambda} \in \sigma(A)$.

9) Si $T \in M_n(\mathbb{C})$ es triangular superior, $t_{ij} = 0$ para $i > j$,

entonces $\sigma(T) = \{t_{ii}; i=1, \dots, n\}$, $\det T = \prod_i t_{ii}$.

10) $p_A(t)$ (cf. 7) $\implies p_A(t) = t^n - (\text{tr } A)t^{n-1} + \dots + (-1)^n \det A$.

Como $p_A(t) = (t - \lambda_1) \dots (t - \lambda_n)$, ver que $\text{tr } A = \sum_i \lambda_i$, $\det A = \prod_i \lambda_i$.

11) Ver que un autovalor λ es de multiplicidad algebraica k , si

$p_A(\lambda) = p'_A(\lambda) = \dots = p_A^{(k-1)}(\lambda) = 0$, $p_A^{(k)}(\lambda) \neq 0$.

12) Si $A \in M_n(\mathbb{R})$, n impar, entonces A tiene un autovalor real.

13) Si $A \in M_n(\mathbb{C})$ entonces $\text{tr } A^m = \sum_j \lambda_j^m$, (sug.: usar T. de Schur).

14) Si A es diagonalizable entonces $q(A)$ también lo es.

Teorema 23. $A \approx D$, D diagonal, equivale a que existe un conjunto de n autovectores de A , linealmente independientes.

Demostración. ($\Rightarrow$). Sea S no singular tal que $S^{-1}AS = D$.

Entonces $AS = SD$. Si escribimos $S = [x_1 \dots x_n]$, donde cada x_j es un vector columna, tendremos $Ax_j = d_{jj} \cdot x_j$. Luego cada x_j

es un autovector, siendo la familia $\{x_1, \dots, x_n\}$ linealmente independiente por ser S no singular.

($\Leftarrow$). Sea $\{x_1, \dots, x_n\}$ una familia linealmente independiente de autovectores correspondientes a los autovalores $\lambda_1, \dots, \lambda_n$. Sea S la matriz $S := [x_1 \dots x_n]$. S es no singular (¿porqué?) y vale $AS = [\lambda_1 x_1 \dots \lambda_n x_n] = SD$, donde D es la matriz diagonal $d_{jj} = \lambda_j$. Luego $S^{-1}AS = D$, QED.

Enunciaremos sin demostrar dos interesantes teoremas.

Teorema 24. Sean A y B diagonalizables. Las proposiciones

i) $AB = BA$

ii) $\exists S$ no singular tal que $S^{-1}AS = D_A$, $S^{-1}BS = D_B$

son equivalentes. (Este teorema vale también para familias conmutantes !)

Teorema 25. Sea $C = \begin{vmatrix} A & 0 \\ 0 & B \end{vmatrix}$, con $A \in M_n$, $B \in M_m$, (suma directa de A y B). C es diagonalizable si y sólo si A y B lo son.

Teorema de Schur. Sea $A \in M_n(\mathbb{C})$. Existe una matriz unitaria U tal que $U^*AU = T$ es una matriz triangular superior tal que $\sigma(A) = \{t_{jj}; j=1, \dots, n\}$. Si $A \in M_n(\mathbb{R})$ y $\sigma(A) \subset \mathbb{R}$, se puede elegir U real.

Demostración. Por inducción sobre n . Para $n=1$ es cierto. Supongamos que vale para $n-1$. Sea $\lambda_1 \in \sigma(A)$ y x_1 un autovector

de norma 1 (esto es, $\sum |x_i|^2 = 1$). Completamos con vectores $z_2, \dots, z_n$ hasta formar una base ortonormal $\{x_1, z_2, \dots, z_n\}$ (método de Gram-Schmidt). Sea $U_1 = [x_1 \ z_2 \ \dots \ z_n]$. U_1 es unitaria y vale

$$U_1^* A U_1 = \begin{vmatrix} x_1^* \\ \vdots \\ z_n^* \end{vmatrix} [\lambda_1 x_1 \ Az_2 \ \dots \ Az_n] = \begin{vmatrix} \lambda_1 & * & * & \dots & * \\ 0 & & & & \\ \vdots & & B & & \\ 0 & & & & \end{vmatrix}.$$

$B \in M_{n-1}(\mathbb{C})$ y vale: $p_A(t) = (t-\lambda_1) \cdot p_B(t)$. Entonces

$\sigma(A) = \sigma(B) \cup \{\lambda_1\}$. Por la hipótesis inductiva $\exists V \in M_{n-1}$, V

unitaria, tal que $V^*BV = T_1 =$ triangular superior, con elementos

de $\sigma(B)$ en la diagonal.
Definamos $V_1 :=$

$$\begin{pmatrix} \frac{1}{0} & - & 0 & . & . & . & 0 \\ \vdots & & & & & & \\ \vdots & & & & v & & \\ 0 & \vdots & & & & & \end{pmatrix}, U := U_1 V_1$$

Si $A \in M_n(\mathbb{R})$ y $\lambda_1 \in \mathbb{R}$ entonces existe un autovector real x_1 y en la demostración anterior se pueden tomar $z_2, \dots, z_n$ reales. Entonces U_1 y B son matrices reales y por la hipótesis inductiva V también lo es, y por lo tanto $U \in M_n(\mathbb{R})$, QED.

2) Demostrar que $\sum_{i < j} t_{ij}^2$ está determinada por A, aunque la matriz

3) Probar que el producto de matrices triangulares superiores del mismo orden es triangular superior.

Teorema 26. Sea $\mathcal{F} \subseteq M_n(\mathbb{C})$ una familia conmutante de matrices, (las matrices de $\mathcal{F}$ conmutan dos a dos). En este caso,

Estamos ahora en condiciones de demostrar el

Probaremos primero un lema.

$$R = \begin{bmatrix} 0 & \vdots & - * \\ 0 & \vdots & T_1 \end{bmatrix} \quad \text{donde } T_1 \text{ es triangular superior } \in M_{n-k}(\mathbb{C}),$$

y T es triangular superior, verificándose $t_{(k+1)(k+1)} = 0$. Entonces $R' = RT$ tiene la misma forma que R , sólo que la matriz 0 en el ángulo superior izquierdo es de orden $(k+1)$.

Demostración. Obsérvese que las primeras k columnas de R son nulas. Luego cada fila de R comienza con k elementos nulos, mientras que las primeras $(k+1)$ columnas de T terminan con $(n-k)$ ceros. Luego RT tiene sus primeras $(k+1)$ columnas nulas, además de ser triangular, QED.

Demostración del T. de Cayley-Hamilton. Sabemos por el teorema de Schur que existe U unitaria tal que $U^*AU = T = \text{triangular}$. Entonces $p_A(A) = p_A(U^*TU) = U^*p_A(T)U$ y bastará probar que $p_A(T) = 0$. Pero $p_A(t) = (t-\lambda_1)\dots(t-\lambda_n)$ implica $p_A(T) = (T-\lambda_1 I)\dots(T-\lambda_n I) = T_1\dots T_n$. Cada T_j es triangular y su elemento $(T_j)_{jj}$ es nulo. Probaremos por inducción que $R_j := T_1\dots T_j$ tiene sus primeras j columnas nulas. En efecto, esto es cierto para $j=1$, y suponiéndolo para j , resulta cierto para $j+1$ como consecuencia del lema. Luego $p_A(T) = R_n = 0$, QED.

Corolario 1 (al T. de Schur). Si $q(t)$ es un polinomio entonces $\sigma(q(A)) = \{ q(\lambda) ; \lambda \in \sigma(A) \} =: q(\sigma(A))$.

Demostración. Si $A \cong T$ entonces $q(A) \cong q(T)$. Pero $q(T)$ tiene en su diagonal principal a $q(\lambda_1), \dots, q(\lambda_n)$. Luego estos son los autovalores de $q(T)$, que por otro lado son los de $q(A)$, QED.

Corolario 2. $\text{tr } A^k = \sum \lambda_j^k$.

Demostración. Sigue del Ej 10, pg. 26 y del corolario anterior con $q(t) = t^k$, QED.

Ejercicio. Si A es equivalente a B entonces $\text{tr } A^k = \text{tr } B^k$.

Observemos que, como $\dim M_n(\mathbb{C})$ es n^2 , las n^2+1 matrices $A^0, A^1, \dots, A^{n^2}$, son linealmente dependientes. Luego, A^{n^2} es combinación lineal de las potencias A^j , $j=0, \dots, n^2-1$ (¿porqué?). Pero el teorema de Cayley-Hamilton nos dice que A^n se puede escribir como combinación lineal de las potencias A^j , $j=1, \dots, (n-1)$. Es más económico!. En otras palabras $A^n = q(A)$, donde $q(t)$ es el

polinomio de grado $< n$, $q(t) := t^n - p_A(t)$.

Corolario 3. Si $k \geq n$ entonces A^k se puede escribir como un polinomio en A de grado $< n$.

Demostración. $A^k = A^{k-n} A^n = A^{k-n} q(A) =$ polinomio de grado $< k$.

Reiterando obtenemos un polinomio de grado menor que n .

Corolario 4. Si $A \in M_n(\mathbb{C})$ y $\det A \neq 0$ entonces existe un polinomio de grado menor que n , $g(t)$, tal que $A^{-1} = g(A)$.

Demostración. Recordemos que $p_A(t) = t^n + a_1 t^{n-1} + \dots + a_n$ donde $a_n = (-1)^n \det A$. Si $\det A \neq 0$ entonces por un lado $\exists A^{-1}$ y por otro lado a_n es no nulo. Multiplicando $0 = p_A(A)$ por A^{-1} obtenemos: $0 = A^{n-1} + a_1 A^{n-2} + \dots + a_{n-1} I + a_n A^{-1}$, es decir, la tesis con $g(t) = (a_n - p_A(t))/(a_n t)$, QED.

Vimos en teorema 23 que si A tiene n autovectores linealmente independientes entonces es diagonalizable mediante una transformación no singular. El teorema de Schur dice algo más y algo menos: la transformación es unitaria pero el resultado que se asegura es la triangularidad. Caso particular del T.23 es el **Teorema 27.** Si $A \in M_n(\mathbb{C})$ tiene n autovalores distintos dos a dos entonces es diagonalizable.

Demostración. Sea $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$, y sean $x^{(1)}, \dots, x^{(n)}$ los correspondientes autovectores. $\{x^{(1)}, \dots, x^{(n)}\}$ es un conjunto linealmente independiente. En efecto, si no fuera así habría un $r > 1$ tal que $\{x^{(1)}, \dots, x^{(r)}\}$ sería linealmente dependiente pero $\{x^{(1)}, \dots, x^{(r-1)}\}$ linealmente independiente. Sean $(\alpha_1, \dots, \alpha_r) \neq (0, \dots, 0)$ tal que $\alpha_1 x^{(1)} + \dots + \alpha_r x^{(r)} = 0$. Luego, aplicando A obtenemos $\alpha_1 \lambda_1 x^{(1)} + \dots + \alpha_r \lambda_r x^{(r)} = 0$. Si multiplicamos la primera igualdad por λ_r y la restamos de la última, resulta $\alpha_1 (\lambda_1 - \lambda_r) x^{(1)} + \dots + \alpha_{r-1} (\lambda_{r-1} - \lambda_r) x^{(r-1)} = 0$. Esto lleva a una contradicción, QED.

Definición 17. 1) El *espacio nulo* de una matriz B es el conjunto de vectores x tales que $Bx = 0$. (Si B es no singular, su espacio nulo constará sólo del vector 0).

2) Si $\lambda \in \sigma(A)$, llamaremos *autoespacio de A* correspondiente al autovector λ al espacio nulo de $(A - \lambda I)$. Lo denotaremos con N_λ . O sea, $N_\lambda := \{x; Ax = \lambda x\} = \{\text{autovectores de A corresp. a } \lambda\} \cup \{0\}$.

3) Se dice que un subespacio $W \subseteq C^n$ es *invariante por A*, si $\forall x \in W, Ax \in W$. O más brevemente, $AW \subseteq W$.

Ejercicios. 1) N_λ es invariante por A.

2) Un subespacio invariante por A *minimal* es el generado por un único autovector de A.

Definición 18. De $A \in M_n(C)$ diremos que es *defectiva* si existe $\lambda \in \sigma(A)$ tal que su multiplicidad geométrica es menor que su multiplicidad algebraica.

Ejercicio. A es diagonalizable si y sólo si no es defectiva.

Definición. Se dice que y es un *autovector a izquierda* de A (correspondiente al autovalor λ), si $y^* A = \lambda y^*$.

Ejercicio. Son equivalentes las proposiciones:

- i) y es autovector a izquierda de A (corresp. a λ)
- ii) y es autovector de A^* correspondiente al autovalor $\bar{\lambda}$
- iii) $\bar{y}$ es autovector de A^T correspondiente a λ .

Recordemos la definición de vectores ortogonales en C^n :

Definición 19. $x, y \in C^n$ son *ortogonales* si $y^* x = \langle x, y \rangle = 0$.

Teorema 28. Sean $\lambda, \mu \in C, \lambda \neq \mu$, y autovector a izquierda de A correspondiente a λ , x autovector de A correspondiente a μ .

Entonces x es ortogonal a y .

Demostración. $y^* A = \lambda y^*, Ax = \mu x$ implican $y^* Ax = \lambda \langle x, y \rangle = \mu \langle x, y \rangle$.

Esto sólo puede valer si $\langle x, y \rangle = 0$, QED.

Ejercicio. ¿Porqué ni 1) ni 2) prueban el T. de Cayley-Hamilton?

1) $p_A(t) = \det |tI - A|$. Luego, $p_A(A) = \det |AI - A| = \det 0 = 0$.

2) $p_A(\lambda) = 0 \forall \lambda \in \sigma(A)$. Por cor. 1, pg.29, $p_A(A)$ tiene por único autovalor a 0. Luego $p_A(A) = 0$.

13. CONSECUENCIAS DEL TEOREMA DE SCHUR.

El radio espectral no es una norma, pero...

Teorema 29. Sea $A \in M_n(C)$. Dado $\varepsilon > 0$, $\exists$ una norma matricial $\| \cdot \|$ tal que $\rho(A) \leq \|A\| \leq \rho(A) + \varepsilon$.

Demostración. La primera desigualdad ya fue probada en T. 16. Por el teorema de Schur, $A = U^* \Delta U$ con Δ triangular, U unitaria. Definamos, para $t > 0$, la matriz D_t como la matriz cuyos elementos diagonales son: $t, t^2, \dots, t^n$ y los restantes $= 0$. Entonces $D_t^{-1} = D_{1/t}$ y AD_t es la matriz cuya columna j -ésima es la columna j -ésima de A multiplicada por t^j , para $j = 1, \dots, n$, mientras que $D_t A$ es la matriz cuya fila j -ésima es la fila j -ésima de A multiplicada por t^j . Por eso $D_t^{-1} \Delta D_t$ es la matriz triangular superior cuyo elemento ij es $\Delta_{ij} t^{j-i}$, siendo los elementos diagonales los autovalores de A . Como los elementos no diagonales tienden a 0 cuando $t \rightarrow +0$, podemos elegir $t > 0$ tal que $\|D_t^{-1} \Delta D_t\|_1 \leq \sup_j (|\lambda_j| + \varepsilon) = \rho(A) + \varepsilon$.

Definamos ahora la siguiente norma (**Ejercicio:** verificar que es una norma matricial):

$$\|B\| := \|D_t^{-1} U B U^* D_t\|_1 = \|S^{-1} B S\|_1 \quad \text{donde } S := U^* D_t. \text{ Entonces}$$

$$\|A\| := \|D_t^{-1} U A U^* D_t\|_1 = \|D_t^{-1} \Delta D_t\|_1 \leq \rho(A) + \varepsilon, \text{ QED.}$$

Definición 20. La matriz A se dice convergente si $A^k \rightarrow 0$, (elemento a elemento).

Teorema 30. A es convergente si y sólo si $\rho(A) < 1$.

Demostración. Sea $\rho(A) < 1$. Por el teorema 29 existe una norma matricial tal que $\|A\| < 1$, y vale $\|A^k\| \leq \|A\|^k \rightarrow 0$. Pero todas las normas son equivalentes, luego, $\|A^k\|_\infty \rightarrow 0$, o sea A es convergente.

Supongamos ahora que $\|A^k\|_\infty \rightarrow 0$. Si $\lambda \in \sigma(A)$ siendo x un autovector correspondiente, entonces $A^k x = \lambda^k x \rightarrow 0$, o sea, $\forall j \quad \lambda^k x_j \rightarrow 0$. Luego $|\lambda| < 1$ y $\rho(A) < 1$, QED.

Corolario 1. Sea $A \in M_n(\mathbb{C})$, $\varepsilon > 0$. $\exists C = C(A, \varepsilon)$ tal que,

$$|(A^k)_{ij}| \leq C.(\rho(A) + \varepsilon)^k \quad \forall i, j.$$

Demostración. Sea $\tilde{A} = \frac{A}{\rho(A) + \varepsilon}$. Entonces $\rho(\tilde{A}) = \frac{\rho(A)}{\rho(A) + \varepsilon} < 1$.

Luego $\tilde{A}$ es convergente y $(\tilde{A}^k)_{ij} \rightarrow 0$ para $k \rightarrow \infty$, $\forall i, j$. Como toda

sucesión convergente es acotada, resulta la existencia de una constante C que acota a $|(\tilde{A}^k)_{ij}| = |(A^k)_{ij}| (\rho(A) + \varepsilon)^{-k}$, QED.

Corolario 2. Sea $\| \cdot \|$ una norma matricial. Entonces

$$\rho(A) = \lim_{k \rightarrow \infty} (\|A^k\|)^{1/k}.$$

Demostración. Por T. de Schur y T. 16, $(\rho(A))^k = \rho(A^k) \leq \|A^k\|$.

Entonces $\rho(A) \leq (\|A^k\|)^{1/k}$. Por otra parte, usando el corolario anterior, $\|\tilde{A}^k\|_\infty \rightarrow 0$ para $k \rightarrow \infty$. Luego, $\|\tilde{A}^k\|_\infty \leq C$ y vale

por la equivalencia de normas, $\|\tilde{A}^k\| \leq \tilde{C}$. Usando la definición de $\tilde{A}$ resulta $(\|A^k\|)^{1/k} \leq (\rho(A) + \varepsilon) \cdot \tilde{C}^{1/k}$. Sigue el corolario, QED.

Veremos que el corolario 2 vale en condiciones más generales.

Definición 21. Diremos que una función f sobre C^n es una *prenorma* si f es a valores reales y verifica:

1. $f(x) \geq 0$
2. $x \neq 0 \Rightarrow f(x) > 0$ (positividad)
3. $f(\alpha x) = |\alpha| f(x)$ (homogeneidad, que implica $f(0) = 0$)
4. $f(x)$ es continua como función sobre $C^n = \mathbb{R}^{2n}$.

Teorema 31. Sean f_1 y f_2 prenormas. Existen constantes positivas m y M tales que $\forall x \quad m \cdot f_1(x) \leq f_2(x) \leq M \cdot f_1(x)$.

Demostración. Sea $\Sigma := \{z \in C^n; \|z\|_2 = 1\}$. Σ es compacto y por el teorema de Weierstrass la función $h(z) := f_2(z)/f_1(z)$, positiva en Σ , tiene un máximo M y un mínimo m en Σ , ambos necesariamente positivos. Esto es $0 < m \leq h(z) \leq M$ para $z \in \Sigma$. La homogeneidad de f_1 y f_2 implica que $h(\alpha z) = h(z)$ para α y z no nulos. En particular, si $\alpha = 1/|z|$ entonces el punto $\alpha z \in \Sigma$ y vale $m \leq h(z) \leq M$ $\forall z \neq 0$; y el corolario sigue, QED.

Corolario 2. Si f es una prenorma en $M_n(C) = C^{n^2}$ entonces

$$\rho(A) = \lim_{k \rightarrow \infty} (f(A^k))^{1/k}.$$

Demostración. Sea $f_1(A) := \|A\|$ una norma matricial. Por cor. 1 existen constantes positivas m, M , tales que

$m\|A^k\| \leq f(A^k) \leq M\|A^k\|$ y $(m\|A^k\|)^{1/k} \leq (f(A^k))^{1/k} \leq (M\|A^k\|)^{1/k}$ y aplicando el corolario 1, teorema 30, obtenemos la tesis, QED.

Ejercicios.

- 1) Sea $A \geq 0$, no nula. Si existe $y > 0$ tal que $Ay = \lambda y$ entonces $\rho(A) = \max_{x>0} \min_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j = \min_{x>0} \max_{1 \leq i \leq n} \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j$, (cf.T.18).
- 2) Sea $A \geq 0$, no nula. Si existe $y > 0$ tal que $Ay = \lambda y$ entonces $\lambda = \rho(A)$.
- 3) Sea $A \geq 0$, no nulo. Si existe $y > 0$ tal que $Ay = \lambda y$ entonces A es semejante a una matriz no negativa con las sumas por filas iguales a una misma constante.
- 4) Si $A > 0$ y x es el vector de Perron de A entonces:

$$\rho(A) = \sum_{i,j=1}^n a_{ij} x_i x_j.$$

14. SERIES DE MATRICES.

Definición 22. Sea $A \in M_n(\mathbb{C})$. A se dice *estrictamente diagonalmente dominante* (más brevemente, $A \in \text{EDD}$) si se cumple que

$$|a_{ii}| > R_i := \sum_{j=1, j \neq i}^n |a_{ij}|, \quad (' \text{ en la } \sum \text{ indica que se omite } j=i).$$

Teorema 32. (Criterio de invertibilidad de Levy-Desplaques).

Si $A \in \text{EDD}$ entonces A es invertible.

Demostración. Sea D matriz diagonal tal que $d_{ii} = a_{ii}$. Tenemos que $D^{-1}A$ es la matriz cuya fila j -ésima es la fila j -ésima de A dividida por a_{jj} . Si definimos $B := I - D^{-1}A$ resulta $b_{jj} = 0$, $b_{ij} = -a_{ij}/a_{ii}$ para $i \neq j$. Calculamos $\|B\|_\infty = \max_i \sum_j |b_{ij}| = \max_i R_i/|a_{ii}| < 1$. Veremos (corolario al teo. 34) que entonces $(I - B)^{-1}$ existe. Pero $A = D(I - B)$, luego A^{-1} existe, QED.

Completamos con algunos resultados sobre series de matrices.

Teorema 33. Sea $A^{(j)} \in M_n(\mathbb{C}) \forall j \in \mathbb{N}$ y $\|\cdot\|$ una norma vectorial.

Si la serie $\sum \|A^{(h)}\|$ es convergente entonces existe $A = \sum A^{(h)} = \lim_{N \rightarrow \infty} (A^{(1)} + \dots + A^{(N)})$ para $N \rightarrow \infty$, (elemento a elemento).

Demostración. Como todas las normas son equivalentes, resulta

$$\|A\|_\infty := \max_{1 \leq i, j \leq n} |a_{ij}| \leq M \|A\|, \quad \forall A \in M_n. \text{ Luego } \sum_h |a_{ij}^{(h)}| \leq M \sum_h \|A^{(h)}\| < \infty \quad \forall i, j=1, \dots, n. \text{ de donde sigue la existencia de}$$

$$\sum_h a_{ij}^{(h)} = a_{ij}, \text{ QED.}$$

Teorema 34. Sea $A \in M_n(\mathbb{C})$, $\|\cdot\|$ una norma matricial, $a_h \geq 0$, $h \in \mathbb{N}$.

Si $\sum a_h \|A\|^h < \infty$ entonces $\sum a_h A^h$ converge.

Demostración. Basta observar que $\sum \|a_h A^h\| \leq \sum a_h \|A\|^h < \infty$ y aplicar el teorema anterior, QED.

Corolario. Sea $A \in M_n(\mathbb{C})$, $\|\cdot\|$ una norma matricial y $\|A\| < 1$.

Entonces $I-A$ es invertible, siendo $(I-A)^{-1} = I + A + \dots + A^h + \dots$

Demostración. La serie $\sum \|A\|^h$ es convergente; por el teorema anterior, existe $B := \lim_{N \rightarrow \infty} (I + A + \dots + A^N)$. Pero entonces $AB = B - I$,

o sea, $(I - A)B = I$, de donde sigue que $B = (I-A)^{-1}$, QED.

Ejercicios. 1) $e^A := \sum_{h=0}^{\infty} A^h / h!$ está bien definida ($A^0 := I$).

2) Sea $A \in M_n(\mathbb{C})$, $\|\cdot\|$ una norma matricial tal que $\|I-A\| < 1$.

Entonces, A es invertible siendo $A^{-1} = \sum_{k=0}^{\infty} (I-A)^k$.

15. TEOREMA DE PERRON, continuación.

Teorema 35 (de Perron, 2ª parte). Sea $A \in M_n(\mathbb{R})$, $A > 0$. Vale:

1) $\lim_{k \rightarrow \infty} \left[\frac{A}{\rho(A)} \right]^k = L$, donde $L := x y^T$ con $x > 0$, autovector de A ,

y $y > 0$, autovector a izquierda de A (ambos para el autovalor $\rho(A)$), normalizados por la condición $x^T y = 1 = y^T x$.

2) $\rho(A)$ tiene multiplicidad algebraica igual a 1.

Demostración. Veamos 1). Recordemos el Teo. 22 que asegura que necesariamente un autovector positivo corresponde a $\rho(A)$. Observemos que $L_{ij} = x_i y_j$ implica que $Lx = xy^T x = x(y^T x) = x$. Luego, x es autovector de L para el autovalor 1. Además $L^2 = (xy^T)(xy^T) = x(y^T x)y^T = xy^T = L$, de donde $L^m = L$ para $m = 2, \dots$ y si $\lambda = \rho(A)$,

$AL = A x y^T = \lambda x y^T = \lambda L = x y^T A = LA$. Entonces $A^m L = LA^m =$

$= (\lambda)^m L$, de donde $(A - \lambda L)^2 = A^2 - 2\lambda^2 L + \lambda^2 L = A^2 - \lambda^2 L$ y por

inducción $(A - \lambda L)^m = A^m - \lambda^m L$. Sea ahora $\mu \neq 0$ autovalor de

$(A - \lambda L)$. Entonces existe $w \neq 0$ tal que $(A - \lambda L)w = \mu w$. Aplicando

L a ambos miembros resulta $0 = \mu Lw$. Entonces, $Lw = 0$, resultando

$Aw = \mu w$. Hemos probado entonces que $\sigma(A - \rho(A)L) \setminus \{0\} \subset \sigma(A)$. Veremos que vale además que: $\sigma(A - \rho(A)L) \setminus \{0\} \subset \sigma(A) \setminus \{\rho(A)\}$. En efecto, sea $\lambda = \rho(A)$. Si λ es autovalor de $(A - \lambda L)$ con autovector w , la demostración anterior nos dice que $\lambda Lw = 0$, luego, $\lambda w = (A - \lambda L)w = Aw$. O sea, w es autovector de A correspondiente a $\lambda = \rho(A)$ y por lo tanto es múltiplo del vector de Perron x : $w = \alpha x$. Pero entonces $Lw = xy^T(\alpha x) = \alpha x \neq 0$, absurdo. Luego, λ no es autovalor de $A - \lambda L$.

La inclusión $\sigma(A - \rho(A)L) \setminus \{0\} \subset \sigma(A) \setminus \{\rho(A)\}$ implica que el radio espectral de $(A - \rho(A)L)$ es menor que $\rho(A)$. O sea $\rho\left(\frac{A}{\rho(A)} - L\right) < 1$. Luego $\left(\frac{A}{\rho(A)} - L\right)^m \rightarrow 0$ por teorema 30.

Pero $\left(\frac{A}{\rho(A)} - L\right)^m = \left(\frac{A}{\rho(A)}\right)^m - L^m$, y 2) sigue pues $L^m = L$.

Veamos 2). Sea $k :=$ multiplicidad algebraica de $\lambda = \rho(A)$. Usando el teorema de Schur, $A = U^* \Delta U$ con Δ triangular superior cuyos elementos diagonales son $\lambda, \dots, \lambda, \lambda_{k+1}, \dots, \lambda_n$, siendo $|\lambda_j| < \lambda$ para $j = k+1, \dots, n$.

Por esta razón $\left(\frac{\Delta}{\lambda}\right)^m$ es triangular superior cuyos elementos diagonales son $1, \dots, 1, (\lambda_{k+1}/\lambda)^m, \dots, (\lambda_n/\lambda)^m$, que para $m \rightarrow \infty$ convergen a $1, \dots, 1, 0, \dots, 0$. Como $A^m = U^* \Delta^m U$, resulta de 1) que:

$$L = U^* \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix} U.$$

Como U y U^* son matrices de rango n y la matriz central tiene rango por lo menos k , resulta que $(\text{rango de } L) \geq k$. Pero como $Lw = (xy^T)w = x(y^T w) = \beta x$ con $\beta = (y^T w)$, sigue que rango de L es igual a 1 y $1 \geq k$, QED.

Enunciamos un teorema que probaremos más adelante.

Teorema de Ky Fan. Sean $A, B \in M_n(\mathbb{C})$, tales que $B \geq |A|$.

Entonces, todo autovalor de A yace en la región:

$$\bigcup_{i=1}^n \{ z \in \mathbb{C} : |z - a_{ii}| \leq \rho(B) - b_{ii} \}.$$

Observación. Si $A \geq_n 0$, podemos aplicar este resultado con $B = A$, resultando $\sigma(A) \subset \bigcup_{i=1}^n \{ z \in \mathbb{C} : |z - a_{ii}| \leq \rho(A) - a_{ii} \}$.

Este resultado es más preciso que $\sigma(A) \subset \{ z \in \mathbb{C} : |z| \leq \rho(A) \}$!

Mencionamos otro resultado interesante:

Teorema 36 (E. Hopf). Si $A > 0$, $\lambda \in \sigma(A) \setminus \{\rho(A)\}$, entonces

$$|\lambda|/\rho(A) \leq \frac{\max \{a_{ij}\} - \min \{a_{ij}\}}{\max \{a_{ij}\} + \min \{a_{ij}\}} < 1.$$

Definición 23. Diremos que una matriz $A \in M_n(\mathbb{C})$ pertenece a la clase E de las *matrices estocásticas* si $A \geq 0$ y $\sum_j a_{ij} = 1$ para $i = 1, \dots, n$.

Si $A \in E$ y $A^T \in E$, diremos que A es *doblemente estocástica*, o más brevemente que $A \in DE$.

Ejercicio. Si U es una matriz unitaria y $a_{ij} := |u_{ij}|^2$, entonces la matriz $[a_{ij}] \in DE$. En particular, toda matriz de permutación P pertenece a DE .

Teorema 37 (Birkhoff). 1) Sea $A \in M_n(\mathbb{C})$, $A \in DE$. Existen matrices de permutación $P_1, \dots, P_N$ y constantes $\alpha_1, \dots, \alpha_N$, $0 < \alpha_j \leq 1$, $\sum \alpha_j = 1$, tales que $A = \alpha_1 P_1 + \dots + \alpha_N P_N$.

2) $N \leq n^2 - 2n + 2$.

Antes de probar el teorema repasaremos algunas nociones sobre conjuntos convexos.

Si $x_1, \dots, x_m \in \mathbb{C}^n$, $\lambda_i > 0$, $\sum \lambda_i = 1$ y $x = \sum_{i=1}^m \lambda_i x_i$ entonces se dice x que es una *combinación convexa* de los x_i .

Un conjunto $K \subset \mathbb{C}^n$ se dice que es *convexo* si dados $x, y \in K$, $0 < \alpha < 1$, entonces $\alpha x + (1-\alpha)y \in K$. Se puede probar que K es convexo sólo si contiene las combinaciones convexas de todo subconjunto finito de K . La convexidad se mantiene por sumas e intersecciones: $K_1 + K_2 := \{z = x + y ; x \in K_1, y \in K_2\}$ es convexo si K_1 y K_2 lo son. Idem $K_1 \cap K_2$.

Por *cápsula convexa* de un conjunto K entendemos el menor conjunto convexo S que contiene a K . La denotamos con $\text{Co}(K)$. Se demuestra que $\text{Co}(K) = \{z ; z = \text{combinación convexa de puntos de } K\}$.

Definición 24. Sea K un convexo cerrado. Un punto $p \in K$ se dice *punto extremal* de K si $p = \alpha x + (1-\alpha)y$, $x, y \in K$, $0 < \alpha < 1$ sólo se cumple si $x = y = p$. Se tiene el siguiente

Teorema 38. Todo convexo compacto (cerrado y acotado) tiene puntos extremales.

El siguiente es uno de los teoremas más importantes del análisis convexo:

Teorema 39 (Krein-Milman). Si K es un convexo compacto y $E :=$ familia de puntos extremales de K entonces $\overline{\text{Co}(E)} = K$.

Demostración del T. de Birkhoff. El conjunto DE de matrices

doblemente estocásticas es convexo, compacto $\subset C^{n^2}$. Las matrices P de permutación son los puntos extremales de DE . En efecto, los elementos de una matriz $A \in DE$ verifican $0 \leq a_{ij} \leq 1$. Entonces si $A, B \in DE$, $0 < \alpha < 1$, $P = \alpha A + (1-\alpha)B$, resulta $p_{ij} = \alpha a_{ij} + (1-\alpha)b_{ij}$ y la única manera que p_{ij} pueda valer 0 o 1 es que $a_{ij} = b_{ij} = p_{ij}$.

Veamos que si $A \in DE$ no es de permutación, no es extremal en DE .

A tiene un elemento $a_{i_1 j_1}$, $0 < a_{i_1 j_1} < 1$. Como $\sum_k a_{i_1 k} = 1$,

hay un $j_2 \neq j_1$ tal que $0 < a_{i_1 j_2} < 1$. Como la suma de la colum-

na j_2 es 1, hay un $i_2 \neq i_1$ tal que $0 < a_{i_2 j_2} < 1$. Seguimos así hasta que aparece un subíndice ya encontrado anteriormente.

Nos quedamos con la secuencia de índices comenzando en el índice repetido cuando aparece por primera vez y terminando cuando aparece por segunda vez. Tendremos así una sucesión de

índices de la forma: $i_1 j_1, i_1 j_2, i_2 j_2, \dots, i_{h-1} j_{h-1}, i_{h-1} j_1$,

dos a dos distintos, tales que los elementos de A

correspondientes a esos índices son positivos y menores que uno.

Sea a el mínimo de estos elementos. Sea B la matriz cuyos elementos $i_k j_k$ son iguales a $+1$, los elementos $i_k j_{k+1}$ son iguales a -1 , ($j_h := j_1$), y los restantes elementos nulos.

Definiendo $A_+ := A + aB$, $A_- := A - aB$ resultan $A_+ \geq 0$, $A_- \geq 0$ y las sumas de sus filas y columnas son $+1$. Luego, A_+ , $A_- \in DE$.

Como $A = \frac{1}{2} A_+ + \frac{1}{2} A_-$ y $A \neq A_+$, sigue que A no es extremal, QED.

NB. Si E es finito entonces $\overline{\text{Co}(E)} = \text{Co}(E)$.

16. LOCALIZACION DE LOS AUTOVALORES.

Vamos a demostrar el teorema de Geršgorin, que nos dice donde están localizados los autovalores de una matriz, $A \in M_n(\mathbb{C})$.

Para ello enunciamos un teorema sobre la dependencia continua de los ceros de un polinomio respecto de sus coeficientes.

Teorema 40. Sean $n \geq 1$ y p un polinomio con ceros λ_i , $i=1, \dots, n$,
 $p(x) = a_n x^n + \dots + a_1 x + a_0$, $a_n \neq 0$, $a_i \in \mathbb{C}$.

Entonces, para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que para todo polinomio $q(x) = b_n x^n + \dots + b_1 x + b_0$ con $|b_i - a_i| < \delta$

hay una forma de numerar sus ceros $\mu_1, \dots, \mu_n$ de manera que se verifique $|\lambda_i - \mu_i| < \varepsilon$.

Corolario. Los autovalores de una matriz $A \in M_n(\mathbb{C})$ dependen continuamente de los elementos de la matriz.

Teorema 41 (Geršgorin). Sea $A \in M_n(\mathbb{C})$ y $R'_i(A) := \sum_{j \neq i} |a_{ij}|$.

Vale:

1) Los autovalores de A están localizados en la región

$$G(A) := \bigcup_{i=1}^n \{ |z - a_{ii}| \leq R'_i(A) \}, \text{ (unión de } n \text{ discos).}$$

2) Si k de estos discos forman una región $G_k(A)$ conexa maximal, entonces en G_k hay exactamente k autovalores.

Demostración. Sea λ un autovalor de A y $x \neq 0$ tal que $Ax = \lambda x$.

Sea p tal que $|x_p| = \max \{ |x_i| ; i=1, \dots, n \} > 0$. Entonces

$$\lambda x_p = \sum_j a_{pj} x_j \text{ implica que } (\lambda - a_{pp})x_p = \sum_{j \neq p} a_{pj} x_j.$$

Usando la desigualdad triangular y la definición de p :

$$|\lambda - a_{pp}| |x_p| \leq R'_p(A) |x_p|. \text{ O sea, para algún } p, |\lambda - a_{pp}| \leq R'_p.$$

Luego, $\lambda \in G(A)$.

2) Supongamos que $G_k = \bigcup_{i=1}^k \{ |z - a_{ii}| \leq R'_i(A) \}$ es conexo,

disjunto de $G(A) \setminus G_k$. Sea D la matriz diagonal tal que

$d_{ii} = a_{ii}$, $i = 1, \dots, n$, y B la matriz definida por $A = D + B$.

Definamos $A_\epsilon := D + \epsilon B$, $0 \leq \epsilon \leq 1$. Tenemos $R'_j(A_\epsilon) = \epsilon R'_j(A)$.

Obsérvese que los discos que forman la región de Geršgorin de A_ϵ tienen el mismo centro que los de $G(A)$ y radio menor. Luego, k discos de los que forman a $G(A_\epsilon)$ están en G_k y $(n-k)$ discos en $G(A) \setminus G_k$. Por el teorema 40 podemos numerar los autovalores de A_ϵ de manera que $\lambda_i(A_\epsilon)$ sea una función continua de ϵ para cada $i = 1, \dots, n$, sobre el intervalo $[0, 1]$.

Luego, dado i , $\lambda_i(A_\epsilon)$ pertenece a G_k para $\epsilon=0$ si y sólo si lo hace para todo ϵ , $0 \leq \epsilon \leq 1$. Como $A_0 = D$, vale que $\lambda_i(A_0) = a_{ii}$. Por lo tanto, para $\epsilon=0$ hay exactamente k raíces en G_k . Entonces lo mismo vale para todo ϵ , en particular para $\epsilon=1$, QED. Obsérvese que lo que hemos hecho es probar que el número de autovalores de A en G_k es una función continua de ϵ a valores enteros, que debe ser constante pues su dominio es conexo.

Definición 25. La región $G(A)$ se denomina la *región de Geršgorin* (para filas) de A . Los discos individuales en $G(A)$ se conocen con el nombre de *discos de Geršgorin*.

Observemos que A y $S^{-1}AS$ tienen los mismos autovalores, aunque sus regiones de Geršgorin son en general diferentes.

Por ejemplo, si S es la matriz diagonal: $s_{ii} = p_i > 0$, se obtiene fácilmente que $(S^{-1}AS)_{ij} = p_j a_{ij} / p_i$.

Aplicando el teorema de Geršgorin a esta matriz y a su traspuesta, se tiene el siguiente

Corolario. a) $\sigma(A) \subset G(S^{-1}AS) = \bigcup_{i=1}^n \{ |z - a_{ii}| \leq \frac{1}{p_i} \sum_j p_j |a_{ij}| \}$

b) $\sigma(A) \subset G((S^{-1}AS)^T) = \bigcup_{j=1}^n \{ |z - a_{jj}| \leq p_j \sum_i (|a_{ij}|/p_i) \}$.

Ejemplo. Sea $A = \begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix}$. Entonces $\sigma(A) = \{1, 2\}$. Si aplicamos el teorema de Geršgorin tenemos que $G(A)$ es el disco de radio 1 y centro en 1. Sin embargo usando el corolario obtenemos una mejor acotación: $\sigma(A) \subset \{ |z - 1| \leq r \} \cup \{2\}$ con $r = p_2/p_1$.

Veamos la siguiente aplicación:

Teorema 42 (Ky-Fan). Sea $A \in M_n(\mathbb{C})$, $B \geq |A|$. Entonces

$$\sigma(A) \subset \bigcup_{i=1}^n \{ |z - a_{ii}| \leq \rho(B) - b_{ii} \}.$$

Demostración. 1) Probaremos que vale la tesis en el caso $B > 0$. En este caso por el teorema de Perron existe $x > 0$ tal que $Bx = \rho(B)x$. Entonces :

$$\sum_j |a_{ij}| x_j \leq \sum_j b_{ij} x_j = (Bx)_i - b_{ii}x_i = (\rho(B) - b_{ii})x_i$$

Por lo tanto, si aplicamos el corolario anterior con $p_i = x_i$ obtenemos la tesis para el caso $B > 0$.

2) Sea $B \geq |A|$. Designemos con E la matriz con todos sus elementos = 1. Sea $B_\epsilon := B + \epsilon E$. Vale $0 < B_\epsilon \geq A$. Por 1)

$$\sigma(A) \subset \bigcup_{i=1}^n \{ |z - a_{ii}| \leq \rho(B_\epsilon) - (b_{ii} + \epsilon) \}. \text{ Pero para } \epsilon \rightarrow 0, \\ \rho(B_\epsilon) - (b_{ii} + \epsilon) \rightarrow \rho(B) - b_{ii}. \text{ Y sigue la tesis, QED.}$$

17. MATRICES NO NEGATIVAS.

Recordemos el

Teorema de Perron. Si $A \in M_n(\mathbb{R})$, $A > 0$, vale:

- a) $\rho(A) > 0$;
- b) $\rho(A) \in \sigma(A)$;
- c) $\exists x > 0 : Ax = \rho(A)x$
- d) $\rho(A)$ es algebraicamente simple;
- e) Si $\lambda \in \sigma(A) \implies |\lambda| < \rho(A)$;
- f) $\left[\frac{A}{\rho(A)} \right]^m \longrightarrow L, L := x y^T, Ax = \rho(A)x, A^T y = \rho(A)y,$
 $x > 0, y > 0, x^T y = 1.$

Vamos a estudiar cuales puntos siguen en pie para $A \geq 0$.

Ejercicio. Recurriendo a las matrices $\begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix}, \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$ probar que, excepto b), el teorema precedente no vale si la hipótesis $A > 0$ se reemplaza por $A \geq 0$.

Probaremos que con modificaciones valen a), d) y c).

Teorema 43. i) Si $A \in M_n(\mathbb{R})$, $A \geq 0$ entonces $\rho(A) \in \sigma(A)$ y $\exists x \geq 0$, no nulo, tal que $Ax = \rho(A)x$.

ii) Si existe un entero $k \geq 1$ tal que $A^k > 0$ entonces $\rho(A) > 0$ y $\exists x > 0$ tal que $Ax = \rho(A)x$.

Demostración. i) Lo haremos aproximando A por la matriz positiva $A(\epsilon) := A + \epsilon \tilde{E}$, $\epsilon > 0$, a la cual se aplica el teorema de Perron.
 Sea $x(\epsilon)$ el vector de Perron de $A(\epsilon)$. Entonces $x(\epsilon) > 0$, $\sum x(\epsilon)_i = \|x(\epsilon)\|_2^2 = 1$, $A(\epsilon)x(\epsilon) = \rho(A(\epsilon)) \cdot x(\epsilon)$. El conjunto $\{x; \|x\|_2 = 1\}$ es compacto por lo que podemos encontrar una sucesión $(\epsilon_k) \searrow 0$ tal que $x(\epsilon_k)_i \rightarrow x_i$ para $k \rightarrow \infty$. Pero entonces $\sum x_i^2 = \lim \|x(\epsilon_k)\|^2 = 1$. Luego, $x \neq 0$. Además, por el teorema 2 $\rho(A) \leq \rho(A(\epsilon_k)) \leq \rho(A(\epsilon_{k-1}))$. Se tiene entonces que la sucesión $\rho(A(\epsilon_k))$ es monótona decreciente y converge a un límite $\rho \geq \rho(A)$. Además, $Ax = \rho x$. Pero entonces $\rho \in \sigma(A)$ y $\rho \leq \rho(A)$. O sea, $\rho = \rho(A)$.
 ii) Sea x un autovector no negativo correspondiente a $\rho = \rho(A)$. Entonces $A^2 x = A(\rho x) = \rho^2 x$, ..., $A^k x = \rho^k x$. Si $A^k > 0$, entonces ρ^k es su raíz de Perron y x es un múltiplo de su autovector de Perron; luego, $\rho > 0$ y $x > 0$, QED.

Ejercicio. Sabemos por teorema 18 que si $A \geq 0$, $x > 0$ y $Ax \leq Bx$, entonces $\rho(A) \leq B$. Usando $A = \begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix}$ y $x = \begin{vmatrix} 1 \\ 0 \end{vmatrix}$, probar que el resultado no vale si $x \geq 0$, $x \neq 0$.

Mostrar también que $\rho(A) \neq \min_{x>0} \max_i \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j$.

Complementamos el teorema 43 (d) y e) del t. de Perron)

Teorema 45. Sea $A \in M_n(\mathbb{C})$, $A \geq 0$, $A^k > 0$. Entonces $\rho(A)$ es algebraicamente simple y todo $\lambda \in \sigma(A) \setminus \{\rho(A)\}$ verifica $|\lambda| < \rho(A)$.

Demostración. Por el teorema de Schur, $\sigma(A^k) = \{\lambda^k; \lambda \in \sigma(A)\}$. Si $\rho(A)$ fuera múltiple para A , $\rho(A)^k$ sería múltiple para A^k , contradiciendo d) del T. de Perron. El resto sigue de e), QED.
 Recordemos que una matriz $A \geq 0$ que verifica $A^k > 0$ tiene la propiedad FC y luego, por el teorema 6, es irreducible. La

recíproca no es cierta: $A = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$ es irreducible pero ninguna de sus potencias es positiva. Caracterizamos a continuación esa clase de matrices irreducibles.

Definición 26. $A \geq 0$ se dice *primitiva* si es irreducible y posee sólo un autovalor de módulo máximo.

Teorema 46. $A \in M_n(\mathbb{R})$, $A \geq 0$, es primitiva si y sólo si existe un número natural $k \geq 1$ tal que $A^k > 0$.

Demostración. $\Leftarrow$ es el contenido del teorema 45.

$\Rightarrow$. Observemos la demostración del teorema 35, 1). (O sea el teorema de Perron 2ª parte: $\lim_{m \rightarrow \infty} \left[\frac{A}{\rho(A)} \right]^m = L := x y^T$). En dicha demostración sólo se usan los siguientes hechos:

- 1) $\rho(A)$ es geométricamente simple,
- 2) $\rho(A)$ es el único autovalor de módulo máximo,
- 3) A (resp. A^T) tiene un autovector $x > 0$ (resp. $y > 0$), correspondiente a $\rho(A)$.

1) y 2) siguen de la Def. 26 y el Teor. 43. En el Teor. 47 se verá que 1)-3) valen para toda matriz irreducible no negativa. Aceptando esto podemos repetir la demostración mencionada y concluir que

$$\lim_{m \rightarrow \infty} \left[\frac{A}{\rho(A)} \right]^m = L := x y^T, \text{ donde } x > 0 \text{ e } y > 0 \text{ implican } L > 0.$$

Entonces para m suficientemente grande $A^m > 0$, QED.

Probaremos ahora los resultados del teorema de Perron que se mantienen para matrices irreducibles.

Teorema 47 (Perron-Frobenius). Sea $A \geq 0$, irreducible. Vale

- a) $\rho(A) > 0$,
- b) $\rho(A) \in \sigma(A)$,
- c) $\exists x > 0$ tal que $Ax = \rho(A)x$,
- d) $\rho(A)$ es algebraicamente simple.

Demostración. a): Las hipótesis implican que $\alpha := \min_{i,j} a_{ij} > 0$. Por teorema 17, i), resulta $\rho(A) \geq \alpha > 0$.

b): Ya lo demostramos en teorema 42, i).

c): Sabemos que $\exists x \geq 0$ tal que $Ax = \rho(A)x$. Luego, $(I + A)^{n-1}x = (1 + \rho(A))^{n-1}x$. Pero por teorema 6, la matriz $(I + A)^{n-1}$ es positiva y por lo tanto $(I + A)^{n-1}x > 0$. Sigue c).

Para probar d) necesitamos un lema.

Lema. Si $A \in M_n(\mathbb{C})$ y $\lambda \in \sigma(A)$ tiene multiplicidad algebraica k entonces $(\lambda + 1) \in \sigma(I + A)$ con multiplicidad algebraica k .

Además, $\rho(I + A) \leq 1 + \rho(A)$. Si $A \geq 0$ entonces $\rho(I + A) = 1 + \rho(A)$.

Demostración del lema. Observemos que $p_A(t) = \det [tI - A] = \det [(t+1)I - (I+A)] = p_{I+A}(t+1)$. O sea, $p_A(t) = \prod (t - \lambda_j) = \prod (s - (\lambda_j+1)) = p_{I+A}(s)$, $s=t+1$.

Entonces, si λ es un cero de multiplicidad k de $p_A(t)$, $\lambda+1$ es un cero de multiplicidad k de $p_{I+A}(s)$. También $\sigma(I+A) = \{\mu = 1+\lambda; \lambda \in \sigma(A)\}$ y por lo tanto $\rho(I+A) \leq 1+\rho(A)$. Si $A \geq 0$ entonces $\rho(A) \in \sigma(A)$ y $1+\rho(A) \in \sigma(I+A)$. Luego, $\rho(I+A) = 1+\rho(A)$, QED.

Demostración de d): Si $\rho(A)$ es de multiplicidad algebraica k entonces $1+\rho(A)$ es de multiplicidad k para $I+A$. Pero $I+A$ es positivo y $1+\rho(A)$ es su autovalor de Perron. Por teorema 35, 2), $k=1$, QED.

Aislamos a continuación un resultado que ya hicimos notar en la demostración del teorema 46.

Teorema 48. Si $A \geq 0$ es una matriz primitiva entonces

$$\lim_{m \rightarrow \infty} \left[\frac{A}{\rho(A)} \right]^m = xy^T =: L, \text{ donde } x > 0 \text{ es autovector de } A$$

para $\rho(A)$, $y > 0$ es autovector de A^T para $\rho(A)$ tal que $y^T x = 1$.

18. GENERALIZACIONES DE ALGUNOS RESULTADOS A MATRICES NO NEGATIVAS.

Sea $A \in M_n(\mathbb{R})$, $A \geq 0$, $\alpha \in \mathbb{R}$, $\alpha \geq 0$.

Teorema 49. i) Supongamos que $\exists x \geq 0$, no nulo, que verifica $Ax \geq \alpha x$. Entonces $\rho(A) \geq \alpha$.

ii) Sea $x \geq 0$, no nulo, tal que $Ax \geq \rho(A)x$. Si $\exists y > 0$ tal que $y^T A = \rho(A)y^T$ entonces vale $Ax = \rho(A)x$.

Demostración. i) Sea $A(\varepsilon) := A + \varepsilon E$. Entonces $A(\varepsilon) > 0$ y existe un vector $y(\varepsilon) > 0$ tal que $y(\varepsilon)^T A(\varepsilon) = \rho(A(\varepsilon))y(\varepsilon)^T$. Además $A(\varepsilon)x - \alpha x > Ax - \alpha x \geq 0$. Multiplicando por $y(\varepsilon)^T$ resulta:

$\rho(A(\varepsilon))y(\varepsilon)^T x - \alpha y(\varepsilon)^T x = (\rho(A(\varepsilon)) - \alpha)y(\varepsilon)^T x > 0$. Luego $(\rho(A(\varepsilon)) - \alpha) > 0$. Como los autovalores de $A(\varepsilon)$ convergen a los de A , resulta que $\rho(A(\varepsilon)) \rightarrow \rho(A)$ e i) queda probado.

ii) Por hipótesis $Ax - \rho(A)x \geq 0$. Si fuera $\neq 0$, entonces multiplicando por y^T resultaría $y^T(Ax - \rho(A)x) > 0$. Pero

$y^T(Ax - \rho(A)x) = (\rho(A) - \rho(A))y^Tx = 0$, contradicción, QED.

Teorema 50. Sea $A \geq 0$. Vale $\rho(A) = \max_{x \geq 0} \min_i \frac{1}{x_i} \sum_j a_{ij} x_j$,
(el mínimo se toma sobre los $x_i \neq 0$ y el máximo sobre los x no nulos).

Demostración. Sea $x \geq 0$, no nulo, y sea $\alpha := \min_i \frac{1}{x_i} \sum_j a_{ij} x_j$
(sólo sobre los i con $x_i \neq 0$). Entonces, $\alpha x \leq Ax$ pues
 $\alpha x_i \leq \sum_j a_{ij} x_j \quad \forall i$. Por el teorema anterior $\alpha \leq \rho(A)$. Como
 $A \geq 0$, existe $x \geq 0$, $x \neq 0$, con $Ax = \rho(A)x$. En consecuencia, el
máximo de los α se alcanza sobre ese x y es $\rho(A)$. El teorema
queda así probado.

Recordemos que el teorema 2 asegura que $|C| \leq D \implies$
 $\rho(C) \leq \rho(|C|) \leq \rho(D)$. Veremos que pasa si esas desigualdades son
igualdades.

Teorema 51. Sean $A, B \in M_n(\mathbb{C})$, $A \geq 0$, irreducible. Supongamos que
 $|B| \leq A$ y $\rho(A) = \rho(B)$. Si $\lambda = e^{i\Phi} \rho(B) \in \sigma(B)$, entonces
existen números reales $\theta_1, \dots, \theta_n$ tales que $B = e^{i\Phi} D A D^{-1}$ con

$$D = \begin{vmatrix} e^{i\theta_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\theta_n} \end{vmatrix} = \begin{vmatrix} x_1/|x_1| & & 0 \\ & \ddots & \\ 0 & & x_n/|x_n| \end{vmatrix}, \quad Bx = \lambda x.$$

Demostración. Existe $x \neq 0$ tal que $Bx = \lambda x$. Entonces,
 $|Bx| = \rho(A)|x| \leq |B||x| \leq A|x|$. Por ii) del teorema 49 (la
irreducibilidad de A asegura la existencia del $y > 0$) sigue
que las desigualdades son igualdades y que $|x|$ es autovector de
 A y también de $|B|$. Por el T. de Perron-Frobenius resulta
 $|x| > 0$. Entonces están bien definidos $e^{i\theta_k} := x_k/|x_k|$ y vale
 $x = D|x|$. Escribimos: $\lambda x = e^{i\Phi} \rho(A)x = e^{i\Phi} \rho(A)D|x| = Bx = BD|x|$.
Luego, $\rho(A)|x| = e^{-i\Phi} D^{-1} B D |x| =: \tilde{B}|x|$.

Observando la matriz $\tilde{B}$ vemos que $|B| = |\tilde{B}|$ y las identidades
anteriores muestran que $A|x| = \tilde{B}|x|$, $|\tilde{B}| \leq A$.

Entonces tenemos $\sum_j a_{ij}|x_j| = \sum_j \tilde{b}_{ij}|x_j| = |\sum_j \tilde{b}_{ij}|x_j|| \quad \forall i$. La

última igualdad implica que $\tilde{b}_{ij}|x_j|$ es real y no negativo para cada i, j . Luego, $\tilde{B} \geq 0$. La primera igualdad, junto con el hecho que $\tilde{B} \leq A$, implican $\tilde{B} = A$, QED.

Corolario. Sea $A \geq 0$, irreducible y $\lambda \in \sigma(A) \setminus \rho(A)$ tal que $|\lambda| = \rho(A)$. Entonces, $A = e^{i\Phi} D A D^{-1}$ con D como en el teorema y $e^{i\Phi} = \lambda / \rho(A)$.

Teorema 52. Sea $A \geq 0$, irreducible. Vale

i) Si $S := \{\lambda_n = \rho(A), \lambda_{n-1}, \dots, \lambda_{n-k+1}\}$ es la familia de autovalores (distintos) de A de módulo $\rho(A)$ entonces cada λ_j es de multiplicidad algebraica uno y

$$S = \{e^{2\pi i p/k} \rho(A); p=0,1,\dots,k-1\}.$$

ii) Si $\lambda \in \sigma(A)$ entonces $\lambda e^{2\pi i p/k} \in \sigma(A)$ para $p=0,1,\dots,k-1$, y ambos tienen la misma multiplicidad.

Corolario 1. k es divisor de n .

Corolario 2. Sea $A \geq 0$, irreducible. Sea $A^m = [a_{ij}^{(m)}]$. Si hay $k \geq 2$ autovalores de módulo máximo entonces $a_{ii}^{(m)} = 0$ si m no es múltiplo de k , $\forall i = 1, \dots, n$. En particular $a_{ii} = 0$.

Demostración. Sea $\Phi = 2\pi/k$, $\lambda = e^{i\Phi} \rho(A)$. $A = e^{i\Phi} D A D^{-1}$ implica $A^m = e^{im\Phi} D A^m D^{-1}$, de donde $a_{ii}^{(m)} = e^{im\Phi} a_{ii}^{(m)}$. Como $e^{im\Phi} \neq 1$, si m/k no es entero, tiene que ser $a_{ii}^{(m)} = 0$, QED.

Demostración del teorema 52. Observemos el siguiente hecho:

Sea $\lambda_{n-u} = e^{i\Phi_u} \rho(A) \in S$. Aplicando el corolario al teorema 51

resulta $A = e^{i\Phi_u} D A D^{-1}$. Entonces, $p_A(t) = M p_{D A D^{-1}}(t \cdot e^{-i\Phi_u}) = M p_A(t \cdot e^{-i\Phi_u})$, M una constante no nula. Luego, $\lambda \in \sigma(A)$ si

y sólo si $\lambda \cdot e^{i\Phi_u} \in \sigma(A)$ y con la misma multiplicidad algebraica.

En particular, $e^{i\Phi_u} S = S$. Esto implica lo siguiente :

dados $u, r \in \{0, \dots, k-1\}$ existe un $q \in \{0, \dots, k-1\}$

tal que $e^{i\Phi_u} \lambda_{n-q} = \lambda_{n-r} = e^{i\Phi_r} \rho(A)$.

Entonces, $\Phi_u + \Phi_q = \Phi_r \pmod{2\pi}$. Esto prueba que $\{\Phi_0, \dots, \Phi_{k-1}\} =$

$\Rightarrow$ G es un grupo abeliano con la operación suma (módulo 2π).

(Observemos que $\Phi_0 = 0$ asegura la existencia del elemento opuesto). Supongamos los Φ ordenados: $0 = \Phi_0 < \Phi_1 < \dots < \Phi_{k-1} < 2\pi$. Entonces $G_1 := \{ h\Phi_1; h \in \mathbb{Z} \}$ (mód 2π) es también un grupo abeliano con la operación suma (módulo 2π), $G_1 \subseteq G$.

Si no fueran coincidentes estos dos grupos, habría un Φ_α tal que $h\Phi_1 < \Phi_\alpha < (h+1)\Phi_1$ y en consecuencia $0 < \Phi_\alpha - h\Phi_1 < \Phi_1$. Pero esto es un absurdo pues $\Phi_\alpha - h\Phi_1 \in G$ y es menor que Φ_1 . Luego $G = G_1$, lo que prueba i).

Para ver ii) basta aplicar la observación al comienzo de la demostración que afirma que $e^{i\Phi \cdot \sigma(A)} \longrightarrow \sigma(A)$, respetando multiplicidades, QED.

Ejercicios. Sea $A \geq 0$, irreducible.

i) Si $a_{ii} > 0$ para algún i entonces $\rho(A)$ es el único autovalor de módulo máximo.

ii) La recíproca no vale: $A = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix}$, $\sigma(A) = \{2, -1, -1\}$.

Enunciamos algunos teoremas complementarios.

Teorema 53. Sea $A \in M_n(\mathbb{R})$, $A \geq 0$, $\rho(A) > 0$. Si $\lambda \in \sigma(A)$, $|\lambda| = \rho(A)$, entonces $\lambda/\rho(A) = e^{i\Phi}$ es una raíz de la unidad: $e^{ik\Phi} = 1$ para cierto k , $1 \leq k \leq n$. Para $p = 0, \dots, k-1$, vale $e^{ip\Phi} \rho(A) \in \sigma(A)$. (Pero puede ser que no sean estos los únicos autovalores de módulo máximo!).

Teorema 54. Sea $A \geq 0$ irreducible. Sean $x > 0$, $y > 0$ tales que $x^T y = 1$. $Ax = \rho(A)x$, $A^T y = \rho(A)y$. Entonces (convergencia en media Cèsaro)

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{m=1}^N \left[\frac{A}{\rho(A)} \right]^m = L := xy^T$$

La velocidad de la convergencia no es exponencial pero existe

$C = C(A) > 0$ tal que $\left\| \frac{1}{N} \sum_{m=1}^N \left[\frac{A}{\rho(A)} \right]^m - L \right\|_\infty \leq \frac{C}{N} \quad \forall N \geq 1$.

El siguiente teorema precisa la caracterización dada en el teorema 46.

Teorema 55. Sea $A \in M_n(\mathbb{R})$; $A \geq 0$. A es primitiva si y sólo si

$$A^{n^2-2n+2} > 0.$$

Teorema 56. Sea $A \in M_n(R)$, $A \geq 0$, irreducible. Sea $\{P_j\}$ la familia de nodos de $\Gamma(A)$. Denotamos con $L_j = \{k_1(j), k_2(j), \dots\}$ al conjunto de las longitudes de todos los caminos dirigidos que comienzan y terminan en P_j , $j = 1, \dots, n$, (o sea de los ciclos).

Sea $g_j :=$ máximo común divisor de los elementos de L_j . Vale:

A es primitiva si y sólo si $g_j = 1$ para todo j .

Demostración. $L_j \neq \emptyset$ pues A es irreducible. Supongamos que A es primitiva. Entonces $A^m > 0$ para cierto m , lo mismo que $A^{m+k} \forall k > 0$.

En consecuencia, $L_j \supset \{m, m+1, \dots\}$ y esto implica $g_j = 1$.

Si por el contrario A no es primitiva entonces A tiene $k \geq 2$ autovalores de módulo máximo y por el corolario 2 del T. 52,

$a_{ii}^{(m)} = 0$ si m no es un múltiplo de k . Luego $L_i \subset \{k, 2k, 3k, \dots\}$ para cada $i = 1, \dots, n$. En consecuencia, g_i es divisible por

k . Luego $g_i \geq k > 1$, QED.