

Complemento sobre el Teorema de Gershgorin

JP

Contents

1	Introducción	1
2	Ideas en las que se basa la prueba de la parte 2)	2
2.1	Comentarios generales	2
2.2	Ideas subyacentes	3
3	Aplicación de estas ideas	4
4	Ejemplos	4
5	La matriz traspuesta	12

1 Introducción

El objetivo de este material es complementar los contenidos de los apuntes del curso (Libro Rojo), en particular ilustrando la parte 2) de la tesis del Teorema de Gershgorin y poniendo énfasis en algunas ideas.

Recordemos lo que el Teorema establece: Tenemos una matriz cuadrada $\mathbb{A} \in \mathcal{M}_{n \times n}(\mathbb{C})$ con **entradas complejas** y para cada fila $i = 1, \dots, n$ de la misma tenemos el disco (o bola cerrada) C_i de centro a_{ii} y radio $r_i = \sum_{j \neq i} |a_{ij}|$.

Esto es, el disco está centrado en el elemento diagonal de la fila y su radio es la suma **de los módulos** del resto de la fila.

En otras palabras, $C_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}$, es el conjunto de todos los **números complejos** cuya distancia a a_{ii} es **menor o igual** a r_i (otra notación es $C_i = B[a_{ii}, r_i]$).

El Teorema hace 2 afirmaciones:

1. Si λ es un valor propio de \mathbb{A} entonces λ no puede estar en cualquier parte del plano complejo, sino que debe hallarse en alguno de los discos (puede que en más de uno).

Esto es, Si λ es un valor propio de \mathbb{A} entonces $\lambda \in \bigcup_{i=1}^n C_i$.

2. Si m de estos discos forman una unión \mathbb{M} que es disjunta de los restantes $n-m$ discos, entonces en esta unión \mathbb{M} hay exactamente m valores propios, **contados con su multiplicidad algebraica** (un valor propio que es raíz múltiple del polinomio característico de \mathbb{A} se cuenta tantas veces como su multiplicidad).

Esto último significa que si p.ej. tenemos una unión \mathbb{M} de 3 discos disjunta con los restantes discos, entonces en alguna parte de \mathbb{M} hay exactamente tres valores propios $\lambda_1, \lambda_2, \lambda_3$, que pueden ser los tres distintos, o puede ser un valor propio que es raíz triple del polinomio característico ($\lambda_1 = \lambda_2 = \lambda_3$), o pueden ser una raíz doble y otra raíz simple ($\lambda_1 = \lambda_2, \lambda_1 \neq \lambda_3$).

Para la demostración remitimos al Libro Rojo.

2 Ideas en las que se basa la prueba de la parte 2)

2.1 Comentarios generales

La prueba de la parte 2) se basa en argumentos de continuidad, externos al curso de GAL.

En efecto, si lo piensan, durante gran parte del curso la noción de distancia no forma parte de la teoría, recién cuando se introduce el concepto de norma podemos hablar de distancias entre puntos (vectores) en un Espacio Vectorial.

Pero ni siquiera entonces aprovechamos para introducir nociones de convergencia, de límite, nunca hablamos de que pueda haber una sucesión de vectores que tiende a otro vector.

No es que no se pueda hacer, de hecho hay toda una rama de las Matemáticas que hace análisis en Espacios Normados, solamente que no es el camino que se sigue en GAL 2.

Volviendo al tema, la idea es construir una familia de matrices que empieza en la matriz diagonal \mathbb{D} que tiene la diagonal principal de \mathbb{A} y termina en la matriz \mathbb{A} .

Esto se hace separando \mathbb{A} como la suma de 2 matrices: la matriz \mathbb{D} que aporta la diagonal principal y la matriz $\mathbb{E} = \mathbb{A} - \mathbb{D}$, que aporta el resto de las entradas.

Esto se hace para cada $\varepsilon \in [0, 1]$ como sigue: $\mathbb{A}_\varepsilon = \mathbb{D} + \varepsilon\mathbb{E}$.

Lo de familia en vez de sucesión se debe a que en lugar de tener una matriz \mathbb{A}_k para cada natural $k \in \mathcal{N}$, tenemos una matriz \mathbb{A}_ε para cada $\varepsilon \in [0, 1]$.

Hablando pronto y mal podemos decir que la familia de matrices $\{\mathbb{A}_\varepsilon, \varepsilon \in [0, 1]\}$ es una sucesión indexada en $[0, 1]$ en vez de ser indexada en \mathcal{N} .

Observaciones

- $\mathbb{A}_0 = \mathbb{D}$ y $\mathbb{A}_1 = \mathbb{A}$
- Todas las matrices \mathbb{A}_ε tienen la misma diagonal principal (que es la diagonal principal de \mathbb{A}).
- A medida que ε se va desplazando desde 0 hacia 1, la matriz \mathbb{A}_ε se va pareciendo más y más a la matriz \mathbb{A} .

Ilustramos esto último tomando $\varepsilon = \{0, 0.2, 0.4, 0.6, 0.8, 1\}$ y la matriz

$$\mathbb{A} = \begin{pmatrix} 1 & -3 & 2 \\ -2 & -4 & 3 \\ 3 & -1 & 2 \end{pmatrix}.$$

Entonces tendremos $\mathbb{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, $\mathbb{E} = \begin{pmatrix} 0 & -3 & 2 \\ -2 & 0 & 3 \\ 3 & -1 & 0 \end{pmatrix}$ y las

matrices correspondientes

$$\mathbb{A}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \mathbb{A}_{0.2} = \begin{pmatrix} 1 & -0.6 & 0.4 \\ -0.4 & -4 & 0.6 \\ 0.6 & -0.2 & 2 \end{pmatrix}, \mathbb{A}_{0.4} = \begin{pmatrix} 1 & -1.2 & 0.8 \\ -0.8 & -4 & 1.2 \\ 1.2 & -0.4 & 2 \end{pmatrix},$$

$$\mathbb{A}_{0.6} = \begin{pmatrix} 1 & -1.8 & 1.2 \\ -1.2 & -4 & 1.8 \\ 1.8 & -0.6 & 2 \end{pmatrix}, \mathbb{A}_{0.8} = \begin{pmatrix} 1 & -2.4 & 1.6 \\ -1.6 & -4 & 2.4 \\ 2.4 & -0.8 & 2 \end{pmatrix}, \mathbb{A}_1 = \begin{pmatrix} 1 & -3 & 2 \\ -2 & -4 & 3 \\ 3 & -1 & 2 \end{pmatrix}$$

- Una forma de verlo es pensar en la matriz como una pizarra o cartel luminoso y que las entradas de \mathbb{A} fuera de la diagonal principal se iluminan más o menos según ε .

Cuando $\varepsilon = 0$, todas las entradas excepto las de la diagonal principal están apagadas, y a medida que ε se va moviendo desde 0 hacia 1 dichas entradas se van iluminando cada vez más, hasta que cuando $\varepsilon = 1$ toda la matriz se encuentra iluminada.

2.2 Ideas subyacentes

1. Lo primero es notar que esta familia de matrices “varía en forma continua con ε ”. En efecto, si tomamos dos valores $\varepsilon, \varepsilon'$ muy cercanos, entonces las matrices \mathbb{A}_ε y $\mathbb{A}_{\varepsilon'}$ tendrán entradas a_{ij} muy parecidas entre sí.

Ambas tienen la misma diagonal principal, y si miramos los elementos fuera de la diagonal principal, tenemos que $\mathbb{A}_\varepsilon(i, j) = \varepsilon a_{ij}$ y que $\mathbb{A}_{\varepsilon'}(i, j) = \varepsilon' a_{ij}$, y por tanto la resta nos da $(\varepsilon - \varepsilon')a_{ij}$.

Es claro que esto tiende a cero cuando $\varepsilon - \varepsilon'$ tiende a cero.

2. Por tanto los polinomios característicos de ambas matrices serán muy parecidos si ambas matrices están “cerca”, dado que las matrices $\mathbb{A}_\varepsilon - \lambda \mathbb{I}$ y $\mathbb{A}_{\varepsilon'} - \lambda \mathbb{I}$ también son muy parecidas en cada entrada, y el determinante es una función continua (consiste en efectuar sumas y productos con los elementos de las matrices, y la suma y el producto son funciones continuas).
3. Por tanto si ambos polinomios característicos son muy parecidos, sus raíces también lo serán (intuitivamente si dos gráficas casi se superponen, los puntos donde cortan el eje x deberían casi coincidir).
4. En suma, **la función** $\lambda(\varepsilon)$ que a cada ε le asigna las raíces del polinomio $\chi_{\mathbb{A}_\varepsilon}(\lambda)$ **es una función continua**, es una curva continua que asocia a cada $\varepsilon \in [0, 1]$ un punto λ en el plano.
5. También notemos que para cada matriz \mathbb{A}_ε , sus discos $C_i(\varepsilon)$ tienen todos el mismo centro (la diagonal principal es siempre la misma) y radios $r_i(\varepsilon)$ que crecen con ε , desde $r_i(0) = 0$ hasta $r_i(1) = r_i$.

6. Por tanto cada disco $C_i(\varepsilon)$ contiene a los discos $C_i(\varepsilon')$ para todo $0 \leq \varepsilon' \leq \varepsilon \leq 1$, o en palabras, las uniones de los discos crecen con ε .

Para $\varepsilon = 0$ cada disco consiste solamente en su centro, y luego van creciendo concéntricamente hasta que cuando $\varepsilon = 1$ alcanza su máximo tamaño, el del disco correspondiente de la matriz original \mathbb{A} .

7. **De lo anterior se deduce que la curva continua que para cada valor propio λ_j de \mathbb{A} empieza en $\lambda_j(0) = a_{jj}$ y termina en $\lambda_j(1) = \lambda_j$ está toda ella contenida en el disco $C_j(1) = C_j$, la bola cerrada de centro a_{jj} y radio $r_j(1) = r_j$.**

3 Aplicación de estas ideas

La idea central es que si $\lambda_j(0) = a_{jj}$ entonces como todos los $\lambda_j(\varepsilon) \in C_j(\varepsilon)$ y todos estos discos tienen centro a_{jj} y crecen con ε , el punto final de la curva continua $\lambda_j(\varepsilon) = \lambda_j(1) = \lambda_j$ (valor propio de \mathbb{A}) tiene que estar en el último de los discos (disco de la matriz \mathbb{A}), $C_j(1) = C_j$.

Por tanto si λ_j es un valor propio de \mathbb{A} ($\lambda_j = \lambda_j(1)$), punto final de una curva continua que empieza en $\lambda_j(0) = a_{jj}$ contenido en la unión \mathbb{M} de m discos de \mathbb{A} que son disjuntos de los restantes $n-m$ discos), entonces λ_j también tiene que estar en esa unión de m discos, ya que si la curva $\lambda_j(\varepsilon)$ comienza en la unión \mathbb{M} de esos m discos y es continua, no puede terminar en los discos de \mathbb{A} que son disjuntos a dicha unión.

Vamos a ilustrar con algunos ejemplos.

4 Ejemplos

Tomaremos una matriz 3×3 y una sucesión de unos pocos $\varepsilon \in [0, 1]$, para c/u de ellos construiremos la matriz \mathbb{A}_ε , hallaremos sus valores propios $\lambda_j(\varepsilon)$ y los correspondientes discos $C_j(\varepsilon)$.

En cada ejemplo tendremos \mathbb{A} , sus valores propios λ_j y la secuencia de valores propios $\lambda_j(\varepsilon)$ que comienza en $\lambda_j(0) = a_{jj}$ y termina en $\lambda_j(1) = \lambda_j$.

Graficaremos la secuencia de discos obtenidos, el valor propio correspondiente (puntos de colores) y el valor propio final (triángulo de colores).

Recordemos que cuando la matriz tiene todas sus entradas reales (parte imaginaria cero), el polinomio característico será un polinomio con coeficientes reales, y de tener una raíz compleja también tendrá su conjugado como raíz.

Ejemplo 1) $\mathbb{A} = \begin{pmatrix} 1 & -3 & 2 \\ -2 & -4 & 3 \\ 3 & -1 & 2 \end{pmatrix}$

Los valores propios de \mathbb{A} son $\lambda_3 = -5.224$, $\lambda_2 = 3.294$, $\lambda_1 = 0.930$. Tomaremos ε entre 0 y 1 con saltos de 0.1.

La figura (1) muestra la familia de discos centrados en $a_{11} = (1, 0)$, $a_{22} = (-4, 0)$, $a_{33} = (2, 0)$ y con radios que varían desde 0 hasta 5, 5 y 4 respectivamente.

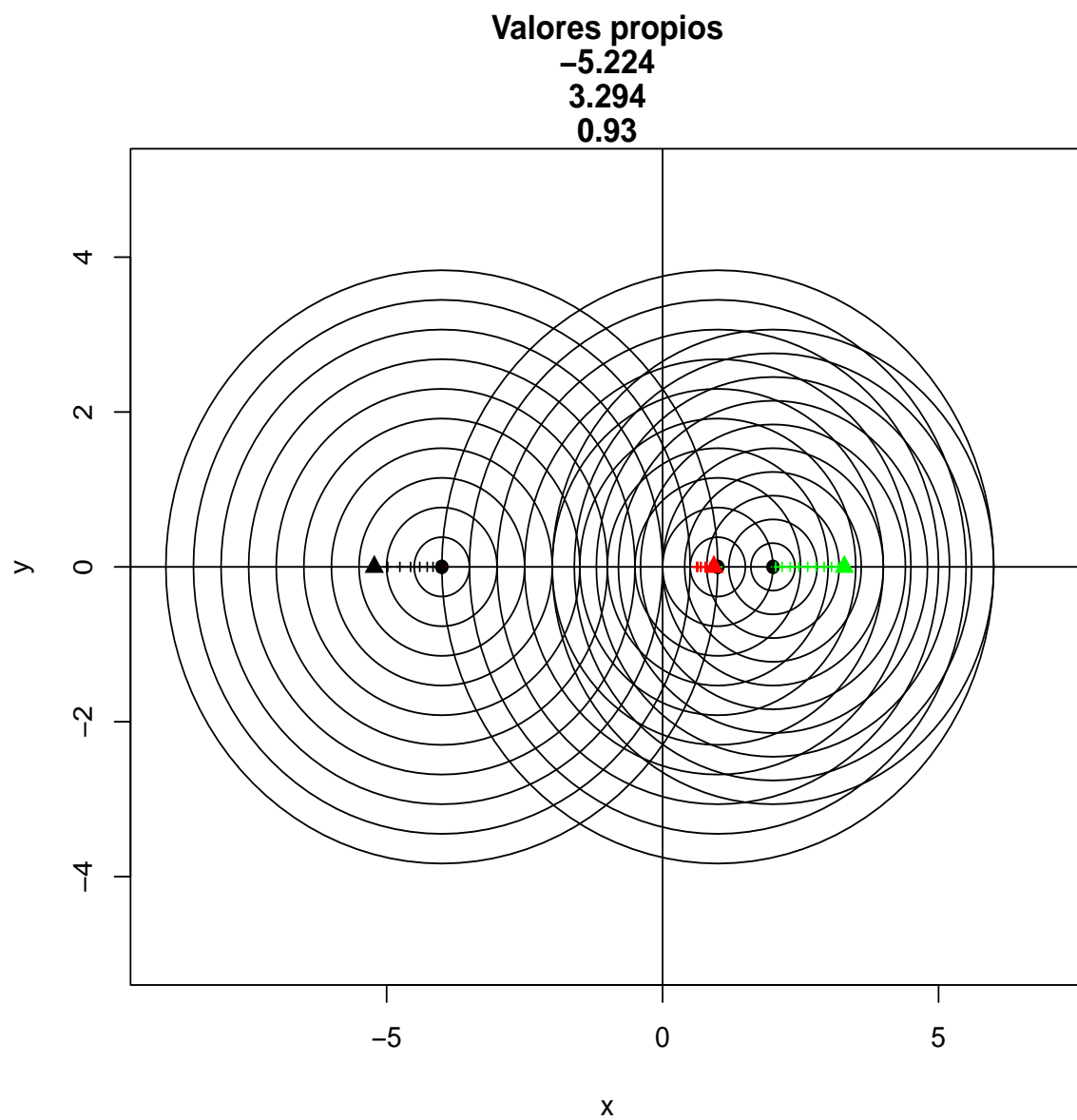


Fig. 1: Ejemplo 1

Vemos p.ej. que la familia de discos de centro $(-4,0)$ tiene una sucesión de valores propios que comienza en -4 y va moviéndose hacia la izquierda (puntos negros) hasta terminar en el valor propio final -5.224 (triángulo negro).

De modo similar los puntos verdes ilustran los valores propios que comienzan en 2 y terminan en 3.294.

Los puntos rojos son la sucesión de valores propios que comienzan en 1 y terminan en 0.930. Notemos que en este caso y a diferencia de las anteriores, la sucesión no es monótona: empieza moviéndose hacia la izquierda para luego retroceder.

La tabla (1) muestra la sucesión de valores propios.

ε	$\lambda_3(\varepsilon)$	$\lambda_2(\varepsilon)$	$\lambda_1(\varepsilon)$
0	-4.00	2.00	1.00
0.1	-4.01	2.05	0.96
0.2	-4.03	2.17	0.87
0.3	-4.08	2.31	0.77
0.4	-4.16	2.47	0.69
0.5	-4.26	2.63	0.64
0.6	-4.40	2.78	0.62
0.7	-4.57	2.93	0.64
0.8	-4.76	3.06	0.70
0.9	-4.98	3.19	0.80
1	-5.22	3.29	0.93

Tab. 1: Sucesión de valores propios del Ejemplo 1

$$\text{Ejemplo 2) } \mathbb{A} = \begin{pmatrix} 2 & -4 & 2 \\ 3 & 7 & -2 \\ 1 & -5 & 3 \end{pmatrix}$$

Los valores propios son $\lambda_1 = 0.655$, $\lambda_2 = 5.672 + 1.179i$, $\lambda_3 = 5.672 - 1.179i$ (uno real y dos complejos conjugados).

La tabla (2) muestra la sucesión de valores propios, mientras que la figura (2) muestra la sucesión de discos y los valores propios en cada disco. La

ε	$\lambda_1(\varepsilon)$	$\lambda_2(\varepsilon)$	$\lambda_3(\varepsilon)$
0	2+0.000i	7+0.000i	3+0.000i
0.1	2+0.000i	7.000+0.000i	3.000+0.000i
0.2	1.981+0.000i	6.995+0.000i	3.024+0.000i
0.3	1.924+0.000i	6.979+0.000i	3.097+0.000i
0.4	1.822+0.000i	6.944+0.000i	3.234+0.000i
0.5	1.680+0.000i	6.881+0.000i	3.439+0.000i
0.6	1.508+0.000i	6.776+0.000i	3.716+0.000i
0.7	1.315+0.000i	6.604+0.000i	4.081+0.000i
0.8	1.106+0.000i	6.294+0.000i	4.600+0.000i
0.9	0.885+0.000i	5.557+0.526i	5.557-0.526i
1	0.655+0.000i	5.672+1.179i	5.672-1.179i

Tab. 2: Sucesión de valores propios del Ejemplo 2

sucesión de valores propios que parte de (2,0) (rojo) se mueve hacia la izquierda, siempre con parte imaginaria cero, hasta el valor propio $\lambda_1 = 0.655$.

La sucesión de valores propios que parte de (3,0) (verde) se mueve hacia la derecha y hacia abajo (sobre el final aparecen partes imaginarias negativas), hasta el valor propio $\lambda_3 = 5.672 - 1.179i$.

La sucesión de valores propios que parte de (7,0) (negro) se mueve hacia la izquierda y hacia arriba, hasta terminar en el valor propio $\lambda_2 = 5.672 + 1.179i$.

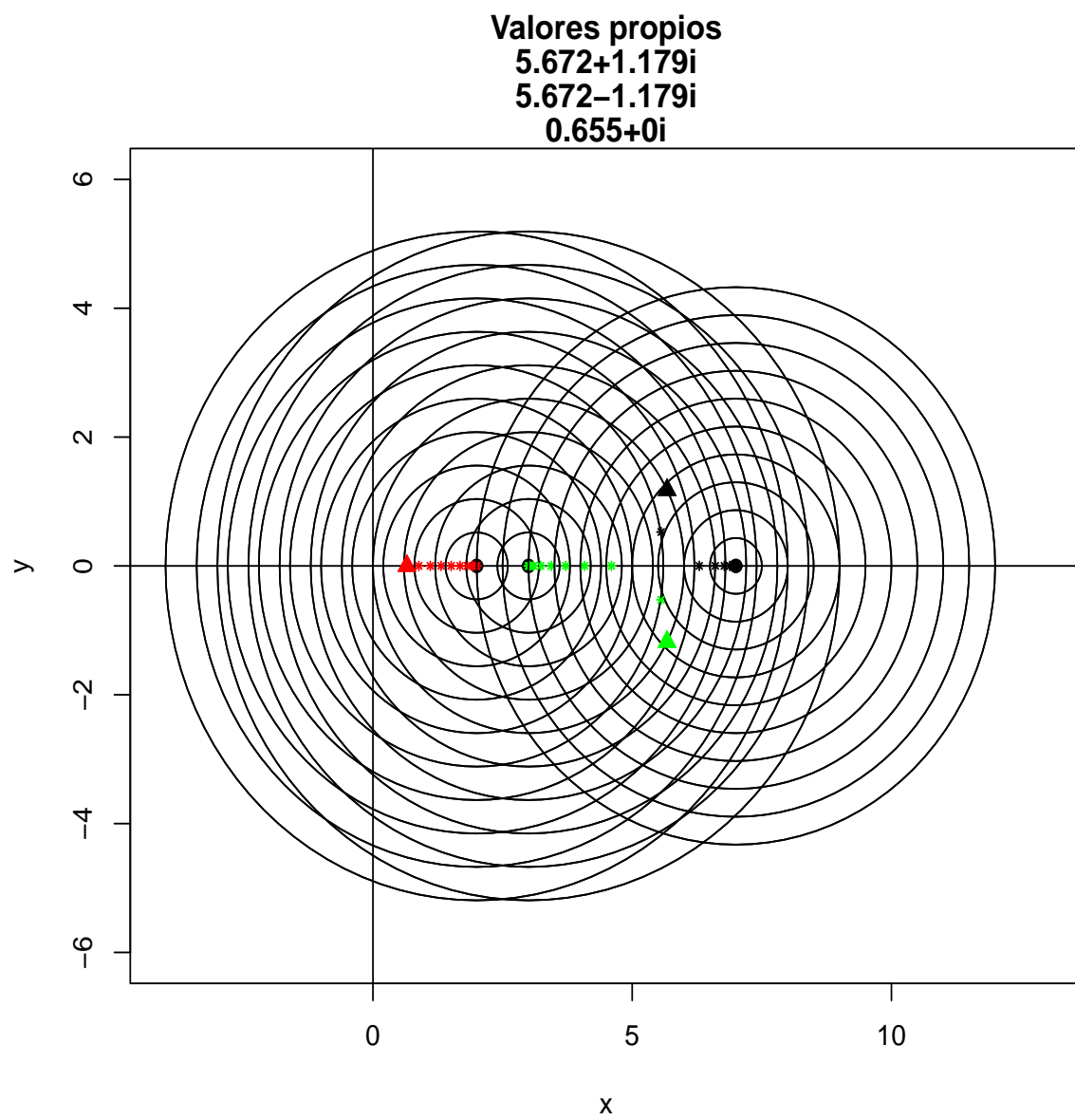


Fig. 2: Ejemplo 2

Ejemplo 3) $\mathbb{A} = \begin{pmatrix} -5 & 1 & -2 \\ -1 & 1 & 1 \\ 2 & -2 & 8 \end{pmatrix}$

Los valores propios son $\lambda_1 = 1.199, \lambda_2 = -4.543, \lambda_3 = 7.344$.

La sucesión de valores propios se muestra en la tabla (3), los discos y la sucesión en la figura (3).

ε	$\lambda_1(\varepsilon)$	$\lambda_2(\varepsilon)$	$\lambda_3(\varepsilon)$
0	1	-5	8
0.1	1.001	-4.995	7.994
0.2	1.005	-4.981	7.976
0.3	1.012	-4.958	7.946
0.4	1.023	-4.925	7.903
0.5	1.037	-4.884	7.847
0.6	1.056	-4.833	7.777
0.7	1.081	-4.774	7.693
0.8	1.112	-4.706	7.593
0.9	1.151	-4.629	7.477
1	1.199	-4.543	7.344

Tab. 3: Sucesión de valores propios del Ejemplo 3

Observaciones:

- (a) Cuando los n discos son disjuntos entre sí (como en el ejemplo 3), podemos aplicar la parte 2) con $m=1$.

En efecto, para $m=1$ cada disco es una unión de un solo elemento (el propio disco), disjunta de los $n-1$ discos restantes, por tanto en cada disco hay exactamente un valor propio.

- (b) Cuando la matriz es de números reales (como en el ejemplo 3), el polinomio característico tendrá coeficientes reales y si admitiera una raíz compleja $z = a + bi$ tendría que admitir también a su conjugada $\bar{z} = a - bi$. Ambos puntos tienen la misma coordenada x y coordenadas y opuestas, son simétricos respecto del eje ox .

Por tanto si los discos están centrados en el eje ox , tanto z como \bar{z} deberían estar **en el mismo disco**, lo que no puede pasar si los n discos son disjuntos (un valor propio en cada disco).

Por tanto **si la matriz es de entradas reales y tiene sus n discos disjuntos, entonces los n valores propios deben ser reales** y distintos, y por tanto la matriz es diagonalizable.

La figura (3) muestra la sucesión de discos comenzando con los de la matriz \mathbb{D} y terminando con los de la matriz \mathbb{A} .

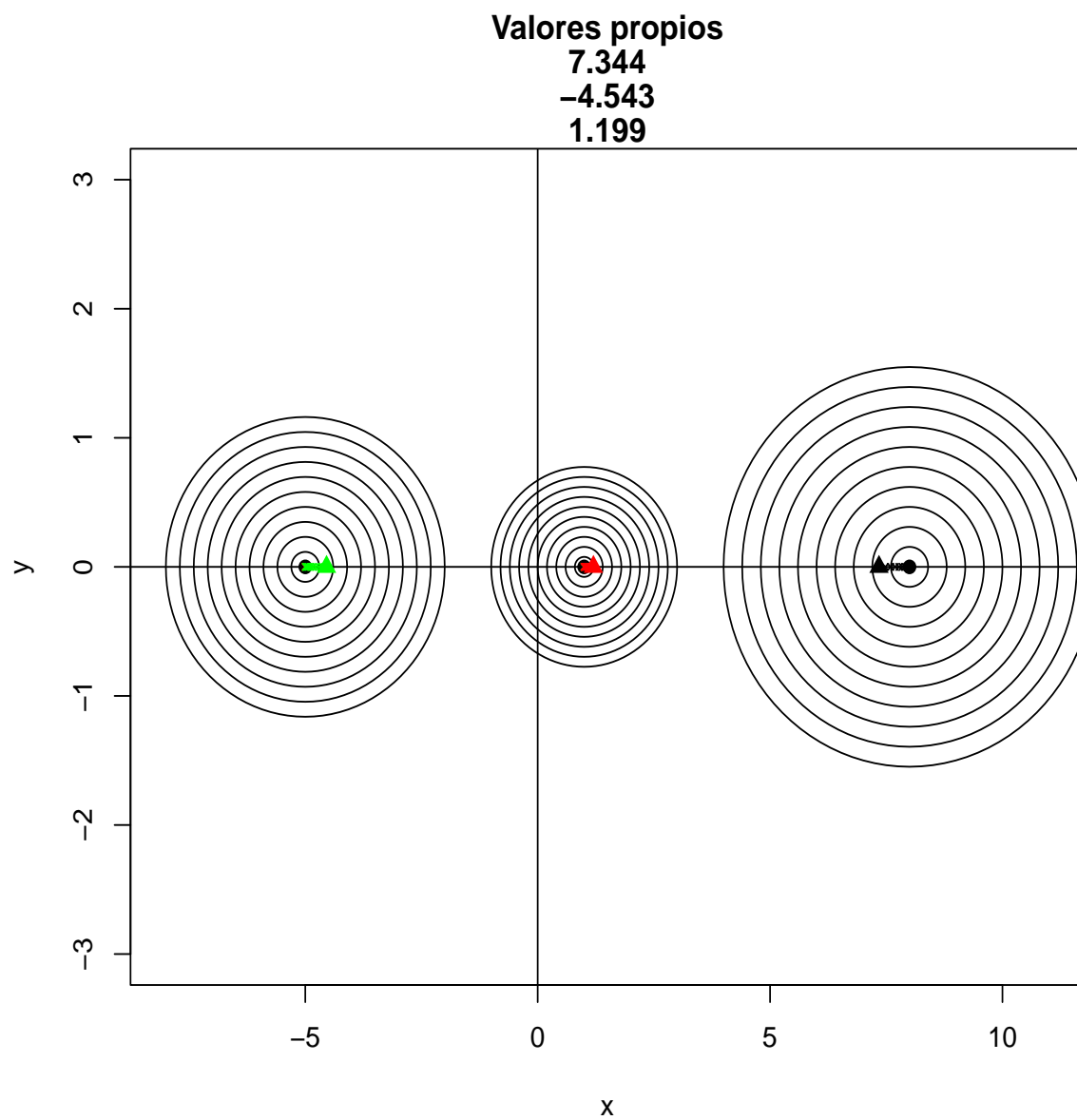


Fig. 3: Ejemplo 3

Ejemplo 4) $A = \begin{pmatrix} 7 & 9 \\ 5 & 5 \end{pmatrix}$

Los valores propios son 12.782 y -0.782.

La sucesión de valores propios se muestra en la tabla (4), los discos y la

sucesión de valores propios en la figura (4).

ε	$\lambda_1(\varepsilon)$	$\lambda_2(\varepsilon)$
0	7	5
0.1	7.204	4.796
0.2	7.673	4.327
0.3	8.247	3.753
0.4	8.864	3.136
0.5	9.500	2.500
0.6	10.147	1.853
0.7	10.801	1.199
0.8	11.459	0.541
0.9	12.120	-0.120
1	12.782	-0.782

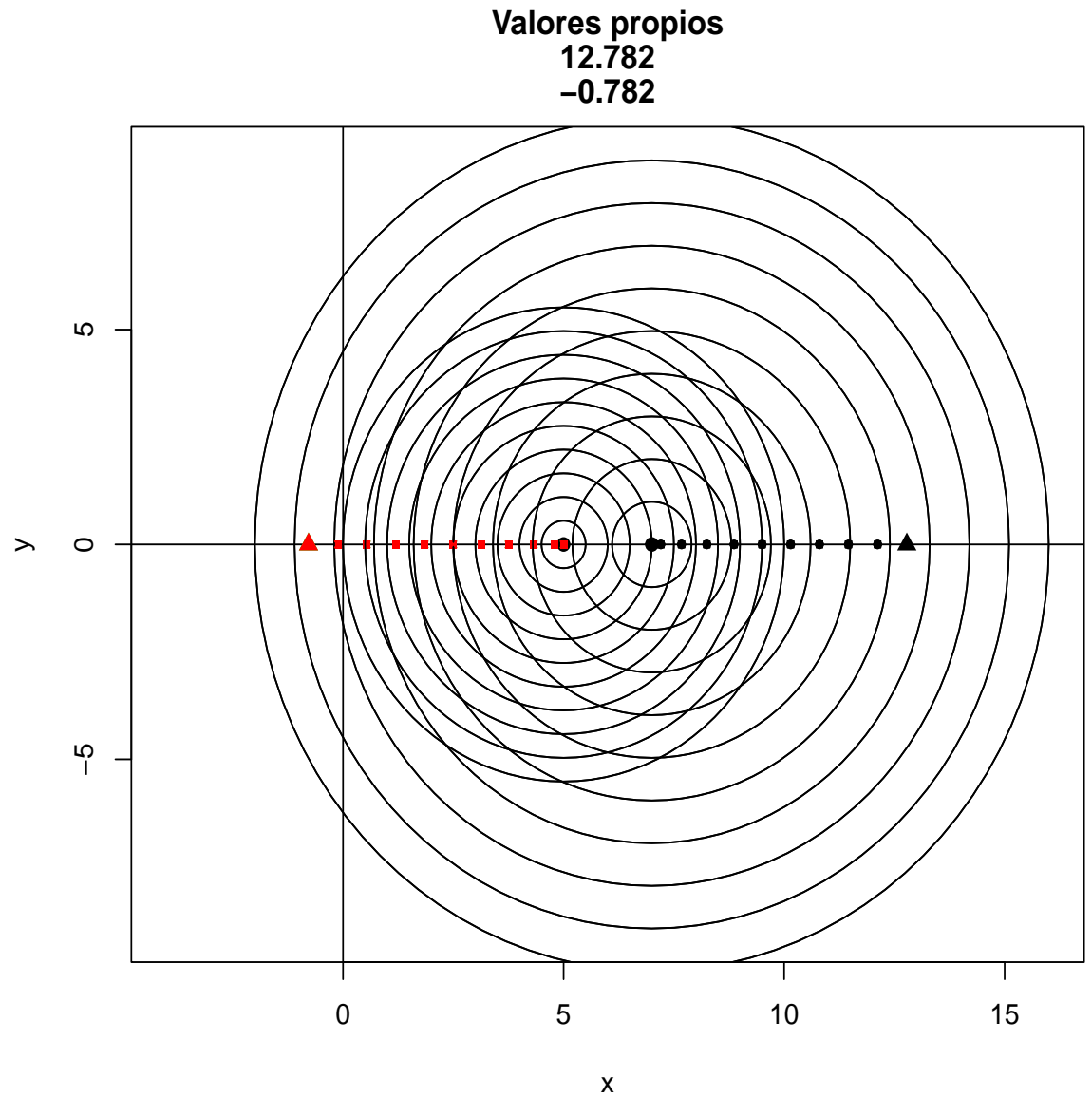
Tab. 4: Sucesión de valores propios del Ejemplo 4

Observación: Si miramos el ejemplo 4, vemos que los dos discos de \mathbb{A} , $C_1 = B[7, 9]$ y $C_2 = B[5, 5]$ se intersectan, y la parte 2 del teorema nos dice que en dicha unión hay 2 valores propios, *pero no nos dice que tengan que estar uno en cada disco*.

En efecto, ambos valores propios terminan estando en la unión, pero ninguno en el disco de centro $(5, 0)$ y radio 5.

Si repasamos la demostración de la parte 2, la continuidad de la curva $\lambda_2(\varepsilon)$ nos permite afirmar que si dicha curva comienza en una unión de 2 discos ($\lambda_2(0) = a_{22} = (5, 0)$), terminará en la misma unión (o sea $\lambda_2(1) = \lambda_2 = -0.782$ estará en la misma unión de esos 2 discos), pero no necesariamente en el mismo disco de centro $a_{22} = (5, 0)$ y radio 5.

Si miramos la figura (4) vemos que $\lambda_2 = -0.782$ (triángulo rojo) está fuera del disco $C_2(1)$ de \mathbb{A} de centro $(5, 0)$ y radio 5, pero dentro de la unión de los discos C_1 y C_2 de \mathbb{A} .



..

Fig. 4: Ejemplo 4: Observar que el disco de centro (5,0) y radio 5 no tiene valor propio dentro (triángulo rojo)

5 La matriz traspuesta

Recordemos que si M es una matriz cuadrada, $\det(M) = \det(M^t)$ y por tanto $\det(A - \lambda I) = \det((A - \lambda I)^t) = \det(A^t - \lambda I^t) = \det(A^t - \lambda I)$.

Entonces \mathbb{A} y \mathbb{A}^t tienen el mismo polinomio característico y por tanto los mismos valores propios.

Llamemos C'_i a los discos de Gershgorin de la matriz \mathbb{A}^t , los mismos tienen los mismos centros que los discos C_i de la matriz \mathbb{A} (ambas matrices tienen la misma diagonal principal), pero los radios r'_i ahora son la suma de los módulos de las filas de \mathbb{A}^t (que son las columnas de \mathbb{A}) sin contar al elemento diagonal.

No hace falta transponer la matriz \mathbb{A} para hallar los discos de \mathbb{A}^t , alcanza con usar las columnas para calcular los radios.

En suma, los valores propios también deben encontrarse en la unión de los discos con centro en los elementos diagonales y radios calculados usando las columnas de \mathbb{A} .

Los valores propios deben hallarse tanto en $U = \bigcup_i C_i$ como en $U' = \bigcup_i C'_i$, (dado que ambas matrices tienen los mismos valores propios), por tanto deben pertenecer a la intersección de ambas uniones.

Como $U \cap U' \subseteq U \cup U'$, estaremos reduciendo la región en la que se encuentran los valores propios.

Ilustraremos este último punto con un ejemplo:

Tomemos $\mathbb{A} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 3 & 3/2 \\ 1 & 1 & 0 \end{pmatrix}$, vamos a hallar los discos de Gershgorin

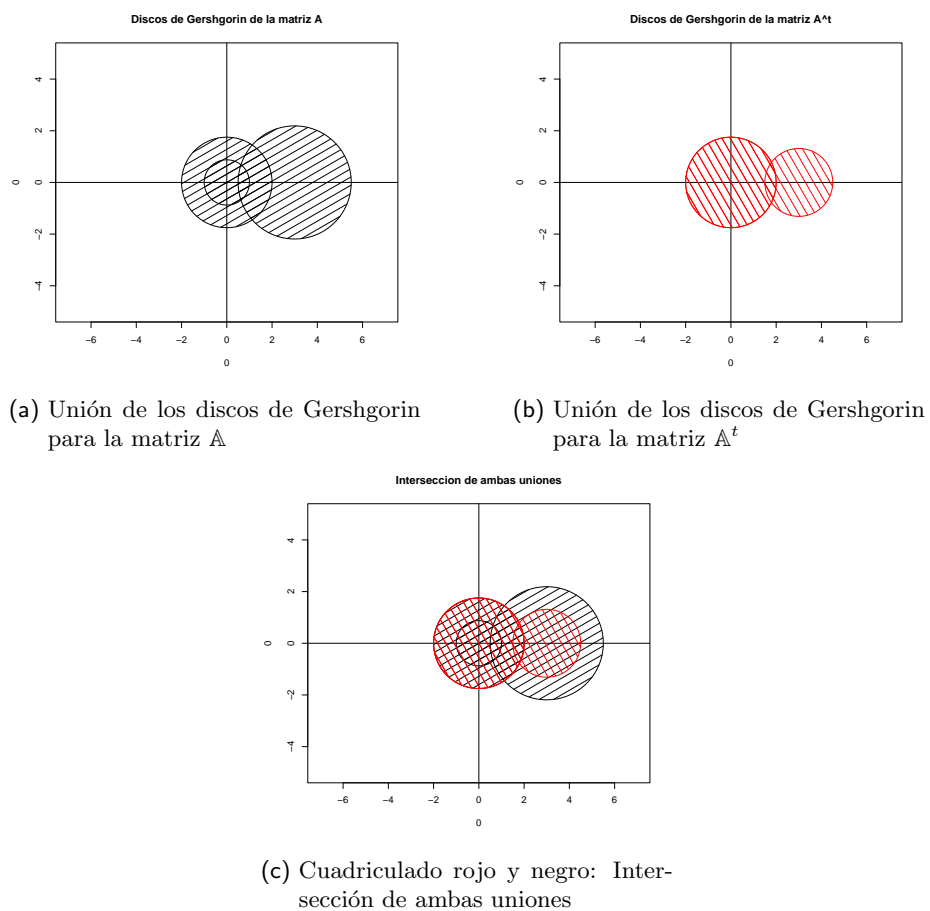
de \mathbb{A} y \mathbb{A}^t y a ver $U = \bigcup_{i=1}^3 C_i$, $U' = \bigcup_{i=1}^3 C'_i$ y $U \cap U'$.

Tenemos que $C_1 = B[0, 1]$ (la bola cerrada de centro $(0,0)$ y radio 1), $C_2 = B[3, 5/2]$, $C_3 = B[0, 2]$ y que $C'_1 = B[0, 2]$, $C'_2 = B[3, 3/2]$, $C'_3 = B[0, 2]$.

Las figuras (5a) y (5b) muestran las uniones de los discos de Gershgorin de \mathbb{A} y \mathbb{A}^t respectivamente, mientras que la figura (5c) muestra la intersección de ambas uniones.

Vemos que al intersectar ambas uniones obtenemos una región del plano menor que la delimitada por cada una de las uniones.

En particular podemos prescindir de parte del anillo originado por las dos bolas centradas en $(3,0)$ (la parte que solamente quedó rayada en un solo color).

Fig. 5: Intersección de las uniones de discos para A y A^t