



UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
ESCUELA PROFESIONAL DE MATEMÁTICA



“Análisis cualitativo del espectro de matrices perturbadas”

Tesis

Para optar el título profesional de

LICENCIADA EN MATEMÁTICAS

Autoras:

BACH.MAT. CAROLINA MINIET CAJO CASTILLO

BACH.MAT. LOURDES DEL PILAR VÁSQUEZ AGUIRRE

Asesor:

M.SC. OSCAR ANTONIO SANTAMARIA SANTISTEBAN

Lambayeque, 07 de noviembre de 2017



UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
ESCUELA PROFESIONAL DE MATEMÁTICA



Los firmantes, por la presente certifican que han leído y recomiendan a la Facultad de Ciencias Físicas y Matemáticas la aceptación de la tesis titulada “Análisis cualitativo del espectro de matrices perturbadas”, presentada por las Bachilleres en Matemáticas: Carolina Miniet Cajo Castillo y Lourdes del Pilar Vásquez Aguirre, en el cumplimiento parcial de los requisitos necesarios para la obtención del título profesional de Licenciado en Matemáticas.

Lic.Mat. Juana Doris Blas Rebaza
Presidenta del Jurado

Lic.Mat. Marco Antonio Martín Peralta Lui
Secretario del Jurado

Lic. Rolando Javier Córdova Descalzi
Vocal del Jurado

Fecha de defensa: Noviembre – 2017



UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
ESCUELA PROFESIONAL DE MATEMÁTICA



TESIS:

**ANÁLISIS CUALITATIVO DEL ESPECTRO DE MATRICES
PERTURBADAS.**

Bach.Mat. Carolina Miniét Cajo Castillo

Autora de la tesis

Bach.Mat. Lourdes del Pilar Vásquez Aguirre

Autora de la tesis

M.Sc. Oscar Antonio Santamaria Santisteban

Asesor

Fecha de defensa: Noviembre – 2017



UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
ESCUELA PROFESIONAL DE MATEMÁTICA



1. Título de tesis

Análisis Cualitativo del Espectro de Matrices Perturbadas.

2. Resolución de aprobación de tesis

N° 727 - 2016 - D/FACFyM

3. Tipo de Investigación

Aplicada

4. Área de Investigación

Ciencias Exactas: Matemática Aplicada.

5. Línea de Investigación

Álgebra matricial y Análisis numérico.

6. Fecha de Inicio

01 de Febrero del 2016.

7. Fecha de Finalización

27 de Marzo del 2017.

Cajo Castillo Carolina Miniét
Código : 101610 J

Vásquez Aguirre Lourdes del Pilar
Código : 084018 D

M. Sc. Oscar A. Santamaria Santisteban
Asesor

Dedicatoria

A Dios.

Por haberme permitido llegar hasta este punto y haberme dado salud para lograr mis objetivos.

A mis padres.

Por haberme apoyado en todo momento, por sus consejos, sus valores, por la motivación constante que me ha permitido ser una persona de bien , pero más que nada por su amor.

Carolina.

A Dios.

Por darme la oportunidad de vivir y por estar conmigo en cada paso que doy.

A mis padres.

Por ser el pilar fundamental en todo lo que soy, en toda mi educación, tanto académica como en la vida.

Lourdes.

Agradecimiento

Primero damos gracias a Dios por permitirnos tener tan buena experiencia dentro de la universidad, a nuestros padres, y a nuestro asesor *Mg. Oscar Santamaría Santisteban* por su esfuerzo y dedicación, gracias a nuestra alma mater por convertirnos en ser unas profesionales en lo que tanto nos apasiona, gracias a cada maestro que hizo parte de este proceso integral de formación, y como recuerdo y prueba viviente en la historia; esta tesis, que perdurará dentro de los conocimientos y desarrollo de las demás generaciones que están por llegar.

Finalmente agradecemos a quienes leen este apartado y más de nuestra tesis, por permitir a nuestras investigaciones y conocimientos, incurrir dentro de su repertorio de información mental.

Resumen

En el presente estudio se analiza la relación entre los autovalores de una matriz dada $A \in \mathbb{C}^{n \times n}$ y su matriz perturbada $A + E \in \mathbb{C}^{n \times n}$. Dicha relación está expresada en los teoremas de Elsner y de Hoffman-Wielandt, en donde además se proporcionan cotas para la perturbación E , dependiendo de la clase de matriz $A \in \mathbb{C}^{n \times n}$. El trabajo se completa detallando una aplicación a un sistema de ranking desarrollado por Google para medir la importancia de una página web, llamada PageRank.

Contenido

Introducción	1
1. Preliminares	5
1.1. Autovalores y Autovectores	5
1.2. Norma de matrices.	6
1.3. Número de condición de matrices. Propiedades.	8
1.4. Algunas matrices especiales. Propiedades.	9
1.4.1. Matriz Diagonalizable	10
1.4.2. Matriz Unitaria	11
1.4.3. Matriz Normal	14
1.4.4. Matriz hermitiana	14
1.4.5. Matriz Antihermitiana	17
1.5. Forma canónica de Jordan.	19
1.6. Algunos resultados adicionales	21
2. Perturbación de matrices	33
2.1. Algunos resultados generales sobre matrices	33
2.1.1. Teoría general de perturbaciones	34
2.2. Algunos resultados de matrices normales.	40
2.3. Matriz normal con perturbación general.	49
2.4. Matriz hermitiana con perturbación antihermitiana.	52
3. Aplicaciones	57
3.1. Page Rank	57
3.1.1. PageRank and the Google Matrix	58
3.1.2. Propiedades de la matriz S	59
3.1.3. Matriz Google	59
3.1.4. Características de la matriz G	60

3.1.5. Estabilidad del PageRank	60
3.1.6. Ejemplo aplicativo	63
Conclusiones	71
Sugerencias	73
Bibliografía	75

Introducción

Esta tesis trata sobre el análisis de perturbaciones para los problemas de autovalores de matrices. La parte más desafiante desde el punto de vista computacional de muchos problemas es a menudo resolver ecuaciones matriciales a gran escala, los más frecuentes son: sistemas lineales de ecuaciones y problemas de autovalores. El rápido desarrollo de la tecnología informática hace posible simulaciones y cálculos de escala cada vez mayor. Es vital que los algoritmos sean adecuados para que la arquitectura de la computación haga pleno uso de la potencia computacional y que entendamos la teoría subyacente para asegurar que las cantidades calculadas sean significativas y confiables. Estas dos cuestiones son la motivación fundamental para este trabajo de tesis.

La teoría de perturbación es una herramienta importante en el análisis numérico, que es de vital utilidad en análisis de estabilidad, estimaciones de errores y desarrollos de algoritmos. Aunque la teoría de perturbación es un tema bien estudiado, todavía hay una serie de fenómenos que no pueden explicarse por los resultados existentes. En esta tesis tratamos de llenar algunas de estas brechas haciendo contribuciones a varios temas en la teoría de perturbación de los autovalores como el análisis de cada tipo de matriz y que perturbaciones se adaptan a su naturaleza. La comprensión de estos fenómenos puede darnos nuevos conocimientos sobre el comportamiento inexplicable de los algoritmos numéricos, proporcionando direcciones para las mejoras.

Además de ser de interés teórico, la teoría de perturbaciones es una herramienta útil que desempeña papeles importantes en muchas aplicaciones. Por ejemplo, se emplea frecuentemente en el análisis de estabilidad de un algoritmo numérico, para examinar si un problema dado está bien condicionado bajo la perturbación, o para errores de delimitación de una solución calculada. Sin embargo, hay una serie de fenómenos que todavía no pueden ser explicados por la teoría existente.

Este trabajo se encuentra organizado en tres capítulos, cuyo contenido se describe a continuación.

En el capítulo I se presentan algunos conceptos del Álgebra lineal, número de condición de matrices, la forma canónica de Jordan y algunos resultados especiales.

En el capítulo II se desarrolla la teoría de perturbación de matrices para cuatro casos: matriz general con perturbación general, matriz normal con perturbación normal, matriz normal con perturbación general y por último matriz hermitiana con perturbación antihermitiana.

En el capítulo III se detalla una aplicación a un sistema de ranking desarrollado por Google para medir la autoridad de una página web, basándose en la cantidad y calidad de los enlaces que apuntan hacia ella, llamada PageRank.

Notación

M_n	el conjunto de matrices complejas de orden $n \times n$
$M_{m,n}$	el conjunto de matrices complejas de orden $m \times n$
$A(i, j)$	entrada de la matriz A en la i -ésima fila y la j -ésima columna
A^T	transpuesta de la matriz A
A^*	conjugada transpuesta de la matriz A
$\mathcal{L}(A)$	espectro de la matriz A , es decir, el conjunto de todos los autovalores de A
$\rho(A)$	radio espectral
$\ \cdot\ $	norma en el espacio vectorial
$\ A\ _\infty$	norma espectral de A
$\ A\ _F$	norma de Frobenius de A
$\ A\ _p$	p -norma Schatten de A
$\ A\ _{(k)}$	k -norma de Fan de A
$W(A)$	rango numérico de A
$w(A)$	radio numérico de A
$A \circ B$	producto Hadamard de A y B
$\det(A)$	determinante de A
$\text{tr}(A)$	traza de A
π_n	el conjunto de permutaciones de $1, 2, \dots, n$
$s(A)$	el conjunto de valores singulares de A
$ A $	$(A^*A)^{1/2}$ o (a_{ij}) si $A = (a_{ij})$
$A^T = A(A^* = A)$	matriz simétrica (Hermitiana)
$A^*A = AA^* = I$	matriz unitaria u ortogonal en el caso real
$A^*A = AA^*$	matriz normal
E	perturbación de A
$\tilde{A} = A + E$	matriz perturbada de A
$\phi_A(\lambda)$	polinomio característico de A

Capítulo 1

Preliminares

En este capítulo introducimos algunos conceptos que usaremos a lo largo de la tesis. Entre ellos, la norma de matrices, número de condición y algunas matrices especiales. A su vez, enunciamos y demostramos algunos resultados que vamos a necesitar para los capítulos siguientes.

1.1. Autovalores y Autovectores

DEFINICIÓN 1.1. Sea A una matriz cuadrada de orden n . Un escalar $\lambda \in \mathbb{C}$ es un **autovalor** de A si existe un vector $v \in \mathbb{C}^n$, $v \neq 0$ tal que $Av = \lambda v$, en este caso v se llama **autovector** de A asociado al autovalor λ .

Se llama **espectro** de A y se denota por $\mathcal{L}(A)$, al conjunto de autovalores distintos de la matriz A .

Geométricamente, $Av = \lambda v$ significa que bajo la transformación por A , los autovectores experimentan sólo cambios en magnitud u orientación. El autovalor λ es simplemente la cantidad de estiramiento o contracción a la que se somete el autovector v cuando se transforma por A .

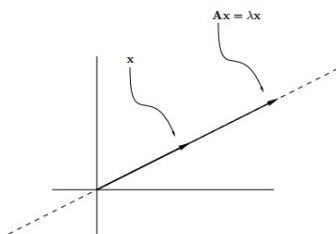


Figura 1.1: Representación gráfica en \mathbb{R}^2

1.2. Norma de matrices.

En las propiedades que definen una norma se considera las operaciones propias de espacio vectorial sobre el cual se definen: el producto por escalares y la suma de vectores (desigualdad triangular). No obstante dado que $\mathbb{F}^{m \times n}$ es un espacio vectorial formado por matrices de orden $m \times n$ con elementos en el cuerpo de los números reales o complejos existe el producto de matrices, lo que obliga a considerar esto en la definición de norma de matrices.

DEFINICIÓN 1.2. Sean μ , ν y ρ normas definidas en $\mathbb{F}^{m \times n}$, $\mathbb{F}^{n \times p}$ y $\mathbb{F}^{m \times p}$, respectivamente.

1. μ , ν y ρ son normas **consistentes** si para todas las matrices $A \in \mathbb{F}^{m \times n}$ y $B \in \mathbb{F}^{n \times p}$ se verifica la desigualdad

$$\rho(AB) \leq \mu(A)\nu(B).$$

En particular una norma ν definida en $\mathbb{F}^{n \times n}$ es consistente si $\nu(AB) \leq \nu(A)\nu(B)$, para todas $A, B \in \mathbb{F}^{n \times n}$.

2. Una norma definida en $\mathbb{F}^{n \times n}$ es una **norma de matriz** si es consistente.

DEFINICIÓN 1.3. El radio espectral de una matriz A es el número

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

TEOREMA 1.1. Sea $\|\cdot\|$ una norma consistente de matriz. Entonces para cualquier matriz A ,

$$\rho(A) \leq \|A\|.$$

Demostración. Ver [5], pág. 67. ■

EJEMPLO 1.1. 1. La norma euclídeana en $\mathbb{F}^{m \times n}$ se define como

$$\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2}.$$

A esta norma se le llama también **norma de Frobenius** y se representa como: $\|A\|_F$ o $\|A\|_2$. Comprobemos a continuación que es una norma de matriz.

Demostración. Suponga que $A \in \mathbb{F}^{m \times n}$ y $B \in \mathbb{F}^{n \times p}$. Entonces $AB \in \mathbb{F}^{m \times p}$ y, llamando a'_j a la j -ésima columna de A^* , se tiene:

$$\|AB\|_F^2 = \sum_{i=1}^m \sum_{j=1}^p |a^i b_j|^2 = \sum_{i=1}^m \sum_{j=1}^p |a'_i{}^* b_j|^2$$

Usamos ahora la desigualdad de Cauchy-Schwartz: $|x^* y| \leq \|x\|_2 \|y\|_2$. Así

$$\begin{aligned} \|AB\|_F^2 &\leq \sum_{i=1}^m \sum_{j=1}^p \|a'_i\|_2^2 \|b_j\|_2^2 \\ &\leq \sum_{k=1}^n \sum_{i=1}^m \|a'_{ki}\|^2 \sum_{k=1}^n \sum_{j=1}^p \|b_{kj}\|^2 \\ &= \sum_{i=1}^n \sum_{k=1}^n \|a_{ik}\|^2 \sum_{k=1}^n \sum_{j=1}^p \|b_{kj}\|^2 \\ &= \|A\|_F^2 \|B\|_F^2. \end{aligned}$$

■

2. En $\mathbb{F}^{n \times n}$, otro ejemplo de norma de matriz, es aquella definida por

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

3. Para cada $p \in \mathbb{N}$ la norma l_p en $\mathbb{F}^{m \times n}$ es definida como:

$$\|A\|_p = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^p \right)^{\frac{1}{p}}.$$

Usando la desigualdad de Hölder se puede demostrar que en $\mathbb{F}^{n \times n}$ esta norma es una norma de matriz si y sólo si $1 \leq p \leq 2$ con $p \in \mathbb{N}$.

1.3. Número de condición de matrices. Propiedades.

DEFINICIÓN 1.4. Se llama número de condición de una matriz no singular A relativo a la norma $\|\cdot\|$ al número

$$k(A) = \|A\| \cdot \|A^{-1}\| = \frac{s_n(A)}{s_1(A)},$$

donde s_n y s_1 es el mayor y menor valor singular de A respectivamente.

PROPOSICIÓN 1.1.

$$k(A) \geq 1 \quad (1.1)$$

para toda matriz A no singular.

Demostración.

$$1 = \|I\| = \|A \cdot A^{-1}\| \leq \|A\| \|A^{-1}\| = k(A).$$

■

DEFINICIÓN 1.5. Una matriz A está **bien condicionada** si $k(A)$ se aproxima a uno y está **mal condicionada** si $k(A)$ es significativamente mayor que uno.

EJEMPLO 1.2. En la siguiente matriz

$$A = \begin{pmatrix} a & -b \\ a + \varepsilon & -b \end{pmatrix}$$

se analizará el comportamiento de su número de condición utilizando la norma de Frobenius.

El determinante de A es $|A| = -ab + b(a + \varepsilon) = b\varepsilon$.

Si $b \neq 0$ y $\varepsilon \neq 0$ entonces $|A| \neq 0$ y, por tanto, A es invertible, siendo su inversa:

$$A^{-1} = \frac{1}{b\varepsilon} \begin{pmatrix} -b & b \\ -a - \varepsilon & a \end{pmatrix}.$$

El número de condición de Frobenius es: $k(A) = \|A\|_F \|A^{-1}\|_F$. Siendo

$$\|A\|_F^2 = a^2 + b^2 + (a + \varepsilon)^2 + b^2 = 2a^2 + 2b^2 + 2a\varepsilon + \varepsilon^2$$

$$\|A^{-1}\|_F^2 = \frac{b^2 + b^2 + (-a - \varepsilon)^2 + a^2}{b^2\varepsilon^2} = \frac{2a^2 + 2b^2 + 2a\varepsilon + \varepsilon^2}{b^2\varepsilon^2}.$$

Entonces

$$k(A)^2 = \frac{(2a^2 + 2b^2 + 2a\varepsilon + \varepsilon^2)^2}{b^2\varepsilon^2}$$

luego

$$k(A) = \frac{|2a^2 + 2b^2 + 2a\varepsilon + \varepsilon^2|}{|b\varepsilon|}.$$

Obsérvese que cuando ε tiende a cero, el número de condición $k(A)$ lo hace a infinito, por lo que la matriz A está mal condicionada.

Por ejemplo, para $a = 10$ y $b = 1$ se tiene que

$$k(A) = \frac{202 + 20\varepsilon + \varepsilon^2}{|\varepsilon|} = \frac{202}{|\varepsilon|} \pm 20 + |\varepsilon|.$$

Si $\varepsilon = 10^{-8}$ entonces $k(A) \simeq 2 \cdot 10^{10}$.

1.4. Algunas matrices especiales. Propiedades.

La importancia de las matrices en álgebra es conocida y existen numerosos teoremas que las caracterizan o que las emplean como herramienta. Pero además, si trabajamos con matrices especiales, esto es, con matrices que cumplen determinadas características, obtenemos otros resultados interesantes o importantes por sus aplicaciones.

Antes de comenzar con estas matrices, vamos a recordar algunas definiciones que utilizaremos en el siguiente capítulo.

DEFINICIÓN 1.6. Sean $A = (a_{ij})$ y $B = (b_{ij})$ matrices en $M_{m,n}$. El producto de Hadamard de A y B es definido por el producto de cada entrada y denotado por $A \circ B$:

$$A \circ B = (a_{ij}b_{ij}) \in M_{m,n}.$$

EJEMPLO 1.3. Para las siguientes matrices

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix} \text{ y } B = \begin{pmatrix} 2 & -1 & 3 & 1 \\ 3 & 4 & -0,5 & 0,1 \end{pmatrix}$$

el producto de Hadamard está dado por

$$A \circ B = \begin{pmatrix} 2 & -2 & 9 & 4 \\ 15 & 24 & -3,5 & 0,8 \end{pmatrix}.$$

1.4.1. Matriz Diagonalizable

DEFINICIÓN 1.7. Un matriz $A \in M_n$ es **diagonalizable**, si existe una matriz invertible P , llamada **matriz de paso**, tal que $P^{-1}AP = D$.

P es una matriz invertible cuyos vectores columna son autovectores de A , y $D = \text{Diag}(\lambda_1, \dots, \lambda_n)$ la matriz diagonal cuyos elementos diagonales son los autovalores de la matriz A .

EJEMPLO 1.4. En la siguiente matriz, se analizará para cuál o cuales valores de $\alpha \in \mathbb{R}$ esta es diagonalizable.

$$A = \begin{pmatrix} \alpha + 3 & \alpha & 0 & 0 \\ -\alpha & -\alpha + 3 & 0 & 0 \\ \alpha & \alpha - 1 & \alpha + 1 & \alpha - 1 \\ -\alpha & 1 - \alpha & 1 - \alpha & 3 - \alpha \end{pmatrix}.$$

Primero calculamos los autovalores de la matriz resolviendo la ecuación característica $|\lambda I - A| = 0$.

$$|\lambda I - A| = \begin{vmatrix} \lambda - \alpha - 3 & -\alpha & 0 & 0 \\ \alpha & \lambda + \alpha - 3 & 0 & 0 \\ -\alpha & 1 - \alpha & \lambda - \alpha - 1 & 1 - \alpha \\ \alpha & \alpha - 1 & \alpha - 1 & \lambda - 3 + \alpha \end{vmatrix}.$$

Operando resulta:

$$|\lambda I - A| = (\lambda - 2)^2(\lambda - 3)^2.$$

Por tanto las soluciones de $|\lambda I - A| = 0$ son:

$$\begin{aligned} \lambda_1 &= 2 & \text{de multiplicidad} & m_1 = 2 \\ \lambda_2 &= 3 & \text{de multiplicidad} & m_2 = 2 \end{aligned}$$

Ahora calculamos el subespacio de autovectores ($H(\lambda)$) correspondiente a cada autovalor.

Para $\lambda_1 = 2$ el subespacio $H(2)$ tiene por ecuaciones: $(A - 2I)v = 0$

$$\begin{pmatrix} -\alpha - 1 & -\alpha & 0 & 0 \\ \alpha & \alpha - 1 & 0 & 0 \\ -\alpha & 1 - \alpha & -\alpha + 1 & 1 - \alpha \\ \alpha & \alpha - 1 & \alpha - 1 & -1 + \alpha \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Obtenemos el siguiente sistema de ecuaciones:

$$\begin{aligned} (-\alpha - 1)x_1 - \alpha x_2 &= 0 \\ \alpha x_1 + (\alpha - 1)x_2 &= 0 \\ -\alpha x_1 + (1 - \alpha)x_2 + (1 - \alpha)x_3 + (1 - \alpha)x_4 &= 0 \\ \alpha x_1 - (1 - \alpha)x_2 - (1 - \alpha)x_3 - (1 - \alpha)x_4 &= 0 \end{aligned}$$

- Si $\alpha = -1$, de la primera y segunda ecuación tenemos: $x_1 = 0, x_2 = 0$. De la tercera y cuarta ecuación tenemos:
 - Caso 1: $\alpha \neq 1$, tenemos que: $x_3 = -x_4$. Por lo que la base de auto-vectores $H(2)$ es $B_{H(2)} = \{(0, 0, -1, 1)\}$. Entonces la dimensión de $H(2)$ es 1, por lo tanto la matriz no es diagonalizable.
 - Caso 2: $\alpha = 1$, tenemos que $x_1 = x_2 = 0$. Por lo que la base de auto-vectores $H(2)$ es $B_{H(2)} = \{(0, 0, 1, 0), (0, 0, 0, 1)\}$. Entonces la dimensión de $H(2)$ es 2 coincidiendo con su multiplicidad, por lo tanto la matriz es diagonalizable.
- Si $\alpha \neq 0$ obtenemos los mismos resultados que en el caso $\lambda = -1$.
- Si $\alpha = -1$ ó $\alpha = 0$, tenemos que: $x_1 = 0, x_2 = 0, x_3 = -x_4$. Por lo que la base de autovectores $H(2)$ es $B_{H(2)} = \{(0, 0, -1, 1)\}$. Entonces la dimensión de $H(2)$ es 1, por lo tanto la matriz no es diagonalizable.

Resumiendo, para el autovalor 2, la matriz únicamente es diagonalizable si $\alpha = 1$. Análogamente se realiza el mismo procedimiento para $\lambda_2 = 3$, teniendo como resultado que la matriz es únicamente diagonalizable si $\alpha = 0$.

Así, las matrices D (autovalores) y P (autovectores por columnas en el orden correspondiente a los autovalores) serán:

$$D = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

y

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & -1 \end{pmatrix}.$$

1.4.2. Matriz Unitaria

Las matrices unitarias son de gran importancia ya que, en muchos casos, son la herramienta principal para obtener distintos resultados en el análisis matricial. A continuación veremos distintas formas de caracterizarlas y cómo intervienen en una descomposición muy especial para matrices cuadradas, este resultado se conoce como el teorema de Schur.

DEFINICIÓN 1.8. Una matriz $U \in M_n$ es unitaria si

$$U^*U = UU^* = I_n,$$

donde U^* es la transpuesta conjugada de U e I_n es la matriz identidad de orden n .

EJEMPLO 1.5. Sea la matriz U de la forma

$$U = \frac{1}{2} \begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix}.$$

Hallamos la transpuesta conjugada de U

$$U^* = \frac{1}{2} \begin{pmatrix} 1-i & 1+i \\ 1+i & 1-i \end{pmatrix}.$$

Entonces el producto es:

$$UU^* = \frac{1}{4} \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} = I_3 \quad \text{y} \quad U^*U = \frac{1}{4} \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} = I_3.$$

Por lo tanto U es unitaria.

OBSERVACIÓN 1.1. El producto de matrices unitarias es unitaria. En general, si $A, B \in M_n$ son matrices unitarias, se tiene:

$$(AB)(AB)^* = (AB)(B^*A^*) = A(BB^*)A^* = AI_nA^* = AA^* = I_n.$$

Por lo tanto AB es unitaria. De forma análoga para BA .

DEFINICIÓN 1.9. a) Dos matrices A y B son unitariamente equivalentes si existe una matriz unitaria $U \in M_n$ tal que $A = U^*BU$. Cuando A es unitariamente equivalente a B denotaremos $A \simeq B$.

b) Cuando $A \in M_n$ es unitariamente equivalente a una matriz diagonal, entonces A es unitariamente diagonalizable.

TEOREMA 1.2. (Teorema de Schur) Si $A \in M_n$ entonces A es unitariamente semejante a una matriz triangular superior. En este caso, los valores propios de A son los elementos diagonales de dicha matriz triangular.

Demostración. La demostración se realizará por inducción sobre n . Si $n = 1$ no hay nada que demostrar porque $A = [a]$ ya es triangular. Suponemos entonces que es válido para matrices de orden $n - 1$. Sea $A \in M_n$ y $\lambda_1 \in \mathcal{L}(A)$. Sea u un vector

propio unitario de A asociado a λ_1 y $U = [u \ U_1]$ una matriz unitaria tal que las columnas de U_1 generan el espacio $\langle u \rangle^\perp$. Entonces

$$U^*AU = \begin{bmatrix} u^* \\ U_1^* \end{bmatrix} A[u \ U_1] = \begin{bmatrix} u^*Au & u^*AU_1 \\ U_1^*Au & U_1^*AU_1 \end{bmatrix}$$

Como $Au = \lambda_1 u$ y u es un vector de norma euclidiana 1, $u^*Au = \lambda_1$. Y como $\langle U_1 \rangle = \langle u \rangle^\perp$, $U_1^*Au = \lambda_1 U_1^*u = 0$. Por lo tanto, si ponemos $B = U_1^*AU$ y $c^* = u^*AU_1$, se tiene que

$$U^*AU = \begin{bmatrix} \lambda_1 & c^* \\ 0 & B \end{bmatrix}.$$

Ahora, B es de orden $(n-1)$ por lo que aplicando la hipótesis inductiva existe una matriz unitaria $V \in M_{n-1}$ tal que $V^*BV = T_1$ es una matriz triangular superior. Entonces

$$\begin{bmatrix} 1 & 0 \\ 0 & V^* \end{bmatrix} U^*AU = \begin{bmatrix} 1 & 0 \\ 0 & V \end{bmatrix} = \begin{bmatrix} \lambda_1 & a^* \\ 0 & T_1 \end{bmatrix} = T$$

donde $a^* = c^*V$, es una matriz triangular superior. Teniendo en cuenta que $\begin{bmatrix} 1 & 0 \\ 0 & V \end{bmatrix}$ es una matriz unitaria y que, por consiguiente, A y T son semejantes, se concluye que $\det(\lambda I_n - A) = \det(\lambda I_n - T) = \prod_{i=1}^n (\lambda - t_{ii})$, es decir, los autovalores de A son los elementos diagonales de T . Esto termina la demostración del teorema. ■

EJEMPLO 1.6. Para la siguiente matriz

$$A = \begin{pmatrix} -149 & -50 & -154 \\ 537 & 180 & 546 \\ -27 & -9 & -25 \end{pmatrix}$$

hallaremos la matriz triangular superior a la que es unitariamente semejante, utilizando el software Scicolab

```
-->schur(A)
ans =

1.    - 7.1118877    - 815.87059
0.      2.          - 55.023631
0.      0.          3.
```

Vemos que los autovalores, que en este caso son 1, 2 y 3, están en la diagonal. El hecho de que los elementos fuera de la diagonal sean tan grandes indica que esta matriz tiene autovalores sensibles si se hacen pequeños cambios a la matriz original.

1.4.3. Matriz Normal

DEFINICIÓN 1.10. Una matriz $A \in M_n$ es normal si conmuta con su traspuesta conjugada, es decir;

$$A^*A = AA^*.$$

Ejemplo:

Sea la matriz A de la forma

$$A = \begin{pmatrix} 5 & -2i & 4 \\ 2i & 8 & -2i \\ 4 & 2i & 5 \end{pmatrix}.$$

Hallamos la traspuesta conjugada de A

$$A^* = \begin{pmatrix} 5 & -2i & 4 \\ 2i & 8 & -2i \\ 4 & 2i & 5 \end{pmatrix}$$

$$\text{Entonces, } AA^* = \begin{pmatrix} 45 & -18i & 36 \\ 2i + 16 & 72 & -18i \\ 36 & 18i & 45 \end{pmatrix} = A^*A$$

Por lo tanto A es normal.

OBSERVACIÓN 1.2. Por definición toda matriz unitaria es normal.

1.4.4. Matriz hermitiana

DEFINICIÓN 1.11. Una matriz Hermitiana (o Hermítica) $A \in M_n$ es una matriz cuadrada de elementos complejos que tiene la característica de ser igual a su propia traspuesta conjugada. Es decir, $A = A^*$.

Se hallará la forma general de una matriz hermitiana de orden 3×3 : Si

$$A = \begin{pmatrix} a_{11} + ib_{11} & a_{12} + ib_{12} & a_{13} + ib_{13} \\ a_{21} + ib_{21} & a_{22} + ib_{22} & a_{23} + ib_{23} \\ a_{31} + ib_{31} & a_{32} + ib_{32} & a_{33} + ib_{33} \end{pmatrix}$$

entonces

$$\bar{A} = \begin{pmatrix} a_{11} - ib_{11} & a_{12} - ib_{12} & a_{13} - ib_{13} \\ a_{21} - ib_{21} & a_{22} - ib_{22} & a_{23} - ib_{23} \\ a_{31} - ib_{31} & a_{32} - ib_{32} & a_{33} - ib_{33} \end{pmatrix}$$

$$A^* = (\bar{A})^T = \begin{pmatrix} a_{11} - ib_{11} & a_{21} - ib_{21} & a_{31} - ib_{31} \\ a_{12} - ib_{12} & a_{22} - ib_{22} & a_{32} - ib_{32} \\ a_{13} - ib_{13} & a_{23} - ib_{23} & a_{33} - ib_{33} \end{pmatrix}$$

A es hermitiana si $A = A^*$, luego

$$a_{11} + ib_{11} = a_{11} - ib_{11} \Rightarrow ib_{11} = -ib_{11} \Rightarrow 2ib_{11} = 0 \Rightarrow b_{11} = 0$$

$$a_{12} + ib_{12} = a_{21} - ib_{21} \Rightarrow a_{12} = a_{21} \text{ y } b_{12} = -b_{21} \text{ y así sucesivamente.}$$

Por lo tanto la forma general de una matriz hermitiana de orden 3×3 es

$$A = \begin{pmatrix} a_{11} & a_{12} + ib_{12} & a_{13} + ib_{13} \\ a_{12} - ib_{12} & a_{22} & a_{23} + ib_{23} \\ a_{13} - ib_{13} & a_{23} - ib_{23} & a_{33} \end{pmatrix}.$$

OBSERVACIÓN 1.3. Los elementos de la diagonal principal de una matriz hermitiana son números reales, ya que tienen que coincidir con sus conjugados.

EJEMPLO 1.7. Las siguientes matrices

$$A = \begin{pmatrix} 2 & 3 - 2i & i \\ 3 + 2i & 5 & 2 + 4i \\ -i & 2 - 4i & 6 \end{pmatrix} \text{ y } B = \begin{pmatrix} 3 & 4 + 5i & 3 - 2i \\ 4 - 5i & 0 & -6 + 5i \\ 3 + 2i & -6 - 5i & 0 \end{pmatrix}$$

son matrices hermitianas.

DEFINICIÓN 1.12. Una matriz hermitiana cuadrada A es semidefinida positiva si $z^*Az > 0$, para todo vector no nulo $z \in \mathbb{C}^n$.

PROPOSICIÓN 1.2. Si $A, B \in M_n$ son matrices hermitianas y $\lambda \in \mathbb{R}$, entonces

1. $A + B$ es hermitiana.
2. λA es hermitiana.

Demostración. 1. Sean $A, B \in M_n$ hermitianas, entonces

$$(A + B)^* = A^* + B^* = A + B \text{ luego } A + B \text{ es hermitiana.}$$

2. Sean $\lambda \in \mathbb{R}$ y $A \in M_n$ hermitiana. Se verifica $\bar{\lambda} = \lambda$, por lo tanto

$$(\lambda A)^* = \bar{\lambda} A^* = \lambda A \text{ luego } \lambda A \text{ es hermitiana.}$$

■

PROPOSICIÓN 1.3. *Todos los autovalores de una matriz hermitiana son reales.*

Demostración. Sea $A \in M_n$ hermitiana y sea $x \in \mathbb{C}^n$ un vector columna. Vamos a demostrar previamente que x^*Ax es un número real. En efecto,

$$\overline{x^*Ax} = \overline{x^t A x} = x^t \overline{A x}.$$

La matriz $x^t \overline{A x}$ es de orden 1×1 , lo cual implica que es simétrica. Es decir,

$$\overline{x^*Ax} = x^t \overline{A x} = (x^t \overline{A x})^t = (\overline{x^t})(\overline{A})^t (x^t)^t = x^* A^* x = x^* A x.$$

Como x^*Ax coincide con su conjugado, es un número real.

Si λ (a priori complejo) es un autovalor de A , existe $x = (x_i) \in \mathbb{C}^n$ un vector columna no nulo tal que $Ax = \lambda x$. Multiplicando a la izquierda por x^* tenemos $x^*Ax = \lambda x^*x$. Teniendo en cuenta que

$$\begin{aligned} x^*x &= \begin{pmatrix} \overline{x_1} & \dots & \overline{x_n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= \overline{x_1}x_1 + \dots + \overline{x_n}x_n \\ &= |x_1|^2 + \dots + |x_n|^2 > 0 \end{aligned}$$

el valor de λ es por tanto $\lambda = \frac{x^*Ax}{x^*x} \in \mathbb{R}$. ■

TEOREMA 1.3. *Sea A una matriz hermitiana de orden n . Entonces los autovectores correspondientes a autovalores distintos de A son ortogonales.*

Demostración. Sean v_1 y v_2 autovectores asociados a autovalores distintos, digamos, λ_1 y λ_2 . Es decir,

$$\begin{aligned} Av_1 &= \lambda_1 v_1 & Av_2 &= \lambda_2 v_2 \\ v_2^* Av_1 &= \lambda_1 v_2^* v_1 & v_1^* Av_2 &= \lambda_2 v_1^* v_2. \end{aligned}$$

Al tomar la transpuesta conjugada de la primera expresión, se tiene

$$\begin{aligned} (v_2^* Av_1)^* &= (\lambda_1 v_2^* v_1)^* \\ v_1^* Av_2 &= \lambda_1 v_1^* v_2. \end{aligned}$$

En la última expresión se usaron los hechos de que $A^* = A$ y λ_1 es real. Luego se tiene

$$\lambda_2 v_1^* v_2 = \lambda_1 v_1^* v_2.$$

Por lo tanto $(\lambda_1 - \lambda_2)v_1^* v_2 = 0$. Pero $\lambda_1 - \lambda_2 \neq 0$, así que $v_1^* v_2 = 0$. Esto es, v_1 y v_2 son ortogonales. ■

1.4.5. Matriz Antihermitiana

DEFINICIÓN 1.13. Una matriz antihermitiana $A \in M_n$ es una matriz cuadrada cuya traspuesta conjugada es el negativo de dicha matriz; es decir,

$$A^* = -A.$$

EJEMPLO 1.8. Sea

$$A = \begin{pmatrix} i & 2 & 1 \\ -2 & -2i & 0 \\ -1 & 0 & -i \end{pmatrix}.$$

Hallando la conjugada traspuesta de A se tiene

$$A^T = \begin{pmatrix} i & -2 & -1 \\ 2 & -2i & 0 \\ 1 & 0 & -i \end{pmatrix} \Rightarrow A^* = \begin{pmatrix} -i & -2 & -1 \\ 2 & 2i & 0 \\ 1 & 0 & i \end{pmatrix}.$$

Por lo tanto $A^* = -A$. Es decir, A es antihermitiana.

TEOREMA 1.4. Los autovalores de una matriz antihermitiana son todos imaginarios puros.

Demostración. Sea A una matriz antihermitiana suponga que $X \neq 0$ es el auto-vector correspondiente al autovalor λ de A entonces tenemos:

$$\begin{aligned} AX &= \lambda X \\ (iA)X &= (i\lambda)X. \end{aligned}$$

Esto muestra que $i\lambda$ es un autovalor de la matriz A . Ahora

$$(iA)^* = -iA^* = -i(-A) = iA.$$

Entonces iA es una matriz hermitiana, por lo tanto $i\lambda$ es un autovalor de una matriz hermitiana, entonces $i\lambda$ será real siempre. Por lo que se prueba que $\lambda = 0$ o λ es puramente imaginario. ■

TEOREMA 1.5. Las matrices antihermitianas son matrices normales. Por lo tanto, son unitariamente diagonalizables y sus autovectores para distintos autovalores son ortogonales.

Demostración.

Primero probaremos que una matriz antihermitiana es una matriz normal: Ya que $A^* = -A$ entonces: $A^*A = -A \cdot -A^* = AA^*$.

Vamos ahora a probar que las matrices normales son unitariamente diagonalizables:

- Supongamos A es normal. Existe una matriz unitaria S y una matriz triangularmente superior U tal que: $A = SUS^*$. Entonces:

$$\begin{aligned}
 U \cdot U^* &= S^* A S (S^* A S)^* \\
 &= S^* A S S^* A^* S \\
 &= S^* A A^* S \\
 &= S^* A^* A S \\
 &= S^* A^* S S^* A S \\
 &= U^* U.
 \end{aligned}$$

Pero si dejamos que $u_{i,j}$ denote la ij -ésima entrada de U entonces la entrada superior izquierda de $U^* U$ es $u_{11} \overline{u_{11}} = |u_{11}|^2$ pero la entrada superior izquierda de $U U^*$ es

$$u_{11} \overline{u_{11}} + u_{12} \overline{u_{12}} + \dots + u_{1n} \overline{u_{1n}} = |u_{11}|^2 + |u_{12}|^2 + \dots + |u_{1n}|^2.$$

Dado que esto es igual $|u_{11}|^2$ y todos los sumandos son número reales no negativos debemos tener $u_{12} = u_{13} = \dots = u_{1n} = 0$. De manera similar, veamos que en la segunda entrada diagonal tenemos $u_{2j} = 0$ para todo $j > 2$. Continuando de esta manera veamos que U tiene que ser diagonal. Vemos que esto se muestra que si A es normal, entonces es diagonalizable con una matriz unitaria.

- Ahora suponga que A es alguna matriz la cual tiene una matriz unitaria S tal que $S^* A S = D$ es diagonal. Note que $DD^* = D^* D$. Entonces

$$\begin{aligned}
 A A^* &= S D S^* (S D S^*)^* \\
 &= S D S^* S D^* S^* \\
 &= S D D^* S^* \\
 &= S D^* D S^* \\
 &= S D^* S^* S D S^* \\
 &= A^* A.
 \end{aligned}$$

En consecuencia, A es normal.

Sea A una matriz antihermitiana, los autovalores distintos λ y μ , y los autovectores $v \neq 0$, $w \neq 0$ tal que

$$Av = \lambda v, \quad Aw = \mu w.$$

$$v^*(Aw) = v^*(\mu w) = \mu(v^*w).$$

Ya que $A^* = -A$ tenemos

$$-v^* A^* w = -(Av)^* w = -(\lambda v)^* w = -\lambda(v^* w).$$

Entonces

$$\lambda(v^* w) = \mu(v^* w).$$

Ya que $\lambda \neq \mu$ entonces $v^* w = 0$. Por lo tanto si A es antihermitiana, entonces los autovectores correspondientes a distintos autovalores son ortogonales, quedando demostrado. ■

TEOREMA 1.6. *Si A es antihermitiana entonces iA es hermitiana.*

Demostración. Sea A una matriz antihermitiana, tenemos

$$(iA)^* = i^* A^* = -i(-A) = iA.$$

Por lo tanto iA es hermitiana. ■

TEOREMA 1.7. *Si A, B es antihermitiana, entonces $aA + bB$ es antihermitiana para todos los escalares reales de a, b .*

Demostración. Sean A, B matrices antihermitianas

$$\begin{aligned} (aA + bB)^* &= (aA)^* + (bB)^* \\ &= a^* A^* + b^* B^* \\ &= -aA + -bB \\ &= -(aA + bB). \end{aligned}$$

Por lo tanto $aA + bB$ es antihermitiana. ■

1.5. Forma canónica de Jordan.

En esta sección se da a las matrices triangulares superiores estructuras más simples. El principal teorema es la descomposición de Jordan, que establece que cada matriz cuadrada compleja es similar a la suma directa de bloques de Jordan, conocida como forma canónica de Jordan o simplemente forma de Jordan.

TEOREMA 1.8. *Para cualquier matriz n -cuadrada A , existen una matriz P n -cuadrada compleja e invertible tal que:*

$$A = P^{-1}(J_1 \oplus J_2 \oplus \dots \oplus J_k)P, \quad (1.2)$$

donde cada J_i , $i = 1, 2, \dots, k$, es un bloque de Jordan.

Demostración. Ver [6], pág. 337. ■

Un bloque de Jordan es una matriz cuadrada de la forma:

$$J_x = \begin{pmatrix} x & 1 & & \\ & x & \ddots & \\ & & \ddots & 1 \\ 0 & & & x \end{pmatrix}. \quad (1.3)$$

Utilizamos las λ -matrices (matrices cuyas entradas son polinomios complejos cuya variable es λ) como una herramienta y usamos operaciones elementales para llevar a las λ -matrices a su forma estándar. A continuación mostraremos que una matriz cuadrada A es similar a su forma de Jordan que se determina por la forma estándar de la λ -matriz $\lambda I - A$.

Realizamos operaciones (adición y multiplicación) en las λ -matrices de la misma manera que lo hacemos para matrices numéricas. Las λ -matrices elementales y las λ -matrices son similarmente definidas a las matrices numéricas. Cualquier matriz numérica cuadrada puede convertirse en una matriz diagonal con 1 y 0 en la diagonal principal por operaciones elementales. Del mismo modo, las λ -matrices se pueden llevar a la forma estándar

$$\begin{pmatrix} d_1(\lambda) & & & 0 \\ & \ddots & & \\ & & d_k(\lambda) & \\ & & & 0 \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix} \quad (1.4)$$

donde $d_i(\lambda) | d_{i+1}(\lambda)$, $i = 1, \dots, k-1$; y cada $d_i(\lambda)$ es 1 o mónico. Por lo tanto para cualquier λ -matriz $A(\lambda)$ existen λ -matrices elementales $P_s(\lambda), \dots, P_1(\lambda)$ y $Q_1(\lambda), \dots, Q_t(\lambda)$ tal que:

$$P_s(\lambda), \dots, P_1(\lambda)A(\lambda)Q_1(\lambda), \dots, Q_t(\lambda) = D(\lambda) \quad (1.5)$$

es la forma estándar (1.4).

Si el $\det A(\lambda)$ es una constante diferente de cero, entonces $(\det A(\lambda))^{-1} \text{adj}(A(\lambda))$ es también una λ -matriz y es la inversa de $A(\lambda)$. Además la matriz es invertible si y solo si su forma estándar (1.4) es la matriz identidad y si, y solo si, es un producto de λ -matrices.

Para la λ -matriz $\lambda I - A$, tenemos $k = n$ y $D(\lambda) = \text{diag}(d_1(\lambda), \dots, d_n(\lambda))$. Los $d_i(\lambda)$ son llamados los factores invariantes de A y los divisores de $d_i(\lambda)$ factorizados en la forma $(\lambda - x)^t$ para alguna constante x y un entero positivo t son los divisores elementales de A .

EJEMPLO 1.9. Para ilustrar esto, veremos el siguiente ejemplo:

$$\begin{pmatrix} -1 & 0 & 0 \\ 1 & 1 & 2 \\ 3 & 0 & 1 \end{pmatrix}. \quad (1.6)$$

Realizamos operaciones elementales en la λ -matriz

$$\lambda I - A = \begin{pmatrix} \lambda + 1 & 0 & 0 \\ -1 & \lambda - 1 & -2 \\ -3 & 0 & \lambda - 1 \end{pmatrix}, \quad (1.7)$$

obtenemos:

$$\lambda I - A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (\lambda + 1)(\lambda - 1)^2 \end{pmatrix}. \quad (1.8)$$

Por lo tanto, los factores invariantes de A son:

$$d_1(\lambda) = 1, d_2(\lambda) = 1, d_3(\lambda) = (\lambda + 1)(\lambda - 1)^2$$

y los divisores elementales de A son: $(\lambda + 1), (\lambda - 1)^2$.

Note en la matriz una suma directa de 2 bloques de Jordan:

$$J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = (-1) \oplus \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \quad (1.9)$$

Tiene los mismos factores invariantes y divisores elementales para A . Llegando a la conclusión que A es similar a esta suma directa.

1.6. Algunos resultados adicionales

Descomposición QR

TEOREMA 1.9. Sea A una matriz de orden $m \times n$ con $m \geq n$. Entonces existe una matriz unitaria Q tal que

$$A = QR,$$

donde R es una matriz triangular superior con elementos no negativos en la diagonal.

Demostración. Sean v_1, v_2, \dots, v_m los vectores columnas de la matriz A (cada v_i es un vector de \mathbb{R}^n). El conjunto $\mathfrak{B} = \{v_1, v_2, \dots, v_m\}$ es base del subespacio S generado por las columnas de A .

$$A = \begin{pmatrix} v_1 & v_2 & \cdots & v_m \end{pmatrix}.$$

Por el método de Gram-Schmidt existe una base $\mathfrak{B}' = \{y_1, y_2, \dots, y_m\}$ de S tal que $[y_1, \dots, y_k] = [v_1, \dots, v_k] \forall k = 1, \dots, m$. Sea R la matriz de cambio de base $((I))_{\mathfrak{B}}$ de la base \mathfrak{B} a la base \mathfrak{B}' , por lo tanto R es una matriz triangular superior. Por otro lado se define la matriz Q cuyas columnas son los vectores y_1, y_2, \dots, y_m , es decir:

$$Q = \begin{pmatrix} y_1 & y_2 & \cdots & y_m \end{pmatrix}.$$

Es claro que por ser $\mathfrak{B}' = \{y_1, y_2, \dots, y_m\}$ un conjunto ortonormal $Q^t Q = I$ $((Q^t Q)_{ij} = \langle y_i, y_j \rangle)$.

Solo resta verificar que $QR = A$:

$$\begin{aligned} QR &= \begin{pmatrix} y_1 & y_2 & \cdots & y_m \end{pmatrix} \begin{pmatrix} \|u_1\| & c_{21}\|u_1\| & \cdots & c_{m1}\|u_1\| \\ 0 & \|u_2\| & \cdots & c_{m2}\|u_2\| \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \|u_m\| \end{pmatrix} \\ &= \begin{pmatrix} \|u_1\|y_1 & c_{21}\|u_1\|y_1 + \|u_2\|y_2 & \cdots & c_{m1}\|u_1\|y_1 + \cdots + \|u_m\|y_m \end{pmatrix} \\ &= \begin{pmatrix} v_1 & v_2 & \cdots & v_m \end{pmatrix} = A \end{aligned}$$

■

TEOREMA 1.10 (Teorema de Rouché). *Sea Ω un abierto acotado del plano complejo \mathbb{C} . Sean φ y ψ son funciones holomorfas en Ω y continuas en $\bar{\Omega}$. Supongamos que se verifica*

$$|\varphi(z) - \psi(z)| < |\varphi(z)| + |\psi(z)|, \quad \forall z \in \partial\Omega.$$

Entonces φ y ψ tiene el mismo número de ceros (contados los órdenes de multiplicidad) en Ω .

Demostración. Observe que la desigualdad de la hipótesis obliga a que los ceros de φ y ψ no puedan estar en la frontera $\partial\Omega$. Comenzamos razonando que los conjuntos $\Re(\varphi)$ y $\Re(\psi)$ son finitos.

Si $\Re(\varphi)$ fuera infinito, al ser $\bar{\Omega}$ compacto, $\Re(\varphi)$ tendría acumulación en $\bar{\Omega}$. Pero como φ no se anula, según hipótesis, en $\partial\Omega$, concluimos que tal punto ha de estar en el interior de Ω . Por lo tanto, φ es idénticamente nula en alguna componente conexa de Ω ; y, por argumentos de continuidad, se anulará en la frontera de esta componente. Pero este conjunto estará en $\partial\Omega$, lo cual es contradictorio.

El razonamiento es análogo para $\Re(\psi)$.

Sea ahora

$$K := \{z \in \bar{\Omega} : |\varphi(z) - \psi(z)| = |\varphi(z)| + |\psi(z)|\}.$$

Como φ y ψ son continuas, se sigue que K es cerrado. Y como es subconjunto de un compacto, $K \subset \overline{\Omega}$, también lo va a ser el propio K .

Además,

$$K \subset \Omega \quad \text{y} \quad \Re(\varphi) \cup \Re(\psi) \subset K.$$

Existe un ciclo Γ con soporte $\Gamma^* \subset \Omega K$, y tal que

$$\text{Ind}_{\Gamma}(z) = 0, \forall z \in \mathbb{C} \setminus \Omega \text{ e } \text{Ind}_{\Gamma}(z) = 1, \forall z \in K.$$

Y aplicando el principio del argumento generalizado a φ y ψ :

$$\frac{1}{i2\pi} \int_{\Gamma} \frac{\varphi'}{\varphi} = \sum_{a \in \Re(\varphi)} \text{Ind}_{\Gamma}(a) m(a) = \sum_{a \in \Re(\varphi)} m(a) = N(\varphi)$$

y

$$\frac{1}{i2\pi} \int_{\Gamma} \frac{\psi'}{\psi} = \sum_{a \in \Re(\psi)} \text{Ind}_{\Gamma}(a) m(a) = \sum_{a \in \Re(\psi)} m(a) = N(\psi),$$

donde $N(\varphi)$ y $N(\psi)$ son, respectivamente, el número de ceros de la citada función, contados el orden de multiplicidad con el que se presentan.

Objetivo: probar que

$$\int_{\Gamma} \frac{\varphi'}{\varphi} = \int_{\Gamma} \frac{\psi'}{\psi}. \quad (1.10)$$

Sea

$$h(z) := \log \left(\frac{\varphi(z)}{\psi(z)} \right), \quad \forall z \in \Omega \setminus K.$$

Es una función bien definida:

$$\frac{\varphi(z)}{\psi(z)} \in]-\infty, 0[\Rightarrow \varphi(z) = -\rho\psi(z) : \rho \geq 0$$

Entonces $|\varphi(z) - \psi(z)| = (1 + \rho)|\psi(z)| = |\varphi(z)| + |\psi(z)| \Rightarrow z \in K$; en conclusión

$$z \in \Omega \setminus K \Rightarrow \frac{\varphi(z)}{\psi(z)} \in \mathbb{C} \setminus]-\infty, 0[.$$

Por tanto,

$$h \in \mathcal{H}(\Omega \setminus K) : h'(z) = \frac{\varphi'(z)}{\varphi(z)} - \frac{\psi'(z)}{\psi(z)}, \forall z \in \Omega \setminus K.$$

Pero, que h' admita primitiva en $\Omega \setminus K$ es equivalente a que

$$\int_{\Gamma} h' = 0,$$

lo cual demuestra (1.10) ■

Teorema de Frobenius- König

DEFINICIÓN 1.14. Para enteros positivos $m \leq n$, sea $S_{m,n}$ el conjunto de inyecciones de $\{1, 2, \dots, m\}$ a $\{1, 2, \dots, n\}$. Sea $A = (a_{ij}) \in M_{m,n}$ con $m \leq n$. Si $\sigma \in S_{m,n}$, entonces la secuencia $a_{1\sigma(1)}, a_{2\sigma(2)}, \dots, a_{m\sigma(m)}$ es llamada transversal de A .

TEOREMA 1.11 (Konig). Para cada $A \in M_{m,n}$, $\delta(A) = \tau(A)$.

Demostración. Usamos inducción en el número $m + n$ de líneas de A . El teorema cumple para el caso $m = 1$ o $n = 1$.

Suponer que $m \geq 2$, $n \geq 2$ y asumimos que el teorema cumple para matrices con el número de líneas $< m + n$. Por definición tenemos $\tau(A) \geq \delta(A)$.

Llamamos a una cobertura mínima propia de A si no consta de todas las m filas de A o de todas las n columnas de A . Distinguimos dos casos:

Caso 1: A no tiene una cobertura mínima propia. En este caso $\delta(A) = \min\{m, n\}$. Sea $A = (a_{ij})$ con $a_{rs} \neq 0$. Denote por $A' \in M_{m-1, n-1}$ la matriz obtenida de A suprimiendo la r -ésima fila y la s -ésima columna de A . Entonces $\delta(A') \leq \min\{m-1, n-1\} = \delta(A) - 1$. Pero si $\delta(A') \leq \delta(A) - 2 = \min\{m-2, n-2\}$, entonces una cobertura mínima de A' más las dos líneas suprimidas daría una cobertura mínima adecuada de A , contradiciendo la suposición. Por lo tanto $\delta(A') = \delta(A) - 1$. Por hipótesis inductiva, $\tau(A') = \delta(A') = \delta(A) - 1$. Considerando la entrada a_{rs} además de $\tau(A')$ entradas de A' , dos de ellos no están en una línea, deducimos que $\tau(A) \geq \tau(A') + 1 = \delta(A)$.

Caso 2: A tiene una cobertura mínima propia. Ahora A tiene una cobertura mínima consistente de p filas y q columnas con $p + q = \delta(A)$ y $p < m$, $q < n$. Si $p = 0$ o $q = 0$, entonces la hipótesis inductiva da el resultado. Supongamos $p, q > 1$. Permutando filas o columnas no cambia el término rango y el rango de línea. Permutamos las filas y las columnas de A para que las p filas y las q columnas ocupan las posiciones iniciales. Entonces A se convierte en una matriz de la forma

$$\begin{bmatrix} * & B \\ C & 0 \end{bmatrix},$$

donde $B \in M_{p, n-q}$, $C \in M_{m-p, q}$. Concluimos que $\delta(B) = p$ y $\delta(C) = q$, porque si uno de ellos es falso, entonces tendríamos $\delta(A) < p + q$. Aplicando la hipótesis inductiva para B y C , obtenemos $\tau(B) = p$ y $\tau(C) = q$. Por lo tanto $\tau(A) \geq \tau(B) + \tau(C) = p + q = \delta(A)$. ■

TEOREMA 1.12 (Teorema de Frobenius- König). Sea $A \in M_{m,n}$ con $m \leq n$. Entonces cada transversal de A contiene a lo más k entradas ceros si solo si A tiene una submatriz cero de orden $r \times s$ con $r + s = n + k$.

Demostración. Suponga que cada transversal de A contiene al menos k entradas ceros. Entonces por el Teorema (1.11), $\delta(A) = \tau(A) \leq m - k$. Así A puede ser cubierto por $m - k$ líneas. Cubriendo estas $m - k$ líneas obtendremos una submatriz cero $r \times s$ de A con $r + s = m + n - (m - k) = n + k$.

Inversamente, supongamos que A tiene una submatriz cero de orden $r \times s$ con $r + s = n + k$. Entonces A puede ser cubierto por

$$(m - r) + (n - s) = m + n - (n + k) = m - k$$

líneas. Por el Teorema (1.11), $\tau(A) = \delta(A) \leq m - k$. Así cada transversal de A contiene al menos $m - (m - k) = k$ entradas ceros. ■

Teorema de Chebyshev

TEOREMA 1.13. Si p_n^* es la mejor aproximación para x^{n+1} en $[-1, 1]$ fuera de P_n , entonces

$$x^{n+1} - p_n^* = 2^{-n} T_{n+1}(x)$$

y así

$$E_n(x^{n+1}; [-1, 1]) = 2^{-n}.$$

Demostración. Sea $f(x) = x^{n+1}$. Si p_n^* es la mejor aproximación para x^{n+1} fuera de P_n y si $E_n(x^{n+1}; [-1, 1]) = \rho$, entonces

$$e(x) = x^{n+1} - p_n^*(x)$$

satisface

$$|e(x_j)| = \|e\| = \rho, \quad j = 0, 1, \dots, n+1.$$

donde $-1 \leq x_0 < x_1 < \dots < x_{n+1} \leq 1$ y $e(x) \in P_{n+1}$.

Ahora, $\rho^2 - e^2(x) \in P_{2n+2}$ y $\rho^2 - e^2(x_j) = 0$, $j = 0, 1, \dots, n+1$. Además, si $-1 < x_j < 1$, entonces $e^2 \leq \rho$ en $[-1, 1]$, $e^2(x)$ tiene un máximo relativo en $x = x_j$ y como

$$\frac{d}{dx} e^2(x) = 0$$

para $x = x_j$. Por lo tanto $x = x_j$

$$\frac{d}{dx} (\rho^2 - e^2(x)) = 0,$$

y cada x_j en el interior de $[-1, 1]$ es un cero de multiplicidad 2, con $\rho^2 - e^2(x)$. Como $\rho^2 - e^2(x)$ no es el polinomio cero, tiene precisamente $2n + 2$ ceros (contando ceros múltiples tantas veces como su multiplicidad); por lo tanto, ni x_0 ni x_{n+1} son puntos interiores de $[-1, 1]$. Es decir, tenemos demostrado que $x_0 = -1$

y $x_{n+1} = 1$ son ceros simples mientras x_1, \dots, x_n son ceros de multiplicidad dos de $\rho^2 - e(x)$. Pero el polinomio

$$(1 - x^2)[e'(x)]^2 \in P_{2n+2},$$

tiene ceros en ± 1 y tiene ceros de multiplicidad dos en x_1, x_2, \dots, x_n . Por lo tanto, $(1 - x^2)[e'(x)]^2$ y $\rho^2 - e^2(x)$ tienen conjuntos idénticos de ceros y por lo tanto, son múltiplos constantes entre sí. La constante se obtiene examinando el coeficiente principal de cada uno, y obtenemos

$$(1 - x^2)[e'(x)]^2 = (n + 1)^2(\rho^2 - e^2(x)).$$

La ecuación diferencial anterior para la función de error, $e(x)$, es fácilmente resuelta. Dado que $e'(x) \in P_n$ y $e'(x_j) = 0$, $j = 1, 2, \dots, n$, sabemos que $e'(x)$ no cambia de signo en $[-1, x_1]$. Suponemos que $e'(x) \geq 0$ en $[-1, x_1]$, entonces la igualdad implica, para $x \in [-1, x_1]$,

$$\frac{e'(x)}{[\rho^2 - e^2(x)]^{1/2}} = \frac{(n + 1)}{(1 - x^2)^{1/2}}.$$

Integrando se tiene

$$\arccos\left(\frac{e(x)}{\rho}\right) = (n + 1)\theta + c,$$

donde $x = \cos \theta$, $x \in [-1, x_1]$, $0 \leq \theta_1 \leq \theta \leq \pi$. por lo tanto obtenemos

$$e(x) = \rho \cos[(n + 1)\theta + c].$$

Ahora $e(-1) = -\rho$ asumiendo que $e'(-1) \geq 0$; así, $\cos[(n + 1)\pi + c] = -1$ y $c = m\pi$, donde $m + n + 1$ es impar. Por lo tanto,

$$e(x) = \pm \rho \cos(n + 1)\theta,$$

$\cos(n + 1)\theta$ es un polinomio de grado $n + 1$ en $x = \cos \theta$; es decir, $\cos(n + 1)\theta \in P_{n+1}$, y su coeficiente líder es 2^n , esto implica que

$$e(x) = 2^{-n} \cos(n + 1)\theta.$$

El polinomio $\cos k\theta$, donde $x = \cos \theta$, $0 \leq \theta \leq \pi$, es llamado el polinomio de Chebyshev de grado k y se escribe

$$T_k(x) = \cos k\theta, \quad k = 0, 1, 2, \dots \quad (1.11)$$

$T_k(x)$ es definido por (1.11) para $-1 \leq x \leq 1$ solamente. Pero como es un polinomio, una vez que conocemos sus coeficientes, sabemos como es una función para todo x (números complejos). Por lo tanto, puesto que $e(x) = 2^{-n}T_{n+1}(x)$ para $-1 \leq x \leq x_1$, la igualdad sigue siendo verdadera para todo x , en particular para $-1 \leq x \leq 1$. ■

El problema resuelto por el teorema anterior es equivalente a encontrar un polinomio de grado k con coeficientes de avance 1 que se desvíe menos de 0 en valor absoluto sobre $[-1, 1]$. El teorema anterior nos dice que este polinomio

$$\tilde{T}_k(x) = 2^{-k+1} T_k(x)$$

es el normalizador (es decir, normalizado de tal manera que su coeficiente principal es 1) polinomial de Chebyshev de grado k y, por lo tanto, para otro polinomio de grado k con coeficientes 1 existe un punto $x \in [-1, 1]$ tal que

$$|p(x)| > \frac{1}{2^{k-1}}. \quad (1.12)$$

Valores singulares

DEFINICIÓN 1.15. Los valores singulares de $A \in M_n$ son definidos por la raíz cuadrada no negativa de los autovalores de A^*A . El valor absoluto de $A \in M_n$ es definido y denotado por $|A| = (A^*A)^{1/2}$. Así los valores singulares de A son los autovalores de $|A|$. Denotaremos los valores singulares de $A \in M_n$ por

$$s_1(A) \geq s_2(A) \geq \dots \geq s_n(A) \text{ y } s(A) = (s_1(A), \dots, s_n(A)).$$

Principio dominante de Fan

Denotamos por $\|x\|$ la norma Euclidiana de $x \in \mathbb{C}^n$.

LEMA 1.1. Sea $A \in M_n$. Entonces para $k = 1, 2, \dots, n$,

$$s_k(A) = \max_{\substack{\Omega \subseteq \mathbb{C}^n \\ \dim \Omega = k}} \min_{\substack{x \in \Omega \\ \|x\| = 1}} \|Ax\| = \min_{\substack{\Omega \subseteq \mathbb{C}^n \\ \dim \Omega = n - k + 1}} \max_{\substack{x \in \Omega \\ \|x\| = 1}} \|Ax\|.$$

De este lema vemos que la norma espectral $\|A\|_\infty$ es igual al mayor valor singular $s_1(A)$.

LEMA 1.2. Sea $A \in M_n$. Entonces para $k = 1, 2, \dots, n$,

$$s_k(A) = \min\{\|A - G\|_\infty \mid \text{rang } G \leq k - 1, G \in M_n\}.$$

Demostración. Por el Lema (1.1),

$$s_k(A) = \min_{\substack{\Omega \subseteq \mathbb{C}^n \\ \dim \Omega = n - k + 1}} \max_{\substack{x \in \Omega \\ \|x\| = 1}} \|Ax\|.$$

Si $\text{rang } G \leq k - 1$, entonces $\dim \ker(G) \geq n - k + 1$. Elegimos un subespacio $\Omega_0 \subseteq \ker(G)$ con $\dim \Omega_0 = n - k + 1$. Tenemos

$$s_k(A) \leq \max_{\substack{x \in \Omega_0 \\ \|x\| = 1}} \|Ax\| = \max_{\substack{x \in \Omega_0 \\ \|x\| = 1}} \|(A - G)x\| \leq \|A - G\|_\infty.$$

Por otro lado, sea $A = U \text{diag}(s_1, \dots, s_n) V$ la descomposición del valor singular de A . Sea

$$G_0 = U \text{diag}(s_1, \dots, s_{k-1}, 0, \dots, 0) V.$$

Entonces $\text{rang } G_0 \leq k - 1$ y $s_k(A) = \|A - G_0\|_\infty$. Esto completa la demostración. ■

TEOREMA 1.14. Sean $A, B \in M_n$ con $1 \leq i, j \leq n$, $i + j - 1 \leq n$. Entonces

$$\begin{aligned} s_{i+j-1}(A + B) &\leq s_i(A) + s_j(B), \\ s_{i+j-1}(AB) &\leq s_i(A)s_j(B). \end{aligned} \tag{1.13}$$

Donde s_j es el j -ésimo valor singular.

Demostración. Por el Lema (1.2), existen $G, H \in M_n$ que satisfacen

$$\begin{aligned} \text{rang } G &\leq i - 1, & \text{rang } H &\leq j - 1, \\ s_i(A) &= \|A - G\|_\infty, & s_j(B) &= \|B - H\|_\infty. \end{aligned}$$

Entonces $\text{rang } (G + H) \leq (i + j - 1) - 1$. Así

$$\begin{aligned} s_{i+j-1}(A + B) &\leq \|A + B - (G + H)\|_\infty \\ &\leq \|A - G\|_\infty + \|B - H\|_\infty \\ &= s_i(A) + s_j(B). \end{aligned}$$

Dado que

$$\begin{aligned} \text{rang } [AH + G(B - H)] &\leq \text{rang } [G(B - H)] \\ &\leq (i + j - 1) - 1, \\ s_{i+j-1}(AB) &\leq \|AB - [AH + G(B - H)]\|_\infty \\ &= \|(A - G)(B - H)\|_\infty \\ &\leq \|A - G\|_\infty \|B - H\|_\infty \\ &= s_i(A)s_j(B). \end{aligned}$$
■

La desigualdad (1.13) es una generalización de

$$\|A + B\|_\infty \leq \|A\|_\infty + \|B\|_\infty \text{ y } \|AB\|_\infty \leq \|A\|_\infty \|B\|_\infty.$$

DEFINICIÓN 1.16. La norma $\|A\|_{(k)} = \sum_{j=1}^k s_j(A)$ para $1 \leq k \leq n$ es llamada k -norma de Fan.

OBSERVACIÓN 1.4. Ver que $\|\cdot\|_{(1)} = \|\cdot\|_\infty$ y $\|\cdot\|_2 = \|\cdot\|_F$. Además la norma $\|\cdot\|_{(n)} = \|\cdot\|_1$ es llamada la norma de la traza.

DEFINICIÓN 1.17. Dada $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_n)$, $\gamma_1 \geq \dots \geq \gamma_n \geq 0$, $\gamma_1 > 0$, define

$$\|A\|_\gamma = \sum_{j=1}^n \gamma_j s_j(A), \quad A \in M_n.$$

Esto es una norma invariante unitaria. Podemos llamarla γ -norma.

TEOREMA 1.15 (Principio Dominante de Fan). Sean $A, B \in M_n$. Si

$$\|A\|_{(k)} \leq \|B\|_{(k)}, \quad k = 1, \dots, n$$

entonces $\|A\| \leq \|B\|$ para cualquier norma invariante unitaria $\|\cdot\|$.

Demostración. Usando la γ -norma y la sumatoria por partes, se tiene

$$\|A\|_\gamma = \sum_{j=1}^{n-1} (\gamma_j - \gamma_{j+1}) \|A\|_{(j)} + \gamma_n \|A\|_{(n)},$$

donde $\gamma = (\gamma_1, \dots, \gamma_n)$, $\gamma_1 \geq \dots \geq \gamma_n \geq 0$, $\gamma_1 > 0$. ■

El principio de dominio de Fan puede ser expresado de manera equivalente como: $\|A\| \leq \|B\|$ para todas las normas invariantes unitarias si y solo si $s(A) \prec_\omega s(B)$.

LEMA 1.3. Sea $T \in M_n$. Entonces

$$T_{(k)} = \min\{\|X\|_1 + k \|Y\|_\infty : T = X + Y; X, Y \in M_n\}$$

para $k = 1, 2, \dots, n$.

Demostración. Si $T = X + Y$, entonces

$$\|T\|_{(k)} \leq \|X\|_{(k)} + \|Y\|_{(k)} \leq \|X\|_1 + \|Y\|_\infty.$$

Por otra parte, sea $T = U \text{diag}(s_1, \dots, s_n) V$ la descomposición en valores singulares con U, V unitarios y $s_1 \geq \dots \geq s_n \geq 0$. Entonces

$$\begin{aligned} X &= U \text{diag}(s_1 - s_k, s_2 - s_k, \dots, s_k - s_k, 0, \dots, 0) V, \\ Y &= U \text{diag}(s_k, \dots, s_k, s_{k+1}, \dots, s_n) V \end{aligned}$$

satisfaciendo $T = X + Y$ y $\|T\|_{(k)} = \|X\|_1 + k \|Y\|_\infty$. ■

DEFINICIÓN 1.18. Sea $\|\cdot\|$ la norma euclidiana en \mathbb{C}^n . El rango numérico de una matriz $A \in M_n$ se define como el conjunto

$$W(A) = \{x^*Ax : \|x\| = 1, x \in \mathbb{C}^n\}.$$

DEFINICIÓN 1.19. El radio numérico de una matriz $A \in M_n$ es definido y denotado por

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

LEMA 1.4. Sea $A \in M_n$. Entonces

$$w(A) \leq \|A\|_\infty \leq 2w(A).$$

Demostración. La primera desigualdad es obvia. Probaremos la segunda. Denotar

$$H = (A + A^*)/2, \quad S = (A - A^*)/2.$$

Entonces H y S son normales y $A = H + S$. Así la norma espectral de una norma de matriz es igual al radio numérico, tenemos

$$\begin{aligned} \|A\|_\infty &= \|H + S\|_\infty \leq \|H\|_\infty + \|S\|_\infty \\ &= w(H) + w(S) \\ &\leq w(A) + w(A^*) \\ &= 2w(A). \end{aligned}$$

■

LEMA 1.5. Sea $A \in M_n$. Entonces $w(A) \leq \|A\|_1 \leq nw(A)$.

Demostración. La primera desigualdad es obvia. Probaremos la segunda. Sea $A = UP$ la descomposición polar con U unitaria y P semidefinida positiva. Entonces $U^*A = P$. Sea $U^* = VDV^*$ la descomposición espectral con V unitaria, $D = \text{diag}(d_1, \dots, d_n)$, $|d_i| = 1$, $i = 1, \dots, n$. Tenemos

$$DV^*AV = V^*PV.$$

Sea $V = (v_1, \dots, v_n)$, $v_i \in \mathbb{C}^n$. Entonces cada v_i es un vector unitario. Por lo

tanto tenemos

$$\begin{aligned}
 \|A\|_1 &= \text{traz} V^* P V = \text{traz} D V^* A V = \left| \sum_{i=1}^n d_i v_i^* A v_i \right| \\
 &\leq \sum_{i=1}^n |d_i v_i^* A v_i| \\
 &= \sum_{i=1}^n |v_i^* A v_i| \\
 &\leq n w(A).
 \end{aligned}$$

■

TEOREMA 1.16. Sea $\|\cdot\|$ una norma invariante unitaria en M_n . Denotar

$$G = \begin{cases} I_{\frac{n}{2}}, & \text{si } n \text{ es par;} \\ \text{Diag}(I_{\frac{n-1}{2}} \otimes Z, 1), & \text{si } n \text{ es impar.} \end{cases}$$

$$\text{cuando } Z = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}.$$

Entonces para cada $A \in M_n$,

$$\|\text{Diag}(1, 0, \dots, 0)\| w(A) \leq \|A\| \leq \|G\| w(A).$$

Demostración. Si $A = 0$, la desigualdad es trivial. Suponer que $A \neq 0$. Sea $A' = A/w(A)$. Entonces por el Lema (1.4), $1 \leq s_1(A') \leq 2$ y por el Lema (1.5), $\sum_{i=1}^n s_i(A') \leq n$. Si n es par, G tiene $n/2$ valores singulares iguales a 2 y $n/2$ valores singulares iguales a 0; si n es impar, G tiene $(n-1)/2$ valores singulares iguales a 1 y $(n-1)/2$ valores singulares iguales a 0. Así

$$s(\text{diag}(1, 0, \dots, 0)) \prec_\omega s(A') \prec_\omega s(G).$$

Por el principio dominante de Fan, la desigualdad cumple para cada norma invariante unitaria. ■

Notar que las desigualdades del teorema anterior cumplen lo siguiente: la primera desigualdad se convierte en igualdad cuando $A = \text{diag}(1, 0, \dots, 0)$ y la segunda cuando $A = G$.

Capítulo 2

Perturbación de matrices

En este capítulo vamos a definir el concepto de perturbación de matrices, el cual depende de que clase de matriz se está utilizando. Para ello, vamos a presentar algunos ejemplos en donde, mediante este concepto, podamos ver que tan sensible es una matriz a perturbaciones.

2.1. Algunos resultados generales sobre matrices

Con el fin de motivar el tema, empecemos por considerar las matrices

$$A_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{y} \quad A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Es claro que ambas matrices tienen autovalores todos iguales a cero. Ahora considere una matriz cualquiera E de orden 4×4 , y considere también las matrices perturbadas

$$\tilde{A}_0 := A_0 + E \quad \text{y} \quad \tilde{A}_1 := A_1 + E.$$

Si, en particular,

$$E = E(\varepsilon) := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \varepsilon & 0 & 0 & 0 \end{bmatrix}$$

entonces los autovalores de \tilde{A}_0 siguen siendo nulos (es decir, iguales a los de A_0), mientras que los autovalores $\tilde{\lambda}$ de \tilde{A}_1 son

$$\tilde{\lambda}_1 = -\sqrt[4]{\varepsilon}, \quad \tilde{\lambda}_2 = -\sqrt[4]{\varepsilon} \, i, \quad \tilde{\lambda}_3 = \sqrt[4]{\varepsilon} \, i \quad \text{y} \quad \tilde{\lambda}_4 = \sqrt[4]{\varepsilon}.$$

Este cambio radical en los autovalores de \tilde{A}_0 y \tilde{A}_1 , a pesar de usar en ambos casos la misma matriz E , nos lleva a desarrollar una teoría de perturbación que tenga en cuenta diferentes clases de matrices.

2.1.1. Teoría general de perturbaciones

En esta sección, A es una matriz de orden $n \times n$, E una perturbación (matriz de orden $n \times n$) de A y $\tilde{A} = A + E$.

DEFINICIÓN 2.1. La multiplicidad algebraica m , de un autovalor λ es la multiplicidad que tiene como raíz del polinomio característico.

TEOREMA 2.1. Sea λ un autovalor de A de multiplicidad algebraica m y sea $\|\cdot\|$ una norma de matriz cualquiera. Entonces para todo $\epsilon > 0$ suficientemente pequeño existe un $\delta > 0$ tal que si $\|E\| < \delta$, el disco $D(\lambda, \epsilon) = \{\xi \in \mathbb{C} : |\xi - \lambda| \leq \epsilon\}$ contiene exactamente m autovalores de \tilde{A} .

Demostración. Sea ϵ tan pequeño tal que $D(\lambda, \epsilon)$ contiene solo el autovalor $\lambda \in A$ y sea $\eta(\xi) = \phi_{\tilde{A}}(\xi) - \phi_A(\xi)$. Por la continuidad del polinomio característico y ya que $\tilde{A} \rightarrow A$ entonces la función $\eta(\xi)$ converge a 0 en el conjunto compacto ∂D . Por otro lado ya que $\phi_A(\xi)$ es diferente de cero en ∂D , entonces $\exists \delta > 0$ tal que $|\eta(\xi)| < |\phi_A(\xi)|$ en ∂D cuando $\|E\| < \delta$. Finalmente por el teorema de Rouché (Teorema 1.10) se demuestra que ϕ_A y $\phi_{\tilde{A}} = \phi_A + \eta$ tienen el mismo número de ceros en D . ■

El Teorema 2.1 afirma que la perturbación $\|E\|$ debe ser pequeña, pero no proporciona una cota en el tamaño de dicha perturbación. Sin embargo, el teorema de Elsner, que veremos luego, nos proporciona cotas explícitas. Pero antes de dar dicho resultado se introducirá la notación para describir la forma en que los autovalores de dos matrices están situados el uno respecto al otro.

DEFINICIÓN 2.2. Sean las matrices A con autovalores $\{\lambda_1, \dots, \lambda_n\}$ y B con autovalores $\{\sigma_1, \dots, \sigma_n\}$.

1) Se llama variación espectral de B con respecto a A al número

$$VE_A(B) = \max_i \min_j |\sigma_i - \lambda_j|.$$

2) La distancia de Hausdorff entre los autovalores de A y B es

$$DH(A, B) = \max\{VE_A(B), VE_B(A)\}.$$

3) La distancia óptima entre los autovalores de A y B es

$$DO(A, B) = \min_{\pi} \{ \max_i |\sigma_{\pi(i)} - \lambda_i| \}$$

donde π es tomado sobre todas las permutaciones de $\{1, 2, \dots, n\}$.

EJEMPLO 2.1. La variación espectral de las matrices vistas en la introducción están dadas por:

$$VE_{A_1}(\tilde{A}_1) = \max_i \min_j |\tilde{\lambda}_i - \lambda_j| = |\sqrt[4]{\epsilon} - 0| = \sqrt[4]{\epsilon}.$$

$$VE_{\tilde{A}_1}(A_1) = \max_i \min_j |\lambda_i - \tilde{\lambda}_j| = |0 + \sqrt[4]{\epsilon}| = \sqrt[4]{\epsilon}.$$

De esto podemos obtener la Distancia de Hausdorff:

$$DH(\tilde{A}_1, A_1) = \max\{VE_{\tilde{A}_1}(A_1), VE_{A_1}(\tilde{A}_1)\} = \sqrt[4]{\epsilon}.$$

La función $VE_A(B)$ no es una métrica, es decir puede ser cero incluso cuando los autovalores de A y B son diferentes.

EJEMPLO 2.2. Si $A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ son matrices de orden 2 con autovalores $\lambda_1 = 0$, $\lambda_2 = 1$ y $\tilde{\lambda}_1 = 0$, $\tilde{\lambda}_2 = 0$ respectivamente, la variación espectral está dada por $VE_A(B) = \max_i \min_j |\tilde{\lambda}_i - \lambda_j| = 0$. Por lo tanto este es un ejemplo claro donde la variación espectral no es una métrica.

Geoméricamente, la función VE tiene la siguiente interpretación: Si

$$D_i = \{\xi : |\xi - \lambda_i| \leq VE_A(B)\}, \quad i = 1, \dots, n,$$

entonces

$$\mathcal{L}(B) \subset \bigcup_{i=1}^n D_i.$$

En otras palabras, los autovalores de B se encuentran en la unión de los discos de radio $VE_A(B)$, centrados en los autovalores de A .

TEOREMA 2.2. (Desigualdad de Hadamard) Si $A = [a_1 \ a_2 \ \dots \ a_n]$ es una matriz de orden n entonces

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

Además la igualdad se cumple si y solo si A tiene una columna cero o A es unitaria.

Demostración. Considere la factorización QR de A , es decir, $A = QR$, con Q una matriz unitaria y R una matriz triangular superior. Entonces $Q^*A = R \equiv \begin{bmatrix} r_1 & r_2 & \cdots & r_n \end{bmatrix}$, y como la pre-multiplicación por una matriz ortogonal no cambia la magnitud del determinante o las normas de la columna de A , tenemos

$$|\det(A)| = |\det(Q)\det(R)| = |\det(R)| \leq \prod_j |\rho_{jj}| \leq \prod_j \|r_j\|_2 = \prod_j \|a_j\|_2. \quad (2.1)$$

donde ρ_{jj} son los elementos de la diagonal de R . Entonces

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

Asumimos que la igualdad se cumple; es decir, $\det(A) = \prod_{j=1}^n \|a_j\|_2$ se tiene que:

- Si $\det(A) = 0$, por la igualdad se tiene que $\prod_{j=1}^n \|a_j\|_2 = 0$. Entonces $\exists i \in \mathbb{N}$ tal que $\|a_i\|_2 = 0$. Esto quiere decir que A tiene una columna nula.
- Si $\det(A) \neq 0$, entonces $\prod_j |\rho_{jj}| = \prod_j \|a_j\|_2$. Por lo tanto $|\rho_{jj}| = \|a_j\|_2$ para todo $j = 1, 2, \dots, n$, entonces A es unitaria.

De los ítems se deduce que si la igualdad se cumple entonces A es unitaria o tiene una columna nula.

Por un lado se asume que si A tiene una columna nula, entonces por propiedad

de determinante se tiene que $\det(A) = 0$, así mismo $\prod_{j=1}^n \|a_j\|_2 = 0$. Por ello

la igualdad se cumple. Por otro lado asumimos que si A es unitaria, entonces

$|\det(A)| = 1$ y $\prod_{j=1}^n \|a_j\|_2 = 1$. De esta manera se cumple la igualdad.

Por lo tanto la igualdad se cumple si y solo si A es unitaria o tiene una columna nula. ■

TEOREMA 2.3 (Teorema de Elsner). *Para cualquier matriz A , de orden $n \times n$ y cualquier perturbación E , se cumple que*

$$DH(A, \tilde{A}) \leq (\|A\|_2 + \|\tilde{A}\|_2)^{1-\frac{1}{n}} \|E\|_2^{\frac{1}{n}}$$

con $\tilde{A} = A + E$.

Demostración. De la Definición 2.2,

$$DH(A, \tilde{A}) = \max\{VE_A(\tilde{A}), VE_{\tilde{A}}(A)\}.$$

Si $C(A, \tilde{A}) := \left(\|A\|_2 + \|\tilde{A}\|_2\right)^{1-\frac{1}{n}} \|E\|_2^{\frac{1}{n}}$ entonces es claro que $C(\tilde{A}, A) = C(A, \tilde{A})$, es decir, la parte derecha de la desigualdad enunciada en el teorema es simétrica en A y \tilde{A} .

Debemos demostrar que $VE_A(\tilde{A}) \leq C(A, \tilde{A})$ y $VE_{\tilde{A}}(A) \leq C(\tilde{A}, A)$, pero debido a la igualdad $C(\tilde{A}, A) = C(A, \tilde{A})$ será suficiente demostrar la primera de estas desigualdades.

Ahora asumimos que el máximo en la variación espectral es alcanzado por el autovalor $\tilde{\lambda}$ de \tilde{A} y sean $\tilde{x}_1, \dots, \tilde{x}_n$ vectores ortonormales tales que $\tilde{A}\tilde{x}_1 = \tilde{\lambda}\tilde{x}_1$. Entonces

$$\begin{aligned} VE_A(\tilde{A})^n &\leq \prod_i |\lambda_i - \tilde{\lambda}| \\ &= \det(A - \tilde{\lambda}I) \\ &\leq \prod_i \|(A - \tilde{\lambda}I)\tilde{x}_i\|_2 \\ &= \|(A - \tilde{A})\tilde{x}_1\|_2 \prod_{i>1} \|(A - \tilde{\lambda}I)\tilde{x}_i\|_2 \\ &\leq \|E\|_2 (\|A\|_2 + \|\tilde{A}\|_2)^{n-1}. \end{aligned}$$

Tomando las raíces enésimas en la desigualdad anterior se tiene

$$VE_A(\tilde{A}) \leq \|E\|_2^{\frac{1}{n}} (\|A\|_2 + \|\tilde{A}\|_2)^{1-\frac{1}{n}}.$$

Por simetría y tomando el máximo obtenemos

$$DH(A, \tilde{A}) \leq \|E\|_2^{\frac{1}{n}} (\|A\|_2 + \|\tilde{A}\|_2)^{1-\frac{1}{n}}.$$

■

EJEMPLO 2.3. Tomando en cuenta el ejemplo de la introducción tenemos,

$$DH(A_1, \tilde{A}_1) = \sqrt[n]{\epsilon} \leq \left(\sqrt{3} + \sqrt{3 + \epsilon^2}\right)^{\left(1-\frac{1}{4}\right)} \sqrt[n]{\epsilon} = (\|A_1\|_2 + \|\tilde{A}_1\|_2)^{1-\frac{1}{n}} \|E\|_2^{\frac{1}{n}},$$

dónde n es el orden de la matriz A_1

LEMA 2.1. Si Γ es una curva continua en el plano complejo con puntos extremos en a y b , y p un polinomio mónico complejo de grado n entonces

$$\max_{z \in \Gamma} |p(z)| \geq 2^{1-2n} |b - a|^n.$$

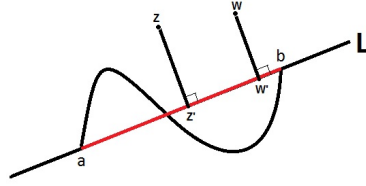


Figura 2.1: Representación gráfica.

Demostración. Sea L la línea recta que pasa por a y b y sea S el segmento de recta que une a a y b :

$$S = \{z : z = a + t(b - a), 0 \leq t \leq 1\}.$$

Para todo punto $z \in \mathbb{C}$, z' denota su proyección ortogonal en L , por lo que $|z - w| \geq |z' - w'|$ para todo $z, w \in \mathbb{C}$. Sean z_i , las raíces de p y $z'_i = a + t_i(b - a)$ con $t_i \in \mathbb{R}$ para todo $i = 1, \dots, n$. Ver figura [2.1]. Además todo $z \in L$ podemos escribirlo como $z = a + t(b - a)$ con $t \in \mathbb{R}$. Por lo tanto

$$\prod_{i=1}^n |z - z'_i| = \prod_{i=1}^n |(t - t_i)(b - a)| = |b - a|^n \left| \prod_{i=1}^n (t - t_i) \right|. \quad (2.2)$$

Del resultado clásico de Chebyshev (1.13) y de (2.2) se sabe que existe un punto $z_0 \in S$ tal que

$$\prod_{i=1}^n |z_0 - z'_i| \geq 2^{1-2n} |b - a|^n. \quad (2.3)$$

Ya que Γ es una curva continua que conecta a y b , entonces la proyección $y'_0 = z_0$ para algún $y_0 \in \Gamma$. Se tiene $|y_0 - z_i| \geq |y'_0 - z'_i| = |z_0 - z'_i|$ para todo $i = 1, \dots, n$. y de (2.3) se tiene

$$|p(y_0)| = \prod_{i=1}^n |y_0 - z_i| \geq 2^{1-2n} |b - a|^n.$$

Tomando el máximo de todos los $z \in \Gamma$ se obtiene

$$\max_{z \in \Gamma} |p(z)| \geq 2^{1-2n} |b - a|^n.$$

■

TEOREMA 2.4. (Bhatia- Elsner-Krause) Sean $A, \tilde{A} \in M_n$, y sea $\|\cdot\|$ la norma espectral. Entonces

$$DO(A, \tilde{A}) \leq 4(\|A\| + \|\tilde{A}\|)^{1-\frac{1}{n}} \|A - \tilde{A}\|^{\frac{1}{n}}. \quad (2.4)$$

Demostración. Sea E_t el espectro de $(1-t)A + t\tilde{A}$ en \mathbb{C} para t entre 0 y 1, y considere Ω la unión de los conjuntos E_t con t corriendo entre 0 y 1. Y Ω' alguna componente conexa de Ω , entonces por homotopía se sabe que, Ω' contiene tanto autovalores de A como de \tilde{A} . Por lo tanto para probar el teorema es suficiente mostrar que si $a, b \in \Omega'$, entonces $|a - b|$ es acotado por la expresión que aparece en el lado derecho de (2.4).

Sin pérdida de generalidad, asumimos que $\|A\| \leq \|\tilde{A}\|$. Sea $\mathcal{L}(A) = \{\lambda_1, \dots, \lambda_n\}$. Por el Lema (2.1), existe un $\lambda \in \Omega'$ tal que

$$|\det(\lambda I - A)| = \prod_{i=1}^n |\lambda - \lambda_i| \geq 2^{1-2n} |b - a|^n. \quad (2.5)$$

Sean X e Y matrices de orden n . Si $\mu \in \mathcal{L}(Y)$, se afirma que

$$|\det(\mu I - X)| \leq \|X - Y\|(\|X\| + \|Y\|)^{n-1}. \quad (2.6)$$

Denotamos por $e_j = (0, \dots, 0, 1, 0, \dots, 0)^T$ el j -ésimo vector base estándar de \mathbb{C}^n . Por el teorema de Schur existe una matriz unitaria U tal que U^*YU es triangular superior y $(U^*YU)(1, 1) = \mu$, entonces $(U^*YU)e_1 = \mu e_1$. Reemplazando X e Y por U^*XU y U^*YU respectivamente (lo cual no cambia la desigualdad 2.6), asumimos que $Ye_1 = \mu e_1$. Ahora usando la norma euclídeana para vectores se observa que

$$\|(\mu I - X)e_1\| = \|(Y - X)e_1\| \leq \|X - Y\|,$$

$$\|(\mu I - X)e_j\| \leq |\mu| + \|X\| \leq \|X\| + \|Y\|, \quad j = 2, \dots, n.$$

Aplicando la desigualdad de Hadamard, se tiene

$$|\det(\mu I - X)| \leq \prod_{i=1}^n \|(\mu I - X)e_i\| \leq \|X - Y\|(\|X\| + \|Y\|)^{n-1},$$

lo cual prueba (2.6).

Sustituyendo $X = A$, $Y = (1-t)A + t\tilde{A}$ y $\mu = \lambda \in \mathcal{L}((1-t)A + t\tilde{A})$ donde $t \in [0, 1]$ en (2.6) se tiene

$$|\det(\lambda I - A)| \leq \|A - \tilde{A}\|(\|A\| + \|\tilde{A}\|)^{n-1}. \quad (2.7)$$

Aquí hemos utilizado la condición $\|A\| \leq \|\tilde{A}\|$, luego combinando (2.5) y (2.7) se obtiene

$$\begin{aligned} |a - b| &\leq 2^{\frac{2n-1}{n}} \|A - \tilde{A}\|^{\frac{1}{n}} (\|A\| + \|\tilde{A}\|)^{1-\frac{1}{n}} \\ &\leq 4 \|A - \tilde{A}\|^{\frac{1}{n}} (\|A\| + \|\tilde{A}\|)^{1-\frac{1}{n}}. \end{aligned}$$

Dado que

$$|a - b| \geq \min_{\pi} \{ \max_i |\tilde{\lambda}_{\pi(i)} - \lambda_i| \} = DO(A, B)$$

el teorema queda demostrado. ■

EJEMPLO 2.4. La matriz

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

tiene autovalores $\lambda(A) = \{0, 1, 1, 2\}$. La matriz perturbada

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

tiene autovalores $\tilde{\lambda}(\tilde{A}) = \{1 - (1 + 2\varepsilon^2)^{1/2}, 1, 1, 1 + (1 + 2\varepsilon^2)^{1/2}\}$ donde $0 \leq \varepsilon < 1$. Efectuando obtenemos que la

$$DO(A, \tilde{A}) = 1 - \sqrt{1 + 2\varepsilon^2}$$

y

$$4(\|A\| + \|\tilde{A}\|)^{1-\frac{1}{n}} \|A - \tilde{A}\|^{\frac{1}{n}} = 4(\sqrt{6} + \sqrt{6 + 4\varepsilon^2})^{3/4} (2\varepsilon)^{1/4}.$$

Por lo tanto obtenemos la desigualdad de Bhatia- Elsner,

$$DO(A, \tilde{A}) \leq 4(\|A\| + \|\tilde{A}\|)^{1-\frac{1}{n}} \|A - \tilde{A}\|^{\frac{1}{n}}.$$

2.2. Algunos resultados de matrices normales.

Un ejemplo de matriz, de orden n , **no diagonalizable** las encontramos en las matrices de la forma

$$A = \begin{pmatrix} 0 & & & \\ \vdots & I_{n-1} & & \\ 0 & & \dots & \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

la cual tiene n autovalores nulos.

Si introducimos en la matriz A una perturbación y en lugar de esta tomamos la matriz

$$\tilde{A} = A + E = \begin{pmatrix} 0 & & & \\ \vdots & & I_{n-1} & \\ 0 & & & \\ \epsilon & 0 & \cdots & 0 \end{pmatrix}$$

con $\epsilon > 0$ muy pequeño, se obtiene como polinomio característico $P_{\tilde{A}}(\lambda) = \lambda^n + (-1)^n \epsilon$ que posee n raíces distintas, por lo que resulta muy sensible a cualquier perturbación que se realice en la matriz, por más pequeña que ésta sea.

Ahora veremos cómo afecta una pequeña perturbación a los autovalores de una matriz diagonalizable.

TEOREMA 2.5. *Sea A una matriz diagonalizable y $\tilde{A} = A + E$ una perturbación de dicha matriz, λ_i los autovalores de la matriz A y $\tilde{\lambda}$ cualquier autovalor de la matriz \tilde{A} con $\lambda_i \neq \tilde{\lambda}$. En las condiciones anteriores se tiene que*

$$\min_i |\tilde{\lambda} - \lambda_i| \leq \|P\| \|P^{-1}\| \|E\| = \|E\| \mathcal{K}(P)$$

donde P representa a la matriz de paso que diagonaliza a la matriz A .

Demostración. Por ser A diagonalizable existe P tal que $P^{-1}AP = D$. Sea x un autovector de \tilde{A} asociado a $\tilde{\lambda}$. Entonces $\tilde{A}x = \tilde{\lambda}x$ o, lo que es lo mismo:

$$(A + E)x = \tilde{\lambda}x$$

Entonces

$$\begin{aligned} (\tilde{\lambda}I - A)x &= Ex \\ (\tilde{\lambda}I - PDP^{-1})x &= Ex \\ P(\tilde{\lambda}I - D)P^{-1}x &= Ex \\ (\tilde{\lambda}I - D)(P^{-1}x) &= P^{-1}Ex \\ (\tilde{\lambda}I - D)(P^{-1}x) &= P^{-1}EP P^{-1}x. \end{aligned}$$

Como $\tilde{\lambda} \neq \lambda_i$ para todo $i = 1, 2, \dots, n$, la matriz $\tilde{\lambda}I - D$ es regular, por lo que

$$P^{-1}x = (\tilde{\lambda}I - D)^{-1}(P^{-1}EP)P^{-1}x$$

y para cualquier norma se tiene

$$\|P^{-1}x\| \leq \|(\tilde{\lambda}I - D)^{-1}\| \|P^{-1}EP\| \|P^{-1}x\|$$

es decir

$$1 \leq \|(\tilde{\lambda}I - D)^{-1}\| \|P^{-1}EP\|.$$

Como $(\tilde{\lambda}I - D)^{-1} = \text{Diag} \left(\frac{1}{\tilde{\lambda} - \lambda_1}, \dots, \frac{1}{\tilde{\lambda} - \lambda_i} \right)$ tenemos que

$$1 \leq \max_i \left(\left| \frac{1}{\tilde{\lambda} - \lambda_i} \right| \| P^{-1} \| \| E \| \| P \| \right)$$

por lo que

$$\min_i |\tilde{\lambda} - \lambda_i| \leq \| P^{-1} \| \| P \| \| E \| = \| E \| k(P).$$

■

Observe que la perturbación hecha en los autovalores depende del número de condición de la matriz de paso P y dado que esta no es única, trataremos de elegir entre todas las posibles matrices de paso, aquella cuyo número de condición sea mínimo.

COROLARIO 2.1. *Si A es unitariamente diagonalizable, el mínimo de la perturbación hecha en los autovalores es menor o igual que la perturbación producida en la matriz. Es decir,*

$$\min_i |\tilde{\lambda} - \lambda_i| \leq \| E \|.$$

Demostración. Según el Teorema 2.5 y dado que P es la matriz de autovectores de A , para cada autovalor $\tilde{\lambda}$ de $\tilde{A} = A + E$ se tiene:

$$\min_i |\tilde{\lambda} - \lambda_i| \leq k(P) \| E \|$$

Si A es unitariamente diagonalizable entonces P es ortogonal; es decir, $P^T P = I$. Para obtener el número de condición $k(P)$, por definición debemos hallar los valores singulares de P , para esto primero debemos obtener los autovalores de $P^T P$. Dado que $P^T P = I$, entonces cada autovalor de $P^T P$ es 1. Por definición de valores singulares se tiene que

$$s_1(P) = s_2(P) = \dots = s_n(P) = 1.$$

Por lo tanto $k(P) = s_n(P)/s_1(P) = 1$, donde $s_n(P)$ es el mayor valor singular y $s_1(P)$ el menor valor singular de la matriz P . Entonces $\min_i |\tilde{\lambda} - \lambda_i| \leq \| E \|$. ■

Por lo tanto, las mejores matrices para el cálculo efectivo de sus autovalores y autovectores son las diagonalizables unitariamente y éstas reciben el nombre de matrices normales.

PROPOSICIÓN 2.1. *Sea T una matriz triangular. Si $T^* T = T T^*$ entonces T es diagonal.*

Demostración. Probaremos que se trata de una matriz diagonal por inducción en el orden de la matriz. Y supondremos que T es triangular superior.

Para $n = 1$ es obvio. Para $n = 2$ es $T = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ entonces $T^* = \begin{pmatrix} \bar{a} & 0 \\ \bar{b} & \bar{c} \end{pmatrix}$.

$$T^*T = \begin{pmatrix} \bar{a} & 0 \\ \bar{b} & \bar{c} \end{pmatrix} \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = \begin{pmatrix} |a|^2 & \bar{a}b \\ a\bar{b} & |b|^2 + |c|^2 \end{pmatrix}$$

$$TT^* = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} \bar{a} & 0 \\ \bar{b} & \bar{c} \end{pmatrix} = \begin{pmatrix} |a|^2 + |b|^2 & b\bar{c} \\ \bar{b}c & |c|^2 \end{pmatrix}.$$

Dado que $T^*T = TT^*$, igualando ambas matrices se obtiene que $|b|^2 = 0$ es decir $b = 0$ y, por tanto, T es diagonal.

Supongamos ahora que el teorema es cierto para cualquier matriz triangular de orden n . Probaremos el resultado para matrices de orden $n + 1$.

$$\text{Sean } T = \begin{pmatrix} a_1 & a_2 & \dots & a_{n+1} \\ 0 & & & \\ \vdots & T_n & & \\ 0 & & & \end{pmatrix} \text{ y } T^* = \begin{pmatrix} \bar{a}_1 & 0 & \dots & 0 \\ \bar{a}_2 & & & \\ \vdots & T_n^* & & \\ a_{n+1}^- & & & \end{pmatrix}$$

$$TT^* = \begin{pmatrix} \sum_{i=1}^{n+1} |a_i|^2 & & \\ & T_n T_n^* & \end{pmatrix} \quad T^*T = \begin{pmatrix} |a_1|^2 & & \\ & T_n^* T_n & \end{pmatrix}.$$

De la igualdad de ambas tenemos que $T^*T = TT^*$ entonces $|a_2|^2 + \dots + |a_{n+1}|^2 = 0$ por lo tanto $a_2 = a_3 = \dots = a_{n+1} = 0$.

Como $T_n^* T_n = T_n T_n^*$, por hipótesis de inducción sabemos que T_n es diagonal y, por lo tanto, T es diagonal. ■

TEOREMA 2.6. Sea $A \in \mathbb{C}^{n \times n}$ una matriz con valores propios $\lambda_1, \lambda_2, \dots, \lambda_n$. A es normal si, y solo si, A es unitariamente diagonalizable.

Demostración. Supongamos que $A \in \mathbb{C}^{n \times n}$ es unitariamente diagonalizable; entonces su conjugada transpuesta es

$$A^* = (UDU^*)^* = (U^*)^* D^* U^* = UD^* U^*.$$

La demostración de que A es una matriz normal es un simple cálculo que sólo usa la conmutatividad de las matrices diagonales y que U^* es la inversa de U .

$$A^*A = UD^*U^*UDU^* = UD^*DU^* = UDD^*U^* = UDU^*UD^*U^* = AA^*.$$

Por lo tanto A es normal.

Recíprocamente por el Teorema de Schur para cada $A \in \mathbb{C}^{n \times n}$ existe una matriz unitaria $U \in \mathbb{C}^{n \times n}$ tal que $U^*AU = T$, una matriz triangular superior. Si A es normal entonces $A^*A = AA^*$, esto es equivalente a $T^*T = TT^*$ y por la Proposición (2.1) se tiene que T es diagonal.

Dado que T es diagonal y por el Teorema de Schur para matrices normales (Teorema 1.2) se tiene una diagonalización ortogonal. Con esto queda demostrado el teorema. ■

El Teorema 2.6 se conoce con el nombre de *Teorema espectral para las matrices normales*. Y éste, a su vez, se expresa diciendo que una matriz $A \in \mathbb{C}^{n \times n}$ es normal si, y sólo si, se puede escribir de la forma

$$A = \sum_{i=1}^n \lambda_i v_i v_i^*$$

con v_1, \dots, v_n vectores ortonormales.

Se trata de otra forma de escribir $U^*AU = \text{Diag}(\lambda_1, \dots, \lambda_n)$ porque si $V = U^*$ entonces V es unitaria y

$$\begin{aligned} A &= V^* \text{Diag}(\lambda_1, \dots, \lambda_n) V = \begin{bmatrix} v_1^* \\ \vdots \\ v_n^* \end{bmatrix} \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix} \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix} \\ &= \lambda_1 v_1 v_1^* + \dots + \lambda_n v_n v_n^*. \end{aligned}$$

TEOREMA 2.7. (Kittaneh) Sean $A, \tilde{A} \in \mathbb{C}^{n \times n}$ matrices normales, entonces

$$\| |A| - |\tilde{A}| \|_2 \leq \|A - \tilde{A}\|_2.$$

Demostración. Notación: $|A| = (A^*A)^{1/2}$ o $(|a_{ij}|)$ si $A = (a_{ij})$.

Sean $A, \tilde{A} \in \mathbb{C}^{n \times n}$ matrices normales. Por el teorema espectral se tiene

$$A = U^* \text{diag}(\lambda_1, \dots, \lambda_n) U \text{ y } \tilde{A} = V^* \text{diag}(\mu_1, \dots, \mu_n) V,$$

donde U y V son matrices unitarias.

Entonces $|A| = U^* \text{diag}(|\lambda_1|, \dots, |\lambda_n|) U$ y $|\tilde{A}| = V^* \text{diag}(|\mu_1|, \dots, |\mu_n|) V$.

Para simplificar hacemos $C = \text{diag}(|\lambda_1|, \dots, |\lambda_n|)$, $D = \text{diag}(|\mu_1|, \dots, |\mu_n|)$ y

$W = (w_{ij}) = UV^*$. Entonces desarrollando se tiene:

$$\begin{aligned}
 \| |A| - |\tilde{A}| \|_2 &= \| U^* C U - V^* D V \|_2 \\
 &= \| C U V^* - U V^* D \|_2 \\
 &= \| C W - W D \|_2 \\
 &= \left(\sum_{i,j=1}^n (|\lambda_i| - |\mu_j|)^2 \cdot |w_{ij}|^2 \right)^{1/2} \\
 &\leq \left(\sum_{i,j=1}^n |\lambda_i - \mu_j|^2 \cdot |w_{ij}|^2 \right)^{1/2} \\
 &= \left(\sum_{i,j=1}^n (|\lambda_i - \mu_j|) |w_{ij}|^2 \right)^{1/2} \\
 &= \| \text{diag}(\lambda_1, \dots, \lambda_n) W - W \text{diag}(\mu_1, \dots, \mu_n) \|_2 \\
 &= \| A - \tilde{A} \|_2.
 \end{aligned}$$

■

COROLARIO 2.2. Sean A y $B \in \mathbb{C}^{n \times n}$ matrices complejas. Entonces

$$\| |A| - |B| \|_2 \leq \sqrt{2} \| A - B \|_2.$$

Demostración. Sean $\hat{A} = \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}$ y $\hat{B} = \begin{pmatrix} 0 & B \\ B^* & 0 \end{pmatrix}$. Entonces \hat{A} y \hat{B} son normales. Aplicando el Teorema anterior para \hat{A} y \hat{B} , tenemos

$$\| |\hat{A}| - |\hat{B}| \|_2^2 = \| |A| - |B| \|_2^2 + \| |A^*| - |B^*| \|_2^2 \leq 2 \| A - B \|_2^2.$$

Por lo tanto $\| |A| - |B| \|_2 \leq \sqrt{2} \| A - B \|_2$.

■

DEFINICIÓN 2.3. Una matriz $A = (a_{ij})$ cuadrada $n \times n$ es doblemente estocástica si $a_{ij} \geq 0$ para $i, j = 1, \dots, n$, y

$$\sum_i a_{ij} = 1, \quad j = 1, \dots, n; \quad \sum_j a_{ij} = 1, \quad i = 1, \dots, n.$$

Al conjunto de matrices doblemente estocásticas $n \times n$ se le denota como Ω_n .

DEFINICIÓN 2.4. Sea K un conjunto convexo en \mathbb{R}^d . Un punto $x \in K$ es llamado **punto extremo** de K si $y, z \in K$, $0 < t < 1$ y $x = ty + (1 - t)z$ implica que $x = y = z$.

El conjunto de todos los puntos extremos de K será denotado por $ex(K)$. El conjunto Ω_n de matrices doblemente estocásticas de orden n es claramente un conjunto convexo compacto.

NOTACIÓN

Π_n es el conjunto de matrices de permutación de orden n .

TEOREMA 2.8. $ex(\Omega_n) = \Pi_n$. Es decir, cada matriz doblemente estocástica es una combinación convexa de matrices de permutación.

Demostración. Primero probaremos que $\Pi_n \subseteq ex(\Omega_n)$. Sea $P \in \Pi_n$, suponemos que $P = tA + (1-t)B$, $0 < t < 1$, $A, B \in \Omega_n$. Dado que A, B son no negativas, para $1 \leq i, j \leq n$, $P(i, j) = 0$, esto implica que $A(i, j) = B(i, j) = 0$, entonces cada fila y columna de A y B tiene como máximo una entrada positiva, pero $A, B \in \Omega_n$. Por lo tanto A, B son matrices de permutación y $A = B = P$. Esto demuestra que $P \in ex(\Omega_n)$ dado que P fue elegido arbitrariamente. Entonces:

$$\Pi_n \subseteq ex(\Omega_n).$$

Ahora demostraremos que cada matriz doblemente estocástica es una combinación convexa de matrices de permutación para eso, utilizaremos inducción sobre el número de entradas positivas. Sea $G \in \Omega_n$, entonces G tiene al menos n entradas positivas ya que sus entradas pueden variar entre 0 y 1 pero con la condición que sume columnas y filas siempre 1. Si G tiene exactamente n entradas positivas, ya que solo serían n veces 1, entonces G es una matriz de permutación.

Primero mostraremos que cada matriz doblemente estocástica tiene al menos una transversal positiva, es decir, una transversal con cada entrada positiva.

- Si $G \in \Omega_n$ no tiene ceros en las entradas, entonces la afirmación se mantiene.
- Si $G \in \Omega_n$ tiene ceros en las entradas, se puede suponer que G tiene una submatriz cero de orden $r \times s$. Permutando filas y columnas si es necesario, tenemos:

$$G = \begin{pmatrix} 0 & C \\ D & E \end{pmatrix}$$

donde $0 \in M_{r,s}$. Dado que la suma de las entradas en cada fila de C y la suma de las entradas en cada columna de D son 1, y la suma de todas las entradas de G es n y $r + s \leq n$, entonces por el Teorema (1.12), G tiene una transversal positiva.

Asumimos que cada matriz doblemente estocástica en Ω_n con al menos $k - 1$ entradas positivas es una combinación convexa de matrices de permutación. Sea $G \in \Omega_n$ con exactamente k entradas positivas, tomamos una transversal positiva de G y suponemos que a es la menor entrada en la transversal. Si $G \notin \Pi_n$, entonces $a < 1$. Sea Q la matriz de permutación con 1's en la transversal correspondiente a la transversal positiva elegida de G y el conjunto $T = \frac{(G - aQ)}{(1 - a)}$, entonces $T \in \Omega_n$, y T tiene al menos un cero más de entrada que G . Por lo tanto T tiene como máximo $k - 1$ entradas positivas y debido a la hipótesis de inducción, T es una combinación convexa de matrices de permutación. Por consiguiente, también lo es G , ya que $G = (1 - a)T + aQ$.

Sea $W \in \Omega_n \setminus \Pi_n$, entonces existen números t_i y matrices de permutación P_i tal que $W = \sum_{i=1}^m t_i P_i$, $0 < t_i < 1$, $i = 1, \dots, m$, $\sum_{i=1}^m t_i = 1$, $m \geq 2$ y el conjunto $F = \frac{(\sum_{j=2}^m t_j P_j)}{(1 - t_1)}$. Entonces tenemos $F \in \Omega_n$, $W = t_1 P_1 + (1 - t_1)F$, $P_1 \neq W$. Por lo tanto $W \notin \text{ex}(\Omega_n)$. Esto prueba que $\text{ex}(\Omega_n) \subseteq \Pi_n$.

$$\therefore \text{ex}(\Omega_n) = \Pi_n.$$

■

TEOREMA 2.9. (Hoffman-Wielandt) Sean A y $\tilde{A} \in \mathbb{C}^{n \times n}$ matrices normales con autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$ y $\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_n$ respectivamente. Entonces existe $r \in S_n$ donde S_n es el conjunto de todas las permutaciones, talque

$$\left(\sum_{i=1}^n |\lambda_i - \tilde{\lambda}_{r(i)}|^2 \right)^{1/2} \leq \|A - \tilde{A}\|_F.$$

Demostración. Considerar la descomposición espectral

$$A = U D_1 U^*, \quad \tilde{A} = V D_2 V^*,$$

donde U, V son unitarios y $D_1 = \text{Diag}(\lambda_1, \dots, \lambda_n)$, $D_2 = \text{Diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_n)$.

Por la invarianza unitaria de la norma tenemos:

$$\begin{aligned} \|A - \tilde{A}\|_2 &= \|U D_1 U^* - V D_2 V^*\|_2 \\ &= \|U D_1 U^* V V^* - U U^* V D_2 V^*\|_2 \\ &= \|U (D_1 U^* V - U^* V D_2) V^*\|_2 \\ &= \|D_1 W - W D_2\|_2 \end{aligned}$$

donde $W = U^* V = (w_{ij})$ es unitaria.

Denotamos $G = (|\lambda_i - \tilde{\lambda}_j|^2)$, $Z = (w_{ij}^2) \in \mathbb{C}^{n \times n}$ y sea $e \in \mathbb{R}^n$ un vector con todas sus componente iguales a 1. Entonces

$$\|A - B\|_F^2 = \sum_{i,j=1}^n |\lambda_i - \tilde{\lambda}_j|^2 |w_{ij}|^2 = e^T (G \circ Z) e.$$

Dado que Z es doblemente estocástica y por el teorema (2.8), Z es una combinación convexa de permutación de matrices, es decir,

$$Z = \sum_{i=1}^{n!} t_i P_i,$$

donde $t_i \geq 0$, $\sum_i t_i = 1$ y los P_i son matrices de permutación.

Suponer que $e^T (G \circ P_k) e = \min\{e^T (G \circ P_i) e | 1 \leq i \leq n!\}$ y P_k corresponde a $\tau \in S_n$. Entonces

$$\begin{aligned} \|A - B\|_F^2 &= \sum_{i=1}^{n!} t_i e^T (G \circ P_i) e \\ &\geq \sum_{i=1}^{n!} t_i e^T (G \circ P_k) e \\ &= e^T (G \circ P_k) e \\ &= \sum_{i=1}^n |\lambda_i - \tilde{\lambda}_{\tau(i)}|^2. \end{aligned}$$

■

EJEMPLO 2.5. La hipótesis que tanto A como \tilde{A} sean normales es necesaria. Veamos el siguiente ejemplo, sean

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}$$

y

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

Observar que A es normal y que \tilde{A} no lo es. Los autovalores de A son 0 y 4 mientras que los autovalores de \tilde{A} son todos ceros. Por lo tanto

$$\left(\sum_{i=1}^n |\lambda_i - \tilde{\lambda}_{\tau(i)}|^2 \right)^{1/2} = 16 > 12 = \|A - \tilde{A}\|_F^2.$$

Este hecho complica la aplicación práctica del teorema de Hoffman-Wielandt, ya que la suma de dos matrices normales no necesariamente es normal. Incluso la suma de una matriz normal y una matriz hermitiana puede no ser normal. Así, la clase de perturbaciones que el teorema puede manejar es estrictamente limitado.

2.3. Matriz normal con perturbación general.

Para poder analizar el caso cuando A es una matriz normal con perturbación general, antes consideremos dos lemas.

LEMA 2.2. Sea $A = (a_{ij}) \in M_n$ una matriz normal. Entonces

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^n (j-i)|a_{ij}|^2 = \sum_{j=1}^{n-1} \sum_{i=j+1}^n (i-j)|a_{ij}|^2. \quad (2.8)$$

Demostración. Sea la matriz particionada :

$$A = \begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix}, A_k \in M_k. \quad (2.9)$$

De $AA^* = A^*A$ tenemos:

$$A_k A_k^* + B_k B_k^* = A_k^* A_k + C_k^* C_k, \quad k = 1, 2, \dots, n-1.$$

Tomando la traza en ambos lados:

$$\|B_k\|_F^2 = \|C_k\|_F^2, \quad k = 1, 2, \dots, n-1.$$

Aplicando sumatoria a ésta igualdad tenemos:

$$\sum_{k=1}^{n-1} \|B_k\|_F^2 = \sum_{k=1}^{n-1} \|C_k\|_F^2.$$

Ya que por la posición de la matriz B_k y C_k tenemos $i < j$ y $j < i$ lo cual por norma de Frobenius es equivalente a:

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

■

Para $X \in M_n$, denotamos la descomposición triangular por $X = L(X) + D(X) + U(X)$, donde $L(X)$ es una matriz triangular estrictamente inferior, $D(X)$ es una matriz diagonal, y $U(X)$ es una matriz triangular estrictamente superior.

LEMA 2.3. *Sea $A \in M_n$ normal. Entonces*

$$\|U(A)\|_F \leq \sqrt{n-1}\|L(A)\|_F, \quad \|L(A)\|_F \leq \sqrt{n-1}\|U(A)\|_F. \quad (2.10)$$

Demostración. Sea $A = (a_{ij})$. Aplicamos el Lema 2.2 y tenemos:

$$\begin{aligned} \|U(A)\|_F^2 &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n |a_{ij}|^2 \\ &\leq \sum_{i=1}^{n-1} \sum_{j=i+1}^n (j-i)|a_{ij}|^2 \\ &= \sum_{j=1}^{n-1} \sum_{i=j+1}^n (i-j)|a_{ij}|^2 \\ &\leq (n-1) \sum_{j=1}^{n-1} \sum_{i=j+1}^n |a_{ij}|^2 = (n-1)\|L(A)\|_F^2 \end{aligned} \quad (2.11)$$

Entonces $\|U(A)\|_F \leq \sqrt{n-1}\|L(A)\|_F$.

De manera similar:

$$\begin{aligned} \|L(A)\|_F^2 &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n |a_{ij}|^2 \\ &\leq \sum_{j=1}^{n-1} \sum_{i=j+1}^n (i-j)|a_{ij}|^2 \\ &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n (j-i)|a_{ij}|^2 \\ &\leq (n-1) \sum_{j=1}^{n-1} \sum_{i=j+1}^n |a_{ij}|^2 = (n-1)\|U(A)\|_F^2 \end{aligned} \quad (2.12)$$

Entonces $\|L(A)\|_F \leq \sqrt{n-1}\|U(A)\|_F$. ■

TEOREMA 2.10. *Sea $A, \tilde{A} \in M_n$ con A normal. Si los autovalores de A y \tilde{A} son $\lambda_1, \dots, \lambda_n$ y $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n$ respectivamente, entonces existe $\tau \in S_n$ tal que*

$$\left(\sum_{i=1}^n |\lambda_i - \tilde{\lambda}_{\tau(i)}|^2 \right)^{1/2} \leq \sqrt{n}\|A - \tilde{A}\|_F. \quad (2.13)$$

Demostración. Por el teorema de Schur, existe una matriz unitaria U tal que $U^*BU = M + T$ es triangular superior, donde $M = \text{diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_n)$ y T es estrictamente triangular superior. Sustituyendo A y \tilde{A} por U^*AU y $U^*\tilde{A}U$ respectivamente, sin pérdida de generalidad podemos suponer que $\tilde{A} = M + T$. Sea $E = A - \tilde{A}$. Entonces

$$A - M = E + T, \quad T = U(A) - U(E), \quad L(A) = L(E). \quad (2.14)$$

Ya que A y M son normales, por el teorema de Hoffman-Wielandt existe un $\tau \in S_n$ tal que:

$$\left(\sum_{i=1}^n |\lambda_i - \tilde{\lambda}_{\tau(i)}|^2 \right)^{1/2} \leq \|A - M\|_F = \|E + T\|_F. \quad (2.15)$$

Usando 2.14 y el Lema 2.10 tenemos

$$\begin{aligned} \|E + T\|_F^2 &= \|E + U(A) - U(E)\|_F^2 \\ &= \|L(E) + D(E) + U(A)\|_F^2 \\ &= \|L(E)\|_F^2 + \|D(E)\|_F^2 + \|U(A)\|_F^2 \\ &\leq \|L(E)\|_F^2 + \|D(E)\|_F^2 + (n-1)\|L(A)\|_F^2 \\ &= \|L(E)\|_F^2 + \|D(E)\|_F^2 + (n-1)\|L(E)\|_F^2 \\ &\leq n\|E\|_F^2 \end{aligned} \quad (2.16)$$

Por tanto $\|E + T\|_F \leq n\|E\|_F$. Combinamos ésta desigualdad (2.15) y que $E = A - \tilde{A}$

$$\left(\sum_{i=1}^n |\lambda_i - \tilde{\lambda}_{\tau(i)}|^2 \right)^{1/2} \leq \|E + T\|_F \leq n\|A - \tilde{A}\|_F.$$

■

EJEMPLO 2.6. Sean las matrices A y \tilde{A} de orden n

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ya que los λ_i y $\tilde{\lambda}_i$ son todos ceros y $\|A - \tilde{A}\|_F = 1$ entonces tenemos:

$$0 \leq \sqrt{n}.$$

Vemos que la constante \sqrt{n} en la desigualdad es la mejor posible.

2.4. Matiz hermitiana con perturbación antihermitiana.

Las matrices hermitianas son una generalización natural de matrices simétricas reales a matrices complejas.

TEOREMA 2.11. Sean A una matriz hermitiana y B una matriz antihermitiana con sus respectivos autovalores satisfaciendo $|\lambda_1| \geq \dots \geq |\lambda_n|$ y $|\sigma_1| \leq \dots \leq |\sigma_n|$. Entonces $|\lambda_i - \sigma_j| \leq \|A - B\|$ para todo $i = 1, 2, \dots, n$.

Demostración. Para cualquier índice j considerar los autovalores $\{\lambda_1, \dots, \lambda_j\}$ y $\{\sigma_1, \dots, \sigma_n\}$ de A y B respectivamente. Elegimos un vector x en la intersección de los subespacios de A y B . Recordar que $\|X^*\| = \|X\|$ para cada matriz X . Dado que A es hermitiana y B es antihermitiana, tenemos

$$\|A - B\| = \|A^* + B^*\| = \|(A + B)^*\| = \|A + B\|.$$

Por lo tanto de cumple la siguiente igualdad,

$$\begin{aligned} \|A - B\|^2 &= \frac{1}{2} \{ \|A - B\|^2 + \|A + B\|^2 \} \\ &\geq \frac{1}{2} \{ \|(A - B)x\|^2 + \|(A + B)x\|^2 \}. \end{aligned}$$

Usando la identidad del paralelogramo y el hecho que λ_j es real y que σ_j es imaginario, tenemos

$$\begin{aligned} &= \|Ax\|^2 + \|Bx\|^2 \\ &\geq |\lambda_j|^2 + |\sigma_j|^2 \\ &= |\lambda_j - \sigma_j|^2. \end{aligned}$$

Por lo tanto $|\lambda_i - \sigma_j| \leq \|A - B\|$ para todo $i = 1, 2, \dots, n$. ■

EJEMPLO 2.7. Sea A una matriz de orden 4 (hermitiana).

$$A = \begin{bmatrix} 6 & 9 & -i \\ 9 & 0 & 1 - 4i \\ i & 1 + 4i & 8 \end{bmatrix}.$$

Utilizando el software Scicoslab obtenemos los autovalores de A

```
--> [v, s]=spec (A)
s =

- 7.0204976      0.      0.
  0.      6.7831886      0.
  0.      0.      14.237309

v =

0.1628368 - 0.5243256i - 0.1019846 + 0.4848031i
- 0.2355796 + 0.7794816i - 0.0088748 + 0.1374838i
0.1883546      0.8576623

0.1187049 - 0.6626018i
0.1086455 - 0.5532868i
0.4784746
```

por lo tanto $\mathcal{L}(A) = \{-7.0204976; 6.7831886; 14.237309\}$.

Sea B una matriz de orden 4 (antihermitiana)

$$B = \begin{bmatrix} 2i & 5 & -i \\ -5 & -2i & 1+2i \\ -i & -1+2i & -2i \end{bmatrix}.$$

Nuevamente utilizamos el software Scicoslab para obtener los autovalores de B

```
--> [v2, s2]=spec (B)
s2 =

5.8068122i      0
0      8.866D-16 - 6.1178783i
0      0

0
0
- 1.998D-16 - 1.6889338i

v2 =

0.7911426      - 0.0483274 + 0.4733521i
- 0.0282822 + 0.5661257i      0.7835908
- 0.1811024 + 0.1414110i - 0.3923159 - 0.0753395i
```

$$\begin{aligned} &0.1802640 + 0.3394121i \\ &0.2504138 + 0.0444992i \\ &0.8874779 \end{aligned}$$

por lo tanto

$$\mathcal{L}(B) = \{5,8068122i; 8,866D - 16 - 6,1178783i; -1,998D - 16 - 1,6889338i\}.$$

Por un lado tenemos

$$\begin{aligned} |\lambda_1 - \sigma_1| &= |14,237309 + 6,1178783i| = 15,496109 \\ |\lambda_2 - \sigma_2| &= |-7,0204976 - 5,8068122i| = 9,1107878 \\ |\lambda_3 - \sigma_3| &= |6,7831886 + 1,6889338i| = 6,9902893. \end{aligned}$$

Por otro lado

$$A - B = \begin{pmatrix} 6 - 2i & 4 & 0 \\ 14 & 2i & -6i \\ 2i & 2 + 2i & 8 + 2i \end{pmatrix}$$

entonces $\|A - B\| = 17,284055$.

Por consiguiente $|\lambda_i - \sigma_i| \leq \|A - B\|$ para $i = 1, 2, 3$.

TEOREMA 2.12. Sea $T = A + iB \in M_n$ con A y B hermitianas. Si α_j y β_j son autovalores de A y B respectivamente con $|\alpha_1| \geq \dots \geq |\alpha_n|$ y $|\beta_1| \geq \dots \geq |\beta_n|$, entonces

$$\| \text{Diag}(\alpha_1 + i\beta_1, \dots, \alpha_n + i\beta_n) \| \leq \sqrt{2} \|T\|$$

para cualquier norma invariante.

Demostración. Dado que los valores singulares de una matriz hermitiana son los valores absolutos de los autovalores, tenemos

$$s(\text{Diag}(\alpha_1 + i\beta_1, \dots, \alpha_n + i\beta_n)) = |s(A) + is(B)|. \quad (2.17)$$

Por el principio dominante de Fan (1.15), es suficiente probar (2.17) para todas las k -normas de Fan, $k = 1, 2, \dots, n$. Primero probaremos que (2.17) se cumple para la norma de las trazas $\|\cdot\|_{(n)}$ y la norma espectral $\|\cdot\|_{(1)}$. El caso $p = 1$ de la desigualdad (1.16) demuestra que (2.17) cumple para la norma de la traza. De

$$A = \frac{T + T^*}{2}, \quad B = \frac{T - T^*}{2i}$$

tenemos $|\alpha_1| = \|A\|_{(1)} \leq \|T\|_{(1)}$ y $|\beta_1| = \|B\|_{(1)} \leq \|T\|_{(1)}$. Por lo tanto $|\alpha_1 + i\beta_1| \leq \sqrt{2}\|T\|_{(1)}$. Así (2.17) cumple para la norma espectral.

Ahora fijando un k arbitrario con $1 \leq k \leq n$. Por el Lema (1.3), existen $X, Y \in M_n$ satisfaciendo $T = X + Y$ y

$$\|T\|_{(k)} = \|X\|_{(n)} + k \|Y\|_{(1)}. \quad (2.18)$$

Sean $X = C + iD$ y $Y = E + iF$ las descomposiciones cartesianas de X y Y respectivamente; es decir, C, D, E, F son matrices hermitianas. Entonces $C + E + i(D + F)$. Dado que la descomposición cartesiana es única, tenemos

$$A = C + E, \quad B = D + F. \quad (2.19)$$

Por los dos casos ya probados de (2.17), tenemos

$$\sqrt{2} \|X\|_{(n)} \geq \Phi(|s(C) + is(D)|) \quad (2.20)$$

$$\sqrt{2} \|Y\|_{(1)} \geq \Phi(|s(E) + is(F)|), \quad (2.21)$$

donde $\Phi_{(k)}$ es la k - norma de Fan en \mathbb{C}^n . Combinando (2.18), (2.20) y (2.21), obtenemos

$$\begin{aligned} \sqrt{2} \|T\|_{(k)} &\geq \Phi_n(|s(C) + is(D)|) + k\Phi(|s(E) + is(F)|) \\ &\geq \sum_{j=1}^k |s_j(C) + is_j(D)| + k|s_1(E) + is_1(F)|. \end{aligned}$$

Así

$$\sqrt{2} \|T\|_{(k)} \geq \sum_{j=1}^k |s_j(C) + is_j(D)| + k|s_1(E) + is_1(F)|. \quad (2.22)$$

Para $1 \leq j \leq n$, por el Teorema (1.14) obtenemos

$$s_j(C + E) \leq s_j(C) + s_1(E), \quad s_j(D + F) \leq s_j(D) + s_1(F). \quad (2.23)$$

Usando (2.23) y (2.19) tenemos

$$\begin{aligned} &\sum_{j=1}^k |s_j(C) + is_j(D)| + k|s_1(E) + is_1(F)| \\ &= \sum_{j=1}^k \{|s_j(C) + is_j(D)| + |s_1(E) + is_1(F)|\} \\ &\geq \sum_{j=1}^k |[s_j(C) + s_1(E)] + i[s_j(D) + s_1(F)]| \\ &\geq \sum_{j=1}^k |s_j(C + E) + is_j(D + F)| \\ &= \Phi_k(|s(A) + is(B)|). \end{aligned}$$

Por lo tanto

$$\sum_{j=1}^k |s_j(C) + is_j(D)| + k|s_1(E) + is_1(F)| \geq \Phi_k(|s(A) + is(B)|). \quad (2.24)$$

Combinando (2.22) y (5.21), obtenemos

$$\sqrt{2} \|T\|_k \geq \Phi_k(|s(A) + is(B)|), \quad k = 1, 2, \dots, n.$$

Esto completa la demostración. ■

Capítulo 3

Aplicaciones

3.1. Page Rank

¿Alguna vez nos hemos preguntado, cómo determina el motor de búsqueda Google el orden en que se muestran las páginas web?.

El ingrediente principal en la determinación de este orden es el vector PageRank, que asigna una puntuación a cada página web, las que tengan altas puntuaciones se muestran primero y las páginas web con puntuaciones bajas se muestran después. El PageRank de una página web se basa en la cantidad y calidad de los enlaces que apuntan hacia ella, las cuáles no depende del contenido de las páginas web. Su importancia se enfatiza en una de las páginas web de Google [20].

“El corazón de nuestro software es PageRank™, un sistema de clasificación de páginas web desarrollado por nuestros fundadores Larry Page y Sergey Brin en la Universidad de Stanford. Y mientras que tenemos docenas de ingenieros trabajando para mejorar cada aspecto de Google sobre una base diaria, PageRank continúa proporcionando la base para todas nuestras herramientas de la búsqueda web”.

El vector PageRank es la distribución estacionaria de una matriz estocástica, la matriz de Google, ésta es a su vez es una combinación convexa de dos matrices estocásticas: una matriz representa la estructura de enlace del gráfico web y una segunda, la matriz de rango uno, imita el comportamiento aleatorio de los internautas y también se puede utilizar para combatir el spam web. Como consecuencia, el PageRank depende principalmente de la estructura de enlace del gráfico web. Analizaremos a continuación la sensibilidad de PageRank a los cambios en la matriz de Google, pero antes modelaremos el algoritmo.

3.1.1. PageRank and the Google Matrix

La estructura del enlace del gráfico web puede representarse matemáticamente como una matriz H . Supongamos que la página web i tiene $l_i > 0$ enlaces externos. Si la página i contiene un enlace en otra página $j \neq i$, entonces $H_{i,j} = 1/l_i$, de lo contrario, $H_{i,j} = 0$. El elemento matricial $H_{i,j}$ representa la probabilidad de que un internauta siga el enlace de la página i a la página j . Si la página web i no tiene enlaces externos, entonces la fila i de H es cero. Entonces la página es llamada nodo colgante, ya que puede ser un archivo pdf o una página cuyos vínculos no se han rastreado.

Para transformar H en una matriz estocástica S , se puede llenar cada fila correspondiente a un nodo colgante con un vector w^T . Eso es:

$$S = H + dw^T$$

donde $d_i = 1$ si la página i no tiene enlaces externos, y $d_i = 0$ de lo contrario, y w es un vector columna, donde $w \geq 0$ y $\|w\|_1 = 1$. Una opción común es fijar a cada elemento $1/n$ en las filas del nudo colgante, donde n es el número de nodos en el gráfico web; en otras palabras $w = \frac{1}{n}\mathbf{1}$, donde $\mathbf{1}$ es el vector columna donde todos sus elementos son 1.



Figura 3.1: Gráfico web.

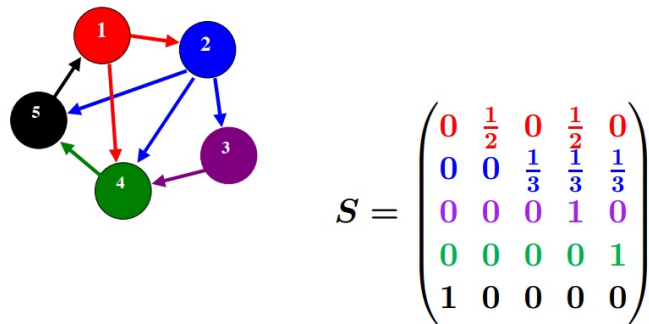


Figura 3.2: El gráfico web como matriz.

3.1.2. Propiedades de la matriz S

- S es estocástica: $0 \leq s_{i,j} \leq 1$, S es estocástica si sus elementos en cada fila suman 1.
- Autovector dominante por la izquierda: $\omega^T S = \omega^T$, $\omega \geq 0$, $\|\omega\|_1 = 1$.
- ω_i es la probabilidad de que el navegante visite la página i .

Pero existen factores en contra:

- S no modela correctamente el comportamiento de navegación.
- Clasifica el orden más bajo y las páginas con rango cero.
- Varios autovalores de magnitud 1.

Entonces ω no es único.

Por lo tanto no es conveniente trabajar con la matriz S . Para remediar estos factores que no nos permitirán estudiar la matriz y llegar al objetivo principal que es ordenar las páginas, cambiaremos la matriz para obtener una más conveniente. A continuación construiremos esa matriz denominada la matriz Google G .

3.1.3. Matriz Google

La matriz de Google es definida como una combinación convexa de S y una matriz de rango 1, es decir:

$$G = \alpha S + (1 - \alpha)\mathbf{1}v^T$$

Donde:

- S es matriz estocástica.
- Factor de amortiguamiento $0 \leq \alpha < 1$, originalmente esta fijado en $\alpha = 0,85$, esto modela la posibilidad de que un navegante salte de una página web a la siguiente sin necesariamente seguir un enlace.
- $\mathbf{1}$ es el vector columna con todas sus entradas igual a 1.
- Vector de personalización $v \geq 0$, $\|v\|_1 = 1$, se utiliza para combatir el spamming de enlaces. Si $v_i = 0$ entonces la página i es una página spam.

3.1.4. Características de la matriz G

- La matriz G es estocástica por filas y, en general, reducible.
- Los autovalores de G son: $1 > \alpha_2(S) \geq \alpha_3(S) \geq \dots$
- Tiene un único autovector dominante por izquierda:

$$\pi^T G = \pi^T, \quad \pi \geq 0, \quad \|\pi\|_1 = 1.$$

Para mayor referencia ver [8].

Por lo tanto, el PageRank vector se define como la distribución estacionaria de G , dónde π_i es el **PageRank** de la página web i .

DEFINICIÓN 3.1. Una distribución de probabilidad $\pi = (\pi_0, \pi_1, \dots)$ es **estacionaria** o invariante para la cadena de Markov con matriz P . Si :

$$\pi = \pi P$$

DEFINICIÓN 3.2. Una distribución de probabilidad es el conjunto de datos que representan los elementos de un espacio muestral con sus respectivas probabilidades.

DEFINICIÓN 3.3. Una cadena de Markov es una serie de eventos, en la cual la probabilidad de que ocurra un evento depende del evento inmediato anterior. En efecto, las cadenas de este tipo de memoria, recuerdan el último evento y esto condiciona las posibilidades de los eventos futuros.

3.1.5. Estabilidad del PageRank

Estudiaremos qué tan sensible es el PageRank π a:

- Cambios en la matriz S .
- Cambios en el factor de amortiguamiento α .
- Cambios en el vector de personalización v .

Cambios en la matriz S

La sensibilidad del vector PageRank π a cambios en S depende del $k(S) \leq \frac{\alpha}{1-\alpha}$. Específicamente, sea $S + E$ una matriz estocástica y establecemos:

$$\tilde{G} = \alpha(S + E) + (1 - \alpha)\mathbf{1}v^T.$$

El vector PageRank perturbado es $\tilde{\pi}$, donde:

$$\tilde{\pi}^T \tilde{G} = \tilde{\pi}^T, \quad \tilde{\pi} \geq 0 \text{ y } \|\tilde{\pi}\|_1 = 1.$$

Obtenemos el error absoluto en $\tilde{\pi}$:

$$\tilde{\pi}^T - \pi^T = \alpha \tilde{\pi}^T E (I - \alpha S)^{-1}, \quad \|\tilde{\pi} - \pi\|_1 \leq \frac{\alpha}{1 - \alpha} \|E\|_\infty.$$

Para el factor de amortiguación original:

$$\text{Si } \alpha = 0,85 \Rightarrow \frac{\alpha}{1 - \alpha} \approx 5,7.$$

Incluso para los factores de amortiguación más grandes la sensibilidad sigue siendo baja:

$$\text{Si } \alpha = 0,99 \Rightarrow \frac{\alpha}{1 - \alpha} = 99$$

Cambios en el factor de amortiguamiento α

La sensibilidad del vector PageRank π a cambios en el factor de amortiguamiento α depende del número de condición que es acotado por $\frac{2}{1-\alpha}$. Específicamente, sea $0 \leq \alpha + \mu < 1$ un factor de amortiguamiento perturbado, y establezcamos:

$$\tilde{G} = (\alpha + \mu)S + (1 - (\alpha + \mu))\mathbf{1}v^T.$$

El vector PageRank perturbado es $\tilde{\pi}$, donde $\tilde{\pi}^T \tilde{G} = \tilde{\pi}^T$, $\tilde{\pi} \geq 0$, y $\|\tilde{\pi}\|_1 = 1$. El error en $\tilde{\pi}$ puede ser acotado por

$$\|\tilde{\pi} - \pi\|_1 \leq \frac{2}{1 - \alpha} |\mu|.$$

La cota del número de condición $\frac{2}{(1-\alpha)}$ es una función creciente en α . Comparando esta cota con las cotas para el número de condición $\frac{\alpha}{(1-\alpha)}$ que vimos anteriormente. Vemos que π es ligeramente más sensible a cambios en el parámetro α que a los cambios en la matriz S . Para el factor de amortiguamiento $\alpha = 0,85$, el número de condición es $\frac{2}{1-\alpha} \approx 13,4$. Para $\alpha = 0,99$, obtenemos $\frac{2}{1-\alpha} = 200$.

Cambios en el vector de personalización v

El vector PageRank π es perfectamente condicionado respecto a los cambios en el vector de personalización v . Específicamente, sea $v + f$ el vector de personalización perturbado con $v + f \geq 0$ y $\|v + f\|_1 = 1$, y establezcamos:

$$\tilde{G} = \alpha S + (1 - \alpha)\mathbf{1}(v + f)^T.$$

El vector PageRank perturbado es $\tilde{\pi}$, donde $\tilde{\pi}^T \tilde{G} = \tilde{\pi}^T$, $\tilde{\pi} \geq 0$, y $\|\tilde{\pi}\|_1 = 1$. La corta del error para $\tilde{\pi}$ contiene un número de condición que es acotado por 1.

$$\|\tilde{\pi} - \pi\|_1 \leq \|f\|_1.$$

3.1.6. Ejemplo aplicativo

Crearemos y trazaremos un gráfico que contiene seis nodos (páginas web) que representan sitios web ficticios. Cada página web estará representada de la siguiente forma:

$\alpha = 1$; $\beta = 2$; $\gamma = 3$; $\delta = 4$; $\epsilon = 5$ y por último $\zeta = 6$.

La matriz s representa el sitio web de inicio y la matriz t representa el sitio web de llegada.

```
Trial>> s = [1 1 2 2 3 3 3 4 5];
t = [2 5 3 4 4 5 6 1 1];
names = {'http://www.example.com/alpha', 'http://www.example.com/beta', ...
         'http://www.example.com/gamma', 'http://www.example.com/delta', ...
         'http://www.example.com/epsilon', 'http://www.example.com/zeta'};
G = digraph(s,t,[],names)

G =

    digraph with properties:

      Edges: [9×1 table]
      Nodes: [6×1 table]

Trial>> plot(G,'Layout','layered', ...
             'NodeLabel',{'alpha','beta','gamma','delta','epsilon','zeta'})
```

Figura 3.3: Programación en matlab .

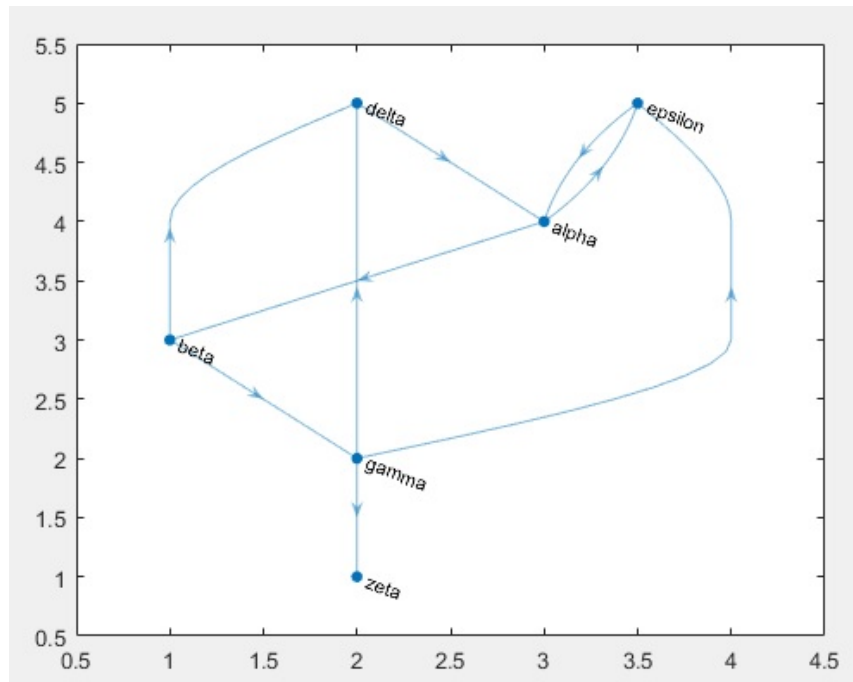


Figura 3.4: Gráfica.

Calculamos el puntaje de PageRank para este gráfico. Utilizamos el factor de amortiguación de 0,85.

```

Trial>> pr = centrality(G, 'pagerank', 'FollowProbability', 0.85)

pr =

    0.3210
    0.1706
    0.1066
    0.1368
    0.2008
    0.0643

```

Veamos las puntuaciones del PageRank y la información de grado para cada página.

```

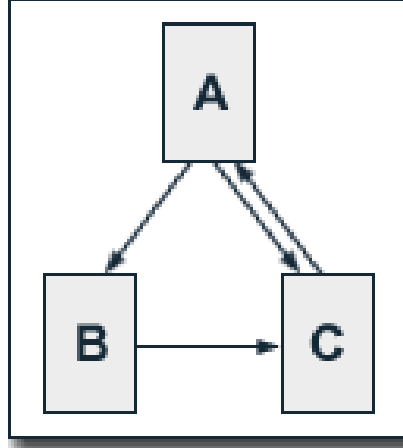
Trial>> G.Nodes.PageRank = pr;
G.Nodes.InDegree = indegree(G);
G.Nodes.OutDegree = outdegree(G);
G.Nodes

ans =

```

Name	PageRank	InDegree	OutDegree
'http://www.example.com/alpha'	0.32098	2	2
'http://www.example.com/beta'	0.17057	1	2
'http://www.example.com/gamma'	0.10657	1	3
'http://www.example.com/delta'	0.13678	2	1
'http://www.example.com/epsilon'	0.20078	2	1
'http://www.example.com/zeta'	0.06432	1	0

EJEMPLO 3.1. Veamos un ejemplo, tenemos un sitio web con tres páginas, A , B y C , donde A enlaza B y C , B enlaza C , y C enlaza A .



Ahora la estructura de los enlaces del gráfico web lo representaremos matemáticamente según la teoría descrita en este capítulo.

Primero obtenemos la matriz $H = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$.

Para transformar la matriz H en una matriz estocástica S , utilizaremos la siguiente fórmula:

$$\begin{aligned}
 S &= H + dw^T \\
 S &= \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 & 1/3 \end{pmatrix} \\
 S &= \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
 S &= \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.
 \end{aligned}$$

Dado que no es conveniente trabajar con la matriz S , hallaremos G que es la

matriz de google,

$$G = \alpha S + (1 - \alpha) \mathbf{1}v^T$$

$$G = 0,85 \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} + 0,15 \times 1/3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$$

$$G = \begin{pmatrix} 0 & 0,425 & 0,425 \\ 0 & 0 & 0,85 \\ 0,85 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0,05 & 0,05 & 0,05 \\ 0,05 & 0,05 & 0,05 \\ 0,05 & 0,05 & 0,05 \end{pmatrix}$$

$$G = \begin{pmatrix} 0,05 & 0,475 & 0,475 \\ 0,05 & 0,05 & 0,9 \\ 0,9 & 0,05 & 0,05 \end{pmatrix}.$$

Observe que la matriz de google G es estocástica por filas, sus autovalores son $1; -0,425 + 0,425i; -0,425 - 0,425i$.

Perturbando el factor de amortiguamiento $\alpha + \mu = 0,85 + 0,05 = 0,9$, obtenemos

$$\tilde{G} = (\alpha + \mu)S + (1 - (\alpha + \mu))\mathbf{1}v^T$$

$$\tilde{G} = 0,9 \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} + 0,1 \times 1/3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$$

$$\tilde{G} = \begin{pmatrix} 0,0333333 & 0,4833333 & 0,4833333 \\ 0,0333333 & 0,0333333 & 0,9333333 \\ 0,9333333 & 0,0333333 & 0,0333333 \end{pmatrix}.$$

Los autovalores de \tilde{G} son $1; -0,45 + 0,45i; -0,45 - 0,45i$.

Analizando las matrices G y \tilde{G} se puede observar que nos encontramos en el primer caso del análisis de perturbación descrita en el capítulo 2, es decir, G es una matriz general con perturbación general.

El teorema de Bhatia Elsner-Krause (Teorema 2.4) nos dice que la distancia óptima entre G y \tilde{G} está acotada por $4(\|G\| + \|\tilde{G}\|)^{1-\frac{1}{3}} \|G - \tilde{G}\|^{\frac{1}{3}}$.

Veamos,

$$\begin{aligned} \|G\|_2 &= \sqrt{\sum_{i=1}^3 \sum_{j=1}^3 |a_{ij}|^2} & , & \quad \|\tilde{G}\|_2 = \sqrt{\sum_{i=1}^3 \sum_{j=1}^3 |\tilde{a}_{ij}|^2} \\ &= 1,44352 & & \quad = 1,48829 \end{aligned}$$

y

$$\begin{aligned}\|G - \tilde{G}\|_2 &= \sqrt{\sum_{i=1}^3 \sum_{j=1}^3 |a_{ij} - \tilde{a}_{ij}|^2} \\ &= 0,4703722\end{aligned}$$

por lo tanto $DO(G, \tilde{G})$ está acotada por 6,37232261.

EJEMPLO 3.2. Se tiene un sitio web con cuatro páginas, tal como se muestra en la siguiente figura.

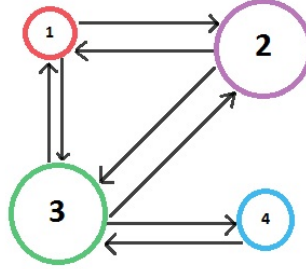


Figura 3.5: Gráfico dirigido de las cuatro páginas web.

Representamos matemáticamente la estructura del sitio web. Lo primero que se debe obtener, según la teoría descrita es la matriz H de orden 4

$$H = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 1/3 & 1/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Transformamos la matriz H en la matriz S , la cual es estocástica por filas,

$$S = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 1/3 & 1/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Hallamos la matriz G (matriz de Google), utilizamos el factor de amortiguación $\alpha = 0,85$ y $n = 4$ el orden de la matriz:

$$G = 0,85 \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 1/3 & 1/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \end{pmatrix} + 0,15 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$$

$$G = \begin{pmatrix} 0,0375 & 0,4625 & 0,4625 & 0,0375 \\ 0,4625 & 0,0375 & 0,4625 & 0,0375 \\ 0,3208333 & 0,3208333 & 0,0375 & 0,3208333 \\ 0,0375 & 0,0375 & 0,8875 & 0,0375 \end{pmatrix}$$

Específicamente, sea $0 \leq \alpha + \mu < 1$ un factor de amortiguamiento perturbado, establezcamos a $\mu = 0,05$ entonces tenemos:

$$\tilde{G} = 0,9 \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 1/3 & 1/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \end{pmatrix} + 0,1 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$$

$$\tilde{G} = \begin{pmatrix} 0,025 & 0,475 & 0,475 & 0,025 \\ 0,475 & 0,025 & 0,475 & 0,025 \\ 0,325 & 0,325 & 0,025 & 0,325 \\ 0,025 & 0,025 & 0,925 & 0,025 \end{pmatrix}$$

Los autovalores de la matriz G son:

```
vals =

    1.         0         0         0
    0    - 0.425         0         0
    0         0    - 0.6194065         0
    0         0         0    0.1944065

vecs =

- 0.5    - 0.7071068    0.2137284    0.3720808
- 0.5     0.7071068    0.2137284    0.3720808
- 0.5     3.641D-16    - 0.5787102    - 0.1784366
- 0.5    - 3.187D-16     0.7574528    - 0.8314277
```

Utilizando Matlab de la misma manera hallamos el espectro de \tilde{G}

```
vals =

    1.         0         0         0
    0    - 0.45         0         0
    0         0    - 0.6558422         0
    0         0         0    0.2058422

vecs =
```

```

- 0.5 - 0.7071068      0.2183880      0.3666639
- 0.5      0.7071068      0.2183880      0.3666639
- 0.5      1.199D-16 - 0.5714315 - 0.1831087
- 0.5 - 2.355D-16      0.7603153 - 0.8352164

```

Analizando las matrices G y \tilde{G} se puede observar que nos encontramos en el primer caso del análisis de perturbación descrita en el capítulo 2.

El teorema de Elsner nos dice que la distancia de Hausdorff entre G y \tilde{G} está acotada por $(\|G\|_2 + \|\tilde{G}\|_2)^{1-\frac{1}{4}} \|E\|_2^{\frac{1}{4}}$.

Veamos,

$$\begin{aligned}\|G\|_2 &= \sqrt{\sum_{i=1}^3 \sum_{j=1}^3 |a_{ij}|^2} & , \quad \|\tilde{G}\|_2 &= \sqrt{\sum_{i=1}^3 \sum_{j=1}^3 |\tilde{a}_{ij}|^2} \\ &= 1,963333269 & & = 2,08\end{aligned}$$

y

$$\begin{aligned}\|E\|_2 &= \sqrt{\sum_{i=1}^3 \sum_{j=1}^3 |a_{ij}|^2} \\ &= 0,0031771\end{aligned}$$

por lo tanto DH está acotada por 0,676959.

Conclusiones

1. El análisis cualitativo del espectro de matrices perturbadas, realizado en esta tesis, ha dado como resultado más significativo, en el análisis de la teoría de perturbaciones para matrices generales, el establecimiento de una cota para la perturbación la cual debe ser pequeña, y el teorema de Elsner el cual nos proporcionó las cotas explícitas; por último el teorema de Bathia-Elsner-Krause establece una cota para la distancia óptima de los autovalores de una matriz.
2. Se estudió la teoría de perturbaciones para matrices normales, teniendo como resultado que la perturbación hecha en los autovalores depende del número de condición de la matriz de paso P , también que las matrices óptimas para el cálculo efectivo de sus autovalores y autovectores son las matrices normales.
3. El teorema de Kittaneh establece una cota para la diferencia de la matriz con su perturbación. Se mostró que las matrices estocásticas pueden ser expresadas como una combinación convexa de matrices de perturbación, este estudio nos ayudó para la aplicación relacionada con el Page rank.
4. Analizamos el teorema de Hoffman-Wielandt con un ejemplo. Se observa que al aplicar el teorema este se complica por el hecho de que la suma de dos matrices normales no es necesariamente una matriz normal. Por lo tanto las perturbaciones que el teorema puede manejar es limitado.
5. En el contexto de la teoría de perturbación se introdujeron entre otros resultados, los casos de matriz normal con perturbación general y de matriz hermitiana con perturbación antihermitiana, y se obtuvo una cota de la diferencia de los autovalores de la matriz y su perturbación. Aunque no se estudia ampliamente, quedan ciertas conjeturas ya que es un tema abierto aún.

6. Se desarrolló la aplicación de Google denominada **PageRank** la cual mediante el estudio de la teoría de perturbación permite poder ordenar las páginas web de Google según la cantidad de enlaces que se encuentran en dicha página. Dado que existe gran cantidad de páginas web, el motor de búsqueda obtiene una matriz demasiado grande, haciéndola sensible al cálculo de autovalores. Estudiamos las propiedades de la matriz de Google y el vector de PageRank, así como también su estabilidad con cambios en la matriz S , cambios en el factor de amortiguamiento y cambios en el vector de personalización v . Por último utilizamos resultados del capítulo II para poder mostrar la consistencia del algoritmo Pagerank en lo que al cálculo de autovalores concierne.

Sugerencias

Una vez concluida la tesis, se considera interesante investigar sobre otros aspectos que se deben tener en cuenta para el análisis de la teoría de perturbación y se propone:

- Analizar con mayor detenimiento la perturbación para matrices hermitianas.
- Generar matrices de prueba y hallar para cada matriz las distintas cotas que se detallaron.
- Extender los estudios expuestos en esta tesis al estudio de otras aplicaciones, tales como el Eigenface, que es el nombre dado a un conjunto de auto-vectores cuando se utilizan en el problema de la visión por computador del reconocimiento de la cara humana.

Bibliografía

- [1] B. Rajendra. *Perturbation Bounds for Matrix Eigenvalues*. Siam, Society for Industrial and Applied Mathematics.
- [2] C.K. Chui. P. Monk. L. Wuytack. *Perturbation theory for matrix equations*. ELSEVIER SCIENCE B.V, 2003.
- [3] C. Francoise. *Eigenvalues of Matrices*. Siam, Society for Industrial and Applied Mathematics, Philadelphia, 2012.
- [4] G. W. Stewart. *Matrix Algorithms*. Volume I: Basic Decompositions. University of Maryland. Siam, Society for Industrial and Applied Mathematics.
- [5] G. W. Stewart. Ji-guang Sun. *Matrix Perturbation Theory*. ACADEMIC PRESS, INC. Harcourt Brace Jovanovich, Publishers, 1990.
- [6] J. P. Lowel. J. Dean. A. S. Thomas. *Elementos de álgebra lineal*. Segunda edición. Editorial Reverté, S. A., 1982.
- [7] R. A. Horn, C. R., Jhonson. *Matrix Analysis, 2nd edn*. Cambridge University Press, Cambridge, UK, 2012.
- [8] N. L. Amy, D. M. Carl. *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton university press.
- [9] R. Bellman. *Introduction to Matrix Analysis*. SIAM, Philadelphia, Reprint of the Second edition, 1995.
- [10] R. Bhatia. *Matrix Analysis*. Springer, New York, 1997.
- [11] R. Bhatia. *Perturbation Bounds for Matrix Eigenvalues*. Longman, Harlow, 1987; 2nd ed., SIAM, Philadelphia, 2007.

- [12] R. Bhatia. *Spectral variation, normal matrices, and Finsler geometry*, Math. Intelligencer, 2007.
- [13] S.S. Aghaian. *Hadamard Matrices and Their Applications*. LNM 1168, Springer, 1985.
- [14] T. Ando. *Majorization, doubly stochastic matrices, and comparison of eigenvalues*. Linear Algebra Appl, 1989.
- [15] T. J. Rivlin. *An Introduction to the Approximation of Functions*. Dover, New York, 1981.
- [16] V. C. Oswaldo. *Análisis complejo y sus aplicaciones*. Instituto de Matemática y Ciencias Afines y Facultad de Ciencias, Universidad Nacional de Ingeniería. Lima.
- [17] Z. Ion. *Análisis matricial aplicado y ampliación de métodos numéricos*. Departamento de Matemática Aplicada y Estadística e Investigación Operativa. Euskal Herriko Unibertsitatea.
- [18] Z. Fuzhen. *Matrix Theory*. Basic Results and Techniques. Second Edition. Springer, 2010.
- [19] Z. Xingzhi. *Matrix Theory*. Graduate Studies in Mathematics. Volume 147. American Mathematical Society.
- [20] [http://www.google.com/ technology/](http://www.google.com/technology/)