

Perturbación espectral de matrices

Juan-Miguel Gracia

17 de abril de 2023

Índice general

Índice general	III
Prólogo	VII
Notaciones	IX
1. Matrices con igual rango	1
1.1. Invariantes de una relación de equivalencia	1
1.2. Equivalencia de matrices rectangulares	2
1.3. Perturbación de la forma canónica	3
1.4. Rango de matrices genéricas	8
1.5. Distancia a las matrices de rango menor	11
1.6. Reducción suave de funciones matriciales	12
1.7. Ejercicios	14
1.8. Notas al Capítulo 1	15
2. Particiones de enteros	17
2.1. Particiones	17
2.2. Mayoración	19
2.3. Dualidad y monotonía	21
2.4. Transformaciones elementales de particiones	27
2.5. Orden lexicográfico	35
2.6. Una identidad sobre particiones	36
2.7. Ejercicios	39
2.8. Notas al Capítulo 2	45
3. Invariantes de semejanza y rangos	47
3.1. Desigualdad de Frobenius	47
3.2. Característica de Segré	49
3.3. Nulidades de potencias	50
3.4. Característica de Weyr	51
3.4.1. Definición alternativa de la característica de Weyr	52
3.5. Relación entre las características de Segré y Weyr	53
3.6. Forma canónica dual	57
3.7. Rangos y particiones de Weyr	62
3.8. Criterio de semejanza mediante rangos	65
3.9. Matrices no derogatorias	69
3.10. Nulidades de un producto	70
3.11. Nulidades de matrices cuasidiagonales por bloques	70
3.12. Rango mínimo de factores invariantes	71
3.13. Ejercicios	73
3.14. Notas al Capítulo 3	74

4. Raíces comunes de polinomios	77
4.1. Resultante de polinomios	77
4.2. Multiplicidades de las raíces	79
4.3. Ejercicios	83
5. Continuidad y aproximación	85
5.1. Continuidad del espectro	85
5.1.1. El espacio métrico de n -tuplas desordenadas	88
5.2. Matrices genéricas	93
5.3. Criterio local de semejanza	94
5.4. Conjetura de Sylvester aproximada	95
5.5. Ejercicios	95
5.6. Notas al Capítulo 5	96
6. Perturbación de la forma de Jordan	97
6.1. Condiciones necesarias	97
6.2. El problema inverso subyacente	101
6.3. La clausura de la órbita de semejanza	112
6.4. Ejercicios	114
6.5. Notas al Capítulo 6	117
7. Forma de Jordan genérica	119
7.1. Introducción	119
7.2. Teorema de Frobenius-König	121
7.3. Menores de la matriz característica y digrafos	121
7.4. Ejercicios	127
7.5. Notas al Capítulo 7	128
8. Entornos de seguridad para la forma de Jordan	129
9. Subespacios invariantes estables	131
10. Jordanización de una función matricial	133
10.1. Puntos de bifurcación	133
10.2. Puntos de bifurcación de la característica de Segré	140
10.3. Funciones matriciales genéricas	141
11. Forma de Jordan probabilística	143
11.1. Referencias comentadas	143
A. Transformaciones elementales de matrices	145
A.1. Matrices sobre un cuerpo	145
A.2. Matrices polinómicas	146
A.3. Forma canónica de Smith	147
A.4. Divisores elementales	148
A.5. Semejanza de matrices	151
A.6. Forma canónica de Jordan	153
A.7. Multiplicación de matrices de bloques	155
A.8. Transformaciones elementales por bloques	156
B. Digrafos y permutaciones	159
B.1. Digrafos	159
B.2. Permutaciones	161
B.3. Digrafos y determinantes	168
B.4. Ejercicios	171

C. Topología	173
C.1. Adherencia, interior, exterior y frontera	173
C.2. Conjuntos diseminados	174
C.3. Estructura de los conjuntos abiertos de \mathbb{R}	175
C.4. Notas al Apéndice C	176
D. Funciones de variable compleja	177
D.1. Representación integral de los ceros	177
D.2. Integrales paramétricas	178
Bibliografía	179
Índice alfabético	183

Prólogo

Este libro es una Monografía sobre perturbación espectral de matrices complejas cuadradas; es decir, analiza la perturbación que experimentan los objetos espectrales (valores propios, forma canónica de Jordan, subespacios invariantes) de una matriz cuadrada cuando los elementos de esta matriz sufren ligeros cambios. También pretende ser un libro de texto que sirva de complemento a los profesores y estudiantes de Álgebra Lineal, Análisis Matricial y Teoría de Matrices de las Escuelas de Ingeniería y de las Facultades de Ciencias.

No sabríamos enfatizar lo suficiente que el alma de la Teoría de Matrices está en las transformaciones elementales en las líneas de una matriz que conservan alguna propiedad de la matriz. Una matriz no es sólo la representación de una aplicación lineal entre espacios vectoriales; para un número creciente de disidentes matemáticos una matriz es un objeto mucho más rico, lleno de propiedades combinatorias que, por regla general, sólo se perciben a ras de tierra. De este modo, al lector que conozca la forma canónica de Jordan de una matriz A mediante construcciones geométricas (subespacios propios, cadenas de Jordan, etc.), le avisamos de que es conveniente que estudie una deducción de esta forma a través de transformaciones elementales en la matriz característica $\lambda I - A$. Para ello, se hallan los factores invariantes de $\lambda I - A$ pasando por transformaciones elementales de esta matriz a su forma canónica de Smith. Después por factorización de estos polinomios en polinomios irreducibles se hallan los divisores elementales.

La idea esencial de la metodología empleada en este libro consiste en expresar los invariantes enteros (por ejemplo, los órdenes de los bloques de Jordan) de la relación de semejanza de matrices mediante los rangos de determinadas matrices. Por eso, el enfoque cambiará de la consideración de la característica de Segré de una matriz (i.e. los órdenes de sus bloques de Jordan) a analizar la característica de Weyr (su dual), que es expresable directamente por medio de rangos de matrices.

El rango de una matriz A aumenta o se mantiene al perturbar ligeramente sus elementos. Además, todo número entero permisible mayor o igual que dicho rango es alcanzado como rango de matrices que están tan próximas a A como se quiera. De estos dos hechos simples se deducen los teoremas principales de esta monografía. Estas son propiedades cualitativas que sólo dependen de la topología del espacio de las matrices. Si miramos a este espacio como espacio métrico, dotado de la distancia inducida por la norma espectral o la norma de Frobenius, resulta que la distancia de la matriz A a la matriz más próxima cuyo rango sea inferior al de A es igual al menor valor singular positivo de A . De aquí se deducen precisiones cuantitativas de los resultados cualitativos antes mencionados.

El Capítulo 1 trata un caso especial, que sirve de introducción a toda la materia del libro; a saber, se estudia cómo varía la forma canónica para la relación de equivalencia entre matrices complejas rectangulares que considera equivalentes dos matrices si tienen el mismo rango.

El libro supone ciertos conocimientos básicos de Teoría de Matrices. Para comodidad del lector nos referiremos a los libros de Lancaster y Tismenetsky: *The Theory of Matrices with Applications*, Second Edition [41], Horn y Johnson: *Matrix Analysis* [37], Golub y Van Loan: *Matrix Computations* [32] y de Gohberg, Lancaster y Rodman: *Invariant Subspaces of Matrices with Applications* [31]. No obstante, los apéndices resumen la información principal que se precisa y el libro puede ser leído por un estudiante de tercer curso ya que requiere

la madurez matemática alcanzada por dos años de contacto con las matemáticas de nivel universitario. Solamente, los capítulos 9 y 10 requieren conocimientos de espacios topológicos y funciones de una variable compleja que van algo más lejos de las definiciones. También requiere saber algunas cosas de topología la Sección 1.6 del Capítulo 1. El libro se ha escrito pensando en el tipo de lector “saltamontes”, que va de aquí para allá buscando lo que le interesa. Por ello, no es preciso haber leído todas las páginas previas para entender el tema que importe. Se ha tratado de facilitar la accesibilidad a cada capítulo, e incluso a cada sección, sin que haya la necesidad de haber estudiado el material precedente del libro. La naturaleza “elemental” de las cuestiones tratadas lo permite; grosso modo sólo se precisa conocer la suma y multiplicación de matrices, las propiedades de los determinantes, las raíces de un polinomio y poco más.

Nuestra experiencia durante años en la consulta de bibliografía de Álgebra Lineal y Teoría de Matrices nos ha permitido observar la referencia casi obligada al libro *Théorie des Matrices* [28],[29] en dos tomos de Felix R. Gantmacher, en todo artículo o libro sobre este tema. Séanos consentido, pues, rendir homenaje a este libro importantísimo, y que lo hagamos extensivo al libro de Richard E. Bellman *Introduction to Matrix Analysis* [10]. Ambos libros tuvieron una influencia decisiva en el autor desde sus años de estudiante.

Notaciones

$[a]$, clase de equivalencia de representante a en un conjunto E dotado de la relación de equivalencia \sim , página 1

$a := b$, a es por definición igual a b , página 1

\mathbb{C} , el cuerpo de los números complejos, página 2

$\mathbb{C}^{p \times q}$, el espacio de las matrices complejas $p \times q$, donde p y q son enteros positivos cualesquiera, página 2

\mathbb{R} , el cuerpo de los números reales, página 2

$M = \begin{pmatrix} 0 & * & 0 \\ 0 & 0 & * \end{pmatrix}$, matriz “genérica”, página 8

$R(f, g)$, matriz resultante de Sylvester de los polinomios $f(\lambda)$ y $g(\lambda)$, página 78

$R(a, b_1, \dots, b_h)$, matriz resultante generalizada de Sylvester de los polinomios $a(\lambda), b_1(\lambda), \dots, b_h(\lambda)$, página 78

$R(f(\lambda), g(\lambda), \lambda)$, es la matriz resultante de Sylvester de $f(\lambda)$ y $g(\lambda)$ cuando se quiera enfatizar la variable independiente λ , página 79

$\mathbb{C}_{\text{sim}}^n$, producto simétrico de \mathbb{C} por sí mismo n veces, página 92

$\mathcal{R}_{\lambda_0}(M)$, subespacio radical de M asociado a λ_0 , página 135

Capítulo 1

Matrices con igual rango

Este capítulo trata de compendiar el libro en su conjunto, mediante la consideración de un caso más sencillo que el estudio de la perturbación de la forma canónica de Jordan. Ésta es una forma canónica para la relación de semejanza de matrices cuadradas. En cambio ahora veremos cómo es afectada la forma canónica para la relación de equivalencia entre matrices rectangulares que considera equivalentes dos matrices si tienen el mismo rango. Usaremos el signo igual para una definición o asignación mediante la notación de Rutishauser. Así pues, $a := b$ se lee “ a es por definición igual a b ”.

1.1. Invariantes de una relación de equivalencia

Recordemos que una relación binaria \sim definida sobre un conjunto cualquiera E se llama *relación de equivalencia* si cumple las propiedades

reflexiva: $a \sim a$ para todo $a \in E$,

simétrica: sean $a, b \in E$, si $a \sim b$, entonces $b \sim a$,

transitiva: sean $a, b, c \in E$, si $a \sim b$ y $b \sim c$, entonces $a \sim c$.

Sea E un conjunto con una relación de equivalencia \sim . Sea $s \in E$. Llamaremos clase de equivalencia de a al subconjunto de los elementos de E equivalente a a , y lo denotamos por $[a]$.

Se llama *invariante* de \sim a toda función f definida sobre E con valores en cualquier conjunto F que sea constante sobre las clases de equivalencia. Con símbolos, un *invariante* de \sim es una función

$$f : E \rightarrow F$$

tal que siempre que $a \sim b$, para $a, b \in E$, se tiene que

$$f(a) = f(b).$$

Un invariante $f : E \rightarrow F$ de la relación de equivalencia \sim se llama *completo* si toma valores distintos en elementos pertenecientes a clases de equivalencia distintas. A saber, el invariante $f : E \rightarrow F$ es completo si

$$a \sim b \iff f(a) = f(b).$$

Sea F_1, \dots, F_n un sistema dado de conjuntos. Se dice que la sucesión finita de funciones (f_1, \dots, f_n) es un *sistema completo de invariantes* de la relación de equivalencia \sim sobre E , si $f_i : E \rightarrow F_i$ es un invariante para cada $i = 1, \dots, n$ y

$$a \sim b \iff f_1(a) = f_1(b), \dots, f_n(a) = f_n(b).$$

Una *forma canónica o normal* de la relación de equivalencia \sim sobre E es un invariante $g : E \rightarrow E$ que verifica las dos condiciones siguientes

- Para todo $a \in E$, $a \sim g(a)$.
- $a, b \in E$, $a \sim b \iff g(a) = g(b)$.

Observación 1.1.1. A veces se llama forma canónica a algo ligeramente diferente. Tal cosa sucede cuando la naturaleza del conjunto E no permite determinar unívocamente el valor de un invariante $g : E \rightarrow E$, pero permite definirlo salvo un número finito de valores. Por ejemplo, la forma canónica de Jordan para la semejanza de matrices cuadradas complejas está determinada salvo permutación de los bloques de Jordan. Por ello, generalizamos así la definición: Sea $\mathcal{P}(E)$ un conjunto de subconjuntos finitos de E formados por elementos equivalentes respecto de \sim . Una *forma canónica* es una aplicación $h : E \rightarrow \mathcal{P}(E)$ que verifica

- Para todo $a \in E$, $h(a) \subset [a]$, donde $[a]$ es la clase de equivalencia de a .
- $a, b \in E$, $a \sim b \iff h(a) = h(b)$.

1.2. Equivalencia de matrices rectangulares

Denotaremos por $\mathbb{C}^{p \times q}$ el espacio de las matrices complejas $p \times q$, donde p y q son dos enteros positivos cualesquiera. Por $O_{p \times q}$ denotaremos la matriz cero $p \times q$; también la denotaremos por O , simplemente, cuando las dimensiones de esta matriz sean claras por el contexto. Por \mathbb{C} (resp. \mathbb{R}) denotaremos el cuerpo de los números complejos (resp. de los números reales). Sean $A, B \in \mathbb{C}^{m \times n}$ dos matrices; decimos que A es *equivalente* a B si existen matrices invertibles $P \in \mathbb{C}^{m \times m}$, $Q \in \mathbb{C}^{n \times n}$ tales que

$$B = PAQ. \quad (1.1)$$

En cuyo caso, denotamos esta relación entre A y B por $A \sim B$. Es fácil comprobar que \sim es una relación de equivalencia sobre $\mathbb{C}^{p \times q}$. Es bien conocido que realizando transformaciones elementales en las filas y/o columnas de una matriz $A \in \mathbb{C}^{m \times n}$ se puede llegar a una matriz de una de las formas siguientes, dependiendo de los valores relativos de r, m y n ,

$$\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}, \quad [I_m, O], \quad \begin{bmatrix} I_n \\ O \end{bmatrix}, \quad I_n, \quad (1.2)$$

donde r es el rango de A e I_r es la matriz unidad o identidad de orden r . Véase la Sección A.1 del Apéndice A o la Sección 2.7 de [41]. Se llama matriz elemental por filas (resp. columnas) de orden m (resp. n) a toda matriz que resulta de efectuar una transformación elemental por filas (resp. columnas) sobre la matriz I_m (resp. I_n). Se llama matriz elemental a toda matriz elemental por filas o por columnas. Toda matriz elemental es invertible. La matriz que resulta de efectuar sobre una matriz M una transformación elemental por filas (resp. columnas) es igual al producto que resulta de premultiplicar (resp. postmultiplicar) M por una matriz elemental por filas (resp. columnas). Dada la matriz $A \in \mathbb{C}^{m \times n}$, por lo dicho antes de la fórmula (1.2) se sigue que existen matrices invertibles $P \in \mathbb{C}^{m \times m}$ y $Q \in \mathbb{C}^{n \times n}$ tales que

$$PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}. \quad (1.3)$$

Sobrentenderemos que el símbolo

$$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$$

denota de ahora en adelante, en este capítulo, la matriz

$$\begin{cases} [I_m, O] & \text{si } r = m \leq n, \\ \begin{bmatrix} I_n \\ O \end{bmatrix} & \text{si } r = n \leq m, \\ I_n & \text{si } r = m = n. \end{cases}$$

El rango de una matriz no varía al pre- o postmultiplicar ésta por una matriz invertible. De aquí que el rango es un invariante para la relación de equivalencia \sim sobre $\mathbb{C}^{m \times n}$. Denotaremos este invariante por $\text{rg} : \mathbb{C}^{m \times n} \rightarrow \mathbb{N}$, donde \mathbb{N} designa el conjunto de los números naturales. Asimismo, teniendo en cuenta (1.3), se sigue que rg es un invariante completo de \sim . Este resultado también puede parafarsearse en términos de una forma canónica. Definimos la aplicación de $\mathbb{C}^{m \times n}$ en sí mismo dada por

$$A \mapsto \begin{pmatrix} I_r & O \\ O & O \end{pmatrix},$$

para cada $A \in \mathbb{C}^{m \times n}$ y donde r es el rango de A . Es claro por (1.3) que esta aplicación es una *forma canónica* para \sim . En general, no hablaremos de manera tan formal y diremos que la matriz

$$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$$

es la “forma canónica” de la matriz A para la equivalencia \sim .

1.3. Perturbación de la forma canónica

En esta sección consideraremos la norma

$$\|X\|_1 := \sum_{i,j} |x_{ij}| \tag{1.1}$$

definida para toda $X = (x_{ij})$ del espacio de matrices $\mathbb{C}^{m \times n}$. En \mathbb{C}^n usaremos la norma

$$\|\mathbf{x}\|_1 := \sum_{i=1}^n |x_i|,$$

para todo $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{C}^n$.

El determinante de una matriz $Y = (y_{ij}) \in \mathbb{C}^{n \times n}$, $\det Y$ ó $|Y|$, se define mediante la suma

$$\sum_{\sigma \in S_n} \varepsilon(\sigma) y_{\sigma(1),1} \cdots y_{\sigma(n),n},$$

siendo S_n el conjunto de las permutaciones de $\{1, \dots, n\}$ y $\varepsilon(\sigma)$ la signatura de la permutación σ . Por lo tanto, el determinante de Y viene expresado como un polinomio en las n^2 variables y_{ij} ; de donde resulta que la función $\det : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}$ es continua. Por el “principio de la inercia” de las funciones continuas resulta que si $\det A \neq 0$ para una matriz A , entonces existe un entorno de A , \mathcal{V} , en $\mathbb{C}^{n \times n}$ tal que para toda $A' \in \mathcal{V}$ se tiene que $\det A' \neq 0$.

Recordemos que una función real f definida sobre un espacio topológico E es semicontinua inferiormente en un punto $a \in E$ si para todo número real $\varepsilon > 0$ existe un entorno \mathcal{V} de a tal que para cada $x \in \mathcal{V}$ se tiene que $f(a) - \varepsilon < f(x)$. Se dice que la función f es semicontinua inferiormente en E si es semicontinua inferiormente en cada punto de E .

La submatriz $p \times q$ de una matriz $A \in \mathbb{C}^{m \times n}$ que corresponde a las filas $i_1 < i_2 < \dots < i_p$ y a las columnas $j_1 < j_2 < \dots < j_q$ de A se denota por

$$A[\alpha|\beta],$$

llamando $\alpha := (i_1, i_2, \dots, i_p)$, $\beta := (j_1, j_2, \dots, j_q)$. Dados los enteros k y n , con $k \leq n$, es costumbre denotar por $Q_{k,n}$ el conjunto de sucesiones crecientes (i_1, \dots, i_k) tales que $1 \leq i_1 < \dots < i_k \leq n$. Con esta notación, $\alpha \in Q_{p,m}$ y $\beta \in Q_{q,n}$.

Proposición 1.3.1. *La función $\text{rg} : \mathbb{C}^{m \times n} \rightarrow \mathbb{R}$ es semicontinua inferiormente en $\mathbb{C}^{m \times n}$.*

DEMOSTRACIÓN. Como la función rango rg sólo toma valores enteros, basta ver que si una matriz $A = (a_{ij}) \in \mathbb{C}^{m \times n}$ tiene rango r , entonces existe un entorno \mathcal{V} de A tal que para cada $A' \in \mathcal{V}$, $\text{rg } A' \geq r$. Ahora bien, si una matriz A tiene rango r , entonces existe en ella un menor de orden r distinto de cero y todos sus menores de orden $r + 1$ son iguales a cero (un menor de A es igual al determinante de una submatriz cuadrada de A). Supongamos que un menor de orden r de A no nulo es el que corresponde a las filas $i_1 < \dots < i_r$ y a las columnas $j_1 < \dots < j_r$ de A ; de acuerdo con la notación previa designamos por

$$\det A[(i_1, \dots, i_r)|(j_1, \dots, j_r)]$$

este menor, que se puede expresar como un polinomio en los elementos $a_{i_k j_\ell}$, donde k y ℓ recorren el conjunto $\{1, \dots, r\}$. Para una matriz cualquiera $Y = (y_{ij}) \in \mathbb{C}^{m \times n}$ consideramos la función

$$f(Y) := \det Y[(i_1, \dots, i_r)|(j_1, \dots, j_r)].$$

Por lo dicho previamente, esta función es continua y $f(A) \neq 0$; por lo tanto existe un entorno, \mathcal{V} , de A , tal que para cada $A' \in \mathcal{V}$ se sigue que $f(A') \neq 0$. De donde, $\text{rg } A' \geq r$. \square

Nos conviene registrar una variación de esta proposición para referencia posterior.

Teorema 1.3.1. *Sea $A \in \mathbb{C}^{m \times n}$ una matriz de rango r , entonces existe un entorno \mathcal{V} de A en $\mathbb{C}^{m \times n}$ tal que para cada $A' \in \mathcal{V}$, la forma canónica de A' , respecto de la equivalencia \sim , es*

$$\begin{pmatrix} I_{r'} & O \\ O & O \end{pmatrix}$$

con $r' \geq r$.

Diremos que este teorema nos indica cómo varía la forma canónica de la matriz A al perturbar ligeramente sus elementos. Es decir nos da la *condición necesaria* que debe verificar la forma canónica de toda matriz suficientemente próxima a A . A continuación, nuestro objetivo es ver si esta condición es *suficiente* y lo que esto quiera decir. Ello es equivalente a precisar los posibles rangos de matrices situadas tan cerca como se quiera de A . Para analizarlo de manera elemental damos los dos lemas siguientes y una proposición; más tarde veremos que el estudio se simplifica de manera extraordinaria apelando a la forma canónica.

Lema 1.3.1. *Todo conjunto abierto no vacío de \mathbb{C}^n contiene n vectores linealmente independientes.*

DEMOSTRACIÓN. Sea \mathcal{V} un conjunto abierto no vacío de \mathbb{C}^n . Demostraremos que \mathcal{V} contiene n vectores linealmente independientes en dos etapas. Primero consideraremos que $\mathbf{0} \in \mathcal{V}$ y luego que $\mathbf{0} \notin \mathcal{V}$.

Caso: $\mathbf{0} \in \mathcal{V}$. Sea $\langle \mathcal{V} \rangle$ el subespacio vectorial de \mathbb{C}^n engendrado por \mathcal{V} , y sea $p = \dim \langle \mathcal{V} \rangle$. Si ocurriera que $p < n$, el conjunto \mathcal{V} contendría p vectores $\mathbf{v}_1, \dots, \mathbf{v}_p$ que formarían una base de $\langle \mathcal{V} \rangle$. Sea $\mathbf{v}_{p+1} \in \mathbb{C}^n$ un vector tal que el conjunto $\{\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{v}_{p+1}\}$ es linealmente independiente. De aquí se sigue que para todo $\lambda \in \mathbb{C}, \lambda \neq 0$, el conjunto

$$\{\mathbf{v}_1, \dots, \mathbf{v}_p, \lambda \mathbf{v}_{p+1}\}$$

es linealmente independiente.

Ahora bien, el conjunto \mathcal{V} es abierto y $\mathbf{0} \in \mathcal{V}$; por lo tanto existe un $r > 0$ tal que la bola abierta de centro $\mathbf{0}$ y radio $r > 0$, $B(\mathbf{0}, r)$, está contenida en \mathcal{V} . Si elegimos un número $\tilde{\lambda} \in \mathbb{C}$ tal que

$$0 < |\tilde{\lambda}| < \frac{r}{\|\mathbf{v}_{p+1}\|_1},$$

entonces

$$\|\tilde{\lambda} \mathbf{v}_{p+1} - \mathbf{0}\|_1 < r;$$

luego $\tilde{\lambda}\mathbf{v}_{p+1} \in \mathcal{V}$; pero el conjunto

$$\{\mathbf{v}_1, \dots, \mathbf{v}_p, \tilde{\lambda}\mathbf{v}_{p+1}\}$$

es linealmente independiente. Lo que contradice a que $\dim \langle \mathcal{V} \rangle = p$. por lo tanto, deberá ser

$$\dim \langle \mathcal{V} \rangle = n.$$

Lo cual equivale a lo que queríamos demostrar.

Caso: $\mathbf{0} \notin \mathcal{V}$. Suponemos ahora que $\mathbf{0} \notin \mathcal{V}$. Sea \mathbf{v}_0 un vector cualquiera de \mathcal{V} . Obviamente, $\mathbf{v}_0 \neq \mathbf{0}$. Como el conjunto \mathcal{V} es abierto, existe un $r > 0$ tal que

$$B(\mathbf{v}_0, r) \subset \mathcal{V}.$$

Si llamamos

$$\mathbf{v}_0 + B(\mathbf{0}, r) := \{\mathbf{v}_0 + \mathbf{x} \mid \mathbf{x} \in B(\mathbf{0}, r)\},$$

se sigue que

$$\mathbf{v}_0 + B(\mathbf{0}, r) = B(\mathbf{v}_0, r).$$

Amplieemos \mathbf{v}_0 hasta obtener una base de \mathbb{C}^n

$$\{\mathbf{w}_1, \dots, \mathbf{w}_{n-1}, \mathbf{v}_0\}.$$

Busquemos escalares no nulos $\tilde{\lambda}_1, \dots, \tilde{\lambda}_{n-1} \in \mathbb{C}$ tales que si definimos $\mathbf{u}_i := \tilde{\lambda}_i \mathbf{w}_i$, se tenga que \mathbf{u}_i pertenezca a $B(\mathbf{0}, r)$ para $i = 1, \dots, n-1$. Es evidente que el conjunto

$$\{\mathbf{u}_1, \dots, \mathbf{u}_{n-1}, \mathbf{v}_0\}$$

es linealmente independiente.

Trasladando los vectores $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ según el vector \mathbf{v}_0 , obtenemos los vectores $\mathbf{u}_1 + \mathbf{v}_0, \dots, \mathbf{u}_{n-1} + \mathbf{v}_0$, que pertenecen a \mathcal{V} . Demostraremos que los n vectores $\mathbf{u}_1 + \mathbf{v}_0, \dots, \mathbf{u}_{n-1} + \mathbf{v}_0, \mathbf{v}_0$ son linealmente independientes, lo que acabará la demostración.

En efecto, sean $\beta_1, \dots, \beta_{n-1}, \beta \in \mathbb{C}$ tales que

$$\beta_1(\mathbf{u}_1 + \mathbf{v}_0) + \dots + \beta_{n-1}(\mathbf{u}_{n-1} + \mathbf{v}_0) + \beta\mathbf{v}_0 = \mathbf{0}.$$

Entonces,

$$\beta_1\mathbf{u}_1 + \dots + \beta_{n-1}\mathbf{u}_{n-1} + (\beta_1 + \dots + \beta_{n-1} + \beta)\mathbf{v}_0 = \mathbf{0};$$

lo que implica que

$$\beta_1 = \dots = \beta_{n-1} = 0, \quad \text{y} \quad (\beta_1 + \dots + \beta_{n-1} + \beta) = 0;$$

de donde, $\beta = 0$.

□

Lema 1.3.2. *Sea*

$$M = \begin{bmatrix} M_1 \\ \vdots \\ M_{n-1} \\ M_n \end{bmatrix} \in \mathbb{C}^{n \times n}$$

una matriz de rango $n-1$, con sus $n-1$ primeras filas M_1, \dots, M_{n-1} linealmente independientes. Dado un $\varepsilon > 0$ existe una matriz fila $\tilde{M}_n \in \mathbb{C}^{1 \times n}$ tal que

$$\|M_n - \tilde{M}_n\|_1 < \varepsilon.$$

y la matriz

$$M = \begin{bmatrix} M_1 \\ \vdots \\ M_{n-1} \\ \tilde{M}_n \end{bmatrix}$$

tiene rango n .

DEMOSTRACIÓN. Sea $M = (m_{ij})$, siendo m_{ij} el elemento i - j de M , ($i, j = 1, \dots, n$).

Definimos la forma lineal $f : \mathbb{C}^n \rightarrow \mathbb{C}$ por la fórmula

$$f(\mathbf{x}) = f(x_1, \dots, x_n) = \begin{vmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & & \vdots \\ m_{n-1,1} & \cdots & m_{n-1,n} \\ x_1 & \cdots & x_n \end{vmatrix}$$

para todo $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{C}^n$.

Desarrollando el determinante que define f por su última fila, queda

$$f(\mathbf{x}) = a_1 x_1 + \cdots + a_n x_n.$$

Como M_1, \dots, M_{n-1} son linealmente independientes, existe algún $i \in \{1, \dots, n\}$ tal que $a_i \neq 0$. Lo que implica que

$$f(0, \dots, \overset{i}{1}, \dots, 0) = a_i \neq 0.$$

Así pues $f \neq 0$ y

$$\dim \text{Ker } f = n - 1. \quad (1.2)$$

Si ocurriera que $f(\mathbf{x}) = 0$ para todo $\mathbf{x} \in B(M_n, \varepsilon)$, por el Lema 1.3.1 existirían n vectores $\mathbf{u}_1, \dots, \mathbf{u}_n \in B(M_n, \varepsilon)$ linealmente independientes y tales que

$$f(\mathbf{u}_1) = 0, \dots, f(\mathbf{u}_n) = 0.$$

Con lo cual se tendría que $n = \dim \text{Ker } f$, lo que contradice a (1.2). Por lo tanto, existe $\tilde{M}_n \in B(M_n, \varepsilon)$ tal que $f(\tilde{M}_n) \neq 0$.

Por lo cual la matriz \tilde{M} , dada por

$$\tilde{M} := \begin{bmatrix} M_1 \\ \vdots \\ M_{n-1} \\ \tilde{M}_n \end{bmatrix},$$

tiene rango n . □

Antes de demostrar la Proposición 1.3.2, permítasenos recordar el siguiente hecho sobre aplicaciones lineales e independencia lineal.

Aserción: Sean V y W dos espacios vectoriales sobre el mismo cuerpo, y sea $f : V \rightarrow W$ una aplicación lineal. Supongamos que $\mathbf{v}_1, \dots, \mathbf{v}_r$ son vectores de V tales que los vectores $f(\mathbf{v}_1), \dots, f(\mathbf{v}_r)$ de W son linealmente independientes. Entonces los vectores $\mathbf{v}_1, \dots, \mathbf{v}_r$ son linealmente independientes.

Proposición 1.3.2. Sea $A \in \mathbb{C}^{m \times n}$ de rango r . Dado un entero s tal que $r \leq s \leq \min(m, n)$ y dado un número real $\varepsilon > 0$, existe una matriz $A' \in \mathbb{C}^{m \times n}$ de rango s tal que

$$\|A' - A\|_1 < \varepsilon.$$

DEMOSTRACIÓN. Vamos a demostrar esta proposición en primer lugar cuando $s = r + 1$. Sean $\alpha = (i_1, \dots, i_r) \in Q_{r,m}$, $\beta = (j_1, \dots, j_r) \in Q_{r,n}$ tales que el menor

$$|A[\alpha|\beta]|$$

es distinto de cero.

Esto implica que las filas A_{i_1}, \dots, A_{i_r} , de la matriz A , son linealmente independientes (basta aplicar la Aserción que precede al caso de la proyección $p : \mathbb{C}^n \rightarrow \mathbb{C}^r$ dada por $(x_1, \dots, x_n) \mapsto (x_{j_1}, \dots, x_{j_r})$). Y también implica que las restantes filas de A son combinación lineal de estas r filas.

Supongamos que $i_r \neq m$ y $j_r \neq n$. Si fuera $i_r = m$ ó $j_r = n$, la demostración que sigue podría modificarse de una manera análoga. Llamemos M a la submatriz $(r+1) \times (r+1)$ de A formada por las filas i_1, \dots, i_r, i_r+1 y las columnas j_1, \dots, j_r, j_r+1 . Es una matriz de rango r pues su menor principal de orden r de la esquina superior izquierda es igual a $|A[\alpha|\beta]|$, que es $\neq 0$, y en A todos los menores de orden $r+1$ son nulos. Aplicando a M el Lema 1.3.2 existe una matriz fila $\tilde{M}_{r+1} \in \mathbb{C}^{1 \times (r+1)}$ tal que:

$$\|(a_{i_r+1, j_1}, \dots, a_{i_r+1, j_r}, a_{i_r+1, j_r+1}) - \tilde{M}_{r+1}\|_1 < \varepsilon,$$

y la matriz $(r+1) \times (r+1)$

$$\tilde{M} = \begin{bmatrix} M_1 \\ \vdots \\ M_r \\ \tilde{M}_{r+1} \end{bmatrix},$$

donde las r primeras filas de \tilde{M} son iguales a las de M , tiene rango $r+1$.

Si $\tilde{M}_{r+1} = (\alpha_1, \dots, \alpha_{r+1})$, cambiando los elementos $a_{i_r+1, j_1}, \dots, a_{i_r+1, j_r+1}$ de A por los números $\alpha_1, \dots, \alpha_{r+1}$, respectivamente, y dejando los demás elementos de A sin cambiar, obtenemos una matriz $\tilde{A} \in \mathbb{C}^{m \times n}$ tal que

$$\|A - \tilde{A}\|_1 = \|(a_{i_r+1, j_1}, \dots, a_{i_r+1, j_r+1}) - (\alpha_1, \dots, \alpha_{r+1})\|_1 < \varepsilon.$$

Además la matriz \tilde{A} tiene rango $r+1$. En efecto,

$$|\tilde{A}[(i_1, \dots, i_r, i_r+1)|(j_1, \dots, j_r, j_r+1)]| = |\tilde{M}| \neq 0.$$

Luego, las filas $\tilde{A}_{i_1}, \dots, \tilde{A}_{i_r}, \tilde{A}_{i_r+1}$ de \tilde{A} son linealmente independientes. Cualquier otra fila \tilde{A}_i de \tilde{A} es combinación lineal de estas $r+1$ filas; de hecho, $\tilde{A}_i = A_i$ es combinación lineal de $A_{i_1} (= \tilde{A}_{i_1}), \dots, A_{i_r} (= \tilde{A}_{i_r})$. Esto prueba que $\text{rg}(\tilde{A}) = r+1$.

Sea ahora s cualquiera, $r < s \leq \min(m, n)$. Sea $t := s - r$, así pues, $r + t = s$. Consideremos la sucesión $r, r+1, r+2, \dots, r+t = s$. Por lo demostrado antes, dada la bola $B(A, \varepsilon)$ de $\mathbb{C}^{m \times n}$, existe una matriz $A^{(1)} (= \tilde{A}) \in \mathbb{C}^{m \times n}$ de rango $r+1$ tal que

$$A^{(1)} \in B(A, \varepsilon).$$

Si $\varepsilon_1 := \varepsilon - \|A - A^{(1)}\|_1$, la bola $B(A^{(1)}, \varepsilon_1)$ está contenida en la bola $B(A, \varepsilon)$ y aplicando lo anterior existe

$$A^{(2)} \in B(A^{(1)}, \varepsilon_1)$$

de rango $r+2$, (así pues existe $A^{(2)} \in B(A, \varepsilon)$ de rango $r+2$).

Tomando luego $\varepsilon_2 := \varepsilon_1 - \|A^{(1)} - A^{(2)}\|_1$, se sigue que

$$B(A^{(2)}, \varepsilon_2) \subset B(A^{(1)}, \varepsilon_1),$$

y, por lo que precede, existe $A^{(3)} \in B(A^{(2)}, \varepsilon_2)$ de rango $r+3$, etc.

De esta manera llegamos, por inducción, a probar la existencia de $A^{(t)} \in B(A, \varepsilon)$ de rango $r+t$. □

Reformulamos ahora la Proposición 1.3.2 para posterior referencia en el siguiente teorema, que tiene una demostración mucho más corta, basada en la forma canónica de A para \sim y que será extendida al caso de la forma canónica de Jordan para la relación de semejanza de matrices en el Capítulo 6.

Teorema 1.3.2. *Sea $A \in \mathbb{C}^{m \times n}$ una matriz de rango r . Sea s un entero tal que $r \leq s \leq \min(m, n)$. Entonces en todo entorno de A existe una matriz A' tal que $\text{rg } A' = s$.*

DEMOSTRACIÓN. Sean $P \in \mathbb{C}^{m \times m}$ y $Q \in \mathbb{C}^{n \times n}$ matrices invertibles tales que

$$PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

Sean $\varepsilon_{r+1}, \varepsilon_{r+2}, \dots, \varepsilon_s$ números complejos no nulos que determinaremos más adelante. Definimos la matriz A' mediante la fórmula

$$A' := P^{-1} \left(\begin{array}{cccc|c} I_r & & & & O \\ & \varepsilon_{r+1} & & & \\ & & \varepsilon_{r+2} & & \\ & & & \ddots & \\ & & & & \varepsilon_s \\ \hline & & & & O \\ & & & & O \end{array} \right) Q^{-1}.$$

Es claro que A' tiene rango s , pues P y Q son matrices invertibles. La norma de matrices $\|\cdot\|_1$ es una norma matricial o submultiplicativa, i.e. para todas matrices $M \in \mathbb{C}^{m \times n}$, $N \in \mathbb{C}^{n \times p}$ se tiene la desigualdad $\|MN\|_1 \leq \|M\|_1 \|N\|_1$, véase [10, Equations (3), pág. 166], [41, Exercise 4, pág. 358]; es decir $\|\cdot\|_1$ no sólo cumple las tres propiedades de una norma en un espacio vectorial, sino que asimismo satisface la desigualdad anterior respecto de la multiplicación de matrices. Teniendo en cuenta este hecho, de la definición de A' deducimos que

$$\|A' - A\|_1 \leq \|P^{-1}\|_1 \|Q^{-1}\|_1 (|\varepsilon_{r+1}| + |\varepsilon_{r+2}| + \dots + |\varepsilon_s|).$$

A continuación vemos que podemos elegir los números $\varepsilon_{r+1}, \varepsilon_{r+2}, \dots, \varepsilon_s$ de valor absoluto tan pequeño como queramos para que A' esté situada en el entorno que hayamos elegido de A .

□

1.4. Rango de matrices genéricas

Supongamos que de una matriz M de dimensiones $m \times n$ conocemos solamente con certeza que en determinadas posiciones tiene ceros. Por ejemplo, que

$$M = \begin{bmatrix} 0 & ? & 0 & ? & 0 \\ 0 & ? & ? & 0 & 0 \\ 0 & 0 & ? & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

donde ? significa que no sabemos qué número figura ahí. A partir de esta información ¿qué propiedades podemos inferir sobre las matrices de este modelo M ? Por ejemplo, ¿qué podemos decir sobre el rango de todas las matrices de modelo M ?

En vez de poner interrogantes ? en el modelo de matriz podemos poner asteriscos o estrellas *,

$$M = \begin{bmatrix} 0 & * & 0 & * & 0 \\ 0 & * & * & 0 & 0 \\ 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

siendo las estrellas * números complejos indeterminados. En términos imprecisos diremos que M es una matriz “genérica”. Para proceder de manera más formal, consideremos el cuerpo $\{0, *\}$ de dos elementos $0, *$ con el álgebra booleana: $* + * = *$, $0 + * = * + 0 = *$, $0 + 0 = 0$, $* \cdot * = *$, $* \cdot 0 = 0 \cdot * = 0$, $0 \cdot 0 = 0$.

Definición 1.4.1. Llamaremos *matriz genérica* $m \times n$ a toda matriz M con elementos en el cuerpo $\{0, *\}$, de dimensiones $m \times n$. En tal caso, denotaremos $M \in \{0, *\}^{m \times n}$.

Si A es una matriz con elementos en un cuerpo cualquiera, denotaremos por $A(i, j)$ el elemento de A situado en el lugar (i, j) . El *soporte* de una matriz genérica $M \in \{0, *\}^{m \times n}$ es el conjunto

$$S_M := \{(i, j) \in \{1, \dots, m\} \times \{1, \dots, n\} \mid M(i, j) = *\}.$$

Si el cardinal de S_M es igual a p , denotemos por $(i_1, j_1), (i_2, j_2), \dots, (i_p, j_p)$ los elementos de S_M ordenados según el orden lexicográfico de $\mathbb{N} \times \mathbb{N}$. Definimos la función matricial asociada a M ,

$$A_M : \mathbb{C}^p \rightarrow \mathbb{C}^{m \times n}$$

que a cada p -tupla $x = (x_1, \dots, x_p)$ de números complejos le asocia la matriz $A_M(x)$ cuyos elementos son

$$\begin{aligned} A_M(x)(i_k, j_k) &:= x_k & \text{para } k = 1, \dots, p, \\ A_M(x)(i, j) &:= 0 & \text{para } (i, j) \notin S_M. \end{aligned} \quad (1.1)$$

Ejemplo 1.4.1. Sea la matriz genérica

$$M = \begin{bmatrix} 0 & * & 0 & * & 0 \\ 0 & * & * & 0 & 0 \\ 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (1.2)$$

entonces su soporte es el conjunto

$$S_M = \{(1, 2), (1, 4), (2, 2), (2, 3), (3, 3)\}.$$

Para cada $x = (x_1, x_2, x_3, x_4, x_5) \in \mathbb{C}^5$, se tiene que

$$A_M(x) = \begin{bmatrix} 0 & x_1 & 0 & x_2 & 0 \\ 0 & x_3 & x_4 & 0 & 0 \\ 0 & 0 & x_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix};$$

esta matriz $A_M(x)$ tiene un menor de orden 3 que es un polinomio no nulo en x_1, x_2, x_3, x_4, x_5 :

$$A_M(x)[(1, 2, 3)|(2, 3, 4)] = \begin{vmatrix} x_1 & 0 & x_2 \\ x_3 & x_4 & 0 \\ 0 & x_5 & 0 \end{vmatrix} = x_2 x_3 x_5.$$

Así pues, para todas las quintuplas $x = (x_1, x_2, x_3, x_4, x_5) \in \mathbb{C}^5$ tales que $x_2 x_3 x_5 \neq 0$, se tiene que el rango de la matriz $A_M(x)$ es igual a 3.

A continuación aparecerán nociones sobre digrafos; pueden verse las definiciones necesarias en la Sección B.1 del Apéndice B. Asociamos a la matriz genérica M un digrafo no ponderado $G(M)$ así: el conjunto de vértices es

$$\{1, 2, \dots, \max(m, n)\},$$

el par (j, i) es un arco de $G(M)$ si $M(i, j) = *$ (¡nótese la inversión de los índices!).

Pretendemos averiguar el rango de “casi todas” las matrices $A_M(x)$, definidas en (1.1), cuando x varía en \mathbb{C}^p . Por “casi todas” entenderemos lo siguiente: existe un polinomio no nulo $f \in \mathbb{C}[x_1, \dots, x_p]$ tal que para todo $x \in \mathbb{C}^p$ que satisfaga $f(x) \neq 0$ la matriz $A_M(x)$ tiene rango constante, digamos r . Recordemos al efecto alguna terminología usada en geometría algebraica. Sea $f(x_1, \dots, x_p)$ un polinomio no nulo en p indeterminadas sobre un cuerpo \mathbb{F} . Denotamos por $V(f)$ el conjunto de vectores (a_1, \dots, a_p) de \mathbb{F}^p tales que $f(a_1, \dots, a_p) = 0$. Al conjunto $V(f)$ le llamaremos una *hipersuperficie algebraica* de \mathbb{F}^p . Un enunciado de la forma “el punto (o elemento) genérico de \mathbb{F}^p tiene una propiedad \mathcal{P} ” o “la propiedad \mathcal{P} , que

un elemento de \mathbb{F}^p puede cumplir o no, es *genérica*” significa precisamente que el conjunto de puntos de \mathbb{F}^p que no cumplen la propiedad \mathcal{P} está contenido en la unión de un número finito de hipersuperficies algebraicas de \mathbb{F}^p . Otra definición más general de propiedad genérica en \mathbb{K}^n o $\mathbb{K}^{m \times n}$ ($\mathbb{K} = \mathbb{R}$ o \mathbb{C}) es la siguiente: La propiedad \mathcal{P} es *genérica* si el subconjunto de puntos de \mathbb{K}^n , o de $\mathbb{K}^{m \times n}$, que verifican \mathcal{P} es abierto y denso.

Rango límite

Definición 1.4.2. Se llama *rango genérico* de la matriz genérica M al máximo

$$\rho_g(M) := \max_{x \in \mathbb{C}^p} \text{rg } A_M(x).$$

Definición 1.4.3. Sea $T = (t_{ij}) \in \mathbb{F}^{m \times n}$ donde \mathbb{F} es un cuerpo cualquiera. Sea $p \leq \max(m, n)$ un entero positivo. Sea (i_1, \dots, i_p) una sucesión no necesariamente creciente de elementos distintos de $\{1, \dots, m\}$ y sean $1 \leq j_1 < \dots < j_p \leq n$. Al vector de \mathbb{F}^p

$$(t_{i_1 j_1}, \dots, t_{i_p j_p})$$

le llamaremos una *subdiagonal* de T de dimensión p .

Definición 1.4.4. Sea $M \in \{0, *\}^{m \times n}$ una matriz genérica. Se llama *rango límite* (en inglés “term rank”) de M al rango de esta matriz, mirada como matriz con elementos en el cuerpo de dos elementos $\{0, *\}$. Lo denotamos por $\rho(M)$. Recordemos que $*$ es el elemento unidad del cuerpo $\{0, *\}$; utilizamos el símbolo $*$ en vez de 1 para sugerir la idea de elemento “genérico”

Así pues, el rango límite de M es el rango de la matriz obtenida al reemplazar todos los elementos $*$ de M por 1 y a continuación calcular el rango de la matriz de ceros y unos obtenida trabajando en el cuerpo $\mathbb{F} = \{0, 1\}$. También podemos decir que $\rho(M)$ es igual a la mayor de las dimensiones de las subdiagonales de M que tienen todas sus componentes iguales a $*$.

La matriz genérica 4×5 $M = (m_{ij})$ del Ejemplo 1.4.1 tiene rango límite 3 pues tiene la subdiagonal

$$(m_{22}, m_{33}, m_{14}) = (*, *, *)$$

de dimensión 3, y M no tiene subdiagonales de dimensión mayor que 3 formadas por elementos iguales a $*$.

Vemos que el rango límite de una matriz genérica M es igual al número maximal de asteriscos ($*$) con no dos $*$ en la misma línea (fila o columna) ¿Cómo podemos caracterizar el rango límite de M en términos de su digrafo $G(M)$? La respuesta es sencilla: $\rho(M)$ es el número maximal de arcos de $G(M)$ con no dos arcos con el mismo origen (fila), ni dos arcos con el mismo final (columna).

El digrafo $G(M)$ para la matriz genérica M del Ejemplo 1.4.1 viene dado en la figura 1.1.

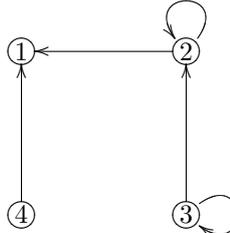


Figura 1.1: Digrafo $G(M)$.

La subdiagonal (m_{22}, m_{33}, m_{14}) de la matriz M del Ejemplo 1.4.1 viene recuadrada a continuación

$$M = \begin{bmatrix} 0 & * & 0 & \boxed{*} & 0 \\ 0 & \boxed{*} & * & 0 & 0 \\ 0 & 0 & \boxed{*} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

y los arcos en el digrafo $G(M)$ correspondientes a los elementos de esta subdiagonal vienen expresados por el subdigrafo de la figura 1.2.

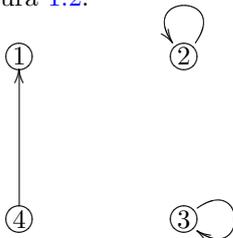


Figura 1.2: Subdigrafo de $G(M)$ de la subdiagonal.

El principal asunto de esta sección es el teorema siguiente.

Teorema 1.4.1 (Edmonds [22]). *El rango genérico y el rango límite de una matriz genérica $M \in \{0, *\}^{m \times n}$ son iguales.*

DEMOSTRACIÓN. El rango genérico de M es igual al rango p de la matriz $A_M(x^0)$ para cierto $x^0 \in \mathbb{C}^p$; sean $\alpha \in Q_{p,m}$ y $\beta \in Q_{p,n}$ tales que el menor de orden p de $A_M(x^0)$

$$\det A_M(x^0)[\alpha|\beta]$$

sea distinto de cero. Como este menor es igual a la suma del producto de las componentes de sus diagonales con el signo adecuado (véase la página 169), se tiene que debe haber una diagonal en la submatriz $A_M(x^0)$ con todas sus componentes no nulas. Así pues, esta diagonal determina en M una subdiagonal con todos sus elementos iguales a $*$. Además, no hay subdiagonales en M con sus elementos iguales a $*$ de mayor dimensión: en efecto, si hubiera una tal subdiagonal de dimensión $q > p$, pensemos en una submatriz cuadrada de orden q de $A_M(x^0)$ que la tuviera como diagonal; el determinante de esta submatriz debería ser 0, pero variando ligeramente los elementos de x^0 obtendríamos un \tilde{x}^0 tal que $\text{rg } A_M(\tilde{x}^0)$ sería $\geq q$, lo que contradice a que $\rho_g(M) = p$. Por lo tanto, $\rho_g(M) = \rho(M)$. \square

1.5. Distancia a las matrices de rango menor

Hemos visto que el rango de una matriz $A \in \mathbb{C}^{m \times n}$, $A \neq 0$, se mantiene o aumenta si perturbamos muy poco los elementos de A , por la Proposición 1.3.1. Pero existen en el espacio $\mathbb{C}^{m \times n}$ matrices de rango inferior a A ; por ejemplo, la matriz nula 0 tiene rango 0. Si $\text{rg } A = r$, el conjunto de matrices de rango menor que r , \mathcal{M}_{r-1} , es un conjunto cerrado. Por lo tanto, tiene sentido el problema de hallar la distancia de A a \mathcal{M}_{r-1} y para qué matriz $A_r \in \mathcal{M}_{r-1}$ se alcanza dicha distancia. Este problema tiene una respuesta elegante por medio de los valores singulares de la matriz A . Para la definición de los valores singulares de una matriz ver el Apéndice ???. Se llama *norma espectral* en el espacio vectorial $\mathbb{C}^{m \times n}$ a la norma

$$\|A\| := \max_{\substack{x \in \mathbb{C}^{n \times 1} \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_2} \tag{1.1}$$

asociada a las normas euclídeas en $\mathbb{C}^{n \times 1}$ y $\mathbb{C}^{m \times 1}$, identificando la matriz A con la aplicación lineal $x \mapsto Ax$ de $\mathbb{C}^{n \times 1}$ en $\mathbb{C}^{m \times 1}$. La norma euclídea en $\mathbb{C}^{n \times 1}$ se define por $\|x\| :=$

$(\sum_{i=1}^n |x_i|^2)^{1/2}$ para cada $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{C}^{n \times 1}$ y la norma euclídea de $\mathbb{C}^{m \times 1}$ se define análogamente. La norma espectral se llama también *norma de operador* y *norma de Hilbert*.

Teorema 1.5.1. *Sea $A \in \mathbb{C}^{m \times n}$ una matriz de rango r . Sea $k, 1 \leq k \leq r$, un entero dado. Sea $\|\cdot\|$ la norma espectral en $\mathbb{C}^{m \times n}$. Sean*

$$U = [u_1, \dots, u_m] \in \mathbb{C}^{m \times m} \text{ y } V = [v_1, \dots, v_n] \in \mathbb{C}^{n \times n}$$

las matrices unitarias tales que

$$U^*AV = \text{diag}(\sigma_1(A), \dots, \sigma_p(A)) \in \mathbb{C}^{m \times n},$$

(descomposición de valores singulares). Sea

$$A_{k-1} := \sum_{i=1}^{k-1} \sigma_i(A) u_i v_i^*.$$

Entonces

$$\min_{\substack{X \in \mathbb{C}^{m \times n} \\ \text{rg } X < k}} \|X - A\| = \|A_{k-1} - A\| = \sigma_k(A) \quad (1.2)$$

donde $\sigma_k(A)$ denota el k -ésimo valor singular de A , ordenados éstos valores en sentido decreciente.

Observación 1.5.1. El teorema también es verdadero poniendo $\text{rg } X = k - 1$ en lugar de $\text{rg } X < k$.

Observación 1.5.2. Es conocido que una matriz $A \in \mathbb{C}^{m \times n}$ tiene rango r si y sólo si sus valores singulares $\sigma_1(A), \dots, \sigma_p(A)$, con $p := \min(m, n)$, satisfacen que

$$\sigma_1(A) \geq \dots \geq \sigma_r(A) > \sigma_{r+1}(A) = \dots = \sigma_p(A) = 0.$$

Corolario 1. *Sea $A \in \mathbb{C}^{m \times n}$ una matriz de rango r . Sea $\|\cdot\|$ la norma espectral en $\mathbb{C}^{m \times n}$. Entonces*

$$\min_{\substack{X \in \mathbb{C}^{m \times n} \\ \text{rg } X < r}} \|X - A\| = \sigma_r(A). \quad (1.3)$$

1.6. Reducción suave de funciones matriciales

Supongamos que para cada t perteneciente a un intervalo (a, b) de la recta real, tenemos la matriz compleja $A(t) \in \mathbb{C}^{m \times n}$ que depende de forma continua de t . Llamemos $r(t) := \text{rg } A(t)$. Por ejemplo, consideremos la función $f(t) := t \text{sen}(\pi/t)$ si $t \in \mathbb{R} \setminus \{0\}$ y $f(0) := 0$ y la función matricial

$$A(t) := \begin{pmatrix} f(t) & 0 \\ 0 & t \end{pmatrix}, \quad t \in \mathbb{R}.$$

Entonces, $r(t) = 2$ para todo t real excepto para $t = 1/n$ con n entero no nulo y para $t = 0$. Además, $r(0) = 0, r(1/n) = 1$. En este ejemplo, r es discontinua en el conjunto de puntos $\mathcal{S} := \{1/n : n \text{ entero no nulo}\} \cup \{0\}$. Observemos que \mathcal{S} tiene infinitos puntos, es un conjunto cerrado y no tiene puntos interiores (i.e. no contiene intervalos abiertos). Como veremos a continuación este comportamiento es típico.

En el caso general, en principio $r(t)$ no es constante, aunque se puede demostrar que $r(t)$ es constante en “casi todos” los puntos de (a, b) . De hecho, por el principio del palomar de Dirichlet y dado que $r(t)$ sólo puede tomar valores enteros dentro del intervalo $0 \leq r(t) \leq \min(m, n)$, el valor de $r(t)$ deberá ser el mismo para infinitos puntos t de (a, b) y $r(t)$ tomará

necesariamente el mismo valor sobre un subconjunto de medida positiva de (a, b) . Incluso más, el valor de $r(t)$ es constante en cada uno de los subintervalos abiertos de una colección finita o numerable de ellos que junto con sus puntos extremos (donde baja el valor de $r(t)$) constituyen todo $[a, b]$. Como $r: (a, b) \rightarrow \mathbb{N}$ es una función con valores enteros se tiene que r es continua en un punto t_0 de (a, b) si y sólo si existe un entorno de t_0 en el que r es constante.

Teorema 1.6.1. *Supongamos que $A: (a, b) \rightarrow \mathbb{C}^{m \times n}$ es una función matricial continua en el intervalo (a, b) . Para cada $t \in (a, b)$ sea $r(t) := \text{rg } A(t)$. Entonces existe una colección finita o numerable de intervalos abiertos (a_i, b_i) disjuntos a pares tales que r es constante en cada (a_i, b_i) , r es discontinua en a_i y b_i , y la adherencia de*

$$\bigcup_i (a_i, b_i)$$

es el intervalo cerrado $[a, b]$.

DEMOSTRACIÓN. En primer lugar, tengamos en cuenta la disyuntiva siguiente:

- O bien $A(t) = O$ para todo $t \in (a, b)$; en cuyo caso, $r(t)$ es idénticamente cero en (a, b) y la conclusión del teorema es evidente tomando $(a_1, b_1) := (a, b)$.
- O bien, existe algún $t_0 \in (a, b)$ tal que $A(t_0) \neq O$.

Supongamos de ahora en adelante que estamos en este segundo caso no trivial.

Sea \mathcal{S} el subconjunto de puntos de (a, b) en los que r es discontinua. Si tomamos un menor fijo de $A(t)$ obtenemos una función continua de t . Sea $\varphi_i(t)$ la suma de los valores absolutos de todos los menores de orden i de $A(t)$. Sin pérdida de generalidad supongamos que $m \leq n$. Entonces las φ_i , $1 \leq i \leq m$, son funciones continuas en (a, b) y si $\varphi_i(t_0) = 0$, entonces $\varphi_j(t_0) = 0$ para $j \geq i$. Nótese que

$$r(t) = \begin{cases} 0 & \text{si } \varphi_1(t) = 0, \\ \text{máx}\{i \mid \varphi_i(t) \neq 0\} & \text{si } \varphi_1(t) \neq 0. \end{cases}$$

Sea $\mathcal{S}_i := \{t \in (a, b) \mid \varphi_i(t) = 0\}$. Observemos que $\mathcal{S}_i \subset \mathcal{S}_{i+1}$ y \mathcal{S}_i es cerrado pues φ_i es continua. Denotemos por $\text{Fr}(\mathcal{S}_i)$ la *frontera* de \mathcal{S}_i . Veamos que para todo i , $1 \leq i \leq m$,

$$\text{Fr}(\mathcal{S}_i) \subset \mathcal{S}.$$

En efecto, si $t_0 \in \text{Fr}(\mathcal{S}_i)$ se sigue que en todo entorno V_{t_0} de t_0 hay puntos $t'_0 \in \mathcal{S}_i$ y $t''_0 \in \mathcal{S}_i^c$. Lo que implica que en todo entorno V_{t_0} de t_0 hay puntos t'_0, t''_0 tales que $\varphi_i(t'_0) = 0$ y $\varphi_i(t''_0) \neq 0$. Por tanto, $r(t'_0) < i$ y $r(t''_0) \geq i$; de donde r es discontinua en t_0 . En consecuencia, $t_0 \in \mathcal{S}$. Por consiguiente,

$$\bigcup_{i=1}^m \text{Fr}(\mathcal{S}_i) \subset \mathcal{S}. \quad (1.1)$$

Comprobemos a continuación que

$$\mathcal{S} \subset \bigcup_{i=1}^m \text{Fr}(\mathcal{S}_i). \quad (1.2)$$

Sea ahora $t_0 \in \mathcal{S}$ y sea $k := r(t_0)$. Por ser t_0 un punto de discontinuidad de r se tiene que $k < m$. Por la semicontinuidad inferior de la función rango y la continuidad de A existe un entorno W_{t_0} de t_0 tal que $r(t) \geq r(t_0)$ siempre que $t \in W_{t_0}$. Como r es discontinua en t_0 tan cerca como queramos de t_0 debe existir un \tilde{t}_0 tal que $r(\tilde{t}_0) > r(t_0)$. Demostraremos que $t_0 \in \text{Fr}(\mathcal{S}_{k+1})$. En efecto, $\varphi_{k+1}(t_0) = 0$ pues $\text{rg } A(t_0) = k$; lo que implica que $t_0 \in \mathcal{S}_{k+1}$. Además, en todo entorno V de t_0 existe un \tilde{t}_0 tal que $\text{rg } A(\tilde{t}_0) > k$; de donde $\varphi_{k+1}(\tilde{t}_0) \neq 0$, i.e. $\tilde{t}_0 \in \mathcal{S}_{k+1}^c$. Por lo tanto, $t_0 \in \text{Fr}(\mathcal{S}_{k+1})$. Así pues,

$$\mathcal{S} = \bigcup_{i=1}^m \text{Fr}(\mathcal{S}_i).$$

Por lo tanto \mathcal{S} es un conjunto cerrado pues la unión de m conjuntos cerrados. \square

Observación 1.6.1. En un espacio topológico E se llama conjunto *diseminado* a todo subconjunto T de E cuya adherencia o clausura tenga interior vacío; en símbolos, $\overset{\circ}{\overline{T}} = \emptyset$. Véase la Definición C.2.1. Así pues, el conjunto \mathcal{S} del Teorema 1.6.1 es diseminado pues $\mathcal{S} = \overline{\mathcal{S}}$. Como el complementario de \mathcal{S} respecto de (a, b) es un abierto \mathcal{S}^c , por el teorema de estructura de los subconjuntos abiertos de la recta real [2, Teorema 3-8, pag. 43], existe una colección numerable de intervalos abiertos (a_i, b_i) , disjuntos dos a dos, cuya unión es igual a \mathcal{S}^c .

Corolario 1. *Sea \mathcal{S} el conjunto de puntos de (a, b) donde r es discontinua. Entonces \mathcal{S} es un conjunto cerrado y diseminado.*

Supongamos ahora que la función matricial A es de clase C^1 . Sabemos que para cada valor de t existen dos matrices invertibles $P_t \in \mathbb{C}^{m \times m}$ y $Q_t \in \mathbb{C}^{n \times n}$ tales que

$$P_t A(t) Q_t = \begin{pmatrix} I_{r(t)} & O \\ O & O \end{pmatrix}; \quad (1.3)$$

la cuestión que aquí nos interesa es la de saber si las aplicaciones $t \mapsto P_t$ y $t \mapsto Q_t$ pueden elegirse continuamente derivables (o de clase C^1). Una respuesta a esta cuestión nos la da el teorema siguiente que proporciona una condición necesaria y suficiente que permite reducir “suavemente” $A(t)$ a su forma canónica $C(t) := \begin{pmatrix} I_{r(t)} & O \\ O & O \end{pmatrix}$. También nos dice cuándo esta forma canónica C es una función de clase C^1 .

Teorema 1.6.2 (Doležal). *Sea $A : (a, b) \rightarrow \mathbb{C}^{m \times n}$ una función matricial de clase C^p . Si el rango de $A(t)$ es constante en (a, b) , digamos igual a k , existen dos funciones matriciales $P : (a, b) \rightarrow \mathbb{C}^{m \times m}$ y $Q : (a, b) \rightarrow \mathbb{C}^{n \times n}$ de clase C^p , tales que para cada $t \in (a, b)$:*

- (i) las matrices $P(t)$ y $Q(t)$ son invertibles,
- (ii) $P(t)A(t)Q(t) = \begin{pmatrix} I_k & O \\ O & O \end{pmatrix}$.

1.7. Ejercicios

Ejercicio 1.1. Probar que en todo entorno de una matriz no nula $A \in \mathbb{C}^{m \times n}$ hay una matriz distinta de A que tiene su mismo rango.

Ejercicio 1.2. Sea $A \in \mathbb{C}^{m \times n}$ una matriz dada y sea $(A_p)_{p=0}^{\infty}$ una sucesión de matrices de $\mathbb{C}^{m \times n}$ que converge hacia A . Demostrar que existe una subsucesión $(A_{p_k})_{k=0}^{\infty}$ de esta sucesión que converge hacia A y que está formada por matrices que tienen todas el mismo rango.

Ejercicio 1.3. Dada una matriz $A \in \mathbb{C}^{n \times n}$, demostrar que tan cerca como se quiera de A existen matrices $B \in \mathbb{C}^{n \times n}$ que tienen todas sus submatrices cuadradas invertibles.

Ejercicio 1.4. Demostrar que no existe una forma canónica continua para la relación de equivalencia \sim en $\mathbb{C}^{p \times q}$. Es decir, probar que no existe ninguna función $c : \mathbb{C}^{p \times q} \rightarrow \mathbb{C}^{p \times q}$ continua en toda matriz $A \in \mathbb{C}^{p \times q}$ tal que: (1) para cada A , $A \sim c(A)$ y (2) $A \sim B \Leftrightarrow c(A) = c(B)$.

Ejercicio 1.5. Sea $A : (a, b) \rightarrow \mathbb{C}^{m \times n}$ una función matricial continua. Para cada $t \in (a, b)$, llamemos $r(t) := \text{rg } A(t)$. Supongamos que $m \leq n$. Sea $t_0 \in (a, b)$ tal que $r(t_0) = m$. Demostrar que r es continua en t_0 .

Indicación. Utilizar una variante del principio del palomar de Dirichlet.

1.8. Notas al Capítulo 1

Relación del Teorema de Doležal con el Teorema del rango del cálculo diferencial de funciones de varias variables reales.

El Teorema 1.6.1 y el ejemplo que le precede están inspirados por la referencia [16, Theorem 10.5.2, pág. 225].

Hay otra forma de definir las matrices genéricas que aparecen en la Sección 1.4 y que es la más usual en la literatura. Ver [14, págs. 294 y 335], [27, pág. 430], [58]. Para citarla aquí, expongamos antes una definición: La familia de números complejos $(z_i)_{i=1,\dots,p}$ se dice que es algebraicamente independiente (sobre \mathbb{Z} , anillo de los números enteros), o que z_1, \dots, z_p son indeterminadas sobre \mathbb{Q} (cuerpo de los números racionales), si la relación

$$\sum_{i_1, \dots, i_p \geq 0} a_{i_1 \dots i_p} z_1^{i_1} \cdots z_p^{i_p} = 0,$$

con $a_{i_1 \dots i_p} \in \mathbb{Z}$, implica que todos los coeficientes $a_{i_1 \dots i_p}$ deben ser iguales a cero; en caso contrario, la familia $(z_i)_{i=1,\dots,p}$ es algebraicamente dependiente. Aquí

$$(a_{i_1 \dots i_p})_{i_1, \dots, i_p \geq 0}$$

es una familia de enteros casi todos nulos (i.e. todos son nulos salvo, a lo más, un número finito).

Como decimos, muchos autores llaman matriz genérica $m \times n$ a una matriz $m \times n$ de números complejos cuyos elementos no nulos forman una familia algebraicamente independiente. Por ejemplo, la matriz de $\mathbb{C}^{4 \times 5}$

$$G = \begin{bmatrix} 0 & z_1 & 0 & z_2 & 0 \\ 0 & z_3 & z_4 & 0 & 0 \\ 0 & 0 & z_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

es “genérica”, según esta definición, si los números complejos z_1, z_2, z_3, z_4, z_5 son algebraicamente independientes. El rango de esta matriz A es igual a 3, pues su menor

$$A[(1, 2, 3)|(2, 3, 4)] = \begin{vmatrix} z_1 & 0 & z_2 \\ z_3 & z_4 & 0 \\ 0 & z_5 & 0 \end{vmatrix} = z_2 z_3 z_5$$

es distinto de cero por la independencia algebraica de z_2, z_3, z_5 .

Otros significados de que la matriz $M \in \{0, *\}^{m \times n}$ tenga la propiedad \mathcal{P} genéricamente es que exista un conjunto $N \subset \mathbb{C}^p$ de medida Lebesgue nula tal que para todo $x \in \mathbb{C}^p \setminus N$, la matriz, definida por la ecuación (1.1) de la Sección 1.4, $A_M(x)$ tiene la propiedad \mathcal{P} . Como caso particular, podríamos considerar que cada número x_k de la p -tupla $x = (x_1, \dots, x_p)$ es un número aleatorio tomado en el intervalo $[0, 1]$ con distribución uniforme ($k = 1, \dots, p$). Entonces la probabilidad de que la matriz $A_M(x)$ tenga la propiedad \mathcal{P} es 1. Esta manera de pensar es muy útil para averiguar propiedades de matrices genéricas con ayuda de un *paquete* de ordenador, como por ejemplo MATLAB. Una de estas propiedades podría ser que el rango genérico r de una matriz genérica M es igual al número maximal de asteriscos (*) con no dos * en la misma línea. El ordenador tomaría al azar una matriz concreta tras otra, todas del mismo “modelo de ceros y no ceros” M , y calcularía su rango, que siempre resultaría ser igual a r .

El Ejercicio 1.3 aparece como el Lema 12.1.6 del libro de *Análisis matricial, 2008* por Jorge Antezana y Demetrio Stojanoff. http://www.iam.conicet.gov.ar/cms/files/u15/Analisis_Matricial_I.pdf.

Capítulo 2

Particiones de enteros

En este capítulo vamos a exponer una serie de resultados elementales sobre particiones de enteros positivos. La palabra partición no debe entenderse aquí como la división de un conjunto en subconjuntos disjuntos, sino como una descomposición de un entero positivo en suma de enteros positivos. Cada una de estas sumas, donde no importa el orden de los sumandos, es una partición.

Los tamaños de los bloques de Jordan asociados a un valor propio λ_0 en la forma normal de Jordan de una matriz cuadrada compleja A forman una partición de la multiplicidad algebraica m_0 de λ_0 como raíz del polinomio característico de A . Ordenaremos los órdenes de estos bloques en sentido decreciente. En la teoría de perturbación de la forma canónica de Jordan interesa sobremanera cómo se perturban estas particiones asociadas a cada valor propio λ_0 cuando los elementos de la matriz A experimentan pequeñas modificaciones.

Como quiera que los órdenes de estos bloques de Jordan asociados a λ_0 no se relacionan fácilmente con rangos de matrices asociadas a A y λ_0 , prestaremos especial atención a la partición dual (o conjugada) de estos órdenes: la que llamaremos partición de Weyr de m_0 asociada a λ_0 , en la característica de Weyr de la matriz A .

Presentaremos la mayoración (o mayorización) de particiones de enteros; tema omnipresente en este libro y en muchas partes de las Matemáticas. Adelantando ideas podemos decir que la mayoración de particiones es una relación de orden y que la perturbación esperable de la partición de Weyr asociada a λ_0 al modificar ligeramente A , expresa una especie de semicontinuidad superior respecto de dicha relación de orden. Este nexo se establece entre la partición de Weyr asociada a λ_0 y la “unión” de las particiones de Weyr asociadas a los valores propios circundantes a λ_0 de las matrices próximas a A .

2.1. Particiones

Dado un entero positivo n , se llama partición de n a cualquier descomposición de n como suma de enteros positivos:

$$n = a_1 + a_2 + \cdots + a_m.$$

Ejemplos:

$$6 = 3 + 2 + 1 \quad \text{y} \quad 6 = 2 + 2 + 2$$

son particiones de 6. Las podríamos denotar por $(3, 2, 1)$ y $(2, 2, 2)$. Otra partición de 6 es $6 = 4 + 2$ ó $(4, 2)$. En nuestro contexto trabajaremos simultáneamente con particiones de enteros diferentes. En él, estos enteros serán las multiplicidades algebraicas de valores propios. Por eso, vamos a considerar una definición más general de partición, que evita utilizar una notación para indicar el número de sumandos, m , de la descomposición.

Definición 2.1.1. Llamaremos *partición* a toda sucesión (finita o infinita) decreciente

$$a = (a_1, a_2, \dots, a_r, \dots)$$

de enteros *casi todos* nulos.

Es decir, que $a_1 \geq a_2 \geq \dots \geq a_r \geq \dots$ y los enteros $a_1, a_2, \dots, a_r, \dots$ son todos nulos excepto un número finito (o cero). Los elementos no nulos a_i se llaman las *partes* de a . La *longitud* de una partición a es el número de sus partes. Denotaremos por $\ell(a)$ la longitud de a . Así pues, $a_j = 0$ para $j > \ell(a)$. La suma de las partes de a es el *peso* de a , denotado por $|a|$,

$$|a| := a_1 + a_2 + \dots + a_{\ell(a)}.$$

Cuando sea preciso, pensaremos que a es una partición del entero $n := |a|$. Es claro que $\ell(a) \leq |a|$.

Si a y b son particiones, su suma $a + b$ es la partición cuya i -ésima componente es $a_i + b_i$. También definimos la unión $a \cup b$ como la partición cuyos componentes son los de a y b reordenados en sentido decreciente.

Por ejemplo, si

$$\begin{aligned} a &:= (5, 4, 4, 2, 2, 0, \dots) \text{ y} \\ b &:= (6, 4, 3, 3, 2, 1, 1, 0, \dots), \end{aligned}$$

entonces

$$a + b = (11, 8, 7, 5, 4, 1, 1, 0, \dots)$$

y

$$a \cup b = (6, 5, 4, 4, 4, 3, 3, 2, 2, 1, 1, 0, \dots).$$

En general, es obvio que la suma y la unión de particiones son operaciones asociativas y conmutativas en el conjunto de las particiones \mathcal{P} ; también es evidente que

$$\ell(a + b) = \max\{\ell(a), \ell(b)\}$$

y

$$\ell(a \cup b) = \ell(a) + \ell(b).$$

Observación 2.1.1. Es de notar que la unión de particiones *no* es exactamente la unión conjuntista, ya que los elementos repetidos deben aparecer repetidos también.

El *diagrama* (de Young o Ferrers) de una partición a se construye de esta manera: disponemos una fila horizontal de a_1 puntos, debajo una fila horizontal de a_2 puntos, etc. Véase en la figura 2.1 el diagrama de la partición $a = (5, 4, 4, 2, 0, \dots)$ que tiene 5 puntos o nodos en la fila superior, 4 en la fila segunda, 4 en la fila tercera y 2 en la fila cuarta.

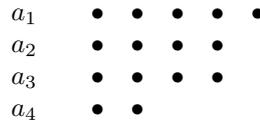


Figura 2.1: Diagrama de a .

Al dibujar estos diagramas adoptamos el convenio, igual al que se utiliza en las matrices, de que el índice de las filas i aumenta de uno en uno yendo de arriba hacia abajo, y el índice de las columnas j aumenta de uno en uno yendo de izquierda a derecha. Así pues, un diagrama de Young puede mirarse como un dibujo de puntos o como una matriz de unos (puntos) y ceros (huecos).

Definición 2.1.2. Si a es una partición dada, definimos la *partición conjugada* de a , \bar{a} , como aquella partición cuya componente i -ésima viene dada por

$$\bar{a}_i := \text{Card}\{j : a_j \geq i\}, \quad i = 1, 2, \dots$$

Por Card denotamos el cardinal. Dicho de otro modo, \bar{a}_i es igual al número de componentes de a que son mayores o iguales que i .

Ejemplo 2.1.1. Si $a = (5, 4, 4, 2, 0, \dots)$, entonces

$$\bar{a} = (4, 4, 3, 3, 1, 0, \dots).$$

La partición conjugada de una partición a es, por tanto, la partición \bar{a} cuyo diagrama se obtiene como el transpuesto del diagrama de a ; es decir, el diagrama obtenido por reflexión respecto de la diagonal principal. De aquí se sigue que \bar{a}_i es el número de puntos que hay en la columna i -ésima del diagrama de a . Por definición de partición conjugada, la primera columna del diagrama de a tiene \bar{a}_1 puntos, la segunda columna tiene \bar{a}_2 puntos, etc. Véase la figura 2.2. En particular, se tiene que $\bar{a}_1 = \ell(a)$ y $a_1 = \ell(\bar{a})$. Obviamente, $\bar{\bar{a}} = a$.

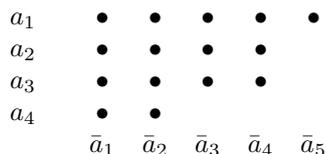


Figura 2.2: Partición conjugada de a .

Consideremos la partición $b = (4, 3, 2, 0, \dots)$ representada por el diagrama de la figura 2.3.



Figura 2.3: Diagrama de b .

Sea $a \cup b = (5, 4, 4, 4, 3, 2, 2, 0, \dots)$ la unión de las particiones $a = (5, 4, 4, 2, 0, \dots)$ y $b = (4, 3, 2, 0, \dots)$, cuyos diagramas están representados por las figuras 2.1 y 2.3, respectivamente. El diagrama de $a \cup b$ se obtiene tomando las filas de los diagramas de a y b y disponiéndolas en orden de longitud decreciente. Véase la figura 2.4

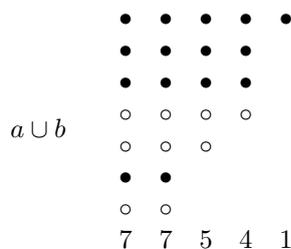


Figura 2.4: Diagrama de $a \cup b$.

Definición 2.1.3. Una partición a se llama *autoconjugada* si es igual a su conjugada, i. e. si $a = \bar{a}$.

Hay exactamente una partición autoconjugada de 6, a saber, $(3, 2, 1, 0, \dots)$. Las particiones autoconjugadas de 9 están ilustradas en la figura 2.5.

2.2. Mayoración

Trataremos de hacer precisa la noción intuitiva de cuándo unos números a_1, a_2, \dots, a_m están *menos dispersos entre sí* o son *más parecidos entre sí* que otros números b_1, b_2, \dots, b_m . Por ejemplo, a_1, a_2, \dots, a_m podrían ser los ingresos anuales de una población de m personas



Figura 2.5: Las particiones autoconjugadas de 9.

y b_1, b_2, \dots, b_m otros posibles ingresos anuales de esa misma población, ¿qué quiere decir que la riqueza que expresa la m -tupla (a_1, a_2, \dots, a_m) está mejor repartida que la que indica la m -tupla (b_1, b_2, \dots, b_m) ? Para ello vamos a definir una relación de orden, muy importante, sobre el conjunto \mathcal{P} de las particiones. A decir verdad, se tratará de dos relaciones de orden.

Definición 2.2.1. Dadas las particiones $a = (a_1, a_2, \dots), b = (b_1, b_2, \dots) \in \mathcal{P}$, escribiremos

$$a \ll b$$

si

$$\sum_{i=1}^k a_i \leq \sum_{i=1}^k b_i \quad \text{para todo } k = 1, 2, \dots$$

Así pues, $a \ll b$ implica que $|a| \leq |b|$.

Sea $m = \max\{\ell(a), \ell(b)\}$. Escribiremos

$$a \prec b$$

si

$$\sum_{i=1}^k a_i \leq \sum_{i=1}^k b_i, \quad \text{para todo } k = 1, \dots, m-1$$

y

$$\sum_{i=1}^m a_i = \sum_{i=1}^m b_i.$$

Es evidente que

$$a \prec b \iff \begin{cases} a \ll b \\ \sum_{i=1}^{\ell(a)} a_i = \sum_{i=1}^{\ell(b)} b_i. \end{cases}$$

Si $a \ll b$ se dice que a está *mayorada (o mayorizada) débilmente* por b . Si $a \prec b$ se dice que a está *mayorada (o mayorizada)* (en sentido estricto) por b .

Es fácil comprobar que \ll y \prec son relaciones de orden (parcial) sobre el conjunto de todas las particiones \mathcal{P} . En efecto,

Para toda $a \in \mathcal{P}$, $a \ll a$ y $a \prec a$. (Propiedad reflexiva).

Si $a, b \in \mathcal{P}$, $a \ll b, b \ll c$ implica $a \ll c$. (Propiedad transitiva).

Si $a, b \in \mathcal{P}$, $a \ll b, b \ll a \Rightarrow a = b$; $a \prec b, b \prec a \Rightarrow a = b$. (Propiedad antisimétrica).

Demostremos algunas de estas propiedades:

- Si $a \prec b, b \prec c \Rightarrow a \ll b, b \ll c$ y

$$\sum_{i=1}^{\ell(a)} a_i = \sum_{i=1}^{\ell(b)} b_i = \sum_{i=1}^{\ell(c)} c_i;$$

lo que implica $a \ll c$ y

$$\sum_{i=1}^{\ell(a)} a_i = \sum_{i=1}^{\ell(c)} c_i;$$

De donde, $a \prec c$. □

- Si $a \prec\prec b$ y $b \prec\prec a$, entonces para todo $k = 1, 2, \dots$,

$$\sum_{i=1}^k a_i \leq \sum_{i=1}^k b_i, \quad \sum_{i=1}^k b_i \leq \sum_{i=1}^k a_i;$$

de aquí que

$$\sum_{i=1}^k a_i = \sum_{i=1}^k b_i.$$

Luego, $a_1 = b_1$, y además para todo $k = 2, \dots$,

$$a_k = \sum_{i=1}^k a_i - \sum_{i=1}^{k-1} a_i = \sum_{i=1}^k b_i - \sum_{i=1}^{k-1} b_i = b_k.$$

De donde deducimos que $a = b$. □

Observación 2.2.1. Si a y b son particiones, entonces $a \prec b$ si y sólo si $a \prec\prec b$ y $|a| = |b|$.

Observación 2.2.2. La relación de mayoración \prec aparece en muchas partes de las Matemáticas (Álgebra, Probabilidades, Estadística, Investigación Operativa)[47]. Arnold [3] ha propuesto llamarle el *orden ubicuo*. Se halla también en Física, Química [7] y Economía (curva de Lorenz sobre las desigualdades del reparto de renta en una población)[47, pág. 5]. Si $a \prec b$, entonces los elementos $a_1, a_2, \dots, a_{\ell(a)}$ están *menos dispersos entre sí* o son *más parecidos entre sí* que los elementos $b_1, b_2, \dots, b_{\ell(b)}$.

2.3. Dualidad y monotonía

El ejemplo de la figura 2.4 ilustra la parte (i) del lema que sigue. Este lema es muy importante; pone de manifiesto la dualidad existente entre \cup y $+$ respecto de la conjugación. La parte (ii) del lema nos dice que la aplicación $a \mapsto \bar{a}$ es monótona decreciente.

Lema 2.3.1. *Si a y b son particiones, entonces*

- (i) $\overline{a \cup b} = \bar{a} + \bar{b}$,
- (ii) $a \prec b \Leftrightarrow \bar{b} \prec \bar{a}$.

DEMOSTRACIÓN. Primera prueba de (i) El diagrama de $a \cup b$ se obtiene tomando las filas de los diagramas de a y b y disponiéndolas en orden de longitud decreciente. De donde la longitud de la columna i -ésima del diagrama de $a \cup b$ es la suma de las longitudes de las columnas i -ésimas de los diagramas de a y de b ; es decir, que

$$\overline{(a \cup b)}_i = \bar{a}_i + \bar{b}_i.$$

Esto demuestra (i).

Segunda prueba de (i) Llamemos $c := a \cup b$ y definamos la función

$$d : \{1, 2, \dots, \ell(a) + \ell(b)\} \rightarrow \mathbb{N}^*$$

dada por

$$d(i) := \begin{cases} a_i & \text{si } i \in \{1, \dots, \ell(a)\}, \\ b_{i-\ell(a)} & \text{si } i \in \{\ell(a) + 1, \dots, \ell(a) + \ell(b)\}. \end{cases}$$

Sabemos que si $\bar{c} = (\bar{c}_1, \bar{c}_2, \dots, \bar{c}_i, \dots)$, entonces para cada $i = 1, \dots, \ell(a) + \ell(b)$,

$$\bar{c}_i := \text{Card} \{j \in \{1, \dots, \ell(a) + \ell(b)\} : c_j \geq i\};$$

y, como

$$\{j \in \{1, \dots, \ell(a) + \ell(b)\} : c_j \geq i\} = \{j \in \{1, \dots, \ell(a) + \ell(b)\} : d(j) \geq i\},$$

se sigue que

$$\bar{c}_i := \text{Card} \{j \in \{1, \dots, \ell(a) + \ell(b)\} : d(j) \geq i\}.$$

Ahora bien,

$$\begin{aligned} \{j \in \{1, \dots, \ell(a) + \ell(b)\} : d(j) \geq i\} = \\ \{k \in \{1, \dots, \ell(a)\} : d(k) \geq i\} \dot{\cup} \{l \in \{\ell(a) + 1, \dots, \ell(a) + \ell(b)\} : d(l) \geq i\}. \end{aligned}$$

donde $\dot{\cup}$ denota la unión *disjunta* de conjuntos; por consiguiente,

$$\bar{c}_i = \bar{a}_i + \bar{b}_i.$$

Esto demuestra (i) por segunda vez.

(ii) Como la conjugada de la partición conjugada de una partición es igual a la propia partición, basta con demostrar la implicación $a \prec b \Rightarrow \bar{b} \prec \bar{a}$. Lo haremos por reducción al absurdo. Supongamos pues que $\bar{b} \not\prec \bar{a}$. Entonces para algún $i \geq 1$ tenemos que

$$\begin{aligned} \bar{b}_1 + \dots + \bar{b}_j \leq \bar{a}_1 + \dots + \bar{a}_j \quad (1 \leq j \leq i-1) \\ \bar{b}_1 + \dots + \bar{b}_i > \bar{a}_1 + \dots + \bar{a}_i; \end{aligned} \tag{2.1}$$

de donde se sigue que $\bar{b}_i > \bar{a}_i$.

Llamemos $\beta := \bar{b}_i$, $\alpha := \bar{a}_i$. De (2.1) se deduce que

$$\bar{b}_{i+1} + \bar{b}_{i+2} + \dots < \bar{a}_{i+1} + \bar{a}_{i+2} + \dots. \tag{2.2}$$

pues $|a| = |b|$, dado que la mayoración $a \prec b$ es estricta. Ahora bien $\bar{b}_{i+1} + \bar{b}_{i+2} + \dots$ es igual al número de puntos en el diagrama de b que están situados a la derecha de la columna i -ésima, y, por lo tanto,

$$\bar{b}_{i+1} + \bar{b}_{i+2} + \dots = \sum_{j=1}^{\beta} (b_j - i),$$

donde se ha tenido en cuenta que β o \bar{b}_i es, por definición, el número de índices j tales que $b_j \geq i$.

Análogamente

$$\bar{a}_{i+1} + \bar{a}_{i+2} + \dots = \sum_{j=1}^{\alpha} (a_j - i).$$

De aquí que por (2.2) tenemos las desigualdades

$$\sum_{j=1}^{\alpha} (a_j - i) > \sum_{j=1}^{\beta} (b_j - i) \geq \sum_{j=1}^{\alpha} (b_j - i), \tag{2.3}$$

donde la desigualdad de la derecha vale porque $\beta > \alpha$ y $b_j \geq i$ para $1 \leq j \leq \beta$ por la definición de \bar{b}_i . Por (2.3) tenemos que

$$a_1 + \dots + a_{\alpha} > b_1 + \dots + b_{\alpha},$$

lo que contradice a que $a \prec b$. □

Otros lemas de interés son los dos siguientes.

Lema 2.3.2. Monotonía. Si $b_1, \dots, b_r, a_1, \dots, a_r$ son particiones tales que

$$b_i \prec\prec a_i, \quad (b_i \prec a_i),$$

respectivamente, para $i = 1, \dots, r$, entonces

$$\bigcup_{i=1}^r b_i \prec\prec \bigcup_{i=1}^r a_i, \quad \left(\bigcup_{i=1}^r b_i \prec \bigcup_{i=1}^r a_i \right),$$

respectivamente.

DEMOSTRACIÓN. Pongamos

$$d := \bigcup_{i=1}^r b_i \quad \text{y} \quad c := \bigcup_{i=1}^r a_i.$$

Para probar que $d \prec\prec c$, tenemos que demostrar que para todo $k \geq 1$

$$\sum_{i=1}^k d_i \leq \sum_{i=1}^k c_i. \quad (2.4)$$

Denotemos por b_{ij} (respectivamente, a_{ij}) la componente j -ésima de b_i (respectivamente, a_i), $i = 1, \dots, r$.

Es claro que d_1, d_2, \dots, d_k son los k términos más altos en las particiones b_1, b_2, \dots, b_r ; así pues, d_1, d_2, \dots, d_k son componentes del principio de algunas particiones b_{i_1}, \dots, b_{i_q} con $1 \leq i_1 < \dots < i_q \leq r$, $1 \leq q \leq k$. Por lo tanto,

$$\sum_{i=1}^k d_i = \sum_{j=1}^{p_1} b_{i_1 j} + \dots + \sum_{j=1}^{p_q} b_{i_q j} \quad (2.5)$$

con $p_1 + \dots + p_q = k$; $k = 1, 2, \dots$.

Como $b_i \prec\prec a_i$ para $i = 1, \dots, r$, tenemos que

$$\sum_{j=1}^{p_t} b_{i_t j} \leq \sum_{j=1}^{p_t} a_{i_t j}, \quad t = 1, \dots, q. \quad (2.6)$$

Dado que

$$\sum_{t=1}^q \sum_{j=1}^{p_t} a_{i_t j}$$

es la suma de k componentes de c , y que estas componentes no son necesariamente las primeras, tenemos que

$$\sum_{t=1}^q \sum_{j=1}^{p_t} a_{i_t j} \leq \sum_{i=1}^k c_i. \quad (2.7)$$

De (2.5), (2.6) y (2.7) deducimos (2.4).

Además, si $b_i \prec a_i$, para $i = 1, \dots, r$ y llamamos

$$n := \max\{\ell(b_1) + \dots + \ell(b_r), \ell(a_1) + \dots + \ell(a_r)\}$$

tenemos que

$$\begin{aligned} \sum_{i=1}^n d_i &= \sum_{j=1}^{\ell(b_1)} b_{1j} + \dots + \sum_{j=1}^{\ell(b_r)} b_{rj} = \\ &= \sum_{j=1}^{\ell(a_1)} a_{1j} + \dots + \sum_{j=1}^{\ell(a_r)} a_{rj} = \sum_{i=1}^n c_i. \end{aligned}$$

Así, queda probado el lema. \square

Lema 2.3.3. Sean $a = (a_1, a_2, \dots)$ y $b = (b_1, b_2, \dots)$ particiones y sean $\bar{a} = (\bar{a}_1, \bar{a}_2, \dots)$ y $\bar{b} = (\bar{b}_1, \bar{b}_2, \dots)$ sus particiones conjugadas. Entonces las dos desigualdades siguientes son equivalentes:

$$(a_1, a_2, \dots) \prec\prec (b_1, b_2, \dots), \quad (2.8)$$

$$(\bar{b}_1, \bar{b}_2, \dots) \prec (\bar{a}_1 + t, \bar{a}_2, \dots), \quad (2.9)$$

donde

$$t := \sum_{i=1}^{\ell(b)} b_i - \sum_{i=1}^{\ell(a)} a_i.$$

Cuando $t = 0$, el Lema 2.3.3 se reduce al Lema 2.3.1 (ii).

DEMOSTRACIÓN. Es claro que si $t = 0$, la equivalencia

$$a \prec\prec b \Leftrightarrow \bar{b} \prec \bar{a} + (t, 0, \dots)$$

es cierta por la parte (ii) del Lema 2.3.1. Por lo tanto, de aquí en adelante supondremos que $t \geq 1$, y utilizaremos una inducción sobre t para probar la equivalencia de (2.8) y (2.9).

Si $t = 1$ definimos una partición c de la manera siguiente:

$$\begin{aligned} c_i &:= a_i \text{ para } i = 1, \dots, \ell(a), \\ c_{\ell(a)+1} &:= 1, \\ c_i &:= 0 \text{ para } i > \ell(a) + 1. \end{aligned}$$

Es fácil ver que $a \prec\prec c \prec b$. Sea \bar{c} la partición conjugada de c . Teniendo en cuenta que $\bar{c}_1 = \ell(c)$ tenemos que $\bar{c}_1 = \bar{a}_1 + 1$ y, por la definición de c , $\bar{c}_i = \bar{a}_i$ para $i = 2, 3, \dots$. Además, por (ii) del Lema 2.3.1, se tiene que $\bar{b} \prec \bar{c}$. En consecuencia, (2.9) se sigue para $t = 1$.

Ahora suponemos que $t > 1$ y que (2.8) implica (2.9) hasta $t - 1$. Definimos una partición c exactamente como para el caso $t = 1$. Entonces

$$\sum_{i=1}^{\ell(c)} c_i - \sum_{i=1}^{\ell(a)} a_i = 1 \quad \text{y} \quad \sum_{i=1}^{\ell(b)} b_i - \sum_{i=1}^{\ell(c)} c_i = t - 1.$$

También es fácil verificar que $a \prec\prec c \prec\prec b$. Por consiguiente, usando la hipótesis de inducción y el caso $t = 1$ tenemos que

$$(\bar{b}_1, \bar{b}_2, \dots) \prec (\bar{c}_1 + t - 1, \bar{c}_2, \dots) = (\bar{a}_1 + t, \bar{a}_2, \dots).$$

Ahora vamos a demostrar que (2.9) implica (2.8). Denotando por $d = (d_1, d_2, \dots)$ la partición conjugada de $(\bar{a}_1 + t, \bar{a}_2, \dots)$ obtenemos que

$$\begin{cases} d_i = a_i & \text{para } i = 1, \dots, \ell(a) (= \bar{a}_1), \\ d_i = 1 & \text{para } i = \ell(a) + 1, \dots, \ell(a) + t, \\ d_i = 0 & \text{para } i > \ell(a) + t. \end{cases} \quad (2.10)$$

Por (ii) del Lema 2.3.1, (2.9) es equivalente a

$$(d_1, d_2, \dots) \prec (b_1, b_2, \dots),$$

y por (2.10) es claro que

$$(a_1, a_2, \dots) \prec\prec (d_1, d_2, \dots).$$

Por consiguiente,

$$(a_1, a_2, \dots) \prec\prec (b_1, b_2, \dots)$$

y el lema queda demostrado. \square

Cancelación de partes iguales

Lema 2.3.4. Sean a y b dos particiones tales que

$$a \prec b.$$

Sea $m := \max\{\ell(a), \ell(b)\}$ y supongamos que existen dos índices $l, k \in \{1, \dots, m\}$, tales que $a_l = b_k$. Entonces, llamando

$$\begin{aligned} a^1 &:= (a_1, \dots, a_{l-1}, a_{l+1}, \dots, a_m), \\ b^1 &:= (b_1, \dots, b_{k-1}, b_{k+1}, \dots, b_m), \end{aligned}$$

se tiene que

$$a^1 \prec b^1.$$

DEMOSTRACIÓN.

Caso 1º. Supongamos que $k < l$. Podemos ver las particiones expandidas para mejor comprensión.

$$\begin{aligned} a_1 \geq \dots \geq a_{k-1} \geq a_k \geq a_{k+1} \geq \dots \geq a_{l-1} \geq \boxed{a_l} \geq a_{l+1} \geq \dots \geq a_m \\ b_1 \geq \dots \geq b_{k-1} \geq \boxed{b_k} \geq b_{k+1} \geq \dots \geq b_{l-1} \geq b_l \geq b_{l+1} \geq \dots \geq b_m \\ b_1^1 \geq \dots \geq b_{k-1}^1 \geq \quad \quad \quad b_k^1 \geq \dots \end{aligned}$$

Entonces

$$\begin{cases} b_j^1 = b_j & \text{si } 1 \leq j \leq k-1, \\ a_j^1 = a_j & \text{si } 1 \leq j \leq l-1. \end{cases}$$

Por lo que para todo $p \in \{1, \dots, k-1\}$,

$$\sum_{j=1}^p a_j^1 = \sum_{j=1}^p a_j \leq \sum_{j=1}^p b_j = \sum_{j=1}^p b_j^1.$$

Cuando consideramos la suma parcial k -ésima de a^1 , resulta

$$\sum_{j=1}^k a_j^1 = \sum_{j=1}^{k+1} a_j - a_{k+1};$$

pero,

$$a_{k+1} \geq a_l = b_k \quad \text{implica} \quad -a_{k+1} \leq -b_k.$$

Por consiguiente,

$$\begin{aligned} \sum_{j=1}^{k+1} a_j - a_{k+1} &\leq \sum_{j=1}^{k+1} a_j - b_k \leq \sum_{j=1}^{k+1} b_j - b_j \\ &= b_1 + \dots + b_{k-1} + b_k + b_{k+1} - b_k \\ &= b_1^1 + \dots + b_{k-1}^1 + b_k^1 = \sum_{j=1}^k b_j^1; \end{aligned}$$

Por ende,

$$\sum_{j=1}^k a_j^1 \leq \sum_{j=1}^k b_j^1.$$

Subcaso 1.1. Si $k + 1 = l$, se tiene para las sumas parciales l -ésimas de a^1 y b^1 que

$$\begin{aligned} \sum_{j=1}^l a_j^1 &= \sum_{j=1}^{l+1} a_j - a_l \leq \sum_{j=1}^{l+1} b_j - b_k = b_1 + \cdots + b_{k-1} + \boxed{b_k} + b_{k+1} + b_{k+2} - \boxed{b_k} \\ &= b_1 + \cdots + b_{k-1} + b_{k+1} + b_{k+2} \\ &= \sum_{j=1}^l b_j^1. \end{aligned}$$

Subcaso 1.2. Si $k + 1 < l$, sea cualquier índice p tal que $k + 1 \leq p < l$; entonces

$$a_p \geq a_l = b_k, \quad \text{lo que implica} \quad -a_p \leq -b_k.$$

De donde,

$$\begin{aligned} \sum_{j=1}^p a_j^1 &= \sum_{j=1}^{p+1} a_j - a_p \leq \sum_{j=1}^{p+1} b_j - b_k = b_1 + \cdots + b_{k-1} + \boxed{b_k} + b_{k+1} + \cdots + b_{p+1} - \boxed{b_k} \\ &= b_1 + \cdots + b_{k-1} + b_{k+1} + \cdots + b_{p+1} \\ &= \sum_{j=1}^p b_j^1. \end{aligned}$$

Si p satisface $l \leq p \leq m - 1$,

$$\sum_{j=1}^p a_j^1 = \sum_{j=1}^{p+1} a_j - a_l \leq \sum_{j=1}^{p+1} b_j - a_l = \sum_{j=1}^{p+1} b_j - b_k = \sum_{j=1}^p b_j^1.$$

Caso 2º. Supongamos ahora que $l \leq k$. Podemos ver las particiones expandidas como antes.

$$\begin{aligned} a_1 \geq \cdots \geq a_{l-1} \geq \boxed{a_l} \geq a_{l+1} \geq \cdots \geq a_{k-1} \geq a_k \geq a_{k+1} \geq \cdots \geq a_m \\ b_1 \geq \cdots \geq b_{l-1} \geq b_l \geq b_{l+1} \geq \cdots \geq b_{k-1} \geq \boxed{b_k} \geq b_{k+1} \geq \cdots \geq b_m \end{aligned}$$

Si $p \in \{1, \dots, l - 1\}$, se tiene que

$$\sum_{j=1}^p a_j^1 = \sum_{j=1}^p a_j \leq \sum_{j=1}^p b_j = \sum_{j=1}^p b_j^1.$$

Para $p = l$, como $a_{l+1} \leq a_l$,

$$a_1^1 + \cdots + a_l^1 = a_1 + \cdots + a_{l-1} + a_{l+1} \leq a_1 + \cdots + a_l + a_l \leq \sum_{j=1}^l b_j = \sum_{j=1}^l b_j^1.$$

Cuando $l < p < k$,

$$a_1^1 + \cdots + a_l^1 + \cdots + a_p^1 = a_1 + \cdots + a_{l-1} + a_{l+1} + \cdots + a_{p+1},$$

y sumando miembro a miembro las desigualdades

$$a_{l+1} \leq a_l, \dots, a_{p+1} \leq a_p,$$

deducimos que

$$a_1 + \cdots + a_{l-1} + a_{l+1} + \cdots + a_{p+1} \leq a_1 + \cdots + a_{l-1} + a_l + \cdots + a_p.$$

Por lo tanto,

$$a_1^1 + \cdots + a_l^1 + \cdots + a_p^1 \leq a_1 + \cdots + a_{l-1} + b_l + \cdots + b_p = \sum_{l=1}^p b_j^1.$$

Finalmente, si $k \leq p \leq m-1$,

$$\begin{aligned} a_1^1 + \cdots + a_p^1 &= a_1 + \cdots + a_{l-1} + a_{l+1} + \cdots + a_k + \cdots + a_{p+1} \\ &= a_1 + \cdots + a_{l-1} + a_l + a_{l+1} + \cdots + a_k + \cdots + a_{p+1} - a_l \\ &\leq b_1 + \cdots + b_{p+1} - a_l = b_1 + \cdots + b_k + \cdots + b_{p+1} - b_k \\ &= b_1 + \cdots + b_{k-1} + b_{k+1} + \cdots + b_{p+1} = b_1^1 + \cdots + b_p^1; \end{aligned}$$

por tanto,

$$a_1^1 + \cdots + a_p^1 \leq b_1^1 + \cdots + b_p^1.$$

Además es obvio que $|a| = |b|$ y $a_l = b_k$ implican que $|a^1| = |b^1|$. Por consiguiente,

$$a^1 \prec b^1.$$

□

Lema 2.3.5. Sean a y b particiones que satisfacen $a \prec b$. Supongamos que existen enteros

$$1 \leq i_1 < i_2 < \cdots < i_k \leq \ell(a), 1 \leq j_1 < j_2 < \cdots < j_k \leq \ell(b)$$

tales que

$$a_{i_1} = b_{j_1}, a_{i_2} = b_{j_2}, \dots, a_{i_k} = b_{j_k}.$$

Llamemos

$$\begin{aligned} a^1 &:= (a_1, \dots, a_{i_1-1}, a_{i_1+1}, \dots, a_{i_k-1}, a_{i_k+1}, \dots, a_{\ell(a)}, 0, \dots), \\ b^1 &:= (b_1, \dots, b_{j_1-1}, b_{j_1+1}, \dots, b_{j_k-1}, b_{j_k+1}, \dots, b_{\ell(b)}, 0, \dots). \end{aligned}$$

Entonces $a^1 \prec b^1$.

DEMOSTRACIÓN. A partir del Lema 2.3.4 se demuestra fácilmente por inducción sobre k . □

2.4. Transformaciones elementales de particiones

Definición 2.4.1. Sean $a = (a_1, a_2, \dots)$ y $c = (c_1, c_2, \dots)$ particiones. Decimos que c se ha obtenido a partir de a por una *transformación elemental* siempre que existan dos índices j y k ($j < k$) tales que

$$c_j = a_j + 1, \quad c_k = a_k - 1 \quad \text{y} \quad c_i = a_i \quad \text{para todo } i \neq j, k.$$

Bajo estas condiciones es evidente que $a \prec c$.

Observación 2.4.1. Para que pueda hacerse esta transformación elemental en la partición

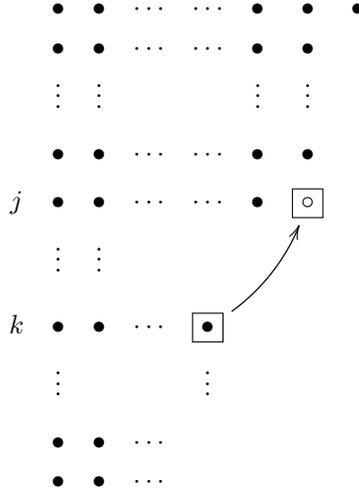
$$a = (a_1, \dots, a_j, \dots, a_k, \dots)$$

es preciso que

$$a_1, \dots, a_{j-1}, a_j + 1, a_{j+1}, \dots, a_{k-1}, a_k - 1, a_{k+1}, \dots$$

siga siendo una sucesión decreciente; es decir, es necesario que

$$a_{j-1} \geq a_j + 1, \quad a_k - 1 \geq a_{k+1}.$$

Figura 2.6: Transformación elemental sobre a .

Si pensamos en el diagrama de a , una transformación elemental sobre a consiste en llevar el último punto de la fila k moviéndolo hacia arriba a la derecha añadiéndolo al último lugar de la fila j . El diagrama resultante debe ser un diagrama de una partición (es decir, debe resultar en escalera sin salientes). Véase la figura 2.6.

Lema 2.4.1. Sean a y b particiones. La partición b se puede obtener de a por una transformación elemental si y sólo si \bar{a} se puede obtener de \bar{b} por una transformación elemental.

DEMOSTRACIÓN. Supongamos que existen índices $j, k, j < k$ tales que

$$\begin{aligned} b_j &= a_j + 1, \\ b_k &= a_k - 1, \\ b_i &= a_i \text{ para todo } i \neq j, k. \end{aligned}$$

Por la Observación 2.4.1,

$$a_{j-1} \geq a_j + 1, \quad a_k - 1 \geq a_{k+1}.$$

Llamemos $l := a_k, m := a_j + 1$. Veremos que

- (I) $l < m$,
- (II) $\bar{a}_l = \bar{b}_l + 1$,
- (III) $\bar{a}_m = \bar{b}_m - 1$,
- (IV) $\bar{a}_i = \bar{b}_i$ para todo $i \neq l, m$.

Esto probará que \bar{a} se obtiene a partir de \bar{b} por una transformación elemental que afecta a las partes de \bar{b} de subíndices l y m , y deja invariantes las otras partes.

- (I) Como $a_j + 1 > a_j \geq a_k$, se sigue que $m > l$.
- (II) Por definición

$$\bar{b}_l = \bar{b}_{a_k} = \text{Card}\{p : b_p \geq a_k\}; \quad (2.1)$$

por ser $b_k = a_k - 1$ se sigue que $b_k < a_k$; mientras que para cada $p \in \{1, \dots, k-1\}, p \neq j$,

$$b_p = a_p \geq a_k;$$

además, $b_j = a_j + 1 \geq a_j \geq a_k$; por tanto,

$$\text{Card}\{p : b_p \geq a_k\} = k - 1. \quad (2.2)$$

Por otra parte,

$$\bar{a}_l = \bar{a}_{a_k} = \{p : a_p \geq a_k\} = k, \quad (2.3)$$

ya que $a_k > a_k - 1 \geq a_{k+1}$. Por (2.1), (2.2) y (2.3), $\bar{b}_l = \bar{a}_l - 1$; de donde

$$\bar{a}_l = \bar{b}_l + 1.$$

(III) Por las definiciones

$$\bar{b}_m = \bar{b}_{a_j+1} = \text{Card}\{p : b_p \geq a_j + 1\};$$

como $a_j + 1 = b_j$, se tiene que

$$\bar{b}_m = \text{Card}\{p : b_p \geq b_j\}. \quad (2.4)$$

Probemos que

$$\text{Card}\{p : b_p \geq b_j\} = j; \quad (2.5)$$

en efecto, es obvio que

$$\text{Card}\{p : b_p \geq b_j\} \geq j;$$

para que este cardinal fuera $> j$ sería preciso que

$$b_j = b_{j+1}.$$

Pero vamos a probar que esta igualdad es imposible, tanto si $j + 1 < j$ (Caso 1) como si $j + 1 = k$ (Caso 2).

Caso 1. Si $j + 1 < k$, entonces

$$a_j + 1 = b_j = b_{j+1} = a_{j+1};$$

lo que implicaría

$$a_j < a_j + 1 = a_{j+1};$$

de donde $a_j < a_{j+1}$, que es absurdo.

Caso 2. Si $j + 1 = k$, entonces

$$a_j + 1 = b_j = b_{j+1} = b_k = a_k - 1;$$

luego

$$\begin{aligned} a_j + 1 &= a_{j+1} - 1, \\ a_j + 2 &= a_{j+1}, \\ a_j < a_j + 2 &= a_{j+1}, \end{aligned}$$

lo que implica $a_j < a_{j+1}$, que es absurdo.

Como $b_j \geq b_{j+1}$ y la igualdad $b_j = b_{j+1}$ es imposible, deberá ser

$$b_j > b_{j+1};$$

por consiguiente

$$\text{Card}\{p : b_p \geq b_j\} = j.$$

Por otra parte,

$$\bar{a}_m = \bar{a}_{a_j+1} = \text{Card}\{p : a_p \geq a_j + 1\}.$$

Demostraremos que

$$\text{Card}\{p : a_p \geq a_j + 1\} = j - 1;$$

en efecto, $a_1 \geq \dots \geq a_{j-1} \geq a_j + 1$ por la Observación 2.4.1; por lo que

$$\text{Card}\{p : a_p \geq a_j + 1\} \geq j - 1.$$

Para que este cardinal fuera $> j - 1$ sería preciso que $a_j \geq a_j + 1$; lo que es imposible pues $a_j < a_j + 1$. Por tanto, $\bar{a}_m = j - 1$, y, en virtud de (2.4) y (2.5), $\bar{b}_m = j = \bar{a}_m + 1$; de donde

$$\bar{a}_m = \bar{b}_m - 1.$$

- (iv) Para todo $i \neq l, m$ queremos demostrar que $\bar{a}_i = \bar{b}_i$. Pero esto es evidente pensando en los diagramas de Young de las particiones a y b . En la Figura 2.7, vemos que en las filas se describe la transformación elemental sobre a que lleva el último punto de la fila k -ésima a continuación del último punto de la fila j -ésima. El número de puntos en la columna i -ésima nos da las partes \bar{a}_i de \bar{a} y \bar{b}_i de \bar{b} . Observamos que todas las columnas, excepto las de índices l y m , tienen los mismos puntos para \bar{b} y \bar{a} .

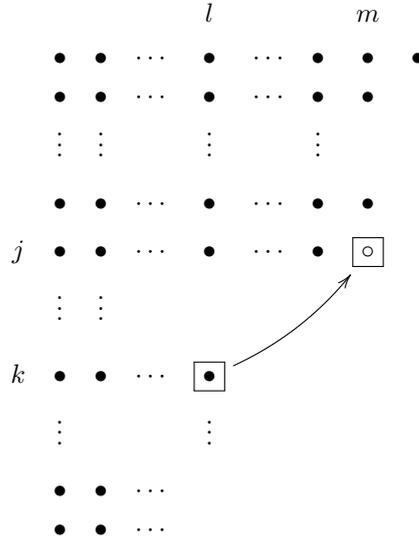


Figura 2.7: Transformando a en b .

La demostración de que si \bar{a} se puede obtener de \bar{b} por una transformación elemental, entonces b se puede obtener de a por una transformación elemental se hace de manera análoga. \square

Si giramos -90° la Figura 2.7 sobre la esquina superior izquierda se obtiene la Figura 2.8 que muestra claramente la reconstrucción de \bar{a} partiendo de \bar{b} al subir el punto \circ .

Subíndices de las partes afectadas

Si la partición a es transformada en la partición b por una transformación elemental, y j, k con $j < k$ son los subíndices de las partes de a que cambian

$$b_j := a_j + 1, \quad b_k := a_k - 1;$$

llamando

$$l := a_k, \quad m := a_j + 1,$$

se sigue que $l < m$ y, además, $\bar{a}_l = \bar{b}_l + 1$, $\bar{a}_m = \bar{b}_m - 1$ y $\bar{a}_i = \bar{b}_i$ para todo $i \neq l, m$.

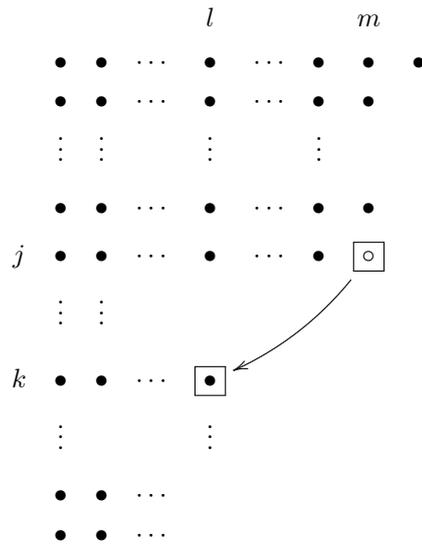


Figura 2.8: Transformando \bar{b} en \bar{a} .

Recíprocamente, supongamos que la partición conjugada \bar{b} es transformada en la partición conjugada \bar{a} por medio de una transformación elemental. En tales condiciones, si l y m son subíndices tales que

$$l < m, \bar{a}_l = \bar{b}_l + 1, \bar{a}_m = \bar{b}_m - 1, \bar{a}_i = \bar{b}_i \text{ para todo } i \neq l, m,$$

llamando

$$j := \bar{a}_m, \quad k := \bar{a}_l + 1,$$

se sigue que

$$j < k, b_j = a_j + 1, b_k = a_k - 1, b_i = a_i \text{ para todo } i \neq j, k.$$

Consideremos a continuación un ejemplo que nos muestra una partición de 16, b , que mayoriza a otra partición a , y también que b puede obtenerse a partir de a efectuado una cadena finita de transformaciones elementales.

Ejemplo 2.4.1. Sean las particiones de 16

$$a = (3, 3, 2, 2, 2, 1, 1, 1, 1, 0, \dots), \quad b = (5, 4, 3, 2, 1, 1, 0, \dots).$$

Se verifica que $a \prec b$ comprobando las desigualdades:

$$\begin{aligned} 3 &\leq 5, \\ 3 + 3 &\leq 5 + 4, \\ 3 + 3 + 2 &\leq 5 + 4 + 3, \\ 3 + 3 + 2 + 2 &\leq 5 + 4 + 3 + 2, \\ 3 + 3 + 2 + 2 + 2 &\leq 5 + 4 + 3 + 2 + 1, \\ 3 + 3 + 2 + 2 + 2 + 1 &\leq 5 + 4 + 3 + 2 + 1 + 1, \\ 3 + 3 + 2 + 2 + 2 + 1 + 1 &\leq 5 + 4 + 3 + 2 + 1 + 1 + 0, \\ 3 + 3 + 2 + 2 + 2 + 1 + 1 + 1 &\leq 5 + 4 + 3 + 2 + 1 + 1 + 0 + 0, \\ 3 + 3 + 2 + 2 + 2 + 1 + 1 + 1 + 1 &= 5 + 4 + 3 + 2 + 1 + 1 + 0 + 0 + 0. \end{aligned}$$

Hay una sucesión finita de transformaciones elementales que lleva a a b . Indiquémoslas:

$$\begin{aligned} a &= (\overset{\downarrow}{3}, 3, 2, 2, 2, 1, 1, \overset{\uparrow}{1}) \prec (\overset{\downarrow}{4}, 3, 2, 2, 2, 1, 1, \overset{\uparrow}{1}) \\ &\prec (5, \overset{\downarrow}{3}, 2, 2, 2, 1, \overset{\uparrow}{1}) \prec (5, 4, \overset{\downarrow}{2}, 2, \overset{\uparrow}{2}, 1) \\ &\prec (5, 4, 3, 2, 1, 1) = b. \end{aligned}$$

Esta cadena de transformaciones es ilustrada mediante la figura múltiple 2.9.

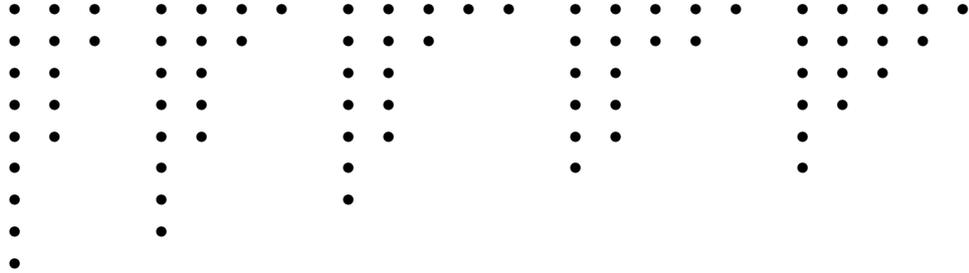


Figura 2.9: Cadena de sucesivas particiones.

Este hecho no es casual como lo muestra el lema importante siguiente.

Lema 2.4.2. Sean $a = (a_1, a_2, \dots)$ y $b = (b_1, b_2, \dots)$ particiones siendo $a \neq b$. Entonces $a \prec b$ si y sólo si b puede ser obtenida a partir de a mediante un número finito de transformaciones elementales.

DEMOSTRACIÓN. Si b se obtiene de a por medio de una sucesión finita de transformaciones elementales, es claro que $a \prec b$.

Recíprocamente, sea $a \prec b$ con $a \neq b$. Llamemos $n := \max\{\ell(a), \ell(b)\}$. Sea $\ell := \max\{i : a_i \neq b_i\}$. Como

$$a_1 + \dots + a_\ell + a_{\ell+1} + \dots + a_n = b_1 + \dots + b_\ell + b_{\ell+1} + \dots + b_n$$

y

$$a_{\ell+1} = b_{\ell+1}, \dots, a_n = b_n,$$

se sigue que

$$a_1 + \cdots + a_\ell = b_1 + \cdots + b_\ell. \quad (2.6)$$

Pero

$$a_1 + \cdots + a_{\ell-1} \leq b_1 + \cdots + b_{\ell-1};$$

más aun,

$$a_1 + \cdots + a_{\ell-1} < b_1 + \cdots + b_{\ell-1}, \quad (2.7)$$

pues si fuera $a_1 + \cdots + a_{\ell-1} = b_1 + \cdots + b_{\ell-1}$, (2.6) implicaría que $a_\ell = b_\ell$, en contradicción con la definición de ℓ . De (2.6) y (2.7) se sigue que

$$a_\ell > b_\ell. \quad (2.8)$$

Debe haber algún índice i tal que $a_i < b_i$, puesto que si ocurriera que $a_i \geq b_i$ para todo $i \in \{1, \dots, n\}$, la desigualdad (2.8) implicaría que

$$a_1 + \cdots + a_n > b_1 + \cdots + b_n,$$

contradiciendo a que $a \prec b$. Así pues, el conjunto $\{i : a_i < b_i\}$ es no vacío. Sea $j := \max\{i : a_i < b_i\}$. Por lo tanto, $a_j < b_j$; lo que implica $a_j \neq b_j$, y esto, a su vez, que $j \leq \ell$; pero $a_\ell \not< b_\ell$, luego deberá ser $j < \ell$. En consecuencia, el conjunto $\{i : j < i, a_i > b_i\}$ es no vacío, puesto que ℓ es uno de sus elementos. Tiene sentido, pues, considerar el elemento

$$k := \min\{i : j < i, a_i > b_i\};$$

así pues, $j < k$.

Con palabras, j es el máximo índice tal que $a_j < b_j$, y k es el mínimo índice mayor que j tal que $a_k > b_k$. Previamente, hemos probado que tales índices máximo y mínimo existen.

Por la elección de j y k ,

$$b_j > a_j \geq a_k > b_k$$

y

$$a_i = b_i \quad \text{para todo } i \in \{j+1, \dots, k-1\}.$$

Definamos la partición $c := (c_1, c_2, \dots)$ como

$$\begin{cases} c_i := b_i & \text{si } i \neq j, k, \\ c_j := b_j - 1, \\ c_k := b_k + 1, \end{cases}$$

Veamos, en primer lugar, que $c_1 \geq c_2 \geq \cdots \geq c_n$.

Las desigualdades triviales son ignoradas. Primero,

$$c_{j-1} = b_{j-1} \geq b_j > b_j - 1 = c_j.$$

De $b_j > a_j \geq a_{j+1} = b_{j+1}$, se sigue que $b_j > b_{j+1}$, y, por tanto, $b_j - 1 \geq b_{j+1}$. Luego, $c_j = b_j - 1 \geq b_{j+1} = c_{j+1}$.

A partir de $b_{k-1} = a_{k-1} \geq a_k > b_k$, se deduce que $b_{k-1} > b_k$; de donde, $b_{k-1} \geq b_k + 1$. Así pues, $c_{k-1} = b_{k-1} \geq b_k + 1 = c_k$.

Por otro lado,

$$c_k = b_k + 1 > b_k \geq b_{k+1} = c_{k+1}.$$

Comprobemos, en segundo lugar, que $c \prec b$:

Para todo $r \in \{1, \dots, j-1\}$, $\sum_{i=1}^r c_i = \sum_{i=1}^r b_i$; además

$$c_1 + \cdots + c_{j-1} + c_j = b_1 + \cdots + b_{j-1} + (b_j - 1) \leq b_1 + \cdots + b_{j-1} + b_j.$$

Para todo $s \in \{j+1, \dots, k-1\}$,

$$c_1 + \cdots + c_s = b_1 + \cdots + (b_j - 1) + b_{j+1} + \cdots + b_s \leq b_1 + \cdots + b_j + b_{j+1} + \cdots + b_s.$$

Para k ,

$$c_1 + \cdots + c_k = b_1 + \cdots + (b_j - 1) + \cdots + (b_k + 1) = b_1 + \cdots + b_j + \cdots + b_k.$$

Para todo $t \in \{k+1, \dots, n\}$,

$$\begin{aligned} c_1 + \cdots + c_t &= b_1 + \cdots + (b_j - 1) + \cdots + (b_k + 1) + \cdots + b_t \\ &= b_1 + \cdots + b_t. \end{aligned}$$

En tercer lugar, probemos que $a < c$:

Para todo $\nu \in \{1, \dots, j-1\}$,

$$\sum_{i=1}^{\nu} c_i = \sum_{i=1}^{\nu} b_i \geq \sum_{i=1}^{\nu} a_i.$$

Para j , como $c_j = b_j - 1 \geq a_j$, se sigue que

$$\sum_{i=1}^j c_i = b_1 + \cdots + b_{j-1} + c_j \geq a_1 + \cdots + a_{j-1} + a_j = \sum_{i=1}^j a_i.$$

Para todo $\nu \in \{j+1, \dots, k-1\}$, dado que $c_i = b_i = a_i$ para $i = j+1, \dots, k-1$, se tiene que

$$\sum_{i=1}^{\nu} c_i \geq \sum_{i=1}^{\nu} a_i.$$

Para k ,

$$c_1 + \cdots + c_{k-1} + c_k = b_1 + \cdots + (b_k + 1) \geq b_1 + \cdots + b_k \geq a_1 + \cdots + a_k.$$

Para todo $\nu > k$,

$$c_1 + \cdots + c_{\nu} = b_1 + \cdots + (b_j - 1) + \cdots + (b_k + 1) + \cdots + b_{\nu} \geq a_1 + \cdots + a_{\nu}$$

y

$$c_1 + \cdots + c_n = b_1 + \cdots + (b_j - 1) + \cdots + (b_k + 1) + \cdots + b_n = a_1 + \cdots + a_n.$$

Ahora, llamemos *discrepancia* de a a b , a la cantidad

$$d(a, b) := \sum_{i=1}^n |b_i - a_i|.$$

Entonces,

$$d(a, c) = \sum_{i=1}^n |c_i - a_i| = \sum_{\substack{i=1 \\ i \neq j, k}}^n |b_i - a_i| + |b_j - a_j - 1| + |b_k - a_k + 1|;$$

ahora bien, $b_j - a_j > 0$ y $a_k - b_k > 0$; luego

$$\begin{aligned} |b_j - a_j - 1| &= |b_j - a_j| - 1 \\ |b_k - a_k + 1| &= |a_k - b_k - 1| = |a_k - b_k| - 1 = |b_k - a_k| - 1. \end{aligned}$$

Por consiguiente,

$$d(a, c) = \sum_{i=1}^n |b_i - a_i| - 2 = d(a, b) - 2.$$

Si $d(a, c) > 0$, partiendo de la partición c y reiterando el proceso un número finito de veces, debemos llegar a la partición a . \square

2.5. Orden lexicográfico

El orden lexicográfico u orden del diccionario es el orden en el que están colocadas las palabras en un diccionario de un idioma. Este orden depende del orden de las letras en el abecedario (o alfabeto) del idioma; en nuestro caso: a, b, c, ..., x, y, z. De igual manera, el orden habitual en el conjunto de los números naturales \mathbb{N} es el orden lexicográfico mirando los números naturales como “palabras” escritas con las “letras” 0, 1, 2, 3, ..., 9. Por ejemplo, el número natural 112 es una “palabra” escrita con las “letras” (o cifras) 1 y 2. Formalizando este concepto podemos dar la definición que sigue.

Definición 2.5.1. Sea A un conjunto ordenado totalmente por una relación de orden \leq . Sea n un entero positivo. Dados dos elementos $(\alpha_1, \alpha_2, \dots, \alpha_n)$ y $(\beta_1, \beta_2, \dots, \beta_n)$ del producto cartesiano de A por sí mismo n veces, A^n , diremos que

$$(\alpha_1, \alpha_2, \dots, \alpha_n) \leq (\beta_1, \beta_2, \dots, \beta_n)$$

si ambas n -tuplas son iguales, o bien $\alpha_1 < \beta_1$, o cuando $\alpha_1 = \beta_1, \dots, \alpha_{i-1} = \beta_{i-1}, \alpha_i \neq \beta_i$, se tiene que $\alpha_i < \beta_i$. Esta relación de orden en A^n se llama el *orden lexicográfico* y también la denotamos por \leq ; es asimismo una relación de orden total.

En el ejemplo del diccionario de un idioma podemos considerar que A es el alfabeto $\{a, b, c, \dots, x, y, z\}$ dotado del orden natural $a < b < c < \dots < x < y < z$. Y que el léxico \mathcal{L} del diccionario es un subconjunto del conjunto $\cup_{k=1}^M A^k$ siendo M el máximo número de letras (o longitud) de una palabra de dicho idioma, donde cada A^k está ordenado lexicográficamente. La relación de orden utilizada en \mathcal{L} es algo diferente de la dada en la Definición 2.5.1: para indicarla se identifica cada palabra de longitud n con una n -tupla de letras; por ejemplo, “zapato” se identifica con $(z, a, p, a, t, o) \in A^6$; dadas dos “palabras” distintas $a = (a_1, a_2, \dots, a_m), b = (b_1, b_2, \dots, b_n) \in \mathcal{L}$ se dice que a es anterior (o menor) que b si $(a_1, a_2, \dots, a_m) \leq (b_1, b_2, \dots, b_m)$ en el orden lexicográfico de A^m cuando $m \leq n$, o $(a_1, a_2, \dots, a_n) \leq (b_1, b_2, \dots, b_n)$ en el orden lexicográfico de A^n en caso contrario.

De modo análogo al representar un número natural por su expresión en la base diez, \mathbb{N} se puede identificar con la unión

$$\bigcup_{k=0}^{\infty} A^k,$$

donde convendremos en que $A := \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ ordenado con el orden natural $0 < 1 < 2 < 3 < 4 < 5 < 6 < 7 < 8 < 9$ y que $A^0 := \{0\}$; estando dotado cada conjunto A^k del orden lexicográfico. La relación de orden natural \leq en \mathbb{N} resulta de considerar que, dados $p, q \in \mathbb{N}$, se tiene que $p \leq q$ si $p \leq q$ en el orden lexicográfico de A^k si ambos p y q pertenecen al mismo A^k , o bien $p \in A^k, q \in A^\ell$, con $k < \ell$. Sin embargo, esto es innecesariamente complicado y resulta más sencillo suponer que los dos números naturales p y q tienen el mismo número de cifras, d , en su desarrollo decimal, añadiendo ceros a la izquierda del que tenga menos cifras, y haciendo la comparación en A^d .

Dado que \mathbb{N} resulta así totalmente ordenado, para cada entero $n > 0$ podemos considerar el producto cartesiano \mathbb{N}^n ordenado por el orden lexicográfico inducido por el orden natural de \mathbb{N} . En particular, si denotamos por \mathcal{P}_n al conjunto de particiones de n , podemos considerar que $\mathcal{P}_n \subset \mathbb{N}^n$ identificando cada partición $a = (a_1, a_2, \dots)$ con la n -tupla (a_1, a_2, \dots, a_n) . De esta manera obtenemos sobre \mathcal{P}_n una relación de orden total: el orden lexicográfico \leq de \mathbb{N}^n . Es importante no confundir este orden \leq en \mathcal{P}_n con el orden de mayoración \prec . De hecho, sobre \mathcal{P}_n vamos a definir a renglón seguido dos nuevas relaciones de orden.

Definición 2.5.2. Sean $a := (a_1, a_2, \dots, a_n), b := (b_1, b_2, \dots, b_n) \in \mathcal{P}_n$. Llamaremos *orden lexicográfico inverso* en \mathcal{P}_n a la relación de orden definida por

$$a \overset{\text{I}}{\leq} b$$

si $a = b$, o bien $a_1 > b_1$, o cuando $a_1 = b_1, \dots, a_{i-1} = b_{i-1}, a_i \neq b_i$, se tiene que $a_i > b_i$.

Así, las particiones de 4 están ordenadas según $\stackrel{I}{\leq}$ de este modo $(4, 0, 0, 0) \stackrel{I}{\leq} (3, 1, 0, 0) \stackrel{I}{\leq} (2, 2, 0, 0) \stackrel{I}{\leq} (2, 1, 1, 0) \stackrel{I}{\leq} (1, 1, 1, 1)$.

Definición 2.5.3. Sean $a := (a_1, a_2, \dots, a_n), b := (b_1, b_2, \dots, b_n) \in \mathcal{P}_n$. Llamaremos *orden lexicográfico transpuesto* en \mathcal{P}_n a la relación de orden definida por

$$a \stackrel{T}{\leq} b$$

si $(a_n, a_{n-1}, \dots, a_2, a_1) \leq (b_n, b_{n-1}, \dots, b_2, b_1)$, donde \leq es el orden lexicográfico habitual de \mathbb{N}^n .

Curiosamente, también $(4, 0, 0, 0) \stackrel{T}{\leq} (3, 1, 0, 0) \stackrel{T}{\leq} (2, 2, 0, 0) \stackrel{T}{\leq} (2, 1, 1, 0) \stackrel{T}{\leq} (1, 1, 1, 1)$. Pero este hecho no es general, como se puede ver con el siguiente ejemplo: Para las particiones $a = (3, 1, 1, 1, 0, 0)$ y $b = (2, 2, 2, 0, 0, 0)$ de 6 se tiene que $a \stackrel{I}{\leq} b$ pues $a_1 = 3 > 2 = b_1$, pero $b \stackrel{T}{\leq} a$ ya que $000222 < 001113$.

Para estos órdenes de \mathcal{P}_n también existe un resultado de dualidad que los relaciona con la conjugación de particiones y que tiene cierta analogía con la parte (ii) del Lema 2.3.1.

Proposición 2.5.1. Sean $a, b \in \mathcal{P}_n$. Entonces

$$a \stackrel{T}{\leq} b \iff \bar{b} \stackrel{I}{\leq} \bar{a}.$$

DEMOSTRACIÓN. Supongamos que $a \stackrel{T}{\leq} b$ y $a \neq b$. Entonces

$$(a_n, a_{n-1}, \dots, a_{i+1}, a_i, \dots, a_1) \leq (b_n, b_{n-1}, \dots, b_{i+1}, b_i, \dots, b_1),$$

donde \leq es el orden lexicográfico de \mathbb{N}^n . Así pues, para algún entero $i \geq 1$ se tiene que

$$a_n = b_n, \dots, a_{i+1} = b_{i+1} \quad \text{y} \quad a_i < b_i.$$

Si llamamos $k := a_i$ y consideramos los diagramas de a y b , vemos que $\bar{a}_1 = \bar{b}_1, \dots, \bar{a}_k = \bar{b}_k$ y que $\bar{a}_{k+1} < \bar{b}_{k+1}$; por lo tanto, $\bar{b} \stackrel{I}{\leq} \bar{a}$. El recíproco se demuestra de manera semejante. \square

2.6. Una identidad sobre particiones

Vamos a exponer a continuación una identidad notable, que relaciona dos particiones con el producto “escalar” de sus particiones conjugadas. Consideremos el producto escalar ordinario en \mathbb{R}^N

$$\mathbf{x} \cdot \mathbf{y} := \sum_{k=1}^N x_k y_k,$$

siendo $\mathbf{x} = (x_1, \dots, x_N)$, $\mathbf{y} = (y_1, \dots, y_N)$ vectores de \mathbb{R}^N . Para multiplicar escalarmente dos particiones las consideraremos como vectores de \mathbb{R}^N para cualquier N mayor o igual que las longitudes de las mismas.

Proposición 2.6.1. Sean $m = (m_1, \dots, m_p, 0, \dots)$ y $n = (n_1, \dots, n_q, 0, \dots)$ particiones de longitudes p y q , respectivamente. Sean \bar{m} y \bar{n} sus particiones conjugadas respectivas, y sea $N := \max(\ell(\bar{m}), \ell(\bar{n}))$. Llamemos $I := \{1, \dots, p\} \times \{1, \dots, q\}$. Entonces

$$\sum_{(i,j) \in I} \min(m_i, n_j) = \bar{m} \cdot \bar{n}, \quad \bar{m}, \bar{n} \in \mathbb{R}^N. \quad (2.1)$$

Antes de dar la demostración analicemos esta identidad en el ejemplo siguiente.

Ejemplo 2.6.1. Sean

$$m = (5, 4, 4, 3, 2, 0, 0, 0, \dots), \quad p = \ell(m) = 5,$$

$$n = (4, 4, 3, 3, 2, 2, 1, 1, 0, \dots), \quad q = \ell(n) = 8.$$

Llamemos $\mu_{ij} := \min(m_i, n_j)$ y sea M la matriz $p \times q$ cuyos elementos son los μ_{ij} . En este caso,

$$M = \begin{pmatrix} 4 & 4 & 3 & 3 & 2 & 2 & 1 & 1 \\ 4 & 4 & 3 & 3 & 2 & 2 & 1 & 1 \\ 4 & 4 & 3 & 3 & 2 & 2 & 1 & 1 \\ 3 & 3 & 3 & 3 & 2 & 2 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 & 1 & 1 \end{pmatrix}.$$

Es fácil comprobar que todos los elementos de esta matriz suman 92:

$$\|M\|_1 = \sum_{(i,j) \in I} \mu_{ij} = 92.$$

Por otro lado,

$$\bar{m} = (5, 5, 4, 3, 1, 0, \dots),$$

$$\bar{n} = (8, 6, 4, 2, 0, 0, \dots);$$

así pues,

$$\bar{m} \cdot \bar{n} = 5 \cdot 8 + 5 \cdot 6 + 4 \cdot 4 + 3 \cdot 2 + 1 \cdot 0 = 40 + 30 + 16 + 6 = 92.$$

Daremos una idea intuitiva de la demostración de esta igualdad sobre este ejemplo. En primer lugar, todos los elementos de M son mayores o iguales que 1. En M hay $5 \cdot 8 = 40$ elementos. Consideramos la suma $\sum_{(i,j) \in I} \mu_{ij}$; si a cada elemento de M le restamos una unidad y aumentamos 40, la suma permanece igual:

$$\sum_{(i,j) \in I} \mu_{ij} = 5 \cdot 8 + \sum_{(i,j) \in I} (\mu_{ij} - 1) = 40 + 52 =$$

$$\text{Card}\{(i, j) \in I : \mu_{ij} \geq 1\} + \sum_{(i,j) \in I} (\mu_{ij} - 1).$$

Sea la matriz $M_1 = (\mu_{ij} - 1)$; M_1 es la matriz obtenida a partir de M restando 1 a cada uno de sus elementos:

$$M_1 = \begin{pmatrix} 3 & 3 & 2 & 2 & 1 & 1 & 0 & 0 \\ 3 & 3 & 2 & 2 & 1 & 1 & 0 & 0 \\ 3 & 3 & 2 & 2 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 2 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \end{pmatrix};$$

obviamente $\|M_1\|_1 = 52$.

A continuación restamos 1 a cada elemento no nulo de M_1 y añadimos este número de unos a la suma de la matriz que resulta:

$$52 = \sum_{(i,j) \in I} (\mu_{ij} - 1) = 30 + \sum_{(i,j) \in I, \mu_{ij} \geq 2} (\mu_{ij} - 2) =$$

$$\text{Card}\{(i, j) \in I : \mu_{ij} \geq 2\} + \sum_{(i,j) \in I, \mu_{ij} \geq 2} (\mu_{ij} - 2) = 30 + 22.$$

Así hemos obtenido la matriz

$$M_2 = \begin{pmatrix} 2 & 2 & 1 & 1 & 0 & 0 & 0 & 0 \\ 2 & 2 & 1 & 1 & 0 & 0 & 0 & 0 \\ 2 & 2 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

cuyo elemento del lugar ij es $\max(\mu_{ij} - 2, 0)$. La norma subuno de M_2 es igual a 22. Restamos 1 a los elementos no nulos de M_2 y sumamos 16 unidades; entonces resulta

$$22 = 16 + \sum_{(i,j) \in I, \mu_{ij} \geq 3} (\mu_{ij} - 3) =$$

$$\text{Card}\{(i, j) \in I : \mu_{ij} \geq 3\} + \sum_{(i,j) \in I, \mu_{ij} \geq 3} (\mu_{ij} - 3) = 16 + 6.$$

Tras este proceso, hemos obtenido la matriz

$$M_3 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Es evidente que $M_3 = (\max(\mu_{ij} - 3, 0))$ y que $\|M_3\|_1 = 6$.

Finalmente, restamos 1 a los elementos no nulos de M_3 y sumamos 6 unidades; se sigue que

$$6 = 6 + 0 = \text{Card}\{(i, j) \in I : \mu_{ij} \geq 4\}.$$

Observemos que $92 = 40 + 30 + 16 + 6$. Es decir, que

$$92 = \sum_{k=1}^4 \text{Card}\{(i, j) \in I : \mu_{ij} \geq k\}.$$

DEMOSTRACIÓN DE LA PROPOSICIÓN 2.6.1. Sea $\mu_{ij} := \min(m_i, n_j)$ para cada $(i, j) \in I$. Vamos a demostrar por inducción sobre t que

$$\sum_{(i,j) \in I} \mu_{ij} = \sum_{k=1}^t \text{Card}\{(i, j) \in I : \mu_{ij} \geq k\} + \sum_{(i,j) \in I, \mu_{ij} \geq t} (\mu_{ij} - t) \quad (2.2)$$

En efecto, para $t = 1$ es obvio que

$$\sum_{(i,j) \in I} \mu_{ij} = pq + \sum_{(i,j) \in I} (\mu_{ij} - 1) =$$

$$\text{Card}\{(i, j) \in I : \mu_{ij} \geq 1\} + \sum_{(i,j) \in I} (\mu_{ij} - 1).$$

Supongamos ahora cierta la fórmula (2.2) para t y demostremos que es cierta para $t + 1$: El conjunto de los elementos $(i, j) \in I$ tales que $\mu_{ij} \geq t$ es igual a la unión del conjunto de los pares $(i, j) \in I$ tales que $\mu_{ij} = t$ con el conjunto de los $(i, j) \in I$ tales que $\mu_{ij} \geq t + 1$. Así pues,

$$\begin{aligned} \sum_{(i,j) \in I, \mu_{ij} \geq t} (\mu_{ij} - t) &= \sum_{(i,j) \in I, \mu_{ij} = t} (\mu_{ij} - t) + \sum_{(i,j) \in I, \mu_{ij} \geq t+1} (\mu_{ij} - t) = \\ &= 0 + \sum_{(i,j) \in I, \mu_{ij} \geq t+1} (\mu_{ij} - t) = \\ &= \sum_{(i,j) \in I, \mu_{ij} \geq t+1} (\mu_{ij} - t - 1) + \text{Card}\{(i, j) \in I : \mu_{ij} \geq t + 1\} = \\ &= \sum_{(i,j) \in I, \mu_{ij} \geq t+1} (\mu_{ij} - (t + 1)) + \text{Card}\{(i, j) \in I : \mu_{ij} \geq t + 1\}. \end{aligned}$$

Sin pérdida de generalidad, supongamos que $m_1 \geq n_1$; entonces poniendo $t = n_1$ en (2.2) resulta que

$$\sum_{(i,j) \in I} \mu_{ij} = \sum_{k=1}^{n_1} \text{Card}\{(i,j) \in I : \mu_{ij} \geq k\}.$$

Pero,

$$\{(i,j) \in I : \mu_{ij} \geq k\} = \{i \in \{1, \dots, p\} : m_i \geq k\} \times \{j \in \{1, \dots, q\} : n_j \geq k\},$$

lo que por la Definición 2.1.2 de partición conjugada nos da

$$\sum_{(i,j) \in I} \mu_{ij} = \sum_{k=1}^{n_1} \bar{m}_k \bar{n}_k.$$

□

2.7. Ejercicios

Ejercicio 2.1. Sean a y b particiones. Demostrar que $|a \cup b| = |a| + |b|$.

Ejercicio 2.2. Toda partición de n en m partes puede ser considerada como una solución en enteros y_i de la ecuación

$$y_1 + y_2 + \dots + y_m = n,$$

con las condiciones $y_1 \geq y_2 \geq \dots \geq y_m \geq 1$.

- (1) Probar que hay tantas particiones de n como soluciones enteras (x_1, x_2, \dots, x_n) de la ecuación

$$x_1 + 2x_2 + \dots + nx_n = n$$

con $x_i \geq 0$ (= número de partes iguales a i).

- (2) Probar que hay tantas particiones de n con m partes como soluciones enteras no negativas (x_1, x_2, \dots, x_n) del sistema de dos ecuaciones

$$\begin{cases} x_1 + 2x_2 + \dots + nx_n = n, \\ x_1 + x_2 + \dots + x_n = m. \end{cases}$$

Ejercicio 2.3. Denotemos por $p_m(n)$ el número de particiones de n que tienen m partes. Demostrar que

(a) $p_{n-2}(n) = 2, \quad n \geq 4.$

(b) $p_{n-3}(n) = 3, \quad n \geq 6.$

(c) $p_2(n) = [n/2]$, donde $[n/2]$ denota el mayor entero $\leq n/2$.

(d) $p_m(n) = p_{m-1}(n-1) + p_m(n-m), \quad 1 < m < n.$

(e) Construir una tabla de valores de $p_m(n)$ para $1 \leq m \leq n, \quad 1 \leq n \leq 7.$

(f) El número de particiones de n es el número

$$p(n) = \sum_{m=1}^n p_m(n).$$

Calcular $p(n)$ para $1 \leq n \leq 7$.

Ejercicio 2.4. Escribir explícitamente

- (a) las 11 particiones de 6.
- (b) las 7 particiones de 6 que tienen a lo más 3 partes.
- (c) las 7 particiones de 6 cuyas partes son menores o iguales que 3.

Ejercicio 2.5. Sea $a = (7, 5, 3, 3, 3)$. Encontrar \bar{a}

- (1) usando diagramas de Young.
- (2) usando la Definición 2.1.2.

Ejercicio 2.6. ¿Cuáles de las particiones siguientes son autoconjugadas?

- (a) $(6, 5, 4, 3, 2, 1)$
- (b) $(5, 4, 4, 2, 1)$
- (c) $(4, 3, 3, 1)$
- (d) $(5, 5, 3, 3)$
- (e) $(6, 4, 2, 2, 1)$
- (f) $(4, 1, 1, 1)$

Ejercicio 2.7. Sean las particiones de 16

$$a = (5, 5, 3, 3) \quad \text{y} \quad b = (7, 5, 4).$$

- Comprobar que $a < b$.
- Hallar una sucesión finita de transformaciones elementales de particiones que transforma a en b . ¿Es única esta sucesión?

Ejercicio 2.8. Se llama *diagrama de Hasse* de un conjunto finito parcialmente ordenado (P, \leq) , o esquema de dominación inmediata, a un diagrama que consiste en nodos que representan los elementos de P y segmentos rectilíneos que unen algunos de estos nodos. A saber, para cada y hay un segmento (hacia abajo) de y a cualquier x si $x < y$ pero no existe ningún $z \in P$ estrictamente entre x e y : $x < z < y$ no se satisface para ningún z .

Por ejemplo, el diagrama de Hasse del conjunto de los divisores positivos de 12 ordenados por la divisibilidad viene dado en la figura 2.10.

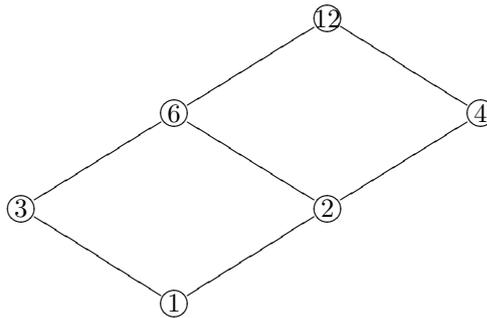


Figura 2.10: Diagrama de Hasse de los divisores de 12.

Trázese el diagrama de Hasse del conjunto de particiones de 6 ordenado por la mayoración estricta \prec .

Ejercicio 2.9. La figura 2.11 muestra el diagrama de Hasse incompleto del conjunto de las 42 particiones de 10 ordenado por la mayoración estricta \prec . Añádanse las particiones que faltan.

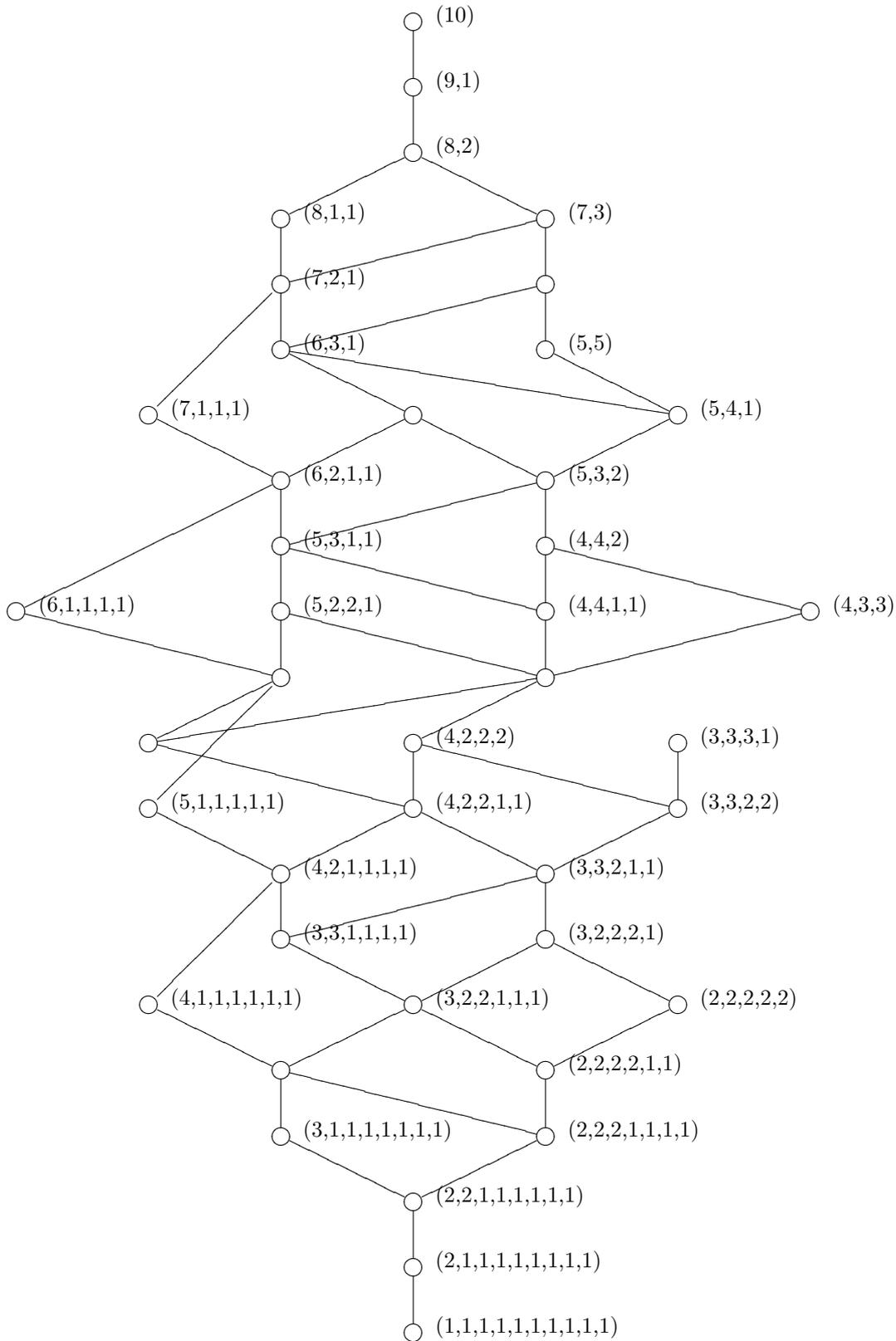


Figura 2.11: Diagrama de Hasse de las particiones de 10.

Ejercicio 2.10. Encontrar todas las particiones de 8 que

- (a) mayoran a $(5, 3)$.
- (b) están mayoradas por $(3, 2, 2, 1)$.

Ejercicio 2.11. Encontrar el menor entero n que tiene tres particiones autoconjugadas.

Ejercicio 2.12. Si $a = (a_1, a_2, \dots)$ es una partición y $\bar{a} = (\bar{a}_1, \bar{a}_2, \dots)$ es su partición conjugada, demostrar que

$$\bar{a}_{i+1} - \bar{a}_i = \text{Card}\{j : a_j = i\}.$$

Ejercicio 2.13. Sean a y b particiones y sea $c := a \cup b$. Demostrar las desigualdades

$$c_{i+j-1} \leq \max\{a_i, b_j\},$$

$$\min\{a_i, b_j\} \leq c_{i+j}.$$

Ejercicio 2.14. Sean a y b particiones. Demostrar que

- (a) Si $a \prec b$, para cualquier partición c , se sigue que $a \cup c \prec b \cup c$.
- (b) Si para alguna partición c se satisface $a \cup c \prec b \cup c$, entonces $a \prec b$.

Ejercicio 2.15. Si a y b son particiones, (i) ¿es cierto que $\overline{a+b} = \bar{a} \cup \bar{b}$? (ii) ¿es cierto que $a \cup b \prec a+b$?

Ejercicio 2.16. Sean $b_1, \dots, b_r, a_1, \dots, a_r$ particiones que satisfacen las condiciones: (1) para cada $i = 1, \dots, r$, $b_i \prec\prec a_i$ y (2) $\cup_{i=1}^r b_i \prec \cup_{i=1}^r a_i$. Demostrar que para todo $i = 1, \dots, r$, $b_i \prec a_i$. ¿Subsiste esta conclusión si retiramos la hipótesis (1)?

Ejercicio 2.17. Dada una partición p , denotemos por $E(p)$ la media aritmética de sus partes. Sean a y b particiones tales que $a \prec b$; demostrar que esto implica que

$$E(a) \leq E(b).$$

Ejercicio 2.18. Demostrar que el número de particiones de n en partes distintas e impares es igual al número de particiones de n autoconjugadas.

Indicación. Utilizar la biyección que a cada partición de n en partes distintas e impares le asocia una partición de n autoconjugada, como sugiere el ejemplo de la figura 2.12, para $n = 11$, en la que cada parte (impar) de la partición $7 + 3 + 1 = 11$ da lugar a un “ángulo recto” en el diagrama de la partición autoconjugada $4 + 3 + 3 + 1 = 11$.

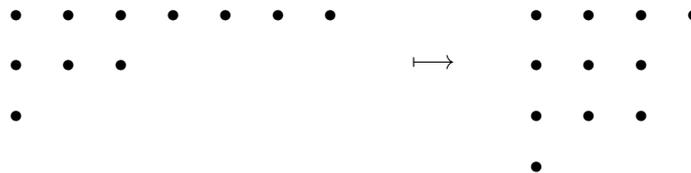


Figura 2.12: Biyección asociada a las particiones autoconjugadas.

Ejercicio 2.19. Mediante la biyección que asocia a una partición su partición conjugada, demostrar que el número, $p(n, m)$, de particiones de n que tienen a lo más m partes es igual al número de particiones de n cuyas partes son menores o iguales que m .

Demostrar también que $p(n, m)$ es igual al número de particiones de $n + m$ cuya parte mayor es igual a m .

Ejercicio 2.20. Denotemos por $p(n, m)$ el número de particiones de n en a lo más m partes; así pues, el número de particiones de n con m partes, $p_m(n)$, es igual a $p(n, m) - p(n, m - 1)$.

Demostrar que si $1 \leq n \leq m$ se tiene que $p(n, m) = p(n)$ y, para $2 \leq m \leq n$:

$$p(n, m) = p(n, m - 1) + p(n - m, m);$$

$$p(n, 1) = 1, \quad p(0, m) := 1.$$

Ejercicio 2.21. Asociando a cada partición $n = a_1 + a_2 + \dots + a_m$ con $a_1 > a_2 > \dots > a_m \geq 1$ la m -tupla $(b_1, b_2, \dots, b_{m-1}, b_m)$ definida por

$$b_1 := a_1 - a_2 - 1, \dots, b_{m-1} := a_{m-1} - a_m - 1, b_m := a_m - 1,$$

demostrar que el número de particiones de n en m partes diferentes es igual al número de particiones de $n - \binom{m+1}{2}$ en a lo más m partes (\Leftrightarrow en partes $\leq m$).

Ejercicio 2.22. Utilizando el diagrama de puntos de una partición, demostrar que el número de particiones a de n cuya partición conjugada $\bar{a} = (\bar{a}_1, \bar{a}_2, \dots)$ verifica que $\bar{a}_1 = \bar{a}_2$, es igual a

$$p^*(n) := p(n) - p(n - 1) = \sum_{m=1}^{\lfloor n/2 \rfloor} p(n - 2m, m),$$

donde $p(n - 2m, m)$ denota el número de particiones de $n - 2m$ en a lo más m partes.

¿Qué significa $\bar{a}_1 = \bar{a}_2$ en términos de las partes de a ? Deducir de esta interpretación que $p^*(n)$ es una sucesión creciente. Por lo tanto, la sucesión $p(n)$ es convexa, es decir que

$$\Delta^2 p(n) := p(n + 2) - 2p(n + 1) + p(n) \geq 0$$

para todo $n \geq 1$.

Ejercicio 2.23.

- (1) Sean $a = (a_1, a_2, \dots)$ y $c = (c_1, c_2, \dots)$ particiones tales que c puede obtenerse de a por una transformación elemental. Demostrar que esto implica que

$$\sum_{i=1}^{\ell(a)} a_i^2 < \sum_{i=1}^{\ell(c)} c_i^2.$$

- (2) Sean $a = (a_1, a_2, \dots)$ y $b = (b_1, b_2, \dots)$ particiones tales que $a \prec\prec b$. Demostrar que entonces

$$\sum_{i=1}^{\ell(a)} a_i^2 \leq \sum_{i=1}^{\ell(b)} b_i^2,$$

dándose la igualdad si y sólo si $a = b$.

Ejercicio 2.24. Sean las particiones de 6 $a = (2, 2, 1, 1, 0, 0)$, $b = (3, 2, 1, 0, 0, 0)$; comprobar que $a \prec b$ y que $b \stackrel{T}{\leq} a, b \stackrel{I}{\leq} a$. Demostrar que, en general, si $a, b \in \mathcal{P}_n$ satisfacen $a \prec b$, entonces se sigue que $b \stackrel{T}{\leq} a$ y $b \stackrel{I}{\leq} a$.

Considerando las particiones $a = (5, 5, 1, 1)$, $b = (6, 3, 3)$ de 12, demostrar que el recíproco no es cierto.

Ejercicio 2.25. Sean $0 \leq n_1 \leq n_2 \leq \dots \leq n_r$ y $0 \leq m_1 \leq m_2 \leq \dots \leq m_s$ sucesiones crecientes de enteros. Demostrar que

$$\sum_{i=1}^r \sum_{j=1}^r \min(n_i, n_j) + \sum_{i=1}^s \sum_{j=1}^s \min(m_i, m_j) - 2 \sum_{i=1}^r \sum_{j=1}^s \min(n_i, m_j) \geq 0$$

y que la igualdad a 0 ocurre si y sólo si $r = s$ y $n_i = m_i$ para todo i .

Ejercicio 2.26. Si $a = (a_1, a_2, \dots, a_m)$ es una partición de un entero positivo en m partes y $\bar{a} = (\bar{a}_1, \bar{a}_2, \dots)$ es su partición conjugada, demostrar que

$$a_1 + 3a_2 + 5a_3 + \dots + (2m-1)a_m = \bar{a}_1^2 + \bar{a}_2^2 + \dots + \bar{a}_{a_1}^2.$$

Ejercicio 2.27. Sean $r_1 \geq r_2 \geq \dots \geq r_m$ y $c_1 \geq c_2 \geq \dots \geq c_n$ enteros no negativos. Una condición necesaria y suficiente para que exista una matriz $m \times n$ de ceros y unos cuyas filas sumen r_1, r_2, \dots, r_m , respectivamente, y cuyas columnas sumen c_1, c_2, \dots, c_n , respectivamente, es que

$$c \prec \bar{r},$$

donde $c := (c_1, c_2, \dots, c_n)$ y $r := (r_1, r_2, \dots, r_m)$. (Teorema de Gale y Ryser, 1957. Véase [47, Theorem C.1, pág. 176]).

- Demostrar que la condición es necesaria.
- Encontrar una matriz 4×6 de ceros y unos cuyas filas sumen 6,5,3,2, respectivamente, y cuyas columnas sumen 4,3,3,3,2,1, respectivamente.
- Encontrar una matriz 3×5 de ceros y unos cuyas filas sumen 4,2,0, respectivamente, y cuyas columnas sumen 2,1,1,1,1, respectivamente.

Ejercicio 2.28. Una partición $a = (a_1, a_2, \dots)$ se puede denotar por

$$a = 1^{\alpha_1} 2^{\alpha_2} \dots a_1^{\alpha_{a_1}},$$

con la que se indica que a tiene α_1 unos, α_2 doses, \dots, α_{a_1} partes iguales a a_1 . Utilizando esta notación dar otra demostración de la Proposición 2.6.1.

Ejercicio 2.29. Sean m, s enteros positivos. Sean $a = (a_1, \dots, a_m)$ y $b = (b_1, \dots, b_m)$ dos m -tuplas de enteros con $a_1 \geq \dots \geq a_m \geq 0, b_1 \geq \dots \geq b_m \geq 0, \sum_{i=1}^m a_i = s, \sum_{i=1}^m b_i = s$. Supongamos que para todas las m -tuplas decrecientes de enteros no negativos (x_1, \dots, x_m) , cuya suma es s , tenemos que

$$(a_1, \dots, a_m) \prec (x_1, \dots, x_m) \text{ si y sólo si } (b_1, \dots, b_m) \prec (x_1, \dots, x_m).$$

Demostrar que

$$(a_1, \dots, a_m) = (b_1, \dots, b_m).$$

Ejercicio 2.30. Sea \mathcal{P} el conjunto de las particiones, y sean las particiones

$$e^{(j)} := (\overbrace{1, \dots, 1}^j, 0, \dots), \quad j = 1, 2, \dots$$

Demostrar que para toda $a \in \mathcal{P}$ existen enteros $\alpha_j \geq 0, j = 1, 2, \dots, \ell(a)$ tales que

$$a = \sum_{j=1}^{\ell(a)} \alpha_j e^{(j)} = \sum_{j=1}^{\infty} \alpha_j e^{(j)}.$$

Ejercicio 2.31. Sea \mathcal{P} el conjunto de las particiones. Sean $a = (a_1, a_2, \dots), b = (b_1, b_2, \dots) \in \mathcal{P}$. Definimos su producto “escalar” mediante

$$a \cdot b := \sum_{i=1}^{\infty} a_i b_i.$$

Demostrar que $a \prec\prec b$ si y sólo si para toda $c \in \mathcal{P}$ se tiene que $a \cdot c \leq b \cdot c$.

Ejercicio 2.32. Sean a y b particiones. Sea $d \in \mathbb{N}$. Demostrar que $\bar{b}_i \geq \bar{a}_{i+d}$ y $\bar{a}_i \geq \bar{b}_{i+d}$ para todo $i > 0$ si y sólo si $|a_i - b_i| \leq d$ para todo $i > 0$.

2.8. Notas al Capítulo 2

La noción de partición de un número entero positivo pertenece tanto a la teoría de los números como a la combinatoria. Edificada a fines del siglo XVIII por Euler, muchos otros grandes matemáticos —Cayley, Gauss, Hardy, Lagrange, Legendre, Littlewood, Rademacher, Ramanujan, Schur y Sylvester— han contribuido al desarrollo de la teoría. La referencia clásica sobre particiones de enteros es el libro de Andrews [1]. Al lector interesado en la teoría combinatoria de particiones le recomendamos el librito de Comtet [18]. La obra estándar sobre mayoración de particiones (y en general sobre mayoración de n -tuplas de números reales) es la de Marshall y Olkin [47]; fue precedida por el libro de Hardy, Littlewood y Pólya [35], que tuvo una enorme influencia.

La segunda prueba del apartado (i) del Lema 2.3.1 es debida a Gorka Armentia, quien me la comunicó en marzo de 2019. La demostración del Lema 2.3.1 aparece en Macdonald [44, págs. 5–7]. En general, este libro ha sido fuente de inspiración para la mayor parte de este capítulo, y hay mucho otro material tomado de él. El Lema 2.3.4 es un resultado debido a Itziar Baragaña, que lo expuso en un seminario celebrado en Vitoria-Gasteiz el 16 de mayo de 2003. La identidad sobre particiones de la Sección 2.6 viene dada en Beitia y Gracia [9, Proposition 1.1].

Los ejercicios propuestos están elegidos con la pretensión de que sean accesibles desde las ideas expuestas o sugeridas en el presente capítulo, algunos son fáciles y otros no tanto, aunque se ha intentado dar las indicaciones suficientes para hacerlos razonables en este contexto. Hay abundantes joyas matemáticas en la bibliografía —el tema se presta a ello— que hemos desechado por los motivos antedichos. Las referencias que se dan a continuación sobre los ejercicios tienen por finalidad principal dar el crédito debido a cada autor, y no el de servir de “libro de soluciones”. El Ejercicio 2.2 viene en Comtet [18], pág. 105, Théorème A. El Ejercicio 2.3 está en Merris [50], pág. 21, nº 1. Bastantes ejercicios numéricos (Ejercicios 2.4, 2.5, 2.6 y 2.10) están tomados o inspirados en los ejercicios del Chapter 1, pág. 21–25, de Merris [50]. El Ejercicio 2.8 ha sido tomado de Merris [50], pág. 4, de James y Kerber [38], pág. 23, y la definición de diagrama de Hasse y el ejemplo de los divisores de 12 está tomada de Melzak [49], pág. 196–197. También han sido tomados de Merris [50], pág. 21–25, los Ejercicios 2.11, 2.12 y 2.15(i). El Ejercicio 2.13 me fue comunicado por J. Queiró y A. Fonseca en Valencia el año 1989. El Ejercicio 2.18 se atribuye a Sylvester (Véase Andrews [1], pág. 14, Example 8); la sugerencia de la demostración está inspirada en Comtet [18], pág. 111, Théorème B. El Ejercicio 2.19 es el Theorem 1.4, pág. 8 de Andrews [1], donde puede verse una generalización interesante, y es el Théorème A, pág. 110 de Comtet [18]. El Ejercicio 2.20 es el Théorème B, pág. 106 de Comtet [18]. El Ejercicio 2.21 es el Théorème C, pág. 116 de Comtet [18]. El Ejercicio 2.22 es una reelaboración del Exercice 3, pág. 125, de Comtet [18]. El Ejercicio 2.23 es una reformulación de los Lemas 1.12 y 1.13, pág. 108, de la Tesis de Beitia [8]. El Ejercicio 2.24 es la Proposition 1.10, pág. 6, de Macdonald [44]. El Ejercicio 2.25 es el Lemma 1, pág. 155, de Byrnes y Gauger [15]. El Ejercicio 2.29 es la Proposition 5 de Gracia, Hoyos, Zaballa [33]. Los Ejercicios 2.30, 2.31 constituyen el Theorem 3.2 de Lippert [42]. El Ejercicio 2.32 es el Lemma 3.2 de Lippert y Strang [43].

Capítulo 3

Invariantes de semejanza y rangos

En este capítulo vamos a exponer una serie de resultados algebraicos previos necesarios para establecer los teoremas sobre la perturbación de la forma canónica de Jordan en el Capítulo 6. Un sistema completo de invariantes para la semejanza de matrices cuadradas complejas está formado por unos invariantes enteros (los órdenes de los bloques de Jordan) y otros invariantes continuos (los valores propios). Aquí insistiremos en la primera clase de invariantes. Nuestro énfasis estará puesto en la característica de Weyr de una matriz cuadrada compleja, que está asociada de manera natural al cálculo de los invariantes enteros de la semejanza mediante rangos de matrices.

También estableceremos un criterio racional de semejanza de dos matrices cuadradas mediante el cálculo de los rangos de tres matrices.

Así pues, la noción de rango de una matriz sigue teniendo una importancia esencial en este capítulo.

3.1. Desigualdad de Frobenius

Vamos a exponer dos resultados sobre rangos de matrices, que nos serán de utilidad más adelante al tratar de la característica de Weyr. Sea \mathbb{F} un cuerpo (conmutativo) cualquiera.

Lema 3.1.1. *Sean matrices cualesquiera $A \in \mathbb{F}^{m \times p}$, $B \in \mathbb{F}^{p \times n}$. Entonces,*

$$\dim B[\text{Ker}(AB)] = \text{rg}(B) - \text{rg}(AB).$$

DEMOSTRACIÓN. Sean $\tau := \dim \text{Ker}(AB)$ y $\sigma := \dim \text{Ker } B$; entonces $\tau = n - \text{rg}(AB)$, $\sigma = n - \text{rg}(B)$.

Si $\mathbf{x} \in \mathbb{F}^{n \times 1}$ y $B\mathbf{x} = 0$, entonces $AB\mathbf{x} = 0$; luego

$$\text{Ker } B \subset \text{Ker}(AB).$$

De aquí que $\sigma \leq \tau$. Sea ahora $\{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_\sigma\}$ una base de $\text{Ker } B$, ampliamos esta base hasta obtener una base $\{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_\sigma, \mathbf{z}_{\sigma+1}, \dots, \mathbf{z}_\tau\}$ de $\text{Ker}(AB)$. Veamos que los vectores $B\mathbf{z}_{\sigma+1}, \dots, B\mathbf{z}_\tau$ son linealmente independientes. En efecto, si se tiene una relación lineal

$$\alpha_{\sigma+1}B\mathbf{z}_{\sigma+1} + \dots + \alpha_\tau B\mathbf{z}_\tau = 0, \quad \text{con } \alpha_{\sigma+1}, \dots, \alpha_\tau \in \mathbb{F},$$

entonces

$$B(\alpha_{\sigma+1}\mathbf{z}_{\sigma+1} + \dots + \alpha_\tau\mathbf{z}_\tau) = 0;$$

por lo tanto existen $\alpha_1, \dots, \alpha_\sigma \in \mathbb{F}$ tales que

$$\alpha_{\sigma+1}\mathbf{z}_{\sigma+1} + \dots + \alpha_\tau\mathbf{z}_\tau = \alpha_1\mathbf{z}_1 + \dots + \alpha_\sigma\mathbf{z}_\sigma.$$

Así pues

$$\alpha_1 = \cdots = \alpha_\sigma = \alpha_{\sigma+1} = \cdots = \alpha_\tau = 0.$$

Por otro lado,

$$B[\text{Ker}(AB)] = \langle Bz_1, \dots, Bz_\sigma, Bz_{\sigma+1}, \dots, Bz_\tau \rangle,$$

donde se ha utilizado la notación

$$\langle \mathbf{x}_1, \dots, \mathbf{x}_r \rangle := \{ \alpha_1 \mathbf{x}_1 + \cdots + \alpha_r \mathbf{x}_r \mid \alpha_i \in \mathbb{F}, \quad i = 1, \dots, r \}$$

con $\mathbf{x}_1, \dots, \mathbf{x}_r$ vectores.

Como $Bz_1 = 0, \dots, Bz_\sigma = 0$, se sigue de lo anterior que

$$B[\text{Ker}(AB)] = \langle Bz_{\sigma+1}, \dots, Bz_\tau \rangle.$$

Por consiguiente, $\{Bz_{\sigma+1}, \dots, Bz_\tau\}$ es una base $B[\text{Ker}(AB)]$. En consecuencia,

$$\begin{aligned} \dim B[\text{Ker}(AB)] &= \tau - \sigma = n - \text{rg}(AB) \\ -n + \text{rg}(B) &= \text{rg}(B) - \text{rg}(AB). \end{aligned}$$

□

Teorema 3.1.1 (Desigualdad de Frobenius). *Sean $A \in \mathbb{F}^{m \times n}, B \in \mathbb{F}^{n \times p}, C \in \mathbb{F}^{p \times q}$ matrices cualesquiera. Entonces,*

$$\text{rg}(AB) + \text{rg}(BC) \leq \text{rg}(B) + \text{rg}(ABC).$$

DEMOSTRACIÓN. Veamos que en el espacio vectorial $\mathbb{F}^{n \times 1}$ se tiene la relación de contenido

$$BC[\text{Ker}(ABC)] \subset B[\text{Ker}(AB)]. \quad (3.1)$$

En efecto, sea $\mathbf{y} \in BC[\text{Ker}(ABC)]$, entonces existe un $\mathbf{x} \in \text{Ker}(ABC)$ tal que $\mathbf{y} = BC\mathbf{x}$. Luego $\mathbf{y} = B(C\mathbf{x})$ con $AB(C\mathbf{x}) = 0$; de donde $\mathbf{y} \in B[\text{Ker}(AB)]$.

De (3.1) se sigue que

$$\dim BC[\text{Ker}(ABC)] \leq \dim B[\text{Ker}(AB)].$$

Aplicando el Lema 3.1.1, primero a A y BC , después a A y B , se tiene que

$$\text{rg}(BC) - \text{rg}(ABC) \leq \text{rg}(B) - \text{rg}(AB).$$

□

Si hacemos $C = 0$ en el Teorema 3.1.1, se tiene que

$$\text{rg}(AB) \leq \text{rg}(B).$$

Si hacemos $A = 0$ en el Teorema 3.1.1, queda

$$\text{rg}(BC) \leq \text{rg}(B).$$

Si ponemos $n = p$ y $B = I_n$ en el Teorema 3.1.1, obtenemos que

$$\text{rg}(A) + \text{rg}(C) \leq n + \text{rg}(AC),$$

o bien,

$$\text{rg}(A) + \text{rg}(C) - n \leq \text{rg}(AC).$$

Por lo tanto, hemos demostrado el corolario siguiente.

Corolario 1 (Ley de Sylvester de la Nulidad). *Sean $A \in \mathbb{F}^{m \times n}$ y $B \in \mathbb{F}^{n \times p}$ matrices cualesquiera. Entonces*

$$\text{rg}(A) + \text{rg}(B) - n \leq \text{rg}(AB) \leq \min\{\text{rg}(A), \text{rg}(B)\}.$$

Estas desigualdades también pueden deducirse directamente sin recurrir a la desigualdad de Frobenius.

3.2. Característica de Segré

Sea $A \in \mathbb{C}^{n \times n}$. Si en la matriz polinómica $\lambda I_n - A \in \mathbb{C}[\lambda]^{n \times n}$ efectuamos transformaciones elementales (ver la Sección A.2 del Apéndice A) para reducirla a forma diagonal, siempre podemos llegar a una matriz de la forma

$$S(\lambda) := \begin{pmatrix} f_1(\lambda) & 0 & \cdots & 0 \\ 0 & f_2(\lambda) & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & f_n(\lambda) \end{pmatrix}$$

donde $f_i(\lambda) \mid f_{i+1}(\lambda)$ para $i = 1, 2, \dots, n-1$. Aquí \mid significa “divide a”.

Este es un resultado de Teoría de Matrices bien conocido; la matriz $S(\lambda)$ se llama la forma normal de Smith de $\lambda I_n - A$. Los polinomios

$$f_1(\lambda), f_2(\lambda), \dots, f_n(\lambda)$$

se llaman los factores invariantes de $\lambda I_n - A$ (o de A). Véase la Sección A.3 del Apéndice A.

Sean $\lambda_1, \dots, \lambda_u \in \mathbb{C}$ los distintos valores propios de A . Entonces es posible descomponer cada factor invariante $f_i(\lambda) \neq 1$ en producto de potencias de la forma

$$(\lambda - \lambda_i)^{n_{ij}} \quad \text{con } n_{ij} \text{ entero } > 0.$$

Como puede verse en el Apéndice A (Secciones A.4 y A.6), todas las potencias de esta clase, que aparecen al descomponer todos los factores invariantes de A distintos de 1, se llaman los *divisores elementales* de A (o de $\lambda I_n - A$). Estos divisores son los polinomios

$$(\lambda - \lambda_i)^{n_{ij}}, \quad j = 1, \dots, r_i,$$

con $n_{i1} \geq n_{i2} \geq \dots \geq n_{ir_i}$ para $i = 1, \dots, u$. Sea $m(\lambda_i)$ o $m(\lambda_i, A)$ la *multiplicidad algebraica* del valor propio λ_i de la matriz A ; es decir, $m(\lambda_i)$ es la multiplicidad de λ_i como raíz del polinomio característico $\det(\lambda I_n - A)$. La partición $(n_{i1}, n_{i2}, \dots, n_{ir_i}, 0, \dots)$ de $m(\lambda_i)$ será llamada *partición de Segré* de λ_i , respecto de A , y será denotada por

$$s(\lambda_i, A).$$

Los números enteros $n_{i1}, n_{i2}, \dots, n_{ir_i}$ se llaman las *multiplicidades parciales* de λ_i . Al número r_i se le llama la *multiplicidad geométrica* de λ_i , y se denota por $\text{mg}(\lambda_i, A)$.

Supongamos ahora que los valores propios $\lambda_1, \lambda_2, \dots, \lambda_u$ de A están ordenados de manera que $m(\lambda_1) \geq \dots \geq m(\lambda_u)$, y, además, si

$$\dots \geq m(\lambda_{i-1}) > m(\lambda_i) = m(\lambda_{i+1}) = \dots = m(\lambda_{i+k-1}) > m(\lambda_{i+k}) \geq \dots$$

entonces se tenga que

$$s(\lambda_i, A) \stackrel{\text{I}}{\leq} s(\lambda_{i+1}, A) \stackrel{\text{I}}{\leq} \dots \stackrel{\text{I}}{\leq} s(\lambda_{i+k-1}, A).$$

Bajo estas condiciones, llamaremos *característica de Segré* de A a la sucesión finita de particiones

$$(s(\lambda_1, A), \dots, s(\lambda_u, A))$$

y la denotaremos por $s(A)$. Esta característica también alude a una estructura de Jordan de orden n . Más concretamente se llama *estructura de Jordan de orden n* a cualquier característica de Segré \mathcal{E} de una matriz cuadrada compleja de orden n ; esto es, \mathcal{E} es igual a una secuencia finita de particiones

$$[s_1, s_2, \dots, s_p]$$

tal que:

(1)

$$\sum_{i=1}^p |s_i| = n,$$

(2)

$$|s_1| \geq |s_2| \geq \cdots \geq |s_p|,$$

(3) cada vez que \geq sea = en (2), digamos que

$$|s_{i-1}| > |s_i| = |s_{i+1}| = \cdots = |s_{i+k-1}| > |s_{i+k}|,$$

se tenga que

$$s_i \stackrel{I}{\leq} s_{i+1} \stackrel{I}{\leq} \cdots \stackrel{I}{\leq} s_{i+k-1}.$$

Por la Proposición A.5.1 de la página 152 en el Apéndice A se tiene que el sistema de divisores elementales de $A \in \mathbb{C}^{n \times n}$ es un sistema completo de invariantes para la relación de semejanza de matrices cuadradas complejas de orden n . Dado que hay que considerar los divisores elementales que estén repetidos tantas veces como aparezcan en diferentes factores invariantes, debe hablarse de *sistema* y no de *conjunto* de divisores elementales. Por lo que precede, otro sistema completo de invariantes de $A \in \mathbb{C}^{n \times n}$ para la semejanza está formado por el *espectro ordenado* (o conjunto de valores propios) $\Sigma(A) = (\lambda_1, \lambda_2, \dots, \lambda_u)$ y la característica de Segré; es decir, que un sistema completo de invariantes de A para la semejanza es $(\Sigma(A), s(A))$.

3.3. Nulidades de potencias

La *nulidad* de una matriz $M \in \mathbb{C}^{m \times n}$ se define por

$$\nu(M) := n - \text{rg}(M).$$

La nulidad de M es la dimensión del núcleo de M , $\text{Ker}(M) := \{x \in \mathbb{C}^{n \times 1} \mid Mx = 0\}$, $\nu(M) = \dim \text{Ker}(M)$. Es evidente que

$$\text{Ker}(M) \subset \text{Ker}(M^2) \subset \cdots \subset \text{Ker}(M^i) \subset \cdots \quad (3.1)$$

Como estos subespacios vectoriales están contenidos en el espacio de dimensión finita $\mathbb{C}^n := \mathbb{C}^{n \times 1}$, la cadena ascendente (3.1) se debe estacionar; es decir, debe existir un entero $p > 0$ tal que

$$\text{Ker}(M^i) = \text{Ker}(M^p) \quad \text{para todo } i \geq p.$$

Además, se puede demostrar que si $k > 0$ es un entero tal que

$$\text{Ker}(M^k) = \text{Ker}(M^{k+1}),$$

entonces

$$\text{Ker}(M^{k+1}) = \text{Ker}(M^{k+2}) = \text{Ker}(M^{k+3}) = \cdots.$$

Por lo tanto, existe un entero positivo mínimo s tal que

$$\text{Ker}(M) \subsetneq \text{Ker}(M^2) \subsetneq \cdots \subsetneq \text{Ker}(M^s) = \text{Ker}(M^{s+1}) = \cdots$$

y, por consiguiente,

$$\nu(M) < \nu(M^2) < \cdots < \nu(M^s) = \nu(M^{s+1}) = \cdots.$$

Llamando $M^0 := I_n$, se tiene que

$$\nu(M^i) \leq \nu(M^{i+1}), \quad \text{para todo } i = 0, 1, 2, \dots \quad (3.2)$$

Recordemos que el peso de la partición $w(\lambda_j, A)$, denotado por $|w(\lambda_j, A)|$, es la suma de sus partes (ver la definición en la página 18):

$$|w(\lambda_j, A)| = m_{j1} + m_{j2} + \cdots + m_{js_j}.$$

Supongamos ahora que los valores propios $\lambda_1, \lambda_2, \dots, \lambda_u$ de A están ordenados de manera que $|w(\lambda_1, A)| \geq \cdots \geq |w(\lambda_u, A)|$, y, además, si

$$\cdots \geq |w(\lambda_{i-1}, A)| > |w(\lambda_i, A)| = |w(\lambda_{i+1}, A)| = \cdots = |w(\lambda_{i+k-1}, A)| > |w(\lambda_{i+k}, A)| \geq \cdots$$

entonces se tenga que

$$w(\lambda_i, A) \stackrel{T}{\geq} w(\lambda_{i+1}, A) \stackrel{T}{\geq} \cdots \stackrel{T}{\geq} w(\lambda_{i+k-1}, A).$$

Bajo estas condiciones, llamaremos *característica de Weyr* de A a la sucesión finita de particiones

$$(w(\lambda_1, A), \dots, w(\lambda_u, A))$$

y la denotaremos por $w(A)$.

3.4.1. Definición alternativa de la característica de Weyr

Siguiendo a Lippert [42] vamos a dar una definición de la característica de Weyr que evita la ordenación $\stackrel{T}{\geq}$ en el conjunto de particiones. Sea \mathcal{P} el conjunto de las particiones. Definamos los conceptos

$$|a| := \sum_{i=1}^{\infty} a_i, \quad a \cdot b := \sum_{i=1}^{\infty} a_i b_i, \quad \|a\|^2 := a \cdot a, \quad a, b \in \mathcal{P}.$$

Sea \mathbb{W} el conjunto de funciones de \mathbb{C} en \mathcal{P} que son nulas en casi todo elemento de \mathbb{C} .

$$\mathbb{W} := \{\mu: \mathbb{C} \rightarrow \mathcal{P}, \mu(\lambda) = (0, 0, \dots) \forall \lambda \in \mathbb{C}, \text{ excepto subconjunto finito de } \lambda\};$$

Es decir, $\mu \in \mathbb{W}$ si existe un subconjunto $F \subset \mathbb{C}$ finito (puede ser vacío) tal que para cada $\lambda \in \mathbb{C} \setminus F$, $\mu(\lambda)$ es la partición cero. Por razones de coherencia notacional llamaremos *espectro* (o *soporte*) de μ al conjunto

$$\Lambda(\mu) := \{\lambda \in \mathbb{C}: \mu(\lambda) \neq (0, 0, \dots)\}.$$

Para cada $\lambda \in \mathbb{C}$ sea $\mu(\lambda) = (\mu(\lambda)_1, \mu(\lambda)_2, \dots) \in \mathcal{P}$.

$$|\mu| := \sum_{\lambda \in \mathbb{C}} |\mu(\lambda)|, \quad \mu \cdot \tilde{\mu} := \sum_{\lambda \in \mathbb{C}} \mu(\lambda) \cdot \tilde{\mu}(\lambda), \quad \|\mu\|^2 := \mu \cdot \mu, \quad \mu, \tilde{\mu} \in \mathbb{W}.$$

De esta manera hemos definido una suerte de producto escalar en el conjunto \mathbb{W} .

Mayorización débil en \mathbb{W}

Para $\mu, \tilde{\mu} \in \mathbb{W}$, definimos $\mu \prec \tilde{\mu}$ si $\mu(\lambda) \prec \tilde{\mu}(\lambda), \forall \lambda \in \mathbb{C}$.

Dada $A \in \mathbb{C}^{n \times n}$ se llama *característica de Weyr* de A al elemento de \mathbb{W} , $w(A)$, definido así:

$$w(A)(\lambda) := \begin{cases} w(\lambda, A) & \text{si } \lambda \in \Lambda(A), \\ (0, 0, \dots) & \text{si } \lambda \notin \Lambda(A). \end{cases}$$

Así pues, con este punto de vista, la característica de Weyr de una matriz $A \in \mathbb{C}^{n \times n}$ es una *función* de \mathbb{C} en \mathcal{P} , nula excepto en los valores propios de A .

3.5. Relación entre las características de Segré y Weyr

Para cada valor propio λ_j de la matriz $A \in \mathbb{C}^{n \times n}$ sean

$$s(\lambda_j, A) = (n_{j1}, n_{j2}, \dots, n_{jr_j}, 0, \dots)$$

y

$$w(\lambda_j, A) = (m_{j1}, m_{j2}, \dots, m_{js_j}, 0, \dots)$$

las particiones de Segré y de Weyr de λ_j , $j = 1, \dots, u$.

¿Cuánto suman las partes de la partición $s(\lambda_j, A)$? La suma $n_{j1} + n_{j2} + \dots + n_{jr_j}$ es igual a la multiplicidad algebraica del valor propio λ_j de A :

$$n_{j1} + n_{j2} + \dots + n_{jr_j} = m(\lambda_j, A). \quad (3.1)$$

Esta igualdad puede argumentarse de varias maneras; he aquí dos de ellas: **1ª.**- Es así porque el producto de los factores invariantes de A

$$f_1(\lambda)f_2(\lambda) \cdots f_n(\lambda)$$

es igual al polinomio característico de A ; **2ª.**- El valor propio λ_j aparece $n_{j1} + n_{j2} + \dots + n_{jr_j}$ veces en la diagonal de la matriz J , forma de Jordan de A , y J es una matriz triangular superior. Véase el Teorema A.6.1 de la página 154. Esto es cierto para cada $j = 1, \dots, u$.

¿Cuánto suman las partes de la partición $w(\lambda_j, A)$? La suma $m_{j1} + m_{j2} + \dots + m_{js_j}$ es igual a $m(\lambda_j, A)$, la multiplicidad algebraica de λ_j . Para demostrar esto nos apoyamos en algunos resultados previos.

Teorema 3.5.1. Sean $A \in \mathbb{C}^{n \times n}$ y $\Lambda(A) = \{\lambda_1, \lambda_2, \dots, \lambda_u\}$. Entonces, para cada $j = 1, \dots, u$, la multiplicidad geométrica de λ_j coincide con la dimensión $\dim \text{Ker}(A - \lambda_j I)$, y la multiplicidad algebraica de λ_j coincide con la dimensión $\dim \text{Ker}[(A - \lambda_j I)^{s_j}]$ ($= \dim \text{Ker}[(A - \lambda_j I)^n]$).

La demostración de este teorema se sigue fácilmente del Teorema 3.5.2. Si el lector desea una demostración directa del Teorema 3.5.1 puede consultar la prueba de la Proposition 2.2.5, pág. 56, de [31].

Así pues,

$$\begin{aligned} \text{mult. geométrica} & \quad \text{mg}(\lambda_j, A) = \nu[(A - \lambda_j I)] \\ \text{mult. algebraica} & \quad m(\lambda_j, A) = \nu[(A - \lambda_j I)^{s_j}]. \end{aligned}$$

Por lo tanto,

$$\begin{aligned} m_{j1} + m_{j2} + \dots + m_{js_j} &= \nu[(A - \lambda_j I)] + \nu[(A - \lambda_j I)^2] - \nu[(A - \lambda_j I)] \\ &+ \dots + \nu[(A - \lambda_j I)^{s_j}] - \nu[(A - \lambda_j I)^{s_j-1}] = \nu[(A - \lambda_j I)^{s_j}] = m(\lambda_j, A). \end{aligned}$$

Acabamos de establecer que

$$|s(\lambda_j, A)| = m(\lambda_j, A) = |w(\lambda_j, A)|. \quad (3.2)$$

¿Es ésta la única relación entre las particiones de Segré y de Weyr de λ_j ? A continuación veremos que la relación entre ambas particiones es mucho más profunda; en efecto, dichas particiones son conjugadas.

Teorema 3.5.2. Sea $A \in \mathbb{C}^{n \times n}$ y α un valor propio de A . Sean

$$s(\alpha, A) = (n_1, n_2, \dots, n_r, 0, \dots)$$

la partición de Segré de α y

$$w(\alpha, A) = (m_1, m_2, \dots, m_s, 0, \dots)$$

la partición de Weyr de α .

Entonces $w(\alpha, A)$ es la partición conjugada de $s(\alpha, A)$.

el número de columnas de ceros de N^2 menos el número de columnas de ceros de N es igual al número de n_j que son ≥ 2 ($j = 1, \dots, r$). Pues si $n_j = 1$, N_j^2 no aumenta ninguna columna de ceros. Así pues,

$$\begin{aligned} m_2 &= \nu(N^2) - \nu(N) = \text{número de columnas de ceros de } N^2 \\ &\quad - \text{número de columnas de ceros de } N \\ &= \text{número de } n_j \geq 2, \quad j \in \{1, \dots, r\}. \end{aligned}$$

En general,

$$m_i = \nu(N^i) - \nu(N^{i-1});$$

- Si $n_j < i$, entonces $N_j^i = 0 = N_j^{i-1}$.
- Si $n_j \geq i$, entonces

$$N_j^i = \begin{pmatrix} 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ \vdots & & & & \ddots & & \vdots \\ & & & & & \ddots & \\ & & & & & & 1 \\ \vdots & & & & & & \vdots \\ 0 & \cdots & & & \cdots & & 0 \end{pmatrix}$$

y esta matriz tiene todos sus elementos iguales a 0, excepto los de los lugares $(1, i)$, $(2, i+1)$, $(3, i+2), \dots, (n_j - i + 1, n_j)$ que valen 1. En particular, si $n_j = i$,

$$N_j^i = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \quad \text{y} \quad N_j^{i-1} = \begin{pmatrix} 0 & \cdots & 1 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

Así pues, la nulidad de N^i aumenta respecto de la de N^{i-1} en tantas unidades como términos n_j haya $\geq i$, $j \in \{1, \dots, r\}$. Luego,

$$m_i = \text{número de } n_j \geq i, \quad j \in \{1, \dots, r\}. \quad (3.3)$$

Es fácil demostrar (3.3) por inducción y no lo haremos explícitamente. La fórmula (3.3) es válida para $i = 1, 2, \dots, n_1$; pues $n_1 \geq n_2 \geq \dots \geq n_r$ implica que no hay ningún n_j mayor o igual que $n_1 + 1$. Por otra parte, $m_s = \nu(N^s) - \nu(N^{s-1}) > 0$ y $\nu(N^s) = \nu(N^{s+1})$; de esto se infiere que $s = n_1$.

Por consiguiente, para todo $i \in \mathbb{N}^* := \{1, 2, 3, \dots\}$, se tiene que

$$m_i = \text{Card}\{j : n_j \geq i\}$$

(habiendo considerado, por convenio, que $n_j := 0$ si $j > r$, y $m_i := 0$ si $i > s$). Por lo tanto,

$$(m_1, m_2, \dots) = \overline{(n_1, n_2, \dots)}.$$

En consecuencia, las longitudes de estas particiones son respectivamente

$$\begin{aligned} \ell((m_1, m_2, \dots)) &= n_1 = s, \\ \ell((n_1, n_2, \dots)) &= m_1 = r. \end{aligned} \quad (3.4)$$

□

Observación 3.5.1. Las particiones $s(\alpha, A)$ y $w(\alpha, A)$ también serán denotadas así

$$s(\alpha, A) =: (s(\alpha, A)_1, s(\alpha, A)_2, \dots)$$

y

$$w(\alpha, A) =: (w(\alpha, A)_1, w(\alpha, A)_2, \dots),$$

con

$$\begin{aligned} s(\alpha, A)_1 &\geq s(\alpha, A)_2 \geq \dots, \\ w(\alpha, A)_1 &\geq w(\alpha, A)_2 \geq \dots; \end{aligned}$$

así pues,

$$\ell(s(\alpha, A)) = w(\alpha, A)_1, \quad \ell(w(\alpha, A)) = s(\alpha, A)_1.$$

Ejemplo 3.5.1. Sea $A \in \mathbb{C}^{20 \times 20}$ con dos valores propios distintos λ_1, λ_2 y característica de Segré

$$s(A) = ((5, 4, 2, 2, 0, \dots), (3, 2, 2, 0, \dots))$$

siendo

$$\begin{aligned} s(\lambda_1, A) &= (5, 4, 2, 2, 0, \dots), \\ s(\lambda_2, A) &= (3, 2, 2, 0, \dots). \end{aligned}$$

Vamos a hallar su característica de Weyr utilizando el Teorema 3.5.2. Hacemos los diagramas de Young para hallar las particiones conjugadas de $s(\lambda_1, A)$ y $s(\lambda_2, A)$. Véase la figura 3.1.

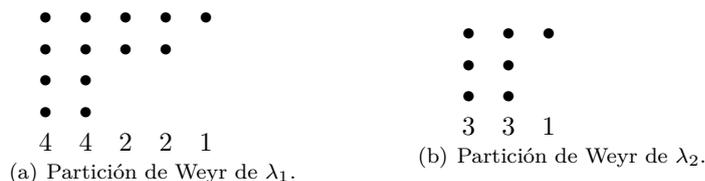


Figura 3.1: Característica de Weyr de A .

Así pues, por el Teorema 3.5.2,

$$\begin{aligned} w(\lambda_1, A) &= (4, 4, 2, 2, 1, 0, \dots), \\ w(\lambda_2, A) &= (3, 3, 1, 0, \dots), \end{aligned}$$

y

$$w(A) = ((4, 4, 2, 1, 1, 0, \dots), (3, 3, 1, 0, \dots)).$$

Antes de finalizar esta sección deseamos hacer notar que la Proposición 2.5.1 de la Sección 2.5 del Capítulo 2 sobre la relación entre los órdenes lexicográficos inverso y transpuesto de las particiones de un entero positivo, implica que para toda matriz $A \in \mathbb{C}^{n \times n}$ cuyos valores propios $\lambda_1, \lambda_2, \dots, \lambda_u$ estén ordenados de modo que $s(A) = (s(\lambda_1, A), \dots, s(\lambda_u, A))$ se tiene que

$$w(A) = (\overline{s(\lambda_1, A)}, \dots, \overline{s(\lambda_u, A)}). \tag{3.5}$$

3.6. Forma canónica dual

Los *parámetros* que describen la forma de Jordan, J , de una matriz $A \in \mathbb{C}^{n \times n}$ son el espectro ordenado $\Sigma(A)$ y la característica de Segré $s(A)$. Véase el Teorema A.6.1 de la página 154, Apéndice A, y el párrafo correspondiente de la página 50. El espectro $\Sigma(A)$ nos

Lema 3.6.1. Si la matriz $A \in \mathbb{C}^{n \times n}$ sólo tiene un valor propio α , con partición de Weyr $w(\alpha, A) = (m_1, m_2, \dots, m_s, 0, \dots)$, donde $m_1 \geq m_2 \geq \dots \geq m_s > 0$, entonces A es semejante a la matriz de orden n

$$W := \begin{pmatrix} \alpha I_{m_s} & E_{m_s} & 0 & \cdots & 0 & 0 \\ 0 & \alpha I_{m_{s-1}} & E_{m_{s-1}} & & & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & & & & \ddots & \\ 0 & 0 & & & \alpha I_{m_2} & E_{m_2} \\ 0 & 0 & & \cdots & 0 & \alpha I_{m_1} \end{pmatrix}$$

donde

$$E_{m_j} := [I_{m_j} \ 0] \in \mathbb{C}^{m_j \times m_{j-1}}, \quad j = 2, 3, \dots, s.$$

DEMOSTRACIÓN. Obviamente, α es también el único valor propio de W . Veamos que W tiene la misma característica de Weyr que A . ambos hechos probarán que W es semejante a A .

Denotemos $w(\alpha, W) = (p_1, p_2, \dots, p_t, 0, \dots)$, con $p_1 \geq p_2 \geq \dots \geq p_t > 0$, la partición de Weyr de α respecto de W . De la relación entre rango y nulidad deducimos que para cada $i = 1, 2, \dots, t$,

$$p_i = \text{rg} [(W - \alpha I)^{i-1}] - \text{rg} [(W - \alpha I)^i].$$

La matriz $W - \alpha I$ es igual a

$$\begin{pmatrix} 0 & E_{m_s} & 0 & \cdots & 0 & 0 \\ 0 & 0 & E_{m_{s-1}} & & & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & & & & \ddots & \\ 0 & 0 & & & 0 & E_{m_2} \\ 0 & 0 & & \cdots & 0 & 0 \end{pmatrix}.$$

Por consiguiente, $\text{rg}(W - \alpha I) = m_s + m_{s-1} + \dots + m_2$. Calculemos el cuadrado de la matriz $W - \alpha I$.

$$(W - \alpha I)^2 = \begin{pmatrix} 0 & 0 & E_{m_s} E_{m_{s-1}} & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & E_{m_{s-1}} E_{m_{s-2}} & & & 0 \\ \vdots & & & & \ddots & & \vdots \\ 0 & & & & & \ddots & \\ 0 & 0 & & & & 0 & E_{m_3} E_{m_2} \\ 0 & 0 & & \cdots & 0 & 0 & 0 \end{pmatrix}.$$

Para $j = 3, 4, \dots, s$, tenemos que

$$E_{m_j} E_{m_{j-1}} = [I_{m_j} \ 0] \begin{bmatrix} I_{m_j} & 0 \\ 0 & I_{m_{j-1}-m_j} \end{bmatrix} = [I_{m_j} \ 0 \ 0],$$

luego $\text{rg}(E_{m_j} E_{m_{j-1}}) = m_j$. Por lo tanto,

$$\text{rg} [(W - \alpha I)^2] = m_s + m_{s-1} + \dots + m_3.$$

De lo obtenido hasta ahora, deducimos

$$\begin{aligned} p_1 &= \text{rg} [(W - \alpha I)^0] - \text{rg} [(W - \alpha I)] = \text{rg} I - \text{rg} [(W - \alpha I)] \\ &= n - (m_s + \dots + m_2) = m_1, \quad (\text{pues } m_1 + m_2 + \dots + m_s = n), \\ p_2 &= \text{rg} [(W - \alpha I)] - \text{rg} [(W - \alpha I)^2] \\ &= m_s + \dots + m_2 - (m_s + \dots + m_3) = m_2. \end{aligned}$$

Avancemos un poco más,

$$\begin{aligned}
 (W - \alpha I)^3 &= (W - \alpha I)^2(W - \alpha I) = \\
 &= \begin{pmatrix} 0 & 0 & E_{m_s}E_{m_{s-1}} & \cdots & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ & & & & \ddots & \\ & & & & & E_{m_3}E_{m_2} \\ 0 & & & & & 0 \\ 0 & & & & & 0 \end{pmatrix} \begin{pmatrix} 0 & E_{m_s} & 0 & \cdots & \cdots & 0 \\ \vdots & & \ddots & & & \vdots \\ & & & \ddots & & \\ & & & & E_{m_3} & 0 \\ 0 & & & & 0 & E_{m_2} \\ 0 & & & & & 0 \end{pmatrix} \\
 &= \begin{pmatrix} 0 & 0 & 0 & E_{m_s}E_{m_{s-1}}E_{m_{s-2}} & \cdots & \cdots & 0 \\ \vdots & & & & \ddots & & \vdots \\ & & & & & \ddots & \\ & & & & & & E_{m_4}E_{m_3}E_{m_2} \\ 0 & & & & & & 0 \\ 0 & & & & & & 0 \\ 0 & & & & & & 0 \end{pmatrix}
 \end{aligned}$$

Para $j = 4, 5, \dots, s$, tenemos que

$$\begin{aligned}
 (E_{m_j}E_{m_{j-1}})E_{m_{j-2}} &= [I_{m_j} \quad 0] [I_{m_{j-2}} \quad 0] = \\
 [I_{m_j} \quad 0] \begin{bmatrix} I_{m_j} & 0 \\ 0 & I_{m_{j-2}-m_j} \end{bmatrix} &= [I_{m_j} \quad 0 \quad 0] \in \mathbb{C}^{m_j \times m_{j-3}};
 \end{aligned}$$

luego

$$\operatorname{rg}(E_{m_j}E_{m_{j-1}}E_{m_{j-2}}) = m_j.$$

Por lo tanto,

$$\operatorname{rg}[(W - \alpha I)^3] = m_s + m_{s-1} + \cdots + m_4.$$

De aquí que

$$\begin{aligned}
 p_3 &= \operatorname{rg}[(W - \alpha I)^2] - \operatorname{rg}[(W - \alpha I)^3] \\
 &= (m_s + \cdots + m_3) - (m_s + \cdots + m_4) = m_3.
 \end{aligned}$$

Por inducción sobre i puede demostrarse que

$$(W - \alpha I)^i = \begin{pmatrix} 0 & \cdots & E_{m_s}E_{m_{s-1}} \cdots E_{m_{s-(i-1)}} & & 0 \\ \vdots & & & \ddots & \\ & & & & E_{m_{i+1}} \cdots E_{m_3}E_{m_2} \\ & & & & \vdots \\ 0 & & & & 0 \end{pmatrix}.$$

También puede probarse por inducción sobre i que para todo $j = i + 1, i + 2, \dots, s$,

$$E_{m_j}E_{m_{j-1}} \cdots E_{m_{j-(i-1)}} = [I_{m_j} \quad 0];$$

por lo tanto,

$$\operatorname{rg}(E_{m_j}E_{m_{j-1}} \cdots E_{m_{j-(i-1)}}) = m_j.$$

De donde,

$$\operatorname{rg}[(W - \alpha I)^i] = m_s + m_{s-1} + \cdots + m_{i+1}.$$

Así pues,

$$\begin{aligned} p_i &= \operatorname{rg} [(W - \alpha I)^{i-1}] - \operatorname{rg} [(W - \alpha I)^i] \\ &= (m_s + m_{s-1} + \cdots + m_{i+1} + m_i) - (m_s + m_{s-1} + \cdots + m_{i+1}) = m_i. \end{aligned}$$

Veremos a continuación la prueba de que $t = s$.

Como α es el único valor propio de A se tiene que

$$\begin{aligned} \operatorname{Ker}(A - \alpha I) &\subsetneq \operatorname{Ker} [(A - \alpha I)^2] \subsetneq \cdots \subsetneq \operatorname{Ker} [(A - \alpha I)^s] \\ &= \operatorname{Ker} [(A - \alpha I)^{s+1}] = \cdots. \end{aligned}$$

Luego

$$\operatorname{Ker} [(A - \alpha I)^s] = \mathbb{C}^n. \quad (3.1)$$

La identidad (3.1) es debida al hecho de que si una matriz $B \in \mathbb{C}^{n \times n}$ tiene los valores propios distintos $\alpha_1, \alpha_2, \dots, \alpha_u$ y para cada $j = 1, \dots, u$ llamamos s_j al entero positivo que satisface

$$\begin{aligned} \operatorname{Ker}(B - \alpha_j I) &\subsetneq \operatorname{Ker} [(B - \alpha_j I)^2] \subsetneq \cdots \subsetneq \operatorname{Ker} [(B - \alpha_j I)^{s_j}] \\ &= \operatorname{Ker} [(B - \alpha_j I)^{s_j+1}] = \cdots, \end{aligned}$$

entonces

$$\mathbb{C}^n = \operatorname{Ker} [(B - \alpha_1 I)^{s_1}] \oplus \cdots \oplus \operatorname{Ker} [(B - \alpha_u I)^{s_u}], \quad (3.2)$$

donde \oplus denota *suma directa* de subespacios. Una demostración de (3.2) puede verse en el libro de Gohberg, Lancaster y Rodman [31], Theorem 2.1.2 en la página 47.

Por (3.1), $\nu[(A - \alpha I)^s] = n$, y, en consecuencia, $\operatorname{rg} [(A - \alpha I)^s] = 0$, lo que equivale a $(A - \alpha I)^s = 0$. Además, $(A - \alpha I)^{s-1} \neq 0$.

Por otro lado, t es el mínimo entero positivo k tal que

$$\operatorname{rg} [(W - \alpha I)^k] = \operatorname{rg} [(W - \alpha I)^{k+1}];$$

lo que equivale a decir que t es el mínimo entero positivo k para el que $(W - \alpha I)^k = 0$, pues si para un $k > 0$ fuera $(W - \alpha I)^k \neq 0$, se tendría que

$$\operatorname{rg} [(W - \alpha I)^k] > \operatorname{rg} [(W - \alpha I)^{k+1}],$$

en virtud de la forma que tienen las potencias de $W - \alpha I$.

Las matrices $W - \alpha I, (W - \alpha I)^2, (W - \alpha I)^3, \dots$ son matrices de *tamaño $s \times s$ de bloques*. Es fácil observar que $(W - \alpha I)^{s-1}$ tiene iguales a cero todos sus s^2 bloques excepto el del lugar $(1, s)$; es decir, el situado en la esquina superior derecha. Este bloque es igual a $E_{m_s} E_{m_{s-1}} \cdots E_{m_{s-(s-2)}}$:

$$(W - \alpha I)^{s-1} = \begin{pmatrix} 0 & \cdots & \cdots & 0 & E_{m_s} E_{m_{s-1}} \cdots E_{m_2} \\ 0 & & & & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix};$$

por lo tanto, $(W - \alpha I)^s = 0$. De aquí que $t = s$. Como $p_i = m_i$ para todo $i = 1, \dots, s$, se sigue la igualdad

$$w(\alpha, W) = w(\alpha, A),$$

entre las particiones de Weyr de α respecto de las matrices W y A . Por consiguiente, W es semejante a A . \square

Definición 3.6.1. Llamaremos *forma canónica dual* de la matriz $A \in \mathbb{C}^{n \times n}$, asociada a su característica de Weyr, a la matriz W del Teorema 3.6.1.

Esta forma canónica se puede “relajar” y admitir otras ordenaciones distintas de los bloques diagonales W_1, W_2, \dots, W_u ; se siguen obteniendo matrices semejantes a A .

3.7. Rangos y particiones de Weyr

Sean $A \in \mathbb{C}^{n \times n}$ y α un valor propio cualquiera de A . Sea $w(\alpha, A) = (m_1, m_2, \dots, m_s, 0, \dots)$, con $m_1 \geq m_2 \geq \dots \geq m_s > 0$, la partición de Weyr de α , respecto de A . Por la definición dada en la página 51, sabemos que una vez conocido α los números m_1, m_2, \dots, m_s , pueden obtenerse mediante el cálculo del rango de las matriz $A - \alpha I$ y sus potencias sucesivas. En efecto,

$$m_k = \text{rg} [(A - \alpha I)^{k-1}] - \text{rg} [(A - \alpha I)^k], \quad (3.1)$$

para $k = 1, \dots, s$. Vamos a ver en esta sección una manera distinta de hallar los números m_1, m_2, \dots, m_s mediante el cálculo de otras matrices relacionadas con α y A . Para cada $k = 1, \dots, s$, sea $P_{\alpha, k}(A)$ la matriz $kn \times kn$ dada por

$$P_{\alpha, k}(A) := \begin{pmatrix} \alpha I - A & 0 & \dots & 0 \\ I & \alpha I - A & & \vdots \\ & I & \alpha I - A & \\ & & \ddots & \ddots \\ 0 & & \dots & \alpha I - A & 0 \\ & & & I & \alpha I - A \end{pmatrix}. \quad (3.2)$$

Por ejemplo, si $k = 3$, se tiene que

$$P_{\alpha, 3}(A) = \begin{pmatrix} \alpha I - A & 0 & 0 \\ I & \alpha I - A & 0 \\ 0 & I & \alpha I - A \end{pmatrix};$$

vamos a hacer transformaciones elementales por bloques por filas y columnas (véase la Sección A.8 de la página 156 del Apéndice A) para transformar esta matriz en la matriz

$$\begin{pmatrix} 0 & 0 & (\alpha I - A)^3 \\ I & 0 & 0 \\ 0 & I & 0 \end{pmatrix}. \quad (3.3)$$

En primer lugar, sumando en $P_{\alpha, 3}(A)$ a la primera fila de bloques la segunda fila de bloques premultiplicada por la matriz $-(\alpha I - A)$ resulta

$$\begin{pmatrix} 0 & -(\alpha I - A)^2 & 0 \\ I & \alpha I - A & 0 \\ 0 & I & \alpha I - A \end{pmatrix};$$

a continuación sumamos a la primera fila de bloques de esta matriz la tercera fila de bloques premultiplicada por $(\alpha I - A)^2$, obteniendo

$$\begin{pmatrix} 0 & 0 & (\alpha I - A)^3 \\ I & \alpha I - A & 0 \\ 0 & I & \alpha I - A \end{pmatrix};$$

ahora sumamos a la segunda columna de bloques la primera postmultiplicada por $-(\alpha I - A)$, lo que da

$$\begin{pmatrix} 0 & 0 & (\alpha I - A)^3 \\ I & 0 & 0 \\ 0 & I & \alpha I - A \end{pmatrix};$$

finalmente, sumamos a la tercera columna de bloques la segunda postmultiplicada por $-(\alpha I - A)$:

$$\begin{pmatrix} 0 & 0 & (\alpha I - A)^3 \\ I & 0 & 0 \\ 0 & I & 0 \end{pmatrix}.$$

En general, haciendo transformaciones elementales por bloques a partir de la matriz $P_{\alpha,k}(A)$ podemos obtener la matriz

$$C_{\alpha,k}(A) := \begin{pmatrix} 0 & \cdots & \cdots & (-1)^{k-1}(\alpha I - A)^k \\ I & 0 & & \vdots \\ & I & 0 & \\ & & \ddots & \ddots \\ 0 & & \cdots & I & 0 \end{pmatrix}; \quad (3.4)$$

esta matriz tiene, por tanto, el mismo rango que $P_{\alpha,k}(A)$. Observamos que $C_{\alpha,k}(A)$ tiene $k - 1$ bloques iguales a la matriz identidad I . Así pues,

$$\text{rg } C_{\alpha,k}(A) = (k - 1)n + \text{rg } [(\alpha I - A)^k];$$

y, consecuentemente,

$$\begin{aligned} \nu[C_{\alpha,k}(A)] &= kn - (k - 1)n - \text{rg } [(\alpha I - A)^k] \\ &= n - \text{rg } [(\alpha I - A)^k] = \nu[(\alpha I - A)^k]; \end{aligned}$$

por consiguiente,

$$\nu[P_{\alpha,k}(A)] = \nu[(\alpha I - A)^k].$$

Teniendo en cuenta, además, las igualdades (3.2) de la Sección 3.4, resulta que para todo $k = 1, \dots, s$ se tiene que

$$\nu[P_{\alpha,k}(A)] = \sum_{i=1}^k m_i. \quad (3.5)$$

Método práctico para hallar la característica de Segré

Supongamos que conocemos los valores propios de la matriz $A \in \mathbb{C}^{n \times n}$. Sea α uno cualquiera de estos valores propios. Calculamos los rangos de la matrices $\alpha I - A$, $(\alpha I - A)^2$, $(\alpha I - A)^3, \dots$. Por la cadena de desigualdades (3.1) de la Sección 3.4, estos rangos van decreciendo estrictamente hasta que su valor se estaciona, digamos para las matrices $(\alpha I - A)^s$ y $(\alpha I - A)^{s+1}$,

$$\begin{aligned} \text{rg } [\alpha I - A] &> \text{rg } [(\alpha I - A)^2] > \text{rg } [(\alpha I - A)^3] > \cdots > \text{rg } [(\alpha I - A)^s] \\ &= \text{rg } [(\alpha I - A)^{s+1}]; \end{aligned}$$

acto seguido, se forman las diferencias

$$\begin{aligned} m_1 &:= n - \text{rg } [\alpha I - A], \\ m_2 &:= \text{rg } [\alpha I - A] - \text{rg } [(\alpha I - A)^2], \\ &\vdots \\ m_s &:= \text{rg } [(\alpha I - A)^{s-1}] - \text{rg } [(\alpha I - A)^s]; \end{aligned}$$

se finaliza hallando la partición conjugada de la partición (m_1, m_2, \dots, m_s) y se obtiene la partición de Segré de α :

$$(n_1, n_2, \dots, n_r) = \overline{(m_1, m_2, \dots, m_s)}.$$

De aquí se sigue que en la diagonal por bloques de la forma de Jordan de A debemos poner los bloques de Jordan de tamaños n_1, n_2, \dots, n_r asociados al valor propio α . Y se hace un proceso análogo para cada uno de los valores propios de A .

Ejemplo 3.7.1. Sea la matriz

$$\begin{pmatrix} 3 & -1 & 2 & 19 \\ 0 & 3 & 1 & -20 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

Su polinomio característico $\det(\lambda I - A)$ es igual a

$$\begin{vmatrix} \lambda - 3 & 1 & -2 & -19 \\ 0 & \lambda - 3 & -1 & 20 \\ 0 & 0 & \lambda - 3 & 0 \\ 0 & 0 & 0 & \lambda + 2 \end{vmatrix} = (\lambda - 3)^3(\lambda + 2);$$

luego los valores propios de A son 3 y -2.

Apliquemos el método para hallar la partición de Segré de 3:

$$3I - A = \begin{pmatrix} 0 & 1 & -2 & -19 \\ 0 & 0 & -1 & 20 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}, \quad \text{rg}(3I - A) = 3;$$

$$(3I - A)^2 = \begin{pmatrix} 0 & 0 & -1 & -75 \\ 0 & 0 & 0 & 100 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 25 \end{pmatrix}, \quad \text{rg}[(3I - A)^2] = 2;$$

$$(3I - A)^3 = \begin{pmatrix} 0 & 0 & 0 & -375 \\ 0 & 0 & 0 & 500 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 125 \end{pmatrix}, \quad \text{rg}[(3I - A)^3] = 1;$$

$$(3I - A)^4 = \begin{pmatrix} 0 & 0 & 0 & -1875 \\ 0 & 0 & 0 & 2500 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 625 \end{pmatrix}, \quad \text{rg}[(3I - A)^4] = 1.$$

Por lo tanto, la partición de Weyr de 3 es

$$(4 - 3, 3 - 2, 2 - 1, 1 - 1, \dots) = (1, 1, 1, 0, \dots).$$

y la partición de Segré de 3 es la partición conjugada de $(1, 1, 1)$, que hallamos mediante su diagrama. Véase la figura 3.2. En consecuencia, $s(3, A) = (3)$. Por lo que en la forma de

$$\begin{array}{c} 1 \quad \bullet \\ 1 \quad \bullet \\ 1 \quad \bullet \\ 3 \end{array}$$

Figura 3.2: Particiones de 3.

Jordan de A hay un solo bloque de Jordan asociado al valor propio 3 y es de orden 3.

Hallemos a continuación, la partición de Segré de -2:

$$-2I - A = \begin{pmatrix} -5 & 1 & -2 & -19 \\ 0 & -5 & -1 & 20 \\ 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{rg}(-2I - A) = 3;$$

$$(-2I - A)^2 = \begin{pmatrix} 25 & -10 & 19 & 115 \\ 0 & 25 & 10 & -100 \\ 0 & 0 & 25 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{rg} [(3I - A)^2] = 3;$$

por consiguiente, la partición de Weyr de -2 es $(4 - 3, 3 - 3, \dots) = (1, 0, \dots)$ y, por ello, la partición de Segré de -2 es (1) según puede deducirse del diagrama en la figura 3.3.



Figura 3.3: Particiones de -2 .

Así pues, en la forma de Jordan de A hay un solo bloque de Jordan asociado al valor propio -2 y es de orden 1.

De todo lo dicho se deduce que la forma de Jordan de A es la matriz

$$J = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

Este método parte de suponer que se tiene el conocimiento *exacto* de los valores propios de A . Veremos en el Capítulo 7 lo que sobrevive de este procedimiento cuando no se disponga del espectro de A con exactitud y sólo se pueda presumir que sabemos quiénes son los elementos de A sin duda alguna. E incluso, se abordará allí lo que pueda decirse de la forma de Jordan de A cuando lo único conocido de A sea que en determinadas posiciones de la matriz hay ceros con toda seguridad.

3.8. Criterio de semejanza mediante rangos

Para comprobar que dos matrices $A, B \in \mathbb{C}^{n \times n}$ son semejantes hay varios procedimientos. Uno de ellos consistiría en calcular sus valores propios (hallando las raíces de su polinomio característico) y luego hallando su característica de Weyr (o de Segré) por cualquiera de los métodos de los rangos de la sección anterior; este procedimiento tiene el inconveniente de que el cálculo de los valores propios es un proceso no (necesariamente) racional, es decir, no siempre puede obtenerse a partir de los elementos de la matriz mediante un número finito de sumas, restas, multiplicaciones y divisiones de números complejos. Una vez conocidos los valores propios α , el proceso de cálculo de los rangos de matrices de la forma $(\alpha I - A)^k$ o $P_{\alpha,k}(A)$ llega a ser un proceso racional pues podrían deducirse los rangos mediante la comprobación de qué menores son iguales a cero o no.

Hay otro procedimiento, enteramente racional, para averiguar si A y B son semejantes que consiste en hallar los factores invariantes de las matrices polinómicas $\lambda I - A$ y $\lambda I - B$ mediante transformaciones elementales en el anillo de polinomios $\mathbb{C}[\lambda]$ (véase la página 146); si los factores invariantes de ambas matrices son iguales, entonces A y B son semejantes. Véase la Sección A.3 y el Teorema A.5.1, ambos del Apéndice A.

En esta sección veremos que hay un criterio racional para decidir si A y B son semejantes que elude tanto el cálculo de los valores propios como el de los factores invariantes. Este criterio se reduce a la determinación del rango de tres matrices complejas fácilmente calculables a partir de A y B . Precisa el conocimiento de la definición del producto de Kronecker o tensorial de dos matrices.

Sean $A \in \mathbb{C}^{m \times m}, B \in \mathbb{C}^{n \times n}$ matrices cuyos valores propios (distintos) son

$$\begin{array}{ll} \lambda_1, \dots, \lambda_s, \lambda_{s+1}, \dots, \lambda_u & \text{los de } A \\ \mu_1, \dots, \mu_s, \mu_{s+1}, \dots, \mu_v & \text{los de } B; \end{array}$$

siendo $\lambda_1 = \mu_1, \dots, \lambda_s = \mu_s$ y $\Lambda(A) \cap \Lambda(B) = \{\lambda_1, \dots, \lambda_s\}$. Sean

$$\begin{aligned} s(\lambda_k, A) &= (m_{k1}, m_{k2}, \dots) \\ s(\lambda_k, B) &= (n_{k1}, n_{k2}, \dots) \\ k &= 1, \dots, s \end{aligned}$$

Sea $\nu(A, B) := \dim\{X \in \mathbb{C}^{m \times n} \mid AX - XB = 0\}$. Es conocido que

$$\nu(A, B) = \sum_{i=1}^u \sum_{j=1}^v \sum_{k=1}^{\nu_i} \sum_{h=1}^{\rho_j} \delta_{ikjh}, \quad (3.1)$$

donde

$$\delta_{ikjh} := \text{grado} \left[\text{m.c.d.}((\lambda - \lambda_i)^{m_{ik}}, (\lambda - \mu_j)^{n_{jh}}) \right]$$

Véanse los libros de Gantmacher [28, Théorème 1, p. 221, Chap. 8, n^o 1], y de Lancaster y Tismenetsky [41, Corollary 1, Section 12.5, p. 421].

Siempre que $\lambda_i \neq \mu_j$ se sigue que

$$\text{m.c.d.}((\lambda - \lambda_i)^{m_{ik}}, (\lambda - \mu_j)^{n_{jh}}) = 1$$

y, por tanto, $\delta_{ikjh} = 0$.

Teniendo en cuenta que en la fórmula (3.1) sólo participan esencialmente los divisores elementales de los valores propios comunes $\lambda_1, \lambda_2, \dots, \lambda_s$, se deduce que

$$\nu(A, B) = \sum_{k=1}^s \sum_{(i,j)} \min(m_{ki}, n_{kj}). \quad (3.2)$$

Como las particiones de Weyr son las particiones conjugadas de las particiones de Segré, por la Proposición 2.6.1 del Capítulo 2 (véase la página 36), se tiene que

$$\nu(A, B) = \sum_{k=1}^s w(\lambda_k, A) \cdot w(\lambda_k, B). \quad (3.3)$$

Esta última formula puede ser expresada con la notación del producto escalar en \mathbb{W} :

$$\nu(A, B) = w(A) \cdot w(B).$$

Véase la Subsección 3.4.1.

Es conocido que cualquier matriz $M \in \mathbb{C}^{n \times n}$ es semejante a su matriz traspuesta M^T . Por tanto, para todo $\mu \in \Lambda(M)$ se tiene que $w(\mu, M) = w(\mu, M^T)$. Por lo tanto, la fórmula (3.3) implica que

$$\nu(A, B) = \nu(A^T, B) = \nu(A, B^T) = \nu(A^T, B^T). \quad (3.4)$$

En este punto es preciso que recordemos una desigualdad relativa al producto escalar que definimos al comienzo de la Sección 2.6 del Capítulo 2. Para todo $\mathbf{x} \in \mathbb{R}^N$ se tiene que $\mathbf{x} \cdot \mathbf{x} \geq 0$, y, además, $\mathbf{x} \cdot \mathbf{x} = 0$ si y sólo si $\mathbf{x} = \mathbf{0}$. Por consiguiente, si $\mathbf{a}, \mathbf{b} \in \mathbb{R}^N$, entonces

$$2\mathbf{a} \cdot \mathbf{b} \leq \mathbf{a} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} \quad (3.5)$$

y se da la igualdad si y sólo si $\mathbf{a} = \mathbf{b}$. En efecto,

$$(\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) = \mathbf{a} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} - 2\mathbf{a} \cdot \mathbf{b};$$

por lo tanto, la desigualdad (3.5) se sigue de que $(\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) \geq 0$. Y $\mathbf{a} = \mathbf{b}$ es equivalente a $(\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) = 0$.

Teorema 3.8.1. Sean $A \in \mathbb{C}^{m \times m}$, $B \in \mathbb{C}^{n \times n}$, entonces

$$\nu(A, B) \leq \frac{1}{2} [\nu(A, A) + \nu(B, B)]. \quad (3.6)$$

Además, vale el signo = si y sólo si $m = n$ y las matrices A y B son semejantes.

DEMOSTRACIÓN. Si $\lambda_k \in \Lambda(A) \cap \Lambda(B)$, entonces

$$2w(\lambda_k, A) \cdot w(\lambda_k, B) \leq w(\lambda_k, A) \cdot w(\lambda_k, A) + w(\lambda_k, B) \cdot w(\lambda_k, B);$$

por lo tanto, con las notaciones anteriores sobre los valores propios de A y B ,

$$\begin{aligned} \sum_{k=1}^s 2w(\lambda_k, A) \cdot w(\lambda_k, B) &\leq \sum_{k=1}^s w(\lambda_k, A) \cdot w(\lambda_k, A) + \sum_{k=1}^s w(\lambda_k, B) \cdot w(\lambda_k, B) \\ &\leq \sum_{k=1}^u w(\lambda_k, A) \cdot w(\lambda_k, A) + \sum_{k=1}^v w(\lambda_k, B) \cdot w(\lambda_k, B). \end{aligned}$$

De donde, por (3.3)

$$\nu(A, B) \leq \frac{1}{2} [\nu(A, A) + \nu(B, B)].$$

Supongamos ahora que

$$\nu(A, B) = \frac{1}{2} [\nu(A, A) + \nu(B, B)]. \quad (3.7)$$

A partir de aquí debemos probar que $m = n$ y que A y B son semejantes; en particular, tenemos que demostrar que A y B tienen los *mismos* valores propios. Sea

$$\Lambda(A) \cap \Lambda(B) = \{\lambda_1 = \mu_1, \dots, \lambda_s = \mu_s\}.$$

Entonces (3.7) implica que

$$\begin{aligned} \sum_{k=1}^s 2w(\lambda_k, A) \cdot w(\lambda_k, B) &= \sum_{k=1}^s w(\lambda_k, A) \cdot w(\lambda_k, A) + \sum_{k=1}^s w(\lambda_k, B) \cdot w(\lambda_k, B) \\ &+ \sum_{j=s+1}^u w(\lambda_j, A) \cdot w(\lambda_j, A) + \sum_{j=s+1}^v w(\mu_j, B) \cdot w(\mu_j, B); \end{aligned}$$

de aquí que

$$\begin{aligned} \sum_{k=1}^s [w(\lambda_k, A) \cdot w(\lambda_k, A) + w(\lambda_k, B) \cdot w(\lambda_k, B) - 2w(\lambda_k, A) \cdot w(\lambda_k, B)] \\ + \sum_{j=s+1}^u w(\lambda_j, A) \cdot w(\lambda_j, A) + \sum_{j=s+1}^v w(\mu_j, B) \cdot w(\mu_j, B) = 0. \end{aligned}$$

Como los sumandos de estas tres sumas son no negativos, por las propiedades antedichas del producto escalar, se sigue que

$$\begin{aligned} \forall k = 1, \dots, s \quad w(\lambda_k, A) &= w(\lambda_k, B) \\ \forall j = s+1, \dots, u \quad w(\lambda_j, A) &= (0, 0, \dots) \quad \text{es decir } \lambda_j \notin \Lambda(A) \\ \forall j = s+1, \dots, v \quad w(\mu_j, B) &= (0, 0, \dots) \quad \text{es decir } \mu_j \notin \Lambda(B). \end{aligned} \quad (3.8)$$

por lo cual, $\Lambda(A) = \Lambda(B) = \{\lambda_1, \dots, \lambda_s\}$. Además, (3.8) demuestra que A y B son semejantes. En particular, se tiene que $m = n$. \square

Corolario 1. Sean $A, B \in \mathbb{C}^{n \times n}$, entonces A es semejante a B si y sólo si

$$\nu(A, A) = \nu(A, B) = \nu(B, B).$$

En varias ocasiones conviene hacer hincapié en que las matrices $m \times n$ son “vectores” de mn componentes. Para ello, se colocan las filas de X en columna, una tras otra. De este modo definimos un operador lineal, vec , del espacio vectorial $\mathbb{C}^{m \times n}$ en el espacio vectorial $\mathbb{C}^{mn \times 1}$:

Definición 3.8.1. El operador vectorizador

$$\text{vec} : \mathbb{C}^{m \times n} \rightarrow \mathbb{C}^{mn \times 1}$$

se define mediante

$$\text{vec}(X) := (x_{11}, x_{12}, \dots, x_{1n}, x_{21}, x_{22}, \dots, x_{2n}, \dots, x_{m1}, x_{m2}, \dots, x_{mn})^T.$$

donde

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix}.$$

Este operador vec tiene la siguiente relación con el producto de matrices:

Lema 3.8.1. Sea A, X, B tres matrices cualesquiera de dimensiones tales que el producto AXB tenga sentido. Entonces

$$\text{vec}(AXB) = (A \otimes B^T) \text{vec}(X), \quad (3.9)$$

donde \otimes denota el producto tensorial o de Kronecker de matrices y T la matriz traspuesta.

Para cualquier entero $n > 0$, denotamos mediante \underline{n} el conjunto $\{1, \dots, n\}$. Previamente, vamos a establecer unas convenciones de notación sobre una forma de designar las filas y columnas de un producto tensorial de dos matrices; supongamos que $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{p \times q}$

$$A = (a_{ij}), B = (b_{hk}), \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad h = 1, \dots, p, \quad k = 1, \dots, q.$$

Sea $P := A \otimes B^T$ (obsérvese que P es el producto tensorial de A por la traspuesta de B). Un elemento genérico de P es de la forma

$$a_{ih}b_{kj}, \quad \text{con } i = 1, \dots, m, \quad j = 1, \dots, n, \quad h = 1, \dots, p, \quad k = 1, \dots, q.$$

Denotaremos las filas y las columnas de P por pares de enteros que ordenaremos lexicográficamente. La matriz P tiene mp filas y nq columnas. Si designamos por

$$P^{(i,j),(h,k)}$$

al elemento de P situado en la fila (i, j) y columna (h, k) , tendremos que

$$P^{(i,j),(h,k)} = a_{ih}b_{kj}$$

Para denotar una submatriz cuadrada de P de orden r indicaremos sus filas y columnas en orden creciente. Así, si $\alpha := ((i_1, j_1), \dots, (i_r, j_r))$ es una sucesión de elementos de $\underline{m} \times \underline{q}$ ordenados $(i_1, j_1) < \dots < (i_r, j_r)$, y $\beta := ((h_1, k_1), \dots, (h_r, k_r))$ es una sucesión de elementos de $\underline{n} \times \underline{p}$ ordenados $(h_1, k_1) < \dots < (h_r, k_r)$, la submatriz de P correspondiente a las filas de α y a las columnas de β la notaremos por $P[\alpha, \beta]$. Como los elementos de las sucesiones α y β son distintos, también hablaremos de los subconjuntos α y β de $\underline{m} \times \underline{q}$ y $\underline{n} \times \underline{p}$, respectivamente.

Ejemplo 3.8.1. Sean $A \in \mathbb{C}^{2 \times 3}$, $B \in \mathbb{C}^{2 \times 4}$, $X \in \mathbb{C}^{3 \times 2}$, entonces

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}, \quad B^T = \begin{pmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \\ b_{13} & b_{23} \\ b_{14} & b_{24} \end{pmatrix}, \quad P = (P^{(i,j),(h,k)}) = A \otimes B^T,$$

con

$$p_{(i,j),(h,k)} = a_{ih}b_{kj}$$

Vectorizando la ecuación $AXB = O_{2 \times 4}$, nos queda $(A \otimes B^T) \text{vec}(X) = 0$. Escribamos desarrollada esta ecuación última para una mejor comprensión de la notación.

$$\begin{matrix} & \begin{matrix} (1,1) & (1,2) & (2,1) & (2,2) & (3,1) & (3,2) \end{matrix} \\ \begin{matrix} (1,1) \\ (1,2) \\ (1,3) \\ (1,4) \\ (2,1) \\ (2,2) \\ (2,3) \\ (2,4) \end{matrix} & \begin{pmatrix} p_{(1,1),(1,1)} & p_{(1,1),(1,2)} & p_{(1,1),(2,1)} & p_{(1,1),(2,2)} & p_{(1,1),(3,1)} & p_{(1,1),(3,2)} \\ p_{(1,2),(1,1)} & p_{(1,2),(1,2)} & p_{(1,2),(2,1)} & p_{(1,2),(2,2)} & p_{(1,2),(3,1)} & p_{(1,2),(3,2)} \\ p_{(1,3),(1,1)} & p_{(1,3),(1,2)} & p_{(1,3),(2,1)} & p_{(1,3),(2,2)} & p_{(1,3),(3,1)} & p_{(1,3),(3,2)} \\ p_{(1,4),(1,1)} & p_{(1,4),(1,2)} & p_{(1,4),(2,1)} & p_{(1,4),(2,2)} & p_{(1,4),(3,1)} & p_{(1,4),(3,2)} \\ p_{(2,1),(1,1)} & p_{(2,1),(1,2)} & p_{(2,1),(2,1)} & p_{(2,1),(2,2)} & p_{(2,1),(3,1)} & p_{(2,1),(3,2)} \\ p_{(2,2),(1,1)} & p_{(2,2),(1,2)} & p_{(2,2),(2,1)} & p_{(2,2),(2,2)} & p_{(2,2),(3,1)} & p_{(2,2),(3,2)} \\ p_{(2,3),(1,1)} & p_{(2,3),(1,2)} & p_{(2,3),(2,1)} & p_{(2,3),(2,2)} & p_{(2,3),(3,1)} & p_{(2,3),(3,2)} \\ p_{(2,4),(1,1)} & p_{(2,4),(1,2)} & p_{(2,4),(2,1)} & p_{(2,4),(2,2)} & p_{(2,4),(3,1)} & p_{(2,4),(3,2)} \end{pmatrix} \end{matrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \quad (3.10)$$

Sean ahora $A \in \mathbb{C}^{m \times m}$, $B \in \mathbb{C}^{n \times n}$ y consideremos la ecuación matricial $AX - XB = O$ cuya incógnita X pertenece a $\mathbb{C}^{m \times n}$. El conjunto de las soluciones de esta ecuación es el núcleo de la aplicación lineal

$$X \mapsto AX - XB$$

del espacio $\mathbb{C}^{m \times n}$ en sí mismo. A la dimensión de este núcleo la venimos denotando por $\nu(A, B)$. Aplicando el Lema 3.8.1 a la ecuación $AX - XB = O$ resulta

$$(A \otimes I_n - I_m \otimes B^T) \text{vec}(X) = 0;$$

por lo que

$$\nu(A, B) = \nu(A \otimes I_n - I_m \otimes B^T) = mn - \text{rg}(A \otimes I_n - I_m \otimes B^T).$$

Por (3.4), se tiene que

$$\nu(A, B) = \nu(A, B^T) = mn - \text{rg}(A \otimes I_n - I_m \otimes B).$$

De aquí se deduce el teorema siguiente.

Teorema 3.8.2. *Dos matrices $A, B \in \mathbb{C}^{n \times n}$ son semejantes si y sólo si*

$$\text{rg}(A \otimes I_n - I_n \otimes A) = \text{rg}(A \otimes I_n - I_n \otimes B) = \text{rg}(B \otimes I_n - I_n \otimes B).$$

3.9. Matrices no derogatorias

Conjetura de Sylvester

Teorema 3.9.1. *Una matriz $A \in \mathbb{C}^{n \times n}$ es no derogatoria si y sólo si las únicas matrices que conmutan con A son los polinomios en A .*

Observación 3.9.1. Esto quiere decir que si $B \in \mathbb{C}^{n \times n}$ satisface $AB = BA$, entonces existe un polinomio $p(\lambda) \in \mathbb{C}[\lambda]$ tal que $B = p(A)$.

3.10. Nulidades de un producto

Lema 3.10.1. Sea $A \in \mathbb{C}^{n \times n}$ y sean los números complejos $\lambda_1, \lambda_2, \dots, \lambda_q$ algunos valores propios distintos de A y p_1, p_2, \dots, p_q enteros no negativos. Entonces

$$\nu \left[\prod_{i=1}^q (\lambda_i I_n - A)^{p_i} \right] = \sum_{i=1}^q \nu [(\lambda_i I_n - A)^{p_i}]. \quad (3.1)$$

Observación 3.10.1. De la ley de Sylvester de la nulidad (Corolario 1 del Teorema 3.1.1) se sigue que la nulidad de un producto de matrices es menor o igual que la suma de las nulidades de estas matrices:

$$\nu(A_1 A_2 \cdots A_q) \leq \nu(A_1) + \nu(A_2) + \cdots + \nu(A_q).$$

El lema presente dice que la igualdad se da para potencias de matrices de la forma $\lambda_i I_n - A$, donde λ_i es un valor propio de A .

Si algunos (o todos) de los números $\lambda_i, i \in \{1, \dots, q\}$ no son valores propios de A la igualdad (3.1) sigue siendo válida. En efecto, en primer lugar, como el producto de las matrices $\alpha I_n - A$ y $\beta I_n - A$ es conmutativo para cualesquiera $\alpha, \beta \in \mathbb{C}$, se tiene que para todos $i, j \in \{1, \dots, q\}$ es cierta la igualdad

$$(\lambda_i I_n - A)^{p_i} (\lambda_j I_n - A)^{p_j} = (\lambda_j I_n - A)^{p_j} (\lambda_i I_n - A)^{p_i};$$

por tanto,

$$\begin{aligned} (\lambda_1 I_n - A)^{p_1} \cdots (\lambda_i I_n - A)^{p_i} \cdots (\lambda_q I_n - A)^{p_q} = \\ (\lambda_1 I_n - A)^{p_1} \cdots (\lambda_{i-1} I_n - A)^{p_{i-1}} (\lambda_{i+1} I_n - A)^{p_{i+1}} \\ \cdots (\lambda_q I_n - A)^{p_q} (\lambda_i I_n - A)^{p_i}. \end{aligned}$$

Así pues, cuando λ_i no es valor propio de A , el rango de

$$(\lambda_1 I_n - A)^{p_1} \cdots (\lambda_i I_n - A)^{p_i} \cdots (\lambda_q I_n - A)^{p_q}$$

es igual al rango de

$$(\lambda_1 I_n - A)^{p_1} \cdots (\lambda_{i-1} I_n - A)^{p_{i-1}} (\lambda_{i+1} I_n - A)^{p_{i+1}} \cdots (\lambda_q I_n - A)^{p_q},$$

y se puede cancelar el factor $(\lambda_i I_n - A)^{p_i}$ en el miembro de la izquierda de (3.1); como $\text{rg}(\lambda_i I_n - A)^{p_i} = n$, la nulidad de esta matriz es cero y no aporta nada a la suma del segundo miembro de (3.1).

3.11. Nulidades de matrices cuasidiagonales por bloques

Consideremos la matriz de la ecuación (3.2) de la Sección 3.7

$$\begin{aligned} P_{\alpha, k}(A) &:= \begin{pmatrix} \alpha I - A & 0 & \cdots & 0 \\ I & \alpha I - A & & \vdots \\ & I & \alpha I - A & \\ & & \ddots & \ddots \\ 0 & & \cdots & \alpha I - A & 0 \\ & & & I & \alpha I - A \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ I_{k-1} & 0 \end{pmatrix} \otimes I + I_k \otimes (\alpha I - A); \end{aligned} \quad (3.1)$$

supongamos que cambiamos en los bloques diagonales de $P_{\alpha,k}(A)$ el número α por diferentes números complejos: α_1 en los primeros k_1 bloques, α_2 en los k_2 bloques siguientes, ..., α_q en los últimos k_q bloques, con $k_1 + k_2 + \dots + k_q = k$

$$P_{\alpha_1, \dots, \alpha_q}^{k_1, \dots, k_q}(A) := \begin{bmatrix} P_{\alpha_1, k_1}(A) & 0 & \dots & 0 & 0 \\ Q_{k_2, k_1}(I) & P_{\alpha_2, k_2}(A) & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & Q_{k_q, k_{q-1}}(I) & P_{\alpha_q, k_q}(A) \end{bmatrix} \quad (3.2)$$

donde para $j = 1, \dots, q-1$,

$$Q_{k_{j+1}, k_j}(I) := \begin{bmatrix} 0 & \dots & 0 & I \\ 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix} \in \mathbb{C}^{(k_{j+1}n) \times (k_j n)}.$$

Mediante transformaciones elementales se puede demostrar el lema siguiente.

Lema 3.11.1.

$$\nu(P_{\alpha_1, \dots, \alpha_q}^{k_1, \dots, k_q}(A)) = \nu(\text{diag}(P_{\alpha_1, k_1}(A), \dots, P_{\alpha_q, k_q}(A))).$$

Por lo tanto,

$$\nu(P_{\alpha_1, \dots, \alpha_q}^{k_1, \dots, k_q}(A)) = \sum_{s=1}^q \nu(P_{\alpha_s, k_s}(A)). \quad (3.3)$$

3.12. Rango mínimo de factores invariantes

Sean $A \in \mathbb{C}^{n \times n}$ y $f_1(\lambda) \mid \dots \mid f_n(\lambda)$ los factores invariantes de A . Es conocido que el último factor invariante, $f_n(\lambda)$ de A es su polinomio mínimo. Por tanto, $f_n(A) = O$; en particular, $\text{rg } f_n(A) = 0$. Veremos en esta sección un teorema de Menge (1932) que extiende este resultado.

Teorema 3.12.1 (Menge). *Sea $A \in \mathbb{C}^{n \times n}$ y sean $f_1(\lambda) \mid \dots \mid f_n(\lambda)$ los factores invariantes de A . Sea $d_i := \text{gr } f_i(\lambda)$ para cada $i = 1, \dots, n$. Entonces,*

$$\text{rg } f_i(A) = \min_{\substack{f(\lambda) \in \mathbb{C}[\lambda] \\ 0 \leq \text{gr } f(\lambda) \leq d_i}} \text{rg } f(A). \quad (3.1)$$

DEMOSTRACIÓN. . En primer lugar, si $B = P^{-1}AP$ con $P \in \mathbb{C}^{n \times n}$ invertible y $f(\lambda) \in \mathbb{C}[\lambda]$ se tiene que $f(B) = P^{-1}f(A)P$; por tanto, $\text{rg } f(B) = \text{rg } f(A)$. Así pues, podemos suponer que la matriz A está en la forma de Jordan $A = \text{diag}(J_1, J_2, \dots, J_u)$ donde $\Lambda(A) = \{\lambda_1, \dots, \lambda_u\}$ y J_j es la suma directa de todos los bloques de Jordan asociados a λ_j , ($j = 1, \dots, u$). Sea

$$s(\lambda_j, A) = (\nu_{j1}, \nu_{j2}, \dots, \nu_{j, n-i+1}, \dots, \nu_{j, n-1}, \nu_{jn})$$

la partición de Segré de λ_j ; esta notación es distinta de la que utilizamos en general, pues aquí es una n -tupla y $\nu_{in} \geq 0$, pudiendo ser igual a cero. Entonces,

$$f_i(\lambda) = (\lambda - \lambda_1)^{\nu_{1, n-i+1}} \dots (\lambda - \lambda_u)^{\nu_{u, n-i+1}}.$$

Por tanto, la nulidad de $f_i(A)$ es igual a

$$\nu[(A - \lambda_1 I)^{\nu_{1, n-i+1}} \dots (A - \lambda_u I)^{\nu_{u, n-i+1}}];$$

lo que por el Lema 3.10.1 implica

$$\nu[f_i(A)] = \nu[(A - \lambda_1 I)^{\nu_{1, n-i+1}}] + \dots + \nu[(A - \lambda_u I)^{\nu_{u, n-i+1}}].$$

Por la Observación 3.10.1 si $f(\lambda)$ es un polinomio cualquiera, se tiene que

$$\nu[f(A)] = \nu[(A - \alpha_1 I)^{p_1}] + \cdots + \nu[(A - \alpha_q I)^{p_q}],$$

donde $\alpha_1, \dots, \alpha_q$ son las raíces de $f(\lambda)$ que son valores propios de A y p_1, \dots, p_q son sus multiplicidades algebraicas respecto de $f(\lambda)$. Si ninguna raíz de $f(\lambda)$ fuera valor propio de A , entonces $\nu[f(A)] = 0$. Por tanto, la demostración de (3.1) se reduce a probar que

$$\nu[f_i(A)] = \max_{\substack{g(\lambda) \in \mathbb{C}[\lambda] \\ \text{gr } g(\lambda) \leq d_i \\ \Lambda(g) \subset \Lambda(A)}} \nu[g(A)], \quad (3.2)$$

donde $\Lambda(g)$ denota el conjunto de raíces del polinomio $g(\lambda)$. Sea $J_k(0)$ el bloque de Jordan de orden k asociado a 0 y sea $p \geq 0$ un entero; es fácil comprobar que

$$\nu[J_k(0)^p] = \min(k, p). \quad (3.3)$$

Sea $g(\lambda) := (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_u)^{m_u}$ con $m_j \geq 0$ entero para cada $j = 1, \dots, u$ y $m_1 + \cdots + m_u \leq d_i$. Es decir, $g(\lambda)$ representa un polinomio cualquiera tal que $\Lambda(g) \subset \Lambda(A)$ y de grado $\leq d_i$. Cuando $m_j = \nu_{j, n-i+1}$, $j = 1, \dots, u$, el polinomio $g(\lambda)$ es el factor invariante $f_i(\lambda)$. Vamos a hallar la nulidad $\nu[g(A)]$ explícitamente en función de los exponentes m_1, \dots, m_u . A partir de ahí comprobaremos que dicha nulidad es máxima cuando $g(\lambda) = f_i(\lambda)$. En primer lugar, por la Observación 3.10.1 se tiene que

$$\nu[g(A)] = \nu[(A - \lambda_1 I)^{m_1}] + \cdots + \nu[(A - \lambda_u I)^{m_u}].$$

Ahora bien, para cada $j = 1, \dots, u$,

$$\nu[(A - \lambda_j I)^{m_j}] = \nu[(J_j - \lambda_j I)^{m_j}];$$

como,

$$\nu[(J_j - \lambda_j I)^{m_j}] = \nu[\text{diag}(J_{\nu_{j1}}(0)^{m_j}, J_{\nu_{j2}}(0)^{m_j}, \dots, J_{\nu_{jn}}(0)^{m_j})],$$

se sigue por (3.3) que

$$\nu[g(A)] = \sum_{j=1}^u [\min(\nu_{j1}, m_j) + \cdots + \min(\nu_{jn}, m_j)].$$

Para cada j la partición conjugada de $(m_j, 0, \dots)$ es

$$(\overbrace{1, \dots, 1}^{m_j}, 0, \dots);$$

por tanto, aplicando la Proposición 2.6.1 del Capítulo 2 a las particiones $s(\lambda_j, A)$ y $(m_j, 0, \dots)$ se tiene que

$$\min(\nu_{j1}, m_j) + \cdots + \min(\nu_{jn}, m_j) = w(\lambda_j, A) \cdot (\overbrace{1, \dots, 1}^{m_j}, 0, \dots, 0);$$

donde este producto escalar \cdot se realiza en \mathbb{R}^n . En consecuencia,

$$\begin{aligned} \nu[g(A)] = w(\lambda_1, A) \cdot (\overbrace{1, \dots, 1}^{m_1}, 0, \dots, 0) &+ w(\lambda_2, A) \cdot (\overbrace{1, \dots, 1}^{m_2}, 0, \dots, 0) \\ &+ \cdots + w(\lambda_u, A) \cdot (\overbrace{1, \dots, 1}^{m_u}, 0, \dots, 0). \end{aligned}$$

Así pues, $\nu[g(A)]$ es la suma de las m_1 partes más grandes de $w(\lambda_1, A)$, más la suma de las m_2 partes más grandes de $w(\lambda_2, A)$, \dots , más la suma de las m_u partes más grandes

de $w(\lambda_u, A)$. En particular, aplicando este resultado cuando cada $m_j = \nu_{j, n-i+1}$, se tiene que $\nu[f_i(A)]$ es la suma de las $\nu_{1, n-i+1}$ partes más grandes de $w(\lambda_1, A)$, más la suma de las $\nu_{2, n-i+1}$ partes más grandes de $w(\lambda_2, A)$, \dots , más la suma de las $\nu_{u, n-i+1}$ partes más grandes de $w(\lambda_u, A)$. Pero, por ser la partición de Segré de λ_j la conjugada de la partición de Weyr de λ_j , se sigue que en $w(\lambda_j, A)$ hay exactamente $\nu_{j, n-i+1}$ partes mayores o iguales que $n - i + 1, j = 1, \dots, u$.

Por tanto, $\nu[f_i(A)]$ es la suma de las partes de la unión

$$\bigcup_{j=1}^u w(\lambda_j, A) \tag{3.4}$$

que son mayores o iguales que $n - i + 1$; en la unión (3.4) hay exactamente

$$\sum_{j=1}^u \nu_{j, n-i+1} = d_i$$

partes mayores o iguales que $n - i + 1$. En consecuencia, $\nu[f_i(A)]$ es la suma de las primeras d_i partes de la unión (3.4). Mientras que $\nu[g(A)]$ es la suma de algunas $m_1 + \dots + m_u (\leq d_i)$ partes de la misma partición unión. Por consiguiente,

$$\nu[f_i(A)] \geq \nu[g(A)].$$

□

3.13. Ejercicios

Ejercicio 3.1. Sean $A \in \mathbb{F}^{n \times n}, B \in \mathbb{F}^{n \times n}$. Demostrar que

$$B(\text{Ker}(AB)) = \text{Ker } A \cap \text{Ker } B.$$

Ejercicio 3.2. Sean $A \in \mathbb{C}^{n \times n}$ y α un valor propio de A . Para cada $k = 1, \dots, n, \dots$ sea $P_{\alpha, k}(A)$ la matriz $kn \times kn$ de bloques definida por la fórmula (3.2) de la Sección 3.7. Demostrar que

$$\nu[P_{\alpha, 1}(A)] \leq \nu[P_{\alpha, 2}(A)] \leq \dots \leq \nu[P_{\alpha, k}(A)] \leq \nu[P_{\alpha, k+1}(A)] \leq \dots$$

Incluso más, demostrar que esta sucesión es estrictamente creciente hasta que se estaciona por primera vez, digamos para el valor $k = p$, y que a partir de ese índice permanece constante:

$$\nu[P_{\alpha, 1}(A)] < \nu[P_{\alpha, 2}(A)] < \dots < \nu[P_{\alpha, p}(A)] = \nu[P_{\alpha, p+1}(A)] = \dots$$

Además debe ser $p \leq n$, y

$$\nu[P_{\alpha, k}(A)] = m(\alpha, A),$$

para todo $k \geq p$.

Ejercicio 3.3. Dada $A \in \mathbb{C}^{n \times n}$, sean $f_1(\lambda) \mid \dots \mid f_n(\lambda)$ los factores invariantes de A , donde \mid denota “divide a”. Para cada $p = 1, \dots, n$, llamemos

$$i_p(A) := \text{Card} \{i \in \{1, \dots, n\} \mid \text{gr}(f_i(\lambda)) \geq p\},$$

siendo $\text{gr}(f_i(\lambda))$ el grado del polinomio $f_i(\lambda)$. Sea $\Lambda(A) = \{\lambda_1, \dots, \lambda_u\}$. Demostrar que

$$i_p(A) = \left(\bigcup_{i=1}^u w(\lambda_i, A) \right)_p;$$

es decir, $i_p(A)$ es igual a la componente p -sima de la unión de las particiones de Weyr de todos los valores propios de A .

Ejercicio 3.4. Sean $\lambda_1, \dots, \lambda_u$ números complejos distintos y sean $A_1 \in \mathbb{C}^{m_1 \times m_1}, \dots, A_u \in \mathbb{C}^{m_u \times m_u}$ matrices tales que $\Lambda(A_1) = \{\lambda_1\}, \dots, \Lambda(A_u) = \{\lambda_u\}$. Sea $A = \text{diag}(A_1, \dots, A_u)$. Demostrar que para cada $i \in \{1, \dots, u\}$,

$$w(\lambda_i, A) = w(\lambda_i, A_i).$$

Ejercicio 3.5. Supongamos que $A \in \mathbb{C}^{n \times n}$, que α es un valor propio de A y definamos $B := \alpha I - A$. Demostrar que 0 es valor propio de B y que

$$w(0, B) = w(\alpha, A).$$

Ejercicio 3.6. Sea $\lambda_0 \in \mathbb{C}$. Sea $B \in \mathbb{C}^{n \times n}$ y definamos $A := \lambda_0 I - B$. Demostrar que

$$\Lambda(A) = \{\lambda_0 - \mu \mid \mu \in \Lambda(B)\}$$

y que para cada $\mu \in \Lambda(B)$,

$$w(\lambda_0 - \mu, A) = w(\mu, B).$$

Ejercicio 3.7. Sean $A \in \mathbb{C}^{m \times m}$ y $B \in \mathbb{C}^{n \times n}$. Utilizando la desigualdad de Cauchy-Schwarz, demostrar que

$$\nu^2(A, B) \leq \nu(A, A)\nu(B, B).$$

Ejercicio 3.8. Sean $A, B \in \mathbb{C}^{n \times n}$ y supongamos que para toda $X \in \mathbb{C}^{n \times n}$ se tiene que

$$\nu(A, X) = \nu(B, X).$$

Demostrar que A es semejante a B .

Ejercicio 3.9. Una matriz $A \in \mathbb{C}^{n \times n}$ se llama nilpotente si existe un entero positivo p tal que $A^p = 0$. Demostrar que A es nilpotente si y sólo si $\Lambda(A) = \{0\}$.

3.14. Notas al Capítulo 3

La desigualdad de Frobenius sobre rangos (Teorema 3.1.1) y el Lema 3.1.1 siguen la exposición del libro clásico de MacDuffee [45, Lemma 8.3 y Theorem 8.3, pág. 11]. La Sección 3.6 sobre la forma canónica dual sigue de cerca la exposición de Zaballa [57].

El Teorema 3.8.1 es debido a Byrnes y Gauger (1977) [15], quienes utilizaron la hipótesis adicional de que A y B tuvieran el mismo polinomio característico. Ya en 1979 fue descubierto que esta hipótesis era redundante, y se podía suprimir sin mengua del criterio. Véase la referencia de Matthews [48], quien dice que esta redundancia fue anticipada por Dixon. Friedland [25] probó este teorema utilizando la propiedad de que la función $\min(x, y)$ es cóncava homogénea sobre \mathbb{R}^2 , pues

$$\begin{aligned} \min(ax, ay) &= a \min(x, y), \quad a \geq 0, \\ \min\left(\frac{x+u}{2}, \frac{y+v}{2}\right) &\geq \frac{1}{2} [\min(x, y) + \min(u, v)]. \end{aligned}$$

Nosotros hacemos una demostración del teorema que no se basa en esta propiedad.

A la hora de ordenar los elementos de una matriz X para formar un vector hay cuatro posibles maneras. Se pueden disponer las filas de X en fila. Es posible colocar las columnas de X en una fila. Es posible colocar las filas de X en una larga columna; nosotros procederemos siempre de esta forma. Y, finalmente, es posible apilar las columnas de X una tras otra; esta última manera es la preferida en el momento de escribir este libro. Aunque las cuatro formas son equivalentes, al leer cada libro o artículo hay que mirar qué forma de vectorizar eligió el autor. Siguiendo la tradición que estuvo de moda en los años 1960, hemos elegido vectorizar una matriz X poniendo sus elementos así

$$\text{vec}(X) = (x_{11}, x_{12}, \dots, x_{1n}, x_{21}, x_{22}, \dots, x_{2n}, \dots, x_{m1}, x_{m2}, \dots, x_{mn})^T.$$

Una razón para hacerlo de esta forma es que al vectorizar la ecuación de Sylvester $AX - XB = O$, resulta $(A \otimes I_n - I_m \otimes B^T) \text{vec}(X) = 0$, quedando denotadas las columnas de $A \otimes I_n - I_m \otimes B^T$ por los pares (h, k) que ordenan de igual manera los elementos x_{hk} de la columna $\text{vec}(X)$.

El Teorema 3.12.1 es debido a Menge (1932); véase la página 74 del libro de MacDuffee [45]. El Ejercicio 3.7 es un resultado debido a Dixon [21], quien probó además que la igualdad $\nu^2(A, B) = \nu(A, A)\nu(B, B)$ es válida si y sólo si A y B son semejantes a sumas directas de una tercera matriz C .

Capítulo 4

Raíces comunes de polinomios

Toda la información posible sobre un polinomio está encerrada en los coeficientes del mismo. Tendría que ser posible contestar a muchas cuestiones sobre el polinomio a partir de condiciones que satisfagan sus coeficientes. En este capítulo mostraremos cómo es posible responder de esta manera a las preguntas de cuántas raíces distintas tiene un polinomio, o cuántas raíces tiene de multiplicidad dada. El método pasa por averiguar cuándo tienen raíces comunes dos (o más) polinomios y, en caso afirmativo, cuántas tienen. Todo ello se hará mediante el cálculo de rangos de matrices construidas con los coeficientes del polinomio o polinomios.

4.1. Resultante de polinomios

Recordemos que el número complejo λ_0 se dice que es una raíz de multiplicidad m del polinomio de coeficientes complejos $f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n$, o que es una raíz m -múltiple, si existe un polinomio $g(\lambda)$ tal que

$$f(\lambda) = (\lambda - \lambda_0)^m g(\lambda), \text{ y } g(\lambda_0) \neq 0.$$

Utilizando la fórmula de Taylor puede demostrarse fácilmente que λ_0 es una raíz de multiplicidad m de $f(\lambda)$ si y sólo si

$$f(\lambda_0) = f'(\lambda_0) = \dots = f^{(m-1)}(\lambda_0) = 0, \quad f^{(m)}(\lambda_0) \neq 0. \quad (4.1)$$

Teorema 4.1.1. Sean

$$\begin{aligned} f(\lambda) &= a_0\lambda^n + a_1\lambda^{n-1} + \dots + a_n \\ g(\lambda) &= b_0\lambda^m + b_1\lambda^{m-1} + \dots + b_m \end{aligned}$$

dos polinomios de grados $n \geq 1, m \geq 1$, con $a_0 \neq 0, b_0 \neq 0$. Una condición necesaria y suficiente para que los polinomios $f(\lambda)$ y $g(\lambda)$ tengan al menos una raíz común es que la matriz cuadrada de orden $n + m$

$$\begin{pmatrix} a_0 & a_1 & \dots & a_n & 0 & 0 & \dots & 0 \\ 0 & a_0 & \dots & a_{n-1} & a_n & 0 & \dots & 0 \\ \dots & \dots \\ 0 & 0 & \dots & a_0 & a_1 & a_2 & \dots & a_n \\ b_0 & b_1 & \dots & b_{m-1} & b_m & 0 & \dots & 0 \\ 0 & b_0 & \dots & b_{m-2} & b_{m-1} & b_m & \dots & 0 \\ \dots & \dots \\ 0 & 0 & \dots & b_0 & b_1 & b_2 & \dots & b_m \end{pmatrix} \quad (4.2)$$

sea singular.

La matriz (4.2) es llamada la *matriz resultante de Sylvester* de f y g . Aun más puede probarse sobre este asunto.

Teorema 4.1.2. *El número de raíces del máximo común divisor de los polinomios*

$$\begin{aligned} f(\lambda) &= a_0\lambda^n + a_1\lambda^{n-1} + \cdots + a_n \\ g(\lambda) &= b_0\lambda^m + b_1\lambda^{m-1} + \cdots + b_m \end{aligned}$$

de grados $n \geq 1, m \geq 1$ contando sus multiplicidades es igual a la nulidad de la matriz

$$R(f, g) := \begin{pmatrix} a_0 & a_1 & \cdots & a_n & 0 & 0 & \cdots & 0 \\ 0 & a_0 & \cdots & a_{n-1} & a_n & 0 & \cdots & 0 \\ \cdots & \cdots \\ 0 & 0 & \cdots & a_0 & a_1 & a_2 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_{m-1} & b_m & 0 & \cdots & 0 \\ 0 & b_0 & \cdots & b_{m-2} & b_{m-1} & b_m & \cdots & 0 \\ \cdots & \cdots \\ 0 & 0 & \cdots & b_0 & b_1 & b_2 & \cdots & b_m \end{pmatrix}. \quad (4.3)$$

Consideremos el polinomio complejo

$$a(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

y h polinomios complejos $b_1(\lambda), \dots, b_h(\lambda)$. Sea $\Pi := \{a(\lambda), b_1(\lambda), \dots, b_h(\lambda)\}$. Supongamos que el máximo grado de los polinomios $b_1(\lambda), \dots, b_h(\lambda)$ es $p \leq n$. Entonces escribamos

$$\begin{aligned} b_1(\lambda) &= b_{1,n-p}\lambda^p + b_{1,n-(p-1)}\lambda^{p-1} + \cdots + b_{1,n-1}\lambda + b_{1n} \\ &\quad \vdots \\ b_i(\lambda) &= b_{i,n-p}\lambda^p + b_{i,n-(p-1)}\lambda^{p-1} + \cdots + b_{i,n-1}\lambda + b_{in} \\ &\quad \vdots \\ b_h(\lambda) &= b_{h,n-p}\lambda^p + b_{h,n-(p-1)}\lambda^{p-1} + \cdots + b_{h,n-1}\lambda + b_{hn} \end{aligned}$$

con $b_{i,n-p} \neq 0$ para al menos un valor de $i \in \{1, \dots, h\}$. Definamos una matriz $p \times (n+p)$ asociada a $a(\lambda)$:

$$S_0 := \begin{bmatrix} 1 & a_1 & a_2 & \cdots & a_n & 0 & \cdots & 0 \\ 0 & 1 & a_1 & \cdots & a_{n-1} & a_n & \cdots & 0 \\ \vdots & & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & a_1 & \cdots & a_{n-1} & a_n \end{bmatrix}$$

y una matriz $n \times (n+p)$ asociada a cada $b_i(\lambda)$

$$S_i := \begin{bmatrix} b_{i,n-p} & b_{i,n-(p-1)} & \cdots & b_{in} & 0 & \cdots & 0 \\ 0 & b_{i,n-p} & \cdots & b_{i,n-1} & b_{in} & \cdots & 0 \\ \vdots & & \ddots & & \ddots & \ddots & \\ 0 & 0 & \cdots & b_{i,n-p} & \cdots & b_{i,n-1} & b_{in} \end{bmatrix}$$

para $i = 1, \dots, h$. Definimos la *matriz resultante generalizada de Sylvester* de los polinomios $a(\lambda), b_1(\lambda), \dots, b_h(\lambda)$ así:

$$R := R(\Pi) := R(a, b_1, \dots, b_h) := \begin{bmatrix} S_0 \\ S_1 \\ \vdots \\ S_h \end{bmatrix} \in \mathbb{C}^{(nh+p) \times (n+p)}. \quad (4.4)$$

Si consideramos los polinomios $a(\lambda)$ y $b_1(\lambda)$ de grados n y p , respectivamente, vemos que

$$R(a(\lambda), b_1(\lambda)) = \begin{bmatrix} S_0 \\ S_1 \end{bmatrix}$$

es la matriz $(n+p) \times (n+p)$ llamada resultante de Sylvester de $a(\lambda)$ y $b_1(\lambda)$. Con estas notaciones se tiene el teorema siguiente, cuya demostración puede verse en el libro de S. Barnett [6, Theorem 1.10 en la pág. 39].

Teorema 4.1.3. *El número de raíces del máximo común divisor de los $h+1$ polinomios*

$$a(\lambda), b_1(\lambda), \dots, b_h(\lambda),$$

contando multiplicidades es igual a la nulidad de la matriz R , definida en (4.4).

Es decir, $\nu(R) := n+p - \text{rg}(R)$, es el número de raíces comunes a estos $h+1$ polinomios en λ . Si $\nu(R) = 0$, o $\text{rg}(R) = n+p$ (rango completo), entonces no hay raíces comunes.

Observación 4.1.1. Si la presencia de otros parámetros en la descripción de los polinomios $f(\lambda)$ y $g(\lambda)$ pudiera dar lugar a confusión, denotaremos la resultante $R(f, g)$ por $R(f, g, \lambda)$ o $R(f(\lambda), g(\lambda), \lambda)$. Análogamente, se usaría $R(a(\lambda), b_1(\lambda), \dots, b_h(\lambda), \lambda)$ en vez de $R(a(\lambda), b_1(\lambda), \dots, b_h(\lambda))$.

4.2. Multiplicidades de las raíces

Dado un polinomio complejo

$$f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n,$$

para cada entero k , $1 \leq k \leq n$, denotemos por N_k (resp. ρ_k) el número de raíces distintas de $f(\lambda)$ de multiplicidad $\geq k$ (resp. $= k$). Obviamente, el número de raíces distintas de $f(\lambda)$ coincide con el número de raíces de $f(\lambda)$ de multiplicidad ≥ 1 , N_1 .

Teorema 4.2.1. *Dado el polinomio complejo*

$$f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n,$$

sea $R(f, f', \dots, f^{(k)})$ la matriz resultante de Sylvester generalizada de los polinomios $f(\lambda), f'(\lambda), \dots, f^{(k)}(\lambda)$, para cada $k = 1, 2, \dots, n, n+1$. Entonces,

$$\begin{aligned} N_k &= \nu(R(f, f', \dots, f^{(k-1)})) - \nu(R(f, f', \dots, f^{(k)})) \\ \rho_k &= \nu(R(f, f', \dots, f^{(k-1)})) - 2\nu(R(f, f', \dots, f^{(k)})) + \nu(R(f, f', \dots, f^{(k+1)})) \end{aligned}$$

para $1 \leq k \leq n$, donde si $k = 1$, $\nu(R(f, f', \dots, f^{(k-1)})) := \nu(R(f))$ y

$$R(f) := \begin{bmatrix} 1 & a_1 & a_2 & \dots & a_n & 0 & \dots & 0 \\ 0 & 1 & a_1 & \dots & a_{n-1} & a_n & \dots & 0 \\ \vdots & & \ddots & \ddots & & & \ddots & \vdots \\ 0 & 0 & \dots & 1 & a_1 & \dots & a_{n-1} & a_n \end{bmatrix}.$$

DEMOSTRACIÓN. La matriz $R(f)$ tiene dimensiones $(n-1) \times (2n-1)$; por lo tanto, $\text{rg } R(f) = n-1$. Lo que implica

$$\nu(R(f)) = n. \tag{4.1}$$

Demostraremos la fórmula para N_k por inducción sobre k . Supongamos que $k = 1$. Entonces el número de raíces distintas, r , de $f(\lambda)$ coincide con N_1 . Sean $\lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{C}$

las raíces distintas de $f(\lambda)$ con multiplicidades respectivas m_1, m_2, \dots, m_r . Entonces se sigue que el número de raíces comunes a $f(\lambda)$ y $f'(\lambda)$, contando sus multiplicidades en el $\text{mcd}(f(\lambda), f'(\lambda))$, es

$$(m_1 - 1) + (m_2 - 1) + \dots + (m_r - 1), \quad (4.2)$$

pues por (4.1), la raíz m_i -múltiple λ_i de $f(\lambda)$ es una raíz $(m_i - 1)$ -múltiple de $f'(\lambda)$, y, por ello, es una raíz de multiplicidad $m_i - 1$ del $\text{mcd}(f(\lambda), f'(\lambda))$, para cada $i = 1, \dots, r$. Por lo tanto,

$$(m_1 - 1) + (m_2 - 1) + \dots + (m_r - 1) = \nu(R(f, f')).$$

De donde,

$$m_1 + m_2 + \dots + m_r - r = \nu(R(f, f')),$$

pero $m_1 + m_2 + \dots + m_r = n$; lo que implica $n - r = \nu(R(f, f'))$; es decir, que $r = n - \nu(R(f, f'))$. Por tanto,

$$r = \nu(R(f)) - \nu(R(f, f'));$$

lo que equivale a

$$N_1 = \nu(R(f)) - \nu(R(f, f')), \quad (4.3)$$

pues $\nu(R(f)) = n$. Esto demuestra la fórmula sobre N_k para $k = 1$.

Supongamos ahora que la fórmula para N_k es cierta para $1, 2, \dots, k$, y probémosla para $k + 1$. Sean $\alpha_1, \alpha_2, \dots, \alpha_{N_{k+1}}$ las raíces de $f(\lambda)$ de multiplicidades $\geq k + 1$; denotemos sus multiplicidades respectivas por $m_1, m_2, \dots, m_{N_{k+1}}$. Entonces para cada $i = 1, \dots, N_{k+1}$, el número α_i es raíz $(m_i - (k + 1))$ -múltiple del polinomio

$$\text{mcd}(f(\lambda), f'(\lambda), \dots, f^{(k+1)}(\lambda));$$

lo que implica

$$(m_1 - (k + 1)) + \dots + (m_{N_{k+1}} - (k + 1)) = \nu(R(f, f', \dots, f^{(k+1)})).$$

Luego,

$$m_1 + \dots + m_{N_{k+1}} - (k + 1)N_{k+1} = \nu(R(f, f', \dots, f^{(k+1)})). \quad (4.4)$$

Pero, por la definición de los números $N_j, j = 1, \dots, n$, el polinomio $f(\lambda)$ tiene $N_k - N_{k+1}$ raíces distintas k -múltiples, $N_{k-1} - N_k$ raíces distintas $(k - 1)$ -múltiples, \dots , $N_2 - N_3$ raíces distintas dobles y $N_1 - N_2$ raíces distintas simples. Por consiguiente, el grado n es igual a la suma de las multiplicidades de las raíces, expresada así:

$$\begin{aligned} n &= (m_1 + \dots + m_{N_{k+1}}) + k(N_k - N_{k+1}) + (k - 1)(N_{k-1} - N_k) \\ &\quad + \dots + 2(N_2 - N_3) + (N_1 - N_2); \end{aligned}$$

de donde,

$$\begin{aligned} m_1 + \dots + m_{N_{k+1}} &= n - kN_k + kN_{k+1} - (k - 1)N_{k-1} \\ &\quad + (k - 1)N_k - (k - 2)N_{k-2} + (k - 2)N_{k-1} \\ &\quad \dots - 2N_2 + 2N_3 - N_1 + N_2 \\ &= n + kN_{k+1} - N_k - N_{k-1} - \dots - N_3 - N_2 - N_1; \end{aligned} \quad (4.5)$$

Ahora, (4.4) y (4.5) implican

$$\begin{aligned} n + kN_{k+1} - N_k - N_{k-1} - \dots - N_2 - N_1 - (k + 1)N_{k+1} \\ = \nu(R(f, f', \dots, f^{(k+1)})); \end{aligned}$$

de donde,

$$n - \sum_{i=1}^k N_i - N_{k+1} = \nu(R(f, f', \dots, f^{(k+1)}));$$

por tanto,

$$n - \sum_{i=1}^k N_i - \nu(R(f, f', \dots, f^{(k+1)})) = N_{k+1}. \quad (4.6)$$

Pero, por la hipótesis de inducción,

$$\begin{aligned} N_1 &= \nu(R(f)) - \nu(R(f, f')), \\ N_2 &= \nu(R(f, f')) - \nu(R(f, f', f'')), \\ &\vdots \\ N_{k-1} &= \nu(R(f, f', \dots, f^{(k-2)})) - \nu(R(f, f', \dots, f^{(k-1)})), \\ N_k &= \nu(R(f, f', \dots, f^{(k-1)})) - \nu(R(f, f', \dots, f^{(k)})); \end{aligned}$$

lo que implica

$$\sum_{i=1}^k N_i = \nu(R(f)) - \nu(R(f, f', \dots, f^{(k)})).$$

De aquí, por (4.6) y (4.1),

$$\nu(R(f, f', \dots, f^{(k)})) - \nu(R(f, f', \dots, f^{(k+1)})) = N_{k+1}.$$

En consecuencia, la fórmula para N_k es cierta para todo $k = 1, \dots, n$.

Cuando $1 \leq k < n$, como $\rho_k := N_k - N_{k+1}$, de esta consecuencia se sigue que

$$\begin{aligned} \rho_k &= \nu(R(f, f', \dots, f^{(k-1)})) - \nu(R(f, f', \dots, f^{(k)})) \\ &\quad - \nu(R(f, f', \dots, f^{(k)})) + \nu(R(f, f', \dots, f^{(k+1)})). \end{aligned}$$

Si $k = n$, entonces $\nu(R(f, f', \dots, f^{(n)})) = 0$, pues no existe raíz común a $f, f', \dots, f^{(n)}$; también $\nu(R(f, f', \dots, f^{(n+1)})) = 0$. Asimismo, $N_{n+1} = 0$. Luego,

$$\rho_n = N_n - N_{n+1} = N_n = \nu(R(f, f', \dots, f^{(n-1)})) - \nu(R(f, f', \dots, f^{(n)}));$$

que equivale a

$$\rho_n = \nu(R(f, f', \dots, f^{(n-1)})) - 2 \cdot 0 + 0.$$

Por consiguiente, la fórmula para ρ_k es cierta para todo $k = 1, \dots, n$. □

Corolario 1. Si $f(\lambda) := \lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n$ es un polinomio complejo, entonces el número de raíces distintas de $f(\lambda)$ es

$$n - \nu(R(f, f')). \quad (4.7)$$

Dado que la matriz $R(f, f')$ tiene $2n - 1$ columnas, se sigue que el número de raíces distintas de $f(\lambda)$ es igual a

$$\text{rg } R(f(\lambda), f'(\lambda), \lambda) - n + 1. \quad (4.8)$$

Observación 4.2.1. La matriz $R(f, f')$ es cuadrada de orden $2n - 1$ y, por las notaciones que estamos usando, viene dada por

$$\begin{bmatrix} 1 & a_1 & \cdots & a_{n-1} & a_n & 0 & \cdots & 0 & 0 \\ 0 & 1 & a_1 & \cdots & a_{n-1} & a_n & \cdots & & 0 \\ \vdots & & \ddots & \ddots & & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & a_1 & \cdots & & a_{n-1} & a_n \\ n & (n-1)a_1 & \cdots & \cdots & \cdots & a_{n-1} & \cdots & 0 & 0 \\ 0 & n & (n-1)a_1 & \cdots & \cdots & & a_{n-1} & \cdots & 0 \\ \vdots & & \ddots & \ddots & & & & \ddots & \\ 0 & 0 & \cdots & n & (n-1)a_1 & \cdots & \cdots & & a_{n-1} \end{bmatrix}.$$

Una sucesión finita o infinita de números reales $(x_k)_{k=0,1,\dots}$ se dice *decreciente* si para todo k , $x_k \geq x_{k+1}$; se dice que $(x_k)_{k=0,1,\dots}$ es *convexa* si para todo k el punto de coordenadas $(k+1, x_{k+1})$ está por debajo de la recta que une los puntos (k, x_k) y $(k+2, x_{k+2})$. Dado que la ecuación de la recta que une estos dos puntos es

$$y = \frac{x_{k+2} - x_k}{2}x - \frac{x_{k+2} - x_k}{2} + x_k,$$

la ordenada del punto de esta recta correspondiente a la abscisa $k+1$ es igual a

$$\begin{aligned} & \frac{x_{k+2} - x_k}{2}(k+1) - \frac{x_{k+2} - x_k}{2} + x_k \\ &= \frac{x_{k+2} - x_k}{2} + x_k = \frac{x_{k+2} + x_k}{2}; \end{aligned}$$

por tanto, la sucesión $(x_k)_{k=0,1,\dots}$ es convexa si y sólo si para todo k

$$x_{k+1} \leq \frac{x_{k+2} + x_k}{2};$$

lo que equivale a

$$0 \leq x_{k+2} - 2x_{k+1} + x_k. \quad (4.9)$$

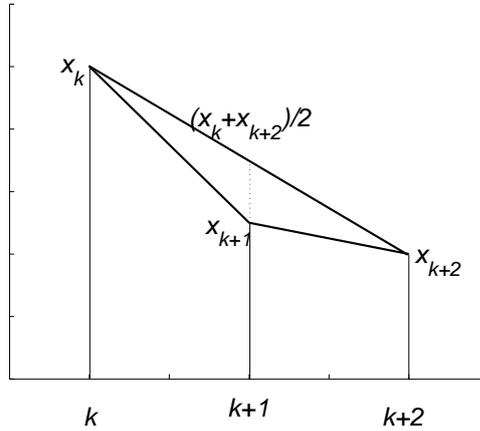


Figura 4.1: Sucesión convexa.

A veces denotaremos por $x(k)$ el término k -ésimo, x_k , de la sucesión x . Las condiciones para que una sucesión sea decreciente o convexa pueden expresarse por medio del **operador diferencial**, Δ , que a una sucesión de números reales x hace corresponder otra sucesión de números reales Δx definida para cada $k = 0, 1, \dots$, por

$$(\Delta x)(k) := x(k+1) - x(k).$$

Aplicando de nuevo el operador Δ a la sucesión Δx se tiene $\Delta^2 x := \Delta(\Delta x)$. Por tanto

$$\begin{aligned} (\Delta^2 x)(k) &= (\Delta(\Delta x))(k) = (\Delta x)(k+1) - (\Delta x)(k) \\ &= x(k+2) - x(k+1) - [x(k+1) - x(k)] = x(k+2) - 2x(k+1) + x(k). \end{aligned}$$

Con estas notaciones se sigue que x es decreciente si y sólo si para todo k ,

$$(\Delta x)(k) \leq 0.$$

Y x es convexa si y sólo si para todo k ,

$$(\Delta^2 x)(k) \geq 0.$$

Resultados completamente análogos a los criterios de decrecimiento y convexidad de una función real de variable real derivable f en un intervalo en términos del signo de sus derivadas f' y f'' , respectivamente.

Corolario 2. La sucesión finita $x_k := \nu(R(f, f', \dots, f^{(k)}))$, para $k = 0, \dots, n$ es decreciente y convexa.

DEMOSTRACIÓN. En efecto, tomando las diferencias primeras y segundas, se tiene que

$$(\Delta x)(k) = -N_{k+1} \leq 0, \quad (\Delta^2 x)(k) = \rho_{k+1} \geq 0, \forall k.$$

□

4.3. Ejercicios

Ejercicio 4.1. Sea $f(\lambda) = a_0 + a_1\lambda + \dots + a_{n-1}\lambda^{n-1} + a_n\lambda^n$ un polinomio de coeficientes complejos y de grado n , cuyas raíces son $\lambda_1, \dots, \lambda_u$ con multiplicidades m_1, \dots, m_u , respectivamente. Sea $\{i_1, \dots, i_v\}$ un subconjunto de $\{1, \dots, u\}$ y sean k_1, \dots, k_v números enteros positivos tales que $k_1 \leq m_{i_1}, \dots, k_v \leq m_{i_v}$. Demostrar que para todo $\varepsilon > 0$ existen números complejos a'_0, a'_1, \dots, a'_n tales que:

- (1) $|a'_0 - a_0| < \varepsilon, |a'_1 - a_1| < \varepsilon, \dots, |a'_n - a_n| < \varepsilon,$
- (2) los números $\lambda_{i_1}, \dots, \lambda_{i_v}$ son raíces del polinomio $g(\lambda) = a'_0 + a'_1\lambda + \dots + a'_{n-1}\lambda^{n-1} + a'_n\lambda^n$ con multiplicidades k_1, \dots, k_v , respectivamente.

Capítulo 5

Continuidad y aproximación

5.1. Continuidad del espectro

Recordemos que la fórmula

$$\|A\| := \sum_{i,j} |a_{ij}|, \quad (5.1)$$

donde $A = (a_{ij}) \in \mathbb{C}^{p \times q}$, define una norma matricial. Esta norma la hemos denotado $\|\cdot\|_1$ en el Capítulo 1. En el capítulo presente omitiremos el subíndice 1 por brevedad. El conjunto $\mathbb{C}^{p \times q}$ es un espacio métrico con la distancia asociada a esta norma.

Si r es un número real positivo y α es un número complejo, por $B(\alpha, r)$ denotaremos la bola abierta en \mathbb{C} de centro α y radio r : $\{z \in \mathbb{C} : |z - \alpha| < r\}$. Si $A \in \mathbb{C}^{p \times q}$, $B_{\mathbb{C}^{p \times q}}(A, r)$ denota la bola abierta en $\mathbb{C}^{p \times q}$ de centro A y radio r : $B_{\mathbb{C}^{p \times q}}(A, r) = \{X \in \mathbb{C}^{p \times q} : \|X - A\| < r\}$. Para cualquier entero $n > 0$, denotamos mediante \underline{n} el conjunto $\{1, \dots, n\}$.

Teorema 5.1.1. *Sea $A \in \mathbb{C}^{n \times n}$ y sea $\Lambda(A) = \{\lambda_1, \lambda_2, \dots, \lambda_u\}$ el conjunto de distintos valores propios de A . Sea $\varepsilon > 0$ un número real tal que los discos*

$$D_1 := B(\lambda_1, \varepsilon), D_2 := B(\lambda_2, \varepsilon), \dots, D_u := B(\lambda_u, \varepsilon)$$

son disjuntos dos a dos. Entonces existe un número $\eta > 0$ tal que para toda $A' \in \mathcal{V} := B_{\mathbb{C}^{n \times n}}(A, \eta)$,

$$\Lambda(A') \subset D_1 \cup D_2 \cup \dots \cup D_u, \quad (5.2)$$

y para todo $j = 1, \dots, u$,

$$\sum_{\mu \in \Lambda(A') \cap B(\lambda_j, \varepsilon)} m(\mu, A') = m(\lambda_j, A). \quad (5.3)$$

DEMOSTRACIÓN. *Primera parte.* Ya que $\Lambda(A) \subset \overline{B(0, \|A\|)}$, los discos D_1, \dots, D_u están contenidos en el disco compacto $D := \overline{B(0, \|A\| + \varepsilon)}$. Sean

$$K := D \setminus (D_1 \cup \dots \cup D_u)$$

y

$$\mathcal{R} := \{(z, X) \in \mathbb{C} \times \mathbb{C}^{n \times n} \mid \det(zI_n - X) \neq 0\}.$$

El subconjunto \mathcal{R} de $\mathbb{C} \times \mathbb{C}^{n \times n}$ es abierto. Sea $z \in K$; ya que $z \notin \Lambda(A)$, tenemos que $(z, A) \in \mathcal{R}$. Por lo tanto, existen un entorno abierto $\Lambda_z \subset \mathbb{C}$ de z y un número real $r_z > 0$ tales que

$$\Lambda_z \times B_{\mathbb{C}^{n \times n}}(A, r_z) \subset \mathcal{R}.$$

Como $\{\Lambda_z\}_{z \in K}$ es un recubrimiento abierto del subconjunto compacto K , existen $z_1, \dots, z_m \in K$ tales que

$$K \subset \Lambda_{z_1} \cup \dots \cup \Lambda_{z_m}.$$

Sea $\eta_1 := \min\{r_{z_1}, \dots, r_{z_m}, \varepsilon\}$ y consideremos cualquier $A' \in B_{\mathbb{C}^{n \times n}}(A, \eta_1)$. Para todo $z \in K$,

$$(z, A') \in \bigcup_{i=1}^m \Lambda_{z_i} \times B_{\mathbb{C}^{n \times n}}(A, r_{z_i}) \subset \mathcal{R};$$

es decir, que $z \notin \Lambda(A')$. En consecuencia, $\Lambda(A') \subset \mathbb{C} \setminus K$. Por otro lado, $\Lambda(A') \subset \overline{B(0, \|A'\|)}$; además

$$\|A'\| \leq \|A' - A\| + \|A\| < \varepsilon + \|A\|.$$

Por consiguiente,

$$\Lambda(A') \subset D \setminus K = D_1 \cup \dots \cup D_u.$$

Esto prueba (5.2).

Segunda parte. Para cada valor propio λ_j de A , ($j = 1, \dots, u$), sean

$$\mu_{j1}, \mu_{j2}, \dots, \mu_{jt_j}$$

los valores propios distintos de A' que hay en $B(\lambda_j, \varepsilon)$. Vamos a demostrar que la suma de las multiplicidades algebraicas de $\mu_{j1}, \mu_{j2}, \dots, \mu_{jt_j}$ es igual a la multiplicidad algebraica de λ_j . Para ello puede ser preciso elegir un número η_1 aun más pequeño. Fijemos nuestra atención en una matriz muy grande asociada a λ_j :

$$P_{\lambda_j, n^2}(A) := \begin{pmatrix} \lambda_j I - A & 0 & \dots & 0 \\ I & \lambda_j I - A & & \vdots \\ & I & \lambda_j I - A & \\ & & \ddots & \ddots \\ 0 & & & \lambda_j I - A & 0 \\ & 0 & \dots & I & \lambda_j I - A \end{pmatrix},$$

que es una matriz de tamaño $n^3 \times n^3$; en esta matriz el número λ_j aparece escrito n^2 veces. Elijamos ahora dos números δ_{1j}, δ_{2j} positivos tales que para todos $A' \in \mathbb{C}^{n \times n}$ y $\mu_1, \dots, \mu_{n^2} \in \mathbb{C}$ que satisfagan

$$\|A' - A\| < \delta_{1j}, |\mu_1 - \lambda_j| < \delta_{2j}, \dots, |\mu_{n^2} - \lambda_j| < \delta_{2j},$$

se tenga que el rango de la matriz $n^3 \times n^3$

$$\begin{pmatrix} \mu_1 I - A' & 0 & \dots & 0 \\ I & \mu_2 I - A' & & \vdots \\ & I & \mu_3 I - A' & \\ & & \ddots & \ddots \\ 0 & & & \mu_{n^2-1} I - A' & 0 \\ & 0 & \dots & I & \mu_{n^2} I - A' \end{pmatrix},$$

sea mayor o igual que el de $P_{\lambda_j, n^2}(A)$.

Sean

$$\delta_1 := \min_{1 \leq j \leq u} \delta_{1j}, \quad \delta_2 := \min_{1 \leq j \leq u} \delta_{2j};$$

si las bolas $B(\lambda_j, \delta_2)$, $j = 1, \dots, u$ no son disjuntas a pares, elegimos un δ_3 , $0 < \delta_3 < \min\{\delta_2, \varepsilon\}$, para que las bolas $B(\lambda_j, \delta_3)$ si lo sean. De acuerdo con lo probado en la Primera parte, podemos elegir un $\eta > 0$ y menor que δ_1 y η_1 tal que $A' \in \mathbb{C}^{n \times n}$, $\|A' - A\| < \eta$ implique que

$$\Lambda(A') \subset \bigcup_{j=1}^u B(\lambda_j, \delta_3).$$

Tomemos $\mathcal{V} := B_{\mathbb{C}^{n \times n}}(A, \eta)$; entonces para cada $A' \in \mathcal{V}$ sean $\mu_{j1}, \mu_{j2}, \dots, \mu_{jt_j}$ los valores propios de A' que están en la bola $B(\lambda_j, \delta_3), j = 1, \dots, u$ (obviamente, $t_j \leq n$). Por (3.2) de la Sección 3.5, (3.5) de la Sección 3.7 y (3.3) de la Sección 3.11 del Capítulo 3 se tiene que

$$\nu\left(P_{\mu_{j1}, \dots, \mu_{jt_j}}^{n, \dots, n}(A')\right) = m(\mu_{j1}, A') + \dots + m(\mu_{jt_j}, A'), \quad (5.4)$$

donde la matriz

$$P_{\mu_{j1}, \dots, \mu_{jt_j}}^{n, \dots, n}(A')$$

viene definida en (3.2) de la Sección 3.11 del Capítulo 3, particularizando $q := t_j, k_1 := n, \dots, k_q := n$ y $\alpha_1 := \mu_{j1}, \dots, \alpha_q := \mu_{jt_j}$. Escogemos ahora un número complejo cualquiera μ que no sea valor propio de A' y que satisfaga la desigualdad $|\mu - \lambda_j| < \delta_3$, entonces por la forma de elegir δ_3 ,

$$\nu\left(P_{\mu_{j1}, \dots, \mu_{jt_j}, \underbrace{\mu, \dots, \mu}_{n-t_j}}^{n, \dots, n}(A')\right) \leq \nu\left(P_{\lambda_j, n^2}(A)\right). \quad (5.5)$$

Pero, mediante transformaciones elementales se puede probar que

$$\nu\left(P_{\mu_{j1}, \dots, \mu_{jt_j}, \underbrace{\mu, \dots, \mu}_{n-t_j}}^{n, \dots, n}(A')\right) = \nu\left(P_{\mu_{j1}, \dots, \mu_{jt_j}}^{n, \dots, n}(A')\right). \quad (5.6)$$

Por consiguiente, por (5.4), (5.5) y (5.6),

$$m(\mu_{j1}, A') + \dots + m(\mu_{jt_j}, A') \leq \nu\left(P_{\lambda_j, n^2}(A)\right). \quad (5.7)$$

Pero por (3.5) de la Sección 3.7 del Capítulo 3,

$$\nu\left(P_{\lambda_j, n^2}(A)\right) = m(\lambda_j, A). \quad (5.8)$$

De (5.7) y (5.8) deducimos

$$m(\mu_{j1}, A') + \dots + m(\mu_{jt_j}, A') \leq m(\lambda_j, A); \quad (5.9)$$

que resumimos así

$$\sum_{s=1}^{t_j} m(\mu_{js}, A') \leq m(\lambda_j, A), \quad j = 1, \dots, u. \quad (5.10)$$

Sumando las u desigualdades (5.10) se obtiene

$$n = \sum_{j=1}^u \sum_{s=1}^{t_j} m(\mu_{js}, A') \leq \sum_{j=1}^u m(\lambda_j, A) = n, \quad (5.11)$$

pues la suma de las multiplicidades algebraicas de todos los valores propios de una matriz $n \times n$ es igual a n . Por lo tanto, todas las desigualdades (5.10) deben ser igualdades

$$\sum_{s=1}^{t_j} m(\mu_{js}, A') = m(\lambda_j, A), \quad j = 1, \dots, u,$$

ya que si al menos una desigualdad de (5.10) fuera estricta se seguiría por (5.11) que $n < n$, lo que sería absurdo. \square

Observación 5.1.1. Acabamos de ver en la demostración del Teorema 5.1.1 que la inclusión (5.2) es válida para un valor de $\eta > 0$ y que la igualdad (5.3) es cierta para un segundo valor de η , quizá más pequeño. Vamos a demostrar a continuación, utilizando teoremas de funciones de una variable compleja, que (5.3) es válida para el primer valor de η .

De hecho, tomando el valor óptimo de ε ; es decir,

$$\varepsilon_0 := \min_{i \neq j} \frac{1}{2} |\lambda_i - \lambda_j|,$$

se tiene que la igualdad de multiplicidades (5.3) es válida para el primer η_0 asociado a este ε_0 . Esto es, aquel $\eta_0 > 0$ tal que $\|A' - A\| < \eta_0$ implica que

$$\Lambda(A') \subset B(\lambda_1, \varepsilon_0) \cup \cdots \cup B(\lambda_u, \varepsilon_0). \quad (5.12)$$

Teorema 5.1.2. *Sea $A \in \mathbb{C}^{n \times n}$ con espectro $\Lambda(A) = \{\lambda_1, \dots, \lambda_u\}$, $u \leq n$. Sea*

$$\varepsilon_0 := \min_{i \neq j} \frac{1}{2} |\lambda_i - \lambda_j|, \quad i, j \in \{1, \dots, u\}.$$

Supongamos que el número real η_0 tiene la propiedad de que $\|A' - A\| < \eta_0$ implica que

$$\Lambda(A') \subset B(\lambda_1, \varepsilon_0) \cup \cdots \cup B(\lambda_u, \varepsilon_0).$$

Si $\|A' - A\| < \eta_0$, entonces para cada $j = 1, \dots, u$ se satisface también la igualdad

$$\sum_{\mu \in \Lambda(A') \cap B(\lambda_j, \varepsilon)} m(\mu, A') = m(\lambda_j, A). \quad (5.13)$$

DEMOSTRACIÓN. Para cada $j = 1, \dots, u$, sea γ_j la circunferencia de centro λ_j y radio ε_0 recorrida en sentido positivo. Para toda matriz $X \in \mathbb{C}^{n \times n}$, definamos por

$$f_X(\lambda) := \det(\lambda I_n - X)$$

el polinomio característico de X . Como los coeficientes de f_X se pueden expresar como suma de menores principales de la matriz X , se tiene que la aplicación $(\lambda, X) \mapsto f_X(\lambda)$ es continua. Sea cualquier matriz $A' \in \mathbb{C}^{n \times n}$ tal que $\|A' - A\| < \eta_0$. Como la bola $B_{\mathbb{C}^{n \times n}}(A, \eta_0)$ es un conjunto convexo (véase[41, Exercise 2, p. 356]), se tiene que el segmento $\{X(t) := A + t(A' - A), t \in [0, 1]\}$ está contenido en dicha bola. Por el Corolario 2 del Teorema D.1.1 del Apéndice D y el Teorema D.2.1 del Apéndice D, para cada $t \in [0, 1]$, se tiene que

$$N_j(t) := \frac{1}{2\pi i} \int_{\gamma_j} \frac{f'_{X(t)}(\lambda)}{f_{X(t)}(\lambda)} d\lambda$$

nos da el número de valores propios de la matriz $X(t)$ que están contenidos en la bola $B(\lambda, \varepsilon_0)$, contando sus multiplicidades. Dicho de otra forma,

$$N_j(t) = \sum_{\mu \in \Lambda(X(t)) \cap B(\lambda_j, \varepsilon_0)} m(\mu, X(t)).$$

Pero $N_j(0) = m(\lambda_j, A)$ y

$$N_j(1) = \sum_{\mu \in \Lambda(A') \cap B(\lambda_j, \varepsilon_0)} m(\mu, A'),$$

además, la función $t \mapsto N_j(t)$ es continua, con valores enteros y el intervalo $[0, 1]$ es un conjunto conexo; por lo tanto, $N_j(0) = N_j(1)$. \square

5.1.1. El espacio métrico de n -tuplas desordenadas

Definición 5.1.1. Decimos que dos n -tuplas

$$(z_1, z_2, \dots, z_n) \text{ y } (w_1, w_2, \dots, w_n)$$

de \mathbb{C}^n son equivalentes si existe una permutación σ de $\{1, \dots, n\}$ tal que

$$w_i = z_{\sigma(i)}, \quad (i = 1, \dots, n).$$

En cuyo caso escribimos

$$(z_1, z_2, \dots, z_n) \sim (w_1, w_2, \dots, w_n).$$

Es fácil comprobar que la relación \sim en \mathbb{C}^n es una relación de equivalencia.

Definición 5.1.2. Llamaremos *n-tupla desordenada de números complejos* a cualquier clase de equivalencia de \sim ; es decir, a todo elemento del conjunto cociente \mathbb{C}^n/\sim . La clase de equivalencia correspondiente a la *n-tupla* $(z_1, z_2, \dots, z_n) \in \mathbb{C}^n$ será denotada por $[(z_1, z_2, \dots, z_n)]$.

Proposición 5.1.1. *Un sistema completo de invariantes para \sim está formado por las n sumas de Newton*

$$s_k(x_1, x_2, \dots, x_n) := x_1^k + x_2^k + \dots + x_n^k,$$

$k = 1, \dots, n$.

Observación 5.1.2. Es obvio que otro sistema completo de invariantes para la relación \sim está constituido por las *n* funciones simétricas elementales. Puede verse la relación explícita que hay entre las sumas de Newton y las funciones simétricas elementales en el Ejercicio 14, § 33, página 608 de Godement [30].

Definición 5.1.3. Sean $\alpha := [(\alpha_1, \alpha_2, \dots, \alpha_n)]$ y $\beta := [(\beta_1, \beta_2, \dots, \beta_n)]$; $\alpha, \beta \in \mathbb{C}^n/\sim$. Llamamos *distancia del emparejamiento óptimo* entre α y β al número

$$d(\alpha, \beta) := \min_{\sigma \in S_n} \max_{1 \leq i \leq n} |\alpha_i - \beta_{\sigma(i)}| \tag{5.14}$$

donde S_n es el grupo de permutaciones de $\{1, \dots, n\}$.

Para abreviar introducimos la notación $\underline{n} := \{1, \dots, n\}$. Con ella, $d(\alpha, \beta)$ se puede expresar así

$$\min_{\sigma \in S_n} \max_{i \in \underline{n}} |\alpha_i - \beta_{\sigma(i)}|.$$

Ejemplo 5.1.1. Sean

$$\begin{aligned} (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7) &:= (3, 3, 5, 5, 7, 8, 3), \\ (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7) &:= (2'9, 5'3, 3'2, 3'1, 8'2, 7'1, 4'9); \end{aligned}$$

entonces

$$\begin{aligned} |\alpha_1 - \beta_1| &= |3 - 2'9| = 0'1, \\ |\alpha_2 - \beta_2| &= |3 - 5'3| = 2'3, \\ |\alpha_3 - \beta_3| &= |5 - 3'2| = 1'8, \\ |\alpha_4 - \beta_4| &= |5 - 3'1| = 1'9, \\ |\alpha_5 - \beta_5| &= |7 - 8'2| = 1'2, \\ |\alpha_6 - \beta_6| &= |8 - 7'1| = 0'9, \\ |\alpha_7 - \beta_7| &= |3 - 4'9| = 1'9, \end{aligned}$$

por lo que

$$\max_{1 \leq i \leq 7} |\alpha_i - \beta_i| = 2'3.$$

Pero, si permutamos $(\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6)$ de acuerdo con la permutación τ de $\{1, \dots, 7\}$ tal que

$$(\beta_{\tau(1)}, \beta_{\tau(2)}, \beta_{\tau(3)}, \beta_{\tau(4)}, \beta_{\tau(5)}, \beta_{\tau(6)}, \beta_{\tau(7)}) = (2'9, 3'2, 5'3, 4'9, 7'1, 8'2, 3'1),$$

se sigue que

$$\max_{1 \leq i \leq 7} |\alpha_i - \beta_{\tau(i)}| = 0'3.$$

Es fácil verificar que, en este ejemplo,

$$d([\alpha_1, \dots, \alpha_7], [(\beta_1, \dots, \beta_7)]) = 0'3.$$

Lema 5.1.1. *La cantidad $d(\alpha, \beta)$ no depende de los representantes elegidos $(\alpha_1, \alpha_2, \dots, \alpha_n) \in \alpha$ y $(\beta_1, \beta_2, \dots, \beta_n) \in \beta$. Además, d es una distancia en \mathbb{C}^n/\sim ; esto es,*

$$\begin{cases} d(\alpha, \beta) \geq 0, \\ d(\alpha, \beta) = 0 \Leftrightarrow \alpha = \beta, \\ d(\alpha, \beta) = d(\beta, \alpha), \\ d(\alpha, \beta) \leq d(\alpha, \gamma) + d(\gamma, \beta), \end{cases}$$

para todos $\alpha, \beta, \gamma \in \mathbb{C}^n/\sim$.

DEMOSTRACIÓN. Veamos en primer lugar que $d(\alpha, \beta)$ está bien definida. Sean

$$(\alpha_1, \dots, \alpha_n), (\beta_1, \dots, \beta_n) \in \mathbb{C}^n$$

y $\sigma_1, \sigma_2 \in S_n$; definamos

$$\begin{aligned} (\alpha'_1, \dots, \alpha'_n) &:= (\alpha_{\sigma_1(1)}, \dots, \alpha_{\sigma_1(n)}), \\ (\beta'_1, \dots, \beta'_n) &:= (\beta_{\sigma_2(1)}, \dots, \beta_{\sigma_2(n)}). \end{aligned}$$

De estas notaciones se siguen las identidades

$$\forall i \in \underline{n}, \quad \alpha_i = \alpha'_{\sigma_1^{-1}(i)}, \quad \beta_i = \beta'_{\sigma_2^{-1}(i)}; \quad (5.15)$$

en efecto, para cada $i \in \underline{n}$ sean $j := \sigma_1^{-1}(i)$ y $k := \sigma_2^{-1}(i)$. De aquí que

$$\alpha'_j = \alpha_{\sigma_1(j)} = \alpha_{\sigma_1(\sigma_1^{-1}(i))} = \alpha_i$$

y

$$\beta'_k = \beta_{\sigma_2(k)} = \beta_{\sigma_2(\sigma_2^{-1}(i))} = \beta_i.$$

Demostraremos que

$$\min_{\sigma \in S_n} \max_{i \in \underline{n}} |\alpha_i - \beta_{\sigma(i)}| = \min_{\sigma \in S_n} \max_{i \in \underline{n}} |\alpha'_i - \beta'_{\sigma(i)}| \quad (5.16)$$

Sea $\pi \in S_n$ una permutación donde se alcanza el mínimo del LHS de (5.16); esto es,

$$\min_{\sigma \in S_n} \max_{i \in \underline{n}} |\alpha_i - \beta_{\sigma(i)}| = \max_{i \in \underline{n}} |\alpha_i - \beta_{\pi(i)}|$$

Por (5.15) se tiene que

$$|\alpha_i - \beta_{\pi(i)}| = \left| \alpha'_{\sigma_1^{-1}(i)} - \beta'_{\sigma_2^{-1}(\pi(i))} \right|;$$

llamando $\tau := \sigma_2^{-1} \circ \pi \circ \sigma_1$ se sigue que

$$\tau(\sigma_1^{-1}(i)) = (\sigma_2^{-1} \circ \pi)(i);$$

por consiguiente

$$|\alpha_i - \beta_{\pi(i)}| = \left| \alpha'_{\sigma_1^{-1}(i)} - \beta'_{\sigma_2^{-1}(\pi(i))} \right| = |\alpha'_j - \beta'_{\tau(j)}|,$$

y los conjuntos

$$\{|\alpha_1 - \beta_{\pi(1)}|, \dots, |\alpha_n - \beta_{\pi(n)}|\}$$

y

$$\{|\alpha'_1 - \beta'_{\tau(1)}|, \dots, |\alpha'_n - \beta'_{\tau(n)}|\}$$

son iguales. Por lo tanto,

$$\max_{i \in \underline{n}} |\alpha_i - \beta_{\pi(i)}| = \max_{i \in \underline{n}} |\alpha'_i - \beta'_{\tau(i)}|.$$

Como τ es un elemento de S_n , deducimos que

$$\max_{i \in \underline{n}} |\alpha_i - \beta_{\pi(i)}| \geq \min_{\sigma \in S_n} \max_{i \in \underline{n}} |\alpha'_i - \beta'_{\sigma(i)}|.$$

De donde,

$$\min_{\sigma \in S_n} \max_{i \in \underline{n}} |\alpha_i - \beta_{\sigma(i)}| \geq \min_{\sigma \in S_n} \max_{i \in \underline{n}} |\alpha'_i - \beta'_{\sigma(i)}|. \quad (5.17)$$

De forma análoga se prueba que

$$\min_{\sigma \in S_n} \max_{i \in \underline{n}} |\alpha'_i - \beta'_{\sigma(i)}| \geq \min_{\sigma \in S_n} \max_{i \in \underline{n}} |\alpha_i - \beta_{\sigma(i)}|. \quad (5.18)$$

Por (5.17) y (5.18) se sigue que

$$\min_{\sigma \in S_n} \max_{i \in \underline{n}} |\alpha_i - \beta_{\sigma(i)}| = \min_{\sigma \in S_n} \max_{i \in \underline{n}} |\alpha'_i - \beta'_{\sigma(i)}|.$$

Lo que implica que

$$d([\alpha_1, \dots, \alpha_n], [\beta_1, \dots, \beta_n]) = d([\alpha'_1, \dots, \alpha'_n], [\beta'_1, \dots, \beta'_n]).$$

Ahora demostraremos que d es una distancia en \mathbb{C}^n / \sim . Sean

$$\alpha := [\alpha_1, \dots, \alpha_n], \beta := [\beta_1, \dots, \beta_n] \in \mathbb{C}^n / \sim.$$

Como $|z| \geq 0$ para todo $z \in \mathbb{C}$, es obvio que $d(\alpha, \beta) \geq 0$. Veamos a continuación por qué $d(\alpha, \beta) = 0$ si y sólo si $\alpha = \beta$: En primer lugar, $d(\alpha, \alpha) = 0$, pues entre las permutaciones que σ recorre en (5.14) está la permutación identidad id . En segundo lugar, si $d(\alpha, \beta) = 0$ se tiene que

$$\min_{\sigma \in S_n} \max_{i \in \underline{n}} |\alpha_i - \beta_{\sigma(i)}| = 0. \quad (5.19)$$

Sea $\sigma_0 \in S_n$ una permutación para la que se alcanza el mínimo de (5.19); entonces

$$\max_{i \in \underline{n}} |\alpha_i - \beta_{\sigma_0(i)}| = 0.$$

Por tanto, $\forall i \in \underline{n}$, $|\alpha_i - \beta_{\sigma_0(i)}| = 0$; de donde

$$(\alpha_1, \dots, \alpha_n) \sim (\beta_1, \dots, \beta_n),$$

lo que implica $\alpha = \beta$.

Ahora probaremos que $d(\alpha, \beta) = d(\beta, \alpha)$. Sea $\sigma_0 \in S_n$ tal que

$$d(\alpha, \beta) = \max_{i \in \underline{n}} |\alpha_i - \beta_{\sigma_0(i)}|$$

Como $\forall i \in \underline{n}$, $i = \sigma_0^{-1}(\sigma_0(i))$, se tiene que $\alpha_i = \alpha_{\sigma_0^{-1}(\sigma_0(i))}$. Por consiguiente,

$$\begin{aligned} d(\alpha, \beta) &= \max_{i \in \underline{n}} \left| \alpha_{\sigma_0^{-1}(\sigma_0(i))} - \beta_{\sigma_0(i)} \right| \\ &= \max_{j \in \underline{n}} |\alpha_{\sigma_0^{-1}(j)} - \beta_j| \geq \min_{\sigma \in S_n} \max_{j \in \underline{n}} |\beta_j - \alpha_{\sigma(j)}| = d(\beta, \alpha); \end{aligned}$$

lo que implica que $d(\alpha, \beta) \geq d(\beta, \alpha)$. Por simetría $d(\beta, \alpha) \geq d(\alpha, \beta)$, y, por consiguiente, $d(\alpha, \beta) = d(\beta, \alpha)$.

Finalmente, probaremos la desigualdad triangular $d(\alpha, \beta) \leq d(\alpha, \gamma) + d(\gamma, \beta)$ para cualesquiera $\alpha, \beta, \gamma \in \mathbb{C}^n / \sim$. Supongamos que $\gamma = [\gamma_1, \dots, \gamma_n]$. Sea $\pi \in S_n$ tal que

$$d(\alpha, \gamma) = \max_{i \in \underline{n}} |\alpha_i - \gamma_{\pi(i)}|;$$

para cada $\sigma \in S_n$ y para cada $i \in \underline{n}$ se tiene que

$$|\alpha_i - \beta_{\sigma(i)}| \leq |\alpha_i - \gamma_{\pi(i)}| + |\gamma_{\pi(i)} - \beta_{\sigma(i)}|;$$

de donde, tomando mínimos al recorrer i el conjunto \underline{n} , se infiere que

$$\begin{aligned} \max_{i \in \underline{n}} |\alpha_i - \beta_{\sigma(i)}| &\leq \max_{i \in \underline{n}} (|\alpha_i - \gamma_{\pi(i)}| + |\gamma_{\pi(i)} - \beta_{\sigma(i)}|) \\ &\leq \max_{i \in \underline{n}} |\alpha_i - \gamma_{\pi(i)}| + \max_{i \in \underline{n}} |\gamma_{\pi(i)} - \beta_{\sigma(i)}|. \end{aligned}$$

Sabemos que si $f: A \rightarrow \mathbb{R}$ es una función que alcanza su mínimo en un conjunto cualquiera A y si $c \in \mathbb{R}$ es una constante, entonces $\min_{x \in A} (c + f(x)) = c + \min_{x \in A} f(x)$. Por lo tanto, como la cantidad

$$\max_{i \in \underline{n}} |\alpha_i - \gamma_{\pi(i)}|$$

no depende de $\sigma \in S_n$, se deduce de las desigualdades anteriores que

$$\min_{\sigma \in S_n} \max_{i \in \underline{n}} |\alpha_i - \beta_{\sigma(i)}| \leq \max_{i \in \underline{n}} |\alpha_i - \gamma_{\pi(i)}| + \min_{\sigma \in S_n} \max_{i \in \underline{n}} |\gamma_{\pi(i)} - \beta_{\sigma(i)}|.$$

Por lo cual, $d(\alpha, \beta) \leq d(\alpha, \gamma) + d(\gamma, \beta)$, pues $\gamma = [(\gamma_{\pi(1)}, \dots, \gamma_{\pi(n)})]$. □

Al conjunto \mathbb{C}^n / \sim se le llama *producto simétrico de \mathbb{C} por sí mismo n veces* y se denota por $\mathbb{C}_{\text{sim}}^n$.

Notación. Si una matriz $A \in \mathbb{C}^{n \times n}$ tiene los valores propios

λ_1 con multiplicidad algebraica m_1

λ_2 con multiplicidad algebraica m_2

\vdots

λ_u con multiplicidad algebraica m_u ,

siendo $m_1 + m_2 + \dots + m_u = n$, a la matriz A le asociamos la n -tupla desordenada

$$\text{Eig } A := [(\overbrace{\lambda_1, \dots, \lambda_1}^{m_1}, \overbrace{\lambda_2, \dots, \lambda_2}^{m_2}, \dots, \overbrace{\lambda_u, \dots, \lambda_u}^{m_u})].$$

El resultado siguiente expresa la continuidad del espectro de una matriz en términos del espacio métrico $(\mathbb{C}_{\text{sim}}^n, d)$.

Teorema 5.1.3. *La aplicación*

$$A \rightarrow \text{Eig } A$$

de $\mathbb{C}^{n \times n}$ en $\mathbb{C}_{\text{sim}}^n$ es continua.

Observación 5.1.3. Sean $A \in \mathbb{C}^{n \times n}$ y $\Lambda(A) = \{\lambda_1, \dots, \lambda_u\}$ el conjunto de sus valores propios. Supongamos que $\delta > 0$ es un número real tal que las bolas abiertas $B(\lambda_i, \delta)$ y $B(\lambda_j, \delta)$ son disjuntas para todos $i, j \in \{1, \dots, u\}$, con $i \neq j$. Si A' es cualquier otra matriz de $\mathbb{C}^{n \times n}$ tal que

$$d(\text{Eig } A, \text{Eig } A') < \delta,$$

entonces

$$\Lambda(A') \subset \cup_{i=1}^u B(\lambda_i, \delta).$$

Sea $A(t)$ una matriz $n \times n$ que depende continuamente del parámetro t . Entonces resulta que la función

$$t \rightarrow \text{Eig } A(t)$$

es continua, por el Teorema 5.1.3 y teniendo en cuenta que la composición de dos aplicaciones continuas es una aplicación continua. Dicho de manera más formal, sea E un espacio topológico y

$$\begin{aligned} A: E &\rightarrow \mathbb{C}^{n \times n} \\ t &\mapsto A(t) \end{aligned}$$

una función matricial continua. Entonces la función

$$\begin{aligned} A: E &\rightarrow \mathbb{C}_{\text{sim}}^n \\ t &\mapsto \text{Eig } A(t) \end{aligned}$$

es continua en cada punto $t_0 \in E$. Esto significa que la distancia de $\text{Eig } A(t)$ a $\text{Eig } A(t_0)$ tiende a cero cuando $t \rightarrow t_0$, para cada t_0 fijado. La continuidad —así formulada— es la continuidad de los valores propios repetidos *como un todo*. es una cuestión diferente saber si es posible o no definir una “parametrización continua” de los valores propios repetidos de $A(t)$; es decir, ¿es posible definir n funciones complejas continuas

$$\mu_1: E \rightarrow \mathbb{C}, \dots, \mu_n: E \rightarrow \mathbb{C},$$

tales que para cada $t \in E$,

$$(\mu_1(t), \dots, \mu_n(t)) \in \text{Eig } A(t)?$$

En general, la respuesta es no. Por ejemplo, la función matricial

$$A(z) := \begin{pmatrix} 0 & z \\ 1 & 0 \end{pmatrix}, \quad z \in \mathbb{C},$$

tiene por valores propios a

$$\mu_1(z) = z^{1/2}, \mu_2(z) = -z^{1/2},$$

que no son funciones continuas en $z = 0$. Sin embargo, la respuesta a la cuestión anterior es afirmativa en los casos siguientes:

1. cuando t varía en un intervalo de la recta real [39, pág. 126, Theorem 5.2], [12, pág. 155];
2. cuando para todo $t \in E$ los valores propios de $A(t)$ son números reales (por ejemplo, esto ocurre si la matriz $A(t)$ es hermítica [53, pág. 39]);
3. cuando para todo $t \in E$ los valores propios de $A(t)$ están situados en una misma recta del plano complejo (por ejemplo, esto ocurre si la matriz $A(t)$ es antihermítica, cuyos valores propios son imaginarios puros);
4. cuando para todo $t \in E$ los valores propios de $A(t)$ están situados en una circunferencia (por ejemplo, esto sucede si $A(t)$ es unitaria [5, pág. 465]).

5.2. Matrices genéricas

Teorema 5.2.1. *Sea $A \in \mathbb{C}^{n \times n}$ y $\varepsilon > 0$ dados. Entonces existe una matriz $B \in \mathbb{C}^{n \times n}$ que tiene sus valores propios simples y tal que $\|B - A\| < \varepsilon$.*

Teorema 5.2.2. *El subconjunto de $\mathbb{C}^{n \times n}$ formado por las matrices que tienen sus valores propios simples es abierto.*

Teorema 5.2.3. *El subconjunto de $\mathbb{C}^{n \times n}$ formado por las matrices no derogatorias es abierto y denso.*

5.3. Criterio local de semejanza

Vimos en el Corolario 1 del Teorema 3.8.1 que dos matrices $A, B \in \mathbb{C}^{n \times n}$ son semejantes si y sólo si

$$\nu(A, A) = \nu(A, B) = \nu(B, B).$$

Vamos a ver que la condición $\nu(A, A) = \nu(A, B)$ es suficiente para garantizar la semejanza cuando la matriz B está suficientemente cerca de A .

Teorema 5.3.1. *Sea $A \in \mathbb{C}^{n \times n}$, entonces existe un entorno \mathcal{V} de A en el espacio $\mathbb{C}^{n \times n}$, tal que si $A' \in \mathcal{V}$ y*

$$\nu(A, A) = \nu(A, A'),$$

las matrices A y A' son semejantes.

DEMOSTRACIÓN. Vectorizando la ecuación matricial $AX - XA = O$, por el Lema 3.8.1, resulta

$$(A \otimes I_n - I_n \otimes A^T)x = 0, \quad (5.1)$$

siendo $x = \text{vec}(X)$. Esto es,

$$x = (x_{11}, x_{12}, \dots, x_{1n}, x_{21}, x_{22}, \dots, x_{2n}, \dots, x_{n1}, x_{n2}, \dots, x_{nn})^T$$

Supongamos que $\text{rg}(A \otimes I_n - I_n \otimes A^T) = r$. Esto significa que existen sucesiones estrictamente crecientes α, β de r elementos de $\underline{n} \times \underline{n}$ tales que

$$|(A \otimes I_n - I_n \otimes A^T)[\alpha|\beta]| \neq 0.$$

Por un argumento de continuidad se puede probar que existe un entorno \mathcal{U} de A en $\mathbb{C}^{n \times n}$ tal que si $A' \in \mathcal{U}$, entonces

$$|(A \otimes I_n - I_n \otimes A'^T)[\alpha|\beta]| \neq 0.$$

De ahora en adelante supondremos que además de ser $A' \in \mathcal{U}$, se tiene que $\nu(A, A) = \nu(A, A')$; lo que es equivalente a

$$\text{rg}(A \otimes I_n - I_n \otimes A^T) = \text{rg}(A \otimes I_n - I_n \otimes A'^T) = r.$$

Tengamos ahora presente la notación preparada en (3.10). Suprimiendo en el sistema de ecuaciones lineales (5.1) las filas que no corresponden a los índices del conjunto α queda un sistema de r ecuaciones con n^2 incógnitas, que es equivalente al de partida; esto es, ambos tienen las mismas soluciones. Después, pasamos al segundo miembro de las ecuaciones los términos asociados a las incógnitas x_{hk} cuyo índice (h, k) pertenece al conjunto β^c , complementario de β en $\underline{n} \times \underline{n}$. Resulta así un sistema de Cramer de r ecuaciones (las de índice $(i, j) \in \alpha$) con r incógnitas (las de índice $(h, k) \in \beta$).

A continuación demos los valores $x_{hk} := \delta_{hk}$ (deltas de Kronecker) a las incógnitas de los segundos miembros; es decir, a aquellas x_{hk} tales que $(h, k) \in \beta^c$. Como $X = I_n$ satisface la ecuación $AX - XB = O$, se tiene que la única solución del sistema de Cramer son los valores $x_{hk} = \delta_{hk}$ para cada $(h, k) \in \beta$.

Si consideramos el sistema

$$(A \otimes I_n - I_n \otimes A'^T)x = 0, \quad (5.2)$$

puede reducirse a un sistema de Cramer de r ecuaciones con r incógnitas. Asignando los mismo valores que antes a las incógnitas presentes en los segundos miembros de las ecuaciones, se tiene una solución única, $x_{hk} = x'_{hk}$, $(h, k) \in \beta$, dada por la regla de Cramer. De este modo, obtenemos una única solución $\mathcal{X}(A')$ de la ecuación $AX - XA' = O$. La función \mathcal{X} definida para toda $A' \in \mathcal{U}$ tal que $\nu(A, A) = \nu(A, A')$ es continua en el punto A ; como $\mathcal{X}(A) = I_n$, tenemos que existe un entorno $\mathcal{V} \subset \mathcal{U}$ de A tal que $|\mathcal{X}(A')| \neq 0$ para todo $A' \in \mathcal{V}$ tal que $\nu(A, A) = \nu(A, A')$. Por consiguiente, $A' \approx A$. \square

5.4. Conjetura de Sylvester aproximada

En la Sección 3.9 del Capítulo 3 vimos un teorema llamado la Conjetura de Sylvester (véase la página 69) que dice que si una matriz cuadrada A es no derogatoria y una matriz B conmuta con A , entonces $B = p(A)$ siendo $p(\lambda)$ un polinomio. Vamos a ver una generalización sorprendente de este hecho, válida para matrices A cualesquiera.

Teorema 5.4.1. *Sea $\|\cdot\|$ una norma matricial cualquiera en el espacio $\mathbb{C}^{n \times n}$. Sean $A, B \in \mathbb{C}^{n \times n}$ matrices que satisfacen la igualdad $AB = BA$. Entonces para cada $\varepsilon > 0$ existen una matriz $A_\varepsilon \in \mathbb{C}^{n \times n}$ y un polinomio $p_\varepsilon(\lambda)$ tales que*

$$\|A_\varepsilon - A\| < \varepsilon \quad y \quad B = p_\varepsilon(A_\varepsilon).$$

5.5. Ejercicios

Ejercicio 5.1. Probar que en todo entorno de una matriz $A \in \mathbb{C}^{n \times n}$ hay una matriz distinta de A que tiene su mismo espectro.

Ejercicio 5.2. Sea $A \in \mathbb{C}^{n \times n}$ una matriz dada y sea $(A_p)_{p=0}^\infty$ una sucesión de matrices de $\mathbb{C}^{n \times n}$ que converge hacia A . ¿Existe una subsucesión $(A_{p_k})_{k=0}^\infty$ que converge hacia A y que está formada por matrices que tienen todas el mismo espectro?

Ejercicio 5.3. Sean $A \in \mathbb{C}^{n \times n}$ y $B \in \mathbb{C}^{m \times m}$ tales que $\Lambda(A) \cap \Lambda(B) = \emptyset$. Demostrar que existen entornos \mathcal{V}_A de A en $\mathbb{C}^{n \times n}$ y \mathcal{V}_B de B en $\mathbb{C}^{m \times m}$ tales que $\forall A' \in \mathcal{V}_A, \forall B' \in \mathcal{V}_B$ se tenga que $\Lambda(A') \cap \Lambda(B') = \emptyset$.

Ejercicio 5.4. Sean $A \in \mathbb{C}^{n \times n}, B \in \mathbb{C}^{m \times m}$ y $C \in \mathbb{C}^{q \times q}$ tales que $\Lambda(A) \cap \Lambda(B) \cap \Lambda(C) = \emptyset$. Demostrar que existen entornos \mathcal{V}_A de A en $\mathbb{C}^{n \times n}, \mathcal{V}_B$ de B en $\mathbb{C}^{m \times m}$ y \mathcal{V}_C de C en $\mathbb{C}^{q \times q}$ tales que $\forall A' \in \mathcal{V}_A, \forall B' \in \mathcal{V}_B, \forall C' \in \mathcal{V}_C$ se tenga que $\Lambda(A') \cap \Lambda(B') \cap \Lambda(C') = \emptyset$.

Ejercicio 5.5. Sean $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ y $x = (x_1, \dots, x_n) \in \mathbb{C}^n$. (i) Probar que podemos elegir x de manera que los números x_1, \dots, x_n sean distintos y los valores propios de

$$\begin{pmatrix} a_{11} + x_1 & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} + x_n \end{pmatrix} \tag{5.1}$$

sean simples. (ii) Demostrar que para todo $\varepsilon > 0$ existe un $x \in \mathbb{C}^n$ tal que $\|x\|_2 < \varepsilon$, los números x_1, \dots, x_n son distintos y la matriz (5.1) tiene los valores propios simples. (iii) ¿Son válidos (i) y (ii) cambiando en todos los sitios el cuerpo \mathbb{C} por el cuerpo \mathbb{R} de los números reales? ¿o por el cuerpo \mathbb{Q} de los números racionales? .

Indicación. Utilizar la desigualdad

$$d(\text{Eig } A, \text{Eig } B) \leq (2n - 1)\|A - B\|,$$

donde la matriz cuadrada A es normal (i.e. $A^*A = AA^*$) y B es una matriz cuadrada arbitraria; $\|\cdot\|$ es la norma espectral; $\text{Eig } A$ y $\text{Eig } B$ son las n -tuplas desordenadas de valores propios de A y B , respectivamente, y d es la distancia del emparejamiento óptimo. Véase (8), pág. 4 de [11].

Ejercicio 5.6. Sea A una matriz cuadrada compleja y sea λ_0 un valor propio de A de multiplicidad algebraica $m(\lambda_0, A) = q$. Sea k un entero positivo tal que $k < q$. Demostrar que tan cerca como se quiera de A existe una matriz X tal que $m(\lambda_0, X) = k$.

Indicación. Sin pérdida de generalidad se puede suponer que la matriz A está en la forma canónica de Jordan.

Ejercicio 5.7. Sea \mathbb{K} uno cualquiera de los cuerpos \mathbb{Q}, \mathbb{R} o \mathbb{C} . Sea $\mathbb{K}^{n \times n}$ el conjunto de todas las matrices $n \times n$ sobre \mathbb{K} . Consideremos las funciones $c_i: \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$ definidas por

$$\det(\lambda I - A) = \lambda^n - c_1(A)\lambda^{n-1} + c_2(A)\lambda^{n-2} - \dots + (-1)^n c_n(A)$$

para toda $A \in \mathbb{K}^{n \times n}$. Recordemos que $c_1(A) = \text{tr } A$ y $c_n(A) = \det A$. Sea $f: \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$ una función continua en todo $\mathbb{K}^{n \times n}$ que es un invariante para la relación de semejanza, i.e. $f(A) = f(B)$ siempre que las matrices A y B de $\mathbb{K}^{n \times n}$ sean semejantes. Demostrar que existe una única función continua $F: \mathbb{K}^n \rightarrow \mathbb{K}$ tal que

$$f(A) = F(c_1(A), \dots, c_n(A))$$

para toda $A \in \mathbb{K}^{n \times n}$.

Ejercicio 5.8. Sea \mathbb{K} uno cualquiera de los cuerpos \mathbb{Q}, \mathbb{R} o \mathbb{C} . Sea $\mathbb{K}^{n \times n}$ el conjunto de todas las matrices $n \times n$ sobre \mathbb{K} . Recordemos que una forma canónica para la relación de semejanza (\approx) en $\mathbb{K}^{n \times n}$ es cualquier función $c: \mathbb{K}^{n \times n} \rightarrow \mathbb{K}^{n \times n}$ tal que: (1) para toda $A \in \mathbb{K}^{n \times n}$, $A \approx c(A)$; (2) si $A, B \in \mathbb{K}^{n \times n}$, $A \approx B$ si y sólo si $c(A) = c(B)$. Demostrar haciendo uso del Ejercicio 5.7 que no existe ninguna forma canónica que sea continua en toda matriz de $\mathbb{K}^{n \times n}$.

Ejercicio 5.9. Sea A una matriz real $n \times n$ cuyos valores propios son reales y simples. Demostrar que existe un entorno \mathcal{V} de A , $\mathcal{V} \subset \mathbb{R}^{n \times n}$, tal que toda matriz $A' \in \mathcal{V}$ tiene sus valores propios simples y reales. Mostrar un ejemplo en el que falle esta conclusión cuando los valores propios de A son reales pero alguno de ellos es múltiple.

Ejercicio 5.10. Sean A, A' dos matrices fijas de $\mathbb{C}^{n \times n}$. Denotemos por $d(\text{Eig } A, \text{Eig } A')$ la distancia del emparejamiento óptimo entre $\text{Eig } A$ y $\text{Eig } A'$. Demostrar que

$$d(\text{Eig } A, \text{Eig } A') = \min\{\varepsilon \geq 0 \mid \Lambda(A') \subset \bigcup_{\alpha \in \Lambda(A)} \mathcal{D}(\alpha, \varepsilon)\}$$

siendo $\mathcal{D}(\alpha, \varepsilon)$ el disco cerrado en \mathbb{C} de centro α y radio ε .

5.6. Notas al Capítulo 5

La demostración del Teorema 5.1.1 elude el teorema de Rouché de funciones de variable compleja, en contra de lo que suele ser habitual; está basada en la semicontinuidad de la función rango, las características de Segré y de Weyr, y la forma canónica de Jordan; no depende de haber establecido previamente los teoremas sobre las condiciones necesarias de la perturbación de la forma de Jordan que se verán en la Sección 6.1 del Capítulo 6. Esta demostración se puede trasladar, sin más, para probar que las raíces de un polinomio dependen con continuidad de sus coeficientes, por medio de las matrices compañeras; por tanto, también aquí podría evitarse el recurso al teorema de Rouché a la hora de demostrar que la función de \mathbb{C}^n en $\mathbb{C}_{\text{sim}}^n$ que envía la n -tupla (a_1, \dots, a_n) de los coeficientes del polinomio $z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n$ a la n -tupla desordenada de sus raíces $(\alpha_1, \dots, \alpha_n)$, es continua [12, pág. 153]. La demostración del Teorema 5.3.1 está inspirada en la del libro de Friedland [Theorem 2.9.5][26]. El Teorema 5.4.1 es debido a Holbrook[36].

Capítulo 6

Perturbación de la forma de Jordan

En este capítulo se estudian las condiciones que necesariamente deben satisfacer las formas canónicas de Jordan de todas las matrices suficientemente próximas a una matriz compleja cuadrada dada A . En este contexto todas las matrices circundantes a A están en pie de igualdad. Podríamos considerar solamente las matrices que además de estar suficientemente cercanas a A cumplieran otras restricciones; por ejemplo estar en una curva que pasa por A en el espacio de matrices; o tener siempre determinados elementos iguales a los correspondientes elementos de A (perturbación estructurada); esta problemática es objeto de intensa investigación en la actualidad, y no puede ser considerada aquí.

También veremos que las condiciones necesarias aludidas arriba son tales que en todo entorno de A hay matrices cuya forma de Jordan las satisface.

6.1. Condiciones necesarias

Recordemos que la fórmula

$$\|A\| := \sum_{i,j} |a_{ij}|, \quad (6.1)$$

donde $A = (a_{ij}) \in \mathbb{C}^{p \times q}$, define una norma matricial. Esta norma la hemos denotado $\|\cdot\|_1$ en el Capítulo 1. En el capítulo presente omitiremos el subíndice 1 por brevedad. El conjunto $\mathbb{C}^{p \times q}$ es un espacio métrico con la distancia asociada a esta norma. Si r es un número real positivo y α es un número complejo, por $B(\alpha, r)$ notaremos la bola abierta en \mathbb{C} de centro α y radio r : $\{z \in \mathbb{C} : |z - \alpha| < r\}$. Para toda matriz $M \in \mathbb{C}^{n \times n}$ y para cada valor propio λ_0 de M , denotaremos por $w(\lambda_0, M)_j$ la componente j -ésima de la partición de Weyr $w(\lambda_0, M)$. Véase la Definición 3.4.1 del Capítulo 3. En general, omitiremos escribir los términos no nulos de las particiones y daremos éstas como tuplas. Para cualquier matriz $M \in \mathbb{C}^{n \times n}$, si μ_1, \dots, μ_v son sus valores propios distintos ($v \leq n$), llamaremos *número adecuado* a la matriz M a cualquier número real $\eta > 0$ tal que

$$\eta < \frac{1}{2} \min_{\substack{i \neq j \\ i, j \in \{1, \dots, v\}}} |\mu_i - \mu_j|;$$

Esta condición implica que las bolas $B(\mu_i, \eta)$, $i \in \{1, \dots, v\}$ son disjuntas dos a dos. Cuando $\eta > 0$ sea un número adecuado a M llamaremos **η -entorno del espectro de M** al conjunto abierto

$$V_\eta(M) := \bigcup_{i=1}^v B(\mu_i, \eta).$$

Recordemos que una matriz $A \in \mathbb{C}^{n \times n}$ es nilpotente si $A^k = 0$ para cierto entero positivo k . Además, se puede probar que A es nilpotente si y sólo si 0 es su único valor propio. Véase la Proposición A.6.1 del Apéndice A.

Teorema 6.1.1. Sean $A \in \mathbb{C}^{n \times n}$ y $\Lambda(A) = \{\lambda_1, \dots, \lambda_u\}$ el conjunto de los valores propios distintos de A . Sea $\eta > 0$ un número adecuado a A . Sea α un valor propio de A . Entonces existe un entorno \mathcal{V} de A en $\mathbb{C}^{n \times n}$ tal que $A' \in \mathcal{V}$ implica que

$$(i) \quad \Lambda(A') \subset \bigcup_{i=1}^u B(\lambda_i, \eta); \quad (6.2)$$

(ii) si $\mu_1, \mu_2, \dots, \mu_t$ son los valores propios de A' que están dentro de $B(\alpha, \eta)$, entonces

$$\bigcup_{i=1}^t w(\mu_i, A') \prec\prec w(\alpha, A). \quad (6.3)$$

DEMOSTRACIÓN. Sea k un número fijo de $\{1, \dots, n\}$. Por la semicontinuidad inferior de la función rango rg (Proposición 1.3.1 del Capítulo 1) en la matriz $(\alpha I - A)^k$, existe un entorno \mathcal{U}_k de esta matriz en $\mathbb{C}^{n \times n}$ tal que para toda $X \in \mathcal{U}_k$ tenemos que

$$\text{rg} [(\alpha I - A)^k] \leq \text{rg } X. \quad (6.4)$$

La continuidad de la suma y del producto de números (y de matrices) nos permite determinar la existencia de un entorno Λ_k de α en \mathbb{C} y otro entorno \mathcal{V}_k de A en $\mathbb{C}^{n \times n}$ tales que si $z_i \in \Lambda_k$ para $i = 1, \dots, k$ y $A' \in \mathcal{V}_k$, entonces

$$\prod_{i=1}^k (z_i I - A') \in \mathcal{U}_k. \quad (6.5)$$

Pongamos

$$\Lambda := \left(\bigcap_{k=1}^n \Lambda_k \right) \cap B(\alpha, \eta), \quad \mathcal{V}_0 := \bigcap_{k=1}^n \mathcal{V}_k.$$

Por la continuidad de los valores propios de una matriz (Teorema 5.1.1 del Capítulo 5), existe un entorno \mathcal{U} de A en $\mathbb{C}^{n \times n}$ tal que $A' \in \mathcal{U}$ implica (6.2) y que los valores propios de A' en $B(\alpha, \eta)$ están contenidos en Λ . Definamos $\mathcal{V} := \mathcal{U} \cap \mathcal{V}_0$; sea $A' \in \mathcal{V}$ y sean μ_1, \dots, μ_t los valores propios de A' que hay en $B(\alpha, \eta)$. De (6.4) y (6.5) deducimos que

$$\text{rg} [(\alpha I - A)^k] \leq \text{rg} \prod_{i=1}^k (z_i I - A'), \quad (6.6)$$

para cada $z_i \in \{\mu_1, \dots, \mu_t\}$, $i = 1, \dots, k$, $k = 1, \dots, n$.

Sean $w(\mu_1, A'), \dots, w(\mu_t, A')$ las particiones de Weyr de μ_1, \dots, μ_t , respecto de A' . Exactamente igual que en la demostración del Lema 2.3.2 del Capítulo 2, si

$$\bigcup_{i=1}^t w(\mu_i, A') = (d_1, d_2, \dots),$$

para cada $k \in \{1, \dots, n\}$ tenemos que

$$\sum_{i=1}^k d_i = \sum_{j=1}^{p_1} w(\mu_{i_1}, A')_j + \dots + \sum_{j=1}^{p_q} w(\mu_{i_q}, A')_j,$$

donde $1 \leq i_1 < \dots < i_q \leq t$, $p_1 + \dots + p_q = k$. Entonces, por (6.6),

$$\operatorname{rg} [(\alpha I - A)^k] \leq \operatorname{rg} \left[\prod_{s=1}^q (\mu_{i_s} I - A')^{p_s} \right]. \quad (6.7)$$

Por las igualdades (3.2) de la página 51 que anteceden a la definición de la partición de Weyr,

$$\sum_{j=1}^{p_s} w(\mu_{i_s}, A')_j = \nu [(\mu_{i_s} I - A')^{p_s}], \quad (s = 1, \dots, q).$$

Así pues,

$$\sum_{s=1}^q \nu [(\mu_{i_s} I - A')^{p_s}] = \sum_{i=1}^k d_i. \quad (6.8)$$

Por el Lema 3.10.1 del Capítulo 3,

$$\sum_{s=1}^q \nu [(\mu_{i_s} I - A')^{p_s}] = \nu \left[\prod_{s=1}^q (\mu_{i_s} I - A')^{p_s} \right]. \quad (6.9)$$

De (6.7), (6.8) y (6.9), deducimos que

$$\nu [(\alpha I - A)^k] \geq \sum_{i=1}^k d_i,$$

y, por las igualdades citadas (3.2) de la página 51, esto implica que

$$\sum_{i=1}^k w(\alpha, A)_i \geq \sum_{i=1}^k d_i.$$

Esta desigualdad vale para todo $k \in \{1, \dots, n\}$. Por el Teorema 3.5.2 del Capítulo 3 y las igualdades (3.4) de la página 56 y (3.2) de la página 53, deducimos que $\ell(w(\alpha, A)) \leq n$ y $\ell(\cup_{i=1}^t w(\mu_i, A')) \leq n$. Por consiguiente se sigue (6.3). \square

Teorema 6.1.2. Sean $A \in \mathbb{C}^{n \times n}$ y $\Lambda(A) = \{\lambda_1, \dots, \lambda_u\}$ el conjunto de los valores propios distintos de A . Sea $\eta > 0$ un número tal que $B(\lambda_i, \eta) \cap B(\lambda_j, \eta) = \emptyset$ si $i \neq j$, $i, j \in \{1, \dots, u\}$.

Entonces existe un entorno \mathcal{V} de A en $\mathbb{C}^{n \times n}$ tal que $A' \in \mathcal{V}$ implica que

(i)

$$\Lambda(A') \subset \bigcup_{i=1}^u B(\lambda_i, \eta);$$

(ii) para todo $i = 1, \dots, u$ si $\mu_{i1}, \dots, \mu_{it_i}$ son los valores propios de A' que hay en $B(\lambda_i, \eta)$, entonces

$$\bigcup_{j=1}^{t_i} w(\mu_{ij}, A') \prec w(\lambda_i, A)$$

o equivalentemente

$$s(\lambda_i, A) \prec \sum_{j=1}^{t_i} s(\mu_{ij}, A').$$

DEMOSTRACIÓN. Por el Teorema 6.1.1, para cada $i \in \{1, \dots, u\}$ existe un entorno \mathcal{U}_i de A en $\mathbb{C}^{n \times n}$ tal que si $A' \in \mathcal{U}_i$, se cumple (i) y que

$$\bigcup_{j=1}^{t_i} w(\mu_{ij}, A') \prec\prec w(\lambda_i, A). \quad (6.10)$$

Llamando $\mathcal{V} := \bigcap_{i=1}^u \mathcal{U}_i$, por (6.10) deducimos que si $A' \in \mathcal{V}$, entonces para cada $i \in \{1, \dots, u\}$,

$$\left| \bigcup_{j=1}^{t_i} w(\mu_{ij}, A') \right| \leq |w(\lambda_i, A)|;$$

o por el Ejercicio 2.1 del Capítulo 2, se tiene

$$\sum_{j=1}^{t_i} |w(\mu_{ij}, A')| \leq |w(\lambda_i, A)|; \quad (6.11)$$

y teniendo en cuenta que la suma de los pesos de las particiones de la característica de Weyr de una matriz cuadrada es igual a su orden, obtenemos

$$n = \sum_{i=1}^u \sum_{j=1}^{t_i} |w(\mu_{ij}, A')| \leq \sum_{i=1}^u |w(\lambda_i, A)| = n. \quad (6.12)$$

Por lo tanto,

$$\sum_{i=1}^u \sum_{j=1}^{t_i} |w(\mu_{ij}, A')| = \sum_{i=1}^u |w(\lambda_i, A)|; \quad (6.13)$$

o, lo que es equivalente,

$$\sum_{i=1}^u \left[|w(\lambda_i, A)| - \sum_{j=1}^{t_i} |w(\mu_{ij}, A')| \right] = 0; \quad (6.14)$$

pero por (6.11) para cada $i = 1, \dots, u$,

$$\left[|w(\lambda_i, A)| - \sum_{j=1}^{t_i} |w(\mu_{ij}, A')| \right] \geq 0;$$

por consiguiente,

$$\left[|w(\lambda_i, A)| - \sum_{j=1}^{t_i} |w(\mu_{ij}, A')| \right] = 0;$$

lo que implica

$$\left| \bigcup_{j=1}^{t_i} w(\mu_{ij}, A') \right| = |w(\lambda_i, A)|;$$

por la Observación 2.2.1 del Capítulo 2, se tiene que

$$\bigcup_{j=1}^{t_i} w(\mu_{ij}, A') \prec w(\lambda_i, A), \quad (i = 1, \dots, u); \quad (6.15)$$

Así pues, queda demostrada la primera parte de (ii).

Tomando particiones conjugadas en la desigualdad (6.15), por el Lema 2.3.1 del Capítulo 2 y el Teorema 3.5.2 del Capítulo 3, se sigue para particiones de Segré que

$$s(\lambda_i, A) \prec \sum_{j=1}^{t_i} s(\mu_{ij}, A'), \quad (i = 1, \dots, u). \quad (6.16)$$

Con esto queda demostrada la segunda parte de (ii). \square

Observación 6.1.1. Teniendo en cuenta el Ejercicio 2.16 del Capítulo 2, si $A' \in \mathcal{V}$, por (6.10) se tiene que para todo $i = 1, \dots, u$

$$\bigcup_{j=1}^{t_i} w(\mu_{ij}, A') \prec w(\lambda_i, A),$$

si y sólo si

$$\bigcup_{i=1}^u \bigcup_{j=1}^{t_i} w(\mu_{ij}, A') \prec \bigcup_{i=1}^u w(\lambda_i, A).$$

También podemos decir que existe un entorno \mathcal{V} de A tal que $A' \in \mathcal{V}$ implica que

$$\bigcup_{\mu \in \Lambda(A')} w(\mu, A') \prec \bigcup_{\alpha \in \Lambda(A)} w(\alpha, A); \quad (6.17)$$

hablando en términos metafóricos podríamos decir que esta desigualdad expresa una especie de semicontinuidad superior de la aplicación

$$X \mapsto \bigcup_{\xi \in \Lambda(X)} w(\xi, X)$$

de $\mathbb{C}^{n \times n}$ en el conjunto de las particiones \mathcal{P} respecto de la relación de orden \prec .

De igual forma, por dualidad, de (6.17) deducimos para las particiones de Segré que $A' \in \mathcal{V}$ implica que

$$\sum_{\alpha \in \Lambda(A)} s(\alpha, A) \prec \sum_{\mu \in \Lambda(A')} s(\mu, A'); \quad (6.18)$$

lo que da lugar a una suerte de aplicación semicontinua inferiormente

$$X \mapsto \sum_{\xi \in \Lambda(X)} s(\xi, X)$$

de $\mathbb{C}^{n \times n}$ en \mathcal{P} respecto de \prec . Aunque debe ser bien entendido que no hemos definido ninguna topología sobre \mathcal{P} .

Corolario 1. Sean $A \in \mathbb{C}^{n \times n}$, $\Lambda(A) = \{\lambda_1, \dots, \lambda_u\}$ y $\eta > 0$ tal que las bolas $B(\lambda_i, \eta)$, $\lambda_i \in \Lambda(A)$, sean disjuntas dos a dos. Entonces existe un entorno \mathcal{V} de A tal que $A' \in \mathcal{V}$ implica que $\Lambda(A') \subset V_\eta(A)$ y

$$\sum_{\mu \in \Lambda(A') \cap B(\lambda_i, \eta)} m(\mu, A') = m(\lambda_i, A), \quad (i = 1, \dots, u).$$

6.2. El problema inverso subyacente

Se llaman problemas directos a los que consisten en hallar los valores de parámetros naturales asociados a un objeto matemático; verbigracia, hallar los valores propios de una matriz cuadrada, hallar las raíces de una ecuación, hallar las soluciones de una ecuación diferencial. Mientras que un problema inverso estriba en la reconstrucción total o parcial del objeto matemático (matriz, ecuación, etc.) dados unos valores de sus parámetros naturales (valores propios, raíces, funciones solución, etc.)

Dada la matriz $A \in \mathbb{C}^{n \times n}$, en esta sección veremos la solución de un problema inverso de matrices consistente en la construcción de una matriz A' tan próxima como se desee a A y de la que hayamos prefijado el número de valores propios que rodean a cada valor propio de A y sus particiones de Weyr. En general, los problemas inversos de matrices sólo tienen naturaleza algebraica, dado que se prescriben condiciones algebraicas de sus parámetros. En

nuestro caso, existe el matiz topológico de la perturbación. Por este carácter de problema inverso algo oculto, lo tildamos de subyacente.

El planteamiento preciso del problema inverso abordado y una respuesta se sustancian en un teorema que es casi-recíproco del Teorema 6.1.2. La demostración del teorema que sigue se hará utilizando la forma canónica dual para la semejanza de matrices asociada a la característica de Weyr (Definición 3.6.1 del Capítulo 3).

Teorema 6.2.1. Sean $A \in \mathbb{C}^{n \times n}$ y $\Lambda(A) = \{\lambda_1, \dots, \lambda_u\}$ el conjunto de los valores propios distintos de A . Sea $\eta > 0$ un número tal que $B(\lambda_i, \eta) \cap B(\lambda_j, \eta) = \emptyset$ si $i \neq j, i, j \in \{1, \dots, u\}$. Para cada $i = 1, \dots, u$ sea t_i un entero positivo dado y sean b_{i1}, \dots, b_{it_i} particiones no nulas dadas.

En cualquier entorno de A existe una matriz A' tal que

(a)

$$\Lambda(A') \subset \bigcup_{i=1}^u B(\lambda_i, \eta),$$

(b) para todo $i = 1, \dots, u$, la matriz A' tiene exactamente t_i valores propios $\mu_{i1}, \dots, \mu_{it_i}$ en la bola $B(\lambda_i, \eta)$, y

$$b_{ij} = w(\mu_{ij}, A') \quad (j = 1, \dots, t_i),$$

si y sólo si

$$\bigcup_{j=1}^{t_i} b_{ij} \prec w(\lambda_i, A) \quad (6.1)$$

para todo $i = 1, \dots, u$.

Antes de probar este teorema, daremos cinco lemas previos que simplifican la demostración de la parte “si” del mismo.

Lema 6.2.1. Sean $A_1, A_2 \in \mathbb{C}^{n \times n}$ matrices semejantes; sea $\Lambda(A_1) = \{\lambda_1, \dots, \lambda_u\}$. Sea $\eta > 0$ tal que $B(\lambda_i, \eta) \cap B(\lambda_j, \eta) = \emptyset$ si $i \neq j, i, j \in \{1, \dots, u\}$. Para cada $i = 1, \dots, u$, sea t_i un entero positivo dado y sean b_{i1}, \dots, b_{it_i} particiones no nulas dadas tales que

$$\bigcup_{j=1}^{t_i} b_{ij} \prec w(\lambda_i, A_1).$$

Entonces las dos afirmaciones siguientes son equivalentes.

(a) Para cualquier $\varepsilon_1 > 0$ existe una matriz A'_1 que satisface $\|A'_1 - A_1\| < \varepsilon_1$ tal que

(a.1) $\Lambda(A'_1) \subset \bigcup_{i=1}^u B(\lambda_i, \eta)$,

(a.2) para todo $i = 1, \dots, u$, la matriz A'_1 tiene exactamente t_i valores propios $\mu_{i1}^{(1)}, \dots, \mu_{it_i}^{(1)}$ en $B(\lambda_i, \eta)$ y

$$b_{ij} = w(\mu_{ij}^{(1)}, A'_1), \quad j = 1, \dots, t_i.$$

(b) Para cualquier $\varepsilon_2 > 0$ existe una matriz A'_2 que satisface $\|A'_2 - A_2\| < \varepsilon_2$ tal que

(b.1) $\Lambda(A'_2) \subset \bigcup_{i=1}^u B(\lambda_i, \eta)$,

(b.2) para todo $i = 1, \dots, u$, la matriz A'_2 tiene exactamente t_i valores propios $\mu_{i1}^{(2)}, \dots, \mu_{it_i}^{(2)}$ en $B(\lambda_i, \eta)$ y

$$b_{ij} = w(\mu_{ij}^{(2)}, A'_2), \quad j = 1, \dots, t_i.$$

DEMOSTRACIÓN. Sea P una matriz $n \times n$ invertible tal que $A_2 = P^{-1}A_1P$. Supongamos que (a) es cierta. Tomemos un $\varepsilon_2 > 0$ cualquiera. Definamos

$$\varepsilon_1 := \frac{\varepsilon_2}{\|P\|\|P^{-1}\|}.$$

Por (a) existe una matriz A'_1 que satisface $\|A'_1 - A_1\| < \varepsilon_1$ y (a.1) y (a.2). Definamos $A'_2 := P^{-1}A'_1P$. Por tanto,

$$\|A'_2 - A_2\| \leq \|P^{-1}\|\|P\|\|A'_1 - A_1\| < \|P^{-1}\|\|P\|\varepsilon_1 = \varepsilon_2.$$

Además, como A'_2 es semejante a A'_1 , entonces A'_2 cumple las condiciones (b.1) y (b.2). Por consiguiente, se satisface (b). La demostración de que (b) implica (a) es análoga. \square

Lema 6.2.2. *Sea cualquier $\delta > 0$. Supongamos probado que para una matriz $A \in \mathbb{C}^{n \times n}$ con $u = 1$, la condición (6.1) es suficiente para garantizar que en todo entorno de A existe una matriz A' que satisface (a) y (b) del Teorema 6.2.1 con $\eta = \delta$. Entonces dicha condición también es suficiente para garantizar la existencia de una tal A' para cualquier matriz A con $u > 1$.*

DEMOSTRACIÓN. Sea $A \in \mathbb{C}^{n \times n}$ con espectro $\Lambda(A) = \{\lambda_1, \dots, \lambda_u\}$ y $u > 1$; llamemos η a un número real positivo que satisfaga que $B(\lambda_i, \eta) \cap B(\lambda_j, \eta) = \emptyset$ si $i \neq j, i, j \in \{1, \dots, u\}$. Utilizando la reducción de A a su forma canónica de Jordan, podemos asegurar que existe una matriz invertible $P \in \mathbb{C}^{n \times n}$ tal que

$$P^{-1}AP = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_u \end{pmatrix}$$

donde para cada $i = 1, \dots, u$, $\Lambda(A_i) = \{\lambda_i\}$; en efecto, basta tomar cada A_i como la matriz diagonal por bloques cuyos bloques diagonales recogen todos los bloques de Jordan asociados al valor propio λ_i de A .

Supongamos que para cada $i = 1, \dots, u$,

$$\bigcup_{j=1}^{t_i} b_{ij} \prec w(\lambda_i, A);$$

entonces,

$$\bigcup_{j=1}^{t_i} b_{ij} \prec w(\lambda_i, P^{-1}AP),$$

pues las particiones de Weyr son invariantes para la relación de semejanza. De aquí que, por el Ejercicio 3.4 del Capítulo 3, para cada $i = 1, \dots, u$,

$$\bigcup_{j=1}^{t_i} b_{ij} \prec w(\lambda_i, A_i);$$

Sea dado ahora un $\varepsilon > 0$; como A_i tiene un solo valor propio, λ_i , por la hipótesis en la bola

$$B\left(A_i, \frac{\varepsilon}{\|P\|\|P^{-1}\|u}\right) \subset \mathbb{C}^{m_i \times m_i}, \quad \text{con } m_i := m(\lambda_i, A),$$

existe una matriz A'_i tal que

$$(I) \quad \Lambda(A'_i) \subset B(\lambda_i, \eta), \text{ y}$$

- (II) A'_i tiene exactamente t_i valores propios $\mu_{i1}, \mu_{i2}, \dots, \mu_{it_i}$ en la bola $B(\lambda_i, \eta)$, con $b_{ij} = w(\mu_{ij}, A'_i)$, para $j = 1, \dots, t_i$.

Definamos ahora la matriz

$$A' = P \begin{pmatrix} A'_1 & 0 & \cdots & 0 \\ 0 & A'_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A'_u \end{pmatrix} P^{-1};$$

entonces

$$\|A' - A\| \leq \|P\| \|P^{-1}\| \sum_{i=1}^u \|A'_i - A_i\| < \|P\| \|P^{-1}\| \frac{\varepsilon u}{\|P\| \|P^{-1}\| u} = \varepsilon.$$

Además,

$$\Lambda(A') = \bigcup_{i=1}^u \Lambda(A'_i);$$

luego, por (I), (II) y el Ejercicio 3.4 del Capítulo 3, para todo $i = 1, \dots, u$,

$$\Lambda(A') \cap B(\lambda_i, \eta) = \{\mu_{i1}, \mu_{i2}, \dots, \mu_{it_i}\}$$

con

$$b_{ij} = w(\mu_{ij}, A'_i) = w(\mu_{ij}, A')$$

para todo $j = 1, \dots, t_i$. □

Lema 6.2.3. *Sea cualquier $\delta > 0$. Supongamos probado que para una matriz $A \in \mathbb{C}^{n \times n}$ con $\Lambda(A) = \{0\}$, la condición (6.1) es suficiente para garantizar que en todo entorno de A existe una matriz A' que satisface (a) y (b) del Teorema 6.2.1 con $\eta = \delta$. Entonces dicha condición también es suficiente para garantizar la existencia de una tal A' para cualquier matriz A con $\Lambda(A) = \{\lambda_1\}$ y $\lambda_1 \neq 0$.*

DEMOSTRACIÓN. Si $\Lambda(A) = \{\lambda_1\}$ y $\lambda_1 \neq 0$, definimos $B := \lambda_1 I - A$; por el Ejercicio 3.5 del Capítulo 3 se tiene que $\Lambda(B) = \{0\}$ y $w(0, B) = w(\lambda_1, A)$. Para cada $\varepsilon > 0$, por la hipótesis, existe una matriz B' tal que $\|B' - B\| < \varepsilon$ donde B' tiene t_1 valores propios $\mu_{11}, \dots, \mu_{1t_1}$ en la bola $B(0, \delta)$ y

$$b_{1j} = w(\mu_{1j}, B'), \quad j = 1, \dots, t_1.$$

Si definimos $A' := \lambda_1 I - B'$, se sigue que

$$\|A' - A\| = \|\lambda_1 I - B' - (\lambda_1 I - B)\| = \|B' - B\| < \varepsilon,$$

y, por el Ejercicio 3.6 del Capítulo 3, $\lambda_1 - \mu_{11}, \dots, \lambda_1 - \mu_{1t_1}$ son exactamente los valores propios que A' tiene en $B(\lambda_1, \delta)$ y

$$w(\lambda_1 - \mu_{ij}, A') = w(\mu_{ij}, B')$$

lo que implica

$$w(\lambda_1 - \mu_{ij}, A') = b_{1j}, \quad j = 1, \dots, t_1. \quad \square$$

Para comprender el enunciado del lema siguiente es conveniente referirse a la Definición 2.4.1 del Capítulo 2 de transformación elemental de una partición.

Lema 6.2.4. *Sea $J \in \mathbb{C}^{n \times n}$ nilpotente en forma de Jordan con $a := w(0, J)$. Sea b_1 una partición tal que b_1 se transforma en a por una transformación elemental de particiones. Supongamos probado que en tal caso existe una matriz nilpotente A' tan próxima como se desee a J tal que*

$$b_1 = w(0, A').$$

Entonces para toda partición b tal que $b \prec a, b \neq a$, existe una matriz nilpotente \tilde{A} tan próxima como queramos de J tal que

$$b = w(0, \tilde{A}).$$

DEMOSTRACIÓN. Por el Lema 2.4.2 del Capítulo 2 existe una sucesión finita de transformaciones elementales de particiones, digamos de q transformaciones, que convierte b en a . Sean estas particiones así obtenidas y numeradas al revés

$$b \prec b_q \prec b_{q-1} \prec \cdots \prec b_2 \prec b_1 \prec a,$$

donde para cada desigualdad \prec la partición que queda a la derecha del signo procede de la partición a la izquierda del signo mediante una transformación elemental.

Sea dado un $\varepsilon > 0$; por la hipótesis existe una matriz nilpotente A'_1 tal que

$$\|A'_1 - J\| < \frac{\varepsilon}{q}, \quad b_1 = w(0, A'_1). \quad (6.2)$$

Sea J_1 la forma canónica de Jordan de A'_1 ; entonces hay una matriz invertible P_1 tal que $A'_1 = P_1 J_1 P_1^{-1}$, y, por ende, $w(0, J_1) = b_1$. Volvemos pues a la situación de partida; aplicando la hipótesis a la desigualdad $b_2 \prec b_1$ existe una matriz A_2 tal que

$$\|A_2 - J_1\| < \frac{\varepsilon}{q \|P_1\| \|P_1^{-1}\|}, \quad w(0, A_2) = b_2;$$

si definimos $A'_2 := P_1 A_2 P_1^{-1}$ se sigue que

$$\|A'_2 - A'_1\| = \|P_1 A_2 P_1^{-1} - P_1 J_1 P_1^{-1}\| \leq \|P_1\| \|P_1^{-1}\| \|A_2 - J_1\| < \frac{\varepsilon}{q}; \quad (6.3)$$

además, $b_2 = w(0, A'_2)$. Por (6.2) y (6.3),

$$\|A'_2 - J\| \leq \|A'_2 - A'_1\| + \|A'_1 - J\| < 2\frac{\varepsilon}{q}.$$

De este modo, supongamos probada por inducción la existencia de una matriz nilpotente A'_k tal que

$$w(0, A'_k) = b_k, \quad \|A'_k - J\| < k\frac{\varepsilon}{q}.$$

Sea J_k la forma de Jordan de A'_k ; entonces existe una matriz invertible P_k tal que

$$A'_k = P_k J_k P_k^{-1};$$

y $w(0, J_k) = b_k$. Por la hipótesis existe una matriz nilpotente A_{k+1} tal que

$$b_{k+1} = w(0, A_{k+1}), \quad \|A_{k+1} - J_k\| < \frac{\varepsilon}{q \|P_k\| \|P_k^{-1}\|}.$$

Llamemos $A'_{k+1} := P_k A_{k+1} P_k^{-1}$; de donde

$$\begin{aligned} \|A'_{k+1} - A'_k\| &\leq \|P_k\| \|P_k^{-1}\| \|A_{k+1} - J_k\| < \frac{\varepsilon}{q} \\ \|A'_{k+1} - A'_k\| &< k\frac{\varepsilon}{q} + \frac{\varepsilon}{q} = (k+1)\frac{\varepsilon}{q} \\ b_{k+1} &= w(0, A'_{k+1}). \end{aligned}$$

Sea A' la matriz obtenida a partir de A , sustituyendo ε en vez de 0 en el lugar de la fila $r - s + 2$ y la columna 1. Entonces A' tiene como único valor propio 0 y $w(0, A')$ es b ya que

$$\begin{aligned} \operatorname{rg}(-A')^{j-1} &= \operatorname{rg}(-A)^{j-1} = r + s - 2(j-1) \\ &\text{para } j = 1, \dots, s, \\ \operatorname{rg}(-A')^{s-1+j} &= 1 + \operatorname{rg}(-A)^{s-1+j} = r + s - 2(s-1) - j \\ &\text{para } j = 1, \dots, r - s + 1, \\ \operatorname{rg}(-A')^{r+1} &= \operatorname{rg}(-A)^{r+1} = 0. \end{aligned}$$

Téngase en cuenta que $-A' = \lambda_1 I - A'$ y $-A = \lambda_1 I - A$, por ser $\lambda_1 = 0$. Ya que $\varepsilon \neq 0$ puede tomarse tan pequeño como se quiera, esta matriz A' tiene las propiedades requeridas.

Ejemplo 6.2.1. Sean $\bar{a} = (3, 2) \prec (4, 1) = \bar{b}$; es decir $r = 3, s = 2$. En este caso $b = (2, 1, 1, 1) \prec (2, 2, 1) = a$ y

$$A = \left[\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

Tomemos

$$A' := \left[\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \varepsilon & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

Entonces $w(0, A') = (a'_1, a'_2, a'_3, a'_4)$ viene dada por

$$\begin{aligned} a'_1 &= \operatorname{rg}(-A')^0 - \operatorname{rg}(-A') = 5 - 3 = 2 = b_1, \\ a'_2 &= \operatorname{rg}(-A') - \operatorname{rg}(-A')^2 = 3 - 2 = 1 = b_2, \\ a'_3 &= \operatorname{rg}(-A')^2 - \operatorname{rg}(-A')^3 = 2 - 1 = 1 = b_3, \\ a'_4 &= \operatorname{rg}(-A')^3 - \operatorname{rg}(-A')^4 = 1 - 0 = 1 = b_4. \end{aligned}$$

Segunda forma de hacer la demostración para $t = 1$. Por el Lema 6.2.4 podemos considerar que a se obtiene de b por una transformación elemental; es decir, que $b = (p-1, q+1) \prec (p, q) = a$ con $p \geq q \geq 1, p \geq 2$, siendo a la partición de Weyr del único valor propio λ_1 de A ; sin pérdida de generalidad podemos suponer que $\lambda_1 = 0$. Suponemos que la matriz nilpotente A es de orden $p+q$.

Sea A' la matriz que se obtiene al poner ε en vez de 0 en el lugar de la fila $p+q$ y la columna $q+p-1$ (o en el lugar de la fila $q+p-1$ y la columna $q+p$, y si además es $p-q > 2$ caben algunas posibilidades más). Entonces A' tiene como único valor propio el cero y tenemos que $\operatorname{rg}(A') = q+1$ y $\operatorname{rg}(A')^2 = 0$. Por lo tanto, si $w(0, A') = (a'_1, a'_2, \dots)$ se tiene que

$$\begin{aligned} a'_1 &= p+q - (q+1) = p-1 = b_1, \\ a'_2 &= q+1 - 0 = q+1 = b_2. \end{aligned}$$

Luego $w(0, A') = b$. Como $\varepsilon \neq 0$ puede tomarse tan pequeño como se quiera, A' tiene las propiedades requeridas.

Ejemplo 6.2.2. Sea $b = (3, 2) \prec (4, 1) = a$; es decir, que $p = 4, q = 1$. En este caso será

$$A = \left[\begin{array}{ccccc} 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

y tomamos

$$A' = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \varepsilon & 0 \end{bmatrix} \quad \text{ó} \quad A' = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \varepsilon \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

En ambos casos, $\Lambda(A') = \{0\}$ y $w(0, A') = (a'_1, a'_2)$ con

$$\begin{aligned} a'_1 &= \text{rg}(A')^0 - \text{rg}(A') = 5 - 2 = 3 = b_1, \\ a'_2 &= \text{rg}(A') - \text{rg}(A')^2 = 2 - 0 = 2 = b_2. \end{aligned}$$

Ejemplo 6.2.3. Sea $b = (5, 4) \prec (6, 3) = a$. La matriz A será

$$A = \left[\begin{array}{ccc|cccc} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{0} & \boxed{0} & \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{0} & 0 & \boxed{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{0} & \boxed{0} & 0 \end{array} \right];$$

la matriz A' se obtiene poniendo $\varepsilon \neq 0$ en lugar de cero en alguno (y sólo alguno) de los lugares marcados con un cuadrado $\boxed{0}$; ya que en cualquiera de las seis opciones se tiene $w(0, A') = (a'_1, a'_2)$ siendo

$$\begin{aligned} a'_1 &= \text{rg}(A')^0 - \text{rg}(A') = 9 - 4 = 5 = b_1, \\ a'_2 &= \text{rg}(A') - \text{rg}(A')^2 = 4 - 0 = 4 = b_2. \end{aligned}$$

Vamos a examinar ahora el caso en que $t > 1$. Entonces la condición (6.1) se reduce a

$$\bigcup_{i=1}^t b_i \prec a; \quad b_i := (b_{i1}, b_{i2}, \dots), i = 1, \dots, t.$$

Por los Lemas 6.2.2 y 6.2.3 podemos seguir suponiendo que $\lambda_1 = 0$ es el único valor propio de A ; esto es, que A es nilpotente. Si llamamos

$$d := \bigcup_{i=1}^t b_i,$$

será $d \prec a$ y por el caso $t = 1$, ya demostrado, podemos encontrar en cualquier entorno de A una matriz B con un único valor propio (el cero) tal que $d = w(0, B)$.

Como sabemos los términos no nulos de d provienen de la “unión” de los términos no nulos de b_1, \dots, b_t reordenados en sentido decreciente. esta reordenación puede hacerse de maneras diferentes. Determinaremos una manera que precisa de qué partición b_i ($i = 1, \dots, t$) procede cada término no nulo de d . Escribimos los términos no nulos de b_1 (en sentido decreciente) y a continuación los de b_2 , y así sucesivamente. En primer lugar, d_1 será escogido de entre los mayores términos, de este modo escritos, el que se encuentre más a la izquierda; es decir,

$$d_1 := b_{i1}$$

tal que

$$b_{i1} = \text{máx}\{b_{ij} : i = 1, \dots, t; j = 1, 2, \dots\}$$

Tomemos

$$A' = \left[\begin{array}{cc|ccc|ccccc} \varepsilon_2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \varepsilon_2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \varepsilon_1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \varepsilon_1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \varepsilon_1 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \varepsilon_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon_2 \end{array} \right];$$

entonces

$$\begin{aligned} w(\varepsilon_1, A') &= (10 - 7, 7 - 7) = (3, 0) = b_1, \\ w(\varepsilon_2, A') &= (10 - 5, 5 - 3, 3 - 3) = (5, 2, 0) = b_2. \end{aligned}$$

Ejemplo 6.2.6. Sean $t = 3, b_1 := (2, 2), b_2 := (2, 1), b_3 := (1, 0)$ y $a := (2, 2, 2, 2)$. Por lo tanto,

$$d := b_1 \cup b_2 \cup b_3 = (2, 2, 2, 1, 1) = (b_{11}, b_{12}, b_{21}, b_{22}, b_{31}),$$

y $d \prec a$. Entonces,

$$A = \left[\begin{array}{cc|cc|} 0 & 0 & 1 & 0 & & \\ 0 & 0 & 0 & 1 & & \\ \hline & & 0 & 0 & 1 & 0 \\ & & 0 & 0 & 0 & 1 \\ \hline & & & & 0 & 0 & 1 & 0 \\ & & & & 0 & 0 & 0 & 1 \\ \hline & & & & & & 0 & 0 \\ & & & & & & 0 & 0 \end{array} \right]; A' = \left[\begin{array}{c|cc|} \varepsilon_3 & 0 & 1 & & \\ \hline & \varepsilon_2 & 0 & 1 & \\ \hline & & \varepsilon_2 & 0 & 1 & 0 \\ & & 0 & \varepsilon_2 & 0 & 1 \\ \hline & & & & \varepsilon_1 & 0 & 1 & 0 \\ & & & & 0 & \varepsilon_1 & 0 & 1 \\ \hline & & & & & & \varepsilon_1 & 0 \\ & & & & & & 0 & \varepsilon_1 \end{array} \right]$$

y

$$\begin{aligned} w(\varepsilon_1, A') &= (8 - 6, 6 - 4, 4 - 4) = (2, 2, 0) = b_1, \\ w(\varepsilon_2, A') &= (8 - 6, 6 - 5, 5 - 5) = (2, 1, 0) = b_2, \\ w(\varepsilon_3, A') &= (8 - 7, 7 - 7) = (1, 0) = b_3. \end{aligned}$$

Ejemplo 6.2.7. Sean $t = 2, b_1 := (2, 1, 1), b_2 := (2, 0, 0)$ y $a := (2, 2, 2)$. Entonces

$$d := b_1 \cup b_2 = (2, 2, 1, 1) = (b_{11}, b_{21}, b_{12}, b_{13})$$

y $d \prec a$. Las matrices A y A' son

$$A = \left[\begin{array}{cc|cc|} 0 & 0 & 1 & 0 & & \\ 0 & 0 & 0 & 1 & & \\ \hline & & 0 & 0 & 1 & 0 \\ & & 0 & 0 & 0 & 1 \\ \hline & & & & 0 & 0 \\ & & & & 0 & 0 \end{array} \right]; A' = \left[\begin{array}{c|cc|} \varepsilon_1 & 0 & 1 & & \\ \hline & \varepsilon_1 & 0 & 1 & \\ \hline & & \varepsilon_2 & 0 & 1 & 0 \\ & & 0 & \varepsilon_2 & 0 & 1 \\ \hline & & & & \varepsilon_1 & 0 \\ & & & & 0 & \varepsilon_1 \end{array} \right].$$

Las particiones de Weyr de ε_1 y ε_2 respecto de A' son:

$$\begin{aligned} w(\varepsilon_1, A') &= (6 - 4, 4 - 3, 3 - 2, 2 - 2) = (2, 1, 1, 0) = b_1, \\ w(\varepsilon_2, A') &= (6 - 4, 4 - 4) = (2, 0) = b_2. \end{aligned}$$

Con esto damos por concluida la demostración del Teorema 6.2.1. \square

La condición (6.1) implica que la suma de los pesos de todas las particiones prescritas b_{ij} debe ser igual al orden n de la matriz:

$$\sum_{i=1}^u \sum_{j=1}^{t_i} |b_{ij}| = n.$$

Observación 6.2.1. Para cada $i = 1, \dots, u$, podríamos preguntarnos si además de t_i y b_{i1}, \dots, b_{it_i} es posible prescribir los valores propios de la matriz A' . Sí se pueden prescribir éstos con tal de que pertenezcan a $\Lambda(A)$. En caso contrario, por la continuidad de los valores propios de A' cuando A' tiene a A , se ve que habría un entorno \mathcal{V} de A tal que toda $X \in \mathcal{V}$ no podría tener los valores propios prescritos que no fueran elementos de $\Lambda(A)$.

Esta observación desvela el problema siguiente: dados $t_i > 0, b_{i1}, \dots, b_{it_i}$ particiones no nulas tales que $\cup_{j=1}^{t_i} b_{ij} \prec w(\lambda_i, A)$, y distintos $z_{i1}, \dots, z_{it_i} \in B(\lambda_i, \eta)$ (para cada $i = 1, \dots, u$), **hallar el mínimo radio** ε de un entorno $B(A, \varepsilon)$ de A tal que en $\overline{B(A, \varepsilon)}$ haya alguna matriz X_0 que satisfaga

$$\Lambda(X_0) = \bigcup_{i=1}^u \{z_{i1}, \dots, z_{it_i}\}$$

$$\forall i = 1, \dots, u, \forall j = 1, \dots, t_i, w(z_{ij}, X_0) = b_{ij}.$$

Conviene enunciar el Teorema 6.2.1 en términos de las características de Segré de las matrices. Teniendo en cuenta el Lema 2.3.1 del Capítulo 2, hemos probado el teorema siguiente.

Teorema 6.2.2. Sean $A \in \mathbb{C}^{n \times n}$ y $\Lambda(A) = \{\lambda_1, \dots, \lambda_u\}$. Sea η un número real positivo tal que $B(\lambda_i, \eta) \cap B(\lambda_j, \eta) = \emptyset, i \neq j, i, j \in \{1, \dots, u\}$. Para cada $i = 1, \dots, u$, sea t_i un entero positivo dado y sean c_{i1}, \dots, c_{it_i} particiones no nulas dadas.

En cualquier entorno de A existe una matriz A' tal que

- (i) $\Lambda(A') \subset \bigcup_{i=1}^u B(\lambda_i, \eta)$,
- (ii) A' tiene exactamente t_i valores propios en $B(\lambda_i, \eta)$, y $c_{ij} = s(\mu_{ij}, A'), j = 1, \dots, t_i, i = 1, \dots, u$,

si y sólo si se cumplen las condiciones

$$s(\lambda_i, A) \prec \sum_{j=1}^{t_i} c_{ij}, \quad (i = 1, \dots, u). \quad (6.4)$$

Para acabar esta sección queremos hacer una observación terminológica: Se dirá que una estructura de Jordan de orden n, \mathcal{E} , es *repelida* por una matriz $A \in \mathbb{C}^{n \times n}$ si hay algún entorno de A tal que ninguna de las matrices situadas en él la tenga como su característica de Segré. Es obvio que el Teorema 6.2.2 nos da condiciones necesarias y suficientes para que \mathcal{E} sea repelida por A .

6.3. La clausura de la órbita de semejanza

Dada una matriz $A \in \mathbb{C}^{n \times n}$ su **órbita de semejanza** es el conjunto

$$[A] := \{P^{-1}AP : P \in \mathbb{C}^{n \times n} \text{ invertible}\};$$

es decir, que la órbita de semejanza es la clase de equivalencia de A respecto de la semejanza de matrices. La clausura topológica (adherencia o cierre) de $[A]$, denotada por $\overline{[A]}$, se define como el menor conjunto cerrado de $\mathbb{C}^{n \times n}$ que contiene a $[A]$. El estudio de la clausura del conjunto $[A]$ tiene estrecha relación con los teoremas de perturbación anteriores. Por la

definición de la clausura, se sigue que una matriz X pertenece a $\overline{[A]}$ si y sólo si existe una sucesión de matrices $(A_k)_{k \in \mathbb{N}}$, $A_k \in \mathbb{C}^{n \times n}$, con $A_k \approx A$ para cada $k \in \mathbb{N}$, que converge a X cuando k tiende a infinito. De manera equivalente, $X \in \overline{[A]}$ si en todo entorno de X hay una matriz semejante a A . El teorema siguiente caracteriza las matrices que son elementos de $\overline{[A]}$.

Teorema 6.3.1. Sean $A \in \mathbb{C}^{n \times n}$ y $\Lambda(A) = \{\lambda_1, \dots, \lambda_u\}$. Entonces $X \in \overline{[A]}$ si y sólo si se cumplen las condiciones siguientes:

(i)

$$\Lambda(A) = \Lambda(X),$$

(ii)

$$w(\lambda_i, A) \prec w(\lambda_i, X), \quad i = 1, \dots, u.$$

DEMOSTRACIÓN. Supongamos que existe una sucesión $A_k \rightarrow X$ cuando $k \rightarrow \infty$ con $A_k \approx A$ para todo k ; en particular, $\Lambda(A_k) = \{\lambda_1, \dots, \lambda_u\}$ para todo k . Para cada $i = 1, \dots, u$, se cumple que $\lambda_i \in \Lambda(X)$; porque si existiera un $j \in \{1, \dots, u\}$ tal que $\lambda_j \notin \Lambda(X)$, entonces podríamos encontrar un número real $\eta > 0$ tal que

$$\eta < \min_{\xi \in \Lambda(X)} |\lambda_j - \xi|$$

y que las bolas $B(\xi, \eta)$, $\xi \in \Lambda(X)$, fueran disjuntas a pares. Por el Teorema 6.1.2 existiría un entorno de X en el que no habría ningún elemento de la sucesión $(A_k)_{k \in \mathbb{N}}$, lo que es una contradicción. Así pues, $\Lambda(A) \subset \Lambda(X)$.

Además $\Lambda(X) \subset \Lambda(A)$, porque si existiera un $\xi_0 \in \Lambda(X)$ tal que $\xi_0 \notin \{\lambda_1, \dots, \lambda_u\}$ llegaríamos a una contradicción puesto que por el Corolario 1 del Teorema 6.1.2 existiría un entorno \mathcal{V} de X tal que si $X' \in \mathcal{V}$ entonces $\Lambda(X') \subset V_\eta(X)$ y

$$\sum_{\xi' \in \Lambda(X') \cap B(\xi_0, \eta)} m(\xi', X') = m(\xi_0, X);$$

condición que no podría cumplir ningún elemento de la sucesión $(A_k)_{k \in \mathbb{N}}$ si hubiéramos elegido $\eta > 0$ con la condición adicional de que $\eta < \min\{|\xi_0 - \lambda_i| : i = 1, \dots, u\}$. Una vez demostrada la condición (i), la (ii) se obtiene a partir de la desigualdad (ii) del Teorema 6.1.2 aplicada a la matriz X , que en dicho teorema se llama A .

Para demostrar el recíproco supongamos que se cumplen las condiciones (i) y (ii). Por la definición de semejanza basta considerar el caso en que X está en forma canónica de Jordan y sólo tiene un valor propio α . Además por el Lema 2.4.2 del Capítulo 2 es suficiente tomar

$$w(\alpha, A) = (p - 1, q + 1, 0, \dots), \quad w(\alpha, X) = (p, q, 0, \dots)$$

con $p \geq q \geq 1, p \geq 2$. Por lo tanto, podemos suponer que

$$X = \text{diag}(J_p(\alpha), J_q(\alpha)).$$

Definimos C_k como la matriz $(p + q) \times (p + q)$ cuyos elementos valen todos 0 excepto el del lugar $(p + 1, 1)$ que vale $1/k$. Entonces, si tomamos $A_k := X + C_k$ tenemos que $A_k \rightarrow X$ cuando $k \rightarrow \infty$ y A_k tiene a α como único valor propio, con la partición de Weyr

$$w(\alpha, A_k) = (p - 1, q + 1, 0, \dots)$$

para todo $k = 0, 1, 2, \dots$. Por consiguiente, la matriz X pertenece a la clausura de la órbita de A . □

6.4. Ejercicios

Ejercicio 6.1. Probar que en todo entorno de una matriz $A \in \mathbb{C}^{n \times n}$ hay una matriz distinta de A que tiene su misma característica de Segré.

Ejercicio 6.2. Sea $A \in \mathbb{C}^{n \times n}$ una matriz dada y sea $(A_p)_{p=0}^\infty$ una sucesión de matrices de $\mathbb{C}^{n \times n}$ que converge hacia A . Demostrar que existe una subsucesión $(A_{p_k})_{k=0}^\infty$ de esta sucesión que converge hacia A y que está formada por matrices que tienen todas la misma característica de Segré.

Indicación. Utilizar una variante del principio del palomar de Dirichlet.

Ejercicio 6.3. Sean $A \in \mathbb{C}^{n \times n}$ y $B \in \mathbb{C}^{m \times m}$ tales que $\Lambda(A) \cap \Lambda(B) \neq \emptyset$. Demostrar que existe una sucesión de pares de matrices

$$(A_p, B_p) \in \mathbb{C}^{n \times n} \times \mathbb{C}^{m \times m}, p = 1, 2, \dots$$

tal que para cada p

$$(1) \|A_p - A\| \leq \frac{1}{p}, \quad \|B_p - B\| \leq \frac{1}{p},$$

$$(2) \Lambda(A_p) \cap \Lambda(B_p) = \emptyset,$$

y las características de Segré de A_p y B_p son constantes:

$$s(A_p) = \mathcal{E}, \quad s(B_p) = \mathcal{F}, \quad p = 1, 2, \dots$$

Ejercicio 6.4. Denotemos

$$L_{nmq} := \mathbb{C}^{n \times n} \times \mathbb{C}^{m \times m} \times \mathbb{C}^{q \times q}.$$

Sea la terna de matrices $(A, B, C) \in L_{nmq}$ tal que

$$\Lambda(A) \cap \Lambda(B) \cap \Lambda(C) \neq \emptyset.$$

Demostrar que existe una sucesión de ternas de matrices $(A_p, B_p, C_p) \in L_{nmq}, p = 1, 2, \dots$ tal que para cada p

$$(1) \max\{\|A_p - A\|, \|B_p - B\|, \|C_p - C\|\} \leq \frac{1}{p}$$

$$(2) \Lambda(A_p) \cap \Lambda(B_p) \cap \Lambda(C_p) = \emptyset$$

y las características de Segré de A_p, B_p y C_p son constantes:

$$s(A_p) = \mathcal{E}, \quad s(B_p) = \mathcal{F}, \quad s(C_p) = \mathcal{G}, \quad p = 1, 2, \dots$$

Ejercicio 6.5. Denotemos

$$L_{nmq} := \mathbb{C}^{n \times n} \times \mathbb{C}^{m \times m} \times \mathbb{C}^{q \times q}.$$

Sea la terna de matrices $(A, B, C) \in L_{nmq}$ tal que

$$\Lambda(A) \cap \Lambda(B) \cap \Lambda(C) \neq \emptyset.$$

$$\Lambda(A) \cap \Lambda(B) \neq \emptyset, \quad \Lambda(B) \cap \Lambda(C) \neq \emptyset, \quad \Lambda(A) \cap \Lambda(C) \neq \emptyset.$$

Demostrar que existe una sucesión de ternas de matrices $(A_p, B_p, C_p) \in L_{nmq}, p = 1, 2, \dots$ tal que para cada p

$$(1) \max\{\|A_p - A\|, \|B_p - B\|, \|C_p - C\|\} \leq \frac{1}{p}$$

(2)

$$\Lambda(A_p) \cap \Lambda(B_p) = \emptyset, \quad \Lambda(B_p) \cap \Lambda(C_p) = \emptyset, \quad \Lambda(A_p) \cap \Lambda(C_p) = \emptyset.$$

y las características de Segré de A_p, B_p y C_p son constantes:

$$s(A_p) = \mathcal{E}, \quad s(B_p) = \mathcal{F}, \quad s(C_p) = \mathcal{G}, \quad p = 1, 2, \dots$$

Ejercicio 6.6. Sea la matriz en forma de Jordan

$$A = \left(\begin{array}{cccc|ccc} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right).$$

Perturbando ligeramente el elemento de una sola posición (i, j) de A obtener una matriz A' tal que:

(a) $s(0, A') = (5, 2)$.

(b) $s(0, A') = (6, 1)$.

(c) $s(0, A') = (7)$.

¿Cuáles son las posiciones posibles en cada caso?

Ejercicio 6.7. Dada la matriz en forma de Jordan

$$A = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right).$$

Hallar las características de Segré de las 25 matrices A' obtenidas al sumar a un solo elemento de A un valor $\varepsilon \neq 0$ suficientemente pequeño. ¿Cambia algo el resultado si se permite alterar dos elementos de A ?

Ejercicio 6.8. Sea la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}.$$

Mostrar que toda matriz suficientemente próxima a A es diagonalizable.

Ejercicio 6.9. Sea la matriz

$$\begin{pmatrix} 3 & -2 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 4 \end{pmatrix}.$$

Describir las posibles características de Weyr de todas las matrices A' suficientemente próximas a A (*características admitidas*). Dar un ejemplo concreto de A' para cada una de las características descritas.

Determinar las características de Weyr *repelidas* por A . Una característica de Weyr de una matriz compleja 3×3 se dice *repelida* por A si hay algún entorno de A tal que ninguna de las matrices situadas en él la tiene como su característica de Weyr.

Ejercicio 6.10. Se llama *conmutante* de una matriz $A \in \mathbb{C}^{n \times n}$ al conjunto de matrices de $\mathbb{C}^{n \times n}$ que conmutan con A . Demostrar que si la forma canónica de Jordan de A tiene k bloques de Jordan, entonces el conjunto de matrices de $\mathbb{C}^{n \times n}$ que tienen k valores propios distintos, es denso en el conmutante de A .

Ejercicio 6.11. Dada una matriz compleja $n \times n$, A , llamemos $i(A)$ al número de factores invariantes de A de grado mayor o igual que 1 (contando los repetidos). Sea k un entero, $2 \leq k \leq n$. Demostrar que el conjunto de matrices $A \in \mathbb{C}^{n \times n}$ tales que $i(A) < k$ es abierto.

Ejercicio 6.12. Dados una matriz $A \in \mathbb{C}^{n \times n}$ y un número entero p , $1 \leq p \leq n$, llamemos $i_p(A)$ al número de factores invariantes de A de grado mayor o igual que p (contando los repetidos). Demostrar que existe un entorno \mathcal{V} de A tal que para toda matriz $A' \in \mathcal{V}$ se tiene que

$$(i_1(A'), i_2(A'), \dots, i_n(A')) < (i_1(A), i_2(A), \dots, i_n(A)),$$

y que, por lo tanto, el conjunto

$$\mathcal{M}_{pk} := \{X \in \mathbb{C}^{n \times n} : i_1(X) + \dots + i_p(X) < k\}$$

es abierto.

Ejercicio 6.13. Supongamos que la matriz $A \in \mathbb{C}^{n \times n}$ es diagonalizable, pero tiene algún valor propio múltiple. Demostrar que tan cerca como queramos de A existe una matriz A' no diagonalizable.

Ejercicio 6.14. Demostrar que en todo entorno de una matriz derogatoria hay matrices no diagonalizables.

Ejercicio 6.15. Una matriz $A \in \mathbb{C}^{n \times n}$ se llama matriz triangular superior de Toeplitz si tiene la siguiente forma

$$A = \begin{bmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} \\ 0 & a_0 & a_1 & \ddots & \vdots \\ & & \ddots & \ddots & \\ \vdots & & & \ddots & a_1 \\ 0 & 0 & 0 & \dots & a_0 \end{bmatrix}.$$

Dada una matriz triangular superior de Toeplitz $A \in \mathbb{C}^{n \times n}$, encontrar todas las posibles características de Segré (o estructuras de Jordan) de las matrices triangulares superiores de Toeplitz que están arbitrariamente próximas a A . ¿Hay características de Segré adicionales si la matriz perturbada no es necesariamente triangular superior de Toeplitz?

Ejercicio 6.16. Resolver el Ejercicio 6.15 para la clase de matrices compañeras de polinomios.

Ejercicio 6.17. Una matriz circulante $n \times n$ es una matriz de la forma

$$A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ a_n & a_1 & \dots & a_{n-1} \\ \vdots & \vdots & & \vdots \\ a_2 & a_3 & \dots & a_1 \end{bmatrix}$$

donde a_1, a_2, \dots, a_n son números complejos. Resolver el Ejercicio 6.15 para la clase de las matrices circulantes.

Ejercicio 6.18. Para una matriz $A \in \mathbb{C}^{n \times n}$, la *altura* de A es el máximo valor de los órdenes de los bloques de Jordan correspondientes a los valores propios α tales que $\text{mg}(\alpha, A) = 1$. Si A no tiene ninguno de estos valores propios, su altura se define igual a 1.

Demostrar o refutar cada uno de los siguientes enunciados (a), (b) y (c):

Para toda matriz $A \in \mathbb{C}^{n \times n}$ existe un $\varepsilon > 0$ tal que cualquier matriz $A' \in \mathbb{C}^{n \times n}$ con $\|A' - A\| < \varepsilon$ tiene la propiedad de que

- (a) la altura de A' es igual a la altura de A ;
- (b) la altura de A' es menor o igual que la altura de A ;
- (c) la altura de A' es mayor o igual que la altura de A .

Ejercicio 6.19. Sea \mathcal{E} una estructura de Jordan de orden n cualquiera. Demostrar que en todo entorno de la matriz cero $O \in \mathbb{C}^{n \times n}$ existen matrices cuya característica de Segré es igual a \mathcal{E} . ¿Para qué otras matrices de $\mathbb{C}^{n \times n}$ ocurre lo mismo?

Ejercicio 6.20. Caracterizar las estructuras de Jordan \mathcal{F} de orden n que poseen la propiedad siguiente: Si una matriz $A \in \mathbb{C}^{n \times n}$ satisface $s(A) = \mathcal{F}$, entonces para toda matriz A' suficientemente próxima a A se tiene que $s(A') = \mathcal{F}$.

Ejercicio 6.21. Dadas dos matrices $A, B \in \mathbb{C}^{n \times n}$, recordemos que $\nu(A, B)$ denota la dimensión del subespacio

$$\{Y \in \mathbb{C}^{n \times n} : AY - YB = O\}$$

de $\mathbb{C}^{n \times n}$ (Sección 3.8). Demostrar que una matriz $C \in \overline{[A]}$ si y sólo si para toda $X \in \mathbb{C}^{n \times n}$ se cumple que $\nu(A, X) = \nu(C, X)$.

6.5. Notas al Capítulo 6

Los teoremas sobre perturbación de la forma de Jordan aparecieron por primera vez en los artículos [46], [19]. Luego han sido recogidos en el libro [31] y en el artículo [34]. El Ejercicio 6.10 está tomado de Holbrook [36]. Los Ejercicios 6.15, 6.16, 6.17, y 6.18 han sido tomados del Capítulo 16, pág. 513, del libro [31] de Gohberg, Lancaster y Rodman. El Ejercicio 6.21 está inspirado en la introducción del artículo [13] de Bongartz.

sabemos también por la experiencia previa al resolver ejercicios de búsqueda de los valores propios de A que 0 es valor propio de A si esta matriz tiene “suficientes” elementos iguales a 0. Un teorema famoso de Frobenius-König dice que una condición suficiente para que esta ocurra es que A tenga una submatriz de ceros $s \times t$ con $s + t$ mayor que el orden de A . Es claro, que la condición no es necesaria, pues hay matrices con ningún elemento nulo y de determinante igual a cero.

Los lugares donde M tiene el signo ? se pueden señalar por medio del digrafo (de Coates) asociado a M . Véase la figura 7.1.

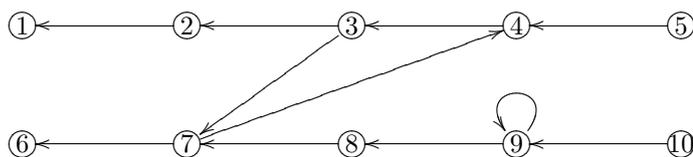


Figura 7.1: Digrafo de Coates de M .

Este digrafo tiene tantos vértices como el orden de M y tiene un arco dirigido del vértice j al vértice i si hay un ? en el lugar (i, j) de la matriz genérica M . La forma canónica genérica de Jordan (7.2) vendrá determinada por el digrafo de la figura 7.1, asociado a M , según veremos más adelante en este capítulo. Este resultado junto con los teoremas de perturbación del Capítulo 6 nos permiten deducir información, exactamente calculable, sobre la forma de Jordan de toda matriz concreta con modelo de ceros (7.1). Veámoslo por medio de un ejemplo; si A es la matriz siguiente de orden 10

$$\begin{bmatrix} 0 & -7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \pi & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & e & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (7.3)$$

En todo entorno de A existen matrices con modelo (7.1); además, A tiene una submatriz cero 5×6 y —por lo tanto— 0 es valor propio de A . En virtud de los teoremas sobre condiciones necesarias de perturbación de la forma canónica de Jordan (Véase el Capítulo 6), se deduce que

$$w(0, M) \ll w(0, A),$$

siendo

$$w(0, M) = \overline{(4, 2, 0, \dots)} = (2, 2, 1, 1, 0, \dots).$$

Esto es consecuencia de la condición (ii)

$$\bigcup_{j=1}^t w(\mu_j, M) \prec w(0, A),$$

del Teorema 6.1.2 del Capítulo 6, siendo μ_1, \dots, μ_t los valores propios “próximos” a 0 de las matrices de modelo M situadas suficientemente cerca de A . Para la matriz A de (7.3) se tiene que

$$w(0, M) \prec w(0, A).$$

Pero a veces la mayoración no es estricta. Esto ocurre, por ejemplo, para la matriz

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad (7.4)$$

donde $w(0, B) = (3, 0, \dots)$, pero las matrices genéricas de modelo

$$M = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix} \quad (7.5)$$

próximas a B tienen un valor propio no nulo cercano a 0 y, en consecuencia, $w(0, M) = (2, 0, \dots)$; aquí,

$$(2, 0, \dots) \ll (3, 0, \dots),$$

pero

$$(2, 0, \dots) \not\prec (3, 0, \dots).$$

7.2. Teorema de Frobenius-König

7.3. Menores de la matriz característica y digrafos

Dada una matriz $A = (a_{ij}) \in \mathbb{F}^{n \times n}$ donde \mathbb{F} es un cuerpo conmutativo, consideremos su matriz característica $\lambda I_n - A$. Los menores de esta matriz son los determinantes de sus submatrices cuadradas; estos menores son, por tanto, polinomios en λ . (En particular, consideraremos al polinomio característico de A ,

$$\det(\lambda I_n - A) = \lambda^n + p_1 \lambda^{n-1} + \dots + p_{n-1} \lambda + p_n, \quad (7.1)$$

como uno de estos menores). El objetivo de la presente sección es describir los coeficientes de estos polinomios en términos del digrafo de Coates, $G(A)$, asociado a la matriz A (véase la Sección B.3 del Apéndice B).

Si hacemos $\lambda = 0$ en la ecuación (7.1), obtenemos que $p_n = \det(-A)$, pero $\det(-A) = (-1)^n \det A$. Es conocido que $p_1 = (-1)^{n-1} \text{tr}(A)$, donde $\text{tr}(A) := a_{11} + a_{22} + \dots + a_{nn}$ denota la *traza* de A . Estas expresiones de p_n y p_1 en función de los elementos de A son casos particulares de una expresión general que nos da el coeficiente ν -ésimo del polinomio característico, p_ν , en términos de los elementos de A :

$$p_\nu = (-1)^{-\nu} \sum_{1 \leq i_1 < i_2 < \dots < i_\nu \leq n} A \begin{pmatrix} i_1 & i_2 & \dots & i_\nu \\ i_1 & i_2 & \dots & i_\nu \end{pmatrix} \quad (7.2)$$

siendo

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_\nu \\ i_1 & i_2 & \dots & i_\nu \end{pmatrix} = \begin{vmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \dots & a_{i_1 i_\nu} \\ a_{i_2 i_1} & a_{i_2 i_2} & \dots & a_{i_2 i_\nu} \\ \vdots & \vdots & & \vdots \\ a_{i_\nu i_1} & a_{i_\nu i_2} & \dots & a_{i_\nu i_\nu} \end{vmatrix} \quad (7.3)$$

el menor principal de A formado por las filas i_1, i_2, \dots, i_ν y las mismas columnas, $\nu = 1, 2, \dots, n$. Puede verse una demostración de este resultado en [28, p. 72] y [41, p. 157, Theorem 2]. Utilizando el digrafo ponderado de Coates de la matriz cuyo determinante aparece en (7.3), podemos expresar

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_\nu \\ i_1 & i_2 & \dots & i_\nu \end{pmatrix}$$

en términos de las diagonales de dicha matriz que tienen todas sus componentes no nulas. Pero estas diagonales se corresponden biunívocamente con las familias de ciclos generadoras de este digrafo (véase la Sección B.3 del Apéndice B). Los vértices de este digrafo son i_1, i_2, \dots, i_ν . Por ello, estas familias de ciclos son, precisamente, las familias de ciclos del digrafo $G(A)$ formadas por ciclos “disjuntos” cuya “unión” de vértices es el conjunto $\{i_1, i_2, \dots, i_\nu\}$. Por el Teorema B.3.1 del Apéndice B cada una de estas familias de ciclos h aporta un sumando al determinante

$$A \begin{pmatrix} i_1 & i_2 & \cdots & i_\nu \\ i_1 & i_2 & \cdots & i_\nu \end{pmatrix},$$

que es igual a su peso $f(h)$ multiplicado por $(-1)^{\nu-d(h)}$, siendo $d(h)$ el número de ciclos “disjuntos” de h , lazos incluidos.

Llamemos ahora, para cada $\nu = 1, 2, \dots, n$, H_ν al conjunto de familias de ciclos de $G(A)$ de longitud ν . Por el Teorema B.3.1 del Apéndice B y la ecuación (7.2) se tiene que

$$p_\nu = (-1)^{-\nu} \sum_{h \in H_\nu} (-1)^{\nu-d(h)} f(h) = \sum_{h \in H_\nu} (-1)^{d(h)} f(h). \quad (7.4)$$

Hemos probado así el teorema siguiente.

Teorema 7.3.1. *Con las notaciones precedentes, el coeficiente p_ν , ($\nu = 1, \dots, n$) del polinomio característico de la matriz $A \in \mathbb{F}^{n \times n}$,*

$$\det(\lambda I_n - A) = \lambda^n + \sum_{\nu=1}^n p_\nu \lambda^{n-\nu}$$

se puede hallar de la manera siguiente. Cada familia de ciclos de longitud ν aporta un sumando a p_ν . El valor numérico del sumando se obtiene mediante el producto de los pesos de los arcos de la familia de ciclos considerada. Este valor debe ser multiplicado por un factor de signo $(-1)^c$ si la familia tiene c ciclos “disjuntos”, lazos inclusive. Si no hubiera ninguna familia de ciclos de longitud ν debe entenderse que $p_\nu = 0$.

^{2A} DEMOSTRACIÓN DEL TEOREMA 7.3.1. Consideremos el digrafo ponderado $G(\lambda I_n - A)$. La novedad de este digrafo respecto del digrafo $G(A)$ es que los pesos de los arcos de $G(\lambda I_n - A)$ son elementos del anillo $\mathbb{F}[\lambda]$; en particular, en cada vértice i hay un lazo con peso $\lambda - a_{ii}$, $i = 1, \dots, n$. Aplicando el Teorema B.3.1 de la página 170 del Apéndice B a la matriz $\lambda I_n - A$ se tiene que

$$\det(\lambda I_n - A) = \sum_{h \in H} (-1)^{n-d(h)} f(h),$$

donde la suma se extiende a todas las familias de ciclos generadoras h del digrafo $G(\lambda I_n - A)$ y $f(h)$ es el peso de h como subdigrafo ponderado de $G(\lambda I_n - A)$. Sea ν un entero, $\nu = 1, \dots, n$. Fijemos nuestra atención en aquellas familias generadoras h que contienen exactamente $n - \nu$ de los lazos de peso $\lambda - a_{ii}$, $i = 1, \dots, n$ y los restantes ciclos de h están formados por arcos (j, i) con valores $-a_{ij}$; para concretar digamos que a h le corresponde el sumando

$$\varepsilon(\sigma) (\lambda - a_{k_1, k_1}) (\lambda - a_{k_2, k_2}) \cdots (\lambda - a_{k_{n-\nu}, k_{n-\nu}}) (-a_{\ell_1, m_1}) (-a_{\ell_2, m_2}) \cdots (-a_{\ell_\nu, m_\nu}) \quad (7.5)$$

en el desarrollo del determinante $\det(\lambda I_n - A)$, donde σ es la permutación

$$\sigma = \begin{pmatrix} k_1 & k_2 & \cdots & k_{n-\nu} & \ell_1 & \ell_2 & \cdots & \ell_\nu \\ k_1 & k_2 & \cdots & k_{n-\nu} & m_1 & m_2 & \cdots & m_\nu \end{pmatrix}$$

de $\{1, \dots, n\}$; es decir que

$$\{1, \dots, n\} = \{k_1, k_2, \dots, k_{n-\nu}\} \dot{\cup} \{\ell_1, \ell_2, \dots, \ell_\nu\}$$

($\dot{\cup}$ denota unión de conjuntos disjuntos) y además (m_1, m_2, \dots, m_ν) es una permutación de $\{\ell_1, \ell_2, \dots, \ell_\nu\}$. El coeficiente de $\lambda^{n-\nu}$ en el desarrollo del polinomio de la expresión (7.5) es igual a

$$\varepsilon(\sigma)(-1)^\nu,$$

y

$$\varepsilon(\sigma) = (-1)^{n-[n-\nu+c]}$$

siendo c el número de ciclos de h con arcos cuyo peso pertenece a \mathbb{F} ; o, lo que es equivalente, c es el número de permutaciones-ciclo en las que se descompone la permutación

$$\begin{pmatrix} \ell_1 & \ell_2 & \cdots & \ell_\nu \\ m_1 & m_2 & \cdots & m_\nu \end{pmatrix}.$$

De aquí que

$$\begin{aligned} \varepsilon(\sigma)(-1)^\nu &= (-1)^{n-[n-\nu+c]}(-1)^\nu \\ &= (-1)^{2\nu-c} = (-1)^{-c} = (-1)^c. \end{aligned}$$

De este modo, la expresión (7.5) es igual a

$$(-1)^c a_{\ell_1, m_1} a_{\ell_2, m_2} \cdots a_{\ell_\nu, m_\nu} \lambda^{n-\nu} + \text{términos de grado } < n - \nu \quad (7.6)$$

Así pues, cada familia de ciclos generadora h de $G(\lambda I_n - A)$ con $n - \nu$ lazos determina biunívocamente una familia de ciclos \tilde{h} en $G(A)$ de longitud ν , de forma que el coeficiente de $\lambda^{n-\nu}$ en (7.6) es igual a

$$(-1)^c f(\tilde{h}), \quad (7.7)$$

siendo $f(\tilde{h})$ el peso de \tilde{h} como subdigrafo ponderado de $G(A)$. Obviamente el número de ciclos de \tilde{h} es también c .

Sumando, a continuación, todas las aportaciones de las familias de ciclos generadoras de $G(\lambda I_n - A)$ con $n - \nu$ lazos, obtenemos el coeficiente p_ν del polinomio

$$\lambda^n + \sum_{\nu=1}^n p_\nu \lambda^{n-\nu}.$$

Pero, por lo antedicho, esto equivale a decir que p_ν se obtiene como suma de los pesos (7.7) aportados por las correspondientes familias de ciclos de longitud ν de $G(A)$, multiplicados por un factor de signo $(-1)^c$, donde c es el número de sus ciclos “disjuntos”. \square

Como $p_n = \det(-A) = (-1)^n \det A$, es de observar que de la fórmula (7.4) para $\nu = n$,

$$p_n = \sum_{h \in H_n} (-1)^{d(h)} f(h)$$

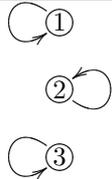
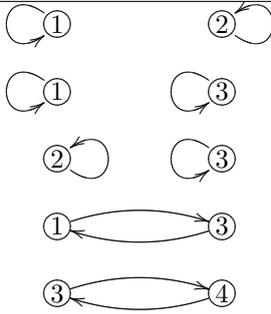
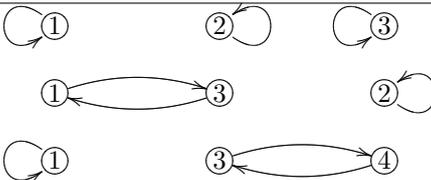
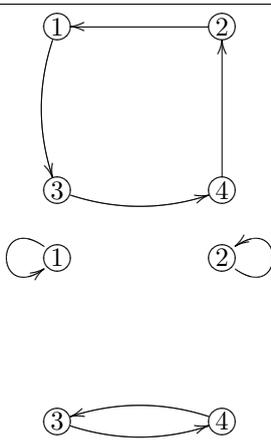
se deduce como caso particular que

$$\det A = \sum_{h \in H_n} (-1)^{n-d(h)} f(h).$$

Ejemplo 7.3.1. Sea la matriz 4×4

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 \\ 0 & a_{22} & 0 & a_{24} \\ a_{31} & 0 & a_{33} & a_{34} \\ 0 & 0 & a_{43} & 0 \end{pmatrix} \quad (7.8)$$

donde los a_{ij} que en ella figuran se consideran elementos distintos de cero. El digrafo $G(A)$ viene dado en la figura 7.2.

Coeficiente p_ν	Sumandos de p_ν	Familias de ciclos correspondientes
p_1	$-a_{11}$ $-a_{22}$ $-a_{33}$	
p_2	$a_{11}a_{22}$ $+a_{11}a_{33}$ $+a_{22}a_{33}$ $-a_{13}a_{31}$ $-a_{34}a_{43}$	
p_3	$-a_{11}a_{22}a_{33}$ $+a_{13}a_{31}a_{22}$ $+a_{11}a_{34}a_{43}$	
p_4	$-a_{31}a_{43}a_{24}a_{12}$ $-a_{11}a_{22}a_{34}a_{43}$	

Cuadro 7.1: Coeficientes del polinomio característico.

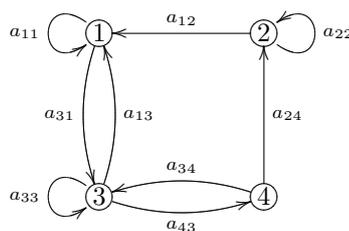


Figura 7.2: Digrafo \$G(A)\$.

Para este Ejemplo 7.3.1, las expresiones algebraicas de \$p_\nu\$ (\$1 \leq \nu \leq 4\$) y las familias de ciclos de \$G(A)\$ correspondientes a cada sumando de \$p_\nu\$ han sido mostradas en la tabla 7.1.

Sea \$A \in \mathbb{F}^{n \times n}\$. Eliminando en la matriz característica \$\lambda I_n - A\$ la fila \$i\$ y la columna \$j\$ se obtiene una matriz \$(n-1) \times (n-1)\$

$$(\lambda I_n - A) \begin{pmatrix} i \\ j \end{pmatrix}^c, \quad (7.9)$$

cuyo determinante \$\Delta_{n-1}^{(ij)}(\lambda)\$ es igual a un polinomio de grado a lo más \$n-1\$:

$$p_1^{(ij)} \lambda^{n-1} + p_2^{(ij)} \lambda^{n-2} + \cdots + p_{n-1}^{(ij)} \lambda + p_n^{(ij)}. \quad (7.10)$$

Sea \$(\lambda I_n - A)(ij)\$ la matriz

$$\begin{pmatrix} \lambda - a_{11} & \cdots & -a_{1,j-1} & 0 & -a_{1,j+1} & \cdots & -a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ -a_{i-1,1} & \cdots & -a_{i-1,j-1} & 0 & -a_{i-1,j+1} & \cdots & -a_{in} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ -a_{i+1,1} & \cdots & -a_{i+1,j-1} & 0 & -a_{i+1,j+1} & \cdots & -a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ -a_{n1} & \cdots & -a_{n,j-1} & 0 & -a_{n,j+1} & \cdots & \lambda - a_{nn} \end{pmatrix}. \quad (7.11)$$

Entonces

$$\det [(\lambda I_n - A)(ij)] = (-1)^{i+j} \Delta_{n-1}^{(ij)}(\lambda);$$

multiplicando los dos miembros de esta última igualdad por \$(-1)^{i+j}\$, se deduce que

$$\Delta_{n-1}^{(ij)}(\lambda) = (-1)^{i+j} \det [(\lambda I_n - A)(ij)]. \quad (7.12)$$

Consideremos el digrafo \$G_1((\lambda I_n - A)(ij))\$ de la matriz \$(\lambda I_n - A)(ij)\$; este digrafo se obtiene a partir del digrafo \$G(\lambda I_n - A)\$ añadiendo un arco de \$j\$ a \$i\$ de peso 1 y suprimiendo todos los arcos de \$G(\lambda I_n - A)\$ que empiezan en \$j\$ y todos los que terminan en \$i\$. Para desarrollar el determinante de \$(\lambda I_n - A)(ij)\$ nos interesan las diagonales de esta matriz que tienen un uno en el lugar \$j\$-ésimo (esto es, las diagonales que contienen al 1 de la posición \$(i, j)\$). Estas diagonales se corresponden precisamente con las familias de ciclos generadoras de \$G_1((\lambda I_n - A)(ij))\$ tales que uno de sus ciclos contiene al arco \$(j, i)\$ de peso 1. A esta familias les llamaremos \$(ij)\$-familias. Para cada \$\nu = 1, \dots, n\$, haciendo consideraciones parecidas a las realizadas en la 2ª Demostración del Teorema 7.3.1 se tiene que cada \$(ij)\$-familia de ciclos generadora \$h\$ de \$G_1((\lambda I_n - A)(ij))\$ con \$n-\nu\$ lazos de peso de la forma \$\lambda - a\$, \$a \in \mathbb{F}\$ y \$c\$ ciclos de peso perteneciente a \$\mathbb{F}\$, aporta el sumando

$$(-1)^{n-[n-\nu+c]} (-1)^{\nu-1} f(\tilde{h}) \lambda^{n-\nu} = (-1)^{c-1} f(\tilde{h}) \lambda^{n-\nu} = (-1)^{c+1} f(\tilde{h}) \lambda^{n-\nu}$$

al determinante \$\det [(\lambda I_n - A)(ij)]\$, siendo \$\tilde{h}\$ la correspondiente familia de ciclos de longitud \$\nu\$ en el digrafo \$G_1^{(ij)}(A)\$ obtenido de \$G(A)\$ al añadirle un arco de \$j\$ a \$i\$ de peso 1 y suprimirle todos

los demás arcos que comienzan (resp. finalizan) en j (resp. i); para indicar que \tilde{h} contiene al arco (j, i) le llamamos (ij) -familia. El número de ciclos de \tilde{h} es igual a c ; denotamos este número por $d(\tilde{h})$.

Todo esto nos indica que $p_\nu^{(ij)}$ se obtiene considerando cada (ij) -familia de ciclos de longitud ν , \tilde{h} , de $G_1^{(ij)}(A)$ y multiplicando su peso $f(\tilde{h})$ por el signo $(-1)^{i+j+d(\tilde{h})+1}$ (ver (7.12)),

$$p_\nu^{(ij)} = \sum_{\tilde{h} \in H_\nu^{(ij)}} (-1)^{i+j+d(\tilde{h})+1} f(\tilde{h}), \quad (7.13)$$

siendo $H_\nu^{(ij)}$ el conjunto de las (ij) -familias de ciclos de $G_1^{(ij)}(A)$ de longitud ν . Si el conjunto $H_\nu^{(ij)}$ es vacío, debe entenderse que $p_\nu^{(ij)} = 0$.

Si eliminamos las filas i_1 e i_2 y las columnas j_1 y j_2 en la matriz $\lambda I_n - A$ (supondremos que $i_1 < i_2$ y $j_1 < j_2$) obtenemos una matriz $(n-2) \times (n-2)$

$$(\lambda I_n - A) \begin{pmatrix} i_1 & i_2 \\ j_1 & j_2 \end{pmatrix}^c, \quad (7.14)$$

cuyo determinante, $\Delta_{n-2}^{(i_1 j_1, i_2 j_2)}(\lambda)$, es igual al polinomio de grado a lo más $n-2$:

$$p_2^{(i_1 j_1, i_2 j_2)} \lambda^{n-2} + \dots + p_{n-1}^{(i_1 j_1, i_2 j_2)} \lambda + p_n^{(i_1 j_1, i_2 j_2)}. \quad (7.15)$$

llamemos $(\lambda I_n - A)(i_1 j_1, i_2 j_2)$ a la matriz $n \times n$ obtenida a partir de $\lambda I_n - A$ cambiando los elementos de sus filas i_1 e i_2 y sus columnas j_1 y j_2 por ceros y poniendo un uno en las posiciones (i_1, j_1) e (i_2, j_2) . Es fácil ver que

$$\begin{aligned} \det [(\lambda I_n - A)(i_1 j_1, i_2 j_2)] &= (-1)^{i_1+j_1+i_2-1+j_2-1} \Delta_{n-2}^{(i_1 j_1, i_2 j_2)}(\lambda) \\ &= (-1)^{i_1+j_1+i_2+j_2} \Delta_{n-2}^{(i_1 j_1, i_2 j_2)}(\lambda). \end{aligned} \quad (7.16)$$

Multiplicando por $(-1)^{i_1+j_1+i_2+j_2}$ los dos miembros de esta ecuación obtenemos

$$\Delta_{n-2}^{(i_1 j_1, i_2 j_2)}(\lambda) = (-1)^{i_1+j_1+i_2+j_2} \det [(\lambda I_n - A)(i_1 j_1, i_2 j_2)]. \quad (7.17)$$

Consideremos el digrafo $G_2((\lambda I_n - A)(i_1 j_1, i_2 j_2))$ de la matriz $(\lambda I_n - A)(i_1 j_1, i_2 j_2)$; este digrafo se deduce a partir del digrafo $G(\lambda I_n - A)$ añadiendo dos arcos de j_1 a i_1 y de j_2 a i_2 , respectivamente, cada uno con peso 1, y suprimiendo todos los demás arcos de $G(\lambda I_n - A)$ que empiezan en j_1 ó j_2 y todos los arcos que acaban en i_1 ó i_2 . Para desarrollar el determinante de $(\lambda I_n - A)(i_1 j_1, i_2 j_2)$ nos interesan las diagonales de esta matriz que tienen un 1 en las posiciones (i_1, j_1) e (i_2, j_2) . Estas diagonales se corresponden exactamente con las familias de ciclos generadoras de $G_2((\lambda I_n - A)(i_1 j_1, i_2 j_2))$ tales que los arcos (j_1, i_1) y (j_2, i_2) con peso 1 están contenidos en uno (o dos) de sus ciclos; les llamaremos $(i_1 j_1, i_2 j_2)$ -familias. Para cada $\nu = 2, \dots, n$, considerando las $(i_1 j_1, i_2 j_2)$ -familias de ciclos generadoras h de $G_2((\lambda I_n - A)(i_1 j_1, i_2 j_2))$ que contienen precisamente $n - \nu$ lazos de peso de la forma $\lambda - a$, con $a \in \mathbb{F}$, y los restantes c ciclos con pesos en \mathbb{F} , se tiene que h aporta el sumando

$$(-1)^{n-[n-\nu+c]} (-1)^{\nu-2} f(\tilde{h}) \lambda^{n-\nu} = (-1)^{c+2} f(\tilde{h}) \lambda^{n-\nu}$$

al determinante $\det [(\lambda I_n - A)(i_1 j_1, i_2 j_2)]$, siendo \tilde{h} la familia de ciclos subordinada por h en el digrafo $G_2^{(i_1 j_1, i_2 j_2)}(A)$ obtenido de $G(A)$ al añadirle dos arcos de j_1 a i_1 y de j_2 a i_2 , de peso 1, y suprimiendo todos los arcos que comienzan en $\{j_1, j_2\}$ y los que finalizan en $\{i_1, i_2\}$. Como la familia \tilde{h} contiene los arcos (j_1, i_1) y (j_2, i_2) se llama también $(i_1 j_1, i_2 j_2)$ -familia.

Por (7.17), cada $(i_1 j_1, i_2 j_2)$ -familia de ciclos \tilde{h} de $G_2^{(i_1 j_1, i_2 j_2)}(A)$ de longitud ν , aporta a $p_\nu^{(i_1 j_1, i_2 j_2)}$ su peso $f(\tilde{h})$ multiplicado por el signo $(-1)^{i_1+j_1+i_2+j_2+d(\tilde{h})+2}$ (véase 7.17), donde $d(\tilde{h})$ es el número de ciclos de \tilde{h} ($d(\tilde{h})$ coincide con el número c de la familia h de antes):

$$p_\nu^{(i_1 j_1, i_2 j_2)} = \sum_{\tilde{h} \in H_\nu^{(i_1 j_1, i_2 j_2)}} (-1)^{i_1+j_1+i_2+j_2+d(\tilde{h})+2} f(\tilde{h}), \quad (7.18)$$

siendo $H_\nu^{(i_1 j_1, i_2 j_2)}$ el conjunto de las $(i_1 j_1, i_2 j_2)$ -familias de ciclos de $G_2^{(i_1 j_1, i_2 j_2)}(A)$ de longitud ν . Si $H_\nu^{(i_1 j_1, i_2 j_2)}$ es un conjunto vacío, se entiende que $p_\nu^{(i_1 j_1, i_2 j_2)} = 0$.

De manera similar puede deducirse el resultado siguiente para los menores de orden $n - k$ para $k = 1, 2, 3, \dots, n - 1$. Sean $1 \leq i_1 < i_2 < \dots < i_k \leq n$, $1 \leq j_1 < j_2 < \dots < j_k \leq n$; llamemos

$$G_k^{(i_1 j_1, i_2 j_2, \dots, i_k j_k)}(A)$$

al digrafo obtenido a partir de $G(A)$ mediante las modificaciones siguientes: se añaden k arcos de peso 1 de j_1 a i_1 , de j_2 a i_2 , \dots y de j_k a i_k ; se eliminan todos los arcos restantes que comienzan en los vértices j_1, j_2, \dots, j_k y todos los restantes que acaban en los vértices i_1, i_2, \dots, i_k .

Teorema 7.3.2. Sea $\Delta_{n-k}^{(i_1 j_1, i_2 j_2, \dots, i_k j_k)}(\lambda)$ el determinante de la matriz $(n - k) \times (n - k)$

$$(\lambda I_n - A) \begin{pmatrix} i_1 & i_2 & \dots & i_k \\ j_1 & j_2 & \dots & j_k \end{pmatrix}^c$$

obtenida al tachar las filas i_1, i_2, \dots, i_k y las columnas j_1, j_2, \dots, j_k de la matriz característica $\lambda I_n - A$. Sea

$$\Delta_{n-k}^{(i_1 j_1, i_2 j_2, \dots, i_k j_k)}(\lambda) = p_k^{(i_1 j_1, i_2 j_2, \dots, i_k j_k)} \lambda^{n-k} + \dots + p_{n-1}^{(i_1 j_1, i_2 j_2, \dots, i_k j_k)} \lambda + p_n^{(i_1 j_1, i_2 j_2, \dots, i_k j_k)}. \quad (7.19)$$

Entonces para cada $\nu = k, k + 1, \dots, n$,

$$p_\nu^{(i_1 j_1, i_2 j_2, \dots, i_k j_k)} = \sum_{\tilde{h} \in H_\nu^{(i_1 j_1, i_2 j_2, \dots, i_k j_k)}} (-1)^{i_1 + j_1 + i_2 + j_2 + \dots + i_k + j_k + d(\tilde{h}) + k} f(\tilde{h}),$$

donde $H_\nu^{(i_1 j_1, i_2 j_2, \dots, i_k j_k)}$ es el conjunto de las $(i_1 j_1, i_2 j_2, \dots, i_k j_k)$ -familias de ciclos de $G_k^{(i_1 j_1, i_2 j_2, \dots, i_k j_k)}(A)$ de longitud ν , $d(\tilde{h})$ es el número de ciclos “disjuntos” de \tilde{h} y $f(\tilde{h})$ es el peso del subdigrafo \tilde{h} de $G_k^{(i_1 j_1, i_2 j_2, \dots, i_k j_k)}(A)$.

7.4. Ejercicios

Ejercicio 7.1. Caracterizar mediante su digrafo de Coates las matrices genéricas $M \in \{0, *\}^{n \times n}$ que son diagonalizables para *casi todo* valor complejo concreto asignado a sus asteriscos (*). ¿Cuándo tienen simples todos sus valores propios estas matrices M genéricamente diagonalizables?

Ejercicio 7.2. Caracterizar mediante su digrafo de Coates las matrices genéricas $M \in \{0, *\}^{n \times n}$ que son no derogatorias para *casi todo* valor complejo concreto asignado a sus asteriscos (*).

Ejercicio 7.3. Para cualquier matriz $A \in \mathbb{C}^{n \times n}$ denotaremos por $i(A)$ el número de factores invariantes de A no constantes, contando repetidos. Nótese que $i(A)$ es la multiplicidad geométrica más grande de los valores propios de A . Dado un entero k , $1 \leq k \leq n$, caracterizar mediante su digrafo de Coates las matrices genéricas $M \in \{0, *\}^{n \times n}$ que satisfacen $i(M) = k$ para *casi todo* valor complejo concreto asignado a sus asteriscos (*). Obsérvese que para el caso $k = 1$ las matrices A que satisfacen la igualdad $i(A) = 1$ son las no derogatorias. Así pues, este ejercicio es una generalización del Ejercicio 7.2 .

Ejercicio 7.4. Sea p un entero, $0 \leq p \leq n$. Para cualquier matriz $A \in \mathbb{C}^{n \times n}$ denotaremos por $i_p(A)$ el número de factores invariantes de A de grado $\geq p$, contando repetidos. Dado un entero k , $1 \leq k \leq n$, caracterizar mediante su digrafo de Coates las matrices genéricas $M \in \{0, *\}^{n \times n}$ que satisfacen $i_p(M) = k$ para *casi todo* valor complejo concreto asignado a sus asteriscos (*). Obsérvese que este ejercicio es una generalización del Ejercicio 7.3 .

Ejercicio 7.5. Sea $M \in \{0, *\}^{n \times n}$ una matriz genérica. Demostrar que existe un entero positivo p tal que $M^p = 0$ (calculando en el cuerpo de dos elementos $\{0, *\}$) si y sólo si el digrafo de Coates $G(M)$ de M es acíclico.

7.5. Notas al Capítulo 7

La primera referencia que me puso sobre la pista del tema de este capítulo fue el artículo de Poljak et al. [\[51\]](#).

Capítulo 8

Entornos de seguridad para la forma de Jordan

En este capítulo damos entornos de seguridad para las condiciones necesarias que debe satisfacer la forma canónica de Jordan bajo pequeñas perturbaciones. Es decir, dada una matriz $A \in \mathbb{C}^{n \times n}$ y un número $\delta > 0$ *adecuado* a A hallaremos un número $\varepsilon > 0$ tal que toda matriz A' con $\|A' - A\| < \varepsilon$ satisfará las condiciones (6.2) y (6.3) del Teorema 6.1.1.

Capítulo 9

Subespacios invariantes estables

Capítulo 10

Jordanización de una función matricial

Sea $A : (a, b) \rightarrow \mathbb{C}^{n \times n}$ una función matricial de *tipo* \mathcal{T} . En este capítulo estudiaremos bajo qué condiciones existe una función matricial $P : (a, b) \rightarrow \text{GL}_n(\mathbb{C})$ de *tipo* \mathcal{T} tal que para todo $t \in (a, b)$, la matriz $P(t)^{-1}A(t)P(t)$ es la forma canónica de Jordan de $A(t)$. Hay varios casos especiales del *tipo* de función: función analítica o de clase C^ω ; función de clase C^p y función suave o de clase C^∞ .

10.1. Puntos de bifurcación

Let $A : (a, b) \rightarrow \mathbb{C}^{n \times n}$ be a matrix function of class C^p defined on an open interval (a, b) of \mathbb{R} . Then for all $t \in (a, b)$, the matrix $A(t)$ is similar to a matrix J_t in Jordan form:

$$A(t) \approx J_t;$$

so, there exists an invertible matrix $P_t \in \mathbb{C}^{n \times n}$ such that

$$P_t^{-1}A(t)P_t = J_t.$$

In the following we try the solution of the

PROBLEM: *Can the function $t \mapsto P_t$ be chosen of class C^p ?*

If such a function $P : (a, b) \rightarrow \text{GL}_n(\mathbb{C})$ exists then $J(t) := P(t)^{-1}A(t)P(t)$, $t \in (a, b)$, must be necessarily a function

$$J : (a, b) \rightarrow \mathbb{C}^{n \times n}$$

of class C^p .

Before giving a partial (but *generic*) answer to this question, we are going to put two examples.

Example 1.- Let $A_1(t) := \begin{pmatrix} t^3 + ti & 1 \\ 0 & |t|^3 + ti \end{pmatrix}$ for every $t \in \mathbb{R}$. The matrix function $A_1 : \mathbb{R} \rightarrow \mathbb{C}^{2 \times 2}$ is of class C^2 . The eigenvalues of $A_1(t)$ are $t^3 + ti$ and $-t^3 + ti$, distinct, if $t < 0$; when $t \geq 0$ $A_1(t)$ has only one eigenvalue $t^3 + ti$. The representative curves in the complex plane of $\alpha_1(t) := t^3 + ti$ and $\alpha_2(t) := |t|^3 + ti$ are given in Figure 10.1 and in Figure 10.2

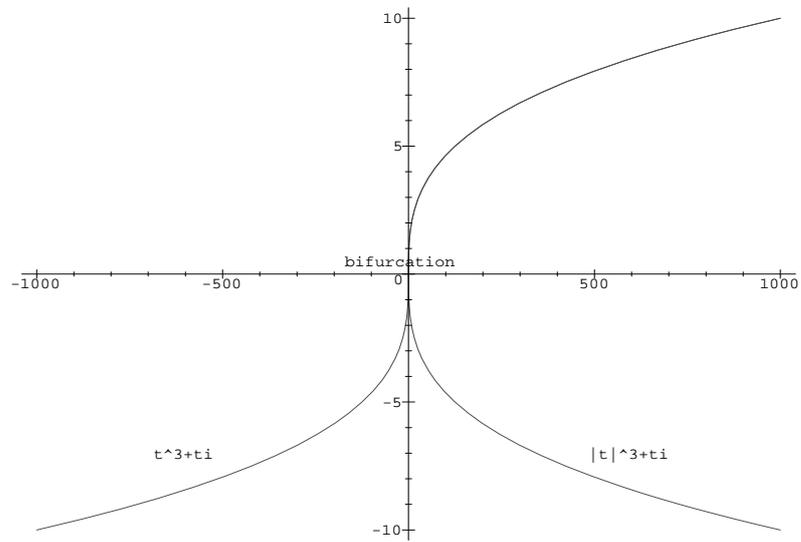


Figura 10.1: Punto de bifurcación en $t_0 = 0$.

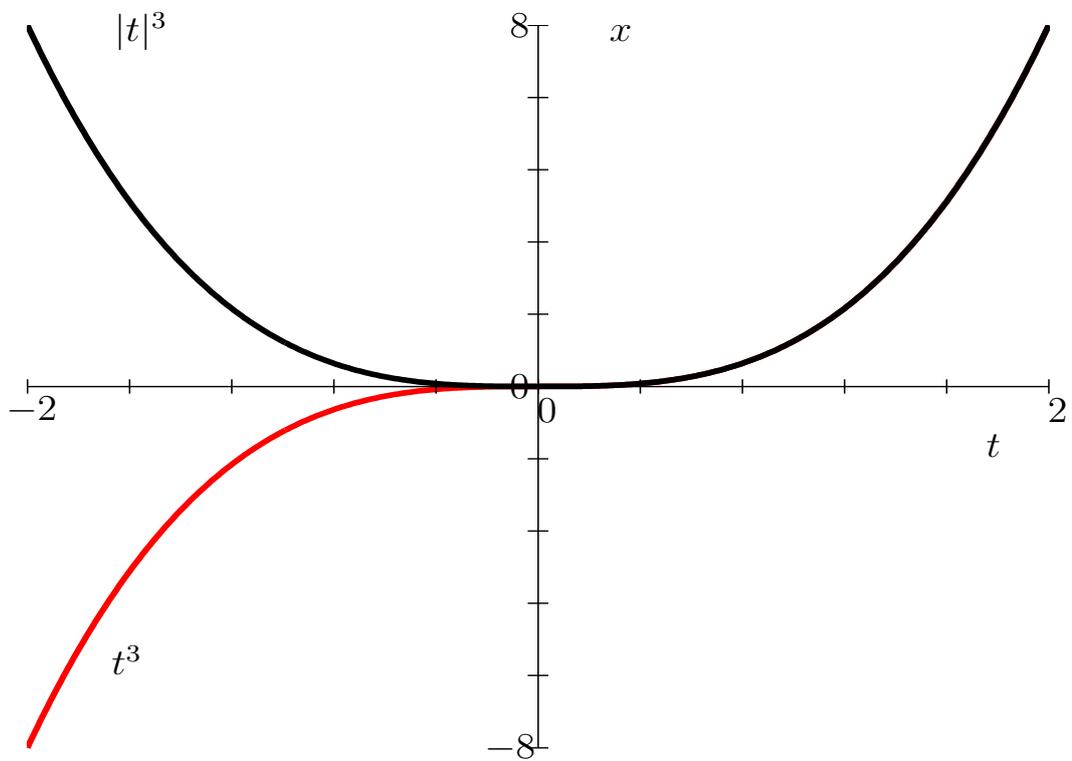


Figura 10.2: Punto de bifurcación en $t_0 = 0$.

The Jordan form of $A_1(t)$ is $J_t := \begin{pmatrix} t^3 + ti & 0 \\ 0 & -t^3 + ti \end{pmatrix}$ if $t < 0$, and $J_t := \begin{pmatrix} t^3 + ti & 1 \\ 0 & t^3 + ti \end{pmatrix}$ if $t \geq 0$. Therefore, given that the entry (1, 2) of J_t jumps from 0 to 1 when t pass through $t = 0$ from left to right, there does not exist a function $P : \mathbb{R} \rightarrow GL_2(\mathbb{C})$ of class C^2 , not even continuous at $t = 0$, such that for all $t \in \mathbb{R}$,

$$P(t)^{-1} A_1(t) P(t) = J_t.$$

Example 2.- Let $A_2(t) = \begin{pmatrix} t^3 + ti & 0 \\ 0 & |t|^3 + ti \end{pmatrix}$. The functions (of class C^2) $\alpha_1(t) := t^3 + ti$ and $\alpha_2(t) := |t|^3 + ti$ are such that the spectrum, $\Lambda(A_2(t))$, is equal to $\{\alpha_1(t), \alpha_2(t)\}$. When $t < 0$, $\alpha_1(t) \neq \alpha_2(t)$; when $t \geq 0$, $\alpha_1(t) = \alpha_2(t)$. But, in this case the constant matrix function $P(t) := I_2$ for all $t \in \mathbb{R}$ is of class C^2 (even of class C^∞) and for all $t \in \mathbb{R}$,

$$I_2^{-1} A_2(t) I_2 = \begin{pmatrix} t^3 + ti & 0 \\ 0 & |t|^3 + ti \end{pmatrix} = \begin{pmatrix} \alpha_1(t) & 0 \\ 0 & \alpha_2(t) \end{pmatrix} =: J_t.$$

Definition 1.- In these examples $t_0 = 0$ is a *point of bifurcation* of the spectrum of A_1 and A_2 . Let us precise this concept. Let $n(t) := \text{Card} \{ \mu \in \mathbb{C} : \det(A(t) - \mu I) = 0 \}$ for every $t \in (a, b)$. We say that a point $t_0 \in (a, b)$ is a *point of bifurcation* of the spectrum of A if in every neighbourhood of t_0 the function n takes two or more different values.

Definición 10.1.1. Sea $A : \Omega \rightarrow \mathbb{C}^{n \times n}$ una función matricial continua, definida en un intervalo abierto Ω de \mathbb{R} . Para cada $t \in \Omega$, sea

$$N_A(t) := \text{Card} \{ z \in \mathbb{C} : \det(zI_n - A(t)) = 0 \}.$$

Decimos que $t_0 \in \Omega$ es un *punto de bifurcación del espectro* de A si en todo entorno de t_0 existe un punto t que satisface $N_A(t) > N_A(t_0)$.

Llamemos $p_A(\lambda, t) := \det(\lambda I_n - A(t))$; de la ecuación (4.8) se sigue que para cada $t \in \Omega$,

$$N_A(t) = \text{rg} R \left(p_A(\lambda, t), \frac{\partial p_A(\lambda, t)}{\partial \lambda}, \lambda \right) - n + 1. \tag{10.1}$$

Por la Proposición 1.3.1, la función $t \mapsto N_A(t)$ es semicontinua inferiormente en cada $t_1 \in \Omega$. Por lo tanto, existe un entorno V_{t_1} de t_1 tal que para todo $t \in V_{t_1}$, se da la desigualdad $N_A(t) \geq N_A(t_1)$. De donde se llega a la siguiente observación.

Observación 10.1.1. Un punto $t_0 \in \Omega$ no es punto de bifurcación del espectro de A si existe un entorno W de t_0 tal que para todo $t \in W$,

$$N_A(t) = N_A(t_0).$$

Definition 2.- If $M \in \mathbb{C}^{n \times n}$ and λ_0 is eigenvalue of M , the *root subspace* of M corresponding to λ_0 is the subspace

$$\mathcal{R}_{\lambda_0}(M) := \text{Ker} (M - \lambda_0 I)^n.$$

Lema 10.1.1. Let $A : (a, b) \rightarrow \mathbb{C}^{n \times n}$ be a matrix function of class C^p . Let $t_1 \in (a, b)$ be such that there exists a neighbourhood $V(t_1)$ of t_1 , $V(t_1) \subset (a, b)$, satisfying that

$$\text{Card} \Lambda(A(t))$$

is constant on $V(t_1)$, and equal to k . Then there exist k different functions $\alpha_i : V(t_1) \rightarrow \mathbb{C}$ of class C^p , $1 \leq i \leq k$, such that: (1) for all $t \in V(t_1)$,

$$\Lambda(A(t)) = \{ \alpha_1(t), \dots, \alpha_k(t) \};$$

(2) moreover, for every $j \in \{1, \dots, k\}$, there exists the limit (in the gap metric)

$$\lim_{t \rightarrow t_1} \mathcal{R}_{\alpha_j(t)}(A(t))$$

and it is equal to $\mathcal{R}_{\alpha_j(t_1)}(A(t_1))$.

Proof. By Lemma 4.2. of [23] there exist k different functions $\alpha_1, \dots, \alpha_k : V(t_1) \rightarrow \mathbb{C}$ of class C^p such that for all $t \in V(t_1)$,

$$\Lambda(A(t)) = \{\alpha_1(t), \dots, \alpha_k(t)\}.$$

Let η be a positive real number such that the open balls $B(\alpha_j(t_1), \eta)$ ($1 \leq j \leq k$) are pairwise disjoint. Then there exists a connected neighbourhood $\mathcal{U}(t_1) \subset V(t_1)$ of t_1 such that for all $t \in \mathcal{U}(t_1)$ $\alpha_j(t) \in B(\alpha_j(t_1), \eta)$ ($1 \leq j \leq k$).

As $A(t)$ has k distinct roots $\alpha_1(t), \dots, \alpha_k(t)$ on $\mathcal{U}(t_1)$, $\mathcal{U}(t_1)$ is connected, and α_i ($i = 1, \dots, k$) are continuous, by Lemma 2.1 of [24] it follows that the algebraic multiplicities of these roots m_1, \dots, m_k are constant on $\mathcal{U}(t_1)$.

By Proposition 2.2.5 of [31], it follows that

$$\dim \mathcal{R}_{\alpha_j(t)}(A(t)) = m_j$$

is constant on $\mathcal{U}(t_1)$ ($1 \leq j \leq k$).

But $\mathcal{R}_{\alpha_j(t)}(A(t)) = \text{Ker}[(A(t) - \alpha_j(t)I)^n]$ for every $t \in \mathcal{U}(t_1)$. Let us take a sequence $\tau_1, \tau_2, \dots, \tau_q, \dots \in \mathcal{U}(t_1)$ converging to t_1 . By Theorem 15.1.3 of [31] we have that the limit

$$\lim_{q \rightarrow \infty} \text{Ker}[(A(\tau_q) - \alpha_j(\tau_q)I)^n]$$

exists and it is equal to

$$\text{Ker}[(A(t_1) - \alpha_j(t_1)I)^n].$$

As this is true for every sequence $\tau_1, \tau_2, \dots, \tau_q, \dots$ of points of $\mathcal{U}(t_1)$ converging to t_1 , we deduce that there exists the limit

$$\lim_{t \rightarrow t_1} \mathcal{R}_{\alpha_j(t)}(A(t))$$

and it is equal to $\mathcal{R}_{\alpha_j(t_1)}(A(t_1))$ ($1 \leq j \leq k$).

□

Lema 10.1.2. Let $A : (a, b) \rightarrow \mathbb{C}^{n \times n}$ of class C^p , let

$$r := \max_{t \in (a, b)} \text{Card} \Lambda(A(t)).$$

Let $\alpha_1, \dots, \alpha_r : (a, b) \rightarrow \mathbb{C}$ be continuous functions such that for every $t \in (a, b)$,

$$\Lambda(A(t)) = \{\alpha_1(t), \dots, \alpha_r(t)\}.$$

Let $t_0 \in (a, b)$ be a point of bifurcation of the spectrum of A .

Then, there exist a number $\epsilon > 0$ and two subsets $\{j_1, \dots, j_u\}$ and $\{k_1, \dots, k_v\}$ of $\{1, \dots, r\}$ such that for every $t \in (t_0 - \epsilon, t_0)$, the numbers $\alpha_{j_1}(t), \dots, \alpha_{j_u}(t)$ are the distinct eigenvalues of $A(t)$, and for every $t \in (t_0, t_0 + \epsilon)$, the numbers $\alpha_{k_1}(t), \dots, \alpha_{k_v}(t)$ are the distinct eigenvalues of $A(t)$.

Proof. It follows from results in [23].

□

A partial solution to the Problem is here:

Teorema 10.1.1. *Let $A : (a, b) \rightarrow \mathbb{C}^{n \times n}$ be a matrix function of class C^p . Let*

$$r := \max_{t \in (a, b)} \text{Card } \Lambda(A(t)).$$

Suppose that the points of bifurcation of the spectrum of A , if there exist, are isolated.

Then, in order that there exists a function $P : (a, b) \rightarrow \mathbb{C}^{n \times n}$ of class C^p such that for every $t \in (a, b)$, the matrix $P(t)$ is invertible and the matrix $P(t)^{-1} A(t) P(t)$ is a Jordan matrix, it is necessary and sufficient that the two following conditions are satisfied:

Condition 1: *For every $k = 1, 2, \dots, n$, the number of $k \times k$ Jordan blocks of $A(t)$ is constant on (a, b) .*

Condition 2: *There exist functions $\alpha_1, \dots, \alpha_r : (a, b) \rightarrow \mathbb{C}$ of class C^p such that:*

(2.1) *for all $t \in (a, b)$, $\Lambda(A(t)) = \{\alpha_1(t), \dots, \alpha_r(t)\}$,*

(2.2) *for each point of bifurcation t_0 of the spectrum of A , let $\epsilon > 0$ and $\{j_1, \dots, j_u\}, \{k_1, \dots, k_v\}$ be the real number and the subsets of $\{1, \dots, r\}$ associated to t_0 as in Lemma 2; then it must hold that*

$$\bigoplus_{i=1}^u \lim_{t \rightarrow t_0^-} \mathcal{R}_{\alpha_{j_i}(t)}(A(t)) = \mathbb{C}^n = \bigoplus_{i=1}^v \lim_{t \rightarrow t_0^+} \mathcal{R}_{\alpha_{k_i}(t)}(A(t)).$$

Proof. It follows the argumentation lines of [55].

□

Revisited Examples

The number 0 is the only bifurcation point of the spectrum of A_1 and A_2 .

If $t < 0$,

$$\begin{aligned} A_1(t) - \alpha_1(t) I_2 &= \begin{pmatrix} 0 & 1 \\ 0 & -2t^3 \end{pmatrix}; \\ A_1(t) - \alpha_2(t) I_2 &= \begin{pmatrix} 2t^3 & 1 \\ 0 & 0 \end{pmatrix}; \end{aligned}$$

from here,

$$\mathcal{R}_{\alpha_1(t)}(A_1(t)) = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle; \quad \mathcal{R}_{\alpha_2(t)}(A_1(t)) = \left\langle \begin{pmatrix} 1 \\ -2t^3 \end{pmatrix} \right\rangle.$$

We deduce that

$$\lim_{t \rightarrow 0^-} \mathcal{R}_{\alpha_1(t)}(A_1(t)) = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle; \quad \lim_{t \rightarrow 0^-} \mathcal{R}_{\alpha_2(t)}(A_1(t)) = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle.$$

But the sum $\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle + \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle$ is not direct and also $\neq \mathbb{C}^2$. So it is lacking the condition 2.2 of Theorem 1. It is not possible to reduce smoothly the matrix function A_1 to Jordan form.

On the other hand, if $t < 0$,

$$\begin{aligned} A_2(t) - \alpha_1(t) I_2 &= \begin{pmatrix} 0 & 0 \\ 0 & -2t^3 \end{pmatrix}; \\ A_2(t) - \alpha_2(t) I_2 &= \begin{pmatrix} 2t^3 & 0 \\ 0 & 0 \end{pmatrix}; \end{aligned}$$

we derive

$$\mathcal{R}_{\alpha_1(t)}(A_2(t)) = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle; \quad \mathcal{R}_{\alpha_2(t)}(A_2(t)) = \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle$$

If $t > 0$,

$$A_2(t) - \alpha_1(t) I_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix};$$

from here,

$$\mathcal{R}_{\alpha_1(t)}(A_2(t)) = \mathbb{C}^2.$$

It is easy too see that

$$\lim_{t \rightarrow 0^-} \mathcal{R}_{\alpha_1(t)}(A_2(t)) \oplus \lim_{t \rightarrow 0^-} \mathcal{R}_{\alpha_2(t)}(A_2(t)) = \mathbb{C}^2 = \lim_{t \rightarrow 0^+} \mathcal{R}_{\alpha_1(t)}(A_2(t)).$$

So, this corroborates that the matrix function A_2 is smoothly reducible to Jordan form.

En primer lugar probemos la aserción: *Un punto de acumulación de puntos de bifurcación es también un punto de bifurcación.*

Sea $\tau_0 \in (a, b)$ un punto de acumulación de puntos de bifurcación del espectro de $A : (a, b) \rightarrow \mathbb{C}^{n \times n}$. Entonces τ_0 es punto de bifurcación del espectro de A . En efecto, en todo entorno (abierto) \mathcal{N} de τ_0 existen puntos de bifurcación t_0 . Luego \mathcal{N} es entorno de t_0 . Por consiguiente, la función n no es constante sobre \mathcal{N} . Como esto ocurre para todo entorno de τ_0 , se sigue que τ_0 es punto de bifurcación.

Así pues, si suponemos que los puntos de bifurcación del espectro de A son *puntos aislados*, se seguirá que el conjunto de estos puntos de bifurcación no puede tener puntos de acumulación que pertenezcan a (a, b) . Pero puede suceder que a o b sean puntos de acumulación de dicho conjunto.

Lema 10.1.3. *Sea $A : (a, b) \rightarrow \mathbb{C}^{n \times n}$ una función matricial de clase C^1 tal que los puntos de bifurcación de su espectro, si existen, son aislados.*

Entonces existen n funciones continuas $\lambda_1, \dots, \lambda_n : (a, b) \rightarrow \mathbb{C}$ tales que para todo $t \in (a, b)$,

$$\Lambda(A(t)) = \{\lambda_1(t), \dots, \lambda_n(t)\}. \quad (10.2)$$

[Pese a la notación de la ecuación (10.2) para algunos valores de t el conjunto $\Lambda(A(t))$ puede tener menos de n elementos].

Si $t'_0 < t_0 < t''_0$, con $t'_0, t_0, t''_0 \in (a, b)$, son tres puntos de bifurcación consecutivos, entonces $\text{Card } \Lambda(A(t))$ es constante en cada uno de los intervalos (t'_0, t_0) y (t_0, t''_0) , digamos que

$$\text{Card } \Lambda(A(t)) = \begin{cases} u & \text{si } t \in (t'_0, t_0), \\ v & \text{si } t \in (t_0, t''_0). \end{cases}$$

Además existen dos subconjuntos $\{j_1, \dots, j_u\}$ y $\{k_1, \dots, k_v\}$ de $\{1, \dots, n\}$ tales que para todo $t \in (t'_0, t_0)$ los números $\lambda_{j_1}(t), \dots, \lambda_{j_u}(t)$ son los u valores propios distintos de $A(t)$, y para todo $t \in (t_0, t''_0)$ los números $\lambda_{k_1}(t), \dots, \lambda_{k_v}(t)$ son los v valores propios distintos de $A(t)$; las funciones $\lambda_{j_1}, \dots, \lambda_{j_u}$ son de clase C^1 en (t'_0, t_0) ; las funciones $\lambda_{k_1}, \dots, \lambda_{k_v}$ son de clase C^1 en (t_0, t''_0) .

Nota.- En este momento, ignoro si existen las derivadas laterales en t_0 :

$$\lambda'_{j_\alpha}(t_0^-) := \lim_{h \rightarrow 0^-} \frac{\lambda_{j_\alpha}(t_0 + h) - \lambda_{j_\alpha}(t_0)}{h}, \quad (\alpha = 1, \dots, u),$$

$$\lambda'_{k_\beta}(t_0^+) := \lim_{h \rightarrow 0^+} \frac{\lambda_{k_\beta}(t_0 + h) - \lambda_{k_\beta}(t_0)}{h}, \quad (\beta = 1, \dots, v).$$

Tampoco sé si existen los límites laterales

$$\lim_{t \rightarrow t_0^-} \lambda'_{j_\alpha}(t), \quad (\alpha = 1, \dots, u),$$

$$\lim_{t \rightarrow t_0^+} \lambda'_{k_\beta}(t), \quad (\beta = 1, \dots, v).$$

DEMOSTRACIÓN DEL LEMA 10.1.3

Por el Corolario VI.1.6, pág. 155 de [12], existen n funciones continuas $\lambda_1, \dots, \lambda_n : (a, b) \rightarrow \mathbb{C}$ tales que para todo $t \in (a, b)$,

$$\Lambda(A(t)) = \{\lambda_1(t), \dots, \lambda_n(t)\}.$$

Por ser t'_0 y t_0 dos puntos de bifurcación del espectro de A consecutivos, se sigue que $\Lambda(A(t))$ tiene un número constante de elementos al variar t en (t'_0, t_0) ; digamos que $\text{Card } \Lambda(A(t)) = u$ para todo $t \in (t'_0, t_0)$. Por el Lema 4.1 de [24] si dos de las funciones $\lambda_1, \dots, \lambda_n$ —digamos λ_i y λ_j — coinciden en un punto τ_0 de (t'_0, t_0)

$$\lambda_i(\tau_0) = \lambda_j(\tau_0),$$

entonces $\lambda_i(t) = \lambda_j(t)$ para todo $t \in (t'_0, t_0)$. Esto permite entresacar el subconjunto $\{\lambda_{j_1}, \dots, \lambda_{j_u}\}$ de $\{\lambda_1, \dots, \lambda_n\}$ formado por las u funciones continuas tales que para todo $t \in (t'_0, t_0)$,

$$\Lambda(A(t)) = \{\lambda_{j_1}(t), \dots, \lambda_{j_u}(t)\}.$$

Por el Lema 2.1 de [24] las multiplicidades algebraicas m_1, \dots, m_u de $\lambda_{j_1}(t), \dots, \lambda_{j_u}(t)$, respectivamente, son constantes en (t'_0, t_0) .

Para cada $\alpha = 1, \dots, u$, veamos que existe la derivada

$$\lambda'_{j_\alpha}(\tau)$$

en cada punto $\tau \in (t'_0, t_0)$ y que la función λ'_{j_α} es continua en τ .

Como $\lambda_{j_1}(\tau), \dots, \lambda_{j_\alpha}(\tau), \dots, \lambda_{j_u}(\tau)$ son u números complejos distintos, existen u discos disjuntos dos a dos centrados en ellos de radio $r > 0$, $B(\lambda_{j_\alpha}(\tau), r)$, $\alpha = 1, \dots, u$. Llamemos c_α a la circunferencia exterior de $B(\lambda_{j_\alpha}(\tau), r)$ orientada en sentido positivo. Por la continuidad de la función λ_{j_α} en τ , existe un entorno $V(\tau) \subset (t'_0, t_0)$ de τ tal que para todo $t \in V(\tau)$, $\lambda_{j_\alpha}(t)$ es la única raíz de

$$f(\lambda, t) := \det(\lambda I - A(t)) = 0$$

que hay dentro de c_α . Por el Teorema B de variable compleja, se tiene que $\forall t \in V(\tau)$,

$$\lambda_{j_\alpha}(t) = \frac{1}{m_\alpha} \frac{1}{2\pi i} \oint_{c_\alpha} \frac{\lambda f'_\lambda(\lambda, t)}{f(\lambda, t)} d\lambda. \quad (10.3)$$

De (10.3) y del Teorema D de variable compleja se sigue que existe $\lambda'_{j_\alpha}(t)$ para todo $t \in V(\tau)$ y que

$$\lambda'_{j_\alpha}(t) = \frac{1}{m_\alpha} \frac{1}{2\pi i} \oint_{c_\alpha} \frac{\partial}{\partial t} \frac{\lambda f'_\lambda(\lambda, t)}{f(\lambda, t)} d\lambda. \quad (10.4)$$

Luego λ'_{j_α} es continua en cada $t \in V(\tau)$.

La demostración se completa de forma análoga. □

Material nuevo. Nueva versión del teorema principal: Theorem 1.1

10.2. Puntos de bifurcación de la característica de Segré

Sea $A : (a, b) \mapsto \mathbb{C}^{n \times n}$ una función matricial de clase C^p en el intervalo abierto (a, b) de \mathbb{R} . Para cada $t \in (a, b)$ sea $s(A(t))$ la característica de Segré de la matriz $A(t)$ (recordar la definición de la página 49). Sea $s(t) := s(A(t))$ para todo $t \in (a, b)$. Así se define una función

$$s : (a, b) \rightarrow \bigcup_{u=1}^n \mathcal{P}^u,$$

con valores en la unión los productos cartesianos $\mathcal{P}^u := \mathcal{P} \times \dots \times \mathcal{P}$ del conjunto \mathcal{P} de particiones de enteros.

Definición 10.2.1. Se dice que el punto $t_0 \in (a, b)$ es un *punto de bifurcación de la característica de Segré* de A si la función s es no constante en todo entorno de t_0 .

Definición 10.2.2. Sea $n(t) := \text{Card} \{ \mu \in \mathbb{C} : \det(A(t) - \mu I) = 0 \}$ para todo $t \in (a, b)$. Decimos que un punto $t_0 \in (a, b)$ es un *punto de bifurcación del espectro* de A si la función n es no constante en todo entorno de t_0 .

Es obvio que si la función n es no constante en un subintervalo (a_1, b_1) de (a, b) , entonces la función s es no constante en (a_1, b_1) . Por consiguiente, si t_0 es un punto de bifurcación del espectro de A , se sigue que t_0 es un punto de bifurcación de la característica de Segré de A . Pero puede haber puntos de bifurcación de la característica de Segré que no son puntos de bifurcación del espectro, como se puede ver en el ejemplo siguiente.

Sea $\lambda_1 : \mathbb{R} \rightarrow \mathbb{C}$ una función de clase C^p y sea $g : \mathbb{R} \rightarrow \mathbb{R}$ la función definida por

$$g(t) := \begin{cases} 0 & \text{para } t \leq 0 \\ e^{-1/t^2} & \text{para } t > 0. \end{cases}$$

Es conocido que esta función g es indefinidamente derivable (o de clase C^∞) en \mathbb{R} . Véase [20, pág. 156]. Definimos la función matricial

$$A(t) := \begin{pmatrix} \lambda_1(t) & 1 & 0 \\ 0 & \lambda_1(t) & g(t) \\ 0 & 0 & \lambda_1(t) \end{pmatrix}$$

Es fácil ver que la característica de Segré de $A(t)$ es $((2, 1))$ si $t \leq 0$ y $((3))$ si $t > 0$. Por lo tanto, 0 es un punto de bifurcación de la característica de Segré de A ; pero está claro que el espectro de A no se bifurca en $t = 0$.

Lema 10.2.1. Sea $A : (a, b) \rightarrow \mathbb{C}^{n \times n}$ una función matricial de clase C^p . Sea

$$r := \max_{t \in (a, b)} \text{Card } \Lambda(A(t)).$$

Si $t_0 \in (a, b)$ no es punto de bifurcación del espectro de A , entonces existen un número $\varepsilon > 0$ y q funciones $\lambda_1, \dots, \lambda_q : (a, b) \rightarrow \mathbb{C}$ de clase C^p , $q \leq r$, tales que para cada $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$:

$$\Lambda(A(t)) = \{\lambda_1(t), \dots, \lambda_q(t)\}$$

y los números $\lambda_1(t), \dots, \lambda_q(t)$ son distintos.

Teorema 10.2.1. Sea $A : (a, b) \rightarrow \mathbb{C}^{n \times n}$ una función matricial de clase C^p . Sea

$$r := \max_{t \in (a, b)} \text{Card } \Lambda(A(t)).$$

Supongamos que los puntos de bifurcación de la característica de Segré de A , si existen, son aislados. Entonces, para que exista una función $P : (a, b) \rightarrow \mathbb{C}^{n \times n}$ de clase C^p tal que para todo $t \in (a, b)$, la matriz $P(t)$ sea invertible y

$$P(t)^{-1} A(t) P(t)$$

sea una matriz en forma canónica de Jordan, es necesario y suficiente que las dos condiciones siguientes sean satisfechas:

Condición 1: La unión de las particiones de la característica de Segré de $A(t)$ es constante en (a, b) .

Condición 2: Existen funciones $\alpha_1, \dots, \alpha_r : (a, b) \rightarrow \mathbb{C}$ de clase C^p tales que:

(2.1) para todo $t \in (a, b)$, $\Lambda(A(t)) = \{\alpha_1(t), \dots, \alpha_r(t)\}$,

(2.2) para cada punto de bifurcación t_0 del espectro de A , sean $\varepsilon > 0$ y $\{j_1, \dots, j_u\}, \{k_1, \dots, k_v\}$ el número real y los subconjuntos de $\{1, \dots, r\}$ asociados a t_0 como en el Lema 1.1; entonces debe ser

$$\bigoplus_{i=1}^u \lim_{t \rightarrow t_0^-} \mathcal{R}_{\alpha_{j_i}(t)}(A(t)) = \mathbb{C}^n = \bigoplus_{i=1}^v \lim_{t \rightarrow t_0^+} \mathcal{R}_{\alpha_{k_i}(t)}(A(t)).$$

(2.3) para cada punto de bifurcación t_0 de la característica de Segré de A que no sea punto de bifurcación del espectro de A , sean $\varepsilon > 0$ y $\{\alpha_{k_1}, \dots, \alpha_{k_q}\}$ el número real y el subconjunto de funciones asociados a t_0 como en el Lema 10.2.1; entonces debe ser

$$\mathbb{C}^n = \bigoplus_{i=1}^q \lim_{t \rightarrow t_0} \mathcal{R}_{\alpha_{k_i}(t)}(A(t)).$$

10.3. Funciones matriciales genéricas

Denotemos por $\mathcal{E}^p := C^p((a, b), \mathbb{C}^{n \times n})$ el conjunto de funciones matriciales $A : (a, b) \rightarrow \mathbb{C}^{n \times n}$ de clase C^p definidas en un intervalo (a, b) de la recta real, de clase C^p . Queremos probar que las funciones matriciales $A \in \mathcal{E}^p$ que tienen aislado el conjunto de puntos de bifurcación de su característica de Segré, forman un subconjunto genérico de \mathcal{E}^p ; es decir, un subconjunto abierto y denso. Para ello, previamente debemos definir una topología sobre \mathcal{E}^p . Hay varias formas posibles de hacerlo.

En \mathcal{E}^p vamos a considerar la topología de la convergencia uniforme sobre conjuntos compactos. Para todo compacto $K \subset (a, b)$ y cada número $i \in \{1, \dots, p\}$, se define

$$N_{K,i}(A) := \max_{t \in K} \|A^{(i)}(t)\|, \quad A \in \mathcal{E}^p,$$

donde $A^{(i)}(t)$ denota el valor de la derivada i -ésima de la función matricial A en el punto t . La función $N_{K,i}$ es una seminorma sobre \mathcal{E}^p .

Al variar K e i , se obtiene una familia de seminormas sobre \mathcal{E}^p . La topología definida sobre \mathcal{E}^p por esta familia se llama la topología de la convergencia compacta para todas las derivadas de orden $\leq p$. Véanse las páginas 179 y 187 de [17].

Teorema 10.3.1. *Sea el espacio $\mathcal{E}^p := C^p((a, b), \mathbb{C}^{n \times n})$, dotado con la topología de la convergencia compacta. Sea el subconjunto \mathcal{B} de \mathcal{E}^p formado por las funciones matriciales cuyos puntos de bifurcación de la característica de Segré son aislados. Entonces,*

- (i) *Para cada $A \in \mathcal{B}$ existen conjuntos compactos K_1, K_2, \dots, K_μ , enteros $i_1, i_2, \dots, i_\mu \in \{1, \dots, p\}$ y números reales $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_\mu > 0$, tales que para toda función matricial $X \in \mathcal{E}^p$ que satisfaga las condiciones*

$$\max_{t \in K_q} \|X^{(i_q)}(t) - A^{(i_q)}(t)\| < \varepsilon_q, \quad q = 1, \dots, \mu,$$

los puntos de bifurcación de la característica de Segré de X son aislados.

- (ii) *Para cada $X \in \mathcal{E}^p$, y para cada selección de conjuntos compactos K_1, K_2, \dots, K_ν , enteros $i_1, i_2, \dots, i_\nu \in \{1, \dots, p\}$ y números reales $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_\nu > 0$, existe una función matricial $B \in \mathcal{B}$ que satisface las condiciones*

$$\max_{t \in K_q} \|X^{(i_q)}(t) - B^{(i_q)}(t)\| < \varepsilon_q, \quad q = 1, \dots, \nu.$$

El enunciado del teorema nos dice que el conjunto \mathcal{B} es abierto y denso.

Capítulo 11

Forma de Jordan probabilística

No creo que pueda desarrollar el tema de este capítulo. Dejaré escrito aquí algo de los indicios que me llevaron a preverlo. Ví un artículo de B. Kågström y sus colaboradores en el que se relataba el siguiente experimento estadístico: dado un haz de matrices reales $m \times n$, $\lambda B - A$, se tomaban 1000 haces al azar a una distancia $\varepsilon = 10^{-3}, 10^{-6}$ de $\lambda B - A$ y se contaban en una tabla el número de casos que en que el haz elegido tenía una determinada estructura de Kronecker. Es sabido que la estructura de Kronecker de un haz está formada por los índices minimales por filas y por columnas, y por los exponentes de sus divisores elementales finitos e infinitos. Sólo hay un número finito de estructuras de Kronecker posibles, determinado por los números naturales m y n . Teniendo en cuenta los teoremas sobre la perturbación de la forma canónica de Kronecker de un haz, los haces que disten menos de ε de $\lambda B - A$ únicamente pueden tener una estructura de Kronecker que no sea repelida por el haz $\lambda B - A$. De este modo cabía hacer una estimación estadística de las probabilidades de alcanzar cada una de esas estructuras de Kronecker no repelidas por $\lambda B - A$. Es obvio que un estudio teórico de estas probabilidades será una cuestión más difícil que las tratadas hasta aquí en este libro.

El tema a esbozar aquí sería hacer una versión probabilística del Teorema 6.2.2 de la página 112 del Capítulo 6. Sea $A \in \mathbb{C}^{n \times n}$ y sea \mathcal{E} es una estructura de Jordan de orden n admisible por A . Dado $\varepsilon > 0$, hallar la probabilidad de que una matriz A' elegida al azar con distribución uniforme en la bola $B(A, \varepsilon) \subset \mathbb{C}^{n \times n}$ tenga como característica de Segré la estructura \mathcal{E} .

11.1. Referencias comentadas

- [1] B. Kågström
- [2] J.W. Demmel: A numerical analyst's Jordan canonical form. Ph. D. Thesis, University of California, Berkeley, 1983.
- [3] F. Chatelin
- [4] A.J. Hetzel, J.S. Liew and K.E. Morrison: The probability that a matrix of integers is diagonalizable, *Amer. Math. Monthly*, **114**, no. 6 (2007) 491–499.
- [5] Z.N. Zhang: The Jordan canonical form of a real random matrix, *Numer. Math. J. Chinese Univ.*, **23** (2001) 363–367. Citado por Hetzel y Morrison, pero está en chino.

Apéndice A

Transformaciones elementales de matrices

La clave en la Teoría de Matrices está en las transformaciones elementales en las líneas (filas o columnas) de una matriz que conservan alguna propiedad que interese de la matriz. De forma más precisa, se trata de realizar transformaciones que conserven algún invariante de una relación de equivalencia en cierto espacio de matrices.

A.1. Matrices sobre un cuerpo

Sea A una matriz $m \times n$ con elementos en el cuerpo \mathbb{C} . Sobre las filas (o columnas) de A podemos realizar una cualquiera de las *transformaciones elementales* siguientes:

- (1) intercambiar dos filas (o columnas);
- (2) multiplicar todos los elementos de una fila (o columna) por algún número complejo no nulo;
- (3) sumar a cualquier fila (o columna) cualquier otra fila (o columna) multiplicada por un número complejo.

Obsérvese que el resultado de estas manipulaciones de las filas y columnas de A puede obtenerse también de otro modo: si la manipulación es sobre las filas (resp. columnas) de A y produce la matriz \tilde{A} , hágase la misma manipulación sobre las filas (resp. columnas) de la matriz identidad I_m (resp. I_n); se obtiene así una matriz E (resp. F) tal que el producto EA (resp. AF) es igual a \tilde{A} . Pueden verse detalles en el libro de Lancaster y Tismenetsky [41, Section 2.7, p. 47]. Las matrices E o F , así obtenidas, se llaman *matrices elementales*, y son siempre matrices invertibles.

Dado un espacio vectorial cualquiera V sobre \mathbb{C} , sean $\mathbf{v}_1, \dots, \mathbf{v}_p \in V$. Consideremos el subespacio vectorial U de V engendrado por estos p vectores. Es bien sabido que el subespacio U no cambia si transformamos su sistema de generadores $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ de alguna de las maneras siguientes:

- (1) $\{\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_i, \dots, \mathbf{v}_p\}$ con $1 \leq i < j \leq p$;
- (2) $\{\mathbf{v}_1, \dots, \alpha \mathbf{v}_i, \dots, \mathbf{v}_p\}$ con $\alpha \in \mathbb{C}, \alpha \neq 0$;
- (3) $\{\mathbf{v}_1, \dots, \mathbf{v}_i + \beta \mathbf{v}_j, \dots, \mathbf{v}_j, \dots, \mathbf{v}_p\}$ con $\beta \in \mathbb{C}$.

Por lo tanto, como el rango de una matriz $A \in \mathbb{C}^{m \times n}$ es igual a la dimensión del subespacio que engendran sus filas (resp. columnas), el rango de A no cambia cuando ejecutamos una transformación elemental sobre sus filas (resp. columnas).

Es conocido que dadas dos matrices $A, B \in \mathbb{C}^{m \times n}$, existen matrices invertibles $P \in \mathbb{C}^{m \times m}, Q \in \mathbb{C}^{n \times n}$ tales que $PAQ = B$ si y sólo si partiendo de A existe una sucesión finita de transformaciones elementales sobre las filas y/o columnas de las matrices que se van obteniendo que acaba convirtiendo A en B . Esto puede hacerse si y sólo si A y B tienen el mismo rango. Si $\text{rg } A = r$, una manera práctica de obtener una P y una Q tales que $PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ es la siguiente: se disponen las matrices I_m, A e I_n de la siguiente forma

$$\begin{array}{c} I_n \\ A \quad I_m \end{array} ;$$

cuando se hacen transformaciones elementales por filas se actúa sobre toda la fila de la matriz $m \times (n+m)$ $[A, I_m]$; cuando se hacen por columnas se cambia toda la columna de la matriz $(n+m) \times n$

$$\begin{bmatrix} I_n \\ A \end{bmatrix}.$$

El objetivo es llevar A a la forma $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$; cuando se llega a ello, las matrices que aparecen en los lugares de I_m e I_n son P y Q , respectivamente. El orden en el que se efectúan las transformaciones elementales no tiene importancia y se pueden intercalar las que se hacen sobre las filas con las que actúan sobre las columnas. En forma esquemática:

$$\begin{array}{c} I_n \\ A \quad I_m \end{array} \longrightarrow \text{transformaciones elementales} \longrightarrow \begin{array}{c} Q \\ \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \quad P \end{array}.$$

Ni que decir tiene que las matrices P y Q , así obtenidas, no son únicas, pues dependen de las transformaciones elementales concretas realizadas. Sobrentenderemos que el símbolo

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

denota la matriz

$$\begin{cases} \begin{bmatrix} I_m, 0 \end{bmatrix} & \text{si } r = m \leq n, \\ \begin{bmatrix} I_n \\ 0 \end{bmatrix} & \text{si } r = n \leq m, \\ I_n & \text{si } r = m = n. \end{cases}$$

A.2. Matrices polinómicas

Sea $\mathbb{C}[\lambda]$ el anillo de los polinomios en una variable λ con coeficientes números complejos. Se dice que un polinomio de grado n es *mónico* si el coeficiente de su término de mayor grado es 1; es decir, si es de la forma

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0.$$

Las matrices $m \times n$ cuyos elementos son polinomios de $\mathbb{C}[\lambda]$ se llaman *matrices polinómicas o polinomiales* (también λ -matrices en libros antiguos) y al conjunto de todas ellas lo denotamos por $\mathbb{C}[\lambda]^{m \times n}$.

Sea $A(\lambda)$ una matriz polinómica $m \times n$ con elementos en el anillo de polinomios $\mathbb{C}[\lambda]$. Sobre las filas (o columnas) de $A(\lambda)$ podemos realizar una cualquiera de las *transformaciones elementales* siguientes:

- (1) intercambiar dos filas (o columnas);
- (2) multiplicar todos los elementos de una fila (o columna) por algún número complejo no nulo;

- (3) sumar a cualquier fila (o columna) cualquier otra fila (o columna) multiplicada por un polinomio de $\mathbb{C}[\lambda]$.

Obsérvese que el resultado de estas manipulaciones de las filas y columnas de $A(\lambda)$ puede obtenerse también de otro modo: si la manipulación es sobre las filas (resp. columnas) de $A(\lambda)$ y produce la matriz $\tilde{A}(\lambda)$, hágase la misma manipulación sobre las filas (resp. columnas) de la matriz identidad I_m (resp. I_n); se obtiene así una matriz $E(\lambda)$ (resp. $F(\lambda)$) tal que el producto $E(\lambda)A(\lambda)$ (resp. $A(\lambda)F(\lambda)$) es igual a $\tilde{A}(\lambda)$. Pueden verse detalles en el libro de Lancaster y Tismenetsky [41, Section 7.3, p. 253]. A las matrices polinómicas $E(\lambda)$ o $F(\lambda)$, así obtenidas, se les llama *matrices polinómicas elementales* y son siempre matrices *unimodulares*, esto quiere decir que sus determinantes son constantes no nulas. Lo cual es condición necesaria y suficiente para que su inversa sea una matriz de polinomios. Si se piensa en una construcción de la matriz inversa de una matriz polinómica cuadrada $U(\lambda)$ que empieza hallando los adjuntos, ... y acaba dividiendo por el determinante de la matriz, se ve fácilmente que la condición es suficiente. Si existe una matriz polinómica $W(\lambda)$ tal que $U(\lambda)W(\lambda) = I$, se sigue que

$$\det U(\lambda) \det W(\lambda) = 1.$$

Así, el producto de los polinomios $\det U(\lambda)$ y $\det W(\lambda)$ es una constante no nula. Esto es posible sólo si cada uno de ellos es una constante no nula.

Vamos a definir una relación de equivalencia en el conjunto $\mathbb{C}[\lambda]^{m \times n}$. Dos matrices $A(\lambda), B(\lambda) \in \mathbb{C}[\lambda]^{m \times n}$ se dicen *equivalentes* si $B(\lambda)$ puede ser obtenida a partir de $A(\lambda)$ mediante una sucesión finita de transformaciones elementales. Así pues, $A(\lambda)$ y $B(\lambda)$ son equivalentes si existen matrices polinómicas elementales $E_1(\lambda), \dots, E_k(\lambda), F_1(\lambda), \dots, F_s(\lambda)$ tales que

$$E_k(\lambda) \cdots E_1(\lambda)A(\lambda)F_1(\lambda) \cdots F_s(\lambda) = B(\lambda),$$

o, equivalentemente,

$$U(\lambda)A(\lambda)V(\lambda) = B(\lambda),$$

donde $U(\lambda) := E_k(\lambda) \cdots E_1(\lambda)$ y $V(\lambda) := F_1(\lambda) \cdots F_s(\lambda)$ son matrices polinómicas unimodulares.

A.3. Forma canónica de Smith

Se llama *rango* de una matriz polinómica $A(\lambda) \in \mathbb{C}^{m \times n}$ al orden del mayor menor que es distinto del polinomio cero. Sea k un entero tal que $1 \leq k \leq r$. Se llama *divisor determinantal* de $A(\lambda)$ de orden k al máximo común divisor mónico de todos los menores de orden k de $A(\lambda)$. Se le denota por $D_k(\lambda)$. Se conviene que $D_0(\lambda) := 1$. Se puede demostrar que el rango y los divisores determinantes son invariantes para la relación de equivalencia de matrices de $\mathbb{C}[\lambda]^{m \times n}$. Véase [41, Propositions 1 and 2, Section 7.5, p. 259].

Si $f(\lambda)$ y $g(\lambda)$ son polinomios de $\mathbb{C}[\lambda]$ se usa la notación $f(\lambda) \mid g(\lambda)$ para indicar que $f(\lambda)$ divide a $g(\lambda)$. En [41, Sections 7.4 and 7.5, p. 256–262] puede verse la demostración del teorema siguiente.

Teorema A.3.1. *Una matriz polinómica $A(\lambda) \in \mathbb{C}[\lambda]^{m \times n}$ de rango r es equivalente a una única matriz $m \times n$ de la forma*

$$\text{diag}[f_1(\lambda), \dots, f_r(\lambda), 0, \dots, 0] = \begin{pmatrix} f_1(\lambda) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & f_2(\lambda) & & 0 & 0 & & 0 \\ \vdots & & \ddots & & \vdots & & \vdots \\ 0 & 0 & & f_r(\lambda) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & & 0 \\ \vdots & & & \vdots & & \ddots & \\ 0 & 0 & \cdots & 0 & 0 & & 0 \end{pmatrix},$$

donde los elementos de la diagonal son polinomios mónicos tales que $f_1(\lambda) \mid f_2(\lambda) \mid \cdots \mid f_r(\lambda)$.

La forma canónica

$$S(\lambda) := \text{diag}[f_1(\lambda), \dots, f_r(\lambda), 0, \dots, 0]$$

se llama *forma canónica o normal de Smith*. Los polinomios $f_1(\lambda), f_2(\lambda), \dots, f_r(\lambda)$ se llaman los *factores invariantes* de $A(\lambda)$. Junto con el rango de $A(\lambda)$ constituyen un sistema completo de invariantes para la equivalencia de matrices polinómicas.

Existe la siguiente relación entre los factores invariantes y los divisores determinantaes de $A(\lambda)$:

$$f_1(\lambda) = \frac{D_1(\lambda)}{D_0(\lambda)}, \quad f_2(\lambda) = \frac{D_2(\lambda)}{D_1(\lambda)}, \quad \dots, \quad f_r(\lambda) = \frac{D_r(\lambda)}{D_{r-1}(\lambda)}. \quad (\text{A.1})$$

De forma análoga a la que hemos indicado en la Sección A.1, una manera de obtener la forma normal de Smith $S(\lambda)$ de una matriz $A(\lambda) \in \mathbb{C}[\lambda]^{m \times n}$ y dos matrices unimodulares $U(\lambda) \in \mathbb{C}[\lambda]^{m \times m}$ y $V(\lambda) \in \mathbb{C}[\lambda]^{n \times n}$ tales que

$$U(\lambda)A(\lambda)V(\lambda) = S(\lambda),$$

es la siguiente: se disponen las matrices $I_m, A(\lambda)$ e I_n de la manera siguiente

$$\begin{array}{cc} I_n & \\ A(\lambda) & I_m \end{array};$$

cuando se hacen transformaciones elementales por filas se actúa sobre toda la fila de la matriz $m \times (n+m)$ $[A(\lambda), I_m]$; cuando se hacen por columnas se cambia toda la columna de la matriz $(n+m) \times n$

$$\begin{bmatrix} I_n \\ A(\lambda) \end{bmatrix}.$$

El objetivo es llevar $A(\lambda)$ a la forma

$$S(\lambda) := \text{diag}[f_1(\lambda), \dots, f_r(\lambda), 0, \dots, 0];$$

cuando se llega a ello, las matrices que aparecen en los lugares de I_m e I_n son $U(\lambda)$ y $V(\lambda)$, respectivamente. El orden en el que se efectúan las transformaciones elementales no tiene importancia y se pueden intercalar las que se hacen sobre las filas con las que actúan sobre las columnas. En forma esquemática:

$$\begin{array}{ccc} I_n & & V(\lambda) \\ A & I_m & \longrightarrow \text{transformaciones elementales} \longrightarrow \begin{array}{cc} S(\lambda) & U(\lambda) \end{array} \end{array}.$$

Las matrices $U(\lambda)$ y $V(\lambda)$, así obtenidas, no son únicas, pues dependen de las transformaciones elementales concretas realizadas.

A.4. Divisores elementales

Los factores invariantes caracterizan la clase de equivalencia de una matriz polinómica $m \times n$ $A(\lambda)$. Sin embargo, si $A(\lambda)$ es una matriz diagonal con bloques rectangulares, i.e. de la forma

$$A(\lambda) = \begin{pmatrix} A_1(\lambda) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_t(\lambda) \end{pmatrix},$$

la forma como dependen los factores invariantes de la matriz $A(\lambda)$ de los factores invariantes de los bloques rectangulares $A_1(\lambda), \dots, A_t(\lambda)$ resulta compleja. Por tanto, en ciertas cuestiones es preferible considerar no los factores invariantes, sino los llamados divisores elementales de la matriz $A(\lambda)$, cuya relación con los divisores elementales de los bloques $A_1(\lambda), \dots, A_t(\lambda)$ es mucho más sencilla.

Sean $f_1(\lambda) \mid \cdots \mid f_r(\lambda)$ los factores invariantes de la matriz $A(\lambda) \in \mathbb{C}[\lambda]^{m \times n}$ de rango r . Supongamos que $\{\lambda_1, \dots, \lambda_u\}$ es el conjunto de todos los números complejos que son raíces de alguno al menos de los polinomios $f_1(\lambda), \dots, f_r(\lambda)$. Por consiguiente, es posible factorizar cada uno de los polinomios $f_1(\lambda), \dots, f_r(\lambda)$ en potencias de la forma $(\lambda - \lambda_k)^p$ donde $k \in \{1, \dots, u\}$ y p es un entero no negativo, eventualmente nulo:

$$\begin{aligned} f_1(\lambda) &= (\lambda - \lambda_1)^{\alpha_{11}} (\lambda - \lambda_2)^{\alpha_{12}} \cdots (\lambda - \lambda_u)^{\alpha_{1u}}, \\ f_2(\lambda) &= (\lambda - \lambda_1)^{\alpha_{21}} (\lambda - \lambda_2)^{\alpha_{22}} \cdots (\lambda - \lambda_u)^{\alpha_{2u}}, \\ &\vdots \\ f_r(\lambda) &= (\lambda - \lambda_1)^{\alpha_{r1}} (\lambda - \lambda_2)^{\alpha_{r2}} \cdots (\lambda - \lambda_u)^{\alpha_{ru}}. \end{aligned}$$

Además, como $f_i(\lambda)$ divide a $f_{i+1}(\lambda)$ para $i = 1, \dots, r - 1$, se tiene que para cada $k = 1, 2, \dots, u$,

$$0 \leq \alpha_{1k} \leq \alpha_{2k} \leq \cdots \leq \alpha_{rk}.$$

Cada factor $(\lambda - \lambda_k)^{\alpha_{ik}}$ que aparece en las factorizaciones de los $f_i(\lambda)$ con $\alpha_{ik} > 0$ se llama un *divisor elemental* de $A(\lambda)$. Un divisor elemental para el que $\alpha_{ik} = 1$ se llama *lineal*; cuando $\alpha_{ik} > 1$ se llama *no lineal*. Si en estas circunstancias un cierto divisor elemental aparece en varios factores invariantes de $A(\lambda)$, lo escribiremos tantas veces como sean los factores invariantes en los que aparece; así pues, debe hablarse del *sistema* de los divisores elementales y no del *conjunto* de divisores elementales de $A(\lambda)$. Para cada número λ_k nos referiremos a los divisores elementales de la forma $(\lambda - \lambda_k)^{\alpha_{ik}}$ como los asociados a λ_k . Obsérvese que sobre \mathbb{C} todos los divisores elementales son potencias de polinomios mónicos de grado uno.

Pueden calcularse los factores invariantes de la matriz polinómica

$$\begin{pmatrix} \lambda & \lambda^2 & 0 & 0 & 0 \\ \lambda^3 & \lambda^5 & 0 & 0 & 0 \\ 0 & 0 & 2\lambda & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

y ver que son $f_1(\lambda) = \lambda, f_2(\lambda) = \lambda, f_3(\lambda) = \lambda^5 - \lambda^4$. De donde, por factorización, se deduce que sus divisores elementales son $\lambda, \lambda, \lambda^4, \lambda - 1$.

Demostraremos que el sistema de divisores elementales de una matriz polinómica $A(\lambda)$ determina todos los factores invariantes no constantes, y que si a esto agregamos las dimensiones de $A(\lambda)$ y su rango, es posible determinar todos los factores invariantes y por tanto su clase de equivalencia.

Teorema A.4.1. *Las dimensiones m, n , el rango r y el sistema de divisores elementales de una matriz $A(\lambda) \in \mathbb{C}[\lambda]^{m \times n}$ determinan completamente todos sus factores invariantes.*

DEMOSTRACIÓN. La demostración queda explicada en términos del ejemplo que se expone a continuación. Supongamos que $A(\lambda)$ es una matriz 6×7 , que su rango es 4, y que sus divisores elementales sean

$$\lambda, \lambda^2, \lambda^2, \lambda + 1, (\lambda + 1)^3, \lambda - 1, \lambda - 1.$$

Como 6 es el mínimo de 6 y 7, $A(\lambda)$ tendrá cuatro factores invariantes

$$f_1(\lambda), f_2(\lambda), f_3(\lambda), f_4(\lambda)$$

y su forma normal de Smith será

$$S(\lambda) = \begin{pmatrix} f_1(\lambda) & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & f_2(\lambda) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & f_3(\lambda) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & f_4(\lambda) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Si desarrollamos $f_1(\lambda), \dots, f_4(\lambda)$ en factores deberían obtenerse los siete divisores elementales señalados. Como $f_4(\lambda)$ es divisible por $f_3(\lambda), f_2(\lambda)$ y $f_1(\lambda)$, tendremos que en $f_4(\lambda)$ aparecen los divisores elementales que están asociados a 0, -1 y 1 y, además, en los grados mayores. Por tanto,

$$f_4(\lambda) = \lambda^2(\lambda + 1)^3(\lambda - 1).$$

Entre los divisores elementales restantes $\lambda, \lambda^2, \lambda + 1, \lambda - 1$ los de grado mayor deben estar en $f_3(\lambda)$; por consiguiente,

$$f_3(\lambda) = \lambda^2(\lambda + 1)(\lambda - 1).$$

A su vez, los de mayor grado de los restantes deben formar $f_2(\lambda)$, a saber, $f_2(\lambda) = \lambda$. Como ya hemos distribuido todos los divisores elementales, tendremos que $f_1(\lambda) = 1$. Es obvio que el procedimiento seguido es aplicable a todos los casos, lo que demuestra el teorema. \square

Corolario 1. *Dos matrices polinómicas complejas $m \times n$ son equivalentes si y sólo si tienen igual rango y los mismos sistemas de divisores elementales.*

Para obtener los divisores elementales de una matriz polinómica que esté en la forma normal de Smith $S(\lambda) = \text{diag}[f_1(\lambda), \dots, f_r(\lambda), 0, \dots, 0]$, basta tomar, de acuerdo con la definición, todos los divisores elementales de sus elementos diagonales $f_1(\lambda), \dots, f_r(\lambda), 0, \dots, 0$. Demostraremos que esta misma regla es aplicable a una matriz polinómica “diagonal” arbitraria.

Lema A.4.1. *El sistema de divisores elementales de una matriz polinómica $m \times n$ de la forma*

$$D(\lambda) := \begin{pmatrix} d_1(\lambda) & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & d_m(\lambda) & 0 & \cdots & 0 \end{pmatrix},$$

donde suponemos que $m \leq n$ sin pérdida de generalidad, es la “unión” de los divisores elementales de sus elementos diagonales $d_1(\lambda), \dots, d_m(\lambda)$.

DEMOSTRACIÓN. Sin limitar la generalidad, podemos suponer que los m polinomios $d_1(\lambda), \dots, d_m(\lambda)$ son mónicos y diferentes de cero. Denotemos por $D_k(\lambda)$ el divisor determinantal de $D(\lambda)$ de orden k . Como los polinomios $d_1(\lambda), \dots, d_m(\lambda)$ son mónicos, tendremos que $D_m(\lambda)$ será igual al menor

$$\begin{vmatrix} d_1(\lambda) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_m(\lambda) \end{vmatrix}$$

de $D(\lambda)$; a saber,

$$D_m(\lambda) = d_1(\lambda) \cdots d_m(\lambda).$$

Pero, por (A.1) de la Sección A.3,

$$D_m(\lambda) = f_1(\lambda) \cdots f_m(\lambda),$$

donde $f_1(\lambda), \dots, f_m(\lambda)$ son los factores invariantes de la matriz $D(\lambda)$. Por tanto si $\lambda_1, \dots, \lambda_u$ son las diversas raíces complejas de al menos alguno de los factores invariantes, veremos que cada uno de los divisores elementales de la matriz $D(\lambda)$ será una potencia de uno de los polinomios $\lambda - \lambda_1, \dots, \lambda - \lambda_u$. Ahora, extraigamos de $d_1(\lambda), \dots, d_m(\lambda)$ la potencia mayor de $\lambda - \lambda_1$ por la que estos polinomios son divisibles. Sea

$$d_i(\lambda) = (\lambda - \lambda_1)^{s_i} c_i(\lambda),$$

donde $c_i(\lambda_1) \neq 0$. Es nuestro deseo demostrar que

$$(\lambda - \lambda_1)^{s_1}, \dots, (\lambda - \lambda_1)^{s_m},$$

es el sistema de divisores elementales de la matriz $D(\lambda)$ que están asociados a λ_1 . Como los divisores elementales de la matriz $A(\lambda)$ no se alteran al permutar sus filas y columnas, dispondremos éstas en tal orden que

$$s_1 \leq s_2 \leq \dots \leq s_m. \tag{A.1}$$

Ahora encontremos la potencia de $\lambda - \lambda_1$ que aparezca en $D_k(\lambda)$. Por definición, $D_k(\lambda)$ es el máximo común divisor mónico de los menores de orden k de la matriz $D(\lambda)$, donde los que son diferentes de cero tienen la forma

$$d_{\nu_1}(\lambda) \cdots d_{\nu_k}(\lambda) = (\lambda - \lambda_1)^{s_{\nu_1} + \dots + s_{\nu_k}} c_{\nu_1}(\lambda) \cdots c_{\nu_k}(\lambda).$$

En virtud de las desigualdades (A.1), la potencia mínima de $\lambda - \lambda_1$ está contenida en el menor

$$d_1(\lambda) \cdots d_k(\lambda) = (\lambda - \lambda_1)^{s_1 + \dots + s_k} c_1(\lambda) \cdots c_k(\lambda).$$

En consecuencia, $D_k(\lambda)$ contiene $\lambda - \lambda_1$ elevado a la potencia $s_1 + \dots + s_k$. Sustituyendo aquí k por $k - 1$, tendremos que $D_{k-1}(\lambda)$ contiene a $\lambda - \lambda_1$ elevado al grado $s_1 + \dots + s_{k-1}$. Pero, por (A.1), el factor invariante $f_k(\lambda)$ es igual al cociente $D_k(\lambda)/D_{k-1}(\lambda)$. Por tanto, $f_k(\lambda)$ contiene a $\lambda - \lambda_1$ exactamente a la potencia s_k . Así pues, los divisores elementales de la matriz $D(\lambda)$ asociados a λ_1 coinciden con los divisores elementales asociados a λ_1 de los elementos diagonales de la matriz $D(\lambda)$. Como nuestro argumento resulta válido para los restantes números $\lambda_2, \dots, \lambda_u$ el lema ha quedado demostrado. \square

Teorema A.4.2. *El sistema de divisores elementales de una matriz polinómica $A(\lambda) \in \mathbb{C}[\lambda]^{m \times n}$ diagonal con bloques rectangulares*

$$A(\lambda) = \begin{pmatrix} A_1(\lambda) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_t(\lambda) \end{pmatrix},$$

es igual a la “unión” de los sistemas de divisores elementales de sus bloques $A_1(\lambda), \dots, A_t(\lambda)$.

DEMOSTRACIÓN. Las transformaciones elementales efectuadas sobre los diversos bloques $A_1(\lambda), \dots, A_t(\lambda)$ se pueden considerar como transformaciones de toda la matriz $A(\lambda)$. Las transformaciones en cuestión no afectan la forma diagonal en bloques de la matriz $A(\lambda)$, y las transformaciones realizadas sobre un solo bloque no afectan la forma de los restantes. Por tanto, todos los bloques $A_1(\lambda), \dots, A_t(\lambda)$ se pueden reducir a su forma normal de Smith $S_1(\lambda), \dots, S_t(\lambda)$ en términos de transformaciones elementales de la matriz $A(\lambda)$.

Por consiguiente, existen matrices unimodulares $U(\lambda) \in \mathbb{C}[\lambda]^{m \times m}$, $V(\lambda) \in \mathbb{C}[\lambda]^{n \times n}$ tales que

$$U(\lambda)A(\lambda)V(\lambda) = \begin{pmatrix} S_1(\lambda) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & S_t(\lambda) \end{pmatrix}.$$

La matriz del miembro de la derecha de esta igualdad es “diagonal”. Por el Lema A.4.1 los divisores elementales de esta matriz se obtienen por la “unión” de los divisores elementales de todos sus elementos diagonales, a saber, en términos de la unión de los divisores elementales de todos los bloques $A_1(\lambda), \dots, A_t(\lambda)$. Lo cual, por el Corolario 1 del Teorema A.4.1, obtenemos que el sistema de divisores elementales de $A(\lambda)$ es la “unión” de los divisores elementales de los bloques $A_1(\lambda), \dots, A_t(\lambda)$. \square

A.5. Semejanza de matrices

Dos matrices $A, B \in \mathbb{C}^{n \times n}$ se llaman *semejantes* si existe una matriz invertible $P \in \mathbb{C}^{n \times n}$ tal que

$$P^{-1}AP = B.$$

Recordemos aquí uno de los teoremas más notables del Álgebra Lineal. La demostración de este teorema puede verse en la referencia [41, Section 7.6, p. 262]. Nos da un criterio racional de semejanza de matrices; por racional entendemos que es un criterio que se puede aplicar efectuando un número finito de operaciones racionales (suma, resta, multiplicación y división).

Teorema A.5.1. *Las matrices $A, B \in \mathbb{C}^{n \times n}$ son semejantes si y sólo si las matrices polinómicas $\lambda I - A$ y $\lambda I - B$ son equivalentes.*

Se llaman factores invariantes y divisores elementales de una matriz constante $A \in \mathbb{C}^{n \times n}$ a los factores invariantes y divisores elementales, respectivamente, de su *matriz característica* $\lambda I - A \in \mathbb{C}[\lambda]^{n \times n}$. En los últimos años algunos autores prefieren evitar cierta confusión y llaman polinomios invariantes de A a los factores invariantes de $\lambda I - A$. Para estos autores sólo tienen factores invariantes las matrices polinómicas.

Forma canónica racional

Existe para cada matriz $A \in \mathbb{C}^{n \times n}$ una matriz semejante a ella que tiene una forma especialmente simple: la llamada forma canónica racional o primera forma normal. Aunque nosotros trabajamos en general en el cuerpo \mathbb{C} es importante observar que esta forma canónica puede ser obtenida en el cuerpo arbitrario \mathbb{F} cuando $A \in \mathbb{F}^{n \times n}$.

La matriz polinómica $\lambda I - A$ tiene rango n , pues el determinante $\det(\lambda I - A)$ es distinto de cero, ya que es el polinomio característico de A . Por consiguiente $\lambda I - A$ tiene n factores invariantes $1, \dots, 1, f_s(\lambda), f_{s+1}(\lambda), \dots, f_n(\lambda)$, donde los $n - s$ últimos suponemos que tienen grado mayor o igual que uno.

Recordemos que se llama *matriz compañera* del polinomio $p(\lambda) = a_0 + a_1\lambda + \dots + a_{n-1}\lambda^{n-1} + \lambda^n$ a la matriz $n \times n$

$$C(p) := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & & & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} & \end{pmatrix}.$$

Es fácil demostrar que el polinomio característico de la matriz $C(p)$ es igual a p :

$$\det(\lambda I - C(p)) = p(\lambda). \quad (\text{A.1})$$

Más aun, todos los menores de $\lambda I - C(p)$ de órdenes $1, 2, \dots, n - 1$ son independientes de λ , por lo tanto, los divisores determinantes

$$D_1(\lambda), D_2(\lambda), \dots, D_{n-1}(\lambda)$$

de $\lambda I - C(p)$ son iguales a 1. Ya que, por definición, $D_n(\lambda) = \det(\lambda I - C(p))$, entonces por (A.1) de la Sección A.3 y (A.1) se tiene que los factores invariantes de la matriz compañera $C(p)$ son $f_1(\lambda) = 1, \dots, f_{n-1}(\lambda) = 1, f_n(\lambda) = p(\lambda)$.

Teorema A.5.2. *Sea $A \in \mathbb{C}^{n \times n}$ y supongamos que los factores invariantes de $\lambda I - A$ de grado ≥ 1 son $f_s(\lambda) \mid f_{s+1}(\lambda) \mid \cdots \mid f_n(\lambda)$. Entonces A es semejante a la matriz diagonal por bloques*

$$C_R(A) := \text{diag}[C(f_s), C(f_{s+1}), \dots, C(f_n)], \quad (\text{A.2})$$

donde $C(f_k)$ denota la matriz compañera asociada al factor invariante $f_k(\lambda)$, $s \leq k \leq n$.

La aplicación $A \mapsto C_R(A)$ es una forma canónica para la relación de semejanza en $\mathbb{C}^{n \times n}$; se le llama la forma canónica racional o primera forma normal. Una demostración de este teorema puede verse en [41, Theorem 2, p. 264, Section 7.6].

Por el Corolario 1 del Teorema A.4.1 y el Teorema A.5.1 se tiene la proposición siguiente.

Proposición A.5.1. *Dos matrices complejas $n \times n$ A y B son semejantes si y sólo si $\lambda I - A$ y $\lambda I - B$ tienen los mismos divisores elementales.*

A.6. Forma canónica de Jordan

Sea α un número complejo. Se llama *bloque de Jordan* de orden q asociado a α a la matriz $q \times q$

$$J_q(\alpha) := \begin{pmatrix} \alpha & 1 & 0 & \cdots & 0 \\ 0 & \alpha & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \alpha & 1 \\ 0 & \cdots & \cdots & 0 & \alpha \end{pmatrix}.$$

¿Cuáles son los divisores elementales de $J_q(\alpha)$? El lema que sigue contesta a esta pregunta.

Lema A.6.1. *La matriz $\lambda I - J_q(\alpha)$ solamente posee el divisor elemental $(\lambda - \alpha)^q$.*

DEMOSTRACIÓN. La matriz $\lambda I - J_q(\alpha)$ tiene la forma

$$\begin{pmatrix} \lambda - \alpha & -1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda - \alpha & -1 & & & 0 \\ & & & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & \\ 0 & & & & \lambda - \alpha & -1 \\ 0 & 0 & \cdots & 0 & \lambda - \alpha \end{pmatrix}.$$

calculemos el divisor determinantal $D_k(\lambda)$ de orden k de la matriz $\lambda I - J_q(\alpha)$. Ante todo, tenemos:

$$D_q(\lambda) = |\lambda I - J_q(\alpha)| = (\lambda - \alpha)^q.$$

Además, $D_{q-1}(\lambda)$ es el máximo común divisor mónico de todos los menores de orden $q - 1$. Pero, entre estos últimos, se encuentra el menor

$$\begin{vmatrix} -1 & 0 & \cdots & 0 & 0 \\ \lambda - \alpha & -1 & & & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & & \ddots & -1 & 0 \\ 0 & 0 & \cdots & \lambda - \alpha & -1 \end{vmatrix} = (-1)^{q-1},$$

que se ha obtenido eliminando la primera columna y la última fila de la matriz $\lambda I - J_q(\alpha)$. Como este menor es igual a ± 1 , $D_{q-1}(\lambda) = 1$. Denotemos por $f_1(\lambda), \dots, f_q(\lambda)$ los factores invariantes de la matriz $\lambda I - J_q(\alpha)$. De las relaciones

$$\begin{aligned} D_{q-1}(\lambda) &= f_1(\lambda) \cdots f_{q-1}(\lambda) = 1, \\ D_q(\lambda) &= f_1(\lambda) \cdots f_{q-1}(\lambda) f_q(\lambda) = (\lambda - \alpha)^q, \end{aligned}$$

se infiere que

$$f_1(\lambda) = \cdots = f_{q-1}(\lambda) = 1, \quad f_q(\lambda) = (\lambda - \alpha)^q.$$

En consecuencia, $\lambda I - J_q(\alpha)$ solamente tiene un divisor elemental, que es igual a $(\lambda - \alpha)^q$. \square

Sea $A \in \mathbb{C}^{n \times n}$. Sean $f_1(\lambda) \mid \cdots \mid f_n(\lambda)$ los factores invariantes de $\lambda I - A$. Sean $\lambda_1, \dots, \lambda_u \in \mathbb{C}$ los distintos valores propios de A . Entonces es posible factorizar cada factor invariante de grado distintos de 1 en producto de potencias de la forma

$$(\lambda - \lambda_i)^{n_{ij}} \quad \text{con } n_{ij} \text{ entero } > 0.$$

Asociado a cada divisor elemental de A (o de $\lambda I - A$) tenemos el bloque de Jordan $J_{n_{ij}}(\lambda_i)$. Considerando el Lema A.6.1, la Proposición A.5.1 y el Teorema A.4.2 se sigue el teorema siguiente.

Teorema A.6.1. Sea $A \in \mathbb{C}^{n \times n}$. Supongamos que $\lambda_1, \dots, \lambda_u$ son los distintos valores propios complejos de A . Sean

$$(\lambda - \lambda_i)^{n_{ij}}, \quad j = 1, \dots, r_i,$$

con $n_{i1} \geq n_{i2} \geq \dots \geq n_{ir_i}$ para $i = 1, \dots, u$, los divisores elementales de $\lambda I - A$ (o de A). Sea $J_{n_{ij}}(\lambda_i)$ el bloque de Jordan asociado al divisor elemental $(\lambda - \lambda_i)^{n_{ij}}$. Sea

$$J_i := \begin{pmatrix} J_{n_{i1}}(\lambda_i) & & & 0 \\ & J_{n_{i2}}(\lambda_i) & & \\ & & \ddots & \\ 0 & & & J_{n_{ir_i}}(\lambda_i) \end{pmatrix}, \quad i = 1, \dots, u.$$

Entonces la matriz A es semejante a la matriz

$$J := \begin{pmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_u \end{pmatrix}.$$

La matriz J se llama la *forma canónica (o normal) de Jordan* de A . Propiamente, no hemos definido una aplicación $A \mapsto J$ por lo que no podemos hablar de una forma canónica en el sentido de la definición que se da en la página 1 del Capítulo 1.

Pero de acuerdo con la Observación 1.1.1 de dicho capítulo, podemos modificar así la definición de forma canónica de Jordan en $\mathbb{C}^{n \times n}$. Para cada sucesión finita de números complejos distintos $\Lambda := (z_1, \dots, z_v)$ y para cada sucesión finita de particiones de enteros

$$\Omega := \left[(n_{11}, \dots, n_{1s_1}), \dots, (n_{v1}, \dots, n_{vs_v}) \right],$$

definimos $F(\Lambda, \Omega)$ como el conjunto de las $(s_1 + \dots + s_v)!$ matrices $n \times n$ diagonales por bloques formadas por las permutaciones de los bloques de Jordan $J_{n_{ij}}(z_i)$ en la matriz diagonal por bloques J :

$$J := \begin{pmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_v \end{pmatrix}.$$

siendo

$$J_i := \begin{pmatrix} J_{n_{i1}}(z_i) & & & 0 \\ & J_{n_{i2}}(z_i) & & \\ & & \ddots & \\ 0 & & & J_{n_{is_i}}(z_i) \end{pmatrix}, \quad i = 1, \dots, v.$$

Sea $\mathcal{P}(\mathbb{C}^{n \times n})$ el conjunto de estos subconjuntos finitos $F(\Lambda, \Omega)$. Sea $s(A)$ la característica de Segré de A y $\Sigma(A) = (\lambda_1, \dots, \lambda_u)$ la sucesión de valores propios distintos de A ordenados de acuerdo con la ordenación de las particiones de $s(A)$. Véase la página 49. Entonces la aplicación $A \mapsto F(\Sigma(A), s(A))$ es una forma canónica en el sentido de la mencionada Observación 1.1.1. Habitualmente se prescinde de este formalismo y se dice que la forma canónica de Jordan está determinada salvo permutaciones de los bloques de Jordan.

Matrices nilpotentes

Definición A.6.1. Se dice que una matriz $A \in \mathbb{C}^{n \times n}$ es *nilpotente* si $A^k = 0$ para cierto entero $k > 0$.

Proposición A.6.1. Una matriz $A \in \mathbb{C}^{n \times n}$ es nilpotente si y sólo si $\Lambda(A) = \{0\}$.

DEMOSTRACIÓN. Si existe un entero $k > 0$ tal que $A^k = 0$ entonces el polinomio mínimo, $m_A(\lambda)$, de A divide a λ^k ; por tanto, $m_A(\lambda) = \lambda^p$, con $p \leq k$. Como un número complejo es valor propio de A si y sólo si es raíz de $m_A(\lambda)$, se sigue que $\Lambda(A) = \{0\}$.

Recíprocamente, si 0 es el único valor propio de A , entonces la forma canónica de Jordan, J , de A es igual a la suma diagonal de bloques de Jordan de la forma $J_p(0)$. Como cada uno de estos bloques es nilpotente, se deduce que J es nilpotente, y, por ende, A es nilpotente. \square

A.7. Multiplicación de matrices de bloques

Con frecuencia es necesario emplear matrices partidas en bloques o submatrices a la hora de sumar y multiplicar matrices. Véase [41, Section 1.6, p. 16], [28, p. 43].

Las operaciones de matrices partidas se efectúan según las mismas reglas formales que el caso de matrices cuyos elementos son números.

Por ejemplo, supongamos que $A_{ij}, \tilde{A}_{ij} \in \mathbb{C}^{m_i \times n_j}$ para $i = 1, 2$ $j = 1, 2, 3$; entonces la suma “por bloques” se obtiene sumando los bloques correspondientes. Si

$$A := \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix}, \quad \tilde{A} := \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} & \tilde{A}_{13} \\ \tilde{A}_{21} & \tilde{A}_{22} & \tilde{A}_{23} \end{pmatrix},$$

entonces puede demostrarse trivialmente que

$$A + \tilde{A} = \begin{pmatrix} A_{11} + \tilde{A}_{11} & A_{12} + \tilde{A}_{12} & A_{13} + \tilde{A}_{13} \\ A_{21} + \tilde{A}_{21} & A_{22} + \tilde{A}_{22} & A_{23} + \tilde{A}_{23} \end{pmatrix}.$$

Detengámonos un poco más en el caso de la multiplicación. Sean $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times p}$ matrices partidas de acuerdo con el esquema siguiente:

$$A = \begin{matrix} & \begin{matrix} n_1 & n_2 & \cdots & n_t \end{matrix} \\ \begin{matrix} m_1 \\ m_2 \\ \vdots \\ m_s \end{matrix} & \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1t} \\ A_{21} & A_{22} & \cdots & A_{2t} \\ \vdots & \vdots & & \vdots \\ A_{s1} & A_{s2} & \cdots & A_{st} \end{pmatrix} \end{matrix}, \quad B = \begin{matrix} \begin{matrix} p_1 & p_2 & \cdots & p_u \end{matrix} \\ \begin{matrix} n_1 \\ n_2 \\ \vdots \\ n_t \end{matrix} & \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1u} \\ B_{21} & B_{22} & \cdots & B_{2u} \\ \vdots & \vdots & & \vdots \\ B_{t1} & B_{t2} & \cdots & B_{tu} \end{pmatrix} \end{matrix}.$$

Sabemos que para que se pueda multiplicar A por B es preciso que el número de columnas de A sea igual al número de filas de B . Así ocurre aquí pues dicho número es $n = \sum_{j=1}^t n_j$ en ambos casos. Para que pueda efectuarse la multiplicación de A por B “por bloques”, es necesario además que el número de columnas de bloques de A sea igual al número de filas de bloques de B ; en este caso, este número es t . Asimismo, las dimensiones verticales n_1, n_2, \dots, n_t de A deben ser iguales a las dimensiones horizontales de B .

Para cada $i = 1, 2, \dots, s$ y cada $j = 1, 2, \dots, u$, definimos

$$C_{ij} := \sum_{k=1}^t A_{ik} B_{kj}.$$

Si C es la matriz de bloques

$$C := \begin{matrix} & \begin{matrix} p_1 & p_2 & \cdots & p_u \end{matrix} \\ \begin{matrix} m_1 \\ m_2 \\ \vdots \\ m_s \end{matrix} & \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1u} \\ C_{21} & C_{22} & \cdots & C_{2u} \\ \vdots & \vdots & & \vdots \\ C_{s1} & C_{s2} & \cdots & C_{su} \end{pmatrix} \end{matrix},$$

entonces puede demostrarse (sic) que

$$C = AB,$$

donde el producto AB se efectúa mirando a A y B como matrices de números.

m_k , $k = 1, \dots, s$, y sustituyendo el bloque diagonal I_{m_i} por la matriz Q . Esto es,

$$H_{(m_1, \dots, m_s)}^{(i)}(Q) := \begin{pmatrix} I_{m_1} & & & & \\ & \ddots & & & \\ & & Q & & \\ & & & \ddots & \\ & & & & I_{m_s} \end{pmatrix}.$$

La transformación (5) equivale a premultiplicar la matriz A por una matriz $E_{(m_1, \dots, m_s)}^{(i,j)}(C)$ obtenida partiendo la matriz identidad de orden m en bloques con bloques diagonales $m_k \times m_k$, $k = 1, \dots, s$, y en la que el bloque 0 que ocupa la posición (i, j) se ha sustituido por la matriz C . Es decir, suponiendo que $i < j$,

$$E_{(m_1, \dots, m_s)}^{(i,j)}(C) := \begin{pmatrix} I_{m_1} & & & & \\ & \ddots & & & \\ & & I_{m_i} & \cdots & C \\ & & & \ddots & \vdots \\ & & & & I_{m_j} \\ & & & & & \ddots \\ & & & & & & I_{m_s} \end{pmatrix}$$

con $C \in \mathbb{C}^{m_i \times m_j}$.

Análogamente para las transformaciones (2), (4) y (6), por columnas, el resultado de la manipulación sobre A equivale a postmultiplicar A por una de las matrices de los tipos

$$Q_{(n_1, \dots, n_t)}^{(i,j)} := \begin{pmatrix} I_{n_1} & & & & \\ & \ddots & & & \\ & & 0 & \cdots & I_{n_i} \\ & & \vdots & & \vdots \\ & & I_{n_j} & \cdots & 0 \\ & & & & & \ddots \\ & & & & & & I_{n_t} \end{pmatrix},$$

$$H_{(n_1, \dots, n_t)}^{(i)}(Q) := \begin{pmatrix} I_{n_1} & & & & \\ & \ddots & & & \\ & & Q & & \\ & & & \ddots & \\ & & & & I_{n_t} \end{pmatrix}$$

y

$$F_{(n_1, \dots, n_t)}^{(i,j)}(C) := \begin{pmatrix} I_{n_1} & & & & \\ & \ddots & & & \\ & & I_{n_i} & & \\ & & \vdots & \ddots & \\ & & C & \cdots & I_{n_j} \\ & & & & & \ddots \\ & & & & & & I_{n_t} \end{pmatrix}$$

con $C \in \mathbb{C}^{n_j \times n_i}$, respectivamente.

Llamaremos a estas matrices

$$P_{(m_1, \dots, m_s)}^{(i,j)}, \quad H_{(m_1, \dots, m_s)}^{(i)}(Q), \quad E_{(m_1, \dots, m_s)}^{(i,j)}(C),$$

$$Q_{(n_1, \dots, n_t)}^{(i,j)}, \quad H_{(n_1, \dots, n_t)}^{(i)}(Q) \quad \text{y} \quad F_{(n_1, \dots, n_t)}^{(i,j)}(C),$$

matrices elementales por bloques. Es evidente que los determinantes de estas matrices son iguales a ± 1 ; excepto los de $H_{(m_1, \dots, m_s)}^{(i)}(Q)$ y $H_{(n_1, \dots, n_t)}^{(i)}(Q)$, que son iguales a $\det Q$. Por tanto al realizar transformaciones elementales por bloques sobre bloques fila (o columna) de una matriz A se obtienen matrices B equivalentes a ella; es decir, que A y B tienen el mismo rango.

Apéndice B

Digrafos y permutaciones

Vamos a dar en este apéndice las primeras definiciones sobre digrafos y la interpretación del determinante de una matriz cuadrada en términos de los ciclos de su digrafo ponderado (Cauchy, Coates). Para ello es preciso utilizar la descomposición de una permutación en producto de permutaciones-ciclo; lo que también revisamos para facilidad del lector.

B.1. Digrafos

La terminología sobre grafos no está estandarizada y varía más de lo que sería de desear de unos autores a otros. Hasta donde llega nuestra experiencia sobre este tema, nos induce a leer con cuidado las definiciones de cada autor. En la literatura, muchos conceptos no son siquiera definidos y se debe apelar a su significado intuitivo. Nuestras definiciones persiguen la claridad formal, conscientes del sacrificio intuitivo del tema; éste aspecto lo fiamos a las figuras que insertamos.

Un *digrafo* (*grafo dirigido u orientado*) D es un par (V, E) donde V es un conjunto finito de elementos llamados *vértices* (*puntos o nodos*) y E es un subconjunto de $V \times V$. Todo par ordenado $(a, b) \in E$ se llama *arco* del digrafo D . Nótese que un digrafo puede contener tanto los arcos (a, b) y (b, a) con a y b vértices distintos como arcos de la forma (a, a) ; éstos últimos arcos se denominan *lazos*. Los vértices a y b de un arco $\alpha = (a, b)$ se llaman extremos del arco α . Además, a es el *vértice inicial u origen* y b es el *vértice final* de α . El número de arcos que salen de un vértice es el *grado saliente* del vértice. El número de arcos que entran en un vértice es el *grado entrante* del vértice. Convenimos que un lazo en un vértice aporta una unidad al grado saliente y también una unidad al grado entrante. Se dice que el arco (a, b) es *incidente* con los vértices a y b .

Un *subdigrafo* del digrafo $D = (V, E)$ es un digrafo $D_1 = (V_1, E_1)$ tal que $V_1 \subset V$ y $E_1 \subset E$. Se dice que el subdigrafo D_1 es un *subdigrafo generador* de D si D_1 tiene el mismo conjunto de vértices que D , es decir, si $V_1 = V$.

Un *paseo* de D es una secuencia finita de vértices $w = (v_0, v_1, \dots, v_p)$ tal que $(v_{i-1}, v_i) \in E$ para $i = 1, 2, \dots, p$ y $p > 0$. El número p es la *longitud* del paseo w . Este número coincide con el número de arcos

$$(v_0, v_1), (v_1, v_2), \dots, (v_{p-1}, v_p)$$

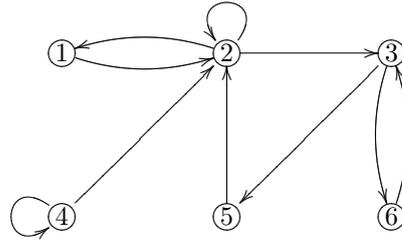
del paseo w ; contando los arcos cuantas veces se pase por ellos para ir desde el vértice v_0 al vértice v_p .

Ejemplo B.1.1. Consideremos el digrafo D representado por la figura B.1.

Aquí $D = (V, E)$ con $V = \{1, 2, 3, 4, 5, 6\}$ y

$$E = \{(1, 2), (2, 1), (2, 2), (2, 3), (3, 5), (3, 6), (4, 2), (4, 4), (5, 2), (6, 3)\}.$$

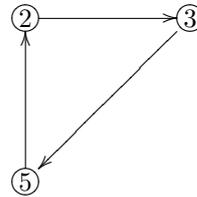
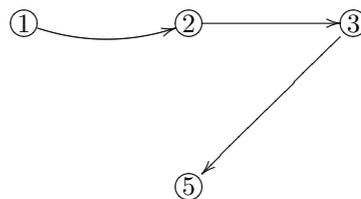
El paseo $w_1 = (1, 2, 3, 6)$ tiene longitud 3 y $w_2 = (1, 2, 2, 1, 2)$ es un paseo de longitud 4.

Figura B.1: Digrafo D .

Un *camino* de D es un paseo $w = (v_0, v_1, \dots, v_p)$ que tiene todos los vértices distintos. Un *ciclo* de D es un paseo $w = (v_0, v_1, \dots, v_p)$ con los vértices v_0, v_1, \dots, v_{p-1} distintos y $v_0 = v_p$. En particular, los lazos (a, a) son ciclos. Tenemos la clasificación siguiente de los paseos:

$$\text{paseos} \begin{cases} \text{caminos,} \\ \text{ciclos,} \\ \text{paseos } w = (v_0, v_1, \dots, v_p) \\ \text{con repeticiones en } v_0, v_1, \dots, v_{p-1} \text{ y/o en } v_1, \dots, v_{p-1}, v_p. \end{cases}$$

Un digrafo se llama *acíclico* si no tiene ciclos. En el digrafo anterior el paseo $C = (2, 3, 5, 2)$ es un ciclo (figura B.2) y el paseo $P = (1, 2, 3, 5)$ es un camino (figura B.3). El número de vértices distintos de un camino P o de un ciclo C será denotado por $|P|$ o $|C|$, respectivamente. Así, en el ejemplo último, $|P| = 4$ y $|C| = 3$. También $|P| = \text{longitud de } P + 1$, $|C| = \text{longitud de } C$. La notación $|X|$ indica el cardinal de un conjunto finito X . Aquí se comete un ligero abuso de notación pues P y C no son conjuntos.

Figura B.2: Ciclo $C = (2, 3, 5, 2)$.Figura B.3: Camino $P = (1, 2, 3, 5)$.

Una *familia de ciclos* de un digrafo D es un subdigrafo h de D que en cada uno de sus vértices el grado saliente y el grado entrante son iguales a uno. Otra forma más intuitiva de expresar esto es decir que una familia de ciclos de D es un subdigrafo h de D que es “unión” de ciclos “disjuntos” dos a dos. Dos ciclos

$$c_1 = (u_0, u_1, \dots, u_{p-1}, u_p) \quad \text{y} \quad c_2 = (v_0, v_1, \dots, v_{q-1}, v_q)$$

se llaman “disjuntos” si sus conjuntos de vértices respectivos $\{u_0, u_1, \dots, u_{p-1}\}$ y $\{v_0, v_1, \dots, v_{q-1}\}$ son disjuntos. La figura B.4 muestra una familia de ciclos del digrafo de la figura B.1.

La *longitud* de una familia de ciclos h es el número de vértices de h ; es, por tanto, igual a la suma de las longitudes de los ciclos que la componen.

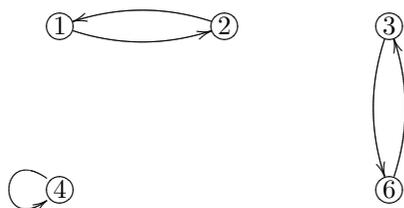


Figura B.4: Familia de ciclos de D .

Una familia de ciclos generadora h del digrafo $D = (V, E)$ es una familia de ciclos h de D que es un subdigrafo generador de D . Es decir, tal que h y D tienen el mismo conjunto de vértices: V . La longitud de una familia de ciclos generadora es igual a $|V|$. Otra manera de definir las familias generadoras de ciclos de D es la siguiente:

Una familia de ciclos generadora h del digrafo D es un subdigrafo generador de D que es “unión” de ciclos “disjuntos” dos a dos (los lazos (a, a) son ciclos). Una familia de ciclos generadora de D se llama también un 1-factor de D . El digrafo de la figura B.1 no tiene familias de ciclos generadoras. Añadiendo el arco $(5, 6)$ a este digrafo obtenemos el digrafo \mathcal{D} (figura B.5)

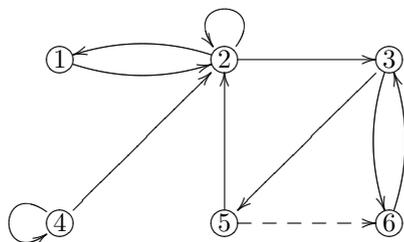


Figura B.5: Digrafo \mathcal{D} .

Una familia de ciclos generadora de \mathcal{D} viene dada en la figura B.6.

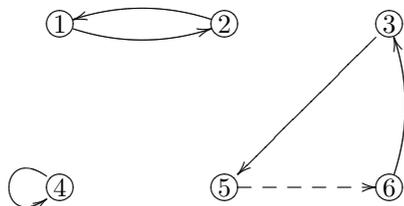


Figura B.6: Familia de ciclos generadora de \mathcal{D} .

Un digrafo ponderado es una terna (V, E, p) en la que $D = (V, E)$ es un digrafo y $p : E \rightarrow \mathcal{R}$ es una aplicación de E en un anillo conmutativo \mathcal{R} , que asocia a cada arco $(\alpha, \beta) \in E$ su peso $p((\alpha, \beta))$.

Se llama peso del digrafo ponderado D , a

$$f(D) := \prod_{(\alpha, \beta) \in E} p((\alpha, \beta)).$$

Necesitaremos también definir la noción de subdigrafo de un digrafo ponderado. Se llama subdigrafo ponderado o subdigrafo, a secas, de un digrafo ponderado $G = (V, E, p)$ a todo digrafo $G_1 = (V_1, E_1, p_1)$ tal que $V_1 \subset V, E_1 \subset E$ y $p_1 = p|_{E_1}$ (restricción de p a E_1).

B.2. Permutaciones

Definición B.2.1. Sea E un conjunto finito de n elementos

$$E = \{a_1, a_2, \dots, a_n\}.$$

Llamaremos *permutación* de E a cualquier aplicación inyectiva

$$\sigma : E \rightarrow E.$$

Como el subconjunto

$$\{\sigma(a_1), \sigma(a_2), \dots, \sigma(a_n)\}$$

de E tiene n elementos, se sigue que σ es suprayectiva y, por lo tanto, biyectiva. Es sabido que el conjunto $S(E)$ de permutaciones de E es un grupo con la operación

$$\sigma\sigma' := \sigma \circ \sigma', \quad \forall \sigma, \sigma' \in S(E)$$

de composición de aplicaciones de E en E ; es decir, que $\sigma\sigma'$ es la aplicación de E en E que envía $x \in E$ a $\sigma(\sigma'(x))$.

Por definición, para cualquier permutación σ , tenemos

$$\sigma\sigma^{-1} = id \text{ y } \sigma^{-1}\sigma = id,$$

donde id es la permutación identidad, esto es, la permutación tal que $id(x) = x$ para todo $x \in E$.

Si $\sigma_1\sigma_2$ son permutaciones de E , entonces la inversa de la permutación

$$\sigma_1\sigma_2$$

es la permutación

$$\sigma_2^{-1}\sigma_1^{-1}.$$

Esto se ve trivialmente por cálculo directo.

Es conocido que el orden del grupo $S(E)$, o cardinal de $S(E)$, es igual a $n!$

Muchas veces una permutación σ de E se denota por

$$\sigma = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ \sigma_{a_1} & \sigma_{a_2} & \cdots & \sigma_{a_n} \end{pmatrix}$$

donde $\sigma_{a_i} := \sigma(a_i)$, $i = 1, 2, \dots, n$. Otra notación utilizada para la permutación σ es

$$\sigma = (\sigma_{a_1}, \sigma_{a_2}, \dots, \sigma_{a_n}).$$

Por ejemplo, la permutación σ de $\{1, 2, 3, 4\}$, dada por

$$\sigma(1) = 3, \sigma(2) = 2, \sigma(3) = 1, \sigma(4) = 4,$$

puede escribirse como

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} \text{ ó } \begin{pmatrix} 2 & 3 & 1 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \text{ ó } (3, 2, 1, 4).$$

Si el cardinal de E , $|E|$, es mayor o igual que 3, entonces el grupo $S(E)$ no es conmutativo; pues supongamos que $p, q \in S(E)$ dejan invariantes todos los elementos de $E \setminus \{a_1, a_2, a_3\}$ y que

$$p = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & \cdots & a_n \\ a_2 & a_3 & a_1 & a_4 & \cdots & a_n \end{pmatrix}, \quad q = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & \cdots & a_n \\ a_2 & a_1 & a_3 & a_4 & \cdots & a_n \end{pmatrix}.$$

Entonces

$$pq = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & \cdots & a_n \\ a_3 & a_2 & a_1 & a_4 & \cdots & a_n \end{pmatrix}$$

$$qp = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & \cdots & a_n \\ a_1 & a_3 & a_2 & a_4 & \cdots & a_n \end{pmatrix},$$

luego $pq \neq qp$.

Definición B.2.2. Una *transposición* de E es una permutación τ de E que intercambia dos elementos de E y deja fijos los demás. Es decir, que existen $a, b \in E$, a, b distintos, tales que

$$\tau(a) = b, \tau(b) = a, \text{ y } \tau(x) = x \text{ para todo } x \in E \setminus \{a, b\}.$$

Llamaremos $J_n := \{1, 2, \dots, n\}$ y S_n al grupo de las permutaciones de J_n .

Definición B.2.3. Diremos que una permutación π de E opera sobre un subconjunto P de E si deja fijo cada elemento de $E \setminus P$; es decir, si

$$\text{para cada } x \in E \setminus P, \quad \pi(x) = x;$$

lo que implica que para todo $y \in P$, $\pi(y) \in P$, ya que π es inyectiva. Obsérvese que esto no implica que para cada $y \in P$, se tenga que $\pi(y) \neq y$.

Proposición B.2.1. Si las permutaciones π_1 y π_2 operan respectivamente sobre los subconjuntos disjuntos P_1 y P_2 de E , entonces

$$\pi_1\pi_2 = \pi_2\pi_1.$$

DEMOSTRACIÓN. Denotemos por X^c el conjunto complementario de un subconjunto X de E , esto es, $X^c := E \setminus X$. Se tiene que

$$E = P_1 \cup P_2 \cup (P_1 \cup P_2)^c.$$

Dado cualquier $x \in E$, o bien $x \in P_1$, ó $x \in P_2$, ó $x \in (P_1 \cup P_2)^c$, y sólo puede darse para cada x una de estas tres posibilidades.

Si $x \in P_1$, entonces $\pi_2(x) = x$; por lo tanto $\pi_1\pi_2(x) = \pi_1(x)$; además, $\pi_2\pi_1(x) = \pi_1(x)$ pues $\pi_1(x) \in P_1$. Luego

$$\pi_1\pi_2(x) = \pi_2\pi_1(x).$$

Si $x \in P_2$, entonces $\pi_1\pi_2(x) = \pi_2(x)$ pues $\pi_2(x) \in P_2$; por otra parte, $\pi_2\pi_1(x) = \pi_2(x)$ pues $\pi_1(x) = x$ ya que $x \notin P_1$. Luego

$$\pi_1\pi_2(x) = \pi_2\pi_1(x).$$

Si $x \in (P_1 \cup P_2)^c$, entonces $x \notin P_1$ y $x \notin P_2$. De $x \notin P_2$ se sigue que $\pi_1\pi_2(x) = \pi_1(x)$ y de $x \notin P_1$ se deduce que $\pi_1(x) = x$. Por estas mismas razones, $\pi_2\pi_1(x) = \pi_2(x) = x$. Luego

$$\pi_1\pi_2(x) = \pi_2\pi_1(x).$$

Así pues, queda probado que

$$\pi_1\pi_2 = \pi_2\pi_1.$$

□

En general, si π_i opera sobre P_i , donde $P_i \subset E$, para cada $i = 1, \dots, k$ y para todos $i, j, i \neq j$, $P_i \cap P_j = \emptyset$, entonces

$$\pi_1\pi_2 \cdots \pi_k = \pi_{i_1}\pi_{i_2} \cdots \pi_{i_k},$$

siendo

$$\begin{pmatrix} 1 & 2 & \cdots & k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}$$

una permutación cualquiera de $\{1, 2, \dots, k\}$.

Si $\alpha : J_n \rightarrow J_n$ es cualquier aplicación de J_n en sí mismo, definimos

$$W(\alpha) := \prod_{1 \leq i < j \leq n} [\alpha(j) - \alpha(i)].$$

Sea σ una permutación de J_n ; como σ es biyectiva

$$i < j \Rightarrow \text{ó } \sigma(i) < \sigma(j), \text{ ó } \sigma(i) > \sigma(j);$$

en el segundo caso diremos que los elementos $\sigma(i)$ y $\sigma(j)$ presentan una *inversión*. Denotemos por $I(\sigma)$ el conjunto

$$\{(i, j) \in J_n \times J_n \mid i < j, \sigma(i) > \sigma(j)\},$$

y consideremos el producto

$$V_n := \prod_{1 \leq i < j \leq n} (j - i);$$

es decir, $V_n = W(id)$, siendo id la aplicación identidad de J_n . Este producto también se puede expresar con más detalle de esta manera

$$V_n = [2 - 1][(3 - 1)(3 - 2)] \cdots [(n - 1)(n - 2) \cdots (n - (n - 1))].$$

El número V_n es un entero positivo.

Consideremos

$$W(\sigma) := \prod_{1 \leq i < j \leq n} [\sigma(j) - \sigma(i)].$$

Como σ es una biyección, cada factor de V_n se encuentra presente en el producto $W(\sigma)$ una vez y sólo una vez, salvo quizás el signo, y se ve que

$$W(\sigma) = (-1)^{|I(\sigma)|} V_n;$$

se define una aplicación ε de S_n en $\{1, -1\}$

$$\varepsilon(\sigma) := (-1)^{|I(\sigma)|}$$

y se llama a $\varepsilon(\sigma)$ la *signatura de la permutación* σ . Es decir, la signatura de σ es igual a (-1) elevado al número de inversiones de σ . Si $\varepsilon(\sigma) = 1$ se dice que σ es una *permutación par*. Si $\varepsilon(\sigma) = -1$ se dice que σ es una *permutación impar*. La signatura de la identidad id es igual a 1.

De la definición de $\varepsilon(\sigma)$ se deduce que

$$W(\sigma) = \varepsilon(\sigma)V_n.$$

Consideremos a continuación dos permutaciones σ_1 y σ_2 de J_n ; busquemos la signatura del producto $\sigma_1\sigma_2$. Tenemos que

$$\begin{aligned} W(\sigma_1\sigma_2) &= \prod_{1 \leq i < j \leq n} [\sigma_1(\sigma_2(j)) - \sigma_1(\sigma_2(i))] \\ &= \varepsilon(\sigma_1) \prod_{1 \leq i < j \leq n} [\sigma_2(j) - \sigma_2(i)] = \varepsilon(\sigma_1)\varepsilon(\sigma_2)V_n. \end{aligned}$$

Pero

$$W(\sigma_1\sigma_2) = \varepsilon(\sigma_1\sigma_2)V_n.$$

Por consiguiente,

$$\varepsilon(\sigma_1\sigma_2) = \varepsilon(\sigma_1)\varepsilon(\sigma_2). \quad (\text{B.1})$$

Veamos cuál es la *signatura* de cualquier *transposición* τ de J_n . Para ello, trataremos de caracterizar los elementos del conjunto de inversiones de τ , $I(\tau)$. Supongamos que $\alpha, \beta \in J_n$, $\alpha < \beta$ y que τ intercambia los números α y β y deja fijos los restantes elementos de J_n . Representemos esta situación así

$$\tau = \begin{pmatrix} 1 & \cdots & \alpha & \cdots & \beta & \cdots & n \\ 1 & \cdots & \beta & \cdots & \alpha & \cdots & n \end{pmatrix},$$

o más detalladamente

$$\tau = \begin{pmatrix} 1 & \cdots & \alpha - 1 & \boxed{\alpha} & \alpha + 1 \cdots & \beta - 1 & \boxed{\beta} & \cdots & n \\ 1 & \cdots & \alpha - 1 & \beta & \alpha + 1 \cdots & \beta - 1 & \alpha & \cdots & n \end{pmatrix}.$$

Los elementos de $I(\tau)$ son los pares siguientes

$$(\alpha, \alpha + 1), (\alpha, \alpha + 2), \dots, (\alpha, \beta - 1), (\alpha, \beta),$$

y

$$(\alpha + 1, \beta), (\alpha + 2, \beta), \dots, (\beta - 1, \beta).$$

En la primera de estas dos líneas hay $\beta - \alpha$ pares y en la segunda hay $\beta - 1 - \alpha$ pares, luego

$$|I(\tau)| = 2(\beta - \alpha) - 1,$$

que es un número impar. En consecuencia,

$$\varepsilon(\tau) = -1.$$

Sea E un conjunto cualquiera de n elementos. Una permutación c de E se llamará *permutación-ciclo* o *permutación cíclica* si existen un entero p ($1 \leq p \leq n$) y p elementos $a_1, a_2, \dots, a_{p-1}, a_p$ de E tales que

$$c(a_1) = a_2, c(a_2) = a_3, \dots, c(a_{p-1}) = a_p,$$

$$c(a_p) = a_1$$

y c deja fijos los elementos de $E \setminus \{a_1, a_2, \dots, a_{p-1}, a_p\}$. Llamando a_{p+1}, \dots, a_n a los elementos de este último conjunto, se puede denotar a la permutación-ciclo c por

$$c = \begin{pmatrix} a_1 & a_2 & \cdots & a_{p-1} & a_p & a_{p+1} & \cdots & a_n \\ a_2 & a_3 & \cdots & a_p & a_1 & a_{p+1} & \cdots & a_n \end{pmatrix}$$

o más simplemente

$$c = (a_1, a_2, \dots, a_p);$$

esta última notación nos indica que cada elemento de E que no figura en ella es dejado fijo por c , cada uno de los elementos que figuran, salvo a_p , tiene por imagen mediante c al siguiente y a_1 es la imagen de a_p . Vemos que

$$c^2 = \begin{pmatrix} a_1 & a_2 & \cdots & a_{p-1} & a_p & a_{p+1} & \cdots & a_n \\ a_3 & a_4 & \cdots & a_1 & a_2 & a_{p+1} & \cdots & a_n \end{pmatrix}$$

y para $1 < k < p$,

$$c^k = \begin{pmatrix} a_1 & a_2 & \cdots & a_{p-1} & a_p & a_{p+1} & \cdots & a_n \\ a_{k+1} & a_{k+2} & \cdots & a_{k-1} & a_k & a_{p+1} & \cdots & a_n \end{pmatrix}$$

y

$$c^p = \begin{pmatrix} a_1 & a_2 & \cdots & a_{p-1} & a_p & a_{p+1} & \cdots & a_n \\ a_1 & a_2 & \cdots & a_{p-1} & a_p & a_{p+1} & \cdots & a_n \end{pmatrix} = id,$$

donde id designa la *permutación idéntica* o *identidad* de E . Por consiguiente, si $p \geq 2$, $c^p = id$ y p es el menor entero $h > 0$ tal que $c^h = id$. Por lo cual, la *permutación-ciclo* $c = (a_1, a_2, \dots, a_p)$ que opera sobre p elementos de un conjunto E es un elemento de orden p del grupo de permutaciones $S(E)$.

Se llama *longitud* de la permutación-ciclo

$$c = (a_1, a_2, \dots, a_p)$$

al número p . Así pues, el orden de c es igual a la longitud de c . Una transposición τ de E que intercambia los elementos distintos a, b de E es igual al ciclo (a, b) , de orden 2. La permutación idéntica id es una permutación-ciclo de orden 1 (por otra parte, id es la única permutación-ciclo de orden 1). Es posible escribir

$$id = (a)$$

para todo $a \in E$; es decir para cada $a \in E$, id es la permutación-ciclo que opera sobre $\{a\}$.

De acuerdo con la Proposición B.2.1, dos permutaciones-ciclo c_1 y c_2 , que operen sobre dos subconjuntos disjuntos de E , conmutan:

$$c_1 c_2 = c_2 c_1.$$

Descomposición de una permutación en producto de permutaciones-ciclo

Sea σ una permutación de un conjunto E de n elementos. Si $\sigma \neq id$, existe algún elemento $a_1 \in E$ tal que $\sigma(a_1) \neq a_1$; llamemos $a_2 := \sigma(a_1)$, $a_3 := \sigma(a_2)$ y sea ℓ_1 el primer índice tal que

$$\sigma(a_{\ell_1}) \in \{a_1, a_2, \dots, a_{\ell_1}\}.$$

Entonces se tiene que $\sigma(a_{\ell_1}) = a_1$; en efecto, si fuera $\sigma(a_{\ell_1}) = a_i$, con $1 < i \leq \ell_1$, se tendría que

$$\sigma(a_{\ell_1}) = \sigma(a_{i-1}),$$

lo que es imposible pues a_{ℓ_1} y a_{i-1} son distintos y σ es una aplicación inyectiva. Por consiguiente, los valores de la permutación-ciclo $c_1 = (a_1, a_2, \dots, a_{\ell_1})$ y de σ coinciden en cada elemento del subconjunto $\{a_1, a_2, \dots, a_{\ell_1}\}$ de E .

Designemos por a_{ℓ_1+1} a un elemento de

$$E \setminus \{a_1, a_2, \dots, a_{\ell_1}\};$$

hallando los transformados sucesivos $a_{\ell_1+2} := \sigma(a_{\ell_1+1})$, $a_{\ell_1+3} := \sigma(a_{\ell_1+2}), \dots$ obtenemos una nueva permutación-ciclo

$$c_2 = (a_{\ell_1+1}, a_{\ell_1+2}, \dots, a_{\ell_1+\ell_2}).$$

Como E es un conjunto finito, al cabo de un número finito de repeticiones de este proceso habremos agotado E . De esta manera habremos definido una partición

$$P_1 = \{a_1, a_2, \dots, a_{\ell_1}\}, P_2 = \{a_{\ell_1+1}, a_{\ell_1+2}, \dots, a_{\ell_1+\ell_2}\}, \dots, \\ P_k = \{a_{\ell_1+\dots+\ell_{k-1}+1}, \dots, a_{\ell_1+\dots+\ell_{k-1}+\ell_k}\}$$

de E tal que los valores de σ y de la permutación-ciclo que opera sobre P_j coinciden en cada elemento de P_j ($j = 1, \dots, k$). Hemos probado, pues, la proposición siguiente.

Proposición B.2.2. *Toda permutación de un conjunto de n elementos es descomponible en producto de permutaciones-ciclo cuya suma de órdenes es n . Esta factorización es única salvo el orden de las permutaciones-ciclo, que conmutan a pares.*

Ejemplo B.2.1.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 5 & 4 & 6 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 5 & 4 & 6 & 3 \end{pmatrix} \\ = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix} \\ = (2, 1)(5, 6, 3)(4) = (2, 1)(5, 6, 3),$$

pues $(4) = id$. Los conjuntos $P_1 = \{2, 1\}$, $P_2 = \{5, 6, 3\}$, $P_3 = \{4\}$ constituyen una partición de $\{1, 2, 3, 4, 5, 6\}$. Los órdenes de las permutaciones-ciclo $(2, 1)$, $(5, 6, 3)$ y (4) son 2, 3 y 1, que suman 6.

Observación B.2.1. Toda permutación-ciclo de orden 1 es idéntica a *id*. Pero en la descomposición anterior es preciso contar *todas* las permutaciones-ciclo de orden 1 si se quiere que la suma de órdenes de las permutaciones-ciclo sea igual a n .

Por otro lado, toda permutación-ciclo es descomponible en producto de transposiciones; en efecto:

$$(a_1, a_2, \dots, a_p) = (a_{p-1}, a_p) \cdots (a_2, a_p)(a_1, a_p). \tag{B.2}$$

(Atención: Es preciso comenzar a aplicar los productos por la derecha).

Ejemplo B.2.2.

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix};$$

es decir,

$$(2, 3, 4, 1) = (4, 1)(3, 1)(2, 1).$$

Pero en este caso, dos cualesquiera de las transposiciones

$$(a_{p-1}, a_p), \dots, (a_2, a_p), (a_1, a_p)$$

no conmutan en general. La Proposición B.2.2 y (B.2) nos permiten deducir la proposición que sigue.

Proposición B.2.3. *Toda permutación de un conjunto finito es descomponible en producto de transposiciones.*

Pero, en general, esta descomposición no es única.

¿Cuál es la signatura de una permutación-ciclo? De las fórmulas (B.1) y (B.2) deducimos que si c es la permutación-ciclo $c = (a_1, a_2, \dots, a_p)$, entonces

$$\varepsilon(c) = (-1)^{p-1}, \tag{B.3}$$

pues la signatura de cada transposición es -1 .

Digrafo permutación

Sea σ una permutación de $\{1, 2, \dots, n\}$. El *digrafo de la permutación* σ es el digrafo $D(\sigma)$ cuyo conjunto de vértices es $\{1, 2, \dots, n\}$ y cuyo conjunto de arcos es $\{(i, \sigma(i)) \mid i = 1, \dots, n\}$. Los ciclos de $D(\sigma)$ corresponden a las permutaciones-ciclo en las que se factoriza σ ; en particular, los lazos (i, i) corresponden a los puntos fijos i de la permutación σ . Así, a la permutación σ del Ejemplo B.2.1 le corresponde el digrafo $D(\sigma)$ igual al digrafo de la figura B.6 de la página 161.

Observamos que un digrafo D cuyo conjunto de vértices es $\{1, \dots, n\}$ es el digrafo de alguna permutación $\sigma \in S_n$ si y sólo si D es una familia de ciclos generadora de D .

En general, si $D = (V, E)$ es un digrafo con $V = \{1, 2, \dots, n\}$, toda familia de ciclos $h = (V_1, E_1)$ de D se puede poner en correspondencia biunívoca con el digrafo de la permutación σ_1 de V_1 dada por el producto de las permutaciones-ciclo asociadas a los ciclos de h . La longitud de h es igual al cardinal de V_1 . Por ejemplo, la familia de ciclos de la figura B.4 de la página 161 se corresponde con el digrafo de la permutación σ_1 de $\{1, 2, 3, 4, 6\}$ dada por

$$\sigma_1 = (1, 2)(3, 6)(4).$$

Matrices de permutación

Una matriz $n \times n$ $P = (p_{ij})$ se llama *matriz de permutación* si sus elementos son todos cero excepto los de una diagonal

$$(p_{\sigma(1),1}, p_{\sigma(2),2}, \dots, p_{\sigma(n),n})$$

que son todos iguales a uno. Es decir, que P tiene todos sus elementos iguales a cero excepto un elemento en cada fila y columna igual a uno. Por ejemplo, la matriz

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

es una matriz de permutación de orden 4. Es claro que hay tantas matrices de permutación de orden n como permutaciones σ de $\{1, \dots, n\}$, es decir, que hay $n!$ matrices de permutación. Es fácil ver que la inversa de una matriz de permutación P es su matriz transpuesta P^T . Por lo tanto, si sobre una matriz cuadrada $A \in \mathbb{C}^{n \times n}$ permutamos sus filas y a continuación permutamos de igual manera sus columnas, se obtiene una matriz \tilde{A} que es semejante a A .

B.3. Digrafos y determinantes

Hay cuatro maneras de asociar un digrafo ponderado a una matriz cuadrada $A = (a_{ij}) \in \mathcal{R}^{n \times n}$, siendo \mathcal{R} un anillo conmutativo. Véase [52, págs.241–246]. De estas cuatro maneras elegimos la de Coates (1959). Aunque se atribuye a este autor, esta manera se puede rastrear hasta Cuachy (1812 y 1815) y Jacobi (1841). El *digrafo ponderado de Coates* $G(A)$ de A se define así: el conjunto de vértices es $\{1, 2, \dots, n\}$, el par (j, i) es un arco de $G(A)$ si $a_{ij} \neq 0$ (¡nótese la inversión de los subíndices!), y el peso del arco (j, i) es el elemento a_{ij} de \mathcal{R} .

Ejemplo B.3.1. Supongamos que todos los elementos a_{ij} que aparecen en la siguiente matriz son distintos de 0.

$$A = \begin{pmatrix} a_{11} & a_{12} & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & a_{24} & a_{25} & 0 \\ 0 & a_{32} & 0 & 0 & 0 & a_{36} \\ a_{41} & 0 & 0 & a_{44} & 0 & 0 \\ 0 & 0 & a_{53} & 0 & 0 & 0 \\ 0 & 0 & a_{63} & 0 & a_{65} & a_{66} \end{pmatrix}.$$

Entonces el digrafo ponderado de Coates de A viene dado por la figura B.7.

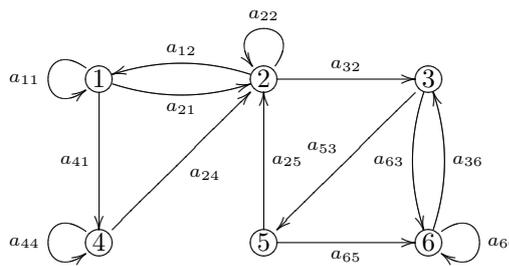


Figura B.7: Digrafo ponderado de Coates de A .

Este digrafo tiene cuatro familias de ciclos generadoras h_i asociadas respectivamente a las permutaciones σ_i de $\{1, 2, 3, 4, 5, 6\}$, $i = 1, 2, 3, 4$. Detallamos estas familias en las cuatro figuras siguientes B.8, B.9, B.10, B.11.

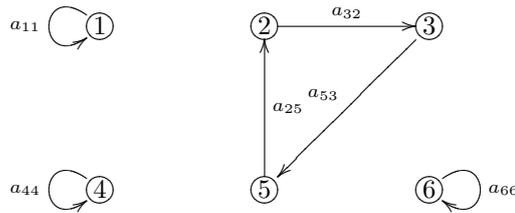


Figura B.8: Familia de ciclos generadora h_1 .

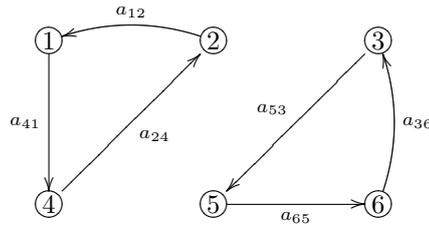


Figura B.9: Familia de ciclos generadora h_2 .

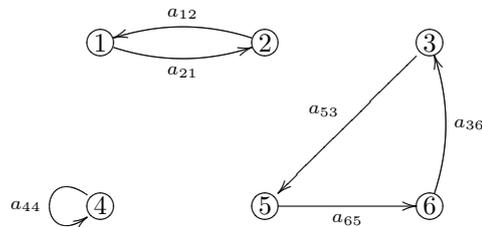


Figura B.10: Familia de ciclos generadora h_3 .

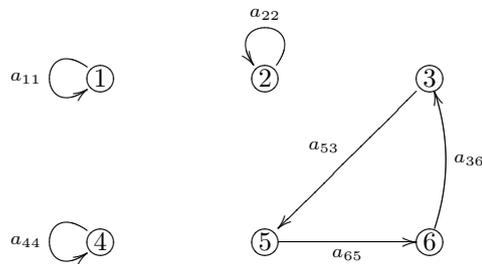


Figura B.11: Familia de ciclos generadora h_4 .

La tabla B.1 nos indica las permutaciones a las que están asociadas estas cuatro familias de ciclos generadoras

Familia	Permutación
h_1	$\sigma_1 = (1)(2, 3, 5)(4)(6)$
h_2	$\sigma_2 = (1, 4, 2)(3, 5, 6)$
h_3	$\sigma_3 = (1, 2)(3, 5, 6)(4)$
h_4	$\sigma_4 = (1)(2)(3, 5, 6)(4)$

Cuadro B.1: Familias de ciclos y sus permutaciones.

Sea $Y = (y_{ij})$ una matriz cuadrada de orden n con elementos en un anillo conmutativo cualquiera \mathcal{R} . Se llama *diagonal* de Y a todo vector de \mathcal{R}^n de la forma

$$(y_{\sigma(1),1}, y_{\sigma(2),2}, \dots, y_{\sigma(n),n})$$

donde σ es una permutación de $\{1, 2, \dots, n\}$. Es decir, una diagonal de Y se obtiene eligiendo n elementos de la matriz de manera que no haya dos en la misma fila o columna. La *diagonal principal* de Y ,

$$(y_{11}, y_{22}, \dots, y_{nn})$$

corresponde a la permutación idéntica id .

Dada una matriz $A = (a_{ij}) \in \mathcal{R}^{n \times n}$, las diagonales de A formadas por n elementos distintos de cero se corresponden biunívocamente con las familias de ciclos generadoras del digrafo (ponderado) de Coates de A . Así pues, las diagonales con todas sus componentes no nulas de la matriz A del Ejemplo B.3.1 de la página 168 vienen dadas en la tabla B.2. En particular, la diagonal señalada con recuadros en la matriz A corresponde a la familia h_3 .

Familia	Permutación	Diagonal	Figura
h_1	$\sigma_1 = (1)(2, 3, 5)(4)(6)$	$(a_{11}, a_{32}, a_{53}, a_{44}, a_{25}, a_{66})$	B.8
h_2	$\sigma_2 = (1, 4, 2)(3, 5, 6)$	$(a_{41}, a_{12}, a_{53}, a_{24}, a_{65}, a_{36})$	B.9
h_3	$\sigma_3 = (1, 2)(3, 5, 6)(4)$	$(a_{21}, a_{12}, a_{53}, a_{44}, a_{65}, a_{36})$	B.10
h_4	$\sigma_4 = (1)(2)(3, 5, 6)(4)$	$(a_{11}, a_{22}, a_{53}, a_{44}, a_{65}, a_{36})$	B.11

Cuadro B.2: Diagonales de A y sus familias de ciclos.

Por lo tanto, el determinante de la matriz A del Ejemplo B.3.1 es igual a la suma de los cuatro productos

$$\varepsilon(\sigma_i) a_{\sigma_i(1),1} a_{\sigma_i(2),2} \cdots a_{\sigma_i(6),6}, \quad i = 1, 2, 3, 4,$$

asociados a las cuatro familias de ciclos generadoras; es decir, es igual a

$$\varepsilon(\sigma_i) f(h_i),$$

siendo $f(h_i)$ el peso del digrafo h_i , $i = 1, 2, 3, 4$. Éste es un hecho general, que pasamos a demostrar en el teorema que sigue. El cual nos da una interpretación combinatoria del determinante de una matriz en términos de los ciclos de su digrafo ponderado de Coates.

Teorema B.3.1. Sean \mathcal{R} un anillo conmutativo y $A \in \mathcal{R}^{n \times n}$. Sea $G(A)$ el digrafo ponderado de Coates de A . Entonces

$$\det A = \sum_{h \in H} (-1)^{n-d(h)} f(h),$$

donde H es el conjunto de las familias de ciclos generadoras de $G(A)$, $f(h)$ es el peso del subdigrafo h de $G(A)$ y $d(h)$ es el número de ciclos “disjuntos” de h (incluidos los lazos).

DEMOSTRACIÓN. Si $c = (i_1, i_2, \dots, i_p)$ es una permutación-ciclo de $\{1, \dots, n\}$, entonces $\varepsilon(c) = (-1)^{p-1}$, siendo $\varepsilon(\sigma)$ la signatura de la permutación σ . Véase (B.3) de la página 167. Por la definición de determinante

$$\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n},$$

donde S_n es el grupo de las permutaciones de $\{1, \dots, n\}$.

Sea $\sigma \in S_n$ y sea $\sigma = c_1 c_2 \cdots c_k$ la factorización canónica de la permutación σ en producto de permutaciones-ciclo, dada por la Proposición B.2.2 de la página 166. Supongamos que ℓ_i es la longitud (u orden) de c_i $i = 1, \dots, k$. En la factorización anterior, cada punto fijo j de σ aporta una permutación (j) de longitud 1; así, algunos de los ℓ_i pueden ser iguales a 1. Entonces,

$$\begin{aligned} \varepsilon(\sigma) &= \\ &= \varepsilon(c_1) \varepsilon(c_2) \cdots \varepsilon(c_k) = (-1)^{\ell_1-1} (-1)^{\ell_2-1} \cdots (-1)^{\ell_k-1} = \\ &= (-1)^{\ell_1+\ell_2+\cdots+\ell_k} (-1)^{-k} = (-1)^n (-1)^{-k} = (-1)^{n-k}. \end{aligned}$$

Por lo tanto,

$$\det A = \sum_{\sigma \in S_n} (-1)^{n-k} a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n} = \sum_{h \in H} (-1)^{n-d(h)} f(h);$$

pues si $h = (\{1, \dots, n\}, \{(i, \sigma(i)) : i = 1, \dots, n\})$, se sigue que

$$f(h) = a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n},$$

ya que

$$p((i, \sigma(i))) = a_{\sigma(i),i}, \quad i = 1, 2, \dots, n.$$

En principio, $|H| \leq |S_n| = n!$ En el Ejemplo B.3.1, $|H| = 4$, que es menor que $24=4!$ □

B.4. Ejercicios

Ejercicio B.1. Demostrar que si σ es una permutación de $1, 2, \dots, n$ y si $\varepsilon(\sigma)$ es su signatura, entonces

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \\ \sigma(1)^2 & \sigma(2)^2 & \cdots & \sigma(n)^2 \\ \vdots & \vdots & \cdots & \vdots \\ \sigma(1)^{n-1} & \sigma(2)^{n-1} & \cdots & \sigma(n)^{n-1} \end{vmatrix} = \varepsilon(\sigma) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & n \\ 1^2 & 2^2 & \cdots & n^2 \\ \vdots & \vdots & \cdots & \vdots \\ 1^{n-1} & 2^{n-1} & \cdots & n^{n-1} \end{vmatrix}.$$

Deducir de aquí otra demostración de que para todas $\sigma_1, \sigma_2 \in S_n$,

$$\varepsilon(\sigma_1 \sigma_2) = \varepsilon(\sigma_1) \varepsilon(\sigma_2).$$

Ejercicio B.2. ¿Son ciertas las proposiciones análogas a las Proposiciones B.2.1, B.2.2 y B.2.3 para las biyecciones de un conjunto infinito sobre sí mismo?

Apéndice C

Topología

Vamos a reseñar aquí ciertas nociones necesarias para seguir más fácilmente el texto.

C.1. Adherencia, interior, exterior y frontera

Sea E un espacio topológico y A un subconjunto de E . Se llama *interior* de A al mayor conjunto abierto contenido en A , o a la unión de todos los conjuntos abiertos contenidos en A , y se denota por $\overset{\circ}{A}$, o $\text{Int}(A)$. A los puntos de $\overset{\circ}{A}$ se les llama puntos interiores de A . Si A^c denota el conjunto complementario de A , respecto de E , al conjunto interior de A^c se llama *exterior* de A y se denota por $\text{Ext}(A)$; así pues,

$$\text{Ext}(A) := \overset{\circ}{A^c}.$$

Los puntos de $\text{Ext}(A)$ son los puntos exteriores a A . Se llama *adherencia o clausura* de A al mínimo conjunto cerrado que contiene a A , o a la intersección de todos los cerrados que contienen a A , y se denota por \overline{A} . Los puntos de \overline{A} se llaman puntos adherentes a A . Por lo tanto, en todo entorno de un punto adherente a A existen puntos de A . Un punto p se llama punto de acumulación de A si en todo entorno de p existen puntos de A , distintos de p . Así pues, un punto es adherente a A si y sólo si es punto de A o es punto de acumulación de A (o ambas cosas a la vez). El conjunto de puntos de acumulación de A se llama el conjunto *derivado* de A y se designa por A' . Un punto frontera de A es aquel que en todo entorno suyo hay puntos de A y puntos de A^c . Se llama *frontera* de A al conjunto de puntos frontera de A y lo denotaremos por $\text{Fr}(A)$ o $\partial(A)$; por consiguiente,

$$\text{Fr}(A) = \overline{A} \cap \overline{A^c}.$$

De donde deducimos que la frontera de un conjunto es la misma que la frontera de su conjunto complementario; es decir, $\text{Fr}(A) = \text{Fr}(A^c)$.

Proposición C.1.1 (Dualidad entre la clausura y el interior).

$$(\overset{\circ}{A})^c = \overline{A^c}, \quad \overline{A}^c = \overset{\circ}{A^c}.$$

DEMOSTRACIÓN. Por definición

$$\overset{\circ}{A} = \bigcup_{i \in I} \Omega_i$$

donde $(\Omega_i)_{i \in I}$ designa la familia de todos los abiertos contenidos en A ; por consiguiente

$$(\overset{\circ}{A})^c = \bigcap_{i \in I} \Omega_i^c$$

siendo $(\Omega_i^c)_{i \in I}$ la familia de cerrados que contienen a A^c ; esta intersección es, por tanto, su clausura:

$$(\mathring{A})^c = \overline{A^c}.$$

La fórmula

$$\overline{A^c} = \widehat{A^c}$$

se deduce de la precedente reemplazando A por A^c . \square

La adherencia tiene las propiedades:

$$\overline{\emptyset} = \emptyset, \quad A \subset \overline{A}, \quad \overline{\overline{A}} = \overline{A}, \quad \overline{A \cup B} = \overline{A} \cup \overline{B}.$$

La propiedad análoga a la última, $\overline{A \cap B} = \overline{A} \cap \overline{B}$ para la intersección, no es necesariamente cierta. Sólo es cierta la relación de contenido

$$\overline{A \cap B} \subset \overline{A} \cap \overline{B}.$$

En efecto si consideramos la recta real \mathbb{R} con la topología ordinaria como el espacio topológico E , y A es el conjunto de los números racionales y B es el de los irracionales, se tiene que $\overline{A} = \mathbb{R}$, $\overline{B} = \mathbb{R}$, pero $A \cap B = \emptyset$; de donde $\overline{A \cap B} = \emptyset$ y $\overline{A} \cap \overline{B} = \mathbb{R}$.

Las propiedades del interior son las siguientes:

$$\mathring{E} = E, \quad \mathring{A} \subset A, \quad \mathring{\mathring{A}} = \mathring{A}, \quad \mathring{A \cap B} = \mathring{A} \cap \mathring{B}.$$

Con el mismo ejemplo de los números racionales e irracionales, se tiene que $\mathring{A} = \emptyset$, $\mathring{B} = \emptyset$, pero $A \cup B = \mathbb{R}$ y, por lo tanto, el contenido en la inclusión

$$\mathring{A} \cup \mathring{B} \subset \mathring{A \cup B}$$

puede ser estricto.

Proposición C.1.2. Para todo subconjunto A de E , se tiene que $\text{Fr}(A) = \overline{A} \setminus \mathring{A}$.

DEMOSTRACIÓN. Se tienen las relaciones

$$\text{Fr}(A) = \overline{A} \cap \overline{A^c}, \text{ y } \overline{A^c} = (\mathring{A})^c, \text{ de donde } \text{Fr}(A) = \overline{A} \cap (\mathring{A})^c,$$

que no es otra cosa que la relación buscada. \square

Corolario 1. Para todo subconjunto cerrado C de E , se tienen las equivalencias:

$$(C = \text{Fr}(C)) \iff (\mathring{C} = \emptyset) \iff (\overline{C^c} = E).$$

Para todo subconjunto A del espacio E se tiene la siguiente partición de E en conjuntos dos a dos disjuntos:

$$E = \mathring{A} \dot{\cup} \text{Fr}(A) \dot{\cup} \text{Ext}(A), \tag{C.1}$$

donde $\dot{\cup}$ significa “unión disjunta”. Además, $\overline{A} = \mathring{A} \dot{\cup} \text{Fr}(A)$. Los conjuntos \mathring{A} y $\text{Ext}(A)$ son abiertos, y $\text{Fr}(A)$ es cerrado.

C.2. Conjuntos diseminados

Definición C.2.1. Sea A un subconjunto de E . Se dice que A es *denso* si $\overline{A} = E$. Se dice que A es *diseminado* (o *ralo*) si el interior de \overline{A} es vacío. Con símbolos A se dice *diseminado* si $\mathring{\overline{A}} = \emptyset$. En inglés a un conjunto diseminado se le llama “nowhere dense”.

Es fácil demostrar que A es diseminado si y sólo si $\overline{A^c}$ es denso, i.e. $\overline{\overline{A^c}} = E$. Esto es equivalente a probar que todo abierto no vacío U tiene una intersección no vacía con $\overline{A^c}$. Asimismo, la unión de un número finito de conjuntos diseminados es diseminado. Para demostrarlo es suficiente demostrar el lema siguiente.

Lema C.2.1. *Si A y B son conjuntos diseminados de un espacio topológico E , entonces el conjunto $A \cup B$ es diseminado.*

DEMOSTRACIÓN. Basta demostrar el Lema para el caso en que los conjuntos diseminados sean cerrados. En efecto, en el caso general, $\overline{A \cup B} = \overline{A} \cup \overline{B}$. Como \overline{A} y \overline{B} son cerrados y diseminados, por lo que ya estaría demostrado, su unión, $\overline{A \cup B}$, es diseminado. Pero esto implica que $\overline{\overline{A \cup B}} = \emptyset$. Pero $\overline{\overline{A \cup B}} = \overline{A \cup B}$; por lo tanto, $\overline{A \cup B} = \emptyset$. Es decir, el conjunto $A \cup B$ es diseminado.

Supongamos ahora que A y B son conjuntos cerrados y diseminados. Esto equivale a que para todo abierto U la intersección $U \cap A^c$ es no vacía. Como $U \cap A^c$ es un abierto no vacío y B^c es denso, se tiene que

$$(U \cap A^c) \cap B^c = U \cap (A^c \cap B^c) \neq \emptyset.$$

Mas, $A^c \cap B^c = (A \cup B)^c$. Por consiguiente, $(A \cup B)^c$ es denso, y, como $\overline{\overline{A \cup B}^c} = (A \cup B)^c$, se tiene que $A \cup B$ es diseminado. \square

La frontera de todo conjunto abierto es diseminada. De esto trata la proposición siguiente, cuya demostración hemos tomado prestada de [56].

Proposición C.2.1. *Si A es un conjunto abierto de un espacio topológico E , entonces su frontera $\text{Fr}(A)$ es un conjunto diseminado.*

DEMOSTRACIÓN. Sea $B := \text{Fr}(A)$. Ya que $B = \overline{A} \cap \overline{A^c}$, B es cerrado; por consiguiente todo lo que necesitamos probar es que $\overset{\circ}{B} = \emptyset$. Primero, observemos que $B = \overline{A} \cap A^c$ pues A es abierto. Ahora utilizamos la propiedad

$$\widehat{C \cap D} = \overset{\circ}{C} \cap \overset{\circ}{D}.$$

Así pues,

$$\overset{\circ}{B} = \overset{\circ}{\overline{A} \cap A^c} = \overset{\circ}{\overline{A}} \cap \overset{\circ}{A^c} \subset \overline{A} \cap \overline{A^c} = \emptyset,$$

ya que $\overset{\circ}{A^c} = \overline{A^c}$ por la Proposición C.1.1. \square

Corolario 1. *La frontera de todo conjunto cerrado es diseminada.*

DEMOSTRACIÓN. Sea C un conjunto cerrado, entonces C^c es abierto. Como $\text{Fr}(C) = \text{Fr}(C^c)$, se sigue que $\text{Fr}(C)$ es un conjunto diseminado. \square

C.3. Estructura de los conjuntos abiertos de \mathbb{R}

Un conjunto de puntos de \mathbb{R} se llama *acotado* si es un subconjunto de algún intervalo finito. Sean Ω un conjunto abierto de \mathbb{R} y (a, b) un intervalo abierto que está contenido en Ω , pero cuyos extremos no pertenecen a Ω . En tal caso (a, b) se llama un *intervalo componente* de Ω .

Teorema C.3.1. *Sea Ω un subconjunto abierto acotado de \mathbb{R} . Entonces, se tiene que:*

- (I) *Cada punto de Ω pertenece exclusivamente a un determinado intervalo componente de Ω .*
- (II) *Los intervalos componentes de Ω constituyen una colección finita o numerable de conjuntos disjuntos dos a dos cuya unión es Ω .*

DEMOSTRACIÓN. Supongamos que $x \in \Omega$. Entonces x está contenido en algún intervalo abierto I contenido en Ω . Existen muchos intervalos I , pero el “mayor” de ellos será el intervalo componente que contiene a x . Para determinar este intervalo componente procedemos así: Sean a el ínfimo de los extremos izquierdos de todos los intervalos I tales que $x \in I \subset \Omega$, y b el supremo de los extremos derechos de dichos intervalos. (Se utiliza aquí la acotación de Ω). Obviamente $x \in (a, b)$. Además fácilmente se ve que $(a, b) \subset \Omega$.

C.4. Notas al Apéndice C

Muchas de las explicaciones de este apéndice están inspiradas en las páginas 11 a 17 del libro de Choquet [17]. El Teorema C.3.1 que describe la estructura de los conjuntos abiertos de \mathbb{R} puede verse en los libros de Apostol [2, Teorema 3-9, pág. 43] y de Kolmogorov y Fomin [40, Theorem 6, pág. 51]. Hay un resultado análogo que describe los conjuntos abiertos de \mathbb{R}^n como unión de una colección numerable de cubos con interiores disjuntos; este resultado puede verse en los libros de Stein [54, Section 3, pág. 16, y Section 1, pág. 167] y de Asplund y Bungart [4, Proposition 4.2.16 en la pág. 166]

$\overset{\circ}{A}$ $\overset{\circ}{A}$

$$\widehat{\overset{\circ}{ABC}} \quad \widehat{\overset{\circ}{A \cap B}}$$

Apéndice D

Funciones de variable compleja

D.1. Representación integral de los ceros

Si Ω es un subconjunto abierto de \mathbb{C} , denotaremos por $H(\Omega)$ al conjunto de las funciones complejas holomorfas o analíticas en Ω . Si $\gamma : [\alpha, \beta] \rightarrow \mathbb{C}$ es una función, a su imagen $\gamma([\alpha, \beta])$ la denotaremos también por γ^* .

Teorema D.1.1. *Sea Ω un abierto de \mathbb{C} . Supongamos que $f, \varphi \in H(\Omega)$ y que $a \in \Omega$ y $r > 0$ son tales que $\overline{B(a, r)} \subset \Omega$. Sea γ la circunferencia orientada positivamente de centro a y radio r . Supongamos que f no tiene ceros en γ^* . Supongamos además que f no es constante sobre $B(a, r)$. Entonces*

- I. *O bien para todo $z \in B(a, r)$, $f(z) \neq 0$.*
- II. *O bien f tiene ceros en $B(a, r)$ (en número finito necesariamente); digamos que estos ceros son z_1, \dots, z_n ; en este caso,*

$$\frac{1}{2\pi i} \int_{\gamma} \varphi(z) \frac{f'(z)}{f(z)} dz = \sum_{k=1}^n \varphi(z_k) m(z_k)$$

siendo $m(z_k)$ el orden del cero que f tiene en $z_k, k = 1, \dots, n$.

Tomando $\varphi(z) = z$, se tiene el corolario siguiente.

Corolario 1. *Sean $f \in H(\Omega)$, Ω abierto de \mathbb{C} , $r > 0$, $\overline{B(a, r)} \subset \Omega$, y γ la circunferencia orientada positivamente de centro a y radio r ; con $0 \notin f(\gamma^*)$. Supongamos que f tiene un único cero z_0 en $B(a, r)$ y que es un cero de orden de multiplicidad m . Entonces*

$$z_0 = \frac{1}{2\pi i m} \int_{\gamma} z \frac{f'(z)}{f(z)} dz.$$

Haciendo $\varphi(z) \equiv 1$, obtenemos este corolario.

Corolario 2. *Sean $f \in H(\Omega)$, Ω abierto de \mathbb{C} , $r > 0$, $\overline{B(a, r)} \subset \Omega$, y γ la circunferencia orientada positivamente de centro a y radio r ; con $0 \notin f(\gamma^*)$. Entonces el número de ceros de f dentro de $B(a, r)$, contando sus multiplicidades, es igual a*

$$\frac{1}{2\pi i} \int_{\gamma} \frac{f'(z)}{f(z)} dz.$$

D.2. Integrales paramétricas

Teorema D.2.1. *Sea γ una circunferencia orientada positivamente en el plano complejo. Sean $t_0, t_1 \in \mathbb{R}, t_0 < t_1$ y*

$$f : \gamma^* \times [t_0, t_1] \rightarrow \mathbb{C}$$

una función continua en $\gamma^ \times [t_0, t_1]$. Definamos*

$$\varphi(t) := \int_{\gamma} f(z, t) dz, \quad t \in [t_0, t_1].$$

Entonces φ es continua en $[t_0, t_1]$.

Teorema D.2.2. *Sea γ una circunferencia orientada positivamente en el plano complejo. Sean $t_1 \in \mathbb{R}$ y*

$$f : \gamma^* \times [t_1, \infty) \rightarrow \mathbb{C}$$

una función continua en $\gamma^ \times [t_1, \infty)$ tal que la derivada parcial*

$$\frac{\partial f}{\partial t} : \gamma^* \times [t_1, \infty) \rightarrow \mathbb{C}$$

existe y es continua en todo punto $(z, t) \in \gamma^ \times [t_1, \infty)$.*

Definamos

$$\varphi(t) := \int_{\gamma} f(z, t) dz, \quad t \in [t_1, \infty).$$

Entonces φ es derivable en $[t_1, \infty)$, para todo t de este intervalo

$$\varphi'(t) = \int_{\gamma} \frac{\partial f}{\partial t}(z, t) dz$$

y φ' es continua en $[t_1, \infty)$.

Bibliografía

- [1] G. E. Andrews. *The Theory of Partitions*. Addison-Wesley, Reading, Massachusetts, 1976. (Citado en la página 45).
- [2] T. M. Apostol. *Análisis matemático*. Reverté, Barcelona, 1960. (Citado en las páginas 14 y 176).
- [3] B. C. Arnold. *Majorization and the Lorenz Order*. Number 43 in Lecture Notes in Statistics. Springer-Verlag, Berlin, New York, 1987. (Citado en la página 21).
- [4] E. Asplund and L. Bungart. *A first course in integration*. Holt, Rinehart and Winston, New York, 1966. (Citado en la página 176).
- [5] F. V. Atkinson. *Discrete and Continuous Boundary Problems*. Academic Press, 1964. (Citado en la página 93).
- [6] S. Barnett. *Polynomials and linear control systems*. Marcel Dekker, Inc., New York, 1983. (Citado en la página 79).
- [7] M. Barysz, D. Bonchev, and O. Mekenyan. Graph-centre, self-returning walks, and critical pressure of alkenes. *Match*, (20):125–139, 1986. (Citado en la página 21).
- [8] M. A. Beitia. *Ecuación Matricial y Producto Tensorial en la Semejanza por Bloques*. PhD thesis, Universidad del País Vasco, Bilbao, 1992. (Citado en la página 45).
- [9] M. A. Beitia and J.-M. Gracia. Sylvester matrix equation for matrix pencils. *Linear Algebra Appl.*, 232:155–197, 1996. (Citado en la página 45).
- [10] R. E. Bellman. *Introduction to Matrix Analysis*. McGraw-Hill, second edition, 1970, 1960. (Citado en las páginas VIII y 8).
- [11] R. Bhatia. *Perturbation Bounds for Matrix Eigenvalues*. Longman Scientific and Technical, Essex CM20 2JE, England, 1987. (Citado en la página 95).
- [12] R. Bhatia. *Matrix analysis*. Springer, 1997. (Citado en las páginas 93, 96 y 139).
- [13] K. Bongartz. Degenerations for representations of tame quivers. *Annales scientifiques de l'É.N.S.*, 28(5):647–668, 1995. (Citado en la página 117).
- [14] R. A. Brualdi and H. J. Ryser. *Combinatorial Matrix Theory*. Cambridge University Press, New York, 1991. (Citado en la página 15).
- [15] C. Byrnes and M. Gauger. Characteristic free, improved decidability criteria for the similarity problem. *Linear Multilinear Algebra*, 5:153–158, 1977. (Citado en las páginas 45 y 74).
- [16] S. L. Campbell and C. D. Meyer. *Generalized inverses of linear transformations*. Pitman, London, 1979. (Citado en la página 15).

- [17] G. Choquet. *Cours d'analyse. Topologie*. Masson, Paris, 1964. Tome II. (Citado en las páginas 142 y 176).
- [18] L. Comtet. *Analyse Combinatoire*. Presses Universitaires de France, 108, Boulevard Saint-Germain, Paris, 1970. Tome Premier. (Citado en la página 45).
- [19] H. den Boer and G. P. A. Thijsse. Semi-stability of sums of partial multiplicities under additive perturbation. *Integral Equations Operator Theory*, 3(1):23–42, 1980. (Citado en la página 117).
- [20] J. Dieudonné. *Calcul Infinitésimal*. Hermann, Paris, 1968. (Citado en la página 140).
- [21] J. Dixon. An isomorphism criterion for modules over a principal ideal domain. *Linear Multilinear Algebra*, 8:69–72, 1979. (Citado en la página 75).
- [22] J. Edmonds. Systems of distinct representatives and linear algebra. *J. R. National Bureau of Standards*, 71B(4):241–245, December 1976. (Citado en la página 11).
- [23] J.-C. Evard. On matrix functions which commute with their derivative. *Linear Algebra Appl.*, 69:145–178, 1985. (Citado en la página 136).
- [24] J.-C. Evard and J.-M. Gracia. On similarities of class C^p and applications to matrix differential equations. *Linear Algebra Appl.*, 137/138:363–386, 1990. (Citado en las páginas 136 y 139).
- [25] S. Friedland. Spectral theory of matrices. Technical Summary Report 2082, Mathematics Research Center. University of Wisconsin-Madison, 610 Walnut Street, Madison, Wisconsin 53706, 1980. (Citado en la página 74).
- [26] S. Friedland. *Matrices*. <http://www2.math.uic.edu/~friedlan/bookm.pdf>, 2010. (Citado en la página 96).
- [27] E. R. Gansner. Acyclic digraphs, young tableaux and nilpotent matrices. *SIAM J. Alg. Disc. Meth.*, 2(4):429–440, December 1981. (Citado en la página 15).
- [28] F. R. Gantmacher. *Théorie de Matrices*, volume 1. Dunod, Paris, 1966. Traduit par Ch. Sarthou. (Citado en las páginas VIII, 66, 121 y 155).
- [29] F. R. Gantmacher. *Théorie des Matrices*, volume 2. Dunod, Paris, 1966. Traduit par Ch. Sarthou. (Citado en la página VIII).
- [30] R. Godement. *Cours D'algèbre*. Hermann, Paris, deuxième édition, 1966. (Citado en la página 89).
- [31] I. Gohberg, P. Lancaster, and L. Rodman. *Invariant subspaces of matrices with applications*. Wiley, 1986. (Citado en las páginas VII, 53, 61, 117 y 136).
- [32] G. Golub and C. van Loan. *Matrix computations*. The Johns Hopkins University Press, third edition, 1996. (Citado en la página VII).
- [33] J.-M. Gracia, I. de Hoyos, and I. Zaballa. A characterization of feedback equivalence. *SIAM J. Control and Optimization*, 28(5):1103–1112, Sept. 1990. (Citado en la página 45).
- [34] J.-M. Gracia, I. de Hoyos, and F.-E. Velasco. Safety neighbourhoods for the invariants of the matrix similarity. *Linear Multilinear Algebra*, 46:25–49, 1999. (Citado en la página 117).
- [35] G. H. Hardy, J. Littlewood, and G. Pólya. *Inequalities*. Cambridge University Press, London and New York, 1st, 2nd edition, 1934, 1952. (Citado en la página 45).

- [36] J. A. Holbrook. Polynomials in a matrix and its commutant. *Linear Algebra Appl.*, 48: 293–301, 1982. (Citado en las páginas 96 y 117).
- [37] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, New York, 1985. (Citado en la página VII).
- [38] G. James and A. Kerber. *The Representation Theory of the Symmetric Group*. Addison-Wesley, 1981. (Citado en la página 45).
- [39] T. Kato. *A short introduction to perturbation theory for linear operators*. Springer, 1982. (Citado en la página 93).
- [40] A. Kolmogorov and S. Fomin. *Introductory real analysis*. Prentice-Hall, Inc., Englewoods Cliffs, N.J., 1970. (Citado en la página 176).
- [41] P. Lancaster and M. Tismenetsky. *The theory of matrices with applications*. Academic Press, second edition, 1985. (Citado en las páginas VII, 2, 8, 66, 88, 121, 145, 147, 152 y 155).
- [42] R. A. Lippert. Fixing multiple eigenvalues by a minimal perturbation. *Linear Algebra Appl.*, 432:1785–1817, 2010. (Citado en las páginas 45 y 52).
- [43] R. A. Lippert and G. Strang. The Jordan forms of AB and BA . *Electronic Journal of Linear Algebra*, 18:281–288, June 2009. (Citado en la página 45).
- [44] I. Macdonald. *Symmetric Functions and Hall Polynomials*. Clarendon Press, Oxford, 1979. (Citado en la página 45).
- [45] C. C. MacDuffee. *The Theory of Matrices*. Chelsea, New York, 1946. (Citado en las páginas 74 y 75).
- [46] A. S. Markus and E. É. Parilis. The change of the Jordan structure of a matrix under small perturbations. *Linear Algebra Appl.*, 54:139–152, 1983. (Citado en la página 117).
- [47] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, 1979. (Citado en las páginas 21, 44 y 45).
- [48] K. Matthews. Applications of the smith normal form: The Byrnes-Gauger theorem. <http://www.numbertheory.org/courses/MP274/>, 1991. (Citado en la página 74).
- [49] Z. A. Melzak. *Mathematical Ideas, Modeling and Applications*. Wiley, New York, 1976. (Citado en la página 45).
- [50] R. Merris. *Multilinear Algebra*. Gordon and Breach, Amsterdam, 1997. (Citado en la página 45).
- [51] S. Poljak and M. Schlegel. Computing the generic jordan canonical form. *Linear Multilinear Algebra*, 28:241–249, 1991. (Citado en la página 128).
- [52] K. Reinschke. *Multivariable Control. A Graph-theoretic Approach*. Number 108 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Berlin, 1988. (Citado en la página 168).
- [53] F. Rellich. *Perturbation Theory of Eigenvalues Problems*. Gordon and Breach, 1969. (Citado en la página 93).
- [54] E. Stein. *Singular integrals and differentiability properties of functions*. Princeton Univ. Press, Princeton, New Jersey, 1970. (Citado en la página 176).
- [55] G. Thijssse. Global holomorphic similarity to a Jordan form. *Results in Mathematics*, 8 (1):78–87, 1985. (Citado en la página 137).

- [56] C. Woo. "boundary of an open set is nowhere dense"(version 5). PlanetMath.org. <http://planetmath.org/BoundaryOfAnOpenSetIsNowhereDense.html>, 2012. (Citado en la página 175).
- [57] I. Zaballa. Sobre formas canónicas conjugadas. In E. M. de Sá and J. Vitória, editors, *Estudos de Matemática. Em homenagem a Luis de Albuquerque*, pages 343–357. Departamento de Matemática da Universidade de Coimbra, 3000 Coimbra, Portugal, 1987. (Citado en la página 74).
- [58] D. Zeilberger. A combinatorial approach to matrix algebra. *Discrete Mathematics*, 56: 61–72, 1985. (Citado en la página 15).

Índice alfabético

- algebraicamente independiente, 15
- arco incidente con un vértice, 159
- autoconjugada, 19
- bifurcación
 - punto de, 140
- bloque de Jordan, 153
- característica de Segré, 49, 154
- característica de Weyr, 52
- circulante, 116
- Coates, 168
- conjugada, 18
- conjunto abierto
 - independencia lineal, 4
- conjunto diseminado, 174
- conjunto ralo, 174
- convergencia compacta
 - topología de la, 142
- descomposición en permutaciones-ciclo, 166
- desigualdad de Frobenius, 48
- desigualdad de Sylvester, 48
- determinante de Vandermonde, 171
- diagonales de una matriz, 122, 125, 126
 - definición, 169
- diagrama de Hasse, 40
- diagrama de Young o Ferrers, 18
- digrafo, 159
 - acíclico, 160
 - arco de un, 159
 - camino de un, 160
 - ciclo de un, 160
 - ciclos
 - familia de, 160
 - familia de ciclos
 - generadora, 161
 - lazo de un, 159
 - paseo de un, 159
 - peso de un, 161
 - ponderado, 161
 - subdigrafo de un, 159
 - vértice de un, 159
- digrafo permutación, 167
- digrafo ponderado
 - de Coates, 121, 168
- discrepancia, 34
- diseminado, 174
- divisor determinantal, 147
- divisor elemental, 149, 152
 - lineal, 149
- emparejamiento óptimo, 89
- equivalencia
 - de matrices polinómicas, 147
- espectro
 - ordenado, 50
- 1-factor, 161
- factor invariante, 148, 152
- familia de ciclos, 122, 160
 - longitud de una, 160
- familia de ciclos generadora, 161
- Ferrers, 18
- forma canónica
 - definición, 1
- forma canónica de Jordan, 154
- forma canónica racional, 152
- forma normal
 - definición, 1
- forma normal de Jordan, 154
- forma normal de Smith, 148
- Frobenius
 - desigualdad sobre rangos, 48
- función
 - semicontinua inferiormente, 3
- genérico, 141
- grado
 - entrante de un vértice, 159
 - saliente de un vértice, 159
- Hasse, 40
- Hilbert
 - norma de, 12
- independencia lineal
 - conjunto abierto, 4

- indeterminadas, 15
- invariante
 - completo, 1
 - definición, 1
- invariantes
 - sistema completo de, 1
- Jordan, *véase también* forma canónica, 1, 154
 - bloque de, 153
 - estructura de, 49
 - estructura repelida, 112
- lexicográfico, 35
- matrices
 - nilpotentes, 154
 - polinómicas, 146
 - polinomiales, 146
 - transformaciones elementales de, 145
 - unimodulares, 147
- λ -matrices, 146
- matrices de bloques
 - transformaciones elementales de, 156
- matrices elementales, 145
 - por bloques, 158
- matrices polinómicas
 - transformaciones elementales de, 146
- matrices polinómicas elementales, 147
- matrices semejantes, 151
- matriz
 - resultante de Sylvester, 78
 - resultante generalizada de Sylvester, 78
- matriz característica, 152
- matriz compañera, 152
- matriz de permutación, 167
- matriz de Toeplitz, 116
- mayoración, 20
- mayoración débil $\prec\prec$, 20
- mayoración estricta \prec , 20
- mayorización, 20
- monotonía de $\prec\prec$ y \prec respecto de \cup , 23
- multiplicidad
 - algebraica, 49
 - geométrica, 49
 - parcial, 49
- número
 - adecuado, 97
- nilpotente, 154
- norma de Hilbert, 12
- norma de operador, 12
- norma espectral, 11
- “nowhere dense”, *véase* diseminado
- nulidad, 50
- orden lexicográfico, 35
- inverso, 35
- transpuesto, 36
- partición, 17
 - autoconjugada, 19
 - conjugada, 18
 - peso de una, 18
- partición de Segré, 49
- partición de Weyr, 51
- permutación
 - impar, 164
 - matriz de, 167
 - par, 164
- permutación, *véase también* familia de ciclos generadora
- permutación cíclica, 165
- permutación-ciclo, 165
- perturbación estructurada, 97
- peso
 - de un arco, 161
- polinomio invariante, 152
- producto
 - simétrico, 92
- punto de bifurcación
 - de la característica de Segré, 140
 - del espectro, 140
 - del espectro de una función matricial, 135
- ralo, 174
- rango, *véase* semicontinuidad del rango
 - de matrices aproximantes, 7
 - desigualdad de Frobenius, 78
 - función semicontinua, 3
 - ley de Sylvester, 48
- rango genérico, 10
- rango límite, 10
- relación de equivalencia, 1
- repelida, 112
- Segré, 154
 - característica de, 49
 - partición de, 49
- semicontinuidad del rango, 3
- signatura
 - de un producto, 171
 - de una permutación, 164
 - transposición, 164
- subdiagonal, 10
- subdigrafo ponderado, 161
- subespacio
 - radical, 135
- suma de particiones, 18
- Sylvester
 - Ley de, 48
 - resultante de, 78

Teorema de Gale y Ryser, [44](#)
Toeplitz, [116](#)
topología
 de la convergencia compacta, [142](#)
transposición, [163](#)

unión de particiones, [18](#)
unimodular, [147](#)

valores singulares, [11](#)
Vandermonde, [171](#)
vectorizador, [68](#)

Weyr
 característica de, [52](#)
 partición de, [51](#)

Young, [18](#)