INFORME TECNICO INTERNO

Nº. 86

INSTITUTO DE MATEMATICA DE BAHIA BLANCA

INMABB (UNS - CONICET)

UNIVERSIDAD NACIONAL DEL SUR

Avda. ALEM 1253 - 8000 BAHIA BLANCA

República Argentina

INFORME TÉCNICO INTERNO

N° 86

INSTITUTO DE MATEMÁTICA DE BAHÍA BLANCA

INMABB (CONICET - UNS)

UNIVERSIDAD NACIONAL DEL SUR

Avda. ALEM 1253 - 8000 BAHIA BLANCA

- 2003 -

INFORME TÉCNICO INTERNO N° 86

La Matriz Positiva y su Espectro

Agnes Benedek y Rafael Panzone

Universidad Nacional del Sur

INMABB

CONICET - UNS

AÑO 2003

PROLOGO

Estas lecciones se iniciaron con un curso optativo que dicté en el primer semestre de 1994. El manuscrito original se formó con los apuntes tomados en clase por el Lic. Edgardo Fernández Stacco. Partiendo de ese material la Dra. Agnes Benedek redactó la versión final:

A. Benedek, E. Fernández y R. Panzone, MATRICES NO NEGATIVAS, INMABB (UNS-CONICET),I:T.I. nº 35,(1994).

Las notas presentes son una variación y ampliación de aquellas.

El material se encuentra aquí o allá en el excelente libro:

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

o bien en,

E. Seneta, NONNEGATIVE MATRICES, Wiley, New York, (1973).

En su mayor parte es presentado también en

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

o en la última versión de este clásico de la teoría de matrices editado por Springer (1986) y titulado

MATRIZENTHEORIE.

El propósito del curso es, primero, 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. En segundo término explorar su espectro de autovalores.

Para su lectura es conveniente tener alguna familiaridad con la teoría de matrices, como la que se logra en los cursos de álgebra, o bien, tener a mano algún texto. A los citados se puede agregar

A. C. Aitken, DETERMINANTES Y MATRICES, Dossat.

R. P.

INDICE GENERAL

INDICE ALFABETICO

INDICE DE TEOREMAS.

LA MATRIZ POSITIVA Y SU ESPECTRO

por Agnes Benedek y Rafael Panzone.

PARTE I

0. DEFINICIONES BASICAS Y NOTACION. *Denotamos* con $M_{n,m}(C)$ al conjunto de matrices de n filas y m columnas cuyos elementos son números complejos. Si $n=m$ escribiremos $M_n(C)$ en lugar de $M_{n,m}(C)$.

DEFINICIÓN 1. 1) Diremos que λ es un *autovalor* de $A \in M_n(C)$ si existe un *autovector* correspondiente λ, es decir, un vector columna $x \neq 0$ tal que

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

2) Diremos que el autovalor es *simple* si dos vectores 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 y unos y tiene en cada fila y en cada columna exactamente un uno.

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

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

b) $n>1$ y existen una matriz de permutación P y un r, $1 \le r \le n-1$, tal que

$$P^T AP = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

con $B \in M_r, D \in M_{n-r}, C \in M_{n-r,r}$. ■

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

Sea P una matriz de permutación y $a_{i\tau(i)}$ el elemento $=1$ en la fila i. Entonces $\tau(i)$ es una permutación. Además, $\det(P) = \pm a_{1\tau(1)} \cdot \ldots \cdot a_{n\tau(n)} = \pm 1$. Si con A^T indicamos la *traspuesta* de A vale que $P^{-1} = P^T$ (ver más adelante); o sea, P es ortogonal.[1]

En una matriz reducible aparecen por lo menos n-1 elementos nulos. Consideremos el sistema lineal $Ax=y$, $A \in M_n$ reducible, $n>1$. Sea $\tilde{A}$ la matriz que aparece en 6): $\tilde{A} = P^T AP = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$. Si definimos $\tilde{x} = P^T x$ e $\tilde{y} = P^T y$ resulta $\tilde{y} = \tilde{A}\,\tilde{x}$. Escribiendo $\tilde{x} = \begin{bmatrix} z \\ U \end{bmatrix}, \tilde{y} = \begin{bmatrix} w \\ V \end{bmatrix}$, el último sistema se reduce a

[1] Cuando sea necesario escribiremos A^{tr} en lugar de A^T.

$$\begin{aligned} Bz + CU &= w \\ DU &= V \end{aligned},$$

que son dos sistemas de orden $n\!-\!r$ y r respectivamente.

1. UN MODELO DE MIGRACION. Comenzamos tratando un modelo donde aparecen naturalmente la *matrices no negativas*. Es un problema de migraciones entre n estados vecinos C_i, i=1,...,n. Suponemos que entre los habitantes no hay ni muertes ni nacimientos y que cada mes una proporción fija $a_{i,j}$ 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,

$a_{i,j} \,.\, p_j^{(m)}$=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,

$\sum_{i=1}^{n} a_{ij} = 1,\ 0 \le a_{ij} \le 1, \quad p_i^{(m+1)} = a_{i1}p_1^{(m)} + a_{i2}p_2^{(m)} + \ldots + a_{in}p_n^{(m)}$. Escribiremos esto en forma matricial. Sea $A = \lfloor a_{ij} \rfloor$ y $p^{(m)} = \begin{bmatrix} p_1^{(m)} \\ \cdot \\ \cdot \\ \cdot \\ p_n^{(m)} \end{bmatrix}$, m=0,1,... . Se tiene entonces:

$p^{(m+1)} = Ap^{(m)} = A(Ap^{(m-1)}) = \ldots = A^{m+1}p^{(0)}$. En el caso más simple, n=2, tenemos, $A = \begin{bmatrix} 1-\alpha & \beta \\ \alpha & 1-\beta \end{bmatrix}$ con $0 \le \alpha, \beta \le 1$. Esta matriz tiene como autovalores $\lambda_1 = 1-\alpha-\beta, \lambda_2 = 1$. Observemos que $1 = |\lambda_2| \ge |\lambda_1|$. Luego, el radio espectral de A verifica $\rho(A) = 1$=autovalor de A. Si además $\alpha, \beta > 0$ entonces A es irreducible. (En efecto, $P^T AP$ se obtiene de A efectuando la misma permutación en filas y columnas.)
Además el autovalor $\rho(A)$ es simple. Supongamos solamente que $\alpha + \beta > 0$. Entonces los autovalores son distintos y se verifica: $x = \begin{bmatrix} \beta \\ \alpha \end{bmatrix}$ es autovector para λ_2=1, $z = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ es autovector para $\lambda_1 = 1-\alpha-\beta$. En estas condiciones A es *diagonalizable*, es decir, existe una matriz no singular S tal que $S^{-1}AS$ es una matriz diagonal: $S^{-1}AS =: \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 1-\alpha-\beta \end{bmatrix}$. Dicha matriz es

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

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 $B = S^{-1}AS$.
A es *diagonalizable* si es semejante a una matriz diagonal. Si S es unitaria (ortogonal) diremos que B es *unitariamente* (*ortogonalmente*) semejante a A. ■

La semejanza $\approx$ es una relación de equivalencia.

Para completar el ejemplo supongamos que $0 < \alpha + \beta < 2$. Esto implica que $|\lambda_1| < 1$. Por tanto, $A = S\Lambda S^{-1}, A^m = S\Lambda^m S^{-1}$. Estamos interesados en el comportamiento asintótico de la población. Se calcula inmediatamente que $\Lambda^m = \begin{bmatrix} 1 & 0 \\ 0 & (1-\alpha-\beta)^m \end{bmatrix}$ de donde

$$lim\, A^m = S\,(lim\, \Lambda^m)\, S^{-1} = S\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} S^{-1} = \frac{1}{\alpha+\beta}\begin{bmatrix} \beta & \beta \\ \alpha & \alpha \end{bmatrix},$$

$$lim\, p^{(m)} = lim\,(A^m p^{(0)}) = (lim\, A^{(m)}) p^{(0)} = \frac{p_1^{(0)} + p_2^{(0)}}{\alpha+\beta}\begin{bmatrix} \beta \\ \alpha \end{bmatrix}.$$

Hemos verificado en este caso los puntos del siguiente teorema.

TEOREMA 1. Sea $A \geq 0$.

1) $\rho(A)$ es un autovalor: $\exists x \neq 0\ tq\ Ax = \rho(A)x$.

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 que $\frac{1}{m}\sum_1^m A^j$ converge (media Cèsaro).■

En efecto, si $\alpha+\beta = 2$, o sea, si $\alpha = 1, \beta = 1$: $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $A^{2j} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $A^{2j+1} = A$.

Luego, $\frac{1}{m}\sum_1^m A^j \rightarrow \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

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 $\forall ij\ a_{ij} > 0$. Análogamente diremos que A es *no negativa*, $A \geq 0$, si $\forall ij\ a_{ij} \geq 0$ y $A \geq B$ si $A - B \geq 0$.

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

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

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

EJERCICIO 1. i) $|A| \geq 0$. $|A| = 0 \Leftrightarrow A = 0$.

ii) $|aA| = |a| \cdot |A|$ si $a \in C$.

iii) $|A+B| \leq |A| + |B|$.
iv) $0 \neq A \geq 0$ no implica $A>0$.
v) $A, B \geq 0, a, b \geq 0 \Rightarrow aA + bB \geq 0$.
vi) $A \geq B, C \geq 0 \Rightarrow A + C \geq B + C$.
vii) $A \geq B, B \geq C \Rightarrow A \geq C$.
viii) Si x es un vector columna, $|Ax| \leq |A||x|$.
ix) $|AB| \leq |A||B|$.
x) $|A^m| \leq |A|^m$.
xi) $0 \leq A \leq B, 0 \leq C \leq D \Rightarrow 0 \leq AC \leq BD$.
xii) $A \geq 0 \Rightarrow A^m \geq 0$; $A > 0 \Rightarrow A^m > 0$.
xiii) $A > 0, 0 \neq x \geq 0 \Rightarrow Ax > 0$.
xiv) $A \geq 0, x > 0, Ax = 0 \Rightarrow A = 0$.
xv) $|A| \leq |B| \Rightarrow \|A\|_2 \leq \|B\|_2$.
xvi) $\||A|\|_2 = \|A\|_2$.

EJERCICIO 2. 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 $A \in M_n$, $\rho(A) = \lim_{m \to \infty} \left(\|A^m\|_2\right)^{1/m}$. Usando este resultado el teorema sigue de x), xi), xv) y xvi) de los ejercicios precedentes. En efecto, $|A^m| \leq |A|^m \leq B^m$ y por tanto $\|A^m\|_2 \leq \||A|^m\|_2 \leq \|B^m\|_2$, QED.

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, $\forall i \;\; a_{ii} \leq \rho(A)$.■

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 tanto $\rho(\hat{A}) \leq \rho(A)$. Observando que $\tilde{A}$ y $\hat{A}$ tienen los mismos autovalores no nulos, $\rho(\tilde{A}) = \rho(\hat{A})$ y el corolario sigue, QED.

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

COROLARIO 3. $0 \leq A < B \Rightarrow \rho(A) < \rho(B)$.■

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 $\rho(\alpha A) \le \rho(B)$. Como $\alpha\rho(A) = \rho(\alpha A)$ tenemos $0 < \rho(A) < \alpha\rho(A) \le \rho(B)$, 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 $E_{i,j}$ el resultado de aplicar a la matriz identidad I la operacion 1), $M_i(c)$ la 2) y $S_{j,i}(c)$ la 3). O sea,

$$E_{i,j} = \begin{array}{cc} \begin{bmatrix} 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0 & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix} & \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \\ \end{matrix} \\ \begin{matrix} \quad & \quad & i & \quad & j & \quad & \quad & \quad \end{matrix} & \end{array} \qquad S_{j,i}(c) = \begin{array}{cc} \begin{bmatrix} 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & c & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix} & \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \\ \end{matrix} \\ \begin{matrix} \quad & \quad & i & \quad & j & \quad & \quad & \quad \end{matrix} & \end{array}$$

$$M_i(c) = \begin{array}{cc} \begin{bmatrix} 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & c & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix} & \begin{matrix} \\ \\ i \\ \\ \\ \\ \\ \\ \end{matrix} \\ \begin{matrix} \quad & \quad & i & \quad & \quad & \quad & \quad & \quad \end{matrix} & \end{array} .$$

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), $S_{j,i}(c)A$ resulta de aplicar 3) a la matriz A. Por otra parte $AE_{i,j}$ es el resultado de intercambiar las columnas i y j de A, $AM_i(c)$ de multiplicar la columna i de A por c y finalmente $AS_{j,i}(c)$ de 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 transposiciones de filas de la matriz identidad, resulta que P se puede escribir como producto de matrices: $P=E(1)E(2)\ldots E(k)$, donde $E(i)=E_{j_i k_i}$. Pero, de $E(j)^T = E(j)^{-1} = E(j)$ resulta $P^T = P^{-1} = E(k)\ldots E(1)$.

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 $B_1 = \{v_1, ..., v_n\}$ 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 medio de la aplicación:

$$x \to [x]_{B_1} = [\alpha_1, ..., \alpha_n]^{tr}.$$

Sea T una transformación lineal de V en sí mismo y $B_2 = \{w_1, ..., w_n\}$ 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: $[Tx]_{B_2} = \left[T \sum_j \alpha_j v_j\right]_{B_2} = \left[\sum_j \alpha_j T v_j\right]_{B_2} = \sum_j \alpha_j [Tv_j]_{B_2}$.

Sea $[Tv_j]_{B_2} = \begin{bmatrix} t_{1j} \\ \cdot \\ t_{nj} \end{bmatrix}$ y ${}_{B_2}[T]_{B_1} := \begin{bmatrix} t_{11} & \cdot & t_{1n} \\ \cdot & \cdot & \cdot \\ t_{n1} & \cdot & t_{nn} \end{bmatrix}$. 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,

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

donde

$${}_{B_2}[I]_{B_1} = \left[[v_1]_{B_2}, ..., [v_n]_{B_2}\right].$$

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

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

Luego, ${}_{B_2}[I]_{B_1}\, {}_{B_1}[I]_{B_2}$=matriz identidad y estas matrices no son singulares. Entonces, ${}_{B_2}[I]_{B_1} = ({}_{B_1}[I]_{B_2})^{-1}$. Ahora es de verificación inmediata que:

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

de donde surgen, respectivamente,

$${}_{B_2}[T]_{B_1} = {}_{B_2}[T]_{B_2}\, {}_{B_2}[I]_{B_1}, \qquad {}_{B_1}[T]_{B_2} = {}_{B_1}[I]_{B_2}\, {}_{B_2}[T]_{B_2}.$$

En consecuencia,

$${}_{B_1}[T]_{B_1} = {}_{B_1}[I]_{B_2}\, {}_{B_2}[T]_{B_2}\, {}_{B_2}[I]_{B_1}.$$

Esta es una transformación de semejanza entre representaciones de T.

Dada una base B_2 y una matriz no singular $S \in M_n$ existe una base B_1 tal que ${}_{B_2}[I]_{B_1} = S$. En efecto, escribiendo $S=[s_1,...,s_n]$ 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 $\{v_1,...,v_n\}$ es linealmente independiente y es la base B_1 requerida.

Sea $V \in M_{n,1}(=C^n)$. En V se define el *producto escalar* $<x,y>$: $<x,y>=y^* \cdot x = \sum x_j \overline{y_j}$.
Una base $B_2 = \{w_1,...,w_n\}$ se dice *ortonormal* si $<w_i, w_j> = \delta_{ij}$. Vale el
TEOREMA 3. Sea $B_2 = \{w_1,...,w_n\}$ una base ortonormal. Entonces $B_1 = \{v_1,...,v_n\}$ es una base ortonormal si y sólo si $S:= {}_{B_2}[I]_{B_1}$ es una matriz unitaria.■
DEMOSTRACIÓN. Sea S la matriz

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdot & s_{1n} \\ s_{21} & s_{22} & \cdot & s_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ s_{n1} & \cdot & \cdot & s_{nn} \end{bmatrix} = [s_1 \ldots s_n].$$

tal que, $s_j = [v_j]_{B_2}$. Como $[w_j]_{B_2} = \begin{array}{c} \begin{bmatrix} 0 \\ \cdot \\ 1 \\ \cdot \\ 0 \end{bmatrix} \end{array} \begin{array}{c} 1 \\ \cdot \\ j \\ \cdot \\ n \end{array}$, $v_j = \sum_i s_{ij} w_i$. Luego,

$$< v_k, v_r > = < \sum_i s_{ik} w_i, \sum_j s_{jr} w_j > = \sum_i s_{ik} \overline{s_{ir}} = \text{elemento } rk \text{ de la matriz } S^*S.$$

Entonces, B_1 es ortonormal si y sólo si $S^*S=I$, QED.
Enunciamos a continuación un teorema que probaremos más adelante.
TEOREMA 4, (Schur). Sea $A \in M_n(C)$ con autovalores $\lambda_1,...,\lambda_n$. Existe una matriz unitaria $U \in M_n(C)$ tal que

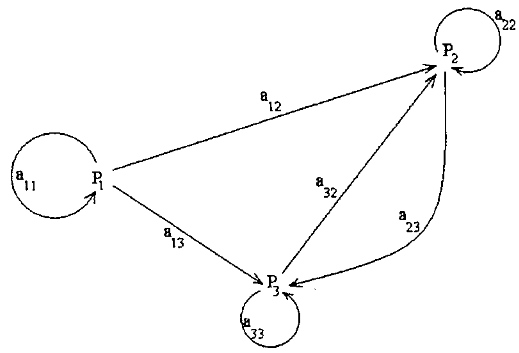
$$U^* AU = T = [t_{ij}]$$

es una matriz triangular superior, (esto es, $t_{ij} = 0$ si $j<i$). Además, $t_{ii} = \lambda_i$.
Si A y los autovalores son reales, U es real.■

5. IRREDUCIBILIDAD Y CONEXION.
DEFINICION 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 $k_1,...,k_m, 1 \leq m \leq n$, tales que $k_1 = p, k_m = q, a_{k_j k_{j+1}} \neq 0$ para $j = 1,...,m-1$.■
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 3. Son equivalentes:
1) $A \in FC$ 2) $|A| \in FC$ 3) $M(A) \in FC$.

DEFINICIÓN 6. Asociamos a cada matriz $A \in M_n$ un grafo $\Gamma(A)$ formado por n nodos $P_1,\ldots,P_n$ 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 3. Para la matriz $\begin{bmatrix} -1 & -2 & 1 \\ 0 & 3 & 4 \\ 0 & 5 & 6 \end{bmatrix}$, $\Gamma(A)$ será el grafo de la siguiente figura.



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

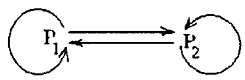
DEFINICIÓN 7. Por camino se entiende una sucesión finita de flechas

$$P_iP_j, P_jP_k, \ldots, P_rP_h.$$

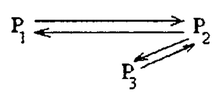
Por longitud de un camino entendemos el número de arcos que lo componen.
Un camino que comienza y termina en un mismo nodo, y no llega dos veces a ningún nodo, se denominará ciclo. Diremos que el grafo $\Gamma(A)$ es fuertemente conexo si para cada par de nodos P_i, P_j distintos hay un camino que comienza 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 4. Las siguientes matrices A y B son irreducibles, $A, B \in FC, C \notin FC$ y no es irreducible.

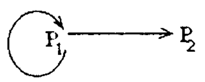
$A = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}$ $\Gamma(A)$



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



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



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 $m \le n-1$. $(|A|^m)_{ij} > 0$ 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 nonegativos de la forma $|a_{ik_1}||a_{k_1k_2}|\ldots|a_{k_{m-1}j}|$, QED.
DEMOSTRACION DEL TEOREMA 6. 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 dónde 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 tramos entre vértices repetidos, en el camino no se repetirán vértices y tendrá longitud $<n$.

c)$\Leftrightarrow$d). $(I+|A|)^{n-1} = I+(n-1)|A|+\ldots+|A|^{n-1} > 0$ si y sólo si para todo par $ij, i \neq j$, hay un h, $1 \leq h \leq n-1$ tal que $\left(|A|^h\right)_{ij} > 0$. 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)$\Leftrightarrow$c), se deja al lector.

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

Veamos que c)$\Rightarrow$a), probando la contrarrecíproca. Si A es reducible existe una matriz de permutación P tal que $P^T\widetilde{A}P = A$ con $\widetilde{A}$ tal que $\widetilde{a}_{ij} = 0$ si $i>r$, $j \leq r$, para cierto r, $1 \leq r < n$. La matriz $|\widetilde{A}|^h$ tiene nulos estos mismos elementos. Entonces $(I+|A|)^{n-1}$ no es positiva ni lo es $P^T(I+|\widetilde{A}|)^{n-1}P = (P^T(I+|\widetilde{A}|)P)^{n-1} = (I+|A|)^{n-1}$.

Veremos que a)$\Rightarrow$d), otra vez probando la contrarrecíproca. Sea $(I+|A|)^{n-1}$ no positiva. Existen entonces $p,q, p \neq q$, tales que $\left(|A|^h\right)_{pq}$=0 para h=1,...,n-1. Entonces no hay camino de P_p a P_q en el grafo $\Gamma(A)$. Separamos los nodos en dos clases:

$$S_1 = \{P_j : P_j = P_q \text{ o existe un camino de } P_j \text{ a } P_q\}, \qquad S_2 = \{P_j : P_j \notin S_1\}.$$

Por hipótesis ambos son no vacíos. Además no hay un camino entre un elemento de S_2 y uno de S_1. O sea, $a_{hj} = 0$ si $P_h \in S_2$ y $P_j \in S_1$. Renumeremos los nodos de manera que

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

donde $\widetilde{P}_j = P_{\sigma(j)}, \sigma$ una permutación de $\{1,\ldots,n\}$. Entonces el grafo $\Gamma(A)$ con los vértices renombrados es el grafo correspondiente a la matriz $\widetilde{A} = \{\widetilde{a}_{ij}\}$ con $\widetilde{a}_{ij} = a_{\sigma(i)\sigma(j)}$, o sea, $PAP^T = \widetilde{A}$, P^T matriz de permutación. Como en ese grafo no hay camino de ningún nodo de S_2 a uno de S_1, vale: $\widetilde{a}_{ij}$=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 ahora pensados como vectores columna. Sabemos que todo espacio vectorial de dimensión n sobre el cuerpo de los complejos es isomorfo a ese espacio. Representamos con x los elementos de C^n: $x = [x_1,\ldots,x_n]^T$.

DEFINICIÓN 8. Llamamos norma en C^n a una función $\|\cdot\| : C^n \to [0,\infty)$ tal que

1) $\|x\| \geq 0$
2) $\|x\| = 0 \Leftrightarrow x = 0$
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 $\|\cdot\|$, [.], sobre C^n son siempre equivalentes. Esto es, existen dos constantes positivas c_1, c_2 tales que para todo $x \in C^n : c_1[x] \leq \|x\| \leq c_2[x]$.■

Demostraremos más adelante este resultado y en un contexto más general.

Sin demostración enunciamos dos teoremas :

TEOREMA 8. Sea B la bola unitaria en C^n respecto a la norma $\|\cdot\|$. Entonces este conjunto B verifica:

a) B es acotado y cerrado, o sea, compacto,

b) B es equilibrado: $c \in C, |c| = 1, x \in B \Rightarrow cx \in B$,

c) B es convexo: si $x,y \in B$ entonces $rx+(1-r)y \in B$, para todo número real $r \in [0,1]$.■

Vale la recíproca. 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 $x+dy \in H$.

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.■

DEFINICIÓN 9. Una norma se dice monótona si $|y| \leq |x| \Rightarrow \|y\| \leq \|x\|$.

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

TEOREMA 10. Una norma es monótona si y sólo si es absoluta.■

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

Supongamos $\|\cdot\|$ absoluta. Sea $\alpha \in [0,1]$. Probaremos primero que: si $x = [x_1,\ldots,x_n]^T \in C^n$ e $y = [x_1,\ldots,x_{k-1},\alpha x_k, x_{k+1},\ldots,x_n]^T$ entonces $\|y\| \leq \|x\|$. En efecto, llamando $z = [x_1,\ldots,x_{k-1},-x_k,x_{k+1},\ldots,x_n]$ tenemos $|x| = |z|$ y por lo tanto $\|x\| = \|z\|$. Pero

$y = \frac{1-\alpha}{2}x + \frac{1+\alpha}{2}z$. Por la desigualdad triangular: $\|y\| \le \|x\|$. De esto sigue fácilmente la monotonía de la norma, QED.

DEFINICIÓN 10. Normas usuales en C^n: $\|x\|_1 := \sum_i |x_i|$, $\|x\|_2 := \left(\sum_i |x_i|^2\right)^{1/2}$. Estos son casos particulares de $\|x\|_p := \left(\sum_i |x_i|^p\right)^{1/p}$ donde el número real $p \in [1, \infty)$. Si $p = \infty$ entronces $\|x\|_\infty := max\{|x_i| : i = 1, \ldots, n\}$.■

7. NORMAS MATRICIALES.

Sea $A \in M_n(C)$. A la matriz A se la puede identificar con un elemento de 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 $M_n(C)$. Repetimos los axiomas 1) a 4) :

1) $|||A||| \ge 0$,

2) $|||A||| = 0 \Leftrightarrow A = 0$,

3) $|||cA||| = |c|\ |||A|||$,

4) $|||A + B||| \le |||A||| + |||B|||$.

DEFINICION 11. Si una norma en $M_n(C)$ verifica, además de 1)-4),

5) $|||AB||| \le |||A|||\ |||B|||$ (submultiplicatividad),

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

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

i) si $A^2 = A \ne 0$, entonces $|||A||| = |||A^2||| \le |||A|||^2$, y en consecuencia $1 \le |||A|||$,

ii) en particular la matriz identidad I verifica $|||I||| \ge 1$,

iii) $1 \le |||I||| = |||A\ A^{-1}||| \le |||A|||\ |||A^{-1}|||$. Luego, $|||A^{-1}||| \ge 1/|||A|||$,

iv) $|||A^k||| \le |||A|||^k$, o equivalentemente, $|||A^k|||^{1/k} \le |||A|||$.

PROPOSICION 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 $\|A\|_\infty := \sup\{|a_{ij}| : i, j = 0, \ldots, n\}$ no lo es.■

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

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

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

$\|\cdot\|_\infty$ no es matricial: sea $J = \begin{vmatrix}1 & 1\\ 1 & 1\end{vmatrix}$; entonces $J^2 = 2J$, $\|J^2\|_\infty = 2 \not\le \|J\|_\infty^2 = 1$

EJERCICIO 4. Probar que $|||A||| := n\|A\|_\infty$ es matricial.

EJERCICIO 5. 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 6. Probar que $\|A\|_p$, $1 \le p \le \infty$, es matricial si y sólo si $1 \le p \le 2$.

A una norma vectorial en C^n, $\|\cdot\|$, se le puede asociar una norma matricial $|||\cdot|||$ en $M_n(C)$ definida para $A \in M_n(C)$ por

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

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

EJERCICIO 7. Probar que

$$|||A||| = max\{\|Ax\| : \|x\| = 1\} = \sup\left\{\frac{\|Ax\|}{\|x\|} : x \ne 0\right\} = \sup\{\|Ax\| : \|x\| \le 1\}.$$

Verificar que $|||\cdot|||$ es una norma vectorial. De aquí, $\|Ax\| \le |||A|||.\|x\|$ y $|||A|||$ es la menor constante k tal que, para todo x de C, $\|Ax\| \le k\|x\|$. Veamos que es una norma matricial. Vale $\|(AB)x\| \le |||A|||\|Bx\| \le |||A||||||B|||\|x\|$, luego $|||AB||| \le |||A|||\,|||B|||$, QED.

OBSERVACION. Si $|||\cdot|||$ es una norma inducida e I es la matriz identidad entonces $|||\,I\,||| = 1$. (Aplíquese el ejercicio precedente).

Denotaremos con $|||\cdot|||_p$ a la norma inducida por $\|\cdot\|_p$. ¿Cuales son las normas inducidas por $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$?

TEOREMA 11. Vale:

i) $|||A|||_1 = \sup_j \sum_i |a_{ij}|$ $\quad (= \|A\|_{1,\infty})$,

ii) $|||A|||_\infty = \sup_i \sum_j |a_{ij}|$ $\quad (= \|A^T\|_{1,\infty})$,

iii) $|||A|||_2 = \sup\{\sqrt{\lambda} : \lambda \in \sigma(A^* A)\}$ (a esta norma se la denomina también *norma espectral* de A).■

DEMOSTRACION. *i*). Designemos con $C(A)$ el miembro derecho de *i*) y escribamos $A = [a_1, \ldots, a_n]$ donde cada a_j es una matriz columna. Tenemos, $\|Ax\|_1 = \|x_1 a_1 + \ldots + x_n a_n\|_1 \le \sum |x_j|\,\|a_j\|_1 \le C(A)\|x\|_1$. Luego, $C(A) \ge |||A|||_1$. 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, $\|Ax\|_\infty=\sup_i\left|\sum_j x_j a_{ij}\right|\leq \|x\|_\infty \sup_i \sum_j |a_{ij}|$. Luego, $C(A)\geq |||A|||_\infty$. Para probar la desigualdad contraria, sea k la fila tal que $C(A)=\sum_j |a_{kj}|$ y sea x el vector tal que $x_j=\dfrac{\overline{a_{kj}}}{|a_{kj}|}$ si $a_{kj}\neq 0$, $x_j=1$ en caso contrario. Entonces, $\|x\|_\infty=1$, mientras que

$\|Ax\|_\infty\geq\left|\sum_j a_{kj}x_j\right|=\sum_j |a_{kj}|$. 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, $(Ax)^*Ax=x^*A^*Ax=\lambda x^*x$, o sea $\lambda=\dfrac{\|Ax\|^2}{\|x\|^2}\geq 0$. Luego, el miembro derecho de *iii*) está bien definido. Lo denotamos con $C(A)$. Para U una matriz unitaria fija. (recuérdese que U mantiene la norma y define una transformación inyectiva) valen las igualdades:

$$|||A|||_2^2=\sup\{\|Ax\|_2^2:\|x\|_2=1\}=\sup\{x^*A^*Ax:\|x\|_2=1\}=\sup\{x^*U^*A^*AUx:\|x\|_2=1\}.$$

Usaremos una matriz U tal que $U^*A^*AU=D$=diagonal. Tal U existe en virtud del teorema 4 ya que una matriz triangular y autoadjunta es necesariamente diagonal. Los elementos diagonales de D son los autovalores (no negativos) de A^*A: $\lambda_1,\dots,\lambda_n$. Entonces, $|||A|||_2^2=\sup\{x^*Dx:\|x\|_2=1\}=\sup\left\{\sum_j \lambda_j|x_j|^2:\|x\|_2=1\right\}=C(A)^2$, QED.

EJERCICIO 8. Sean U y V unitarias. Entonces, $|||UAV|||_2=|||A|||_2$.

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

DEMOSTRACION. 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

$|||AX|||=|\lambda|.|||X|||\leq |||A|||.|||X|||$. Como $X\neq 0$, $\rho(A)=|\lambda|\leq |||A|||$, QED.

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

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

8. MATRICES UNITARIAS, HERMITIANAS Y NORMALES.

DEFINICION 12. Dados n vectores $\{x_i : i=1,...,n\} \subset C^n$ se dice que ellos forman una base ortonormal si $x_i^* x_j = \delta_{ij}$.■

DEFINICION 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) Para todo x de C^n vale: $(Ux)^*(Ux)=x^*x$, (o sea U mantiene la longitud euclídea).■

Dejaremos al lector probar las equivalencias de este teorema ya que la mayoría de ellas sigue aplicando la definición. 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 detA $\neq 0$ y $AB=I$ (o sea $B = A^{-1}$).

En efecto, $Ax=0 \Rightarrow BAx=0=x$. 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 $A B_d = I$ y se verifica $B=B(A B_d)=(BA) B_d = B_d$. Luego, $B = A^{-1}$.

EJEMPLO 7. Si $n=2$ las matrices ortogonales se pueden escribir como

$$T(\theta) = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix} \quad \text{o} \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} T(\theta), \qquad \theta \in [0, 2\pi].$$

EJERCICIO 9. Las matricas 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 $U_{k_j} \to U$, que también es unitaria.

2) Si A es unitariamente equivalente a B, esto es, $\exists U$ unitaria tal que $A = U^* BU$, entonces $\sum_{ij} |a_{ij}|^2 = \sum_{ij} |b_{ij}|^2$, es decir, $\|A\|_2 = \|B\|_2$.■

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

DEFINICION 14. Se denomina polinomio característico de A a

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

Vale, $\det(tI - A) = (-1)^n \det(A - tI)$. Observando el desarrollo del determinante se pueden precisar algunos coeficientes de este polinomio:

$$p_A(t)=t^n-tr(A)t^{n-1}+\ldots+(-1)^n\det A,$$

(tr(A)=traza de $A:=\sum a_{ii}$). Si $\lambda_1,\ldots,\lambda_n$ son los ceros de este polinomio, i.e., los autovalores de A, entonces:

$$p_A(t)=(t-\lambda_1)\ldots(t-\lambda_n).$$

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, y por el teorema 4, iguales a la suma de sus autovalores (!)

DEMOSTRACION de 2) del TEOREMA 14. $\sum_{ij}|a_{ij}|^2=\sum_j(\sum_i\overline{a_{ij}}a_{ij})=\operatorname{tr}(A^*A)=$

$$=\operatorname{tr}(U^*B^*UU^*BU)=\operatorname{tr}(U^*B^*BU)=\operatorname{tr}(B^*B)=\sum_{ij}|b_{ij}|^2,$$ QED.

DEFINICION 15. 1) Se dice que $A\in M_n(C)$ es normal si $A^*A=AA^*$.

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

3) A se dice unitariamente diagonalizable si $\exists U$ unitaria tal que $U^*AU=D$=matriz diagonal . En símbolos $A\overset{u}{\approx}D$.

4) A se dice ortogonalmente diagonalizable, en símbolos $A\overset{o}{\approx}D$, si $\exists O$ ortogonal tal que O^TAO=matriz diagonal.■

Vale el siguiente *teorema espectral* para matrices normales:

TEOREMA 15. Las siguientes afirmaciones son equivalentes:

a) A es normal,

b) $A\overset{u}{\approx}D$,

c) $\|A\|_2^2=\sum_{ij}|a_{ij}|^2=\sum_{j=1}^n|\lambda_j|^2$ donde las 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 $A\in M_n(R)$ entonces $A\overset{o}{\approx}D$.■

NOTA. Recordemos que *iii*) del teorema 11 dice que la norma matricial inducida por $\|\cdot\|_2$, $\|\!|\cdot|\!\|_2$, verifica $\|\!|A|\!\|_2^2=\rho(A^*A)$ y A^*A es hermitiana.

EJERCICIO 10. Sea A hermitiana tal que $x^*Ax\geq 0\quad\forall x\in C^n$. Entonces valen:

1) $\sigma(A)\subset[0,\infty)$; 2) tr$A$=0$\Rightarrow A$=0.■

9. PROBLEMAS.

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

1) Demostrar que si $\lfloor a_{ij} \rfloor = A \in M_n(R)$, $\sum_j a_{ij} = 1$ para $i = 1,\dots,n$, 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 $f(t)$ son iguales las sumas de los elementos de las filas de $f(A)$.

2) $A \in M_n(R)$. Si $\lambda \in \sigma(A) \cap R$, $0 \neq x \in C^n$ tal que $Ax = \lambda x$, $x = u + iv$, $u, v \in R^n$, 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 las n variables $x_1,\dots,x_n$ se define como

$$S_k(x_1,\dots,x_n) = \sum_{1 \le i_1 < i_2 < \dots < i_k \le n} \prod_{j=1}^{k} x_{i_j} .$$

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 $A \in M_n(C)$ y $p_A(t) = (t - \lambda_1)\dots(t - \lambda_n)$ el polinomio característico de A. Vale:

$$p_A(t) = t^n - S_1(\lambda_1,\dots,\lambda_n)t^{n-1} + S_2(\lambda_1,\dots,\lambda_n)t^{n-2} - \dots \pm S_n(\lambda_1,\dots,\lambda_n).$$

5) Hay $\binom{n}{k}$ *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 $\lambda_1,\dots,\lambda_n$. 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

$$\frac{1}{t - \lambda_i} = \frac{1}{t} + \frac{\lambda_i}{t^2} + \frac{\lambda_i^2}{t^3} + \dots$$

Por tanto,

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

Vale: $p'(t) = p(t).f(t)$. De esta igualdad sigue que:

$$na_n t^{n-1} + (n-1)a_{n-1}t^{n-2} + \dots + a_1 = (a_n t^n + \dots + a_0)(nt^{-1} + \mu_1 t^{-2} + \mu_2 t^{-3} + \dots).$$

Luego, valen las *identidades de Newton*, si $1 \le k \le n$:

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

y también valen las siguiente igualdades, si $k = 1, 2, \dots$:

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

7) Asumiendo que $\operatorname{tr}(A^k) = \mu_k(A)$ probar que son *equivalentes* las afirmaciones:

i) $\operatorname{tr}(A^k) = \operatorname{tr}(B^k)$ 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, $p_A(t) = p_B(t)$ si y sólo si $\mu_k(A) = \mu_k(B)$, k= 1,2,...).

8) Si $A \in M_n(C)$ 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) Sea T triangular y A unitariamente equivalente a T. Entonces T es normal si y sólo si A lo es.

11) Si T es normal entonces T es diagonal (cf. Teorema 15).

LA MATRIZ POSITIVA Y SU ESPECTRO
PARTE II

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

LEMA 2. 1) $A \in M_n(R)$, $A \geq 0$. Si $\sum_j a_{ij} = c$, independiente de i, entonces $\rho(A) = \|A\|_\infty$, la norma inducida por $\|\cdot\|_\infty$.

2) Si la suma por columnas es independiente de j, $\sum_i a_{ij} = c$, entonces $\rho(A) = \|A\|_1$. ■

DEMOSTRACION. 1) Vale, $A\begin{vmatrix}1\\ \vdots\\ 1\end{vmatrix} = c\begin{vmatrix}1\\ \vdots\\ 1\end{vmatrix}$. Luego, c es autovalor. Entonces $\rho(A) \geq c = \|A\|_\infty$. Como $\|\cdot\|_\infty$ es matricial del teorema 16 obtenemos $\|A\|_\infty \geq \rho(A)$.

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

TEOREMA 17. Sea $A \in M_n(R)$, $A \geq 0$. Entonces:

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

ii) $\inf_j \sum_i a_{ij} \leq \rho(A) \leq \sup_j \sum_i a_{ij}$. ■

DEMOSTRACION. Las cotas superiores resultan aplicando el teorema 16. Sea $\alpha > 0$. Definamos $b_{ij} := \alpha \dfrac{a_{ij}}{\sum_k a_{ik}}$. Entonces $b_{ij} \leq a_{ij}$, o sea, $0 \leq B \leq A$. B verifica la hipOtesis del lema anterior con $c = \alpha$. Luego, $\alpha = \rho(B) \leq \rho(A)$ y se probO *i*). La cota inferior en *ii*) sigue de *i*) usando nuevamente que $\rho(A) = \rho(A^T)$, QED.

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

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

3) Si A es irreducible, $A \geq 0$, entonces $\rho(A) > 0$.■

En efecto, si A es irreducible entonces no puede tener una fila de ceros, QED.

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

i) $\inf_i (\sum_j a_{ij}x_j)/x_i \leq \rho(A) \leq \sup_i (\sum_j a_{ij}x_j)/x_i$,

ii) $\inf_j (\sum_i a_{ij}/x_i)x_j \leq \rho(A) \leq \sup_j (\sum_i a_{ij}/x_i)x_j$.■

DEMOSTRACION. Sea S la matriz diagonal $s_{ii} = x_i$. Entonces S^{-1} es la matriz diagonal $(S^{-1})_{ii} = 1/x_i$. Tenemos, $(S^{-1}AS)_{ij} = \dfrac{a_{ij}x_j}{x_i}$. Como $\rho(S^{-1}AS) = \rho(A)$, obtenemos el teorema 18 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 $\rho(A) < \beta$.■

DEMOSTRACION. Las hipótesis del corolario implican que

$$\alpha \le \min_i(\sum_j a_{ij}x_j)/x_i \le \max_i(\sum_j a_{ij}x_j)/x_i \le \beta$$

y se aplica el teorema 18. Si las desigualdades fueran estrictas entonces también se verificarían con $\alpha+\varepsilon$ y $\beta-\varepsilon$ y tendríamos $\alpha<\alpha+\varepsilon\le\rho(A)\le\beta-\varepsilon<\beta$, QED.

DEFINICION 16. Dada una matriz $A\in M_n(C)$, diremos que λ_1 es un autovalor de *multiplicidad algebraica L*, si $p_A(t)=(t-\lambda_1)^L(t-\lambda_2)\ldots$, $\lambda_1\ne\lambda_2$, $\lambda_1\ne\lambda_3,\ldots$.

Diremos que λ_1 es de *multiplicidad geométrica K*, si *K*=la dimensión del subespacio $\Lambda=\{x: Ax=\lambda_1 x\}$.■

En general no coinciden y vale: *mult geom*(λ_1)≤*mult alg*(λ_1) En efecto, sean $\alpha_1,\ldots,\alpha_K$ vectores linealmente independientes que generan Λ. Sean $\beta_{K+1},\ldots,\beta_n$ elementos de C^n tales que $\{\alpha_1,\ldots,\alpha_K,\beta_{K+1},\ldots,\beta_n\}$ es una base en C^n. Si definimos $S=[\alpha_1\cdots\alpha_K\ \beta_{K+1}\cdots\beta_n]$:=la matriz cuyas columnas son los vectores α_i,β_j, resulta S no singular y $AS=[\lambda_1\alpha_1\cdots\lambda_1\alpha_K\ A\beta_{K+1}\cdots A\beta_n]$. Como $S^{-1}S=I$, tenemos

$B:=S^{-1}AS=\begin{vmatrix}\lambda_1 I & \vdots & * \\ \cdots & \vdots & \cdots \\ 0 & \vdots & C\end{vmatrix}$ donde la matriz identidad $I\in M_K$. Entonces,

$p_A(t)=p_B(t)=(t-\lambda_1)^K\det[tI-C]$. Luego, $K\le L$, QED.

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

DEMOSTRACION. En el corolario al teorema 17 se vió que $\rho(A)>0$. Sea $\lambda\in\sigma(A)$ tal que $|\lambda|=\rho(A)$. Entonces existe x distinto de 0 tal que $Ax=\lambda x$ y vale: $\rho(A)|x|=|\lambda||x|=|\lambda x|=|Ax|\le A|x|$. Definamos $0\le y:=A|x|-\rho(A)|x|$. Supongamos que $0\ne y$. Entonces, $0<Ay=A(A|x|)-\rho(A)A|x|$. Si $w:=A|x|$ entonces *w*>0 verifica $\rho(A)z<Az$. Luego, por el corolario al teorema 18, $\rho(A)<\rho(A)$, absurdo. En consecuencia *y*=0, $|x|$>0 y por tanto el teorema queda probado con *z*=$|x|$, QED.

TEOREMA 20. Sea *A*>0, $\lambda\in\sigma(A)$, $\lambda\ne\rho(A)$. Entonces $|\lambda|<\rho(A)$.■

DEMOSTRACION. Sea $\lambda\in\sigma(A)$ y $\lambda\ne\rho(A)$, con autovector *x*. Supongamos que $|\lambda|=\rho(A)$. En la demostración del teorema 19 se probó que $|x|$ es autovector correspondiente a ρ (*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, existe θ tal que, $x_i=|x_i|e^{i\theta}$. Esto es lo mismo que $x=|x|e^{i\theta}$. Entonces $|x|$ es también autovector para λ lo cual implica que λ es real, o sea, igual a ρ (*A*), absurdo, QED.

TEOREMA 21. $\rho=\rho$ (*A*) tiene multiplicidad geométrica igual a uno.■

DEMOSTRACION. Supongamos que *z,w*, sean autovectores correspondientes al autovalor ρ. Por la demostración anterior sabemos que existen θ,φ tales que:

$$e^{-i\theta}z = p > 0, e^{-i\varphi}w = q > 0.$$

Sea $0<\beta := \min_j \frac{q_j}{p_j}$ y sea $r \in R^n$ definido por $r_j := q_j - \beta p_j$. Entonces, $r \geq 0$ verifica $Ar = \rho r$, de donde se deduce que 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, $w = \beta e^{i(\varphi-\theta)}z$, QED.

Recolectando resultados parciales, y ya demostrados, tenemos,
TEOREMA de Perron (1ª parte) Sea $A \in M_n(C)$, A>0. Vale:
a) $0 < \rho(A) \in \sigma(A)$,
b) $\exists x > 0 : Ax = \rho(A)x$,
c) si $\rho(A) \neq \lambda \in \sigma(A)$ entonces $|\lambda| < \rho(A)$,
d) $\rho\,(A)$ es un autovalor de multiplicidad geométrica uno.■
COROLARIO. Sea A>0. $\exists_1 x_0 > 0$ tal que $Ax_0 = \rho(A)x_0$ y $\sum_j (x_0)_j = 1$.■
A este vector se lo denomina *autovector de Perron*.

Dada A>0, A^T es también positiva. Sea y_0 su autovector de Perron:
$A^T y_0 = \rho(A^T)y_0 = \rho(A)y_0$. Trasponiendo resulta

$$y_0^T A = \rho(A)y_0^T.$$

Se dice entonces que y_0 es el *vector de Perron a izquierda* de A. A $\rho(A)$ se la denomina *la raíz de Perron*.
Damos un criterio para encontrar la raíz de Perron de A.
TEOREMA 22. Sea A positiva. Si $x \geq 0$ es un autovector correspondiente a λ entonces $\lambda = \rho(A)$ y x>0. O sea, x es el autovector de Perron salvo por un factor positivo.■
DEMOSTRACION. 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 el corolario del teorema 18 obtenemos $\lambda = \rho(A)$, QED.

11. **PROBLEMAS**.
1) Sabemos que si $|||\cdot|||$ es una norma matricial y $\lambda \in \sigma(A), A \in M_n(C)$, entonces $|\lambda| \leq |||A|||$. Sea $f(z)$ un polinomio no constante; entonces existe un polinomio mónico $p(z)$ tal que $f(z)=Kz^k p(z)$, $K\neq 0$, $p(z)=z^n + a_{n-1}z^{n-1} + \ldots + a_0$, $a_0 \neq 0$. Obviamente, para acotar las raíces de $f(z)$ basta acotar las de $p(z)$.
La matriz *acompañante* de $p(z)$ se define como:

$$(1) \qquad C=C(p)=\begin{bmatrix} -a_{n-1} & -a_{n-2} & . & . & -a_1 & -a_0 \\ 1 & 0 & . & . & 0 & 0 \\ 0 & 1 & . & . & 0 & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & 1 & 0 \end{bmatrix}.$$

Ella verifica:

$$p_C(z) = p(z)$$

como se ve calculando $\det(zI - C(p))$ por cofactores de la 1ª columna. Luego, $p(z')=0$ implica $|z'| \le \|C(p)\|$ cualquiera sea la norma matricial sobre $M_n(C)$. En particular, si se usa $\|\cdot\| = \|\cdot\|_1$ vale (*cota de Cauchy*),

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

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

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

2) Sean

$$S = \begin{bmatrix} 0 & 0 & . & . & 0 & 0 \\ 1 & 0 & . & . & 0 & 0 \\ 0 & 1 & . & . & 0 & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & 1 & 0 \end{bmatrix}, \qquad R = \begin{bmatrix} -a_{n-1} & -a_{n-2} & . & . & -a_1 & -a_0 \\ 0 & 0 & . & . & 0 & 0 \\ 0 & 0 & . & . & 0 & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & 0 & 0 \end{bmatrix}$$

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

Para demostrar esto último probar primero que

$$\|A\|_2^2 = \|A^*A\|_2 = \|AA^*\|_2$$

usando que $\|B\|_2 = max\{|y^*Bx| : \|y\|_2 = \|x\|_2 = 1\}$.

Entonces,

$$\|C(p)\|_2^2 = \|C^*C\|_2 = \|(S+R)^*(S+R)\|_2 = \|S^*S + R^*R\|_2 \le \|S^*S\|_2 + \|R^*R\|_2.$$

Luego, vale (*cota de Carmichael y Mason*),

(4) $$|z'| \le (1 + |a_0|^2 + \cdots + |a_{n-1}|^2)^{1/2}.$$

3) Sea $q(z) = (z-1)p(z) = z^{n+1} + (a_{n-1} - 1)z^n + (a_{n-2} - a_{n-1})z^{n-1} + \cdots + (a_0 - a_1)z - a_0$.

Como $1 \le |a_0| + |a_1 - a_0| + \cdots + |a_{n-1} - a_{n-2}| + |1 - a_{n-1}|$,

(5) $$|z'| \le |a_0| + |a_1 - a_0| + \cdots |a_{n-1} - a_{n-2}| + |1 - a_{n-1}|.$$

4) **TEOREMA** (Kakeya). Sea $f(z) = a_n z^n + \ldots + a_1 z + a_0$, $a_n \ge a_{n-1} \ge \cdots \ge a_0 \ge 0$; entonces $f(z') = 0 \Rightarrow |z'| \le 1$. ■

(Sugerencia: usar (5)).

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

$$|z'| \ge \frac{|a_0|}{(1 + |a_0|^2 + \cdots + |a_n|^2)^{1/2}}, \quad (\textit{Carmichael y Mason}).$$

12. AUTOVALORES Y AUTOVECTORES.

Supongamos $A^T = A$, $A \in M_n(R)$. Consideremos $x^T A x = \sum_{ij} a_{ij} x_i x_j$ pensado como función numérica sobre $\{ x \in R^n : x^T x = 1 \}$. Como este conjunto es cerrado y acotado, la función contínua $x^T A x$ alcanza su *máximo* λ allí, o sea, $\exists y : y^T A y = \lambda$ donde $y^T y = 1$, mientras $x^T Ax \leq \lambda$ si $x^T x = 1$. O sea, la función $f(x) := x^T Ax - \lambda x^T x \leq 0$ para todo $x \in R^n$ verifica $f(y) = 0$. Además, las derivadas parciales de $f(x)$ respecto de x_j verifican: $\forall\, j \;\; 0 = \frac{\partial f}{\partial x_j}(y) = \frac{\partial}{\partial x_j} \sum \left(a_{ik} x_i x_k - \lambda \sum x_k^2 \right)\Big|_{x=y} = 2\left(\sum_k a_{jk} y_k - \lambda y_j \right)$.

Vale entonces que $Ay = \lambda y$, o sea,

PROPOSICION. λ es autovalor de A e y es uno de sus autovectores. Más aún, λ es el mayor autovalor de A.■

EJERCICIOS. 1) Si $q(t)$ es un polinomio entonces $Ax = \lambda x \;\; x \neq 0 \Rightarrow q(A)x = q(\lambda)x$. Luego, $\sigma(q(A)) \supset \{ q(\lambda) : \lambda \in \sigma(A) \} = q(\sigma(A))$, (ver corol. 1 al teor. de Schur).

2) Sea $A \in M_n(C)$. A es singular si y sólo si $0 \in \sigma(A)$. (sugerencia: det$A$=0 si y sólo si $\exists\, x \neq 0$ tal que Ax=0).

3) Si $\exists A^{-1}$ (es decir, A no es 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 $\lambda = 0$ ò 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 $\sigma(A) \subset R$. (sugerencia: conjugar la identidad $x^* A x = \lambda x^* x$.)

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

8) Si $A \in M_n(R)$ entonces $p_A(t)$ tiene coeficientes reales. Ver que si λ es un autovalor no real, $\bar{\lambda} \in \sigma(A)$.

9) Si $T \in M_n(C)$ es triangular superior, es decir, t_{ij}=0 para $i > j$, entonces $\sigma(T) = \{ t_{ii} : i = 1, \ldots, n \}$, $\det T = \prod_i t_{ii}$.

10) $p_A(t)$ (cf. 7)) verifica $p_A(t) = t^n - (tr\, A)t^{n-1} + \cdots + (-1)^n \det A)$. Como $p_A(t) = (t - \lambda_1)\ldots(t - \lambda_n)$, ver que $tr\, A = \sum_i \lambda_i$, $\det A = \prod_i \lambda_i$.

11) Probar que un autovalor λ es de multiplicidad algebraica k si $p_A(\lambda) = p'_A(\lambda) = \cdots = p_A^{k-1}(\lambda) = 0$, $p_A^{(k)}(\lambda) \neq 0$.

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

13) Si $A \in M_n(C)$ entonces $tr\, A^m = \sum_j \lambda_j^m$,(sugerencia: usar teorema 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. ■

DEMOSTRACION. ($\Rightarrow$). Sea S no singular tal que $S^{-1}AS = D$. Entonces AS=SD. Si escribimos $S = [x_1 \ldots x_n]$, donde cada x_i es un vector columna, tendremos $A x_j = d_{jj} x_j$. Luego, cada x_j es un autovector, siendo la familia $\{x_1, \ldots, x_n\}$ linealmente independiente por ser S no singular.

($\Leftarrow$). Sea $\{x_1, \ldots, x_n\}$ una familia linealmente independiente de autovectores correspondientes a los autovalores $\lambda_1, \cdots, \lambda_n$. Sea S la matriz S:= $[x_1 \cdots x_n]$. S es no singular y vale $AS=[\lambda_1 x_1 \cdots \lambda_n x_n]=SD$, 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*. Es decir, familias de matrices que conmutan dos a dos.

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

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

DEMOSTRACION. Por inducción sobre n. Para $n=1$ es cierto. Supongamos que vale para $n-1$. Sea $\lambda_1\in\sigma(A)$ y $x_1=\begin{vmatrix} x_{11} \\ \vdots \\ x_{n1} \end{vmatrix}$ un autovector de norma 1 (esto es, $\sum_j |x_{j1}|^2=1$).

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 $U_1=[x_1\ z_2 \cdots z_n]$. U_1 es unitaria y vale,

$$U_1^*AU_1=\begin{bmatrix} x_1^* \\ \cdot \\ \cdot \\ z_n^* \end{bmatrix}[\lambda_1 x_1, Az_2,\cdots,Az_n]=\begin{bmatrix} \lambda_1 & * & \cdot & * \\ 0 & & & \\ \cdot & & B & \\ 0 & & & \end{bmatrix},$$

$B\in M_{n-1}(C)$ y $p_A(t)=(t-\lambda_1)p_B(t)$. Entonces, $\sigma(A)=\sigma(B)\cup\{\lambda_1\}$. Por la hipótesis inductiva existe $V\in M_{n-1}(C)$, V unitaria, tal que $V^*BV=T_1$=matriz triangular superior, con elementos de $\sigma(B)$ en la diagonal. *Definamos*,

$$V_1:=\begin{bmatrix} 1 & 0 & . & 0 \\ 0 & & & \\ . & & V & \\ 0 & & & \end{bmatrix}, \quad U:=U_1V_1.$$

Entonces, $U^*AU=V_1^*U_1^*AU_1V_1=$

$$=\begin{bmatrix} 1 & 0 & . & 0 \\ 0 & & & \\ . & & V^* & \\ 0 & & & \end{bmatrix}\begin{bmatrix} \lambda_1 & * & . & * \\ 0 & & & \\ . & & B & \\ 0 & & & \end{bmatrix}\begin{bmatrix} 1 & 0 & . & 0 \\ 0 & & & \\ . & & V & \\ 0 & & & \end{bmatrix}=\begin{bmatrix} \lambda_1 & * & . & * \\ 0 & & & \\ . & & V^*BV & \\ 0 & & & \end{bmatrix},$$

es la matriz T requerida. Si $A \in M_n(R)$ y $\lambda_1 \in R$ entonces existe un autovector real x_1 y en la demostración anterior se pueden tomar $z_2, \cdots, z_n$ reales. Entonces, U_1 y B son matrices reales y por la hipótesis inductiva V tambien lo es. Por tanto, U es real, QED.

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

$$\sigma(q(A)) = \{q(\lambda) : \lambda \in \sigma(A)\} =: q(\sigma(A)).\blacksquare$$

DEMOSTRACION. Si $A \overset{u}{\approx} T$ entonces $q(A) \overset{u}{\approx} q(T)$. Pero $q(T)$ tiene en su diagonal principal a $q(\lambda_1),\ldots,q(\lambda_n)$. Luego, estos son los autovalores de $q(T)$ que por otro lado son los de $q(A)$, QED.

EJERCICIOS. 1) Demostrar que el teorema de Schur vale con T triangular inferior.
2) Demostrar que $\sum_{i<j} t_{ij}^2$ está determinada por A aunque la matriz T del teorema de Schur no está unívocamente determinada.
3) Probar que el producto de matrices triangulares superiores del mismo orden es triangular superior.

Enunciamos sin demostración un complemento del teorema de Schur.

TEOREMA 26. Sea $\mathcal{F} \subset M_n(C)$ una familia conmutante de matrices. Entonces existe U unitaria tal que para toda A en $\mathcal{F}$, U^*AU es triangular superior.■

Estamos ahora en condiciones de demostrar el

TEOREMA de Cayley-Hamilton. Toda matriz $A \in M_n(C)$ satisface su propio polinomio característico. Esto es, $p_A(A) = 0$.■

Probaremos primero un lema.

LEMA 3. Sean R y $T \in M_n(C)$ de la siguiente forma: T es triangular superior verificando $t_{k+1,k+1} = 0$ y

$$R = \begin{bmatrix} 0 & * \\ 0 & T_1 \end{bmatrix},$$

donde T_1 es triangular superior en $M_{n-k}(C)$. Entonces, $R' = RT$ tiene la misma forma que R sólo que la matriz 0 en el ángulo superior izquierdo es de orden $k+1$.■

DEMOSTRACION. 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.

DEMOSTRACION DEL T. DE CAYLEY-HAMILTON. Sabemos por el teorema de Schur que existe U unitaria tal que $U^*AU =$ triangular. Entonces, $p_A(A) = p_A(U^*TU) = = U^*p_A(T)U$ y bastará probar que $p_A(T) = 0$. Pero $p_A(t) = (t - \lambda_1 I)\ldots(t - \lambda_n I)$ implica $p_A(T) = (T - \lambda_1 I)\ldots(T - \lambda_n I) = T_1 \ldots T_n$. Cada T_j es triangular y su elemento $(T_j)_{jj}$ es nulo. Probaremos por inducción que $R_j := T_1 \ldots T_j$ 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 2 (al T. de Schur). $\operatorname{tr} A^k = \sum \lambda_j^k$.■

DEMOSTRACION. Sigue del ejercicio 10, de los que preceden al teorema 23 y del corolario 1 con $q(t)=t^k$, QED.

OBSERVACION. Si A es equivalente a B entonces $\operatorname{tr} A^k = \operatorname{tr} B^k$. Como $\dim M_n(C) = n^2$, las n^2+1 matrices $A^0, A^1, \ldots, A^{n^2}$ son linealmente dependientes. Luego, A^{n^2} es combinación lineal de las potencias A^j, $j = 0,\ldots,n^2-1$. Pero el teorema de Cayley-Hamilton nos dice que A^n se puede escribir como combinación lineal de las potencias A^j, $j = 0,\ldots,n-1$. (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 $<n$.■

DEMOSTRACION. $A^k = A^{k-n}A^n = A^{k-n}q(A) =$ polinomio de grado $<k$. Reiterando, obtenemos un polinomio de grado menor que n, QED.

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

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

Vimos en el 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 puede elegirse unitaria pero se asegura sólo la triangularidad. Caso particular del T.23 es el siguiente

TEOREMA 27. Si $A \in M_n(C)$ tiene n autovalores *distintos* dos a dos entonces es diagonalizable.■

DEMOSTRACION. Sea $\sigma(A) = \{\lambda_1, \ldots \lambda_n\}$ y sean $x^{(1)}, \ldots, x^{(n)}$ los correspondientes autovectores. $\{x^{(1)}, \ldots, x^{(n)}\}$ es un conjunto linealmente independiente. En efecto, si no fuera así habría un $r > 1$ tal que $\{x^{(1)}, \ldots, x^{(r)}\}$ es linealmente dependiente pero $\{x^{(1)}, \ldots, x^{(r-1)}\}$ linealmente independiente. Sean $(\alpha_1, \ldots, \alpha_r) \neq 0$ tal que

$$\alpha_1 x^{(1)} + \cdots + \alpha_r x^{(r)} = 0 .$$

Luego, $\alpha_r \neq 0$ y $(\alpha_1, \ldots, \alpha_{r-1}) \neq 0$. Aplicando A a la igualdad precedente obtenemos

$$\alpha_1 \lambda_1 x^{(1)} + \cdots + \alpha_r \lambda_r x^{(r)} = 0 .$$

Si multiplicamos la primera igualdad por λ_r y lo restamos de esta última resulta $(\lambda_1 - \lambda_r)\alpha_1 x^{(1)} + \cdots + (\lambda_{r-1} - \lambda_r)\alpha_{r-1} x^{(r-1)} = 0$. Esto lleva a una contradicción, QED.

DEFINICION 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, $N_\lambda := \{x : Ax = \lambda x\} = \{$autovectores de A correspondientes a $\lambda\} \cup \{0\}$.

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

EJERCICIO. 1) N_λ es invariante por A.

2) Un subespacio no trivial, invariante por *A*, minimal, es generado por un único autovector de A.

DEFINICION 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.

DEFINICION 19. 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 correspondiente a λ,

ii) *y* es autovector de A^* correspondiente al autovalor $\overline{\lambda}$,

iii) $\bar{y}$ es autovector de A^T correspondiente a λ.

Recordemos la definición de vectores ortogonales en C^n:

DEFINICION 20. $x, y \in C^n$ se dicen ortogonales si $y^*x=<x,y>=0$.

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.■

DEMOSTRACION. $y^*A = \lambda y^*, Ax = \mu x$ implican $y^*Ax = \lambda\langle x,y\rangle = \mu\langle x,y\rangle$. Esto sólo puede valer si $\langle x,y\rangle$=0, QED.

EJERCICIOS sobre el teorema de Cayley-Hamilton. ¿Dónde fallan los siguientes argumentos?

1) $p_A(t) = \det|tI - A|$. Luego, $p_A(A) = \det|AI - A|$=det0=0,

2) $p_A(\lambda) = 0 \; \forall \lambda \in \sigma(A)$. Por el corolario 1 precedente $p_A(A)$ 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 $\varepsilon > 0$ existe una norma matricial $|||\cdot|||$ tal que $\rho(A) \leq |||A||| \leq \rho(A) + \varepsilon$.■

DEMOSTRACION La primera desigualdad ya fue probada en el teorema 12 para cualquier norma matricial.

Por el teorema de Schur, $A=U^*\Delta U$ con Δ triangular, U unitaria. Definamos, para t>0, la matriz Δ_t como la matriz cuyos elementos diagonales son: $t, t^2, \ldots, t^n$ y los restantes nulos. Entonces, $\Delta_t^{-1} = \Delta_{1/t}$ y $A\Delta_t$ es la matriz cuya columna *j*-ésima es la columna *j*-ésima de A multiplicada por t^j para *j*=1,...,*n*, mientras que $\Delta_t A$ es la matriz cuya fila *j*-ésima es la fila *j*-ésima de *A* multiplicada por t^j. Por eso $\Delta_t^{-1}\Delta\Delta_t$ es la matriz tiangular 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 $t \to +0$ podemos elegir t>0 tal que

$$|||\Delta_t^{-1}\Delta\Delta_t|||_1 \leq \sup_j\left(|\lambda_j| + \varepsilon\right) = \rho(A) + \varepsilon.$$

Definamos ahora la siguiente norma matricial:

$$|||B||| := \left\|\Delta_t^{-1} UBU^* \Delta_t\right\|_1 = \left\|S^{-1}BS\right\|_1 \text{ donde } S := U^* \Delta_t \,.$$

Entonces,

$$|||A||| = \left\|\Delta_t^{-1} UAU^* \Delta_t\right\|_1 = \left\|\Delta_t^{-1} \Delta \Delta_t\right\|_1 , \qquad \text{QED.}$$

DEFINICION 21. Diremos que la matriz A es convergente si $A^k \to 0$ elemento a elemento.■

Entonces, A es convergente si y sólo si $\left\|A^k\right\|_\infty \to 0$.

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

DEMOSTRACION Sea $\rho(A)<1$.Por el teorema 29 existe una norma matricial tal que $|||A|||<1$ y vale $|||A^k||| \le |||A|||^k \to 0$. Pero todas las normas son equivalentes, luego $\left\|A^k\right\|_\infty \to 0$. O sea, A es convergente.

Supongamos ahora que $\left\|A^k\right\|_\infty \to 0$. Si $\lambda \in \sigma(A)$ y x es uno de sus autovectores entonces $A^k x = \lambda^k x \to 0$, o sea, $\forall j \ \lambda^k x_j \to 0$. Luego, $|\lambda|<1$ y $\rho(A)<1$, QED.

COROLARIO 1. Sea $A \in M_n(C)$, $\varepsilon > 0$. $\exists C = C(A,\varepsilon)$ tal que $\forall i,j$

$$\left|(A^k)_{ij}\right| \le C(\rho(A)+\varepsilon)^k \,.■$$

DEMOSTRACION Sea $\tilde{A} = \dfrac{A}{\rho(A)+\varepsilon}$. Entonces $\rho(\tilde{A}) = \dfrac{\rho(A)}{\rho(A)+\varepsilon} < 1$. En consecuencia, $\tilde{A}$ es convergente y $(\tilde{A}^k)_{ij} \to 0$ para $k \to \infty, \forall i,j$. Como toda sucesión convergente es acotada existe una constante C que acota a $\left|(\tilde{A}^k)_{ij}\right| = \dfrac{(A^k)_{ij}}{(\rho(A)+\varepsilon)^k}$, QED.

COROLARIO 2. Sea $|||.|||$ una norma matricial. Entonces,

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

DEMOSTRACION Por el teorem de Schur y el teorema 12, $(\rho(A))^k = \rho(A^k) \le |||A^k|||$. Entonces $\rho(A) \le |||A^k|||^{1/k}$. Por otra parte, usando el corolario anterior, $\left\|\tilde{A}^k\right\|_\infty \to 0$ para $k \to \infty$. Luego, $\left\|\tilde{A}^k\right\|_\infty \le C$ y vale por la equivalencia de normas, $|||\tilde{A}^k||| \le \tilde{C}$. Usando la definición de $\tilde{A}$ resulta $|||A^k|||^{1/k} \le (\rho(A)+\varepsilon)\tilde{C}^{1/k}$. Sigue el corolario, QED.

Veremos que el corolario 2 vale en condiciones más generales.

DEFINICION 22. Diremos que una función f sobre C^n es una *prenorma* si f es a valores reales y verifica:

1. $f(x) \ge 0$,
2. si x es no nulo entonces $f(x)>0$ (positividad),
3. $f(\alpha x)=|\alpha| f(x)$ (homogeneidad, que implica $f(0)=0$),
4. $f(x)$ es continua como función sobre $C^n = R^{2n}$.■

TEOREMA 31. Sean f_1, f_2 prenormas. Existen constantes positivas m y M tales que para todo x, $mf_1(x) \leq f_2(x) \leq Mf_1(x)$. ■

DEMOSTRACIÓN. Sea $\Sigma := \{z \in C^n : \|z\|_2 = 1\}$. Σ es un conjunto compacto y por el teorema de Weierstrass la función $h(z) := \dfrac{f_2(z)}{f_1(z)}$, 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 en Σ. La homogeneidad de f_i, i=1,2, implica que $h(az)$=$h(z)$ para a y z no nulos. Y el corolario sigue inmediatamente, QED.

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

$$\rho(A) = \lim_{k\to\infty} \left(f(A^k)\right)^{1/k}. \blacksquare$$

DEMOSTRACIÓN. Sea $f_1(A) := \|\|A\|\|$ una norma matricial. Por el teorema 31, aplicado en C^{n^2}, existen constantes positivas m, M, tales que $m\|\|A^k\|\| \leq f(A^k) \leq M\|\|A^k\|\|$. Osea,

$$(m\|\|A^k\|\|)^{1/k} \leq (f(A^k))^{1/k} \leq (M\|\|A^k\|\|)^{1/k}.$$

Usando el corolario 2 del teorema 30 obtenemos la tesis, QED.

EJERCICIOS.

1) Sea $A \geq 0$, no nula. Si existe y>0 tal que $Ay = \lambda y$ entonces $\lambda = \rho(A)$.

2) Sea $A \geq 0$, no nula. Si existe y>0 tal que $Ay = \lambda y$ entonces (cf. teor. 18 y su corolario),

$$\rho(A) = \max_{x>0} \min_{1\leq i\leq n} \frac{1}{x_i} \sum_{J=1}^{n} a_{ij} x_j = \min_{x>0} \max_{1\leq i\leq n} \frac{1}{x_i} \sum_{J=1}^{n} a_{ij} x_j.$$

3) Sea $A \geq 0$, no nula. Si existe y>0 tal que $Ay = \lambda y$ entonces A es semejante a una matriz no negativa con las sumas por filas iguales.

4) Si A>0 y x es el vector de Perron de A entonces: $\rho(A) = \sum_{ij} a_{ij} x_i x_j$.

14. SERIES DE MATRICES.

DEFINICIÓN 23. Sea $A \in M_n(C)$. A se dice *estrictamente diagonalmente dominante* (más brevemente, $A \in EDD$) si se cumple que $|a_{ii}| > R_i := \sum_{j=1}^{n}{}' |a_{ij}| := \sum_{j\neq i} |a_{ij}|$. ■

TEOREMA 32. (criterio de invertibilidad de Levy-Desplaques). Si $A \in EDD$ 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 $\|\|B\|\|_\infty = \max_i \sum_j |b_{ij}| = \max_i \dfrac{R_i}{|a_{ii}|} < 1$. Veremos en el corolario al teorema 34 que en este caso $(I - B)^{-1}$ existe. Pero A=$D(I$-$B)$. Luego, A^{-1} existe, QED.

Completamos con algunos resultados sobre series de matrices.

TEOREMA 33. Sea $A^{(j)} \in M_n(C)$ $\forall j \in N$ y $\|\cdot\|$ una norma vectorial. Si la serie $\sum \|A^{(h)}\|$ es convergente entonces existe $A = \sum A^{(h)} = \lim_{N\to\infty} \sum_{h=1}^{N} A^{(h)}$ ■.

DEMOSTRACIÓN. Como todas las normas son equivalentes resulta $\|A\|_\infty := \sup\{|a_{ij}| : 1 \le i, j \le n\} \le M\|A\|$ $\forall A \in M_n$. Luego, $\sum_h |a_{ij}^{(h)}| \le M \sum \|A^{(h)}\| < \infty$ $\forall i, j = 1, \ldots, n$, de donde sigue la existencia de $\sum_h a_{ij}^{(h)} = a_{ij}$, QED.

TEOREMA 34. Sea $A \in M_n(C)$, $|||\cdot|||$ una norma matricial, $a_h \ge 0$, $h \in N$. Si $\sum a_h |||A|||^h < \infty$ entonces $\sum a_h A^h$ converge. ■

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

COROLARIO. Sea $A \in M_n(C)$, $|||\cdot|||$ una norma matricial y $|||A||| < 1$. Entonces $I - A$ es invertible siendo $(I - A)^{-1} = I + A + \cdots + A^h + \cdots$ ■

DEMOSTRACIÓN. La serie $\sum |||A|||^h$ es convergente, luego, por el teorema 34, existe $B := \lim_{N\to\infty}(I + A + \cdots + A^N)$. Pero entonces $AB = B - I$, o sea, $(I - A)B = I$, de donde sigue que $B = (I - A)^{-1}$, (ver observación a continuación del teorema 13), QED.

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

2) Sea $A \in M_n(C)$, $|||\cdot|||$ una norma matricial tal que $|||I - A||| < 1$. Entonces A es invertible, siendo $A^{-1} = \sum_{k=0}^{\infty} (I - A)^k$.

LA MATRIZ POSITIVA Y SU ESPECTRO

PARTE III

15. TEOREMA DE PERRON, continuación.

TEOREMA 35 (de Perron, 2ª parte). Sea $A \in M_n(C)$. Vale:

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

DEMOSTRACION. Veamos 1). Recordemos el teorema 22 que afirma que necesariamente un autovector positivo corresponde a $\rho(A)$. Observemos que $L_{ij} = x_i y_j$ implica que $Lx = xy^T x = x(y^T x) = x$. O sea, 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=1,2,... y si $\lambda = \rho(A)$, $AL = Axy^T = \lambda xy^T = \lambda L = x(\lambda y^T) = xy^T A = LA$. Entonces,

$$A^m L = LA^m = \lambda^m L = \rho(A)^m L,$$

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 $0 = \mu L w$, o sea, Lw = 0, resultando $Aw = \mu w$. Hemos probado entonces que $\sigma(A-\rho(A)L)\setminus\{0\} \subset \sigma(A)$.

Veamos que vale además:

$$\sigma(A-\rho(A)L)\setminus\{0\} \subset \sigma(A)\setminus\{\rho(A)\}.$$

En efecto, sea $\lambda = \rho(A)$. Si λ es autovalor de $A - \lambda L$ con autovector w, la demostración anterior nos dice que $\lambda L w = 0$. Nuevamente $\lambda w = (A-\lambda L)w = Aw$ y w es autovector de A correspondiente a $\lambda = \rho(A)$. Por tanto w es múltiplo del vector de Perron x. O sea, $w = \alpha x$. Pero entonces, $Lw = xy^T(\alpha x) = \alpha x \neq 0$, contradicción. 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 $\rho(A)$. Luego, $\rho\left(\frac{A}{\rho(A)} - L\right) < 1$ y por el teorema 30, $\left(\frac{A}{\rho(A)} - L\right)^m \to 0$. Pero, $\left(\frac{A}{\rho(A)}\right)^m - L = \left(\frac{A}{\rho(A)} - L\right)^m \to 0$ y 1) queda demostrada.

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

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

$$L = U^* \begin{matrix} k\left\{\vphantom{\begin{matrix}1\\ \ddots\\ 1\end{matrix}}\right. \\ \\ \\ \\ \end{matrix} \begin{vmatrix} 1 & & & & & \\ & \ddots & & & * & \\ & & 1 & & & \\ \hline & 0 & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{vmatrix} U = U^*\Lambda U .$$

Como U y U^* son matrices de rango n y la matriz Λ tiene rango por lo menos k, resulta que (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 uno y $1 \geq k$, QED.

Enunciamos un teorema que probaremos más adelante utilizando el T. de Gershgorin..
TEOREMA (Ky Fan). Sean A, $B \in M_n(C)$ tales que $B \geq |A|$. Entonces todo autovalor de A yace en la región: $\bigcup\limits_{i=1}^{n}\{z \in C : |z - a_{ii}| \leq \rho(B) - b_{ii}\}$.■

OBSERVACION. Si $A \geq 0$, podemos aplicar este resultado con B=A, resultando

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

Este resultado es más preciso que $\sigma(A) \subset B_{\rho(A)}(0)$.
Mencionamos otro resultado interesante:
TEOREMA 36 (E. Hopf). Si A>0, $\rho(A) \neq \lambda \in \sigma(A)$ entonces

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

DEFINICION 24. Diremos que una matriz $A \in M_n(C)$ pertenece a la clase E de las matrices estocásticas si $A \geq 0$ y $\sum\limits_j a_{ij} = 1$ para i=1,...,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 $[a_{ij}] \in DE$. Toda matriz de permutación P pertenece a DE
TEOREMA 37 (Birkhoff). 1) Sea $A \in M_n(C)$, $A \in DE$. Existen matrices de permutación $P_1, \cdots, P_N$ y constantes $a_1, \cdots, a_N$, $0 < a_j \leq 1$, $\sum a_j = 1$ tales que $A = a_1 P_1 + \cdots + a_N P_N$.
2) $N \leq n^2 - 2n + 2 = (n-1)^2 + 1$.■

Antes de probar el teorema repasaremos algunas nociones sobre *conjuntos convexos*.
Si $x_1, \cdots, x_m \in C^n, \lambda_i > 0, \sum \lambda_i = 1$ y $x = \sum\limits_{i=1}^{m} \lambda_i x_i$ entonces x se dice que es una *combinación convexa* de los x_i. Un conjunto $K \subset C^n$ se dice que es *convexo* si dados $x, y \in K$ y un número a, $0<a<1$, entonces $ax + (1-a)y \in K$. Se puede probar que K es convexo 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 $K_1 \cap K_2$.

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

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

DEFINICION 25. Sea K un convexo cerrado. Un punto $p \in K$ se dice *punto extremal* de K si $p = ax + (1-a)y,\ x, y \in K$, $0<a<1$ 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 $Co(E)=K$.■

DEMOSTRACION del teorema de Birkhoff. El conjunto DE de matrices doblemente estocásticas es un conjunto convexo compacto de 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 $0 \le a_{ij} \le 1$. Entonces, si $A,B \in DE$, $0<a<1$, $P=aA+(1-a)B$, resulta $p_{ij} = aa_{ij} + (1-a)b_{ij}$ 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 entonces no es extremal en DE. A tiene un elemento $a_{i_1 j_1} \in (0,1)$. Como $\sum_k a_{i_1 k} = 1$, hay un $j_2 \ne j_1$ tal que $0 < a_{i_1 j_2} < 1$. Como la suma de la columna j_2 es 1, hay un i_2 distinto de i_1 tal que $0 < a_{i_2 j_2} < 1$. Seguimos así hasta que aparezca 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:

$$i_1 j_1, i_1 j_2, i_2 j_2, \cdots, i_{h-1} j_{h-1}, i_{h-1} j_1,$$

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 $A_+ \ge 0, A_- \ge 0$ y las sumas de sus filas y columnas son iguales a $+1$. Luego, $A_+, A_- \in DE$. Como $A=\frac{1}{2}A_+ + \frac{1}{2}A_-$ sigue que A no es extremal. El teorema de Birkhoff sigue ahora del teorema de Krein-Milman. QED.

16. LOCALIZACION DE LOS AUTOVALORES.

Vamos a demostrar el teorema de Gershgorin que nos dice por donde están localizados los autovalores de una matriz, $A \in M_n(C)$. Para ello enunciamos un teorema sobre la dependencia continua de los ceros de un polinomio respecto de sus coeficientes.

TEOREMA 40. Sean $n \ge 1$ y p un polinomio con ceros λ_i, i=1,...,n,

$$p(x)= a_n x^n + \cdots + a_1 x + a_0,\quad a_n \ne 0.$$

Entonces, para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que, para todo polinomio

$$q(x)= b_n x^n + \cdots + b_1 x + b_0,\quad |b_i - a_i| < \delta,$$

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

COROLARIO. Los autovalores de una matriz $A \in M_n(C)$ dependen continuamente de los elementos de la matriz.■

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

1) Los autovalores de A están localizados en la región

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

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

DEMOSTRACION. 1) Sea λ un autovalor de A y $x \neq 0$ tal que $Ax = \lambda x$. Sea p tal que $|x_p| = \text{máx } \{|x_i| : i = 1, \ldots, n\}$. Entonces, $\lambda x_p = \sum_j a_{pj} x_j$ implica que

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

Usando la desigualdad triangular y la definición de p: $|\lambda - a_{pp}||x_p| \leq R'_p(A)|x_p|$. O sea, para algún p, $|\lambda - a_{pp}| \leq R'_p(A)$. 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 $d_{ii} = a_{ii}$, $i = 1, \ldots, n$ y B la matriz definida por $A = D + B$. Definamos, $A_\varepsilon := D + \varepsilon B$, $0 \leq \varepsilon \leq 1$. Tenemos $R'_j(A_\varepsilon) = \varepsilon R'_j(A)$. Obsérvese que los discos que forman la región de Gershgorin de A_ε tienen el mismo centro que los de $G(A)$ y radio menor. Luego, k discos de los que forman a $G(A_\varepsilon)$ 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 $\lambda_i(A_\varepsilon)$ sea una función continua de ε para cada $i = 1, \ldots, n$, sobre el intervalo $[0,1]$. Pero $A_0 = D$ y podemos suponer que $\lambda_i(A_0) = a_{ii}$.

Luego, si $\lambda_i(A_0) \in G_k$ entonces $\lambda_i(A_\varepsilon)$ pertenece a G_k para todo ε, $0 \leq \varepsilon \leq 1$. En consecuencia, para $\varepsilon = 0$ hay exactamente k raíces en G_k. Pero entonces lo mismo vale para todo ε, en particular para $\varepsilon = 1$. En efecto, el número de autovalores de A_ε en G_k es una función continua de ε (corolario del teorema 40), pero siendo a valores enteros debe ser constante pues su dominio es conexo, QED.

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

Observemos que A y $S^{-1}AS$ tienen los mismos autovalores aunque sus regiones de Gershgorin son en general diferentes. Por ejemplo, si S es la matriz diagonal $s_{ii} = p_i > 0$ se obtiene que $(S^{-1}AS)_{ij} = p_j a_{ij} / p_i$. Aplicando el teorema de Gershgorin a esta matriz y a su traspuesta, se llega al siguiente

COROLARIO. a) $\sigma(A) \subset G(S^{-1}AS) = \bigcup_{i=1}^{n} \left\{ |z - a_{ii}| \le \frac{1}{p_i} \sum_{k \ne i} p_k |a_{ik}| \right\}$

b) $\sigma(A) \subset G((S^{-1}AS)^T) = \bigcup_{i=1}^{n} \left\{ |z - a_{ii}| \le p_i \sum_{k \ne i} |a_{ki}| / p_k \right\}$. ■

EJEMPLO. Sea $A=\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$. Entonces $\sigma(A)=\{1,2\}$. Si aplicamos el teorema de Gershgorin tenemos que $G(A)$ es el disco de radio 1 y centro en 1. Sin embargo usando el corolario obtenemos una acotación que puede ser mejor:

$$\sigma(A) \subset \left\{ z : |z-1| \le \frac{p_2}{p_1} \right\} \cup \{2\}.$$

Veamos la siguiente aplicación:

TEOREMA 42 (Ky-Fan). Sea $A \in M_n(C)$, $B \ge |A|$. Entonces

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

DEMOSTRACION. 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 :

$$\sum_{j \ne i} |a_{ij}| x_j \le \sum_{j \ne i} |b_{ij}| x_j = (Bx)_i - b_{ii} x_i = (\rho(B) - b_{ii}) x_i .$$

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

2) El caso $B \ge |A|$ se puede reducir a este caso considerando la matriz positiva $B_\varepsilon := B + \varepsilon I$, que también verifica $B_\varepsilon \ge A$. Por 1)

$$\sigma(A) \subset \bigcup_{I=1}^{n} \{ |z - a_{ii}| \le \rho(B_\varepsilon) - (b_{ii} + \varepsilon) \}.$$

Pero para $\varepsilon \to 0$, $\rho(B_\varepsilon) - (b_{ii} + \varepsilon) \to \rho(B) - b_{ii}$, y sigue la tesis, QED.

17. MATRICES NO NEGATIVAS, MATRICES IRREDUCIBLES Y MATRICES PRIMITIVAS. TEOREMA DE PERRON-FROBENIUS.

Recordemos el

TEOREMA 43 (de Perron). Si $A \in M_n(C)$, $A > 0$, 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 $\lambda \in \sigma(A)$ y $\lambda \ne \rho(A)$, entonces $|\lambda| < \rho(A)$,

f) $\left[\frac{A}{\rho(A)} \right]^m \xrightarrow[m \to \infty]{} L$, $L := x y^T$, donde $Ax = \rho(A)x$, $A^T y = \rho(A) y$,

$x > 0$, $y > 0$ $x^T y = 1$. ■

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

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), c) y d). Comencemos con a) y c).
TEOREMA 44. *i*) Si $A \in M_n(C)$, $A \geq 0$ entonces $\rho(A) \in \sigma(A)$ y $\exists\, x \geq 0$, $x \neq 0$, 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 $Ax = \rho(A)x$.■
DEMOSTRACION. *i*) Lo probaremos aproximando A por la matriz positiva $A(\varepsilon) := A + \varepsilon \Xi$ donde Ξ tiene todos sus elementos iguales a 1 y $\varepsilon > 0$, a la cual se aplica el teorema de Perron. Sea $x(\varepsilon)$ el vector de Perron de $A(\varepsilon)$. Entonces, $x(\varepsilon) > 0$, $\sum x(\varepsilon)_i = \|x(\varepsilon)\|_1 = 1$. El conjunto $\{ x : \|x\|_1 = 1 \}$ es compacto en C^n. Podemos entonces encontrar una sucesión $\{\varepsilon_k\} \downarrow 0$ tal que $x(\varepsilon_k)_i \to x_i \geq 0$ para todo i y $k \to \infty$. Pero $\sum x_i = \lim_{k \to \infty} \|x(\varepsilon_k)\|_1 = 1$. Por tanto $x \neq 0$. Por otra parte, por el teorema 2,

$$\rho(A) \leq \rho(A(\varepsilon_k)) \leq \rho(A(\varepsilon_{k-1})).$$

Luego, la sucesión $\rho(A(\varepsilon_k))$ 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, por el teorema 22, ρ^k es su raíz de Perron y x es un múltiplo de su autovector de Perron. En consecuencia, $\rho > 0$ y $x > 0$, QED.
EJERCICIO. Sabemos por el teorema 18 que si $A \geq 0$, $x>0$ y $Ax \leq \beta x$ entonces $\rho(A) \leq \beta$. Usando $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ y $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ probar que el resultado no vale para $x \geq 0$, $x \neq 0$. Mostrar también que $\rho(A) \neq \min_{x>0} \max_i \frac{1}{x_i} \sum_{j=1}^n a_{ij} x_j$.
Complementamos a continuación el teorema 44.
TEOREMA 45. Sea $A \in M_n(C)$, $A \geq 0, A^k > 0$. Entonces, $\rho(A)$ es algebraicamente simple y todo $\lambda \in \sigma(A) \setminus \{\rho(A)\}$ verifica $|\lambda| < \rho(A)$.■
DEMOSTRACION. Sabemos que $\sigma(A^k) = \{\lambda^k : \lambda \in \sigma(A)\}$. En consecuencia, si $\rho = \rho(A)$ fuera múltiple para A, ρ^k sería múltiple para A^k, contradiciendo d) del teorema 43 de Perron. El resto sigue de e) de ese teorema, QED.

NB. Observemos que una matriz A que verifica $A \geq 0$, $A^k > 0$, tiene la propiedad *FC*. Luego, por el teorema 6, es irreducible. La recíproca no es cierta: $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ es irreducible pero ninguna de sus potencias es positiva.
Caracterizamos a continuación esa clase de matrices irreducibles.
DEFINICION 26. $A \geq 0$ se dice *primitiva* si es irreducible y posee un solo autovalor de módulo máximo.■

TEOREMA 46. La matriz $A \geq 0$ es primitiva si y sólo si existe un número natural k>0 tal que $A^k > 0$.■

DEMOSTRACION. $\Leftarrow$. Es el contenido del teorema 45 (ver NB a continuación del mismo).

$\Rightarrow$. Repasemos la demostración de 1) teorema 35. O sea, del teorema de Perron 2ª parte: $\lim_{m\to\infty}\left[\frac{A}{\rho(A)}\right]^m = L := x\,y^T$. En dicha demostración sólo se usan los siguientes hechos: 1) $\rho(A)$>0 es un autovalor 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 $\rho(A)$. En el próximo teorema de Perron-Frobenius se verá que 1)-3) valen para toda matriz irreducible no negativa, en particular para matrices primitivas. Aceptando esto podemos repetir la demostración mencionada y concluir que $\lim_{m\to\infty}\left[\frac{A}{\rho(A)}\right]^m = L := x\,y^T$ donde $x > 0$ e $y > 0$ implican $L > 0$. Entonces para m suficientemente grande $A^m > 0$, QED.

Probaremos ahora los resultados del teorema de Perron que se mantienen válidos 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 $A\,x = \rho(A)\,x$,

d) $\rho(A)$ es algebraicamente simple.■

DEMOSTRACION. a) La hipótesis implica que $\alpha := \min_i \sum_j a_{ij} > 0$. Por teorema 17, i), resulta $\rho(A) \geq \alpha > 0$.

b) Ya lo demostramos en i) del teorema 44.

c) $\exists\, x \geq 0$, $x \neq 0$, tal que $Ax = \rho(A)x$, (cf. i) T.44). Luego, $(I + A)^{n-1}x = (1 + \rho(A))^{n-1}x$. Pero por c) del teorema 6, la matriz $(I + A)^{n-1}$ es positiva y por tanto $(I + A)^{n-1}x > 0$. En consecuencia, x>0.

Para probar d) necesitamos un lema.

LEMA. Si $A \in M_n(C)$ y $\lambda \in \sigma(A)$ tiene multiplicidad algebraica k entonces $1 + \lambda \in \sigma(I + A)$ con multiplicidad algebraica k. Además, $\rho(I + A) \leq 1 + \rho(A)$. Si $A \geq 0$ entonces $\rho(I + A) = 1 + \rho(A)$.■

DEMOSTRACION DEL LEMA. Observemos que

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

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 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 $\rho(A)$ fuera de multiplicidad algebraica k entonces $1+\rho(A)$ sería de multiplicidad algebraica k para $I+A$. Pero $I+A$ es positiva y $1+\rho(A)$ es su autovalor de Perron. Por teorema 35, 2), $k=1$, QED.

Agregamos a continuación un resultado, e) del siguiente teorema, que complementa al teorema de Perron-Frobenius en el caso de matrices primitivas (v. Def. 26).
TEOREMA 48. Si $A \geq 0$ es una matriz primitiva entonces valen:
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,
e) $\lim_{m\to\infty}\left[\frac{A}{\rho(A)}\right]^m = x\, y^T =: L$ donde $x > 0$ es autovector de A para $\rho(A)$, $y > 0$ es autovector de A^T para $\rho(A)$ tal que $y^T x = 1$.■
DEMOSTRACION. Para los puntos a)-d) ver teorema 47 y definición 26. Para e) recordar que ya notamos su validez en la demostración del teorema 46, QED.

18. EXTENSION DE ALGUNOS RESULTADOS A MATRICES NO NEGATIVAS. AUTOVALORES DEL MISMO MODULO.

Sea $A \in M_n(R), A \geq 0, \alpha \in R, \alpha \geq 0$ (v. Corol. del Teor. 18).
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 $Ax = \rho(A) x$.■
DEMOSTRACION. *i*) Sea $A(\varepsilon) := A + \varepsilon \Xi$, (definición en el teorema 44) . 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 $A(\varepsilon)x - \alpha\, x > A\, x - \alpha\, x \geq 0$. Multiplicando por $y(\varepsilon)^T$ resulta:
$$0 < y(\varepsilon)^T (A(\varepsilon) - \alpha) x = (\rho(A(\varepsilon)) - \alpha) y(\varepsilon)^T x.$$
Luego, $\rho(A(\varepsilon)) - \alpha > 0$. Como los autovalores de $A(\varepsilon)$ convergen a los de A resulta que $\rho(A(\varepsilon)) \to \rho(A)$ e *i*) queda probado.
ii) Por hipótesis, $A\, x - \rho(A) x \geq 0$. Si fuera $\neq 0$, multiplicando por y^T resultaría $y^T(A\, x - \rho(A) x) > 0$. Pero $y^T(A\, x - \rho(A) x) = (\rho(A) - \rho(A)) y^T\, x = 0$, 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).■

DEMOSTRACION. 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, $\alpha\, x_i \leq \sum_j a_{ij} x_j \quad \forall i$, o sea, $\alpha\, x \leq A\, x$. Por el teorema anterior $\alpha \leq \rho(A)$. Si $A \geq 0$, existe $x \geq 0$, $x \neq 0$, con $A\, x = \rho(A) x$, (cf. teorema 44). Entonces el máximo de los α se alcanza sobre ese x y es $\rho(A)$, QED.
NB. Compárese el Teor. 50 con el ejercicio a continuación del Teor. 44.

Recordemos que el teorema 2 afirma que $|C| < D \Rightarrow \rho(C) \le \rho(|C|) \le \rho(D)$. Veremos que pasa si esas desigualdades son igualdades.

TEOREMA 51. Sean $A, B \in M_n(C)$, $A \ge 0$, irreducible. Supongamos que $|B| \le A$ y $\rho(A) = \rho(B)$. Si $\lambda = e^{i\phi}\rho(B) \in \sigma(B)$ entonces existen números reales $\theta_1, \ldots, \theta_n$ tales que $B = e^{i\phi}DAD^{-1}$ con D matriz diagonal, $D = \begin{bmatrix} e^{i\theta_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\theta_n} \end{bmatrix}$. O sea, $|A| = |B|$ y

$$B = [b_{kl}] = e^{i\phi}\left[e^{i(\theta_k - \theta_l)}a_{kl}\right]. \blacksquare$$

DEMOSTRACION. Existe $x \ne 0$ tal que $Bx = \lambda x$. Entonces, $|Bx| = \rho(A)|x| \le |B||x| \le$ $\le A|x|$. 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 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,

$$\rho(A)|x| = e^{-i\phi}D^{-1}BD|x| = \tilde{B}|x|. \quad \tilde{B} := e^{-i\phi}D^{-1}BD.$$

Observando la matriz $\tilde{B} = e^{-i\phi}D^{-1}BD$ vemos que $|B| = |\tilde{B}|$ y las identidades anteriores muestran que $\rho(A)|x| = A|x| = \tilde{B}|x|$, $|\tilde{B}| \le A$. Entonces tenemos,

$$\sum_j a_{ij}|x_j| = \sum_j \tilde{b}_{ij}|x_j| = \left|\sum_j \tilde{b}_{ij}|x_j|\right| \quad \forall i.$$

La última igualdad implica que $\tilde{b}_{ij}|x_j|$ es real y no negativo para cada *ij*. Luego, $\tilde{B} \ge 0$. La primera igualdad junto con $|x| > 0$ y $\tilde{B} \le A$ implican $\tilde{B} = A$, QED.

COROLARIO. Sea $A \ge 0$, irreducible y $\rho(A) \ne \lambda \in \sigma(A)$ tal que $|\lambda| = \rho(A)$. Entonces, $A = e^{i\Phi}DAD^{-1}$ con D como en el teorema y $e^{i\Phi} = \dfrac{\lambda}{\rho(A)}$. O sea,

$$a_{kl} \ne 0 \Rightarrow \frac{\lambda e^{i(\theta_k - \theta_l)}}{\rho(A)} = 1, \ \Phi = \theta_l - \theta_k. \blacksquare$$

TEOREMA 52. Sea $A \ge 0$, irreducible. Vale

i) Si $S := \{\lambda_n = \rho(A), \lambda_{n-1}, \cdots, \lambda_{n-k+1}\}$ es la familia de autovalores (distintos) de A de módulo $\rho(A)$ entonces cada λ_j es de multiplicidad algebraica uno y $S = \{e^{2\pi ip/k}\rho(A) : p = 0, \ldots, k-1\}$.

ii) Si $\lambda \in \sigma(A)$ entonces $\lambda e^{2\pi ip/k} \in \sigma(A)$ para $p = 0, \ldots, k-1$ y con la misma multiplicidad.$\blacksquare$

DEMOSTRACION. Notemos el siguiente hecho. Sea $\lambda_{n-u} = e^{i\phi_u}\rho(A) \in S$. Aplicando el corolario al teorema 51 resulta $A = e^{i\phi_u}DAD^{-1}$. Entonces $p_A(t)$ es igual, salvo por un factor constante, a $p_{DAD^{-1}}(e^{-i\phi_u}t) = p_A(e^{-i\phi_u}t)$. Luego,

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

En particular, $e^{i\phi_u}S = S$. Esto implica que

(^) dados $u,r\in\{0,...,k-1\}$ existe un $q\in\{0,\ldots,k-1\}$ tal que $e^{i\phi_u}\lambda_{n-q}=\lambda_{n-r}=e^{i\phi_r}\rho(A)$. Entonces, $\phi_u+\phi_q=\phi_r(\text{mod }2\pi)$. Esto prueba que $\{\phi_0=0,\phi_1,\ldots,\phi_{k-1}\}=:G$ es un grupo abeliano con la operación suma $(\text{mod }2\pi)$.
Supongamos los ϕ ordenados: $0=\phi_0<\phi_1<\cdots<\phi_{k-1}<2\pi$. Entonces, $G_1:=\{h\phi_1 : h\in Z\}$ $(\text{mod }2\pi)$ es también un grupo abeliano con la operación suma (módulo 2π) y $G_1\subset G$. Si estos dos grupos no fueran coincidentes habría un ϕ_q tal que $h\phi_1<\phi_q<(h+1)\phi_1$ y en consecuencia $0<\phi_q-h\phi_1<\phi_1$. Pero esto es un absurdo pues $\phi_q-h\phi_1\in G$ y es menor que ϕ_1. Luego $G=G_1$, lo que prueba *i*).
Para demostrar *ii*) basta aplicar (*), QED.
COROLARIO 1. Sea $A\geq 0$, irreducible. Entonces k es divisor de $\#\{\lambda\in\sigma(A):\lambda\neq 0\}$. Si A es no singular entonces k es divisor de n.■
COROLARIO 2. Sea $A\geq 0$, irreducible. Sea $A^m=\left[a_{ij}^{(m)}\right]$. Si hay $k\geq 2$ autovalores de módulo máximo, entonces, $\forall\, i=1,\ldots,n$, si m no es múltiplo de k, $a_{ii}^{(m)}=0$. En particular, $a_{ii}=0$. ■
DEMOSTRACION. Sea $\phi=2\pi/k$, $\lambda=e^{i\phi}\rho(A)$. $A=e^{i\phi}DAD^{-1}$ implica $A^m=e^{im\phi}DA^mD^{-1}$ 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.
EJERCICIO. 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 $A\in M_n(R)$, $A\geq 0$, $\rho(A)>0$. Si $\lambda\in\sigma(A)$ y $|\lambda|=\rho(A)$ 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 $p=0,\ldots,k-1$ vale $e^{ip\phi}\rho(A)\in\sigma(A)$. ■
NB. Pero puede ocurrir que no sean estos últimos los únicos autovalores de módulo máximo.
TEOREMA 54. i) Sea $A\geq 0$, irreducible. Sean $x>0$, $y>0$ tales que $x^Ty=1$ $Ax=\rho(A)x$, $A^Ty=\rho(A)y$. Entonces (convergencia en media Cèsaro),

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

ii) Existe $C=C(A)>0$ tal que

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

iii) Si además A es primitiva y $1>r>|\lambda|/\rho(A), \forall\lambda\in\sigma(A), \lambda\neq\rho(A)$, entonces existe $C=C(A,r)$ tal que $\|(A/\rho(A))^m-L\|_\infty\leq Cr^m, m=1,2,\cdots$.■

El siguiente teorema debido a Wielandt precisa la caracterización dada en el teorema 46.

TEOREMA 55. Sea $A \in M_n(R)$, $A \geq 0$. A es primitiva si y sólo si $A^{n^2-2n+2} > 0$.■

El exponente, en general, no puede disminuirse. Pero vale el siguiente resultado,

TEOREMA 56. Sea $A \in M_n(R)$, $A \geq 0$, irreducible. Si todos los elementos de la diagonal de A son positivos entonces $A^{n-1} > 0$ (y A es primitiva).■

DEMOSTRACION. Sea $a = min\{a_{11},..a_{nn}\}$ y D la matriz diagonal cuyos elementos diagonales coinciden con los de A. Sea $B = A - D \geq 0$. Por ser A irreducible también lo es B. Luego, $A \geq aI + B = a\left(I + \frac{B}{a}\right)$. Por tanto, $A^{n-1} \geq a^{n-1}\left(I + \frac{B}{a}\right)^{n-1} > 0$, QED.

TEOREMA 57. Sea $A \in M_n(R)$, $A \geq 0$, irreducible. Sea $\{P_j\}$ la familia de nodos de $\Gamma(A)$. Denotamos con $L_j = \{k_1(j), k_2(j), \ldots\}$ al conjunto de las longitudes de todos los caminos dirigidos que comienzan y terminan en P_j, $j = 1, \ldots, n$ (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.■

DEMOSTRACION. $L_j \neq \emptyset$ pues A es irreducible. Supongamos que A es primitiva. Entonces, $A^m > 0$ para cierto m. A no tiene ninguna columna de ceros pues A^T es irreducible. Por tanto, $A^{m+k} > 0$ $\forall k > 0$. En consecuencia, $L_j \supset \{m, m+1, \ldots\}$ lo cual implica que $g_j = 1$.

Si por el contrario A no es primitiva entonces A tiene $k \geq 2$ autovalores de módulo máximo. Por el corolario 2 del teorema 52, si m no es un múltiplo de k entonces $a_{ii}^{(m)} = 0$. Luego, $L_j \subset \{k, 2k, 3k, \ldots\}$ para cada $j = 1, \ldots, n$. En consecuencia, g_j es divisible por k. Luego, $g_j \geq k > 1$, QED.

COROLARIO. Sea $A \in M_n(R)$, $A \geq 0$, irreducible. Si A no es primitiva entonces, para todo j, $g_j > 1$ ■

NB. Un teorema de Romanovsky afirma que si $A \geq 0$ es irreducible entonces los g_j son todos iguales al número de autovalores de A de módulo máximo.

DEFINICION 27. Sea $A \in M_n(R)$, $A \geq 0$, primitiva. $\gamma(A)$ representa al índice de primitividad de A que es el menor k tal que $A^k > 0$.■

TEOREMA 58. Sean $A \in M_n(R)$, $A \geq 0$, primitiva y s la longitud del ciclo más corto en $\Gamma(A)$. Entonces $\gamma(A) \leq n + s(n-2)$.■

Aceptaremos este resultado para presentar una prueba del teorema de Wielandt.

DEMOSTRACION DEL TEOREMA 55. Basta ver que $\gamma(A) \leq n^2 - 2n + 2$ y esto para n>1. Como A es irreducible hay ciclos en $\Gamma(A)$. Si el más corto tuviera longitud s y s=n se presentaría una contradicción con el teorema 57. Entonces s<n y tenemos

$$\gamma(A) \leq n + s(n-2) \leq n + (n-1)(n-2) = n^2 - 2n + 2, \quad \text{QED.}$$

Si algún elemento de la diagonal de A fuera positivo entonces s=1 y tendríamos $\gamma(A) \leq 2n - 2$. Este resultado se mejora con el siguiente que implica al teorema 56.

TEOREMA 59. Sea $A \in M_n(R)$, $A \geq 0$, irreducible. Si A posee d elementos diagonales positivos entonces $\gamma(A) \leq 2n - d - 1$.■

LA MATRIZ POSITIVA Y SU ESPECTRO
PARTE IV

19. FORMAS CANONICAS. Supongamos dada una familia de entes en la que hay definida una relación de equivalencia. Diremos que tenemos una *forma canónica* asociada si en cada clase de equivalencia es posible elegir un (o esencialmente un) objeto. La idea es que a los fines que interesan debería bastar con utilizar la forma canónica equivalente al elemento en estudio de la que se espera sea más manejable. En nuestro caso la familia es $M_n(C)$, o bien $M_n(R)$, y la equivalencia se define por la semejanza entre matrices: $A \approx B \Leftrightarrow \exists S : \det S \neq 0 \, y \, B = SAS^{-1}$. Una forma canónica en esta situación es la llamada *forma de Jordan* que analizaremos. Si la relación $\propto$ se definiera por la posesión del mismo polinomio carácterístico nos encontraríamos con que $A \approx B \Rightarrow A \propto B$, aunque la recíproca no es cierta. Si la relación fuera $\approx$, en cada clase de equivalencia tendríamos demasiadas matrices triangulares superiores para utilizar como representante de la clase (cf. ejemplo 3) siguiente). La forma de Jordan es una de ellas pero es esencialmente única. Si bien las matrices triangulares superiores tienen demasiados elementos, $n(n+1)/2$, para distinguir la semejanza, las matrices diagonales, que son de Jordan, no tienen los necesarios para hacerlo. En efecto,
PROPOSICION. Una matriz de Jordan no es necesariamente diagonalizable.■

DEFINICION 28. i) Un bloque de Jordan $J_k(\lambda)$ es una matriz triangular superior de orden k de la forma, $J_1(\lambda) = [\lambda]$,

(1)
$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & . & . & 0 \\ 0 & \lambda & 1 & 0 & . & 0 \\ 0 & 0 & \lambda & 1 & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & . & \lambda & 1 \\ 0 & 0 & 0 & . & . & \lambda \end{bmatrix}.$$

ii) Una matriz J de Jordan de orden n es de la forma,

(2)
$$J = \begin{bmatrix} J_{n_1}(\lambda_1) & 0 & . & 0 \\ 0 & J_{n_2}(\lambda_2) & . & 0 \\ . & . & . & . \\ 0 & 0 & . & J_{n_h}(\lambda_k) \end{bmatrix}, \qquad n_1 + n_2 + \ldots + n_h = n.$$

iii) Una matriz A se dice *no derogatoria* si todos sus autovalores tienen multiplicidad geométrica uno.■
Desafortunadamente la forma de Jordan no es estable para el cálculo, lo mismo que el rango de una matriz, pero si se la conoce se tiene mucha información sobre la matriz dada originalmente.
J tiene $n-h$ unos. Sus autovalores $\lambda_1, \lambda_2, \cdots$ no son necesariamente dos a dos distintos.
Veamos la proposición. Sea $J = J_k(\lambda)$ un bloque de Jordan de orden k mayor que uno. Si fuera diagonalizable y semejante a la matriz diagonal D tendríamos $J = SDS^{-1}$ y $D = \lambda I$. Luego, $0 \neq J - D = S(\lambda I)S^{-1} - D = \lambda I - D = 0$.
Si e_i representa al versor columna que tiene un 1 en el lugar i vemos que $J_n(\lambda)$ tiene un solo autoversor: e_1. O sea, es no derogatoria. Y la matriz J en ii), tiene k autovectores

linealmente independientes de multiplicidad 1 correspondientes a $\lambda_1, \lambda_2, \cdots, \lambda_h$: $e_1, e_{n_1+1}, e_{n_1+n_2+1}, \cdots$. Es decir, hay sólo un autovector por bloque. En efecto,

$$(3) \qquad J_n(\lambda)x = (\lambda I + J_n(0))x = \lambda x + J_n(0)\begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{bmatrix} + \begin{bmatrix} x_2 \\ \cdot \\ x_n \\ 0 \end{bmatrix}.$$

O sea, x es un autovector de $J_n(\lambda)$ para el autovalor λ si es un múltiplo de e_1. Por otra parte, e_1 es el único autovector para ese autovalor salvo por un factor no nulo. Obsérvese que $(J_n(0))^n = 0$ y que $p_{J_n(\lambda)}(t) = (t-\lambda)^n$.

EJEMPLOS. 1) Sea $A(\varepsilon) = \begin{bmatrix} \varepsilon & 0 \\ 1 & 0 \end{bmatrix}$. Su matriz de Jordan es $J_\varepsilon = \begin{bmatrix} 0 & 0 \\ 0 & \varepsilon \end{bmatrix}$ si $\varepsilon \neq 0$. Pero la matriz de Jordan de $A(0)$ no es J_0.

2) $B(\varepsilon) = \begin{bmatrix} \varepsilon & 0 \\ 0 & 0 \end{bmatrix}$ tiene rango 1 para todo $\varepsilon \neq 0$ pero su límite, $B(0)$, tiene rango 0.

3) $V := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. $V \approx \begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix}, a > 0$. No existe C tal que $C^2 = V$. En efecto, $\sigma(C) = \{0\}$ y la forma de Jordan, J, de C tiene ceros en la diagonal, por lo que $C^2 = SJ^2S^{-1} = 0$.

El siguiente resultado mejora al teorema de Schur.

TEOREMA 60. Sea $A \in M_n(C)$ con autovalores distintos $\lambda_1, \cdots \lambda_k$ de multiplicidades $n_1, \cdots, n_k$. Entonces A es semejante, $A = STS^{-1}$, a una matriz triangular superior T de la forma

$$(4) \qquad T = \begin{bmatrix} T_1 & 0 & . & 0 \\ 0 & T_2 & . & 0 \\ . & . & . & . \\ 0 & 0 & 0 & T_k \end{bmatrix}$$

donde $T_i \in M_{n_i}$ es triangular superior con sus elementos diagonales iguales a λ_i.

Si A es real lo mismo que sus autovalores entonces S puede elegirse real.■

DEMOSTRACION. Usando el teorema de Schur podemos suponer $A=T=T^{(1)}$, triangular superior, y los $\lambda_1, \cdots \lambda_k$ ordenados en la diagonal de A, agrupados los iguales. Dado el índice rs, sea E la matriz cuyo elemento rs sea 1 y los restantes nulos. Supongamos $r < s$. Como $(I-aE)(I+aE)=I$ resulta $I - aE = (I+aE)^{-1}$. Entonces $T^{(2)} := (I-aE)T^{(1)}(I+aE)$ tiene los mismos elementos que $T^{(1)}$ salvo eventualmente los de índice $rs, r(s+1), \cdots, (r-1)s, (r-2)s, \cdots$. Además, $t_{rs}^{(2)} = t_{rs} + a(t_{rr} - t_{ss})$. Luego, si $t_{rr} \neq t_{ss}$, podemos elegir a tal que $t_{rs}^{(2)} = 0$. Repitiendo este procedimiento se llega en un paso j a una matriz $T^{(j)}$ de la forma (4), QED.

TEOREMA 61. Una matriz triangular superior con elementos diagonales iguales a λ es semejante a una suma directa de bloques de Jordan con elementos diagonales iguales a λ. Es decir, a una matriz de Jordan J cuyos elementos diagonales son todos iguales a λ.■

No demostraremos esta última proposición.

TEOREMA 62 (de Jordan). Sea $A \in M_n(C)$.

i) Existe una matriz no singular S tal que $A = SJS^{-1}$ con J una matriz de Jordan:

$$(5) \qquad J = J(A) = \begin{bmatrix} J_{n_1}(\lambda_1) & 0 & . & 0 \\ 0 & J_{n_2}(\lambda_2) & . & 0 \\ . & . & . & .. \\ 0 & 0 & . & J_{n_k}(\lambda_k) \end{bmatrix}, \qquad n = n_1 + \ldots + n_k .$$

ii) Los autovalores λ_j no son necesariamente distintos.

iii) Si A es real con autovalores reales entonces S puede elegirse real.

iv) La matriz de Jordan $J=J(A)$ es única salvo por una permutación de los bloques diagonales.■

DEMOSTRACION. i)-iii) son consecuencias inmediatas de los teoremas 60 y 61. No demostraremos el punto iv), QED.

20. CONSECUENCIAS DEL TEOREMA DE JORDAN. Los elementos iguales a 1 en J pueden modificarse. En efecto, vale el

TEOREMA 63. Sea $A \in M_n(C)$.

i) Dado $\varepsilon > 0$ existe una matriz no singular R tal que $J = R^{-1}AR$ y

$$(6) \qquad J = J(\varepsilon) = \begin{bmatrix} J_{n_1}(\lambda_1, \varepsilon) & 0 & . & 0 \\ 0 & J_{n_2}(\lambda_2, \varepsilon) & . & 0 \\ . & . & . & . \\ 0 & 0 & . & J_{n_k}(\lambda_k, \varepsilon) \end{bmatrix} \qquad \text{con}$$

$$(7) \qquad J_k(\lambda, \varepsilon) = \begin{bmatrix} \lambda & \varepsilon & 0 & . & . & 0 \\ 0 & \lambda & \varepsilon & 0 & . & 0 \\ 0 & 0 & \lambda & \varepsilon & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & . & \lambda & \varepsilon \\ 0 & 0 & 0 & . & . & \lambda \end{bmatrix} \in M_k(C) .$$

ii) Si A y $\sigma(A)$ son reales entonces R puede elegirse real.■

DEMOSTRACION. Sea D la matriz diagonal: $diag(1, \varepsilon, \cdots, \varepsilon^{n-1})$. Supongamos que $W=S^{-1}AS$ sea una forma de Jordan asociada a la matriz A y $R=SD$. Entonces, el elemento ij de W es igual a $\sum_{k,l} (S^{-1})_{ik} A_{kl} S_{lj}$ y el de $J := (D^{-1}S^{-1})A(SD)$ igual a

$\sum_{k,l} \frac{(S^{-1})_{ik}}{\varepsilon^{i-1}} A_{kl} S_{lj} \varepsilon^{j-1}$. En consecuencia, J y W tienen los mismos elementos diagonales y los mismos ceros fuera de la diagonal mientras que $J_{j,j+1} = W_{j,j+1}\varepsilon = 1.\varepsilon$, QED.

El siguiente teorema de aproximación, que es un corolario del teorema 63, afirma que las matrices diagonalizables son densas en el espacio de matrices.

TEOREMA 64. Sean $A = [a_{ij}] \in M_n(C)$, $\delta > 0$. Existe $A(\delta) \in M_n(C)$, diagonalizable y con autovalores distintos, tal que

$$(8) \qquad \sum_{ij} \left| a_{ij} - a_{ij}(\delta) \right|^2 < \delta . \quad ■$$

DEMOSTRACION. Sean $A = RJ(A)R^{-1}$ y $J' = J'(A)$ una matriz obtenida de $J(A)$ modificando ligeramente sus elementos diagonales y de manera que estos sean dos a dos distintos: $\sum |\lambda_j - \lambda_j'| < \varepsilon$. Definamos: $A(\delta) := RJ'R^{-1}$, $N(\delta) = R(J(A) - J')R^{-1}$. $A(\delta)$ tiene autovalores distintos, los de J', por lo que es diagonalizable y vale, $A - A(\delta) = N(\delta)$ y $\sum |N_{ij}|^2 < \delta$ si ε es adecuadamente elegido, QED.

Si bien la forma triangular a la que se llega por medio del teorema de Schur es más imprecisa que la de Jordan tiene al menos dos ventajas: la matriz de semejanza es unitaria y pueden triangularizarse simultáneamente todas las matrices de una familia conmutante. Una consecuencia de este hecho es el siguiente

TEOREMA 65. Sean $A, B \in M_n(C)$. Si $AB=BA$ entonces

(9) $$\sigma(A+B) \subset \sigma(A) + \sigma(B). \ \blacksquare$$

DEMOSTRACION. Sea $\{\lambda_1, \cdots, \lambda_n\}$ la familia de autovalores de A y $\{\mu_1, \cdots, \mu_n\}$ la de B. Existe una matriz unitaria, U, tal que $UAU^{-1} = R$, $UBU^{-1} = S$ son triangulares superiores (v. teor. 26). Por tanto, y luego de reordenar los autovalores de la familia $\{\mu_1, \cdots, \mu_n\}$, que podemos suponer indicados de la misma forma, de $U(A+B)U^{-1} = R + S$ vemos que

(10) $$\{\lambda_i + \mu_i : i =,1,\ldots,n\} = \sigma(A+B),$$

pues $A+B$ y $R+S$ tienen los mismos autovalores, QED.

NB. En general, en (9) no vale la igualdad y si A y B no conmutan puede no verificarse la inclusión.

Recurriendo al teorema de Jordan puede darse una demostración del siguiente

TEOREMA 66. Sea $A \in M_n(C), B \in M_n(C), AB = BA$. Si A es no derogatoria entonces existe un polinomio $p(t)$, y de grado a lo sumo $n-1$, tal que $B = p(A)$.$\blacksquare$

21. ESTRUCTURA DE LA MATRIZ DE JORDAN Y EL POLINOMIO MINIMAL.

El número de bloques k en la forma de Jordan (5) asociada a la matriz A es el número de autovectores linealmente independientes de A (y de J). En particular, si $k = n$, A es diagonalizable.

La multiplicidad geométrica de un autovalor λ es igual al número de bloques que tienen a λ en su diagonal.

El número de veces que aparece λ en la diagonal de J es su multiplicidad algebraica.

Por tanto, el conocimiento de los autovalores con sus multiplicidades no es suficiente para determinar la matriz de Jordan. Es necesario conocer también el tamaño de los bloques correspondientes a cada autovalor.

Como $(J_k(\lambda) - \lambda I)^j$ es nula si $j = k$ y no lo es si $j < k$ puede verse que el orden de los bloques está relacionado con los rangos de ciertas potencias.

EJEMPLO. $A = J_k(0)^2 = \begin{bmatrix} 0 & J_{k-1}(0) \\ 0 & 0 \end{bmatrix}$ tiene por forma de Jordan a $W = \begin{bmatrix} J_m(0) & 0 \\ 0 & J_m(0) \end{bmatrix}$ si $k = 2m$ y a $\begin{bmatrix} J_m(0) & 0 \\ 0 & J_{m-1}(0) \end{bmatrix}$ si $k = 2m-1$.

En efecto, observemos en primer lugar que A y W tienen los mismos autovalores. Sea $k=2m$; entonces $rango J_k(0)^s = k - s$ si $s < k$, $rango J_k(0)^s = 0$ si $s \ge k$. En particular,

$rango A^h = 2\sup\{(m-h),0\} = rango W^h$. Pero esta situación no puede presentarse si la forma de Jordan de A no tiene exactamente dos bloques del mismo orden. El caso $k = 2m-1$ se deja al lector.

Dada la matriz $A \in M_n(C)$, su polinomio característico *aniquila* a A, es decir, $p_A(A) = 0$ como afirma el teorema de Cayley-Hamilton. $p_A(t)$ es de grado n y mónico. Habrá entonces un polinomio $q(t)$, mónico, de grado mínimo, necesariamente $\leq n$, tal que $q(A) = 0$. Recurriendo al algoritmo de Euclides se prueba fácilmente el

TEOREMA 67. Existe un único polinomio mónico de grado mínimo, $q_A(t)$, que aniquila a A. Su grado es a lo sumo n y $q_A(t)$ divide a $p_A(t)$ y a todo polinomio que aniquila a A.■

Llamaremos a $q_A(t)$ *el polinomio minimal de A*. Observemos que matrices semejantes tienen los mismos polinomios característico y minimal. En efecto, si $A = SBS^{-1}$ entonces $q_B(A) = Sq_B(B)S^{-1} = 0$ y $q_A | q_B$, por lo que $q_A \equiv q_B$.

Además, $p_A(\lambda) = 0 \Rightarrow q_A(\lambda) = 0$. En efecto, si $Ax = \lambda x$, $x \neq 0$, entonces $0 = q_A(A)x = q_A(\lambda)x$. Por tanto, $q_A(\lambda) = 0$. En consecuencia, si $p_A(t) = \prod_{i=1}^{m}(t-\lambda_i)^{s_i}$ con los λ_i distintos y $\sum s_i = n$ entonces

(11) $$q_A(t) = \prod_{i=1}^{m}(t-\lambda_i)^{r_i} \text{ con } 1 \leq r_i \leq s_i.$$

Sean $A = SJS^{-1}, J = J_n(\lambda)$, (v. (1)). El polinomio característico de A es $(t-\lambda)^n$ y coincide con su polinomio minimal pues $(J-\lambda I)^k \neq 0$ si $1 \leq k < n$, $((J-\lambda I)^k = 0$ si $k \geq n)$.

Apliquemos este resultado. Si $A = SJS^{-1}$ y $J = J(\lambda)$ la matriz de Jordan de A con

$$J = \begin{bmatrix} J_r(\lambda) & 0 \\ 0 & J_s(\lambda) \end{bmatrix}, r \geq s,$$

entonces $q_J(t) = (t-\lambda)^r$, con r el orden del bloque más grande correspondiente a λ.

En esta forma se obtiene el

TEOREMA 68. Sea A una matriz de orden n cuyos autovalores distintos son $\lambda_1, \cdots, \lambda_m$. Entonces, $q_A(t) = \prod_{i=1}^{m}(t-\lambda_i)^{r_i}$ donde r_i es el orden del bloque más grande de A correspondiente al autovalor λ_i.■

De este resultado sigue que A es diagonalizable si y sólo si r_i=1 para todo λ_i. De otra forma, una matriz A cuyos autovalores distintos son $\lambda_1, \cdots, \lambda_m$ es diagonalizable si y sólo si $q(A) = 0$ donde $q(t) = \prod_{i=1}^{m}(t-\lambda_i)$.

Para una matriz A de orden n vale que su polinomio minimal es a lo sumo de grado n que es el grado de su polinomio característico. Sin embargo vale el (cf. (1), §11),

TEOREMA 69. i) Todo polinomio mónico $p(t) = t^n + a_{n-1}t^{n-1} + \ldots + a_0$ es polinomio minimal y polinomio característico de su matriz acompañante $C = C(p)$.

ii) A es semejante a $C(p_A(t))$ si y sólo si $q_A(t) \equiv p_A(t)$. ■

22. LA MAYORIZACION. La noción de mayorización es importante pues en la teoría de matrices aparece naturalmente como la relación significativa y precisa entre dos conjuntos de números reales, por ejemplo, entre las componentes de dos vectores.

DEFINICION 29. Si las familias de números reales $\alpha = \{ a_1, \cdots, a_n \}$ y $\beta = \{ b_1, \cdots, b_n \}$ tales que $\sum a_i = \sum b_i$, se ordenan como $a_{i_1} \le \ldots \le a_{i_n}$ y $b_{j_1} \le \ldots \le b_{j_n}$, diremos que β *mayoriza* a α, $\beta \succ \alpha$, si para todo k=1,...n, $\sum_{h=1}^{k} b_{j_h} \ge \sum_{h=1}^{k} a_{i_h}$.■

Esto equivale a decir que la suma de k elementos distintos de la familia α tiene un mínimo que es menor o igual que la suma de k elementos distintos de β, cualesquiera sean ellos. El resultado siguiente es una caracterización de la relación de mayoración.

TEOREMA 70. Son equivalentes:

a) $\beta \succ \alpha$

b) Existe una matriz doblemente estocástica S tal que $\beta = S\alpha$.■

DEMOSTRACION. $a) \Rightarrow b)$ será demostrada más adelante. Veamos que $b) \Rightarrow a)$. Sea $\Sigma \in M_n$ doblemente estocástica tal que $\beta = \Sigma\alpha$. Sean y y x vectores ordenados en forma creciente tales que $y = P\beta, x = Q\alpha$, P y Q matrices de permutación. Entonces $y = (P\Sigma Q^T)x$ donde $S = P\Sigma Q^T$ es doblemente estocástica como se deduce del teorema de Birkhoff recordando que el producto de matrices de permutación es de permutación. Si probamos que $y \succ x$ entonces tendremos que $\beta \succ \alpha$. Debemos ver que $\sum_{i=1}^{k}(y_i - x_i) \ge 0$ para todo k. Pero $\sum_{i=1}^{k}(y_i - x_i) = \sum_{i=1}^{k}\sum_{j=1}^{n} s_{ij}x_j - \sum_{i=1}^{k} x_i = \sum_{j=1}^{n}\sum_{i=1}^{k} s_{ij}x_j - \sum_{i=1}^{k} x_i$.

Definamos, $w_j^{(k)} = \sum_{i=1}^{k} s_{ij}$. Entonces, $0 \le w_j^{(k)} \le 1, \sum_{j=1}^{n} w_j^{(k)} = k$. En consecuencia tenemos,

$$\sum_{i=1}^{k}(y_i - x_i) = \sum_{j=1}^{n} w_j^{(k)} x_j - \sum_{i=1}^{k} x_i + x_k(k - \sum_{j=1}^{n} w_j^{(k)}) = -\sum_{j=1}^{n} w_j^{(k)}(x_k - x_j) + \sum_{j=1}^{k}(x_k - x_j) =$$

$$= \sum_{j=1}^{k}(1 - w_j^{(k)})(x_k - x_j) - \sum_{j=k+1}^{n} w_j^{(k)}(x_k - x_j) = \sum_{j=1}^{k}(1 - w_j^{(k)})(x_k - x_j) + \sum_{j=k+1}^{n} w_j^{(k)}(x_j - x_k).$$

Como las componentes del vector x están ordenadas, las dos últimas sumas son no negativas, QED

Cuando una matriz de permutación P se aplica a un vector v se obtiene un vector $w = Pv$ cuyas componentes son una permutación de las de v. Por otra parte sabemos que una matriz doblemente estocástica es una combinación convexa de matrices de permutación: $S = \sum_{1}^{N} p_i P_i, p_i > 0, \sum_{1}^{N} p_i = 1$. De b) del teorema 70 se deduce entonces que si $\beta \succ \alpha$ entonces β es una combinación convexa de permutaciones de α. Naturalmente, si esto ocurre, construímos con las permutaciones de α una matriz doblemente estocástica S de manera que $\beta = S\alpha$. Tenemos así expresada en forma ligeramente diferente la proposición b) del teorema precedente:

$\beta \succ \alpha$ si y sólo si β es una combinación convexa de permutaciones de α.

Veremos un resultado debido a Schur, teorema 75, para matrices hermitianas: $A = A^*$, que afirma que el conjunto de los elementos diagonales de A mayoriza al conjunto de sus autovalores. Recordemos que estas matrices tienen autovalores reales lo mismo que sus elementos de la diagonal principal y que vale $trA = \sum a_{jj} = \sum \lambda_j$.

Citamos a continuación algunos teoremas sobre los autovalores de este tipo de matrices.

TEOREMA 71 (de Rayleigh-Ritz). Sea A hermitiana. Si se ordenan los autovalores de A, $\lambda_1 \leq \cdots \leq \lambda_n$, valen,

i) $\lambda_1 x^* x \leq x^* Ax \leq \lambda_n x^* x$ para todo vector complejo x,

ii) $\lambda_n = \max\limits_{x^*x=1} x^* Ax$,

iii) $\lambda_1 = \min\limits_{x^*x=1} x^* Ax$. ∎

TEOREMA 72 (del minimax de Courant-Fischer). Sea A hermitiana. Suponemos ordenados los autovalores de A: $\lambda_1 \leq \cdots \leq \lambda_n$. Entonces, para $x \in C^n, x^* x = 1$,

a) $\lambda_k = \min\limits_{w_1,\ldots,w_{n-k}} \max\limits_{x \perp w_1,\ldots,w_{n-k}} x^* Ax$

b) $\lambda_k = \max\limits_{w_1,\ldots,w_{k-1}} \min\limits_{x \perp w_1,\ldots,w_{k-1}} x^* Ax$. ∎

TEOREMA 73 (Weyl). Sean A, B hermitianas y supongamos ordenados de menor a mayor los autovalores de A, B y A+B. Entonces,

i) $\lambda_{j+k-n}(A+B) \leq \lambda_j(A) + \lambda_k(B)$, $\quad 1 \leq j,k \leq n,\ j+k \geq n+1$,

ii) $\lambda_j(A) + \lambda_k(B) \leq \lambda_{j+k-1}(A+B)$, $\quad 1 \leq j,k \leq n,\ j+k \leq n+1$. ∎

A la siguiente proposición i) se la cita como el *teorema de entrelazamiento de autovalores* (con la matriz orlada).

TEOREMA 74. i) Sea a un número real. Dada la matriz hermitiana A consideremos su matriz orlada

$$\hat{A} = \begin{bmatrix} A & y \\ y^* & a \end{bmatrix}. \tag{12}$$

$\hat{A}$ es hermitiana y sus autovalores están vinculados con los de A en la siguiente forma:

$$\hat{\lambda}_1 \leq \lambda_1 \leq \hat{\lambda}_2 \leq \lambda_2 \leq \cdots \leq \lambda_n \leq \hat{\lambda}_{n+1}. \tag{13}$$

ii) Sea dada una familia de números reales como en (13). Sea $D = diag[\lambda_1,\ldots,\lambda_n]$. Existe un número real a y un vector $y \in R^n$ tal que $\{\hat{\lambda}_1,\ldots,\hat{\lambda}_{n+1}\}$ es la familia de autovalores de la matriz simétrica real $\hat{A} = \begin{bmatrix} D & y \\ y^T & a \end{bmatrix} \in M_{n+1}(R)$. ∎

DEMOSTRACION. Solamente probaremos i). Sea $1 \leq k \leq n$ y $\hat{x} = \begin{bmatrix} x \\ t \end{bmatrix}, x \in C^n, t \in C$.

También $\hat{w} = \begin{bmatrix} w \\ u \end{bmatrix}, w \in C^n, u \in C$. Usando a) del teorema 72 obtenemos

$$\hat{\lambda}_{k+1} = \min_{\hat{w}_1,\ldots,\hat{w}_{(n+1)-(k+1)}} \max_{0 \neq \hat{x} \perp \hat{w}_1,\ldots,\hat{w}_{(n+1)-(k+1)}} \frac{\hat{x}^* \hat{A} \hat{x}}{\hat{x}^* \hat{x}} \geq \min_{\hat{w}_1,\ldots,\hat{w}_{(n+1)-(k+1)}} \max_{0 \neq \hat{x} \perp \hat{w}_1,\ldots,\hat{w}_{(n+1)-(k+1)}, e_{n+1}} \frac{\hat{x}^* \hat{A} \hat{x}}{\hat{x}^* \hat{x}} = \vartheta.$$

La última expresión es $\geq \lambda_k$. En efecto,

$\hat{x}^* \hat{A} \hat{x} = \begin{bmatrix} x^* & \bar{t} \end{bmatrix} \begin{bmatrix} A & y \\ y^* & a \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} = \begin{bmatrix} x^* A + \bar{t} y^* & x^* y + \bar{t} a \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix}$ es igual a $x^* Ax$ si $t = 0$ que es el valor de t cuando $\hat{x} \perp e_{n+1}$. Por tanto, $\vartheta \geq \min_{w_1,\ldots,w_{n-k}} \max_{0 \neq x \perp w_1,\ldots,w_{n-k}} \frac{x^* Ax}{x^* x} = \lambda_k$ y $\hat{\lambda}_{k+1} \geq \lambda_k$.

Usando b) del teorema de Courant-Fischer se demuestra que $\hat{\lambda}_k \leq \lambda_k$. Como $1 \leq k \leq n$ sigue i) del teorema, QED.

TEOREMA 75 (Schur). Sea $A \in M_n(C)$, hermitiana. Entonces, $\beta = \{a_{jj}\} \succ \alpha = \{\lambda_j\}$.■

DEMOSTRACION. Por inducción. El resultado es obvio para n=1. Supongamos $n > 1$ y los autovalores ordenados de menor a mayor. Recordemos que $\sum a_{jj} = \sum \lambda_j$. La matriz B=$P^T AP$, con P la permutación que lleva la columna y la fila donde se encuentra el elemento diagonal más grande a la posición n-ésima, es autoadjunta y tiene los mismos autovalores que A pues $P^T = P^{-1}$. Sea A' la submatriz principal de B obtenida eliminando la fila y la columna n-ésima. Entonces B es la matriz orlada de A'. Por el teorema de entrelazamiento de los autovalores tenemos: $\lambda_1 \leq \lambda'_1 \leq \lambda_2 \leq \lambda'_2 \leq \cdots \leq \lambda_n$. Luego, $\sum_1^k \lambda_i \leq \sum_1^k \lambda'_i$ si $k < n$. Por hipótesis inductiva, $\beta' \succ \alpha'$. Por tanto, $\sum_1^k \lambda_i \leq \min_k \sum_{j=1}^k a'_{m_j m_j} = \min_k \sum_{j=1}^k a_{m_j m_j}$, QED.

El siguiente resultado confirma la relación íntima del concepto de mayorización con las matrices hermitianas, especialmente la existente entre la diagonal principal y los autovalores.

TEOREMA 76. Sean $a = [a_i]$, $\lambda = [\lambda_i]$, $i = 1,\ldots,n$, vectores reales (ordenados en forma creciente). Si $a \succ \lambda$ entonces existe una matriz simétrica real, $A = [a_{ij}] \in M_n(R)$, con autovalores { λ_i } tal que para todo i, $a_{ii} = a_i$. ■

En la demostración del teorema haremos uso de la siguiente proposición auxiliar.

LEMA AUXILIAR. Sean $n \geq 2$ y $a = [a_i]$, $b = [b_i]$, vectores de R^n con sus componentes ordenadas. Si $b \succ a$ entonces hay un vector c en R^{n-1} tal que

$$a_1 \leq c_1 \leq a_2 \leq c_2 \leq \ldots \leq c_{n-1} \leq a_n$$

y con $\hat{b} \succ c$ donde $\hat{b} = [b_1,\ldots,b_{n-1}]^T \in R^{n-1}$.■

DEMOSTRACION. Si n=2 el resultado es trivial. Sea $n > 2$. Definamos $K = \{ d = [d_1,\ldots,d_{n-1}]^T \} \subset R^{n-1}$ como el conjunto de vectores d que verifican

$$a_1 \leq d_1 \leq a_2 \ldots \leq d_{n-1} \leq a_n, \qquad \sum_{i=1}^k d_i \leq \sum_{i=1}^k b_i,\ k = 1,2,\ldots,n-2.$$

El conjunto K es compacto y convexo. Es no vacio pues $\hat{a} = [a_1,\ldots,a_{n-1}]$ está en K. Definamos en K la función $f(d) = \sum_1^{n-1} d_i$. Vale $f(\hat{a}) \leq f(\hat{b})$. Si encontramos un d tal que $f(\hat{b}) \leq f(d)$, una combinación convexa, c, entre a y d, pertenecerá a K y verificará $f(c) = f(\hat{b})$. Este vector satisfará entonces la tesis del teorema. Por ser f continua en K, existe $m \in K$ donde f toma su máximo. Este será el d buscado. En efecto, m verifica

I) $m_k \leq a_{k+1}$, $k = 1,2,\ldots,n-1$, II) $\sum_1^k m_i \leq \sum_1^k b_i$, $k = 1,2,\ldots,n-2$.

Supongamos estrictas todas las desigualdades de II). En este caso no puede haber una desigualdad estricta en I) pues si no podría aumentarse una componente de m para obtener m' en K, aún satisfaciendo I) y II) pero con $f(m')>f(m)$, contradicción. O sea, si en II) se da siempre el signo < entonces en I) debe ser siempre =.

Luego, $m=[a_2,...,a_n]$ y vale

$$f(m)=(\sum_1^n a_i)-a_1=(\sum_1^n b_i)-a_1=(\sum_1^{n-1} b_i)+b_n-a_1\geq\sum_1^{n-1} b_i=f(\hat{b}).$$

Entonces podemos elegir $d=m$.

Si no todas las desigualdades de II) son estrictas habrá por lo menos un signo =. Sea r el k más grande donde aparece en II) la igualdad, es decir,

$$\sum_1^r m_i=\sum_1^r b_i,\quad \sum_1^k m_i<\sum_1^k b_i \text{ para } k=r+1,...,n-2.$$

En esta situción y por las mismas razones que antes, debemos tener

$$m_k=a_{k+1} \text{ para. } k=r+1,...,n-1.$$

Luego, $f(m)=\sum_1^r m_i+\sum_{r+1}^{n-1} m_i=\sum_1^r b_i+\sum_{r+2}^n a_i=\sum_1^{n-1} b_i-\sum_{r+1}^{n-1} b_i+\sum_1^n a_i-\sum_1^{r+1} a_i$. Por tanto,

$$f(m)=\sum_1^{n-1} b_i+(\sum_1^{r+1} b_i-\sum_1^{r+1} a_i)+\sum_{r+2}^n b_i-\sum_{r+1}^{n-1} b_i\geq\sum_1^{n-1} b_i+\sum_{r+2}^n b_i-\sum_{r+1}^{n-1} b_i\geq\sum_1^{n-1} b_i=f(\hat{b}),\quad \text{QED.}$$

DEMOSTRACION DEL TEOREMA 76. Por inducción. Es trivial si $n=1$. Suponemos demostrado el teorema para todos los vectores a y λ de dimensión $n-1$. Por el lema auxiliar sabemos que existe un vector $\gamma\in R^{n-1}$ tal que $\hat{a}=[a_1...a_{n-1}]^T\succ\gamma$ y

$$\lambda_1\leq\gamma_1\leq\lambda_2\leq\gamma_2\leq...\leq\gamma_{n-1}\leq\lambda_n.$$

Por la hipótesis inductiva existe una matriz simétrica real $B=[b_{ij}]\in M_{n-1}(R), b_{ii}=a_i$, con autovalores $[\gamma_i]$. Sea $\Gamma=diag[\gamma_1,...,\gamma_{n-1}]$. Existe entonces una matriz ortogonal Q tal que $B=Q\Gamma Q^T$. Por ii) del teorema 74 hay una matriz simétrica real $\hat{A}=\begin{bmatrix}\Gamma & y\\ y^T & \alpha\end{bmatrix}$ $\in M_n(R), y\in R^{n-1},\alpha\in R$ con autovalores $[\lambda_i]$. Como $a\succ\lambda,\hat{a}\succ\gamma$ tenemos que $tr\hat{A}=\sum\lambda_i=\sum_1^n a_i=tr\Gamma+\alpha=\alpha+\sum_1^{n-1} a_i$ por lo que $\alpha=a_n$. Pero

$$A:=\begin{bmatrix}Q & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\Gamma & y\\ y^T & \alpha\end{bmatrix}\begin{bmatrix}Q^T & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}Q\Gamma Q^T & Qy\\ (Qy)^T & \alpha\end{bmatrix}=\begin{bmatrix}B & Qy\\ (Qy)^T & a_n\end{bmatrix}$$

tiene los autovalores $[\lambda_i]$ y los elementos de a en la diagonal principal, QED.

Veamos una aplicación de este resultado.

DEMOSTRACION DE QUE $a)\Rightarrow b)$ EN EL TEOREMA 70. Por el teorema 76 sabemos que existe una matriz simétrica real $B=[b_{ij}]$ tal que $\forall i\ b_{ii}=b_i$ y $\{a_1,\cdots,a_n\}$ es la familia de autovalores de B (cf. definición 29). Existe una matriz ortogonal O tal que $B=ODO^*=ODO^T$ con $D=diag\{a_1,...,a_n\}$. Sea $S=[s_{ij}]$, $s_{ij}=o_{ij}^2$. S es doblemente estocástica y vale $b_{ii}=\sum_k s_{ik}a_k=(S\alpha)_i$, QED.

La siguiente proposición es una extensión del teorema 74.

TEOREMA 77 (principio de inclusión). Sean $A \in M_n(C)$, hermitiana y $1 \leq r \leq n$. Sea A_r una submatriz principal de A (obtenida de A eliminando $n-r$ filas y las columnas con los mismos índices). Para todo k tal que $1 \leq k \leq r$ tenemos

(14) $$\lambda_k(A) \leq \lambda_k(A_r) \leq \lambda_{k+n-r}(A). \blacksquare$$

No demostraremos este principio aunque podría hacerse utilizando solamente el teorema de Courant-Fischer.

Un corolario del mismo es el siguiente teorema 78.

(H) Sean $A \in M_n(C)$ hermitiana, $1 \leq r \leq n$ y $u_1,...,u_r$, vectores ortonormales en C^n. Sea $B_r := [u_i^* A u_j] \in M_r$. Esta matriz es igual a U^*AU donde $U \in M_{n,r}$, $U = [u_1\, u_2 \cdots u_r]$ con $U^*U = I \in M_r$; o sea, $B_r = \begin{bmatrix} u_1^* \\ . \\ u_r^* \end{bmatrix} [Au_1 \cdots Au_r]$. Con estas hipótesis tenemos el

TEOREMA 78 (de separación de Poincaré). Para los autovalores ordenados vale

(15) $$\lambda_k(A) \leq \lambda_k(B_r) \leq \lambda_{k+n-r}(A), \quad k = 1,2,...,r. \blacksquare$$

DEMOSTRACION. Si $r = n$ el teorema no dice nada pues B_n tiene los mismos autovalores que A. Sea $r < n$. Entonces, extendiendo la familia $\{u_1,...,u_r\}$ a la familia ortonormal $\{u_1,...,u_r, u_{r+1},...,u_n\}$ podemos definir la matriz unitaria $U = [u_1\, u_2 \cdots u_n]$. En este caso $B_r = [u_i^* A u_j]$ es una submatriz principal de $B_n = U^*AU$. La tesis sigue ahora del teorema 77, QED.

Sumando las desigualdades (15) obtenemos el siguiente corolario al teorema 78.

COROLARIO. Supongamos (H). Entonces, para $U \in M_{n,r}$,

(16) $$\lambda_1(A) + ... + \lambda_r(A) = \min_{U^*U = I \in M_r} trU^*AU$$

(17) $$\lambda_{n-r+1}(A) + ... + \lambda_n(A) = \max_{U^*U = I \in M_r} trU^*AU. \blacksquare$$

DEMOSTRACION. Veamos (16). (15) implica que

$$\lambda_1(A) + ... + \lambda_r(A) \leq \min_{U^*U = I \in M_r} trU^*AU.$$

Si en $U = [u_1\, u_2 \cdots u_r]$ elegimos autovectores ortonormales correspondientes a los r autovalores menores de A tendremos: $B_r = \begin{bmatrix} u_1^* \\ . \\ u_r^* \end{bmatrix} [\lambda_1 u_1 \cdots \lambda_r u_r] = \begin{bmatrix} \lambda_1 & . & 0 \\ . & . & . \\ 0 & . & \lambda_r \end{bmatrix}$ cuya traza es $\lambda_1(A) + ... + \lambda_r(A)$ y vale la igualdad en (16). Análogamente se prueba (17), QED.

Volviendo al tema de la mayorización tenemos el siguiente complemento al teorema de Weyl.

TEOREMA 79. Sean A, B hermitianas y supongamos ordenados de menor a mayor los autovalores de A, B y $A+B$. Entonces, si $\lambda(A)$ denota al vector de autovalores ordenados de A, vale

(18) $$\lambda(A+B) \succ \lambda(A) + \lambda(B). \blacksquare$$

DEMOSTRACION. Usando (16) obtenemos

$$\sum_1^k \lambda_i(A+B) = \min_{U^*U = I \in M_k} tr(U^*(A+B)U) = \min_{U^*U = I \in M_k} tr\,(U^*AU + U^*BU).$$

La última expresión es igual a $\min_{U^*U=I\in M_k}(trU^*AU+trU^*BU)\geq$

$$\geq \min_{U^*U=I\in M_k} trU^*AU + \min_{U^*U=I\in M_k} trU^*BU = \sum_1^k \lambda_i(A)+\sum_1^k \lambda_i(B).$$

Para $k=n$, $U\in M_n$ y $trU^*AU=trA$. Como $tr(A+B)=trA+trB$, se da la igualdad entre las dos expresiones precedentes para $k=n$, QED.

Otra consecuencia del teorema de Courant-Fischer es el siguiente

TEOREMA 80. Sea $A\in M_n(C)$ una matriz hermitiana. Sea $S=S_k$ un subespacio de C^n de dimensión positiva *k*. Supongamos ordenados los autovalores de A. Si existe una constante *b* tal que $x^*Ax\geq bx^*x$ para todo *x* de *S* entonces

(19) $$\lambda_n\geq\cdots\geq\lambda_{n-k+1}\geq b.$$

Si existe una constante *a* tal que $x^*Ax\leq ax^*x$ para todo *x* de *S* entonces

(20) $$a\geq\lambda_k\geq\cdots\geq\lambda_1. \ ■$$

DEMOSTRACION. Supongamos que la familia $V=\{u_1,\ldots,u_{n-k}\}$ esté compuesta por vectores ortonormales que generan $S^\perp$. Entonces, por b) del teorema 72, tenemos

$$b\leq \min_{0\neq x\in S}\frac{x^*Ax}{x^*x}=\min_{0\neq x\in\perp V}\frac{x^*Ax}{x^*x}\leq \max_{w_1,\ldots,w_{n-k}}\ \min_{0\neq x\perp w_1,\ldots,w_{n-k}}\frac{x^*Ax}{x^*x}=\lambda_{n-k+1}.$$

Utilizando a) del mismo teorema se prueba (20), QED.

COROLARIO. Si $x^*Ax\geq 0$ en *S* entonces por lo menos *k* autovalores son ≥ 0 y si $x^*Ax>0$ en ese subespacio entonces hay por lo menos *k* autovalores >0.

23. OTRA RELACION DE EQUIVALENCIA.

DEFINICION 29. Las matrices *A*, *B* se dicen **congruentes* si existe una matriz no singular *S* tal que $B=SAS^*$.■

La *congruencia define una relación de equivalencia.

DEFINICION 30. Sea $A\in M_n(C)$ una matriz hermitiana. La *inercia* de A es la tripla ordenada $i(A)=(i_+(A),i_-(A),i_0(A))$ donde $i_+(A)$ es el número de autovalores positivos contados según su multiplicidad e $i_-(A)$ es el número de autovalores negativos también contados según su multiplicidad. $i_0(A)=n-r=$número de autovalores nulos donde $r=rango(A)=i_++i_-$. El *signo* de A es i_+-i_-.■

TEOREMA 81 (ley de inercia de Sylvester). Sean $A,B\in M_n(C)$, hermitianas. $i(A)=i(B)$ si y sólo si A y B son *congruentes.■

DEMOSTRACION. Supongamos que los autovalores positivos vienen indicados antes que los negativos y estos antes que los nulos. Sabemos que $A=UGU^*$ con $G=diag[\lambda_1,\ldots,\lambda_n]$. Sea $D=diag\left[+\sqrt{\lambda_1},\ldots,+\sqrt{\lambda_{i_+}},+\sqrt{-\lambda_{i_++1}},\ldots,+\sqrt{-\lambda_{i_++i_-}},1,\ldots,1\right]$ una matriz diagonal con todos sus elementos diagonales positivos. Entonces,

$$G=DYD=D\begin{bmatrix}I_+&0&0\\0&I_-&0\\0&0&0\end{bmatrix}D \text{ donde } I_+=I\in M_{i_+(A)},\ I_-=-I\in M_{i_-(A)},\ Y\in M_n(R).$$

Vale, $A=UGU^*=SYS^*$ con $S=UD$. $Y=Y(A)$ es la *matriz de inercia* de A y S es no singular. Es decir, toda matriz hermitiana es *congruente con su matriz de inercia. La demostración del teorema se reduce entonces a probar que A es *congruente con B si y sólo si tienen la misma matriz de inercia. En verdad sólo debemos probar que $A=SBS^*\Rightarrow Y(A)=Y(B)$. Si A y B son *congruentes entonces tienen el mismo rango por lo que basta ver que $i_+(A)=i_+(B)$. Sea $\{v_1,\dots,v_{i_+(A)}\}$ una familia ortonormal de autovectores de A correspondientes a los autovalores $\{\lambda_1(A),\dots,\lambda_{i_+(A)}\}$. Sea E el espacio generado por estos. Si $x=a_1v_1+\dots+a_{i_+}v_{i_+}$ entonces $x^*Ax=\sum_1^{i_+}|a_j|^2\lambda_j>0$. Luego, $0<x^*SBS^*x=(S^*x)^*B(S^*x)$. Por lo que $y^*By>0$ en todo y no nulo de S^*E. Del teorema 80 sigue ahora que $i_+(B)\geq i_+(A)$ y por tanto que $i_+(B)=i_+(A)$, QED.

24. MATRICES DEFINIDAS POSITIVAS. Sea $A\in M_n(C)$ una matriz hermitiana. Como $\overline{x^*Ax}=x^TA^T\overline{x}=(x^*Ax)^T=x^*Ax\in C$, vemos que para todo x, x^*Ax es real. Esto *caracteriza* a las matrices hermitianas. En efecto, si escribimos $B=\dfrac{A+A^*}{2},C=\dfrac{A-A^*}{2i}$ tenemos que B y C son hermitianas y $A=B+iC$. Si x^*Ax es real entonces $x^*Cx=0$ para todo x y aplicando el teorema 80 todos sus autovalores son nulos. Como existe una matriz unitaria U tal que $C=U^*DU$ con D la matriz diagonal de los autovalores de C resulta que $C=0$. Luego, $A=B$ hermitiana.

DEFINICION 31. Una matriz $A\in M_n(C)$ se dice *definida positiva* si $x\neq 0$ implica $x^*Ax>0$. Se dice *semidefinida positiva* si $x\neq 0$ implica $x^*Ax\geq 0$. Idem para *definida negativa* y *semidefinida negativa*. Cuando no pertenece a ninguna de estas categorías se dice que A es *indefinida*.■

Por lo dicho sabemos que una matriz definida positiva (semidefinida positiva) debe ser hermitiana pues x^*Ax es real. Si los autovalores $\{d_j\}$ de A, hermitiana, son positivos (no negativos) y $x\neq 0$ entonces para cierta U unitaria vale

$$x^*Ax=x^*U^*DUx=y^*Dy=\sum d_i|y_i|^2>0\,(\geq 0)\text{ con } y=Ux.$$

O sea, $\forall x\neq 0\;\; x^*Ax>0\,(\geq 0)$ *si y sólo si* $D>0(\geq 0)$. Por tanto su espectro, $\sigma(A)$, es positivo (no negativo) si y sólo si A es definida positiva (semidefinida positiva).

Luego, la traza y el determinante de una matriz definida positiva son positivos.

NB. Como muestra el ejemplo $M=\begin{bmatrix}1&2\\2&1\end{bmatrix}$ una matriz puede ser positiva sin ser siquiera semidefinida positiva. En efecto, $\sigma(M)=\{-1,3\}$.

Si A es definida positiva también lo son $\overline{A},A^*,A^T$. También A^{-1} es definida positiva pues $x=Ay\Rightarrow x^*A^{-1}x=y^*A^*y$.

Toda submatriz principal de una matriz definida positiva es definida positiva como se ve fácilmente. Por tanto los menores principales de la matriz son positivos. Por esta razón también los elementos de la diagonal de un matriz definida positiva son positivos.

Y vale, $a_{ii}a_{jj}>|a_{ij}|^2$.

El conjunto de las matrices definidas positivas es un cono positivo en el espacio vectorial de todas las matrices. Más generalmente, las combinaciones lineales con coeficientes no negativos de matrices semidefinidas positivas son semidefinidas positivas.

Si $K \in M_{n,m}$ entonces A definida positiva implica que $K^*AK \in M_m$ es semidefinida positiva. Pero por ser A definida positiva tenemos $x^*K^*AKx > 0 \Leftrightarrow Kx \neq 0$. Entonces $K^*AKx = 0 \Rightarrow Kx = 0$ pues en ese caso $x^*K^*AKx = 0$. O sea, $K^*AKx = 0 \Leftrightarrow Kx = 0$. Es decir, K^*AK y K tienen el mismo espacio nulo como aplicaciones de C^m (en C^m y C^n respectivamente). Por tanto tienen el mismo rango: $rango(K^*AK) = rango(K)$. En consecuencia, K^*AK es definida positiva si y sólo si *rango*(K)=m.

EJERCICIOS. 1) La matriz *antisimétrica* $B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $B^T = -B$ es tal que si $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ entonces $x^*Bx = 2\,\mathrm{Im}(\overline{x_1}x_2)$.

Si x es real vale $x^TBx = 0$, real, pero B no es simétrica.

2) Supongamos que para $A \in M_n(C), x \in C^n$, pidiéramos $\mathrm{Re}(x^*Ax) > 0$ para todo x no nulo. Se prueba que esto ocurre si y sólo si la *parte hermitiana* de A, $H(A)$, es definida positiva donde $H(A) = \frac{1}{2}(A + A^*)$. Si $S(A) = \frac{1}{2}(A - A^*)$ entonces $A = H(A) + S(A)$, (cf. inicio §24). $S(A)$ verifica $S(A)^* = -S(A)$, o sea , es *antihermitiana*:.

25. RAICES Y CARACTERIZACIONES DE MATRICES DEFINIDAS POSITIVAS .

DEFINICION 32. Sea $V = \{v_1, \dots, v_k\}$ una familia de vectores en un espacio vectorial E con producto interior $\langle .,. \rangle$. La *matriz de Gram* de V *con respecto a* $\langle .,. \rangle$ es la matriz $G = [g_{ij}] \in M_k,\ g_{ij} = \langle v_j, v_i \rangle$.■

Si $k = n =$ dimensión de E entonces $G[x_1, \dots, x_n]^T = \left[\left\langle \sum_1^n x_j v_j, v_1 \right\rangle, \dots, \left\langle \sum_1^n x_j v_j, v_n \right\rangle\right]^T$.

Sea $A \in M_n(C)$ una matriz hermitiana. Vale el

TEOREMA 82. I) A es definida positiva si y sólo si todos sus autovalores son positivos. A es semidefinida positiva si y sólo si todos sus autovalores son no negativos.
II) A es definida positiva si y sólo si una familia de n menores principales de ordenes decrecientes de n a 1, encajada (es decir, cada menor es el determinante de una submatriz principal de la matriz del anterior) es positiva.
III) A es definida positiva si y sólo si $A = B^*B$ con B no singular, o sea, A es *congruente con la identidad.
IV) A es semidefinida positiva con rango r si y sólo si hay n vectores en C^n, $W = \{w_1, \dots, w_n\}$, conteniendo exactamente r vectores linealmente independientes tales que A es la matriz de Gram de W con respecto al producto interior euclídeo.■

Las matrices definidas positivas dentro de la familia de las matrices hermitianas representan aproximadamente el papel de los números positivos en el cuerpo de los complejos. Esa semejanza se comprueba en el siguiente teorema.

TEOREMA 83. Sean A semidefinida positiva y k un entero. Entonces existe una única matriz hermitiana semidefinida positiva, B, tal que $B^k = A$. Además, $B = q(A)$ donde q es un polinomio, por lo que B conmuta con A. A y B tienen el mismo rango, por lo que una es definida positiva si la otra lo es. También, una es real si lo es la otra.■

26. FACTORIZACIONES DE MATRICES COMPLEJAS. LA FORMA POLAR.

Sea $A \in M_n(C)$. La matriz hermitiana (semidefinida positiva) AA^* puede diagonalizarse por medio de una matriz unitaria X: $AA^* = XMX^*$ de manera que $M = diag[\mu_1,\ldots,\mu_n]$, $\mu_j \geq \mu_{j+1} \geq 0$. Si escribimos $X = [x_1 \cdots x_n]$, x_k vector columna, tendremos: $AA^* x_h = \mu_h x_h$.

Sea $\mu_h = \lambda_h^2$, $\lambda_h \geq 0$. Las raíces de los autovalores de M resultan ordenadas de la siguiente forma: $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k > \lambda_{k+1} = \ldots = 0$ $(k \leq n)$ y podemos escribir $M = \Lambda^2$, $\Lambda = diag[\lambda_1,\ldots,\lambda_n]$ con Λ *unívocamente determinada* por AA^*, o sea, por A.

Supongamos A *no singular*. Si se verifica que $\mu_j > \mu_{j+1}$ para $j = 1,\ldots,n-1$, entonces los x_j quedan *determinados salvo por un factor* $d_j = e^{i\theta_j}$, de módulo 1, pues $\|x_j\|_2 = 1$. O sea, si W es unitaria y vale $AA^* = WMW^*$ entonces $W = XD$ donde $D = diag[d_1,\ldots,d_n]$. Y todo W de esta forma sirve pues $WMW^* = X(DMD^*)X^* = XMX^* = AA^*$.

Si $\mu_j = \mu_{j+1}$ entonces fijamos una familia ortonormal de autovectores correspondientes a μ_j, de dimensión la multiplicidad de μ_j, y esto para cada autovalor no simple. Entonces, nuevamente, los x_j quedan determinados salvo por un factor de módulo uno.

En ambos casos: $AA^* = X\Lambda^2 X^*$ con $M = \Lambda^2$ (no singular).

Definamos: $Y = \Lambda^{-1} X^* A$. Por lo dicho Y queda *determinada* por X. Luego, $YY^* = \Lambda^{-1} X^*(AA^* X)\Lambda^{-1} = \Lambda^{-1} X^*(X\Lambda^2)\Lambda^{-1} = \Lambda^{-1} X^* X\Lambda = \Lambda^{-1} I\Lambda = I$. Es decir, Y tiene sus *filas ortonormales* por lo que es *unitaria*.

Si $A \in M_n(R)$ entonces X puede elegirse *real* por lo que resulta Y *real*.

En consecuencia, si $\det A \neq 0$, existen $X = X(A), XX^* = I$, $Y = Y(A), YY^* = I$, ambas reales si A es real, tales que

(21) $$A = X\Lambda Y.$$

Escribamos en esta situación: $A = X\Lambda Y = (X\Lambda X^*)(XY) = PU$. Entonces, si $x \neq 0$,

$$x^* Px = (x^* X)\Lambda(X^* x) = v^* \Lambda v = \sum \lambda_k |v_k|^2 > 0,$$

y P resulta *definida positiva*. Por otra parte, $P^2 = X\Lambda^2 X^* = AA^*$, o sea, $P = (AA^*)^{1/2}$, la *única raíz cuadrada definida positiva* de AA^* (cf. T. 83).

Además, por ser producto de dos matrices unitarias, U es *unitaria* y como $U = P^{-1}A$, U queda *unívocamente determinada*.

Si A es real, P y U son *reales* pues X e Y pudieron haberse elegido reales.

Llegamos así a la *forma polar* de la matriz A,

(22) $$A = PU,\ P = \sqrt{AA^*}, UU^* = I.$$

La matriz semidefinida positiva P que aparece en (22) se denomina *valor absoluto* (*por izquierda*) de A.

Si A es no singular veamos que $AA^* = A^* A \Leftrightarrow PU = UP$. Es decir, A es normal si y sólo si P y U *conmutan*.

Si U y P conmutan, teniendo en cuenta que $P^* = P$ y que $UU^* = I$, resulta fácilmente que A es normal. Supongamos que $AA^* = A^* A$. Entonces, $P^2 = U^* P^2 U = (U^* PU)^2$. Esta es una matriz definida positiva lo mismo que las matrices P y U^*PU que son raíces cuadradas de P^2. Por el teorema 83 tenemos $P = U^*PU$ por lo que P y U conmutan.

Recolectando resultados, que son *válidos aún para A singular*, obtenemos,
TEOREMA 84. Sea $A \in M_n(C), rango(A) = k \leq n$.

i) Existen una matriz unitaria $X \in M_n$, una matriz diagonal Λ con elementos diagonales $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k > \lambda_{k+1} = \ldots = \lambda_n = 0$ y una matriz $Y \in M_n$, unitaria, tales que

(23) $$A = X\Lambda Y.$$

ii) La matriz Λ está unívocamente determinada y $\sigma(AA^*) = \{\lambda_j^2 : j = 1,\ldots,n\}$.

iii) Las columnas de X son autovectores normalizados de AA^*.

iv) Si AA^* tiene autovalores distintos entonces X está determinada salvo por un factor $D = diag[d_1,\ldots,d_n], |d_j| = 1$, *i.e.*, $A = X\Lambda Y = W\Lambda Y'$, $W = XD$.

v) Dada X, la matriz Y queda unívocamente determinada si $rango(A) = n$.

vi) Si A es real, X e Y pueden elegirse reales.

vii) A puede factorearse en la forma $A=PU$ con P semidefinida positiva y U unitaria.

viii) Si A es normal entonces P y U conmutan.

ix) A puede factorearse en la forma $A=VQ$ con Q semidefinida positiva y V unitaria.■

(ix) resulta de factorear A^* y recordar que si U es unitaria entonces U^* es unitaria.)

NB. Es importante destacar que el teorema 84 puede extenderse a matrices rectangulares: $A \in M_{m,n}(C)$.

Si en (23) escribimos $Y=W^*$ tendremos la *descomposición en valores singulares* de la matriz $A \in M_n(C)$,

(24) $$A = X\Lambda W^*.$$

Entonces $A^* = W\Lambda X^*$ por lo que las columnas de W son autovectores de A^*A. Los elementos diagonales de Λ se denominan *valores singulares de* A y las columnas de X (W) *autovectores singulares a izquierda* (*derecha*) de A.
Como $\sigma(P)^2 = \sigma(P^2) = \sigma(AA^*)$ resulta que los autovalores de la matriz P definida no negativa son los elementos de la diagonal de Λ.

Si $W=X$ entonces A es hermitiana. Por otra parte sabemos que si A es normal entonces es unitariamente diagonalizable: $A = UDU^*$ donde $D = diag[\alpha_1,\ldots,\alpha_n]$, $\sigma(A) = \{\alpha_j\}$ por lo que $AA^* = UDD^*U$. En consecuencia, si reordenamos los elementos diagonales de DD^* (multiplicando por una adecuada matriz de permutación que es una matriz unitaria pues es real y $P^T = P^{-1}$) obtenemos $DD^* = \Lambda^2$. O sea,

(25) $$(DD^*)_{ii} = \alpha_i\overline{\alpha_i} = |\alpha_i|^2 = \mu_i = \lambda_i^2.$$

Es decir, la factorización (24), $A = X\Lambda W^*$, generaliza a matrices arbitrarias la factorización $A = UDU^*$ para matrices normales, utilizando la raíz cuadrada no negativa λ_i de $|\alpha_i^2|$ en lugar del autovalor α_i.

Si tuviéramos la forma polar $A=PU$, como P es hermitiana existe V unitaria tal que $P=V\Lambda V^*$, Λ diagonal. Entonces obtendríamos la descomposición en valores singulares del siguiente modo: $A=PU=V\Lambda(V^*U)=V\Lambda W^*$.

Sea A no singular. Entonces $A=PU=ZQ$, $U = P^{-1}A, Z = AQ^{-1}$, U y Z unitarias. La descomposición $A = V\Lambda W^*$ puede escribirse como $A=(V\Lambda V^*)(VW^*)=PU$. Análogamente, $A=(VW^*)(W\Lambda W^*)=ZQ$ de donde surge que $U=Z$. En general es $P \neq Q$, pero, por lo ya visto, $P = Q \Leftrightarrow A$ *es normal*.

EJERCICIO. Si $A := V\Sigma W^*$ con V y W unitarias, $\Sigma = diag[\sigma_1,...,\sigma_n]$, $\sigma_1 \geq ... \geq \sigma_n > 0$, entonces $V\Sigma W^*$ es la descomposición en valores singulares de la matriz A.
(Sugerencia: Si la descomposición es de la forma $A = V'\Sigma'W'^*$ entonces $\Sigma = \Sigma'$. Como $AA^*V = V\Sigma$ resulta $V' = VD$, D diagonal con autovalores de módulo uno.)

NB. Es posible *trasplantar desigualdades* en los valores singulares de matrices a desigualdades en los autovalores y recíprocamente. Por ejemplo, sea A no singular con Σ, V, W, matrices como en el ejercicio precedente. Sean

$$\tilde{A} = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}, \ \tilde{U} = \begin{bmatrix} V & -V \\ W & W \end{bmatrix}.$$

Entonces $\tilde{U}^* = \begin{bmatrix} V^* & W^* \\ -V^* & W^* \end{bmatrix}$. La matriz $\tilde{U}/\sqrt{2} \in M_{2n}(C)$ es unitaria. Vale,

$2\tilde{A} = \tilde{U}\begin{bmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{bmatrix}\tilde{U}^*$. En consecuencia, los autovalores de $\tilde{A}$ son precisamente los valores singulares de A y sus opuestos.

TEOREMA 85. Sean $A \in M_n(C)$ y $\|.\|$ una norma en este espacio. Entonces, para todo $\varepsilon > 0$ existe una matriz B con valores singulares distintos y positivos tal que $\|A - B\| < \varepsilon$.■

DEMOSTRACION. Basta demostrarlo para la norma euclídea. Sea $V\Sigma W^*$ una descomposición a valores singulares de A. Sea $\Sigma' = diag[\sigma_1 + d,...,\sigma_n + nd]$.
Si todos los valores singulares de A son iguales entonces para cualquier d>0 los autovalores de Σ' son distintos y positivos. En caso contrario, eligiendo $0 < nd <$ la menor diferencia positiva entre los σ_k, llegamos a la misma conclusión que los elementos de Σ' son distintos y positivos. Pongamos $B = V\Sigma'W^*$. Los valores singulares de B son los elementos de la diagonal de Σ' y $B = B(d) \to A$ si $d \downarrow 0$. En este caso, $\|A - B(d)\|_2 = \|\Sigma - \Sigma'(d)\|_2 \to 0$, QED.

OBSERVACIONES. a) Como los autovalores de una matriz son unitariamente invariantes, los valores singulares lo son también.
b) Si V es unitaria entonces A y AV tienen el mismo valor absoluto por la izquierda. En efecto, si $A = PU$ es la forma polar de A, entonces la forma polar de AV es $AV = PW$ con $W := UV$.
c) Sea P el valor absoluto por la izquierda de A. Entonces,

(25) $$\operatorname{Re} trA = \operatorname{Re} \sum \alpha_i \leq \sum |\alpha_i| = \sum \sigma_i = trP.$$

Como P es el valor absoluto por la izquierda tanto de A como de AV para V unitaria, resulta que la desigualdad (25) se mantiene si reemplazamos A por AV:

(25') $$\operatorname{Re} tr(AV) \leq \sum \sigma_i = trP \quad \forall\ V \text{ unitaria}.$$

Sin embargo, si ponemos V=U^* tendremos $AV = P$, por lo que para esa V:

$$\operatorname{Re} tr(AV) = trP.$$

O sea, V=U^* *maximiza* el primer miembro de (25') y hace a AV *semidefinida positiva*.

LA MATRIZ POSITIVA Y SU ESPECTRO
PARTE V

27. MATRICES DEFINIDAS POSITIVAS, PROPIEDADES Y DESIGUALDADES.
El producto ordinario de matrices definidas positivas no es definida positiva, en general. Por ejemplo, $AB=\begin{bmatrix}2 & 1\\1 & 1\end{bmatrix}\begin{bmatrix}2 & 1\\1 & 3\end{bmatrix}=\begin{bmatrix}5 & 5\\3 & 4\end{bmatrix}=C$. A y B son definidas positivas pero C no lo es pues es real pero no simétrica, aunque tiene autovalores positivos. El *producto de Hadamard*, $A\circ B$, en cambio, es definida positiva si A y B lo son.
DEFINICION 33. Sean $A=[a_{ij}]$, $B=[b_{ij}]$. El producto de Hadamard de A y B es $A\circ B=[a_{ij}b_{ij}]$.■

El ejemplo precedente puede todavía clarificarse notando que el producto usual de dos matrices hermitianas es hermitiana si y sólo si ellas conmutan. En cambio, el producto de Hadamard de hermitianas es siempre hermitiana.
TEOREMA 86 (T. de Schur del producto). a) Si A y B en $M_n(C)$ son semidefinidas positivas entonces $A\circ B$ es semidefinida positiva.
b) Si A y B son definidas positivas entonces $A\circ B$ es definida positiva.■
DEMOSTRACION. a) Sea rango(A)=k. Existe U unitaria tal que $A=U\Lambda U^*$ donde Λ es una matriz diagonal con los autovalores de A en su diagonal. Supongamos que los autovalores λ_j son positivos para $j=1,\dots,k$. Consideremos el vector $v_i=+\sqrt{\lambda_i}u_i$ donde $U=[u_1..u_n]$. Entonces, $\{v_1,\dots,v_k\}$ es una *familia ortogonal* y vale

(1) $$A=v_1v_1^*+\dots+v_kv_k^*.$$

Si m=rango(B) entonces $B=w_1w_1^*+\dots+w_mw_m^*$ y obtenemos

(2) $$A\circ B=\sum_{i,j=1}^{k,m}(v_i\circ w_j)(v_i\circ w_j)^*=\sum_{i,j=1}^{k,m}u_{ij}u_{ij}^*.$$

Cada uno de los sumandos es de la forma hh^*, h un vector, y es por lo tanto una matriz semidefinida positiva por lo que (2) es semidefinida positiva.
b) En este caso $k=m=n$ por lo que $\{v_1,\dots,v_n\}$ y $\{w_1,\dots,w_n\}$ son bases de C^n. Si $A\circ B$ fuera singular existiría x no nulo tal que $(A\circ B)x=0$. Por tanto,

$$x^*(A\circ B)x=\sum_{i,j=1}^{n}x^*(u_{ij}u_{ij}^*)x=\sum_{i,j=1}^{n}\left|x^*u_{ij}\right|^2=0.$$

Luego, cada término de la última suma se anula por lo que, para todo i,j,

$$\left|x^*u_{ij}\right|^2=\left|x^*(v_i\circ w_j)\right|^2=\left|(x\circ\overline{v}_i)^*w_j\right|^2=0.$$

Luego, el último parentesis es cero para todo i. Pero $x\circ\overline{v}_i=0\Rightarrow v_i^*x=0\Rightarrow x^*v_i=0$, para todo i, por lo que x*=x=0. Por tanto, $A\circ B$ es no singular, QED.

TEOREMA 87 (Fejér). A es semidefinida positiva si y sólo si $\sum_{i,j=1}^{n}a_{ij}b_{ij}\ge 0$ para toda $B=[b_{ij}]$ semidefinida positiva.■

DEMOSTRACION. Supongamos A y B semidefinidas positivas y $x=[1,\dots,1]^{tr}$. Por el teorema de Schur $A\circ B$ es semidefinida positiva por lo que

$$0\le x^*(A\circ B)x=\sum_{i,j=1}^{n}a_{ij}b_{ij}.$$

Recíprocamente, si $\sum_{i,j=1}^{n} a_{ij}b_{ij} \geq 0$ para toda $B=[b_{ij}]$ semidefinida positiva utilizando

$B=[\bar{x}_i x_j]=\begin{bmatrix}\bar{x}_1\\ \vdots \\ \bar{x}_n\end{bmatrix}[x_1 \cdots x_n]$ con $[x_1 \cdots x_n]^{tr} \in C^n$ obtenemos

$0 \leq \sum_{i,j=1}^{n} a_{ij}b_{ij} = \sum_{i,j=1}^{n} a_{ij}\bar{x}_i x_j = x^* Ax$ y A es semidefinida positiva, QED.

TEOREMA 88 (desigualdad de Minkowski). Sean A y B definidas positivas. Entonces,

(3) $$(\det(A+B))^{1/n} \geq (\det(A))^{1/n} + (\det(B))^{1/n}. \blacksquare$$

DEMOSTRACION. Recordemos que $A^{1/2}, A^{-1/2}$ son definidas positivas y que *C=HBH* es definida positiva si *H* es hermitiana no singular. Además, notemos que *B* y *C* se determinan mutuamente: $B=H^{-1}CH^{-1}$.

Entonces, (3) vale si y sólo si

$(\det A^{-1/2})^{1/n}(\det(A+B))^{1/n}(\det A^{-1/2})^{1/n} \geq (\det I)^{1/n} + (\det(A^{-1/2}BA^{-1/2}))^{1/n}$. O sea, si y sólo si $(\det(I+A^{-1/2}BA^{-1/2}))^{1/n} \geq 1 + (\det(A^{-1/2}BA^{-1/2}))^{1/n}$. De lo dicho sigue que (3) es válida si y sólo si, para toda *B* definida positiva,

(4) $$(\det(I+B))^{1/n} \geq 1 + (\det(B))^{1/n}$$

Si $\{\lambda_k\}$ es la familia de autovalores (positivos) de *B*, $\{1+\lambda_k\}$ será la familia de autovalores de *I+B*. (4) se escribe entonces como

(5) $$\prod_{i=1}^{n}(1+\lambda_i) \geq (1+\sqrt[n]{\lambda_1 \ldots \lambda_n})^n.$$

Pero

(6) $(1+\sqrt[n]{\lambda_1 \ldots \lambda_n})^n = \sum \binom{n}{j}\left(\prod \lambda_i{}^j\right)^{1/n}$,

(7) $\prod_{i=1}^{n}(1+\lambda_i) = 1+\sum \lambda_i + \sum_{i<k}\lambda_i\lambda_k + \sum_{i<k<h}\lambda_i\lambda_k\lambda_h + \ldots = \sum_{0}^{n}\Sigma(j)$.

$\Sigma(j)$ representa la suma en (7) con términos de j factores. Ella tiene $\binom{n}{j}$ sumandos y el producto de todos ellos es igual a $\prod \lambda_k^{\binom{n-1}{j-1}} = \left(\prod \lambda_k\right)^{\binom{n-1}{j-1}}$. De la desigualdad entre la media geométrica y la aritmética sigue que

$\dfrac{\Sigma(j)}{\binom{n}{j}} \geq \left(\prod \lambda_k\right)^{\binom{n-1}{j-1}/\binom{n}{j}} = \left(\prod \lambda_k\right)^{\frac{j}{n}}$, por lo que $\Sigma(j) \geq \binom{n}{j}\left(\prod \lambda_k{}^j\right)^{\frac{1}{n}}$ y $(7) \geq (6)$, QED.

El teorema de Minkowski, como el siguiente debido a Hadamard, versan sobre *desigualdades* de expresiones *numéricas* asociadas a las matrices que, en general, no tienen relación con las desigualdades matriciales que puedan presentarse directamente entre ellas.

TEOREMA 89 (desigualdad de Hadamard). Sea $A=[a_{ij}] \in M_n$ definida positiva. Entonces,

(8) $$\det A \le \prod_1^n a_{ii}.$$

La igualdad se presenta si y sólo si A es diagonal.■

DEMOSTRACION. Por ser definida positiva A es no singular y vale $a_{ii} > 0$ para todo i. Sea $D = diag\lfloor a_{11}^{-1/2},...,a_{nn}^{-1/2}\rfloor$. Luego,

$$\det A \le a_{11}...a_{nn} \Leftrightarrow \det DAD \le 1.$$

Por tanto, bastará probar (8) para matrices B (=DAD) definidas positivas, con elementos diagonales iguales a 1. En este caso (8) sigue de la relación entre media geométrica y media aritmética. En efecto, si $\{\lambda_n\}$ son los autovalores (positivos) de B,

(9) $$\det B = \prod_1^n \lambda_i \le \left(\frac{\sum \lambda_i}{n}\right)^n = \left(\frac{trB}{n}\right)^n = 1 = \prod_1^n b_{ii}.$$

En la desigualdad geométrico-aritmética anterior vale la igualdad si y sólo si $\lambda_i = 1$ $\forall i$, esto es, si y sólo si B, hermitiana, es unitariamente diagonalizable a I. Pero esto ocurre si y sólo si B=I. O sea: en (9) $\det B = 1$ si y sólo si $B = I$.

Luego, el signo = se presenta en (8) si y sólo si $DAD =: B = I$, y esto ocurre si y sólo si $A = D^{-1}D^{-1}$ = matriz diagonal, QED.

La desigualdad de Hadamard puede escribirse así: $(\det A)\prod_1^n 1 \le \det A \circ I$.

Se generaliza de la siguiente forma que no demostraremos.

TEOREMA 90 (desigualdad de Oppenheim). Si A y B son definidas positivas entonces

(10) $$(\det A)\prod_1^n b_{ii} \le \det A \circ B.■$$

Una consecuencia del teorema 89 es el siguiente resultado también llamado *desigualdad de Hadamard*.

TEOREMA 91. Sea $B = \lfloor b_{ij} \rfloor \in M_n$ arbitraria, no singular. Entonces,

(11) $$|\det B| \le \prod_{i=1}^n \left(\sum_{j=1}^n |b_{ij}|^2\right)^{1/2}.$$

La igualdad se presenta si y sólo si las filas de B son ortogonales.■

DEMOSTRACION. Sea A:=BB* y apliquemos (8):

$$|\det A| = |\det BB^*| = |\det B|^2 \le \prod \sum_j |b_{ij}b^*_{ji}| = \prod_i \sum_j |b_{ij}|^2.$$

Sigue (11) extrayendo la raíz cuadrada a esta expresión. Las filas de B son ortogonales exactamente cuando A es diagonal, QED.

Si hubiéramos usado A:=B*B habríamos obtenido $|\det B| \le \prod_{j=1}^n \left(\sum_{i=1}^n |b_{ij}|^2\right)^{1/2}$ con igualdad si y sólo si las columnas de B son ortogonales.

TEOREMA 92 (Ostrowski-Taussky). Sea A tal que $\frac{A+A^*}{2}$ sea definida positiva. Entonces, $\det((A+A^*)/2) \le |\det A|$. Vale la igualdad si y sólo si A es hermitiana.■

No demostraremos este resultado. Veremos a continuación algunas *desigualdades matriciales* que se originan en la definida positividad.

28. DESIGUALDADES MATRICIALES.

DEFINICION 34. Sean A y B matrices hermitianas. Escribimos $A \trianglerighteq B$ si $A-B$ es semidefinida positiva y $A \succ B$ si $A-B$ es definida positiva.■

$\trianglerighteq$ define un orden parcial (no total) entre las matrices hermitianas pues es una relación reflexiva y transitiva. Vale $(A \trianglerighteq B) \wedge (B \trianglerighteq A) \Rightarrow A = B$; en efecto, la matriz hermitiana $A-B$ debe tener todos sus autovalores nulos. Sin embargo *no* vale que $A \trianglerighteq B$ y $A \neq B \quad \Rightarrow A \succ B$, (ejemplo: $A = \begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix}, B = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$).

Sea T arbitraria. Entonces,

(12) $$A \trianglerighteq B \Rightarrow T^*AT \trianglerighteq T^*BT.$$

En efecto, $x^*(T^*AT - T^*BT)x = (x^*T^*)(A-B)(Tx) \geq 0$.

Sea T no singular. Entonces,

(13) $$A \succ B \Rightarrow T^*AT \succ T^*BT.$$

En efecto, $x \neq 0 \Leftrightarrow Tx \neq 0$, por lo que $x^*(T^*AT - T^*BT)x = (Tx)^*(A-B)(Tx) > 0$ si $x \neq 0$.

TEOREMA 93. a) Sean A definida positiva y B semidefinida positiva. Entonces,

(14) $$A \trianglerighteq B \Leftrightarrow \rho(BA^{-1}) \leq 1, \quad A \succ B \Leftrightarrow \rho(BA^{-1}) < 1.$$

Sean A y B definidas positivas. Valen,

b) $A \trianglerighteq B \Leftrightarrow B^{-1} \trianglerighteq A^{-1}$,

c) $A \trianglerighteq B \Rightarrow \det A \geq \det B, tr\, A \geq tr\, B$,

d) $A \trianglerighteq B \Rightarrow \lambda_j(A) \geq \lambda_j(B)$ si los respectivos autovalores son ordenados en el mismo orden creciente.■

DEMOSTRACION. a) Existe una matriz K no singular tal que $KAK^* = I$. La matriz hermitiana KBK^* es diagonalizable por medio de una matriz unitaria U: $UKBK^*U^*=D$ matriz diagonal real. Como $UKAK^*U^* = UU^* = I$ hemos probado que existe una matriz no singular C tal que $A=CIC^*$ y $B=CDC^*$. [1]

Entonces, $BA^{-1} = CDC^*C^{*-1}C^{-1} = CDC^{-1}$. Por tanto los autovalores de BA^{-1} son los elementos diagonales de D. Veamos la primera de las desigualdades (14). $A \trianglerighteq B$ si y sólo si $C(I-D)C^*$ es semidefinida positiva y esto sucede si y sólo si $I - D \trianglerighteq 0$. Pero esto equivale a $I - D \geq 0$. Es decir, ningún autovalor de BA^{-1} debe superar a 1 y sigue que $A \trianglerighteq B \Leftrightarrow \rho(BA^{-1}) \leq 1$. Análogamente se llega a la segunda desigualdad (14).

Las siguientes proposiciones b), c) y d) son consecuencias de a).

b) En este caso existe B^{-1}. De $A \trianglerighteq B \Leftrightarrow \rho(BA^{-1}) \leq 1$ obtenemos

$$B^{-1} \trianglerighteq A^{-1} \Leftrightarrow \rho(A^{-1}B) \leq 1.$$

Pero $\rho(A^{-1}B) = \rho(BA^{-1})$ como puede verse, por ejemplo, usando el Corolario 2 del Teorema 30, por lo que $A \trianglerighteq B \Leftrightarrow B^{-1} \trianglerighteq A^{-1}$.

c) sigue de d). d) Es una consecuencia del teorema de Weyl. Probemos directamente las desigualdades. Basta ver que si A es definida positiva y B semidefinida positiva los autovalores ordenados verifican $\lambda_k(A+B) \geq \lambda_k(A)$.

---

[1] Éste es un caso particular del siguiente teorema: si A y B son *hermitianas* y H es una combinación lineal real de ellas, *definida positiva*, entonces existe una matriz *no* singular, C, tal que C^*AC y C^*BC son ambas *diagonales*. Obsérvese que si quisiéramos diagonalizar simultáneamente por semejanza necesitaríamos que las matrices conmuten.

$$\lambda_k(A) = \min_{w_1,\dots,w_{n-k}\in C^n} \max_{x\perp w_1,\dots,w_{n-k}} \frac{x^*Ax}{x^*x} \le \min_{w_1,\dots,w_{n-k}\in C^n} \max_{x\perp w_1,\dots,w_{n-k}} \left(\frac{x^*Ax}{x^*x} + \lambda_1(B)\right) \text{ pues}$$

$\lambda_i(B) \ge 0$. Por el teorema de Rayleigh-Ritz la última expresión es menor o igual a

$$\min_{w_1,\dots,w_{n-k}\in C^n} \max_{x\perp w_1,\dots,w_{n-k}} \left(\frac{x^*Ax}{x^*x} + \frac{x^*Bx}{x^*x}\right) = \min_{w_1,\dots,w_{n-k}\in C^n} \max_{x\perp w_1,\dots,w_{n-k}} \left(\frac{x^*(A+B)x}{x^*x}\right) = \lambda_k(A+B),$$

QED.

Las matrices definidas positivas como factores tienen un comportamiento muy interesante.

TEOREMA 94. Sean A definida positiva y B hermitiana. Entonces,
i) AB es diagonalizable con autovalores reales,
ii) inercia(AB)=inercia(B).■

DEMOSTRACION. $A^{-1/2}(AB)A^{1/2} = A^{1/2}BA^{1/2} = H$ es hermitiana, por lo tanto diagonalizable: $H = SDS^{-1}$, D diagonal real con los autovalores de H en su diagonal. Luego, AB es semejante a la matriz diagonal D. En particular tenemos: inercia(AB)=inercia(D).
Por otra parte, $H = A^{1/2}BA^{1/2} = A^{1/2}B(A^{1/2})^*$ pues $A^{1/2}$ es definida positiva. Por el teorema de Sylvester H y B tienen la misma tripla de inercia e igual a inercia(D) que es igual a inercia(AB) como vimos, QED.

NB. Lo notable en este análisis es que cualquier matriz C diagonalizable a una D real es el producto de una matriz definida positiva y una matriz hermitiana.
En efecto, $C = SDS^{-1} = S(S^*S^{*-1})DS^{-1} = (SS^*)((S^{-1})^*DS^{-1}) = AB$ donde $A = SS^*$ es definida positiva y $B = (S^{-1})^*DS^{-1}$ es hermitiana, qed.

El punto iii) del siguiente teorema es conocido como *desigualdad de Fischer* y generaliza a la desigualdad de Hadamard, la que puede deducirse de aquella.

TEOREMA 95. Sea $H = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix}$ hermitiana, A y C cuadradas, no vacias. Entonces,
i) H es definida positiva si y sólo si A es definida positiva y $C \succ B^*A^{-1}B$.
ii) H es definida positiva si y sólo si A es definida positiva y $\rho(B^*A^{-1}BC^{-1}) < 1$.
iii) Si H es definida positiva entonces $\det H \le (\det A)(\det C)$.

DEMOSTRACION. i) $\equiv$ ii) por el teorema 93.

i) Supongamos H definida positiva. Entonces, A, C y $H^{-1} = \begin{bmatrix} M & N \\ Q & R \end{bmatrix}$ son definidas positivas. A partir de $\begin{bmatrix} M & N \\ Q & R \end{bmatrix}\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$ obtenemos

(15) $$H^{-1} = \begin{bmatrix} (A-BC^{-1}B^*)^{-1} & N \\ Q & (C-B^*A^{-1}B)^{-1} \end{bmatrix}.$$

En efecto, por ejemplo, de $\begin{cases} MA+NB^* = I \\ MB+NC = 0 \end{cases}$ despejamos N y obtenemos $M(A-BC^{-1}B^*) = I$.

Por tanto, $A-BC^{-1}B^{*}, (A-BC^{-1}B^{*})^{-1}, C-B^{*}A^{-1}B, (C-B^{*}A^{-1}B)^{-1}$ son definidas positivas. En consecuencia, $C \succ B^{*}A^{-1}B$. Veamos la recíproca. Supongamos A y $C-B^{*}A^{-1}B$ definidas positivas. Vale

$$(16) \qquad \begin{bmatrix} I & 0 \\ -B^{*}(A^{-1})^{*} & I \end{bmatrix} \begin{bmatrix} A & B \\ B^{*} & C \end{bmatrix} \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & C-B^{*}A^{-1}B \end{bmatrix} = K.$$

Esto muestra que H y K son *congruentes. Luego, como K es definida positiva también lo es H.

iii) De (16) obtenemos $\det H = \det K = (\det A)(\det(C-B^{*}A^{-1}B))$. Pero $C=(C-B^{*}A^{-1}B)+B^{*}A^{-1}B$ por lo que $C \rhd C-B^{*}A^{-1}B \rhd 0$. Luego, por el teorema 93, $\det(C-B^{*}A^{-1}B) \leq \det C$, QED.