Análisis Matricial.

Jorge Antezana y Demetrio Stojanoff

Índice General

1 Preliminares.

1.1 Generalidades

A lo largo de este trabajo consideraremos a $\mathbb{C}$ o $\mathbb{R}$ como cuerpo de escalares. Notaremos $\mathcal{M}_n(\mathbb{C}) = \mathcal{M}_n = \mathbb{C}^{n\times n}$, las matrices cuadradas de $n \times n$ sobre $\mathbb{C}$. Análogamente, $\mathcal{M}_n(\mathbb{R}) = \mathbb{R}^{n\times n}$. Para denotar las entradas de una matriz $\in \mathcal{M}_n(\mathbb{C})$, usaremos indistintamente, por conveniencia del contexto, las notaciones $A = (A_{ij})_{ij}$ o $A = (a_{ij})_{ij}$.

Los vectores de $\mathbb{C}^n$ serán pensados como vectores **columna**, es decir que identificamos $\mathbb{C}^n$ con $\mathbb{C}^{n\times 1}$. Sin embargo, a lo largo del texto los describiremos como una fila (estilo $x = (x_1, \dots, x_n) \in \mathbb{C}^n$), para ahorrar espacio.

Si $A \in \mathcal{M}_n(\mathbb{C})$, lo pensaremos también como un operador (o transformación lineal) sobre $\mathbb{C}^n$ por multiplicación: si $x \in \mathbb{C}^n$, entonces $A(x) = Ax$ (el producto standard de matrices). Algunas veces pensaremos a ciertas matrices como operando en espacios vectoriales más generales. Por ejemplo, si $\mathcal{S} \subseteq \mathbb{C}^n$ es un subespacio y $A \in \mathcal{M}_n(\mathbb{C})$ verifica que $A(\mathcal{S}) \subseteq \mathcal{S}$, entonces se puede pensar a A (o su restricción a $\mathcal{S}$) como un operador en $\mathcal{S}$. En tal caso diremnos que "pensamos" a $A|_{\mathcal{S}} \in L(\mathcal{S})$, donde $L(\mathcal{S})$ denotaría al conjunto de operadores lineales en $\mathcal{S}$.

En $\mathbb{C}^n$ consideraremos el producto interno (o escalar) común, dado por

$$\langle x, y\rangle = \sum_{k=1}^{n} x_k \, \overline{y_k} \ , \quad x, y \in \mathbb{C}^n.$$

Es claro que $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ verifica las propiedades de un tal producto: Dados $v, v, w \in \mathbb{C}^n$ y $\lambda \in \mathbb{C}$, entonces

1. $\langle v, v\rangle \geq 0$ y $\langle v, v\rangle = 0$ si y sólo si $v = 0$.
2. $\langle u, v\rangle = \overline{\langle v, u\rangle}$.
3. $\langle v, (u + w)\rangle = \langle v, u\rangle + \langle v + w\rangle$.
4. $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$.

Dado $x \in \mathbb{C}^n$, definiremos su **norma Euclídea**, a la usual:

$$\|x\| = \|x\|_2 = \langle x, x\rangle^{1/2} = (\sum_{k=1}^{n} |x_k|^2)^{1/2}.$$

Muchas veces consideraremos otras normas de vectores y matrices. Por ello damos una definición general:

Definición 1.1.1. Sea $K = \mathbb{C}$ o $\mathbb{R}$ y $\mathbb{V}$ un K-espacio vectorial. Una **norma** en $\mathbb{V}$ es una función $N : \mathbb{V} \to \mathbb{R}$ que verifica las siguientes condiciones: Dados $v, v \in \mathbb{V}$ y $\lambda \in K$,

1. $N(v) \geq 0$ y $N(v) = 0$ si y sólo si $v = 0$.
2. $N(u + v) \leq N(u) + N(v)$.
3. $N(\lambda v) = |\lambda| \ N(v)$. $\triangle$

Cuando una tal norma proviene de un producto interno, diremos que el par $(\mathbb{V}, N)$, o bien $(\mathbb{V}, \langle \cdot, \cdot\rangle)$ es un espacio de Hilbert (ojo, acá se asume que $\dim \mathbb{V} < \infty$). Usualmente usaremos letras H o K para tales espacios y notaremos por $L(H, K)$ al espacio de operadores lineales de H en K. Si $H = K$, escribimos $L(H)$ en lugar de $L(H, H)$. Si $A \in L(H, K)$ notaremos por $Nu(A)$ a su núcleo y $R(A)$ a su imagen. Además notaremos $r(A) = \dim R(A)$ al rango (columna) de A.

Definición 1.1.2. Sea H un espacio de Hilbert. Los vectores $x_1, ..., x_k \in H$ forman un conjunto **ortogonal** cuando $\langle x_i, x_j \rangle = 0$, si $i \neq j$. Si además los vectores están normalizados, es decir $\|x\|^2 = \langle x_i, x_i \rangle = 1$, $\leq i \leq k$, entonces el conjunto se dice **ortonormal**. Usaremos las siglas **bon** para denotar a un base ortonormal de H. $\triangle$

Definición 1.1.3. Sean H y K espacios de Hilbert y sea $A \in L(H, K)$. Se llama **adjunto** de A al único operador $A^* \in L(K, H)$ que satisface

$$\langle Ax, z \rangle_K = \langle x, A^* z \rangle_H \ , \quad x \in H, \ z \in K.$$

$\triangle$

Observación 1.1.4. Si, en particular, $A \in L(H)$, entonces A^* también está en $L(H)$. Para una bon fija $\mathcal{B} = \{e_1, \ldots, e_n\}$ de H, se identifica a los operadores de $L(H)$ con matrices en $\mathcal{M}_n(\mathbb{C})$ vía

$$A_{ij} = \langle Ae_j, e_i \rangle \ , \quad 1 \leq i, j \leq n.$$

Con esta identificación, la matriz de A^* es la traspuesta conjugada de la matriz de A. Es decir que $A^*_{ij} = \overline{A_{ji}}$, $1 \leq i, j \leq n$.

Definición 1.1.5. Dado $A \in L(H)$ un operador en un espacio de Hilbert, decimeos que A es:

1. **Hermitiano** si $A = A^*$.
2. **anti-Hermitiano** si $A = -A^*$.
3. **unitario** si $AA^* = A^*A = I$.
4. **normal** si $AA^* = A^*A$.
5. **definido positivo** si $\langle Ax, x \rangle > 0$ para todo $x \in H$. En tal caso de escribe $A > 0$.
6. **semidefinido positivo** si $\langle Ax, x \rangle \geq 0$ para todo $x \in H$. En tal caso de escribe $A \geq 0$.

Los mismos nombres tendrán las matrices de $\mathcal{M}_n(\mathbb{C})$, al ser pensadas cono operadores en $H = \mathbb{C}^n$ con el producto escalar y la norma usuales. Además usaremos las siguientes notaciones:

1. $\mathcal{H}(n) = \{A \in \mathcal{M}_n(\mathbb{C}) : A = A^*\}$.
2. $\mathcal{U}(n) = \{U \in \mathcal{M}_n(\mathbb{C}) : U \text{ es unitaria } \}$.
3. $\mathcal{M}_n(\mathbb{C})^+ = \{A \in \mathcal{M}_n(\mathbb{C}) : A \geq 0\} \subseteq \mathcal{H}(n)$.
4. $\mathcal{G}l(n) = \{A \in \mathcal{M}_n(\mathbb{C}) : A \text{ es invertible } \}$. $\triangle$

Definición 1.1.6. Se llama espectro de una matriz $A \in \mathcal{M}_n(\mathbb{C})$ al conjunto de todos los autovalores de A:

$$\sigma(A) = \{\lambda \in \mathbb{C} : \lambda \text{ es autovalor de } A\} = \{\lambda \in \mathbb{C} : Nu(A - \lambda I) \neq \{0\} \ \}.$$

Notar que $A \in \mathcal{G}l(n)$ si y sólo si $0 \notin \sigma(A)$. $\triangle$

Observación 1.1.7. Se sabe que los autovalores de $A \in \mathcal{M}_n(\mathbb{C})$ son las raíces del polinomio característico de A. Este polinomio tiene grado n, pero puede tener raíces múltiples, por lo que $\sigma(A)$ puede tener menos de n elementos (en tanto conjunto, sus elementos sólo pueden contarse de a uno). Muchas veces es necesario usar a cada $\lambda \in \sigma(A)$ tantas veces como multiplicidad tiene como raíz del característico. Para hacer eso, diremos que

$$\text{“los autovalores de } A \text{ son } \lambda_1, \dots, \lambda_n\text{”} ,$$

donde estaremos repitiendo cada autovalor de A tantas veces como multiplicidad tiene en el polinomio característico. Por eso quedan n.

Definición 1.1.8. El **radio espectral** de una matriz $A \in \mathcal{M}_n(\mathbb{C})$ se define como

$$\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}.$$

△

1.2 Matrices Unitarias

Las matrices unitarias son de gran importancia ya que, en muchos casos, son la herramienta principal para obtener distintos resultados en el análisis matricial. En esta sección veremos distintas formas de caracterizarlas y veremos también cómo intervienen en una descomposición muy especial para matrices cuadradas dada por el teorema de Schur.

Definición 1.2.1. Se dice que A es unitariamente equivalente a B y se nota $A \simeq B$ si existe $U \in \mathcal{U}(n)$ tal que $A = U^*BU$. △

Teorema 1.2.2. *Si $A = [a_{ij}]$ y $B = [b_{ij}] \in \mathcal{M}_n(\mathbb{C})$ son unitariamente equivalentes, entonces*

$$\sum_{i,j=1}^{n} |b_{ij}|^2 = \sum_{i,j=1}^{n} |a_{ij}|^2.$$

Demostración. Se usa que $\sum\limits_{i,j=1}^{n} |a_{ij}|^2 = tr(A^*A)$ (Ejercicio: verificarlo). Entonces es suficiente probar que $tr(B^*B) = tr(A^*A)$. Sabemos que $B = U^*AU$ entonces

$$tr(B^*B) = tr(U^*A^*UU^*AU) = tr(U^*A^*AU) = tr(A^*A).$$

La última igualdad se deduce de que $tr(XY) = tr(YX)$ para todo par $X, Y \in \mathcal{M}_n(\mathbb{C})$. □

Teorema 1.2.3. *Si $U \in \mathcal{M}_n(\mathbb{C})$, las siguientes afirmaciones son equivalentes:*

a) U es unitaria.

b) U es no singular y $U^* = U^{-1}$.

c) $UU^* = I$.

d) U^* *es unitaria.*

e) Las columnas de U *forman una bon.*

f) Las filas de U *forman una bon.*

g) Para todo $x \in \mathbb{C}^n$, $\|Ux\|_2 = \|x\|_2$ *(o sea,* U *es una isometría).*

Demostración. Ejercicio. Para probar que g) implica lo demás, se usa que, si $A \in \mathcal{M}_n(\mathbb{C})$ y $\langle Ax, x\rangle = 0$ para todo $x \in \mathbb{C}^n$, entonces $A = 0$. Ojo que esto es falso en $\mathcal{M}_n(\mathbb{R})$. □

El siguiente resultado es de gran importancia ya que afirma que cualquier matriz $A \in \mathcal{M}_n(\mathbb{C})$ es unitariamente equivalente a una matriz triangular superior que contiene en su diagonal los autovalores de A.

Teorema 1.2.4 (Triangularización Unitaria de Schur). *Dada* $A \in \mathcal{M}_n(\mathbb{C})$ *con autovalores* $\lambda_1, ..., \lambda_n$ *dispuestos en cualquier orden (y contados con multiplicidad). Entonces existen matrices* $U \in \mathcal{U}(n)$ *y* $T \in \mathcal{M}_n(\mathbb{C})$ *triangular superior (o sea que* $t_{ij} = 0$ *si* $i > j$*) que verifican:*

1. $t_{ii} = \lambda_i$, $1 \le i \le n$ *y*
2. $A = UTU^*$.

Si $A \in \mathcal{M}_n(\mathbb{R})$*, el teorema sigue valiendo (entre matrices reales) si* $\sigma(A) \subseteq \mathbb{R}$*. Además, si* $B \in \mathcal{M}_n(\mathbb{C})$ *conmuta con* A*, se las puede triangular a ambas con la misma matriz unitaria, pero respetando un orden dado para los autovalores en uno solo de los casos (hay que elegir* A *o* B*).*

Demostración. La prueba del teorema la realizaremos por inducción sobre la dimensión n. Si $n = 1$, el resultado es trivial. Si $n > 1$, tomemos $x_1 \in Nu(A - \lambda_1 I)$ con $\|x_1\| = 1$. Completamos a una bon con vectores $x_2, \ldots, x_n$, y los ponemos en las columnas de una matriz U_1, que resulta unitaria. Es fácil ver que

$$U_1^* A U_1 = \begin{pmatrix} \lambda_1 & * \\ 0 & A_2 \end{pmatrix},$$

donde $A_2 \in \mathcal{M}_{n-1}(\mathbb{C})$, con autovalores $\lambda_2, \ldots, \lambda_n$. Por HI se triangula a A_2 con una matriz unitaria $V \in \mathcal{M}_{n-1}(\mathbb{C})$, que se extiende a otra unitaria $U_2 \in \mathcal{M}_n(\mathbb{C})$ poniendo un 1 en el lugar $1, 1$ y ceros en los costados. Es fácil verificar que $U = U_1 U_2$ triangula a A de la forma requerida.

El caso real sale igual. Notar que el $x_1 \in \mathbb{R}^n$ existe si $\lambda_1 \in \mathbb{R}$. El caso de dos matrices que conmutan se deduce de que $Nu(A - \lambda_1 I)$ es invariante para B, por lo que el vector x_1 se puede elegir como un autovector de B (aunque no se sabe cuales de los autovalores de B pueden elegirse). El resto de la prueba sigue igual. □

Corolario 1.2.5. *Sea* $A \in \mathcal{M}_n(\mathbb{C})$ *con autovalores* $\lambda_1, ..., \lambda_n$ *dispuestos en cualquier orden. Entonces* $\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i$ *y* $\det A = \prod_{i=1}^{n} \lambda_i$.

Demostración. Por el teorema sabemos que podemos escribir $A = U^*TU$, con $U \in \mathcal{U}(n)$ y T triangular superior, con los autovalores de A en su diagonal. Luego $\operatorname{tr} A = \operatorname{tr} T$ y $\det A = \det T$. □

Corolario 1.2.6. *Sea* $U \in \mathcal{U}(n)$. *Entonces* $|\det U| = 1$.

Demostración. Basta verificar que $\sigma(U) \subseteq \{z \in \mathbb{C} : |z| = 1\}$, lo que es un ejercicio fácil. □

1.3 Matrices normales

Recordemos que una matriz $A \in \mathcal{M}_n(\mathbb{C})$ es normal si $A^*A = AA^*$, es decir si A conmuta con su adjunta.

Definición 1.3.1. Si $A \in \mathcal{M}_n(\mathbb{C})$ es unitariamente equivalente a una matriz diagonal, entonces se dice que A es **unitariamente diagonalizable**. △

Notacón: Si $a = (a_1, \dots, a_n) \in \mathbb{C}^n$, denotaremos por $\operatorname{diag}(a)$ a la matriz diagonal

$$\operatorname{diag}(a) = \operatorname{diag}(a_1, \dots, a_n) = \begin{pmatrix} a_1 & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & a_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{C}).$$

Teorema 1.3.2. *Si* $A = [a_{ij}] \in \mathcal{M}_n(\mathbb{C})$ *tiene autovalores* $\lambda_1, ..., \lambda_n$, *las siguientes condiciones son equivalentes:*

a) A es normal.

b) A es unitariamente diagonalizable.

c) $\sum_{i,j=1}^{n} |a_{ij}|^2 = \sum_{i=1}^{n} |\lambda_i|^2$

d) Para todo $x \in \mathbb{C}^n$, $\|Ax\| = \|A^*x\|$.

Demostración. La equivalencia entre a) y d) la obtenemos del siguiente hecho: para cada $x \in \mathbb{C}^n$,

$$\|Ax\|^2 = \langle A^*Ax, x\rangle \quad \text{y} \quad \|A^*x\|^2 = \langle AA^*x, x\rangle.$$

Recordar que si $B \in \mathcal{M}_n(\mathbb{C})$ cumple que $\langle Bx, x\rangle = 0$ para todo $x \in \mathbb{C}^n$, entonces $B = 0$.
Por el Teorema de Schur existe una matriz unitaria U y una matriz triangular superior T tal que $T = U^*AU$. Además como A es normal, T es normal, es decir $T^*T = TT^*$ y

esto más el hecho de que T es triangular superior implican que T es una matriz diagonal (Ejercicio: verificarlo). Entonces A es unitariamente diagonalizable. Por lo tanto a) implica b). Recíprocamente si A es unitariamente diagonalizable, $A = U^*TU$ con U unitaria y T diagonal, por lo tanto T es normal. Entonces A es normal. Luego queda demostrado que a) y b) son equivalentes.

Si A es unitariamente diagonalizable, $A = U^*DU$ con U unitaria y D diagonal y por el teorema (1.2.2) tenemos que $\sum_{i,j=1}^{n} |a_{ij}|^2 = \sum_{i=1}^{n} |\lambda_i|^2$.
Asumamos la condición c). El teorema de Schur nos dice que A es unitariamente equivalente a una matriz T triangular superior, entonces

$$\sum_{i=1}^{n} |\lambda_i|^2 = \sum_{i,j=1}^{n} |a_{ij}|^2 = \sum_{i=1}^{n} |\lambda_i|^2 + \sum_{i<j} |t_{ij}|^2.$$

Por lo tanto $t_{ij} = 0$ para $i < j$ y T es diagonal. $\square$

Corolario 1.3.3. *Si $A \in \mathcal{M}_n(\mathbb{C})$ es normal, entonces $\|A\|_{sp} = \rho(A)$. Recordemos que $\|A\|_{sp} = \max\{\|Ax\| : \|x\| = 1\}$.*

Demostración. Es fácil verificar que matrices unitariamente equivalentes tienen la misma norma espectral. Por el Teorema 1.3.2, A es unitariamente equivalente a $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Finalmente, una fácil cuenta muestra que la norma espectral de una matriz diagonal es el máximo de los módulos de sus elementos diagonales. $\square$

1.4 Matrices Hermitianas

La práctica nos enseña que, por lo general, no es fácil calcular los autovalores de una matriz, pero en muchos casos es suficiente saber que los autovalores están en un intervalo especificado. El objetivo de esta sección es estudiar algunas de las principales características que distinguen a las matrices Hermitianas y conocer principios variacionales que se utilizan para localizar el espectro de una matriz Hermitiana sin la necesidad de conocer los autovectores asociados en forma exacta. Recordemos las notaciones $\mathcal{H}(n) = \{A \in \mathcal{M}_n(\mathbb{C}) : A = A^*\}$ y $\mathcal{M}_n(\mathbb{C})^+ = \{A \in \mathcal{H}(n) : A \geq 0\}$.

Teorema 1.4.1. *Si $A \in \mathcal{H}(n)$, entonces A es unitariamente diagonalizable y $\sigma(A) \subseteq \mathbb{R}$.*

Demostración. Se deduce del Teorema 1.3.2 $\square$

Definición 1.4.2. Sea $A \in \mathcal{H}(n)$. Por el Teorema anterior, $\sigma(A) \subseteq \mathbb{R}$. Por lo tanto, sus autovalores pueden ordenarse usando el orden de $\mathbb{R}$. En adelante usaremos las siguientes notaciones:

1. $\lambda(A) = (\lambda_1(A), \ldots, \lambda_n(A))$ es el vector de autovalores de A ordenados en forma **creciente**.

2. $\mu(A) = (\mu_1(A), \dots, \mu_n(A))$ es el vector de autovalores de A ordenados en forma **decreciente**, es decir $\mu_k(A) \geq \mu_{k+1}(A)$, $1 \leq k \leq n-1$. También $\mu_k(A) = \lambda_{n-k+1}(A)$.

3. Se llamarán
$$\lambda_{min}(A) = \lambda_1(A) = \mu_n(A) = \min \sigma(A)$$
y, análogamente,
$$\lambda_{max}(A) = \lambda_n(A) = \mu_1(A) = \max \sigma(A) .$$

Así, cuando escribamos $\lambda_i(A)$ o, directamente λ_i (si el contexto es claro) estaremos asumiendo que al enumerar los autovalores de A lo hemos hecho en forma creciente. Y en forma decreciente si escibimos $\mu_i(A)$ o μ_i. $\triangle$

Para matrices generales la única caracterización conocida de sus autovalores es que son las raíces del polinomio característico de la matriz. Pero cuando las matrices son Hermitianas, el hecho de poder establecer un orden entre ellos nos permite obtener caracterizaciones más interesantes. El próximo teorema describe el máximo y el mínimo de los autovalores de la matriz en función de las expresiones $\dfrac{\langle Ax, x\rangle}{\langle x, x\rangle}$ $(0 \neq x \in \mathbb{C}^n)$, conocidas como cocientes de Rayleig-Ritz.

Teorema 1.4.3 (Rayleigh-Ritz). *Sea $A \in \mathcal{H}(n)$. Entonces*

1. *Para todo $x \in \mathbb{C}^n$, $\lambda_{min}(A)\|x\|^2 \leq \langle Ax, x\rangle \leq \lambda_{max}(A)\|x\|^2$.*

2. $\lambda_{max}(A) = \lambda_n(A) = \max\limits_{x\neq 0} \dfrac{\langle Ax, x\rangle}{\langle x, x\rangle} = \max\limits_{\|x\|=1} \langle Ax, x\rangle$.

3. $\lambda_{min}(A) = \lambda_1(A) = \min\limits_{x\neq 0} \dfrac{\langle Ax, x\rangle}{\langle x, x\rangle} = \min\limits_{\|x\|=1} \langle Ax, x\rangle$.

En particular, $A \in \mathcal{M}_n(\mathbb{C})^+$ si y sólo si $\lambda_{min}(A) \geq 0$.

Demostración. Como A es Hermitiana, es unitariamente diagonalizable, es decir $A = U^*DU$, con $U \in \mathcal{U}(n)$ unitaria y $D = \operatorname{diag}(\lambda_1(A), ..., \lambda_n(A))$. Notar que $\langle Ax, x\rangle = \langle D\,Ux, Ux\rangle$, $x \in \mathbb{C}^n$. Como U es una isometría,
$$U\Big(\{x \in \mathcal{M}_n(\mathbb{C}) : \|x\| = 1\}\Big) = \{x \in \mathcal{M}_n(\mathbb{C}) : \|x\| = 1\}$$
y, además, $\lambda_i(D) = \lambda_i(A)$, $1 \leq i \leq n$, es fácil ver que es suficiente probar el Teorema para matrices diagonales. Pero, para $x \in \mathbb{C}^n$,
$$\langle Dx, x\rangle = \sum_{i=1}^{n} \lambda_i(A)|x_i|^2,$$
con lo que el resultado se torna inmediato. $\square$

1.5 Principio minimax.

Recordemos que, dada $A \in \mathcal{H}(n)$, denotamos a los autovalores de A (contados con multiplicidad) como $\lambda_1(A) \leq \lambda_2(A) \leq ... \leq \lambda_n(A)$.

Teorema 1.5.1 (Courant-Fisher). *Sea $A \in \mathcal{H}(n)$ y sea k un entero, $1 \leq k \leq n$. Las letras $\mathcal{M}$ y $\mathcal{S}$ las usaremos para denotar subespacios de $\mathbb{C}^n$, y $\mathcal{M}_1$ denotará a los elementos de $\mathcal{M}$ de norma uno. Entonces,*

$$\lambda_k(A) = \min_{\dim \mathcal{M}=k} \max_{x \in \mathcal{M}_1} \langle Ax, x\rangle = \max_{\dim \mathcal{S}=n-k+1} \min_{x \in \mathcal{S}_1} \langle Ax, x\rangle.$$

Demostración. Como en el Teorema anterior, se puede suponer (s.p.g.) que A es diagonal, dado que existe $U \in \mathcal{U}(n)$ tal que, si $D = \operatorname{diag}(\lambda_1(A), ..., \lambda_n(A))$, entonces $A = U^*DU$. Como U es unitaria, permuta subespacios de una dimensión fija, y preserva sus vectores de norma uno. Notar que $\langle Ax, x\rangle = \langle D\, Ux, Ux\rangle$ y $\lambda(A) = \lambda(D)$. Así que suponemos, directamente, que $A = D$. Entonces, dado $x \in \mathbb{C}^n$,

$$\langle Ax, x\rangle = \sum_{j=1}^{n} \lambda_j(A)|x_j|^2. \tag{1}$$

Dado $1 \leq r \leq n$ notemos por $\mathcal{H}_r$ al subespacio generado por $e_1, \dots, e_r$, los primeros elementos de la base canónica de $\mathbb{C}^n$. Si $\dim \mathcal{S} = n-k+1$, entonces $\mathcal{S} \cap \mathcal{H}_k \neq \{0\}$. Pero si $x \in (\mathcal{S} \cap \mathcal{H}_k)_1$, entonces por (1), $\langle Ax, x\rangle \leq \lambda_k(A)$. Esto prueba que

$$\lambda_k(A) \geq \max_{\dim \mathcal{S}=n-k+1} \min_{x \in \mathcal{S}_1} \langle Ax, x\rangle.$$

La otra desigualdad resulta de elegir $\mathcal{S} = (\mathcal{H}_{k-1})^{\perp}$, y la otra fórmula se demuestra en forma análoga. $\square$

Observación 1.5.2. La versión tradicional de las fórmulas de Courant-Fisher sería la siguiente:

$$\min_{w_1, w_2, ..., w_{n-k} \in \mathbb{C}^n} \max_{\substack{x \neq 0, x \in \mathbb{C}^n \\ x \perp w_1, w_2, ..., w_{n-k}}} \frac{x^*Ax}{x^*x} = \lambda_k(A) \tag{2}$$

$$\max_{w_1, w_2, ..., w_{k-1} \in \mathbb{C}^n} \min_{\substack{x \neq 0, x \in \mathbb{C}^n \\ x \perp w_1, w_2, ..., w_{k-1}}} \frac{x^*Ax}{x^*x} = \lambda_k(A) \tag{3}$$

Teorema 1.5.3 (Teorema de Weyl). *Sean $A, B \in \mathcal{H}(n)$ y $j \in \{1, \dots, n\}$. Entonces:*

$$\lambda_j(A) + \lambda_1(B) \leq \lambda_j(A+B) \leq \lambda_j(A) + \lambda_n(B)$$

Demostración. Notar que, por el Teorema 1.4.3, para todo $x \in \mathbb{C}^n$ tal que $\|x\| = 1$,

$$\langle Ax, x\rangle + \lambda_1(B) \leq \langle Ax, x\rangle + \langle Bx, x\rangle \leq \langle Ax, x\rangle + \lambda_n(B).$$

Por lo tanto el teorema se puede deducir de las fórmulas de Courant-Fischer. $\square$

Corolario 1.5.4. *Sean $A, B \in \mathcal{H}(n)$ tales que $A \leq B$. Entonces, para todo $1 \leq j \leq n$,*

$$\lambda_j(A) \leq \lambda_j(B).$$

Demostración. Basta aplicar el Teorema de Weyl al par A y $B - A \geq 0$, y notar que, por el Teorema 1.4.3, $\lambda_n(B - A) \geq 0$. □

Como consecuencia del Teorema de Courant-Fisher obtenemos el llamado **teorema de entrelace de Cauchy**. Este teorema relaciona los autovalores de una matriz Hermitiana con sus correspondientes submatrices principales.

Teorema 1.5.5 (Entrelace de Cauchy). *Sea $A \in \mathcal{H}(n)$, r un entero tal que $1 \leq r \leq n$ y $A_r \in \mathcal{M}_{n-1}(\mathbb{C})$ la submatriz principal de A obtenida al borrar la fila y la columna r-ésimas de A. Entonces para cada $1 \leq k \leq n-1$,*

$$\lambda_k(A) \leq \lambda_k(A_r) \leq \lambda_{k+1}(A).$$

Es decir que

$$\lambda_1(A) \leq \lambda_1(A_r) \leq \lambda_2(A) \leq ... \leq \lambda_{n-1}(A_r) \leq \lambda_n(A).$$

Demostración. Supongamos, por simplicidad, que $r = n$. Fijemos un k tal que $1 \leq k \leq n-1$. Trabajando sólo con los subespacios $\mathcal{M} \subseteq \mathcal{H}_{n-1}$ que tienen $\dim \mathcal{M} = k$ (que son menos que todos los subespacio de dimensión k de $\mathbb{C}^n$), obtenemos

$$\lambda_k(A_n) = \min_{\substack{\dim \mathcal{M}=k \\ \mathcal{M} \subseteq \mathcal{H}_{n-1}}} \max_{x \in \mathcal{M}_1} \langle A_n x, x \rangle \geq \min_{\dim \mathcal{M}=k} \max_{x \in \mathcal{M}_1} \langle Ax, x \rangle = \lambda_k(A),$$

porque si $x \in \mathcal{H}_{n-1}$, entonces $\langle Ax, x \rangle = \langle A_n x, x \rangle$, dado que $x_n = 0$. Análogamente, tomando subespacios $\mathcal{S} \subseteq \mathcal{H}_{n-1}$ tales que $\dim \mathcal{S} = n - k$, como $n - k = n - (k+1) + 1$ y a la ves, $n - k = (n-1) - k + 1$, obtenemos

$$\lambda_k(A_n) = \max_{\substack{\dim \mathcal{S}=n-k \\ \mathcal{S} \subseteq \mathcal{H}_{n-1}}} \min_{x \in \mathcal{S}_1} \langle A_n x, x \rangle \leq \max_{\dim \mathcal{S}=n-k} \min_{x \in \mathcal{S}_1} \langle Ax, x \rangle = \lambda_{k+1}(A).$$

□

Observación 1.5.6. En forma análoga se puede probar versiones más generales del Teorema anterior:

1. Sea $A \in \mathcal{H}(n)$, r un entero tal que $1 \leq r \leq n$ y $A_r \in \mathcal{H}(r)$ cualquier submatriz principal de A obtenida al borrar $n - r$ filas y las correspondientes columnas de A. Entonces para cada entero k tal que $1 \leq k \leq r$, se tiene

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

2. Más en general aún, si $P \in \mathcal{M}_n(\mathbb{C})^+$ es un proyector autoadjunto (o sea ortogonal) sobre un subespacio $\mathcal{S}$ de $\dim \mathcal{S} = r$, entonces al producto PAP se lo puede pensar como un operador en el espacio de Hilbert $\mathcal{S}$ (para que su vector de autovalores tenga sólo r cordenadas, sacando los $n - r$ ceros que "sobran"). Entonces se obtienen desigualdades análogas: para cada entero k tal que $1 \leq k \leq r$, se tiene

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

△

Ejercicio 1.5.7. Escribir explícitamente cómo quedan los resultados de esta sección (minimax, teorema de Weyl y los tres entrelaces) en función de los vectores $\mu(A)$ ordenados decrecientemente. $\triangle$

Ahora, como corolario del Teorema de entrelace, veremos una caracterización de positividad de matrices en términos de submatrices principales. Para ello necesitaremos del siguiente resultado previo.

Lema 1.5.8. *Toda submatriz principal de una matriz definida positiva, es definida positiva.*

Demostración. Ejercicio. $\square$

Teorema 1.5.9. *Si $A \in \mathcal{H}(n)$, entonces A es definida positiva (i.e. $A > 0$) si y sólo si*

$$\det A_i > 0 \quad \textit{para todo} \;\; 1 \le i \le n,$$

donde A_i es la submatriz principal de A obtenida con las primeras i filas y columnas de A.

Demostración. Llamemos A_i a la submatriz principal de orden i de A determinada por las primeras i filas y columnas de A. Por el Teorema anterior si A es definida positiva, A_i es definida positiva para i:1,...,n. Esto significa que los autovalores de cada A_i son todos positivos y por lo tanto el determinante de cada una de ellas también lo es.
Para demostrar la recíproca utilizaremos inducción sobre la dimensión n y el Teorema del entrelace de Cauchy 1.5.5. Como $\det(A_1) > 0$ y A_1 es de orden 1, A_1 es definda positiva. Supongamos entonces A_k definida positiva para algún $k < n$. Por lo tanto todos los autovalores de A_k son positivos. La desigualdad de entrelace nos dice, entonces, que todos los autovalores de A_{k+1} son positivos, excepto quizás el menor de ellos. Pero el producto de los autovalores de A_{k+1} coincide con el determinante de A_{k+1}, el cual es positivo por hipótesis. Por lo tanto, todos los autovalores de A_{k+1} (anche el primero) son positivos. Y, del hecho de que $A_{k+1} \in \mathcal{H}(k+1)$, se deduce que A_{k+1} es definida positiva. Finalmente, como $A_n = A$, hemos probado que A es definida positiva. $\square$

Ejercicio (difícil): Probar que, dada $A \in \mathcal{H}(n)$, entonces $A \ge 0$ si y sólo si para todo $J \subseteq \{1, \dots, n\}$, $\det A_J \ge 0$, donde $A_J = (a_{ij})_{i,j \in J} \in \mathcal{H}(|J|)$.

1.6 Descomposición polar y valores singulares

Antes de definir los conceptos que dan título a esta sección, necesitamos hacer las siguientes observaciones:

1. Dada $A \in \mathcal{M}_n(\mathbb{C})$, es fácil ver que $A^*A \ge 0$. En efecto, si $x \in \mathbb{C}^n$, entonces $\langle A^*Ax, x\rangle = \|Ax\|^2 \ge 0$.

2. Además, dada una matriz $B \in \mathcal{M}_n(\mathbb{C})^+$, existe una única matriz $C \in \mathcal{M}_n(\mathbb{C})^+$ tal que $C^2 = B$. En efecto, Se debe escribir $B = UDU^*$, con $U \in \mathcal{U}(n)$ y $D = \operatorname{diag}(\lambda_1(B), \dots, \lambda_n(B)) \in \mathcal{M}_n(\mathbb{C})^+$. Se toma

$$D' = \operatorname{diag}\left(\lambda_1(B)^{1/2}, \dots, \lambda_n(B)^{1/2}\right)$$

Esto es posible por la frase final del Teorema 1.4.3. Finalmente se define $C = UD'U^*$. Es claro que $C \in \mathcal{M}_n(\mathbb{C})^+$ y que $C^2 = B$. La unicidad es algo más complicada y se deja como ejercicio.

3. Dada $B \in \mathcal{M}_n(\mathbb{C})^+$ llamaremos $B^{1/2}$ a su "raiz cuadrada positiva" en el sentido del ítem anterior.

Definición 1.6.1. Sea $A \in \mathcal{M}_n(\mathbb{C})$,

1. Llamaremos "módulo de A" a la matriz

$$|A| = (A^*A)^{1/2} \in \mathcal{M}_n(\mathbb{C})^+.$$

2. Llamaremos **valores singulares** de A a los autovalores de $|A|$ ordenados en forma decreciente, notándolos $s_1(A) \geq \cdots \geq s_n(A) \geq 0$.

3. Llamaremos $s(A) = (s_1(A), \ldots, s_n(A)) = \mu(|A|)$ y $\Sigma(A)$ a la matriz diagonal

$$\Sigma(A) = \operatorname{diag}(s(A)) = \begin{pmatrix} s_1(A) & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & s_n(A) \end{pmatrix}.$$

$\triangle$

Ejemplos 1.6.2. Sea $A \in \mathcal{M}_n(\mathbb{C})$.

1. Si $A \geq 0$, entonces $A = |A|$ y $s(A) = \mu(A)$.

2. Si $A \in \mathcal{H}(n)$, entonces $s(A) = |\mu(A)|$, salvo el orden. En efecto, $A^*A = A^2$, luego los autovalores de $|A|$ son los módulos de los de A (si $a \in \mathbb{R}$, $(a^2)^{1/2} = |a|$).

3. En general, los autovalores y los valores singulares de una misma matriz pueden ser bien distintos. Por ejemplo, si A es un bloque nilpotente de Jordan en $\mathcal{M}_n(\mathbb{C})$ (i.e. $Ae_k = e_{k+1}$, $1 \leq k \leq n-1$ y $Ae_n = 0$), entoces $\sigma(A) = \{0\}$ porque $A^n = 0$, pero $s(A) = (1, \ldots, 1, 0)$, porque A^*A es un proyector de rango $n-1$.

Teorema 1.6.3 (Descomposicioń polar y en valores singulares). *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces*

1. *Para todo $x \in \mathbb{C}^n$, se verifica que $\|Ax\| = \|\,|A|x\|$.*

2. *Existe una matriz unitaria $U \in \mathcal{M}_n(\mathbb{C})$ tal que $A = U|A|$.*

3. *Existen dos matrices unitarias $V, W \in \mathcal{M}_n(\mathbb{C})$ tales que*

$$A = W\Sigma(A)V.$$

Demostración. Ejercicio. $\square$

1.7 Normas en $\mathcal{M}_n(\mathbb{C})$.

Se estudiarán en esta sección distintas normas en el espacio vectorial de matrices $\mathcal{M}_n(\mathbb{C})$. Muchas de estas normas son útiles en diversas desigualdades matriciales específicas. Pero no olvidemos que, como $\dim \mathcal{M}_n(\mathbb{C}) < \infty$, es un resultado conocido que todas las normas en $\mathcal{M}_n(\mathbb{C})$ son equivalentes.

En los siguientes ejemplos definiremos las normas más clásicas para matrices. Dejaremos como ejercicio para el lector la verificación (en algunos casos altamente no trivial) de que son, efectivamente, normas.

Ejemplos 1.7.1.

1. La norma espectral $\|\cdot\|$ definida del siguiente modo

$$\|A\| = \|A\|_{sp} = \max_{\|x\|=1} \|Ax\| = s_1(A),$$

   donde la última igualdad surge de que $\|A\|_{sp} = \|\,|A|\,\|_{sp} = \rho(|A|)$.

2. Las normas de Schatten. Dado $1 \le p < \infty$

$$\|A\|_p = \left(\sum_{n=1}^{\infty} s_i\,(A)^p \right)^{1/p} .$$

   La $\|\cdot\|_2$ se llama **norma de Frobenius**. Ella verifica que

$$\|A\|_2^2 = \operatorname{tr} A^*A = \sum_{i,j=1}^{n} |a_{ij}|^2$$

   y proviene del producto escalar en $\mathcal{M}_n(\mathbb{C})$ definido por $\langle A, B \rangle = \operatorname{tr} B^*A$.

3. Las normas Ky-Fan. Dado $k \in \{1, \dots, n\}$

$$\|A\|_{(k)} = \sum_{i=1}^{k} s_i\,(A)\,.$$

   Notar que $\|A\|_{(1)} = \|A\|_{sp}$ y $\|A\|_{(n)} = \|A\|_1$ (de Schatten).

4. Toda norma N en $\mathbb{C}^n$ induce una norma $|||\cdot|||_N$ en $\mathcal{M}_n(\mathbb{C})$ del siguiente modo:

$$|||A|||_N = \max_{N(x)=1} N(Ax).$$

   Estas normas satisfacen que:

   (a) $|||I|||_N = 1$

   (b) $\rho(A) \le |||A|||_N$

   (c) $|||AB|||_N \le |||A|||_N \; |||B|||_N$.

5. $|||\cdot|||_1$ es la inducida por la norma $\|\cdot\|_1$ en $\mathbb{C}^n$. Puede demostrarse que

$$|||A|||_1 = \max_i \|C_i(A)\|_1$$

donde $C_i(A)$ simboliza el vector formado por la columna i-ésima de A.

6. Analogamente, $|||\cdot|||_\infty$ es la inducida por la norma $\|\cdot\|_\infty$ en $\mathbb{C}^n$. Puede demostrarse que

$$|||A|||_\infty = \max_i \|F_i(A)\|_1$$

donde $F_i(A)$ simboliza el vector formado por la fila i-ésima de A.

Definición 1.7.2. Una norma $\|\cdot\|$ en $\mathcal{M}_n(\mathbb{C})$ se llama:

1. Matricial: si $\|AB\| \leq \|A\|\;\|B\|$
2. Unitariamente invariante (NUI): si $\|UAV\| = \|A\|$ para todo $U, V \in \mathcal{U}(n)$. $\triangle$

El siguiente es un ejemplo de una norma que no es matricial

Ejemplo 1.7.3. Definamos

$$N_\infty(A) = \max_{ij} |a_{ij}|,$$

y consideremos la matriz

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Entonces $A^2 = 2A$, y en consecuencia $N_\infty(A^2) = 2$ mientras que $N_\infty(A) = 1$. El lector interesado puede demostrar que $n.N_\infty(\cdot)$ sí es una norma matricial en $\mathcal{M}_n(\mathbb{C})$.

Teorema 1.7.4. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ y $\|\cdot\|$ una norma matricial. Entonces:*

1. *$\|A - I\| < 1$ implica que $A \in \mathcal{G}l(n)$ y $A^{-1} = \sum_{n=0}^{\infty} (I - A)^n$*
2. *Si $B \in \mathcal{G}l(n)$ y $\|B - A\| \leq \|B^{-1}\|^{-1}$, entonces, $A \in \mathcal{G}l(n)$.*

Demostración. Comencemos notando que $1 \Rightarrow 2$. En efecto,

$$\|B^{-1}A - I\| = \|B^{-1}(A - B)\| \leq \|B^{-1}\|\;\|A - B\| < 1.$$

Por lo tanto $B^{-1}A$ es inversible y $A = B(B^{-1}A)$ también lo es. Para demostrar (1), llamemos $C = I - A$. Entonces, $\|C\| < 1$ y $\|C^m\| \leq \|C\|^m$, en consecuencia

$$\sum_{k=0}^{\infty} \|C^k\| \leq \sum_{k=0}^{\infty} \|C\|^k = \frac{1}{1 - \|C\|}.$$

Luego, $\sum_{k=1}^{\infty} C^k$ converge. Por otro lado,

$$A \sum_{k=0}^{N} C^k = A \sum_{k=0}^{N} (I - A)^k = 1 - C^{N+1} \xrightarrow[N \to \infty]{} 1$$

y analogamente $\left(\sum_{k=0}^{\infty} C^k \right) A = 1$. □

Lema 1.7.5. *Sea $P \in \mathbb{C}[x]$ y $A \in \mathcal{M}_n(\mathbb{C})$. Entonces,*

$$\sigma\left(P(A)\right) = P(\sigma\left(A\right)) := \{P(\lambda) : \ \lambda \in \sigma\left(A\right)\}.$$

Demostración. Ejercicio. Se sugieren dos maneras de hacerlo: una aplicando el Toerema 1.2.4. La otra con cuentas de polinomios (que es la que sirve en dimensión infinita). □

Proposición 1.7.6. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ y $\| \cdot \|$ una norma matricial. Entonces $\rho(A) \leq \|A\|$. Más aún, $\|A^m\|^{1/m} \geq \rho(A)$ para todo $m \in \mathbb{N}$.*

Demostración. Sea $\lambda \in \sigma\left(A\right)$, $x \in \mathbb{C}^n$ tal que $x \neq 0$ y $Ax = \lambda x$. Llamemos X a la matriz cuyas columnas son todas iguales al vector x. Luego, $AX = \lambda X$, y por ende

$$|\lambda| \, \|X\| = \|AX\| \leq \|A\| \, \|X\|,$$

de donde se deduce que $|\lambda| \leq \|A\|$. Como el autovalor era cualquiera, es tiene que $\rho(A) \leq \|A\|$. Además, por el Lema 1.7.5, se tiene que $\rho(A^m) = \rho(A)^m$. Por lo tanto, usando la parte ya probada, obtenemos que $\rho(A) \leq \|A^m\|^{1/m}$.

□

Observación 1.7.7. Dada una norma matricial $\| \cdot \|$ y una matriz inversible S, entonces, $\|A\|_S := \|SAS^{-1}\|$ es también matricial.

Proposición 1.7.8. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ y $\varepsilon > 0$. Entonces, existe una norma matricial N en $\mathcal{M}_n(\mathbb{C})$ tal que $N(A) \leq \rho(A) + \varepsilon$.*

Demostración. Sea $A = UTU^*$ con T una matriz triangular superior y $U \in \mathcal{U}(n)$. Luego, $T = U^*AU$. Sea $D_s = \operatorname{diag}(s, s^2, \ldots, s^n)$. Entonces, $D_s T D_s^{-1} = (t_{ij} s^{i-j})$ y si $s \to \infty$

$$D_s T D_s^{-1} \xrightarrow{\text{En cualquier norma}} \operatorname{diag}\left(\lambda_1\left(A\right), \ldots, \lambda_n\left(A\right)\right).$$

Como $\|\operatorname{diag}\left(\lambda_1\left(A\right), \ldots, \lambda_n\left(A\right)\right)\|_{sp} = \rho(A)$, entonces, $\|D_s T D_s^{-1}\|_{sp} \xrightarrow[s \to \infty]{} \rho(A)$. Luego, existe $s \in \mathbb{R}$ tal que $\|D_s T D_s^{-1}\|_{sp} < \rho(A) + \varepsilon$. Sea N la norma definida del siguiente modo:

$$N(B) = \|D_s U^* B U D_s^{-1}\|_{sp}\,, \quad B \in \mathcal{M}_n(\mathbb{C}).$$

Por construcción, esta norma satisface la propiedad buscada. □

Corolario 1.7.9. *Si $A \in \mathcal{M}_n(\mathbb{C})$, entonces $A^m \to 0$ si y sólo si $\rho(A) < 1$.*

Demostración. Es claro que $\rho(A) < 1$ si y sólo si $\rho(A^m) = \rho(A)^m \xrightarrow[m\to\infty]{} 0$. Como $\rho(A^m) \le \|A^m\|_{sp}$, una implicación es clara. Para probar la otra, supongamos que $\rho(A) < 1$ y elijamos una norma matricial N tal que $\rho(A) \le N(A) < 1$. Tal norma existe por la Proposición 1.7.8. Como N es matricial, deducimos que $N(A^m) \le N(A)^m \xrightarrow[m\to\infty]{} 0$.

□

Teorema 1.7.10. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces, para toda norma $\|\cdot\|$ en $\mathcal{M}_n(\mathbb{C})$,*

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

Demostración. Si $\|\cdot\|$ es una norma matricial, se tiene $\rho(A) \le \|A^m\|^{1/m}$ para todo $m \in \mathbb{N}$. Sea $\varepsilon > 0$ y $B = \dfrac{A}{\rho(A)+\varepsilon}$. Entonces, $\rho(B) < 1$ y $\|B^m\| \to 0$. En consecuencia existe un $m_0 \in \mathbb{N}$ tal que, para todo $m \ge m_0$, se tiene $\|A^m\| < (\rho(A)+\varepsilon)^m$, lo que prueba el Teorema. El mismo resultado vale para normas no matriciales, por ser todas las normas equivalentes. □

Ejercicio 1.7.11. Probar que si N es una norma matricial,

$$\rho(A) = \inf_{m\in\mathbb{N}} N(A^m)^{1/m}.$$

Observación 1.7.12. Todos los resultados de esta sección, a partir del Teorema 1.7.4, son también ciertos en álgebras de Banach, donde las normas son matriciales por definición. El único resultado propio de matrices es el Corolario 1.7.9.

Sin embargo, hemos incluido las demostraciones anteriores porque tienen un buen sabor matricial, salvo el Teorema 1.7.4, que tiene la prueba standard (y no creo que pueda mejorarse). Nótese que se asumió implícitamente que $\mathcal{M}_n(\mathbb{C})$ es un espacio completo, porque usamos que una serie absolutamente sumable es convergente.

2 Mayorización y matrices Hermitianas

2.1 Definiciones y propiedades básicas

Notaciones: Sea $x \in \mathbb{R}^n$, $x = (x_1, ..., x_n)$. Notaremos $x^\downarrow$ y $x^\uparrow$ a los vectores obtenidos al reordenar las cordenadas de x en forma decreciente y creciente respectivamente. Es decir, por ejemplo, que

$$x_1^\downarrow = \max_i x_i \ , \quad x_1^\downarrow + x_2^\downarrow = \max_{i\ne j} x_i + x_j \ , \quad \text{etc.}$$

Por ejemplo, si x es el vector de autovalores de una matriz $A \in \mathcal{H}(n)$, entonces $x^\downarrow = \mu(A)$ y $x^\uparrow = \lambda(A)$. Otra notación: dado $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, escribiremos

$$\operatorname{tr} x = \sum_{i=1}^n x_i.$$

Con estas notaciones podemos dar la definición de relación de mayorización,

Definición 2.1.1. Sean $x, y \in \mathbb{R}^n$. Se dice que x **es mayorizado por** y, y se nota $x \prec y$ si se verifica que $\operatorname{tr} x = \operatorname{tr} y$ y

$$\sum_{j=1}^{k} x_j^{\downarrow} \ \leq \ \sum_{j=1}^{k} y_j^{\downarrow}, \quad 1 \leq k < n, \tag{4}$$

$\triangle$

Observación 2.1.2. Dado que $\sum\limits_{j=1}^{k} x_j^{\uparrow} = \sum\limits_{j=1}^{n} x_j - \sum\limits_{j=1}^{n-k} x_j^{\downarrow}$, la relación $x \prec y$ también es equivalente a las condiciones: $\operatorname{tr} x = \operatorname{tr} y$ y

$$\sum_{j=1}^{k} x_j^{\uparrow} \ \geq \ \sum_{j=1}^{k} y_j^{\uparrow}, \quad 1 \leq k < n, \tag{5}$$

Si sólo se cumple la condición (4), se dice que x es **submayorizado** por y y se nota $x \prec_w y$. Por el contrario, si sólo se cumple la condición (5), se dice que x es **supramayorizado** por y y se nota $x \prec^w y$. $\triangle$

Ejemplos 2.1.3. 1. Si $x \in \mathbb{R}^n$, cumple que $x_i \geq 0$ y $\sum\limits_{i=1}^{n} x_i = 1$, entonces

$$(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}) \prec (x_1, x_2, ..., x_n) \prec (1, 0, ..., 0). \tag{6}$$

Una manera de verificarlo es suponer que existe $k < n$ tal que $\sum\limits_{i=1}^{k} x_i^{\downarrow} < k/n$. En tal caso, debe suceder que $x_k^{\downarrow} < 1/n$ y, por lo tanto, también $\sum\limits_{i=k+1}^{n} x_i^{\downarrow} < (n-k)/n$. Pero esto contradice el hecho de que

$$\sum_{j=1}^{n} \frac{1}{n} = 1 = \sum_{i=1}^{n} x_i.$$

La otra mayorización es obvia.
2. Si $x \prec y$ e $y \prec x$, entonces x e y difieren en una permutación. $\triangle$

Ejercicio 2.1.4. Sean $x, y \in \mathbb{R}^n$. Probar que $x \prec_w y$ si y sólo si existe $u \in \mathbb{R}^n$ tal que

$$x \leqslant u \prec y \ ,$$

donde $x \leqslant u$ significa que $x_i \leq y_i$, para todo $1 \leq i \leq n$. Se sugiere hacer inducción sobre n, y hacer el paso inductivo "agrandando" la cordenada $x_1^{\downarrow}$ hasta donde se pueda (o se iguala alguna de las sumas parciales, o se llega hasta $y_1^{\downarrow}$).

Existe una relación muy estrecha entre las relaciones de mayorización y las matrices doblemente estocásticas.
Notaciones: Si $A \in \mathcal{M}_n(\mathbb{C})$,

1. Llamaremos $F_i(A)$ a la fila i-ésima de A y $C_j(A)$ a la columna j-ésima.
2. Notaremos $A \geqslant 0$ si $A_{ij} \geq 0$, es decir, si A tiene entradas no negativas (no el lo mismo que $A \geq 0$).

Definición 2.1.5. Una matriz $A \in \mathcal{M}_n(\mathbb{R})$ se denomina **doblemente estocástica** si $A \geqslant 0$ y, para todo $1 \leq i \leq n$, se verifica que

$$\operatorname{tr} F_i(A) = 1 \quad \text{y} \quad \operatorname{tr} C_i(A) = 1.$$

Al conjunto de matrices doble estocásticas en $\mathcal{M}_n(\mathbb{C})$ lo denotaremos por medio de $\mathcal{DS}(n)$.
$\triangle$

Ejercicio 2.1.6. Probar que $A \in \mathcal{DS}(n)$ si y sólo si

$$A \geqslant 0\ , \quad Ae = e \quad \text{y} \quad A^*e = e,$$

donde $e = (1, 1, \ldots, 1) \in \mathbb{R}^n$.

Teorema 2.1.7. *$A \in \mathcal{DS}(n)$ si y sólo si $Ax \prec x$ para todo $x \in \mathbb{R}^n$.*

Demostración. Supongamos primero que $Ax \prec x$ para todo x. Sea $\mathcal{E} = \{e_1, \ldots, e_n\}$ la base canónica de $\mathbb{C}^n$. Entonces, como $Ae_i \prec e_i$, $1 \leq i \leq n$, podemos deducir que $A \geqslant 0$ y, para $1 \leq i \leq n$,

$$1 = \operatorname{tr} e_i = \operatorname{tr} Ae_i = \operatorname{tr} C_i(A).$$

Por otra parte, tomemos $e = (1, 1, \ldots, 1)$. Usando el Ejemplo 2.1.3, como $Ae \prec e$ y todas las cordenadas de e son iguales, deducimos que

$$e = Ae = (\operatorname{tr} F_1(A), \ldots, \operatorname{tr} F_n(A)).$$

Recíprocamente, supongamos que $A \in \mathcal{DS}(n)$ y llamemos $y = Ax$. Queremos probar que $y \prec x$. Se puede suponer que las cordenadas de x y de y están ordenadas en forma decreciente, porque si $P, Q \in \mathcal{U}(n)$ son matrices de permutación, entonces $QAP \in \mathcal{DS}(n)$ (esto se formaliza fácil con el Ejercicio anterior). Además, como $y = Ax$, para $1 \leq k \leq n$, tenemos

$$\sum_{j=1}^{k} y_j = \sum_{j=1}^{k} \sum_{i=1}^{n} a_{ij} x_i .$$

Si llamamos $t_i = \sum\limits_{j=1}^{k} a_{ij}$, entonces $0 \leq t_i \leq 1$ y $\sum\limits_{i=1}^{n} t_i = k$. Luego,

$$\begin{aligned}
\sum_{j=1}^{k} y_j - \sum_{j=1}^{k} x_j &= \sum_{j=1}^{k}\sum_{i=1}^{n} a_{ij}x_i - \sum_{i=1}^{k} x_i = \sum_{i=1}^{n} t_i x_i - \sum_{i=1}^{k} x_i \\
&= \sum_{i=1}^{n} t_i x_i - \sum_{i=1}^{k} x_i + (k - \sum_{i=1}^{n} t_i)x_k \\
&= \sum_{i=1}^{k} t_i x_i + \sum_{i=k+1}^{n} t_i x_i - \sum_{i=1}^{k} x_i + kx_k - \sum_{i=1}^{k} t_i x_k - \sum_{i=k+1}^{n} t_i x_k \\
&= \sum_{i=1}^{k}(t_i - 1)x_i + \sum_{i=1}^{k} x_k - \sum_{i=1}^{k} t_i x_k + \sum_{i=k+1}^{n} t_i(x_i - x_k) \\
&= \sum_{i=1}^{k}(t_i - 1)x_i + \sum_{i=1}^{k}(1 - t_i)x_k + \sum_{i=k+1}^{n} t_i(x_i - x_k) \\
&= \sum_{i=1}^{k}(t_i - 1)(x_i - x_k) + \sum_{i=k+1}^{n} t_i(x_i - x_k) \leq 0,
\end{aligned}$$

pues los dos términos del último renglón son sumas de términos no positivos. Por lo tanto $\sum_{j=1}^{k} y_j \leq \sum_{j=1}^{k} x_j$ para $1 \leq k \leq n$ y además, cuando $k = n$, obtenemos la igualdad por el hecho de que $A \in \mathcal{DS}\,(n)$. Así, $y \prec x$. □

Ejemplo 2.1.8. Como motivación del siguiente resultado, supongamos que $(y_1, y_2) \prec (x_1, x_2)$, $y_1 \geq y_2$ y $x_1 \geq x_2$. Entonces

$$x_2 \leq y_2 \leq y_1 \leq x_1 \qquad \text{y} \qquad y_1 + y_2 = x_1 + x_2.$$

Sea $\lambda \geq 0$ tal que $y_1 = \lambda x_1 + (1 - \lambda)x_2$. Entonces,

$$\begin{aligned}
y_2 &= y_1 + y_2 - y_1 = x_1 + x_2 - y_1 = x_1 + x_2 - \big(\lambda x_1 + (1 - \lambda)x_2\big) \\
&= (1 - \lambda)x_1 + \lambda x_2.
\end{aligned}$$

y por lo tanto $(y_1, y_2) = \lambda(x_1, x_2) + (1 - \lambda)(x_2, x_1)$. △

Teorema 2.1.9. *Sean $x, y \in \mathbb{R}^n$. Entonces, son equivalentes:*

1. *$y \prec x$*
2. *y es una combinación convexa de permutaciones de x*
3. *Existe $A \in \mathcal{DS}\,(n)$ tal que $y = Ax$.*

Demostración. Claramente $2 \Rightarrow 3$ y por el Teorema 2.1.7 se tiene que $3 \Rightarrow 1$. Luego, solo resta probar que $1 \Rightarrow 2$. Lo haremos por inducción sobre la dimensión n. Para $n = 1$ es trivial y el caso $n = 2$ fue probado en el Ejemplo 2.1.8. Sea $n > 2$. Sin perdida de generalidad podemos suponer que los vectores estan ordenados en forma decreciente. Luego,

$x_n \leq y_n \leq y_1 \leq x_1$. Sea $k > 1$ tal que $x_k \leq y_1 \leq x_{k-1}$ y $\lambda \geq 0$ tal que $y_1 = \lambda x_1 + (1-\lambda)x_k$. Sea P_0 la permutación tal que

$$P_0(x_1, \ldots, x_n) = (x_k, \ldots, x_{k-1}, x_1, x_{k+1}, \ldots, x_n).$$

Definamos $z = \lambda x + (1-\lambda)P_0(x)$. Sean $y' = (y_2, \ldots, y_n)$ y $z' = (z_2, \ldots, z_n)$.

Vamos a probar que $y' \prec z'$: Como $\operatorname{tr}(z') = \operatorname{tr}(z) - y_1 = \operatorname{tr}(x) - y_1$ y $\operatorname{tr}(y') = \operatorname{tr}(y) - y_1 = \operatorname{tr}(x) - y_1$, se tiene que $\operatorname{tr}(y') = \operatorname{tr}(z')$. Si $m \leq k-1$, entonces, como $y_1 \leq x_{k-1}$,

$$\sum_{i=2}^{m} z_i = \sum_{i=2}^{m} x_i \geq (m-1)x_{k-1} \geq (m-1)y_1 \geq \sum_{i=2}^{m} y_i.$$

Por otro lado, si $m \geq k$.

$$\begin{aligned}\sum_{i=2}^{m} z_i &= \sum_{i=2}^{k-1} x_i + (1-\lambda)x_1 + \lambda x_k + \sum_{i=k+1}^{m} x_i \\ &= \sum_{i=1}^{m} x_i + \lambda x_1 - (1-\lambda)x_k \\ &= \sum_{i=1}^{m} x_i - y_1 \geq \sum_{i=1}^{m} y_i - y_1 = \sum_{i=2}^{m} y_i.\end{aligned}$$

Por hipótesis inductiva $y' = \sum_{i=1}^{s} \mu_i P_{\sigma_i}(z')$ para ciertos $\mu_i \geq 0$ que suman uno, y para ciertas permutaciones $\sigma_i \in S_{n-1}$ (pensadas como biyecciones del conjunto $\{2, 3, \ldots, n\}$). Llamemos también $\sigma_i \in S_n$ la extensión de la permutación σ_i a $\mathbb{I}_n$ poniendo $\sigma_i(1) = 1$. Luego $y = \sum_{i=1}^{s} \mu_i P_{\sigma_i}(z)$. Pero entonces

$$y = \sum_{i=1}^{s} \mu_i P_{\sigma_i}(z') = \sum_{i=1}^{s} \lambda \mu_i P_{\sigma_i}(x) + \sum_{i=1}^{s} (1-\lambda)\mu_i P_{\sigma_i} P_0(x),$$

que es una combinación convexa de permutaciones de x. $\square$

Observación 2.1.10. Un error típico en cuentas con mayorización, es olvidarse de ordenar a los vectores antes de sumar sus "primeras" k cordenadas. De hecho, esto sucede en la prueba anterior con el vector z'. Por suerte no es grave en este caso, porque z' está del lado de "los mayores", y lo grave es no reordenar del lado de "los menores". Más explícitamente, si $x, y \in \mathbb{R}^n$, como

$$\sum_{i=1}^{k} y_i \leq \sum_{i=1}^{k} y_i^{\downarrow},$$

es imprescindible ordenar a y para verificar que $y \prec x$, pero no para verificar que $x \prec y$. Notar que, en la prueba de la relación $y' \prec z'$, el vector y' ya venía ordenado correctamente. $\triangle$

Teorema 2.1.11. *Sean $x, y \in \mathbb{R}^n$. Entonces, son equivalentes:*

1. $y \prec x$
2. $\operatorname{tr}(f(y)) \leq \operatorname{tr}(f(x))$ *para toda función convexa* $f : \mathbb{R} \to \mathbb{R}$. *Se usa la notación* $f(x)$ *para el vector* $(f(x_1), \dots, f(x_n))$.
3. $\displaystyle\sum_{i=1}^{n} |y_i - t| \leq \sum_{i=1}^{n} |x_i - t|$ *para todo* $t \in \mathbb{R}$.

Analogamente, son equivalentes

1. $y \prec_w x$ *(submayorización)*
2. $\operatorname{tr}(f(y)) \leq \operatorname{tr}(f(x))$ *para toda función* $f : \mathbb{R} \to \mathbb{R}$ *convexa y no decreciente.*
3. $\displaystyle\sum_{i=1}^{n} (y_i - t)^+ \leq \sum_{i=1}^{n} (x_i - t)^+$ *para todo* $t \in \mathbb{R}$.

Demostración. Sólo probaremos la primer parte, puesto que los argumentos para probar la segunda son similares. Supongamos que $y \prec x$. Entonces, existe una matriz doble estocástica $A = (a_{ij})$ tal que $y = Ax$. Luego, dada una función convexa f, se tiene que

$$\operatorname{tr}(f(y)) = \sum_{i=1}^{n} f(y_i) = \sum_{i=1}^{n} f\left(\sum_{j=1}^{n} a_{ij} x_j\right) \leq \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} f(x_j) = \sum_{i=j}^{n} f(x_j)$$

Es claro que $2 \Rightarrow 3$, así que probemos que $3 \Rightarrow 1$. Supongamos que los vectores x e y están ordenados de forma decreciente (ni 3 ni 1 depende del orden de las cordenadas). Tomando $t < x_n$, se verifica que $\operatorname{tr}(y) \leq \operatorname{tr}(x)$, y tomando $t > x_1$, que $\operatorname{tr}(y) \geq \operatorname{tr}(x)$. Por otro lado, dado $x \in \mathbb{R}$, se tiene que $x^+ = (x + |x|)/2$. Luego

$$\begin{aligned}
\sum_{i=1}^{n} (y_i - t)^+ &= \frac{\sum_{i=1}^{n} (y_i - t) + \sum_{i=1}^{n} |y_i - t|}{2} \\
&\leq \frac{\sum_{i=1}^{n} (x_i - t) + \sum_{i=1}^{n} |x_i - t|}{2} \\
&= \sum_{i=1}^{n} (x_i - t)^+.
\end{aligned}$$

Tomando $t = x_k$, resulta $\displaystyle\sum_{i=1}^{n} (x_i - t)^+ = \sum_{i=1}^{k} (x_i - t)^+ = \sum_{i=1}^{k} x_i - kt$. Por lo tanto

$$\begin{aligned}
\sum_{i=1}^{k} y_i - kt = \sum_{i=1}^{k} (y_i - t) &\leq \sum_{i=1}^{k} (y_i - t)^+ \leq \sum_{i=1}^{n} (y_i - t)^+ \\
&\leq \sum_{i=1}^{n} (x_i - t)^+ = \sum_{i=1}^{k} x_i - kt,
\end{aligned}$$

lo cual muestra que $\displaystyle\sum_{i=1}^{k} y_i \leq \sum_{i=1}^{k} x_i$. $\square$

Corolario 2.1.12. *Sean $x, y \in \mathbb{R}^n$ tales que $x \prec y$ y sea $f : \mathbb{R} \to \mathbb{R}$ una función convexa. Entonces, $(f(x_1), \dots, f(x_n)) \prec_w (f(y_1), \dots, f(y_n))$.*

Demostración. Ejercicio. □

Teorema 2.1.13 (Birkhoff). *El conjunto de matrices doble estocásticas $\mathcal{DS}(n)$ es convexo y sus puntos extremales son las matrices de permutación. Es decir que toda $A \in \mathcal{DS}(n)$ es combinación convexa de matrices de permutación.*

Demostración. Hay dos demostraciones usuales, ambas complicadas, divertidas y polémicas. Lamentablemente, no hay espacio en este curso para incluirlas. Pero se puede contar una de ellas: el primer paso es probar que si $A \in \mathcal{DS}(n)$, entonces existe $\sigma \in S_n$ tal que $a_{i\sigma(i)} > 0$, para todo $1 \le i \le n$. Esto suele llamarse que "A tiene una diagonal no nula". Suena razonable, si filas y columnas deben sumar 1, pero la prueba no es fácil, y es conocida como Teorema de "los casamientos" de Hall.

El siguiente paso es hacer inducción en el número de entradas no nulas de A (desde n hasta n^2; ¿porqué vale $\mathrm{P}(n)$?). El paso inductivo se hace restándole a A la matriz de permutación de σ, multiplicada por $a = \min a_{i\sigma(i)}$. Esto queda con entradas positivas (una menos) y "dobleestocatizable" (todo suma justo $1 - a$). □

2.2 Aplicaciones a matrices Hermitianas.

Teorema 2.2.1 (Teorema de mayorización de Schur).
Sea $A \in \mathcal{H}(n)$. Notamos $\mathrm{d}(A) = (a_{11}, \dots, a_{nn}) \in \mathbb{R}^n$ y $\lambda(A) \in \mathbb{R}^n$ el vector de los autovalores de A. Entonces,

$$\mathrm{d}(A) \prec \lambda(A).$$

Demostración. Para demostrar que $\mathrm{d}(A) \prec \lambda(A)$ vamos a probar que $\mathrm{d}(A) = B\ \lambda(A)$, para cierta $B \in \mathcal{DS}(n)$. Como $A \in \mathcal{H}(n)$, si $D = \mathrm{diag}(\lambda(A))$, existe $U \in \mathcal{U}(n)$ tal que $A = U^*DU$. Mediante cuentas elementales de matrices, se puede verificar que cada entrada de A tiene la forma: dados $1 \le i, j \le n$,

$$a_{ij} = \sum_{k=1}^{n} \overline{u}_{ki} \lambda_k u_{kj} \ , \qquad \text{en particular,} \qquad a_{ii} = \sum_{k=1}^{n} \lambda_k |u_{ki}|^2.$$

Consideremos ahora la matriz $B = (|u_{ji}|^2)_{ij}$ que, por ser U unitaria, cumple $B \in \mathcal{DS}(n)$. Además

$$B\lambda(A) = \begin{bmatrix} |u_{11}|^2 & \cdots & |u_{n1}|^2 \\ \vdots & \ddots & \vdots \\ |u_{1n}|^2 & \cdots & |u_{nn}|^2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} \sum\limits_{k=1}^{n} |u_{k1}|^2 \lambda_k \\ \vdots \\ \sum\limits_{k=1}^{n} |u_{kn}|^2 \lambda_k \end{bmatrix} = \mathrm{d}(A).$$

Luego, el Teorema 2.1.9 completa la demostración. □

Como corolario de este teorema encontramos una nueva caracterización para los autovalores de una matriz Hermitiana. En este caso, para la suma de los k-primeros autovalores.

Corolario 2.2.2 (Principio del Máximo de Ky Fan). *Sea $A \in \mathcal{H}(n)$. Entonces para todo $1 \le k \le n$,*

$$\sum_{j=1}^{k} \mu_j(A) = \max \sum_{j=1}^{k} \langle Ax_j, x_j \rangle,$$

donde el máximo se toma sobre todas las k-uplas ortonormales $\{x_1, ..., x_k\}$ en $\mathbb{C}^n$.

Demostración. Fijemos $1 \le k \le n$. Sea $\{x_1, ..., x_k\}$ una k-upla cualquiera de vectores ortonormales. Sea $U \in \mathcal{U}(n)$ tal que sus primeras k columnas sean los vectores dados. Notemos $B = U^*AU$, que verifica $\mu(B) = \mu(A)$ y, además, $\sum\limits_{j=1}^{k} \langle Ax_j, x_j \rangle = \sum\limits_{j=1}^{k} b_{jj}$. Por el Teorema de mayorización de Schur 2.2.1,

$$\sum_{j=1}^{k} \langle Ax_j, x_j \rangle = \sum_{j=1}^{k} b_{jj} \le \sum_{j=1}^{k} \mu_j(A).$$

Si consideramos en particular una k-upla ortonormal $\{x_1, ..., x_k\}$ compuesta por autovectores de A correspondientes a los autovalores $\mu_1(A), \dots, \mu_k(A)$, obtenemos

$$\sum_{j=1}^{k} \langle Ax_j, x_j \rangle = \sum_{j=1}^{k} \langle x_j, \mu_j(A) x_j \rangle = \sum_{j=1}^{k} \mu_j(A).$$

De esta manera vemos que se alcanza la igualdad cuando se toma el máximo sobre todas las k-uplas de vectores ortonormales. $\square$

Como una consecuencia del principio máximo de Ky-Fan, obtenemos una interesante relación de mayorización entre los autovalores de las matrices Hermitianas A , B y $A + B$.

Teorema 2.2.3. *Sean A y $B \in \mathcal{H}(n)$. Sean $\mu(A)$, $\mu(B)$ y $\mu(A+B)$ los vectores que contienen los autovalores de A, B y $A+B$ respectivamente, ordenados de manera decreciente. Entonces*

$$\mu(A + B) \prec \mu(A) + \mu(B)$$

Demostración. Por el principio del máximo de Ky-Fan, para $k : 1, ..., n$, podemos escribir

$$\begin{aligned}
\sum_{j=1}^{k} \mu_j(A + B) &= \max \sum_{j=1}^{k} \langle x_j, (A + B)x_j \rangle \\
&= \max \{ \sum_{j=1}^{k} \langle x_j, Ax_j \rangle + \sum_{j=1}^{k} \langle x_j, Bx_j \rangle \} \\
&\le \max \sum_{j=1}^{k} \langle x_j, Ax_j \rangle + \max \sum_{j=1}^{k} \langle x_j, Bx_j \rangle \\
&= \sum_{j=1}^{k} \mu_j(A) + \sum_{j=1}^{k} \mu_j(B).
\end{aligned}$$

Finalmente, para $k = n$ hay igualdad, porque $\mathrm{tr}(A + B) = \mathrm{tr}\ (A) + \mathrm{tr}(B)$. $\square$

Ejercicio 2.2.4. Dada $C \in \mathcal{M}_n(\mathbb{C})$, sea

$$\widehat{C} = \begin{pmatrix} 0 & C \\ C^* & 0 \end{pmatrix} \in \mathcal{M}_{2n}(\mathbb{C}).$$

Probar que $\sigma\left(\widehat{C}\right) = \{\pm s_i\,(C)\}$ (con las mismas multiplicidades). Es decir,

$$\mu(\widehat{C}) = (s_1(C), \cdots, s_n(C), -s_n(C), \cdots, -s_1(C)). \tag{7}$$

$\triangle$

Corolario 2.2.5. *Sean $A, B \in \mathcal{H}(n)$. Entonces*

$$s(A+B) \prec_w s(A) + s(B).$$

Demostración. Notar que $\widehat{A+B} = \widehat{A} + \widehat{B}$. Aplicando el Ejercicio anterior y las desigualdades resultantes de la relación $\mu(\widehat{A+B}) \prec \mu(\widehat{A}) + \mu(\widehat{B})$ para $1 \leq k \leq n$, se obtiene que $s(A+B) \prec_w s(A) + s(B)$. $\square$

Observación 2.2.6. Notar que el resultado anterior es necesario para verificar que las normas $\|\cdot\|_{(k)}$ de Ky Fan, definidas en el Ejemplo 1.7.1, cumplen la desigualdad triangular.

2.3 Normas unitariamente invariantes.

Definición 2.3.1. Una norma $|||\cdot|||$ en $\mathcal{M}_n(\mathbb{C})$ se dice que es una norma unitariamente invariante (NUI) , si cumple que

$$|||UAV||| = |||A|||$$

para toda $A \in \mathcal{M}_n(\mathbb{C})$ y $U, V \in \mathcal{U}(n)$. Notar que, en tal caso, por el Teorema 1.6.3 se tiene que $|||A||| = |||\Sigma(A)|||$. $\triangle$

Definición 2.3.2. Dada una norma $N(\cdot)$ unitariamente invariante, se define la función $g_N : \mathbb{R}^n \to \mathbb{R}$ como

$$g_N(x_1, \ldots, x_n) = N\left(\operatorname{diag}\left(x_1, \ldots, x_n\right)\right)$$

$\triangle$

Proposición 2.3.3. *Sea N una NUI. Entonces, para todo $x \in \mathbb{R}^n$,*

1. *g_N es una norma en $\mathbb{R}^n$.*
2. *$g_N(x_1, \ldots, x_n) = g_N(|x_1|, \ldots, |x_n|)$.*
3. *$g_N(x_1, \ldots, x_n) = g_N(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$, para toda $\sigma \in S_n$.*

Observación 2.3.4. Una función $f : \mathbb{R}^n \to \mathbb{R}$ que cumple los ítems 1, 2 y 3 de la Proposición anterior se denomina **gauge simétrica**.

Demostración.

1. Ejercicio para el lector.

2. Sea $x_j = \omega_j |x_j|$ donde $w_j = e^{i\theta_j}$. Luego, como $\operatorname{diag}(\omega_1, \ldots, \omega_n) \in \mathcal{U}(n)$,

$$\begin{aligned} g_N(|x_1|, \ldots, |x_n|) &= N(\operatorname{diag}(|x_1|, \ldots, |x_n|)) \\ &= N(\operatorname{diag}(\omega_1, \ldots, \omega_n) \operatorname{diag}(|x_1|, \ldots, |x_n|)) \\ &= N(\operatorname{diag}(x_1, \ldots, x_n)) \\ &= g_N(\operatorname{diag}(x_1, \ldots, x_n)). \end{aligned}$$

3. Sea P_σ la matriz unitaria tal que

$$P_\sigma \operatorname{diag}(x_1, \ldots, x_n) P_\sigma^{-1} = \operatorname{diag}\left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right).$$

   Entonces,

$$\begin{aligned} g_N(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) &= N(\operatorname{diag}\left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right)) \\ &= N(P_\sigma \operatorname{diag}(x_1, \ldots, x_n) P_\sigma^{-1}) \\ &= N(\operatorname{diag}(x_1, \ldots, x_n)) \\ &= g_N(\operatorname{diag}(x_1, \ldots, x_n)). \end{aligned}$$

□

Lema 2.3.5. *Si g es una función gauge simetrica, entonces, g es monótona, es decir, si $|x_i| \leq |y_i|$ para todo $i \in \{1, \ldots, n\}$, entonces, $g(x) \leq g(y)$.*

Demostración. Sin pérdida de generalidad, podemos suponer que $x, y \in \mathbb{R}^n_+$. Por un argumento inductivo, es suficiente verificar que si $t \in [0,1]$, entonces, $g(y_1, \ldots, ty_k, \ldots, y_n) \leq g(y_1, \ldots, y_k, \ldots, y_n)$. En efecto,

$$\begin{aligned} g(y_1, \ldots, ty_k, \ldots, y_n) &= g\left(\left(\frac{1+t}{2}y_1, \ldots, \frac{1+t}{2}y_k, \ldots, \frac{1+t}{2}y_n\right) + \right. \\ &\quad + \left.\left(\frac{1-t}{2}y_1, \ldots, \frac{1-t}{2}(-y_k), \ldots, \frac{1-t}{2}y_n\right)\right) \\ &\leq \frac{1+t}{2}\, g(y) + \frac{1-t}{2}\, g(y_1, \ldots, -y_k, \ldots, y_n) \\ &= \frac{1+t}{2}\, g(y) + \frac{1-t}{2}\, g(y) = g(y). \end{aligned}$$

□

Teorema 2.3.6. *Sea g es una función gauge simétrica y $x, y \in \mathbb{R}^n_+$ tales que $x \prec_w y$. Entonces, $g(x) \leq g(y)$.*

Demostración. Como $x \prec_w y$, por el Ejercicio 2.1.4, existe $u \in \mathbb{R}^n$ tal que $x \leqslant u \prec y$. Por el Lema anterior, $g(x) \leq g(u)$. Dado $\sigma \in S_n$, notemos $y_\sigma = (y_{\sigma(1)}, \dots, y_{\sigma(n)})$. Por el Teorema 2.1.9, existe $A \subseteq S_n$ tal que $u = \sum_{\sigma \in A} \lambda_\sigma y_\sigma$ para ciertos λ_σ tales que $\lambda_\sigma \in [0,1]$ y $\sum_{\sigma \in A} \lambda_\sigma = 1$. Luego

$$g(u) = g\left(\sum_{\sigma \in A} \lambda_\sigma y_\sigma\right) \leq \sum_{\sigma \in A} \lambda_\sigma g(y_\sigma) = \sum_{\sigma \in A} \lambda_\sigma g(y) = g(y).$$

□

Teorema 2.3.7.

1. *Si N es una NUI, entonces, g_N es una función gauge simetrica.*

2. *Si g es una función gauge simetrica, entonces,*

$$\|A\|_g = g(s_1(A), \dots, s_n(A)) \ , \quad A \in \mathcal{M}_n(\mathbb{C})$$

es una NUI en $\mathcal{M}_n(\mathbb{C})$.

Demostración. 1. Esto es la Proposición 2.3.3.

2. Sólo demostraremos la desigualdad triangular. Las demás propiedades quedan como ejercicio para el lector. Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Como $s(A+B) \prec_w s(A) + s(B)$, se tiene

$$\begin{aligned} \|A+B\|_g &= g(s(A+B)) \leq g(s(A) + s(B)) \\ &\leq g(s(A)) + g(s(B)) = \|A\|_g + \|B\|_g. \end{aligned}$$

□

Teorema 2.3.8. *Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Entonces, $N(A) \leq N(B)$ para toda NUI "N", si y sólo si $\|A\|_{(k)} \leq \|B\|_{(k)}$ para $k = 1, \dots, n$.*

Demostración. Es consecuencia del Teorema 2.3.6 para funciones gauge simétricas, y del Teorema 2.3.7. En efecto, si $\|A\|_{(k)} \leq \|B\|_{(k)}$ para $k = 1, \dots, n$, entonces $s(A) \prec_w s(B)$. Por lo tanto $g(s(A)) \leq g(s(B))$ para toda función gauge simetrica. La recíproca es evidente. □

Corolario 2.3.9. *Sean $A, B \in \mathcal{M}_n(\mathbb{C})^+$ tales que $A \leq B$. Entonces, $N(A) \leq N(B)$ para toda norma unitariamente invariante N.*

Demostración. Es consecuencia del Corolario 1.5.4, porque

$$s_k(A) = \mu_k(A) \leq \mu_k(B) = s_k(B) \ , \quad 1 \leq k \leq n.$$

□

2.4 Mayorización de matrices Hermitianas

Hasta el momento sólo hemos visto resultados relacionados con la mayorización de vectores. Pero a cada matriz $A \in \mathcal{H}(n)$ se le puede asociar el vector $\mu(A) \in \mathbb{R}^n$ formado por todos los autovalores de A. Esto permite la siguiente definición,

Definición 2.4.1. Si A, $B \in \mathcal{H}(n)$, se dice que A está mayorizada por B y se escribe $A \prec B$ si se verifica que $\mu(A) \prec \mu(B)$. Es decir, $A \prec B$ si $\operatorname{tr} A = \operatorname{tr} B$ y

$$\sum_{j=1}^{k} \mu_j(A) \quad \leq \quad \sum_{j=1}^{k} \mu_j(B), \quad 1 \leq k \leq n, . \tag{8}$$

$\triangle$

Definición 2.4.2. Sea $P \in \mathcal{H}(n)$ un proyector (o sea $P = P^2 = P^*$) y $A \in \mathcal{M}_n(\mathbb{C})$. Se define el **pinching** de A como la matriz

$$C_P(A) := PAP + (I - P)A(I - P).$$

Por ejemplo, si P proyecta sobre las primeras k cordenadas en $\mathbb{C}^n$, si

$$A = \begin{pmatrix} B & C \\ D & E \end{pmatrix} \quad \Rightarrow \quad C_P(A) = \begin{pmatrix} B & 0 \\ 0 & E \end{pmatrix},$$

donde los bloques tienen los tamaños adecuados (por ejemplo, $B \in \mathcal{M}_k(\mathbb{C})$).

Proposición 2.4.3. *Sea $A \in \mathcal{H}(n)$ y $P \in \mathcal{H}(n)$ un proyector. Entonces*

$$C_P(A) \prec A.$$

Demostración. Sea $U = P - (I - P) \in \mathcal{U}(n)$. Es fácil ver que

$$2\ C_P(A) = A + UAU = A + UAU^{-1}.$$

Pero, como $\mu(UAU^{-1}) = \mu(A)$, por el Teorema 2.2.3 se tiene

$$2\ \mu(C_P(A)) \prec \mu(A) + \mu(UAU^{-1}) = 2\ \mu(A),$$

por lo que $C_P(A) \prec A$. $\square$

Ejercicio: Clarificar en qué sentido la Proposición 2.4.3 (o mejor su extensión por inducción a k bloques), es una generalización del Toerema de mayorización de Schur.

2.5 Teoremas de Lindskii y sus aplicaciones.

En esta sección usaremos la siguiente notación: Si $B \in \mathcal{M}_n(\mathbb{C})$, llamaremos

$$s(B) = (s_1(B), \ldots, s_n(B)) = \mu(|B|),$$

al vector de valores singulares de B, ordenados en forma **decreciente**.

Teorema 2.5.1 (Lidskii 1). *Sean $A, B \in \mathcal{H}(n)$. Entonces*

$$\mu(A) - \mu(B) \prec \mu(A - B) \prec \mu(A) - \lambda(B).$$

Ejercicio 2.5.2.

1. Decir porqué es incorrecta la siguiente prueba de la primera parte del Teorema de Lidskii 1: Si $A, B \in \mathcal{H}(n)$, por el Teorema 2.2.3, $\mu(A) \prec \mu(A - B) + \mu(B)$. Por lo tanto, para todo $1 \le k \le n$,

$$\sum_{j=1}^{k} \mu_j(A) - \mu_j(B) \le \sum_{j=1}^{k} \mu_j(A - B).$$

   Deducimos entonces que $\mu(A) - \mu(B) \prec \mu(A - B)$. La igualdad, para $k = n$, sale tomando trazas.

2. Demostrar (bien) la otra parte del Teorema. $\triangle$

Teorema 2.5.3 (Lidskii 2). *Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Entonces*

$$|s(A) - s(B)| \prec_w s(A - B).$$

Una ves resuelto el Ejercicio 2.5.2, se entenderá porque es necesaria (y suficiente) esta versión más técnica del Teorema de Lidskii:

Teorema 2.5.4 (Lidskii 3). *Sean $A, B \in \mathcal{H}(n)$ y $1 \le i_1 < \ldots < i_k \le n$. Entonces*

$$\sum_{j=1}^{k} \mu_{i_j}(A + B) \le \sum_{j=1}^{k} \mu_{i_j}(A) + \sum_{j=1}^{k} \mu_j(B).$$

Lema 2.5.5. *Sea $A \in \mathcal{H}(n)$ y $\mathcal{S} \subseteq \mathbb{C}^n$ un subespacio de dimensión n-1. Si $A_{\mathcal{S}} = P_{\mathcal{S}} A P_{\mathcal{S}} : \mathcal{S} \to \mathcal{S}$ es el comprimido de A a $\mathcal{S}$, entonces:*

1. $\mu_1(A) \ge \mu_1(A_{\mathcal{S}}) \ge \mu_2(A) \ge \mu_2(A_{\mathcal{S}}) \ge \ldots$.
2. *Sea $v_1, \ldots, v_n$ una bon de autovectores de A, asociados a $\mu_1(A), \ldots, \mu_n(A)$ (respectivamente).*
   a. *Si $v_1, \ldots, v_k \in \mathcal{S}$, entonces, $\mu_i(A) = \mu_i(A_{\mathcal{S}})$ para $1 \le i \le k$.*
   b. *Si $v_k, \ldots, v_n \in \mathcal{S}$, entonces, $\mu_{i-1}(A_{\mathcal{S}}) = \mu_i(A)$, $k \le i \le n$.*

Demostración. Se deduce de la Observación 1.5.6 (y su demostración). $\square$

Demostración del Teorema de Lidskii 3. Lo haremos por inducción sobre n. El caso $n = 1$ es trivial. Sea $n > 1$. Considermos una bon $\{u_1, \ldots, u_n\}$ de autovectores de A asociados a $\mu(A)$ y $\{w_1, \ldots, w_n\}$ una bon de autovectores de $A + B$ asociados a $\mu(A + B)$. Dividiremos la prueba en casos.

Caso 1: $(i_k < n)$ Sea $\mathcal{S}$ el subespacio generador por los vectores $w_1, \ldots, w_{n-1}$. Por el item 2 del lema, para $k = 1, \ldots, n-1$ se tiene que $\mu_k(A+B) = \mu_k((A+B)_{\mathcal{S}})$. Luego

$$\begin{aligned}\sum_{j=1}^{k} \mu_{i_j}(A+B) &= \sum_{j=1}^{k} \mu_{i_j}((A+B)_{\mathcal{S}}) \\ &\le \sum_{j=1}^{k} \mu_{i_j}(A_{\mathcal{S}}) + \sum_{j=1}^{k} \mu_j(B_{\mathcal{S}}) \\ &\le \sum_{j=1}^{k} \mu_{i_j}(A) + \sum_{j=1}^{k} \mu_j(B),\end{aligned}$$

donde, en la primera deisgualdad, se usó la HI para $A_{\mathcal{S}}$ y $B_{\mathcal{S}}$.

Caso 2: $(1 < i_1)$ Sea $\mathcal{S}$ el subespacio generador por $u_2, \ldots, u_n$. Entonces,

$$\begin{aligned}\sum_{j=1}^{k} \mu_{i_j}(A+B) &\le \sum_{j=1}^{k} \mu_{i_j-1}(A_{\mathcal{S}} + B_{\mathcal{S}}) \\ &\le \sum_{j=1}^{k} \mu_{i_j-1}(A_{\mathcal{S}}) + \sum_{j=1}^{k} \mu_j(B) \\ &\le \sum_{j=1}^{k} \mu_{i_j}(A) + \sum_{j=1}^{k} \mu_j(B).\end{aligned}$$

Caso 3: $(i_1 = 1,\ i_k = n)$ Sean $1 < l_1 < \ldots < l_{n-k} < n$ tales que $\{n - l_1 + 1, \ldots, n - l_{n-k} + 1\} = \{i_1, \ldots, i_k\}^c$. Aplicamos el caso 1 o 2 para $l_1, \ldots, l_{n-k}$, $-A$, $-B$ y $-A-B$. (Notar que $\mu_j(-C) = -\lambda_j(C) = -\mu_{n-j+1}(C)$). Se tiene que

$$\sum_{j=1}^{n-k} \mu_{l_j}(-A-B) \le \sum_{j=1}^{n-k} \mu_{l_j}(-A) + \sum_{j=1}^{n-k} \mu_j(-B).$$

Luego

$$-\sum_{j=1}^{n-k} \mu_{n-l_j+1}(A+B) \le -\sum_{j=1}^{n-k} \mu_{n-l_j+1}(A) - \sum_{j=1}^{n-k} \mu_{n-j+1}(B),$$

y por lo tanto

$$\sum_{j=1}^{n-k} \mu_{n-l_j+1}(A+B) \ge \sum_{j=1}^{n-k}\ muavin - l_j + 1A + \sum_{j=1}^{n-k} \mu_{n-j+1}(B).$$

Finalmente, como $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ se tiene que

$$\sum_{j=1}^{k} \mu_{i_j}(A+B) \le \sum_{j=1}^{k} \mu_{i_j}(A) + \sum_{j=1}^{k} \mu_j(B).$$

$\square$

Demostración del Teorema de Lidskii 1. Reemplazando en el tercer Teorema de Lidskii $A+B$ por A, A por B y B por $A-B$ se tiene que

$$\max_{1\leq i_1<\ldots<i_k\leq n}\left\{\sum_{j=1}^{k}\mu_{i_j}(A)-\mu_{i_j}(B)\right\}\leq\sum_{j=1}^{k}\mu_j(A-B)\,. \tag{9}$$

Entonces, $\mu(A)-\mu(B)\prec\mu(A-B)$. $\square$

Demostración del Teorema de Lidskii 2. Si recordamos el Ejercicio 2.2.4, obtenemos, para $1\leq k\leq n$,

$$\max_{1\leq i_1<\ldots<i_k\leq 2n}\left\{\sum_{j=1}^{k}\mu_{i_j}\left(\widehat{A}\right)-\mu_{i_j}\left(\widehat{B}\right)\right\}=\sum_{j=1}^{k}|s(A)-s(B)|_j^{\downarrow}\ ,$$

porque si $1\leq j\leq n$, entonces, por la fórmula (7),

$$\mu_j\left(\widehat{A}\right)-\mu_j\left(\widehat{B}\right)=s_j(A)-s_j(B)\quad \text{y}$$

$$\mu_{2n-j+1}\left(\widehat{A}\right)-\mu_{2n-j+1}\left(\widehat{B}\right)=-(s_j(A)-s_j(B)).$$

Por otro lado, aplicando la fórmula (9) a $\widehat{A}$, $\widehat{B}$ y a $\widehat{A-B}=\widehat{A}-\widehat{B}$, se tiene

$$\max_{1\leq i_1<\ldots<i_k\leq 2n}\left\{\sum_{j=1}^{k}\mu_{i_j}\left(\widehat{A}\right)-\mu_{i_j}\left(\widehat{B}\right)\right\}\leq\sum_{j=1}^{k}\mu_j\left(\widehat{A-B}\right)=\sum_{j=1}^{k}s_j(A-B)\,,$$

y podemos concluir que $\displaystyle\sum_{j=1}^{k}|s(A)-s(B)|_j^{\downarrow}\leq\sum_{j=1}^{k}s_j(A-B)\,.$ $\square$

Corolario 2.5.6. *Sea $|||\cdot|||$ una NUI en $\mathcal{M}_n(\mathbb{C})$ y $A,B\in\mathcal{M}_n(\mathbb{C})$. Entonces*

$$|||\Sigma(A)-\Sigma(B)|||\leq|||A-B|||.$$

Demostración. Notar que $s(\Sigma(A)-\Sigma(B))=|s(A)-s(B)|^{\downarrow}$. Por lo tanto el Teorema de Lidskii 2 implica que, para $1\leq k\leq n$,

$$\|\Sigma(A)-\Sigma(B)\|_{(k)}\leq\|A-B\|_{(k)}\ .$$

Luego se aplica el Teorema 2.3.8. $\square$

Corolario 2.5.7. *Sea $|||\cdot|||$ una NUI en $\mathcal{M}_n(\mathbb{C})$ y $A\in\mathcal{G}l(n)$. Sea U la única matriz unitaria tal que $A=U|A|$ (i.e. $U=A|A|^{-1}\in\mathcal{U}(n)$). Entonces*

$$d_{|||\cdot|||}(A,\mathcal{U}(n))=|||A-U|||=|||\Sigma(A)-I|||.$$

Demostración. Sea $V \in \mathcal{U}(n)$. Entonces $\Sigma(V) = I$ y, por el Corolario anterior, $|||A - V||| \geq |||\Sigma(A) - I|||$. Por otra parte, sea $W \in \mathcal{U}(n)$ tal que $|A| = W\Sigma(A)W^*$. Entonces

$$A - U = UW\Sigma(A)W^* - UWW^* = UW(\Sigma(A) - I)W^*.$$

Dado que $|||\cdot|||$ es una NUI, resulta que $|||A - U||| = |||\Sigma(A) - I|||$. □

Ejercicio 2.5.8. Sea $A \in \mathcal{M}_n(\mathbb{C})$, y sea $\mathcal{M}_n(\mathbb{C})_1$ el conjunto de matrices en $\mathcal{M}_n(\mathbb{C})$ de rango uno (o cero). Sea

$$\Sigma_1(A) = \operatorname{diag}(0, s_2(A), \ldots, s_n(A)).$$

Probar que si N es una NUI, $d_N(A, \mathcal{M}_n(\mathbb{C})_1) = N(\Sigma_1(A))$ y se alcanza en la matriz

$$A_1 = UW \operatorname{diag}(s_1(A), 0, \ldots, 0)\, W^*,$$

donde $A = UW\Sigma(A)W^*$. Probar, aparte, que A_1 no depende de la matriz $U \in \mathcal{U}(n)$ elegida para realizar la descomposición polar de A, pero sí puede depender de W (si $s_1(A)$ tiene multiplicidad mayor que uno para $|A|$).

Mostrar que otra manera de encontrar A_1 es tomando $A_1 = s_1(A)yx^*$, donde x es un vector tal que $\|x\|_2 = 1$ y $\|Ax\|_2 = \|A\|_{sp} = s_1(A)$, e $y = Ux$.

Generalizar el resultado al conjunto de matrices de rango a lo sumo k.

3 Desigualdades de matrices.

3.1 Producto de Hadamard

Definición 3.1.1. Dadas $A, B \in \mathcal{M}_{mn}(\mathbb{C})$ se define el producto de Hadamard $A \circ B$ como la matriz

$$A \circ B = [a_{ij}\ b_{ij}]_{ij}.$$

Notar que este producto tiene sentido tanto para matrices como para vectores. △

Teorema 3.1.2 (Teorema de Schur). *Sean A, $B \in \mathcal{M}_n(\mathbb{C})^+$, entonces $A \circ B \in \mathcal{M}_n(\mathbb{C})^+$. Además, si $A > 0$ y $B > 0$, entonces $A \circ B > 0$.*

Demostración. La segunda parte se deduce de la primera. En efecto, si $A > 0$ y $B > 0$, existen números $a, b > 0$ tales que $A \geq aI$ y $B \geq bI$. Entonces, aplicando dos veces el caso que aún no hemos probado, obtendríamos

$$A \circ B \geq aI \circ B \geq aI \circ bI = abI > 0.$$

Supongamos entonces que A, $B \in \mathcal{M}_n(\mathbb{C})^+$. Es fácil ver que deben existir vectores $v_i, w_i \in \mathbb{C}^n$, $1 \leq i \leq n$, tales que

$$A = \sum_{i=1}^n v_i v_i^* \quad \text{y} \quad B = \sum_{i=1}^n w_i w_i^*.$$

En efecto, basta tomar como $v_i = C_i(A^{1/2})$ (resp. $w_i = C_i(B^{1/2})$) y hacer algunas cuentas de matrices. Notar que si $x \in \mathbb{C}^n$, entonces $xx^* = (x_k \overline{x_l})_{kl}$. Por lo tanto,

$$A \circ B = \sum_{i,j=1}^n v_i v_i^* \circ w_j w_j^* = \sum_{i,j=1}^n (v_i \circ w_j)(v_i^* \circ w_j^*) \geq 0,$$

porque $v_i^* \circ w_j^* = (v_i \circ w_j)^*$. □

3.2 Determinantes

Teorema 3.2.1 (Desigualdad de Hadamard). *Si $A \in \mathcal{M}_n(\mathbb{C})^+$, entonces*

$$\det A \leq \prod_{i=1}^{n} a_{ii}.$$

Si $A > 0$, la igualdad vale si y sólo si A es diagonal.

Demostración. Podemos suponer que $A > 0$, y entonces $a_{ii} > 0$, $1 \leq i \leq n$. Consideramos la matriz diagonal

$$D = \operatorname{diag}\left(a_{11}^{1/2}, \dots, a_{nn}^{1/2}\right).$$

Entonces $B = D^{-1}AD^{-1} = (a_{ii}^{-1/2} a_{jj}^{-1/2} a_{ij})_{ij}$ tiene unos en la diagonal. Además,

$$\det B = \det A \det(D)^{-2} = \det A \prod_{i=1}^{n} a_{ii}^{-1}.$$

Por lo tanto, sería suficiente mostrar que $\det B \leq 1$. Aplicando la desigualdad aritmético-geométrica [1] obtenemos,

$$\det(B) = \prod_{i=1}^{n} \lambda_i(B) \leq \Big[\frac{1}{n}\sum_{i=1}^{n} \lambda_i(B)\Big]^n = (\operatorname{tr}\frac{B}{n})^n = 1.$$

y esto prueba el resultado. Con respecto a la igualdad, si la hubiera en la desigualdad aritmético-geométrica, entonces los números involucrados deben ser todos iguales. Es decir que todos los $\lambda_i(B) = 1$. Pero entonces, como $B \geq 0$, debe ser $B = I$, o sea $A = D^2$. $\square$

Corolario 3.2.2. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces*

$$|\det A| \leq \prod_{i=1}^{n} \|C_i(A)\|_2.$$

Demostración. Se aplica la desigualdad de Haramard a la matriz $B = A^*A \geq 0$. Notar que $\det B = |\det A|^2$ y que $B_{ii} = \|C_i(A)\|_2^2$, para $1 \leq i \leq n$. $\square$

Lema 3.2.3. *Sean $A \in \mathcal{G}l(n)^+$, $\alpha(A) = \frac{detA}{detA_{11}}$ donde A_{11} es la submatriz principal de orden $n-1$ de A que resulta de borrar la primer fila y columna de A y $E_{11} \in \mathcal{M}_n(\mathbb{C})$ es la matriz cuya entrada $(1,1)$ es 1 y las demás son todas nulas. Entonces $A - tE_{11} \geq 0$ si y sólo si $t \leq \alpha(A)$.*

Demostración. Es fácil ver, desarrollando por la primera columna, que

$$\det(A - tE_{11}) = \det A - t \det A_{11}. \tag{10}$$

[1] Si $a_1, \dots, a_n > 0$, entonces $(\prod_{i=1}^{n} a_i)^{1/n} \leq 1/n \sum_{i=1}^{n} a_i$. Se demuestra rápidamente usando que el log es una función cóncava.

Luego, $\det(A - tE_{11}) \geq 0$ si y sólo si $t \leq \alpha(A)$. Por otro lado, todas las demás submatrices principales de $A - tE_{11}$ obtenidas con las últimas i filas y columnas, son las mismas que las respectivas de A. Por lo tanto, el determinante de cada una de ellas es positivo. Luego, por el Teorema 1.5.9 (hecho desde abajo), tenemos el resultado para desigualdades estrictas. El caso general sale tomando límite. $\square$

Teorema 3.2.4 (Desigualdad de Oppenheim). *Si $A, B \in \mathcal{M}_n(\mathbb{C})^+$, entonces*

$$\det A \prod_{i=1}^{n} b_{ii} \leq \det A \circ B$$

Demostración. Si $\det A = 0$, el resultado se deduce del Teorema de Schur 3.1.2, que asegura que $A \circ B \geq 0$. Supongamos, entonces, que $A > 0$. La demostración la realizaremos por inducción sobre n. Si $n = 1$, el resultado es inmediato. Sea $n \geq 2$ y supongamos el resultado válido para todas las matrices de dimensión $n - 1$. Entonces, con las notaciones del Lema 3.2.3, sabemos que

$$\det A_{11} \prod_{i=2}^{n} b_{ii} \leq \det A_{11} \circ B_{11}.$$

Por el Lema 3.2.3, si $\alpha = (\det A_{11})^{-1} \det A$, entonces $A - \alpha E_{11} \geq 0$. El Teorema de Schur 3.1.2 dice que $(A - \alpha E_{11}) \circ B \geq 0$. Aplicando la fórmula (10), como $E_{11} \circ B = b_{11} E_{11}$ y $(A \circ B)_{11} = A_{11} \circ B_{11}$, resulta que

$$0 \leq \det(A \circ B - \alpha E_{11} \circ B) = \det A \circ B - \alpha b_{11} \det(A_{11} \circ B_{11}).$$

Aplicando la hipótesis inductiva, obtenemos

$$\det A \circ B \geq \alpha b_{11} \ \det \ A_{11} \circ B_{11} \geq \alpha b_{11} \det A_{11} \prod_{i=2}^{n} b_{ii} = \det A \prod_{i=1}^{n} b_{ii}$$

y el teorema queda demostrado. $\square$

Teorema 3.2.5 (Desigualdad de Fisher). *Si $A, P \in \mathcal{M}_n(\mathbb{C})^+$, con P un proyector, entonces*

$$\det A \leq \det \ C_P(A).$$

Recordamos que $C_P(A) = PAP + (I - P)A(I - P)$.

Demostración. Supongamos que $\dim R(P) = k$. Conjugando a P y a A con alguna matriz unitaria (lo que no cambia los determinantes), podemos suponer que $R(P)$ es el subespacio generado por los primeros k elementos de la base canónica de $\mathbb{C}^n$. O, lo que es lo mismo, que $P = \operatorname{diag}(1, \dots, 1, 0, \dots, , 0)$, donde los unos llegan hasta el lugar k.

Dado $r \in \mathbb{N}$, llamemos $E_r \in \mathcal{M}_r(\mathbb{C})^+$ a la matriz con todas sus entradas iguales a 1. Notar que $E_r \geq 0$ porque $0 \leq E_r^* E_r = E_r^2 = r E_r$. Consideremos la matriz de bloques

$$B = \begin{pmatrix} E_k & 0 \\ 0 & E_{n-k} \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})^+,$$

que verifica que $A \circ B = C_P(A)$. Ahora, aplicando la desigualdad de Oppenheim, tenemos que

$$\det A = \det A \prod_{i=1}^{n} b_{ii} \leq \det A \circ B = \det C_P(A).$$

$\square$

De los resultados anteriores obtenemos la siguiente relación para el determinante del producto convencional de matrices y el producto de Hadamard.

Teorema 3.2.6. *Si A, $B \in \mathcal{M}_n(\mathbb{C})^+$, entonces*

$$\det A \ \det B \leq \det A \circ B.$$

Demostración. El Teorema se deduce de las desigualdades de Hadamard y de Oppenheim. En efecto,

$$\det A \ \det B \leq \det A \prod_{i=1}^{n} b_{ii} \leq \det A \circ B.$$

$\square$

3.3 Partes reales

Si $A \in \mathcal{M}_n(\mathbb{C})$, se llama **parte real** de A a

$$\operatorname{Re} A = \frac{A + A^*}{2}.$$

Si $x \in \mathbb{C}^n$, $\operatorname{Re} x$ es el vector de partes reales de sus cordenadas.

Proposición 3.3.1 (Ky Fan). *Dada $A \in \mathcal{M}_n(\mathbb{C})$, sea $\mu(A) \in \mathbb{C}^n$ el vector de autovalores de A en algún orden. Entonces*

$$\operatorname{Re} \mu(A) \prec \mu(\operatorname{Re} A)$$

Demostración. Ordenemos al vector $\mu(A)$ de tal modo que

$$\operatorname{Re} \mu_1(A) \geq \operatorname{Re} \mu_2(A) \geq \ldots \geq \operatorname{Re} \mu_n(A).$$

Sea $\{x_1, \ldots, x_n\}$ una base ortonormal respecto a la cual A es una matriz triangular superior (que existe por el Teorema 1.2.4). En particular, $\langle Ax_i, x_i \rangle = \mu_i(A)$. Dado $1 \leq k \leq n$, si $\mathcal{S}$ es el subespacio generado por $x_1, \ldots, x_k$ y P el proyector ortogonal sobre S. Entonces

$$\begin{aligned}
\sum_{j=1}^{k} \operatorname{Re} \mu_j(A) &= \sum_{j=1}^{k} \langle \operatorname{Re} A x_j, x_j \rangle = \operatorname{tr} P \operatorname{Re} A P \\
&= \sum_{j=1}^{k} \mu_j(P \operatorname{Re}(A) P) \leq \sum_{j=1}^{k} \mu_j(C_P(\operatorname{Re} A)) \leq \sum_{j=1}^{k} \mu_j(\operatorname{Re} A).
\end{aligned}$$

En las últimas desigualdades se usó que $\mu_j\,(P\operatorname{Re}(A)P) \in \sigma\,(C_P(\operatorname{Re} A))$, para $1 \le j \le k$ y, además, la Proposición 2.4.3. La igualdad para $k = n$ se debe a que

$$\operatorname{Re}\operatorname{tr}(A) = \frac{\operatorname{tr}(A) + \overline{\operatorname{tr}(A)}}{2} = \operatorname{tr}\left(\frac{A + A^*}{2}\right) = \operatorname{tr}(\operatorname{Re} A).$$

□

Proposición 3.3.2 (Kittaneh '95). *Sean $A, B \in \mathcal{M}_n(\mathbb{C})$ tales que $AB \in \mathcal{H}(n)$. Entonces,*

$$\|AB\| \le \|\operatorname{Re}(BA)\|$$

Demostración. Comencemos notando que los autovalores de BA son los mismos que los de AB (Ejercicio: verificarlo) y por ende son todos reales. Luego, usando la Proposición 3.3.1, obtenemos que

$$\mu(BA) = \operatorname{Re}(\mu(BA)) \prec \mu(\operatorname{Re}(BA)).$$

Es fácil ver que esto implica que

$$\rho(AB) = \rho(BA) = \max_{1\le k\le n} |\mu_k(BA)| \le \max_{1\le k\le n} |\mu_k(\operatorname{Re} BA)| = \rho(\operatorname{Re}(BA)), \tag{11}$$

porque si $x \prec y$, entonces $y_n^\downarrow \le x_n^\downarrow \le x_1^\downarrow \le y_1^\downarrow$. Pero $\max_k |x_k| = \max\{x_1^\downarrow, -x_n^\downarrow\}$, y lo mismo para y. Por lo tanto, Como AB y $\operatorname{Re} BA \in \mathcal{H}(n)$,

$$\|AB\| = \rho(AB) \le \rho(\operatorname{Re}(BA)) = \|\operatorname{Re}(BA)\|.$$

□

Proposición 3.3.3 (Corach-Porta-Recht, '93). *Sean $T, S \in \mathcal{H}(n)$ y supongamos que S es inversible. Entonces,*

$$\|STS^{-1} + S^{-1}TS\| \ge 2\,\|T\|$$

Demostración. Aplicar la desigualdad de Kittaneh a $A = TS^{-1}$ y $B = S$. □

Observación 3.3.4. En realidad, la desigualdad de Kittaneh, y por ende también la CPR, son ciertas para toda NUI. Esto se hace generalizando la ecuación (11) a una mayorización débil

$$s(AB) = |\mu(AB)|^\downarrow = |\mu(BA)|^\downarrow \prec_w |\mu(\operatorname{Re} BA)|^\downarrow = s(\operatorname{Re} BA),$$

lo que saldría usando Ky Fan 3.3.1 y que $f(x) = |x|$ es una función convexa. Se propone la verificación como ejercicio, dado que el Corolario 2.1.12 también lo es.

Notar que se está usando que $\mu(AB) = \mu(BA)$, es decir, no sólo tienen los mismos autovalores, sino las mismas multiplicidades. Esto es cierto, y es otro ejercicio para el lector. △

4 Teoría de Perrón-Frobenius.

4.1 Matrices de entradas positivas.

A lo largo de este capítulo, usaremos la siguiente notación: dadas $A, B \in \mathcal{M}_n(\mathbb{C})$ y $x \in \mathbb{C}^n$:

1. $A \geqslant 0$ si $A_{ij} \geq 0$, $1 \leq i, j \leq n$.
2. $A > 0$ si $A_{ij} > 0$, $1 \leq i, j \leq n$.
3. $|A| = (|a_{ij}|)_{ij}$ y analogamente $|x| = (|x_1|, \dots, |x_n|)$.
4. $A \leqslant B$ si $a_{ij} \leq b_{ij}$, $1 \leq i, j \leq n$.
5. El vector $(1, 1, \dots, 1)$ será denotado por medio de $\mathbf{e}$.

El objetivo de este capítulo es la demostración del Teorema de Perrón, para matrices de entradas estrictamente positivas y sus generalizaciones a matrices de entradas no negativas.

Teorema 4.1.1 (Perrón). *Sea $A > 0$. Entonces vale:*

1. $\rho(A) > 0$ *y* $\rho(A) \in \sigma(A)$.
2. *Existe* $x \in \mathbb{R}^n$ *tal que* $x > 0$ *y* $Ax = \rho(A)x$.
3. *Si* $y \geqslant 0$, $y \neq 0$ *y* $Ay = \lambda y$, *entonces,* $\lambda = \rho(A)$ *e* $y > 0$.
4. $\rho(A)$ *es raíz simple del polinomio característico de* A.
5. *Si* $\lambda \in \sigma(A)$ *y* $\lambda \neq \rho(A)$, *entonces,* $|\lambda| < \rho(A)$.
6. *Si* $\rho(A) = 1$, *entonces,* $A^n \xrightarrow[n\to\infty]{} L = xy^t$ *donde* $x, y > 0$ *son vectores tales que* $\langle x, y\rangle = 1$, $Ax = x$, *y* $A^t y = y$.

Antes de demostrar el Teorema dee Perrón, presentaremos varios resultados generales para matrices $A \geqslant 0$, su radio espectral y los autovectores correspondientes.

Proposición 4.1.2. *Si* $0 \leqslant A \leqslant B$, *entonces,* $\rho(A) \leq \rho(B)$.

Demostración. Como $0 \leqslant A \leqslant B$, entonces, para todo $n \geq 1$ se tiene que $0 \leqslant A^n \leqslant B^n$. Por lo tanto $\|A^n\|_2^{1/n} \leq \|B^n\|_2^{1/n}$ y, tomando límite, se obtiene la desigualdad buscada . □

Corolario 4.1.3. *Si* $0 < A$, *entonces* $\rho(A) > 0$.

Demostración. Si $0 < A$, existe $\varepsilon > 0$ tal que $\varepsilon I \leqslant A$. Así $\rho(A) \geq \rho(\varepsilon I) = \varepsilon > 0$. □

Corolario 4.1.4. *Sean* $A \geqslant 0$, $J \subseteq \{1, \dots, n\}$ *y* $A_J = (a_{ij})_{i,j \in J}$. *Entonces* $\rho(A_J) \leq \rho(A)$.

Demostración. Basta extender A_J a $\mathcal{M}_n(\mathbb{C})$ poniendo ceros en las entradas que le faltan, y aplicar la Proposición 4.1.2. □

Observación 4.1.5. Recordemos que, dada $A \in \mathcal{M}_n(\mathbb{C})$,

$$\|\|A\|\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_{1\le i\le n} \|F_i(A)\|_1$$
$$\|\|A\|\|_1 = \max_{\|x\|_1=1} \|Ax\|_1 = \max_{1\le i\le n} \|C_i(A)\|_1 = \|\|A^t\|\|_\infty .$$

Si $A \geqslant 0$, entonces $\|F_i(A)\|_1 = \operatorname{tr}(F_i(A))$ y $\|C_i(A)\|_1 = \operatorname{tr}(C_i(A))$. △

A continuación vienen tres Lemas que sirven para ubicar el radio espectral de una matriz $A \geqslant 0$ usando la Observación anterior:

Lema 4.1.6. *Sea $A \geqslant 0$. Si $\operatorname{tr}(F_i(A))$ es constante, entonces, $\rho(A) = \|\|A\|\|_\infty$.*

Demostración. La desigualdad $\rho(A) \le \|\|A\|\|_\infty$ vale siempre. En nuestro caso, como $A\mathbf{e} = \|\|A\|\|_\infty \mathbf{e}$, se tiene que $\|\|A\|\|_\infty \le \rho(A)$. □

Lema 4.1.7. *Sean $A \geqslant 0$,*

$$\alpha = \max_{1\le i\le n} \operatorname{tr}(F_i(A)) = \|\|A\|\|_\infty \quad y \quad \beta = \min_{1\le i\le n} \operatorname{tr}(F_i(A)).$$

Entonces $\beta \le \rho(A) \le \alpha$.

Demostración. La desigualdad $\rho(A) \le \alpha$ es conocida, por lo tanto sólo verificaremos que $\beta \le \rho(A)$. Podemos suponer que $\beta > 0$. Definamos entonces la matriz $B \in \mathcal{M}_n(\mathbb{C})$ cuyas filas son:

$$F_i(B) = \frac{\beta}{\operatorname{tr}(F_i(A))} F_i(A).$$

De este modo, $\operatorname{tr}(F_i(B)) = \beta$ para todo $i \ge 1$. En consecuencia, usando el Lema 4.1.6 y la Proposición 4.1.2, $\beta = \rho(B) \le \rho(A)$, ya que, por su construcción, $0 \leqslant B \leqslant A$. □

Lema 4.1.8. *Sea $A \geqslant 0$, $x > 0$ e $y = Ax$. Notemos*

$$\beta = \min_{1\le i\le n} \frac{y_i}{x_i} \quad y \quad \alpha = \max_{1\le i\le n} \frac{y_i}{x_i} .$$

Entonces $\beta \le \rho(A) \le \alpha$.

Demostración. Sea $D = \operatorname{diag}(x)$. Entonces, por cuentas elementales, obtenemos que $D^{-1}AD = (x_i^{-1}x_j\ a_{ij})_{ij}$. Además $D^{-1}AD \geqslant 0$ y $\rho(D^{-1}AD) = \rho(A)$. Por otro lado se tiene que

$$\operatorname{tr}(F_i(D^{-1}AD)) = x_i^{-1} \sum_{j=1}^{n} a_{ij}x_j = \frac{(Ax)_i}{x_i} = \frac{y_i}{x_i} .$$

Luego basta aplicar el Lema 4.1.7. □

Teorema 4.1.9. *Sean $A \geqslant 0$ y $x > 0$.*

1. *Si existen $\alpha, \beta \in \mathbb{R}$ tales que $\beta x \leqslant Ax \leqslant \alpha x$, entonces $\beta \le \rho(A) \le \alpha$.*

2. Si x es un autovector de A (y es $x > 0$), entonces $Ax = \rho(A)x$.

Demostración. La primera parte se deduce inmediatamente del Lema 4.1.8. Supongamos que $Ax = \lambda x$. Como $Ax \geqslant 0$, debe cumplirse que $\lambda \geq 0$, en particular $\lambda \in \mathbb{R}$. Luego se verifican las hipótesis de la primera parte con $\alpha = \beta = \lambda$. $\square$

Observaciones: 4.1.10.

1. Las desigualdades anteriores (para α y para β) son independientes.
2. Además, si $x > 0$, $\beta x < Ax$ implica $\beta < \rho(A)$ y $Ax < \alpha x$ implica $\rho(A) < \alpha$. En efecto, si en los Lemas 4.1.7 y 4.1.8 se toma β estrictamente menor que los mínimos correspondientes, se obtiene $\beta < \rho(A)$.

A partir de este momento $A > 0$. Notar que esto implica que si $0 \leqslant x$ y $x \neq 0$, entonces $Ax > 0$. Este hecho se usará reiteradas veces.

Proposición 4.1.11. *Sea $A > 0$ y $\lambda \in \sigma(A)$ tal que $|\lambda| = \rho(A)$. Si $y \neq 0$ cumple que $Ay = \lambda y$, entonces:*

1. $|y| > 0$
2. $A|y| = \rho(A)|y|$

Demostración. Sea $x = |y|$, entonces por la desigualdad triangular,

$$\rho(A)x = |\lambda|x = |\lambda y| = |Ay| \leqslant A|y| = Ax.$$

Sea $z = Ax - \rho(A)x \geqslant 0$. Supongamos que $z \neq 0$. Entonces $Az > 0$. Si llamamos $u = Ax > 0$, $Az = Au - \rho(A)u > 0$. Por lo tanto, $Au > \rho(A)u$ y, por la Observación 4.1.10, se obtiene la contradicción $\rho(A) > \rho(A)$. Dado que esta provino de suponer que $z \neq 0$, se tiene que $z = 0$ y por ende $Ax = \rho(A)x$. Notar que esto implica que $|y| = x > 0$. $\square$

Corolario 4.1.12. *Si $A > 0$, entonces, $\rho(A) \in \sigma(A)$ y existe $x > 0$ tal que $Ax = \rho(A)x$.*

Proposición 4.1.13. *Sean $A > 0$ y $\lambda \in \sigma(A)$ tales que $|\lambda| = \rho(A)$. Si $y \neq 0$ cumple que $Ay = \lambda y$, entonces, existe $\theta \in [0, 2\pi)$ tal que $|y| = e^{i\theta} y$.*

Demostración. Por la Proposición 4.1.11 $A|y| = \rho(A)|y|$. Además $|Ay| = |\lambda y| = \rho(A)|y|$. En consecuencia, $|Ay| = A|y|$. Luego, existe $\theta \in [0, 2\pi)$ tal que $y_j = e^{i\theta}|y_j|$ para todo j, porque vale la igualdad en la desigualdad triangular (basta mirar las primeras cordenadas de $|Ay|$ y $A|y|$). $\square$

Corolario 4.1.14. *Si $A > 0$, entonces $\rho(A)$ es el único autovalor de módulo máximo.*

Corolario 4.1.15. *Sea $A > 0$. Entonces* $\dim Nu(A - \rho(A)I) = 1$.

Demostración. Sean $x, y \in Nu(A - \rho(A)I)$. Probaremos que son linealmente dependientes. Por la Proposición 4.1.13 se puede suponer que $x > 0$ e $y > 0$. Sea $\beta = \min\limits_i x_i/y_i$, y definamos $z = x - \beta y$. Como $x_i - \beta y_i \geq x_i - (x_i/y_i)y_i = 0$, se tiene que $z \geq 0$.

Dado que $Az = \rho(A)z$, si $z \neq 0$, entonces, $z > 0$. Pero, si $\beta = x_k/y_k$, entonces, la coordenada k-ésima de z es cero. El absurdo, proviene de suponer que $z \neq 0$, por lo tanto, $z = 0$ y $x = \beta y$. $\square$

Corolario 4.1.16. *Sea $A > 0$ y $x \in R^n$ tal que $x \geqslant 0$, $x \neq 0$ y $Ax = \lambda x$. Entonces, $\lambda = \rho(A)$ y $x > 0$.*

Demostración. Como $Ax > 0$, y por ende $x > 0$, se puede aplicar el Teorema 4.1.9. □

Teorema 4.1.17. *Sea $A > 0$ tal que $\rho(A) = 1$. Sean $x, y \in R^n$ tales que $\langle x, y\rangle = 1$, $x, y > 0$, $Ax = x$ y $A^t y = y$. Entonces, $A^m \xrightarrow[n\to\infty]{} xy^t$.*

Demostración. Sea $L = xy^t = (x_i y_j)_{ij}$

1. $L^2 = L^m = L$: En efecto, $L^2 = xy^t xy^t = x\langle x,\, y\rangle y^t = xy^t = L$.

2. $AL = LA = L$: $Axy^t = xy^t = xy^t A$.

3. $(A - L)^m = A^m - L$: Razonemos por inducción sobre m. Para $m = 1$ trivial.
$$\begin{aligned}(A-L)^{k+1} &= (A-L)(A-L)^k = (A-L)(A^k - L)\\ &= A^{k+1} - AL - LA^k + L = A^{k+1} - L - L + L\\ &= A^{k+1} - L.\end{aligned}$$

4. $\sigma(A - L) \setminus \{0\} \subseteq \sigma(A) - \{1\}$ (En particular, $\rho(A - L) < 1$.): Sea $\lambda \in \mathbb{C}$ $\lambda \neq 0$ y $z \in \mathbb{C}^n$ $z \neq 0$ tales que $(A - L)z = \lambda z$. Como $L(L - A) = 0$, se tiene que
$$Lz = \frac{1}{\lambda}L(\lambda z) = \frac{1}{\lambda}L(L-A)z = 0.$$
Luego, $Az = \lambda z$ y por lo tanto $\lambda \in \sigma(A)$. Si $\lambda = 1$, entonces, $z = \mu x$ y por lo tanto $(A - L)x = x$. Pero $Ax = x$ y $Lx = xy^t x = x$, en consecuencia, $(A - L)x = 0 = x$, absurdo.

5. Como el único $\lambda \in \sigma(A)$ con $|\lambda| = 1$ es $\lambda = 1$, se tiene que $\rho(A - L) < 1$ y por ende $A^m - L = (A - L)^m \to 0$ cuando $m \to \infty$.

□

Final de la demostración del Teorema de Perrón. Sólo falta verificar que $\rho(A)$ es raíz simple del polinomio característico de A. Supongamos, sin perdida de generalidad, que $\rho(A) = 1$. En cierta base de $\mathbb{C}^n$, es decir, para cierta $S \in \mathcal{G}l(n)$,
$$SAS^{-1} = \begin{pmatrix} J_1 & 0 & 0\\ \vdots & \ddots & \vdots\\ 0 & 0 & J_r\end{pmatrix},$$
donde J_1 es el bloque de Jordan asociado al autovalor 1 de A. Entonces $J_1 = I_k + N_k$, donde I_k es la matriz identidad de $\mathcal{M}_k(\mathbb{C})$ y $N_k \in \mathcal{M}_k(\mathbb{C})$ es cierta matriz extrictamente triangular superior. Luego
$$J_1^m = \begin{pmatrix} 1 & * & \cdots & * & *\\ 0 & 1 & \cdots & * & *\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & *\\ 0 & 0 & \cdots & 0 & 1\end{pmatrix} \xrightarrow[m\to\infty]{} SLS^{-1},$$
pero SLS^{-1} tiene rango 1. En consecuencia, $J_1 = 1$. □

4.2 Matrices de entradas no negativas.

El Teorema de Perrón falla en general si $A \geqslant 0$ pero $A \not> 0$. Por ejemplo, si

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

entonces $A^m = A$ o I, según m sea impar o par. Además, $\sigma(A) = \{1, -1\}$. En este caso el autovector asociado al 1 es positivo estricto (es $\mathbf{e}$). Pero eso no pasa para la matriz $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Es más, todas las partes del Teorema pueden hacerse fallar tomando matrices diagonales de bloques adecuadas (Ejercicio). Sin embargo, hay una extensa clase de matrices no negativas en las que el Teorema sigue valiendo:

Definición 4.2.1. Sea $A \in \mathcal{M}_n(\mathbb{C})$ tal que $A \geqslant 0$. Diremos que A es una matriz irreducible si existe un $m \geq 1$ tal que $A^m > 0$. △

Teorema 4.2.2. *Sea $A \geqslant 0$ una matriz irreducible. Entonces valen:*

1. *$\rho(A) > 0$ y $\rho(A) \in \sigma(A)$.*
2. *Existe $x \in \mathbb{R}^n$ tal que $x > 0$ y $Ax = \rho(A)x$.*
3. *Si $y \geqslant 0$, $y \neq 0$ y $Ay = \lambda y$, entonces, $\lambda = \rho(A)$ e $y > 0$.*
4. *$\rho(A)$ es raíz simple del polinomio característico de A.*
5. *Si $\lambda \in \sigma(A)$ y $\lambda \neq \rho(A)$, entonces, $|\lambda| < \rho(A)$.*
6. *Si $\rho(A) = 1$, entonces, $A^m \to L = xy^t$ donde $x, y > 0$ son vectores tales que $\langle x, y\rangle = 1$, $Ax = x$ y $A^t y = y$.*

Demostración. Sea $m \in \mathbb{N}$ tal que $A^m > 0$. Por el Lema 1.7.5,

$$\sigma(A^m) = \{\lambda^m : \lambda \in \sigma(A)\}.$$

Por el Teorema 4.1.1 aplicado a A^m, concluimos que $\rho(A) = \rho(A^m)^{1/m} > 0$. Sea $\lambda \in \sigma(A)$ tal que $|\lambda| = \rho(A)$ y $x \in \mathbb{C}^n$ tal que $Ax = \lambda$. Entonces, por lo anterior, $\lambda^m = \rho(A^m)$ y $A^m x = \rho(A^m)x$. De ahí podemos deducir que algún múltiplo y de x cumple que $y > 0$, y por ello $\lambda = \rho(A)$ y $Ay = \rho(A)y$.

Además, cada $\lambda^m \in \sigma(A^m)$ posee una multiplicidad en el polinomio característico de A^m mayor o igual que la de λ en el de A (esto se ve fácil triangulando con el Teorema 1.2.4). Por lo tanto $\rho(A)$ posee multiplicidad algebraica uno como autavalor de A. Razonamientos similares permiten concluir que $\rho(A)$ es el único autovalor de módulo máximo, y también la condición 3.

Sólo falta probar la condición 5: Supongamos sin pérdida de generalidad que $\rho(A) = 1$. Dado que en la demostración del Teorema de Perrón no se usó que A fuera estrictamente positiva para obtenerlas, siguen valiendo las sigientes relaciones:

$$(A - L)^m = A^m - L \tag{12}$$

$$\sigma(A - L) \setminus \{0\} \subset \sigma(A) - \{1\} \tag{13}$$

De (13) se deduce que $\rho(A - L) < 1$, por lo tanto $(A - L)^m \xrightarrow[m\to\infty]{} 0$, mientras que usando (12) se deduce que $A^m \to L$. □

Referencias

[1] R. Bhatia; *Matrix Analysis*, Springer, Estados Unidos, 1997.

[2] A. Benedek y R. Panzone; *La matriz positiva y su espectro*, Informe Técnico interno No.86, INMABB, Bahía Blanca, 2003.

[3] M. C. González; *Relaciones de Mayorización para el Producto de Hadamard*, Tesis de licenciatura, Depto. Mat. FCEA-UNC, Neuquén, 2003.

[4] R. Horn- C. Johnson; *Matrix Analysis*, Cambridge University Press, Estados Unidos, 1985.

[5] R. Horn- C. Johnson; *Topics in Matrix Analysis*, Cambridge University Press, Estados Unidos, 1991.