

Teoría espectral

Ingeniería Biomédica Curso: Matemáticas I

Fernando Chamizo <https://matematicas.uam.es/~fernando.chamizo/>

1. Autovalores y autovectores

En un espacio vectorial V de dimensión finita n , una vez fijada una base, un endomorfismo $f : V \rightarrow V$ viene determinado por una matriz $A \in \mathcal{M}_n(\mathbb{R})$, esto es, por n^2 números. En algunas aplicaciones se produce una simplificación considerable cuando A es diagonal y, por tanto, está determinada por solo n números. En cierto modo, un endomorfismo con matriz diagonal en n dimensiones se desacopla en n endomorfismos en dimensión 1 tan tontos como multiplicar por un número. La *teoría espectral* estudia cuándo tal desacople es posible, incluso en el contexto de dimensión infinita, mucho más difícil, que excede a este curso y es crucial en física cuántica.

Comenzamos con una definición que recoge todo lo que vamos a hacer, pero que te resultará, seguramente, demasiado matemática hasta que veamos su realización con matrices. Dado un endomorfismo $f : V \rightarrow V$ se dice que $\vec{v} \in V - \{\vec{0}\}$ es un *vector propio* o *autovector* de f si $\vec{v} \in \text{Ker}(f - \lambda \text{Id})$ para cierto $\lambda \in K$. A este λ se le llama *valor propio* o *autovalor*.

Recordando la definición de núcleo, los vectores propios no son otra cosa que los vectores no nulos tales que $f(\vec{v}) = \lambda\vec{v}$. En este curso nos vamos a fijar solo en el caso $V = K^n$ y, por tanto, $f(\vec{x}) = A\vec{x}$ con $A \in \mathcal{M}_n(K)$. Aunque un matemático muy rigorista hablaría siempre de los autovectores y autovalores del endomorfismo f , con un ligero abuso de notación, casi todo el mundo (matemáticos incluidos) habla de los autovectores y autovalores de la matriz cuadrada A . Seguiremos esta política y en lo sucesivo se aplicarán estos conceptos solo a matrices. Así pues, un vector propio o autovector de una matriz $A \in \mathcal{M}_n(K)$ es un vector no nulo $\vec{v} \in K^n$ que cumple $A\vec{v} = \lambda\vec{v}$ y, en ese caso, se dice que λ es el valor propio o autovalor correspondiente a \vec{v} de la matriz A .

Los autovectores son soluciones no triviales del sistema homogéneo $(A - \lambda I)\vec{x} = \vec{0}$ que tiene el mismo número de ecuaciones y de incógnitas, por tanto, podemos hallar los autovalores imponiendo $\det(A - \lambda I) = 0$ que equivale a que el sistema sea compatible indeterminado y podemos hallar los autovectores resolviendo los sistemas correspondientes. La ecuación de los autovalores $\det(A - \lambda I) = 0$ se denomina *ecuación característica* y $\det(A - \lambda I)$ es el *polinomio*

característico. Es fácil convencerse de que realmente es un polinomio y de que su grado es n cuando $A \in \mathcal{M}_n(K)$.

Como primer ejemplo, hallemos los autovalores y autovectores de

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

Primero calculamos el polinomio característico:

$$\det(A - \lambda I) = \begin{vmatrix} 3 - \lambda & 2 \\ 1 & 4 - \lambda \end{vmatrix} = (3 - \lambda)(4 - \lambda) - 2 = \lambda^2 - 7\lambda + 10.$$

Al resolver $\lambda^2 - 7\lambda + 10 = 0$ obtenemos los autovalores $\lambda_1 = 2$ y $\lambda_2 = 5$ (la ordenación es indiferente). Para $\lambda_1 = 2$ los vectores propios son los $\vec{v} \in \mathbb{R}^2 - \{\vec{0}\}$ que satisfacen

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \vec{v} = \vec{0}, \quad \text{lo que implica} \quad \vec{v} = \mu \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad \mu \neq 0.$$

De la misma manera, para $\lambda_2 = 5$ se tiene

$$\begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix} \vec{v} = \vec{0} \quad \Longrightarrow \quad \vec{v} = \mu \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mu \neq 0.$$

Con ello hemos hallado todos los autovectores.

Para cada autovalor λ el conjunto de soluciones de $(A - \lambda I)\vec{x} = \vec{0}$, llamado *autoespacio* de λ , es necesariamente infinito, todos sus elementos excepto $\vec{0}$ son autovectores. Un abuso de notación muy habitual es restringir el nombre *autovectores* (o *vectores propios*) a los elementos de bases fijadas de los autoespacios. Así en el ejemplo anterior muchos dirían que $(2, -1)^t$ es “el” autovector correspondiente a $\lambda_1 = 2$, sobreentendiendo que el resto son múltiplos suyos, o dirían que la matriz A tiene dos autovectores que son $(2, -1)^t$ y $(1, 1)^t$. Este abuso de notación es difícil de aceptar por los matemáticos ya que hay infinitos autovectores y no hay una manera canónica de especificarlos. Sin embargo está tan extendido que es sensato hacer alguna concesión.

Una ecuación polinómica sobre \mathbb{C} siempre tiene raíces en \mathbb{C} , pero en \mathbb{R} esto no es cierto en general y una matriz real puede tener autovalores complejos. Consideremos, por ejemplo,

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

No tiene autovalores (ni, por tanto, autovectores) en \mathbb{R} porque la ecuación característica es $\lambda^2 + 1 = 0$. Sin embargo, en \mathbb{C} , $\lambda_1 = i$ y $\lambda_2 = -i$ son valores propios válidos y para λ_1

$$\begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \vec{v} = \vec{0} \quad \Longrightarrow \quad \vec{v} = \mu \begin{pmatrix} i \\ 1 \end{pmatrix}, \quad \mu \neq 0.$$

Mientras que para λ_2

$$\begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \vec{v} = \vec{0} \quad \Longrightarrow \quad \vec{v} = \mu \begin{pmatrix} -i \\ 1 \end{pmatrix}, \quad \mu \neq 0.$$

Por cierto, los cálculos para λ_2 son innecesarios si uno se fija en que basta conjugar los de λ_1 .

Por tanto, $\vec{v}_1 = (i, 1)^t$ y $\vec{v}_2 = (-i, 1)^t$ son vectores propios válidos de A y la gente de a pie diría que son “los” vectores propios de A .

Cuanta mayor sea la dimensión, mayor puede ser el número de autovalores. Para practicar, hallemos los valores propios de

$$A = \begin{pmatrix} 4 & -2 & -5 \\ -2 & 1 & 2 \\ 4 & -2 & -5 \end{pmatrix}.$$

Nuestro objetivo es, entonces, resolver $|A - \lambda I| = 0$. El cálculo del polinomio característico $|A - \lambda I|$ se puede llevar a cabo con la definición de determinante o con la regla de Sarrus. Sin embargo, con un poco de ingenio la cuentas se reducen utilizando las propiedades de los determinantes. Partimos de restar a la primera fila la tercera $|A - \lambda I|$

$$\stackrel{=}{f_1 \mapsto f_1 - f_3} \begin{vmatrix} -\lambda & 0 & \lambda \\ -2 & 1 - \lambda & 2 \\ 4 & -2 & -5 - \lambda \end{vmatrix} = -\lambda \begin{vmatrix} 1 & 0 & -1 \\ -2 & 1 - \lambda & 2 \\ 4 & -2 & -5 - \lambda \end{vmatrix} \stackrel{=}{\substack{f_2 \mapsto f_2 + 2f_1 \\ f_3 \mapsto f_3 - 4f_1}} \begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 - \lambda & 0 \\ 0 & -2 & -1 - \lambda \end{vmatrix}$$

y desarrollando por la primera columna, $|A - \lambda I| = -\lambda(1 - \lambda)(-1 - \lambda)$. De donde los autovalores son 0, 1 y -1 .

Las raíces múltiples del polinomio característico inducen algunas complicaciones que serán fundamentales después. Consideremos el siguiente ejemplo en dimensión 3:

$$A = \begin{pmatrix} 5 & 4 & 2 \\ -6 & -5 & -3 \\ 6 & 6 & 4 \end{pmatrix}.$$

Tras algunos cálculos, obtenemos que el polinomio característico es

$$\begin{vmatrix} 5 - \lambda & 4 & 2 \\ -6 & -5 - \lambda & -3 \\ 6 & 6 & 4 - \lambda \end{vmatrix} = -\lambda^3 + 4\lambda^2 - 5\lambda + 2 = -(\lambda - 1)^2(\lambda - 2).$$

De esta forma, tenemos $\lambda_1 = 1$ (doble) y $\lambda_2 = 2$. Aplicando reducción de Gauss al sistema $(A - \lambda_1 I)\vec{x} = \vec{0}$, se tiene

$$A - \lambda_1 I \xrightarrow{f_1 \mapsto f_1/2} \begin{pmatrix} 2 & 2 & 1 \\ -6 & -6 & -3 \\ 6 & 6 & 3 \end{pmatrix} \xrightarrow{\substack{f_2 \mapsto f_2 + 3f_1 \\ f_3 \mapsto f_3 - 3f_1}} \begin{pmatrix} 2 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Tomando $x_2 = \mu_1$, $x_3 = \mu_2$, como es habitual, se obtiene $x_1 = -\mu_1 - \mu_2/2$. Es decir, las soluciones son

$$\vec{x} = \mu_1(-1, 1, 0)^t + \mu_2(-1/2, 0, 1)^t.$$

Entonces hay dos autovectores asociados a λ_1 , $\vec{v}_1 = (-1, 1, 0)^t$, $\vec{v}_2 = (-1/2, 0, 1)^t$, con el abuso de notación. En realidad hay todo un subespacio de dimensión dos salvo el vector nulo. Por cierto, los autovectores siempre se pueden multiplicar por constantes no nulas, así pues es lícito escoger $\vec{v}_2 = (1, 0, -2)^t$.

Para λ_2 , resolviendo el sistema $(A - 2I)\vec{x} = \vec{0}$ se obtiene que los autovectores son los múltiplos no nulos de $\vec{v}_3 = (2, -3, 3)^t$.

Los ejemplos con matrices complejas no entrañan dificultades especiales más allá de que seguramente no estamos muy entrenados para resolver ecuaciones algebraicas con coeficientes complejos. En realidad tampoco lo estamos completamente para las de coeficientes reales y siempre los ejercicios deben estar muy preparados para que la ecuación característica sea asequible. Solo por dar un ejemplo, consideremos

$$A = \begin{pmatrix} 1 + 2i & -1 \\ -2 & 1 - i \end{pmatrix} \quad \text{que cumple} \quad |A - \lambda I| = \lambda^2 - (2 + i)\lambda + 1 + i.$$

Aplicando la solución de la ecuación de segundo grado, se obtiene $\lambda = \frac{2+i \pm i}{2}$, por tanto, los autovalores son $\lambda_1 = 1$ y $\lambda_2 = 1 + i$. Los autovectores son las soluciones no triviales de los sistemas:

$$\begin{pmatrix} 2i & -1 \\ -2 & -i \end{pmatrix} \vec{x} = \vec{0} \quad \text{y} \quad \begin{pmatrix} i & -1 \\ -2 & -2i \end{pmatrix} \vec{x} = \vec{0}.$$

En el primer sistema podemos tomar $\vec{v}_1 = (1, 2i)^t$ y en el segundo $\vec{v}_2 = (1, i)^t$ como soluciones no triviales.

2. Diagonalización

Una vez que hemos visto varios ejemplos, vamos al concepto fundamental que recoge la motivación inicial.

Se dice que una matriz $A \in \mathcal{M}_n(K)$ es *diagonalizable* sobre K si existe una base de K^n formada por autovectores. Se puede demostrar que *los autovectores correspondientes a autovalores distintos son siempre linealmente independientes* así que comprobar que $A \in \mathcal{M}_n(K)$ es diagonalizable se reduce a contar si hay n autovectores, con el abuso de notación que no gusta a los matemáticos de identificar el número de autovectores (que, en rigor, es infinito) con el número de autovectores linealmente independientes.

Si $A \in \mathcal{M}_n(K)$ es diagonalizable sobre K , entonces con la base de autovectores $\{\vec{v}_1, \dots, \vec{v}_n\}$ podemos formar una matriz cuadrada C poniéndolos en las columnas. Si λ_j es el autovalor

de \vec{v}_j ; entonces

$$AC = A(\vec{v}_1 \dots \vec{v}_n) = A(\lambda_1 \vec{v}_1 \dots \lambda_n \vec{v}_n) = (\vec{v}_1 \dots \vec{v}_n) \begin{pmatrix} \lambda_1 & 0 & 0 & \dots \\ 0 & \lambda_2 & 0 & \dots \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} = CD.$$

La matriz C es invertible, pues $\text{rg}(C) = n$ ya que sus columnas son linealmente independientes, y se concluye

$$A = CDC^{-1} \quad C = (\vec{v}_1 \dots \vec{v}_n) \quad \text{y} \quad D = \text{diag}(\lambda_1, \dots, \lambda_n)$$

donde diag indica la matriz diagonal construida a partir de sus argumentos.

Para entender por qué esta relación es interesante, consideremos $\vec{x} \mapsto \vec{y} = A\vec{x}$, que es un endomorfismo de K^n . Con el cambio $\vec{x} = C\vec{x}'$, $\vec{y} = C\vec{y}'$, la ecuación $\vec{y} = A\vec{x}$ se convierte en $C\vec{y}' = AC\vec{x}'$ que, usando la relación $A = CDC^{-1}$, equivale a $\vec{y}' = D\vec{x}'$. En conclusión, las matrices diagonalizables corresponden a endomorfismos que adquieren matrices diagonales tras un cambio de variable, lo que explica la notación.

Repasemos los ejemplos vistos hasta ahora en los que hemos calculado los vectores propios:

$$A_1 = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 5 & 4 & 2 \\ -6 & -5 & -3 \\ 6 & 6 & 4 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 1+2i & -1 \\ -2 & 1-i \end{pmatrix}.$$

Se tiene que A_1 es diagonalizable sobre \mathbb{R} (y sobre \mathbb{C}), porque hay dos autovectores $\vec{v}_1 = (2, -1)$, $\vec{v}_2 = (1, 1)$. De hecho si solo nos preguntasen si es diagonalizable, con ver que hay dos autovalores no harían falta más cuentas, pues para cada uno hay al menos un autovector. Por la misma razón, A_2 y A_4 son diagonalizables sobre \mathbb{C} . No lo son sobre \mathbb{R} porque en el primer caso hay autovalores complejos (que necesariamente dan autovectores complejos) y en el segundo porque la matriz de partida ya es compleja. Los cálculos sí serían necesarios para ver que A_3 es diagonalizable porque solo hay dos autovalores. Ya vimos que podíamos extraer dos de ellos para λ_1 y uno más para λ_3 . Según nuestros cálculos, una base válida de vectores propios es:

$$\mathcal{B} = \left\{ \vec{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \vec{v}_3 = \begin{pmatrix} 2 \\ -3 \\ 3 \end{pmatrix} \right\}.$$

En definitiva, todos los ejemplos que hemos visto son diagonalizables ya sea en \mathbb{R} o al menos en \mathbb{C} si hay autovalores que no son reales o la matriz no es real. Por ilustrar con un par de

ejemplos la relación $A = CDC^{-1}$, según los cálculos anteriores, para A_1 esta relación se podría escribir como

$$A_1 = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}^{-1},$$

mientras que para A_3 asegura

$$A_3 = \begin{pmatrix} 5 & 4 & 2 \\ -6 & -5 & -3 \\ 6 & 6 & 4 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 2 \\ 1 & 0 & -3 \\ 0 & -2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} -1 & 1 & 2 \\ 1 & 0 & -3 \\ 0 & -2 & 3 \end{pmatrix}^{-1}.$$

La ordenación de los autovectores debe ser coherente con la de los autovalores. Es decir, si ponemos en la columna i de C el autovector correspondiente a λ_j entonces debemos tomar necesariamente $d_{ii} = \lambda_j$.

La clave para que no todas las matrices sean diagonalizables sobre \mathbb{C} está en los autovalores múltiples, pues un resultado asegura que *un factor $(\lambda - \lambda_j)^{n_j}$ en el polinomio característico implica que el autoespacio de λ_j tiene a lo más dimensión n_j* (a lo más hay n_j autovectores con el abuso de notación), pero nada asegura que sea exactamente n_j .

Consideremos por ejemplo

$$A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$$

que tiene polinomio característico $\lambda^2 - 4\lambda + 4 = (\lambda - 2)^2$, lo que da lugar a un único valor propio $\lambda_1 = 2$, y al resolver $(A - 2I)\vec{x} = 0$ todas las soluciones son $\mu(-1, 1)^t$. No podemos extraer más que un autovector (y sus múltiplos).

Un ejemplo en dimensión 3 es:

$$A = \begin{pmatrix} 5 & 3 & 1 \\ -6 & -3 & -1 \\ 6 & 4 & 2 \end{pmatrix} \quad \text{que tiene} \quad |A - \lambda I| = -(\lambda - 1)^2(\lambda - 2).$$

A pesar de que el polinomio característico es el mismo que en un ejemplo diagonalizable anterior, dando lugar a $\lambda_1 = 1$, $\lambda_2 = 2$; la matriz A no es diagonalizable, porque con λ_2 solo obtenemos un autovector (y sus múltiplos) y λ_1 no provee dos autovectores independientes ya que aplicando eliminación de Gauss a $A - \lambda_1 I$

$$\begin{pmatrix} 4 & 3 & 1 \\ -6 & -4 & -1 \\ 6 & 4 & 1 \end{pmatrix} \xrightarrow{f_3 \rightarrow f_3 + f_2} \begin{pmatrix} 4 & 3 & 1 \\ -6 & -4 & -1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{f_2 \rightarrow 2f_2 + 3f_1} \begin{pmatrix} 4 & 3 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

lo que prueba que tiene rango dos y, por tanto, la solución de $(A - \lambda_1 I)\vec{x} = \vec{0}$ solo tiene un parámetro libre.

Para decidir si una matriz $A \in \mathcal{M}_n(K)$ es diagonalizable, no hace falta calcular los vectores propios, basta contarlos (estrictamente, contar los que son linealmente independientes). Según lo dicho sobre las multiplicidades y recordando que n menos el rango da el número de soluciones linealmente independientes de un sistema homogéneo, se tiene

Una matriz $A \in \mathcal{M}_n(K)$ es diagonalizable sobre K si y solo si $|A - \lambda I|$ tiene todas sus raíces λ_j en K y se cumple que $n - \text{rg}(A - \lambda_j I)$ coincide con la multiplicidad de λ_j .

La condición de que las raíces estén en K solo es relevante cuando $K = \mathbb{R}$, ya que las raíces de un polinomio están siempre en \mathbb{C} y pueden ser o no reales. Como la suma de las multiplicidades da el grado del polinomio una alternativa algo artificial es:

Una matriz $A \in \mathcal{M}_n(K)$ es diagonalizable sobre K si y solo si $\sum (n - \text{rg}(A - \lambda_j I))$ donde la suma es sobre todos los autovalores (distintos) λ_j que están en K .

En palabras, se pide que la suma de las dimensiones de los autoespacios es igual a la dimensión del espacio total.

Terminemos este apartado sobre diagonalización con un resultado teórico de importancia capital en física y matemáticas y, por extensión, en algunos de sus usos en ingeniería.

Teorema espectral. Si $A \in \mathcal{M}_n(\mathbb{R})$ es simétrica, es decir, $A = A^t$, entonces es diagonalizable sobre \mathbb{R} y siempre se puede escoger una base ortonormal de autovectores.

Podría cambiarse ortonormal por ortogonal, porque la única diferencia está en si normalizamos o no lo hacemos. El enunciado se puede parafrasear diciendo que la C que aparece al diagonalizar siempre se puede escoger de forma que sea una matriz ortogonal.

Hay un resultado más general que incluye a las matrices $A \in \mathcal{M}_n(\mathbb{C})$ que cumplen $A = \overline{A}^t$, pero no lo veremos aquí.

Una consecuencia fácil del resultado anterior, aunque no totalmente inmediata si hay autovalores con varios vectores propios (independientes) es:

Si $A \in \mathcal{M}_n(\mathbb{R})$ es simétrica, los autovectores correspondientes a autovalores distintos son siempre ortogonales.

Tanto esta consecuencia como el propio teorema espectral son muy sorprendentes. A pesar de que la demostración no es complicada, suena muy extraño que la simetría de una matriz real obligue a que todos los valores propios sean reales, que haya suficientes vectores propios para diagonalizar y que además sean perpendiculares.

Exploremos un ejemplo de dimensión 2 y otro de dimensión 3. El propósito es solo maravillarse de que estas propiedades se cumplan, los cálculos serán como los hechos hasta ahora.

Vamos a hallar una base ortonormal en la que la A diagonalice y una ortogonal en la que lo haga B donde

$$A = \begin{pmatrix} 2 & 6 \\ 6 & -7 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

En el primer caso el polinomio característico es $|A - \lambda I| = \lambda^2 + 5\lambda + 50$ y, resolviendo la ecuación, $\lambda_1 = 5$ y $\lambda_2 = -10$. De $(A - 5I)\vec{x} = \vec{0}$ se obtiene el vector propio $\vec{v}_1 = (2, 1)^t$ (y sus múltiplos). De la misma forma, $(A + 10I)\vec{x} = \vec{0}$ conduce a $\vec{v}_2 = (-1, 2)^t$. Son ortogonales en la línea de la consecuencia del teorema espectral. La base buscada se obtiene normalizando: $\mathcal{B} = \{\vec{u}_1, \vec{u}_2\}$ con $\vec{u}_1 = \frac{1}{\sqrt{5}}(2, 1)^t$ y $\vec{u}_2 = \frac{1}{\sqrt{5}}(-1, 2)^t$.

Para B , una manera más o menos rápida de obtener el polinomio característico es

$$|B - \lambda I| \stackrel{c_1 \mapsto c_1 - c_3}{=} \begin{vmatrix} 1 - \lambda & 1 & 1 \\ 0 & 4 - \lambda & 1 \\ \lambda - 1 & 1 & 2 - \lambda \end{vmatrix} \stackrel{f_3 \mapsto f_1 + f_3}{=} \begin{vmatrix} 1 - \lambda & 1 & 1 \\ 0 & 4 - \lambda & 1 \\ 0 & 2 & 3 - \lambda \end{vmatrix}.$$

Desarrollando por la primera columna, $|B - \lambda I| = (1 - \lambda)(\lambda^2 - 7\lambda + 10) = (1 - \lambda)(\lambda - 2)(\lambda - 5)$ y los autovalores son $\lambda_1 = 1$, $\lambda_2 = 2$ y $\lambda_3 = 5$. Resolviendo los sistemas $(B - \lambda_j I)\vec{x} = \vec{0}$ se obtienen los autovectores (salvo multiplicar por constantes)

$$\vec{v}_1 = (-1, 0, 1)^t, \quad \vec{v}_2 = (1, -1, 1)^t \quad \text{y} \quad \vec{v}_3 = (1, 2, 1)^t.$$

En sintonía con el resultado, salen automáticamente ortogonales y conforman la base buscada.

3. Sucesiones y potencias

La diagonalización de matrices simplifica los cálculos que aparecen en ciertas aplicaciones. Aquí veremos una de ellas, desde un punto de vista teórico, que guarda relación con temas que forman parte de los planes de estudios de ingeniería.

Muchos algoritmos están basados en la repetición: la salida del algoritmo, o parte de ella, se toma como nueva entrada cierto número de veces. Si tal algoritmo viene dado por una aplicación lineal $K^n \rightarrow K^n$, surge el problema de determinar la sucesión de vectores $\vec{x}_1, \vec{x}_2, \dots$ definidos mediante

$$(1) \quad \vec{x}_{k+1} = A\vec{x}_k \quad \text{con} \quad k = 0, 1, 2, \dots$$

donde $A \in \mathcal{M}_n(K)$ y $\vec{x}_0 \in K^n$ están fijados.

En principio, esto es muy fácil: $\vec{x}_1 = A\vec{x}_0$, $\vec{x}_2 = A\vec{x}_1 = A^2\vec{x}_0$ y, en general, $\vec{x}_k = A^k\vec{x}_0$. Esta fórmula es exacta, pero demasiado teórica. No nos da intuición acerca del comportamiento a la larga de la sucesión, lo cual es a menudo fundamental para responder a algunas preguntas naturales. Deseamos una fórmula más explícita, a través de un cálculo sencillo de A^k , y ahí es donde entra la diagonalización.

Sea $A \in \mathcal{M}_n(K)$ diagonalizable. Si $C \in \mathcal{M}_n(K)$ tiene por columnas una base de vectores propios y D es la matriz diagonal compuesta por los valores propios respectivos, $A^k = CD^kC^{-1}$. En particular, la solución de (1) es $\vec{x}_k = CD^kC^{-1}\vec{x}_0$.

El punto importante a observar es que el cálculo de D^k es sencillo, pues se reduce a elevar a k cada elemento de la diagonal.

La prueba del resultado es poco más que repetir nuestras conclusiones sobre diagonalización. Sabemos $A = CDC^{-1}$. De aquí, $A^2 = CDC^{-1}CDC^{-1} = CD^2C^{-1}$, $A^3 = A^2A = CD^2C^{-1}CDC^{-1} = CD^3C^{-1}$ y así sucesivamente.

Comencemos con un ejemplo en el que sería fácil deducir el resultado experimentalmente. Queremos resolver (1) para

$$A = \begin{pmatrix} -11 & 6 \\ -20 & 11 \end{pmatrix} \quad \text{y} \quad \vec{x}_0 \in \mathbb{R}^2 \text{ arbitrario.}$$

El polinomio característico es $(\lambda^2 - 121) + 120 = \lambda^2 - 1$, por tanto, tenemos dos valores propios $\lambda_1 = -1$ y $\lambda_2 = 1$. Vectores propios respectivos son $\vec{v}_1 = (3, 5)^t$ y $\vec{v}_2 = (1, 2)^t$. Escribiendo $\vec{x}_0 = (a, b)^t$, de acuerdo con el resultado anterior,

$$\vec{x}_k = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} (-1)^k & 0 \\ 0 & 1^k \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}^{-1} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} (6(-1)^k - 5)a + 3(1 - (-1)^k)b \\ 10((-1)^k - 1)a + (6 - 5(-1)^k)b \end{pmatrix}$$

donde para la última igualdad se han operado las matrices. Más sencillo es darse cuenta de que $D^2 = I$, porque los autovalores son ± 1 . Por tanto, $D^k = D$, y equivalentemente $A^k = A$, para k impar mientras que $D^k = I$, y $A^k = I$, para k par. Por consiguiente,

$$\vec{x}_k = A\vec{x}_0 = \begin{pmatrix} -11a + 6b \\ -20a + 11b \end{pmatrix} \quad \text{si } k \text{ es impar} \quad \text{y} \quad \vec{x}_k = \vec{x}_0 \quad \text{si } k \text{ es par.}$$

Esto es lo mismo que la fórmula anterior escrito de forma menos aparatosa.

El siguiente ejemplo sería más difícil de resolver experimentalmente:

$$\begin{cases} x_{k+1} = 8x_k + 6y_k \\ y_{k+1} = -3x_k - y_k \end{cases} \quad \text{con } x_0 = y_0 = 1.$$

En forma matricial, escribiendo $\vec{x}_k = (x_k, y_k)^t$,

$$\vec{x}_{k+1} = A\vec{x}_k \quad \text{con} \quad A = \begin{pmatrix} 8 & 6 \\ -3 & -1 \end{pmatrix} \quad \text{y} \quad \vec{x}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

El polinomio característico es $|A - \lambda I| = \lambda^2 - 7\lambda + 10$, por tanto, los valores propios resultan $\lambda_1 = 2$, $\lambda_2 = 5$. Para calcular vectores propios respectivos, debemos hallar soluciones de los sistemas

$$(A - \lambda_1 I)\vec{x} = \begin{pmatrix} 6 & 6 \\ -3 & -3 \end{pmatrix} \vec{x} = \vec{0} \quad \text{y} \quad (A - \lambda_2 I)\vec{x} = \begin{pmatrix} 3 & 6 \\ -3 & -6 \end{pmatrix} \vec{x} = \vec{0}.$$

Elecciones naturales son $\vec{v}_1 = (-1, 1)^t$ y $\vec{v}_2 = (-1, 1)^t$, con lo que la sucesión \vec{x}_k responde a la fórmula

$$\vec{x}_k = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2^k & 0 \\ 0 & 5^k \end{pmatrix} \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Operando, se tiene

$$\vec{x}_k = \begin{pmatrix} 2 \cdot 5^k - 2^k & 2 \cdot 5^k - 2^{k+1} \\ -5^k + 2^k & -5^k + 2^{k+1} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \cdot 5^k - 3 \cdot 2^k \\ -2 \cdot 5^k + 3 \cdot 2^k \end{pmatrix}.$$

Es decir, $x_k = 4 \cdot 5^k - 3 \cdot 2^k$, $y_k = -2 \cdot 5^k + 3 \cdot 2^k$. Cambiando $(1, 1)^t$ por un vector genérico, se tendría la solución para cualquier valor inicial.

Hallemos ahora una fórmula general para el vector \vec{x}_k determinado por (1) con

$$A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix} \quad \text{y} \quad \vec{x}_0 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

La matriz está medio diagonalizada lo que hace que sea sencillo calcular el polinomio característico $(\lambda^2 - 5\lambda + 6)(5 - \lambda)$. Con ello, los valores propios son $\lambda_1 = 2$, $\lambda_2 = 3$ y $\lambda_3 = 5$. Son distintos y entonces A es diagonalizable. Algunos cálculos llevan a que vectores propios correspondientes a estos valores propios son $\vec{v}_1 = (2, 1, 0)^t$, $\vec{v}_2 = (1, 1, 0)^t$ y $\vec{v}_3 = (0, 0, 1)^t$. El resultado implica

$$\vec{x}_k = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2^k & 0 & 0 \\ 0 & 3^k & 0 \\ 0 & 0 & 5^k \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 2^{k+1} - 3^k \\ 2^k - 3^k \\ -5^k \end{pmatrix}.$$

A modo de comprobación, para $k = 0$ está claro que obtenemos \vec{x}_0 y para $k = 1$ el resultado coincide con $A\vec{x}_0$.

Un pequeño truco cuando \vec{x}_0 toma un valor explícito, que evita los tediosos cálculos de la inversa, es que $C^{-1}\vec{x}_0$ es la solución del sistema $C\vec{x} = \vec{x}_0$. Así, en el ejemplo anterior $C^{-1}(1, 0, -1)^t$ es $(1, -1, -1)^t$ porque esta es la solución de $2x_1 + x_2 = 1$, $x_1 + x_2 = 0$, $x_3 = -1$.

4. La forma canónica de Jordan

Teniendo en cuenta problemas como el del apartado anterior, la pregunta natural es qué hacer con las matrices que no son diagonalizables. La respuesta es que casi se diagonalizan y su forma casi diagonal todavía sirve de algo en diversas situaciones, por ejemplo, para el cálculo de potencias, aunque no entraremos en ello.

Concretamente, hay una teoría matemática que permite expresar cualquier $A \in \mathcal{M}_n(\mathbb{C})$ como

$$A = CJC^{-1} \quad \text{con } C, J \in \mathcal{M}_n(\mathbb{C})$$

donde J es diagonal si A es diagonalizable, mientras que si no lo es, J es casi diagonal en el sentido de que sus elementos fuera de la diagonal se anulan excepto que algunos $J_{i+1,i}$ pueden valer 1, además solo pueden aparecer si $J_{ii} = J_{i+1,i+1}$. Esta matriz J que es diagonal o casi diagonal, se dice que es la *forma canónica de Jordan*. La teoría general que explica dónde situar los unos y cómo construir C se complica bastante según crece la dimensión y es de limitado interés para la mayoría de los ingenieros. En este curso nos limitamos a los casos de dimensión 2 y 3 en los que el problema es fácil, sobre todo a la hora de hallar J . En lo sucesivo nos restringiremos al caso no diagonalizable, pues el diagonalizable ya sabemos tratarlo.

En dimensión 2 la determinación de la forma canónica de Jordan es inmediata, solo hay una posibilidad. Si $A \in \mathcal{M}_2(\mathbb{C})$ no es diagonalizable entonces

$$J = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}$$

donde λ_1 es el autovalor de A . Necesariamente es único, pues si hubiera dos también habría dos autovectores linealmente independientes y, por tanto, sería diagonalizable.

Si $A \in \mathcal{M}_3(\mathbb{C})$ no es diagonalizable, puede tener uno o dos autovalores. Si tiene dos, λ_1 y λ_2 , uno de ellos, digamos λ_1 , será de multiplicidad dos. En esa situación, la forma canónica de Jordan es

$$J = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

Si hay solo un autovalor λ_1 entonces resultan dos posibilidades:

$$\text{o bien } J = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix} \quad \text{o bien } J = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}.$$

La primera ocurre cuando $\text{rg}(A - \lambda_1 I) = 1$ y la segunda cuando $\text{rg}(A - \lambda_1 I) = 2$.

Analicemos los siguientes ejemplos, para convencernos de que es sencillo decidir cuál es la forma canónica J :

$$A_1 = \begin{pmatrix} 0 & 4 \\ -1 & -4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 & 1 \\ 3 & -2 & -1 \\ -4 & 4 & 3 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 2 & 3 & -1 \\ 0 & 2 & 0 \\ 0 & -1 & 2 \end{pmatrix}.$$

Con cálculos bastante rápidos se deduce que los polinomios característicos son respectivamente $(\lambda + 2)^2$, $(1 - \lambda)^2(\lambda + 1)$ y $(2 - \lambda)^3$, resultando entonces un solo autovalor para A_1 y A_3 y dos para A_2 . Como $A_1 + 2I$ no es la matriz nula, A_1 tiene un solo vector propio, salvo múltiplos, y estamos en el caso no diagonalizable del resultado en dimensión 2 con $\lambda_1 = -2$. Para A_2 , hay dos autovalores, el de multiplicidad dos es $\lambda_1 = 1$ y como $(A_2 - I)\vec{x} = \vec{0}$ tiene un conjunto de soluciones de dimensión 1, de nuevo no es diagonalizable y se tiene el primer caso dentro de dimensión 3. Finalmente, A_3 solo tiene el autovalor $\lambda_1 = 2$ y $\text{rg}(A_3 - 2I) = 2$ lo que muestra que no es diagonalizable y que es del último tipo discutido en dimensión 3.

En definitiva, las formas canónicas respectivas de los ejemplos anteriores son:

$$J_1 = \begin{pmatrix} -2 & 1 \\ 0 & -2 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{y} \quad J_3 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Veamos otros ejemplos ligeramente singulares:

$$A_4 = \begin{pmatrix} 5 & -3 & -1 \\ 3 & -1 & -1 \\ 0 & 0 & 2 \end{pmatrix}, \quad A_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -4 & -1 \end{pmatrix}, \quad A_6 = \begin{pmatrix} i & -1 \\ -1 & -i \end{pmatrix}.$$

La matriz A_4 solo tiene un autovalor, $\lambda_1 = 2$, y $\text{rg}(A_4 - 2I) = 1$ lo que prueba que no es diagonalizable y es del segundo caso de dimensión 3. Por otro lado, A_5 tiene dos autovalores, $\lambda_1 = 1$ y $\lambda_2 = -1$. Como existen dos vectores independientes con $(A_5 - I)\vec{v} = \vec{0}$ entonces es diagonalizable (porque a estos le podemos añadir un autovector correspondiente a λ_2 para formar una base). Por tanto, la forma canónica de Jordan es la D del caso diagonalizable. La matriz A_6 es solo para ilustrar que el caso complejo no alberga ninguna dificultad adicional, aparte de nuestra posible falta de práctica con estos números. El polinomio característico es λ^2 , por tanto, tenemos un solo autovalor $\lambda_1 = 0$. Está claro que A_6 tiene un solo vector propio, salvo múltiplos, por consiguiente, no es diagonalizable y la forma canónica de Jordan es la del resultado en dimensión 2.

Para resumir, las formas canónicas de Jordan respectivas son:

$$J_4 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad J_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{y} \quad J_6 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

La determinación de la matriz C en el caso no diagonalizable, hay ciertas recetas que cobran sentido si uno entiende la teoría general, que escapa de los contenidos del curso. La columnas de C se dice que forman una *base de Jordan*. El hecho de que C sea invertible equivale a que verdaderamente formen una base de \mathbb{C}^n .

Si

$$J = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix} \quad \text{entonces} \quad C = (\vec{v}_1 | \vec{v}_2) \quad \text{con} \quad \begin{cases} \vec{v}_2 \text{ tal que } (A - \lambda_1 I)\vec{v}_2 \neq \vec{0}, \\ \vec{v}_1 = (A - \lambda_1 I)\vec{v}_2. \end{cases}$$

La condición $(A - \lambda_1 I)\vec{v}_2 \neq \vec{0}$ da mucha libertad para escoger \vec{v}_2 , por ello hay infinidad de posibles matrices C , todas ellas válidas.

Para dimensión 3, comenzamos con el caso de dos autovalores. Si

$$J = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} \quad \text{entonces} \quad C = (\vec{v}_1 | \vec{v}_2 | \vec{v}_3) \quad \text{con} \quad \begin{cases} \vec{v}_2 \text{ tal que } (A - \lambda_1 I)\vec{v}_2 \neq \vec{0} \\ \quad \text{y } (A - \lambda_1 I)^2 \vec{v}_2 = \vec{0}, \\ \vec{v}_1 = (A - \lambda_1 I)\vec{v}_2, \\ \vec{v}_3 \text{ autovector para } \lambda_2. \end{cases}$$

Los dos casos con un solo autovalor se tratan de la siguiente manera. El primero es muy similar al que acabamos de discutir, de hecho se podrían unificar. Si

$$J = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix} \quad \text{entonces} \quad C = (\vec{v}_1 | \vec{v}_2 | \vec{v}_3) \quad \text{con} \quad \begin{cases} \vec{v}_2 \text{ tal que } (A - \lambda_1 I)\vec{v}_2 \neq \vec{0}, \\ \vec{v}_1 = (A - \lambda_1 I)\vec{v}_2, \\ \vec{v}_3 \text{ autovector para } \lambda_1 \\ \quad \text{no proporcional a } \vec{v}_1. \end{cases}$$

Finalmente, si

$$J = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix} \quad \text{entonces} \quad C = (\vec{v}_1 | \vec{v}_2 | \vec{v}_3) \quad \text{con} \quad \begin{cases} \vec{v}_3 \text{ tal que } (A - \lambda_1 I)^2 \vec{v}_3 \neq \vec{0}, \\ \vec{v}_2 = (A - \lambda_1 I)\vec{v}_3, \\ \vec{v}_1 = (A - \lambda_1 I)\vec{v}_2. \end{cases}$$

Veamos el cálculo de C en algunos de los ejemplos anteriores. Para A_1 , el vector $\vec{v}_2 = (1, 0)^t$ cumple $(A_1 + 2I)\vec{v}_2 = (2, -1)^t \neq \vec{0}$ y entonces C con columnas $(2, -1)^t$ y $(1, 0)^t$ es una elección válida. Si uno quisiera comprobar el resultado habría que verificar:

$$\begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 4 \\ -1 & -4 \end{pmatrix}.$$

En realidad, numéricamente es más simple y equivalente comprobar $A_1\vec{v}_1 = \lambda_1\vec{v}_1$ y $A_1\vec{v}_2 = \vec{v}_1 + \lambda_1\vec{v}_2$

Con la elección $\vec{v} = (1, 1)^t$ habríamos obtenido una matriz C con columnas $(6, -3)^t$ y $(1, 1)^t$ que también es válida. El primer vector debe ser autovector y por ello ha resultado un múltiplo del calculado antes.

En el caso de A_2 , el valor propio de multiplicidad dos es $\lambda_1 = 1$ y se tiene

$$(A_2 - \lambda_1 I)^2 = \begin{pmatrix} 0 & 0 & 0 \\ -8 & 8 & 4 \\ 8 & -8 & -4 \end{pmatrix} \quad \text{y} \quad (A_2 - \lambda_1 I)^2 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \vec{0}, \quad (A_2 - \lambda_1 I) \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \neq \vec{0}.$$

Por tanto, podemos tomar $\vec{v}_2 = (1, 0, 2)^t$, $\vec{v}_1 = (A_2 - I)\vec{v}_2 = (1, 1, 0)^t$ y \vec{v}_3 un vector propio correspondiente a $\lambda_2 = -1$, por ejemplo $\vec{v}_3 = (0, -1, 1)^t$, lo que lleva a

$$C = (\vec{v}_1 | \vec{v}_2 | \vec{v}_3) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 2 & 1 \end{pmatrix}$$

La comprobación ahora de que se cumple $A_2 = CJC^{-1}$ sería un poco larga (y, por tanto, no muy aconsejable para verificar el resultado).

Para A_3 el único valor propio era $\lambda_1 = 2$ y había una única celda de Jordan. El cálculo de $(A_3 - 2I)^2$ lleva a una matriz cuyo único elemento no nulo es el segundo de la primera fila, por tanto, podemos tomar $\vec{v}_3 = (0, 1, 0)^t$, $\vec{v}_2 = (A_3 - 2I)\vec{v}_3 = (3, 0, -1)^t$ y $\vec{v}_1 = (A_3 - 2I)\vec{v}_2 = (1, 0, 0)^t$.

Como último ejemplo de base de Jordan, hallemos una para A_4 . Recordemos que era de la segunda forma del resultado en \mathbb{C}^3 . La diferencia con respecto a A_2 es que ahora $\lambda_1 = \lambda_2$, el único autovalor es 2. La matriz $(A_4 - 2I)^2$ es nula, en sintonía con lo dicho sobre las potencias de celdas de Jordan y entonces no hay otra restricción sobre \vec{v}_2 más que $(A_4 - 2I)\vec{v}_2 \neq \vec{0}$ (que no sea autovector). Por ejemplo, tomemos $\vec{v}_2 = (1, 0, 0)^t$ que da lugar a $\vec{v}_1 = (A_4 - 2I)\vec{v}_2 = (3, 3, 0)^t$. Como \vec{v}_3 hay que escoger un autovector que no sea múltiplo de \vec{v}_1 , por ejemplo $(1, 0, 3)^t$.

El resto de los ejemplos antes enunciados no aportan ninguna novedad.