Mayorización

Jorge Antezana y Demetrio Stojanoff

26 de julio de 2006

Índice

1. Mayorización. **2**

1.1. Notaciones. 2
1.2. Definiciones y propiedades básicas 3
1.3. Mayorización y funciones convexas 7
1.4. Birkhoff, Hall y los casamientos 9

2. Aplicaciones al Análisis Matricial. **11**

2.1. Generalidades 11
2.2. Autoadjuntas 14
2.3. Descomposición polar y valores singulares 16
2.4. Teorema de Schur y aplicaciones 18
2.5. Partes reales 20
2.6. Teorema de Schur-Horn 21
2.7. Normas unitariamente invariantes. 24
2.8. Mayorización de matrices Hermitianas 27
2.9. Teoremas de Lindskii y sus aplicaciones. 30

1. Mayorización.

1.1. Notaciones.

A lo largo de este trabajo consideraremos a $\mathbb{C}$ o $\mathbb{R}$ como cuerpo de escalares. Llamaremos $\mathbb{R}_+$ (resp. $\mathbb{R}_+^*$) al conjunto de números reales no negativoas (resp. positivos). Dado $n \in \mathbb{N}$, usaremos el símbolo $\mathbb{I}_n$ para denotar al conjunrto $\{1, 2, \ldots, n\} \subseteq \mathbb{N}$. 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, $\mathcal{M}_n(\mathbb{R}) = \mathbb{R}^{n\times n}$. Llamaremos $\mathcal{M}_{n,m}(\mathbb{C}) = \mathbb{C}^{n\times m}$, al espacio de matrices rectangulares. Para denotar las entradas de una matriz $A \in \mathcal{M}_{n,m}(\mathbb{C})$, usaremos indistintamente, por conveniencia del contexto, las notaciones $A = (A_{ij})_{i,j\in\mathbb{I}_n}$ o $A = (a_{ij})_{i,j\in\mathbb{I}_n}$. Las columns y filas de A se pueden pensar como vectores de $\mathbb{C}^n$. Se usará la notación $C_i(A)$ (respectivamente $F_i(A)$) para denotar a la i-ésima columna (resp. fila) de A.

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 $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$), para ahorrar espacio. Por ejemplo, si $A \in \mathcal{M}_n(\mathbb{C})$ e $i \in \mathbb{I}_n$, entonces $C_i(A) = (a_{1i}, a_{2i}, \ldots, a_{ni}) \in \mathbb{C}^n$.

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 $\mathcal{S}$.

1.2. 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

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

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 $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, escribiremos

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

Con estas notaciones podemos dar la definición de mayorización:

Definición 1.2.1. 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

$$\operatorname{tr} y = \operatorname{tr} x \quad \text{y} \quad \sum_{j=1}^k x_j^\downarrow \ \leq \ \sum_{j=1}^k y_j^\downarrow, \quad 1 \leq k < n, \tag{1}$$

▲

Observación 1.2.2. Dado que $\sum\limits_{j=1}^k x_j^\uparrow = \operatorname{tr} x - \sum\limits_{j=1}^{n-k} x_j^\downarrow$, la relación $x \prec y$ también es equivalente a las condiciones:

$$\operatorname{tr} y = \operatorname{tr} x \quad \text{y} \quad \sum_{j=1}^k x_j^\uparrow \ \geq \ \sum_{j=1}^k y_j^\uparrow, \quad 1 \leq k < n, \tag{2}$$

Si sólo se cumple la condición (1), se dice que x es **submayorizado** por y y se nota $x \prec_w y$. Por el contrario, si sólo se cumple la condición (2), se dice que x es **supramayorizado** por y y se nota $x \prec^w y$. ▲

Ejemplos 1.2.3. 1. Sea $x \in \mathbb{R}^n$ tal que $x_i \geq 0$, $i \in \mathbb{I}_n$. Llamemos $a = \operatorname{tr} x$. Entonces

$$\left(\frac{a}{n}\ ,\ \frac{a}{n}\ , \ldots ,\ \frac{a}{n}\right) \prec (x_1, x_2, \ldots, x_n) \prec (a, 0, \ldots, 0). \tag{3}$$

Una manera de verificarlo es suponer que existe $k < n$ tal que $\sum\limits_{i=1}^k x_i^\downarrow < \frac{ka}{n}$. En tal caso, debe suceder que $x_k^\downarrow < \frac{a}{n}$ y, por lo tanto, también $\sum\limits_{i=k+1}^n x_i^\downarrow < \frac{(n-k)a}{n}$. Pero esto contradice el hecho de que $\sum\limits_{j=1}^n \frac{a}{n} = a = \sum\limits_{i=1}^n x_i$. La otra mayorización es obvia.
2. Si $x \prec y$ e $y \prec x$, entonces $x^\downarrow = y^\downarrow$ y, por lo tanto, x e y difieren en una permutación. ▲

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 1.2.4. Una matriz $A \in \mathcal{M}_n(\mathbb{R})$ se denomina **doblemente estocástica** si $A \geqslant 0$ y, para todo $i \in \mathbb{I}_n$, 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 $\mathcal{DS}\,(n)$. ▲

Ejercicio 1.2.5. 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 $e = (1, 1, \dots, 1) \in \mathbb{R}^n$. Deducir que $\mathcal{DS}\,(n)$ es un conjunto convexo.

Teorema 1.2.6. *$A \in \mathcal{DS}\,(n)$ 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 $\mathcal{E} = \{e_1, \dots, e_n\}$ la base canónica de $\mathbb{C}^n$. Entonces, como $Ae_i \prec e_i$, $i \in \mathbb{I}_n$, podemos deducir que $A \geqslant 0$ y

$$1 = \operatorname{tr} e_i = \operatorname{tr} Ae_i = \operatorname{tr} C_i(A) \qquad \text{para todo} \qquad i \in \mathbb{I}_n\ .$$

Por otra parte, tomemos $e = (1, 1, \dots, 1)$. Usando el Ejemplo 1.2.3, como $Ae \prec e$ y todas las cordenadas de e son iguales, deducimos que

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

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 (ver Observación 1.2.8), entonces $QAP \in \mathcal{DS}\,(n)$ (esto se formaliza fácil usando el Ejercicio 1.2.5). Además, como $y = Ax$,

$$\sum_{j=1}^{k} y_j = \sum_{j=1}^{k} \sum_{i=1}^{n} a_{ji} x_i = \sum_{i=1}^{n} \Big(\sum_{j=1}^{k} a_{ji} \Big) x_i \qquad \text{para todo} \qquad k \in \mathbb{I}_n\ . \tag{4}$$

Si llamamos $t_i = \sum\limits_{j=1}^{k} a_{ji}$, para $i \in \mathbb{I}_n$, entonces

$$0 \le t_i \le \operatorname{tr} C_i(A) = 1 \qquad \text{y} \qquad \sum_{i=1}^{n} t_i = \sum_{j=1}^{k} \operatorname{tr} F_i(A) = k\ .$$

Luego, aplicando la ecuación (4),

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

pues los dos términos del último renglón son sumas de términos no positivos. Por lo tanto $\sum\limits_{j=1}^{k} y_j \leq \sum\limits_{j=1}^{k} x_j$ para todo $k \in \mathbb{I}_n$ y además, cuando $k = n$, obtenemos la igualdad, usando la ecuación (4) (para $k = n$, los $t_i = 1$). Así, $y \prec x$. ■

Ejemplo 1.2.7. Como motivación del siguiente resultado, supongamos que

$$(y_1, y_2) \prec (x_1, x_2), \ y_1 \geq y_2 \ \text{ y } \ x_1 \geq x_2.$$

Entonces $x_2 \leq y_2 \leq y_1 \leq x_1$ y $y_1 + y_2 = x_1 + x_2$. 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 $(y_1, y_2) = \lambda(x_1, x_2) + (1-\lambda)(x_2, x_1)$. ▲

Observación 1.2.8. Sea $n \in \mathbb{N}$. Recordar que $\mathbb{I}_n = \{1, 2, \ldots, n\}$.

1. Llamaremos S_n al n-ésimo grupo simétrico, es decir
$$S_n = \{\sigma : \mathbb{I}_n \to \mathbb{I}_n \ : \ \sigma \ \text{ es biyectiva } \} ,$$
con el producto dado por la composición de funciones.
2. Dados $\sigma \in S_n$ y $x \in \mathbb{C}^n$, llamaremos $x_\sigma = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$.
3. Dada $\sigma \in S_n$, llamaremos $P_\sigma \in \mathcal{M}_n(\mathbb{C})$ al operador dado por
$$P_\sigma \ x \ = \ x_\sigma \ , \quad x \in \mathbb{C}^n \ .$$
Observar que, si $\tau \in S_n$, entonces $P_\sigma P_\tau = P_{\sigma\tau}$. Por otra parte, P_σ opera en la base canónica por $P_\sigma(e_k) = e_{\sigma^{-1}(k)}$, $k \in \mathbb{I}_n$.
4. El grupo $\mathcal{U}_\mathcal{P}(n) = \{P_\sigma : \sigma \in S_n\}$ está incluido en $\mathcal{U}(n)$, dado que cada P_σ es claramente isométrico. Por lo tanto, para cada $\sigma \in S_n$, $P_{\sigma^{-1}} = P_\sigma^{-1} = P_\sigma^* = P_\sigma^T$.
5. Observar que $\mathcal{U}_\mathcal{P}(n) \subseteq \mathcal{DS}\,(n)$. En efecto, dada $\sigma \in S_n$,
$$C_k(P_\sigma) = P_\sigma(e_k) = e_{\sigma^{-1}(k)} \quad \text{ y } \quad F_k(P_\sigma) = C_k(P_\sigma^T) = C_k(P_{\sigma^{-1}}) = e_{\sigma(k)} \ , \tag{5}$$
para todo $k \in \mathbb{I}_n$. Más adelante veremos que $\mathcal{U}_\mathcal{P}(n)$ es ni más ni menos que el conjuntos de puntos extremales de $\mathcal{DS}\,(n)$. ▲

Teorema 1.2.9. *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. Como $\mathcal{DS}\,(n)$ es convexo y $\mathcal{U}_\mathcal{P}(n) \subseteq \mathcal{DS}\,(n)$, obtenemos que $2 \Rightarrow 3$. Por el Teorema 1.2.6 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 1.2.7. 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 matriz de permutación que verifica

$$P_0(x_1, \ldots, x_n) = (x_k, x_2, \ldots, x_{k-1}, x_1, x_{k+1}, \ldots, x_n).$$

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 $\sigma_i \in S_n$ a la extensión de la permutación σ_i a $\mathbb{I}_n$, poniendo $\sigma_i(1) = 1$. Luego, como $z_1 = y_1$, se tiene $y = \sum_{i=1}^{s} \mu_i P_{\sigma_i}(z)$. 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. ■

Observación 1.2.10. Un error típico al tratar de demostrar mayorización entre dos vectores, es olvidarse de ordenarlos 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. ▲

Corolario 1.2.11. *Sean $w, z \in \mathbb{R}^m$ tales que $w \prec z$, y sean $x, y \in \mathbb{R}^k$ tales que $x \prec y$. Entonces los vectores $(x, w), (y, z) \in \mathbb{R}^{k+m}$ cumplen que $(x, w) \prec (y, z)$.*

Demostración. Por el Teorema 1.2.9, existen $A \in \mathcal{DS}(k)$ y $B \in \mathcal{DS}(m)$ tales que $Ay = x$ y $Bz = w$. Pero si consideramos

$$C = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \in \mathcal{M}_{k+m}(\mathbb{C}),$$

es fácil ver que $C \in \mathcal{DS}(k+m)$ y que $C(y, z) = (x, w)$. ■

Ejercicio 1.2.12. Sean $x, u \in \mathbb{R}^n$ tales que $x \leqslant u$, i.e., $x_i \leq u_i$, para todo $i \in \mathbb{I}_n$.

1. Probar que también se cumple que $x^\downarrow \leqslant u^\downarrow$.
2. Deducir que $x \prec_w u$. ▲

Proposición 1.2.13. *Sean $x, y \in \mathbb{R}^n$. Entonces $x \prec_w y$ si y sólo si existe $u \in \mathbb{R}^n$ tal que*

$$x \leqslant u \prec y \ .$$

Demostración. Antes que nada, es claro que si el tal u existe, entonces $x \prec_w y$ (por el Ejercicio 1.2.12 y la transitividad de $\prec_w$). Para probar la recíproca, podemos asumir que $\operatorname{tr} x < \operatorname{tr} y$, porque sinó basta tomar $u = x$. Asumiremos también que x e y están ordenados en forma decreciente, dado que una ves encontrado el u para este caso, luego se lo puede reordenar igual que a x, lo que preserva la relación $x \leqslant u$ y no afecta la relación $u \prec y$.

Se hará inducción sobre n. Si $n = 1$, el resultado es trivial (en ese caso $\prec$ significa igualdad). Si $n > 1$, cosideramos dos casos:

Caso 1: Si existe $k \in \mathbb{I}_{n-1}$ tal que $\sum\limits_{i=1}^{k} x_i = \sum\limits_{i=1}^{k} y_i$, llamamos $a = (x_1, \dots, x_k)$ y $b = (y_1, \dots, y_k) \in \mathbb{R}^k$. Es claro que $a \prec b$. Por otra parte, si llamamos $w = (x_{k+1}, \dots, x_n)$ y $z = (y_{k+1}, \dots, y_n) \in \mathbb{R}^{n-k}$, es también claro que $w \prec_w z$, porque están bien ordenados y, si $r \in \mathbb{I}_{n-k}$, entonces

$$\sum_{i=1}^{r} z_i - \sum_{i=1}^{r} w_i = \sum_{i=1}^{k+r} y_i - \sum_{i=1}^{k+r} x_i \geq 0.$$

Ahora bien, por hipótesis inductiva, debe existir $v \in \mathbb{R}^{n-k}$ tal que $w \leqslant v \prec z$. Entonces basta tomar $u = (a, v) \in \mathbb{R}^n$ que cumple lo pedido, porque $w \leqslant v$ implica que $x = (a, w) \leqslant (a, v) = u$. Por otra parte, como $a \prec b$ y $v \prec z$, el Corolario 1.2.11 dice que $u = (a, v) \prec (b, z) = y$.

Caso 2: Supongamos que $d = \min\limits_{k \in \mathbb{I}_n} \left\{ \sum\limits_{i=1}^{k} y_i - \sum\limits_{i=1}^{k} x_i \right\} > 0$, y que se realiza en cierto $k_0 \in \mathbb{I}_n$. Tomemos $v = x + d e_1$, es decir que agrandamos en d la primera cordenada de x. Observar que v está ordenado decrecientemente, por estarlo x. Por ser d quien es, es claro que $x \leqslant v$ y que $v \prec_w y$. Pero claramente v cae en el Caso 1 (sumando hasta k_0 , y si k_0 era n, bingo). Entonces existe $u \in \mathbb{R}^n$ tal que $x \leqslant v \leqslant u \prec y$. ■

Ejercicio 1.2.14. Probar que, dados $x, y \in \mathbb{R}^n$, entoncecs $x \prec_w y$ si y sólo si existe $v \in \mathbb{R}^n$ tal que $x \prec v \leqslant y$.

1.3. Mayorización y funciones convexas

Dado $\mathbb{I} \subseteq R$, una función $f : \mathbb{I} \to \mathbb{R}$, y un vector $x \in \mathbb{I}^n$, notaremos por

$$f(x) = (f(x_1), \dots, f(x_n)) \in \mathbb{R}^n.$$

Por otra parte, cuando $\mathbb{I}$ es un intervalo, decimos que f es *convexa* si, dados $a, b \in \mathbb{I}$ y $\lambda \in [0, 1]$, se cumple que

$$f\big(\lambda a + (1-\lambda) b\big) \leq \lambda f(a) + (1-\lambda) f(b).$$

La función f se dice *cóncava* si $-f$ es convexa.

Teorema 1.3.1. *Sean $x, y \in \mathbb{R}^n$. Sea $\mathbb{I} \subseteq \mathbb{R}$ un intervalo (semirrecta o todo $\mathbb{R}$) tal que $x, y \in \mathbb{I}^n$. Entonces, son equivalentes:*

1. $y \prec x$

2. $\operatorname{tr} f(y) \le \operatorname{tr} f(x)$ *para toda función convexa* $f : \mathbb{I} \to \mathbb{R}$.

3. $\displaystyle\sum_{i=1}^n |y_i - t| \le \sum_{i=1}^n |x_i - t|$ *para todo* $t \in \mathbb{R}$.

Análogamente, son equivalentes

1'. $y \prec_w x$ *(submayorización)*

2'. $\operatorname{tr} f(y) \le \operatorname{tr} f(x)$ *para toda función* $f : \mathbb{R} \to \mathbb{R}$ *convexa y* **no decreciente**.

3'. $\displaystyle\sum_{i=1}^n (y_i - t)^+ \le \sum_{i=1}^n (x_i - t)^+$ *para todo* $t \in \mathbb{R}$.

Demostración. Sólo probaremos la primer parte, puesto que los argumentos para probar la segunda son similares (para 1' $\to$ 2', se aplica 1 $\to$ 2 y la Proposición 1.2.13, que será útil para funciones no decrecientes). Supongamos que $y \prec x$. Entonces, por el Teorema 1.2.9, $y = \displaystyle\sum_{i=1}^s \lambda_i P_{\sigma_i}(x)$ para ciertos $\lambda_i \ge 0$ que suman uno, y para ciertas $\sigma_i \in S_n$. Luego $f\big(\sum_{i=1}^s \lambda_i P_{\sigma_i} x\big) \leqslant \sum_{i=1}^s \lambda_i f\big(P_{\sigma_i} x\big)$ (en cada cordenada), y

$$\operatorname{tr} f(y) = \operatorname{tr} f\big(\sum_{i=1}^s \lambda_i P_{\sigma_i} x\big) \le \operatorname{tr} \sum_{i=1}^s \lambda_i f\big(P_{\sigma_i} x\big) = \sum_{i=1}^s \lambda_i \operatorname{tr} P_{\sigma_i}\big(f(x)\big) = \operatorname{tr} f(x) \ .$$

La implicación $2 \to 3$ (respectivamente, $2' \to 3'$) se deduce de que la función $x \mapsto |x - t|$ (resp. $x \mapsto (x-t)^+$) es convexa (resp. convexa no decreciente) para todo $t \in \mathbb{R}$. Probemos $3 \to 1$. Supongamos que los vectores x e y están ordenados de forma decreciente (ni 3 ni 1 depende del orden de las cordenadas). Sean $M = \max\{x_1, y_1\}$ y $m = \min\{x_n, y_n\}$. Tomando $t > M$, se tiene que

$$\sum_{i=1}^n |y_i - t| = \sum_{i=1}^n t - y_i = kt - \sum_{i=1}^n y_i \ ,$$

y lo mismo para x. Luego la desigualdad 3 para estos valores de t implica que $\operatorname{tr} x \le \operatorname{tr} y$. Análogamente, la desigualdad 3 para valores de t tales que $t < m$ implica que $\operatorname{tr} y \le \operatorname{tr} x$. Luego $\operatorname{tr} x = \operatorname{tr} y$. Por otro lado, dado $x \in \mathbb{R}$, se tiene que $2x^+ = x + |x|$. Luego

$$\begin{aligned} 2\sum_{i=1}^n (y_i - t)^+ &= \sum_{i=1}^n (y_i - t) + \sum_{i=1}^n |y_i - t| = \operatorname{tr} y - nt + \sum_{i=1}^n |y_i - t| \\ &\le \operatorname{tr} x - nt + \sum_{i=1}^n |y_i - t| = \sum_{i=1}^n (x_i - t) + \sum_{i=1}^n |x_i - t| \\ &= 2\sum_{i=1}^n (x_i - t)^+ . \end{aligned}$$

Deducimos que basta probar 3' $\to$ 1'. Fijemos $k \in \mathbb{I}_n$. Tomando $t = x_k$, resulta que

$$\sum_{i=1}^n (x_i - t)^+ = \sum_{i=1}^k (x_i - t)^+ = \sum_{i=1}^k x_i - kt \ .$$

Por lo tanto

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

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

Corolario 1.3.2. *Sea $\mathbb{I} \subseteq \mathbb{R}$ un intervalo. Sea $g : \mathbb{I} \to \mathbb{R}$ una función convexa (resp. convexa creciente). Entonces, dados $x, y \in \mathbb{I}^n$,*

$$x \prec y \quad (\textit{resp. } x \prec_w y) \qquad \Longrightarrow \qquad g(x) \prec_w g(y).$$

En particualr $x \prec y \implies |x| \prec_w |y|$, y también $x \prec_w y \implies x^+ \prec_w y^+$.

Demostración. Sea $f : \mathbb{R} \to \mathbb{R}$ convexa no decreciente. Es fácil ver, entonces, que $f \circ g$ es una función convexa. Por el Teorema 1.3.1, si $x \prec y$, entonces

$$\operatorname{tr} f(g(x)) = \operatorname{tr} f \circ g\ (x) \leq \operatorname{tr} f \circ g\ (y) = \operatorname{tr} f(g(y)).$$

Pero por el mismo Teorema (en su segunda parte), como lo anterior vale para toda $f : \mathbb{R} \to \mathbb{R}$ convexa no decreciente, deducimos que $g(x) \prec_w g(y)$.

Si g es creciente y $x \prec_w y$, por el Corolario 1.2.13 existe $u \in \mathbb{R}^n$ tal que $x \leqslant u \prec y$. Luego, por el caso anterior, $g(x) \leqslant g(u) \prec_w g(y)$. Para concluir que $g(x) \prec_w g(y)$ basta aplicar el Ejercicio 1.2.12. ∎

Corolario 1.3.3. *Sean $x, y \in \mathbb{R}^n$, tales que $x > 0$ e $y > 0$. Entonces, se tiene que*

$$x \prec y \qquad \Longrightarrow \qquad \prod_{i=1}^{n} x_i \geq \prod_{i=1}^{n} y_i\ .$$

Demostración. Sea $g(t) = -\log t$, que es una función convexa (pero decreciente), definida en $\mathbb{I} = (0, +\infty)$. Por el Corolario 1.3.2, si $x \prec y$, entonces $g(x) \prec_w g(y)$. En particular,

$$-\log \prod_{i=1}^{n} x_i = -\sum_{i=1}^{n} \log x_i \leq -\sum_{i=1}^{n} \log y_i = -\log \prod_{i=1}^{n} y_i\ ,$$

de lo que se concluye que $\prod_{i=1}^{n} x_i \geq \prod_{i=1}^{n} y_i$. ∎

Corolario 1.3.4. *Sean $x, y \in \mathbb{R}^n$, tales que $x > 0$ e $y > 0$. Si se cumple que*

$$\prod_{i=1}^{k} x_i \leq \prod_{i=1}^{k} y_i\ , \qquad \textit{para todo} \quad k \in \mathbb{I}_n\ , \tag{6}$$

entonces $x \prec_w y$.

Demostración. Notar que la ecuación (6) equivale a decir que $\log(x) \prec_w \log(y)$. Como la función $t \mapsto \exp t := e^t$ es convexa y creciente, podemos aplicar el Corolario 1.3.2 y deducir que $x = \exp \log x \prec_w \exp \log y = y$. ∎

1.4. Birkhoff, Hall y los casamientos

Sean $V = \{v_1, \ldots, v_n\}$ y $M = \{m_1, \ldots, m_n\}$ dos conjuntos de n elementos. Pensaremos que V es un conjunto de varones (humanos) y M de mujeres. Dada una relación $C \subseteq V \times M$, diremos que v_i "conoce a"m_j (puede usarse tambien "tiene onda con", o "gusta de") si $(v_i, m_j) \in C$. El llamado *problema de los casamientos* (PC) consiste en encontrar condiciones sobre C que aseguren que exista $f : V \to M$ biyectiva, tal que $\mathrm{Gr}(f) \subseteq C$. Si pensamos que cada v se casa con $f(v)$, el problema se traduce a poder casar todos los varones de V con mujeres de M (sin bigamia) y que todas las parejitas sean felices (se gusten mutuamente). Para describir esas condiciones pongamos un poco de notación: Dado $J \subseteq \mathbb{I}_n$, llamaremos $V_J = \{v_j : j \in J\}$, y

$$\begin{aligned} M_J &= \{m_i \in M : (v_j, m_i) \in C \quad \text{para algún} \quad j \in J\} \\ &= \{ \text{ chicas conocidas por algún muchacho de } V_J\}. \end{aligned}$$

Observar que $M_J = \pi_M([V_J \times M] \cap C)$, donde π_M es la proyección sobre M. Como siempre, se abrevia $M_i = M_{\{i\}}$. Es evidente que si el PC tiene solución, debe cumplirse que $|M_J| \geq |J|$ para todo $J \subseteq \mathbb{I}_n$, porque $f(V_J)$ debería estar incluido en M_J. Pero mucho menos claro es que vale la recíproca:

Teorema 1.4.1 (El problema de los casamientos de Hall)**.** *El PC tiene solución para una relación* $C \subseteq V \times M$ *si y sólo si*

$$|M_J| \geq |J| \quad \textit{para todo} \quad J \subseteq \mathbb{I}_n \ . \tag{7}$$

Demostración. Probaremos la suficiencia por inducción sobre n. Todo es fácil si $n = 1$ (ese es el problema de la pareja de náufragos). Si $n > 1$, separemos dos casos:
Caso 1: Supongamos que tenemos una condición mejor que (7), a saber,

$$|M_J| \geq |J| + 1 \quad \text{para todo} \quad J \subseteq \mathbb{I}_n \ , \ \ \emptyset \neq J \neq \mathbb{I}_n \ . \tag{8}$$

En tal caso fijamos al vago v_n de M y lo casamos con una chica que conozca ($m_j \in M_n$). Veamos ahora que, si $J = \mathbb{I}_{n-1}$, los conjuntos V_J y $M \setminus \{m_j\}$ cumplen la condición (7) (para aplicarles la HI). En efecto, notar que si $I \subseteq J$, entonces la ecuación (8) asegura que $|M_I \cap M \setminus \{m_j\}| \geq |M_I| - 1 \geq |I|$. En otras palabras, dados k muchachos, entre todos deben conocer al menos k chicas todavía solteras. Por HI, tenemos una biyección entre V_J y $M \setminus \{m_j\}$ con gráfico contenido en C, que se puede extender, mandando $n \mapsto j$, a todo V sobre todo M.
Caso 2: Si existe un $J \subseteq \mathbb{I}_n$ tal que

$$\emptyset \neq J \neq \mathbb{I}_n \quad \text{y} \quad |M_J| = |J| = k < n \ , \tag{9}$$

por HI podemos definir una biyección $f_1 : V_J \to M_J$ con $\mathrm{Gr}(f_1) \subseteq C$. Por otra parte, por la igualdad (9), es fácil ver que los conjuntos que quedan, V_{J^c} y $M \setminus M_J$ cumplen también la condición (7). En efecto, si $I \subseteq J^c$ tiene $|I| = r$, observemos que $M_{I \cup J} \setminus M_J = M_I \setminus M_J$ (las que no conocen los de J deben conocerlas los de I). Pero

$$|M_{I \cup J} \setminus M_J| \geq |M_{I \cup J}| - |M_J| \geq (r + k) - k = r \ .$$

Luego $|M_I \setminus M_J| \geq r = |I|$. Otra forma de verlo es la siguiente: casamos k pibes que conocían justo k chicas. Dados r de los solteros, junto con los casados conocían al menos $k + r$ chicas, por lo que los r solteros conocían ellos a todas las solteras de este grupo (por lo menos r), porque los k novios solo conocían a las k que se casaron con ellos. Aplicamos nuevamente la HI para definir otra biyección $f_2 : V_{J^c} \to M \setminus M_J$, también con gráfico dentro de C. Pegando ambas funciones, encontramos la biyección buscada. ■

Definición 1.4.2. Sea $A \in \mathcal{M}_n(\mathbb{C})$.

1. Dada $\sigma \in S_n$, llamamos al vector $(a_{1\sigma(1)}, \ldots, a_{n\sigma(n)})$ una *diagonal* de A. Notar que las diagonales tienen exactamente un elemento de cada fila y uno de cada columna de A.

2. Decimos que *A tiene una diagonal sin ceros* si alguna de las diagonales antes definidas tiene todas sus cordenadas no nulas.

Corolario 1.4.3 (König-Frobenius)**.** *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces son equivalentes:*

1. *Toda diagonal de A tiene ceros.*

2. *Existe subconjuntos $I, J \subseteq \mathbb{I}_n$ tales que $|I| + |J| > n$ y la submatriz $A_{IJ} \equiv 0$, es decir que $a_{ij} = 0$ para todo par $(i, j) \in I \times J$.*

Demostración. Coonsideremos los conjuntos $M = V = \mathbb{I}_n$ y la relación

$$C = \{(i, j) \in \mathbb{I}_n \times \mathbb{I}_n : a_{ij} \neq 0\}.$$

Es claro que A tiene alguna diagonal sin ceros si y sólo si el PC tiene solución para la relación C. Que A no tenga ninguna diagonal sin ceros equivale, por el Teorema 1.4.1, a que exista $I \subseteq \mathbb{I}_n$ tal que $|M_I| < |I| = k$. Observar que

$$K := \mathbb{I}_n \setminus M_I = \{j \in \mathbb{I}_n : a_{ij} = 0 \quad \text{para todo} \quad i \in I\}$$

es el mayor de los conjuntos J de índices tales que $A_{IJ} \equiv 0$. Además, si $|K| = r$, entonces $k + r > n$ si y sólo si $n - r = |M_I| < k$. Y esto concluye la prueba. ■

Corolario 1.4.4. *Si $A \in \mathcal{DS}(n)$, entonces A debe tener alguna diagonal sin ceros.*

Demostración. Supongamos que no. Reordenando filas y columnas de A (multiplicando por matrices de permutacion) podemos suponer, por el Corolario 1.4.3, que existen $k, r \in \mathbb{I}_n$ tales que $k + r > n$ y que $a_{ij} = 0$ si $i \in \mathbb{I}_k$ y $j \in \mathbb{I}_r$. En otras palabras, que existen $P, Q \in \mathcal{U}_\mathcal{P}(n)$ tales que

$$PAQ = \begin{bmatrix} 0_{k\times r} & B \\ C & D \end{bmatrix} \in \mathcal{DS}(n) \ ,$$

donde $0_{k\times r}$ es la matriz nula de $\mathcal{M}_{k,r}(\mathbb{C})$. En tal caso, las k filas de B deben tener traza uno, lo mismo que las r columnas de C. Pero entonces la suma de todas las entradas de PAQ (las de D son no negativas) debería sumar estrictamente más que n. Pero esto contradice el hecho de que $PAQ \in \mathcal{DS}(n)$. ■

Teorema 1.4.5 (Birkhoff)**.** *El conjunto de matrices doble estocásticas $\mathcal{DS}(n)$ es convexo y sus puntos extremales son el conjunto $\mathcal{U}_\mathcal{P}(n)$ de matrices de permutación. Es decir que toda $A \in \mathcal{DS}(n)$ es combinación convexa de matrices de permutación.*

Demostración. Es fácil ver que si $P \in \mathcal{U}_\mathcal{P}(n)$, entonces es extremal en $\mathcal{DS}(n)$. Luego basta ver que toda $A \in \mathcal{DS}(n)$ es combinación convexa de matrices de permutación.

Sea $A \in \mathcal{DS}(n)$. Notaremos $k(A) = \big|\{(i, j) \in \mathbb{I}_n \times \mathbb{I}_n : a_{ij} \neq 0\}\big|$. Probaremos el resultado inducción en $k(A)$. Observar que $n \leq k(A) \leq n^2$, y que $k(A) = n$ si y sólo si $A \in \mathcal{U}_\mathcal{P}(n)$, por lo que lo afirmado es trivial en este caso. Supongamos que $k(A) > n$.

Por el Corolario 1.4.4 existe $\sigma \in S_n$ tal que $a_{i\sigma(i)} > 0$, para todo $i \in \mathbb{I}_n$. Sea $P = P_\sigma \in \mathcal{U}_\mathcal{P}(n)$ la matriz asociada a la permutación σ. Por la ecuación (5), $P_{ij} \neq 0$ si y sólo si $j = \sigma(i)$. Sea $a = \min\limits_{i\in\mathbb{I}_n} a_{i\sigma(i)}$. Notar que, por el hecho de que $k(A) > n$, se debe cumplir que $0 < a < 1$. Es fácil ver, entonces, que $B = \dfrac{A - aP}{1 - a} \in \mathcal{DS}(n)$. Finalmente, se observa que $A = aP + (1 - a)B$, y se aplica la hipótesis inductiva, ya que $k(B) < k(A)$. En efecto, si $a_{ij} = 0$, entonces $P_{ij} = 0$, por lo que también $b_{ij} = 0$. Esto dice que $k(B) \leq k(A)$. Por otra parte, si $a = a_{i\sigma(i)} \neq 0$, entonces $b_{i\sigma(i)} = 0$, con lo que $k(B) < k(A)$ como se afirmó. ■

2. Aplicaciones al Análisis Matricial.

2.1. Generalidades

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. $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$, pero $\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle$.

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 2.1.1. comun]norma 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 $u, v \in \mathbb{V}$ y $\lambda \in K$,

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

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 comun]espacio!de Hilbert espacio de Hilbert (ojo, acá se asume que $\dim \mathbb{V} < \infty$, sino hay que pedir que $\mathbb{V}$ sea completo). Usualmente usaremos letras $\mathbb{H}$ o $\mathbb{K}$ para tales espacios y notaremos por $L(\mathbb{H}, \mathbb{K})$ al espacio de operadores lineales de $\mathbb{H}$ en $\mathbb{K}$ (acotados, si $\dim \mathbb{H} = \infty$). Si $\mathbb{H} = \mathbb{K}$, escribimos $L(\mathbb{H})$ en lugar de $L(\mathbb{H}, \mathbb{H})$. Si $A \in L(\mathbb{H}, \mathbb{K})$ notaremos por $\ker A$ a su comun]núcleo núcleo y $R(A)$ a su imagen. Además notaremos $\operatorname{rk}(A) = \dim R(A)$ al rango (columna) de A.

Definición 2.1.2. Sea $\mathbb{H}$ un espacio de Hilbert. Los vectores $x_1, ..., x_k \in \mathbb{H}$ forman un conjunto **ortogonal** cuando $\langle x_i, x_j\rangle = 0$, si $i \neq j$. Si además los vectores están normalizados, es decir $\|x_i\|^2 = \langle x_i, x_i\rangle = 1$ ($i \in \mathbb{I}_k$), entonces el conjunto se dice **ortonormal**. Usaremos las siglas **b.o.n.** para denotar a una base ortonormal de $\mathbb{H}$. ▲

Definición 2.1.3. Sean $\mathbb{H}$ y $\mathbb{K}$ espacios de Hilbert y sea $A \in L(\mathbb{H}, \mathbb{K})$. Se llama **adjunto** de A al único operador $A^* \in L(\mathbb{K}, \mathbb{H})$ que satisface

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

Observación 2.1.4. Si, en particular, $A \in L(\mathbb{H})$, entonces A^* también está en $L(\mathbb{H})$. Para una b.o.n. fija $\mathcal{B} = \{e_1, \dots, e_n\}$ de $\mathbb{H}$, se identifica a los operadores de $L(\mathbb{H})$ con matrices en $\mathcal{M}_n(\mathbb{C})$ vía

$$A_{ij} = \langle Ae_j, e_i\rangle \ , \quad i, j \in \mathbb{I}_n \ .$$

Con esta identificación, la matriz de A^* es la traspuesta conjugada de la matriz de A. Es decir que $A^*_{ij} = \overline{A_{ji}}$, $i, j \in \mathbb{I}_n$. ▲

Definición 2.1.5. Dado $A \in L(\mathbb{H})$ un operador en un espacio de Hilbert, decimos que A es:

1. **Hermitiano**comun]operador!hermitiano si $A = A^*$.
2. **anti-Hermitiano** comun]operador!anti-hermitiano si $A = -A^*$.

3. **unitario** comun]operador!unitario si $AA^* = A^*A = I$.

4. **normal**comun]operador!normal si $AA^* = A^*A$.

5. **definido positivo** comun]operador!definido positivo si $\langle Ax, x\rangle > 0$ para todo $x \in \mathbb{H}$. En tal caso de escribe $A > 0$.

6. **semidefinido positivo**comun]operador!semidefinido positivo si $\langle Ax, x\rangle \geq 0$ para todo $x \in \mathbb{H}$. En tal caso de escribe $A \geq 0$.

Los mismos nombres tendrán las matrices de $\mathcal{M}_n(\mathbb{C})$, al ser pensadas como operadores en $\mathbb{H} = \mathbb{C}^n$ con el producto escalar y la norma usuales. Además usaremos las siguientes notaciones: comun]matriz!hermitiana comun]matriz!anti-hermitiana comun]matriz!unitaria comun]matriz!normal comun]matriz!definida positiva comun]matriz!semidefinida positiva

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. $\mathcal{G}l\,(n) = \{A \in \mathcal{M}_n(\mathbb{C}) : A \text{ es invertible }\}$. ▲

Definición 2.1.6. Se llama espectro de una matriz $A \in \mathcal{M}_n(\mathbb{C})$ al conjunto de todos los autovalores de A:

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

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

Observación 2.1.7. 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

"los autovalores de A son $\lambda_1, \dots, \lambda_n$",

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

Definición 2.1.8. Sea $A \in \mathcal{M}_n(\mathbb{C})$.

1. El **radio numérico** de A se define como

$$w(A) = \max\{\ |\langle Ax, x\rangle|\ :\ x \in \mathbb{C}^n,\ \|x\| = 1\ \}\ .$$

2. El **radio espectral** de A se define como

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

3. La **norma espectral** de A es su norma como operador, inducida por la norma euclídea de $\mathbb{C}^n$. Es decir,

$$\|A\|_{sp} = \max\{\|Ax\| : x \in \mathbb{C}^n, \|x\| = 1\} = \min\{C \geq 0 : \|Ax\| \leq C\|x\|\ ,\ \ x \in \mathbb{C}^n\}\ .$$

Observar que $\rho(A) \le w(A) \le \|A\|_{sp}$. Tomando la matriz $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, se ve que las desiguadades pueden ser estrictas. En efecto,

$$\rho(A) = 0 \ , \quad w(A) = 1/2 \quad \text{y} \quad \|A\|_{sp} = 1 \ .$$

Ejercicio: verificarlo. ▲

Definición 2.1.9. Una matriz $T \in \mathcal{M}_n(\mathbb{C})$ se llama *triangular superior* si verifica que $T_{ij} = 0$ para $i > j$. Similarmente se definen matrices estrictamente triangulares superiores ($T_{ij} = 0$ si $i \ge j$), triangulares inferiores y estrictamente triangulares inferiores. Denotaremos por $\mathcal{TS}(n)$ al conjunto de matrices triangulares superiores en $\mathcal{M}_n(\mathbb{C})$. ▲

Teorema 2.1.10 (Schur 1)**.** 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 $T \in \mathcal{TS}(n)$ que verifican:

1. $T_{ii} \ = \ \lambda_i$, para todo $i \in \mathbb{I}_n$ 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 para en uno solo de los casos (hay que elegir A o B). ■

Corolario 2.1.11. Sea $A \in \mathcal{M}_n(\mathbb{C})$ con autovalores $\lambda_1, ..., \lambda_n$. Entonces

$$\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i \qquad y \qquad \det A = \prod_{i=1}^{n} \lambda_i \ .$$

Corolario 2.1.12. Sea $U \in \mathcal{U}(n)$. Entonces $|\det U| = 1$.

Corolario 2.1.13. Sea $P \in \mathbb{C}[x]$ y $A \in \mathcal{M}_n(\mathbb{C})$. Entonces,

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

2.2. Autoadjuntas

Teorema 2.2.1. Si $A \in \mathcal{H}(n)$, entonces existen $D \in \mathcal{H}(n)$ diagonal, y $U \in \mathcal{U}(n)$ tales que $A = UDU^*$. Por lo tanto, $\sigma(A) \subseteq \mathbb{R}$.

Definición 2.2.2. 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. Escribiremos $\lambda(A) = (\lambda_1(A), \ldots, \lambda_n(A))$ para denotar al vector de autovalores de A ordenados en forma **creciente**, es decir $\lambda_k(A) \le \lambda_{k+1}(A)$, $k \in \mathbb{I}_{n-1}$.
2. $\mu(A) = (\mu_1(A), \ldots, \mu_n(A))$ será el vector de autovalores de A ordenados en forma **decreciente**, es decir $\mu_k(A) \ge \mu_{k+1}(A)$, $k \in \mathbb{I}_{n-1}$. 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. ▲

Teorema 2.2.3 (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, si $A \in \mathcal{H}(n)$, $A \geq 0$ si y sólo si $\lambda_{min}(A) \geq 0$ si y sólo si $\sigma(A) \subseteq \mathbb{R}_+$.

Teorema 2.2.4 (Courant-Fisher)**.** Sea $A \in \mathcal{H}(n)$ y sea $k \in \mathbb{I}_n$. 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,

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

Teorema 2.2.5 (Teorema de Weyl)**.** *Sean $A, B \in \mathcal{H}(n)$ y $j \in \mathbb{I}_n$. 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 2.2.3, 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. ■

Corolario 2.2.6. *Sean $A, B \in \mathcal{H}(n)$ tales que $A \leq B$. Entonces, para todo $j \in \mathbb{I}_n$,*

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

Demostración. Aplicar el Teorema de Weyl a A y $B - A$, y notar que $\lambda_1(B - A) \geq 0$. ■

Definición 2.2.7. Sean $A \in \mathcal{M}_n(\mathbb{C})$ y $J \subseteq \mathbb{I}_n$. Si J tiene k elementos, notaremos

$$A_J = \{a_{ij}\}_{i,j\in J} \in \mathcal{M}_k(\mathbb{C}).$$

Es decir que A_J es la matriz cuadrada resultante de borrar de A las filas y columnas con índices fuera de J. En particuar, para cada $r \in \mathbb{I}_n$, llamaremos

$$A_r = A_{\mathbb{I}_n\setminus\{r\}} = \{a_{ij}\}_{i\neq r\neq j} \subseteq \mathcal{M}_{n-1}(\mathbb{C}), \tag{10}$$

a la submatriz principal obtenida de borrar la fila y la columna r-ésimas de A. ▲

Teorema 2.2.8 (Entrelace de Cauchy)**.** Sea $A \in \mathcal{H}(n)$, $r \in \mathbb{I}_n$ y $A_r \in \mathcal{M}_{n-1}(\mathbb{C})$ la submatriz principal de A obtenida como en la ecuación (10). Entonces

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

para cada $k \in \mathbb{I}_{n-1}$. Es decir que

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

Ejercicio 2.2.9. Escribir explícitamente cómo quedan los resultados de esta sección (minimax, teorema de Weyl y entrelace) en función de los vectores $\mu(A)$ ordenados decrecientemente. ▲

Recordemos que $A \in \mathcal{M}_n(\mathbb{C})^+$ si $\langle Ax, x\rangle \geq 0$ para todo $x \in \mathbb{C}^n$.

Definición 2.2.10. Dadas $A, B \in \mathcal{H}(n)$, se dice que $A \leq B$ si se tiene que $B - A \in \mathcal{M}_n(\mathbb{C})^+$, o sea si $\langle Ax, x\rangle \geq \langle Bx, x\rangle$ para todo $x \in \mathbb{C}^n$. ▲

Proposición 2.2.11. Sean A, B y $C \in \mathcal{M}_n(\mathbb{C})$. Entonces

1. $A \in \mathcal{M}_n(\mathbb{C})^+$ si y sólo si $A \in \mathcal{H}(n)$ y $\sigma(A) \subseteq \mathbb{R}_+$.
2. $A \in \mathcal{G}l(n)^+$ si y sólo si $A \in \mathcal{H}(n)$ y $\sigma(A) \subseteq \mathbb{R}_+^*$.
3. Si $A = B^*B$ entonces $A \in \mathcal{M}_n(\mathbb{C})^+$.
4. Si $A, B \in \mathcal{H}(n)$ y $A \leq B$, entonces $C^*AC \leq C^*BC$.

Teorema 2.2.12. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces $A \in \mathcal{M}_n(\mathbb{C})^+$ si y sólo si existe $B \in \mathcal{M}_n(\mathbb{C})$ tal que $A = B^*B$. Además existe una única matriz $B \in \mathcal{M}_n(\mathbb{C})^+$ tal que $A = B^*B = B^2$.

2.3. Descomposición polar y valores singulares

Definición 2.3.1. Dada $A \in \mathcal{M}_n(\mathbb{C})^+$, llamaremos $A^{1/2}$ a la única raiz cuadrada de A en $\mathcal{M}_n(\mathbb{C})^+$, que existe (y es única) por el Teorema 2.2.12.

Definición 2.3.2. Sea $A \in \mathcal{M}_n(\mathbb{C})$,

1. Llamaremos comun]módulo de A "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 $s_1(A) \geq \cdots \geq s_n(A) \geq 0$. Notar que, por el Corolario 2.1.13, $s_i(A) = \mu_i(|A|) = \mu_i(A^*A)^{1/2}$, $i \in \mathbb{I}_n$.
3. Llamaremos $s(A) = (s_1(A), \ldots, s_n(A)) = \mu(|A|)$ y $\Sigma(A)$ a la matriz diagonal
$$\Sigma(A) = \operatorname{diag}(s(A)) = \begin{pmatrix} s_1(A) & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & s_n(A) \end{pmatrix}.$$

▲

Ejemplos 2.3.3. 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 $s(A) = |\mu(A)|^{\downarrow}$. En efecto, $A^*A = A^2$, luego los autovalores de $|A|$ son los módulos de los de A. (si $a \in \mathbb{R}$, entonces $(a^2)^{1/2} = |a|$). Más aún, si A es normal, sigue valiendo que $s(A) = |\mu(A)|^{\downarrow}$ (Ejercicio: verificarlo).
3. En general (fundamentalmente, si A no es normal), 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. $Ae_k = e_{k+1}$, $k \in \mathbb{I}_{n-1}$ 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 2.3.4 (Descomposición 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 $U \in \mathcal{U}(n)$ tal que se tiene la factorización

$$A = U\ |A|\ ,$$

que es la llamada **descomposición polar** (DP) de A, aunque no es siempre única.

3. Cualquier $U \in \mathcal{U}(n)$ que cumpla $A = U|A|$, verifica que $UA^*AU^* = AA^*$, y por lo tanto

$$U|A|U^* = |A^*| \quad \text{y} \quad A = |A^*|U.$$

Esto dice que U^* es un unitario admisible para la DP de A^*.

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

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

Además, las columnas $C_i(V)$ forman una b.o.n. de autovectores de $|A|$ (y A^*A), y las columnas $C_i(W)$ forman una b.o.n. de autovectores de $|A^*|$ (y AA^*).

Observación 2.3.5 (Partes positiva y negativa de una matriz autoadjunta)**.** Sea $A \in \mathcal{H}(n)$, y tomemos su DP $A = U|A|$, con $U \in \mathcal{U}(n)$. Luego se verifican las siguientes propiedades:

1. Si diagonalizamos a A en una base ortonormal, luego $A^*A = A^2$, $|A|$ y U son diagonales en esa base, y por lo tanto conmutan entre ellos (y con A).

2. Si se asume que U opera como la identidad en $\ker A = R(A)^\perp$, entonces U es diagonal con ± 1's en la diagonal (en aquella base ortonormal). Por lo tanto

$$U^* = U = U^{-1} \qquad \text{y} \qquad -I \le U \le I\ .$$

3. Podemos deducir que $-|A| \le A \le |A|$. En efecto, $|A|^{1/2}U|A|^{1/2} = A$, y

$$-|A| = -|A|^{1/2}I|A|^{1/2} \le |A|^{1/2}U|A|^{1/2} \le |A|^{1/2}I|A|^{1/2} = |A|.$$

4. Luego, si denotamos

$$A_+ = \frac{A + |A|}{2} \qquad \text{y} \qquad A_- = \frac{|A| - A}{2}\ ,$$

se prueba fácilmente que

Ambas matrices $A_+\ , A_- \in \mathcal{M}_n(\mathbb{C})^+$. $A = A_+ - A_-$ y $|A| = A_+ + A_-$. $A_+A_- = A_-A_+ = 0$.

Es fácil ver que A_+ y A_- son las únicas matrices que cumples las tres propiedades anteriores. Se las llama partes positiva y negativa de la matriz autoadjunta A.

5. Otras propiedades que verifican A_+ y A_- son:

   1. $AA_+ = A_+A = (A_+)^2$ (idem con A_-).
   2. $\mu_k(A_+) = \max\{\mu_k(A), 0\}$, $k \in \mathbb{I}_n$.
   3. $\mu_k(A_-) = -\min\{\mu_{n-k+1}(A), 0\}$, $k \in \mathbb{I}_n$.

Observación 2.3.6. Existe una versión de la caracterización minimax, de Courant-Fisher, para los valores singulares de una matriz $A \in \mathcal{M}_n(\mathbb{C})$ (no necesariamente autoadjunta), muy útil para probar desigualdades. Con las mismas notaciones (para subespacios) que en el Teorema 2.2.4, se tiene que

$$s_k(A) = \max_{\dim \mathcal{M} = k} \ \min_{x \in \mathcal{M}_1} \|Ax\| = \min_{\dim \mathcal{S} = n-k+1} \ \max_{x \in \mathcal{S}_1} \|Ax\|. \tag{11}$$

En efecto, basta notar que $\|Ax\| = \langle A^*Ax, x\rangle^{1/2}$ y que $s_k(A) = \mu_k(A^*A)^{1/2}$. Luego se aplica el Teorema 2.2.4 (y el Ejercicio 2.2.9) para A^*A. De la fórmula (11) podemos deducir la siguiente importante desigualdad: Sean $A, C \in \mathcal{M}_n(\mathbb{C})$. Entonces, para todo $k \in \mathbb{I}_n$, se tiene que

$$s_k(AC) \le \|A\|_{sp}\ s_k(C)\ . \tag{12}$$

Se usa que $\|ACx\| \le \|A\|_{sp}\ \|Cx\|$, $x \in \mathbb{C}^n$. En particular esto implica la desigualdad

$$\operatorname{tr}|AC| \le \|A\|_{sp}\ \operatorname{tr}|C|\ ,$$

que será usada más adelante. ▲

2.4. Teorema de Schur y aplicaciones

Teorema 2.4.1 (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 $D = \operatorname{diag}(\lambda(A))$, 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 $i, j \in \mathbb{I}_n$,

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

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 1.2.9 completa la demostración. ■

Observación 2.4.2. Otra demostración del Teorema mayorización de Schur puede hacerse por inducción, aplicando el Teorema de entrelace de Cauchy 2.2.8. Para ello basta reordenar la diagonal de A, conjugandola por una matriz de permutación, lo que no cambia sus autovalores. ▲

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.

Proposición 2.4.3 (Principio del Máximo de Ky Fan)**.** *Sea $A \in \mathcal{H}(n)$. Entonces para todo $k \in \mathbb{I}_n$, se tiene que*

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

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

Demostración. Fijemos $k \in \mathbb{I}_n$. 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 2.4.1,

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

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

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

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

Ejercicios 2.4.4. Sea $A \in \mathcal{H}(n)$. Identificando las k-uplas ortonormales de $\mathbb{C}^n$ con bases de rangos de proyectores, y con columnas de isometrías de $\mathbb{C}^k$ en $\mathbb{C}^n$, probar que:

1. Si, para $k \in \mathbb{I}_n$, notamos $\mathcal{P}_k(n) = \{P \in \mathcal{H}(n) : P^2 = P \text{ y } \operatorname{rk}(P) = k\}$, entonces
$$\sum_{j=1}^{k} \mu_j(A) = \max_{P \in \mathcal{P}_k(n)} \operatorname{tr} PAP \ . \tag{13}$$

2. Para $k \in \mathbb{I}_n$, notamos $\mathcal{U}_k(n) = \{U \in \mathcal{M}_{n,k}(\mathbb{C}) : U^* U = I_k\}$, es decir, el espacio de isometrías de $\mathbb{C}^k$ en $\mathbb{C}^n$. Entonces
$$\sum_{j=1}^{k} \mu_j(A) = \max_{U \in \mathcal{U}_k(n)} \operatorname{tr} U^* AU \ .$$

▲

Ejercicio 2.4.5. Sea $A \in \mathcal{M}_n(\mathbb{C})$.

Teorema 2.4.6. *Sean A y $B \in \mathcal{H}(n)$. Entonces*

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

Demostración. Por la fórmula (13), para $k : 1, ..., n$, podemos escribir

$$\begin{aligned} \sum_{j=1}^{k} \mu_j(A+B) &= \max_{P \in \mathcal{P}_k(n)} \operatorname{tr} P(A+B)P \\ &\leq \max_{P \in \mathcal{P}_k(n)} \operatorname{tr} PAP + \max_{P \in \mathcal{P}_k(n)} \operatorname{tr} PBP \\ &= \sum_{j=1}^{k} \mu_j(A) + \sum_{j=1}^{k} \mu_j(B). \end{aligned}$$

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

Recordar que, dada $C \in \mathcal{M}_n(\mathbb{C})$, entonces si

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

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

$$\mu(\widehat{C}) = (s_1(C), \cdots, s_n(C), -s_n(C), \cdots, -s_1(C)). \tag{14}$$

Corolario 2.4.7. *Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Entonces*

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

Demostración. Notar que $\widehat{A+B} = \widehat{A} + \widehat{B}$. Por la observación anterior, y las desigualdades resultantes de la relación $\mu(\widehat{A+B}) \prec \mu(\widehat{A}) + \mu(\widehat{B})$ para $k \in \mathbb{I}_n$, se tiene que $s(A+B) \prec_w s(A) + s(B)$. ■

Observación 2.4.8. Notar que el resultado anterior es necesario para verificar que las normas $\|\cdot\|_{(k)}$ de Ky Fan, definidas por

$$\|A\|_{(k)} = \sum_{i=1}^{k} s_i(A) \ , \quad A \in \mathcal{M}_n(\mathbb{C}) \ ,$$

cumplen la desigualdad triangular.

2.5. Partes reales

Definición 2.5.1. Si $A \in \mathcal{M}_n(\mathbb{C})$, se llama **parte real** de A a

$$\operatorname{Re} A = \frac{A + A^*}{2} \in \mathcal{H}(n).$$

Si $x \in \mathbb{C}^n$, notaremos $\operatorname{Re} x \in \mathbb{R}^n$ al vector de las partes reales de sus cordenadas.

Proposición 2.5.2 (Fan-Hoffman)**.** *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces*

1. $\mu_k(\operatorname{Re} A) \leqslant \mu_k(|A|) = s_k(A)$, *para todo* $k \in \mathbb{I}_n$.
2. *Existe* $U \in \mathcal{U}(n)$ *tal que* $\operatorname{Re} A \leq U|A|U^*$.

Demostración. Sean $x_1, \ldots, x_n$ y $w_1, \ldots, w_n$ bases ortonormales de $\mathbb{C}^n$, formadas por autovectores de $\operatorname{Re} A$ (resp. A^*A) asociadas a $\mu(\operatorname{Re} A)$ (resp. $\mu(A^*A)$). Fijemos k entre 1 y n. Sea

$$x \in \operatorname{span}\{x_1, \ldots, x_k\} \cap \operatorname{span}\{w_k, \ldots, w_n\} \ ,$$

un vector unitario (debe existir por las dimensiones de los subespacios). Entonces, por Courant-Fisher 2.2.4 y la Observación 2.3.6,

$$\begin{aligned} \mu_k(\operatorname{Re}(A)) \quad &\leq \langle \operatorname{Re} Ax, x\rangle = \operatorname{Re}\langle Ax, x\rangle \leq |\langle Ax, x\rangle| \\ &\leq \|Ax\| = \langle A^*Ax, x\rangle^{1/2} \leq \mu_k(A^*A)^{1/2} = \mu_k(|A|) = s_k(A). \end{aligned}$$

La segunda parte se deduce de la primera, dado que $\operatorname{diag}(\mu(\operatorname{Re} A)) \leq \Sigma(A)$. ■

Proposición 2.5.3 (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, y tal que $\langle Ax_i, x_i\rangle = \mu_i(A)$ (que existe por el Teorema 2.1.10). Dado $k \in \mathbb{I}_n$, por el Principio del máximo de Ky Fan (Proposición 2.4.3), se tiene que

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

La igualdad para $k = n$ se debe a que

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

■

Corolario 2.5.4. *Si* $A \in \mathcal{M}_n(\mathbb{C})$ *cumple que* $A + A^* > 0$*, entonces*

$$\sigma(A) \subseteq \{z \in \mathbb{C} : \operatorname{Re} z > 0\} \ .$$

En realidad, se puede cambiar $\operatorname{Re} z > 0$ *por* $\mu_n(\operatorname{Re} A) \leq \operatorname{Re} z \leq \mu_1(\operatorname{Re} A)$.

2.6. Teorema de Schur-Horn

2.6.1. Sea $x \in \mathbb{C}^n$ con $\|x\| = 1$ (a estos vectores los llamaremos *unitarios*). Entonces la matriz $P_x = xx^* = (x_i \overline{x_j})_{ij} \in \mathcal{M}_n(\mathbb{C})$ es el proyector ortogonal sobre el subespacio span $\{x\}$. En efecto, notar que dado $z \in \mathbb{C}^n$,

$$P_x z = (xx^*)z = x(x^*z) = \langle z, x\rangle x,$$

que es la conocida fórmula de dicho proyector. Por lo tanto, si $\mathcal{B} = \{x_1, \dots, x_n\}$ es una b.o.n. de $\mathbb{C}^n$, vale que

$$z = \sum_{i=1}^{n} \langle z, x_i\rangle x_i \ , \quad z \in \mathbb{C}^n \quad \Longrightarrow \quad I = \sum_{i=1}^{n} x_i x_i^* \ . \tag{15}$$

▲

Proposición 2.6.2. *Sea* $A \in \mathcal{M}_n(\mathbb{C})$*. Se tiene que* $A \in \mathcal{M}_n(\mathbb{C})^+$ *si y sólo si existen* $x_1, \dots, x_r \in \mathbb{C}^n$ *unitarios y* $\lambda_i, \dots, \lambda_r \in \mathbb{R}_+$ *tales que*

$$A = \sum_{i=1}^{r} \lambda_i x_i x_i^* \ , \qquad \text{o sea que} \qquad A = \sum_{i=1}^{r} \lambda_i P_{x_i} \ . \tag{16}$$

Demostración. Por un lado es claro que si A cumple (16), entonces $A \geq 0$. Recíprocamente, si $A \in \mathcal{M}_n(\mathbb{C})^+$, sea $\mathcal{B} = \{x_1, \dots, x_n\}$ es una b.o.n. de autovectores de A asociados a los autovectores $\mu_1(A), \dots, \mu_n(A)$. Usando la ecuación (15), si $z \in \mathbb{C}^n$,

$$Az = A\Big(\sum_{i=1}^{n} \langle z, x_i\rangle x_i\Big) = \sum_{i=1}^{n} \langle z, x_i\rangle A x_i = \sum_{i=1}^{n} \mu_i(A)\langle z, x_i\rangle x_i = \Big(\sum_{i=1}^{n} \mu_i(A) x_i x_i^*\Big) z \ .$$

Luego A cumple (16). ■

Observación 2.6.3. Notar que la mínima cantidad r de proyectores de rango uno que puede usarse para obtener una representación de A como en (16) es $r = \operatorname{rk}(A)$. Además, si $A \in \mathcal{M}_n(\mathbb{C})^+$ cumple (16), definiendo $y_i = \lambda_i^{1/2} x_i$, se tiene que las matrices $y_i y_i^*$ no son más proyectores, pero sí se verifica que $A = \sum_{i=1}^{r} y_i y_i^*$. ▲

Es natural preguntarse, dado $A \in \mathcal{M}_n(\mathbb{C})^+$ y $r \geq \operatorname{rk}(A)$, para qué sucesiones $\lambda_i, \dots, \lambda_r$ en $\mathbb{R}_+$ se puede obtener para A una representación como (16). Este problema está íntimamente relacionado con el llamado Teorema de Schur-Horn. Recordemos que si $A \in \mathcal{M}_n(\mathbb{C})$, llamamos $\mathrm{d}(A) = (a_{11}, \dots, a_{nn}) \in \mathbb{C}^n$.

Proposición 2.6.4. *Sean* $c \in \mathbb{R}_+^n$ *y* $A \in \mathcal{M}_n(\mathbb{C})^+$*. Son equivalentes:*

1. *Existen vectores unitarios* $x_1, \dots, x_n \in \mathbb{C}^n$ *tales que* $A = \sum_{j=1}^{n} c_j \ x_j \ x_j^*$.

2. Existe $B \in \mathcal{H}(n)$ tal que $\mathrm{d}(B) = c$ *y* $\mu(B) = \mu(A)$*, o sea que* $B \simeq A$.

Demostración. Si se verifica 1, sea $X \in \mathcal{M}_n(\mathbb{C})$ definida por $C_k(X) = c_k^{1/2} x_k$, $k \in \mathbb{I}_n$. Veamos que $XX^* = A$. En efecto, si llamamos $X_k \in \mathcal{M}_n(\mathbb{C})$ a la matriz cuya columna k-ésima es $C_k(X)$, pero las demás son nulas, se tiene que

$$X = \sum_k X_k \ , \quad X_k X_j^* = 0 \quad \text{ si } \quad j \neq k \quad \text{y} \quad X_k X_k^* = c_k \ x_k x_k^* .$$

Es claro que todo esto implica que $XX^* = A$. Por lo tanto $\mu(A) = \mu(XX^*) = \mu(X^*X)$. Si $B = X^*X$, es fácil ver, además, que $B_{ii} = c_i \|x_i\|^2 = c_i$, $i \in \mathbb{I}_n$, lo que prueba 2. Recíprocamente, si $B \in \mathcal{H}(n)$ cumple $\mu(B) = \mu(A)$ y $\mathrm{d}\,(B) = c$, sea $U \in \mathcal{U}(n)$ tal que $U^*AU = B$ (se puede hacer, pasando por $\mathrm{diag}\,(\mu(A))$). Consideremos la matriz $X = A^{1/2}U$. Entonces $X^*X = B$ y $XX^* = A^{1/2}U\ U^*A^{1/2} = A$, mientras que $c_i = B_{ii} = \|C_i(X)\|^2$. Basta ahora definir $x_i = \dfrac{C_i(X)}{\|C_i(X)\|}$, y se verifica como antes que

$$A = XX^* = \sum_{j=1}^n c_j \ x_j \ x_j^* \ ,$$

lo que prueba 1. ∎

Teorema 2.6.5. *Sean $b, c \in \mathbb{R}^n$ Entonces son equivalentes:*

1. *$c \prec b$.*
2. *Existe $B \in \mathcal{H}(n)$ tal que $\mathrm{d}\,(B) = c$ y $\mu(B) = b^{\downarrow}$.*

Si, además, b y c tienen entradas no negativas, lo anterior equivale a

3. *Existen vectores unitarios $x_1, \dots, x_n \in \mathbb{C}^n$ tales que*

$$\mathrm{diag}\,(b) = \sum_{j=1}^n c_j \ x_j \ x_j^* .$$

Demostración. Antes que nada, notemos que se puede suponer que $b = b^{\downarrow}$ y $c = c^{\downarrow}$, porque las tres condiciones son invariantes por permutaciones (en el caso de 2 y 3, vía conjugar con matrices de permutación adecuadas). Notemos $A = \mathrm{diag}\,(b)$. El Teorema de Schur 2.4.1 muestra que 2 implica 1. La Proposición 2.6.4 muestra que 2 y 3 son equivalentes, cuando b y c tienen entradas no negativas.

Verificaremos, en principio, que 1 implica 3 en el caso en que b y c tienen entradas estrictamente positivas. Lo haremos por inducción en n. Si $n = 1$ no hay nada que probar. Sea $n > 1$. Como $b_1 \geq c_1 \geq b_n$, sea $k \in \mathbb{I}_n$ tal que $b_k \geq c_1 \geq b_{k+1}$.

Se afirma: existe $x_1 \in \mathrm{span}\,\{e_k, e_{k+1}\}$ de norma uno, tal que $A_1 = A - c_1 x_1 x_1^*$ tiene rango a lo sumo $n - 1$. En efecto, para ver que el tal x_1 existe, definimos

$$x(t) = \cos \frac{t\,\pi}{2}\ e_k + \sin \frac{t\,\pi}{2}\ e_{k+1} \quad \text{y} \quad A(t) = A - c_1 x(t) x(t)^* \ , \quad t \in [0, 1].$$

Entonces la curva $d(t) = \det A(t)$ es continua. Pero $d(1) \leq 0 \leq d(0)$ porque $A(0)$ y $A(1)$ son matrices diagonales, con un sólo elemento diagonal no positivo (es $b_{k+1} - c_1$) en el caso de $A(1)$, y con todos no negativos (anche $b_k - c_1$) en el de $A(0)$. Luego basta tomar $x_1 = x(t)$ para algún $t \in [0, 1]$ tal que $d(t) = 0$.

Es claro que existe una b.o.n. $\{y_1, y_2\}$ de $\mathrm{span}\,\{e_k, e_{k+1}\}$ tal que $A_1 y_1 = (b_k + b_{k+1} - c_1) y_1$ y $A_1 y_2 = 0$. Luego la matriz de A_1 en la bon $\mathcal{B} = \{y_2, e_1, \dots, e_{k-1}, y_1, e_{k+2}, \dots, e_n\}$ queda

$$\Big[A_1\Big]_{\mathcal{B}} = \mathrm{diag}\,(0, b_1, \dots, b_{k-1}, (b_k + b_{k+1} - c_1), b_{k+2}, \dots, b_n)\,. \tag{17}$$

Sean $a, d \in \mathbb{R}^{n-1}$, dados por $a = (c_2, \dots, c_n)$ y

$$d = (b_1, \dots, b_{k-1}, (b_k + b_{k+1} - c_1), b_{k+2}, \dots, b_n).$$

Notar que, como $b_k \geq c_1 \geq b_{k+1}$, entoncecs $d_{k-1} = b_{k-1} \geq b_k \geq d_k \geq b_{k+1} \geq b_{k+2} = d_{k+1}$. Para aplicar la HI al asunto anterior, deberíamos probar que $a \prec d$. En efecto, si $r \leq k$,

$$\sum_{i=1}^{r-1} a_i = \sum_{i=2}^{r} c_i \leq \sum_{i=1}^{r-1} c_i \leq \sum_{i=1}^{r-1} b_i = \sum_{i=1}^{r-1} d_i.$$

Si $k+1 \leq r \leq n-1$,

$$\sum_{i=1}^{r} c_i \leq \sum_{i=1}^{r} b_i \Longrightarrow \sum_{i=1}^{r-1} a_i = \sum_{i=2}^{r} c_i \leq \sum_{i=1}^{k-1} b_i + (b_k + b_{k+1} - c_1) + \sum_{i=k+2}^{r} b_i = \sum_{i=1}^{r-1} d_i$$

y las trazas andan bien porque $c \prec b$. En consecuencia, $a \prec d$. Sean, por HI, $n-1$ vectores unitarios $z_2, \dots, z_n \in \mathbb{C}^{n-1}$ tales que $\operatorname{diag}(d) = \sum_{j=2}^{n} c_j \; z_j z_j^* \in \mathcal{M}_{n-1}(\mathbb{C})^+$. Luego

$$\sum_{j=2}^{n} c_j \; (0, z_j)(0, z_j)^* = \operatorname{diag}(0, d) = \Big[A_1\Big]_{\mathcal{B}} \in \mathcal{M}_n(\mathbb{C})^+. \tag{18}$$

Si definimos $x_2, \dots, x_n \in \mathbb{C}^n$ tales que las cordenadas de cada x_j en $\mathcal{B}$ sean $(0, z_j)$, resulta que son también unitarios (los z_j lo son y $\mathcal{B}$ es una bon). Traduciendo la ecuación (18) (pensada en la base $\mathcal{B}$) a cordenadas en la base canónica, obtenemos

$$A_1 = \sum_{j=2}^{n} c_j \; x_j x_j^* \in \mathcal{M}_n(\mathbb{C})^+$$

y, por lo tanto,

$$A = A_1 + c_1 x_1 x_1^* = \sum_{j=1}^{n} c_j \; x_j x_j^* \; .$$

Concluímos que 1 implica 3, si b y c tienen entradas estricatamente positivas.

De lo anterior podemos deducir que $1 \leftrightarrow 2$ en ese caso $(b, c > 0)$, porque $1 \rightarrow 3$ y $3 \leftrightarrow 2$. Pero esto se generaliza sin dificultad al caso general $(b, c \in \mathbb{R}^n)$ usando que, si $e = (1, \dots, 1) \in \mathbb{R}^n$, entonces para todo $m \in \mathbb{R}$ se tiene que $x + me \prec y + me$ si y sólo si $x \prec y$. Y que dada $B \in \mathcal{H}(n)$, entonces $\mathrm{d}(B + mI) = \mathrm{d}(B) + me$ y $\mu(B + mI) = \mu(B) + me$.

Finalmente, probado $1 \leftrightarrow 2$ en general, ahora por la Proposición 2.6.4, ya sabemos que 3 es equivalente a ellas si b y c tienen entradas no negativas. ■

Teorema 2.6.6. *Sea $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ y $A \in \mathcal{H}(n)$ tal que $\mu(A) = a^{\downarrow}$. Entonces,*

$$\big\{ \, \mathrm{d}\,(\, UAU^*) : U \in \mathcal{U}(n) \big\} = \{x \in \mathbb{R}^n : x \prec a\}.$$

Demostración. Si $B = UAU^*$, con $U \in \mathcal{U}(n)$, entoces $\mu(A) = \mu(B) = a^{\downarrow}$. Luego por el Teorema de mayorización de Schur 2.4.1, se tiene que $\mathrm{d}(B) \prec a$.

Recíprocamente, si $x \in \mathbb{R}^n$ cumple $x \prec a$, por el Teorema 2.6.5 existe $B \in \mathcal{H}(n)$ tal que $\mathrm{d}(B) = x$ y $\mu(B) = a^{\downarrow}$. Por lo tanto debe existir $U \in \mathcal{U}(n)$ tal que $B = UAU^*$. Luego $x \in \{\mathrm{d}\,(UAU^*) : U \in \mathcal{U}(n)\}$. ■

Corolario 2.6.7. *Sean $A \in \mathcal{M}_n(\mathbb{C})^+$ y $c \in \mathbb{R}_+^m$, con $m \geq n$. Entonces existen proyectores autoadjuntos $P_1, \dots, P_m$ de rango uno, tales que*

$$A = \sum_{k=1}^{m} c_k P_k \qquad \Longleftrightarrow \qquad c \prec (\mu(A), 0, \dots, 0) := \tilde{\mu}(A) \in \mathbb{R}_+^m.$$

Demostración. Sea $A_1 = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{M}_m(\mathbb{C})^+$. Luego $\mu(A_1) = \tilde{\mu}(A)$. Por el Teorema 2.6.6, $c \prec \tilde{\mu}(A)$ si y sólo si existe $U \in \mathcal{U}(m)$ tal que $\mathrm{d}\,(UA_1U^*) = c$. Es claro que si $P = \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{M}_m(\mathbb{C})^+$, entonces $UA_1U^* = UPA_1PU^*$. Por lo tanto, si llamamos $U_1 \in \mathcal{M}_{m,n}(\mathbb{C})$ a la parte no nula de UP, entonces $U_1AU_1^* = UA_1U^*$. Notar que $U_1^*U_1 = I_n$. Si definimos $T = A^{1/2}U_1^* \in \mathcal{M}_{n,m}(\mathbb{C})$, y $x_i = C_i(T) \in \mathbb{C}^n$ para $i \in \mathbb{I}_m$, se tiene que $UA_1U^* = T^*T$, por lo que $\|x_i\|^2 = c_i$, $i \in \mathbb{I}_m$. Por otro lado,

$$A = TT^* = \sum_{k=1}^{m} x_k x_k^* = \sum_{k=1}^{m} c_k P_k \ ,$$

donde $P_k = c_k^{-1} x_k x_k^*$ es proyector autoadjunto de rango uno, para $k \in \mathbb{I}_m$ (si $c_k = 0$, puede tomarse como P_k cualquier cosa). La recíproca se prueba definiendo $T \in \mathcal{M}_{n,m}(\mathbb{C})$ tal que tenga columnas $x_k = c_k^{1/2} y_k$, donde $y_k y_k^* = P_k$, $k \in \mathbb{I}_m$. El hecho de que $A = TT^*$ implica que existe una $U_1 \in \mathcal{M}_{m,n}(\mathbb{C})$ que cumple $T = A^{1/2}U_1^*$, $U_1^*U_1 = I_n$ y $\mathrm{d}\,(U_1AU_1^*) = \mathrm{d}\,(T^*T) = c$. Luego se extiende U_1 a una $U \in \mathcal{U}(m)$, con lo que $U_1AU_1^* = UA_1U^*$. ■

Observación 2.6.8. El resultado anterior resulve el problema planteado en el párrafo anterior a la Proposición 2.6.4, al menos para el caso $r \geq n$. Es fácil ver, usando el Teorema 2.6.5, que si $A \in \mathcal{M}_n(\mathbb{C})^+$, $\mathrm{rk}(A) \leq r < n$ y $c \in \mathbb{R}_+^r$, entonces la condición necesaria y suficiente para que A pueda ser representado $A = \sum_{k=1}^{r} c_k P_k$ para ciertos proyectores P_k de rango uno, es que $\mu(A) \succ (c, 0, \dots, 0) \in \mathbb{R}^n$. ▲

2.7. Normas unitariamente invariantes.

Definición 2.7.1. 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 2.3.4 se tiene que $|\|A\|| = |\|\Sigma(A)\||$. ▲

Definición 2.7.2. Dada una norma $N(\cdot)$ unitariamente invariante, se define la función $g_N : \mathbb{R}^n \to \mathbb{R}$ como

$$g_N(x_1, \dots, x_n) = N\left(\mathrm{diag}\,(x_1, \dots, x_n)\right)$$

▲

Proposición 2.7.3. *Sea N una NUI. Entonces:*

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

Observación 2.7.4. 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. Se deduce de que la aplicación $\mathbb{C}^n \ni x \mapsto \mathrm{diag}\,(x) \in \mathcal{M}_n(\mathbb{C})$ es lineal e inyectiva.

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 2.7.5. *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 2.7.6. *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 la Proposición 1.2.13, 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 $y_\sigma = (y_{\sigma(1)}, \ldots, y_{\sigma(n)})$. Por el Teorema 1.2.9, 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 2.7.7.

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 2.7.3.

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 2.7.8. *Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Entonces son equivalentes:*

1. *$N(A) \leq N(B)$ para toda norma unitariamente invariante N.*

2. *$\|A\|_{(k)} \leq \|B\|_{(k)}$ para todo $k \in \mathbb{I}_n$.*

3. *$s(A) \prec_w s(B)$.*

Demostración. Es consecuencia del Teorema 2.7.6 para funciones gauge simétricas, y del Teorema 2.7.7. En efecto, si $\|A\|_{(k)} \leq \|B\|_{(k)}$ para todo $k \in \mathbb{I}_n$, 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 2.7.9. *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. Aplicando el Corolario 2.2.6, obtenemos que

$$s_k(A) = \mu_k(A) \leq \mu_k(B) = s_k(B) \ , \quad \text{para todo} \quad k \in \mathbb{I}_n \ .$$

Luego basta aplicar el Corolario 2.7.8. ■

Corolario 2.7.10. *Sea N una NUI en $\mathcal{M}_n(\mathbb{C})$. Dadas $A, B \in \mathcal{M}_n(\mathbb{C})$, se tiene que*

1. *$N(AB) \leq \|A\|_{sp} \ N(B)$.*

Además, si N es normalizada (i.e., $N(E_{11}) = 1$),

2. *$\|A\|_{sp} \leq N(A) \leq \|A\|_1 = \operatorname{tr}|A|$.*

3. *N es una norma matricial.*

Demostración. 1. Se deduce de la desigualdad $s_k(AB) \leq \|A\|_{sp} \ s_k(B)$, $k \in \mathbb{I}_n$, vista en la fórmula (12), y del Corolario 2.7.8.

2. Sea g la función gauge simetrica asociada a N. Como N está normalizada, entonces $g(e_k) = 1$ para todo $k \in \mathbb{I}_n$. Luego,

$$\|A\|_{sp} = s_1(A) = g(s_1(A), 0, \dots, 0) \leq g(s(A)) = N(A).$$

Análogamente, $N(A) = g(s(A)) \leq \sum_{k=1}^{n} s_k(A) g(e_k) = \sum_{k=1}^{n} s_k(A) = \|A\|_1$.

3. Es claro usando lo anterior. ■

Proposición 2.7.11. *Sea $g : \mathbb{R}^n \to \mathbb{R}$, una función convexa e invariante por permutaciones, es decir que se $x \in \mathbb{R}^n$ y $P \in S_n$, entonces $g(x) = g(Px)$. Entonces, dados $x, y \in \mathbb{R}^n$,*

$$x \prec y \quad \Longrightarrow \quad g(x) \le g(y).$$

En particular, esto se verifica si g es una función gauge simetrica.

Demostración. Por el Teorema 1.2.9, si $x \prec y$, existen $P_1, \dots, P_m \in S_n$ y $\lambda \in \mathbb{R}^m$ tales que $\lambda \prec e_1 \in \mathbb{R}^m$ y $x = \sum_{i=1}^m \lambda_i P_i y$. Entonces

$$g(x) = g\Big(\sum_{i=1}^m \lambda_i P_i y\Big) \le \sum_{i=1}^m \lambda_i g(P_i y) = y.$$

Notar que si g es una fgs, es convexa por ser una norma (homogeneidad + DT). ■

Corolario 2.7.12. *Dadas $A, B \in \mathcal{H}(n)$. Entonces*

$$\mu(A) \prec \mu(B) \quad \Longrightarrow \quad s(A) \prec_w s(B).$$

Demostración. Como $A \in \mathcal{H}(n)$, se tiene que $s(A) = |\,\mu(A)|^{\downarrow}$. Por lo tanto, si g es una fgs, $g(s(A)) = g(\mu(A))$. Lo mismo pasa para B, y el resultado se deduce de la Porposición 2.7.11, aplicado a las fgs's $g_k(x) = \sum_{i=1}^k x_i^{\downarrow}$, $k \in \mathbb{I}_n$. ■

Corolario 2.7.13. *Dadas $A, B \in \mathcal{H}(n)$, si $\mu(A) \prec \mu(B)$ entonces $N(A) \le N(B)$ para toda norma unitariamente invariante N.*

Demostración. Se deduce del Corolario 2.7.12 y del Teorema 2.7.7. ■

Ejercicio 2.7.14. Probar que, si $x \in \mathbb{R}^n$, entonces, para cada $k \in \mathbb{I}_n$,

$$\sum_{i=1}^k x_i^{\downarrow} = \text{mín } \big\{ \|y\|_1 + k\|z\|_\infty : x = y + z \big\}.$$

2.8. 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 2.8.1. 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

$$\sum_{j=1}^k \mu_j(A) \;\le\; \sum_{j=1}^k \mu_j(B), \quad 1 \le k \le n \tag{19}$$

y $\operatorname{tr} A = \operatorname{tr} B$. ▲

Definición 2.8.2. 1. 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 $\mathbb{C}^n$, entonces

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

donde los bloques tienen los tamaños adecuados (por ejemplo, $B \in \mathcal{M}_k(\mathbb{C})$). $C_P(A)$ tiene siempre esa pinta, si uno trabaja en cordenadas de una b.o.n. que empiece generando $R(P)$ y termine generando $\ker P$.

2. Más generalmente, un sistema de proyectores en $\mathcal{M}_n(\mathbb{C})$ es una conjunto

$$\mathcal{P} = \{P_1, \dots, P_r\} \subseteq \mathcal{H}(n) \ ,$$

donde los P_i son proyectores no nulos tales que

$$P_i P_j = 0 \quad \text{si} \quad i \neq j \quad \text{y} \quad \sum_{i=1}^{r} P_i = I \ .$$

Notar que un proyector $P \in \mathcal{H}(n)$ define un sistema de dos proyectores $\mathcal{P} = \{P, I - P\}$.

3. Dado un sistema de proyectores $\mathcal{P} = \{P_1, \dots, P_r\}$ en $\mathcal{M}_n(\mathbb{C})$, se define el pinching asociado:

$$C_{\mathcal{P}} : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C}) \ , \quad \text{dado por} \quad C_{\mathcal{P}}(A) = \sum_{i=1}^{r} P_i A P_i \ , \quad A \in \mathcal{M}_n(\mathbb{C}),$$

que también puede verse como una compresión a bloques diagonales (operando en una b.o.n. adecuada). Notar que se tiene la siguiente factorización:

$$C_{\mathcal{P}} = C_{P_1} \circ C_{P_2} \circ \cdots \circ C_{P_r} \tag{20}$$

y lo mismo en cualquier otro orden entre los C_{P_i}. ▲

Ejercicios 2.8.3.

1. Si $A \in \mathcal{H}(n)$ y $\sigma(A) = \{\lambda_1, \dots, \lambda_r\}$ (todos distintos), entonces definiendo P_i como el proyector sobre $\ker(A - \lambda_i I)$, $i \in \mathbb{I}_r$, se obtiene un sistema de proyectores que verifica que $A = \sum_{i=1}^{r} \lambda_i P_i$.

2. Dado un sistema de proyectores $\mathcal{P}$ en $\mathcal{M}_n(\mathbb{C})$ y una matriz $A \in \mathcal{M}_n(\mathbb{C})$, se tiene que $C_{\mathcal{P}}(A) = A$ si y sólo si A conmuta con todos los P_i de $\mathcal{P}$. O sea, si A es diagonal de bloques. Verificar que eso sucede en el ejemplo anterior.

3. Probar que, dado un sistema de proyectores $\mathcal{P}$ en $\mathcal{M}_n(\mathbb{C})$, el operador pinching $C_{\mathcal{P}}$ verifica las siguientes propiedades:

   Es lineal, idempotente (i.e., $C_{\mathcal{P}} \circ C_{\mathcal{P}} = C_{\mathcal{P}}$) y $R(C_{\mathcal{P}})$ es el subespacio de matrices que conmutan con todos los P_i, $i \in \mathbb{I}_r$. Reduce normas y preserva trazas. Si $A \in \mathcal{M}_n(\mathbb{C})$ es autoadjunto (resp. positivo) entonces $C_{\mathcal{P}}(A)$ es autoadjunto (resp. positivo). Por lo tanto, en general, $C_{\mathcal{P}}(A^*) = C_{\mathcal{P}}(A)^*$. ▲

Proposición 2.8.4. *Sea $A \in \mathcal{H}(n)$ y $\mathcal{P}$ un sistema proyectores en $\mathcal{H}(n)$. Entonces*

$$C_{\mathcal{P}}(A) \prec A.$$

4. *Demostración.* Por la ecuación (20), basta considerar el caso de pinchings de un solo proyector $P \in \mathcal{H}(n)$ (o sea, el sistema $\mathcal{P} = \{P, I - P\}$). 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 2.4.6 se tiene

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

por lo que $C_P(A) \prec A$. ■

Ejercicio 2.8.5.

1. Clarificar en qué sentido la Proposición 2.8.4 es una generalización del Teorema de mayorización de Schur.
2. Dados $x, y, z, w \in \mathbb{R}^n$ tales que $x = x^\downarrow$, $y = y^\downarrow$, $z = z^\downarrow$ y $w = w^\downarrow$, probar que
$$z \prec w \quad \text{y} \quad x \prec y \implies x + z \prec y + w \ .$$
¿Es cierto si no estan ordenados?
3. Deducir del Teorema 2.4.6 (+ inducción) que, si $A_1, \dots, A_m \in \mathcal{H}(n)$, entonces
$$\mu\Big(\sum_{k=1}^m A_k\Big) \prec \sum_{k=1}^m \mu(A_k).$$

▲

Definición 2.8.6. Dado un espacio vectorial $\mathbb{V}$, si $\mathcal{C} \subseteq \mathbb{V}$, llamaremos **conv** $(\mathcal{C})$ a la cápsula convexa de $\mathcal{C}$:
$$\operatorname{conv}(\mathcal{C}) = \Big\{ \sum_{k=1}^m \lambda_k b_k \ : \ m \in \mathbb{N},\ b_k \in \mathcal{C},\ \lambda \in \mathbb{R}^m \text{ y } \lambda \prec (1, 0, \dots, 0) \Big\}.$$
es decir, el conjunto de todas las combinaciones convexas de elementos de $\mathcal{C}$. ▲

El siguiente teorema da una caracterización, intrínseca de matrices, de la mayorización matricial:

Teorema 2.8.7. *Sea $A \in \mathcal{H}(n)$. Denotemos por*
$$\mathcal{U}(A) = \{UAU^* : U \in \mathcal{U}(n)\} = \{B \in \mathcal{H}(n) : \mu(B) = \mu(A)\}$$
la órbita unitaria de A. Entonces,
$$\{T \in \mathcal{H}(n) : T \prec A\} = \operatorname{conv}(\mathcal{U}(A)\,)\,. \tag{21}$$
O sea que $T \prec A$ si y sólo si T es combinación convexa de conjugados unitarios de A.

Demostración. Para $k \in \mathbb{I}_m$, sean $U_k \in \mathcal{U}(n)$ y $\lambda_k \in [0, 1]$, tales que $\sum_k \lambda_k = 1$. Consideremos
$$T = \sum_{k=1}^m \lambda_k \ U_k A U_k^* \in \operatorname{conv}(\mathcal{U}(A)\,)\,.$$
Por el Ejercicio 2.8.5,
$$\mu(T) \prec \sum_{k=1}^m \mu(\lambda_k \ U_k A U_k^*) = \sum_{k=1}^m \lambda_k \ \mu(A) = \mu(A) \implies T \prec A.$$
Recíprocamente, sea $T \in \mathcal{H}(n)$ tal que $\mu(T) \prec \mu(A)$. Notaremos $a = \mu(A)$. Entonces, por el Teorema 1.2.9, existe $\mathcal{C} \subseteq S_n$ tal que, si notamos $a_\sigma = (a_{\sigma(1)}, \dots, a_{\sigma(n)})$, $\sigma \in \mathcal{C}$, entonces
$$\mu(T) = \sum_{\sigma \in \mathcal{C}} \lambda_\sigma \ a_\sigma$$
para ciertos $\lambda_\sigma \in [0, 1]$ tales que $\sum_{\sigma \in \mathcal{C}} \lambda_\sigma = 1$. Sea, para cada $\sigma \in \mathcal{C}$, $P_\sigma \in \mathcal{U}_{\mathcal{P}}(n)$ la matriz asociada (definida en la base canónica por $P_\sigma(e_k) = e_{\sigma^{-1}(k)}$). Notemos $D = \operatorname{diag}(a)$. Entonces $P_\sigma^* D P_\sigma = \operatorname{diag}(a_\sigma)$. Por lo tanto,
$$\operatorname{diag}(\mu(T)) = \sum_{\sigma \in \mathcal{C}} \lambda_\sigma \ P_\sigma^* D P_\sigma.$$
Finalmente, si $V, W \in \mathcal{U}(n)$ hacen que $A = V^* D V$ y $T = W \operatorname{diag}(\mu(T))\, W^*$, entonces
$$T = W \operatorname{diag}(\mu(T))\, W^* = \sum_{\sigma \in \mathcal{C}} \lambda_\sigma \ W P_\sigma^* D P_\sigma W^* = \sum_{\sigma \in \mathcal{C}} \lambda_\sigma \ (W P_\sigma^* V) \ A \ (V^* P_\sigma W^*).$$
Luego $T \in \operatorname{conv}(\{UAU^* : U \in \mathcal{U}(n)\}) = \operatorname{conv}(\mathcal{U}(A)\,)$. ■

Observación 2.8.8. Notar que el Teorema 2.8.7 permite generalizar el Corolario 2.7.13 en el siguiente sentido: Si $A, B \in \mathcal{H}(n)$, por la fórmula (21) se tiene que

$$A \prec B \quad \Longrightarrow \quad N(A) \leq N(B)$$

para toda norma N que verifique $N(UCU^*) = N(C)$, $C \in \mathcal{M}_n(\mathbb{C})$, $U \in \mathcal{U}(n)$. Estas normas se llaman débilmente unitariamente invariantes (NDUI). Notar que, por ejemplo,

$$w(C) = \text{máx}\left\{|\langle Cx, x\rangle| : \|x\| = 1\right\}, \quad C \in \mathcal{M}_n(\mathbb{C}),$$

que se llama *radio numérico* de C, es una tal norma, pero no es NUI. Otro ejemplo de este tipo es $M(C) = N(C) + |\operatorname{tr} C|$, que es NDUI para cualquier norma unitariamente invariante N. ▲

Observación 2.8.9. Para el caso de pinchings, se tiene un resultado más general que la Proposición 2.8.4: Dado un sistema de proyectores $\mathcal{P} = \{P_1, \ldots, P_r\}$ en $\mathcal{H}(n)$, se tiene que, para toda NDUI N,

$$N(C_{\mathcal{P}}(A)) \leq N(A) \quad , \qquad \text{para toda} \qquad A \in \mathcal{M}_n(\mathbb{C}) \ , \tag{22}$$

es decir, incluso para A **no autoadjunta**. En efecto, por la ecuación (20) basta probarlo para un solo proyector $P \in \mathcal{H}(n)$. En tal caso, sea $U = P - (I - P) \in \mathcal{U}(n)$. Es fácil ver que

$$2\ C_P(A) = A + UAU = A + UAU^*.$$

Con esto la desigualdad, al tomar N, es clara. Otra forma de verlo es observar que, en el caso general, siempre se verifica que $C_{\mathcal{P}}(A) \in \operatorname{conv}(\{UAU^* : U \in \mathcal{U}(n)\})$, aunque A no sea autoadjunta. Esto se hace eligiendo las matrices unitarias y diagonales de bloques (para $\mathcal{P}$), con $\pm\ I_{R(P_i)}$ en cada bloque. ▲

2.9. Teoremas de Lindskii y sus aplicaciones.

Recordemos que, si $B \in \mathcal{M}_n(\mathbb{C})$, llamamos $s(B) = (s_1(B), \ldots, s_n(B)) = \mu(|B|)$, al vector de valores singulares de B, ordenados en forma **decreciente**.

Teorema 2.9.1 (Lidskii 1). *Sean $A, B \in \mathcal{H}(n)$. Entonces*

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

Ejercicio 2.9.2.

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 2.4.6, $\mu(A) \prec \mu(A - B) + \mu(B)$. Por lo tanto, para todo $k \in \mathbb{I}_n$,

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

   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. ▲

Teorema 2.9.3 (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 2.9.2, se entenderá porque es necesaria (y suficiente) la siguiente versión más técnica del Teorema de Lidskii. Obsrvar que puede verse como una generalización natural del Teorema de Weyl 2.2.5 (que, de hecho, es lo que se usa en su prueba).

Teorema 2.9.4 (Lidskii 3)**.** *Sean* $A, B \in \mathcal{H}(n)$*,* $k \in \mathbb{I}_n$ *y* $J \subseteq \mathbb{I}_n$ *con* $|J| = k$*. Entonces*

$$\sum_{j\in J} \mu_j(A) + \sum_{i=1}^{k} \lambda_i(B) \le \sum_{j\in J} \mu_j(A+B) \le \sum_{j\in J} \mu_j(A) + \sum_{i=1}^{k} \mu_i(B). \tag{23}$$

Demostración. Probaremos, en principio, la desiguadad de la derecha en (23). Sin pérdida de generalidad podemos suponer que $\mu_k(B) = 0$. En efecto, si probaramos el resultado para $B - \mu_k(B)I$ (que cumple lo pedido) y A, podríamos deducir inmediatamente la deigualdad de la derecha en (23), dado que

$$\mu_j(A+B) = \mu_j(A+B-\mu_k(B)I) + \mu_k(B) \ \text{ y } \ \mu_i(B) = \mu_i(B-\mu_k(B)I) + \mu_k(B) \ ,$$

para todo $j \in J$ e $i \in \mathbb{I}_n$ (sobrará $k\mu_k(B)$ en ambos términos, y se puede cancelar). Sea $B = B_+ - B_-$ la descomposición de B en partes positiva y negativa, definida en la Observación 2.3.5. Como $\mu_k(B) = 0$, se tiene que:

$$\operatorname{tr}(B_+) = \sum_{j=1}^{k} \mu_j(B).$$

Por el teorema de Weyl, como $A + B \le A + B_+$, deducimos que $\mu_j(A+B) \le \mu_j(A+B_+)$ para todo $j \in \mathbb{I}_n$. En consecuencia,

$$\sum_{j\in J} \Big(\mu_j(A+B) - \mu_j(A)\Big) \le \sum_{j\in J} \Big(\mu_j(A+B_+) - \mu_j(A)\Big).$$

Finalmente, usando nuevamente el teorema de Weyl, resulta que $\mu_j(A+B_+) \ge \mu_j(A)$, para todo $j \in \mathbb{I}_n$. Por lo tanto

$$\begin{aligned}
\sum_{j\in J} \Big(\mu_j(A+B_+) - \mu_{i_j}(A)\Big) &\le \sum_{j=1}^{n} \Big(\mu_j(A+B_+) - \mu_j(A)\Big) \\
&= \operatorname{tr}(A+B_+) - \operatorname{tr}(A) = \operatorname{tr}(B_+) \\
&= \sum_{j=1}^{k} \mu_j(B) \ .
\end{aligned}$$

Esto prueba la desiguadad de la derecha en la ecuación (23). La otra desigualdad se deduce de la anterior, pero aplicada al conjunto $J' = \{n - j + 1 : j \in J\}$ y las matrices $-A$ y $-B$. Se usa que

$$\mu_r(-C) = -\mu_{n-r+1}(C) = -\lambda_r(C) \ ,$$

para cualquier $C \in \mathcal{H}(n)$ y para todo $r \in \mathbb{I}_n$. ■

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

$$\sum_{j=1}^{k} \big(\mu(A) - \mu(B)\big)_j^{\downarrow} = \max_{J\subseteq \mathbb{I}_n \ |J|=k} \left\{ \sum_{j\in J} \mu_j(A) - \mu_j(B) \right\} \le \sum_{j=1}^{k} \mu_j(A-B), \tag{24}$$

para todo $k \in \mathbb{I}_n$. Entonces, $\mu(A) - \mu(B) \prec \mu(A - B)$. ■

Demostración del Teorema de Lidskii 2. Usando la fórmula (14), si $k \in \mathbb{I}_n$, entonces

$$\max_{J\subseteq \mathbb{I}_{2n} \ |J|=k} \left\{ \sum_{j\in J} \mu_j\left(\widehat{A}\right) - \mu_j\left(\widehat{B}\right) \right\} = \sum_{j=1}^{k} \big|s(A) - s(B)\big|_j^{\downarrow} \ ,$$

porque si $j \in \mathbb{I}_n$, entonces, por la fórmula (14),

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

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

Por otro lado, aplicando la fórmula (24) a $\widehat{A}$, $\widehat{B}$ y a $\widehat{A-B} = \widehat{A} - \widehat{B}$, se tiene

$$\max_{J\subseteq \mathbb{I}_{2n}\ |J|=k} \left\{ \sum_{j\in J} \mu_j\left(\widehat{A}\right) - \mu_j\left(\widehat{B}\right) \right\} \le \sum_{j=1}^{k} \mu_j\left(\widehat{A-B}\right) = \sum_{j=1}^{k} s_j\left(A-B\right),$$

y podemos concluir que $\displaystyle\sum_{j=1}^{k} \left|s(A)-s(B)\right|_j^{\downarrow} \le \sum_{j=1}^{k} s_j\left(A-B\right)$ para todo $k \in \mathbb{I}_n$. ■

Corolario 2.9.5. *Sean N una NUI en $\mathcal{M}_n(\mathbb{C})$ y $A, B \in \mathcal{M}_n(\mathbb{C})$. Entonces*

$$N\big(\Sigma(A) - \Sigma(B)\big) \le N(A-B) \ .$$

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

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

Luego se aplica el Teorema 2.7.8. ■

Corolario 2.9.6. *Sean N 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*

$$\mathrm{d}_{_N}(A, \mathcal{U}(n)) = N(A-U) = N(\Sigma(A) - I) \ .$$

Demostración. Sea $V \in \mathcal{U}(n)$. Entonces $\Sigma(V) = I$ y, por el Corolario 2.9.5, $N(A-V) \ge N(\Sigma(A)-I)$. 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 N es una NUI, resulta que $N(A-U) = N(\Sigma(A) - I)$. ■

Ejercicio 2.9.7. 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

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

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

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

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.

Referencias

[1] T. Ando, Totally positive matrices, Linear Algebra Appl. 90 (1987), 165-219.

[2] R. Bhatia; *Matrix Analysis*, Springer, New York, 1997.

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

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

[5] R. Horn y C. Johnson; *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[6] R. Horn y C. Johnson; *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[7] P. D. Lax, Linear Algebra, Springer Verlag, Berlín, 1998.

[8] L. Mirsky, An introduction to Linear Algebra, Clarendon Press, Oxford, 1963.

[9] M. L. Metha, Matrix Theory, 2a Ed., Hindustan Publishing Co. 1989.

[10] R. Bellman, Introduction to Matrix Analysis, 2a Ed., McGraw-Hill, New York, 1970.

[11] W. F. Donoghue, Jr., Monotone matrix functions and analytic continuation, Springer-Verlag, Berlín, 1974.

[12] A. W. Marshall y I. Olkin, Inequalities: Theory of Mayorization and its Applications, Academic Press, New York, 1979.

[13] B. Simon, *Trace ideals and their applications*, London Mathematical Society Lecture Note Series, 35, Cambridge University Press, Cambridge-New York, 1979.

Ejercicios 2.9.8. Sean $A, B \in \mathcal{H}(n)$. Entonces

1. $\mu(A) + \lambda(B) \prec \mu(A+B) \prec \mu(A) + \mu(A)$.

2. Dados $x, y \in \mathbb{R}^n$, se tiene que
$$x^\downarrow - y^\downarrow \prec x - y \prec x^\downarrow - y^\uparrow \quad \text{y} \quad x^\downarrow + y^\uparrow \prec x + y \prec x^\downarrow + y^\downarrow .$$
Por lo tanto $g(x^\downarrow - y^\downarrow) \le g(x-y)$ para toda fgs.

3. Si, para $C \in \mathcal{H}(n)$, llamamos $E_d(C) = \operatorname{diag}(\mu(C))$ y $E_c(C) = \operatorname{diag}(\lambda(C))$, entonces
$$N\big(E_d(A) - E_d(B)\big) \le N(A-B) \le N\big(E_d(A) - E_c(B)\big) \quad \text{y}$$
$$N\big(E_d(A) + E_c(B)\big) \le N(A+B) \le N\big(E_d(A) + E_d(B)\big).$$

4. **Hoffman- andt** Sean $A, B \in \mathcal{M}_n(\mathbb{C})$ matrices normales. Sean $\mu(A)$ y $\mu(A)$ sus vectores de autovalores en algún orden. Entonces existe $\sigma \in S_n$ tal que
$$\sum_{k=1}^{n} \left|\mu_i(A) - \mu_{\sigma(i)}(B)\right|^2 \le \|A - B\|_2^2 .$$

5. Dada $A \in \mathcal{M}_n(\mathbb{C})$, probar que
$$\|A\|_{(k)} = \text{mín } \left\{ \|B\|_1 + k\|C\|_{sp} : A = B + C \right\}.$$

6. Usar lo anterior para dar una nueva demostración del Teorema 2.9.3 (Lidskii 2), mostrando previamente que, dadas $A, B \in \mathcal{H}(n)$,

   $\|\mu(A) - \mu(B)\|_\infty \leq \|A - B\|$. $\|\mu(A) - \mu(B)\|_1 \leq \|A - B\|_1 = \|A - B\|_{(n)}$.

7o Mostrar que el Teorema 2.9.3 implica el Teorema 2.9.1 (Lidskii 1).

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. $\mu_k(A) \geq \mu_k(A_{\mathcal{S}}) \geq \mu_{k+1}(A)$, para $k \in \mathbb{I}_{n-1}$.
2. Sea $v_1, \dots, v_n$ una b.o.n. de autovectores de A, asociados a $\mu_1(A), \dots, \mu_n(A)$ (respectivamente).
   a. Si $v_1, \dots, v_k \in \mathcal{S}$, entonces $\mu_i(A) = \mu_i(A_{\mathcal{S}})$ para $i \in \mathbb{I}_k$.
   b. Si $v_k, \dots, v_n \in \mathcal{S}$, entonces $\mu_{i-1}(A_{\mathcal{S}}) = \mu_i(A)$, $k \leq i \leq n$.

Probar el Teorema de Lidskii 3 (por inducción sobre n) usando el ejercicio anterior, y considerando independientemente los casos:

1. $i_k < n$,
2. $1 < i_1$,
3. $i_1 = 1$, $i_k = n$. ▲