Universidad Autónoma del Estado de Hidalgo
Instituto de Ciencias Básicas e Ingeniería
Área Académica de Matemáticas y Física

Cercanía de las Matrices Casi Normales al Conjunto de las Matrices Normales

Tesis que para obtener el título de

Licenciado en Matemáticas Aplicadas

presenta

Diana Xochitl Canales Licona

bajo la dirección de

Dr. Rubén A. Martínez Avendaño

Pachuca, Hidalgo. Agosto de 2009.

Resumen

¿Será cierto que matrices que son casi normales están cerca del conjunto de las matrices normales? La respuesta es afirmativa. De hecho, Huaxin Lin [8] mostró que para todo $\epsilon > 0$ existe $\delta > 0$ tal que, para cualquier entero positivo n y cualquier matriz A en el álgebra $\mathcal{M}_n(\mathbb{C})$ de matrices de tamaño $n \times n$ con entradas en los números complejos que satisfaga $\|A^*A - AA^*\| < \delta$, existe una matriz normal N tal que $\|A - N\| < \epsilon$. El principal resultado de este trabajo es probar una versión más débil del teorema de Lin con δ dependiente de n.

Abstract

Are almost normal matrices near normal matrices? The answer is affirmative. In fact, Huaxin Lin [8] showed the following. For any $\epsilon > 0$ there exists $\delta > 0$ such that, for any n and any matrix A in the algebra $\mathcal{M}_n(\mathbb{C})$ of $n \times n$ complex matrices with $\|A^*A - AA^*\| < \delta$, there is a normal matrix N with $\|A - N\| < \epsilon$. The main result in this work is to show a weak version of Lin's theorem with δ chosen dependent of n.

Por su amor infinito, su apoyo incondicional, y por los innumerables sacrificios que han hecho por sus hijos...

a mis padres

Agradecimientos

Por su colosal afecto y apoyo, estoy muy agradecida para con mi amada familia: mis padres, mi hermanita Miriam, mis amados hermanos Rene y Hugo, y mis sobrinitos Joshua e Ivan. Ellos son el principal motor de mi vida.

También, quiero expresar mi inmensa gratitud y respeto hacia los profesores de la LIMA por sus muy valiosas enseñanzas, y por el aliento que nos brindarón en cada desafío. Todos ellos son fuente de inspiración. En particular, quiero agradecer al Dr. Sampedro, pues en sus consejos siempre encontré dirección.

Tengo una deuda especial con el Jurado que revisó este trabajo. Gracias por sus correcciones, comentarios y sugerencias.

Entre mis más profundas deudas, esta la que tengo con el Dr. Rubén por todo su apoyo, por su agradable amistad, y en especial por los gratos momentos profundos y significativos que fueron luz y claridad en mi vida. Estos momentos me alentarón para continuar y además me devolvierón esperanza y dirección. Mi eterna gratitud al Dr. Rubén. Ha sido un gran honor haber realizado este trabajo bajo su dirección.

Finalmente, quiero agradecer a la Universidad Autónoma del Estado de Hidalgo por todo su apoyo. A Viry, Annita, Blanca, Lis y Thalía por haberme permitido hospedar en su casa. A Yola, Fanny, Hortencia, Luis Ángel, Adán, Carlos (el gato), Ceci, Fernando (el mayita), al Dr. Edgar, al Dr. Fede y a mi mejor amigo Miguel Ángel (el Pepo) por su valiosa amistad y por los momentos tan agradables que pasamos a lo largo de la licenciatura. A la Mtra. Olga Rocha por revisar la ortografía de este trabajo. Y por último quiero agradecer al servicio que ofrece el Laboratorio de Matemáticas Computacionales del CIMA y a todas las personas que me apoyarón siempre. A todos gracias.

ÍNDICE GENERAL

Introducción

En 1969 Peter Rosenthal [10] se planteó la siguiente pregunta: ¿será cierto que matrices que casi conmutan están cerca de matrices que conmutan? Más precisamente, se preguntó si para todo $\epsilon > 0$ y para cualesquiera matrices A y B de tamaño $n \times n$ con entradas en los números complejos que satisfagan $\|AB - BA\| < \epsilon$, entonces existen matrices A' y B' que conmutan ($A'B' = B'A'$) y además que están "cerca" (en una manera técnica, la cual no mencionaremos aquí) de A y B. Rosenthal consideraba que el crecimiento de este tipo de cercanía con respecto a n podría ser importante para aplicaciones al estudio de operadores compactos. Por ejemplo, comenta que la existencia de tal cercanía podría implicar (si ésta no crece tan rápido con respecto a n) la existencia de subespacios invariantes comunes para ciertas clases de operadores compactos.

En 1976, P. R. Halmos [5] no sólo retoma esta pregunta al establecer problemas de aproximación para operadores, sino que además plantea el problema concreto para matrices que casi conmutan con su adjunta de la siguiente manera: ¿será cierto que para todo $\epsilon > 0$ existe un $\delta > 0$ tal que, para cualquier matriz A de tamaño $n \times n$ que satisfaga $\|A\| < 1$ y $\|A^*A - AA^*\| < \delta$, la distancia de A al conjunto de las matrices normales es menor que ϵ? Recuérdese que una matriz A se dice normal si conmuta con su adjunta, es decir, si $A^*A = AA^*$. Este último problema estuvo abierto por muchos años, durante los cuales fue estudiado extensamente desde el punto de vista de la teoría de operadores y de las álgebras de operadores, hasta que Huaxin Lin [8] en 1995 lo resolvió afirmativamente. Lin demostró además que δ se puede escoger independientemente de la dimensión de la matriz. La importancia de obtener este δ independiente de la dimensión, estriba en el hecho de que este resultado tiene aplicaciones en el problema de si un operador normal se puede aproximar por un operador normal con espectro finito. La prueba de Lin es más bien técnica y necesita fuertemente de la teoría de álgebras C^*. En 1996 Peter Friis y Mikael Rørdam [4] simplificaron y extendieron la prueba de Lin, y en 1999 Friis [3] refina su prueba para dar una nueva demostración del Teorema de Brown-Douglas-Fillmore sobre operadores esencialmente normales.

El motivo que nos condujo a escribir este trabajo fue preguntarnos si podríamos encontrar una expresión explícita para δ (aunque ésta dependa de la dimensión de la matriz).

Para iniciar con nuestra busqueda, en el Capítulo 1 introduciremos algunos conceptos básicos necesarios para comenzar a abordar el problema. También estableceremos el escenario donde se desarrollará nuestro análisis del problema, el espacio vectorial de las matrices de tamaño $n \times n$ con entradas en los números complejos $\mathcal{M}_n(\mathbb{C})$. En este espacio definiremos la norma de Frobenius y la norma usual de operadores, y mostraremos algunas de sus propiedades que serán útiles a lo largo de este trabajo. Además veremos que todas las normas definidas en $\mathcal{M}_n(\mathbb{C})$ son equivalentes. Finalizaremos este capítulo presentando al conjunto de las matrices normales y mostrando que es cerrado con la métrica dada por normas $\|\cdot\|_1$ que cumplen $\|X^*\|_1 = \|X\|_1$ para todo $X \in \mathcal{M}_n(\mathbb{C})$.

En el Capítulo 2 calcularemos la distancia explícita de una matriz arbitraria de tamaño 2×2 al conjunto de las matrices normales $\mathcal{N}$ con las métricas dadas por la norma de Frobenius y la norma usual de operadores. Para esto, primero encontraremos la distancia de ciertas matrices de tamaño 2×2 al conjunto de las matrices normales.

En general, encontrar la distancia de una matriz de dimensión mayor a dos al conjunto de las matrices normales, resulta ser un cálculo complicado la mayor parte de las veces. Por tal razón, en el Capítulo 3 presentaremos un resultado de Henrici [6], el cual va a ser de utilidad pues nos proporcionará una cota superior para esta distancia con la métrica dada, principalmente, por la norma de Frobenius.

En el Capítulo 4 concluiremos este trabajo, primero notando que de una observación de Higham [7] se puede ver que “las matrices *casi* hermitianas están *cerca* de las matrices hermitianas”, y generalizando este hecho a un tipo de matrices que llamaremos λ-hermitianas. Luego, veremos que “las matrices *casi* normales están *cerca* del conjunto de las matrices normales”.

Al final de este trabajo presentamos en un apéndice las condiciones que deben satisfacer las matrices B y C para que $\langle X, Y \rangle_1 := \text{traza}(Y^* BXC)$ defina un producto interno y para que la norma $\|\cdot\|_1 := \sqrt{\langle \cdot, \cdot \rangle_1}$ cumpla $\|X^*\|_1 = \|X\|_1$ para todo $X \in \mathcal{M}_n(\mathbb{C})$.

CAPÍTULO 1

Operadores en $\mathbb{C}^n$

En la Sección 1.1 introduciremos algunos conceptos básicos y estableceremos el escenario donde se desarrollará nuestro análisis del problema, el espacio vectorial $\mathcal{M}_n(\mathbb{C})$. En este espacio definiremos la norma de Frobenius (Sección 1.2) y la norma usual de operadores (Sección 1.3), y mostraremos algunas de sus propiedades que serán de utilidad a lo largo de este trabajo. Además, veremos que todas las normas definidas en $\mathcal{M}_n(\mathbb{C})$ son equivalentes (Sección 1.4). Por último, en la Sección 1.5 presentaremos al conjunto de las matrices normales y veremos que es cerrado con la métrica dada por ciertas normas definidas en $\mathcal{M}_n(\mathbb{C})$.

1.1

Operadores

Las siguientes definiciones son bien conocidas y las presentamos aquí para justificar notación.

Definición 1.1. Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

una matriz de tamaño $n \times m$. Definimos la *matriz transpuesta de A* como la matriz de tamaño $m \times n$

$$A^T := \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix},$$

esto es, la matriz cuya i-ésima fila es la i-ésima columna de A.

Definición 1.2. Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}.$$

Definimos la *matriz conjugada de A* como

$$\overline{A} := \begin{pmatrix} \overline{a}_{11} & \overline{a}_{12} & \cdots & \overline{a}_{1m} \\ \overline{a}_{21} & \overline{a}_{22} & \cdots & \overline{a}_{2m} \\ \vdots & \vdots & & \vdots \\ \overline{a}_{n1} & \overline{a}_{n2} & \cdots & \overline{a}_{nm} \end{pmatrix},$$

esto es, la matriz cuyas entradas son las mismas de A conjugadas.

La siguiente matriz surge al combinar estas definiciones.

Definición 1.3. Sea A una matriz de tamaño $n \times m$. Definimos la *matriz adjunta de A* como $A^* := \overline{A^T}$.

Se puede verificar que (con la suma, producto escalar y producto usual de matrices) la operación “tomar adjunta” cumple las siguientes propiedades:

1. $(A^*)^* = A$.
2. $(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} B^*$.
3. $(AB)^* = B^* A^*$.

En este trabajo vamos a llamar *vectores* a las matrices que constan de una columna. Al conjunto de vectores de tamaño $n \times 1$ lo vamos a denotar por $\mathbb{C}^n$ y lo consideraremos como espacio vectorial sobre $\mathbb{C}$ con la suma y producto por escalar usual de vectores.

Definición 1.4. Definimos un *operador* como una función $T : \mathbb{C}^n \to \mathbb{C}^n$ tal que para todo $x, y \in \mathbb{C}^n$ y $\alpha, \beta \in \mathbb{C}$ se satisface $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$.

Como es bien conocido, dada una base de $\mathbb{C}^n$ podemos representar al operador mediante una matriz de tamaño $n \times n$ (ver [1, pág. 56-57]), y si la base está fija, cada matriz $n \times n$ representa a un operador (ver [1, pág. 134]). Así que a partir de aquí trabajaremos con la base canónica de $\mathbb{C}^n$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

e identificaremos a los operadores con matrices de tamaño $n \times n$.

Al conjunto que consta de las matrices de tamaño $n \times n$ lo vamos a denotar por $\mathcal{M}_n(\mathbb{C})$ y lo consideraremos como espacio vectorial sobre $\mathbb{C}$ con la suma y producto por escalar usual de matrices.

Definición 1.5. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Decimos que A es una matriz *triangular superior* si es de la forma:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Definición 1.6. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Decimos que A es una *matriz diagonal* si es de la forma:

$$\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Definición 1.7. Definimos la *matriz identidad*, la cual denotaremos por I, como la matriz diagonal tal que cada entrada de la diagonal principal es 1.

Definición 1.8. Sea $\mathcal{V}$ un espacio vectorial sobre $\mathbb{C}$. Un *producto interno* en $\mathcal{V}$ es una función $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{C}$ que satisface las siguientes propiedades:

- $\langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$, para todo $u, v, w \in \mathcal{V}$ y $\alpha, \beta \in \mathbb{C}$.
- $\langle v, w \rangle = \overline{\langle w, v \rangle}$, para todo $v, w \in \mathcal{V}$.
- $\langle v, v \rangle \geq 0$, para todo $v \in \mathcal{V}$; y $\langle v, v \rangle = 0$ si y sólo si $v = 0$.

En este trabajo usaremos el producto interno en $\mathbb{C}^n$ que se obtiene de la siguiente proposición.

Proposición 1.9. *Sean $x, y \in \mathbb{C}^n$, digamos $x = (x_1, ..., x_n)^T$ e $y = (y_1, ..., y_n)^T$. Entonces la función:*

$$\langle x, y\rangle := x_1\bar{y_1} + x_2\bar{y_2} + \cdots + x_n\bar{y_n} = (\bar{y_1}, \bar{y_2}, ..., \bar{y_n}) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = y^*x,$$

define un producto interno en $\mathbb{C}^n$.

Demostración. Ver [1, pág. 18]. ∎

El siguiente lema será de utilidad más adelante.

Lema 1.10. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ y sean $x, y \in \mathbb{C}^n$. Entonces $\langle Ax, y\rangle = \langle x, A^*y\rangle$.*

Demostración. Por definición $\langle x, A^*y\rangle = (A^*y)^*x$. Luego, por la propiedad 1 y 3 de matrices adjuntas se cumple que $(A^*y)^*x = y^*(A^*)^*x = y^*Ax$. Esta última igualdad es precisamente la definición de $\langle Ax, y\rangle$, como queríamos. ∎

Definición 1.11. Sea $\mathcal{V}$ un espacio vectorial sobre $\mathbb{C}$. Una *norma* en $\mathcal{V}$ es una función $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ que satisface las siguientes propiedades:

- $\|v\| \geq 0$, para todo $v \in \mathcal{V}$; y $\|v\| = 0$ si y sólo si $v = 0$.
- $\|\alpha v\| = |\alpha|\|v\|$, para todo $v \in \mathcal{V}$ y $\alpha \in \mathbb{C}$.
- $\|v + w\| \leq \|v\| + \|w\|$, para todo $v, w \in \mathcal{V}$ (desigualdad del triángulo).

Sea $\mathcal{V}$ un espacio vectorial sobre $\mathbb{C}$. Se demuestra (ver [2, pág. 3]) que para cada producto interno $\langle \cdot, \cdot \rangle$ en $\mathcal{V}$, podemos definir una norma en $\mathcal{V}$, a saber: $\|v\| := \sqrt{\langle v, v\rangle}$, para todo $v \in \mathcal{V}$.

Definición 1.12. Llamaremos *norma euclidiana* a la norma dada por el producto interno en $\mathbb{C}^n$, esto es: $\|x\| := \sqrt{\langle x, x\rangle} = \sqrt{x^*x}$, para todo $x \in \mathbb{C}^n$.

Proposición 1.13 (Desigualdad de Cauchy-Schwarz)**.** *Sean $x, y \in \mathbb{C}^n$. Entonces*

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\|,$$

con igualdad si y sólo si x e y son linealmente dependientes (ver [1, pág. 19]).

Definición 1.14. Decimos que $U \in \mathcal{M}_n(\mathbb{C})$ es una matriz *unitaria*, si $U^*U = I$.

De la definición se sigue que $UU^* = I$ y además que U es biyectiva.

Lema 1.15. *Sea $U \in \mathcal{M}_n(\mathbb{C})$ una matriz unitaria. Entonces, para todo $x \in \mathbb{C}^n$ se cumple $\|Ux\| = \|x\|$.*

Demostración. Por definición $\|Ux\|^2 = \langle Ux, Ux\rangle$. Luego, por el Lema 1.10 y como U es unitaria $\langle Ux, Ux\rangle = \langle x, U^*Ux\rangle = \langle x, Ix\rangle = \|x\|^2$. ∎

Este lema establece que la norma del vector que surgue al aplicarle una matriz unitaria a un vector cualquiera, es igual a la norma del vector original. Por esta razón, las matrices unitarias también son llamadas *isometrías*.

El siguiente teorema será de gran utilidad a lo largo de este trabajo.

Teorema 1.16 (Teorema de Schur)**.** *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces existe una matriz unitaria U tal que U^*AU es una matriz triangular superior.*

Demostración. Procederemos por inducción sobre el tamaño de la matriz.

Si $n = 1$, claramente A es una matriz triangular superior.

Si $n > 1$, supondremos que el teorema se cumple para matrices de tamaño $(n-1)\times(n-1)$. Sea λ_1 un eigenvalor de A, y sea x_1 un eigenvector de A correspondiente a λ_1 con $\|x_1\| = 1$. Entonces usando el algoritmo de Gram-Schmidt [1, pág. 27] podemos obtener vectores $x_2, ..., x_n$ tales que $\{x_1, x_2, ..., x_n\}$ es una base ortonormal de $\mathbb{C}^n$ (recordemos que un conjunto de vectores son ortonormales si cada uno tiene norma 1 y el producto interno entre dos cualesquiera de ellos es cero). Así, colocando estos vectores como columnas en una matriz $U_1 = [x_1\ x_2\ \cdots\ x_n]$ tenemos

$$
\begin{aligned}
U_1^*AU_1 &= \begin{pmatrix} —x_1^*— \\ —x_2^*— \\ \vdots \\ —x_n^*— \end{pmatrix} A \begin{pmatrix} | & | & & | \\ x_1 & x_1 & \cdots & x_1 \\ | & | & & | \end{pmatrix} \\
&= \begin{pmatrix} x_1^*Ax_1 & x_1^*Ax_2 & \cdots & x_1^*Ax_n \\ x_2^*Ax_1 & x_2^*Ax_2 & \cdots & x_2^*Ax_n \\ \vdots & \vdots & & \vdots \\ x_n^*Ax_1 & x_n^*Ax_2 & \cdots & x_n^*Ax_n \end{pmatrix},
\end{aligned}
$$

y dado que $Ax_1 = \lambda_1 x_1$, se sigue que:

$$
\begin{aligned}
U_1^* A U_1 &= \begin{pmatrix} x_1^*\lambda_1 x_1 & x_1^* A x_2 & \cdots & x_1^* A x_n \\ x_2^*\lambda_1 x_1 & x_2^* A x_2 & \cdots & x_2^* A x_n \\ \vdots & \vdots & & \vdots \\ x_n^*\lambda_1 x_1 & x_n^* A x_2 & \cdots & x_n^* A x_n \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 \|x_1\|^2 & x_1^* A x_2 & \cdots & x_1^* A x_n \\ \lambda_1 \langle x_1, x_2 \rangle & x_2^* A x_2 & \cdots & x_2^* A x_n \\ \vdots & \vdots & & \vdots \\ \lambda_1 \langle x_1, x_n \rangle & x_n^* A x_2 & \cdots & x_n^* A x_n \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & x_1^* A x_2 & \cdots & x_1^* A x_n \\ 0 & x_2^* A x_2 & \cdots & x_2^* A x_n \\ \vdots & \vdots & & \vdots \\ 0 & x_n^* A x_2 & \cdots & x_n^* A x_n \end{pmatrix},
\end{aligned}
$$

pues $\{x_1, x_2, ..., x_n\}$ es una base ortonormal. De esta manera podemos escribir

$$
U_1^* A U_1 = \begin{pmatrix} \lambda_1 & Y_1 \\ 0 & A_1 \end{pmatrix}
$$

en forma de bloque. Luego, dado que A_1 es una matriz de tamaño $(n-1)\times(n-1)$ sabemos, por hipótesis de inducción, que existe una matriz unitaria W de tamaño $(n-1)\times(n-1)$, tal que $T_1 := W^* A_1 W$ es una matriz triangular superior. Así, definiendo la matriz unitaria $U_2 = \begin{pmatrix} 1 & 0 \\ 0 & W \end{pmatrix}$ en forma de bloque, la matriz $U = U_1 U_2$ es una matriz unitaria, y además cumple

$$
\begin{aligned}
U^* A U &= (U_1 U_2)^* A (U_1 U_2) \\
&= U_2^* (U_1^* A U_1) U_2 \\
&= \begin{pmatrix} 1 & 0 \\ 0 & W^* \end{pmatrix} \begin{pmatrix} \lambda_1 & Y_1 \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & W \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & Y_1 W \\ 0 & W^* A W \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & Y_1 W \\ 0 & T_1 \end{pmatrix}.
\end{aligned}
$$

Esto muestra que $U^* A U$ es una matriz triangular superior como se afirmó. ■

Notemos que las entradas de la diagonal principal de la matriz triangular del Teorema de Schur son los eigenvalores de A incluyendo las posibles multiplicidades.

Definición 1.17. Decimos que $A \in \mathcal{M}_n(\mathbb{C})$ es una matriz *normal*, si $A^* A = A A^*$.

El siguiente lema será útil.

Lema 1.18. *Sea $T \in \mathcal{M}_n(\mathbb{C})$ una matriz triangular superior. Entonces T es una matriz normal si y sólo si T es una matriz diagonal.*

Demostración. Con un cálculo podemos ver que si T es diagonal, entonces T es normal.

Ahora supongamos que T es una matriz normal de la forma

$$T = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ 0 & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & m_{nn} \end{pmatrix}.$$

Las entradas de la diagonal de $T^*T - TT^*$ cumplen

$$\begin{aligned} & \left[\begin{pmatrix} \bar{m}_{11} & 0 & \cdots & 0 & \cdots & 0 \\ \bar{m}_{12} & \bar{m}_{22} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ \bar{m}_{1i} & \bar{m}_{2i} & \cdots & \bar{m}_{ii} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ \bar{m}_{1n} & \bar{m}_{2n} & \cdots & \bar{m}_{in} & \cdots & \bar{m}_{nn} \end{pmatrix} \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1i} & \cdots & m_{1n} \\ 0 & m_{22} & \cdots & m_{2i} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & m_{ii} & \cdots & m_{in} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & m_{nn} \end{pmatrix} \right]_{ii} \\ - & \left[\begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1i} & \cdots & m_{1n} \\ 0 & m_{22} & \cdots & m_{2i} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & m_{ii} & \cdots & m_{in} \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & m_{nn} \end{pmatrix} \begin{pmatrix} \bar{m}_{11} & 0 & \cdots & 0 & \cdots & 0 \\ \bar{m}_{12} & \bar{m}_{22} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ \bar{m}_{1i} & \bar{m}_{2i} & \cdots & \bar{m}_{ii} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ \bar{m}_{1n} & \bar{m}_{2n} & \cdots & \bar{m}_{in} & \cdots & \bar{m}_{nn} \end{pmatrix} \right]_{ii} \\ = & \sum_{k=1}^{i} |m_{ki}|^2 - \sum_{k=i}^{n} |m_{ik}|^2. \end{aligned}$$

Luego, como por hipótesis T es una matriz normal, $T^*T - TT^* = 0$, entonces

$$0 = [T^*T - TT^*]_{11} = -(|m_{12}|^2 + |m_{13}|^2 + |m_{14}|^2 + \cdots + |m_{1n}|^2),$$

lo cual ocurre si y sólo si $|m_{1k}| = 0$, para todo $k \in \{2, 3, ..., n\}$. Por lo que entonces

$$\begin{aligned} 0 &= [T^*T - TT^*]_{22} \\ &= |m_{12}|^2 - (|m_{23}|^2 + |m_{24}|^2 + \cdots + |m_{2n}|^2) \\ &= -(|m_{23}|^2 + |m_{24}|^2 + \cdots + |m_{2n}|^2), \end{aligned}$$

pero esto ocurre si y sólo si $|m_{2k}| = 0$, para todo $k \in \{3, 4, ..., n\}$. Pero entonces

$$\begin{aligned} 0 &= [T^*T - TT^*]_{33} \\ &= |m_{13}|^2 + |m_{23}|^2 - (|m_{34}|^2 + \cdots + |m_{3n}|^2) \\ &= -(|m_{34}|^2 + \cdots + |m_{3n}|^2). \end{aligned}$$

De igual manera, esto ocurre si y sólo si $|m_{3k}| = 0$, para todo $k \in \{4, ..., n\}$.

Continuando con este proceso podemos concluir que T, es en efecto, una matriz diagonal. ∎

El Teorema de Schur aplicado a matrices normales se llama Teorema Espectral.

Corolario 1.19 (Teorema Espectral). *Sea $A \in \mathcal{M}_n(\mathbb{C})$ una matriz normal. Entonces existe una matriz unitaria U tal que U^*AU es una matriz diagonal.*

Demostración. Por el Teorema de Schur sabemos que existe una matriz unitaria U tal que $T := U^*AU$ es una matriz triangular superior.

Luego, de la hipótesis sobre A, se tiene:

$$\begin{aligned} T^*T &= (U^*AU)^*(U^*AU) \\ &= U^*A^*UU^*AU \\ &= U^*A^*AU \\ &= U^*AA^*U \\ &= U^*AUU^*A^*U \\ &= (U^*AU)(U^*AU)^* \\ &= TT^*. \end{aligned}$$

Es decir, la matriz triangular superior T es una matriz normal. Por lo tanto, aplicando el Lema 1.18, se tiene que T es diagonal. ∎

La siguiente definición será útil.

Definición 1.20. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Definimos la *traza de A* como la suma de los elementos de la diagonal principal de A.

El siguiente lema muestra algunas propiedades de la traza.

Lema 1.21. *Sean $A, B \in \mathcal{M}_n(\mathbb{C})$ y $\alpha, \beta \in \mathbb{C}$. Entonces*

1. $\operatorname{traza}(\alpha A + \beta B) = \alpha \operatorname{traza}(A) + \beta \operatorname{traza}(B)$.

2. $\operatorname{traza}(AB) = \operatorname{traza}(BA)$.

3. $\operatorname{traza}(A^*) = \overline{\operatorname{traza}(A)}$.

Demostración. Con un cálculo se puede verificar que la parte 1 se cumple.

Para verificar la parte 2 notemos que las entradas de la diagonal de AB son

$$\left[\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2i} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ii} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{ni} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1i} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2i} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{ii} & \cdots & b_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{ni} & \cdots & b_{nn} \end{pmatrix}\right]_{ii}$$
$$= \sum_{k=1}^{n} a_{ik} b_{ki}.$$

Análogamente las entradas de la diagonal de la matriz BA son $\sum_{k=1}^{n} b_{ik}a_{ki}$. Así, traza$(BA) = \sum_{i=1}^{n}\sum_{k=1}^{n} b_{ik}a_{ki} = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ki}b_{ik} =$ traza(AB).

Para verificar la propiedad 3 primero notemos que traza$(A^T) =$ traza(A), por lo tanto traza$(A^*) =$ traza$(\bar{A}^T) =$ traza$(\bar{A}) = \overline{\text{traza}(A)}$. ∎

Definición 1.22. Sean $A, B \in \mathcal{M}_n(\mathbb{C})$. Decimos que *$A$ y B son unitariamente equivalentes*, si existe una matriz unitaria U tal que $A = U^*BU$.

Se puede verificar que dos matrices unitariamente equivalentes tienen la misma traza. Más aún, tenemos el siguiente corolario.

Corolario 1.23. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ y sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ sus eigenvalores incluyendo las posibles multiplicidades. Entonces* traza$\,(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

Demostración. Por el Teorema de Schur sabemos que existe U unitaria tal que U^*AU es triangular, y además las entradas de la diagonal principal son los eigenvalores de A incluyendo las posibles multiplicidades. Entonces $\lambda_1 + \lambda_2 + \cdots + \lambda_n =$ traza$\,(U^*AU)$. Luego, por la propiedad 2 del Lema 1.21 se cumple que traza$\,(U^*(AU)) =$ traza$\,((AU)U^*) =$ traza$\,(A)$. ∎

Definición 1.24. Decimos que la norma $\|\cdot\|$ definida en $\mathcal{M}_n(\mathbb{C})$ es *unitariamente invariante* si para toda U unitaria, se cumple $\|U^*AU\| = \|A\|$ para todo $A \in \mathcal{M}_n(\mathbb{C})$.

En las siguientes dos secciones presentaremos normas unitariamente invariantes definidas en $\mathcal{M}_n(\mathbb{C})$.

1.2

Norma de Frobenius

Iniciaremos esta sección con el siguiente lema.

Lema 1.25. *Sea $A \in \mathcal{M}_n(\mathbb{C})$, digamos*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Entonces $\text{traza}(A^*A) = \sum_{ij} |a_{ij}|^2$.

Demostración. Como

$$\begin{aligned} [A^*A]_{jj} &= \left[\begin{pmatrix} \bar{a}_{11} & \bar{a}_{21} & \cdots & \bar{a}_{j1} & \cdots & \bar{a}_{n1} \\ \bar{a}_{12} & \bar{a}_{22} & \cdots & \bar{a}_{j2} & \cdots & \bar{a}_{n2} \\ \vdots & \vdots & & \vdots & & \vdots \\ \bar{a}_{1j} & \bar{a}_{2j} & \cdots & \bar{a}_{jj} & \cdots & \bar{a}_{nj} \\ \vdots & \vdots & & \vdots & & \vdots \\ \bar{a}_{1n} & \bar{a}_{2n} & \cdots & \bar{a}_{jn} & \cdots & \bar{a}_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jj} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix} \right]_{jj} \\ &= \sum_{i=1}^{n} |a_{ij}|^2, \end{aligned}$$

tenemos que $\text{traza}(A^*A) = \sum_{j=1}^{n} \sum_{i=1}^{n} |a_{ij}|^2 = \sum_{ij} |a_{ij}|^2$. ∎

Lema 1.26. *Sea $\langle A, B \rangle := \text{traza}(B^*A)$, para $A, B \in \mathcal{M}_n(\mathbb{C})$. Entonces $\langle \cdot, \cdot \rangle$ define un producto interno en $\mathcal{M}_n(\mathbb{C})$.*

Demostración. Verificaremos que las condiciones de la Definición 1.8 se satisfacen.

Sean $A, B, C \in \mathcal{M}_n(\mathbb{C})$ y $\alpha, \beta \in \mathbb{C}$. Entonces por el Lema 1.21 se cumple

$$\begin{aligned} \langle \alpha A + \beta B, C \rangle &= \text{traza}(C^*(\alpha A + \beta B)) \\ &= \text{traza}(\alpha C^* A + \beta C^* B) \\ &= \alpha \, \text{traza}(C^* A) + \beta \, \text{traza}(C^* B) \\ &= \alpha \langle A, C \rangle + \beta \langle B, C \rangle . \end{aligned}$$

Por la propiedad 3 del Lema 1.21 y las propiedades 3 y 1 de matrices adjuntas $\overline{\langle B, A \rangle} = \overline{\text{traza}(A^*B)} = \text{traza}((A^*B)^*) = \text{traza}(B^*A) = \langle A, B \rangle$.

Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces $\langle A, A\rangle = \text{traza}(A^*A) = \sum_{ij} |a_{ij}|^2$ (Lema 1.25). Por lo tanto $\langle A, A\rangle \geq 0$ con igualdad si y sólo si $A = 0$. ∎

Definición 1.27. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Definimos la *norma de Frobenius (o euclidiana) de* A como $\|A\|_F := \sqrt{\sum_{ij} |a_{ij}|^2}$.

Así, por el Lema 1.25 y el Lema 1.26 tenemos $\|A\|_F = \sqrt{\text{traza}(A^*A)} = \sqrt{\langle A, A\rangle}$. Esto implica que la norma de Frobenius es una norma. En lo que resta de esta sección mostraremos algunas de sus propiedades.

Lema 1.28. *Sea* $A \in \mathcal{M}_n(\mathbb{C})$. *Entonces* $\|A^*\|_F = \|A\|_F$.

Demostración. Es inmediato de la definición de la norma de Frobenius. ∎

Proposición 1.29. *Sean* $A, B \in \mathcal{M}_n(\mathbb{C})$. *Entonces* $\|AB\|_F \leq \|A\|_F\|B\|_F$.

Demostración. Sean

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \text{ y } B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix},$$

y sean $a_i := (\bar{a}_{i1}, \bar{a}_{i2}, \cdots, \bar{a}_{in})^T$ y $b_i := (b_{1i}, b_{2i}, \cdots, b_{ni})^T$ para todo $i \in \{1, 2, ..., n\}$.

Entonces por la desigualdad de Cauchy-Schwarz

$$\begin{aligned}
\|AB\|_F^2 &= \left\| \begin{pmatrix} \langle b_1, a_1\rangle & \langle b_2, a_1\rangle & \cdots & \langle b_n, a_1\rangle \\ \langle b_1, a_2\rangle & \langle b_2, a_2\rangle & \cdots & \langle b_n, a_2\rangle \\ \vdots & \vdots & & \vdots \\ \langle b_1, a_n\rangle & \langle b_2, a_n\rangle & \cdots & \langle b_n, a_n\rangle \end{pmatrix} \right\|_F^2 \\
&= \sum_{ij} | \langle b_j, a_i\rangle |^2 \\
&\leq \sum_{ij} \|b_j\|^2 \|a_i\|^2 \\
&= \sum_i \|a_i\|^2 \sum_j \|b_j\|^2 \\
&= \sum_{ij} |\bar{a}_{ij}|^2 \sum_{ij} |b_{ij}|^2 \\
&= \sum_{ij} |a_{ij}|^2 \sum_{ij} |b_{ij}|^2 \\
&= \|A\|_F^2 \|B\|_F^2,
\end{aligned}$$

como queríamos. ∎

La igualdad en esta proposición no siempre ocurre. A continuación presentamos un ejemplo.

Ejemplo. Sean $A = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$. Entonces $\|BA\|_F = \left\| \begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix} \right\|_F = \sqrt{28}$ y $\|A\|_F \|B\|_F = \sqrt{6}\sqrt{6} = 6$. En este caso, la desigualdad es estricta.

Proposición 1.30. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces $\|A^*A\|_F \leq \|A\|_F^2$.*

Demostración. El resultado se sigue de la Proposición 1.29 y del Lema 1.28. ∎

Notemos que en el ejemplo anterior $B = A^*$, lo cual hace de éste también un ejemplo para ilustrar que la igualdad de esta proposición no siempre ocurre.

Lema 1.31. *Sea $U \in \mathcal{M}_n(\mathbb{C})$ una matriz unitaria. Entonces $\|U\|_F = \sqrt{n}$.*

Demostración. Se sigue de que $\|U\|_F^2 = \langle U, U\rangle = \text{traza}(U^*U) = \text{traza}(I) = n$. ∎

La siguiente proposición muestra que la norma de Frobenius es unitariamente invariante.

Proposición 1.32. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ y sea U una matriz unitaria. Entonces $\|U^*AU\|_F = \|A\|_F$.*

Demostración. Se sigue de $\|U^*AU\|_F^2 = \langle U^*AU, U^*AU\rangle = \text{traza}((U^*AU)^*(U^*AU)) = \text{traza}(U^*A^*UU^*AU) = \text{traza}(U^*A^*AU) = \text{traza}(A^*AUU^*) = \text{traza}(A^*A) = \|A\|_F^2$. ■

Lema 1.33. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ y sea $x \in \mathbb{C}^n$. Entonces $\|Ax\| \leq \|A\|_F\|x\|$, con igualdad si y sólo si el rango de A es uno y $\bar{x}$ está en el espacio fila de A.*

Demostración. Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

y sea $a_i = (a_{i1}, a_{i2}, \cdots, a_{in})^T$, para todo $i \in \{1, 2, ..., n\}$. Entonces por la desigualdad de Cauchy-Schwarz

$$\begin{aligned}
\|Ax\|^2 &= \left\| \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right\|^2 \\
&= \left\| \begin{pmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \vdots \\ x_1a_{n1} + x_2a_{n2} + \cdots + x_na_{nn} \end{pmatrix} \right\|^2 \\
&= \left\| \begin{pmatrix} \langle a_1, \bar{x}\rangle \\ \langle a_2, \bar{x}\rangle \\ \vdots \\ \langle a_n, \bar{x}\rangle \end{pmatrix} \right\|^2 \\
&= |\langle a_1, \bar{x}\rangle|^2 + |\langle a_2, \bar{x}\rangle|^2 + \cdots + |\langle a_n, \bar{x}\rangle|^2 \\
&\leq \|a_1\|^2\|\bar{x}\|^2 + \|a_2\|^2\|\bar{x}\|^2 + \cdots + \|a_n\|^2\|\bar{x}\|^2 \\
&= (\|a_1\|^2 + \|a_2\|^2 + \cdots + \|a_n\|^2)\|x\|^2 \\
&= \|A\|_F^2\|x\|^2,
\end{aligned}$$

con igualdad si y sólo si $\bar{x}$ es linealmente dependiente a cada fila de A, es decir, si $\bar{x} = \alpha_1 a_1 = \alpha_2 a_2 = \cdots = \alpha_n a_n$, para algunas constantes α_i, con $i \in \{1, 2, ..., n\}$. Esto implica que cada fila de A es un múltiplo de $\bar{x}$. Sin embargo, esto ocurre si y sólo si el rango de A es uno y $\bar{x}$ está en el espacio fila de A. ■

1.3

Norma Usual de Operadores

Sea $A \in \mathcal{M}_n(\mathbb{C})$. Dado que el operador A y la norma euclidiana son funciones continuas, la imagen del conjunto compacto $\{x \in \mathbb{C}^n : \|x\| \leq 1\}$, bajo la composición de A y la norma euclidiana, es compacto. Es decir, el conjunto $\{\|Ax\| : \|x\| \leq 1\}$ es compacto, y por lo tanto cerrado y acotado. Esto implica que el $\sup\{\|Ax\| : \|x\| \leq 1\}$ es un número finito y se alcanza. A continuación justificaremos cómo el $\sup\{\|Ax\| : \|x\| \leq 1\}$ equipa a $\mathcal{M}_n(\mathbb{C})$ con una norma.

Definición 1.34. Sea $A \in \mathcal{M}_n(\mathbb{C})$. Definimos la *norma usual del operador A* como $\|A\| := \sup\{\|Ax\| : \|x\| \leq 1\}$.

A diferencia de la norma de Frobenius, la norma usual de operadores no proviene de un producto interno (justificaremos esto al final de la sección). Sin embargo, sí es una norma.

Proposición 1.35. *La norma usual de operadores es una norma en $\mathcal{M}_n(\mathbb{C})$.*

Demostración. Verificaremos que las condiciones de la Definición 1.11 se cumplen.

Claramente, $\|A\| = \sup\{\|Ax\| : \|x\| \leq 1\} \geq 0$ para todo $A \in \mathcal{M}_n(\mathbb{C})$. Notemos que si $\|A\| = 0$, entonces $Ax = 0$ para todo $x \in \{y \in \mathbb{C}^n : \|y\| \leq 1\}$. Luego, por linealidad $Ax = 0$, para todo $x \in \mathbb{C}^n$. Esto nos permite concluir que $A = 0$. Así, $\|A\| = 0$ si y sólo si $A = 0$.

Observemos que para todo $A \in \mathcal{M}_n(\mathbb{C})$ y $\alpha \in \mathbb{C}$

$$\begin{aligned}
\|\alpha A\| &= \sup\{\|(\alpha A)x\| : \|x\| \leq 1\} \\
&= \sup\{|\alpha|\|Ax\| : \|x\| \leq 1\} \\
&= |\alpha| \sup\{\|Ax\| : \|x\| \leq 1\} \\
&= |\alpha|\|A\|.
\end{aligned}$$

Para demostrar que la norma usual de operadores satisface la desigualdad del triángulo primero notemos que, para todo $A, B \in \mathcal{M}_n(\mathbb{C})$, por la desigualdad del triángulo de la norma euclidiana en $\mathbb{C}^n$:

$$\begin{aligned}
\|A + B\| &= \sup\{\|(A + B)x\| : \|x\| \leq 1\} \\
&= \sup\{\|Ax + Bx\| : \|x\| \leq 1\} \\
&\leq \sup\{\|Ax\| + \|Bx\| : \|x\| \leq 1\}.
\end{aligned}$$

Además,

$$\sup\{\|Ax\| + \|Bx\| : \|x\| \leq 1\} \leq \sup\{\|Ax\| : \|x\| \leq 1\} + \sup\{\|Bx\| : \|x\| \leq 1\}.$$

Esto prueba que $\|A+B\| \leq \|A\|+\|B\|$. ■

A continuación mostraremos algunas propiedades que, al igual que la norma de Frobenius, satisface la norma usual de operadores. Comencemos notando que la norma usual de operadores satisface la primera parte del Lema 1.33.

Lema 1.36. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces $\|Ay\| \leq \|A\| \cdot \|y\|$ para todo $y \in \mathbb{C}^n$.*

Demostración. En el caso $y=0$ la desigualdad es trivial. Por otro lado, si $y \neq 0$ basta notar que $\left\|A\left(\frac{y}{\|y\|}\right)\right\| \leq \sup\{\|Ax\| : \|x\| \leq 1\} = \|A\|$. ■

Lema 1.37. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces existe $x_0 \in \mathbb{C}^n$, con $\|x_0\| = 1$ tal que $\|A\| = \|Ax_0\|$.*

Demostración. Primero supongamos $A \neq 0$. Como mencionamos al inicio de esta sección, el $\sup\{\|Ax\| : \|x\| \leq 1\}$ se alcanza, es decir, existe un $x_0 \in \mathbb{C}^n$ con $\|x_0\| \leq 1$ tal que $\|A\| = \|Ax_0\|$. Luego, por el lema anterior $\|A\| \leq \|A\| \cdot \|x_0\|$, y por lo tanto $1 \leq \|x_0\|$. Así, $\|x_0\| = 1$. Si $A = 0$, el resultado es trivial. ■

Lema 1.38. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces $\|A^*\| = \|A\|$.*

Demostración. Iniciaremos probando que $\|A\| \leq \|A^*\|$. Por el Lema 1.10 y por la desigualdad de Cauchy-Schwarz, $\|Ax\|^2 = \langle Ax, Ax\rangle = \langle x, A^*Ax\rangle \leq \|x\| \cdot \|A^*Ax\|$. Además, $\|A^*Ax\| \leq \|A^*\| \cdot \|Ax\|$ (Lema 1.36). Así, $\|Ax\|^2 \leq \|x\| \cdot \|A^*\| \cdot \|Ax\|$ y por lo tanto $\|Ax\| \leq \|x\| \cdot \|A^*\|$. Esto implica que $\|A\| = \sup\{\|Ax\| : \|x\| \leq 1\} \leq \sup\{\|x\| \cdot \|A^*\| : \|x\| \leq 1\} = \|A^*\| \sup\{\|x\| : \|x\| \leq 1\} = \|A^*\|$.

Para probar que $\|A^*\| \leq \|A\|$, basta aplicar el argumento anterior a A^*. ■

Proposición 1.39. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces $\|A^*A\| = \|A\|^2$.*

Demostración. Primero mostraremos que $\|A\|^2 \leq \|A^*A\|$. Por el Lema 1.10 y por la desigualdad de Cauchy-Schwarz, $\|Ax\|^2 = \langle Ax, Ax\rangle = \langle x, A^*Ax\rangle \leq \|x\| \cdot \|A^*Ax\|$. Además, $\|A^*Ax\| \leq \|A^*A\| \cdot \|x\|$ (Lema 1.36). Así, $\|Ax\|^2 \leq \|x\|^2 \cdot \|A^*A\|$. Esto implica que $\|A\|^2 = \sup\{\|Ax\|^2 : \|x\| \leq 1\} \leq \sup\{\|x\|^2 \cdot \|A^*A\| : \|x\| \leq 1\} = \|A^*A\| \sup\{\|x\|^2 : \|x\| \leq 1\} = \|A^*A\|$.

Para probar $\|A^*A\| \leq \|A\|^2$, notemos que $\|A^*Ax\| \leq \|A^*\| \cdot \|Ax\|$ (Lema 1.36). Por lo tanto $\|A^*A\| = \sup\{\|A^*Ax\| : \|x\| \leq 1\} \leq \sup\{\|A^*\| \cdot \|Ax\| : \|x\| \leq 1\} = \|A^*\| \sup\{\|Ax\| : \|x\| \leq 1\} = \|A^*\|\|A\|$. El resultado se sigue del Lema 1.38. ■

Lema 1.40. *Sea $U \in \mathcal{M}_n(\mathbb{C})$ una matriz unitaria. Entonces $\|U\| = 1$.*

Demostración. Por el Lema 1.15 se cumple $\|U\| = \sup\{\|Ux\| : \|x\| \leq 1\} = \sup\{\|x\| : \|x\| \leq 1\} = 1$. ∎

Al igual que la norma de Frobenius, la norma usual de operadores es unitariamente invariante.

Proposición 1.41. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ y sea U unitaria. Entonces $\|U^*AU\| = \|A\|$.*

Demostración. Iniciaremos probando que $\|AU\| = \|A\|$. Como U es suprayectiva se cumple que para todo $x \in \mathbb{C}^n$ existe $y \in \mathbb{C}^n$ tal que $x = Uy$. Así, $\|x\| = \|Uy\|$. Luego, como U es una isometría (Lema 1.15) tenemos $\|x\| = \|y\|$, y por lo tanto $\|A\| = \{\|Ax\| : \|x\| \leq 1\} = \{\|AUy\| : \|y\| \leq 1\} = \|AU\|$.

Ahora probaremos que $\|U^*B\| = \|B\|$, para todo $B \in \mathcal{M}_n(\mathbb{C})$. Aplicando lo que acabamos de probar a B^*, tenemos $\|B^*U\| = \|B^*\|$. Luego, por el Lema 1.38 el lado izquierdo de esta igualdad cumple $\|B^*U\| = \|(B^*U)^*\| = \|U^*B\|$ y el lado derecho es $\|B\|$. Así, $\|U^*B\| = \|B\|$.

Esto implica que $\|U^*(AU)\| = \|AU\| = \|A\|$ como queríamos. ∎

Sea $A \in \mathcal{M}_n(\mathbb{C})$. El operador A "estira" cualquier vector $x \in \mathbb{C}^n$ a una longitud no mayor que $\|A\| \cdot \|x\|$, pues $\|Ax\| \leq \|A\| \cdot \|x\|$ (Lema 1.36). Esto indica que $\|A\|$ puede ser pensada como el factor más grande por el cual A "estira" cualquier vector. Más aún, se tiene la siguiente proposición.

Proposición 1.42. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces*

$$\|A\| = \inf\{k \geq 0 : \|Ax\| \leq k\|x\|, \forall x \in \mathbb{C}^n\}.$$

Demostración. Llamemos Ψ al conjunto $\{k \geq 0 : \|Ax\| \leq k\|x\|, \forall x\}$, el cual es no vacío, pues $\|A\| \in \Psi$ (Lema 1.36), y además es acotado inferiormente. Denotemos por $\alpha \geq 0$ a este ínfimo. Esto implica que $\alpha \leq \|A\|$.

Para demostrar la igualdad entre α y $\|A\|$, basta mostrar que $\alpha \geq \|A\|$. Procederemos por contradicción. Supongamos $\alpha < \|A\|$. Entonces existe $c \in \Psi$ tal que $\alpha \leq c < \|A\|$. Por el Lema 1.37 sabemos que existe $x_0 \in \mathbb{C}^n$, con $\|x_0\| = 1$ tal que $\|A\| = \|Ax_0\|$ y como c cumple $\|Ax_0\| \leq c\|x_0\| = c$, entonces $\|A\| \leq c$. Sin embargo, esto contradice que $c < \|A\|$. ∎

Calcular la norma usual de un operador ya sea con la primera definición que dimos o con esta nueva manera de enunciarla, no siempre es fácil ya que el cálculo puede volverse complicado y tedioso. El siguiente teorema nos proporciona otra alternativa para calcular la norma usual de un operador dado.

Definición 1.43. Decimos que $A \in \mathcal{M}_n(\mathbb{C})$ es una *matriz autoadjunta o hermi-*

tiana, si $A^* = A$.

Teorema 1.44. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ una matriz autoadjunta. Entonces*

$$\|A\| = \text{máx}\{|\lambda| : \ \lambda \textit{ es eigenvalor de } A\}.$$

Demostración. Primero notemos que como A es una matriz autoadjunta y por lo tanto normal, el Teorema Espectral afirma que existe una matriz unitaria U tal que U^*AU es diagonal. Además, las entradas de la diagonal de U^*AU son los eigenvalores de A incluyendo las posibles multiplicidades, digamos $\lambda_1, \lambda_2, \ldots, \lambda_n$. Así, por la Proposición 1.41 se cumple que

$$\begin{aligned}
\|A\| &= \|U^*AU\| \\
&= \sup\{\|U^*AUx\| : \|x\| \leq 1\} \\
&= \sup\left\{\left\|\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right\| : \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} \leq 1\right\} \\
&= \sup\left\{\sqrt{|\lambda_1|^2|x_1|^2 + |\lambda_2|^2|x_2|^2 + \cdots + |\lambda_n|^2|x_n|^2} : \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} \leq 1\right\}.
\end{aligned}$$

Sin pérdida de generalidad supondremos $|\lambda_1| = \text{máx}\{|\lambda| : \ \lambda \text{ es eigenvalor de } A\}$. Entonces

$$\sqrt{|\lambda_1|^2|x_1|^2 + |\lambda_2|^2|x_2|^2 + \cdots + |\lambda_n|^2|x_n|^2} \leq |\lambda_1|\sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2},$$

lo cual implica

$$\sup\left\{\sqrt{|\lambda_1|^2|x_1|^2 + |\lambda_2|^2|x_2|^2 + \cdots + |\lambda_n|^2|x_n|^2} : \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} \leq 1\right\} \leq |\lambda_1|,$$

equivalentemente $\|A\| \leq \text{máx}\{|\lambda| : \ \lambda \text{ es eigenvalor de } A\}$.

Por otro lado, observemos que

$$|\lambda_i| \in \left\{\sqrt{|\lambda_1|^2|x_1|^2 + |\lambda_2|^2|x_2|^2 + \cdots + |\lambda_n|^2|x_n|^2} : \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} \leq 1\right\},$$

pues $|\lambda_i| = \sqrt{|\lambda_1|^2 \cdot 0^2 + \cdots + |\lambda_i|^2 \cdot 1^2 + \cdots + |\lambda_n|^2 \cdot 0^2}$, para todo $i \in \{1, 2, \ldots, n\}$. Por lo tanto $\text{máx}\{|\lambda| : \lambda \text{ es eigenvalor de } A\} \leq \|A\|$. Esto muestra lo deseado. ∎

Corolario 1.45. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces*

$$\|A\| = \sqrt{\text{máx}\{|\lambda| : \ \lambda \textit{ es eigenvalor de } A^*A\}}.$$

Demostración. Como A^*A es autoadjunta y $\|A\| = \sqrt{\|A^*A\|}$ (Proposición 1.39), por el Teorema 1.44 tenemos $\|A\| = \sqrt{\text{máx}\{|\lambda| : \ \lambda \text{ es eigenvalor de } A^*A\}}$. ∎

Vamos a terminar esta sección observando que la norma usual de operadores no proviene de un producto interno, pues no satisface la ley del paralelogramo (ver [11, pág. 80]). El siguiente ejemplo muestra esto para $n = 2$.

Ejemplo. Sean $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Entonces

$$\begin{aligned}
\|A+B\|^2 &= \left\| \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right\|^2 \\
&= \left\| \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right\| \\
&= \left\| \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \right\| \\
&= \text{máx}\left\{ |\lambda| : \ \lambda \text{ es eigenvalor de } \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \right\} \\
&= \frac{3+\sqrt{5}}{2},
\end{aligned}$$

y

$$\begin{aligned}
\|A-B\|^2 &= \left\| \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \right\|^2 \\
&= \left\| \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \right\| \\
&= \left\| \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \right\| \\
&= \text{máx}\left\{ |\lambda| : \ \lambda \text{ es eigenvalor de } \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \right\} \\
&= \frac{3+\sqrt{5}}{2}.
\end{aligned}$$

Luego, como $\|A\| = 1 = \|B\|$ la ley del paralelogramo no se cumple.

En el caso $n > 2$ se puede verificar que, de manera análoga al caso $n = 2$, la norma usual de operadores no proviene de un producto interno.

1.4

Equivalencia de Normas

Iniciaremos notando que por la Proposición 1.42 y el Lema 1.33 tenemos $\|A\| \leq \|A\|_F$. En la siguiente proposición veremos cuando ocurre la igualdad.

Proposición 1.46. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces $\|A\| = \|A\|_F$ si y sólo si el rango de A es uno.*

Demostración. Supongamos que el rango de A es uno y sea $x \neq 0$ tal que $\bar{x}$ está en el espacio fila de A. Entonces $\|A\|_F\|x\| = \|Ax\|$ (Proposición 1.33) y además $\|Ax\| \leq \|A\| \cdot \|x\|$ (Lema 1.36). Esto implica que $\|A\|_F \leq \|A\|$ y por lo tanto $\|A\| = \|A\|_F$.

Ahora supongamos $\|A\| = \|A\|_F$. Como existe $x_0 \in \mathbb{C}^n$ con $\|x_0\| = 1$ tal que $\|Ax_0\| = \|A\|$ (Lema 1.37) tenemos $\|Ax_0\| = \|A\|_F = \|A\|_F\|x_0\|$, pero por el Lema 1.33 esto ocurre sólo si el rango de A es uno. ∎

Como se verá a continuación, esta relación que existe entre la norma usual de operadores y la norma de Frobenius no es casualidad.

Definición 1.47. Sean $\|\cdot\|_1$ y $\|\cdot\|_2$ normas definidas en el espacio vectorial $\mathcal{V}$. Decimos que $\|\cdot\|_1$ y $\|\cdot\|_2$ son *equivalentes*, si existen $c_1, c_2 > 0$ tales que $c_1\|v\|_1 \leq \|v\|_2 \leq c_2\|v\|_1$ para todo $v \in V$.

La siguiente proposición muestra que la norma usual de operadores y la norma de Frobenius son equivalentes.

Proposición 1.48. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces $\|A\| \leq \|A\|_F \leq \sqrt{n}\|A\|$.*

Demostración. Como observamos al inicio de esta sección, se cumple $\|A\| \leq \|A\|_F$.

Para probar la desigualdad $\|A\|_F \leq \sqrt{n}\|A\|$, notemos que $\|A\|_F^2 = \text{traza}(A^*A)$, y que si λ_1,λ_2,. . .,λ_n son los eigenvalores de A^*A, incluyendo las posibles multiplicidades, entonces por los Corolarios 1.23 y 1.45 tenemos que $\text{traza}(A^*A) = \lambda_1 + \cdots + \lambda_n \leq |\lambda_1| + \cdots + |\lambda_n| \leq n \max\{|\lambda_1|, \ldots, |\lambda_n|\} = n\|A^*A\| = n\|A\|^2$. ∎

Observemos que de la Proposición 1.46 podemos concluir que la desigualdad $\|A\| \leq \|A\|_F$ es óptima, es decir, no existe mejor constante $c_1 > 1$ tal que $c_1\|A\| \leq \|A\|_F$ para toda $A \in \mathcal{M}_n(\mathbb{C})$. Por los Lemas 1.31 y 1.40 podemos ver que la desigualdad $\|A\|_F \leq \sqrt{n}\|A\|$ también es óptima, es decir, no existe mejor constante $c_2 < \sqrt{n}$ tal que $\|A\|_F \leq c_2\|A\|$ para toda $A \in \mathcal{M}_n(\mathbb{C})$.

En lo que resta de esta sección mostraremos que en realidad todas las normas definidas en espacios vectoriales con dimensión finita, y en partícular en $\mathcal{M}_n(\mathbb{C})$, son equivalentes.

Lema 1.49. *Sea $\mathcal{V}$ un espacio vectorial. Entonces todas las normas definidas en $\mathcal{V}$ son funciones continuas.*

Demostración. Sea $\|\cdot\|$ una norma definida en $\mathcal{V}$. Dado $\epsilon > 0$, basta elegir $0 < \delta \leq \epsilon$. Pues si $\|v - w\| < \delta$ entonces, como $|(|v\| - \|w\|)| \leq \|v - w\|$, tenemos $|(|v\| - \|w\|)| < \epsilon$. ∎

Notación 1.50. Sea $\mathcal{V}$ un espacio vectorial y sea $\|\cdot\|$ una norma definida en $\mathcal{V}$. Denotamos a la *bola cerrada unitaria con la métrica dada por* $\|\cdot\|$ como $\mathcal{D} := \{v \in \mathcal{V} : \|v\| \leq 1\}$.

Lema 1.51. *Sea $\mathcal{V}$ un espacio vectorial de dimensión finita. Entonces la bola unitaria cerrada con la métrica dada por cualquier norma es un conjunto compacto.*

Demostración. El resultado se sigue del Teorema de Heine-Borel (ver [11, pág. 44]), pues por hipótesis $\mathcal{V}$ es de dimensión finita y además la bola cerrada unitaria es cerrada y acotada. ∎

Teorema 1.52. *Sea $\mathcal{V}$ un espacio vectorial de dimensión finita. Entonces todas las normas definidas en $\mathcal{V}$ son equivalentes.*

Demostración. Sean $\|\cdot\|_1$ y $\|\cdot\|_2$ normas definidas en $\mathcal{V}$. Llamaremos $S_{\|\cdot\|_2}$ al conjunto $\{v \in \mathcal{V} : \|v\|_2 = 1\}$. Como $\|\cdot\|_1$ es una función continua (Lema 1.49) y $S_{\|\cdot\|_2}$ es un conjunto compacto (pues $S_{\|\cdot\|_2}$ es un subconjunto cerrado de $\mathcal{D} = \{v \in \mathcal{V} : \|v\|_2 \leq 1\}$, el cual es compacto por el Lema 1.51) la norma $\|\cdot\|_1$ alcanza su mínimo y su máximo en $S_{\|\cdot\|_2}$. Sean $c_1 = \text{mín}\left\{\|w\|_1 : w \in S_{\|\cdot\|_2}\right\}$ y $c_2 = \text{máx}\left\{\|w\|_1 : w \in S_{\|\cdot\|_2}\right\}$. Entonces, para $v \in \mathcal{V}$ distinto de cero $\frac{v}{\|v\|_2} \in S_{\|\cdot\|_2}$ y

$$
\begin{aligned}
\|v\|_1 &= \|v\|_2 \left\| \left(\frac{v}{\|v\|_2} \right) \right\|_1 \geq \|v\|_2 \, \text{mín}\left\{\|w\|_1 : w \in S_{\|\cdot\|_2}\right\} = \|v\|_2 c_1 \\
\|v\|_1 &= \|v\|_2 \left\| \left(\frac{v}{\|v\|_2} \right) \right\|_1 \leq \|v\|_2 \, \text{máx}\left\{\|w\|_1 : w \in S_{\|\cdot\|_2}\right\} = \|v\|_2 c_2,
\end{aligned}
$$

es decir, $\|v\|_2 c_1 \leq \|v\|_1 \leq \|v\|_2 c_2$. En el caso $v = 0$, la desigualdad es trivial. ∎

Notemos que el teorema anterior incluye la Proposición 1.48, pues $\mathcal{M}_n(\mathbb{C})$ es un espacio vectorial de dimensión finita. La ventaja de nuestra proposición es que nos proporciona explícitamente las constantes c_1 y c_2 de la Definición 1.47. Estas constantes serán de utilidad en lo que sigue.

1.5

Conjunto de las Matrices Normales $\mathcal{N}$

Definición 1.53. Sea $\|\cdot\|$ una norma definida en el espacio vectorial $\mathcal{V}$ y sea $\mathcal{W}$ un subconjunto de $\mathcal{V}$. Para cada $v \in \mathcal{V}$ definimos la *distancia de v a $\mathcal{W}$ con la métrica dada por la norma* $\|\cdot\|$ como $\text{dist}(v, \mathcal{W}) := \text{ínf}\{\|v - w\| : w \in \mathcal{W}\}$.

En general, si $\mathcal{V}$ es de dimensión finita y $\mathcal{W}$ es cerrado, entonces la distancia de cualquier elemento fuera de $\mathcal{W}$ tiene distancia positiva a $\mathcal{W}$. Veamos esto.

Lema 1.54. *Sea $\|\cdot\|$ una norma definida en el espacio vectorial de dimensión finita $\mathcal{V}$, y sea $\mathcal{W}$ un subconjunto cerrado de $\mathcal{V}$. Entonces para toda $v \notin \mathcal{W}$ se cumple que* $\operatorname{dist}(v, \mathcal{W}) = \operatorname{ínf}\{\|v - w\| : w \in \mathcal{W}\}$ *es positiva y se alcanza.*

Demostración. Notemos que si $\operatorname{dist}(v, \mathcal{W}) = 0$, entonces existe una sucesión $\{w_k\}$ en el conjunto $\mathcal{W}$ tal que $\|v - w_k\| \to 0$ cuando $k \to \infty$. Esto implica que la sucesión $\{w_k\}$ converge a v, pero esto es una contradicción pues por hipótesis $\mathcal{W}$ es cerrado y $v \notin \mathcal{W}$. Así que $\operatorname{dist}(v, \mathcal{W}) > 0$.

Ahora mostremos que la distancia de v a $\mathcal{W}$ se alcanza. Sea $r := \operatorname{dist}(v, \mathcal{W})$ y sea $\{w_n\}$ una sucesión de $\mathcal{W}$ tal que $\|v - w_n\| \to r$. Entonces para todo $\epsilon > 0$ existe un número natural n_0 tal que para $n \geq n_0$ se cumple que $|\|v - w_n\| - r| < \epsilon$. Esto implica que la sucesión $\{w_n\}$, la cual está acotada en un espacio de dimensión finita, tiene una subsucesión convergente. Luego, como por hipotésis $\mathcal{W}$ es un conjunto cerrado con la métrica dada por la norma $\|\cdot\|$, tenemos que esta subsucesión converge a algún elemento de $\mathcal{W}$, mostrando así lo deseado. ∎

Notación 1.55. Sea $\|\cdot\|$ una norma definida en el espacio vectorial $\mathcal{V}$. Para cada $v \in \mathcal{V}$ y $r > 0$ denotamos a la *bola abierta centrada en v de radio r con la métrica dada por la norma* $\|\cdot\|$, como $\mathcal{B}_r(v) := \{w \in \mathcal{V} : \|v - w\| < r\}$.

Lema 1.56. *Sea $\|\cdot\|$ una norma definida en $\mathcal{M}_n(\mathbb{C})$ tal que $\|X^*\| = \|X\|$ para todo $X \in \mathcal{M}_n(\mathbb{C})$. Entonces la función $\phi : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C})$ definida como $\phi(A) := A^*A - AA^*$ es continua con la métrica dada por la norma $\|\cdot\|$.*

Demostración. Primero probaremos que $\varphi(A) := A^*$ es una función continua con la métrica dada por la norma $\|\cdot\|$, es decir, que la imagen inversa de bolas abiertas es un conjunto abierto. Sea $r > 0$. Por la propiedad 2 de matrices autoadjuntas $\varphi^{-1}(\mathcal{B}_r(A^*)) := \{B \in \mathcal{M}_n(\mathbb{C}) : \|A^* - B^*\| < r\} = \{B \in \mathcal{M}_n(\mathbb{C}) : \|(A - B)^*\| < r\}$ y como por hipótesis $\|(A - B)^*\| = \|A - B\|$ entonces $\varphi^{-1}(\mathcal{B}_r(A^*)) = \mathcal{B}_r(A)$.

El hecho de que ϕ es una función continua se sigue de notar que es la suma de funciones continuas. ∎

Denotemos por $\mathcal{N}$ al conjunto de matrices normales.

Teorema 1.57. *Sea $\|\cdot\|$ una norma definida en $\mathcal{M}_n(\mathbb{C})$ tal que $\|X^*\| = \|X\|$ para todo $X \in \mathcal{M}_n(\mathbb{C})$. Entonces el conjunto de las matrices normales $\mathcal{N}$ es cerrado con la métrica dada por la norma $\|\cdot\|$.*

Demostración. Basta mostrar que toda sucesión en $\mathcal{N}$ que converge, converge a un elemento de $\mathcal{N}$. Sea $\{A_n\}_{n=1}^{\infty}$ una sucesión de matrices normales tal que $\lim_{n\to\infty} A_n = A$, y sea ϕ como en el Lema 1.56. Entonces ϕ es una función continua con la métrica dada por la norma $\|\cdot\|$ y

$$\begin{aligned} 0 &= \lim_{n\to\infty}(A_n^*A_n - A_nA_n^*) \\ &= \lim_{n\to\infty}\phi(A_n) \\ &= \phi(\lim_{n\to\infty} A_n) \\ &= \phi(A) \\ &= A^*A - AA^*, \end{aligned}$$

es decir, A es una matriz normal. ∎

Notemos que, por el Lema 1.38 y el Lema 1.28, el teorema anterior indica que $\mathcal{N}$ es un conjunto cerrado con la métrica dada por la norma usual de operadores y la norma de Frobenius. Así, por el Lema 1.54, para cada $A \notin \mathcal{N}$ existen $N_1, N_2 \in \mathcal{N}$ tales que $\operatorname{dist}(A,\mathcal{N}) = \operatorname{dist}(A, N_1)$ y $\operatorname{dist}_F(A,\mathcal{N}) = \operatorname{dist}_F(A, N_2)$.

Lema 1.58. *Sea $\|\cdot\|$ una norma unitariamente invariante definida en $\mathcal{M}_n(\mathbb{C})$. Entonces para cada $A \in \mathcal{M}_n(\mathbb{C})$ existe una matriz triangular superior T que satisface* $\operatorname{dist}_{\|\cdot\|}(A,\mathcal{N}) = \operatorname{dist}_{\|\cdot\|}(T,\mathcal{N})$.

Demostración. Aplicando el Teorema de Schur 1.16 y dado que la norma es unitariamente invariante se obtiene el resultado. ∎

Notemos que, como la norma usual de operadores y la norma de Frobenius son unitariamente invariantes (Proposición 1.41 y 1.32), el lema anterior indica que para toda $A \in \mathcal{M}_n(\mathbb{C})$ existe una matriz triangular superior T tal que $\operatorname{dist}(A,\mathcal{N}) = \operatorname{dist}(T,\mathcal{N})$ y $\operatorname{dist}_F(A,\mathcal{N}) = \operatorname{dist}_F(T,\mathcal{N})$.

En general, calcular la distancia de $\mathcal{N}$ a un elemento $A \in \mathcal{M}_n(\mathbb{C})$ es una tarea difícil. En el siguiente capítulo abordaremos el caso para las matrices 2×2.

CAPÍTULO 2

Distancia de Matrices 2×2 a $\mathcal{N}$

En este capítulo, calcularemos la distancia de una matriz 2×2 arbitraria al conjunto de las matrices normales $\mathcal{N}$ con las métricas dadas por la norma de Frobenius y la norma usual de operadores.

Para obtener este resultado, primero calcularemos la distacia de la matriz $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ al conjunto de las matrices diagonales, y luego al conjunto de las matrices normales de la forma $\begin{pmatrix} 0 & x \\ y & 0 \end{pmatrix}$. Este último es un cálculo importante, pues en este conjunto están las matrices normales para las cuales se alcanza la distacia de $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ a $\mathcal{N}$, tanto con la norma de Frobenius (Sección 2.1) como con la norma usual de operadores (Sección 2.2). Al final de cada una de estas secciones calcularemos la distacia de $\begin{pmatrix} \lambda & 1 \\ 0 & \mu \end{pmatrix}$ a $\mathcal{N}$. En la Sección 2.3 obtendremos explícitamente la distancia de una matriz arbitraria de tamaño 2×2 a $\mathcal{N}$ con la métrica dada por ambas normas.

2.1

Distancia Con la Norma de Frobenius

El siguiente lema será útil.

Lema 2.1. *Sea $x \in \mathbb{C}^n$. Entonces $|x-1|^2 + |x|^2 \geq \frac{1}{2}$, con igualdad si y sólo si $x = \frac{1}{2}$.*

Demostración. El resultado se sigue de observar que

$$\begin{aligned} |x|^2 + |x-1|^2 &= |x|^2 + |x|^2 + 1 - 2\,\mathrm{Re}(x) \\ &= 2\,\mathrm{Re}^2(x) + 2\,\mathrm{Im}^2(x) + 1 - 2\,\mathrm{Re}(x) \\ &\geq 2\,\mathrm{Re}^2(x) - 2\,\mathrm{Re}(x) + 1, \end{aligned}$$

y que la función real $f(u) = 2u^2 - 2u + 1$ tiene un mínimo global en $u = \frac{1}{2}$. ■

Iniciaremos calculando la distancia de la matriz $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ a un subconjunto de $\mathcal{N}$, el conjunto de las matrices diagonales.

Proposición 2.2. *La distancia con la norma de Frobenius de* $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ *al conjunto de las matrices diagonales es 1. Además, esta distancia se alcanza para la matriz cero.*

Demostración. El resultado se sigue de notar que

$$\left\| \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} w & 0 \\ 0 & z \end{pmatrix} \right\|_F = \left\| \begin{pmatrix} w & -1 \\ 0 & z \end{pmatrix} \right\|_F = \sqrt{|w|^2 + |z|^2 + 1}.$$

■

En la siguiente proposición calcularemos la distancia de $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ al conjunto de las matrices normales de la forma $\begin{pmatrix} 0 & x \\ y & 0 \end{pmatrix}$.

Proposición 2.3. *La distancia con la norma de Frobenius de* $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ *al conjunto de las matrices normales de la forma* $\begin{pmatrix} 0 & x \\ y & 0 \end{pmatrix}$ *es* $\frac{1}{\sqrt{2}}$. *Además, esta distancia se alcanza para las matrices* $\begin{pmatrix} 0 & \frac{1}{2} \\ y & 0 \end{pmatrix}$ *donde* $|y| = \frac{1}{2}$.

Demostración. El resultado se sigue de observar que toda matriz normal de la forma $\begin{pmatrix} 0 & x \\ y & 0 \end{pmatrix}$ cumple $|x| = |y|$, y que entonces por el Lema 2.1

$$\left\| \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & x \\ y & 0 \end{pmatrix} \right\|_F^2 = \left\| \begin{pmatrix} 0 & x-1 \\ y & 0 \end{pmatrix} \right\|_F^2 = |x-1|^2 + |y|^2 = |x-1|^2 + |x|^2 \geq \frac{1}{2},$$

con igualdad si y sólo si $x = \frac{1}{2}$ y por lo tanto $|y| = \frac{1}{2}$. ■

Como veremos a continuación, las matrices para las cuales se alcanza la distancia en la proposición anterior son las mismas para las cuales se alcanza la distancia de $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ a $\mathcal{N}$.

Proposición 2.4. *La distancia con la norma de Frobenius de* $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ *al conjunto de las matrices normales* $\mathcal{N}$ *es* $\frac{1}{\sqrt{2}}$. *Además, esta distancia se alcanza para las matrices normales* $\begin{pmatrix} 0 & \frac{1}{2} \\ y & 0 \end{pmatrix}$ *donde* $|y| = \frac{1}{2}$.

Demostración. Basta notar que toda matriz normal $\begin{pmatrix} w & x \\ y & z \end{pmatrix}$ cumple $|x| = |y|$, y que entonces por el Lema 2.1

$$\begin{aligned} \left\| \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} w & x \\ y & z \end{pmatrix} \right\|_F &= \left\| \begin{pmatrix} w & x-1 \\ y & z \end{pmatrix} \right\|_F \\ &= \sqrt{|w|^2 + |x-1|^2 + |y|^2 + |z|^2} \\ &\geq \sqrt{|x-1|^2 + |y|^2} \\ &= \sqrt{|x-1|^2 + |x|^2} \\ &\geq \frac{1}{\sqrt{2}}, \end{aligned}$$

con igualdad si y sólo si $w = z = 0$ y $x = \frac{1}{2}$. ∎

Ahora calcularemos la distancia de las matrices de la forma $\begin{pmatrix} \lambda & 1 \\ 0 & \mu \end{pmatrix}$ a $\mathcal{N}$.

Proposición 2.5. *La distancia con la norma de Frobenius de* $\begin{pmatrix} \lambda & 1 \\ 0 & \mu \end{pmatrix}$ *al conjunto de las matrices normales* $\mathcal{N}$ *es* $\frac{1}{\sqrt{2}}$. *Más aún, si* $\lambda = \mu$, *entonces esta distancia se alcanza para las matrices normales* $\begin{pmatrix} \lambda & \frac{1}{2} \\ y & \mu \end{pmatrix}$ *con* $|y| = \frac{1}{2}$. *En el caso* $\lambda \neq \mu$ *se alcanza para la matriz normal* $\begin{pmatrix} \lambda & \frac{1}{2} \\ y & \mu \end{pmatrix}$ *donde* $y = \frac{1}{2}(\lambda - \mu)(\overline{\lambda - \mu})^{-1}$.

Demostración. Como toda matriz normal $\begin{pmatrix} w & x \\ y & z \end{pmatrix}$ cumple $|x| = |y|$, por el Lema 2.1 tenemos

$$\begin{aligned} \left\| \begin{pmatrix} \lambda & 1 \\ 0 & \mu \end{pmatrix} - \begin{pmatrix} w & x \\ y & z \end{pmatrix} \right\|_F &= \left\| \begin{pmatrix} w-\lambda & x-1 \\ y & z-\mu \end{pmatrix} \right\|_F \\ &= \sqrt{|w-\lambda|^2 + |x-1|^2 + |y|^2 + |z-\mu|^2} \\ &\geq \sqrt{|x-1|^2 + |y|^2} \\ &= \sqrt{|x-1|^2 + |x|^2} \\ &\geq \frac{1}{\sqrt{2}}, \end{aligned}$$

con igualdad si y sólo si $w = \lambda$, $z = \mu$ y $x = \frac{1}{2}$.

Mediante un cálculo se puede verificar que si $\lambda = \mu$, entonces las matrices normales $\begin{pmatrix} \lambda & \frac{1}{2} \\ y & \mu \end{pmatrix}$ con $|y| = \frac{1}{2}$ cumplen igualdad en la desigualdad anterior. En el caso $\lambda \neq \mu$, la igualdad se cumple para la matriz normal $\begin{pmatrix} \lambda & \frac{1}{2} \\ y & \mu \end{pmatrix}$ donde $y = \frac{1}{2}(\lambda - \mu)(\overline{\lambda - \mu})^{-1}$. Claramente, éstas son las únicas matrices normales para las que se alcanza la igualdad. ■

2.2

Distancia Con la Norma Usual de Operadores

Al igual que en la sección anterior, comenzaremos calculando la distacia de la matriz $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ al conjunto de las matrices diagonales.

Proposición 2.6. *La distancia con la norma usual de operadores de* $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ *al conjunto de las matrices diagonales es 1. Además, esta distancia se alcanza para la matriz cero.*

Demostración. Sea $\begin{pmatrix} w & 0 \\ 0 & z \end{pmatrix}$ una matriz diagonal arbitraria. Entonces por el Corolario 1.45 del Capítulo 1

$$\begin{aligned}
\left\| \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} w & 0 \\ 0 & z \end{pmatrix} \right\| &= \left\| \begin{pmatrix} w & -1 \\ 0 & z \end{pmatrix} \right\| \\
&= \sqrt{\left\| \begin{pmatrix} \bar{w} & 0 \\ -1 & \bar{z} \end{pmatrix} \begin{pmatrix} w & -1 \\ 0 & z \end{pmatrix} \right\|} \\
&= \sqrt{\left\| \begin{pmatrix} |w|^2 & -\bar{w} \\ -w & |z|^2+1 \end{pmatrix} \right\|} \\
&= \sqrt{\text{máx} \left\{ |\lambda| : \ \lambda \text{ es eigenvalor de } \begin{pmatrix} |w|^2 & -\bar{w} \\ -w & |z|^2+1 \end{pmatrix} \right\}} \\
&= \sqrt{\left| \frac{|w|^2+|z|^2+1}{2} + \frac{\sqrt{(|w|^2-|z|^2-1)^2+4|w|^2}}{2} \right|}.
\end{aligned}$$

Notemos que el primer término de la suma dentro de la raíz cumple $\frac{|w|^2+|z|^2+1}{2} \geq \frac{1}{2}$. El segundo término satisface $\frac{\sqrt{(|w|^2-|z|^2-1)^2+4|w|^2}}{2} \geq \frac{1}{2}$ pues

$$(|w|^2 - |z|^2 - 1)^2 + 4|w|^2 \geq 1,$$

ya que

$$\begin{aligned}(|w|^2-|z|^2-1)^2+4|w|^2-1 &= |w|^4+|z|^4+1-2|w|^2|z|^2-2|w|^2+2|z|^2+4|w|^2-1\\ &= |w|^4+|z|^4-2|w|^2|z|^2+2|w|^2+2|z|^2\\ &= (|w|^2-|z|^2)^2+2(|w|^2+|z|^2)\geq 0.\end{aligned}$$

De aquí que

$$\sqrt{\left|\frac{|w|^2+|z|^2+1}{2}+\frac{\sqrt{(|w|^2-|z|^2-1)^2+4|w|^2}}{2}\right|}\geq 1.$$

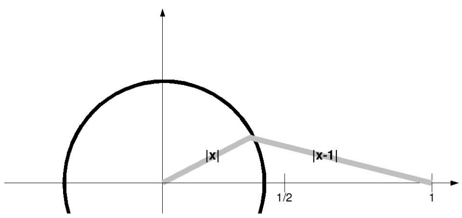
Observemos que la igualdad ocurre si y sólo si $w=z=0$. ■

El siguiente lema será útil.

Lema 2.7. *Sea $x\in\mathbb{C}$. Entonces* $\max\{|x|,|x-1|\}\geq\frac{1}{2}$, *con igualdad si y sólo si $x=\frac{1}{2}$.*

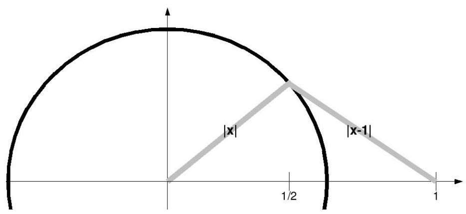
Demostración. Se puede verificar algebraicamente que este máximo es mayor o igual que $\frac{1}{2}$ (por ejemplo, usando el Lema 2.1). Sin embargo, preferimos la siguiente demostración geométrica.

1. Si $|x|<\frac{1}{2}$. Entonces $\frac{1}{2}<|x-1|$, y por lo tanto $\max\{|x|,|x-1|\}=|x-1|>\frac{1}{2}$. Esto lo podemos observar en la siguiente figura



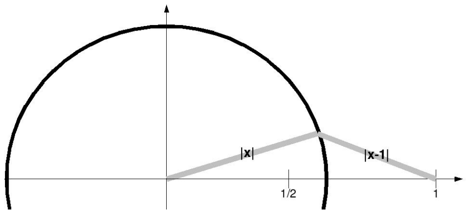
2. Si $|x|>\frac{1}{2}$. Entonces pueden suceder los siguientes tres casos:

   - $|x|=|x-1|$. Notemos que esto ocurre si y sólo si $\mathrm{Re}(x)=\frac{1}{2}$. En la siguiente figura representamos este hecho.



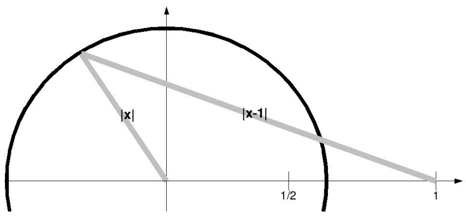
     En este caso $\max\{|x|,|x-1|\}=|x-1|=|x|>\frac{1}{2}$.

   - $|x|>|x-1|$. Con ayuda de la siguiente figura podemos ver que esto ocurre si y sólo si $\mathrm{Re}(x)\in\left(\frac{1}{2},|x|\right]$,



así que entonces $\max\{|x|, |x-1|\} = |x| > \frac{1}{2}$.

- $|x| < |x-1|$. En la siguiente figura se puede obervar que esto ocurre si y sólo si $\mathrm{Re}(x) \in \left[-|x|, \frac{1}{2}\right)$.

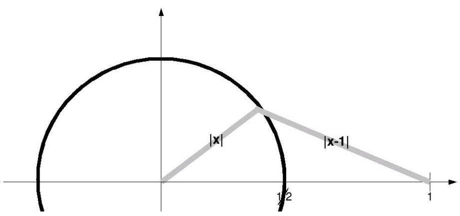


Así, $\max\{|x|, |x-1|\} = |x-1| > \frac{1}{2}$.

En cualquiera de estas tres posibilidades que ocurren cuando $|x| > \frac{1}{2}$, y al igual que en el caso en que $|x| < \frac{1}{2}$, tenemos

$$\max\{|x|, |x-1|\} > \frac{1}{2}.$$

3. Por último supongamos que $|x| = \frac{1}{2}$. Entonces, como podemos ver en la siguiente figura $|x| \leq |x-1|$, con igualdad si y sólo si $x = \frac{1}{2}$.



Así que en este caso $\max\{|x|, |x-1|\} = |x-1| \geq \frac{1}{2}$.

Con esto concluimos lo deseado. ■

Proposición 2.8. *La distancia con la norma usual de operadores de* $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ *al conjunto de las matrices normales de la forma* $\begin{pmatrix} 0 & x \\ y & 0 \end{pmatrix}$ *es* $\frac{1}{2}$. *Además, esta distancia se alcanza para las matrices normales* $\begin{pmatrix} 0 & \frac{1}{2} \\ y & 0 \end{pmatrix}$ *donde* $|y| = \frac{1}{2}$.

Demostración. El resultado se sigue de observar que toda matriz normal $\begin{pmatrix} 0 & x \\ y & 0 \end{pmatrix}$ cumple $|x| = |y|$, y que entonces por el Corolario 1.45 y el Lema 2.7

$$\begin{aligned}
\left\|\begin{pmatrix} 0 & x \\ y & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right\| &= \left\|\begin{pmatrix} 0 & x-1 \\ y & 0 \end{pmatrix}\right\| \\
&= \sqrt{\left\|\begin{pmatrix} 0 & \bar{y} \\ \overline{x-1} & 0 \end{pmatrix}\begin{pmatrix} 0 & x-1 \\ y & 0 \end{pmatrix}\right\|} \\
&= \sqrt{\left\|\begin{pmatrix} |y|^2 & 0 \\ 0 & |x-1|^2 \end{pmatrix}\right\|} \\
&= \text{máx}\{|y|, |x-1|\} \\
&= \text{máx}\{|x|, |x-1|\} \\
&\geq \frac{1}{2},
\end{aligned}$$

con igualdad si y sólo si $x = \frac{1}{2}$ y por lo tanto $|y| = \frac{1}{2}$. ∎

Como veremos a continuación, las matrices para las cuales se alcanza la distancia en la proposición anterior son las mismas para las cuales se alcanza la distancia de $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ a $\mathcal{N}$.

Proposición 2.9. *La distancia con la norma usual de operadores de* $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ *al conjunto de las matrices normales* $\mathcal{N}$ *es* $\frac{1}{2}$*. Además, esta distancia se alcanza para las matrices normales de la forma* $\begin{pmatrix} 0 & \frac{1}{2} \\ y & 0 \end{pmatrix}$ *donde* $|y| = \frac{1}{2}$.

Demostración. Como toda matriz normal $\begin{pmatrix} w & x \\ y & z \end{pmatrix}$ cumple $|x| = |y|$, por el Corolario 1.45 y el Lema 2.1 tenemos

$$\begin{aligned}
\left\|\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} w & x \\ y & z \end{pmatrix}\right\|^2 &= \left\|\begin{pmatrix} w & x-1 \\ y & z \end{pmatrix}\right\|^2 \\
&= \left\|\begin{pmatrix} \bar{w} & \bar{y} \\ \overline{x-1} & \bar{z} \end{pmatrix}\begin{pmatrix} w & x-1 \\ y & z \end{pmatrix}\right\| \\
&= \left\|\begin{pmatrix} |w|^2 + |y|^2 & \bar{w}(x-1) + \bar{y}z \\ \overline{w(x-1) + y\bar{z}} & |x-1|^2 + |z|^2 \end{pmatrix}\right\|
\end{aligned}$$

$$
\begin{aligned}
&= \text{máx}\left\{|\lambda| : \ \lambda \text{ es eigenvalor de } \begin{pmatrix} |w|^2+|y|^2 & \bar{w}(x-1)+\bar{y}z \\ w(x-1)+y\bar{z} & |x-1|^2+|z|^2 \end{pmatrix}\right\} \\
&= \frac{|w|^2+|y|^2+|x-1|^2+|z|^2}{2} + \frac{\sqrt{(|w|^2+|y|^2-|x-1|^2-|z|^2)^2+4\left|\bar{w}(x-1)+\bar{y}z\right|^2}}{2} \\
&\geq \frac{|w|^2+|y|^2+|x-1|^2+|z|^2}{2} \\
&\geq \frac{|y|^2+|x-1|^2}{2} \\
&= \frac{|x|^2+|x-1|^2}{2} \\
&\geq \frac{1}{4},
\end{aligned}
$$

con igualdad cuando $w = z = 0$ y $x = \frac{1}{2}$, pues en este caso

$$
\frac{|w|^2+|y|^2+|x-1|^2+|z|^2}{2} + \frac{\sqrt{(|w|^2+|y|^2-|x-1|^2-|z|^2)^2+4\left|\bar{w}(x-1)+\bar{y}z\right|^2}}{2} = \frac{1}{4}.
$$

Por lo tanto $\left\|\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} w & x \\ y & z \end{pmatrix}\right\| = \frac{1}{2}$ para todo matriz normal de la forma $\begin{pmatrix} w & x \\ y & z \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} \\ y & 0 \end{pmatrix}$ con $|y| = \frac{1}{2}$. ■

Proposición 2.10. *La distancia con la norma usual de operadores de* $\begin{pmatrix} \lambda & 1 \\ 0 & \mu \end{pmatrix}$ *al conjunto de las matrices normales* $\mathcal{N}$ *es* $\frac{1}{2}$*. Más aún, si* $\lambda = \mu$*, entonces esta distancia se alcanza para las matrices normales* $\begin{pmatrix} \lambda & \frac{1}{2} \\ y & \mu \end{pmatrix}$ *con* $|y| = \frac{1}{2}$*. En el caso* $\lambda \neq \mu$ *se alcanza para la matriz normal* $\begin{pmatrix} \lambda & \frac{1}{2} \\ y & \mu \end{pmatrix}$ *donde* $y = \frac{1}{2}(\lambda-\mu)(\overline{\lambda-\mu})^{-1}$*.*

Demostración. Como toda matriz normal $\begin{pmatrix} w & x \\ y & z \end{pmatrix}$ cumple $|x| = |y|$, por el Corolario 1.45 y el Lema 2.1

$$
\left\|\begin{pmatrix} \lambda & 1 \\ 0 & \mu \end{pmatrix} - \begin{pmatrix} w & x \\ y & z \end{pmatrix}\right\|^2 = \left\|\begin{pmatrix} w-\lambda & x-1 \\ y & z-\mu \end{pmatrix}\right\|^2
$$

$$
\begin{aligned}
&= \frac{|w-\lambda|^2+|y|^2+|x-1|^2+|z-\mu|^2}{2} + \\
&\quad \frac{\sqrt{(|w-\lambda|^2+|y|^2-|x-1|^2-|z-\mu|^2)^2+4\left|\overline{(w-\lambda)}(x-1)+\bar{y}(z-\mu)\right|^2}}{2} \\
&\geq \frac{|y|^2+|x-1|^2}{2} \\
&= \frac{|x|^2+|x-1|^2}{2} \\
&\geq \frac{1}{4},
\end{aligned}
$$

con igualdad cuando $w=\lambda$, $z=\mu$ y $x=\frac{1}{2}$.

Mediante un cálculo se puede verificar que si $\lambda=\mu$, entonces las matrices normales $\begin{pmatrix} \lambda & \frac{1}{2} \\ y & \mu \end{pmatrix}$ con $|y|=\frac{1}{2}$ cumplen igualdad arriba. En el caso $\lambda\neq\mu$, la igualdad se cumple para la matriz normal $\begin{pmatrix} \lambda & \frac{1}{2} \\ y & \mu \end{pmatrix}$ donde $y=\frac{1}{2}(\lambda-\mu)(\overline{\lambda-\mu})^{-1}$. Claramente, éstas son las únicas matrices normales para las que se alcanza la igualdad. ∎

La proposición anterior se puede demostrar recurriendo a la Proposición 1.48.

Proposición 2.11. *La distancia con la norma usual de operadores de* $\begin{pmatrix} \lambda & 1 \\ 0 & \mu \end{pmatrix}$ *al conjunto de las matrices normales* $\mathcal{N}$ *es* $\frac{1}{2}$.

Demostración. Por la Proposición 1.48

$$
\operatorname{dist}_F\left(\begin{pmatrix} \lambda & 1 \\ 0 & \mu \end{pmatrix}, \mathcal{N}\right) \leq \sqrt{2}\operatorname{dist}\left(\begin{pmatrix} \lambda & 1 \\ 0 & \mu \end{pmatrix}, \mathcal{N}\right)
$$

y como $\operatorname{dist}_F\left(\begin{pmatrix} \lambda & 1 \\ 0 & \mu \end{pmatrix}, \mathcal{N}\right)=\frac{1}{\sqrt{2}}$ (Proposición 2.5), tenemos

$$
\frac{1}{2} \leq \operatorname{dist}\left(\begin{pmatrix} \lambda & 1 \\ 0 & \mu \end{pmatrix}, \mathcal{N}\right).
$$

El resultado se sigue de notar que esta distancia se alcanza, en el caso $\lambda=\mu$, para las matrices normales de la forma $\begin{pmatrix} \lambda & \frac{1}{2} \\ y & \mu \end{pmatrix}$ donde $|y|=\frac{1}{2}$. En el caso $\lambda\neq\mu$, para la matriz normal $\begin{pmatrix} \lambda & \frac{1}{2} \\ y & \mu \end{pmatrix}$ donde $y=\frac{1}{2}(\lambda-\mu)(\overline{\lambda-\mu})^{-1}$. ∎

2.3

Caso General de 2×2

Sea $A \in \mathcal{M}_2(\mathbb{C})$ una matriz arbitraria. El Teorema de Schur 1.16 nos asegura que existe una matriz unitaria U tal que $T := U^*AU$ es una matriz triangular, digamos $T = \begin{pmatrix} \lambda & x \\ 0 & \mu \end{pmatrix}$. Notemos que si $x = 0$, entonces T es una matriz diagonal y por lo tanto normal. Así, por el Lema 1.58 la norma usual de operadores cumple $\operatorname{dist}(A, \mathcal{N}) = \operatorname{dist}(T, \mathcal{N}) = 0$, lo cual implica que A es normal.

Si $x \neq 0$, entonces

$$\begin{aligned}
\operatorname{dist}(A, \mathcal{N}) &= \operatorname{dist}(T, \mathcal{N}) \\
&= \operatorname{dist}\left(\begin{pmatrix} \lambda & x \\ 0 & \mu \end{pmatrix}, \mathcal{N}\right) \\
&= \operatorname{dist}\left(x\begin{pmatrix} \frac{\lambda}{x} & 1 \\ 0 & \frac{\mu}{x} \end{pmatrix}, \mathcal{N}\right) \\
&= |x| \operatorname{dist}\left(\begin{pmatrix} \frac{\lambda}{x} & 1 \\ 0 & \frac{\mu}{x} \end{pmatrix}, \mathcal{N}\right) \\
&= \frac{|x|}{2}.
\end{aligned}$$

De igual manera tenemos que $\operatorname{dist}_F(A, \mathcal{N}) = \frac{|x|}{\sqrt{2}}$.

Claramente podemos observar que del valor de $|x|$ depende qué tan cerca se encuentra la matriz A a $\mathcal{N}$. Así, mientras $|x|$ sea "casi" cero entonces T, y por lo tanto A, será "casi" normal. Esta es razón por la cual $|x|$ es lo que en el siguiente capítulo Henrici [6] llama la *"desviación para la normalidad de A"*.

CAPÍTULO 3

Cota Superior para la Distancia de Matrices $n \times n$ a $\mathcal{N}$

En el capítulo anterior encontramos explícitamente la distancia de un elemento arbitrario de $\mathcal{M}_2(\mathbb{C})$ al conjunto de las matrices normales $\mathcal{N}$ con las métricas dadas por la norma de Frobenius y la norma usual de operadores. Sin embargo, esta tarea se torna más tortuosa para $n > 2$. Esto es debido a que requiere del cálculo explícito de eigenvalores de matrices de tamaño general y condiciones de normalidad. Así que, en general dado un elemento de $\mathcal{M}_n(\mathbb{C})$, si no podemos obtener explícitamente su distancia a $\mathcal{N}$, por lo menos nos gustaría tener una cota superior de esta distancia. Con este objetivo en mente hemos escrito este capítulo, el cual está basado en un trabajo de Henrici [6].

Comenzaremos con la Sección 3.1 introduciendo el concepto de *desviación para la normalidad* de una matriz triangular superior, y nos ocuparemos en encontrar una cota superior para esta desviación. En la Sección 3.2 veremos cómo esta cota nos ayudará a encontrar la cota superior para la *desviación para la normalidad* de matrices no necesariamente triangulares. Gracias a esta cota, en la Sección 3.3 presentaremos una cota superior para la distancia de un elemento de $\mathcal{M}_n(\mathbb{C})$ a $\mathcal{N}$.

3.1

Desviación Para la Normalidad

Sea $T \in \mathcal{M}_n(\mathbb{C})$ de la forma

$$T = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ 0 & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & m_{nn} \end{pmatrix}.$$

Observemos que al escribir $T = D + M$, donde

$$D = \begin{pmatrix} m_{11} & 0 & \cdots & 0 \\ 0 & m_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & m_{nn} \end{pmatrix}$$

y $M := T - D$, podemos ver que mientras $\|M\|_F^2 = \|T\|_F^2 - \|D\|_F^2$ sea "casi" cero entonces T será "casi" diagonal y por lo tanto "casi" normal (Proposición 1.18). Por esta razón decimos que $\|M\|_F$, es *"la desviación para la normalidad de T".*

En esta sección nos ocuparemos de encontrar una cota superior para *"la desviación para la normalidad de T".*

Iniciaremos definiendo γ_{ij} como la entrada ij de la matriz $T^*T - TT^*$, es decir,

$$T^*T - TT^* =: \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} \end{pmatrix}.$$

El siguiente lema nos proporciona explícitamente los elementos de la diagonal de esta matriz en términos de las entradas de M.

Lema 3.1. *Los elementos de la diagonal de $T^*T - TT^*$ son de la forma*

$$\gamma_{ii} = \sum_{k=1}^{i-1} |m_{ki}|^2 - \sum_{k=i+1}^{n} |m_{ik}|^2,$$

donde la notación anterior indica que $\gamma_{11} = -\sum_{k=2}^{n} |m_{1k}|^2$ y $\gamma_{nn} = \sum_{k=1}^{n-1} |m_{kn}|^2$.

Demostración. Como $T = D + M$, entonces

$$\begin{aligned} T^*T - TT^* &= (D^* + M^*)(D + M) - (D + M)(D^* + M^*) \\ &= D^*D + D^*M + M^*D + M^*M - DD^* - DM^* - MD^* - MM^*, \end{aligned}$$

y dado que las matrices diagonales son normales, es decir, $D^*D = DD^*$, tenemos

$$T^*T - TT^* = D^*M + M^*D + M^*M - DM^* - MD^* - MM^*.$$

Así, los elementos de la diagonal de $T^*T - TT^*$ cumplen

$$\begin{aligned} \gamma_{ii} &= [T^*T - TT^*]_{ii} \\ &= [D^*M]_{ii} + [M^*D]_{ii} + [M^*M]_{ii} - [DM^*]_{ii} - [MD^*]_{ii} - [MM^*]_{ii}. \end{aligned}$$

Luego, como los elementos de la diagonal de

$$D^*M = \begin{pmatrix} \overline{m}_{11} & 0 & 0 & \cdots & 0 \\ 0 & \overline{m}_{22} & 0 & \cdots & 0 \\ 0 & 0 & \overline{m}_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \overline{m}_{nn} \end{pmatrix} \begin{pmatrix} 0 & m_{12} & m_{13} & \cdots & m_{1n} \\ 0 & 0 & m_{23} & \cdots & m_{2n} \\ 0 & 0 & 0 & \cdots & m_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

y de

$$DM^* = \begin{pmatrix} m_{11} & 0 & 0 & \cdots & 0 \\ 0 & m_{22} & 0 & \cdots & 0 \\ 0 & 0 & m_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & m_{nn} \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ \overline{m}_{12} & 0 & 0 & \cdots & 0 \\ \overline{m}_{13} & \overline{m}_{23} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \overline{m}_{1n} & \overline{m}_{2n} & \overline{m}_{3n} & \cdots & 0 \end{pmatrix}$$

son cero, y por lo tanto los elementos de la diagonal de $M^*D = (D^*M)^*$ y de $MD^* = (DM^*)^*$ también son cero, tenemos que

$$\begin{aligned}
\gamma_{ii} &= [T^*T - TT^*]_{ii} \\
&= [M^*M]_{ii} - [MM^*]_{ii} \\
&= \left[\begin{pmatrix} 0 & ... & 0 & ... & 0 & 0 \\ \overline{m}_{12} & ... & 0 & ... & 0 & 0 \\ \vdots & & \vdots & & \vdots & \vdots \\ \overline{m}_{1i} & ... & \overline{m}_{i-1,i} & ... & 0 & 0 \\ \vdots & & \vdots & & \vdots & \vdots \\ \overline{m}_{1,n} & ... & \overline{m}_{i-1,n} & ... & \overline{m}_{n-1,n} & 0 \end{pmatrix} \begin{pmatrix} 0 & m_{12} & ... & m_{1i} & ... & m_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & ... & m_{i-1,i} & ... & m_{i-1,n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & ... & 0 & ... & m_{n-1,n} \\ 0 & 0 & ... & 0 & ... & 0 \end{pmatrix} \right]_{ii} \\
&- \left[\begin{pmatrix} 0 & m_{12} & ... & m_{1i} & ... & m_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & ... & m_{i,i+1} & ... & m_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & ... & 0 & ... & m_{n-1,n} \\ 0 & 0 & ... & 0 & ... & 0 \end{pmatrix} \begin{pmatrix} 0 & ... & 0 & ... & 0 & 0 \\ \overline{m}_{12} & ... & 0 & ... & 0 & 0 \\ \vdots & & \vdots & & \vdots & \vdots \\ \overline{m}_{1i} & ... & \overline{m}_{i,i+1} & ... & 0 & 0 \\ \vdots & & \vdots & & \vdots & \vdots \\ \overline{m}_{1,n} & ... & \overline{m}_{in} & ... & \overline{m}_{n-1,n} & 0 \end{pmatrix} \right]_{ii} \\
&= \sum_{k=1}^{i-1} |m_{ki}|^2 - \sum_{k=i+1}^{n} |m_{ik}|^2,
\end{aligned}$$

como queríamos. ∎

La siguiente proposición presenta una primera cota superior para $\|M\|_F$ en términos de los γ_{ii}.

Proposición 3.2. *Sean T, M y γ_{ii} como antes. Entonces*

$$\|M\|_F^2 \leq \gamma_{22} + 2\gamma_{33} + \cdots + (n-1)\gamma_{nn} = \sum_{i=1}^{n} (i-1)\gamma_{ii}, \tag{3.1}$$

con igualdad si y sólo si T es de la forma

$$\begin{pmatrix} m_{11} & m_{12} & 0 & \cdots & 0 & 0 \\ 0 & m_{22} & m_{23} & \cdots & 0 & 0 \\ 0 & 0 & m_{33} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & m_{n-1,n-1} & m_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & m_{nn} \end{pmatrix}.$$

Demostración. Primero observemos que por el lema anterior

$$\begin{aligned}
\sum_{i=1}^{n}(i-1)\gamma_{ii} &= \sum_{i=1}^{n-1}(i-1)\gamma_{ii} + (n-1)\gamma_{nn} \\
&= \sum_{i=1}^{n-1}(i-1)\left[\sum_{k=1}^{i-1}|m_{ki}|^2 - \sum_{k=i+1}^{n}|m_{ik}|^2\right] + (n-1)\sum_{k=1}^{n-1}|m_{kn}|^2 \\
&= \sum_{i=1}^{n-1}\sum_{k=1}^{i-1}(i-1)|m_{ki}|^2 - \sum_{i=1}^{n-1}\sum_{k=i+1}^{n}(i-1)|m_{ik}|^2 + \sum_{k=1}^{n-1}(n-1)|m_{kn}|^2 \\
&= \sum_{i=1}^{n}\sum_{k=1}^{i-1}(i-1)|m_{ki}|^2 - \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(i-1)|m_{ij}|^2,
\end{aligned}$$

y como

$$\sum_{i=1}^{n}\sum_{k=1}^{i-1}(i-1)|m_{ki}|^2 = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(j-1)|m_{ij}|^2,$$

entonces

$$\begin{aligned}
\sum_{i=1}^{n}(i-1)\gamma_{ii} &= \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(j-1)|m_{ij}|^2 - \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(i-1)|m_{ij}|^2 \\
&= \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(j-i)|m_{ij}|^2.
\end{aligned}$$

Así,

$$\begin{aligned}
\sum_{i=1}^{n}(i-1)\gamma_{ii} - \|M\|_F^2 &= \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(j-i)|m_{ij}|^2 - \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}|m_{ij}|^2 \\
&= \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(j-i-1)|m_{ij}|^2,
\end{aligned}$$

y como es mayor o igual que 0, se tiene la desigualdad (3.1).

En cuanto a la igualdad, notemos que cuando $j = i+1$, entonces $j-i-1 = 0$ y por lo tanto

$$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(j-i-1)|m_{ij}|^2 \quad = \quad \sum_{i=1}^{n-1}\sum_{j=i+2}^{n}(j-i-1)|m_{ij}|^2.$$

Esto nos permite concluir que la igualdad en (3.1) ocurre si y sólo si $m_{ij} = 0$ para todo $j > i+1$. ∎

Ahora proporcionaremos una cota superior para $\sum_{i=1}^{n}(i-1)\gamma_{ii}$. Pero antes, observemos que por el Lema 1.21

$$\sum_{i=1}^{n}\gamma_{ii} = \mathrm{traza}(T^*T - TT^*) = 0. \tag{3.2}$$

Proposición 3.3. *Sea γ_{ii} como antes. Entonces*

$$\left(\sum_{i=1}^{n}(i-1)\gamma_{ii}\right)^2 \leq \frac{n^3-n}{12}\sum_{i=1}^{n}\gamma_{ii}^2,$$

con igualdad si y sólo si $\gamma_{ii} = a\frac{2i-n-1}{2}$ para algún número real a.

Demostración. Primero notemos que por la ecuación (3.2)

$$\sum_{i=1}^{n}(i-1)\gamma_{ii} = \sum_{i=1}^{n}(i-1)\gamma_{ii} - c\sum_{i=1}^{n}\gamma_{ii} = \sum_{i=1}^{n}(i-1-c)\gamma_{ii},$$

para cualquier constante c. Así, por la desigualdad de Cauchy-Schwarz

$$\left(\sum_{i=1}^{n}(i-1)\gamma_{ii}\right)^2 \leq \sum_{i=1}^{n}(i-1-c)^2\sum_{i=1}^{n}\gamma_{ii}^2, \tag{3.3}$$

con igualdad si y sólo si $\gamma_{ii} = a(i-1-c)$ para algún número real a.

Como tenemos una desigualdad distinta para cada constante c, buscamos la que nos de la mejor estimación. Esto se logra observando que

$$\begin{aligned}\sum_{i=1}^{n}(i-1-c)^2 &= \sum_{i=1}^{n}\left[(i-1)^2 - 2(i-1)c + c^2\right]\\ &= \frac{(n-1)n(2n-1)}{6} - cn(n-1) + c^2n,\end{aligned}$$

y que la función $f(x) = \frac{(n-1)n(2n-1)}{6} - xn(n-1) + x^2n$ alcanza su mínimo en $x = \frac{n-1}{2}$. De aquí que la desigualdad (3.3) toma la siguiente forma

$$\left(\sum_{i=1}^{n}(i-1)\gamma_{ii}\right)^2 \leq \sum_{i=1}^{n}\left(i-1-\frac{n-1}{2}\right)^2\sum_{i=1}^{n}\gamma_{ii}^2,$$

con igualdad si y sólo si $\gamma_{ii} = a(i-1-\frac{n-1}{2}) = a\frac{2i-n-1}{2}$ para algún número real a.

El resultado se sigue de observar que

$$\begin{aligned}\sum_{i=1}^{n}\left(i-1-\frac{n-1}{2}\right)^2 &= \sum_{k=0}^{n-1}\frac{[2k-(n-1)]^2}{4}\\ &= \frac{1}{4}\sum_{k=0}^{n-1}\left[4k^2+(n-1)^2-4k(n-1)\right]\\ &= \frac{1}{12}[n^3-n].\end{aligned}$$

■

3.2

Cota de Henrici

Consideremos el caso en que $A \in \mathcal{M}_n(\mathbb{C})$ no es necesariamente triangular superior. Vamos a definir la *"desviación para la normalidad de A"*, o como Henrici [6] la llama *"F-departure from normality of A"*, como $\sqrt{\|A\|_F^2 - \sum|\lambda_i|^2}$ donde $\lambda_1, \lambda_2, ..., \lambda_n$ son los eigenvalores de A incluyendo las posibles multiplicidades. En el siguiente teorema presentaremos la cota superior que ofrece Henrici [6] para esta desviación, y además veremos en la demostración que esta definición coincide con la de la sección anterior cuando A es una matriz triangular. La ventaja de esta cota es que se puede cálcular en términos elementales y se reduce a cero si A es normal.

Teorema 3.4. *Sea $A \in \mathcal{M}_n(\mathbb{C})$ y sean $\lambda_1, \lambda_2, ..., \lambda_n$ sus eigenvalores incluyendo las posibles multiplicidades. Entonces*

$$\left(\|A\|_F^2 - \sum_{i=1}^{n}|\lambda_i|^2\right)^{\frac{1}{2}} \leq \left(\frac{n^3-n}{12}\right)^{\frac{1}{4}}\sqrt{\|A^*A - AA^*\|_F}, \tag{3.4}$$

con igualdad si y sólo si A es normal o si A es una matriz unitariamente equivalente a una matriz de la forma

$$\begin{pmatrix} \lambda & \alpha_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & \alpha_2 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & \alpha_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix},$$

donde $\lambda \in \mathbb{C}$ *y* $|\alpha_k|^2 = ak(n-k)$ *para algún* $a > 0$ *y para* $k \in \{1, 2, ..., n-1\}$.

Demostración. Por el Teorema de Schur 1.16 sabemos que existe U unitaria, tal que $T := U^*AU$ es una matriz triangular superior, digamos

$$T = \begin{pmatrix} \lambda_1 & m_{12} & m_{13} & \cdots & m_{1,n-1} & m_{1n} \\ 0 & \lambda_2 & m_{23} & \cdots & m_{2,n-1} & m_{2n} \\ 0 & 0 & \lambda_3 & \cdots & m_{3,n-1} & m_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{n-1} & m_{n-1,n} \\ 0 & 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix},$$

donde los elementos de la diagonal son los eigenvalores de A incluyendo las posibles multiplicidades. Como en la sección anterior, escribamos $T = D + M$, donde

$$D := \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

y $M := T - D$.

Notemos que como la norma de Frobenius es unitariamente invariante

$$\|A\|_F^2 = \|U^*AU\|_F^2 = \|T\|_F^2 = \|M + D\|_F^2 = \|M\|_F^2 + \|D\|_F^2 = \|M\|_F^2 + \sum_{i=1}^{n} |\lambda_i|^2.$$

Por lo tanto

$$\left(\|A\|_F^2 - \sum_{i=1}^{n} |\lambda_i|^2\right)^{\frac{1}{2}} = \|M\|_F.$$

Esto nos permite ver que $\|M\|_F$ no depende de la elección de la matriz unitaria del Teorema de Schur 1.16, y además justifica el uso del nombre de "*desviación para la normalidad de* A".

Así, por la Proposición 3.2 y la Proposición 3.3

$$\begin{aligned} \left(\|A\|_F^2 - \sum_{i=1}^{n} |\lambda_i|^2\right)^{\frac{1}{2}} &= \|M\|_F \\ &\leq \left(\sum_{i=1}^{n} (i-1)\gamma_{ii}\right)^{\frac{1}{2}} \\ &\leq \left(\frac{n^3 - n}{12} \sum_{i=1}^{n} \gamma_{ii}^2\right)^{\frac{1}{4}}, \end{aligned}$$

donde γ_{ii} es la i-ésima entrada de la diagonal de $T^*T - TT^*$.

Luego, como $A = UTU^*$ entonces

$$\begin{aligned} A^*A - AA^* &= (UTU^*)^*(UTU^*) - (UTU^*)(UTU^*)^* \\ &= U(T^*T - TT^*)U^*, \end{aligned}$$

y debido a que la norma de Frobenius es unitariamente invariante,

$$\|A^*A - AA^*\|_F^2 = \|U(T^*T - TT^*)U^*\|_F^2 = \|T^*T - TT^*\|_F^2 = \sum_{ij} \gamma_{ij}^2 \geq \sum_{i=1}^{n} \gamma_{ii}^2. \quad (3.5)$$

De esto se sigue que

$$\left(\|A\|_F^2 - \sum_{i=1}^{n} |\lambda_i|^2\right)^{\frac{1}{2}} \leq \left(\frac{n^3 - n}{12}\|A^*A - AA^*\|_F^2\right)^{\frac{1}{4}},$$

o equivalentemente,

$$\left(\|A\|_F^2 - \sum_{i=1}^{n} |\lambda_i|^2\right)^{\frac{1}{2}} \leq \left(\frac{n^3 - n}{12}\right)^{\frac{1}{4}} \sqrt{\|A^*A - AA^*\|_F},$$

es decir, la desigualdad (3.4).

En cuanto a la igualdad en (3.4), esta ocurre sólo si hay igualdad simultáneamente en la Proposición 3.2, la Proposición 3.3 y la ecuación (3.5).

Primero, supongamos que se tiene igualdad en la Proposición 3.2, es decir, que $m_{ij} = 0$ para todo $j > i + 1$. Entonces por el Lema 3.1

$$\begin{aligned} \gamma_{11} &= -|m_{12}|^2 \\ \gamma_{22} &= |m_{12}|^2 - |m_{23}|^2 \\ \gamma_{33} &= |m_{23}|^2 - |m_{34}|^2 \\ &\vdots \\ \gamma_{kk} &= |m_{k-1,k}|^2 - |m_{k,k+1}|^2 \\ &\vdots \\ \gamma_{nn} &= |m_{n-1,n}|^2. \end{aligned}$$

De aquí se sigue que

$$\sum_{i=1}^{k} \gamma_{ii} = -|m_{k,k+1}|^2, \text{ para todo } k \in \{1, 2, ..., n-1\}\ .$$

Si también suponemos igualdad en la Proposición 3.3, es decir, $\gamma_{ii} = a\frac{2i-n-1}{2}$ para algún número real a, entonces mediante un desarrollo y simplificación directa

$$\begin{aligned} \sum_{i=1}^{k} \gamma_{ii} &= \sum_{i=1}^{k} a\frac{2i - n - 1}{2} \\ &= \frac{a}{2}k(k - n). \end{aligned}$$

Esto implica que $|m_{k,k+1}|^2 = \frac{a}{2}k(n-k)$, para todo $k \in \{1, 2, ..., n-1\}$, y por lo tanto que $a \geq 0$.

Notemos que en el caso $a = 0$ tenemos $m_{k,k+1} = 0$ para todo $k \in \{1, 2, ..., n-1\}$, por lo cual T, y entoncen también A, es una matriz normal.

Ahora supongamos igualdad en la ecuación (3.5), es decir que

$$
\begin{aligned}
&T^*T - TT^* = \\
&= \begin{pmatrix} \bar{\lambda}_1 & 0 & 0 & \cdots & 0 \\ \overline{m}_{12} & \bar{\lambda}_2 & 0 & \cdots & 0 \\ 0 & \overline{m}_{23} & \bar{\lambda}_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \bar{\lambda}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & m_{12} & 0 & \cdots & 0 \\ 0 & \lambda_2 & m_{23} & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix} \\
&\quad - \begin{pmatrix} \lambda_1 & m_{12} & 0 & \cdots & 0 \\ 0 & \lambda_2 & m_{23} & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} \bar{\lambda}_1 & 0 & 0 & \cdots & 0 \\ \overline{m}_{12} & \bar{\lambda}_2 & 0 & \cdots & 0 \\ 0 & \overline{m}_{23} & \bar{\lambda}_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \bar{\lambda}_n \end{pmatrix} \\
&= \begin{pmatrix} |\lambda_1|^2 & \bar{\lambda}_1 m_{12} & 0 & \cdots & 0 \\ \lambda_1 \overline{m}_{12} & |m_{12}|^2 + |\lambda_2|^2 & \bar{\lambda}_2 m_{23} & \cdots & 0 \\ 0 & \lambda_2 \overline{m}_{23} & |m_{23}|^2 + |\lambda_3|^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \bar{\lambda}_{n-1} m_{n-1,n} \\ 0 & 0 & 0 & \cdots & |m_{n-1,n}|^2 + |\lambda_n|^2 \end{pmatrix} \\
&\quad - \begin{pmatrix} |\lambda_1|^2 + |m_{12}|^2 & \bar{\lambda}_2 m_{12} & 0 & \cdots & 0 \\ \lambda_2 \overline{m}_{12} & |m_{23}|^2 + |\lambda_2|^2 & \bar{\lambda}_3 m_{23} & \cdots & 0 \\ 0 & \lambda_3 \overline{m}_{23} & |m_{34}|^2 + |\lambda_3|^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \bar{\lambda}_n m_{n-1,n} \\ 0 & 0 & 0 & \cdots & |\lambda_n|^2 \end{pmatrix} \\
&= \begin{pmatrix} -|m_{12}|^2 & (\overline{\lambda_1 - \lambda_2}) m_{12} & 0 & \cdots & 0 \\ (\lambda_1 - \lambda_2)\overline{m}_{12} & |m_{12}|^2 - |m_{23}|^2 & (\overline{\lambda_2 - \lambda_3}) m_{23} & \cdots & 0 \\ 0 & (\lambda_2 - \lambda_3)\overline{m}_{23} & |m_{23}|^2 - |m_{34}|^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & (\overline{\lambda_{n-1} - \lambda_n}) m_{n-1,n} \\ 0 & 0 & 0 & \cdots & |m_{n-1,n}|^2 \end{pmatrix},
\end{aligned}
$$

es una matriz diagonal. Entonces

$$(\lambda_k - \lambda_{k+1})\overline{m}_{k,k+1} = 0, \text{ para todo } k \in \{1, 2, ..., n-1\}.$$

Si además suponemos $a > 0$, entonces $|m_{k,k+1}| \neq 0$ para todo $k \in \{1, 2, ..., n-1\}$, y por lo tanto $\lambda_k = \lambda_{k+1}$ para todo $k \in \{1, 2, ..., n-1\}$.

En conclusión, si hay igualdad en 3.4, entonces A es normal o A es una matriz unitariamente equivalente a una matriz de la forma

$$\begin{pmatrix} \lambda & \alpha_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & \alpha_2 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & \alpha_{n-1} \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix},$$

donde $\lambda \in \mathbb{C}$ y $|\alpha_k|^2 = ak(n-k)$, para algún $a > 0$ y para $k \in \{1, 2, ..., n-1\}$. El recíproco es claramente cierto. ■

3.3

Cota Superior para la Distancia de Matrices $n \times n$ a $\mathcal{N}$

A continuación veremos cómo la cota de Henrici nos ofrece una cota superior para la distancia de una matriz al conjunto de las matrices normales con la métrica dada por la norma de Frobenius.

Teorema 3.5. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces*

$$\mathrm{dist}_F(A, \mathcal{N}) \leq \left(\frac{n^3 - n}{12}\right)^{\frac{1}{4}} \sqrt{\|A^*A - AA^*\|_F}.$$

Demostración. Sean T, M y D como en la prueba del Teorema 3.4. Entonces por el Lema 1.58 y la cota de Henrici

$$\begin{aligned} \mathrm{dist}_F(A, \mathcal{N}) &= \mathrm{dist}_F(T, \mathcal{N}) \\ &= \inf\{\|T - N\|_F : N \in \mathcal{N}\} \\ &= \inf\{\|D + M - N\|_F : N \in \mathcal{N}\} \\ &\leq \|M\|_F \\ &\leq \left(\frac{n^3 - n}{12}\right)^{\frac{1}{4}} \sqrt{\|A^*A - AA^*\|_F}. \end{aligned}$$

■

A continuación presentamos la versión de este teorema con respecto a la norma usual de operadores.

Teorema 3.6. *Sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces*

$$\operatorname{dist}(A, \mathcal{N}) \leq \left(\frac{n(n^3-n)}{12}\right)^{\frac{1}{4}} \sqrt{\|A^*A - AA^*\|}.$$

Demostración. Como antes $\operatorname{dist}(A, \mathcal{N}) \leq \|M\|$. El resultado se sigue de la Proposición 1.48 y la cota de Henrici, pues

$$\begin{aligned}
\|M\| &\leq \|M\|_F \\
&\leq \left(\frac{n^3-n}{12}\right)^{\frac{1}{4}} \sqrt{\|A^*A - AA^*\|_F} \\
&\leq \left(\frac{n^3-n}{12}\right)^{\frac{1}{4}} \sqrt{\sqrt{n}\|A^*A - AA^*\|} \\
&= \left(\frac{n(n^3-n)}{12}\right)^{\frac{1}{4}} \sqrt{\|A^*A - AA^*\|}.
\end{aligned}$$

■

En general, podemos escsribir la versión del Teorema 3.5 con respecto a cualquier norma.

Teorema 3.7. *Sea $\|\cdot\|_1$ una norma definida en $\mathcal{M}_n(\mathbb{C})$, y sea $A \in \mathcal{M}_n(\mathbb{C})$. Entonces existe $c > 0$ tal que*

$$\operatorname{dist}_{\|\cdot\|_1}(A, \mathcal{N}) \leq c \left(\frac{n^3-n}{12}\right)^{\frac{1}{4}} \sqrt{\|A^*A - AA^*\|_1}.$$

Demostración. Sean $c_1, c_2 > 0$ tales que $c_1\|A\|_F \leq \|A\|_1 \leq c_2\|A\|_F$ para todo $A \in M_n(\mathbb{C})$ (Teorema 1.52). Entonces por el Teorema 3.5.

$$\begin{aligned}
\operatorname{dist}_{\|\cdot\|_1}(A, \mathcal{N}) &\leq c_2 \operatorname{dist}_F(A, \mathcal{N}) \\
&\leq c_2 \left(\frac{n^3-n}{12}\right)^{\frac{1}{4}} \sqrt{\|A^*A - AA^*\|_F} \\
&\leq c_2 \left(\frac{n^3-n}{12}\right)^{\frac{1}{4}} \sqrt{\frac{1}{c_1}\|A^*A - AA^*\|_1} \\
&= \frac{c_2}{\sqrt{c_1}} \left(\frac{n^3-n}{12}\right)^{\frac{1}{4}} \sqrt{\|A^*A - AA^*\|_1}.
\end{aligned}$$

■

Sea $A \in \mathcal{M}_n(\mathbb{C})$. De estos teoremas se puede observar que si $\|A^*A - AA^*\|_1$ es casi cero, es decir, si A es *casi* normal, entonces A está *cerca* del conjunto de las matrices normales. En el siguiente capítulo estudiaremos esta idea.

CAPÍTULO 4

Matrices *Casi* Normales Están *Cerca* de Matrices Normales

Iniciaremos este capítulo notando que de una observación de Higham [7] se puede ver que las matrices *casi* hermitianas están *cerca* de las matrices hermitianas. Generalizaremos este hecho para un tipo de matrices que llamaremos λ-hermitianas (Sección 4.1). En la Sección 4.2 veremos que gracias a los resultados del capítulo anterior podemos concluir que las matrices *casi* normales están *cerca* del conjunto de las matrices normales $\mathcal{N}$. Más precisamente, dado $n > 1$ un entero y $\epsilon > 0$, presentaremos una expresión explícita para $\delta > 0$ (resultado principal de este trabajo) tal que para cualquier $A \in \mathcal{M}_n(\mathbb{C})$ que satisfaga $\|A^*A - AA^*\| < \delta$ la distancia de A a $\mathcal{N}$ es menor que ϵ.

4.1

Matrices λ-Hermitianas

Definición 4.1. Sea $\lambda \in \mathbb{C}$. Decimos que $A \in \mathcal{M}_n(\mathbb{C})$ es una matriz λ-*hermitiana*, si $A^* = \lambda A$.

Notemos que cuando $\lambda = 1$ las matrices λ-hermitianas son simplemente las matrices hermitianas o autoadjuntas.

Claramente la matriz cero siempre es λ-hermitiana, para todo $\lambda \in \mathbb{C}$.

Proposición 4.2. *Sea $\lambda \in \mathbb{C}$ y $A \neq 0$ una matriz λ-hermitiana. Entonces $|\lambda| = 1$.*

Demostración. Como $A^* = \lambda A$, por el Lema 1.38, tenemos $\|A\| = \|A^*\| = \|\lambda A\| =$

$|\lambda| \cdot \|A\|$. Luego, dado que $A \neq 0$, se sigue $|\lambda| = 1$. ■

Notemos que si $\lambda = e^{i\theta}$ y A es una matriz λ-hermitiana, entonces $H := e^{i\frac{\theta}{2}}A$ es una matriz autoadjunta, pues

$$H^* = (e^{i\frac{\theta}{2}}A)^* = e^{-i\frac{\theta}{2}}A^* = e^{-i\frac{\theta}{2}}(e^{i\theta}A) = e^{i\frac{\theta}{2}}A = H.$$

Por lo tanto, $A = e^{-i\frac{\theta}{2}}H$. Esto implica que toda matriz λ-hermitiana es un múltiplo de alguna matriz autoadjunta.

Sea $\lambda \in \mathbb{C}$. Denotaremos por $\lambda\mathbf{H}$ al conjunto de las matrices λ-hermitianas.

Proposición 4.3. *Sea $\lambda \in \mathbb{C}$ tal que $|\lambda| = 1$. Entonces*

1. *$A \in \lambda\boldsymbol{H}$ si y sólo si $A^* \in \bar{\lambda}\boldsymbol{H}$.*
2. *Para todo $A \in \mathcal{M}_n(\mathbb{C})$ tenemos $A = \frac{A+\bar{\lambda}A^*}{2} + \frac{A-\bar{\lambda}A^*}{2}$ con $\left(\frac{A+\bar{\lambda}A^*}{2}\right) \in \lambda\boldsymbol{H}$ y $\left(\frac{A-\bar{\lambda}A^*}{2}\right) \in (-\lambda)\boldsymbol{H}$.*

Demostración. Con un cálculo se puede verificar que la parte 1 se cumple.

Para probar $\frac{A+\bar{\lambda}A^*}{2} \in \lambda\mathbf{H}$, basta observar que $(A + \bar{\lambda}A^*)^* = A^* + \lambda A = \lambda(A + \bar{\lambda}A^*)$. De manera análoga, $\frac{A-\bar{\lambda}A^*}{2} \in (-\lambda)\mathbf{H}$. ■

En particular cuando $\lambda = 1$, todo elemento $A \in \mathcal{M}_n(\mathbb{C})$ se puede escribir como $A = \frac{A+A^*}{2} + \frac{A-A^*}{2}$, donde $\left(\frac{A+A^*}{2}\right) \in \mathbf{H}$ y $\left(\frac{A-A^*}{2}\right) \in (-1)\mathbf{H}$. Además, por una observación de Higham [7] sabemos que:

$$\operatorname{dist}_F(A, \mathbf{H}) = \left\| \frac{A - A^*}{2} \right\|_F,$$

la cual se alcanza para la matriz $\frac{A+A^*}{2} \in \mathbf{H}$. Esto podría sugerir que $A = \frac{A+A^*}{2} + \frac{A-A^*}{2}$ es una descomposición ortogonal para $\mathcal{M}_n(\mathbb{C})$ con el producto interno del Lema 1.26 (ver [1, pág. 33]). Sin embargo, no es así, aunque en ocasiones $\frac{A+A^*}{2}$ y $\frac{A-A^*}{2}$ puedan ser ortogonales. Esto se debe a que el complemento ortogonal de $\mathbf{H}$ consta únicamente de la matriz cero.

En general $A = \frac{A+\bar{\lambda}A^*}{2} + \frac{A-\bar{\lambda}A^*}{2}$ no es una descomposición ortogonal. Sin embargo, usando un método similar al de Higham [7] tenemos la siguiente proposición.

Proposición 4.4. *Sea $|\lambda| = 1$ y sea $\|\cdot\|_1$ una norma definida en $\mathcal{M}_n(\mathbb{C})$ tal que para todo $X \in \mathcal{M}_n(\mathbb{C})$ se cumple $\|X^*\|_1 = \|X\|_1$. Entonces para cada $A \in \mathcal{M}_n(\mathbb{C})$ tenemos $\operatorname{dist}_{\|\cdot\|_1}(A, \lambda\boldsymbol{H}) = \dfrac{\|A - \bar{\lambda}A^*\|_1}{2}$.*

Demostración. Primero notemos que para todo $B \in \lambda\mathbf{H}$

$$\|A - B\|_1 = \left\| \frac{A + \bar{\lambda}A^*}{2} - B + \frac{A - \bar{\lambda}A^*}{2} \right\|_1 \leq \left\| \frac{A + \bar{\lambda}A^*}{2} - B \right\|_1 + \left\| \frac{A - \bar{\lambda}A^*}{2} \right\|_1,$$

y dado que $\frac{A+\overline{\lambda}A^*}{2} \in \lambda\mathbf{H}$ entonces

$$\begin{aligned}
\mathrm{dist}_{\|\cdot\|_1}(A, \lambda\mathbf{H}) &= \text{ínf}\{\|A - B\|_1 : B \in \lambda\mathbf{H}\} \\
&\leq \text{ínf}\left\{\left\|\frac{A+\overline{\lambda}A^*}{2} - B\right\|_1 + \left\|\frac{A-\overline{\lambda}A^*}{2}\right\|_1 : B \in \mathbf{H}\right\} \\
&= \frac{\|A - \overline{\lambda}A^*\|_1}{2}.
\end{aligned}$$

Por otro lado, para todo $B \in \lambda\mathbf{H}$ la norma $\|\cdot\|_1$ cumple

$$\left\|\frac{A-\overline{\lambda}A^*}{2}\right\|_1 = \left\|\frac{A-\overline{\lambda}A^*}{2} - \frac{B-\overline{\lambda}B^*}{2}\right\|_1 \leq \left\|\frac{A-B}{2}\right\|_1 + |\lambda|\left\|\frac{(A-B)^*}{2}\right\|_1 = \|A-B\|_1.$$

Así, $\dfrac{\|A-\overline{\lambda}A^*\|_1}{2} \leq \text{ínf}\{\|A-B\|_1 : B \in \lambda\mathbf{H}\} = \mathrm{dist}_{\|\cdot\|_1}(A, \lambda\mathbf{H})$. ■

Esta proposición indica que las matrices *casi* λ-hermitianas están *cerca* de las matrices λ-hermitianas. Este hecho se cumple también sin la hipótesis de que la norma $\|\cdot\|_1$ satisfaga $\|X^*\|_1 = \|X\|_1$ para todo $X \in \mathcal{M}_n(\mathbb{C})$. Veamos esto.

Teorema 4.5. *Sea $\|\cdot\|_1$ una norma definida en $\mathcal{M}_n(\mathbb{C})$, sea $\lambda \in \mathbb{C}$ con $|\lambda| = 1$, y sea $\epsilon > 0$. Entonces $\delta = 2\epsilon$ cumple que, para toda $A \in \mathcal{M}_n(\mathbb{C})$ que satisfaga $\|A^* - \lambda A\|_1 < \delta$, existe $B \in \lambda\boldsymbol{H}$ tal que $\|A - B\|_1 < \epsilon$.*

Demostración. Sea $A \in \mathcal{M}_n(\mathbb{C})$ tal que $\|A^* - \lambda A\|_1 < \delta$. Entonces $\frac{A+\overline{\lambda}A^*}{2} \in \lambda\mathbf{H}$ cumple

$$\left\|A - \frac{A+\overline{\lambda}A^*}{2}\right\|_1 = \frac{\|A-\overline{\lambda}A^*\|_1}{2} < \frac{\delta}{2} = \epsilon,$$

como queríamos. ■

Notese que en este teorema, δ no depende de la dimensión de la matriz.

4.2

Matrices Normales

Nos gustaría saber hasta donde se pueden extender los resultados del Capítulo 3. Una forma sería extender a matrices que cumplan $A^*A = \lambda AA^*$ para algún λ. Lamentablemente no se obtiene nada nuevo. Veamos esto.

Proposición 4.6. *Sea $\lambda \in \mathbb{C}$ y sea $A \neq 0$ tal que $A^*A = \lambda AA^*$. Entonces $\lambda = 1$.*

Demostración. Como $A^*A = \lambda AA^*$, por el Lema 1.21 se cumple

$$\begin{aligned} 0 &= \text{traza}(A^*A - \lambda AA^*) \\ &= \text{traza}(A^*A) - \lambda\,\text{traza}(AA^*) \\ &= \text{traza}(A^*A) - \lambda\,\text{traza}(A^*A) \\ &= (1-\lambda)\,\text{traza}(A^*A) \\ &= (1-\lambda)\|A\|_F^2. \end{aligned}$$

Luego, como por hipótesis $A \neq 0$, necesariamente $\lambda = 1$. ∎

Notemos que si $\lambda = 1$ entonces las matrices en esta proposición son normales.

A continuanción veremos que las matrices *casi* normales están *cerca* del conjunto de las matrices normales $\mathcal{N}$.

Teorema 4.7 (Huaxin Lin [8]). *Dado $\epsilon > 0$, existe $\delta > 0$ tal que para todo entero $n > 1$ y para cualquier matriz $A \in \mathcal{M}_n(\mathbb{C})$ que cumpla $\|A^*A - AA^*\| < \delta$, existe $N \in \mathcal{N}$ que satisface $\|A - N\| < \epsilon$.*

No abordaremos la demostración de este teorema pues es más bien técnica y necesita fuertemente de la teoría de álgebras C^*. Observemos que en particular, Lin demostró que δ se puede escoger independientemente de n.

En lo que resta de esta sección nos ocuparemos en dar una expresión explícita para este δ aunque dependa de n. Comenzaremos con el caso $n = 2$.

Teorema 4.8. *Sea $\epsilon > 0$. Entonces $\delta_F = 2\sqrt{2}\epsilon^2$ cumple que, para cualquier matriz $A \in \mathcal{M}_2(\mathbb{C})$ que satisfaga $\|A^*A - AA^*\|_F < \delta_F$ existe $N \in \mathcal{N}$ tal que $\|A - N\|_F < \epsilon$.*

Demostración. Sea $A \in \mathcal{M}_2(\mathbb{C})$ tal que $\|A^*A - AA^*\|_F < \delta_F$. Entonces, por el Teorema de Schur 1.16 existe U unitaria tal que $T := U^*AU$ es triangular superior, digamos $T = \begin{pmatrix} \lambda & x \\ 0 & \mu \end{pmatrix}$. Por lo tanto, como la norma de Frobenius es unitarimente invariante (Proposición 1.32), se cumple:

$$\begin{aligned} \|A^*A - AA^*\|_F &= \|U^*(A^*A - AA^*)U\|_F \\ &= \|T^*T - TT^*\|_F \\ &= \left\| \begin{pmatrix} -|x|^2 & x\overline{(\lambda-\mu)} \\ \overline{x}(\lambda-\mu) & |x|^2 \end{pmatrix} \right\|_F \\ &= \sqrt{2}|x|\sqrt{|x|^2 + |\lambda-\mu|^2}. \end{aligned}$$

Luego, como $\text{dist}_F(A, \mathcal{N}) = \text{dist}_F(T, \mathcal{N}) = \frac{|x|}{\sqrt{2}}$ (Sección 2.3) tenemos

$$\text{dist}_F(A,\mathcal{N})^2 = \frac{|x|^2}{2} \leq \frac{\sqrt{2}|x|\sqrt{|x|^2 + |\lambda-\mu|^2}}{\sqrt{2}\cdot 2} = \frac{\|A^*A - AA^*\|_F}{\sqrt{2}\cdot 2} < \frac{\delta_F}{\sqrt{2}\cdot 2} = \epsilon^2.$$

El resultado se sigue de notar que $\mathcal{N}$ es cerrado (Teorema 1.57), pues por el Lema 1.54 existe $N \in \mathcal{N}$ tal que $\|A - N\|_F = \mathrm{dist}_F(A, \mathcal{N}) < \epsilon$. ■

A continuación presentamos la versión de este teorema con la norma usual de operadores.

Teorema 4.9. *Sea $\epsilon > 0$. Entonces $\delta_{op} = 4\epsilon^2$ cumple que, para cualquier matriz $A \in \mathcal{M}_2(\mathbb{C})$ que satisfaga $\|A^*A - AA^*\| < \delta_{op}$ existe $N \in \mathcal{N}$ tal que $\|A - N\| < \epsilon$.*

Demostración. Análogamente a la prueba anterior, y como la norma usual de operadores es unitarimente invariante (Proposición 1.41), se cumple:

$$\|A^*A - AA^*\| \;=\; \left\| \begin{pmatrix} -|x|^2 & x(\overline{\lambda - \mu}) \\ \overline{x}(\lambda - \mu) & |x|^2 \end{pmatrix} \right\| = |x|\sqrt{|x|^2 + |\lambda - \mu|^2}.$$

Luego, como $\mathrm{dist}(A, \mathcal{N}) = \mathrm{dist}(T, \mathcal{N}) = \frac{|x|}{2}$ (Sección 2.3) tenemos

$$\mathrm{dist}(A, \mathcal{N})^2 = \frac{|x|^2}{4} \leq \frac{|x|\sqrt{|x|^2 + |\lambda - \mu|^2}}{4} = \frac{\|A^*A - AA^*\|}{4} < \frac{\delta_{op}}{4} = \epsilon^2.$$

De manera análoga a la prueba anterior, el resultado se sigue del Teorema 1.57 y del Lema 1.54. ■

El Teorema 4.8 se puede demostrar para un δ diferente usando los resultados que obtuvimos en el Capítulo 3 en lugar de los del Capítulo 2. Más precisamente: dado $\epsilon > 0$, por el Teorema 3.5 se sigue que $\delta_H = \sqrt{2}\epsilon^2$ cumple que para cualquier $A \in \mathcal{M}_2(\mathbb{C})$ con $\|A^*A - AA^*\|_F < \delta_H$,

$$\mathrm{dist}_F(A, \mathcal{N}) \leq \left(\frac{2^3 - 2}{12}\right)^{\frac{1}{4}} \sqrt{\|A^*A - AA^*\|_F} < \left(\frac{1}{2}\right)^{\frac{1}{4}} \sqrt{\delta_H} = \epsilon.$$

Observemos que $\delta_H < \delta_F$. Por lo tanto δ_F ofrece una mejor estimación.

Terminamos este capítulo mostrando los resultados principales de este trabajo. Vamos a verificar que una matriz casi normal de tamaño general está cerca de una matriz normal y presentaremos expresiones explícitas para δ.

Teorema 4.10. *Sea $n > 1$ un número entero y $\epsilon > 0$. Entonces $\delta = \epsilon^2 \left(\frac{n^3 - n}{12}\right)^{-\frac{1}{2}}$ cumple que, para cualquier $A \in \mathcal{M}_n(\mathbb{C})$ que satisfaga $\|A^*A - AA^*\|_F < \delta$ existe $N \in \mathcal{N}$ tal que $\|A - N\|_F < \epsilon$.*

Demostración. Sea $A \in \mathcal{M}_n(\mathbb{C})$ tal que $\|A^*A - AA^*\|_F < \delta$. Por el Teorema 3.5

$$\mathrm{dist}_F(A, \mathcal{N}) \leq \left(\frac{n^3 - n}{12}\right)^{\frac{1}{4}} \sqrt{\|A^*A - AA^*\|_F} < \left(\frac{n^3 - n}{12}\right)^{\frac{1}{4}} \sqrt{\delta} = \epsilon.$$

Luego, como $\mathcal{N}$ es cerrado (Teorema 1.57), por el Lema 1.54 existe $N \in \mathcal{N}$ tal que $\|A - N\|_F = \text{dist}_F(A, \mathcal{N}) < \epsilon$. ■

A continuación presentamos la versión de este teorema con la norma usual de operadores.

Teorema 4.11. *Sea $n > 1$ un número entero y $\epsilon > 0$. Entonces $\delta = \epsilon^2 \left(\frac{n(n^3-n)}{12}\right)^{-\frac{1}{2}}$ cumple que, para cualquier $A \in \mathcal{M}_n(\mathbb{C})$ que satisfaga $\|A^*A - AA^*\| < \delta$ existe $N \in \mathcal{N}$ tal que $\|A - N\| < \epsilon$.*

Demostración. Sea $A \in \mathcal{M}_n(\mathbb{C})$ tal que $\|A^*A - AA^*\| < \delta$. Por el Teorema 3.6

$$\text{dist}(A, \mathcal{N}) \leq \left(\frac{n(n^3-n)}{12}\right)^{\frac{1}{4}} \sqrt{\|A^*A - AA^*\|} < \left(\frac{n(n^3-n)}{12}\right)^{\frac{1}{4}} \sqrt{\delta} = \epsilon.$$

Luego, como $\mathcal{N}$ es cerrado (Teorema 1.57), por el Lema 1.54 existe $N \in \mathcal{N}$ tal que $\|A - N\| = \text{dist}(A, \mathcal{N}) < \epsilon$. ■

En general, usando el Teorema 1.52 tenemos el siguiente resultado.

Teorema 4.12. *Sea $n > 1$ un entero, y sea $\|\cdot\|_1$ una norma definida en $\mathcal{M}_n(\mathbb{C})$. Entonces dado $\epsilon > 0$, existe $c > 0$ (el cual puede depender de n) tal que $\delta = \dfrac{\epsilon^2}{c^2}\left(\frac{n^3-n}{12}\right)^{-\frac{1}{2}}$ cumple que, para todo $A \in \mathcal{M}_n(\mathbb{C})$ que satisfaga $\|A^*A - AA^*\|_1 < \delta$ tenemos $\text{dist}_{\|\cdot\|_1}(A, \mathcal{N}) < \epsilon$. Más aún, si $\|X^*\|_1 = \|X\|_1$ para todo $X \in \mathcal{M}_n(\mathbb{C})$, entonces existe $N \in \mathcal{N}$ con $\|A - N\|_1 < \epsilon$.*

Demostración. Sea $A \in \mathcal{M}_n(\mathbb{C})$ tal que $\|A^*A - AA^*\|_1 < \delta$. Entonces por el Teorema 3.7, existe $c > 0$ tal que

$$\text{dist}_{\|\cdot\|_1}(A, \mathcal{N}) \leq c\left(\frac{n^3-n}{12}\right)^{\frac{1}{4}} \sqrt{\|A^*A - AA^*\|_1} < c\left(\frac{n^3-n}{12}\right)^{\frac{1}{4}} \sqrt{\delta} = \epsilon.$$

Notemos que si además $\|X^*\|_1 = \|X\|_1$ para todo $X \in \mathcal{M}_n(\mathbb{C})$, entonces $\mathcal{N}$ es cerrado (Teorema 1.57). Así, por el Lema 1.54 existe $N \in \mathcal{N}$ tal que $\|A - N\|_1 = \text{dist}_{\|\cdot\|_1}(A, \mathcal{N}) < \epsilon$. ■

Apéndice

Como vimos en el Capítulo 1, el conjunto $\mathcal{N}$ es cerrado con la métrica dada por cualquier norma $\|\cdot\|$ que satisfaga $\|X^*\| = \|X\|$, para todo $X \in \mathcal{M}_n(\mathbb{C})$ (Teorema 1.57). Esta hipótesis también es necesaria para la generalización de la observación de Higham [7] (Proposición 4.4), y para que se alcance la distancia en el Teorema 4.12. En este Apéndice vamos a caracterizar algunas normas provenientes de un cierto producto interno que satisfacen esta condición.

En general, se sabe que dado un producto interno $\langle X, Y\rangle_1$ definido en $\mathcal{M}_n(\mathbb{C})$, existe una única transformación lineal $\mathcal{A} : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C})$ tal que $\langle X, Y\rangle_1 = \langle \mathcal{A}(X), Y\rangle = \operatorname{traza}(Y^*\mathcal{A}(X))$, para todo $X, Y \in \mathcal{M}_n(\mathbb{C})$ (ver [2, pág. 31]).

En particular, se puede verificar que $\mathcal{A} : \mathcal{M}_n(\mathbb{C}) \to \mathcal{M}_n(\mathbb{C})$ definida como $\mathcal{A}(X) := \sum_i B_i X C_i$ puede definir un producto interno en $\mathcal{M}_n(\mathbb{C})$ restringiendo las condiciones sobre los B_i y C_i.

En este Apéndice vamos a mostrar las condiciones que debe satisfacer la transformación lineal de la forma $\mathcal{A}(X) := BXC$ para que $\langle X, Y\rangle_1 = \operatorname{traza}(Y^*\mathcal{A}(X))$ defina un producto interno y para que $\|\cdot\|_1 := \sqrt{\langle \cdot, \cdot\rangle_1}$ cumpla $\|X^*\|_1 = \|X\|_1$, para todo $X \in \mathcal{M}_n(\mathbb{C})$. El caso general involucra teoría de representaciones de productos tensoriales en algebras C^* y está fuera del alcance de este trabajo (ver [12]).

Definición A.1. Decimos que $A \in \mathcal{M}_n(\mathbb{C})$ es una matriz *positiva definida*, si es autoadjunta y todos sus eigenvalores son estrictamente positivos.

Proposición A.2. *Sean B y C matrices positivas definidas. Entonces la función $\langle X, Y\rangle_1 := \operatorname{traza}(Y^*BXC)$ para $X, Y \in \mathcal{M}_n(\mathbb{C})$, define un producto interno en $\mathcal{M}_n(\mathbb{C})$.*

Demostración. Verificaremos que las condiciones de la Definición 1.8 se satisfacen.

Con un cálculo podemos ver que $\langle \alpha X + \beta Y, Z\rangle_1 = \alpha \langle X, Z\rangle_1 + \beta \langle Y, Z\rangle_1$, para todo $X, Y, Z \in \mathcal{M}_n(\mathbb{C})$ y $\alpha, \beta \in \mathbb{C}$.

Como B y C son matrices positivas definidas y por lo tanto autoadjuntas, por el Lema 1.21 se cumple $\overline{\langle Y, X\rangle_1} = \overline{\text{traza}(X^*BYC)} = \text{traza}(C^*Y^*B^*X) = \text{traza}(CY^*BX) = \text{traza}(Y^*BXC) = \langle X, Y\rangle_1$.

Por último, notemos que el Teorema Espectral 1.19 indica la existencia de matrices unitarias U y V tales que $B' := U^*BU$ y $C' := V^*CV$ son diagonales. Así, por el Lema 1.21 se cumple

$$\begin{aligned}
\langle X, X\rangle_1 &= \text{traza}(X^*BXC) \\
&= \text{traza}(X^*UB'U^*XVC'V^*) \\
&= \text{traza}((V^*X^*U)B'(U^*XV)C') \\
&= \text{traza}((U^*XV)^*B'(U^*XV)C') \\
&= \text{traza}(Y^*B'YC'),
\end{aligned}$$

donde $Y = U^*XV$. Luego, como

$$\begin{aligned}
Y^*B'YC' &= \begin{pmatrix} \bar{y}_{11} & \bar{y}_{21} & \cdots & \bar{y}_{n1} \\ \bar{y}_{12} & \bar{y}_{22} & \cdots & \bar{y}_{n2} \\ \vdots & \vdots & & \vdots \\ \bar{y}_{1n} & \bar{y}_{2n} & \cdots & \bar{y}_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \\
& \quad \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{pmatrix} \begin{pmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_n \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1\bar{y}_{11} & \lambda_2\bar{y}_{21} & \cdots & \lambda_n\bar{y}_{n1} \\ \lambda_1\bar{y}_{12} & \lambda_2\bar{y}_{22} & \cdots & \lambda_n\bar{y}_{n2} \\ \vdots & \vdots & & \vdots \\ \lambda_1\bar{y}_{1n} & \lambda_2\bar{y}_{2n} & \cdots & \lambda_n\bar{y}_{nn} \end{pmatrix} \begin{pmatrix} \mu_1 y_{11} & \mu_2 y_{12} & \cdots & \mu_n y_{1n} \\ \mu_1 y_{21} & \mu_2 y_{22} & \cdots & \mu_n y_{2n} \\ \vdots & \vdots & & \vdots \\ \mu_1 y_{n1} & \mu_2 y_{n2} & \cdots & \mu_n y_{nn} \end{pmatrix},
\end{aligned}$$

tenemos

$$\begin{aligned}
\text{traza}(Y^*B'YC') = \; & \lambda_1\mu_1|y_{11}|^2 + \lambda_2\mu_1|y_{21}|^2 + \cdots + \lambda_n\mu_1|y_{n1}|^2 + \\
& \lambda_1\mu_2|y_{12}|^2 + \lambda_2\mu_2|y_{22}|^2 + \cdots + \lambda_n\mu_2|y_{n2}|^2 + \\
& + \cdots + \\
& \lambda_1\mu_n|y_{1n}|^2 + \lambda_2\mu_n|y_{2n}|^2 + \cdots + \lambda_n\mu_n|y_{nn}|^2.
\end{aligned}$$

El resultado se sigue de notar que como B y C son positivas definidas por hipótesis, se tiene $\lambda_i > 0$ y $\mu_j > 0$ y por lo tanto $\text{traza}(Y^*B'YC') \geq 0$. Esto implica que $\langle X, X\rangle_1 \geq 0$. Notemos también que $\langle X, X\rangle_1 = 0$ si y sólo si $y_{ij} = 0$ para todo $i, j \in \{1, 2, ..., n\}$, pero esto ocurre si y sólo si $X = 0$. ∎

Como veremos en los siguientes lemas, estos son básicamente los únicos ejemplos de transformaciones lineales de la forma $\mathcal{A}(X) = BXC$ que definen un producto interno en $\mathcal{M}_n(\mathbb{C})$.

Notación A.3. Denotaremos por E_{ij} a la matriz en $\mathcal{M}_n(\mathbb{C})$ con 1 en la entrada ij y cero en las demás entradas.

Lema A.4. *Sea* $\langle X,Y\rangle_1 := \text{traza}(Y^*BXC)$ *para* $X,Y\in\mathcal{M}_n(\mathbb{C})$ *una función tal que* $\overline{\langle Y,X\rangle_1} = \langle X,Y\rangle_1$ *para todo* $X,Y\in\mathcal{M}_n(\mathbb{C})$*. Entonces* B *y* C *se pueden elegir autoadjuntas.*

Demostración. Primero notemos que por hipótesis y por el Lema 1.21 se cumple

$$\begin{aligned} 0 &= \overline{\langle Y,X\rangle_1} - \langle X,Y\rangle_1 \\ &= \text{traza}(Y^*[B^*XC^* - BXC]) \\ &= \langle B^*XC^* - BXC, Y\rangle, \end{aligned}$$

donde $\langle\cdot,\cdot\rangle$ es el producto interno del Lema 1.26. Como esto ocurre para todo $Y\in\mathcal{M}_n(\mathbb{C})$, necesariamente $B^*XC^* - BXC = 0$.

De esta manera, si $X = E_{ij}$, entonces

$$\begin{aligned} BE_{ij} &= \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix} \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1_{ij} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & \cdots & b_{1i} & \cdots & 0 \\ 0 & 0 & \cdots & b_{2i} & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & b_{ni} & \cdots & 0 \end{pmatrix}; \end{aligned}$$

donde la columna distinta de cero es la columna j. Así,

$$\begin{aligned} BE_{ij}C &= \begin{pmatrix} 0 & 0 & \cdots & b_{1i} & \cdots & 0 \\ 0 & 0 & \cdots & b_{2i} & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & b_{ni} & \cdots & 0 \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix} \\ &= \begin{pmatrix} b_{1i}c_{j1} & b_{1i}c_{j2} & \cdots & b_{1i}c_{jn} \\ b_{2i}c_{j1} & b_{2i}c_{j2} & \cdots & b_{2i}c_{jn} \\ \vdots & \vdots & & \vdots \\ b_{ni}c_{j1} & b_{ni}c_{j2} & \cdots & b_{ni}c_{jn} \end{pmatrix}, \end{aligned}$$

es decir, la entrada st de $BE_{ij}C$ es $b_{si}c_{jt}$. De manera análoga, la entrada st de $B^*E_{ij}C^*$ es $\bar{b}_{is}\bar{c}_{tj}$.

Así, si $B^*E_{ij}C^* = BE_{ij}C$, entonces $\bar{b}_{is}\bar{c}_{tj} = b_{si}c_{jt}$. Luego, como $C \neq 0$ existe $c_{jt}\neq 0$ tal que podemos definir $e^{i\theta} := \dfrac{c_{jt}}{\bar{c}_{tj}}$. Por lo tanto $\bar{b}_{is} = e^{i\theta}b_{si}$ para todo $i,s\in\{1,2,...,n\}$, es decir, $B^* = e^{i\theta}B$. De manera análoga se tiene que $C^* = e^{-i\theta}C$.

El resultado se sigue de notar que si $\alpha = e^{i\frac{\theta}{2}}$, entonces αB y $\bar{\alpha}C$ son autoadjuntas y por lo tanto $\langle X,Y\rangle_1 = \text{traza}(Y^*BXC) = \text{traza}(Y^*(\alpha B)X(\bar{\alpha}C))$. ∎

Lema A.5. *Sean B y C matrices autoadjuntas, y sea $\langle X, Y\rangle_1 := \mathrm{traza}(Y^*BXC)$ para $X, Y \in \mathcal{M}_n(\mathbb{C})$ una función tal que $\langle X, X\rangle_1 \geq 0$ para todo $X \in \mathcal{M}_n(\mathbb{C})$ y $\langle X, X\rangle_1 = 0$ si y sólo si $X = 0$. Entonces las matrices B y C son positivas definidas o las matrices $-B$ y $-C$ son positivas definidas.*

Demostración. De manera análoga a la Proposición A.2, tenemos que existen matrices unitarias U y V tales que $B' := U^*BU$ y $C' := V^*CV$ son diagonales. Por lo cual, si $Y = U^*XV$

$$\begin{aligned}\langle X, X\rangle_1 &= \lambda_1\mu_1|y_{11}|^2 + \lambda_2\mu_1|y_{21}|^2 + \cdots + \lambda_n\mu_1|y_{n1}|^2 + \\ &\quad \lambda_1\mu_2|y_{12}|^2 + \lambda_2\mu_2|y_{22}|^2 + \cdots + \lambda_n\mu_2|y_{n2}|^2 + \\ &\quad + \cdots + \\ &\quad \lambda_1\mu_n|y_{1n}|^2 + \lambda_2\mu_n|y_{2n}|^2 + \cdots + \lambda_n\mu_n|y_{nn}|^2.\end{aligned}$$

Así, si $X \neq 0$, entonces $\langle X, X\rangle_1 > 0$ y por lo tanto $\lambda_i\mu_j > 0$ para todo $i, j \in \{1, 2, ..., n\}$. Esto implica que $\lambda_i \neq 0$ y $\mu_j \neq 0$.

Luego, si $\lambda_i > 0$ para algún i, entonces $\mu_j > 0$ para todo $j \in \{1, 2, ..., n\}$. Sin embargo, esto a su vez implica $\lambda_i > 0$ para todo $i \in \{1, 2, ..., n\}$ y por lo tanto que B y C son matrices positivas definidas. De manera análoga, si $\lambda_i < 0$ para algún i podemos concluir $\lambda_i < 0$ y $\mu_j < 0$ para todo $i, j \in \{1, 2, ..., n\}$ y por lo tanto que $-B$ y $-C$ son positivas definidas. ∎

Así, para que $\langle X, Y\rangle_1 := \mathrm{traza}(Y^*BXC)$ para $X, Y \in \mathcal{M}_n(\mathbb{C})$ defina un producto interno en $\mathcal{M}_n(\mathbb{C})$, las matrices B y C se pueden escoger positivas definidas. Esto es debido a que si $-B$ y $-C$ son positivas definidas, entonces $\langle X, Y\rangle_1 = \mathrm{traza}(Y^*BXC) = \mathrm{traza}(Y^*(-B)X(-C))$.

Concluiremos este apéndice mostrando las condiciones que deben cumplir B y C para que las normas $\|\cdot\|_1 := \sqrt{\langle\cdot,\cdot\rangle_1}$ cumplan $\|X^*\|_1 = \|X\|_1$ para todo $X \in \mathcal{M}_n(\mathbb{C})$.

Teorema A.6. *Sea $\langle X, Y\rangle_1 := \mathrm{traza}(Y^*BXC)$ para $X, Y \in \mathcal{M}_n(\mathbb{C})$ un producto interno en $\mathcal{M}_n(\mathbb{C})$, y sea $\|\cdot\|_1 := \sqrt{\langle\cdot,\cdot\rangle_1}$. Entonces $\|X^*\|_1 = \|X\|_1$ para todo $X \in \mathcal{M}_n(\mathbb{C})$ si y sólo si $C = \gamma B$ (con γ necesariamente positivo).*

Demostración. Observemos que si $C = \gamma B$ para algún $\gamma > 0$, entonces por el Lema 1.21 tenemos que para todo $X \in \mathcal{M}_n(\mathbb{C})$ se cumple:

$$\begin{aligned}\|X^*\|_1^2 &= \mathrm{traza}(X^*BX(\gamma B)) \\ &= \mathrm{traza}(X(\gamma B)X^*B) \\ &= \mathrm{traza}(XBX^*(\gamma B)) \\ &= \|X\|_1^2.\end{aligned}$$

Ahora supongamos que se cumple $\|X^*\|_1 = \|X\|_1$ para todo $X \in \mathcal{M}_n(\mathbb{C})$. Entonces para $X = E_{jk}$ tenemos $\|E_{jk}\|_1^2 = \mathrm{traza}(E_{jk}^*BE_{jk}C) = \mathrm{traza}(E_{kj}BE_{jk}C)$, donde

$$
\begin{aligned}
E_{kj}B &= \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1_{kj} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ b_{j1} & b_{j2} & \cdots & b_{jn} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.
\end{aligned}
$$

Notemos que la fila distinta de cero de esta matriz es la fila k. Análogamente, tenemos que la fila distinta de cero de la matriz $E_{jk}C$ es la fila j, y por lo tanto

$$
\begin{aligned}
E_{kj}BE_{jk}C &= \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ b_{j1} & b_{j2} & \cdots & b_{jn} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ c_{k1} & c_{k2} & \cdots & c_{kn} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ b_{jj}c_{k1} & b_{jj}c_{k2} & \cdots & b_{jj}c_{kn} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},
\end{aligned}
$$

donde la fila distinta de cero es la fila k. Así $\text{traza}(E_{kj}BE_{jk}C) = b_{jj}c_{kk}$. De manera análoga se muestra que

$$\|E_{jk}^*\|_1^2 = \text{traza}(E_{jk}BE_{jk}^*C) = \text{traza}(E_{jk}BE_{kj}C) = b_{kk}c_{jj}.$$

Luego, como $\|E_{jk}^*\|_1 = \|E_{jk}\|_1$ para todo $j, k \in \{1, 2, ..., n\}$, tenemos

$$b_{kk}c_{jj} = b_{jj}c_{kk}, \text{ para todo } j, k \in \{1, 2, ..., n\}. \tag{A.1}$$

Por otro lado observemos que si $X = E_{jk} + \alpha E_{st}$ para algún $\alpha \in \mathbb{C}$, entonces

$$
\begin{aligned}
\|E_{jk} + \alpha E_{st}\|_1^2 &= \text{traza}((E_{jk} + \alpha E_{st})^* B(E_{jk} + \alpha E_{st})C) \\
&= \text{traza}((E_{kj} + \bar{\alpha} E_{ts})B(E_{jk} + \alpha E_{st})C),
\end{aligned}
$$

donde

$$(E_{kj}+\bar{\alpha}E_{ts})B = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1_{kj} & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & \bar{\alpha}1_{ts} & \cdots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ b_{j1} & b_{j2} & \cdots & b_{jn} \\ \vdots & \vdots & & \vdots \\ \bar{\alpha}b_{s1} & \bar{\alpha}b_{s2} & \cdots & \bar{\alpha}b_{sn} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Notemos que las filas distintas de cero de esta matriz son las filas k y t. Análogamente, tenemos que las filas distintas de cero de la matriz $(E_{jk}+\alpha E_{st})C$ son las filas j y s. Así,

$$(E_{kj}+\bar{\alpha}E_{ts})B(E_{jk}+\alpha E_{st})C = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ b_{j1} & b_{j2} & \cdots & b_{jn} \\ \vdots & \vdots & & \vdots \\ \bar{\alpha}b_{s1} & \bar{\alpha}b_{s2} & \cdots & \bar{\alpha}b_{sn} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ c_{k1} & c_{k2} & \cdots & c_{kn} \\ \vdots & \vdots & & \vdots \\ \alpha c_{t1} & \alpha c_{t2} & \cdots & \alpha c_{tn} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ b_{jj}c_{k1}+\alpha b_{js}c_{t1} & b_{jj}c_{k2}+\alpha b_{js}c_{t2} & \cdots & b_{jj}c_{kn}+\alpha b_{js}c_{tn} \\ \vdots & \vdots & & \vdots \\ \bar{\alpha}b_{sj}c_{k1}+|\alpha|^2 b_{ss}c_{t1} & \bar{\alpha}b_{sj}c_{k2}+|\alpha|^2 b_{ss}c_{t2} & \cdots & \bar{\alpha}b_{sj}c_{kn}+|\alpha|^2 b_{ss}c_{tn} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

donde las filas distintas de cero son las filas k y t. Por lo tanto,

$$\operatorname{traza}((E_{kj}+\bar{\alpha}E_{ts})B(E_{jk}+\alpha E_{st})C) = b_{jj}c_{kk}+\alpha b_{js}c_{tk}+\bar{\alpha}b_{sj}c_{kt}+|\alpha|^2 b_{ss}c_{tt}.$$

De manera análoga se puede mostrar

$$\begin{aligned}\|(E_{jk}+\alpha E_{st})^*\|_1^2 &= \text{traza}((E_{jk}+\alpha E_{st})B(E_{kj}+\bar{\alpha}E_{ts})C)\\ &= b_{kk}c_{jj}+\bar{\alpha}b_{kt}c_{sj}+\alpha b_{tk}c_{js}+|\alpha|^2 b_{tt}c_{ss}.\end{aligned}$$

Esto nos permite concluir que, como $\|(E_{jk}+\alpha E_{st})^*\|_1 = \|E_{jk}+\alpha E_{st}\|_1$, se tiene

$$b_{kk}c_{jj}+\bar{\alpha}b_{kt}c_{sj}+\alpha b_{tk}c_{js}+|\alpha|^2 b_{tt}c_{ss} = b_{jj}c_{kk}+\alpha b_{js}c_{tk}+\bar{\alpha}b_{sj}c_{kt}+|\alpha|^2 b_{ss}c_{tt}.$$

Luego, como $b_{kk}c_{jj}+|\alpha|^2 b_{tt}c_{ss} = b_{jj}c_{kk}+|\alpha|^2 b_{ss}c_{tt}$ por la ecuación (A.1), entonces

$$\bar{\alpha}b_{kt}c_{sj}+\alpha b_{tk}c_{js} = \alpha b_{js}c_{tk}+\bar{\alpha}b_{sj}c_{kt}.$$

Además como B y C son autoadjuntas y por lo tanto $b_{kt}=\bar{b}_{tk}$ y $c_{sj}=\bar{c}_{js}$ para todo $j,k,s,t\in\{1,2,...,n\}$, tenemos

$$\text{Re}(\alpha b_{tk}c_{js}) = \text{Re}(\alpha b_{js}c_{tk}) \text{ para todo } \alpha\in\mathbb{C},$$

lo cual implica que $b_{tk}c_{js}=b_{js}c_{tk}$.

Como $B\neq 0$ existe $b_{tk}\neq 0$ y por lo tanto podemos definir $\gamma := \dfrac{c_{tk}}{b_{tk}}$. De aquí se sigue la relación $c_{js}=\gamma b_{js}$ para todo $s,t\in\{1,2,...,n\}$, es decir, $C=\gamma B$.

Por último notemos que, como B y C son autoadjuntas y distintas de cero

$$\bar{\gamma}B = C^* = C = \gamma B.$$

Esto implica que $\bar{\gamma}=\gamma$ y por lo tanto que $\gamma\in\mathbb{R}$. El resultado se sigue de observar que, como B y C son positivas definidas necesariamente $\gamma>0$. ∎

Este teorema nos permite concluir que, para que la norma $\|\cdot\|_1 := \sqrt{\langle\cdot,\cdot\rangle_1}$ cumpla $\|X^*\|_1=\|X\|_1$ para todo $X\in\mathcal{M}_n(\mathbb{C})$, las matrices B y C se pueden escoger esencialmente iguales, pues $\|X\|_1^2 = \text{traza}(Y^*BXC) = \text{traza}(Y^*BX(\gamma B)) = \text{traza}(Y^*(\sqrt{\gamma}B)X(\sqrt{\gamma}B))$. Esto nos produce una familia de ejemplos que cumplen la hipótesis principal del Teorema 1.57, de la Proposición 4.4 y del Teorema 4.12.

BIBLIOGRAFÍA

[1] Tom M. Apostol, *Cálculus Vol. 2. Cálculo con Funciones de Varias Variables y Algebra Lineal, con Aplicaciones a las Ecuaciones Diferenciales y las Probabilidades*, Reverté (2006).

[2] J. B. Conway, *A Course in Functional Analysis*, Springer (1990).

[3] P. Friis, *Normal Elements with Finite Spectrum In C^*-Algebras Of Real Rank Zero*, Ph.D. Dissertation, University of Toronto (1999).

[4] P. Friis and M. Rørdam, *Almost Commuting Self-adjoint Matrices, a Short Proof of Huaxin Lin's Theorem*, J. Reine Angew. Math **479** (1996), 121-131.

[5] P. R. Halmos, *Some Unsolved Problems of Unknown Depth About Operators on Hilbert Spaces*, Proceedings of the Royal Society of Edinburgh **76A** (1976), 67-76.

[6] P. Henrici, *Bounds for Iterates, Inverses, Spectral Variation and Fields of Values of Non-normal Matrices*, Numer. Math. **4** (1962), 24-40.

[7] N. J. Higham, *Matrix Nearness Problems and Applications*, Applications of matrix theory (Bradford, 1988), 1-27, Inst. Math. Appl. Conf. Ser. New Ser., 22, Oxford Univ. Press, New York (1989).

[8] H. Lin, *Almost Commuting Self-adjoint Matrices and Applications*, Operator algebras and their applications (Waterloo, ON, 1994/1995),193-233, Fields Inst. Commun. 13, Amer. Math. Soc., Providence, RI (1997).

[9] W. K. Nicholson, *Linear Algebra with Applications*, PWS Publishing Company Toronto (1995).

[10] P. Rosenthal, *Are Almost Commuting Matrices Near Commuting Matrices?*, The American Mathematical Monthly **76** (1969), 925-926.

[11] W. Rudin, *Real and Complex Analysis*, Mc. Graw Hill (1987).

[12] N. E. Wegge-Olsen, *K-Theory and C^* Algebras*, Oxford Univ. Press (1993).