Jorge Antezana y Demetrio Stojanoff

Análisis Matricial

Buenos Aires, Octubre de 2008.

Prefacio

El análisis matricial (AM) es una continuación natural del álgebra lineal, pero considerando que el cuerpo de escalares son los números reales o los complejos, y con una mirada basada en la problemática de las teorías de espacios de Hilbert y de sus operadores acotados.

Muchos problemas del análisis, la geometría diferencial, el análisis funcional, los sistemas dinámicos, la física teórica y otras importantes teorías, pueden "bajarse" al caso de matrices. Y eso sin mencionar todas las ramas de la matemática aplicada, que suelen tener a este tipo de reducciones como herramienta principal. En general, poder reducir y reformular un problema al caso matricial es un éxito, porque es mucho más viable resolver un problema de matrices que el problema de origen.

Por todo lo anterior, mencionar aplicaciones del AM es innecesario. Cualquier matemático quiere poder aplicarlo, y trata sistemáticamente de hacerlo. Porque es un contexto donde las cuentas se pueden hacer (o la mayoría cree a priori que deberían poder hacerse). Más aún, cuando la reducción a matrices de un problema P sigue siendo difícil, se puede concluir que P tenía una dificultad intrínseca. Pero con dificultad o sin ella, el tema es cómo resolver P en matrices.

Para poder hacerlo, hay que desarrollar a fondo una teoría de matrices, o al menos una extensa serie de herramientas para trabajar con ellas, que pueda resolver los inmumerables problemas que le "caen" de arriba. Podría decirse que eso es el AM.

Lo más interesante del AM es que es el contexto más basico (un alumno de segundo año de la licenciatura ya puede entender la mayoría de los enunciados) en el que se pueden plantear problemas matemáticos bien difíciles, muchos de ellos no resueltos aún. Pero para entender a fondo este tipo de problemas, sus ramificaciones, y las técnicas que se suelen aplicar para resolverlos, hace falta hacer un curso específico de AM, que pueda ser atendido tanto por matemáticos formados como por estudiantes de la licenciatura. Otra particularidad remarcable, es que con ese solo basamento, alcanza para leer y entender (y porqué no *hacer*) una gran cantidad de publicaciones actuales. Un típico trabajo final para un curso de AM, es estudiar un paper de los últimos 2 o 3 años del tema. Y en muchos casos, con los contenidos de este texto se tienen (casi) todas las herramientas para poder entenderlo a fondo.

Por otra parte, como toda teoría matemática, el AM tiene su problemática propia. El tema más típicamente matricial son las desigualdades, que involucran normas, autovalores, valores singulares, determinantes, trazas, etc. El estudio de desigualdades de matrices y operadores es de una gran sutileza y forma una especie de mundo aparte. Sus especialistas son unos tipos especiales, una especie de gremio de artesanos. Las técnicas que se usan suelen ser intrincadas y de un gran ingenio. Se aplican ideas de toda la matemática, pero la teoría tiene sus reglas propias y toda una gama de herramientas y métodos espcecíficos. Una de esas herramientas, fundamental y no muy conocida, es otro tema central para el AM: La teoría de mayorización (de vectores y matrices), y sus múltiples ramificaciones. Esta teoría, elemental pero difícil, está poco difundida entre los matemáticos, por lo que ha sido y sigue siendo redescubierta innumerables veces en distintas áreas, muchas veces con terminologías ad hoc. Si bien la mayorización aparece como una forma de comparar vectores de $\mathbb{R}^n$, cuando se la piensa en vectores de autovalores o de valores singulares, se percibe rápidamente que es una noción intrínsecamente relacionada con la teoría de matrices. Estos dos aspectos: mayorización y desigualdades, son desarrollados con profundidad en este texto.

Una rama muy diferenciada del AM, la de matrices de entradas no negativas, llamada teoría de Perron y Frobenuis, podría tener el rango de área independiente. De ella daremos un capítulo con las bases principales de la teoría, y otro capítulo exponiendo una de sus ramas: las matrices totalmente positivas.

Este libro es el resultado de más de una decena de cursos, dictados en diversos departamentos de matemática (FCEN-UBA, FI-UBA, FCE-UNC y, sobre todo, en la FCE-UNLP) y en varios congresos, en los últimos años. Es importante aclarar que se asumen como conocidos (y no se exponen en este texto) todos los contenidos de un curso inicial de álgebra lineal. Para comodidad del lector, y para fijar notaciones y prerrequisitos, se enumeran al pricipio del primer capítulo todas las nociones y resultados específicos de un tal curso que serán usados a lo largo del texto. Cualquier libro de álgebra lineal (y hay miles) sirve como referencia para los mismos. Si me dan a elegir, yo recomiendo el de K. Hoffman y R. Kuntze [6] para algebristas, y el de P. D. Lax [10] para analistas.

Debo mencionar que este libro está fuertemente basado en cuatro excelentes textos que son la bibliografía básica en el tema: los dos tomos Matrix Analysis [7] y Topics of Matrix Analysis [8] de R. Horn y C. Johnson, el reciente libro [4] y, sobre todo, el maravilloso libro Matrix Analysis [3], ambos de R. Bhatia. Sin embargo, hay varios aspectos que lo diferencian. Por un lado, el presente texto está pensado como base para un curso elemental, y organizado efectivamente para que todo el material pueda darse en un cuatrimestre. Por otro lado, hay fuertes diferencias en la manera de encarar muchos de los temas, y se incluyen resultados más modernos y numerosas pruebas simplificadas de resultados clásicos, en base a publicaciones recientes o al aporte de alumnos, ayudantes y profesores de todos los cursos antes mencionados.

Los temas elegidos son solo una pequeña parte de la teoría, pero son la base principal sobre la que se edifican la mayoría de las áreas no incuidas en el texto. Hay muchas técnicas de análisis, funciones analíticas y geometría diferencial que suelen ser efectivas para problemas de matrices. Ese tipo de interacciones no están incluidos porque el texto está pensado para una audiencia no necesariamente experta. Una referencia escencial para esa clase de recursos son los libros mencionado de R. Bhatia [3] y [4]. Otra teoría aparte, poco desarrollada aquí, es la de perturbaciones de matrices (autovalores, autovectores, etc). Sobre estos temas, se podrían mencionar varios capítulos de [3], y también el monumental tratado de T. Kato [9]. Tampoco se hace mucho incapié en este texto en los métodos algorítmicos, que vendrían a ser la otra pata de la teoría. Bajado un problema a matrices, hay dos alternativas: resolverlo teóricamente (para ese lado va este texto) o resolverlo aproximando, dado que en matrices se puede (si no son muy grandes). La bibliografía sobre aproximación mediante algoritmos es inmensa, y nos limitaremos a citar el excelente tratado Matrix computations [2] de G. Golub y C. F. Van Loan, y la bibliografía que allí aparece. La mayoría de las herramientas necesarias para los algoritmos mencionados, y muchos de los procedimientos específicos que ellos usan, sí están expuestos en el texto; pero sin hacer hincapié en la óptica de la velocidad de convergencia o la robustez ante perturbaciones, sinó en la problemática teórica que presentan. Otros temas que no tratamos son los de matrices diagonalizables, polinomios minimales y formas canónicas, en particular la forma de Jordan. Esto es porque ese tipo de resultados no se usarán en el resto del texto, y porque suelen estar incluidos en un buen curso básico de álgebra lineal. Los dos libros antes mencionados ([6] y [10]) dan excelentes tratamientos de estos temas.

Muchos de los resultados que expondremos siguen siendo válidos en contextos más generales que las matrices reales o complejas. Por ejemplo matrices a coeficientes en cuerpos generales o en anillos, álgebras de Banach, operadores en espacios de Banach o de Hilbert, ágebras de operadores (C^* y de von Neumann). Esto sucede particularmente con resultados de los capítulos 1 (en las secciones 5, 7 y 9), 3, 6, 7, 8 (sección 3), 9, 10 y 12. La decisión que tomamos para presentarlos fue dar demostraciónes específicas para el caso matricial y, por lo general, mencionar luego los contextos donde siguen valiendo, y las técnicas diferentes para cada uno de ellos. La principal razón que justifica este enfoque es que el libro busca ser autocontenido en un nivel elemental, y que las teorías mencionadas son muy variadas, lo que haría muy difícil dar las demostraciones generales sin largas secciones introductorias de cada una de ellas. Además, las pruebas para el caso matricial suelen ser muchísimo más simples y breves, y brindan un manejo interesante de las técnicas propias del AM. Por otra parte, opinamos que es muy útil el enfrentarse con una primera versión de enunciados complicados en un ámbito menos complicado, para después poder entender el significado de esos enunciados en los contextos más específicos (además de su nueva demostración).

Sin embargo, este enfoque tiene un límite. Por lo tanto, una parte importante del AM hemos decidido desarrollarlo en el ambiente más general de operadores en espacios de Hilbert. Se seleccionaron para esa parte aquellos resultados cuyas pruebas difieren poco al aumentar la generalidad, y que forman una rama imporante de la teoría de operadores, aunque mantengan un espíritu claramente matricial. Sin embargo, ese trabajo se realizará en un segundo volumen, dado que el contenido del presente libro ya es suficiente para un curso cuatrimestral, y porque la segunda parte requiere una introducción específica de espacios de Hilbert que no consideramos necesaria para este texto puramente matricial.

Los contenidos del libro están suficientemente explicitados en los títulos de las secciones del índice. A continuación haremos algunos comentarios sobre el enfoque aplicado en cada capítulo. Como se dijo anteriormente, al principio del capítulo 1 se enumera una serie de notaciones y resultados del álgebra lineal elemental. En la sección 5 se presentan varias fórmulas elemetales, pero no demasiado conocidas, para operar con matrices. De particular importancia es el manejo de matrices de bloques y las técnicas para operar con ellas. Luego se presenta el teorema de Schur que muestra la equivalencia unitaria de toda matriz con una triangular superior. Este teorema, si bien suele estar incluido en los textos elementales, es presentado en detalle porque será de importancia clave para numerosos resultados a lo largo de todo el libro. El capítulo culmina con tres secciones de resultados elementales, que también serán muy usados más adelante: polinomios aplicados a matrices, descomposición QR y las propiedades básicas de las matrices de rango uno.

Los capítulos 2 y 3, sobre matrices normales, autoadjuntas y positivas, empiezan con material básico, desarrollan en detalle las propiedades variacionales de los autovalores de matrices autoadjuntas, y dan una versión finitodimensional de los principales resultados de la teoría de operadores en espacios de Hilbert, pero con las notaciones tradicionales del AM. Se propone un estudio exhaustivo de las propiedades y caracterizaciones de las matrices definidas positivas, dado que suelen ser las protagonistas de las más interesantes desigualdades que se estudiarán más adelante. Por otra parte, muchos problemas generales de matrices pueden reducirse al caso positivo, a traves de yeites como tomar partes reales e imaginarias (ahí se cae en las autoadjuntas) y luego positivas y negativas, usando matrices de bloques de 2×2, o bien usando la descomposición polar.

Los capítulos 4 y 5 tratan sobre mayorización, primero en su versión vectorial, y después en sus primeras aplicaciones a las matrices. El tratamiento es muy detallado, porque consideramos que es un tema poco difundido, y que es sumamente útil en muchas ramas de la matemática, además de ser escencial para el AM. El capítulo 6, sobre monotonía y convexidad de operadores, incluye una introducción al cálculo funcional para matrices autoadjuntas, en el estilo del de operadores en espacios de Hilbert, pero con pruebas ad hoc. Luego se dan las principales caracterizaciones y propiedades de las funciones mencionadas, que son herramientas escenciales para estudiar desigualdades de matrices y operadores. Este capítulo esta fuertemente basado en la exposición de estos temas que se hace en el libro de Bhatia [3]. Sin embargo, hay importantes diferencias de enfoque, se presentan muchas pruebas diferentes, y la selección de resultados presentados es distinta.

En el capítulo 7 se da una introducción básica a la teoría de productos tensoriales y alternados, como siempre con pruebas adecuadas al contexto matricial. Esta teoría, por ser bastante ardua de exponer, suele aparecer mencionada sin mucho detalle en los libros, en función de poder aplicar los recursos que brinda (escenciales para entender las propiedades de los determinantes y como herramienta para probar desigualdades) sin meterse en camisa de once varas. Aquí intentamos dar una exposición detallada y (casi) autocontenida, dado que el contexto matricial lo permite sin que el esfuerzo sea excesivo, y porque en capítulos posteriores necesitaremos trabajar con propiedades muy específicas del determinante de matrices y submatrices. El tema es que una buena presentación de los productos alternados permite justificar completamente todas esas propiedades, trabajo que se inicia en la sección 3, y se continúa en el capítulo 12.

El capítulo 8 trata sobre productos de Hadamard. Aquí también el tratamiento es muy detallado, porque es un área que aparece poco en los tratados del tema, es un tema fuerte de investigación dentro del AM, y tiene además muy intereantes aplicaciones en otras disciplinas. Se presenta una pueba completa del teorema de Haagerup que caracteriza la norma del operador de multiplicación (de Hadamard) por una matriz fija, relativa a la norma espectral de matrices.

El capítulo 9 presenta una serie de importantes desigualdades de matrices, y puede pensarse como lugar en el que se concentran las técnicas y desarrollos realizados en los capítulos anteriores. La lista no es exhaustiva, pero da una idea de las principales líneas de la teoría, y presenta la mayoría de las técnicas usuales que se utilizan para mostrar este tipo de desigualdades.

En el capítulo 10 se estudian las principales propiedades del rango y del radio numéricos de matrices. Los teoremas más importantes que desarrollamos son el de Hausdorff-Toeplitz sobre la convexidad del rango numérico, y el de T. Ando sobre caracterizaciones matriciales del radio numérico.

Los últimos tres capítulos enfocan la teoría de Perron-Frobenius sobre las matrices de entradas positivas, y las totalmente positivas. En el 11 se exponen los resultados clásicos sobre matrices estrictamente positivas, no negativas e irreducibles. En el 12 se introducen los complementos de Schur y numerosas técnicas con determinandtes que, además de tener un interés propio, son la herramienta clave para desarrollar, en el capítulo 13, una introducción a la teoría de matrices totalmente positivas. Esta capítulo se basa en un trabajo de T. Ando [20], y está escrito utilizando como punto de partida al trabajo final de A. Iglesias para un curso de AM dictado en La Plata.

Todos los capítulos tienen una última sección de ejercicios. Se proponen además numerosos ejercicios a lo largo del texto de cada capítulo. Al principio de las secciones finales se los reenumeran, agreagándose a continuación series de ejercicios nuevos.

Qerríamos agradecer a Gustavo Corach por haber iniciado y habernos incluido en el trabajo de investigación de nuestro grupo del IAM en los temas de Análisis Matricial. También va nuestro agradecimiento a Celeste González, a partir de cuyo trabajo [25] se comenzo a escribir la primera versión de este libro, a Pedro Massey, que nos aportó invalorables comentarios e ideas (además de muchos ejercicios), y a Agustín Iglesias e Ivan Angiono, de quienes hemos tomado algunos fragmentos de texto. También agradecemos a los alumnos de los distintos cursos que hemos dictado en estos años, que han aportado una infinidad de sugerencias, correcciones e ideas.

Índice General

Capítulo 1

Preliminares

1.1 Generalidades

1.1.1. Para empezar, enumeraremos las notaciones y convenciones más básicas, sobre vectores y matrices, que usaremos a lo largo de todo el texto:

1. Usaremos a $\mathbb{C}$ o $\mathbb{R}$ como cuerpo de escalares.
2. Llamaremos $\mathbb{R}_+$ al conjunto de números reales no negativos, y $\mathbb{R}_+^*$ al conjunto de números reales positivos.
3. Dado $n \in \mathbb{N}$, usaremos el símbolo $\mathbb{I}_n$ para denotar al conjunto $\{1, 2, \ldots, n\} \subseteq \mathbb{N}$.
4. Llamaremos $\mathcal{M}_{n,m}(\mathbb{C}) = \mathbb{C}^{n\times m}$, al espacio de matrices rectangulares de $n \times m$.
5. Cuando $n = m$, notaremos $\mathcal{M}_n(\mathbb{C}) = \mathcal{M}_n = \mathcal{M}_{n,n}(\mathbb{C})$, a las matrices cuadradas de $n\times n$ sobre $\mathbb{C}$.
6. Para matrices reales escribiremos $\mathcal{M}_{n,m}(\mathbb{R}) = \mathbb{R}^{n\times m}$ y $\mathcal{M}_n(\mathbb{R}) = \mathcal{M}_{n,n}(\mathbb{R})$.
7. Para denotar las entradas de una matriz $A \in \mathcal{M}_{n,m}(\mathbb{C})$, usaremos indistintamente, por conveniencia del contexto, las notaciones $A = (A_{ij})_{\substack{i\in\mathbb{I}_n \\ j\in\mathbb{I}_m}}$ o $A = (a_{ij})_{\substack{i\in\mathbb{I}_n \\ j\in\mathbb{I}_m}}$.
8. Dada $A \in \mathcal{M}_{n,m}(\mathbb{C})$, denotaremos por $A^T \in \mathcal{M}_{m,n}(\mathbb{C})$ a su matriz traspuesta, dada por $A^T_{ij} = A_{ji}$, para $i \in \mathbb{I}_n$ y $j \in \mathbb{I}_m$.
9. Dado $n \in \mathbb{N}$, denotaremos por $I \in \mathcal{M}_n(\mathbb{C})$, o bien I_n , si es que hace falta aclarar su tamaño, a la matriz identidad, dada por $I_{ij} = 1$ si $i = j$ e $I_{ij} = 0$ si $i \neq j$.
10. La suma y producto de matrices (cuando sus tamaños lo permitan) se hacen con las definiciones usuales del álgebra lineal. Por ejemplo, si $A \in \mathcal{M}_{n,m}(\mathbb{C})$ y $B \in \mathcal{M}_{m,r}(\mathbb{C})$, entonces $AB \in \mathcal{M}_{n,r}(\mathbb{C})$ y sus entradas son

$$(AB)_{ij} = \sum_{k=1}^{m} A_{ik}\ B_{kj}\ , \qquad \text{para todo } i \in \mathbb{I}_n \text{ y todo } j \in \mathbb{I}_r\ . \tag{1.1}$$

11. Dada $A \in \mathcal{M}_n(\mathbb{C})$, diremos que A es inversible si existe $A^{-1} \in \mathcal{M}_n(\mathbb{C})$, la única matriz que cumple que $AA^{-1} = A^{-1}A = I$. Denotaremos por

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

que es un grupo (de Lie) con la multiplicación usual de matrices. Su neutro es I_n .

12. Asumiremos como conocidas las propiedades del "determinante", que denotaremos $\det : \mathcal{M}_n(\mathcal{A}) \to \mathcal{A}$, para cualquier $n \in \mathbb{N}$ y cualquier anillo conmutativo $\mathcal{A}$. En el Capítulo 7 sobre productos tensoriales, se darán definiciones precisas, y se demostrarán la mayoría de dichas propiedades. En el Capítulo 12 se profundizará ese estudio. Sin embargo, usaremos desde ahora esas propiedades, ad referendum de sus pruebas (esperemos que no haya círculos muy viciosos).

13. Dada $A \in \mathcal{M}_n(\mathbb{C})$, consideremos la matriz $xI_n - A \in \mathcal{M}_n(\mathbb{C}[x])$. El polinomio característico de A está dado por la fórmula $P_A(x) = \det(xI - A) \in \mathbb{C}[x]$. Es un polinomio mónico de grado n.

14. Dada $A \in \mathcal{M}_n(\mathbb{C})$, su traza es el número $\operatorname{tr} A = \sum\limits_{i=1}^{n} A_{ii}$. Usaremos el hecho conocido de que, si $B \in \mathcal{M}_n(\mathbb{C})$, entonces $\operatorname{tr} AB = \operatorname{tr} BA$.

15. Sea $A \in \mathcal{M}_{n,m}(\mathbb{C})$. Las columnas de A se pueden pensar como vectores de $\mathbb{C}^n$, y sus filas, como vectores de $\mathbb{C}^m$. Se usará la notación $C_i(A) \in \mathbb{C}^n$ (respectivamente $F_i(A) \in \mathbb{C}^m$) para denotar a la i-ésima columna (resp. fila) de A.

16. Los vectores de $\mathbb{C}^n$ serán pensados como vectores **columna**, es decir que identificamos $\mathbb{C}^n$ con $\mathbb{C}^{n\times 1}$. Sin embargo, a lo largo del texto los describiremos como una fila (estilo $x = (x_1, \dots, x_n) \in \mathbb{C}^n$), para ahorrar espacio. Por ejemplo, si $A \in \mathcal{M}_n(\mathbb{C})$ e $i \in \mathbb{I}_n$, entonces $C_i(A) = (a_{1i} , a_{2i} , \dots , a_{ni}) \in \mathbb{C}^n$.

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

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

Por ejemplo, si tomamos $\mathbb{1} = (1, \dots, 1) \in \mathbb{C}^n$, entonces $\operatorname{diag}(e) = I_n$.

18. Por otra parte, si $A \in \mathcal{M}_n(\mathbb{C})$, llamaremos $\mathrm{d}(A) \in \mathbb{C}^n$ al la diagonal de A pensada como vector, i.e. $\mathrm{d}(A) = (A_{11} , \dots , A_{nn})$. ▲

1.1.2 (Matrices y operadores)**.** Enumeraremos a continuación las propiedades de las matrices cuando son vistas como transformaciones lineales:

1. Dados dos $\mathbb{C}$-espacios vectoriales $\mathbb{V}$ y $\mathbb{W}$, llamaremos $L(\mathbb{V}, \mathbb{W})$ al $\mathbb{C}$-espacio vectorial de transformaciones lineales $T : \mathbb{V} \to \mathbb{W}$. Si $\mathbb{V} = \mathbb{W}$, ponemos $L(\mathbb{V}) = L(\mathbb{V}, \mathbb{V})$.

2. Dado un $\mathbb{C}$-espacio vectorial $\mathbb{V}$ y un conjunto $X \subseteq \mathbb{V}$, denotaremos por $\operatorname{Gen}\{X\}$ al subespacio de $\mathbb{V}$ generado por X. Si $X = \{x_1, \dots, x_m\}$, escribiremos también $\operatorname{Gen}\{X\} = \operatorname{Gen}\{x_1, \dots, x_m\}$.

3. Si $A \in \mathcal{M}_{n,m}(\mathbb{C})$, la pensaremos también como un elemento de $L(\mathbb{C}^m, \mathbb{C}^n)$ actuando por multiplicación: si $x \in \mathbb{C}^m = \mathcal{M}_{m,1}(\mathbb{C})$, entonces $A(x) = A \cdot x \in \mathbb{C}^n$, usando el producto de la Eq. (1.1). En este contexto usaremos las notaciones conocidas:
$$\ker A = \{x \in \mathbb{C}^m : Ax = 0\} \quad \text{y} \quad R(A) = A(\mathbb{C}^m) = \operatorname{Im}(A) .$$

4. Se denotará por $\mathcal{E} = \{e_1, \dots, e_m\}$ a la base canónica de $\mathbb{C}^m$. A veces seremos más explícitos, poniendo $\mathcal{E}_m = \{e_1^{(m)}, \dots, e_m^{(m)}\}$, para aclarar el contexto.

5. Si $A \in \mathcal{M}_{n,m}(\mathbb{C})$, entonces se tiene que
$$Ae_i^{(m)} = C_i(A) \in \mathbb{C}^n , \quad \text{para todo} \quad i \in \mathbb{I}_m .$$
Por lo tanto, tenemos que $R(A) = \operatorname{Gen}\{C_1(A), \dots, C_m(A)\}$.

6. Por el teorema de la dimensión, si $A \in \mathcal{M}_n(\mathbb{C}) \simeq L(\mathbb{C}^n)$, entonces
$$A \in \mathcal{G}l(n) \iff \ker A = \{0\} \iff R(A) = \mathbb{C}^n .$$

7. El **rango** de $A \in \mathcal{M}_{n,m}(\mathbb{C})$ es $\operatorname{rk}(A) = \dim R(A) = \dim \operatorname{Gen}\{C_1(A), \dots, C_m(A)\}$. Más adelante, en la Observación 3.7.4 (ver también el Ejercicio 1.1.15), veremos que coincide con el rango fila de A, que es la $\dim \operatorname{Gen}\{F_1(A), \dots, F_n(A)\}$.

8. 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 diremos que "pensamos" a $A|_{\mathcal{S}} \in L(\mathcal{S})$. ▲

Espacios de Hilbert finitodimensionales

1.1.3. En $\mathbb{C}^n$ consideraremos el producto interno (o escalar) común, dado por
$$\langle x, y \rangle = \sum_{k=1}^{n} x_k \, \overline{y_k} , \quad x, y \in \mathbb{C}^n . \tag{1.2}$$

Es claro que $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ verifica las propiedades que definen a 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:

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

A x se lo llama **unitario**, si $\|x\| = 1$. Muchas veces consideraremos otras normas de vectores y matrices. Por ello damos una definición general: ▲

Definición 1.1.4. 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. $N(\lambda v) = |\lambda|\; N(v)$. ▲

Definición 1.1.5. Sean $\mathbb{V}$ un K-espacio vectorial, con $K = \mathbb{C}$ o $\mathbb{R}$, y N una norma en $\mathbb{V}$.

1. Cuando N proviene de un producto interno $\langle\cdot,\cdot\rangle$, diremos que

   $$\text{el par } (\mathbb{V}, N)\ , \quad \text{o bien} \quad (\mathbb{V}, \langle\cdot,\cdot\rangle) \quad \text{es un } K\text{-espacio de Hilbert} .$$

   Cuando $K = \mathbb{C}$, también diremos que $\mathbb{V}$ es un "espacio de Hilbert" a secas. Ojo, acá se asume que $\dim \mathbb{V} < \infty$. Sinó hay que pedir que $\mathbb{V}$ sea completo.
2. 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 $\dim \mathbb{H} = \infty$).
3. Si $\mathbb{H} = \mathbb{K}$, escribimos $L(\mathbb{H})$ en lugar de $L(\mathbb{H}, \mathbb{H})$.
4. Si $A \in L(\mathbb{H}, \mathbb{K})$ notaremos por $\ker A$ a su núcleo y $R(A)$ a su imagen. ▲

Definición 1.1.6. Sea $\mathbb{H}$ un espacio de Hilbert.

1. Dados $x, y \in \mathbb{H}$, decimos que son ortogonales, y escribimos $x \perp y$ si $\langle x, y\rangle = 0$.
2. Sea $X \subseteq \mathbb{H}$. Denotaremos por $X^\perp = \{y \in \mathbb{H} : y \perp x \text{ para todo } x \in X\}$, al subespacio *ortogonal* a X.
3. Los vectores $x_1, \dots, x_k \in \mathbb{H}$ forman un conjunto **ortogonal** cuando $\langle x_i, x_j\rangle = 0$, para todo $i \neq j$.
4. 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**.

5. Usaremos las siglas **BON** para denotar a una base ortonormal de $\mathbb{H}$. Por ejemplo, la base canónica $\mathcal{E}_n$ es una BON de $\mathbb{C}^n$ con el producto interno de la Eq. (1.2). ▲

Definición 1.1.7. 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

$$\langle Ax, z\rangle_K = \langle x, A^* z\rangle_H \ , \quad x \in \mathbb{H},\ z \in \mathbb{K}. \tag{1.3}$$

La demostración de que A^* existe es un resultado básico de la teoría de espacios de Hilbert. En el caso finitodimensional, se puede construir a A^* usando BONes, como veremos. ▲

1.1.8 (Propiedades de la adjunta)**.** Sean $A, B \in L(\mathbb{H})$. Usando la Eq. (1.3) (y la unicidad) se verifican fácilmente las siguientes propiedades:

1. Supongamos que $\dim \mathbb{H} = n$. Si para cualquier BON fija $\mathcal{B} = \{v_1, \dots, v_n\}$ de $\mathbb{H}$, se identifica a los operadores de $L(\mathbb{H})$ con matrices en $\mathcal{M}_n(\mathbb{C})$ vía

$$A_{ij} = \langle Av_j, v_i\rangle \ , \quad i, j \in \mathbb{I}_n \ ,$$

   entonces la matriz de A^* es $\overline{A^T}$, la traspuesta conjugada de la matriz de A. En otras palabras, $A^*_{ij} = \overline{A_{ji}}$, para todo par $i, j \in \mathbb{I}_n$.

2. $(A^*)^* = A$.

3. Dado $\lambda \in \mathbb{C}$, se tiene que $(\lambda A + B)^* = \overline{\lambda}\ A^* + B^*$.

4. $\ker A = R(A^*)^\perp$ y, si $\dim \mathbb{H} < \infty$, también $R(A) = (\ker A^*)^\perp$.

5. $(AB)^* = B^*\ A^*$ e $I^* = I$.

6. A es inversible si y sólo si A^* es inversible. En tal caso, $(A^*)^{-1} = (A^{-1})^*$. ▲

Definición 1.1.9. Dado $A \in L(\mathbb{H})$ un operador en un espacio de Hilbert, decimos que A es:

1. **Hermitiano** si $A = A^*$.

2. **anti-Hermitiano** si $A = -A^*$.

3. **unitario** si $AA^* = A^*A = I$.

4. **normal** si $AA^* = A^*A$.

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

6. **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. En este contexto recordar que, si $A \in \mathcal{M}_n(\mathbb{C})$, entonces $A^* = \overline{A^T}$. Además usaremos las siguientes notaciones:

1. $\mathcal{H}(n) = \{A \in \mathcal{M}_n(\mathbb{C}) : A = A^*\}$.
2. $\mathcal{U}(n) = \{U \in \mathcal{M}_n(\mathbb{C}) : U$ es unitaria $\}$.
3. $\mathcal{N}(n) = \{N \in \mathcal{M}_n(\mathbb{C}) : N$ es normal $\}$.
4. $\mathcal{M}_n(\mathbb{C})^+ = \{A \in \mathcal{M}_n(\mathbb{C}) : A \geq 0\}$.
5. $\mathcal{G}l(n) = \{A \in \mathcal{M}_n(\mathbb{C}) : A$ es invertible $\}$ y $\mathcal{G}l(n)^+ = \mathcal{G}l(n) \cap \mathcal{M}_n(\mathbb{C})^+$. ▲

Lema 1.1.10 (Polarización)**.** *Sea $\mathbb{H}$ un $\mathbb{C}$-espacio de Hilbert y sea $A \in L(\mathbb{H})$. Entonces*

$$\langle Ax, y\rangle = \frac{1}{4} \sum_{k=1}^{4} i^k \left\langle A\,(x + i^k\, y), (x + i^k\, y) \right\rangle \quad \textit{para todo par } x, y \in \mathbb{H} \ . \tag{1.4}$$

Demostración. La cuenta es directa, y se deja como ejercicio. ■

Proposición 1.1.11. *Sea $\mathbb{H}$ es un $\mathbb{C}$-espacio de Hilbert y sea $A \in L(\mathbb{H})$. Luego, las siguientes condiciones son equivalentes:*

1. $A = 0$.
2. $\langle Ax, y\rangle = 0$ *para todo par* $x, y \in \mathbb{C}^n$.
3. $\langle Az, z\rangle = 0$ *para todo* $z \in \mathbb{C}^n$.

Demostración. Es claro que $1 \to 2 \to 3$. La implicación $3 \to 2$ es consecuencia directa de la fórmula de polarización (1.4). Si asumimos 2 y, para cualquier $x \in \mathbb{H}$, tomamos $y = Ax$, obtenemos que $\|Ax\|^2 = \langle Ax, Ax\rangle = \langle Ax, y\rangle = 0$, por lo que $A = 0$. ■

Observación 1.1.12. Es importante señalar que el ítem 3 no implica a los otros en el caso de que $\mathbb{H}$ sea sólo un $\mathbb{R}$-espacio de Hilbert. Observar que no puede valer la polarización. Peor aún, tomando como $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in L(\mathbb{R}^2)$ (una rotación de 90 grados), es claro que $\langle Ax, x\rangle = 0$ para todo $x \in \mathbb{R}^2$. Veamos ahora una aplicación a matrices: ▲

Corolario 1.1.13. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces*

1. $A \in \mathcal{H}(n)$ *si y sólo si* $\langle Az, z\rangle \in \mathbb{R}$ *para todo* $z \in \mathbb{C}^n$.
2. $\mathcal{M}_n(\mathbb{C})^+ \subseteq \mathcal{H}(n)$.

Demostración. Si $A \in \mathcal{H}(n)$ y $z \in \mathbb{C}^n$, entonces $\langle Az, z\rangle = \langle z, Az\rangle = \overline{\langle Az, z\rangle}$. Si suponemos que $\langle Az, z\rangle \in \mathbb{R}$ para todo $z \in \mathbb{C}^n$, entonces

$$\langle Az, z\rangle = \overline{\langle Az, z\rangle} = \langle z, Az\rangle = \langle A^*z, z\rangle \ , \ \forall\, z \in \mathbb{C}^n \implies \langle\, (A - A^*)\, z, z\rangle = 0 \ , \ \forall\, z \in \mathbb{C}^n \ .$$

Por la Proposición 1.1.11, deducimos que $A = A^*$. Por la definición de $\mathcal{M}_n(\mathbb{C})^+$, si se tiene que $A \in \mathcal{M}_n(\mathbb{C})^+$, entonces $\langle Az, z\rangle \in \mathbb{R}_+ \subseteq \mathbb{R}$ para todo $z \in \mathbb{C}^n$. ■

Ejercicio 1.1.14. Sean $A, \in \mathcal{M}_{m,n}(\mathbb{C})$. Entonces se tiene que

$$A^*A \in \mathcal{M}_n(\mathbb{C})^+ \quad \text{y} \quad (A^*A)_{i,j} = \langle\, C_j(A)\, , \, C_i(A)\, \rangle \ , \quad \text{para todo par } i, j \in \mathbb{I}_n \ .$$

En particular, $\operatorname{tr}(A^*A) = \sum_{j=1}^{n} \|C_j(A)\|^2 = \sum_{i,j=1}^{n} |a_{ij}|^2$. ▲

Ejercicio 1.1.15. Sean $A, \in \mathcal{M}_{m,n}(\mathbb{C}) \simeq L(\mathbb{C}^m, \mathbb{C}^n)$. Mostrar que entonces

$$\ker A = \operatorname{Gen}\{F_1(A), \dots, F_n(A)\,\}^{\perp} \subseteq \mathbb{C}^n \ .$$

Deducir que $\operatorname{rk}_{F(A)} := \dim \operatorname{Gen}\{F_1(A), \dots, F_n(A)\,\} = \operatorname{rk}(A)$, o sea que los rangos fila y columna de A coinciden. ▲

1.2 El espectro

Definición 1.2.1. Se llama espectro de una matriz $A \in \mathcal{M}_n(\mathbb{C})$ al conjunto de todos los autovalores de A:

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

que es un subconjunto finito y no vacío de $\mathbb{C}$. ▲

1.2.2 (Propiedades del espectro de matrices)**.** Sea $A \in \mathcal{M}_n(\mathbb{C})$. Valen:

1. $\lambda \in \sigma\,(A)$ si y sólo si existe $x \in \mathbb{C}^n$ tal que $x \neq 0$ y $Ax = \lambda x$.
2. Si $\mu \in \mathbb{C}$, entonces $\sigma(A + \mu I) = \sigma(A) + \mu = \{\lambda + \mu : \lambda \in \sigma(A)\}$.
3. $A \in \mathcal{G}l(n)$ si y sólo si $0 \notin \sigma\,(A)$. Más aún, $\lambda \notin \sigma\,(A)$ si y sólo si $A - \lambda I \in \mathcal{G}l(n)$.
4. Sea $P_A(x) \in \mathbb{C}[x]$ el polinomio característico de A. Luego $\lambda \in \sigma(A)$ si y sólo si $P_A(\lambda) = 0$, o sea que $\sigma(A)$ es el conjunto de raíces de $P_A(x)$.
5. Como $\operatorname{gr}(P_A) = n$, se tiene que $0 < |\sigma(A)| \leq n$.
6. $\sigma(A^*) = \overline{\sigma(A)}$. En efecto, usando 1.1.8, tenemos que
$$A - \lambda I \notin \mathcal{G}l(n) \iff (A - \lambda I)^* = A^* - \overline{\lambda}\ I \notin \mathcal{G}l(n)\ .$$
7. Si $A \in \mathcal{G}l(n)$, entonces $\sigma\,(A^{-1}) = \sigma\,(A)^{-1} = \{\lambda^{-1} : \lambda \in \sigma(A)\}$. En efecto, es consecuencia de la igualdad $\ker(A - \lambda I) = \ker(A^{-1} - \lambda^{-1} I)$ (Ejercicio: probarla). ▲

Observación 1.2.3. Vimos que los autovalores de $A \in \mathcal{M}_n(\mathbb{C})$ son las raíces del polinomio característico $P_A(x) = \det(xI - A)$ y que $\operatorname{gr}(P_A) = n$. Pero P_A 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, factorizamos en $\mathbb{C}[x]$ a $P_A(x) = \prod_{i=1}^{n}(x-\lambda_i) = \prod_{i=1}^{n}(x-\lambda_i(A)\,)$, y diremos que

$$\text{“los autovalores de } A \text{ son } \lambda_1\,,\ldots,\lambda_n\text{”} \quad , \qquad \text{o} \qquad \text{“}\lambda(A) = (\lambda_1(A),\ldots,\lambda_n(A)\,)\text{”} \ ,$$

donde estaremos repitiendo cada autovalor de A tantas veces como multiplicidad tiene como raíz de P_A, y disponiéndolos en algún orden de $\mathbb{C}$ fijado previamente (por ejemplo, el lexicográfico en las coordenadas polares, con el cero al final). Por eso quedan n. Al vector $\lambda(A) \in \mathbb{C}^n$ se lo llama “vector de autovalores de A”. ▲

Observación 1.2.4. Sean $A, B \in \mathcal{M}_n(\mathbb{C})$ y $S \in \mathcal{G}l\,(n)$ tales que $B = SAS^{-1}$. Luego B difiere de A en un cambio de base. Se suele decir que A y B son *similares* y se nota $A \sim B$. Por las propiedades del determinante, se tiene que

$$P_B(x) = \det(xI - SAS^{-1}) = \det S(xI - A)S^{-1} = \det(xI - A) = P_A(x) \ ,$$

por lo que $\lambda(A) = \lambda(B)$ y también $\sigma(A) = \sigma(B)$. ▲

Definición 1.2.5. Sea $A \in \mathcal{M}_n(\mathbb{C})$.

1. El **radio numérico** de A se define como
$$w(A) = \text{máx}\{\ |\langle Ax, x\rangle| \ : \ x \in \mathbb{C}^n,\ \|x\| = 1\ \} \ .$$
2. El **radio espectral** de A se define como
$$\rho(A) = \text{máx}\{\ |\lambda| : \lambda \in \sigma\,(A)\} \ .$$
3. La **norma espectral** de A es su norma como operador, inducida por la norma euclídea de $\mathbb{C}^n$. Es decir,
$$\|A\|_{sp} = \text{máx}\{\|Ax\| : x \in \mathbb{C}^n, \|x\| = 1\} = \text{mín}\{C \geq 0 : \|Ax\| \leq C\|x\| \ , \quad x \in \mathbb{C}^n\} \ .$$
4. La **norma 2 o norma Frobenius** de A es su norma euclídea, si la pensamos como un vector largo. Por el Ejercicio 1.1.14, tenemos que
$$\|A\|_2^2 = \sum_{i,j=1}^{n} |a_{ij}|^2 = \operatorname{tr}(A^*A) \ . \tag{1.5}$$

Observar que, analizando los autovectores (unitarios) de A, se muestra fácilmente que

$$\rho(A) \leq w(A) \leq \|A\|_{sp} \quad . \tag{1.6}$$

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. ▲

1.3 Matrices unitarias

Recordemos que $U \in \mathcal{M}_n(\mathbb{C})$ es unitaria si $UU^* = U^*U = I$, y que $\mathcal{U}(n)$ denota al conjunto de matrices unitarias en $\mathcal{M}_n(\mathbb{C})$.

Teorema 1.3.1. *Si $U \in \mathcal{M}_n(\mathbb{C})$, las siguientes afirmaciones son equivalentes:*

1. $U \in \mathcal{U}(n)$.
2. $U^* \in \mathcal{U}(n)$.
3. $U \in \mathcal{G}l(n)$ *y* $U^{-1} = U^*$.
4. *U preserva el producto escalar, o sea que*
$$\langle Ux, Uy\rangle = \langle x, y\rangle \ , \qquad \text{para todo par } x, y \in \mathbb{C}^n \ .$$
5. *Si $\mathcal{B}$ es una BON de $\mathbb{C}^n$, entonces $U(\mathcal{B})$ también lo es.*
6. *Las columnas de U forman una BON de $\mathbb{C}^n$.*
7. *Las filas de U forman una BON de $\mathbb{C}^n$.*
8. *Para todo $x \in \mathbb{C}^n$, se tiene que $\|Ux\| = \|x\|$ (o sea que U es una isometría).*

Además, $\mathcal{U}(n)$ es un **grupo** *con la multiplicación de matrices.*

Demostración. Ejercicio. Se sugiere usar el Ejercicio 1.1.14 y, para probar que 8 implica todo lo demás, usar la Proposición 1.1.11. ∎

Definición 1.3.2. Dadas $A, B \in \mathcal{M}_n(\mathbb{C})$, se dice que A es unitariamente equivalente a B y se nota
$$A \cong B \quad \text{si existe} \quad U \in \mathcal{U}(n) \quad \text{tal que} \quad A = U^*BU \ .$$
Observar que, como $\mathcal{U}(n)$ es un grupo, se tiene que $\cong$ es una relación de equivalencia. ▲

Teorema 1.3.3. *Sean A y $B \in \mathcal{M}_n(\mathbb{C})$ tales que $A \cong B$. Entonces*
$$\lambda(A) = \lambda(B) \quad , \quad \|A\|_2 = \|B\|_2 \quad , \quad \|A\|_{sp} = \|B\|_{sp} \quad y \quad A \text{ es } \star \ \text{ si y sólo si } B \text{ es } \star \ ,$$
donde $\star$ puede significar: Hermitiana, anti-Hermitiana, normal, defînda positiva o unitaria.

Demostración. La primera igualdad se sigue de la Observación 1.2.4. Sea $U \in \mathcal{U}(n)$ tal que $B = UAU^*$. Entonces, por la Eq. (1.5),
$$\|B\|_2^2 = \operatorname{tr}(B^*B) = \operatorname{tr}(UA^*U^*\ UAU^*) = \operatorname{tr}(UA^*AU^*) = \operatorname{tr}(A^*A) = \|A\|_2^2 \ ,$$
donde la penúltima igualdad se deduce del hecho de que $\operatorname{tr}(XY) = \operatorname{tr}(YX)$ para todo par $X, Y \in \mathcal{M}_n(\mathbb{C})$. Con respecto a las normas espectrales,
$$\|B\|_{sp} = \max_{\|x\|=1} \|UA(U^*x)\| = \max_{\|y\|=1} \|Ay\| = \|A\|_{sp} \ \ ,$$
ya que $U^*\big(\{x \in \mathbb{C}^n : \|x\| = 1\}\big) = \{y \in \mathbb{C}^n : \|y\| = 1\}$, porque U y U^* son isometrías sobreyectivas. Las afirmaciones sobre $\star$ se prueban directamente de las definiciones, porque $B^* = (UAU^*)^* = UA^*U^*$, con la misma $U \in \mathcal{U}(n)$. ∎

1.4 Matrices triangulares

Definición 1.4.1. Sea $T \in \mathcal{M}_n(\mathbb{C})$. Diremos que

1. T es *triangular superior* (abreviamos TS) si verifica que $T_{ij} = 0$ para $i > j$. Es decir que T tiene ceros por debajo de su diagonal.
2. T es *estrictamente* TS si $T_{ij} = 0$ para $i \geq j$. Acá también $\mathrm{d}(T) = 0$.
3. Análogamente se definen las matrices triangulares inferiores y estrictamente triangulares inferiores.
4. Denotaremos por $\mathcal{TS}(n) = \{\, T \in \mathcal{M}_n(\mathbb{C}) : T$ es triangular superior $\}$. ▲

1.4.2 (Propiedades de las matrices triangulares)**.** Tenemos las siguientes propiedades (enumeraremos los resultados, y las pruebas no escritas quedarán como ejercicio para el lector):

1. Sea $\mathcal{E} = \{e_1, \ldots, e_n\}$ la base canónica de $\mathbb{C}^n$. Notemos $\mathbb{H}_k = \mathrm{Gen}\,\{e_1, \ldots, e_k\}$, para cada $k \in \mathbb{I}_n$, y $\mathbb{H}_0 = \{0\}$. Dada $T \in \mathcal{M}_n(\mathbb{C})$ se tiene que
$$T \in \mathcal{TS}(n) \iff T(\mathbb{H}_k) \subseteq \mathbb{H}_k\,, \quad \text{para todo } \ k \in \mathbb{I}_n\,, \tag{1.7}$$
y T es estrictamente TS $\iff T(\mathbb{H}_k) \subseteq \mathbb{H}_{k-1}$, para todo $k \in \mathbb{I}_n$.
2. Usando la Eq. (1.7) sale fácil que $\mathcal{TS}(n)$ es un subanillo de $\mathcal{M}_n(\mathbb{C})$. Es decir que
$$T_1\,,\ T_2 \in \mathcal{TS}(n) \implies T_1 + T_2 \ \text{ y } \ T_1 T_2 \in \mathcal{TS}(n)\,. \tag{1.8}$$
También son subanillos los otros conjuntos de matrices triangulares, pero las estrictamente triangulares no tienen uno.
3. Además, dadas $S,\ T \in \mathcal{TS}(n)$, se tiene que
$$\mathrm{d}\,(ST) = \mathrm{d}\,(S) \cdot \mathrm{d}\,(T) := (S_{11} \cdot T_{11}\,, \ldots, S_{nn} \cdot T_{nn})\,. \tag{1.9}$$
Esto se muestra directamente calculando $(ST)_{ii}$ por la fórmula (1.1).
4. Si $T \in \mathcal{M}_n(\mathbb{C})$ es estrictamente triangular, entonces $T^n = 0$.
5. Si $T \in \mathcal{TS}(n)$, entonces su polinomio característico cumple que
$$P_T(x) = \det\,(xI - T) = \prod_{i=1}^{n}\,(x - T_{ii}) \implies \lambda(T) = \mathrm{d}(T) \text{ (salvo el orden)}\,. \tag{1.10}$$
Lo mismo pasa para su transpuesta T^T, que es una trianagular inferior genérica. La prueba es por inducción, desarrollando el determinante por la primera columna.
6. Si $T \in \mathcal{TS}(n)$, entonces $\mathrm{tr}\, T = \sum\limits_{i=1}^{n} T_{ii}$ y $\det T = \prod\limits_{i=1}^{n} T_{ii}$. Para la parte del determinante, se sugiere mirar la prueba de la Proposición 1.5.4.

7. Si $T \in \mathcal{TS}(n)$ es inversible, entonces
$$T^{-1} \in \mathcal{TS}(n) \quad \text{y} \quad \mathrm{d}\left(T^{-1}\right) = \mathrm{d}\,(T)^{-1} := (\mathrm{d}_1(T)^{-1}\,, \dots, \mathrm{d}_n(T)^{-1}\,)\ . \tag{1.11}$$
En efecto, por la Eq. (1.7) y el hecho de que $T \in \mathcal{G}l\,(n)$, sabemos que
$$T(\mathbb{H}_k) = \mathbb{H}_k \implies T^{-1}(\mathbb{H}_k) = \mathbb{H}_k\ , \quad \text{para todo} \ \ k \in \mathbb{I}_n \implies T^{-1} \in \mathcal{TS}(n)\ .$$
La igualdad $\mathrm{d}\,(T^{-1}) = \mathrm{d}\,(T)^{-1}$ se deduce ahora de la Eq. (1.9). ▲

1.5 Herramientas para operar con matrices

En esta sección veremos varias estrategias para operar con matrices:

1.5.1. Sean $A \in \mathcal{M}_{n,m}(\mathbb{C})$ y $B \in \mathcal{M}_{m,r}(\mathbb{C})$. Enumeraremos los resultados, y las pruebas no escritas quedarán como ejercicio para el lector.

1. La entrada $(AB)_{ij}$ es el producto de la $F_i(A) \in \mathcal{M}_{1,m}(\mathbb{C})$, por la $C_j(B) \in \mathcal{M}_{m,1}(\mathbb{C})$.
2. Si $x = (x_1\,, \dots, x_m) \in \mathbb{C}^m$, entonces $Ax = A\Big(\sum\limits_{i=1}^{m} x_i\, e_i\Big) = \sum\limits_{i=1}^{m} x_i\ C_i(A)$.
3. El producto $AB \in \mathcal{M}_{n,r}(\mathbb{C})$ representa la composición $A \circ B : \mathbb{C}^r \to \mathbb{C}^n$, cuando se las piensa como operadores. Sus columnas se describen por la acción de A en las columnas de B:
$$C_i(AB) = A \cdot\ C_i(B)\ \in \mathbb{C}^n\ , \qquad \text{para todo } i \in \mathbb{I}_r\ . \tag{1.12}$$
Esto se puede probar directamente por la fórmula (1.1) del producto, o bien observando que $C_i(AB) = (AB)e_i^{(r)} = A \cdot (\ Be_i^{(r)}) = A \cdot\ C_i(B)$, lo que se sigue de pensar a AB como una composición.
4. Análogamente puede verse que
$$F_i(AB) = F_i(A) \cdot\ B\ \in \mathbb{C}^r\ , \qquad \text{para todo } i \in \mathbb{I}_n\ . \tag{1.13}$$
5. Si alguna $C_i(B) = 0$, entonces $C_i(AB) = 0$ (con el mismo i). Análogamente, si se tiene que $F_i(A) = 0$, entonces $F_i(AB) = 0$. Esto se usa para identificar los ideales (a izq. y der.) del anillo $\mathcal{M}_n(\mathbb{C})$, cuando $n = m$.
6. Fijemos una columna $C_i(A)$. Alguna entrada de $C_i(A)$ aparece al calcular cualquier entrada de AB. Pero siempre multiplicada por alguna entrada de la $F_i(B)$ (recordar la fórmula (1.1): $(AB)_{st} = \sum\limits_{i=1}^{m} A_{si}\ B_{it}$).
7. Por lo tanto, si $F_i(B) = 0$, entonces podemos cambiar la $C_i(A)$ sin que ello afecte al producto AB. Análogamente, si $C_i(A) = 0$, se puede cambiar impunemente $F_i(B)$.
8. Sean $\lambda \in \mathbb{C}^n$ y $\mu \in \mathbb{C}^m$. Si $D_1 = \mathrm{diag}\,(\lambda) \in \mathcal{M}_n(\mathbb{C})$ y $D_2 = \mathrm{diag}\,(\mu) \in \mathcal{M}_m(\mathbb{C})$, entonces
$$D_1A = \big(\lambda_i\ a_{ij}\big)_{\substack{i\in\mathbb{I}_n\\ j\in\mathbb{I}_m}}\ , \quad AD_2 = \big(a_{ij}\ \mu_j\big)_{\substack{i\in\mathbb{I}_n\\ j\in\mathbb{I}_m}} \quad \text{y} \quad D_1AD_2 = \big(\lambda_i\ \mu_j\ a_{ij}\big)_{\substack{i\in\mathbb{I}_n\\ j\in\mathbb{I}_m}}\ . \tag{1.14}$$

Bloques

1.5.2. Sea $k \in \mathbb{I}_n$, tal que $0 < k < n$ y llamemos $I = \mathbb{I}_k$ y $J = \mathbb{I}_n \setminus \mathbb{I}_k = \{k+1, \dots, n\}$. Dada $A \in \mathcal{M}_n(\mathbb{C})$ notaremos su representación en bloques como sigue:

$$A = \begin{bmatrix} A_I & A_{IJ} \\ A_{JI} & A_J \end{bmatrix} \begin{array}{l} \mathbb{C}^k \\ \mathbb{C}^{n-k} \end{array} , \qquad \text{donde} \qquad A_{IJ} = (A_{rl})_{\substack{r \in I \\ l \in J}} \in \mathcal{M}_{k,\, n-k}(\mathbb{C}) \ ,$$

y en forma similar se definen $A_I \in \mathcal{M}_k(\mathbb{C})$, $A_{JI} \in \mathcal{M}_{n-k,\, k}(\mathbb{C})$ y $A_J \in \mathcal{M}_{n-k}(\mathbb{C})$. Más adelante (en la sección 2.4 y, sobre todo en los Capítulos 12 y 13) se usará una notación más detallada, tipo $A_{IJ} = A[I|J] = A[I|I)$ y así. Para fijar ideas observemos que, por ejemplo,

$$A \in \mathcal{TS}(n) \iff A_I \in \mathcal{TS}(k) \ , \ \ A_J \in \mathcal{TS}(n-k) \quad \text{y} \quad A_{JI} = 0. \tag{1.15}$$

Es extremadamente útil el hecho de que esta notación es consistente con las operaciones de matrices. Por ejemplo: Si $B \in \mathcal{M}_n(\mathbb{C})$:

1. $A + B = \begin{bmatrix} A_I + B_I & A_{IJ} + B_{IJ} \\ A_{JI} + B_{JI} & A_J + B_J \end{bmatrix} \begin{array}{l} \mathbb{C}^k \\ \mathbb{C}^{n-k} \end{array}$,

2. $A^* = \begin{bmatrix} A_I^* & A_{JI}^* \\ A_{IJ}^* & A_J^* \end{bmatrix} \begin{array}{l} \mathbb{C}^k \\ \mathbb{C}^{n-k} \end{array}$, una especie de transpuesta adjuntada. Luego

$$A = A^* \iff A_I \in \mathcal{H}(k) \ , \ \ A_J \in \mathcal{H}(n-k) \quad \text{y} \quad A_{JI} = A_{IJ}^* \ .$$

3. La más importante es la fórmula del producto:

$$AB = \begin{bmatrix} A_I B_I + A_{IJ} B_{JI} & A_I B_{IJ} + A_{IJ} B_J \\ A_{JI} B_I + A_J B_{JI} & A_{JI} B_{IJ} + A_J B_J \end{bmatrix} \begin{array}{l} \mathbb{C}^k \\ \mathbb{C}^{n-k} \end{array} , \tag{1.16}$$

que reproduce la fórmula del producto de matrices en $\mathcal{M}_2(\mathbb{C})$. Observar que los tamaños de todos los productos que aparecen son los adecuados al bloque donde viven. La prueba de esta fórmula es straightforward. Hay que usar la fórmula (1.1) para $(AB)_{ij}$, dividiendo cada suma en sus primeros k sumandos, y en los otros $n-k$.

4. Por ejemplo, cualquiera sea $C \in \mathcal{M}_{k,\, n-k}(\mathbb{C})$, se tiene que

$$A = \begin{bmatrix} I_k & C \\ 0 & I_{n-k} \end{bmatrix} \in \mathcal{G}l(n) \quad \text{y, además} \ , \quad A^{-1} = \begin{bmatrix} I_k & -C \\ 0 & I_{n-k} \end{bmatrix} \in \mathcal{G}l(n) \ . \tag{1.17}$$

Análogamente, si $X \in \mathcal{M}_{n-k,\, k}(\mathbb{C})$, luego $\begin{bmatrix} I_k & 0 \\ X & I_{n-k} \end{bmatrix}^{-1} = \begin{bmatrix} I_k & 0 \\ -X & I_{n-k} \end{bmatrix}$.

5. Otra aplicación: Si $U \in \mathcal{U}(k)$ y $V \in \mathcal{U}(n-k)$, entonces

$$W = \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \in \mathcal{U}(n) \quad \text{y} \quad W\,A\,W^* = \begin{bmatrix} U\,A_I\,U^* & U\,A_{IJ}\,V^* \\ V\,A_{JI}\,U^* & V\,A_J\,V^* \end{bmatrix} \ . \tag{1.18}$$

6. Si uno parte a $\mathbb{I}_n$ en r mitades (con $r \geq 3$), para cada $A \in \mathcal{M}_n(\mathbb{C})$ se pueden definir sus $r \times r$ bloques relativos a esa partición. Y valen propiedades análogas a las del caso 2×2. En particular, vale algo similar a la fórmula (1.16), pero imitando el producto en $\mathcal{M}_r(\mathbb{C})$. ▲

Observación 1.5.3. Sea $\mathcal{S} \subseteq \mathbb{C}^n$ un subespaceio con $\dim \mathcal{S} = k$. Todo lo que se hizo recién se puede generalizar exactamente igual a una representación de $\mathcal{M}_n(\mathbb{C})$ en matrices de 2×2 de bloques, pero teniendo en cuenta la descomposición $\mathbb{C}^n = \mathcal{S} \oplus \mathcal{S}^\perp$. El caso anterior correspondería a tomar $\mathcal{S} = \text{Gen}\{e_1, \dots, e_k\}$. La manera más económica de verlo es tomar una BON $\{v_1, \dots, v_n\} = \mathcal{B}$ de $\mathbb{C}^n$ tal que $\mathcal{S} = \text{Gen}\{v_1, \dots, v_k\}$ (por lo que $\mathcal{S}^\perp = \text{Gen}\{v_{k+1}, \dots, v_n\}$). Tomando coordenadas de las matrices en la base $\mathcal{B}$, el laburo anterior se extrapola a cualquier descomposición. Pedimos que $\mathcal{B}$ sea una BON y no una base cuaquiera (que empiece por una base de $\mathcal{S}$) para que valgan las fórmulas relativas a A^*, en particular (1.18). Las demás valen tomando coordenadas en cualquer base de aquel tipo. La notación que usaremos para estas representaciones es

$$A = \begin{array}{c} \begin{bmatrix} A_{\mathcal{S}} & A_{\mathcal{S},\mathcal{S}^\perp} \\ A_{\mathcal{S}^\perp,\,\mathcal{S}} & A_{\mathcal{S}^\perp} \end{bmatrix} \end{array} \begin{array}{c} \mathcal{S} \\ \mathcal{S}^\perp \end{array} .$$

Observar que, si $P_{\mathcal{S}}$ es el proyector ortogonal sobre $\mathcal{S}$, entonces $P_{\mathcal{S}} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{array}{c} \mathcal{S} \\ \mathcal{S}^\perp \end{array}$. Además, $A_{\mathcal{S}} = P_{\mathcal{S}} A P_{\mathcal{S}} = P_{\mathcal{S}} A\big|_{\mathcal{S}}$, o sea pensado en $L(\mathcal{S})$ (sin los tres ceros). Al operador $A_{\mathcal{S}} \in L(\mathcal{S})$ se lo llama **la compresión de A a $\mathcal{S}$**. Su matriz concreta (de $k \times k$) depende de la BON $\mathcal{B}$ elegida, pero en tanto operador en $L(\mathcal{S})$, nuestro $A_{\mathcal{S}}$ sólo depende del subsepacio $\mathcal{S}$. Fórmulas semejantes se tienen para los otros tres bloques de A. ▲

Proposición 1.5.4. *Sea $n \geq 2$, $A \in \mathcal{M}_n(\mathbb{C})$ y $\mathcal{S} \subseteq \mathbb{C}^n$ un subespacio propio. Si*

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} \begin{array}{c} \mathcal{S} \\ \mathcal{S}^\perp \end{array} \implies \det A = \det B \ \det D \quad y \quad P_A(x) = P_C(x) P_D(x) . \tag{1.19}$$

Por lo tanto $\lambda(A) = \big(\lambda(B), \lambda(D)\big)$. Una fórmula similar vale si A es triangular inferior de bloques (para $\mathcal{S}$).

Demostración. Eligiendo bases de $\mathcal{S}$ y $\mathcal{S}^\perp$, y usando que la similaridad no cambia ni el det ni el característico, podemos asumir que $\mathcal{S} = \text{Gen}\{e_1, \dots, e_k\}$, donde $k = \dim \mathcal{S}$. Más aún, ya que elegimos una BON cualquiera de $\mathcal{S}$, podemos suponer su primer elemento era un autovector $x_1 \in \mathcal{S}$ de B para cierto $\lambda \in \sigma(B)$. Observar que entonces también se tiene que $Ax_1 = \lambda x_1$. Al tomar matrices, queda que

$$A = \begin{bmatrix} \lambda & \star \\ 0 & A_1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} \lambda & * \\ 0 & B_1 \end{bmatrix} \quad \text{con} \quad A_1 = \begin{bmatrix} B_1 & C_1 \\ 0 & D \end{bmatrix} \in \mathcal{M}_{n-1}(\mathbb{C}) .$$

Ojo que si k era uno, queda que $A_1 = D$ y que $B = [\lambda]$ (en ese caso $*$, B_1 y C_1 no existen). Si $k > 1$, desarrollando por la primer comumna, queda que

$$\det A = \lambda \det A_1 \ , \ \det B = \lambda \det B_1 \ , \ P_A(x) = (x - \lambda)\, P_{A_1}(x) \ \text{ y } \ P_B(x) = (x - \lambda)\, P_{B_1}(x) .$$

Las dos últimas salen porque $xI_n - A = \begin{bmatrix} x-\lambda & -\star \\ 0 & xI_{n-1} - A_1 \end{bmatrix}$, y lo mismo para B. Haciendo ahora inducción en $n \geq 2$ (o en k, va por gustos), estamos hechos. Otra manera de probarlo es vía la definición con permutaciones de $\mathbf{S}_n$, porque las permutaciones que no *pasan* por el bloque nulo de abajo, son todas las del tipo $\sigma \in \mathbf{S}_k \times \mathbf{S}_{n-k}$. Este camino queda como ejercicio. ■

Como ejemplo del uso de estas técnicas, mostraremos a continuación la relación que hay entre el espectro del producto de dos matrices, en sus dos órdenes posibles. Se suguiere tratar de probar el siguiente enunciado directamente, para ver cuanto más fácil puede hacerse con la técnica de bloques, y la primera aparición del famoso truco de 2×2.

Proposición 1.5.5. *Dadas $A, B \in \mathcal{M}_n(\mathbb{C})$, entonces $\sigma(AB) = \sigma(BA)$. Más aún, AB y BA tienen el mismo polinomio característico, por lo que $\lambda(AB) = \lambda(BA) \in \mathbb{C}^n$.*

Demostración. Por la Eq. (1.17), sabemos que la matriz

$$M = \begin{bmatrix} I & A \\ 0 & I \end{bmatrix} \in \mathcal{G}l(2n) \ , \quad \text{y que} \quad M^{-1} = \begin{bmatrix} I & -A \\ 0 & I \end{bmatrix} .$$

Además, usando la ecuación (1.16), nos queda que

$$\begin{aligned} M^{-1} \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} M &= \begin{bmatrix} I & -A \\ 0 & I \end{bmatrix} \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \begin{bmatrix} I & A \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ B & 0 \end{bmatrix} \begin{bmatrix} I & A \\ 0 & I \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix} . \end{aligned}$$

Usando la Proposición 1.5.4, podemos deducir que $P_{AB} = P_{BA}$, porque si se tienen dos polinomios $P, Q \in \mathbb{C}[x]$ que cumplen $x^n P(x) = x^n Q(x)$, entonces $P = Q$. ■

Observación 1.5.6. Sean $A \in \mathcal{M}_{n,m}(\mathbb{C})$ y $B \in \mathcal{M}_{m,n}(\mathbb{C})$ con $m > n$. Con casi la misma prueba que la Proposición 1.5.5 puede mostrarse que $\sigma(BA) = \sigma(AB) \cup \{0\}$, puesto que sus polinomios característicos cumplen $P_{BA}(x) = x^{m-n} P_{AB}(x)$. ▲

1.6 El Teorema de Schur y sus corolarios

El siguiente resultado, el primero de los varios debidos a Schur que enunciaremos, es sumamente útil, y será usado sistemáticamente en todo este trabajo.

Teorema 1.6.1 (Schur 1). *Sea $A \in \mathcal{M}_n(\mathbb{C})$ con vector de autovalores $\lambda(A) = (\lambda_1, \dots, \lambda_n)$, dispuestos en cualquier orden prefijado. Entonces*

1. *Existen matrices $U \in \mathcal{U}(n)$ y $T \in \mathcal{TS}(n)$ que verifican:*

   a. $A = U\,T\,U^*$.

b. $d(T) = \lambda(A)$, *i.e.* $T_{ii} = \lambda_i$ *, para todo* $i \in \mathbb{I}_n$.

2. *Si* $A \in \mathcal{M}_n(\mathbb{R})$, *el teorema sigue valiendo (con* $U \in \mathcal{M}_n(\mathbb{R})$*) siempre que* $\sigma(A) \subseteq \mathbb{R}$.

3. *Si* $B \in \mathcal{M}_n(\mathbb{C})$ *conmuta con* A, *existen* $U \in \mathcal{U}(n)$, *y* T_1 , $T_2 \in \mathcal{TS}(n)$ *tales que*
$$A = U\,T_1\,U^* \ , \quad B = U\,T_2\,U^* \quad (\textit{con } \textbf{la misma } U) \quad y \quad d(T_1) = \lambda(A) \ .$$
La $d(T_2)$ *tendrá a los autovalores de* B, *pero en un orden que no podremos elegir.*

Demostración. La prueba la realizaremos por inducción sobre la dimensión n. Si $n = 1$, el resultado es trivial. Si $n > 1$, tomemos $x_1 \in \ker(A - \lambda_1 I)$ con $\|x_1\| = 1$. Completamos a una BON de $\mathbb{C}^n$ con vectores $x_2, \dots, x_n$, y los ponemos en las columnas de una matriz U_1 . Por el Teorema 1.3.1, $U_1 \in \mathcal{U}(n)$. Como $U_1(e_1) = x_1$ y $U_1^*(x_1) = e_1$, es fácil ver que
$$C_1(U_1^* A U_1) = U_1^* A U_1 \ e_1 = \lambda_1 \, e_1 \quad \Longrightarrow \quad U_1^* A U_1 = \begin{array}{c} \left[\begin{array}{cc} \lambda_1 & * \\ 0 & A_2 \end{array}\right] \end{array} \begin{array}{l} \mathbb{C} \\ \mathbb{C}^{n-1} \end{array} \ ,$$
donde $A_2 \in \mathcal{M}_{n-1}(\mathbb{C})$. Por la Observación 1.2.4, sus polinomios caractrísticos cumplen
$$P_A(x) = P_{U_1^* A U_1}(x) = (x - \lambda_1) P_{A_2}(x) \quad \Longrightarrow \lambda(A_2) = (\lambda_2, \dots, \lambda_n) \in \mathbb{C}^{n-1} \ .$$
Por HI, existen $V \in \mathcal{U}(n-1)$ y $T_2 \in \mathcal{TS}(n-1)$ tales que $V^* A_2 V = T_2$ y $\mathrm{d}(T_2) = \lambda(A_2)$. Podemos extender V a otra matriz $U_2 = \begin{bmatrix} 1 & 0 \\ 0 & V \end{bmatrix} \in \mathcal{U}(n)$. Sea $U = U_1 U_2$. Entonces, usando las ecuaciones (1.18) y (1.15) sobre productos de matrices de bloques, nos queda que
$$\begin{aligned} U^* A U = U_2^* \,(U_1^* \, A \, U_1) \, U_2 \quad &= \begin{bmatrix} 1 & 0 \\ 0 & V^* \end{bmatrix} \begin{bmatrix} \lambda_1 & * \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & V \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & \star \\ 0 & V^* A_2 V \end{bmatrix} = \begin{bmatrix} \lambda_1 & \star \\ 0 & T_2 \end{bmatrix} = T \in \mathcal{TS}(n) \ , \end{aligned}$$
y se tiene que $\mathrm{d}(T) = (\lambda_1 , \mathrm{d}(T_2)\,) = (\lambda_1 , \lambda(A_2)\,) = \lambda(A)$.

El caso real sale igual. Notar que se puede elegir $x_1 \in \mathbb{R}^n$ siempre que $\lambda_1 \in \mathbb{R}$. El caso de dos matrices que conmutan, se deduce de que $\ker(A - \lambda_1 I)$ es invariante para B (cuenta fácil, ya que conmutan), por lo que el vector x_1 se puede elegir como un autovector de B actuando en $\ker(A - \lambda_1 I)$ (no se sabe cuales de los autovalores de B en $\mathbb{C}^n$ pueden elegirse ahí). El resto de la prueba sigue igual, usando que las matrices achicadas A_2 y B_2 siguen conmutando. ■

Corolario 1.6.2. *Sea* $A \in \mathcal{M}_n(\mathbb{C})$ *con vector de autovalores* $\lambda(A) = (\lambda_1, \dots, \lambda_n)$. *Entonces*
$$\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i \qquad y \qquad \det A = \prod_{i=1}^{n} \lambda_i \ .$$

Demostración. Por el Teorema 1.6.1 sabemos que podemos escribir $A = U^* T U$, donde se tiene que $U \in \mathcal{U}(n)$, $T \in \mathcal{TS}(n)$ y $\mathrm{d}(T) = \lambda(A)$. Luego $\operatorname{tr} A = \operatorname{tr} T$ y $\det A = \det T$. ■

Corolario 1.6.3. *Sea* $U \in \mathcal{U}(n)$. *Entonces* $|\det U| = 1$.

Demostración. Basta notar que $\sigma(U) \subseteq \{z \in \mathbb{C} : |z| = 1\}$, dado que U es una isometría. ■

Observación 1.6.4. Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. En general, no se tiene la menor idea de qué pueden ser los espectros $\sigma(A+B)$ y $\sigma(AB)$. Sin embargo, cuando A y B conmutan, el Teorema 1 de Schur nos da alguna información al respecto. El siguiente resultado vale también el álgebras de Banach de dimensión infinita, pero con una prueba mucho más sofisticada. ▲

Corolario 1.6.5. *Sean* $A, B \in \mathcal{M}_n(\mathbb{C})$*, tales que* $AB = BA$. *Entonces*

1. $\sigma(A+B) \subseteq \sigma(A) + \sigma(B) = \{\lambda + \mu : \lambda \in \sigma(A) \text{ y } \mu \in \sigma(B)\}$.
2. $\sigma(AB) \subseteq \sigma(A) \cdot \sigma(B) = \{\lambda \cdot \mu : \lambda \in \sigma(A) \text{ y } \mu \in \sigma(B)\}$.

Más aún, existen ciertas ordenaciones de los vectores de autovalores $\lambda(A)$ *y* $\lambda(B)$ *tales que (operando en esos órdenes),* $\lambda(A+B) = \lambda(A) + \lambda(B)$ *y*

$$\lambda(AB) = \lambda(A) \cdot \lambda(B) = (\lambda_1(A)\lambda_1(B), \dots, \lambda_n(A)\lambda_n(B)) \ . \tag{1.20}$$

Demostración. Probaremos solamente la igualdad (1.20). Las cuentas para $\lambda(A+B)$ son iguales (y más fáciles). Por el Teorema 1 de Schur 1.6.1, existen $U \in \mathcal{U}(n)$, y $T_1 \,,\ T_2 \in \mathcal{TS}(n)$ tales que $A = U T_1 U^*$, $B = U T_2 U^*$, $\mathrm{d}(T_1) = \lambda(A)$ y $\mathrm{d}(T_2) = \lambda(B)$, aunque los órdenes en que aparecen $\lambda(A)$ y $\lambda(B)$ no lo sabemos. Pero en esos órdenes, tenemos que

$$T_1 \cdot T_2 \in \mathcal{TS}(n) \implies \lambda(T_1 T_2) = \mathrm{d}(T_1 T_2) = \mathrm{d}(T_1) \cdot \mathrm{d}(T_2) = \lambda(A) \cdot \lambda(B) \ ,$$

por las fórmulas (1.8), (1.10) y (1.9). Pero $AB = (U T_1 U^*)(U T_2 U^*) = U(T_1 T_2)U^*$. Luego, por el Teorema 1.3.3, y en el orden que había, $\lambda(T_1 T_2) = \lambda(AB)$. ■

Corolario 1.6.6. *Sea* $A \in \mathcal{M}_n(\mathbb{C})$ *con vector de autovalores* $\lambda(A) = (\lambda_1, \dots, \lambda_n)$. *Entonces*

$$\lambda(A^*) = \overline{\lambda(A)} = (\overline{\lambda_1}, \dots, \overline{\lambda_n}) \ .$$

Esto generaliza la igualdad $\sigma(A^*) = \overline{\sigma(A)}$ *ya vista.*

Demostración. Sean $U \in \mathcal{U}(n)$ y $T \in \mathcal{TS}(n)$, con $\mathrm{d}(T) = \lambda(A)$, tales que $A = U^* T U$. Luego $A^* = U^* T^* U$, por lo que $\lambda(A^*) = \lambda(T^*)$. Pero T^* es triangular inferior, así que también se tiene que $\lambda(T^*) = \mathrm{d}(T^*) = \overline{\mathrm{d}(T)} = \overline{\lambda(A)}$. ■

1.7 Polinomios y matrices

Observación 1.7.1. Si $A \in \mathcal{M}_n(\mathbb{C})$ y $P(x) = \sum\limits_{k=0}^{m} b_k\, x^k \in \mathbb{C}[x]$, entonces P se puede evaluar en A de la siguiente manera:

$$P(A) = \sum_{k=0}^{m} b_k\, A^k \in \mathcal{M}_n(\mathbb{C}) \ ,$$

dado que las potencias (enteras) A^k se definen con el producto de matrices, y viven en $\mathcal{M}_n(\mathbb{C})$. Además, se tiene las siguientes propiedades:

1. Como las potencias de A conmutan entre sí, se deduce fácilmente que la aplicación $E_A : \mathbb{C}[x] \to \mathcal{M}_n(\mathbb{C})$ dada por $P \mapsto P(A)$ (o sea la evaluación en A) es un morfismo de anillos. Por lo tanto, si se tiene una factorización $P = QR$, con $Q, R \in \mathbb{C}[x]$, entonces $P(A) = Q(A)R(A)$, ahora con el producto de matrices.

2. Si $S \in \mathcal{G}l(n)$, entonces $(SAS^{-1})^k = SA^kS^{-1}$, para todo $k \in \mathbb{N}$. Luego, es fácil ver que

$$P(SAS^{-1}) = S \cdot P(A) \cdot S^{-1} \ , \quad \text{para todo } P \in \mathbb{C}[x] \ . \tag{1.21}$$

3. Si $T \in \mathcal{TS}(n)$, hemos visto que $T^2 \in \mathcal{TS}(n)$ y que $\mathrm{d}\,(T^2) = (T_{11}^2, \dots, T_{nn}^2) = \mathrm{d}\,(T)^2$. Esto se extinede a potencias enteras, por inducción. Por lo tanto,

$$P(T) \in \mathcal{TS}(n) \ \text{ y } \ P(T)_{ii} = P(T_{ii}) \text{ para todo } \ i \in \mathbb{I}_n \ , \tag{1.22}$$

o sea que $\mathrm{d}\,(P(T)\,) = P(\mathrm{d}\,(T)\,) = P(\mathrm{d}\,(T)\,) = (P(T_{11})\,, \dots, P(T_{nn})\,)$. ▲

Corolario 1.7.2. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ con vector de autovalores $\lambda(A) = (\lambda_1\,, \dots, \lambda_n)$. Entonces,*

$$\lambda(P(A)\,) = P(\lambda(A)\,) := (P(\lambda_1)\,, \dots, P(\lambda_n)\,) \quad \textit{para todo } P \in \mathbb{C}[x] \ .$$

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

Demostración. Supongamos que $T \in \mathcal{TS}(n)$. Recordemos de la Eq. (1.10) que $\mathrm{d}\,(T) = \lambda(T)$. Por la Eq. (1.22), sabemos que $P(T) \in \mathcal{TS}(n)$ y $\mathrm{d}\,(P(T)\,) = P(\mathrm{d}\,(T)\,)$. Luego,

$$\lambda(P(T)\,) \stackrel{(1.10)}{=} \mathrm{d}(P(T)\,) = P(\mathrm{d}\,(T)\,) = P(\lambda(T)\,) \ ,$$

lo que prueba el Corolario en este caso. Si $A \in \mathcal{M}_n(\mathbb{C})$ es cualquier matriz, sean

$$U \in \mathcal{U}(n) \ \text{ y } \ T \in \mathcal{TS}(n) \quad \text{tales que} \quad A = U^*TU \ \text{ y } \ \lambda(A) = \mathrm{d}(T) = \lambda(T) \ .$$

Por la Eq. (1.21), tenemos que $P(A) = U^*P(T)U$. Luego, por la Observación 1.2.4 (que decía que cambiar de base no cambia el vector λ) y el caso anterior, sabemos que

$$\lambda(P(A)\,) = \lambda(P(T)\,) = P(\lambda(T)\,) = P(\lambda(A)\,) \ .$$

■

Se sugiere otra manera de hacerlo, aplicando cuentas de polinomios. Por ejemplo, factorizar el polinomio $Q(x) = P(x) - \mu$, para $\mu \in P(\sigma\,(A)\,)$ o $\mu \in \sigma\,(P(A)\,)$, y analizar qué pasa con $Q(A)$. Esta prueba es la que sirve en dimensión infinita. Pero tiene el defecto que no da información sobre multiplicidades.

Corolario 1.7.3 (Hamilton-Cayley)**.** *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Luego $P_A(A)$ es la matriz nula.*

Demostración. Por el Teorema 1.6.1, la Eq. (1.21) y la Observación 1.2.4, sabemos que existen $U \in \mathcal{U}(n)$ y $T \in \mathcal{TS}(n)$ tales que $UTU^* = A$, $P_T(x) = P_A(x)$ y $P_A(A) = UP_T(T)U^*$. Luego basta probar que $P_T(T) = 0$ para matrices $T \in \mathcal{TS}(n)$. En este caso, por la Eq. (1.10), sabemos que

$$P_T(x) = \prod_{i=1}^{n} (x - T_{ii}) \implies P_T(T) = \prod_{i=1}^{n} (T - T_{ii}\ I) .$$

Llamemos $T_i = (T - T_{ii}\ I)$ y $\mathbb{H}_i = \text{Gen}\{e_1, \dots, e_i\}$, para cada $i \in \mathbb{I}_n$. Todas las T_i estan en $\mathcal{TS}(n)$, comnutan entre sí, y cumplen que $(T_i)_{ii} = 0$. Luego, si $\mathbb{H}_0 = \{0\}$, se tiene que

$$T_i(\mathbb{H}_i) = T_i\big(\mathbb{H}_{i-1} \oplus \text{Gen}\{e_i\}\big) = T_i\big(\mathbb{H}_{i-1}\big) + \text{Gen}\{Te_i\} \subseteq \mathbb{H}_{i-1} \quad \text{para todo} \quad i \in \mathbb{I}_n \ .$$

Si al producto de ellas $P_T(T) = \prod\limits_{i \in \mathbb{I}_n} T_i$ lo ordenamos así: $T_1 T_2 \dots T_n$ (total las T_i conmutan), vemos inmediatamente que

$$\Big[\prod_{i \in \mathbb{I}_n} T_i\Big](\mathbb{C}^n) = \Big[\prod_{i \in \mathbb{I}_n} T_i\Big](\mathbb{H}_n) \subseteq \Big[\prod_{i \in \mathbb{I}_{n-1}} T_i\Big](H_{n-1}) \subseteq \dots \subseteq T_1 T_2(\mathbb{H}_2) \subseteq T_1(\mathbb{H}_1) = \{0\} .$$

O sea que $P_T(T) = 0$. ■

Observación 1.7.4. El Teorema de Hamilton Cayley vale para matrices en cualquier cuerpo. La prueba anterior es simple, pero para generalizarla hace falta subir a una clausura algebráica (o aun cuerpo de descomposición de $P_A(x)$), porque necesita del Teorema 1 de Schur, que solo vale para cuerpos algebráicamente cerrados (se necesitan los $\lambda_i(A)$, que son las raices que factorizan a $P_A(x)$). A nadie se le ocurra postular la siguiente prueba general:

$$P_A(A) = \det(A \cdot I - A) = \det 0 = 0,$$

porque es desastrosamente errónea. En palabras del maestro Enzo Gentile, "es peor que pegarle una patada a una vieja en la cara" (sic). ▲

1.8 QR

Otra aplicación importante de las matrices triangulares es el denominado método QR.

1.8.1. Repaso: Recordemos lo visto en los ítems 6 y 7 de 1.5.1. Sea $A \in \mathcal{M}_n(\mathbb{C})$ y fijemos una columna $C_i(A)$. Observar que alguna entrada de $C_i(A)$ aparece al calcular cualquier entrada de AB. Pero siempre multiplicada por alguna entrada de la $F_i(B)$ (recordar que $(AB)_{st} = \sum\limits_{i=1}^{m} A_{si}\ B_{it}$). Por lo tanto, si $F_i(B) = 0$, entonces podemos cambiar a piacere la $C_i(A)$ sin que ello afecte al producto AB. Análogamente, si $C_i(A) = 0$, se puede cambiar impunemente a $F_i(B)$. ▲

Teorema 1.8.2. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces existen $Q \in \mathcal{U}(n)$ y $R \in \mathcal{TS}(n)$ tales que*

a. $A = QR$.

b. $R_{jj} \geq 0$, *para todo* $j \in \mathbb{I}_n$.

Si $A \in \mathcal{G}l(n)$, *entonces tales* Q *y* R *son únicas.*

Demostración. **Caso** $A \in \mathcal{G}l(n)$**:** Por el método de Gramm-Schmidt, si notamos $x_k = C_k(A)$, existe una BON $\mathcal{B} = \{u_1, \dots, u_n\}$ tal que $\text{Gen}\{x_1, \dots, x_k\} = \text{Gen}\{u_1, \dots, u_k\}$, para todo $k \in \mathbb{I}_n$. Además, por la construcción de $\mathcal{B}$ por Gramm-Schmidt,

$$u_j = \frac{x_j - \sum\limits_{i=1}^{j-1} \langle x_j, u_i \rangle \ u_i}{\|x_j - \sum\limits_{i=1}^{j-1} \langle x_j, u_i \rangle \ u_i\|} \implies x_j = \sum_{i=1}^{j} r_{ij} \ u_i \ , \ \text{con } r_{jj} = \| \ x_j - \sum_{i=1}^{j-1} \langle x_j, u_i \rangle \ u_i \ \| > 0 \ . \quad (1.23)$$

Tomemos $Q \in \mathcal{U}(n)$ con columnas $C_i(Q) = u_i$, $i \in \mathbb{I}_n$. Y tomemos $R \in \mathcal{TS}(n)$ dada por $R = (r_{ij})_{i,\, j \in \mathbb{I}_n}$, donde ponemos $r_{ij} = 0$ cuando $i > j$. Como se vió en 1.5.1, tenemos que

$$C_j(QR) = Q \cdot C_j(R) = \sum_{i=1}^{n} r_{ij} \ C_i(Q) = \sum_{i=1}^{j} r_{ij} \ u_i = x_j = C_j(A) \ , \qquad \text{para todo } j \in \mathbb{I}_n \ .$$

Por lo tanto $A = QR$.

Unicidad: Si hubiera otro par Q', R' cumpliendo las hipótesis, llamemos $C_i(Q') = v_i$ para cada $i \in \mathbb{I}_n$. Es fácil ver que $\text{Gen}\{x_1, \dots, x_k\} = \text{Gen}\{v_1, \dots, v_k\}$, $k \in \mathbb{I}_n$ (usar la Eq. (1.7)). De ahí se deduce que existen constantes c_i tales que $|c_i| = 1$ y $v_i = c_i u_i$ para todo $i \in \mathbb{I}_n$. Como $r'_{ii} > 0$, de la Eq. (1.23) y del hecho de que $A = Q' \, R'$, se deduce que

$$x_i = \sum_{s=1}^{i} r'_{si} \ v_s = \sum_{s=1}^{i} r'_{si} \ c_s u_s \implies r'_{ii} \ c_i = r_{ii} \implies c_i = \frac{r_{ii}}{r'_{ii}} > 0 \implies c_i = 1 \ ,$$

para todo $i \in \mathbb{I}_n$. Luego $Q = Q'$ y, por lo tanto $R' = Q^* A = R$.

Caso general: Si $A \notin \mathcal{G}l(n)$, el proceso es similar, salvo que, cada vez que aparece un

$$x_k \in \text{Gen}\{x_1, \dots, x_{k-1}\} \ ,$$

se pone $u_k = 0$ en la $C_k(Q)$, y, en la Eq. (1.23), ponemos $r_{kj} = 0$ para todo $j \in \mathbb{I}_n$, dado que el $u_k = 0$ no aporta para generar a los x_j $(j \geq k)$. Luego $F_k(R) = 0$.

Así queda que $R \in \mathcal{TS}(n)$, $A = QR$ y $r_{ii} \geq 0$ para todo $i \in \mathbb{I}_n$, pero $Q \notin \mathcal{U}(n)$. Esto se arregla de la siguiente mantera: se cambian las $C_k(Q) = u_k = 0$ del proceso anterior por una BON de $R(A)^\perp = \text{Gen}\{u_j : u_j \neq 0\}^\perp$ (observar que la cantidad es la correcta). Como se vió en el repaso 1.8.1 (o bien en 1.5.1), al multiplicar la Q cambiada por R, cada una de las nuevas $C_k(Q)$ sólo opera con la respectiva $F_k(R) = 0$. Luego sigue pasando que $A = QR$, pero ahora $Q \in \mathcal{U}(n)$. ■

Ejercicio 1.8.3. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Usando QR, probar que

$$|\det A| \ \leq \ \prod_{i=1}^{n} \| \ C_i(A) \ \|_2 \ , \qquad (1.24)$$

y que son iguales si y sólo si $A^* A$ es diagonal (o sea que R lo es). Se sugiere interpretarlo también como un cálculo de volúmenes. ▲

1.9 Matrices de rango uno

Recordemos que, si $A \in \mathcal{M}_{n,m}(\mathbb{C})$, notamos $\operatorname{rk}(A) = \dim R(A)$. A continuación daremos una caracterización muy util de las matrices con rango uno.

Definición 1.9.1. Dados $x \in \mathbb{C}^n$ e $y \in \mathbb{C}^m$ consideremos la matriz

$$x \odot y := xy^* = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \cdot [\, \overline{y_1}\,, \dots, \,\overline{y_m}\,] = (x_i\, \overline{y_j}\,)_{\substack{i \in \mathbb{I}_n \\ j \in \mathbb{I}_m}} \in \mathcal{M}_{n,m}(\mathbb{C}) \ . \tag{1.25}$$

Observar que $x \odot y$ actúa en $\mathbb{C}^m$ de la siguiente manera:

$$x \odot y(z) = (xy^*)\, z = x\,(y^*\, z) = \langle z, y \rangle\, x \qquad \text{para todo } z \in \mathbb{C}^m \ . \tag{1.26}$$

Por lo tanto, se tiene que $R(x \odot y) \subseteq \operatorname{Gen}\{x\}$, por lo que $\operatorname{rk}(x \odot y) \leq 1$. ▲

Por ejemplo, si $A \in \mathcal{M}_{n,m}(\mathbb{C})$ cumple que su única columna no nula es $C_k(A)$, entonces se ve fácilmente que $A = C_k(A) \odot e_k^{(m)}$, tanto por la Eq. (1.25) como por la Eq. (1.26). Observar que todo $\mathcal{M}_{n,m}(\mathbb{C})$ es el span de este tipo de matrices, porque se tiene la igualdad

$$A = \sum_{k=1}^{m} C_k(A) \odot e_k^{(m)} = \sum_{j=1}^{n} e_j^{(n)} \odot F_j(A) \quad , \quad \text{para toda } A \in \mathcal{M}_{n,m}(\mathbb{C}) \ . \tag{1.27}$$

La segunda igualdad se sigue de un argumento similar al anterior.

Proposición 1.9.2.

1. *Si $A \in \mathcal{M}_{n,m}(\mathbb{C})$ tiene $\operatorname{rk}(A) \leq 1$, existen $x \in \mathbb{C}^n$ e $y \in \mathbb{C}^m$ tales que $A = x \odot y$.*
2. *$\mathcal{M}_{n,m}(\mathbb{C}) = \operatorname{Gen}\{x \odot y : x \in \mathbb{C}^n \text{ e } y \in \mathbb{C}^m\}$.*

Demostración.

1. Sea $x \in \mathbb{C}^n$ tal que $R(A) = \operatorname{Gen}\{x\}$. Luego existe una funcional lineal

$$\varphi : \mathbb{C}^m \to \mathbb{C} \quad \text{ tal que } Az = \varphi(z) \cdot x \,, \text{ para todo } z \in \mathbb{C}^m \ .$$

Es sabido que existe un único $y \in \mathbb{C}^m$ tal que $\varphi(z) = \varphi_y(z) = \langle z, y \rangle$, para todo $z \in \mathbb{C}^m$ (basta poner $y_i = \overline{\varphi(e_i)}$, para cada $i \in \mathbb{I}_m$). Luego, por la Eq. (1.26), podemos concluir que $A = x \odot y$.

2. Se deduce de la Eq (1.27). ■

1.9.3. Estudiaremos a continuación las propiedades de las matrices $x \odot y$. Enumeraremos los resultados, y las pruebas no escritas quedarán como ejercicio para el lector. Tomemos dos vectores $x \in \mathbb{C}^n$ e $y \in \mathbb{C}^m$. Luego:

1. La norma espectral: $\|x \odot y\|_{sp} = \|x\| \; \|y\|$, ya que $\|\varphi_y\|_{sp} = \max\limits_{\|z\|=1} |\varphi_y(z)| = \|y\|$.
2. El adjunto: $(x \odot y)^* = (xy^*)^* = y\, x^* = y \odot x$.
3. Si $A \in \mathcal{M}_n(\mathbb{C})$, se tiene que $A \cdot (x \odot y) = A \cdot x \cdot y^* = (Ax) \odot y$.
4. Si $B \in \mathcal{M}_m(\mathbb{C})$, entonces $(x \odot y) \cdot B = x \odot (B^* y)$ (usar 2 y 3, o la Eq. (1.26)).
5. Dados $v \in \mathbb{C}^m$ y $w \in \mathbb{C}^k$, se tiene que $(x \odot y) \cdot (v \odot w) = \langle v, y\rangle \cdot x \odot w \in \mathcal{M}_{n,k}(\mathbb{C})$.

A partir de ahora supondremos que $n = m$, o sea que $x \odot y \in \mathcal{M}_n(\mathbb{C})$.

6. El espectro: El único autovalor de $x \odot y$ que puede ser no nulo es $\lambda_1 = \langle x, y\rangle$ (adivinen quién es el autovector). Más aún, $\lambda(x \odot y) = (\langle x, y\rangle, 0, \dots, 0)$. Para verlo basta tomar la matriz de $x \odot y$ en una base que empiece por x, y usar la Proposición 1.5.4.
7. Si $\|x\| = 1$, entonces $x \odot x = P_x$ es el proyector ortogonal sobre $\operatorname{Gen}\{x\}$. En efecto, observar que $x \odot x(z) = \langle z, x\rangle \cdot x$, la conocida fórmula de dicho proyector (se usa el hecho de que $z - \langle z, x\rangle \cdot x \in \{x\}^\perp$).
8. En general, si $x \neq 0$, el proyector $P_x = \dfrac{x}{\|x\|} \odot \dfrac{x}{\|x\|} = \dfrac{1}{\|x\|^2}\, x \odot x$.
9. Autoadjuntos: Se tiene que $A \in \mathcal{H}(n)$ si y sólo si A se descompone como una *suma algebráica* (i.e. con $\pm$ 1) de matrices $x_i \odot x_i$ (elegir los x_i entre los autovectores de A y esperar hasta el Teorema 2.2.1).
10. Positivos: $A \in \mathcal{M}_n(\mathbb{C})^+$ si y sólo si A se descompone como una **suma** de matrices $x_i \odot x_i$ (ver Proposición 3.5.6 de bastante más adelante). ▲

1.10 Ejercicios

Ejercicios que aparecen en el texto

1.10.1 (Polarización, Lema 1.1.10)**.**

Sea $\mathbb{H}$ un $\mathbb{C}$-espacio de Hilbert y sea $A \in L(\mathbb{H})$. Entonces

$$\langle Ax, y\rangle = \frac{1}{4} \sum_{k=1}^{4} i^k \left\langle A\,(x + i^k\, y), (x + i^k\, y) \right\rangle \quad \text{para todo par } x, y \in \mathbb{H} \; .$$

1.10.2. Demostrar los 7 items de 1.4.2 (sobre matrices triangulares).

1.10.3. Demostrar los 8 items de 1.5.1 y los 6 items de 1.5.2 (sobre matrices de bloques).

1.10.4. Demostrar la Proposición 1.5.4 usando la definición del determinante con permutaciones de $\mathbf{S}_n$. Usar que las permutaciones que no *pasan* por el bloque nulo de abajo, son todas las del tipo $\sigma \in \mathbf{S}_k \times \mathbf{S}_{n-k}$.

1.10.5. Demostrar los 10 items de 1.9.3 (sobre matrices tipo $x \odot y$).

1.10.6. Sean $A, \in \mathcal{M}_{m,n}(\mathbb{C})$. Entonces se tiene que

$$A^*A \in \mathcal{M}_n(\mathbb{C})^+ \quad \text{y} \quad (A^*A)_{i,j} = \langle\, C_j(A)\,,\, C_i(A)\,\rangle\,, \qquad \text{para todo par } i,j \in \mathbb{I}_n\ .$$

En particular, $\operatorname{tr}(A^*A) = \sum\limits_{j=1}^{n} \|C_j(A)\|^2 = \sum\limits_{i,j=1}^{n} |a_{ij}|^2$.

1.10.7. Sean $A, \in \mathcal{M}_{m,n}(\mathbb{C}) \simeq L(\mathbb{C}^m, \mathbb{C}^n)$. Mostrar que entonces

$$\ker A = \operatorname{Gen}\{F_1(A), \dots, F_n(A)\}^{\perp} \subseteq \mathbb{C}^n\ .$$

Deducir que $\operatorname{rk}_F(A) := \dim \operatorname{Gen}\{F_1(A), \dots, F_n(A)\} = \operatorname{rk}(A)$, o sea que los rangos fila y columna de A coinciden.

1.10.8. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Usando QR, probar que

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

y que son iguales si y sólo si A^*A es diagonal (o sea que R lo es). Se sugiere interpretarlo también como un cálculo de volúmenes.

Ejercicios nuevos

1.10.9. Mostrar que una matriz diagonalizable A satisface una ecuación polinomial de grado igual al $|\sigma(A)|$, y no menor.

1.10.10. Usar el Teorema 1 de Schur para probar que si, $A \in \mathcal{M}_n(\mathbb{C})$ tiene vector de autovalores $\lambda(A) = (\lambda_1\,, \dots, \lambda_n)$, entonces $\operatorname{tr} A^k = \sum\limits_{i=1}^{n} \lambda_i^k$, para todo $k \in \mathbb{N}$.

1.10.11. Deducir que, si $A, B \in \mathcal{M}_n(\mathbb{C})$ cumplen que $\operatorname{tr} A^k = \operatorname{tr} B^k$ para todo $k \in \mathbb{N}$, entonces $\lambda(A) = \lambda(B)$ (si usamos el mismo convenio para ordenarlos).

1.10.12. Dadas $A, B \in \mathcal{M}_n(\mathbb{C})$, notemos $C = AB - BA$. Probar que si C conmuta con A, entonces C es nilpotente.

1.10.13 (Triangulares)**.** Si $T \in \mathcal{TS}(n)$ es inversible, probar que

$$T^{-1} \in \mathcal{TS}(n) \quad \text{y} \quad \operatorname{d}\left(T^{-1}\right) = \operatorname{d}(T)^{-1} = \lambda(T)^{-1}\ .$$

1.10.14. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Demostrar:

1. Para todo $\varepsilon > 0$, existe una matriz diagonalizable D_ε tal que $\|A - D\|_{sp} \leq \varepsilon$.

2. Para todo $\varepsilon > 0$ existe una matriz inversible S tal que $T = SAS^{-1}$ es una matriz triangular superior que satisface
$$\sum_{i=1}^{n-1} \sum_{j>i} |T_{ij}|^2 \leq \varepsilon.$$
Notar que se suman los cuadrados de los módulos de las entradas que están estrictamente sobre la diagonal, por lo que puede decirse que T esta muy próxima a ser diagonal.

1.10.15. El objetivo de este ejercicio es probar que las "matrices de matrices", comúnmente llamadas *matrices de bloques*, se comportan del mismo modo que las matrices con entradas escalares. Una **matriz de bloques** es una matriz A, de $n \times m$, tal que sus entradas son matrices: para cada i, j,
$$A_{ij} \in \mathcal{M}_{n_i \times m_j}(\mathbb{C}).$$
Para que tenga sentido multiplicarlas como matrices de bloques, es decir que valga la fórmula
$$(A \cdot B)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} \ ,$$
hay que restringirse a conjuntos de matrices donde las condiciones sobre los tamaños de los bloques son más específicas. Hallar estas condiciones. Explicitar en el caso $m = n = 2$.

1.10.16. Considerar la matriz de bloques
$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22}, \end{bmatrix}, \quad A_{ii} \in \mathcal{M}_{n_i}(\mathbb{C})$$
Mostrar que $\sigma(A) = \sigma(A_{11}) \cup \sigma(A_{22})$. Si la matriz es triangular de bloques, es decir
$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \quad A_{ii} \in \mathcal{M}_{n_i}(\mathbb{C}) \ , \quad i = 1, 2,$$
¿qué se puede decir del $\sigma\,(A)$?.

1.10.17. Explicar cuál es el error en cada una de las siguientes "demostraciones" (falsas) del Teorema de Hamilton-Cayley:

1. Como $p_A(\lambda) = 0$ para cada autovalor λ de A, y como los autovalores de $q(A)$ son los $q(\lambda)$ para cualquier polinomio, se sigue que los autovalores de $p_A(A)$ son todos nulos; por lo tanto, $p_A(A) = 0$.

2. Como $p_A(t) = \det(tI - A)$, $p_A(A) = \det(AI - A) = \det(A - A) = \det 0 = 0$. Por lo tanto, $p_A(A) = 0$.

1.10.18. Sea $\mathbb{E}_n \in \mathcal{M}_n(\mathbb{C})$ la matriz cuyas entradas son todas iguales a 1. Hallar los autovalores de $\mathbb{E}_2$ y $\mathbb{E}_3$. Generalizar para $\mathbb{E}_n$.

1.10.19. Probar que cualquier familia de matrices que conmutan dos a dos tiene un autovector común a todas, mediante los siguientes pasos:

1. Probar que si $A, B \in \mathcal{M}_n(\mathbb{C})$ conmutan, entonces tienen un autovector en común.
2. Si $\mathcal{F} = \{A_1, \ldots, A_m\}$ es una familia finita de matrices que conmutan dos a dos, usar inducción para probar que hay un autovector común para todos.
3. Si la familia tiene cardinal no finito, encontrar algún curro para que dé.

1.10.20. Sean $A, B \in \mathcal{M}_n(\mathbb{C})$, y suponer que una de las dos es no singular. Si AB es diagonalizable, mostrar que BA es diagonalizable. Hallar un contraejemplo si A y B son singulares.

1.10.21. Sean $x, y, z, w \in \mathbb{C}^n$ todos vectores unitarios. Probar que

$$\langle x, y\rangle = \langle z, w\rangle \quad \Longrightarrow \qquad \text{existe } \ U \in \mathcal{U}(n) \ \text{ tal que } \ Ux = z \quad \text{y} \quad Uy = w \ .$$

1.10.22. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Una factorización $A = BC$ con $B, C \in \mathcal{M}_n(\mathbb{C})$ es llamada

1. LU-factorización si B es triangular inferior y $C \in \mathcal{TS}(n)$.
2. UL-factorización si C es triangular inferior y $B \in \mathcal{TS}(n)$.

Probar que siempre existen tales factorizaciones.

1.10.23. Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Definamos las transformaciones lineales

$$L_A \text{ y } R_B : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C}) \quad \text{ dadas por } L_A(X) = AX \text{ y } R_B(X) = XB \ , \ X \in \mathcal{M}_n(\mathbb{C}) \ .$$

1. Probar que $\sigma(L_A) = \sigma(A)$ y que $\sigma(R_B) = \sigma(B)$.
2. Probar que $\sigma(L_A - R_B) = \{\lambda - \mu : \lambda \in \sigma(A) \text{ y } \mu \in \sigma(B)\}$.
3. Deducir que las siguientes condiciones son equivalentes:
   (a) Para todo $Y \in \mathcal{M}_n(\mathbb{C})$, existe un único $X \in \mathcal{M}_n(\mathbb{C})$ tal que $AX - XB = Y$.
   (b) $\sigma(A) \cap \sigma(B) = \varnothing$.

1.10.24 (Proceso QR)**.** Sea $A \in \mathcal{G}l(n)$. Asumiremos que todos los autovalores de A tienen módulos distintos. Definiremos recursivamente tres sucesiones

$$\{A_m\}_{m\in\mathbb{N}} \quad \text{en } \mathcal{M}_n(\mathbb{C}) \ , \quad \{Q_m\}_{m\in\mathbb{N}} \quad \text{en } \mathcal{U}(n) \ , \quad \text{y} \quad \{R_m\}_{m\in\mathbb{N}} \quad \text{en } \mathcal{TS}(n) \ ,$$

donde todas las factorizaciones que haremos serán la única QR del Teorema 1.8.2:

1. Pongamos $A_1 = A = Q_1 \ R_1$.

2. Definimos $A_2 = R_1\, Q_1$ y lo factorizamos $A_2 = Q_2\, R_2$.

k. Definido $A_k = R_{k-1}\, Q_{k-1}$, lo factorizamos $A_k = Q_k\, R_k$, y definimos $A_{k+1} = R_k\, Q_k$.

m. Así seguimos definiendo y factorizando para todo $m \in \mathbb{N}$.

Probar que estas sucesiones cumplen lo siguiente.

(a) $A_2 = Q_1^* A Q_1$ y $A_3 = Q_2^*\, A_2\, Q_2 = Q_2^*\, Q_1^*\, A\, Q_1\, Q_2$.

(b) Dado $m \in \mathbb{N}$, sea $U_m = \prod\limits_{k=1}^{m} Q_k \in \mathcal{U}(n)$. Entonces $A_{m+1} = U_m^*\ A\ U_m$.

(c) Se cumple que $Q_m \xrightarrow[m\to\infty]{} I$. Más aún, $U_m \xrightarrow[m\to\infty]{} U \in \mathcal{U}(n)$.

(d) $R_m \xrightarrow[m\to\infty]{} T \in \mathcal{TS}(n)$, y también $A_m = Q_m\, R_m \xrightarrow[m\to\infty]{} T$.

(e) $T = U^* A U$, por lo que $\lambda(T) = \lambda(A)$.

Este proceso es fácil de aplicar, porque hacer QR es barato computacionalmente. Observar es un algoritmo para realizar el Teorema 1 de Schur 1.6.1, por lo que que permite calcular los autovalores de A, cosa que en general es bien complicada. Sin embargo, las pruebas de los items (c), (d) y (e) son bastante difíciles, y se enuncian más a título informativo que como verdadero ejercicio. Sugerimos asumir (c) y probar todo lo demás. ▲

Capítulo 2

Matrices normales y Hermitianas

2.1 Matrices normales

2.1.1. Repaso: Sean $A \in \mathcal{M}_n(\mathbb{C})$ y $a = (a_1, \dots, a_n) \in \mathbb{C}^n$.

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

2. Si $a = (a_1, \dots, a_n) \in \mathbb{C}^n$, recordemos que $\operatorname{diag}(a)$ denota la matriz diagonal

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

3. El Teorema 1.3.3 decía que si $B \in \mathcal{M}_n(\mathbb{C})$ cumple que $A \cong B$, entonces

$$\lambda(A) = \lambda(B) \ , \ \|A\|_2 = \|B\|_2 \ , \ \|A\|_{sp} = \|B\|_{sp} \quad \text{y} \quad A \text{ es } \star \text{ si y sólo si } B \text{ es } \star \ , \tag{2.1}$$

con $\star$ = Hermitiana, anti-Hermitiana, normal, definda positiva o unitaria. ▲

Teorema 2.1.2. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ con vector de autovalores $\lambda(A) = (\lambda_1 , \dots , \lambda_n)$. Entonces las siguientes condiciones son equivalentes:*

1. *A es normal.*

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

3. *$A \cong D$ para cierta matriz* **diagonal** *$D \in \mathcal{M}_n(\mathbb{C})$.*

4. *$A \cong \operatorname{diag}(\lambda(A))$.*

5. *Existe una BON $\mathcal{B} = \{v_1 , \dots , v_n\}$ de $\mathbb{C}^n$ tal que $Av_i = \lambda_i \ v_i$ para todo $i \in \mathbb{I}_n$.*

6. $\|A\|_2^2 = \sum\limits_{i=1}^{n} |\lambda_i|^2$

Demostración. 1 → 2: Para cada $x \in \mathbb{C}^n$, se tiene que

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

2 → 3: Por el Teorema 1 de Schur 1.6.1, existen $U \in \mathcal{U}(n)$ y $T \in \mathcal{TS}(n)$ tales que $T = U^*AU$. Luego, también se tiene que $\|Ty\| = \|T^*y\|$, para todo $y \in \mathbb{C}^n$. Aplicando esto a la base canónica, se deduce inductivamente (fila por fila) que T debe ser diagonal.

3 → 4: Si $A \cong D$, con D diagonal, la Eq. (2.1) asegura que $\lambda(D) = \lambda(A)$. Pero como D es diagonal, $\lambda(D) = \mathrm{d}(D)$ salvo el orden. Conjugando con una matriz de permutación (o sea que tiene a la base canónica en algún otro orden en sus columnas, por lo que es unitaria), reordenamos la diagonal de D, obteniendo que $D \cong \mathrm{diag}(\lambda(A))$.

4 ↔ 5: Llamemos $D = \mathrm{diag}(\lambda(A))$. Si existe $U \in \mathcal{U}(n)$ tal que $D = U^*AU$, tomemos $\mathcal{B} = \{C_1(U), \dots, C_n(U)\}$. Como $AU = UD$, la fórmula (1.12) decía que $AC_i(U) = \lambda_i\, C_i(U)$ para todo $i \in \mathbb{I}_n$. Recíprocamente, si existe $\mathcal{B}$ como en 5, tomar la $U \in \mathcal{U}(n)$ dada por $C_i(U) = v_i$, para todo $i \in \mathbb{I}_n$, y hacer la misma cuenta.

4 → 6: Si $A \cong \mathrm{diag}(\lambda(A))$, la Eq. (2.1) muestra que $\|A\|_2^2 = \|\mathrm{diag}(\lambda(A))\|_2^2 = \sum\limits_{i=1}^{n} |\lambda_i|^2$.

6 → 1: Si $U \in \mathcal{U}(n)$ y $T \in \mathcal{TS}(n)$ cumplen $T = U^*AU$ y $\mathrm{d}(T) = \lambda(A)$, entonces

$$\sum_{i=1}^{n} |\lambda_i|^2 = \|A\|_2^2 = \|T\|_2^2 = \sum_{i=1}^{n} |\lambda_i|^2 + \sum_{i<j} |t_{ij}|^2.$$

Por lo tanto $t_{ij} = 0$ para $i < j$, o sea que T es diagonal, y por ende normal. Por la Eq. (2.1), también A debe ser $\star$ (o sea normal). ■

Definición 2.1.3. Sea $A \in \mathcal{M}_n(\mathbb{C})$ una matriz normal. Diremos que

$$\mathcal{B} = \{v_1, \dots, v_n\} \quad \text{es una BON } \textbf{adaptada} \text{ a } \lambda(A)$$

si $\mathcal{B}$ verifica el ítem 4 del Teorema 2.1.2, es decir que $\mathcal{B}$ es una BON de $\mathbb{C}^n$, y $Av_i = \lambda_i(A)\, v_i$ para todo $i \in \mathbb{I}_n$. ▲

Hemos visto en la Eq. (1.6) que si $A \in \mathcal{M}_n(\mathbb{C})$, entonces $\rho(A) \le w(A) \le \|A\|_{sp}$, y que en general estas desigualdades pueden ser estrictas. Pero no es así cuando A es normal:

Corolario 2.1.4. *Si $A \in \mathcal{M}_n(\mathbb{C})$ es normal, entonces $\|A\|_{sp} = w(A) = \rho(A)$.*

Demostración. Sea $\lambda(A) = (\lambda_1, \dots, \lambda_n)$ el vector de autovalores de A. Llamemos $D = \mathrm{diag}(\lambda(A))$. Por el Teorema 2.1.2, existe $U \in \mathcal{U}(n)$ tal que $A = UDU^*$. Por un lado, la Eq. (2.1) asegura que $\|A\|_{sp} = \|D\|_{sp}$ y que $\rho(A) = \rho(D)$, pues tienen el mismo espectro. Por otro lado, si $x = (x_1, \dots x_n) \in \mathbb{C}^n$ es unitario (i.e. $\|x\| = 1$), entonces

$$\|Dx\|^2 = \sum_{i=1}^{n} |\lambda_i|^2 |x_i|^2 \le \max_{i \in \mathbb{I}_n} |\lambda_i|^2 \ \cdot \ \sum_{i=1}^{n} |x_i|^2 = \rho(A)^2 \ ,$$

por lo que $\|A\|_{sp} = \|D\|_{sp} \le \rho(A)$. Las otras desigualdades las vimos en la Eq. (1.6). ■

Corolario 2.1.5. *Sean $A, B \in \mathcal{M}_n(\mathbb{C})$ matrices normales. Si tomamos un orden fijo en $\mathbb{C}$ para ordenar los vectores $\lambda(A)$ y $\lambda(B)$, se tiene que*

$$\lambda(A) = \lambda(B) \iff \text{existe } U \in \mathcal{U}(n) \text{ tal que } B = UAU^* \quad , \ \text{i.e.} \quad A \cong B \ .$$

En otras palabras, si definimos la órbita unitaria $\mathcal{U}(A) := \{\ UAU^* : U \in \mathcal{U}(n)\}$, entonces

$$\mathcal{U}(A) = \{\ B \in \mathcal{N}(n) \ : \ \lambda(B) = \lambda(A)\ \} \ .$$

Demostración. La Eq. (2.1) asegura que si $A \cong B$, entonces $\lambda(A) = \lambda(B)$. Recíprocamente, si $D = \operatorname{diag}(\lambda(A)) = \operatorname{diag}(\lambda(B))$, el Teorema 2.1.2 dice que $A \cong D \cong B$. ■

2.2 Matrices Hermitianas

Por lo general no es fácil calcular los autovalores de una matriz. Pero en muchos casos es suficiente saber que ellos están en un intervalo especificado. En el resto de este Capítulo estudiaremos algunas de las principales características que distinguen a las matrices Hermitianas, en particular los principios variacionales que se utilizan para localizar su espectro, sin la necesidad de conocer los autovectores asociados en forma exacta. Recordemos las notaciones

$$\mathcal{H}(n) = \{A \in \mathcal{M}_n(\mathbb{C}) : A = A^*\} \quad \text{y} \quad \mathcal{M}_n(\mathbb{C})^+ = \{A \in \mathcal{H}(n) : A \geq 0\} \ .$$

Teorema 2.2.1. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Luego son equivalentes:*

1. *$A \in \mathcal{H}(n)$.*
2. *A es normal y $\sigma(A) \subseteq \mathbb{R}$.*
3. *$\lambda(A) \in \mathbb{R}^n$ y existe una base ortonormal $\mathcal{B}$ adaptada a $\lambda(A)$.*
4. *$\lambda(A) \in \mathbb{R}^n$ y $A \cong \operatorname{diag}(\lambda(A))$, i.e. existe $U \in \mathcal{U}(n)$ tal que $U^*AU = \operatorname{diag}(\lambda(A))$.*
5. *$\langle Ax, x\rangle \in \mathbb{R}$ para todo $x \in \mathbb{C}^n$.*

Demostración. Por el Corolario 1.6.6, se tiene que $\lambda(A^*) = \overline{\lambda(A)}$. Por lo tanto, si $A \in \mathcal{H}(n)$, vemos que $\sigma(A) \subseteq \mathbb{R}$. El resto se deduce del Teorema 2.1.2 y del hecho de que una matriz diagonal en $\mathcal{M}_n(\mathbb{R})$ debe ser autoajunta. La equivalencia entre el ítem 5 y los demás se sigue del Corolario 1.1.13. ■

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 $\lambda_k(A) \leq \lambda_{k+1}(A)$, $k \in \mathbb{I}_{n-1}$.

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

3. Se llamarán

$$\lambda_{\text{mín}}(A) = \lambda_1(A) = \mu_n(A) = \text{mín}\, \sigma(A) \quad \text{y} \quad \lambda_{\text{máx}}(A) = \lambda_n(A) = \mu_1(A) = \text{máx}\, \sigma(A) \ .$$

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 . ▲

Proposición 2.2.3. *Sea $A \in \mathcal{H}(n)$. Entonces se tiene que*

$$\|A\|_{sp} = \rho(A) = \text{máx}\{\lambda_n(A), -\lambda_1(A)\} \ . \tag{2.2}$$

Demostración. Como $\mathcal{H}(n) \subseteq \mathcal{N}(n)$, la igualdad $\|A\|_{sp} = \rho(A)$ se sigue del Corolario 2.1.4. La otra se deduce de que $\sigma(A) \subseteq [\lambda_1(A), \lambda_n(A)] \subseteq \mathbb{R}$, y contiene a los bordes. ■

2.3 Principio minimax

Para matrices generales la única caracterización conocida de sus autovalores es que son las raíces del polinomio característico de la matriz. Pero cuando las matrices son Hermitianas, el hecho de poder establecer un orden entre ellos nos permite obtener caracterizaciones más interesantes. Los próximos teoremas describen al vector $\lambda(A)$, para $A \in \mathcal{H}(n)$, en función de las expresiones $\dfrac{\langle Ax, x\rangle}{\langle x, x\rangle}$, para $x \in \mathbb{C}^n \setminus \{0\}$, conocidas como cocientes de Rayleig-Ritz.

Teorema 2.3.1 (Rayleigh-Ritz)**.** *Sea $A \in \mathcal{H}(n)$. Entonces*

1. *Para todo $x \in \mathbb{C}^n$ se tiene que $\lambda_{\text{mín}}(A)\|x\|^2 \leq \langle Ax, x\rangle \leq \lambda_{\text{máx}}(A)\|x\|^2$.*

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

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

En particular, si $A \in \mathcal{H}(n)$, tenemos que

$$A \in \mathcal{M}_n(\mathbb{C})^+ \iff \lambda_{\text{mín}}(A) \geq 0 \iff \sigma(A) \subseteq \mathbb{R}_+ \ .$$

Demostración. Sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una BON de $\mathbb{C}^n$ adaptada a $\lambda(A)$, o sea que $Av_i = \lambda_i(A)v_i$ para todo $i \in \mathbb{I}_n$. Por lo tanto, dado $x \in \mathbb{C}^n$, se tiene que

$$x = \sum_{i=1}^n \langle x, v_i\rangle \ v_i \ , \quad \|x\|^2 = \sum_{i=1}^n |\langle x, v_i\rangle|^2 \quad \text{y} \quad \langle Ax, x\rangle = \sum_{i=1}^n \lambda_i(A) \, |\langle x, v_i\rangle|^2 \ . \tag{2.3}$$

Luego, si asumimos que $\|x\| = 1$, tenemos que

$$\langle Ax, x\rangle = \sum_{i=1}^{n} \lambda_i(A)\,|\,\langle x, v_i\rangle|^{\,2} \leq \lambda_n(A) \sum_{i=1}^{n} |\,\langle x, v_i\rangle|^{\,2} = \lambda_n(A) = \langle Av_n, v_n\rangle \ .$$

Análogamente, $\langle Ax, x\rangle = \sum\limits_{i=1}^{n} \lambda_i(A)\,|\,\langle x, v_i\rangle|^{\,2} \geq \lambda_1(A) = \langle Av_1\,, v_1\rangle$. Es claro que estas desigualdades muestran los tres ítems a probar. ■

Observación 2.3.2. Dada $A \in \mathcal{H}(n)$, las caracterizaciones del Teorema anterior se pueden reescribir de la siguiente forma: $\lambda_1(A)\, I \leq A \leq \lambda_n(A)\, I$,

$$\lambda_1(A) = \text{máx}\{\lambda \in \mathbb{R} : \lambda\, I \leq A\} \quad \text{ y } \quad \lambda_n(A) = \text{mín}\{\lambda \in \mathbb{R} : A \leq \lambda\, I\} \ .$$

En efecto, para mostrarlo basta recordar que dados $B, C \in \mathcal{H}(n)$, vale que

$$B \leq C \iff \langle B\,x\,, x\rangle \leq \langle C\,x\,, x\rangle \text{ para todo } x \text{ unitario en } \mathbb{C}^n \ .$$ ▲

Notaciones: En el resto de esta sección usaremos las siguientes convenciones:

1. Las letras $\mathcal{M}$ y $\mathcal{S}$ denotarán subespacios de $\mathbb{C}^n$.
2. Dado $\mathcal{M} \subseteq \mathbb{C}^n$, escribiremos

   $$\mathcal{M}_1 = \{x \in \mathcal{M} : \|x\| = 1\}$$

   al conjunto de elementos de $\mathcal{M}$ de norma uno. ▲

Teorema 2.3.3 (Courant-Fisher). *Sea $A \in \mathcal{H}(n)$ y sea $k \in \mathbb{I}_n$. Entonces,*

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

Demostración. Sea $\mathcal{B} = \{v_1\,, \dots, v_n\}$ una BON de $\mathbb{C}^n$ adaptada a $\lambda(A)$. Como en la prueba del Teorema 2.3.1, cualquier $x \in \mathbb{C}^n$ verifica la Eq. (2.3). Dado $r \in \mathbb{I}_n$, notemos por $\mathbb{H}_r = \text{Gen}\,\{v_1\,, \dots, v_r\}$ y $\mathbb{K}_r = \text{Gen}\,\{v_r\,, \dots, v_n\}$. Notar que $\dim \mathbb{H}_r = r$ y $\dim \mathbb{K}_r = n - r + 1$. Por la Eq. (2.3) vemos que, si $x \in \mathbb{K}_k$,

$$\langle Ax, x\rangle = \sum_{i=k}^{n} \lambda_i(A)\,|\,\langle x, v_i\rangle|^{\,2} \implies \lambda_k(A) = \min_{x\in(\mathbb{K}_k)_1} \langle Ax, x\rangle \leq \max_{\dim \mathcal{S}=n-k+1}\ \min_{x\in\mathcal{S}_1} \langle Ax, x\rangle \ .$$

Por otro lado, si $\dim \mathcal{S} = n - k + 1$, entonces $\mathcal{S} \cap \mathbb{H}_k \neq \{0\}$. Pero si $y \in (\mathcal{S} \cap \mathbb{H}_k)_1$, la Eq. (2.3) asegura que $\langle Ay, y\rangle = \sum\limits_{i=1}^{k} \lambda_i(A)\,|\,\langle y, v_i\rangle|^{\,2}$ y que $\|y\|^2 = \sum\limits_{i=1}^{k} |\,\langle y, v_i\rangle|^{\,2} = 1$. Luego

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

La otra fórmula se demuestra en forma análoga: el $\min\limits_{\dim \mathcal{M}=k}$ se alcanza en $\mathcal{M} = \mathbb{H}_k$, y cualquier otro tal $\mathcal{M}$ cumple que $\mathcal{M} \cap \mathbb{K}_k \neq \{0\}$. ■

Observación 2.3.4. La versión tradicional de las fórmulas de Courant-Fisher sería la siguiente:

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

Teorema 2.3.5 (Teorema de Weyl)**.** *Sean $A, B \in \mathcal{H}(n)$. Entonces:*

$$\lambda_j(A) + \lambda_1(B) \le \lambda_j(A+B) \le \lambda_j(A) + \lambda_n(B) \qquad \textit{para todo } j \in \mathbb{I}_n \ . \tag{2.4}$$

Demostración. Por el Teorema 2.3.1, para todo $x \in \mathbb{C}^n$ tal que $\|x\| = 1$, se tiene que

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

Por lo tanto el teorema se puede deducir de las fórmulas de Courant-Fischer. ■

Observación 2.3.6. Una reformulación del Teorema de Weyl, que es bastante común en sus aplicaciones, es la siguiente: Sean $C, D \in \mathcal{H}(n)$, entonces:

$$\lambda_1(C-D) \le \lambda_j(C) - \lambda_j(D) \le \lambda_n(C-D) \ , \qquad \text{para todo} \quad j \in \mathbb{I}_n \ . \tag{2.5}$$

Para mostrarla, basta tomar $A = D$ y $B = C - D$, observar que ambos viven en $\mathcal{H}(n)$, que $A + B = C$ y, por último, aplicar la Eq. (2.4). ▲

Corolario 2.3.7. *Sean $A, B \in \mathcal{H}(n)$ tales que $A \le B$, i.e. $B - A \in \mathcal{M}_n(\mathbb{C})^+$. Entonces*

$$\lambda_j(A) \le \lambda_j(B) \qquad \textit{para todo } j \in \mathbb{I}_n \ .$$

Demostración. Llamemos $C = B - A$. Por el Teorema 2.3.5, tenemos que

$$\lambda_j(A) + \lambda_1(C) \le \lambda_j(A + C) = \lambda_j(A + (B - A)) = \lambda_j(B) \ .$$

Por otra parte, como $C \in \mathcal{M}_n(\mathbb{C})^+$, entonces $\lambda_1(C) = \min_{\|x\|=1} \langle Cx, x\rangle \ge 0$. ■

Una consecuencia importante del Teorema de Weyl es el hecho de que, entre las autoadjuntas, matrices muy cercanas tienen autovalores muy cercanos. Y con cotas bien claras:

Corolario 2.3.8. *Sean $A, B \in \mathcal{H}(n)$. Entonces:*

$$\|\lambda(A) - \lambda(B)\|_\infty := \max_{j\in\mathbb{I}_n} |\lambda_j(A) - \lambda_j(B)| \le \rho(A - B) = \|A - B\|_{sp} \ .$$

Demostración. Por el Teorema de Weyl, en su versión dada por la Eq. (2.5), se tiene que

$$\lambda_1(A-B) \le \lambda_j(A) - \lambda_j(B) \le \lambda_n(A-B) \ , \quad \text{para todo} \quad j \in \mathbb{I}_n \ .$$

Por lo tanto, aplicando la Proposición 2.2.3, se obtiene que

$$\|\lambda(A) - \lambda(B)\|_\infty \le \text{máx}\Big\{|\lambda| : \lambda \in \big[\lambda_1(A-B),\ \lambda_n(A-B)\big]\Big\} = \rho(A-B) \ . \qquad ■$$

Ejercicio 2.3.9 (Aronszajn)**.** Demostrar las siguientes afirmaciones:

1. Dados S_1, S_2 y S_3 subespacios de $\mathbb{C}^n$, probar que

$$\dim(S_1 \cap S_2 \cap S_3) \ge \dim S_1 + \dim S_2 + \dim S_3 - 2n \ .$$

2. Sean $A, B \in \mathcal{H}(n)$. Dados $i, j \in \mathbb{I}_n$ tales que $i + j \le n+1$, se tiene que

$$\mu_{i+j-1}(A+B) \le \mu_i(A) + \mu_j(B) \ . \qquad ▲$$

2.4 Entrelace de Cauchy

Una consecuencia directa del Teorema de Courant-Fisher es el llamado teorema de entrelace de Cauchy, que relaciona los autovalores de una matriz Hermitiana con los de sus submatrices principales. Antes fijaremos nuevas notaciones para estas submatrices:

Definición 2.4.1. Sean $A \in \mathcal{M}_n(\mathbb{C})$ y $J \subseteq \mathbb{I}_n$. Si J tiene k elementos, notaremos

$$A[J] = \{a_{ij}\}_{i,j \in J} \in \mathcal{M}_k(\mathbb{C}) \quad \text{y} \quad A(J) = \{a_{ij}\}_{i,j \notin J} \in \mathcal{M}_{n-k}(\mathbb{C}) \ .$$

Si el contexto lo permite, a veces abreviaremos $A[J] = A_J$, como en la sección 1.5. Con esa convención, se tiene que $A(J) = A_{\mathbb{I}_n \setminus J}$. Observar que $A[J]$ es la matriz cuadrada resultante de borrar de A las filas y columnas con índices fuera de J. Para cada $r \in \mathbb{I}_n$, llamaremos

$$A_r = A(\{r\}) = \{a_{ij}\}_{i \ne r \ne j} \in \mathcal{M}_{n-1}(\mathbb{C}) \ , \tag{2.6}$$

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

Teorema 2.4.2 (Entrelace de Cauchy)**.** *Sean $A \in \mathcal{H}(n)$, $r \in \mathbb{I}_n$ y $A_r \in \mathcal{M}_{n-1}(\mathbb{C})$ la submatriz principal de A obtenida como en la Eq.* (2.6) *. Entonces*

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

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

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

Demostración. Supongamos, por simplicidad, que $r = n$. Los otros casos se prueban exactamente igual, pero con notaciones algo más engorrosas. Fijemos un $k \in \mathbb{I}_{n-1}$. Sea

$$\mathbb{H}_{n-1} = \{e_n\}^\perp = \text{Gen}\{e_1, \dots, e_{n-1}\} = \{x \in \mathbb{C}^n : x_n = 0\} .$$

Si $x \in \mathbb{H}_{n-1}$, notaremos $x_0 = (x_1, \dots, x_{n-1}) \in \mathbb{C}^{n-1}$ a su parte significativa. Observar que $\langle A_n\, x_0, x_0\rangle = \sum\limits_{i,j=1}^{n-1} A_{ij}\, x_j\, \overline{x_i} = \langle Ax, x\rangle$, dado que $x_n = 0$. Trabajando sólo con los subespacios $\mathcal{M} \subseteq \mathbb{H}_{n-1}$ que tienen $\dim \mathcal{M} = k$ (que son menos que los subespacios de dimensión k de todo $\mathbb{C}^n$, pero se identifican con todos los de $\mathbb{C}^{n-1}$), obtenemos, del Teorema 2.3.3, que

$$\lambda_k(A_n) = \min_{\substack{\dim \mathcal{M}=k \\ \mathcal{M}\subseteq\, \mathbb{H}_{n-1}}} \max_{x\in\mathcal{M}_1} \langle A_n\, x_0, x_0\rangle \geq \min_{\substack{\dim \mathcal{M}=k \\ \mathcal{M}\subseteq \mathbb{C}^n}} \max_{x\in\mathcal{M}_1} \langle Ax, x\rangle = \lambda_k(A).$$

Tomemos ahora subespacios $\mathcal{S} \subseteq \mathbb{H}_{n-1}$ tales que $\dim \mathcal{S} = n-k$. Como $n-k = n-(k+1)+1$ y a la ves, $n-k = (n-1)-k+1$, obtenemos

$$\lambda_k(A_n) = \max_{\substack{\dim \mathcal{S}=n-k \\ \mathcal{S}\subseteq\, \mathbb{H}_{n-1}}} \min_{x\in\mathcal{S}_1} \langle A_n\, x_0, x_0\rangle \leq \max_{\dim \mathcal{S}=n-k} \min_{x\in\mathcal{S}_1} \langle Ax, x\rangle = \lambda_{k+1}(A) ,$$

lo que prueba el teorema. ■

Observación 2.4.3. En forma análoga se puede probar versiones más generales del Teorema anterior: Dado $A \in \mathcal{H}(n)$,

1. Si $J \subseteq \mathbb{I}_n$ cumple que $|J| = r$, entonces para cada $k \in \mathbb{I}_r$, se tiene
$$\lambda_k(A) \leq \lambda_k\big(A[J]\big) \leq \lambda_{k+n-r}(A) .$$
Observar que, si $r = n-1$, entonces $k+n-r = k+1$, como en el Teorema 2.4.2

2. Más en general aún, si $P \in L(\mathbb{H})^+$ es un proyector autoadjunto (o sea ortogonal) sobre un subespacio $\mathcal{S}$ de $\dim \mathcal{S} = r$, entonces al producto PAP se lo puede pensar como un operador en el espacio de Hilbert $\mathcal{S}$ (para que su vector de autovalores tenga sólo r coordenadas, sacando los $n-r$ ceros que "sobran"). A esta compresión se la denota $A_{\mathcal{S}} = PAP\big|_{\mathcal{S}} \in L(\mathcal{S})$. Entonces se obtienen desigualdades análogas:
$$\lambda_k(A) \leq \lambda_k(A_{\mathcal{S}}) \leq \lambda_{k+n-r}(A) \quad , \qquad \text{para cada } k \in \mathbb{I}_r .$$
En efecto, basta cambiar coordenadas a una BON de $\mathbb{C}^n$ cuyos primeros r vectores generen $\mathcal{S}$. En esa base, $A_{\mathcal{S}} = A[\mathbb{I}_r]$ y se aplica el caso anterior. ▲

Ejercicio 2.4.4. Escribir explícitamente cómo quedan los resultados de esta sección (la fórmula minimax, el teorema de Weyl y los tres entrelaces) en función de los vectores $\mu(A)$ ordenados decrecientemente. ▲

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

Lema 2.4.5. *Sea $A \in \mathcal{G}l(n)^+$. Entonces $A[J] \in \mathcal{G}l(r)^+$, para todo $J \subseteq \mathbb{I}_n$ con $|J| = r$.*

Demostración. Si $\mathbb{H}_J = \operatorname{Gen}\{e_i : i \in J\}$ y $0 \neq x \in \mathbb{H}_J$, llamemos $x_J \in \mathbb{C}^r$ al vector resultante de sacarle a x los ceros fuera de J. Entonces

$$0 < \langle Ax, x\rangle = \sum_{i,j\in\mathbb{I}_n} A_{ij}\, x_j\, \overline{x_i} = \sum_{i,j\in J} A_{ij}\, x_j\, \overline{x_i} = \langle A[J]\, x_J\, , x_J\rangle \ .$$

Como tales x_J recorren todo $\mathbb{C}^r \setminus \{0\}$, vemos que $A[J] \in \mathcal{G}l(r)^+$. ■

Teorema 2.4.6. *Si $A \in \mathcal{H}(n)$, entonces las siguientes condiciones son equivalentes:*

1. *A es definida positiva (i.e. $A \in \mathcal{G}l(n)^+$).*
2. *Si llamamos $A_{[k]} = A[\mathbb{I}_k] = \{a_{ij}\}_{i,j\in\mathbb{I}_k} \in \mathcal{M}_k(\mathbb{C})$, entonces*
$$\det A_{[k]} > 0 \quad \textit{para todo} \;\; k \in \mathbb{I}_n \,,$$
3. *Si llamamos $A_{(k)} = A(\mathbb{I}_k) = \{a_{ij}\}_{i,j>k} \in \mathcal{M}_{n-k}(\mathbb{C})$, entonces*
$$\det A_{(k)} > 0 \quad \textit{para todo} \;\; k \in \mathbb{I}_{n-1} \cup \{0\} \,,$$

Demostración. El Lema 2.4.5 dice que $1 \to 2$ y 3, porque es claro que si $B \in \mathcal{G}l(r)^+$, entonces $\det B > 0$. La recíproca se prueba por inducción sobre n. Si $n = 1$, entonces tenemos que $A = A_{[1]} = \det(A_{[1]}) > 0$. Si $n > 1$, es claro que la condición 2 se verifica también para $A_{[n-1]}$, porque tiene las mismas submatrices involucradas. Por hipótesis inductiva, tenemos que $A_{[n-1]} \in \mathcal{G}l(n-1)^+$, y por el Teorema 2.3.1 sabemos que $0 < \lambda_1(A_{[n-1]})$. El Teorema del entrelace de Cauchy 2.4.2 nos asegura que $\lambda_2(A) \geq \lambda_1(A_{[n-1]}) > 0$. Luego

$$0 < \prod_{k=2}^{n} \lambda_k(A) \quad \text{y tambén} \quad 0 < \det A_{[n]} = \det A = \prod_{k=1}^{n} \lambda_k(A) \ .$$

De ahí deducimos que $\lambda_1(A) > 0$. Usando el Teorema 2.3.1, podemos concluir rápidamente que $\langle Ax, x\rangle > 0$ para todo $x \neq 0$, o sea que $A \in \mathcal{G}l(n)^+$. La prueba de la equivalencia con el ítem 3 es exactamente igual, pero usando para la inducción a $A_{(1)} \in \mathcal{M}_{n-1}(\mathbb{C})$. ■

Ejercicio 2.4.7 (difícil). Probar que, dada $A \in \mathcal{H}(n)$, entonces

$$A \in \mathcal{M}_n(\mathbb{C})^+ \quad \iff \quad \det A[J] \geq 0 \quad \text{para todo} \quad J \subseteq \mathbb{I}_n \ .$$

Se suguiere inducionar en n. Luego tomar un J de tamaño máximo para que $\det A[J] \neq 0$ y aplicar varias veces la fórmula $\det(B + \varepsilon E_{ii}) = \det B + \varepsilon \det B(i)$, como en la Eq. (8.3). ▲

2.5 Ejercicios

Ejercicios del texto

2.5.1. Sea $A \in \mathcal{H}(n)$ y sea $k \in \mathbb{I}_n$. Probar que

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

2.5.2. Sean $A, B \in \mathcal{H}(n)$. Demostrar las siguientes afirmaciones:

1. Para todo $j \in \mathbb{I}_n$, se tiene que

$$\mu_j\,(A) + \mu_n\,(B) \le \mu_j\,(A+B) \le \mu_j\,(A) + \mu_1\,(B) \tag{2.7}$$

2. Dados S_1, S_2 y S_3 subespacios de $\mathbb{C}^n$, ellos cumplen que

$$\dim(S_1 \cap S_2 \cap S_3) \ge \dim S_1 + \dim S_2 + \dim S_3 - 2n \ .$$

3. Dados $i, j \in \mathbb{I}_n$ tales que $i + j \le n + 1$, se tiene que

$$\mu_{i+j-1}(A+B) \le \mu_i(A) + \mu_j(B) \ .$$

2.5.3 (Aronszajn)**.** Sea $A = \begin{array}{cc} \begin{bmatrix} C & X \\ X^* & D \end{bmatrix} & \begin{array}{c} k \\ n-k \end{array} \end{array} \in \mathcal{H}(n)$. Probar que

$$\mu_{i+j-1}(A) \le \mu_i(C) + \mu_j(D) \qquad \text{para todo par } i \in \mathbb{I}_k \ , \ j \in \mathbb{I}_{n-k} \ .$$

2.5.4. Dado $A \in \mathcal{H}(n)$, mostrar que:

1. Si $J \subseteq \mathbb{I}_n$ cumple que $|J| = r$, entonces para cada $k \in \mathbb{I}_r$, se tiene

$$\mu_k(A) \ge \mu_k\big(A[J]\big) \ge \mu_{k+n-r}(A) \ .$$

   En particular, si $r \in \mathbb{I}_n$, entonces $\mu_k(A) \ge \mu_k\big(A_r\big) \ge \mu_{k+1}(A)$ para todo $k \in \mathbb{I}_{n-1}$.

2. Sea $P \in L(\mathbb{H})^+$ es un proyector autoadjunto (o sea ortogonal) sobre un subespacio $\mathcal{S}$ de $\dim \mathcal{S} = r$. Sea $A_{\mathcal{S}} = PAP\big|_{\mathcal{S}} \in L(\mathcal{S})$, la compresión de A a $\mathcal{S}$. Luego

$$\mu_k(A) \ge \mu_k\big(A_{\mathcal{S}}\big) \ge \mu_{k+n-r}(A) \quad , \qquad \text{para cada } k \in \mathbb{I}_r \ .$$

2.5.5 (**Ejercicio difícil).** Probar que, dada $A \in \mathcal{H}(n)$, entonces

$$A \in \mathcal{M}_n(\mathbb{C})^+ \quad \iff \quad \det A[J] \ge 0 \qquad \text{para todo} \qquad J \subseteq \mathbb{I}_n \ .$$

▲

Ejercicios nuevos

2.5.6. Mostrar que A es normal si y sólo si sus partes real e imaginaria conmutan.

2.5.7. Sea $A \in \mathcal{M}_n(\mathbb{C})$ y $p(t)$ un polinomio.

1. Probar que si A es normal entonces $p(A)$ también lo es.
2. Si $p(A)$ es normal, ¿ puede asegurarse que A lo sea?.

2.5.8.

1. Mostrar que si A es similar a una matriz unitaria, entonces A^{-1} es similar a A^*.
2. Considerar la matriz
$$\begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}$$
y mostrar que el conjunto de matrices que son similares a una matriz unitaria es un subcojunto propio del conjunto de matrices A para las que A^{-1} es similar a A^*.

2.5.9. Sea A una matriz normal. Mostrar:

1. La matriz A es autoadjunta si y sólo si todos sus autovalores son reales.
2. La matriz A es unitaria si y sólo si todos sus autovalores tienen módulo 1.
3. Si la matriz A es nilpotente, entonces, $A = 0$.

2.5.10.

1. Mostrar que dos matrices normales son similares sii son unitariamente equivalentes.
2. Mostrar que A es normal sii conmuta con una cierta matriz normal con autovalores distintos.

2.5.11. Dada $A \in \mathcal{M}_n(\mathbb{C})$, probar que A es normal si y sólo si hay un polinomio p de grado a lo sumo $n-1$ tal que $A^* = p(A)$. Notar que esto da una buena explicación “intuitiva” de por qué una matriz normal conmuta con su adjunto. Si A es real, mostrar que se puede elegir p con coeficientes reales, de manera que $A^T = p(A)$.

2.5.12. Sea $A \in \mathcal{H}(n)$ y $S \in \mathcal{M}_n(\mathbb{C})$. Mostrar que SAS^* es autoadjunta. Si S es invertible, SAS^{-1} ¿ es autoadjunta ?

2.5.13 (*). A lo largo de este ejercicio consideraremos la traza normalizada de modo que $\operatorname{tr}(I) = 1$, es decir, dada una matriz A de $n \times n$, $\operatorname{tr}(A) = \dfrac{1}{n} \sum_{k=1}^{n} A_{ii}$.

Sean $A, B \in \mathcal{H}(n)$. Demostrar:

1. $\operatorname{tr}(AB)^2 \le \operatorname{tr}(A^2B^2)$.
2. Si $A \neq 0$, entonces, $\dfrac{\operatorname{r}(A)}{n} \ge \dfrac{(\operatorname{tr} A)^2}{\operatorname{tr} A^2}$.

(Pista: Usar la desigualdad de Jensen.)

2.5.14. Sea $A \in \mathcal{M}_n(\mathbb{C})$ una matriz normal. Probar que $w(A) = \rho(A) = \|A\|_{sp}$.

2.5.15 (**Gersgorin**)**.** Sea $A \in \mathcal{M}_n(\mathbb{C})$. Para cada $i \in \mathbb{I}_n$, sea $R_i = \sum\limits_{j \neq i} |a_{ij}|$. Mostrar que

$$\sigma(A) \subseteq \bigcup_{i \in \mathbb{I}_n} \{ z \in \mathbb{C} : |z - a_{ii}| \le R_i \} \ .$$

Deducir que, si $R_i < |a_{ii}|$ para todo $i \in \mathbb{I}_n$, entonces $A \in \mathcal{G}l(n)$. Una tal matriz suele ser llamada "diagonal dominante".

2.5.16 (Este es bien difícil)**.** Sea $A \in \mathcal{H}(n)$. Para cada $i \in \mathbb{I}_n$, mostrar que si

$$r_i = \left(\sum_{j \neq i} |a_{ij}|^2 \right)^{1/2} \quad \Longrightarrow \quad \sigma(A) \cap [a_{ii} - r_i \,, \ a_{ii} + r_i] \neq \varnothing \ .$$

Esto mejora al Ejercicio anterior en dos aspectos: Primero observar que cada $r_i \le R_i$. Además, ubica al menos un autovalor en cada disquito (acá son intervalitos).

Capítulo 3

Matrices definidas positivas

3.1 Propiedades básicas

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

Definición 3.1.1. 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 3.1.2. *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* $C^*AC \leq C^*BC$.

Demostración.

1. Si $A \in \mathcal{H}(n)$ y $\sigma(A) \subseteq \mathbb{R}_+$, el Teorema 2.3.1 asegura que

$$0 \leq \lambda_1(A)\|x\|^2 \leq \langle Ax, x\rangle \quad \text{para todo } x \in \mathbb{C}^n \quad \Longrightarrow \quad A \in \mathcal{M}_n(\mathbb{C})^+ .$$

   Por el Corolario 1.1.13, sabemos que $\mathcal{M}_n(\mathbb{C})^+ \subseteq \mathcal{H}(n)$ (para que $A \in \mathcal{H}(n)$ bastaba que $\langle Ax, x\rangle \in \mathbb{R}$ para todo $x \in \mathbb{C}^n$). Por el Teorema 2.3.1, se tiene que, si $A \in \mathcal{M}_n(\mathbb{C})^+$, entonces $\lambda_1(A) = \min\limits_{\|x\|=1} \langle Ax, x\rangle \geq 0$, por lo que $\sigma(A) \subseteq \mathbb{R}_+$.

2. Sólo difiere del caso anterior en que en ambas situaciones $0 \notin \sigma(A)$. Notar que si $A > 0$, como la bola de $\mathbb{C}^n$ es compacta, existe un $\varepsilon > 0$ tal que $\lambda_1(A) = \min\limits_{\|x\|=1} \langle Ax, x\rangle \geq \varepsilon$. Luego $0 \notin \sigma(A)$, o sea que A es inversible.

3. Para todo $x \in \mathbb{C}^n$ tenemos que

$$\langle B^*Bx, x\rangle = \langle Bx, Bx\rangle = \|Bx\|^2 \geq 0.$$

Por lo tanto $B^*B \in \mathcal{M}_n(\mathbb{C})^+$.

4. Si $B - A \geq 0$ y $x \in \mathbb{C}^n$, entonces $\langle C^*(B-A)Cx, x\rangle = \langle (B-A)Cx, Cx\rangle \geq 0$. Luego $C^*(B-A)C \geq 0$, es decir $C^*AC \leq C^*BC$. ■

Teorema 3.1.3. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces:*

1. *$A \in \mathcal{M}_n(\mathbb{C})^+$ si y sólo si existe $B \in \mathcal{M}_n(\mathbb{C})$ tal que $A = B^*B$.*

2. *En tal caso, existe una* **única** *matriz $B \in \mathcal{M}_n(\mathbb{C})^+$ tal que $A = B^*B = B^2$.*

Demostración. Sabemos que si $A = B^*B$ entonces $A \in \mathcal{M}_n(\mathbb{C})^+$. Luego basta probar que si $A \in \mathcal{M}_n(\mathbb{C})^+$, entonces existe una raiz cuadrada $B \in \mathcal{M}_n(\mathbb{C})^+$ para A, y que la tal B es única. Escribamos $A = UDU^*$, con $U \in \mathcal{U}(n)$ y $D = \operatorname{diag}(\lambda(A)) \in \mathcal{M}_n(\mathbb{C})^+$. Se toma

$$D^{1/2} := \operatorname{diag}\left(\lambda_1(A)^{1/2}, \ldots, \lambda_n(A)^{1/2}\right) \in \mathcal{M}_n(\mathbb{C})^+ .$$

Esto es posible por la Proposición 3.1.2. Es claro que $(D^{1/2})^2 = D$. Finalmente se define $B = UD^{1/2}U^*$. Luego $B \in \mathcal{M}_n(\mathbb{C})^+$ y $B^2 = A$. La unicidad es algo más complicada. El problema es que la matriz $U \in \mathcal{U}(n)$ que diagonaliza a A no es única. Sea otra $C \in \mathcal{M}_n(\mathbb{C})^+$ tal que $C^2 = A$. Entonces, como C y A conmutan, el Teorema 1 de Schur 1.6.1 asegura que

$$\text{existe } V \in \mathcal{U}(n) \text{ tal que } V^*AV = D \text{ y } V^*CV \in \mathcal{M}_n(\mathbb{C})^+ \text{ es diagonal} .$$

Para lo anterior se usa que $\mathcal{N}(n) \cap \mathcal{TS}(n)$ consta de las matrices diagonales, como se vió en la prueba del Teorema 2.1.2. Como $(V^*CV)^2 = V^*AV = D$, es claro que $V^*CV = D^{1/2}$ (entre diagonales la unicidad es trivial). Por otro lado,

$$UDU^* = VDV^* = A \implies (V^*U)D = D(V^*U) .$$

Aquí usaremos que $D^{1/2}$ se puede escribir como $P(D)$ para cierto polinomio $P \in \mathbb{R}[X]$. En efecto, basta elegir un P tal que $P(\lambda_i(A)) = \lambda_i(A)^{1/2}$, para todo $i \in \mathbb{I}_n$. Pero entonces V^*U conmuta con $P(D) = D^{1/2}$. Por ello $B = UD^{1/2}U^* = VD^{1/2}V^* = C$. ■

Observación 3.1.4 (El Grammiano)**.** Dada una matriz $B \in \mathcal{M}_n(\mathbb{C})$, llamemos $f_i = C_i(B)$, $i \in \mathbb{I}_n$. Notar que, entonces,

$$G(f_1, \ldots, f_n) := \Big(\langle f_i, f_j \rangle \Big)_{i,j=1}^{n} = B^*B \in \mathcal{M}_n(\mathbb{C})^+.$$

La matriz anterior es conocida como matriz de Gramm (o Grammiano) de $f_1, \ldots, f_n$. Del Teorema 3.1.3 deducimos que una matriz es semi definida positiva si y sólo si es una matriz de Gramm. Y que es definida positiva si y sólo si es una matriz de Gramm de un sistema linealmente independiente (en nuestro caso, esto equivale a que $B \in \mathcal{G}l(n)$).

Los mismos resultados son ciertos (la prueba es una ligera variación de la del Teorema 3.1.3) si la n-unpa $f_1, \dots, f_n$ vive en cualquier espacio de Hilbert $\mathbb{H}$ (anche infinitodimensional, porque en tal caso B es un operador $B : \mathbb{C}^n \to \mathbb{H}$, que es automáticamente continuo). Notar que la matriz de Gramm sigue estando en $\mathcal{M}_n(\mathbb{C})^+$ (y en $\mathcal{G}l(n)^+$ si el sistema es LI), donde n es el número de vectores en cuestión. ▲

Corolario 3.1.5 (Cholewsky)**.** *Sea $A \in \mathcal{M}_n(\mathbb{C})^+$. Entonces existe $T \in \mathcal{TS}(n)$ tal que $T_{ii} \geq 0$ para todo i, y tal que $A = T^*T$.*

Demostración. Sea $B \in \mathcal{M}_n(\mathbb{C})$ tal que $A = B^*B$. Sea $B = QT$, $Q \in \mathcal{U}(n)$, $T \in \mathcal{TS}(n)$, una descomposición de B como en el Teorema 1.8.2. Entonces $A = T^*Q^*QT = T^*T$. ■

3.2 Descomposición polar y valores singulares

Definición 3.2.1. 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 3.1.3. ▲

Observación 3.2.2. Sea $A \in \mathcal{M}_n(\mathbb{C})^+$. En la prueba del Teorema 3.1.3 se muestran las dos maneras usuales de describir a $A^{1/2}$:

1. Si $A = U \operatorname{diag}(\lambda(A))\, U^*$, con $U \in \mathcal{U}(n)$, se tiene que $A^{1/2} = U \operatorname{diag}\left(\lambda(A)^{1/2}\right) U^*$.
2. $A^{1/2} = P(A)$ para culquier $P \in \mathbb{C}[x]$ tal que $P(\lambda) = \lambda^{1/2}$ para todo $\lambda \in \sigma(A)$. ▲

Definición 3.2.3. Sea $A \in \mathcal{M}_n(\mathbb{C})$,

1. Llamaremos "módulo de A" a la matriz

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

2. Llamaremos **valores singulares** de A a los autovalores de $|A|$ ordenados en forma decreciente, notándolos $s_1(A) \geq \cdots \geq s_n(A) \geq 0$. Notar que, por el Corolario 1.7.2,

$$s_i(A) = \mu_i(|A|) = \mu_i(A^*A)^{1/2}, \qquad \text{para todo } i \in \mathbb{I}_n \ . \tag{3.1}$$

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

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

Observar que $|A| \cong \Sigma(A)$. ▲

Ejemplos 3.2.4. Sea $A \in \mathcal{M}_n(\mathbb{C})$.

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

2. Si $A \in \mathcal{M}_n(\mathbb{C})$ es normal, entonces $s(A) = |\lambda(A)|$, salvo el orden. En efecto, si $A = U \operatorname{diag}(\lambda(A))\, U^*$ para cierto $U \in \mathcal{U}(n)$, entonces

$$A^*A = U \operatorname{diag}\left(\overline{\lambda(A)}\right) \operatorname{diag}(\lambda(A))\, U^* = U \operatorname{diag}\left(|\lambda(A)|^2\right) U^* .$$

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 $s(A) = (1, \dots, 1, 0)$, porque A^*A es un proyector de rango $n-1$. ▲

Teorema 3.2.5 (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. *En particular, se tiene que* $\|A\|_{sp} = \|\,|A|\,\|_{sp} = \rho(|A|) = s_1(A) = \mu_1(A^*A)^{1/2}$.

3. *Existe una matriz unitaria* $U \in \mathcal{U}(n)$ *tal que*

$$A = U|A| ,$$

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

4. *Cualquier* $U \in \mathcal{U}(n)$ *que cumpla* $A = U|A|$, *verifica que*

$$A^* = U^*|A^*| , \quad AA^* = UA^*AU^* , \quad U|A|U^* = |A^*| \quad y \quad A = |A^*|\, U .$$

Esto dice que U^* *es un unitario admisible para la DP de* A^*. *O sea que* A *tiene una* **descomposición polar a derecha** $A = |A^*|\,U$ *con* **el mismo** U *que la otra.*

5. *Existen* $V, W \in \mathcal{U}(n)$ *tales que* $A = W\Sigma(A)V^*$.

6. *Las columnas* $C_i(V)$ *forman una BON de autovectores de* $|A|$ *(y* A^*A*), y las columnas* $C_i(W)$ *forman una BON de autovectores de* $|A^*|$ *(y* AA^**).*

Demostración.

1. Dado $x \in \mathbb{C}^n$, se tiene que

$$\|Ax\|^2 = \langle Ax, Ax \rangle = \langle A^*Ax, x \rangle = \langle |A|^2 x, x \rangle = \langle |A|x, |A|x \rangle = \|\,|A|x\|^2.$$

2. Se deduce de lo anterior, de la definición de norma espectral y del Corolario 2.1.4.

3. Podemos definir (con buena definición) una isometría suryectiva

$$U_1 : R(|A|) \to R(A) \;\text{ dada por }\; U_1(|A|x) = Ax, \text{ para todo } x \in \mathbb{C}^n \; .$$

De hecho, $|A|x = |A|y \iff x - y \in \ker |A| = \ker A \iff Ax = Ay$. Como

$$\dim R(A)^\perp = n - \dim R(A) = \dim \ker(A) = \dim \ker(|A|) = \dim R(|A|)^\perp \; ,$$

podemos extender la isometría U_1 a una matriz unitaria $U \in \mathcal{U}(n)$, operando isométricamente desde $R(|A|)^\perp$ sobre $R(A)^\perp$. Por la definición de U, se tiene que $A = U|A|$.

4. Notar que $AA^* = U|A|^2U^* = UA^*AU^*$. Sea $P(x) \in \mathbb{C}[x]$ tal que $P(\lambda) = \lambda^{1/2}$, para todo $\lambda \in \sigma(AA^*) = \sigma(A^*A)$ (acabamos de ver que $AA^* \cong A^*A$). Luego

$$|A^*| = (AA^*)^{1/2} = P(AA^*) = UP(A^*A)U^* = U|A|U^*.$$

Luego $A = U|A| = U|A|U^*U = |A^*|U$, por lo que también $A^* = U^*|A^*|$.

5. Sea $V \in \mathcal{U}(n)$ tal que $|A| = V\Sigma(A)V^*$. Si llamamos $W = UV \in \mathcal{U}(n)$, tenemos que

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

6. Notar que $\Sigma(A) = V^*|A|V$, por lo que cada $C_i(V)$ es un autovector de $|A|$, y todas las columnas de V forman una bon, por ser V unitaria. La prueba para W es similar, dado que también $\Sigma(A)^2 = W^*AA^*W$. ■

Existe una versión de la caracterización minimax de Courant-Fisher 2.3.3 para los valores singulares de una matriz:

Proposición 3.2.6. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ (no necesariamente autoadjunta). Con las mismas notaciones (para subespacios) que en el Teorema 2.3.3, se tiene que*

$$s_k(A) = \max_{\dim \mathcal{M}=k} \; \min_{x \in \mathcal{M}_1} \|Ax\| = \min_{\dim \mathcal{S}=n-k+1} \; \max_{x \in \mathcal{S}_1} \|Ax\| \quad \textit{para todo } k \in \mathbb{I}_n \; . \tag{3.2}$$

Demostración. 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 2.3.3 (y el Ejercicio 2.4.4 para traducirlo a μ'es) para A^*A. ■

Corolario 3.2.7. *Dadas $A, C \in \mathcal{M}_n(\mathbb{C})$, para todo $k \in \mathbb{I}_n$ se tiene que*

$$s_k(AC) \le \|A\|_{sp} \; s_k(C) \quad . \qquad \textit{En particular} \qquad \operatorname{tr}|AC| \le \|A\|_{sp} \; \operatorname{tr}|C| \; . \tag{3.3}$$

Demostración. Se usa la Eq. (3.2) para calcular $s_k(AC)$ y $s_k(C)$, junto con la siguiente desigualdad: $\|ACx\| \le \|A\|_{sp} \, \|Cx\|$, para todo $x \in \mathbb{C}^n$. ■

3.3 Parte positiva y parte negativa

Fijemos una matriz **autoadjunta** $A \in \mathcal{H}(n)$, y tomemos $A = U|A|$, con $U \in \mathcal{U}(n)$, una DP de A. Supongamos, además, que U opera como la identidad en $\ker A = \ker |A| = R(|A|)^{\perp}$. Una tal U existe por la construcción hecha en el Teorema 3.2.5, y además es única (Ejercicio: mostrar ambas cosas). Luego se verifican las siguientes propiedades:

1. Si $\mathcal{B} = \{v_1, \dots, v_n\}$ es una BON de $\mathbb{C}^n$ adaptada a $\mu(A)$, luego $A^*A = A^2$, $|A|$, $|A|^{1/2}$ y U son diagonales en la base $\mathcal{B}$. Por lo tanto conmutan entre ellos (y con A).

2. En la base $\mathcal{B}$, la matriz de U es diagonal con ± 1's en la diagonal. Más específicamente,

$$U v_k = v_k \quad \text{si} \quad \mu_k(A) \geq 0 , \quad \text{y} \quad U v_k = -v_k \quad \text{si } \mu_k(A) < 0 , \tag{3.4}$$

   dado que $|A|\, v_k = (\mu_k^2(A)\,)^{1/2}\, v_k = |\mu_k(A)|\, v_k$ para todo $k \in \mathbb{I}_n$. Por lo tanto,

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

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

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

4. Luego, si denotamos

$$A_+ = \frac{A + |A|}{2} \qquad \text{y} \qquad A_- = \frac{|A| - A}{2} , \tag{3.5}$$

   se prueba fácilmente que

   (a) Ambas matrices $A_+ , A_- \in \mathcal{M}_n(\mathbb{C})^+$.

   (b) $A = A_+ - A_-$ y $|A| = A_+ + A_-$.

   (c) $A_+A_- = A_-A_+ = 0$.

   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:

   (a) $AA_+ = A_+A = (A_+)^2$ (idem con A_-).

   (b) $(-A)_+ = A_-$ y $(-A)_- = A_+$.

   (c) Por la definición de A_+ y la fórmula (3.4), se tiene que

$$\mu_k(A_+) = \text{máx} \, \{ \, \mu_k(A) \, , \, 0 \, \} \quad , \quad \text{para todo } k \in \mathbb{I}_n \ . \tag{3.6}$$

   (d) $\mu_k(A_-) = \mu_k(\,(-A)_+) = \text{máx}\{\mu_k(-A), 0\} = -\,\text{mín}\{\mu_{n-k+1}(A), 0\}$, $k \in \mathbb{I}_n$.

6. Si $A = B - C$ con $B, C \in \mathcal{M}_n(\mathbb{C})^+$, entonces se tiene que
$$\mu_k(A_+) \leq \mu_k(B) \quad \text{y} \quad \mu_k(A_-) \leq \mu_k(C) \ , \quad \text{para todo } k \in \mathbb{I}_n \ . \tag{3.7}$$
En efecto, si $A < 0$, entonces $A_+ = 0$ y la primera desigualdad es obvia. Si $\mu_1(A) \geq 0$, sea $p = \max\{k \in \mathbb{I}_n : \mu_k(A) \geq 0\}$. Luego, como $B = C + A \geq A$, se tiene que
$$\mu_k(A_+) = \mu_k(A) \leq \mu_k(B) \ \text{ para } k \in \mathbb{I}_p \quad \text{y} \quad \mu_k(A_+) = 0 \leq \mu_k(B) \ \text{ para } k > p \ ,$$
por el Teorema de Weyl 2.3.5. La otra desigualdad en (3.7) se deduce de lo anterior aplicado a la igualdad $-A = C - B$, dado que $(-A)_+ = A_-$. ▲

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

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

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

Ejemplos 3.4.1. 1. La norma espectral $\|\cdot\|_{sp}$, definida del siguiente modo
$$\|A\| = \|A\|_{sp} = \max_{\|x\|=1} \|Ax\| = s_1(A),$$
donde la última igualdad surge de que $\|A\|_{sp} = \|\ |A|\ \|_{sp} = \rho(|A|)$.

2. Las normas de Schatten. Dado $1 \leq p < \infty$
$$\|A\|_p = \left(\sum_{i=1}^n s_i\,(A)^p \right)^{1/p} = (\operatorname{tr} |A|^p\,)^{1/p} \ .$$
La $\|\cdot\|_2$ se llama **norma de Frobenius**. Ella verifica que
$$\|A\|_2^2 = \operatorname{tr} A^*A = \sum_{i,j=1}^n |a_{ij}|^2$$
y proviene del producto escalar en $\mathcal{M}_n(\mathbb{C})$ definido por $\langle A, B\rangle = \operatorname{tr} B^*A$.

3. Las normas Ky-Fan. Dado $k \in \{1, \dots, n\}$
$$\|A\|_{(k)} = \sum_{i=1}^k s_i\,(A) \,.$$
Notar que $\|A\|_{(1)} = \|A\|_{sp}$ y $\|A\|_{(n)} = \|A\|_1$ (de Schatten).

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

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

Estas normas satisfacen que:

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

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

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

Ejercicio 3.4.2. Consideremenos en $\mathbb{C}^n$ las siguientes normas:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \quad \text{y} \quad \|x\|_\infty = \max_{i \in \mathbb{I}_n} |x_i| \quad , \qquad \text{para todo } x \in \mathbb{C}^n \ ,$$

conocidas como las normas 1 e ∞. Como en el Ejemplo anterior, ellas inducen en $\mathcal{M}_n(\mathbb{C})$ las siguientes normas matriciales: Dada $A \in \mathcal{M}_n(\mathbb{C})$,

$$|||A|||_\infty = \max_{\|x\|_\infty = 1} \|Ax\|_\infty \quad \text{y} \quad |||A|||_1 = \max_{\|x\|_1 = 1} \|Ax\|_1 \ .$$

Probar que estas normas pueden calcularse efectivamente mediante las fórmulas:

$$|||A|||_\infty = \max_{i \in \mathbb{I}_n} \|F_i(A)\|_1 \quad \text{y} \quad |||A|||_1 = \max_{i \in \mathbb{I}_n} \|C_i(A)\|_1 \ . \tag{3.8}$$

para toda $A \in \mathcal{M}_n(\mathbb{C})$. ▲

Definición 3.4.3. Una norma $\|\cdot\|$ en $\mathcal{M}_n(\mathbb{C})$ se llama:

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

Ejemplo 3.4.4. Sea $N_\infty(A) = \max_{ij \in \mathbb{I}_n} |a_{ij}|$, para $A \in \mathcal{M}_n(\mathbb{C})$. Sean $\mathbb{1}_n = (1, \dots, 1) \in \mathbb{C}^n$ y $\mathbb{E}_n = \mathbb{1}_n \odot \mathbb{1}_n$. Como $\langle \mathbb{1}_n \, , \mathbb{1}_n \rangle = n$, tenemos que

$$\mathbb{E}_n^2 = \mathbb{1}_n \odot \mathbb{1}_n \cdot \mathbb{1}_n \odot \mathbb{1}_n = \langle \mathbb{1}_n \, , \mathbb{1}_n \rangle \; \mathbb{1}_n \odot \mathbb{1}_n = n\mathbb{E}_n \implies N_\infty(\mathbb{E}_n^2) = n \ ,$$

mientras que $N_\infty(\mathbb{E}_n) = 1$. O sea que N_∞ **no** es matricial en $\mathcal{M}_n(\mathbb{C})$ para ningún $n > 1$. El lector interesado puede demostrar que $n \, N_\infty(\cdot)$ sí es una norma matricial en $\mathcal{M}_n(\mathbb{C})$. ▲

Teorema 3.4.5. *Sea* $\|\cdot\|$ *una norma* **matricial** *en* $\mathcal{M}_n(\mathbb{C})$. *Dada* $A \in \mathcal{M}_n(\mathbb{C})$ *se tiene que*

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

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

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

Si valiera 1, luego $B^{-1}A$ sería inversible y $A = B(B^{-1}A)$ también debería serlo. Para probar el ítem 1, llamemos $C = I - A$. Tenemos que

$$\|C\| < 1 \implies \|C^m\| \leq \|C\|^m \ \ \forall \ \ m \in \mathbb{N} \implies \sum_{k=0}^{\infty} \|C^k\| \leq \sum_{k=0}^{\infty} \|C\|^k = \frac{1}{1 - \|C\|} \quad .$$

Luego, la serie $\displaystyle\sum_{k=1}^{\infty} C^k$ converge. En particular, $C^k \xrightarrow[k\to\infty]{} 0$. Luego

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

Analogamente $\left(\sum\limits_{k=0}^{\infty} C^k\right) A = 1$. ■

Proposición 3.4.6. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ y $\| \cdot \|$ una norma matricial. Entonces $\rho(A) \leq \|A\|$. Más aún, $\|A^m\|^{1/m} \geq \rho(A)$ para todo $m \in \mathbb{N}$.*

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

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

de donde se deduce que $|\lambda| \leq \|A\|$. Como el autovalor era cualquiera, $\rho(A) \leq \|A\|$. Además, por el Corolario 1.7.2, se tiene que $\sigma(A^m) = \sigma(A)^m$, y entonces también $\rho(A^m) = \rho(A)^m$. Por lo tanto, usando la parte ya probada, obtenemos que $\rho(A) \leq \|A^m\|^{1/m}$. ■

Observación 3.4.7. Dada una norma matricial $\| \cdot \|$ en $\mathcal{M}_n(\mathbb{C})$ y una matriz $S \in \mathcal{G}l(n)$, la fórmula $\|A\|_S := \|SAS^{-1}\|$, $A \in \mathcal{M}_n(\mathbb{C})$, define otra norma matricial. ▲

Proposición 3.4.8. *Dados $A \in \mathcal{M}_n(\mathbb{C})$ y $\varepsilon > 0$, existe una norma matricial $N_{A,\varepsilon}$ en $\mathcal{M}_n(\mathbb{C})$ tal que $N_{A,\varepsilon}(A) \leq \rho(A) + \varepsilon$.*

Demostración. Sea $A = UTU^*$ con T una matriz triangular superior y $U \in \mathcal{U}(n)$. Luego, $T = U^*AU$. Sea $D_s = \operatorname{diag}(s, s^2, \ldots, s^n)$. Entonces, $(D_s T D_s^{-1})_{ij} = t_{ij} \; s^{i-j}$ para todo par $i, j \in \mathbb{I}_n$ (eso fue visto en la Eq. (1.14)). Por lo tanto,

$$D_s T D_s^{-1} \xrightarrow[s\to\infty]{\text{En cualquier norma}} \operatorname{diag}(\lambda_1(A), \ldots, \lambda_n(A)) \, .$$

Como $\|\operatorname{diag}(\lambda_1(A), \ldots, \lambda_n(A))\|_{sp} = \rho(A)$, entonces, $\|D_s T D_s^{-1}\|_{sp} \xrightarrow[s\to\infty]{} \rho(A)$. Luego debe existir un $s_0 \in \mathbb{R}$ tal que $\|D_{s_0} \, T \, D_{s_0}^{-1}\|_{sp} < \rho(A) + \varepsilon$. Consideremos ahora la norma

$$N_{A,\varepsilon} = \big(\| \cdot \|_{sp}\big)_{D_{s_0} U^*} \, , \qquad \text{o sea} \qquad N_{A,\varepsilon}(B) = \|D_{s_0} \, U^* \, B \, U \, D_{s_0}^{-1}\|_{sp} \; , \quad B \in \mathcal{M}_n(\mathbb{C}).$$

Luego $N_{A,\varepsilon}(A) = \|D_{s_0} \, U^* \, A \, U \, D_{s_0}^{-1}\|_{sp} = \|D_{s_0} \, T \, D_{s_0}^{-1}\|_{sp} < \rho(A) + \varepsilon$. ■

Corolario 3.4.9. *Si* $A \in \mathcal{M}_n(\mathbb{C})$*, entonces* $A^m \to 0$ *si y sólo si* $\rho(A) < 1$.

Demostración. Es claro que $\rho(A) < 1$ si y sólo si $\rho(A^m) = \rho(A)^m \xrightarrow[m\to\infty]{} 0$. Usando que $\rho(A^m) \le \|A^m\|_{sp}$ para todo $m \in \mathbb{N}$, una implicación es clara. Para probar la otra, supongamos que $\rho(A) < 1$. Por la Proposición 3.4.8, existe una norma matricial N tal que $\rho(A) \le N(A) < 1$. Como N es matricial, deducimos que $N(A^m) \le N(A)^m \xrightarrow[m\to\infty]{} 0$. ■

Teorema 3.4.10. *Sea* $A \in \mathcal{M}_n(\mathbb{C})$*. Entonces, para cualquier norma* $\|\cdot\|$ *en* $\mathcal{M}_n(\mathbb{C})$*,*

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

Demostración. Supongamos primero que $\|\cdot\|$ es una norma matricial. Por la Proposición 3.4.6, sabemos que se tiene $\rho(A) \le \|A^m\|^{1/m}$ para todo $m \in \mathbb{N}$. Fijado $\varepsilon > 0$, consideremos la matriz $B = \dfrac{A}{\rho(A)+\varepsilon}$. Como $\rho(B) < 1$, el Corolario 3.4.9 asegura que $\|B^m\| \to 0$. En consecuencia existe un $m_0 \in \mathbb{N}$ tal que, para todo $m \ge m_0$, se verifica

$$\|B^m\| < 1 \ , \quad \text{es decir que} \quad \|A^m\| < (\rho(A)+\varepsilon)^m \implies \|A^m\|^{1/m} < \rho(A)+\varepsilon \ ,$$

lo que prueba el Teorema en este caso. El mismo resultado vale para normas no matriciales, por ser todas las normas equivalentes. ■

Ejercicio 3.4.11. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Si N es una norma matricial en $\mathcal{M}_n(\mathbb{C})$, mostrar que $\rho(A) = \inf_{m\in\mathbb{N}} N(A^m)^{1/m}$. Más aún, probar que en tal caso, $N(A^m)^{1/m} \underset{m\to\infty}{\searrow} \rho(A)$. ▲

Observación 3.4.12. Todos los resultados de esta sección, a partir del Teorema 3.4.5, son también ciertos en álgebras de Banach, donde las normas son matriciales por definición. El único resultado propio de matrices es la Proposición 3.4.8, que nos permite dar una prueba fácil de la fórmula del radio espectral (Teorema 3.4.10). Esta fórmula vale también en dimensión infinita, y la prueba usa herramientas de análisis complejo. El curro es mostrar que la llamada *resolvente*, que es la función

$$\rho_A : \mathbb{C} \setminus \sigma(A) \to \mathcal{G}l(n) \quad \text{dada por} \quad \rho_A(z) = (zI - A)^{-1} \ , \quad z \in \mathbb{C} \setminus \sigma(A) \ ,$$

es analítica. La fórmula dada surge del radio de convergencia de su serie de potencias alrededor del "infinito". Sin embargo, hemos incluido las demostraciones anteriores porque tienen un buen sabor matricial, salvo el Teorema 3.4.5, que tiene la prueba standard (y no creo que pueda mejorarse). Nótese que se asumió implícitamente que $\mathcal{M}_n(\mathbb{C})$ es un espacio completo, porque usamos que una serie absolutamente sumable es convergente. ▲

3.5 Algunas caracterizaciones

A continuación daremos algunas caracterizaciones fáciles de la positividad y la contractividad de matrices. Al final incluimos una mini-introducción al producto de Hadamard, mostrando el Teorema de Schur 2. A lo largo de ésta Sección abreviaremos $\|A\|_{sp} = \|A\|$. Usaremos la Proposición 1.5.5 que dice que dadas $A, B \in \mathcal{M}_n(\mathbb{C})$, entonces $\sigma(AB) = \sigma(BA)$. El primer enunciado resume varias caracterizaciones ya mencionadas de esta norma.

Lema 3.5.1. *Sea* $A \in \mathcal{M}_{n,m}(\mathbb{C})$ *entonces*

$$s_1(A) = \|A\| = \|\,|A|\,\| = \rho(|A|) = \rho(A^*A)^{1/2} = \|A^*A\|^{1/2} = \|AA^*\|^{1/2} \ . \tag{3.9}$$

Demostración. Como $|A|$ y $A^*A \in \mathcal{H}(n)$, la Proposición 2.1.4 asegura que

$$\|\,|A|\,\| = \rho(|A|) = s_1(A) \quad \text{ y que } \quad \rho(A^*A)^{1/2} = \|A^*A\|^{1/2} \ .$$

Las igualdades $\|A\| = \|\,|A|\,\| = s_1(A)$ se deducen de que $\|Ax\| = \|\,|A|x\|$ para todo $x \in \mathbb{C}^n$ (ítem 1 del Teorema 3.2.5). La igualdad $\rho(|A|) = \rho(A^*A)^{1/2}$ se sigue del Corolario 1.7.2, usando que $|A|^2 = A^*A$. Observar que $\rho(A^*A) = \rho(AA^*)$ porque $\sigma(A^*A) = \sigma(AA^*)$. ■

Proposición 3.5.2. *Sea* $A \in \mathcal{H}(n)$*, entonces* $-\|A\|\ I \le A \le \|A\|\ I$ *. Más aún,*

$$-\lambda I \le A \le \lambda\ I \iff \|A\| \le \lambda \iff \rho(A) \le \lambda \ ,$$

para cialquier $\lambda \in \mathbb{R}_+^*$.

Demostración. Notar que si $A \in \mathcal{H}(n)$, entonces $\|A\| = \rho(A) = \text{máx}\{\mu_1(A), -\mu_n(A)\}$. Por lo tanto, $\rho(A) \le \lambda \iff -\lambda \le \mu_n(A) \le \mu_1(A) \le \lambda$. Por el Teorema 2.3.1, tenemos que

$$-\lambda \le \mu_n(A) = \min_{\|x\|=1} \langle Ax, x\rangle \iff -\lambda I \le A \quad \text{ y además}$$

$$\max_{\|x\|=1} \langle Ax, x\rangle = \mu_1(A) \le \lambda \iff A \le \lambda I \ . \quad ■$$

Proposición 3.5.3. *Dada* $A \in \mathcal{M}_n(\mathbb{C})$*, se tienen las equivalencias*

$$\|A\| = s_1(A) \le 1 \iff |A| \le I \iff AA^* \le I \iff A^*A \le I \ . \tag{3.10}$$

Demostración. Es consecuencia del Lema 3.5.1 y de la Proposición 3.5.2. ■

Proposición 3.5.4. *Si* $A \in \mathcal{M}_n(\mathbb{C})^+$ *y* $B \in \mathcal{G}l(n)^+$*, entonces*

$$A \le B \iff \|A^{1/2}B^{-1/2}\| \le 1 \iff \rho(AB^{-1}) \le 1 \ . \tag{3.11}$$

Demostración. Notemos que

$$A \le B \iff B^{-1/2}AB^{-1/2} \le I \iff (A^{1/2}B^{-1/2})^*A^{1/2}B^{-1/2} \le I \ .$$

Luego se aplica la Proposición 3.5.3 y el hecho de que $\sigma\left(B^{-1/2}AB^{-1/2}\right) = \sigma\left(AB^{-1}\right)$. ■

3.5.5. Sea $x \in \mathbb{C}^n$ con $\|x\| = 1$ (a estos vectores los llamaremos *unitarios*). Entonces, como vimos en 1.9.3, la matriz $P_x = x \odot x = xx^* = (x_i\overline{x_j})_{ij} \in \mathcal{M}_n(\mathbb{C})^+$ es el proyector ortogonal sobre el subespacio Gen $\{x\}$. Por lo tanto, si $\mathcal{B} = \{x_1, \dots, x_n\}$ es una BON de $\mathbb{C}^n$, vale que

$$z = \sum_{i=1}^n \langle z, x_i\rangle x_i \quad \text{para todo } \ z \in \mathbb{C}^n \qquad \Longrightarrow \qquad I = \sum_{i=1}^n x_i \odot x_i \ , \tag{3.12}$$

por lo que $\{x_i \odot x_i \, : i \in \mathbb{I}_n\}$ es un sistema de proyectores (ver Definición 5.4.2). ▲

Proposición 3.5.6. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Las siguientes condiciones son equivalentes:*

1. $A \in \mathcal{M}_n(\mathbb{C})^+$.
2. *Existen* $y_1, \dots, y_r \in \mathbb{C}^n$ *tales que* $A = \sum\limits_{i=1}^{r} y_i \odot y_i = \sum\limits_{i=1}^{r} y_i \cdot y_i^*$.

Demostración. La implicación $2 \to 1$ es clara, porque cada matriz $y_i \cdot y_i^* \in \mathcal{M}_n(\mathbb{C})^+$. Recíprocamente, si $A \in \mathcal{M}_n(\mathbb{C})^+$, sea $\mathcal{B} = \{x_1, \dots, x_n\}$ es una BON de $\mathbb{C}^n$ adaptada a $\mu(A)$. Usando la ecuación (3.12), para todo $z \in \mathbb{C}^n$ se tiene que

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

Luego basta elegir $y_i = \mu_i(A)^{1/2}\, x_i$ para aquellos $i \in \mathbb{I}_n$ tales que $\mu_i(A) > 0$. ■

3.6 El producto de Hadamard

Definición 3.6.1. Dadas $A, B \in \mathcal{M}_{n,m}(\mathbb{C})$, su producto de Hadamard $A \circ B$ es la matriz

$$A \circ B = \big(a_{ij}\ b_{ij}\big)_{\substack{i \in \mathbb{I}_n \\ j \in \mathbb{I}_m}} \in \mathcal{M}_{n,m}(\mathbb{C}) \ .$$

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

A este producto de matrices, también llamado producto de Schur, le dedicaremos un capítulo entero, porque tiene interesantísimas aplicaciones dentro y fuera de la teoría del Análisis Matricial. Pero vamos adelantando un resultado al respecto (otro teorema de Schur), porque es elemental y compete a las matrices positivas.

Teorema 3.6.2 (Teorema 2 de Schur)**.** *Sean $A,\ B \in \mathcal{M}_n(\mathbb{C})^+$, entonces $A \circ B \in \mathcal{M}_n(\mathbb{C})^+$. Además, si $A, B \in \mathcal{G}l(n)^+$, entonces $A \circ B \in \mathcal{G}l(n)^+$.*

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

$$A \circ B \geq aI \circ B \geq aI \circ bI = ab\, I \in \mathcal{G}l(n)^+ \ . \tag{3.13}$$

Supongamos entonces que $A,\ B \in \mathcal{M}_n(\mathbb{C})^+$. Por la Proposición 3.5.6 (ver también 1.9.3), deben existir vectores $v_i \in \mathbb{C}^n$, $i \in \mathbb{I}_r$, tales que $A = \sum\limits_{i=1}^{r} v_i\, v_i^*$. Como el producto $\circ$ es distributivo, basta mostrar que $v\, v^* \circ B \in \mathcal{M}_n(\mathbb{C})^+$ para todo $v \in \mathbb{C}^n$ y toda $B \in \mathcal{M}_n(\mathbb{C})^+$. Y para ver esto, alcanza con hacer la siguiente cuenta:

$$v\, v^* \circ B \ = \ \Big(v_i\, \overline{v_j}\ B_{ij}\Big)_{i,j \in \mathbb{I}_n} \ = \ \operatorname{diag}(v)\ B\ \operatorname{diag}(v)^* \in \mathcal{M}_n(\mathbb{C})^+ \ ,$$

donde la igualdad del medio se testea haciendo la cuenta, o mirando la Eq. (1.14). ■

Corolario 3.6.3. *Sean* $A,\ B \in \mathcal{M}_n(\mathbb{C})^+$, *entonces*

1. $\mu_n(A)\mu_n(B) \leq \mu_n(A \circ B)$.
2. $\|A \circ B\|_{sp} = \mu_1(A \circ B) \leq \mu_1(A)\mu_1(B) = \|A\|_{sp}\, \|B\|_{sp}$.

Demostración. La primera desigualdad se deduce de la ecuación (3.13), pero usando que $A \geq \mu_n(A)I$ y $B \geq \mu_n(B)I$. La segunda, de una cuenta análoga, pero aplicando ahora las desigualdades $A \leq \mu_1(A)I$ y $B \leq \mu_1(B)I$ (tadas fueron vistas en la Observación 2.3.2). ∎

Corolario 3.6.4. *Si* $A \in \mathcal{M}_n(\mathbb{C})^+$, *entonces* $B = \left(|A_{ij}|^2\right)_{i,j\in\mathbb{I}_n} \in \mathcal{M}_n(\mathbb{C})^+$.

Demostración. Se deduce de que $A^T = \overline{A} = \left(\,\overline{A_{ij}}\,\right)_{i,j\in\mathbb{I}_n} \in \mathcal{M}_n(\mathbb{C})^+$. ∎

Ejercicio 3.6.5. Mostrar que el resultado anterior falla si uno no eleva los módulos al cuadrado. En otras palabras, se debe encontrar un ejemplo de una matriz $A \in \mathcal{M}_n(\mathbb{C})^+$ tal que $B = \left(|A_{ij}|\right)_{i,j\in\mathbb{I}_n} \notin \mathcal{M}_n(\mathbb{C})^+$. Observar que hay que buscar para $n \geq 3$. ▲

Corolario 3.6.6. *Si* $A \in \mathcal{M}_n(\mathbb{C})^+$ *y* $P(x) \in \mathbb{R}[x]$ *tiene coeficientes no negativos, entonces*

$$P_\circ(A) := \Big(\, P(A_{ij})\,\Big)_{i,j\in\mathbb{I}_n} \in \mathcal{M}_n(\mathbb{C})^+ \ .$$

Demostración. Por una inducción directa, podemos ver que $A^{[k]} = A \circ A \circ \cdots \circ A \in \mathcal{M}_n(\mathbb{C})^+$ (se multiplica k veces) para todo $k \in \mathbb{N}$. Después se usa lo que cumple $P(x)$. ∎

Ejercicio 3.6.7. Probar que, si $A \in \mathcal{M}_n(\mathbb{C})^+$, entonces $e_\circ^A := \Big(\, e^{A_{ij}}\,\Big)_{i,j\in\mathbb{I}_n} \in \mathcal{M}_n(\mathbb{C})^+$. ▲

3.7 El famoso truco 2×2

Cuando se trabaja con operadores y matrices, muchas veces una cuenta inmanejable termina saliendo "mágicamente" y en un par de renglones, con el famoso truco de matrices de bloques de 2×2. Ver, por ejemplo, la Proposición 1.5.5, y tratar de probarla de otra manera. En esta sección juntaremos varios resultados que aportan técnicas para usar dicho método. Para operar entre matrices de bloques, se usarán sistemáticamente los resultados desarrollados en la Sección 1.5. En particular, las fórmulas (1.16), (1.17) y (1.18).

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

$$A \in \mathcal{M}_n(\mathbb{C})^+ \iff A^{(2)} = \begin{bmatrix} A & A \\ A & A \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})^+.$$

En efecto, Si tomamos la matriz

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} -I & I \\ I & I \end{bmatrix} \in \mathcal{U}(2n),$$

cuentas elementales muestran que $U = U^* = U^{-1}$, y que

$$UA^{(2)}U^* = UA^{(2)}U = \begin{bmatrix} 0 & 0 \\ 0 & 2A \end{bmatrix} .$$

Ahora sí es claro que $A \geq 0$ si y sólo si $A^{(2)} \geq 0$. Dejamos como ejercicio la verificación de que si $A^{(k)} \in \mathcal{M}_{kn}(\mathbb{C})$ se define en foma semejante a $A^{(2)}$, entonces $A \geq 0$ si y sólo si $A^{(k)} \geq 0$. ▲

3.7.2. Si $A \in \mathcal{M}_n(\mathbb{C})$, entoces $B = \begin{bmatrix} |A^*| & A \\ A^* & |A| \end{bmatrix} \geq 0$. En efecto, sea $U \in \mathcal{U}(n)$ tal que $A = U|A|$. Entonces

$$\begin{aligned} 0 \ \leq \begin{bmatrix} U & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} |A| & |A| \\ |A| & |A| \end{bmatrix} \begin{bmatrix} U^* & 0 \\ 0 & I \end{bmatrix} &= \begin{bmatrix} U|A| & U|A| \\ |A| & |A| \end{bmatrix} \begin{bmatrix} U^* & 0 \\ 0 & I \end{bmatrix} \\ = \begin{bmatrix} U|A|U^* & U|A| \\ |A|U^* & |A| \end{bmatrix} &= \begin{bmatrix} |A^*| & A \\ A^* & |A| \end{bmatrix} , \end{aligned}$$

dado que $U|A|U^* = |A^*|$. El mismo resultado sigue valiendo si $A \in \mathcal{M}_{nm}(\mathbb{C})$, o sea si A es rectangular. En ese caso $B \in \mathcal{M}_{n+m}(\mathbb{C})^+$ (Ejercicio). ▲

Proposición 3.7.3. *Sea $A \in \mathcal{M}_{n,m}(\mathbb{C})$, y llamemos $r = \text{mín}\{n, m\}$. Luego*

$$s_k(A^*) = s_k(A) \quad \textit{para todo} \quad k \in \mathbb{I}_r \ . \tag{3.14}$$

Demostración. Como vimos en la Observación 1.5.6, $\mu(AA^*) = \mu(A^*A)$ salvo una cola de $m - n$ (o $n - m$) ceros. Usando el Corolario 1.7.2 (o sea que $\lambda(P(A)) = P(\lambda(A))$ para todo polinomio P) y la definición de $|A|$, vemos que $\mu(A^*A) = \mu(|A|^2) = \mu(|A|)^2$. De ahí sale que $s(A) = s(A^*)$ salvo los ceros finales. Esto muestra la fórmula (3.14). ■

Observación 3.7.4 (El rango)**.** Recordemos que, si $A \in \mathcal{M}_{n,m}(\mathbb{C})$ decimos que

$$\text{rk}\, A = \dim R(A) = \dim \text{Gen}\, \{C_1(A), \dots, C_m(A)\} \ ,$$

lo que usualmente se llama **rango columna** de A. Llamemos $r = \text{mín}\{n, m\}$. Sea $U \in \mathcal{U}(n)$ tal que $A = U|A|$. Es fácil ver que

$$\text{rk } A = \text{rk } |A| = \text{rk } \Sigma(A) = \ \text{máx}\{k \in \mathbb{I}_r \ : \ s_k(A) \neq 0\} \ . \tag{3.15}$$

El **rango fila** de A es, con esta definición, la $\dim \text{Gen}\, \{F_1(A), \dots, F_n(A)\} = \text{rk}\, A^T = \text{rk}\, A^*$. Por la Proposición 3.7.3, $s(A^*) = s(A)$ (salvo ceros finales). Luego la fórmula (3.15) muestra que ambos rangos son iguales. ▲

Proposición 3.7.5. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces*

$$\widehat{A} := \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix} \cong \begin{bmatrix} \Sigma(A) & 0 \\ 0 & -\Sigma(A) \end{bmatrix} \in \mathcal{H}(2n) \ .$$

En particular, $\sigma(\widehat{A}) = \{\pm s_i(A)\}$ (con las mismas multiplicidades). Es decir,

$$\mu(\widehat{A}) = (s_1(A), \cdots, s_n(A), -s_n(A), \cdots, -s_1(A)). \tag{3.16}$$

Demostración. Sean $U, V \in \mathcal{U}(n)$ tales que $\Sigma(A) = VAU^* = UA^*V^*$. Es fácil ver que

$$W = \frac{1}{\sqrt{2}} \begin{bmatrix} V & U \\ -V & U \end{bmatrix} \in \mathcal{U}(2n).$$

Entonces

$$\begin{aligned} W\widehat{A}\,W^* &= \frac{1}{2} \begin{bmatrix} UA^* & VA \\ UA^* & -VA \end{bmatrix} \begin{bmatrix} V^* & -V^* \\ U^* & U^* \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} UA^*V^* + VAU^* & VAU^* - UA^*V^* \\ UA^*V^* - VAU^* & -VAU^* - UA^*V^* \end{bmatrix} \\ &= \begin{bmatrix} \Sigma(A) & 0 \\ 0 & -\Sigma(A) \end{bmatrix}, \end{aligned}$$

como queríamos demostrar. ■

Proposición 3.7.6. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces

$$\|A\|_{sp} \leq 1 \quad \iff \quad M = \begin{bmatrix} I & A \\ A^* & I \end{bmatrix} \geq 0.$$

Demostración. Notar que $M = I_{2n} + \widehat{A}$. Usando que $\sigma(\widehat{A}) = \sigma(-\widehat{A})$ (por la Proposición 3.7.5) y luego el Teorema de Rayleigh-Ritz 2.3.1 (o la Observación 2.3.2), obtenemos que

$$\begin{bmatrix} I & A \\ A^* & I \end{bmatrix} \geq 0 \iff I_{2n} + \widehat{A} \geq 0 \iff -\widehat{A} \leq I_{2n} \iff -I_{2n} \leq \widehat{A} \leq I_{2n}\ .$$

Por la Proposición 3.5.2, esto equivale a que $\|A\|_{sp} = s_1(A) = \rho(\widehat{A}) = \|\widehat{A}\|_{sp} \leq 1$. ■

Observación 3.7.7. Notar que la Proposición 3.7.6 sigue siendo válida si A es rectangular, por el simple recurso de agregarle ceros (arriba o a la derecha) para que A quede cuadrada, lo que no cambia su norma. ▲

3.7.8. Si $A, B \in \mathcal{M}_n(\mathbb{C})^+$, entonces son equivalentes

1. $A \leq B$.
2. La matriz $M = \begin{bmatrix} B & A \\ A & B \end{bmatrix} \geq 0$.

En efecto, si $B \in \mathcal{G}l(n)^+$, entonces $M \geq 0$ si y sólo si

$$0 \leq \begin{bmatrix} B^{-1/2} & 0 \\ 0 & B^{-1/2} \end{bmatrix} \begin{bmatrix} B & A \\ A & B \end{bmatrix} \begin{bmatrix} B^{-1/2} & 0 \\ 0 & B^{-1/2} \end{bmatrix} = \begin{bmatrix} I & B^{-1/2}AB^{-1/2} \\ B^{-1/2}AB^{-1/2} & I \end{bmatrix},$$

lo que, por la Proposición 3.7.6, equivale a que $\|A^{1/2}B^{-1/2}\|^2 = \|B^{-1/2}AB^{-1/2}\| \leq 1$. Por la Proposición 3.5.2, se tiene que $\|A^{1/2}B^{-1/2}\| \leq 1$ si y sólo si $A \leq B$. Un ejercicio fácil es deducir que la equivalencia sigue valiendo si no se pide que B sea inversible (si uno cambia B por $B + \varepsilon I$, entonces M pasa a ser $M + \varepsilon I_{2n}$). ▲

3.7.9. Sea $A \in \mathcal{M}_n(\mathbb{C})$ una contracción, es decir que $\|A\|_{sp} \leq 1$.

1. Se tiene que $A^*(I - AA^*)^{1/2} = (I - A^*A)^{1/2} A^*$.

2. Entonces las matrices

$$\begin{bmatrix} A & (I - AA^*)^{1/2} \\ (I - A^*A)^{1/2} & -A^* \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} A & -(I - AA^*)^{1/2} \\ (I - A^*A)^{1/2} & A^* \end{bmatrix}$$

son unitarias en $\mathcal{M}_{2n}(\mathbb{C})$.

En efecto, observar que $A^*(I - AA^*) = A^* - A^*AA^* = (I - A^*A)A^*$. Por inducción vemos que $A^*(I - AA^*)^k = (I - A^*A)^k A^*$ para todo $k \in \mathbb{N} \cap \{0\}$. Luego se usa que a $(I - AA^*)^{1/2}$ se lo puede realizar como un polinomio en $(I - AA^*)$, y lo mismo para $(I - A^*A)$, con el mismo polinomio, dado que tienen el mismo espectro. La verificación la segunda parte es directa, y se deja como ejercicio. ▲

3.7.10. Sea $M \in \mathcal{H}(n)$, y representémosla por bloques $M = \begin{array}{c} \begin{bmatrix} A & C \\ C^* & B \end{bmatrix} \end{array} \begin{array}{l} k \\ n-k \end{array}$.

1. Si $A = \lambda I_k$ y $B = \mu I_{n-k}$ para ciertos $\lambda, \mu \in \mathbb{R}_+^*$, se tiene que

$$M = \begin{bmatrix} \lambda I_k & C \\ C^* & \mu I_{n-k} \end{bmatrix} \geq 0 \iff C\,C^* \leq \lambda\,\mu\, I_k \iff \|C\|^2 \leq \lambda\,\mu\ .$$

2. En el caso general, dado $\varepsilon > 0$, existe $\lambda > 0$ tal que

$$\begin{bmatrix} A & C \\ C^* & B \end{bmatrix} \leq \begin{bmatrix} A + \varepsilon I_k & 0 \\ 0 & \lambda I_{n-k} \end{bmatrix}.$$

En efecto, si $M = \begin{bmatrix} \lambda I_k & C \\ C^* & \mu I_{n-k} \end{bmatrix}$, conjugandola con $D = \begin{bmatrix} \lambda^{-1/2} I_k & 0 \\ 0 & \mu^{-1/2} I_{n-k} \end{bmatrix}$, caemos en el caso de la Proposición 3.7.6 (para $C \in \mathcal{M}_{k,n-k}(\mathbb{C})$ rectangular, donde también es cierto). Luego, por las citas que se indican sobre los símbolos,

$$\begin{aligned} M \geq 0 \quad &\iff DMD = \begin{bmatrix} I_k & \lambda^{-1/2}\mu^{-1/2} C \\ \lambda^{-1/2}\mu^{-1/2} C^* & I_{n-k} \end{bmatrix} \geq 0 \\ &\overset{\text{Prop. 3.7.6}}{\iff} \lambda^{-1}\,\mu^{-1} \|C\|^2 \overset{\text{Lema 3.5.1}}{=} \lambda^{-1}\,\mu^{-1} \|C\,C^*\| \leq 1 \\ &\overset{\text{Prop. 3.5.3}}{\iff} C\,C^* \leq \lambda\,\mu\, I_k\ . \end{aligned}$$

Para la segunda parte, basta tomar $\lambda \geq \dfrac{\|C\|^2}{\varepsilon} + \|B\|$. En efecto, en ese caso,

$$B \leq \|B\| I_{n-k} \implies \lambda\, I_{n-k} - B \geq \big(\lambda - \|B\|\big) I_{n-k} \geq \frac{\|C\|^2}{\varepsilon}\, I_{n-k}\ ,$$

y, si llamamos $m = n - k$, se aplica el caso anterior a la matriz

$$0 \leq \begin{bmatrix} \varepsilon I_k & -C \\ -C^* & \frac{\|C\|^2}{\varepsilon} I_m \end{bmatrix} \leq \begin{bmatrix} \varepsilon I_k & -C \\ -C^* & \lambda\, I_m - B \end{bmatrix} = \begin{bmatrix} A + \varepsilon I_k & 0 \\ 0 & \lambda I_m \end{bmatrix} - \begin{bmatrix} A & C \\ C^* & B \end{bmatrix}\ . \qquad ▲$$

3.8 Cortocircuitos

Lema 3.8.1. *Sean $D, A \in \mathcal{M}_n(\mathbb{C})^+$. Luego, las siguientes condiciones son equivalentes:*

1. $D \leq A$.
2. $\|D^{1/2}x\| \leq \|A^{1/2}x\|$ *para todo* $x \in \mathbb{C}^n$.
3. *Existe* $C \in \mathcal{M}_n(\mathbb{C})$ *tal que* $\|C\|_{sp} \leq 1$ *y* $D^{1/2} = CA^{1/2}$.

Demostración. Observar que $\|D^{1/2}x\|^2 = \langle D^{1/2}x\,, D^{1/2}x\rangle = \langle Dx\,, x\rangle$ y lo mismo vale para A. Esto da la equivalencia $1 \leftrightarrow 2$. El hecho de que $3 \to 1$ se deduce de que $\|C\|_{sp} \leq 1 \Rightarrow C^*C \leq I$ (por la Proposición 3.5.3). Asumamos ahora que vale 2. Entonces $\ker A^{1/2} \subseteq \ker D^{1/2}$. Luego, podemos definir (con buena definición) la función

$$C_0 : R(A^{1/2}) \to R(D^{1/2}) \quad \text{dada por} \quad C_0(A^{1/2}x) = D^{1/2}x \ , \quad \text{para cualquier } x \in \mathbb{C}^n \ .$$

Es fácil ver que C_0 es lineal. Extendámosla a una $C \in \mathcal{M}_n(\mathbb{C})$ poniendo $C|_{\ker A^{1/2}} \equiv 0$. Ahora podemos verificar sin dificultades que $D^{1/2} = CA^{1/2}$ y, por la condición 2, el hecho de que $\|C\|_{sp} \leq 1$ (acá se usa que $\ker A^{1/2} = R(A^{1/2})^\perp$). ■

Notación: Recordar que, si $\mathcal{M} \subseteq \mathbb{C}^n$ es un subespacio, denotamos por $P_\mathcal{M} \in \mathcal{M}_n(\mathbb{C})^+$ al proyector ortogonal sobre $\mathcal{M}$. Observar que $0 \leq P_\mathcal{M} \leq I$, que $P_\mathcal{M}(\mathcal{M}^\perp) = \{0\}$ y que $P_\mathcal{M}\, x = x$ para todo $x \in \mathcal{M}$.

Teorema 3.8.2. *Sea $A \in \mathcal{M}_n(\mathbb{C})^+$ y $\mathcal{S} \subseteq \mathbb{C}^n$ un subespacio. Sea*

$$\mathcal{M}(A, \mathcal{S}) := \{D \in \mathcal{M}_n(\mathbb{C})^+ : \ D \leq A \quad y \quad R(D) \subseteq \mathcal{S}\} \ . \tag{3.17}$$

Consiredemos el subespacio $\mathcal{M} = A^{-1/2}(\mathcal{S})$ y la matriz $T = A^{1/2}P_\mathcal{M}A^{1/2}$. Entonces,

1. $T \in \mathcal{M}(A, \mathcal{S})$.
2. *Para cualquier $D \in \mathcal{M}(A, \mathcal{S})$, se cumple que $D \leq T$.*

En otras palabras, $T = A^{1/2}P_\mathcal{M}A^{1/2}$ es el máximo de $\mathcal{M}(A, \mathcal{S})$ en el orden usual de $\mathcal{H}(n)$.

Demostración. Observar que $T = A^{1/2}P_\mathcal{M}A^{1/2} \leq A^{1/2}\, I\, A^{1/2} = A$. Además, se tiene que $R(T) \subseteq A^{1/2}(\mathcal{M}) \subseteq \mathcal{S}$. Luego $T \in \mathcal{M}(A, \mathcal{S})$. Si $D \in \mathcal{M}(A, \mathcal{S})$, en particular $D \leq A$. Por el Lema 3.8.1, debe existir una contracción C tal que $D^{1/2} = CA^{1/2}$, o sea que $D^{1/2} = A^{1/2}C^*$.

Como $A^{1/2}(R(C^*)\,) = R(D^{1/2}) \subseteq \mathcal{S}$, deducimos que $R(C^*) \subseteq \mathcal{M}$, o sea $P_\mathcal{M}C^* = C^*$. Usando que $C^*C \leq I$ (porque $\|C\|_{sp} \leq 1$), podemos deducir que $C^*C = P_\mathcal{M}C^*CP_\mathcal{M} \leq P_\mathcal{M}$. Luego

$$D = D^{1/2}D^{1/2} = A^{1/2}C^*CA^{1/2} \leq A^{1/2}P_\mathcal{M}A^{1/2} = T \ ,$$

lo cual muestra que $T = \max \mathcal{M}(A, \mathcal{S})$. ■

Definición 3.8.3. Sean $A \in \mathcal{M}_n(\mathbb{C})^+$ y $\mathcal{S} \subseteq \mathbb{C}^n$, un subespacio. Llamaremos **shorted** de A al subespacio $\mathcal{S}$, y lo notaremos $\Sigma(A, \mathcal{S})$, al máximo del conjunto $\mathcal{M}(A, \mathcal{S})$.

En la siguiente proposición, recopilamos una serie de resultados más o menos inmediatos de la definición y la demostración del Teorema 3.8.2.

Proposición 3.8.4. *Sean $A \in \mathcal{M}_n(\mathbb{C})^+$ y $\mathcal{S} \subseteq \mathbb{C}^n$, un subespacio. Entonces:*

1. $\Sigma(A, \mathcal{S}) \leq A$.
2. *Para todo $\alpha \in \mathbb{R}_+$, se tiene que* $\Sigma(\alpha A, \mathcal{S}) = \alpha \Sigma(A, \mathcal{S})$.
3. *Si $B \in \mathcal{M}_n(\mathbb{C})^+$ cumple que $A \leq B$, entonces*
$$\mathcal{M}(A, \mathcal{S}) \subseteq \mathcal{M}(B, \mathcal{S}) \quad \text{y por lo tanto} \quad \Sigma(A, \mathcal{S}) \leq \Sigma(B, \mathcal{S}) \ .$$
4. *Si $\mathcal{S} \subseteq \mathcal{T} \subseteq \mathbb{C}^n$, entonces $\mathcal{M}(A, \mathcal{S}) \subseteq \mathcal{M}(A, \mathcal{T})$ y $\Sigma(A, \mathcal{S}) \leq \Sigma(A, \mathcal{T})$.*
5. *Si $R(A) \subseteq \mathcal{S}$, entonces $\Sigma(A, \mathcal{S}) = A$.*
6. $\Sigma(\Sigma(A, \mathcal{S}), \mathcal{S}) = \Sigma(A, \mathcal{S})$.

Demostración. Ejercicio. ■

Proposición 3.8.5. *Sean $A \in \mathcal{M}_n(\mathbb{C})^+$ y $\mathcal{S}, \mathcal{T} \subseteq \mathbb{C}^n$, dos subespacios . Entonces*

$$\Sigma(\Sigma(A, \mathcal{S}), \mathcal{T}) = \Sigma(A, \mathcal{S} \cap \mathcal{T}) \ .$$

Demostración. Consideremos los conjuntos

$$\mathcal{M}(A, \mathcal{S} \cap \mathcal{T}) = \{D : \ 0 \leq D \leq A \ \ R(D) \subseteq \mathcal{S} \cap \mathcal{T}\}$$
$$\mathcal{M}(\Sigma(A, \mathcal{T}), \mathcal{S}) = \{D : \ 0 \leq D \leq \Sigma(A, \mathcal{T}) \ \ R(D) \subseteq \mathcal{S}\}.$$

Probaremos que estos conjuntos son iguales y por ende sus máximos también lo son. En efecto, sea $D \in \mathcal{M}(A, \mathcal{S} \cap \mathcal{T})$, entonces se tiene que

$$R(D) \subseteq \mathcal{T} \quad \text{y} \quad D \leq A \implies D \leq \Sigma(A, \mathcal{T}) \ , \quad \text{y también que} \quad R(D) \subseteq \mathcal{S} \ .$$

En consecuencia, $D \in \mathcal{M}(\Sigma(A, \mathcal{T}), \mathcal{S})$. Reciprocamente, si $D \in \mathcal{M}(\Sigma(A, \mathcal{T}), \mathcal{S})$ entonces $D \leq \Sigma(A, \mathcal{T})$, lo cual muestra que $R(D) \subseteq \mathcal{T}$ y en consecuencia $R(D) \subseteq \mathcal{S} \cap \mathcal{T}$. Pero como $D \leq \Sigma(A, \mathcal{T}) \leq A$ se tiene que $D \in \mathcal{M}(A, \mathcal{S} \cap \mathcal{T})$. ■

Cálculo matricial del shorted: complementos de Schur

Se busca dar una expresión "en coordenadas" del shorted $\Sigma(A, \mathcal{S})$. Para ello necesitamos seguir haciendo cuentas del estilo 2×2, que son interesantes en sí mismas.

Proposición 3.8.6. *Sea $\mathcal{S} \subseteq \mathbb{C}^n$ un subespacio, y sea* $M = \begin{array}{c} \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \end{array} \begin{array}{c} \mathcal{S} \\ \mathcal{S}^\perp \end{array} \in \mathcal{M}_n(\mathbb{C})$. *Entonces $M \in \mathcal{M}_n(\mathbb{C})^+$ si y sólo si se verifican las siguientes condiciones:*

1. *$A \in L(\mathcal{S})^+$ y $D \in L(\mathcal{S}^\perp)^+$.*
2. *Existe una* **contracción** *$C \in L(\mathcal{S}^\perp, \mathcal{S})$ tal que $B = A^{1/2} C D^{1/2}$.*

En tal caso se tiene que $R(B) \subseteq R(A)$ y que $R(B^*) \subseteq R(D)$.

Demostración. Si se cumplen las condiciones pedidas, se observa que

$$M = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} = \begin{bmatrix} A^{1/2} & 0 \\ 0 & D^{1/2} \end{bmatrix} \begin{bmatrix} I_{\mathcal{S}} & C \\ C^* & I_{\mathcal{S}^\perp} \end{bmatrix} \begin{bmatrix} A^{1/2} & 0 \\ 0 & D^{1/2} \end{bmatrix} \in \mathcal{M}_n(\mathbb{C})^+ ,$$

por la Proposición 3.7.6 y la Observación 3.7.7. Si suponemos que $M \geq 0$, es claro que $A \in L(\mathcal{S})^+$ y $D \in L(\mathcal{S}^\perp)^+$. Asumamos que ambas son inversibles. Haciendo la cuenta anterior al reves, si llamamos $C = A^{-1/2} B D^{-1/2}$ se tiene que

$$\begin{bmatrix} I_{\mathcal{S}} & C \\ C^* & I_{\mathcal{S}^\perp} \end{bmatrix} = \begin{bmatrix} A^{-1/2} & 0 \\ 0 & D^{-1/2} \end{bmatrix} \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \begin{bmatrix} A^{-1/2} & 0 \\ 0 & D^{-1/2} \end{bmatrix} \in \mathcal{M}_n(\mathbb{C})^+ .$$

Luego queda que $\|C\|_{sp} \leq 1$ y B $= A^{1/2} C D^{1/2}$. El caso general sale aplicando lo anterior a las matrices $M + \frac{1}{n} I$. Se toma la sucesión $C_n = (A + \frac{1}{n} I_{\mathcal{S}})^{-1/2} B \, (D + \frac{1}{n} I_{\mathcal{S}^\perp})^{-1/2}$, que consta de contracciones. Como la bola de $\mathcal{M}_n(\mathbb{C})$ es compacta, hay una subsucesión $C_{n_k} \xrightarrow[k \to \infty]{} C$, donde C es también una contracción. Ahora observamos que, para todo $k \in \mathbb{N}$,

$$B = (A + \frac{1}{n_k} I_{\mathcal{S}})^{1/2} \, C_{n_k} \, (D + \frac{1}{n_k} I_{\mathcal{S}})^{1/2} \xrightarrow[k \to \infty]{} A^{1/2} C D^{1/2} ,$$

donde la continuidad de tomar raices cuadradas se deduce de que todas las matrices de cada sucesión se diagonalizan en la misma base. ■

Proposición 3.8.7. *Sean $A \in \mathcal{G}l(k)^+$, $C \in \mathcal{M}_m(\mathbb{C})^+$ y $B \in \mathcal{M}_{k,m}(\mathbb{C})$. Sea*

$$M = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \begin{array}{c} \mathbb{C}^n \\ \mathbb{C}^m \end{array} \in \mathcal{M}_{k+m}(\mathbb{C}) .$$

Entonces se verifican las siguientes propiedades:

1. *$M \in \mathcal{G}l(m + k)^+ \iff B^* A^{-1} B < C$ (o sea que $C - B^* A^{-1} B \in \mathcal{G}l(m)^+$).*
2. *$M \in \mathcal{M}_{k+m}(\mathbb{C})^+ \iff B^* A^{-1} B \leq C$.*

3. Si $M \in \mathcal{G}l(m+k)$, entonces $M^{-1} = \begin{bmatrix} * & * \\ * & (C - B^*A^{-1}B)^{-1} \end{bmatrix}$.

Demostración. Sea $X = -A^{-1}B \in \mathcal{M}_{km}(\mathbb{C})$. Entonces, haciendo cuentas elementales, obtenemos que

$$\begin{bmatrix} I_k & 0 \\ X^* & I_m \end{bmatrix} M \begin{bmatrix} I_k & X \\ 0 & I_m \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & C - B^*A^{-1}B \end{bmatrix} ,$$

lo que prueba 1 y 2. Por otra parte, como $\begin{bmatrix} I_k & X \\ 0 & I_m \end{bmatrix}^{-1} = \begin{bmatrix} I_k & -X \\ 0 & I_m \end{bmatrix}$, deducimos que

$$\begin{bmatrix} I_k & -X \\ 0 & I_m \end{bmatrix} M^{-1} \begin{bmatrix} I_k & 0 \\ -X^* & I_m \end{bmatrix} = \begin{bmatrix} A^{-1} & 0 \\ 0 & (C - B^*A^{-1}B)^{-1} \end{bmatrix}$$

y, por lo tanto, que

$$M^{-1} = \begin{bmatrix} I_k & X \\ 0 & I_m \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ 0 & (C - B^*A^{-1}B)^{-1} \end{bmatrix} \begin{bmatrix} I_k & 0 \\ X^* & I_m \end{bmatrix} = \begin{bmatrix} * & * \\ * & (C - B^*A^{-1}B)^{-1} \end{bmatrix} ,$$

lo que prueba la parte 3. ■

Ejercicio 3.8.8. Dar otra prueba de la Proposición 3.8.7, vía la Proposición 3.8.6. ▲

Corolario 3.8.9. *Sean $M \in \mathcal{M}_n(\mathbb{C})^+$ y $\mathcal{S} \subseteq \mathbb{C}^n$, un subespacio. Supongamos que*

$$M = \begin{array}{c} \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \end{array} \begin{array}{l} \mathcal{S}^\perp \\ \mathcal{S} \end{array} \qquad \text{y que la compresión} \qquad A = M_{\mathcal{S}^\perp} \in \mathcal{G}l(\mathcal{S}^\perp)^+ ,$$

o sea que $\langle Mx\,, x\rangle > 0$ para todo $x \in \mathcal{S}^\perp \setminus \{0\}$. Luego se tiene que

1. $\Sigma\,(M, \mathcal{S}) = \begin{bmatrix} 0 & 0 \\ 0 & C - B^*A^{-1}B \end{bmatrix} \begin{array}{l} \mathcal{S}^\perp \\ \mathcal{S} \end{array}$.

2. $M \in \mathcal{G}l\,(n)^+ \iff$ *existe un $\lambda \in \mathbb{R}_+^*$ tal que $\lambda\, P_{\mathcal{S}} \le \Sigma\,(M, \mathcal{S})$.*

Demostración. Pongamos que $\dim \mathcal{S} = m$ y llamemos $k = n - m = \dim \mathcal{S}^\perp$. Trabajando en una BON que empiece generando a $\mathcal{S}^\perp$, podemos asumir que $\mathcal{S} = \operatorname{Gen}\{e_{k+1}\,, \dots, e_n\}$ y que estamos en las condiciones de la Proposición 3.8.7 (en particular, que $A \in \mathcal{G}l(k)^+$). Si ahora llamamos $T = \begin{bmatrix} 0 & 0 \\ 0 & C - B^*A^{-1}B \end{bmatrix} \begin{array}{l} \mathcal{S}^\perp \\ \mathcal{S} \end{array}$, es claro que $R(T) \subseteq \mathcal{S}$ y que

$$\begin{aligned} M - T &= \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & C - B^*A^{-1}B \end{bmatrix} \\ &= \begin{bmatrix} A & B \\ B^* & B^*A^{-1}B \end{bmatrix} = \begin{bmatrix} A^{1/2} & 0 \\ B^*A^{-1/2} & 0 \end{bmatrix} \begin{bmatrix} A^{1/2} & A^{-1/2}B \\ 0 & 0 \end{bmatrix} \ge 0 , \end{aligned}$$

por lo que $T \leq M$ y $T \in \mathcal{M}(M, \mathcal{S})$. Tomemos un $Q = \begin{bmatrix} 0 & 0 \\ 0 & D \end{bmatrix} \begin{matrix} \mathcal{S}^\perp \\ \mathcal{S} \end{matrix} \in \mathcal{M}(M, \mathcal{S})$. Luego se tiene que $M - Q = \begin{bmatrix} A & B \\ B^* & C - D \end{bmatrix} \in \mathcal{M}_n(\mathbb{C})^+$. Por la Proposición 3.8.7 vemos que

$$B^* A^{-1} B \leq C - D \implies D \leq C - B^* A^{-1} B \implies Q \leq T \ .$$

Así que $T = \Sigma\,(M, \mathcal{S})$. La segunda parte surge de que

$$C - B^* A^{-1} B \in \mathcal{G}l(\mathcal{S})^+ \iff \Sigma\,(M, \mathcal{S}) = \begin{bmatrix} 0 & 0 \\ 0 & C - B^* A^{-1} B \end{bmatrix} \geq \begin{bmatrix} 0 & 0 \\ 0 & \lambda\, I_{\mathcal{S}} \end{bmatrix} = \lambda\, P_{\mathcal{S}}$$

para cierto $\lambda \in \mathbb{R}_+^*$, volviendo a aplicar la Proposición 3.8.7. ■

Observación 3.8.10. La defnción más tradicional del shorted de un $M \in \mathcal{M}_n(\mathbb{C})^+$ se suele hacer usando el complemento de Schur, según la fórmula que se desarrolla en el Capítulo 12 y se muestra en el Corolario 3.8.9. La ventaja de la definición que surge del Teorema 3.8.2 es que no necesita que haya una submatriz inversible. Sin embargo, utilizando seudoinversas de Moore-Penrose, se puede dar una fórmula estilo la del Corolario 3.8.9 para cualquier $M \in \mathcal{M}_n(\mathbb{C})^+$, reemplazando $B^* A^{-1} B$ por $B^* A^\dagger B$, que en dimensión finita siempre existe (ver el Ejercicios 3.9.20 y 3.9.30). Más allá de esto, la simpleza de las pruebas de las Proposiciones 3.8.4 y 3.8.5 da una muestra de que el enfoque basado en maximizar el conjunto $\mathcal{M}(M, \mathcal{S})$ tiene fuertes ventajas metodológicas.

Todos los resultados de esta sección siguen siendo válidos en el contexto de operadores acotados en espacios de Hilbert. Las pruebas son muy similares, pero necesitan técnicas específicas, sobre todo para mostrar el Lema 3.8.1 y la Proposición 3.8.6 (ojo con la bola compacta). Estos temas se expondrán, con mucha mayor profundidad, en el tomo II. ▲

3.9 Ejercicios

Ejercicios del texto

3.9.1. Mostrar que las normas definidas en 3.4.1 son, efectivamente, normas.

3.9.2. Consideremenos en $\mathbb{C}^n$ las siguientes normas:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \quad \text{y} \quad \|x\|_\infty = \max_{i \in \mathbb{I}_n} |x_i| \quad , \qquad \text{para todo } x \in \mathbb{C}^n \ ,$$

que inducen en $\mathcal{M}_n(\mathbb{C})$ las sendas normas matriciales: Dada $A \in \mathcal{M}_n(\mathbb{C})$,

$$|||A|||_\infty = \max_{\|x\|_\infty = 1} \|Ax\|_\infty \quad \text{y} \quad |||A|||_1 = \max_{\|x\|_1 = 1} \|Ax\|_1 \ .$$

Probar que estas normas pueden calcularse efectivamente mediante las fórmulas:

$$|||A|||_\infty = \max_{i \in \mathbb{I}_n} \|F_i(A)\|_1 \quad \text{y} \quad |||A|||_1 = \max_{i \in \mathbb{I}_n} \|C_i(A)\|_1 \ .$$

para toda $A \in \mathcal{M}_n(\mathbb{C})$.

3.9.3. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Si N es una norma matricial en $\mathcal{M}_n(\mathbb{C})$, mostrar que

$$\rho(A) = \inf_{m\in\mathbb{N}} N(A^m)^{1/m} \quad \text{y que} \quad N(A^m)^{1/m} \underset{m\to\infty}{\searrow} \rho(A) \ .$$

3.9.4. Encontrar una matriz $A \in \mathcal{M}_n(\mathbb{C})^+$ tal que $B = \big(|A_{ij}|\big)_{i,j\in\mathbb{I}_n} \notin \mathcal{M}_n(\mathbb{C})^+$. Observar que hay que buscar para $n \geq 3$.

3.9.5. Probar que, si $A \in \mathcal{M}_n(\mathbb{C})^+$, entonces $e_\circ^A := \left(e^{A_{ij}}\right)_{i,j\in\mathbb{I}_n} \in \mathcal{M}_n(\mathbb{C})^+$.

3.9.6. Si $A \in \mathcal{M}_{n,m}(\mathbb{C})$, entoces $B = \begin{bmatrix} |A^*| & A \\ A^* & |A| \end{bmatrix} \in \mathcal{M}_{n+m}(\mathbb{C})^+$. Se suguiere aplicar 3.7.2 agregándole ceros a A para que quede cuadrada.

3.9.7. Sea $A \in \mathcal{M}_{n,m}(\mathbb{C})$ una contracción, es decir que $\|A\|_{sp} \leq 1$. Probar que

1. $A^*(I_n - AA^*)^{1/2} = (I_m - A^*A)^{1/2}A^* \in \mathcal{M}_{m,n}(\mathbb{C})$.
2. Las matrices

$$\begin{bmatrix} A & (I-AA^*)^{1/2} \\ (I-A^*A)^{1/2} & -A^* \end{bmatrix} \ \text{y} \ \begin{bmatrix} A & -(I-AA^*)^{1/2} \\ (I-A^*A)^{1/2} & A^* \end{bmatrix}$$

son unitarias en $\mathcal{M}_{n+m}(\mathbb{C})$.

3.9.8. Demostrar los 6 items de la Proposición 3.8.4.

Ejercicios nuevos

3.9.9. Sea $A \in \mathcal{M}_{n,m}(\mathbb{C})$. Probar que, para todo $k \in \mathbb{I}_n$, se tiene que

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

3.9.10. Consideremos los conjuntos $R_k(n) = \{T \in \mathcal{M}_{nm}(\mathbb{C}) : \operatorname{rk} T \leq k\}$. Mostrar que

$$A \in \mathcal{M}_{nm}(\mathbb{C}) \implies s_k(A) = \min_{T\in R_{k-1}} \|A - T\| \quad \text{para todo } k \in \mathbb{I}_n \ .$$

3.9.11. Mostrar que si $A, H \in \mathcal{M}_{nm}(\mathbb{C})$, y $\operatorname{rk} H = k$, entonces

$$s_j(A) \geq s_{j+k}(A+H) \ , \quad \text{para todo} \quad j \in \mathbb{I}_{n-k} \ .$$

3.9.12. Mostrar que, para cualquier $A \in \mathcal{M}_{nm}(\mathbb{C})$ y para cada $k \in \mathbb{I}_n$, vale que

$$\sum_{j=1}^k s_j(A) = \max \left| \sum_{j=1}^k \langle Ax_j \ , y_j\rangle \right| ,$$

donde el máximo se toma sobre todas las $k-$uplas ortonormales $x_1 , \dots , x_k$ e $y_1 , \dots , y_k$.

3.9.13. Sean $A, B, C \in \mathcal{M}_n(\mathbb{C})$ tal que $\|A\|_{sp} \le 1$. Se definen

$$D_A = (I - A^*A)^{1/2} \qquad \text{y} \qquad D_{A^*} = (I - AA^*)^{1/2} \ .$$

1. Suponiendo que $\|A\|_{sp} < 1$, verificar:
   (a) Si $K = D_A^{-1}C^*$ y $L = D_{A^*}^{-1}B$, entonces
   $$KK^* \le 1 \text{ (resp. } LL^* \le 1) \iff A^*A + C^*C \le 1 \text{ (resp. } AA^* + BB^* \le 1).$$
   (b) Demostrar que $\begin{bmatrix} I & A \\ A^* & I \end{bmatrix}^{-1} = \begin{bmatrix} D_{A^*}^{-2} & -D_{A^*}^{-1}AD_A^{-1} \\ -D_A^{-1}A^*D_{A^*}^{-1} & D_A^{-2} \end{bmatrix}$.
   (c) Sea $X \in \mathcal{M}_n(\mathbb{C})$. Demostar que las matrices
   $$\begin{bmatrix} I & 0 & A & B \\ 0 & I & C & X \\ A^* & B^* & I & 0 \\ C^* & X^* & 0 & I \end{bmatrix} \qquad \text{y} \qquad \begin{bmatrix} I & A & 0 & B \\ A^* & I & C^* & 0 \\ 0 & C & I & X \\ B^* & 0 & X^* & I \end{bmatrix}$$
   son conjugadas y ambas positivas.
2. **(Parrot)** Demostrar que las siguientes afirmaciones son equivalentes:
   (a) Existe $X \in \mathcal{M}_n(\mathbb{C})$ tal que $\left\| \begin{bmatrix} A & B \\ C & X \end{bmatrix} \right\|_{sp} \le 1$.
   (b) $\left\| \begin{bmatrix} A & B \end{bmatrix} \right\|_{sp} \le 1$ y $\left\| \begin{bmatrix} A \\ C \end{bmatrix} \right\|_{sp} \le 1$.

3.9.14 (Fejer)**.** Sea $A \in \mathcal{M}_n(\mathbb{C})$. Probar que

$$A \in \mathcal{M}_n(\mathbb{C})^+ \iff \sum_{i,j \in \mathbb{I}_n} A_{ij}B_{ij} \ge 0 \quad \text{para toda} \quad B \in \mathcal{M}_n(\mathbb{C})^+ \ .$$

3.9.15. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Probar que

$$\operatorname{Re} A \in \mathcal{M}_n(\mathbb{C})^+ \implies \operatorname{Re}(A \circ B) \in \mathcal{M}_n(\mathbb{C})^+ \quad \text{para toda} \quad B \in \mathcal{M}_n(\mathbb{C})^+ \ .$$

3.9.16. Sea $\mathbb{A} = \{A_k\}_{k \in \mathbb{N}}$ una sucesión en $\mathcal{H}(n)$ tal que, para todo $k \in \mathbb{N}$,

1. $A_{k+1} \le A_k$ (es decir que $\mathbb{A}$ es decreciente)
2. Existe $B \in \mathcal{H}(n)$ tal que $B \le A_k$ (o sea que $\mathbb{A}$ es acotada inferiormente). Observar que esto equivale pedir que la sucesión $\{\|A_k\|_{sp}\}_{k \in \mathbb{N}}$ sea acotada.

Entonces existe $A = \inf\limits_{k \in \mathbb{N}} A_k = \lim\limits_{k \in \mathbb{N}} A_k \in \mathcal{H}(n)$. Es decir que $A_k \xrightarrow[k \to \infty]{} A$, que $A \le A_k$ para todo $k \in \mathbb{N}$ y que, si un $C \in \mathcal{H}(n)$ es menor que todos los A_k, entonces también $C \le A$. Probar un resultado similar si $\mathbb{A}$ es creciente y acotada superiormente.

Se suguiere definir $\langle Ax \,, x \rangle = \lim\limits_{k \in \mathbb{N}} \langle A_k x \,, x \rangle$ para cada $x \in \mathbb{C}^n$, y extrapolar a $\langle Ax \,, y \rangle$ usando polarización. Otro argumento (bien finitodimensional) sería diagonalizar a todas, aplicar el Teorema de Weyl 2.3.5 y usar que $\mathcal{U}(n)$ es compacto, para construir A y ver que el límite de arriba camina, y que es un ínfimo por hipótesis.

Pseudoinversas y proyecciones oblicuas

Definición 3.9.17. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Se dice que una matriz $B \in \mathcal{M}_n(\mathbb{C})$ es una pseudoinversa de A si verifica que $ABA = A$ y $BAB = B$. ▲

Observación 3.9.18. En algunos libros, las matrices que cumplen las condiciones de la definición anterior se denominan g-inversas reflexivas. ▲

3.9.19. Sea $\mathcal{S}$ un subespacio de $\mathbb{C}^n$. Demostrar que todas las proyecciones oblicuas sobre $\mathcal{S}$ poseen la siguiente representación matricial: Dada $Q \in \mathcal{M}_n(\mathbb{C})$, se tiene que

$$Q^2 = Q \quad \text{y} \quad R(Q) = \mathcal{S} \quad \iff \quad \text{existe } X \in L(\mathcal{S}^\perp, \mathcal{S}) \text{ tal que} \quad Q = \begin{bmatrix} I_{\mathcal{S}} & X \\ 0 & 0 \end{bmatrix} \begin{array}{l} \mathcal{S} \\ \mathcal{S}^\perp \end{array} .$$

Ya que estamos, probar que $\|Q\|_{sp}^2 = 1 + \|X\|_{sp}^2$.

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

1. Probar que para cualquier pseudoinversa B de A se cumple que

$$(AB)^2 = AB \quad , \quad (BA)^2 = BA \quad , \quad R(AB) = R(A) \quad \text{y} \quad \ker BA = \ker A \ .$$

2. Dadas dos proyecciones oblicuas (o no) P, Q tales que $R(P) = R(A)$ y $\ker Q = \ker A$, probar que existe una **unica** pseudoinversa B (de A) tal que $AB = P$ y $BA = Q$.

Definición 3.9.21. Dada $A \in \mathcal{M}_n(\mathbb{C})$, se definen:

1. La pseudoinversa de Moore-Penrose $A^\dagger$ de A, como la única que cumple que las proyecciones $AA^\dagger$ y $A^\dagger A \in \mathcal{H}(n)$ i.e., son ortogonales.

2. El módulo mínimo reducido de A como:

$$\gamma(T) := \min \left\{ \|Ax\| : \quad x \in \ker A^\perp \quad \|x\| = 1 \right\} \ .$$ ▲

3.9.22. Sea $A \in \mathcal{G}l(n)$. Probar que, en ese caso, $A^\dagger = A^{-1}$ y $\gamma(A) = \|A^{-1}\|_{sp}^{-1}$.

3.9.23. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Demostrar lo siguiente:

1. $(A^*)^\dagger = (A^\dagger)^*$ y $(A^\dagger)^\dagger = A$.

2. $A^\dagger = A^*(AA^*)^\dagger$.

3. Sea $B \in \mathcal{M}_n(\mathbb{C})$ tal que

$$R(B) = \ker A^\perp \ , \quad \ker AB = \ker B \quad \text{y} \quad R(AB) = R(A) \ .$$

Entonces $(AB)^\dagger = B^\dagger A^\dagger$.

3.9.24. Sea $\{A_n\} \subseteq \mathcal{M}_n(\mathbb{C})$ una sucesión de matrices que poseen el mismo rango. Supongamos que $\{A_n\} \xrightarrow[n\to\infty]{} L \in \mathcal{M}_n(\mathbb{C})$ y que L posee el mismo rango que las A_n .

1. Probar que $A_n^\dagger \xrightarrow[n\to\infty]{} L^\dagger$.
2. Dar un ejemplo de una sucesión $\{A_n\} \xrightarrow[n\to\infty]{} L$ pero tal que $\{A_n^\dagger\} \not\xrightarrow[n\to\infty]{} L^\dagger$.

3.9.25. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Probar:

1. $A^\dagger = \lim\limits_{\varepsilon\to 0} A^*(\varepsilon I + AA^*)^{-1}$.
2. $A^\dagger = \alpha \sum\limits_{n=0}^{\infty} (I - \alpha A^*A)^n A^* = \alpha \sum\limits_{n=0}^{\infty} A^*(I - \alpha AA^*)^n$ para cualquier $0 < \alpha < 2/\|A\|_{sp}^2$.

3.9.26. Sea $A \in \mathcal{M}_n(\mathbb{C})^+$. Probar que $A = \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{matrix} R(A) \\ \ker A \end{matrix}$. Deducir que

$$A^\dagger = \begin{bmatrix} A_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{matrix} R(A) \\ \ker A \end{matrix} \qquad \text{y que} \qquad \gamma(A) = \min\{\lambda \in \sigma(A) : \lambda \neq 0\} \ .$$

3.9.27. Sea $A \in \mathcal{M}_n(\mathbb{C})$ tal que su descomposición en valores singulares es $A = W\Sigma(A)V^*$. Expresar $A^\dagger$ en función de W, $\Sigma(A)$, y V. ¿Qué relación hay entre los valores singulares de A y los de $A^\dagger$?

3.9.28. Dada $A \in \mathcal{M}_n(\mathbb{C})$ demostrar las siguientes propiedades:

1. $\gamma(A)^2 = \gamma(A^*A) = \gamma(AA^*) = \gamma(A^*)^2$.
2. $\gamma(A) = \|A^\dagger\|^{-1}$.

3.9.29. Sean $A, B \in \mathcal{M}_n(\mathbb{C})^+$. Probar que $A \leq B \iff \rho(AB^\dagger) \leq 1$.

3.9.30. Sean $M \in \mathcal{M}_n(\mathbb{C})^+$ y un subespacio $\mathcal{S} \subseteq \mathbb{C}^n$ tales que $M = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \begin{matrix} \mathcal{S}^\perp \\ \mathcal{S} \end{matrix}$. Probar que $\Sigma(M, \mathcal{S}) = \begin{bmatrix} 0 & 0 \\ 0 & C - B^*A^\dagger B \end{bmatrix} \begin{matrix} \mathcal{S}^\perp \\ \mathcal{S} \end{matrix}$.

3.9.31. Sean $A \in \mathcal{M}_n(\mathbb{C})^+$ y $\mathcal{S} \subseteq \mathbb{C}^n$ un subespacio. Dado $x \in \mathcal{S}$, se tiene que

$$\left\langle \Sigma(A, \mathcal{S}) \begin{bmatrix} 0 \\ x \end{bmatrix}, \begin{bmatrix} 0 \\ x \end{bmatrix} \right\rangle = \inf \left\{ \left\langle A \begin{bmatrix} y \\ x \end{bmatrix}, \begin{bmatrix} y \\ x \end{bmatrix} \right\rangle \ : \ y \in \mathcal{S}^\perp \right\} \ .$$

3.9.32. Sean $A \in \mathcal{M}_n(\mathbb{C})^+$ y $\mathcal{S} \subseteq \mathbb{C}^n$ un subespacio. Entonces existen únicos

$$F \text{ y } G \in \mathcal{M}_n(\mathbb{C})^+ \quad \text{tales que} \quad A = F + G \ , \quad R(F^{1/2}) \subseteq \mathcal{S} \quad \text{y} \quad R(G^{1/2}) \cap \mathcal{S} = \{0\} \ .$$

Más aun, $F = \Sigma(A, \mathcal{S})$ y $G = A - \Sigma(A, \mathcal{S})$.

3.9.33. $\mathcal{S} \subseteq \mathbb{C}^n$ un subespacio. Consideremos la casimatriz $M = \begin{array}{c} \begin{bmatrix} D & B \\ B^* & ? \end{bmatrix} \end{array} \begin{array}{c} \mathcal{S}^\perp \\ \mathcal{S} \end{array}$. Sea

$$\mathcal{P}(M, \mathcal{S}) = \left\{ X \in L(\mathcal{S})^+ \ : \ \begin{bmatrix} D & B \\ B^* & X \end{bmatrix} \in \mathcal{M}_n(\mathbb{C})^+ \right\} .$$

Probar que $\mathcal{P}(M, \mathcal{S}) \neq \varnothing$ si y sólo si $R(B) \subseteq R(D^{1/2})$. Más aún, probar que en tal caso existe $X_0 = \text{mín}\, \mathcal{P}(M, \mathcal{S})$, e identificarlo.

Definición 3.9.34. Dada una matriz $A \in \mathcal{M}_n(\mathbb{C})^+$, consideraremos en $\mathbb{C}^n$ el pseudoproducto interno $\langle \cdot, \ \cdot \rangle_A$ definido por

$$\langle x, \ y \rangle_A = \langle Ax, \ y \rangle \quad , \qquad \text{para todo par} \quad x \ , \ y \ \in \ \mathbb{C}^n \ . \qquad \blacktriangle$$

3.9.35. Sea $A \in \mathcal{M}_n(\mathbb{C})^+$ y $\mathcal{S}$ un subespacio de $\mathbb{C}^n$. Demostrar las siguientes propiedades:

1. $\mathcal{S}^{\perp_A} = A^{-1}(\mathcal{S}^\perp) = A(\mathcal{S})^\perp$.
2. $T \in \mathcal{M}_n(\mathbb{C})$ es A-autoadjunto si y sólo si $TA = A^*T$.
3. El conjunto de las proyecciones A-autoadjuntas con rango $\mathcal{S}$, que denotaremos
$$\mathcal{P}(A, \, \mathcal{S}) = \{Q \in \mathcal{M}_n(\mathbb{C}) : \ Q^2 = Q, \ AQ = Q^*A \text{ y } R(Q) = \mathcal{S}\} \neq \varnothing \ .$$

3.9.36. Sea $A \in \mathcal{M}_n(\mathbb{C})^+$ y $\mathcal{S}$ un subespacio de $\mathbb{C}^n$. Si $A = \begin{bmatrix} a & b \\ b^* & c \end{bmatrix} \begin{array}{c} \mathcal{S}^\perp \\ \mathcal{S} \end{array}$, probar que la proyección $P_{A, \mathcal{S}}$ definida por

$$P_{A, \mathcal{S}} = \begin{bmatrix} 1 & a^\dagger b \\ 0 & 0 \end{bmatrix} \begin{array}{c} \mathcal{S}^\perp \\ \mathcal{S} \end{array} \in \mathcal{M}_n(\mathbb{C})$$

satisface que $P_{A, \mathcal{S}} \in \mathcal{P}(A, \, \mathcal{S})$ y encima $\|P_{A, \mathcal{S}}\|_{sp} = \min\limits_{Q \in \mathcal{P}(A, \, \mathcal{S})} \|Q\|_{sp}$.

3.9.37. Dada una un proyección $Q \in \mathcal{M}_n(\mathbb{C})$, no necesariamente ortogonal, construir una matriz $A \in \mathcal{G}l(n)^+$ de modo que Q resulte A-autoadjunta.

3.9.38. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Dada una matriz $W \in \mathcal{G}l(n)^+$, demostrar que $B = (A^*WA)^\dagger A^*D$ es una pseudoinversa de A tal que AB es una proyección ortogonal respecto al producto interno $\langle \cdot, \ \cdot \rangle_W$ y BA lo es respecto al producto interno usual.

3.9.39. Dado $T \in \mathcal{M}_n(\mathbb{C})$ y $A \in \mathcal{G}l(n)^+$, encontrar una expresión para la Moore-Penrose de T respecto al $\langle \cdot, \ \cdot \rangle_A$.

3.9.40. Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Encontrar $C \in \mathcal{M}_n(\mathbb{C})$, tal que

$$\|AC - B\|_{sp} = \min_{X \in \mathcal{M}_n(\mathbb{C})} \|AX - B\|_{sp} \quad .$$

Ahora reemplazar la norma espectral por alguna otra NUI y encontrar C para dicha norma. ¿Qué conclusión puede sacar?, ¿puede extender su conclusión para otras normas unitariamente invariantes?.

3.9.41. Dadas $A, B \in \mathcal{M}_n(\mathbb{C})$, se dice que $A \leq^* B$ si $BA^* = AA^*$ y $B^*A = A^*A$. Demostrar:

1. $A \leq^* B \iff A^\dagger A = B^\dagger A$ y $AA^\dagger = AB^\dagger$.
2. $A^\dagger = \max\limits_{\leq^*} \{B \in \mathcal{M}_n(\mathbb{C}) : \; BAB = B, \; (AB)^* = AB, \text{ y } (BA)^* = BA\}$.
3. $A^\dagger = \min\limits_{\leq^*} \{B \in \mathcal{M}_n(\mathbb{C}) : \; ABA = A, \; (AB)^* = AB, \text{ y } (BA)^* = BA\}$.

3.9.42 (Ljance-Ptak.)**.** Sea $E \in \mathcal{M}_n(\mathbb{C})$ una proyección oblicua. Si $P, Q \in \mathcal{M}_n(\mathbb{C})$ designan las proyecciones ortogonales al rango y nucleo de E respectivamente, probar que

$$\|PQ\|_{sp} < 1 \quad \text{y} \quad \text{que} \quad \|E\|^2 = \frac{1}{1 - \|PQ\|_{sp}^2} \quad .$$

Capítulo 4

Mayorización

Es este capítulo expondremos las nociones básicas de mayorización, y su relación con combinaciones convexas de permutaciones, matrices doblemente estocáscticas y funciones convexas. En la sección 3 mostraremos el teorema de Birkhoff que asegura que las matrices de permutación son el conjunto de puntos extremales de las matrices doblemente estocáscticas. En la última sección introduciremos las propiedades básicas de la mayorización logarítmica.

4.1 Definiciones y caracterizaciones

Notaciones: Sea $x = (x_1, ..., x_n) \in \mathbb{R}^n$.

1. Notaremos $x^\downarrow$ y $x^\uparrow$ a los vectores obtenidos al reordenar las coordenadas 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 se tendrá que $x^\downarrow = \mu(A)$ y $x^\uparrow = \lambda(A)$.

2. Denotaremos por $\mathbb{1} = (1, 1, \ldots, 1) \in \mathbb{R}^n$, al vector con todas sus entradas iguales a uno. Si hace falta aclarar el tamaño, escribiremos $\mathbb{1}_n$.

3. Escribiremos $\operatorname{tr} x = \langle x, \mathbb{1} \rangle = \sum\limits_{j=1}^{n} x_i$.

Con estas notaciones podemos dar la definición de mayorización:

Definición 4.1.1. Sean $x, y \in \mathbb{R}^n$.

1. Se dice que y **mayoriza a** x, y se nota $x \prec y$ si se verifica que

   $$\operatorname{tr} y = \operatorname{tr} x \ , \quad \text{y además} \quad \sum_{j=1}^{k} x_j^\downarrow \ \leq \ \sum_{j=1}^{k} y_j^\downarrow \quad \text{para todo } k \in \mathbb{I}_n \ . \tag{4.1}$$

2. Dado que $\sum\limits_{j=1}^{k} x_j^\uparrow = \operatorname{tr} x - \sum\limits_{j=1}^{n-k} x_j^\downarrow$ para todo $k \in \mathbb{I}_n$, la relación $x \prec y$ equivale a que

$$\operatorname{tr} y = \operatorname{tr} x \ , \quad \text{y además} \quad \sum_{j=1}^{k} x_j^\uparrow \ \geq \ \sum_{j=1}^{k} y_j^\uparrow \quad \text{para todo } k \in \mathbb{I}_n \ . \tag{4.2}$$

3. Si sólo se cumple la segunda condición (4.1) (o sea que se permite que $\operatorname{tr} x < \operatorname{tr} y$), se dice que x está **submayorizado** por y y se nota $x \prec_w y$.

4. Si sólo se cumple la segunda condición (4.2) (aca se permite que $\operatorname{tr} x > \operatorname{tr} y$), se dice que x está **supramayorizado** por y y se nota $x \prec^w y$. ▲

Ejemplos 4.1.2.

1. Sea $x \in \mathbb{R}^n$. Llamemos $a = \operatorname{tr} x$. Entonces $\frac{a}{n}\mathbb{1} = \left(\frac{a}{n} \ , \ \frac{a}{n} \ , \dots , \ \frac{a}{n}\right) \prec x$. En efecto, supongamos que existiera un $k \in \mathbb{I}_n$ tal que $\sum\limits_{i=1}^{k} x_i^\downarrow < \frac{ka}{n}$. En tal caso,

$$x_k^\downarrow \ \leq \ \frac{1}{k}\sum_{i=1}^{k} x_i^\downarrow \ < \ \frac{a}{n} \quad \Longrightarrow \quad \sum_{i=k+1}^{n} x_i^\downarrow < \frac{(n-k)a}{n} \quad \Longrightarrow \quad \sum_{i=1}^{n} x_i^\downarrow < a \ .$$

2. Si $x \in \mathbb{R}_+^n$, entonces $x \prec (\operatorname{tr} x, 0, \dots, 0)$.

3. Sean $x, y \in \mathbb{R}^n$. Si sucediera que $x \prec y$ e $y \prec x$, entonces es fácil ver que $x^\downarrow = y^\downarrow$. Por lo tanto, x e y sólo difieren en una permutación. ▲

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

1. Diremos $A \geqslant 0$ si $A_{ij} \geq 0$ para todo par $i \in \mathbb{I}_n$, $j \in \mathbb{I}_m$. En otras palabras, $A \geqslant 0$ si A tiene entradas no negativas. Ojo con el simbolito, no el lo mismo escribir

$$A \geqslant 0 \text{ (entradas positivas) que } A \geq 0 \text{ (semidefinida positiva)} \ .$$

2. Si $x, y \in \mathbb{R}^n$, pondremos $x \geqslant y$ si se cumple que $x_i \geq y_i$ para todo $i \in \mathbb{I}_n$. También escribiremos que $x > 0$ si $x \in \mathbb{R}_+^{*\,n}$. ▲

Definición 4.1.3. Una matriz $A \in \mathcal{M}_n(\mathbb{R})$ se denomina **doblemente estocástica** si

$$A \geqslant 0 \quad , \quad \operatorname{tr} F_i(A) = 1 \quad \text{ y } \quad \operatorname{tr} C_i(A) = 1 \quad \text{ para todo } i \in \mathbb{I}_n \quad .$$

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

Ejercicio 4.1.4. Sea $A \in \mathcal{M}_n(\mathbb{R})$. Probar que $A \in \mathcal{DS}(n)$ si y sólo si

$$A \geqslant 0 \ , \quad A\,\mathbb{1} = \mathbb{1} \quad \text{y} \quad A^*\,\mathbb{1} = \mathbb{1},$$

donde $\mathbb{1} = (1, 1, \dots, 1) \in \mathbb{R}^n$. Deducir que $\mathcal{DS}(n)$ es un conjunto convexo, y que dadas dos matrices $A, B \in \mathcal{DS}(n)$, entonces también $AB \in \mathcal{DS}(n)$. ▲

Observación 4.1.5 (Matrices de permutación). Sea $n \in \mathbb{N}$.

1. Llamaremos $\mathbf{S}_n$ al n-ésimo grupo simétrico, es decir

$$\mathbf{S}_n = \{\sigma : \mathbb{I}_n \to \mathbb{I}_n \ : \ \sigma \text{ es biyectiva } \} ,$$

   con el producto dado por la composición de funciones.

2. Dados $\sigma \in \mathbf{S}_n$ y $x \in \mathbb{C}^n$, llamaremos $x_\sigma = (x_{\sigma(1)}, \dots, x_{\sigma(n)})$.

3. Dada $\sigma \in \mathbf{S}_n$, llamaremos $P_\sigma : \mathbb{C}^n \to \mathbb{C}^n$ al operador dado por

$$P_\sigma \, x \ = \ x_\sigma \ , \quad x \in \mathbb{C}^n \ .$$

   Es claro que esta función es lineal, por lo que pensamos a $P_\sigma \in \mathcal{M}_n(\mathbb{C})$ como su matriz en la base canónica.

4. Observar que, dado $x \in \mathbb{R}^n$, existen $\sigma, \tau \in \mathbf{S}_n$ tales que $x^\downarrow = P_\sigma \, x$ y $x^\uparrow = P_\tau \, x$.

5. Dada $\sigma \in \mathbf{S}_n$, las columnas de P_σ están dadas por

$$C_k(P_\sigma) = P_\sigma(e_k) = (e_k)_\sigma = e_{\sigma^{-1}(k)} \quad , \quad k \in \mathbb{I}_n \ .$$

6. Dadas $\sigma, \tau \in \mathbf{S}_n$, entonces $P_\sigma P_\tau = P_{\sigma\tau}$, porque $(x_\tau)_\sigma = x_{\sigma\tau}$ para todo $x \in \mathbb{C}^n$.

7. El grupo $\mathcal{U}_\mathcal{P}(n) = \{P_\sigma : \sigma \in \mathbf{S}_n\}$ está incluido en $\mathcal{U}(n)$, dado que cada P_σ es claramente isométrico. Por lo tanto, para cada $\sigma \in \mathbf{S}_n$, $P_{\sigma^{-1}} = P_\sigma^{-1} = P_\sigma^* = P_\sigma^T$.

8. Por otra parte, también se tiene que $\mathcal{U}_\mathcal{P}(n) \subseteq \mathcal{DS}(n)$. En efecto, dada $\sigma \in \mathbf{S}_n$,

$$C_k(P_\sigma) = P_\sigma(e_k) = e_{\sigma^{-1}(k)} \quad \text{y} \quad F_k(P_\sigma) = C_k(P_\sigma^T) = C_k(P_{\sigma^{-1}}) = e_{\sigma(k)} \ , \tag{4.3}$$

   para todo $k \in \mathbb{I}_n$. Otra forma de verlo es notando que $P_\sigma\,\mathbb{1} = P_\sigma^T\,\mathbb{1} = \mathbb{1}$. Más adelante veremos que $\mathcal{U}_\mathcal{P}(n)$ es ni más ni menos que el conjunto de puntos extremales de $\mathcal{DS}(n)$.

9. Dadas $A \in \mathcal{M}_n(\mathbb{C})$ y $P_\sigma \in \mathcal{U}_\mathcal{P}(n)$, para cada $k \in \mathbb{I}_n$ se tiene que

$$F_k(P_\sigma A) = F_{\sigma(k)}(A) \quad , \quad C_k(AP_\sigma) = C_{\sigma^{-1}(k)}(A) \quad \text{y} \quad \mathrm{d}\,(P_\sigma A P_\sigma^*) = \mathrm{d}\,(A)_\sigma \ . \tag{4.4}$$

   En efecto, para todo $i \in \mathbb{I}_n$, tenemos que $C_i(P_\sigma A) = P_\sigma(C_i(A)) = C_i(A)_\sigma$. Luego $(P_\sigma A)_{ki} = A_{\sigma(k)i}$, para todo $i \in \mathbb{I}_n$. Esto prueba la primera igualdad.

   La seguna sale aplicando la de las filas a $(AP_\sigma)^T = P_{\sigma^{-1}} A^T$. La de las diagonales sale porque cada

$$(P_\sigma A P_\sigma^*)_{kk} = \langle P_\sigma A P_\sigma^* \, e_k \, , e_k \rangle = \langle A P_\sigma^* \, e_k \, , P_\sigma^* \, e_k \rangle = \langle A \, e_{\sigma(k)} \, , \, e_{\sigma(k)} \rangle = A_{\sigma(k)\sigma(k)} \ .$$

   En resumen, multplicar por P_σ a izquierda permuta las filas, hacerlo a derecha permuta las columnas (con σ^{-1}), y conjugar con P_σ permuta la diagonal de las matrices. ▲

Teorema 4.1.6. *Sea $A \in \mathcal{M}_n(\mathbb{R})$. Luego se tiene que*

$$A \in \mathcal{DS}\,(n) \quad \iff \quad Ax \prec x \quad \textit{para todo } x \in \mathbb{R}^n \ .$$

Demostración. Supongamos que $Ax \prec x$ para todo x. Sea $\mathcal{E} = \{e_1\,, \dots, e_n\}$ la base canónica de $\mathbb{C}^n$. Para cada $i \in \mathbb{I}_n$, se tiene que $C_i(A) = Ae_i \prec e_i$. Esto implica que $A \geqslant 0$ y que

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

Por otra parte, sabemos que $A\mathbb{1} \prec \mathbb{1}$. Pero por el Ejemplo 4.1.2, como todas las coordenadas de $\mathbb{1}$ son iguales, deducimos que $\mathbb{1} \prec A\mathbb{1}$. Y como no vale la pena permutar al $\mathbb{1}$, queda que

$$\mathbb{1} = A\,\mathbb{1} = (\operatorname{tr} F_1(A), \dots, \operatorname{tr} F_n(A)\,) \ .$$

Recíprocamente, supongamos que $A \in \mathcal{DS}\,(n)$ y llamemos $y = Ax$. Queremos probar que $y \prec x$. Se puede suponer que las coordenadas de x y de y están ordenadas en forma decreciente, porque si $P, Q \in \mathcal{U}_{\mathcal{P}}(n)$ son matrices de permutación (ver Observación 4.1.5), entonces $QAP \in \mathcal{DS}\,(n)$ (por el Ejercicio 4.1.4). Por otra parte, como $y = Ax$,

$$\sum_{j=1}^{k} y_j = \sum_{j=1}^{k}\sum_{i=1}^{n} a_{ji}\, x_i = \sum_{i=1}^{n}\Big(\sum_{j=1}^{k} a_{ji}\Big) x_i \qquad \text{para todo} \qquad k \in \mathbb{I}_n \ . \tag{4.5}$$

Fijemos un $k \in \mathbb{I}_n$. Si para cada $i \in \mathbb{I}_n$ llamamos $t_i = \sum\limits_{j=1}^{k} a_{ji}$, 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 Eq. (4.5),

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

pues los dos sumandos 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 $k \in \mathbb{I}_n$. Por último, observar que la Eq. (4.5) muestra que $\operatorname{tr} y = \operatorname{tr} x$ (para $k = n$, todos los $t_i = \operatorname{tr} C_i(A) = 1$). Así, llegamos a que $Ax = y \prec x$. ■

Ejemplo 4.1.7. Como motivación del siguiente resultado, supongamos que

$$x, y \in \mathbb{R}^2 \quad \text{cumplen que} \quad y = (y_1, y_2) \prec x = (x_1, x_2) \ , \quad y_1 \geq y_2 \ \text{ y } \ x_1 \geq x_2 \ .$$

En este caso, esto significa que

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

Luego, debe existir un $\lambda \in [0, 1]$ 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 $y = (y_1, y_2) = \lambda(x_1, x_2) + (1 - \lambda)(x_2, x_1) = \lambda x + (1 - \lambda)P_\tau\, x$, donde $\tau \in \mathbf{S}_2$ es la permutación no trivial. ▲

Teorema 4.1.8. *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 4.1.6 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 4.1.7. 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 $\tau \in \mathbf{S}_n$ la trasposición que permuta 1 con k. Luergo $P_\tau \in \mathcal{U}_{\mathcal{P}}(n)$ verifica que

$$P_\tau\, x = (x_k\,, x_2\,, \dots, x_{k-1}\,, x_1\,, x_{k+1}\,, \dots, x_n) \ .$$

Definamos $z = \lambda x + (1 - \lambda)P_\tau\, x$. Observar que $z_1 = \lambda\, x_1 + (1 - \lambda)\, x_k = y_1$. Sean

$$y' = (y_2\,, \dots, y_n) \quad \text{ y } \quad z' = (z_2\,, \dots, z_n) \in \mathbb{R}^{n-1} \ .$$

Vamos a probar que $y' \prec z'$: Como $z_1 = y_1$ y tr z = tr x = tr y, se deduce fácilmente 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 $m \geq k$.

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

Luego $y' \prec z'$ y, por hipótesis inductiva, $y' = \sum\limits_{i=1}^{s} \mu_i \, P_{\sigma_i} \, z'$ para ciertos $\mu_i \geq 0$ que suman uno, y para ciertas permutaciones $\sigma_i \in S_{n-1}$ (pensadas como biyecciones del conjunto $\{2, 3, \dots, n\}$). Llamemos también $\sigma_i \in \mathbf{S}_n$ a la extensión de cada permutación σ_i a todo $\mathbb{I}_n$, poniendo $\sigma_i(1) = 1$. Luego, como $z_1 = y_1$, se tiene que $y = \sum\limits_{i=1}^{s} \mu_i \, P_{\sigma_i} \, z$. Pero entonces

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

que es una combinación convexa de permutaciones de x. ■

Observación 4.1.9. Un error típico al tratar de demostrar mayorización entre dos vectores, es olvidarse de ordenarlos antes de sumar sus “primeras” k coordenadas. De hecho, esto sucede en la prueba anterior con los vectores z' e y'. 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 $x \prec y$. En la prueba de la relación $y' \prec z'$, el vector y' ya venía ordenado correctamente. ▲

Corolario 4.1.10. *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 4.1.8, 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 $C(y, z) = (x, w)$. ■

Lema 4.1.11. Sean $x, u \in \mathbb{R}^n$ tales que $x \leqslant u$. Entonces se tiene que

$$x^\downarrow \leqslant u^\downarrow \qquad \text{y} \qquad x \prec_w u \ .$$

Demostración. Es claro que $x_1^\downarrow = \max\limits_{i \in \mathbb{I}_n} x_i \le \max\limits_{i \in \mathbb{I}_n} u_i = u_1^\downarrow$. Si ambos máximos se alcanzan en la misma coordenada de x y de u, un argumento inductivo permite concluir que $x^\downarrow \leqslant u^\downarrow$. Sinó, supongamos que $x_1^\downarrow = x_k$ mientras que $u_1^\downarrow = u_j$. Si llamamos $y \in \mathbb{R}^n$ al resultado de permutar las cordenadas k y j de x, sigue pasando que $y \leqslant u$, porque

$$y_j = x_k \le u_k \le u_1^\downarrow = u_j \qquad \text{mientras que} \qquad y_k = x_j \le x_1^\downarrow = x_k \le u_k \ .$$

Por el caso anterior, $x^\downarrow = y^\downarrow \leqslant u^\downarrow$, y por lo tanto $x \prec_w u$. ■

Proposición 4.1.12. *Sean $x, y \in \mathbb{R}^n$. Entonces*

$$x \prec_w y \quad \iff \quad \text{existe } \ u \in \mathbb{R}^n \ \text{ tal que } \ x \leqslant u \prec y \ .$$

Demostración. Antes que nada, es claro que si el tal u existe, entonces $x \prec_w y$ (por el Lema 4.1.11 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, y $\prec_w$ equivale a $\leqslant$). Si $n > 1$, cosideramos dos casos:

Caso 1: Supongamos que existe $k \in \mathbb{I}_{n-1}$ tal que $\sum\limits_{i=1}^{k} x_i = \sum\limits_{i=1}^{k} y_i$. En tal caso, llamaremos $a = (x_1, \dots, x_k)$ y $b = (y_1, \dots, y_k) \in \mathbb{R}^k$. Como x e y están ordenados, 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

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

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 $x = (a, w) \leqslant (a, v) = u$. Por otra parte, como $a \prec b$ y $v \prec z$, el Corolario 4.1.10 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 coordenada 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 4.1.13. Probar que, dados $x, y \in \mathbb{R}^n$, entoncecs

1. $x \prec_w y$ si y sólo si existe $v \in \mathbb{R}^n$ tal que $x \prec v \leqslant y$.
2. $x \prec^w y$ si y sólo si $-y \prec_w -x$ si y sólo si existe $w \in \mathbb{R}^n$ tal que $y \geqslant w \prec x$. ▲

4.2 Mayorización y funciones convexas

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

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

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 4.2.1. *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. $\operatorname{tr} f(y) \leq \operatorname{tr} f(x)$ *para toda función convexa* $f : \mathbb{I} \to \mathbb{R}$.
3. $\displaystyle\sum_{i=1}^{n} |y_i - t| \leq \sum_{i=1}^{n} |x_i - t|$ *para todo* $t \in \mathbb{R}$.

Análogamente, son equivalentes

1'. $y \prec_w x$ *(submayorización)*

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

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

Demostración. Sólo probaremos la primer parte, puesto que los argumentos para probar la segunda son similares (para 1' $\to$ 2', se aplica 1 $\to$ 2 y la Proposición 4.1.12, que será útil para funciones no decrecientes). Supongamos que $y \prec x$. Entonces, por el Teorema 4.1.8, $y = \displaystyle\sum_{i=1}^{s} \lambda_i P_{\sigma_i}(x)$ para ciertos $\lambda_i \geq 0$ que suman uno, y para ciertas $\sigma_i \in \mathbf{S}_n$. Luego $f\big(\sum\limits_{i=1}^{s} \lambda_i P_{\sigma_i} x\big) \leqslant \sum\limits_{i=1}^{s} \lambda_i f\big(P_{\sigma_i} x\big)$ (en cada coordenada), y

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

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 coordenadas). 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

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

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 $\sum_{i=1}^{k} y_i \le \sum_{i=1}^{k} x_i$. ■

Corolario 4.2.2. *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 4.2.1, si $x \prec y$, entonces

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

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 4.1.12 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 Lema 4.1.11. ■

Corolario 4.2.3. *Sean $x, y \in \mathbb{R}^n$, tales que $x > 0$ e $y > 0$. Entonces, se tiene que*

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

Demostración. Sea $g(t) = -\log t$, que es una función convexa (pero decreciente), definida en $\mathbb{I} = (0, +\infty)$. Por el Corolario 4.2.2, 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 $\prod\limits_{i=1}^{n} x_i \geq \prod\limits_{i=1}^{n} y_i$. ■

4.3 Birkhoff, Hall y los casamientos

El objetivo de esta sección es probar el teorema de Birkhoff que asegura que $\mathcal{U}_{\mathcal{P}}(n)$ es el conjunto de puntos extremales de $\mathcal{DS}(n)$ y por ende (ya sea por el teormea de Krein Millman, o porque va a salir a mano), que toda $A \in \mathcal{DS}(n)$ es combinación convexa de matrices de permutación. Observar que este hecho va a explicar una parte del Teorema 4.1.8. La herramienta clave, en función de poder probar lo anterior inductivamente, son dos resultados combinatorios que son interesantes por sí mismos: El teorema de los casamientos de Hall, y el criterio de König-Frobenius sobre existencia de diagonales sin ningún cero para matrices. Empecemos con las "notaciones" para 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

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

Es claro que esta notación es machista, pero en la época de Hall nadie cuestionaba esas cosas. 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 4.3.1 (El problema de los casamientos de Hall)**.** *El PC tiene solución para una relación $C \subseteq V \times M$ si y sólo si*

$$|M_J| \geq |J| \qquad \textit{para todo} \qquad J \subseteq \mathbb{I}_n \ . \tag{4.6}$$

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 (4.6), a saber,

$$|M_J| \geq |J| + 1 \qquad \text{para todo} \qquad J \subseteq \mathbb{I}_n \ , \quad \emptyset \neq J \neq \mathbb{I}_n \ . \tag{4.7}$$

En tal caso fijamos al vago v_n de V y lo casamos con una chica que conozca ($m_j \in M_n$, o sea que $(n, j) \in C$). Veamos ahora que, si $J = \mathbb{I}_{n-1}$, los conjuntos V_J y $M \setminus \{m_j\}$ cumplen la condición (4.6) (para aplicarles la HI). En efecto, notar que si $I \subseteq J$, entonces la Eq. (4.7) asegura que $|M_I \cap M \setminus \{m_j\}| \geq |M_I| - 1 \geq |I|$. En otras palabras, dados k de los muchachos restantes, 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

$$\emptyset \neq J \neq \mathbb{I}_n \qquad \text{y} \qquad |M_J| = |J| = k < n \ , \tag{4.8}$$

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 (4.8), es fácil ver que los conjuntos que quedan, V_{J^c} y $M \setminus M_J$ cumplen también la condición (4.6). 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 4.3.2. Sea $A \in \mathcal{M}_n(\mathbb{C})$.

1. Dada $\sigma \in \mathbf{S}_n$, llamamos al vector $(a_{1\sigma(1)}, \dots, a_{n\sigma(n)})$ 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 coordenadas no nulas. ▲

Corolario 4.3.3 (König-Frobenius). *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces son equivalentes:*

1. *Toda diagonal de A tiene ceros.*
2. *Existen 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. Consideremos 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 4.3.1, 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 4.3.4. *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 4.3.3, 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 4.3.5 (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 $n \le k(A) \le n^2$, 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 4.3.4 existe $\sigma \in \mathbf{S}_n$ 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 Eq. (4.3), $P_{ij} \neq 0$ 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 $k(B) \le k(A)$. 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. ■

Observación 4.3.6. El Teorema de Birkhoff está intimamente relacionado con el Teorema 4.1.8. Sin embargo, no se implican mutuamente. En efecto, el Teorema 4.3.5 da como novedad la implicación 3 $\implies$ 2 del Teorema 4.1.8. Pero esta salía implícita por el rulo de auqella prueba, y el Teorema 4.3.5 no dice que si $y \prec x$ entonces haya una $A \in \mathcal{DS}(n)$ tal que $Ax = y$. Mirándolos ahora al reves, la implicación 3 $\implies$ 2 del Teorema 4.1.8 dice que **para cada** $x \in \mathbb{R}^n$, hay una combinación convexa de permutaciones que hace, en ese x, lo mismo que hace A. Pero no que haya una que sirva para todos los x a la vez. ▲

4.4 Mayorización logarítmica

Definición 4.4.1. 1. Sean $x, y \in \mathbb{R}_+^n$. Escribiremos $x \underset{\log}{\prec_w} y$ si

$$\prod_{i=1}^{k} x_i^\downarrow \le \prod_{i=1}^{k} y_i^\downarrow \quad \text{para todo} \quad k \in \mathbb{I}_n \,. \tag{4.9}$$

2. Si $x, y > 0$, escribimos $x \underset{\log}{\prec} y$ si se cumple que $x \underset{\log}{\prec_w} y$, y además $\prod_{i=1}^{n} x_i = \prod_{i=1}^{n} y_i$. ▲

Observación 4.4.2. Sean $x, y \in \mathbb{R}_+^{*n}$. Si $x \underset{\log}{\prec} y$ entonces, como en el caso de la mayorización común, se cumplen desigualdades invesas para las entradas mas pequeñas de x e y. Es decir,

$$\prod_{i=k}^{n} x_i^\downarrow \ge \prod_{i=k}^{n} y_i^\downarrow \quad \text{para todo} \quad k \in \mathbb{I}_n \,. \tag{4.10}$$

Esto no vale si solo $x \underset{\log}{\prec_w} y$, porque usa la igualdad de los productos hasta n. ▲

Proposición 4.4.3. *Sean $x, y \in \mathbb{R}_+^n$.*

1. *Si $x \underset{\log}{\prec_w} y$, entonces $x^p \prec_w y^p$ para todo $p \in \mathbb{R}_+$.*
2. *Si $x, y > 0$, entonces $x \underset{\log}{\prec} y$ implica que $x^p \prec_w y^p$ para todo $p \in \mathbb{R}$.*

Demostración.

2. Supongamos que $x > 0$ e $y > 0$. Luego $x \underset{\log}{\prec} y$ implica que $\log x \prec \log y$. Como la función $t \mapsto e^{pt}$ es convexa para todo $p \in \mathbb{R}$, deducimos lo afirmado en el item 2. a partir del Corolario 4.2.2.

1. Si $x, y \in \mathbb{R}_+^n$ y $x \underset{\log}{\prec_w} y$, supongamos que x e y están ordenados en forma decreciente. Observar que, si $k_1 = \max\{j \in \mathbb{I}_n : x_j > 0\}$, entonces la condición (4.9) implica que $y_{k_1} > 0$. Podemos suponer, entonces, que $k_1 = n$, porque las desigualdades que definen la relación $x^p \prec_w y^p$ serán automáticas para $k > k_1$. Estamos casi en el caso anterior, dado que la única diferencia es que tenemos $\log x \prec_w \log y$ en lugar de $\log x \prec \log y$. Pero esto es suficiente si suponemos que $p \in \mathbb{R}_+$, porque en tal caso se tiene que la función $t \mapsto e^{pt}$ es convexa *creciente*, y se aplica nuevamente el Corolario 4.2.2. ■

Observación 4.4.4. 1. El caso más usado de la Proposición 4.4.3 es cuando $p = 1$. Es decir que, si $x, y \in \mathbb{R}_+^n$, entonces $x \underset{\log}{\prec_w} y$ implica $x \prec_w y$. Esto será sumamente util cuando se lo aplique a desigualdades con valores singulares de matrices, usando técnicas de productos alternados. Observar que, en este caso, el Corolario 4.2.3 nos dice que, si hubese quedado $x \prec y$, debía cumplirse que $x \underset{\log}{\prec} y$.

2. Por otra parte, la Proposición 4.4.3 se puede generalizar, sin cambiar la prueba, si remplazamos las funciones $f(t) = t^p$ por cualquier función $f : \mathbb{R}_+ \to \mathbb{R}$ tal que la aplicación $t \mapsto f(e^t)$ resulte convexa (o convexa *creciente*). Notar que, en el caso demostrado, se usaba esa propiedad para la función $t \mapsto (e^t)^p = e^{pt}$. ▲

4.5 Ejercicios

Ejercicios del texto

4.5.1. Sea $A \in \mathcal{M}_n(\mathbb{R})$. Probar que $A \in \mathcal{DS}(n)$ si y sólo si

$$A \geqslant 0 \ , \quad A\,\mathbb{1} = \mathbb{1} \quad \text{y} \quad A^*\,\mathbb{1} = \mathbb{1},$$

Deducir que $\mathcal{DS}(n)$ es un conjunto convexo, y que dadas dos matrices $A, B \in \mathcal{DS}(n)$, entonces también $AB \in \mathcal{DS}(n)$.

4.5.2. Probar los 9 items de la Observación 4.1.5.

4.5.3. Probar que, dados $x, y \in \mathbb{R}^n$, entoncecs

1. $x \prec_w y$ si y sólo si existe $v \in \mathbb{R}^n$ tal que $x \prec v \leqslant y$.
2. $x \prec^w y$ si y sólo si $-y \prec_w -x$ si y sólo si existe $w \in \mathbb{R}^n$ tal que $y \geqslant w \prec x$. ▲

Ejercicios nuevos

4.5.4. Sea $x \in \mathbb{R}^n$. Probar que

$$\sum_{i=1}^{k} x_i^{\downarrow} = \min\ \left\{ \|y\|_1 + k\|z\|_\infty : x = y + z \right\} \quad \text{para todo } k \in \mathbb{I}_n \ .$$

4.5.5. Sean $x, y \in \mathbb{R}^n$. Demostrar las siguientes propiedades:

1. $x \prec y$ si y sólo si $x \prec_w y$ junto con $x \prec^w y$.
2. Si $\alpha > 0$, entonces
$$x \prec_w y \implies \alpha x \prec_w \alpha y \quad \text{y} \quad x \prec^w y \implies \alpha x \prec^w \alpha y \ .$$
3. $x \prec_w y \iff -x \prec^w -y$.
4. Para cualquier $\alpha \in \mathbb{R}$ no nulo, se tiene que $x \prec y \iff \alpha x \prec \alpha y$.

4.5.6. Una transformación lineal A en $\mathbb{C}^n$ se dice que *preserva positividad* si lleva vectores de coordenadas no negativas en vectores de coordenadas no negativas. Se dice que *preserva la traza* si $\operatorname{tr} Ax = \operatorname{tr} x$ para todo x. Se dice *unital* si $A\mathbb{1} = \mathbb{1}$.

Dada $A \in \mathcal{M}_n(\mathbb{C})$, probar que $A \in \mathcal{DS}\,(n)$ si y sólo si la transformación lineal A preserva positividad, preserva la traza, y es unital. Mostrar que A preserva la traza si y sólo si A^* es unital.

4.5.7. Sea $A \in \mathcal{DS}\,(n)$.

1. Sea $x \in \mathbb{R}^n$. Si $|x| = (|x_1|, \dots, |x_n|)$, demostrar que
$$|Ax| \leqslant A(|x|) \qquad (\leqslant \text{ significa coordenada a coordenada}) \ .$$
2. Demostrar que $1 \in \sigma(A)$ y que 1 admite un autovector de coordenadas positivas.
3. Demostrar que $1 = \rho(A) = \|A\|_{sp}$.

4.5.8. Sean $x, y \in \mathbb{R}^n$. Probar que

1. $x^\downarrow + y^\uparrow \prec x + y \prec x^\downarrow + y^\downarrow$.
2. Si $x, y \in \mathbb{R}^n_+$, entonces $x^\downarrow \circ y^\uparrow \prec_w x \circ y \prec_w x^\downarrow \circ y^\downarrow$.
3. Si $x, y \in \mathbb{R}^n_+$, entonces $(x, y) \prec (x + y, 0)$ en $\mathbb{R}^{2n}$.
4. Probar que son equivalentes
   a. $x \prec_w y$
   b. Existe $w \in \mathbb{R}^n$ tal que $x \prec w \leq y$.

4.5.9. Sea $f : \mathbb{R}_+ \to \mathbb{R}_+$ cóncava y tal que $f(0) = 0$. Probar que:

1. f es creciente (¿es continua?).
2. Si $c, d \in \mathbb{R}_+$, entonces $f(c + d) \leq f(c) + f(d)$.

3. Si $x, y \in \mathbb{R}^n_+$,
$$x \prec y \quad \Longrightarrow \quad \sum_{i \in \mathbb{I}_n} f(x_i) \geq \sum_{i \in \mathbb{I}_n} f(y_i).$$
Más aún, la misma desigualdad sigue valiendo si $x \prec^w y$.

4. Si $x, y \in \mathbb{R}^n$,
$$\sum_{i \in \mathbb{I}_n} f(|x_i + y_i|) \leq \sum_{ij} f(|x_i|) + \sum_{i \in \mathbb{I}_n} f(|y_i|) \ .$$

4.5.10. Sean $x, y, u \in \mathbb{R}^n$ tales que sus coordenadas estan ordenadas en forma decreciente. Probar que:

1. $x \prec y \quad \Longrightarrow \quad \langle x, \ u \rangle \leq \langle y, \ u \rangle$.
2. Si solo pedimos que $x \prec_w y$ pero tenemos que $u \in \mathbb{R}^n_+$, entonces $\langle x, \ u \rangle \leq \langle y, \ u \rangle$.

Capítulo 5

Mayorización de autovalores y valores singulares

En este capítulo estudiaremos numerosas apariciones de relaciones de mayorización entre vectores reales asociados a matrices, particularmente a vectores de autovalores, de valores singulares y a diagonales de matrices. Mostraremos el teorema de Schur-Horn, que caracteriza **todas** las posibles diagonales de matrices en la órbita unitaria de una matriz $A \in \mathcal{H}(n)$ como aquellos vectores $d \in \mathbb{R}^n$ tales que $d \prec \mu(A)$. Luego nos concentraremos en una caracterización de las normas unitariamente invariantes, y daremos un criterio para mostrar desigualdades para cualquiera de esas normas, en términos de mayorización débil de vectores de valores singulares. Después daremos numerosas propiedades relacionadas con la mayorización aplicada a matrices autoadjuntas, que serán herramientas imprescindibles para atacar distintas desigualdades en los siguientes capítulos. Culminaremos con una prueba bastante reciente y asombrosamente simple del famoso teorema de Lidskii, sobre relaciones de mayorización entre vectores de autovalores de **restas** de matrices autoadjuntas, lo que da un expectacular avance sobre el teorema de Weyl 5.1.6, que hace lo mismo con las sumas.

El teorema de Lidskii fue particularmente famoso porque fue anunciado sin dar detalles de la prueba (en los años 60). Esto generó polémicas, que dieron lugar a numerosas demostraciones independientes (en el libro de Bhatia [3] aparecen cuatro, todas bien complicadas). Sin embargo, la que presentamos aquí (debida a Chi-Kwong Li y R. Mathias [28]) es particularmente simple, y usa como única herramienta importante el teorema de Weyl, que por muchos años fue considerado escencialmente más débil. Finalizaremos dando algunas de las importantes aplicaciones del Teorema de Lidskii.

5.1 Aplicaciones a matrices Hermitianas

Teorema 5.1.1 (Teorema de mayorización de Schur 3)**.** *Sea $A \in \mathcal{H}(n)$. Recoredemos la notación $\mathrm{d}\,(A) = (a_{11}, \dots, a_{nn}) \in \mathbb{R}^n$. Entonces, se tiene que*

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

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

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

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 4.1.8 completa la demostración. ■

Observación 5.1.2. Otra demostración del Teorema mayorización de Schur puede hacerse por inducción, aplicando el Teorema de entrelace de Cauchy 2.4.2. Para ello basta reordenar la diagonal de A, conjugandola por una matriz de permutación como se hace en la Observación 4.1.5, lo que no cambia sus autovalores. ▲

Observación 5.1.3. La matriz $B = (|u_{ji}|^2)_{ij} \in \mathcal{DS}(n)$ que aparece en el teorema anterior (para cierta $U \in \mathcal{U}(n)$), es un tipo especial de matriz doble estocástica. A tales matrices se las llama **ortoestocásticas**. Este simpático nombre, que proviene de construir elementos de $\mathcal{DS}(n)$ a partir de matrices unitarias (que en el caso real se llaman ortogonales), está bien elegido porque no toda matriz $A \in \mathcal{DS}(n)$ tiene la suerte de provenir de una matriz unitaria. Por ejemplo $A = \dfrac{1}{2}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \in \mathcal{DS}(3)$, pero no es ortoestocástica, porque no hay modo de que sus columnas provengan de vectores ortogonales. ▲

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

Proposición 5.1.4 (Principio del Máximo de Ky Fan)**.** *Sea $A \in \mathcal{H}(n)$. Entonces*

$$\sum_{j=1}^{k} \mu_j(A) = \text{máx} \sum_{j=1}^{k} \langle Ax_j \, , x_j \rangle \ , \qquad \textit{para todo } k \in \mathbb{I}_n \quad ,$$

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 ortonormal cualquiera. Sea $U \in \mathcal{U}(n)$ tal que sus primeras k columnas sean los vectores dados. Sea $B = U^*AU$. Luego

$$\mu(B) = \mu(A) \quad \text{y} \quad \sum_{j=1}^{k} b_{jj} = \sum_{j=1}^{k} \langle Be_j \, , e_j \rangle = \sum_{j=1}^{k} \langle A\,Ue_j \, , Ue_j \rangle = \sum_{j=1}^{k} \langle Ax_j \, , x_j \rangle \ .$$

Pero, por el Teorema de mayorización de Schur 5.1.1, se tiene que

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

Para ver la otra desigualdad, tomemos $\mathcal{B} = \{v_1\,, ..., v_n\}$ una BON de $\mathbb{C}^n$ adaptada a $\mu(A)$. Luego $\{v_1, ..., v_k\}$ es una k-upla ortonormal y

$$\text{máx} \sum_{j=1}^{k}\langle Ax_j\,, x_j\rangle \ge \sum_{j=1}^{k}\langle Av_j\,, v_j\rangle = \sum_{j=1}^{k}\langle \mu_j(A)\ v_j\,, v_j\rangle = \sum_{j=1}^{k} \mu_j(A) \ ,$$

como queríamos demostrar. ■

Ejercicios 5.1.5. 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 $\mathcal{P}_k(n) = \{P \in \mathcal{H}(n) : P^2 = P \text{ y } \mathrm{rk}(P) = k\}$, entonces

$$\sum_{j=1}^{k} \mu_j(A) = \max_{P\in\mathcal{P}_k(n)} \ \mathrm{tr}\, PAP \ . \tag{5.1}$$

2. Para todo $k \in \mathbb{I}_n$, se tiene que

$$\sum_{j=1}^{k} \mu_j(A) = \max_{U\in\mathcal{U}_k(n)} \ \mathrm{tr}\, U^*AU \ , \tag{5.2}$$

donde $\mathcal{U}_k(n) = \{U \in \mathcal{M}_{n,k}(\mathbb{C}) : U^*\,U = I_k\}$ es el espacio de isometrías de $\mathbb{C}^k$ en $\mathbb{C}^n$. ▲

Teorema 5.1.6 (Weyl). *Sean A y $B \in \mathcal{H}(n)$. Entonces*

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

Es importante el hecho de que, en la suma $\mu(A) + \mu(B)$, se asume que ambos vectores están ordenados de la misma forma.

Demostración. Por la fórmula (5.1), para todo $k \in \mathbb{I}_n$ podemos escribir

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

La igualdad para $k = n$ surge de que $\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$. ■

Recordar que en la Proposición 3.7.5 probamos que, dada $C \in \mathcal{M}_n(\mathbb{C})$, entonces si

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

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

$$\mu(\widehat{C}) = (s_1(C), \cdots, s_n(C), -s_n(C), \cdots, -s_1(C)\,) \ . \tag{5.3}$$

Corolario 5.1.7. *Sean $A, B \in \mathcal{H}(n)$. Entonces*

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

Demostración. Notar que $\widehat{A + B} = \widehat{A} + \widehat{B}$. Por la Eq. (5.3) y las desigualdades resultantes de la relación $\mu(\widehat{A + B}) \prec \mu(\widehat{A}) + \mu(\widehat{B})$ para los n primeros índices $k \in \mathbb{I}_n \subseteq \mathbb{I}_{2n}$, se tiene que $s(A + B) \prec_w s(A) + s(B)$. ■

Observación 5.1.8. Notar que el resultado anterior es necesario para verificar que las normas $\|\cdot\|_{(k)}$ de Ky Fan, para cada $k \in \mathbb{N}$, definidas en el Ejemplo 3.4.1, cumplen la desigualdad triangular. ¿Les había salido el ejercicio? ▲

5.2 Teorema de Schur-Horn

5.2.1. Sea $x \in \mathbb{C}^n$ con $\|x\| = 1$ (a estos vectores los llamaremos *unitarios*). Entonces, como vimos en 1.9.3, la matriz $P_x = x \odot x = xx^* = (x_i\overline{x_j})_{ij} \in \mathcal{M}_n(\mathbb{C})$ es el proyector ortogonal sobre el subespacio $\operatorname{Gen}\{x\}$. Por lo tanto, si $\mathcal{B} = \{x_1, \dots, x_n\}$ es una BON de $\mathbb{C}^n$, vale que

$$z = \sum_{i=1}^{n} \langle z, x_i\rangle x_i \quad \text{para todo} \ \ z \in \mathbb{C}^n \qquad \Longrightarrow \qquad I = \sum_{i=1}^{n} x_i \odot x_i \ . \tag{5.4}$$

▲

Proposición 5.2.2. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Se tiene que $A \in \mathcal{M}_n(\mathbb{C})^+$ si y sólo si existen vectores unitarios $x_1, \dots, x_r \in \mathbb{C}^n$, y números $\lambda_1, \dots, \lambda_r \in \mathbb{R}_+$ tales que*

$$A = \sum_{i=1}^{r} \lambda_i \ x_i \odot x_i \ , \qquad \textit{o sea que} \qquad A = \sum_{i=1}^{r} \lambda_i \ P_{x_i} \ . \tag{5.5}$$

Demostración. Por un lado es claro que si A cumple (5.5), entonces $A \geq 0$. Recíprocamente, si $A \in \mathcal{M}_n(\mathbb{C})^+$, sea $\mathcal{B} = \{x_1, \dots, x_n\}$ es una BON de $\mathbb{C}^n$ adaptada a $\mu(A)$. Usando la ecuación (3.12), para todo $z \in \mathbb{C}^n$ se tiene que,

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

Luego A cumple (5.5). ■

Observación 5.2.3. Notar que la mínima cantidad r de proyectores de rango uno que puede usarse para obtener una representación de A como en (5.5) es $r = \operatorname{rk}(A)$. Además, si $A \in \mathcal{M}_n(\mathbb{C})^+$ cumple (5.5), definiendo $y_i = \lambda_i^{1/2} x_i$, se tiene que las matrices $y_i \odot y_i$ no son más proyectores, pero sí son positivos y se verifica que $A = \sum_{i=1}^{r} y_i \odot y_i$. ▲

Es natural preguntarse, dado $A \in \mathcal{M}_n(\mathbb{C})^+$ y $r \geq \operatorname{rk}(A)$, para qué sucesiones $\lambda_1, \dots, \lambda_r$ en $\mathbb{R}_+$ se puede obtener para A una representación como (5.5). 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 5.2.4. *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* $A = \sum_{j=1}^{n} c_j \; x_j \odot x_j$.
2. Existe $B \in \mathcal{M}_n(\mathbb{C})^+$ tal que $\mathrm{d}\,(B) = c$ *y* $\mu(B) = \mu(A)$*, o sea que* $B \cong A$.

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 $XX^* = A$: Para cada $k \in \mathbb{I}_n$, notemos por $X_k \in \mathcal{M}_n(\mathbb{C})$ a la matriz tal que $C_k(X_k) = C_k(X)$, pero $C_j(X_k) = 0$ si $j \neq k$. Luego, se tiene que

$$X = \sum_{k \in \mathbb{I}_n} X_k \ , \quad X_k X_j^* = 0 \quad \text{si} \quad j \neq k \quad \text{y} \quad X_k X_k^* = c_k \; x_k x_k^* = c_k \; x_k \odot x_k \ .$$

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 $B \in \mathcal{M}_n(\mathbb{C})^+$ 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 $\operatorname{diag}(\mu(A))$). 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

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

lo que prueba 1. ■

Ahora sí, podemos probar la recíproca del Teorema 3 de Schur , sobre mayorizacón entre la diagonal y los autovectores. Este resultado se debe a R. Horn, y poniéndolos juntos se los conoce como el Teorema de Schur-Horn, que tiene importantes aplicaciones y generalizaciones a operadores en espacios de Hilbert y a álgebras de operadores. Como ingrediente extra (clave para la prueba) se incluye una tercera condición equivalente (en el caso positivo), que tiene que ver con lo que veníamos viendo, o sea el expresar una matriz $A \in \mathcal{M}_n(\mathbb{C})^+$ como suma de múltiplos de proyectores unidimensionales.

Teorema 5.2.5. *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* $x_1, \ldots, x_n \in \mathbb{C}^n$ *tales que*

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

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, usando la Observación 4.1.5). Notemos $A = \operatorname{diag}(b)$. El Teorema 3 de Schur 5.1.1 muestra que 2 implica 1. La Proposición 5.2.4 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 $b_1 \geq c_1 \geq b_n$, podemos tomar un $k \in \mathbb{I}_{n-1}$ tal que $b_k \geq c_1 \geq b_{k+1}$.

Se afirma: existe $x_1 \in \operatorname{Gen}\{e_k, e_{k+1}\}$ 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 BON $\{y_1, y_2\}$ de $\operatorname{Gen}\{e_k, e_{k+1}\}$ 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 $\mathcal{B} = \{y_2, e_1, \ldots, e_{k-1}, y_1, e_{k+2}, \ldots, e_n\}$ queda

$$\Big[A_1\Big]_{\mathcal{B}} = \operatorname{diag}\left(0, b_1, \ldots, b_{k-1}, (b_k + b_{k+1} - c_1), b_{k+2}, \ldots, b_n\right) \ . \tag{5.6}$$

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

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

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 $k+1 \le r \le n-1$,

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

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

$$\sum_{j=2}^{n} c_j \; (0, z_j)(0, z_j)^* = \operatorname{diag}(0, d) = \Big[A_1 \Big]_{\mathcal{B}} \in \mathcal{M}_n(\mathbb{C})^+. \tag{5.7}$$

Si definimos $x_2, \dots, x_n \in \mathbb{C}^n$ tales que las coordenadas 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 (5.7) (pensada en la base $\mathcal{B}$) a coordenadas 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 $1 \to 3$ y $3 \leftrightarrow 2$ ($2 \to 1$ era el Teorema 3 de Schur). Pero esto se generaliza sin dificultad al caso general (con b y c cualesquiera en $\mathbb{R}^n$) usando que para todo $m \in \mathbb{R}$ se tiene que

$$x \,+\, m\,\mathbb{1} \;\prec\; y \,+\, m\,\mathbb{1} \quad \Longleftrightarrow \quad x \,\prec\, y \;,$$

y que dada $B \in \mathcal{H}(n)$, entonces $\operatorname{d}(B + mI) = \operatorname{d}(B) + m\mathbb{1}$ y $\mu(B + mI) = \mu(B) + m\mathbb{1}$.

Finalmente, probado $1 \leftrightarrow 2$ en general, ahora por la Proposición 5.2.4, ya sabemos que 3 es equivalente a ellas si b y c tienen entradas no negativas. ■

Teorema 5.2.6. *Sea $a \in \mathbb{R}^n$ y $A \in \mathcal{H}(n)$ tal que $\mu(A) = a^{\downarrow}$. Entonces,*

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

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 5.1.1, se tiene que $\operatorname{d}(B) \prec a$.

Recíprocamente, si $x \in \mathbb{R}^n$ cumple $x \prec a$, por el Teorema 5.2.5 existe $B \in \mathcal{H}(n)$ tal que $\operatorname{d}(B) = x$ y $\mu(B) = a^{\downarrow}$. Por lo tanto debe existir $U \in \mathcal{U}(n)$ tal que $B = UAU^*$. Luego $x \in \{\operatorname{d}(UAU^*) : U \in \mathcal{U}(n)\}$. ■

Corolario 5.2.7. *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 5.2.6, $c \prec \tilde{\mu}(A)$ si y sólo si existe $U \in \mathcal{U}(m)$ tal que $\mathrm{d}\,(UA_1U^*) = c$. Es claro que si

$$P = \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{M}_m(\mathbb{C})^+ , \qquad \text{entonces} \qquad UA_1U^* = UPA_1PU^* .$$

Luego, 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 todo lo siguiente:

$$T = A^{1/2}U_1^* \quad , \quad U_1^*U_1 = I_n \quad \text{y} \quad \mathrm{d}\,(U_1AU_1^*) = \mathrm{d}\,(T^*T) = c \ .$$

Luego se extiende U_1 a una $U \in \mathcal{U}(m)$, con lo que $U_1AU_1^* = UA_1U^*$. ■

Observación 5.2.8. El resultado anterior resulve el problema planteado en el párrafo anterior a la Proposición 5.2.4, al menos para el caso $r \geq n$. Es fácil ver, usando el Teorema 5.2.5, 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 $\mu(A) \succ (c, 0, \dots, 0) \in \mathbb{R}^n$. ▲

5.3 Normas unitariamente invariantes

Definición 5.3.1. Dada una norma N en $\mathcal{M}_n(\mathbb{C})$, decimos que N es una norma unitariamente invariante (NUI) , si cumple que

$$N(UAV) \ = \ N(A) \qquad \text{para toda } A \in \mathcal{M}_n(\mathbb{C}) \text{ y todo par } U, V \in \mathcal{U}(n) \ .$$

En tal caso, el Teorema 3.2.5 dice que $N(A) = N(\Sigma(A)\,)$ para toda $A \in \mathcal{M}_n(\mathbb{C})$. ▲

Definición 5.3.2. Sea N una NUI en $\mathcal{M}_n(\mathbb{C})$. Consideremos la función

$$g_N : \mathbb{C}^n \to \mathbb{R}_+ \quad \text{dada por} \quad g_N(x) = N(\operatorname{diag}(x))$$

para todo $x \in \mathbb{C}^n$. ▲

Proposición 5.3.3. *Sea N una NUI en $\mathcal{M}_n(\mathbb{C})$ y sea $x \in \mathbb{C}^n$. Entonces:*

1. *g_N es una norma en $\mathbb{C}^n$.*
2. *$g_N(x) = g(|x|) := g_N(|x_1|, \dots, |x_n|)$.*
3. *$g_N(x) = g(x_\sigma) = g_N(x_{\sigma(1)}, \dots, x_{\sigma(n)})$, para toda $\sigma \in \boldsymbol{S}_n$.*

Demostración.

1. Se deduce de que la aplicación $\mathbb{C}^n \ni x \mapsto \operatorname{diag}(x) \in \mathcal{M}_n(\mathbb{C})$ es lineal e inyectiva.
2. Sea $x_j = \omega_j\,|x_j|$ donde $w_j = e^{i\,\theta_j}$. Como $W = \operatorname{diag}(\omega_1, \dots, \omega_n) \in \mathcal{U}(n)$, tenemos que
$$g_N(|x|) = N(\operatorname{diag}(|x|)) = N(W^*\operatorname{diag}(x)) = N(\operatorname{diag}(x)) = g_N(x) .$$
3. Sea $P_\sigma \in \mathcal{U}_{\mathcal{P}}(n)$ la matriz asociada a σ. Luego $P_\sigma \operatorname{diag}(x)\, P_\sigma^* = \operatorname{diag}(x_\sigma)$. Entonces,
$$g_N(x_\sigma) = N(P_\sigma \operatorname{diag}(x)\, P_\sigma^*) = N(\operatorname{diag}(x)) = g_N(x) .$$ ■

Definición 5.3.4. Una función $f : \mathbb{C}^n \to \mathbb{R}$ que cumple los ítems 1, 2 y 3 de la Proposición anterior se denomina **gauge simétrica**. Abreviaremos esto escribiendo que f es una **fgs**. ▲

Lema 5.3.5. *Si g es una* fgs, *entonces, g es monótona en el siguiente sentido: Si se cumple que $|x_i| \le |y_i|$ para todo $i \in \mathbb{I}_n$, entonces, $g(x) \le g(y)$.*

Demostración. Por la Proposición 5.3.3, podemos suponer que $x, y \in \mathbb{R}_+^n$. Por un argumento inductivo, es suficiente verificar que si $t \in [0,1]$ y $k \in \mathbb{I}_n$, entonces

$$g(y_1, \dots, t\,y_k, \dots, y_n) \le g(y_1, \dots, y_k, \dots, y_n) .$$

En efecto, si tomamos ese $x = (y_1, \dots, t\,y_k, \dots, y_n)$, entonces

$$\begin{aligned} g(x) &= g\left(\frac{1+t}{2}\; y \;+\; \frac{1-t}{2}\; (y_1, \dots, -y_k, \dots, y_n)\right) \\ &\le \frac{1+t}{2}\; g(y) \;+\; \frac{1-t}{2}\; g(y_1, \dots, -y_k, \dots, y_n) \\ &= \frac{1+t}{2}\; g(y) \;+\; \frac{1-t}{2}\; g(y) \;=\; g(y) , \end{aligned}$$ ■

Este Lema nos permitirá mostrar la relación clave que existe entre estas fgs's y la mayorización débil de vectores.

Teorema 5.3.6. *Sea g es una función gauge simétrica en $\mathbb{C}^n$ y sean $x, y \in \mathbb{R}^n_+$ tales que $x \prec_w y$. Entonces, $g(x) \leq g(y)$.*

Demostración. Como $x \prec_w y$, la Proposición 4.1.12 nos asegura que existe $u \in \mathbb{R}^n$ tal que $x \leqslant u \prec y$. Ahora bien, el Lema 5.3.5 garantiza que $g(x) \leq g(u)$ (recordar que $x \geqslant 0$).

Por otro lado, como $u \prec y$, el Teorema 4.1.8 nos dice que $u = \sum_{\sigma \in \mathbf{S}_n} \lambda_\sigma y_\sigma$, para ciertos $\lambda_\sigma \in [0,1]$ tales que $\sum_{\sigma \in \mathbf{S}_n} \lambda_\sigma = 1$. Luego

$$g(u) = g\left(\sum_{\sigma \in \mathbf{S}_n} \lambda_\sigma \ y_\sigma \right) \leq \sum_{\sigma \in \mathbf{S}_n} \lambda_\sigma \ g(y_\sigma) = \sum_{\sigma \in \mathbf{S}_n} \lambda_\sigma \ g(y) \ = \ g(y) \ . \qquad \blacksquare$$

Teorema 5.3.7.

1. *Si N es una NUI en $\mathcal{M}_n(\mathbb{C})$, entonces, g_N es una* fgs *en $\mathbb{C}^n$.*
2. *Si g es una* fgs *en $\mathbb{C}^n$, entonces la función $N_g : \mathcal{M}_n(\mathbb{C}) \to \mathbb{R}_+$ dada por*

   $$N_g(A) = g(s_1(A), \ldots, s_n(A)) \ , \qquad para \quad A \in \mathcal{M}_n(\mathbb{C}) \ ,$$

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

Demostración.

1. Esto es la Proposición 5.3.3.
2. Sólo demostraremos la desigualdad triangular. Las demás propiedades quedan como ejercicio para el lector. Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Luego

   $$N_g(A+B) \ = \ g(s(A+B)) \ \leq \ g(s(A)+s(B)) \ \leq \ g(s(A))+g(s(B)) \ = \ N_g(A)+N_g(B) \ ,$$

   ya que $s(A+B) \prec_w s(A) + s(B)$ y podemos usar el Teorema 5.3.6. ■

Teorema 5.3.8 (Ky Fan)**.** *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 de los Teoremas 5.3.6 y 5.3.7. En efecto, observar que

$$\|A\|_{(k)} \leq \|B\|_{(k)} \quad \text{para todo} \quad k \in \mathbb{I}_n \quad \iff \quad s(A) \prec_w s(B) \ ,$$

y en tal caso se tiene que $g(s(A)) \leq g(s(B))$ para toda fgs. La recíproca es evidente. ■

Ahora saldamos otra deuda contraida en el Ejemplo 3.4.1:

Corolario 5.3.9. *Para todo $n \in \mathbb{N}$ y todo $p \in [1\,,\ \infty)$, la norma p de Schatten, dada por*

$$\|A\|_p = \left(\sum_{i=1}^{n} s_i\,(A)^p \right)^{1/p} = (\operatorname{tr}|A|^p\,)^{1/p} \quad para\ A \in \mathcal{M}_n(\mathbb{C}) \quad ,$$

es efectivamente una norma en $\mathcal{M}_n(\mathbb{C})$, y además es NUI.

Demostración. Se usa el Teorema 5.3.7 y el hecho de que la norma p usual en $\mathbb{R}^n$ es una función gauge simétrica. ■

Corolario 5.3.10. *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 2.3.7, obtenemos que

$$0 \leq A \leq B \implies s_k(A) = \mu_k(A) \leq \mu_k(B) = s_k(B)\ , \quad \text{para todo} \quad k \in \mathbb{I}_n\ .$$

Luego basta aplicar el Teorema 5.3.8. ■

Corolario 5.3.11. *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 $s_k(AB) \leq \|A\|_{sp}\ s_k(B)$ para todo $k \in \mathbb{I}_n$, vista en la fórmula (3.3), y del Teorema 5.3.8.

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

$$x \prec y \quad \implies \quad g(x) \leq g(y)\,.$$

En particular, esto se verifica si g es una fgs.

Demostración. Si $x \prec y$, el Teorema 4.1.8 asegura que $x = \sum\limits_{\sigma \in \mathbf{S}_n} \lambda_\sigma P_\sigma y$. para ciertos $\lambda_\sigma \in [0,1]$ tales que $\sum\limits_{\sigma \in \mathbf{S}_n} \lambda_\sigma = 1$. Entonces

$$g(x) = g\Big(\sum_{\sigma \in \mathbf{S}_n} \lambda_\sigma P_\sigma y \Big) \le \sum_{\sigma \in \mathbf{S}_n} \lambda_\sigma \; g(P_\sigma y) = y \ .$$

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

Corolario 5.3.13. *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, se tiene que $g(s(A)) = g(\mu(A))$. Lo mismo pasa para B, y el resultado se deduce de la Porposición 5.3.12, aplicado a las fgs's $g_k(x) = \sum\limits_{i=1}^{k} x_i^{\downarrow}$, para $k \in \mathbb{I}_n$. ■

Corolario 5.3.14. *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 5.3.13 y del Teorema 5.3.7. ■

5.4 Mayorización de matrices Hermitianas

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

Definición 5.4.1. Si A, $B \in \mathcal{H}(n)$, se dice que A está mayorizada por B y se escribe $A \prec B$ si se verifica que $\mu(A) \prec \mu(B)$. Es decir, $A \prec B$ si

$$\sum_{j=1}^{k} \mu_j(A) \le \sum_{j=1}^{k} \mu_j(B), \quad 1 \le k \le n \quad \text{ y } \quad \operatorname{tr} A = \operatorname{tr} B \ . \qquad \blacktriangle$$

Definición 5.4.2. Sea $A \in \mathcal{M}_n(\mathbb{C})$.

1. Dado un proyector $P \in \mathcal{H}(n)$ (o sea $P = P^2 = P^*$), se define el **pinching** de A como

   $$C_P(A) := PAP + (I-P)A(I-P) \ \in \ \mathcal{H}(n) \ .$$

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

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

   donde los bloques tienen los tamaños adecuados (por ejemplo, $B \in \mathcal{M}_k(\mathbb{C})$). La matriz de $C_P(A)$ tiene siempre esa pinta, si uno trabaja en coordenadas de una BON 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 BON adecuada). Notar que se tiene la siguiente factorización:
$$C_{\mathcal{P}} = C_{P_1} \circ C_{P_2} \circ \cdots \circ C_{P_r} , \tag{5.8}$$
y lo mismo en cualquier otro orden entre los C_{P_i}. ▲

Ejercicios 5.4.3.

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 $A = \sum\limits_{i=1}^{r} \lambda_i P_i$.

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:
   (a) 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, $i \in \mathbb{I}_r$.
   (b) Reduce normas espectrales y preserva trazas (si no sale, ver la Proposición 5.4.9).
   (c) 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 5.4.4. *Sea $A \in \mathcal{H}(n)$ y $\mathcal{P}$ un sistema proyectores en $\mathcal{H}(n)$. Entonces*
$$C_{\mathcal{P}}(A) \prec A.$$

Demostración. Por la Eq. (5.8), 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 $U = P - (I - P) = 2\,P - I \in \mathcal{H}(n)$. Es fácil ver que, si $R(P) = \mathcal{S}$, entonces se tiene que

$$U = P - (I - P) = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{matrix} \mathcal{S} \\ \mathcal{S}^\perp \end{matrix} \in \mathcal{U}(n) \implies 2\ C_P(A) = A + UAU = A + UAU^* \ . \quad (5.9)$$

Pero, como $\mu(UAU^*) = \mu(A)$, por el Teorema de Weyl 5.1.6 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 5.4.5.
1. Clarificar en qué sentido la Proposición 5.4.4 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

$$z \prec w \quad \text{y} \quad x \prec y \quad \implies \quad x + z \prec y + w\ .$$

¿Es cierto si no estan ordenados?

3. Deducir del Teorema 5.1.6 (+ 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 5.4.6. Dado un espacio vectorial $\mathbb{V}$ y un subconjunto $\mathcal{C} \subseteq \mathbb{V}$, 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 5.4.7. *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,

$$\{T \in \mathcal{H}(n) : T \prec A\} = \operatorname{conv}[\mathcal{U}(A)]\,. \quad (5.10)$$

O sea que $T \prec A$ si y sólo si T es combinación convexa de conjugados unitarios de A.

Demostración. Tomemos una matriz

$$T = \sum_{k=1}^{m} \lambda_k \; U_k A U_k^* \; \in \; \operatorname{conv}\left[\mathcal{U}(A)\right] \; ,$$

donde los $U_k \in \mathcal{U}(n)$ y $\lambda \in \mathbb{R}_+^m$ cumple que $\lambda \prec e_1^{(m)}$. Por el Ejercicio 5.4.5,

$$\mu(T) \prec \sum_{k=1}^{m} \mu(\lambda_k \; U_k A U_k^*) = \sum_{k=1}^{m} \lambda_k \; \mu(A) = \mu(A) \implies T \prec A \; .$$

Recíprocamente, sea $T \in \mathcal{H}(n)$ tal que $\mu(T) \prec \mu(A)$. Notaremos $a = \mu(A)$. Con las notaciones de la Observación 4.1.5, el Teorema 4.1.8 dice que

$$\mu(T) = \sum_{\sigma \in \mathbf{S}_n} \lambda_\sigma \; P_\sigma \, a = \sum_{\sigma \in \mathbf{S}_n} \lambda_\sigma \; a_\sigma$$

para ciertos $\lambda_\sigma \in [0,1]$ tales que $\sum\limits_{\sigma \in \mathbf{S}_n} \lambda_\sigma = 1$. Notemos $D = \operatorname{diag}(a)$. Por la Eq. (4.4) de la Observación 4.1.5, se tiene que $P_\sigma D P_\sigma^* = \operatorname{diag}(a_\sigma)$. Por lo tanto,

$$\operatorname{diag}(\mu(T)\,) = \sum_{\sigma \in \mathbf{S}_n} \lambda_\sigma \; P_\sigma D P_\sigma^* \; .$$

Finalmente, si $V, W \in \mathcal{U}(n)$ hacen que $A = V^* D V$ y $T = W \; \operatorname{diag}(\mu(T)\,)\, W^*$, entonces

$$T = W \; \operatorname{diag}(\mu(T)\,)\, W^* = \sum_{\sigma \in \mathbf{S}_n} \lambda_\sigma \; W P_\sigma D P_\sigma^* W^* = \sum_{\sigma \in \mathbf{S}_n} \lambda_\sigma \; (W P_\sigma V) \; A \; (V^* P_\sigma^* W^*) \; .$$

Luego $T \in \operatorname{conv}\left[\{UAU^* : U \in \mathcal{U}(n)\}\right] = \operatorname{conv}\left[\mathcal{U}(A)\,\right]$. ■

Observación 5.4.8. Notar que el Teorema 5.4.7 permite generalizar el Corolario 5.3.14 en el siguiente sentido: Si $A, B \in \mathcal{H}(n)$, por la fórmula (5.10) se tiene que

$$A \prec B \quad \implies \quad N(A) \le N(B)$$

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 NUI N. Para el caso de pinchings, se tiene un resultado más general que la Proposición 5.4.4: ▲

Proposición 5.4.9. *Sean $A \in \mathcal{M}_n(\mathbb{C})$ (no necesariamente autoadjunta) y $\mathcal{P}$ un sistema proyectores en $\mathcal{H}(n)$. Entonces*

$$N\big(C_{\mathcal{P}}(A)\big) \le N(A) \quad , \qquad \textit{para toda norma } N \textit{ que sea } \text{NDUI} \textit{ en } \mathcal{M}_n(\mathbb{C}) \quad . \tag{5.11}$$

Demostración. Por la Eq. (5.8) basta probarlo para un solo proyector $P \in \mathcal{H}(n)$. En tal caso, sea $U = P - (I - P) = 2\,P - I \in \mathcal{U}(n)$. Por la Eq. (5.9), se tiene que

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

Con esto, la desigualdad al tomar N es clara. ■

Observación 5.4.10. Otra forma de verlo es observar que, en el caso general, siempre se verifica que $C_{\mathcal{P}}(A) \in \operatorname{conv}\left[\{UAU^* : U \in \mathcal{U}(n)\}\right]$, 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 (ver el ejercicio 5.6.4). ▲

Proposición 5.4.11. *Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Sea N una NUI. Entonces se tienen las desigualdades*

$$\frac{1}{2}\;N\begin{bmatrix} A+B & 0\\ 0 & A+B\end{bmatrix} \le N\begin{bmatrix} A & 0\\ 0 & B\end{bmatrix} \le N\begin{bmatrix} |A|+|B| & 0\\ 0 & 0\end{bmatrix}\ .$$

Demostración. La primera desigualdad se deduce de que

$$\begin{bmatrix} B & 0\\ 0 & A\end{bmatrix} \cong \begin{bmatrix} A & 0\\ 0 & B\end{bmatrix}\ , \qquad \text{vía la matriz} \qquad \begin{bmatrix} 0 & I\\ I & 0\end{bmatrix} \in \mathcal{U}(n)\ .$$

Para probar la segunda, notar que, si $A = U|A|$ y $B = V|B|$ con $U, V \in \mathcal{U}(n)$, entonces

$$\begin{bmatrix} A & 0\\ 0 & B\end{bmatrix} = \begin{bmatrix} U & 0\\ 0 & V\end{bmatrix}\begin{bmatrix} |A| & 0\\ 0 & |B|\end{bmatrix} \implies N\begin{bmatrix} A & 0\\ 0 & B\end{bmatrix} = N\begin{bmatrix} |A| & 0\\ 0 & |B|\end{bmatrix}\ ,$$

por lo que podemos suponer que $A, B \in \mathcal{M}_n(\mathbb{C})^+$. En tal caso, si $C = A^{1/2}$, $D = B^{1/2}$, y

$$T = \begin{bmatrix} C & 0\\ D & 0\end{bmatrix}\ , \qquad \text{entonces} \qquad \begin{bmatrix} A+B & 0\\ 0 & 0\end{bmatrix} = T^*T\ .$$

Por otra parte, $TT^* \cong T^*T$, por lo que $N(TT^*) = N(T^*T)$. Además,

$$TT^* = \begin{bmatrix} C^2 & CD\\ DC & D^2\end{bmatrix} = \begin{bmatrix} A & CD\\ DC & B\end{bmatrix}.$$

Luego $\begin{bmatrix} A & 0\\ 0 & B\end{bmatrix}$ es un pinching de TT^*, y por la fórmula (5.11), en la Proposición 5.4.9, se tiene que $N\begin{bmatrix} A & 0\\ 0 & B\end{bmatrix} \le N(TT^*)$. ■

5.5 Teoremas de Lindskii y sus aplicaciones

El teorema de Lidskii tiene tres versiones equivalentes. Comenzaremos enunciando las tres, y luego iremos armando las pruebas.

Teorema 5.5.1 (Lidskii 1)**.** *Sean* $A, B \in \mathcal{H}(n)$*. Entonces*

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

Ejercicio 5.5.2.

1. Decir porqué es incorrecta la siguiente prueba de la primera parte del Teorema de Lidskii 1: Si $A, B \in \mathcal{H}(n)$, por el Teorema 5.1.6, $\mu(A) \prec \mu(A - B) + \mu(B)$. 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. Si no ven la pifiada, leer la Observación 4.1.9.

2. Demostrar (bien) la otra parte del Teorema (¿quien era $\mu(-B)$?). ▲

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

Teorema 5.5.3 (Lidskii 2)**.** *Sean* $A, B \in \mathcal{M}_n(\mathbb{C})$*. Entonces*

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

Una ves resuelto el Ejercicio 5.5.2, 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 2.3.5 (que, de hecho, es lo que se usa en su prueba). La prueba, asombrosamente simple comparada con las que históricamente la precedieron, necesita un mínimo repaso: Si $A \in \mathcal{H}(n)$, se definen $A_+ = \frac{A+|A|}{2}$ y $A_- = \frac{|A|-A}{2}$. Se probaba en la Sección 3.3 que ambas están en $\mathcal{M}_n(\mathbb{C})^+$, que $A = A_+ - A_-$ y que para todo $k \in \mathbb{I}_n$ se tiene

$$\mu_k(A_+) = \text{máx} \ \{ \ \mu_k(A) \ , \ 0 \ \} \quad \text{ y } \quad \mu_k(A_-) = -\,\text{mín}\{\mu_{n-k+1}(A), 0\} \ . \tag{5.12}$$

Teorema 5.5.4 (Lidskii 3)**.** *Sean* $A, B \in \mathcal{H}(n)$*,* $k \in \mathbb{I}_n$ *y* $J \subseteq \mathbb{I}_n$ *con* $|J| = k$*. Entonces*

$$\sum_{j \in J} \mu_j(A) + \sum_{i=1}^{k} \lambda_i(B) \leq \sum_{j \in J} \mu_j(A + B) \leq \sum_{j \in J} \mu_j(A) + \sum_{i=1}^{k} \mu_i(B) \, . \tag{5.13}$$

Demostración. Probaremos, en principio, la desiguadad de la derecha en (5.13). 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 desigualdad de la derecha en (5.13), 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 descriptas en el repaso previo (sino, ver la Eq. (3.5)). Como $\mu_k(B) = 0$, aplicando la Eq. (5.12) se tiene que

$$\mu_j(B_+) = \text{máx } \{ \ \mu_j(B) \ , \ 0 \ \} \quad , \quad \text{para todo } j \in \mathbb{I}_n \qquad \Longrightarrow \qquad \operatorname{tr}(B_+) = \sum_{j=1}^{k} \mu_j(B) \ .$$

Por el teorema de Weyl, más especificamente el Corolario 2.3.7, el hecho de que

$$A + B \leq A + B_+ \implies \mu_j(A+B) \leq \mu_j(A+B_+) \quad \text{ para todo } j \in \mathbb{I}_n \, .$$

En consecuencia,

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

Por el Corolario 2.3.7, también $\mu_j(A+B_+) \geq \mu_j(A)$, para todo $j \in \mathbb{I}_n$. Por lo tanto

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

Esto prueba la desiguadad de la derecha en la Eq. (5.13). La otra 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 $r \in \mathbb{I}_n$. ■

Observación 5.5.5. Una formulación equivalente del Teorema 5.5.4 que también se usa mucho es la siguiente: Dadas $A, B \in \mathcal{H}(n)$, $k \in \mathbb{I}_n$ y $J \subseteq \mathbb{I}_n$ con $|J| = k$, se tiene que

$$\sum_{j \in J} \mu_j\,(A) - \mu_j(B) \leq \sum_{i=1}^{k} \mu_i\,(A - B) \leq \sum_{j \in J} \mu_j\,(A) - \lambda_j(B) \ . \tag{5.14}$$

En efecto, basta aplicar la Eq. (5.13) a las matrices B y $A - B$, y pasar restando. ▲

Demostración del Teorema de Lidskii 1. Usando la formulación (5.14) del tercer Teorema de Lidskii, obtenemos

$$\sum_{j=1}^{k} \big(\mu(A) - \mu(B)\big)_j^{\downarrow} = \max_{\substack{J \subseteq \mathbb{I}_n \\ |J| = k}} \left\{ \sum_{j \in J} \mu_j\,(A) - \mu_j\,(B) \right\} \leq \sum_{j=1}^{k} \mu_j\,(A - B) \, , \tag{5.15}$$

para todo $k \in \mathbb{I}_n$. Como las trazas están bien, sale que $\mu(A) - \mu(B) \prec \mu(A-B)$. ■

Demostración del Teorema de Lidskii 2. Veamos en principio que, si $k \in \mathbb{I}_n$, entonces

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

En efecto, por la fórmula (3.16) de Proposición 3.7.5, si $j \in \mathbb{I}_n$ se tiene que

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

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

Por otro lado, aplicando la fórmula (5.15) o la (5.14) a $\widehat{A}$, $\widehat{B}$ y a $\widehat{A-B} = \widehat{A} - \widehat{B}$, vemos que

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

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 5.5.6. *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

$$\|\Sigma(A) - \Sigma(B)\|_{(k)} \le \|A-B\|_{(k)} \qquad \text{para todo } \ k \in \mathbb{I}_n \ .$$

Luego se aplica el Teorema 5.3.8. ■

Corolario 5.5.7. *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*

$$\mathrm{d}_N(A, \mathcal{U}(n)\,) = N(A-U) = N(\Sigma(A) - I) = g_N(s(A) - \mathbb{1}) .$$

Demostración. Sea $V \in \mathcal{U}(n)$. Entonces $\Sigma(V) = I$. Por el Corolario 5.5.6, tenemos que $N(A-V) \ge N(\Sigma(A) - I)$. 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 5.5.8. Sea $\mathcal{M}_n(\mathbb{C})_1$ el conjunto de matrices en $\mathcal{M}_n(\mathbb{C})$ de rango uno (o cero). Dada $A \in \mathcal{M}_n(\mathbb{C})$, llamemos $\Sigma_1(A) = \operatorname{diag}(0, s_2(A), \dots, s_n(A))$.

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

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

donde $A = UW\Sigma(A)W^*$. Probar, aparte, que A_1 no depende de la matriz $U \in \mathcal{U}(n)$ elegida para realizar la descomposición polar de A, pero sí puede depender de W (si $s_1(A)$ tiene multiplicidad mayor que uno para $|A|$). Mostrar que otra manera de encontrar A_1 es tomando $A_1 = s_1(A) x \odot y = s_1(A) y x^*$, donde x es un vector unitario tal que $\|Ax\| = \|A\|_{sp} = s_1(A)$, e $y = Ux$. Generalizar el resultado al conjunto de matrices de rango a lo sumo k. ▲

5.6 Ejercicios

Ejercicios del texto

5.6.1. Sea $A \in \mathcal{H}(n)$. Para cada $k \in \mathbb{I}_n$, notamos

$$\mathcal{P}_k(n) = \{P \in \mathcal{H}(n) : P^2 = P \;\text{ y }\; \operatorname{rk}(P) = k\} \quad \text{y} \quad \mathcal{U}_k(n) = \{U \in \mathcal{M}_{n,k}(\mathbb{C}) : U^* U = I_k\} .$$

Probar que, para todo $k \in \mathbb{I}_n$, se tiene que

$$\sum_{j=1}^{k} \mu_j(A) = \max_{P \in \mathcal{P}_k(n)} \operatorname{tr} PAP = \max_{U \in \mathcal{U}_k(n)} \operatorname{tr} U^* A U .$$

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

5.6.3. Dado un sistema de proyectores $\mathcal{P} = \{P_1, \dots, P_r\}$ en $\mathcal{M}_n(\mathbb{C})$, probar que se tiene la siguiente factorización de su pinching:

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

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

5.6.4. Sean $A \in \mathcal{M}_n(\mathbb{C})$ y $\mathcal{P} = \{P_1, \dots, P_r\} \subseteq \mathcal{M}_n(\mathbb{C})$ un sistema de proyectores. Probar que

$$C_{\mathcal{P}}(A) = 2^{-n} \sum_{J \subseteq \mathbb{I}_n} U_J A U_J = 2^{-n} \sum_{J \subseteq \mathbb{I}_n} U_J A U_J{}^* ,$$

donde cada $U_J = \sum_{k \in J} P_k - \sum_{k \notin J} P_k \in \mathcal{U}(n)$. Deducir que $C_{\mathcal{P}}(A) \in \operatorname{conv}[\{UAU^* : U \in \mathcal{U}(n)\}]$.

5.6.5. 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.

5.6.6. 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:

1. 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 , $i \in \mathbb{I}_r$.
2. Reduce normas espectrales y preserva trazas (si no sale, ver la Proposición 5.4.9).
3. 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, $C_\mathcal{P}(A^*) = C_\mathcal{P}(A)^*$.

5.6.7. Dados $x, y, z, w \in \mathbb{R}^n$, todos ordenados en forma decreciente, probar que

$$z \prec w \quad \text{y} \quad x \prec y \quad \Longrightarrow \quad x + z \prec y + w \ .$$

¿Es cierto si no estan ordenados?

5.6.8. Si $A_1, \dots, A_m \in \mathcal{H}(n)$, probar que $\mu\Big(\sum\limits_{k=1}^{m} A_k \Big) \prec \sum\limits_{k=1}^{m} \mu(A_k)$.

5.6.9. Hacer el Ejercicio 5.5.8

Normas débilmente unitariamente invariantes (NDUI's)

Definición 5.6.10. Una norma N en $\mathcal{M}_n(\mathbb{C})$ es una NDUI si cumple que

$$N(A) = N(UAU^*) \quad \text{para toda} \quad A \in \mathcal{M}_n(\mathbb{C}) \quad \text{y toda} \quad U \in \mathcal{U}(n) \ .$$

5.6.11. Probar que las siguientes normas son NDUI's.

1. El radio numérico.
2. $N(A) = \|A\| + |\operatorname{tr} A|$.
3. $N(A) = \displaystyle\int_{\mathcal{U}(n)} m(UAU^*)dU$, donde $m(\cdot)$ una norma en $\mathcal{M}_n(\mathbb{C})$ y dU refiere a la medida de Haar (normalizada) de $\mathcal{U}(n)$.

5.6.12. Sea N una NDUI. Demostrar que

1. Dadas $A, B \in \mathcal{M}_n(\mathbb{C})$, si $A \prec B$, entonces $N(A) \le N(B)$.
2. Dado un sistema de proyectores $\mathcal{P}$ en $\mathbb{C}^n$, se tiene que

$$N(\mathcal{C}_\mathcal{P}(A)\,) \le N(A) \qquad \text{para toda} \quad A \in \mathcal{M}_n(\mathbb{C}) \ .$$

5.6.13. Probar que si N es una NDUI, entonces $N(A) \ge \dfrac{|\operatorname{tr} A|}{n}\, N(I)$. Si no sale a mano, esperar hasta el Corolario 10.2.7.

Ejercicios nuevos

5.6.14. Dadas $A, B \in \mathcal{M}_n(\mathbb{C})^+$, mostrar que

$$\mu(A) \cdot \lambda(B) \prec \mu(AB) \prec \mu(A) \cdot \mu(B) \ .$$

Si sólo tenemos que $A, B \in \mathcal{H}(n)$, entonces mostrar que

$$\langle\, \mu(A)\,,\, \lambda(B)\,\rangle \leq \operatorname{tr} AB \leq \langle\, \mu(A)\,,\, \mu(B)\,\rangle \ .$$

5.6.15. Dada $A \in \mathcal{M}_n(\mathbb{C})$, probar que

1. $\|A\|_{(k)} = \text{mín } \Big\{ \|B\|_1 + k\|C\|_{sp} : A = B + C \ \Big\}$.
2. Usar lo anterior para dar una nueva prueba del Teorema 5.5.3 (Lidskii 2), mostrando previamente que, dadas $A, B \in \mathcal{H}(n)$,
   (a) $\|\mu(A) - \mu(B)\|_\infty \leq \|A - B\|$.
   (b) $\|\mu(A) - \mu(B)\|_1 \leq \|A - B\|_1 = \|A - B\|_{(n)}$.
3. Mostrar que el Teorema 5.5.3 implica el Teorema 5.5.1 (Lidskii 1).

5.6.16. Sea $A \in \mathcal{H}(n)$

1. Sea $\mathcal{S} \subseteq \mathbb{C}^n$ 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:
   (a) $\mu_k(A) \geq \mu_k(A_\mathcal{S}) \geq \mu_{k+1}(A)$, para todo $k \in \mathbb{I}_{n-1}$.
   (b) Sea $v_1, \dots, v_n$ una BON adaptada a $\mu(A)$.
      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 $\mu_{i-1}\,(A_\mathcal{S}) = \mu_i\,(A)$, $k \leq i \leq n$.
2. Probar el Teorema de Lidskii 3 (por inducción sobre n) usando el ejercicio anterior, y considerando independientemente los casos:
   (a) $i_k < n$,
   (b) $1 < i_1$,
   (c) $i_1 = 1$, $i_k = n$. ▲

5.6.17. 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 $g(x^\downarrow - y^\downarrow) \leq g(x - y)$ 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

$$N\big(E_d(A) - E_d(B)\big) \le N(A - B) \le N\big(E_d(A) - E_c(B)\big) \quad \text{y}$$

$$N\big(E_d(A) + E_c(B)\big) \le N(A + B) \le N\big(E_d(A) + E_d(B)\big).$$

5.6.18 (Hoffman-Weilandt y agregados). Sean $A, B \in \mathcal{M}_n(\mathbb{C})$ matrices normales. Sean $\mu(A)$ y $\mu(A)$ sus vectores de autovalores en algún orden.

1. Demostrar que existe una matriz doble estocástica $D \in \mathcal{DS}(n)$ tal que

$$\|A - B\|_2^2 = \sum_{i,j} |\mu_i(A) - \mu_j(B)|^2 \; D_{ij} \ .$$

2. Probar la siguiente desigualdades:

$$\min_{\sigma \in \mathbf{S}_n} \sum_{i=1}^{n} |\mu_i(A) - \mu_{\sigma(i)}(B)|^2 \le \|A - B\|_2^2 \le \max_{\sigma \in \mathbf{S}_n} \sum_{i=1}^{n} |\mu_i(A) - \mu_{\sigma(i)}(B)|^2 \ .$$

3. Sea $B \in \mathcal{M}_n(\mathbb{R})$ diagonal y definamos

$$\mathcal{U}(B) = \{UBU^* : U \in \mathcal{U}(n)\} \subset \mathcal{H}(n).$$

Demostar que dada $C \in \mathcal{M}_n(\mathbb{C})$, la distancia $d_2(C, \mathcal{U}(B))$ (calculada en norma 2, la de Frobenius), se realiza en una matriz diagonal $D \in \mathcal{U}(B)$.

5.6.19. Consideremos el pinching $\mathcal{C} : M_{2n}(\mathbb{C}) \to M_n(\mathbb{C}) \oplus M_n(\mathbb{C})$ dado por

$$\mathcal{C} \begin{bmatrix} X & Y \\ Z & W \end{bmatrix} = \begin{bmatrix} X & 0 \\ 0 & W \end{bmatrix}$$

1. Si $A \in M_{2n}(\mathbb{C})^+$ entonces: existe $U \in \mathcal{U}(2n)$ tal que $\mathcal{C}(U^*AU) = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}$ si y solo si existe unitarios U_1, $U_2 \in \mathcal{U}(2n)$ tales que

$$A = U_1^* \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix} U_1 + U_2^* \begin{bmatrix} 0 & 0 \\ 0 & C \end{bmatrix} U_2 \ .$$

Sugerencia: primero notar que $\mathcal{C}(U^*AU) \ge 0$; representar a A como XX^* con cierta $X \in M_{2n}(\mathbb{C})$ y recordar que XX^* y X^*X son unitariamente equivalentes.

2. Usar lo anterior para probar que para A, $B \in M_n(\mathbb{C})^+$ se tiene que

$$(\mu(A + B), 0) \succ (\mu(A), \mu(B)) \ .$$

Sugerencia: verificar que $\begin{bmatrix} A + B & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ & B \end{bmatrix} \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} \ .$

5.6.20. Probar que para cualquier $A \in \mathcal{M}_{n,m}(\mathbb{C})$ vale:

$$s_j(A) = \max_{\dim \mathcal{S} = j} \ \min_{x \in \mathcal{S}, \|x\|=1} \|Ax\| = \min_{\dim \mathcal{T} = n-j+1} \ \max_{x \in \mathcal{T}, \|x\|=1} \|Ax\|$$

para cualquier $j \in \mathbb{I}_n$.

5.6.21. Para $j \in \{0\} \cup \mathbb{I}_n$, sea

$$R_j = \{T \in \mathcal{M}_{n,m}(\mathbb{C}) \ : \ \operatorname{rk}(T) \le j\} \ .$$

Mostrar que para cualquier $j \in \mathbb{I}_n$, se tiene que

$$s_j(A) = \min_{T \in R_{j-1}} \|A - T\| \ .$$

5.6.22. Mostrar que si $A, H \in \mathcal{M}_{n,m}(\mathbb{C})$, y H tiene rango k, entonces

$$s_j(A) \ge s_{j+k}(A + H) \ , \qquad \text{para todo} \qquad j \in \mathbb{I}_{n-k} \quad .$$

5.6.23. Mostrar que para cualquier $A \in \mathcal{M}_{n,m}(\mathbb{C})$ y para cada $k \in \mathbb{I}_n$, vale

$$\sum_{j=1}^{k} s_j(A) = \max \left| \sum_{j=1}^{k} \langle Ax_j, y_j \rangle \right| ,$$

donde el máximo se toma sobre todas las $k-$uplas ortonormales $x_1, \dots, x_k$ e $y_1, \dots, y_k$.

5.6.24. Sea $A \in \mathcal{H}(n)$.

1. Demostrar que

$$\sum_{j=1}^{k} \mu_j(A) = \max_{U \in \mathcal{U}_k(n)} \operatorname{tr}(UAU^*) \qquad \text{y} \qquad \sum_{j=1}^{k} \lambda_j(A) = \min_{U \in \mathcal{U}_k(n)} \operatorname{tr}(UAU^*) \ .$$

2. Por otro lado, si A es positiva demostrar:

$$\prod_{j=1}^{k} \mu_j(A) = \max_{U \in \mathcal{U}_k(n)} \det(UAU^*) \qquad \text{y} \qquad \prod_{j=1}^{k} \lambda_j(A) = \min_{U \in \mathcal{U}_k(n)} \det(UAU^*) \ .$$

Recordar que $\mathcal{U}_k(n) = \{U \in \mathcal{M}_{n,k}(\mathbb{C}) : U^* U = I_k\}$ es el espacio de isometrías de $\mathbb{C}^k$ en $\mathbb{C}^n$.

5.6.25. Sean $A, B \in \mathcal{M}_n(\mathbb{C})^+$. Probar que

$$\lambda^{\downarrow}(A) \cdot \lambda^{\uparrow}(B) \prec \lambda(AB) \prec \lambda^{\downarrow}(A) \cdot \lambda^{\downarrow}(B).$$

Si sólo pedimos que $A, B \in \mathcal{H}(n)$, mostrar que

$$\langle \lambda^{\downarrow}(A), \lambda^{\uparrow}(B) \rangle \le tr(AB) \le \langle \lambda^{\downarrow}(A), \lambda^{\downarrow}(B) \rangle.$$

5.6.26. Sean N una NUI y $A \in \mathcal{M}_n(\mathbb{C})$. Probar que si $\mu(A)$ es el vector de los autovalores de A y $\text{Eig}(A) = \text{diag}(\mu(A)) \in \mathcal{M}_n(\mathbb{C})$, entonces

$$N(\text{Eig}(A)) = \text{ínf } \{\|SAS^{-1}\| : S \in \mathcal{G}l(n)\}.$$

¿Qué matrices A verifican que el ínfimo es un mínimo para toda NUI? Observar que la conclusión anterior no vale para cada NUI sola. Diga algunas en las que sí vale y otras donde no.

5.6.27 (Mayo conjunta). Dadas $X, Y \in \mathcal{M}_{n,m}(\mathbb{R})$ decimos que

- $Y \prec_s X$ si existe $D \in \mathcal{DS}(n)$ tal que $DX = Y$
- $Y \prec_p X$ si $Yv \prec Xv$ (mayorización usual en $\mathbb{R}^n$) para todo $v \in \mathbb{R}^m$
- $Y \prec_w X$ si existe D estocástica por filas tal que $DX = Y$.

Esta es una manera compacta de describir lo que podríamos llamar *mayorización conjunta* (fuerte, puntual y débil) de m vectores de R^n; a saber, las columnas de X y de Y.

1. Probar que $Y \prec_w X \iff$ las filas de Y son combinaciones convexas de las de X.
2. Probar que $\prec_s \implies \prec_p \implies \prec_w$, pero no valen las recíprocas (el contraejemplo de $\prec_p \implies \prec_s$ es opcional, porque es bastante complicadito).
3. Si $n = 2$ (solo dos filas, o mejor dicho, muchos vectores de $\mathbb{R}^2$), o $m = n$ y las matrices son inversibles (esto significa "bases"), entonces sí vale que $\prec_p \implies \prec_s$.
4. Supongamos que $m = n$ y $Y \prec_w X$. Entonces $|\det X| \geq |\det Y|$. Además, si $|\det X| = |\det Y| \neq 0$, entonces existe una matriz $P \in \mathbf{S}_n$ tal que $Y = PX$.

Normas duales

Definición 5.6.28. Sean Φ una norma en $\mathbb{C}^n$ y N una en $\mathcal{M}_n(\mathbb{C})$. Se definen sus normas del siguiente modo: dados $x \in \mathbb{C}^n$ y $A \in \mathcal{M}_n(\mathbb{C})$, ponemos

$$\Phi'(x) = \sup_{\Phi(x)=1} |\langle x,\ y\rangle| \quad \text{ y } \quad N'(A) = \sup_{|||B|||=1} |\operatorname{tr}(A^*B)| \ .$$

Las normas duales aparecen como las normas de los operadores adjuntos de los vectores o matrices, cuando se los piensa como funcionales según indica el siguiete ejercicio:

5.6.29. 1. Sea ϕ una funcional lineal en $\mathbb{C}^n$. Probar que existe un único

$$x \in \mathbb{C}^n \quad \text{tal que} \quad \phi(y) = \phi_x(y) = \langle y, x\rangle \ , \quad \text{para todo } y \in \mathbb{C}^n \ .$$

2. Sea φ una funcional lineal en $\mathcal{M}_{n,m}(\mathbb{C})$. Probar que existe una única

$$X \in \mathcal{M}_{n,m}(\mathbb{C}) \quad \text{tal que} \quad \varphi(A) = \varphi_X(A) = tr(X^*A) \ , \quad \text{para todo } A \in \mathcal{M}_{mn}(\mathbb{C}) \ .$$

Usando esas identificaciones, definir para las funcionales en $\mathcal{M}_n(\mathbb{C})$, las nociones de adjunta, autoadjunta y positiva, en función de cómo actua en las matrices. Después comparar lo que haya hecho con la Definición 8.3.1.

5.6.30. Sean M una norma en $\mathbb{C}^n$ y N una en $\mathcal{M}_n(\mathbb{C})$. Ellas inducen normas en las funcionales lineales. Dados $x \in \mathbb{C}^n$ y $X \in \mathcal{M}_n(\mathbb{C})$, mostrar que

$$M'(x) = \|\phi_x\|_M = \max_{M(y)=1} |\phi_x(y)| \quad \text{y} \quad N'(X) = \|\varphi_X\|_N = \max_{N(A)=1} |\varphi_X(A)| \ .$$

5.6.31. Sean ϕ y Ψ normas en $\mathbb{C}^n$, Demostrar:

1. $|\langle x,\ y\rangle| \leq \Phi(x)\Phi'(y)$ para todo $x, y \in \mathbb{C}^n$.
2. Si $\Phi(x) \leq c\Psi(x)$ para cierto $c > 0$ y para todo $x \in \mathbb{C}^n$, entonces, $\Phi' \geq c^{-1}\Psi'$

5.6.32. Mostrar que para toda norma Φ en $\mathbb{C}^n$, $\Phi'' = \Phi$, es decir que es igual a la norma dual de su norma dual.

5.6.33. Sea Φ una función gauge simetrica en $\mathbb{C}^n$.

1. Mostrar que Φ' es tambien una fgs.
2. Sea N_Φ la NUI en $\mathcal{M}_n(\mathbb{C})$ asociada a Φ. Probar que $N'_\Phi = N_{\Phi'}$.

5.6.34. Demostrar que

1. $\|A\|'_p = \|A\|_q$, donde $1 \leq p \leq \infty$ y $\frac{1}{p} + \frac{1}{q} = 1$.
2. La única NUI que coincide con su dual es la norma 2.

5.6.35. Dados $k \in \mathbb{I}_n$ y $A \in \mathcal{M}_n(\mathbb{C})$, probar que $\|A\|'_{(k)} = \max\{\|A\|_{(1)}, \frac{1}{k}\|A\|_{(n)}\}$.

5.6.36. Sean p, q, r números reales positivos tales que $1/r = 1/p + 1/q$.

1. Mostrar que para toda función gauge simétrica Φ se tiene que

$$\Phi(|x \circ y|^r)^{1/r} \leq \Phi(|x|^p)^{1/p}\Phi(|y|^q)^{1/q}.$$

2. Probar que para toda norma unitariamente invariante $|||\cdot|||$ se verifica que:

$$||||AB|^r|||^{1/r} \leq ||||A|^p|||^{1/p}||||B|^q|||^{1/q}.$$

Capítulo 6

Funciones monótonas y convexas de operadores

6.1 Cálculo funcional

Notaciones: Diremos que un subconjunto $\mathbb{I} \subseteq \mathbb{R}$ es un *intervalo*, cuando $\mathbb{I}$ es un conjunto convexo (i.e., si $\mathbb{I}$ es un intervalo abierto, semiabierto o cerrado; acotado, semirrecta o todo $\mathbb{R}$). Dado un intervalo $\mathbb{I} \subseteq \mathbb{R}$, llamaremos

$$\mathcal{H}_{\mathbb{I}}(n) = \left\{ A \in \mathcal{H}(n) \ : \ \sigma(A) \subseteq \mathbb{I} \right\} .$$

Definición 6.1.1. Sea $\mathbb{I} \subseteq \mathbb{R}$ un intervalo, y sea $f : \mathbb{I} \to \mathbb{C}$ una función cualquiera. Fijada $A \in \mathcal{H}_{\mathbb{I}}(n)$, se define

$$f(A) = P(A),$$

donde $P \in \mathbb{C}[x]$ verifica que $P(\lambda) = f(\lambda)$ para todo $\lambda \in \sigma(A)$. La definición es buena, porque si $Q \in \mathbb{C}[x]$ cumple lo mismo que P, entonces, por el Corolario 1.7.2,

$$\sigma(P(A) - Q(A)) = \sigma((P - Q)(A)) = (P - Q)(\sigma(A)) = \{0\},$$

y esto, dado que $P(A) - Q(A)$ es normal, implica que $P(A) = Q(A)$. ▲

Observación 6.1.2. Sean $\mathbb{I} \subseteq \mathbb{R}$ un intervalo, $f : \mathbb{I} \to \mathbb{C}$ una función y $A \in \mathcal{H}_{\mathbb{I}}(n)$. Es fácil ver que, si A es diagonl, es decir $A = \operatorname{diag}(x)$ para cierto $x \in \mathbb{I}^n \subseteq \mathbb{R}^n$, entonces

$$f(A) = \operatorname{diag}(f(x)) .$$

Y de esto puede deducirse que, si $B \in \mathcal{H}_{\mathbb{I}}(n)$ y $U \in \mathcal{U}(n)$ son tales que

$$B = U\operatorname{diag}(\mu(B))\, U^* \implies f(B) = U\operatorname{diag}(f(\mu(B)))\, U^*.$$

Otra manera de ver este cálculo es la siguiente: Sea $\sigma(A) = \{\lambda_1, \dots, \lambda_k\}$ (sin repetición). Llamemos $\mathcal{S}_i = \ker(A - \lambda_i I)$, y P_i a los proyectores ortogonales sobre $\mathcal{S}_i$, para $i \in \mathbb{I}_k$.

Luego $\mathcal{P} = \{P_1, \dots, P_k\}$ es un sistema de proyectores en $\mathcal{H}(n)$ (o sea que son autoadjuntos, ortogonales 2 a 2 y que suman I) que verifica que

$$A = \sum_{i=1}^{k} \lambda_i \; P_i \; . \quad \text{Entonces se tiene que} \quad f(A) = \sum_{i=1}^{k} f(\lambda_i) \; P_i \; . \tag{6.1}$$

Por otra parte, notar que este cálculo no está bien definido en matrices que nos son autoadjuntas (en realidad, si no son normales). Por ejemplo, si $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, entonces los polinomios $P(t) = t$ y $Q(t) = t^2$ coinciden en $\sigma(A) = \{0\}$, pero $P(A) = A$ mientras que $Q(A) = 0$. ▲

Ejercicios 6.1.3. Verificar que el cálculo funcional cumple las siguientes propiedades: Sean $\mathbb{I} \subseteq \mathbb{R}$ un intervalo, $f, g : \mathbb{I} \to \mathbb{C}$ dos funciones y $A \in \mathcal{H}_{\mathbb{I}}(n)$. Entonces

1. $(f \pm g)(A) = f(A) \pm g(A)$ y $fg(A) = f(A)g(A)$.
2. $\sigma(f(A)) = \{f(\lambda) : \lambda \in \sigma(A)\}$. Más aún, $\mu(f(A)) = f(\mu(A))^{\downarrow}$.
3. $f(A)$ siempre es una matrix normal.
4. $f(t) \in \mathbb{R}$ para todo $t \in \mathbb{I}$ si y sólo si $f(A) \in \mathcal{H}(n)$ para toda $A \in \mathcal{H}_{\mathbb{I}}(n)$.
5. $f(B) \geq 0$ para toda $B \in \mathcal{H}_{\mathbb{I}}(n)$ si y sólo si $f(t) \geq 0$ para todo $t \in \mathbb{I}$.
6. Si $U \in \mathcal{U}(n)$, entonces $f(UAU^*) = Uf(A)U^*$.
7. Si la matriz de A en alguna BON tiene la forma
$$A = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix} , \qquad \text{entonces} \qquad f(A) = \begin{bmatrix} f(B) & 0 \\ 0 & f(C) \end{bmatrix} .$$
8. Si una sucesión $(f_m)_{m \in \mathbb{N}}$ de funciones definidas en $\mathbb{I}$ convergen **puntualmente** a f, entonces $f_m(B) \xrightarrow[m \to \infty]{} f(B)$ para toda $B \in \mathcal{H}_{\mathbb{I}}(n)$.
9. Si tiene sentido la composición $h \circ f$, entonces $g \circ f(A) = h(f(A))$. ▲

Ejemplos 6.1.4.

1. Si $f : \mathbb{R}_+^* \to \mathbb{R}$ esta dada por $f(t) = t^{-1}$, entonces $f(A) = A^{-1}$ para toda $A \in \mathcal{G}l(n)^+$.
2. Si $A \in \mathcal{M}_n(\mathbb{C})^+$, entonces $A^{1/2} = f(A)$, donde $\mathbb{I} = \mathbb{R}_+$, $f(t) = \sqrt{t}$ y $A^{1/2}$ es la única raiz cuadrada positiva de A definida en la Proposición 3.1.3.
3. Si $A \in \mathcal{H}(n)$, entonces $e^A := \exp(A) = \displaystyle\sum_{m=0}^{\infty} \frac{A^m}{m!}$.
4. Si $A \in \mathcal{G}l(n)^+$, entonces existe $B = \log A$, que es la única matriz autoadjunta que verifica la fórmula $e^B = A$. En efecto, $B = \log A$ está bien definida, y cumple que $e^B = A$ por 9 del Ejercicio 6.1.3. La unicidad se deduce de la fórmula (6.1). ▲

6.1.1 Continuidad del cálculo funcional

Proposición 6.1.5. *Sea $\mathbb{I}$ un intervalo y $f : \mathbb{I} \to R$ una función. Entonces*

1. *$\mathcal{H}_{\mathbb{I}}(n)$ es un subconjunto convexo de $\mathcal{H}(n)$.*
2. *Si $\mathbb{I}$ es es un intervalo abierto, entonces $\mathcal{H}_{\mathbb{I}}(n)$ es abierto en $\mathcal{H}(n)$.*
3. *Si $\varepsilon > 0$ y $g : \mathbb{I} \to \mathbb{R}$ es otra función tal que*
$$\|f - g\|_{\mathbb{I},\infty} := \sup\big\{|f(t) - g(t)| : t \in \mathbb{I}\big\} < \varepsilon \ ,$$
entonces $\|f(A) - g(A)\| < \varepsilon$ para toda $A \in \mathcal{H}_{\mathbb{I}}(n)$.
4. *Si f es* **continua**, *dada una sucesión $(A_m)_{m\in\mathbb{N}}$ en $\mathcal{H}_{\mathbb{I}}(n)$ tal que $A_m \xrightarrow[m\to\infty]{} A \in \mathcal{H}_{\mathbb{I}}(n)$, se tiene que $f(A_m) \xrightarrow[m\to\infty]{} f(A)$.*

Demostración.

1. Sean $A, B \in \mathcal{H}_{\mathbb{I}}(n)$. Dado $\lambda \in [0, 1]$, el Teorema 5.1.6 asegura que
$$x = \mu\big(\lambda A + (1 - \lambda)B\big) \prec \lambda\mu(A) + (1 - \lambda)\mu(B) = y \ .$$
Por lo tanto $x_i \in [y_n\,, y_1] \subseteq \mathbb{I}$ (porque $\mathbb{I}$ es convexo), para todo $i \in \mathbb{I}_n$.
2. Sea $A \in \mathcal{H}_{\mathbb{I}}(n)$, y sea $\varepsilon > 0$ tal que $(\mu_n(A) - \varepsilon\,,\ \mu_1(A) + \varepsilon) \subseteq \mathbb{I}$. Si $B \in \mathcal{H}(n)$ y $\|A - B\| < \varepsilon$, entonces, para todo $x \in \mathbb{C}^n$ con $\|x\| = 1$, se tiene que
$$\big|\langle Ax, x\rangle - \langle Bx, x\rangle\big| \le \|A - B\|\,\|x\|^2 < \varepsilon \ .$$
Luego, por el Teorema 2.3.1, deducimos que
$$\mu_n(A) - \varepsilon < \mu_n(B) \qquad \text{y} \qquad \mu_1(B) < \mu_1(A) + \varepsilon \ ,$$
por lo que $\sigma(B) \subseteq \mathbb{I}$.
3. Notar que $\sigma(f(A) - g(A)) = (f - g)\,\sigma(A)$ y $\|f(A) - g(A)\| = \rho(f(A) - g(A)\,)$.
4. Sea $(a, b) \subseteq \mathbb{R}$ un intervalo abierto tal que
$$\sigma(A) \subseteq (a, b) \cap \mathbb{I} \subseteq [a', b'] = \mathbb{J} \subseteq \mathbb{I} \ .$$
Por el item 2, existe $m_0 \in \mathbb{N}$ tal que $A_m \in \mathcal{H}_{(a,b)}(n) \cap \mathcal{H}_{\mathbb{I}}(n) \subseteq \mathcal{H}_{\mathbb{J}}(n)$, para todo $m \ge m_0$. Por el teorema de Weierstrass (en el intervalo cerrado $\mathbb{J}$), dado $\varepsilon > 0$, existe $P \in \mathbb{C}[x]$ tal que $\|f - P\|_{\mathbb{J},\infty} < \varepsilon$. Entonces, por el item 3, si $m \ge m_0$,
$$\begin{aligned}\|f(A) - f(A_m)\| &\le \|f(A) - P(A)\| + \|P(A) - P(A_m)\| + \|P(A_m) - f(A_m)\| \\ &< 2\,\varepsilon + \|P(A) - P(A_m)\| \xrightarrow[m\to\infty]{} 2\,\varepsilon \ ,\end{aligned}$$
porque el resultado es claro para polinomios. Luego $\|f(A) - f(A_m)\| \xrightarrow[m\to\infty]{} 0$. ∎

Observación 6.1.6. El item 1 de la Proposición 6.1.5 se puede generalizar de la siguiente forma: Dado un conjunto abierto $\mathcal{V} \subseteq \mathbb{C}$, el conjunto

$$\mathcal{M}_n(\mathbb{C})_{\mathcal{V}} \;=\; \{\; A \in \mathcal{M}_n(\mathbb{C}) \;:\; \sigma\,(A) \subseteq \mathcal{V} \;\}$$

es abierto. Más aún, el mismo resultado es cierto cambiando $\mathcal{M}_n(\mathbb{C})$ por $L(\mathbb{H})$, para cualquier espacio de Hilbert $\mathbb{H}$, aún con dimesión infinita. La demostración se deja como ejercicio. ▲

Observación 6.1.7. La noción de cálculo funcional para autoadjuntos que hemos presentado, es una traducción al caso matricial del cálculo funcional continuo para operadores autoadjuntos en espacios de Hilbert. Las únicas diferencias en el caso general son:

1. Las funciones a evaluar deben ser continuas.
2. No existen, en general, polinomios que coincidan con una f dada en todo el espectro del operador elegido (o del intervalo $\mathbb{I}$), por lo que se usa el teorema de Weierstrass para definir $f(A)$ (la buena definición se prueba como en el item 2 de la Proposición 6.1.5).
3. La convergencia útil entre funciones no es la puntual, sino la uniforme en compactos (notar que coinciden en conjuntos finitos).

Todos los demás resultados y ejercicios presentados en esta sección (salvo las menciones específicas de vectores de autovalores, como la Observación 6.1.2) son ciertos en el caso general, con las mismas pruebas, luego de adaptarlas mínimamente a operadores en espacios de Hilbert. La única que necesita más cuidado es la identidad $\sigma\,(f(A)) = f(\sigma\,(A))$, que es fácil para polinomios, pero requiere argumentos especiales para funciones continuas en general. También son ciertos en general los resultados de las próximas dos secciones, dado que las nociones de monotonía y convexidad de operadores se reducen al caso de matrices (siempre que valga para matrices de cualquier tamaño). ▲

6.1.2 Diferenciabilidad del cálculo funcional

En la Proposición 6.1.5 hemos visto que, si $\mathbb{I}$ un intervalo y $f : \mathbb{I} \to \mathbb{R}$ una función continua, entonces la aplicación $\mathbf{f} : \mathcal{H}_{\mathbb{I}}(n) \to \mathcal{H}(n)$ dada por $A \mapsto f(A)$, $A \in \mathcal{H}_{\mathbb{I}}(n)$, es también continua. En caso de que $\mathbb{I}$ sea abierto y que f sea de clase C^1, veremos que $\mathbf{f}$ es diferenciable, y mostraremos cómo calcular sus derivadas direccionales. Sin embargo, como una demostración completa de estos resultados necesita un desarrollo analítico bastante extenso, solo daremos los enunciados y un esbozo de las demostraciones, dejando ciertos pasos técnicos sin demostrar. Para una prueba completa de los resultados de esta sección, remitimos al Capítulo V del libro de Bhatia [3].

Daremos además un resultado probado por Dalekiĭi y Kreĭn [23], [24] (ver también [8] o [3]), el cual provee una herramienta importante para derivar curvas de matrices producidas con el cálculo funcinal, que puede interpretarse como una especie de regla de la cadena. Más adelante, este resultado nos permitirá encontrar una caracterización de las denominadas funciones monótonas de operadores. Para simplificar su enuciado usaremos el producto de

Hadamard o de Schur de matrices, el cual será estudiado con más detalle en el Capítulo 8. Recordar (de la Sección 3.5) que, dadas $A, B \in \mathcal{M}_{n,m}(\mathbb{C})$, se define el producto de Hadamard $A \circ B$ como la matriz $A \circ B = \big(a_{ij}\ b_{ij}\big)_{\substack{i\in\mathbb{I}_n\\ j\in\mathbb{I}_m}} \in \mathcal{M}_{n,m}(\mathbb{C})$.

Definición 6.1.8. Sea $\mathbb{I}$ un intervalo abierto de $\mathbb{R}$ y $f : \mathbb{I} \to \mathbb{R}$ una función de clase C^1.

1. Denotaremos por $f^{[1]}$ a la función definida sobre $\mathbb{I} \times \mathbb{I}$ dada por

$$f^{[1]}(x,y) = \begin{cases} \dfrac{f(y)-f(x)}{y-x} & \text{si } x \neq y \\ \\ f'(x) & \text{si } x = y \end{cases} .$$

   A esta función se la denomina *primera diferencia dividida* de f.

2. Si $D = \operatorname{diag}(d_1, \dots, d_n) \in \mathcal{M}_n(\mathbb{C})$ es una matriz diagonal, llamaremos

$$f^{[1]}(D) \ = \ \big(f^{[1]}(d_i, d_j)\big)_{i,j\in\mathbb{I}_n} \in \mathcal{M}_n(\mathbb{C}) \ .$$

▲

Notationes: Recordemos que, dada $g : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ (U abierto), se dice que g es diferenciable en $x_0 \in U$ si existe una matriz $Dg_{x_0} \in \mathcal{M}_{mn}(\mathbb{C})$ (llamada derivada o diferencial de g en x_0 , y que debe tener las derivadas parciales de g en sus columnas) que cumple

$$\frac{\|g(x_0+h) - g(x_0) - Dg_{x_0} \cdot h\|}{\|h\|} \xrightarrow[h\to 0]{} 0 \ . \tag{6.2}$$

En tal caso se tiene que, para todo $h \in \mathbb{R}^n$, la derivada direccional

$$\frac{\partial}{\partial h}\, g(x_0) := \frac{d}{dt}\Big|_{t=0} g(x_0 + th) = Dg_{x_0} \cdot h \ .$$

Observar que si $\mathbb{I}$ es un intervalo abierto, entonces $\mathcal{H}_{\mathbb{I}}(n)$ es abierto en $\mathcal{H}(n)$, que es un $\mathbb{R}$-espacio vectorial que identificaremos con un $\mathbb{R}^M$. Luego podemos aplicar las nociones anteriores, pero reemplazando x_0 y h por matrices adecuadas.

Teorema 6.1.9. *Sean $\mathbb{I} \subseteq \mathbb{R}$ un intervalo abierto y $f : \mathbb{I} \to \mathbb{R}$ una función de clase C^1. Entonces su extensión $f : \mathcal{H}_{\mathbb{I}}(n) \to \mathcal{H}(n)$ es diferenciable en todo punto $A \in \mathcal{H}_{\mathbb{I}}(n)$. Si tomamos coordenadas en las que A sea diagonal, se tiene que*

$$Df_A(H) = f^{[1]}\big(A\big) \circ H \ , \quad \text{para todo } H \in \mathcal{H}(n) \ . \tag{6.3}$$

Es decir que dados $B \in \mathcal{H}_{\mathbb{I}}(n)$ y $U \in \mathcal{U}(n)$ tales que $A = UBU^*$ es diagonal, entonces

$$Df_B(H) = U^*\Big(f^{[1]}\big(A\big) \circ UHU^*\Big)U \ , \quad \text{para todo } H \in \mathcal{H}(n) \ , \tag{6.4}$$

Demostración. Mostremos el resultado, en principio, para funciones polinómicas. En este contexto, por linealidad podemos asumir que $f(x) = x^m$, para $m \in \mathbb{N} \cup \{0\}$. Observar que, en tal caso, $f^{[1]}\big(a, b\big) = \sum\limits_{k=1}^{m} a^{k-1}\, b^{m-k}$ para todo $a\,, b \in \mathbb{R}$ (incluso si $a = b$). Además,

$$Df_A(H) = \frac{d}{dt}\Big|_{t=0} f(A + tH) = \frac{d}{dt}\Big|_{t=0} (A + tH)^m = \sum_{k=1}^{m} A^{k-1} H A^{m-k} \ .$$

Si ahora usamos que $A = \operatorname{diag}(a_1\,, \dots, a_n)$, nos queda lo que queríamos:

$$Df_A(H) = \sum_{k=1}^{m} A^{k-1} H A^{m-k} = \Big[\sum_{k=1}^{m} a_i^{k-1}\ a_j^{m-k}\ H_{ij}\Big]_{i,j \in \mathbb{I}_n} = f^{[1]}\big(A\big) \circ H \ .$$

Luego, si f es un polinomio y $B \in \mathcal{H}_{\mathbb{I}}(n)$ no es diagonal, uno puede diagonalizar a B con una $U \in \mathcal{U}(n)$, derivar ahí y desdiagonalizar. Usando que $f(U(B + H)U^*) = Uf(B + H)U^*$ para todo $H \in \mathcal{H}(n)$ pequeño (para que $B + H \in \mathcal{H}_{\mathbb{I}}(n)$), no es difícil ver que

$$Df_B(H) = U^* \Big(f^{[1]}\big(UBU^*\big) \circ UHU^*\Big) U \ , \qquad \text{para todo } H \in \mathcal{H}(n) \ . \tag{6.5}$$

por el método directo de calcular el cociente incremental, como en la Eq. (6.2). En particular, el término de la derecha **no depende** de la U que diagonalice a B.

Sea ahora f una función de clase C^1 en $\mathbb{I}$. Usando el teorema de Weierstrass se puede construir una sucesión $(P_m)_{m \in \mathbb{N}}$ de polinomios que aproximan uniformemente tanto a f como a f' en cualquier subintervalo cerrado prefijado de $\mathbb{J}$. Es fácil ver que ${P_m}^{[1]}\big(A\big) \xrightarrow[m \to \infty]{} f^{[1]}\big(A\big)$. Fijemos ahora $H \in \mathcal{H}(n)$ pequeño, y $U \in \mathcal{U}(n)$ tal que $A = UBU^*$ es diagonal. Llamemos

$$\mathcal{D}f_B(H) = U^* \Big(f^{[1]}\big(A\big) \circ UHU^*\Big) U$$

(para ese $U \in \mathcal{U}(n)$ fijo), al cadidato a derivada. Hay que mostrar que el cociente incremental

$$\frac{\|f(B + H) - f(B) - \mathcal{D}f_B(H)\|_2}{\|H\|_2} \xrightarrow[H \to 0]{} 0 \ . \tag{6.6}$$

Esto probaría que f es diferenciable en B, que su derivada $Df_B(H) = \mathcal{D}f_B(H)$ (o sea que (6.4) es cierta), que su fórmula no depende del U elegido, y que se cumple la Eq. (6.3), para el caso en que B ya era diagonal (tomando $U = I$).

La prueba de (6.6) es una ardua acotación, de la que sólo mostraremos sus ideas principales. Se hace intercalando términos que involucran a los polinomios P_m . En efecto, si uno fija un $\varepsilon > 0$, encuentra un $m \in \mathbb{N}$ tal que tres cantidades:

- $\|f(B + H) - f(B) - (P_m(B + H) - P_m(B)\,)\|_2$,
- $\|P_m(B + H) - P_m(B) - \mathcal{D}(P_m)_B(H)\|_2$ y

- $\|\mathcal{D}f_B(H) - \mathcal{D}(P_m)_B(H)\|_2$

se pueden hacer menores que $\varepsilon\|H\|_2$, siempre que H sea chico. Luego uno se olvida del m y queda que el cociente de (6.6) es menor que 3ε para un tal H. Observar que la tercera vale a partir de un m para todo H. La primera se puede hacer válida para todos los m grandes (y para cualquier H tal que $B + H \in \mathcal{H}_{\mathbb{J}}(n)$), por un argumento que depende del teorema del valor medio y de la convergencia de las matrices ${P_m}^{[1]}\big(A\big)$ (más bien de que sean una sucesión de Cauchy). Finalmente, la segunda es la que pide H chico, tamaño que depende del m, pero este m se puede elegir de modo que se cumplan las otras dos. ∎

Corolario 6.1.10 (Dalekiĭi y Kreĭn)**.** Sean $\mathbb{I}, \mathbb{J} \subseteq \mathbb{R}$ dos intervalos abiertos y consideremos un curva de clase C^1 $\gamma : \mathbb{I} \to \mathcal{H}_{\mathbb{J}}(n)$. Sea $f : \mathbb{J} \to \mathbb{R}$ otra función de clase C^1. Entonces

1. La curva que llamaremos $f \bullet \gamma : \mathbb{I} \to \mathcal{H}(n)$ dada por $f \bullet \gamma(t) = f\big(\gamma(t)\big)$, vía el cálculo funcional, es también de clase C^1.

2. Supongamos que $\gamma(t_0) = \operatorname{diag}(a_1, \ldots, a_n)$ para cierto $t_0 \in \mathbb{I}$. Entonces se cumple la siguiente fórmula:
$$(f \bullet \gamma)'(t_0) = f^{[1]}\big(\gamma(t_0)\big) \circ \gamma'(t_0) \ . \tag{6.7}$$

Demostración. La suavidad de $f \bullet \gamma$ se deduce de la diferenciablidad de $f : \mathcal{H}_{\mathbb{J}}(n) \to \mathcal{H}(n)$ (y de la la suavidad de γ). La regla de la cadena y la fórmula (6.7) se deducen también del Teorema 6.1.9, puesto que $(f \bullet \gamma)'(t_0) = Df_{\gamma(t_0)}\big(\gamma'(t_0)\big) = f^{[1]}\big(\gamma(t_0)\big) \circ \gamma'(t_0)$. ∎

6.2 Funciones monótonas de operadores

Definición 6.2.1. Sea $\mathbb{I} \subseteq \mathbb{R}$ un intervalo y $f : \mathbb{I} \to \mathbb{R}$, una función. Diremos que f es *monótona de operadores* (MOP) si, para todo $n \in \mathbb{N}$ y $A, B \in \mathcal{H}_{\mathbb{I}}(n)$, se tiene que

$$A \leq B \quad \Longrightarrow f(A) \leq f(B) \ .$$

Notar que, tomando $n = 1$, se ve que f debe ser monótona en el sentido usual. ▲

Ejemplos 6.2.2.

1. Dados $a, b \in \mathbb{R}$, la funcion $f(t) = a + bt$ es MOP si y sólo si $b \geq 0$.

2. $f(t) = t^2$ no es monótona de operadores (en ningún intervalo $\mathbb{I} \subseteq [0, +\infty)$ con más de un punto). En efecto, tomando $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, se ve que $A \leq B$, pero
$$A^2 = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \not\leq \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} = B^2 \ .$$
El ejemplo se puede cambiar, de acuerdo al intervalo $\mathbb{I}$, tomando $C = aI + \varepsilon A$ y $D = aI + \varepsilon B$, para constantes $a \in \mathbb{I}^\circ$ y $\varepsilon > 0$ convenientes. Notar que las entradas $2, 2$ de C^2 y D^2 siguen coincidiendo.

3. $f(t) = -t^{-1}$ es MOP en $\mathbb{I} = (0, +\infty)$. En efecto, si $0 < A \le B \in \mathcal{M}_n(\mathbb{C})^+$, entonces $0 < B^{-1/2}AB^{-1/2} \le I$. Luego $\mu_1(B^{-1/2}AB^{-1/2}) \le 1$, o sea
$$\mu_n((B^{-1/2}AB^{-1/2})^{-1}) \ge 1 \implies (B^{-1/2}AB^{-1/2})^{-1} = B^{1/2}A^{-1}B^{1/2} \ge I \ ,$$
por lo que $A^{-1} \ge B^{-1}$. ▲

Ejercicio 6.2.3. Probar que

1. La suma y la composición (cuando tenga sentido) de MOP´s es MOP.
2. Dada una matriz $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{M}_2(\mathbb{R})$, con $d \ne 0$, definamos la función
$$f_M(t) = \frac{a + bt}{c + dt} \ , \quad t \ne \frac{-c}{d} \ .$$
Entonces f_M es MOP en $(\frac{-c}{d}, +\infty)$ si y sólo si $\det M \le 0$. Notar que
$$f_M(t) = \frac{b}{d} \ + \ \frac{\det M}{cd + d^2\, t} \ .$$
Por lo tanto, si $\det M < 0$, entonces f_M es composición de MOP´s. Pero si f_M fuera MOP y $\det M > 0$, podría deducirse que $t \mapsto 1/t$ es MOP. ▲

Proposición 6.2.4. *La función $f(t) = t^{1/2}$ es MOP en $\mathbb{I} = [0, +\infty)$.*

Demostración. Sean $A, B \in \mathcal{M}_n(\mathbb{C})^+$ tales que $A \le B$. Supongamos, en principio, que $B > 0$. Entonces, por la Proposición 3.5.4,
$$1 \ge \|A^{1/2}B^{-1/2}\|_{sp} \ge \rho(A^{1/2}B^{-1/2}) = \rho(B^{-1/4}A^{1/2}B^{-1/4}) \implies I \ge B^{-1/4}A^{1/2}B^{-1/4} \ .$$
Por lo tanto $B^{1/2} \ge A^{1/2}$. Si B no es inversible, para cada $\varepsilon > 0$ se toma la matriz $B + \varepsilon I > 0$. Luego $A^{1/2} \le (B + \varepsilon I)^{1/2}$ para todo $\varepsilon > 0$. Como $(t + \varepsilon)^{1/2} \xrightarrow[\varepsilon \to 0]{\text{puntualmente}} t^{1/2} = f(t)$,
$$\langle A^{1/2}x, x\rangle \ \le \ \Big\langle (B + \varepsilon I)^{1/2}x, x\Big\rangle \xrightarrow[\varepsilon \to 0]{} \langle B^{1/2}x, x\rangle \ , \quad \text{para todo} \quad x \in \mathbb{C}^n \ .$$
Deducimos que $A^{1/2} \le B^{1/2}$. ■

Ejercicio 6.2.5. Rellenar los detalles de la siguiente prueba alternativa de la Proposición 6.2.4, que se basa en un resultado del Capítulo 9: Suponemos que $0 < A < B$. Entonces definimos la función
$$C : [0, 1] \to \mathcal{G}l(n)^+ \ , \quad \text{dada por} \quad C(t) = A + t(B - A) \ , \quad t \in [0, 1] \ .$$
Sea $R(t) = C(t)^{1/2}$, $t \in [0, 1]$. Entonces
$$R(t)^2 = C(t) \quad \implies \quad \dot{R}(t)R(t) + R(t)\dot{R}(t) = \dot{C}(t) = B - A > 0 \ , \quad t \in [0, 1] \ ,$$
donde el *punto* denota derivada respecto de t. Por la Observación 9.1.5, como $R(t) > 0$ y $\dot{C}(t) = S(R, \dot{R}) > 0$, entonces, $\dot{R}(t) > 0$ para todo $t \in [0, 1]$. Luego R es *creciente* y, en particular, $A^{1/2} = R(0) < R(1) = B^{1/2}$. ▲

Teorema 6.2.6. *Las funciones $f(t) = t^r$ son MOP's en $\mathbb{I} = [0, +\infty)$, para todo $r \in [0, 1]$. En otras palabras, si $0 \le A \le B \in \mathcal{M}_n(\mathbb{C})^+$, entonces $A^r \le B^r$ para todo $0 \le r \le 1$.*

Demostración. Supongamos, en principio, que $0 < A \le B \in \mathcal{G}l(n)^+$ y que r es diádico, es decir que $r = k/2^m$, para $k \in \mathbb{I}_{2^m}$. En este caso probaremos, por inducción en m, que $A^r \le B^r$. En efecto, si $m = 1$, ya lo sabemos por la Proposición 6.2.4.

Si suponemos el hecho cierto para todo número $j/2^m$, tomemos $r = k/2^{m+1}$. Si $k \le 2^m$, entonces $k/2^m \le 1$. Por la hipótesis inductiva y la Proposición 6.2.4, se tiene que

$$A^{k/2^m} \le B^{k/2^m} \quad \Longrightarrow \quad A^r = (A^{k/2^m})^{1/2} \le (B^{k/2^m})^{1/2} = B^r \ .$$

Si $k > 2^m$, usaremos que $B^{-1} \le A^{-1}$. Por tanto, $B^r A^{-1} B^r \ge B^r B^{-1} B^r = B^{2r-1}$. Luego, como $0 < 2r - 1 = \dfrac{k - 2^m}{2^m} \le 1$, por la hipótesis inductiva tenemos que

$$\begin{aligned}(A^{-1/2} B^r A^{-1/2})^2 = A^{-1/2} B^r A^{-1} B^r A^{-1/2} \quad &\ge A^{-1/2} B^{2r-1} A^{-1/2} \\ &\ge A^{-1/2} A^{2r-1} A^{-1/2} = A^{2(r-1)} \ .\end{aligned}$$

Aplicando la Proposición 6.2.4, deducimos que $A^{-1/2} B^r A^{-1/2} \ge A^{r-1}$, y por ello $B^r \ge A^r$.

Si r no es diádico, tomemos una sucesión de diádicos $r_m \xrightarrow[m\to\infty]{} r$. Como las funciones $f_m(t) = t^{r_m} \xrightarrow[m\to\infty]{\text{puntualmente}} f(t) = t^r$, deducimos que $B^r \ge A^r$, también en este caso.

Finalmente, si $A \not> 0$, como $(A + \frac{1}{m} I)^r \le (B + \frac{1}{m} I)^r$ para todo $m \in \mathbb{N}$ y la función $t \mapsto t^r$ es continua, aplicando la Proposición 6.1.5 obtenemos que

$$A^r = \lim_{m\to\infty} \left(A + \frac{1}{m}\ I\right)^r \le \lim_{m\to\infty} \left(B + \frac{1}{m}\ I\right)^r = B^r \ ,$$

lo que prueba la desigualdad en el caso general. ∎

Lema 6.2.7. *Sea $A \in \mathcal{G}l(n)^+$. Entonces*

$$\lim_{h\to 0} \frac{A^h - I}{h} = \log A \ .$$

Demostración. Observar que, para todo $t \in (0, +\infty)$, se verifica que

$$\lim_{h\to 0} \frac{t^h - 1}{h} = \lim_{h\to 0} \frac{e^{h\ \log t} - 1}{h} = \log t \ .$$

Por lo tanto las funciones $f_h(t) = \dfrac{t^h - 1}{h} \xrightarrow[h\to 0]{\text{puntualmente}} g(t) = \log\ t$ en todo $(0, +\infty)$. Aplicando el item 8 del Ejercicio 6.1.3, se culmina la prueba. ∎

Proposición 6.2.8. *La función $f(t) = \log t$ es MOP en $\mathbb{I} = (0, +\infty)$. En otras palabras, dados $A \le B$ ambos en $\mathcal{G}l(n)^+$, se tiene que $\log A \le \log B$.*

Demostración. Se deduce del Lema 6.2.7. En efecto, tomando h con valores en $(0,1)$, por el Teorema 6.2.6 se tiene que

$$\log A = \lim_{h\to 0^+} \frac{A^h - I}{h} \leq \lim_{h\to 0^+} \frac{B^h - I}{h} = \log B \ . \qquad \blacksquare$$

Para finalizar daremos una caracterización de las MOPs en terminos de la primera diferencia dividida de f, la cual puede interpretarse como análogo matricial al resultado clásico de cálculo que dice que una función de una variable real derivable es no-decreciente si y sólo si su derivada es no-negativa.

Teorema 6.2.9. *Sea $\mathbb{I}$ un intervalo abierto de $\mathbb{R}$, $f : \mathbb{I} \to \mathbb{R}$ una función de clase C^1. Entonces, las siguientes afirmaciones son equivalente:*

1. *f es MOP.*
2. *Para todo $n \in \mathbb{N}$ y toda matriz diagonal $D \in \mathcal{H}_{\mathbb{I}}(n)$, se tiene que $f^{[1]}(D) \in \mathcal{M}_n(\mathbb{C})^+$.*

Demostración. 1 $\to$ 2. Sea $D = \operatorname{diag}(d_1, \dots, d_n) \in \mathcal{H}_{\mathbb{I}}(n)$. Recordemos que por medio de $A \circ B$ denotamos al producto de Hadamard, i.e., el producto entrada a entrada. Sea $E_n = \mathbb{1}_n \mathbb{1}_n^* \in \mathcal{M}_n(\mathbb{C})^+$. Respecto a este producto, $\mathbb{E}_n$ se comporta como la identidad. Sea $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{H}(n)$ dada por $\gamma(t) = D + t\,\mathbb{E}_n$. Observar que para todo $t \geq 0$ se tiene que $D + t\,\mathbb{E}_n \geq D$. Por la Proposición 6.1.5, $\sigma(D + t\,\mathbb{E}_n) \subseteq \mathbb{I}$ para valores pequeños de t. Luego, para dichos valores de t, tiene sentido hacer $f \circ \gamma$. Más aún, como f es de clase C^1 y γ es suave, podemos derivar la curva $f \circ \gamma$, y por el Teorema 6.1.9 obtenemos

$$\frac{d}{dt} f(D + t\,\mathbb{E}_n)\bigg|_{t=0} = Df_D(\mathbb{E}_n) = f^{[1]}(D) \circ \gamma'(0) = f^{[1]}(D) \circ \mathbb{E}_n = f^{[1]}(D) \, .$$

Usando que $\mathbb{E}_n \in \mathcal{M}_n(\mathbb{C})^+$ y que f es MOP, se tiene que el cociente incremental

$$\frac{f(D + t\,\mathbb{E}_n) - f(D)}{t} \in \mathcal{M}_n(\mathbb{C})^+ \ , \qquad \text{para todo } t \in (-\varepsilon, \varepsilon) \ ,$$

lo cual se preserva al tomar límite. Por ende, $f^{[1]}(D) \in \mathcal{M}_n(\mathbb{C})^+$.

2 $\to$ 1. Sean $A, B \in \mathcal{H}_{\mathbb{I}}(n)$ tales que $A \leq B$, y definamos la curva $\gamma(t) = (1-t)A + tB$, para $t \in [0,1]$. Como $\mathcal{H}_{\mathbb{I}}(n)$ es convexo (Proposición 6.1.5), $\gamma(t) \in \mathcal{H}_{\mathbb{I}}(n)$ para todo $t \in [0,1]$. Luego, la nueva curva $\eta(t) = f(\gamma(t)\,)$ está bien definida. El primer paso será probar que para todo $t \in (0,1)$ se tiene que $\eta'(t) \in \mathcal{M}_n(\mathbb{C})^+$. Para ello, fijemos un $t \in (0,1)$ cualquiera. Sin pérdida de generalidad podemos suponer que $\gamma(t)$ es diagonal (sino se conjuga con una unitaria). Luego, por el Corolario 6.1.10 se tiene que

$$\eta'(t) = f^{[1]}(\gamma(t)\,) \circ \gamma'(t) = f^{[1]}(\gamma(t)\,) \circ (B - A) \in \mathcal{M}_n(\mathbb{C})^+ \ ,$$

donde usamos que $A \leq B$, que $f^{[1]}(\gamma(t)\,) \in \mathcal{M}_n(\mathbb{C})^+$ y luego el Teorema 2 de Schur 3.6.2 (producto $\circ$ de positivas es positiva). Ahora bien, fijemos $x \in \mathbb{C}^n$. Por la linealidad de la función $A \mapsto \langle Ax, x\rangle$ se tiene que la función $g(t) = \langle \eta(t)x, x\rangle$ es continua en el $[0,1]$, derivable en $(0,1)$. Más aún, $g'(t) = \langle \eta'(t)x, \, x\rangle$ para todo $t \in (0,1)$. Pero entonces, por lo que acabamos de ver, g es creciente en el $[0,1]$. En consecuencia

$$\langle f(A)x, x\rangle = g(0) \le g(1) = \langle f(B)x, \, x\rangle \, .$$

Como $x \in \mathbb{C}^n$ es arbitrario, $f(A) \le f(B)$, lo cual concluye la demostración. ■

Para ilustrar como se utiliza esta caracterización para demostrar que una función es monótona de operadores, probaremos que la función $f(t) = \tan(t)$ es MOP en el itervalo $(-\pi, \pi)$. Para ello, necesitaremos un par de lemas previos.

Lema 6.2.10. *Sea $d = (d_1, \dots, d_n) \in \mathbb{R}^n$. Entonces la matricz*

$$K_e(d) \;=\; \left(\, e^{i(d_j - d_i)} \, \right)_{i,j \in \mathbb{I}_n} \;\in \mathcal{M}_n(\mathbb{C})^+ \;.$$

Demostración. En efecto, si tomamos $\mathbb{E} = \mathbb{1}\,\mathbb{1}^* \in \mathcal{M}_n(\mathbb{C})^+$ (la matriz de cuyas entradas son todas iguales a 1) y $U = \operatorname{diag}\left(e^{id_1}, \dots, e^{id_n}\right)$, entonces $K_e(d) = U^*\mathbb{E}U \in \mathcal{M}_n(\mathbb{C})^+$. ■

Lema 6.2.11. *Sea $d = (d_1, \dots, d_n) \in \mathbb{R}^n$. Entonces la matriz*

$$K_s(d) \;=\; \left(\frac{\operatorname{sen}(d_j - d_i)}{d_j - d_i} \right)_{i,j \in \mathbb{I}_n} \;\in \mathcal{M}_n(\mathbb{C})^+ \;,$$

donde, para abreviar notaciones, estamos aceptando la convención de que $\frac{\operatorname{sen} 0}{0} = 1$.

Demostración. Este lema se deduce del anterior si recordamos la siguiente identidad, la cual puede testearse a mano muy facilmente:

$$\frac{\operatorname{sen} a}{a} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{iat} \, dt \quad \text{ para todo } a \in \mathbb{R} \;\; (\text{incluso si } a = 0) \; . \tag{6.8}$$

En efecto, dado $x = (x_1, \dots, x_n) \in \mathbb{C}^n$ se tiene que $\langle K_s(d)\, x, \, x\rangle = \displaystyle\sum_{i,j=1}^{n} \frac{\operatorname{sen}(d_j - d_i)}{d_j - d_i} \, \overline{x}_i \, x_j$.

Pero por la formula integral (6.8) y el lema anterior se tiene que

$$2\pi \sum_{i,j=1}^{n} \frac{\operatorname{sen}(d_j - d_i)}{d_j - d_i} \, \overline{x}_i \, x_j = \int_{-\pi}^{\pi} \sum_{i,j=1}^{n} e^{i(d_j - d_i)t} \, \overline{x}_i \, x_j \, dt = \int_{-\pi}^{\pi} \langle K_e(d)\, x, \, x\rangle \; dt \ge 0 \; .$$ ■

Proposición 6.2.12. *La función $f(t) = \tan(t)$ es MOP en $\mathbb{I} = (-\pi/2 \, , \pi/2)$.*

Demostración. Sea $D = \operatorname{diag}(d_1, \dots, d_n) \in \mathcal{H}_{\mathbb{I}}(n)$ y la matriz de diferencias divididas

$$\tan^{[1]}(D)_{ij} = \begin{cases} \dfrac{\tan(d_j) - \tan(d_i)}{d_j - d_i} & \text{si } d_i \neq d_j \\ \sec^2(d_i) & \text{si } d_i = d_j \end{cases} .$$

Usando la identidad $\tan(x) - \tan(y) = \dfrac{\operatorname{sen}(x - y)}{\cos(x)\cos(y)}$ se tiene que

$$\tan^{[1]}(D)_{ij} = \left[\frac{1}{\cos(d_i)} \left(\frac{\operatorname{sen}(d_j - d_i)}{d_j - d_i} \right) \frac{1}{\cos(d_j)} \right]_{i,j \in \mathbb{I}_n} = \sec(D)\, K_s(d) \sec(D)^* \in \mathcal{M}_n(\mathbb{C})^+ \; ,$$

por el Lema 6.2.11. Con esta información, el Teorema 6.2.9 garantiza que $f(t) = \tan(t)$ es MOP en $\mathbb{I} = (-\pi, \pi)$. ■

6.3 Funciones convexas de operadores

Recordemos que la Proposición 6.1.5 asegura que, si $\mathbb{I} \subseteq \mathbb{R}$ es un intervalo, $\mathcal{H}_{\mathbb{I}}(n)$ es comvexo.

Definición 6.3.1. Sea $\mathbb{I} \subseteq \mathbb{R}$ un intervalo y $f : \mathbb{I} \to \mathbb{R}$, una función. Diremos que f es *convexa de operadores* ($\cup$OP) si, para todo $n \in \mathbb{N}$, $\lambda \in [0,1]$ y $A, B \in \mathcal{H}_{\mathbb{I}}(n)$, se tiene

$$f\big(\lambda A + (1-\lambda)B\big) \quad \leq \quad \lambda f(A) + (1-\lambda) f(B) \ . \tag{6.9}$$

Notar que, tomando $n = 1$, se ve que f debe ser convexa en el sentido usual. Diremos que f es cóncava de operadores ($\cap$OP) si $-f$ es $\cup$OP. ▲

Observación 6.3.2. Si $f : \mathbb{I} \to \mathbb{R}$ es **continua**, para verificar que es convexa de operadores, es suficiente probar que

$$f\Big(\frac{A+B}{2}\Big) \quad \leq \quad \frac{f(A)+f(B)}{2} \ ,$$

para todo par $A, B \in \mathcal{H}_{\mathbb{I}}(n)$ (y todo $n \in \mathbb{N}$). En efecto, esta condicion implica que f cumple la Eq. (6.9) para todo λ diádico en $[0,1]$. Esto se prueba por inducción. Por ejemplo,

$$f\Big(\frac{1}{4}\ A + \frac{3}{4}\ B \Big) = f\Big(\frac{\frac{A+B}{2} + B}{2} \Big) \leq \frac{f\big(\frac{A+B}{2}\big) + f(B)}{2}$$

$$\leq \frac{\frac{f(A)+f(B)}{2} + f(B)}{2} = \frac{1}{4}\ f(A) + \frac{3}{4}\ f(B) \ .$$

Como f es continua, la Proposición 6.1.5 dice que (6.9) se cumple para todo $\lambda \in [0,1]$. ▲

Ejemplos 6.3.3.

1. Dados $a, b \in \mathbb{R}$, se tiene que la funcion $f(t) = a + bt$ es $\cup$OP (y $\cap$OP).
2. $f(t) = t^2$ **sí** es $\cup$OP en $[0, +\infty)$. En efecto, dados $A, B \in \mathcal{M}_n(\mathbb{C})^+$, se tiene que

$$\frac{A^2+B^2}{2} - \Big(\frac{A+B}{2}\Big)^2 = \frac{1}{4}\big(A^2 + B^2 - AB - BA\big) = \frac{1}{4}(A-B)^2 \ .$$

   Como f es continua, esto prueba que es $\cup$OP.
3. $f(t) = t^3$ **no** es $\cup$OP en $[0, +\infty)$. En efecto, una cuenta elemental muestra que, si

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{entonces} \quad \frac{A^3+B^3}{2} - \Big(\frac{A+B}{2}\Big)^3 = \begin{bmatrix} 6 & 1 \\ 1 & 0 \end{bmatrix} ,$$

   que no es positiva. Tampoco puede ser $\cup$OP en ningún intervalo $\mathbb{I} \subseteq [0, +\infty)$.

4. $f(t) = t^{-1}$ es $\cup$OP en $\mathbb{I} = (0, +\infty)$. En efecto, si $A, B \in \mathcal{G}l(n)^+$, entonces
$$\frac{A^{-1}+B^{-1}}{2} - \Big(\frac{A+B}{2}\Big)^{-1} = \frac{(A^{-1}-B^{-1})(A^{-1}+B^{-1})^{-1}(A^{-1}-B^{-1})}{2} \in \mathcal{M}_n(\mathbb{C})^+ \ .$$
En efecto, esto se deduce de la identidad
$$2(A+B)^{-1} = A^{-1}(A^{-1}+B^{-1})^{-1}B^{-1} + B^{-1}(A^{-1}+B^{-1})^{-1}A^{-1} \ .$$
Como f es continua, lo que vimos muestra que es $\cup$OP. ▲

Ejercicio 6.3.4. Probar que

1. La suma y la composición (cuando tenga sentido) de $\cup$OP´s es $\cup$OP.
2. Dada una matriz $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{M}_2(\mathbb{R})$, con $d \neq 0$, definamos la función
$$f_M(t) = \frac{a+bt}{c+dt} \ , \quad t \neq \frac{-c}{d} \ .$$
Entonces f_M es $\cup$OP en $\mathbb{I} = (\frac{-c}{d}, +\infty)$ si y sólo si $\det M \geq 0$. Por otra parte, f es $\cap$OP en $\mathbb{I}$ si y sólo si f es MOP en $\mathbb{I}$ si y sólo si $\det M \leq 0$.
3. Sea $\mathbb{I} \subseteq \mathbb{R}$ un intervalo tal que $0 \in \mathbb{I}$. Sean $f(t) = |t|$ y $g(t) = t \vee 0$, $t \in \mathbb{I}$. Entonces f no es $\cup$OP y g no es $\cup$OP ni MOP. ▲

Definición 6.3.5. Sean $A \in \mathcal{M}_n(\mathbb{C})$ y $P \in \mathcal{H}(n)$ un proyector con $R(P) = \mathcal{S}$. Llamaremos *compresión* de A a $\mathcal{S}$, al operador
$$A_{\mathcal{S}} : \mathcal{S} \to \mathcal{S} \quad \text{dado por} \quad A_{\mathcal{S}}(x) = PA\,x \ , \quad x \in \mathcal{S} \ .$$
Notar que $A_{\mathcal{S}} = PAP\big|_{\mathcal{S}}$ pensado en $L(\mathcal{S})$. En coordenadas de una BON de $\mathbb{C}^n$ tal que la matriz de P en esa base sea $P = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$, se tiene que
$$PAP = \begin{bmatrix} A_{\mathcal{S}} & 0 \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad C_P(A) = \begin{bmatrix} A_{\mathcal{S}} & 0 \\ 0 & A_{\mathcal{S}^\perp} \end{bmatrix} \ ,$$
donde las matrices (grandes) viven, ahora, en $\mathcal{M}_n(\mathbb{C})$. ▲

Recordemos que, para cada $k \in \mathbb{I}_n$, 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 $\mathbb{C}^n$.

Teorema 6.3.6. *Sea $\mathbb{I} \subseteq \mathbb{R}$ un intervalo y $f : \mathbb{I} \to \mathbb{R}$, una función. Son equivalentes:*

1. *f es convexa de operadores.*

2. *Para todo $n \in \mathbb{N}$ y para todo sistema de proyectores $\mathcal{P}$ en $\mathcal{H}(n)$, se tiene que*
$$f\big(C_{\mathcal{P}}(A)\big) \le C_{\mathcal{P}}(f(A)) \quad \textit{para todo } A \in \mathcal{H}_{\mathbb{I}}(n) \ .$$
3. *Dados $n \in \mathbb{N}$, $A \in \mathcal{H}_{\mathbb{I}}(n)$ y $\mathcal{S} \subseteq \mathbb{C}^n$ un subespacio, se tiene que $f(A_{\mathcal{S}}) \le f(A)_{\mathcal{S}}$.*
4. *Dados $k, n \in \mathbb{N}$ tales que $k \le n$, $A \in \mathcal{H}_{\mathbb{I}}(n)$ y $V \in \mathcal{U}_k(n)$, se verifica que*
$$f(V^*AV) \le V^*f(A)V \ .$$

Demostración. Antes que nada, observar que $C_{\mathcal{P}}(A) \subseteq \mathcal{H}_{\mathbb{I}}(n)$ por la Proposición 5.4.4 (o la Eq. (6.10) de abajo) y el hecho de que $\mathcal{H}_{\mathbb{I}}(n)$ es convexo. De ahi se deduce que, si $\dim \mathcal{S} = k$, entonces $A_{\mathcal{S}} \in \mathcal{H}_{\mathbb{I}}(k)$ y tiene sentido calcular $f(A_{\mathcal{S}})$, incluso si $0 \notin \mathbb{I}$.

$1 \to 2$. Como otras veces, por la Eq. (5.8), podemos suponer (s.p.g.) que trabajamos con un solo proyector $P \in \mathcal{H}(n)$. Observar que, dado $A \in \mathcal{H}_{\mathbb{I}}(n)$, se tiene que
$$C_P(A) = \frac{A + UAU^*}{2} \ , \quad \text{con} \quad U = 2P - I \in \mathcal{U}(n) \ . \tag{6.10}$$

Por lo tanto, si asumimos que f es $\cup$OP, entonces
$$f\big(C_P(A)\big) \le \frac{f(A) + f(UAU^*)}{2} = \frac{f(A) + Uf(A)U^*}{2} = C_P\big(f(A)\big) \ .$$

$2 \to 3$. Basta mirar los bloques $1,1$ de la desigualdad $f\big(C_{P_{\mathcal{S}}}(A)\big) \le C_{P_{\mathcal{S}}}\big(f(A)\big)$.

$3 \to 4$. Llamemos $\mathcal{S} = R(V) \subseteq \mathbb{C}^n$ y $P = P_{\mathcal{S}}$. Entonces se tiene que $V^*AV = V^*(PAP)V$. Por lo tanto, si denotamos $V_0 : \mathbb{C}^k \to \mathcal{S}$, al mismo V correstringido a su imagen, tenemos que V_0 es unitario y que $V^*AV = V_0^*A_{\mathcal{S}}V_0 \in \mathcal{H}_{\mathbb{I}}(k)$. Además
$$f(V^*AV) = f(V_0^*A_{\mathcal{S}}V_0) = V_0^*f(A_{\mathcal{S}})V_0 \quad \text{y} \quad V^*f(A)V = V_0^*f(A)_{\mathcal{S}}V_0 \ ,$$
por lo que
$$f(V^*AV) \le V^*f(A)V \iff f(A_{\mathcal{S}}) \le f(A)_{\mathcal{S}} \ .$$

$4 \to 1$. Dados $A, B \in \mathcal{H}_{\mathbb{I}}(n)$, consideremos el operador $T = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \in \mathcal{H}_{\mathbb{I}}(2n)$. Dado $\lambda \in [0,1]$, sean $\mu = 1 - \lambda$ y $V = \begin{bmatrix} \lambda^{\frac12} I_n \\ \mu^{\frac12} I_n \end{bmatrix} \in \mathcal{U}_n(2n)$. Una cuenta directa pueba que
$$V^*TV = \lambda A + \mu B \quad \text{y que} \quad V^*f(T)V = \lambda f(A) + \mu f(B) \ ,$$
usando que $f(T) = \begin{bmatrix} f(A) & 0 \\ 0 & f(B) \end{bmatrix}$. Por todo ello, si f cumple 4, se tiene que
$$f\big(\lambda A + \mu B\big) = f(V^*TV) \le V^*f(T)V = \lambda f(A) + \mu f(B). \qquad \blacksquare$$

Corolario 6.3.7. *Sean $A \in \mathcal{G}l(n)^+$ y $P \in \mathcal{H}(n)$, un proyector. Entonces*

1. *$C_P(A) \in \mathcal{G}l(n)^+$ y $C_P(A)^{-1} \leq C_P(A^{-1})$.*

2. *Si $\mathcal{S} = R(P)$ entonces, $A_{\mathcal{S}}{}^{-1} \leq (A^{-1})_{\mathcal{S}}$. Es decir que, pensados como operadores en $L(\mathcal{S})$, se tiene que*
$$(PAP)^{-1} \leq PA^{-1}P. \tag{6.11}$$

Demostración. Se deduce de que $\mathcal{G}l(n)^+ = \mathcal{H}_{(0,+\infty)}(n)$ y del Teorema 6.3.6, dado que $t \mapsto t^{-1}$ es $\cup$OP en $(0, +\infty)$, como se ha observado en el Ejemplo 6.3.3.4. ■

Observación 6.3.8. Una versión más detallada de la desigualdad (6.11) se deduce de la teoría de complementos de Schur. En la Proposición 3.8.7 vimos que, si $\mathcal{S} = R(P)$,

$$A = \begin{bmatrix} a & b \\ b^* & c \end{bmatrix} \begin{matrix} \mathcal{S} \\ \mathcal{S}^\perp \end{matrix} \ , \implies (A^{-1})_{\mathcal{S}}{}^{-1} = (PA^{-1}P)^{-1} = a - bc^{-1}b^* \ .$$

En particular, también así se muestra que $(PA^{-1}P)^{-1} \leq a = A_{\mathcal{S}}$. ▲

Proposición 6.3.9. *Sea $\mathbb{I} \subseteq \mathbb{R}$ un intervalo tal que $0 \in \mathbb{I}$ y sea $f : \mathbb{I} \to \mathbb{R}$, una función. Entonces son equivalentes:*

1. *f es convexa de operadores y $f(0) \leq 0$.*

2. *Dados $n \in \mathbb{N}$, $A \in \mathcal{H}_{\mathbb{I}}(n)$ y $P \in \mathcal{H}(n)$, un proyector, se cumple que*
$$f(PAP) \leq Pf(A)P \ ,$$
pensados como matrices en $\mathcal{M}_n(\mathbb{C})$.

Demostración. Como $0 \in \mathbb{I}$, es fácil ver que $PAP \in \mathcal{H}_{\mathbb{I}}(n)$. Sea $\mathcal{S} = R(P)$. Entonces, en coordenadas de una BON de $\mathbb{C}^n$ tal que $P = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$, se tiene que

$$PAP = \begin{bmatrix} A_{\mathcal{S}} & 0 \\ 0 & 0 \end{bmatrix} \implies f(PAP) = \begin{bmatrix} f(A_{\mathcal{S}}) & 0 \\ 0 & f(0)I_{\mathcal{S}^\perp} \end{bmatrix} \ .$$

Por otra parte, en la misma BON,

$$Pf(A)P = \begin{bmatrix} f(A)_{\mathcal{S}} & 0 \\ 0 & 0 \end{bmatrix} \ .$$

Por lo tanto, las condiciones 1 y 2 son equivalentes, por serlo 1 y 3 del Teorema 6.3.6 (es decir, que $f(A_{\mathcal{S}}) \leq f(A)_{\mathcal{S}}$ para todo el mundo, equivale a que f sea $\cup$OP). ■

Ejercicio 6.3.10. Sea $\mathbb{I} \subseteq \mathbb{R}$ un intervalo tal que $0 \in \mathbb{I}$ y sea $f : \mathbb{I} \to \mathbb{R}$, una función. Entonces son equivalentes:

1. f es convexa de operadores y $f(0) \leq 0$.

2. Dados $n \in \mathbb{N}$, $A \in \mathcal{H}_{\mathbb{I}}(n)$, se cumple que
$$f(C^*AC) \le C^* f(A) C.$$
para todo $C \in \mathcal{M}_n(\mathbb{C})$ tal que $\|C\| \le 1$.

3. Dados $n \in \mathbb{N}$, $A, B \in \mathcal{H}_{\mathbb{I}}(n)$ y $C, D \in \mathcal{M}_n(\mathbb{C})$ tales que $C^*C + D^*D \le I$, se verifica
$$f(C^*AC) + f(D^*BD) \le C^* f(A) C + D^* f(B) D.$$

Se sugiere usar las matrices unitarias construidas en 3.7.9. ▲

Corolario 6.3.11. *La función $f(t) = t^r$ es $\cup OP$ en $\mathbb{I} = [0, +\infty)$, para todo $r \in [1, 2]$.*

Demostración. Como $f(0) = 0$, por la Proposición 6.3.9, bastaría probar que
$$(PAP)^r \le PA^rP \ , \quad \forall \ \ n \in \mathbb{N} \, , \ A \in \mathcal{M}_n(\mathbb{C})^+ \ \text{ y } \ P \ \text{ un proyector en } \ \mathcal{H}(n) \ .$$
En efecto, como $0 \le P \le I$, tenemos que $0 \le A^{1/2}PA^{1/2} \le A$. Sea $g(t) = t^{r-1}$. Por el Teorema 6.2.6, g es MOP (porque $0 \le r - 1 \le 1$). Luego $(A^{1/2}PA^{1/2})^{r-1} \le A^{r-1}$, por lo que
$$PA^{1/2}\big(A^{1/2}PA^{1/2}\big)^{r-1}A^{1/2}P \le PA^{1/2}A^{r-1}A^{1/2}P = PA^rP \ .$$
Finalmente, observar que para todo $k \in \mathbb{N} \cup \{0\}$ se tiene que
$$PA^{1/2}\big(A^{1/2}PA^{1/2}\big)^{k}A^{1/2}P = (PAP)^{k+1} \ .$$
Por lo tanto, para todo polinomio $Q \in \mathbb{C}[x]$ vale la igualdad
$$PA^{1/2}Q\big(A^{1/2}PA^{1/2}\big)A^{1/2}P = Q(PAP) \ PAP \ . \, ,$$
De ahí deducimos que una igualdad similar valdrá para toda $f : [0, +\infty) \to \mathbb{R}$ (eligiendo un $Q \in \mathbb{C}[x]$ que coincida con f en $\sigma\left(A^{1/2}PA^{1/2}\right) = \sigma\left(PAP\right)$). En particular,
$$PA^rP \ge PA^{1/2}\big(A^{1/2}PA^{1/2}\big)^{r-1}A^{1/2}P = (PAP)^r \ . \qquad ■$$

Proposición 6.3.12. *Sea $f : [0, +\infty) \to [0, +\infty)$ una función* **continua**. *Entonces se tiene que f es MOP si y sólo si f es $\cap OP$ (i.e., $-f$ es $\cup OP$).*

Demostración. Supongamos que f es MOP, y sea $\mathcal{S} \subseteq \mathbb{C}^n$ un subespacio. Notemos $P = P_{\mathcal{S}}$. Como $f(0) \ge 0$ por hipótesis, usando la Proposición 6.3.9 bastaría probar que
$$Pf(A)P \le f(PAP) \ , \qquad \text{para toda} \qquad A \in \mathcal{M}_n(\mathbb{C})^+ \ ,$$
ya que, en tal caso, podríamos deducir que $-f$ es $\cup$OP. Para hacerlo, llamemos $Q = I - P$ y tomemos las matrices $U = \begin{bmatrix} P & Q \\ Q & P \end{bmatrix} \in \mathcal{U}(2n)$ y $T = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})^+$. Como se vió en 3.7.10, para todo $\varepsilon > 0$ existe $\lambda > 0$ tal que
$$UTU = UTU^* = \begin{bmatrix} PAP & PAQ \\ QAP & QAQ \end{bmatrix} \le \begin{bmatrix} PAP + \varepsilon I & 0 \\ 0 & \lambda I \end{bmatrix} \ .$$

Como $f(0) \geq 0$, se tiene que $T' = \begin{bmatrix} f(A) & 0 \\ 0 & 0 \end{bmatrix} \leq \begin{bmatrix} f(A) & 0 \\ 0 & f(0)I \end{bmatrix} = f(T)$. Reemplazando T por T' y usando que f es MOP, obtenemos que

$$\begin{bmatrix} Pf(A)P & Pf(A)Q \\ Qf(A)P & Qf(A)Q \end{bmatrix} = UT'U^* \leq Uf(T)U^* = f(UTU^*) \leq \begin{bmatrix} f(PAP + \varepsilon I) & 0 \\ 0 & f(\lambda)I \end{bmatrix} .$$

En particular, se cumple que $Pf(A)P \leq f(PAP + \varepsilon I)$ para todo $\varepsilon > 0$. Como f es continua, por la Proposición 6.1.5 se tiene que $Pf(A)P \leq f(PAP)$, como necesitábamos.

Para probar la recíproca, tomemos $A \leq B$ ambos en $\mathcal{M}_n(\mathbb{C})^+$.

$$\text{Dado } \lambda \in (0,1), \text{ podemos escribir} \qquad \lambda B = \lambda A + (1-\lambda)\frac{\lambda}{1-\lambda}(B - A) .$$

Si f es cóncava (y con valores positivos), esto nos da que

$$f(\lambda B) \geq \lambda f(A) + (1-\lambda) f\Big(\frac{\lambda}{1-\lambda}(B-A)\Big) \geq \lambda f(A) , \qquad \text{para todo} \qquad \lambda \in (0,1) .$$

Tomando $\lambda \to 1^-$, del hecho de que f es continua podemos deducir que $f(B) \geq f(A)$. ■

Corolario 6.3.13. *Sea $f(t) = t^r$, definida en $\mathbb{I} = [0, +\infty)$. Si $r > 1$, f* **no** *es MOP.*

Demostración. Si lo fuera, debería ser $\cap$OP. Pero, como función escalar, es convexa. ■

Corolario 6.3.14.

1. *Las funciones $t \mapsto t^r$, para $r \in (0,1)$, son $\cap OP$ en $[0, +\infty)$.*
2. *$f(t) = \log t$ es $\cap OP$ en $(0, +\infty)$.*

Demostración. La primera parte se deduce de los Teoremas 6.2.6 y 6.3.12. Para probar la concavidad de operadores del logaritmo, fijemos un subespacio $\mathcal{S}$ de $\mathbb{C}^n$. Dada $A \in \mathcal{G}l(n)^+$, por lo anterior sabemos que

$$(A^r)_{\mathcal{S}} \leq (A_{\mathcal{S}})^r , \quad r \in \ (0,1) .$$

Luego, por el Lema 6.2.7,

$$\begin{aligned} \log(A_{\mathcal{S}}) &= \lim_{r \to 0^+} \frac{(A_{\mathcal{S}})^r - I_{\mathcal{S}}}{r} \quad \geq \quad \lim_{r \to 0^+} \frac{(A^r)_{\mathcal{S}} - I_{\mathcal{S}}}{r} \\ &= \left(\lim_{r \to 0^+} \frac{A^r - I_n}{r} \right)_{\mathcal{S}} = (\log A)_{\mathcal{S}} . \end{aligned}$$

Por el Teorema 6.3.6, deducimos que log es $\cap$OP. ■

Observación 6.3.15. El Teorema 6.3.6 da un criterio heurístico para dilucidar qué funciones crecientes pueden ser MOP's: por lo menos deberían ser cóncavas de números. Esto es coherente con los ejemplos: t^r para $r \leq 1$, $-t^{-1}$, $\log t$ son todas cóncavas.

Sin embargo hay que tener mucho cuidado. Porque el Teorema 6.3.6 pide que la f, además de tomar valores positivos, debe estar definida en **toda** la semirrecta $[0, +\infty)$, incluido el cero, y hasta el infinito. Esto se ve claramente mirando bien la prueba, porque uno hace tender ε a cero, por lo que λ se va a infinito, y uno necesita poder tomar $f(\lambda)$. Y para demostrar la implicación MOP $\implies$ $\cap$OP, se usa también que exista $f(0)$. (ejercicio: probar $\cap$OP $\implies$ MOP para f no definida en 0. Ojo con $B - A$). Por ello el Teorema no se puede aplicar directamente a los ejemplos $-t^{-1}$ y $\log t$ (para ver que $\log t$ es $\cap$OP hizo falta el razonamiento de recién, pasando por t^r).

Pero la cosa es más grave si el dominio de f se termina antes del infinito. Ahí el criterio heurístico (que los autores difundíamos muy confiados hasta que fuimos despabilados por unos alumnos despiertos) es directamente **erróneo**. Para convencerse, basta recordar (de la Proposición 6.2.12) que la función $f : [0, \pi/2) \to [0, +\infty)$ dada por $f(t) = \tan t$ es MOP, siendo a la vez **convexa** como función numérica en esa mitad de su dominio. ▲

6.4 Ejercicios

Ejercicios que aparecen en el texto

6.4.1. Verificar que el cálculo funcional cumple las siguientes propiedades: Sean $\mathbb{I} \subseteq \mathbb{R}$ un intervalo, $f, g : \mathbb{I} \to \mathbb{C}$ dos funciones y $A \in \mathcal{H}_{\mathbb{I}}(n)$. Entonces

1. $(f \pm g)(A) = f(A) \pm g(A)$ y $fg(A) = f(A)g(A)$.
2. $\sigma(f(A)) = \{f(\lambda) : \lambda \in \sigma(A)\}$. Más aún, $\mu(f(A)) = f(\mu(A))^{\downarrow}$.
3. $f(A)$ siempre es una matrix normal.
4. $f(t) \in \mathbb{R}$ para todo $t \in \mathbb{I}$ si y sólo si $f(A) \in \mathcal{H}(n)$ para toda $A \in \mathcal{H}_{\mathbb{I}}(n)$.
5. $f(B) \geq 0$ para toda $B \in \mathcal{H}_{\mathbb{I}}(n)$ si y sólo si $f(t) \geq 0$ para todo $t \in \mathbb{I}$.
6. Si $U \in \mathcal{U}(n)$, entonces $f(UAU^*) = Uf(A)U^*$.
7. Si la matriz de A en alguna BON tiene la forma
$$A = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}, \qquad \text{entonces} \qquad f(A) = \begin{bmatrix} f(B) & 0 \\ 0 & f(C) \end{bmatrix}.$$
8. Si una sucesión $(f_m)_{m\in\mathbb{N}}$ de funciones definidas en $\mathbb{I}$ convergen **puntualmente** a f, entonces $f_m(B) \xrightarrow[m\to\infty]{} f(B)$ para toda $B \in \mathcal{H}_{\mathbb{I}}(n)$.
9. Si tiene sentido la composición $h \circ f$, entonces $g \circ f(A) = h(f(A))$.

6.4.2. Probar las siguientes afirmaciones.

1. Si $f : \mathbb{R}_+^* \to \mathbb{R}$ esta dada por $f(t) = t^{-1}$, entonces $f(A) = A^{-1}$ para toda $A \in \mathcal{G}l(n)^+$.

2. Si $A \in \mathcal{M}_n(\mathbb{C})^+$, entonces $A^{1/2} = f(A)$, donde $\mathbb{I} = \mathbb{R}_+$, $f(t) = \sqrt{t}$ y $A^{1/2}$ es la única raiz cuadrada positiva de A definida en la Proposición 3.1.3.

3. Si $A \in \mathcal{H}(n)$, entonces $e^A := \exp(A) = \displaystyle\sum_{m=0}^{\infty} \frac{A^m}{m!}$.

4. Si $A \in \mathcal{G}l(n)^+$, entonces existe $B = \log A$, que es la única matriz autoadjunta que verifica la fórmula $e^B = A$.

6.4.3. Completar los detalles de la demostración del Teorema 6.1.9. En particular, con las notaciones de allí, mostrar que dado $\varepsilon > 0$, existe $m_0 \in \mathbb{N}$ tal que, para todo $m \geq m_0$,

$$\|f(B+H) - f(B) - (P_m(B+H) - P_m(B))\|_2 \leq \varepsilon \|H\|_2$$

para todo $H \in \mathcal{H}(n)$ tal que $B + H \in \mathcal{H}_{\mathbb{J}}(n)$. Se sugiere acotar el incremento de la función $P_k - P_m$ usando su diferencial en un punto intermedio del segmento entre B y $B + H$, y que esas diferenciales convergen uniformemente a cero.

6.4.4. Probar que

1. La suma y la composición (cuando tenga sentido) de MOP´s es MOP.

2. Dada una matriz $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{M}_2(\mathbb{R})$, con $d \neq 0$, definamos la función

$$f_M(t) = \frac{a + bt}{c + dt} \ , \quad t \neq \frac{-c}{d} \ .$$

   Entonces f_M es MOP en $(\frac{-c}{d}, +\infty)$ si y sólo si $\det M \leq 0$.

6.4.5. Sea $\gamma : [a, b] \to \mathcal{H}(n)$ una curva suave tal que $\dot{\gamma}(t) \in \mathcal{M}_n(\mathbb{C})^+$ para todo $t \in (a, b)$. Probar que γ es *creciente*, en el sentido de que $t \leq s \implies \gamma(t) \leq \gamma(s)$, en el orden de $\mathcal{H}(n)$. En particular, deducir que $\gamma(a) \leq \gamma(b)$. Se suguiere chusmear el Teorema 6.2.9.

6.4.6. Rellenar los detalles de la siguiente prueba alternativa de la Proposición 6.2.4: Supongamos que $A < B$, ambos en $\mathcal{G}l(n)^+$. Entonces definimos la función

$$C : [0, 1] \to \mathcal{G}l(n)^+ \ , \quad \text{dada por} \quad C(t) = A + t(B - A) \ , \quad t \in [0, 1] \ .$$

Sea $R(t) = C(t)^{1/2}$, $t \in [0, 1]$. Entonces

$$R(t)^2 = C(t) \quad \implies \quad \dot{R}(t)R(t) + R(t)\dot{R}(t) = \dot{C}(t) = B - A > 0 \ , \quad t \in [0, 1] \ ,$$

donde el *punto* denota derivada respecto de t. Por la Observación 9.1.5, como $R(t) > 0$ y $\dot{C}(t) = S(R, \dot{R}) > 0$, entonces, $\dot{R}(t) > 0$ para todo $t \in [0, 1]$. Luego R es *creciente* y, en particular, $A^{1/2} = R(0) < R(1) = B^{1/2}$.

6.4.7. Probar que

1. La suma y la composición (cuando tenga sentido) de ∪OP´s es ∪OP.

2. Dada una matriz $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathcal{M}_2(\mathbb{R})$, con $d \neq 0$, definamos la función

$$f_M(t) = \frac{a+bt}{c+dt} \ , \quad t \neq \frac{-c}{d} \ .$$

   Entonces f_M es ∪OP en $\mathbb{I} = (\frac{-c}{d}, +\infty)$ si y sólo si $\det M \geq 0$. Por otra parte, f es ∩OP en $\mathbb{I}$ si y sólo si f es MOP en $\mathbb{I}$ si y sólo si $\det M \leq 0$.

3. Sea $\mathbb{I} \subseteq \mathbb{R}$ un intervalo tal que $0 \in \mathbb{I}$. Sean $f(t) = |t|$ y $g(t) = t \vee 0$, $t \in \mathbb{I}$. Entonces f no es ∪OP y g no es ∪OP ni MOP.

6.4.8. Sea $\mathbb{I} \subseteq \mathbb{R}$ un intervalo tal que $0 \in \mathbb{I}$ y sea $f : \mathbb{I} \to \mathbb{R}$, una función. Entonces son equivalentes:

1. f es convexa de operadores y $f(0) \leq 0$.

2. Dados $n \in \mathbb{N}$, $A \in \mathcal{H}_{\mathbb{I}}(n)$, se cumple que

$$f(C^*AC) \leq C^*f(A)C. \tag{6.12}$$

   para todo $C \in \mathcal{M}_n(\mathbb{C})$ tal que $\|C\| \leq 1$.

3. Dados $n \in \mathbb{N}$, $A, B \in \mathcal{H}_{\mathbb{I}}(n)$ y $C, D \in \mathcal{M}_n(\mathbb{C})$ tales que $C^*C + D^*D \leq I$, se verifica

$$f(C^*AC) + f(D^*BD) \leq C^*f(A)C + D^*f(B)D.$$

Se sugiere usar las matrices unitarias construidas en 3.7.9.

Ejercicios nuevos

6.4.9. Mostrar que todas las definciones y propiedades del cálculo funcional no necesitan que el dominio de f sea un intervalo. En particular, verificar que si $U \subseteq \mathbb{R}$ es abierto, entonces

1. $\mathcal{H}_U(n) := \{A \in \mathcal{H}(n) : \sigma(A) \subseteq U\}$ es abierto en $\mathcal{H}(n)$.

2. Si $f : U \to \mathbb{R}$ es C^1, su extensión $f : \mathcal{H}_U(n) \to \mathcal{H}(n)$ es diferenciable.

3. El Teorema 6.1.9 y el Corolario 6.1.10 siguen siendo válidos en este contexto.

6.4.10. Sea $A \in \mathcal{H}(n)$ y sea $\gamma : (-1, 1) \to \mathcal{H}(n)$ una curva C^1 tal que $\gamma(0) = A$. Sea $\lambda \in \sigma(A)$ una raiz **simple** de $P_A(x)$, y sea $x_0 \in \ker(A - \lambda I)$ un autovector unitario. Probar que existe un $\varepsilon > 0$ y una curva suave $x : (-\varepsilon, \varepsilon) \to \mathbb{C}^n$ tal que

1. $x(0) = x_0$.
2. $x(t)$ es autovector de $\gamma(t)$ para todo $t \in (-\varepsilon, \varepsilon)$.
3. La función $(-\varepsilon, \varepsilon) \ni t \mapsto \lambda(t)$, que da el autovalor asociado a cada $x(t)$, es suave.
4. Todos los $\lambda(t)$ son autovectores simples de $\gamma(t)$.

Mostrar, además, que $\|\dot{x}(0)\| \le \mathrm{d}\left(\lambda\,, \sigma(A) \setminus \{\lambda\}\,\right)^{-1} \|\dot{\gamma}(0) \cdot x_0\|$.

Sugerencia: Tomar un abierto $U \supseteq \sigma(A)$ que separe a λ del resto de $\sigma(A)$, y definir allí la función f que vale uno cerca de λ y cero en el resto de U. Tomar $g(t) = f(\gamma(t)\,)$, para los $t \in (-\varepsilon, \varepsilon)$ tales que $\gamma(t) \in \mathcal{H}_U(n)$, observar que cada $g(t)$ es un proyector autoadjunto de rango uno (por el Corolario 2.3.8) y usar la Eq. (6.1) para ver qué proyector es. Definir entonces $x(t) = g(t) \cdot x_0$. Para obtener $\lambda(t)$, buscar una coordenada no nula de x_0 y dividir ahí (o tomar $\lambda(t) = \operatorname{tr} g(t)\gamma(t)$). Para acotar la norma, diagonalizar adecuadamente a A y luego usar el Corolario 6.1.10.

6.4.11. Sean $U(t) = \begin{bmatrix} \cos t & -\operatorname{sen}|t| \\ \operatorname{sen}|t| & \cos t \end{bmatrix}$ y $A(t) = U(t) \begin{bmatrix} 1+t^2 & 0 \\ 0 & 1-t^2 \end{bmatrix} U(t)^*$ para $t \in \mathbb{R}$. Mostrar que la curva $A(t)$ es suave cerca de 0, pero como $A(0) = I$ tiene multiplicidades, **no** hay curvas **suaves** $x(t)$ a valores en $\mathbb{R}^n$ que cumplan lo mismo que en el ejercicio anterior.

Ojo: La curva $x(t) = U(t)e_1$ no es suave, pero hay que ver que no puede haber otra suave.

6.4.12. Sea $\mathbb{I} \subseteq \mathbb{R}$ un intervalo tal que $0 \in \mathbb{I}$ y sea $f : \mathbb{I} \to \mathbb{R}$, una función tal que $f(0) \le 0$. Entonces son equivalentes:

1. f es convexa (a secas, no pedimos $\cup$OP).
2. Dados $n \in \mathbb{N}$ y $A \in \mathcal{H}_{\mathbb{I}}(n)$, se cumple que
$$f(C^*AC) \prec_w C^*f(A)C.$$
para todo $C \in \mathcal{M}_n(\mathbb{C})$ tal que $\|C\|_{sp} \le 1$. Comparar con la Eq. (6.12).

Sugerencia: Probar que si f es convexa entonces $f\big(\,\langle Ax,\ x\rangle\,\big) \le \langle f(A)x,\ x\rangle$ para todo vector unitario x y toda matriz $A \in \mathcal{H}_{\mathbb{I}}(n)$.

6.4.13. Sea $\mathbb{I} \subseteq \mathbb{R}$ un intervalo tal que $0 \in \mathbb{I}$ y sea $f : \mathbb{I} \to \mathbb{R}$, una función convexa **creciente** tal que $f(0) \le 0$. Dados $n \in \mathbb{N}$ y $A \in \mathcal{H}_{\mathbb{I}}(n)$, probar que para todo $i \in \mathbb{I}_n$ se verifica que
$$\lambda_i\left(f(C^*AC)\,\right) \le \lambda_i\left(C^*f(A)C\right).$$
donde $C \in \mathcal{M}_n(\mathbb{C})$ es una contracción. Dar un contraejemplo si la función no es creciente. Sugerencia: Usar minimax.

6.4.14. Sea $C \in \mathcal{M}_n(\mathbb{C})$ una contracción y $A \in \mathcal{M}_n(\mathbb{C})^+$. Demostrar que dados $r, s \in \mathbb{R}$ tales que $1 \le r \le s$, entonces $(C^*A^rC)^{1/r} \le (C^*A^sC)^{1/s}$.

6.4.15. Sea $A, B \in \mathcal{M}_n(\mathbb{C})^+$, $r \in (1, +\infty)$ y $\alpha \in [0, 1]$. Probar que

$$(A^r + B^r)^{\frac{1}{r}} \geq \alpha^{1-\frac{1}{r}} \; A + (1-\alpha)^{1-\frac{1}{r}} \; B \; .$$

Sugerencia: Analizar separadamente los casos $\alpha = 0$ o 1 y $\alpha \in (0, 1)$. Usar que $t \mapsto t^{\frac{1}{r}}$ es tanto MOP como $\cap$OP. Si no sale, chusmear el Lema 9.5.3.

6.4.16. Sean $A, B \in \mathcal{M}_n(\mathbb{C})^+$. Probar que la función

$$[1\,, +\infty) \ni p \longmapsto \left(\frac{A^p + B^p}{2} \right)^{\frac{1}{p}}$$

es creciente, relativa al orden $\leq$ de $\mathcal{M}_n(\mathbb{C})^+$.
Sugerencia: Dados $r, q \in [1\,, +\infty)$ tales que $r < q$, aplicar el ejercicio anterior para los números $r = \frac{q}{r} > 1$ y $\alpha = \frac{1}{2}$. Si no sale, chusmear la Proposición 9.5.6.

Capítulo 7

Productos tensoriales y alternados

7.1 Producto tensorial de a dos

Comenzaremos fijando ciertas convenciones de notación adecuadas para esta teoría:

1. Para cada $n \in \mathbb{N}$, llamaremos $\mathbb{H}_n = \mathbb{C}^n$ con el producto interno usual. Denotaremos por $e_1^{(n)}, \ldots, e_n^{(n)}$ a los vectores de la base canónica de $\mathbb{H}_n$.

2. Llamaremos $\mathbb{H}_n \otimes \mathbb{H}_k$ al espacio de funcionales $F : \mathbb{H}_n \times \mathbb{H}_k \to \mathbb{C}$ *bilineales* (i.e., lineales en cada coordenada), pensado como $\mathbb{C}$-espacio vectorial de la forma natural.

3. Dada $F \in \mathbb{H}_n \otimes \mathbb{H}_k$, le asociamos la matriz $F = \left[F(e_i^{(n)}, e_j^{(k)})\right] \in \mathcal{M}_{n,k}(\mathbb{C})$. Luego

$$F(x, y) \ = \ \sum_{i=1}^{n} \sum_{j=1}^{k} x_i \ y_j \ F_{ij} \ = \ x^T \ F \ y \ , \quad x \in \mathbb{H}_n \ , \quad y \in \mathbb{H}_k \ .$$

   Esto muestra que podemos identificar naturalmente $\mathbb{H}_n \otimes \mathbb{H}_k$ con $\mathcal{M}_{n,k}(\mathbb{C})$.

4. Esto permite, ademas, definir el producto interno natural en $\mathbb{H}_n \otimes \mathbb{H}_k$. Dadas $F, G \in \mathbb{H}_n \otimes \mathbb{H}_k$, las identificamos con sus matrices en $\mathcal{M}_{n,k}(\mathbb{C})$, definimos

$$\langle F, G \rangle = \operatorname{tr} G^* F = \sum_{(i,j) \in \mathbb{I}_n \times \mathbb{I}_k} F_{ij} \ \overline{G_{ij}} \quad . \tag{7.1}$$

5. Dados $x \in \mathbb{H}_n$ e $y \in \mathbb{H}_k$, notaremos por $x \otimes y \in \mathbb{H}_n \otimes \mathbb{H}_k$, al llamado *tensor elemental*, dado por

$$x \otimes y(u, v) = \langle u, \overline{x} \rangle \langle v, \overline{y} \rangle \ , \quad u \in \mathbb{H}_n \, , \ v \in \mathbb{H}_k \ . \tag{7.2}$$

   Observar que $\langle u, \overline{x} \rangle \langle v, \overline{y} \rangle = x^T u \cdot y^T v = u^T \big(x y^T\big) v$. Por lo tanto, la matriz de $x \otimes y$ es $x y^T \in \mathcal{M}_{n,k}(\mathbb{C})$. Por lo tanto, no toda $F \in \mathbb{H}_n \otimes \mathbb{H}_k$ es elemental, pero sí sucede que toda F es *suma* de tensores elementales, porque las matrices del tipo $x y^T$ son todas las de rango uno. Observar que la aplicación

$$\mathbb{H}_n \times \mathbb{H}_k \ni (x, y) \longmapsto x \otimes y \in \mathbb{H}_n \otimes \mathbb{H}_k$$

es bilineal. Además, a partir de la Eq. (7.1), vemos la fórmula

$$\langle x \otimes y \ , \ u \otimes v \rangle = \operatorname{tr}\left((uv^T)^* xy^T\right) = \operatorname{tr}\left(u^* x \cdot y^T \overline{v}\right) = \langle x, u \rangle \cdot \langle y, v \rangle \ , \tag{7.3}$$

para $x, u \in \mathbb{H}_n$, $y, v \in \mathbb{H}_k$.

6. Se puede deducir que el conjunto

$$\mathcal{E}_{n,k} = \{e_i^{(n)} \otimes e_j^{(k)} : i \in \mathbb{I}_n \ , \ j \in \mathbb{I}_k\} \sim \{E_{ij} \in \mathcal{M}_{n,k}(\mathbb{C}) : i \in \mathbb{I}_n \ , \ j \in \mathbb{I}_k\}$$

   es una base ortonormal de $\mathbb{H}_n \otimes \mathbb{H}_k$, que llamaremos base canónica. La consideraremos ordenada alfabéticamente (leyéndola por $filas$).

7. Dados $A \in L(\mathbb{H}_n)$ y $B \in L(\mathbb{H}_k)$, podemos definir el operador $A \otimes B \in L(\mathbb{H}_n \otimes \mathbb{H}_k)$, a través de la fórmula $A \otimes B(F) = AFB^T$, para $F \in \mathbb{H}_n \otimes \mathbb{H}_k$, pensado como una matriz en $\mathcal{M}_{n,k}(\mathbb{C})$. En particular,

$$A \otimes B(x \otimes y) \simeq Axy^T B^T = (Ax) \cdot (By)^T \simeq Ax \otimes By \ , \quad x \in \mathbb{H}_n \ , \quad y \in \mathbb{H}_k \ .$$

   Observar que esta ecuación no define a $A \otimes B$ en todos los elementos de $\mathbb{H}_n \otimes \mathbb{H}_k$, pero sí lo caracteriza completamente (por ser lineal).

8. El producto tensorial de matrices verifica las siguientes propiedades:

   (a) Sean $I_n \in \mathcal{M}_n(\mathbb{C})$ y $I_k \in \mathcal{M}_k(\mathbb{C})$. Entonces $I_n \otimes I_k$ es la identidad de $\mathbb{H}_n \otimes \mathbb{H}_k$.

   (b) $(\alpha A_1 + A_2) \otimes B = \alpha(A_1 \otimes B) + A_2 \otimes B$, para todo $\alpha \in \mathbb{C}$.

   (c) $(A \otimes B)^* = A^* \otimes B^*$.

   (d) $(A_1 \otimes B_1)(A_2 \otimes B_2) = A_1A_2 \otimes B_1B_2$.

   (e) Si existen A^{-1} y B^{-1}, entonces $A^{-1} \otimes B^{-1} = (A \otimes B)^{-1}$. En particular, si $A \in \mathcal{U}(n)$ y $B \in \mathcal{U}(k)$, entonces $A \otimes B \in \mathcal{U}(nk)$.

   (f) $A \otimes B \geq 0$ si $A \geq 0$ y $B \geq 0$. Más aún, $|A \otimes B| = |A| \otimes |B|$. Se usa el Teorema 3.1.3 y la unicidad de la raiz cuadrada positiva.

Observación 7.1.1. Dados $A \in L(\mathbb{H}_n)$ y $B \in L(\mathbb{H}_k)$, la matriz de $A \otimes B$ en la base canónica de $\mathbb{H}_n \otimes \mathbb{H}_k$ (ordenada por filas) es el llamado producto de Kronecker de A y B que se define como la matriz por bloques

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \dots & a_{nn}B \end{bmatrix} \in \mathcal{M}_{nk}(\mathbb{C}) \ . \tag{7.4}$$

La verificación es sumamente tediosa, pero podemos dar un esbozo: La base canonica de $\mathbb{H}_n \otimes \mathbb{H}_k$ se ordena así:

$$e_1^{(n)} \otimes e_1^{(k)}, \dots, e_1^{(n)} \otimes e_k^{(k)}, \ e_2^{(n)} \otimes e_1^{(k)}, \dots, e_2^{(n)} \otimes e_k^{(k)}, \dots\dots, e_n^{(n)} \otimes e_1^{(k)}, \dots, e_n^{(n)} \otimes e_k^{(k)} \ .$$

Luego, el vector $e_i^{(n)} \otimes e_r^{(k)}$ se ubica en el lugar $k(i-1)+r$ de la base canónica. Fijemos un par $i, j \in \mathbb{I}_n$. Como el bloque de $k \times k$ ubicado en el jugar (i,j) involucra las filas entre $k(i-1)+1$ y ki, y a las columnas $k(j-1)+1$ y kj, se escribe

$$\Big(\big\langle A \otimes B(e_j^{(n)} \otimes e_s^{(k)}) \ , \ e_i^{(n)} \otimes e_r^{(k)} \big\rangle \Big)_{r,s \in \mathbb{I}_k} = \big(\langle A e_j^{(n)}, e_i^{(n)} \rangle \cdot \langle B e_s^{(k)}, e_r^{(k)} \rangle \big)_{r,s \in \mathbb{I}_k} = a_{ij} B \ ,$$

como se afirmaba. ▲

Proposición 7.1.2. *Sean $A \in \mathcal{M}_n(\mathbb{C})$ y $B \in \mathcal{M}_m(\mathbb{C})$. Si los autovalores de A son la n-upla $\{\lambda_1, ..., \lambda_n\}$, y los de B son $\{\mu_1, ..., \mu_m\}$, entonces los autovalores de $A \otimes B$ son*

$$\{ \ \lambda_{(i,j)} \}_{(i,j) \in \mathbb{I}_n \times \mathbb{I}_m} \ , \qquad \textit{donde} \qquad \lambda_{(i,j)} = \lambda_i \, \mu_j \ ,$$

todos contados con multiplicidad.

Demostración. Aplicando el Teorema 1 de Schur 1.6.1, si $A = UT_1U^*$ y $B = VT_2V^*$, con U, V unitarias y T_1 , T_2 triangulares superiores, entonces

$$A \otimes B = U \otimes V \cdot T_1 \otimes T_2 \cdot (U \otimes V)^* \ ,$$

por lo que $\sigma(A \otimes B) = \sigma(T_1 \otimes T_2)$ (con multiplicidades). Por la representación matricial de $T_1 \otimes T_2$ como producto de Kronecker (que queda también triangular superior, pero con los productos $\lambda_i \mu_j$ en su diagonal), se obtiene la igualdad anunciada. ■

Corolario 7.1.3. *Sean $A \in \mathcal{M}_n(\mathbb{C})$ y $B \in \mathcal{M}_m(\mathbb{C})$.*

1. $\|A \otimes B\|_{sp} = \|A\|_{sp} \cdot \|B\|_{sp}$.
2. *Más aún, los valores singulares de $A \otimes B$ son*

$$s(A \otimes B) = \{ \ s_i(A) s_j(B) \ \}_{(i,j) \in \mathbb{I}_n \times \mathbb{I}_m}$$

contados con multiplicidad, y ordenados en forma decreciente.

Demostración. Se sigue de que $|A \otimes B| = |A| \otimes |B|$ y de la Proposición 7.1.2. ■

7.2 Potencias tensoriales

Una cuenta muy engorrosa, aunque elemental, muestra que se tiene un isomorfismo natural entre $(\mathbb{H}_n \otimes \mathbb{H}_k) \otimes \mathbb{H}_r$ y $\mathbb{H}_n \otimes (\mathbb{H}_k \otimes \mathbb{H}_r)$, identificando a ambos con las funciones *trilineales* en $\mathbb{H}_n \times \mathbb{H}_k \times \mathbb{H}_r$. La clave es observar que, dados $x \in \mathbb{H}_n$, $y \in \mathbb{H}_k$ y $z \in \mathbb{H}_r$, a los tensores elementales $(x \otimes y) \otimes z$ & $x \otimes (y \otimes z)$ se los puede identificar con la misma funcional trilineal, por una fórmula semejante a (7.2). Es decir, que el producto tensorial es *asociativo*. Esto permite definir productos de varios espacios, sin necesidad de aclarar el orden en que se los define. Lamentablemente, en ese contexto se pierde la representación de las funciones multilineales como matrices, a menos que se quiera pensar en matrices de muchas dimensiones. Usaremos particularmente la asociatividad para definir *potencias*, en el sentido tensorial, de un mismo espacio $\mathbb{H}_n$, y de operadores en $\mathbb{H}_n$. Damos, a continuación, el listado de notaciones y resultados que se siguen naturalmente (y que se prueban planarmente por inducción usando lo anterior y la asociatividad):

7.2.1. Sean $n, k \in \mathbb{N}$.

1. Notaremos $\bigotimes^k \mathbb{H}_n$, llamado *espacio k-tensorial* sobre $\mathbb{H}_n$, al producto tensorial de $\mathbb{H}_n$ por sí mismo k veces. Los elementos de $\bigotimes^k \mathbb{H}_n$ se pueden pensar como funcionales k-multilineales $F : \mathbb{H}_n^k \to \mathbb{C}$.

2. Dados $x_1, \cdots, x_k \in \mathbb{H}_n$, se define el k-tensor elemental $x_1 \otimes x_2 \cdots \otimes x_k$ por la fórmula

$$x_1 \otimes \cdots \otimes x_k \ (u_1, \cdots, u_k) = \prod_{i=1}^{k} \langle u_i, \overline{x}_i \rangle \ , \quad (u_1, \cdots, u_k) \in \mathbb{H}_n^k \ . \tag{7.5}$$

   Luego todo elemento de $\bigotimes^k \mathbb{H}_n$ es suma de k-tensores elementales.

3. El producto interno sobre $\bigotimes^k \mathbb{H}_n$, definido inductivamente en todo par de elementos de $\bigotimes^k \mathbb{H}_n$, está determinado por el producto de k-tensores:

$$\Big\langle x_1 \otimes x_2 \cdots \otimes x_k \ , \ y_1 \otimes y_2 \cdots \otimes y_k \Big\rangle = \prod_{i=1}^{k} \langle x_i, y_i \rangle \ , \tag{7.6}$$

   para $x_1, \ldots, \ x_k, \ y_1, \ldots, \ y_k \in \mathbb{H}_n$.

4. La aplicación $\mathbb{H}_n^k \ni (x_1, \ldots, x_k) \mapsto x_1 \otimes \cdots \otimes x_k \in \bigotimes^k \mathbb{H}_n$ es k-multilineal.

5. La base canónica ortonormal de $\bigotimes^k \mathbb{H}_n$ es, por definición,

$$\Big\{ e_{\alpha_1}^{(n)} \otimes e_{\alpha_2}^{(n)} \otimes \cdots \otimes e_{\alpha_k}^{(n)} \ : \ \alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k) \in \mathbb{I}_n^k \Big\}.$$

   Luego $\dim \bigotimes^k \mathbb{H}_n = n^k$. ▲

7.2.2. Todo operador $A : \mathbb{H}_m \to \mathbb{H}_n$ induce un operador de $\bigotimes^k A : \bigotimes^k \mathbb{H}_m \to \bigotimes^k \mathbb{H}_n$, llamado *potencia k-tensorial* de A, determinado por la fórmula

$$\bigotimes^k A \ (x_1 \otimes x_2 \cdots \otimes x_k) = Ax_1 \otimes Ax_2 \otimes \cdots \otimes Ax_k \ , \tag{7.7}$$

para $x_1, \ldots, \ x_k \in \mathbb{H}_m$. Se tienen las siguientes propiedades:

a. Dados $A \in L(\mathbb{H}_m, \mathbb{H}_n)$ y $B \in L(\mathbb{H}_n, \mathbb{H}_r)$,

$$\bigotimes^k (AB) = \bigotimes^k A \cdot \bigotimes^k B \ . \tag{7.8}$$

b. $(\bigotimes^k A)^* = \bigotimes^k A^*$.

c. Si $A \in \mathcal{G}l(n)$, entonces $\bigotimes^k A^{-1} = (\bigotimes^k A)^{-1}$. En particular $\bigotimes^k A$ es unitaria si $A \in \mathcal{U}(n)$.

d. Si $A \in \mathcal{M}_n(\mathbb{C})^+$, entonces $\bigotimes^k A \geq 0$, porque $A = C^*C$ para algún $C \in \mathcal{M}_n(\mathbb{C})$, y $\otimes^k A = (\otimes^k C)^* \otimes^k C$.

e. Más generalmente, si $A \in \mathcal{M}_n(\mathbb{C})$, se tiene que $|\bigotimes^k A| = \bigotimes^k |A|$.

f. Si los autovalores de A son $\lambda_1, \dots, \lambda_n$, los autovalores (resp. valores singulares) de $\bigotimes^k A$ son

$$\Big\{ \prod_{j=1}^{k} \lambda_{i_j} \ : \ (i_1, \dots, i_k) \in \mathbb{I}_n^k \Big\} \ \Big(\text{resp. } \Big\{ \prod_{j=1}^{k} s_{i_j}(A) \ : \ (i_1, \dots, i_k) \in \mathbb{I}_n^k \Big\} \Big) \ ,$$

contados con multiplicidad. En particular se tiene que

$$\rho\Big(\bigotimes^k A \Big) = \rho(A)^k \qquad \text{y} \qquad \Big\| \bigotimes^k A \Big\|_{sp} = \|A\|_{sp}^k \ . \qquad \blacktriangle$$

7.3 Productos alternados y determinantes

Sea $\mathbf{S}_k$ el *grupo simétrico* de grado k, esto es el grupo de todas la permutaciones de $\mathbb{I}_k$. Cada $\pi \in \mathbf{S}_k$ da lugar a un operador linear $P_\pi^{(n)} \in \mathcal{U}(\bigotimes^k \mathbb{H}_n)$, por la siguiente fórmula: Si pensamos a los elementos de $\bigotimes^k \mathbb{H}_n$ como funcionales k- multilineales $F : \mathbb{H}_n^k \to \mathbb{C}$, se define

$$P_\pi^{(n)}(F) \ (x_1, x_2, \cdots, x_k) = F(x_{\pi(1)}, x_{\pi(2)}, \cdots, x_{\pi(k)}) \ , \quad (x_1, x_2, \cdots, x_k) \in \mathbb{H}_n^k \ . \qquad (7.9)$$

Dados $x_1, \dots, \ x_k \in \mathbb{H}_n$, es fácil ver, usando la fórmula (7.5), que

$$P_\pi^{(n)}(x_1 \otimes x_2 \cdots \otimes x_k) = x_{\pi^{-1}(1)} \otimes x_{\pi^{-1}(2)} \cdots \otimes x_{\pi^{-1}(k)} \ . \qquad (7.10)$$

Observación 7.3.1. El hecho de que $P_\pi^{(n)}$ sea unitario se puede probar mostrando primero que $(P_\pi^{(n)})^* = P_{\pi^{-1}}^{(n)}$ (esto puede hacerse usando solo los k-tensores elementales). Despues, ahí si por definición, se ve que $P_{\pi^{-1}}^{(n)} = (P_\pi^{(n)})^{-1}$. $\blacktriangle$

Definiremos a continuación las nociones básicas de productos alternados

Definición 7.3.2. Sea $n \in \mathbb{N}$ y $k \in \mathbb{I}_n$. Llamaremos espacio *k-alternado* (o k-ésimo *Grassmann*) sobre $\mathbb{H}_n$, al subespacio de $\bigotimes^k \mathbb{H}_n$ dado por

$$\Lambda^k \mathbb{H}_n = \Big\{ \ F \in \bigotimes^k \mathbb{H}_n \ : \ P_\pi^{(n)} F = \operatorname{sgn}(\pi) \ F \quad \text{para toda} \quad \pi \in \mathbf{S}_k \ \Big\} \ ,$$

donde $\operatorname{sgn}(\pi) = \pm 1$ de acuerdo a si π es una permutación par o impar. Los elementos de $\Lambda^k \mathbb{H}_n$ se llaman k-*tensores alternados*. Se considera a $\Lambda^k \mathbb{H}_n$ como un espacio de Hilbert con el producto interno de $\bigotimes^k \mathbb{H}_n$.

Observación 7.3.3. Notaremos por $\mathbf{P}_k^n$ a la proyección ortogonal de $\bigotimes^k \mathbb{H}_n$ sobre $\Lambda^k\mathbb{H}_n$. Es fácil ver que $\mathbf{P}_k^n$ está dada por la fórmula

$$\mathbf{P}_k^n = \frac{1}{k!} \sum_{\pi \in \mathbf{S}_k} \operatorname{sgn}(\pi)\ P_\pi^{(n)} \ . \tag{7.11}$$

En efecto, como cada $P_\pi^{(n)} \in \mathcal{U}(\bigotimes^k \mathbb{H}_n)$, entonces $(P_\pi^{(n)})^* = (P_\pi^{(n)})^{-1} = P_{\pi^{-1}}^{(n)}$. Por lo tanto $(\mathbf{P}_k^n)^* = \mathbf{P}_k^n$, ya que al adjuntarlo tan solo se reordena la suma (se usa $\operatorname{sgn}(\pi^{-1}) = \operatorname{sgn}(\pi)$). Por otro lado, como para todo par $\sigma, \pi \in \mathbf{S}_k$ se tiene que

$$\operatorname{sgn}(\sigma\pi) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\pi) \quad \text{ y } \quad P_{\sigma\pi}^{(n)} = P_\sigma^{(n)}\ P_\pi^{(n)} \ ,$$

podemos deducir que $R(\mathbf{P}_k^n) \subseteq \Lambda^k\mathbb{H}_n$. Finalmente, es claro, a partir de la definición de $\Lambda^k\mathbb{H}_n$, que $\mathbf{P}_k^n(F) = F$ para toda $F \in \Lambda^k\mathbb{H}_n$. ▲

Definición 7.3.4. Dados $x_1, \dots,\ x_k \in \mathbb{H}_n$, se define el k-tensor alternado elemental :

$$x_1 \wedge x_2 \cdots \wedge x_k := \mathbf{P}_k^n\ (x_1 \otimes x_2 \ldots \otimes x_k) = \frac{1}{k!} \sum_{\pi \in \mathbf{S}_k} \operatorname{sgn}(\pi)\ x_{\pi(1)} \otimes x_{\pi(2)} \cdots \otimes x_{\pi(k)} \ ,$$

también llamado *producto alternado* de la k-upla ordenada $x_1, x_2 \ldots, x_k$. ▲

Observación 7.3.5. Enumeraremos aquí algunas propiedades de los k-tensores elementales:

1. Notar que, como $\Lambda^k\mathbb{H}_n = \mathbf{P}_k^n\big(\bigotimes^k \mathbb{H}_n\big)$, y los k-tensores elementales generan $\bigotimes^k \mathbb{H}_n$, podemos asegurar que los k-tensores alternados elementales generan $\Lambda^k\mathbb{H}_n$.

2. Usando el ítem 5 de 7.2.1 y el hecho de que $\mathbf{P}_k^n$ es lineal, podemos deducir que la aplicación $(x_1\,, \dots, x_k) \mapsto x_1 \wedge \cdots \wedge x_k$ es k-multilineal.

3. Por otra parte, dados $x_1, x_2 \ldots, x_k \in \mathbb{H}_n$ y $\pi \in \mathbf{S}_k$, se sigue de las definiciones que

$$x_{\pi(1)} \wedge x_{\pi(2)} \wedge \cdots \wedge x_{\pi(k)} = \operatorname{sgn}(\pi)\ x_1 \wedge x_2 \cdots \wedge x_k \ . \tag{7.12}$$

   En resumen, $(x_1\,, \dots, x_k) \mapsto x_1 \wedge \cdots \wedge x_k$ es una aplicación k-multilineal **alternada**.

4. De la fórmula (7.12) puede deducirse que, si existen $x_i = x_j$ con $i \neq j$, entonces $x_1 \wedge \cdots \wedge x_k = 0$ (usando la transposición $\tau = (i, j) \in \mathbf{S}_k$, cuyo $\operatorname{sgn}(\tau) = -1$).

5. Más aún, esto implica que si el conjunto $\{x_1\,, \dots, x_k\}$ es linealmente dependiente, su produco alternado debe ser nulo. xEsto se usará en la subsección siguiente. ▲

Productos alternados y el determinante

En los capítulos anteriores se hizo uso libre de la función determinante

$$\mathcal{M}_n(\mathbb{C}) \ni A \mapsto \det A \in \mathbb{C} \ ,$$

y de sus conocidas propiedades. Dar una exposición completa y formal de dichos resultados es algo que uno siempre trata de evitar, porque es un asunto complicado y poco amigable. Sin embargo, con la teoría de productos alternados a mano, esto está bastante cerca, por lo que trataremos de dar las dos definiciones usuales, mostrar su equivalencia, y dar pruebas de sus propiedades más importantes. Por lo tanto, en esta sección supondremos que nos olvidamos lo que sabemos (y hemos usado) al respecto. Empecemos por una de las definiciones. Asumimos conocida cierta teoría básica de grupos de permutaciones, como hemos hecho hasta ahora.

Definición 7.3.6. Sea $A = (a_{ij})_{i,j\in\mathbb{I}_n} \in \mathcal{M}_n(\mathbb{C})$. Definimos su determinante por la fórmula

$$\det A = \sum_{\sigma\in \mathbf{S}_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} a_{j,\sigma(j)} \in \mathbb{C} . \tag{7.13}$$

Con la misma fórmula se define el determinante de matrices a coeficientes en cualquier anillo (como en $\mathbb{C}[X]$, lo que se usa para definir el polinomio característico de una matriz). ▲

7.3.7. A continuación enumeraremos una serie de propiedades que se siguen fácilmente de esta definición del determinante. Las pruebas que no estén escritas deben considerarse como ejercicios: Sea $A \in \mathcal{M}_n(\mathbb{C})$.

1. $\det A^T = \det A$ y $\det A^* = \overline{\det A}$. Acá se usa solamente que $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$.
2. $\det I = 1$, ya que el sumando $\prod\limits_{j=1}^{n} I_{j,\sigma(j)} = 0$ para toda $\sigma \neq Id$.
3. Si todas las diagonales de A tienen algún cero (en el sentido de la Definición 4.3.2), entonces $\det A = 0$. Por ejemplo (usando el Teorema 4.3.3) esto sucede si existe una submatriz nula de tamaño $k \times r$ con $k + r > n$.
4. En particular, si exite alguna $F_i(A) = 0$ (o bien una columna), entonces $\det A = 0$.
5. Dada $\sigma \in \mathbf{S}_n$ y $P_\sigma \in \mathcal{U}_{\mathcal{P}}(n)$ su matriz de permutación asociada, entonces se tiene que $\det P_\sigma = \operatorname{sgn}(\sigma)$. Esto sale por que la única diagonal sin ceros de P_σ es la producida por la misma σ, como se ve en la Eq. (4.4).
6. Si $T \in \mathcal{TS}(n)$, entonces $\det T = \prod\limits_{i\in\mathbb{I}_n} T_{ii}$. Esto se ha usado sistemáticamente en los capítulos anteriores, y se lo justificó "desarrollando" por la primera columna. Eso no es incorrecto (ver el Ejercicio 7.5.11 o la Eq. (12.13)), pero sale más directo con la Definición 7.3.6, porque la única diagonal sin ceros de T (si es que hay una) es la producida por $\sigma = Id$.
7. La función $\mathcal{M}_n(\mathbb{C}) \ni A \mapsto \det A \in \mathbb{C}$ es continua (más aún, es de clase C^∞), debido a que es un polinomio de grado n en los coeficientes de A. ▲

Para dar la segunda definición y probar las principales propiedades del determinante, necesitamos desarrollar un poco la teoría de productos alternados. La relación clave entre estos y la fórmula (7.13) para el determinante es lo siguiente:

Proposición 7.3.8. *Sean* $x_1, \dots, x_k, y_1, \dots, y_k \in \mathbb{H}_n$. *Entonces*

$$\big\langle x_1 \wedge x_2 \wedge \cdots \wedge x_k \ , \ y_1 \wedge y_2 \wedge \cdots \wedge y_k \big\rangle = \frac{1}{k!} \ \det \ \big(\langle x_i, y_j \rangle \big)_{i,j \in \mathbb{I}_k} \ . \tag{7.14}$$

Demostración. Es consecuencia de las ecuaciones (7.6) y (7.11), vía la Definición 7.3.6 para el determinante. En efecto, si $D \langle x_1 \wedge \cdots \wedge x_k, y_1 \wedge \cdots \wedge y_k \rangle$, entonces

$$\begin{aligned} D &= \langle \mathbf{P}_k^n \ x_1 \otimes \cdots \otimes x_k \ , \mathbf{P}_k^n \ y_1 \otimes \cdots \otimes y_k \rangle = \langle x_1 \otimes \cdots \otimes x_k \ , \mathbf{P}_k^n \ y_1 \otimes \cdots \otimes y_k \rangle \\ &= \frac{1}{k!} \sum_{\sigma \in \mathbf{S}_k} \operatorname{sgn}(\sigma) \ \big\langle x_1 \otimes \cdots \otimes x_k \ , \ y_{\sigma(1)} \otimes \cdots \otimes y_{\sigma(k)} \big\rangle \\ &= \frac{1}{k!} \sum_{\sigma \in \mathbf{S}_k} \operatorname{sgn}(\sigma) \ \prod_{i=1}^{k} \big\langle x_i, y_{\sigma(i)} \big\rangle = \frac{1}{k!} \ \det \ \big(\langle x_i, y_j \rangle \big)_{i,j \in \mathbb{I}_k} \ , \end{aligned}$$

donde la segunda igualdad surge de que $\mathbf{P}_k^n$ es un proyector autoadjunto. ■

Potencia alternada de una matriz

Observación 7.3.9. Por las ecuaciones (7.7) y (7.10), si $A \in L(\mathbb{H}_m \, , \ \mathbb{H}_n)$, su k-potencia tensorial "mezcla" $P_\pi^{(n)}$ y $P_\pi^{(m)}$ en el siguiente sentido:

$$P_\pi^{(n)} \left(\bigotimes^k A \right) = \left(\bigotimes^k A \right) P_\pi^{(m)} \quad \text{para toda} \quad \pi \in \mathbf{S}_k \ .$$

Entonces, por la fórmula (7.11), $\bigotimes^k A$ mezcla las proyecciones $\mathbf{P}_k^n$ y $\mathbf{P}_k^m$:

$$\mathbf{P}_k^n \left(\bigotimes^k A \right) = \left(\bigotimes^k A \right) \mathbf{P}_k^m . \tag{7.15}$$

Por lo tanto $\bigotimes^k A\big(\Lambda^k \mathbb{H}_m\big) \subseteq \Lambda^k \mathbb{H}_n$. ▲

Definición 7.3.10. Sea $A \in L(\mathbb{H}_n \, , \mathbb{H}_m)$. La restricción de $\bigotimes^k A$ al espacio alternado $\Lambda^k \mathbb{H}_n$ es llamada la *k-potencia exterior* de A, y denotada por

$$\Lambda^k A \in L(\Lambda^k \mathbb{H}_n \, , \ \Lambda^k \mathbb{H}_m) \ .$$

Por la Eq. (7.7) y la Observación 7.3.5, la k-potencia exterior $\Lambda^k A$ está determinada por la fórmula:

$$\Lambda^k A \ (x_1 \wedge x_2 \wedge \cdots \wedge x_k) = A x_1 \wedge A x_2 \wedge \cdots \wedge A x_k , \tag{7.16}$$

para toda k-upla $x_1, \dots, x_k$ en $\mathbb{H}_n$. ▲

Observación 7.3.11. Si I_n es la identidad de $\mathbb{H}_n$, entonces $\Lambda^k I_n = I_{\Lambda^k \mathbb{H}_n}$. Por otra parte, se sigue de (7.8) o de (7.16) que, si $A \in L(\mathbb{H}_n, \mathbb{H}_m)$ y $B \in L(\mathbb{H}_m, \mathbb{H}_r)$, entonces

$$\Lambda^k (AB) = \Lambda^k A \cdot \Lambda^k B \quad \text{y} \quad (\Lambda^k A)^* = \Lambda^k A^* \ . \tag{7.17}$$

Cuando $n = m$, i.e. $A \in L(\mathbb{H}_n)$, la Eq. (7.15) dice que $\Lambda^k \mathbb{H}_n$ reduce a $\bigotimes^k A$, por lo que se tiene una identidad matricial del tipo $\bigotimes^k A = \begin{array}{c} \begin{bmatrix} \Lambda^k A & 0 \\ 0 & * \end{bmatrix} \end{array} \begin{array}{c} \Lambda^k \mathbb{H}_n \\ \left(\Lambda^k \mathbb{H}_n\right)^\perp \end{array}$. De ahi se deducen fácilmente las siguientes propiedades:

a. Si $A \in \mathcal{G}l(n)$, $\Lambda^k A^{-1} = (\Lambda^k A)^{-1}$.

b. $\Lambda^k A$ es unitaria si $A \in \mathcal{U}(n)$.

c. $\Lambda^k A \geq 0$ si $A \geq 0$. Además $|\Lambda^k A| = \Lambda^k |A|$. ▲

Definición 7.3.12. 1. Sea $n \in \mathbb{N}$ y $k \in \mathbb{I}_n$. Notamos por $Q_{k,n}$ al conjunto de sucesiones estrictamente crecientes de k enteros elegidos en $\mathbb{I}_n$:

$$Q_{k,n} = \left\{ \alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k) \in \mathbb{I}_n^k \ : \ 1 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_k \leq n \ \right\} .$$

Otra manera de verlo es $Q_{k,n} = \left\{ J \subseteq \mathbb{I}_n : |J| = k \right\}$, si pensamos a los conjuntos J ordenados en forma creciente. Luego $|Q_{k,n}| = \binom{n}{k}$.

2. Sean $A \in \mathcal{M}_{n,m}(\mathbb{C})$, $\alpha \in Q_{k,n}$ y $\beta \in Q_{l,m}$. Entonces denotaremos por $A[\alpha|\beta]$ a la submatriz de $k \times l$ de A dada por

$$A[\alpha|\beta] = \left(A_{\alpha_i \beta_j} \right)_{(i,j) \in \mathbb{I}_k \times \mathbb{I}_l} \in \mathcal{M}_{k,l}(\mathbb{C}) \ .$$

Cuando $\alpha = \beta$, $A[\alpha|\beta]$ se abreviará como $A[\alpha]$. Si $\alpha = \mathbb{I}_n$ (resp. $\beta = \mathbb{I}_m$), notaremos $A[\alpha|\beta] = A[-|\beta]$ (resp. $A[\alpha|\beta] = A[\alpha|-]$).

3. Dada $\alpha \in Q_{k,n}$, usaremos la abreviación:

$$\mathbf{e}_\alpha^\wedge \ = \ \mathbf{e}_\alpha^{(n)\,\wedge} := e_{\alpha_1}^{(n)} \wedge e_{\alpha_2}^{(n)} \wedge \cdots \wedge e_{\alpha_k}^{(n)} \in \Lambda^k \mathbb{H}_n \ .$$

A continuación veremos que forman una BON de $\Lambda^k \mathbb{H}_n$. ▲

Proposición 7.3.13. *El conjunto*

$$\mathcal{E}_{k,n}^\wedge = \{ \sqrt{k!} \ \mathbf{e}_\alpha^\wedge : \alpha \in Q_{k,n} \} \tag{7.18}$$

es una BON de $\Lambda^k \mathbb{H}_n$. *Por lo tanto, tenemos que* $\dim \Lambda^k \mathbb{H}_n = |Q_{k,n}| = \binom{n}{k}$.

Demostración. El hecho de que $\mathcal{E}_{k,n}^\wedge$ genera $\Lambda^k \mathbb{H}_n$ se deduce de los ítems 1, 2 y 3 de la Observación 7.3.5 (la Eq. (7.12) permite ordenar las coordenadas). Por otro lado, si $\alpha, \beta \in Q_{k,n}$ no son iguales, es fácil ver que la matriz $\left(\left\langle e_{\alpha_i}, e_{\beta_j} \right\rangle \right)_{i,j \in \mathbb{I}_k}$ debe tener una fila nula (la de algún $\alpha_i \notin \beta$). Luego, por la Proposición 7.3.8 y el ítem 4 de 7.3.7, se tiene que $\langle \mathbf{e}_\alpha^\wedge, \mathbf{e}_\beta^\wedge \rangle = 0$. Finalmente, como $\det I_k = 1$, llegamos a que $\mathcal{E}_{k,n}^\wedge$ es una BON. ■

Proposición 7.3.14. *Sea* $A \in \mathcal{M}_{n,m}(\mathbb{C})$. *Identifiquemos* $\Lambda^k A \in L(\Lambda^k \mathbb{H}_m, \Lambda^k \mathbb{H}_n)$ *con su matriz en las bases* $\mathcal{E}^{\wedge}_{k,m}$ *y* $\mathcal{E}^{\wedge}_{k,n}$. *Dados* $\alpha \in Q_{k,n}$ *y* $\beta \in Q_{k,m}$, *se tiene que*

$$\left(\Lambda^k A\right)_{\alpha,\beta} = \det A[\alpha|\beta] \ . \tag{7.19}$$

Demostración. De acuerdo a las ecuaciones (7.14) y (7.18), se tiene que

$$\begin{aligned} \left(\Lambda^k A\right)_{\alpha,\beta} &= \left\langle \Lambda^k A \sqrt{k!}\ \mathbf{e}^{(m)}_\beta , \sqrt{k!}\ \mathbf{e}^{(n)}_\alpha \right\rangle = k! \left\langle \Lambda^k A\ \mathbf{e}^{(m)}_\beta , \mathbf{e}^{(n)}_\beta \right\rangle \\ &= \det \left[\left\langle A e_{\beta_j}, e_{\alpha_i} \right\rangle \right]_{i,j \in \mathbb{I}_k} = \det A[\alpha|\beta] \ , \end{aligned}$$

donde la última igualdad se sigue de que $\langle A e_{\beta_j}, e_{\alpha_i} \rangle = A_{\alpha_i \beta_j}$. ■

Determinantes

Miremos qué es $\Lambda^n \mathbb{H}_n$, o sea el caso $k = n$. Como $\mathbb{I}_n$ es el único elemento de $Q_{n,n}$, la Proposición 7.3.13 asegura que el vector $\mathbf{e}_n = \sqrt{n!}\ \mathbf{e}^{\wedge}_{\mathbb{I}_n} = \sqrt{n!}\ e_1 \wedge e_2 \wedge \cdots \wedge e_n$ es una BON de $\Lambda^n \mathbb{H}_n$. O sea que $\Lambda^n \mathbb{H}_n = \mathbb{C} \cdot \mathbf{e}_n$. Dados $x_1, \dots, x_n \in \mathbb{H}_n$, si tomamos la matriz $X \in \mathcal{M}_n(\mathbb{C})$ dada por $F_i(X) = x_i$, queda

$$\sqrt{n!}\ \ x_1 \wedge x_2 \wedge \cdots \wedge x_n = \det X \cdot \mathbf{e}_n \ . \tag{7.20}$$

En efecto, si abreviamos $a = x_1 \wedge \cdots \wedge x_n$, entonces $a = \langle a, \mathbf{e}_n \rangle \cdot \mathbf{e}_n$. Ahí podemos aplicar la Proposición 7.3.8, ya que $\langle x_i, e_j \rangle = X_{ij}$, para todo par $i, j \in \mathbb{I}_n$.

La fórmula (7.20) es la segunda definición para el $\det X$, en el sentido de que $X \mapsto \det X$ es la única función n-multilineal alternada (en las filas de X) tal que $\det I_n = 1$ (esto vale porque la matriz X asociada a $e_1 \wedge \cdots \wedge e_n$ es $X = I_n$). La Proposición 7.3.8 muestra que es equivalente a la de diagonales y permutaciones.

Si en la Proposición 7.3.14 consideramos el caso $k = n = m$, e identificamos $L(\Lambda^n \mathbb{H}_n)$ con $\mathbb{C}$ (vía $zI \mapsto z$), tenemos que

$$\Lambda^n A = \det A \quad \text{ para toda } \quad A \in \mathcal{M}_n(\mathbb{C}) \ . \tag{7.21}$$

Observar que esto brinda un camino directo para probar la igualdad $\det AB = \det A \cdot \det B$, que no se ve tan fácil vía la Definición 7.3.6.

Proposición 7.3.15. *Sean* $A, B \in \mathcal{M}_n(\mathbb{C})$, *Entonces se tiene que*

1. $\det AB = \det A \cdot \det B$.
2. $\det A \neq 0$ *si y sólo si* $A \in \mathcal{G}l(n)$.
3. $\mathcal{G}l(n)$ *es abierto y denso en* $\mathcal{M}_n(\mathbb{C})$.

Demostración. Lo primero se deduce de que $\Lambda^n AB = \Lambda^n A \cdot \Lambda^n B$ vía la fórmula (7.21). Si $A \in \mathcal{G}l(n)$, tenemos que $\det A \cdot \det A^{-1} = \det I_n = 1 \neq 0$. Si $A \notin \mathcal{G}l(n)$, entonces sus columnas deben ser un conjunto linealmente dependiente (porque $\ker A \neq \{0\}$). Luego se aplica la Eq. (7.20) y el último punto de la Observación 7.3.5 a la matriz A^T, y el hecho de que $\det A = \det A^T$, como asegura 7.3.7. Como $A \mapsto \det A$ es continua, el item 2 implica que $\mathcal{G}l(n) = \det^{-1}\{z \in \mathbb{C} : z \neq 0\}$ es abierto en $\mathcal{M}_n(\mathbb{C})$. La densidad podría probarse usando la multilinealidad de $A \mapsto \det A$, pero sale más fácil viendo que, para cualqueir $A \in \mathcal{M}_n(\mathbb{C})$, existen matrices $A + \varepsilon I \in \mathcal{G}l(n)$ para ε arbitrariamente pequeño. ■

Corolario 7.3.16. *Sean k, $n \in \mathbb{N}$.*

1. *Un conjunto $\{x_1, x_2, \dots, x_k\} \subseteq \mathbb{H}_n$ es linealmente independiente si y sólo si el producto alternado $x_1 \wedge x_2 \wedge \cdots \wedge x_k \neq 0$.*
2. *El espacio $\Lambda^k \mathbb{H}_n \neq \{0\}$ si y sólo si $k \leq n$.*

Demostración. Sea $X \in \mathcal{M}_{n,k}$ dada por $C_i(X) = x_i$, $i \in \mathbb{I}_k$. Luego $\{x_1, x_2, \dots, x_k\}$ es linealmente independiente si y sólo si $\ker X = \{0\}$. Esto, a su ves, equivale a que $X^*X \in \mathcal{G}l(k)$ (porque $\ker X^*X = \ker X$). Pero, por la Proposición 7.3.8, tenemos que

$$X^*X = \big(\langle x_j, x_i \rangle \big)_{i,j \in \mathbb{I}_k} = \big(\langle x_i, x_j \rangle \big)_{i,j \in \mathbb{I}_k} \implies \det X^*X = k!\, \|x_1 \wedge x_2 \wedge \cdots \wedge x_k\|^2 .$$

Luego aplicamos la Proposición 7.3.15. La segunda parte se deduce inmediatamente de la primera, porque en $\mathbb{H}_n$ puede haber, a lo sumo, n vectores linealmente independientes. ■

Observación 7.3.17. Recordemos que, si $A \in \mathcal{M}_{n,m}(\mathbb{C})$ decimos que

$$\mathrm{rk} A = \dim R(A) = \dim \mathrm{Gen}\,\{C_1(A), \dots, C_m(A)\} ,$$

es el rango columna de A. Como $\Lambda^k A\,(\mathbf{e}_\alpha^\wedge) = C_{\alpha_1}(A) \wedge \cdots \wedge C_{\alpha_k}(A)$, para todo $\alpha \in Q_{k,m}$, el Corolario 7.3.16 muestra que

$$\mathrm{rk}\ A = \max\{k \in \mathbb{N} \ : \ \Lambda^k A \neq 0\} \tag{7.22}$$

(ver también el Corolario 7.4.3 de más adelante). Usando que $\Lambda^k A^* = (\Lambda^k A)^*$, la fórmula (7.22) da otra prueba de que $\mathrm{rk}\, A^* = \mathrm{rk}\, A$. ▲

El siguiente resultado generaliza la Proposición 7.3.15 a determinantes de submatrices:

Corolario 7.3.18 (Fórmula de Cauchy-Binnet)**.** *Dadas $A \in \mathcal{M}_{n,r}(\mathbb{C})$ y $B \in \mathcal{M}_{r,m}(\mathbb{C})$, sea $k \leq \min\{n, r, m\}$. Luego, para cada par $\alpha \in Q_{k,n}$, $\beta \in Q_{k,m}$ se tiene que*

$$\det(AB)[\alpha|\beta] = \sum_{\omega \in Q_{k,r}} \det A[\alpha|\omega] \det B[\omega|\beta] . \tag{7.23}$$

Demostración. Por la ley de multiplicación (7.17) y la Proposición 7.3.14, tenemos que

$$\begin{aligned} \det(AB)[\alpha|\beta] &= (\Lambda^k AB)_{\alpha\beta} = \left(\Lambda^k A \cdot \Lambda^k B\right)_{\alpha\beta} \\ &= \sum_{\omega \in Q_{k,r}} \left(\Lambda^k A\right)_{\alpha\omega} \left(\Lambda^k B\right)_{\omega\beta} \\ &= \sum_{\omega \in Q_{k,r}} \det A[\alpha|\omega] \det B[\omega|\beta] \ , \end{aligned}$$

lo que prueba lo afirmado. ■

Observación 7.3.19. La versión más clásica de la Fórmula de Cauchy Binnet es la siguiente: Sean $A \in \mathcal{M}_{k,r}(\mathbb{C})$ y $B \in \mathcal{M}_{r,k}(\mathbb{C})$, con $k \le r$. Luego,

$$\det AB = \sum_{\omega \in Q_{k,r}} \det A[-|\omega] \det B[\omega|-] \ , \tag{7.24}$$

que resulta de operar con todas las submatrices cuadradas de tamaño máximo de A (eligiendo columnas) y de B (eligiendo las mismas filas). Es claro que (7.24) se deduce del Corolario 7.3.18. También vale la recíproca, porque dadas A y B como en el Corolario 7.3.18, las matrices $A_0 = A[\alpha, -] \in \mathcal{M}_{k,r}(\mathbb{C})$ y $B_0 = B[-, \beta] \in \mathcal{M}_{r,k}(\mathbb{C})$ cumplen que

$$A_0 B_0 = (AB)[\alpha|\beta] \ , \quad A_0[-|\omega] = A[\alpha|\omega] \quad \text{y} \quad B_0[\omega|-] = B[\omega|\beta] \ , \quad \omega \in Q_{k,r} \ ;$$

por lo que (7.23) para A y B se reduce a (7.24) para A_0 y B_0. ▲

Proposición 7.3.20. *Sean $A, B \in \mathcal{G}l(n)^+$ y $\lambda \in [0, 1]$. Entonces*

$$\det\big(\lambda A + (1-\lambda)B\big) \ge (\det A)^\lambda \, (\det B)^{1-\lambda} \ .$$

Es decir, la aplicación $\mathcal{G}l(n)^+ \ni A \mapsto \log\det A$ es cóncava.

Demostración. Sea $C = B^{-1}A$. Como $\sigma(C) = \sigma\big(B^{-1/2}AB^{-1/2}\big)$ (con multiplicidades), podemos llamar $\mu(C) = \mu(B^{-1/2}AB^{-1/2}) \in \mathbb{R}_+^{*\,n}$. Además,

$$\det\big(\lambda A + (1-\lambda)B\big) = \det B\,\big(\lambda B^{-1}A + (1-\lambda)I\big) = \det B \, \det\big(\lambda C + (1-\lambda)I\big).$$

Luego basta probar que

$$\det\big(\lambda C + (1-\lambda)I\big) \ge (\det A)^\lambda \, (\det B)^{1-\lambda} \, \det B^{-1} = (\det A)^\lambda (\det B)^{-\lambda} = (\det C)^\lambda \ .$$

En otras palabras, basta ver que

$$\prod_{i=1}^n \big(\lambda\mu_i(C) + 1 - \lambda\big) \ge \prod_{i=1}^n \mu_i(C)^\lambda \ .$$

Finalmente, veremos que $\big(\lambda\mu_i(C) + 1 - \lambda\big) \ge \mu_i(C)^\lambda$ para cada $i \in \mathbb{I}_n$, con lo cual el resultado quedaría probado. En efecto, dado $c > 0$, la función $f(t) = c^t$ es convexa en todo $\mathbb{R}$. Notar que $f(0) = 1$ y $f(1) = c$. Por lo tanto

$$\big(\lambda c + 1 - \lambda\big) = \lambda f(1) + (1-\lambda) f(0) \ge f(\lambda 1 + (1-\lambda)0) = f(\lambda) = c^\lambda \ .$$

Aplicando lo anterior a cada $c = \mu_i(C)$, obtenemos el resultado. ■

Ejercicio 7.3.21. 1. Sea $H \in \mathcal{G}l(n)^+$ (en principio real). Entonces,

$$\sqrt{\frac{\pi^n}{\det H}} \;=\; \int_{\mathbb{R}^n} e^{-\langle Hx,x\rangle}\; dx \ .$$

Para probarlo, hacer un cambio de variables $y = Ux$, para $U \in \mathcal{U}(n)$ tal que $UHU^* = \operatorname{diag}(\mu(H))$. Como U es unitaria, la integral no cambia. Luego usar que $\int_{\mathbb{R}^n} e^{-at^2}\, dt = \pi^{1/2}a^{-1/2}$. De paso, esto prueba que $e^{-\langle Hx,x\rangle}$ es integrable en $\mathbb{R}^n$ (notar que el cambio de variables manda bolas en bolas). ¿Vale lo mismo para matrices complejas?

2. Probar la Proposición 7.3.20 usando lo anterior (y la desigualdad de Hölder !).

7.4 Propiedades útiles de los productos alternados

El siguiente resultado, si bien es algo técnico, es la llave para la caracterización completa de los autovalores de un producto alternado:

Lema 7.4.1. *Sea $T \in \mathcal{TS}(n)$, con los números $\lambda_1, \dots, \lambda_n$ en su diagonal. Sea $k \in \mathbb{I}_n$.*

1. *En términos de la BON $\mathcal{E}^{\wedge}_{k,n}$ de la Eq.* (7.18), *ordenada lexicográficamente, la matriz de $\Lambda^k T$ es, también, triangular superior.*

2. *Para cada $J \in Q_{k,n}$, se tiene que* $(\Lambda^k T)_{JJ} \;=\; \prod\limits_{i\in J} \lambda_i$.

Demostración. Sean $I, J \in Q_{k,n}$ tales que $I > J$. Debemos probar que

$$(\Lambda^k T)_{IJ} = \det T[I,J] = 0 \ ,$$

donde la primea igualdad sabemos que es cierta por la Proposición 7.3.14. Si $I = (\alpha_1, \cdots, \alpha_k)$ y $J = (\beta_1, \dots, \beta_k)$ (vectores ordenados en forma *creciente*), debe existir algún $j \in \mathbb{I}_k$ tal que $\alpha_j > \beta_j$ (sino valdría que $I \le J$ en el lexicográfico). Por lo tanto,

$$\alpha_i > \beta_r \quad \text{para todo par } (i,r) \text{ tal que} \quad 1 \le r \le j \le i \le k \ .$$

Como $T \in \mathcal{TS}(n)$, tenemos que $T_{\alpha_i\beta_r} = 0$ para todos esos pares. Es decir que $T[I,J]$ tiene una submatriz nula de tamaño $(k-j+1) \times j$. Aplicando König-Frobenius (Corolario 4.3.3), deducimos que $T[I,J]$ no tiene ninguna diagonal sin ceros. Esto implica que $\det T[I,J] = 0$, como se afirmó. Por otra parte, si $J \in Q_{k,n}$, por la Proposición 7.3.14 sabemos que

$$(\Lambda^k T)_{JJ} = \det T[J] = \prod_{i\in J} \lambda_i \ ,$$

puesto que $T[J] \in \mathcal{TS}(k)$. ■

Teorema 7.4.2. *Sea* $A \in \mathcal{M}_n(\mathbb{C})$ *con vector de autovalores* $\lambda(A) = (\lambda_1(A), \dots, \lambda_n(A))$. *Sea* $k \in \mathbb{I}_n$. *Entonces los autovalores de* $\Lambda^k A$ *están dados por*

$$\lambda_J(\Lambda^k A) = \prod_{i \in J} \lambda_i(A) \ , \quad J \in Q_{k,n} \ ,$$

contados con multiplicidad.

Demostración. Es similar al caso de los productos tensoriales (Proposición 7.1.2). Se aplica el Teorema 1 de Schur 1.6.1 y el hecho de que $\Lambda^k U$ es unitaria si $U \in \mathcal{U}(n)$, pero usando ahora el Lema 7.4.1. ■

Corolario 7.4.3. *Sea* $A \in \mathcal{M}_n(\mathbb{C})$. *Sea* $k \in \mathbb{I}_n$. *Entonces los valores singulares de* $\Lambda^k A$ *son*

$$s\big(\Lambda^k A\big) = \big\{ \ s_J\big(\Lambda^k A\big) \ \big\}_{J \in Q_{k,n}} = \Big\{ \ \prod_{i \in J} s_i(A) \ \Big\}_{J \in Q_{k,n}} \ ,$$

contados con multiplicidad, y ordenados en forma decreciente. Además, si ordenamos a los autovalores $\lambda_1(A), \dots, \lambda_n(A)$ *de* A *con módulos decrecientes, se tiene que*

$$\rho\big(\Lambda^k A\big) = \prod_{i=1}^{k} |\lambda_i(A)| \qquad y \qquad \big\|\Lambda^k A\big\|_{sp} = s_1\big(\Lambda^k A\big) = \prod_{i=1}^{k} s_i(A) \ .$$

Demostración. Se deduce del Teorema 7.4.2 y del hecho de que $\big|\Lambda^k A\big| = \Lambda^k |A|$. ■

A continuación veremos algunas propiedades funtoriales de los productos alternados, que serán necesarias para las aplicaciones a desigualdades.

Proposición 7.4.4. *Sea* $A \in \mathcal{M}_n(\mathbb{C})$ *y* $k \in \mathbb{I}_n$.

1. *Si* $A_m \xrightarrow[m \to \infty]{} A$, *entonces* $\Lambda^k A_m \xrightarrow[m \to \infty]{} \Lambda^k A$.
2. *Si* $A \in \mathcal{M}_n(\mathbb{C})^+$ *entonces, para todo* $r \in \mathbb{R}_+$ *se tiene que*

$$(\Lambda^k A)^r = \Lambda^k (A^r) \ .$$

3. *Si* A *es alguna de estas cosas:*
   a. *Idempotente (i.e.,* $A^2 = A$*),*
   b. *Proyector (i.e.,* $A^2 = A = A^*$*),*
   c. *Isometría parcial (i.e.,* AA^* *y* A^*A *son proyectores),*
   d. *Autoadjunto, normal o unitario,*

   entonces $\Lambda^k A$ *es del mismo tipo.*

4. *Si $A = U|A|$ es una DP de A, entonces*

$$\Lambda^k A = \Lambda^k U \ \Lambda^k |A| \tag{7.25}$$

es una descomposición polar de $\Lambda^k A$.

Demostración.

1. Por la fórmula (7.19), para todo par $\alpha, \beta \in Q_{k,n}$, tenemos que

$$(\Lambda^k A_m)_{\alpha\beta} = \det A_m[\alpha|\beta] \xrightarrow[m\to\infty]{} \det A[\alpha|\beta] = (\Lambda^k A)_{\alpha\beta} \ .$$

Observar que el determinante de una matriz es un polinomio en sus entradas, por lo que la función $B \mapsto \det B$ es continua.

2. Como $(\Lambda^k A)^2 = \Lambda^k(A^2)$, la veracidad del enunciado cuando $r \in \mathbb{N}$ se deduce a través de una simple inducción. Recordando que $(\Lambda^k A)^{-1} = \Lambda^k(A^{-1})$, es claro que también vale si $r \in \mathbb{Z}$. Para extender el resultado a los $r \in \mathbb{Q}$, basta notar que si $m \in \mathbb{N} - \{0\}$ entonces:

$$(\Lambda^k A^{1/m})^m = \Lambda^k[(A^{1/m})^m] = \Lambda^k A \ .$$

Finalmente, el caso general se obtiene por continuidad (y el item 1.).

3. Todas estas propiedades se deducen directamente de las propiedades vistas en la Observación 7.3.11.

4. Ya hemos visto (o podemos deducir de lo anterior) que $|\Lambda^k A| = \Lambda^k |A|$. Como $\Lambda^k U$ es isometría parcial, y la igualdad (7.25) se tiene que cumplir a partir de que $A = U|A|$, entonces (7.25) es una DP de $\Lambda^k A$. ∎

7.5 Ejercicios

Ejercicios del texto

7.5.1. Probar que el producto tensorial de matrices verifica las siguientes propiedades: Fijemos $A \in \mathcal{M}_n(\mathbb{C})$ y $B \in \mathcal{M}_k(\mathbb{C})$. Se considera que $A \otimes B \in L(\mathbb{H}_n \otimes \mathbb{H}_k) \cong \mathcal{M}_{nm}(\mathbb{C})$, y en $\mathbb{H}_n \otimes \mathbb{H}_k$ se usa el producto escalar definido en las Eqs. (7.1) y (7.3).

1. Sean $I_n \in \mathcal{M}_n(\mathbb{C})$ y $I_k \in \mathcal{M}_k(\mathbb{C})$. Entonces $I_n \otimes I_k$ es la identidad de $\mathbb{H}_n \otimes \mathbb{H}_k$.

2. $(\alpha A_1 + A_2) \otimes B = \alpha(A_1 \otimes B) + A_2 \otimes B$, para todo $\alpha \in \mathbb{C}$.

3. $(A \otimes B)^* = A^* \otimes B^*$.

4. $(A_1 \otimes B_1)(A_2 \otimes B_2) = A_1 A_2 \otimes B_1 B_2$.

5. Si existen A^{-1} y B^{-1}, entonces $A^{-1} \otimes B^{-1} = (A \otimes B)^{-1}$. En particular, si $A \in \mathcal{U}(n)$ y $B \in \mathcal{U}(k)$, entonces $A \otimes B \in \mathcal{U}(nk)$.

6. $A \otimes B \geq 0$ si $A \geq 0$ y $B \geq 0$. Más aún, $|A \otimes B| = |A| \otimes |B|$. Se usa el Teorema 3.1.3 y la unicidad de la raiz cuadrada positiva.

7.5.2. Completar los detalles de la prueba de la Eq. (7.4).

7.5.3. Completar los detalles de la definición inductiva del espacio $\bigotimes^k \mathbb{H}_n$, probar los 5 items de 7.2.1 y los 6 de 7.2.2.

7.5.4. Dados $n, k \in \mathbb{N}$ y $\pi \in \mathbf{S}_k$, tomemos el operador de permutación $P_\pi^{(n)} \in L(\bigotimes^k \mathbb{H}_n)$, definido en la Eq. (7.9). Probar las siguentes propiedades:

1. La fórmula (7.10) sobre como actúa $P_\pi^{(n)}$ en los tensores elementales.

2. Mostrar que $P_\pi^{(n)}$ es untario. Más aún, mostrar que $(P_\pi^{(n)})^* = P_{\pi^{-1}}^{(n)} = (P_\pi^{(n)})^{-1}$.

3. Sea $\mathbf{P}_k^n$ la proyección ortogonal de $\bigotimes^k \mathbb{H}_n$ sobre $\Lambda^k \mathbb{H}_n$. Completar los detalles de la demostración de la Eq. (7.11):

$$\mathbf{P}_k^n = \frac{1}{k!} \sum_{\pi \in \ \mathbf{S}_k} \operatorname{sgn}(\pi) \ P_\pi^{(n)} \ .$$

7.5.5. Dar los detalles de las pruebas de los 5 items de la Observación 7.3.5, sobre las propiedades de los k-tensores elementales.

7.5.6. Probar todos los resultados enunciados en la Observación 7.3.11, sobre las propiedades de las k-potencias alternadas (o exteriores) de matrices.

7.5.7. Probar los 7 items de 7.3.7 (sobre determinantes).

7.5.8. Probar que $\mathcal{G}l(n)$ es denso en $\mathcal{M}_n(\mathbb{C})$.

7.5.9. Hacer el Ejercicio 7.3.21.

Ejercicios nuevos

7.5.10. Sea $\sigma \in \mathbf{S}_n$ y $P_\sigma \in \mathcal{U}_{\mathcal{P}}(n)$ su matriz de permutación asociada, definida en la Observación 4.1.5. Probar que $\det P_\sigma = \operatorname{sgn}(\sigma)$ de las tres maneras propuestas:

1. Usando la Definición 7.3.6 de una (esto es parte del Ejercicio anterior).

2. Mostrar que si $\tau \in \mathbf{S}_n$ es una trasposición, entonces $\det P_\tau = -1$, y usar que sigma es producto de trasposiciones, y que la flecha $\sigma \mapsto \det P_\sigma$ es un morfismo.

3. Usando que P_σ tiene filas $F_i(P_\sigma) = e_{\sigma(i)}$, $i \in \mathbb{I}_n$ (por la Eq. (4.3)), y aplicando luego las ecuaciones (7.12) y (7.20).

4. Alguna otra que se les ocurra.

7.5.11. Demostrar el algoritmo usual para calcular $\det A$, para $A \in \mathcal{M}_n(\mathbb{C})$, desarrollando por alguna fila o columna de A. Por ejemplo la fila r-ésima:

$$\det A = \sum_{i \in \mathbb{I}_n} (-1)^{r+i} \; A_{r,\,i} \; \det A(r|\,i) \ . \tag{7.26}$$

Se sugiere usar la multilinealidad de $B \mapsto \det B$ (tanto para filas como para columnas) y calcular $\det B$ en el caso de que alguna fila o columna de B esté en la base canónica de $\mathbb{C}^n$. Otra opción es esperar hasta la Eq. (12.13).

7.5.12. Sea $t = (t_1, \dots, t_n) \in \mathbb{C}^n$. Se llama *matriz de Vandermonde* de t a

$$V(t) = \left(t_i^{j-1}\right)_{i,j \in \mathbb{I}_n} = \begin{bmatrix} 1 & t_1 & \dots & t_1^{n-1} \\ 1 & t_2 & \dots & t_2^{n-1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & t_n & \dots & t_n^{n-1} \end{bmatrix} \in \mathcal{M}_n(\mathbb{C}) \ .$$

Probar que $\det V(t) \ = \ \prod\limits_{i<j} (t_j - t_i)$, por lo que $V(t) \in \mathcal{G}l\,(n)$ si los t_i son todos distintos.

7.5.13. Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Probar:

1. $(s(A), s(B)) \prec (s(|A| + |B|), 0)$ en $\mathbb{R}^{2n}$.

2. $(s(A), s(B)) \prec^w (s(A + B), 0)$ en $\mathbb{R}^{2n}$.

3. Sea $F : \mathcal{M}_n(\mathbb{C}) \to \mathbb{R}_+$ dada por $F(C) = \sum_i f(s_i(C))$, para una $f : \mathbb{R}_+ \to \mathbb{R}_+$ cóncava tal que $f(0) = 0$. Probar que F es subaditiva, o sea
$$F(A + B) \le F(A) + F(B).$$

4. Si $z \in \mathbb{C}$, $\det(I + zA) = \sum\limits_{k=0}^{n} z^k \operatorname{tr}(\Lambda^k A)$.

5. $P_A(x) = x^n + \sum\limits_{k=0}^{n-1} (-1)^{n-k} \operatorname{tr}(\Lambda^{n-k} A) \ x^k$.

6. $\det(I + |A + B|) \le \det(I + |A|) \det(I + |B|)$.

7. $|\det(I + A)| \le \det(I + |A|)$.

8. $|\det(I + A + B)| \le \det(I + |A|) \det(I + |B|)$.

Productos simétricos

Sean $k, n \in \mathbb{N}$. Recordemos que, dados $x_1, \dots, x_k \in \mathbb{H}_n$ y $\pi \in \mathbf{S}_k$, tenemos que

$$P_\pi^{(n)}(x_1 \otimes x_2 \cdots \otimes x_k) = x_{\pi^{-1}(1)} \otimes x_{\pi^{-1}(2)} \cdots \otimes x_{\pi^{-1}(k)} \ ,$$

donde $P_\pi^{(n)}$ es un operador unitario que cumple $(P_{\pi^{-1}}^{(n)}) = (P_\pi^{(n)})^{-1}$. Dado $k \in \mathbb{I}_n$. Llamaremos espacio *k-simétrico* sobre $\mathbb{H}_n$, al subespacio de $\bigotimes^k \mathbb{H}_n$ dado por

$$\vee^k \mathbb{H}_n = \Big\{ \ F \in \bigotimes^k \mathbb{H}_n \ : \ P_\pi^{(n)} F = F \quad \text{para toda} \quad \pi \in \mathbf{S}_k \ \Big\} \ ,$$

Los elementos de $\vee^k \mathbb{H}_n$ se llaman k-*tensores simétricos*. Se considera a $\vee^k \mathbb{H}_n$ como un espacio de Hilbert con el producto interno de $\bigotimes^k \mathbb{H}_n$. Dados $x_1, \dots, x_k \in \mathbb{H}_n$, se define el k-tensor simétrico elemental:

$$x_1 \vee x_2 \cdots \vee x_k := \frac{1}{k!} \sum_{\pi \in \ \mathbf{S}_k} x_{\pi(1)} \otimes x_{\pi(2)} \cdots \otimes x_{\pi(k)} \ \in \ \vee^k \mathbb{H}_n \ .$$

7.5.14. Dada $A \in \mathcal{M}_n(\mathbb{C})$, definimos su **permanente** por la fórmula

$$\operatorname{per} A \ = \sum_{\sigma \in \mathbf{S}_n} \prod_{j=1}^n a_{j,\sigma(j)} \in \mathbb{C} \ . \tag{7.27}$$

Es decir que es como el determinante, pero sin signos negativos.

1. Probar que, si $T \in \mathcal{TS}(n)$, entonces $\operatorname{per} T = \prod_{i \in \mathbb{I}_n} T_{ii}$. En particular, $\operatorname{per} I_n = 1$.
2. Si $A \geqslant 0$, mostrar que $\operatorname{per} A = 0 \iff$ existen 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 $(i,j) \in I \times J$.
3. Deducir que si $A \in \mathcal{DS}(n)$, entonces $\operatorname{per} A \neq 0$.
4. Si $B, C \in \mathcal{M}_n(\mathbb{C})^+$ cumplen que $C \leq B$, probar que $0 \leq \operatorname{per} C \leq \operatorname{per} B$.

7.5.15. Sean $x_1, \dots, x_k, y_1, \dots, y_k \in \mathbb{H}_n$. Probar que

$$\big\langle x_1 \vee x_2 \vee \cdots \vee x_k \ , \ y_1 \vee y_2 \vee \cdots \vee y_k \big\rangle = \frac{1}{k!} \operatorname{per} \big(\langle x_i, y_j \rangle \big)_{i,j \in \mathbb{I}_k} \ .$$

7.5.16. Dados $x_1, \dots, x_k, y_1, \dots, y_k \in \mathbb{H}_n$, llamemos

$$G(x,y) \ = \ \big(\langle x_i, y_j \rangle \big)_{i,j \in \mathbb{I}_k} \ \in \mathcal{M}_k(\mathbb{C})$$

a la matriz que se usó en el Ejercicio anterior.

1. Probar que $|\det G(x,y)|^2 \leq \det G(x,x) \, \det G(y,y)$

2. También que $|\text{per } G(x,y)|^2 \leq \text{per } G(x,x) \, \text{per } G(y,y)$.

3. Traducir a que si $A, B \in \mathcal{M}_{n,k}(\mathbb{C})$, entonces

$$|\text{per } A^*B|^2 \leq \text{per } A^*A \text{ per } B^*B$$

4. (Otro teorema de Schur) Si $A \in \mathcal{M}_n(\mathbb{C})^+$, per $A \geq \det A$ (se sugiere usar el teorema de Cholewsky, Corolario 3.1.5).

7.5.17. Sea $A \in \mathcal{G}l(n)^+$. Llamemos

$$r_i = \text{tr}(F_i(A)) \ , \quad i \in \mathbb{I}_n \quad \text{ y } \quad s = r_1 + \cdots + r_n = \langle A\mathbb{1} \, , \mathbb{1} \rangle \ .$$

1. Probar que $s^n \cdot \text{per}\, A \geq n! \cdot \prod\limits_{ij \in \mathbb{I}_n} |r_i|^2$.

2. Deducir que, si $A \in \mathcal{DS}\,(n) \cap \mathcal{G}l(n)^+$, entonces per $A \geq n! \; n^{-n}$.

Capítulo 8

Producto de Hadamard

8.1 Propiedades básicas

Recordemos algunas nociones adelantadas en la Sección 3.5

Definición 8.1.1. Dadas $A, B \in \mathcal{M}_{n,m}(\mathbb{C})$ se define el producto de Hadamard $A \circ B$ como la matriz

$$A \circ B = \big(a_{ij}\ b_{ij}\big)_{\substack{i \in \mathbb{I}_n \\ j \in \mathbb{I}_m}} \in \mathcal{M}_{n,m}(\mathbb{C}) \ .$$

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

Teorema 8.1.2 (Teorema 2 de Schur)**.** *Sean $A,\ B \in \mathcal{M}_n(\mathbb{C})^+$, entonces $A \circ B \in \mathcal{M}_n(\mathbb{C})^+$. Además, si $A > 0$ y $B > 0$, entonces $A \circ B > 0$.*

Demostración. Ya fue demostrado en 3.6.2

Corolario 8.1.3. *Sean $A,\ B \in \mathcal{M}_n(\mathbb{C})^+$, entonces*

1. $\mu_n(A)\mu_n(B) \leq \mu_n(A \circ B)$.
2. $\|A \circ B\| = \mu_1(A \circ B) \leq \mu_1(A)\mu_1(B) = \|A\|\,\|B\|$.

Demostración. Ejercicio. ■

Ahora empezamos a mostrar novedades sobre el producto de Hadamard.

Proposición 8.1.4. *Sea $\mathcal{S} = \text{Gen}\,\{e_i \otimes e_i : i \in \mathbb{I}_n\} \subseteq \mathbb{H}_n \otimes \mathbb{H}_n$. Identificaremos $L(\mathcal{S})$ con $\mathcal{M}_n(\mathbb{C})$ en la manera obvia. Definamos el operador lineal*

$$\Phi : L(\mathbb{H}_n \otimes \mathbb{H}_n) \to \mathcal{M}_n(\mathbb{C}) \quad \textit{dado por} \quad \Phi(T) = T_{\mathcal{S}}\ , \quad T \in L(\mathbb{H}_n \otimes \mathbb{H}_n)\ .$$

Entonces, dados $A, B \in \mathcal{M}_n(\mathbb{C})$, se verifica que $\Phi(A \otimes B) = A \circ B$.

Demostración. Representemos $A \otimes B$ como producto de Kronecker, como en la Observación 7.1.1. Con las notaciones de submatrices de la Definición 7.3.12, es fácil ver que, si tomamos $\alpha = \big\{(1,1),(2,2),\dots,(n,n)\big\} \subseteq \mathbb{I}_n \times \mathbb{I}_n$, entonces

$$\Phi(A \otimes B) = (A \otimes B)_{\mathcal{S}} = (A \otimes B)[\alpha] = A \circ B \ ,$$

como se afirmaba. ■

Definición 8.1.5. Dada $A \in \mathcal{M}_{n,m}(\mathbb{C})$, llamaremos

$$C(A) = \max_{i \in \mathbb{I}_m} \|C_i(A)\|_2 \qquad \text{y} \qquad F(A) = \max_{i \in \mathbb{I}_n} \|F_i(A)\|_2 \ .$$

Notar que estos números pueden, también, caracterizarse por las fórmulas

$$C(A)^2 = \|A^*A \circ I_m\|_{sp} \qquad \text{y} \qquad F(A)^2 = \|AA^* \circ I_n\|_{sp} \ .$$

Por lo tanto $C(A) \le \|A\|_{sp}$ y $F(A) \le \|A\|_{sp}$. ▲

Proposición 8.1.6. *Sean $A, B \in \mathcal{M}_{n,m}(\mathbb{C})$. Entonces*

$$\|A \circ B\|_{sp} \le C(A)F(B) \le \|A\|_{sp}\, \|B\|_{sp} \ .$$

Demostración. Para cada par $x \in \mathbb{C}^m,\ y \in \mathbb{C}^n$ de vectores unitarios, se tiene que

$$\begin{aligned}
\big|\langle A \circ B\, x, y\rangle\big|^2 &= \Big| \sum_{i,j} a_{ij}\, b_{ij}\, x_j\, \overline{y_i} \Big|^2 = \Big| \sum_{i,j} (a_{ij}\, x_j)\,(b_{ij}\, \overline{y_i}) \Big|^2 \\
&\le \Big(\sum_{i,j} |a_{ij}|^2\, |x_j|^2 \Big) \Big(\sum_{i,j} |b_{ij}|^2\, |y_i|^2 \Big) \qquad \text{(por Cauchy-Schwarz)} \\
&= \Big(\sum_j |x_j|^2 \sum_i |a_{ij}|^2 \Big) \Big(\sum_i |y_j|^2 \sum_j |b_{ij}|^2 \Big) \\
&\le C(A)^2 \|x\|^2\ F(B)^2 \|y\|^2 = C(A)^2\, F(B)^2 \ .
\end{aligned}$$

Como $\|A \circ B\|_{sp} = \max \big\{ |\langle A \circ B\, x, y\rangle| \ : \|x\| = \|y\| = 1 \big\}$, el resultado está demostrado. ■

Observación 8.1.7. Sean $A \in \mathcal{G}l(n)^+$ y $J \subseteq \mathbb{I}_n$, con $|J| = k$. Luego se tiene que

$$A[J] = (a_{ij})_{i,j \in J} \in \mathcal{G}l(k)^+ \qquad \text{y} \qquad A^{-1}[J] \ge A[J]^{-1} \ .$$

En efecto, esto es un caso particular del Corolario 6.3.7 (o de la Proposición 3.8.7). ▲

Proposición 8.1.8. *Sean $A, B \in \mathcal{G}l(n)^+$. Entonces se verifica que*

1. $(A \circ B)^{-1} \le A^{-1} \circ B^{-1}$.
2. $A \circ A^{-1} \ge I \ge (A \circ A^{-1})^{-1}$.

Demostración. Se deduce de la Observación 8.1.7 y de la Proposición 8.1.4, ya que que $A \circ B$ es una submatriz principal de $A \otimes B$, mientras que $A^{-1} \circ B^{-1}$ lo es de $A^{-1} \otimes B^{-1} = (A \otimes B)^{-1}$, para los mismos índices. ■

8.2 La norma de un multiplicador Hadamard

Definición 8.2.1. Fijemos $A \in \mathcal{M}_n(\mathbb{C})$, y definamos el operador de multiplicación

$$M_A : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C}) \quad \text{dado por} \quad M_A(B) = A \circ B \ , \quad B \in \mathcal{M}_n(\mathbb{C}) \ .$$

Fijada una norma N en $\mathcal{M}_n(\mathbb{C})$, denotaremos $K_N(A)$ a la norma inducida para M_A:

$$\begin{aligned} K_N(A) &= \max \left\{ N(A \circ B) : B \in \mathcal{M}_n(\mathbb{C}) \text{ es tal que } N(B) = 1 \right\} \\ &= \min \left\{ k \geq 0 : N(A \circ B) \leq k\, N(B) \quad \text{para toda} \quad B \in \mathcal{M}_n(\mathbb{C}) \right\} \ . \end{aligned}$$

En el caso de que N sea la norma espectral, escribiremos K_A en lugar de $K_{\|\cdot\|_{sp}}(A)$. ▲

Observación 8.2.2. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Si N es una norma unitariamente invariante tal que $N(E_{11}) = 1$, entonces

$$\max_{i,j} |a_{ij}| \leq K_N(A) \ .$$

En efecto, notar que para todo $i, j \in \mathbb{I}_n$ se tiene

$$N(A \circ E_{ij}) = |a_{ij}| N(E_{ij}) = |a_{ij}| \quad \text{y} \quad N(E_{ij}) = 1 \ .$$

Por otra parte, para la norma espectral, de la Proposición 8.1.6 se puede deducir que

$$K_A \leq \min \left\{ C(A), F(A) \right\} \leq \|A\|_{sp} \ .$$

En efecto, notar que para toda $B \in \mathcal{M}_n(\mathbb{C})$, tenemos que

$$\|A \circ B\| \leq C(A) F(B) \leq C(A) \|B\| \quad \text{y} \quad \|A \circ B\| \leq F(A) C(B) \leq F(A) \|B\|,$$

ya que $\|C_i(B)\|_2 = \|Be_i\| \leq \|B\|$ para todo $i \in \mathbb{I}_n$. Análogamente, $F(B) \leq \|B\|$. ▲

Ejercicios 8.2.3.

1. Si $N = \|\cdot\|_2$ (la norma de Frobenius), entonces

   $$K_N(A) = \max_{i,j \in \mathbb{I}_n} |a_{ij}| \quad , \quad A \in \mathcal{M}_n(\mathbb{C}) \ .$$

   Notar que $(\mathcal{M}_n(\mathbb{C}), N)$ es un espacio de Hilbert, M_A es un operador diagonal, y $K_N(A) = \|M_A\|_{sp}$.

2. Algo más complicado es probar que, para cualquier $A \in \mathcal{M}_n(\mathbb{C})$,

   $$K_{\|\cdot\|_1}(A) = K_A \ .$$

   Debe usarse que $\|\cdot\|_1$ es la norma "dual" de la espectral (esto es del mismo tipo que $(\ell^1)^* = \ell^\infty$, pensada en los valores singulares), y que el operador "adjunto" de M_A es el mismo M_A. Esto último se deduce de la identidad

   $$\operatorname{tr}\left((A \circ B) C^t \right) = \sum_{i,j \in \mathbb{I}_n} a_{ij}\ b_{ij}\ c_{ij} = \operatorname{tr}\left(B (A \circ C)^t \right) \ ,$$

   donde se identifica a $\mathcal{M}_n(\mathbb{C})$ con $\mathcal{M}_n(\mathbb{C})'$ a través de la aplicación $C \mapsto \varphi_C = \operatorname{tr}(\,\cdot\, C^t)$ (ver los Ejercicios 5.6.29 al 5.6.33) . ▲

Teorema 8.2.4. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Dada una factorización $A = D^*B$, con $B, D \in \mathcal{M}_n(\mathbb{C})$, se verifica que*

$$K_A \leq C(D)C(B) \ .$$

Demostración. Consideremos la siguiente matriz:

$$P = \begin{bmatrix} D^* & 0 \\ B^* & 0 \end{bmatrix} \begin{bmatrix} D & B \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} D^*D & A \\ A^* & B^*B \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})^+ \ .$$

Fijemos $M \in \mathcal{M}_n(\mathbb{C})$ tal que $\|M\|_{sp} \leq 1$. Luego $\begin{bmatrix} I & M \\ M^* & I \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})^+$, y por el Teorema 2 de Schur 8.1.2, tenemos que

$$P_M = P \circ \begin{bmatrix} I & M \\ M^* & I \end{bmatrix} = \begin{bmatrix} I \circ D^*D & M \circ A \\ (M \circ A)^* & I \circ B^*B \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})^+ \ .$$

Como $(D^*D)_{ii} = \|C_i(D)\|^2$, $i \in \mathbb{I}_n$, podemos deducir que $I \circ D^*D \leq C(D)^2 I$, y análogamente se ve que $I \circ B^*B \leq C(B)^2 I$. Por ende,

$$P_M \leq \begin{bmatrix} C(D)^2 I & M \circ A \\ (M \circ A)^* & C(B)^2 I \end{bmatrix} = R_M \ .$$

Conjugando con $F = \begin{bmatrix} C(D)^{-1} I & 0 \\ 0 & C(B)^{-1} I \end{bmatrix}$, obtenemos que

$$F R_M F = \begin{bmatrix} I & C(D)^{-1}C(B)^{-1}(M \circ A) \\ C(D)^{-1}C(B)^{-1}(M \circ A)^* & I \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})^+ \ .$$

Esto nos dice que $\|C(D)^{-1}C(B)^{-1}(M \circ A)\|_{sp} = C(D)^{-1}C(B)^{-1}\|M \circ A\|_{sp} \leq 1$, o sea

$$\|M\|_{sp} \leq 1 \implies \|M \circ A\|_{sp} \leq C(D)C(B) \ .$$

En otras palabras, $K_A \leq C(D)C(B)$. ■

Corolario 8.2.5 (Schur 4)**.** *Sea $A \in \mathcal{M}_n(\mathbb{C})^+$. Entonces $K_A = \text{máx}\{A_{ii} : \ i \in \mathbb{I}_n\}$.*

Demostración. Notemos $M = \text{máx}\{A_{ii} : i \in \mathbb{I}_n\}$. Hemos visto que $M \leq K_A$ (porque $A_{ii} = \|A \circ E_{ii}\|$). Por otra parte, como $A \in \mathcal{M}_n(\mathbb{C})^+$, sabemos que existe $B \in \mathcal{M}_n(\mathbb{C})$ tal que $A = B^*B$. Es fácil ver que, en tal caso, $A_{ii} = \|C_i(B)\|^2$ para todo $i \in \mathbb{I}_n$. Esto dice que $M = C(B)^2$. Por el Teorema 8.2.4 deducimos que $K_A \leq C(B)^2 = M$. ■

8.3 Funcionales positivas

El teorema de Haagerup (1983) dice que, dado $A \in \mathcal{M}_n(\mathbb{C})$, existe una factorización $A = D^*B$, como en el Teorema 8.2.4, tal que se obtiene la igualdad $K_A = C(D)C(B)$. Su formulación y demostración original utilizaba profundas nociones y resultados de álgebras de operadores.

Esto motivó que, desde el ámbito de los especialistas en análisis matricial, fueran apareciendo numerosas pruebas simplificadas del teorema de Haagerup. De todas ellas hemos seleccionado la obtenida por Paulsen, Power y Smith en [29]. Necesitaremos, sin embargo, adaptar al contexto de matrices ciertas nociones y resultados elementales de álgebras de operadores. Fundamentalmente, propiedades y criterios de existencia de funcionales positivas.

Definición 8.3.1. Sea $\mathcal{S} \subseteq \mathcal{M}_n(\mathbb{C})$ un subespacio cerrado por adjunción, es decir, $T \in \mathcal{S}$ si y sólo si $T^* \in \mathcal{S}$. Una *funcional* en $\mathcal{S}$ es una aplicación lineal $\varphi : \mathcal{S} \to \mathbb{C}$. Se definen los siguientes tipos de funcionales:

1. Notamos φ^* (*adjunta* de φ) a la funcional dada por $\varphi^*(A) = \overline{\varphi(A^*)}$, $A \in \mathcal{S}$.
2. Decimos que φ es *autoadjunta* si $\varphi^* = \varphi$. Es decir, si $\varphi(A^*) = \overline{\varphi(A)}$, $A \in \mathcal{S}$.
3. La funcional φ se llama *positiva* si $\varphi(A) \geq 0$ cuando $A \in \mathcal{S} \cap \mathcal{M}_n(\mathbb{C})^+$.
4. Se considera la norma inducida en las funcionales por la norma espectral de las matrices. Es decir $\|\varphi\| = \text{máx}\{|\varphi(A)| : A \in \mathcal{S}\,,\ \|A\|_{sp} = 1\}$. ▲

Ejercicios 8.3.2. Sea $\mathcal{S} \subseteq \mathcal{M}_n(\mathbb{C})$ un subespacio cerrado por adjunción.

1. Sea φ una funcional en $\mathcal{S}$. Probar que
   (a) $\|\varphi\| = \|\varphi^*\|$.
   (b) φ es autoadjunta si y sólo si $\varphi(A) \in \mathbb{R}$ para toda $A \in \mathcal{S} \cap \mathcal{H}(n)$.
   (c) Si φ es positiva, entonces es también autoadjunta.
   (d) Toda funacional autoadjunta en $\mathcal{S}$ es resta de dos positivas.

   Se usa que si $A \in \mathcal{S}$, entonces $\operatorname{Re} A \in \mathcal{S}$ e $\operatorname{Im} A \in \mathcal{S}$.

2. Dada $B \in \mathcal{M}_n(\mathbb{C})$, se define la siguiente funcional en $\mathcal{M}_n(\mathbb{C})$:

   $$\varphi_B : \mathcal{M}_n(\mathbb{C}) \to \mathbb{C} \quad \text{dada por} \quad \varphi_B(A) = \langle A, B\rangle = \operatorname{tr}(AB^*)\ , \quad A \in \mathcal{M}_n(\mathbb{C}).$$

   Verificar que
   (a) Para toda funcional φ en $\mathcal{M}_n(\mathbb{C})$ existe una única matriz $B \in \mathcal{M}_n(\mathbb{C})$ tal que $\varphi = \varphi_B$.
   (b) Dados $x, y \in \mathbb{C}^n$ consideremos la matriz $x \odot y = xy^* \in \mathcal{M}_n(\mathbb{C})$, definida en la sección 1.9. Se tiene que $\varphi_B(xy^*) = \langle x, By\rangle$.
   (c) $(\varphi_B)^* = \varphi_{B^*}$, y por lo tanto φ_B es autoadjunta si y sólo si $B \in \mathcal{H}(n)$.
   (d) φ_B es positiva si y sólo si $B \in \mathcal{M}_n(\mathbb{C})^+$. ▲

Proposición 8.3.3. *Sea $B \in \mathcal{M}_n(\mathbb{C})$. Entonces*

1. $|\operatorname{tr} B| \leq \operatorname{tr} |B|$.

2. $\operatorname{tr} B = \operatorname{tr}|B|$ *si y sólo si* $B \in \mathcal{M}_n(\mathbb{C})^+$.

3. $\|\varphi_B\| = \|B\|_1 = \operatorname{tr}|B|$.

Demostración.

1. Sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una bon de vectores propios de $|B|$ asociada a $s(B)$. Luego, si $B = U|B|$ es la DP de B, con $U \in \mathcal{U}(n)$,

$$|\operatorname{tr} B| = \left| \sum_{k=1}^{n} \langle U|B|v_k, v_k\rangle \right| \leq \sum_{k=1}^{n} s_k(B)\big|\langle Uv_k, v_k\rangle\big| \leq \sum_{k=1}^{n} s_k(B) = \operatorname{tr}|B| \ .$$

2. Si $\operatorname{tr} B = \operatorname{tr}|B|$, entonces

$$\sum_{k=1}^{n} s_k(B) = \operatorname{tr}|B| = \operatorname{tr} B = \sum_{k=1}^{n} \langle Bv_k, v_k\rangle = \sum_{k=1}^{n} s_k(B)\langle Uv_k, v_k\rangle \ .$$

Dado que $|\langle Uv_k, v_k\rangle| \leq 1$ para todo $k \in \mathbb{I}_n$, por el caso en que se obtiene igualdad en la desigualdad de Cauchy Schwarz, todos los números complejos $\langle Uv_k, v_k\rangle$ deben tener el mismo argumento. Como la suma da un número positivo, se debe verificar que $\langle Uv_k, v_k\rangle = 1$ para todo $k \in \mathbb{I}_n$. Pero un unitario con unos en la diagonal (en nuestro caso la matriz de U en la base $\mathcal{B}$) debe ser la identidad. De ello se deduce que $U = I$ y que $B = |B| \in \mathcal{M}_n(\mathbb{C})^+$. La recíproca es obvia.

3. Notar que $\operatorname{tr}|B| = \operatorname{tr}(U|B|U^*) = \operatorname{tr}(UB^*) = \varphi_B(U) \leq \|\varphi_B\|$. Por otro lado, por el item anterior y el Corolario 5.3.11,

$$|\operatorname{tr}(AC)| \leq \operatorname{tr}|AC| \leq \|A\|_{sp} \operatorname{tr}|C| \ ,$$

para todo par $A, C \in \mathcal{M}_n(\mathbb{C})$. Entonces

$$|\varphi_B(A)| = |\operatorname{tr}(A|B|U^*)| = |\operatorname{tr}(U^*A\,|B|)| \leq \|U^*A\|_{sp} \operatorname{tr}|B| = (\operatorname{tr}|B|)\,\|A\|_{sp} \ ,$$

para toda $A \in \mathcal{M}_n(\mathbb{C})$. Por lo tanto $\|\varphi_B\| \leq \operatorname{tr}|B|$. ■

Teorema 8.3.4. *Sea $\mathcal{S} \subseteq \mathcal{M}_n(\mathbb{C})$ un subespacio cerrado por adjunción tal que $I \in \mathcal{S}$. Sea φ una funcional en $\mathcal{S}$. Luego las siguientes condiciones son equivalentes:*

1. φ *es positiva.*

2. $\|\varphi\| = \varphi(I)$.

3. Existe $B \in \mathcal{M}_n(\mathbb{C})^+$ *tal que* φ *es la restricción de* φ_B *a* $\mathcal{S}$.

Demostración.

$1 \to 2$ Sea $A \in \mathcal{S}$. Si $A = A^*$, se tiene que $-\|A\|_{sp}\, I \leq A \leq \|A\|_{sp}\, I$. Luego, si φ es positiva en $\mathcal{S}$, tenemos

$$-\|A\|_{sp}\, \varphi(I) \leq \varphi(A) \leq \|A\|_{sp}\, \varphi(I) \implies |\varphi(A)| \leq \|A\|_{sp}\, \varphi(I)\ .$$

Si $A \neq A^*$, sea $\theta \in [0, 2\pi)$ tal que $\varphi(A) = e^{i\theta}|\varphi(A)|$, o sea que $\varphi(e^{-i\theta}A) = |\varphi(A)|$. Llamenos $A_0 = e^{-i\theta}A$. Como φ es autoadjunta y $\varphi(A_0) \in \mathbb{R}$, deducimos que $\varphi(A_0) = \varphi(\operatorname{Re} A_0)$. Por todo esto,

$$|\varphi(A)| = \varphi(A_0) = \varphi(\operatorname{Re} A_0) \leq \|\operatorname{Re} A_0\|_{sp}\, \varphi(I) \leq \|A_0\|_{sp}\, \varphi(I) = \|A\|_{sp}\, \varphi(I)\ .$$

Luego $\|\varphi\| \leq \varphi(I)$. La otra desigualdad es obvia ($\|I\|_{sp} = 1$).

$2 \to 3$ Sea φ una funcional en $\mathcal{S}$ tal que $\|\varphi\| = \varphi(I)$. Por el teorema de Hahn Banach (en dimensión finita se lo puede probar por inducción con la prueba tradicional), existe una extensión Φ de φ a todo $\mathcal{M}_n(\mathbb{C})$ que tiene la misma norma. Luego existe $B \in \mathcal{M}_n(\mathbb{C})$ tal que $\Phi = \varphi_B$. Por la Proposición 8.3.3, deducimos que

$$\operatorname{tr}|B| = \|\varphi_B\| = \|\Phi\| = \|\varphi\| = \varphi(I) = \varphi_B(I) = \operatorname{tr} B\ ,$$

y por lo tanto $B \in \mathcal{M}_n(\mathbb{C})^+$.

$3 \to 1$ Sea $B \in \mathcal{M}_n(\mathbb{C})^+$ tal que φ es la restricción de φ_B a $\mathcal{S}$. Si $A \in \mathcal{S} \cap \mathcal{M}_n(\mathbb{C})^+$, tenemos que

$$\varphi(A) = \operatorname{tr} AB = \operatorname{tr} B^{1/2}AB^{1/2} \geq 0\ ,$$

porque $B^{1/2}AB^{1/2} \in \mathcal{M}_n(\mathbb{C})^+$ y la funcional tr es positiva. ■

Corolario 8.3.5. Sea $\mathcal{S} \subseteq \mathcal{M}_n(\mathbb{C})$ un subespacio cerrado por adjunción tal que $I \in \mathcal{S}$, y φ una funcional positiva en $\mathcal{S}$. Luego existe Φ funcional positiva en $\mathcal{M}_n(\mathbb{C})$, con la misma norma, tal que φ es la restricción de Φ a $\mathcal{S}$. ■

8.4 Matrices incompletas

Sea $J \subseteq \mathbb{I}_n \times \mathbb{I}_n$. Una matriz *incompleta* asociada al conjunto J es un "cacho" de matriz $A = (a_{ij})_{i,j \in J}$. O sea que no se pone nada en las entradas $(i,j) \notin J$. Una matriz $B \in \mathcal{M}_n(\mathbb{C})$ es una *completación* de A si $b_{ij} = a_{ij}$ para todo $(i,j) \in J$.

Definición 8.4.1. Sea $J \subseteq \mathbb{I}_n \times \mathbb{I}_n$.

1. Llamaremos $S_J \subseteq \mathcal{M}_n(\mathbb{C})$ al subespacio

$$S_J = \big\{ C \in \mathcal{M}_n(\mathbb{C}) : c_{ij} = 0 \quad \text{para todo} \quad (i,j) \notin J \big\}.$$

2. Si A está definida solo en J, y $C \in S_J$, denotaremos $A \circ C = B \circ C$, donde $B \in \mathcal{M}_n(\mathbb{C})$ es cualquier completación de A. Notar que, como $C \in S_J$, la definición no depende de la completación elegida.

3. Diremos que J cumple (P) si

   (a) $(i,j) \in J \implies (j,i) \in J$,

   (b) $(i,i) \in J$ para todo $i \in \mathbb{I}_n$.

   En otras palabas, si J es simétrico y contiene a la diagonal (reflexivo). ▲

Existen numerosos resultados sobre matrices incompletas, fundamentalmente relativos a preguntas del tipo: ¿que debe cumplir A para que se la pueda completar a una matriz que cumpla una propiedad dada?

Un ejemplo de este tipo de resultados, es el llamado teorema de Parrot, que describe algunos casos de matrices incompletas que pueden completarse a una contracción. Una versión de aquel resultado aparece en el Ejercicio 3.9.13.

El siguiente teorema da una respuesta al problema de cuando se puede completar una casi-matriz A para que quede positiva (semidefinida), siempre que el conjunto J en el que está definida tenga la propiedad (P). Observemos que si $B \in \mathcal{M}_n(\mathbb{C})^+$ es una completación de un tal A, entonces, por el Teorema 2 de Schur 3.6.2, debe cumplirse que

$$A \circ C = B \circ C \in \mathcal{M}_n(\mathbb{C})^+ \quad \text{para toda} \quad C \in S_J \cap \mathcal{M}_n(\mathbb{C})^+ \ . \tag{8.1}$$

Esto nos da una condición necesaria sobre A para que pueda existir una completacion positiva. Esta condición sería muy pobre si J no cumple (P), porque en tal caso habría muy pocas matrices en $S_J \cap \mathcal{M}_n(\mathbb{C})^+$. Pero veremos que, si J cumple (P), entonces la condición es también suficiente:

Teorema 8.4.2. *Supongamos que $J \subseteq \mathbb{I}_n \times \mathbb{I}_n$ cumple (P). Sea $A = (a_{ij})_{i,j\in J}$ una matriz definida solo en J. Luego las siguientes condiciones son equivalentes:*

1. *Existe una completación B de A tal que $B \in \mathcal{M}_n(\mathbb{C})^+$.*

2. *Para toda matriz $C \in S_J \cap \mathcal{M}_n(\mathbb{C})^+$ se verifica $A \circ C \in \mathcal{M}_n(\mathbb{C})^+$.*

Demostración. En la Eq. (8.1) vimos que la ida es consequencia del Teorema 2 de Schur. Supongamos entonces que A cumple 2. Sea $\varphi_A : S_J \to \mathbb{C}$ la funcional definida por

$$\varphi_A(C) = \sum_{(i,j)\in J} \overline{a_{ij}}\ c_{ij} \ , \quad C = (c_{ij}) \in S_J \ .$$

Verifiquemos ahora que φ_A es positiva en S_J . En efecto, si $C \in S_J \cap \mathcal{M}_n(\mathbb{C})^+$, luego también $\overline{C} = C^T \in S_J \cap \mathcal{M}_n(\mathbb{C})^+$. Por hipótesis $A \circ \overline{C} \in \mathcal{M}_n(\mathbb{C})^+$. Si llamamos $e = (1, \dots, 1) \in \mathbb{R}^n$, entonces

$$0 \le \langle\ (A \circ \overline{C})\ e, e\rangle = \sum_{(i,j)\in J} a_{ij}\ \overline{c_{ij}} = \overline{\varphi_A(C)} = \varphi_A(C),$$

por lo que φ_A es positiva. Observar que S_J verifica las hipótesis del Teorema 8.3.4 (es cerrado por adjunción e $I \in S_J$), gracias a que J cumple (P). Luego, obtenemos una matriz $B \in$

$\mathcal{M}_n(\mathbb{C})^+$ tal que $\varphi_B\big|_{S_J} = \varphi_A$. Notar que, si $(i,j) \in J$, entonces $E_{ij} \in S_J$. Por otra parte, es fácil ver que $\operatorname{tr}(BE_{ij}) = b_{ji} = \overline{b_{ij}}$. Luego,

$$\overline{b_{ij}} = \operatorname{tr}(BE_{ij}) = \varphi_B(E_{ij}) = \varphi_A(E_{ij}) = \overline{a_{ij}} \quad , \quad (i,j) \in J \ .$$

Eso dice que B es una completación positiva de A. ■

8.5 El teorema de Haagerup

Lema 8.5.1. *Sean $T \in \mathcal{M}_n(\mathbb{C})$ y $\lambda, \mu \in \mathbb{R}_+^{*\,n}$. Notemos $D_1 = \operatorname{diag}(\lambda), D_2 = \operatorname{diag}(\mu) \in \mathcal{G}l(n)^+$ y $L \in \mathcal{M}_n(\mathbb{C})$ la matriz con entradas $L_{ij} = \lambda_i^{-1/2}\mu_j^{-1/2}$. Entonces*

$$M = \begin{bmatrix} D_1 & T \\ T^* & D_2 \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})^+ \iff \|L \circ T\| \le 1 \ .$$

Demostración. Observar que

$$\begin{bmatrix} D_1^{-1/2} & 0 \\ 0 & D_2^{-1/2} \end{bmatrix} M \begin{bmatrix} D_1^{-1/2} & 0 \\ 0 & D_2^{-1/2} \end{bmatrix} = \begin{bmatrix} I & D_1^{-1/2} T D_2^{-1/2} \\ D_2^{-1/2} T^* D_1^{-1/2} & I \end{bmatrix} .$$

Luego, por la Proposición 3.7.6, $M \in \mathcal{M}_{2n}(\mathbb{C})^+$ si y sólo si $\|D_1^{-1/2} T D_2^{-1/2}\| \le 1$. El resultado queda probado con sólo observar que $D_1^{-1/2} T D_2^{-1/2} = L \circ T$. ■

Teorema 8.5.2. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Luego las siguientes condiciones son equivalentes:*

1. *$K_A \le 1$, es decir $\|A \circ C\| \le \|C\|$ para todo $C \in \mathcal{M}_n(\mathbb{C})$. .*
2. *Existen $X, Y \in \mathcal{M}_n(\mathbb{C})^+$ tales que*
   (a) *$X \circ I \le I$ e $Y \circ I \le I$.*
   (b) *La matriz $N = \begin{bmatrix} X & A \\ A^* & Y \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})^+$.*
3. *Existen $B, D \in \mathcal{M}_n(\mathbb{C})$ tales que*
   (a) *$A = D^* B$.*
   (b) *$C(B) \le 1$ y $C(D) \le 1$.*

Demostración. $1 \to 2$: Sea $J \subseteq \mathbb{I}_{2n} \times \mathbb{I}_{2n}$ dado por

$$J = \{(i,i) : i \in \mathbb{I}_{2n}\} \cup \{(i, n+j) : i,j \in \mathbb{I}_n\} \cup \{(n+i, j) : i,j \in \mathbb{I}_n\}.$$

Observar que J cumple (P). Consideremos la matriz P de tamaño $2n \times 2n$, definida sólo en J, dada por

$$P = \begin{bmatrix} D & A \\ A^* & D \end{bmatrix} , \qquad \text{donde} \qquad D = \begin{bmatrix} 1 & & ? \\ & \ddots & \\ ? & & 1 \end{bmatrix} ,$$

que es una matriz de tamaño $n \times n$ definida solamente en la diagonal.

Clamor: Si $M \in S_J \cap \mathcal{M}_{2n}(\mathbb{C})^+$, entonces $P \circ M \in \mathcal{M}_{2n}(\mathbb{C})^+$.

En efecto, $M = \begin{bmatrix} D_1 & T \\ T^* & D_2 \end{bmatrix}$, donde $T \in \mathcal{M}_n(\mathbb{C})$, y $D_1 = \operatorname{diag}(\lambda)\,, D_2 = \operatorname{diag}(\mu)$ son matrices diagonales positivas en $\mathcal{M}_n(\mathbb{C})$. Si suponemos que D_1, D_2 son estrictamente positivas, y notamos $L \in \mathcal{M}_n(\mathbb{C})$ la matriz con entradas $L_{ij} = \lambda_i^{-1/2}\mu_j^{-1/2}$, el Lema 8.5.1 nos dice que, como $M \in \mathcal{M}_{2n}(\mathbb{C})^+$, entonces $\|L \circ T\| \le 1$. Observar que

$$P \circ M = \begin{bmatrix} D_1 & A \circ T \\ (A \circ T)^* & D_2 \end{bmatrix} .$$

Como $K_A \le 1$, tenemos que $\|L \circ (A \circ T)\| = \|A \circ (L \circ T)\| \le 1$. Usando nuevamente el Lema 8.5.1, deducimos que $P \circ M \in \mathcal{M}_{2n}(\mathbb{C})^+$. El caso general (sin suponer que D_1 y D_2 son inversibles) se deduce del anterior, tomando la sucesión

$$M_m = M + \frac{1}{m} I_{2n} \quad \text{en } S_J \;. \quad \text{Entonces} \quad \mathcal{M}_{2n}(\mathbb{C})^+ \ni P \circ M_m \xrightarrow[m \to \infty]{} P \circ M \;.$$

Como $\mathcal{M}_{2n}(\mathbb{C})^+$ es cerrado, el clamor queda demostrado. Por el Teorema 8.4.2, tenemos que existe una completación N de P tal que

$$N = \begin{bmatrix} X & A \\ A^* & Y \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})^+ \quad \text{y, por lo tanto,} \quad X \circ I = Y \circ I = I \;.$$

Luego las matricecs $X, Y \in \mathcal{M}_n(\mathbb{C})^+$ cumplen lo pedido.

$2 \to 3$: Como $N \in \mathcal{M}_{2n}(\mathbb{C})^+$, por el teorema de Cholewsky (Corolario 3.1.5), existe una matriz $K \in \mathcal{M}_{2n}(\mathbb{C})$ triangular superior tal que $N = K^*K$. Si la escribimos en bloques de $n \times n$,

$$K = \begin{bmatrix} D & B \\ 0 & G \end{bmatrix} \quad \Longrightarrow K^*K = \begin{bmatrix} D^*D & D^*B \\ B^*D & B^*B + G^*G \end{bmatrix} = \begin{bmatrix} X & A \\ A^* & Y \end{bmatrix} = N.$$

El resultado se sigue de que $A = D^*B$ y, como $X \circ I \le I$ e $Y \circ I \le I$, entonces

$$C(D)^2 = \|D^*D \circ I\| = \|X \circ I\| \le 1 \qquad \text{y}$$

$$C(B)^2 = \|B^*B \circ I\| \le \|(B^*B + G^*G) \circ I\| = \|X \circ I\| \le 1 \;.$$

La implicación $3 \to 1$ fué probada en el Teorema 8.2.4. ■

Corolario 8.5.3 (Teorema de Haagerup (1983)). *Sea* $A \in \mathcal{M}_n(\mathbb{C})$. *Entonces*

$$K_A = \text{mín}\left\{ \; C(B)C(D) \;:\; B, D \in \mathcal{M}_n(\mathbb{C}) \quad y \quad A = D^*B \; \right\} .$$

Demostración. Si $K_A = 0$, entonces por la Observación 8.2.2, se tiene que $A = 0$ y el resultado es trivial. Si $K_A > 0$, una desigualdad se deduce del Teorema 8.2.4 y, para probar la otra, basta cambiar A por $K_A^{-1}A$ y aplicar $1 \to 3$ del Teorema 8.5.2. ■

Corolario 8.5.4. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Notemos $A^{(k)} \in \mathcal{M}_{kn}(\mathbb{C})$ la matriz con $k \times k$ bloques de $n \times n$ iguales a A. Entonces $K_A = K_{A^{(k)}}$.*

Demostración. Es evidente que $K_A \le K_{A^{(k)}}$ (trabajando con matrices de $n \times n$ rellenadas con ceros). Recíprocamente, si $B, D \in \mathcal{M}_n(\mathbb{C})$ cumplen que $A = D^*B$ y $K_A = C(B)C(D)$, entonces

$$A^{(k)} = \begin{bmatrix} A & \dots & A \\ A & \dots & A \\ \vdots & \vdots & \vdots \\ A & \dots & A \end{bmatrix} = \begin{bmatrix} D & \dots & D \\ 0 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & 0 \end{bmatrix}^* \begin{bmatrix} B & \dots & B \\ 0 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & 0 \end{bmatrix} = D_k^* B_k \ .$$

Pero es claro que $C(B_k) = C(B)$ y $C(D_k) = C(D)$, dado que tienen las mismas columnas (salvo ceros). Así, $K_{A^{(k)}} \le C(B)C(D) = K_A$. ■

8.6 Determinantes

Teorema 8.6.1 (Desigualdad de Hadamard). *Si $A \in \mathcal{M}_n(\mathbb{C})^+$, entonces*

$$\det A \le \det\,(A \circ I) = \prod_{i=1}^{n} a_{ii} \quad .$$

La igualdad vale si y sólo si A es diagonal.

Demostración. Podemos suponer que $A > 0$, y entonces $a_{ii} > 0$, para todo $i \in \mathbb{I}_n$. Consideramos la matriz diagonal

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

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

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

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

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

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

[1] Si $a_1, \dots, a_m > 0$, entonces $\prod_{i=1}^{m} a_i^{\frac{1}{m}} \le \frac{1}{m} \sum_{i=1}^{m} a_i$. Sale usando que el log es una función cóncava.

Corolario 8.6.2. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces*

$$|\det A| \leq \prod_{i=1}^{n} \|C_i(A)\|_2 \ . \tag{8.2}$$

Demostración. Se aplica la desigualdad de Haramard a la matriz $B = A^*A \geq 0$. Notar que $\det B = |\det A|^2$ y que $B_{ii} = \|C_i(A)\|_2^2$, para todo $i \in \mathbb{I}_n$. ∎

Ejercicio 8.6.3. Veremos tres demostraciones alternativas de estas desigualdades.

1. Probar el Teorema 8.6.1 usando el Teorema 3 de Schur 5.1.1 y el Corolario 4.2.3.
2. Probar que el Corolario 8.6.2 implica la desigualdad de Hadamard.
3. Probar el Corolario 8.6.2 usando la descomposición QR (Teorema 1.8.2) de $A \in \mathcal{M}_n(\mathbb{C})$. Observar que (8.2) es trivial para matrices triangulares.
4. Probar el Corolario 8.6.2 usando la interpretación del determinante como un área o volumen.

Lema 8.6.4. *Sean $A \in \mathcal{G}l(n)^+$ y $\alpha(A) = \frac{\det A}{\det A_{11}}$, donde $A_{11} = (a_{ij})_{2\leq i,j\leq n} \in \mathcal{M}_{n-1}(\mathbb{C})$. Sea $E_{11} = e_1 e_1^t \in \mathcal{M}_n(\mathbb{C})$. Entonces $A - tE_{11} \geq 0$ si y sólo si $t \leq \alpha(A)$.*

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

$$\det(A - tE_{11}) = \det A - t \det A_{11}. \tag{8.3}$$

Luego, $\det(A - tE_{11}) \geq 0$ si y sólo si $t \leq \alpha(A)$. Por otro lado, todas las demás submatrices principales de $A - tE_{11}$ obtenidas con las últimas i filas y columnas, son las mismas que las respectivas de A. Por lo tanto, el determinante de cada una de ellas es positivo. Luego, por el Teorema 2.4.6 (hecho desde abajo), tenemos el resultado para desigualdades estrictas. El caso general sale tomando límite. ∎

Teorema 8.6.5 (Desigualdad de Oppenheim)**.** *Si $A, B \in \mathcal{M}_n(\mathbb{C})^+$, entonces*

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

Demostración. Si $\det A = 0$, el resultado se deduce del Teorema 2 de Schur 8.1.2, que asegura que $A \circ B \geq 0$. Supongamos, entonces, que $A > 0$. La demostración la realizaremos por inducción sobre n. Si $n = 1$, el resultado es inmediato. Sea $n \geq 2$ y supongamos el resultado válido para todas las matrices de dimensión $n - 1$. Entonces, con las notaciones del Lema 8.6.4, sabemos que

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

Por el Lema 8.6.4, si $\alpha = (\det A_{11})^{-1} \det A$, entonces $A - \alpha E_{11} \geq 0$. El Teorema 2 de Schur 8.1.2 dice que $(A - \alpha E_{11}) \circ B \geq 0$. Aplicando la fórmula (8.3), como $E_{11} \circ B = b_{11} E_{11}$ y $(A \circ B)_{11} = A_{11} \circ B_{11}$, resulta que

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

Aplicando la hipótesis inductiva, obtenemos

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

y el teorema queda demostrado. ■

Teorema 8.6.6 (Desigualdad de Fisher)**.** *Sea $A \in \mathcal{M}_n(\mathbb{C})^+$, y sea $\mathcal{P}$ un sistema de proyectores en $\mathcal{H}(n)$. Entonces*

$$\det A \leq \det(C_{\mathcal{P}}(A)).$$

Recordamos que $C_{\mathcal{P}}(A) = \sum_{i=1}^{r} P_i A P_i$, si $\mathcal{P} = \{P_1, \dots, P_r\}$.

Demostración. Por la Eq. (5.8), basta probar el caso $\mathcal{P} = \{P, I - P\}$, para $P \in \mathcal{H}(n)$ un proyector. Supongamos que $\dim R(P) = k$. Conjugando a P y a A con alguna matriz unitaria (lo que no cambia los determinantes), podemos suponer que $R(P)$ es el subespacio generado por los primeros k elementos de la base canónica de $\mathbb{C}^n$. O, lo que es lo mismo, que $P = \operatorname{diag}(1, \dots, 1, 0, \dots, , 0)$, donde los unos llegan hasta el lugar k.

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

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

que verifica que $A \circ B = C_P(A)$. Aplicando la desigualdad de Oppenheim, tenemos que $\det A = \det A \prod_{i=1}^{n} b_{ii} \leq \det A \circ B = \det C_P(A)$. ■

Observación 8.6.7. Otra demostración del Teorema anterior puede hecerse usando las Proposiciones 5.4.4 y 4.2.3. En efecto, con las notaciones de 8.6.6, como $\mu(C_{\mathcal{P}}(A)) \prec \mu(A)$, si $\mu_n(A) > 0$, entonces también $\mu_n(C_{\mathcal{P}}(A)) > 0$ y

$$\det A = \prod_{i=1}^{n} \mu_i(A) \leq \prod_{i=1}^{n} \mu_i(C_{\mathcal{P}}(A)) = \det C_{\mathcal{P}}(A) \ .$$

Si $\mu_n(A) = 0$, entonces $\det A = 0$, pero $C_{\mathcal{P}}(A) \geq 0$, por lo que $\det C_{\mathcal{P}}(A) \geq 0$. ▲

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

Teorema 8.6.8. *Si A, $B \in \mathcal{M}_n(\mathbb{C})^+$, entonces*

$$\det A\ B \le \det A \circ B.$$

Demostración. El Teorema se deduce de las desigualdades de Hadamard y de Oppenheim. En efecto, $\det A\ B = \det A\ \det B \le \det A \prod_{i=1}^{n} b_{ii} \le \det A \circ B$. ∎

8.7 Ejercicios

Ejercicios del texto

8.7.1. Dada $A \in \mathcal{M}_{n,m}(\mathbb{C})$, mostrar que

$$C(A) = \max_{i \in \mathbb{I}_m} \|C_i(A)\|_2 = \|A^*A \circ I_m\|_{sp}^{1/2} \qquad \text{y} \qquad F(A) = \max_{i \in \mathbb{I}_n} \|F_i(A)\|_2 = \|AA^* \circ I_n\|_{sp}^{1/2} \ .$$

Deducir que $\max\big\{C(A)\,,\ F(A)\big\} \le \|A\|_{sp}$.

8.7.2. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Porbar las siguientes afirmaciones:

1. Si $N = \|\cdot\|_2$ (la norma de Frobenius), entonces
$$K_N(A) = \max_{i,j} |a_{ij}| \quad , \quad A \in \mathcal{M}_n(\mathbb{C}) \ .$$
2. Dadas $B, C \in \mathcal{M}_n(\mathbb{C})$, se cumple que
$$\operatorname{tr}\big((A \circ B)C^T\big) = \sum_{i,j} a_{ij}\ b_{ij}\ c_{ij} = \operatorname{tr}\big(B(A \circ C)^T\big) \ .$$
3. Probar que el operador "adjunto" de $M_A \in L(\mathcal{M}_n(\mathbb{C}))$ es el mismo M_A, idenficando $\mathcal{M}_n(\mathbb{C})'$ con $\mathcal{M}_n(\mathbb{C})$, a través de la aplicación
$$\mathcal{M}_n(\mathbb{C}) \ni C \ \longmapsto\ \varphi_C = \operatorname{tr}(\,\cdot\, C^T) \in \mathcal{M}_n(\mathbb{C})' \ .$$
4. Probar que $K_{\|\cdot\|_1}(A) = K_A$.

8.7.3. Sea $\mathcal{S} \subseteq \mathcal{M}_n(\mathbb{C})$ un subespacio cerrado por adjunción (i.e. $T \in \mathcal{S} \implies T^* \in \mathcal{S}$).

1. Sea φ una funcional en $\mathcal{S}$ (usaremos notaciones de la Definición 8.3.1). Probar que
   (a) $\|\varphi\| = \|\varphi^*\|$.
   (b) φ es autoadjunta si y sólo si $\varphi(A) \in \mathbb{R}$ para toda $A \in \mathcal{S} \cap \mathcal{H}(n)$.
   (c) Si φ es positiva, entonces es también autoadjunta.
   (d) Toda funacional autoadjunta en $\mathcal{S}$ es resta de dos positivas.

Se usa que si $A \in \mathcal{S}$, entonces $\operatorname{Re} A \in \mathcal{S}$ e $\operatorname{Im} A \in \mathcal{S}$.

2. Dada $B \in \mathcal{M}_n(\mathbb{C})$, se define la siguiente funcional en $\mathcal{M}_n(\mathbb{C})$:

$$\varphi_B : \mathcal{M}_n(\mathbb{C}) \to \mathbb{C} \quad \text{dada por} \quad \varphi_B(A) = \langle A, B\rangle = \operatorname{tr}(AB^*) \ , \quad A \in \mathcal{M}_n(\mathbb{C}).$$

Verificar que

(a) Para toda funcional φ en $\mathcal{M}_n(\mathbb{C})$ existe una única matriz $B \in \mathcal{M}_n(\mathbb{C})$ tal que $\varphi = \varphi_B$.

(b) Dados $x, y \in \mathbb{C}^n$ consideremos la matriz $x \odot y = xy^* \in \mathcal{M}_n(\mathbb{C})$, definida en la sección 1.9. Se tiene que $\varphi_B(xy^*) = \langle x, By\rangle$.

(c) $(\varphi_B)^* = \varphi_{B^*}$, y por lo tanto φ_B es autoadjunta si y sólo si $B \in \mathcal{H}(n)$.

(d) φ_B es positiva si y sólo si $B \in \mathcal{M}_n(\mathbb{C})^+$.

8.7.4 (Hahn Banach finito)**.** Sea $\mathcal{S} \subseteq \mathcal{M}_n(\mathbb{C})$ un subespacio, y sea $\varphi : \mathcal{S} \to \mathbb{C}$ una funcional lineal. Si $\|\varphi\| = \text{máx}\{|\varphi(A)| : A \in \mathcal{S} \text{ y } \|A\|_{sp} = 1\}$, existe una extensión Φ de φ a todo $\mathcal{M}_n(\mathbb{C})$ que tiene la misma norma.

8.7.5. Si $a_1, \dots, a_n > 0$, entonces $\prod\limits_{i=1}^{n} a_i^{\frac{1}{n}} \le \frac{1}{n} \sum\limits_{i=1}^{n} a_i$.

8.7.6. Distintas pruebas de la desigualdad de Hadamard:

1. Probar el Teorema 8.6.1 usando el Teorema 3 de Schur 5.1.1 y el Corolario 4.2.3.

2. Probar que el Corolario 8.6.2 implica la desigualdad de Hadamard.

3. Probar el Corolario 8.6.2 usando la descomposición QR (Teorema 1.8.2) de $A \in \mathcal{M}_n(\mathbb{C})$. Observar que (8.2) es trivial para matrices triangulares.

4. Probar el Corolario 8.6.2 usando la interpretación del determinante como un área o volumen.

8.7.7. Dar los detalles de la prueba del Lema 8.6.4.

Ejercicios nuevos

8.7.8. Sean $x, y \in \mathbb{C}^n$ y $G \in \mathcal{M}_n(\mathbb{C})$. Probar que

$$G \circ x \odot y \ = \ \operatorname{diag}(x) \ G \operatorname{diag}(y)^* \ .$$

Definición 8.7.9. Dada $G \in \mathcal{M}_n(\mathbb{C})^+$, se define:

1. El índice **minimal** de G como

$$I(G) = \max\{\lambda \ge 0 \ : \ G \circ B \ge \lambda B \quad \text{para todo} \quad B \in \mathcal{M}_n(\mathbb{C})^+\} \ .$$

2. Dada una norma N en $\mathcal{M}_n(\mathbb{C})$, se define el índice N de Hadamard para G como

$$I_N(G) = \max \left\{ \lambda \geq 0 \, : \, N(G \circ B) \geq \lambda N(B) \quad \text{para todo} \quad B \in \mathcal{M}_n(\mathbb{C})^+ \right\}$$

$$= \min \left\{ N(G \circ B) \, : \, B \in \mathcal{M}_n(\mathbb{C})^+ \text{ y } N(B) = 1 \right\} .$$

El índice de G asociado con la norma espectral $\|\cdot\| = \|\cdot\|_{sp}$ se denota $I_{sp}(G)$, mientras que el asociado a la norma Frobenius $\|\cdot\|_2$ será denotado por $I_2(G)$. ▲

8.7.10. Sean $G \in \mathcal{M}_n(\mathbb{C})^+$, $\mathbb{1} = (1, \dots, 1) \in \mathbb{C}^n$ y $E = \mathbb{1} \cdot \mathbb{1}^T$. Sea N una norma.

1. $I(G) \neq 0$ si y sólo si $\mathbb{1} \in R(G)$. Si $y \in \mathbb{C}^n$ cumple que $Gy = \mathbb{1}$, entonces

$$I(G) = \Big(\sum_{i=1}^{m} y_i \Big)^{-1} = \langle y, \mathbb{1} \rangle^{-1} = \rho(G^\dagger E)^{-1} = \min \left\{ \langle Gz, z \rangle \, : \, \sum_{i=1}^{n} z_i = 1 \right\}$$

   Y si $G > 0$, se tiene que $I(G) = \Big(\sum\limits_{i,j=1}^{n} (G^{-1})_{ij} \Big)^{-1} = \dfrac{\det G}{\det(G + E) - \det G}$.

2. $I(G) \leq I_N(G)$ para cualquier norma unitariamente invariante N.

3. $I_N(G) \neq 0 \iff G \circ I \neq 0 \iff G_{ii} \neq 0$ para todo $i \in \mathbb{I}_n$.

4. Si $D = \operatorname{diag}(d) \in \mathcal{G}l(n)^+$ es **diagonal**, $I_N(D) = N'(D^{-1})^{-1}$. En part.

$$I(D) = I_{sp}(D) = \Big(\sum_{i=1}^{n} d_i^{-1} \Big)^{-1} \quad \text{e} \quad I_2(D) = \Big(\sum_{i=1}^{n} d_i^{-2} \Big)^{-1/2} .$$

5. Los indices I_2 e I_{sp} se alcanzan en matrices $B \in \mathcal{M}_n(\mathbb{C})^+$ de rango 1. O sea,

$$I_2(G) = \min_{\|x\|=1} \|G \circ xx^*\|_2 \quad \text{e} \quad I_{sp}(G) = \min_{\|y\|=1} \|G \circ yy^*\| .$$

   Más aún, ambos minimos se alcanzan en vectores $x \geqslant 0$ (o $y \geqslant 0$).

6. $I_{sp}(A) = \inf \{ I_{sp}(D) : A \leq D \text{ y } D \text{ es diagonal} \}$.

7. Si $x \in \mathbb{C}^n$, entonces $I_{sp}(x \odot x) = \min\limits_{i \in \mathbb{I}_n} |x_i|^2$.

8. Sea $A = \begin{bmatrix} a & b \\ \bar{b} & c \end{bmatrix} \in \mathcal{M}_2(\mathbb{C})^+$. Probar que

   (a) Si $|b| < \min\{a, c\}$, entonces $I_{sp}(A) = \frac{ac - |b|^2}{a + c - 2|b|}$.

   (b) Si $|b| \geq \min\{a, c\}$, se tiene que $I_{sp}(A) = \min\{a, c\}$.

Capítulo 9

Algunas desigualdades de matrices

9.1 Partes reales

Definición 9.1.1. 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 coordenadas. ▲

Proposición 9.1.2 (Fan-Hoffman)**.** *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces*

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

Demostración. Sean $x_1, \dots, x_n$ y $w_1, \dots, w_n$ bases ortonormales de $\mathbb{C}^n$, formadas por autovectores de $\operatorname{Re} A$ (resp. A^*A) adaptadas a $\mu(\operatorname{Re} A)$ (resp. $\mu(A^*A)$). Dado $k \in \mathbb{I}_n$, sea

$$x \in \operatorname{Gen}\{x_1, \dots, x_k\} \cap \operatorname{Gen}\{w_k, \dots, w_n\} ,$$

un vector unitario (debe existir por las dimensiones de los subespacios). Entonces, por el Teorema de Courant-Fisher 2.3.3 y la Proposición 3.2.6,

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

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

Proposición 9.1.3 (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 1 de Schur 1.6.1). Dado $k \in \mathbb{I}_n$, por el Principio del máximo de Ky Fan (Proposición 5.1.4), se tiene que

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

Para $k = n$ hay igualdad porque $\operatorname{Re}\operatorname{tr}(A) = \dfrac{\operatorname{tr}A + \overline{\operatorname{tr}A}}{2} = \operatorname{tr}\dfrac{A + A^*}{2} = \operatorname{tr}\operatorname{Re}A$. ■

Corolario 9.1.4. 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 $\mu_n(\operatorname{Re}A) \leq \operatorname{Re} z \leq \mu_1(\operatorname{Re}A)$.

Observación 9.1.5. Sean $A, B \in \mathcal{H}(n)$. Se llama **producto simetrizado** de A y B a

$$S = S(A, B) = AB + BA \in \mathcal{H}(n) .$$

Supongamos que $A > 0$ y $S = S(A, B) > 0$. Notar que, si $C = A^{-1/2}BA^{1/2}$,

$$0 < A^{-1/2}SA^{-1/2} = A^{1/2}BA^{-1/2} + A^{-1/2}BA^{1/2} = \operatorname{Re} C .$$

Por el Corolario 9.1.4, se tiene que $\sigma(C) = \sigma(B) \subseteq \mathbb{R}_+^*$. Como $B \in \mathcal{H}(n)$, debe ser $B > 0$. Si $A \in \mathcal{M}_n(\mathbb{C})^+$ no es inversible, notar que dado $\varepsilon > 0$ bien chico, se tiene que

$$S(A + \varepsilon I, B) = S(A, B) + 2\varepsilon B > 0 \quad (\text{porque } \mathcal{G}l(n)^+ \text{ es abierto en } \mathcal{H}(n)) .$$

Luego se aplica el caso anterior, y tambien $A \geq 0 \;+\; S(A, B) > 0 \implies B > 0$. ▲

Ejercicio 9.1.6. Sean $A, B \in \mathcal{H}(n)$.

1. Probar que, para cada $x \in \mathbb{C}^n$, se tiene $\langle S(A, B)x, x\rangle = 2\operatorname{Re}\langle Ax, Bx\rangle$.
2. Dar un ejemplo de matrices positivas A y B tales que $S(A, B) \not\geq 0$. ▲

Proposición 9.1.7 (Kittaneh '95)**.** *Sean $A, B \in \mathcal{M}_n(\mathbb{C})$ tales que $AB \in \mathcal{H}(n)$. Entonces,*

$$|||AB||| \leq |||\operatorname{Re} BA|||$$

para toda NUI $|||\cdot|||$ en $\mathcal{M}_n(\mathbb{C})$.

Demostración. Comencemos notando que los autovalores de BA son los mismos que los de AB y por ende son todos reales. Más aún, en la Proposición 1.5.5 vimos que $\mu(AB) = \mu(BA)$. Luego, usando la Proposición 9.1.3, obtenemos que

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

Como AB y $\operatorname{Re} BA \in \mathcal{H}(n)$, podemos aplicar el Corolario 5.3.14 (usando que $t \mapsto |t|$ es convexa) y deducir que $s(AB) = |\mu(AB)|^{\downarrow} \prec_w |\mu(\operatorname{Re} AB)|^{\downarrow} = s(\operatorname{Re} AB)$, por lo que $|||AB||| \leq |||\operatorname{Re}(BA)|||$ para toda NUI. ■

Proposición 9.1.8 (Corach-Porta-Recht, '93)**.** *Sean $T, S \in \mathcal{H}(n)$ y supongamos que S es inversible. Entonces,*

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

para toda NUI $|||\cdot|||$ *en* $\mathcal{M}_n(\mathbb{C})$.

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

Ejercicios 9.1.9.

1. Usando el famoso truco de las matrices de 2×2, extender la desigualdad CPR a cualquier $T \in \mathcal{M}_n(\mathbb{C})$, no necesariamente autoadjunta. Se sugiere usar las matrices
$$\widehat{T} = \begin{bmatrix} 0 & T \\ T^* & 0 \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})$$
y una adecuada $S_1 \in \mathbb{H}(2n)$ invertible. Ojo con los $s_k(\widehat{T})$, que son los de T, pero repetidos dos veces cada uno.

2. Con el mismo truco, probar también que, si $T \in \mathcal{M}_n(\mathbb{C})$ y $S \in \mathcal{H}(n)$ es inversible, entonces
$$|||STS + S^{-1}TS^{-1}||| \geq 2\ |||T|||$$
para toda NUI $|||\cdot|||$ en $\mathcal{M}_n(\mathbb{C})$.

3. Verificar, además, que la constante 2 es óptima en el primer caso (fijando S y moviendo todos los $T \in \mathcal{M}_n(\mathbb{C})$ o $\mathcal{H}(n)$), pero no siempre lo es en el segundo. ¿Para qué matrices S lo será? (esto último es difícil, pero es fácil encontrar familias razonablemente grandes de ejemplos donde vale, al menos para la norma espectral).

4. Otra manera de probar la desigualdad CPR (la original) es
   (a) Primero reducir al caso en que S es diagonal.
   (b) Después escribir $STS^{-1} + S^{-1}TS$ como un producto de Hadamard.
   (c) Verificar que la matriz que multiplica Hadamard, luego de "pasarla dividiendo", es semi definida positiva.
   (d) Aplicar el siguiente resultado: Si $A \geq 0$, entonces para toda $B \in \mathcal{M}_n(\mathbb{C})$ y para toda nui $|||\cdot|||$ en $\mathcal{M}_n(\mathbb{C})$, se tiene que
   $$|||A \circ B|||\ \leq\ \max\{\, a_{ii} : i \in \mathbb{I}_n \,\}\ |||B|||\ .$$
   Esto es conocido como el Teorema de Schur (ver Corolario 8.2.5, Schur 4).

9.2 Desigualdad de Thompson

Observación 9.2.1. A diferencia del módulo de números, el de matrices no cumple la desigualdad triangular. O sea que existen matrices A, B tales que $|A+B| \not\leq |A|+|B|$ (Ejercicio: encontrar un par así en $\mathcal{M}_2(\mathbb{C})$). Esto sucede porque sus partes unitarias pueden mezclar tamaños en forma aleatoria. Lo mejor que se tiene para ese lado es el siguiente resultado, donde uno corrige ese problema: ▲

Teorema 9.2.2 (Thompson). *Dadas $A, B \in \mathcal{M}_n(\mathbb{C})$, existen $U, V \in \mathcal{U}(n)$ tales que*

$$|A+B| \leq U|A|U^* + V|B|V^* \ . \tag{9.1}$$

Demostración. Hagamos la descomposición polar $A+B = W|A+B|$, con $W \in \mathcal{U}(n)$. Entonces

$$|A+B| = W^*(A+B) = \operatorname{Re}\big(W^*(A+B)\big) = \operatorname{Re} W^*A + \operatorname{Re} W^*B \ . \tag{9.2}$$

Por otra parte, por la Proposicion 9.1.2 (Fan-Hoffman), existen $U, V \in \mathcal{U}(n)$ tales que

$$\operatorname{Re} W^*A \leq U|W^*A|U^* = U|A|U^* \quad \text{y} \quad \operatorname{Re} W^*B \leq U|W^*B|U^* = U|B|U^* \ ,$$

porque $(W^*A)^*W^*A = A^*W^*WA = A^*A$ y entonces $|W^*A| = |A|$ (lo mismo para B). ■

En el caso de la desigualdad triangular numérica, la igualdad se da si y sólo si ambos números poseen igual argumento. Algo similar vale en el caso matricial:

Teorema 9.2.3. *Dadas $A, B \in \mathcal{M}_n(\mathbb{C})$, las siguientes afirmaciones son equivalentes:*

1. *Sólo la igualdad puede darse en la ecuación* (9.1).
2. *Existe $W \in \mathcal{U}(n)$ tal que $A = W|A|$ y también $B = W|B|$.*

Demostración.

$1 \Rightarrow 2$ Sea $A+B = W|A+B|$ la descomposición polar de $A+B$, con $W \in \mathcal{U}(n)$. Veremos que $WA \geq 0$ y $WB \geq 0$. Como en la Eq. (9.2), se tiene que

$$|A+B| = \operatorname{Re} W^*A + \operatorname{Re} W^*B \ .$$

Llamemos $C = W^*A$ y $D = W^*B$. Siguiendo el razonamiento anterior, por la Proposicion 9.1.2 (Fan-Hoffman), existen $U, V \in \mathcal{U}(n)$ tales que $\operatorname{Re} W^*A = \operatorname{Re} C \leq U|A|U^*$ y $\operatorname{Re} D \leq V|B|V^*$. Ahora bien, la hipótesis de que sólo puede darse la igualdad en (9.1) fuerza a que $\operatorname{Re} C = U|A|U^*$ y $\operatorname{Re} D = V|B|V^*$. Por lo tanto,

$$\operatorname{tr}\left[(\operatorname{Re} C)^2\right] = \operatorname{tr}|A|^2 = \operatorname{tr} A^*A = \frac{\operatorname{tr} AA^* + \operatorname{tr} A^*A}{2} = \frac{\operatorname{tr} CC^* + \operatorname{tr} C^*C}{2} \ . \tag{9.3}$$

Observar que $4\operatorname{tr}\left[(\operatorname{Re} C)^2\right] = \operatorname{tr} CC^* + \operatorname{tr} C^*C + \operatorname{tr} C^2 + \operatorname{tr}(C^*)^2$, por lo que la Eq. (9.3) se traduce como $\operatorname{tr} CC^* + \operatorname{tr} C^*C = \operatorname{tr} C^2 + \operatorname{tr}(C^*)^2$. Luego

$$\operatorname{tr}\left[(C-C^*)(C^*-C)\right] = \operatorname{tr} CC^* + \operatorname{tr} C^*C - \operatorname{tr} C^2 - \operatorname{tr}(C^*)^2 = 0 \ .$$

Esto muestra que $C = W^*A \in \mathcal{H}(n)$. Luego $W^*A = \operatorname{Re} W^*A = U|A|U^* \in \mathcal{M}_n(\mathbb{C})^+$. Analogamente se prueba que $W^*B \in \mathcal{M}_n(\mathbb{C})^+$.

2 ⇒ 1 Supongamos ahora que $A = W|A|$ y $B = W|B|$ para la misma $W \in \mathcal{U}(n)$. Luego $A + B = W(|A| + |B|) \implies |A + B| = |A| + |B|$. Si vale (9.1) para algún par $U, V \in \mathcal{U}(n)$, entonces

$$|A| + |B| \leq U|A|U^* + V|B|V^* \implies M = U|A|U^* + V|B|V^* - |A| - |B| \in \mathcal{M}_n(\mathbb{C})^+ \ .$$

Luego, la matriz $M \in \mathcal{M}_n(\mathbb{C})^+$ y tiene traza nula, o sea que $M = 0$. Esto muestra que sólo la igualdad puede complirse en (9.1). ■

9.3 Desigualdad aritmético-geométrica en matrices

Recordemos la siguiente desigualdad numérica, que ya había aparecido en el Teorema 8.6.1: Dados $a_1, \dots, a_m \in \mathbb{R}_+^*$ y $\lambda_1 \dots, \lambda_m \in [0,1]$ tales que $\sum\limits_{i \in \mathbb{I}_m} \lambda_i = 1$, se cumple que

$$\prod_{i=1}^m a_i^{\lambda_i} \ \leq \ \sum_{i=1}^m \lambda_i \, a_i \ . \tag{9.4}$$

Es la llamada desigualdad aritmético-geométrica, y se demuestra rápidamente usando que el logaritmo es una función creciente y cóncava (de números). Como se hará en la mayoría de las secciones que siguen, daremos versiones matriciales de ella. Pero en principio sólo para el caso $m = 2$ y $\lambda_1 = \lambda_2 = \frac{1}{2}$. Algunas páginas más adelante (Teorema 9.4.1), mostraremos una versión más general (sólo se asume que $m = 2$), que es también conocida como la desigualdad de Young. Igual damos una prueba de este caso, porque usa una técnica interesante que es bueno ver cómo funciona.

Proposición 9.3.1. *Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Entonces*

$$s_i(AB^*) \leq \frac{1}{2} \ s_i(A^*A + B^*B) \qquad \textit{para todo } i \in \mathbb{I}_n \quad .$$

Demostración. Sea $X = \begin{bmatrix} A & 0 \\ B & 0 \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})$. Cuentas elementales muestran que

$$X^*X = \begin{bmatrix} A^*A + B^*B & 0 \\ 0 & 0 \end{bmatrix} \qquad \text{y} \qquad XX^* = \begin{bmatrix} AA^* & AB^* \\ BA^* & BB^* \end{bmatrix} \ .$$

Sean $P = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$ y $U = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} = 2P - I_{2n} \in \mathcal{U}(2n)$. Luego

$$\widehat{AB^*} = \begin{bmatrix} 0 & AB^* \\ BA^* & 0 \end{bmatrix} = XX^* - C_P(XX^*) = \frac{1}{2} \left(XX^* - UXX^*U^* \right) \ .$$

Tomemos la descomposición $\widehat{AB^*} = \widehat{AB^*}^+ - \widehat{AB^*}^-$. Observar que ambas matrices XX^* y $UXX^*U^* \in \mathcal{M}_{2n}(\mathbb{C})^+$. Luego, la Eq. (3.7) y el ítem 5.b de la Sección 3.3 aseguran que

$$s_i(AB^*) = \mu_i\big(\widehat{AB^*}\big) = \mu_i\big(\widehat{AB^*}^+\big) \leq \frac{1}{2} \ \mu_i(XX^*) = \frac{1}{2} \ \mu_i(X^*X) = \frac{1}{2} \ s_i(A^*A + B^*B) \ ,$$

para todo $i \in \mathbb{I}_n$. ■

El siguiente resultado es más fino que el anterior, y no se generaliza tan fácilmente (salvo para la norma Frobenius, ver Teorema 9.4.7).

Proposición 9.3.2. *Sean $A, B, X \in \mathcal{M}_n(\mathbb{C})$. Entonces se tiene que*

$$|||AXB^*||| \leq \frac{1}{2}\ |||A^*AX + XB^*B||| \ , \tag{9.5}$$

para toda NUI $||| \cdot |||$ en $\mathcal{M}_n(\mathbb{C})$.

Demostración. Debemos dividir la prueba en tres pasos:
Paso 1: Supondremos que $A = B \in \mathcal{H}(n)$ y también $X \in \mathcal{H}(n)$. Por la Proposición 9.1.7,

$$|||AXA||| \leq |||\operatorname{Re} XA^2||| = \frac{1}{2}\ |||A^2X + XA^2||| \ .$$

Paso 2: Ahora X es cualquiera, pero $A, B \in \mathcal{H}(n)$: Tomemeos

$$T = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \in \mathcal{H}(2n) \quad \text{ e } \quad Y = \begin{bmatrix} 0 & X \\ X^* & 0 \end{bmatrix} \in \mathcal{H}(2n) \ .$$

Por el Paso 1, $s(TYT) \prec_w \frac{1}{2}\ s(T^2Y + YT^2)$. Pero cuentas fáciles muestran que

$$TYT = \begin{bmatrix} 0 & AXB \\ BX^*A & 0 \end{bmatrix} = \widehat{A^*XB} \quad \text{y análogamente} \quad T^2Y + YT^2 = \widehat{A^2X} + \widehat{XB^2} \ .$$

Luego podemos deducir que $s(AXB) \prec_w \frac{1}{2}\ s(A^2X + XB^2) = s\big(\frac{1}{2}\ [A^2X + XB^2]\big)$.

Paso 3: El caso general. Tomemos descomposiciones polares $A = U|A|$ y $B = V|B|$, con $U, V \in \mathcal{U}(n)$. Notar que $A^*AX + XB^*B = |A|^2X + X|B|^2$, mientras que

$$|||AXB^*||| = |||\ U\,|A|\ X\ |B|\,V^*\ ||| = |||\ |A|\ X\ |B|\ |||,$$

con lo que la desigualdad (9.5) queda demostrada en general a partir del Paso 2. ■

9.4 Desigualdades de Young para matrices

La desigualdad de Young clásica dice que si $a, b \in \mathbb{R}_+$ y $p, q \in [1, +\infty)$, entonces

$$\frac{1}{p} + \frac{1}{q} = 1 \implies ab \ \leq \ \frac{a^p}{p} + \frac{b^q}{q} \ , \tag{9.6}$$

con igualdad si y sólo si $a^p = b^q$. Observar que escrita así, es un refraseo de la desigualdad aritmético geométrica (9.4). Primero daremos una versión de (9.6) para valores singulares de matrices, que generaliza la Proposición 9.3.1.

Teorema 9.4.1 (Ando [19]). *Sean $p, q \in [1, +\infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$. Entonces para todo par de matrices $A, B \in \mathcal{M}_n(\mathbb{C})$ y todo $j \in \mathbb{I}_n$, se tiene que*

$$s_j\left(AB^*\right) \le s_j\left(\frac{|A|^p}{p} + \frac{|B|^q}{q}\right) \ , \tag{9.7}$$

o equivalentemente, que existe $U \in \mathcal{U}(n)$ tal que $U|AB^*|U^* \le \dfrac{|A|^p}{p} + \dfrac{|B|^q}{q}$.

Antes de demostrar el teorema, necesitamos algunos pasos técnicos:

Lema 9.4.2. *Sean $Q \in \mathcal{M}_n(\mathbb{C})$ una proyección ortogonal y $X \in \mathcal{M}_n(\mathbb{C})^+$. Entonces*

$$QX^rQ \le (QXQ)^r \quad \text{para } 0 < r \le 1 \quad \text{y} \quad QX^rQ \ge (QXQ)^r \quad \text{para } 1 \le r \le 2 \ .$$

Demostración. Puesto que $f(t) = t^r$ es cóncava de operadores para $r \in [0, 1]$ y convexa de operadores para $r \in [1, 2]$, este lema es un respaso de la Proposición 6.3.9. ■

El paso clave para probar el teorema, vía el cáculo de los autovalores con el principio minimax del capítulo 2, es el siguiente resultado tecniquísimo:

Lema 9.4.3. *Sean $p \in (1, 2]$ y $q \in [2, +\infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$. Sean $A \in \mathcal{M}_n(\mathbb{C})^+$, $B \in \mathcal{G}l(n)^+$ y $k \in \mathbb{I}_n$. Sean $\mathcal{B} = \{v_1, \dots, v_n\}$ una BON de $\mathbb{C}^n$ adaptada a $\mu(|AB|)$, y $\mathcal{S}_k =$ Gen$\{v_1, \dots, v_k\}$. Llamemos P al proyector ortogonal sobre $\mathcal{S}_k$ y Q al proyector ortogonal sobre $\mathcal{M} := R(B^{-1}P) = B^{-1}(\mathcal{S}_k)$. Si abreviamos $\mu = \mu_k(|AB|)$, se tiene que*

$$\mu\ Q \le \frac{QA^pQ}{p} + \frac{QB^qQ}{q} \quad . \tag{9.8}$$

Demostración. Por la definición de Q se tienen las siguientes igualdades:

$$QB^{-1}P = B^{-1}P \quad \text{y} \quad PB^{-1}Q = PB^{-1} \ . \tag{9.9}$$

Por otra parte, sabemos que $B(R(Q)\,) = B(\mathcal{M}) = \mathcal{S}_k = R(P)$, por lo que

$$PBQ = BQ \quad \text{y} \quad QBP = QB \ . \tag{9.10}$$

Luego, juntado esta última igualdad con (9.9)

$$(QB^2Q)(B^{-1}PB^{-1}) = QBPB^{-1} = Q\,.$$

Análogamente, se ve que $(B^{-1}PB^{-1})(QB^2Q) = Q$, lo cual muestra que la inversa de QB^2Q "dentro de $\mathcal{M}$" es $B^{-1}PB^{-1}$. Usando quien es $\mathcal{S}_k$, vemos que $\mu\,P \le |AB|$. Luego,

$$(BA^2B)^{1/2} = |AB| \ge \mu P \implies BA^2B \ge \mu^2 P \implies A^2 \ge \mu^2 B^{-1}PB^{-1} \ ,$$

donde vale elevar al cuadrado por que $|AB|$ y P conmutan. Como $p \in (1, 2]$, tenemos que la función $f(t) = t^{\frac{p}{2}}$ es monótona de operadores. Luego, usando (9.9), se ve que

$$A^p \ge \mu^p (B^{-1}PB^{-1})^{\frac{p}{2}} \implies QA^pQ \ge \mu^p Q(B^{-1}PB^{-1})^{\frac{p}{2}}Q = \mu^p (B^{-1}PB^{-1})^{\frac{p}{2}} \ .$$

Como QB^2Q es la inversa en $\mathcal{M}$ de $B^{-1}PB^{-1}$, y en $\mathcal{M}^\perp$ todo es nulo, tenemos que

$$QA^pQ \geq \mu^p(QB^2Q)^{\frac{-p}{2}} \qquad \text{(todo pensado en } L(\mathcal{M}) \text{) .} \tag{9.11}$$

Para probar (9.8), primeramente consideremos el caso $q \in [2,4]$. Por el Lema 9.4.2

$$QB^qQ \geq (QB^2Q)^{\frac{q}{2}} . \tag{9.12}$$

Luego, juntando (9.11) y (9.12) se tiene que

$$\begin{aligned} \frac{QA^pQ}{p} + \frac{QB^qQ}{q} &\geq \frac{\mu^p\ (QB^2Q)^{\frac{-p}{2}}}{p} + \frac{(QB^2Q)^{\frac{q}{2}}}{q} \\ &\overset{\star}{\geq} \mu\ (QB^2Q)^{\frac{-1}{2}}\ (QB^2Q)^{\frac{1}{2}} = \mu\ Q\ , \end{aligned}$$

donde $\overset{\star}{\geq}$ se puede probar usando la desigualdad de Young numérica, puesto que $\mu(QB^2Q)^{-1/2}$ y $(QB^2Q)^{1/2}$ conmutan entre sí. Esto concluye a demostración para este caso. Supongamos ahora que $q \in (4,\infty)$. Sea $s = \frac{q}{2}$. Entonces $0 < \frac{2}{s} < 1$, y $\frac{q}{s} = 2$. Por el Lema 9.4.2 se tiene

$$Q\ B^q\ Q = Q\ (B^s)^{\frac{q}{s}}\ Q \geq (Q\ B^s\ Q)^{\frac{q}{s}} \quad \text{y} \quad (Q\ B^s\ Q)^{\frac{2}{s}} \geq Q\ B^2\ Q\ . \tag{9.13}$$

Por lo tanto, usando que $f(t) = t^{\frac{p}{2}}$ es MOP, se tiene que

$$(QB^sQ)^{\frac{p}{s}} \geq (QB^2Q)^{\frac{p}{2}} \implies (QB^sQ)^{\frac{-p}{s}} \leq (QB^2Q)^{\frac{-p}{2}} \ \text{ en } L(\mathcal{M})\ .$$

Combinando esta desigualdad con (9.11) se obtiene

$$QA^pQ \geq \mu^p(QB^sQ)^{\frac{-p}{s}} ,$$

y luego combinándola con (9.13) resulta

$$\begin{aligned} \frac{QA^pQ}{p} + \frac{QB^qB}{q} &\geq \mu^p\ \frac{(QB^sQ)^{\frac{-p}{s}}}{p} + \frac{(QB^sQ)^{\frac{q}{s}}}{q} \\ &\geq \mu\ (QB^sQ)^{\frac{-1}{s}}\ (QB^sQ)^{\frac{1}{s}}\ =\ \mu Q\ , \end{aligned}$$

donde nuevamente en la segunda desigualdad se ha usado la versión numérica de la desigualdad de Young. ■

Demostración del Teorema 9.4.1: Probaremos la Eq. (9.7), mientras que la segunda formulación queda como ejercicio para el lector. Supongamos primero que $A, B \in \mathcal{M}_n(\mathbb{C})^+$. En tal caso tenemos que $|AB| = (BA^2B)^{1/2}$ y la ecuación (9.7) puede reescribirse como

$$\mu_j\left((BA^2B)^{1/2}\right) = \mu_j\left(BA^2B\right)^{1/2} \leq \mu_j\left(\frac{A^p}{p} + \frac{B^q}{q}\right) \quad \text{para todo} \quad j \in \mathbb{I}_n\ . \tag{9.14}$$

Como $\mu_j\,(BA^2B) = \mu_j\,(AB^2A)$ para todo $j \in \mathbb{I}_n$, los papeles de A y B son simétricos, razón por la cual podemos suponer que $p \in (1,2]$ y $q \in [2,\infty)$. Más aún, apelando a las tecnicas usuales de continuidad, podemos también asumir que $B > 0$. Dicho todo esto, fijemos $k \in \mathbb{I}_n$ y llamemos $\mu = \mu_k\left(BA^2B\right)^{1/2} = \mu_k(|AB|)$. Sean $\mathcal{B} = \{v_1, \dots, v_n\}$ una BON de $\mathbb{C}^n$ adaptada a $\mu(|AB|)$, y $\mathcal{S}_k = \operatorname{Gen}\{v_1, \dots, v_k\}$. Llamemos P al proyector ortogonal sobre $\mathcal{S}_k$ y Q al proyector ortogonal sobre $\mathcal{M} := R(B^{-1}P) = B^{-1}(\mathcal{S}_k)$. Entonces, el Lema 9.4.3 dice que

$$\mu\, Q \le \frac{QA^pQ}{p} + \frac{QB^qQ}{q} \quad \Longrightarrow \quad \min_{\substack{x\in\mathcal{M}\\ \|x\|=1}} \left\langle \left(\frac{A^p}{p} + \frac{B^q}{q}\right) x\,,x \right\rangle \ge \mu = \mu_k(|AB|)\ .$$

Observar que $\dim \mathcal{S}_k = \dim \mathcal{M} = k$. Luego, utilizando el principio minimax (Teorema 2.3.3) para calcular $\mu_k\left(\frac{A^p}{p} + \frac{B^q}{q}\right)$, la desigualdad (9.14) queda demostrada en este caso.

El caso general se deduce de lo anterior por la siguiente observación: Dadas $A, B \in \mathcal{M}_n(\mathbb{C})$, si hacemos la descomposición polar $B = V|B|$ con $V \in \mathcal{U}(n)$, se tiene que

$$|AB^*|^2 = BA^*AB^* = B|A|^2B^* = V\big(|B|\,|A|^2\,|B|\big)V^* = V\big|\,|A|\,|B|\,\big|^2V^*\ .$$

De ahí podemos deducir que los vectores $s(AB^*) = \mu(|AB^*|) = \mu(|A|\,|B|)$. Volviendo a mirar la Eq. (9.7) se ve que lo anterior hace suficiente probar el caso positivo, cosa que ya hicimos. ■

Ejercicio 9.4.4. Mostrar con un ejemplo que la desigualdad (9.7) deja de ser cierta si en el miembro izquierdo se quita la estrella en B. (Ayuda: basta considerar matrices de 2×2 y el caso $p = q = 2$). ▲

Corolario 9.4.5. *Sean $p, q \in [1, +\infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$ y sea N una NUI en $\mathcal{M}_n(\mathbb{C})$. Entonces para todo par de matrices $A, B \in \mathcal{M}_n(\mathbb{C})$, se tiene que*

$$N(AB^*) \le N\Big(\ \frac{|A|^p}{p} + \frac{|B|^q}{q}\ \Big)\ .$$

Demostración. Evidente a partir del Teorema 9.4.1, porque $\leqslant\ \Longrightarrow\ \prec_w$. ■

Las desigualdades de Hirzallah-Kittaneh

Cuando uno extiende desigualdades numéricas a matriciales, aparede un ingrediente nuevo: Dado que las matrices no conmutan, si uno multiplica por una tercera matrix X, cómo afecta esto a la desigualdad? Y de qué lado hay que multiplicar cada factor para que la desigualdad se preserve?

Este tipo de análisis y generalizaciones aparecerán seguido en lo que resta del Capítulo. Daremos a continuación una primera versión que camina para Young, aunque solamente para la norma de Frobenius. Lo primero que apareció al respecto, es un resultado de Bhatia y Davis [22]: Dados $A, B \in \mathcal{M}_n(\mathbb{C})^+$, $X \in \mathcal{M}_n(\mathbb{C})$, y p, q conjugados, se tiene que

$$\|AXB\|_2 \ \le\ \Big\|\ \frac{A^pX}{p} + \frac{XB^q}{q}\ \Big\|_2\ . \tag{9.15}$$

Sin embargo, mostraremos un resultado más fino, debido a Hirzallah y Kittaneh [26], que además determina completamente los casos en que en (9.15) pudiera darse la igualdad. Para ello, comenzamos por demostrar unas desigualdades numéricas:

Lema 9.4.6. *Sean $a, b \in \mathbb{R}_+^*$, y $p, q \in [1, +\infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$. Si $r = \text{máx}\{p, q\}$, entonces*

$$\left(\frac{a^p}{p} + \frac{b^q}{q}\right)^2 \geq \frac{1}{r^2}\ (a^p - b^q)^2 + a^2 b^2. \tag{9.16}$$

Demostración. Primero observar que si $p = q = 2$, tenemos en realidad una igualdad. Asumamos que $q = r > p$, o sea que $q \in [2, +\infty)$. Vía cuentas elementales, se ve que

$$\left(\frac{a^p}{p} + \frac{b^q}{q}\right)^2 - \frac{1}{q^2}\,(a^p - b^q)^2 = a^p \left(\left(\ 1 - \frac{2}{q}\ \right)\ a^p + \frac{2}{q}\ b^q\right)\ ,$$

donde se usa la igualdad $1 - \frac{2}{q} = \frac{1}{p} - \frac{1}{q} = \left(\frac{1}{p} - \frac{1}{q}\right)\left(\frac{1}{p} + \frac{1}{q}\right) = \frac{1}{p^2} - \frac{1}{q^2}$. Ahora, usando la desigualdad de Young clásica (o la aritmético geométrica), tenemos que

$$\left(1 - \frac{2}{q}\right) a^p + \frac{2}{q}\ b^q \geq a^{p\,(1 - \frac{2}{q}\,)}\ b^{q\,\frac{2}{q}} = a^{\frac{q-p}{q}}\ b^2\ ,$$

ya que $p(1 - \frac{2}{q}\,) = p(\,\frac{1}{p} - \frac{1}{q}\,) = 1 - \frac{p}{q}$. Además, usando el hecho de que $\frac{p}{q} = p - 1$, vemos que $p + (\frac{q-p}{q}) = 2$. Así llegamos finalmente a que

$$\left(\frac{a^p}{p} + \frac{b^q}{q}\right)^2 - \frac{1}{q^2}\,(a^p - b^q)^2 \geq a^p\ a^{\frac{q-p}{q}}\ b^2\ =\ a^2 b^2 \quad \Longrightarrow \quad \text{Eq. (9.16)}\ .$$

De manera análoga, si $p > q$, queda que $\left(\dfrac{a^p}{p} + \dfrac{b^q}{q}\right)^2 \geq \dfrac{1}{p^2}\,(a^p - b^q)^2 + a^2 b^2$. ■

Teorema 9.4.7. *Sean $A, B \in \mathcal{M}_n(\mathbb{C})^+$, $X \in \mathcal{M}_n(\mathbb{C})$, y $p, q \in [1, +\infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$. Si $r = \text{máx}\{p, q\}$, entonces*

$$\|AXB\|_2^2\ +\ \frac{1}{r^2}\,\|A^pX - XB^q\|_2^2\ \leq\ \left\|\ \frac{A^pX}{p} + \frac{XB^q}{q}\ \right\|_2^2 . \tag{9.17}$$

Demostración. Como $A, B \in \mathcal{M}_n(\mathbb{C})^+$, existen $U, V \in \mathcal{U}(n)$ tales que $A = UD_1U^*$ y $B = VD_2V^*$, donde $D_1 = \text{diag}(\mu(A)\,)$ y $D_2 = \text{diag}\,(\mu(B)\,)$. Llamemos $Y = U^*XV$, $\lambda = \mu(A)$ y $\mu = \mu(B)$. Entonces tenemos que

$$\begin{aligned}
\frac{A^pX}{p} + \frac{XB^q}{q} \ &= \ \frac{U\,D_1^p\,U^*X}{p} + \frac{X\,V\,D_2^q\,V^*}{q} \ = \ U\left(\frac{D_1^p\,Y}{p} + \frac{Y\,D_2^q}{q}\right)V^* \\
&= \ U\ \left[\left(\frac{\lambda_i^p}{p} + \frac{\mu_j^q}{q}\right) y_{ij}\right]_{i,j\in\mathbb{I}_n} V^*,
\end{aligned}$$

$$A^pX - XB^q = U\,(D_1^pY - YD_2^q)\,V^* = U\ \left[\ \left(\lambda_i^p - \mu_j^q\right) y_{ij}\right]_{i,j\in\mathbb{I}_n} V^* \quad \text{y}$$

$$AXB = U(D_1 Y D_2)V^* = U\Big[\lambda_i \ \mu_j \ y_{ij}\Big]_{i,j\in\mathbb{I}_n} V^*.$$

Por la desigualdad (9.16) aplicada a cada par $a = \lambda_i$ y $b = \mu_j$, tenemos que

$$\begin{aligned}\left\| \frac{A^pX}{p} + \frac{XB^q}{q}\right\|_2^2 &= \sum_{i,j=1}^n \left(\frac{\lambda_i^p}{p} + \frac{\mu_j^q}{q}\right)^2 |y_{ij}|^2 \\ &\geq \frac{1}{r^2}\sum_{i,j=1}^n \left(\lambda_i^p - \mu_j^q\right)^2 |y_{ij}|^2 + \sum_{i,j=1}^n \lambda_i^2\mu_j^2\,|y_{ij}|^2 \\ &= \frac{1}{r^2}\,\|A^pX - XB^q\|_2^2 + \|AXB\|^2 \ ,\end{aligned}$$

lo que completa la prueba del teorema. ■

Observación 9.4.8. Para el caso $p = q = 2$, la desigualdad (9.17) es en realidad una igualdad. Esto se observa en la demostración del Teorema 9.4.7, ya que la desigualdad que se presenta allí es en realidad una igualdad, como se observó en el Lema 9.4.6. ▲

A partir de lo anterior pueden encontrarse condiciones necesarias y suficientes para que se satisfagan las igualdades en (9.15) y (9.7), como se demuestra a continuación.

Corolario 9.4.9. *Sean A, B, X, p y q como en el Teorema 9.4.7. Se tiene que*

$$\left\| \frac{A^pX}{p} + \frac{XB^q}{q}\right\|_2 = \|AXB\|_2 \quad \Longleftrightarrow \quad A^pX = XB^q \ . \tag{9.18}$$

Demostración. Si se tiene que $\left\| \frac{A^pX}{p} + \frac{XB^q}{q}\right\|_2 = \|AXB\|_2$, la desigualdad (9.17) asegura que que $\|A^pX - XB^q\|_2 = 0$, o sea que $A^pX = XB^q$.

Asumamos que $A^pX = XB^q$. Como antes, tenemos $A = UD_1U^*$, $B = VD_2V^*$, donde $D_1 = \operatorname{diag}(\mu(A))$ y $D_2 = \operatorname{diag}(\mu(B))$. Luego

$$A^pX = UD_1^pU^*X = U\left[D_1^p\,(U^*XV)\right]V^* \quad \text{y} \quad XB^q = XVD_2^qV^* = U\left[(U^*XV)\,D_2^q\right]V^* \ ,$$

con lo cual, llamando $Y = U^*XV$, tendremos que $D_1^p\ Y = Y\ D_2^q$, es decir que

$$\lambda_i^p\ y_{ij} \ = \ y_{ij}\ \mu_j^q \implies \lambda_i\ y_{ij} = y_{ij}\ \mu_j^{\frac{q}{p}} \quad \text{para todos los } i, j \in \mathbb{I}_n \ .$$

Llegamos a que $D_1\ Y = Y\ D_2^{\frac{q}{p}}$, y por ende $AX = XB^{\frac{q}{p}}$. Así,

$$AXB = XB^{\frac{q}{p}}\ B = XB^q = \frac{XB^q}{p} + \frac{XB^q}{q} = \frac{A^pX}{p} + \frac{XB^q}{q} \ .$$

Vemos que vale la igualdad entre las matrices, que implica la igualdad de sus normas. ■

Corolario 9.4.10. *Sean* $p, q \in [1, +\infty)$ *tales que* $\frac{1}{p} + \frac{1}{q} = 1$. *Entonces*

$$s_j\left(\frac{A^pX}{p} + \frac{XB^q}{q}\right) = s_j(AB) \quad \text{para todo } j \in \mathbb{I}_n \qquad \Longleftrightarrow \qquad A^p = B^q \ .$$

Demostración. Es consecuencias del Corolario 9.4.9 aplicado a $X = I$. Observar que la condición $A^p = B^q$ implicaba igualdad de matrices. ■

Corolario 9.4.11. *Sean* $p, q \in [1, +\infty)$ *tales que* $\frac{1}{p} + \frac{1}{q} = 1$. *Entonces*

$$U\left(\frac{A^pX}{p} + \frac{XB^q}{q}\right)U^* = |AB| \quad \text{para alguna } U \in \mathcal{U}(n) \qquad \Longleftrightarrow \qquad A^p = B^q \ .$$

Demostración. Es lo mismo que el Corolario 9.4.10. ■

Observación 9.4.12. Sean $a, b \in \mathbb{R}_+^*$, y $p, q \in [1, +\infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$. Luego

$$\left(\frac{a^p}{p} + \frac{b^q}{q}\right)^2 \ \geq \ \left(\frac{a^p}{p} + \frac{b^q}{q} - ab\right)^2 + a^2b^2 \quad , \tag{9.19}$$

$$\left(\frac{a^p}{p} + \frac{b^q}{q} + ab\right)^2 \ \geq \ \left(\frac{a^p}{p} + \frac{b^q}{q} - ab\right)^2 + 4a^2b^2 \quad \text{y} \tag{9.20}$$

$$\frac{1}{s}\,|a^p - b^q| \ \geq \ \frac{a^p}{p} + \frac{b^q}{q} - ab \ , \tag{9.21}$$

donde $s = \text{mín}\{p, q\}$. En efecto, para probar (9.19) recordamos que $\alpha = \frac{a^p}{p} + \frac{b^q}{q} \geq ab$, con lo cual $\alpha^2 \geq \alpha^2 - 2\alpha ab + 2a^2b^2 = (\alpha - ab)^2 + a^2b^2$. También, $(\alpha + ab)^2 - (\alpha - ab)^2 = 4\alpha ab \geq 4a^2b^2$, y obtenemos (9.20). Para obtener (9.21), si $a^p \geq b^q$, también tenemos $a \geq b^{q/p}$, y como $1/s - 1/p \geq 0$,

$$\left(\frac{1}{s} - \frac{1}{p}\right)a^p + ab \geq \left(\frac{1}{s} - \frac{1}{p}\right)b^q + b^{q/p}b = \left(\frac{1}{s} + \frac{1}{q}\right)b^q \implies$$

$$\implies \frac{1}{s}\,|a^p - b^q| \geq \frac{a^p}{p} + \frac{b^q}{q} - ab.$$

Si $b^q \geq a^p$, la demostración es análoga. Partiendo de estas desigualdades y siguiendo el razonamiento de la demostración de (9.17), podemos mostrar que si $A, B \in \mathcal{M}_n(\mathbb{C})^+$, $X \in \mathcal{M}_n(\mathbb{C})$, y $p, q \in [1, +\infty)$ cumplen que $\frac{1}{p} + \frac{1}{q} = 1$, entonces

$$\left\|\frac{A^pX}{p} + \frac{XB^q}{q}\right\|_2^2 \ \geq \ \left\|\frac{A^pX}{p} + \frac{XB^q}{q} - ABX\right\|_2^2 + \|AXB\|_2^2 \ , \tag{9.22}$$

$$\left\|\frac{A^pX}{p} + \frac{XB^q}{q} + ABX\right\|_2^2 \ \geq \ \left\|\frac{A^pX}{p} + \frac{XB^q}{q} - ABX\right\|_2^2 + \|AXB\|_2^2 \ , \tag{9.23}$$

$$\frac{1}{s}\,\|A^pX - XB^q\|_2 \ \geq \ \left\|\frac{A^pX}{p} + \frac{XB^q}{q} - AXB\right\|_2 \ , \tag{9.24}$$

donde $s = \text{mín}\{p, q\}$. Notar que (9.22) también es más fuerte que (9.15). Además, (9.24) puede usarse para demostrar de manera directa los corolarios anteriores. ▲

Ejercicio 9.4.13. Sean $A, B, C, D \in \mathcal{M}_n(\mathbb{C})^+$ tales que $AD = DA$ y $BC = CB$. Sean $X \in \mathcal{M}_n(\mathbb{C})$ y $p, q \in [1, +\infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$. Entonces

$$\left\| \frac{1}{p} A^p X C^p + \frac{1}{q} D^q X B^q \right\|_2^2 \geq \frac{1}{r^2} \|A^p X C^p - D^q X B^q\|_2^2 + \|ADXCB\|_2^2 \ . \tag{9.25}$$

donde $r = \text{máx}\{p, q\}$. En particular, si $C = D = I$ se recupera (9.17). ▲

Ejercicio 9.4.14. Sean $A, B, C, D \in \mathcal{M}_n(\mathbb{C})$ tales que $AD = DA$, $A^*D = DA^*$, $BC = CB$ y $B^*C = CB^*$. Sean $X \in \mathcal{M}_n(\mathbb{C})$ y $p, q \in [1, +\infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$. Entonces

$$\left\| \frac{1}{p} |A|^p X |C|^p + \frac{1}{q} |D|^q X |B|^q \right\|_2^2 \geq \frac{1}{r^2} \| |A|^p X |C|^p - |D|^q X |B|^q \|_2^2 + \|ADXC^*B^*\|_2^2 \ .$$

donde $r = \text{máx}\{p, q\}$. En particular, si $C = D = I$, tenemos

$$\left\| \frac{1}{p} |A|^p X + \frac{1}{q} X |B|^q \right\|_2^2 \geq \frac{1}{r^2} \| |A|^p X - X |B|^q \|_2^2 + \|AXB^*\|_2^2 \,, \tag{9.26}$$

que es la extensión natural de (9.17) a matrices cualesquiera. Extender las desigualdades (9.22)-(9.24) de manera análoga. ▲

9.5 Desigualdades tipo Hölder para matrices

Observación 9.5.1 (Caso numérico). . Dados $p, q \in (1, +\infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$, la forma más simple de la desigualdad numérica de Hölder se escribe

$$(|a|^p + |b|^p)^{\frac{1}{p}} (|c|^q + |d|^q)^{\frac{1}{q}} \geq |ac + bd| \ , \tag{9.27}$$

para todo cuarteto $a, b, c, d \in \mathbb{C}$. Más aún, podemos escribir:

$$(|a|^p + |b|^p)^{\frac{1}{p}} = \max\{|ac + bd| : |c|^q + |d|^q = 1\} \ , \tag{9.28}$$

que tiene que ver con la frase "el dual de ℓ^p es ℓ^q." ▲

Proposición 9.5.2. *Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Entonces se verifica que*

$$(A^*A + B^*B)^{1/2} = \text{máx}\left\{ |C^*A + D^*B| \ : \ C, D \in \mathcal{M}_n(\mathbb{C}) \ \text{ y } \ C^*C + D^*D \leq I \right\}. \tag{9.29}$$

Demostración. Dadas $C, D \in \mathcal{M}_n(\mathbb{C})$ matrices cualesquiera, tenemos que

$$\begin{bmatrix} A^*A + B^*B & A^*C + B^*D \\ C^*A + D^*B & C^*C + D^*D \end{bmatrix} = \begin{bmatrix} A & C \\ B & D \end{bmatrix}^* \begin{bmatrix} A & C \\ B & D \end{bmatrix} \geq 0 \ .$$

Recordemos que la Proposición 3.8.6 asegura que, dadas $X \in \mathcal{M}_n(\mathbb{C})^+$ e $Y \in \mathcal{M}_n(\mathbb{C})$,

$$\begin{bmatrix} X & Y^* \\ Y & I \end{bmatrix} \geq 0 \iff X \geq Y^*Y \geq 0 \,. \tag{9.30}$$

Por lo tanto, si $C^*C + D^*D \leq I$, tenemos que $A^*A + B^*B \geq |C^*A + D^*B|^2$. Usando que $f(t) = t^{1/2}$ es monótona de operadores, llegamos a que

$$(A^*A + B^*B)^{1/2} \geq |C^*A + D^*B| \,. \tag{9.31}$$

Más aún, cuando $A^*A + B^*B \in \mathcal{G}l(n)^+$, si consideramos

$$C = A(A^*A + B^*B)^{-1/2} \quad \text{y} \quad D = B(A^*A + B^*B)^{-1/2} \,,$$

obtenemos una igualdad en la Eq. (9.31). Cuando $A^*A + B^*B \notin \mathcal{G}l(n)^+$, observar que

$$\mathcal{S} = \ker(A^*A + B^*B) = \ker(A^*A) \cap \ker(B^*B) = \ker A \cap \ker B \,.$$

Sean $A_1 = A|_{\mathcal{S}^\perp} \oplus I|_{\mathcal{S}}$, $B_1 = B|_{\mathcal{S}^\perp} \oplus I|_{\mathcal{S}}$. Luego, $A_1^*A_1 + B_1^*B_1 \in \mathcal{G}l(n)^+$. Tomando

$$C = A_1(A_1^*A_1 + B_1^*B_1)^{-1/2} \,, \ D = B_1(A_1^*A_1 + B_1^*B_1)^{-1/2} \,,$$

una cuenta fácil mustra que también obtenemos una igualdad en la Eq. (9.31). ■

Lema 9.5.3. *Sea $A, B \in \mathcal{M}_n(\mathbb{C})^+$, $p \in (1, +\infty)$ y $\alpha \in [0, 1]$. Entonces se tiene que*

$$(A^p + B^p)^{\frac{1}{p}} \geq \alpha^{1-\frac{1}{p}} \ A + (1-\alpha)^{1-\frac{1}{p}} \ B \,.$$

Demostración. Cuando $\alpha = 0$, tenemos que $A^p \geq 0$, con lo cual $A^p + B^p \geq B^p$. Luego, como $f(t) = t^{1/p}$ es monótona de operadores, sale que $(A^p + B^p)^{1/p} \geq B$. Caso análogo ocurre para $\alpha = 1$. Consideremos entonces $0 < \alpha < 1$, y llamemos $\beta = 1 - \alpha$. Usando que $f(t) = t^{1/p}$ también es cóncava de operadores, tenemos que

$$(A^p + B^p)^{\frac{1}{p}} = \left[\alpha \left(\alpha^{\frac{-1}{p}} A\right)^p + \beta \left(\beta^{\frac{-1}{p}} B\right)^p\right]^{\frac{1}{p}} \geq \alpha^{1-\frac{1}{p}} \ A + \beta^{1-\frac{1}{p}} \ B \,.$$ ■

Teorema 9.5.4. *Sean A, B, C y $D \in \mathcal{M}_n(\mathbb{C})$, $p, q \in [2, +\infty)$ y $r \in (1, +\infty]$ que cumplen la ecuación $\frac{1}{p} + \frac{1}{q} = 1 - \frac{1}{r}$. Para todo $\alpha \in (0, 1)$, si llamamos $\beta = 1 - \alpha$, se tiene que*

$$|C|^q + |D|^q \leq I \implies \Big(|A|^p + |B|^p\Big)^{\frac{2}{p}} \geq \left|\alpha^{\frac{1}{r}} \ C^*A + \beta^{\frac{1}{r}} \ D^*B\right|^2 \,. \tag{9.32}$$

Demostración. El caso $r = \infty$ (o sea, $p = q = 2$), fue visto en la Proposición 9.5.2, con lo cual asumimos $r < \infty$. Dado que $(1/2 - 1/p) + (1/2 - 1/q) = 1/r$, tenemos que:

$$\begin{bmatrix} \alpha^{1-\frac{2}{p}} A^*A + \beta^{1-\frac{2}{p}} B^*B & \alpha^{\frac{1}{r}} A^*C + \beta^{\frac{1}{r}} B^*D \\ \alpha^{\frac{1}{r}} C^*A + \beta^{\frac{1}{r}} D^*B & \alpha^{1-\frac{2}{q}} C^*C + \beta^{1-\frac{2}{q}} D^*D \end{bmatrix} =$$

$$\begin{bmatrix} \alpha^{\frac{1}{2}-\frac{1}{p}} A & \alpha^{\frac{1}{2}-\frac{1}{q}} C \\ \beta^{\frac{1}{2}-\frac{1}{p}} B & \beta^{\frac{1}{2}-\frac{1}{q}} D \end{bmatrix}^* \begin{bmatrix} \alpha^{\frac{1}{2}-\frac{1}{p}} A & \alpha^{\frac{1}{2}-\frac{1}{q}} C \\ \beta^{\frac{1}{2}-\frac{1}{p}} B & \beta^{\frac{1}{2}-\frac{1}{q}} D \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})^+ \,.$$

De acuerdo al Lema 9.5.3, tenemos que

$$\alpha^{1-\frac{2}{p}} A^*A + \beta^{1-\frac{2}{p}} B^*B \le (|A|^p + |B|^p)^{\frac{2}{p}} \quad \text{y} \quad \alpha^{1-\frac{2}{q}} C^*C + \beta^{1-\frac{2}{q}} D^*D \le (|C|^q + |D|^q)^{\frac{2}{q}} \ ,$$

con lo cual concluimos que $\begin{bmatrix} (|A|^p + |B|^p)^{\frac{2}{p}} & \alpha^{\frac{1}{r}} A^*C + \beta^{\frac{1}{r}} B^*D \\ \alpha^{\frac{1}{r}} C^*A + \beta^{\frac{1}{r}} D^*B & (|C|^q + |D|^q)^{\frac{2}{q}} \end{bmatrix} \ge$

$$\ge \begin{bmatrix} \alpha^{1-\frac{2}{p}} A^*A + \beta^{1-\frac{2}{p}} B^*B & \alpha^{\frac{1}{r}} A^*C + \beta^{\frac{1}{r}} B^*D \\ \alpha^{\frac{1}{r}} C^*A + \beta^{\frac{1}{r}} D^*B & \alpha^{1-\frac{2}{p}} C^*C + \beta^{1-\frac{2}{p}} D^*D \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})^+ \ .$$

Usando la Eq. (9.30) y el hecho de que $|C^*|^q + |D^*|^q \le I$, estamos hechos. ■

Definición 9.5.5. Dadas $C, D \in \mathcal{M}_n(\mathbb{C})^+$, escribimos $C \preccurlyeq D$ si $C^m \le D^m$, para todo $m \in \mathbb{N}$. Esta relación es un orden parcial denominado **orden espectral**.

Proposición 9.5.6. *Sean $A, B \in \mathcal{M}_n(\mathbb{C})^+$. Entonces se verifica:*

1. *La función $[1\,, +\infty) \ni p \longmapsto \left(\frac{A^p+B^p}{2}\right)^{\frac{1}{p}}$ es creciente (relativa al orden $\le$ de $\mathcal{M}_n(\mathbb{C})^+$).*
2. *El siguiente límite existe y da el $\preccurlyeq$-supremo de A y B:*

$$A \vee B = \lim_{p\to\infty} \left(A^p + B^p\right)^{\frac{1}{p}} = \min_{\preccurlyeq} \left\{ \ C \in \mathcal{M}_n(\mathbb{C})^+ : A \preccurlyeq C \ \ y \ \ B \preccurlyeq C \ \right\} \ .$$

Demostración. Sean $r, q \in [1\,, +\infty)$ tales que $r < q$. Si aplicamos el Lema 9.5.3 para los números $\alpha = \frac{1}{2}$ y $p = \dfrac{q}{r} > 1$, resulta que

$$(A^q + B^q)^{\frac{r}{q}} = \left((A^r)^{\frac{q}{r}} + (B^r)^{\frac{q}{r}}\right)^{\frac{r}{q}} \ge \frac{1}{2^{1-\frac{r}{q}}} \ \left(A^r + B^r\right) \ .$$

Luego, usando que $t^{\frac{1}{r}}$ es MOP y multiplicando por $2^{\frac{-1}{q}}$, llegamos a que

$$\left(\frac{A^q + B^q}{2}\right)^{\frac{1}{q}} \ge \left(\frac{A^r + B^r}{2}\right)^{\frac{1}{r}} \ .$$

Sea $M = \max\{\mu_1(A), \mu_1(B)\}$. Para cualquier $p \in [1, +\infty)$ tenemos que

$$M^p \ge \max\{\mu_1(A^p), \mu_1(B^p)\} = \max\{\|A^p\|_{sp}\,, \|B^p\|_{sp}\} \ge \Big\|\frac{A^p + B^p}{2}\Big\|_{sp} \ ,$$

y por lo tanto $M\,I \ge \left(\frac{A^p+B^p}{2}\right)^{\frac{1}{p}}$. En resumen, la función $p \longmapsto \left(\frac{A^p+B^p}{2}\right)^{\frac{1}{p}}$ es creciente y acotada superiormente. Aplicando el Ejercicio 3.9.16 vemos que debe existir el límite

$$A \vee B = \sup_{p\in[1\,,+\infty)} \left(\frac{A^p + B^p}{2}\right)^{\frac{1}{p}} = \lim_{p\to\infty} \left(\frac{A^p + B^p}{2}\right)^{\frac{1}{p}} = \lim_{p\to\infty} \left(A^p + B^p\right)^{\frac{1}{p}} \ .$$

Fijemos $m \in \mathbb{N}$. Tomando $p > m$, se tiene que $t^{\frac{m}{p}}$ es MOP, y así

$$A^m = (A^p)^{\frac{m}{p}} \leq (A^p + B^p)^{\frac{m}{p}} = \left((A^p + B^p)^{\frac{1}{p}}\right)^m \xrightarrow[p\to\infty]{} (A \vee B)^m \ .$$

Análogamente se muestra que $(A \vee B)^m \geq B^m$. Esto vale para cualquier $m \in \mathbb{N}$, y así $A, B \preccurlyeq A \vee B$. Sea ahora $C \in \mathcal{M}_n(\mathbb{C})^+$ tal que $A, B \preccurlyeq C$. Para cada par $m, k \in \mathbb{N}$ tenemos que $A^{km}, B^{km} \leq C^{km}$, y usando que $t^{\frac{1}{k}}$ es MOP, resulta que $\left(\frac{A^{km}+B^{km}}{2}\right)^{1/k} \leq C^m$. Luego

$$(A \vee B)^m = \lim_{p\to\infty} \left(\frac{A^p + B^p}{2}\right)^{\frac{m}{p}} = \lim_{k\to\infty} \left(\frac{A^{km} + B^{km}}{2}\right)^{\frac{1}{k}} \leq C^m \ , \quad \text{para todo } m \in \mathbb{N} \ .$$

Así, $A \vee B \preccurlyeq C$. Luego, $A \vee B$ es el supremo de A, B con respecto al orden espectral. ■

Corolario 9.5.7. *Sean $A, B, C, D \in \mathcal{M}_n(\mathbb{C})$. Fijado $p \in [2, \infty)$, tomemos el $q \in (1, 2]$ tal que se tenga la igualdad $\frac{1}{p} + \frac{1}{q} = 1$. Entonces*

$$\left(\|C\|^q + \|D\|^q\right)^{\frac{2}{q}} \left(|A|^p + |B|^p\right)^{\frac{2}{p}} \geq |CA + DB|^2 \ . \tag{9.33}$$

Por lo tanto, se tiene que $\left(\|C\| + \|D\|\right)^2 \left(|A| \vee |B|\right)^2 \geq |CA + DB|^2$.

Demostración. Fijemos $p \in [2, \infty)$. Supongamos en principio que $\|C\|, \|D\| \leq 1$. En tal caso, se tiene que $|C^*| \leq I \implies |C^*| \preccurlyeq I$ y lo mismo para D. Por la Proposición 9.5.6,

$$\left(\frac{|C^*|^t + |D^*|^t}{2}\right)^{\frac{1}{t}} \leq |C^*| \vee |D^*| \leq I \quad \text{ para todo } \quad t \in [1, +\infty) \ . \tag{9.34}$$

Vamos a aplicar el Teorema 9.5.4. Para ello, tomemos $t \in [2, +\infty)$ y $r \in (1, +\infty]$ tales que $\frac{1}{p} + \frac{1}{t} = 1 - \frac{1}{r}$. Dados $\alpha \in [0, 1]$ y $\beta = 1 - \alpha$, la Eq. (9.34) y el teorema citado aseguran que

$$\left(|A|^p + |B|^p\right)^{\frac{2}{p}} \geq 2^{\frac{1}{t}} \ \left|\alpha^{\frac{1}{r}} \ CA + \beta^{\frac{1}{r}} \ DB\right|^2 \ .$$

Haciendo $t \to \infty$ (con lo cual $r \to q$), tenemos que $(|A|^p + |B|^p)^{\frac{2}{p}} \geq \left|\alpha^{\frac{1}{q}} CA + \beta^{\frac{1}{q}} DB\right|^2$.

Si ahora suponemos más que antes: que $\|C\|^q + \|D\|^q = 1$, entonces podemos elegir $\alpha = \|C\|^q$ y $\beta = \|D\|^q = 1 - \alpha$. Reemplazando $C' = \alpha^{\frac{-1}{q}} C$ y $D' = \beta^{\frac{-1}{q}} D$, queda que $\|C'\| = \|D'\| = 1$ (si alguna era nula, su prima también la hacemos nula). Por el caso anterior obtenemos que

$$\|C\|^q + \|D\|^q = 1 \implies (|A|^p + |B|^p)^{2/p} \geq \left|\alpha^{\frac{1}{q}} C'A + \beta^{\frac{1}{q}} D'B\right|^2 = |CA + DB|^2 \ . \tag{9.35}$$

En el caso general, dadas $C, D \in \mathcal{M}_n(\mathbb{C})$ (alguna de las dos no nulas), consideramos

$$E = (\|C\|^q + \|D\|^q)^{-1/q} \, C \quad \text{ y } \quad F = (\|C\|^q + \|D\|^q)^{-1/q} \, D \ ,$$

con lo cual $\|E\|^q + \|F\|^q = 1$. Por la Eq. (9.35) nos queda que

$$(|A|^p + |B|^p)^{\frac{2}{p}} \geq |EA + FB|^2 = \frac{|CA + DB|^2}{(\|C\|^q + \|D\|^q)^{\frac{2}{q}}} \implies \quad \text{la Eq. (9.33)} \ .$$

La otra desigualdad se obtiene de lo anterior, considerando el límite para $p \to \infty$. ■

9.6 La técnica alternativa

En esta sección repasaremos y mezclaremos una serie de definiciones y resultados de los Capítulos 4 y 7 que usaremos seguido en el resto de este Capítulo. Primero enunciamos los relativos a las propiedades espectrales de los productos alternados de matrices (Sección 7.3):

Teorema 9.6.1. Sea $A \in \mathcal{M}_n(\mathbb{C})$ con autovalores $\lambda_1(A), \dots, \lambda_n(A)$. Sea $k \in \mathbb{I}_n$. Entonces los autovalores de $\Lambda^k A$ están dados por

$$\lambda_J(\Lambda^k A) = \prod_{i \in J} \lambda_i(A) , \quad J \in Q_{k,n} ,$$

contados con multiplicidad. ■

Corolario 9.6.2. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Sea $k \in \mathbb{I}_n$. Entonces los valores singulares de $\Lambda^k A$ son

$$s\big(\Lambda^k A\big) = \left\{ \ s_J\big(\Lambda^k A\big) \ \right\}_{J \in Q_{k,n}} = \left\{ \ \prod_{i \in J} s_i(A) \ \right\}_{J \in Q_{k,n}} ,$$

contados con multiplicidad, y ordenados en forma decreciente. Además, si ordenamos a los autovalores $\lambda_1(A), \dots, \lambda_n(A)$ de A con módulos decrecientes, se tiene que

$$\rho\big(\Lambda^k A\big) = \prod_{i=1}^{k} |\lambda_i(A)| \quad \text{y} \quad \big\|\Lambda^k A\big\|_{sp} = s_1\big(\Lambda^k A\big) = \prod_{i=1}^{k} s_i(A) . \qquad ■$$

Ahora vienen los resultados que relacionan la mayorización logarítmica con la usual (ver Sección 4.4): Recordemos que, dados $x, y \in \mathbb{R}^n_+$, escribimos $x \underset{\log}{\prec_w} y$ si

$$\prod_{i=1}^{k} x_i^{\downarrow} \ \le \ \prod_{i=1}^{k} y_i^{\downarrow} \quad \text{para todo} \quad k \in \mathbb{I}_n . \tag{9.36}$$

Si $x > 0$ e $y > 0$, escribimos $x \underset{\log}{\prec} y$ si $x \underset{\log}{\prec_w} y$ y, además, $\prod_{i=1}^{n} x_i = \prod_{i=1}^{n} y_i$.

Observación 9.6.3. Sean $x, y \in \mathbb{R}^n$ tales que $x > 0$ e $y > 0$. Si $x \underset{\log}{\prec} y$ entonces, como en el caso de la mayorización común, se cumplen desigualdades invesas para las entradas mas pequeñas de x e y. Es decir que

$$\prod_{i=k}^{n} x_i^{\downarrow} \ \ge \ \prod_{i=k}^{n} y_i^{\downarrow} \quad \text{para todo} \quad k \in \mathbb{I}_n . \tag{9.37}$$

Tener cuidado, que esto no vale si sólo se tiene que $x \underset{\log}{\prec_w} y$, porque se usa la igualdad de los productos hasta n. ▲

Proposición 9.6.4. Sean $x, y \in \mathbb{R}_+^n$.

1. Si $x \underset{\log}{\prec_w} y$, entonces $x^p \prec_w y^p \quad$ para todo $\quad p \in \mathbb{R}_+$.

2. Si $x > 0$ e $y > 0$, entonces $x \underset{\log}{\prec} y$ implica que $x^p \prec_w y^p$ para todo $p \in \mathbb{R}$. ■

Observación 9.6.5. 1. El caso más usado de la Proposición 9.6.4 es cuando $p = 1$. Es decir que, si $x, y \in \mathbb{R}_+^n$, entonces $x \underset{\log}{\prec_w} y$ implica $x \prec_w y$. Esto será sumamente util cuando se lo aplique a desigualdades con valores singulares de matrices, usando técnicas de productos alternados. Observar que, en este caso, el Corolario 4.2.3 nos dice que, si hubese quedado $x \prec y$, debía cumplirse que $x \underset{\log}{\prec} y$.

2. Por otra parte, la Proposición 9.6.4 se puede generalizar, sin cambiar la prueba, si remplazamos las funciones $f(t) = t^p$ por cualquier función $f : \mathbb{R}_+ \to \mathbb{R}$ tal que la aplicación $t \mapsto f(e^t)$ resulte convexa (o convexa *creciente*). Notar que, en el caso demostrado, se usaba esa propiedad para la función $t \mapsto (e^t)^p = e^{pt}$. ▲

9.7 Primeras aplicaciones

En esta sección veremos tres desigualdades importantes que se prueban en forma directa usando la técnica de extrapolar una desigualdad conocida, pero aplicada a los productos alternados, y luego deducir una mayorización vía la Proposición 9.6.4:

Desigualdad de Weyl

Proposición 9.7.1 (Mayorante de Weyl). *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces, si $\mu(A)$ denota el vector de autovalores de A, se tiene que*

1. $|\mu(A)| \underset{log}{\prec} s(A)$.

2. *Para todo $p \geq 0$ se tiene que* $|\mu(A)|^p \prec_w s(A)^p$.

Demostración. Verifiquemos las desigualdades de la fórmula (9.36) para los vectores $|\mu(A)|$ y $s(A)$. Para $k = 1$, basta notar que $\rho(A) \leq \|A\|_{sp}$. El caso $k \geq 1$ se obtiene considerando la desigualdad anterior para los tensores alternados $\Lambda^k(A)$. En efecto, por el Corolario 9.6.2, tenemos que

$$\prod_{i=1}^{k} |\mu(A)|_i^{\downarrow} = \rho(\Lambda^k A) \leq \|\Lambda^k A\|_{sp} = \prod_{i=1}^{k} s_i(A) \ , \qquad \text{para todo} \qquad k \in \mathbb{I}_n \ .$$

La igualdad para $k = n$ se deduce de que $|\det A| = (\det A^* A)^{1/2} = \det |A|$. La segunda parte se deduce de la Proposición 9.6.4. ■

Desigualdad de B. Simon

La que sigue es una variante de la Proposición 9.1.7 (desigualdad de Kittaneh '95):

Proposición 9.7.2. *Sean $A, B \in \mathcal{M}_n(\mathbb{C})$ tales que AB es normal. Entonces*

$$|||AB||| \leq |||BA|||$$

para toda NUI $|||\cdot|||$ en $\mathcal{M}_n(\mathbb{C})$.

Demostración. Como AB es normal, se tiene que $\|AB\|_{sp} = \rho(AB)$. Luego

$$s_1(AB) = \|AB\|_{sp} = \rho(AB) = \rho(BA) \leq \|BA\|_{sp} = s_1(BA) \ .$$

Aplicando esta desigualdad a $\Lambda^k AB$ $(1 \leq k \leq n)$, que también es normal, obtenemos que

$$s(AB) \underset{\log}{\prec_w} s(BA) \ , \qquad \text{i.e.} \qquad \prod_{i=1}^{k} s_i(AB) \leq \prod_{i=1}^{k} s_i(BA) \quad , \quad k \in \mathbb{I}_n \ .$$

Por la Proposición 9.6.4, deducimos que $s(AB) \prec_w s(BA)$. ∎

Desigualdad de Horn

Proposición 9.7.3 (Teorema de Horn)**.** *Sean $A, B \in \mathcal{G}l(n)$. Sea $\mu(AB)$ el vector de autovalores de AB ordenado con módulos decrecientes, i.e. $|\mu(AB)| = |\mu(AB)|^{\downarrow}$. Entonces*

$$|\mu(AB)| \underset{log}{\prec} s(AB) \underset{log}{\prec} s(A)s(B) = (s_1(A)s_1(B), \dots, s_n(A)s_n(B)\,) \ .$$

En particular, si $A, B \in \mathcal{G}l(n)^+$, para todo $k \in \mathbb{I}_n$ se tiene que

$$\prod_{i=1}^{k} \mu_i(AB) \leq \prod_{i=1}^{k} \mu_i(A)\mu_i(B) \quad \text{ y } \quad \prod_{i=k}^{n} \mu_i(AB) \geq \prod_{i=k}^{n} \mu_i(A)\mu_i(B) \ . \tag{9.38}$$

Demostración. Las relación $|\mu(AB)| \underset{\log}{\prec} s(AB)$ se deduce de la Proposición 9.7.1. Por otro lado, el Corolario 9.6.2 asegura que, para todo $k \in \mathbb{I}_n$,

$$\prod_{i=1}^{k} s_i\,(AB) = \|\Lambda^k AB\|_{sp} = \ \|\Lambda^k A \,\cdot\, \Lambda^k B\|_{sp} \leq \ \|\Lambda^k A\|_{sp}\,\|\Lambda^k B\|_{sp} = \ \prod_{i=1}^{k} s_i\,(A)\,s_i\,(B) \ .$$

Además, como $|\det C| = \det |C|$ para cualquier $C \in \mathcal{M}_n(\mathbb{C})$, se tiene que

$$\prod_{i=1}^{n} s_i\,(AB) = \det |AB| = |\det AB| = |\det A|\ |\det B| = \det |A|\ \det |B| = \prod_{i=1}^{n} s_i\,(A)\,s_i\,(B) \ .$$

La Eq. (9.38) se deduce de lo anterior y de la Observación 9.6.3, ya que podemos usar que $\mu(AB) = \mu(A^{1/2}BA^{1/2}) \in \mathbb{R}_+^{*n}$, $\mu(A) = s(A)$ y $\mu(B) = s(B)$. ∎

Ejercicio 9.7.4. Dadas $A, B \in \mathcal{G}l(n)^+$, probar que $\mu(A^{1/2}BA^{1/2})^2 \underset{\log}{\prec} \mu(AB^2A)$. Comparar con el Teorema 9.9.4 de más adelante. ▲

9.8 La exponencial

Generalidades

Sea $A \in \mathcal{M}_n(\mathbb{C})$. Recordemos que la exponencial de A es la matriz

$$e^A = \exp(A) = \sum_{m=0}^{\infty} \frac{A^m}{m!} . \tag{9.39}$$

La serie converge absolutamente, porque

$$\sum_{m=0}^{\infty} \frac{\|A^m\|}{m!} \le \sum_{m=0}^{\infty} \frac{\|A\|^m}{m!} = e^{\|A\|} < \infty .$$

Por medio de una prueba similar a la del Corolario 1.7.2 (usando el Teorema 1 de Schur 1.6.1), se puede ver que, si $\lambda(A) = (\lambda_1(A), \dots, \lambda_n(A))$, entonces

$$\lambda(e^A) = e^{\lambda(A)} := (e^{\lambda_1(A)}, \dots, e^{\lambda_n(A)}) . \tag{9.40}$$

En particular, esto dice que $\sigma\left(e^A\right) = e^{\sigma(A)}$ y que $e^A \in \mathcal{G}l(n)$ para toda $A \in \mathcal{M}_n(\mathbb{C})$. Para hacer esto se usa que, al igual que con los polinomios, se tiene que

$$e^{SAS^{-1}} = Se^AS^{-1} , \quad \text{para toda } S \in \mathcal{G}l(n) .$$

Por último, no es difícil verificar con el mismo tipo de argumentos que

$$e^A = \lim_{m\to\infty} \left(I + \frac{A}{m}\right)^m . \tag{9.41}$$

Observación 9.8.1. Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Por la teoría general de series de potencias (con variables que conmutan, para usar el binomio de Newton), se puede ver que

$$\text{si} \quad AB = BA \implies e^{A+B} = e^A\, e^B .$$

Entre otras cosas, esto sirve para probar que $(e^A)^{-1} = e^{-A}$, porque $e^A\, e^{-A} = e^{A-A} = I$. En forma similar puede verse que, si $A \in \mathcal{H}(n)$, entonces

$$e^A \in \mathcal{G}l(n)^+ \quad \text{y} \quad (e^A)^{1/2} = e^{\frac{A}{2}} . \tag{9.42}$$

Más aún, cuando $A \in \mathcal{H}(n)$, se tiene que $e^A = f(A)$, donde $f(t) = e^t$, $t \in \mathbb{R}$, en el sentido del cálculo funcional para autoadjuntas visto en el Capítulo 6. Esto puede probarse diagonalizando a A o bien tomando límite de polinomios en la fórmula (9.39). Aplicando los resultados conocidos para dicho cálculo (en particular 6.1.3), se tienen las siguientes propiedades: Si $A \in \mathcal{H}(n)$, entonces

1. $e^A \in \mathcal{G}l(n)^+$ y $A = \log e^A$.
2. $(e^A)^r = e^{rA}$ para todo $r \in \mathbb{R}$.
3. Si $A > 0$, entonces $A = e^{\log A}$.
4. Más aún, $A^r = e^{r \log A}$ para todo $r \in \mathbb{R}$. ▲

Fórmula de Lie-Trotter

Lamentablemente, cuando A y B no conmutan, la cosa se pone mucho más brava, y es muy dificil encontrar las propiedades de e^{A+B} en función de las de A y B. La única herramienta que se tiene, y que se usa sistemáticamente para este problema, es la famosa fórmula de Lie-Trotter que mostraremos a continuación.

Teorema 9.8.2. *Dadas A y $B \in \mathcal{M}_n(\mathbb{C})$, se tiene la siguiente fórmula:*

$$e^{A+B} \;=\; \lim_{m\to\infty} \left(e^{\frac{A}{m}}\; e^{\frac{B}{m}}\right)^m \;. \tag{9.43}$$

Demostración. Dadas X, $Y \in \mathcal{M}_n(\mathbb{C})$, mediante una típica cuenta telescópica puede verse que $X^m - Y^m = \sum\limits_{j=0}^{m-1} X^{m-1-j}(X-Y)Y^j$ para todo $m \in \mathbb{N}$. Luego,

$$\begin{aligned} \|X^m - Y^m\| &= \|\sum_{j=0}^{m-1} X^{m-1-j}(X-Y)Y^j\| \\ &\le \sum_{j=0}^{m-1} \|X^{m-1-j}\|\;\|X-Y\|\;\|Y^j\| \\ &\le \|X-Y\| \sum_{j=0}^{m-1} M^{m-1-j}\, M^j = m\;\|X-Y\|\; M^{m-1}\;, \end{aligned} \tag{9.44}$$

donde $M = \text{máx}(\|X\|, \|Y\|)$. Consideremos ahora las matrices

$$X_m = e^{\frac{A+B}{m}}\;, \quad Y_m = e^{\frac{A}{m}}\; e^{\frac{B}{m}}\;, \qquad \text{para cada } m \in \mathbb{N}\;.$$

Observar que $M_m = \text{máx}(\|X_m\|, \|Y_m\|) \le e^{\frac{\|A\|+\|B\|}{m}}$. Por la desigualdad (9.44),

$$\|X_m^m - Y_m^m\| \le \; m\,\|X_m - Y_m\|\; e^{\frac{m-1}{m}\,(\|A\|+\|B\|)} \le \; m\,\|X_m - Y_m\|\; e^{\|A\|+\|B\|}\;, \quad m \in \mathbb{N}.$$

Luego del desarrollo en series de la función exponencial, obtenemos que $X_m - Y_m =$

$$\begin{aligned} &= 1 + \frac{A+B}{m} + \sum_{k=2}^{\infty} \frac{(A+B)^k}{m^k\; k!} - \left\{\left[1 + \frac{A}{m} + \sum_{k=2}^{\infty} \frac{A^k}{m^k\; k!}\right]\left[1 + \frac{B}{m} + \sum_{k=2}^{\infty} \frac{B^k}{m^k\; k!}\right]\right\} \\ &= \frac{1}{m^2}\left[\sum_{k=2}^{\infty} \frac{(A+B)^k}{m^{k-2}\; k!} - AB - \left(I + \frac{A}{m}\right)\sum_{k=2}^{\infty} \frac{B^k}{m^{k-2}\; k!} - \sum_{k=2}^{\infty} \frac{A^k}{m^{k-2}\; k!}\; e^{\frac{B}{m}}\right] \\ &= \frac{1}{m^2}\; C_m(A,B)\;. \end{aligned}$$

Si mostráramos que $C(A,B) = \sup\limits_{m\in\mathbb{N}} \|C_m(A,B)\| < \infty$, entonces tendríamos que

$$\|X_m^m - Y_m^m\| \le m\,\|X_m - Y_m\|\; e^{\|A\|+\|B\|} \le \frac{C(A,B)\; e^{\|A\|+\|B\|}}{m} \quad \xrightarrow[m\to\infty]{} \; 0\;.$$

Afortunadamente, las constantes $\|C_m(A,B)\|$ se pueden acotar con las series de las normas. Aparecen varios sumandos, todos ellos elementales. Por ejemplo, $\|e^{\frac{B}{m}}\| \le e^{\|\frac{B}{m}\|} \le e^{\|B\|}$ y

$$\left\| \sum_{k=2}^{\infty} \frac{A^k}{m^{k-2}\ k!} \right\| \le \sum_{k=2}^{\infty} \frac{\|A^k\|}{m^{k-2}\ k!} \le \sum_{k=2}^{\infty} \frac{\|A^k\|}{k!} \le e^{\|A\|} \quad \text{para todo } m \in \mathbb{N} \ .$$

Por lo tanto, $\|e^{A+B} - (e^{\frac{A}{m}} e^{\frac{B}{m}})^m\| = \|X_m^m - Y_m^m\| \xrightarrow[m\to\infty]{} 0 \implies \lim_{m\to\infty} (e^{\frac{A}{m}} e^{\frac{B}{m}})^m = e^{A+B}$. ■

Otras fórmulas relacionadas

Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Conjugando la fórmula (9.43) con la sucesión $e^{\frac{B}{2m}}$ (que tiende a I), obtenemos esta otra versión de Lie-Trotter

$$e^{A+B} = \lim_{m\to\infty} e^{\frac{B}{2m}} \left(e^{\frac{A}{m}}\ e^{\frac{B}{m}} \right)^m e^{-\frac{B}{2m}} = \lim_{m\to\infty} \left(e^{\frac{B}{2m}}\ e^{\frac{A}{m}}\ e^{\frac{B}{2m}} \right)^m , \tag{9.45}$$

que es mucho más adecuada en caso de que $A, B \in \mathcal{H}(n)$, para que la cosa pase entre matrices positivas.

Ejercicios 9.8.3.

1. Si suponemos que $A, B \in \mathcal{H}(n)$, puede obtenerse la siguiente nueva fórmula:

$$e^{A+B} = \lim_{t\to 0} \left(e^{\frac{tB}{2}}\ e^{tA}\ e^{\frac{tB}{2}} \right)^{\frac{1}{t}} . \tag{9.46}$$

   Observar que $e^{\frac{tB}{2}}\ e^{tA}\ e^{\frac{tB}{2}} \in \mathcal{G}l(n)^+$ por lo que $\left(e^{\frac{tB}{2}}\ e^{tA}\ e^{\frac{tB}{2}} \right)^{\frac{1}{t}}$ tiene sentido para todo $t \neq 0$. Además, $\left(e^{\frac{tB}{2}}\ e^{tA}\ e^{\frac{tB}{2}} \right)^{-1} = e^{\frac{-tB}{2}}\ e^{-tA}\ e^{\frac{-tB}{2}}$, por lo que el caso $t < 0$ no crea problemas.
   Sug: Desarrollar el producto de las tres series asociadas a $e^{\frac{tB}{2}}\ e^{tA}\ e^{\frac{tB}{2}}$ y aplicarles la serie de potencias de $-1 < x \mapsto \log(1+x)$. Despues seguir pasos similares a los de la prueba del Teorema 9.8.2.

2. Extendiendo el argumento que utilizamos para probar el Teorema 9.8.2 probar que, dadas $A_1, A_2, \dots, A_k \in \mathcal{M}_n(\mathbb{C})$, se tiene que

$$\exp\left(\sum_{i=1}^{k} A_i \right) = \lim_{m\to\infty} \left[e^{\frac{A_1}{m}}\ e^{\frac{A_2}{m}} \dots e^{\frac{A_k}{m}} \right]^m . \tag{9.47}$$

9.9 Desigualdades de Araki y Cordes

Proposición 9.9.1. *Sean $A, B \in \mathcal{G}l(n)^+$. Entonces, para todo $r \in (0,1)$*

$$\mu_1(A^r B^r) \le \mu_1(AB)^r \ .$$

Demostración. Podemos asumir que $\mu_1(AB) = 1$, porque si $\alpha^2 = \mu_1(AB)$, basta cambiar A, B por $\alpha^{-1}A$ y $\alpha^{-1}A$, dado que $\alpha \neq 0$ y la desigualdad a probar es homogénea. Luego debemos verificar que $\mu_1(A^r B^r) \leq 1$. En efecto,

$$\mu_1(AB) = \mu_1\left(A^{1/2}BA^{1/2}\right) = 1 \quad \Longrightarrow \quad A^{1/2}BA^{1/2} \leq I \quad \Longrightarrow \quad B \leq A^{-1} \ .$$

Como $r \in (0,1)$, el Teorema 6.2.6 dice que $f(x) = x^r$ es monótona de operadores. Luego

$$B^r \leq A^{-r} \quad \Longrightarrow \quad A^{r/2}B^rA^{r/2} \leq I \quad \Longrightarrow \quad \mu_1(A^rB^r) = \mu_1\left(A^{r/2}B^rA^{r/2}\right) \leq 1 \ ,$$

como se aseveró. ■

Proposición 9.9.2 (Cordes '87 [5]). *Sean $A, B \in \mathcal{M}_n(\mathbb{C})^+$. Entonces*

$$\|A^rB^r\|_{sp} \leq \|AB\|_{sp}^r \quad \textit{para todo } r \in [0,1] \ . \tag{9.48}$$

Demostración. Como antes, supondremos que $\|AB\|_{sp} = 1$. En tal caso

$$\begin{aligned} 1 = \|AB\|_{sp} = \|AB^2A\|_{sp} \quad & \Longrightarrow AB^2A \leq 1 \Longrightarrow B^2 \leq A^{-2} \\ & \Longrightarrow B^{2r} \leq A^{-2r} \quad \text{(por el Teorema 6.2.6)} \\ & \Longrightarrow A^rB^{2r}A^r \leq 1 \Longrightarrow \|A^rB^{2r}A^r\|_{sp} \leq 1 \\ & \Longrightarrow \|A^rB^r\|_{sp} \leq 1 \ . \end{aligned}$$

El caso general sale por homogeneidad. ■

Dadas$A, B \in \mathcal{H}(n)$, recordar que escribimos $A \prec B$ en lugar de $\mu(A) \prec \mu(B)$ para notar la mayorización entre los autovalores de A y los de B.

Definición 9.9.3.

1. Sean $A, B \in \mathcal{M}_n(\mathbb{C})^+$. Escribiremos $A \underset{\log}{\prec_w} B$ si $\mu(A) \underset{\log}{\prec_w} \mu(B)$. Es decir, si

$$\prod_{i=1}^{k} \mu_i(A) \leq \prod_{i=1}^{k} \mu_i(B) \quad \text{para todo} \quad k \in \mathbb{I}_n \ . \tag{9.49}$$

2. Si $A, B \in \mathcal{G}l(n)^+$, decimos que $A \underset{\log}{\prec} B$ si $\det A = \det B$ y $A \underset{\log}{\prec_w} B$.

Observar que, $A \underset{\log}{\prec} B \iff \log A \prec \log B$, ya que la función $t \mapsto e^t$ es creciente. ▲

Teorema 9.9.4 (Araki '90 [21]). *Sean $A, B \in \mathcal{M}_n(\mathbb{C})^+$. Para todo $r \in (0,1)$ se tiene que*

$$(A^{r/2}B^rA^{r/2})^{1/r} \underset{log}{\prec} A^{1/2}BA^{1/2}$$

Demostración. Fijemos un $r \in (0,1)$. Como

$$\sigma(A^{r/2}B^rA^{r/2}) = \sigma(A^rB^r) \quad \text{ y } \quad \sigma(A^{1/2}BA^{1/2}) = \sigma(AB) \ ,$$

basta ver que

$$\prod_{j=1}^k \mu_j(A^rB^r)^{1/r} \le \prod_{j=1}^k \mu_j(AB) \ , \quad k \in \mathbb{I}_n \quad \text{ y } \quad \prod_{j=1}^n \mu_j(A^rB^r)^{1/r} = \prod_{j=1}^n \mu_j(AB) \ .$$

Aplicando la Proposición 9.9.1 a $\Lambda^k A$ y $\Lambda^k B$, y usando la Proposición 7.4.4 se obtiene

$$\mu_1^{1/r}((\Lambda^k A)^r(\Lambda^k B)^r) = \mu_1^{1/r}(\Lambda^k A^r\,\Lambda^k B^r) \le \mu_1(\Lambda^k A\,\Lambda^k B) \ .$$

Como $(\Lambda^k A)(\Lambda^k B) = \Lambda^k(AB)$, podemos deducir que

$$\prod_{j=1}^k \mu_j^{1/r}(A^rB^r) = \mu_1^{1/r}(\Lambda^k(A^rB^r)) \le \mu_1(\Lambda^k(AB)) = \prod_{j=1}^k \mu_j(AB) \ .$$

La igualdad en el caso $k = n$ se obtiene tomando determinantes. ■

Corolario 9.9.5. *Sea $|||\cdot|||$ una NUI en $\mathcal{M}_n(\mathbb{C})$. Dados $A, B \in \mathcal{M}_n(\mathbb{C})^+$, se tiene que*

$$|||A^rB^rA^r||| \le |||(ABA)^r||| \quad \text{para todo } r \in (0,1) \ .$$

Demostración. Aplicando el Teorema 9.9.4 a las matrices A^2 y B y la Proposición 4.4.3, obtenemos

$$\prod_{j=1}^k s_j(A^rB^rA^r) \le \prod_{j=1}^k s_j(ABA)^r \ \ (k \in \mathbb{I}_n) \implies s(A^rB^rA^r) \prec_w s(\,(ABA)^r) \ . \tag{9.50}$$

Se usa que $s_i(\,(ABA)^r) = \mu_i(\,(ABA)^r) = \mu_i(ABA)^r = s_i(ABA)^r$, para todo $i \in \mathbb{I}_n$. ■

Observación 9.9.6. En las Proposiciones 9.9.1 y 9.9.2, el Teorema 9.9.4 y el Corolario 9.9.5, valen las desigualdades inversas si uno considera exponentes $t \ge 1$. En efecto, basta aplicar lo conocido a las matrices A^t y B^t, con el exponente $r = t^{-1}$. En el caso del Corolario 9.9.5, se puede corregir el exponente externo en la Eq. (9.50) ▲

Proposición 9.9.7. *Dadas $A, B \in \mathcal{G}l(n)^+$, se cumple que*

$$\log A + \log B \prec \log(A^{1/2}BA^{1/2}) \ . \tag{9.51}$$

Demostración. Por la fórmula de Lie-Trotter, en la versión dada por la fórmula (9.46),

$$e^{\log A + \log B} = \lim_{r\to 0} \ (A^{r/2}B^rA^{r/2})^{1/r}$$

Por el Teorema de Araki, $(A^{r/2}B^rA^{r/2})^{1/r} \underset{\log}{\prec} (A^{1/2}BA^{1/2})$ para todo $r \in (0,1)$. Aplicando el Corolario 2.3.8, obtenemos que

$$e^{\log A + \log B} \underset{\log}{\prec} (A^{1/2}BA^{1/2}) \implies \log A + \log B \prec \log(A^{1/2}BA^{1/2}) \ .$$ ■

Ejercicio 9.9.8. Usando el Corolario 9.9.5 para las NUI's $\|A\|_p = (\operatorname{tr}|A|^p)^{1/p}$, mostrar la famosa desigualdad de Araki-Lieb-Thirring: Dadas matrices $A, B \in \mathcal{G}l(n)^+$, se tiene que

$$\operatorname{tr}\left[(B^{1/2}AB^{1/2})^{st}\right] \le \operatorname{tr}\left[(B^{t/2}A^tB^{t/2})^s\right] \ ,$$

para todo par de números $s, t \ge 1$. ▲

9.10 Desigualades entre exponenciales

Si $z \in \mathbb{C}$, uno tiene que $|e^z| = e^{\operatorname{Re} z} = |e^{\operatorname{Re} z}|$. Veamos qué pasa en matrices:

Proposición 9.10.1. *Sea $|||\cdot|||$ una NUI en $\mathcal{M}_n(\mathbb{C})$. Para toda $A \in \mathcal{M}_n(\mathbb{C})$, se tiene que*

$$|||e^A||| \le |||e^{\operatorname{Re} A}||| \ .$$

Demostración. Usando que $\|B^m\|_{sp} \le \|B\|_{sp}^m$ y que $s_1(B)^2 = \|B\|_{sp}^2 = \|B^*B\|_{sp}$, obtenemos que

$$s_1(B^m)^2 \le s_1(B)^{2m} \quad \Longrightarrow \quad s_1(\,(B^*)^m B^m) \le s_1(B^*B)^m \ ,$$

para todo $m \in \mathbb{N}$ y toda $B \in \mathcal{M}_n(\mathbb{C})$. Aplicando esto a $\Lambda^k B$, tenemos más:

$$\prod_{i=1}^{k} s_i(\,(B^*)^m B^m) \le \prod_{i=1}^{k} s_i(B^*B)^m \quad \text{para todo } k \in \mathbb{I}_n \text{ y todo } m \in \mathbb{N} \ .$$

Eligiendo ahora $B = \exp \frac{A}{m}$ y aplicando Lie-Trotter llegamos a que

$$\prod_{i=1}^{k} s_i(\,e^{A^*} e^{A}) \le \prod_{i=1}^{k} s_i\big([e^{\frac{A^*}{m}} e^{\frac{A}{m}}]^m\big) \xrightarrow[m\to\infty]{} \prod_{i=1}^{k} s_i\big(e^{A^*+A}\big) \quad \text{para todo } k \in \mathbb{I}_n \ .$$

Tomando raices cuadradas, llegamos a que $s(e^A) \underset{\log}{\prec_w} s(e^{\operatorname{Re} A})$. Usando ahora la Proposición 4.4.3, tenemos finalmente que $s(e^A) \prec_w s(e^{\operatorname{Re} A})$. ■

Ejercicio 9.10.2. Encontrar $A \in \mathcal{M}_2(\mathbb{C})$ tal que $|||e^A||| < |||e^{\operatorname{Re} A}|||$ para toda NUI. ▲

Recordemos que si $C, D \in \mathcal{M}_n(\mathbb{C})^+$, entonces $\sigma(CD) \subseteq \mathbb{R}_+$ y más aún, al vector de autovalores $\mu(CD) = \mu(D^{1/2}CD^{1/2}) \in \mathbb{R}_+^n$, se lo puede pensar ordenado en forma decreciente. En particular, se tiene que $\operatorname{tr} CD \in \mathbb{R}_+$.

Lema 9.10.3. *Dadas $A, B \in \mathcal{M}_n(\mathbb{C})$, se tiene que*

$$|\lambda(e^{A+B})| \prec_w \mu\,(e^{\operatorname{Re} A}\ e^{\operatorname{Re} B}) \quad y \quad |\operatorname{tr}\, e^{A+B}| \le \operatorname{tr}\,(e^{\operatorname{Re} A}\ e^{\operatorname{Re} B}) \ .$$

Demostración. Fijado $k \in \mathbb{I}_n$, definamos las siguiente función continua:

$$f_k : \mathcal{M}_n(\mathbb{C}) \to \mathbb{R}_+ \quad \text{dada por} \quad f_k(X) = \sum_{i=1}^{k} |\lambda(X)|_i^{\downarrow} \ , \quad \text{para } X \in \mathcal{M}_n(\mathbb{C}) \ .$$

Notar que todas estas f_k cumplen que, para todo par $X, Y \in \mathcal{M}_n(\mathbb{C})$ y todo $m \in \mathbb{N}$,

$$f_k(XY) = f_k(YX) \quad \text{y} \quad f_k(X^{2m}) \le f_k([X^*X]^m) \ . \tag{9.52}$$

En efecto, observar que para todo $r \in \mathbb{I}_n$ y todo $m \in \mathbb{N}$, se tiene que

$$\rho(\Lambda^r X^{2m}) = \rho(\Lambda^r X)^{2m} \le \|\Lambda^r X\|_{sp}^{2m} = \|\Lambda^r X^* X\|_{sp}^{m} = \rho(\Lambda^r [X^*X]^m) \ .$$

Por la Proposición 9.6.4, deducimos que $|\lambda(X^{2m})| \prec_w \mu([X^*X]^m)$, o sea la desigualdad de (9.52) para todo $k \in \mathbb{I}_n$. La igualdad vale porque $\lambda(XY) = \lambda(YX)$. Por lo tanto,

$$\begin{aligned} f_k\big((XY)^{2^m} \big) &\le f_k\big([\,(XY)^*(XY)\,]^{2^{m-1}} \big) = f_k\big([\,Y^*X^*XY\,]^{2^{m-1}} \big) \\ &= f_k\big([\,(X^*X)\,(YY)^*\,]^{2^{m-1}} \big) \ . \end{aligned}$$

donde la última igualdad vale porque $[\,Y^*X^*XY\,]^{2^{m-1}}$ difiere de $[\,X^*XYY^*\,]^{2^{m-1}}$ tan sólo en "pasar" el primer Y^* al final. Repitiendo esto, se llega a que

$$f_k\big((XY)^{2^m} \big) \le f_k\big([\,(X^*X)^2 (YY^*)^2\,]^{2^{m-2}} \big) \le \cdots \le f_k\big((X^*X)^{2^{m-1}} (YY^*)^{2^{m-1}} \big)$$

Pongamos ahora $X = \exp \frac{A}{2^m}$, $Y = \exp \frac{B}{2^m}$ y $M_m = 2^m$. Queda

$$f_k\big((e^{\frac{A}{M_m}} \ e^{\frac{B}{M_m}})^{M_m} \big) \le f_k\big((e^{\frac{A^*}{M_m}} \, e^{\frac{A}{M_m}})^{\frac{1}{2}M_m} \ (e^{\frac{B}{M_m}} \, e^{\frac{B^*}{M_m}})^{\frac{1}{2}M_m} \big) \ .$$

Tomando límite $m \to \infty$, y usando Lie-Trotter (y que las f_k son continuas), tenemos que

$$f_k(e^{A+B}) \le f_k(e^{\frac{A^*+A}{2}} \ e^{\frac{B+B^*}{2}}) = f_k(e^{\operatorname{Re} A} \ e^{\operatorname{Re} B}) \ .$$

Como esto vale para todo $k \in \mathbb{I}_n$, llegamos a que $|\lambda(e^{A+B})| \prec_w \mu(e^{\operatorname{Re} A} \ e^{\operatorname{Re} B})$. La otra desigualdad se prueba usando que la función $f(X) = |\operatorname{tr} X|$ también cumple las condiciones de (9.52) (sale usando que $f(Y) \le f_n(Y)$, pero coinciden si $Y \in \mathcal{M}_n(\mathbb{C})^+$), y haciendo el resto de la cuenta igual. ■

Observación 9.10.4. Si una función $f : \mathcal{M}_n(\mathbb{C}) \to \mathbb{R}$ es continua y cumple las condiciones

$$f(XY) = f(YX) \quad \text{ y } \quad |f(X^{2m})| \le f([XX^*]^m) \tag{9.53}$$

para todo $m \in \mathbb{N}$ y todo par $X, Y \in \mathcal{M}_n(\mathbb{C})$, decimos que f es de la **clase T**. La sutil diferencia con la Eq. (9.52) es que aca no pedimos que f sea positiva, pero no ponemos módulo en el último término de (9.53). Toda la cuenta del Lema 9.10.3 puede rehacerse para una tal función, con lo que se obtiene la desigualdad más general: Si f es de la clase T,

$$f(e^{A+B}) \le f(e^{\operatorname{Re} A} \ e^{\operatorname{Re} B}) \ ,$$

para todo par $A, B \in \mathcal{M}_n(\mathbb{C})$. ▲

Proposición 9.10.5. *Sea $|||\cdot|||$ una NUI en $\mathcal{M}_n(\mathbb{C})$. Dadas $A, B \in \mathcal{H}(n)$, se tiene que*

$$|||\, e^{A+B} ||| \le |||\, e^A \ e^B ||| \ .$$

Demostración. Por el Lema 9.10.3 y el hecho de que $A, B \in \mathcal{H}(n)$, tenemos que

$$s(e^{A+B}) = \mu(e^{A+B}) = |\mu(e^{A+B})| \prec_w \mu(e^A e^B) \prec_w s(e^A e^B) \ ,$$

donde la última mayorización sale de la Proposición 9.7.1 (mayorante de Weyl). ■

Proposición 9.10.6 (Desigualdad de Golden-Thompson)**.** *Si* $A, B \in \mathcal{H}(n)$*, entonces*

$$\operatorname{tr} e^{A+B} \leq \operatorname{tr}\, e^A \; e^B \ . \tag{9.54}$$

Demostración. Es consecuencia directa del Lema 9.10.3. ■

Ejercicio 9.10.7. Usando la Proposición 9.10.5, mostrar que

$$\operatorname{tr}\, e^{A+B} \leq \operatorname{tr}\Big(\left[e^B e^{2A} e^B\right]^{1/2}\Big) \quad \text{ para } A, B \in \mathcal{H}(n) \ .$$

¿Esto es mejor o peor que Golden-Thompson? ▲

9.11 Desigualdad de Ando-Johnson-Bapat

Logaritmos

Proposición 9.11.1. *Sean* $A \in \mathcal{G}l(n)^+$ *y* $B \in \mathcal{G}l(m)^+$*. Entonces se verifica que*

$$\log(A \otimes B) = (\log A) \otimes I_m + I_n \otimes (\log B) \ .$$

Demostración. Supongamos que $A = \sum\limits_{i=1}^{k} \mu_i P_i$ y que $B = \sum\limits_{j=1}^{h} \lambda_j Q_j$. Luego

$$A \otimes B = \sum_{i=1}^{k} \sum_{j=1}^{h} \mu_i \lambda_j P_i \otimes Q_j \ .$$

Notemos que $(P_i \otimes Q_j)_{(i,j)}$ es un sisitema de proyectores para $\mathbb{C}^n \otimes \mathbb{C}^m$. Luego, si $u \otimes v \in \mathbb{C}^n \otimes \mathbb{C}^m$ y abreviamos $L = \log(A \otimes B)(u \otimes v)$, se tiene que

$$\begin{aligned}
L &= \sum_{i=1}^{k} \sum_{j=1}^{h} \log(\mu_i \lambda_j)[P_i \otimes Q_j(u \otimes v)] \\
&= \sum_{i=1}^{k} \left[\log(\mu_i) \sum_{j=1}^{h} P_i(u) \otimes Q_j(v)\right] + \sum_{j=1}^{h} \left[\log(\lambda_j) \sum_{i=1}^{k} P_i(u) \otimes Q_j(v)\right] \\
&= \sum_{i=1}^{k} \log(\mu_i) \left[P_i(u) \otimes \left(\sum_{j=1}^{h} Q_j(v)\right)\right] + \sum_{j=1}^{h} \log(\lambda_j) \left[\left(\sum_{i=1}^{k} P_i(u)\right) \otimes Q_j(v)\right] \\
&= \sum_{i=1}^{k} \log(\mu_i)[P_i(u) \otimes v] + \sum_{j=1}^{h} \log(\lambda_j)[u \otimes Q_j(v)] \\
&= \left(\sum_{i=1}^{k} \log(\mu_i) P_i(u)\right) \otimes v + u \otimes \left(\sum_{j=1}^{h} \log(\lambda_j) Q_j(v)\right) \\
&= [(\log A) \otimes I_m]\ (u \otimes v) + [I_n \otimes (\log B)]\ (u \otimes v) \ .
\end{aligned}$$

La igualdad en el resto de los elementos de $\mathbb{C}^n \otimes \mathbb{C}^m$ se obtiene por linealidad. ■

Corolario 9.11.2. *Sean* $A, B \in \mathcal{G}l(n)^+$*. Entonces*

$$\log(A \circ B) \geq (\log A + \log B) \circ I_n \ .$$

Demostración. Consideremos la función $\Phi : L(\mathbb{H}_n \otimes \mathbb{H}_n) \to \mathcal{M}_n(\mathbb{C})$ definida en la Proposición 8.1.4. Recordar que $\Phi(T) = T_{\mathcal{S}}$ $(T \in L(\mathbb{H}_n \otimes \mathbb{H}_n)\ )$ para cierto subespacio $\mathcal{S} \subseteq \mathbb{H}_n \otimes \mathbb{H}_n$, y que $\Phi(A \otimes B) = A \circ B$, para $A, B \in \mathcal{M}_n(\mathbb{C})$. Además, por el Corolario 6.3.14, la función $t \mapsto \log T$ es $\cap$OP en $(0, +\infty)$. Luego, aplicando el Teorema 6.3.6 deducimos que

$$\Phi(\log X) = (\log X)_{\mathcal{S}} \leq \log(X_{\mathcal{S}}) = \log \Phi(X) \quad \text{ para todo } X \in \mathcal{G}l(\mathbb{H}_n \otimes \mathbb{H}_n)^+ \ .$$

Ahora bien, por la Proposición 9.11.1, $\log A \otimes B = (\log A) \otimes I_n + I_n \otimes (\log B)$, así que

$$\begin{aligned} \log(A \circ B) = \log \Phi(A \otimes B) &\geq \Phi\big[\log(A \otimes B)\big] \\ &= \Phi\big[(\log A) \otimes I_n + I_n \otimes (\log B)\big] \\ &= (\log A) \circ I_n + I_n \circ (\log B) \ , \end{aligned}$$

como se afirmaba. ■

La desigualdad

Ahora daremos la prueba obtenida por T. Ando de la que fue llamada muchos años "conjetura de Johnson-Bapat":

Teorema 9.11.3. *Sean* $A, B \in \mathcal{G}l(n)^+$*. Entonces*

$$\prod_{i=k}^{n} \mu_i(A \circ B) \geq \prod_{i=k}^{n} \mu_i(AB) \quad \text{ para todo } \quad k \in \mathbb{I}_n \ .$$

Demostración. Por el Corolario 9.11.2 sabemos que

$$\log(A \circ B) \geq (\log A + \log B) \circ I_n \ .$$

Por el Teorema de Weyl 2.3.5, para todo $k \in \mathbb{I}_n$, tenemos que

$$\log\left(\prod_{i=k}^{n} \mu_i(A \circ B)\right) = \sum_{i=k}^{n} \mu_i(\log(A \circ B)\,) \geq \sum_{i=k}^{n} \mu_i(\,(\log A + \log B) \circ I_n) \ .$$

De acuerdo al Teorema de mayorización de Schur 5.1.1,

$$(\log A + \log B) \circ I_n \prec \log A + \log B,$$

y entonces

$$\sum_{i=k}^{n} \mu_i(\,(\log A + \log B) \circ I_n) \geq \sum_{i=k}^{n} \mu_i(\log A + \log B) \ , \quad k \in \mathbb{I}_n \ .$$

Por la Proposición 9.9.7, $\log A + \log B \prec \log(A^{1/2}BA^{1/2})$. Luego

$$\sum_{i=k}^{n} \mu_i(\log A + \log B) \geq \sum_{i=k}^{n} \mu_i(\log(A^{1/2}BA^{1/2})) = \log\left(\prod_{i=k}^{n} \mu_i(AB)\right) , \quad k \in \mathbb{I}_n .$$

Combinando estos resultados obtenemos que

$$\prod_{i=k}^{n} \mu_i(A \circ B) \geq \prod_{i=k}^{n} \mu_i(AB) , \quad k \in \mathbb{I}_n ,$$

como se quería demostrar. ■

Observación 9.11.4. Este teorema mejora el resultado de Bapat-Sunder:

$$\prod_{i=k}^{n} \mu_i(A \circ B) \geq \prod_{i=k}^{n} \mu_i(A)\mu_i(B) \quad \text{para todo} \quad k \in \mathbb{I}_n .$$

pues, de acuerdo al Teorema de Horn (Proposición 9.7.3), se tiene que

$$\prod_{i=k}^{n} \mu_i(AB) \geq \prod_{i=k}^{n} \mu_i(A)\mu_i(B) \quad \text{para todo} \quad k \in \mathbb{I}_n .$$

También verificaremos la siguiente variante del Teorema 9.11.3: ▲

Teorema 9.11.5. *Sean $A, B \in \mathcal{G}l(n)^+$. Entonces*

$$\prod_{i=k}^{n} \mu_i(A \circ B) \geq \prod_{i=k}^{n} \mu_i(AB^t) \quad \text{para todo} \quad k \in \mathbb{I}_n .$$

Demostración. Como $\log B^t = (\log B)^t$, tenemos que

$$(\log B^t) \circ I = (\log B)^t \circ I = (\log B) \circ I.$$

Por lo tanto, podemos reemplazar $\{\log A + \log B\} \circ I$ por $\{\log A + \log B^t\} \circ I$ en la demostración anterior, con lo que queda demostrado el teorema. ■

De los Teoremas 9.11.3 y 9.11.5 se deducen los clásicos teoremas de Fiedler que aseguran que, para $A, B \in \mathcal{G}l(n)^+$,

$$A \circ B \geq \mu_n(AB)I \quad y \quad A \circ B \geq \mu_n(AB^t)I$$

9.12 Medias de operadores positivos

Definición 9.12.1. Cuando $0 \leq \alpha \leq 1$, la α-media de $A, B \in \mathcal{G}l(n)^+$ es lo siguiente

$$A \#_\alpha B = A^{1/2}(A^{-1/2}BA^{-1/2})^\alpha A^{1/2}.$$

En particular, la *media geométrica* $A\#B$ es la $\frac{1}{2}$-media, o sea,

$$A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2} A^{1/2}.$$

Proposición 9.12.2. *Sean $A, B \in \mathcal{G}l(n)^+$. Entonces $A^{-1/2}(A\#B)B^{-1/2} \in \mathcal{U}(n)$.*

Demostración. Sea $T = A^{-1/2}B^{1/2}$. Consideremos su descomposición polar a derecha, dada por $T = |T^*|U$, con $U \in \mathcal{U}(n)$. Luego

$$\begin{aligned} A^{-1/2}(A\#B)B^{-1/2} &= A^{-1/2}(A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2})B^{-1/2} \\ &= (A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}B^{-1/2} \\ &= (TT^*)^{1/2}\, T^{-1} = |T^*|\, T^{-1} = U \in \mathcal{U}(n)\ , \end{aligned}$$

como queríamos demostrar. ■

Lema 9.12.3. *Sean $A \in \mathcal{G}l(n)^+$, $B \in \mathcal{M}_n(\mathbb{C})^+$ y $C \in \mathcal{M}_n(\mathbb{C})$. Entonces*

$$\begin{pmatrix} A & C \\ C^* & B \end{pmatrix} \geq 0 \iff B \geq C^*A^{-1}C\ .$$

Demostración. Para abreviar llamemos $X = -A^{-1}C$. Entonces $\begin{pmatrix} I & X \\ 0 & I \end{pmatrix}$ es inversible. Un cálculo sencillo nos muestra que

$$\begin{pmatrix} I & 0 \\ X^* & I \end{pmatrix}\begin{pmatrix} A & C \\ C^* & B \end{pmatrix}\begin{pmatrix} I & X \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & B - C^*A^{-1}C \end{pmatrix}. \tag{9.55}$$

Entonces de aquí resulta claro el enunciado. ■

Proposición 9.12.4. *Dados $A, B \in \mathcal{G}l(n)^+$, la media geométrica $A\#B$ es la mayor matriz autoadjunta del siguiente conjunto:*

$$\Omega = \left\{ C \in \mathcal{H}(n) : \quad \begin{bmatrix} A & C \\ C & B \end{bmatrix} \geq 0 \right\}.$$

Demostración. Observemos que como $(A\#B)A^{-1}(A\#B) = A^{1/2}(A^{-1/2}BA^{-1/2})A^{1/2} = B$, entonces por el Lema 9.12.3, la matriz

$$\begin{pmatrix} A & A\#B \\ A\#B & B \end{pmatrix}$$

es semidefinida positiva. Luego, $A\#B \in \Omega$. Para demostrar que es efectivamente el máximo tomemos $C \in \Omega$ arbitrario. Entonces el Lema 9.12.3 nos dice que $B \geq CA^{-1}C$. Por lo tanto,

$$(A^{-1/2}CA^{-1/2})^2 = A^{-1/2}\,(CA^{-1}C)\,A^{-1/2} \leq A^{-1/2}BA^{-1/2}$$

y por el Teorema 6.2.6, tenemos

$$A^{-1/2}CA^{-1/2} \leq |\,A^{-1/2}CA^{-1/2}\,| \leq (A^{-1/2}BA^{-1/2})^{1/2}.$$

Luego

$$C \leq A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2} = A\#B,$$

lo cual demuestra que $A\#B$ es el máximo del conjunto Ω. ■

Corolario 9.12.5. *Sean $A, B \in \mathcal{G}l(n)^+$. Entonces*

$$\begin{pmatrix} A & A\#B \\ A\#B & B \end{pmatrix} \geq 0 \qquad y \qquad \begin{pmatrix} B & A\#B \\ A\#B & A \end{pmatrix} \geq 0\ .$$

Demostración. Es inmediato a partir de la Proposición 9.12.4. ■

Observación 9.12.6. Puede probarse que $(A\#B)^2 \underset{\log}{\prec} A^{1/2}BA^{1/2}$, lo que implica que

$$\prod_{i=k}^{n} \mu_i(A\#B)^2 \geq \prod_{i=k}^{n} \mu_i(AB) \quad \text{para todo} \quad k \in \mathbb{I}_n\ .$$

En consecuencia, el siguiente resultado mejora el Teorema 9.11.3. ▲

Teorema 9.12.7. *Sean $A, B \in \mathcal{G}l(n)^+$. Entonces*

$$\prod_{i=k}^{n} \mu_i(A \circ B) \geq \prod_{i=k}^{n} \mu_i(A\#B)^2 \quad \textit{para todo} \quad k \in \mathbb{I}_n\ .$$

Demostración. Por la Proposición 9.12.5 y el Teorema de mayorización de Schur 5.1.1

$$\begin{bmatrix} A & A\#B \\ A\#B & B \end{bmatrix} \circ \begin{bmatrix} B & A\#B \\ A\#B & A \end{bmatrix} = \begin{bmatrix} A \circ B & A\#B \circ A\#B \\ A\#B \circ A\#B & A \circ B \end{bmatrix} \geq 0\ .$$

Ahora, usando 3.7.8, se deduce que $A \circ B \geq (A\#B) \circ (A\#B)$ y por ende

$$\prod_{i=k}^{n} \mu_i(A \circ B) \geq \prod_{i=k}^{n} \mu_i[(A\#B) \circ (A\#B)] \quad \text{para todo} \quad k \in \mathbb{I}_n\ .$$

Finalmente, aplicando el Teorema 9.11.3 al miembro de la derecha se obtiene

$$\prod_{i=k}^{n} \mu_i[(A\#B) \circ (A\#B)] \geq \prod_{i=k}^{n} \mu_i(A\#B)^2 \quad \text{para todo} \quad k \in \mathbb{I}_n\ .$$

completanto la demostración. ■

9.13 Ejercicios

Ejercicios del texto

9.13.1. Sean $A, B \in \mathcal{H}(n)$.

1. Probar que, para cada $x \in \mathbb{C}^n$, se tiene $\langle S(A,B)x, x\rangle = 2\operatorname{Re}\langle Ax, Bx\rangle$.
2. Dar un ejemplo de matrices positivas A y B tales que $S(A,B) \not\geq 0$.

9.13.2. Encontrar dos matrices $A, B \in \mathcal{M}_2(\mathbb{C})$ tales que $|A+B| \not\leq |A| + |B|$.

9.13.3. Mostrar que la desigualdad (9.7) deja de ser cierta si en el miembro izquierdo se quita la estrella en B. (Ayuda: basta considerar matrices de 2×2 y el caso $p = q = 2$).

9.13.4. Sean $A, B, C, D \in \mathcal{M}_n(\mathbb{C})^+$ tales que $AD = DA$ y $BC = CB$. Sean $X \in \mathcal{M}_n(\mathbb{C})$ y $p, q \in [1, +\infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$. Entonces

$$\left\|\frac{1}{p}A^pXC^p + \frac{1}{q}D^qXB^q\right\|_2^2 \geq \frac{1}{r^2}\left\|A^pXC^p - D^qXB^q\right\|_2^2 + \left\|ADXCB\right\|_2^2 .$$

donde $r = \max\{p, q\}$.

9.13.5. Sean $A, B, C, D \in \mathcal{M}_n(\mathbb{C})^+$ tales que $AD = DA$, $A^*D = DA^*$, $BC = CB$ y $B^*C = CB^*$. Sean $X \in \mathcal{M}_n(\mathbb{C})$ y $p, q \in [1, +\infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$. Entonces

$$\left\|\frac{1}{p}\,|A|^p\,X\,|C|^p + \frac{1}{q}\,|D|^q\,X\,|B|^q\right\|_2^2 \geq \frac{1}{r^2}\,\left\||A|^p\,X\,|C|^p - |D|^q\,X\,|B|^q\right\|_2^2 + \left\|ADXC^*B^*\right\|_2^2 \ .$$

donde $r = \max\{p, q\}$. En particular, si $C = D = I$, mostrar que

$$\left\|\frac{1}{p}\,|A|^p\,X + \frac{1}{q}X\,|B|^q\right\|_2^2 \geq \frac{1}{r^2}\,\left\||A|^p\,X - X\,|B|^q\right\|_2^2 + \left\|AXB^*\right\|_2^2 \ ,$$

que es la extensión natural de (9.17) a matrices cualesquiera. Extender las desigualdades (9.22)-(9.24) de manera análoga.

9.13.6. Dadas $A, B \in \mathcal{G}l(n)^+$, probar que $\mu(A^{1/2}BA^{1/2})^2 \underset{\log}{\prec} \mu(AB^2A)$.

9.13.7. Dada $A \in \mathcal{M}_n(\mathbb{C})$, probar que

$$e^A = \lim_{m\to\infty}\left(I + \frac{A}{m}\right)^m \ . \tag{9.56}$$

9.13.8. Encontrar $A \in \mathcal{M}_2(\mathbb{C})$ tal que $|||e^A||| < |||e^{\operatorname{Re} A}|||$ para toda NUI.

9.13.9. 1. Dades $A, B \in \mathcal{H}(n)$, mostrar la siguiente nueva fórmula:

$$e^{A+B} = \lim_{t \to 0} \left(e^{\frac{tB}{2}} \, e^{tA} \, e^{\frac{tB}{2}}\right)^{\dfrac{1}{t}} .$$

Observar que $e^{\frac{tB}{2}} \, e^{tA} \, e^{\frac{tB}{2}} \in \mathcal{G}l(n)^+$ por lo que $\left(e^{\frac{tB}{2}} \, e^{tA} \, e^{\frac{tB}{2}}\right)^{\dfrac{1}{t}}$ tiene sentido para todo $t \neq 0$. Además, $\left(e^{\frac{tB}{2}} \, e^{tA} \, e^{\frac{tB}{2}}\right)^{-1} = e^{\frac{-tB}{2}} \, e^{-tA} \, e^{\frac{-tB}{2}}$, por lo que el caso $t < 0$ no crea problemas.

Sug: Desarrollar el producto de las tres series asociadas a $e^{\frac{tB}{2}} \, e^{tA} \, e^{\frac{tB}{2}}$ y aplicarles la serie de potencias de $-1 < x \mapsto \log(1 + x)$. Despues seguir pasos similares a los de la prueba del Teorema 9.8.2.

2. Extendiendo el argumento que utilizamos para probar el Teorema 9.8.2 probar que, dadas $A_1, A_2, \dots, A_k \in \mathcal{M}_n(\mathbb{C})$, se tiene que

$$\exp\left(\sum_{i=1}^{k} A_i\right) = \lim_{m \to \infty} \left[e^{\frac{A_1}{m}} \, e^{\frac{A_2}{m}} \, \dots \, e^{\frac{A_k}{m}}\right]^m .$$

9.13.10 (Desigualdad de Araki-Lieb-Thirring)**.** Si $A, B \in \mathcal{G}l(n)^+$, probar que

$$\operatorname{tr}\left[(B^{1/2}AB^{1/2})^{st}\right] \leq \operatorname{tr}\left[(B^{t/2}A^tB^{t/2})^s\right] , \quad \text{para todo par } s, t \geq 1 .$$

Sug: Usar el Corolario 9.9.5 para las NUI's $\|A\|_p = (\operatorname{tr}|A|^p)^{1/p}$.

Ejercicios nuevos

9.13.11. Sean $A, B, C \in \mathcal{M}_n(\mathbb{C})$. Demostrar:

1. $e^{A+B} = \lim\limits_{r \to 0} (e^{rA/2} e^{rB} e^{rA/2})^{1/r}$
2. Si $A, B \geq 0$ y $r \in (0, 1)$, entonces $(A^{r/2}B^rA^{r/2})^{1/r} \prec A^{1/2}BA^{1/2}$
3. Nuevamente bajo el supuesto que $A, B \geq 0$, se tiene que para toda norma unitariamente invariante $|||\cdot|||$:
$$|||A^{r/2}B^rA^{r/2}|||^{1/r} \underset{r \downarrow 0}{\longrightarrow} |||e^{A+B}|||,$$
en forma **decreciente**.
4. Si $A, B > 0$, entonces $|||\log(A) + \log(B)||| \leq |||\log(A^{1/2}BA^{1/2})|||$ para toda norma unitariamente invariante $|||\cdot|||$.
5. (Golden-Thompson) Si $A, B \in \mathcal{H}(n)$, entonces $\operatorname{tr}(e^{A+B}) \leq \operatorname{tr}(e^A \, e^B)$.
6. Dar un ejemplo que muestre que la desigualdad $\operatorname{tr}(e^{A+B+C}) \leq \operatorname{tr}(e^Ae^Be^C)$ es falsa.

Onda Cauchy-Schwarz

9.13.12. Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Probar que, para todo $r \in (0, +\infty)$,

$$s^r(AB) \underset{\log}{\prec_w} s^r(A)s^r(B) \implies s^r(AB) \prec_w s^r(A)s^r(B) \ .$$

9.13.13. Sean φ una fgs, y $p, q \in [1, +\infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$. Dados $x, y \in \mathbb{R}^n$, se tiene que

$$\varphi(|x \cdot y|) \leq \varphi(|x|^p)^{\frac{1}{p}} \cdot \varphi(|y|^q)^{\frac{1}{q}} \ .$$

Sug: usando Hölder (de números, en las coordenadas) mostrar que, para todo $t > 0$ se tiene

$$\varphi(|x \cdot y|) \leq \frac{t^p \, \varphi(|x|^p)}{p} + \frac{\varphi(|y|^q)}{q \, t^q} \ .$$

Luego calcular el mínimo de esas cosas.

9.13.14. Sea N una NUI en $\mathcal{M}_n(\mathbb{C})$.

1. Sean $p, q \in [1, +\infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$. Mostrar que

$$N(AB) \leq N(|A|^p)^{\frac{1}{p}} \cdot N(|B|^q)^{\frac{1}{q}} \quad \text{para todo par } A, B \in \mathcal{M}_n(\mathbb{C}) \ .$$

2. Más aún, si tomamos p, q y r positivos tales que $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, probar que

$$N(|AB|^r)^{\frac{1}{r}} \leq N(|A|^p)^{\frac{1}{p}} \cdot N(|B|^q)^{\frac{1}{q}} \quad \text{para todo par } A, B \in \mathcal{M}_n(\mathbb{C}) \ .$$

3. Deducir que $N(|AB|^{1/2}) \leq N(A)^{1/2} \cdot N(B)^{1/2}$ y que $N(AB) \leq N(A^*A)^{1/2} \cdot N(B^*B)^{1/2}$.

4. Otra mejora: Dadas $A, B, X \in \mathcal{M}_n(\mathbb{C})$ se tiene que

$$N(AXB^*)^2 \leq N(A^*A\,X) \cdot N(X\,B^*B) \ .$$

5. Peor todavía: si $r \in \mathbb{R}_+^*$, entonces debe valer que

$$N(\,|AXB^*|^r\,)^2 \leq N(\,|A^*A\,X|^r\,) \cdot N(\,|X\,B^*B|^r\,) \ .$$

6. Ahora deducir que, si $s \in [0, 1]$, entonces se cumple que

$$N(A^sXB^{1-s}) \leq N(AX)^s \, N(XB)^{1-s} \quad \text{y que} \quad N(A^sXB^s) \leq N(X)^{1-s} \, N(AXB)^s \ .$$

Observar que tomando $X = I$ en la de la derecha, se obtiene una generalización de la desigualdad de Cordes (9.48) a todas las NUI's (porque $\|I\|_{sp} = 1$).

Capítulo 10

Rango y Radio Numéricos

En este capítulo consideraremos espacios de Hilbert $\mathbb{H}$ finito deimensionales (con su producto interno y la norma euclídea). Por lo tanto, puede pensarse que $\mathbb{H} = \mathbb{C}^n$, e identificar el álgebra de operadores $L(\mathbb{H})$ con $\mathcal{M}_n(\mathbb{C})$. La mayoría de los resultados de este capítulo valen también para el caso infinitodimensional, salvo el hacho de que en varias de las igualdades e inclusiones que aparecen, el rango numérico (o las cápsulas convexas) deberían cambiarse por sus clausuras. Una versión general aparece en el tomo II de este libro.

10.1 Definiciones y propiedades básicas

Definición 10.1.1. Sea $A \in L(\mathbb{H})$.

1. El **Rango numérico** de A es el conjunto

$$W(A) = \{\langle Ax, x\rangle \ : \ x \in \mathbb{H}, \ \|x\| = 1 \ \} \ .$$

2. Recordemos que el **radio numérico** de A se define como

$$w(A) = \max_{\lambda \in W(A)} |\lambda| = \max\{ \ |\langle Ax, x\rangle| \ : \ x \in \mathbb{H}, \ \|x\| = 1 \ \}$$

   y que define una norma en $L(\mathbb{H})$ (si el cuerpo es $\mathbb{C}$). ▲

10.1.2. A continuación enumeraremos una serie de propiedades elementales del rango y radio numéricos que se siguen fácilmente sus definiciones. Las pruebas que no estén escritas deben considerarse como ejercicios. Sea $A \in L(\mathbb{H})$.

1. Se cumplen las siguientes desigualdades:

$$\rho(A) \leq w(A) \leq \|A\|_{sp} \ .$$

   La segunda se deduce de Cauchy-Schwarz. La primera se deduce de la fórmula (10.1) de un poco más abajo.

2. Tomando $T = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}$, se ve que las desiguadades pueden ser estrictas. En efecto, es claro que $\rho(T) = 0$ y $\|T\|_{sp} = 2$. Por otra parte, como la función

$$f : \{z \in \mathbb{C} : |z| \le 1\} \to \mathbb{R} \quad \text{ dada por } \quad f(z) = 2|z|\,(1 - |z|^2)^{1/2}$$

alcanza el máximo $f(z) = 1$ cuando $|z|^2 = 1/2$, podemos deducir que $w(T) = 1$.

3. Vimos en el Corolario 2.1.4 que, si A es **normal**, entonces $\rho(A) = w(A) = \|A\|_{sp}$.

4. Recordemos que $\operatorname{Re} A = \dfrac{A + A^*}{2}$ e $\operatorname{Im} A = \dfrac{A - A^*}{2i}$. Se tiene que

$$\|\operatorname{Re} A\|_{sp} = w(\operatorname{Re} A) = \max_{\|x\|=1} |\operatorname{Re}\langle Ax\,,x\rangle| \le w(A) \ ,$$

y lo mismo vale para $\operatorname{Im} A$. Luego $\|A\|_{sp} \le 2 \cdot w(A)$.

5. Dado $B \in L(\mathbb{H})$ se cumple que $W(A + B) \subseteq W(A) + W(B)$.

6. Dado $\lambda \in \mathbb{C}$, se tiene que $W(A + \lambda I) = W(A) + \lambda$ y $W(\lambda \cdot A) = \lambda \cdot W(A)$.

7. Si $U \in \mathcal{U}(\mathbb{H})$, entonces $W(UAU^*) = W(A)$ y $w(UAU^*) = w(A)$.

8. $W(A)$ es compacto (por serlo la cáscara de la bola unidad de $\mathbb{H}$). ▲

Proposición 10.1.3. *Sea $A \in L(\mathbb{H})$. Entonces*

$$\sigma(A) \subseteq W(A) \ . \tag{10.1}$$

Demostración. Si λ es autovalor de A, es claro que $\lambda = \langle Ax, x\rangle \in W(A)$ para cualquier autovector unitario x asociado a λ. ■

10.2 El Teorema de Hausdorff Töeplitz

El Teorema de Hausdorff Töeplitz dice que, para todo $A \in L(\mathbb{H})$, se cumple que $W(A)$ es **convexo**. Para probarlo se necesitan una serie de reducciones. La principal es ver que basta probarlo para matrices en $\mathcal{M}_2(\mathbb{C})$ (esto lo veremos en la prueba del Teorema). Pero aún entre ellas, necesitamos dos reducciones especiales:

Lema 10.2.1. *Dada $A \in \mathcal{M}_2(\mathbb{C})$, existe $U \in \mathcal{U}(2)$ tal que,*

$$B = U^*AU = \begin{bmatrix} c & a \\ b & c \end{bmatrix} \ , \qquad \textit{con} \qquad c = \frac{tr\,A}{2} \ .$$

Demostración. Cambiando A por $A - \dfrac{\operatorname{tr} A}{2}\, I$, podemos suponer que $\operatorname{tr} A = 0$ y tratar de hacer que la diagonal de B sea nula. Si $\sigma(A) = \{0\}$, esto es fácil (por el Teorema 1 de Schur 1.6.1). Sinó, $\sigma(A) = \{\lambda, -\lambda\}$ con $\lambda \neq 0$. Sean x_1 y x_2 autovectores unitarios asociados a λ y $-\lambda$, respectivamente. Tomemos la curva $x(t) = e^{it}x_1 + x_2$, para $t \in [0, 2\pi]$. Observar que $x(t) \neq 0$, por que x_1 y x_2 son LI. Entonces,

$$\begin{aligned} \langle Ax(t), x(t)\rangle &= \lambda - \lambda + e^{it}\langle Ax_1, x_2\rangle + e^{-it}\langle Ax_2, x_1\rangle \\ &= \lambda e^{it}\langle x_1, x_2\rangle - \lambda e^{-it}\langle x_2, x_1\rangle = 2i\lambda \operatorname{Im}\left(e^{it}\langle x_1, x_2\rangle\right) . \end{aligned}$$

Eligiendo $t_0 \in [0, 2\pi]$ tal que $e^{it_0}\langle x_1, x_2\rangle \in \mathbb{R}$, tenemos que $\langle Ax(t_0), x(t_0)\rangle = 0$, con $x(t_0) \neq 0$. Normalizando a $x(t_0)$, completando a una BON de $\mathbb{C}^2$, y tomando $U \in \mathcal{U}(2)$ tal que tenga a esa BON en sus columnas, obtenemos

$$B = U^*AU = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix} , \quad \text{con} \quad a, b \in \mathbb{C} ,$$

donde $B_{22} = 0$ porque $B_{11} = 0 = \operatorname{tr} B$. ■

Lema 10.2.2. *Dada $B \in \mathcal{M}_2(\mathbb{C})$ con diagonal nula, existen $V \in \mathcal{U}(2)$ y $w \in \mathbb{C}$ con $|w| = 1$ tales que,*

$$w \cdot VBV^* = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix} , \quad \textit{con} \quad a \geq 0 \quad \textit{y} \quad b \geq 0 .$$

Demostración. Si $B = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}$, tomando $V = \begin{bmatrix} u & 0 \\ 0 & 1 \end{bmatrix} \in \mathcal{U}(2)$ y $w \in \mathbb{C}$ con $|w| = 1$, tenemos que

$$w \cdot VB_1V^* = \begin{bmatrix} 0 & wu\,a \\ w\overline{u}\,b & 0 \end{bmatrix} .$$

Si $a = e^{i\theta_1}|a|$ y $b = e^{i\theta_2}|b|$, tomando $u = e^{\frac{i}{2}(\theta_2 - \theta_1)}$ y $w = e^{\frac{i}{2}(\theta_2 + \theta_1)}$, se obtiene que

$$w \cdot VB_1V^* = \begin{bmatrix} 0 & |a| \\ |b| & 0 \end{bmatrix} ,$$

como deseábamos. ■

Teorema 10.2.3 (Hausdorff-Töeplitz)**.** *Sea $A \in L(\mathbb{H})$. Entonces $W(A)$ es convexo.*

Demostración. Sean $\alpha, \beta \in W(A)$ distintos, y sean $x, y \in \mathbb{H}$ unitarios tales que $\langle Ax, x\rangle = \alpha$ y $\langle Ay, y\rangle = \beta$. Tomemos $\mathcal{B}_0 = \{v_1, v_2\}$ una BON de $\mathcal{S} = \operatorname{Gen}\{x, y\}$. Consideremos la compresión $A_{\mathcal{S}} \in L(\mathcal{S})$. La matriz de $A_{\mathcal{S}}$ en la base $\mathcal{B}_0$ es $B = (\langle Av_j, v_i\rangle)_{i,j \in \mathbb{I}_2} \in \mathcal{M}_2(\mathbb{C})$. Dado $z = (z_1, z_2) \in \mathbb{C}^2$, se tiene que

$$w = z_1v_1 + z_2v_2 \in \mathcal{S} , \quad \|w\| = \|z\|_2 \quad \text{y} \quad \langle Bz, z\rangle = \langle Aw, w\rangle ,$$

por lo que $\alpha, \beta \in W(B)$ y, para probar que las combinaciones convexas de α y β están en $W(A)$, basta verificar que están en $W(B)$. En otras parabras, alcanza con probar el teorema en el caso de que $A \in \mathcal{M}_2(\mathbb{C})$. Para ello, por los Lemas 10.2.1 y 10.2.2, se puede asumir que

$$A = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix} , \quad \text{con} \quad a \geq 0 \quad \text{y} \quad b \geq 0 ,$$

puesto que $W(C+\lambda I) = W(C) + \lambda$ y $W(u \cdot VCV^*) = u \cdot W(C)$ para cualquier $C \in \mathcal{M}_2(\mathbb{C})$, $\lambda \in \mathbb{C}$, $V \in \mathcal{U}(2)$ y $u \in \mathbb{C}$ con $|u| = 1$. Obervar que los cambios inducidos por las reducciones anteriores (translaciones y rotaciones) no perturban el hecho de que $W(A)$ sea convexo. Veremos que en este caso,

$$W(A) = \Big\{ t\big(\ (a+b)\cos\theta + i(a-b)\,\mathrm{sen}\,\theta\big)\ :\ t \in [0,1/2] \quad \text{y} \quad \theta \in [0,2\pi]\ \Big\}\ , \tag{10.2}$$

que es una elipse (o eventualmente un segmento) centrada en el origen, y por lo tanto convexa. En efecto, dado $z \in \mathbb{C}^2$ con $\|z\| = 1$, como

$$\langle Az, z\rangle = \big\langle Ae^{i\alpha}z, e^{i\alpha}z\big\rangle \qquad \text{para todo} \qquad \alpha \in \mathbb{R}\ ,$$

podemos suponer que $z = (t, (1-t^2)^{1/2}e^{i\theta})$ para $t \in [0,1]$ y $\theta \in [0,2\pi]$. En tal caso, cuentas elementales muestran que

$$\langle Az, z\rangle = t(1-t^2)\big[\ (a+b)\cos\theta + i(a-b)\,\mathrm{sen}\,\theta\,\big]\ .$$

Observar que los números $t(1-t^2)$ recorren el intervalo $[0,1/2]$ cuando $t \in [0,1]$. Esto prueba la fórmula (10.2). ■

Corolario 10.2.4. *Sea* $A \in L(\mathbb{H})$.

1. *En general se cumple que* $\mathrm{conv}\,[\sigma\,(A)\,] \subseteq W(A)$.
2. *Si* A *es* **normal**, *entonces* $\mathrm{conv}\,[\sigma\,(A)\,] = W(A)$.

Demostración. La inclusión $\sigma\,(A) \subseteq W(A)$ ya fue vista en la Proposición 10.1.3. Pero por el Teorema 10.2.3, sabemos que esa incusón arrastra a la cápsula convexa. Si A es normal, sea $\{x_1,\dots,x_n\}$ una BON de $\mathbb{H}$ formada por autovectores de A asociados a sus autovalores $\lambda_1\,,\dots,\lambda_n$. Si $x \in \mathbb{H}$ tiene $\|x\| = 1$, entonces

$$\langle Ax, x\rangle = \left\langle A\sum_{k=1}^n \langle x, x_k\rangle\, x_k\,, \sum_{k=1}^n \langle x, x_k\rangle\, x_k \right\rangle = \sum_{k=1}^n |\,\langle x, x_k\rangle\,|^2\,\lambda_k \in \mathrm{conv}\,[\sigma\,(A)\,]\ ,$$

porque $\sum\limits_{k=1}^n |\,\langle x, x_k\rangle\,|^2 = \|x\|^2 = 1$. Por lo tanto $W(A) \subseteq \mathrm{conv}\,[\sigma\,(A)]$. ■

Definición 10.2.5. 1. Dados dos espacios de Hilbert $\mathbb{H}$ y $\mathbb{K}$, notamos

$$\mathbb{H} \oplus \mathbb{K} = \{(x,y) : x \in \mathbb{H}\ ,\ y \in \mathbb{K}\}\ ,$$

que es un espacio de Hilbert con el producto $\big\langle (x_1\,,y_1)\,,(x_2\,,y_2)\big\rangle = \langle x_1, x_2\rangle + \langle y_1, y_2\rangle$.

2. Dados $A \in L(\mathbb{H})$ y $B \in L(\mathbb{K})$, se define el operador

$$A \oplus B \in L(\mathbb{H} \oplus \mathbb{K}) \qquad \text{dado por} \qquad A \oplus B(x,y) = (Ax, By),\ \ (x,y) \in \mathbb{H} \oplus \mathbb{K}\ .$$

3. Matricialmente tenemos que

$$A \oplus B := \begin{array}{c} \left[\begin{array}{cc} A & 0 \\ 0 & B \end{array}\right] \end{array} \begin{array}{c} \mathbb{H} \\ \mathbb{K} \end{array} \in L(\mathbb{H} \oplus \mathbb{K}) \ .$$

4. En forma similar se definen sumas directas de muchos espacios de Hilbert y de muchos operadores en ellos. ▲

Corolario 10.2.6. Sean $A \in L(\mathbb{H})$ y $B \in L(\mathbb{K})$. Entonces

$$W(A \oplus B) = \operatorname{conv}\left[W(A) \cup W(A) \right] \quad \text{ y } \quad w(A \oplus B) = \text{máx}\{w(A), w(B)\} \ . \tag{10.3}$$

Idem con muchos bloques diagonales.

Demostración. La inclusión $W(A) \cup W(A) \subseteq W(A \oplus B)$ se testea inmediatamente usando vectores con una coordenada nula. Por el Teorema 10.2.3, esto arrastra a la cápsula convexa de $W(A) \cup W(A)$. Recíprocamente, dados $x \in \mathbb{H}$ e $y \in \mathbb{K}$ no nulos tales que $\|(x,y)\|^2 = \|x\|^2 + \|y\|^2 = 1$, tenemos que

$$\begin{aligned} \langle A \oplus B(x,y), (x,y) \rangle &= \langle Ax, x \rangle + \langle By, y \rangle \\ &= \|x\|^2 \left\langle A \frac{x}{\|x\|} , \frac{x}{\|x\|} \right\rangle + \|y\|^2 \left\langle B \frac{y}{\|y\|} , \frac{y}{\|y\|} \right\rangle \ , \end{aligned}$$

quien claramente pertenece a $\operatorname{conv}\left[W(A) \cup W(A) \right]$. ■

Corolario 10.2.7. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces existe $U \in \mathcal{U}(n)$ tal que, si $B = U^*AU$, luego $B_{ii} = \dfrac{\operatorname{tr} A}{n}$ para todo $i \in \mathbb{I}_n$.*

Demostración. Cambiando A por $A - \dfrac{\operatorname{tr} A}{n} I$, lo que debemos probar es que, si $\operatorname{tr} A = 0$, entonces podemos conseguir $U \in \mathcal{U}(n)$ tal que la diagonal de U^*AU sea nula. Lo probaremos por inducción en n. Para $n = 1$ es trivial. Observar que el caso $n = 2$ es el Lema 10.2.1. Si $n > 2$, aplicando el Corolario 10.2.4 obtenemos que

$$0 = \frac{\operatorname{tr} A}{n} \in \operatorname{conv}\left[\sigma\left(A\right)\right] \subseteq W(A) \ .$$

Luego existe un vector unitario $x \in \mathbb{C}^n$ tal que $\langle Ax, x \rangle = 0$. Completando $\{x\}$ a una BON de $\mathbb{C}^n$ que lo tenga como primer elemento, y tomando $U_1 \in \mathcal{U}(n)$ la matriz con esa BON en sus columnas, obtenemos que

$$C = U_1^* A U_1 = \left[\begin{array}{cc} 0 & * \\ * & D \end{array}\right] \begin{array}{c} 1 \\ n-1 \end{array} \ ,$$

porque $C_{11} = \langle Ax, x \rangle = 0$. Como $D \in \mathcal{M}_{n-1}(\mathbb{C})$ cumple que $\operatorname{tr} D = 0$, podemos aplicar la hipótesis inductiva y encontrar $V \in \mathcal{U}(n-1)$ tal que la diagonal de V^*DV sea nula. Definiendo $U_2 = 1 \oplus V \in \mathcal{U}(n)$ y $U = U_1 U_2 \in \mathcal{U}(n)$, se ve que

$$U^*AU = U_2^* C U_2 = \left[\begin{array}{cc} 1 & 0 \\ 0 & V^* \end{array}\right] \left[\begin{array}{cc} 0 & * \\ * & D \end{array}\right] \left[\begin{array}{cc} 1 & 0 \\ 0 & V \end{array}\right] = \left[\begin{array}{cc} 0 & *V \\ V^* * & V^*DV \end{array}\right] \ ,$$

que tiene la diagonal nula. ■

10.3 Caracterizaciones

Observación 10.3.1. Sea $W \subseteq \mathbb{C}$ un conjunto convexo compacto, y sea $z_0 \notin W$. Entonces existe un único $w_0 \in W$ tal que

$$d(z_0, W) = |z_0 - w_0| = d > 0 \ .$$

El tal w_0 existe (y es único) por la teoría usual de espacios de Hilbert, usando que W es convexo y cerrado (para una cuenta ad hoc, ver el Ejercicio 10.5.2). Más aún, si $x_0 = \overline{z_0 - w_0}$, entonces $\mathrm{Re}\,(x_0 w_0) + d^2 = \mathrm{Re}\,(x_0 z_0)$ y

$$\mathrm{Re}\,(x_0 z) \le \mathrm{Re}\,(x_0 w_0) \quad \text{ para todo } \quad z \in W \ . \tag{10.4}$$

Esto se deduce de que w_0 es la *proyección ortogonal* de z_0 sobre W, y de que el producto escalar en $\mathbb{C}$ pensado como $\mathbb{R}^2$ está dado por $\langle z \,, w \rangle = \mathrm{Re}\,(\overline{w}\, z)$. Observar que la recta $\{z \in \mathbb{C} : \mathrm{Re}\,[x_0(z - w_0)] = 0\}$ es ortogonal a $z_0 - w_0$ y pasa por w_0 . ▲

Teorema 10.3.2. *Sea $A \in L(\mathbb{H})$. Entonces*

$$\begin{aligned} W(A) &= \Big\{ z \in \mathbb{C} : |z - \lambda| \le w(A - \lambda I) \quad \text{para todo} \quad \lambda \in \mathbb{C} \Big\} \\ &= \Big\{ z \in \mathbb{C} : |z - \lambda| \le \|A - \lambda I\|_{sp} \quad \text{para todo} \quad \lambda \in \mathbb{C} \Big\} \ . \end{aligned}$$

Demostración. Notemos $W = W(A)$, X al conjunto de arriba e Y al conjunto de abajo. Es claro, por las definiciones, que $X \subseteq Y$. Usando que $W(A) - \lambda = W(A - \lambda I)$, es fácil ver que $W \subseteq X$. En lo que sigue probaremos que $Y \subseteq W$: Supongamos que $z_0 \notin W$, y sea w_0 la proyección ortogonal de z_0 sobre W (como en la Observación 10.3.1).

Para facilitar las cuentas, rotaremos y transladaremos el problema para que $z_0 \in \mathbb{R}_+$, $w_0 = 0$ y $W \subseteq \{z \in \mathbb{C} : \mathrm{Re}\, z \le 0\}$. Para hacerlo, sean

$$d = |z_0 - w_0| = d(z_0 \,, W) > 0 \ , \quad \text{ y } \quad B = e^{-i\theta}(A - w_0 I) \in L(\mathbb{H}) \ ,$$

donde $z_0 - w_0 = e^{i\theta} d$. Luego $W_B = W(B) = e^{-i\theta}(W(A) - w_0)$ y si tomamos el conjunto

$$Y_B = \big\{ z \in \mathbb{C} : |z - \lambda| \le \|B - \lambda I\|_{sp} \,, \ \lambda \in \mathbb{C} \big\}, \quad \text{ entonces } \quad Y_B = e^{-i\theta}(Y - w_0) \ .$$

Por lo tanto, para ver que $z_0 \notin Y$, alcanza probar que $d = e^{-i\theta}(z_0 - w_0) \notin Y_B$. Observar que, como la función $x \mapsto e^{-i\theta}(x - w_0)$ preserva distancias, la proyección de d a W_B es, ahora, $e^{-i\theta}(w_0 - w_0) = 0$. Además, como $d = d - 0 > 0$, si $z \in W_B$, la Eq. (10.4) dice que

$$\mathrm{Re}\,(z\, \overline{d}\,) = d\, \mathrm{Re}\, z \le 0 \quad \Longrightarrow \quad \mathrm{Re}\, z \le 0 \ ,$$

como queríamos. Ahora, dados $x \in \mathbb{C}^n$ con $\|x\| = 1$ y $m \in \mathbb{N}$, tendremos que

$$\begin{aligned} \|(B + mI)x\|^2 &= \big\langle (B + mI)x \,, (B + mI)x \big\rangle \\ &= \|Bx\|^2 + m^2 + 2m\, \mathrm{Re}\, \langle Bx, x\rangle \\ &\le \|B\|_{sp}^2 + m^2 \ , \end{aligned}$$

porque $\langle Bx, x\rangle \in W_B$. Es decir que $\|B+mI\|_{sp} \le (\|B\|_{sp}^2 + m^2)^{1/2}$. Por otro lado, es fácil ver que $(\|B\|_{sp}^2 + m^2)^{1/2} - m \xrightarrow[m\to\infty]{} 0$. Por lo tanto, debe existir $m \in \mathbb{N}$ tal que

$$\|B+mI\|_{sp} - m \le (\|B\|_{sp}^2 + m^2)^{1/2} - m < d \ .$$

En otras palabras, para ese m se tiene que

$$\|B+mI\|_{sp} < d + m = |d+m| \ ,$$

por lo que $d \notin Y_B$ y entonces $z_0 \notin Y$. Resumiendo, vimos que si $z_0 \notin W$, entonces $z_0 \notin Y$, o sea que $Y \subseteq W$. ∎

A continuación damos un adelanto del libro II. Se trata de una caracterización obtenida por T. Ando [17] del radio numérico, que es tan útil como la caracetrización de la norma espectral dada en la Proposición 3.7.6. Su prueba necesita bastantes desarrollos específicos que se trabajarán en el libro II, por lo que no creemos necesario reproducirlos aquí. Sin embargo lo enunciamos ahora, en su versión matricial, porque es uno de los criterios básicos para trabajar con rangos numéricos.

Teorema 10.3.3 (Ando 1973). *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Son equivalentes:*

1. $w(A) \le 1$.
2. *Para todo $\theta \in [0, 2\pi)$ se tiene que* $\operatorname{Re}(e^{i\theta}A) \le I$.
3. *Existen $C \in \mathcal{M}_n(\mathbb{C})$ e $Y \in \mathcal{M}_n(\mathbb{C})^+$ tales que*
$$Y \le I \quad , \quad \|C\|_{sp} \le 2 \quad y \quad (I-Y)^{1/2}\, C\, Y^{1/2} \;=\; A \ .$$
4. *Existe $Y \in \mathcal{M}_n(\mathbb{C})^+$ tal que* $M_1 = \begin{bmatrix} I-Y & A^*/2 \\ A/2 & Y \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})^+$.
5. *Existe $L \in \mathcal{H}(n)$ tal que* $M_2 = \begin{bmatrix} I+L & A^* \\ A & I-L \end{bmatrix} \in \mathcal{M}_{2n}(\mathbb{C})^+$.

Demostración. Si vale la condición 2, como existen $x \in \mathbb{C}^n$ unitario y $\theta \in [0, 2\pi)$ tales que

$$w(A) = |\langle Ax, x\rangle| = e^{i\theta}\langle Ax, x\rangle = \langle (e^{i\theta}A)\, x, x\rangle = \langle \operatorname{Re}(e^{i\theta}A)\, x, x\rangle \le \langle I\, x, x\rangle = 1 \ ,$$

tenemos que $2 \to 1$. Por otro lado, la implicación $1 \to 2$ es bien fácil. Siguiendo, tenemos que la equivalencia $3 \leftrightarrow 4$ no es otra cosa que el Teorema 3.8.6. Veamos que $4 \to 5$: Supongamos que $M_1 \in \mathcal{M}_{2n}(\mathbb{C})^+$,, para cierto $Y \in \mathcal{M}_n(\mathbb{C})^+$. Luego, si $L = I - 2Y \in \mathcal{H}(n)$, se tiene que

$$I + L = 2(I-Y) \text{ y } I - L = 2Y \quad \Longrightarrow \quad \begin{bmatrix} I+L & A^* \\ A & I-L \end{bmatrix} = 2\begin{bmatrix} I-Y & A^*/2 \\ A/2 & Y \end{bmatrix} \ge 0 \ .$$

Ahora veamos que $5 \to 2$: Dados $x \in \mathbb{C}^n$ y $\theta \in [0, 2\pi)$, tomemos el vector $y = -e^{i\theta}x$. Luego, si asumimos que L cumple la condición 5, tendremos que

$$2\|x\|^2 = \left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\|^2 \ge \left\langle \begin{bmatrix} -L & -A \\ -A^* & L \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix} \right\rangle = 2\left\langle \operatorname{Re}(e^{i\theta}A)\, x, x\right\rangle \ . \tag{10.5}$$

Luego $\operatorname{Re}(e^{i\theta}A) \le I$. Lamentablemente, la implicación $1 \to 4$ quedará para el libro II. ∎

10.4 Comparación con NUI's

Proposición 10.4.1. *Sea $A \in L(\mathbb{H})$. Entonces*

$$\frac{1}{2}\,\|A\| \le w(A) \le \|A\| \ .$$

Admás, las constantes 1 *y* 1/2 *son óptimas para la desigualdad anterior.*

Demostración. Tomemos partes real e imaginaria: $A = \operatorname{Re} A + i \operatorname{Im} A$. Luego

$$w(A) \le \|A\| \le \|\operatorname{Re} A\| + \|\operatorname{Im} A\| = w(\operatorname{Re} A) + w(\operatorname{Im} A) \le 2 \cdot w(A) \ ,$$

donde la última desigualdad se deduce de que $W(\operatorname{Re} A) = \{\operatorname{Re} z : z \in W(A)\}$, por lo que $w(\operatorname{Re} A) \le w(A)$, y lo mismo para $\operatorname{Im} A$. La optimalidad de las constantes 1 y 1/2 se ve tomando las matrices E_{11} y $2E_{21}$. ∎

Proposición 10.4.2 (Marcus-Sandy 1985)**.** *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces*

$$\frac{1}{n}\sum_{i=1}^{n} s_i(A) \ \le \ w(A) \le \sum_{i=1}^{n} s_i(A) = \|A\|_1 \ .$$

Admás, las constantes 1 *y* $1/n$ *son óptimas para la desigualdad anterior.*

Demostración. Tomemos la descomposición polar $A = U|A|$, con $U \in \mathcal{U}(n)$. Conjugando con otra matriz unitaria (lo que no cambia ni $w(A)$ ni $\|A\|_1$), podemos suponer que que U es diagonal. Observar que $|VAV^*| = V|A|V^*$, si $V \in \mathcal{U}(n)$. Pongamos $U = \operatorname{diag}(w_1, \dots w_n)$, con $|w_i| = 1$, $i \in \mathbb{I}_n$. Veremos que, en este caso, el número $\displaystyle\frac{1}{n}\sum_{i=1}^{n} s_i(A)$ es superado por alguno de los módulos de los elementos diagonales de A. En efecto, si notamos $\{e_1, \dots, e_n\}$ a la base canónica de $\mathbb{C}^n$, y llamamos $P = |A|$, dado $k \in \mathbb{I}_n$ tenemos que

$$|A_{kk}| = |\langle Ae_k, e_k\rangle| = |\langle UPe_k, e_k\rangle| = |\langle Pe_k, U^*e_k\rangle| = |w_k \langle Pe_k, e_k\rangle| = |P_{kk}| = P_{kk} \ ,$$

donde la última igualdad surge de que $P \in \mathcal{M}_n(\mathbb{C})^+$. Por otra parte,

$$\sum_{k=1}^{n} P_{kk} = \operatorname{tr} P = \operatorname{tr}|A| = \sum_{i=1}^{n} s_i(A) \implies \frac{1}{n}\sum_{i=1}^{n} s_i(A) \le P_{kk} = |\langle Ae_k, e_k\rangle| \le w(A)$$

para algún $k \in \mathbb{I}_n$. Para ver que 1 y $1/n$ son óptimas, tomar las matrices E_{11} e I. ∎

Definición 10.4.3. Llamemos $A = \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix} \in \mathcal{M}_2(\mathbb{C})$ y $V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \in \mathcal{U}(2)$. Dado $n \in \mathbb{N}$, si $n = 2m$ consideremos las matrices diagonales de bloques

$$C_n = A \oplus A \oplus \cdots \oplus A = \sum_{k=1}^{m} 2E_{2k,2k-1} \in \mathcal{M}_n(\mathbb{C}) \quad \text{y}$$

$$U_n = V \oplus V \oplus \cdots \oplus V = \sum_{k=1}^{m} E_{2k,2k-1} + E_{2k-1,2k} \in \mathcal{U}(n) \ .$$

Si $n = 2m+1$ e I_1 denota a la "matriz" $[\ 1\ ] \in \mathcal{M}_1(\mathbb{C})$,

$$C_n = A \oplus \cdots \oplus A \oplus I_1 = E_{n,n} + \sum_{k=1}^{m} 2E_{2k,2k-1} \in \mathcal{M}_n(\mathbb{C}) \quad \text{y}$$

$$U_n = V \oplus \cdots \oplus V \oplus I_1 = E_{n,n} + \sum_{k=1}^{m} E_{2k,2k-1} + E_{2k-1,2k} \in \mathcal{U}(n) \ .$$

Como $w(A) = w(I_1) = 1$, la Eq. (10.3) asegura que $w(C_n) = 1$ para todo $n \in \mathbb{N}$. ▲

Los resultados anteriores fueron usados por C.R. Johnson y C.K. Li [27] para calcular, para N una NUI fija en $\mathcal{M}_n(\mathbb{C})$, las mejores constantes m y M tales que

$$m \cdot N(T) \le w(T) \le M \cdot N(T) \quad \text{para todo} \quad T \in \mathcal{M}_n(\mathbb{C}) \ .$$

Proposición 10.4.4 (Johnson-Li 1988)**.** *Sea N una NUI en $\mathcal{M}_n(\mathbb{C})$. Luego*

$$N(C_n)^{-1}\ N(T) \ \le w(T) \ \le \ N(E_{11})^{-1}\ N(T) \quad \textit{para toda} \quad T \in \mathcal{M}_n(\mathbb{C}) \ .$$

Además, las constantes $N(C_n)^{-1}$ y $N(E_{11})^{-1}$ son óptimas.

Demostración. Fijemos $T \in \mathcal{M}_n(\mathbb{C})$. Si $T = 0$, el resultado es claro. Si no, sea $A = w(T)^{-1}\, T$, que tiene $w(A) = 1$. Por las Proposiciones 10.4.1 y 10.4.2 se tiene que

$$1 \le s_1(A) \le 2 \quad \text{y} \quad \sum_{k=1}^{n} s_k(A) \le n \ .$$

Observar que $s(E_{11}) = e_1$ y $s(C_n) = v_n$, donde

$$v_n = \begin{cases} (2,\dots,2,0,\dots,0) = \displaystyle\sum_{k=1}^{m} 2e_k & \text{si} \quad n = 2m \\ \\ (2,\dots,2,1,0,\dots,0) = \displaystyle\sum_{k=1}^{m} 2e_k + e_{m+1} & \text{si} \quad n = 2m+1 \ . \end{cases} \tag{10.6}$$

Por lo tanto, $s(E_{11}) \prec_w s(A) \prec_w s(C_n)$. Como N es una NUI, el Teorema de Ky Fan 5.3.8 dice que que

$$N(E_{11}) \ \le \ N(A) = \frac{N(T)}{w(T)} \ \le \ N(C_n) \ .$$

Invirtiendo y multiplicando por $N(T)$, se obtienen las desigualdades buscadas. Tomando $T = C_n$ y $T = E_{11}$ y observando que $w(C_n) = w(E_{11}) = 1$, podemos deducir que las constantes dadas son óptimas. ■

Proposición 10.4.5. *Sea N es una NUI en $\mathcal{M}_n(\mathbb{C})$. Entonces*

$$w(T) \leq N(T) \ , \ \text{para toda } T \in \mathcal{M}_n(\mathbb{C}) \implies \|T\|_{sp} \leq N(T) \ , \ \text{para toda } T \in \mathcal{M}_n(\mathbb{C}) \ .$$

Demostración. Observar que $\|T\|_{sp} = \|\,|T|\,\|_{sp} = w(|T|) \leq N(|T|) = N(T)$. ■

El siguiente teorema es el contenido del paper de T. Ando [18]:

Teorema 10.4.6 (Ando 2005)**.** Si definimos la norma

$$N_0(T) = \text{máx}\left\{\frac{\|T\|_{sp}}{2} \ , \ \frac{\|T\|_1}{n}\right\} \qquad \text{para} \qquad T \in \mathcal{M}_n(\mathbb{C}) \ ,$$

se tiene que

1. N_0 es una NUI.
2. $N_0(T) \leq w(T)$ para todo $T \in \mathcal{M}_n(\mathbb{C})$.
3. N_0 es la **mayor** NUI en $\mathcal{M}_n(\mathbb{C})$ tal que $N(T) \leq w(T)$ para todo $T \in \mathcal{M}_n(\mathbb{C})$. Es decir, si N es una NUI en $\mathcal{M}_n(\mathbb{C})$,

$$N(T) \leq w(T) \ , \ \ \forall \ \ T \in \mathcal{M}_n(\mathbb{C}) \quad \implies \quad N(T) \leq N_0(T) \ , \ \ \forall \ \ T \in \mathcal{M}_n(\mathbb{C}) \ .$$

Demostración. Los dos primeros items son claros de lo anterior. Fijemos $T \in \mathcal{M}_n(\mathbb{C})$. Como la desigualdad a probar es entre normas unitariamente invariantes, podemos asumir que $T = \Sigma(T) = \text{diag}\,(s_1(T), \dots, s_n(T))$. Más aún, supongamos que

$$N_0(T) = \text{máx}\left\{ \ \frac{s_1(T)}{2} \ , \ \frac{1}{n}\sum_{i=1}^{n} s_i(T) \ \right\} = 1 \ .$$

En este caso deberíamos probar que $N(T) \leq 1$. Las desigualdades resultantes de la igualdad anterior implican que $s(T) \prec_w v_n$, donde v_n es el vector definido en la Eq. (10.6). Tomemos C_n y U_n las matrices de la Definición 10.4.3. Notar que $U_n C_n = B_n$, donde

$$B_n = \text{diag}\,(2, 0, 2, 0, \dots, 2, 0) \qquad \text{o bien} \qquad B_n = \text{diag}\,(2, 0, 2, 0, \dots, 2, 0, 1) \ .$$

Observar que $s(B) = v_n$. Luego, como $s(T) \prec_w v_n = s(B_n)$ y N es una NUI, el Teorema de Ky Fan 5.3.8 nos dice que que $N(T) \leq N(B_n)$, con lo que bastaría probar que $N(B_n) \leq 1$. Por otra parte, en la Definición 10.4.3 se ve que $w(C_n) = 1$ para todo $n \in \mathbb{N}$. Como $U \in \mathcal{U}(n)$ y N es una NUI, tenemos que

$$N(T) \leq N(B_n) = N(U_n C_n) = N(C_n) \leq w(C_n) = 1 = N_0(T) \ .$$

Observar que recién al final usamos la hipótesis sobre N. ■

10.5 Ejercicios

Ejercicios que aparecen en el texto

10.5.1. Sea $A \in L(\mathbb{H})$. Probar las siguientes afirmaciones:

1. $\rho(A) \le w(A) \le \|A\|$.
2. Tomando $T = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, se tiene que $\rho(T) = 0$, $w(T) = 1/2$ y $\|T\| = 1$, por lo que las desiguadades de arriba pueden ser estrictas.
3. Si A es normal, entonces $\rho(A) = w(A) = \|A\|_{sp}$.
4. $\|\operatorname{Re} A\|_{sp} = w(\operatorname{Re} A) = \max\limits_{\|x\|=1} |\operatorname{Re} \langle Ax\,, x\rangle| \le w(A)$ y lo mismo vale para $\operatorname{Im} A$.
5. $\|A\|_{sp} \le 2 \cdot w(A)$.
6. Dado $B \in L(\mathbb{H})$ se cumple que $W(A+B) \subseteq W(A) + W(B)$.
7. Dado $\lambda \in \mathbb{C}$, se tiene que $W(A+\lambda I) = W(A) + \lambda$ y $W(\lambda \cdot A) = \lambda \cdot W(A)$.
8. Si $U \in \mathcal{U}(\mathbb{H})$, entonces $W(UAU^*) = W(A)$ y $w(UAU^*) = w(A)$.
9. $W(A)$ es compacto. ▲

10.5.2. Sea $W \subseteq \mathbb{C}^n$ un conjunto convexo compacto no vacío, y sea $v_0 \in \mathbb{C}^n \setminus W$. Entonces:

1. Existe un **único** $w_0 \in W$ tal que

$$0 < d = \|v_0 - w_0\| = d(v_0, W) := \min_{w \in W} \|v_0 - w\| \ .$$

   Para mostrar la unicidad, se suguiere asumir que $v_0 = 0$ y usar que, dados $x, y \in \mathbb{C}^n$ vale la igualdad del paralelogramo, que dice que $\|x-y\|^2 + \|x+y\|^2 = 2(\|x\|^2 + \|y\|^2)$.

2. Probar que el hiperplano $H = w_0 + \{v_0 - w_0\}^{\perp_{\mathbb{R}}}$ separa a W de v_0 en el sentido de que

$$\operatorname{Re}\big\langle v_0, v_0 - w_0\big\rangle > \operatorname{Re}\big\langle w_0, v_0 - w_0\big\rangle \ge \operatorname{Re}\big\langle w, v_0 - w_0\big\rangle \quad \text{para todo} \quad w \in W \ .$$

Nota: Lo anterior vale en espacios de Hilbert generales, y alcamza con que W sea cerrado (no hace falta compacto). Sin embargo, proponemos este caso especial porque es lo que se usa para la Observación 10.3.1.

Ejercicios nuevos

10.5.3. Sean $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n, \gamma \in \mathbb{C}$, todos de módulo uno. Son equivalentes:

1. Existen $A, B \in \mathcal{U}(n)$ tales que $\lambda(A) = \vec{\alpha}$, $\lambda(B) = \vec{\beta}$ y además $\gamma \in \sigma(BA)$.
2. Existe un $p \in \operatorname{conv}[\alpha_1, \dots, \alpha_n] \cap \operatorname{conv}\left[\gamma\overline{\beta_1}, \dots, \gamma\overline{\beta_n}\right]$.

Sugerencia: Usar el ejercicio 1.10.21

10.5.4. Sea $w'(A) = \max\limits_{w(B)=1} |\operatorname{tr} A^*B|$, la norma dual del radio numérico en $\mathcal{M}_n(\mathbb{C})$. Denotamos por $B_{w'} = \{A \in \mathcal{M}_n(\mathbb{C}) : w'(A) \leq 1\}$ a su bola unidad. Probar que

$$B_{w'} = \operatorname{conv}\left[\{x \odot x : x \in \mathbb{C}^n \ \text{ y } \ \|x\| = 1\}\right] ,$$

o sea que los puntos extremales de $B_{w'}$ son los proyectores de rk uno.

Sugerencia: Cuanto vale $\operatorname{tr}\big(B \cdot x \odot x\big)$?

Capítulo 11

Teoría de Perron-Frobenius

Repasemos algunas notaciones que usaremos sistemáticamente en los siguientes capítulos. Dados $A, B \in \mathcal{M}_{n,m}(\mathbb{R})$ y $x \in \mathbb{R}^n$, diremos que

1. $A \geqslant 0$, si $A_{ij} \geq 0$ para todo par $i \in \mathbb{I}_n$, $j \in \mathbb{I}_m$.

2. $A > 0$, si $A_{ij} > 0$ para todo par $i \in \mathbb{I}_n$, $j \in \mathbb{I}_m$.

3. Las mismas notaciones ($x \geqslant 0$ o $x > 0$) se usarán para vectores.

4. Denotaremos por

$$\mathrm{MP}_{n,m} = \{A \in \mathcal{M}_{n,m}(\mathbb{R}) : A \geqslant 0\} \quad \text{y} \quad \mathrm{MEP}_{n,m} = \{A \in \mathcal{M}_{n,m}(\mathbb{R}) : A > 0\} .$$

   para matrices cuadradas, abreviaremos MP_n y MEP_n.

5. $|A| = (|a_{ij}|)_{\substack{i \in \mathbb{I}_n \\ j \in \mathbb{I}_m}}$ y analogamente $|x| = (|x_1|, \dots, |x_n|)$.

6. $A \leqslant B$, si $B - A \in \mathrm{MP}_{n,m}$. O sea que $a_{ij} \leq b_{ij}$ para todo $i, j \in \mathbb{I}_n$. Análogamente, escribiremos $A < B$ siempre que $B - A \in \mathrm{MEP}_{n,m}$.

7. El vector $(1, 1, \dots, 1) \in \mathbb{R}^n$ será denotado por medio de $\mathbb{1}$.

Advertencia: Hay overlaps de notaciones entre lo anterior y las que solemos usar para matrices definidas positivas. Esto es lamentable, pero necesario; porque buscar otras complicaría notablemente la exposición. Las convenciones que usaremos de ahora en más serán las siguientes:

1. Mantendremos la notación $A \geq 0$ (resp. $B \geq A$) para decir que $A \in \mathcal{M}_n(\mathbb{C})^+$ (resp. $B - A \in \mathcal{M}_n(\mathbb{C})^+$). Observar que los símbolos $\geq$ y $\geqslant$ son diferentes.

2. Al escribir $A > 0$ o $B > A$ solamente aludiremos a los signos de las entradas. Para evitar confuciones, si $A \in \mathcal{M}_n(\mathbb{C})$ es definida positiva, no usaremos la notación $A > 0$, sinó $A \in \mathcal{G}l(n)^+$ (o que $B - A \in \mathcal{G}l(n)^+$).

3. $|A|$ sólo se usará para los módulos de las entradas. El viejo módulo se escribirá $(A^*A)^{1/2}$.

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

11.1 Matrices de entradas positivas

Empezamos esta sección con el objetivo final de la misma: el teorema de Perron para matrices de entradas estrictamente positivas. La idea es anunciar de entrada las propiedades principales de tales matrices y, además, dar una orientación estratégica a los numerosos resultados parciales (aunque muchos de ellos son son interesantes de por sí) que iremos probando para llegar a una demostración completa del teorema.

Teorema 11.1.1 (Teorema de Perron). *Sea $A \in \mathrm{MEP}_n$, es decir que $A > 0$. Entonces se verifican las siguientes propiedades:*

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

$$x > 0 \ , \ y > 0 \ , \ \langle x, y \rangle = 1 \ , \ Ax = x \quad \text{y} \quad A^T y = y \ .$$

Antes de demostrar el Teorema de Perron, presentaremos varios resultados generales para matrices $A \in \mathrm{MP}_n$, su radio espectral y los autovectores correspondientes. En lo que sigue de la sección, y salvo mención explícita al contrario, asumiremos que todas las matrices mencionadas con las letras A y B estarán en $\mathcal{M}_n(\mathbb{R})$, para algún $n \in \mathbb{N}$.

Proposición 11.1.2. *Sean $A, B \in \mathrm{MP}_n$ tales que $A \leqslant B$. Entonces, $\rho(A) \leq \rho(B)$.*

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

Corolario 11.1.3. *Sea $A \in \mathrm{MEP}_n$. Entonces se cunple que $\rho(A) > 0$.*

Demostración. Como $A > 0$, existe un $\varepsilon > 0$ tal que $\varepsilon I \leqslant A$. Así $\rho(A) \geq \rho(\varepsilon I) = \varepsilon > 0$. ■

Corolario 11.1.4. *Sean $A \in \mathrm{MP}_n$, $J \subseteq \mathbb{I}_n$ y $A[J] = (a_{ij})_{i,j \in J}$. Entonces $\rho(A[J]) \leq \rho(A)$.*

Demostración. Basta extender $A[J]$ a MP_n poniendo ceros en las entradas que le faltan, y aplicar la Proposición 11.1.2. ■

Observación 11.1.5. Recordemos (ver Ejercicio 3.4.2) que, dada $A \in \mathcal{M}_n(\mathbb{C})$,

$$|||A|||_\infty = \max_{\|x\|_\infty = 1} \|Ax\|_\infty = \max_{i \in \mathbb{I}_n} \|F_i(A)\|_1 \ .$$

Por otra parte, observar que si $A \in \mathrm{MP}_n$, entonces $\|F_i(A)\|_1 = \operatorname{tr} F_i(A)$. ▲

A continuación vienen tres Lemas que sirven para ubicar el radio espectral de una matriz $A \in \mathrm{MP}_n$ usando la Observación anterior:

Lema 11.1.6. *Sea $A \in \mathrm{MP}_n$. Supongamos que $\mathbb{1} \in \mathbb{R}^n$ es un autovector de A. Entonces el autovalor asociado es $|||A|||_\infty$ y además $\rho(A) = |||A|||_\infty$.*

Demostración. La desigualdad $\rho(A) \le |||A|||_\infty$ vale siempre, porque $|||\cdot|||_\infty$ es matricial y podemos aplicar la Proposición 3.4.6. Por otro lado, si $A\mathbb{1} = \lambda \mathbb{1}$, entonces

$$\|F_i(A)\|_1 = \operatorname{tr} F_i(A) = \langle F_i(A) \, , \mathbb{1} \rangle = (A\,\mathbb{1})_i = \lambda \quad \text{ para todo } i \in \mathbb{I}_n \ .$$

Por la Observación 11.1.5, podemos deducir que $\lambda = |||A|||_\infty$. Finalmente, el hecho de que $|||A|||_\infty \in \sigma(A)$ implica que $|||A|||_\infty \le \rho(A)$. ■

Lema 11.1.7. *Sea $A \in \mathrm{MP}_n$. Llamemos*

$$\alpha = \max_{i \in \mathbb{I}_n} \operatorname{tr} F_i(A) = |||A|||_\infty \quad \text{ y } \quad \beta = \min_{i \in \mathbb{I}_n} \operatorname{tr} F_i(A) \ .$$

Entonces se verifica que $\beta \le \rho(A) \le \alpha$.

Demostración. La desigualdad $\rho(A) \le \alpha$ es conocida (Proposición 3.4.6), por lo tanto sólo probaremos que $\beta \le \rho(A)$. Podemos suponer que $\beta > 0$, porque sinó todo es fácil. Definamos entonces la matriz $B \in \mathrm{MP}_n$ cuyas filas son:

$$F_i(B) = \frac{\beta}{\operatorname{tr} F_i(A)} \ F_i(A) \leqslant F_i(A) \quad \text{ para todo } i \in \mathbb{I}_n \ .$$

De este modo, $\operatorname{tr} F_i(B) = \beta$ para todo $i \ge 1$. En consecuencia, usando el Lema 11.1.6 y la Proposición 11.1.2, $\beta = \rho(B) \le \rho(A)$, ya que, por su construcción, $0 \leqslant B \leqslant A$. ■

Lema 11.1.8. *Sean $A \in \mathrm{MP}_n$ y $x > 0$. Notemos $y = Ax$. Entonces se tiene que*

$$\beta = \min_{i \in \mathbb{I}_n} \frac{y_i}{x_i} \ \le \ \rho(A) \quad \text{ y } \quad \alpha = \max_{i \in \mathbb{I}_n} \frac{y_i}{x_i} \ \ge \ \rho(A) \ .$$

Demostración. Sea $D = \operatorname{diag}(x)$. Entonces, por cuentas elementales, obtenemos que

$$D^{-1}AD = (x_i^{-1} x_j \ a_{ij})_{ij \in \mathbb{I}_n} \in \mathrm{MP}_n \ .$$

Además $\rho(D^{-1}AD) = \rho(A)$ (porque el espectro no cambia). Por otro lado,

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

Luego basta aplicar el Lema 11.1.7. ■

Teorema 11.1.9. *Sea $A \in \mathrm{MP}_n$ y fijemos un vector $x > 0$.*

1. *Dados $\alpha, \beta \in \mathbb{R}$, se tiene que*

$$\beta x \leqslant Ax \implies \beta \le \rho(A) \quad y \quad Ax \leqslant \alpha x \implies \rho(A) \le \alpha \ .$$

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

Demostración. La primera parte se deduce inmediatamente del Lema 11.1.8. Supongamos que $Ax = \lambda x$. Como $Ax \geqslant 0$, debe cumplirse que $\lambda \ge 0$, en particular $\lambda \in \mathbb{R}$. Luego se verifican las hipótesis de la primera parte con $\alpha = \beta = \lambda$. ■

Observación 11.1.10. En las condiciones del Teorema 11.1.9, también vale que

$$\text{si } A \in \mathrm{MP}_n \quad \text{y} \quad x > 0 \ , \quad \beta x < Ax \implies \beta < \rho(A) \quad \text{y} \quad Ax < \alpha x \implies \rho(A) < \alpha \ .$$

En efecto, si en los Lemas 11.1.7 y 11.1.8 se toma β estrictamente menor que los mínimos correspondientes β_0, se obtiene $\beta < \beta_0 \le \rho(A)$. Lo mismo para α. ▲

Observación 11.1.11. Sean $A \in \mathrm{MEP}_n$ y $x \in \mathbb{R}^n \setminus \{0\}$. Notar que, si $x \geqslant 0$, entonces tiene que valer que $Ax > 0$. Este hecho se usará reiteradas veces.

Corolario 11.1.12. *Sean $A \in \mathrm{MEP}_n$, $x \in \mathbb{R}^n \setminus \{0\}$ y $\lambda \in \mathbb{C}$ tales que $x \geqslant 0$ y $Ax = \lambda x$. Entonces $\lambda = \rho(A)$ y $x > 0$. Otra manera de decirlo es que si un autovector de A es no negativo, en realidad debía ser positivo y corresponder al radio espectral.*

Demostración. Por la Observación 11.1.11 sabemos que $Ax > 0$, y por ende $x > 0$. Entonces se puede aplicar el Teorema 11.1.9. ■

Proposición 11.1.13. *Sean $A \in \mathrm{MEP}_n$ y $\lambda \in \sigma(A)$ un autovalor de módulo máximo, o sea que $|\lambda| = \rho(A)$. Dado un autovector $y \in \mathbb{C}^n \setminus \{0\}$ para λ, es decir que $Ay = \lambda y$, entonces:*

$$|y| > 0 \qquad y \qquad A\,|y| \ = \ \rho(A)\,|y| \ .$$

Demostración. Llamemos $x = |y|$. Por la desigualdad triangular, se tiene que

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

Sea $z = Ax - \rho(A)x \geqslant 0$. Queremos mostrar que $z = 0$. Supongamos que eso no pasa. Entonces, por la Observación 11.1.11, tenemos que $Az > 0$. Si ahora llamamos

$$u = Ax \ , \qquad \text{entonces} \qquad Az = A(u - \rho(A)x) = Au - \rho(A)u > 0 \ .$$

Por lo tanto tenemos que $u > 0$ y $Au > \rho(A)u$. Aplicando la Observación 11.1.10, se obtiene la contradictoria desigualdad $\rho(A) > \rho(A)$. Dado que esta provino de suponer que $z \neq 0$, ahora sabemos que $z = 0$ y por ende $Ax = \rho(A)x$. Notar que, como $Ax > 0$, esto implica que $|y| = x > 0$. ■

Corolario 11.1.14. Si $A \in \mathrm{MEP}_n$, entonces

$$\rho(A) \in \sigma(A) \quad \text{ y existe un } \quad x \in \mathbb{R}^n \ \text{ tal que } x > 0 \ \text{ y } \ Ax = \rho(A)x \ . \qquad \blacksquare$$

Proposición 11.1.15. *Sean $A \in \mathrm{MEP}_n$ y $\lambda \in \sigma(A)$ tales que $|\lambda| = \rho(A)$. Si $y \in \mathbb{C}^n \setminus \{0\}$ cumple que $Ay = \lambda y$, entonces, existe $\theta \in [0, 2\pi)$ tal que $y = e^{i\theta}|y|$, por lo que $\lambda = \rho(A)$.*

Demostración. Por la Proposición 11.1.13 sabenos que $A|y| = \rho(A)|y|$. Además

$$|Ay| = |\lambda y| = \rho(A)|y| \implies A|y| = |Ay| \ .$$

Mirando las primeras coordenadas, tenemos que $\sum\limits_{i \in \mathbb{I}_n} A_{1j}\,|y_j| = \Big|\sum\limits_{i \in \mathbb{I}_n} A_{1j}\,y_j\Big|$. Luego vale la igualdad en la desigualdad triangular, y todos los y_j deben apuntar para el mismo lado. O sea que debe existir un $\theta \in [0, 2\pi)$ tal que $y_j = e^{i\theta}|y_j|$ para todo $j \in \mathbb{I}_n$. $\blacksquare$

Corolario 11.1.16. Si $A \in \mathrm{MEP}_n$, entonces $\rho(A)$ es el único autovalor de módulo máximo.

Corolario 11.1.17. *Sea $A \in \mathrm{MEP}_n$. Entonces* $\dim \ker(A - \rho(A)I) = 1$.

Demostración. Sean $x, y \in \ker(A - \rho(A)I)$. Probaremos que son linealmente dependientes. Por la Proposición 11.1.15 se puede suponer que $x > 0$ e $y > 0$. Sea $\beta = \min\limits_{i \in \mathbb{I}_n} \dfrac{x_i}{y_i}$, y definamos $z = x - \beta y$. Como cada $x_i - \beta y_i \geq x_i - \frac{x_i}{y_i}\, y_i = 0$, se tiene que $z \geq 0$.

Dado que $Az = \rho(A)z$, si sucesidese que $z \neq 0$, entonces se tendría que $z > 0$. Pero, si tomamos un $k \in \mathbb{I}_n$ tal que $\beta = \frac{x_k}{y_k}$, entonces la coordenada k-ésima de z debería ser nula. Este absurdo proviene de suponer que $z \neq 0$. Por lo tanto, $z = 0$ y $x = \beta y$. $\blacksquare$

El siguiente resultado, que describe el límite de las potencias de una matriz $A \in \mathrm{MP}_n$ que cumple ciertas hipótesis, será prontamente aplicado para probar el item 6 del Teorema de Perron. Lo enunciaremos pidiendo lo esctrictamente necesario que debe cumplir A para que la tesis pueda probarse. Esto complica su formulación, pero igual es conveniente para poder aplicarlo luego a matrices primitivas, en las que todas las hipótesis que pedimos se verifican.

Proposición 11.1.18. *Sea $A \in \mathrm{MP}_n$ con $\rho(A) = 1$. Supongamos que A cumple que:*

1. $\dim \ker(A - I) = 1$.
2. $1 \in \sigma(A)$ *es el único autovalor de módulo máximo.*
3. *Existen $x, y \in \mathbb{R}^n$ tales que*

$$x > 0 \ , \ y > 0 \ , \ \langle x, y\rangle = 1 \ , \ Ax = x \quad \text{ y } \quad A^T y = y \ .$$

Entonces, se tiene que $A^m \xrightarrow[m \to \infty]{} xy^T$.

Demostración. Llamemos $L = xy^T = (x_i y_j)_{i,j \in \mathbb{I}_n}$. Este L es, en realidad, el proyector espectral asociado al $1 \in \sigma(A)$. Esto no lo probaremos ahora, pero es útil tenerlo en cuenta para entender las propiedades de L que veremos a continuación:

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

2. $AL = LA = L$. Esto se deduce de que $Axy^T = xy^T = xy^T A$.

3. $(A - L)^m = A^m - L$, para todo $m \in \mathbb{N}$.

   Para mostrarlo, razonemos por inducción sobre m. El caso $m = 1$ es trivial. Además,

$$\begin{aligned} (A-L)^{m+1} &= (A-L)(A-L)^m = (A-L)(A^m - L) && \text{(por la HI)} \\ &= A^{m+1} - AL - LA^k + L = A^{m+1} - L - L + L \\ &= A^{m+1} - L \ . \end{aligned}$$

4. $\sigma(A - L) \setminus \{0\} \subseteq \sigma(A) - \{1\}$. En particular, se tiene que $\rho(A - L) < 1$.

   En efecto, sean $\lambda \in \mathbb{C} \setminus \{0\}$ y $z \in \mathbb{C}^n \setminus \{0\}$ tales que $(A - L)z = \lambda z$. Entonces

$$Lz = \frac{1}{\lambda} L(\lambda z) = \frac{1}{\lambda} L(L - A)z = 0 \ ,$$

   por que en 1 y 2 vimos que $L(L - A) = 0$. Luego $Az = \lambda z$ y por lo tanto $\lambda \in \sigma(A)$. Si tuviéramos que $\lambda = 1 = \rho(A)$, entonces $x \in \text{Gen}\{z\}$ (recordar que $\dim \ker(A - I) = 1$), lo que nos diría que $(A - L)x = x$. Pero $Ax = x$ y $Lx = xy^T x = x$. En consecuencia uno tendría que $(A - L)x = 0 = x$, lo que no vale.

5. Como el único $\lambda \in \sigma(A)$ con $|\lambda| = 1$ es $\lambda = 1$, se tiene que $\rho(A - L) < 1$. Entonces el Corolario 3.4.9 sirve para afirmar que $A^m - L = (A - L)^m \xrightarrow[m\to\infty]{} 0$. ■

Final de la demostración del Teorema de Perron

Recordemos el enunciado que escribimos al principio de la sección:

Teorema de Perron 11.1.1. Sea $A \in \text{MEP}_n$. Entonces se verifica que

1. $\rho(A) > 0$ y $\rho(A) \in \sigma(A)$.

2. Existe un $x \in \mathbb{R}^n$ tal que $x > 0$ y $Ax = \rho(A)x$.

3. Dado $y \in \mathbb{R}^n \setminus \{0\}$, si $y \geqslant 0$ y $Ay = \lambda y$, entonces $\lambda = \rho(A)$ e $y > 0$.

4. $\rho(A)$ es raíz simple del polinomio característico de A.

5. Si $\lambda \in \sigma(A)$ y $\lambda \neq \rho(A)$, entonces $|\lambda| < \rho(A)$.

6. Si $\rho(A) = 1$, entonces $A^m \xrightarrow[m\to\infty]{} L = xy^T$, donde $x, y \in \mathbb{R}^n$ son vectores tales que

$$x > 0 \ , \ y > 0 \ , \ \langle x, y \rangle = 1 \ , \ Ax = x \quad \text{ y } \quad A^T y = y \ .$$

Demostración. Los items 1 y 2 fueron vistos en los Corolarios 11.1.3 y 11.1.14. El item 3 se probó en el Corolario 11.1.12. El item 5 es el Corolario 11.1.16. El item 6 se deduce de la Proposición 11.1.18. Observar que ya hemos visto (aquí se usa el Corolario 11.1.17) que si $A \in \mathrm{MEP}_n$, entonces A cumple las tres condiciones que pide la Proposición 11.1.18.

Sólo falta verificar el item 4, que dice que $\rho(A)$ es raíz simple de $P_A(x) = \det(xI - A) \in \mathbb{C}[x]$. Con las notaciones del resto del libro esto significa que, si tomamos el vector $\lambda(A) \in \mathbb{C}^n$ de autovalores de A (que cuenta las multiplicidades como raices de P_A) con un orden en el que los módulos decrezcan, entonces $\lambda_1(A) = \rho(A)$ pero $|\lambda_2(A)| < \rho(A)$ (acá se usa el item 5).

Supongamos, sin perdida de generalidad, que $\rho(A) = 1$. Apliquémosle a A el Teorema 1 de Schur 1.6.1, considerando en $\lambda(A)$ el orden mencionado. Luego tendremos $U \in \mathcal{U}(n)$ y $T \in \mathcal{TS}(n)$ tales que $U^*AU = T$ y $\mathrm{d}\,(T) = \lambda(A)$. Por otra parte,

$$T^m = U^*A^mU \xrightarrow[m\to\infty]{} U^*LU = M \ .$$

Observar que todos los $T^m \in \mathcal{TS}(n)$, por lo que también $M \in \mathcal{TS}(n)$. Además, se tiene que $\operatorname{rk} M = \operatorname{rk} L = 1$. Sin embargo, como $T \in \mathcal{TS}(n)$, sabemos que

$$(T^m)_{ii} = (T_{ii})^m = \lambda_i(A)^m \quad \text{ para todo } i \in \mathbb{I}_n \text{ y todo } m \in \mathbb{N} \ .$$

Para cada $i \in \mathbb{I}_n$ tal que $\lambda_i(A) = 1$, podemos deducir que $M_{ii} = 1$. Al estar $M \in \mathcal{TS}(n)$, es fácil ver que su rk será, por lo menos, el número de unos que tenga en la diagonal. Como sabemos que $\operatorname{rk} M = 1$, deducimos que tan solo $\lambda_1(A) = 1$ y los demás tienen menor módulo (porque sus potencias deben converger a cero), como se quería demostrar. ■

Definición 11.1.19. Sea $A \in \mathrm{MEP}_n$. El único vector $x \in \mathbb{R}^n$ tal que

$$Ax = \rho(A)x \ , \quad x > 0 \quad \text{y} \quad \operatorname{tr} x = 1 \ ,$$

se llamará vector de Perron de A. ▲

11.2 Matrices de entradas no negativas

El Teorema de Perron falla en general si $A \in \mathrm{MP}_n$ pero $A \not> 0$. Por ejemplo, si

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

entonces $A^m = A$ o I, según m sea impar o par. Además, $\sigma(A) = \{1, -1\}$. En este caso el autovector asociado al 1 es positivo estricto (es $\mathbb{1}$). Pero eso no pasa si tomamos la matriz $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Es más, todas las partes del Teorema (salvo una) pueden hacerse fallar tomando matrices diagonales de bloques adecuadas (Ejercicio). La que se salva es la siguiente:

Proposición 11.2.1. *Sea $A \in \mathrm{MP}_n$. Entonces*

1. *$\rho(A) \in \sigma(A)$.*

2. *Existe $x \in \mathbb{R}^n \setminus \{0\}$ tal que $x \geqslant 0$ y $Ax = \rho(A)x$.*

Demostración. Sea $\mathbb{E} = \mathbb{1}\,\mathbb{1}^T \in \mathrm{MEP}_n$ (todas las entradas de $\mathbb{E}$ son iguales a 1). Dado $\varepsilon > 0$, tenemos que $A_\varepsilon = A + \varepsilon\,\mathbb{E} \in \mathrm{MEP}_n$. Por la Proposición 11.1.2, si $0 < \varepsilon' < \varepsilon$, entonces

$$\rho(A) \leq \rho(A_{\varepsilon'}) \leq \rho(A_\varepsilon).$$

Llamemos $x_\varepsilon > 0$ al vector de Perron de cada A_ε, normalizado para que $\operatorname{tr} x_\varepsilon = 1$. Como la bola de $\mathbb{R}^n$ es compacta, se puede tomar una sucesión decreciente $\varepsilon_m \underset{m\to\infty}{\searrow} 0$ tal que, si llamamos $A_m = A_{\varepsilon_m}$ y $x_m = x_{\varepsilon_m}$, entonces existen $M \in \mathbb{R}$ y $x \in \mathbb{R}^n$ tales que

$$\rho(A_m) \underset{m\to\infty}{\searrow} M \geq \rho(A) \quad \text{y} \quad x_m \xrightarrow[m\to\infty]{} x \geqslant 0 \ .$$

Observar que $\operatorname{tr} x = 1$, por lo que $x \neq 0$. Además, $A_m x_m = \rho(A_m) x_m \xrightarrow[m\to\infty]{} Mx$ y, como $A_m \xrightarrow[m\to\infty]{} A$, entonces $A_m x_m \xrightarrow[m\to\infty]{} Ax$. Por lo tanto deducimos que $Ax = Mx$, con $M \geq \rho(A)$. Luego $M = \rho(A)$ y $x \geqslant 0$ es un autovector. ■

Matrices primitivas

Definición 11.2.2. Sea $A \in \mathrm{MP}_n$. Diremos que A es una matriz *primitiva* si existe un $m \in \mathbb{N}$ tal que $A^m \in \mathrm{MEP}_n$. ▲

Las matrices primitivas son casi tan buenas como las de MEP_n. Veamos que cumplen el Teorema de Perron tutti, que enunciamos por tercera vez.

Teorema 11.2.3. *Sea $A \in \mathrm{MP}_n$ una matriz* **primitiva**. *Entonces valen:*

1. *$\rho(A) > 0$ y $\rho(A) \in \sigma(A)$.*

2. *Existe un $x \in \mathbb{R}^n$ tal que $x > 0$ y $Ax = \rho(A)x$.*

3. *Dado $y \in \mathbb{R}^n \setminus \{0\}$, si $y \geqslant 0$ y $Ay = \lambda y$, entonces $\lambda = \rho(A)$ e $y > 0$.*

4. *$\rho(A)$ es raíz simple del polinomio característico de A.*

5. *Si $\lambda \in \sigma(A)$ y $\lambda \neq \rho(A)$, entonces, $|\lambda| < \rho(A)$.*

6. *Si $\rho(A) = 1$, entonces $A^m \xrightarrow[m\to\infty]{} L = xy^T$, donde $x, y \in \mathbb{R}^n$ son vectores tales que*

$$x > 0 \ , \ y > 0 \ , \ \langle x, y \rangle = 1 \ , \ Ax = x \quad y \quad A^T y = y \ .$$

Demostración. Sea $m \in \mathbb{N}$ tal que $A^m > 0$. Por el Corolario 1.7.2,

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

Por el Teorema 11.1.1 aplicado a A^m, concluimos que $\rho(A) = \rho(A^m)^{1/m} > 0$. Sea $\lambda \in \sigma(A)$ tal que $|\lambda| = \rho(A)$ y sea $y \in \mathbb{C}^n \setminus \{0\}$ tal que $Ay = \lambda y$. Entonces

$$A^m y = \lambda^m y \quad \text{y} \quad |\lambda|^m = \rho(A^m) \implies \lambda^m = \rho(A^m) \quad \text{y} \quad A^m y = \rho(A^m) y \ .$$

Por el Teorema 11.1.1 aplicado a A^m, podemos deducir que algún $x \in \operatorname{Gen}\{y\}$ cumple que $x > 0$, y por ello $\lambda = \rho(A)$ y $Ax = \rho(A)x$.

Además, cada $\lambda^m \in \sigma(A^m)$ posee una multiplicidad en el polinomio característico de A^m mayor o igual que la de λ en el de A (esto se ve fácil triangulando con el Teorema 1.6.1). Por lo tanto $\rho(A)$ posee multiplicidad algebráica uno como autavalor de A. Razonamientos similares permiten concluir que $\rho(A)$ es el único autovalor de módulo máximo (item 5), y también la condición 3. Finalmente, con los items anteriores ya demostrados, estamos en condiciones de asegurar que A cumple las hipótesis de la Proposición 11.1.18, lo que prueba el item 6. ■

Observación 11.2.4. Dada una matriz $A \in \mathrm{MP}_n$, para saber si es primitiva hace falta calcular muchas potencias A^m hasta que caiga en MEP_n. Obviamente hace falta un teorema que diga hasta donde es necesario probar. Algo del tipo: Dado $n \in \mathbb{N}$, existe un $M(n) \in \mathbb{N}$ (que uno debería poder calcular) tal que toda $A \in \mathrm{MP}_n$ que sea primitiva debe cumpir que $A^m > 0$ para algún $m \le M(n)$. Esta teoría existe, y se calculan los $M(n)$ óptimos. Pero las cuentas son muy complicadas y no las desarrollaremos aquí.

El lector interesado puede buscar data al respecto en el libro de Horn-Johnson [7]. Sin embargo, con una hipótesis razonable (si $A \in \mathrm{MP}_n$ cumple que $\mathrm{d}(A) > 0$), sale mucho más facilmente que la constante $M(n) = n - 1$ sirve. Obsrvar que en tal caso, una vez que $A^m > 0$, eso sigue pasando para las potencias mayores (lo que no es cierto para todas las primitivas). Esperen algunos renglones y verán. ▲

Matrices irreducibles

Definición 11.2.5. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Decimos que:

1. A es **reducible** si existe $P \in \mathcal{U}_{\mathcal{P}}(n)$, una matriz de permutación, tal que

$$PAP^{-1} = \begin{array}{cc} \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} & \begin{array}{c} k \\ n-k \end{array} \end{array} ,$$

   donde $1 \le k \le n-1$ y $B \in \mathcal{M}_k(\mathbb{R})$. Otra manera de decirlo es que existe un $J \subseteq \mathbb{I}_n$ tal que $1 \le |J| < n$ (o sea que es propio) que cumpla que

$$A\big(\operatorname{Gen}\{e_j : j \in J\}\big) \subseteq \operatorname{Gen}\{e_j : j \in J\} \ . \tag{11.1}$$

   Se dice que A es **irreducible** si no es reducible.

2. Denotemos momentáneamente por $V_n = \{(p,q) \in \mathbb{I}_n^2 : p \neq q\}$, al conjunto de pares de índices **distintos** en $\mathbb{I}_n$.

3. Decimos que un par $(p,q) \in V_n$ se **conecta** por A (o que A conecta p con q), si existen $p = i_0$, $i_1, \ldots, i_m = q$ en $\mathbb{I}_n$ tales que $a_{i_{k-1}i_k} \neq 0$ para todo $k \in \mathbb{I}_m$.

   Observar que se puede suponer que todos los i_k son distintos entre sí, porque si hubiera repeticiones, una parte de la sucesión se podría borrar (los intermedios entre los dos repetidos), quedando otra sucesión más corta que seguiría conectando a p con q. Por lo tanto, puede suponerse que $m \leq n-1$.

4. A es **fuertemente conexa** (FC) si todo par $(p,q) \in V_n$ se conecta por A. ▲

Lema 11.2.6. *Sea $A \in \mathrm{MP}_n$. Dado un par $(p,q) \in V_n$, son equivalentes:*

1. *El par (p,q) se conecta por A.*

2. *Existe $1 \leq m \leq n-1$ tal que la entrada $(A^m)_{p\,q} > 0$.*

Demostración. Basta notar que, como mostraría una inducción adecuada,

$$(A^m)_{p\,q} = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_{m-1}=1}^{n} a_{p\ i_1} \cdot \prod_{k \in \mathbb{I}_{m-2}} a_{i_k\, i_{k+1}} \cdot a_{i_{m-1}\, q} \ ,$$

y que todos estos términos son no negativos. En caso de que alguno de esos sumandos no se anule, les sacamos aquellos términos que vivan en la diagonal de A, y nos queda una sucesión que conecta p con q. Recordar que si A conectaba a (p,q), entonces existe alguna sucesión de no más de n índices que los conecta. ■

Ejemplo 11.2.7. Ahorita vamos a ver que irreducible es lo mismo que FC (se lo enunciará pra matrices de MP_n , pero obviamente eso es irrelevante). Veamos una serie de ejemplos donde se ve independientemente que pasa lo mismo: Sea $A \in \mathcal{M}_n(\mathbb{C})$ tal que $F_i(A) = 0$, para algún $i \in \mathbb{I}_n$. Tomemos cualquier $\sigma \in \mathbf{S}_n$ tal que $\sigma(i) = n$, y $P_\sigma \in \mathcal{U}_\mathcal{P}(n)$ su matriz asociada. Por la Eq. (4.3), se tiene que $F_n(P_\sigma A) = 0$. Como multiplicar del otro lado permuta sólo sus ceros, también vale que $F_n(P_\sigma A P_\sigma^{-1}) = 0$. O sea que A es reducible.

Veámoslo desde otro punto de vista: Si $F_i(A) = 0$, entonces a i no se lo puede conectar con ningún otro $j \in \mathbb{I}_n \setminus \{i\}$, porque todos los a_{ik} son nulos. Luego A no es FC. Ya que estamos dejamos un pequeño ejercicio: $A \in \mathcal{M}_n(\mathbb{C})$ es reducible si y sólo si A^T lo es. Por lo tanto, lo anterior vale también para columnas nulas. ▲

Proposición 11.2.8. *Sea $A \in \mathrm{MP}_n$. Entonces son equivalentes:*

1. *A es irreducible.*

2. *A es FC.*

3. *$(I + A)^{n-1} > 0$.*

4. $I + A$ es primitiva.

En particular se tiene que, si A es primitiva, entonces es irreducible y FC.

Demostración. $2 \leftrightarrow 3$: Por el Lema anterior, es claro que 3 implica 2, porque conectar por A es lo mismo que conectar por $I+A$, dado que los elementos de la diagonal no se usan para las conexiones. Recíprocamente, por el teorema del binomio de Newton, se tiene que $(I+A)^{n-1}$ es combinación lineal, a coeficientes positivos, de las potencias A^k, $0 \le k \le n-1$. Luego, si A es FC, el Lema 11.2.6 asegura que todas las entradas de $(I+A)^{n-1}$ (afuera de la diagonal) deben ser estrictamente positivas. Además, $(I+A)^{n-1} \geqslant I^{n-1} = I$.

$1 \to 2$: Si A no es FC, existe un par $(p,q) \in V_n$ que no se conecta por A. Sean

$$J_1 = \{ i \in \mathbb{I}_n \setminus \{p\} : A \text{ conecta al par } (p,i) \ \} \cup \{p\} \quad \text{ y } \quad J_2 = \mathbb{I}_n \setminus J_1 \ .$$

Entonces $p \in J_1$ y $q \in J_2$, por lo que ambos son no vacíos. En particular, $a_{ij} = 0$ si $i \in J_1$ y $j \in J_2$ (sino el par (p,j) sería conectado por A, pasando por i). Si reordenamos $\mathbb{I}_n$ poniendo primero a J_2 y luego a J_1, encontraremos una matriz $P \in \mathcal{U}_{\mathcal{P}}(n)$ de permutación tal que

$$PAP^{-1} = \begin{array}{cc} \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} & \begin{array}{c} J_2 \\ J_1 \end{array} \end{array} .$$

$3 \to 4$: Obvio.

$4 \to 1$: Si A es reducible, sea $P \in \mathcal{U}_{\mathcal{P}}(n)$ tal que $PAP^{-1} = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$. Entonces

$$P(I+A)^m P^{-1} = (I + PAP^{-1})^m = \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \notin \text{MEP}_n \qquad \text{para todo} \quad m \in \mathbb{N} \ .$$

Por lo tanto ningúna potencia $(I+A)^m \in \text{MEP}_n$, o sea que $I+A$ no es primitiva. ■

Teorema 11.2.9 (Teorema de Perron-Frobenius)**.** *Sea $A \in \text{MP}_n$, y asumamos que A es* **irreducible**. *Entonces se verifica que*

1. *$\rho(A) > 0$ y $\rho(A) \in \sigma(A)$.*
2. *Existe $x > 0$ tal que $Ax = \rho(A)x$.*
3. *$\rho(A)$ es raíz simple del polinomio característico de A.*

Demostración. Como A es ireducible, el Ejemplo 11.2.7 nos dice que A no puede tener ninguna fila nula. Usando el Lema 11.1.7, tenemos que

$$\rho(A) \ \geq \ \beta \ = \ \min_{i \in \mathbb{I}_n} \ \operatorname{tr} F_i(A) \ > \ 0 \ .$$

Por otra parte, por la Proposición 11.2.1, $\rho(A) \in \sigma(A)$ (para esto alcanza con el hecho de que $A \in \text{MP}_n$). Además, $\sigma(I+A) = 1 + \sigma(A)$. Más aún, por el Teorema 1 de Schur 1.6.1, se tiene que $\lambda(I+A) = \mathbb{1} + \lambda(A)$ (contando multiplicidades, y en algún orden). Por lo tanto

$\rho(I + A) = 1 + \rho(A)$ (porque el máximo está a la derecha y no en la tercera posición). Como $I + A$ es primitiva, si denotamos por x al vector de Perron de $I + A$, entonces tenemos que

$$x > 0 \quad \text{y} \quad Ax = (I + A - I)\, x = (1 + \rho(A)\,)\, x - x = \rho(A)\, x \ .$$

Por último, la igualdad $\lambda(I+A) = \mathbb{1}+\lambda(A)$ dice que cada $\lambda_i(A) = \rho(A)$ produce un $\lambda_i(I+A) = 1 + \rho(A)$. Como de éstos hay uno solo, sale el item 3. ■

A continuación presentamos dos resultados sobre matrices irreducibles de MP_n que son muy útiles, pero que quedaron medio aislados:

Corolario 11.2.10. *Sean $A \in \mathrm{MP}_n$ irreducible y $x \in \mathbb{R}^n \setminus \{0\}$. Si se tiene que*

$$x \geqslant 0 \quad y \quad Ax \geqslant \rho(A)x \implies x > 0 \quad y \quad Ax = \rho(A)x \ .$$

Demostración. Como A es irreducible, también A^T lo es (¿porque?). Por el Teorema de Perron-Frobenius existe un vector $y > 0$ tal que $A^T y = \rho(A)y$, o sea que $y^T A = \rho(A)y^T$. Por otra parte, sabemos que $Ax - \rho(A)x \geqslant 0$. Si sucediera que $Ax - \rho(A)x \neq 0$, entonces

$$y^T > 0 \implies 0 < y^T(Ax - \rho(A)x) = y^T Ax - \rho(A)y^T x = \rho(A)y^T x - \rho(A)y^T x = 0.$$

Esta contradicción nos convence de que $Ax = \rho(A)x$. Usando lo anterior, el hecho de que $x > 0$ puede deducirse ahora del Teorema de Perron-Frobenius. ■

Proposición 11.2.11. *Sean $A, B \in \mathrm{MP}_n$ tales que A es irreducible y $B \leqslant A$. Si además asumimos que $B \neq A$, entonces $\rho(B) < \rho(A)$.*

Demostración. La Proposición 11.2.1 nos dice que existe un $x \in \mathbb{R}^n \setminus \{0\}$ tal que $x \geqslant 0$ y $Bx = \rho(B)x$. Supongamos que $\rho(B) = \rho(A)$. En tal caso, por el Corolario 11.2.10,

$$x \geqslant 0 \quad \text{y} \quad A \geqslant B \implies Ax \geqslant Bx = \rho(B)x = \rho(A)x \implies Ax = \rho(A)x \quad \text{y} \quad x > 0 \ .$$

Por lo tanto $Ax = \rho(A)x = \rho(B)x = Bx$, o sea que $(A - B)x = 0$. Sin embargo, esto es imposible porque $A \neq B$, $A - B \geqslant 0$ y $x > 0$. La contradicción provino de suponer que $\rho(B) = \rho(A)$. Luego $\rho(B) < \rho(A)$. ■

Observación 11.2.12. Sea $A \in \mathrm{MP}_n$ una matriz irreducible. En este caso, $\rho(A)$ no es, necesariamente, el único autovector de módulo máximo. En efecto, tomando

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} , \qquad \text{se tiene que } A \text{ es irreducible porque } I + A > 0, \text{ pero } \sigma(A) = \{1, -1\} \ .$$

En general, puede verse que los otros autovalores de módulo máximo en el $\sigma(A)$ son los siguientes: $\omega_1\, \rho(A)\,, \ldots, \omega_{k-1}\, \rho(A)$, donde los ω_i son las raíces k-ésimas de la unidad, para cierto $k \leq n$. En el caso anterior, k era 2. El lector interesado puede buscar más información al respecto en el libro de A. Benedek y R. Panzone [1], el de Horn y Johnson [7] o en los siguientes ejercicios. ▲

Ejercicio 11.2.13. Sea $A \in \mathrm{MP}_n$. Probar que:

1. A es primitiva si y sólo si A es irreducible y $\rho(A)$ es el único autovector de módulo máximo de A.
2. Si A es irreducible y semidefinida positiva (o sea que A es irreducible, $A \geq 0$ y $A \geqslant 0$), entonces A es primitiva. ▲

Ejercicio 11.2.14. Sean $B \in \mathcal{M}_n(\mathbb{C})$ y $A \in \mathrm{MP}_n$ una matriz irreducible.

1. Supongamos que $|B| \leqslant A$, $\rho(A) = \rho(B)$ y $\lambda = e^{i\phi}\rho(B)$ es un autovalor de B de módulo máximo. Entonces, existen números reales $\theta_1, \dots, \theta_n$ tales que
$$B = e^{i\phi} D \ A \ D^{-1}$$
donde $D = \operatorname{diag}\left(e^{i\theta_1}, \dots, e^{i\theta_n}\right)$.
2. Supongamos que $\rho(A) = 1$ y sea $\mathcal{S} = \{\lambda_1 , \dots , \lambda_k\} = \{\lambda \in \sigma(A) : |\lambda| = 1\}$.
   (a) Pongamos que cada $\lambda_j = e^{i\phi_j}\rho(A)$. Probar que $\mu \in \sigma\,(A) \iff e^{-i\phi_j}\mu \in \sigma\,(A)$.
   (b) Concluir a partir del item anterior que $\mathcal{S}$ es un grupo abeliano.
   (c) Probar que $\mathcal{S} = G_k = \{e^{\frac{2\pi i p}{k}} \ : \ p \in \mathbb{I}_k\}$ y que cada $\lambda_p \in \mathcal{S}$ tiene multiplicidad algebraica igual a uno.
   (d) Mostrar que si A es no singular y n es primo, entonces, $\rho(A)$ es el único autovalor de módulo máximo, o bien A posee n autovalores distintos. ▲

Ejemplo 11.2.15. Sea $J_n \in \mathcal{M}_n(\mathbb{R})$ el bloque de Jordan de tamaño $n \times n$ (con $n \geq 2$). Es decir que $J_n e_1 = 0$ y $J_n e_k = e_{k-1}$, para $2 \leq k \leq n$. Llamemos

$$A = J + J^T = \begin{bmatrix} 0 & 1 & 0 & 0 & .. & .. & .. & 0 \\ 1 & 0 & 1 & 0 & .. & .. & .. & 0 \\ 0 & 1 & 0 & 1 & .. & .. & .. & 0 \\ 0 & 0 & 1 & \ddots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 1 & 0 & 0 \\ 0 & .. & .. & .. & 1 & 0 & 1 & 0 \\ 0 & .. & .. & .. & 0 & 1 & 0 & 1 \\ 0 & .. & .. & .. & 0 & 0 & 1 & 0 \end{bmatrix} \in \mathcal{H}(n) \ ,$$

que actúa en $\mathbb{R}^n$ por $Ax = (x_2 \, , \, x_1 + x_3 \, , \, x_2 + x_4 \, , \, \dots \, , \, x_{n-2} + x_n \, , \, x_{n-1})$, $x \in \mathbb{R}^n$. No es difícil verificar que A es irreducible, ya sea mostrando que $(I + A)^{n-1} > 0$, o viendo que satisface la definición de ser FC (con la diagonal de arriba si $p < q$ y con la de abajo si $q < p$). También puede probarse que $A \in \mathcal{G}l\,(n)$ si y sólo si n es par. Esta matriz es muy famosa y es, tal ves, la primera matriz a la que se le calcularon todos los autovalores y autovectores. Esto lo hizo Lagrange en 1759, para resolver un sistema de ecuaciones diferenciales ordinarias asociado al problema de la cuerda que vibra. Sus autovalores son, en orden decreciente,

$$\mu_k(A) = 2\cos\frac{k\pi}{n+1} \ , \quad 1 \leq k \leq n \ ,$$

por lo que $\|A\| = \rho(A) = 2\cos\frac{\pi}{n+1}$. Notar que $\mu_n(A) = -\mu_1(A)$, luego A no es primitiva. Además, si $n+1 = 2k$ (es decir, si n es impar), entonces $\mu_k(A) = 2\cos\pi/2 = 0$, lo que prueba lo antedicho. Los autovectores asociados son, respectivamente,

$$x_k = \left(\operatorname{sen}\frac{k\pi}{n+1} \ , \ \operatorname{sen}\frac{2k\pi}{n+1} \ , \ \dots \ , \ \operatorname{sen}\frac{nk\pi}{n+1} \right) , \quad 1 \le k \le n \ .$$

Notar que el único con entradas positivas es x_1 , porque $\frac{n\pi}{n+1}$ no llegó aún a π. La verificación de lo anterior es tediosa pero elemental. Se basa en las fórmulas del seno y coseno de sumas y restas, y en que $\operatorname{sen}(\pi - t) = \operatorname{sen} t$ y $\cos(\pi - t) = -\cos t$.

A es el prototipo de matriz tridiagonal o de Jacobi. En realidad cumple que $I+A$ es totalmente positiva, lo que justifica (más bien digamos que sugirió) las propiedades de sus autovalores y autovectores, como se verá en el Capítulo 13.

Veamos que $Ax_1 = \mu_1(A)x_1$, lo que nos dirá que $\rho(A) = 2\cos\frac{\pi}{n+1}$ y que x_1 es el vector de Perron-Frobenius de A. En efecto, se tienen dos casos: para las entradas 1 y n:

$$\begin{aligned} A(x_1)_1 &= \operatorname{sen}\tfrac{2\pi}{n+1} = 2\cos\tfrac{\pi}{n+1} \ \operatorname{sen}\tfrac{\pi}{n+1} \quad \text{y} \\ A(x_1)_n &= \operatorname{sen}\tfrac{(n-1)\pi}{n+1} = \cos\tfrac{\pi}{n+1} \ \operatorname{sen}\tfrac{n\pi}{n+1} \ - \ \cos\tfrac{n\pi}{n+1} \ \operatorname{sen}\tfrac{\pi}{n+1} \\ &= 2\cos\tfrac{\pi}{n+1} \ \operatorname{sen}\tfrac{n\pi}{n+1} \ . \end{aligned}$$

Para las entradas $2 \le k \le n-1$ se tiene que $A(x_1)_k = (x_1)_{k+1} + (x_1)_{k-1}$. Pero

$$\begin{aligned} (x_1)_{k+1} = \operatorname{sen}\tfrac{(k+1)\pi}{n+1} &= \cos\tfrac{\pi}{n+1} \ \operatorname{sen}\tfrac{k\pi}{n+1} + \cos\tfrac{k\pi}{n+1} \ \operatorname{sen}\tfrac{\pi}{n+1} \quad \text{y} \\ (x_1)_{k-1} = \operatorname{sen}\tfrac{(k-1)\pi}{n+1} &= \cos\tfrac{\pi}{n+1} \ \operatorname{sen}\tfrac{k\pi}{n+1} - \cos\tfrac{k\pi}{n+1} \ \operatorname{sen}\tfrac{\pi}{n+1} \ . \end{aligned}$$

Sumando se obtiene la fórmula buscada. Los números $c_m = 2\cos\frac{\pi}{m}$ para $m \ge 3$, que aparececn como normas de las matrices anteriores, son muy importantes en varias ramas de la matemática. Por ejemplo, aparecen en la teoría del índice de V. Jones. Tienen la siguiente particularidad: Sea $\mathcal{N}(\mathbb{Z}) \subseteq \mathbb{R}$ el conjuntos de normas espectrales de matrices de cualquier tamaño (incluso rectangulares) con entradas en $\mathbb{Z}$. Entonces

$$\mathcal{N}(\mathbb{Z}) \cap (0,2) = \{ \ 2 \ \cos\frac{\pi}{m} \ : m \ge 3\}.$$

Notar que realizamos todos estos valores con las matrices cuadradas anteriores. Sin embargo, se los puede realizar con matrices más pequeñas. En efecto, si $n = 2k$, sea $B \in \mathcal{M}_k(\mathbb{Z})$ dada por $B = I_k + J_k$. Entonces la matriz

$$\widehat{B} = \begin{bmatrix} 0 & B \\ B^T & 0 \end{bmatrix} \in \mathcal{H}(n)$$

difiere de la matriz A del principio sólo en una reordenación de la base canónica (poniendo los pares primero y los impares al final). Es decir que existe una matriz de permutación $P \in \mathcal{U}_{\mathcal{P}}(n)$ tal que $PAP^{-1} = \widehat{B}$. Por lo tanto

$$\|B\| = s_1(B) = \mu_1(\widehat{B}) = \|\widehat{B}\| = \|A\| = c_{n+1} \quad .$$

Por eso era que $-\mu_{n-j+1}(A) = \mu_j(A) = s_j(B)$, para todo $j \in \mathbb{I}_k$ (ver Proposición 3.7.5). Algo similar puede hacecrse si $n = 2k+1$, tomando $B' = (B, e_k) \in \mathcal{M}_{k,\,k+1}(\mathbb{Z})$. ▲

11.3 Ejercicios

Ejercicios que aparecen en el texto

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

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

entonces $A^m = A$ o I, según m sea impar o par. Además, $\sigma(A) = \{1, -1\}$. En este caso el autovector asociado al 1 es positivo estricto (es $\mathbb{1}$). Pero eso no pasa para la matriz $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Es más, todas las partes del Teorema (salvo una) pueden hacerse fallar tomando matrices diagonales de bloques adecuadas

11.3.2. Sea $A \in \mathrm{MP}_n$. Probar que:

1. A es primitiva si y sólo si A es irreducible y $\rho(A)$ es el único autovector de módulo máximo de A.

2. Si A es irreducible y semidefinida positiva (o sea que A es irreducible, $A \geq 0$ y $A \geqslant 0$), entonces A es primitiva. ▲

11.3.3. Sean $B \in \mathcal{M}_n(\mathbb{C})$ y $A \in \mathrm{MP}_n$ una matriz irreducible.

1. Supongamos que $|B| \leqslant A$, $\rho(A) = \rho(B)$ y $\lambda = e^{i\phi}\rho(B)$ es un autovalor de B de módulo máximo. Entonces, existen números reales $\theta_1, \ldots, \theta_n$ tales que
$$B = e^{i\phi} D\ A\ D^{-1}$$
donde $D = \operatorname{diag}\left(e^{i\theta_1}, \ldots, e^{i\theta_n}\right)$.

2. Supongamos que $\rho(A) = 1$ y sea $\mathcal{S} = \{\lambda_1, \ldots, \lambda_k\} = \{\lambda \in \sigma(A) : |\lambda| = 1\}$.
   (a) Pongamos que cada $\lambda_j = e^{i\phi_j}\rho(A)$. Probar que $\mu \in \sigma(A) \iff e^{-i\phi_j}\mu \in \sigma(A)$.
   (b) Concluir a partir del item anterior que $\mathcal{S}$ es un grupo abeliano.
   (c) Probar que $\mathcal{S} = G_k = \{e^{\frac{2\pi i p}{k}} : p \in \mathbb{I}_k\}$ y que cada $\lambda_p \in \mathcal{S}$ tiene multiplicidad algebraica igual a uno.
   (d) Mostrar que si A es no singular y n es primo, entonces, $\rho(A)$ es el único autovalor de módulo máximo, o bien A posee n autovalores distintos.

11.3.4. Completar las pruebas de lo enunciado en la Observación 11.2.15.

Ejercicios nuevos

11.3.5. Sean $A \in \mathcal{M}_n(\mathbb{R})$ y $x \in \mathbb{R}^n$. Probar que:

1. Si $A > 0$ y $x \geqslant 0$, pero $x \neq 0$, entonces $Ax > 0$.
2. Si $A \geqslant 0$ y $x > 0$, entonces $Ax = 0 \iff A = 0$. Más aún, $(Ax)_k = 0 \iff F_k(A) = 0$.
3. Si $A > 0$ y es inversible, entonces $A^{-1} \notin \mathrm{MP}_n$.
4. Si $A \geqslant 0$ y es inversible, entonces $A^{-1} \geqslant 0 \iff A$ tiene exactamente una entrada no nula por columna.

11.3.6. Si $A \in \mathrm{MP}_n$ posee un autovalor positivo, probar que A es similar a una matriz de MP_n tal que la traza se sus filas es constante. ¿Cual es esa constante?.

11.3.7. Sea $A \in \mathrm{MP}_n$. Demostrar que

$$\rho(A) = \max_{x>0} \min_{i\in\mathbb{I}_n} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij} x_j = \min_{x>0} \max_{i\in\mathbb{I}_n} \frac{1}{x_i} \sum_{j=1}^{n} a_{ij} x_j \quad .$$

11.3.8. Sea $A \in \mathrm{MP}_n$. Probar que

$$\lim_{m\to\infty} \Big(\operatorname{tr} F_i(A^m)\Big)^{1/m} = \rho(A) \qquad \text{para cualquier } i \in \mathbb{I}_n \ .$$

11.3.9. Sea $A \in \mathcal{M}_n(\mathbb{R})$ tal que las entradas fuera de la diagonal de A son no negativas[1]. Mostrar que A posee un autovalor real $r(A)$ tal que $r(A) \geq \operatorname{Re}(\lambda)$ para todo $\lambda \in \sigma(A)$.

11.3.10. Sean $\sigma \in \mathbf{S}_n$ y $P_\sigma \in \mathcal{U}_\mathcal{P}(n) \subseteq \mathrm{MP}_n$ su matriz asociada. Decir que debe cumplir σ para que P_σ sea irreducible. Se recomienda mirar la Eq. (11.1). De paso, calcular $\sigma(P_\sigma)$.

11.3.11. Probar que si A es una matriz doble estocástica reducible, entonces existe una permutación $P_\sigma \in \mathcal{U}_\mathcal{P}(n)$ tal que

$$P_\sigma A P_\sigma^{-1} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}.$$

[1] estas matrices se conocen con el nombre de esencialmente no-negativas.

Capítulo 12

Complementos de Schur y determinantes

12.1 Notaciones y definiciones

Recordemos las notaciones asociadas a submatrices vistas en Capítulos anteriores:

1. Sea $n \in \mathbb{N}$ y $k \in \mathbb{I}_n$. Notamos por $Q_{k,n}$ al conjunto de sucesiones estrictamente crecientes de k enteros elegidos en $\mathbb{I}_n$:
$$Q_{k,n} = \left\{ \alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k) \in \mathbb{I}_n^k \ : \ 1 \le \alpha_1 < \alpha_2 < \cdots < \alpha_k \le n \ \right\}.$$
Otra manera de verlo es $Q_{k,n} = \left\{ J \subseteq \mathbb{I}_n : |J| = k \right\}$, si pensamos a los conjuntos J ordenados en forma creciente. Luego $|Q_{k,n}| = \binom{n}{k}$.

2. Dado $\alpha \in Q_{k,n}$, denotaremos por $\alpha' = \mathbb{I}_n \setminus \alpha \in Q_{n-k\,,n}$ a su complemento (ordenado convenientemente).

3. Sean $A \in \mathcal{M}_{n,m}(\mathbb{C})$, $\alpha \in Q_{k,n}$ y $\beta \in Q_{r,m}$. Entonces denotaremos por $A[\alpha|\beta]$ a la submatriz de $k \times r$ de A dada por
$$A[\alpha|\beta] = \left(A_{\alpha_i \beta_j} \right)_{(i,j) \in \mathbb{I}_k \times \mathbb{I}_r} \in \mathcal{M}_{k,r}(\mathbb{C}) \ .$$
Llamaremos $A(\alpha|\beta) = A[\alpha'|\beta'] \in \mathcal{M}_{n-k\,,\,m-r}(\mathbb{C})$. Análogamente se definen
$$A[\alpha|\beta) = A[\alpha|\beta'] \in \mathcal{M}_{k\,,\,m-r}(\mathbb{C}) \quad \text{y} \quad A(\alpha|\beta] = A[\alpha'|\beta] \in \mathcal{M}_{n-k\,,\,r}(\mathbb{C}) \ .$$

4. Cuando $\alpha = \beta$, $A[\alpha|\alpha]$ se abreviará como $A[\alpha]$ y $A(\alpha|\alpha) = A(\alpha)$. Si $\alpha = \mathbb{I}_n$ (resp. $\beta = \mathbb{I}_m$), notaremos $A[\alpha|\beta] = A[-|\beta]$ (resp. $A[\alpha|\beta] = A[\alpha|-]$).

5. Dadas $A \in \mathcal{M}_{n,r}(\mathbb{C})$ y $B \in \mathcal{M}_{r,m}(\mathbb{C})$, sea $k \le \min\{n, r, m\}$. Luego, para cada par $\alpha \in Q_{k,n}$, $\beta \in Q_{k,m}$ se tiene la fórmula de Cauchy Binnet para AB:
$$\det{(AB)}[\alpha|\beta] = \sum_{\omega \in Q_{k,r}} \det A[\alpha|\omega] \det B[\omega|\beta] \ . \tag{12.1}$$

6. Dada $\alpha \in Q_{k,n}$, usaremos la abreviación:

$$\mathbf{e}_\alpha^\wedge \;=\; \mathbf{e}_\alpha^{(n)\,\wedge} := e_{\alpha_1}^{(n)} \wedge e_{\alpha_2}^{(n)} \wedge \cdots \wedge e_{\alpha_k}^{(n)} \in \Lambda^k \mathbb{H}_n \ .$$

Luego, por la multilinealidad de la función $\mathbb{H}_n^k \ni (x_1, \ldots, x_k) \mapsto x_1 \wedge \cdots \wedge x_k$ (ver 7.2.1, item 3 y Definición 7.3.4), y por la Eq. (7.18), el conjunto

$$\mathcal{E}_{k,n}^\wedge = \{\sqrt{k!}\ \mathbf{e}_\alpha^\wedge : \alpha \in Q_{k,n}\}$$

es una base ortonormal de $\Lambda^k \mathbb{H}_n$, y se la llama base canónica. Por lo tanto, tenemos que $\dim \Lambda^k \mathbb{H}_n = |Q_{k,n}| = \binom{n}{k}$.

El complemento de Schur

Definición 12.1.1. Sea $A \in \mathcal{M}_n(\mathbb{C})$, $k \in \mathbb{I}_n$ y $\alpha, \beta \in Q_{k,n}$.

1. Supongamos que $A[\alpha|\beta]$ es inversible. En tal caso definimos el *complemento de Schur* de $A[\alpha|\beta]$ en A, como la matriz

$$A/[\alpha|\beta] = A(\alpha|\beta) - A(\alpha|\beta] \cdot A[\alpha|\beta]^{-1} \cdot A[\alpha|\beta) \in \mathcal{M}_{n-k}(\mathbb{C}) \ , \tag{12.2}$$

indexada por α' y β'.

2. Si $\alpha = \beta$, escribiremos A/α en lugar de $A/[\alpha|\alpha]$. ▲

Observación 12.1.2. Sean $A \in \mathcal{M}_n(\mathbb{C})^+$ y $\alpha \in Q_{k,n}$. Si $A[\alpha] \in \mathcal{G}l(k)^+$ y consideramos el subespacio $\mathcal{S} = \operatorname{Gen}\{e_j : j \notin \alpha\}$, entonces el Corolario 3.8.9 dice que

$$\begin{bmatrix} A/\alpha & 0 \\ 0 & 0 \end{bmatrix} \begin{array}{l} \mathcal{S} \\ \mathcal{S}^\perp \end{array} = \begin{bmatrix} A(\alpha) - A(\alpha|\alpha]\ A[\alpha]^{-1}\ A(\alpha|\alpha]^* & 0 \\ 0 & 0 \end{bmatrix} \begin{array}{l} \mathcal{S} \\ \mathcal{S}^\perp \end{array} = \Sigma\,(A, \mathcal{S}) \ .$$ ▲

Definición 12.1.3. Sea $\alpha \in Q_{k,n}$.

1. Llamaremos $\operatorname{tr} \alpha = \sum_{i=1}^{k} \alpha_i$.

2. Llamaremos $\operatorname{sgn}(\alpha)$ al signo de la permutación $\pi_\alpha \in \mathbf{S}_n$ dada por $\pi_\alpha(\alpha_i) = i$ para $i \in \mathbb{I}_k$, y $\pi_\alpha(\alpha'_j) = k + j$ para $j \in \mathbb{I}_{n-k}$. Es decir que π_α pone a los buenos al principio y a los malos al final, preservando sus órdenes. Por lo tanto,

$$\operatorname{sgn}(\alpha) = \prod_{i=1}^{k} (-1)^{\alpha_i - i} = (-1)^r \ , \quad \text{con} \quad r = \operatorname{tr} \alpha \ - \ \frac{k(k+1)}{2} \ . \tag{12.3}$$

En efecto, se puede ver que

$$\pi_\alpha = (k, \ldots, \alpha_k) \ldots (2, \ldots, \alpha_2)(1, 2, \ldots, \alpha_1),$$

donde $(a_1, \dots, a_r)$ denota al r-ciclo asociado. Esto se deduce de que el primer ciclo (que consta de $\alpha_1 - 1$ trasposiciones) manda α_1 al lugar 1. El segundo (que consta de $\alpha_2 - 2$ trasposiciones) manda α_2 al lugar 2 (observar que $(1, 2, \dots, \alpha_1)$ no movió a α_2) y deja a α_1 en el lugar 1. Se sigue así hasta mandar α_k al lugar k. Lo demás (los valores que toma en α') queda armado para producir la permutación π_α, porque se mantuvo su orden interno, y van a donde deben ir. Por ejemplo, los índices $1, \dots, \alpha_1 - 1$ están en α' (si $\alpha_1 > 1$) y se "corren" un lugar a la derecha por el primer ciclo. Luego, los "lugares" $2, \dots, \alpha_2 - 1$ están ocupados por más elementos de α' y se vuelven a corren con el segundo ciclo (manteniendo su orden original). Luego de aplicar los k ciclos, quedan todos los de α' ordenaditos y al final.

3. Sea $T_\alpha \in \mathcal{U}_{\mathcal{P}}(n)$, la matriz de permutación asociada a π_α, dada por

$$\begin{cases} T_\alpha e_i = e_{\alpha_i} & \text{si} \quad i = 1, \dots, k \\ T_\alpha e_{k+j} = e_{\alpha'_j} & \text{si} \quad j = 1, \dots n - k \ . \end{cases} \tag{12.4}$$

Tenemos entonces que $\det T_\alpha = \operatorname{sgn}(\pi_\alpha) = \operatorname{sgn}(\alpha)$. ▲

El siguiente resultado generaliza la Proposición 3.8.7 y el Corolario 3.8.9 a matrices y bloques cualesquiera (siempre que sean cuadrados).

Teorema 12.1.4. *Sean $A \in \mathcal{M}_n(\mathbb{C})$, $k \in \mathbb{I}_n$ y $\alpha, \beta \in Q_{k,n}$. Se tiene que*

$$A[\alpha|\beta] \in \mathcal{G}l(k) \implies \det A = \operatorname{sgn}(\alpha) \ \operatorname{sgn}(\beta) \ \det A[\alpha|\beta] \ \det A/[\alpha|\beta] \ . \tag{12.5}$$

Si también $A \in \mathcal{G}l(n)$, entonces $A/[\alpha|\beta] \in \mathcal{G}l(n-k)$ y

$$\Big(A/[\alpha|\beta]\Big)^{-1} = A^{-1}(\beta|\alpha) \ . \tag{12.6}$$

Demostración. Empecemos con el caso particular $\alpha = \beta = \mathbb{I}_k$. En este caso, se puede aplicar una argumento igual al de la prueba de la Proposición 3.8.7, que mostraremos brevemente: Un cálculo elemental prueba que A admite la factorización

$$A = \begin{bmatrix} I_n[\alpha] & 0 \\ A(\alpha|\alpha]A[\alpha]^{-1} & I_n(\alpha) \end{bmatrix} \begin{bmatrix} A[\alpha] & 0 \\ 0 & A/\alpha \end{bmatrix} \begin{bmatrix} I_n[\alpha] & A[\alpha]^{-1}A[\alpha|\alpha) \\ 0 & I_n(\alpha) \end{bmatrix} . \tag{12.7}$$

A partir de esto podemos deducir sin problemas la Eq. (12.5), porque los factores de la derecha y de la izquierda en el lado derecho de Eq. (12.7) tienen determinante 1, mientras que el factor central tiene determinante igual a $\det A[\alpha] \ \det(A/\alpha)$. También, Eq. (12.6) es consecuencia de la Eq. (12.7), tomando inversas de ambos lados.

Para probar el caso general, consideremos las matrices T_α, T_β definidas en Eq. (12.4). Llamemos $B = T_\alpha^{-1} A T_\beta$. Usando la Eq. (12.4) vemos que, como matrices de sus tamaños,

$$B[\mathbb{I}_k] = A[\alpha|\beta] \ , \ B(\mathbb{I}_k) = A(\alpha|\beta) \ , \ B(\mathbb{I}_k|\mathbb{I}_k] = A(\alpha|\beta] \quad \text{y} \quad B[\mathbb{I}_k|\mathbb{I}_k) = A[\alpha|\beta) \ . \tag{12.8}$$

Mostraremos la primera igualdad de la Eq. (12.8), ya que las demás se muestran de manera análoga: dados $i, j \in \mathbb{I}_k$, tenemos que

$$B[\mathbb{I}_k]_{ij} = (T_\alpha^{-1} A T_\beta)[\mathbb{I}_k]_{ij} \quad = \langle T_\alpha^{-1} A T_\beta \, e_j \, , e_i \rangle \; = \; \langle A T_\beta \, e_j \, , T_\alpha \, e_i \rangle$$

$$= \left\langle A e_{\beta_j} \, , e_{\alpha_i} \right\rangle \; = \; A_{\alpha_i \beta_j} \; = \; A[\alpha|\beta]_{ij} \; .$$

Observar que las igualdades de (12.8) aseguran que $B/[\mathbb{I}_k] = A/[\alpha|\beta]$. Luego

$$\operatorname{sgn}(\alpha) \; \operatorname{sgn}(\beta) \det A = \det B = \; \det B[\mathbb{I}_k|\mathbb{I}_k] \; \det B/[\mathbb{I}_k|\mathbb{I}_k] = \; \det A[\alpha|\beta] \; \det A/[\alpha|\beta] \; ,$$

ya que $\det T_\alpha = \operatorname{sgn}(\alpha)$. Finalmente, la Eq. (12.6) resulta de la relación:

$$A^{-1}(\beta|\alpha) = (T_\beta^{-1} A^{-1} T_\alpha)(\mathbb{I}_k) = B^{-1}(\mathbb{I}_k) = \big(B/[\mathbb{I}_k] \big)^{-1} = \big(A/[\alpha|\beta] \big)^{-1} \; . \qquad \blacksquare$$

En el siguiente enunciado veremos que toda matriz $A \in \mathcal{M}_n(\mathbb{C})$ puede ser aproximada tanto como se quiera por matrices tales que todas sus submatrices cuadradas son inversibles. Esto será usado para obtener varias identidades de determinantes a partir del Teorema 12.1.4.

Lema 12.1.5. *Dada $A \in \mathcal{M}_n(\mathbb{C})$ y $\varepsilon > 0$, existe $B \in \mathcal{M}_n(\mathbb{C})$ tal que*

1. *$\|A - B\| < \varepsilon$, donde $\| \cdot \|$ es una norma en $\mathcal{M}_n(\mathbb{C})$.*
2. *Todas las submatrices cuadradas de B son inversibles.*

Demostración. La cantidad total de submatrices cuadradas de A es $M = \sum\limits_{k=1}^{n} \binom{n}{k}^2$. Consideremos la función $\phi : \mathcal{M}_n(\mathbb{C}) \to \mathbb{C}^M$ que asigna a cada $B \in \mathcal{M}_n(\mathbb{C})$ la lista de los determinantes de todas sus submatrices cuadradas, en algún orden prefijado. Notar que ϕ es una función continua. Llamemos

$$\Gamma = \phi^{-1}\Big(\; \big\{ a \in \mathbb{C}^M : \; a_i \neq 0 \quad \text{ para todo } \quad i \in \mathbb{I}_M \big\} \; \Big).$$

Para probar el Lema, basta ver que Γ es denso en $\mathcal{M}_n(\mathbb{C})$. Llamemos, para cada $i \in \mathbb{I}_M$,

$$\Gamma_i = \phi^{-1}\Big(\; \big\{ a \in \mathbb{C}^M : a_i \neq 0 \big\} \; \Big).$$

Es claro que todos estos conjuntos son abiertos y que $\Gamma = \bigcap\limits_{i=1}^{M} \Gamma_i$. Por otra parte, cada Γ_i es denso en $\mathcal{M}_n(\mathbb{C})$, porque $\mathcal{G}l(k)$ es denso en $\mathcal{M}_k(\mathbb{C})$ para todo $k \in \mathbb{N}$. Por ejemplo, si $\phi(A)_1 = \det A$, entonces $\Gamma_1 = \mathcal{G}l(n)$. El resultado se sigue de que una intersección finita de abiertos densos es densa. En efecto, si U y V son abiertos densos y Z es abierto, entonces $Z \cap U$ es un abierto no vacío. Entonces $(Z \cap U) \cap V = Z \cap (U \cap V) \neq \varnothing$ para todo abierto Z. Por lo tanto $U \cap V$ es denso. Por inducción, quien dice 2, dice M. $\blacksquare$

12.2 Identidades asociadas a determinantes

Teorema 12.2.1 (Identidad de Jacobi). *Sea $A \in \mathcal{G}l(n)$. Entonces*

$$\det A^{-1}[\alpha|\beta] \;=\; \operatorname{sgn}(\alpha)\,\operatorname{sgn}(\beta)\,\frac{\det A(\beta|\alpha)}{\det A}\ , \qquad \text{para todo par} \qquad \alpha,\beta \in Q_{k,n}\ . \tag{12.9}$$

Demostración. Se sigue de las ecuaciones (12.6) y (12.5), aplicadas a α' y β':

$$\det A^{-1} \;=\; \operatorname{sgn}(\alpha)\,\operatorname{sgn}(\beta)\,\det A^{-1}[\alpha|\beta]\,\det A^{-1}/[\alpha|\beta] \qquad \Longrightarrow$$

$$\Longrightarrow \qquad \det A^{-1}[\alpha|\beta] \;=\; \frac{\operatorname{sgn}(\alpha)\operatorname{sgn}(\beta)}{\det A}\cdot(\det A^{-1}/[\alpha|\beta])^{-1} = \frac{\operatorname{sgn}(\alpha)\,\operatorname{sgn}(\beta)}{\det A}\cdot\det A(\beta|\alpha)\ ,$$

lo que culmina la prueba. ■

Observación 12.2.2. Cuando $k = 1$, la Eq. (12.9) induce la identidad:

$$(A^{-1})_{ij} = (-1)^{i+j}\ \frac{\det A(j|i)}{\det A}\ , \quad i,j \in \mathbb{I}_n\ , \tag{12.10}$$

conocida como la regla de Cramer. ▲

12.2.3. Sean $J_n = \operatorname{diag}(1,-1,1,-1,\dots,(-1)^{n-1}) \in \mathcal{U}(n)$, y $\omega,\alpha \in Q_{k,n}$. Luego

$$\det J_n[\alpha|\omega] = \delta_{\alpha\beta}\ \operatorname{sgn}(\alpha)(-1)^{\frac{k(k-1)}{2}}\ ,$$

donde $\delta_{\alpha,\beta} = 1$ o 0 de acuerdo a si $\alpha = \beta$ o $\alpha \neq \beta$. En efecto, si $\alpha \neq \omega$, en $J_n[\alpha|\omega]$ hay una columna de ceros, y por lo tanto su determinante es cero. Cuando $\alpha = \omega$ tenemos que, si p_α denota al número de elementos pares de α, entonces

$$p_\alpha \equiv \sum_{i=1}^{k}(\alpha_i - 1) = \operatorname{tr}\alpha - k \quad (\text{módulo}\ \ 2)\ .$$

Luego

$$\det J_n[\alpha] = (-1)^{p_\alpha} = (-1)^{\operatorname{tr}\alpha - k} = \operatorname{sgn}(\alpha)\cdot(-1)^{\frac{k(k+1)}{2}-k} = \operatorname{sgn}(\alpha)(-1)^{\frac{k(k-1)}{2}}\ ,$$

lo que culmina la prueba. Si $A \in \mathcal{G}l(n)$, suele notarse $A^{\#} = J_nA^{-1}J_n$ a la llamada *inversión* de A. La siguiente igualdad se sigue de la Eq. (12.9), usando (12.1) (Cauchy-Binnet):

$$\det(J_nA^{-1}J_n)[\alpha|\beta] = \frac{\det A(\beta|\alpha)}{\det A} \quad \text{para} \quad \alpha,\beta \in Q_{k,n}\ , \tag{12.11}$$

dado que $\det J_n[\alpha]\cdot\operatorname{sgn}(\alpha) = \det J_n[\beta]\cdot\operatorname{sgn}(\beta) = (-1)^{k(k-1)/2}$. ▲

12.2.4. La siguiente igualdad es válida para toda $A \in \mathcal{M}_n(\mathbb{R})$: dados $\alpha, \beta \in Q_{k,n}$,

$$\sum_{\omega \in Q_{k,n}} \operatorname{sgn}(\omega) \ \det A[\alpha|\omega] \ \det A(\beta|\omega) = \delta_{\alpha,\beta} \ \operatorname{sgn}(\beta) \ \det A \ . \tag{12.12}$$

De hecho, cuando A es inversible, por Eq. (12.9), el lado izquierdo de la Eq. (12.12) da

$$\operatorname{sgn}(\beta) \ \det A \sum_{\omega \in Q_{k,n}} \det A[\alpha|\omega] \ \det A^{-1}[\omega|\beta] = \operatorname{sgn}(\beta) \ \det A \ \det I_n[\alpha|\beta] \ ,$$

por la fórmula de Cauchy Binnet (12.1). El caso no inversible se deduce por continuidad. Observar que, tomando $k = 1$, y fijando cualquier $r \in \mathbb{I}_n$ como $\alpha = \beta$, nos queda el algoritmo usual propuesto en el Ejercicio 7.5.11 (y usando n veces) desarrollando por la fila r:

$$\det A = \sum_{i \in \mathbb{I}_n} \operatorname{sgn}\{r\} \operatorname{sgn}\{i\} \ A_{r,i} \ \det A(r|i) = \sum_{i \in \mathbb{I}_n} (-1)^{r+i} \ A_{r,i} \ \det A(r|i) \ . \tag{12.13}$$

El desarrollo por columnas sale aplicando esto a A^T. ▲

12.2.5. Dados $\alpha, \beta \in Q_{k,n}$ y, además, $\omega, \tau \in Q_{l,n}$ tales que $\omega \subseteq \alpha', \tau \subseteq \beta'$, sean

$$\mu = \alpha \cup \omega = (\mu_1, \mu_2, \dots, \mu_{k+l}) \quad \text{y} \quad \nu = \beta \cup \tau = (\nu_1, \nu_2, \dots, \nu_{k+l}) \in Q_{k+l,n} \ .$$

Existen entonces γ y $\sigma \in Q_{k,k+l}$ tales que $\alpha_i = \mu_{\gamma_i}$ y $\beta_i = \nu_{\sigma_i}$, $i \in \mathbb{I}_k$. Luego definimos

$$\operatorname{sgn} \frac{\alpha}{\alpha \cup \omega} = \operatorname{sgn}(\gamma) = (-1)^{\operatorname{tr}\gamma - \frac{k(k+1)}{2}} \ , \quad \operatorname{sgn} \frac{\beta}{\beta \cup \tau} = \operatorname{sgn}(\sigma) \ . \tag{12.14}$$

Con estas notaciones, se tiene la siguiente versión local del Teorema 12.1.4:

$$\det A[\alpha|\beta] \det \Big((A/[\alpha|\beta])[\omega|\tau] \Big) = \operatorname{sgn} \frac{\alpha}{\alpha \cup \omega} \ \operatorname{sgn} \frac{\beta}{\beta \cup \tau} \ \det A[\alpha \cup \omega|\beta \cup \tau] \tag{12.15}$$

En efecto, consideremos la matriz $B = (a_{\mu_i \nu_j})_{i,j \in \mathbb{I}_{k+l}} \in \mathcal{M}_{k+l}(\mathbb{R})$. Entonces vemos que Eq. (12.15) coincide con Eq. (12.5) para B, γ, σ en lugar de A, α, β, respectivamente. De hecho, como $B = A[\mu|\nu] = A[\alpha \cup \omega|\beta \cup \tau]$, entonces $B[\gamma|\sigma] = A[\alpha|\beta]$ y, por otra parte,

$$\begin{aligned} B/[\gamma|\sigma] &= B(\gamma|\sigma) - B(\gamma|\sigma] \, B[\gamma|\sigma]^{-1} \, B[\gamma|\sigma) \\ &= A(\alpha|\beta)[\omega|\tau] - A(\alpha|\beta] \, A[\alpha|\beta]^{-1} \, A[\alpha|\beta)[\omega|\tau] \\ &= (A/[\alpha|\beta])[\omega|\tau] \ . \end{aligned}$$

Una consecuencia inmediata es la siguiente caracterización de las entradas de un complemento de Schur: Dados $\alpha, \beta \in Q_{k,n}$, se tiene

$$\{A/[\alpha|\beta]\,\}_{(\alpha'_i,\beta'_j)} = \operatorname{sgn} \frac{\alpha}{\alpha \cup \{\alpha'_i\}} \ \ \operatorname{sgn} \frac{\beta}{\beta \cup \{\beta'_j\}} \ \frac{\det A[\alpha \cup \{\alpha'_i\}|\beta \cup \{\beta'_j\}]}{\det A[\alpha|\beta]} \tag{12.16}$$

Corolario 12.2.6. *Sea* $A \in \mathcal{M}_n(\mathbb{R})$ *y sea* $r \in \mathbb{I}_n$ *tal que* $A_{rr} \neq 0$. *Entonces, para todo* $\alpha \in Q_{k,n}$ *tal que* $r \notin \alpha$*, se tiene que*

$$(A/[r])[\alpha] \ = A[\{r\} \cup \alpha]/[r] \ .$$

Todas esas letras significan que las submatrices principales de $A/[r]$ sólo dependen de las entradas correspondientes de A (y no de las demás).

Demostración. Dados $i, j \in \alpha$, la fórmula (12.16) asegura que

$$\begin{aligned} \Big((A/[r])[\alpha] \Big)_{ij} &= \Big((A/[r] \Big)_{ij} \\ &= \operatorname{sgn} \frac{r}{\{i,r\}} \operatorname{sgn} \frac{r}{\{j,r\}} \frac{\det A[\{i,r\}|\{j,r\}]}{A_{rr}} \\ &= \Big(A[\{r\} \cup \alpha]/[r] \Big)_{ij} \ , \end{aligned}$$

por lo ambas matrices coinciden. ■

12.2.7 (Identidad de Sylvester). Dados $A \in \mathcal{M}_n(\mathbb{R})$ y $\alpha, \beta \in Q_{k,n}$, se cumple que

$$\det \Big(\big\{ \det A[\alpha \cup \{\alpha'_i\} | \beta \cup \{\beta'_j\}] \big\}_{i,j \in \mathbb{I}_{n-k}} \Big) = \det A \cdot \det A[\alpha|\beta]^{n-k-1} \tag{12.17}$$

Para probarlo, tomemos los números

$$\varepsilon_i = \operatorname{sgn} \ \frac{\alpha}{\alpha \cup \{\alpha'_i\}} \quad \text{y} \quad \rho_j = \operatorname{sgn} \ \frac{\beta}{\beta \cup \{\beta'_j\}} \ , \qquad \text{para todo } i, j \in \mathbb{I}_{n-k} \ .$$

Por la Eq. (12.16), vemos que el lado izquierdo de la Eq. (12.17) es igual a

$$\det \Big[\det A[\alpha|\beta] \, (A/[\alpha|\beta])_{ij} \ \varepsilon_i \, \rho_j \Big] =$$

$$\det A[\alpha|\beta]^{n-k} \ \det \Big[\operatorname{diag}(\varepsilon_1 \,, \ldots, \varepsilon_{n-k}) \, A/[\alpha|\beta] \operatorname{diag}(\rho_1 \,, \ldots, \rho_{n-k}) \Big] =$$

$$\det A[\alpha|\beta]^{n-k-1} \ \det A[\alpha|\beta] \det(A/[\alpha|\beta]) \prod_{i=1}^{n-k} \operatorname{sgn} \ \frac{\alpha}{\alpha \cup \{\alpha'_i\}} \cdot \prod_{i=1}^{n-k} \operatorname{sgn} \ \frac{\beta}{\beta \cup \{\beta'_j\}} \ .$$

La fórmula (12.17) se sigue ahora de la Eq. (12.5) y del siguiente resultado: ▲

Lema 12.2.8. *Sea* $\alpha \in Q_{k,n}$. *Entonces*

$$\prod_{i=1}^{n-k} \operatorname{sgn} \frac{\alpha}{\alpha \cup \{\alpha'_i\}} \quad = \operatorname{sgn}(\alpha) \ . \tag{12.18}$$

Demostración. Recordemos la definición de los signos de α y de $\frac{\alpha}{\alpha\cup\{\alpha'_i\}}$, para cada entrada $i \in \mathbb{I}_{n-k}$. En ambos casos, es calcular el signo de la permutación que manda a los buenos al principio y los malos al final. En otras palabras, están determinados por la cantidad de trasposiciones necesarias para efectuar tales ordenamientos. Pero miremos el siguiente proceso: empezamos por el último elemento de α' y lo corremos hasta el final (si no estaba allí). Al penúltimo, lo corremos hasta justo despues de α_k . Y seguimos así con todos los de α' hasta llegar al primero. El número de trasposiciones en cada paso es justo el que determina el correspondiente sgn $\frac{\alpha}{\alpha\cup\{\alpha'_i\}}$, porque los α_j que tiene arriba quedaron pegados entre sí, al haber sacado antes los de α' mayores que α'_i . Pero al final de todo mandamos prolijamente a todo α' hasta el final, por lo que la suma total da el $\operatorname{sgn}(\alpha)$. ■

Observación 12.2.9. La prueba anterior es un poco charlada, pero créanme que ustedes lo preferirían así, antes que tener que leer una cuenta explícita. Igual daremos una versión guiada de esa cuenta en los ejercicios. Como aval de la prueba anterior, se verá allí que la suma de los exponentes de la productoria de (12.18) da **igual** (y no sólo congruente) al exponente de -1 en $\operatorname{sgn}(\alpha)$, según las Eq's (12.3) y (12.14). ▲

12.3 Un poco más de complementos de Schur

Recordemos que, si $A \in \mathcal{M}_n(\mathbb{C})$, $k \in \mathbb{I}_n$ y $\alpha, \beta \in Q_{k,n}$ satisfacen que $A[\alpha|\beta]$ es inversible, el complemento de Schur $A[\alpha|\beta]$ en A es la matriz

$$A/[\alpha|\beta] = A(\alpha|\beta) - A(\alpha|\beta] \cdot A[\alpha|\beta]^{-1} \cdot A[\alpha|\beta) \in \mathcal{M}_{n-k}(\mathbb{C}) \ ,$$

indexada por α' y β'. Llamemos $\mathbb{C}^{\beta'} = \operatorname{Gen}\{e_j : j \in \beta'\}$ y $\mathbb{C}^{\beta} = \operatorname{Gen}\{e_k : k \in \beta\}$. Representemos $\mathbb{C}^n = \mathbb{C}^{\beta'} \oplus \mathbb{C}^{\beta}$, poniendo que un vector $\mathbb{C}^n \ni x = x_{\beta'} + x_\beta$. Análogamente $\mathbb{C}^n = \mathbb{C}^{\alpha'} \oplus \mathbb{C}^{\alpha}$. Observar que $A[\alpha|\beta]$ opera desde $\mathbb{C}^\beta$ hacia $\mathbb{C}^\alpha$, lo que denotaremos $A[\alpha|\beta] : \mathbb{C}^\beta \to \mathbb{C}^\alpha$. Por lo tanto $A[\alpha|\beta]^{-1} : \mathbb{C}^\alpha \to \mathbb{C}^\beta$. Los otros cachos de A operan en forma coherente, por ejemplo $A[\alpha|\beta) : \mathbb{C}^{\beta'} \to \mathbb{C}^\alpha$ y así. Con estas notaciones, podemos pensar que $A/[\alpha|\beta] : \mathbb{C}^{\beta'} \to \mathbb{C}^{\alpha'}$.

Proposición 12.3.1. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ y supongamos que $A[\alpha|\beta]$ es inversible para ciertos $\alpha, \beta \in Q_{k,n}$. Definamos*

$$P(A,\alpha,\beta) \in \mathcal{M}_n(\mathbb{C}) \quad \text{dado por} \quad P(A,\alpha,\beta)\ x = x_{\beta'} - A[\alpha|\beta]^{-1}A[\alpha|\beta)\, x_{\beta'} \quad ,$$

para $x = x_{\beta'} + x_\beta \in \mathbb{C}^n$. Entonces, si abreviamos $P(A,\alpha,\beta) = P$, se tiene que

1. $P^2 = P$.
2. $\ker P = \operatorname{Gen}\{e_j : j \in \beta\} = \mathbb{C}^\beta$.
3. $(AP)(\alpha|\beta) = A/[\alpha|\beta]$ *y las demás coordenadas de AP son nulas.*

Demostración. Observar que $Px = Px_{\beta'}$ y que $-A[\alpha|\beta]^{-1}A[\alpha|\beta)\, x_{\beta'} \in \mathbb{C}^\beta$ para todo $x \in \mathbb{C}^n$. Por ello es claro que $P^2 = P$ y que $\ker P = \mathbb{C}^\beta$, porque los sumandos que definen a P no interfieren entre sí, por lo que nunca se anula en $\mathbb{C}^{\beta'} \setminus \{0\}$.

De lo anterior, deducimos que si $x = x_\beta \in \mathbb{C}^\beta$, entonces $APx = 0$. Esto dice que $(AP)[\alpha|\beta]$ y $(AP)(\alpha|\beta]$ son matrices nulas, porque son las partes de AP que actúan en $\mathbb{C}^\beta$ y van a lugares que no interfieren entre sí. Por lo tanto, para cualquier $x \in \mathbb{C}^n$, se tiene que

$$\begin{aligned} APx &= APx_{\beta'} = \big[A(\alpha|\beta) + A[\alpha|\beta)\big]x_{\beta'} - \big[A(\alpha|\beta] + A[\alpha|\beta]\,\big]A[\alpha|\beta]^{-1}A[\alpha|\beta)\,x_{\beta'} \\ &= \Big(A(\alpha|\beta) - A(\alpha|\beta]A[\alpha|\beta]^{-1}A[\alpha|\beta)\Big)\,x_{\beta'} = A/[\alpha|\beta]\,x_{\beta'} \in \mathbb{C}^{\alpha'} \ . \end{aligned}$$

Esto muestra que $(AP)[\alpha|\beta) \equiv 0$ y que $(AP)(\alpha|\beta) = A/[\alpha|\beta]$. ■

Corolario 12.3.2. *Sean $A \in \mathcal{M}_n(\mathbb{C})$ y $\alpha, \beta \in Q_{k,n}$ tales que $A[\alpha|\beta]$ es inversible.*

1. *Para todo $x_{\beta'} \in \mathbb{C}^{\beta'}$ existe un $x_\beta \in \mathbb{C}^\beta$ tal que $A/[\alpha|\beta]\,x_{\beta'} \ = \ A(x_{\beta'} + x_\beta)$.*
2. *Sea $Q \in \mathcal{M}_n(\mathbb{C})$ tal que*

$$Q^2 = Q \quad , \quad \ker Q = \mathbb{C}^\beta \qquad y \qquad R(A \cdot Q) \subseteq \mathbb{C}^{\alpha'} \ . \tag{12.19}$$

   Entonces se tiene que $Q = P(A, \alpha, \beta)$.

Demostración. Sea $P = P(A, \alpha, \beta)$ la proyección de la Proposición 12.3.1. Luego

$$A/[\alpha|\beta]\,x_{\beta'} \ = \ APx_{\beta'} = A\,x_{\beta'} - A\Big(A[\alpha|\beta]^{-1}A[\alpha|\beta)\,x_{\beta'}\Big) \ .$$

Luego basta tomar $x_\beta = -A\ , A[\alpha|\beta]^{-1}A[\alpha|\beta)\,x_{\beta'}$. Si me dan ahora un Q que cumple (12.19), entonces $Q^2 = Q$ y $\ker Q = \mathbb{C}^\beta$. Luego, como en el Ejercicio 3.9.19, se puede ver que

$$Qx = Qx_{\beta'} = x_{\beta'} + Q[\beta, \beta)x_{\beta'} \ , \qquad \text{para todo} \quad x = x_{\beta'} + x_\beta \in \mathbb{C}^n \ .$$

El hecho de que $R(AQ) \subseteq \mathbb{C}^{\alpha'}$ indica que $(AQ)[\alpha|\beta) = 0$. De lo anterior, uno deduce que

$$0 = (AQ)[\alpha|\beta) = A[\alpha|\beta) + A[\alpha|\beta]Q[\beta, \beta) \implies A[\alpha|\beta]Q[\beta, \beta) = -A[\alpha|\beta) \ .$$

Como $A[\alpha|\beta]$ es inversible, tenemos que $Q[\beta, \beta) = -A[\alpha|\beta]^{-1}A[\alpha|\beta)$, o sea que $Q = P$. ■

El siguiente teorema es un resultado análogo a la Proposición 3.8.5.

Teorema 12.3.3. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ y supongamos que $A[\alpha|\beta]$ es inversible para ciertos $\alpha, \beta \in Q_{k,n}$. Sean $\omega, \tau \in Q_{r,n}$, tales que $\omega \subseteq \alpha'$ y $\tau \subseteq \beta'$. Entonces*

1. *$(A/[\alpha|\beta])[\omega|\tau] \in \mathcal{G}l\,(r)$ si y sólo si $A[\alpha \cup \omega|\beta \cup \tau] \in \mathcal{G}l\,(k + r)$.*
2. *En este caso se cumple la siguiente igualdad de matrices:*

$$\Big(A/[\alpha|\beta]\Big)/[\omega|\tau] = A/[\alpha \cup \omega|\beta \cup \tau] \ . \tag{12.20}$$

Demostración. El item 1 se deduce en forma inmediata de la Eq. (12.15). Por la Proposición 12.3.1, tenemos tres proyectores: $P(A, \alpha, \beta)$, $P(A/[\alpha|\beta], \omega, \tau)$ y $P(A, \alpha \cup \omega, \beta \cup \tau)$ tales que

$$A/[\alpha|\beta] = A \cdot P(A, \alpha, \beta) \ , \ A/[\alpha \cup \omega|\beta \cup \tau] = A \cdot P(A, \alpha \cup \omega, \beta \cup \tau) \quad \text{ y}$$

$$\Big(A/[\alpha|\beta]\Big)/[\omega|\tau] = A/[\alpha|\beta] \cdot P(A/[\alpha|\beta], \omega, \tau) = A \cdot P(A, \alpha, \beta) \cdot P(A/[\alpha|\beta], \omega, \tau) \ , \tag{12.21}$$

salvo los ceros. Ahora bien, $\ker P(A, \alpha, \beta) = \mathbb{C}^\beta$, mientras que $A/[\alpha|\beta]$ opera sólo en $\mathbb{C}^{\beta'}$ por lo que, si pensamos a $P(A/[\alpha|\beta], \omega, \tau) \in \mathcal{M}_n(\mathbb{C})$ con ceros fuera de β', se tiene que $\ker P(A/[\alpha|\beta], \omega, \tau) = \mathbb{C}^{\beta\cup\tau}$. En particular, como $\mathbb{C}^\beta \subseteq \mathbb{C}^{\beta\cup\tau}$, se tiene que

$$P(A/[\alpha|\beta], \omega, \tau)\big(I - P(A, \alpha, \beta)\big) = 0 \implies P(A/[\alpha|\beta], \omega, \tau)P(A, \alpha, \beta) = P(A/[\alpha|\beta], \omega, \tau)$$

y la matriz $Q = P(A, \alpha, \beta) \cdot P(A/[\alpha|\beta], \omega, \tau) \in \mathcal{M}_n(\mathbb{C})$ cumple que

$$Q^2 = Q \quad , \quad \ker Q = \mathbb{C}^{\beta\cup\tau} \quad \text{ y } \quad R(A \cdot Q) \subseteq \mathbb{C}^{(\alpha\cup\omega)'} \ , \tag{12.22}$$

donde igualdad del medio es un ligero ejercicio, y la última inclusión surge de que, como dice la Eq. (12.21), se tiene que $\Big(A/[\alpha|\beta]\Big)/[\omega|\tau] = AQ$. Ahora, la Eq. (12.22) asegura, vía el Corolario 12.3.2, que $Q = P(A, \alpha \cup \omega, \beta \cup \tau)$, por lo que

$$\Big(A/[\alpha|\beta]\Big)/[\omega|\tau] \ = \ AQ \ = A/[\alpha \cup \omega|\beta \cup \tau] \ . \qquad \blacksquare$$

Observación 12.3.4. Otra forma de probar la fórmula (12.20), con técnicas más parecidas a las del resto de este Capítulo, sería calculando las coordenadas como determinantes por medio de las ecuaciones (12.15) y (12.16). En efecto, sean $i \in (\alpha \cup \omega)'$ y $j \in (\beta \cup \tau)'$, llamemos $\mu = \{i\}$ y $\nu = \{j\}$ y obsevemos que, en el lado izquierdo de (12.20) tenemos

$$\begin{aligned} \big(\,(A/[\alpha|\beta])/[\omega|\tau]\,\big)_{i,j} &= \operatorname{sgn} \tfrac{\omega}{\omega\cup\mu} \cdot \operatorname{sgn} \tfrac{\tau}{\tau\cup\nu} \ \frac{\det(A/[\alpha|\beta])[\omega \cup \mu|\tau \cup \nu]}{\det\,(A/[\alpha|\beta])[\omega|\tau]} \\ &= \varepsilon \ \frac{\det A[\alpha \cup \omega \cup \mu|\beta \cup \tau \cup \nu]}{\det A[\alpha \cup \omega|\beta \cup \tau]} \ , \quad \text{ donde} \end{aligned}$$

$$\varepsilon \ = \operatorname{sgn} \tfrac{\omega}{\omega\cup\mu} \cdot \operatorname{sgn} \tfrac{\alpha}{\alpha\cup\omega} \cdot \operatorname{sgn} \tfrac{\alpha}{\alpha\cup\omega\cup\mu} \cdot \operatorname{sgn} \tfrac{\tau}{\tau\cup\nu} \cdot \operatorname{sgn} \tfrac{\beta}{\beta\cup\tau} \cdot \operatorname{sgn} \tfrac{\beta}{\beta\cup\tau\cup\nu} \ .$$

La misma entrada del lado derecho de (12.20) es igual a

$$\Big(A/[\alpha \cup \omega|\beta \cup \tau]\Big)_{ij} \ = \ \operatorname{sgn} \tfrac{\alpha\cup\omega}{\alpha\cup\omega\cup\mu} \ \operatorname{sgn} \tfrac{\beta\cup\tau}{\beta\cup\tau\cup\nu} \ \frac{\det A[\alpha \cup \omega \cup \mu|\beta \cup \tau \cup \nu]}{\det A[\alpha \cup \omega|\beta \cup \tau]} \ .$$

Por lo tanto, ambas matrices tienen todas sus coordenadas iguales, salvo los signos. Para ver que ellos coniciden, bastaría verificar que, para todo $i \in (\alpha \cup \omega)'$ y todo $j \in (\beta \cup \tau)'$,

$$\varepsilon \ = \operatorname{sgn} \tfrac{\alpha\cup\omega}{\alpha\cup\omega\cup\mu} \ \operatorname{sgn} \tfrac{\beta\cup\tau}{\beta\cup\tau\cup\nu} \ , \quad \text{ donde } \mu = \{i\} \text{ y } \nu = \{j\} \ . \tag{12.23}$$

Esto sale usando la Eq. (12.14) y un sinnúmero de cuentas que, con gran alegría, le dejamos al lector interesado como ejercicio. $\blacktriangle$

12.4 Ejercicios

Ejercicios que aparecen en el texto

12.4.1. Completar los detalles de la prueba de la Eq. (12.3).

12.4.2. Dar una prueba numérica de la Eq. (12.18). Es decir: Dada $\alpha \in Q_{k,n}$ probar que $\prod\limits_{i=1}^{n-k} \operatorname{sgn} \frac{\alpha}{\alpha \cup \{\alpha'_i\}} = \operatorname{sgn}(\alpha)$. Se suguieren los siguientes pasos:

1. Para cada $i \in \mathbb{I}_{n-k}$, sea $\gamma^i \in Q_{k,k+1}$, definido como en 12.2.5 para α y $\alpha \cup \{\alpha'_i\}$. Recordemos que en la Eq. (12.14) se vió que $\operatorname{sgn} \frac{\alpha}{\alpha \cup \{\alpha'_i\}} = \operatorname{sgn}(\gamma^i) = (-1)^{\operatorname{tr} \gamma^i - \frac{k(k+1)}{2}}$, donde $\gamma^i \in \mathbf{S}_{k+1}$ manda α'_i al final de $\alpha \cup \{\alpha'_i\}$. El tema es ver en qué lugar de entre las entradas de α que ubicado el α'_i . Pongamos, por conveniencia, que $\alpha_0 = 0$ y $\alpha_{k+1} = \infty$. En tal caso, mostrar que
$$\operatorname{tr}\ \gamma^i = \frac{(k+2)(k+1)}{2} - j\ , \qquad \text{cuando se tiene que} \qquad \alpha_{j-1} < \alpha'_i < \alpha_j\ .$$
2. Deducir que $\operatorname{sgn} \frac{\alpha}{\alpha \cup \{\alpha'_i\}} = \operatorname{sgn}(\gamma^i) = (-1)^{k-j+1}$ si $\alpha_{j-1} < \alpha'_i < \alpha_j$.
3. Contemos cuantos hay de cada tipo: probar que
$$|\{i \in \mathbb{I}_{n-k} : \alpha'_i < \alpha_1\}| = \alpha_1 - 1 \quad , \quad |\{i \in \mathbb{I}_{n-k} : \alpha_k < \alpha'_i\}| = n - \alpha_k \qquad \text{y}$$
$$|\{i \in \mathbb{I}_{n-k} : \alpha_{j-1} < \alpha'_i < \alpha_j\}| = \alpha_j - \alpha_{j-1} - 1 \qquad \text{cuando } 2 \le j \le k\ .$$
4. Ahora sí, calcular el exponente de la productoria:
$$\begin{aligned} \sum_{i=1}^{n-k} \left(\operatorname{tr} \gamma^i - \frac{k(k+1)}{2} \right) &= \sum_{j=1}^{k} (k-j+1)(\alpha_j - \alpha_{j-1} - 1) + 0 \cdot (n - \alpha_k) \\ &= \sum_{j=1}^{k} \alpha_j - \sum_{j=1}^{k} (k-j+1) = \operatorname{tr} \alpha - \frac{k(k+1)}{2}\ . \end{aligned}$$
5. Usando la Eq. (12.3), concluir la prueba de la fórmula (12.18).

12.4.3. Dar una prueba de la fórmula (12.20) basandose en el camino delineado en la Observación 12.3.4. En otras palabras, probar la Eq. (12.23). Se sugieren dos caminos. El primero es encontrar permutaciones que tengan esos signos y que hagan lo mismo, como en la prueba del Lema 12.2.8. El segundo es contar todo a lo bestia, y mostrar que son congruentes módulo 2. En ambos casos, se puede reducir el trabajo a probar dos identidades más cortas y similares, a saber: Dados $\alpha \in Q_{k,n}$, $\omega \in Q_{r,n}$ y $\mu = \{i\}$, todos disjuntos,
$$\operatorname{sgn} \frac{\alpha}{\alpha \cup \omega} \cdot \operatorname{sgn} \frac{\omega}{\omega \cup \mu} \cdot \operatorname{sgn} \frac{\alpha}{\alpha \cup \omega \cup \mu} = \operatorname{sgn} \frac{\alpha \cup \omega}{\alpha \cup \omega \cup \mu}\ ,$$
y lo mismo para β, τ y ν. Por ejemplo, se puede mandar $\omega \cup \mu$ al final, después mandar a μ al último lugar, y después volver a mezclar a ω con α.

Ejercicios nuevos

Notación: Sea $A \in \mathcal{M}_{m,n}(\mathbb{C})$ tal que $m > n$ y $\operatorname{rk}(A) = n$.

1. Dado $I \subseteq \mathbb{I}_m$ con $|I| = r$ y dado $b \in \mathbb{C}^m$, denotaremos por b_I al vector de C^r que se obtiene dejando sólo las coordenadas de b que pertenecen a I.
2. $J(A) := \{I \subseteq \mathbb{I}_m : |I| = n \text{ y } \det A[I] \neq 0\}$. Observar que $\operatorname{rk}(A) = n \implies J(A) \neq \varnothing$.

12.4.4 (Ben Tal - Teboulle)**.** Dada $A \in \mathcal{M}_{m,n}(\mathbb{C})$ tal que $m > n$ y $rk(A) = n$, sea $\mathbf{c} \in \mathbb{C}^n$ la solución del problema de cuadrados mínimos

$$\min_{x \in \mathbb{C}^n} \|Ax - b\|^2 \quad \text{para un } b \in \mathbb{C}^m \text{ fijo }.$$

Si para cada $I \in J(A)$ denotamos $\mathbf{c}_I = A[I]^{-1} b_I$, probar que $\mathbf{c}$ pertenece a la cápsula convexa de los $\mathbf{c}_I$. Se suguiere probar que $\mathbf{c}$ es la única solución de la denomina ecuación normal

$$A^* A x = A^* b.$$

Luego, usar la regla de Cramer y la fórmula de Cauchy-Binet.

Capítulo 13

Matrices totalmente positivas

Este capítulo está basado en un trabajo de T. Ando [20], aparecido en 1987, y está escrito utilizando como punto de partida un trabajo de A. Iglesias. Las herramientas fundamentales para las demostraciones son los intrincados resultados del Capítulo anterior.

13.1 Definiciones y criterios de positividad total

En esta sección introducimos las nociones de regularidad de signo y positividad total.

Definición 13.1.1. 1. Llamaremos *sucesión de signatura* a una sucesión

$$\varepsilon = (\varepsilon_i)_{i\in\mathbb{N}} \in \{-1,1\}^{\mathbb{N}} \ , \quad \text{es decir, tal que} \quad \varepsilon_i = \pm 1 \quad \text{para todo} \quad i \in \mathbb{N} \ .$$

2. Dado $\lambda \in \mathbb{R}$ con $|\lambda| = 1$, notaremos $\lambda\varepsilon$ a la sucesión de signatura $\lambda\varepsilon = (\lambda\varepsilon_i)_{i\in\mathbb{N}}$.

3. Si τ es otra sucesión de signatura, llamaremos $\tau\varepsilon = (\tau_i \ \varepsilon_i)_{i\in\mathbb{N}}$. ▲

Definición 13.1.2. Sean $A \in \mathcal{M}_{n,m}(\mathbb{R})$ y ε una sucesión de signatura. Sea $r = \min\{n, m\}$.

1. Decimos que A es de *signo regular con signatura* ε, y abreviaremos diciendo que A es ε-RS si, en la base canónica $\mathcal{E}^{\wedge}_{k,n} = \{\sqrt{k!}\ \mathbf{e}^{\wedge}_{\alpha} : \alpha \in Q_{k,n}\}$de $\Lambda^k\mathbb{H}_n$, se tiene que

$$\varepsilon_k \cdot \Lambda^k A \geqslant 0 \quad \text{para todo} \quad k \in \mathbb{I}_r \ , \tag{13.1}$$

   La regularidad de signo de A es equivalente a la condición

$$\varepsilon_k \cdot a_{\beta_1} \wedge a_{\beta_2} \wedge \cdots \wedge a_{\beta_k} \geqslant 0, \quad \text{para} \quad \beta \in Q_{k,m} \ \ k \in \mathbb{I}_r \ , \tag{13.2}$$

   o, por la Eq. (7.19), en forma de determinantes,

$$\varepsilon_k \det A[\alpha|\beta] \geq 0 \quad \text{para} \quad \alpha \in Q_{k,n} \ , \ \beta \in Q_{k,m} \quad k \in \mathbb{I}_r \ , \tag{13.3}$$

2. A se dice *estrictamente de signo regular con signatura* ε (A es ε-ERS) si, en la Eq. (13.1) (o, equivalentemente, en la Eq. (13.2) o la Eq. (13.3)), reemplazamos $\geqslant$ por $>$.

3. Decimos que A es *totalmente positiva* (y abreviaremos TP) si es ε-RS respecto de la sucesión $\varepsilon \equiv 1$, es decir, si
$$\Lambda^k A \geqslant 0, \quad k \in \mathbb{I}_r \ , \tag{13.4}$$
o equivalentemente si
$$a_{\beta_1} \wedge a_{\beta_2} \wedge \cdots \wedge a_{\beta_k} \geqslant 0, \qquad \text{para} \qquad \beta \in Q_{k,m} \quad k \in \mathbb{I}_r \ , \tag{13.5}$$
es decir, si
$$\det A[\alpha|\beta] \geq 0 \qquad \text{para} \qquad \alpha \in Q_{k,n} \ , \ \beta \in Q_{k,m} \quad k \in \mathbb{I}_r \ . \tag{13.6}$$

4. A se dice estrictamente totalmente positiva (ETP) si $\geqslant$ es reemplazado por $>$ en las ecuaciones (13.4), (13.5) o (13.6). ▲

Para testear la regularidad de signo de A se requiere chequear los signos de un número muy grande de determinantes. Pero si el rango de A es conocido, en particular si A es inversible, el número necesario de determinantes a chequear puede ser considerablemente reducido. La prueba de este resultado, que es bastante complicada, se posterga a un Apéndice al final del Capítulo. Esto se debe a que su aplicación es clave en el desarrollo de la teoría y la construcción de ejemplos, pero estas aplicaciones son de un grado mucho menor de dificultad. Una ves apreciado el efecto devastador del siguiente Teorema, es probable que el lector afronte con mayor entusiasmo la difícil lectura de su demostración.

Definición 13.1.3. Sea $n \in \mathbb{N}$, $k \in \mathbb{I}_n$ y $\alpha \in Q_{k,n}$. La **dispersión** de α es el número
$$d(\alpha) = \alpha_k - \alpha_1 - (k-1) = \sum_{i \in \mathbb{I}_{k-1}} \alpha_{i+1} - \alpha_i - 1 \ ,$$
con la convención de que $d(\alpha) = 0$ para los $\alpha \in Q_{1,n}$. Observar que $d(\alpha) = 0$ si y sólo si las entradas de α son consecutivas, i.e. $\alpha_{i+1} = \alpha_i + 1$ para todo $i \in \mathbb{I}_{k-1}$. ▲

Teorema 13.1.4. Sea $A \in \mathcal{M}_{n,m}(\mathbb{R})$ con $\operatorname{rk} A = r$, y sea ε una sucesión de signatura.

1. Para que A sea ε-RS, es suficiente que las Eqs. (13.2) o (13.3) se verifiquen en los casos en que $d(\beta) \leq m - r$.

2. En particular, si las Eqs. (13.5) o (13.6) son válidas en esos casos, A es TP. ■

Ahora pasemos a los criterios para la regularidad de signo estricta. El número de determinantes se reduce aún más. La prueba de este criterio también se dará en el Apéndice.

Teorema 13.1.5. Sean $A \in \mathcal{M}_{n,m}(\mathbb{R})$ y ε una sucesión de signatura.

1. Para que A sea ε-ERS es suficiente que, para todo $k \in \mathbb{I}_{\text{mín}(n,m)}$,
$$\varepsilon_k \det A[\alpha|\beta] > 0 \qquad \text{para} \qquad \alpha \in Q_{k,n}, \ \beta \in Q_{k,m} \qquad \text{tales que} \qquad d(\alpha) = d(\beta) = 0 \ .$$

2. En particular A es ETP si

$$\det A[\alpha|\beta] > 0 \quad \text{para} \quad \alpha \in Q_{k,n},\ \beta \in Q_{k,m} \quad \text{tales que} \quad d(\alpha) = d(\beta) = 0 \ . \qquad \blacksquare$$

Ejemplo 13.1.6 (Vandermonde)**.** Sea $t = (t_1, \dots, t_n) \in \mathbb{R}^n$. Se llama *matriz de Vandermonde* de t a $V(t) \in \mathcal{M}_n(\mathbb{R})$ dada por $V(t)_{ij} = t_i^{j-1}$. O sea que

$$V(t) = \begin{bmatrix} 1 & t_1 & \dots & t_1^{n-1} \\ 1 & t_2 & \dots & t_2^{n-1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & t_n & \dots & t_n^{n-1} \end{bmatrix} . \quad \text{Es conocido que} \qquad \det V(t) = \prod_{i<j} (t_j - t_i) \ . \qquad (13.7)$$

La prueba es un ejercicio tradicional de inducción (ver Ejercicio 7.5.12). Supongamos que $0 < t_1 < \dots < t_n$. Entonces $V(t)$ es ETP.

En efecto, observar en principio que $V(t) \in \mathcal{G}l(n)$. Luego, para probar que $V(t)$ es ETP, el Teorema 13.1.5 nos dice que basta ver que $\det V(t)[\alpha \,|\, \beta] > 0$ para los pares α, $\beta \in Q_{k,n}$, tales que $d(\alpha) = d(\beta) = 0$. Si $\beta = (r+1, r+2, \dots, r+k)$ y llamamos $t_\alpha = (t_{\alpha_1}, \dots, t_{\alpha_k})$, entonces se ve fácilmente que

$$\det V(t)[\alpha \,|\, \beta] = \prod_{i=1}^{k} t_{\alpha_i}^r \ \cdot \ \det V(t_\alpha) > 0 \ .$$

El argumento clave es que, gracias a que $d(\beta) = 0$, la submatriz $V(t)[\alpha \,|\, \beta]$ tiene en sus filas, potencias **consecutivas** de los t_{α_i} , por lo que, dividiendo a cada fila por $V(t)[\alpha \,|\, \beta]_{i,1} = t_{\alpha_i}^r$, se obtiene la matriz $V(t_\alpha)$ que es tambien una matriz de Vandermonde de una k-upla ordenada. La positividad del determinante en este caso se deduce de la fórmula (13.7). Observar que la ETPcidad se mantendría si uno inventara matrices de Vandermonde rectangulares (donde no coincidan necesariamente el número de potencias y de números t_i) pero siempre pidiendo que el vector t tenga entradas estrictamente crecientes. $\blacktriangle$

Corolario 13.1.7. *Una matriz $A \in \mathcal{G}l(n)$ triangular inferior es TP si* $\det A[\alpha|1, 2, \dots, k] \geq 0$ *para cada $k \in \mathbb{I}_n$ y cada $\alpha \in Q_{k,n}$.*

Demostración. Sea A triangular inferior. Como el $\operatorname{rk} A = n$, de acuerdo al Teorema 13.1.4, basta mostrar que $\det A[\alpha|\beta] \geq 0$, para $\alpha, \beta \in Q_{k,n}$, con $d(\beta) = 0$. Si $\alpha_1 < \beta_1$, entonces $\det A[\alpha|\beta] = 0$ por ser A triangular inferior. Si $\alpha_1 \geq \beta_1$, sea $\tau = \{1, 2, \dots, \beta_1 - 1\}$. Por ser A triangular inferior, es claro que $A[\tau|\beta] \equiv 0$. Entonces, por hipótesis,

$$\begin{aligned} 0 \ &\leq \ \det A[\alpha \cup \tau | 1, 2, \dots, \beta_k] = \det A[\alpha \cup \tau | \beta \cup \tau] \\ &= \ \det A[\tau] \det A[\alpha|\beta] = \prod_{i=1}^{\beta_1 - 1} a_{ii} \ \det A[\alpha|\beta]. \end{aligned}$$

Aplicando la hipótesis a los conjuntos $\alpha = \mathbb{I}_r$, se obtiene que $\prod_{i=1}^{r} a_{ii} \geq 0$ para todo $r \in \mathbb{I}_n$. Pero como $\det A = \prod_{i=1}^{n} a_{ii} \neq 0$, se sigue que $\det A[\alpha|\beta] \geq 0$. $\blacksquare$

La prueba del Teorema 13.1.5, combinada con el Corolario 13.1.7, genera el siguiente criterio alternativo:

Corolario 13.1.8. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ triangular inferior. Entonces es TP si se verifica que* $\det A[\alpha|1,2,\dots,k] > 0$ *para cada $k \in \mathbb{I}_n$ y cada $\alpha \in Q_{k,n}$, con $d(\alpha) = 0$.*

Demostración. Ejercicio (hace falta ver la prueba del Teorema 13.1.5). ■

Definición 13.1.9. Una matriz $A \in \mathcal{M}_n(\mathbb{C})$ es llamada una *matriz de Jacobi* (o *tridiagonal*) si $a_{ij} = 0$ siempre que $|i - j| > 1$. ▲

Teorema 13.1.10. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz de Jacobi. Supongamos que*

1. *$A \geqslant 0$.*
2. *Para todo $k \in \mathbb{I}_n$ y $\alpha \in Q_{k,n}$ tal que $d(\alpha) = 0$, se tiene que* $\det A[\alpha] \geq 0$.

Entonces A es TP y, para cualquier $t = (t_1, t_2, \dots, t_n) \in {\mathbb{R}_+^*}^n$,

$$\det\big(A + \operatorname{diag}(\,t\,)\;\big) \geq \det A + \prod_{i=1}^{n} t_i \ . \tag{13.8}$$

Demostración. Por inducción en n. La afirmación es trivial para $n = 1$. Supongmos que el Teorema es válido para $n - 1$ en lugar de n. Consideremos primero el caso en que $\det A > 0$. Por el Teorema 13.1.4, tenemos que chequear

$$\det A[\alpha|\beta] \geq 0 \quad , \qquad \text{para} \qquad \alpha, \beta \in Q_{k,n} \qquad \text{con} \qquad d(\beta) = 0.$$

Para $k = n$, esto es la hipótesis. Para $k \leq n - 1$, usaremos la hipótesis inductiva, que asegura que las matrices $A(1)$ y $A(n)$ son TP. Supongamos que $1 \notin \beta$. Si $1 \notin \alpha$, entonces $A[\alpha|\beta]$ es submatriz de $A(1)$ y $\det A[\alpha|\beta] \geq 0$. Si $1 \in \alpha$, entonces la primera fila de $A[\alpha|\beta]$ es $(a_{1,\beta_1}, 0, \dots, 0)$. Luego

$$\det A[\alpha|\beta] = a_{1,\beta_1} \det A\big[\{\alpha_2, \dots, \alpha_k\}|\{\beta_2, \dots, \beta_k\}\big] \geq 0 \ ,$$

porque la última matriz también vive dentro de $A(1)$. El análisis es similar si $1 \in \beta$, porque en tal caso, como $d(\beta) = 0$, debe verificarse que $n \notin \beta$, y puede usarse que $A(n)$ es TP. Por lo anterior, deducimos que A es TP en este caso (i.e., $\det A > 0$).

Supongamos ahora que $a_{11} > 0$ (es fácil ver que este caso es suficiente). Veamos que, en tal caso, $A/\{1\}$ cumple las hipótesis del Teorema. En efecto, es fácil ver que $A/\{1\}$ difiere de $A(1)$ sólo en la entrada $2,2$ (es decir la 1,1 si la numeráramos de corrido). Por lo tanto, dado $\alpha \subseteq \{2, \dots, n\}$ con $d(\alpha) = 0$, si $2 \notin \alpha$ entonces $\det(A/\{1\}[\alpha]) = \det(A(1)[\alpha]) \geq 0$. Si $2 \in \alpha$, por la Eq. (12.15) se tiene que

$$\det(A/\{1\}[\alpha]) = \det(A/\{1\}[2, 3, \dots, k]) = \frac{\det A[1, 2, \dots, k]}{a_{11}} \geq 0 \ .$$

La hipótesis inductiva nos asegura que $A/\{1\}$ es TP y

$$\det\Big(A/\{1\} + \operatorname{diag}(t_2,\ldots,t_n)\Big) \ge \det A/\{1\} + \prod_{i=2}^{n} t_i \ .$$

Por lo tanto, por el Teorema 12.1.4 y el hecho de que $A \geqslant 0$, se tiene que

$$\begin{aligned}
\det\Big(A+\operatorname{diag}(t)\Big) &= (a_{11}+t_1)\det\Big((A+\operatorname{diag}(t)\,)/\{1\}\Big)\\
&= (a_{11}+t_1)\det\Big(A/\{1\}+\operatorname{diag}(t_2+\frac{a_{12}a_{21}}{a_{11}}-\frac{a_{12}a_{21}}{a_{11}+t_1},\ldots,t_n)\Big)\\
&\ge a_{11}\det A/\{1\}+\prod_{i=1}^{n} t_i = \det A + \prod_{i=1}^{n} t_i \ .
\end{aligned}$$

Resta ver que A es TP. Para ello basta observar que, para todo $\varepsilon > 0$, la matriz $A+\varepsilon I$ tiene $\det(A+\varepsilon I) \ge \varepsilon^n > 0$ y cumple las hipótesis del Teorema (para ambas cosas se usa la fórmula (13.8), que fue probada para A, y vale para los los $A[\alpha]$ con $d(\alpha)=0$ por la hipótesis inductiva). Luego $A+\varepsilon I$ es TP por el argumento del principio. Como estas matrices convergen a A, ella también debe ser TP. ■

Corolario 13.1.11. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ de Jacobi y TP. Entonces, dado $t \in \mathbb{R}^n_+$, se tiene que $A + \operatorname{diag}(t)$ es también TP.*

Demostración. Se sigue del Teorema 13.1.10, aplicado a las submatrices principales, que $A + \operatorname{diag}(t)$ es una matriz de Jacobi positiva con menores principales no negativos. ■

Corolario 13.1.12. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ de Jacobi tal que $A \geqslant 0$. Entonces exite $\xi \in \mathbb{R}$ tal que $\xi I + A$ es TP.*

Demostración. Ejercicio. ■

Concluímos esta sección con un teorema de aproximación de una matriz TP con otras ETPs.

Teorema 13.1.13. *Toda matriz ε-RS puede ser aproximadada arbitrariamente cerca por matrices ε-ERS con la misma signatura. En particular, toda matriz TP puede ser aproximada arbitrariamente cerca por matrices estrictamente TPs.*

Demostración. Sea $A \in \mathcal{M}_{n,m}$ una matriz ε-RS. Podemos asumir que $n = m$, considerando $[A, 0]$ o $[\frac{0}{A}]$ si es necesario. Como veremos en la Sección 8, existe una sucesión $\{G_p\}$ de matrices n-cuadradas ETPs tales que $G_p \xrightarrow[p\to\infty]{} I_n$. Ahora procedamos por inducción hacia atrás en $k = \operatorname{rk} A$. Notemos que la Eq. (7.17) implica

$$\varepsilon_i \cdot \Lambda^i(G_pAG_p) > 0 \quad \text{si} \quad i \le \operatorname{rk} A \quad \text{y} \quad \Lambda^i(G_pAG_p) = 0 \quad \text{si} \quad i > \operatorname{rk} A \ . \tag{13.9}$$

Esto se deduce de que, dadas matrices $X, Y, Z \in \mathcal{M}_n(\mathbb{R})$ tales que $X > 0$, $Z > 0$ pero $0 \ne Y \geqslant 0$, entonces $XYZ > 0$. Cuando $\operatorname{rk} A = n$, la afirmación se sigue inmediatamente de la Eq. (13.9). Asumamos que la afirmación es cierta para todas las matrices regulares de signo de rango $k+1$. Si $\operatorname{rk} A = k$, tomemos un p para el que $B := G_p A G_p$ está suficientemente cerca de A. De acuerdo a las ecuaciones (13.9) y (7.19), B tiene la propiedad

$$\varepsilon_i \det B[\alpha|\beta] > 0 \qquad \text{para} \qquad \alpha, \beta \in Q_{i,n} \quad i \in \mathbb{I}_k \ . \tag{13.10}$$

Sea

$$\delta = \min_{1 \le i \le k} \frac{\min \left\{ \, |\det B[\alpha|\beta]| : \alpha, \beta \in Q_{i,n} \, \right\}}{\max \left\{ \, |\det B[\omega|\tau]| : \omega, \tau \in Q_{i-1,n} \, \right\}} \ , \qquad \text{donde} \qquad \det B[\varnothing] = 1 \ .$$

Fijemos $0 < t < \delta$ y consideremos la matriz $C = B + t\varepsilon_k \varepsilon_{k+1} E_{11}$. Dados α, $\beta \in Q_{r,n}$, desarrollando por la primera columna, se tiene que

$$\det C[\alpha|\beta] = \det B[\alpha|\beta] + t\varepsilon_k\varepsilon_{k+1} \det B[\alpha \setminus \{1\}|\beta \setminus \{1\}] \quad \text{si} \quad 1 \in \alpha \cap \beta \ ,$$

$$\text{y} \qquad \det C[\alpha|\beta] = \det B[\alpha|\beta] \qquad \text{en otro caso} \ .$$

Para ver que C es ε-RS se consideran tres casos: submatrices de tamaños $r \le k$ (ahí se usa que $t < \delta$ y el sumando extra no puede cambiar signos), $r > k+1$ (ahí da todo cero porque $\operatorname{rk} B = \operatorname{rk} A = k$) o $r = k+1$. Por ejemplo, tomando α, $\beta \in Q_{k+1,n}$ tales que $1 \in \alpha \cap \beta$, se ve que $\varepsilon_{k+1} \det C[\alpha|\beta] > 0$, porque $\det B[\alpha|\beta] = 0$ pero $\varepsilon_k \det B[\alpha \setminus \{1\}|\beta \setminus \{1\}] > 0$, por la Eq. (13.10). En particular, esto muestra que $\operatorname{rk} C = k+1$. Para t chicos, C está suficientemente cerca de B, y por lo tanto de A. Ahora, por hipótesis inductiva C puede ser aproximada arbitrariamente cerca por matrices estrictamente regulares de signo con signatura ε. Esto completa la inducción. ■

13.2 Permanencia de la positividad total

Esta sección está dedicada a métodos canónicos de producción de matrices TP nuevas a partir de otras dadas. Es claro que si A es de ε-RS, también lo es su adjunta $A^* = A^T$.

Teorema 13.2.1. *Si $A \in \mathcal{M}_{n,m}(\mathbb{R})$ es ε_A-RS y $B \in \mathcal{M}_{m,l}(\mathbb{R})$ es ε_B-RS, entonces:*

1. *El producto AB es ε-RS, con $\varepsilon = \varepsilon_A \cdot \varepsilon_B$.*
2. *En este caso, AB se convierte en ε-ERS si*
   - *(a) A es ε_A-ERS y $\operatorname{rk} B = l$, o si*
   - *(b) $\operatorname{rk} A = n$ y B es ε_B-ERS.*
3. *Si A y B son ETP, tambien lo es AB.*

Demostración. Los ítems 1 y 3 son consecuencia inmediata de las ecuaciones (7.17) o (12.1) (Cauchy-Binnet). El ítem 2 se deduce de los siguientes hechos:

* Si $C > 0$ y $D \geqslant 0$ no tiene columnas nulas (y se puede multiplicar), entonces $CD > 0$.

* Si $\operatorname{rk} B = l$, las columnas de B son LI. Luego para todo $k \leq l$ y todo $\beta \in Q_{k,l}$, se tiene que $\operatorname{rk} B[-|\beta] = k$, por lo que debe existir un $\alpha \in Q_{k,m}$ tal que $\det A[\alpha|\beta] \neq 0$. ■

La suma de dos matrices TPs no es en general TP. Por lo tanto, es poco común que una matriz $A \in \mathcal{M}_n(\mathbb{R})$ genere un *semigrupo TP de un parámetro*. Esto significaría que e^{tA} sea TP para todo $t > 0$. La excepción la dan las matrices de Jacobi TP:

Teorema 13.2.2. *Sea $A \in \mathcal{M}_n(\mathbb{R})$. Son equivalentes:*

1. *e^{tA} es TP para todo $t > 0$.*
2. *$A = \xi I + B$ para algún $\xi \in \mathbb{R}$ y una matriz $B \in \mathcal{M}_n(\mathbb{R})$ que es de Jacobi y TP.*

Demostración. Supongamos primero que A es de la forma mencionada. Entonces, como

$$e^{tA} = e^{\xi t} e^{tB} = e^{\xi t} \lim_{p\to\infty} \left(I + \frac{tB}{p} \right)^p ,$$

por la Eq. (9.41), la positividad total de e^{tA} resulta del Teorema 13.2.1, ya que B es una matriz de Jacobi y es TP, así que $I + \frac{t}{p} B$ sigue siendo TP por el Corolario 13.1.11.

Supongamos recíprocamente que e^{tA} es TP para todo $t > 0$. Por el Corolario 13.1.12, basta mostrar que A es una matriz real de Jacobi con elementos no negativos fuera de la diagonal. Usando el desarrollo en serie $e^{tA} = \sum\limits_{k\in\mathbb{N}} \frac{t^k A^k}{k!}$, es fácil ver que

$$A = \lim_{t\to 0} \frac{e^{tA} - I}{t} \quad \text{o equivalentemente que} \quad \lim_{t\to 0} \frac{I + tA - e^{tA}}{t} = 0 . \tag{13.11}$$

Como $e^{tA} \geqslant 0$, esto muestra que todas las entradas no diagonales de A son no negativas. Veamos que $a_{ij} = 0$ si $|i - j| > 1$. Por ejemplo, si $i + 1 < j$, entonces

$$\det \Big(e^{tA}[i, i+1|i+1, j] \Big) \geq 0 \quad \text{para todo} \quad t > 0 ,$$

lo que, vía la Eq. (13.11), implica que

$$0 \leq \lim_{t\to 0} \det \frac{I + tA}{t} \big[i, i+1|i+1, j\big] = \lim_{t\to 0} \{t a_{i,i+1} a_{i+1,j} - (1 + t a_{i+1,i+1}) a_{ij}\} = -a_{ij} .$$

El caso $j + 1 < i$ es análogo. Luego A es de Jacobi y $\rho I + A \geqslant 0$ para algún $\rho \in \mathbb{R}$. Para encontrar una B que sea TP, basta usar el Corolario 13.1.12. ■

Teorema 13.2.3. *Sea $A \in \mathcal{M}_{n,m}(\mathbb{R})$ ε-RS. Sean $\alpha \in Q_{k,n}$ y $\beta \in Q_{k,m}$. Entonces,*

1. *$A[\alpha|\beta]$ es ε-RS.*
2. *Si $n = m$, $d(\alpha) = 0$ y $A(\alpha) \in \mathcal{G}l(n-k)$, entonces A/α' es ε_α-RS, donde la sucesión de signos $\varepsilon_\alpha = (\varepsilon_{n-k}\ \varepsilon_{n-k+i})_{i\in\mathbb{I}_k}$.*

3. *Si $n = m$ y A es inversible, entonces $A^{\#} = J_n A^{-1} J_n$ es ε_J-RS, donde $\varepsilon_J = (\varepsilon_n \varepsilon_{n-i})_{i\in\mathbb{N}}$, con la convención de que $\varepsilon_j = 1$ si $j \le 0$.*

4. *En particular, si A es TP, también lo serán $A[\alpha|\beta]$, A/α' y $J_n A^{-1} J_n$.*

Además, los mismos resultados valen reemplazando "regular de signo" por "estrictamente regular de signo" (o TP por ETP) en todos los casos.

Demostración. Fijemos $\omega,\ \tau \in Q_{p,n}$ tales que $\tau \subseteq \beta$, $\omega \subseteq \alpha$.

1. Es trivial, ya que

$$\varepsilon_p \det A[\alpha|\beta][\omega|\tau] = \varepsilon_p \det A[\omega|\tau] \ge 0 \qquad (\text{resp. } > 0) \quad .$$

2. Supongamos ahora que $\tau \subseteq \alpha$. Se sigue de la la Eq. (12.15) que

$$\det\big(A/\alpha'\big)[\omega|\tau] = \operatorname{sgn}\ \frac{\alpha'}{\alpha' \cup \omega} \ \cdot\ \operatorname{sgn}\ \frac{\alpha'}{\alpha' \cup \tau} \ \cdot\ \frac{\det A[\alpha' \cup \omega|\alpha' \cup \tau]}{\det A[\alpha']} \ .$$

   Notar que $\det A[\alpha' \cup \omega|\alpha' \cup \tau]$ tiene signo ε_{n-k+p} y $\det A(\alpha)$ tiene signo ε_{n-k}. Pero como $d(\alpha) = 0$, se ve fácilmente que $\operatorname{sgn}(\alpha'/\alpha' \cup \omega) = \operatorname{sgn}(\alpha'/\alpha' \cup \tau)$ (ambos dependen sólo de cuantos elementos de α' están después del bloque α).

3. Observar que $\varepsilon_n \det A > 0$ y, por la Eq. (12.11), tenemos que

$$\det\big(J_n A^{-1} J_n\big)\,[\alpha|\beta] = \frac{\det A(\beta|\alpha)}{\det A} \ ,$$

   donde $\varepsilon_{n-k} \det A(\beta|\alpha) = \varepsilon_{n-k} \det A[\beta'|\alpha'] \ge 0$ (resp > 0).

Las últimas afirmaciones se deducen de lo anterior. ■

En lo que sigue, se usará varias veces el Corolario 12.2.6, cuya fórmula repasaremos para comodidad del lector: Sea $A \in \mathcal{M}_n(\mathbb{R})$ y sea $r \in \mathbb{I}_n$ tal que $A_{rr} \ne 0$. Entonces, para todo $\alpha \in Q_{k,n}$ tal que $r \notin \alpha$, se tiene que

$$(A/[r])[\alpha] \ = A[\{r\} \cup \alpha]/[r] \ . \tag{13.12}$$

Proposición 13.2.4 (Pinching)**.** *Sea $B \in \mathcal{M}_m(\mathbb{R})$ una matriz TP. Entonces*

$$\det B \le \ \det B[\mathbb{I}_k] \det B(\mathbb{I}_k) \qquad \textit{para todo} \qquad k \in \mathbb{I}_{m-1} \ . \tag{13.13}$$

Demostración. Probaremos la Eq. (13.13) por inducción en m. Para $m = 2$, tenemos que

$$b_{12} \ge 0 \quad \text{ y } \quad b_{21} \ge 0 \implies \det B = b_{11}b_{22} - b_{12}b_{21} \le b_{11}b_{22} \ .$$

Asumamos que la afirmación es cierta para todos los casos de orden menor que m. Asumamos que $k > 1$ y que $b_{11} > 0$. Por la Eq. (12.5) se tiene que

$$\det B[\mathbb{I}_k] \ \det B(\mathbb{I}_k) = b_{11} \det B[\mathbb{I}_k]/\{1\} \ \det B(\mathbb{I}_k) = b_{11} \det B/\{1\}[2,\dots,k] \ \det B(\mathbb{I}_k) \ ,$$

donde la última igualdad se deduce de la Eq. (13.12), que decía que

$$B/\{1\}[\alpha] = B[\{1\} \cup \alpha]/\{1\} \qquad \text{para todo } \alpha \text{ tal que } 1 \notin \alpha \qquad .$$

Como la matriz $B[1, k{+}1, k{+}2, \dots, m]$ es de orden menor que m y es TP, la hipótesis inductiva nos asegura que

$$\begin{aligned} b_{11} \det B(\mathbb{I}_k) &\geq \det B[1, k+1, k+2, \dots, m] \\ &= b_{11} \det B[1, k+1, k+2, \dots m]/\{1\} \\ &= b_{11} \det B/\{1\}(\mathbb{I}_k) \ , \end{aligned}$$

Usando nuevamente la hipótesis inductiva en la matriz $B/\{1\}$ que es de orden $m-1$, y es TP por el Teorema 13.2.3, obtenemos

$$\begin{aligned} \det B[\mathbb{I}_k] \ \det B(\mathbb{I}_k) &= b_{11} \det \Big(B/\{1\}[2,\dots,k] \Big) \ \det B(\mathbb{I}_k) \\ &\geq b_{11} \det \Big(B/\{1\}[2,\dots,k] \Big) \ \det B/\{1\}(\mathbb{I}_k) \\ &\geq b_{11} \det \Big(B/\{1\}[2,\dots,m] \Big) = b_{11} \det B/\{1\} = \det B \ . \end{aligned}$$

Si $k = 1$, asumiendo ahora que $b_{mm} > 0$, tenemos que

$$\det B = b_{mm} \ \det(B/\{m\}) \leq b_{mm} \det(B/\{m\}[1]) \det(B/\{m\}[2,\dots,m-1])$$

Ahora, $\det(B/\{m\}[1]) = b_{11} - \dfrac{b_{1m}b_{m1}}{b_{mm}} \leq b_{11}$ por ser $B \geqslant 0$. Además, por la Eq. (13.12), tenemos que $B/\{m\}[2,\dots,m-1] = B[2,\dots,m]/\{m\} = B(1)/\{m\}$. Así,

$$\det B \quad \leq b_{11} b_{mm} \det B(1)/\{m\} = b_{11} \det B(1) = \det B[1] \det B(1) \ .$$

Los casos en que $b_{11} = 0$ o $b_{mm} = 0$ se pueden anlizar a mano (mostrando que en tal caso $\det B = 0$ por tener una fila o columna nula) o bien cambiando b_{11} por $b_{11} + \frac{1}{n}$ (con lo que B sigue siendo TP) y tomando límite. ■

Corolario 13.2.5. *Si $A \in \mathcal{G}l(n)$ es TP, entonces*

$$\det A[\alpha] > 0 \qquad \textit{para cada} \qquad k \in \mathbb{I}_n \qquad \textit{y cada} \qquad \alpha \in Q_{k,n} \ . \tag{13.14}$$

En particular, en este caso siempre existen los complementos de Schur $A/[\alpha]$.

Demostración. Por inducción en n. El caso $n = 1$ es trivial. Asumamos que la afirmación vale para $n - 1$. Por la Proposición 13.2.4, tenemos que $0 < \det A \leq a_{11} \det A(1)$. Luego $A(1)$ es TP e inversible y $a_{11} > 0$. Si $\alpha_1 > 1$, entonces $\alpha \subseteq \mathbb{I}_n \setminus \{1\}$ y la desigualdad $\det A[\alpha] > 0$ se sigue de la hipótesis inductiva aplicada a $A(1)$. Si $\alpha_1 = 1$, la Eq. (13.12) aplicada a $\alpha \setminus \{1\}$ con $r = 1$, muestra que

$$\det A[\alpha] = a_{11} \det A[\alpha]/\{1\} = a_{11} \det \ A/\{1\}[\alpha \setminus \{1\}] \ .$$

Entonces la desigualdad $\det A[\alpha] > 0$ se deduce de la hipótesis inductiva aplicada a $A/\{1\}$, que es inversible por el Teorema 12.1.4 y es TP por el Teorema 13.2.3. ■

13.3 Factorizaciones LU y UL

Una factorización $A = BC$ es llamada una *LU-factorización* (resp, UL-factorización) si B (resp. C) es triangular inferior y C (resp. B) es triangular superior.

Teorema 13.3.1. *Sea $A \in \mathcal{M}_{n,m}(\mathbb{R})$ TP con $n \geq m$. Entonces A admite una LU-factorización $A = A_L A_U$ y una UL-factorización $A = \tilde{A}_L \tilde{A}_U$, donde las matrices triangulares $A_L \in \mathcal{M}_n(\mathbb{R})$, $\tilde{A}_U \in \mathcal{M}_m(\mathbb{R})$ y A_U, $\tilde{A}_L \in \mathcal{M}_{n,m}(\mathbb{R})$ son todas TPs.*

Para demostrar el Teorema necesitamos dos largos Lemas técnicos: Para ellos necesitaremos una notación nueva: Dados $x_1, \dots, x_m \in \mathbb{R}^n$, denotaremos por $X = [x_1, \dots, x_m] \in \mathcal{M}_{nm}(\mathbb{R})$ a la matriz cuyas columnas están dadas por $C_i(X) = x_i$, para todo $i \in \mathbb{I}_m$.

Lema 13.3.2. *Sea $A = (a_{ij})_{i,j\in\mathbb{I}_n} = [a_1, \dots, a_n] \in \mathcal{M}_n(\mathbb{R})$ una matriz TP. Si $a_{1k} \neq 0$, entonces también es TP la matriz $B = [b_1, \dots, b_n] \in \mathcal{M}_n(\mathbb{R})$ definida por*

$$b_i = a_i \quad \text{si} \quad i \in \mathbb{I}_k \quad \text{y} \quad b_i = a_i - \frac{a_{1i}}{a_{1k}}\, a_k\ , \quad \text{si} \quad i \in \mathbb{I}_n \setminus \mathbb{I}_k\ .$$

Demostración. Por el Teorema 13.1.13 podemos asumir que $\det A > 0$. Como obviamente $\det A = \det B$, de acuerdo al Teorema 13.1.4 basta mostrar que

$$b_i \wedge b_{i+1} \wedge \cdots \wedge b_j \geqslant 0 \qquad \text{para} \qquad 1 \leq i \leq j \leq n, \tag{13.15}$$

i.e., la positividad para los α tales que $d(\alpha) = 0$. Si $j \leq k$ o $i \leq k \leq j$, entonces

$$b_i \wedge b_{i+1} \wedge \cdots \wedge b_j = a_i \wedge a_{i+1} \wedge \cdots \wedge a_j\ ,$$

y la Eq. (13.15) es válida porque A es TP. Si $k < i$, consideremos la matriz

$$C = [a_k, a_{k+1}, \dots, a_n, 0, \dots, 0] \in \mathcal{M}_n(\mathbb{R})\ ,$$

que es TP, por serlo A. Se ve fácilmente de la definición de $C/\{1\}$ que

$$M = [b_{k+1}, b_{k+2}, \dots, b_n, 0, \dots, 0] = \begin{bmatrix} 0 \\ C/\{1\} \end{bmatrix} \in \mathcal{M}_{n,\,n-1}(\mathbb{R})\ .$$

En efecto, observar que todas las primeras coordenadas de los b_j $(j \in \mathbb{I}_n \setminus \mathbb{I}_k)$ son nulas, por lo que la primera fila va bien. Si $j \in \mathbb{I}_{n-k}$ e $i \in \mathbb{I}_{n-1}$, entonces

$$(C/\{1\})_{ij} = c_{i+1\,,\,j+1} - \frac{c_{i+1\,,\,1}\, c_{1\,,\,j+1}}{c_{11}} = a_{i+1,\ k+j} - \frac{a_{1,\ k+j}}{a_{1\,,\,k}}\, a_{i+1,\ k} = \big(b_{k+j}\big)_{i+1} = M_{i+1\,,\,j}\ .$$

En los demás casos, debe ser $(C/\{1\})_{ij} = 0$ (y la otra también). Ahora la ecuación (13.15) se deduce de la positividad total de $C/\{1\}$ y de M, garantizada por el Teorema 13.2.3 ∎

Lema 13.3.3. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz TP. Entonces existen C y $S \in \mathcal{M}_n(\mathbb{R})$, ambas TP, tales que $C[1,1] \equiv 0$ (i.e., $F_1(C) = c_{11}\, e_1$), S es triangular superior y $A = CS$.*

Demostración. Para cada $j \in \mathbb{I}_{n-1}$, consederemos la matriz

$$T_j := [e_1 , e_2 , \dots , e_{j-1} , 0, e_j , e_{j+1} , \dots , e_{n-1}] = \begin{bmatrix} I_{j-1} & 0 \\ 0 & N_{n-j+1} \end{bmatrix} \in \mathcal{M}_n(\mathbb{R}) \ ,$$

donde $N_{n-j+1} \in \mathcal{M}_{n-j+1}(\mathbb{R})$ es el bloque de Jordan con los unos arriba (Jacobi nos robó la J). Observar que $T_1 = N_n$ el bloque de Jordan tutti. Cada T_j es una matriz de Jacobi positiva y triangular superior. Luego las T_j son TP por el Teorema 13.1.10.

Estas matrices nos permitirán "correr" hacia la izquierda las columnas no nulas de A. Si $a_1 = a_2 = \dots = a_{k-1} = 0$ pero $a_k \neq 0$, entonces

$$A = [a_k, a_{k+1}, \dots , a_n, 0, \dots , 0] \ T_1^{k-1},$$

y la matriz $A_1 = [a_k, a_{k+1}, \dots , a_n, 0, \dots , 0]$ es TP. Si seguimos sin columnas nulas hasta a_{k+p} y a_{k+p+r} es la primera columna no nula de las que quedan, como antes obtenemos que

$$[a_{k+p+1}, \dots , a_n, 0, \dots , 0] = [a_{k+p+r}, \dots , a_n, 0, \dots , 0] \ T_1^{r-1}$$

o, mirando como dan las cuentas,

$$A_1 = [a_k, \dots , a_{k+p}, a_{k+p+r}, \dots , a_n, 0, \dots , 0] \ T_{p+2}^{r-1} \ .$$

Es decir,

$$A = [a_k, \dots , a_{k+p}, a_{k+p+r}, \dots , a_n, 0, \dots , 0] \ T_{p+2}^{r-1} \ T_1^{k-1} \ .$$

Observar que todo queda TP, y $T_{p+2}^{r-1} \ T_1^{k-1}$ es triangular superior. Aplicando entonces este procedimiento finitas veces, obtenemos que $A = BT$, donde T triangular superior y TP y $B = [b_1, b_2, \dots , b_n]$ es una matriz TP tal que $b_i = 0$ implica que $b_j = 0$, para $j > i$. Si $B[1|1) \neq 0$, tomemos el mayor i para el cual $b_{1i} \neq 0$. Afirmamos que $b_{1,i-1} \neq 0$. En efecto, si $b_{1,i-1} = 0$, para todo $j \in \mathbb{I}_n$, tendríamos

$$\det B[1, j|i-1, i] = b_{1i-1}b_{ji} - b_{1i}b_{ji-1} = -b_{1i}b_{ji-1} \geq 0 \ ,$$

lo que implicaría que todos los $b_{j,i-1} = 0$, o sea $b_{i-1} = 0$, lo que contradice que $b_i \neq 0$. Entonces B admite una factorización $B = DU_1$, donde

$$D := [b_1, \dots , b_{i-1}, b_i - \frac{b_{1,i}}{b_{1,i-1}} \ b_{i-1}, b_{i+1}, \dots , b_n] \quad \text{y}$$

$$U_1 := [e_1, \dots , e_{i-1}, \frac{b_{1,i}}{b_{1,i-1}} \ e_{i-1} + e_i, e_{i+1}, \dots , e_n] \ .$$

Notemos que ahora $D_{1i} = 0$. Por otro lado, U_1 es una matriz de Jacobi positiva, triangular superior, y por lo tanto TP por el Teorema 13.1.10. La positividad total de D se sigue del Lema 13.3.2, porque, como $b_{1,j} = 0$ si $j > i$, entonces

$$b_j = b_j - \frac{b_{1,j}}{b_{1,i-1}} \ b_{i-1} \ , \qquad \text{para} \qquad j = i+1, i+2, \dots , n \ .$$

Repitiendo este procedimiento, llegamos a una factorizacióin $B = CU_pU_{p-1}\cdots U_1$, donde cada U_i es triangular superior y TP mientras que C es una matriz TP tal que $C[1|1) = 0$, como se buscaba. Ahora basta tomar $S = U_pU_{p-1}\cdots U_1T$ que es como se pedía. ■

Demostración del Teorema 13.3.1: Considerando la matriz $[A, 0] \in \mathcal{M}_n(\mathbb{R})$, podemos confinar la prueba al caso $n = m$. Además, usando conversiones, basta tratar sólo la factorización LU. Cuando $n = 1$, es trivial. Asumamos que la afirmación es cierta para $n-1$ en lugar de n. Para conseguir hacer el paso inductivo, alcanzaría con probar que existen R, F y $S \in \mathcal{M}_n(\mathbb{R})$, todas TP, con S es triangular superior, R es triangular inferior y

$$F = \begin{bmatrix} f_{11} & 0 \\ 0 & F(1) \end{bmatrix} , \quad \text{tales que} \quad A = RFS ,$$

porque en tal caso se factoriza $F(1)$ por hipótesis inductiva, y se agrandan las matrices triangulares (y TP) de $\mathcal{M}_{n-1}(\mathbb{R})$ obtenidas, poniéndoles $(f_{11})^{1/2}$ en el lugar $1, 1$. Recordar que producto de triangulares es triangular, y lo mismo con TPs (por el Teorema 13.2.1).

Pero por el Lema 13.3.3, existen S como antes y $C \in \mathcal{M}_n(\mathbb{R})$ tal que C es TP, $C[1|1) \equiv 0$ y $A = CS$. Y por el mismo Lema, existen R como arriba y $F \in \mathcal{M}_n(\mathbb{R})$ tal que F es TP, $F(1|1] \equiv 0$ y $C^T = F^T R^T$. Observar que multiplicar por una triangular superior a derecha, solo cambia a la primera columna multiplicándola por un escalar, asi que F hereda lo bueno que tenía C (ceros en la primera fila) pero gana ceros en la primera columna, quedando como queríamos. ■

Corolario 13.3.4. *Toda $A \in \mathcal{G}l(n)$ triangular superior (inferior) y TP es producto de un cierto número de matrices de Jacobi triangulares superiores (inferiores) y TPs.*

Demostración. Por inducción en n. El caso $n = 1$ es trivial. Asumamos que la afirmación es cierta para $n-1$, y sea $A \in \mathcal{G}l(n)$ triangular superior (inferior) y TP. Por el Lema 13.3.3 (y su prueba), tenemos una factorización $A = DS$ con S un producto de matrices de Jacobi triangulares superiores, todas ellas TP, y D también TP, con $D[1, 1) \equiv 0$. Pero como A es triangular superior, también se da que $D(1, 1] \equiv 0$. Luego

$$D = \begin{bmatrix} d_{11} & 0 \\ 0 & D(1) \end{bmatrix}$$

y $D(1) \in \mathcal{G}l(n-1)$ es totalmente positva y triangular superior. Por hipótesis inductiva tenemos $D(1) = \hat{W}_1\hat{W}_2\cdots\hat{W}_s$ para algunas matrices de Jacobi TPs y triangulares superiores $\hat{W}_i \in \mathcal{G}l(n-1)$, $i \in \mathbb{I}_s$. Sean

$$W_i = \begin{bmatrix} d_{11}^{1/2} & 0 \\ 0 & \hat{W}_i \end{bmatrix} , \quad i \in \mathbb{I}_s .$$

Entonces $A = W_1W_2\cdots W_sS$ es una factorización como la buscada. ■

Definición 13.3.5. Además de la relación de orden usual $A \geqslant B$ entre matrices en $\mathcal{M}_n(\mathbb{R})$, introduzcamos una más fuerte: Diremos que $A \overset{t}{\geqslant} B$ si $\Lambda^k A \geqslant \Lambda^k B$ para todo $k \in \mathbb{N}$. En otras palabras, si

$$\det A[\alpha|\beta] \geq \det B[\alpha|\beta] \quad \text{para todo} \quad k \in \mathbb{I}_n \quad \text{y} \quad \alpha, \beta \in Q_{k,n} . \tag{13.16}$$

En esta notación, $A \overset{t}{\geqslant} 0$ significa que A es TP. ▲

Observación 13.3.6. Es claro que la relación $A \overset{t}{\geqslant} B$ implica que $A[\alpha|\beta] \overset{t}{\geqslant} B[\alpha|\beta]$ para cualquier par $\alpha, \beta \in Q_{k,n}$, pero no que $A - B \overset{t}{\geqslant} 0$, como puede verse con las matrices

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Tampoco $A \overset{t}{\geqslant} B \overset{t}{\geqslant} 0$ implica que $A/\{1\} \overset{t}{\geqslant} B/\{1\}$ o que $J_n B^{-1} J_n \overset{t}{\geqslant} J_n A^{-1} J_n$. ▲

Teorema 13.3.7. *Si $A \in \mathcal{M}_n(\mathbb{R})$ es TP, y $\alpha = \mathbb{I}_k$ o $\alpha = \{k, k+1, \dots, n\}$, entonces, si $A(\alpha)$ es inversible, se cumple que*

$$A[\alpha] \overset{t}{\geqslant} A/\alpha' \ . \tag{13.17}$$

Demostración. Prueba para el caso $\alpha = \mathbb{I}_k$: Fijemos $\omega, \tau \in Q_{l,n}$ tales que $\omega \subseteq \alpha$ y $\tau \subseteq \alpha$. Deberíamos probar que $\det \big(A[\alpha]\big)[\omega|\tau] \geq \det A/\alpha'[\omega|\tau]$. Observar que la Eq. (12.15) nos dice que

$$\det A/\alpha' \ [\omega|\tau] \ = \ \frac{A[\alpha' \cup \omega|\alpha' \cup \tau]}{\det A(\alpha)} \ ,$$

porque los dos signos que intervienen en la Eq. (12.15) son iguales en este caso, por ser α' quien es. Por lo tanto, bastaría probar que

$$\det A[\omega \cup \alpha'|\tau \cup \alpha'] \leq \det \big(A[\alpha]\big)[\omega|\tau] \ \cdot \ \det A(\alpha) = \det A[\omega|\tau] \ \cdot \ \det A(\alpha) \ .$$

Consideremos $A[\omega \cup \alpha'|\tau \cup \alpha'] \in \mathcal{M}_{n-k+l}(\mathbb{R})$ con su numeración de corrido en $\mathbb{I}_{n-k+l}$. Como $\omega, \tau \subseteq \alpha = \mathbb{I}_k$, entonces

$$A[\omega|\tau] = A[\alpha' \cup \omega|\alpha' \cup \tau] \ [1, \dots, l] \quad \text{y} \quad A(\alpha) = A[\alpha' \cup \omega|\alpha' \cup \tau] \ [l+1, \dots, n-k+l] \ ,$$

y el resultado se deduce de la Proposición 13.2.4. ■

Las factorizaciones LU y UL en el Teorema 13.3.1 dan lugar a otras desigualdades.

Teorema 13.3.8. *Si $A \in \mathcal{M}_n(\mathbb{R})$ es TP, y $\alpha = \mathbb{I}_k$ o $\alpha = \mathbb{I}_n \setminus \mathbb{I}_k$, entonces, si $A(\alpha)$ es inversible,*

$$A[\alpha] - A/\alpha' \overset{t}{\geqslant} 0 \tag{13.18}$$

Demostración. Prueba para el caso $\alpha = \mathbb{I}_n \setminus \mathbb{I}_k = \{k+1, k+2, \dots, n\}$: Sea $A = A_L A_U$ una factorización LU con A_L y A_U TP, garantizada por el Teorema 13.3.1. Entonces, por las propiedades de A_L y A_U, se tiene que $A(\alpha) = A_L(\alpha) A_U(\alpha)$ (por lo que ambas submatrices son inversibles), $A(\alpha|\alpha] = A_L(\alpha) A_U(\alpha|\alpha]$ y $A[\alpha|\alpha) = A_L[\alpha|\alpha) A_U(\alpha)$. Por lo tanto,

$$\begin{aligned} A[\alpha] - A/\alpha' &= A[\alpha|\alpha) A(\alpha)^{-1} A(\alpha|\alpha] \\ &= A_L[\alpha|\alpha) A_U(\alpha) (A_L(\alpha) A_U(\alpha))^{-1} A_L(\alpha) A_U(\alpha|\alpha] \\ &= A_L[\alpha|\alpha) A_U(\alpha|\alpha]. \end{aligned}$$

Como $A_L[\alpha|\alpha)$ y $A_U(\alpha|\alpha]$ son TPs, también lo es su producto $A_L[\alpha|\alpha) A_U(\alpha|\alpha]$. La prueba para el caso $\alpha = \mathbb{I}_k$ se hace usando una factoriación UL. ■

13.4 Matrices oscilatorias

Una matriz $A \in \mathcal{M}_n(\mathbb{R})$ se dice *oscilatoria* (abreviaremos OSC) si es TP y una cierta potencia A^p es ETP. Las OSC juegan el rol de las primitivas en el Capítulo 11. En esta sección presentaremos un criterio simple para que una matriz TP sea OSC. Observemos que una matriz OSC es inversible, y su adjunta es también OSC. por lo tanto, por el Corolario 13.2.5, si $A \in \mathcal{M}_n(\mathbb{R})$ es OSC, entonces se tiene que $\det A[\alpha] > 0$ para todo $\alpha \in Q_{k,n}$.

Teorema 13.4.1. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz OSC. Entonces*

1. *$A^{\#} = J_n A^{-1} J_n$ es OSC.*
2. *$A[\alpha]$ y A/α' son OSC's para cada $\alpha \in Q_{k,n}$ tal que $d(\alpha) = 0$.*

Demostración. Supongamos que A es TP y A^p es ETP.

1. El Teorema 13.2.3 asegura que $J_n A^{-1} J_n$ es TP y que $(J_n A^{-1} J_n)^p = J_n (A^p)^{-1} J_n$ es ETP. Así, $J_n A^{-1} J_n$ es OSC.
2. Probemos primero que $A[\alpha]$ es OSC para el caso $\alpha = \mathbb{I}_{n-1}$. Sea $B = A[\mathbb{I}_{n-1}] = A(n)$. Tomemos $\beta, \tau \in Q_{k,n-1}$, y sean $\mu = \beta \cup \{n\}$ y $\nu = \tau \cup \{n\}$. Por la fórmula de Cauchy-Binnet (12.1), el hecho de que $\det A^p[\mu|\nu] > 0$ implica que existe una sucesión

$$\left(\omega^{(i)}\right)_{i=0}^{p} \quad \text{en} \quad Q_{k+1,n} \quad \text{tal que } \omega^0 = \mu, \ \omega^{(p)} = \nu \quad , \quad \text{y} \qquad \prod_{i=1}^{p} \det A[\omega^{(i-1)}|\omega^{(i)}] > 0 \ .$$

Llamemos $\tilde{\omega}^{(i)} \in Q_{k,n-1}$ a la k-upla obtenida eliminando la última componente de $\omega^{(i)}$. Como $A[\omega^{(i-1)}|\omega^{(i)}]$ es TP con determinante positivo, por la Eq. (13.13) se tiene que

$$\det B[\tilde{\omega}^{(i-1)}|\tilde{\omega}^{(i)}] > 0 \ , \qquad \text{para todo} \qquad i \in \mathbb{I}_p \ .$$

Entonces, nuevamente por la positividad total de B y la Eq. (12.1),

$$\det B^p[\beta|\tau] \geq \prod_{i=1}^{p} \det B[\tilde{\omega}^{(i-1)}|\tilde{\omega}^{(i)}] > 0 \ ,$$

lo que prueba que B^p es ETP. El caso $A[2, 3, \ldots, n]$ se trata de manera análoga. El resto de los casos ($\alpha \in Q_{k,n}$ con $k < n-1$ y $d(\alpha) = 0$) ahora se pueden probar por inducción en n, dado que $\alpha \subseteq \mathbb{I}_{n-1}$ o $\alpha \subseteq \mathbb{I}_n \setminus \{1\}$.

Veamos que A/α' es OSC: Observar que $J_n A^{-1} J_n[\alpha] = J_k A^{-1}[\alpha] J_k$, dado que $d(\alpha) = 0$ (puede aparecer un $-J_k$, pero se cancela). Por los casos anteriores, sabemos que

$$J_n A^{-1} J_n \text{ es OSC} \quad \Longrightarrow \quad J_k A^{-1}[\alpha] J_k = J_n A^{-1} J_n[\alpha] \text{ es OSC}$$

$$\Longrightarrow \quad (A^{-1}[\alpha])^{-1} \text{ es OSC} \ .$$

Pero la Eq. (12.6) dice que $A/\alpha' = (A^{-1}[\alpha])^{-1}$. ∎

Lo siguiente da un criterio para la “oscilatoriedad”.

Teorema 13.4.2. *Sea* $A = (a_{ij})_{i,j\in \mathbb{I}_n} \in \mathcal{M}_n(\mathbb{R})$ *una matriz TP. Entonces*

$$A \text{ es OSC} \quad \Longleftrightarrow \quad A \in \mathcal{G}l(n) \ , \quad a_{i,\,i+1} > 0 \quad \text{y} \quad a_{i+1,\,i} > 0 \ , \quad \text{para todo} \quad i \in \mathbb{I}_{n-1} \ . \tag{13.19}$$

Demostración. Supongamos que A es OSC. Por el Teorema 13.4.1,

$$B := A[i, i+1] = \begin{bmatrix} a_{i,\,i} & a_{i,\,i+1} \\ a_{i+1,\,i} & a_{i+1,\,i+1} \end{bmatrix}$$

debe ser OSC. Luego $B^p > 0$ para algún p. Pero esto es posible sólo cuando $a_{i,i+1} > 0$ y $a_{i+1,i} > 0$, ya que si alguno de los dos se anulara, entonces B y todas sus potencias serían triangulares. La otra implicación será probada como consecuencia de un resultado más general, hacia el fin de la sección. ■

Corolario 13.4.3. *Sean* $A, B \in \mathcal{M}_n(\mathbb{R})$*, ambas TP. Si* A *es OSC y* B *es inversible, entonces* AB *y* BA *son OSC.*

Demostración. Observar que $A \in \mathcal{G}l(n)$, por se OSC. Luego AB y $BA \in \mathcal{G}l(n)$. Además, como B es inversible y TP, entonces $b_{ii} > 0$ para todo $i \in \mathbb{I}_n$ (Corolario 13.2.5). Por lo tanto AB y BA satisfacen la condición (13.19), ya que

$$(AB)_{i,\,i+1} = \sum_{j=1}^{n} a_{i,\,j}\, b_{j,\,i+1} \geq a_{i,\,i+1}\, b_{i+1\,,\,i+1} > 0 \quad \text{para todo} \quad i \in \mathbb{I}_{n-1} \ .$$

La prueba para $(AB)_{i+1\,,\,i}$ y las entradas correspondientes de BA es análoga. ■

El siguiente teorema presenta una extensión de la condición (13.19) para matrices OSCs.

Proposición 13.4.4. *Supongamos que* $A \in \mathcal{G}l(n)$ *y es TP. Si* A *satisface la Eq.* (13.19)*, entonces* $\det A[\alpha|\beta] > 0$ *para cada par* $\alpha, \beta \in Q_{k,n}$ *tal que*

$$|\alpha_i - \beta_i| \leq 1 \quad \text{y} \quad \max\{\alpha_i\,,\, \beta_i\} < \min\{\alpha_{i+1}\,,\, \beta_{i+1}\} \quad \text{para todo} \quad i \in \mathbb{I}_k \ , \tag{13.20}$$

donde usamos la convención $\alpha_{k+1} = \beta_{k+1} = n+1$.

Demostración. Haremos la prueba por inducción en k. El caso $k = 1$ se sigue del Corolario 13.2.5 (si $\alpha = \beta$) y la suposición (13.19). Fijemos $k > 1$, y supongamos que la afirmación es cierta para cada par en $Q_{k-1,n}$ que satisface las hipótesis. Tomemos un par $\alpha, \beta \in Q_{k,n}$ que cumpla la Eq. (13.20). Si $d(\alpha) = d(\beta) = 0$, entoces la Eq. (13.20) impica que $\alpha = \beta$. Luego $\det A[\alpha|\beta] = \det A[\alpha|\alpha] > 0$, nuevamente por el Corolario 13.2.5. Ahora, asumiendo que $d(\beta) > 0$, sea $B = A[\alpha|\beta] = [b_{\beta_1}, b_{\beta_2}, \dots b_{\beta_k}]$, donde cada $b_{\beta_i} \in \mathbb{R}^\alpha \cong \mathbb{R}^k$. Supongamos que $\det B = 0$. En principio, por la hipótesis inductiva, sabemos que

$$\det B[\alpha_1\,, \dots, \alpha_{k-1}|\beta_1\,, \dots, \beta_{k-1}] > 0 \quad \text{y} \quad \det B[\alpha_2\,, \dots, \alpha_k|\beta_2\,, \dots, \beta_k] > 0 \ .$$

Junto con la positividad total de B, esto implica que

$$0 \neq b_{\beta_1} \wedge b_{\beta_2} \wedge \ldots \wedge b_{\beta_{k-1}} \geqslant 0 \quad \text{y} \quad 0 \neq b_{\beta_2} \wedge b_{\beta_3} \wedge \ldots \wedge b_{\beta_k} \geqslant 0 \ . \tag{13.21}$$

Entonces el hecho de que $\det B = 0$ garantiza que $b_{\beta_k} \in \operatorname{Gen}\{b_{\beta_i} : i \in \mathbb{I}_{k-1}\}$, y además

$$b_{\beta_k} = \sum_{i=1}^{k-1} \xi_i \ b_{\beta_i} \quad \text{para ciertos } \xi_i \in \mathbb{R} \text{ , donde } \xi_1 \neq 0 \tag{13.22}$$

(sinó $\{b_{\beta_i} : i \in \mathbb{I}_k \setminus \{1\}\ \}$ sería LD). Sustituyamos b_{β_k} en la Eq. (13.21) para obetener

$$(-1)^{k-2} \xi_1 \, b_{\beta_1} \ \wedge b_{\beta_2} \ \wedge \cdots \wedge b_{\beta_{k-1}} \geqslant 0 \ . \tag{13.23}$$

Como $d(\beta) > 0$, el conjunto ordenado $\gamma := \{j \notin \beta : \beta_1 < j < \beta_k\}$ es no vacío. Mostremos que, para cada $j \in \gamma$, la α-proyección b_j de $C_j(A)$ cumple que $b_j \in \operatorname{Gen}\left\{b_{\beta_1}, b_{\beta_2}, \ldots, b_{\beta_{k-1}}\right\}$. Es decir que

$$b_j \wedge b_{\beta_1} \wedge b_{\beta_2} \wedge \ldots \wedge b_{\beta_{k-1}} = 0 \tag{13.24}$$

Para esto, tomemos i tal que $\beta_i < j < \beta_{i+1}$. Entonces, como $A[\alpha|\beta \cup \{j\}]$ es TP,

$$b_{\beta_1} \ldots \wedge b_{\beta_i} \wedge b_j \wedge b_{\beta_{i+1}} \wedge \ldots \wedge b_{\beta_{k-1}} \geqslant 0 \quad \text{y} \quad b_{\beta_2} \ldots \wedge b_{\beta_i} \wedge b_j \wedge b_{\beta_{i+1}} \wedge \ldots \wedge b_{\beta_k} \geqslant 0. \tag{13.25}$$

Ahora sustituyamos la expresión (13.22) para b_{β_k} en la Eq. (13.25) para obtener

$$(-1)^{k-1} \xi_1 \, b_{\beta_1} \ \wedge \ldots \wedge b_{\beta_i} \wedge b_j \wedge b_{\beta_{i+1}} \wedge \ldots \wedge b_{\beta_{k-1}} \geqslant 0. \tag{13.26}$$

es claro que las ecuaciones (13.21) y (13.23) versus (13.25) y (13.26) son consistentes sólo si la igualdad se da en la Eq. (13.26). Pero como $\xi_1 \neq 0$, solo queda que la Eq. (13.24) sea válida. El argumento muestra que $\operatorname{rk}\big(A[\alpha|\beta \cup \gamma]\big) = k-1$. Si lo que pasaba era que $d(\alpha) > 0$, consideremos el conjunto ordenado $\tau := \{i \notin \alpha : \alpha_1 < i < \alpha_k\}$. El argumento anterior, aplicado a los vectores filas (o a A^T), dice que $\operatorname{rk} A[\alpha \cup \tau|\beta \cup \gamma] = k-1$.

Por último, se sigue de la Eq. (13.20) y de que $d(\beta) > 0$ o $d(\alpha) > 0$, que existe algún $\omega \in Q_{k,n}$ tal que $d(\omega) = 0$, $\omega \subseteq \alpha \cup \tau$, y $\omega \subseteq \beta \cup \gamma$. Pero, como $\operatorname{rk} A[\alpha \cup \tau|\beta \cup \gamma] = k-1$, se debe cumplir que $\det A[\omega] = 0$, lo que contradice al Corolario 13.2.5. Esto completa la prueba en todos los casos. ■

La “ida” del Teorema 13.4.2 se verá como consecuencia del siguiente resultado más general.

Teorema 13.4.5. *Sean $A_1, \ldots, A_p \in \mathcal{M}_n(\mathbb{R})$, inversibles y TPs, con $p \geq n-1$. Si cada A_i satisface la Eq.* (13.19)*, entonces el producto $A_1 A_2 \cdots A_p$ es ETP.*

Demostración. Por el Teorema 13.1.5 basta mostrar que

$$\det(A_1 A_2 \cdots A_p)[\alpha|\beta] > 0 \quad \text{para} \quad \alpha, \beta \in Q_{k,n} \quad \text{tales que} \quad d(\alpha) = d(\beta) = 0 \ .$$

Asumamos que $\beta_1 \geq \alpha_1$ y sea $\omega^{(0)} = \alpha$. Definamos $\omega^{(l)} \in Q_{k,n}$, para $l \in \mathbb{I}_{p-1}$, como

$$\omega_i^{(l)} = \text{mín } \left\{\beta_i \, , \, \alpha_i + \text{máx}\{l + i - k, 0\}\ \right\} \ , \quad i \in \mathbb{I}_k \ .$$

Es fácil ver que $\omega^{(p)} = \beta$ y que cada par $\omega^{(l-1)}, \omega^{(l)}$ satisface la Eq. (13.20). Usando la Proposición 13.4.4, podemos deducir que $\det A_l[\omega^{(l-1)}|\omega^{(l)}] > 0$, para todo $l \in \mathbb{I}_p$. Por lo tanto, se sigue de la Eq. (12.1) (Cahuchy-Binnet) y la positividad total que

$$\det(A_1 A_2 \cdots A_p)[\alpha|\beta] \ \geq \ \prod_{l=1}^{p} \det A_l[\omega^{(l-1)}|\omega^{(l)}] > 0 \ ,$$

lo que prueba el Teorema. ■

Corolario 13.4.6. *Sea $A \in \mathcal{M}_n(\mathbb{R})$.*

1. *Si A es OSC, entonces A^{n-1} es ETP.*
2. *Si A es ε-RS, es inversible y cumple que*

$$a_{ii} \neq 0 \quad \textit{para} \quad i \in \mathbb{I}_n \quad \textit{y} \quad a_{i\,,\,i+1}a_{i+1\,,\,i} > 0 \ , \quad \textit{para} \quad i \in \mathbb{I}_{n-1} \ ,$$

   entonces $A^{2(n-1)}$ es ETP.

Demostración.

1. Se sigue inmediatamente del Teorema 13.4.5.
2. A^2 es TP por ser A ε-RS. Por la hipótesis se tiene que $(A^2)_{i\,,\,i+1} > 0$ y $(A^2)_{i+1\,,\,i} > 0$ para todo $i \in \mathbb{I}_{n-1}$. Entonces satisface la Eq. (13.19), y podemos usar la parte 1. ■

13.5 Variación de signos

Esta sección está dedicada a caracterizaciones de la regularidad de signo de una matriz en términos de algunas propiedades de disminución de variación del operador lineal que ésta induce.

Definición 13.5.1. Sea $x \in \mathbb{R}^n$. Dada una sucesión de signatura ε, decimos que

1. ε es una sucesión de signo de x si para todo $i \in \mathbb{I}_n$ se cumple que $\varepsilon_i x_i = |x_i|$.
2. En tal caso, el número de cambios de signo de x asociado a ε, denotado por $C(\varepsilon)$, es el número de índices $i \in \mathbb{I}_{n-1}$ tales que $\varepsilon_i\, \varepsilon_{i+1} < 0$. Es decir que

$$C(\varepsilon) = \frac{1}{2} \ \sum_{i=1}^{n-1} (1 - \varepsilon_i\, \varepsilon_{i+1}) \ .$$

3. La *máxima variación de signos* $V_+(x)$ (resp. mínima variación de signos $V_-(x)$) es el máximo (resp. mínimo) de los valores $C(\varepsilon)$, cuando ε recorre todas las sucesiones de signo de x. Vemos que

$$0 \leq V_-(x) \leq V_+(x) \leq n-1 \quad \text{para todo} \quad x \in \mathbb{R}^n \ .$$

Si ninguna componente de x se anula, x tiene una única sucesión de signo, y por lo tanto $V_-(x) = V_+(x)$. Este valor común es llamado la variación de signo exacta y denotado por $V(x)$. ▲

Observación 13.5.2. Sea $x \in \mathbb{R}^n$.

1. Ojo que x puede tener variación de signo exacta, aunque tenga coordenadas nulas. Por ejemplo si $x = (1, 0, -1)$, entonces $V_-(x) = V_+(x) = 1$.

2. Pero si x tiene variación de signo exacta, es fácil ver que

   (a) $x_1 \neq 0 \neq x_n$.

   (b) Si $x_i = 0$, entonces $x_{1-1}\, x_{i+1} < 0$.

   (c) En particular, x no puede tener dos coordenadas nulas seguidas.

3. Si $\alpha \in Q_{k,n}$ y x_α es la α-proyección de x a $\mathbb{R}^\alpha$, entonces

$$V_-(x_\alpha) \leq V_-(x) \quad \text{y} \quad V_+(x_\alpha) \leq V_+(x) \ .$$

En efecto, si ε es una sucesión de signo de x, entonces ε_α lo será para x_α (y todas se consiguen así). Pero es fácil ver que $C(\varepsilon_\alpha) \leq C(\varepsilon)$. ▲

Proposición 13.5.3. *Sean $a_1, a_2, \dots, a_m \in \mathbb{R}^n$, linealmente independientes, con $n > m$. Entonces son equivalentes:*

(1) $V_+(b) \leq m-1$*, para todo* $b \in \operatorname{Gen}\{a_1, a_2, \dots, a_m\} \setminus \{0\}$.

(2) $a = a_1 \wedge a_2 \wedge \cdots \wedge a_m$ *es estrictamente definido (como vector), i. e.* $\pm a > 0$.

Demostración. Sea $A = [a_1, a_2, \dots, a_m] \in \mathcal{M}_{n,m}(\mathbb{R})$. Es claro que la condición $a > 0$ equivale a que $\det A[\alpha|-] > 0$ para todo $\alpha \in Q_{m,n}$. Para ver la suficiencia, supongamos que $a > 0$ y que existe $b \in \operatorname{Gen}\{a_1, a_2, \dots, a_m\} \setminus \{0\}$ tal que $V_+(b) \geq m$. Es fácil ver que existe $\alpha \in Q_{m+1,n}$ tal que la α-proyección de b tiene variación máxima m. Como $\det A[\beta|-] > 0$ para todo $\beta \in Q_{m,n}$, en particular aquellos $\beta \subseteq \alpha$, deducimos que las α-proyecciones a_i' de a_i para $i \in \mathbb{I}_m$ también cumplen $a_1' \wedge a_2' \wedge \cdots \wedge a_m' > 0$ y que $b_\alpha \in \operatorname{Gen}\{a_1', \dots, a_m'\}$. Por lo tanto, considerando la α-proyección si es necesario, podemos suponer que

$$n = m+1 \quad \text{y que} \quad (-1)^{i-1} b_i \geq 0 \ , \quad \text{para todo} \quad i \in \mathbb{I}_n \ .$$

Como los $\mathbf{e}^\wedge_{(i)} = e_1 \wedge \cdots \wedge e_{i-1} \wedge e_{i+1} \wedge \cdots \wedge e_n$, $i \in \mathbb{I}_n$ forman una base completa ortogonal de $\Lambda^m \mathbb{R}^n$, tenemos que

$$a_1 \wedge \cdots \wedge a_m = \sum_{i=1}^n \xi_i \, \mathbf{e}^\wedge_{(i)} \ , \quad \text{donde} \quad \xi_i = \sqrt{m}\, \langle a_1 \wedge \cdots \wedge a_m \, , \mathbf{e}^\wedge_{(i)} \rangle = \det A(\, i\,|-] > 0 \ ,$$

para todo $i \in \mathbb{I}_n$.

Como $b \in \text{Gen}\{a_1, a_2, \ldots, a_m\}$ y, además, $b = \sum_{i=1}^{n} b_i\, e_i$, tenemos que

$$0 = b \wedge a_1 \wedge a_2 \wedge \cdots \wedge a_m = \left\{ \sum_{i=1}^{n} (-1)^{i-1} \xi_i \; b_i \right\} \cdot e_1 \wedge e_2 \wedge \cdots \wedge e_n \ ,$$

porque $e_i \wedge \mathbf{e}^{\wedge}_{(j)} = \delta_{ij}(-1)^{i-1} e_1 \wedge e_2 \wedge \cdots \wedge e_n$. Pero las condiciones (ya verificadas) $\xi_i > 0$ y $(-1)^{i-1} b_i \geq 0$, $i \in \mathbb{I}_n$ implican entonces que $b_i = 0$, $i \in \mathbb{I}_n$, o sea $b = 0$, una contradicción. Esto completa la prueba de la suficiencia.

Probemos ahora la necesidad. Como vimos al principio, bastaría verificar que

$$\det A[\alpha|-] \det A[\beta|-] > 0 \quad \text{para todo par} \quad \alpha\ ,\ \beta \in Q_{m,n}\ .$$

Fijados α y β, podemos unirlos por una sucesión $\alpha = \omega^{(0)},\ \omega^{(1)}, \ldots, \omega^{(k)} = \beta$ en $Q_{m,n}$ que verifica la siguiente propiedad: para cada $i \in \mathbb{I}_k$ existe

$$\tau^{(i)} \in Q_{m+1,n} \quad \text{tal que} \quad \omega^{(i-1)} \subseteq \tau^{(i)} \quad \text{y} \quad \omega^{(i)} \subseteq \tau^{(i)}\ .$$

Observar que la desigualdad $\det A[\alpha|-] \det A[\beta|-] > 0$ se sigue de las desigualdades

$$\det A[\omega^{(i-1)}] \det A[\omega^{(i)}|-] > 0 \quad , \qquad 1 \leq i \leq k\ , \tag{13.27}$$

dado que

$$\operatorname{sgn}\left(\det A[\alpha|-] \det A[\beta|-]\right) = \operatorname{sgn}\left(\prod_{i=1}^{k} \det A[\omega^{(i-1)}] \det A[\omega^{(i)}|-] \right)\ .$$

Considerando, en el caso i-ésimo de la Eq. (13.27), la proyección sobre $\tau^{(i)}$, podemos asumir nuevamente que $n = m + 1$, ya que estamos trabajando en dos subconjuntos de $\tau^{(i)}$ y, sobre todo, porque la hipótesis $V_+(b) \leq m - 1$, para todo $b \in \text{Gen}\{a_1, a_2, \ldots, a_m\} \setminus \{0\}$ se preserva al proyectar sobre $\tau^{(i)}$, por la Observación 13.5.2. Ahora, como antes,

$$a_1 \wedge a_2 \wedge \cdots \wedge a_m = \sum_{i=1}^{n} \xi_i \; \mathbf{e}^{\wedge}_{(i)} \qquad \text{con} \qquad \xi_i = \det A(\{i\}|-]\ .$$

Si $\xi_i = 0$ para algún i, entonces $e_i \in \text{Gen}\{a_1, a_2, \ldots, a_m\}$. Pero en tal caso tendríamos que $V_+(e_i) = n - 1 = m$, lo que contradice la condición (1). Además, si no todos los ξ_j tienen el mismo signo, entonces $\xi_l \xi_{l+1} < 0$ para algún l. Entonces, si $b = \xi_{l+1} e_l + \xi_l e_{l+1}$, tenemos como antes que

$$b \wedge a_1 \wedge a_2 \wedge \cdots \wedge a_m = \Big((-1)^{l-i} \xi_{l+1} \xi_l + (-1)^l \xi_l \xi_{l+1} \Big) \;\; e_1 \wedge e_2 \wedge \ldots \wedge e_n = 0\ ,$$

por lo que $b \in \text{Gen}\{a_1, a_2, \ldots, a_m\}$. Pero como $\xi_l \xi_{l+1} < 0$, tenemos $V_+(b) = n - 1 = m$, lo que también contradice la condición (1). Luego todos los signos de ξ_i son iguales, es decir, se cumple (2). ∎

Lema 13.5.4. *Sea $x \in \mathbb{R}^n$. Entonces*

$$V_+(x) + V_-(J_n x) = V_-(x) + V_+(J_n x) = n - 1 \ . \tag{13.28}$$

Demostración. Cuando ε recorre todas las sucesiones de signo de x, $J_n\varepsilon$ recorre todas las sucesiones de signo de $J_n x$. Observar que

$$\begin{aligned} C(\varepsilon) + C(J_n\varepsilon) &= \frac{1}{2}\sum_{i=1}^{n-1}(1 - \varepsilon_i\varepsilon_{i+1}) + \frac{1}{2}\sum_{i=1}^{n-1}(1 - (J_n\varepsilon)_i(J_n\varepsilon)_{i+1}) \\ &= \frac{1}{2}\left[\sum_{1=1}^{n-1}(1 - \varepsilon_i\varepsilon_{i+1}) + (1 - (-1^{i-1})\varepsilon_i(-1)^i\varepsilon_i)\right] \\ &= \frac{1}{2}\sum_{i=1}^{n-1}(1 - \varepsilon_i\varepsilon_{i+1}) + (1 + \varepsilon_i\varepsilon_{i+1}) = n - 1 \ , \end{aligned}$$

lo que muestra la Eq. (13.28). ■

Teorema 13.5.5. *Sea $\mathcal{M}$ un subespacio de $\mathbb{R}^n$ tal que $0 < \dim \mathcal{M} < n$. Entonces, las siguientes dos condiciones son equivalentes:*

(1) $V_+(x) \le \dim \mathcal{M} - 1$ para todo $x \in \mathcal{M} \setminus \{0\}$.

(2) $V_-(y) \ge \dim \mathcal{M}$ para todo $y \in \mathcal{M}^\perp \setminus \{0\}$.

Demostración. Tomemos bases completas ortonormales

$$\{a_1, a_2, \dots, a_m\} \quad \text{para} \quad \mathcal{M}, \quad \text{y} \quad \{a_{m+1}, a_{m+2}, \dots, a_n\} \quad \text{para} \quad \mathcal{M}^\perp \ .$$

Si $A = [a_1, a_2, \dots, a_n]$, entonces A es unitaria y podemos asumir que $\det A = 1$. Por la Proposición 13.5.3 (y su prueba), la condición (1) es equivalente a que $\det A[\alpha|\mathbb{I}_m]$ sea no nulo y tenga el mismo signo para todo $\alpha \in Q_{m,n}$. Como $A \in \mathcal{U}(n)$ y $\det A = 1$,

$$\begin{aligned} \det(J_n A J_n)(\alpha|\mathbb{I}_m) &= \det(J_n(A^*)^{-1}J_n)(\alpha|\mathbb{I}_m) = \det A^*[\mathbb{I}_m|\alpha] \\ &= \det A[\alpha|\mathbb{I}_m] \ , \end{aligned}$$

por la Eq. (12.11). Luego $\det(J_n A J_n)[\tau|m+1, \dots, n]$ es no nulo y tiene el mismo signo para todo $\tau \in Q_{n-m,n}$. Llamemos $b_i = J_n A J_n e_i = (-1)^i J_n a_i$, para $i > m$. La condición anterior es equivalente a que $b_{m+1} \wedge b_{m+2} \wedge \cdots \wedge b_n > 0$ (o bien < 0). Por lo tanto, también $J_n a_{m+1} \wedge J_n a_{m+2} \wedge \cdots \wedge J_n a_m$ es estrictamente definida. Entonces, nuevamente por la Proposición 13.5.3, obtenemos que $V_+(J_n y) \le n - m - 1$ para todo $y \in \mathcal{M}^\perp \setminus \{0\}$. Aplicando el Lema 13.5.4, deducimos la condición (2). La implicación (2) $\Rightarrow$ (1) se prueba igual. ■

Una versión local del Teorema 13.5.5 da la siguiente caracterización de la regularidad de signo estricta en términos de una propiedad de disminución de variación.

Teorema 13.5.6. *Sea $A \in \mathcal{M}_{n,m}(\mathbb{R})$, con $n \geq m$. Entonces A es ε-ERS si y sólo si el operador lineal $A : \mathbb{R}^m \to \mathbb{R}^n$ disminuye la variación de signos, en el sentido de que*

$$V_+(Ax) \leq V_-(x) \quad \text{para todo} \quad x \in \mathbb{R}^m \setminus \{0\} \tag{13.29}$$

Demostración. Supongamos que $A = [a_1, a_2, \dots, a_m]$ es ε-ERS. Tomemos $x \in \mathbb{R}^m \setminus \{0\}$, y llamemos $k = V_-(x)$. Entonces existen $\beta, \omega \in Q_{k+1,m}$ tales que

$$\beta_i \leq \omega_i < \beta_{i+1} \quad \text{para todo} \quad i \in \mathbb{I}_k \quad \text{y} \quad \omega_k < \beta_{k+1} \leq \omega_{k+1} \ ,$$

que describen los signos de x de la siguiente manera:

- $\beta_1 = \min\{j \in \mathbb{I}_n : x_j \neq 0\}$ y $\omega_{k+1} = \max\{j \in \mathbb{I}_n : x_j \neq 0\}$.
- Las componentes de x tienen signo constante (no nulo en los bordes) para todo j entre β_i y ω_i .
- $x_j = 0$ si $\omega_i < j < \beta_{i+1}$ para algún i.
- Para cada $i \in \mathbb{I}_k$, hay un cambio de signo entre x_{ω_i} y la siguinete entrada no nula de x, que es $x_{\beta_{i+1}}$.

Si, para cada $i \in \mathbb{I}_{k+1}$, llamamos

$$b_i = \sum_{\beta_i \leq j \leq \omega_i} x_j a_j \quad \Longrightarrow \quad Ax = \sum_{i=1}^{k+1} b_i \ , \quad \text{ya que} \quad Ax = \sum_{i=1}^{n} x_i a_i \ ,$$

y $x_j = 0$ para aquellos j que no aparecen en algún b_i . Ahora la regularidad estricta de signo de A implica que, si vamos eligiendo $k+1$-tuplas $(j_1, \dots, j_{k+1})$ tales que $\beta_i \leq j_i \leq \omega_i$, $i \in \mathbb{I}_{k+1}$, tenemos que

$$\varepsilon_{k+1} \cdot a_{j_1} \wedge a_{j_2} \wedge \cdots \wedge a_{j_{k+1}} > 0 \ .$$

Por lo tanto, si $\gamma = \varepsilon_{k+1} \cdot \operatorname{sgn} \left(\prod_{i=1}^{k+1} x_{\omega_i} \right)$, se tiene que

$$\gamma \ \cdot \ b_1 \wedge b_2 \wedge \cdots \wedge b_{k+1} > 0 \ ,$$

dado que es una suma de vectores tales que todas sus coordenadas tienen signo γ. Entonces el Lema 13.5.3 dice que

$$V_+(Ax) = V_+ \left(\sum_{i=1}^{k+1} b_i \right) \leq k = V_-(x) \ ,$$

lo que prueba la Eq. (13.29).

Supongamos recíprocamente que $A = [a_1, a_2, \dots, a_m]$ satisface la condición (13.29). Sean $\omega \in Q_{k,m}$ y $x \in \mathbb{R}^m$ tales que $x_\omega = \sum\limits_{i=1}^{k} x_i e_{\omega_i} \neq 0$. Por otra parte, puede obsevarse que

$Ax_\omega = \sum_{i=1}^{k} x_i a_{\omega_i} \in \operatorname{Gen}\{a_{\omega_1}, \dots, a_{\omega_k}\}$. Es claro que $V_-(x_\omega) \le k-1$, de donde por hipótesis, $V_+(Ax_\omega) \le V_-(x_\omega) \le k-1$. Esto dice que

$$V_+(y) \le k-1 \quad \text{para todo} \quad y \in \operatorname{Gen}\{a_{\omega_1}, \dots, a_{\omega_k}\} \setminus \{0\} .$$

Entonces se sigue del Lema 13.5.3 que $a_{\omega_1} \wedge a_{\omega_2} \wedge \cdots \wedge a_{\omega_k}$ es estrictamente definida. Por lo tanto, A será ε-ERS si el signo $a_{\omega_1} \wedge a_{\omega_2} \wedge \cdots \wedge a_{\omega_k}$ depende sólo de k. Para $k = m$ esto es trivial. Fijemos $1 \le k \le m-1$ y tomemos $\alpha, \beta \in Q_{k,m}$. Como en la prueba del Lema 13.5.3, existe una sucesión $\alpha = \omega^{(0)},\ \omega^{(p)}, \dots, \omega^{(r)} = \beta$ en $Q_{k,n}$ que verifica la siguiente propiedad: para cada $i \in \mathbb{I}_r$, existe

$$\tau^{(i)} \in Q_{k+1,n} \quad \text{tal que} \quad \omega^{(i-1)} \subseteq \tau^{(i)} \quad \text{y} \quad \omega^{(i)} \subseteq \tau^{(i)} .$$

Por lo tanto, basta probar que, para cada $\tau \in Q_{k+1,m}$ e $i \in \mathbb{I}_{k+1}$, se cumple que

$$a_{\tau_1} \wedge \cdots \wedge a_{\tau_{i-1}} \wedge a_{\tau_{i+1}} \wedge \cdots \wedge a_{\tau_{k+1}} \quad \text{y} \quad a_{\tau_1} \wedge \cdots \wedge a_{\tau_i} \wedge a_{\tau_{i+2}} \wedge \cdots \wedge a_{\tau_{k+1}}$$

tienen el mismo signo. Mediante un argumento de continuidad esto será establecido si

$$a_{\tau_1} \wedge \cdots \wedge a_{\tau_{i-1}} \wedge \{(1-t)a_{\tau_i} + t a_{\tau_{i+1}}\} \wedge a_{\tau_{i+2}} \wedge \cdots \wedge a_{\tau_{k+1}}$$

es estrictamente definido para cada $0 < t < 1$. Y esto se deduce del Lema 13.5.3, vía la Eq. (13.29), porque, si $x \in \mathbb{R}^k \setminus \{0\}$, entonces

$$V_-\left(\sum_{j=1}^{i-1} x_j e_{\tau_j} + x_i((1-t)e_{\tau_i} + t e_{\tau_{i+1}}) + \sum_{j=i+2}^{k+1} x_{j-1} e_{\tau_j} \right) \le k-1 ,$$

puesto que los signos de las coordenadas τ_i y τ_{i+1} son iguales. ∎

Lema 13.5.7. *Sea $x \in \mathbb{R}^n$. Si una sucesión $x_p \xrightarrow[p\to\infty]{} x$, entonces*

$$V_-(x) \le \liminf_{p\to\infty} V_-(x_p) \quad y \quad \limsup_{p\to\infty} V_+(x_p) \le V_+(x) . \tag{13.30}$$

Demostración. Si llamamos $J = \{i \in \mathbb{I}_n : x_i \ne 0\}$, debe existir algún $p_0 \in \mathbb{N}$ tal que

$$|(x_p)_j - x_j| < \min\{|x_i| : i \in J\} \quad \text{para todo} \quad j \in J \quad \text{y} \quad p \ge p_0 .$$

Entonces, $\operatorname{sgn}(x_p)_j = \operatorname{sgn}(x_j)$ para $j \in J$. Luego, si $p \ge p_0$, toda sucesión de signo ε_p para x_p es también una sucesión de signo para x, porque los signos (de x y x_p) coinciden en J, mientras que en los $i \notin J$ no hay problemas, porque $x_i = 0$. Resumiendo, para todo $p \ge p_0$ debe cumplirse que $V_-(x) \le V_-(x_p)$ para $p \ge p_0$ (porque x tiene más sucesiones de signo que x_p). O sea que $V_-(x) \le \liminf_{p\to\infty} V_-(x_p)$. La otra desigualdad se prueba igual. ∎

La regularidad de signo está caracterizada por una propiedad de variación de signo más débil.

Corolario 13.5.8. *Sea $A \in \mathcal{M}_{n,m}(\mathbb{R})$ con $\operatorname{rk} A = m$. Entonces A es ε-RS si y sólo si*

$$V_-(Ax) \le V_-(x) \quad \text{para todo} \quad x \in \mathbb{R}^m \setminus \{0\} . \tag{13.31}$$

Demostración. Como veremos en la Sección 7, existe una sucesión $(G_p)_{p\in\mathbb{N}}$ en $\mathcal{G}l(n)$ de matrices ETPs , tales que $G_p \xrightarrow[p\to\infty]{} I_n$. Supongamos primero que A es ε-RS. Como $\operatorname{rk}A = m$, el Teorema 13.2.1 asegura que G_pA es ε-ERS para todo $p \in \mathbb{N}$. Además $G_pA \xrightarrow[p\to\infty]{} A$. Entonces el Teorema 13.5.6 garantiza que

$$V_+(G_pAx) \le V_-(x) \quad \text{para todo} \quad x \in \mathbb{R}^m \setminus \{0\} \ .$$

Luego, por la Eq. (13.30), tenemos que

$$V_-(Ax) \le \liminf_{p\to\infty} \le V_+(G_{p_0}Ax) \le V_-(x) \ ,$$

lo que muestra la Eq. (13.31). Supongamos ahora que la Eq. (13.31) es válida. Por el Teorema 13.5.6, aplicado a G_p , como A es inyectiva,

$$V_+(G_p(Ax)\,) \le V_-(Ax) \le V_-(x) \quad \text{para todo} \quad p \in \mathbb{N} \quad \text{y} \quad x \in \mathbb{R}^m \setminus \{0\} \ ,$$

El Teorema 13.5.6 (al reves) muestra que G_pA debe ser ε-ERS para todo $p \in \mathbb{N}$. Tomando límite, vemos que A debe ser ε-RS. ■

Usando la relación de dualidad (13.28), podemos hablar de algunas propiedades de aumento de signo.

Corolario 13.5.9. *Sea $A \in \mathcal{M}_{n,m}(\mathbb{R})$ con $\operatorname{rk}A = m$. Entonces J_nAJ_m es ERS (respectivamente, RS) si y sólo si*

$$n - m + V_+(x) \le V_-(Ax) \quad (\text{ resp.} \quad V_+(Ax)\,) \ ,$$

para todo $x \in \mathbb{R}^m \setminus \{0\}$. ■

Cuando $n = m$, la regularidad de signo admite varias caracterizaciones equivalentes.

Teorema 13.5.10. *Sea $A \in \mathcal{G}l(n)$. Entonces las siguientes condiciones son equivalentes:*

1. *A es regular de signo.*
2. *$V_+(Ax) \le V_+(x)$ para todo $x \in \mathbb{R}^n \setminus \{0\}$.*
3. *$V_-(Ax) \le V_+(x)$ para todo $x \in \mathbb{R}^n \setminus \{0\}$.*
4. *$V_-(Ax) \le V_-(x)$ para todo $x \in \mathbb{R}^n \setminus \{0\}$.*

Demostración.

$1 \to 2$: Si A es regular de signo (e inversible), también lo es $J_nA^{-1}J_n$ por el Teorema 13.2.3. Entonces 13.5.10 se sigue del Corolario 13.5.9, reemplazando x por Ax y A por A^{-1} (en nuestro caso, $n - m = 0$).

$2 \to 3$: Trivial.

$3 \to 4$: Uusaremos la sucesión G_p que aparece en la prueba del Corolario 13.5.8. El mismo argumento de la prueba de $1 \to 2$, muestra que,

$$V_+(G_pAx) \le V_-(x) \quad \text{para todo} \quad p \in \mathbb{N} \quad \text{y} \quad x \in \mathbb{R}^n \setminus \{0\} ,$$

porque G_pA es ERS. Luego, como $G_pA\,x \xrightarrow[p\to\infty]{} Ax$, el Lema 13.5.7 nos da que

$$V_-(Ax) \le \liminf V_-(G_pAx) \le \liminf V_+(G_pAx) \le V_-(x) \ , \quad x \in \mathbb{R}^n \setminus \{0\} \ .$$

$4 \to 1$: Se sigue del Corolario 13.5.8. ∎

13.6 Totalmente Perron-Frobenius

En esta sección estudiaremos propiedades espectrales de las matrices regulares de signo o totalmente positivas. La herramienta clave para esto son los resultados de Perron y Frobenius para matrices positivas. Recordemos la parte más elemental del teorema de Perron-Frobenius:

Observación 13.6.1. Sea $A \in \mathcal{M}_n(\mathbb{R})$ tal que $A \geqslant 0$.

1. Llamaremos $\lambda_1(A), \dots, \lambda_n(A)$ a sus autovalores, ordenados de forma que

$$|\lambda_1(A)| \ge |\lambda_2(A)| \ge \cdots \ge |\lambda_n(A)| \ .$$

2. El mayor autovalor de A es real y no negativo, i.e., $\rho(A) = \lambda_1(A) \ge 0$, y hay un autovector positivo $u_1 \geqslant 0$ correspondiente a $\lambda_1(A)$.

3. Si $A > 0$, entonces

$$\lambda_1(A) > |\lambda_2(A)| \quad \text{y} \quad \ker(A - \lambda_1(A)I) = \operatorname{Gen}\{u_1\} \ ,$$

para cierto $u_1 > 0$.

Teorema 13.6.2. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz ε-ERS. Entonces todos los autovalores de A son reales y distintos. Más aún,*

$$\frac{\varepsilon_m}{\varepsilon_{m-1}} \ \lambda_m(A) > |\lambda_{m+1}(A)| \ , \qquad \text{para todo} \quad m \in \mathbb{I}_n \, , \tag{13.32}$$

donde usamos la convención $\varepsilon_0 = 1$ y $\lambda_{n+1}(A) = 0$. Además, los correspondientes autovectores $u_1, u_2, \dots, u_n$ pueden ser elegidos en $\mathbb{R}^n$, y de modo tal que

$$u_1 \wedge u_2 \wedge \cdots \wedge u_m > 0 \quad \text{(como vector)} \ , \ \text{ para todo} \quad m \in \mathbb{I}_n \, . \tag{13.33}$$

Demostración. La prueba se hará por inducción en m. El caso $m = 1$ es consecuencia de la Observación 13.6.1 porque $\varepsilon_1 A > 0$ por hipótesis. Supongamos que el resultado es cierto para $1 \le i \le m-1$. Como $\varepsilon_m \cdot \Lambda^m A > 0$, la Observación 13.6.1 dice que

$$0 < \rho(\varepsilon_m \cdot \Lambda^m A) = \lambda_1(\varepsilon_m \cdot \Lambda^m A) = \varepsilon_m \prod_{i=1}^{m} \lambda_i(A) \ .$$

Por la hipótesis inductiva, que en particular nos dice que $\frac{\varepsilon_i}{\varepsilon_{i-1}}\ \lambda_i(A) = |\lambda_i(A)| > 0$, para $i < m$, y del hecho de que $\rho(\varepsilon_m \cdot \Lambda^m A)$ es el único autovalor de $\varepsilon_m \cdot \Lambda^m A$ de módulo máximo, deducimos que

$$\varepsilon_m \prod_{i=1}^{m} \lambda_i(A) = \prod_{i=1}^{m} \frac{\varepsilon_i}{\varepsilon_{i-1}}\ \lambda_i(A) = \frac{\varepsilon_m}{\varepsilon_{m-1}}\ \lambda_m(A) \prod_{i=1}^{m-1} |\lambda_i(A)| > |\lambda_{m+1}(A)| \prod_{i=1}^{m-1} |\lambda_i(A)| \ .$$

Luego la Eq. (13.32) se cumple para m. Ahora, como $\lambda_m(A)$ es real, u_m puede ser elegido real. Por lo tanto tenemos que $u_1 \wedge u_2 \wedge \cdots \wedge u_m$ es autovector no nulo de $\varepsilon_m \cdot \Lambda^m A$ correspodiente a $\lambda_1(\varepsilon_m \cdot \Lambda^m A)$, y tiene coordenadas reales. Entonces, por la Observación 13.6.1, tomando $\xi = 1$ o bien $\xi = -1$, tenemos que $\xi \cdot u_1 \wedge u_2 \wedge u_2 \wedge \cdots \wedge u_m > 0$. Ahora, reemplazando a u_m por ξu_m en caso de que sea necesario, obtenemos la Eq. (13.33) para todo $m \in \mathbb{I}_n$. ■

Los autovectores reales $\{u_1, u_2, \dots, u_n\}$ conseguidos en el Teorema 13.6.2 posee propiedades oscilatorias interesantes. Para sus formulaciones, necesitamos algunas definiciones.

Definición 13.6.3. Sea $x \in \mathbb{R}^n$.

1. Notaremos por $x(t) : [1, n] \to \mathbb{R}^n$ a la función linear a trozos

$$x(t) = (k + 1 - t)x_k + (t - k)x_{k+1} \qquad \text{si} \qquad k \le t \le k + 1 \ , \quad k \in \mathbb{I}_{n-1} \ . \tag{13.34}$$

   Observar que $x(t)$ es continua, lineal a trozos, y que $x(j) = x_j$ para todo $j \in \mathbb{I}_n$.

2. Los **nodos** de $x(t)$ son las raíces de la ecuación $x(t) = 0$, ordenados de manera creciente (si son finitos, i.e., si no hay dos coordenadas nulas consecutivas de x).

3. Diremos que dos sucesiones ordenadas $\xi_1 < \xi_2 < \cdots < \xi_k$ y $\eta < \eta_2 \cdots < \eta_{k+1}$ están **entrelazadas** si se cumple que

$$\eta_k < \xi_k < \eta_{k+1} \ , \qquad \text{para todo} \qquad k \in \mathbb{I}_k \ . \qquad \blacktriangle$$

Teorema 13.6.4. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz ε-ERS. Sean $u_1, \dots, u_n$ sus autovectores reales, correspondientes a los autovalores $\lambda_k(A)$, $k \in \mathbb{I}_n$ (ordenados con módulos decrecientes). Entonces*

1. *La variación de signo de cada u_k es exacta. Más aún,*

$$V(u_k) = k - 1 \ , \qquad \text{para todo} \qquad k \in \mathbb{I}_n \ . \tag{13.35}$$

2. *Además, los nodos de $u_k(t)$ y los de $u_{k+1}(t)$ están entrelazados.*

Demostración. Fijemos $k \in \mathbb{I}_n$. Por el Teorema 13.6.2 podemos asumir que $u_1 \wedge u_2 \wedge \cdots \wedge u_k > 0$. Luego, por el Lema 13.5.3, sabemos que $V_+(u_k) \leq k-1$. Consideremos $J_n A^{-1} J_n$, la cual es nuevamente ERS por el Teorema 13.2.3. Como $J_n u_k$ es un autovector de $J_n A^{-1} J_n$ correspondiente a $1/\lambda_k(A) = \lambda_{n-k+1}(J_n A^{-1} J_n)$, el argumento anterior da que $V_+(J_n u_k) \leq n-k$. Por la Eq. (13.28), deducimos que

$$V_- u_k = n-1-V_+(J_n u_k) \geq k-1 \geq V_+(u_k) \ ,$$

lo que prueba el item 1. Para probar el item 2, necesitamos varios pasos previos:

Clamor 1: Para todo $k \in \mathbb{I}_{n-1}$ y todo $(\xi, \zeta) \in \mathbb{R}^2 \setminus \{0\}$, se cumple que

$$V_+(\xi u_k + \zeta u_{k+1}) - 1 \leq V_-(\xi u_k + \zeta u_{k+1}) \ . \tag{13.36}$$

En efecto, como $u_1 \wedge \cdots \wedge u_k \wedge u_{k+1} > 0$ (o < 0), la Proposición 13.5.3 garantiza que, si llamamos $z = \xi u_k + \zeta u_{k+1}$, entonces $V_+(z) \leq (k+1)-1 = k$. Aplicando el mismo argumento a $J_n u_n, \ldots, J_n u_{k+1}, J_n u_k$, que son los primeros $n-k+1$ autovectores de las matriz ERS $J_n A^{-1} J_n$, obtenemos, vía la Eq. (13.28), que

$$V_+(J_n z) \leq n-k \implies V_-(z) \geq k-1 \geq V_+(z) - 1 \ ,$$

lo que termina de mostrar la Eq. (13.36). ▲

Sean $x(t) = u_k(t)$ e $y(t) = u_{k+1}(t)$. Por la Eq. (13.35) y la Observación 13.5.2, si un nodo de e $x(t)$ o de $y(t)$ es entero, la coordenada correspondiente de u_k o u_{k+1} es nula, no es ni la primera ni la última, y las dos adyacentes son no nulas y de signos opuestos. Por lo tanto, $x(t)$ tiene $k-1$ nodos e $y(t)$ tiene k nodos.

Clamor 2: Sean $(\xi, \zeta) \in \mathbb{R}^2 \setminus \{0\}$ y $j \in \mathbb{I}_n \setminus \{1, n\}$. Si $\xi x(j) + \zeta y(j) = 0$, entonces

$$\Big(\xi x(j-1) + \zeta y(j-1)\Big) \Big(\xi x(j+1) + \zeta y(j+1)\Big) \ < \ 0 \ . \tag{13.37}$$

En efecto, como en la Observación 13.5.2, si $z \in \mathbb{R}^n$ cumple que $V_+(z) - 1 \leq V_-(z)$, entonces $z_j = 0$ implica que $j = 1$, $j = n$, o $z_{j-1} z_{j+1} < 0$. Luego basta aplicar lo anterior y la Eq. (13.36) al vector $z = \xi u_k + \zeta u_{k+1}$. ▲

Ahora vamos a por el item 2: Sean $t_1 < t_2 < \cdots < t_k$ los nodos de $y(t)$. Entonces bastaría probar que, para todo $l \in \mathbb{I}_{k-1}$, hay al menos un nodo de $x(t)$ en el intervalo abierto (t_l, t_{l+1}).

Clamor 3: Supongamos que $x(t) > 0$ para todo $t \in (t_l, t_{l+1})$. Entonces

$$x(t_l) \neq 0 \neq x(t_{l+1}) \quad \text{ y, por ende, } \quad 0 < \alpha = \min\{x(t) \ , \ t \in [t_l, t_{l+1}] \ \} \ .$$

Supongamos, por ejemplo, que $x(t_l) = 0$. Tomemos $i \in \mathbb{N}$ tal que $i-1 < t_l < i$ (o bien $j-1$ y $j+1$ si $t_l = j$ es entero). Como $x(t)$ es lineal en $[i-1, i]$, tenemos que $x(i-1)x(i) < 0$ (resp. $x(j-1)x(j+1) < 0$, en este caso por la ecuación (13.37)). Tomando

$$\xi = -\frac{y(i)}{x(i)} \quad \left(\text{ resp. } \xi = -\frac{y(j+1)}{x(j+1)} \right) \quad ,$$

se tiene que $\xi x(t)+y(t)$ se anula en el intervalo $[i-1,i]$, ya que $\xi x(i)+y(i) = \xi x(t_l)+y(t_l) = 0$, y es una función lineal en $[i-1,i]$ (resp. se anula en $[j,j+1]$ por las mismas razones). Pero esto contradice la Eq. (13.37). ▲

Recta final: Por la definición de nodos, $y(t)$ es definido, supongamos que ≥ 0, en el intervalo $[t_l, t_{l+1}]$. Sea η el mínimo de los $\eta > 0$ para los cuales $z_\eta(t) = x(t) - \eta y(t)$ tiene un nodo $s \in [t_l, t_{l+1}]$. Observar que $s \neq t_l$ porque $y(t_l) = 0 \neq x(t_l)$. Por lo mismo $s \neq t_{l+1}$. Ahora, por la minimalidad de η, tenemos que $z_\eta(t) \geq 0$ en $[t_l, t_{l+1}]$. Pero como $z_\eta(t) = (u_k - \eta u_{k+1})(t)$ es lineal en los intervalos $[j, j+1]$, $j \in \mathbb{I}_{n-1}$, esto es posible sólo cuando $s \in \mathbb{N}$, o cuando $z_\eta(t)$ se anula en todo el intervalo $[j, j+1]$ que contiene a s. Pero cada una de estas psibilidades produce una contradicción con la Eq. (13.37), en el primer caso porque $z_\eta(t)$ no cambia de signo, en el otro porque $z_\eta(t)$ tiene dos ceros enteros consecutivos. ■

Si $A \in \mathcal{M}_n(\mathbb{R})$ es estrictamente regular de signo, su adjunta A^* es también estrictamente regular de signo. Por el Teorema 13.6.2, los autovectores reales $\{v_1, v_2, \dots, v_n\}$ de A^* pueden ser elegidos de forma tal que

$$v_1 \wedge v_2 \wedge \cdots \wedge v_k > 0 \ , \qquad \text{para todo} \qquad k \in \mathbb{I}_n \ .$$

Las propiedades (13.33) y (13.38) de los autovectores de A y A^* caracterizan en algún sentido la regularidad de signo estricta.

Teorema 13.6.5. *Si $A \in \mathcal{M}_n(\mathbb{R})$ es inversible, tiene n autovalores reales de distinto módulo, y los autovectores reales u_k de A y v_k de A^*, correspondientes a $\lambda_k(A) = \lambda_k(A^*)$, son elegidos de forma tal que satisfagan las ecuaciones*

$$u_1 \wedge u_2 \wedge \cdots \wedge u_k > 0 \qquad y \qquad v_1 \wedge v_2 \wedge \cdots \wedge v_k > 0, \qquad \textit{para todo} \qquad k \in \mathbb{I}_n \ , \tag{13.38}$$

entonces alguna potencia de A es estrictamente regular de signo.

Demostración. Notemos $\lambda_k = \lambda_k(A)$, para $k \in \mathbb{I}_n$. Sean

$$U = [u_1 \, , u_2 \, , \dots , u_n] \qquad \text{y} \qquad V = [v_1 \, , v_2 \, , \dots , v_n] \ .$$

Entonces U y V son inversibles. Si notamos $\Lambda = \operatorname{diag}(\lambda_1 \, , \lambda_2 \, , \dots , \lambda_n)$, entonces

$$A = U \cdot \Lambda \cdot U^{-1} \quad \text{y} \quad A^* = V \cdot \Lambda \cdot V^{-1} \ . \tag{13.39}$$

Es fácil ver que $\langle u_i, v_j \rangle = 0$ para $i \neq j$. Esto dice que V^*U es diagonal. Sea $\rho \in \mathbb{C}^n$ tal que

$$\operatorname{diag}(\rho)^{-1} = V^*U \qquad \Longrightarrow \qquad U^{-1} = \operatorname{diag}(\rho) \, V^* \ . \tag{13.40}$$

Se tiene que $\rho > 0$ (es decir que sus entradas son positivas), ya que para todo $k \in \mathbb{I}_n$,

$$0 < k! \ \langle u_1 \wedge u_2 \wedge \cdots \wedge u_k, v_1 \wedge u_2 \wedge \cdots \wedge v_k \rangle = \det V^*U[\mathbb{I}_k] = \prod_{i=1}^{k} \rho_i^{-1} \ .$$

Por las ecuaciones (12.1) (Cauchy-Binnet), (13.39) y (13.40), y varias reducciones elementales, podemos ver que, para todo $p \in \mathbb{N}$ y para todo par $\alpha, \beta \in Q_{k,n}$,

$$
\begin{aligned}
\det A^p[\alpha|\beta] &= \det(U\Lambda^p U^{-1})[\alpha|\beta] = \sum_{\omega \in Q_{k,n}} \det U[\alpha|\omega] \cdot \left(\prod_{i=1}^{k} \lambda_{\omega_i}\right)^p \cdot \det U^{-1}[\omega|\beta] \\
&= \sum_{\omega \in Q_{k,n}} \det U[\alpha|\omega] \cdot \left(\prod_{i=1}^{k} \lambda_{\omega_i}\right)^p \cdot \left(\prod_{i=1}^{k} \rho_{\omega_i}\right) \cdot \det V[\beta|\omega] \\
&= \left(\prod_{i=1}^{k} \lambda_i\right)^p \cdot \left(\prod_{i=1}^{k} \rho_i\right) \det U[\alpha|\mathbb{I}_k] \det V[\beta|\mathbb{I}_k] \quad + \\
&\quad + \sum_{\substack{\omega \in Q_{k,n} \\ \omega \neq \mathbb{I}_k}} \det U[\alpha|\omega] \cdot \left(\prod_{i=1}^{k} \lambda_{\omega_i}\right)^p \left(\prod_{i=1}^{k} \rho_{\omega_i}\right) \cdot \det V[\beta|\omega].
\end{aligned}
$$

Notar que, como $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0$, entonces

$$
\prod_{i=1}^{k} |\lambda_i| > \prod_{i=1}^{k} |\lambda_\omega| \quad \text{para todo} \quad \omega \in Q_{k,n} \setminus \mathbb{I}_k \ ,
$$

mientras que las ecuaciones de (13.38) implican que

$$
U[\alpha|\mathbb{I}_k] > 0 \quad \text{y} \quad V[\beta|\mathbb{I}_k] > 0 \quad \text{para todo} \quad k \in \mathbb{N} \quad \text{y} \quad \alpha, \beta \in Q_{k,n} \ .
$$

Entonces para un p suficientemente grande, $\det A^p[\alpha|\beta]$ es no nulo y tiene el mismo signo que $\left(\prod_{i=1}^{k} \lambda_i\right)^p$ para cada $\alpha, \beta \in Q_{k,n}$. Entonces A^p es ERS. ■

Ahora compararemos los autovalores de A con los de $A[\alpha]$, para un α adecuado. El siguiente Teorema generaliza un hecho que sabíamos para $A \in \mathcal{H}(n)$ y $\alpha \in Q_{n-1,n}$ (entrelace de Cauchy) y que habíamos mencionado para también para $Q_{k,n}$ (ver Teorema 2.4.2).

Teorema 13.6.6. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz ETP. Dados $k \in \mathbb{N}$ y $\alpha \in Q_{k,n}$ con componentes consecutivas (i.e., tal que $d(\alpha) = 0$), se tiene que, para todo $j \in \mathbb{I}_k$,*

$$
\lambda_j(A) > \lambda_j(A[\alpha]) > \lambda_{n+j-k}(A) \tag{13.41}
$$

y

$$
\lambda_j(A) > \lambda_j(A/\alpha') > \lambda_{n+j-k}(A) \ . \tag{13.42}
$$

Demostración. El caso que concentra la dificulatad es cuando $k = n-1$, donde $\alpha = \mathbb{I}_{n-1}$ o bien $\alpha = \mathbb{I}_n \setminus \{1\}$. Supongamos que $\alpha = \mathbb{I}_{n-1}$, y llamemos $B = A[\alpha]$. Observar que $\lambda(A)$ consta de las n raices distintas del polinomio característico $P_A(t) = \det(tI - A)$. Análogamente $\lambda(B)$ consta de las $n-1$ raices distintas del polinomio característico $P_B(t) = \det(tI - B)$. Llamemos

$\lambda_i = \lambda_i(A)$, $i \in \mathbb{I}_n$ y notemos $A_t = tI_n - A$ y $B_t = tI_n - B$, $t \in \mathbb{R}$. Para mostrar la Eq. (13.41) para este α, basta ver que

$$P_B(\lambda_i)P_B(\lambda_{i+1}) < 0 \ , \qquad \text{para todo} \qquad i \in \mathbb{I}_{n-1} \ . \tag{13.43}$$

Consideremos los vectores x_t, con parámetro real t, definidos por

$$x_t := \left[(-1)^{n+i} \det A_t[\alpha|i)\right]_{i \in \mathbb{I}_n} \ .$$

Entonces la Regla de Cramer (12.10) muestra que, para todo $t \notin \sigma(A)$, se tiene la igualdad $x_t = d_A(t)A_t^{-1}e_n$. Luego, $A_t x_t = P_A(t)e_n$ para esos t, pero por continuidad,

$$A_t x_t = 0 \quad \text{si} \quad t \in \sigma(A) \quad \Longrightarrow \quad Ax_{\lambda_j} = \lambda_j x_{\lambda_j} \ , \quad \text{para} \quad j \in \mathbb{I}_n \ . \tag{13.44}$$

La n-ésima componente $x_t(n)$ de x_t coincide con $P_B(t)$, mientras que la primera componente $x_t(1)$ admite la representación

$$x_t(1) = \sum_{j=2}^{n} t^{n-j} \sum_{\substack{\omega \in Q_{j,n} \\ \omega_1 = 1, \omega_j = n}} \det A[\omega - \{n\}|\omega - \{1\}]. \tag{13.45}$$

Esto último sale del hecho de que

$$A_t[\mathbb{I}_{n-1}|2,\dots,n] = \begin{bmatrix} -a_{12} & -a_{23} & \dots & -a_{1,n-1} & -a_{1,n} \\ t - a_{22} & -a_{23} & \dots & -a_{2,n-1} & -a_{2,n} \\ -a_{32} & t - a_{33} & \dots & -a_{3,n-1} & -a_{3,n} \\ \vdots & & & & \vdots \\ -a_{n-1,2} & -a_{n-1,3} & \dots & t - a_{n-1,n-1} & -a_{n-1,n} \end{bmatrix}$$

y viendo que subdeterminantes le correponden a cada potencia de t.

Clamor: Se cumple que $x_t(1) > 0$ para todo $t > 0$.

En efecto, como A es TP, la Eq. (13.45) muestra que $x_t(1)$ es un polinomio en t con coeficientes no negativos. Luego bastaría mostrar que existe un $\lambda > 0$ tal que $x_\lambda(1) \neq 0$ (así ya sabríamos que algún coeficiente no se anula). Para ello usaremos que, como $P_B(t)$ tiene sólo $n-1$ raices, existe $j \in \mathbb{I}_n$ tal que $x_{\lambda_j}(n) = P_B(\lambda_j) \neq 0$. Por la Eq. (13.44), x_{λ_j} es un autovector vector no nulo (porque $x_{\lambda_j}(n) \neq 0$) de A, que es ETP. Luego, por el Teorema 13.6.4, x_{λ_j} tiene variación exacta, por lo que su primer componente $x_{\lambda_j}(1) \neq 0$. ▲

Aplicando el Clamor a los otros λ_i , junto con la Eq. (13.44), concluimos que x_{λ_i} es el i-ésimo autovector de A con $x_{\lambda_i}(1) > 0$, para todo $i \in \mathbb{I}_n$. Luego se sigue del Teorema 13.6.4 que la n-ésima componente tiene signo $(-1)^{i-1}$. Esto establece la Eq. (13.43), porque $x_{\lambda_i}(n) = P_B(\lambda_i)$, $i \in \mathbb{I}_n$. Para $\alpha = \{2, 3, \dots, n\}$, tomamos nuevamente $B = A[\alpha]$ y ahora

$$y_t = \left[(-1)^{1+i} \det A_t[\alpha|i)\right]_{i \in \mathbb{I}_n} \ .$$

En este caso tenemos $A_t y_t = P_A(t)e_1$, lo que también implica que $Ay_{\lambda_j} = \lambda_j y_{\lambda_j}$. Aquí, $y_t(1)$ coincide con $P_B(t)$, mientras que la última admite un representación como la anterior. Entonces la primera tiene signo $(-1)^{i-1}$ y obtenemos así la Eq. (13.41) para este α.

El caso $k < n-1$ se probará por inducción descendente. Supongamos que la Eq. (13.41) es cierta para cierto $k > 1$ y tomemos $\alpha \in Q_{k-1,n}$ con $d(\alpha) = 0$. Supongamos que $\alpha = \{i, i+1, \dots, i+k-2\}$ con $i+k-1 \le n$. Llamemos $\beta = \alpha \cup \{i+k-1\}$. Aplicando el caso anterior a la matriz ETP $A[\beta] \in \mathcal{M}_k(\mathbb{R})$, obtenemos que

$$\lambda_j(A[\beta]) > \lambda_j(A[\alpha]) > \lambda_{j+1}(A[\beta]) \ , \quad \text{para todo} \quad j \in \mathbb{I}_{k-1} \ .$$

Por otro lado, la hipótesis inductiva asegura que

$$\lambda_j(A) > \lambda_j(A[\beta]) > \lambda_{n+j-k}(A) \ , \quad \text{para todo} \quad j \in \mathbb{I}_k \ .$$

Combinando estas desigualdades, se obtien la Eq. (13.41) para el caso $k-1$, lo que completa la inducción. Resta ver el caso en que $i+k-2 = n$, i. e. $\alpha = \{i, i+1, \dots, n\}$, que se obtiene de manera análoga, aplicando el segundo argumento a la matriz $A[\alpha \cup \{i-1\}]$. Veamos ahora la Eq. (13.42): Sabemos que $J_n A^{-1} J_n$ es también ETP por el Teorema 13.2.3. Además, por el Teorema 12.1.4, se ve fácilmente que $(J_n A^{-1} J_n)[\alpha] = J_\alpha (A/\alpha')^{-1} J_\alpha$. Observemos que $\sigma\left(J_n A^{-1} J_n\right) = \sigma\left(A\right)^{-1}$. Luego,

$$\frac{1}{\lambda_j(A)} = \lambda_{n-j+1}(J_n A^{-1} J_n) \qquad \text{y} \qquad \frac{1}{\lambda_j(A/\alpha')} = \lambda_{k-j+1}((J_n A^{-1} J_n)[\alpha]) \ .$$

Aplicando la Eq. (13.41) a $J_n A^{-1} J_n$, tenemos que

$$\lambda_j(J_n A^{-1} J_n) > \lambda_j(J_n A^{-1} J_n[\alpha]) > \lambda_{n+j-k}(J_n A^{-1} J_n) \ .$$

Luego

$$\begin{aligned} \lambda_j(A) = \lambda_{n-j+1}(J_n A^{-1} J_n)^{-1} \ &> \ \lambda_{k-j+1}(J_n A^{-1} J_n[\alpha])^{-1} = \lambda_j(A/\alpha') \\ &> \ \lambda_j(J_n A^{-1} J_n)^{-1} = \lambda_{n+j-k}(A) \ , \end{aligned}$$

lo que completa la prueba. ∎

Con la ayuda del Teorema de aproximación 13.1.13, algunos de los resulatdos anteriores pueden ser genarlizados al caso en que A es regular de signo o TP.

Corolario 13.6.7. Si $A \in \mathcal{M}_n(\mathbb{R})$ es regular de signo con signatura ε, entonces todos sus autovalores son reales, y

$$\frac{\varepsilon_k}{\varepsilon_{k-1}} \lambda_k(A) > 0 \ , \quad \text{para todo} \quad k = 1,\ 2, \dots,\ \text{rk}A \ .$$

Si A es TP, entonces para cada $k \in \mathbb{I}_n$ y cada $\alpha \in Q_{k,n}$ tal que $d(\alpha) = 0$, se tiene que

$$\lambda_j(A) \ge \lambda_j(A[\alpha]) \ge \lambda_{n+j-k}(A) \ ,$$

para todo $j \in \mathbb{I}_n$. ∎

Dado $x = (x_i) \in \mathbb{R}^n$, denotemos por $x^\downarrow$ a su reordenación decreciente:

$$x_1^\downarrow \geq x_2^\downarrow \geq \cdots \geq x_n^\downarrow \quad \text{ y } \quad x_i^\downarrow = x_{\pi(i)} \quad \text{para alguna} \quad \pi \in \mathbf{S}_n \ . \tag{13.46}$$

Teorema 13.6.8. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ una matriz TP. Entonces*

$$\operatorname{diag}(A) \prec \lambda(A) \ .$$

Demostración. Dada una matriz $L \in \mathcal{M}_m(\mathbb{R})$, llamemos $\delta^*(L) = \operatorname{diag}(L)^\downarrow$, a su diagonal reordenada. Probaremos el teorema por inducción en n. El caso $n = 1$, es trivial. Asumamos que el teorema es cierto en $\mathcal{M}_{n-1}(\mathbb{R})$. Como $\operatorname{tr}\lambda(A) = \operatorname{tr} A = \operatorname{tr}\operatorname{diag}(A)$, y $\lambda_i(A) = \lambda_i^\downarrow(A)$, $i \in \mathbb{I}_n$, basta mostrar que

$$\sum_{i=1}^k \lambda_i(A) \geq \sum_{i=1}^k \delta_i^*(A) \quad \text{ para } \quad k \in \mathbb{I}_{n-1} \ . \tag{13.47}$$

Sean $p, q \in \mathbb{I}_n$ tales que $A_{11} = \delta_p^*(A)$ y $A_{nn} = \delta_q^*(A)$. Tomando la conversión de A en caso de que sea necesario, podemos asumir que $p > q$. Sea $B = A(n)$ y $C = A(1)$. Como $B, C \in \mathcal{M}_{n-1}(\mathbb{R})$ son ambas TP, la hipótesis inductiva determina que

$$\sum_{i=1}^k \lambda_i(B) \geq \sum_{i=1}^k \delta_i^*(B) \quad \text{ y } \quad \sum_{i=k}^{n-1} \lambda_i(C) \leq \sum_{i=k}^{n-1} \delta_i^*(C) \ , \quad k \in \mathbb{I}_{n-1} \ . \tag{13.48}$$

Observar que $\delta_i^*(B) = \delta_i^*(A)$ para $1 \leq i \leq p$ (aca se usa que $p > q$). Por el Corolario 13.6.7, $\lambda_i(A) \geq \lambda_i(B)$, para $i \in \mathbb{I}_{n-1}$. Luego, las desigualdades de (13.48) implican la Eq. (13.47) para los casos $1 \leq k \leq p$. Veamos ahora que

$$\sum_{i=k+1}^n \lambda_i(A) \leq \sum_{i=k+1}^n \delta_i^*(A) \ , \quad \text{ para } \quad k > p \ . \tag{13.49}$$

En efecto, como $\delta_i^*(A) = \delta_i^*(C)$ si $p + 1 \leq i \leq n$, y $\lambda_{i-1}(C) \geq \lambda_i(A)$, para todo $i \in \mathbb{I}_{n-1}$ (por el Corolario 13.6.7), observamos que la Eq. (13.48) implica (13.49), y por ende también la Eq. (13.47) para estos valores de k. ■

13.7 Algunos ejemplos

En esta sección presentamos algunos ejemplos de matrices TPs y la caracterización de estas matrices.

13.7.1. [Núcleos totalmente positivos] La mayor parte de las matrices TPs no triviales surgen de la restricción de núcleos totalmente positivos a conjuntos finitos adecuados. Daremos a continuación algunas fórmulas de producción de núcleos totalmente positivos: Sean Γ, Λ conjuntos totalmente ordenados (en general, subconjuntos de $\mathbb{R}$ o $\mathbb{Z}$).

1. Una función a valores reales $K(s,t)$ para $s \in \Gamma$, $t \in \Lambda$ es un **núcleo totalmente positivo** (TP) si la matriz $[K(s_i,t_j)]_{i,j\in\mathbb{I}_n}$ es TP para todo $n \in \mathbb{N}$ y toda elección
$$s_1 < s_2 < \ldots < s_n \quad \text{en } \Gamma \quad \text{y} \quad t_1 < t_2 < \ldots < t_n \quad \text{en } \Lambda \ .$$
La positividad total estricta de un núcleo se define análogamente.
2. Si $K(s,t)$ es TP y $f(s), g(t)$ son funciones positivas en Γ y Λ respectivamente, entonces el núcleo $f(s)K(s,t)g(t)$ es TP.
3. Si $K(s,t)$ es TP y $\phi(s)$ es un operador monótonamente creciente de un conjunto totalmente ordenado Γ_1 a Γ, y $\psi(t)$ es un operador monótonamente creciente de un conjunto totalmente ordenado Λ_1 a Λ, entonces $K(\phi(s),\psi(t))$ es un núcleo TP en $\Gamma_1 \times \Lambda_1$.
4. Si dos núcleos $L(s,t)$ y $M(s,t)$ son TPs y $d\sigma(\cdot)$ es una medida en Γ, entonces el núcleo
$$K(u,v) := \int_T L(s,u)M(s,v)d\sigma(s) \ , \qquad \text{para} \ \ u,v \in \Lambda \ , \tag{13.50}$$
es TP en $\Lambda \times \Lambda$, si la integral existe. Esto es sólo una modificación del Teorema 13.2.1.

Pasemos ahora a la construcción de ejemplos concretos

1. El núcleo $L(k,t) = t^k$, definido en en $\mathbb{N}_0 \times \mathbb{R}_+$ es TP. Esto es una consecuencia de la positividad total de las matrices de Vandermonde, vista en el Ejemplo 13.1.6.
2. Dados $k \in \mathbb{I}_n$ y $\alpha \in \mathbb{R}_+^{n+1}$, el núcleo $K(s,t) = \sum\limits_{k=0}^{n} \alpha_k s^k t^k$ es TP en $\mathbb{R}_+ \times \mathbb{R}_+$. En efecto, a $K(s,t)$ se lo puede realizar como una composición del tipo (13.50), con dos copias del núcleo L del item anterior (con la medida en $\mathbb{N}_0$ dada por los α_k).
3. Para cualquier $\sigma > 0$ el núcleo $K(s,t) = \exp(\sigma st)$ es TP en $\mathbb{R}_+ \times \mathbb{R}_+$, ya que es un límite de núcleos del item anterior (con $\alpha_k = \sigma^k/k!$).
4. El núcleo $K(s,t) = \exp[-\sigma(s-t)^2]$ es ETP en $\mathbb{R}_+ \times \mathbb{R}_+$, porque
$$\exp\left[-\sigma(s-t)^2\right] = \exp(-\sigma s^2)\exp(2\sigma st)\exp(-\sigma t^2) \ , \qquad \text{para todo par} \ \ s\,,t \in \mathbb{R}_+ \ .$$
5. Por lo tanto, para todo $n \in \mathbb{N}$ y todo $p \in \mathbb{R}_+^*$, la matriz
$$G_p = \Big(\exp\left[-\,p\,(i-j)^2\right]\Big)_{i,j\,\in\mathbb{I}_n} \in \mathcal{M}_n(\mathbb{R})$$
es ETP por el item 3. Además, se tiene que $G_p \xrightarrow[p\to\infty]{} I_n$. Esta sucesión (tomando $p \in \mathbb{N}$) ha sido usada varias veces en secciones anteriores.

6. Para cada $0 < \lambda < 1$ y $0 \neq p \in \mathbb{R}$, consideremos el *promedio pesado* en $\mathbb{R}_+ \times \mathbb{R}_+$
$$M_{\lambda,p} := \{\lambda s^p + (1-\lambda)t^p\}^{1/p}. \tag{13.51}$$
Entonces $M_{\lambda,p}(s,t)$ o $1/M_{\lambda,p}(s,t)$ es TP de acuerdo a si $p < 0$ o $p > 0$. Esto se sigue de la observación de que para cualquier $\gamma > 0$
$$\frac{1}{(s+t)^\gamma} = \frac{1}{\Gamma(\gamma)} \int_{-\infty}^{0} e^{us} e^{ut} \frac{du}{|u|^{1-\gamma}}, \tag{13.52}$$
donde $\Gamma(\cdot)$ es la función gamma, y el núcleo $\exp(us)$ es TP en $\mathbb{R}_+ \times \mathbb{R}_+$.

7. El núcleo $K(s,t) = \text{mín}\{s,t\}$ es TP en $\mathbb{R}_+ \times \mathbb{R}_+$, porque
$$K(s,t) = \lim_{p \to -\infty} M_{\lambda,p}(s,t) \tag{13.53}$$

8. Si $f(t), g(t)$ son funciones positivas en $\mathbb{R}_+$ tales que $h(t) = \dfrac{f(t)}{g(t)}$ es no decreciente, entonces el núcleo
$$K(s,t) = f\big(\text{ mín}\{s,t\}\big) \cdot g\big(\text{ máx}\{s,t\}\big)$$
es TP en $\mathbb{R}_+ \times \mathbb{R}_+$. En efecto, es fácil ver que se la puede reescribir como
$$\begin{aligned} K(s,t) &= \text{mín}\{h(s),h(t)\}\; g\big(\text{ mín}\{s,t\}\big)\; g\big(\text{ máx}\{s,t\}\big) \\ &= g(s) \cdot \text{mín}\{h(s),h(t)\} \cdot g(t)\ . \end{aligned}$$

9. Dado $\sigma > 0$, poniendo $g(t) = \exp(-\sigma t)$ y $f(t) = \exp(\sigma t)$ (con lo que nos queda $h(t) = \exp(2\sigma t)$, que es creciente), obtenemos que el núcleo del item 8,
$$K(s\,,t) = \exp\big(\sigma \text{ mín}\{s,t\} - \sigma \text{ máx}\{s,t\}\big) = \exp\left(-\sigma|s-t|\right) \quad \text{ es TP en } \mathbb{R}_+ \times \mathbb{R}_+\ .$$

10. Sean $\{b_i\}_{i\in\mathbb{I}_n}$ y $\{c_i\}_{i\in\mathbb{I}_n}$ dos sucesiones en $\mathbb{R}_+^*$. Entones la matriz
$$\mathcal{M}_n(\mathbb{R}) \ni \big[\, b_{\text{mín}(i,j)}\, c_{\text{máx}(i,j)} \,\big]_{i\,,j\,\in\mathbb{I}_n} \text{ es TP} \quad \Longleftrightarrow \quad \frac{b_1}{c_1} \le \frac{b_2}{c_2} \le \cdots \le \frac{b_n}{c_n}\ .$$
Esto se sigue inmediatamente del item 8, ya que podemos considerar las funciones
$$f(t) = b_i\,, \quad \text{si } i-1 \le t < i \quad \text{y} \quad g(t) = c_i\ , \quad \text{si } i-1 \le t < i\ .$$
Una matriz de este tipo es llamada *matriz de Green.* ▲

13.7.2 (Matriz de Hurwitz)**.** Un conocido teorema de A. Hurwitz dice que un polinomio $p(z) = d_0 z^n + d_1 z^{n-1} + \ldots + d_n$ a coeficientes reales ($d_0 > 0$) tiene todos sus ceros en semiplano abierto $\operatorname{Re} z < 0$ si y sólo si la matriz

$$H_p = \left[\, d_{2j-i} \,\right]_{i,j \in \mathbb{I}_n} = \begin{bmatrix} d_1 & d_3 & d_5 & d_7 & d_9 & \cdots & 0 \\ d_0 & d_2 & d_4 & d_6 & d_8 & \cdots & 0 \\ 0 & d_1 & d_3 & d_5 & d_7 & \cdots & 0 \\ 0 & d_0 & d_2 & d_4 & d_6 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & d_n \end{bmatrix} \in \mathcal{M}_n(\mathbb{R}) \ , \tag{13.54}$$

donde ponemos que $d_k = 0$ para $k < 0$ o $k > n$, tiene menores principales positivos:

$$\det H_p\,[1, 2, \ldots, k] > 0 \ , \qquad \text{para todo } k \in \mathbb{I}_n \quad . \tag{13.55}$$

Un tal polinomio $p(z)$ es llamado un *polinomio de Hurwitz* y la matriz H es la matriz de Hurwitz asociada a él.

Mostraremos, por inducción en n, que en tal caso la matriz de Hurwitz es TP. Observar que $d_1 > 0$ para cualquier n, por la Eq. (13.55). Luego el caso $n = 1$ es trivial. Supongamos que es cierto para $n - 1$. Tomemos una $H_p \in \mathcal{M}_n(\mathbb{R})$ para un buen polinomio p. Llamemos $G = H/\{1\} \in \mathcal{M}_{n-1}(\mathbb{R})$, indexada en $\{2, 3, \ldots, n\}$. La Eq. (12.15) nos asegura que

$$d_1 \det \Big(H/\{1\}[\{2, \ldots, k\}] \Big) = \det H_p[\mathbb{I}_k] > 0 \qquad \text{para todo } k \in \mathbb{I}_n \setminus \{1\} \quad . \tag{13.56}$$

Sean $g_j = F_j(G) \in \mathbb{R}^{n-1}$, para $j = 2, 3, \ldots, n$. Llamemos $c = \frac{d_0}{d_1}$. Entonces las matriz $T \in \mathcal{M}_{n-1}(\mathbb{R})$, indexada en $\{2, 3, \ldots, n\}$, cuyas filas $f_j = F_j(T)$ están definidos por

$$f_2 = g_2 \ , \quad f_{2j-1} = g_{2j-1} \quad , \qquad \text{y} \qquad f_{2j} = g_{2j} - c\, g_{2j-1} \ \text{ para } \ j \geq 2 \ , \tag{13.57}$$

también tiene menores principales positivos. Haciendo la cuenta vemos que T es una matriz de la forma (13.54) con $n - 1$ en lugar de n, y d'_j en lugar de d_j , donde

$$d'_{2j} = d_{2j+1} \qquad \text{y} \qquad d_{2j-1} = d_{2j} - c\, d_{2j+1} \ , \qquad \text{para } \ j = 0, 1, 2, \ldots \tag{13.58}$$

Por la hipótesis inductiva, tenemos que T es TP, por lo que también lo es

$$\tilde{T} := \begin{bmatrix} 0 & 0 \\ 0 & T \end{bmatrix} \in \mathcal{M}_n(\mathbb{R}) \ .$$

Haciendo algunas cuentas, podemos deducir de la Eq. (13.58) que

$$H_p[1, 2, \ldots, n-2] = H_p(n-1, n) = \left(\left\{ S + \frac{c}{2}(I_n - J_n) \right\} S \tilde{T} S^* \right)(n-1, n), \tag{13.59}$$

donde $S = [0, e_1, e_2, \ldots, e_{n-1}]$. Las matrices S y S^* son TPs, y lo es la matriz triangular superior $S + \frac{c}{2}(I_n - J_n)$. Ahora la positividad total de H_p sale de la Eq. (13.59) por los Teoremas 13.2.1 y 13.1.4. ▲

13.7.3 (Matrices de Toeplitz). Para una sucesión (bi-)infinita $\{a_n : -\infty < n < \infty\}$, la matriz $(a_{i-j})_{i,j\in\mathbb{N}}$ es llamada su matriz de Toeplitz, y la función $f(z) = \sum\limits_{-\infty}^{\infty} a_n z^n$, su función generadora. Un matriz de Toeplitz es TP si y sólo si su función generadora es de la forma

$$f(z) = Cz^k \exp\left(\gamma_1 z + \frac{\gamma_{-1}}{z}\right) \cdot \frac{\prod\limits_1^{\infty}(1+\alpha_n z)\prod\limits_1^{\infty}\left(1+\frac{\rho_n}{z}\right)}{\prod\limits_1^{\infty}(1-\beta_n z)\prod\limits_1^{\infty}\left(1-\frac{\delta_n}{z}\right)} ,$$

donde k es un entero, $C \geq 0$, $\gamma_1, \gamma_{-1} \geq 0$ y $\alpha_n, \beta_n, \rho_n, \delta_n \geq 0$ son tales que

$$\sum_1^{\infty}(\alpha_n + \beta_n + \rho_n + \delta_n) < \infty .$$

Cuando $a_n = 0$ para $n < 0$, la matriz de Toeplitz es TP si y sólo si su función generadora es de la forma

$$f(z) = Ce^{\gamma z} \frac{\prod\limits_1^{\infty}(1+\alpha_n z)}{\prod\limits_1^{\infty}(1-\beta_n z)} ,$$

donde $C \geq 0$, $\gamma \geq 0$, y $\alpha_n, \beta_n \geq 0$ son tales que $\sum\limits_1^{\infty}(\alpha_n + \beta_n) < \infty$. Las pruebas de estos hechos, basadas fuertemente en la teoría de funciones analíticas están más allá del alcance de este trabajo. Cuando es aplicada a un polinomio la caracterización anterior implica que el polinomio $p(z) = d_0 z^n + d_1 z^{n-1} + \ldots + d_n$ $(d_0 > 0)$ tiene todos sus ceros en eje real no negativo si y sólo si la matriz infinita $(d_{n+j-i})_{i,j\in\mathbb{N}}$ es TP, donde $d_k = 0$ para $k < 0$ o $k > n$. Notemos que la matriz de Hurwitz H_p introducida antes es una submatriz de T, más precisamente $H_p = T[n+1, n+2, \ldots, 2n|2, 4, \ldots, 2n]$. ▲

13.7.4 (Función de frecuencia de Pólya). Una función $f(t)$ en $(-\infty, \infty)$ es llamada una *función de frecuencia de Pólya* si el núcleo $K(s,t) := f(s-t)$ es TP. La siguiente caracterización se debe a Schoenberg (1953), $f(t)$ es una función de frecuencia de Pólya si y sólo si su transformada bilátera de Laplace existe en una tira abierta que contenga al eje imaginario y tiene la forma

$$\int_{-\infty}^{\infty} e^{-st} f(s)ds = C\exp(\gamma t^2 + \delta t) \cdot \prod_1^{\infty} \frac{\exp(\alpha_n t)}{1+\alpha_n t},$$

donde $C > 0$, $\gamma \geq 0$, δ y α_n son reales tales que $0 < \sum\limits_1^{\infty}|\alpha_n|^2 < \infty$. La prueba de este resultatdo está más alla del alcance de este trabajo. ▲

13.8 Apéndice: La prueba del criterio clave

Las pruebas de los criterios de positividad total se basan en intrincados cálculos de determinantes que permiten usar un argumento inductivo. Para ello usaremos fuertemente los resultados de la sección 2 del Capítulo 12. Recordemos, dado que las usaremos bastante, algunas ecuaciones de allí:

Dados $\alpha, \beta \in Q_{k,n}$ y, además, $\omega, \tau \in Q_{l,n}$ tales que $\omega \subseteq \alpha', \tau \subseteq \beta'$, sean

$$\mu = \alpha \cup \omega = (\mu_1, \mu_2, \dots, \mu_{k+l}) \quad \text{y} \quad \nu = \beta \cup \tau = (\nu_1, \nu_2, \dots, \nu_{k+l}) \in Q_{k+l,n} \ .$$

Existen entonces γ y $\sigma \in Q_{k,k+l}$ tales que $\alpha_i = \mu_{\gamma_i}$ y $\beta_i = \nu_{\sigma_i}$, $i \in \mathbb{I}_k$. Luego definimos

$$\operatorname{sgn} \frac{\alpha}{\alpha \cup \omega} = \operatorname{sgn}(\gamma) = (-1)^{\operatorname{tr} \gamma - \frac{k(k+1)}{2}} \ , \quad \operatorname{sgn} \frac{\beta}{\beta \cup \tau} = \operatorname{sgn}(\sigma) \ . \tag{13.60}$$

Repasemos la Eq. (12.15): Sea $A \in \mathcal{M}_n(\mathbb{R})$. Dados $\alpha, \beta \in Q_{k,n}$ y además, $\omega, \tau \in Q_{l,n}$ tales que $\omega \subseteq \alpha', \tau \subseteq \beta'$, entonce se tiene que

$$\det A[\alpha|\beta] \det \Big((A/[\alpha|\beta])[\omega|\tau] \Big) = \operatorname{sgn} \frac{\alpha}{\alpha \cup \omega} \ \operatorname{sgn} \frac{\beta}{\beta \cup \tau} \ \det A[\alpha \cup \omega | \beta \cup \tau] \tag{13.61}$$

Una consecuencia inmediata es la siguiente caracterización de las entradas de un complemento de Schur, vista como (12.16): Dados $\alpha, \beta \in Q_{k,n}$, se tiene

$$\{A/[\alpha|\beta]\}_{(\alpha'_i, \beta'_j)} = \operatorname{sgn} \frac{\alpha}{\alpha \cup \{\alpha'_i\}} \ \operatorname{sgn} \frac{\beta}{\beta \cup \{\beta'_j\}} \ \frac{\det A[\alpha \cup \{\alpha'_i\} | \beta \cup \{\beta'_j\}]}{\det A[\alpha|\beta]} \tag{13.62}$$

La siguiente igualdad, vista en (12.12), es válida para toda $A \in \mathcal{M}_n(\mathbb{R})$: Dado $\beta \in Q_{k,n}$,

$$\sum_{\omega \in Q_{k,n}} \operatorname{sgn}(\omega) \ \det A[\omega|\beta] \ \det A(\omega|\beta) = \operatorname{sgn}(\beta) \ \det A \ . \tag{13.63}$$

Identidad de Sylvester (Eq. (12.17)): Dados $A \in \mathcal{M}_n(\mathbb{R})$ y $\alpha, \beta \in Q_{k,n}$, se cumple que

$$\det \Big(\big\{ \det A[\alpha \cup \{\alpha'_i\} | \beta \cup \{\beta'_j\}] \big\}_{i,j \in \mathbb{I}_{n-k}} \Big) = \det A \cdot \det A[\alpha|\beta]^{n-k-1} \tag{13.64}$$

Necesitamos además el siguiente resultado específico:

Lema 13.8.1. *Sea $A \in \mathcal{M}_n(\mathbb{R})$. Dados $\alpha \in Q_{n-1,n}$ y $\omega \in Q_{n-2,n}$ tales que $\omega \subseteq \alpha$, se tiene*

$$\det A[\omega|1,n) \det A[\alpha|q) = \det A[\omega|1,q) \det A[\alpha|n) + \det A[\omega|q,n) \det A[\alpha|1) \ , \tag{13.65}$$

para todo $1 < q < n$. Notar que asumimos que $n \geq 3$.

Demostración. Fijemos $p \in \omega$ y sean $\mu = \omega \setminus \{p\}$ y $\nu = \{1, q, n\}'$. Además sean $\{m\} = \alpha \setminus \omega$. Dividiendo ambos lados de la Eq. (13.65) por $\det A[\mu|\nu]^2$ tenemos que el lado izquierdo de la (eventual) igualdad queda

$$\frac{\det A[\mu \cup \{p\} | \nu \cup \{q\}] \ \det A[\mu \cup \{p, m\} | \nu \cup \{1, n\}]}{\det A[\mu|\nu]^2} = \diamond$$

y el derecho,

$$\frac{\det A[\mu \cup \{p\}|\nu \cup \{n\}] \ \det A[\mu \cup \{p,m\}|\nu \cup \{1,q\}]}{\det A[\mu|\nu]^2} \quad +$$

$$\frac{\det A[\mu \cup \{p\}|\nu \cup \{1\}] \ \det A[\mu \cup \{p,m\}|\nu \cup \{q,n\}]}{\det A[\mu|\nu]^2} = \diamond\diamond \ .$$

Llamemos $B = A/[\mu|\nu]$. Notar que, por las ecuaciones (13.61) y (13.62), se tiene que

$$\diamond \ = \operatorname{sgn} \tfrac{\mu}{\mu \cup \{p\}} \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{q\}} \ \operatorname{sgn} \tfrac{\mu}{\mu \cup \{p,m\}} \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{1,n\}} \ B_{p,q} \ \det \Big(\ B[p,m|1,n] \Big) \ ,$$

Por otra parte, $\diamond\diamond = \varepsilon_2 \ B_{p,n} \ \det B[p,m|1,q] + \varepsilon_3 \ B_{p,1} \ \det B[p,m|q,n]$, donde

$$\varepsilon_2 \ = \operatorname{sgn} \tfrac{\mu}{\mu \cup \{p\}} \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{n\}} \ \operatorname{sgn} \tfrac{\mu}{\mu \cup \{p,m\}} \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{1,q\}} \quad \text{y}$$

$$\varepsilon_3 \ = \operatorname{sgn} \tfrac{\mu}{\mu \cup \{p\}} \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{1\}} \ \operatorname{sgn} \tfrac{\mu}{\mu \cup \{p,m\}} \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{q,n\}} \ .$$

Sacando como factor a $\operatorname{sgn} \tfrac{\mu}{\mu \cup \{p\}} \ \operatorname{sgn} \tfrac{\mu}{\mu \cup \{p,m\}}$, se ve que la Eq. (13.65) es equivalente a la siguiente relación: $\operatorname{sgn} \tfrac{\nu}{\nu \cup \{q\}} \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{1,n\}} \ B_{p,q} \ \det B[p,m|1,n] =$

$$\operatorname{sgn} \tfrac{\nu}{\nu \cup \{n\}} \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{1,q\}} \ B_{p,n} \ \det B[p,m|1,q] + \operatorname{sgn} \tfrac{\nu}{\nu \cup \{1\}} \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{q,n\}} \ B_{q,1} \ \det B[p,m|q,n] \ .$$

Por otra parte, usando la Eq. (13.60), una cuidadosa cuenta muestra que

$$\operatorname{sgn} \tfrac{\nu}{\nu \cup \{q\}} \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{1,n\}} \ = \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{n\}} \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{1,q\}} \ = \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{1\}} \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{q,n\}} \ . \tag{13.66}$$

En efecto, observar que por la definición usando permutaciones tenemos que

$$\operatorname{sgn} \tfrac{\nu}{\nu \cup \{q\}} \ = \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{q,n\}} \ , \quad \operatorname{sgn} \tfrac{\nu}{\nu \cup \{1\}} \ = \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{1,n\}} \ , \quad \text{y que} \quad \operatorname{sgn} \tfrac{\nu}{\nu \cup \{n\}} \ = \ 1 \ .$$

Luego las igualdades de (13.66) surgen de que $\operatorname{sgn} \tfrac{\nu}{\nu \cup \{1,q\}} = \operatorname{sgn} \tfrac{\nu}{\nu \cup \{1\}} \ \operatorname{sgn} \tfrac{\nu}{\nu \cup \{q\}}$. Usando ahora la Eq. (13.66), nos queda que la Eq. (13.65) es equivalente a la relación:

$$B_{p,q} \ \det \ B[p,m|1,n] = B_{p,n} \ \det B[p,m|1,q] + B_{p,1} \ \det B[p,m|q,n] \ ,$$

que se verifica fácilmente para cualquier matrix B (notar que son determinantes de matrices de 2×2). ■

Sean $A \in \mathcal{M}_{n,m}(\mathbb{R})$ y ε una sucesión de signatura. Sea $r = \min\{n,m\}$. Recordemos las definiciones: A es ε-RS (resp. ε-ERS) si

$$\varepsilon_k \det A[\alpha|\beta] \geq 0 \quad (\text{resp.} > 0) \quad \text{para todo} \quad k \in \mathbb{I}_r \ , \quad \alpha \in Q_{k,n} \ , \quad \beta \in Q_{k,m} \ . \tag{13.67}$$

Recordemos el enunciado del Teorema 13.1.4:

Teorema 13.1.4 Sea $A \in \mathcal{M}_{n,m}(\mathbb{R})$ con $\operatorname{rk} A = r$, y sea ε una sucesión de signatura.

1. Para que A sea ε-RS es suficiente que, para todo $k \in \mathbb{I}_{\min\{n,m\}}$ y $\alpha \in Q_{k,n}$,

$$\varepsilon_k \det A[\alpha|\beta] \geq 0 \quad \text{para} \quad \beta \in Q_{k,m} \quad \text{tal que} \quad d(\beta) \leq m - r \ . \tag{13.68}$$

2. En particular, A es TP si $\det A[\alpha|\beta] \geq 0$ en esos casos.

Demostración. Observar que cualquier β con $d(\beta) \neq 0$ cumple que $d(\beta) \leq m-2$. Eso hace innecesario estudiar los casos en que $r \leq 2$, en particular si $n \leq 2$ o $m \leq 2$. Asumamos entonces que $n\,, m \geq 3$. Probaremos la la Eq. (13.67) por inducción en k, asumiendo que A cumple la condición (13.68). Cuando $k=1$, la Eq. (13.67) es cierta porque $d(\beta)=0$ para cualquier $\beta \in Q_{1,m}$. Supongamos que se cumple la Eq. (13.67) para todos los $j<k$ pero no para k. Luego deben existir un $\beta \in Q_{k,m}$ y un $\alpha \in Q_{k,n}$ tales que

$$\varepsilon_k \det A[\alpha|\beta] < 0 \ , \tag{13.69}$$

y que $d(\beta)$ es el mínimo para los β que cumplen tal cosa. En particular tenemos que

$$l = d(\beta) > m - r \ . \tag{13.70}$$

Afirmamos el tal β cumple que para todo $p \notin \beta$ tal que $\beta_1 < p < \beta_k$, debe pasar que

$$a_p \wedge a_{\beta_2} \wedge \cdots \wedge a_{\beta_{k-1}} = 0 \ . \tag{13.71}$$

En el caso $k=2$ esto sería que $a_p = 0$. Observar que si (13.71) fuera cierta, tendríamos que $\dim \operatorname{Gen}\{a_j : \beta_1 \leq j \leq \beta_k\} \leq k$, y estamos considerando $k+l$ columnas. Por lo tanto $r = \operatorname{rk} A \leq m-l$, lo que contradiría la Eq. (13.70). Esta contradicción mostraría que la Eq. (13.67) es válida para $d(\beta)=k$.

Así que vamos a por la fórmula (13.71): Para esto fijemos un p como antes, y llamemos $\tau = \{\beta_2, \beta_3, \ldots, \beta_{k-1}\}$. Una reformulación de (13.71) es decir que para todo tal p vale que el $\operatorname{rk}(A[-|\tau \cup \{p\}]\,) \leq k-2$. Dado $\omega \in Q_{k-1,n}$, con $\omega \subseteq \alpha$, el Lema 13.8.1 nos dice que

$$\det A[\omega|\tau \cup \{p\}] \det A[\alpha|\tau \cup \{\beta_1, \beta_k\}] =$$

$$\det A[\omega|\tau \cup \{\beta_k\}] \det A[\alpha|\tau \cup \{\beta_1, p\}] + \det A[\omega|\tau \cup \{\beta_1\}] \det A[\alpha|\tau \cup \{p, \beta_k\}] \tag{13.72}$$

Como $\tau \cup \{\beta_1, \beta_k\} = \beta$, $d(\tau \cup \{\beta_1, p\}) \leq l-1$ y $d(\tau \cup \{p, \beta_k\}) \leq l-1$, se sigue de la Eq. (13.69), la hipótesis inductiva y la propiedad minimal de l que la identidad arriba mencionada sólo puede ser válida cuando

$$\det A[\omega|\tau \cup \{p\}] = 0 \ , \qquad \text{para todo} \qquad \omega \in Q_{k-1,n} \ , \ \ \omega \subseteq \alpha \ , \tag{13.73}$$

pues el lado derecho de la igualdad (13.72) tiene signo $\varepsilon_{k-1}\,\varepsilon_k$ (o 0) y en el izquierdo hay, por hipótesis, un factor cumple que $\varepsilon_k \det A[\alpha|\beta] < 0$ y el otro $\varepsilon_{k-1} \det A[\omega|\tau \cup \{p\}] \geq 0$. Por otro lado, si $k \geq 3$, al calcular $\det A[\alpha|\beta] \neq 0$ vía la Eq. (13.63), vemos que existe un $\gamma \in Q_{k-2\,,n}$ tal que $\gamma \subseteq \alpha$ y $\det A[\gamma|\tau] \neq 0$. Luego, para probar que $\operatorname{rk} A[-|\tau \cup \{p\}] \leq k-2$, sería suficiente mostrar que todo vector fila de $A[-|\tau \cup \{p\}]$ es una combinación lineal de los vectores fila con índices en γ, o equivalentemente que

$$\det A[\gamma \cup \{q\}|\tau \cup \{p\}] = 0 \ , \qquad \text{para todo} \qquad q \in \mathbb{I}_n \setminus \gamma \ . \tag{13.74}$$

En el caso $k=2$ el tal $\gamma = \varnothing$ (al igual que τ), pero es claro que (13.74) equivale a que $a_p = 0$, y el resto de la cuenta funciona. Cuando $q \in \alpha$, (13.74) se deduce de la Eq. (13.73), ya que

$\gamma \cup \{q\} \in Q_{k-1,n}$ y está dentro de α. Fijemos ahora un $q \notin \alpha$. Sean $\mu = \{\mu_1, \mu_2, \mu_3\} = (\alpha \setminus \gamma) \cup \{q\}$, y $\nu = \{\beta_1, p, \beta_k\}$. Consideremos la matriz

$$B \in \mathcal{M}_3(\mathbb{R}) \quad \text{dada por} \quad b_{ij} = \det A[\gamma \cup \{\mu_i\} | \tau \cup \{\nu_j\}] \ , \quad \text{para} \quad i\,,j \in \mathbb{I}_3 \ .$$

Entonces por hipótesis inductiva todos los b_{ij} tienen el mismo signo ε_{k-1} y, por la Identidad de Sylvester (13.64), todos los subdeterminantes de matrices 2×2 de $B[-|1)$ y $B[-|3)$ tienen el mismo signo $\varepsilon_{k-2}\,\varepsilon_k$. Por otro lado, la Eq. (13.73) implica que $b_{i,2} = 0$ siempre que $\mu_i \neq q$. Luego la Eq. (13.74) equivale a que $C_2(B) = b_2 = 0$. Si $q = \mu_1$, tendríamos

$$B = \begin{bmatrix} \det A[\gamma \cup \{q\}|\tau \cup \{\beta_1\}] & \det A[\gamma \cup \{q\}|\tau \cup \{p\}] & \det A[\gamma \cup \{q\}|\tau \cup \{\beta_k\}] \\ \det A[\gamma \cup \{\mu_2\}|\tau \cup \{\beta_1\}] & 0 & \det A[\gamma \cup \{\mu_2\}|\tau \cup \{\beta_k\}] \\ \det A[\gamma \cup \{\mu_3\}|\tau \cup \{\beta_1\}] & 0 & \det A[\gamma \cup \{\mu_3\}|\tau \cup \{\beta_k\}] \end{bmatrix} ,$$

con todas las entradas del mismo signo. Si $b_2 \neq 0$, las condiciones anteriores sólo son consistentes cuando $b_{2,1} = b_{3,1} = 0$ o bien $b_{2,3} = b_{3,3} = 0$. Esto es así porque los de la izquierda producen determinantes (de 2×2) de un signo y los de la derecha del signo contrario, cosa solo permitida si del lado malo (el signo que no concuerda con $\varepsilon_{k-2}\,\varepsilon_k$) son ambos cero. En el caso de que $q = \mu_3$ pasa lo mismo ($b_{1,1} = b_{2,1} = 0$ o bien $b_{1,3} = b_{2,3} = 0$).

Aplicando nuevamente la Eq. (13.64) tendríamos que, si $\alpha \setminus \gamma = \{a_1\, a_2\}$, entonces

$$\det \begin{bmatrix} \det A[\gamma \cup \{a_1\}|\tau \cup \{\beta_1\}] & \det A[\gamma \cup \{a_1\}|\tau \cup \{\beta_k\}] \\ \det A[\gamma \cup \{a_2\}|\tau \cup \{\beta_1\}] & \det A[\gamma \cup \{a_2\}|\tau \cup \{\beta_k\}] \end{bmatrix} = \det A[\alpha|\beta] \det A[\gamma|\tau] \quad (13.75)$$

es nulo, mientras que $\det A[\gamma|\tau] \neq 0$. Llegamos a que $\det A[\alpha|\beta] = 0$, lo que no vale.

Supongamos ahora que $q = \mu_2$. Queda

$$B = \begin{bmatrix} \det A[\gamma \cup \{\mu_1\}|\tau \cup \{\beta_1\}] & 0 & \det A[\gamma \cup \{\mu_1\}|\tau \cup \{\beta_k\}] \\ \det A[\gamma \cup \{q\}|\tau \cup \{\beta_1\}] & \det A[\gamma \cup \{q\}|\tau \cup \{p\}] & \det A[\gamma \cup \{q\}|\tau \cup \{\beta_k\}] \\ \det A[\gamma \cup \{\mu_3\}|\tau \cup \{\beta_1\}] & 0 & \det A[\gamma \cup \{\mu_3\}|\tau \cup \{\beta_k\}] \end{bmatrix} ,$$

Ahora debe pasar que, si $b_{22} \neq 0$, entonces $b_{1,1} = b_{3,3} = 0$ o bien $b_{1,3} = b_{3,1} = 0$. Esto sale porque $\det B[1,2|1,2]$ y $\det B[1,2|2,3]$ deben tener signo $\varepsilon_{k-2}\,\varepsilon_k$, pero deberían ser opuestos, porque todos los b_{ij} tienen el mismo signo. Si por ejemplo el malo es el de la derecha, debe pasar que $b_{1,3} = 0$. Y la misma idea obligaría a que $b_{3,1} = 0$, por lo que B tendrá una diagonal con tres tipos del mismo signo. Pero en tal caso, la matriz de (13.75), que es $B[1,3|1,3]$, sería diagonal y su determinante tendría signo $\varepsilon_{k-2}\,\varepsilon_k$, por lo que el de $\det A[\alpha|\beta]$ sería ε_k . Minga. En el caso opuesto ($b_{1,1} = b_{3,3} = 0$), un razonamiento semejante lleva a la misma conclusón absurda. Así $b_2 = 0$, lo que establece la validez de la Eq. (13.71). Ya habíamos visto que ello muestra que la Eq. (13.67) es válida para $d(\beta) = k$, lo que completa la inducción. ■

Recordemos el enunciado del Teorema 13.1.5:

Teorema 13.1.5 Sean $A \in \mathcal{M}_{n,m}(\mathbb{R})$ y ε una sucesión de signatura.

1. Para que A sea ε-ERS es suficiente que, para todo $k \in \mathbb{I}_{\min(n,m)}$,

$$\varepsilon_k \det A[\alpha|\beta] > 0 \quad \text{para} \quad \alpha \in Q_{k,n},\ \beta \in Q_{k,m} \quad \text{tales que} \quad d(\alpha) = d(\beta) = 0 \ .$$

2. En particular, A es ETP si $\det A[\alpha|\beta] > 0$ es esos casos.

Demostración. Probemos las desigualdades

$$\varepsilon_k \det A[\alpha|\beta] > 0 \quad \text{para} \quad \alpha \in Q_{k,n} \ , \ \beta \in Q_{k,m} \ , \ k \in \mathbb{I}_{\text{mín}(n,m)} \ , \tag{13.76}$$

por inducción en k. Cuando $k = 1$, esto es trivial porque $d(\alpha) = d(\beta) = 0$ para $\alpha \in Q_{1,n}$ y $\beta \in Q_{1,m}$. Asumamos que la Eq. (13.76) es cierta con $k-1$ en lugar de k. Primero fijemos un $\alpha \in Q_{k,n}$ con $d(\alpha) = 0$, y probemos la Eq. (13.76) para este α por inducción en $l = d(\beta)$. Cuando $l = 0$, esto se sigue de la hipótesis del teorema. Supongamos que $\varepsilon_k \det A[\alpha|\gamma] > 0$ siempre que $\gamma \in Q_{k,m}$ y $d(\gamma) \le l-1$. Sea $\beta \in Q_{k,m}$ con $d(\beta) = l$. Entonces existe p tal que

$$\beta_1 < p < \beta_k \ , \quad d(\tau \cup \{\beta_1 , p\}) \le l-1 \quad \text{y} \quad d(\tau \cup \{p , \beta_k\}) \le l-1 \ ,$$

donde $\tau = \{\beta_2 , \dots , \beta_{k-1}\}$. Se sigue de la Eq. (13.65), como en la Eq. (13.72) de la prueba del Teorema 13.1.4,

$$\det A[\omega|\tau \cup \{p\}] \det A[\alpha|\tau \cup \{\beta_1, \beta_k\}] \ =$$

$$\det A[\omega|\tau \cup \{\beta_k\}] \det A[\alpha|\tau \cup \{\beta_1, p\}] + \det A[\omega|\tau \cup \{\beta_1\}] \det A[\alpha|\tau \cup \{p, \beta_k\}]$$

para cualquier $\omega \in Q_{k-1,n}$ tal que $\omega \subseteq \alpha$. Usando las dos hipótesis inductivas vemos que el lado de la derecha es no nulo con signo $\varepsilon_{k-1}\,\varepsilon_k$, mientras que $\det A[\omega|\tau \cup \{p\}]$ en el lado izquierdo es no nulo con signo ε_{k-1}. Por lo tanto la igualdad es consistente sólo cuando $\varepsilon_k \det A[\alpha|\beta] > 0$. Esto prueba la Eq. (13.76) para los $\alpha \in Q_{k,n}$ con $d(\alpha) = 0$. Luego fijamos cualquier $\beta \in Q_{k,m}$ y hacemos una inducción similar sobre $l = d(\alpha)$, dado que el caso $d(\alpha) = 0$ es lo que probamos antes. Hay que usar la Eq. (13.65) para filas, que se deduce de la usual tomando traspuestas de las matrices involucradas. Así podemos concluir que la Eq. (13.76) es cierta en general. ■

13.9 Ejercicios

13.9.1. Sea $A \in \mathcal{M}_n(\mathbb{R})$ triangular inferior. Entonces es TP si se verifica que

$$\det A[\alpha|1, 2, \dots , k] > 0 \quad \text{para cada } k \in \mathbb{I}_n \ \text{ y } \ \text{cada } \ \alpha \in Q_{k,n} \quad \text{con } d(\alpha) = 0 \ .$$

13.9.2. Sea $A \in \mathcal{M}_n(\mathbb{R})$ tal que A es de Jacobi y $A \geqslant 0$ (entradas positivas). Probar que exite un $\xi \in \mathbb{R}$ tal que $\xi\, I + A$ es TP.

13.9.3. Sea $A \in \mathcal{M}_n(\mathbb{R})$ la matriz del Ejemplo 11.2.15. Porbar que, si $\lambda \in \mathbb{R}$, entonces

$$\lambda\, I + A \quad \text{es TP} \qquad \iff \qquad \lambda \ge 1 \ .$$

Cotejar las propiedades de sus autovalores y autovectores con los resultados desarrollados en las secciones 13.5 y 13.6.

13.9.4. Probar detalladamente el Teorema 13.2.1 que decía:
Sean $A \in \mathcal{M}_{n,m}(\mathbb{R})$ y $B \in \mathcal{M}_{m,l}(\mathbb{R})$. Probar que entonces

1. Si A es ε_A-RS y B es ε_B-RS, el producto AB es ε-RS, con $\varepsilon = \varepsilon_A \cdot \varepsilon_B$.
2. En este caso, AB se convierte en ε-ERS si
   (a) A es ε_A-ERS y $\operatorname{rk} B = l$, o si
   (b) $\operatorname{rk} A = n$ y B es ε_B-ERS.
3. Si A y B son ETP, tambien lo es AB.

13.9.5. Si dos núcleos $L, M : \Gamma \to \Lambda$ son TPs y $d\sigma(\cdot)$ es una medida positiva en Γ, entonces

$$K(u,v) := \int_T L(s,u)M(s,v)d\sigma(s) \ , \quad \text{para} \ \ u,v \in \Lambda \ , \quad \text{es un núcleo TP.} \tag{13.77}$$

Se sugiere replantearlo para que se pueda deducir del Teorema 13.2.1, en principio para medidas concentradas en finitos átomos.

13.9.6. Verificar la veracidad de los otros $4+10$ items del apartado 13.7.1, donde se muestran los más interesantes ejemplos de matrices TP.

Bibliografía

Libros

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

[2] G. Golub y C. F. Van Loan, Matrix computations (Third edition) Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, 1996.

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

[4] R. Bhatia; *Positive Definite Matrices*, Princeton Series in Applied Mathematics, Princeton University Press, 2006.

[5] O. Cordes, *Spectral theory of linear differential operators and comparison algebras*, London Mathematical Society Lecture Note Series, 76. Cambridge University Press, Cambridge, 1987

[6] K. Hoffman y R. Kunze, Linear algebra, Prentice-Hall Mathematics Series, Prentice-Hall, Inc., Englewood Cliffs, N.J. 1961.

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

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

[9] T. Kato, Perturbation theory for linear operators, Reprint of the 1980 edition, Classics in Mathematics, Springer-Verlag, Berlin, 1995.

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

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

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

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

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

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

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

Papers

[17] T. Ando; *Structure of Operators with Numerical Radius One*, Acta Sci. Math (Szeged) 34 (1973), 11-15.

[18] T. Ando; *Unitarily invariant norms related to the numerical radius*, Linear Algebra and its Applications, In Press, Corrected Proof, Available online 22 April 2005.

[19] T. Ando; *Matrix Young inequalities*, Operator theory in function spaces and Banach lattices, 33–38, Oper. Theory Adv. Appl., 75, Birkhäuser, Basel, 1995.

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

[21] H. Araki *On an inequality of Lieb and Thirring*, Lett. Math. Phys. 19 (1990), no. 2, 167-170.

[22] R. Bhatia y C. Davis, *More matrix forms of the arithmetic-geometric mean inequality*, SIAM J. Matrix Anal. Appl. 14 (1993), no. 1, 132–136.

[23] J. L. Daleckiĭ, S. G. Kreĭn, *Formulas of differentiation according to a parameter of functions of Hermitian operators*, (Russian) Doklady Akad. Nauk SSSR (N.S.) 76, (1951). 13–16.

[24] J. L. Daleckiĭ, S. G. Kreĭn, *Integration and differentiation of functions of Hermitian operators and applications to the theory of perturbations*, (Russian) Voronež. Gos. Univ. Trudy Sem. Funkcional. Anal. 1956 (1956), no. 1, 81–105.

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

[26] O. Hirzallah and F, Kittaneh, *Matrix Young inequalities for the Hilbert-Schmidt norm*, Linear Algebra Appl. 308 (2000), 77-84.

[27] C.R. Johnson, C.K. Li, *Inequalities relating unitarily invariant norms and the numerical radius*, Linear and Multilinear Algebra 23 (1988) 183-191.

[28] Chi-Kwong Li; R. Mathias *The Lidskii-Mirsky-Wielandt theorem - additive and multiplicative versions*, Numer. Math. 81 (1999), no. 3, 377–413.

[29] V. I. Paulsen, S. C. Power and R.R. Smith, *Schur products and matrix completions*, J. Funct. Anal. 85 (1989), 151-178.

[30] E. L. Pekarev, *Shorts of operators and some extremal problems*, Acta Sci. Math. (Szeged) 56 (1992), 147-163.

Índice alfabético

Notaciones y abreviaturas

Se enumeran las principales notaciones y abreviaturas del libro, por orden de aparición:

Capítulo 1

$\mathbb{I}_n = \{1, 2, \dots, n\}$.
$\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$ y $\mathbb{R}_+^* = \{x \in \mathbb{R} : x > 0\}$.
$\mathcal{M}_n(\mathbb{C}) = \mathbb{C}^{n\times n}$ y $\mathcal{M}_{n,m}(\mathbb{C}) = \mathbb{C}^{n\times m}$
$\mathcal{M}_n(\mathbb{R}) = \mathbb{R}^{n\times n}$ y $\mathcal{M}_{n,m}(\mathbb{R}) = \mathbb{R}^{n\times m}$
$\mathcal{G}l(n) = \{A \in \mathcal{M}_n(\mathbb{C}) : A$ es inversible $\}$
$P_A(x) = \det(xI - A) \in \mathbb{C}[x]$ es el polinomio característico de $A \in \mathcal{M}_n(\mathbb{C})$.
$\operatorname{tr} A = \sum\limits_{i=1}^{n} A_{ii}$ para una $A \in \mathcal{M}_n(\mathbb{C})$.
$C_i(A) = (a_{1i}, a_{2i}, \dots, a_{ni}) \in \mathbb{C}^n$ es la i-ésima columna de $A \in \mathcal{M}_{n,m}(\mathbb{C})$.
$F_j(A) = (a_{j1}, a_{j2}, \dots, a_{jm}) \in \mathbb{C}^n$ es la j-ésima fila de $A \in \mathcal{M}_{n,m}(\mathbb{C})$.
$\mathrm{d}\,(A) = (A_{11}, \dots, A_{nn}) \in \mathbb{C}^n$, la diagonal de $A \in \mathcal{M}_n(\mathbb{C})$.
$\operatorname{diag}(a) = \operatorname{diag}(a_1, \dots, a_n) = \begin{bmatrix} a_1 & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & a_n \end{bmatrix} \in \mathcal{M}_n(\mathbb{C})$, para $a \in \mathbb{C}^n$.
$\mathcal{E}_m = \{e_1^{(m)}, \dots, e_m^{(m)}\}$ es la base canónica de $\mathbb{C}^m$.
$\mathbb{1} = \mathbb{1}_n = \sum\limits_{k=1}^{n} e_k^{(n)} = (1, \dots, 1) \in \mathbb{C}^n$.
$\mathbb{E}_n = \mathbb{1}_n \odot \mathbb{1}_n \in \mathcal{M}_n(\mathbb{C})^+$, la matriz de puros unos.
$\operatorname{Gen}\{X\} = \operatorname{Gen}\{x_1, \dots, x_m\}$ es el subespacio generado por $X = \{x_1, \dots, x_m\}$.
$\ker A = \{x \in \mathbb{C}^m : Ax = 0\}$ y $R(A) = A(\mathbb{C}^m) = \operatorname{Im}(A) \subseteq \mathbb{C}^n$, para $A \in \mathcal{M}_{n,m}(\mathbb{C})$.
$\operatorname{rk}(A) = \dim R(A) = \dim \operatorname{Gen}\{C_1(A), \dots, C_m(A)\}$, para $A \in \mathcal{M}_{n,m}(\mathbb{C})$.
$\langle x, y\rangle = \sum\limits_{k=1}^{n} x_k\, \overline{y_k}$, $x, y \in \mathbb{C}^n$.
$\|x\| = \|x\|_2 = \langle x, x\rangle^{1/2} = \left(\sum\limits_{k=1}^{n} |x_k|^2 \right)^{1/2}$ para $x \in \mathbb{C}^n$.
$L(\mathbb{H}, \mathbb{K})$ es el espacio de operadores lineales de $\mathbb{H}$ en $\mathbb{K}$ (dos espcios de Hilbert).
BON : base ortonormal.
$A^* = \overline{A^T} \in \mathcal{M}_{m,n}(\mathbb{C})$ la adjunta de $A \in \mathcal{M}_{n,m}(\mathbb{C})$.
$\mathcal{H}(n) = \{A \in \mathcal{M}_n(\mathbb{C}) : A = A^*\}$, matrices autoadjuntas.
$\mathcal{U}(n) = \{U \in \mathcal{M}_n(\mathbb{C}) : U^*U = I\}$, matrices unitarias.
$\mathcal{N}(n) = \{N \in \mathcal{M}_n(\mathbb{C}) : N^*N = NN^*\}$, matrices normales.
$\mathcal{M}_n(\mathbb{C})^+ = \{A \in \mathcal{M}_n(\mathbb{C}) : A \geq 0\} \subseteq \mathcal{H}(n)$, semidefinidas positivas.
$\mathcal{G}l(n)^+ = \{A \in \mathcal{M}_n(\mathbb{C}) : A > 0\} = \mathcal{G}l(n) \cap \mathcal{M}_n(\mathbb{C})^+$, definidas positivas.

$\sigma(A) = \{\lambda \in \mathbb{C} : \ker(A - \lambda I) \neq \{0\}\}$, el espectro de $A \in \mathcal{M}_n(\mathbb{C})$.
$\lambda(A) = (\lambda_1(A), \dots, \lambda_n(A))$ los n autovalores (con multiplicidad) de $A \in \mathcal{M}_n(\mathbb{C})$.
$w(A) = \text{máx}\{ |\langle Ax, x\rangle| \ : \ x \in \mathbb{C}^n, \ \|x\| = 1 \}$, el radio numérico de $A \in \mathcal{M}_n(\mathbb{C})$.
$\rho(A) = \text{máx}\{ |\lambda| : \lambda \in \sigma(A)\}$, el radio espectral de $A \in \mathcal{M}_n(\mathbb{C})$.
$\|A\|_{sp} = \text{máx}\{\|Ax\| : x \in \mathbb{C}^n, \|x\| = 1\} = \text{mín}\{C \geq 0 : \|Ax\| \leq C\|x\| , \ \ x \in \mathbb{C}^n\}$.
$\|A\|_2^2 = \sum\limits_{i,j=1}^{n} |a_{ij}|^2 = \operatorname{tr}(A^*A)$, la norma Frobenius de $A \in \mathcal{M}_n(\mathbb{C})$.
$A \cong B$ (unitariamente equivalentes) si existe $U \in \mathcal{U}(n)$ tal que $A = U^*BU$.
$\mathcal{TS}(n) = \{ T \in \mathcal{M}_n(\mathbb{C}) : T_{ij} = 0$ para $i \geq j\}$ las triangulares superiores.
$A[I|J] = A_{IJ} = (A_{rl})_{\substack{r\in I\\ l\in J}} \in \mathcal{M}_{k,m}(\mathbb{C})$, para $A \in \mathcal{M}_n(\mathbb{C})$, $I, J \subseteq \mathbb{I}_n$ con $|J| = k$, $|K| = m$.
$A[I|J) = A[I|\mathbb{I}_n \setminus J]$ y $A(I|J] = A[\mathbb{I}_n \setminus I|J]$.
$A_r = A(\{r\}) = \{a_{ij}\}_{i\neq r\neq j} \in \mathcal{M}_{n-1}(\mathbb{C})$, para $A \in \mathcal{M}_n(\mathbb{C})$ y $r \in \mathbb{I}_n$.
QR es la factorización $A = QR$ con $Q \in \mathcal{U}(n)$ y $R \in \mathcal{TS}(n)$ tal que $R_{jj} \geq 0$, para todo $j \in \mathbb{I}_n$.
$x \odot y = xy^* = (x_i\, \overline{y_j})_{\substack{i\in\mathbb{I}_n\\ j\in\mathbb{I}_m}} \in \mathcal{M}_{n,m}(\mathbb{C})$, para $x \in \mathbb{C}^n$ e $y \in \mathbb{C}^m$.
$\mathbf{S}_n = \{\sigma : \mathbb{I}_n \to \mathbb{I}_n$ biyectiva $\}$, el n-grupo simetrico.
L_A y $R_B : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C})$ dadas por $L_A(X) = AX$ y $R_B(X) = XB$, para $X \in \mathcal{M}_n(\mathbb{C})$.

Capítulo 2

$\lambda(A) \in \mathbb{R}^n$ es el vector **creciente** de autovalores de $A \in \mathcal{H}(n)$.
$\mu(A) \in \mathbb{R}^n$ es el vector **decreciente** de autovalores de $A \in \mathcal{H}(n)$.
$\lambda_{\text{mín}}(A) = \lambda_1(A) = \mu_n(A) = \text{mín}\,\sigma(A)$ y $\lambda_{\text{máx}}(A) = \lambda_n(A) = \mu_1(A) = \text{máx}\,\sigma(A)$.

Capítulo 3

$B \leq C \iff \langle B\,x\,, x\rangle \leq \langle C\,x\,, x\rangle$ para todo x unitario en $\mathbb{C}^n$ (con $B, C \in \mathcal{H}(n)$).
$\mathcal{M}_1 = \{x \in \mathcal{M} : \|x\| = 1\}$ para un subespacio $\mathcal{M} \subseteq \mathbb{C}^n$.
$P_{\mathcal{S}} \in \mathcal{M}_n(\mathbb{C})$ es la proyección ortogonal sobre un subespacio $\mathcal{S} \subseteq \mathbb{C}^n$.
$A_{\mathcal{S}} = P_{\mathcal{S}} A P_{\mathcal{S}}\big|_{\mathcal{S}} \in L(\mathcal{S})$, la compresión de $A \in \mathcal{M}_n(\mathbb{C})$ a un subespacio $\mathcal{S} \subseteq \mathbb{C}^n$.
$A_{[k]} = A[\mathbb{I}_k] = \{a_{ij}\}_{i,j\in\mathbb{I}_k} \in \mathcal{M}_k(\mathbb{C})$ y
$A_{(k)} = A(\mathbb{I}_k) = \{a_{ij}\}_{i,j>k} \in \mathcal{M}_{n-k}(\mathbb{C})$, ambos para $A \in \mathcal{M}_n(\mathbb{C})$ y $k \in \mathbb{I}_n$.
$A^{1/2} \in \mathcal{M}_n(\mathbb{C})^+$ es la raiz cuadrada de $A \in \mathcal{M}_n(\mathbb{C})^+$.
$|A| = (A^*A)^{1/2}$, el módulo de $A \in \mathcal{M}_n(\mathbb{C})$.
$s_i(A) = \mu_i(|A|) = \mu_i(A^*A)^{1/2}$, los valores singulares de $A \in \mathcal{M}_n(\mathbb{C})$, para $i \in \mathbb{I}_n$.
$s(A) = (s_1(A), \dots, s_n(A)) = \mu(|A|) \in \mathbb{R}_+^n$ y
$\Sigma(A) = \operatorname{diag}(s(A)) \in \mathcal{M}_n(\mathbb{C})^+$, para $A \in \mathcal{M}_n(\mathbb{C})$.
$A = U|A| = |A^*|U$ es una descomposición polar de $A \in \mathcal{M}_n(\mathbb{C})$ si $U \in \mathcal{U}(n)$.
$A = W\Sigma(A)V^*$ es una descomposición en valores singulares de $A \in \mathcal{M}_n(\mathbb{C})$ si $W, V \in \mathcal{U}(n)$.
$A^+ = \frac{|A|+A}{2}$ y $A^- = \frac{|A|-A}{2}$ son las partes positiva y negativa de $A \in \mathcal{H}(n)$.
$\|A\|_p = \left(\sum\limits_{i=1}^{n} s_i(A)^p\right)^{1/p} = (\operatorname{tr}|A|^p)^{1/p}$ (norma de Schatten) para $A \in \mathcal{M}_n(\mathbb{C})$ y $1 \leq p < \infty$.
$\|A\|_{(k)} = \sum\limits_{i=1}^{k} s_i(A)$ (norma de Ky Fan) para $A \in \mathcal{M}_n(\mathbb{C})$ y $k \in \mathbb{I}_n$.
$|||A|||_N = \max\limits_{N(x)=1} N(Ax)$ la norma matricial inducida en $\mathcal{M}_n(\mathbb{C})$ por una norma N en $\mathbb{C}^n$.
$A \circ B = \big(a_{ij}\ b_{ij}\big)_{\substack{i\in\mathbb{I}_n\\ j\in\mathbb{I}_m}} \in \mathcal{M}_{n,m}(\mathbb{C})$ el producto de Hadamard (o Schur) de $A, B \in \mathcal{M}_{n,m}(\mathbb{C})$.

$\widehat{A} = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix} \cong \begin{bmatrix} \Sigma(A) & 0 \\ 0 & -\Sigma(A) \end{bmatrix} \in \mathcal{H}(2n)$ para $A \in \mathcal{M}_n(\mathbb{C})$.

$\mathcal{M}(A, \mathcal{S}) = \{D \in \mathcal{M}_n(\mathbb{C})^+ : \ D \le A \quad \text{y} \quad R(D) \subseteq \mathcal{S}\}$ para $A \in \mathcal{M}_n(\mathbb{C})^+$.

$\Sigma\,(A, \mathcal{S}) = \max\limits_{\le} \mathcal{M}(A, \mathcal{S})$ el shorted de $A \in \mathcal{M}_n(\mathbb{C})^+$ a un subespacio $\mathcal{S} \subseteq \mathbb{C}^n$.

$A^\dagger$ la seudoinversa de Moore-Penrose de $A \in \mathcal{M}_n(\mathbb{C})$.

$\gamma(A) = \min \left\{ \|Ax\| : \quad x \in \ker A^\perp \quad \|x\| = 1 \right\}$ el módulo mínimo de $A \in \mathcal{M}_n(\mathbb{C})$.

$\langle x, \ y \rangle_A = \langle Ax, \ y \rangle$ para $A \in \mathcal{M}_n(\mathbb{C})^+$ y $x \ , \ y \ \in \ \mathbb{C}^n$.

$\mathcal{P}(A, \ \mathcal{S}) = \{Q \in \mathcal{M}_n(\mathbb{C}) : \ Q^2 = Q, \ AQ = Q^*A \text{ y } R(Q) = \mathcal{S}\}$ para $A \in \mathcal{M}_n(\mathbb{C})^+$.

$A \le^* B$ si $BA^* = AA^*$ y $B^*A = A^*A$ (orden $*$), para $A, B \in \mathcal{M}_n(\mathbb{C})$.

Capítulo 4

$x^\downarrow$ y $x^\uparrow$ los reordenados de $x \in \mathbb{R}^n$ en forma decreciente y creciente.

$\operatorname{tr} x = \langle x, \mathbb{1} \rangle = \sum\limits_{j=1}^{n} x_i$, para $x \in \mathbb{C}^n$.

$x \prec y$ si $y \in \mathbb{R}^n$ mayoriza a $x \in \mathbb{R}^n$.

$x \prec_w y$ (resp. $x \prec^w y$) si $y \in \mathbb{R}^n$ submayoriza (supramayoriza) a $x \in \mathbb{R}^n$.

$A \geqslant B$ si $A_{ij} \ge B_{ij}$ para todo par $i \in \mathbb{I}_n$, $j \in \mathbb{I}_m$, con $A, B \in \mathcal{M}_{n,m}(\mathbb{R})$.

$x \geqslant y$ si $x_i \ge y_i$ para todo $i \in \mathbb{I}_n$, con $x, y \in \mathbb{R}^n$.

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

$\mathcal{DS}\,(n) = \{A \in \mathcal{M}_n(\mathbb{C}) : A \geqslant 0 \ , \ \operatorname{tr} F_i(A) = 1 \text{ y } \operatorname{tr} C_i(A) = 1 \text{ para todo } i \in \mathbb{I}_n\}$.

$x_\sigma = (x_{\sigma(1)}, \dots, x_{\sigma(n)})$, para $\sigma \in \mathbf{S}_n$ y $x \in \mathbb{C}^n$.

$P_\sigma \in \mathcal{U}(n)$ la matriz de permutación dada por $P_\sigma \ x \ = \ x_\sigma$, para $\sigma \in \mathbf{S}_n$ y $x \in \mathbb{C}^n$.

$\mathcal{U}_\mathcal{P}(n) = \{P_\sigma : \sigma \in \mathbf{S}_n \ \} \subseteq \mathcal{U}(n)$.

$\mathbb{I}$ denota un intervalo en $\mathbb{R}$.

$f(\vec{x}) = (f(x_1), \dots, f(x_n)\,) \in \mathbb{R}^n$, para una función $f : \mathbb{I} \to \mathbb{R}$, y un vector $\vec{x} \in \mathbb{I}^n$.

$x \underset{\log}{\prec} y$ log-mayorización (con productos), para $x, y \in {\mathbb{R}_+^*}^n$ $(x, y > 0)$.

$x \underset{\log}{\prec_w} y$ log-mayorización débil, para $x, y \in \mathbb{R}_+^n$.

Capítulo 5

$\mathcal{P}_k(n) = \{P \in \mathcal{H}(n) : P^2 = P \text{ y } \operatorname{rk}(P) = k\}$, los proyectores ortogonales de rango k, para $k \in \mathbb{I}_n$.

$\mathcal{U}_k(n) = \{U \in \mathcal{M}_{n,k}(\mathbb{C}) : U^*\,U = I_k\}$, el espacio de isometrías de $\mathbb{C}^k$ en $\mathbb{C}^n$.

NUI : norma unitariamente invariante.

$g_N : \mathbb{C}^n \to \mathbb{R}_+$ dada por $g_N(x) = N\,(\operatorname{diag}(x)\,)$ para N una NUI en $\mathcal{M}_n(\mathbb{C})$ y $x \in \mathbb{C}^n$.

fgs : función gauge simétrica.

$A \prec B$ si $\mu(A) \prec \mu(B)$, para $A, B \in \mathcal{H}(n)$.

$C_P(A) = PAP + (I - P)A(I - P)$ el pinching de $A \in \mathcal{M}_n(\mathbb{C})$ por $P \in \mathcal{P}_k(n)$.

$C_\mathcal{P}(A) = \sum\limits_{i=1}^{r} P_iAP_i$ el pinching de A por el sistema de proyectores $\mathcal{P} = \{P_1, \dots, P_r\} \subseteq \mathcal{H}(n)$.

$\operatorname{conv}[\mathcal{C}] = \left\{ \ \sum\limits_{k=1}^{m} \lambda_k b_k \ : \ m \in \mathbb{N}, \ b_k \in \mathcal{C}, \ \lambda \in \mathbb{R}^m \text{ y } \lambda \prec (1, 0, \dots, 0) \right\}$, la cápsula convexa de $\mathcal{C}$.

$\mathcal{U}(A) = \{UAU^* : U \in \mathcal{U}(n)\} = \{B \in \mathcal{H}(n) : \mu(B) = \mu(A)\}$, la órbita unitaria de $A \in \mathcal{H}(n)$.

NDUI : norma débilmente unitariamente invariante.

Capítulo 6

$\mathcal{H}_{\mathbb{I}}(n) = \big\{ A \in \mathcal{H}(n) \ : \ \sigma(A) \subseteq \mathbb{I} \ \big\}$.

$f(A)$: el cálculo funcional de $A \in \mathcal{H}_{\mathbb{I}}(n)$ por $f : \mathbb{I} \to \mathbb{R}$.

$e^A = \exp(A) = \sum\limits_{m=0}^{\infty} \frac{A^m}{m!}$, la exponencial de $A \in \mathcal{M}_n(\mathbb{C})$.

$\|f - g\|_{\mathbb{I},\infty} := \sup \big\{ |f(t) - g(t)| : t \in \mathbb{I} \big\}$.

$f^{[1]}(x, y)$ es la primera diferencia dividida de una función $f : \mathbb{I} \to \mathbb{R}$ de clase C^1.

$Dg_{x_0} \in \mathcal{M}_{mn}(\mathbb{C})$ es la derivada o diferencial de $g : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ (U abierto) en $x_0 \in U$.

$f \bullet \gamma(t) = f\big(\gamma(t)\big)$ es la composición de una curva γ y el cálculo funcional por f.

MOP : función monótona de operadores.

$\cup$**OP** : función convexa de operadores.

$\cap$**OP** : función cóncava de operadores.

Capítulo 7

$\mathbb{H}_n = \mathbb{C}^n$ con su producto interno.

$\mathbb{H}_n \otimes \mathbb{H}_k = \{$ funcionales $F : \mathbb{H}_n \times \mathbb{H}_k \to \mathbb{C}$ bilineales $\}$ el producto tensorial.

$x \otimes y = xy^T \in \mathbb{H}_n \otimes \mathbb{H}_k$ es el tensor elemental, dado por $x \otimes y(u, v) = \langle u, \overline{x} \rangle \langle v, \overline{y} \rangle$, $u \in \mathbb{H}_n$, $v \in \mathbb{H}_k$.

$\mathcal{E}_{n,k} = \{ e_i^{(n)} \otimes e_j^{(k)} : i \in \mathbb{I}_n \, , \ j \in \mathbb{I}_k \} \sim \{ E_{ij} \in \mathcal{M}_{n,k}(\mathbb{C}) : i \in \mathbb{I}_n \, , \ j \in \mathbb{I}_k \}$, la BON de $\mathbb{H}_n \otimes \mathbb{H}_k$.

$A \otimes B (x \otimes y) = Ax \otimes By$, $x \in \mathbb{H}_n$, $y \in \mathbb{H}_k$, con $A \in L(\mathbb{H}_n)$ y $B \in L(\mathbb{H}_k)$.

$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \dots & a_{nn}B \end{bmatrix} \in \mathcal{M}_{nk}(\mathbb{C})$, el producto de Kroneker de $A \in L(\mathbb{H}_n)$ y $B \in L(\mathbb{H}_k)$.

$\bigotimes^k \mathbb{H}_n$ es el espacio k-tensorial sobre $\mathbb{H}_n$, el producto tensorial de $\mathbb{H}_n$ por sí mismo k veces.

$x_1 \otimes \dots \otimes x_k \ (u_1, \cdots, u_k) = \prod_{i=1}^k \ \langle u_i, \overline{x}_i \rangle$, los k-tensores elementales.

$\bigotimes^k A : \bigotimes^k \mathbb{H}_m \to \bigotimes^k \mathbb{H}_n$, la potencia k-tensorial de A, dada por $\bigotimes^k A \ (x_1 \otimes \dots \otimes x_k) = Ax_1 \otimes \dots \otimes Ax_k$.

$P_\pi^{(n)} \in \mathcal{U}(\bigotimes^k \mathbb{H}_n)$ dado por $P_\pi^{(n)}(F) \ (x_1, \cdots, x_k) = F(x_{\pi(1)}, \cdots, x_{\pi(k)})$, para $\pi \in \mathbf{S}_n$.

$\Lambda^k \mathbb{H}_n = \big\{ \ F \in \bigotimes^k \mathbb{H}_n \ : \ P_\pi^{(n)} F = \operatorname{sgn}(\pi) \ F \quad \text{para toda} \quad \pi \in \mathbf{S}_k \ \big\}$, el espacio k-alternado sobre $\mathbb{H}_n$.

$\mathbf{P}_k^n = \frac{1}{k!} \sum\limits_{\pi \in \mathbf{S}_k} \operatorname{sgn}(\pi)$, la proyección ortogonal de $\bigotimes^k \mathbb{H}_n$ sobre $\Lambda^k \mathbb{H}_n$.

$x_1 \wedge \dots \wedge x_k = \mathbf{P}_k^n \ (x_1 \otimes \ldots \otimes x_k) = \frac{1}{k!} \sum\limits_{\pi \in \mathbf{S}_k} \operatorname{sgn}(\pi) \ x_{\pi(1)} \otimes \dots \otimes x_{\pi(k)}$, el k tensor alternado.

$\Lambda^k A \in L(\Lambda^k \mathbb{H}_n \, , \ \Lambda^k \mathbb{H}_m)$ dado por $\Lambda^k A \ (x_1 \wedge \dots \wedge x_k) = Ax_1 \wedge \dots \wedge Ax_k$, k-potencia alternada de A.

$Q_{k,n} = \big\{ \alpha = (\alpha_1, \alpha_2, \cdots, \alpha_k) \in \mathbb{I}_n^k \ : \ 1 \le \alpha_1 < \alpha_2 < \dots < \alpha_k \le n \ \big\} \sim \{ J \subseteq \mathbb{I}_n \, : \, |J| = k \}$.

$\alpha' = \mathbb{I}_n \setminus \alpha \in Q_{n-k,n}$, el complemento de un $\alpha \in Q_{k,n}$.

$\mathbf{e}_\alpha^\wedge \ = \ \mathbf{e}_\alpha^{(n) \, \wedge} = e_{\alpha_1}^{(n)} \wedge e_{\alpha_2}^{(n)} \wedge \dots \wedge e_{\alpha_k}^{(n)} \in \Lambda^k \mathbb{H}_n$, para $\alpha \in Q_{k,n}$.

$\mathcal{E}_{k,n}^\wedge = \{ \sqrt{k!} \ \mathbf{e}_\alpha^\wedge : \alpha \in Q_{k,n} \}$, la BON de $\Lambda^k \mathbb{H}_n$.

$\det A \ = \ \sum\limits_{\sigma \in \mathbf{S}_n} \operatorname{sgn}(\sigma) \ \prod\limits_{j=1}^n \ a_{j,\sigma(j)} = \Lambda^n A$, para $A \in \mathcal{M}_n(\mathbb{C})$.

$\operatorname{per} A \ = \ \sum\limits_{\sigma \in \mathbf{S}_n} \ \prod\limits_{j=1}^n \ a_{j,\sigma(j)} \in \mathbb{C}$, la permanente de A.

$V(t) = \Big(t_i^{j-1} \Big)_{i,j \in \mathbb{I}_n} = \begin{bmatrix} 1 & t_1 & \dots & t_1^{n-1} \\ 1 & t_2 & \dots & t_2^{n-1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & t_n & \dots & t_n^{n-1} \end{bmatrix} \in \mathcal{M}_n(\mathbb{C})$, el Vandermonde de $t = (t_1, \dots, t_n) \in \mathbb{C}^n$.

Capítulo 8

$C(A) = \max\limits_{i \in \mathbb{I}_m} \|C_i(A)\|_2$ y $F(A) = \max\limits_{i \in \mathbb{I}_n} \|F_i(A)\|_2$, para $A \in \mathcal{M}_{n,m}(\mathbb{C})$.
$K_N(A) = \max\limits_{N(B)=1} N(A \circ B) = \min\big\{k : N(A \circ B) \leq k\, N(B)\ ,\quad B \in \mathcal{M}_n(\mathbb{C})\big\}$, con N norma en $\mathcal{M}_n(\mathbb{C})$.
$K_A = K_{\|\cdot\|_{sp}}(A)$, para $A \in \mathcal{M}_n(\mathbb{C})$.
φ^* la adjunta de una funcional $\varphi : \mathcal{S} \subseteq \mathcal{M}_n(\mathbb{C}) \to \mathbb{C}$, dada por $\varphi^*(A) = \overline{\varphi(A^*)}$, $A \in \mathcal{S}$.
$\varphi_B : \mathcal{M}_n(\mathbb{C}) \to \mathbb{C}$ dada por $\varphi_B(A) = \langle A, B\rangle = \operatorname{tr}(AB^*)$, con $B \in \mathcal{M}_n(\mathbb{C})$.
$S_J = \big\{C \in \mathcal{M}_n(\mathbb{C}) : c_{ij} = 0 \quad \text{para todo} \quad (i,j) \notin J\big\}$, para $J \subseteq \mathbb{I}_n \times \mathbb{I}_n$.
$J \subseteq \mathbb{I}_n \times \mathbb{I}_n$ cumple (P) si $(i,j) \in J \implies (j,i) \in J$ y también $(i,i) \in J$ para todo $i \in \mathbb{I}_n$.

$$A^{(k)} = \mathbb{E}_k \otimes A = \begin{bmatrix} A & \dots & A \\ \vdots & \vdots & \vdots \\ A & \dots & A \end{bmatrix} \in \mathcal{M}_{kn}(\mathbb{C}), \text{ para } A \in \mathcal{M}_n(\mathbb{C}).$$

Capítulo 9

$\operatorname{Re} A = \frac{A+A^*}{2} \in \mathcal{H}(n)$ e $\operatorname{Im} A = \frac{A-A^*}{2i}$, para $A \in \mathcal{M}_n(\mathbb{C})$.
$S(A,B) = AB + BA \in \mathcal{H}(n)$ es el producto simetrizado de $A, B \in \mathcal{H}(n)$.
$C \preccurlyeq D$ si $C^m \leq D^m$, para todo $m \in \mathbb{N}$, con $C, D \in \mathcal{M}_n(\mathbb{C})^+$.
$A \vee B = \lim\limits_{p\to\infty} \big(A^p + B^p\big)^{\frac{1}{p}} = \min\limits_{\preccurlyeq} \big\{\ C \in \mathcal{M}_n(\mathbb{C})^+ : A \preccurlyeq C \ \text{ y } \ B \preccurlyeq C\ \big\}$, para $A, B \in \mathcal{M}_n(\mathbb{C})^+$.
$f : \mathcal{M}_n(\mathbb{C}) \to \mathbb{R}$ es **clase T** si es continua, $f(XY) = f(YX)$ y $|f(X^{2m})| \leq f([XX^*]^m)$, $\forall\, m, X, Y$.
$A \#_\alpha B = A^{1/2}(A^{-1/2}BA^{-1/2})^\alpha A^{1/2}$, para $\alpha \in [0,1]$, $A, B \in \mathcal{M}_n(\mathbb{C})^+$.
$A\#B = A\#_{\frac{1}{2}}B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$.

Capítulo 10

$W(A) = \{\langle Ax, x\rangle \ :\ x \in \mathbb{C}^n,\ \|x\| = 1\ \}$ es el rango numérico de $A \in \mathcal{M}_n(\mathbb{C})$.

Capítulo 11

$\mathrm{MP}_{n,m} = \{A \in \mathcal{M}_{n,m}(\mathbb{R}) : A \geqslant 0\}$, matrices de entradas positivas.
$\mathrm{MEP}_{n,m} = \{A \in \mathcal{M}_{n,m}(\mathbb{R}) : A > 0\}$, matrices de entradas estrictamente positivas.
$V_n = \{(p,q) \in \mathbb{I}_n^2 : p \neq q\}$.
FC : matriz fuertemente convexa.

Capítulo 12

$A/[\alpha|\beta] = A(\alpha|\beta) - A(\alpha|\beta] \cdot A[\alpha|\beta]^{-1} \cdot A[\alpha|\beta) \in \mathcal{M}_{n-k}(\mathbb{C})$, el complemento de Schur de $A \in \mathcal{M}_n(\mathbb{C})$.
$\operatorname{sgn}(\alpha) = \prod\limits_{i=1}^{k} (-1)^{\alpha_i - i} = (-1)^r$ con $r = \operatorname{tr}\alpha \ - \ \frac{k(k+1)}{2}$, para un $\alpha \in Q_{k,n}$.
$J_n = \operatorname{diag}\big(1, -1, 1, -1, \dots, (-1)^{n-1}\big) \in \mathcal{U}(n)$.
$\operatorname{sgn} \frac{\alpha}{\alpha\cup\omega} = \operatorname{sgn}(\gamma)$, donde γ manda α al principio de $\alpha \cup \omega$.

Capítulo 13

$\varepsilon = (\varepsilon_i)_{i\in\mathbb{N}} \in \{-1,1\}^{\mathbb{N}}$ es una sucesión de signatura.
Si τ es otra sucesión de signatura, llamaremos $\tau\varepsilon = (\tau_i\ \varepsilon_i)_{i\in\mathbb{N}}$.

A es ε-**RS** si es de signo regular con signatura ε.

A es ε-**ERS** si es estrictamente de signo regular con signatura ε.

A es **TP** si es totalmente positiva, o sea que A es ε-RS respecto de la sucesión $\varepsilon \equiv 1$.

A es **ETP** si es estrictamente totalmente positiva, o sea que A es ε-ERS respecto de la sucesión $\varepsilon \equiv 1$.

$d(\alpha) = \alpha_k - \alpha_1 - (k-1) = \sum\limits_{i \in \mathbb{I}_{k-1}} \alpha_{i+1} - \alpha_i - 1$, es la dispersión de $\alpha \in Q_{k,n}$.

LU-factorización: $A = LU$ con L triangular inferior y U triangular superior.

UL-factorización: $A = UL$ con L triangular inferior y U triangular superior.

$[x_1 , \dots , x_m] = X \in \mathcal{M}_{nm}(\mathbb{C})$ si $C_i(X) = x_i \in \mathbb{C}^n$ para cada $i \in \mathbb{I}_m$.

$A \overset{t}{\geqslant} B$ si $\Lambda^k A \geqslant \Lambda^k B$, i.e. $\det A[\alpha|\beta] \geq \det B[\alpha|\beta]$ para todo $k \in \mathbb{I}_n$ y $\alpha, \beta \in Q_{k,n}$.

A es **OSC** si es TP y una cierta potencia A^p es ETP.

$V_+(x)$ la máxima variación de signos de $x \in \mathbb{R}^n$.

$V_-(x)$ la mínima variación de signos de $x \in \mathbb{R}^n$.

$G_p = \Big(\exp \big[-p\,(i-j)^2 \big] \Big)_{i,j\, \in \mathbb{I}_n} \in \mathcal{M}_n(\mathbb{R})$.

$H_p = \big[\, d_{2j-i} \,\big]_{i,j\, \in \mathbb{I}_n} \in \mathcal{M}_n(\mathbb{R})$, matriz de Hurwitz $p(z) = d_0 z^n + d_1 z^{n-1} + \dots + d_n$.