9. Autovalores y Autovectores

Sea V un espacio vectorial sobre el cuerpo K y sea $F : V \to V$ un operador lineal. Un escalar $\lambda \in K$ es un *autovalor* de F si existe $v \in V$, con $v \neq 0$, tal que

$$F(v) = \lambda v \qquad (v \neq 0)$$

En tal caso v es un *autovector* de F correspondiente al autovalor λ. Sinónimo de autovalor es *valor propio* (en inglés "eigenvalue", donde "eigen" proviene del alemán y significa propio) y sinónimo de autovector es *vector propio* ("eigenvector" en inglés).
La acción de F sobre un autovector es pues la multiplicación por un escalar. Por ejemplo, si $V = \mathbb{R}^n$, esto implica que $v \neq 0$ es autovector de F si y sólo si $w = F(v)$ tiene la misma dirección que v (aunque no necesariamente el mismo sentido) o es nulo. Como veremos, el conocimiento del conjunto de los autovalores y autovectores de un operador permite conocer su estructura en detalle.

El concepto de autovalor y autovector de un operador lineal tiene una importancia fundamental en Física, especialmente en Mecánica Cuántica. Mencionemos que en la misma los observables físicos corresponden a operadores lineales (de características determinadas) en un cierto espacio (el de los vectores que representan los estados del sistema) y las cantidades medibles son precisamente los autovalores de dichos operadores. E inmediatamente después de una medición, el sistema cuántico queda en un estado que es autovector del operador correspondiente al autovalor medido.
El concepto de autovalor y autovector es también fundamental para el tratamiento de sistemas descriptos por ecuaciones diferenciales lineales acopladas, tales como osciladores armónicos acoplados, ya sean clásicos o cuánticos, donde permite entender el concepto de desacoplamiento y modo normal. Otras aplicaciones incluyen, como veremos, la determinación de los ejes principales de inercia en un cuerpo rígido y la resolución de sucesiones definidas mediante relaciones recursivas lineales, tales como la famosa sucesión de Fibonacci.

Veamos algunas consecuencias inmediatas de tal definición:
10.1) El conjunto de autovectores de F correspondientes a un determinado autovalor λ, junto con el vector nulo 0, forma un subespacio de V denominado *espacio propio* o autoespacio ("eigenspace" en inglés) de F correspondiente al autovalor λ, que denotaremos como $V_F(\lambda)$ o simplemente, $V(\lambda)$.
En efecto, si $v \neq 0$ es autovector de F corresp. al autovalor λ, y $\alpha \in K$, $F(\alpha v) = \alpha F(v) = \alpha\lambda v = \lambda(\alpha v)$ $\forall$ α, por lo que αv es también autovector de F corresp. a λ si $\alpha \neq 0$
Y si v_1 y v_2 son autovectores corresp. al mismo autovalor λ, $F(v_1 + v_2) = F(v_1) + F(v_2) = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2)$, por lo que $v_1 + v_2$ es también autovector corresp. al autovalor λ (si $v_1 + v_2 \neq 0$).
La dimensión de $V_F(\lambda)$ es como mínimo uno (ya que al menos existe un autovector no nulo).

9.2) El espacio propio $V_F(\lambda)$, con λ autovalor de F, es el núcleo del operador $F - \lambda I$:

$$V_F(\lambda) = N[F - \lambda I]$$

donde I denota el operador identidad en V $(I(v) = v \ \forall \ v \in V)$.
En efecto, si $F(v) = \lambda v \Rightarrow 0 = F(v) - \lambda v = (F - \lambda I)(v)$, y si $(F - \lambda I)(v) = 0 \Rightarrow F(v) = \lambda v$.
Por lo tanto $\lambda \in K$ es autovalor de F si y sólo si

$$N[F - \lambda I] \neq \{0\}$$

En particular, $\lambda = 0$ es autovalor de F si y sólo si $N(F) \neq \{0\}$, es decir, si y sólo si F no es monomorfisomo. En tal caso $V_F(0) = N(F)$.

9.3) Si V es de dimensión finita n, $\lambda \in K$ es autovalor si y sólo si

$$\text{Det}[F - \lambda I] = 0$$

En efecto, si λ es autovalor, $N[F - \lambda I] \neq \{0\}$, por lo que $F - \lambda I$ no es monomorfismo. La matriz que representa a $F - \lambda I$ debe ser entonces singular en cualquier base, y por lo tanto, su determinante nulo. Por otro lado, si $\text{Det}[F - \lambda I] = 0$, $F - \lambda I$ no es monomorfismo, y por lo tanto existe $v \neq 0$ tal que $(F - \lambda I)(v) = 0$.
Recordemos que si e es una base de V,

$$\text{Det}[F - \lambda I] = |[F]_e - \lambda I_n|$$

donde $[F]_e \equiv [F]_e^e$ es la matriz que representa a F en dicha base, $I_n = [I]_e$ es la matriz identidad de $n \times n$ y $|A| = \text{Det}A$ denota el determinante de la matriz A. $\text{Det}[F - \lambda I]$ es *independiente* de la base elegida e:

$$|[F]_{e'} - \lambda I_n| = |S^{-1}[F]_e S - \lambda I_n| = |S^{-1}([F]_e - \lambda I_n)S| = |S^{-1}||[F]_e - \lambda I_n||S| = |[F]_e - \lambda I_n|$$

ya que $[I]_{e'} = [I]_e = I_n$ y $|S^{-1}| = 1/|S|$. Los autovalores de F en un espacio de dimensión finita n se obtienen entonces como las raíces pertenecientes al cuerpo K del polinomio

$$P(\lambda) = \text{Det}[F - \lambda I]$$

denominado *polinomio característico*, que es de grado n (pues $[F - \lambda I]_e$ es de $n \times n$). La ecuación $P(\lambda) = 0$ se denomina *ecuación caracerística* y posee, por lo tanto, a lo sumo n *raíces distintas*, que en general pueden ser complejas. Serán autovalores si pertenecen al cuerpo K. Si $K = \mathbb{C} \Rightarrow$ *toda raíz* de $P(\lambda)$ es autovalor.

9.4) *Teorema: Los autovectores de F correspondientes a autovalores distintos son LI*
Demostraremos el teorema por inducción. Para $n = 1$, v_1 autovector es LI pues es no nulo (mejor comprensión se logra comenzando con $n = 2$: Si $v_1 \neq 0$ y $v_2 \neq 0$ son autovectores correspondientes a autovalores *distintos* $\Rightarrow$ son LI pues de lo contrario $v_2 = \alpha v_1$, y correspondería por 9.1 al mismo autovalor que v_1). Supongamos ahora que $v_1, \ldots, v_{k-1}$ son autovectores LI, con autovalores $\lambda_1, \ldots, \lambda_{k-1}$, y que $v_k \neq 0$ es autovector con autovalor λ_k, siendo λ_k distinto a todos los anteriores. Si

$$0 = \alpha_1 v_1 + \ldots + \alpha_{k-1} v_{k-1} + \alpha_k v_k$$

entonces

$$0 = F(0) = \alpha_1 F(v_1) + \ldots + \alpha_{k-1} F(v_{k-1}) + \alpha_k F(v_k) = \alpha_1 \lambda_1 v_1 + \ldots \alpha_{k-1} \lambda_{k-1} v_{k-1} + \alpha_k \lambda_k v_k$$

Restando ahora la primera ecuación mult. por λ_k,

$$0 = (\lambda_1 - \lambda_k)\alpha_1 v_1 + \ldots + (\lambda_{k-1} - \lambda_k)\alpha_{k-1} v_{k-1}$$

que implica $\alpha_1 = \ldots = \alpha_{k-1} = 0$ por ser $v_1, \ldots v_{k-1}$ LI y $\lambda_k \neq \lambda_i$ para $i = 1, \ldots, k-1$. Por lo tanto, también $\alpha_k = 0$ y entonces $v_1, \ldots, v_k$ son L.I.

Nótese que la demostración es igualmente válida si los $\lambda_1, \ldots, \lambda_{k-1}$ no son todos distintos (pero sí distintos a λ_k) siempre que los $v_1, \ldots, v_{k-1}$ sean LI
Por lo tanto, *ningún elemento $v_k \neq 0$ de $V(\lambda_k)$ puede ser generado por autovectores correspondientes a autovalores distintos de λ_k.*

9.5) Si $\lambda_1 \neq \lambda_2 \Rightarrow V_F(\lambda_1) \cap V_F(\lambda_2) = \{0\}$
Es consecuencia inmediata del último párrafo. Si $v \in V_F(\lambda_2)$ y $v \in V_F(\lambda_1)$, entonces $F(v) = \lambda_1 v = \lambda_2 v$, por lo que $(\lambda_1 - \lambda_2)v = 0$ y por lo tanto $v = 0$. Esto implica que $V_F(\lambda_1) + V_F(\lambda_2) = V_F(\lambda_1) \oplus V_F(\lambda_2)$.
Del mismo modo, la intersección de $V_F(\lambda_k)$ con la suma $V_F(\lambda_1) + \ldots + V_F(\lambda_{k-1})$ es $\{0\}$ si λ_k es distinto a todos los autovalores anteriores. Si así no fuese existiría un vector $v \in V_F(\lambda_k)$ con $v \neq 0$ que puede ser escrito como combinación lineal de autovectores correspondientes a autovalores distintos, pero esto es imposible por el teorema 9.4 anterior.
Podemos entonces escribir la suma de espacios propios como suma directa $V_F(\lambda_1) \oplus \ldots \oplus V_F(\lambda_k)$.

9.6) Un operador F en un espacio V de dimensión finita se dice *diagonalizable* si existe una base formada por autovectores de F.
En tal caso, denotando la base como $e' = (e'_1, \ldots, e'_n)$, con $F(e'_i) = \lambda_i e'_i$, $i = 1, \ldots, n$, la matriz que representa a F en dicha base es *diagonal*:

$$[F]_{e'} = \begin{pmatrix} & \ldots & \\ [F(e'_1)]_{e'} & \ldots & [F(e'_n)]_{e'} \\ & \ldots & \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

Recíprocamente, si $[F]_{e'}$ es diagonal $\Rightarrow F(e'_i) = \lambda_i e'_i$ y e' es necesariamente una base de autovectores. Si F es diagonalizable y e es una base arbitraria de V, tenemos

$$[F]_{e'} = S^{-1}[F]_e S$$

con $[F]_{e'}$ *diagonal* y $S = [I]_e^{e'}$ la matriz de cambio de base, *por lo que existe una matriz no singular* S *tal que* $S^{-1}[F]_e S$ *es diagonal*. La columna i de S es el vector de componentes $[e'_i]_e$ del autovector e'_i corresp. al autovalor λ_i en la base original e. Esta es la forma de construir la matriz diagonalizante S.

Consecuencia Importante de 9.4: Si $P(\lambda)$ *posee* n *raíces distintas* $\lambda_i \in K$, $\Rightarrow F$ *es diagonalizable.*
En efecto, en tal caso existirán n autovectores LI (uno por cada autovalor) que serán por tanto base de V. La dimensión de cada espacio propio $V_F(\lambda_i)$ es en este caso 1, que es igual a la multiplicidad de cada raíz.

9.7) Teorema: *F es diagonalizable si y sólo si i) todos las raíces de $P(\lambda)$ pertenecen al cuerpo y ii) la dimensión del espacio propio $V_F(\lambda_i)$ correspondiente a la raíz λ_i es igual a la multiplicidad m_i de dicha raíz.*
La dimensión del espacio propio $d_i = \dim V_F(\lambda_i)$ es el máximo número de autovectores LI que pueden obtenerse para un mismo autovalor λ_i, y se denomina también *multiplicidad geométrica* de λ_i.
Demostración: Supongamos F diagonalizable. En una base e' en la que $[F]_{e'}$ es diagonal, tenemos

$$P(\lambda) = (\lambda_1 - \lambda)\ldots(\lambda_n - \lambda)$$

La multiplicidad m_i de una raíz λ_i será pues igual al número de veces que λ_i se repite en la diagonal. Pero este número es igual al número de vectores de la base e' que tienen a λ_i como autovalor, que es precisamente la dimensión del espacio propio. Nótese también que d_i es el número de filas nulas de $[F]_{e'} - \lambda_i I_n$.
Por otro lado, si la dimensión de cada espacio propio es igual a la multiplicidad m_i de la raíz λ_i, la suma directa de todos los espacios propios correspondientes a autovalores *distintos*, $V_F(\lambda_1) \oplus \ldots \oplus V_F(\lambda_k)$ tendrá dimensión $d_1 + \ldots + d_k = m_1 + \ldots + m_k = n$ (ya que la suma de todas las multiplicidades es igual al grado del polinomio), por lo que será igual al espacio V. Existe entonces una base formada por autovectores de F.

9.8) En general, la dimensión del espacio propio $V_F(\lambda_i)$ puede ser igual o menor que la multiplicidad de la raíz λ_i: dim $V_F(\lambda_i) \leq m_i$.
En efecto, eligiendo una base e donde los primeros $d_i = \dim V_F(\lambda_i)$ elementos formen una base de $V_F(\lambda_i)$, las primeras d_i columnas de $[F]_e$ tendrán elementos no nulos sólo en la diagonal y por lo tanto $P(\lambda) = \text{Det}[[F]_e - \lambda I_n] = (\lambda_i - \lambda)^{d_i} Q(\lambda)$, con $Q(\lambda)$ un polinomio de grado $n - d_i$, por lo que m_i será como mínimo d_i ($m_i = d_i$ si $Q(\lambda_i) \neq 0$ y $m_i > d_i$ si $Q(\lambda_i) = 0$).
Si $d_i < m_i$, *F no es diagonalizable* (aún tomando como cuerpo $\mathbb{C}$).

Ejemplo 1 (Casos Triviales): Si $I : V \to V$ es el operador identidad $\Rightarrow I(v) = v \; \forall \; v \in V$ por lo que su único autovalor es 1. El espacio propio correspondiente es V ($V_F(1) = V$).
Si V es de dimensión finita n, puede llegarse a la misma conclusión notando que

$$P(\lambda) = \text{Det}[I - \lambda I] = |I_n - \lambda I_n| = |(1 - \lambda)I_n| = (1 - \lambda)^n$$

$\lambda = 1$ es pués la única raíz de $P(\lambda)$, y posee multiplicidad n, que es igual a la dimensión del espacio propio (el mismo V). I es por lo tanto diagonalizable (caso trivial).

Si $0 : V \to V$ es el operador nulo $\Rightarrow 0(v) = 0 = 0v \; \forall \; v \in V$ por lo que su único autovalor es 0, y el espacio propio correspondiente es V ($V_F(0) = V$).
Si V es de dimensión $n \Rightarrow [0]_e = 0$ (la matriz nula) en cualquier base y por lo tanto $P(\lambda) = \text{Det}[0 - \lambda I_n] = |-\lambda I_n| = (-\lambda)^n$, por lo que 0 es la única raíz, con multiplicidad n. 0 es también trivialm. diagonalizable.

Ejemplo 2: Sea $F : \mathbb{R}^2 \to \mathbb{R}^2$ la reflexión respecto del eje x. Si $e = (e_1, e_2)$ es la base canónica, con $e_1 = (1, 0)$, $e_2 = (0, 1)$, sabemos que $F(e_1) = e_1$, $F(e_2) = -e_2$. Por lo tanto e_1 es autovector de F con autovalor 1 y e_2 autovector de F con autovalor -1. No pueden existir otros autovalores pues la dimensión de V es 2. $V_F(1)$ es entonces el espacio generado por e_1, es decir, el conjunto de vectores $(x, 0)$, con $x \in \mathbb{R}$, sobre los que F actúa como identidad, y $V_F(-1)$ el generado por e_2, es decir, el conjunto de vectores $(0, y)$, con $y \in \mathbb{R}$, para los que la acción de F es la inversión de sentido.
Podemos obtener el mismo resultado a partir de la representación matricial

$$[F]_e = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

que ya es diagonal, por lo que los autovalores son 1 y -1: Tenemos $P(\lambda) = |[F]_e - \lambda I_2| = (1 - \lambda)(-1 - \lambda)$, siendo entonces las raíces ± 1.

Ejemplo 3: Sea $F:\mathbb{R}^2\to\mathbb{R}^2$ la reflexión respecto de la recta de ecuación $y=x$, dada por (recordar ejemplo dado) $F(x,y)=(y,x)$. Si $e'=(e'_1,e'_2)$, con $e'_1=(1,1)$, $e'_2=(-1,1)$, tenemos $F(e'_1)=e'_1$, $F(e'_2)=-e'_2$, por lo que los autovalores son nuevamente 1 y -1, con $V_F(1)$ el espacio generado por e'_1 y $V_F(-1)$ el espacio generado por e'_2. Se obtiene entonces

$$[F]_e=\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad [F]_{e'}=\begin{pmatrix}1&0\\0&-1\end{pmatrix}=S^{-1}[F]_eS\,,\quad S=\begin{pmatrix}1&-1\\1&1\end{pmatrix}$$

El pol. característico es $P(\lambda)=|[F]_e-\lambda I_2|=\lambda^2-1=(1-\lambda)(-1-\lambda)=|[F]_{e'}-\lambda I_2|$ y sus raíces ± 1.

Ejemplo 4: **Operador de Proyección:** Si $P^2=P\Rightarrow$ los únicos autovalores posibles de P son 0 y 1, con $V_P(0)=N(P)$ si $N(P)\neq\{0\}$ y $V_P(1)=I(P)$ si $I(P)\neq\{0\}$.
Dem.: Hemos visto que si $P^2=P\Rightarrow P$ es un proyector sobre $I(P)$ en la dirección de $N(P)$, con $I(P)\oplus N(P)=V$ y $P(v)=v$ si $v\in I(P)$ y $P(v)=0=0v$ si $v\in N(P)$. Por lo tanto, los autovalores son: 1 (si $I(P)\neq\{0\}$, es decir, si $P\neq 0$) y 0 (si $N(P)\neq\{0\}$, es decir, si $P\neq I_V$). P es pues dagonalizable, siendo una base de autovectores la formada por la unión de una base de la imagen $I(P)$ y una del núcleo $N(P)$. P es pues siempre diagonalizable.
Estas propiedades pueden también demostrarse directamente: Si $P(v)=\lambda v$, con $v\neq 0\Rightarrow P^2(v)=P(P(v))=P(\lambda v)=\lambda P(v)=\lambda^2 v$, pero también $P^2(v)=P(v)=\lambda v$, por lo que $\lambda^2=\lambda$, es decir $\lambda(\lambda-1)=0$. Esto implica $\lambda=0$ o $\lambda=1$. Si $\exists\, v\neq 0$ tal que $P(v)=0\Rightarrow 0$ es autovalor y $V_P(0)=N(P)$. Si $\exists\, v\neq 0$ tal que $P(v)=v\Rightarrow 1$ es autovalor y $V_P(1)=I(P)$. En efecto, si $v\neq 0$ y $P(v)=v\Rightarrow v\in I(P)$, y si $v\in I(P)$ $\Rightarrow v=P(v')$ y $P(v)=P(P(v'))=P^2(v')=P(v')=v$, por lo que $v\in V_P(1)$.

Ejemplo 6: **Potencias de operadores lineales.** Si v es autovector de F con autovalor $\lambda\Rightarrow v$ es también autovector de F^k con autovalor λ^k para cualquier $k>0$ natural, y también para cualquier $k<0$ entero si F es invertible (o sea, automorfismo), en cuyo caso $\lambda\neq 0$ (pues necesariamente $N(F)=\{0\}$; ver 9.4).
Dem.: Si $k=2$, $F^2(v)=F(F(v))=F(\lambda v)=\lambda F(v)=\lambda^2 v$. La demostración para $k>2$ es análoga y puede hacerse fácilmente por inducción: $F^k(v)=F(F^{k-1}(v))=F(\lambda^{k-1}v)=\lambda^{k-1}F(v)=\lambda^k v$.
Si $k=-1$, $F^k=F^{-1}$ es la inversa de F y $v=F^{-1}F(v)=F^{-1}(\lambda v)=\lambda F^{-1}(v)$, por lo que $F^{-1}(v)=\lambda^{-1}v$. El resultado para $F^{-k}\equiv(F^{-1})^k$ y $k>1$ es entonces obvio: $F^k(v)=\lambda^k v\;\forall\, k\in\mathbb{Z}$ (véase también 9.4)

Ejemplo 7: Si λ es un autovalor de $F\Rightarrow\alpha\lambda+c$ es autovalor del operador $\alpha F+cI$, con $\alpha, I\in K$ e I el op. identidad: En efecto, si $F(v)=\lambda v$, $v\neq 0\Rightarrow(\alpha F+cI)(v)=\alpha F(v)+cv=(\alpha\lambda+c)v$.

Ejemplo 8: Sea $D^2:V\to V$ el operador derivada segunda en el espacio V de funciones $f:[0,a]\to\mathbb{R}$, $a>0$, que son derivables a cualquier orden y satisfacen $f(0)=f(a)=0$ (V es de dimensión infinita).
La ecuación $D^2(f)=\lambda f$ conduce a la ecuación diferencial $f''=\lambda f$, cuya solución es $f(x)=c_1\cos(sx)+c_2\sin(sx)$ con $\lambda=-s^2$. La condición de contorno $f(0)=f(1)=0$ implica $c_1=0$ y $\sin(sa)=0$, o sea, $s=n\pi/a$, con $n>0$ natural. Por lo tanto, los autovalores son $\lambda_n=-n^2\pi^2/a^2$ y los autovectores (llamados en este caso autofunciones) $f_n(x)=c_n\sin(n\pi x/a)$, con $n=1,2,\ldots$ y $c_n\neq 0$ arbitrario.
Ejemplo 9: La **ecuación de Schrödinger** estacionaria de la mecánica cuántica es

$$H|\Psi\rangle=E|\Psi\rangle$$

donde H es un operador lineal, $|\Psi\rangle$ un vector no nulo que representa el estado del sistema y E la energía del estado. Es pues una ecuación de autovalores: La energía E representa un autovalor de H y el vector $|\Psi\rangle$ el autovector correspondiente de H. Para una partícula en una dimensión, H toma la forma

$$H=-\frac{\hbar^2}{2m}D^2+V(x)$$

donde D es el operador derivada anterior, $\hbar=h/(2\pi)$, con h la constante de Planck, m la masa de la partícula y $V(x)$ el potencial, con $|\Psi\rangle\to\psi(x)$, siendo $\psi(x)$ la "función de onda" (en realidad, $\psi(x)=\langle x|\Psi\rangle$, y la forma anterior de H es su representación efectiva en la base de autoestados $|x\rangle$ del operador posición: $\langle x|H|\Psi\rangle=(-\frac{\hbar^2}{2m}D^2+V(x))\psi(x))$. H es el operador energía pues el impulso está representado por el operador $p=-i\hbar D$, por lo que el primer término de H es el operador energía cinética $P^2/(2m)$.
Si $V(x)=0$ para $x\in[0,a]$ y $V(x)=\infty$ para $x>a$ o $x<0$, la Ec. de Schrödinger se reduce al problema del ej. 8: Tenemos $H=-\frac{\hbar^2}{2m}D^2$ para $x\in[0,a]$, con $\psi(x)=0$ para $x\geq a$ o $x\leq 0$, debiendo ser continua. Los autovalores son por lo tanto $E_n=\frac{\hbar^2n^2\pi^2}{2ma^2}$ y los autovectores $\psi_n(x)=c_n\sin(n\pi x/a)$.

10. Autovalores y Autovectores de Matrices

Todas las definiciones y propiedades anteriores se aplican igualmente al cálculo de autovalores y autovectores de matrices cuadradas, que pueden ser siempre consideradas como la representación de un cierto operador lineal (en un espacio vectorial de dimensión n) en una cierta base. Consideraremos en lo sucesivo $K = \mathbb{C}$.
Dada una matriz A de $n \times n$, λ es autovalor de A sii

$$|A - \lambda I_n| = 0$$

donde $|\ldots|$ denota el determinante. El polinomio

$$P(\lambda) = |A - \lambda I_n| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ & & \ldots & \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} \qquad (10.1)$$

se denomina *polinomio característico* de la matriz A y es de grado n en λ, por lo que posee a lo sumo n raíces distintas.
Un vector columna $X = \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix}$ de $n \times 1$ es *autovector* de A correspondiente al autovalor λ si $X \neq 0$ y

$$AX = \lambda X$$

El conjunto de autovectores X corresp. a λ puede obtenerse resolviendo el sistema homogéneo

$$(A - \lambda I_n)X = 0$$

10.1) Las matrices semejantes poseen el mismo polinomio característico y por lo tanto los mismos autovalores. Recordemos que A es semejante o similar a B si $A = S^{-1}BS$, con S de $n \times n$ no singular.
En efecto $P_A(\lambda) = |A - \lambda I_n| = |S^{-1}BS - \lambda I_n| = |S^{-1}(B - \lambda I_n)S| = |B - \lambda I_n| = P_B(\lambda)$

Los autovectores no son en general los mismos: Si $AX = \lambda X$ y $A = S^{-1}BS$, el correspondiente autovector de B es SX, como el lector podrá fácilmente demostrar. Esto puede verse también directamente a partir del cambio de base asociado.

10.2) A y A^t (t denota traspuesta) poseen el mismo polinomio característico y por lo tanto los mismos autovalores (pero no necesariamente los mismos autovectores). Como $|B| = |B^t|$ $\forall$ B de $n \times n$,

$$P_{A^t}(\lambda) = |A^t - \lambda I_n| = |(A - \lambda I_n)^t| = |A - \lambda I_n| = P_A(\lambda)$$

10.3) Si A es real $\Rightarrow$ $P(\lambda)$ es real y por lo tanto sus raíces complejas aparecerán en pares conjugados:
Si λ es una raíz compleja, $0 = [P(\lambda)]^* = P(\lambda^*)$.
Además, si X es autovector de A con autovalor λ y A es real, entonces X^* es autovector de A con autovalor λ^*: Como $AX = \lambda X \Rightarrow (AX)^* = AX^* = \lambda^* X^*$.

10.4) A es una matriz singular si y sólo si A posee al menos un autovalor nulo.
Si A es singular $\Rightarrow$ $|A| = 0$ y por lo tanto, $|A - 0I_n| = 0$, por lo que 0 es autovalor. Análogamente, si $|A - 0I_n| = 0 \Rightarrow |A| = 0$ y por lo tanto, A es singular. Esto puede también deducirse directamente de 10.5

Un operador $F : V \to V$ en un espacio de dimensión finita tendrá pues un autovalor nulo si y sólo si no es un automorfismo. Notemos también que si $\mathrm{Det}[F] = 0$, el núcleo $N(F)$ no es otra cosa que el espacio propio correspondiente el autovalor 0: $N(F) = \{v | F(v) = 0\} = \{v | F(v) = 0.v\} = V_F(0)$

10.5) El determinante de una matriz es igual al producto de *todos* sus autovalores (reales y complejos, y elevados a sus respectivas multiplicidades):

$$|A| = \lambda_1 \lambda_2 \ldots \lambda_n$$

En efecto, si $\lambda_1,\ldots,\lambda_n$ son las raíces de $P(\lambda)$, podemos escribir (utilizando la factorización en términos de raíces y notando que el término de grado n es $(-1)^n\lambda^n$)

$$P(\lambda)=|A-\lambda I_n|=(\lambda_1-\lambda)(\lambda_2-\lambda)\ldots(\lambda_n-\lambda) \qquad (10.2)$$

Por lo tanto, $|A|=P(0)=\lambda_1.\lambda_2\ldots\lambda_n$.
Esto implica que en un espacio de dimensión finita *el determinante de un operador lineal F es el producto de todos sus autovalores.*

10.6) La traza de una matriz es igual a la suma de todos sus autovalores (reales y complejos, y repetidos tantas veces como indica su multiplicidad):

$$\mathrm{Tr}A=\sum_{i=1}^{n}a_{ii}=\sum_{i=1}^{n}\lambda_i$$

A partir de la expresión (9.2) para $P(\lambda)$, vemos que el término de grado $n-1$ en λ es $(-\lambda)^{n-1}(\lambda_1+\ldots+\lambda_n)$, mientras que a partir de (9.1) vemos que el mismo es $(-\lambda)^{n-1}(a_{11}+\ldots+a_{nn})$. Como ambos son idénticos, se obtiene el resultado deseado.
Esto implica que la traza de un operador F es la suma de todos sus autovalores.

10.7 Diagonalización de Matrices

Una matriz A es diagonalizable si existe S no singular tal que

$$S^{-1}AS=A',\ \ \text{con}\ \ A'\ \text{diagonal}:\ A'=\begin{pmatrix}\lambda_1 & 0 & \ldots & 0\\ 0 & \lambda_2 & \ldots & 0\\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n\end{pmatrix}$$

es decir, si es semejante a una matriz diagonal.
En tal caso los elementos diagonales son necesariamente los autovalores de A, ya que

$$|A-\lambda I_n|=|A'-\lambda I_n|=(\lambda_1-\lambda)(\lambda_2-\lambda)\ldots(\lambda_n-\lambda)$$

y la columna i de S es autovector correspondiente al autovalor λ_i, pues $AS=SA'$:

$$S=\begin{pmatrix} & \ldots & \\ X_1 & \ldots & X_n\\ & \ldots & \end{pmatrix},\ \ \text{con}\ \ AX_i=\lambda_i X_i,\ i=1,\ldots,n,\ \ X_i=\begin{pmatrix}x_{1i}\\ \ldots\\ x_{ni}\end{pmatrix}$$

Análogamente, si existen n vectores columna X_i LI tales que $AX_i=\lambda_i X_i,\quad i=1,\ldots,n$ entonces A es diagonalizable, con S la matriz de columnas X_i (que será invertible pues los X_i son LI y por lo tanto $|S|\neq 0$)

La dimensión del espacio propio correspondiente al autovalor λ_i es la dimensión del espacio nulo de $|A-\lambda_i I_n|$:

$$d_i=\dim V(\lambda_i)=\dim N[A-\lambda_i I_n]=n-R(A-\lambda_i I_n)$$

donde R denota el rango.
Notemos que si $P(\lambda)$ posee n raíces distintas $\Rightarrow$ A es diagonalizable, pues en tal caso existirán n vectores columna X_i LI tales que $AX_i=\lambda_i X_i$.
Si A es diagonalizable, resulta evidente que $|A|=|A'|=\lambda_1\ldots\lambda_n$ y que $\mathrm{Tr}A=\mathrm{Tr}A'=\lambda_1+\ldots+\lambda_n$, ya que el determinante y la traza de matrices semejantes son idénticas ($|S^{-1}AS|=|A|$, $\mathrm{Tr}S^{-1}AS=\mathrm{Tr}ASS^{-1}=\mathrm{Tr}A$).

Ejemplo 1: Consideremos la matriz

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

que corresponde a la representación en la base canónica de la reflexión respecto de la recta $y=x$. Los autovalores se obtienen de la ecuación

$$|A-\lambda I_2|=\begin{vmatrix}-\lambda & 1\\ 1 & -\lambda\end{vmatrix}=1-\lambda^2=0$$

Por lo tanto, $\lambda = \pm 1$. El autovector X_1 correspondiente a $\lambda_1 = 1$ se obtiene resolviendo el sistema homogéneo $(A - \lambda_1 I_2)X_1 = 0$, es decir,

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

que conduce a $x = y$. Los autovectores son entonces de la forma $x\binom{1}{1}$, con $x \neq 0$ y $V(1)$ es el espacio generado por $\binom{1}{1}$. Este corresponde al espacio generado por e'_1 en el ej. 3 anterior ($[e'_1]_e = \binom{1}{1}$).
Notemos que la matriz anterior tiene rango 1, por lo que $\dim V(1) = 2 - 1 = 1$.
El autovector correspondiente a $\lambda_2 = -1$ se obtiene resolviendo el sistema $(A - \lambda_2 I_2)X_2 = 0$, es decir,

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

que conduce a $x = -y$. Los autovectores son entonces de la forma $x\binom{1}{-1}$, con $x \neq 0$, y $V(-1)$ es el espacio generado por $\binom{1}{-1}$. Este corresponde al espacio generado por e'_2 en el ej. 3 anterior ($[e'_2]_e = \binom{1}{-1}$).
Una matriz de autovectores es entonces

$$S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{con} \quad S^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

y se verifica

$$S^{-1}AS = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Notemos que se cumple $|A| = -1 = \lambda_1 \lambda_2$ y $\mathrm{Tr}A = 0 = \lambda_1 + \lambda_2$.

Ejemplo 2: Consideremos

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

La ec. característica es

$$|A - \lambda I_2| = \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 = 0$$

por lo que el único autovalor es $\lambda = 1$ con multiplicidad $m = 2$. No obstante, la matriz

$$A - 1I_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

posee rango 1, por lo que $\dim V(1) = 2 - 1 = 1 < 2$. Por lo tanto, esta matriz *no es diagonalizable*, ya que no existe una base de autovectores de la misma. La ecuación

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

conduce a $y = 0$, por lo que los autovectores son de la forma $x\binom{1}{0}$ y $V(1)$ es el espacio generado por $\binom{1}{0}$. No existe otro autovector LI de $\binom{1}{0}$. De todos modos, se cumple $|A| = 1 = 1.1$ y $\mathrm{Tr}A = 2 = 1 + 1$. Nótese que A es no singular ($|A| = 1 \neq 0$). La condición de no diagonalizable nada tiene que ver con la singularidad.

Cabe destacar, no obstante, que la matriz

$$B = \begin{pmatrix} 1 & 1 \\ \varepsilon & 1 \end{pmatrix}$$

es diagonalizable $\forall\, \varepsilon \neq 0$, ya que en tal caso la ecuación

$$|B - \lambda I_2| = (1-\lambda)^2 - \varepsilon = 0$$

posee siempre 2 raíces distintas: $\lambda = 1 \pm \sqrt{\varepsilon}$.
Esta conclusión es general: Si A no es diagonalizable podemos siempre encontrar una matriz B arbitrariamente próxima a A (es decir, cuyos elementos difieran de los de A en menos de ε, con $\varepsilon > 0$ arbitrario) tal que B es diagonalizable.

Ejemplo 3: Sea $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Tenemos

$$|A - \lambda I_3| = \begin{pmatrix} 1-\lambda & 1 & 1 \\ 0 & 2-\lambda & 1 \\ 0 & 0 & 1-\lambda \end{pmatrix} = (1-\lambda)(2-\lambda)(1-\lambda)$$

por lo que las raíces de $P(\lambda)$ son $\lambda_1 = 1$ y $\lambda_2 = 2$, con multiplicidades 2 y 1 respectivamente. Si $\lambda = \lambda_1$,

$$|A - 1I_3| = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

posee rango 1, por lo que $\dim V(1) = \dim N(A - 1I_3) = 3 - 1 = 2$, igual a la multiplicidad de λ_1. La matriz es por lo tanto *diagonalizable* ya que necesariamente $\dim V(2) = 1$. El sistema $(A - 1I_3)X = 0$ conduce a $y + z = 0$, es decir, $y = -z$, con z y x arbitrarios, por lo que los autovectores para $\lambda_1 = 1$ son de la forma $\begin{pmatrix} x \\ -z \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$. Para $\lambda_2 = 2$,

$$|A - 2I_3| = \begin{pmatrix} -1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

posee rango 2. El sistema $(A - 2I_3)X = 0$ conduce a $x = y$, con $z = 0$, por lo que los autovectores son de la forma $x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$. Una matriz de autovectores es por lo tanto

$$S = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \text{con} \quad S^{-1} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$

Se verifica entonces

$$A' = S^{-1}AS = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Nótese que el orden de los autovalores en A' corresponde al orden de los autovectores (columnas) en S.

Ejemplo 4: (Para hacer en la práctica) Sea $F : \mathbb{R}^2 \to \mathbb{R}^2$ el operador de rotación de ángulo θ (antihorario), con $[F]_e = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$, siendo e la base canónica. Es obvio que F no puede tener autovalores reales, ya que no existe $v \neq 0$ tal que $F(v) = \lambda v$. No es por lo tanto diagonalizable para $K = \mathbb{R}$.
Sin embargo, $[F]_e$ tiene autovalores complejos $\lambda = e^{\pm i\theta} = \cos\theta \pm i \sin\theta$ y es por lo tanto *diagonalizable* si se lo considera como $F : \mathbb{C}^2 \to \mathbb{C}^2$, con $K = \mathbb{C}$. Determine el lector los autovectores y compruebe que existe S tal que $S^{-1}[F]_e S$ es diagonal!

Ejemplo 5: Consideremos el operador de proyección ortogonal P sobre el plano $x + y = 0$ en $\mathbb{R}^3$ (o sea, proyección sobre este plano en la dirección del eje z). Si $e' = (e'_1, e'_2, e'_3)$, con $e'_1 = (e_1 + e_2)/\sqrt{2}$, $e'_2 = (-e_1 + e_2)/\sqrt{2}$, $e'_3 = e_3$, con $e = (e_1, e_2, e_3)$ la base canónica, entonces $P(e'_1) = 0$, $P(e'_2) = e'_2$, $P(e'_3) = e'_3$ y por lo tanto, $[P]_{e'} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Esta es pues una base de autovectores de P.
En la base canónica, obtenemos en cambio

$$[P]_e = S[P]_{e'}S^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \text{con} \quad S = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} / \sqrt{2}$$

y $S^{-1}=S^t$. En esta base $[P]_e$ no es diagonal, aunque sigue cumpliendo que $[P]_e^2=[P]_e$. Se deja como ejercicio verificar explícitamente que los autovalores de la matriz $[P]_e$ son 0 y 1, y que una base de autovectores es precisamente e' (aunque por su puesto no es la única), de modo que S es una matriz diagonalizante de $[P]_e$, que verifica $S^{-1}[P]_eS=[P]_{e'}$, con $[P]_{e'}$ diagonal.

Ejemplo 6: **Autovalores de una matrix general de** 2×2. Si $A=\begin{pmatrix} a & b\\ c & d\end{pmatrix}=\frac{a+d}{2}I_2+\begin{pmatrix}\frac{a-d}{2} & b\\ c & -\frac{a-d}{2}\end{pmatrix}$, obtenemos fácilmente, a partir de $|A-\lambda I_2|=0$, que los autovalores son

$$\lambda_\pm=\frac{a+d}{2}\pm\sqrt{(\frac{a-d}{2})^2+bc}=\frac{1}{2}\mathrm{Tr}[A]\pm\sqrt{(\frac{\mathrm{Tr}[A]}{2})^2-\mathrm{Det}[A]}$$

donde $\mathrm{Tr}[A]=a+d$ es la traza de A y $\mathrm{Det}[A]=ad-bc$ su determinante. La última expresión puede obtenerse directamente de resolver el sistema $\left\{\begin{array}{rcl}\lambda_++\lambda_- & = & \mathrm{Tr}[A]\\ \lambda_+\lambda_- & = & \mathrm{Det}[A]\end{array}\right.$
Los dos autovalores quedan pues completamente determinados por la traza y el determinante.

Ejemplo 7: **Corrección de primer orden en los autovalores**.
Consideremos una matriz $B=A+\delta A$ de $n\times n$, siendo A diagonalizable y $\delta A=\varepsilon M$ una perturbación (ε es un parámetro suficientemente pequeño y M una matriz de $n\times n$ arbitraria).
Sea $S=(X_1,\ldots,X_n)$ una matriz de autovectores de A, tal que $S^{-1}AS=A'$ con A' diagonal ($A'_{ij}=\lambda_i\delta_{ij}$, con $AX_i=\lambda_iX_i$). Definiendo $\delta A'=S^{-1}(\delta A)S=\varepsilon S^{-1}MS$ (perturbación en la base en que A es diagonal) y notando que $|B-\lambda I|=|S^{-1}BS-\lambda I|$, obtenemos

$$|B-\lambda I|=|D+\delta A'-\lambda I|=\begin{vmatrix}\lambda_1+\delta A'_{11}-\lambda & \delta A'_{12} & \ldots & \delta A'_{1n}\\ \delta A'_{21} & \lambda_2+\delta A'_{22}-\lambda & \ldots & \delta A'_{2n}\\ & & \ldots & \\ \delta A'_{n1} & \delta A'_{n2} & \ldots & \lambda_n+\delta A'_{nn}-\lambda\end{vmatrix}$$

Consideremos primero el caso en que los autovalores λ_i de A son todos distintos. Para $\lambda=\lambda_i+\delta\lambda_i$ con $\delta\lambda_i$ una corrección de orden ε al autovalor λ_i, obtenemos entonces

$$|B-(\lambda_i+\delta\lambda_i)I|=(\delta A'_{ii}-\delta\lambda_i)\prod_{j\neq i}(\lambda_j-\lambda_i)+O(\varepsilon^2)$$

donde el primer término es el de mayor orden ($O(\varepsilon)$), y los restantes de orden $O(\varepsilon^2)$ o mayor. Por lo tanto, la ec. $|B-\lambda I|=0$ conduce a

$$\delta\lambda_i=\delta A'_{ii}+O(\varepsilon^2),\quad \text{con } \delta A'_{ii}=(S^{-1}\delta AS)_{ii}=\varepsilon\sum_{j,k}S^{-1}_{ij}M_{jk}S_{ki}$$

es decir, los $\delta\lambda_i$ son los términos diagonales de δA en la base en que A es diagonal. Como $\sum_j S^{-1}_{ij}S_{ji}=1$, si la columna i de S (el autovector X_i) se multiplica por α, la fila de i de S^{-1} se multiplica por $1/\alpha$, para mantener la igualdad anterior. Por lo tanto, la corrección $\delta A'_{ii}$ es, como debe ser, independiente de la base elegida del espacio propio, es decir de la elección del autovector $X_i\neq 0$ en el espacio propio.

En el caso general, si el espacio propio asociado a un autovalor λ_i tiene dimensión d_i (se dice entonces que tiene degeneración d_i), la corrección a λ_i son los *autovalores* de δA en el espacio propio asociado a λ_i, pues $|A-\lambda I|=|(\delta A')_i-\delta\lambda_iI_{d_i}|\prod_{\lambda_j\neq\lambda_i}(\lambda_j-\lambda_i)^{m_j}+O(\varepsilon^{d_i+1})$, con $\delta A'_i$ la matriz $\delta A'$ restringida al espacio propio asociado a λ_i y m_j la multiplicidad (algebraica) del autovalor λ_j. Se deben pues obtener los autovalores de $\delta A'_i$ (matriz de $d_i\times d_i$). El nivel degenerado λ_i se desdobla normalmente en varios niveles. Se dice entonces que se rompe la degeneración.
Es importante que A sea diagonalizable. De lo contrario, el ej. 2 anterior muestra que en el caso no-diagonalizable, la corrección puede ser por ej. de orden $\sqrt{\varepsilon}$.

10.8 Evaluación de Potencias y Series de Matrices

La diagonalización es muy conveniente para evaluar potencias y series de matrices (de $n \times n$) u operadores. En primer lugar, si

$$A = SA'S^{-1} \qquad (10.3)$$

(A semejante a A') se cumple, para k natural,

$$A^k = SA'^k S^{-1}$$

ya que $A^2 = (SA'S^{-1})(SA'S^{-1}) = SA'^2 S^{-1}$ y en general (por inducción) $A^k = AA^{k-1} = SA'S^{-1}SA'^{k-1}S^{-1} = SA'^k S^{-1}$. Análogamente, para funciones definidas por series de potencias $f(u) = \sum_{k=0}^{\infty} a_k u^k$ convergentes $\forall\ u \in \mathbb{C}$,

$$f(A) = \sum_{k=0}^{\infty} a_k A^k = \sum_{k=0}^{\infty} a_k SA'^k S^{-1} = S[\sum_{k=0}^{\infty} a_k A'^k]S^{-1} = Sf(A')S^{-1}$$

Notemos que $f(A)$ está bien definido pues $|(A^k)_{ij}| \leq (mn)^k/n$, donde m el mayor elemento de la matriz ($|A_{ij}| \leq m\ \forall\ i,j$) y n la dimensión. Esto implica que la serie matricial converge absolutamente si la serie converge absolutamente $\forall\ u$ ($|(f(A))_{ij}| \leq f(mn)/n$). En particular,

$$\exp[At] = \sum_{k=0}^{\infty} \frac{(At)^k}{k!} = S \exp[A't]S^{-1}$$

Finalmente, si A es invertible (en cuyo caso A' es también invertible, como el lector debe reconocer inmediatamente) se cumple $A^{-1} = SA'^{-1}S^{-1}$ y en general

$$A^{-k} = SA'^{-k}S^{-1}$$

Si A es *diagonalizable*, $S^{-1}AS = A'$ con A' diagonal. Por lo tanto $A = SA'S^{-1}$. Podemos entonces utilizar las expresiones anteriores con A' diagonal y S una matriz de autovectores, en cuyo caso la evaluación resulta inmediata pues

$$A' = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \quad \Rightarrow \quad (A')^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_n^k \end{pmatrix}$$

para k natural. Esto implica

$$f(A') = \sum_{k=0}^{\infty} a_k (A')^k = \begin{pmatrix} f(\lambda_1) & 0 & \dots & 0 \\ 0 & f(\lambda_2) & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & f(\lambda_n) \end{pmatrix}$$

En particular,

$$\exp[A't] = \begin{pmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & e^{\lambda_n t} \end{pmatrix}$$

Además, si A es invertible, sus autovalores son todos no nulos y es fácil ver que

$$(A')^{-1} = \begin{pmatrix} \lambda_1^{-1} & 0 & \dots & 0 \\ 0 & \lambda_2^{-1} & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_n^{-1} \end{pmatrix}$$

y por lo tanto

$$(A')^{-k} = \begin{pmatrix} \lambda_1^{-k} & 0 & \dots & 0 \\ 0 & \lambda_2^{-k} & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_n^{-k} \end{pmatrix}$$

Por ejemplo, en el caso del ej. 1 anterior se obtiene

$$\exp[At]=\exp[\begin{pmatrix}0&1\\1&0\end{pmatrix}t]=\frac{1}{2}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\begin{pmatrix}e^t&0\\0&e^{-t}\end{pmatrix}\begin{pmatrix}1&1\\1&-1\end{pmatrix}=\begin{pmatrix}\cosh(t)&\sinh(t)\\\sinh(t)&\cosh(t)\end{pmatrix}$$

y en el ej. 3 anterior,

$$A^n=S\begin{pmatrix}1^n&0&0\\0&1^n&0\\0&0&2^n\end{pmatrix}S^{-1}=\begin{pmatrix}1&2^n-1&2^n-1\\0&2^n&2^n-1\\0&0&1\end{pmatrix}$$

$$\exp[At]=S\begin{pmatrix}e^t&0&0\\0&e^t&0\\0&0&e^{2t}\end{pmatrix}S^{-1}=\begin{pmatrix}e^t&e^{2t}-e^t&e^{2t}-e^t\\0&e^{2t}&e^{2t}-e^t\\0&0&e^t\end{pmatrix}$$

Ejemplo: **Sucesión de Fibonacci**. Está definida por la relación recursiva *lineal*

$$a_{n+1}=a_n+a_{n-1},\quad n\geq 1$$

con $a_0=0$, $a_1=1$.
La expresión explícita de a_n puede obtenerse fácilmente planteando el problema en forma matricial. Resolveremos en realidad el problema para valores iniciales generales a_0, a_1. Tenemos, para $n\geq 1$,

$$\begin{pmatrix}a_{n+1}\\a_n\end{pmatrix}=\begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}a_n\\a_{n-1}\end{pmatrix}$$

Por lo tanto, para $n\geq 1$ y definiendo $A=\begin{pmatrix}1&1\\1&0\end{pmatrix}$,

$$\begin{pmatrix}a_{n+1}\\a_n\end{pmatrix}=A^n\begin{pmatrix}a_1\\a_0\end{pmatrix}$$

La evaluación de A^n puede efectuarse mediante su diagonalización. Los autovalores de A son los números aureos

$$\lambda_\pm=\frac{1\pm\sqrt{5}}{2}$$

que satisfacen $\lambda^2=\lambda+1$, con autovectores $v_\pm\propto\binom{\lambda_\pm}{1}$. Podemos entonces escribir $A=SA'S^{-1}$ con $S=\begin{pmatrix}\lambda_+&\lambda_-\\1&1\end{pmatrix}$ y $A'=\begin{pmatrix}\lambda_+&0\\0&\lambda_-\end{pmatrix}$. Por lo tanto, $A^n=S(A')^nS^{-1}$ y se obtiene finalmente (se dejan las cuentas para el lector)

$$a_n=[(\lambda_+^n-\lambda_-^n)a_1-(\lambda_+^n\lambda_--\lambda_-^n\lambda_+)a_0]/\sqrt{5}$$

En el caso usual de Fibonacci, $a_0=0$, $a_1=1$ y $a_n=(\lambda_+^n-\lambda_-^n)/\sqrt{5}$. Como $\lambda_+=1.618$, $\lambda_-=-0.618$, el término dominante para n grande es el proporcional a λ_+^n.

Un tratamiento equivalente consiste en expresar el vector inicial $\binom{a_1}{a_0}$ como combinación lineal de los autovectores de A: $A^n\binom{a_1}{a_0}=A^n[c_+\binom{\lambda_+}{1}+c_-\binom{\lambda_-}{1}]=c_+\lambda_+^n\binom{\lambda_+}{1}+c_-\lambda_-^n\binom{\lambda_-}{1}$, de donde $a_n=\lambda_+^nc_++\lambda_-^nc_-$. Como $\binom{c_+}{c_-}=S^{-1}\binom{a_1}{a_0}$, se obtiene $c_+=a_1-\lambda_-a_0$, $c_-=-a_1+\lambda_+a_0$, obteniéndose el resultado anterior.

El mismo método se puede aplicar para toda sucesión definida por una relación recursiva fija *lineal*:

$$a_{n+1}=\alpha_0a_n+\alpha_1a_{n-1}+\ldots+\alpha_ka_{n-k}$$

para $n\geq k$, con $a_0,\ldots,a_k$ dados, que conduce a

$$\begin{pmatrix}a_{n+1}\\a_n\\\ldots\\a_{n-k+1}\end{pmatrix}=\begin{pmatrix}\alpha_0&\alpha_1&\ldots&\alpha_k\\1&0&\ldots&0\\0&1&\ldots&0\\&&\ldots&\\0&\ldots&1&0\end{pmatrix}\begin{pmatrix}a_n\\a_{n-1}\\\ldots\\a_{n-k}\end{pmatrix}=\begin{pmatrix}\alpha_0&\alpha_1&\ldots&\alpha_k\\1&0&\ldots&0\\0&1&\ldots&0\\&&\ldots&\\0&\ldots&1&0\end{pmatrix}^{n-k+1}\begin{pmatrix}a_k\\a_{k-1}\\\ldots\\a_0\end{pmatrix}$$

La sucesión geométrica elemental $a_n=\alpha_0^na_0$ corresponde al caso $k=0$ ($a_{n+1}=\alpha_0a_n$ si $n\geq 0$).

10.9 Desacoplamiento de Sistemas de Ecuaciones Diferenciales Lineales

Como otra aplicación, consideremos por ejemplo el sistema de ecuaciones diferenciales ordinarias lineales de primer orden

$$\frac{dX}{dt} = AX$$

con X de $n \times 1$ y A de $n \times n$, con elementos *constantes* (o sea, independientes del tiempo). Suponiendo A diagonalizable, tenemos $A = SA'S^{-1}$, con A' diagonal y S la matriz de autovectores. Por lo tanto $dX/dt = SA'S^{-1}X$, lo que implica

$$dX'/dt = A'X', \quad X' = S^{-1}X,$$

Como A' es diagonal, el sistema en las variables X' está *desacoplado*, y es de fácil resolución. Tenemos, para las componentes x'_i de X', las ecuaciones desacopladas

$$dx'_i/dt = \lambda_i x'_i, \quad i = 1, \ldots, n$$

donde λ_i son los autovalores de A, cuya solución es $x'_i = c_i e^{\lambda_i t}$. Finalmente, se obtiene

$$X(t) = SX'(t) = \sum_{i=1}^{n} c_i V_i e^{\lambda_i t},$$

donde V_i denota los autovectores de A (las columnas de S). Esto constituye la solución *general* del sistema de primer orden, conteniendo n constantes arbitrarias c_i que pueden determinarse a partir de las condiciones iniciales $x_i(0)$.

El procedimiento usualmente utilizado en Física e Ingeniería para llegar a esta solución es plantear una solución del tipo $X(t) = Ve^{\lambda t}$ con V constante. La ec. $dX/dt = AX$ implica entonces $\lambda V = AV$, por lo que V debe ser autovector de A con autovalor λ. La solución general se obtiene luego como combinación lineal arbitraria de estas soluciones particulares. Este procedimiento es en realidad correcto para encontrar la solución general sólo en el caso de matrices A diagonalizables.

Nótese también que el mismo método puede utilizarse para resolver sistemas análogos de segundo orden

$$\frac{d^2X}{dt^2} = AX$$

Sólo es necesario reemplazar $c_i e^{\lambda_i t}$ por $c_i^+ e^{\sqrt{\lambda_i} t} + c_i^- e^{-\sqrt{\lambda_i} t}$ en la solución general anterior.

Ejemplo 1: Consideremos el sistema de tres ecuaciones diferenciales acopladas de primer orden,

$$dx/dt = x + y + z, \quad dy/dt = 2y + z, \quad dz/dt = z$$

donde x, y, z son funciones de t, el cual puede escribirse en forma matricial como

$$\frac{d}{dt}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

o sea, $dv/dt = Av$, siendo A la matriz del ej. 3 anterior y $v = (x, y, z)^t$. Por lo tanto, utilizando las matrices S y S^{-1} de dicho ejemplo,

$$\frac{d}{dt}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = S\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}S^{-1}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

y mult. a izq. por S^{-1}, se llega a

$$\frac{d}{dt}\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}, \quad \text{con} \quad \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = S^{-1}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y - z \\ -z \\ y + z \end{pmatrix}$$

el cual es un sistema de tres ecuaciones dif. lineales *desacopladas*:

$$dx'/dt = x', \quad dy'/dt = y', \quad dz'/dt = 2z'$$

que es equivalente al original. La solución del sistema desacoplado es muy fácil de obtener:

$$x' = c_1 e^t, \; y' = c_2 e^t, \; z' = c_3 e^{2t}$$

Finalmente

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = S \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x' + z' \\ -y' + z' \\ y' \end{pmatrix} = c_1 e^t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 e^t \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} + c_3 e^{2t} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 e^t + c_3 e^{2t} \\ -c_2 e^t + c_3 e^{2t} \\ c_2 e^t \end{pmatrix}$$

Ejemplo 2: Sistema de dos resortes acoplados (recordar dibujo). Ecuación de movimiento:

$$\frac{d^2}{dt^2} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -\frac{1}{m} \begin{pmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_1 + k_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Resuelto en clase. Detalles a cargo del lector. Sólo recordamos que las frecuencias propias $\omega_i = \sqrt{\lambda_i}$ (con λ_i los autovalores de la matriz $\left(\begin{smallmatrix} k_1+k_2 & -k2 \\ -k2 & k_1+k_2 \end{smallmatrix}\right)/m$ son $\omega_1 = \sqrt{(k_1 + 2k_2)/m}$, $\omega_2 = \sqrt{k_1/m}$, con $V_1 \propto \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$, $V_2 \propto \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$.

11. Matrices hermíticas y reales simétricas

Un caso especial muy importante para la física es aquel de matrices de $n \times n$ *hermíticas* (o hermitianas o autoadjuntas), que son las que satisfacen $A^\dagger = A$, donde $A^\dagger \equiv A^{t*}$ denota la matriz traspuesta y conjugada (matriz adjunta):

$$A = A^\dagger \;\Leftrightarrow\; A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{12}^* & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{1n}^* & a_{2n}^* & \ldots & a_{nn} \end{pmatrix}$$

con a_{ii} real para $i = 1, \ldots, n$. Si todos los elementos de A son *reales* $\Rightarrow$ $A^\dagger = A^t$ y la condición de A hermítica equivale a A *simétrica* ($A^t = A$).

11.1) *Teorema:* Si A de $n \times n$ es una matriz hermítica sus autovalores λ_i son todos reales y los autovectores X_i correspondientes a autovalores distintos son ortogonales respecto del producto escalar usual para vectores complejos: Si $AX_i = \lambda_i X_i$, $AX_j = \lambda_j X_j \Rightarrow X_i^\dagger X_j = 0$ si $\lambda_i \neq \lambda_j$, donde

$$X_i^\dagger X_j = (x_{1i}^* \ldots x_{ni}^*) \begin{pmatrix} x_{1j} \\ \ldots \\ x_{nj} \end{pmatrix} = x_{1i}^* x_{1j} + \ldots x_{ni}^* x_{nj}$$

Demostración: Sea X_i de $n \times 1$ autovector de A con autovalor λ_i (por lo tanto $X_i \neq 0$). Multiplicando la igualdad $AX_i = \lambda_i X_i$ a izquierda por $X_i^\dagger = X_i^{t*}$ se obtiene

$$X_i^\dagger A X_i = \lambda_i X_i^\dagger X_i \qquad (11.1)$$

con

$$X_i^\dagger X_i = (x_{1i}^* \ldots x_{ni}^*) \begin{pmatrix} x_{1i} \\ \ldots \\ x_{ni} \end{pmatrix} = x_{1i}^* x_{1i} + \ldots x_{ni}^* x_{ni} = |x_{1i}|^2 + \ldots + |x_{ni}|^2 > 0$$

Trasponiendo y conjugando la igualdad (11.1) se obtiene, notando que $(AB)^\dagger = B^\dagger A^\dagger$,

$$X_i^\dagger A^\dagger X_i = \lambda_i^* X_i^\dagger X_i \qquad (11.2)$$

Pero como $A^\dagger = A$, esto implica, comparando con (11.1), que $\lambda_i X_i^\dagger X_i = \lambda_i^* X_i^\dagger X_i$, o sea,

$$0 = (\lambda_i - \lambda_i^*) X_i^\dagger X_i$$

Como $X_i^\dagger X_i > 0 \Rightarrow \lambda_i = \lambda_i^*$, es decir, λ_i real.
Del mismo modo, si $AX_i = \lambda_i X_i$, $AX_j = \lambda_j X_j$, multiplicando a izquierda la primer ecuación por $X_j^\dagger$ y la segunda por $X_i^\dagger$ se obtiene

$$X_j^\dagger A X_i = \lambda_i X_j^\dagger X_i, \quad X_i^\dagger A X_j = \lambda_j X_i^\dagger X_j$$

Trasponiendo y conjugando la pimera de estas ecuaciones se obtiene $X_i^\dagger A^\dagger X_j = \lambda_i^* X_i^\dagger X_j$, es decir, $X_i^\dagger A X_j = \lambda_i X_i^\dagger X_j$ pues $A^\dagger = A$ y $\lambda_i = \lambda_i^*$ (ya demostrado). Por lo tanto, $\lambda_i X_i^\dagger X_j = \lambda_j X_i^\dagger X_j$, o sea,

$$0 = (\lambda_i - \lambda_j) X_i^\dagger X_j$$

por lo que $X_i^\dagger X_j = 0$ si $\lambda_i \neq \lambda_j$.

Puede demostrarse (lo haremos más adelante) que *toda* matriz hermítica A es *diagonalizable*, y que siempre existen n autovectores X_i LI y además ortogonales que satisfacen $AX_i = \lambda_i X_i$, con $X_i^\dagger X_j = 0$ si $i \neq j$ (aún cuando $\lambda_i = \lambda_j$).

En tal caso, eligiendo autovectores normalizados tales que $X_i^\dagger X_i = 1$ para $i = 1, \ldots, n$, se obtiene

$$X_i^\dagger X_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Por lo tanto la matriz de autovectores $S = (X_1 \ldots X_n)$ satisface

$$S^\dagger S = I_n$$

pues $(S^\dagger S)_{ij} = X_i^\dagger X_j = \delta_{ij}$. La inversa de S es pues directamente la matriz adjunta: $S^{-1} = S^\dagger$.

En resumen, si $A = A^\dagger$, $\exists$ S tal que $S^{-1} = S^\dagger$ y $S^\dagger A S = A'$, con A' diagonal y real: $A_{ij} = \lambda_i \delta_{ij}$

Matrices antihermíticas: Si $A^\dagger = -A$, se dice que A es antihermítica. En tal caso, $B = -iA$ resulta hermítica (pues $B^\dagger = iA^\dagger = -iA = B$), lo que implica, como $A = iB$, que A será también diagonalizable, con autovectores ortogonales si corresponden a autovalores distintos, pero con autovalores *imaginarios* en lugar de reales: Si $BX_i = \lambda_i X_i \Rightarrow AX_i = (i\lambda_i) X_i$.

Matrices reales simétricas: Para el caso particular de *matrices reales*, los resultados anteriores implican que *los autovalores de matrices reales simétricas ($A^\dagger = A^t = A$) son todos reales.* Los autovectores pueden pués elegirse reales, y por lo tanto, serán ortogonales respecto del producto escalar usual: Si $AX_i = \lambda_i X_i$ y $AX_j = \lambda_j X_j \Rightarrow$

$$X_i^t X_j = x_{1i} x_{1j} + \ldots + x_{ni} x_{nj} = 0 \text{ si } \lambda_i \neq \lambda_j$$

En tal caso, eligiendo autovectores normalizados (tales que $X_i^t X_i = 1$) la inversa de la matriz $S = (X_1, \ldots, X_n)$ será directamente la traspuesta:

$$S^t S = I_n \qquad (A = A^t \text{ real}, \ X_i \text{ real para } i = 1, \ldots, n)$$

En resumen, si $A = A^t$, con A real, $\exists$ S real tal que $S^{-1} = S^t$ y $S^t A S = A'$, con A' diagonal y real: $A'_{ij} = \lambda_i \delta_{ij}$. Más aun, puede elegirse siempre S tal que $\det S = +1$ (Si $S^{-1} = S^t \Rightarrow \det S = \pm 1$) en cuyo caso S corresponde a una rotación, como veremos más adelante

Matrices reales antisimétricas: Si A es real y $A^t = -A$, nuevamente $B = -iA$ resulta hermítica ($B^\dagger = iA^t = -iA = B$) y por lo tanto, $A = iB$ será diagonalizable en el cuerpo de los complejos, con autovalores *imaginarios* y autovectores ortogonales si corresponden a autovalores distintos.

Ejemplo 1: Sea

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

A es una matriz real simétrica si v es real. Tenemos

$$|A - \lambda I_2| = (1 - \lambda)^2 - v^2$$

por lo que los autovalores son $\lambda = 1 \pm v$, reales. Para $\lambda_1 = 1 + v$, puede verse fácilmente que el autovector es de la forma $X_1 = x_1 \binom{1}{1}$, mientras que para $\lambda_2 = 1 - v$, es de la forma $X_2 = x_2 \binom{-1}{1}$. Por lo tanto, podemos elegir $x_1 = x_2 = 1/\sqrt{2}$, para que $X_1^t X_1 = X_2^t X_2 = 1$. Se verifica además $X_1^t X_2 = \frac{1}{2}(1, 1)\binom{-1}{1} = 0$. Por lo tanto

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad S^{-1} = S^t = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

con

$$S^t A S = \begin{pmatrix} 1+v & 0 \\ 0 & 1-v \end{pmatrix}$$

Ejemplo 2: Sea

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

con v real. A es una matriz hermítica ($A^\dagger = A$). Tenemos

$$|A - \lambda I_2| = (1-\lambda)^2 - (iv)(-iv) = (1-\lambda^2) - v^2$$

por lo que los autovalores son nuevamente $\lambda = 1 \pm v$, reales. Para $\lambda_1 = 1+v$, puede verse fácilmente que el autovector es de la forma $X_1 = x_1\binom{i}{1}$, mientras que para $\lambda_2 = 1-v$, es de la forma $X_2 = x_2\binom{-i}{1}$. Por lo tanto, podemos elegir $x_1 = x_2 = 1/\sqrt{2}$, para que $X_1^\dagger X_1 = X_2^\dagger X_2 = 1$. Se verifica además $X_1^\dagger X_2 = \frac{1}{2}(-i, 1)\binom{-i}{1} = (-1+1)/2 = 0$. Por lo tanto

$$S = \frac{1}{\sqrt{2}}\begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}, \quad S^{-1} = S^\dagger = \frac{1}{\sqrt{2}}\begin{pmatrix} -i & 1 \\ i & 1 \end{pmatrix}$$

con

$$S^\dagger A S = \begin{pmatrix} 1+v & 0 \\ 0 & 1-v \end{pmatrix}$$

Ejemplo 3: Consideremos la ecuación

$$ax^2 + 2bxy + cy^2 = d$$

con coeficientes y variables reales. Podemos escribirla en forma matricial como

$$X^t A X = d, \quad X = \binom{x}{y}, \quad X^t = (x, y), \quad A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

La matriz A es real y simétrica, siendo por lo tanto siempre diagonalizable. Existe entonces una matriz ortogonal de autovectores S ($S^{-1} = S^t$), con $\mathrm{Det}\, S = 1$, tal que $S^{-1}AS = S^t AS = A'$, con A' diagonal: $A' = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}$, siendo $\lambda_\pm$ los autovalores de A (las raíces de $(a-\lambda)(b-\lambda) - b^2 = 0$). En tal caso, si $X = SX'$, tenemos

$$X^t A X = X'^t S^t A S X' = X'^t A' X' = \lambda_+ x'^2 + \lambda_- y'^2 = d$$

Si $d > 0$, vemos entonces que la gráfica de la ecuación en las variables x', y' será una elipse si $\lambda_\pm$ son ambos mayores que 0, y una hipérbola si $\lambda_+\lambda_- < 0$, con ejes principales x', y' en ambos casos. Como la transformación corresponde a una rotación (eligiendo el orden de autovectores tal que $\mathrm{Det}S = +1$), la ecuación original corresponderá si $|A| \neq 0$ a una elipse o hipérbola con ejes principales rotados (como consecuencia del término cruzado $2bxy$). El ángulo de inclinación θ entre los ejes x' y x puede obtenerse a partir de la matriz S escribiéndola en la forma $S = [I]_e^{e'} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Para más detalles ver ejemplo resuelto en clase o en práctica.

Ejemplo 4: **Tensor de Inercia**. El tensor de inercia de un cuerpo rígido respecto de un origen O es

$$I_O = \sum_\nu m_\nu \begin{pmatrix} r_\nu^2 - x_\nu^2 & -x_\nu y_\nu & -x_\nu z_\nu \\ -y_\nu x_\nu & r_\nu^2 - y_\nu^2 & -y_\nu z_\nu \\ -z_\nu x_\nu & -z_\nu y_\nu & r_\nu^2 - z_\nu^2 \end{pmatrix} = \sum_\nu m_\nu (X_\nu^t X_\nu I_3 - X_\nu X_\nu^t)$$

donde $X_\nu^t = (x_\nu, y_\nu, z_\nu)$ y $r_\nu^2 = X_\nu^t X_\nu = x_\nu^2 + y_\nu^2 + z_\nu^2$. I_O queda pues representado por una matriz real *simétrica*. Frente a una rotación del sistema de coordenadas, $X_\nu = SX'_\nu$, con $\mathrm{Det}S = 1$, $S^t S = I_3$, se obtiene

$$I_O = \sum_\nu m_\nu (X'^t_\nu S^t S X'_\nu I_3 - SX'_\nu X'^t_\nu S^t) = S[\sum_\nu m_\nu (X'^t_\nu X'_\nu I_3 - X'_\nu X'^t_\nu)]S^t = SI'_O S^t$$

o sea, $I'_O = S^t I_O S$ con I'_O el tensor de inercia en el sistema rotado. Como I_O es real simétrica, existirá una matriz ortogonal de rotación S (matriz de autovectores normalizados y ordenados tal que $S^t S = I_3$ y $|S| = 1$) tal que I'_O sea *diagonal*. Esta matriz determinará los 3 *ejes principales de inercia*, y los autovalores de I_O serán los *momentos principales de inercia*. Si el vector velocidad angular Ω coincide con alguna de estas direcciones, el vector momento angular (dado en general por $L_O = I_O\Omega$) será proporcional a Ω.

Ejemplo 5: **Sistema general de n resortes acoplados**. El movimiento de tal conjunto esta descripto por un sistema de ecuaciones de segundo orden del tipo (recordar discusión de clase)

$$\frac{m_i d^2 x_i}{dt^2} = -\sum_{j=1}^{n} k_{ij} x_j, \quad i = 1, \ldots, n$$

donde x_i es la posición de la partícula i (medida desde la posición de equilibrio), $m_i > 0$ su masa y $k_{ij} = k_{ji}$. Podemos reescribir tal sistema en forma matricial como

$$M \frac{d^2 X}{dt^2} = -KX$$

donde M es una matriz diagonal de elementos m_i ($M_{ij} = m_i \delta_{ij}$), $X = (x_1, \ldots, x_n)^t$ y K la matriz de elementos k_{ij}. Definiendo la matriz diagonal $M^{1/2}$ de elementos $\sqrt{m_i}$ ($(M^{1/2})_{ij} = \sqrt{m_i}\delta_{ij}$) podemos reescribir tal sistema como $M^{1/2} M^{1/2} \frac{d^2 X}{dt^2} = -KX$, y por lo tanto, multiplicando a izquierda por $M^{-1/2} = (M^{1/2})^{-1}$ (matriz diagonal de elementos $(M^{-1/2})_{ij} = \frac{1}{\sqrt{m_i}}\delta_{ij}$), como

$$\frac{d^2 Y}{dt^2} = -\tilde{K} Y, \text{ donde } \tilde{K} = M^{-1/2} K M^{-1/2}, \;\; Y = M^{1/2} X$$

(de forma que $y_i = \sqrt{m_i} x_i$). La ventaja esta forma matricial es que la matriz $\tilde{K}$ es real **simétrica** ($\tilde{K}^t = \tilde{K}$) y por lo tanto siempre **diagonalizable**. Existe entonces una matriz ortogonal S tal que $S^t \tilde{K} S = \tilde{K}'$, con K' **diagonal**, de elementos $\tilde{K}'_{ij} = \lambda_i \delta_{ij}$, siendo λ_i los autovalores de $\tilde{K}$. Por lo tanto, escribiendo $\tilde{K} = S \tilde{K}' S^t$, el sistema original resulta equivalente a

$$\frac{d^2 Y'}{dt^2} = -\tilde{K}' Y', \quad \text{donde } Y' = S^t Y$$

Esto representa, dado que $\tilde{K}'$ es diagonal, un sistema de n resortes **desacoplados**:

$$\frac{d^2 y'_i}{dt^2} = -\lambda_i y'_i, \;\; i = 1, \ldots, n$$

La solución general de c/u de estas ecuaciones es, para $\lambda_i \neq 0$, $y'_i(t) = Ae^{i\omega_i t} + Be^{i\omega_i t} = C\cos(\omega_i t + \phi)$, donde $\omega_i = \sqrt{\lambda_i}$ **son las frecuencias propias de vibración del sistema**. Las variables $y'_i = \sum_{j=1}^n S_{ji}\sqrt{m_j} x_j$ (o sea, $y'_i = (S^t Y)_i$) se denominan **modos normales de vibración**. Notemos que las frecuencias propias son las raíces de los autovalores de la matriz $M^{-1/2} K M^{-1/2}$, los cuales, en virtud de la propiedad 12.1 siguiente, coinciden con los de la matriz $M^{-1}K$. Por lo tanto, la conocida fórmula $\omega = \sqrt{k/m}$ para la frecuencia angular de un oscilador armónico se generaliza a $\omega_i = \sqrt{(M^{-1}K)_i}$, donde $(M^{-1}K)_i$ denota aquí el iésimo autovalor de la matriz $M^{-1}K$. Puede demostrarse que si la matriz K es definida positiva (definición que veremos luego y que corresponde a un sistema estable) entonces todos los autovalores de $\tilde{K}$ son positivos.

Ejemplo 6: **Problema generalizado de autovalores:** La ecuación

$$AX = \lambda BX$$

donde A y B son matrices de $n \times n$, B es no singular y $X \neq 0$ es un vector columna, define un problema generalizado de autovalores. Es obviamente equivalente al problema estándar $B^{-1}AX = \lambda X$, es decir, a la obtención de los autovalores y autovectores de la matriz $B^{-1}A$. Los autovalores pueden pues obtenerse como $|B^{-1}A - \lambda I| = 0$, equivalente a

$$|A - \lambda B| = 0$$

y los espacios propios pueden obtenerse como $\text{Nu}(B^{-1}A - \lambda I)$, equivalente a $\text{Nu}(A - \lambda B)$.

Si $A^\dagger = A$ y $B^\dagger = B$, $B^{-1}A$ no es general hermítica ($(B^{-1}A)^\dagger = AB^{-1}$). Si B es definida positiva (es decir, todos sus autovalores λ_i^B son positivos), es posible reescribir el problema generalizado como un problema de autovalores *hermítico y estándar*, mediante el mismo método utilizado en el ej. 5:
Si $AX = \lambda BX \Rightarrow B^{-1/2} A B^{-1/2} B^{1/2} X = \lambda B^{1/2} X$, por lo que el problema se reduce al problema estándar

$$\tilde{A}\tilde{X} = \lambda \tilde{X}$$

con $\tilde{A} = B^{-1/2} A B^{-1/2}$ una matriz *hermítica* ($\tilde{A}^\dagger = \tilde{A}$) y $\tilde{X} = B^{1/2}\tilde{X}$. Aqui, si $B = S_B B' S_B^\dagger$, con B' diagonal ($B'_{ij} = \lambda_i^B \delta_{ij}$, $\lambda_i^B > 0$), $B^{1/2} = S_B B'^{1/2} S_B^\dagger$, con $(B'^{1/2})_{ij} = \sqrt{\lambda_i^B}\delta_{ij}$, siendo $B^{-1/2}$ su inversa.

Por lo tanto, los autovalores λ serán *reales*, y pueden obtenerse de $|\tilde{A}-\lambda I|=0$, equivalente a $|A-\lambda B|=0$, mientras que los autovectores $\tilde{X}$ correspondientes a autovalores distintos serán ortogonales: $\tilde{X}_i^\dagger\tilde{X}_j=0$ si $\lambda_i\neq j\lambda_j$. Esto implica que los autovectores X del problema original serán ortogonales para *un producto escalar modificado dependiente de B:* $\tilde{X}_i^\dagger\tilde{X}_j=X_i^\dagger BX_j=0$ si $\lambda_i\neq\lambda_j$ y $\tilde{X}_i=B^{1/2}X_i$. Veremos en las próximas secciones la definición de producto escalar con más detalle. Al ser $\tilde{A}$ diagonalizable y $B^{1/2}$ no singular, tanto el conjunto $\{\tilde{X}_1,\ldots,\tilde{X}_n\}$ como $\{X_1,\ldots,X_n\}$ formarán una base de $\mathbb{C}^n$.

Más aun, dado que $\tilde{A}$ es hermítica, existirá una matriz no singular de autovectores normalizados y ortogonales $\tilde{S}=(\tilde{X}_1,\ldots,\tilde{X}_n)$ tal que $\tilde{S}^\dagger\tilde{S}=I$ y $\tilde{S}^\dagger\tilde{A}S=\tilde{A}'$ con $\tilde{A}'$ diagonal: $\tilde{A}'_{ij}=\lambda_i\delta_{ij}$. Esto implica que $S=B^{-1/2}\tilde{S}=(X_1,\ldots,X_n)$ satisface simultáneamente

$$S^\dagger AS=A',\quad S^\dagger BS=I$$

con $A'=\tilde{A}'$ diagonal: $A'_{ij}=\lambda_i\delta_{ij}$. Los autovalores generalizados λ_i pueden obtenerse directamente como las raíces de $|A-\lambda B|=0$, mientras que los correspondientes autovectores X_i (columnas de S) de la ecuación $(A-\lambda_iB)X_i=0$. La existencia de tal S queda pues garantizada en el caso $A^\dagger=A$ y $B^\dagger=B$, con B definida positiva ($\lambda_i^B>0\ \forall\ i$).

12. Otras propiedades importantes

12.1) Sean $F:V\to V$, $G:V\to V$ dos operadores sobre el mismo espacio V de dimensión finita. Entonces los autovalores de FG *son los mismos que los de* GF, aún cuando $FG\neq GF$

Demostración: En efecto, si $FG(v)=\lambda v$ con $v\neq 0\Rightarrow GF(G(v))=G(FG(v))=G(\lambda v)=\lambda G(v)$. Si $G(v)\neq 0\Rightarrow\lambda$ es también autovalor de GF con autovector $G(v)$. Si $G(v)=0\Rightarrow$ necesariamente $\lambda=0$ (pues $FG(v)=F(G(v))=F(0)=0$) y por lo tanto $0=\mathrm{Det}[FG]=\mathrm{Det}[GF]$. Esto implica que 0 es también autovalor de GF. Análogamente se muestra que todo autovalor de GF es autovalor de FG.

Esta propiedad es entonces también válida para matrices. Si A,B son dos matrices de $n\times n\Rightarrow$ los autovalores de AB son los mismos que los de BA, aún si $AB\neq BA$.

12.2) Si $[F,G]=FG-GF=0$ y $v\neq 0$ es autovector de F con autovalor $\lambda\Rightarrow G(v)\in V_F(\lambda)$, es decir, $G(v)$ es también autovector de F con el mismo autovalor si $G(v)\neq 0$.

Demostración: $F(G(v))=FG(v)=GF(v)=G(\lambda v)=\lambda G(v)$, por lo que $G(v)\in V_F(\lambda)$.

Además, si V es de dimensión finita n y los autovalores de F *son todos distintos* $\Rightarrow$ todo autovector de F es también autovector de G, pues en tal caso $V_F(\lambda)$ es de dimensión 1 para todo autovalor λ y por lo tanto, necesariamente $G(v)=\alpha v$.

Todos los operadores que conmutan con F quedan en este caso directamente representados por matrices *diagonales* en la base en que $[F]_e$ es diagonal, y son por lo tanto diagonalizables.

En general, si F y G son ambos diagonalizables y $[F,G]=0\Rightarrow$ existe una base común donde F y G son simultáneamente diagonales. Solo hay que diagonalizar F, y luego diagonalizar G dentro de cada espacio propio de F.

13. Subespacios Invariantes

Sea $F : V \to V$ un operador lineal y sea $W \subset V$ un subespacio de V. W es un subespacio invariante bajo la acción de F (se dice también invariante bajo F o por F) si $F(W) \subset W$, es decir, si $\forall\, v \in W$, $F(v) \in W$.

Ejemplos triviales son $W = V$ y $W = \{0\}$, que son subespacios invariantes para todo operador lineal $F : V \to V$ (pues $F(V) \subset V$ y $F(0) = 0$).
También el núcleo $N(F)$ y la imagen $I(F)$ son siempre invariantes por F ($F(N(F)) = \{0\} \subset N(F)$, y si $v \in I(F)$, $F(v) \in I(F)$).
Resulta asimismo trivial reconocer que si $F = \alpha I$, con I el operador identidad $\Rightarrow$ cualquier subespacio $W \subset V$ es invariante por F, ya que si $v \in W$, $F(v) = \alpha v \in W$.
Como otro ejemplo común consideremos el proyector P sobre S_1 en la dirección de S_2, con $S_1 \oplus S_2 = V$. Es obvio que S_1 es invariante por P pues si $v_1 \in S_1 \Rightarrow P(v_1) = v_1 \in S_1$ ($P(S_1) = S_1 = I(P)$). Más aún, cualquier subespacio de S_1 es también invariante por P.

13.1) Si λ es autovalor de $F \Rightarrow$ el espacio propio $V(\lambda)$ es un subespacio invariante por F: Si $v \in V(\lambda)$ $\Rightarrow F(v) = \lambda v \in V(\lambda)$.
13.2) La suma de espacios propios $V(\lambda_1) \oplus V(\lambda_2)$ de un mismo operador F es también invariante por F. Si $v \in V(\lambda_1) \oplus V(\lambda_2)$, $v = v_1 + v_2$, con $F(v_i) = \lambda_i v_i$, $i = 1, 2$ y por lo tanto $F(v) = \lambda_1 v_1 + \lambda_2 v_2 \in V(\lambda_1) \oplus V(\lambda_2)$. Análogamente, la suma de espacios propios $V(\lambda_1) \oplus \ldots \oplus V(\lambda_k)$ es invariante por F.
13.3) Si $[F, G] = 0$ y W es un subespacio invariante por F, $G(W)$ es también invariante por F.
En efecto, si $v \in W$, $w = F(v) \in W$ y por lo tanto, $F(G(v)) = FG(v) = GF(v) = G(F(v)) = G(w) \in G(W)$.
13.4) Si se escribe a V como suma directa de subespacios invariantes,
$V = S_1 \oplus S_2 \oplus \ldots \oplus S_k$, existe una base (aquella formada por la unión de las bases de cada subespacio invariante) en la que la matriz que representa a $[F]_e$ estará bloqueada en la forma

$$[F]_e = \begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & A_k \end{pmatrix}$$

donde A_i es una matriz de dimensión $d_i \times d_i$, con $d_i = \dim S_i$, $i = 1, \ldots, k$ y $\sum_{i=1}^{k} d_i = n$.
En efecto, si $(e_{i1}, \ldots, e_{id_i})$ es una base de S_i, $F(e_{ij}) \in S_i$, por lo que $F(e_{ij}) = \sum_{l=1}^{d_i} (A_i)_{lj} e_{il}$ para $j = 1, \ldots, d_i$. Análogamente, si existe una base en la que $[F]_e$ tiene la forma de bloques anterior $\Rightarrow V = S_1 \oplus S_2 \oplus \ldots \oplus S_k$, con S_i invariante por F para $i = 1, \ldots, k$. Basta con considerar S_i como el subespacio generado por los vectores de la base correspondientes a cada bloque.

Los autovalores de F pueden pues obtenerse directamente diagonalizando cada bloque A_i: Como $\det F = \prod_i \det A_i$ (demostración dada en clase), el polinomio característico resulta $|F - \lambda I| = \prod_{i=1}^{k} |A_i - \lambda I_{d_i}|$, por lo que sus racies serán las raíces de cada término, es decir, los autovalores de cada bloque A_i. Y los autovectores correspondientes pertenecerán al subespacio invariante asociado a A_i (detalles dados en clase). El conocimiento de subespacios invariantes posibilita pues grandes simplificaciones cuando se tiene que diagonalizar matrices de grandes dimensiones.

14. Forma Canónica de Jordan

Surge ahora la pregunta sobre cuál es la forma más simple en que pueden escribirse los operadores (o matrices) no diagonalizables. El siguiente teorema nos da la respuesta:
Teorema: Sea $F : V \to V$ un operador lineal en un espacio vectorial V de dimensión finita n sobre $\mathbb{C}$. Entonces existe una base $e = (e_{11}, e_{12}, \ldots, e_{1d_1}, \ldots, e_{k1}, e_{k2}, \ldots, e_{kd_k})$, con $\sum_{i=1}^{k} d_i = n$, en la que

$$\begin{array}{rcl} F(e_{i1}) & = & \lambda_i e_{i1} \\ F(e_{ij}) & = & \lambda_i e_{ij} + e_{i,j-1}, \;\; j = 2, \ldots, d_i \end{array} \;, \quad i = 1, \ldots, k$$

o sea, $F(e_{11}) = \lambda_1 e_{11}$, $F(e_{12}) = \lambda_1 e_{12} + e_{11}$, $\ldots$, $F(e_{1d_1}) = \lambda_1 e_{1d_1} + e_{1,d_1-1}$ y similar para $i \geq 1$. El caso diagonalizable corresponde a $d_i = 1 \; \forall \; k$, en cuyo caso $k = n$. Los parámetros λ_i no son necesariamente distintos y son los autovalores de F, como demostraremos a continuación.
La matriz $[F]_e \equiv [F]_e^e$ en esta base toma entonces la forma de bloques

$$[F]_e = A = \begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & A_k \end{pmatrix}$$

donde A_i son matrices de $d_i \times d_i$ de la forma

$$A_i = \begin{pmatrix} \lambda_i & 1 & 0 & \ldots & 0 \\ 0 & \lambda_i & 1 & \ldots & 0 \\ & & & \ldots & \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & \lambda_i \end{pmatrix} = \lambda_i I_{d_i} + J_{d_i} \,, \quad J_{d_i} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ & & & \ldots & \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

con I_{d_i} la matriz identidad de $d_i \times d_i$.
Cada subespacio $S_i = \overline{(e_{i1}, e_{i2}, \ldots, e_{id_i})}$ es claramente invariante por F, ya que $F(e_{ij}) \in S(\lambda_i)$.
Es claro también que los escalares λ_i, $i = 1, \ldots, k$, son los autovalores de F, pues

$$P(\lambda) = \mathrm{Det}[F - \lambda I] = |A_1 - \lambda I_{d_1}| \ldots |A_k - \lambda I_{d_k}| = (\lambda_1 - \lambda)^{d_1} \ldots (\lambda_k - \lambda)^{d_k}$$

posee como únicas raíces a $\lambda_1, \ldots, \lambda_k$.

Cada submatriz A_i posee un único autovalor λ_i de multiplicidad d_i ($|A_i - \lambda I_{d_i}| = (\lambda_i - \lambda)^{d_i}$), pero el espacio propio correspondiente es *de dimensión 1*: $\dim N[A_i - \lambda_i I_{d_i}] = \dim N[J_{d_i}] = 1$ pues $\mathrm{Rango}(J_{d_i}) = d_i - 1$. Por lo tanto, la submatriz A_i no es diagonalizable si $d_i > 1$. Cada subespacio S_i contiene entonces *un único subespacio propio* de dimensión 1, que es el generado por e_{i1}, y el número total de autovectores LI de F es $k \leq n$ (uno por cada S_i).

Notemos que $(F - \lambda_i I)e_{ij} = e_{i,j-1}$ para $j > 1$, con $(F - \lambda_i I)e_{i1} = 0$, por lo que aplicando m veces el op. $(F - \lambda_i I)$ sobre e_{ij} resulta

$$(F - \lambda_i I)^m e_{ij} = \left\{ \begin{array}{cc} e_{i,j-m} & m < j \\ 0 & m \geq j \end{array} \right.$$

Los operadores no diagonalizables en espacios finitos se caracterizan pues por la existencia de vectores e_{ij} no nulos tales que $(F - \lambda_i I)^j e_{ij} = 0$ pero $(F - \lambda_i I) e_{ij} \neq 0$ si $j > 1$. Si F es diagonalizable tales vectores no existen. Notemos que conociendo e_{id_i}, los restantes vectores e_{ij} pueden obtenerse como

$$e_{ij} = (F - \lambda_i I)^{d_i - j} e_{id_i} = (F - \lambda_i I) e_{i,j+1}, \;\; j = 1, \ldots, d_i - 1$$

La ecuación previa implica también que $(F - \lambda_i I)^{d_i}(e_{ij}) = 0, \;\; j = 1, \ldots, d_i$. Por lo tanto, la matriz $J_{d_i} = A_i - \lambda_i I_{d_i}$ es *nilpotente*:

$$(J_{d_i})^{d_i} = 0$$

donde 0 denota la mariz nula de $d_i \times d_i$. Por ejemplo,

$$J_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \;\; J_2^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \;\; J_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \;\; J_3^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \;\; J_3^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

La evaluación de potencias del operador puede entonces realizarse sin mayor dificultad, ya que

$$A^m = \begin{pmatrix} A_1^m & 0 & \dots & 0 \\ 0 & A_2^m & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & A_k^m \end{pmatrix}$$

donde, teniendo en cuenta que I_{d_i} conmuta con J_{d_i},

$$A_i^m = (\lambda_i I_{d_i} + J_{d_i})^m = \lambda_i^m I_{d_i} + m\lambda_i^{m-1} J_{d_i} + \frac{m(m-1)}{2!}\lambda_i^{m-2} J_{d_i}^2 + \ldots + J_{d_i}^m$$

para $i = 1, \ldots, k$. Esta expansión contiene a lo sumo d_i términos ya que $(J_{d_i})^r = 0$ si $r \geq d_i$.
En general, si $p(t)$ es un polinomio de grado m,

$$p(t) = p(\lambda_i)1 + p'(\lambda_i)(t - \lambda_i) + \ldots + \frac{p^{(m)}(\lambda_i)}{m!}(t - \lambda_i)^m$$

se obtiene

$$p(A_i) = p(\lambda_i) I_{d_i} + p'(\lambda_i) J_{d_i} + \ldots + \frac{p^{(d_i-1)}(\lambda_i)}{(d_i-1)!} J_{d_i}^{d_i-1}$$

Además, la forma de Jordan es muy conveniente para la evaluación de exponenciales:

$$\exp[A_i t] = \exp[\lambda_i I_{d_i} t + J_{d_i} t] = \exp[\lambda_i I_{d_i} t]\exp[J_{d_i} t] = e^{\lambda_i t}\left(I_{d_i} + J_{d_i} t + \ldots + J_{d_i}^{d_i-1}\frac{t^{d_i-1}}{(d_i-1)!}\right)$$

Por lo tanto, $B(t) = \exp[A_i t]$ será una matriz triangular con elementos $B_{kj} = e^{\lambda_i t} t^{j-k}/(j-k)!$ si $k \leq j$ y $B_{kj} = 0$ si $k > j$.

Para obtener la representación de Jordan se puede, una vez obtenidos los k autovalores λ_i y autovectores e_{i1}, $i = 1, \ldots, k$, resolver las ecuaciones inhomogéneas $F(e_{ij}) = \lambda_i e_{ij} + e_{i,j-1} \quad j = 2, \ldots, d_i$, es decir, trabajando en forma matricial en la base e,

$$AX_{i1} = \lambda_i X_{i1}, \;\; AX_{ij} = \lambda_i X_{ij} + X_{i,j-1}, \quad j = 2, \ldots, d_i, \;\; i = 1, \ldots, k$$

que no poseen solución única. Otra forma más eficiente es partir de e_{id_i}, es decir, encontrar un vector X_{id_i} que satisfaga

$$(A - \lambda_i I)^{d_i} X_{id_i} = 0, \;\; (A - \lambda_i I)^{d_i-1} X_{id_i} \neq 0$$

Los vectores restantes del bloque pueden obtenerse como

$$X_{ij} = (A - \lambda_i I)^{d_i-j} X_{id_i}, \;\; j = 1, \ldots, d_i - 1$$

Ejemplo:

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

Tenemos $|A - \lambda I_3| = (1-\lambda)(2-\lambda)^2$, por lo que las raíces son $\lambda_1 = 1$, con multiplicidad 1, y $\lambda_2 = 2$, con multiplicidad 2. Como

$$A - 2I_3 = \begin{pmatrix} -1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

posee rango 2, $N[A - 2I_3]$ posee dimensión 1, por lo que A no es diagonalizable.
Para $\lambda_2 = 2$ el sistema homogéneo $(A - \lambda I_3)X = 0$ posee la solución general $x = y$, $z = 0$, de modo que el autovector es de la forma $x(1,1,0)^t$. Eligiendo $X_{11} = (1,1,0)^t$, el vector X_{12} puede obtenerse resolviendo

$$\begin{pmatrix} -1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

que da como resultado $z=1$, $x=y$. Podemos elegir entonces $X_{12}=(0,0,1)^t$. Finalmente, $A-1I_3=\begin{pmatrix}0&1&1\\0&1&1\\0&0&1\end{pmatrix}$, por lo que $(A-I_3)X_{31}=0$ conduce a $X_{31}=x(1,0,0)^t$. Obtenemos entonces

$$S=\begin{pmatrix}1&0&1\\1&0&0\\0&1&0\end{pmatrix},\quad \text{con } S^{-1}=\begin{pmatrix}0&1&0\\0&0&1\\1&-1&0\end{pmatrix}$$

y finalmente la forma de Jordan

$$A'=S^{-1}AS=\begin{pmatrix}2&1&0\\0&2&0\\0&0&1\end{pmatrix}$$

Puede comenzarse también determinando X_{12} a partir de las condiciones $(A-2I_3)X_{12}\neq 0$, $(A-2I_3)^2X_{12}=0$, y obtener luego X_{11} como $(A-2I_3)X_{12}$ (hecho así en clase).

Se obtiene también

$$\exp[A't]=\begin{pmatrix}\exp[t\begin{pmatrix}2&1\\0&2\end{pmatrix}]&0\\0&\exp(t)\end{pmatrix}=\begin{pmatrix}e^{2t}&te^{2t}&0\\0&e^{2t}&0\\0&0&e^t\end{pmatrix}$$

ya que $\exp[t\begin{pmatrix}2&1\\0&2\end{pmatrix}]=\exp[t(2I_2+J_2)]=\exp[2tI_2]\exp[tJ_2]=e^{2t}(I_2+tJ_2)=e^{2t}\begin{pmatrix}1&t\\0&1\end{pmatrix}$, pues $J_2^2=0$.

Resolución general de sistemas de ecuaciones lineales de primer orden. La solución del sistema

$$\frac{dX}{dt}=AX$$

con A constante pero no diagonalizable, puede obtenerse a partir de la forma canónica de Jordan. Tenemos, para $X(0)=X_0$ y $A=SA'S^{-1}$,

$$X=\exp[At]X_0=S\exp[A't]C=\sum_{i=1}^{k}e^{\lambda_i t}\sum_{m=1}^{d_i}c_{im}\sum_{j=0}^{m-1}V_{i,m-j}\frac{t^j}{j!}$$

donde $V_{i,m-j}$ denota las columnas de S [$S=(V_{11},V_{12},\ldots,V_{1,d_1},\ldots,V_{k,1},\ldots,V_{k,d_k})$, de forma que $S\exp[A't]=(e^{\lambda_1 t}V_{11},e^{\lambda_1 t}(V_{12}+tV_{11}),e^{\lambda_1 t}(V_{13}+tV_{12}+\frac{t^2}{2!}V_{11}),\ldots)$] y $C=S^{-1}X_0=(c_{11},c_{12},\ldots,c_{1d_1},\ldots)^t$ un vector de constantes determinadas por el vector inicial $X_0=X(0)$. Por ejemplo,

$$X=e^{\lambda t}[c_1V_1+c_2(V_2+V_1t)+c_3(V_3+V_2t+V_1t^2/2!)+\ldots]$$

en el caso de un solo bloque, con $V_j=(A-\lambda I)^{d-j}V_d$ y d la dimensión del bloque.

15. Polinomios Anuladores y Teorema de Cayley-Hamilton

Sea $F:V\to V$ un operador lineal en un espacio vectorial V de dimensión finita n.
Si $p(\lambda)=a_0+a_1\lambda+\ldots+a_m\lambda^m$ es un polinomio de grado m, podemos asociar a $p(\lambda)$ el operador lineal

$$p(F)=a_0I+a_1F+\ldots+a_mF^m$$

La matriz que representa al operador $p(F)$ en una base e es

$$\begin{aligned}[p(F)]_e&=a_0[I]_e+a_1[F]_e+\ldots+a_m[F^m]_e=a_0I_n+a_1[F]_e+\ldots+a_m[F]_e^m\\&=p([F]_e)\end{aligned}$$

donde hemos asociado $F^0=I$ y $[F]_e^0=I_n$. Además, si escribimos $p(\lambda)$ en términos de sus m raíces λ_i

$$p(\lambda)=a_m(\lambda-\lambda_1)(\lambda-\lambda_2)\ldots(\lambda-\lambda_m)$$

entonces

$$p(F) = a_m(F - \lambda_1 I)(F - \lambda_2 I)\ldots(F - \lambda_m I)$$

ya que las potencias F^k de F conmutan todas entre si $\forall\ j \geq 0$.

Un polinomio p se dice que es *anulador* de F si $p(F) = 0$ (o sea, si $p(F)$ es el operador nulo). Dado que la dimension del espacio de operadores lineales $H = \{F : V \to V,\ F \text{ lineal}\}$ en un espacio vectorial V de dimensión finita n es n^2, es claro que el conjunto $(I = F^0, F, F^2, \ldots, F^{n^2})$ es LD (pues son $n+1$). y que por lo tanto, existen siempre n^2+1 constantes $c_0, \ldots, c_n$ no todas nulas tales que $c_0 I + c_1 F + \ldots c_n F^{n^2} = 0$. Esto muestra en forma básica que siempre existe un polinomio anulador de F.

No obstante, el siguiente teorema (denominado teorema de Cayley-Hamilton) muestra que el mismo *polinomio característico asociado a F, que es de grado n, es siempre un polinomio anulador de F.*

Teorema: Si $p(\lambda) = \text{Det}[F - \lambda I]$ es el polinomio característico de $F \Rightarrow p(F) = 0$.
Como el polinomio característico es de grado n ($p(\lambda) = a_0 + a_1\lambda + \ldots + a_n\lambda^n$ con $a_n = (-1)^n \neq 0$), el teorema implica que es siempre posible expresar F^n en términos de potencias de grado $\leq n-1$: Como $p(F) = 0 \Rightarrow$

$$F^n = -(a_0 + a_1 F + \ldots a_{n-1}F^{n-1})/a_n$$

Por lo tanto, *cualquier* potencia F^k con $k \geq n$ puede expresarse en términos de potencias de grado $\leq n-1$.

Demostraremos el teorema a partir de la forma canónica de Jordan. No obstante, en el caso en que F es diagonalizable, el teorema es obvio, ya que en tal caso existe una base e en la que $[F]_e$ es diagonal,

$$[F]_e = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

donde λ_i, $i = 1, \ldots, n$ son los autovalores de F, y

$$[p(F)]_e = p([F])_e = \begin{pmatrix} p(\lambda_1) & 0 & \ldots & 0 \\ 0 & p(\lambda_2) & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & p(\lambda_n) \end{pmatrix} \qquad (15.1)$$

para cualquier polinomio p. Pero si p es el polinomio característico, $p(\lambda_i) = 0$ para $i = 1, \ldots, n$ y por lo tanto $[p(F)]_e = p([F]_e) = 0$ (matriz nula). Esto implica a su vez $p(F) = 0$ (operador nulo).

En el caso general, utilizando la base en la que $[F]_e$ tiene la forma canónica de Jordan, tenemos, para cualquier polinomio p,

$$[p(F)]_e = p([F]_e) = \begin{pmatrix} p(A_1) & 0 & \ldots & 0 \\ 0 & p(A_2) & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & p(A_k) \end{pmatrix}$$

donde A_i, $i = 1, \ldots, k$ son matrices de $d_i \times d_i$ que satisfacen $(A_i - \lambda_i I_{d_i})^{d_i} = 0$ (matriz nula).
Recordemos ahora que el polinomio característico de F está dado por

$$p(\lambda) = |[F]_e - \lambda I_n| = |A_1 - \lambda I_{d_1}| \ldots |A_k - \lambda I_{d_k}| = (\lambda_1 - \lambda)^{d_1} \ldots (\lambda_k - \lambda)^{d_k}$$

Por lo tanto,

$$p(A_i) = (\lambda_1 I_{d_i} - A_i)^{d_1} \ldots (\lambda_i I_{d_i} - A_i)^{d_i} \ldots (\lambda_k I_{d_i} - A_i)^{d_k} = 0$$

pues $(\lambda_i I_{d_i} - A_i)^{d_i} = 0$. Esto implica $[p(F)]_e = 0$ y entonces $p(F) = 0$. Se cumple pues, para cualquier base e', $p([F]_{e'}) = [p(F)]_{e'} = [0]_{e'} = 0$.

El teorema vale por consiguiente para matrices A generales de $n \times n$. Si $p(\lambda) = |A - \lambda I_n|$ es el polinomio característico asociado a A (de grado n) $\Rightarrow p(A) = 0$ (la matriz nula de $n \times n$).
Para matrices A diagonalizables el resultado es evidente, ya que en tal caso $A = SA'S^{-1}$, con A' diagonal, y por lo tanto $p(A) = p(SA'S^{-1}) = Sp(A')S^{-1}$, pero $p(A')$ tiene la forma 15.1 y es por lo tanto la matriz nula.

Escribiendo

$$p(F) = c_n F^n + c_{n-1} F^{n-1} + \ldots + c_1 F + c_0 I$$

en el caso del polinomio característico tenemos $c_n = (-1)^n \neq 0$ y $c_0 = \mathrm{Det}[F]$. Por lo tanto, como $p(F) = 0$, podemos escribir

$$F^n = -(c_{n-1} F^{n-1} + \ldots c_1 F + c_0 I)/c_n$$

de modo que F^n (y por lo tanto cualquier potencia F^k, con $k \geq n$ natural) puede escribirse en términos de los operadores $F^{n-1}, \ldots, F, I$. Más aún, si F es invertible, $c_0 \neq 0$ y multiplicando la expresión anterior por F^{-1} se obtiene

$$F^{-1} = -(c_n F^{n-1} + c_{n-1} F^{n-2} + \ldots c_1 I)/c_0$$

de modo que también F^{-1} (y por tanto cualquier potencia F^{-k}, $k > 0$ natural) puede escribirse en términos de $F^{n-1}, \ldots, F, I$.

Cabe destacar que el polinomio característico no es necesariamente el polinomio anulador de grado mínimo. Sí lo es en el caso de autovalores todos distintos o, en general, en el caso de bloques de Jordan con autovalores todos distintos.
Si F es diagonalizable, el polinomio anulador de grado mínimo es simplemente $P_m(\lambda) = \prod_i (\lambda - \lambda_i)$, donde la productoria es sobre *autovalores distintos*.
En el caso general, el polinomio anulador de grado mínimo será $P_m(\lambda) = \prod_i (\lambda - \lambda_i)^{d_i}$, donde la productoria es nuevamente sobre autovalores distintos y d_i es la dimensión del mayor bloque de Jordan asociado a λ_i. $P_m(\lambda)$ es pues de grado $\leq n$.

Ejemplo 1: Sea

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

El polinomio característico es

$$p(\lambda) = |A - \lambda I_2| = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1$$

Se cumple entonces

$$p(A) = A^2 - I_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Esto muestra simplemente que $A^2 = I_2$ y que por lo tanto, $A^k = I_2$ si k es par y $A^k = A$ si k impar.

Ejemplo 2:

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

El polinomio característico es

$$p(\lambda) = |A - \lambda I_3| = (1-\lambda)(2-\lambda)^2 = -\lambda^3 + 5\lambda^2 - 8\lambda + 4$$

El teorema implica entonces que

$$-A^3 + 5A^2 - 8A + 4I_3 = 0$$

donde $A^2 = A.A$, $A^3 = A.A.A$ (producto matricial), como es fácil verificar. Por lo tanto, $A^3 = 5A^2 - 8A + 4I_3$ y $A^{-1} = (A^2 - 5A + 8I)/4$. Cualquier potencia A^k con k entero puede expresarse en términos de I_3, A y A^2.

Ej. 3: Matriz A de $n \times n$ de rango 1, con $n \geq 2$. Dado que A y $A' = S^{-1}AS$ poseen el mismo rango y traza, tenemos, si A' es la forma canónica de Jordan de A, $r(A') = 1$ y $\mathrm{Tr}\, A' = \mathrm{Tr}\, A$, por lo que A' tiene sólo una fila no nula. Si $\mathrm{Tr}\, A \neq 0$, la única posibilidad es que A sea diagonalizable y tenga un único autovalor no nulo igual a $\mathrm{Tr}\, A$, siendo los restantes nulos, mientras que si $\mathrm{Tr}\, A = 0$, la única posibilidad para A' es un bloque de Jordan de dimensión 2 de la forma $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, siendo todos los autovalores nulos y A no diagonalizable. Toda matriz de rango 1 es necesariamente de la forma $A = bc^t$, con b y c vectores columna no nulos (de $n \times 1$), como el lector puede fácilmente demostrar, con $\mathrm{Tr}\, A = c^t b = \mathbf{c} \cdot \mathbf{b}$. Se deja como problema hallar los autovectores asociados y el polinomio minimal en ambos casos.

16. Demostración

Daremos aquí un resumen de la demostración de la forma canónica de Jordan. En primer lugar, sabemos que todo operador lineal $F : V \to V$, con V de dimensión finita n, posee un polinomio anulador $P(x)$, tal que $P(F) = 0$ (o sea, $P(F)(v) = 0 \ \forall \ v \in V$). Existirá entonces un polinomio anulador de grado mínimo $P_m(x) = a_0 x + a_1 x + \ldots + a_m x^m$ (polinomio minimal), tal que $P_m(F) = a_0 F + a_1 F + \ldots + a_m F^m = 0$.
1) λ es raíz de $P_m(F)$ si y sólo si λ es autovalor de F. Esto indica que las raíces del polinomio minimal y el polinomio característico son las mismas. Sólo la multiplicidad puede ser diferente.
Dem.: Si λ es autovalor de $F \Rightarrow \exists \ v \neq 0$ tal que $F(v) = \lambda v$, y en tal caso $P_m(F)(v) = P_m(\lambda)v = 0$, por lo que $P_m(\lambda) = 0$, es decir, λ es raíz de $P_m(x)$.
Si λ es raíz de $P_m(x) \Rightarrow P_m(x) = Q_{m-1}(x)(x - \lambda)$. En tal caso $P_m(F)(v) = Q_{m-1}(F)(F - \lambda I)(v) = 0$ $\forall \ v \in V$, por lo que necesariamente $\exists \ v \neq 0$ tal que $(F - \lambda I)(v) = 0$, es decir, λ es autovalor de F y v autovector asociado (si tal vector no existiese tendríamos $Q_{m-1}(F)(v) = 0 \ \forall \ v \in V$ y el polinomio minimal sería $Q_{m-1}(F)$, de grado $m - 1 < m$, en contradicción con la hipótesis).
2) Si $P_m(x) = Q_1(x)Q_2(x)$, con $Q_1(x)$ y $Q_2(x)$ polinomios sin raíces comunes y $P_m(F) = Q_1(F)Q_2(F) = 0$ $\Rightarrow V = N(Q_1(F)) \oplus N(Q_2(F))$, donde $N(Q_i(F))$ $(i = 1, 2)$ denota el núcleo de $Q_i(F)$. Los subespacios $N(Q_i(F))$ son además invariantes por F.
Dem.: Al no tener raíces comunes, existen polinomios $A_1(x)$, $A_2(x)$ t.q. $1 = A_1(x)Q_1(x) + A_2(x)Q_2(x)$, o sea,

$$I = A_1(F)Q_1(F) + A_2(F)Q_2(F)$$

Por lo tanto, $\forall \ v \in V$, $v = A_1(F)Q_1(F)(v) + A_2(F)Q_2(F)(v) = v_1 + v_2$, con $v_i = A_i(F)Q_i(F)$. Pero $v_1 \in N(Q_2(F))$ pues $Q_2(F)A_1(F)Q_1(F)(v) = A_1(F)Q_1(F)Q_2(F)(v) = 0$, y análogamente, $v_2 \in N(Q_1(F))$. Esto muestra que $V = N(Q_2(F)) + N(Q_1(F))$.
Además, si $v \in N(Q_1(F))$ y $v \in N(Q_2(F)) \Rightarrow v = A_1(F)Q_1(F)(v) + Q_2(F)Q_2(F)(v) = 0$. Esto muestra que $V = N(Q_1(F)) \oplus N(Q_2(F))$.
Finalmente, si $v \in N(Q_1(F)) \Rightarrow v = A_2(F)Q_2(F)(v)$ y en tal caso $Q_1(F)F(v) = A_2(F)Q_1(F)Q_2(F)F(v) = 0$, por lo que $F(v) \in N(Q_1(F))$. Análogamente, si $v \in N(Q_2(F)) \Rightarrow F(v) \in N(Q_2(F))$, por lo que ambos núcleos son invariantes por F.
3) Generalizando, si

$$P_m(x) = (x - \lambda_1)^{d_1} \ldots (x - \lambda_k)^{d_k}$$

con $\lambda_i \neq \lambda_j$ si $i \neq j$ (las raíces distintas de $P_m(x)$) y $P_m(F) = 0 \Rightarrow V = V_1 \oplus \ldots \oplus V_k$, con $V_i = N(F - \lambda_i I)^{d_i}$. El espacio completo V puede pues subdividirse en k subespacios invariantes, núcleos de $\tilde{F}_i^{d_i}$, donde $\tilde{F}_i = (F - \lambda_i I)$. Podemos pues construir una base de V formada por las bases de V_i.
4) Para construir una base de V_i, notemos que debe existir un vector $v \neq 0$ tal que $\tilde{F}_i^{d_i}(v) = 0$ pero $\tilde{F}_i^{d_i - 1}(v) \neq 0$ (de lo contrario $P_m(F)$ no sería el polinomio minimal). En tal caso, los d_i vectores no nulos

$$e_{ij} = \tilde{F}_i^{d_i - j}(v), \quad j = 1, \ldots, d_i$$

(o sea $e_{id_i} = v$, $e_{ij} = \tilde{F}_i(e_{i,j+1})$, $j = 1, \ldots, d_i - 1$) son LI. Dem.: Si

$$c_1 \tilde{F}_i^{d_i - 1}(v) + c_2 \tilde{F}_i^{d_i - 2}(v) + \ldots + c_{d_i - 1}\tilde{F}_i(v) + c_{d_i} v = 0$$

aplicando $\tilde{F}_i^{d_i - 1}$ al segundo miembro obtenemos $c_{d_i}\tilde{F}_i^{d_i - 1}(v) = 0$, por lo que $c_{d_i} = 0$. Luego, aplicando sucesivamente $\tilde{F}_i^j$, con $j = d_i - 1, \ldots, 0$, vemos que $c_j = 0$ para $j = 1, \ldots, d_i$. Notemos además que $\tilde{F}_i(e_{i1}) = \tilde{F}_i^{d_i}(v) = 0$, o sea, $(F - \lambda_i I)e_{i1} = 0$, por lo que e_{i1} es autovector de F con autovalor λ_i. Tenemos pues, en el subespacio S_i generado por los d_i vectores $B_i = (e_{i1}, \ldots, e_{id_i})$,

$$[\tilde{F}_i]_{B_i} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ & & & \ldots & \\ 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & \ldots & 0 & 0 \end{pmatrix} = J_{d_i}, \quad \text{es decir,} \quad [F_i]_{B_i} = [\tilde{F}_i]_{B_i} + \lambda_I I_{d_i} = \begin{pmatrix} \lambda_i & 1 & 0 & \ldots & 0 \\ 0 & \lambda_i & 1 & \ldots & 0 \\ & & & \ldots & \\ 0 & 0 & \ldots & \lambda_i & 1 \\ 0 & 0 & \ldots & 0 & \lambda_i \end{pmatrix}$$

Se obtiene así un bloque de Jordan de dimensión d_i (grado del término correspondiente $(x - \lambda_i)^{d_i}$ del polinomio minimal).

Puede existir otro vector $v \in N(\tilde{F}_i^{d_i})$ tal que $\tilde{F}^{d_i-1}(v) \neq 0$ pero $\tilde{F}_i^{d_i}(v) = 0$ *y que no pertenezca al espacio generado por los vectores de B_i*. Este vector generaría otro bloque de Jordan de la misma dimensión con el mismo autovalor λ_i. En general, pueden surgir vectores $v \in N(\tilde{F}_i^{d_i})$ que no pertenezcan a los subespacios generados por el conjunto de vectores anteriores y que satisfagan $\tilde{F}_i^{r-1}(v) = 0$ pero $\tilde{F}_i^{r}(v) = 0$, con $r \leq d_i$, que generarán otros bloques de Jordan de dimensión $r \leq d_i$ con el mismo autovalor. La dimension total de $N(F - \lambda_i I)^{d_i}$ será así la multiplicidad $m_i \geq d_i$ de λ_i en el polinomio característico.
Si $d_i = 1$ los bloques son de dimensión 1 y los vectores correspondientes *autovectores* de F con autovalor λ_i. Este es el caso donde F es *diagonalizable* en el subespacio asociado a λ_i, es decir, donde la dimensión de $N(F - \lambda_i I)$ coincide con la multiplicidad de λ_i como raíz del polinomio característico.
Por lo tanto, si F es diagonalizable, el polinomio minimal es $P_m(x) = (x - \lambda_1) \ldots (x - \lambda_k)$.