El Teorema Espectral y Formas Canónicas*

G.A. Raggio**

Resumen

Luego de una brevísima reseña del álgebra lineal se demuestran: El teorema espectral en su forma general y la concomitante forma canónica de Jordan en el caso de espacios vectoriales complejos de dimensión finita; y el teorema de triangularización unitaria de Schur para espacios vectoriales complejos con producto escalar. El caso real también se trata.

1. Operadores lineales sobre un espacio vectorial de dimensión finita

Todo el material de esta primera sección se ofrece como repetitorio (y para fijar parte de la notación); las demostraciones pueden estudiarse en alguno de los buenos de los cientos de libros de álgebra lineal. Algunos son realmente patéticos. Es dificl elegir un puñado y recomendarlos pero hay uno que es mi preferido; es el libro *Linear Algebra and its Applications*[1] de Peter D. Lax. Lo mejor que puede hacer para su formación intelectual es tirar estas notas y leer el libro de Lax donde encontrará todo lo que se ve en estas notas y muchísimo mas.

Un **espacio vectorial** complejo (resp. real) V es un conjunto –cuyos elementos se denominan **vectores**– con una operación binaria $+$ y un producto $\cdot$ por números complejos (resp. reales), tales que: (1) $V \times V \ni (x, y) \mapsto x + y \in V$ satisface todas las propiedades de la suma usual; a saber: es comutativa $x + y = y + x$, es asociativa $x + (y + z) = (x + y) + z$, hay un elemento neutro $0 \in V$ tal que $x + 0 = x$ y para todo $x \in V$ hay $y \in V$ tal que $x + y = 0$; (2) $\mathbb{C} \times V \ni (\alpha, x) \mapsto \alpha \cdot x \in V$ (resp. $\mathbb{R} \times V \ni (\alpha, x) \mapsto \alpha \cdot x \in V$), satisface $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$ y $1 \cdot x = x$; y además (3) $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$, y $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$. En lo que sigue listamos las definiciones y –sin dar la demostración– los resultados básicos necesarios para el estudio de las transformaciones lineales entre espacios vectoriales.

El término **escalar** denota alternativamente un número complejo o un número real, según corresponda al espacio vectorial complejo o real que se considera.

*Estas notas fueron elaboradas para la edición 2006 del curso *Métodos Matemáticos de la Física* del tercer año de la Licenciaturas en Física y en Astronomía de FaMAF. Se agradece la comunicación de errores, comentarios, etc. a: *raggio@famaf.unc.edu.ar*

**FaMAF, Universidad Nacional de Córdoba.

[1] John Wiley & Sons, 2^{nd} edition, New York .

Dependencia lineal, bases y dimensión, subespacios Un conjunto $S \subset V$ se dice **linealmente dependiente** si existe un número finito de vectores $x_1, x_2, \cdots, x_n \in S$ y escalares $\alpha_1, \alpha_2, \cdots, \alpha_n$ no todos nulos tales que $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0$. Un conjunto que no es linealmente dependiente se dice **linealmente independiente**. Un vector $x \in V$ es **combinación lineal** de los vectores $x_1, x_2, \cdots, x_n \in V$ si existen escalares $\alpha_1, \alpha_2, \cdots, \alpha_n$ tales que $x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots \alpha_n x_n$. Si $C \subset V$, denotamos con $\mathsf{lin}\{C\}$ al conjunto de las combinaciones lineales de los elementos de C.

Una **base**[2] B de V es un conjunto de vectores linealmente independientes de V tales que todo $x \in V$ es combinación lineal de vectores en B ($\mathsf{lin}(B) = V$). Las **componentes** de $x \in V$ en la base B son los escalares x_j tales que $x = \sum_j x_j b_j$ con $b_j \in B$ (¡la suma es finita!).

Teorema 1.1 *Las componentes de un vector en una base son unívocas.*

Teorema 1.2 *Todo espacio vectorial admite una base. Si V admite una base finita entonces todas las bases de V constan de la misma cantidad de elementos.*

Si V admite base finita se dice que V es de dimensión finita n donde n es el número de elementos de cualquier base de V. En caso contrario se dice que V es de dimensión infinita.

Teorema 1.3 *Si V es un espacio vectorial complejo (resp. real) de dimensión finita n y $B = \{b_1, b_2, \cdots, b_n\}$ es una base de V entonces*

$$\Phi_B(x) = (x_1, x_2, \cdots, x_n) \ ,$$

donde $x_1, x_2, \cdots, x_n$ son las componentes de x en la base B, es una transformación lineal inyectiva y suryectiva de V en $\mathbb{C}^n$ (resp. de V en $\mathbb{R}^n$), cuya inversa es la aplicación lineal

$$\Phi_B^{-1}(x_1, x_2, \cdots, x_n) = \sum_{j=1}^{n} x_j b_j \ .$$

Un **subsespacio** $E \subset V$ es un conjunto tal que toda combinación lineal de vectores de E está en E: $\mathsf{lin}E = E$. En tanto y en cuanto un subespacio E es un espacio vectorial sobre los mismos escalares que V, su dimensión está definida y la anotamos $\dim(E)$. Si V es de dimensión finita, todo subespacio también es de dimensión finita.

Teorema 1.4 *Si E y F son subespacios de V entonces $E \cap F$ es un subespacio y* $\dim(\mathsf{lin}\{E \cup F\}) + \dim(E \cap F) = \dim(E) + \dim(F)$.

Dos subespacios E y F de V se dicen **complementarios** si $E \cap F = \{0\}$ y todo $x \in V$ es combinación lineal de un vector de E con uno de F ($\mathsf{lin}(E \cup F) = V$).

Teorema 1.5 *Si $E, F \subset V$ son subespacios complementarios entonces cualquiera sea $x \in V$ hay un único $e \in E$ y un único $f \in F$ tales que $x = e + f$.*

[2] Más precisamente una base de Hamel.

Suma directa Si U y V son espacios vectoriales sobre los mismos escalares, entonces la **suma directa** $U \oplus V$ es $U \times V$ con la siguiente suma y multiplicación por escalares

$$\alpha \cdot (u_1, v_1) + (u_2, v_2) = (\alpha u_1 + u_2, \alpha u_2 + v_2) \ .$$

Teorema 1.6 *$U \oplus V$ es un espacio vectorial.*

En el contexto de subespacios $E, F \subset V$ complementarios, se escribe $V = E \oplus F$. Esta notación es consistente con la definición de suma directa si se identifica a E con $E \times \{0\}$ y a F con $\{0\} \times F$.

De ahora en más:

$$\boxed{\mathcal{H} \text{ es un espacio vectorial complejo o real de dimensión finita } n}$$

Operadores Un **operador** A es una transformación lineal de $\mathcal{H}$ en $\mathcal{H}$; vale decir $A(f + \alpha \cdot g) = Af + \alpha \cdot (Ag)$. Si A es un operador, $ker(A) := \{x \in \mathcal{H} : \ Ax = 0\}$ es un subespacio llamado **núcleo** de A. A es **inyectivo** si $x \neq y$ implica que $Ax \neq Ay$ o, lo que es lo mismo, $ker(A) = \{0\}$. La **imagen** de A es el subespacio $\mathcal{R}_A := \{Af : \ f \in \mathcal{H}\}$. A es **suryectivo** si $\mathcal{R}_A = \mathcal{H}$. Si A es inyectivo y suryectivo entonces hay un único operador B tal que $AB = BA = \mathbf{1}$ donde $\mathbf{1}f = f$ para todo $f \in \mathcal{H}$. En tal caso, decimos que A es **invertible** y anotamos A^{-1} para el operador inverso.

Teorema 1.7 *Si A es un operador entonces $\dim(ker(A)) + \dim(\mathcal{R}_A) = n$. Todo vector de $\mathcal{H}$ puede escribirse como suma de un vector en $ker(A)$ y un vector en $\mathcal{R}_A$.*

Corolario 1.7.1 *Un operador A inyectivo sobre $\mathcal{H}$ de dimensión finita es suryectivo y por ende invertible.*

Una **proyección** P es un operador sobre $\mathcal{H}$ tal que $P^2 = P$.

Teorema 1.8 *Si P es una proyección entonces $ker(P)$ y $\mathcal{R}_P$ son subespacios complementarios. Viceversa, si E y F son subsespacios complementarios, entonces existe una única proyección Q tal que $\mathcal{R}_Q = E$ y $ker(Q) = F$.*

Si A y B son operadores sobre $\mathcal{H}$ y α es un escalar, el operador $A + \alpha B$ está definido por $(A + \alpha B)f := Af + \alpha BF$, que es manifiestamente un operador sobre $\mathcal{H}$. Esto indica que el conjunto de los operadores de $\mathcal{H}$ en $\mathcal{H}$ forman un espacio vectorial sobre los mismos escalares que $\mathcal{H}$; escribimos $\mathcal{L}(\mathcal{H})$ para este espacio vectorial. En $\mathcal{L}(\mathcal{H})$ hay también una multiplicación natural definida por la aplicación sucesiva de operadores: $(AB)f := A(Bf)$. En general $AB \neq BA$ –analice ejemplos– por lo cual la multiplicación es no comutativa. El elemento neutro para la multiplicación es el operador identidad, que escribimos $\mathbf{1}$, dado por $\mathbf{1}f := f$. La suma y el producto de operadores tienen la propiedad distributiva usual, con lo cual $\mathcal{L}(\mathcal{H})$ es un álgebra no comutativa. ¿Cual es la dimensión de $\mathcal{L}(\mathcal{H})$? Si $\mathcal{H}$ es de dimensión n, y $B = \{b_1, b_2, \cdots, b_n\}$ una base de $\mathcal{H}$, considere los n^2 operadores dados por

$$A^{(j,k)}f := f_k b_j \ , \quad j, k \in \{1, 2, \cdots, n\} \ ,$$

donde f_k es la k-ésima componente de $f \in \mathcal{H}$ respecto de la base B. Es inmediato verificar que $A^{(j,k)}$ son lineales. Si A es un operador arbitrario sobre $\mathcal{H}$, entonces cualquiera sea

$f \in \mathcal{H}$ hay n escalares $a(f;j)$, $j = 1, 2, \cdots, n$, que son las componentes de Af en la base B, i.e.,

$$Af = \sum_{j=1} a(f;j) b_j \ . \tag{1}$$

Ya que $f = \sum_{j=1}^n f_j b_j$, tenemos

$$Af = \sum_{j=1}^n f_j A b_j = \sum_{j=1}^n f_j \sum_{k=1}^n a(b_j;k) b_k = \sum_{k=1}^n \left(\sum_{j=1}^n a(b_j;k) f_j \right) b_k \ . \tag{2}$$

Esto muestra que

$$A = \sum_{k=1}^n \sum_{j=1}^n a(b_j;k) A^{(k,j)} \ ,$$

o sea que todo operador es combinación lineal de los $A^{(j,k)}$. Supongamos que $A = \sum_{k=1}^n \sum_{j=1}^n \alpha_{j,k} \ A^{(k,j)} = 0$ para escalares $a_{j,k}$. Entonces, cualquiera sea $m = 1, 2, \cdots, n$, tendremos

$$0 = Ab_m = \sum_{k=1}^n \sum_{j=1}^n \alpha_{j,k} A^{k,j} b_m = \sum_{k=1}^n \sum_{j=1}^n \alpha_{j,k} (b_m)_j b_k = \sum_{k=1}^n \alpha_{m,k} b_k \ ;$$

y como B es una base deducimos que $\alpha_{m,k} = 0$ para todo $k, m \in \{1, 2, \cdots, n\}$. Por lo tanto $\{A^{j,k} : \ j, k \in \{1, 2, \cdots, n\}\}$ es una base de $\mathcal{L}(\mathcal{H})$ y la dimensión buscada es n^2.

Las ecuaciones (1) y (2) junto a la unicidad de las componentes de un vector respecto de una base, nos permiten calcular las componentes de Af en términos de las componentes de f:

$$(Af)_j = a(f;j) = \sum_{k=1}^n a(b_k;j) f_k \ .$$

Esta ecuación es la base del modelo matricial y el cálculo concomitante para operadores. Si definimos una matriz $n \times n$ $[A]$ tal que el elemento de la j-ésima fila y la k-ésima columna de esta matriz es el escalar

$$A_{j,k} := a(b_k;j) \ ;$$

tendremos

$$(Af)_j = \sum_{k=1}^n A_{j,k} f_k \ .$$

Recordamos la regla (convencial pero casi universalmente usada) par multiplicar matrices. Si M es una matriz $(p \times q)$ y N es una matriz $(q \times r)$, entonces la matriz producto MN es la matriz $(p \times r)$ cuyos elementos son

$$(MN_{j,k} = \sum_{\ell=1}^q M_{j,\ell} N_{\ell,k} \ , \quad j = 1, 2, \cdots, p \ , \quad k = 1, 2, \cdots, r \ .$$

Dada una base $B = \{b_1, b_2, \cdots, b_n\}$ de $\mathcal{H}$, identificamos a cada vector $f \in \mathcal{H}$ (Teorema 1.3) con la matriz $n \times 1$

$$f \longleftrightarrow \begin{bmatrix} f_1 \\ f_2 \\ \cdot \\ \cdot \\ f_n \end{bmatrix} \; ;$$

entonces

$$Af \longleftrightarrow [A] \begin{bmatrix} f_1 \\ f_2 \\ \cdot \\ \cdot \\ f_n \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^n A_{1,k} f_k \\ \sum_{k=1}^n A_{2,k} f_2 \\ \cdot \\ \cdot \\ \sum_{k=1} A_{n,k} f_k \end{bmatrix} .$$

Si A es un operador sobre $\mathcal{H}$, escribiremos $[A]$ para la matriz (cuadrada de dimensión $n \times n$ con elementos escalares) asociada a A via alguna base de $\mathcal{H}$. Si es necesario por algún motivo recalcar de que base B se tratare, escribimos $[A]_B$. Observese que la correspondencia $A \longleftrightarrow [A]$ mediada por una base cualquiera es biyectiva, que $[A + \alpha B] = [A] + \alpha[B]$, y que $[AB] = [A][B]$.

Los resultados que siguen se formulan a veces en términos de operadores, otras veces en términos de matrices asociadas a los operadores (siempre mediadas por una base). El lector deberá elegir la versión que le gusta o la que necesita y hacer la transformación pertinente.

2. El espectro de un operador

Proposición 2.1 *Sea A un operador sobre $\mathcal{H}$ y $\lambda \in \mathbb{C}$ (resp. $\lambda \in \mathbb{R}$). Las siguientes condiciones son equivalentes:*

1. *existe un vector $f \in \mathcal{H}$ no nulo tal que $Af = \lambda f$;*
2. *$A - \lambda \mathbf{1}$ no es inyectivo;*
3. *para alguna base de $\mathcal{H}$ la matriz $[A]$ asociada a A satisface $det([A] - \lambda \mathbf{1}) = 0$;*
4. *cualquiera sea la base de $\mathcal{H}$, se tiene $det([A] - \lambda \mathbf{1}) = 0$ para la matriz $[A]$ asociada con A.*

El polinomio de grado n, $\chi_A : \mathbb{C} \to \mathbb{C}$ (resp. $\chi_A : \mathbb{R} \to \mathbb{R}$), definido por

$$\chi_A(\zeta) := det([A] - \zeta \mathbf{1})$$

no depende de la base de $\mathcal{H}$ usada para definir $[A]$.

<u>Demostración</u>: Aunque está a priori definido para la matriz $[A]$ asociada con alguna base de $\mathcal{H}$, χ_A no depende de la base elegida. En efecto, si $[A]_B$ y $[A]_{B'}$ son las matrices asociadas con A en las bases B y B' respectivamente de $\mathcal{H}$, entonces hay un operador T invertible tal que $T(B) = B'$, y entonces $[A]_{B'} = [T^{-1}AT]_B$ y por ende

$$det([A]_{B'} - \zeta \mathbf{1}) = det([T^{-1}AT]_B - \zeta \mathbf{1}) = det([T]_B^{-1}[A]_B[T]_B - \zeta \mathbf{1})$$

$$= det([T]_B^{-1}\{[A]_B - \zeta\mathbf{1}\}[T]_B) = det([T]_B^{-1})det([A]_B - \zeta\mathbf{1})det([T]_B)$$
$$= \frac{1}{det([T]_B)}det([A]_B - \zeta\mathbf{1})det([T]_B) = det([A]_B - \zeta\mathbf{1})\ .$$

La equivalencia de 3. y 4., es consecuencia de esto.
Las condiciones 1. y 2. son claramente equivalentes.
La equivalencia de 1. y 3. se trata en el curso de álgebra lineal. El lema que sigue (que usa las propiedades básicas del determinante) lo prueba. □

Lema 1 *Si A es un operador sobre $\mathcal{H}$ entonces es invertible si y solo si $det[A] \neq 0$, donde $[A]$ es la matriz asociada a $[A]$ via cualquier base de $\mathcal{H}$.*

Demostración: Si A es invertible entonces $1 = det([\mathbf{1}]) = det([AA^{-1}]) = det([A][A^{-1}]) = det([A])det([A^{-1}]$ y por lo tanto $det([A]) \neq 0$. Si A no es invertible, entonces en vistas del corolario 1.7.1, A no es inyectivo y hay $f \in \mathcal{H}$ no nulo con $Af = 0$. Sea $\{f, b_2, \cdots, b_n\}$ una base de $\mathcal{H}$; la matriz correspondiente es

$$[A] = \begin{bmatrix} 0 & * & \cdots & * \\ 0 & * & \cdots & * \\ . & . & \cdots & . \\ 0 & * & \cdots & * \end{bmatrix},$$

y tiene su primera columna nula. La estrella $*$ en una matriz denotará algún escalar (que no viene al caso especificar). Ya que la determinate es suma de $n!$ sumandos que contienen un elemento de cada fila y de cada columna de $[A]$, deducimos que $det([A]) = 0$. □

Un escalar que satisface la condición 1., se denomina **autovalor** de A, y cualquier $f \in \mathcal{H}$ con $f \neq 0$ y $Af = \lambda f$, se denomina **autovector** de A al autovalor λ. El **espectro** de A, anotado $\sigma(A)$, es el conjunto de los autovalores de A. Si $\lambda \in \sigma(A)$, entonces $\mathcal{E}_\lambda := \{f \in \mathcal{H} :\ Af = \lambda f\}$ es el **autoespacio** de A asociado con λ; es un subespacio de $\mathcal{H}$[3]. La **multiplicidad geométrica** de $\lambda \in \sigma(A)$ es la dimensión del autoespacio de A asociado a λ.

Proposición 2.2 *Si A es un operador, los autovectores a autovalores distintos son linealmente independientes.*

Demostración: Sean $\lambda_1, \lambda_2, \cdots, \lambda_k$, k distintos autovalores de A y $0 \neq f_j$, $j = 1, 2, \cdots, k$; autovectores correspondientes. Si $\{f_1, \cdots, f_k\}$ es linealmente dependiente, entonces hay un $1 \leq p < k$ tal que $\{f_1, f_2, \cdots, f_p\}$ es linealmente independiente y $f_{p+1} = \sum_{j=1}^p \alpha_j f_j$. Operando con A sobre esta identidad, obtenemos $\lambda_{p+1} f_{p+1} = \sum_{j=1}^p \alpha_j \lambda_j f_j$. Además, $\lambda_{p+1} f_{p+1} = \lambda_{p+1} \sum_{j=1}^p \alpha_j f_j$. Entonces, $0 = \sum_{j=1}^p (\lambda_{p+1} - \lambda_j)\alpha_j f_j$, y la independencia lineal de $\{f_1, \cdots, f_p\}$ indica que $(\lambda_{p+1} - \lambda_j)\alpha_j = 0$ para $j = 1, 2, \cdots, p$. Luego, $\alpha_j = 0$ para $j = 1, 2, \cdots, p$; y esto implica que $f_{p+1} = 0$, contrariando la hipótesis de que f_{p+1} es un autovector. □

[3] $0 \in \mathcal{E}_\lambda$ pero no es un autovector.

El polinomio χ_A se denomina **polinomio característico** de A. El teorema fundamental del álgebra implica –ya que χ_A tiene raices– que el espectro de un operador no es vacio si el espacio $\mathcal{H}$ es complejo. Veamos una demostración de este hecho crucial que no procede via el polinomio característico.

Teorema 2.1 *Si $\mathcal{H}$ es complejo el espectro de un operador sobre $\mathcal{H}$ no es vacio.*

<u>Demostración</u>: Tome un $0 \neq f \in \mathcal{H}$; el conjunto $\{f, Af, A^2f, \cdots, A^nf\}$ es linealmente dependiente. Sea k el menor número tal que $\{f, Af, \cdots, A^kf\}$ es linealmente independiente. Entonces, $k \geq 0$, y hay números complejos $a_o, a_1, \cdots, a_k$ no todos nulos tales que $a_of + a_1Af + \cdots + a_kA^kf + A^{k+1}f = 0$. Sea $P : \mathbb{C} \to \mathbb{C}$ el polinomio $p(z) := a_o + a_1z + \cdots + a_kz^k + z^{k+1}$. Por el teorema fundamental del álgebra, $p(z) = (z - \zeta_1)(z - \zeta_2)\cdots(z - \zeta_{k+1})$, donde $\zeta_1, \cdots, \zeta_{k+1}$ son las raices de p. Entonces, $0 = a_of + a_1Af + \cdots + a_kA^kf + A^{k+1}f = (a_o + a_1A + \cdots + a_kA^k + A^{k+1})f = (A - \zeta_1\mathbf{1})(A - \zeta_2\mathbf{1})\cdots(A - \zeta_{k+1}\mathbf{1})f$; luego alguno de los $k+1$ operadores $A - \zeta_j\mathbf{1}$, $j = 1, 2, \cdots, k+1$, no es inyectivo. □

Este resultado es falso en el caso real;

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

tiene espectro vacio. En efecto, si

$$\begin{bmatrix} a \\ b \end{bmatrix}$$

fuere autovector al autovalor $\lambda \in \mathbb{R}$ entonces: $b = \lambda a$ y $-a = \lambda b$ de donde $-a = \lambda^2 a$; con lo cual, o bien $a = 0$ en cuyo caso $b = 0$ y no hay autovector, o sino $\lambda^2 = -1$ lo que es imposible en los reales.

2.1. El polinómio minimal.

Sea $A : \mathcal{H} \to \mathcal{H}$ un operador. Considere las potencias $\{\mathbf{1}, A, A^2, A^3, \cdots\}$ en $\mathcal{L}(\mathcal{H})$ cuya dimensión como espacio vectorial complejo es n^2. Debe existir entonces una potencia p-ésima de A, A^p, con $p \in \mathbb{N}$ tal que $\{\mathbf{1}, A, A^2, A^3, \cdots, A^{p-1}\}$ es linealmente independiente y $\{\mathbf{1}, A, A^2, A^3, \cdots, A^p\}$ es en cambio linealmente dependiente. Sean $z_o, z_1, \cdots, z_p$ los escalares no todos nulos tales que

$$z_o\mathbf{1} + z_1A + z_2A^2 + \cdots + z_pA^p = 0 \ .$$

Si $z_p = 0$, la independencia lineal de $\{\mathbf{1}, A, A^2, A^3, \cdots, A^{p-1}\}$ indica que $z_0 = z_1 = \cdots = z_{p-1} = 0$; por lo tanto $z_p \neq 0$ y dividiendo por z_p podemos suponer que $z_p = 1$. La determinación práctica de los z_j se hace, por ejemplo, con el método de eliminación de Gauss. El **polinómio minimal** de A es el polinómio $\chi_A^o : \mathbb{C} \to \mathbb{C}$ (resp. $\chi_A^o : \mathbb{R} \to \mathbb{R}$) de grado p definido por

$$\chi_A^o(\zeta) := z_o + z_1\zeta + z_2\zeta^2 + \cdots + \zeta^p \ , \ \zeta \in \mathbb{C} \ .$$

En virtud de su definición, se tiene

$$\chi_A^o(A) = 0 \ .$$

Consideremos el siguiente ejemplo simple.

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

El polinomio característico es $\chi_A(\zeta) = (1-\zeta)^3$. Tenemos:

$$A^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} ;$$

y tendremos

$$0 = z_0\mathbf{1} + z_1 A + A^2 = \begin{pmatrix} z_0+z_1+1 & 0 & 0 \\ 0 & z_0+z_1+1 & z_1+2 \\ 0 & 0 & z_0+z_1+1 \end{pmatrix} ,$$

si $z_0 = 1$, $z_1 = -2$. Luego,

$$\chi_A^o(\zeta) = 1 - 2\zeta + \zeta^2 = (1-\zeta)^2 .$$

Proposición 2.3 *Si $q : \mathbb{C} \to \mathbb{C}$ es un polinomio entonces $q(A) = 0$ si y sólo si el polinomio minimal de A es un divisor de q.*

<u>Demostración</u>: Si $q(z) = \chi_A^o(z) r(z)$ para algún polinomio r, entonces $q(A) = \chi_A^o(A) r(A) = 0$. Suponga que $q(A) = 0$. En virtud de la definición del polinomio minimal, el grado de q es mayor o igual al de χ_A^o y hay entonces polinomios t y r tales que $q(z) = \chi_A^o(z)t(z) + r(z)$ donde el grado de r es estrictamente menor al de χ_A^o. Entonces, $0 = q(A) = \chi_A^o(A)t(A) + r(A) = r(A)$ y esto implica que r se anula idénticamente pues sino r sería un polinomio que anula a A de menor grado que χ_A^o. $\square$

Proposición 2.4 *El espectro de A coincide con las raíces de su polinomio minimal.*

<u>Demostración</u>: Si $\lambda \in \sigma(A)$, hay $0 \neq f \in \mathcal{H}$ tal que $Af = \lambda f$ y, a fortiori, $A^k f = \lambda^k f$ para todo $k \in \mathbb{N}$. Luego, $0 = \chi_A^o(A) f = \chi_A^o(\lambda) f$ y necesariamente $\chi_A^o(\lambda) = 0$.
Si μ es raíz de χ_A^o entonces $\chi_A^o(z) = (z-\mu)t(z)$ donde t es un polinomio. Tenemos $0 = \chi_A^o(A) = (A - \mu\mathbf{1})t(A)$ y por ende $(A-\mu\mathbf{1})t(A)g = 0$ para todo $g \in \mathcal{H}$. Ya que el grado de t es inferior al de χ_A^o, $t(A) \neq 0$, y existe entonces $g_o \in \mathcal{H}$ tal que $t(A)g_o \neq 0$. Entonces $f := t(A)g_o$ satisface $(A - \mu\mathbf{1})f = 0$, o sea que f es autovector de A al autovalor μ. $\square$

3. Análisis de un operador

Un subespacio $V \subset \mathcal{H}$ se dice **invariante** para el operador A si $A(V) \subset V$. El análisis de un operador se verá considerablemente simplificado si conseguimos encontrar subespacios invariantes de dimensión lo más chica posible.

Si A es un operador sobre $\mathcal{H}$ que es complejo, $\sigma(A)$ no es vacio y hay por ende un autovalor λ y un autovector $f \in \mathcal{H}$ a este autovalor. El subespacio $\mathsf{lin}\{f\}$ es invariante y $Af \in \mathsf{lin}\{f\}$. Si conseguimos identificar un vector g linealmente independiente de f pero tal que Ag dependa linealmente de f y de g, i.e., $Ag \in \mathsf{lin}\{f, g\}$; entonces el operador A restringido a $\mathsf{lin}\{f, g\}$ toma la forma

$$\begin{bmatrix} \lambda & * \\ 0 & * \end{bmatrix},$$

en la base $\{f, g\}$. Si este proceso se puede continuar, o sea hay h tal que $\{f, g, h\}$ es linealmente independiente pero $Ah \in \mathsf{lin}\{f, g, h\}$, entonces la restricción de A a este subespacio tridimensional tendrá la forma matricial

$$\begin{bmatrix} \lambda & * & * \\ 0 & * & * \\ 0 & 0 & * \end{bmatrix},$$

en la base $\{f, g, h\}$. Etc. Que esto es posible es el resultado siguiente:

Teorema 3.1 *(Teorema de Triangularización) Si $\mathcal{H}$ es complejo y A es un operador sobre $\mathcal{H}$, existe una base $\{f_1, f_2, \cdots, f_n\}$ de $\mathcal{H}$ tal que $Af_j = \mathsf{lin}\{f_1, f_2, \cdots, f_j\}$ para $j = 1, 2, \cdots, n$.*

Demostración: La demostración procede por inducción en la dimensión n de $\mathcal{H}$.
Si $\mathcal{H}$ es unidimensional, $\mathcal{H} = \mathbb{C}$, la afirmación es verdadera. Repetimos la notación y la argumentación del preambulo del teorema. Hay un autovalor λ con correspondiente autovector f. El subespacio $V := \mathcal{R}_{A-\lambda\mathbf{1}}$ tiene dimensión p inferior a $\mathcal{H}$ y es invariante para A. Sea B la restricción de A a V. La hipótesis inductiva indica que hay una base $\{f_1', f_2', \cdots, f_p'\}$ de V que cumple con las afirmaciones del teorema. Sean $f_j := f_j'$ para $j = 1, 2, \cdots, p$, y $f_j := g_{j-p}$ donde $\{g_1, g_2, \cdots, g_{n-p}\}$ es un conjunto de vectores de $\mathcal{H}$ que son linealmente independientes y no están en V. Entonces, $\{f_1, \cdots, f_n\}$ es una base de $\mathcal{H}$ y se tiene

$$Af_j \in \mathsf{lin}\{f_1, f_2, \cdots, f_j\}$$

para $j = 1, 2, \cdots, p$; y también, para $j = p + 1, \cdots, n$,

$$Af_j = \underbrace{(A - \lambda\mathbf{1})f_j}_{\in V = lin\{f_1, \cdots, f_p\}} + \lambda f_j \in \mathsf{lin}\{f_1, \cdots, f_p, f_j\} \subset \mathsf{lin}\{f_1, \cdots, f_j\}\,. \square$$

Corolario 3.1.2 *Si $\mathcal{H}$ es complejo y A es un operador sobre $\mathcal{H}$, existe una base de $\mathcal{H}$ respecto de la cual la matriz $[A]$ asociada tiene forma triangular superior*

$$[A] = \begin{bmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \cdot & \cdot & \cdot & \cdots & * \\ 0 & 0 & 0 & \cdots & * \end{bmatrix}.$$

Suponga que $[A]$ es triangular superior. El polinomio característico de A (o $[A]$) factoriza según $\chi_A(z) = (z-\lambda_1)(z-\lambda_2)\cdots(z-\lambda_n)$ donde los autovalores $\lambda_1, \lambda_2, \cdots, \lambda_n$, repetidos tantas veces como lo indica su multiplicidad algebraica, son los elementos diagonales de $[A]$. Ahora,

$$\chi_A([A]) = ([A]-\lambda_1\mathbf{1})([A]-\lambda_2\mathbf{1})\cdots([A]-\lambda_n\mathbf{1})\ . \tag{3}$$

Considere los dos primeros factores a la izquierda,

$$([A]-\lambda_1\mathbf{1})([A]-\lambda_2\mathbf{1})$$

$$= \begin{bmatrix} 0 & * & * & \cdots & * \\ 0 & \lambda_2-\lambda_1 & * & \cdots & * \\ 0 & 0 & \lambda_3-\lambda_1 & \cdots & * \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & \lambda_n-\lambda_1 \end{bmatrix} \times$$

$$\begin{bmatrix} \lambda_1-\lambda_2 & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & \lambda_3-\lambda_2 & \cdots & * \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & \lambda_n-\lambda_2 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & (\lambda_3-\lambda_1)(\lambda_3-\lambda_2) & \cdots & * \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & (\lambda_n-\lambda_1)(\lambda_n-\lambda_2) \end{bmatrix}.$$

Luego, para los primeros tres factores

$$([A]-\lambda_1\mathbf{1})([A]-\lambda_2\mathbf{1})([A]-\lambda_3\mathbf{1})$$

$$= \begin{bmatrix} 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & (\lambda_3-\lambda_1)(\lambda_3-\lambda_2) & \cdots & * \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & (\lambda_n-\lambda_1)(\lambda_n-\lambda_2) \end{bmatrix}$$

$$\times \begin{bmatrix} \lambda_1-\lambda_3 & * & * & \cdots & * \\ 0 & \lambda_2-\lambda_3 & * & \cdots & * \\ 0 & 0 & 0 & \cdots & * \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & (\lambda_n-\lambda_3) \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & \cdots & * \\ 0 & 0 & 0 & \cdots & * \\ 0 & 0 & 0 & \cdots & * \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & (\lambda_n-\lambda_1)(\lambda_n-\lambda_2)(\lambda_n-\lambda_3) \end{bmatrix}.$$

En cada multiplicación generamos una nueva fila cero. Si esto continua, terminariamos con la matriz 0.

Lema 2 *Si C y D son matrices cuadradas de dimensión $n \times n$ que son ambas de forma triangular superior y, además $C_{j,k} = 0$ para $1 \leq j, k \leq m < n$ y $D_{m+1,m+1} = 0$, entonces $(CD)_{j,k} = 0$ para $1 \leq j, k \leq m + 1$.*

Demostración: Ya que el producto de matrices de forma triangular superior es nuevamente una matriz triangular superior, basta calcular los elementos $(CD)_{j,k}$ con $m + 1 \geq k \geq j$. Tenemos

$$(CD)_{j,k} = \sum_{\ell=1}^{n} C_{j,\ell} D_{\ell,k} = \sum_{\ell=j}^{n} C_{j,\ell} D_{\ell,k} = \sum_{\ell=j}^{k} C_{j,\ell} D_{\ell,k} \ ;$$

si $j \leq k \leq m$ la hipótesis sobre C nos $C_{j,\ell} = 0$ para todo $1 \leq \ell \leq k$, y por ende $(CD)_{j,k} = 0$. Si $j \leq k = m + 1$ entonces la hipótesis sobre C nos da $C_{j,\ell} = 0$ para $1 \leq \ell \leq m$ y la hipótesis sobre D nos da $D_{m+1,m+1} = 0$; luego nuevamente $(CD)_{j,k} = 0$. $\square$

Este lema y el Teorema de Triangularización nos permiten demostrar el

Teorema 3.2 *(Teorema de Cayley-Hamilton) Para todo operador A sobre un espacio vectorial complejo, $\chi_A(A) = 0$.*

Demostración: La demostración ya fué comenzada en las consideraciones previas al lema. Si $n = 1$ no hay nada que demostrar. Suponga entonces que $n \geq 2$. Podemos suponer que $[A]$ es triangular superior. Los dos primeros factores de (3) satisfacen las condiciones del Lema con $m = 1$, lo cual nos dice que $([A] - \lambda_1 \mathbf{1})([A] - \lambda_2 \mathbf{1})$ es triangular superior con las dos primeras columnas nulas (como vimos). Con $C = ([A] - \lambda_1 \mathbf{1})([A] - \lambda_2 \mathbf{1})$, $D = ([A] - \lambda_3 \mathbf{1})$ y $m = 2$ se satisfacen las condiciones del Lema, luego $C_3 := ([A] - \lambda_1 \mathbf{1})([A] - \lambda_2 \mathbf{1})([A] - \lambda_3 \mathbf{1})$ tiene las primeras 3 columnas nulas. Repitiendo este procedimiento, en el paso k, $C_k := ([A] - \lambda_1 \mathbf{1})([A] - \lambda_2 \mathbf{1}) \cdots ([A] - \lambda_k \mathbf{1})$ es triangular superior con las primeras k columnas nulas y $(A - \lambda_{k+1} \mathbf{1})$ es triangular superior con el elemento diagonal $k+1$ nulo, por lo cual el Lema da que C_{k+1} es triangular superior con las primeras $k + 1$ columnas nulas. $\square$

3.1. Reducción

Un par (E, F) de subespacios complementarios de $\mathcal{H}$ **reduce** al operador A, si E y F son ambos invariantes para A. En tal caso, si $h = e \oplus f$ con $e \in E$ y $f \in F$, se tendrá $Ah = Ae + Af$ con $Ae \in E$ y $Af \in F$. Definiendo A_E y A_F respectivamente por $A_E e := Ae$ y $A_F f := Af$, tendremos $Ah = A_E e + A_F f$ y, recordando que la descomposición de un vector arbitrario de $\mathcal{H}$ en sus componentes en E y en F respectivamente es únivoca, tenemos $A = A_E \oplus A_F$. Esto nos permite reducir el análisis de A al de A_E y A_F.
Veamos un ejemplo simple de un operador que no se puede reducir. En $\mathbb{C}^2$, sea

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} . \tag{4}$$

Es inmediato verificar que los únicos subespacios invariantes para A son $\{0\}$, $\mathcal{H}$ y el autoespacio al autovalor 0 (el núcleo),

$$\mathcal{E}_0 = \left\{ \begin{bmatrix} z \\ 0 \end{bmatrix} : \ z \in \mathbb{C} \right\} ,$$

que coincide con $\mathcal{R}_A$ (!). Cualquier subespacio unidimensional de $\mathcal{H}$ complementario a $\mathcal{E}_0$ no es invariante; por lo tanto no hay forma de reducir a A.

3.2. Los núcleos de A^p

Si A es un operador, y $f \in \mathcal{H}$ es tal que $A^p f = 0$, entonces obviamente $A^k f = 0$, para todo $k \geq p$; luego $ker(A^p) \subset ker(A^k)$. Pero, ya que la dimensión de $\mathcal{H}$ es finita, la dimensión de $ker(A^k)$ no puede crecer indefinidamente. Además

Proposición 3.1 *Si A es un operador y $\dim(ker(A^p)) = \dim(ker(A^{p+1}))$ entonces $ker(A^m) = ker(A^p)$ para todo $m \geq p$.*

Demostración: De $ker(A^p) \subset ker(A^{p+1})$ se deduce que si $\dim(ker(A^p)) = \dim(ker(A^{p+1}))$, entonces $ker(A^p) = ker(A^{p+1})$. Suponga que $f \in ker(A^{p+2})$; entonces $0 = A^{p+2}f = A^{p+1}(Af)$ lo cual implica que $Af \in ker(A^{p+1}) = ker(A^p)$ o sea que $0 = A^p Af = A^{p+1}f$ y $f \in ker(A^{p+1}) = ker(A^p)$.

Si A es inyectivo, el **índice** de A es 0; en caso contrario, $ker(A) \neq \{0\}$, y el índice de A es el mínimo número natural k tal que $ker(A^{k+1}) = ker(A^k)$ pero $ker(A^{k-1}) \neq ker(A^k)$. Denominamos con $\iota(A)$ el índice del operador A.

Proposición 3.2 *Si A es un operador entonces $ker(A^n) = ker(A^{n+1})$, y $\iota(A) \leq n$.*

Demostración: Sea k el índice de A y suponga que $k > n$. Entonces, para $j = 0, 1, 2, \cdots, n$ se tiene $ker(A^j) \subset ker(A^{j+1})$ pero $ker(A^j) \neq ker(A^{j+1})$ y, por lo tanto, $\dim(ker(A^{j+1}) \geq \dim(ker(A^j)) + 1$. Entonces, con $\dim(ker(A^0)) = 0$, $\dim(ker(A^{n+1})) \geq n + 1$ y esto es imposible ya que $\dim(\mathcal{H}) = n$. Luego, $k \leq n$, y por la proposición anterior, $ker(A^n) = ker(A^{n+1})$. □

3.3. Invertible $\oplus$ nilpotente

Un operador sobre $\mathcal{H}$ se dice **nilpotente** si alguna potencia de él se anula. El ejemplo canónico es (4). Claramente, un autovalor de un operador nilpotente es necesariamente nulo.

Proposición 3.3 *Si A es un operador nilpotente entonces su índice es igual al menor número natural m tal que $A^m = 0$ pero $A^{m-1} \neq 0$. Se tiene $A^n = 0$.*

Demostración: Si A es inyectivo ($\iota(A) = 0$) entonces no es nilpotente. Sea m el mínimo natural con $A^m = 0$, y $A^{m-1} \neq 0$; entonces $m \geq 1$. La desigualdad $k := \iota(A) \leq m$ se desprende de la definición del índice. Si $k < m$ entonces la proposición 3.1 implica $ker(A^k) = ker(A^m) = \mathcal{H}$ en cuyo caso m no es minimal. Por lo tanto $k = m$ y la afirmación $A^n = 0$ se desprende de la proposición 3.2. □

Teorema 3.3 *Si A es un operador de índice k los subespacios $E := ker(A^k)$ y $F := \mathcal{R}_{A^k}$ son complementarios y reducen a A. La restricción de A a F es inyectiva, mientras que la restricción A_E de A a E es nilpotente con $(A_E)^k = 0$. Esta reducción de A es unívoca.*

Demostración: Si A es inyectivo, entonces $F = \mathcal{H}$ y $E = \{0\}$ y no hay nada más que decir. Suponga que $k \geq 1$ y $f \in E \cap F$. Entonces, $A^k f = 0$ y hay $g \in \mathcal{H}$ con $f = A^k g$. Luego $A^{2k} g = A^k f = 0$, pero por definición del índice, $ker(A^{2k}) = ker(A^k)$, y por ende $g \in ker(A^k)$ con lo cual $f = A^k g = 0$. Por otro lado, cualquiera sea $x \in \mathcal{H}$, tenemos que $x = u + v$ con $u \in E$ y $v \in F$. Esto demuestra que $\mathcal{H} = E \oplus F$.
Si $x \in E$ entonces $A^k(Ax) = A^{k+1}x = 0$ con lo cual $A(E) \subset E$ y E es invariante y, además, $A^k x = 0$ por lo cual $(A_E)^k = 0$. Si $y \in F$, entonces $y = A^k z$ con $z \in \mathcal{H}$ y $Ay = A^k(Az) \in F$, o sea que F es invariante. Además, si $0 = Ay = A(A^k z) = A^{k+1} z$, entonces $z \in ker(A^{k+1}) = ker(A^k)$, o sea $y = A^k z = 0$. La restricción A_F es por ende inyectiva.

3.4. Los subespacios espectrales $ker((A - \lambda\mathbf{1})^n)$

Para un operador A definimos $\mathcal{H}(\lambda) := ker((A - \lambda\mathbf{1})^n)$, para $\lambda \in \sigma(A)$. Aquí $\lambda \in \sigma(A)$ nos garantiza que $H(\lambda)$ no es trivial ya que $\{0\} \neq \mathcal{E}_\lambda \subset \mathcal{H}(\lambda)$. Estos son los llamados **subespacios espectrales**, por lo que sigue. Demostramos que cada subespacio espectral es invariante para A y que $\mathcal{H}$ es la suma directa de los espacios espectrales de A. Esto nos entrega el Teorema Espectral. Conviene tener en mente la situación que se presenta para el operador (4), donde hay un sólo espacio espectral.

Teorema 3.4 *Para todo $\lambda \in \sigma(A)$, $\mathcal{H}(\lambda)$ es invariante para A y para la restricción A_λ de A a este subespacio, se tiene $A_\lambda = \lambda\mathbf{1} + D_\lambda$ donde D_λ es un operador nilpotente. La dimensión de $\mathcal{H}(\lambda)$ es igual a la multiplicidad de λ como raíz del polinomio característico de A.*

Demostración: Sea k el índice de $B := A - \lambda\mathbf{1}$, entonces $H(\lambda) = ker(B^k)$. Por la proposición 3.3, este subespacio es invariante para B y por ende para A; además $D_\lambda = A_\lambda - \lambda\mathbf{1}$ es nilpotente con lo cual λ es el único autovalor de A_λ. También $\mathcal{R}_B$ es invariante para B y por ende para A; y λ no es autovalor de la restricción C de A a este subespacio. Ya que

$$A = A_\lambda \oplus C$$

y $\chi_A(z) = det(A_\lambda - z\mathbf{1}) det(C - z\mathbf{1})$; y $det(C - z\mathbf{1})$ no tiene a λ entre sus raices, la multiplicidad de λ como raíz de χ_A es igual a aquella como raíz de χ_{A_λ}. La demostración se completa observando que λ es el único autovalor de A_λ lo que implica que λ es la única raız de χ_{A_λ} y, por el Teorema Fundamental del Álgebra, su multiplidad es la dimensión de $\mathcal{H}(\lambda)$. $\square$

La **multiplicidad algebraica** de $\lambda \in \sigma(A)$ es la multiplicidad de λ como raíz de χ_A, o como acabamos de ver, la dimensión de $ker((A - \lambda\mathbf{1})^k)$ donde k es el índice de A. Esta dimensión es también aquella de $ker((A - \lambda\mathbf{1})^n)$.

Proposición 3.4 *Si $\lambda, \mu \in \sigma(A)$ son distintos, entonces $\mathcal{H}(\lambda) \cap \mathcal{H}(\mu) = \{0\}$.*

Demostración: Suponga que $f \in \mathcal{H}(\lambda) \cap \mathcal{H}(\mu)$ no es cero. Sea p el menor número natural tal que $(A - \lambda\mathbf{1})^p f = 0$ y $(A - \lambda\mathbf{1})^{p-1} f \neq 0$. Entonces $n \geq p \geq 1$, y

$$0 = (A - \mu\mathbf{1})^n f = (\{A - \lambda\mathbf{1}\} + (\lambda - \mu)\mathbf{1})^n f$$

$$= \sum_{j=0}^{n} \binom{n}{j} (\lambda - \mu)^{n-j} (A - \lambda \mathbf{1})^j f$$

$$= \sum_{j=0}^{p-1} \binom{n}{j} (\lambda - \mu)^{n-j} (A - \lambda \mathbf{1})^j f \ .$$

Operando con $(A - \lambda \mathbf{1})^{p-1}$ a ambos lados de la igualdad, obtenemos

$$0 = \sum_{j=0}^{p-1} \binom{n}{j} (\lambda - \mu)^{n-j} (A - \lambda \mathbf{1})^{j+p-1} f = (\lambda - \mu)^n (A - \lambda \mathbf{1})^{p-1} f \ ;$$

y, como $\lambda \neq \mu$, se tiene $(A - \lambda \mathbf{1})^{p-1} f = 0$, lo que contradice la definición de p. Por lo tanto, la suposición $f \neq 0$ es falsa.

Teorema 3.5 *Si $\mathcal{H}$ es complejo y A es un operador sobre $\mathcal{H}$, entonces $\mathcal{H}$ es la suma directa de los subespacios $\mathcal{H}(\lambda)$, $\lambda \in \sigma(A)$.*

Demostración: Usamos el teorema 3.4. Tomamos $\lambda_1 \in \sigma(A)$ que no es vacio. Entonces $\mathcal{H} = \mathcal{H}_\lambda \oplus \mathcal{R}_1$ donde $\mathcal{R}_1$ es la imagen de $(A - \lambda_1 \mathbf{1})^{k_1}$ siendo $k_1 := \iota(A - \lambda_1 \mathbf{1})$. Ya que $\mathcal{R}_1$ es invariante para A por la proposición 3.3, la restricción A_1 de A a $\mathcal{R}_1$ tiene un autovalor $\lambda_2 \neq \lambda_1$ y $H(\lambda_2) \subset \mathcal{R}_1$. Continuando este procedimiento hasta agotar $\sigma(A)$, tendremos $\sum_{\lambda \in \sigma(A)} \dim(\mathcal{H}_\lambda) = n$, lo que demuestra la afirmación. $\square$

Juntando todo lo expuesto obtenemos el siguiente resultado fundamental:

Teorema 3.6 *(Teorema Espectral) Si $\mathcal{H}$ es complejo y A es un operador sobre $\mathcal{H}$, entonces para cada $\lambda \in \sigma(A)$ hay una proyección P_λ y un operador nilpotente D_λ tales que:*

1. $\mathcal{R}_{P_\lambda} = \mathcal{H}(\lambda)$; $ker(P_\lambda) = \bigoplus_{\lambda \neq \mu \in \sigma(A)} \mathcal{H}(\mu)$;
2. $P_\lambda P_{\lambda'} = 0$ *si* $\lambda \neq \lambda'$;
3. $P_\lambda A = A P_\lambda = P_\lambda A P_\lambda = \lambda P_\lambda + D_\lambda$;
4. $P_\lambda D_\lambda = D_\lambda P_\lambda = D_\lambda$;
5. $\sum_{\lambda \in \sigma(A)} P_\lambda = \mathbf{1}$;
6. $A = \sum_{\lambda \in \sigma(A)} \{\lambda P_\lambda + D_\lambda\}$.

Las siguientes observaciones son relevantes:

- La **descomposición espectral** de A dada por 6., no está escrita como suma directa. Pero, ya que $\oplus \mathcal{H}(\lambda)$ reduce a P_λ y a D_λ, se puede escribir

  $$A = \bigoplus_{\lambda \in \sigma(A)} \{\lambda \mathbf{1} + \widehat{D}_\lambda\} \ ,$$

  donde $\widehat{D}_\lambda$ es la restricción de D_λ a $\mathcal{H}(\lambda)$.

- La proyección P_λ llamada **proyección espectral**, queda determinada por todos los nucleos $ker((A-\mu\mathbf{1})^n)$ y no solamente por el nucleo $ker((A-\lambda\mathbf{1})^n)$, pues el núcleo de P_λ es
$$\bigoplus_{\lambda\neq\mu\in\sigma(A)} \mathcal{H}(\mu)\ .$$
En el caso de A autoadjunto respecto de un producto escalar sobre $\mathcal{H}$, las proyecciones espectrales son autoadjuntas y están determinadas univocamente por $\mathcal{H}(\lambda)$ pues su núcleo es el complemento ortogonal de $\mathcal{H}(\lambda)$.

  Una vez determinada P_λ, se tiene $D_\lambda = AP_\lambda - \lambda P_\lambda$

Ahora tenemos la suficiente información sobre un operador arbitrario para relacionar a su polinomio minimal con la estructura del operador. Dejamos la demostración como ejercicio para el lector.

Teorema 3.7 *La multiplicidad del autovalor λ de A como raíz del polinomio minimal de A es el índice de $A-\lambda\mathbf{1}$.*

3.5. Forma canónica de Jordan

Los elementos constitutivos de un operador son entonces los operadores de la forma
$$\lambda P + D$$
donde $\lambda\in\mathbb{C}$, P es una proyección, y D es un nilpotente con $PD = DP = D$. ¿Cuál es la forma matricial más simple posible para un operador elemental de este tipo? Claramente, eligiendo una base arbitraria de $\mathcal{R}_P$ y completando, también arbitrariamente, a una base de $\mathcal{H}$, la matriz asociada $[P]$ será $\mathbf{1}\oplus 0$:
$$[P] = \begin{bmatrix} \mathbf{1} & 0 \\ 0 & 0 \end{bmatrix}\ .$$
Particionando a D acorde a $\mathcal{H} = \mathcal{R}_P \oplus ker(P)$,
$$[D] = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}\ ;$$
entonces
$$PD - DP = \begin{bmatrix} D_{11} & D_{12} \\ -D_{21} & 0 \end{bmatrix}\ ,$$
y la condición $PD = DP = D$ implica que
$$[D] = \begin{bmatrix} D_{11} & 0 \\ 0 & 0 \end{bmatrix}\ ;$$
La libertad que tenemos en la elección de la base de $\mathcal{R}_P$ nos permitirá establecer una forma canónica (por simple) de $[D]$.

Teorema 3.8 *Si D es un operador nilpotente y $\dim(ker(D)) = m$, entonces $m\geq 1$ y hay:*

1. m números naturales $k_1, \cdots, k_m$ no nulos, tales que $k_1 + k_2 + \cdots + k_m = n$; y $\iota(D) = k_1 \geq k_2 \geq \cdots \geq k_m$;

2. m vectores x_1, x_2, $\cdots$, x_m de $\mathcal{H}$ tales que: (i) k_j es el menor número natural que satisface $D^{k_j} x_j = 0$, para $j = 1, 2, \cdots, m$; y (ii)

$$\{x_1, Dx_1, \cdots, D^{k_1-1}x_1, x_2, Dx_2, \cdots, D^{k_2-1}x_2,$$

$$x_3, \cdots, x_m, Dx_m, \cdots, D^{k_m-1}x_m\}$$

es una base de $\mathcal{H}$.

Si C es un operador nilpotente tal que los números m, k_1, k_2, $\cdots$, k_m asociados a C son idénticos a los de D entonces hay un operador invertible T tal que $T^{-1}CT = D$.

Demostración: La demostración procede por inducción en la dimensión n de $\mathcal{H}$. Si $\mathcal{H}$ es unidimensional entonces $D = 0$, $k = m = 1$ y la afirmación es verdadera. Dado D nilpotente con $\iota(D) = k$, $\mathcal{R}_D$ es invariante para D y la restricción B de D a $\mathcal{R}$ es nilpotente. Ya que D no es inyectivo, la dimensión de $\mathcal{R}_D$ es estrictamente menor a n. Si $y \in \mathcal{R}_D$, entonces $y = Dx$ con $x \in \mathcal{H}$, y $B^{k-1}y = B^{k-1}Dx = D^k x = 0$, con lo cual $p = \iota(B)$ no supera a $k-1 \leq n-1$. Por la hipótesis inductiva hay, $t = \dim(ker(B))$ números naturales $p_1, p_2, \cdots, p_t$ tales que: $p = p_1 \geq p_2 \geq \cdots \geq p_t$; y $p_1 + p_2 + \cdots + p_t = n - m$. Y, hay t vectores $y_1, y_2, \cdots, y_t$ en $\mathcal{R}_D$ tales que

$$\{y_1, Dy_1, \cdots, D^{p_1-1}y_1, y_2, Dy_2, \cdots, D^{p_2-1}y_2,$$

$$y_3, \cdots, y_m, Dy_m, \cdots, D^{p_t-1}y_t\}$$

es una base de $\mathcal{R}_D$; y, para todo $j = 1, 2, \cdots, t$, $D^s y_j = 0$ si y sólo si $s \geq p_j$. Ahora, en general tendremos $\mathcal{R}_D \cap ker(D) \neq \{0\}$, pero sea $\mathcal{K}$ un subespacio de $ker(D)$ que sea complementario a $\mathcal{R}_D \cap ker(D)$ en $ker(D)$

$$ker(D) = \underbrace{(\mathcal{R}_D \cap ker(D))}_{ker(B)} \oplus \mathcal{K}\ .$$

Entonces $\mathcal{H} = \mathcal{R}_D \oplus \mathcal{K}$ y, ya que $D(\mathcal{K}) = \{0\}$, estos subespacios complementarios reducen a D.
Como $y_1, y_2, \cdots, y_t \in \mathcal{R}_D$, hay $x_1, x_2, \cdots, x_t \in \mathcal{H}$ tales que $y_j = Dx_j$, para $j = 1, 2, \cdots, t$. Sea $\{x_{t+1}, \cdots, x_m\}$ una base de $\mathcal{K}$. Se tiene $Dx_j = 0$ para $j \geq t + 1$. Ahora, para $j = 1, 2, \cdots, t$, $D^{s+1}x_j = D^s y_j = 0$ si y sólo si $s \geq p_j$. Definamos

$$k_j := \begin{cases} p_j + 1 & , \quad \text{si } j = 1, 2, \cdots, t \\ 1 & , \quad \text{si } j = t+1, \cdots, m \end{cases} .$$

Entonces, $D^s x_j \neq 0$ si y sólo si $s \leq k_j - 1$, de acuerdo con 2.(i).

Considere

$$\{x_1, Dx_1, \cdots, D^{k_1-1}x_1, x_2, Dx_2, \cdots, D^{k_2-1}x_2,$$

$$x_3, \cdots, x_t, Dx_t, \cdots, D^{k_t-1}x_t, x_{t+1}, x_{t+2}, \cdots, x_m\}\ .$$

Hay aquí $k_1+k_2+\cdots+k_t+(m-t)=(p_1+1)+\cdots+(p_t+1)+(m-t)=p_1+\cdots p_t+m=(n-m)+m=n$ vectores, como se pide en 1. Veamos que son linealmente independientes. Supongamos que

$$0=\sum_{j=1}^{m}\sum_{\ell=0}^{k_j-1}\alpha_{j,\ell}D^{\ell}x_j \ . \tag{5}$$

Multiplicando por D,

$$0=\sum_{j=1}^{m}\sum_{\ell=0}^{k_j-1}\alpha_{j,\ell}D^{\ell}Dx_j=\sum_{j=1}^{t}\sum_{\ell=0}^{k_j-1}\alpha_{j,\ell}D^{\ell}y_j=\sum_{j=1}^{t}\sum_{\ell=0}^{k_j-1}\alpha_{j,\ell}D^{\ell}y_j \ ,$$

y de las propiedades de $y_1,y_2,\cdots,y_t$, deducimos que $a_{j,\ell}=0$ para $j=1,2,\cdots,t$ y $0\le\ell\le k_j-2=p_j-1$. Luego, (5) se reduce a

$$0=\left(\sum_{j=1}^{t}\alpha_{j,k_j-1}D^{k_j-1}x_j\right)+\sum_{j=t+1}^{m}a_{j,0}x_j \ .$$

Aquí la primera suma está en $\mathcal{R}_D\cap ker(D)$ y la segunda en $\mathcal{K}$. Por lo tanto, cada suma debe anularse por separado. Entonces

$$0=\sum_{j=1}^{t}\alpha_{j,k_j-1}D^{k_j-1}x_j=\sum_{j=1}^{t}\alpha_{j,k_j-1}D^{p_j-1}y_j \ ;$$

pero $\{D^{p_1-1}y_1,D^{p_2-1}y_2,\cdots,D^{p_t-1}y_t\}$ es una base de $ker(B)$ con lo cual $\alpha_{j,k_j-1}=0$ para $j=1,2,\cdots,t$. También,

$$0=\sum_{j=t+1}^{m}a_{j,0}x_j$$

implica que $a_{j,0}=0$ para $j=t+1,t+2,\cdots,m$ ya que $\{x_{t+1},\cdots,x_m\}$ es una base de $\mathcal{K}$. Hemos demostrado que el conjunto de los x_j con las potencias $D^sx_j\neq0$ forman una base de $\mathcal{H}$, de acuerdo con lo que se pide en 2.(ii). Observando que $k_1\ge k_2\ge\cdots\ge k_t\ge 1=k_{t+1}=k_{t+2}=\cdots=k_m$, y que $D^{k_1}x_j=0$ para $j=1,2,\cdots,m$, pero $D^{k_1-1}x_1\neq0$, deducimos que $k_1=k$ es el índice de D.

Si C es nilpotente y tiene $(m,k_1,\cdots,k_m)$ idéntico al correspondiente vector de D, considere los vectores $\{v_1,v_2,\cdots v_m\}$ asociados con C. Definiendo, $T\sum_{j=1}^{m}\sum_{\ell=0}^{m_j}\alpha_{j,\ell}D^{\ell}x_j=\sum_{j=1}^{m}\sum_{\ell=0}^{m_j}\alpha_{j,\ell}D^{\ell}v_j$, T es invertible y $T^{-1}CT=D$. Esto completa la demostración. □

Teorema 3.9 *Si el operador nilpotente D está caracterizado por los números m, k_1, k_2, $\cdots$, k_m, entonces para $p=1,2,\cdots$,*

$$\dim(ker(D^p))-\dim(ker(D^{p-1}))=\begin{cases}0 & , \ si\ k_1<p\\ \max\{j\in\{1,2,\cdots,m\}:k_j\ge p\} & , \ si\ k_1\ge p\end{cases}$$

o sea la cantidad de valores de k_j que no son menores que p. Entonces,

$$\begin{aligned}&\dim(ker(D^p))-\dim(ker(D^{p-1}))-\big(\dim(ker(D^{p+1}))-\dim(ker(D^p))\big)\\&=2\dim(ker(D^p))-\dim(ker(D^{p+1}))-\dim(ker(D^{p-1}))\end{aligned}$$

es la cantidad de valores de k_j que son iguales a p.

Demostración: Sean $f_{1j} = D^{k_1-j}x_1$ para $j = 1, 2, \cdots, k_1$; $f_{2j} = D^{k_2-j}x_2$, $j = 1, 2, \cdots, k_2$; etc., o sea $f_{\ell j} = D^{k_\ell - j}x_\ell$, $j = 1, 2, \cdots, k_\ell$ para $\ell = 1, 2, \cdots, m$. Donde x_ℓ son los m vectores del teorema anterior. Respecto de la base $\{f_{11}, f_{12}, \cdots, f_{1k_1}, f_{21}, \cdots, f_{2k_2}, \cdots, f_{m,k_m}\}$, la matriz de D es suma directa de m bloques $[D_\ell]$ de dimensión $k_\ell \times k_\ell$

$$[D_\ell] = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

cuyos únicos elementos no nulos son los $k_\ell - 1$, unos en la primera diagonal superior. Es inmediato verificar que

$$\dim(ker([D_\ell]^p)) = \begin{cases} p & , \quad \text{si } p \le k_\ell \\ k_\ell & , \quad \text{si } p > k_\ell \end{cases} .$$

Entonces

$$\dim(ker(D^p)) = \sum_{\ell=1}^{m} \dim(ker([D_\ell]^p)) = \sum_{\ell=1;\, k_\ell \ge p}^{m} p + \sum_{\ell=1;\, k_\ell < p}^{m} k_\ell .$$

Luego,

$$\begin{aligned} \dim(ker(D^p)) - \dim(ker(D^{p-1})) &= p \left(\sum_{\ell=1;\, k_\ell \ge p}^{m} \right) + \sum_{\ell=1;\, k_\ell < p}^{m} k_\ell \\ &-(p-1) \left(\sum_{\ell=1;\, k_\ell \ge p-1}^{m} \right) - \sum_{\ell=1;\, k_\ell < p-1}^{m} k_\ell \\ &= p \left(\sum_{\ell=1;\, k_\ell \ge p}^{m} \right) + \sum_{\ell=1;\, k_\ell < p-1}^{m} k_\ell + \sum_{\ell=1;\, k_\ell = p-1}^{m} k_\ell \\ -(p-1) \left(\sum_{\ell=1;\, k_\ell = p-1}^{m} \right) - (p-1) \left(\sum_{\ell=1;\, k_\ell \ge p}^{m} \right) &- \sum_{\ell=1;\, k_\ell < p-1}^{m} k_\ell = \sum_{\ell=1;\, k_\ell \ge p}^{m} . \end{aligned}$$

Teniendo en cuenta que $k_1 \ge k_2 \ge \cdots \ge k_m$, esto demuestra la primera afirmación. La segunda es consecuencia inmediata. □

El Teorema Espectral junto al análisis de los operadores nilpotentes conducen a la forma canónica de Jordan.

Para cualquier $k \in \mathbb{N}$, y $\lambda \in \mathbb{C}$, $J_k(\lambda)$ es la **matriz de Jordan** $k \times k$ dada por

(6)
$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix} .$$

Teorema 3.10 *(Forma canónica de Jordan[4]) Sea $\mathcal{H}$ complejo y sea A un operador. Denotese con $m(\lambda)$ y $g(\lambda)$ la multiplicidad algebraica, respectivamente geométrica, del autovalor λ de A. Hay una base B de $\mathcal{H}$, tal que*

$$[A] = \bigoplus_{\lambda\in\sigma(A)} B(\lambda) \ ,$$

donde, para cada $\lambda \in \sigma(A)$, $B(\lambda)$ es una matriz cuadrada $m(\lambda)\times m(\lambda)$ compuesta de $g(\lambda)$ matrices de Jordan $J_{d_1(\lambda)}(\lambda)$, $J_{d_2(\lambda)}(\lambda)$, $\cdots$, $J_{d_{g(\lambda)}(\lambda)}(\lambda)$

$$B(\lambda) = \bigoplus_{j=1}^{g(\lambda)} J_{d_j(\lambda)}(\lambda) \ .$$

El número

$$(7) \qquad 2\dim(ker((A-\lambda\mathbf{1})^m)) - \dim(ker((A-\lambda\mathbf{1})^{m-1})) - \dim(ker((A-\lambda\mathbf{1})^{m+1}))$$

es no negativo para todo $m \in \mathbb{N}$ y $\lambda \in \sigma(A)$, e igual a la cantidad de matrices $J_m(\lambda)$ que aparecen en $B(\lambda)$.
Esto determina las dimensiones, $d_j(\lambda)$, de las matrices de Jordan que aparecen en el bloque $B(\lambda)$ asociado con λ, y la cantidad de veces que aparecen en el bloque. La matriz de Jordan de dimensión maximal tiene dimensión igual al índice de $A-\lambda\mathbf{1}$ restringido a $\mathcal{H}(\lambda)$.

Si las multiplicidades algebraicas de los autovalores de una matriz A son a lo sumo 3, la forma canónica de Jordan se obtiene a partir de las multiplicidades algebraicas y geométricas solamente; se puede prescindir de la formula (7). Si en cambio A tiene un autovalor de multiplicidada algebraica $m \geq 4$ y la multiplicidad geométrica g correspondiente satisface $1 < g < m-1$, no podemos prescindir de (7). El siguiente ejemplo sirve para ilustrar el procedimiento a seguir y las afirmaciones hechas.

Considere, para $a \in \mathbb{C}$ las cuatro matrices

$$A = \begin{bmatrix} a & 0 & 0 & -1 \\ -1 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & -1 & a \end{bmatrix} \ , \ B = \begin{bmatrix} a & 0 & 0 & -1 \\ -1 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{bmatrix} \ ,$$

$$C = \begin{bmatrix} a & 0 & 0 & 0 \\ -1 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & -1 & a \end{bmatrix} \ , \ D = \begin{bmatrix} a & 0 & 0 & 0 \\ -1 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{bmatrix} \ .$$

Si en estas cuatro matrices permutamos la primera columna con la segunda columna, la primera fila con la segunda fila y, luego, la tercera columna con la cuarta columna y la tercera fila con la cuarta fila, obtenemos –en los cuatro casos– matrices triangulares superiores cuyos elemntos diagonales son siempre a. El polinomio característico de estas matrices es entonces $(\zeta - a)^4$, con lo cual a es el único autovalor y tiene multiplicidad algebraica 4. Calculamos

$$A - a\mathbf{1} = \begin{bmatrix} 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} \ , \ B - a\mathbf{1} = \begin{bmatrix} 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \ ,$$

[4]La primera demostración se debe a M.E. Camille Jordan (1838-1922) y es del año 1870.

$$C - a\mathbf{1} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} , \quad D - a\mathbf{1} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} .$$

Usando el teorema 1.7, obtenemos

$$dim(ker(A - a\mathbf{1})) = 1 \ , \ \ dim(ker(B - a\mathbf{1})) = 2 \ ,$$

$$dim(ker(C - a\mathbf{1})) = 2 \ , \ \ dim(ker(D - a\mathbf{1})) = 3 \ .$$

Calculamos

$$(A - a\mathbf{1})^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} , \quad (B - a\mathbf{1})^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} ,$$

$$(C - a\mathbf{1})^2 = (D - a\mathbf{1})^2 = 0 \ .$$

Luego,

$$dim(ker(A - a\mathbf{1})^2) = 2 \ , \ \ dim(ker(B - a\mathbf{1})^2) = 3 \ ,$$

$$dim(ker(C - a\mathbf{1})^2) = dim(ker(D - a\mathbf{1})^2) = 4 \ .$$

Siguiendo,

$$(A - a\mathbf{1})^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} , \quad (B - a\mathbf{1})^2 = 0 \ ;$$

$$dim(ker(A - a\mathbf{1})^3) = 3 \ , \ \ dim(ker(B - a\mathbf{1})^2) = 4 \ .$$

Finalmente,

$$(A - a\mathbf{1})^4 = 0 \ , \ \ dim(ker(A - a\mathbf{1})^4) = 4 \ .$$

Podemos entonces con la fórmula (7) determinar las formas canónicas de Jordan de A, B, C, y D como sigue. Abreviamos $d(X, m) := dim(ker(X - a\mathbf{1})^m)$, con $X = A, B, C, D$.

m	$d(A,m)$	$d(B,m)$	$d(C,m)$	$d(D,m)$
0	0	0	0	0
1	1	2	2	3
2	2	3	4	4
3	3	4	4	4
≥ 4	4	4	4	4

La forma canónica de Jordan de A tiene entonces:

- $2d(A,1) - d(A,0) - d(A,2) = 0$; o sea ningún bloque $J_1(a)$;
- $2d(A,2) - d(A,1) - d(A,3) = 4 - 1 - 3 = 0$; o sea, ningún bloque $J_2(a)$;
- $2d(A,3) - d(a,2) - d(A,4) = 6 - 2 - 4 = 0$; o sea ningún bloque $J_3(a)$;
- $2d(A,4) - d(A,3) - d(A,5) = 8 - 3 - 4 = 1$; o sea un bloque $J_4(a)$.

Vale decir que la forma canónica de Jordan de A es $J_4(a)$ y la multiplicidad geométrica del autovalor a es 1.
Analogamente,

$$2d(B,1) - d(B,0) - d(B,2) = 1 \ , \ \ 2d(B,2) - d(B,1) - d(B,3) = 0 \ ,$$
$$2d(B,3) - d(B,2) - d(B,4) = 1 \ ;$$

y por lo tanto la forma canónica de Jordan de B es

$$J_1(a) \oplus J_3(a)$$

y la multiplicidad geométrica de a es 2.
Similarmente,

$$2d(C,1) - d(C,0) - d(C,2) = 0 \ , \ \ 2d(C,2) - d(C,1) - d(C,3) = 2 \ ,$$

con lo cual

$$J_2(a) \oplus J_2(a)$$

es la forma canónica de Jordan de C y 2 es la multiplicidad geométrica de a. Por último,

$$2d(D,1) - d(D,0) - d(D,2) = 2 \ , \ \ 2d(D,2) - d(D,1) - d(D,3) = 1 \ ,$$

y la forma canónica de Jordan de D es:

$$J_1(a) \oplus J_1(a) \oplus J_2(a) \ ,$$

cuyo autovalor a tiene multiplicidad geométrica 3. Observese que tanto B como C tienen multiplicidad geométrica 2 de su autovalor y la aplicación de (7) (y solamente esta) distingue los dos casos posibles:

$$J_1(a) \oplus J_3(a) \ , \ \ J_2(a) \oplus J_2(a) \ .$$

El último caso en dimensión 4 con multiplicidad 4 del autovalor es aquel donde la multiplicidad geométrica es 4 en cuyo caso la matriz es diagonalizable y la forma canónica de Jordan es $J_1(a) \oplus J_1(a) \oplus J_1(a) \oplus J_1(a)$.

4. Operadores sobre espacios vectoriales con producto escalar

En ese caso, anotando con $\langle \cdot, \cdot \rangle$ al producto escalar que será lineal en la segunda componente cuando $\mathcal{H}$ es complejo, $\mathcal{H}$ tiene la rica estructura geométrica de un espacio euclideo. Escribimos $\|\cdot\|$ para la norma asociada al producto escalar; i.e., $\|f\| := \sqrt{\langle f, f\rangle}$. Decimos que $\mathcal{M} \subset \mathcal{H}$ es **ortogonal** a $\mathcal{N} \subset \mathcal{H}$, y anotamos $\mathcal{M} \perp \mathcal{N}$, si $\langle f, g\rangle = 0$ para todo $f \in \mathcal{M}$ y todo $g \in \mathcal{N}$. Si $\mathcal{M} \subset \mathcal{H}$, entonces el **complemento ortogonal** de $\mathcal{M}$ es $\mathcal{M}^\perp := \{f \in \mathcal{H} : \ \langle f, g\rangle = 0 \ , \ \ \text{para todo } g \in \mathcal{M}\}$. Una **base ortonormal** es una base cuyos vectores tienen todos norma 1 y son dos-a-dos ortogonales.

Recordamos que a todo operador A sobre $\mathcal{H}$ está asociado su **operador adjunto**, anotado A^* y definido por

$$\langle g, A^* f\rangle = \langle Ag, f\rangle \ , \ \ \text{para todo } f, g \in \mathcal{H} \ .$$

Un operador U sobre $\mathcal{H}$ que satisface cualquiera de las siguientes condiciones equivalentes[5]

[5] Aquí se usa la finitud de la dimensión de $\mathcal{H}$.

1. U es isométrico, i.e., $\|Uf\| = f$ para todo $f \in \mathcal{H}$;
2. $\langle Uf, Ug\rangle = \langle f, g\rangle$, para todo $f, g \in \mathcal{H}$;
3. $U^*U = \mathbf{1}$;
4. $UU^* = \mathbf{1}$;
5. U es invertible y $U^{-1} = U^*$;

se dice **unitario**. Claramente, los operadores unitarios y sólo ellos, preservan toda la estructura geométrica euclidea de $\mathcal{H}$. Si B es una base ortonormal de $\mathcal{H}$ y T es invertible, entonces $T(B)$ es una base ortonormal si y sólo si T es unitario.

Teorema 4.1 *(Teorema de triangularización unitaria de Schur) Si $\mathcal{H}$ es complejo y A es un operador sobre $\mathcal{H}$, existe una base ortonormal de $\mathcal{H}$ tal que $[A]$ es triangular superior.*

Demostración: Ya que A tiene un autovalor λ_1, hay un autovector $x \in \mathcal{H}$ correspondiente. Sea $x_1 = \|x\|^{-1}x$ y $\{x_1, y_2, \cdots, y_n\}$ una base ortonormal de $\mathcal{H}$. Particionando a la matriz de A en esta base de acuerdo a $\mathcal{H} = \mathbb{C} \cdot x_1 \oplus \{x_1\}^\perp$, tenemos

$$[A] = \begin{bmatrix} \lambda_1 & * \\ 0 & [A]_1 \end{bmatrix},$$

donde $[A]_1$ es una matriz $(n-1) \times (n-1)$. Pero $[A]_1$ tiene un autovalor λ_2 con correspondiente autvector normalizado $x_2 \in \{x_1\}^\perp$. Si $\{x_1, x_2, y_3, \cdots, y_n\}$ es una base ortonormal de $\mathcal{H}$ y particionamos la matriz de A en esta base de acuerdo a $\mathbb{C} \cdot x_1 \oplus \mathbb{C} \cdot x_2 \oplus \{x_1, x_2\}^\perp$, tendremos

$$[A] = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & [A]_2 \end{bmatrix},$$

donde la matriz $[A]_2$ es $(n-2)\times(n-2)$. Repitiendo este procedimiento, en la $(n-1)$-ésima repetición, producimos una matriz $[A]_{n-1}$ de dimensión 1×1 y hemos obtenido una base ortonormal de $\mathcal{H}$ tal que $[A]$ es triangular superior. $\square$

Recuerde que un operador A se dice **normal** si $A^*A = AA^*$.

Lema 3 *Si la matriz cuadrada A es triangular superior entonces es normal si y sólo si es diagonal.*

Demostración: La afirmación es correcta si la dimensión de $\mathcal{H}$ es 1. Suponemos que $n \geq 2$. La triangularidad superior de A es $A_{j,k} = 0$ si $j < k$. En tal caso $(A^*)_{j,k} = \overline{A_{k,j}} = 0$ si $k < j$, o sea que A^* es triangular inferior. Entonces

$$\sum_{\ell=\max\{j,k\}}^{n} \overline{A_{\ell,j}} A_{\ell,k} = \sum_{\ell=1}^{n} \overline{A_{\ell,j}} A_{\ell,k} = (A^*A)_{j,k}$$

$$= (AA^*)_{j,k} = \sum_{\ell=1}^{n} A_{j,\ell}\overline{A_{k,\ell}} = \sum_{\ell=1}^{\min\{j,k\}} A_{j,\ell}\overline{A_{k,\ell}} \,.$$

En particular, para $j = k$ obtenemos luego de restar $|A_{j,j}|^2$, que

$$\sum_{\ell=j+1}^{n} |A_{\ell,j}|^2 = \sum_{\ell=1}^{j-1} |A_{j,\ell}|^2 \ ,$$

sobrentendiendo que cuando $j = n$ la suma de la izquierda no aparece, y cuando $j = 1$ la suma de la derecha tampoco. Para $j = n$ obtenemos

$$0 = \sum_{\ell=1}^{n-1} |A_{n,\ell}|^2 \ ,$$

vale decir que $A_{n,\ell} = 0$ si $\ell \neq n$; con $j = n-1$, y lo recién visto, obtenemos

$$0 = |A_{n,n-1}|^2 = \sum_{\ell=1}^{n-2} |A_{n-1,\ell}|^2 \ ;$$

luego $A_{n-1,\ell} = 0$ si $\ell \neq n-1$. Prosiguiendo hasta agotar las filas de A obtenemos que A es diagonal si es normal. La afirmación reciproca es trivialmente cierta. $\square$

Si A es normal y U es unitario entonces $(U^*AU)^*(U^*AU) = U^*A^*UU^*AU = U^*A^*AU$ $= U^*AA^*U = U^*AUU^*A^*U = (U^*AU)(U^*AU)^*$; o sea que U^*AU es normal. Esta observación y los dos resultados anteriores demuestran

Teorema 4.2 *(Teorema espectral para operadores normales) Si A es un operador sobre un espacio vectorial complejo $\mathcal{H}$ entonces existe una base ortonormal de $\mathcal{H}$ que consiste de autovectores de A, i.e. $[A]$ es diagonal, si y sólo si A es normal.*

Como corolario inmediato de este resultado se obtiene

Teorema 4.3 *Si A es un operador normal sobre un espcio vectorial complejo entonces:*

1. *Para todo $\lambda \in \sigma(A)$ la multiplicidad geométrica y la multiplicidad algebraica de λ son iguales;*
2. *Los autoespacios a autovalores distintos son ortogonales;*
3. *Los proyectores espectrales a los subespacios espectrales son ortoproyectores;*
4. *$A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda$*

Si bien no hay nada que objetar a la demostración de este resultado via el teorema de triangularización unitaria de Schur, esta no permite ver la rica estructura geométrica subyacente. Por este motivo, y porque nos servirá en el caso en el cual $\mathcal{H}$ es real, damos a continuación una demostración que reposa sobre el Teorema Espectral 3.6. Para esto analizamos en profundidad las consecuencias de la normalidad de un operador.

4.1. Operadores normales. Estructura.

Proposición 4.1 *Para todo operador A sobre $\mathcal{H}$ se tiene $(A^*)^* = A$, y $ker(A^*)$ es el complemento ortogonal de $\mathcal{R}_A$.*

Demostración: La propiedad $(A^*)^* = A$ es consecuencia de la definición. Si $f \in ker(A^*)$, entonces $\langle f, Ag\rangle = \langle A^* f, g\rangle = 0$ para todo $g \in \mathcal{H}$, con lo cual f es ortogonal a la imagen de A, o sea $ker(A^*) \subset \mathcal{R}_A^{\perp}$. Si $f \in \mathcal{R}_A^{\perp}$ entonces $0 = \langle f, Ag\rangle = \langle A^* f, g\rangle$ para todo $g \in \mathcal{H}$, de donde $A^* f = 0$; lo que demuestra la otra inclusión. $\square$

Proposición 4.2 *Si $\mathcal{M}$ es un subespacio de $\mathcal{H}$ y $\mathcal{N}$ es su complemento ortogonal, entonces $\mathcal{M}$ es invariante para A si y sólo si $\mathcal{N}$ es invariante para A^*.*

Demostración: Si $Af \in \mathcal{M}$ para todo $f \in \mathcal{M}$ y $g \in \mathcal{N}$, entonces $\langle A^* g, f\rangle = \langle g, Af\rangle = 0$ para todo $f \in \mathcal{M}$ con lo cual $A^* g \in \mathcal{N}$. Si, reciprocamente, $Ag \in \mathcal{N}$ para todo $g \in \mathcal{N}$, entonces por lo recién demostrado, $Af = (A^*)^* f \in \mathcal{N}^{\perp} = \mathcal{M}$ para todo $f \in \mathcal{M}$. $\square$

Proposición 4.3 *El operador A sobre $\mathcal{H}$ es normal si y sólo si $\|Af\| = \|A^* f\|$ para todo $f \in \mathcal{H}$.*

Demostración: Si $A^*A = AA^*$ entonces $\langle Af, Ag\rangle = \langle f, A^*Ag\rangle = \langle f, AA^*g\rangle = \langle A^*f, A^*g\rangle$ cualesquiera sean $f, g \in \mathcal{H}$. Luego, con $g = f$, $\|Af\| = \|A^*f\|$.
Para demostrar la suficiencia de la condición, usamos la igualdad de polarización. En el caso complejo,

$$\langle f, (A^*A - AA^*)g\rangle = \frac{1}{4}\left(\langle f+g, (A^*A - AA^*)(f+g)\rangle - \langle f-g, (A^*A - AA^*)(f-g)\rangle\right.$$

$$\left. + i\langle f+ig, (A^*A - AA^*)(f+ig)\rangle - i\langle f-ig, (A^*A - AA^*)(f-ig)\rangle\right) \ ;$$

y cada uno de los sumandos a la derecha de la igualdad se anula si $\|Ah\| = \|A^*h\|$, cualquiera sea h. En el caso real, si $B^* = B$ se tiene

$$\langle f, Bg\rangle = \frac{1}{4}\left(\langle f+g, B(f+g)\rangle - \langle f-g, B(f-g)\rangle\right) \ ;$$

y ya que $A^*A - AA^*$ es autoadjunto,

$$\langle f, (A^*A - AA^*)g\rangle = \frac{1}{4}\left(\langle f+g, (A^*A - AA^*)(f+g)\rangle - \langle f-g, (A^*A - AA^*)(f-g)\rangle\right) \ .$$

Nuevamente cada uno de los sumandos a la derecha de la igualdad se anula si $\|Ah\| = \|A^*h\|$. $\square$

Proposición 4.4 *Si A es normal y el subespacio $\mathcal{M}$ es invariante para A entonces el complemento ortogonal de $\mathcal{M}$ es invariante para A.*

Demostración: Sea $\{e_1, e_2, \cdots, e_m\}$ una base ortonormal de $\mathcal{M}$ y $\{f_1, f_2, \cdots, f_p\}$ una base ortonormal de $\mathcal{M}^\perp$. Tenemos, ya que A es normal,

$$\sum_{k=1}^{m} |\langle e_j, Ae_k\rangle|^2 + \sum_{k=1}^{p} |\langle e_j, Af_k\rangle|^2$$

$$= \sum_{k=1}^{m} |\langle A^* e_j, e_k\rangle|^2 + \sum_{k=1}^{p} |\langle A^* e_j, f_k\rangle|^2 = \|A^* e_j\|^2 = \|Ae_j\|^2 \ .$$

Por otro lado, ya que $\mathcal{M}$ es invariante para A,

$$\|Ae_j\|^2 = \sum_{k=1}^{m} |\langle e_k, Ae_j\rangle|^2 \ .$$

Luego,

$$\sum_{j=1}^{m} \|Ae_j\|^2 = \sum_{j,k=1}^{m} |\langle e_k, Ae_j\rangle|^2 \ ,$$

y también

$$\sum_{j=1}^{m} \|Ae_j\|^2 = \sum_{j=1}^{m} \left\{ \sum_{k=1}^{m} |\langle e_j, Ae_k\rangle|^2 + \sum_{k=1}^{p} |\langle e_j, Af_k\rangle|^2 \right\}$$

$$= \sum_{j,k=1}^{m} |\langle e_k, Ae_j\rangle|^2 + \sum_{j=1}^{m}\sum_{k=1}^{p} |\langle e_j, Af_k\rangle|^2 \ .$$

Por lo tanto,

$$\sum_{j=1}^{m}\sum_{k=1}^{p} |\langle e_j, Af_k\rangle|^2 = 0 \ ,$$

de donde $\langle e_j, Af_k\rangle = 0$ y por ende $Af_k \in \mathcal{M}^\perp$. Ya que $\{f_1, f_2, \cdots, f_p\}$ es base (ortonormal) de $\mathcal{M}^\perp$, deducimos que $\mathcal{M}^\perp$ es invariante para A. $\Box$

Hemos aislado una de las propiedades básicas de los operadores normales:

$\boxed{\text{Si } A \text{ es normal entonces todo subespacio invariante para } A \text{ lo es también para } A^*}$.

Proposición 4.5 *Si A es normal y nilpotente entonces $A = 0$.*

Demostración: Sea k el índice de A. Sea $B = A^*A$ que es autoadjunto. Entonces $B^k = (A^*A)^k = (A^*)^k A^k = 0$ luego B es nilpotente. Sea m el índice de B. Si $B \neq 0$, tenemos $m \geq 2$ y hay $f \in \mathcal{H}$ con $B^{m-1}f \neq 0$. Entonces $2m - 2 \geq m$ y

$$0 \neq \|B^{m-1}f\|^2 = \langle B^{m-1}f, B^{m-1}f\rangle = \langle f, (B^*)^{m-1}B^{m-1}f\rangle = \langle f, B^{2m-2}f\rangle = 0 \ .$$

Por lo tanto $m < 2$ y por ende $m = 1$, o sea $B = 0$. Pero entonces, para todo $g \in \mathcal{H}$,

$$\|Ag\|^2 = \langle Ag, Ag\rangle = \langle g, Bg\rangle = 0 \ ,$$

y por lo tanto $Ag = 0$, o sea $A = 0$. $\Box$

Proposición 4.6 *Si A es normal entonces $ker(A^*) = ker(A)$ y $\mathcal{R}_A = \mathcal{R}_{A^*}$.*

Demostración: Si $Af = 0$ entonces $\|A^*f\|^2 = \langle A^*f, A^*f\rangle = \langle f, AA^*f\rangle = \langle f, A^*Af\rangle = 0$ y por ende $A^*f = 0$, lo que demuestra que $ker(A) \subset ker(A^*)$. Pero entonces $ker(A) \subset ker(A^*) \subset ker((A^*)^*) = ker(A)$ de donde la igualdad de los núcleos. La igualdad de las imágenes es consecuencia de $\mathcal{R}_A = ker(A^*)^\perp$. $\square$

Proposición 4.7 *Si A es normal, $\lambda \in \sigma(A)$ y $Af = \lambda f$, entonces $A^*f = \overline{\lambda}f$. Además, el índice de $A - \lambda\mathbf{1}$ es 1.*

Demostración:

$$\|(A^* - \overline{\lambda}\mathbf{1})f\|^2 = \langle (A^* - \overline{\lambda}\mathbf{1})f, (A^* - \overline{\lambda}\mathbf{1})f\rangle$$
$$= \langle f, (A - \lambda\mathbf{1})(A^* - \overline{\lambda}\mathbf{1})f\rangle = \langle f, (A^* - \overline{\lambda}\mathbf{1})(A - \lambda\mathbf{1})f\rangle = 0 \ ,$$

luego $A^*f = \overline{\lambda}f$.
Ahora, si $g \in ker((A - \lambda\mathbf{1})^2)$, entonces con $h := (A - \lambda\mathbf{1})g$, tendremos $(A - \lambda\mathbf{1})h = 0$ y, por lo recién demostrado, $(A^* - \overline{\lambda}\mathbf{1})h = 0$. Luego,

$$\|h\|^2 = \langle (A - \lambda\mathbf{1})g, h\rangle = \langle g, (A^* - \overline{\lambda}\mathbf{1})h\rangle = 0 \ ,$$

y por lo tanto, $h = 0$; luego $(A - \lambda\mathbf{1})g = 0$ o sea que $g \in ker(A - \lambda\mathbf{1})$. Esto demuestra que $ker((A - \lambda\mathbf{1})^2) \subset ker(A - \lambda\mathbf{1})$; y ya que la inclusión inversa es trivial, $ker((A - \lambda\mathbf{1})^2) = ker(A - \lambda\mathbf{1})$ y el índice es 1. $\square$

Proposición 4.8 *Si A es normal, los subespacios espectrales $\mathcal{H}(\lambda)$ coinciden con los autoespacios $\{f \in \mathcal{H} : \ Af = \lambda f\}$, y son dos-a-dos ortogonales.*

Demostración: Por la proposición anterior $\mathcal{H}(\lambda) = ker(A - \lambda\mathbf{1}) = \{f \in \mathcal{H} : \ Af = \lambda f\}$. Si $Af = \lambda f$, y $Ag = \mu g$ con $\lambda \neq \mu$, entonces, también por la proposición anterior,

$$(\lambda - \mu)\langle f, g\rangle = \langle \overline{\lambda}f, g\rangle - \langle f, \mu g\rangle = \langle A^*f, g\rangle - \langle f, Ag\rangle = 0 \ ,$$

y esto tiene a $\langle f, g\rangle = 0$ como consecuencia. $\square$

De estos resultados se obtiene inmediatamente el teorema 4.3 del teorema espectral general 3.6.

5. El caso de un espacio vectorial real

Cuando el espacio vectorial subyacente $\mathcal{H}$ es real el análisis de un operador es más tedioso. El motivo algebraico es que un polinomio con coeficientes reales no tiene porque tener soluciones. Transportando esto, un operador sobre un espacio vectorial real puede tener espectro vacio. El ejemplo canónico es

$$\left[\begin{array}{cc} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{array}\right] \ ,$$

una simple rotación por un ángulo ϕ en $\mathbb{R}^2$, cuyo espectro es vacio salvo cuando $\phi = 0, \pi$ (mod 2π).

Los teoremas de triangularización 3.1 y de triangularización unitaria de Schur 4.1, se demostrarón usando la existencia de autovalores. De hecho ambos son inválidos en el caso real. En efecto si para alguna base, la matriz asociada a un operador A es triangular superior, entonces A tiene todos los autovalores que se necesitan: los elementos diagonales de la matriz.
La forma canónica de Jordan no se puede lograr para un operador sobre un espacio vectorial real. Hay una versión real que descompone en términos de matrices de Hessenberg (vease: R.A. Horn, and C.R. Johnson: *Matrix Analysis*. Cambridge University Press, Cambridge 1985; p. 150-154).

Cuando el espacio vectorial real $\mathcal{E}$ está munido de un producto escalar real, lo llamamos **espacio euclideo** y, por supuesto, todas las nociones geométricas tienen sentido y coinciden con las del caso complejo. Cuando el espacio es euclideo, es costumbre llamar **simétrico** a un operador autoadjunto; y **ortogonal** a un operador unitario.

El siguiente resultado da la forma canónica en el caso de operadores normales.

Teorema 5.4 *(Teorema espectral para operadores normales) Si $\mathcal{E}$ es un espacio euclideo, y A es un operador normal sobre $\mathcal{E}$ entonces hay una base ortonormal de $\mathcal{E}$, y una operador ortogonal T sobre $\mathcal{E}$ tal que,*

$$[T^*AT] = A_1 \oplus A_2 \oplus \cdots A_k \ ,$$

donde las matrices A_1, A_2, $\cdots$, A_k son reales y de dimensión (1×1), o bien reales y de dimensión (2×2) de la forma

$$\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \ , \tag{8}$$

con $\beta \neq 0$[6].

Daremos la demostración más adelante. Presentamos una consecuencia inmediata que surge al particularizar a operadores simétricos; para ellos no pueden aparecer los bloques 2×2 (8) salvo para $\beta = 0$ en cuyo caso son bloques 1×1.

Teorema 5.5 *(Teorema espectral para operadores simétricos) Si $\mathcal{E}$ es un espacio euclideo y A es un operador simétrico, entonces $\sigma(A) \subset \mathbb{R}$ y para cada $\lambda \in \sigma(A)$ existe un orto-proyector P_λ tal que $P_\lambda P_\mu = 0$ si $\lambda \neq \mu$ y*

$$A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda \ .$$

Si $[A]$ es la matriz asociada a A via una base ortonormal, entonces hay un operador ortogonal T tal que $[T^*AT]$ es diagonal.

[6]Los autovalores de esta matriz, vista como matriz compleja, son $\alpha \pm i\beta$.

La aplicación de este resultado a las formas cuadráticas en $\mathbb{R}^n$ es inmediata; la incluimos dado su uso frecuente.

Teorema 5.6 *Si* $Q : \mathbb{R}^n \to \mathbb{R}$ *es una forma cuadrática, o sea*

$$Q(x) = \sum_{j,k=1}^{n} q_{j,k} x_j x_k \ , \quad x = (x_1, x_2, \cdots, x_n) \ ,$$

con $q_{j,k} = q_{k,j} \in \mathbb{R}$, *entonces hay una matriz* $(n \times n)$, $[T]$ *que es ortogonal y tal que con*

$$[y](x) := T[x]$$

se tiene

$$Q(x) = \sum_{j=1}^{n} \kappa_j (y_j(x))^2 \ ,$$

donde $\kappa_1, \kappa_2, \cdots, \kappa_n$ *son los autovalores de la matriz* $(q_{j,k})$ *repetidos de acuerdo a su multiplicidad.*

El resto de esta sección se dedica a la demostración del teorema 5.4. Primeramente observamos que la definición de un autovalor para el caso real es idéntica a la del caso complejo y que los resultados de la §2., han sido formulados para incluir el caso real, e.g. la proposición 2.1 es válida. El polinomio característico del operador A es el polinomio de grado n con coeficientes reales

$$\chi_A(\lambda) := det(A - \lambda \mathbf{1}) \ .$$

Observamos que si $a, b \in \mathbb{R}$ satisfacen $a^2 < b$ entonces $x^2 + 2ax + b$ no tiene raices reales. Todo polinomio con coeficientes reales factoriza en polinomios de forma $x - \lambda$, donde λ es raíz del polinomio, y en polinomios de grado 2 sin raices reales.

Proposición 5.9 *Si* $p(x) = a_o + a_1 x + \cdots + a_n x^n$ *con* $a_n \neq 0$ *y* $a_k \in \mathbb{R}$ *para* $k = 0, 2, \cdots, n$, *entonces*

$$p(x) = a_n \left(\prod_{j=1}^{r} (x - \lambda_j)^{m_j} \right) \left(\prod_{j=1}^{s} (x + 2a_j x + b_j)^{\ell_j} \right) \ ,$$

donde $\lambda_1, \lambda_2, \cdots, \lambda_r \in \mathbb{R}$ *son las distintas raices reales de* p *de multiplicidades* $m_1, m_2, \cdots, m_r$*; y donde* $a_j, b_j \in \mathbb{R}$ *satisfacen* $a_j^2 < b_j$ *y los números naturales* ℓ_j *satisfacen* $\ell_1 + \ell_2 + \cdots + \ell_s = n - m_1 - m_2 - \cdots - m_r$.

Demostración: Esto es consecuencia inmediata del Teorema fundamental del Álgebra. Visto como polinomio de una variable compleja, p tiene raices $\lambda_1, \cdots, \lambda_p$ con $1 \leq p \leq n$ de multiplicidades $m_1, m_2, \cdots, m_p$, y

$$p(z) = a_n \prod_{j=1}^{p} (z - \lambda_j)^{m_j} \ .$$

Si $\lambda \in \mathbb{C}$ es raíz de p entonces $\overline{\lambda}$ también es raíz, ya que $p(\overline{\lambda}) = \sum_{k=0}^{n} a_k (\overline{\lambda})^k = \overline{\sum_{k=0}^{n} a_n \lambda^k} = \overline{p(\lambda)} = 0$. Si $\lambda \in \mathbb{C}$ es raíz de parte imaginaria no nula, tendremos

$$(z - \lambda)(z - \overline{\lambda}) = z^2 - 2Re(\lambda) z + |\lambda|^2 \ ,$$

y $Re(\lambda)^2 < Re(\lambda)^2 + Im(\lambda)^2 = |\lambda|^2$. Por lo tanto,

$$p(x) = a_n \left(\prod_{1\leq j\leq p\ ,\ \lambda_j\in\mathbb{R}} (x-\lambda)^{m_j} \right) \left(\prod_{1\leq j\leq p\ ,\ Im(\lambda_j)\neq 0} (x^2 - 2Re(\lambda_j)x + |\lambda_j|^2)^{m_j} \right) .$$

□

Lema 4 *Si $a, b \in \mathbb{R}$, y $x^2 + 2ax + b$ es un factor del polinomio minimal p de A entonces $A^2 + 2aA + b\mathbf{1}$ no es inyectivo.*

Demostración: Si $x^2 + 2ax + b$ es factor del polinomio minimal de A entonces

$$0 = p(A) = q(A)(A^2 + 2aA + b\mathbf{1}) .$$

Si $A^2 + 2aA + b\mathbf{1}$ fuere inyectivo, multiplicando por derecha con $(A^2 + 2aA + b\mathbf{1})^{-1}$ deducimos que $q(A) = 0$ lo que contradice la minimalidad de p. □

Lema 5 *Si A es un operador sobre un espacio real que no tiene autovalores, entonces hay un subsepacio bidimensional invariante para A.*

Demostración: Hay $a, b \in \mathbb{R}$ con $a^2 < b$ tal que $A^2 + 2aA + b\mathbf{1}$ no es inyectivo. Luego, hay $0 \neq f \in ker(A^2 + 2aA + b\mathbf{1})$. Sea $F := lin\{f, Af\}$. Entonces, ya que $A^2 f = -2aAf - bf$, concluimos que F es invariante para A.
Si $cf + dAf = 0$ y $d \neq 0$, entonces $Af = -(c/d)f$ lo que contradice la ausencia de autovalores de A. Por lo tanto $d = 0$ y entonces $c = 0$ ya que $f \neq 0$. Esto demuestra que F tiene dimensión 2. □

Supongamos que $\mathcal{H}$ que es real, está munido de un producto escalar $\langle\cdot,\cdot\rangle$. Si A es un operador normal cuyo espectro es no vacio, consideremos el subespacio $\mathcal{K} := \oplus_{\lambda\in\sigma(A)}\mathcal{H}(\lambda)$ de $\mathcal{H}$ formado por la suma directa de los subespacios espectrales de A.

Lema 6 *$\mathcal{K}$ y $\mathcal{K}^\perp$ son invariantes para A y la restricción de A a $\mathcal{K}^\perp$ que llamamos B no tiene autovalores.*

Demostración: Por la proposición 4.8, $\mathcal{K}$ es invariante para A. Por la proposición 4.7, cada $\mathcal{H}(\lambda)$ es invariante para A^* luego $\mathcal{K}$ también lo es. Aplicando la proposición 4.2, $\mathcal{K}^\perp$ es invariante para A. Está claro que B no tiene autovalores, pues si los tuviera estos serían autovalores de A y estarían contemplados en $\mathcal{K}$. □

Existen entonces $a_j, b_j \in \mathbb{R}$ con $a_j^2 < b_j$ tales que

$$\chi_B^o(x) = (x^2 + 2a_1x + b_1)^{m_1} \cdots (x^2 + 2a_px + b_p)^{m_p} .$$

Consideramos los pares ordenados (a_j, b_j) distintos, $j = 1, 2, \cdots, p$. Por el lema 4 se tiene que $B^2 + 2a_jB + b_j\mathbf{1}$ no es inyectivo.

Lema 7 *El índice de $B^2 + 2a_jB + b_j\mathbf{1}$ es uno.*

Demostración: Esto es la proposición 4.7 aplicada a $C := B^2 + 2a_jB + b_j\mathbf{1}$ y su autovalor 0. $\square$

Sea

$$\mathcal{K}_j := ker((B^2 + 2a_jB + b_j\mathbf{1}))\ ,\ \ j = 1, 2, \cdots, p\ .$$

Lema 8 *Cada $\mathcal{K}_j$ es invariante para B y para B^*. Si $j \neq k$, $\mathcal{K}_j$ es ortogonal a $\mathcal{K}_k$.*

Demostración: Sea $C_j := B^2 + 2a_tB + b_t\mathbf{1}$, que es normal. Si $f \in \mathcal{K}_j$ entonces $C_jBf = BC_jf = 0$ y $C_jB^*f = B^*C_jf = 0$; luego $\mathcal{K}_j$ es invariante para B y para B^*. Por la proposición 4.2, $\mathcal{K}_j^\perp$ es invariante para B.
Si $f \in \mathcal{K}_j \cap \mathcal{K}_k$, donde $j \neq k$, tenemos $C_jf = C_kf = 0$ y de aquí, $B^2f = -2a_jBf - b_jf = -2a_kBf - b_kf$ luego $2(a_k - a_j)Bf = (b_j - b_k)f$. Si $a_k = a_j$, entonces $(b_j - b_k)f = 0$ implica que o bien $b_j = b_k$ en cuyo caso se contradice $j \neq k$; o bien $f = 0$. Si $a_j \neq a_k$, entonces $Bf = ((b_j - b_k)/2(a_k - a_j))f$, y ya que B no tiene autovalores, deducimos que $f = 0$. Si $f \in \mathcal{K}$, entonces $f = x \oplus y$ donde $x \in \mathcal{K}_j$ y $y \in \mathcal{K}_j^\perp$. Entonces $C_kx = B^2x + 2a_kBx + b_kx \in \mathcal{K}_j$ ya que $\mathcal{K}_j$ es invariante para B. También $C_ky = B^2y + 2a_kBy + b_ky \in \mathcal{K}_j^\perp$ pues este subespacio es invariante para B. Luego, $0 = C_kx \oplus C_ky$ implica $C_kx = 0$ y entonces $x \in \mathcal{K}_j \cap \mathcal{K}_k$, de donde deducimos que $x = 0$. Esto demuestra que $\mathcal{K}_k \perp \mathcal{K}_j$. $\square$

Lema 9 $\oplus_{j=1}^p \mathcal{K}_j = \mathcal{K}^\perp$.

Demostración: Sea $\mathcal{M}$ el complemento ortogonal de $\oplus_{j=1}^p \mathcal{K}_j$ en $\mathcal{K}^\perp$. Entonces, ya que $\oplus_{j=1}^p \mathcal{K}_j$ es invariante para B y para B^*, la proposición 4.2 indica que $\mathcal{M}$ es invariante para B y B^*. La restricción D de B a $\mathcal{M}$ es normal y no tiene autovalores. El polinomio característico χ_D es entonces producto de factores cuadráticos sin raices reales. Pero cualquiera de estos factores es factor del polinomio característico de B y debe ser uno de los factores $x^2 + 2a_jx + b_j$ con $a_j^2 < b_j$ asociado con $\mathcal{K}_j$. $\square$

Lema 10 *$\mathcal{K}_j$ tiene dimensión par, y existe una base ortonormal de $\mathcal{K}_j$ tal que para la restricción B_j de B a $\mathcal{K}_j$ se tiene*

$$[B_j] = \bigoplus_{\ell=1}^{m_j} \begin{bmatrix} -a_j & \sqrt{b_j - a_j^2} \\ -\sqrt{b_j - a_j^2} & -a_j \end{bmatrix}\ .$$

Demostración: $\mathcal{K}_j$ siendo invariante para B, la restricción B_j a $\mathcal{K}_j$ puede verse como operador de $\mathcal{K}_j$ en si mismo. Por el lema 5, hay un subespacio bidimensional F_1 de $\mathcal{K}_j$ que es invariante para B_j. Por las proposiciones 4.4 y 4.2, F_1 es invariante para B_j^*. Sea $\{x, y\}$ una base ortonormal de F_1. La restricción de B_j a F_1 tiene la matriz asociada

$$C = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}\ .$$

Ya que B_j es normal, obtenemos $\beta^2 = \gamma^2$ y $\alpha(\beta - \gamma) = \delta(\beta - \gamma)$. Si $\beta = \gamma$, entonces C es diagonal y tiene autovalores lo que contradice la propiedad de que B no tiene autovalores. Por lo tanto $\gamma = -\beta$ con $\beta \neq 0$ y esto conlleva que $\alpha = \gamma$. Además, $C^2 + 2a_jC + b_j\mathbf{1} = 0$

implica que $\alpha = -a_j$ y que $|\beta| = \sqrt{b_j - a_j^2}$. Intercambiando x con y si fuere necesario, obtenemos

$$C = \begin{bmatrix} -a_j & \sqrt{b_j - a_j^2} \\ -\sqrt{b_j - a_j^2} & -a_j \end{bmatrix} .$$

Podemos ahora proceder análogamente con la restricción de B_j al complemento ortogonal de F_1 en $\mathcal{K}_j$, para obtener un segundo subespacio F_2 de dimensión 2 que es invariante para B_j, etc. Esto puede repetirse hasta agotar a $\mathcal{K}_j$ que, por ende, debe tener dimensión par. $\square$

Índice alfabético