UNIVERSIDAD NACIONAL DE LA PLATA
FACULTAD DE CIENCIAS EXACTAS
DEPARTAMENTO DE MATEMÁTICA

Trabajo de Tesis Doctoral:

LA DISPERSIÓN ESPECTRAL Y SUS APLICACIONES EN EL ANÁLISIS MATRICIAL Y LA TEORÍA DE OPERADORES

Tesista: Sebastián Gonzalo Zárate

Director: Dr. Pedro G. Massey

Co-director: Dr. Demetrio Stojanoff

2022

A Clau

Agradecimientos

Quisiera comenzar agradeciendole a Clau, por tantos años juntos, por su amor y aliento en todo momento. A mis viejos que me apoyaron para que realice una carrera universitaria y siempre fomentaron mi desarrollo personal.

A mis directores, Pedro y Demetrio, por ser excelentes personas y ayudarme ampliamente en el trabajo que implica realizar un doctorado; también porque me han enseñado muchísimo y me han dado oportunidades para seguir creciendo en este camino.

A Pedro Marun que me ha ayudado mucho a lo largo de estos años, con tantas charlas de matemática y también por brindarme su amistad.

A los "Bibliotekids", con quienes hemos compartido la oficina estos años, gracias por tantas risas mates de por medio.

Además, quiero agradecerles a Fede, Nico, Nico y Agus, por todos estos años de amistad; que sigan ocurriendo a lo largo de los años.

A mi "hermana"Noe, con quien nos hemos hecho muy amigos, gracias por todo este tiempo compartido.

Por otro lado, quisiera dedicarle unas palabras a las instituciones públicas y gratuitas de educación, que han permitido mi formación de manera integral. Acceder a una educación de calidad, pública y sin arancelamiento es crucial para el desarrollo de cada individuo y también del país. Por este motivo, le agradezco a la Facultad de Ciencias Exactas de La Plata, por permitirme realizar mi formación de grado y posgrado, con excelentes docentes que ponen su empeño para enseñar de la mejor forma posible; y a CONICET, por otorgarme una beca para realizar mi doctorado, en particular al Instituto Argentino de Matemática, por la calidez y fraternidad de sus integrantes.

Índice general

1. Preliminares **9**
- 1.1. Mayorización de vectores en $\mathbb{R}^d$ . . . 10
- 1.2. Normas unitariamente invariantes . . . 13
- 1.3. Geometría relativa entre dos subespacios . . . 14
- 1.4. Submayorización de sucesiones c_0 . . . 17
- 1.5. Operadores compactos autoadjuntos y mayorización . . . 19

2. Valores de Ritz de matrices autoadjuntas **21**
- 2.1. Cotas de mayorización mixtas para el error de Rayleigh-Ritz . . . 22
- 2.2. El teorema $\tan\Theta$: cotas a posteriori de error cuadrático mejoradas . . . 26
- 2.3. Aplicaciones: cotas de mayorización a priori para valores de Ritz . . . 32

3. Dispersión espectral de matrices autoadjuntas **37**
- 3.1. Propiedades básicas de la dispersión espectral . . . 37
- 3.2. Una desigualdad clave para la dispersión espectral . . . 40
- 3.3. Reformulaciones de la desigualdad clave . . . 46
- 3.4. Conmutadores generalizados y conjugaciones unitarias . . . 47
- 3.5. Equivalencias entre desigualdades . . . 51

4. Submatrices principales, ángulos y dispersión espectral **55**
- 4.1. Variación de submatrices principales . . . 55
- 4.2. Prueba del Teorema 4.1.2 . . . 57
- 4.3. Prueba del Teorema 4.1.4 . . . 61
- 4.4. Una aplicación: sobre la variación absoluta de los valores de Ritz . . . 65

5. Dispersión espectral de operadores compactos y autoadjuntos **69**
- 5.1. Dispersión espectral de operadores autoadjuntos . . . 69
- 5.2. Una desigualdad clave: el caso compacto . . . 73
- 5.3. Desigualdades de tipo AG y submayorización . . . 74

Introducción

La teoría de perturbaciones constituye un tópico de estudio dentro del análisis matricial y la teoría de operadores (ver [46, 26, 50]). Asociados a este tópico, se encuentran los temas clásicos de álgebra lineal numérica y teoría de la aproximación. En este contexto, podemos pensar en el estudio de la sensibilidad de los valores de Ritz y los cocientes de Rayleigh de matrices autoadjuntas, es decir, los cambios en los autovalores de compresiones de matrices autoadjuntas, que es un campo de investigación sólido y bien establecido en estas áreas (ver [4, 12, 30, 31, 32, 34, 38, 52, 58]).

Podríamos comenzar describiendo el siguiente tema de estudio: dada una matriz compleja y autoadjunta A de tamaño $d \times d$ e isometrías X, Y de tamaño $d \times k$, con rangos $\mathcal{X}$ e $\mathcal{Y}$ respectivamente, es de interés obtener cotas superiores para la variación absoluta de los valores de Ritz

$$|\lambda(X^* A\, X) - \lambda(Y^* A\, Y)| = (\, |\lambda_i(X^* A\, X) - \lambda_i(Y^* A\, Y)|\,)_{i \in \mathbb{I}_k} \in \mathbb{R}^k_{\geq 0}\,, \tag{1}$$

donde $X^* A\, X$ e $Y^* A\, Y$ son matrices complejas y autoadjuntas de tamaño $k \times k$ que se conocen como cocientes de Rayleigh de A y $\lambda(X^* A\, X)$, $\lambda(Y^* A\, Y) \in \mathbb{R}^k$ son los autovalores (contando multiplicidades y reordenados en forma no creciente) también conocidos como valores de Ritz de A.

Usualmente, las cotas para la variación absoluta de los valores de Ritz se obtienen en términos de residuos $R_X = A\, X - P_{\mathcal{X}}\, A\, X$ y $R_Y = A\, Y - P_{\mathcal{Y}}\, A\, Y$, donde $P_{\mathcal{X}}$ y $P_{\mathcal{Y}}$ denotan los proyectores ortogonales sobre los subespacios $\mathcal{X}$ e $\mathcal{Y}$ respectivamente. Notemos que los residuos, miden que tan invariante resulta un subespacio con respecto a A; por ejemplo, en el caso que $\mathcal{X}$ es A-invariante, tenemos que $R_X = 0$.
También se obtienen cotas en términos de los ángulos principales entre los subespacios $\mathcal{X}$ e $\mathcal{Y}$, denotados $\Theta(\mathcal{X}, \mathcal{Y}) \in [0, \pi/2]^k$, que permiten medir la distancia entre dos subespacios; por ejemplo puede considerarse el ángulo más grande entre $\mathcal{X}$ e $\mathcal{Y}$ como distancia entre los subespacios (ver [47] donde se detallan varias métricas angulares).

En particular, las cotas superiores se clasifican de acuerdo a qué parametros se utilizan para acotar la variación de los valores de Ritz (ver [58]). En efecto, las cotas *a priori* o autónomas son aquellas que se obtienen en términos de los ángulos principales; las cotas *a posteriori* o no autónomas son aquellas obtenidas en términos de (los valores singulares de) residuos, y las cotas mixtas son aquellas obtenidas en términos de ángulos principales y residuos. Conviene destacar que las cotas obtenidas a través de los ángulos principales, podrían no estar disponibles de forma directa en la práctica. Por otro lado, las cotas a posteriori están basadas en el cálculo de los valores singulares de las matrices de residuos. Más aún, las cotas basadas en residuos (a posteriori y mixtas) son particularmente convenientes cuando uno de los subespacios, digamos $\mathcal{X}$ es A-invariante (en este caso $R_X = 0$) a diferencia de las cotas a priori.

La formulación mediante el análisis matricial del problema de sensibilidad anterior, permite aplicar esta teoría en diferentes áreas de investigación tales como: comparación de grafos [32] en términos de análisis espectral de los grafos; distinción de señales en el procesamiento de señales, donde los valores de Ritz sirven como signos armónicos para diferenciar subespacios; métodos de elementos finitos (FEM) [30], para la aproximación de subespacios correspondientes a modos fundamentales; y análisis matricial, por ejemplo para cotas para autovalores después de perturbaciones aditivas de la matriz. Además, las cotas para la variación de los valores de Ritz juegan un papel central en el análisis de algoritmos para la aproximación

simultánea de autovalores basados en métodos de Rayleigh-Ritz (ver [46, 50]). El papel crucial de la submayorización en la obtención de cotas para la variación de los valores de Ritz (reconocido en el artículo [32]) es bien sabido; esta relación de preorden parcial es una herramienta poderosa de comparación entre vectores de entradas reales. En este contexto, las cotas en términos de submayorización implican toda una familia de desigualdades con respecto a normas unitariamente invariantes y respecto a la clase de funciones convexas no decrecientes ([37]).

Dada una matriz autoadjunta A de tamaño $d \times d$, podemos mencionar como medida vectorial de la dispersión de sus autovalores a la *dispersión espectral* de A dada por

$$\mathrm{Spr}^+(A) = (\lambda_i(A) - \lambda_{d-i+1}(A))_{i \in \mathbb{I}_{[d/2]}} \in \mathbb{R}^{[d/2]}\,,$$

donde $[d/2]$ denota la parte entera de $d/2$.

Este concepto fue considerado en la literatura por Argentati y Knyazev en [30], como un factor dentro de la cota superior a priori para la variación absoluta de los valores de Ritz de A. Además, los autores propusieron la siguiente conjetura autónoma que continúa abierta

$$|\lambda(X^* A\, X) - \lambda(Y^* A\, Y)| \prec_w \mathrm{Spr}^+(A)\ \mathrm{sen}(\Theta(\mathcal{X}, \mathcal{Y}))\,, \tag{2}$$

donde $\mathrm{sen}(\Theta(\mathcal{X}, \mathcal{Y}))$ denota el vector cuyas entradas son los senos de los ángulos principales entre $\mathcal{X}$ e $\mathcal{Y}$, $\prec_w$ denota la relación de submayorización y el producto de vectores se considera entrada a entrada.

Más aún, en caso que $\mathcal{X}$ sea A-invariante, los autores conjeturaron que

$$|\lambda(X^* A\, X) - \lambda(Y^* A\, Y)| \prec_w \mathrm{Spr}^+(A)\ \mathrm{sen}(\Theta(\mathcal{X}, \mathcal{Y}))^2\,. \tag{3}$$

Estas conjeturas son extensiones naturales de los resultados obtenidos en [31] (para el caso $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = 1$).

Continuando con el desarrollo de problemas de sensibilidad (de los valores de Ritz) y motivados por trabajos previos (ver [4, 30, 31, 32]) recientemente Knyazev y Zhu [33] han formulado una serie de cotas superiores a posteriori y mixtas para la variación absoluta de los valores de Ritz. En este contexto, los autores han conjeturado que: si $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$ son subespacios de dimensión k tales que $\Theta_1(\mathcal{X}\,, \mathcal{Y}) < \frac{\pi}{2}$, entonces

$$|\lambda(X^* A\, X) - \lambda(Y^* A\, Y)| \prec_w \frac{s(P_{\mathcal{Y}}\, R_X) + s(P_{\mathcal{X}}\, R_Y)}{\cos(\Theta(\mathcal{X}\,, \mathcal{Y}))} \qquad \text{y} \tag{4}$$

$$|\lambda(X^* A\, X) - \lambda(Y^* A\, Y)| \prec_w [s(P_{\mathcal{X}+\mathcal{Y}}\, R_X) + s(P_{\mathcal{X}+\mathcal{Y}}\, R_Y)]\ \tan(\Theta(\mathcal{X}\,, \mathcal{Y}))\,, \tag{5}$$

donde el cociente se considera entrada a entrada, la suposición $\Theta_1(\mathcal{X}\,, \mathcal{Y}) < \frac{\pi}{2}$ asegura que el divisor es no nulo y donde $s(Z) \in \mathbb{R}^d$ denota el vector de valores singulares de una matriz compleja Z, contando multiplicidades y ordenados en forma no creciente.

Utilizando técnicas similares a las propuestas por Knyazev y Zhu, relacionadas con desigualdades de submayorización, expondremos una prueba de esta conjetura en la tesis. Más aún, estos resultados impulsan el desarrollo de una serie de desigualdades que permiten obtener cotas a priori para la variación absoluta de los valores de Ritz (ver [38]).

Por ejemplo, es posible obtener para ángulos pequeños ($\Theta_1(\mathcal{X}, \mathcal{Y}) < \pi/4$) la desigualdad

$$|\lambda(X^* A\, X) - \lambda(Y^* A\, Y)| \prec_w (2\sqrt{2})\ \mathrm{Spr}^+(A)\ \mathrm{sen}(\Theta(\mathcal{X}, \mathcal{Y}))\,,$$

cuyo orden de aproximación coincide con el conjeturado por Knyazev y Zhu.

Por otro lado, podríamos pensar en obtener cotas no autónomas para la variación absoluta de los valores de Ritz. En este sentido, Knyazev y Zhu proponen (a modo de conjetura en [33]) una mejora para una versión reciente de Nakatsukasa del teorema de la tangente de Davis y Kahan [17]. Nakatsukasa (2009,

[43]) describe una versión relajada (y mejorada) del teorema de la tangente, que sirve como estimación a posteriori para la tangente de los ángulos entre $\mathcal{X}$ e $\mathcal{Y}$ en términos de residuos como sigue: sea $X_\perp$ una isometría de rango $\mathcal{X}^\perp$, entonces

$$||\tan(\Theta(\mathcal{X},\mathcal{Y}))|| \leq \frac{||R_Y||}{\delta}\,,$$

donde $||\cdot||$ denota una norma unitariamente invariante y δ es una constante de separación de los espectros de Y^*AY y $X_\perp^* A\, X_\perp$.

La propuesta conjeturada por Knyazev y Zhu, establecía que la cota superior $||R_Y||$ podía reemplazarse por $||P_{\mathcal{X}+\mathcal{Y}}\, R_Y||$, donde $P_{\mathcal{X}+\mathcal{Y}}$ denota el proyector ortogonal sobre $\mathcal{X}+\mathcal{Y}$. En este sentido, probaremos no sólo que esto es posible, sino también que la constante de separación $0<\delta$ puede reemplazarse por otra constante $0<\delta\leq\delta'$ de forma que se obtiene la desigualdad

$$||\tan(\Theta(\mathcal{X},\mathcal{Y}))|| \leq \frac{||P_{\mathcal{X}+\mathcal{Y}} R_Y||}{\delta'}\,,$$

que resulta ajustada; incluso en casos donde la desigualdad clásica no lo es.

El desarrollo de alguno de los resultados anteriormente propuestos, sugiere la necesidad de estudiar con detalle desigualdades (de mayorización) que involucren a la dispersión espectral, teniendo como objetivo aplicarlas en la obtención de cotas a priori para la variación absoluta de los valores de Ritz; y también en posibles pruebas para las conjeturas (2) y (3). Curiosamente, la dispersión espectral no ha sido estudiada en la literatura, luego de su introducción en [30]. Por este motivo, desarrollaremos una serie de propiedades y desigualdades de (sub)mayorización para la dispersión espectral, que se relacionan con problemas de interés dentro del análisis matricial.

Por ejemplo, probaremos la siguiente desigualdad de submayorización, que relaciona los valores singulares de un bloque antidiagonal B de una matriz autoadjunta A con la dispersión espectral de A: si $d=n+r$ y consideramos A una matriz autoadjunta de $d\times d$ por bloques

$$A=\begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix}\begin{matrix}\mathbb{C}^n\\ \mathbb{C}^r\end{matrix} \qquad \text{entonces} \qquad 2\,s(B)\prec_w \operatorname{Spr}^+(A)\,. \tag{6}$$

La desigualdad (6) (que resulta ajustada) ha sido clave para el desarrollo de desigualdades que involucran a la dispersión espectral; además, está conectada con el trabajo de Tao [51], donde él prueba que para una matriz semidefinida positiva A por bloques

$$A=\begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix}\begin{matrix}\mathbb{C}^n\\ \mathbb{C}^r\end{matrix} \qquad \text{vale que} \qquad 2\,s_i(B)\leq s_i(A\oplus A) \qquad \text{para} \qquad i\in\mathbb{I}_n\,. \tag{7}$$

Por otro lado, siguiendo los trabajos [6, 51] resulta que la desigualdad (6) está también relacionada con la desigualdad de Zhan para valores singulares de la diferencia de matrices semidefinidas positivas (ver [54, 55, 56]). Motivados por estos resultados, veremos que la desigualdad (6) resulta equivalente a la desigualdad

$$s(A_1-A_2)\prec_w \operatorname{Spr}^+(A_1\oplus A_2)\,, \tag{8}$$

para matrices complejas autoadjuntas A_1 y A_2 de tamaño $d\times d$.

Además, podemos señalar otro problema asociado; resulta que la ecuación (8) puede ser extendida al contexto de conmutadores generalizados como sigue: dadas A_1, A_2 y X matrices complejas de tamaño $d\times d$ tales que A_1 y A_2 son autoadjuntas, entonces

$$s(A_1\,X-X\,A_2)\prec_w s(X)\,\operatorname{Spr}^+(A_1\oplus A_2)\,, \tag{9}$$

donde el producto entre vectores se considera entrada a entrada. Esta última desigualdad, conecta a su vez estas ideas con una serie de trabajos que abordan desigualdades (más generales) para valores singulares de conmutadores generalizados (ver [22, 27, 28, 29, 53]).

Es conveniente remarcar, que las desigualdades propuestas ((6), (8) y (9)), constituyen cotas novedosas para matrices autoadjuntas; este tipo de desigualdades sólo estaban disponibles en el caso semidefinido positivo (sin utilizar la dispersión espectral). En este sentido, la dispersión espectral permite extender desigualdades conocidas para matrices semidefinidas positivas que no son válidas en el caso autoadjunto; a desigualdades de submayorización para matrices autoadjuntas. Si bien las desigualdades que se obtienen son mas débiles, proveen mejores cotas para normas unitariamente invariantes.

Estos problemas que parecen netamente matriciales, están fuertemente vinculados con los problemas de sensibilidad de los valores de Ritz. Por un lado, la desigualdad de Lidskii permite realizar la siguiente estimación:

$$|\lambda(X^*A\,X) - \lambda(Y^*A\,Y)| \prec_w s(X^*A\,X - Y^*A\,Y)\,. \tag{10}$$

Por lo tanto, es posible obtener cotas superiores para la variación absoluta de los valores de Ritz a través de cotas superiores para

$$s(X^*AX - Y^*AY) \in \mathbb{R}^k\,. \tag{11}$$

Motivados por ideas geométricas, dada la isometría X de tamaño $d \times k$ y rango $R(X) = \mathcal{X}$ y el subespacio $\mathcal{Y}$, elegiremos una isometría $Y = Y_r$ conveniente, para obtener cotas superiores para la ecuación (11) en términos de relaciones de submayorización que involucren a los ángulos principales entre $\mathcal{X}$ e $\mathcal{Y}$, y a la dispersión espectral de A. En el desarrollo de estas ideas, tendrán un rol importante las desigualdades (6) y (9), que permitirán, mediante la elección de curvas con ciertas propiedades minimales, obtener las siguientes cotas superiores

$$s(X^*A\,X - Y^*A\,Y) \prec_w \Theta(\mathcal{X},\mathcal{Y})\,\mathrm{Spr}^+(A)\,,$$

y en el caso que $\mathcal{X}$ es A-invariante

$$s(X^*A\,X - Y^*A\,Y) \prec_w \Theta(\mathcal{X},\mathcal{Y})^2\,\mathrm{Spr}^+(A)\,.$$

Una vez obtenidas estas cotas superiores, aplicaremos la desigualdad de Lidskii aditiva (10), para derivar las siguientes cotas superiores para la variación absoluta de los valores de Ritz

$$|\lambda(X^*AX) - \lambda(Y^*AY)| \prec_w \Theta(\mathcal{X},\mathcal{Y})\,\mathrm{Spr}^+(A)\,. \tag{12}$$

Más aún, si $\mathcal{X}$ es A-invariante entonces

$$|\lambda(X^*AX) - \lambda(Y^*AY)| \prec_w \Theta(\mathcal{X},\mathcal{Y})^2\,\mathrm{Spr}^+(A)\,. \tag{13}$$

Si bien estas estimaciones no resuelven las conjeturas planteadas por Argentati y Knyazev, constituyen nuevas desigualdades para la variación absoluta de los valores de Ritz. Más aún, el tipo de desigualdades obtenidas previamente en la literatura ([30],[33]) no involucran de forma directa a la dispersión espectral; de hecho algunas de estas desigualdades sólo son válidas para ciertas elecciones de subespacios $\mathcal{X}$ y no para subespacios arbitrarios.

Paralelamente, es sabido que el desarrollo de desigualdades que involucran operadores compactos es un tópico central en teoría de operadores (ver [2, 6, 9, 24, 27, 29, 54, 55] entre otros). Por esto, y considerando el desarrollo de las desigualdades matriciales anteriores (como (6)), podemos pensar en extrapolar este tipo

de desigualdades al contexto de operadores compactos y autoadjuntos en un espacio de Hilbert separable de dimensión infinita.

Notemos por $\mathcal{H}$ a un espacio de Hilbert separable de dimensión infinita; $L(\mathcal{H})$ al espacio de operadores actuando en $\mathcal{H}$ y $K(\mathcal{H})$ al ideal de operadores compactos en $L(\mathcal{H})$. Para poner las ideas en contexto, describiremos algunas desigualdades para valores singulares de operadores compactos que motivan los resultados posteriores. Así, comenzamos por la desigualdad Aritmético-Geométrica (AG) de Bhatia y Kittaneh [9], que establece lo siguiente: dados dos operadores compactos A y B, entonces

$$2\, s_i(A\, B^*) \leq s_i(A^*A + B^*B)\ , \quad \text{para} \quad i \in \mathbb{N}\,, \tag{14}$$

donde $s(C) = (s_i(C))_{i\in\mathbb{N}}$ denota la sucesión de valores singulares de un operador $C \in K(\mathcal{H})$, ordenados en forma no creciente. Como consecuencia, se obtiene que dados dos operadores acotados $S, C \in L(\mathcal{H})$ tales que $C^*C + S^*S \leq I$ y un operador compacto y positivo $E \in K(\mathcal{H})$ entonces

$$2\, s_i(S\, E\, C^*) \leq s_i(E)\ , \quad \text{para} \quad i \in \mathbb{N}\,, \tag{15}$$

tomando $A = SE^{1/2}$ y $E^{1/2}C^* = B^*$ en la ecuación (14). En este sentido, Corach, Porta y Recht, motivados por su estudio de la geometría en el contexto de álgebra de operadores, obtuvieron en [14] la siguiente desigualdad con respecto a una norma unitariamente invariante $N(\cdot)$

$$N(T) \leq \frac{1}{2} N(STS^{-1} + S^{-1}TS)\,, \tag{16}$$

donde $T \in K(\mathcal{H})$ y $S \in L(\mathcal{H})$ es un operador acotado autoadjunto e inversible. Posteriormente, Bhatia y Davis probaron en [8] la siguiente desigualdad de tipo AG con respecto a una norma unitariamente invariante

$$N(A^*XB) \leq \frac{1}{2} N(AA^*X + XBB^*)\,, \tag{17}$$

donde $X \in K(\mathcal{H})$ y A, $B \in L(\mathcal{H})$ son operadores acotados. En este caso, resulta que las desigualdades (16) y (17) son equivalentes.

Este tipo de desigualdades, han sido estudiadas y extendidas en distintos contextos ([1, 5, 6]); de hecho resultan estar relacionadas con conceptos profundamente geométricos (ver [2, 14, 15, 16]). Por otro lado, las desigualdades de tipo AG están relacionadas con la desigualdad de Zhan para valores singulares (nuevamente, ver [54, 55, 56]), que pueden enunciarse para operadores compactos y positivos como: dados $C, D \in K(\mathcal{H})$ operadores compactos y positivos, entonces

$$s_i(C - D) \leq s_i(C \oplus D)\ , \quad \text{para} \quad i \in \mathbb{N}\,, \quad \text{donde} \quad C \oplus D = \begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix}. \tag{18}$$

De hecho, las desigualdades (18) y (14) resultan equivalentes (ver [6, 55, 56]).

En este contexto, motivados por las ideas matriciales previas (ver [38, 39, 40]) introducimos la noción de *dispersión espectral* de un operador autoadjunto en el álgebra $\mathcal{A} = K(\mathcal{H}) + \mathbb{C}\, I$ formada por perturbaciones compactas de múltiplos del operador identidad actuando en $\mathcal{H}$. A partir de esto, obtendremos desigualdades en términos de (sub)mayorización ($\prec_w$) con respecto a la dispersión espectral en el contexto general de operadores compactos y autoadjuntos.

Para describir la dispersión espectral de un operador compacto y autoadjunto $A \in K(\mathcal{H})$, consideramos la escala espectral completa de A, como la (única) sucesión $\lambda(A) = (\lambda_i(A))_{i\in\mathbb{Z}_0}$ (donde $\mathbb{Z}_0 = \mathbb{Z} \setminus \{0\}$) tal que sus entradas son autovalores de A (o ceros), contando multiplicidades y ordenados de forma que

$$0 \geq \lambda_{-i-1}(A) \geq \lambda_{-i}(A) \quad \text{y} \quad \lambda_{-i}(A) \leq \lambda_{-i-1}(A) \leq 0 \leq \lambda_{i+1}(A) \leq \lambda_i(A) \quad \text{para} \quad i \in \mathbb{N}\,.$$

Entonces la *dispersión espectral* de $A \in K(\mathcal{H})$, notada también $\mathrm{Spr}^+(A) \in c_0(\mathbb{N})$, está dada por

$$\mathrm{Spr}^+(A) = (\,\lambda_i(A) - \lambda_{-i}(A)\,)_{i\in\mathbb{N}}\,.$$

Aplicaremos entonces esta definición para obtener una versión análoga de la desigualdad (6) para operadores compactos y la relacionaremos con las siguientes desigualdades del tipo AG: dados $S, C \in L(\mathcal{H})$ tales que $C^*C + S^*S = P$, donde P es una proyección ortogonal, y dado un operador compacto y autoadjunto $E \in K(\mathcal{H})$ entonces

$$2\,s(SEC^*) \prec_w \mathrm{Spr}^+(E)\,. \tag{19}$$

Esta desigualdad, sustituye a la ecuación (15) en el contexto de submayorización. Además, permite derivar otras desigualdades de tipo AG, que involucren la dispersión espectral; por ejemplo, dados A, $B \in L(\mathcal{H})$ y $E \in K(\mathcal{H})$ un operador compacto y autoadjunto, veremos que

$$s(A\,E\,B^*) \prec_w \frac{1}{2}\,\mathrm{Spr}^+\big(\,(A^*A + B^*B)^{1/2}\,E\,(A^*A + B^*B)^{1/2}\,\big)\,. \tag{20}$$

Observemos que la dispersión espectral de un operador compacto y autoadjunto es en general más grande (entrada a entrada) que los valores singulares; sin embargo, desigualdades de este tipo (como (20)) sólo estaban disponibles para operadores compactos positivos (de hecho las desigualdades entrada a entrada (14) y (18) no son válidas para operadores compactos autoadjuntos) de forma que la dispersión espectral permite extender estas desigualdades al caso de operadores compactos y autoadjuntos. A su vez, este tipo de desigualdades permitiran obtener desigualdades para normas unitariamente invariantes, asociadas a ideales simétricamente normados de $L(\mathcal{H})$ (ver [19]) .

Organización general de la tesis

En el Capítulo 1 daremos los preliminares correspondientes al análisis matricial y teoría de operadores. En particular, describiremos los conceptos de (sub)mayorización de vectores y sucesiones, daremos desigualdades clásicas como las desigualdades de Lidskii y Weyl y mencionaremos la relación entre (sub)mayorización y el concepto de normas unitariamente invariantes. También describiremos la descomposición de Halmos ([20]) para dos subespacios y la aplicaremos para describir rotaciones directas entre dos subespacios $\mathcal{X}$ e $\mathcal{Y}$ de $\mathbb{C}^d$.

El Capítulo 2 lo dedicaremos al estudio de la variación absoluta de los valores de Ritz. En particular, describiremos con detalle las técnicas necesarias para probar las conjeturas (4) y (5) planteadas por Knyazev y Zhu en [33]. Además, en la Sección 2.2 revisaremos el teorema de la tangente de Davis y Kahan [17]; en particular abordaremos la versión relajada de Nakatsukasa [43] y daremos una versión mejorada que permitirá por ejemplo obtener una cota cuadrática para el error a posteriori para aproximación simultanea de autovalores de A mediante valores de Ritz. Sobre el final en la Sección 2.3, definiremos la dispersión espectral de una matriz autoadjunta y describiremos algunas cotas a priori para la variación absoluta de los valores de Ritz, relacionadas con las conjeturas (2) y (3) de Argentati y Knyazev [30].

El desarrollo de esta sección, nos motivará para describir en el Capítulo 3 una serie de desigualdades matriciales que involucran a la dispersión espectral de matrices autoadjuntas. En particular, en la Sección 3.1 describiremos las propiedades básicas de la dispersión espectral, varias de estas con respecto a la relación de (sub)mayorización. El desarrollo de estas propiedades permitirá en la Sección 3.2, describir la desigualdad (6) que resulta clave para el desarrollo de otras desigualdades que involucran a la dispersión espectral. De hecho, desarrollaremos también en esta sección una desigualdad relacionada con la desigualdad de Zhan [54] para matrices positivas y una desigualdad para estimar los ángulos principales entre dos subespacios en términos de la dispersión espectral. En la Sección 3.4 estableceremos la desigualdad (9) y otras desigualdades para conmutadores y conmutadores generalizados. Por último, en la Sección 3.5 estableceremos una serie de equivalencias entre la desigualdad (6) y algunas desigualdades expuestas durante

el capítulo, que demostrarán la versatilidad de estos resultados y también permitirá tener aplicaciones directas en el capítulo posterior.

El Capítulo 4 retoma el estudio de sensibilidad de los valores de Ritz. En la Sección 4.1 establecemos los resultados sobre el proceso de extracción de submatrices $(A, X, \mathcal{Y}) \mapsto Y_r^* A Y_r$, donde $Y_r = Y_r(X, \mathcal{Y})$ está definido por $Y_r = U\,X$ para una rotación directa U de $\mathcal{X}$ en $\mathcal{Y}$. Dichos resultados permitirán en las Secciones 4.2 y 4.3 probar los resultados principales de este capítulo, dedicados a estimar el vector de valores singulares $s(X^*AX - Y_r^*AY_r)$ en términos de los ángulos principales $\Theta(\mathcal{X}, \mathcal{Y})$ y la dispersión espectral $\mathrm{Spr}^+(A)$. Las técnicas que utilizaremos para probar estos resultados se basan en estimaciones matriciales obtenidas en el Capítulo 3. En la Sección 4.4 aplicaremos los resultados previamente desarrollados para obtener las cotas superiores para la variación absoluta de los valores de Ritz (12) y (13). Notemos que estas desigualdades no resuelven las conjeturas (2) y (3). Sin embargo, las cotas obtenidas en las ecuaciones (12) y (13) resultan comparables con las conjeturas anteriores. Más aún, probaremos que las cotas superiores que obtendremos resultan ajustadas.

El Capítulo 5.1 estará dedicado al desarrollo de desigualdades de (sub)mayorización para operadores compactos y autoadjuntos, extendiendo lo expuesto en el Capítulo 3. Primero, introduciremos la definición de dispersión espectral para operadores compactos y autoadjuntos y daremos algunas propiedades básicas. En la Sección 5.2 describiremos una desigualdad formalmente análoga (6), que resultará importante para el desarrollo de las desigualdades posteriores. Además, daremos una versión (usando submayorización) de la desigualdad de Zhan (18) pero para operadores compactos y autoadjuntos. En la Sección 5.3 abordaremos una serie de desigualdades del tipo Aritmético-Geométrica y probaremos, utilizando una desigualdad formalmente análoga a (6), las desigualdades (19) y (20).

Capítulo 1

Preliminares

Comenzaremos mencionando la notación y los resultados clásicos a utilizar; primero daremos los preliminares correspondientes al análisis matricial (ver [7, 25]) y luego los preliminares para teoría de operadores (ver [19]) .

Consideramos $\mathcal{M}_{d,k}(\mathbb{C})$ el espacio de matrices complejas de tamaño $d \times k$ y notamos $\mathcal{M}_{d,d}(\mathbb{C}) = \mathcal{M}_d(\mathbb{C})$ al álgebra de matrices complejas de tamaño $d \times d$. Además utilizaremos con frecuencia los siguientes subconjuntos de $\mathcal{M}_d(\mathbb{C})$:

- $\mathcal{H}(d)$ el conjunto de matrices autoadjuntas;
- $\mathcal{M}_d(\mathbb{C})^+$ el cono de matrices semidefinidas positivas;
- $\mathcal{G}l(d)$ el grupo de matrices inversibles;
- $\mathcal{U}(d)$ el grupo de matrices unitarias;
- $\mathcal{G}l(d)^+ = \mathcal{G}l(d) \cap \mathcal{M}_d(\mathbb{C})^+$.

Para $d \in \mathbb{N}$, notamos $\mathbb{I}_d = \{1, \dots, d\}$. Dado un vector $x \in \mathbb{C}^d$ denotamos por $D_x = \mathrm{diag}(x)$ a la matriz diagonal en $\mathcal{M}_d(\mathbb{C})$ cuya diagonal es x. Dado $x = (x_i)_{i \in \mathbb{I}_d} \in \mathbb{R}^d$ denotamos por $x^\downarrow = (x_i^\downarrow)_{i \in \mathbb{I}_d}$ al vector que se obtiene reordenando las entradas de x en orden no creciente. De forma similar denotamos por $x^\uparrow = (x_i^\uparrow)_{i \in \mathbb{I}_d}$ al vector que se obtiene reordenando las entradas de x en forma no decreciente. Además usamos las notaciones $(\mathbb{R}^d)^\downarrow = \{x \in \mathbb{R}^d \ : \ x = x^\downarrow\}$, $\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} : x \geq 0\}$ y $(\mathbb{R}^d_{\geq 0})^\downarrow = \{x \in \mathbb{R}^d_{\geq 0} \ : \ x = x^\downarrow\}$. Para $r \in \mathbb{N}$, notamos $\mathbb{1}_r = (1, \dots, 1) \in \mathbb{R}^r$. En general utilizaremos la abreviatura BON para referirnos a "base ortonormal" y SON para referirnos a "sistema ortonormal" .

Por otro lado, dado un subespacio $\mathcal{Z} \subset \mathbb{C}^d$, notamos por $\mathcal{L}(\mathcal{Z})$ al espacio de operadores lineales actuando en $\mathcal{Z}$. Dada una matriz $A \in \mathcal{H}(d)$ notamos por $\lambda(A) = (\lambda_i(A))_{i \in \mathbb{I}_d} \in (\mathbb{R}^d)^\downarrow$ al vector de autovalores de A contando multiplicidades y ordenado en forma no creciente. Para $B \in \mathcal{M}_d(\mathbb{C})$, $s(B) = \lambda(|B|)$ denota el vector de valores singulares de B, es decir, los autovalores de $|B| = (B^*B)^{1/2} \in \mathcal{M}_d(\mathbb{C})^+$.

Cabe destacar que dada una matriz $A \in \mathcal{M}_d(\mathbb{C})$ existen unitarios $U, V \in \mathcal{U}(d)$ tales que

$$A = U \, \Sigma_A \, V^* \, ,$$

donde $\Sigma_A \in \mathcal{M}_d(\mathbb{C})^+$ denota la matriz diagonal $\mathrm{diag}(s(A))$. Esta descomposición es llamada *descomposición en valores singulares* de A.

Las operaciones aritméticas entre vectores se realizarán entrada a entrada, es decir, si $x = (x_i)_{i \in \mathbb{I}_d}$ e $y = (y_i)_{i \in \mathbb{I}_d} \in \mathbb{R}^d$ entonces

- $x + y = (x_i + y_i)_{i \in \mathbb{I}_d}$;

- $x\,y = (x_i\,y_i)_{i\in\mathbb{I}_d}$;

y asumiendo que $y_i \neq 0$, para $i \in \mathbb{I}_d$

- $\frac{x}{y} = (x_i/y_i)_{i\in\mathbb{I}_d}$.

Más aún, si asumimos que x, $y \in \mathbb{R}^d$ entonces escribimos $x \leq y$ siempre que $x_i \leq y_i$, para $i \in \mathbb{I}_d$.

1.1. Mayorización de vectores en $\mathbb{R}^d$

En lo que sigue daremos la definición de submayorización y mayorización entre vectores de $\mathbb{R}^d$, así como algunas propiedades elementales y resultados básicos. La mayorización resulta ser un tópico de gran interés dentro del análisis matricial (ver [7, 25]), entre tantas cosas porque permite obtener desigualdades para normas unitariamente invariantes (como la norma espectral).

Definición 1.1.1. Sean x, $y \in \mathbb{R}^d$. Decimos que x está *submayorizado* por y, y notamos $x \prec_w y$, si

$$\sum_{i=1}^{j} x_i^{\downarrow} \leq \sum_{i=1}^{j} y_i^{\downarrow} \quad \text{para} \quad j \in \mathbb{I}_d\,.$$

Si $x \prec_w y$ y $\operatorname{tr} x \overset{\text{def}}{=} \sum_{i=1}^{d} x_i = \sum_{i=1}^{d} y_i = \operatorname{tr} y$, entonces decimos que x está *mayorizado* por y, y notamos $x \prec y$. △

Observación 1.1.2. Sean $x \in \mathbb{R}^d_{\geq 0}$ e $y \in \mathbb{R}^h_{\geq 0}$ dos vectores de entradas no negativas (y de diferentes tamaños). En este caso extendemos la definición de submayorización entre x e y, suma y producto entre x e y en el siguiente sentido:

Sea 0_n el vector nulo de $\mathbb{R}^n$ y $\ell := \max\{d\,,\,h\}$.

1. decimos que x está submayorizado por y si

$$x \prec_w y \quad \text{si} \quad \begin{cases} (x\,,\,0_{h-d}) \prec_w \quad y & \text{para} \quad d < h \\ x \quad \prec_w (y\,,\,0_{d-h}) & \text{para} \quad h < d \end{cases}, \tag{1.1}$$

2. De forma similar definimos $x + y$ y $x\,y \in \mathbb{R}^{\ell}_{\geq 0}$, agregando ceros al final del vector correspondiente, para que las dimensiones de ambos vectores coincidan.

△

Dada $f : [a\,,\,b] \to \mathbb{R}$, donde $[a\,,\,b] \subset \mathbb{R}$ es un intervalo y $z = (z_i)_{i\in\mathbb{I}_d} \in [a\,,\,b]^d$ notamos $f(z) = (f(z_i))_{i\in\mathbb{I}_d} \in \mathbb{R}^d$.

Observación 1.1.3. Sea $[a\,,\,b] \subset \mathbb{R}$ un intervalo y $f : [a\,,\,b] \to \mathbb{R}$ una función convexa. Entonces

1. Si x, $y \in [a\,,\,b]^d$ satisface $x \prec y$, entonces $f(x) \prec_w f(y)$.
2. Si x, $y \in [a\,,\,b]^d$ sólo satisface que $x \prec_w y$ pero f es también no decreciente en $[a\,,\,b]$, entonces $f(x) \prec_w f(y)$.

△

Los resultados a continuación corresponden a múltiples desigualdades clásicas de mayorización, que involucran autovalores y valores singulares (las pruebas detalladas pueden verse en [7, 25]). Comenzamos con algunas desigualdades para valores singulares; dada $A \in \mathcal{M}_d(\mathbb{C})$ notamos por $\mathrm{re}(A) = \frac{A+A^*}{2} \in \mathcal{H}(d)$.

Teorema 1.1.4. *Sean C, $D \in \mathcal{H}(d)$. Entonces,*

1. *Desigualdad de Lidskii aditiva: $\lambda(C) - \lambda(D) \prec \lambda(C-D) \prec \lambda(C) - \lambda^{\uparrow}(D)$;*
2. *$|\lambda(C) - \lambda(D)| \prec_w s(C-D)$;*
3. *sea $\mathcal{P} = \{P_i\}_{i=1}^r$ un sistema de proyectores (proyecciones mutuamente ortogonales en $\mathbb{C}^d$ tales que $\sum_{i=1}^r P_i = I$). Si*
$$\mathcal{C}_{\mathcal{P}}(D) \stackrel{def}{=} \sum_{i=1}^{r} P_i\, D\, P_i \implies \lambda(\mathcal{C}_{\mathcal{P}}(D)) \prec \lambda(D)\ . \tag{1.2}$$
4. *Más aún, si $C, D \in \mathcal{M}_d(\mathbb{C})^+$:*
$$\lambda(C)\,\lambda^{\uparrow}(D) \prec_w \lambda(CD) \prec_w \lambda(C)\,\lambda(D)\,.$$

□

La desigualdad superior del ítem 1. del teorema anterior, es también conocida como desigualdad de Weyl aditiva para autovalores, por lo que muchas veces nos referiremos a dicha desigualdad de este modo.

Teorema 1.1.5. *Sean C, $D \in \mathcal{M}_d(\mathbb{C})$. Entonces,*

1. *desigualdad aditiva de Weyl: $s(C+D) \prec_w s(C) + s(D)$;*
2. *desigualdad multiplicativa de Weyl: $s(CD) \prec_w s(C)\,s(D)$;*
3. *$s(re(C)) \prec_w s(C)$;*
4. *Más aún, si $CD \in \mathcal{H}(d)$ entonces $s(CD) \prec_w s(re(DC))$.*

□

El siguiente lema condensa algunas desigualdades conocidas de mayorización (y submayorización) que son consecuencia de los teoremas anteriores.

Lema 1.1.6. Sean x, y, $z \in \mathbb{R}^d$. Entonces,

1. $x^{\downarrow} + y^{\uparrow} \prec x + y \prec x^{\downarrow} + y^{\downarrow}$;
2. Si z, $w \in (\mathbb{R}^d)^{\downarrow}$, $x \prec z$ y $y \prec w$ entonces $x + y \prec z + w$.

   Más aún, si $x(\cdot)$, $z(\cdot) : [0,1] \to \mathbb{R}^d$ son funciones continuas tales que $x(t) \prec_w z(t) = (z(t))^{\downarrow}$ para $t \in [0,1]$, entonces
$$\int_0^1 x(t)\ dt \prec_w \int_0^1 z(t)\ dt\,.$$

Si además asumimos que x, y, $z \in \mathbb{R}^d_{\geq 0}$ entonces,

3. $x^{\downarrow}\, y^{\uparrow} \prec_w x\, y \prec_w x^{\downarrow}\, y^{\downarrow}$;

4. Si $x \prec_w y$ y $y,\, z \in (\mathbb{R}^d_{\geq 0})^{\downarrow}$ entonces $x\, z \prec_w y\, z$.

Demostración. Las pruebas de estos resultados pueden encontrarse en [7], con excepción de la última parte del ítem 2. Para ver esto, para $n \in \mathbb{N}$ consideramos la partición regular $\{t_0 = 0 < t_1 < \ldots < t_n = 1\}$ de $[0,1]$, de forma que $t_j = \frac{j}{n}$ para $j \in \{0\} \cup \mathbb{I}_n$ y $\Delta_j = t_j - t_{j-1} = \frac{1}{n} = \Delta_n$, para $j \in \mathbb{I}_n$. Por la primera parte del ítem 2. tenemos que, para $n \in \mathbb{N}$,

$$\sum_{j=1}^{n} x(t_j)\, \Delta_n \prec_w \sum_{j=1}^{n} z(t_j)\, \Delta_n\,.$$

Como $x(\cdot)$, $z(\cdot)$ son funciones continuas, resulta que

$$\int_0^1 x(t)\; dt = \lim_n \sum_{j=1}^{n} x(t_j)\, \Delta_n \prec_w \lim_n \sum_{j=1}^{n} z(t_j)\, \Delta_n = \int_0^1 z(t)\; dt\,,$$

donde hemos usado que: si $(u_n)_{n\in\mathbb{N}}$ y $(v_n)_{n\in\mathbb{N}} \in \mathbb{R}^d$ son sucesiones que convergen a u y $v \in \mathbb{R}^d$ respectivamente, y tales que $u_n \prec_w v_n$ para $n \in \mathbb{N}$, entonces $u \prec_w v$. □

Observación 1.1.7. Sean $x, y \in \mathbb{R}^d$. Si $x \leq y$ entonces,

$$x^{\downarrow} \leq y^{\downarrow} \quad \text{y} \quad x \prec_w y\,.$$

En particular, si A, $P \in \mathcal{M}_d(\mathbb{C})$ son tales que P es una proyección, $s(PAP) \leq s(A)$ por lo que $s(PAP) \prec_w s(A)$. △

Definición 1.1.8. *Sean $A \in \mathcal{H}(d)$ y $r \in \mathbb{I}_d$. Definimos $A_r \in \mathcal{M}_{d-1}(\mathbb{C})$ la submatriz principal de A que se obtiene quitando la fila y columa r-ésima de A, es decir, si $A = (a_{ij})_{i,j\in\mathbb{I}_d}$, entonces $A_r = (a_{ij})_{i,j\in\mathbb{I}_d\setminus\{r\}}$.*

Teorema 1.1.9 (Desigualdades de entrelace)**.** *Sean $A \in \mathcal{H}(d)$, $r \in \mathbb{I}_d$ y $A_r \in \mathcal{M}_{d-1}(\mathbb{C})$ la submatriz principal de A obtenida en la Definición 1.1.8. Entonces*

$$\lambda_k(A) \geq \lambda_k(A_r) \geq \lambda_{k+1}(A)\,, \qquad \textit{para cada} \quad k \in \mathbb{I}_{d-1}\,.$$

Es decir que

$$\lambda_1(A) \geq \lambda_1(A_r) \geq \lambda_2(A) \geq \ldots \geq \lambda_{n-1}(A) \geq \lambda_{d-1}(A_r) \geq \lambda_d(A)\,.$$

□

Proposición 1.1.10. Sea $1 \leq k < d$ y sea $E \in \mathcal{M}_{k,(d-k)}(\mathbb{C})$. Entonces

$$\hat{E} = \begin{pmatrix} 0 & E \\ E^* & 0 \end{pmatrix} \in \mathcal{H}(d) \qquad \text{y} \qquad \lambda(\hat{E}) = (s(E), -s(E^*)^{\downarrow}) \in (\mathbb{R}^d)^{\downarrow}\,.$$

□

1.2. Normas unitariamente invariantes

Recordemos la definición de norma unitariamente invariante:

Definición 1.2.1. Una norma N en $\mathcal{M}_d(\mathbb{C})$ es *unitariamente invariante* (NUI) si $N(UAV) = N(A)$, para todo $A \in \mathcal{M}_d(\mathbb{C})$ y $U, V \in \mathcal{U}(d)$. △

Sea $A \in \mathcal{M}_d(\mathbb{C})$. Algunos ejemplos de NUIs son la norma espectral

$$\|A\|_{sp} = s_1(A)\,,$$

y las normas Ky Fan (de Schatten) para $p \geq 1$,

$$||A||_{(p)} = \left(\sum_{j=1}^{d} s_j(A)^p\right)^{1/p}.$$

En particular, para $p = 2$ se tiene la norma Frobenius,

$$||A||_F = \left(\sum_{j=1}^{d} s_j(A)^2\right)^{1/2}.$$

Además, recordamos que N es una NUI *estrictamente convexa* en $\mathcal{M}_d(\mathbb{C})$ si para cualesquiera $A, B \in \mathcal{M}_d(\mathbb{C})$, con $A \neq B$ tales que $N(A) = N(B)$ y $t \in (0,1)$, se tiene que

$$N(t\,A + (1-t)\,B) < N(A) = N(B)\,.$$

Definición 1.2.2. Sea $f : \mathbb{C}^d \to \mathbb{R}$. Diremos que f es una *función gauge simétrica* si:

1. f es una norma sobre $\mathbb{C}^d$.
2. $f(x) = f(|x|)$.
3. Dada una permutación $\sigma : \mathbb{I}_d \to \mathbb{I}_d$, $f(x) = f(x_\sigma) = (f(x_{\sigma_1}), ..., f(x_{\sigma_d}))$.

△

Observación 1.2.3. La relación entre NUIs y funciones gauge simétricas es estrecha; sea N una NUI en $\mathcal{M}_d(\mathbb{C})$, la función

$$g_N : \mathbb{C}^d \to \mathbb{R}_{\geq 0}\,, \quad \text{dada por} \quad g_N(x) = N(\mathrm{diag}(x))\,,$$

resulta una función gauge simétrica.

Recíprocamente, sea $f : \mathbb{C}^d \to \mathbb{R}$ una función gauge simétrica, entonces la función $N_f : \mathcal{M}_d(\mathbb{C}) \to \mathbb{R}_{\geq 0}$ dada por

$$N_f(A) = f(s(A)) \quad \text{para} \quad A \in \mathcal{M}_d(\mathbb{C})\,,$$

es una NUI.

De esta forma si N es una NUI y g_N es la función gauge simétrica asociada, podemos considerar $g_N(s(A)) = N(A)$ para cualquier matriz $A \in \mathcal{M}_d(\mathbb{C})$.

El siguiente resultado relaciona los conceptos de submayorización y NUIs.

Teorema 1.2.4. *Dadas A, $B \in \mathcal{M}_d(\mathbb{C})$ los siguientes enunciados son equivalentes:*

1. *Para toda norma unitariamente invariante N en $\mathcal{M}_d(\mathbb{C})$ tenemos que $N(A) \leq N(B)$.*
2. *$s(A) \prec_w s(B)$.* □

1.3. Geometría relativa entre dos subespacios

A continuación describiremos la definición de ángulos principales y su relación con las rotaciones directas entre subespacios $\mathcal{X}$ e $\mathcal{Y}$ de $\mathbb{C}^d$. Además, daremos una caracterización de la descomposición de Halmos para dos subespacios. Esta descomposición permite describir a $\mathbb{C}^d$ en términos de los subespacios $\mathcal{X}$ e $\mathcal{Y}$.

Ángulos principales y rotaciones directas

Notación 1.3.1. Dado $1 \leq k \leq d$ notamos por $\mathcal{I}(k,d)$ al conjunto de isometrías $X \in \mathcal{M}_{d,\,k}(\mathbb{C})$, es decir, tales que $X^*X = I_k$ donde I_k denota la identidad de $\mathcal{M}_k(\mathbb{C})$. Notemos que $\mathcal{I}(k,d)$ puede ser identificado de forma natural con la variedad compleja de Stiefel. Más aún, si $\mathcal{X} \subset \mathbb{C}^d$ es un subespacio k dimensional de $\mathbb{C}^d$ notamos $\mathcal{I}_{\mathcal{X}}(k,d)$ a las isometrías $X \in \mathcal{M}_{d,\,k}(\mathbb{C})$ con rango $R(X) = \mathcal{X}$.

△

Definición 1.3.2. Sean $\mathcal{X}$, $\mathcal{Y} \subset \mathbb{C}^d$ subespacios tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$. Consideramos $X \in \mathcal{I}_{\mathcal{X}}(k,d)$, $Y \in \mathcal{I}_{\mathcal{Y}}(k,d)$ y $X_\perp \in \mathcal{I}_{\mathcal{X}^\perp}(d-k,d)$.
Los *ángulos principales* $\Theta(\mathcal{X}\,,\,\mathcal{Y}) = (\theta_j)_{j\in\mathbb{I}_k} \in ([0,\pi/2]^k)^\downarrow$ entre $\mathcal{X}$ e $\mathcal{Y}$ están definidos por

$$\cos(\theta_j) = s_{k-j+1}(X^*Y) \quad \text{o también por} \quad \operatorname{sen}(\theta_j) = s_j(Y^*X_\perp) \quad \text{para} \quad j \in \mathbb{I}_k\,, \tag{1.3}$$

(ver [47] para mas detalles). △

En [17] Davis y Kahan introducen la noción fundamental de rotación directa de $\mathcal{X}$ en $\mathcal{Y}$ que describimos a continuación.

Definición 1.3.3. Un unitario $U \in \mathcal{U}(d)$ es una rotación directa de $\mathcal{X}$ en $\mathcal{Y}$ si $U\mathcal{X} = \mathcal{Y}$ y si existen $C_0 \in \mathcal{M}_k(\mathbb{C})^+$, $C_1 \in \mathcal{M}_{d-k}(\mathbb{C})^+$ y $S_0 \in \mathcal{M}_{d-k,k}(\mathbb{C})$ tales que

$$W^*U\,W = \begin{bmatrix} C_0 & -S_0^* \\ S_0 & C_1 \end{bmatrix} \in \mathcal{U}(d)\,, \tag{1.4}$$

para alguna matriz unitaria $W \in \mathcal{U}(d)$, cuyas primeras k columnas forman una BON de $\mathcal{X}$. En este caso podemos escribir $U = e^{i\,Z}$, donde $Z \in \mathcal{H}(d)$ es tal que

$$\lambda(Z) = (\Theta(\mathcal{X}\,,\,\mathcal{Y})^*, -\,\Theta(\mathcal{X}\,,\,\mathcal{Y})^*, 0)^\downarrow \in (\mathbb{R}^d)^\downarrow\,,$$

con $\Theta(\mathcal{X}\,,\,\mathcal{Y})^* = (\theta_i)_{i\in\mathbb{I}_m} \in ((0,\pi/2]^m)^\downarrow$ denotando el vector de ángulos principales positivos y por lo tanto $m \leq [d/2]$.

△

Observación 1.3.4. Las rotaciones directas poseen cierta propiedad de extremalidad, que juega un rol central en el estudio de propiedades métricas en la variedad de Grassmann de todos los subespacios k-dimensionales de $\mathbb{C}^d$. De hecho, Davis y Kahan prueban en [17] que si $V \in \mathcal{U}(d)$ es tal que $V\,\mathcal{X} = \mathcal{Y}$ y U es una rotación directa de $\mathcal{X}$ en $\mathcal{Y}$ entonces

$$2\,\operatorname{sen}(\theta_j/2) = s_j((1-U)|_{\mathcal{X}}) \leq s_j((1-V)|_{\mathcal{X}}) \quad \text{para} \quad j \in \mathbb{I}_d\,. \tag{1.5}$$

Por otro lado si $T \in \mathcal{H}(d)$ es tal que $\lambda_j(T) \in [-\pi,\pi]$, para $j \in \mathbb{I}_d$, y $V = e^{-iT}$ entonces, las desigualdades de entrelace (1.1.9) muestran que $s_j((1-V)|_{\mathcal{X}}) \leq s_j(1-V)$, para $j \in \mathbb{I}_d$.

Como $\lambda_j(T)/2 \in [-\pi/2,\pi/2]$ resulta que $|\operatorname{sen}(\lambda_j(T)/2)| = \operatorname{sen}(|\lambda_j(T)|/2)$; entonces, un argumento directo muestra que $s_j(1-V) = 2\,\operatorname{sen}(s_j(T)/2)$, para $j \in \mathbb{I}_d$. Luego, se puede deducir del argumento de Davis y Kahan (1.5) que

$$\theta_j = s_j(Z) \leq s_j(T) \quad \text{para} \quad j \in \mathbb{I}_k\,. \tag{1.6}$$

△

En términos generales, la propiedad que caracteriza a las rotaciones directas $U \in \mathcal{U}(d)$ de $\mathcal{X}$ en $\mathcal{Y}$ es el hecho de que son *rotaciones minimales* de $\mathbb{C}^d$, que mueven $\mathcal{X}$ en $\mathcal{Y}$.

Descomposición de Halmos

Comenzaremos haciendo una breve descripción de la descomposición de Halmos para dos subespacios (ver [13, 20]). El objetivo es obtener un refinamiento en la estructura por bloques (1.4) de una rotación directa U de $\mathcal{X}$ en $\mathcal{Y}$.

Dados dos subespacios $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$ tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$, podemos descomponer a $\mathbb{C}^d$ de la siguiente forma:

$$\mathbb{C}^d = (\mathcal{X} \cap \mathcal{Y}) \oplus (\mathcal{X} \cap \mathcal{Y}^\perp) \oplus (\mathcal{X}^\perp \cap \mathcal{Y}) \oplus (\mathcal{X}^\perp \cap \mathcal{Y}^\perp) \oplus \mathcal{G}\,, \tag{1.7}$$

donde

$$\mathcal{G} \stackrel{\text{def}}{=} \left[(\mathcal{X} \cap \mathcal{Y}) \oplus (\mathcal{X} \cap \mathcal{Y}^\perp) \oplus (\mathcal{X}^\perp \cap \mathcal{Y}) \oplus (\mathcal{X}^\perp \cap \mathcal{Y}^\perp) \right]^\perp,$$

representa la *parte genérica* entre $\mathcal{X}$ e $\mathcal{Y}$.

Más aún, como estamos en un contexto de dimensión finita, las dimensiones de $\mathcal{X}^\perp \cap \mathcal{Y}$ y $\mathcal{X} \cap \mathcal{Y}^\perp$ coinciden; en efecto,

$$\dim(\mathcal{X} + \mathcal{Y}^\perp) = \dim(\mathcal{X}) + \dim(\mathcal{Y}^\perp) - \dim(\mathcal{X} \cap \mathcal{Y}^\perp) = d - \dim(\mathcal{X} \cap \mathcal{Y}^\perp)\,.$$

Luego

$$\dim((\mathcal{X} + \mathcal{Y}^\perp)^\perp) = \dim(\mathcal{X}^\perp \cap \mathcal{Y}) = d - (d - \dim(\mathcal{X} \cap \mathcal{Y}^\perp))\,,$$

de forma que

$$p \stackrel{\text{def}}{=} \dim \mathcal{X} \cap \mathcal{Y}^\perp = \dim \mathcal{X}^\perp \cap \mathcal{Y}\,.$$

Además como prueba Halmos (también puede verse en [13, Teorema 1.1]), tenemos que

$$r \stackrel{\text{def}}{=} \dim \mathcal{X} \cap \mathcal{G} = \dim \mathcal{Y} \cap \mathcal{G}\,.$$

Por otro lado, como $\mathcal{X}^\perp \cap \mathcal{G}$ y $\mathcal{Y} \cap \mathcal{G}$ son complementos (algebraicos) de $\mathcal{X} \cap \mathcal{G}$ como subespacios de $\mathcal{G}$, resulta que $\dim \mathcal{X}^\perp \cap \mathcal{G} = \dim \mathcal{Y} \cap \mathcal{G}$.

Por lo tanto tenemos que

$$\mathcal{X} = (\mathcal{X} \cap \mathcal{Y}) \oplus (\mathcal{X} \cap \mathcal{Y}^\perp) \oplus (\mathcal{X} \cap \mathcal{G}) \quad \text{y} \quad \mathcal{Y} = (\mathcal{X} \cap \mathcal{Y}) \oplus (\mathcal{X}^\perp \cap \mathcal{Y}) \oplus (\mathcal{Y} \cap \mathcal{G})\,.$$

Basados en esta descomposición de Halmos introducimos tres subespacios que nos serán útiles mas adelante:

$$\mathcal{S}_1 = (\mathcal{X} \cap \mathcal{Y}^\perp) \oplus (\mathcal{X} \cap \mathcal{G})\,,\ \mathcal{S}_2 = (\mathcal{X}^\perp \cap \mathcal{Y}) \oplus (\mathcal{X}^\perp \cap \mathcal{G}) \quad \text{y} \quad \mathcal{S}_3 = (\mathcal{X}^\perp \cap \mathcal{Y}) \oplus (\mathcal{Y} \cap \mathcal{G})\,.$$

Estos subespacios permiten reescribir a los subespacios $\mathcal{X}$ e $\mathcal{Y}$ como sigue:

$$\mathcal{X} = (\mathcal{X} \cap \mathcal{Y}) \oplus \mathcal{S}_1 \quad \text{y} \quad \mathcal{Y} = (\mathcal{X} \cap \mathcal{Y}) \oplus \mathcal{S}_3\,.$$

Observación 1.3.5. A partir de la descomposición de Halmos (1.7), podemos decir que una rotación directa U, fija el subespacio común $\mathcal{X} \cap \mathcal{Y}$ y el subespacio $\mathcal{X}^\perp \cap \mathcal{Y}^\perp$. Por lo tanto, U actúa de una manera trivial (es decir, como la identidad) en $(\mathcal{X} \cap \mathcal{Y}) \oplus (\mathcal{X}^\perp \cap \mathcal{Y}^\perp)$. Por otro lado, U rota el subespacio $\mathcal{S}_1$ sobre $\mathcal{S}_3$.

Con el fin de obtener una representación matricial por bloques de U, notamos además que

$$\big[\,(\mathcal{X}\cap\mathcal{Y})\oplus(\mathcal{X}^\perp\cap\mathcal{Y}^\perp)\,\big]^\perp=\mathcal{S}_1\oplus\mathcal{S}_2\quad\text{ con }\quad\dim\mathcal{S}_1=\dim\mathcal{S}_2\,.$$

Por lo tanto para describir U, es natural considerar la descomposición ortogonal $\mathbb{C}^d=(\mathcal{X}\cap\mathcal{Y})\oplus\mathcal{S}_1\oplus\mathcal{S}_2\oplus(\mathcal{X}^\perp\cap\mathcal{Y}^\perp)$. En lo que sigue describiremos con detalle las ideas anteriores.

Consideremos los ángulos $\Theta'=\Theta(\mathcal{S}_1\,,\,\mathcal{S}_3)\in[0,\pi/2]^{p+r}$, y las matrices diagonales

$$C=\operatorname{diag}(\,\cos(\Theta')\,)\qquad\text{y}\qquad S=\operatorname{diag}(\,\operatorname{sen}(\Theta')\,)\in\mathcal{M}_{p+r}(\mathbb{C})^+\,.$$

De la descomposición (1.7) observamos que los subespacios $\mathcal{X}\cap\mathcal{Y}^\perp$ y $\mathcal{X}^\perp\cap\mathcal{Y}$ son ortogonales; luego los ángulos entre dichos subespacios p-dimensionales son $\frac{\pi}{2}\,\mathbb{1}_p$. Por lo tanto, como U rota $\mathcal{S}_1$ en $\mathcal{S}_3$, debe rotar $\mathcal{X}\cap\mathcal{Y}^\perp$ en $\mathcal{X}^\perp\cap\mathcal{Y}$, y en consecuencia las primeras p entradas de Θ' se corresponden con dichos ángulos.

Así, una matriz $U\in\mathcal{U}(d)$ es una rotación directa de $\mathcal{X}$ en $\mathcal{Y}$ si existe una BON $\mathcal{B}$ de $\mathbb{C}^d$ obtenida por yuxtaposición de bases ortonormales para $\mathcal{X}\cap\mathcal{Y}$, $\mathcal{S}_1$, $\mathcal{S}_2$ y $\mathcal{X}^\perp\cap\mathcal{Y}^\perp$, de forma que la representación matricial por bloques de U con respecto a $\mathcal{B}$ y a estos subespacios es:

$$\begin{pmatrix} I & 0 & 0 & 0\\ 0 & C & -S & 0\\ 0 & S & C & 0\\ 0 & 0 & 0 & I\end{pmatrix}\begin{matrix}\mathcal{X}\cap\mathcal{Y}\\ \mathcal{S}_1\\ \mathcal{S}_2\\ \mathcal{X}^\perp\cap\mathcal{Y}^\perp\end{matrix}\,. \tag{1.8}$$

Más aún, si $s=\dim\mathcal{X}\cap\mathcal{Y}$, entonces

$$\Theta'=\Big(\,\frac{\pi}{2}\,\mathbb{1}_p\,,\,\theta_1\,,\,\ldots\,,\,\theta_r\,\Big)\qquad\text{y}\qquad\Theta(\mathcal{X}\,,\,\mathcal{Y})=(\Theta',0_s)\,, \tag{1.9}$$

donde $\theta_1,...,\theta_r\in(0,\frac{\pi}{2})$, representan los ángulos entre $\mathcal{X}\cap\mathcal{G}$ y $\mathcal{Y}\cap\mathcal{G}$.

Por otro lado, a partir de la representación matricial por bloques para U, podemos obtener una representación para la matriz autoadjunta Z tal que $e^{i\,Z}=U$. En efecto, consideremos $D_{\Theta'}$ la matriz diagonal cuya diagonal es Θ'. En vistas de la Definición 1.3.3 podemos observar que los autovalores de Z cumplen que $\lambda(Z)=(\Theta',-\Theta',0)^\downarrow\in(\mathbb{R}^d)^\downarrow$ (porque Θ' denota el vector de ángulos principales positivos). Combinando este hecho con la Proposición 1.1.10, podemos pensar que Z tiene la siguiente estructura antidiagonal con respecto a $\mathcal{B}$:

$$\begin{pmatrix} 0 & 0 & 0 & 0\\ 0 & 0 & D_{\Theta'} & 0\\ 0 & D_{\Theta'} & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{matrix}\mathcal{X}\cap\mathcal{Y}\\ \mathcal{S}_1\\ \mathcal{S}_2\\ \mathcal{X}^\perp\cap\mathcal{Y}^\perp\end{matrix}\,. \tag{1.10}$$

En este caso, puede corroborarse que $e^{i\,Z}=U$ (respetando la base $\mathcal{B}$), recordando que $e^{i\,Z}=\sum_{k=0}^{\infty}\frac{(i\,Z)^k}{k!}$.

Por último cabe destacar que esta observación (y la descomposición de Halmos) permiten concluir que siempre es posible construir rotaciones directas en $\mathbb{C}^d$. △

El siguiente resultado [47, Teorema 3.1], describe una versión de submayorización de la desigualdad triangular para ángulos principales.

Teorema 1.3.6. *Sean $\mathcal{X},\mathcal{Y},\mathcal{Z}\subset\mathbb{C}^d$ subespacios tales que $\dim(\mathcal{X})=\dim(\mathcal{Y})=\dim(\mathcal{Z})=k$, entonces*

$$\Theta(\mathcal{X},\mathcal{Y})\prec_w\Theta(\mathcal{X},\mathcal{Z})+\Theta(\mathcal{Z},\mathcal{Y})\,.$$

A continuación describiremos los preliminares correspondientes a teoría de operadores (para mas detalles ver [19, 36, 49]).

Notación. Notaremos $\mathcal{H}$ a un espacio de Hilbert separable de dimensión infinita y $L(\mathcal{H})$ al álgebra de operadores lineales sobre $\mathcal{H}$. En este caso, $K(\mathcal{H}) \subset L(\mathcal{H})$ denota el ideal de operadores compactos y $\mathcal{U}(\mathcal{H}) \subset L(\mathcal{H})$ denota el grupo de operadores unitarios actuando en $\mathcal{H}$. Además notaremos por $L(\mathcal{H})^{sa}$ y $K(\mathcal{H})^{sa}$ a los subespacios reales de operadores autoadjuntos y compactos y autoadjuntos, respectivamente. También, notaremos por $L(\mathcal{H})^+$ al cono de operadores positivos y $K(\mathcal{H})^+ = L(\mathcal{H})^+ \cap K(\mathcal{H})$.

Como antes, notaremos $\mathbb{I}_k = \{1, ..., k\} \subset \mathbb{N}$, para $k \in \mathbb{N}$ y $\mathbb{I}_\infty = \mathbb{N}$.

1.4. Submayorización de sucesiones c_0

En lo que sigue utilizaremos $\mathbb{M}$ para denotar $\mathbb{N}$ o $\mathbb{Z}_0 \stackrel{\text{def}}{=} \mathbb{Z} \setminus \{0\}$. Consideramos el espacio $c_0(\mathbb{M})$ de sucesiones $a = (a_n)_{n\in\mathbb{M}}$ tales que para todo $\epsilon > 0$, el conjunto $\{n \in \mathbb{M} : |a_n| > \epsilon\}$ es finito. Utilizaremos ambos conjuntos de índices, $\mathbb{N}$ y $\mathbb{Z}_0$, para describir reordenamientos de valores singulares y autovalores de operadores compactos autoadjuntos (como en [36]). Dado $k \in \mathbb{N}$, denotamos por $\mathcal{P}_k(\mathbb{M})$ al conjunto de subconjuntos finitos de $\mathbb{M}$ de exactamente k elementos. Además, notamos $\mathbb{R}_{\leq 0} = \{x \in \mathbb{R} : x \leq 0\}$.

Definición 1.4.1. Sea $a = (a_n)_{n\in\mathbb{M}} \in c_0(\mathbb{M}) \cap \mathbb{R}^{\mathbb{M}}$ una sucesión real. Definimos los siguientes reordenamientos asociados con a:

1. Si $a \in \mathbb{R}^{\mathbb{M}}_{\geq 0}$, definimos $a^\downarrow \in c_0(\mathbb{N}) \cap \mathbb{R}^{\mathbb{N}}_{\geq 0}$ determinado de forma única (recursiva) por las condiciones

$$a^\downarrow_1 = \max_{k\in\mathbb{M}} a_k \quad \text{y} \quad \sum_{k=1}^{n} a^\downarrow_k = \max_{\mathbb{F}\in\mathcal{P}_n(\mathbb{M})} \sum_{i\in\mathbb{F}} a_i \quad \text{para todo} \quad n \in \mathbb{N}\,.$$

   Por ejemplo, si $a^\downarrow_1 = a_r$, entonces $a^\downarrow_2 = \max_{k\neq r} a_k$ y así sucesivamente. Luego $a^\downarrow_{n+1} \leq a^\downarrow_n$ para todo $n \in \mathbb{N}$.

2. Si $a \in \mathbb{R}^{\mathbb{M}}_{\leq 0}$, definimos $a^\uparrow = -(-a)^\downarrow \in \mathbb{R}^{\mathbb{N}}_{\leq 0}$. Por lo tanto

$$\sum_{k=1}^{n} a^\uparrow_k = \min_{\mathbb{F}\in\mathcal{P}_n(\mathbb{M})} \sum_{i\in\mathbb{F}} a_i \quad \text{y} \quad a^\uparrow_{n+1} \geq a^\uparrow_n \quad \text{para todo} \quad n \in \mathbb{N}\,.$$

3. En general, para una sucesión $a \in c_0(\mathbb{M}) \cap \mathbb{R}^{\mathbb{M}}$ definimos a^+ y $a^- \in c_0(\mathbb{M}) \cap \mathbb{R}^{\mathbb{M}}$ dados por:

$$a^+_k = \begin{cases} a_k & \text{si} \quad a_k \geq 0 \\ 0 & \text{si} \quad a_k < 0 \end{cases} \quad \text{para} \quad k \in \mathbb{M}\,,$$

   y $a^- = -(-a)^+ = a - a^+$, la parte negativa de a.

4. Finalmente definimos $a^{\uparrow\downarrow} \in c_0(\mathbb{Z}_0)$ dado por

$$a^{\uparrow\downarrow}_n = \begin{cases} (a^-)^\uparrow_{-n} & \text{si} \quad n < 0 \\ (a^+)^\downarrow_n & \text{si} \quad n > 0 \end{cases} \quad \text{para} \quad n \in \mathbb{Z}_0\,.$$

   Cabe destacar que incluso si a tiene una cantidad finita de entradas positivas (o negativas), las partes positiva y negativa de $a^{\uparrow\downarrow}$ son (por construcción) sucesiones infinitas. △

Observación 1.4.2. Notemos que los reordenamientos de sucesiones $a \in c_0(\mathbb{M})$ dados en la Definición 1.4.1 no poseen entradas nulas cuando a tiene una cantidad infinita de entradas positivas y negativas. Por lo tanto, en general, $a^{\uparrow\downarrow}$ no es una permutación de a. Por otro lado, en caso que $a \in c_0(\mathbb{M})$ tenga sólo k valores positivos (o negativos), entonces $a_i^{\uparrow\downarrow} = 0$ para $k < i \in \mathbb{N}$ (o $-k < -i$). △

Definición 1.4.3. Sean $a = (a_n)_{n\in\mathbb{M}_1} \in c_0(\mathbb{M}_1)$, $b = (b_n)_{n\in\mathbb{M}_2} \in c_0(\mathbb{M}_2)$ sucesiones reales, donde $\mathbb{M}_j = \mathbb{N}$ o $\mathbb{M}_j = \mathbb{Z}_0$, para $j = 1, 2$ (pero posiblemente $\mathbb{M}_1 \neq \mathbb{M}_2$). Decimos que a está submayorizado por b y lo notamos

$$a \prec_w b \quad \text{si} \quad \sum_{k=1}^{n} a_k^{\uparrow\downarrow} \leq \sum_{k=1}^{n} b_k^{\uparrow\downarrow} \quad \text{para todo} \quad n \in \mathbb{N}\,.$$

Decimos que a está mayorizado por b y notamos

$$a \prec b \quad \text{si} \quad a \prec_w b \quad \text{y} \quad -a \prec_w -b\,. \qquad \triangle$$

Observación 1.4.4. Con la notación de la Definición 1.4.3 y usando la Definición 1.4.1 puede verse que

$$a \prec_w b \quad \text{si} \quad \max_{\mathbb{F}\in\mathcal{P}_n(\mathbb{M}_1)} \sum_{i\in\mathbb{F}} a_i^+ \leq \max_{\mathbb{F}\in\mathcal{P}_n(\mathbb{M}_2)} \sum_{i\in\mathbb{F}} b_i^+\,, \quad n \in \mathbb{N}\,.$$

Más aún, también obtenemos que $a \prec b$ si

$$\max_{\mathbb{F}\in\mathcal{P}_n(\mathbb{M}_1)} \sum_{i\in\mathbb{F}} a_i^+ \leq \max_{\mathbb{F}\in\mathcal{P}_n(\mathbb{M}_2)} \sum_{i\in\mathbb{F}} b_i^+ \quad \text{y} \quad \min_{\mathbb{F}\in\mathcal{P}_n(\mathbb{M}_1)} \sum_{i\in\mathbb{F}} a_i^- \geq \min_{\mathbb{F}\in\mathcal{P}_n(\mathbb{M}_2)} \sum_{i\in\mathbb{F}} b_i^-\,, \quad n \in \mathbb{N}\,.$$

△

Además consideramos las siguientes construcciones. Primero, dadas dos sucesiones $a = (a_n)_{n\in\mathbb{N}} \in c_0(\mathbb{N})$ y $b = (b_n)_{n\in\mathbb{N}} \in c_0(\mathbb{N})$, construímos la sucesión $(a, b) \in c_0(\mathbb{Z}_0)$ determinada por

$$(a\,,\,b)_n = \begin{cases} a_{-n} & \text{si} \quad n < 0 \\ b_n & \text{si} \quad n > 0 \end{cases} \quad \text{para} \quad n \in \mathbb{Z}_0\,. \tag{1.11}$$

Por ejemplo, si $c \in c_0(\mathbb{M}) \cap \mathbb{R}^{\mathbb{M}}$ entonces $c^{\uparrow\downarrow} = (c^{\uparrow}, c^{\downarrow})$ (ver Definición 1.4.1). Finalmente, dadas $a = (a_n)_{n\in\mathbb{N}} \in \ell^\infty(\mathbb{M})$ y $b = (b_n)_{n\in\mathbb{M}} \in c_0(\mathbb{M})$ consideramos

$$a \cdot b \in c_0(\mathbb{M}) \quad \text{dada por} \quad (a \cdot b)_n = a_n\, b_n \quad , \quad \text{para} \quad n \in \mathbb{M}\,, \tag{1.12}$$

donde $\mathbb{M} = \mathbb{N}$ o $\mathbb{M} = \mathbb{Z}_0$.

El siguiente resultado describe varias propiedades elementales de (sub)mayorización entre sucesiones reales. Las pruebas pueden obtenerse adaptando las ideas del caso finito dimensional.

Lema 1.4.5. Sean $x\,,\, y\,,\, z\,,\, w \in c_0(\mathbb{M}) \cap \mathbb{R}^{\mathbb{M}}$ sucesiones reales. Entonces,

1. $x + y \prec x^{\uparrow\downarrow} + y^{\uparrow\downarrow}$;
2. Si $x \prec y$ entonces $|x| \prec_w |y|$;
3. Si $x \prec z$ y $y \prec w$ entonces $x + y \prec z + w$.

Si asumimos que $\mathbb{M} = \mathbb{N}$, entonces

4. Si $x \prec_w y$ y $z \prec_w w$ entonces $(x, z) \prec_w (y, w)$.

Si además asumimos que $x\,,\, y\,,\, z \in c_0(\mathbb{N})$ son sucesiones no negativas entonces,

5. $x \cdot y \prec_w x^{\downarrow} \cdot y^{\downarrow}$.
6. Si $x \prec_w y$ e $y,\, z \in c_0(\mathbb{N})^{\downarrow}$ entonces $x \cdot z \prec_w y \cdot z$. □

1.5. Operadores compactos autoadjuntos y mayorización

Definición 1.5.1. Sea $A \in K(\mathcal{K}, \mathcal{H})$ una transformación compacta. Entonces consideramos

1. la sucesión no creciente $s(A) = (s_n(A))_{n\in\mathbb{N}} \in c_0(\mathbb{N})$ de sus valores singulares.

   Si además asumimos que $\mathcal{H} = \mathcal{K}$ y $A = A^*$, entonces consideramos:

2. La sucesión de autovalores de A (contando multiplicidad y en algún orden) $\mu(A) \in c_0(\mathbb{N})$;

3. La sucesión ordenada $\lambda(A) \stackrel{\text{def}}{=} \mu(A)^{\uparrow\downarrow} \in c_0(\mathbb{Z}_0)$, llamada *escala espectral completa* de A (también llamada "sistema completo de autovalores"por Markus en [36]). △

Observación 1.5.2. Sea $A \in K(\mathcal{H})^{sa}$ un operador compacto y autoadjunto, y sean $\mu(A) \in c_0(\mathbb{N})$ y $\lambda(A) \in c_0(\mathbb{Z}_0)$ como en la Definición 1.5.1. Notemos que considerando $A^+, A^- \in K(\mathcal{H})^+$ las partes positiva y negativa de A, entonces $A = A^+ - A^-$ y

$$\lambda_i(A) = \lambda_i(A^+) \quad \text{y} \quad \lambda_{-i}(A) = -\lambda_i(A^-) \quad \text{para todo} \quad i \in \mathbb{N}\,. \tag{1.13}$$

Sea $|\mu(A)| \in c_0(\mathbb{N})$ (respectivamente $|\lambda(A)| \in c_0(\mathbb{Z}_0)$) la sucesión obtenida reemplazando los valores de $\mu(A)$ (respectivamente $\lambda(A)$) con sus valores absolutos, de forma que $|\mu(A)|$ y $|\lambda(A)|$ son sucesiones no negativas. Entonces,

$$s(A) = |\mu(A)|^{\downarrow} \quad \text{y similarmente} \quad s(A) = |\lambda(A)|^{\downarrow}\,. \tag{1.14}$$

Por otro lado, tenemos que

$$s(A) = \big((\lambda_j(A))_{j\in\mathbb{N}}\,,\, (\lambda_j(-A))_{j\in\mathbb{N}} \big)^{\downarrow} \in c_0(\mathbb{N})\,. \tag{1.15}$$

En lo que sigue, para $1 \le k \le \infty$, notaremos

$$\mathcal{P}_k(\mathcal{H}) = \{P \in L(\mathcal{H}) \,:\, P^2 = P^* = P\,,\ \mathrm{rk}(P) = k\}\,,$$

al subconjunto de $L(\mathcal{H})$ de todas las proyecciones (ortogonales) de rango k. △

Observación 1.5.3 (Courant-Fischer)**.** A partir de la observación anterior y teniendo en cuenta la definición de escala espectral completa 1.5.1, daremos otra forma de caracterizar a dicha sucesión.

Dado un operador compacto y autoadjunto $A \in K(\mathcal{H})^{sa}$, el principio del Minimax de Courant-Fischer, da una caracterización para la escala espectral completa de A (ver [19]). Dado $0 \le k$ denotamos por $\mathcal{G}(k, \mathcal{H})$ a la variedad de Grassman de subespacios k dimensionales de $\mathcal{H}$. Entonces:

$$\lambda_i(A) = \min_{S\in\mathcal{G}(i-1,\mathcal{H})} \max_{\substack{x\in S^\perp \\ ||x||=1}} \langle A\,x, x\rangle \quad \text{y} \quad \lambda_{-i}(A) = \max_{S\in\mathcal{G}(i-1,\mathcal{H})} \min_{\substack{x\in S^\perp \\ ||x||=1}} \langle A\,x, x\rangle\,, \quad \text{para} \quad i \in \mathbb{N}\,.$$

En particular, notemos que si $B \in K(\mathcal{H})^{sa}$ es un operador de rango finito, con n autovalores positivos y m autovalores negativos $(m, n \ge 0)$, entonces $\lambda_i(B) = 0$ para $i > n$ y $\lambda_{-i}(B) = 0$ para $i > m$ (como indicaba la Observación 1.4.2).

△

Los siguientes resultados son clásicos y corresponden a relaciones espectrales entre operadores compactos (ver [3, 19, 36, 49]).

Teorema 1.5.4. Sean A, $B \in K(\mathcal{H})^{sa}$. Entonces, tenemos las siguientes desigualdades:

1. Desigualdades de entrelace: si $P \in \mathcal{P}_k(\mathcal{H})$ $(1 \leq k \leq \infty)$ entonces
$$\lambda_j(A) \geq \lambda_j(PAP) \quad \text{y} \quad \lambda_{-j}(A) \leq \lambda_{-j}(PAP) \quad \text{para} \quad j \in \mathbb{I}_k\,. \tag{1.16}$$
2. Para todo $k \in \mathbb{N}$,
$$\sum_{i=1}^{k} \lambda_i(A) = \max_{P \in \mathcal{P}_k} \operatorname{tr}(P\,A\,P) \quad \text{y} \quad \sum_{i=1}^{k} \lambda_{-i}(A) = \min_{P \in \mathcal{P}_k} \operatorname{tr}(P\,A\,P)\,.$$
3. Desigualdad de Weyl para la escala espectral: $\lambda(A+B) \prec\!\!\prec \lambda(A) + \lambda(B)$.
4. Desigualdad de Lidskii [36, Teorema 5.1]: $\lambda(A) - \lambda(B) \prec\!\!\prec \lambda(A-B)$. □

También recordamos las desigualdades aditiva y multiplicativa de Weyl para valores singulares:

Teorema 1.5.5. *Dados A, $B \in K(\mathcal{H})$, siguiendo la notación de la ecuación (1.12), tenemos que*
$$s(A+B) \prec\!\!\prec_w s(A) + s(B) \quad y \quad s(AB) \prec\!\!\prec_w s(A) \cdot s(B)\,. \tag{1.17}$$
□

Además tenemos las siguientes desigualdades: Dado $A \in K(\mathcal{H})$, entonces
$$s_j(D\,A\,E) \leq \|D\|\,\|E\|\,s_j(A) \quad \text{para todo} \quad j \in \mathbb{N} \quad \text{y} \quad D, E \in L(\mathcal{H})\,. \tag{1.18}$$

Recordamos también la siguiente caracterización:

Proposición 1.5.6. Sea $E \in K(\mathcal{H})^{sa}$ tomemos $\hat{E} = \begin{bmatrix} 0 & E \\ E^* & 0 \end{bmatrix} \in K(\mathcal{H} \oplus \mathcal{H})$. Entonces
$$\lambda(\hat{E}) = \big(\, s(E)\,, -s(E^*)\,\big)^{\uparrow\downarrow} = \big(\, s(E)\,, -s(E)\,\big)^{\uparrow\downarrow}\,.$$
□

Ideales simétricamente normados

Las relaciones de submayorización tienen un papel central en el estudio de *ideales de operadores simétricamente normados*, desarrollados con detalle en [19].

Observación 1.5.7. Un ideal de operadores simétricamente normado $\mathcal{C}$ de $L(\mathcal{H})$ es un ideal bilátero junto con una norma simétrica $N(\cdot)$, es decir, una norma con las siguientes propiedades adicionales:

1. Dados $A \in \mathcal{C}$ y $D, E \in L(\mathcal{H})$, entonces $N(D\,A\,E) \leq \|D\|\,N(A)\,\|E\|$;
2. Para todo operador A tal que $\operatorname{rk}(A) = 1$, $N(A) = \|A\| = s_1(A)$.

Más aún, cualquier norma simétrica resulta unitariamente invariante, como consecuencia del ítem 1. Es decir, si $U, V \in \mathcal{U}(\mathcal{H})$ entonces $N(UAV) = N(A)$, para todo $A \in \mathcal{C}$. Además, estas normas inducen las llamadas funciones gauge simétricas (o simétricamente normantes en [19]) g_N para sucesiones, de forma que $N(A) = g_N(s(A)\,)$.

En este contexto, tenemos que: dados $B \in \mathcal{C}$ y $A \in K(\mathcal{H})$ entonces
$$s(A) \prec\!\!\prec_w s(B) \implies A \in \mathcal{C} \quad \text{y} \quad N(A) = g_N(s(A)\,) \leq g_N(s(B)\,) = N(B)\,. \tag{1.19}$$
Por simplicidad, en lo que sigue llamaremos a este tipo de normas $N(\cdot)$, *norma unitariamente invariante*. Como ejemplo de este tipo de normas, mencionamos las normas p-Schatten , para $1 \leq p < \infty$ asociadas a los ideales simétricamente normados de Schatten en $L(\mathcal{H})$. △

Capítulo 2

Valores de Ritz de matrices autoadjuntas

En este capítulo obtendremos cotas superiores a priori (autónomas), a posteriori (no autónomas) y de tipo mixto para la variación absoluta de los valores de Ritz de matrices autoadjuntas en términos de submayorización. Algunos de nuestros resultados demuestran conjeturas recientes de [30, 33, 58] que permiten extender varios resultados conocidos para subespacios unidimensionales a subespacios arbitrarios.

El capítulo se divide en tres partes; primero demostraremos la [33, Conjetura 2.1] que establece cotas de tipo mixto vinculadas al error en la variación (absoluta) de los valores del Ritz. En la segunda parte, haremos una revisión al teorema tan Θ de Nakatsukasa [43] bajo condiciones relajadas y obtendremos una versión mejorada de este resultado. Como consecuencia daremos una estimación del error cuadrático a posteriori para la variación de los valores del Ritz que mejoran varias cotas conocidas.

Finalmente, estableceremos conexiones entre las cotas de tipo mixto de la primera parte y algunas cotas a priori para la variación de los valores de Ritz conjeturadas en [30, 31].

Comenzamos recordando la siguiente:

Notación 2.0.1. En esta sección utilizaremos la siguiente notación:

1. $\mathcal{X}$, $\mathcal{Y} \subset \mathbb{C}^d$ denotan subespacios de dimensión k. Notamos $\mathcal{I}_{\mathcal{X}}(k,d)$ a las isometrías $X \in \mathcal{M}_{d,\,k}(\mathbb{C})$ con rango $R(X) = \mathcal{X}$. De esta forma fijamos $X \in \mathcal{I}_{\mathcal{X}}(k,d)$ e $Y \in \mathcal{I}_{\mathcal{Y}}(k,d)$.

2. $\Theta(\mathcal{X}\,,\,\mathcal{Y}) \in (\mathbb{R}^k_{\geq 0})^{\downarrow}$ denota el vector de ángulos principales entre los subespacios $\mathcal{X}$ e $\mathcal{Y}$; en este caso,

$$\cos(\Theta^{\uparrow}(\mathcal{X}\,,\,\mathcal{Y})) = s(X^*Y) = (s_1(X^*Y), \ldots, s_k(X^*Y)) \in (\mathbb{R}^k_{\geq 0})^{\downarrow}.$$

3. Para una matriz (fija) autoadjunta $A \in \mathcal{H}(d)$ consideramos $R_X = A\,X - X\,X^*A\,X \in \mathcal{M}_{d,k}(\mathbb{C})$ y de forma similar R_Y para Y. Notemos que

$$R_X = A\,X - X\,X^*A\,X = A\,X - P_{\mathcal{X}}A\,X = P_{\mathcal{X}^\perp}A\,X \in \mathcal{M}_{d,k}(\mathbb{C})\,,$$

   donde $P_{\mathcal{X}} \in \mathcal{M}_d(\mathbb{C})$ denota el proyector ortogonal sobre $\mathcal{X}$ y $\mathcal{X}^\perp$ denota el complemento ortogonal de $\mathcal{X}$. Consideramos de forma análoga estas notaciones para $\mathcal{Y}$. En particular $\mathcal{X}$ es un subespacio A-invariante si y sólo si $R_X = 0$.

4. Sea $X_\perp \in \mathcal{I}_{\mathcal{X}^\perp}(d-k,d)$. Entonces, la matriz $(X, X_\perp) \in \mathcal{U}(d)$ y

$$\tilde{A} = (X, X_\perp)\ \ A\ \ (X, X_\perp)^* = \begin{bmatrix} X^*A\,X & R_X^*\,X_\perp \\ X_\perp^*\,R_X & X_\perp^*A\,X_\perp \end{bmatrix}.$$

   Cabe destacar que, como $R_X = (I - P_{\mathcal{X}})\,R_X$, entonces $s(R_X) = s(X_\perp^*\,R_X)$, por lo que podemos pensar a R_X (salvo por una isometría) como el bloque $(2,1)$ de $\tilde{A}$, en la representación matricial de (el conjugado unitario de A) $\tilde{A}$ como antes. $\triangle$

2.1. Cotas de mayorización mixtas para el error de Rayleigh-Ritz

Utilizaremos la Notación 2.0.1; más aún en este apartado asumiremos que $\mathcal{X}$ e $\mathcal{Y}$ son tales que $\Theta_1(\mathcal{X}, \mathcal{Y}) < \frac{\pi}{2}$, es decir que $X^*Y \in \mathcal{G}l(k)$ es inversible.

El primer resultado propone una cota de submayorización para el error en la distancia entre tiras de autovalores de matrices autoadjuntas:

Teorema 2.1.1. *Sean C, $D \in \mathcal{H}(d)$. Entonces,*

1. *si $T \in \mathcal{G}l(d)^+$, entonces*
$$s(C-D) \prec_w s(T^{-1})\, s(CT-TD)\,. \tag{2.1}$$
2. *Si $T \in \mathcal{G}l(d)$, entonces*
$$|\lambda(C)-\lambda(D)| \prec_w s(T^{-1})\; s(CT-TD). \tag{2.2}$$

.

Demostración. Probaremos primero el ítem 1. Como T es positiva e inversible, usando el Teorema 1.1.5 (ítem 2) tenemos que

$$\begin{aligned} s(C-D) &= s(CT^{\frac12}T^{-\frac12} - T^{-\frac12}T^{\frac12}D) = s(T^{-\frac12}(T^{\frac12}CT^{\frac12} - T^{\frac12}DT^{\frac12})T^{-\frac12}) \\ &\prec_w s(T^{-\frac12})^2\, s(T^{\frac12}CT^{\frac12} - T^{\frac12}DT^{\frac12}) = s(T^{-1})\, s(T^{\frac12}(C-D)\,T^{\frac12})\,. \end{aligned}$$

Por el Teorema 1.1.5 (ítems 3. y 4.) y el hecho de que $\mathrm{re}(DT) = \mathrm{re}(TD)$ tenemos que

$$\begin{aligned} s(T^{\frac12}(C-D)T^{\frac12}) &\prec_w s(\mathrm{re}[(C-D)T]) = s(\mathrm{re}(CT-DT)\,) = s(\mathrm{re}(CT) - \mathrm{re}(DT)\,) \\ &= s(\mathrm{re}(CT) - \mathrm{re}(TD)\,) = s(\mathrm{re}(CT-TD)\,) \\ &\prec_w s(CT-TD), \end{aligned} \tag{2.3}$$

Luego utilizando las desigualdades previas y el Lema 1.1.6 tenemos que

$$s(C-D) \prec_w s(T^{-1})\, s(CT-TD)\,. \tag{2.4}$$

Ahora bien para probar el ítem 2., consideramos la representación, $T = U\Sigma V^*$, donde U, $V \in \mathcal{U}(d)$ son matrices unitarias y $\Sigma \in \mathcal{M}_d(\mathbb{C})$ es la matriz diagonal cuya diagonal es $s(T) \in \mathbb{R}^d_{\geq 0}$ (esta representación corresponde a la descomposición en valores singulares de T); notemos que Σ es definida positiva e inversible.
Más aún, $T^{-1} = V^*\Sigma^{-1}U$, de forma que $s(T^{-1}) = s(\Sigma^{-1})$. Luego, usando el ítem 2. del Teorema 1.1.4 y la ecuación (2.1) (ítem 1.) tenemos que

$$\begin{aligned} |\lambda(C)-\lambda(D)| = |\lambda(U^*CU) - \lambda(V^*DV)| &\prec_w s(U^*CU - V^*DV) \\ &\prec_w s(\Sigma^{-1})\, s(U^*CU\Sigma - \Sigma V^*DV) \\ &= s(T^{-1})\, s(U^*(C\,U\,\Sigma\,V^* - U\,\Sigma\,V^*D)\,V) \\ &= s(T^{-1})\, s(U^*(CT-TD)V) \\ &= s(T^{-1})\, s(CT-TD)\,. \end{aligned}$$

□

El siguiente resultado describe los autovalores de la compresión de $P_{\mathcal{Y}^\perp}$ con respecto a $\mathcal{X}$, en términos de $\Theta(\mathcal{X}, \mathcal{Y})$.

Teorema 2.1.2 ([30]). *Sean $\mathcal{X}$, $\mathcal{Y} \subset \mathbb{C}^d$ tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$. Entonces*

$$\lambda(P_{\mathcal{X}} P_{\mathcal{Y}^\perp} P_{\mathcal{X}}) = s(P_{\mathcal{X}} P_{\mathcal{Y}^\perp} P_{\mathcal{X}}) = s^2(P_{\mathcal{Y}} P_{\mathcal{X}^\perp}) = s^2(P_{\mathcal{X}^\perp} P_{\mathcal{Y}}) = (\operatorname{sen}^2(\Theta(\mathcal{X},\mathcal{Y})), 0_{d-k}).$$

Demostración. Primero notemos que la descomposición de Halmos de $\mathbb{C}^d$ en términos de $\mathcal{X}$ e $\mathcal{Y}$ permite describir a los proyectores $P_{\mathcal{X}}$ y $P_{\mathcal{Y}^\perp}$ como sigue (ver Observación 1.3.5):

$$P_{\mathcal{X}} = \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \mathcal{X}\cap\mathcal{Y} \\ \mathcal{S}_1 \\ \mathcal{S}_2 \\ \mathcal{X}^\perp\cap\mathcal{Y}^\perp \end{matrix} \quad \text{y} \quad P_{\mathcal{Y}^\perp} = (I - P_{\mathcal{Y}}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & S^2 & -C\,S & 0 \\ 0 & -S\,C & C^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \mathcal{X}\cap\mathcal{Y} \\ \mathcal{S}_1 \\ \mathcal{S}_2 \\ \mathcal{X}^\perp\cap\mathcal{Y}^\perp \end{matrix},$$

donde hemos considerado una rotación directa $U \in \mathcal{U}(d)$ de $\mathcal{X}$ en $\mathcal{Y}$ (ver ecuación (1.8)), de forma que $P_{\mathcal{Y}} = U^*\, P_{\mathcal{X}}\, U$.

Luego

$$P_{\mathcal{X}}\, P_{\mathcal{Y}^\perp}\, P_{\mathcal{X}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & S^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \mathcal{X}\cap\mathcal{Y} \\ \mathcal{S}_1 \\ \mathcal{S}_2 \\ \mathcal{X}^\perp\cap\mathcal{Y}^\perp \end{matrix},$$

por lo que $\lambda(P_{\mathcal{X}}\, P_{\mathcal{Y}^\perp}\, P_{\mathcal{X}}) = (\lambda(S^2), 0_{d-k}) = (\operatorname{sen}^2(\Theta(\mathcal{X},\mathcal{Y})), 0_{d-k})$.

Por otro lado, utilizando la representación por bloques para $P_{\mathcal{Y}}$ y $P_{\mathcal{X}^\perp}$ se obtiene que $\lambda(P_{\mathcal{X}^\perp}\, P_{\mathcal{Y}}\, P_{\mathcal{X}^\perp}) = s^2(P_{\mathcal{X}^\perp}\, P_{\mathcal{Y}}) = (\operatorname{sen}^2(\Theta(\mathcal{X},\mathcal{Y})), 0_{d-k})$.

□

Antes de probar uno de los resultados principales de esta sección, recordamos el siguiente

Lema 2.1.3 ([4]). Bajo las condiciones de la Notación 2.0.1,

$$s(P_{\mathcal{X}} R_Y) \prec_w s(P_{\mathcal{X}+\mathcal{Y}}\, R_Y)\operatorname{sen}(\Theta(\mathcal{X},\mathcal{Y}))\,. \tag{2.5}$$

Demostración. Notemos que

$$\begin{aligned} s(P_{\mathcal{X}}\, R_Y) &= s(Y^* A\, P_{\mathcal{Y}^\perp}\, P_{\mathcal{X}}) = s(Y^* A\, P_{\mathcal{Y}^\perp}\, P_{\mathcal{Y}^\perp}\, P_{\mathcal{X}}) \\ &= s(Y^* A\, P_{\mathcal{Y}^\perp}\, P_{\mathcal{X}+\mathcal{Y}}\, P_{\mathcal{Y}^\perp}\, P_{\mathcal{X}})\,, \end{aligned}$$

pues $\mathcal{X}, \mathcal{Y} \subset \mathcal{X} + \mathcal{Y}$ (como subespacios) lo que implica que

$$P_{\mathcal{Y}^\perp}\, P_{\mathcal{X}} = (P_{\mathcal{X}} - P_{\mathcal{Y}}\, P_{\mathcal{X}}) = P_{\mathcal{X}+\mathcal{Y}}\,(P_{\mathcal{X}} - P_{\mathcal{Y}}\, P_{\mathcal{X}}) = P_{\mathcal{X}+\mathcal{Y}}\, P_{\mathcal{Y}^\perp}\, P_{\mathcal{X}}\,.$$

Luego por el Teorema 1.1.5 ítem 2. y el Teorema 2.1.2

$$\begin{aligned} s(P_{\mathcal{X}}\, R_Y) &= s(Y^* A\, P_{\mathcal{Y}^\perp}\, P_{\mathcal{X}+\mathcal{Y}}\, P_{\mathcal{Y}^\perp}\, P_{\mathcal{X}}) \prec_w s(Y^* A\, P_{\mathcal{Y}^\perp}\, P_{\mathcal{X}+\mathcal{Y}})\, s(P_{\mathcal{X}}\, P_{\mathcal{Y}^\perp}) \\ &\prec_w s(P_{\mathcal{X}+\mathcal{Y}}\, R_Y)\operatorname{sen}(\Theta(\mathcal{X},\mathcal{Y}))\,, \end{aligned}$$

donde hemos omitido los 0_{d-k} en el vector del $\operatorname{sen}(\Theta(\mathcal{X},\mathcal{Y}))$, siguiendo la Observación 1.1.2.

□

El siguiente resultado corresponde a la [33, Conjetura 2.1] (ver también el Corolario 2.1.6 debajo).

Teorema 2.1.4. *Bajo la Notación 2.0.1, si* $\Theta_1(\mathcal{X}\,,\,\mathcal{Y}) < \frac{\pi}{2}$ *entonces*

$$|\lambda(X^*A\,X) - \lambda(Y^*A\,Y)| \prec_w \frac{s(P_{\mathcal{Y}}\,R_X) + s(P_{\mathcal{X}}\,R_Y)}{\cos(\Theta(\mathcal{X}\,,\,\mathcal{Y}))} \qquad y \tag{2.6}$$

$$|\lambda(X^*A\,X) - \lambda(Y^*A\,Y)| \prec_w [s(P_{\mathcal{X}+\mathcal{Y}}\,R_X) + s(P_{\mathcal{X}+\mathcal{Y}}\,R_Y)]\,\tan(\Theta(\mathcal{X}\,,\,\mathcal{Y}))\,. \tag{2.7}$$

Observación 2.1.5. Notemos que la condición $\Theta_1(\mathcal{X}\,,\,\mathcal{Y}) < \frac{\pi}{2}$, asegura que la matriz $T = X^*Y$ es inversible. Además, ésta condición simplifica algunos aspectos de la descomposición de Halmos; por ejemplo, asegura que $\mathcal{X} \cap \mathcal{Y}^\perp = \{0\}$ y también $\mathcal{X}^\perp \cap \mathcal{Y} = \{0\}$ dado que las dimensiones de ambos subespacios coinciden. En consecuencia, $\mathbb{C}^d = (\mathcal{X} \cap \mathcal{Y}) \oplus (\mathcal{X}^\perp \cap \mathcal{Y}^\perp) \oplus \mathcal{G}$. △

Demostración. Consideremos $T = X^*Y$, por la observación anterior tenemos que $T \in \mathcal{M}_k(\mathbb{C})$ es inversible. Usando el Teorema 2.1.1 tenemos que

$$|\lambda(X^*A\,X) - \lambda(Y^*A\,Y)| \prec_w s(T^{-1})\; s\left((X^*A\,X)\,T - T\,(Y^*A\,Y)\right)\,. \tag{2.8}$$

Luego por construcción

$$s(T^{-1}) = \frac{1}{\cos(\Theta(\mathcal{X},\mathcal{Y}))} \in (\mathbb{R}^k_{>0})^{\downarrow}\,. \tag{2.9}$$

Siguiendo el argumento en [33, Teorema 4.1] notamos que

$$\begin{aligned}(X^*A\,X)T - T(Y^*A\,Y) &= (X^*A\,X)X^*Y - X^*Y(Y^*A\,Y) = X^*A\,P_{\mathcal{X}}Y - X^*P_{\mathcal{Y}}A\,Y\\ &= X^*A\,(I - P_{\mathcal{X}^\perp})Y - X^*(I - P_{\mathcal{Y}^\perp})A\,Y\\ &= X^*A\,Y - X^*A\,P_{\mathcal{X}^\perp}Y - X^*A\,Y + X^*P_{\mathcal{Y}^\perp}A\,Y\\ &= -X^*A\,P_{\mathcal{X}^\perp}Y + X^*P_{\mathcal{Y}^\perp}A\,Y\,.\end{aligned}$$

Usando que $s(C) = s(C^*)$ para $C \in \mathcal{M}_k(\mathbb{C})$, resulta que

$$s(X^*A\,P_{\mathcal{X}^\perp}Y) = s(Y^*P_{\mathcal{X}^\perp}A\,X) = s(P_{\mathcal{Y}}P_{\mathcal{X}^\perp}A\,X) = s(P_{\mathcal{Y}}R_X) \in (\mathbb{R}^k_{\geq 0})^{\downarrow}\,.$$

Análogamente $s(X^*P_{\mathcal{Y}^\perp}A\,Y) = s(P_{\mathcal{X}}R_Y)$. Los hechos anteriores y la propiedad subaditiva de los valores singulares (ítem 1. del Teorema 1.1.5) implican que

$$s((X^*A\,X)T - T(Y^*A\,Y)) = s(-X^*A\,P_{\mathcal{X}^\perp}Y + X^*P_{\mathcal{Y}^\perp}A\,Y) \prec_w s(P_{\mathcal{X}}R_Y) + s(P_{\mathcal{Y}}R_X)\,. \tag{2.10}$$

Ahora bien combinando las ecuaciones (2.9) y (2.10) con (2.8), y utilizando el ítem 4. de Lema 1.1.6, obtenemos (2.6).

Para probar (2.7) recordemos que el Lema 2.1.3 establece que

$$s(P_{\mathcal{X}}R_Y) \prec_w s(P_{\mathcal{X}+\mathcal{Y}}\,R_Y)\,\mathrm{sen}(\Theta(\mathcal{X},\mathcal{Y}))\,. \tag{2.11}$$

Como las entradas de estos vectores están ordenadas en forma no creciente, por el ítem 2. del Lema 1.1.6 tenemos que

$$s(P_{\mathcal{X}}R_Y) + s(P_{\mathcal{Y}}R_X) \prec_w \left(s(P_{\mathcal{X}+\mathcal{Y}}\,R_Y) + s(P_{\mathcal{X}+\mathcal{Y}}\,R_X)\right)\,\mathrm{sen}(\Theta(\mathcal{X},\mathcal{Y}))\,. \tag{2.12}$$

Por lo tanto, combinando las ecuaciones (2.6) y (2.12) con el Lema 1.1.6 probamos (2.7).

□

El hecho de que la ecuación (2.6) implica (2.7) ya había sido probado en [33]; de forma que hemos incluido la prueba por completitud.

Corolario 2.1.6. *Considerando la Notación 2.0.1, asumimos que* $\Theta_1(\mathcal{X}, \mathcal{Y}) < \frac{\pi}{2}$. *Si* $\mathcal{X}$ *es* A*-invariante entonces*

$$|\lambda(X^* A X) - \lambda(Y^* A Y)| \prec_w \frac{s(P_{\mathcal{X}} R_Y)}{\cos(\Theta(\mathcal{X}, \mathcal{Y}))} \quad y \tag{2.13}$$

$$|\lambda(X^* A X) - \lambda(Y^* A Y)| \prec_w s(P_{\mathcal{X}+\mathcal{Y}} R_Y) \tan(\Theta(\mathcal{X}, \mathcal{Y})) . \tag{2.14}$$

Demostración. Notemos que si $\mathcal{X}$ is A-invariante, entonces $R_X = 0$. En efecto $P_{\mathcal{X}^\perp} A X = 0$ pues $R(A X) \subset \mathcal{X}$. Luego el resultado se obtiene a partir del Teorema 2.1.4.

□

Es natural preguntarse si las cotas obtenidas en los resultados anteriores pueden mejorarse. El siguiente ejemplo mostrará que las cotas de submayorización obtenidas en el Teorema 2.1.4 y el Corolario 2.1.6 son ajustadas.

Ejemplo 2.1.7. Sea $\lambda = (a, b, c, d) \in \mathbb{R}^4$, donde $a < b < c < d$ y consideremos $A \in \mathcal{H}(4)$ dada por $A = D_\lambda$, es decir, A es la matriz diagonal con diagonal λ.

Consideramos además el subespacio A-invariante $\mathcal{X} = \text{span}\{e_1, e_2\}$ generado por los dos primeros elementos de la base canónica de $\mathbb{C}^4$. Para $\theta \in (0, \pi/2)$, sea $f_\theta = \cos\theta\, e_2 + \text{sen}\,\theta\, e_3$ y tomemos $\mathcal{Y}_\theta = \text{span}\{e_1, f_\theta\}$. Entonces, los ángulos principales están dados por $\Theta(\mathcal{X}, \mathcal{Y}_\theta) = (\theta, 0)$.

Por otro lado consideramos las representaciones matriciales para $X, X_\perp$ e Y con respecto a la base canónica de $\mathbb{C}^4$

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad , \quad X_\perp = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{y} \quad Y_\theta = \begin{pmatrix} 1 & 0 \\ 0 & \cos\theta \\ 0 & \text{sen}\,\theta \\ 0 & 0 \end{pmatrix} .$$

Luego, un cálculo directo permite establecer que $\lambda(X^* A X) = (b, a)$ y que $\lambda(Y_\theta^* A Y_\theta) = (b\cos^2\theta + c\,\text{sen}^2\theta, a)$. Más aún,

$$R_{Y_\theta} = \begin{pmatrix} 0 & 0 \\ 0 & (b-c)\cos\theta\,\text{sen}^2\theta \\ 0 & (c-b)\cos^2\theta\,\text{sen}\,\theta \\ 0 & 0 \end{pmatrix} \quad , \quad P_{\mathcal{X}} R_{Y_\theta} = \begin{pmatrix} 0 & 0 \\ 0 & (b-c)\cos\theta\,\text{sen}^2\theta \\ 0 & 0 \\ 0 & 0 \end{pmatrix} .$$

Por lo tanto, $s(P_{\mathcal{X}} R_{Y_\theta}) = ((c-b)\cos\theta\,\text{sen}^2\theta, 0)$. Ahora bien,

$$|\lambda(X^* A X) - \lambda((Y_\theta)^* A Y_\theta)| = ((c-b)\,\text{sen}^2\theta, 0) \quad \text{y} \tag{2.15}$$

$$\frac{s(P_{\mathcal{X}} R_{Y_\theta})}{\cos(\Theta(\mathcal{X}, \mathcal{Y}_\theta))} = ((c-b)\,\text{sen}^2\theta, 0) . \tag{2.16}$$

Es decir, la ecuación (2.13) del Corolario 2.1.6 resulta ser una igualdad en este caso. Esto muestra además que la ecuación (2.6) es ajustada, dado que la ecuación (2.13) resulta ser un caso particular (cuando $\mathcal{X}$ es A-invariante).

Por otro lado notemos que $\mathcal{X} + \mathcal{Y}_\theta = \text{span}\{e_1, e_2, e_3\}$. Luego, como $P_{\mathcal{X}+\mathcal{Y}_\theta} R_{Y_\theta} = R_{Y_\theta}$ y $s(R_{Y_\theta}) = ((c-b)\cos\theta\,\text{sen}\,\theta, 0)$,

$$s(P_{\mathcal{X}+\mathcal{Y}_\theta} R_{Y_\theta}) \tan(\Theta(\mathcal{X}, \mathcal{Y}_\theta)) = ((c-b)\,\text{sen}^2\theta, 0) . \tag{2.17}$$

Entonces las ecuaciones (2.15) y (2.17) muestran que la ecuación (2.14) del Corolario 2.1.6 es una igualdad en este caso. Esto muestra que la ecuación (2.7) es ajustada, dado que la ecuación (2.14) es un caso particular (cuando $\mathcal{X}$ es A-invariante). △

Observación 2.1.8 (Relaciones con trabajos anteriores)**.** En el caso vectorial, cuando $\mathcal{X}$ e $\mathcal{Y}$ son espacios uno dimensionales, el Teorema 2.1.4 implica las cotas superiores de [58, Teorema 3.7], que es uno de los resultados principales de ese trabajo (ver también el Corolario 2.2.11 y la Observación 2.2.12).

Por otro lado en [33] Knyazev y Zhu obtienen varias cotas para la variación absoluta de los valores de Ritz. Usando la Notación 2.0.1, los autores prueban (ver [33, Teorema 4.2 y Corolario 4.4]) que

$$|\lambda(X^*A\,X - Y^*A\,Y)|^2 \prec_w \frac{\{s(P_{\mathcal{Y}}\ R_X) + s(P_{\mathcal{X}}\ R_Y)\}^2}{\cos^2(\Theta(\mathcal{X}\,,\,\mathcal{Y}))} \qquad \text{y} \tag{2.18}$$

$$|\lambda(X^*A\,X) - \lambda(Y^*A\,Y)|^2 \prec_w \{s(P_{\mathcal{X}+\mathcal{Y}}\ R_X) + s(P_{\mathcal{X}+\mathcal{Y}}\ R_Y)\}^2\ \tan^2(\Theta(\mathcal{X}\,,\,\mathcal{Y}))\,. \tag{2.19}$$

Usando que $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ dada por $f(x) = x^2$ es una función convexa creciente, la Observación 1.1.3 muestra que las ecuaciones (2.18) y (2.19) son consecuencia de las ecuaciones (2.6) y (2.7) del Teorema 2.1.4.
Similarmente, usando que $\cos\Theta_1(\mathcal{X},\mathcal{Y}) = \cos\Theta_{\text{máx}}(\mathcal{X},\mathcal{Y}) \leq \cos\Theta_i(\mathcal{X},\mathcal{Y})$, para $i \in \mathbb{I}_k$, tenemos que el Teorema 2.1.4 implica los [33, Teoremas 4.1 y 4.3].

En [33] los autores muestran que sus resultados son aplicables en varias situaciones como: cotas de mayorización de primer orden y cuadráticas; cotas para autovalores de perturbaciones aditivas de una matriz. Las observaciones anteriores muestran que nuestros resultados también son aplicables en estos contextos. Más aún, el Teorema 2.1.4 permite formalizar los argumentos relacionados con cotas para autovalores de perturbaciones aditivas de una matriz y en particular con cotas para autovalores después de remover los bloques no diagonales de una matriz, desarrollado en [33, Sección 5] (ver la discusión detallada en ese trabajo). △

Además, las cotas del Teorema 2.1.4 pueden ser utilizadas para realizar un análisis detallado y obtener mejores órdenes de convergencia para algoritmos iterativos relacionados con el método de Rayleigh-Ritz (ver [46, 50, 57]).

2.2. El teorema tan Θ: cotas a posteriori de error cuadrático mejoradas

En esta sección revisaremos la extensión de Nakatsukasa del teorema tan Θ de Davis y Kahan [17]. La motivación es estudiar una versión mejorada de dicho teorema, conjeturada en [33] (ver Corolario 2.2.9).

Primero recordamos las condiciones de separación para el resultado de Nakatsukasa. Como antes, seguiremos la Notación 2.0.1.

Definición 2.2.1. Sean $A \in \mathcal{H}(d)$ y $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$ subespacios con $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$, tales que $\mathcal{X}$ es A-invariante. Sean $[X, X_\perp], [Y, Y_\perp] \in \mathcal{U}(d)$ matrices unitarias tales que las columnas de (las matrices de $d \times k$) X e Y forman bases ortonormales para $\mathcal{X}$ e $\mathcal{Y}$ respectivamente. Dado $\delta > 0$ decimos que $(A\,,\,\mathcal{X}\,,\,\mathcal{Y}\,,\,\delta)$ satisface la propiedad de separación de Davis-Kahan-Nakatsukasa (DKN) si existe $a \leq b$ tal que

1. $\lambda_i(X_\perp^* A X_\perp) = \lambda_i(P_{\mathcal{X}^\perp}\, A\, P_{\mathcal{X}^\perp}) \in [a, b]$, para $i \in \mathbb{I}_{d-k}$;
2. $\lambda_i(Y^*AY) = \lambda_i(P_{\mathcal{Y}}\, A\, P_{\mathcal{Y}}) \in (\infty, a - \delta] \cup [b + \delta, \infty)$, para $i \in \mathbb{I}_k$. △

Observación 2.2.2. En los siguientes resultados utilizaremos la notación $||\cdot||$ para denotar una NUI. En este contexto, consideraremos también la notación $||x|| := g_{||\cdot||}(x)$para $x \in \mathbb{C}^d$, donde $g_{||\cdot||}$ denota la función gauge simétrica asociada a $||\cdot||$ (ver Observación 1.2.3).

El siguiente resultado es el teorema de Nakatsukasa tan Θ (bajo condiciones relajadas).

Teorema 2.2.3 ([43]). *Sean $A \in \mathcal{H}(d)$, $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$ y sea $\delta > 0$ tal que $(A\,,\,\mathcal{X}\,,\,\mathcal{Y}\,,\,\delta)$ satisface la propiedad de separación DKN. Entonces, $\Theta_1(\mathcal{X},\mathcal{Y}) < \pi/2$ y*

$$\delta \,\|\tan(\Theta(\mathcal{X}\,,\,\mathcal{Y}))\| \leq \|R_Y\|\,,$$

para toda norma unitariamente invariante $\|\cdot\|$. Equivalentemente, $\delta\,\tan(\Theta(\mathcal{X}\,,\,\mathcal{Y})) \prec_w s(R_Y)$.

Observación 2.2.4. El Teorema 2.2.3 requiere el conocimiento de la matriz A completa para acotar superiormente la norma del vector $\tan(\Theta(\mathcal{X},\mathcal{Y}))$. Por lo tanto, sería interesante acotar la norma del vector $\tan(\Theta(\mathcal{X},\mathcal{Y}))$ superiormente (sólo) en términos del operador autoadjunto $A_{\mathcal{X}+\mathcal{Y}} = P_{\mathcal{X}+\mathcal{Y}}\,A|_{\mathcal{X}+\mathcal{Y}} \in \mathcal{L}(\mathcal{X}+\mathcal{Y})$ (definido de forma obvia). En el siguiente resultado mostramos que el teorema tan Θ mencionado anteriormente permite obtener este resultado. Además, también veremos que es posible describir la propiedad de separación DKN para $(A_{\mathcal{X}+\mathcal{Y}}, \mathcal{X}, \mathcal{Y})$, que resulta ser más general que la propiedad de separación DKN para $(A, \mathcal{X}, \mathcal{Y})$, para la cual se cumple el teorema tan Θ; argumentando en términos de desigualdades de entrelace, demostraremos que estas condiciones de separación para $A_{\mathcal{X}+\mathcal{Y}}$ proporcionan mejores constantes de separación que las condiciones de separación DKN para la matriz A. △

Formalizaremos el contenido de la observación anterior en el siguiente resultado. Primero recordemos algunos detalles de la descomposición de Halmos.

Observación 2.2.5. Sean $\mathcal{X}$, $\mathcal{Y} \subset \mathbb{C}^d$ subespacios con $\dim\mathcal{X} = \dim\mathcal{Y} = k$. La descomposición de Halmos en dos subespacios permitía describir a $\mathbb{C}^d$ de la siguiente forma:

$$\mathbb{C}^d = (\mathcal{X}\cap\mathcal{Y}) \oplus (\mathcal{X}\cap\mathcal{Y}^\perp) \oplus (\mathcal{X}^\perp\cap\mathcal{Y}) \oplus (\mathcal{X}^\perp\cap\mathcal{Y}^\perp) \oplus \mathcal{G}\,, \tag{2.20}$$

donde

$$\mathcal{G} \overset{\text{def}}{=} \left[\,(\mathcal{X}\cap\mathcal{Y}) \oplus (\mathcal{X}\cap\mathcal{Y}^\perp) \oplus (\mathcal{X}^\perp\cap\mathcal{Y}) \oplus (\mathcal{X}^\perp\cap\mathcal{Y}^\perp)\,\right]^\perp,$$

representaba la parte genérica entre $\mathcal{X}$ e $\mathcal{Y}$. Cada uno de estos cinco subespacios (que podrían ser nulos) reducen a las proyecciones $P_{\mathcal{X}}$ y $P_{\mathcal{Y}}$. Más aún, los subespacios $\mathcal{X}_g = \mathcal{X}\cap\mathcal{G}$ e $\mathcal{Y}_g = \mathcal{Y}\cap\mathcal{G}$ están en *posición genérica* de forma que $\mathcal{G} = \mathcal{X}_g + \mathcal{Y}_g$ y $\mathcal{X}_g \cap \mathcal{Y}_g = \{0\}$. △

Teorema 2.2.6. *Sea $A \in \mathcal{H}(d)$, y sean $\mathcal{X}$, $\mathcal{Y} \subset \mathbb{C}^d$ tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$. Sea $A_{\mathcal{X}+\mathcal{Y}} = S^*AS \in \mathcal{H}(p)$, donde $S \in \mathcal{I}_{\mathcal{X}+\mathcal{Y}}(p,d)$. Entonces,*

1. *Si $\delta > 0$ es tal que $(A\,,\,\mathcal{X}\,,\,\mathcal{Y}\,,\,\delta)$ satisface la propiedad de separación DKN entonces existe $\delta\,' \geq \delta$ tal que $(A_{\mathcal{X}+\mathcal{Y}}\,,\,S^*\mathcal{X}\,,\,S^*\mathcal{Y}\,,\,\delta\,')$ satisface la propiedad de separación DKN.*
2. *Si $\delta\,' > 0$ es tal que $(A_{\mathcal{X}+\mathcal{Y}}\,,\,S^*\mathcal{X}\,,\,S^*\mathcal{Y}\,,\,\delta')$ satisface la propiedad de separación DKN, entonces*

$$\delta\,'\,\|\tan(\Theta(\mathcal{X}\,,\,\mathcal{Y}))\| \leq \|A_{\mathcal{X}+\mathcal{Y}}\,Y_S - Y_S\,(Y_S^*A_{\mathcal{X}+\mathcal{Y}}\,Y_S)\| = \|P_{\mathcal{X}+\mathcal{Y}}\;R_Y\|\,, \tag{2.21}$$

para toda norma unitariamente invariante $\|\cdot\|$, donde $Y_S = S^*Y \in \mathcal{M}_{p,k}(\mathbb{C})$.

Demostración. Probaremos primero el ítem 1. Sean $X\,, Y \in \mathcal{M}_{d,k}(\mathbb{C})$ tales que sus columnas forman bases ortonormales para $\mathcal{X}$ e $\mathcal{Y}$ respectivamente. Por hipótesis, existe $a \leq b$ tal que, para $i \in \mathbb{I}_{d-k}$ y $j \in \mathbb{I}_k$,

$$\lambda_i(X_\perp^*AX_\perp) \in [a,b] \qquad \text{y} \qquad \lambda_j(Y^*AY) \in (\infty, a-\delta] \cup [b+\delta, \infty)\,,$$

donde $X_\perp \in \mathcal{I}_{\mathcal{X}^\perp}(d-k,d)$. Sea $\mathcal{Z} = \mathcal{X}+\mathcal{Y}$ notemos que $S \in \mathcal{M}_{d,p}(\mathbb{C})$ es una isometría de $\mathbb{C}^p$ en $\mathcal{Z}$. Más aún, la matriz $S^*A\,S \in \mathcal{H}(p)$. De forma similar, $X_S = S^*X$, $Y_S = S^*Y \in \mathcal{M}_{p,k}(\mathbb{C})$ son isometrías de $\mathbb{C}^k$ en $S^*\mathcal{X}$, $S^*\mathcal{Y} \subseteq \mathbb{C}^p$, respectivamente. Ahora bien, consideramos los subespacios

$$\mathcal{X}\cap\mathcal{Y} \quad , \quad \mathcal{X}_g = \mathcal{G}\cap\mathcal{X} \qquad \text{y} \qquad \mathcal{X}_{g^\perp} = \mathcal{G}\ominus\mathcal{X}_g\,,$$

(ver Observaciones 2.2.5 y 1.3.5), por el Teorema 2.2.3 tenemos que $\Theta_1(\mathcal{X},\mathcal{Y}) < \pi/2$ de forma que, $\mathcal{X}^\perp \cap \mathcal{Y} = \{0\} = \mathcal{X} \cap \mathcal{Y}^\perp$. Luego,

$$\mathcal{X} = (\mathcal{X}\cap\mathcal{Y}) \oplus \mathcal{X}_g \quad , \quad \mathcal{Z} = (\mathcal{X}\cap\mathcal{Y}) \oplus \mathcal{X}_g \oplus \mathcal{X}_{g^\perp} \quad \text{y} \quad \mathcal{X}_{g^\perp} = \mathcal{Z} \ominus \mathcal{X}\,.$$

Sea $X' \in \mathcal{M}_{d,(p-k)}(\mathbb{C})$ tal que sus columnas forman una BON de $\mathcal{X}_{g^\perp} \subset \mathcal{X}^\perp$. Entonces, $X'_S = S^*\,X' \in \mathcal{M}_{p,(p-k)}(\mathbb{C})$ es una isometría de $\mathbb{C}^{p-k}$ en $S^*\mathcal{X}_{g^\perp} = (S^*\mathcal{X})^\perp \subseteq \mathbb{C}^p$.

Para chequear la propiedad de separación DKN para $(A_{\mathcal{X}+\mathcal{Y}}\,,\, S^*\mathcal{X}\,,\, S^*\mathcal{Y})$ consideramos los autovalores de

$$(X'_S)^*(S^*A\,S)\,X'_S = (X')^*\,S\,S^*\,A\,S\,S^*\,X' = (X')^*\,A\,X' \in \mathcal{H}(p-k)\,,$$

donde $SS^* = P_{\mathcal{Z}} \in \mathcal{M}_d(\mathbb{C})$, $P_{\mathcal{Z}}\,X' = X'$ y $(X')^*\,P_{\mathcal{Z}} = (X')^*$. Luego vemos que

$$\lambda_i((X'_S)^*(S^*A\,S)\,X'_S) = \lambda_i(P_{\mathcal{X}_{g^\perp}}\,A\,P_{\mathcal{X}_{g^\perp}}) \quad \text{para} \quad i \in \mathbb{I}_{p-k}\,.$$

Además como $\mathcal{X}_{g^\perp} \subset \mathcal{X}^\perp$, tenemos que $P_{\mathcal{X}_{g^\perp}}\,A\,P_{\mathcal{X}_{g^\perp}}$ es una compresión de $P_{\mathcal{X}^\perp}\,A\,P_{\mathcal{X}^\perp}$. Luego, las desigualdades de entrelace 1.1.9 muestran que si $\lambda_i((P_{\mathcal{X}^\perp}\,A\,P_{\mathcal{X}^\perp})) \in [a,b]$, para $i \in \mathbb{I}_{d-k}$, entonces

$$\lambda_i(P_{\mathcal{X}_{g^\perp}}\,A\,P_{\mathcal{X}_{g^\perp}}) \in [a,b] \quad \text{para} \quad i \in \mathbb{I}_{p-k}\,. \tag{2.22}$$

Por otro lado notemos que

$$Y_S^*\,(S^*A\,S)\,Y_S = Y^*P_{\mathcal{Z}}\,A\,P_{\mathcal{Z}}\,Y = Y^*A\,Y\,,$$

dado que, como antes, $SS^* = P_{\mathcal{Z}}$, $P_{\mathcal{Z}}Y = Y$ y $Y^*P_{\mathcal{Z}} = Y^*$. Por lo tanto, tenemos que

$$\lambda_i(Y_S^*\,(S^*A\,S)\,Y_S) = \lambda_i(Y^*A\,Y) \in (\infty, a-\delta] \cup [b+\delta,\infty) \quad \text{para} \quad i \in \mathbb{I}_k\,. \tag{2.23}$$

Luego el ítem 1. es consecuencia de las ecuaciones (2.22) y (2.23) y de que $S^*\mathcal{X} \subseteq \mathbb{C}^p$ es, por construcción, un subespacio $A_{\mathcal{X}+\mathcal{Y}}$ invariante.

Por otro lado para probar el ítem 2. fijamos una NUI $||\cdot||$. Usando que $\mathcal{X}$, $\mathcal{Y} \subset \mathcal{Z}$ y el hecho de que S^* es una isometría de $\mathcal{Z}$ en $\mathbb{C}^p$, vemos que $\Theta(\mathcal{X},\mathcal{Y}) = \Theta(S^*\mathcal{X}, S^*\mathcal{Y})$. Luego aplicando el teorema de Nakatsukasa tan Θ (Teorema 2.2.3) a la matriz autoadjunta $S^*AS \in \mathcal{H}(p)$ y los subespacios $S^*\mathcal{X}$, $S^*\mathcal{Y} \subseteq \mathbb{C}^p$ vemos que

$$\delta'\,\|\tan(\Theta(\mathcal{X}\,,\,\mathcal{Y}))\| \leq \|A_{\mathcal{X}+\mathcal{Y}}\,Y_S - Y_S\,(Y_S^*A_{\mathcal{X}+\mathcal{Y}}\,Y_S)\|\,,$$

donde $Y_S = S^*Y \in \mathcal{M}_{p,k}(\mathbb{C})$ es una isometría de $\mathbb{C}^k$ en $S^*\mathcal{Y}$. Además notamos que

$$\begin{aligned} A_{\mathcal{X}+\mathcal{Y}}\,Y_S - Y_S\,(Y_S^*A_{\mathcal{X}+\mathcal{Y}}\,Y_S) &= S^*A\,S\,S^*Y - S^*Y\,(Y^*S(S^*A\,S)S^*Y) \\ &= S^*\,(A\,Y - Y\,(Y^*A\,Y))\,, \end{aligned}$$

donde hemos usado que $SS^* = P_{\mathcal{Z}}$, $P_{\mathcal{Z}}\,Y = Y$ y $Y^*\,P_{\mathcal{Z}} = Y^*$. Por lo tanto

$$\|A_{\mathcal{X}+\mathcal{Y}}\,Y_S - Y_S\,(Y_S^*A_{\mathcal{X}+\mathcal{Y}}\,Y_S)\| = \|P_{\mathcal{Z}}\,(A\,Y - Y\,(Y^*A\,Y))\| = \|P_{\mathcal{X}+\mathcal{Y}}\,R_Y\|\,.$$

□

Observación 2.2.7. Utilizando la notación del Teorema 2.2.6 y la Observación 1.2.4, la ecuación (2.21) es equivalente a la siguiente relación de submayorización

$$\delta'\,\tan(\Theta(\mathcal{X}\,,\,\mathcal{Y})) \prec_w s(A_{\mathcal{X}+\mathcal{Y}}\,Y_S - Y_S\,(Y_S^*A_{\mathcal{X}+\mathcal{Y}}\,Y_S)) = s(P_{\mathcal{X}+\mathcal{Y}}\;R_Y)\,,$$

en términos de la constante de separación δ' para $A_{\mathcal{X}+\mathcal{Y}} = S^*A\,S$, $S^*\mathcal{X}$ y $S^*\mathcal{Y}$. △

Considerando la notación del Teorema 2.2.6, sea $\delta > 0$ tal que $(A\,,\,\mathcal{X}\,,\,\mathcal{Y}\,,\,\delta)$ satisface la propiedad de separación DKN. Dada una NUI $||\cdot||$, el Teorema 2.2.3 permite acotar superiormente $\|\tan(\Theta(\mathcal{X},\mathcal{Y}))\|$ por

$$\|\tan(\Theta(\mathcal{X},\mathcal{Y}))\| \leq \frac{\|R_Y\|}{\delta}\,. \tag{2.24}$$

Por otro lado, el ítem 2. del Teorema 2.2.6 establece que, existe $\delta' \geq \delta > 0$ tal que $(A_{\mathcal{X}+\mathcal{Y}}\,,\,S^*\mathcal{X}\,,\,S^*\mathcal{Y}\,,\,\delta')$ satisface la propiedad de separación DKN, de forma que obtenemos la cota superior

$$\|\tan(\Theta(\mathcal{X},\mathcal{Y}))\| \leq \frac{\|P_{\mathcal{X}+\mathcal{Y}}\,R_Y\|}{\delta\,'}\,. \tag{2.25}$$

Dado que $\|P_{\mathcal{X}+\mathcal{Y}}\,R_Y\| \leq \|R_Y\|$ y que $\delta \leq \delta\,'$, notamos que la cota superior de la ecuación (2.25) mejora la cota clásica (2.24).

Para comparar en mas detalle estos resultados consideramos el siguiente

Ejemplo 2.2.8. Sea $\tilde{\lambda} = (a,b,d,c) \in \mathbb{R}^4$, donde $a < b < c < d$ y sea $\tilde{A} \in \mathcal{H}(4)$ dada por $\tilde{A} = D_{\tilde{\lambda}}$. A los efectos de este ejemplo, consideramos el parámetro real $c \in (b,d)$ como una variable (y fijamos a, b, d).

Sean $\mathcal{X}$, $\mathcal{Y}_\theta \subset \mathbb{C}^4$ como en el Ejemplo 2.1.7, es decir, $\mathcal{X} = \text{span}\{e_1,\, e_2\}$ e $\mathcal{Y}_\theta = \text{span}\{e_1,\, f_\theta\}$. Recordemos que $\Theta(\mathcal{X},\mathcal{Y}_\theta) = (\theta, 0)$. En particular, en este caso $\tan(\Theta(\mathcal{X},\mathcal{Y}_\theta)) = (\tan\theta, 0)$.

Por lo tanto $\mathcal{X} + \mathcal{Y}_\theta = \text{span}\{e_1,\, e_2,\, e_3\}$.

Ahora bien, sean

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad , \quad X_\perp = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{y} \quad Y_\theta = \begin{pmatrix} 1 & 0 \\ 0 & \cos\theta \\ 0 & \text{sen}\,\theta \\ 0 & 0 \end{pmatrix}\,.$$

Entonces, tenemos que $\lambda(Y_\theta^*\tilde{A}Y_\theta) = (b\,\cos^2\theta + d\,\text{sen}^2\,\theta, a)$, mientras que $\lambda(X_\perp^*\tilde{A}\,X_\perp) = (d,c)$. Luego, tomando $\theta_0(c) = \theta_0 = \arcsin\left(\sqrt{\frac{c-b}{d-b}}\right)$ y

$$\delta_\theta = c - (b\,\cos^2\theta + d\,\text{sen}^2\,\theta) > 0 \quad \text{para} \quad 0 < \theta < \theta_0\,,$$

resulta que $(\tilde{A},\mathcal{X},\mathcal{Y}_\theta,\delta_\theta)$ satisface la propiedad de separación DKN, δ_θ es la constante de separación óptima (más grande) y la propiedad de separación vale sólo para $0 < \theta < \theta_0$.

Nuevamente, cálculos simples permiten obtener que $s(R_{Y_\theta}) = ((d-b)\cos\theta\,\text{sen}\,\theta, 0)$.

En este caso, la ecuación (2.24) del Teorema 2.2.3 se traduce en

$$\tan\theta \leq \frac{(d-b)\cos\theta\,\text{sen}\,\theta}{c - (b\,\cos^2\theta + d\,\text{sen}^2\,\theta)} \quad \text{para} \quad 0 < \theta < \theta_0\,. \tag{2.26}$$

Notemos que $\lim_{c\to b^+}\theta_0 = 0$ es decir, el rango de θ para el cual podemos aplicar la cota de la ecuación (2.26) tiende a ser pequeño. En el caso límite, cuando $b = c$ (tenemos multiplicidad de autovalores) no podemos aplicar la cota de la ecuación (2.26) (porque la constante de separación es $\delta_0 = 0$). Finalmente si consideramos el caso límite para el cual θ es pequeño, entonces la cota superior es comparable con la cota $(\frac{d-b}{c-b})\,\tan\theta$ $(> \tan\theta)$.

Por otro lado, $\mathcal{X} + \mathcal{Y}_\theta \ominus \mathcal{X} = \mathbb{C}\,e_3$, es el subespacio generado por e_3. En este caso tomando $X' = (0,0,1,0)^t$, resulta que $\lambda((X'_S)^*\tilde{A}\,X'_S) = d$. Esto implica que, si consideramos $\delta'_\theta = d - (b\,\cos^2\theta + d\,\text{sen}^2\,\theta) = (d-b)\,\cos^2\theta > 0$, para $\theta \in (0,\pi/2)$, obtenemos que $(\tilde{A}_{\mathcal{X}+\mathcal{Y}_\theta}, S^*\mathcal{X}, S^*\mathcal{Y}_\theta, \delta'_\theta)$ satisface la propiedad de

separación DKN, donde $S \in \mathcal{M}_{4,3}(\mathbb{C})$ es la matriz cuyas columnas son los tres primeros elementos de la base canónica. En este caso tenemos que

$$\frac{s_1(P_{\mathcal{X}+\mathcal{Y}_\theta}\, R_{Y_\theta})}{\delta'_\theta} = \frac{(d-b)\cos\theta\ \operatorname{sen}\theta}{(d-b)\ \cos^2\theta} = \tan\theta\,,$$

y por lo tanto, la cota superior de la ecuación (2.25) coincide con $\tan\theta$, donde $\tan(\Theta(\mathcal{X}, \mathcal{Y}_\theta)) = (\tan\theta, 0)$, es decir, la cota superior es ajustada. Además notemos que esta cota es aplicable para todo $\theta \in (0, \pi/2)$. △

El siguiente resultado fue conjeturado en [33] por Knyazev y Zhu.

Corolario 2.2.9. *Sean $A \in \mathcal{H}(d)$, $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$ y $\delta > 0$ tales que $(A\,, \mathcal{X}\,, \mathcal{Y}\,, \delta)$ satisface la propiedad de separación DKN. Entonces,*

$$\delta\, \|\tan(\Theta(\mathcal{X}\,, \mathcal{Y}))\| \leq \|P_{\mathcal{X}+\mathcal{Y}}\ R_Y\|\,,$$

para toda NUI $\|\cdot\|$.

Demostración. Sea $S \in \mathcal{M}_{d,p}(\mathbb{C})$ tal que sus columnas forman una BON para $\mathcal{X}+\mathcal{Y}$. Por el ítem 1. del Teorema 2.2.6, existe $\delta' \geq \delta$ tal que $(S^*A\,S\,,\ S^*\mathcal{X}\,,\ S^*\mathcal{Y}\,,\ \delta')$ satisface la propiedad de separación DKN. Por el ítem 2. de este resultado, tenemos que

$$\delta\, \|\tan(\Theta(\mathcal{X}\,, \mathcal{Y}))\| \leq \delta'\, \|\tan(\Theta(\mathcal{X}\,, \mathcal{Y}))\| \leq \|P_{\mathcal{X}+\mathcal{Y}}\ R_Y\|\,.$$

□

Finalmente, obtenemos la siguiente cota cuadrática para el error a posteriori para aproximación simultanea de autovalores de A mediante valores de Ritz correspondientes a cocientes de Rayleigh, para los cuales la propiedad de separación DKN es válida.

Teorema 2.2.10. *Sean $A \in \mathcal{H}(d)$, $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$ y $\delta > 0$ tales que $(A\,,\ \mathcal{X}\,,\ \mathcal{Y}\,,\ \delta)$ satisface la propiedad de separación DKN. Entonces, para toda NUI $\|\cdot\|$ tenemos que*

$$\|\lambda(X^*A\,X) - \lambda(Y^*A\,Y)\| \leq \frac{\|P_{\mathcal{X}+\mathcal{Y}}\, R_Y\|^2}{\delta}\,.$$

Demostración. Este resultado es una consecuencia directa del Corolario 2.1.6 y el Teorema 2.2.6. □

El Teorema 2.2.10 permite obtener la siguiente extensión del [58, Teorema 5.3] (ver Observación 2.2.12 debajo) que constituye una cota de mayorización cuadrática a posteriori para el error de aproximación simultanea para autovalores consecutivos.

Corolario 2.2.11. *Sean $A \in \mathcal{H}(d)$ e $\mathcal{Y} \subset \mathbb{C}^d$ un subespacio k-dimensional, tales que:*

1. *$\lambda_1(Y^*AY) < \lambda_j(A)$, donde $j \in \mathbb{I}_{d-k}$ es el mayor índice que realiza la desigualdad;*
2. *$\lambda_i(Y^*AY) \geq \lambda_{i+j}(A)$, para $i \in \mathbb{I}_k$.*

Sea $\mathcal{U}$ el subespacio A-invariante generado por los autovectores asociados a $\lambda_i(A)$, para $1 \leq i \leq j$ y sea $\mathcal{X} = (I - P_{\mathcal{U}})\mathcal{Y}$. Si $\eta = \lambda_j(A) - \lambda_1(Y^*AY) > 0$ entonces

$$\|(\lambda_{i+j}(A))_{i\in\mathbb{I}_k} - \lambda(Y^*A\,Y)\| \leq \frac{\|P_{\mathcal{X}+\mathcal{Y}}\, R_Y\|^2}{\eta}\,,$$

para toda NUI $\|\cdot\|$.

Demostración. Sea $\mathcal{V} = \mathcal{U} + \mathcal{Y}$, notemos que $\mathcal{U} \cap \mathcal{Y} = \{0\}$ (por la condición del ítem 1.); luego, $p = \dim(\mathcal{V}) = \dim(\mathcal{U}) + k$ es decir $j = \dim(\mathcal{U}) = p - k$.

Más aún, $\mathcal{V} \ominus \mathcal{U} = (I - P_{\mathcal{U}})\mathcal{Y} = \mathcal{X}$; en particular, $\dim(\mathcal{X}) = \dim(\mathcal{Y})$ y $\mathcal{V} \ominus \mathcal{X} = \mathcal{U}$. Además notemos que $\Theta_1(\mathcal{X}, \mathcal{Y}) < \pi/2$, de otra forma tendríamos que $\mathcal{U} \cap \mathcal{Y} \neq \{0\}$, pues $\mathcal{V} \ominus \mathcal{X} = \mathcal{U}$.

Sea $V \in \mathcal{I}_{\mathcal{V}}(p, d)$ y tomemos $A_V = V^* A V \in \mathcal{H}(p)$. De forma similar consideramos X, $Y \in \mathcal{M}_{d,k}(\mathbb{C})$, $U \in \mathcal{M}_{d,p-k}(\mathbb{C})$ tales que sus columnas forman bases ortonormales de $\mathcal{X}$, $\mathcal{Y}$ y $\mathcal{U}$ respectivamente; tomemos $X_V = V^* X$, $Y_V = V^* Y \in \mathcal{M}_{p,k}(\mathbb{C})$ y $U_V = V^* U \in \mathcal{M}_{p,p-k}(\mathbb{C})$. Entonces las columnas de U_V generan $\mathcal{U}_V \subset \mathbb{C}^p$ que es un subespacio invariante de A_V.

En particular, las columnas de X_V generan $\mathcal{X}_V \subset \mathbb{C}^p$, que también es un subespacio invariante de A_V. En este caso $\mathcal{X}_V^{\perp} = \mathcal{U}_V$ y $\Theta_1(\mathcal{X}_V, \mathcal{Y}_V) = \Theta_1(\mathcal{X}, \mathcal{Y}) < \pi/2$, donde $\mathcal{Y}_V \subset \mathbb{C}^p$ es el subespacio generado por las columnas de Y_V. Además por construcción,

$$\lambda_i(Y_V^* A_V\, Y_V) = \lambda_i(Y^* A\, Y) \quad \text{para} \quad i \in \mathbb{I}_k\,.$$

Como $\mathcal{X} \subset \mathcal{U}^{\perp}$ las desigualdades de entrelace 1.1.9 para compresiones de matrices autoadjuntas y el ítem 2. de la hipótesis, permiten obtener que para $i \in \mathbb{I}_k$,

$$\lambda_i(X_V^* A_V\, X_V) = \lambda_i(X^* A\, X) \leq \lambda_i(A_{U_\perp}) = \lambda_{j+i}(A) \leq \lambda_i(Y_V^* A_V\, Y_V)\,, \tag{2.27}$$

donde $U_\perp \in \mathcal{I}_{\mathcal{U}^\perp}(d-j, d)$. Por otro lado, por hipótesis, $(A_V, \mathcal{X}_V, \mathcal{Y}_V, \eta)$ satisface la propiedad de separación DKN (recordar que $\mathcal{X}_V^{\perp} = \mathcal{U}_V$). Por lo tanto, el Teorema 2.2.10 permite concluir que

$$\|\lambda(X_V^* A_V\, X_V) - \lambda(Y_V^* A_V\, Y_V)\| \leq \frac{\|\, P_{\mathcal{X}_V + \mathcal{Y}_V}\,(A_V\, Y_V - Y_V\,(Y_V^* A_V\, Y_V))\,\|^2}{\eta}\,. \tag{2.28}$$

Ahora bien, la ecuación (2.27) establece que

$$|(\lambda_{i+j}(A))_{i \in \mathbb{I}_k} - \lambda(Y_V^* A_V\, Y_V)| \prec_w |\lambda(X_V^* A_V\, X_V) - \lambda(Y_V^* A_V\, Y_V)|\,.$$

Además, un argumento análogo al de la prueba del Teorema 2.2.6 muestra que

$$\|\, P_{\mathcal{X}_V + \mathcal{Y}_V}\,(A_V\, Y_V - Y_V\,(Y_V^* A_V\, Y_V))\,\| = \|\, P_{\mathcal{X}+\mathcal{Y}}\, R_Y\,\|\,.$$

Luego el resultado se obtiene combinando los dos hechos anteriores con la ecuación (2.28) y la Observación 1.2.4.

□

Observación 2.2.12. Notamos que la hipótesis en el ítem 1. del Corolario 2.2.11 es que exista un autovalor β de A tal que $\lambda_1(Y^* A Y) < \beta$. En efecto, en este caso podemos aplicar las desigualdades de entrelace 1.1.9 para obtener que $\lambda_i(Y^* A Y) \geq \lambda_{d-k+i}(A)$, para $i \in \mathbb{I}_k$. Entonces, $\beta = \lambda_j(A)$ para algún $1 \leq j \leq d - k$.

Por otro lado, la hipótesis del ítem 2. es bastante restrictiva y difícil de chequear en general. Sin embargo, mencionamos dos casos en los cuales las hipótesis del Corolario 2.2.11 pueden ser chequeadas fácilmente :

1. Si se cumple la hipótesis del ítem 1. para $j = d - k$, por las desigualdades de entrelace tenemos que

$$\lambda_i(Y^* A Y) \geq \lambda_{i+d-k}(A) \quad \text{para} \quad i \in \mathbb{I}_k\,,$$

y por lo tanto las hipótesis del ítem 2. se cumplen automáticamente.

2. En el caso $k = 1$, es decir si $\mathcal{Y} = \mathbb{C}\, y$ para algún vector unitario $y \in \mathbb{C}^d$, las hipótesis se convierten en la existencia de un $j \in \mathbb{I}_{d-1}$ tal que $\lambda_{j+1}(A) \leq \langle A\, y,\, y\rangle < \lambda_j(A)$; entonces, el Corolario 2.2.11 implica que
$$0 \leq \langle Ay,\, y\rangle - \lambda_{j+1}(A) \leq \frac{\|P_{\mathcal{X}+\mathcal{Y}}(Ay - \langle Ay,\, y\rangle\, y)\|}{\lambda_j(A) - \langle Ay,\, y\rangle}\,,$$
donde $\mathcal{X} = \mathbb{C}\, x$, para $x = (I - P_U)y \in \mathbb{C}^d$; que coincide con [58, Teorema 5.3]. De hecho, como es mencionado en [58], el Corolario 2.2.11 codifica varias cotas conocidas para estimaciones de autovalores, incluso en el caso $k = 1$. △

2.3. Aplicaciones: cotas de mayorización a priori para valores de Ritz

En esta sección estableceremos una relación entre las cotas mixtas para el error en los valores de Ritz, obtenidas en la Sección 2.1 y también con algunas cotas de mayorización a priori consideradas en [30, 31].

La siguiente definición corresponde al concepto de dispersión espectral, que estará presente en los capítulos siguientes. Sin embargo, no ahondaremos todavía en su rol y sus propiedades. En [30], A. Knyazev et.al. consideraron como medida vectorial de la dispersión del espectro de una matriz autoadjunta A a lo que llamamos *dispersión espectral completa* de A.

Definición 2.3.1. Sea $A \in \mathcal{H}(d)$. Definimos la *dispersión espectral completa* de A como el vector
$$\mathrm{Spr}(A) = \lambda(A) + \lambda(-A) = \big(\lambda_i(A) - \lambda_i^{\uparrow}(A)\,\big)_{i\in\mathbb{I}_d} \in \mathbb{R}^d\,. \tag{2.29}$$

Notemos que $\mathrm{Spr}(A) \in \mathbb{R}^d$ es un vector antisimétrico, es decir $\mathrm{Spr}_i(A) = -\mathrm{Spr}_{d-i+1}(A)$, para $i \in \mathbb{I}_d$.

Además definimos la *dispersión espectral* de A como
$$\mathrm{Spr}^+(A) = (\lambda_i(A) - \lambda_{d-i+1}(A))_{i\in\mathbb{I}_{[d/2]}} = \big(\lambda_i(A) - \lambda_i^{\uparrow}(A)\,\big)_{i\in\mathbb{I}_{[d/2]}} \in \mathbb{R}^{[d/2]}_{\geq 0}\,, \tag{2.30}$$
donde $[d/2]$ denota la parte entera de $d/2$.

△

Observación 2.3.2. Sean $A \in \mathcal{H}(d)$ y $\mathcal{Z} \subset \mathbb{C}^d$ un subespacio de $\dim(\mathcal{Z}) = p$. La dispersión espectral de A relativa a $\mathcal{Z}$, notada por $\mathrm{Spr}(A\,,\,\mathcal{Z})$, está dada por
$$\mathrm{Spr}(A, \mathcal{Z}) = \lambda(A_{\mathcal{Z}}) - \lambda^{\uparrow}(A_{\mathcal{Z}}) = (\lambda_i(A_{\mathcal{Z}}) - \lambda_{p-i+1}(A_{\mathcal{Z}}))_{i\in\mathbb{I}_p} \in (\mathbb{R}^p)^{\downarrow}\,,$$
donde $A_{\mathcal{Z}} = P_{\mathcal{Z}}\, A|_{\mathcal{Z}} \in \mathcal{L}(\mathcal{Z})$ es un operador autoadjunto (definido de la forma usual). Notemos que si $\mathcal{Z} = \mathbb{C}^d$, entonces $\mathrm{Spr}(A, \mathbb{C}^d) = \mathrm{Spr}(A)$.

Por un lado, sean $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$ subespacios tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$. En este caso, si $s = \dim(\mathcal{X} \cap \mathcal{Y})$, entonces $\dim(\mathcal{X} + \mathcal{Y}) = 2k - s$ y las primeras $[\frac{2k-s}{2}]$ entradas del vector $\mathrm{Spr}(A, \mathcal{X} + \mathcal{Y})$ son no negativas. Por otro lado, los ángulos principales $\Theta(\mathcal{X}, \mathcal{Y})$ tienen $k - s$ entradas no nulas. De esta forma, como $[\frac{2k-s}{2}] \geq k - s$, podemos considerar la siguiente igualdad de vectores
$$(\mathrm{Spr}_i(A, \mathcal{X} + \mathcal{Y})\ \mathrm{sen}(\Theta_i(\mathcal{X}, \mathcal{Y}))\,)_{i\in\mathbb{I}_k} = \mathrm{Spr}^+(A, \mathcal{X} + \mathcal{Y})\ \mathrm{sen}(\Theta(\mathcal{X}, \mathcal{Y}))\,.$$

Notemos que $\mathrm{Spr}^+(A, \mathcal{X}+\mathcal{Y})$ tiene sólo $[\frac{2k-s}{2}]$ entradas, por lo que el producto $\mathrm{Spr}^+(A, \mathcal{X}+\mathcal{Y})\ \mathrm{sen}(\Theta(\mathcal{X}, \mathcal{Y}))$ debe entenderse en términos de la Observación 1.1.2.

Este argumento permite describir las conjeturas de Argentati y Knyazev (2.31) y (2.32) a continuación, en términos de la dispersión espectral Spr^+.

△

Observación 2.3.3 (Cotas a priori para variaciones en los valores de Ritz: conjeturas y trabajos previos)**.** Sea $A \in \mathcal{H}(d)$ y sean $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$ subespacios tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$. En [30] los autores conjeturaron que la siguiente cota de submayorización para la variación absoluta de los valores de Ritz es válida:

$$|\lambda(X^*AX) - \lambda(Y^*AY)| \prec_w \operatorname{Spr}^+(A, \mathcal{X}+\mathcal{Y}) \operatorname{sen}(\Theta(\mathcal{X}, \mathcal{Y}))\,. \tag{2.31}$$

Más aún, en caso que $\mathcal{X}$ sea A-invariante, los autores conjeturaron que

$$|\lambda(X^*AX) - \lambda(Y^*AY)| \prec_w \operatorname{Spr}^+(A, \mathcal{X}+\mathcal{Y}) \operatorname{sen}(\Theta(\mathcal{X}, \mathcal{Y}))^2\,. \tag{2.32}$$

Estas conjeturas son extensiones naturales de los resultados obtenidos en [31] (para el caso $k = 1$).

Aunque la [30, Conjetura 2.1.] afirma la validez de las ecuaciones (2.31) y (2.32) para subespacios arbitrarios $\mathcal{X}$ y $\mathcal{Y}$ tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y})$, este tipo de cotas son importantes cuando $\mathcal{Y}$ es una (pequeña) perturbación del subespacio $\mathcal{X}$. En este caso la validez de las ecuaciones (2.31) y (2.32) asegurarían distintos ordenes de aproximación de X^*AX por Y^*AY en términos de ángulos principales y de la dispersión espectral de A (es decir, considerando A, $\mathcal{X}$ e $\mathcal{Y}$ variables). Observemos que este tipo de resultados tendrían aplicaciones inmediatas en el estudio de estabilidad numérica y convergencia de métodos iterativos relacionados con algoritmos del tipo Rayleigh-Ritz.

En [30, Teorema 2.1.] los autores probaron que, en general,

$$|\lambda(X^*AX) - \lambda(Y^*AY)| \prec_w (\lambda_{\text{máx}}(A_{\mathcal{X}+\mathcal{Y}}) - \lambda_{\text{mín}}(A_{\mathcal{X}+\mathcal{Y}})) \operatorname{sen}(\Theta(\mathcal{X}, \mathcal{Y}))\,, \tag{2.33}$$

mientras que cuando $\mathcal{X}$ es A-invariante,

$$|\lambda(X^*AX) - \lambda(Y^*AY)| \prec_w (\lambda_{\text{máx}}(A_{\mathcal{X}+\mathcal{Y}}) - \lambda_{\text{mín}}(A_{\mathcal{X}+\mathcal{Y}})) \operatorname{sen}(\Theta(\mathcal{X}, \mathcal{Y}))^2\,, \tag{2.34}$$

donde $A_{\mathcal{X}+\mathcal{Y}} = P_{\mathcal{X}+\mathcal{Y}}\, A|_{\mathcal{X}+\mathcal{Y}} \in \mathcal{L}(\mathcal{X}+\mathcal{Y})$; más aún, en [30, Teorema 2.2.] probaron que en el caso particular en que $\mathcal{X}$ es el subespacio A-invariante correspondiente a los k autovalores más grandes de A, entonces

$$0 \leq \lambda(X^*AX) - \lambda(Y^*AY) \prec_w (\lambda_i(A_{\mathcal{X}+\mathcal{Y}}) - \lambda_{\text{mín}}(A_{\mathcal{X}+\mathcal{Y}}))_{i\in\mathbb{I}_k} \operatorname{sen}(\Theta(\mathcal{X}, \mathcal{Y}))^2\,. \tag{2.35}$$

Notar que la ecuación (2.35) constituye una cota mas fuerte que la de la ecuación (2.34); sin embargo es mas débil que la cota de la ecuación (2.32), dado que $\operatorname{Spr}_i(A, \mathcal{X}+\mathcal{Y}) \leq \lambda_i(A_{\mathcal{X}+\mathcal{Y}}) - \lambda_{\text{mín}}(A_{\mathcal{X}+\mathcal{Y}})$, para $i \in \mathbb{I}_k$.
Además, notemos que las cotas de las ecuaciones (2.33) y (2.34) resultan mucho más débiles que las conjeturadas en (2.31) y (2.32); en efecto, la cota $\lambda_{\text{máx}}(A_{\mathcal{X}+\mathcal{Y}}) - \lambda_{\text{mín}}(A_{\mathcal{X}+\mathcal{Y}})$ no es sensible a los autovalores intermedios de $A_{\mathcal{X}+\mathcal{Y}}$. Por ejemplo, tomemos $A \in \mathcal{H}(2d)$ tal que $\lambda(A) = (\mathbb{1}_{2d-1}, -1)$ y $\mathcal{X}+\mathcal{Y} = \mathbb{C}^d$. En este caso $\lambda_{\text{máx}}(A) - \lambda_{\text{mín}}(A) = 2$ mientras que $\operatorname{Spr}^+(A) = (2, 0_{n-1})$, de forma que las cotas que proponen las conjeturas (2.31) y (2.32) resultan notablemente mas pequeñas.

En lo que sigue aplicaremos el Teorema 2.1.4 para obtener resultados relacionados con las conjeturas de [30] descriptas en las ecuaciones (2.31) y (2.32).

Para obtener estos resultados probaremos primero la siguiente

Proposición 2.3.4. *Sea $A \in \mathcal{H}(d)$ y sean $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$ tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$. Entonces*

$$s(P_{\mathcal{X}}\, R_Y) \prec_w \operatorname{Spr}^+(A, \mathcal{X}+\mathcal{Y}) \operatorname{sen}(\Theta(\mathcal{X}, \mathcal{Y}))\,. \tag{2.36}$$

Demostración. Dada una matriz $A \in \mathcal{M}_d(\mathbb{C})$, consideramos $A \in \mathcal{L}(\mathbb{C}^d)$ (el operador que induce A de forma natural). Esto permite pensar que $A \in \mathcal{L}(\mathbb{C}^d)$ tiene distintas representaciones matriciales por bloques con respecto a descomposiciones ortogonales $\mathbb{C}^d = \mathcal{V} \oplus \mathcal{V}^\perp$, eligiendo subespacios adecuados $\mathcal{V} \subset$

$\mathbb{C}^d$. Utilizaremos el siguiente argumento: sea $\mathcal{Z} = \mathcal{X}+\mathcal{Y}$ con $\dim(\mathcal{Z}) = p$, consideremos las representaciones matriciales con respecto a la descomposición $\mathbb{C}^d = \mathcal{Z} \oplus \mathcal{Z}^\perp$:

$$P_{\mathcal{X}} = \begin{pmatrix} P^{\mathcal{X}} & 0 \\ 0 & 0 \end{pmatrix} \ , \ P_{\mathcal{Y}} = \begin{pmatrix} P^{\mathcal{Y}} & 0 \\ 0 & 0 \end{pmatrix} \quad \text{y} \quad A = \begin{pmatrix} A_{\mathcal{Z}} & * \\ * & * \end{pmatrix},$$

donde $P^{\mathcal{X}}$, $P^{\mathcal{Y}}$, $A_{\mathcal{Z}} = P_{\mathcal{Z}} A|_{\mathcal{Z}} \in \mathcal{L}(\mathcal{Z})$ son operadores autoadjuntos. En este caso tenemos que

$$P_{\mathcal{X}}\,(A\,P_{\mathcal{Y}} - P_{\mathcal{Y}}\,A\,P_{\mathcal{Y}}) = \begin{pmatrix} P^{\mathcal{X}}\,(A_{\mathcal{Z}}\,P^{\mathcal{Y}} - P^{\mathcal{Y}}\,A_{\mathcal{Z}}\,P^{\mathcal{Y}}) & 0 \\ 0 & 0 \end{pmatrix}.$$

Por otro lado, cálculos simples permiten ver que

$$(s(P_{\mathcal{X}} R_Y), 0_{d-k}) = s(P_{\mathcal{X}}\,(A\,P_{\mathcal{Y}} - P_{\mathcal{Y}}\,A\,P_{\mathcal{Y}})) \in (\mathbb{R}^d_{\geq 0})^{\downarrow}.$$

Luego, $(s(P_{\mathcal{X}} R_Y), 0_{p-k}) = s(P^{\mathcal{X}}\,(A_{\mathcal{Z}}\,P^{\mathcal{Y}} - P^{\mathcal{Y}}\,A_{\mathcal{Z}}\,P^{\mathcal{Y}})) = s(P^{\mathcal{X}}(I_{\mathcal{Z}} - P^{\mathcal{Y}})\,A_{\mathcal{Z}}\,P^{\mathcal{Y}})$. En consecuencia, podemos asumir que $\mathbb{C}^d = \mathcal{Z} = \mathcal{X} + \mathcal{Y}$ y probar que

$$(s(P_{\mathcal{X}}\,R_Y), 0_{d-k}) = s(P_{\mathcal{X}}\,(P_{\mathcal{Y}^\perp} A\,P_{\mathcal{Y}})) \prec_w (\mathrm{Spr}^+(A)\,\mathrm{sen}(\Theta(\mathcal{X},\mathcal{Y})), 0_{d-k})\,. \tag{2.37}$$

Ahora bien por el ítem 2. del Teorema 1.1.5,

$$s(P_{\mathcal{X}} P_{\mathcal{Y}^\perp} A P_{\mathcal{Y}}) = s(P_{\mathcal{X}} P_{\mathcal{Y}^\perp} P_{\mathcal{Y}^\perp} A P_{\mathcal{Y}}) \prec_w s(P_{\mathcal{X}} P_{\mathcal{Y}^\perp})\, s(P_{\mathcal{Y}^\perp} A P_{\mathcal{Y}}). \tag{2.38}$$

Por un lado por el Teorema 2.1.2, tenemos que $s(P_{\mathcal{X}} P_{\mathcal{Y}^\perp}) = (\mathrm{sen}(\Theta(\mathcal{X},\mathcal{Y})), 0_{d-k})$. Por otro lado, consideramos la matriz inducida por la descomposición $\mathbb{C}^d = \mathcal{Y} \oplus \mathcal{Y}^\perp$:

$$A = \begin{pmatrix} A_{11} & A_{21}^* \\ A_{21} & A_{22} \end{pmatrix} \text{ y tomamos } A_1 := \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} \ , \ A_2 := \begin{pmatrix} 0 & A_{21}^* \\ A_{21} & 0 \end{pmatrix}, \tag{2.39}$$

luego $A = A_1 + A_2$. Bajo estas consideraciones, $A_1 = P_{\mathcal{Y}}\,A\,P_{\mathcal{Y}} + P_{\mathcal{Y}^\perp}\,A\,P_{\mathcal{Y}^\perp}$ y por lo tanto por el ítem 3. del Teorema 1.1.4, $\lambda(A_1) \prec \lambda(A)$. En consecuencia

$$-\lambda^{\uparrow}(A_1) \prec -\lambda^{\uparrow}(A)\,. \tag{2.40}$$

Usando la desigualdad aditiva de Lidskii sobre $A_2 = A - A_1$ (ver ítem 1. Teorema 1.1.4)

$$\lambda(A_2) \prec \lambda(A) - \lambda^{\uparrow}(A_1)\,. \tag{2.41}$$

Combinando las ecuaciones (2.40) y (2.41), tenemos que

$$\lambda(A_2) \prec \lambda(A) - \lambda^{\uparrow}(A) = \mathrm{Spr}(A) \in \mathbb{R}^d\,. \tag{2.42}$$

Más aún, la Proposición 1.1.10, asegura que $\lambda(A_2) = (s(A_{21}), -s(A_{21}^*))^{\downarrow}$; en particular, $s(A_{21}) = (\lambda_i(A_2))_{i\in\mathbb{I}_k}$. Además, tenemos que $s(P_{\mathcal{Y}^\perp} A P_{\mathcal{Y}}) = (s(A_{21}), 0_{d-k})$; y por lo tanto

$$s(P_{\mathcal{Y}^\perp} A P_{\mathcal{Y}}) = (s(A_{21}), 0_{d-k}) = ((\lambda_i(A_2))_{i\in\mathbb{I}_k}, 0_{d-k}) \prec_w ((\mathrm{Spr}_i(A))_{i\in\mathbb{I}_k}, 0_{d-k})\,, \tag{2.43}$$

donde $\mathrm{Spr}(A) = (\mathrm{Spr}_i(A))_{i\in\mathbb{I}_d}$. Combinando las ecuaciones (2.38) y (2.43) junto con el Lema 1.1.6 (y utilizando la Observación 2.3.2) obtenemos que

$$s(P_{\mathcal{X}} P_{\mathcal{Y}^\perp} A P_{\mathcal{Y}}) \prec_w (\mathrm{Spr}^+(A)\,\mathrm{sen}(\Theta(\mathcal{X},\mathcal{Y})), 0_{d-k}) \in (\mathbb{R}^d_{\geq 0})^{\downarrow}.$$

Finalmente, el resultado se obtiene de la última relación de mayorización considerando sólo las primeras k entradas de cada vector.

□

Teorema 2.3.5. *Sean $A \in \mathcal{H}(d)$ y $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$ subespacios tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$. Si $\Theta_1(\mathcal{X}, \mathcal{Y}) < \frac{\pi}{2}$, Entonces*

$$|\lambda(X^* A\, X) - \lambda(Y^* A\, Y)| \;\prec_w\; \frac{2\,\mathrm{Spr}^+(A, \mathcal{X}+\mathcal{Y})\ \mathrm{sen}(\Theta(\mathcal{X}, \mathcal{Y}))}{\cos(\Theta(\mathcal{X}\,,\,\mathcal{Y}))}\,. \tag{2.44}$$

Demostración. El Teorema 2.1.4 establece que

$$|\lambda(X^* A\, X) - \lambda(Y^* A\, Y)| \prec_w \frac{s(P_{\mathcal{X}} R_Y) + s(P_{\mathcal{Y}} R_X)}{\cos(\Theta(\mathcal{X}, \mathcal{Y}))}\,.$$

Luego la Proposición (2.3.4) junto con el Lema 1.1.6 implican que

$$\frac{s(P_{\mathcal{X}} R_Y) + s(P_{\mathcal{Y}} R_X)}{\cos(\Theta(\mathcal{X}, \mathcal{Y}))} \prec_w \frac{2\,\mathrm{Spr}^+(A, \mathcal{X}+\mathcal{Y})\ \mathrm{sen}(\Theta(\mathcal{X}, \mathcal{Y}))}{\cos(\Theta(\mathcal{X}\,,\,\mathcal{Y}))}\,.$$

Por lo tanto el resultado se obtiene a partir de estas dos desigualdades. □

El siguiente resultado ilustra la dependencia de $s(P_{\mathcal{X}} R_Y)$ con respecto a $\mathrm{sen}(\Theta(\mathcal{X}, \mathcal{Y}))$ en el caso que $\mathcal{X}$ es A-invariante.

Proposición 2.3.6. *Sean $A \in \mathcal{H}(d)$ y $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$ subespacios tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$. Suponga-mos que $\mathcal{X}$ es A-invariante. Entonces,*

$$s(P_{\mathcal{X}} R_Y) \prec_w 2\ (\lambda_i(A_{\mathcal{X}+\mathcal{Y}}) - \lambda_{\mathrm{mín}}(A_{\mathcal{X}+\mathcal{Y}}))_{i\in\mathbb{I}_k}\ \ \mathrm{sen}^2(\Theta(\mathcal{X}, \mathcal{Y}))\,. \tag{2.45}$$

Demostración. Siguiendo el argumento en la prueba de la Proposición 2.3.4, podemos asumir que $\mathbb{C}^d = \mathcal{X} + \mathcal{Y}$. Bajo esta suposición, consideramos primero el caso $A \in \mathcal{M}_d(\mathbb{C})^+$ tal que $\lambda_{\mathrm{mín}}(A) = 0$, para probar que

$$s(P_{\mathcal{X}} R_Y) \prec_w 2\,(\lambda_i(A))_{i\in\mathbb{I}_k}\ \mathrm{sen}^2(\Theta(\mathcal{X}, \mathcal{Y}))\,. \tag{2.46}$$

En efecto, como $\mathcal{X}$ es A-invariante podemos escribir $A = P_{\mathcal{X}} A P_{\mathcal{X}} + P_{\mathcal{X}^\perp} A P_{\mathcal{X}^\perp}$. Con esta descomposición en mente y utilizando el hecho de que $(s(P_{\mathcal{X}} R_Y), 0_{d-k}) = s(P_{\mathcal{X}} P_{\mathcal{Y}^\perp} A P_{\mathcal{Y}})$, tenemos que

$$\begin{aligned} s(P_{\mathcal{X}} P_{\mathcal{Y}^\perp} A\, P_{\mathcal{Y}}) &= s(P_{\mathcal{X}} P_{\mathcal{Y}^\perp} P_{\mathcal{X}} A\, P_{\mathcal{X}} P_{\mathcal{Y}} + P_{\mathcal{X}} P_{\mathcal{Y}^\perp} P_{\mathcal{X}^\perp} A\, P_{\mathcal{X}^\perp} P_{\mathcal{Y}}) \\ &\prec_w s(P_{\mathcal{X}} P_{\mathcal{Y}^\perp} P_{\mathcal{X}} A\, P_{\mathcal{X}} P_{\mathcal{Y}}) + s(P_{\mathcal{X}} P_{\mathcal{Y}^\perp} A\, P_{\mathcal{X}^\perp} P_{\mathcal{Y}}) \ \stackrel{\mathrm{def}}{=}\ M\ . \end{aligned}$$

Ahora bien el ítem 2. del Teorema 1.1.5 y el hecho de que $0_d \le s(P_{\mathcal{X}}\, P_{\mathcal{Y}}) \le \mathbb{1}_d$, junto con el Teorema 2.1.2, aseguran que

$$\begin{aligned} M &\prec_w s(P_{\mathcal{X}} P_{\mathcal{Y}^\perp} P_{\mathcal{X}})\, s(A) + s(P_{\mathcal{X}} P_{\mathcal{Y}^\perp})\, s(A)\, s(P_{\mathcal{X}^\perp} P_{\mathcal{Y}}) \\ &\prec_w 2\,\lambda(A)\,(\mathrm{sen}^2(\Theta(\mathcal{X}, \mathcal{Y})), 0_{d-k}) \in (\mathbb{R}^d_{\ge 0})^{\downarrow}\,, \end{aligned}$$

pues $A \in \mathcal{M}_d(\mathbb{C})^+$ es semidefinida positiva. Por lo tanto la ecuación (2.46) se deduce del argumento anterior.

En general, para $A \in \mathcal{H}(d)$ consideramos la matriz auxiliar $\tilde{A} = A - \lambda_{\mathrm{mín}(A)}\, I \in \mathcal{M}_d(\mathbb{C})^+$. Notemos que en este caso,

$$R_Y(\tilde{A}) = \tilde{A}\, Y - Y(Y^* \tilde{A}\, Y) = A\, Y - Y(Y^* A\, Y) = R_Y\,,$$

y $\lambda(\tilde{A}) = \lambda(A) - \lambda_{\mathrm{mín}(A)}\, \mathbb{1}_d$, de forma que $\lambda_{\mathrm{mín}}(\tilde{A}) = 0$. Luego el resultado se obtiene de los hechos previos y de la ecuación (2.46) aplicada a $\tilde{A}$. □

Teorema 2.3.7. *Sean $A \in \mathcal{H}(d)$ y $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$ subespacios tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$ y supongamos que $\mathcal{X}$ es A-invariante. Si $\Theta_1(\mathcal{X}, \mathcal{Y}) < \frac{\pi}{2}$, entonces*

$$|\lambda(X^* A\, X) - \lambda(Y^* A\, Y)| \prec_w \frac{2\ (\lambda_i(A_{\mathcal{X}+\mathcal{Y}}) - \lambda_{\text{mín}}(A_{\mathcal{X}+\mathcal{Y}}))_{i\in\mathbb{I}_k}\ \operatorname{sen}^2(\Theta(\mathcal{X}, \mathcal{Y}))}{\cos(\Theta(\mathcal{X}\,, \mathcal{Y}))}\,. \tag{2.47}$$

Demostración. El resultado es consecuencia directa del Corolario 2.1.6 y la Proposición 2.3.6 con un argumento similar al de la prueba del Teorema 2.3.5. □

Corolario 2.3.8. *Sean $A \in \mathcal{H}(d)$ y $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$ subespacios tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$. Si $\Theta_1(\mathcal{X}, \mathcal{Y}) < \frac{\pi}{2}$, entonces*

$$|\lambda(X^* A\, X) - \lambda(Y^* A\, Y)| \prec_w \frac{2}{\cos(\Theta_1(\mathcal{X}, \mathcal{Y}))}\ \operatorname{Spr}(A, \mathcal{X} + \mathcal{Y})\ \operatorname{sen}(\Theta(\mathcal{X}, \mathcal{Y}))\,.$$

Más aún, si asumimos que $\mathcal{X}$ es A-invariante, entonces

$$|\lambda(X^* A\, X) - \lambda(Y^* A\, Y)| \prec_w \frac{2}{\cos(\Theta_1(\mathcal{X}, \mathcal{Y}))}\ (\lambda_i(A_{\mathcal{X}+\mathcal{Y}}) - \lambda_{\text{mín}}(A_{\mathcal{X}+\mathcal{Y}}))_{i\in\mathbb{I}_k}\ \operatorname{sen}^2(\Theta(\mathcal{X}, \mathcal{Y}))\,.$$

□

Terminamos esta sección con algunas observaciones sobre las relaciones entre los Teoremas 2.3.5 y 2.3.7, el Corolario 2.3.8 y las cotas conjeturadas en (2.31) y (2.32). Como ya mencionamos en la Observación 2.3.3, las cotas de las ecuaciones (2.31) y (2.32) serían particularmente relevantes en caso de que $\mathcal{Y}$ sea una (pequeña) perturbación de $\mathcal{X}$, es decir, en caso de que $\mathcal{X}$ y $\mathcal{Y}$ son subespacios cercanos (por ejemplo si $\Theta_1(\mathcal{X}, \mathcal{Y})$ es chico). Para simplificar la exposición, supongamos que $\Theta_1(\mathcal{X}, \mathcal{Y}) \leq \pi/4$. Esta suposición se cumple en una serie de situaciones significativas, algunas de ellas descriptas en [33, Sección 5.2.]. Podemos mencionar por ejemplo, una versión del Teorema sen(2θ) de Davis y Kahan [17, Teorema 8.2], que asegura bajo ciertas condiciones que $\Theta_1(\mathcal{X}, \mathcal{Y}) \leq \pi/4$.

En este caso, si $A \in \mathcal{H}(d)$ entonces el Corolario 2.3.8 implica que

$$|\lambda(X^* A\, X) - \lambda(Y^* A\, Y)| \prec_w (2\sqrt{2})\ \operatorname{Spr}(A, \mathcal{X} + \mathcal{Y})\ \operatorname{sen}(\Theta(\mathcal{X}, \mathcal{Y}))\,. \tag{2.48}$$

Por lo tanto, bajo las suposiciones anteriores ($\Theta_1(\mathcal{X}, \mathcal{Y}) \leq \pi/4$), la cota superior de la ecuación (2.48) tiene el orden de aproximación conjeturado (cuando se considera A así como los subespacios $\mathcal{X}$ e $\mathcal{Y}$ como variables), salvo por el factor constante $2\sqrt{2}$.

Si además asumimos que $\mathcal{X}$ es A-invariante, entonces por el mismo resultado obtenemos que

$$|\lambda(X^* A\, X) - \lambda(Y^* A\, Y)| \prec_w (2\sqrt{2})\ (\lambda_i(A_{\mathcal{X}+\mathcal{Y}}) - \lambda_{\text{mín}}(A_{\mathcal{X}+\mathcal{Y}}))_{i\in\mathbb{I}_k}\ \operatorname{sen}^2(\Theta(\mathcal{X}, \mathcal{Y}))\,. \tag{2.49}$$

Nuevamente, la cota superior en la ecuación (2.49) tiene el orden de aproximación conjeturado (cuando se considera A así como los subespacios $\mathcal{X}$ e $\mathcal{Y}$ como variables), salvo por el factor constante $2\sqrt{2}$. Además, observemos que esta cota es válida para un subespacio A-invariante arbitrario $\mathcal{X}$ (a diferencia de la cota en (2.35) de [30] que es válida sólo para elecciones especiales de subespacios A-invariantes $\mathcal{X}$).

Capítulo 3

Dispersión espectral de matrices autoadjuntas

Dedicaremos este capítulo a describir con más detalle el concepto de dispersión espectral de matrices autoadjuntas (ver Definición 2.3.1); motivados por la [30, Conjetura 2.1.] (ver ecuaciones (2.31) y (2.32)) que está descripta en términos de la dispersión espectral de una matriz autoadjunta A; y por otro lado, por la aparición de cierta forma necesaria, de la dispersión espectral como estimación en términos de submayorización para $s(P_{\mathcal{Y}} A\, P_{\mathcal{Y}^\perp})$ (ver Proposición 2.3.4).

Aunque resulta un concepto natural, la dispersión espectral de matrices autoadjuntas parece no haber sido considerada en la literatura, salvo por su introducción en [30]. Así, comenzaremos con una sección donde presentamos algunos resultados básicos relacionados con esta noción. Después de esto, consideraremos una desigualdad de submayorización para la dispersión espectral que tendrá un papel clave en este capítulo (pero que también impactará de forma directa en el capítulo siguiente). Obtendremos además algunas consecuencias de esta desigualdad, relacionadas con la desigualdad de Zhan [54] y la noción de rotación directa entre subespacios de Davis-Kahan [17] (ver también [47], donde se desarrollan métricas angulares sobre la variedad de Grassmann de subespacios k-dimensionales de $\mathbb{C}^d$).

3.1. Propiedades básicas de la dispersión espectral

En esta sección presentaremos múltiples propiedades básicas de la dispersión espectral y describiremos la relación entre esta noción y los valores singulares (y también con otras nociones del análisis matricial).

Primero recordemos la definición de dispersión espectral

Definición 2.3.1. Sea $A \in \mathcal{H}(d)$. Definimos la *dispersión espectral completa* de A como el vector

$$\operatorname{Spr}(A) = \lambda(A) + \lambda(-A) = \big(\lambda_i(A) - \lambda_i^{\uparrow}(A)\big)_{i\in\mathbb{I}_d} \in \mathbb{R}^d\,. \tag{3.1}$$

Recordemos que $\operatorname{Spr}(A) \in \mathbb{R}^d$ es un vector antisimétrico, es decir $\operatorname{Spr}_i(A) = -\operatorname{Spr}_{d-i+1}(A)$, para $i \in \mathbb{I}_d$.

Además definimos la *dispersión espectral* de A como

$$\operatorname{Spr}^+(A) = (\lambda_i(A) - \lambda_{d-i+1}(A))_{i\in\mathbb{I}_{[d/2]}} = \big(\lambda_i(A) - \lambda_i^{\uparrow}(A)\big)_{i\in\mathbb{I}_{[d/2]}} \in \mathbb{R}^{[d/2]}_{\geq 0}\,, \tag{3.2}$$

donde $[d/2]$ denota la parte entera de $d/2$. △

Observación 3.1.1 (Propiedades inmediatas)**.** Sea $d = 2k$ o $2k+1$ de forma que $[d/2] = k$ y $A \in \mathcal{H}(d)$. Notemos $\lambda(A) = (\lambda_i)_{i\in\mathbb{I}_d}$ y $\lambda^{\uparrow}(A) = (\lambda_i^{\uparrow})_{i\in\mathbb{I}_d}$.

1. Para todo $t \in \mathbb{R}$ tenemos que

$$\mathrm{Spr}(A - tI) = \mathrm{Spr}(A) \qquad \text{y} \qquad \mathrm{Spr}^+(A - tI) = \mathrm{Spr}^+(A)\ . \tag{3.3}$$

Esto es mas bien evidente dado que $\lambda(A - t\,I) = \lambda(A) - t\,\mathbb{1}_d$.

2. Para $r \in \mathbb{I}_k$, es conocido que (ver [7])

$$\sum_{i\in\mathbb{I}_r} \lambda_i(A) \;\; = \text{máx}\left\{ \sum_{i\in\mathbb{I}_r} \langle A\,x_i\,,\,x_i\rangle : \{x_i\}_{i\in\mathbb{I}_r} \text{ es un SON} \right\} \qquad \text{y}$$

$$-\sum_{i\in\mathbb{I}_r} \lambda_i^{\uparrow}(A) \;\; = \sum_{i\in\mathbb{I}_r} \lambda_i(-A) = \text{máx}\left\{ -\sum_{i\in\mathbb{I}_r} \langle A\,y_i\,,\,y_i\rangle : \{y_i\}_{i\in\mathbb{I}_r} \text{ es un SON} \right\}.$$

Entonces para cada $r \in \mathbb{I}_k$ tenemos que

$$\sum_{i\in\mathbb{I}_r} \mathrm{Spr}_i(A) = \text{máx}\left\{ \sum_{i\in\mathbb{I}_r} \langle A\,x_i\,,\,x_i\rangle - \langle A\,y_i\,,\,y_i\rangle : \{x_i\}_{i\in\mathbb{I}_r} \text{ y } \{y_i\}_{i\in\mathbb{I}_r} \text{ son SON} \right\}. \tag{3.4}$$

3. Recordamos que como $A \in \mathcal{H}(d)$, $|\lambda(A)|^{\downarrow} = s(A)$. Luego $\text{máx}\{|\lambda_i(A)|, |\lambda_i^{\uparrow}(A)|\} \le s_i(A)$, para $i \in \mathbb{I}_k$; por lo tanto podemos concluir que

$$\mathrm{Spr}_i^+(A) \le |\lambda_i| + |\lambda_i^{\uparrow}| \le 2s_i(A) \quad \text{para todo} \quad i \in \mathbb{I}_k\ . \tag{3.5}$$

En el caso positivo tenemos que

$$A \in \mathcal{M}_d(\mathbb{C})^+ \implies \mathrm{Spr}_i(A) \le \lambda_i = s_i(A) \quad \text{para todo} \quad i \in \mathbb{I}_k\ . \tag{3.6}$$

4. Por otro lado notemos que

$$\mathrm{Spr}_i^+(A) = \lambda_i - \lambda_{d-i+1} = \underbrace{\lambda_i - \lambda_{k+1}}_{\ge 0} + \underbrace{\lambda_{k+1} - \lambda_{d-i+1}}_{\ge 0} \quad \text{para} \quad i \in \mathbb{I}_k\,.$$

Esto implica que

$$s(A - \lambda_{k+1}\,I) = \big(\,(\lambda_i - \lambda_{k+1}\,)_{i\in\mathbb{I}_k}\,,\,(\lambda_{k+1} - \lambda_{k+i}\,)_{i\in\mathbb{I}_{d-k}}\,\big)^{\downarrow} \prec \mathrm{Spr}^+(A) \in \mathbb{R}^k_{\ge 0}\ . \tag{3.7}$$

5. Además, si $B \in \mathcal{H}(d)$ entonces

$$\lambda(A) \prec \lambda(B) \in \mathbb{R}^d \implies \quad \mathrm{Spr}(A) \prec \mathrm{Spr}(B) \implies \mathrm{Spr}^+(A) \prec_w \mathrm{Spr}^+(B)\ , \tag{3.8}$$

como consecuencia directa de la Definición 2.3.1 y el Lema 1.1.6. En efecto, si $\lambda(A) \prec \lambda(B)$ entonces

$$\sum_{j=1}^{m} \lambda_j(A) \le \sum_{j=1}^{m} \lambda_j(B)\,, \quad \text{para} \quad m \in \mathbb{I}_d\,,$$

y la igualdad de trazas (junto con la desigualdad anterior) implica que

$$\sum_{j=1}^{m} \lambda_j^{\uparrow}(A) \ge \sum_{j=1}^{m} \lambda_j^{\uparrow}(B)\,, \quad \text{para} \quad m \in \mathbb{I}_d\,.$$

Luego

$$\sum_{j=1}^{m} \mathrm{Spr}_j^+(A) = \sum_{j=1}^{m} \lambda_j(A) - \lambda_j^{\uparrow}(A) \le \sum_{j=1}^{m} \lambda_j(B) - \lambda_j^{\uparrow}(B) = \sum_{j=1}^{m} \mathrm{Spr}_j^+(B) \quad \text{para} \quad m \in \mathbb{I}_k\,.$$

6. Consideremos una isometría $Z \in \mathcal{M}_{d,r}(\mathbb{C})$ es decir, $Z^*Z = I_r$, para algún $r \in \mathbb{I}_d$. Luego, $Z^*AZ \in \mathcal{M}_r(\mathbb{C})$ es una submatriz principal de un conjugado unitario de A. Por lo tanto podemos aplicar las desigualdades de entrelace 1.1.9 para obtener

$$\lambda_i^{\uparrow}(A) \le \lambda_i^{\uparrow}(Z^*AZ) \quad \text{y} \quad \lambda_i(Z^*AZ) \le \lambda_i(A) \quad \text{para} \quad i \in \mathbb{I}_r\,.$$

En consecuencia,

$$\mathrm{Spr}_i^+(Z^*AZ) \le \mathrm{Spr}_i^+(A) \quad \text{para} \quad i \in \mathbb{I}_{[r/2]} \overset{(1.1)}{\Longrightarrow} \mathrm{Spr}^+(Z^*AZ) \prec_w \mathrm{Spr}^+(A)\,. \tag{3.9}$$

Notemos que si bien el vector $\mathrm{Spr}^+(Z^*AZ)$ tiene sólo $[r/2]$ entradas, mientras que $\mathrm{Spr}^+(A)$ tiene $[d/2]$ entradas, la comparación (3.9) puede hacerse mediante la Observación 1.1.2.

7. Si $U \in \mathcal{U}(d)$, por la desigualdad de Lidskii 1.1.4 tenemos que

$$\lambda(A - U^*AU) \prec \lambda(A) + \lambda(-U^*AU) = \lambda(A) - \lambda^{\uparrow}(A) = \mathrm{Spr}(A)\,.$$

Más aún, si N es una NUI tenemos que

$$\text{máx}\left\{\, N(A - U^*AU) \;:\; U \in \mathcal{U}(d) \,\right\} = N(D_{\mathrm{Spr}(A)})\,.$$

Es decir, $|\mathrm{Spr}(A)|^{\downarrow} = \mathrm{Spr}^+(A \oplus A) = (\mathrm{Spr}^+(A)\,,\, \mathrm{Spr}^+(A))^{\downarrow}$ (agregando un 0 si d es impar) permite calcular el diámetro de la órbita unitaria de A (con respecto a cualquier NUI). En este sentido el vector $\mathrm{Spr}(A)$ puede ser considerado como una medida vectorial del diámetro de la órbita unitaria de A. △

Proposición 3.1.2. *Sea $A \in \mathcal{H}(d)$. Entonces*

$$\frac{1}{2}\,|\mathrm{Spr}(A)|^{\downarrow} = \frac{1}{2}\,\mathrm{Spr}^+(A \oplus A) \prec_w s(A)\,, \tag{3.10}$$

donde $A \oplus A = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \in \mathcal{H}(2d)$.

Demostración. Notemos por $\lambda(A) = (\lambda_i)_{i\in\mathbb{I}_d} \in (\mathbb{R}^d)^{\downarrow}$. Entonces para $1 \le i \le k = [\frac{d}{2}]$,

$$\mathrm{Spr}_{2i-1}^+(A \oplus A) = \mathrm{Spr}_{2i}^+(A \oplus A) = \mathrm{Spr}_i^+(A) = \lambda_i - \lambda_{d-i+1} \le |\lambda_i| + |\lambda_{d-i+1}|\,.$$

Por un lado tenemos que $\mathrm{Spr}^+(A \oplus A)$ es igual a

$$\mathrm{Spr}^+(A \oplus A) = (\mathrm{Spr}^+(A)\,,\, \mathrm{Spr}^+(A))^{\downarrow} \in \mathbb{R}^d \quad \text{si} \quad d = 2k\,,$$

$$\mathrm{Spr}^+(A \oplus A) = (\mathrm{Spr}^+(A)\,,\, \mathrm{Spr}^+(A)\,,\, 0)^{\downarrow} \in \mathbb{R}^d \quad \text{si} \quad d = 2k+1\,.$$

Por lo tanto para chequear que $\mathrm{Spr}^+(A \oplus A) \prec_w 2\,s(A)$ es suficiente probar que

$$\sum_{i\in\mathbb{I}_{2r}} \mathrm{Spr}_i^+(A \oplus A) \le 2 \sum_{i\in\mathbb{I}_{2r}} s_i(A) \quad \text{para} \quad r \in \mathbb{I}_k\,,$$

para los casos pares $2r$. En efecto, supongamos que la desigualdad anterior vale para $2r$, con $r \in \mathbb{I}_k$. Luego

$$\sum_{i\in\mathbb{I}_{2r+1}} \mathrm{Spr}_i^+(A \oplus A) = \sum_{i\in\mathbb{I}_{2r}} \mathrm{Spr}_i^+(A \oplus A) + \mathrm{Spr}_{2r+1}^+(A \oplus A) \overset{(3.5)}{\le} 2 \sum_{i\in\mathbb{I}_{2r}} s_i(A) + 2\,s_{2r+1}(A) \quad \text{para} \quad r \in \mathbb{I}_k\,,$$

pues $\mathrm{Spr}_{2r+1}^+(A \oplus A) = \mathrm{Spr}_{r+1}^+(A) \le 2\,s_i(A)$.

Luego para el caso $2r$, si $r \in \mathbb{I}_k$, tenemos que

$$\sum_{i\in\mathbb{I}_{2r}} \mathrm{Spr}^+(A\oplus A) \leq \sum_{i\in\mathbb{I}_{2r}} \mathrm{Spr}_i^+(A) \overset{(3.5)}{\leq} 2 \sum_{i\in\mathbb{I}_{2r}} s_i(A)\,,$$

dado que $s(A) = (|\lambda_i|)^{\downarrow}_{i\in\mathbb{I}_d}$, y $\sum\limits_{i\in\mathbb{I}_{2r}} s_i(A) = \max\{\sum\limits_{j\in\mathbb{F}} |\lambda_j| : |\mathbb{F}| = 2r\}$. Notemos que la ecuación (3.10) es consecuencia de este hecho.

Por otro lado, si $A \in \mathcal{H}(d)$ entonces, por los ítems 1. y 4. de la Observación 3.1.1 y la Proposición 3.1.2,

$$\frac{1}{2}\,\mathrm{Spr}^+(A\oplus A) \prec s\big(\,A - \lambda_k(A)\,I\,\big) \overset{(3.7)}{\prec} \mathrm{Spr}^+(A)\ . \tag{3.11}$$

□

El siguiente resultado corresponde a una versión aditiva de la desigualdad de Lidskii para la dispersión espectral.

Proposición 3.1.3. *Sean $A, B \in \mathcal{H}(d)$. Entonces*

$$\mathrm{Spr}(A) - \mathrm{Spr}(B) \prec \mathrm{Spr}(A-B) \prec \mathrm{Spr}(A) - \mathrm{Spr}^{\uparrow}(B) = \mathrm{Spr}(A) + \mathrm{Spr}(B)\ . \tag{3.12}$$

Demostración. Por la desigualdad de Lidskii 1.1.4 y el ítem 1. del Lema 1.1.6,

$$\begin{aligned}\mathrm{Spr}(A) - \mathrm{Spr}(B) \quad &= \lambda(A) - \lambda(B) + \lambda(-A) - \lambda(-B)\\ &\overset{\text{Lidskii}}{\prec} \lambda(A-B) + \lambda(-(A-B)) = \mathrm{Spr}(A-B)\ .\end{aligned}$$

Para la otra desigualdad, notemos que $\mathrm{Spr}^{\uparrow}(B) = \lambda^{\uparrow}(B) - \lambda(B) = -\mathrm{Spr}(B)$. Por lo tanto

$$\begin{aligned}\mathrm{Spr}(A-B) \quad &= \lambda(A-B) + \lambda(B-A)\\ &\overset{\text{Lidskii}}{\prec} \lambda(A) - \lambda^{\uparrow}(B) + \lambda(B) - \lambda^{\uparrow}(A) = \mathrm{Spr}(A) - \mathrm{Spr}^{\uparrow}(B)\ ,\end{aligned}$$

utilizando nuevamente el ítem 1. del Lema 1.1.6.

□

3.2. Una desigualdad clave para la dispersión espectral

Como mencionamos anteriormente en el inicio del capítulo, una de las motivaciones centrales de este capitulo es describir desigualdades para la dispersión espectral, en particular, el siguiente resultado constituye una mejora de la primera desigualdad de submayorización obtenida en la Proposición 2.3.4.

Teorema 3.2.1. *Sea* $A = \begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^r \end{matrix} \in \mathcal{H}(k+r)$. *Entonces*

$$2\,s(B) \prec_w \mathrm{Spr}^+(A)\ . \tag{3.13}$$

Demostración. Consideremos $U = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^r \end{matrix} \in \mathcal{U}(k+r)$. Entonces

$$s(UA - AU) = s(A - U^*AU) = |\lambda(A - U^*AU)|^{\downarrow}\,.$$

Por la desigualdad de Lidskii (ítem 1.) 1.1.4 tenemos que

$$\lambda(A - U^*AU) \prec \lambda(A) - \lambda^{\uparrow}(U^*AU) = \lambda(A) - \lambda^{\uparrow}(A)\,.$$

Combinando la ecuación anterior con la Observación 1.1.3 para la función convexa $f(x) = |x|$ tenemos que

$$s(UA - AU) = |\lambda(A - U^*AU)|^{\downarrow} \prec_w |\lambda(A) - \lambda^{\uparrow}(A)|^{\downarrow} = \left(\operatorname{Spr}^+(A)\,,\, \operatorname{Spr}^+(A)\right)^{\downarrow},$$

(o $\left(\operatorname{Spr}^+(A)\,,\, \operatorname{Spr}^+(A)\,,\, 0\right)^{\downarrow}$ si $k+r$ es impar). Usando la Proposición 1.1.10 y notando que

$$UA - AU = \begin{bmatrix} 0 & 2\,B \\ -2\,B^* & 0 \end{bmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^r \end{matrix} \implies s(UA - AU) = 2\left(s(B)\,,\, s(B)\,,\, 0_{|k-r|}\right)^{\downarrow},$$

donde $s(B)$ es un vector de tamaño $\min\{r\,,\,k\} \leq [\frac{k+r}{2}]$ (usando que $\operatorname{rk}(B) \leq \min\{r\,,\,k\}$).

Luego concluímos que

$$2\left(s(B)\,,\, s(B)\,,\, 0_{|k-r|}\right)^{\downarrow} \prec_w \left(\operatorname{Spr}^+(A)\,,\, \operatorname{Spr}^+(A)\right)^{\downarrow} \implies 2\,s(B) \prec_w \operatorname{Spr}^+(A)\,. \tag{3.14}$$

□

Observación 3.2.2. Notemos que el teorema anterior permite entre otras cosas mejorar los resultados de la Sección 2.3, removiendo un factor 2 que aparece en los enunciados de los Teoremas 2.3.5 y 2.3.7 y de sus corolarios. Además, este teorema será muy importante en el Capítulo 4, pues permitirá obtener estimaciones para la variación absoluta de los valores de Ritz de una matriz $A \in \mathcal{H}(d)$ (ver Teorema 4.1.2 y su prueba en 4.2).

Por otro lado, es importante notar que la desigualdad de la ecuación (3.13) es ajustada. En efecto, consideremos $(k = r)$

$$\text{si} \quad A = \begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^k \end{matrix} \implies \lambda(A) = (s(B)\,,\, -s(B))^{\downarrow} \quad \text{(ver Proposición 1.1.10)} \quad .$$

Luego en este caso tenemos la igualdad $2\,s(B) = \operatorname{Spr}^+(A)$. △

Observación 3.2.3. Sea $A = \begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^r \end{matrix} \in \mathcal{M}_{k+r}(\mathbb{C})^+$. Y. Tao prueba en [51] que

$$2s_j(B) \leq \lambda_j(A) = s_j(A) \quad \text{para todo} \quad j \in \mathbb{I}_{k+r}\,. \tag{3.15}$$

Notemos que en el caso positivo, vale que $\operatorname{Spr}^+(A) \leq \lambda(A) = s(A)$. Sin embargo, la desigualdad

$$2s_j(B) \leq \operatorname{Spr}^+_j(A) \quad \text{para todo} \quad j \in \mathbb{I}_{[\frac{k+r}{2}]}\,, \tag{3.16}$$

no es válida para matrices autoadjuntas; incluso falla en el caso en que A es semidefinida positiva. Por ejemplo, tomemos

$$A = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 1 & 0 & 1 & 3 \end{pmatrix} \in \mathcal{M}_{2+2}(\mathbb{C})^+\,.$$

Entonces $s(B) = (1, 1)$ y $\lambda(A) = (4{,}61\,,\ 2{,}61\,,\ 2{,}38\,,\ 0{,}39)$. Por lo tanto, la ecuación (3.16) falla para $j = 2$. En particular, el Teorema 3.2.1 no implica la desigualdad de Tao. Así mismo, la desigualdad de Tao tampoco implica el Teorema 3.2.1, como lo sugiere la ecuación (3.11) porque $A - \lambda_{[\frac{k+r}{2}]}(A)\, I \notin \mathcal{M}_{k+r}(\mathbb{C})^+$.

Por otro lado, si $A \in \mathcal{M}_{k+r}(\mathbb{C})^+$ tanto el resultado de Tao, como el antes expuesto son aplicables. En este caso, si N es una NUI y $\mu = (\mathrm{Spr}^+(A),\, 0) \in \mathbb{R}^{k+r}$, entonces

$$2\, N \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix} \leq N(D_\mu) \leq N(A)\ , \tag{3.17}$$

donde D_μ es la matriz diagonal con diagonal principal μ. En efecto, como $\mu \leq \lambda(A) = s(A)$, entonces

$$2\,(s(B),\, 0) \overset{(3.2.1)}{\prec_w} (\mathrm{Spr}^+(A),\, 0) \prec_w s(A)\,,$$

y la ecuación (3.17) es consecuencia de estas relaciones. Señalamos además que en el caso (genérico) $A \in \mathcal{G}l(k+r)^+$ obtenemos una desigualdad estricta $N(D_\mu) < N(A)$, Para una NUI estrictamente convexa N. Finalmente, notemos que el Teorema 3.2.1 es aplicable en un caso general, cuando A es autoadjunta (y no sólo semidefinida positiva). △

Corolario 3.2.4. *Sea* $A = \begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^r \end{matrix} \in \mathcal{H}(k+r)$. *Entonces*

$$\mathrm{Spr}^+ \begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix} \prec_w \mathrm{Spr}^+(A) \qquad y \qquad \mathrm{Spr}^+ \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \prec_w \mathrm{Spr}^+(A)\ . \tag{3.18}$$

Demostración. Por la Proposición 1.1.10 tenemos que $\mathrm{Spr}^+ \begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix} = 2\, s(B)$, donde hemos removido algunos ceros finales para igualar las dimensiones, en vistas de la Observación 1.1.2. Aplicando la ecuación (3.13), obtenemos la primera desigualdad (3.18). La segunda se obtiene de las ecuaciones (3.8) y (1.2). □

En lo que sigue desarrollaremos dos aplicaciones distintas para la desigualdad (3.13), una relacionada con la desigualdad de Zhan [54] y otra relacionada con rotaciones directas entre subespacios [17, 47] .

Observación 3.2.5. En [54] Zhan prueba que dados $A_1,\ A_2 \in \mathcal{M}_d(\mathbb{C})^+$ entonces

$$s_i(A_1 - A_2) \leq s_i(A_1 \oplus A_2) \quad \text{para todo} \quad i \in \mathbb{I}_d\,. \tag{3.19}$$

Estas desigualdades fundamentales son a su vez equivalentes a otros resultados centrales del análisis matricial (ver [6, 51, 55, 56]). Por otro lado, estas desigualdades entrada a entrada entre valores singulares, permiten obtener desigualdades de operadores de la forma $A_1 - A_2 \leq V^*(A_1 \oplus A_2)V$, para contracciones apropiadas $V \in \mathcal{M}_{2d,d}(\mathbb{C})$. Más aún, la ecuación (3.19) implica la relación de submayorización (más débil) $s(A_1 - A_2) \prec_w s(A_1 \oplus A_2)$, (ver ecuación (1.1)).

En caso que $C,\ D \in \mathcal{H}(d)$ son arbitrarios, entonces la desigualdad anterior no es válida. Sin embargo, pequeñas modificaciones en los argumentos de [54] implican las desigualdades

$$s_i(C - D) \ \leq\ 2\ s_i(C \oplus D) \quad \text{para todo} \quad i \in \mathbb{I}_d\ . \tag{3.20}$$

Las desigualdades anteriores resultan ajustadas, incluso para matrices autoadjuntas $C,\ D \in \mathcal{H}(d)$ (tomando $C \in \mathcal{M}_d(\mathbb{C})^+$ y $D = -C$). Como antes, estas desigualdades entrada a entrada entre valores singulares, permiten obtener desigualdades de submayorización de operadores relacionadas.

Finalmente, si estamos interesados en obtener desigualdades para normas unitariamente invariantes, es posible obtener mejores cotas superiores (para el caso positivo) mediante extensiones de las ecuaciones (3.19) y (3.20) para matrices autoadjuntas, como mostramos en el siguiente resultado. △

Teorema 3.2.6. *Sean A_1, $A_2 \in \mathcal{H}(d)$ matrices autoadjuntas. Entonces*

$$s(A_1 - A_2) \prec_w \mathrm{Spr}^+ (A_1 \oplus A_2)\,. \tag{3.21}$$

Demostración. Tomemos $A_1\,,\, A_2 \in \mathcal{H}(d)$. Sean

$$Z = \frac{1}{\sqrt{2}} \begin{bmatrix} I & I \\ -I & I \end{bmatrix} \in \mathcal{U}(2d) \qquad \text{y} \qquad T = \frac{1}{2} \begin{bmatrix} A_1 + A_2 & A_1 - A_2 \\ A_1 - A_2 & A_1 + A_2 \end{bmatrix} \in \mathcal{H}(2d)\,.$$

Luego

$$ZTZ^* = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} = A_1 \oplus A_2\,;$$

y por lo tanto, por el Teorema 3.2.1 vemos que $s(A_1 - A_2) \prec_w \mathrm{Spr}^+(T) = \mathrm{Spr}^+(A_1 \oplus A_2)$.

□

Observación 3.2.7. Sean A_1, $A_2 \in \mathcal{M}_d(\mathbb{C})^+$. Como mencionamos anteriormente, en este caso tenemos que $\mathrm{Spr}_i^+(A_1 \oplus A_2) \leq s_i(A_1 \oplus A_2)$, para $i \in \mathbb{I}_d$. Usando este hecho y el Teorema 3.2.6 obtenemos que

$$N(A_1 - A_2) \leq N(D_{\mathrm{Spr}^+(A_1 \oplus A_2)}) \leq N(D_{(s_i(A_1 \oplus A_2))_{i \in \mathbb{I}_d}})\,,$$

para cualquier norma unitariamente invariante N. Estas desigualdades, constituyen una mejora (incluso en el caso positivo) con respecto a las desigualdades para NUIs que pueden obtenerse a partir de (3.20). Por otro lado, notemos que las desigualdades entrada a entrada

$$s_i(A_1 - A_2) \leq \mathrm{Spr}_i^+ (A_1 \oplus A_2) \qquad \text{para} \qquad i \in \mathbb{I}_d\,,$$

son falsas incluso en el caso semidefinido positivo. En efecto, tomemos $A_1 = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$ y $A_2 = 3\,I$. Entonces $s(A_1 - A_2) = (2\,,\, 2)$ pero

$$\lambda\,(A_1 \oplus A_2) = (5\,,\, 3\,,\, 3\,,\, 1) \implies \mathrm{Spr}^+\,(A_1 \oplus A_2) = (4\,,\, 0)\,.$$

Además, remarcamos que las cotas obtenidas en los Teoremas 3.2.1 y 3.2.6 corresponden a vectores invariantes por traslaciones de la forma $M \mapsto M + \lambda\, I$ de las matrices M involucradas, a diferencia de las cotas superiores basadas en valores singulares, como las de las ecuaciones (3.15) y (3.19). Como los vectores considerados en estos teoremas son invariantes bajo las traslaciones correspondientes, consideramos que las cotas superiores en términos de la dispersión espectral son muy adecuadas en este contexto.

Finalmente, conviene notar que la desigualdad (3.21) es ajustada; en efecto tomamos $A_1 \in \mathcal{M}_d(\mathbb{C})^+$ arbitrario y consideramos $A_2 = -A_1 \in \mathcal{H}(d)$. Entonces $s(A_1 - A_2) = 2\,\lambda(A_1) = \mathrm{Spr}^+(A_1 \oplus A_2)$. △

Observación 3.2.8. Dada una función en $C^2(\mathbb{R})$ $g : \mathbb{R} \to \mathbb{R}$, podemos aproximarla en $x_0 = 0$ mediante su polinomio de Taylor de primer orden como:

$$g(x) = g(0) + g'(0)\,x + R_\varepsilon(x)\,,$$

donde $R_\varepsilon(x) = g''(\varepsilon)\,\frac{x^2}{2}$ cumple que $\lim_{x \to 0} \frac{R_\varepsilon(x)}{x} = 0$, para cierto $\varepsilon \in [0, x]$.

Haremos uso de este hecho múltiples veces durante este capítulo, en el contexto de funciones en $C^2(\mathbb{R})$ aplicadas a vectores de $\mathbb{R}^d$, es decir, si $z \in \mathbb{R}^d$, $g(z) = (g(z_i))_{i \in \mathbb{I}_d}$; y también mediante el cálculo funcional continuo.

Fijado $m \in \mathbb{N}$, sea $x_m \in \mathbb{R}^k$ tal que $\lim_{m\to\infty} x_m = 0 \in \mathbb{R}^k$ (entrada a entrada). Luego, considerando $g : \mathbb{R} \to \mathbb{R}$ en $C^2(\mathbb{R})$ y utilizando la aproximación anterior, tenemos que

$$g(x_m) = g(0) + g'(0)\, x_m + O(m)\,,$$

donde $O(m) =: R_\varepsilon(x_m)$ es tal que $\lim_{m\to\infty} m\, O(m) = 0$.

Además si $Z_m \in \mathcal{H}(d)$ es tal que $\lim_{m\to\infty} Z_m = 0 \in \mathcal{M}_d(\mathbb{C})$, entonces

$$g(Z_m) = g(0) + g'(0)\, Z_m + O_1(m)\,,$$

donde $O_1(m) =: R_\varepsilon(Z_m)$ es tal que $\lim_{m\to\infty} m\, O_1(m) = 0$. △

Observación 3.2.9. El siguiente resultado está relacionado con las desigualdades de la ecuación (1.6) con la misma idea que los casos anteriores, obtener mejores cotas superiores (por la Proposición 3.1.2), con respecto a un orden mas débil (submayorización en lugar de desigualdades entrada a entrada). △

Teorema 3.2.10. *Sean $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{C}^d$ subespacios tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$ y sea $Z \in \mathcal{H}(d)$ tal que $e^{iZ}\mathcal{X} = \mathcal{Y}$. Entonces se verifica la siguiente relación de submayorización*

$$\Theta(\mathcal{X}\,,\,\mathcal{Y}) \prec_w \frac{1}{2}\,\mathrm{Spr}^+(Z)\ . \tag{3.22}$$

Demostración. Sea $X \in \mathcal{M}_{d,k}(\mathbb{C})$ una isometría tal que $R(X) = \mathcal{X}$. Consideramos la curva suave $Y(\cdot) : [0,1] \to \mathcal{M}_{d,k}(\mathbb{C})$ dada por $Y(t) = e^{i\,t\,Z}X$, para $t \in [0,1]$; además consideramos $\mathcal{Y}(t) = R(Y(t)) \subseteq \mathbb{C}^d$, para $t \in [0,1]$. Notemos que $Y(0) = X$, $\mathcal{Y}(0) = \mathcal{X}$ y $\mathcal{Y}(1) = \mathcal{Y}$.

Como cada $Y(t)$ es una isometría, la función $\Theta(\cdot) : [0,1] \to [0,\pi/2]^k$ dada por $\Theta(t) = \Theta(\mathcal{X}\,,\,\mathcal{Y}(t)) = \arccos(s^{\uparrow}(X^*Y(t)\,)\,)$ para $t \in [0,1]$, es continua y cumple que $\Theta(0) = 0$.

Dado $m \in \mathbb{N}$, consideramos los subespacios $\mathcal{Y}(\frac{j}{m}) = R(Y(\frac{j}{m}))$, para $j = 0, ..., m$. Recordemos que el Teorema 1.3.6 establece una versión de la desigualdad triangular en términos de submayorización para ángulos principales; aplicando este resultado tenemos que

$$\Theta(\mathcal{X}\,,\,\mathcal{Y}) \quad \prec_w \sum_{j=0}^{m-1} \Theta(\mathcal{Y}(\tfrac{j}{m})\,,\,\mathcal{Y}(\tfrac{j+1}{m}))\,. \tag{3.23}$$

Notemos que $Y(t+h) = e^{i\,t\,Z}\,Y(h)$ con $e^{i\,t\,Z} \in \mathcal{U}(d)$, para t, h, $t+h \in [0,1]$; en consecuencia, tenemos que $\Theta(\mathcal{Y}(\frac{j}{m})\,,\,\mathcal{Y}(\frac{j+1}{m})) = \Theta(\mathcal{Y}(0)\,,\,\mathcal{Y}(\frac{1}{m}))$, para cada $j \in \mathbb{I}_{m-1}$. En lo que sigue probaremos que,

$$\Theta(\mathcal{X}\,,\,\mathcal{Y}(\tfrac{1}{m})) \quad \prec_w \tfrac{1}{2m}\ \mathrm{Spr}^+(X) + O(m) \quad \text{con} \quad \lim_{m\to\infty} m\, O(m) = 0\,. \tag{3.24}$$

En efecto, como $\frac{d}{dt}\,\mathrm{sen}(t)|_{t=0} = 1$ y $\mathrm{sen}(0) = 0$ vemos que

$$\Theta(\mathcal{X},\,\mathcal{Y}(\tfrac{1}{m})) \quad = \mathrm{sen}(\Theta(\mathcal{X},\,\mathcal{Y}(\tfrac{1}{m}))) + O_1(m) \quad \text{con} \quad \lim_{m\to\infty} m\, O_1(m) = 0\,, \tag{3.25}$$

utilizando una aproximación de primer orden de tipo Taylor (ver Observación 3.2.8).

Sea $X_\perp \in \mathcal{M}_{d,d-k}(\mathbb{C})$ una isometría con rango $R(X_\perp) = \mathcal{X}^\perp$. Para todo $m \in \mathbb{N}$, como $Y(\frac{1}{m})$ es una isometría, la ecuación (1.3) (definición de ángulos principales) asegura que

$$\mathrm{sen}(\Theta(\mathcal{X}\,,\,\mathcal{Y}(\tfrac{1}{m}))) \quad = s(Y(\tfrac{1}{m})^*X_\perp) = s(X\,e^{\frac{i}{m}\,Z}\,X_\perp) \in ([0,1]^k)^{\downarrow}\,.$$

Más aún, como $X\, e^{i\,t\,Z}\, X_\perp|_{t=0} = 0$ y $\frac{d}{dt}(X\, e^{i\,t\,Z}\, X_\perp)|_{t=0} = i\, X\, Z\, X_\perp$, tenemos que

$$X\, e^{\frac{i}{m}\,Z}\, X_\perp = \frac{i}{m}\, X\, Z\, X_\perp + O_2(m) \quad \text{con} \quad \lim_{m\to\infty} m\, O_2(m) = 0\,. \tag{3.26}$$

Luego, combinando las ecuaciones (3.13), (3.25) y (3.26) junto con la desigualdad de Weyl (Teorema 1.1.5) obtenemos que

$$\Theta(\mathcal{X},\, \mathcal{Y}(\tfrac{1}{m})) \;\prec_w \tfrac{1}{2m} s\big(\, 2\, X\, Z\, X_\perp\,\big) + s\big(\, O_2(m)\,\big) + O_1(m)$$

$$\overset{(3.13)}{\prec_w}\; \tfrac{1}{2m}\, \mathrm{Spr}^+(Z) + O(m)\,,$$

que prueba que la ecuación (3.24) vale. En consecuencia, utilizando la ecuación (3.23) tenemos que

$$\Theta(\mathcal{X}\,,\, \mathcal{Y}) \prec_w \frac{1}{2}\, \mathrm{Spr}^+(Z) + m\, O(m)\,.$$

y el resultado se obtiene tomando $m \to \infty$.

□

Como consecuencia del Teorema 3.2.10 damos una versión mas fuerte del [47, Teorema 5.1].

Corolario 3.2.11. *Sean $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{C}^d$ subespacios tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$. Sea $R \in \mathcal{H}(d)$ tal que $e^{iR}\mathcal{X} = \mathcal{Y}$. Si $U = e^{i\,Z} \in \mathcal{U}(d)$ es una rotación directa de $\mathcal{X}$ en $\mathcal{Y}$, para $Z \in \mathcal{H}(d)$, entonces*

$$s(Z) \prec_w \frac{1}{2}|\mathrm{Spr}(R)| \prec_w s(R)\,.$$

Demostración. Si $U \in \mathcal{U}(d)$ es una rotación directa de $\mathcal{X}$ en $\mathcal{Y}$ entonces, hemos visto (ver Definición 1.3.3) que $U = e^{i\,Z}$ para $Z \in \mathcal{H}(d)$ tal que

$$\lambda(Z) = (\Theta(\mathcal{X}\,,\, \mathcal{Y})^*, -\,\Theta(\mathcal{X}\,,\, \mathcal{Y})^*, 0)\;, \quad \text{donde} \quad \Theta(\mathcal{X}\,,\, \mathcal{Y})^* = (\theta_i)_{i\in\mathbb{I}_r} \in \mathbb{R}^r\,,$$

denota el vector de ángulos principales positivos, para algún $r \leq [d/2]$. Luego, $s(Z) = (\Theta(\mathcal{X}\,,\, \mathcal{Y})^*, \Theta(\mathcal{X}\,,\, \mathcal{Y})^*, 0)^{\downarrow} \in (\mathbb{R}^d)^{\downarrow}$. Por otro lado, por el Teorema 3.2.10 tenemos que

$$s(Z) = (\Theta(\mathcal{X}\,,\, \mathcal{Y})^*, \Theta(\mathcal{X}\,,\, \mathcal{Y})^*, 0) \prec_w \frac{1}{2}(\mathrm{Spr}^+(R)\,,\, \mathrm{Spr}^+(R))^{\downarrow} = \frac{1}{2}|\mathrm{Spr}(R)|^{\downarrow}\,.$$

Por último por la Proposición 3.1.2 obtenemos que $\frac{1}{2}|\mathrm{Spr}(R)| \prec_w s(R)$ y el resultado es consecuencia de estas desigualdades de mayorización.

□

Observación 3.2.12. Sea $A = \begin{bmatrix} a_1 & b \\ \bar{b} & a_2 \end{bmatrix} \in \mathcal{H}(2)$. Entonces

$$\mathrm{Spr}^+(A) = \lambda_1(A) - \lambda_2(A) = \big[\,(\mathrm{tr}\; A)^2 - 4\,\det\, A\big]^{1/2} = \big(\,(a_1 - a_2)^2 + 4|b|^2\,\big)^{1/2} \in \mathbb{R}_{\geq 0}\;. \tag{3.27}$$

De forma general, si consideramos $A = \begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix} \begin{matrix} \mathbb{C}^k \\ \mathbb{C}^k \end{matrix} \in \mathcal{H}(2k)$, es natural preguntarse si la desigualdad

$$\lambda\left(\big[\,(A_1 - A_2)^2 + 4\, B^* B\big]^{1/2}\right) \prec_w \mathrm{Spr}^+(A)\,, \tag{3.28}$$

es válida. Notemos que esta desigualdad constituiría una mejora de los Teoremas 3.2.1 y 3.2.6. Sin embargo, la ecuación (3.28) no es cierta en general. Por ejemplo,

$$A = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & 1 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix} \begin{matrix} \mathbb{C}^2 \\ \\ \mathbb{C}^2 \end{matrix} \quad \text{cumple que} \quad \mathrm{Spr}^+(A) = (6{,}2714\,,\, 1{,}6339)\;,$$

y $\lambda\left(\left[\,(A_1-A_2)^2+4\,B^*B\right]^{1/2}\right)=(4{,}7599\,,\,3{,}3680)$. Pero

$$\operatorname{tr}\left(\operatorname{Spr}^{+}(A)\,\right)=7{,}9053<8{,}1279=\operatorname{tr}\left(\left[\,(A_1-A_2)^2+4\,B^*B\right]^{1/2}\right)\ .$$

En lo que sigue proponemos una versión débil de la ecuación (3.28) pero que vale en general (para matrices autoadjuntas). △

Proposición 3.2.13. *Sea* $A=\begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix}\begin{matrix}\mathbb{C}^k\\ \mathbb{C}^k\end{matrix}\in\mathcal{H}(2k)$. *Entonces*

$$\lambda\big(\,(A_1-A_2)^2+4\operatorname{Re}(B)^2\,\big)\prec_w\operatorname{Spr}^2(A)\ . \tag{3.29}$$

Demostración. Consideramos la matriz unitaria $U=\begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}\in\mathcal{U}(2k)$, y conjugamos A, obteniendo

$$U^*A\,U=\begin{bmatrix} A_2 & -B^* \\ -B & A_1 \end{bmatrix}.$$

Además construímos

$$D\overset{\text{def}}{=}A-U^*AU=\begin{bmatrix} A_1-A_2 & 2\operatorname{Re}(B) \\ 2\operatorname{Re}(B) & A_2-A_1 \end{bmatrix}.$$

Por la desigualdad de Weyl (ítem 1. del Teorema 1.1.4),

$$\lambda(D)\prec\lambda(A)+\lambda(-A)=(\operatorname{Spr}(A)\,,\,-\operatorname{Spr}^{\uparrow}(A)\,)\in(\mathbb{R}^{2k}_{\geq 0})^{\downarrow}\ .$$

Notemos que

$$D^2=\begin{bmatrix} (A_1-A_2)^2+4\operatorname{Re}(B)^2 & * \\ * & (A_1-A_2)^2+4\operatorname{Re}(B)^2 \end{bmatrix}.$$

Si $E\overset{\text{def}}{=}(A_1-A_2)^2+4\mathrm{Re}(B)^2$, entonces utilizando la Observación 1.1.3 y el Teorema 1.1.4, tenemos que

$$(\lambda(E)\,,\,\lambda(E)\,)^{\downarrow}\overset{(1.2)}{\prec}\lambda(D^2)\prec_w(\operatorname{Spr}^{+}(A)^2,\operatorname{Spr}^{+}(A)^2)\,, \tag{3.30}$$

dado que $x\mapsto x^2$ es una función convexa. Luego la ecuación (3.29) es consecuencia de la ecuación (3.30).

□

Es conveniente notar que la Proposición 3.2.13 no implica el Teorema 3.2.1 ni el Teorema 3.2.6 (puesto que la relación $\prec_w$ no se preserva tomando raíz cuadrada).

3.3. Reformulaciones de la desigualdad clave

En lo que sigue veremos una serie de resultados que son consecuencia de la desigualdad clave dada en el Teorema 3.2.1. Luego probaremos que esta desigualdad resulta equivalente a varias de las desigualdades que deducimos aqui (ver Teorema 3.5.1 debajo).

3.4. Conmutadores generalizados y conjugaciones unitarias

Sean $A, Z \in \mathcal{H}(d)$, consideremos la curva suave (de unitarios) $U(t) = e^{i\,t\,Z}$ para $t \in [0,1]$. Los conmutadores de la forma $i\,(AZ - ZA) \in \mathcal{H}(d)$ aparecen naturalmente al trabajar con derivaciones de $U^*(t)\,A\,U(t) = e^{-i\,t\,Z}\,A\,e^{i\,t\,Z}$ en $t = 0$. En efecto

$$(U^*(t)\,A\,U(t))'\,|_{t=0} = -e^{-i\,t\,Z}\,i\,Z\,A\,U(t) + U^*(t)\,A\,e^{i\,t\,Z} i\,Z\,|_{t=0} = i\,(AZ - ZA)\,.$$

A su vez, estimaciones para $s(AZ - ZA)$ en términos de relaciones de submayorización, permiten obtener cotas para los valores singulares de conmutadores generalizados de la forma $A\,W - W\,B$, donde $A, B \in \mathcal{H}(d)$ y $W \in \mathcal{M}_d(\mathbb{C})$.

Ahora bien, en el Capítulo 4 estaremos interesados en estimar $s(X^*A\,X - Y^*A\,Y)$ donde $X, Y \in \mathcal{M}_{d,k}(\mathbb{C})$ son isometrías tales que $R(X) = \mathcal{X}$ y $R(Y) = \mathcal{Y}$. En particular, si $\mathcal{X}$ es un subespacio A-invariante, las técnicas que permiten estimar dicho vector requieren cotas para $s(A\,W - W\,B)$ (ver Sección 4.3).

Por este motivo, comenzamos esta sección con la siguiente desigualdad para valores singulares de conmutadores generalizados en términos de la dispersión espectral.

Teorema 3.4.1. *Sean $A, X \in \mathcal{H}(d)$. Entonces*

$$\lambda\left(i\,(A\,X - X\,A)\right) \prec_w \frac{1}{2}\,\mathrm{Spr}^+(A)\;\mathrm{Spr}^+(X)\,. \tag{3.31}$$

Demostración. Si $k \in \mathbb{I}_d$, entonces existe una proyección $P \in \mathcal{H}(d)$ con $\mathrm{rk}\,P = k$ tal que

$$\sum_{j=1}^{k} \lambda_j\left(i\,(A\,X - X\,A)\right) = \mathrm{tr}\,\left(i\,(A\,X - X\,A)\,P\right).$$

Más aún,

$$\mathrm{tr}\,\left(i\,(AX - XA)\,P\right) = \mathrm{tr}\,\left(i\,(XP - PX)\,A\right) \le \mathrm{tr}\,\left(\lambda(i\,(XP - PX))\lambda(A)\right),$$

pues $\mathrm{tr}(ZW) \le \mathrm{tr}(\lambda(Z)\,\lambda(W))$, para $Z, W \in \mathcal{H}(d)$. En efecto, para $Z, W \in \mathcal{M}_d(\mathbb{C})^+$ el resultado es consecuencia de la desigualdad de Weyl 1.1.5 (ítem 2.).

Si $Z, W \in \mathcal{H}(d)$, entonces consideramos $Z - \beta\,I \in \mathcal{M}_d(\mathbb{C})^+$ y $W - \mu\,I \in \mathcal{M}_d(\mathbb{C})^+$ para ciertos $\beta, \mu \in \mathbb{R}$. Luego

$$\begin{aligned}\mathrm{tr}((Z - \beta\,I)\,(W - \mu\,I)) &= \mathrm{tr}(Z\,W) - \mathrm{tr}(\beta\,W) - \mathrm{tr}(\mu\,Z) + \mathrm{tr}(\beta\,\mu\,I)\\ &\le \mathrm{tr}(\lambda(Z - \beta\,I)\,\lambda(W - \mu\,I))\\ &= \mathrm{tr}(\lambda(Z)\,\lambda(W)) - \mathrm{tr}(\lambda(\mu\,Z)) - \mathrm{tr}(\lambda(\beta\,W)) + \mathrm{tr}(\lambda(\mu\,\beta))\,,\end{aligned}$$

de lo que se deduce la desigualdad.

Ahora bien, consideremos la descomposición matricial por bloques inducida por P:

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^* & X_{22} \end{bmatrix} \quad \text{y} \quad P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \implies i\,(XP - PX) = i\begin{bmatrix} 0 & -X_{12} \\ X_{12}^* & 0 \end{bmatrix}.$$

Notemos $B = X_{12} \in \mathcal{M}_{k,d-k}(\mathbb{C})$; luego por la Proposición 1.1.10

$$\lambda(i\,(XP - PX)) = (s(B)\,,\,-s(B))^{\downarrow}\,.$$

Por otro lado el Teorema 3.2.1 implica que $s(B) \prec_w \frac{1}{2}\,\mathrm{Spr}^+(X)$. Este hecho junto con el ítem 4. del Lema 1.1.6 muestran que

$$s(B)\,\mathrm{Spr}^+(A) \prec_w \frac{1}{2}\,\mathrm{Spr}^+(X)\,\mathrm{Spr}^+(A)\,. \tag{3.32}$$

Más aún, sea $k' \leq k \leq d$ el número de valores singulares no nulos de B, entonces

$$\begin{aligned}
\operatorname{tr}\left(\lambda(i(XP-PX))\,\lambda(A)\right) &= \sum_{j=1}^{k'} s_j(B)\,\lambda_j(A) - \sum_{j=1}^{k'} s_j(B)\,\lambda_j^{\uparrow}(A) \\
&= \sum_{j=1}^{k'} s_j(B)\operatorname{Spr}_j(A) \\
&\leq \sum_{j=1}^{k'} s_j(B)\operatorname{Spr}_j^{+}(A) \\
&\overset{(3.32)}{\leq} \frac{1}{2}\sum_{j=1}^{k'} \operatorname{Spr}_j^{+}(X)\operatorname{Spr}_j^{+}(A)\,,
\end{aligned}$$

considerando que, $\operatorname{Spr}_j(A) \leq \operatorname{Spr}_j^{+}(A)$ para $j \in \mathbb{I}_d$ (completando con ceros a $\operatorname{Spr}^{+}(A)$). Combinando los argumentos previos obtenemos que

$$\sum_{j=1}^{k} \lambda_j\left(i\,(AX-XA)\right) \leq \frac{1}{2}\sum_{j=1}^{k'} \operatorname{Spr}_j^{+}(X)\operatorname{Spr}_j^{+}(A) \leq \frac{1}{2}\sum_{j=1}^{k} \operatorname{Spr}_j^{+}(X)\operatorname{Spr}_j^{+}(A)\,,$$

para cada $k \in \mathbb{I}_d$ (ver Observación 1.1.2), lo que prueba la ecuación (3.31). □

Aunque el Teorema 3.4.1 contiene mucha información, su enunciado es bastante técnico. El siguiente resultado, que es una consecuencia del Teorema 3.4.1, es más claro y tiene varias implicaciones directas (ver Corolario 3.4.4 debajo).

Teorema 3.4.2. *Sean $A, X \in \mathcal{H}(d)$. Entonces*

$$s(AX-XA) \prec_w \frac{1}{2}\operatorname{Spr}^{+}(A\oplus A)\operatorname{Spr}^{+}(X\oplus X)\,. \tag{3.33}$$

Demostración. Aplicando la ecuación (3.31) sobre $A, -X \in \mathcal{H}(d)$, tenemos que

$$\lambda(-i\,(AX-XA)) \prec_w \frac{1}{2}\operatorname{Spr}^{+}(A)\operatorname{Spr}^{+}(-X) = \frac{1}{2}\operatorname{Spr}^{+}(A)\operatorname{Spr}^{+}(X)\,,$$

donde hemos usado que $\operatorname{Spr}^{+}(-X) = \operatorname{Spr}^{+}(X)$.

Ahora bien, consideramos $0 \leq r \leq d$ tal que $(\lambda_j(i\,(AX-XA)))_{j\in\mathbb{I}_r}$ denota los autovalores no negativos de $i\,(AX-XA)$ (ordenados en forma no creciente). De esta forma

$$\begin{aligned}
s(AX-XA) &= \left((\lambda_j(i\,(AX-XA)))_{j\in\mathbb{I}_r}\,,\,(\lambda_j(-i\,(AX-XA))_{j\in\mathbb{I}_{d-r}}\right)^{\downarrow} \\
&\prec_w \frac{1}{2}\left(\operatorname{Spr}^{+}(A)\operatorname{Spr}^{+}(X)\,,\,\operatorname{Spr}^{+}(A)\operatorname{Spr}^{+}(X)\right)^{\downarrow} \\
&= \frac{1}{2}\operatorname{Spr}^{+}(A\oplus A)\operatorname{Spr}^{+}(X\oplus X)\,,
\end{aligned}$$

que prueba la ecuación (3.33). □

Observación 3.4.3. Notemos que el teorema anterior propone una simetría en la estimación del conmutador $AX - XA$. Más aún, en caso que X o A estén cerca de un múltiplo de la identidad, uno esperaría que $s(AX - XA)$ sea de alguna forma pequeño y este es el caso; fijado $\varepsilon > 0$, supongamos que $X = \mu\, I + X_0$ donde $X_0 \in \mathcal{H}(d)$ es tal que $||X|| < \varepsilon$ y $\mu \in \mathbb{R}$. Entonces

$$s(AX - XA) = s(AX_0 - X_0A) \prec_w \varepsilon\, \mathrm{Spr}^+(A \oplus A)\,,$$

donde hemos estimado $\mathrm{Spr}^+(X_0 \oplus X_0)$ por $2\,\varepsilon\, \mathbb{1}_d$, dado que $\lambda(X_0) \subset [-\varepsilon, \varepsilon]$.

Por otro lado, el teorema anterior permite deducir la siguiente desigualdad de submayorización (ver (3.34) debajo) para conmutadores generalizados de la forma $AZ - ZB$, para $A, B \in \mathcal{H}(d)$ y $Z \in \mathcal{M}_d(\mathbb{C})$. Estimaciones para este tipo de conmutadores (en términos de sus valores singulares) son necesarias para deducir resultados sobre la variación absoluta de los valores de Ritz de una matriz A en el caso que $\mathcal{X}$ es un subespacio A-invariante (ver el Teorema 4.1.4 y su prueba en 4.3).

Corolario 3.4.4. *Sean $A_1\,,\ A_2 \in \mathcal{H}(d)$ y $X \in \mathcal{M}_d(\mathbb{C})$. Entonces*

$$s(A_1\, X - X A_2) \prec_w s(X)\ \mathrm{Spr}^+(A_1 \oplus A_2)\,. \tag{3.34}$$

Demostración. Consideremos las matrices $C = A_1X - XA_2 \in \mathcal{M}_d(\mathbb{C})$,

$$B = A_1 \oplus A_2 = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \in \mathcal{H}(2d) \qquad \text{y} \qquad Z = \begin{bmatrix} 0 & X \\ -X^* & 0 \end{bmatrix} \in i\,\mathcal{H}(2d)\ .$$

Notemos que

$$B\ Z - Z\ B = \begin{bmatrix} 0 & A_1X - XA_2 \\ X^*A_1 - A_2X^* & 0 \end{bmatrix} = \begin{bmatrix} 0 & C \\ C^* & 0 \end{bmatrix} \in \mathcal{H}(2d)\,,$$

Luego

$$\begin{aligned} s(i\,(B\,Z - Z\,B)) &= s(B\,(i\,Z) - (i\,Z)\,B) \\ &\overset{3.4.2}{\prec_w} \frac{1}{2}\mathrm{Spr}^+(i\,Z \oplus i\,Z)\ \mathrm{Spr}^+(B \oplus B) \\ &\overset{3.1.2}{\prec_w} s(Z)\,\mathrm{Spr}^+(B \oplus B)\,. \end{aligned}$$

Ahora bien, notemos que

$$s(B\,Z - Z\,B) = (s(A_1\,X - X\,A_2), s(A_1\,X - X\,A_2))^{\downarrow}\,,$$

$$s(Z) = (s(X), s(X))^{\downarrow} \qquad \text{y} \qquad \mathrm{Spr}^+(B \oplus B) = (\mathrm{Spr}^+(A_1 \oplus A_2), \mathrm{Spr}^+(A_1 \oplus A_2))^{\downarrow}\,.$$

Combinando los hechos anteriores se obtiene la ecuación (3.34). □

Observación 3.4.5. La estimación del corolario anterior puede compararse con la obtenida por Kittaneh en [27] para operadores compactos y positivos en un espacio de Hilbert separable (estimaciones que en particular son válidas para matrices) que enunciamos a continuación: Sean $A, B \in \mathcal{M}_d(\mathbb{C})^+$ y $X \in \mathcal{M}_d(\mathbb{C})$. Entonces

$$s_i(A\,X - X\,B) \leq ||X||\, s_i(A \oplus B)\,, \qquad \text{para} \qquad i \in \mathbb{I}_d\,. \tag{3.35}$$

En este sentido, el Corolario 3.4.4 es válido para el caso autoadjunto, donde la ecuación (3.35) no vale en general. Además propone como cota superior a la dispersión espectral de $A \oplus B$ que en el caso positivo es más chica que $s(A \oplus B)$ (en términos de submayorización).

Observación 3.4.6. Analizando la prueba del Teorema 3.4.2, vemos que el resultado es consecuencia del Teorema 3.2.1. Por otro lado, el Corolario 3.4.4 extiende al Teorema 3.2.6 (tomando $X = I$ en la ecuación (3.34)); en particular la desigualdad (3.34) es ajustada (ver la Observación 3.2.7). △

Corolario 3.4.7. *Sean A, $X \in \mathcal{H}(d)$. Entonces*

$$s(AX - XA) \prec_w \operatorname{Spr}^+(X)\, \operatorname{Spr}^+(A \oplus A)\,. \tag{3.36}$$

Si $X \in \mathcal{M}_d(\mathbb{C})^+$, entonces

$$s(AX - XA) \prec_w \frac{1}{2}\,\|X\|\, \operatorname{Spr}^+(A \oplus A)\,. \tag{3.37}$$

Demostración. Para la primera desigualdad, como $X \in \mathcal{H}(d)$,

$$\frac{1}{2}\operatorname{Spr}^+(X \oplus X) \prec_w \operatorname{Spr}^+(X)\,,$$

(ver ecuación (3.11)). Luego el resultado es consecuencia del Teorema 3.4.2 y el Lema 1.1.6. Para el caso $X \in \mathcal{M}_d(\mathbb{C})^+$, consideramos $Y = X - \lambda I \in \mathcal{H}(d)$ tomando $\lambda = ||X||/2$. Luego $AY - YA = AX - XA$ y por el Corolario 3.4.4

$$s(AX - XA) \prec_w s(X - \lambda I)\, \operatorname{Spr}^+(A \oplus A)\,.$$

Ahora bien, como $s(Y) \leqslant \|Y\|\mathbb{1} = \frac{\|X\|}{2}\mathbb{1}$ tenemos que $s(Y) \prec_w \frac{\|X\|}{2}\mathbb{1}$. Por lo tanto, combinando estos hechos con la ecuación anterior y aplicando el Lema 1.1.6 obtenemos la ecuación (3.37). □

Corolario 3.4.8. *Sean A, $B \in \mathcal{M}_d(\mathbb{C})^+$ y $X \in \mathcal{M}_d(\mathbb{C})$. Entonces*

$$s(AX - XB) \prec_w s(X)\; s(A \oplus B)\,. \tag{3.38}$$

En particular para cualquier norma unitariamente invariante N,

$$N(AX - XB) \leq \|X\|\, N(A \oplus B)\,. \tag{3.39}$$

Demostración. La prueba se deduce del Corolario 3.4.4 y recordando que para $C \in \mathcal{M}_{2d}(\mathbb{C})^+$, $\operatorname{Spr}(C) \leqslant \lambda(C) = s(C)$. □

Sean A, $X \in \mathcal{H}(d)$ y sea $U := e^{i\,X} \in \mathcal{U}(d)$. Como fue mencionado en el ítem 7. de la Observación 3.1.1 (o como consecuencia de la ecuación (3.21) del Teorema 3.2.6) tenemos que

$$s(A - UAU^*) \prec_w \operatorname{Spr}^+(A \oplus UAU^*) = \operatorname{Spr}^+(A \oplus A) = |\operatorname{Spr}(A)|\,.$$

Sin embargo, en el caso en que X está cerca de un múltiplo (real) de la identidad, entonces $U = e^{i\,X}$ también está cerca de la identidad; luego, en este caso uno esperaría que A y U^*AU también estén cerca. Similarmente, en caso de que A esté cerca de un múltiplo (real) de la identidad, uno esperaría que A y U^*AU estén cerca. El siguiente resultado provee una estimación cuantitativa que aborda estas situaciones.

Teorema 3.4.9. Sean A, $X \in \mathcal{H}(d)$ y $U = e^{i\,X} \in \mathcal{U}(d)$. Entonces

$$s(A - U^*AU) \prec_w \frac{1}{2}\operatorname{Spr}^+(X \oplus X)\, \operatorname{Spr}^+(A \oplus A)\,.$$

Demostración. Sea $U(\cdot) : [0,1] \to \mathcal{H}(d)$ una función suave dada por $A(t) = e^{-i\,t\,X}\,A\,e^{i\,t\,X}$, para $t \in [0,1]$. Notemos que $A(0) = A$ y $A(1) = U^*AU$; usando la desigualdad de Weyl para valores singulares (ítem 1. del Teorema 1.1.5)

$$s(A - U^*AU) \prec_w \sum_{j=0}^{m-1} s\big(\, A(\tfrac{j}{m}) - A(\tfrac{j+1}{m})\,\big) \quad \text{para todo} \quad m \in \mathbb{N}\,. \tag{3.40}$$

Además $A(t+h) = e^{-i\,t\,X}\,A(h)\,e^{i\,t\,X}$ con $e^{i\,t\,X} \in \mathcal{U}(d)$, para t, h, $t+h \in [0,1]$. Esto permite asegurar que

$$s(A(\tfrac{j}{m}) - A(\tfrac{j+1}{m})) \;\; = s(A - A(\tfrac{1}{m})) \quad \text{para} \quad j \in \mathbb{I}_{m-1}\,. \tag{3.41}$$

Por otro lado, como $A - A(0) = 0$ y $\frac{d}{dt}A(t)|_{t=0} = i\,(AX - XA)$ tenemos que (ver Observación 3.2.8)

$$s(A - A(\tfrac{1}{m})) \;\; = \tfrac{1}{m}\,s(AX - XA) + O(m) \quad \text{con} \quad \lim_{m\to\infty} m\,O(m) = 0\,. \tag{3.42}$$

Luego por el Teorema 3.4.2 tenemos que

$$s(A - A(\tfrac{1}{m})) \;\; \prec_w \tfrac{1}{m}\,\tfrac{1}{2}\,\mathrm{Spr}^+(X \oplus X) \;\; \mathrm{Spr}^+(A \oplus A) + O(m)\,. \tag{3.43}$$

Por lo tanto, combinando las ecuaciones (3.40), (3.41) y (3.43) tenemos que, para un m suficientemente grande,

$$s(A - U^*AU) \prec_w \frac{1}{2}\,\mathrm{Spr}^+(X \oplus X) \;\; \mathrm{Spr}^+(A \oplus A) + m\,O(m)\,.$$

El resultado se obtiene tomando $m \to \infty$ en la expresión anterior.

□

Observación 3.4.10. Inspeccionando la prueba del Teorema 3.4.9, observamos que este resultado es consecuencia del Teorema 3.4.2 mediante su aplicación en la ecuación (3.43).

3.5. Equivalencias entre desigualdades

Y. Tao prueba en [51] que la desigualdad (3.15) es equivalente a la desigualdad de Zhan (3.19). Más aún, podemos ver que la desigualdad de Kittaneh (3.35) generaliza a la desigualdad de Zhan. Basados en estas ideas, veremos la equivalencia entre varias de las desigualdades principales obtenidas en las secciones 3.2 y 3.4. En particular ya hemos probado las desigualdades involucradas en el teorema siguiente (3.5.1); sin embargo cada una de las desigualdades es de interés en si misma y se enfocan en contextos distintos.

Teorema 3.5.1. *Las siguientes desigualdades son equivalentes:*

1. *Dados A_1, $A_2 \in \mathcal{H}(d)$ entonces*

$$s(A_1 - A_2) \prec_w \mathrm{Spr}^+\,(A_1 \oplus A_2)\,. \tag{3.44}$$

2. *Dado $A = \begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix} \begin{matrix} \mathbb{C}^d \\ \mathbb{C}^d \end{matrix} \in \mathcal{H}(2d)$ entonces*

$$2\,s(B) \prec_w \mathrm{Spr}^+(A)\,. \tag{3.45}$$

3. *Dados $A, X \in \mathcal{H}(d)$, entonces*

$$s(A\,X - X\,A) \prec_w \frac{1}{2}\,\mathrm{Spr}^+(X \oplus X) \;\; \mathrm{Spr}^+(A \oplus A)\,. \tag{3.46}$$

4. Dados A, $X \in \mathcal{H}(d)$ y $U = e^{i\,X} \in \mathcal{U}(d)$ entonces

$$s(A - U^*AU) \prec_w \frac{1}{2}\,\mathrm{Spr}^+(X \oplus X)\;\mathrm{Spr}^+(A \oplus A)\,. \tag{3.47}$$

5. Dados $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{C}^d$ y $Z \in \mathcal{H}(d)$ tales que $e^{iZ}\mathcal{X} = \mathcal{Y}$, entonces

$$\Theta(\mathcal{X}\,,\,\mathcal{Y}) \prec_w \frac{1}{2}\,\mathrm{Spr}^+(Z)\;. \tag{3.48}$$

Demostración. 1. $\implies$ 2.: Sea $B = U|B|$ la descomposición polar de B, donde $U \in \mathcal{U}(d)$. Si

$$W = \begin{bmatrix} U & 0 \\ 0 & I \end{bmatrix} \in \mathcal{U}(2d) \quad \text{entonces} \quad W^*AW = \begin{bmatrix} U^*A_1U & U^*B \\ B^*U & A_2 \end{bmatrix} = \begin{bmatrix} U^*A_1U & |B| \\ |B| & A_2 \end{bmatrix}\;.$$

Ahora bien, como $\mathrm{Spr}(A) = \mathrm{Spr}(W^*AW)$ y $s(B) = s(|B|)$, para probar la ecuación (3.45) podemos asumir que $B \in \mathcal{H}(d)$. En este caso, tomando $R = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \in \mathcal{U}(2d) \cap \mathcal{H}(2d)$, vemos que

$$RAR = \begin{bmatrix} A_2 & B \\ B & A_1 \end{bmatrix} \implies \frac{A + RAR}{2} = \begin{bmatrix} \frac{A_1+A_2}{2} & B \\ B & \frac{A_1+A_2}{2} \end{bmatrix}\;.$$

Por la desigualdad de Weyl (ítem 1. del Teorema 1.1.4),

$$\lambda\left(\frac{A + RAR}{2}\right) \prec \frac{\lambda(A) + \lambda(RAR)}{2} = \lambda(A) \overset{(3.8)}{\implies} \mathrm{Spr}^+(\frac{A + RAR}{2}) \prec_w \mathrm{Spr}^+(A)\,.$$

En consecuencia, para probar la ecuación (3.45) podemos asumir que $B \in \mathcal{H}(d)$ y que $A_1 = A_2$.

Además tomemos $Z = \frac{1}{\sqrt{2}} \begin{bmatrix} I & I \\ -I & I \end{bmatrix} \in \mathcal{U}(2d)$. En este caso tenemos que $A = \begin{bmatrix} A_1 & B \\ B & A_1 \end{bmatrix}$ y por lo tanto

$$Z^*AZ = \begin{bmatrix} A_1 - B & 0 \\ 0 & A_1 + B \end{bmatrix} \qquad \text{y} \qquad \mathrm{Spr}^+(Z^*AZ) = \mathrm{Spr}^+(A)\;. \tag{3.49}$$

Luego $2\,s(B) = s\big(\,[A_1 - B] - [A_1 + B]\,\big) \overset{(3.44)}{\prec_w} \mathrm{Spr}^+\big(\,[A_1 - B] \oplus [A_1 + B]\,\big) \overset{(3.49)}{=} \mathrm{Spr}^+(A)$.

Por otro lado ya hemos probado que (ver Observaciones 3.4.6 y 3.4.10) 2. $\implies$ 3. $\implies$ 4.

4. $\implies$ 3. Sean $A, X \in \mathcal{H}(d)$. Consideramos $D(t) = A - e^{-i\,t\,X}\,A\,e^{i\,t\,X}$, para $t \in [0, 1]$. Entonces, $D(\cdot)$ es una función suave tal que $D(0) = 0$ y $D'(0) = -i\,(AX - XA)$. Luego por la desigualdad de Weyl para valores singulares tenemos que (y nuevamente un argumento de tipo Taylor)

$$s(t\,(AX - XA)) = s(A - e^{-i\,t\,X}\,A\,e^{i\,t\,X}) + O(t) \quad \text{con} \quad \lim_{t \to 0^+} \frac{O(t)}{t} = 0\,.$$

Luego por la ecuación (3.47) vemos que

$$s(AX - XA) \prec_w \frac{1}{2}\,\mathrm{Spr}^+(X \oplus X)\;\mathrm{Spr}^+(A \oplus A) + \frac{O(t)}{t}\,.$$

Luego 3. se obtiene tomando el límite $t \to 0^+$. Como 3. implica el Corolario 3.4.4, que implica 1. (tomando $X = I$) vemos que 1. - 4. son equivalentes. Además, ya hemos probado(en la prueba del Teorema 3.2.10) que 2. $\implies$ 5. Por lo tanto, bastaría probar que 5. $\implies$ 2.

En efecto, fijemos $A \in \mathcal{H}(2d)$ como en 2. Sea $\mathcal{X} = \mathbb{C}^d \oplus 0 \subset \mathbb{C}^d \oplus \mathbb{C}^d$ y sea $\mathcal{Y}(t) = e^{i\,t\,A}\mathcal{X}$, para $t \in [0,1]$. Sean

$$X = \begin{bmatrix} I_d \\ 0 \end{bmatrix}\,,\; X_\perp = \begin{bmatrix} 0 \\ I_d \end{bmatrix} \in \mathcal{M}_{2d,d}(\mathbb{C}) \qquad \text{y} \qquad Y(t) = e^{i\,t\,A}X \qquad \text{para} \qquad t \in [0,1]\,.$$

Luego, por la ecuación (1.3), tenemos que

$$\operatorname{sen}(\Theta(\mathcal{X}\,,\,\mathcal{Y}(t))) = s(Y(t)^*\,X_\perp) = s(X\,e^{i\,t\,A}\,X_\perp)\,.$$

Como $Y^*(0)\,X_\perp = 0$ y $\frac{d}{dt}(Y^*(t)\,X_\perp)|_{t=0} = i\,X\,A\,X_\perp = i\,B \in \mathcal{M}_d(\mathbb{C})$ obtenemos que (ver Observación 3.2.8)

$$t\,s(B) = \operatorname{sen}(\Theta(\mathcal{X}\,,\,\mathcal{Y}(t))) + O_1(t) \qquad \text{con} \qquad \lim_{t\to 0^+} \frac{O_1(t)}{t} = 0\,.$$

Por otro lado, como $\Theta(\mathcal{X}\,,\,\mathcal{Y}(t))$ es una función continua de $t \in [0,1]$ tal que $\Theta(\mathcal{X}\,,\,\mathcal{Y}(0)) = 0$, tenemos que

$$\operatorname{sen}(\Theta(\mathcal{X}\,,\,\mathcal{Y}(t))) = \Theta(\mathcal{X}\,,\,\mathcal{Y}(t)) + O_2(t) \qquad \text{con} \qquad \lim_{t\to 0^+} \frac{O_2(t)}{t} = 0\,.$$

Los hechos anteriores junto con la ecuación (3.48) de 5. prueban que

$$t\,s(B) \prec_w \frac{t}{2}\,\operatorname{Spr}^+(A) + O_1(t) + O_2(t) \implies 2\,s(B) \prec_w \operatorname{Spr}^+(A) + \frac{2}{t}\,(O_1(t) + O_2(t))\,.$$

Por lo tanto 2. se obtiene de la desigualdad anterior tomando el límite $t \to 0^+$. □

Observación 3.5.2. Cabe destacar que el ítem 2. del Teorema 3.5.1 implica el Teorema 3.2.1 en su forma general (cuando el bloque B puede ser rectangular). Este hecho es consecuencia de la ecuación (3.9) en la Observación 3.1.1. △

Capítulo 4

Submatrices principales, ángulos y dispersión espectral

El objetivo de esta sección es describir una serie de desigualdades para la variación absoluta de los valores de Ritz, muy relacionadas con la [30, Conjetura 2.1]. En este desarrollo, algunas de las desigualdades de mayorización previamente obtenidas en la Sección 3.3 serán cruciales.

En la Sección 4.1 establecemos los resultados sobre el proceso de extracción de submatrices $(A, X, \mathcal{Y}) \mapsto Y_r^* A Y_r$, donde $Y_r = Y_r(X, U)$ está definido por $Y_r = U\, X$ para una rotación directa U de $\mathcal{X}$ en $\mathcal{Y}$. En este caso, acotaremos el vector de valores singulares $s(X^*AX - Y_r^*AY_r)$ en términos de los ángulos principales $\Theta(\mathcal{X}, \mathcal{Y})$ y la dispersión espectral $\mathrm{Spr}^+(A)$. En la Sección 4.4 aplicaremos los resultados previamente desarrollados para obtener cotas superiores para la variación absoluta de los valores de Ritz.

4.1. Variación de submatrices principales

Observación 4.1.1. Comenzaremos recordando algunas notaciones y hechos que utilizaremos para desarrollar los argumentos principales. Siguiendo la Notación 2.0.1, $\mathcal{X}, \mathcal{Y} \subset \mathbb{C}^d$, denotan subespacios k-dimensionales de $\mathbb{C}^d$. Además, $\mathcal{I}_{\mathcal{X}}(k, d)$ corresponde al conjunto de isometrías $X \in \mathcal{M}_{d,\,k}(\mathbb{C})$ con rango $R(X) = \mathcal{X}$.

Por otro lado, recordamos que los *ángulos principales* entre $\mathcal{X}$ e $\mathcal{Y}$, notados $\Theta(\mathcal{X}, \mathcal{Y}) = (\theta_j)_{j \in \mathbb{I}_k} \in ([0, \pi/2]^k)^{\downarrow}$, están dados por

$$\cos(\Theta) = (\cos(\theta_j))_{j \in \mathbb{I}_k} = (s_{k-j+1}(X^*Y))_{j \in \mathbb{I}_k}\ , \tag{4.1}$$

donde $X \in \mathcal{I}_{\mathcal{X}}(k, d)$ e $Y \in \mathcal{I}_{\mathcal{Y}}(k, d)$ (ver Definición 1.3.2).

Por último, recordamos que la descomposición de Halmos permite describir $\mathbb{C}^d$ en términos de los subespacios $\mathcal{X}$ e $\mathcal{Y}$. Asumiendo que (ver la prueba del Teorema 4.1.2 debajo)

$$\mathcal{X}^{\perp} \cap \mathcal{Y}^{\perp} = \{0\}\,,$$

recordamos que podemos considerar los subespacios de $\mathbb{C}^d$:

$$\mathcal{S}_1 = (\mathcal{X} \cap \mathcal{Y}^{\perp}) \oplus (\mathcal{X} \cap \mathcal{G})\ ,\ \ \mathcal{S}_2 = (\mathcal{X}^{\perp} \cap \mathcal{Y}) \oplus (\mathcal{X}^{\perp} \cap \mathcal{G}) \quad \text{y} \quad \mathcal{S}_3 = (\mathcal{X}^{\perp} \cap \mathcal{Y}) \oplus (\mathcal{Y} \cap \mathcal{G})\ .$$

De esta forma,

$$\mathbb{C}^d = (\mathcal{X} \cap \mathcal{Y}) \oplus \mathcal{S}_1 \oplus \mathcal{S}_2\,. \tag{4.2}$$

Sea U una rotación directa de $\mathcal{X}$ en $\mathcal{Y}$, consideremos $p = \dim(\mathcal{X} \cap \mathcal{Y}^{\perp})$ y $r = \dim(\mathcal{X} \cap \mathcal{G})$. En este caso, existen bases ortonormales para $\mathcal{S}_1$ y $\mathcal{S}_2$, y matrices diagonales semidefinidas positivas

$$C = \mathrm{diag}(\cos(\Theta')) \quad \text{y} \quad S = \mathrm{diag}(\mathrm{sen}(\Theta')) \in \mathcal{M}_{p+r}(\mathbb{C})^+\,,$$

donde $\Theta' = (\frac{\pi}{2}\,\mathbb{1}_p\,,\,\theta_1,\ldots,\theta_r)$ denotan los ángulos principales entre $\mathcal{S}_1$ y $\mathcal{S}_3$, tales que

$$U = \begin{pmatrix} I & 0 & 0 \\ 0 & C & -S \\ 0 & S & C \end{pmatrix}. \tag{4.3}$$

Esta descomposición matricial se obtiene a partir de la BON construida por yuxtaposición de bases ortonormales para $\mathcal{X}\cap\mathcal{Y}$, $\mathcal{S}_1$ y $\mathcal{S}_2$. Así $(\theta_1,\ldots,\theta_r)\in(0,\pi/2)^r$ denota los ángulos principales entre $\mathcal{X}\cap\mathcal{G}$ e $\mathcal{Y}\cap\mathcal{G}$. Más aún, en este caso si $s=\dim\mathcal{X}\cap\mathcal{Y}$, entonces

$$\Theta' = \Big(\,\frac{\pi}{2}\,\mathbb{1}_p\,,\,\theta_1\,,\,\ldots\,,\,\theta_r\,\Big) \qquad \text{y} \qquad \Theta(\mathcal{X}\,,\,\mathcal{Y}) = (\Theta',0_s)\,.$$

△

A continuación enunciamos los Teoremas 4.1.2 y 4.1.4, cuyas pruebas daremos posteriormente en las secciones 4.2 y 4.3.

Teorema 4.1.2. *Sean A, $B\in\mathcal{H}(d)$ y sean $\mathcal{X}$, $\mathcal{Y}\subset\mathbb{C}^d$ subespacios tales que $\dim(\mathcal{X})=\dim(\mathcal{Y})=k$. Sean $U=U(\mathcal{X},\mathcal{Y})$ una rotación directa de $\mathcal{X}$ en $\mathcal{Y}$ y $\Theta=\Theta(\mathcal{X},\mathcal{Y})\in([0,\pi/2]^k)^{\downarrow}$ los ángulos principales entre $\mathcal{X}$ e $\mathcal{Y}$ definidos en (4.1). Dado $X\in\mathcal{I}_{\mathcal{X}}(k,d)$, si consideramos*

$$Y_r = Y_r(X,U) \overset{def}{=} UX\in\mathcal{I}_{\mathcal{Y}}(k,d)\,,$$

entonces tenemos que

$$s(X^*\,A\,X - Y_r^*\,B\,Y_r) \prec_w s(A-B) + \Theta(\mathcal{X},\mathcal{Y})\,\left(\frac{\mathrm{Spr}^+(A)+\mathrm{Spr}^+(B)}{2}\right). \tag{4.4}$$

□

Observación 4.1.3. El teorema anterior corresponde a una situación general de dos matrices autoadjuntas A, $B\in\mathcal{H}(d)$, que permite estimar los valores singulares de $X^*\,A\,X - Y_r^*\,B\,Y_r$ en términos de dos cantidades; por un lado $s(A-B)$ detecta la cercanía de A y B y por otro $\Theta(\mathcal{X},\mathcal{Y})$ y $\frac{1}{2}\,(\mathrm{Spr}^+(A)+\mathrm{Spr}^+(B))$ proveen un control mediante la distancia entre los subespacios $\mathcal{X}$ e $\mathcal{Y}$, en términos de los ángulos principales y la dispersión espectral de A y B respectivamente. Notemos que en el caso que A y B están cerca de múltiplos de la identidad, el término $s(A-B)$ es el que tiene mayor peso en la estimación, al igual que cuando los subespacios $\mathcal{X}$ e $\mathcal{Y}$ están cerca.

△

El siguiente resultado corresponde al caso invariante (cuando $\mathcal{X}$ es A-invariante), que permite obtener cotas superiores más fuertes para pequeñas perturbaciones $\mathcal{Y}$ de $\mathcal{X}$.

Teorema 4.1.4. *Bajo las hipótesis del Teorema 4.1.2, y asumiendo además que el subespacio k-dimensional $\mathcal{X}$ es A-invariante, tenemos que*

$$s(X^*\,A\,X - Y_r^*\,A\,Y_r) \prec_w \Theta^2(\mathcal{X},\mathcal{Y})\,\mathrm{Spr}^+(A)\,. \tag{4.5}$$

□

Observación 4.1.5. Notemos que el teorema anterior permite mejorar la estimación para el caso $A = B$ del Teorema 4.1.2, admitiendo una cota de orden cuadrático con respecto a los ángulos $\Theta(\mathcal{X}, \mathcal{Y})$. Esta mejora, que se debe a la condición adicional de que $\mathcal{X}$ es A-invariante, se obtiene a través del refinamiento de las técnicas utilizadas para probar 4.1.2.

Por otro lado, este caso permitirá obtener cotas superiores para la variación absoluta de los valores de Ritz de A a través de la desigualdad de Lidskii

$$|\lambda(X^* A\, X) - \lambda(Y_r^* A\, Y_r)| \prec_w s(X^* A\, X - Y_r^* A\, Y_r)\,.$$

△

4.2. Prueba del Teorema 4.1.2

En lo que sigue describiremos las técnicas y los resultados necesarios para probar los Teoremas 4.1.2 y 4.1.4. Para esto recordamos la siguiente notación:

1. A, $B \in \mathcal{H}(d)$.
2. $\mathcal{X}$, $\mathcal{Y} \subset \mathbb{C}^d$, son subespacios tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$ y $X \in \mathcal{I}_{\mathcal{X}}(k, d)$.
3. $U = U(\mathcal{X}, \mathcal{Y}) \in \mathcal{U}(d)$ denota una rotación directa de $\mathcal{X}$ en $\mathcal{Y}$.
4. $Y_r = Y_r(X, U) \stackrel{\text{def}}{=} UX \in \mathcal{I}_{\mathcal{Y}}(k, d)$.
5. $\Theta(\mathcal{X}\,, \mathcal{Y}) \in ([0, \pi/2]^k)^{\downarrow}$ denota los ángulos principales entre $\mathcal{X}$ e $\mathcal{Y}$.

El enfoque para probar el Teorema 4.1.2 se basa en el siguiente argumento: vamos a considerar las curvas suaves

$$L(\cdot) : [0, 1] \to \mathcal{H}(d) \quad \text{e} \quad Y_r(\cdot) : [0, 1] \to \mathcal{I}(k, d)\,, \tag{4.6}$$

tales que $L(0) = A$, $L(1) = B$, $Y_r(0) = X$ e $Y_r(1) = Y_r$. Además, consideramos la curva suave

$$\gamma : [0, 1] \to \mathcal{H}(k) \quad \text{dada por} \quad \gamma(t) = Y_r(t)^*\, L(t)\, Y_r(t) \quad \text{para} \quad t \in [0, 1]\,. \tag{4.7}$$

Una vez construída $\gamma(\cdot)$ aplicaremos el siguiente resultado.

Proposición 4.2.1. *Sea $\gamma : [0, 1] \to \mathcal{M}_{m,\ell}(\mathbb{C})$ una curva suave tal que $\gamma(0) = C$ y $\gamma(1) = D$. Entonces*

$$s(D - C) \prec_w \int_0^1 s(\gamma'(t))\; dt\,.$$

Demostración. Notemos que por el teorema fundamental del cálculo tenemos que

$$D - C = \gamma(1) - \gamma(0) = \int_0^1 \gamma'(t)\; dt\,. \tag{4.8}$$

Por otro lado, para $n \in \mathbb{N}$ consideramos la partición regular $\{t_0 = 0 < t_1 < \ldots < t_n = 1\}$ de $[0, 1]$ de forma que $t_j = \frac{j}{n}$, para $j \in \{0\} \cup \mathbb{I}_n$ y $\Delta_j = t_j - t_{j-1} = \frac{1}{n} = \Delta_n$, para $j \in \mathbb{I}_n$. Luego,

$$\int_0^1 \gamma'(t)\; dt = \lim_{n\to\infty} \sum_{j=1}^{n} \gamma'(t_j)\; \Delta_n \quad \text{y} \quad \int_0^1 s(\gamma'(t))\; dt = \lim_{n\to\infty} \sum_{j=1}^{n} s(\gamma'(t_j))\; \Delta_n\,, \tag{4.9}$$

donde utilizamos que $\gamma'(t) \in \mathcal{M}_{m,\ell}(\mathbb{C})$ y $s(\gamma'(t)) \in \mathbb{R}^m$, para $t \in [0,1]$, son continuas (por hipótesis y por la continuidad de los valores singulares). Luego por la desigualdad de Weyl aditiva (ver Teorema 1.1.5) tenemos que

$$s(\sum_{j=1}^{n} \gamma'(t_j)\ \Delta_n) \prec_w \sum_{j=1}^{n} s(\gamma'(t_j))\ \Delta_n\,. \tag{4.10}$$

Finalmente el resultado se obtiene combinando las ecuaciones (4.8), (4.9) y (4.10), junto con la continuidad de los valores singulares y el siguiente hecho: si $(u_n)_{n\in\mathbb{N}}$ y $(v_n)_{n\in\mathbb{N}} \in \mathbb{R}^m$ son sucesiones que convergen a u y $v \in \mathbb{R}^m$ respectivamente y tales que $u_n \prec_w v_n$ para cada $n \in \mathbb{N}$, entonces $u \prec_w v$. □

Si bien los argumentos considerados anteriormente son válidos para cualquier elección de curvas suaves $L(\cdot)$ e $Y_r(\cdot)$ como en la ecuación (4.6), estamos particularmente interesados en elecciones que permitan obtener mejores cotas superiores. Por lo tanto, buscamos aquellas curvas $L(\cdot)$ e $Y_r(\cdot)$ para las cuales la curva asociada $\gamma(\cdot)$ de la ecuación (4.7) es *minimal* en cierto sentido.

Es importante señalar que no vamos a estudiar el problema de minimalidad asociado, simplemente elegiremos $L(\cdot)$ e $Y_r(\cdot)$ de forma que tengan propiedades de minimalidad por separado y usaremos estas elecciones para construir $\gamma(\cdot)$.

Para $L(\cdot)$ hay una elección natural; el segmento que une A y B, es decir,

$$L(t) = (1-t)\,A + t\,B \quad \text{para} \quad t \in [0,1]\,. \tag{4.11}$$

A continuación construiremos $Y_r(t)$, basados en los conceptos previamente descriptos en la descomposición de Halmos. En función del argumento anterior, definimos

$$U(t) = \begin{pmatrix} I & 0 & 0 \\ 0 & C(t) & -S(t) \\ 0 & S(t) & C(t) \end{pmatrix} \in \mathcal{U}(d)\,, \tag{4.12}$$

donde $S(\cdot)$, $C(\cdot) :\ [0\,,\,1] \to \mathcal{H}(p+r)$ estan dados por

$$S(t) = \operatorname{diag}(\operatorname{sen}(t\,\Theta')) \quad \text{y} \quad C(t) = \operatorname{diag}(\cos(t\,\Theta')) \in \mathcal{M}_{p+r}(\mathbb{C})^+\,. \tag{4.13}$$

Así $U(0) = I$ y $U(1) = U$. A partir de la construcción explícita de $U(t)$ definimos

$$Y_r(t) = U(t)\,X = \begin{pmatrix} I & 0 \\ 0 & C(t) \\ 0 & S(t) \end{pmatrix} \quad \text{con} \quad X = \begin{pmatrix} I & 0 \\ 0 & I \\ 0 & 0 \end{pmatrix}\,, \tag{4.14}$$

donde la representación matricial por bloques es con respecto a la descomposición $\mathbb{C}^d = (\mathcal{X} \cap \mathcal{Y}) \oplus \mathcal{S}_1 \oplus \mathcal{S}_2$, como antes.

En este caso, resulta que

$$\gamma(t) = Y_r(t)^*\,L(t)\,Y_r(t) = X^*\,(\,U(t)^*\,L(t)\,U(t)\,)\,X \quad \text{para} \quad t \in [0,1]\,. \tag{4.15}$$

Notemos que, derivando la ecuación anterior obtenemos

$$\gamma'(t) = X^*\,U(t)^*\,L'(t)\,U(t)\,X + X^*\,(\,(U'(t))^*\,L(t)\,U(t) + U(t)^*\,L(t)\,U'(t)\,)\,X\,. \tag{4.16}$$

El primer término de la ecuación (4.16) puede ser estimado utilizando la desigualdad de Lidskii. Esto nos deja con el análisis del segundo término. Usando identidades trigonométricas obtenemos que $U(t+h) = U(t)\,U(h)$ para t, h, $t+h \in [0,1]$ y también $U(t)^* = U(-t)$, para $t \in [0,1]$; por lo tanto $U'(t) = U(t)\,U'(0)$ y de forma similar $(U'(t))^* = -U'(0)\,U(t)^*$. En consecuencia, considerando

$$A(t) = U(t)^*\,L(t)\,U(t)\,, \tag{4.17}$$

resulta que

$$X^* \left((U'(t))^* L(t) U(t) + U(t)^* L(t) U'(t) \right) X = X^* \left(A(t) U'(0) - U'(0) A(t) \right) X \,. \tag{4.18}$$

Derivando la ecuación (4.12) obtenemos que

$$U'(0) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -D_{\Theta'} \\ 0 & D_{\Theta'} & 0 \end{pmatrix} \in i \cdot \mathcal{H}(d) \,,$$

donde $D_{\Theta'}$ es la matriz diagonal con diagonal $\Theta' = (\frac{\pi}{2} \mathbb{1}_p \,, \theta_1, \ldots, \theta_r)$ definida anteriormente y $i \cdot \mathcal{H}(d)$ denota al espacio de matrices antihermitianas.

Consideremos la representación matricial de $A(t) = (A_{ij}(t))_{i,j=1}^3$ con respecto a la descomposición de la ecuación (4.2). Un cálculo directo, utilizando la descomposición matricial por bloques anterior, muestra que

$$X^* \left(A(t) U'(0) - U'(0) A(t) \right) X = 2 \, \mathrm{Re} \left(\begin{pmatrix} 0 & A_{13}(t) \\ 0 & A_{23}(t) \end{pmatrix} \begin{pmatrix} 0 \\ D_{\Theta'} \end{pmatrix} \right) . \tag{4.19}$$

Observación 4.2.2. Manteniendo la notación anterior y utilizando la desigualdad multiplicativa de Weyl (Teorema 1.1.5) tenemos que

$$s(X^* \left(A(t) U'(0) - U'(0) A(t) \right) X) \prec_w 2 \, \Theta \, s \left(\begin{pmatrix} A_{13}(t) \\ A_{23}(t) \end{pmatrix} \right) , \tag{4.20}$$

donde hemos usado que $\Theta = (\Theta', 0_s) \in (\mathbb{R}^k_{\geq 0})^{\downarrow}$ y la Observación 1.1.2. Ahora bien, podemos considerar la siguiente representación por bloques para $A(t)$

$$A(t) = \left(\begin{array}{cc|c} A_{11}(t) & A_{12}(t) & A_{13}(t) \\ A_{21}(t) & A_{22}(t) & A_{23}(t) \\ \hline A_{31}(t) & A_{32}(t) & A_{33}(t) \end{array} \right) , \tag{4.21}$$

con respecto a la descomposición

$$\mathbb{C}^d = \left[\, (\mathcal{X} \cap \mathcal{Y}) \oplus \mathcal{S}_1 \, \right] \oplus \mathcal{S}_2 \,.$$

Luego la matriz $\begin{pmatrix} A_{13}(t) \\ A_{23}(t) \end{pmatrix}$ puede verse como un bloque anti-diagonal de $A(t)$. Por lo tanto, por el Teorema 3.2.1 aplicado a $A(t)$ y el Lema 1.1.6, tenemos que

$$s(X^* \left(A(t) U'(0) - U'(0) A(t) \right) X) \prec_w 2 \, \Theta \, s \left(\begin{pmatrix} A_{13}(t) \\ A_{23}(t) \end{pmatrix} \right) \prec_w \Theta \, \mathrm{Spr}^+(A(t)) \,. \tag{4.22}$$

△

Ahora estamos en condiciones de probar el Teorema 4.1.2 que enunciamos nuevamente a continuación.

Teorema 4.1.2. Sean A, $B \in \mathcal{H}(d)$ y sean $\mathcal{X}$, $\mathcal{Y} \subset \mathbb{C}^d$ subespacios tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$. Sean $U = U(\mathcal{X}, \mathcal{Y})$ una rotación directa de $\mathcal{X}$ en $\mathcal{Y}$ y $\Theta = \Theta(\mathcal{X}, \mathcal{Y}) \in ([0, \pi/2]^k)^{\downarrow}$ los ángulos principales entre $\mathcal{X}$ e $\mathcal{Y}$ definidos en (4.1). Dado $X \in \mathcal{I}_{\mathcal{X}}(k, d)$, si consideramos

$$Y_r = Y_r(X, U) \stackrel{\text{def}}{=} UX \in \mathcal{I}_{\mathcal{Y}}(k, d) \,,$$

entonces tenemos que

$$s(X^* \, A \, X - Y_r^* \, B \, Y_r) \prec_w s(A - B) + \Theta(\mathcal{X}, \mathcal{Y}) \left(\frac{\mathrm{Spr}^+(A) + \mathrm{Spr}^+(B)}{2} \right) .$$

Demostración. Supongamos primero que $\mathcal{X}^\perp \cap \mathcal{Y}^\perp = \{0\}$. Sea $\gamma(\cdot) : [0,1] \to \mathcal{H}(k)$ definida como en la ecuación (4.15), donde $Y_r(\cdot)$ está definido como en la ecuación (4.14). Notemos que por construcción γ es una curva suave tal que $\gamma(0) = X^* A X$ y $\gamma(1) = Y_r^* B Y_r$. Por la Proposición 4.2.1 tenemos que

$$s(X^* A X - Y_r^* B Y_r) \prec_w \int_0^1 s(\gamma'(t))\, dt\,. \tag{4.23}$$

Usando la ecuación (4.16) y la desigualdad de Weyl (ver Teorema 1.1.5), tenemos que

$$\begin{aligned} s(\gamma'(t)) \quad &\prec_w s(X^* U(t)^* L'(t) U(t) X) \\ &+ s(X^* ((U'(t))^* L(t) U(t) + U(t)^* L(t) U'(t)) X)\,, \end{aligned} \tag{4.24}$$

donde $L(t) = (1-t) A + t B$, para $t \in [0,1]$ como antes. Luego $L'(t) = B - A$ y por lo tanto

$$s(X^* U(t)^* L'(t) U(t) X) = s(X^* U(t)^* (B - A) U(t) X) \prec_w s(A - B)\,,$$

donde hemos usado la Observación 1.1.7. Los hechos anteriores junto con el ítem 2. del Lema 1.1.6 muestran que

$$\int_0^1 s(X^* U(t)^* L'(t) U(t) X)\, dt \prec_w s(A - B)\,. \tag{4.25}$$

Notemos que $A(t)$ y $L(t)$ son conjugados unitarios, por lo tanto tienen la misma dispersión espectral. Luego, por la Proposición 3.1.3 tenemos que

$$\mathrm{Spr}^+(A(t)) = \mathrm{Spr}^+(L(t)) \prec (1-t)\,\mathrm{Spr}^+(A) + t\,\mathrm{Spr}^+(B)\,. \tag{4.26}$$

Por las ecuaciones (4.22) y (4.26) y usando el ítem 2. del Lema 1.1.6,

$$\begin{aligned} &\int_0^1 s(X^* ((U'(t))^* L(t) U(t) + U(t)^* L(t) U'(t)) X)\, dt \prec_w \\ &\Theta\left(\mathrm{Spr}^+(A) \int_0^1 (1-t)\, dt + \mathrm{Spr}^+(B) \int_0^1 t\, dt\right) = \frac{1}{2}\,\Theta\,(\,\mathrm{Spr}^+(A) + \mathrm{Spr}^+(B)\,)\,. \end{aligned}$$

Usando esta última relación de submayorización, junto con las ecuaciones (4.24) y (4.25) obtenemos

$$\int_0^1 s(\gamma'(t))\, dt \prec_w s_j(A-B)_{j=1}^k + \Theta\left(\frac{\mathrm{Spr}^+(A) + \mathrm{Spr}^+(B)}{2}\right). \tag{4.27}$$

Luego el resultado se obtiene combinando las ecuaciones(4.23) y (4.27).

En caso que $\mathcal{X}^\perp \cap \mathcal{Y}^\perp \neq \{0\}$, consideramos el subespacio $\mathcal{Z} = \mathcal{X} + \mathcal{Y} \subset \mathbb{C}^d$ y $Z \in \mathcal{I}_{\mathcal{Z}}(m,d)$, donde $m = \dim(\mathcal{Z})$.

Sean $A' = Z^* A Z$, $B' = Z^* B Z \in \mathcal{H}(m)$ y sean $X' = Z^* X$ e $Y' = Z^* Y_r \in \mathcal{I}(k,m)$. Si $\mathcal{X}' = R(X')$ e $\mathcal{Y}' = R(Y')$ denotan los rangos de X' e Y', tenemos que $(\mathcal{X}')^\perp \cap (\mathcal{Y}')^\perp = \{0\}$. Además, $(X')^* Y' = (X^* Z)(Z^* Y_r) = X^* P_{\mathcal{Z}} Y_r = X^* Y_r$ (donde $P_{\mathcal{Z}} \in \mathcal{H}(d)$ denota la proyección ortogonal sobre $\mathcal{Z} \subset \mathbb{C}^d$). Luego, concluímos que $\Theta(\mathcal{X}', \mathcal{Y}') = \Theta(\mathcal{X}, \mathcal{Y}) \in (\mathbb{R}^k)^\downarrow$, por definición de ángulos principales.

Este argumento permite aplicar la primera parte de la prueba a las matrices $A', B' \in \mathcal{H}(m)$, la isometría X' y el subespacio $\mathcal{Y}'$. En este caso, consideramos una isometría $Y_r' = U' X'$, donde $U' \in \mathcal{U}(m)$ es una rotación directa de $\mathcal{X}'$ en $\mathcal{Y}'$. Ahora bien, notemos que $Z^* U Z$ es una rotación directa de $\mathcal{X}'$ en $\mathcal{Y}'$ de forma que podemos tomar $Y_r' = Z^* U Z Z^* X = Z^* Y_r = Y'$. Luego, obtenemos la siguiente desigualdad

$$s((X')^* A' X' - (Y')^* B' Y') \prec_w s_j(A' - B')_{j=1}^k + \Theta\left(\frac{\mathrm{Spr}^+(A') + \mathrm{Spr}^+(B')}{2}\right). \tag{4.28}$$

Más aún, notemos que

$$(X')^*\,A'\,X' = X^*\,P_{\mathcal{Z}}\,A\,P_{\mathcal{Z}}\,X = X^*\,A\,X \qquad \text{y}$$
$$(Y')^*\,B'\,Y' = Y_r^*\,P_{\mathcal{Z}}\,B\,P_{\mathcal{Z}}\,Y_r = Y_r^*\,B\,Y_r\,.$$

Por lo tanto,

$$s((X')^*\,A'\,X' - (Y')^*\,B'\,Y') = s(X^*\,A\,X - Y_r^*\,B\,Y_r)\,. \tag{4.29}$$

Por otro lado, usando la Observación 1.1.7 (o las desigualdades de entrelace) vemos que

$$s(A'-B') = s(Z^*\,(A-B)\,Z) \prec_w s(A-B) \qquad \text{y}$$
$$\mathrm{Spr}^+(A') \prec_w \mathrm{Spr}^+(A) \;\;,\;\; \mathrm{Spr}^+(B') \prec_w \mathrm{Spr}^+(B)\,.$$

Finalmente, el resultado es consecuencia de las desigualdades de submayorización anteriores, combinadas con las ecuaciones (4.28) y (4.29). □

4.3. Prueba del Teorema 4.1.4

Notación 4.3.1. Repasamos la notación a utilizar para la prueba del Teorema 4.1.4. Consideramos:

1. $A \in \mathcal{H}(d)$.
2. $\mathcal{X}$, $\mathcal{Y} \subset \mathbb{C}^d$, son subespacios tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$ y $\mathcal{X}$ es A-invariante.
3. $U = U(\mathcal{X},\mathcal{Y}) \in \mathcal{U}(d)$, es una rotación directa de $\mathcal{X}$ en $\mathcal{Y}$.
4. $Y_r = Y_r(X,U) \overset{\text{def}}{=} UX \in \mathcal{I}_{\mathcal{Y}}(k,d)$. △

La técnica para probar este resultado es similar a la utilizada en la sección anterior. Primero probaremos un caso particular del Teorema 4.1.4 cuando $k = \dim(\mathcal{X}) = \dim(\mathcal{Y}) \leq d/2$ (ver Proposiciones 4.3.2 y 4.3.3 debajo). Luego, reduciremos el caso general a este caso particular previamente desarrollado. En lo que sigue consideramos la descomposición

$$\mathbb{C}^d = \mathcal{X} \oplus \mathcal{X}^\perp \quad \text{asumiendo que} \quad \mathcal{X} \cap \mathcal{Y} = \{0\} \quad \text{y} \quad \mathcal{X}^\perp \cap \mathcal{Y}^\perp = \{0\}\,.$$

Notemos que en este caso

$$A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} \quad \text{es decir} \quad A_{12} = A_{21}^* = 0\,. \tag{4.30}$$

Dada una rotación directa U de $\mathcal{X}$ en $\mathcal{Y}$, consideramos la descomposición por bloques

$$U(t) = \begin{pmatrix} C(t) & -S(t) \\ S(t) & C(t) \end{pmatrix}\,,$$

donde $S(\cdot)$, $C(\cdot) :\; [0\,,\,1] \to \mathcal{H}(p+r)$ están dados por

$$S(t) = \mathrm{diag}(\,\mathrm{sen}(t\,\Theta)\,) \qquad \text{y} \qquad C(t) = \mathrm{diag}(\,\cos(t\,\Theta)\,) \in \mathcal{M}_{p+r}(\mathbb{C})^+\,.$$

En este contexto $\Theta = \Theta(\mathcal{X},\mathcal{Y}) = (\frac{\pi}{2}\,\mathbb{1}_p\,,\,\theta_1,\ldots,\theta_r) \in ([0,\pi/2]^k)^\downarrow$ donde, como antes,

$$p = \dim(\mathcal{X}\cap\mathcal{Y}^\perp) = \dim(\mathcal{X}^\perp\cap\mathcal{Y}) \qquad \text{y} \qquad r = \dim(\mathcal{X} \ominus (\mathcal{X}\cap\mathcal{Y}^\perp)\,) = k - p\,.$$

En este caso, definimos

$$Y_r(t) = U(t)\,X = \begin{pmatrix} C(t) \\ S(t) \end{pmatrix} \quad \text{con} \quad X = \begin{pmatrix} I \\ 0 \end{pmatrix}\,, \tag{4.31}$$

donde consideramos la representación por bloques con respecto a la descomposición $\mathbb{C}^d = \mathcal{X} \oplus \mathcal{X}^\perp$. En este caso, tomamos

$$\gamma(t) = Y_r(t)^* \, A \, Y_r(t) = X \, (\, U(t)^* \, A \, U(t) \,) \, X = X^* \, A(t) \, X \quad \text{para} \quad t \in [0,1] \, , \tag{4.32}$$

donde $A(t) = U(t)^* \, A \, U(t)$. Derivando,

$$\gamma'(t) = X^* \, A'(t) \, X = X^* \, (A(t) \, U'(0) - U'(0) \, A(t)) \, X \, , \tag{4.33}$$

donde hemos usado que $U(t+h) = U(t) \, U(h)$, para t, h, $t+h \in [0,1]$, de forma que $U'(t) = U(t) \, U'(0)$ y similarmente $(U'(t))^* = -U'(0) \, U(t)^*$. De forma análoga al desarrollo en la ecuación (4.19), usando las representaciones matriciales por bloques anteriores y tomando $A(t) := (A_{ij}(t))_{i,j=1}^2$, tenemos que

$$\gamma'(t) = X^* \, (A(t) \, U'(0) - U'(0) \, A(t)) \, X = 2 \, \mathrm{Re}(\, A_{12}(t) \, D_\Theta) \, . \tag{4.34}$$

Recordemos que $\mathcal{X}$ es un subespacio A-invariante, de forma que la representación por bloques de A es la expuesta en la ecuación (4.30). Luego, si consideramos la representación por bloques $A(t) = U(t)^* \, A \, U(t) = (A_{ij}(t))_{i,j=1}^2$ tenemos que,

$$A_{12}(t) = S(t) \, A_{22} \, C(t) - C(t) \, A_{11} \, S(t) \quad \text{para} \quad t \in [0,1] \, . \tag{4.35}$$

La ecuación (4.35) muestra que $\mathcal{X}$ podría no ser $A(t)$-invariante, para $t \in (0,1]$; sin embargo, basados en el hecho de que $\mathcal{X}$ es A-invariante, probaremos que es posible obtener una cota superior conveniente para el crecimiento de los valores singulares de $\gamma'(t)$, para $t \in (0,1]$. En efecto, usando la ecuación (4.35) vemos que

$$2 \, \mathrm{Re}(A_{12}(t) \; D_\Theta) = 2 \, \mathrm{Re}(\, (S(t) \, A_{22} \, C(t) - C(t) \, A_{11} \, S(t) \,) \, D_\Theta) \, .$$

Más aún, como $S(t)$, $C(t)$ y D_Θ conmutan entre si, entonces

$$2 \, \mathrm{Re}(A_{12}(t) \; D_\Theta) = 2 \, \mathrm{Re}(\, S(t)(A_{22} \, D_\Theta - D_\Theta \, A_{11}) \, C(t) \,) \, . \tag{4.36}$$

Usando los ítems 2. y 3. del Teorema 1.1.5 junto con las ecuaciones (4.34) y (4.36) obtenemos que

$$s(\gamma'(t) \,) = 2 \, s \, (\, \mathrm{Re}(A_{12}(t) \; D_\Theta) \,) \prec_w 2 \, t \, \Theta \, s(A_{22} \, D_\Theta - D_\Theta \, A_{11}) \, , \tag{4.37}$$

donde hemos usado que $s(S(t) \,) = \lambda(S(t) \,) = \mathrm{sen}(t \, \Theta) \leq t \, \Theta$ (entrada a entrada) y que $0 \leq C(t) \leq I$.

Proposición 4.3.2. *Bajo la Notación 4.3.1, supongamos que $\mathcal{X} \cap \mathcal{Y} = \{0\}$. Entonces*

$$s(X^* \, A \, X - Y_r^* \, A \, Y_r) \prec_w \Theta^2 \, \mathrm{Spr}^+(A) \, , \tag{4.38}$$

Demostración. Consideramos primero el caso $\mathcal{X}^\perp \cap \mathcal{Y}^\perp = \{0\}$. Bajo esta suposición, consideramos la curva $\gamma(t)$ para $t \in [0,1]$ dada en la ecuación (4.32). Como $\gamma(0) = X^*AX$ y $\gamma(1) = Y_r^* \, A \, Y_r$, la Proposición 4.2.1 muestra que

$$s(X^*AX - Y_r^* \, A \, Y_r) \prec_w \int_0^1 s(\gamma'(t)) \; dt \, . \tag{4.39}$$

Además, por la ecuación (4.37), tenemos que

$$s(\gamma'(t)) \prec_w 2 \, t \, \Theta \, s(A_{22} \, D_\Theta - D_\Theta \, A_{11}) \, . \tag{4.40}$$

Ahora bien, el Corolario 3.4.4 establece que

$$s(A_{22} \, D_\Theta - D_\Theta \, A_{11}) \prec_w \Theta \; \mathrm{Spr}^+(A_{11} \oplus A_{22}) = \Theta \; \mathrm{Spr}^+(A) \, , \tag{4.41}$$

donde hemos usado que $A = A_{11} \oplus A_{22}$ por hipótesis ($\mathcal{X}$ es A-invariante). Por lo tanto, combinando las ecuaciones (4.40) y (4.41),

$$s(\gamma'(t)) \prec_w 2\,t\,\Theta^2\ \mathrm{Spr}^+(A) \overset{(4.39)}{\Longrightarrow} s(X^*\,A\,X - Y_r^*\,A\,Y_r) \prec_w \Theta^2\,\mathrm{Spr}^+(A)\ ,$$

donde hemos usado el Lema 1.1.6. Finalmente, para el caso $\mathcal{X}^\perp \cap \mathcal{Y}^\perp \neq \{0\}$ podemos realizar un argumento análogo al utilizado en el final de la prueba del Teorema 4.1.2, para reducirlo al caso $\mathcal{X}^\perp \cap \mathcal{Y}^\perp = \{0\}$ y probar la desigualdad (4.38).

□

Proposición 4.3.3. *Utilizando la Notación 4.3.1 y asumiendo que $\mathcal{X}$, $\mathcal{Y} \subset \mathbb{C}^d$ son subespacios tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k \leq d/2$. Entonces tenemos que*

$$s(X^*\,A\,X - Y_r^*\,A\,Y_r) \prec_w \Theta^2\,\mathrm{Spr}^+(A)\,. \tag{4.42}$$

Demostración. Asumamos que $k = \dim(\mathcal{X}) = \dim(\mathcal{Y}) \leq d/2$ y que $\mathcal{X}$ e $\mathcal{Y}$ son subespacios arbitrarios tales que $\mathcal{X}$ es A-invariante. Si $\mathcal{X} \cap \mathcal{Y} = \{0\}$ entonces la ecuación (4.42) es consecuencia de la Proposición 4.3.2. En caso que $\mathcal{X} \cap \mathcal{Y} \neq \{0\}$ consideramos la siguiente

Afirmación: para todo $t \in [0, \pi/2]$ existe un subespacio k-dimensional $\mathcal{Y}(t) \subset \mathbb{C}^d$ y una rotación directa $W(t) \in \mathcal{U}(d)$ de $\mathcal{X}$ en $\mathcal{Y}(t)$ de forma que $\mathcal{X} \cap \mathcal{Y}(t) = \{0\}$ para $t \in (0, \pi/2]$,

$$\lim_{t\to 0^+} P_{\mathcal{Y}(t)} = P_{\mathcal{Y}} \in \mathcal{H}(d) \qquad \text{y} \qquad \lim_{t\to 0^+} W(t) = U\,.$$

Asumamos primero que la afirmación es cierta (por un momento). Sea $Y_{r,t} = W(t)\,X$, para $t \in [0, \pi/2]$ y notemos que $\lim_{t\to 0^+} Y_{r,t} = U\,X = Y_r$ (como en el enunciado del teorema). Dado que $\mathcal{X} \cap \mathcal{Y}(t) = \{0\}$ para $t \in (0, \pi/2]$, por la Proposición 4.3.2 tenemos que

$$s(X^*AX - Y_{r,t}^*\,A\,Y_{r,t}) \prec_w \Theta(\mathcal{X}, \mathcal{Y}(t))^2\,\mathrm{Spr}^+(A) \quad \text{para} \quad t \in (0, \pi/2]\,.$$

Por continuidad de los valores singulares (y por lo tanto de los ángulos principales) concluimos que

$$\begin{aligned} s(X^*AX - Y_r^*AY_r) &= \lim_{t\to 0^+} s(X^*AX - Y_{r,\,t}^*\,A\,Y_{r,\,t}) \prec_w \lim_{t\to 0^+} \Theta(\mathcal{X}, \mathcal{Y}(t))^2\,\mathrm{Spr}^+(A) \\ &= \Theta(\mathcal{X}, \mathcal{Y})^2\,\mathrm{Spr}^+(A)\,. \end{aligned}$$

Prueba de la afirmación. Si $\dim(\mathcal{X} \cap \mathcal{Y}) = s \geq 1$, notamos que

$$\dim(\mathcal{X}^\perp \cap \mathcal{Y}^\perp) = \dim(\mathcal{X}^\perp) + \dim(\mathcal{Y}^\perp) - \dim(\mathcal{X} \cap \mathcal{Y})^\perp \geq s\,.$$

Por otro lado, tenemos las descomposiciones

$$\mathcal{X} = (\mathcal{X} \cap \mathcal{Y}) \oplus (\mathcal{X} \cap \mathcal{Y}^\perp) \oplus (\mathcal{X} \cap \mathcal{G}) \quad \text{e} \quad \mathcal{Y} = (\mathcal{X} \cap \mathcal{Y}) \oplus (\mathcal{X}^\perp \cap \mathcal{Y}) \oplus (\mathcal{Y} \cap \mathcal{G})\,,$$

donde $\mathcal{G}$ denota la parte genérica de los subespacios $\mathcal{X}$ e $\mathcal{Y}$ (ver descomposición de Halmos). Sean $\Theta' = (\frac{\pi}{2}\,\mathbb{1}_p\,, \theta_1\,\ldots, \theta_r)$ los ángulos principales entre $\mathcal{S}_1 := (\mathcal{X} \cap \mathcal{Y}^\perp) \oplus (\mathcal{X} \cap \mathcal{G})$ y $(\mathcal{X}^\perp \cap \mathcal{Y}) \oplus (\mathcal{Y} \cap \mathcal{G})$, donde $p = \dim(\mathcal{X} \cap \mathcal{Y}^\perp) = \dim(\mathcal{X}^\perp \cap \mathcal{Y})$ y $r = \dim(\mathcal{X} \cap \mathcal{G}) = \dim(\mathcal{Y} \cap \mathcal{G})$; entonces, $\Theta = \Theta(\mathcal{X}\,, \mathcal{Y}) = (\Theta'\,, 0_s) \in [0, \pi/2]^k$, donde $k = p + r + s$.

Sea $U \in \mathcal{U}(d)$ una rotación directa de $\mathcal{X}$ en $\mathcal{Y}$. Entonces existe una BON $\mathcal{B}$ para $\mathbb{C}^d$, compatible con la descomposición

$$\mathbb{C}^d = (\mathcal{X} \cap \mathcal{Y}) \oplus \mathcal{S}_1 \oplus \mathcal{S}_2 \oplus (\mathcal{X}^\perp \cap \mathcal{Y}^\perp)\,, \tag{4.43}$$

donde $\mathcal{S}_2 = (\mathcal{X}^\perp \cap \mathcal{Y}) \oplus (\mathcal{X}^\perp \cap \mathcal{G})$, tal que

$$U = \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & C & -S & 0 \\ 0 & S & C & 0 \\ 0 & 0 & 0 & I \end{pmatrix}, \tag{4.44}$$

$$\text{donde} \quad C = \operatorname{diag}(\cos(\Theta')) \quad \text{y} \quad S = \operatorname{diag}(\operatorname{sen}(\Theta')) \in \mathcal{M}_{p+r}(\mathbb{C})^+$$

son matrices diagonales semidefinidas positivas.

Sean $\{v_j\}_{j\in\mathbb{I}_s}$ vectores en $\mathcal{B}$ que forman una BON de $\mathcal{X}\cap\mathcal{Y}$ y $\{w_j\}_{j\in\mathbb{I}_s}$ los vectores en $\mathcal{B}$ que forman un sistema ortonormal en $\mathcal{X}^\perp \cap \mathcal{Y}^\perp$ (usando que $\dim \mathcal{X}^\perp \cap \mathcal{Y}^\perp \geq s$). Luego tomamos $u_j(t) = \cos(t)\, v_j + \operatorname{sen}(t)\, w_j$, para $j \in \mathbb{I}_s$ y $\mathcal{S}(t)$ el subespacio generado por el sistema ortonormal $\{u_j(t)\}_{j\in\mathbb{I}_s}$, para $t \in [0, \pi/2]$. Además, consideramos

$$\mathcal{Y}(t) = \mathcal{S}(t) \oplus (\mathcal{X}^\perp \cap \mathcal{Y}) \oplus (\mathcal{Y} \cap \mathcal{G}) \quad \text{para} \quad t \in [0, \pi/2]\,.$$

Por construcción, $\dim(\mathcal{Y}(t) = k)$ y para $0 < t \leq \theta_r$ tenemos que

$$\Theta(\mathcal{X}, \mathcal{Y}(t)) = (\theta_j(t))_{j\in\mathbb{I}_k} = (\Theta',\, t\,\mathbb{1}_s) \in (0, \pi/2]^k\,.$$

En particular, $\mathcal{X} \cap \mathcal{Y}(t) = \{0\}$ para $0 < t < \theta_r$. Además $\lim_{t\to 0^+} P_{\mathcal{Y}(t)} = P_{\mathcal{Y}}$, pues $\mathcal{S}(0) = \mathcal{X} \cap \mathcal{Y}$.

Consideremos la descomposición ortogonal

$$\mathbb{C}^d = \mathcal{Z}_1 \oplus \mathcal{Z}_2 \oplus \mathcal{T}\,, \tag{4.45}$$

donde $\mathcal{T} = (\mathcal{X}^\perp \cap \mathcal{Y}^\perp) \ominus \mathcal{S}(\frac{\pi}{2})$,

$$\mathcal{Z}_1 = (\mathcal{X}\cap\mathcal{Y}) \oplus (\mathcal{X}\cap\mathcal{Y}^\perp) \oplus (\mathcal{X}\cap\mathcal{G}) \quad \text{y} \quad \mathcal{Z}_2 = (\mathcal{X}^\perp\cap\mathcal{Y}) \oplus (\mathcal{X}^\perp\cap\mathcal{G}) \oplus \mathcal{S}(\frac{\pi}{2})\,.$$

Ahora bien, a partir de la descomposición dada en la ecuación (4.45), construímos la matriz por bloques $W(t)$ dada por

$$W(t) = \begin{pmatrix} C_1(t) & -S_1(t) & 0 \\ S_1(t) & C_1(t) & 0 \\ 0 & 0 & I \end{pmatrix} \quad \text{para} \quad t \in [0, \pi/2]\,, \tag{4.46}$$

donde $C_1(t) = \operatorname{diag}(\cos(t)\,\mathbb{1}_s, \cos(\Theta'))$ y $S_1(t) = \operatorname{diag}(\operatorname{sen}(t)\,\mathbb{1}_s, \operatorname{sen}(\Theta'))$. Luego, $W(\cdot)$ es una función continua tal que $W(t)$ es una rotación directa de $\mathcal{X}$ en $\mathcal{Y}(t)$, para $t \in [0, \pi/2]$, y tal que $W(0) = U$ (comparar las ecuaciones (4.44) y (4.46)). Estos hechos prueban la afirmación.

□

Ahora estamos en condiciones de probar el Teorema 4.1.4, que enunciamos nuevamente a continuación.

Teorema 4.1.4. Bajo las hipótesis del Teorema 4.1.2, y asumiendo además que el subespacio k-dimensional $\mathcal{X}$ es A-invariante, tenemos que

$$s(X^*\,A\,X - Y_r^*\,A\,Y_r) \prec_w \Theta^2(\mathcal{X}, \mathcal{Y})\,\operatorname{Spr}^+(A)\,. \tag{4.47}$$

Demostración. En caso que $k \leq d/2$, el resultado es consecuencia de la Proposición 4.3.3. Por otro lado, si $k > d/2$ consideramos los subespacios $\mathcal{X}$, $\mathcal{Y}$ inmersos en $\mathbb{C}^{d'} = \mathbb{C}^d \oplus \mathbb{C}^{(d'-d)}$ para cierto $k \leq d'/2$, tomando $\mathcal{X}' = \mathcal{X} \oplus 0_{(d'-d)} \subset \mathbb{C}^{d'}$ y $\mathcal{Y}' = \mathcal{Y} \oplus 0_{(d'-d)} \subset \mathbb{C}^{d'}$. Notemos que en este caso $\Theta(\mathcal{X}', \mathcal{Y}') = (\Theta(\mathcal{X}, \mathcal{Y}), 0_{(d'-d)})$.

De forma similar, insertamos X, A y U obteniendo $X' = \begin{pmatrix} X \\ 0 \end{pmatrix} \in \mathcal{I}_{\mathcal{X}'}(k, d')$,

$$A' = \begin{pmatrix} A & 0 \\ 0 & \lambda_h(A)\, I \end{pmatrix} \in \mathcal{H}(d') \quad \text{y} \quad U' = \begin{pmatrix} U & 0 \\ 0 & I \end{pmatrix} \in \mathcal{U}(d')\,,$$

donde $h = [\frac{d+1}{2}]$ (parte entera). En este caso U' es una rotación directa de $\mathcal{X}'$ en $\mathcal{Y}'$ tal que

$$Y_r' = U'X' = \begin{pmatrix} UX \\ 0 \end{pmatrix} = \begin{pmatrix} Y_r \\ 0 \end{pmatrix}\,.$$

Luego tenemos que

$$(X')^* A'\, X' - (Y_r')^* A'\, Y_r' = \begin{pmatrix} X^*AX - Y_r^*AY_r \\ 0 \end{pmatrix}\,.$$

Por lo tanto, por la Proposición 4.3.3 resulta que

$$s(X^*AX - Y_r^*AY_r) = s((X')^* A'\, X' - (Y_r')^* A'\, Y_r') \prec_w \Theta(\mathcal{X}', \mathcal{Y}')^2 \operatorname{Spr}^+(A')\,.$$

Finalmente, por construcción,

$$\operatorname{Spr}^+(A') = (\operatorname{Spr}^+(A), 0_{h'}) \implies \Theta(\mathcal{X}', \mathcal{Y}')^2 \operatorname{Spr}^+(A') = \Theta(\mathcal{X}, \mathcal{Y})^2 \operatorname{Spr}^+(A)\,,$$

donde $h' = [\frac{d'}{2}] - [\frac{d}{2}] \geq 1$. □

4.4. Una aplicación: sobre la variación absoluta de los valores de Ritz

Como ya hemos mencionado anteriormente, una de las motivaciones pare el estudio de la variación absoluta de los valores de Ritz de $A \in \mathcal{H}(d)$ alrededor del subespacio $\mathcal{X}$ de $\mathbb{C}^d$, es que dicha variación provee de una medida indirecta de la calidad del subespacio $\mathcal{X}$ como posible subespacio invariante de A. Como ejemplo de dicho fenómeno, recordamos la siguiente desigualdad mencionada en la Sección 2.1: sea $P_{\mathcal{X}+\mathcal{Y}}$ la proyección ortogonal sobre el subespacio $\mathcal{X} + \mathcal{Y}$ de $\mathbb{C}^d$, sean $R_X = AX - X(X^*AX)$ y $R_Y = AY - Y(Y^*AY)$ los residuos de A en X e Y. Si $\mathcal{X}$ e $\mathcal{Y}$ se encuentran en posición aguda (es decir, de forma que $\theta_j < \pi/2$, para $j \in \mathbb{I}_k$) entonces el Corolario 2.1.6 establece que

$$|\lambda(X^*AX) - \lambda(Y^*AY)| \prec_w [s(P_{\mathcal{X}+\mathcal{Y}}\, R_X) + s(P_{\mathcal{X}+\mathcal{Y}}\, R_Y)]\, (\tan(\theta_j))_{j=1}^k\,. \tag{4.48}$$

Ahora bien, cuando $\mathcal{X}$ e $\mathcal{Y}$ están cerca, los residuos de A en X e Y son comparables; más aún, como podemos comparar $\tan(\theta_j) \approx \theta_j$, para $j \in \mathbb{I}_k$, concluímos que la desigualdad de la ecuación (4.48) provee una cota superior de primer orden para la variación absoluta de los valores de Ritz de A alrededor de $\mathcal{X}$, en términos del residuo R_X (donde la distancia entre los subespacios es medida en términos de los ángulos principales). Por lo tanto, si R_X es pequeño, entonces la variación absoluta también resulta pequeña.

En este sentido estamos interesados en obtener cotas superiores *autónomas* para la variación absoluta de los valores de Ritz de $A \in \mathcal{H}(d)$ asociados con los subespacios k-dimensionales $\mathcal{X}$ e $\mathcal{Y}$, en términos de los ángulos principales entre $\mathcal{X}$ e $\mathcal{Y}$, y la dispersión espectral de A. En este contexto, Knyazev y Argentati conjeturaron (ver [30, Conjetura 2.1]) las cotas superiores de las ecuaciones (2.31) (caso general) y (2.32) (caso $\mathcal{X}$ es A-invariante). El siguiente resultado confirma parcialmente dichas conjeturas.

Teorema 4.4.1. *Sea $A \in \mathcal{H}(d)$ y sean $X, Y \in \mathcal{I}(k, d)$, con rangos $\mathcal{X} = R(X)$ e $\mathcal{Y} = R(Y)$ tales que $\dim(\mathcal{X}) = \dim(\mathcal{Y}) = k$. Si $\Theta = \Theta(\mathcal{X}, \mathcal{Y})$ entonces:*

1. *tenemos que*

$$|\lambda(X^*AX) - \lambda(Y^*AY)| \prec_w \Theta \operatorname{Spr}^+(A)\,. \tag{4.49}$$

2. Más aún, si $\mathcal{X}$ es A-invariante, tenemos que

$$|\lambda(X^*AX) - \lambda(Y^*AY)| \prec_w \Theta^2 \operatorname{Spr}^+(A)\,. \tag{4.50}$$

Demostración. Sea $U \in \mathcal{U}(d)$ una rotación directa de $\mathcal{X}$ en $\mathcal{Y}$ y sea $Y_r = U\,X$ como en el Teorema 4.1.2. Como $R(Y_r) = R(Y)$ entonces $V = Y_r^*Y \in \mathcal{U}(k)$ y tenemos que

$$\lambda(Y^*A\,Y) = \lambda(V\,Y^*A\,Y\,V^*) = \lambda(Y_r^*\,P_{\mathcal{Y}}\,A\,P_{\mathcal{Y}}\,Y_r) = \lambda(Y_r^*\,A\,Y_r) \in \mathbb{R}^k\,.$$

Por la desigualdad de Lidskii (Teorema 1.1.4) y el Teorema 4.1.2 tenemos que

$$\begin{aligned} |\lambda(X^*AX) - \lambda(Y^*AY)| &= |\lambda(X^*AX) - \lambda(Y_r^*\,A\,Y_r)| \\ &\prec_w |\lambda(X^*AX - Y_r^*A\,Y_r)|^{\downarrow} \\ &= s(X^*AX - Y_r^*A\,Y_r) \prec_w \Theta\operatorname{Spr}^+(A)\,. \end{aligned}$$

Para probar el ítem 2. debemos realizar un argumento análogo y utilizar el Teorema 4.1.4.

Observación 4.4.2. Comparando las Conjeturas (2.31) y (2.32) con el Teorema 4.4.1, notamos que no aparece como factor $\operatorname{Spr}^+(A, \mathcal{X}+\mathcal{Y})$ sino $\operatorname{Spr}^+(A)$. Esto que a priori parecería indicar una peor estimación en términos de la dispersión espectral, es en realidad un detalle técnico.

En efecto, inspeccionando la prueba del Teorema 4.1.2 para el caso $A = B$, podemos observar la aparición en las estimaciones finales de una matriz $A' = Z^*A\,Z$, que se corresponde con la compresión $A_{\mathcal{X}+\mathcal{Y}} = P_{\mathcal{X}+\mathcal{Y}}A_{|\mathcal{X}+\mathcal{Y}}$, que a su vez produce estimaciones en términos de $\operatorname{Spr}^+(A') = \operatorname{Spr}^+(A, \mathcal{X}+\mathcal{Y})$. De esta forma, las estimaciones obtenidas pueden interpretarse sin pérdida de generalidad en términos de A.

Por lo tanto la única diferencia entre las Conjeturas (2.31) y (2.32) y el Teorema 4.4.1, recae en las cotas ligeramente más grandes Θ (respectivamente Θ^2) en lugar de $\operatorname{sen}(\Theta)$ (respectivamente $\operatorname{sen}^2(\Theta)$). Estos números se encuentran asintóticamente cerca cuando $\Theta \to 0^+$ y pueden ser comparados globalmente mediante la constante $\pi/2$.

Sin embargo, las Conjeturas (2.31) y (2.32) han sido testeadas numéricamente de forma exhaustiva. Para explorar nuestros resultados, consideramos los siguientes ejemplos, que prueban que tanto las Conjeturas (2.31) y (2.32) y el Teorema 4.4.1 son ajustadas. △

Ejemplo 4.4.3. Consideremos $a > b > 0$ y sea

$$A = \begin{pmatrix} 0 & 0 & a & 0 \\ 0 & 0 & 0 & b \\ a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \end{pmatrix} \in \mathcal{H}(4)\,.$$

Si $\{e_1, e_2, e_3, e_4\}$ denota la base canónica de $\mathbb{C}^4$, sea $\mathcal{X} = \operatorname{Span}\{e_1, e_2\}$. Por otro lado, dado $\theta \in [0, \pi/2]$ sean

$$f_1(\theta) = \cos(\theta)\,e_1 + \operatorname{sen}(\theta)\,e_3 \qquad \text{y} \qquad f_2(\theta) = \cos(\theta)\,e_2 + \operatorname{sen}(\theta)\,e_4\,. \tag{4.51}$$

Además, tomamos $\mathcal{Y}(\theta) = \operatorname{Span}\{f_1(\theta),\ f_2(\theta)\}$, para $\theta \in [0, \pi/2]$. Luego, mediante un cálculo directo puede verse que $\Theta(\mathcal{X}\,,\,\mathcal{Y}(\theta)) = (\theta,\,\theta)$. Por lo tanto, las isometrías X, $Y_r(\theta)$ del Teorema 4.1.2 (asociadas a los subespacios $\mathcal{X}$ e $\mathcal{Y}(\theta)$) están dadas por

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{y} \qquad Y_r(\theta) = \begin{pmatrix} \cos(\theta) & 0 \\ 0 & \cos(\theta) \\ \operatorname{sen}(\theta) & 0 \\ 0 & \operatorname{sen}(\theta) \end{pmatrix}\,. \tag{4.52}$$

Un cálculo directo permite ver que

$$X^*AX = 0 \qquad \text{y} \qquad Y_r(\theta)^*AY_r(\theta) = \begin{pmatrix} a\,\operatorname{sen}(2\theta) & 0 \\ 0 & b\,\operatorname{sen}(2\theta) \end{pmatrix}.$$

Por lo tanto, para $\theta \in [0, \pi/2]$,

$$|\lambda(X^*AX) - \lambda(Y_r(\theta)^*AY_r(\theta))| = |\lambda(X^*AX - Y_r(\theta)^*AY_r(\theta))| = \operatorname{sen}(2\theta)\,(a, b)\,,$$

y en particular, la desigualdad de Lidskii es una igualdad en este caso. Por otro lado, resulta que $\lambda(A) = (a, b, -b, -a) \in (\mathbb{R}^4)^{\downarrow}$ que muestra que $\operatorname{Spr}^+(A) = 2\,(a, b)$. Por lo tanto, las desigualdades de los Teoremas 4.1.2 (con $A = B$) y 4.4.1 (ítem 1.) se convierten en

$$\operatorname{sen}(2\theta)\,(a, b) \prec_w 2\,(\theta\,,\,\theta)\,(a, b) = 2\,\theta\,(a, b)\,,$$

donde la relación de submayorización anterior es equivalente a las siguientes desigualdades

$$\operatorname{sen}(2\theta)\,a \le 2\,\theta\,a \qquad \text{y} \qquad \operatorname{sen}(2\theta)\,(a + b) \le 2\,\theta\,(a + b)\,,$$

que son ajustadas; esto puede verse considerando $\theta \to 0^+$. Más aún, obtenemos que

$$\lim_{\theta \to 0^+} \frac{|\lambda(X^*AX) - \lambda(Y_r(\theta)^*AY_r(\theta))|}{\Theta(\mathcal{X},\, \mathcal{Y}(\theta))\operatorname{Spr}^+(A)} = \mathbb{1}_2\,,$$

considerando el cociente entrada a entrada. Este hecho muestra que nuestra cota superior para la variación absoluta de los valores de Ritz es ajustada. △

Ejemplo 4.4.4. Consideremos $a > b > 0$ y sea

$$A = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \in \mathcal{H}(4)\,.$$

Tomamos la base canónica $\{e_1, e_2, e_3, e_4\}$ de $\mathbb{C}^4$ y consideramos $\mathcal{X} = \operatorname{Span}\{e_1, e_2\}$. Notemos que en este caso $\mathcal{X}$ es un subespacio A-invariante. Por otro lado, dado $\theta \in [0, \pi/2]$ sea $\mathcal{Y}(\theta) = \operatorname{Span}\{f_1(\theta), f_2(\theta)\}$, donde los $f_j(\theta)$ están dados como en la ecuación (4.51). Luego como en el ejemplo anterior, tenemos que $\Theta(\mathcal{X}\,,\,\mathcal{Y}(\theta)) = (\theta,\, \theta)$. En este caso, las isometrías X, $Y_r(\theta)$ como en el Teorema 4.1.4 están dadas por la ecuación (4.52). Luego, realizando algunos cálculos se puede ver que

$$X^*AX = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \qquad \text{y} \qquad Y_r(\theta)^*AY_r(\theta) = \begin{pmatrix} a\,\cos^2(\theta) & 0 \\ 0 & b\,\cos^2(\theta) \end{pmatrix}.$$

Por lo tanto, para $\theta \in [0, \pi/2]$,

$$|\lambda(X^*AX) - \lambda(Y_r(\theta)^*AY_r(\theta))| = |\lambda(X^*AX - Y_r(\theta)^*AY_r(\theta))| = \operatorname{sen}^2(\theta)\,(a, b)\,,$$

y en particular la desigualdad de Lidskii es una igualdad en este caso. Más aún, resulta que $\lambda(A) = (a, b, 0, 0) \in (\mathbb{R}^4)^{\downarrow}$ que prueba que $\operatorname{Spr}^+(A) = (a, b)$. En consecuencia los Teoremas 4.1.4 y 4.4.1 (ítem 2.) se convierten en

$$\operatorname{sen}^2(\theta)\,(a, b) \prec_w (\theta\,,\,\theta)^2\,(a, b) = \theta^2\,(a, b)\,,$$

donde la relación de submayorización anterior es equivalente a las desigualdades

$$\operatorname{sen}^2(\theta)\,a \le \theta^2\,a \qquad \text{y} \qquad \operatorname{sen}^2(\theta)\,(a + b) \le \theta^2\,(a + b)\,,$$

que resultan ajustadas; esto puede verse considerando $\theta \to 0^+$. Notemos que además se obtiene que

$$\lim_{\theta \to 0^+} \frac{|\lambda(X^*AX) - \lambda(Y_r(\theta)^*AY_r(\theta))|}{\Theta(\mathcal{X}, \mathcal{Y}(\theta))^2\operatorname{Spr}^+(A)} = \mathbb{1}_2\,.$$

Esto muestra que nuestra cota superior para la variación absoluta de los valores de Ritz es ajustada. △

Observación 4.4.5. Consideremos la notación del Teorema 4.4.1. Uno puede preguntarse por qué nuestras cotas no coinciden con las cotas conjeturadas en la Observación 2.3.3. Ahora que hemos descrito nuestras técnicas en detalle, podemos dar una opinión sobre este tema. El hecho de que nuestras cotas dependan de $\Theta(\mathcal{X},\mathcal{Y})$, que desde un punto de vista geométrico corresponde a la longitud del arco entre subespacios, es una consecuencia de la elección de curvas que conducen a la construcción de $\gamma(t)$ como en las ecuaciones (4.15) y (4.32), basados en rotaciones directas. Por otro lado $\operatorname{sen}(\Theta(\mathcal{X},\mathcal{Y}))$, que desde un punto de vista geométrico corresponde a la longitud cordal entre subespacios, parece estar asociado con curvas más cortas $\tilde{\gamma}(t)$. Aun así, la importancia de nuestra elección $\gamma(t)$ es que su derivada $\gamma'(t)$ conduce a desigualdades de valores singulares que nos permiten reducir el problema al nivel infinitesimal.

Por ejemplo, notemos que la curva más corta $\mu(t) = (1-t)X^*A\,X - t\,Y^*A\,Y$, $t \in [0,1]$, que une $X^*A\,X$ y $Y^*A\,Y$ no proporciona tal reducción, ya que $\mu'(t) = Y^*A\,Y - X^*A\,X$, para $t \in [0,1]$. Sin embargo, la técnica geométrica considerada aquí es bastante flexible por lo que podríamos utilizarla en el futuro para seguir trabajando en estos problemas. △

Como comentario final, notemos que es natural preguntarse si las estimaciones de los Teoremas 4.1.2 y 4.1.4 pueden ser usadas para obtener estimaciones para la distancia entre compresiones (que podemos pensar como aproximaciones de rango bajo de A) dadas por $P_{\mathcal{X}}\,A\,P_{\mathcal{X}}$ y $P_{\mathcal{Y}}\,A\,P_{\mathcal{Y}}$. Resulta que no es el caso. En efecto, en el caso trivial cuando $A = I$ tenemos que $P_{\mathcal{X}}\,I\,P_{\mathcal{X}} - P_{\mathcal{Y}}\,I\,P_{\mathcal{Y}} = P_{\mathcal{X}} - P_{\mathcal{Y}}$; pero $\operatorname{Spr}^+(I) = 0$ es el vector nulo, por lo que no hay esperanzas de obtener una cota superior para $P_{\mathcal{X}}\,A\,P_{\mathcal{X}} - P_{\mathcal{Y}}\,A\,P_{\mathcal{Y}}$ en términos de $\operatorname{Spr}^+(A)$ en general. Además, en general no hay dependencia de $P_{\mathcal{X}}\,A\,P_{\mathcal{X}} - P_{\mathcal{Y}}\,A\,P_{\mathcal{Y}}$ en términos de $\Theta(\mathcal{X},\mathcal{Y})^2$ cuando $\mathcal{X}$ es A-invariante (nuevamente tomando $A = I$ puede verse este hecho).

Capítulo 5

Dispersión espectral de operadores compactos y autoadjuntos

Teniendo en cuenta lo desarrollado en el Capítulo 3, es natural preguntarse si el concepto de dispersión espectral puede extrapolarse al contexto de operadores compactos y autoadjuntos en un espacio de Hilbert separable de dimensión infinita. En este sentido, desarrollaremos primero el concepto de dispersión espectral y veremos algunas propiedades básicas, con el objetivo de revisar la desigualdad Aritmético-Geométrica y otras desigualdades asociadas. Así probaremos una desigualdad clave (que podría decirse análoga al Teorema 3.2.1), que permitirá obtener una versión de la desigualdad Aritmético-Geométrica (AG) de Bhatia y Kittaneh [9] y consecuencias relacionadas con la desigualdad de Zhan (ver [54, 55]).

5.1. Dispersión espectral de operadores autoadjuntos

En esta sección introduciremos las propiedades básicas para la dispersión espectral. En particular, $\mathcal{H}$ denotará un espacio de Hilbert separable (de dimensión infinita) y trabajaremos con el álgebra de perturbaciones compactas de múltiplos de la identidad $\mathcal{A}(\mathcal{H}) = K(\mathcal{H}) \oplus \mathbb{C}\, I$. Además probaremos una desigualdad clave para operadores autoadjuntos $A \in \mathcal{A}(\mathcal{H})$ que dará una cota superior para los valores singulares de cualquier bloque antidiagonal de A (con respecto a la relación de submayorización).

Definición y propiedades básicas

En esta sección trabajaremos con la C^*-subálgebra unital $\mathcal{A}(\mathcal{H}) = K(\mathcal{H}) \oplus \mathbb{C}\, I \subseteq L(\mathcal{H})$ y con su parte autoadjunta $\mathcal{A}(\mathcal{H})^{sa} = \mathcal{A}(\mathcal{H}) \cap L(\mathcal{H})^{sa}$. En lo que sigue, construiremos la dispersión espectral completa en términos de un operador arbitrario $A \in \mathcal{A}(\mathcal{H})^{sa}$, basados en el caso compacto considerado en la Definición 1.5.1.

Observación 5.1.1. Dado $A \in \mathcal{A}(\mathcal{H})^{sa}$ remarcamos que existe un único $\mu \in \mathbb{R}$ tal que

$$A_0 \overset{\text{def}}{=} A - \mu\, I \in K(\mathcal{H})^{sa}\,.$$

En efecto, si $A \in \mathcal{A}(\mathcal{H})^{sa}$, entonces $A = \mu\, I + A_0$ donde A_0 es compacto y $\mu \in \mathbb{R}$. Supongamos que existen $\alpha \in \mathbb{R}$ y $A_1 \in K(\mathcal{H})$ tales que $A = \alpha\, I + A_1$. Luego

$$\alpha\, I + A_1 = \mu\, I + A_0\,,$$

y por lo tanto $A_1 - A_0 = (\mu - \alpha)\, I$. Sin embargo, $A_1 - A_0 \in K(\mathcal{H})$, pero $r\, I \notin K(\mathcal{H})$ para $r \neq 0$ (de otra forma $K(\mathcal{H}) = B(\mathcal{H})$). En consecuencia $\mu = \alpha$ y $A_1 = A_0$.

En lo que sigue, $A_0 \in K(\mathcal{H})^{sa}$ denotará siempre al único operador compacto asociado a $A \in \mathcal{A}(\mathcal{H})^{sa}$ como antes.

En estos términos, podemos construír la escala espectral completa de A en función de la escala espectral completa de $A_0 \in K(\mathcal{H})^{sa}$ (ver Definición 1.5.1) de la siguiente forma: consideramos la escala espectral completa de A dada por

$$\lambda(A) \stackrel{\text{def}}{=} \lambda(A_0) + \mu\, \mathbb{1} \in \ell^\infty(\mathbb{Z}_0)\,, \tag{5.1}$$

donde $\mathbb{1} \in \ell^\infty(\mathbb{Z}_0)$ es la sucesión cuyas entradas son iguales a 1. Notemos que

$$\lambda_{-n}(A) \leq \lambda_{-(n+1)}(A) \leq \mu \leq \lambda_{n+1}(A) \leq \lambda_n(A) \quad , \quad \text{para todo} \quad n \in \mathbb{N}\,. \qquad \triangle$$

La siguiente noción está motivada por la dispersión espectral de matrices autoadjuntas, introducida en [30] (que hemos usado en los capítulos anteriores).

Definición 5.1.2. Dado $A \in \mathcal{A}(\mathcal{H})^{sa}$ definimos la *dispersión espectral completa* de A, notada $\mathrm{Spr}(A) \in c_0(\mathbb{Z}_0)$ como la sucesión: si $A = A_0 + \mu\, I$ con $A_0 \in K(\mathcal{H})^{sa}$, entonces

$$\mathrm{Spr}(A) \stackrel{\text{def}}{=} (\,\lambda_i(A) - \lambda_{-i}(A)\,)_{i\in\mathbb{Z}_0} = \lambda(A_0) + \lambda(-A_0)\,. \tag{5.2}$$

Además, consideramos la *dispersión espectral* de A, como la parte no negativa de $\mathrm{Spr}(A) \in c_0(\mathbb{N})$:

$$\mathrm{Spr}^+(A) \stackrel{\text{def}}{=} \big(\,\mathrm{Spr}_i(A)\,\big)_{i\in\mathbb{N}}\,. \tag{5.3}$$

$\triangle$

Observación 5.1.3. Por construcción, la dispersión espectral de un operador en $\mathcal{A}(\mathcal{H})^{sa}$ es una medida vectorial de la dispersión de sus autovalores. Es conveniente aclarar, que la definición aquí introducida de dispersión espectral es esencialmente una extensión del caso matricial.

Por otro lado, siguiendo la Observación 1.5.3, consideremos $x = (2,1) \in \mathbb{R}^2$ y la matriz $A = \mathrm{diag}(2,1) \in \mathcal{M}_2(\mathbb{C})$. Asociado a A, podemos construir un operador compacto de rango finito cuya representación por bloques es $\tilde{A} = A \oplus 0_{\ell^2(\mathbb{N})}$.

Luego, siguiendo la Definición 2.3.1, $\mathrm{Spr}^+(A) = 1$. Sin embargo, la ecuación (5.2) establece que $\mathrm{Spr}^+(\tilde{A}) = (2,1,0,...)$.

Esta diferencia sugiere que las propiedades de la dispersión espectral para operadores compactos pueden ser distintas a las propiedades de la dispersión espectral de matrices, lo que motiva los siguientes resultados. $\triangle$

En el siguiente resultado describimos varias propiedades básicas de la dispersión espectral en $\mathcal{A}(\mathcal{H})^{sa}$.

Proposición 5.1.4. *Sean $A, B \in \mathcal{A}(\mathcal{H})^{sa}$. Las siguientes propiedades son válidas:*

1. $\mathrm{Spr}(A) \in c_0^{\uparrow\downarrow}(\mathbb{Z}_0)$ *es antisimétrico* $(\mathrm{Spr}_{-j}(A) = -\mathrm{Spr}_j(A))$ *y* $\mathrm{Spr}^+(A) \in c_0^{\downarrow}(\mathbb{N}) \cap \mathbb{R}^{\mathbb{N}}_{\geq 0}$.
2. *La dispersión espectral es invariante bajo traslaciones, es decir, para cualquier $c \in \mathbb{R}$,*

$$\mathrm{Spr}(A + c\, I) = \mathrm{Spr}(A)\,.$$

3. *Para $c \in \mathbb{R}$ tenemos que* $\mathrm{Spr}(c\, A) = |c|\, \mathrm{Spr}(A)$. *En particular,* $\mathrm{Spr}^+(A) = \mathrm{Spr}^+(-A)$.

Demostración. Los ítems 1. y 2. son consecuencia directa de la definición de dispersión espectral (ver ecuación (5.2)). El ítem 3. es consecuencia del siguiente hecho: dado $A \in \mathcal{A}(\mathcal{H})^{sa}$ entonces $\lambda(cA) = c\,\lambda(A)$ si $c \geq 0$ y $\lambda(cA) = c\,(\lambda_{-i}(A))_{i\in\mathbb{Z}_0}$ si $c < 0$. $\square$

El siguiente resultado describe varias relaciones entre la dispersión espectral y los valores singulares de un operador compacto y autoadjunto. Notemos que algunas de las desigualdades propuestas son formalmente análogas a las vistas en la Observación 3.1.1 y la Proposición 3.1.2. Sin embargo, algunas propiedades de la Proposición 5.1.5 (como el caso positivo del ítem 1. y el ítem 6.) sólo son válidas en el caso compacto.

Proposición 5.1.5. *Sean A, $B \in K(\mathcal{H})^{sa}$.*

1. *Las siguientes desigualdades entrada a entrada son válidas:*

$$\mathrm{Spr}_i^+(A) = |\lambda_i(A)| + |\lambda_{-i}(A)| \leq 2s_i(A) \quad \text{para todo} \quad i \in \mathbb{N}\,. \tag{5.4}$$

En el caso positivo, tenemos que:

$$A \in K(\mathcal{H})^+ \implies \mathrm{Spr}_i^+(A) = \lambda_i(A) = s_i(A) \quad \text{para todo} \quad i \in \mathbb{N}\,. \tag{5.5}$$

2. *Sea $A \oplus A \in K(\mathcal{H} \oplus \mathcal{H})^{sa}$ dado por $A \oplus A = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$. Entonces*

$$\frac{1}{2}\,\mathrm{Spr}^+(A \oplus A) \prec_w s(A) \overset{(1.14)}{=} |\lambda(A)|^{\downarrow}\,. \tag{5.6}$$

3. *Sin embargo, para $A \in K(\mathcal{H})^{sa}$ tenemos que $s(A) \prec_w \mathrm{Spr}^+(A)$.*

4. *Si $\lambda(A) \prec \lambda(B)$ entonces $\mathrm{Spr}^+(A) \prec_w \mathrm{Spr}^+(B)$.*

5. *Si $P = P^* = P^2$ entonces, para todo $i \in \mathbb{N}$,*

$$\mathrm{Spr}_i^+(PAP) \leq \mathrm{Spr}_i^+(A) \implies \mathrm{Spr}^+(PAP) \prec_w \mathrm{Spr}^+(A)\,. \tag{5.7}$$

6. *Con el orden para los autovalores definido en 1.5.1, vale que*

$$\lambda(A \oplus 0_{\mathcal{H}}) = \lambda(A) \implies \mathrm{Spr}^+(A \oplus 0_{\mathcal{H}}) = \mathrm{Spr}^+(A)\,. \tag{5.8}$$

7. *Desigualdad aditiva para la dispersión espectral:*

$$\mathrm{Spr}(A) - \mathrm{Spr}(B) \prec \mathrm{Spr}(A - B) \prec \mathrm{Spr}(A) + \mathrm{Spr}(B)\,.$$

Demostración. Como $s(A) = |\lambda(A)|^{\downarrow}$ (ver la Definición 1.4.1 y la Observación 1.5.2), tenemos que

$$\text{máx}\{|\lambda_i(A)|, |\lambda_{-i}(A)|\} \leq s_i(A) \quad \text{para todo} \quad i \in \mathbb{N}\,. \tag{5.9}$$

Esto prueba el ítem 1. Para probar el ítem 2. notemos que

$$\sum_{i=1}^{n} \mathrm{Spr}_i^+(A \oplus A) = \begin{cases} 2\sum_{i=1}^{k} \mathrm{Spr}_i^+(A) & \text{si} \quad n = 2k \\ 2\sum_{i=1}^{k} \mathrm{Spr}_i^+(A) + \mathrm{Spr}_{k+1}^+(A) & \text{si} \quad n = 2k+1 \end{cases}.$$

Además, recordemos que $\sum_{i=1}^{n} s_i(A) = \text{máx}\{\sum_{i\in\mathbb{F}} |\lambda_i(A)| \ : \ \mathbb{F} \subset \mathbb{Z}_0\,,\ |\mathbb{F}| = n\}$. Entonces, usando que

$$\mathrm{Spr}_{k+1}^+(A) = |\lambda_{k+1}(A)| + |\lambda_{-(k+1)}(A)| \leq 2\ \text{máx}\{|\lambda_{k+1}(A)|, |\lambda_{-(k+1)}(A)|\}$$

y que

$$2\sum_{i=1}^{k}\mathrm{Spr}_i^+(A) = 2\sum_{i=1}^{k}|\lambda_i(A)| + |\lambda_{-i}(A)|\,,$$

se obtiene la relación de la ecuación (5.6).

Para probar el ítem 3. notemos que $s(A) \overset{(1.14)}{=} |\lambda(A)|^{\downarrow}$, por lo que, para $k \in \mathbb{N}$ obtenemos

$$\sum_{i=1}^{k} s_i(A) = \sum_{i=1}^{k_1}\lambda_i(A) + \sum_{i=1}^{k_2}|\lambda_{-i}(A)|$$

pára ciertos $0 \leq k_1$, k_2 tales que $k_1 + k_2 = k$. Luego,

$$\sum_{i=1}^{k} s_i(A) \leq \sum_{i=1}^{k}\lambda_i(A) + |\lambda_{-i}(A)| = \sum_{i=1}^{k}\mathrm{Spr}_i^+(A)\,.$$

Para probar el ítem 4., fijamos $k \in \mathbb{N}$. Como $\lambda(A) \prec\!\!\prec \lambda(B)$, entonces

$$\sum_{i=1}^{k}\lambda_i(A) \leq \sum_{i=1}^{k}\lambda_i(B) \qquad \text{y} \qquad \sum_{i=1}^{k}\lambda_{-i}(A) \geq \sum_{i=1}^{k}\lambda_{-i}(B)\,.$$

Por lo tanto,

$$\sum_{i=1}^{k}\mathrm{Spr}_i(A) = \sum_{i=1}^{k}\lambda_i(A) - \lambda_{-i}(A) \leq \sum_{i=1}^{k}\lambda_i(B) - \lambda_{-i}(B) = \sum_{i=1}^{k}\mathrm{Spr}_i(B)\,.$$

Para probar el ítem 5., aplicando las desigualdades de entrelace (1.16), obtenemos

$$\lambda_j(A) \geq \lambda_j(PAP) \geq 0 \qquad \text{y} \qquad \lambda_{-j}(A) \leq \lambda_{-j}(PAP) \leq 0 \qquad \text{para} \qquad j \in \mathbb{I}_k\,,$$

donde $1 \leq k = \mathrm{rk}(P) \leq \infty$. En particular, vemos que $\mathrm{Spr}_j^+(PAP) \leq \mathrm{Spr}_j^+(A)$, para $j \in \mathbb{I}_k$. Notemos que, para $j \notin \mathbb{I}_k$, tenemos que $\lambda_j(PAP) = \lambda_{-j}(PAP) = 0$ y entonces también vale que $\mathrm{Spr}_j^+(PAP) = 0 \leq \mathrm{Spr}_j^+(A)$. Luego la relación de submayorización $\mathrm{Spr}^+(PAP) \prec\!\!\prec_w \mathrm{Spr}^+(A)$ se obtiene de estos hechos.

El ítem 6. se obtiene notando que agregar ceros en los autovalores no afecta en el orden de $\lambda(A)$ dado en la Definición 1.5.1 (ver también Observación 1.5.3).

Para probar el ítem 7. notemos que por la desigualdad de Lidskii aditiva y el ítem 3. del Lema 1.4.5

$$\begin{aligned}\mathrm{Spr}(A) - \mathrm{Spr}(B) \quad &= \lambda(A) - \lambda(B) + \lambda(-A) - \lambda(-B) \\ &\overset{\text{Lidskii}}{\prec\!\!\prec} \lambda(A-B) + \lambda(-(A-B)) = \mathrm{Spr}(A-B)\,.\end{aligned}$$

Para la otra desigualdad, utilizamos la desigualdad de Weyl aditiva. En efecto, notemos que $\mathrm{Spr}(A-B) = \lambda(A-B) + \lambda(-(A-B))$. Por lo que

$$\begin{aligned}\mathrm{Spr}(A-B) \quad &= \lambda(A-B) + \lambda(B-A) \\ &\overset{\text{Weyl}}{\prec\!\!\prec} \lambda(A) + \lambda(-B) + \lambda(B) + \lambda(-A) = \mathrm{Spr}(A) + \mathrm{Spr}(B)\,,\end{aligned}$$

Nuevamente utilizando el ítem 3. del Lema 1.4.5. □

Observación 5.1.6. En lo que sigue obtendremos algunas desigualdades que involucran a la dispersión espectral de operadores compactos. En este contexto algunos de estos resultados serán formalmente análogos a aquellos previamente obtenidos en [39] para la dispersión espectral de matrices. Por este motivo, recordamos que la dispersión espectral de una matriz $A \in \mathcal{H}(d)$ no coincide con la dispersión espectral del operador obtenido insertando A como un operador compacto de rango finito $\tilde{A} \in K(\mathcal{H})^{sa}$ (ver Observación 5.1.3).

Esta distinción lleva por ejemplo a que las ecuaciones (5.5), (5.8) y la desigualdad del ítem 3. de la Proposición 5.1.5 (es decir, $s(A) \prec_w \mathrm{Spr}^+(A)$) fallen en el contexto matricial. Por lo tanto, las pruebas que exhibimos son necesarias en el contexto de operadores compactos. △

5.2. Una desigualdad clave: el caso compacto

En lo que sigue enunciamos una versión para operadores compactos y autoadjuntos del Teorema 3.2.1.

Teorema 5.2.1. *Sea* $A \in \mathcal{A}(\mathcal{H} \oplus \mathcal{K})^{sa}$ *tal que* $A = \begin{bmatrix} A_1 & B \\ B^* & A_2 \end{bmatrix} \begin{matrix} \mathcal{H} \\ \mathcal{K} \end{matrix}$ *es la representación por bloques de* A*. Entonces,* $B \in K(\mathcal{K}, \mathcal{H})$ *(por construcción) y*

$$2\, s(B) \prec_w \mathrm{Spr}^+(A)\,. \tag{5.10}$$

Demostración. Consideramos $U = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{matrix} \mathcal{H} \\ \mathcal{K} \end{matrix} \in \mathcal{U}(\mathcal{H} \oplus \mathcal{K})$. Entonces $UA - AU \in K(\mathcal{H} \oplus \mathcal{K})$ y

$$s(UA - AU) = s(A - U^*AU) \overset{(1.14)}{=} |\lambda(A - U^*AU)|^{\downarrow}\,.$$

Sea $A_0 \in K(\mathcal{H})^{sa}$ el operador asociado a A como en la Observación 5.1.1. Por la desigualdad de Weyl del Teorema 1.5.4, tenemos que

$$\lambda(A - U^*AU) = \lambda(A_0 - U^*A_0\,U) \prec \lambda(A_0) + \lambda(-U^*A_0\,U) = \mathrm{Spr}(A_0) = \mathrm{Spr}(A)\,.$$

Usando el ítem 2. del Lema 1.4.5 deducimos que

$$s(UA - AU) = |\lambda(A - U^*AU)|^{\downarrow} \prec_w |\mathrm{Spr}(A)|^{\downarrow} = \big(\mathrm{Spr}^+(A)\,,\, \mathrm{Spr}^+(A)\big)^{\downarrow}\,,$$

donde $\big(\mathrm{Spr}^+(A)\,,\, \mathrm{Spr}^+(A)\big) \in c_0(\mathbb{Z}_0)$ es construído como en la ecuación (1.11). Luego, mediante un cálculo directo y utilizando la Proposición 1.1.10 y la ecuación (1.14) vemos que

$$UA - AU = \begin{bmatrix} 0 & 2\,B \\ -2\,B^* & 0 \end{bmatrix} \begin{matrix} \mathcal{H} \\ \mathcal{K} \end{matrix} \implies s(UA - AU) = 2\,\big(s(B)\,,\, s(B)\big)^{\downarrow}\,,$$

y concluímos que

$$2\,\big(s(B)\,,\, s(B)\big)^{\downarrow} \prec_w \big(\mathrm{Spr}^+(A)\,,\, \mathrm{Spr}^+(A)\big)^{\downarrow} \implies 2\,s(B) \prec_w \mathrm{Spr}^+(A)\,. \qquad \square$$

El siguiente resultado está relacionado con la desigualdad de Zhan para valores singulares de la diferencia de operadores compactos y positivos (ver [54, 55]).

Teorema 5.2.2. *Sean* $A, B \in K(\mathcal{H})^{sa}$*. Entonces, si notamos* $A \oplus B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \in K(\mathcal{H} \oplus \mathcal{H})^{sa}$*,*

$$s(A - B) \prec_w \mathrm{Spr}^+(A \oplus B)\,. \tag{5.11}$$

Luego para cualquier norma unitariamente invariante N*,*

$$N(A - B) \leq g_N\big(\mathrm{Spr}^+(A \oplus B)\big)\,. \tag{5.12}$$

Demostración. Consideremos los operadores

$$Z = \frac{1}{\sqrt{2}} \begin{bmatrix} I & I \\ -I & I \end{bmatrix} \in \mathcal{U}(\mathcal{H} \oplus \mathcal{H}) \quad \text{y} \quad C = \frac{1}{2} \begin{bmatrix} A+B & A-B \\ A-B & A+B \end{bmatrix} \in K(\mathcal{H} \oplus \mathcal{H})^{sa} .$$

Notemos que $Z\,C\,Z^* = A \oplus B \in K(\mathcal{H} \oplus \mathcal{H})$. Entonces por el Teorema 5.2.1,

$$s(A-B) \prec_w \mathrm{Spr}^+(C) = \mathrm{Spr}^+(A \oplus B) . \tag{5.13}$$

Finalmente, por la ecuación (1.19) sabemos que (5.11) $\implies$ (5.12). □

Observación 5.2.3. En [54], Zhan obtiene las desigualdades para valores singulares $s_j(A-B) \leq s_j(A \oplus B)$, para $j \in \mathbb{N}$, donde A y B son operadores compactos y positivos. En este caso las desigualdades de Zhan fallan para operadores compactos y autoadjuntos (por ejemplo tomando A compacto y positivo y considerando $B = -A$).

En este contexto, el Teorema 5.2.2 provee una desigualdad de submayorización que resulta ser una extensión para operadores compactos y autoadjuntos de la desigualdad de Zhan. Por otro lado, el Teorema 5.2.2 no puede ser mejorado a una desigualdad entrada a entrada entre valores singulares. En efecto, sean

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \oplus 0_{\ell^2(\mathbb{N})} \quad \text{y} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \oplus 0_{\ell^2(\mathbb{N})} .$$

Entonces $\lambda(A \oplus B) = \big((\dots, 0, -1), (3, 1, 1, 0, \dots) \big)$, de forma que

$$s(A-B) = (2, 2, 0 \dots) \quad \text{y} \quad \mathrm{Spr}^+(A \oplus B) = (4, 1, 1, 0 \dots) ,$$

lo que prueba que $s_2(A-B) > \mathrm{Spr}_2^+(A \oplus B)$.

De forma similar, la desigualdad de submayorización obtenida en la ecuación (5.10) no puede ser mejorada a una desigualdad entrada a entrada. Por ejemplo tomemos

$$A = \begin{pmatrix} 1/2 & 1/2 & 0 & 2 \\ 1/2 & 1/2 & 2 & 0 \\ 0 & 2 & 1/2 & 1/2 \\ 2 & 0 & 1/2 & 1/2 \end{pmatrix} \oplus 0_{\ell^2(\mathbb{N})} .$$

Entonces $\lambda(A) = \big((..., 0, -1, -2), (3, 2, 0, ...) \big)$ y $\mathrm{Spr}^+(A) = (5, 3, 0, ...)$. Tomando el bloque $A_{1,2}$ como $B = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$ tenemos que $s(B) = (2, 2, 0, ..)$ pero

$$4 = 2\,s_2(B) > \mathrm{Spr}_2^+(A) = 3 . \qquad \triangle$$

5.3. Desigualdades de tipo AG y submayorización

Sean $A, B \in K(\mathcal{H})$, La desigualdad Aritmético-Geométrica de Bhatia y Kittaneh establece que :

$$2\,s_i(A\,B^*) \leq s_i(A^*A + B^*B) , \quad \text{para} \quad i \in \mathbb{N} . \tag{5.14}$$

Ahora bien, supongamos que C, $S \in L(\mathcal{H})$ son tales que $C^*C + S^*S \leq I$ y sea $E \in K(\mathcal{H})^+$. Entonces, podemos aplicar la desigualdad AG para concluir que, para $j \in \mathbb{I}_k$

$$s_j(SEC^*) = s_j\big((SE^{1/2})(CE^{1/2})^* \big) \leq \frac{1}{2} s_j \left(E^{1/2}(S^*S + C^*C)E^{1/2} \right) \leq \frac{1}{2} s_j(E) . \tag{5.15}$$

La ecuación (5.15) fue obtenida en [5], donde también se prueba que resulta ser más fuerte que las desigualdades de tipo AG obtenidas en [1], (ver también [21, 23] donde se prueban desigualdades de tipo AG para funciones convexas y para normas unitariamente invariantes). En particular (ver Observación 1.5.7) para cualquier norma unitariamente invariante N tenemos que

$$N(SEC^*) \leq \frac{1}{2} N(E)\,. \tag{5.16}$$

Más aún, de las ecuaciones (5.15) y (5.16) es posible derivar desigualdades de tipo AG más generales. Sin embargo, las ecuaciones (5.15) y (5.16) fallan para operadores autoadjuntos $E \in K(\mathcal{H})^{sa}$ como vemos en el siguiente:

Ejemplo 5.3.1. Como en ejemplos previos, utilizaremos que una matriz $A \in \mathcal{M}_n(\mathbb{C})$ puede ser insertada como un operador de rango finito en $K(\mathcal{H})$ (pensando en el operador $\tilde{A} = A \oplus 0_{\ell^2(\mathbb{N})}$), lo que permite construir contraejemplos fácilmente usando matrices. Ahora bien, vemos que la ecuación (5.16) es falsa si $E \in K(\mathcal{H})^{sa} \setminus K(\mathcal{H})^{+}$. En efecto, consideremos $N(X) = \|X\|_2 = (\mathrm{tr}(X^*X))^{1/2}$ la norma Frobenius (que es unitariamente invariante). Sean

$$S = \begin{bmatrix} \mathrm{sen}(\pi/3) & 0 \\ 0 & \mathrm{sen}(\pi/5) \end{bmatrix},\ E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} \cos(\pi/3) & 0 \\ 0 & \cos(\pi/5) \end{bmatrix}.$$

entonces $C^*C + S^*S = I$ y $\|S\,E\,C^*\|_2 \approx 0{,}7598$ mientras que $\frac{1}{2}\|E\|_2 = \frac{1}{2}$. △

En lo que sigue obtendremos una generalización de la ecuación (5.16) para operadores autoadjuntos $E \in K(\mathcal{H})^{sa}$; estas desigualdades son nuevas, en el sentido de que involucran cotas superiores en términos de la dispersión espectral de operadores autoadjuntos en $K(\mathcal{H})$. Notemos que, si bien estos resultados están basados en relaciones (más débiles) de submayorización, implican desigualdades para normas unitariamente invariantes como en la ecuación (5.16).

Proposición 5.3.2. *Sean C, $S \in L(\mathcal{H})$ tales que $C^*C + S^*S = P$, donde $P = P^2$ y sea $E \in K(\mathcal{H})^{sa}$. Entonces*

$$2\,s(SEC^*) \prec_w \mathrm{Spr}^{+}(E)\,. \tag{5.17}$$

Demostración. Considerando la descomposición polar de $S = U\,|S|$ y $C = V\,|C|$ y la ecuación (1.18), podemos asumir que S, $C \in L(\mathcal{H})^{+}$, pues

$$s_i(S\,E\,C^*) = s_i(U\,|S|\,E\,|C|\,V^*) \leq ||U||\,||V^*||\,s_i(|S|\,E\,|C|) = s_i(|S|\,E\,|C|)\ \forall i \in \mathbb{N}\,,$$

y

$$s_i(|S|\,E\,|C|) = s_i(U^*S\,E\,C^*V) \leq ||U^*||\,||V||\,s_i(S\,E\,C^*) = s_i(S\,E\,C^*)\ \forall i \in \mathbb{N}\,,$$

de forma que $s(S\,E\,C^*) = s(|S|\,E\,|C|)$.

En este caso, notemos que

$$P(S^2 + C^2)P = P = S^2 + C^2 \implies (S^2 - PS^2P) + (C^2 - PC^2P) = 0\,,$$

por lo que $S^2 = PS^2P$ y $C^2 = PC^2P$. Luego tomando $S = (PS^2P)^{1/2}$ y $C = (PC^2P)^{1/2}$ vemos que $C = CP = PCP$ y $S = SP = PSP$. En consecuencia, $S(PEP)C = SEC$ y usando la ecuación (5.7),

$$\mathrm{Spr}^{+}(PEP) \prec_w \mathrm{Spr}^{+}(E)\,.$$

Por lo tanto, podemos restringirnos a $B(R(P))$ y asumir que $P = I$. En este caso si $C^2 + S^2 = I$, obtenemos que C y $S = (I - C^2)^{1/2}$ conmutan. Este hecho implica que $U = \begin{bmatrix} C & -S \\ S & C \end{bmatrix} \in \mathcal{U}(\mathcal{H} \oplus \mathcal{H})$. Más aún,

$$U(E \oplus 0_{\mathcal{H}})U^* = \begin{bmatrix} CEC & CES \\ SEC & SES \end{bmatrix} \in K(\mathcal{H} \oplus \mathcal{H})^{sa}$$

$$\overset{(5.10)}{\Longrightarrow} 2\,s(SEC) \prec_w \mathrm{Spr}^+(U(E\oplus 0_{\mathcal{H}})U^*) = \mathrm{Spr}^+(E\oplus 0_{\mathcal{H}}) \overset{(5.8)}{=} \mathrm{Spr}^+(E)\,,$$

donde hemos usado el Teorema 5.2.1 y la ecuación (5.8) de la Proposición 5.1.5. □

Observación 5.3.3. Con la notación de la Proposición 5.3.2, si $E \in K(\mathcal{H})^+$ entonces $\mathrm{Spr}^+(E) \overset{(3.6)}{=} s(E)$, Luego, la Proposición 5.3.2 implica que $N(SEC) \leq \frac{1}{2}\,N(E)$ y por lo tanto recuperamos la ecuación (5.16).

En caso que $E \in K(\mathcal{H})^{sa} \setminus K(\mathcal{H})^+$ entonces resulta que $s(E) \prec \mathrm{Spr}^+(E)$ de forma estricta, es decir, si N es una norma unitariamente invariante estrictamente convexa entonces

$$N(E) < N(\mathrm{Spr}^+(E))\,.$$

Por ejemplo, si consideramos $N(X) = \|X\|_2 = (\mathrm{tr}(X^*X))^{1/2}$ entonces tenemos que

$$\|E\|_2^2 = \sum_{j\in\mathbb{Z}_0} \lambda_j(E)^2 < \sum_{j\in\mathbb{N}} (\lambda_j(E) - \lambda_{-j}(E))^2 = \|\mathrm{Spr}^+(E)\|_2^2\,,$$

para todo operador autoadjunto de Hilbert-Schmidt $E \notin K(\mathcal{H})^+$. △

En lo que sigue, enunciamos un resultado conocido de R. Douglas [18] que establece un criterio para factorizar operadores que necesitaremos más adelante.

Teorema 5.3.4 (Douglas)**.** Sean $A\,,\ B \in L(\mathcal{H})$. Entonces las siguientes condiciones son equivalentes:

1. $R(A) \subseteq R(B)$.
2. Existe $\lambda \in \mathbb{R}_{\geq 0}$ tal que $AA^* \leq \lambda\, BB^*$.
3. Existe $C \in L(\mathcal{H})$ tal que $A = BC$.

En este caso, existe una única solución reducida

$$C \in B(\mathcal{H}) \quad \text{tal que} \quad A = BC \quad \text{y} \quad R(C) \subseteq \overline{R(B^*)} = \ker B^{\perp}\ . \tag{5.18}$$

□

El siguiente resultado ilustra una desigualdad de tipo AG aplicable a normas unitariamente invariantes.

Teorema 5.3.5. *Sean $A,\ B \in L(\mathcal{H})$ y sea $E \in K(\mathcal{H})^{sa}$. Entonces*

$$s(A\,E\,B^*) \prec_w \frac{1}{2}\ \mathrm{Spr}^+\left(\,(A^*A + B^*B)^{1/2}\,E\,(A^*A + B^*B)^{1/2}\,\right). \tag{5.19}$$

Demostración. Por el Teorema 5.3.4, como $A^*A \leq A^*A + B^*B$ (y B^*B también), las ecuaciones lineales de operadores

$$A = S\,(A^*A + B^*B)^{1/2} \qquad \text{y} \qquad B = C\,(A^*A + B^*B)^{1/2}\,, \tag{5.20}$$

admiten soluciones únicas $S,\ C \in L(\mathcal{H})$ que cumplen

$$\ker(A^*A + B^*B)^{1/2} = R(\,(A^*A + B^*B)^{1/2})^{\perp} \overset{(5.18)}{\subseteq} R(S^*)^{\perp} \cap R(C^*)^{\perp} = \ker(S) \cap \ker(C)\ .$$

En consecuencia, si $z \in R(\,(A^*A + B^*B)^{1/2})$ y $x \in \mathcal{H}$ es tal que $z = (A^*A + B^*B)^{1/2}x$ entonces

$$\begin{aligned}\langle (S^*S + C^*C)z, z\rangle \ &= \|S(A^*A + B^*B)^{1/2}x\|^2 + \|C(A^*A + B^*B)^{1/2}x\|^2 \\ &\overset{(5.20)}{=} \|A\,x\|^2 + \|B\,x\|^2 = \langle (A^*A + B^*B)\,x\,,\,x\rangle = \|z\|^2\ .\end{aligned}$$

Por otro lado, si $z \in \ker(AA^* + BB^*)^{1/2}$ entonces $(S^*S + C^*C)z = 0$. Luego, si consideramos P la proyección ortogonal sobre la clausura de $R((A^*A + B^*B)^{1/2})$, entonces los argumentos previos muestran que $\langle (S^*S + C^*C)z, z\rangle = \langle Pz, z\rangle$, para $z \in \mathcal{H}$; y por lo tanto, $S^*S + C^*C = P$. Más aún

$$AEB^* = S\,(A^*A + B^*B)^{1/2}\,E\,(A^*A + B^*B)^{1/2}\,C^*\,.$$

Ahora aplicando la Proposición 5.3.2 sobre $A\,E\,B^*$ (siguiendo la expresión anterior) obtenemos la ecuación (5.19). □

Ejemplo 5.3.6. Sean A, B y E como en el Teorema 5.3.5, pero supongamos que $E \geq 0$. Como $s\left((A^*A + B^*B)^{1/2}\,E\,(A^*A + B^*B)^{1/2}\right) = s\left(E^{1/2}\,(A^*A + B^*B)\,E^{1/2}\right)$ y $\mathrm{Spr}^+(C) = s(C)$ para $C \in K(\mathcal{H})^+$, entonces la ecuación (5.14) y su consecuente (5.15), que son desigualdades valor singular a valor singular, resultan más fuertes que la ecuación (5.19) en el caso positivo. Más aún, la desigualdad de la ecuación (5.19) es un sustituto de (5.15) en el caso autoadjunto no positivo.

Sin embargo, para el caso autoadjunto, la ecuación (5.19) no puede ser mejorada a una desigualdad valor singular a valor singular como la ecuación (5.15). En efecto, consideramos

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1 \end{pmatrix} \quad \text{y} \quad E = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

Luego, $F = A^*A + B^*B = \begin{pmatrix} \frac{13}{4} & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & \frac{13}{4} \end{pmatrix}$ es tal que

$$\lambda(F^{\frac{1}{2}}E\,F^{\frac{1}{2}}) = \left((\dots, 0, \tfrac{-13}{4}), (\tfrac{39}{4}, 2, 0, \dots)\right) \quad \text{y} \quad \mathrm{Spr}^+(F^{\frac{1}{2}}E\,F^{\frac{1}{2}}) = (13, 2, 0, \dots).$$

Más aún, $s(A\,E\,B^*) \approx (4,74\,,\, 1,58\,,\, 1\,,\, 0....)$, lo que prueba que

$$3,16 \approx 2\,s_2(A\,E\,B^*) > \mathrm{Spr}_2^+(F^{\frac{1}{2}}E\,F^{\frac{1}{2}}) = 2\,. \qquad \triangle$$

El siguiente resultado complementa la Proposición 5.3.2.

Proposición 5.3.7. *Sean C, $S \in L(\mathcal{H})^+$ tales que $C^2 + S^2 = P$, con $P = P^2$ y sean E_1, $E_2 \in K(\mathcal{H})^{sa}$. Entonces*

$$s(SE_1C + CE_2S) \prec_w \frac{1}{2}\,\mathrm{Spr}^+(E_1 \oplus -E_2)\,. \tag{5.21}$$

Demostración. Como en la prueba de la Proposición 5.3.2, podemos asumir que $P = I$ y construir una matriz unitaria por bloques $U = \begin{bmatrix} C & -S \\ S & C \end{bmatrix} \in \mathcal{U}(\mathcal{H} \oplus \mathcal{H})$. Entonces

$$U(E_1 \oplus -E_2)U^* = \begin{bmatrix} CE_1C - SE_2S & CE_1S + SE_2C \\ SE_1C + CE_2S & SE_1S - CE_2C \end{bmatrix} \in K(\mathcal{H} \oplus \mathcal{H})^{sa}$$

$$\overset{(5.10)}{\Longrightarrow} 2\,s(SE_1C + CE_2S) \prec_w \mathrm{Spr}^+(U(E_1 \oplus -E_2)U^*) = \mathrm{Spr}^+(E_1 \oplus -E_2)\,,$$

Donde utilizamos el Teorema 5.2.1. □

Por otro lado, si $E \in K(\mathcal{H})^{sa}$ entonces

$$\mathrm{Spr}^+(E \oplus -E) = \left(2\,|\lambda(E)|\right)^{\downarrow} = 2\;s(E)\;.$$

Usamos este hecho en el siguiente

Corolario 5.3.8. *Sean C, $S \in L(\mathcal{H})^+$ tales que $C^2 + S^2 = P$, con $P = P^2$ y sea $E \in K(\mathcal{H})^{sa}$. Entonces*

$$s(\,\mathrm{Re}(SEC)\,) \prec_w \frac{1}{4}\,\mathrm{Spr}^+(E \oplus -E) = \frac{1}{2}\,s(E)\,. \tag{5.22}$$

Demostración. Notemos que $\mathrm{Re}(SEC) = \frac{SEC+CES}{2}$ y por lo tanto el resultado es consecuencia de la Proposición 5.3.7 con $E_1 = E_2 = E$. □

Terminaremos esta sección recopilando varios resultados equivalentes expuestos anteriormente. Más aún, cada uno de estos puede ser interpretado en términos de desigualdades para normas unitariamente invariantes.

Teorema 5.3.9. *Las siguientes desigualdades son equivalentes: dados E, $F \in K(\mathcal{H})^{sa}$, $X \in K(\mathcal{H})$ y A, B, $P \in L(\mathcal{H})$, con $P = P^* = P^2$.*

1. $2\,s(P\,E\,(I-P)) \prec_w \mathrm{Spr}^+(E)$;
2. $s(A\,E\,B^*) \prec_w \frac{1}{2}\,\mathrm{Spr}^+((A^*A + B^*B)^{1/2}\,E\,(A^*A + B^*B)^{1/2})$;
3. $s(E - F) \prec_w \mathrm{Spr}^+(E \oplus F)$.

Demostración. Inspeccionando la prueba del Teorema 5.3.5 vemos que 1. → 2. (vía la Proposición 5.3.2). Por otro lado, tomando $A = P$ y $B = I - P$ en 2. vemos que $A^*A + B^*B = I$ y recuperamos el ítem 1. Esto prueba la equivalencia entre 1. y 2.

Para probar que 3. → 1.podemos asumir que $\mathcal{H} = \mathcal{K} \oplus \mathcal{K}$ y que $P = P_{\mathcal{K}\oplus\{0\}}$. Entonces

$$E = \begin{bmatrix} E_1 & B \\ B^* & E_2 \end{bmatrix} \begin{matrix} \mathcal{K} \\ \mathcal{K} \end{matrix} \quad \text{con} \quad s(PE(I-P)) = s(B) \in c_0(\mathbb{N})^{\downarrow}\,.$$

Sea $B = U|B|$ la descomposición polar de B. Entonces $U \in B(\mathcal{K})$ es una isometría parcial. Si construimos la isometría parcial $W \in L(\mathcal{H}) = B(\mathcal{K} \oplus \mathcal{K})$ dada por

$$W = \begin{bmatrix} U & 0 \\ 0 & I_{\mathcal{K}} \end{bmatrix} \quad \text{entonces} \quad W^*EW = \begin{bmatrix} U^*E_1U & U^*B \\ B^*U & E_2 \end{bmatrix} = \begin{bmatrix} U^*E_1U & |B| \\ |B| & E_2 \end{bmatrix}\,.$$

Entonces, por el ítem 5. de la Proposición 5.1.5 vemos que $\mathrm{Spr}^+(W^*EW) \prec_w \mathrm{Spr}^+(E)$ y $s(B) = s(|B|)$. Por lo tanto, para probar el ítem 1. podemos asumir que $B \in K(\mathcal{K})^{sa}$. En este caso, tomando el operador unitario y autoadjunto $R = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \in B(\mathcal{K} \oplus \mathcal{K})$, tenemos que

$$RER = \begin{bmatrix} E_2 & B \\ B & E_1 \end{bmatrix} \implies \frac{E + RER}{2} = \begin{bmatrix} \frac{E_1+E_2}{2} & B \\ B & \frac{E_1+E_2}{2} \end{bmatrix}\,.$$

Por el ítem 3. del Teorema 1.5.4 (desigualdad de Weyl) y el ítem 4. de la Proposición 5.1.5

$$\lambda\left(\frac{E + RER}{2}\right) \prec \frac{\lambda(E) + \lambda(RER)}{2} = \lambda(E) \Longrightarrow \mathrm{Spr}^+\left(\frac{E + RER}{2}\right) \prec_w \mathrm{Spr}^+(E)\,.$$

En consecuencia, para probar el ítem 1. podemos asumir que $B \in K(\mathcal{K})^{sa}$ y $E_1 = E_2$.

Ahora bien, tomamos $Z = \frac{1}{\sqrt{2}}\begin{bmatrix} I & I \\ -I & I \end{bmatrix} \in B(\mathcal{K} \oplus \mathcal{K})$ que resulta unitario. Por construcción resulta que

$$E = \begin{bmatrix} E_1 & B \\ B & E_1 \end{bmatrix} \implies Z^*EZ = \begin{bmatrix} E_1 - B & 0 \\ 0 & E_1 + B \end{bmatrix}\,,\ \mathrm{Spr}^+(Z^*EZ) = \mathrm{Spr}^+(E)\,. \tag{5.23}$$

Finalmente, usando el ítem 5. tenemos que

$$2\,s(B) = s\big(\,[E_1 - B] - [E_1 + B]\,\big) \prec_w \mathrm{Spr}^+(\,[E_1 - B] \oplus [E_1 + B]\,) \overset{(5.23)}{=} \mathrm{Spr}^+(E)\,.$$

Esto prueba que 3. → 1.

Por otro lado 1. → 3. fue probado en el Teorema 5.2.2.

Bibliografía

[1] H. Albadawi, Singular value and arithmetic-geometric mean inequalities for operators. Ann. Funct. Anal. 3 (2012), no. 1, 10-18.

[2] E. Andruchow, G. Corach, D. Stojanoff, Geometric operator inequalities. Linear Algebra Appl. 258 (1997), 295-310.

[3] J. Antezana, P. Massey, M. Ruiz, D. Stojanoff, The Schur-Horn theorem for operators and frames with prescribed norms and frame operator. Illinois J. Math. 51 (2007), no. 2, 537-560.

[4] M.E. Argentati, A.V. Knyazev, C.C. Paige, I. Panayotov, Bounds on changes in Ritz values for a perturbed invariant subspace of a Hermitian matrix. SIAM J. Matrix Anal. Appl. 30 (2008), no. 2, 548-559.

[5] W. Audeh, Generalizations for singular value and arithmetic-geometric mean inequalities of operators. J. Math. Anal. Appl. 489 (2020), no. 2, 124184, 8pp.

[6] W. Audeh, F. Kittaneh, Singular value inequalities for compact operators. Linear Algebra Appl. 437 (2012), no. 10, 2516-2522.

[7] R. Bhatia, Matrix analysis, 169, Springer-Verlag, New York, 1997.

[8] R. Bhatia and C. Davis, More matrix form of the arithmetic geometric mean inequality, SIAM J. Matrix Anal. Appl. 14 (1993), 132-136.

[9] R. Bhatia, F. Kittaneh, On the singular values of a product of operators. SIAM J. Matrix Anal. Appl. 11 (1990), no. 2, 272-277.

[10] R. Bhatia, F. Kittaneh, The matrix arithmetic-geometric mean inequality revisited. Linear Algebra Appl. 428 (2008), no. 8-9, 2177-2191.

[11] C. M. Bishop, Pattern recognition and machine learning. Information Science and Statistics. Springer, New York, 2006.

[12] N. Bosner, Z. Drmač, Subspace gap residuals for Rayleigh-Ritz approximations. SIAM J. Matrix Anal. Appl. 31 (2009), no. 1, 54-67.

[13] A. Böttcher, I.M. Spitkovsky, A gentle guide to the basics of two projections theory. Linear Algebra Appl. 432 (2010), no. 6, 1412-1459.

[14] G. Corach, H. Porta, L. Recht, An operator inequality, Linear Algebra Appl. 142 (1990) 153–158.

[15] G. Corach, H. Porta, L. Recht, Geodesics and operator means in the space of positive operators. Internat. J. Math. 4 (1993), no. 2, 193-202.

[16] G. Corach, H. Porta, L. Recht, The geometry of the space of self-adjoint invertible elements in a C^*-algebra. Integral Equations Operator Theory 16 (1993), no. 3, 333-359.

[17] C. Davis, W.M. Kahan, The rotation of eigenvectors by a perturbation. III. SIAM J. Numer. Anal. 7 1970 1-46.

[18] R.G. Douglas, On majorization, factorization, and range inclusion of operators on Hilbert space. Proc. Amer. Math. Soc. 17 (1966), 413-415.

[19] I.C. Gohberg, M.G. Krein, Introduction to the theory of linear nonself-adjoint operators in Hilbert space. Am. Math. Soc., Providence (1969).

[20] P.R. Halmos, Two subspaces. Trans. Amer. Math. Soc. 144 1969 381-389.

[21] O. Hirzallah, Inequalities for sums and products of operators, Linear Algebra Appl. 407 (2005) 32-42.

[22] O. Hirzallah, Commutator inequalities for Hilbert space operators. Linear Algebra Appl. 431 (2009), no. 9, 1571-1578.

[23] O. Hirzallah, Singular values of convex functions of operators and the arithmetic-geometric mean inequality. J. Math. Anal. Appl. 433 (2016), no. 2, 935-947.

[24] O. Hirzallah, F. Kittaneh, Singular values,norms, and commutators. Linear Algebra and its Applications 432 (2010) 1322–1336.

[25] R.A. Horn, C.R: and Johnson, Matrix analysis. Second Edition, Cambridge University Press, Cambridge, 2012.

[26] T. Kato, Perturbation theory for linear operators, Springer-Verlag Berlin Heidelberg New York 1980.

[27] F. Kittaneh, Inequalities for commutators of positive operators. J. Funct. Anal. 250 (2007), no. 1, 132-143.

[28] F. Kittaneh, Norm inequalities for commutators of Hermitian operators. Integral Equations Operator Theory 62 (2008), no. 1, 129-135.

[29] F. Kittaneh, Singular value inequalities for commutators of Hilbert space operators. Linear Algebra Appl. 430 (2009), no. 8-9, 2362-2367.

[30] A.V. Knyazev, M.E. Argentati, Rayleigh-Ritz majorization error bounds with applications to FEM. SIAM J. Matrix Anal. Appl. 31 (2009), no. 3, 1521-1537.

[31] A.V. Knyazev, M.E. Argentati, On proximity of Rayleigh quotients for different vectors and Ritz values generated by different trial subspaces. Linear Algebra Appl. 415 (2006), no. 1, 82-95.

[32] A.V. Knyazev, M.E. Argentati, Majorization for changes in angles between subspaces, Ritz values, and graph Laplacian spectra. SIAM J. Matrix Anal. Appl. 29 (2006/07), no. 1, 15-32.

[33] A.K. Knyazev, P. Zhu, Rayleigh-Ritz majorization error bounds of mixed type. SIAM J. Matrix Anal. Appl. 38 (2017), no. 1, 30-49.

[34] C.-K. Li, R.-C. Li, A note on eigenvalues of perturbed Hermitian matrices. Linear Algebra Appl. 395 (2005), 183-190.

[35] C. K. Li, M. Roy, The Lidskii-Mirsky-Wielandt theorem additive and multiplicative versions, Numerische Mathematik 81 (1999): 377-413.

[36] A.S.Markus, The eigen- and singular values of the sum and product of linear operators, 1964 Russ. Math. Surv. 19 R02

[37] A.W. Marshall, I. Olkin, B.C. Arnold, Inequalities: theory of majorization and its applications. Second edition. Springer Series in Statistics. Springer, New York, 2011.

[38] P Massey, D. Stojanoff, S. Zárate, Majorization bounds for Ritz values of self-adjoint matrices. SIAM J. Matrix Anal. Appl. 41 (2020), no. 2, 554-572.

[39] P. Massey, D. Stojanoff, S. Zárate, The spectral spread of Hermitian matrices, Linear Algebra Appl. 616 (2021), 19-44.

[40] P. Massey, D. Stojanoff, S. Zárate, Absolute variation of Ritz values, principal angles and spectral spread, SIAM J. Matrix Anal. Appl., Vol. 42 (2021) no. 4, 1451-1757.

[41] P. Massey, D. Stojanoff, S. Zárate, Norm inequalities for the spectral spread of Hermitian operators (2021), arXiv:2106.09092v1.

[42] R. Mathias, Quadratic residual bounds for the Hermitian eigenvalue problem. SIAM J. Matrix Anal. Appl. 19 (1998), no. 2, 541-550.

[43] Nakatsukasa, Y., The tan θ theorem with relaxed conditions. Linear Algebra Appl. 436 (2012), no. 5, 1528-1534.

[44] M. A. Nielsen, I.L. Chuang, Quantum computation and quantum information. Cambridge University Press, Cambridge, 2000.

[45] E. Ovtchinnikov, Cluster robust error estimates for the Rayleigh-Ritz approximation. II. Estimates for eigenvalues. Linear Algebra Appl. 415 (2006), no. 1, 188-209.

[46] B.N. Parlett, The symmetric eigenvalue problem. Corrected reprint of the 1980 original, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998.

[47] L. Qiu, Y. Zhang, C.K. Li, Unitarily invariant metrics on the Grassmann space. SIAM J. Matrix Anal. Appl. 27 (2005), no. 2, 507-531.

[48] Y. Saad, Numerical Methods for Large Eigenvalue Problems, Revised Edition, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.

[49] Barry Simon, Trace Ideals and their Applications, London Mathematical Society Lecture Note Series, (1979).

[50] G.W. Stewart, J. Sun, Matrix perturbation theory, Computer Science and Scientific Computing, Academic Press Inc., Boston, MA, 1990.

[51] Y. Tao, More results on singular value inequalities of matrices, Linear Algebra Appl. 416 (2006) 724-729.

[52] Z. Teng, L. Lu, R.-C. Li, Cluster-robust accuracy bounds for Ritz subspaces. Linear Algebra Appl. 480 (2015), 11-26.

[53] Y.Q. Wang, H.K. Du, Norms of commutators of Hermitian operators. J. Math. Anal. Appl. 342 (2008), no. 1, 747-751.

[54] X. Zhan, Singular values of differences of positive semidefinite matrices. SIAM J. Matrix Anal. Appl. 22 (2000), no. 3, 819-823.

[55] X. Zhan, Matrix inequalities. Lecture Notes in Mathematics, 1790. Springer-Verlag, Berlin, 2002.

[56] X. Zhan, On some matrix inequalities, Linear Algebra and its Applications 376 (2004) 299-303.

[57] P. Zhu, Angles between subspaces and the Rayleigh-Ritz method, Ph.D thesis, University of Colorado Denver, 2012.

[58] P. Zhu, M.E. Argentati, A.V. Knyazev, Bounds for the Rayleigh quotient and the spectrum of self-adjoint operators. SIAM J. Matrix Anal. Appl. 34 (2013), no. 1, 244-256.