

Capítulo 5

Teoría espectral

5.1. 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 es una simplificación deseable que A sea diagonal y por tanto determinada solo por n números. Sin conocer siquiera esas aplicaciones, suena creíble que sea un objetivo a buscar si pensamos en que un sistema lineal de ecuaciones con matriz diagonal es trivial, porque todas las ecuaciones están desacopladas. De la misma forma, 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, que no se trata en este curso.

Nos preguntamos, por tanto, si existe alguna base en la que la matriz de un endomorfismo f dado sea diagonal y en ese caso diremos que es f es *diagonalizable*.

El hecho de que exista una base $\{\vec{v}_1, \dots, \vec{v}_n\}$ en la que nuestro endomorfismo f se *diagonalice*, esto es, adquiera una matriz diagonal, quiere decir que al utilizar coordenadas \vec{x} en dicha base

$$(5.1) \quad \vec{x} \xrightarrow{f} D\vec{x} \quad \text{con} \quad D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Escrito de otro modo, $f(\vec{v}_j) = \lambda_j \vec{v}_j$. Asignemos nombres a esta situación: Dado un espacio vectorial V sobre K y un endomorfismo $f : V \rightarrow V$, se dice que $\vec{v} \in V - \{\vec{0}\}$ es un *vector propio* o *autovector* si $f(\vec{v}) = \lambda \vec{v}$ para algún $\lambda \in K$. A este λ se le llama *valor propio* o *autovalor*.

Un pequeño abuso de notación muy habitual, que no evitaremos en este capítulo, es hablar de los vectores y valores propios de una matriz cuadrada $A \in \mathcal{M}_n(K)$, suponiendo que uno se refiere a los del endomorfismo de K^n dado por $f(\vec{x}) = A\vec{x}$. De la misma manera, se dice que A es *diagonalizable* si existe una base $\{\vec{v}_1, \dots, \vec{v}_n\}$,

necesariamente formada por autovectores, en la que $f(\vec{x}) = A\vec{x}$ adquiere una matriz diagonal. Según lo que sabemos de cambios de base, si C es la matriz cuyas columnas son las coordenadas de los \vec{v}_j en la base canónica (o en la que esté referida A), entonces

$$(5.2) \quad C^{-1}AC = D$$

para D como en (5.1) y $A\vec{v}_j = \lambda_j\vec{v}_j$.

Si conocemos un valor propio λ , hallar los vectores propios correspondientes equivale a resolver $f(\vec{x}) = \lambda\vec{x}$ lo cual, como casi todo en este curso, se reduce a un sistema de ecuaciones lineales. Por ser más concretos, si $f(\vec{x}) = A\vec{x}$ y λ es autovalor, los autovectores respectivos son las soluciones no nulas de

$$(5.3) \quad (A - \lambda I)\vec{x} = \vec{0}.$$

Desde un punto de vista más teórico, los autovectores correspondientes a λ son los vectores no nulos de $\text{Ker}(f - \lambda \text{Id})$ con Id el endomorfismo identidad (el que deja todo fijo). A este núcleo se llama *autoespacio* de λ .

El problema es entonces encontrar un método para hallar los valores propios. El siguiente resultado lo traduce en resolver una ecuación algebraica.

Proposición 5.1.1. *Dado un espacio vectorial V sobre K de dimensión finita y un endomorfismo f , se tiene que $\lambda \in K$ es autovalor si y solo si resuelve la ecuación $\det(A - \lambda I) = 0$ donde A es la matriz de f en una base fijada.*

Por supuesto lo habitual es usar la base canónica. A la ecuación $\det(A - \lambda I) = 0$ se le llama *ecuación característica* y a $\det(A - \lambda I)$ *polinomio característico*. Es fácil convencerse de que realmente es un polinomio y que su grado es $\dim V$.

Demostración. La condición necesaria y suficiente para que $\lambda \in K$ sea autovalor es que existan autovectores respectivos, esto es, $\text{Ker}(f - \lambda \text{Id}) \neq \{\vec{0}\}$. Dicho de otro modo, que el sistema homogéneo (5.3) tenga solución no trivial, lo que equivale a que el rango de $A - \lambda I$ sea menor que el número de incógnitas (Teorema 1.2.2) y esto a que su determinante sea no nulo (Teorema 3.3.4). \square

Por ejemplo, estudiemos si es diagonalizable el endomorfismo $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definido por

$$f(\vec{x}) = A\vec{x} \quad \text{con} \quad A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}.$$

Primero hallamos 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 $\lambda_2 = 5$. 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 autovectores¹. El endomorfismo f (o la matriz A) es diagonalizable ya que tenemos la base $\mathcal{B} = \{(2, -1)^t, (1, 1)^t\}$ formada por autovectores. La comprobación es que, si cambiamos a la base \mathcal{B} , de acuerdo con (5.2),

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

Aunque aquí lo eludiremos, es habitual abusar de la notación y restringir el nombre *autovectores* 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. Por supuesto, hay ambigüedad en la posible base en la que se diagonaliza un endomorfismo.

Una ecuación polinómica sobre \mathbb{C} siempre tiene raíces en \mathbb{C} pero en \mathbb{R} esto no es cierto en general. Así, el endomorfismo $f : K^2 \rightarrow K^2$

$$f(\vec{x}) = A\vec{x} \quad \text{con} \quad A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

no tiene autovalores (ni por tanto autovectores) cuando $K = \mathbb{R}$ porque la ecuación característica es $\lambda^2 + 1 = 0$, por tanto en este caso no es diagonalizable. Sin embargo, si $K = \mathbb{C}$, $\lambda_1 = i$ y $\lambda_2 = -i$ son vectores 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 consiguiente f se diagonaliza en la base $\mathcal{B} = \{(i, 1)^t, (-i, 1)^t\}$.

Vemos entonces que un problema para que un endomorfismo $f : V \rightarrow V$ no sea diagonalizable es que V sea un espacio vectorial sobre $K = \mathbb{R}$ y no haya suficientes autovalores. El problema podría persistir cuando $K = \mathbb{C}$ si hay raíces múltiples de la ecuación característica. La multiplicidad no depende de la base, es decir los polinomios característicos están asociados a los endomorfismos, más que a sus matrices. Esto es fácil pero no obvio.

¹Quizá te llame la atención que las columnas de $A - \lambda_1 I$ sean autovectores para λ_2 y las de $A - \lambda_2 I$ lo sean para λ_1 . Esto es un hecho general del caso de dimensión 2 con $\lambda_1 \neq \lambda_2$. Se deja al lector interesado que investigue la razón.

Lema 5.1.2. *El polinomio característico no depende de la base elegida.*

Demostración. Si $A' = C^{-1}AC$

$$|A' - \lambda I| = |C^{-1}AC - \lambda C^{-1}C| = |C^{-1}||A - \lambda I||C| = |A - \lambda I|$$

donde se ha aplicado la propiedad multiplicativa de los determinantes. \square

Incluso si tenemos tantos autovalores como $\dim V$, cabe preguntarse si los autovectores respectivos gozan de la independencia lineal requerida para que sean parte de una base. El siguiente resultado asegura que sí.

Proposición 5.1.3. *Si $\vec{v}_1, \dots, \vec{v}_n$ son autovectores de $f : V \rightarrow V$ con autovalores distintos, entonces son linealmente independientes.*

Demostración. Supongamos que no lo fueran y llegaremos a una contradicción. Si uno de los vectores propios es combinación lineal de los otros, los podemos reordenar de forma que se tenga:

$$\vec{v}_1 = \mu_2 \vec{v}_2 + \dots + \mu_m \vec{v}_m \quad \text{con } \mu_j \neq 0.$$

Además, consideraremos que m es lo menor posible. Si λ_j es el autovalor que corresponde a \vec{v}_j entonces para $m > 2$

$$\vec{v}_1 = \frac{f(\vec{v}_1) - \lambda_m \vec{v}_1}{\lambda_1 - \lambda_m} = \mu_2 \frac{\lambda_2}{\lambda_1 - \lambda_m} \vec{v}_2 + \dots + \mu_{m-1} \frac{\lambda_{m-1}}{\lambda_1 - \lambda_m} \vec{v}_{m-1},$$

lo que contradice que m sea mínimo. Si $m = 2$, el mismo argumento lleva a la contradicción $\vec{v}_1 = \vec{0}$. \square

Con esto ya estamos en condiciones de dar un resultado teórico que resuelve el problema que nos habíamos planteado.

Teorema 5.1.4. *Un endomorfismo $f : V \rightarrow V$, con V de dimensión finita, es diagonalizable si y solo si*

$$\sum_{i=1}^m \dim \text{Ker}(f - \lambda_i \text{Id}) = \dim V$$

donde $\lambda_1, \dots, \lambda_m$ son los autovalores (distintos) de f .

Demostración. Si $\mathcal{B}_i = \{\vec{v}_{i1}, \dots, \vec{v}_{id_i}\}$ es una base de $\text{Ker}(f - \lambda_i \text{Id})$, de dimensión d_i , entonces $\mathcal{B} = \bigcup_{i=1}^m \mathcal{B}_i$ consta de $\sum_{i=1}^m d_i$ elementos que son vectores propios linealmente independientes por la Proposición 5.1.3, en particular $\sum_{i=1}^m d_i \leq \dim V$. Si se da la igualdad, forman una base (Proposición 2.2.4) y entonces f es diagonalizable. Recíprocamente, si existe una base de autovectores, como cada uno está en algún $\text{Ker}(f - \lambda_i \text{Id})$, necesariamente se tiene la igualdad de dimensiones. \square

Los núcleos anteriores son no triviales por definición de autovector y se tiene la consecuencia inmediata:

Corolario 5.1.5. Dada $A \in \mathcal{M}_n(K)$ si $\det(A - \lambda I) = 0$ tiene n raíces distintas, entonces A es diagonalizable.

Por ejemplo, si nos tomamos el trabajo de hacer los cálculos,

$$A = \begin{pmatrix} -1 & 1 & -3 \\ 4 & 0 & 4 \\ -1 & -1 & 1 \end{pmatrix} \implies |A - \lambda I| = -\lambda^3 + 4\lambda = -\lambda(\lambda^2 - 4).$$

Por tanto los autovalores son $\lambda = -2, 0, 2$ y la matriz A es diagonalizable.

Consideremos $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dado por

$$f(\vec{x}) = A\vec{x} \quad \text{con} \quad A = \begin{pmatrix} 5 & 4 & 2 \\ -6 & -5 & -3 \\ 6 & 6 & 4 \end{pmatrix}.$$

Con 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 $\lambda_2 = 2$. Los rangos de $A - \lambda_1 I$ y de $A - \lambda_2 I$ son respectivamente

$$\text{rg} \begin{pmatrix} 4 & 4 & 2 \\ -6 & -6 & -3 \\ 6 & 6 & 3 \end{pmatrix} = 1 \quad \text{y} \quad \text{rg} \begin{pmatrix} 3 & 4 & 2 \\ -6 & -7 & -3 \\ 6 & 6 & 2 \end{pmatrix} = 2.$$

Por tanto $\dim \text{Ker}(f - \lambda_1 \text{Id}) = 2$ y $\dim \text{Ker}(f - \lambda_2 \text{Id}) = 1$, y el Teorema 5.1.4 asegura que el endomorfismo f es diagonalizable. En realidad el segundo rango nos lo podíamos haber ahorrado porque por definición de autovalor sabemos que la segunda dimensión es al menos uno.

Hay otra razón por la cual el segundo rango nos lo podríamos haber ahorrado y es que se puede demostrar, como veremos más adelante, que $\dim \text{Ker}(f - \lambda_i \text{Id})$ es a lo más la multiplicidad de la raíz $\lambda = \lambda_i$ en la ecuación característica.

Si se nos pidiera una base \mathcal{B} en la que diagonaliza habría que resolver los sistemas $(A - \lambda_1 I)\vec{x} = \vec{0}$ y $(A - \lambda_2 I)\vec{x} = \vec{0}$ para obtener bases $\mathcal{B}_1 = \{\vec{v}_1, \vec{v}_2\}$ y $\mathcal{B}_2 = \{\vec{v}_3\}$ de los núcleos. Una solución válida es

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

Por ahora 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. Esta es la situación común pero hay contraejemplos. Así, el endomorfismo $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dado por

$$f(\vec{x}) = A\vec{x} \quad \text{con} \quad A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$$

tiene polinomio característico $\lambda^2 - 4\lambda + 4 = (\lambda - 2)^2$, lo que da lugar a un único valor propio $\lambda_1 = 2$, y para que fuera diagonalizable necesitaríamos $\dim \text{Ker}(f - 2\text{Id}) = 2$ lo que requiere $\text{rg}(A - 2I) = 0$, que obviamente es falso.

Un ejemplo en dimensión 3 es $f(\vec{x}) = A\vec{x}$ con

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

A pesar de que el polinomio característico es el mismo que antes, dando lugar a $\lambda_1 = 1$, $\lambda_2 = 2$; no es diagonalizable, porque $\dim \text{Ker}(f - \lambda_2\text{Id}) = 1$ pero λ_1 no provee dos autovectores independientes:

$$\dim \text{Ker}(f - \lambda_1\text{Id}) = 3 - \text{rg} \begin{pmatrix} 4 & 3 & 1 \\ -6 & -4 & -1 \\ 6 & 4 & 1 \end{pmatrix} = 3 - 2 = 1.$$

Los ejemplos que involucran números complejos 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 preparados para que la ecuación característica sea asequible. Solo por dar un ejemplo, consideremos el endomorfismo $F : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ dado por

$$f(\vec{x}) = A\vec{x} \quad \text{con} \quad A = \begin{pmatrix} 1 + 2i & -1 \\ -2 & 1 - i \end{pmatrix}$$

y estudiemos si es diagonalizable y, en caso afirmativo, una base en que se diagonaliza. Se tiene

$$|A - \lambda I| = \begin{vmatrix} 1 + 2i - \lambda & -1 \\ -2 & 1 - i - \lambda \end{vmatrix} = \lambda^2 - (2 + i)\lambda + 1 + i.$$

En principio debemos resolver una ecuación cuadrática con números complejos, quizá no recuerdes o no hayas aprendido nunca cómo hacer raíces cuadradas de números complejos, por ello seguiremos un camino alternativo, siempre basado en que el problema está preparado. A simple vista se observa que $\lambda_1 = 1$ es solución y, dividiendo entre $\lambda - 1$ o de forma más sencilla², se sigue que el otro autovalor es $\lambda_2 = 1 + i$. El Corolario 5.1.5 asegura que es diagonalizable y para hallar la base debemos calcular autovectores dando soluciones no triviales de los sistemas $(A - \lambda_i I)\vec{x} = \vec{0}$ que son:

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

En realidad no hay nada que resolver, como sabemos que hay infinitas soluciones, una de las ecuaciones es prescindible y basta con ajustar una solución de la otra. Por ejemplo, en el primer sistema podemos tomar $\vec{v}_1 = (1, 2i)^t$ y en el segundo $\vec{v}_2 = (1, i)^t$; y una posibilidad para la base buscada es $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$.

²Por ejemplo usando que para $\lambda^2 - B\lambda + C$ siempre C es el producto de raíces.

Exprimiendo el silicio [opcional]. El comando básico de `matlab/octave` para calcular valores y vectores propios es `eig`. Tiene un comportamiento un poco peculiar: si lo usamos tal cual o recuperamos su salida asignándosela a una variable, por ejemplo `la = eig(A)`, entonces genera los valores propios, pero si se la asignamos a dos variables como `[C,D] = eig(A)` entonces en `D` estará la forma diagonal y en `C` el cambio de base, cuyas columnas son autovectores. Así pues, para comprobar el primer ejemplo que vimos, podemos usar:

```

1 % Matriz original
2 A = [3,2; 1,4];
3 % Muestra los autovalores
4 eig(A)
5 % D es la forma diagonal y C el cambio de base
6 [C,D] = eig(A)
7 % Primer autovector de la base
8 C(:,1)
9 % Segundo autovector de la base
10 C(:,2)
11 % Comprobación: esto recupera la matriz original
12 C*D*inv(C)

```

Una particularidad es que siempre escoge las columnas de `C` de manera que sean de norma uno, por ello los autovectores resultan un múltiplo de los que hemos elegido nosotros. Si cambiamos la matriz por `[3,1; -1,1]`, aparentemente el programa también la diagonaliza lo cual está en contradicción con lo que habíamos probado. Esto se debe a que las matrices no diagonalizables en \mathbb{C}^n están infinitamente cerca de otras que sí lo son y los errores de redondeo engañan a `matlab/octave`. La `C` que ofrece como resultado está muy próxima a ser singular y su inversa tiene elementos del orden de 30 millones.

Aquí va un pequeño programa que se explica a sí mismo con tres instrucciones para hacer los cálculos de esta sección con `sagemath`:

```

1 # Matriz original
2 A = matrix(2,2,[3,2,1,4])
3 # El polinomio característico
4 print(A.characteristic_polynomial())
5 # Los autovalores
6 print(A.eigenvalues())
7 # Los autovalores, autovectores y multiplicidades
8 print(A.eigenvectors_right())

```

Ciertamente, es un poco raro el nombre `eigenvectors_right`. Se debe a que existe también un `eigenvectors_left` que calcula los autovectores de la traspuesta, lo que correspondería a multiplicar vectores fila a la izquierda. En el programa anterior la salida del comando, escrita en una línea, es la lista

$$[(5, [(1, 1)], 1), (2, [(1, -1/2)], 1)]$$

En cada elemento de la lista, el primer número es el autovalor, la sublista que sigue una base del autoespacio y el tercer número la multiplicidad del autovalor en el polinomio característico. Así para la matriz $20I_2$ obtendríamos `[(20, [(1, 0), (0, 1)], 2)`. Este tipo de información es suficiente para saber si es diagonalizable o no. Por ejemplo para la matriz `[3,1; -1,1]` tendríamos `[(2, [(1, -1)], 2)]` y como solo hay un

vector en la base del autoespacio (menos que la multiplicidad), no es diagonalizable. Existe también un comando `eigenspaces_right` que da los autoespacios y sus dimensiones.

Con `sagemath` el análogo de `eig` es `eigenmatrix_right` que siempre produce un par de matrices. Su primer elemento es la matriz diagonal y el segundo la matriz de cambio de base (al revés que en `matlab/octave`). Cuando se aplica a una matriz no diagonalizable las columnas de esta matriz correspondientes a autovectores ausentes se completan con cero. Así una manera de saber si una matriz es diagonalizable o no, es con un código como el siguiente:

```

1 # Matriz original
2 A = matrix(3,3, [5,3,1, -6,-3,-1, 6,4,2])
3
4 [D, C] = A.eigenmatrix_right()
5 # Si C tiene columnas nulas no es diagonalizable
6 if C.determinant()==0: print('No es diagonalizable')
7 else: print('Si es diagonalizable')
```

Tal como está, corresponde al último ejemplo no diagonalizable que hemos visto.

5.2. El teorema espectral

Los cambios de base utilizados para diagonalizar no tienen nada de particular porque dada cualquier base es posible construir un endomorfismo que tenga a sus elementos como vectores propios. Nos planteamos una pregunta, en principio, un poco rebuscada: en un espacio vectorial con producto escalar, queremos saber cuándo ese cambio de base preserva las distancias. En \mathbb{R}^3 , geoméricamente esto significa que, salvo simetrías, el cambio corresponde a girar la cabeza, lo que muestra que la pregunta tampoco es tan antinatural.

Para no introducir el concepto de *aplicación adjunta*³, vamos a trabajar todo el rato con matrices de $\mathcal{M}_n(K)$, especialmente de $\mathcal{M}_n(\mathbb{C})$. Según sabíamos del capítulo anterior, preservar distancias es lo mismo que tener una matriz unitaria, y estas matrices, son justamente las que tienen columnas que forman una base ortonormal, por la Proposición 4.4.3. Por tanto, nuestra pregunta se traduce en saber cuándo dada $A \in \mathcal{M}_n(\mathbb{C})$ existe $U \in \mathcal{M}_n(\mathbb{C})$ unitaria tal que

$$U^{-1}AU = D \quad \text{con } D \text{ diagonal.}$$

Despejando, $A = UDU^{-1} = UDU^\dagger$, donde la última igualdad viene de la propiedad $U^\dagger U$ que caracteriza las matrices unitarias. Trasponiendo y conjugando $A^\dagger = UD^\dagger U^\dagger$. De aquí,

$$AA^\dagger = UDU^\dagger UD^\dagger U^\dagger = UDD^\dagger U^\dagger \quad \text{y} \quad A^\dagger A = UD^\dagger U^\dagger UDU^\dagger = UD^\dagger DU^\dagger.$$

³Por si tienes curiosidad: Para $f : V \rightarrow V$ donde V tiene un producto escalar, se dice que $g : V \rightarrow V$ es la *aplicación adjunta* si $\langle \vec{v}, f(\vec{w}) \rangle = \langle g(\vec{v}), \vec{w} \rangle$ para todo $\vec{v}, \vec{w} \in V$. El concepto es tan relevante que se generaliza al caso de aplicaciones lineales generales (no necesariamente endomorfismos) incluso sin que haya un producto escalar.

Para una matriz diagonal, $DD^\dagger = D^\dagger D$, así que ambos productos coinciden.

Llamemos *matriz normal* a una matriz $A \in \mathcal{M}_n(\mathbb{C})$ que verifica $AA^\dagger = A^\dagger A$. El cálculo anterior muestra que las matrices que diagonalizan en una base ortonormal de \mathbb{C}^n (con el producto escalar usual) son normales. El teorema que da nombre a esta sección⁴ afirma que el recíproco también es cierto.

Teorema 5.2.1 (Teorema espectral). *Una matriz $A \in \mathcal{M}_n(\mathbb{C})$ es normal si y solo si existe una base ortonormal de \mathbb{C}^n en la que diagonaliza.*

Demostración. Lo que tenemos que ver es que si A es normal entonces existe U unitaria tal que $U^{-1}AU$ es diagonal. Como estamos en \mathbb{C} , siempre existe un autovalor $\lambda_1 \in \mathbb{C}$ y, normalizando, un autovector correspondiente \vec{u} con $\|\vec{u}\| = 1$. El proceso de Gram-Schmidt permite partir de \vec{u} y añadir nuevos vectores hasta completar una base ortonormal \mathcal{B} de \mathbb{C}^n . En \mathcal{B} , la matriz A pasará a ser A' de la forma

$$(5.4) \quad A' = \left(\begin{array}{c|ccc} \lambda_1 & c_2 & \dots & c_n \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array} \right)$$

porque \vec{u} tiene coordenadas $\vec{e}_1 = (1, 0, \dots, 0)^t$ en \mathcal{B} y se debe cumplir $A'\vec{e}_1 = \lambda_1\vec{e}_1$. Se tiene $A' = U_1^{-1}AU_1 = U_1^\dagger AU_1$ con $U_1 \in \mathcal{M}_n(\mathbb{C})$ unitaria, porque \mathcal{B} es ortonormal, y $AA^\dagger = A^\dagger A$ implica que A' también es normal. Si calculamos el primer elemento de $A'(A')^\dagger$ resulta ser $|\lambda_1|^2 + \sum_{j=2}^n |c_j|^2$ mientras que el de $(A')^\dagger A'$ es $|\lambda_1|^2$. Como deben coincidir, se concluye $c_2 = \dots = c_n = 0$.

Es fácil ver que B es normal, por serlo A' , por tanto podemos repetir el argumento cambiando A por B y concluir que existe $V_1 \in \mathcal{M}_{n-1}(\mathbb{C})$ unitaria tal que $B' = V_1^{-1}BV_1$ cumple $b'_{1j} = b'_{j1} = 0$ para $j \neq 1$, como en (5.4). Esto implica, como antes, que si B'' es la matriz B suprimiendo la primera fila y la primera columna, B'' es normal y se tiene

$$U_2^{-1}U_1^{-1}AU_1U_2 = \left(\begin{array}{cc|ccc} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & B'' & \\ 0 & 0 & & & \end{array} \right) \quad \text{con} \quad U_2 = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & V_1 & \\ 0 & & & \end{array} \right)$$

donde U_2 es unitaria, porque V_1 lo es.

Repetiendo el proceso n veces, se sigue que existen $U_1, \dots, U_n \in \mathcal{M}_n(\mathbb{C})$ unitarias tales que $U^{-1}AU$ es diagonal con $U = U_1U_2 \dots U_n$. Para terminar, basta observar que el producto de matrices unitarias es también una matriz unitaria. \square

⁴Con cierta razón, alguien podría disentir de la terminología manteniendo que este no es el teorema espectral sino un teorema espectral para aplicaciones normales. De acuerdo, pero es uno de los teoremas espectrales más importantes. Todavía más, la expresión “teoría espectral” del título de este capítulo fue introducida por D. Hilbert en relación con un resultado que extiende a dimensión infinita uno de los corolarios de nuestro teorema espectral.

Para deducir las dos consecuencias más empleadas del Teorema 5.2.1, necesitamos un resultado auxiliar bastante sorprendente pero con una prueba muy breve.

Proposición 5.2.2. *Si una matriz $A \in \mathcal{M}_n(\mathbb{C})$ verifica $A = A^\dagger$, sus autovalores son reales.*

Demostración. Si \vec{v} es un autovector con autovalor λ , cada una de las siguientes igualdades es sencilla:

$$\lambda \|\vec{v}\|^2 = (A\vec{v})^t \vec{v} = \vec{v}^t A^t \vec{v} = \vec{v}^t \overline{(A^\dagger \vec{v})} = \vec{v}^t \overline{(A\vec{v})} = \bar{\lambda} \|\vec{v}\|^2,$$

y los números reales son los únicos de \mathbb{C} que coinciden con sus conjugados. \square

Para poner en valor la prueba anterior, veamos en qué se manifiesta que una matriz simétrica arbitraria de $\mathcal{M}_2(\mathbb{R})$ tenga valores propios reales cuando hacemos los cálculos a mano. Debemos resolver

$$\begin{vmatrix} a - \lambda & b \\ b & c - \lambda \end{vmatrix} = 0 \quad \text{que equivale a} \quad \lambda^2 - (a + c)\lambda + b^2 = 0.$$

El discriminante de esta ecuación es $(a + c)^2 - 4b^2$ que claramente es no negativo y por tanto nunca lleva a soluciones complejas. Dar una prueba similar en el caso 3×3 sería muchísimo más difícil.

Las consecuencias anunciadas más que dos son una, porque el caso de real se deduce del complejo, simplemente hay una diferencia en la nomenclatura. Se llaman *matrices hermíticas* a las matrices $A \in \mathcal{M}_n(\mathbb{C})$ que verifican $A = A^\dagger$, como en el resultado anterior. Cuando $A \in \mathcal{M}_n(\mathbb{R})$ esto es mismo que $A = A^t$ y se tienen las *matrices simétricas* (reales).

Corolario 5.2.3. *Una matriz hermítica siempre tiene autovalores reales y es diagonalizable en una base ortonormal de \mathbb{C}^n .*

Corolario 5.2.4. *Una matriz real simétrica siempre tiene autovalores reales y es diagonalizable en una base ortonormal de \mathbb{R}^n .*

Algo menos importante, pero aún así destacable es:

Corolario 5.2.5. *Una matriz $U \in \mathcal{M}_n(\mathbb{C})$ unitaria es diagonalizable en una base ortonormal de \mathbb{C}^n y tiene autovalores de módulo uno.*

Demostración. Ya sabemos $UU^\dagger = U^\dagger U = I$, por tanto es normal y el Teorema 5.2.1 asegura que existen V unitaria y D diagonal con $V^{-1}UV = D$. De $V^{-1} = V^\dagger$ se deduce $D^\dagger = V^{-1}U^\dagger V$ y por tanto $DD^\dagger = I$, así que los elementos de la diagonal de D tienen módulo uno. \square

La doble implicación en el enunciado del Teorema 5.2.1 no deja lugar a que diagonalicemos matrices no normales en bases ortonormales. El siguiente resultado asegura que, no obstante, podemos conseguir en ellas muchos ceros en los elementos de la matriz. La prueba está en realidad contenida en la del teorema espectral.

Proposición 5.2.6. *Dada $A \in \mathcal{M}_n(\mathbb{C})$, existe una base ortonormal en la que se transforma en una matriz triangular superior T , esto es, con $t_{ij} = 0$ para $i > j$.*

Demostración. Los primeros pasos de la prueba del Teorema 5.2.1 muestran que existe $U_1 \in \mathcal{M}_n(\mathbb{C})$ tal que $A' = U_1^{-1}AU_1$ es de la forma indicada en (5.4), esto es, los elementos de la primera columna bajo la diagonal se anulan. Repitiendo el argumento con B , como se hizo en esa prueba, tenemos otra matriz unitaria $U_1 \in \mathcal{M}_n(\mathbb{C})$ tal que $U_2^{-1}U_1^{-1}AU_1U_2$ tiene un aspecto similar pero ahora también los elementos de la segunda columna bajo la diagonal se anulan. A la larga, llegamos a que existen $U_1, \dots, U_n \in \mathcal{M}_n(\mathbb{C})$ unitarias tales que $U^{-1}AU$ es triangular superior con $U = U_1U_2 \cdots U_n$.

En resumen, la prueba es la misma que la del Teorema 5.2.1 salvo que, como no suponemos que A sea normal, no tenemos en general $c_2 = \cdots = c_n = 0$. \square

Una consecuencia no directa de este último resultado, completa un cabo suelto de la sección anterior que no tiene nada que ver con productos escalares.

Corolario 5.2.7. *Sea $f : K^n \rightarrow K^n$ un endomorfismo con $K = \mathbb{R}$ o \mathbb{C} . Si λ es un valor propio de f con multiplicidad m en el polinomio característico, entonces*

$$\dim \text{Ker}(f - \lambda \text{Id}) \leq m.$$

Demostración. Digamos que $f(\vec{x}) = A\vec{x}$ en la base canónica. Por la Proposición 5.2.6 existe U tal que $U^{-1}AU = T$ es triangular superior y, por el Lema 5.1.2, $|A - \lambda I| = |T - \lambda I|$. Las soluciones de $|T - \lambda I| = 0$ son los elementos t_{ii} de la diagonal con la multiplicidad dada por el número de veces que aparecen (por la parte final de la Proposición 3.1.2). Por otro lado, resolviendo el sistema $(T - \lambda I)\vec{x} = \vec{0}$ por eliminación de Gauss-Jordan para uno de estos $\lambda = t_{ii}$ que aparezca m veces en la diagonal, está claro que a lo más hay m columnas que no son columnas pivote. La Proposición 2.3.4 asegura que el número de estas columnas es la dimensión del núcleo. \square

Para $K = \mathbb{C}$ el polinomio característico tiene n raíces contando multiplicidades, esto es, la suma de las multiplicidades es n . Recordando el Teorema 5.1.4 se deduce:

Corolario 5.2.8. *Una matriz $A \in \mathcal{M}_n(\mathbb{C})$ es diagonalizable si y solo si a cada valor propio le corresponden tantos vectores propios linealmente independientes como su multiplicidad.*

Los resultados de esta sección anticipan que hay diagonalización en ciertos casos e incluso que se produce en una base ortonormal, pero el método para calcular dicha base sigue siendo el que conocemos, por eso no hay mucho lugar para ejemplos realmente nuevos. Si no existen multiplicidades mayores que 1, no gozamos de más libertad para escoger los autovectores que multiplicar por constantes y lo único que tendremos que hacer para que la base sea ortonormal es ajustar la norma a uno. Las multiplicidades mayores conllevan demasiada libertad y es necesario poner algo de nuestra parte para alcanzar la ortonormalidad.

Ilustremos esta situación con los siguientes ejemplos de matrices simétricas reales, sin entrar en los detalles de los cálculos:

$$A = \begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 5 & 2 & 4 \\ 2 & 8 & -2 \\ 4 & -2 & 5 \end{pmatrix}.$$

El polinomio característico de A es $\lambda^2 - 25\lambda$ de modo que tenemos los autovalores $\lambda_1 = 0$ y $\lambda_2 = 25$. Para el primero, los autovectores son proporcionales a $(4, -3)^t$ y para el segundo, a $(3, 4)^t$. Automáticamente son ortogonales, en sintonía con la teoría. Construir una base ortonormal solo requiere normalizarlos para llegar a $\mathcal{B}_A = \{(4/5, -3/5)^t, (3/5, 4/5)^t\}$. Por otro lado, el polinomio característico de B es $\lambda^3 - 18\lambda^2 + 81\lambda = \lambda(\lambda - 9)^2$. Así pues $\lambda_1 = 0$ con multiplicidad 1 y $\lambda_2 = 9$ con multiplicidad 2. Para λ_1 los vectores propios son múltiplos de $(2, -1, -2)^t$ mientras que para λ_2 son todos los $(x, y, z)^t \neq \vec{0}$ que satisfacen $2x - y - 2z = 0$. Tenemos por tanto $(1, 0, 1)^t$ y $(0, 2, -1)^t$, que forman una base del espacio de soluciones pero no son ortogonales. Con un paso de Gram-Schmidt, reemplazamos el último vector por $(1, 4, -1)^t$ y basta normalizar el resultado para obtener

$$\mathcal{B}_B = \left\{ \frac{1}{3}(2, -1, -2)^t, \frac{1}{\sqrt{2}}(1, 0, 1)^t, \frac{1}{3\sqrt{2}}(1, 4, -1)^t \right\}.$$

Expriumiendo el silicio [opcional]. Los autovalores y autovectores se siguen calculando con `matlab/octave` y `sagemath` de la misma manera sean o no las matrices normales. Como `eig` en `matlab/octave` normaliza los vectores, para matrices simétricas reales y hermíticas, en general, la base de autovectores será ortonormal. Esto no es así con `sagemath`.

Aprovechemos este apartado para ilustrar una situación que aparece sobre todo en física. Digamos que tenemos una matriz simétrica real y modificamos ligeramente sus elementos, preservando la simetría, ¿cómo se modifican sus valores y vectores propios? La teoría de la perturbación responde a esta pregunta. Un caso muy particular es que, cuando los valores propios son distintos, si la modificación consiste en sumar una misma cantidad pequeña ϵ a cada elemento, entonces cada autovalor λ pasa a ser aproximadamente $\lambda + (\sum_j v_j)^2 \epsilon$ donde v_j son las coordenadas del autovector correspondiente.

El siguiente código `matlab/octave` genera matrices simétricas A al azar con elementos en $[0, 1]$ y comprueba la aproximación para $\epsilon = \text{pert}$.

```

1 % Dimensión
2 N = 3;
3 % Perturbación
4 pert = 0.01
5
6 % Matriz simétrica al azar
7 A = rand(N);
8 A = (A+A')/2;
9
10 % Valores y vectores propios de A y A perturbada
11 [V,D] = eig(A);
12 [Vp,Dp] = eig(A+pert);
13
```

```

14 % Muestra los resultados
15 for j = 1:N
16     disp('-----')
17     fprintf('lambda_%d=%%.8f\n',j,D(j,j))
18     fprintf('lambda_%d_perturbado_%d=%%.8f\n',j,Dp(j,j))
19     fprintf('Aproximación_%d=%%.8f\n',D(j,j) + pert*sum(V(:,j))^2)
20 end

```

Con $\epsilon = 0.01$, como aparece en el programa, en una de las ejecuciones se obtuvieron los resultados correspondientes a la siguiente tabla, lo que da una idea de la bondad de la aproximación:

	original	perturbado	aproximación
λ_1	-0.15537117	-0.15530942	-0.15530780
λ_2	0.06759992	0.06869853	0.06872612
λ_3	1.20275030	1.23158995	1.23156073

La utilidad de este tipo de resultados es que a veces uno tiene la solución exacta para cierto problema de autovalores y autovectores y quiere resolver otro numéricamente parecido. En física se emplea sobre todo en el caso de dimensión infinita.

5.3. Algunas aplicaciones

Como se mencionó al principio del capítulo, un motivo para diagonalizar un endomorfismo o una matriz es simplificar los cálculos que parecen en ciertas aplicaciones. Aquí veremos dos de ellas, desde un punto de vista teórico, que guardan relación con temas que forman parte de los planes de estudios de los ingenieros.

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 $K^n \rightarrow K^n$, la pregunta natural es determinar la sucesión de vectores $\vec{x}_1, \vec{x}_2, \dots$ definidos mediante

$$(5.5) \quad \vec{x}_{n+1} = A\vec{x}_n \quad \text{con } n = 0, 1, 2, \dots$$

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, $\vec{x}_n = A^n\vec{x}_0$. 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 decidir la eficiencia del algoritmo. Deseamos una fórmula más explícita, a través de un cálculo sencillo de A^n , y ahí es donde entra la diagonalización.

Proposición 5.3.1. *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, $A^n = CD^nC^{-1}$. En particular, la solución de (5.5) es $\vec{x}_n = CD^nC^{-1}\vec{x}_0$.*

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

La prueba del resultado es poco más que repetir nuestras conclusiones sobre diagonalización.

Demostración. Sabíamos $C^{-1}AC = D$ por (5.2) y, por tanto, $A = CDC^{-1}$. 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. \square

Comencemos con un ejemplo en el que sería fácil deducir el resultado experimentalmente. Queremos resolver (5.5) 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,

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

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 $D^n = D$, y equivalentemente $A^n = A$, para n impar mientras que $D^n = I$, y $A^n = I$, para n par. Por consiguiente,

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

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

Hallemos ahora una fórmula general para el vector \vec{x}_n determinado por (5.5) 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 $\vec{v}_3 = (0, 0, 1)^t$. La Proposición 5.3.1 implica

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

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

Como último ejemplo vamos a buscar una fórmula para los *números de Fibonacci*⁵ definidos por la recurrencia

$$F_{n+2} = F_{n+1} + F_n \quad \text{con} \quad F_0 = 0, \quad F_1 = 1.$$

Es decir, partiendo de 0 y 1 cada número de Fibonacci se obtiene sumando los dos anteriores. Los primeros son 0, 1, 1, 2, 3, 5, 8, 13, 21... y no es difícil sospechar un crecimiento exponencial. En principio, parece que la ecuación que los define escapa de las posibilidades de la Proposición 5.3.1 porque no está en forma matricial, pero hay un método para conseguirlo que es fácil de generalizar. Escribimos la relación anterior como

$$\vec{x}_{n+1} = A\vec{x}_n \quad \text{con} \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{y} \quad \vec{x}_n = \begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix}.$$

La segunda ecuación es tautológica, dice que $F_{n+1} = F_{n+1}$, es solo un artificio para introducir una matriz cuadrada a la que aplicar la Proposición 5.3.1.

Los autovalores de A son $\lambda_{\pm} = (1 \pm \sqrt{5})/2$ y dos autovectores respectivos son $(\lambda_{\pm}, 1)^t$. Entonces se tiene

$$\vec{x}_n = \begin{pmatrix} \lambda_+ & \lambda_- \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_+^n & 0 \\ 0 & \lambda_-^n \end{pmatrix} \begin{pmatrix} \lambda_+ & \lambda_- \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

porque $\vec{x}_0 = (F_1, F_0)^t = (1, 0)^t$. Operar esto da un poco de pereza pero si lo hacemos y nos fijamos en la segunda coordenada de \vec{x}_n obtendremos una flamante fórmula explícita para los números de Fibonacci:

$$F_n = \frac{\lambda_+^n - \lambda_-^n}{\sqrt{5}} = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right).$$

Como $-1 < \lambda_- < 0$, para n grande λ_-^n es despreciable y $F_n \approx \lambda_+^n / \sqrt{5}$. Por ejemplo,

$$\frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^{20} = 6765.000029 \dots \quad \text{y} \quad F_{20} = 6765.$$

La segunda aplicación que estudiaremos es relativa a la solución de sistemas lineales de ecuaciones diferenciales. Es decir, ecuaciones del tipo:

$$(5.6) \quad \vec{x}'(t) = A\vec{x}(t), \quad \text{con} \quad \vec{x}(0) = \vec{x}_0$$

⁵El nombre deriva de que fueron introducidos, en los albores del siglo XIII, por Leonardo de Pisa apodado Fibonacci. Hoy en día tienen cierta presencia en la cultura popular por su aparición en diferentes novelas y películas.

donde A es una matriz cuadrada constante, $\vec{x}(t)$ es una función vectorial dependiendo de una variable real t y $\vec{x}'(t)$ es su derivada. La mayor parte de las veces $A \in \mathcal{M}_n(\mathbb{R})$ y $\vec{x}(t) \in \mathbb{R}^n$, pero la teoría sirve igualmente para el caso complejo.

En los cursos de matemáticas se prueba que (5.6) tiene solución única aunque no nos preocuparemos por ello aquí.

Proposición 5.3.2. Si $A \in \mathcal{M}_n(\mathbb{C})$ es diagonalizable, $\vec{x}(t) = C \exp(Dt)C^{-1}\vec{x}_0$ es solución de (5.6), donde se ha usado la notación de la Proposición 5.3.1 y $\exp(Dt)$ es la matriz diagonal obtenida al reemplazar d_{ii} en D por $e^{d_{ii}t}$.

Definiendo⁶ $\exp(At) = C \exp(Dt)C^{-1}$, la solución es $\vec{x}(t) = \exp(At)\vec{x}_0$ y esta fórmula es bastante sugestiva porque formalmente verifica (5.6) si no se tienen escrúpulos sobre derivar matrices.

Demostración. Sea $\vec{y} = C^{-1}\vec{x}$, entonces la ecuación $\vec{x}'(t) = A\vec{x}(t)$ se escribe como $(C\vec{y})'(t) = AC\vec{y}(t)$. Las constantes son irrelevantes al derivar, por tanto también equivale a $C\vec{y}'(t) = AC\vec{y}(t)$. Recordando (5.2), $\vec{y}'(t) = D\vec{y}(t)$, lo cual se desacopla en n ecuaciones diferenciales en una dimensión:

$$y_i'(t) = d_{ii}y_i(t) \quad \text{con } i = 1, 2, \dots, n.$$

Cada una de ellas tiene solución $y_i(t) = c_i e^{d_{ii}t}$ con c_i una constante arbitraria, así que $\vec{y}(t) = \exp(Dt)\vec{c}$ con \vec{c} un vector constante arbitrario y hemos probado que $\vec{x}(t) = C\vec{y} = C \exp(At)\vec{c}$ resuelve $\vec{x}'(t) = A\vec{x}(t)$. Para que se cumpla la condición $\vec{x}(0) = \vec{x}_0$ basta tomar $\vec{c} = C^{-1}\vec{x}_0$. \square

Supongamos que queremos resolver

$$\vec{x}' = A\vec{x} \quad \text{con } A = \begin{pmatrix} 7 & -10 \\ 3 & -4 \end{pmatrix} \quad \text{y } \vec{x}(0) = \begin{pmatrix} 7 \\ 4 \end{pmatrix}.$$

La ecuación característica es $\lambda^2 - 3\lambda + 2 = 0$ y produce los valores propios $\lambda_1 = 2$ y $\lambda_2 = 1$. Como $\{(2, 1)^t, (5, 3)^t\}$ es una base de vectores propios correspondiente, la Proposición 5.3.2 conduce a la solución,

$$\vec{x}(t) = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} e^{2t} & 0 \\ 0 & e^t \end{pmatrix} \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 7 \\ 4 \end{pmatrix} = \begin{pmatrix} 2e^{2t} + 5e^t \\ e^{2t} + 3e^t \end{pmatrix}.$$

Muchas veces estas ecuaciones diferenciales no se presentan en forma vectorial. Por ejemplo, un enunciado podría ser resolver

$$\begin{cases} x' = -11x + 6y, \\ y' = -20x + 11y, \end{cases} \quad \text{con } x(0) = y(0) = 1.$$

⁶En realidad lo habitual es definir $\exp(At)$ como la serie infinita $\sum_{n=0}^{\infty} A^n t^n / n!$ que da la solución tanto si A es diagonalizable como si no. Esto permite establecer una conexión, que no analizaremos, entre las dos aplicaciones vistas aquí, pudiéndose deducir ambas de que A^n tenga una fórmula sencilla en una base adecuada.

Con la notación anterior, esto es $\vec{x}'(t) = A\vec{x}(t)$ donde $\vec{x} = (x, y)^t$, $\vec{x}_0 = (1, 1)^t$ y A es la matriz del primer ejemplo de solución de (5.5). Tomando de ese ejemplo los autovalores y autovectores, se tiene

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} e^{-t} & 0 \\ 0 & e^t \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3e^{-t} - 2e^t \\ 5e^{-t} - 4e^t \end{pmatrix}.$$

A modo de comprobación, verifiquemos la primera ecuación del enunciado. Se tiene $x'(t) = -3e^{-t} - 2e^t$ y esto coincide con $-11(3e^{-t} - 2e^t) + 6(5e^{-t} - 4e^t)$.

En analogía con el estudio anterior de los números de Fibonacci y dentro de las ecuaciones básicas que maneja un ingeniero o un físico, consideremos la siguiente ecuación de un *oscilador armónico*:

$$y''(t) + 4y(t) = 0 \quad \text{con} \quad y(0) = 3, \quad y'(0) = 2.$$

El truco para convertir esta ecuación diferencial en vectorial y así estar en disposición de aplicar la Proposición 5.3.2, es similar al empleado con la recurrencia que definía los números de Fibonacci. Tomamos $\vec{x}(t) = (y'(t), y(t))^t$ y entonces la ecuación equivale a

$$\vec{x}' = A\vec{x} \quad \text{con} \quad A = \begin{pmatrix} 0 & -4 \\ 1 & 0 \end{pmatrix} \quad \text{y} \quad \vec{x}(0) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

De nuevo la segunda ecuación no aporta información, simplemente indica $y' = y'$, es un artificio para completar la forma matricial. La ecuación característica es $\lambda^2 + 4 = 0$ y debemos ir a \mathbb{C} para resolverla. Los valores propios son $\lambda_1 = -2i$ y $\lambda_2 = 2i$. Una base de autovalores respectivos es $\{(-2i, 1)^t, (2i, 1)^t\}$ con lo que tenemos

$$\begin{pmatrix} y'(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} -2i & 2i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e^{-2it} & 0 \\ 0 & e^{2it} \end{pmatrix} \begin{pmatrix} -2i & 2i \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Solo estamos interesados en la segunda coordenada, que nos da $y(t)$. Si nos tomamos el trabajo de operar, resulta

$$y(t) = \frac{3+i}{2}e^{-2it} + \frac{3-i}{2}e^{2it} = 3\cos(2t) + \sin(2t),$$

donde para la segunda igualdad se ha sustituido $e^{\pm 2it} = \cos(2t) \pm i\sin(2t)$.

A primera vista, parece un milagro que los números complejos se simplifiquen y el resultado sea real. La explicación analítica es la existencia y unicidad de soluciones de ecuaciones diferenciales reales mientras que la explicación algebraica es la invariancia por la conjugación. Si A es una matriz real, al conjugar un par autovalor y autovector, se obtiene otro par válido. De ahí que $C \exp(Dt)C^{-1}$ no cambie al conjugar, ni por tanto la solución de $\vec{x}(t)$ de (5.6).

Un último apunte es que seguro que has aprendido o aprenderás un método más rápido para resolver ecuaciones diferenciales de la forma $y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y = 0$, que incluyen el oscilador armónico. El propósito del cálculo anterior es más ilustrar la generalidad que la eficiencia.

Exprimiendo el silicio [opcional]. Dada en `matlab/octave` una matriz cuadrada A , su exponencial se calcula con `expm(A)`. La m es de *matrix* porque `exp(A)` calcula la exponencial de los elementos por separado (desconozco qué utilidad puede tener esto). El siguiente código evalúa `expm(A*t)*x0` para muchos valores de t en un intervalo o equivalentemente la solución en esos valores del sistema correspondiente de ecuaciones diferenciales. Con ellos se dibuja una aproximación de las gráficas de las componentes de la solución.

Tal como está, la matriz es la del segundo ejemplo y por tanto las gráficas aproximan a las de las funciones $f_1(t) = 3e^{-t} - 2e^t$ y $f_2(t) = 5e^{-t} - 4e^t$.

```

1 % Matriz del sistema
2 A = [-11,6; -20,11]
3
4 % Vector inicial
5 x0 = [1; 1]
6
7 % Intervalo
8 inter = linspace(0,1,300);
9
10 sol = [];
11 for t = inter
12     sol = [sol, expm(A*t)*x0];
13 end
14
15 % Dibuja las componentes de la solución
16 plot(inter,sol')
```

La definición de `inter` significa que estamos en el intervalo $[0, 1]$ tomando 300 puntos intermedios. El bucle en cada paso añade a `sol` la columna correspondiente a la solución en t . El comando `plot` dibuja el resultado. En él aparece `sol'` en lugar de `sol` porque este comando dibuja las columnas como gráficas separadas.

Con `sagemath` la exponencial de una matriz A se calcula mediante `exp(A)`.

5.4. La forma canónica de Jordan

Nos planteamos ahora qué hacer con los endomorfismos que no son diagonalizables. La respuesta es que casi se diagonalizan y su forma casi diagonal todavía sirve de algo en aplicaciones como las vistas antes, aunque no entraremos en este último punto. La teoría general es complicada y de limitado interés para un ingeniero. Por ello nos restringiremos a los casos de dimensión 2 y 3. Como antes, para que las cosas sean más tangibles, consideraremos matrices en vez de endomorfismos. Toda la sección radica en los dos siguientes resultados y el algoritmo (Proposición 5.4.3) para hallar las bases a las que se refieren:

Teorema 5.4.1. *Dada $A \in \mathcal{M}_2(\mathbb{C})$ siempre existe una base de \mathbb{C}^2 en la que adquiere una de las siguientes formas:*

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \text{o bien} \quad \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}.$$

El primer caso es el diagonalizable que ya conocíamos, lo que aporta de nuevo el resultado es que las no diagonalizables lo son salvo por un 1 fuera de la diagonal.

Teorema 5.4.2. Dada $A \in \mathcal{M}_3(\mathbb{C})$ siempre existe una base de \mathbb{C}^3 en la que adquiere una de las siguientes formas:

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \quad \text{o bien} \quad \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} \quad \text{o bien} \quad \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}.$$

De nuevo el primer caso es el diagonalizable y la novedad es que, si no lo es, basta añadir uno o dos unos fuera de la diagonal en ciertas posiciones.

Tanto para $n = 2$ como para $n = 3$, estas matrices diagonales o casi diagonales se dice que son la *forma canónica de Jordan* de la matriz A .

Estos son casos particulares del *teorema de Jordan* que asegura que para cualquier endomorfismo $f : \mathbb{C}^n \rightarrow \mathbb{C}^n$ existe una base en la su matriz adquiere una *forma canónica de Jordan* del tipo

$$\begin{pmatrix} J_{n_1}(\lambda_1) & O & O & \dots & O \\ O & J_{n_2}(\lambda_2) & O & \dots & O \\ O & O & J_{n_3}(\lambda_3) & \dots & O \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ O & O & O & \dots & J_{n_k}(\lambda_k) \end{pmatrix} \quad \text{donde} \quad J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & \lambda & 1 \\ 0 & \dots & \dots & \dots & 0 & \lambda \end{pmatrix}.$$

Cada uno de los bloques $J_{n_i}(\lambda_i)$ se dice que es una *celda de Jordan*. La forma canónica de Jordan es única salvo reordenar estas celdas. Una notación tan abigarrada hace pensar que la prueba de este teorema es bastante pesada. Así es en la mayoría de los textos. Una excepción es la ingeniosa prueba breve por inducción de [6]. El hecho de que existan demostraciones breves del teorema de Jordan no se traduce en que el algoritmo para hallar la forma canónica de Jordan o la *base de Jordan* en la que se alcanzan sea sencillo. Nosotros aquí solo nos ocuparemos de los casos de dimensiones 2 y 3, y no veremos el algoritmo general ni la prueba del teorema de Jordan. Se deja al lector interesado que consulte en la bibliografía (por ejemplo [6], [13], [22], [14]).

La gracia de la forma canónica de Jordan es que se salvan parcialmente cosas empleadas en las aplicaciones⁷. Concretamente, hay fórmulas para resolver sistemas lineales de ecuaciones diferenciales si la matriz está en forma canónica de Jordan y fórmulas para potencias de estas matrices. Todo queda un poco más feo pero es totalmente explícito. A título de curiosidad, cuando elevamos a n las matrices de los casos no diagonales del Teorema 5.4.1 y del Teorema 5.4.1, se obtiene, respectivamente,

$$\begin{pmatrix} \lambda_1^n & n\lambda_1^{n-1} \\ 0 & \lambda_1^n \end{pmatrix}, \quad \begin{pmatrix} \lambda_1^n & n\lambda_1^{n-1} & 0 \\ 0 & \lambda_1^n & 0 \\ 0 & 0 & \lambda_2^n \end{pmatrix}, \quad \begin{pmatrix} \lambda_1^n & n\lambda_1^{n-1} & \frac{n(n-1)}{2}\lambda_1^{n-2} \\ 0 & \lambda_1^n & n\lambda_1^{n-1} \\ 0 & 0 & \lambda_1^n \end{pmatrix}.$$

⁷Mi opinión es que los matemáticos sobrevaloramos el interés de la forma canónica de Jordan porque en muchas aplicaciones estamos en el caso diagonalizable (por ejemplo, gracias al teorema espectral) y porque toda matriz se vuelve diagonalizable con una perturbación pequeña.

En las dimensiones 2 y 3, cubiertas por los teoremas anteriores, es fácil decidir a cuál de las formas canónicas corresponde una matriz: basta contar de cuántos vectores propios (linealmente independientes) disponemos.

Para dimensión 2, es de la segunda forma si y solo si no es diagonalizable (es decir, si no hay dos autovectores para formar una base).

Para dimensión 3, cuando no es diagonalizable hay dos posibilidades. La segunda forma se obtendrá cuando tenemos dos vectores propios linealmente independientes y la tercera cuando solo hay uno.

Analicemos los siguientes ejemplos, para convencernos de que es sencillo:

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

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 Teorema 5.4.1 con $\lambda_1 = -2$. Para A_2 , el autovalor de multiplicidad dos es $\lambda_1 = 1$ y como $(A_2 - I)\vec{v} = \vec{0}$ tiene un conjunto de soluciones de dimensión 1 y, por supuesto, también lo tiene $(A_2 - \lambda_2 I)\vec{v} = \vec{0}$ con $\lambda_2 = -1$, obtenemos la segunda forma del Teorema 5.4.2 con estos λ_j . Finalmente, A_3 solo tiene el autovalor $\lambda_1 = 2$ y $(A_3 - 2I)\vec{v} = \vec{0}$ solo tiene una solución y sus múltiplos, por tanto se tiene la tercera forma del Teorema 5.4.2.

El número de autovalores no determina la forma de Jordan porque ya hemos visto que no determina ni siquiera si es diagonalizable o no. Veamos más ejemplos:

$$A_4 = \begin{pmatrix} -4 & 9 & 3 \\ -4 & 8 & 2 \\ 0 & 0 & 2 \end{pmatrix}, \quad A_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -4 & -1 \end{pmatrix}, \quad A_6 = \begin{pmatrix} i & -1 \\ -1 & -i \end{pmatrix}.$$

La matriz A_4 solo tiene un autovalor, $\lambda_1 = 2$, pero la dimensión del espacio de soluciones de $(A_4 - 2I)\vec{v} = \vec{0}$ es dos, entonces disponemos de dos autovectores independientes y la forma canónica de Jordan debe ser la segunda del Teorema 5.4.2 con $\lambda_1 = \lambda_2 = 2$. 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 primera del Teorema 5.4.2 con $\lambda_1 = \lambda_2 = 1$ y $\lambda_3 = -1$. 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, porque no es la matriz nula. En conclusión tenemos como forma canónica de Jordan la segunda del Teorema 5.4.1 con $\lambda_1 = 0$.

Algo bien distinto es encontrar la base en la que se obtiene la forma canónica de Jordan, en el caso no diagonalizable. El siguiente resultado da un algoritmo para los casos de dimensión baja que nos ocupan.

Proposición 5.4.3. Sea $A \in M_n(\mathbb{C})$ no diagonalizable con $n = 2$ o $n = 3$. Con las siguientes elecciones, se obtiene una base de Jordan $\{\vec{v}_1, \dots, \vec{v}_n\}$ válida en cada uno de los casos.

1) Si $n = 2$, escogemos $\vec{v}_2 \neq \vec{0}$ que no sea vector propio y $\vec{v}_1 = (A - \lambda_1 I)\vec{v}_2$ donde λ_1 es el (único) autovalor de A .

2) Si $n = 3$ y la forma canónica es la segunda del Teorema 5.4.2, escogemos \vec{v}_2 que no sea vector propio tal que $(A - \lambda_1 I)^2 \vec{v}_2 = \vec{0}$ y $\vec{v}_1 = (A - \lambda_1 I)\vec{v}_2$, donde λ_1 es el autovalor de A con multiplicidad mayor que uno. Completamos con un \vec{v}_3 autovector de A que no sea múltiplo de \vec{v}_1 .

3) Si $n = 3$ y la forma canónica es la tercera del Teorema 5.4.2, escogemos \vec{v}_3 tal que $(A - \lambda_1 I)^2 \vec{v}_3 \neq \vec{0}$, $\vec{v}_2 = (A - \lambda_1 I)\vec{v}_3$ y $\vec{v}_1 = (A - \lambda_1 I)\vec{v}_2$, donde λ_1 es el autovalor de A .

Este resultado parece arbitrario y difícil de recordar. Quizá dé algo de luz que cuando tenemos una sola celda de Jordan $J_n(\lambda)$ siempre se cumple $(A - \lambda I)^n = O$ y se consigue una base $\{\vec{v}_1, \dots, \vec{v}_n\}$ tomando $(A - \lambda I)^{n-1} \vec{v}_n \neq \vec{0}$ y recursivamente $\vec{v}_{j-1} = (A - \lambda I)^{n-1} \vec{v}_j$. El resultado anterior es la aplicación de esta idea a cada una de las celdas de Jordan.

Calculemos una base de Jordan para algunos de los ejemplos anteriores. Se tiene que $\vec{v} = (1, 0)^t$ cumple $(A + 2I)\vec{v} \neq \vec{0}$ y una base posible es $\{(A + 2I)\vec{v}, \vec{v}\}$, esto es, $\{(2, -1)^t, (1, 0)^t\}$. Si uno quisiera comprobar el resultado habría que verificar:

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

En el caso de A_2 , hay una celda de Jordan de dimensión dos correspondiente a $\lambda_1 = 1$ y se tiene (la última condición asegura que el vector indicado no es autovector)

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

Por tanto obtenemos una base de Jordan $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ si tomamos $\vec{v}_2 = (1, 0, 2)^t$, $\vec{v}_1 = (A - I)\vec{v}_2 = (1, 1, 0)^t$ y \vec{v}_3 un vector propio correspondiente a $\lambda_2 = -1$, por ejemplo $\vec{v}_3 = (0, -1, 1)^t$. Si quisiéramos hacer ahora la comprobación, sería un poco larga y consistiría en verificar

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 2 & 1 \end{pmatrix}^{-1} A_2 \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

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 $\vec{v}_3 = (0, 1, 0)^t$ y completar la base con

$$\vec{v}_2 = (A_3 - 2I)\vec{v}_3 = (3, 0, -1)^t \quad \text{y} \quad \vec{v}_1 = (A_3 - 2I)\vec{v}_2 = (1, 0, 0)^t.$$

Como último ejemplo de base de Jordan, hallemos una para A_4 . Recordemos que era de la segunda forma del Teorema 5.4.2. No entraña más complicación que ahora $\lambda_1 = \lambda_2$, el único autovalor era 2. La matriz $(A_4 - 2I)^2$ es nula, en sintonía con lo dicho antes sobre las celdas de Jordan y entonces no hay otra restricción sobre \vec{v}_2 más que no sea autovector. Por ejemplo, tomemos $\vec{v}_2 = (1, 0, 0)^t$ que da lugar a $\vec{v}_1 = (A_4 - 2I)\vec{v}_2 = (-6, -4, 0)^t$. Como \vec{v}_3 hay que escoger un autovector que no sea múltiplo de \vec{v}_1 , por ejemplo $(1, 0, 2)^t$. En este caso se cumple

$$\begin{pmatrix} -6 & 1 & 1 \\ -4 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}^{-1} A_4 \begin{pmatrix} -6 & 1 & 1 \\ -4 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

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

Vamos con la prueba de los resultados. Es poco ilustrativa en el sentido de que difícilmente se entrevé una forma de generalizarla a dimensiones superiores para llegar al teorema de Jordan.

Demostración del Teorema 5.4.1. Si A tiene dos autovalores $\lambda_1 \neq \lambda_2$ entonces por el Corolario 5.1.5, se tiene la primera forma. Si solo hay un autovalor λ_1 , es que el polinomio característico es $(\lambda - \lambda_1)^2$ y la Proposición 5.2.6 asegura que existe $U \in \mathcal{M}_2(\mathbb{C})$ unitaria tal que $U^{-1}AU$ es triangular superior. Es decir, hay una base ortonormal $\{\vec{v}_1, \vec{v}_2\}$ en la que $A\vec{v}_1 = \lambda_1\vec{v}_1$ y $A\vec{v}_2 = \lambda_1\vec{v}_2 + b\vec{v}_1$ para algún $b \in \mathbb{C}$. Si $b = 0$ es de la primera forma y si $b \neq 0$, en la base $\{\vec{v}_1, b^{-1}\vec{v}_2\}$ se tiene la segunda forma porque $A(b^{-1}\vec{v}_2) = \lambda_1 b^{-1}\vec{v}_2 + \vec{v}_1$. \square

Demostración del Teorema 5.4.2. Si A tiene tres autovalores entonces es diagonalizable (de la primera forma) por el Corolario 5.1.5.

Consideremos ahora que hay dos autovalores λ_1 y λ_2 . Digamos que λ_1 es el de multiplicidad 2 y λ_2 el de multiplicidad uno. Sean \vec{v}_1, \vec{v}_3 autovectores respectivos y \vec{v}_2 linealmente independiente. Se cumple $A\vec{v}_2 = a\vec{v}_1 + b\vec{v}_2 + c\vec{v}_3$ para ciertos $a, b, c \in \mathbb{C}$ y, añadiendo a esto $A\vec{v}_1 = \lambda_1\vec{v}_1$, $A\vec{v}_3 = \lambda_2\vec{v}_3$, el polinomio característico de A en la base $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es $(\lambda_1 - \lambda)(b - \lambda)(\lambda_2 - \lambda)$. De aquí se deduce $b = \lambda_1$. Definiendo $\vec{w}_2 = \vec{v}_2 + c\vec{v}_3/(\lambda_1 - \lambda_2)$ se tiene $A\vec{w}_2 = a\vec{v}_1 + \lambda_1\vec{w}_2$. Si $a = 0$, hemos diagonalizado en la base $\{\vec{v}_1, \vec{w}_2, \vec{v}_3\}$ y si $a \neq 0$, como $A(a^{-1}\vec{w}_2) = \lambda_1 a^{-1}\vec{w}_2 + \vec{v}_1$, se obtiene la segunda forma canónica del teorema en la base $\{\vec{v}_1, a^{-1}\vec{w}_2, \vec{v}_3\}$.

Por último consideremos el caso en que hay un solo autovalor λ_1 , necesariamente de multiplicidad 3. La Proposición 5.2.6 implica que existe una base (de hecho ortonormal) $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ en la que

$$(5.7) \quad A\vec{v}_1 = \lambda_1\vec{v}_1, \quad A\vec{v}_2 = \lambda_1\vec{v}_2 + a\vec{v}_1, \quad A\vec{v}_3 = \lambda_1\vec{v}_3 + b\vec{v}_1 + c\vec{v}_2.$$

Si a, b y c fueran simultáneamente nulos, claramente A sería diagonalizable, por tanto descartamos ese caso.

Si $ac \neq 0$ entonces $\mathcal{B} = \{\vec{w}_1, \vec{w}_2, \vec{v}_3\}$ con $\vec{w}_1 = ac\vec{v}_1$, $\vec{w}_2 = b\vec{v}_1 + c\vec{v}_2$ es también una base (por ser linealmente independientes). Además se tiene $A\vec{w}_1 = \lambda_1\vec{w}_1$, $A\vec{w}_2 = \lambda_1\vec{w}_2 + \vec{w}_1$, $A\vec{v}_3 = \lambda_1\vec{v}_3 + \vec{w}_2$, lo que prueba que en \mathcal{B} se alcanza la tercera forma canónica del teorema.

Si $a = 0$, \vec{v}_1 y \vec{v}_2 y sus combinaciones lineales son vectores propios, en particular lo es $\vec{w}_1 = b\vec{v}_1 + c\vec{v}_2$, que cumple $A\vec{v}_3 = \lambda_1\vec{v}_3 + \vec{w}_1$. Si elegimos cualquier otro autovector \vec{w}_2 linealmente independiente con \vec{w}_1 , en la base $\{\vec{w}_1, \vec{v}_3, \vec{w}_2\}$ se alcanza la segunda forma.

Finalmente, si $c = 0$ y $a \neq 0$, usando (5.7), $\vec{w}_3 = b\vec{v}_2 - a\vec{v}_3$ es autovector y es linealmente independiente de $\vec{w}_1 = a\vec{v}_1$ y \vec{v}_2 . Como $A\vec{v}_2 = \lambda_1\vec{v}_2 + \vec{w}_1$, en $\{\vec{w}_1, \vec{v}_2, \vec{w}_3\}$ se alcanza de nuevo la segunda forma. \square

Demostración de la Proposición 5.4.3. Probemos cada apartado por separado.

1) Por definición $A\vec{v}_2 = \lambda_1\vec{v}_2 + \vec{v}_1$ y $\vec{v}_1 \neq \vec{0}$ (ya que \vec{v}_2 no es autovector) y solo falta comprobar que $A\vec{v}_1 = \lambda_1\vec{v}_1$ es autovector. Esto equivale a $(A - \lambda_1 I)\vec{v}_1 = \vec{0}$. Si C es la matriz de cambio de base a la forma canónica de Jordan $J = J_2(\lambda_1)$, se tiene $A = CJC^{-1}$ y se verifica

$$(A - \lambda_1 I)\vec{v}_1 = (A - \lambda_1 I)^2\vec{v}_2 = (C(J - \lambda_1 I)C^{-1})^2\vec{v}_2 = C(J - \lambda_1 I)^2C^{-1}\vec{v}_2.$$

La última igualdad se sigue escribiendo el cuadrado como un producto por sí mismo. Finalmente, un cálculo prueba $(J - \lambda_1 I)^2 = O$.

2) El vector \vec{v}_1 es no nulo porque \vec{v}_2 no es vector propio y, por construcción, $(A - \lambda_1 I)\vec{v}_1 = \vec{0}$, de donde \vec{v}_1 es autovector. Con ello, $A\vec{v}_1 = \lambda_1\vec{v}_1$, $A\vec{v}_2 = \lambda_1\vec{v}_2 + \vec{v}_1$ y $A\vec{v}_3 = \lambda_2\vec{v}_3$ (no se excluye $\lambda_2 = \lambda_1$). Sabemos que \vec{v}_1 y \vec{v}_3 son linealmente independientes, por tanto, para verificar que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es base solo resta comprobar que $\vec{v}_2 = \mu_1\vec{v}_1 + \mu_3\vec{v}_3$ es imposible. Si $\lambda_2 = \lambda_1$ es obvio porque la parte derecha es autovector y \vec{v}_2 no lo es. Si $\lambda_2 \neq \lambda_1$, aplicando a la igualdad $(A - \lambda_2 I)(A - \lambda_1 I)$, se obtiene en el lado derecho $\vec{0}$, mientras que aplicado a \vec{v}_2 resulta $(A - \lambda_2 I)\vec{v}_1 = (\lambda_1 - \lambda_2)\vec{v}_1$ que es no nulo.

3) Este caso es similar al primero, salvo que $(A - \lambda_1 I)\vec{v}_1 = \vec{0}$ se deduce ahora de que $J = J_3(\lambda_1)$ verifica $(J - \lambda_1 I)^3 = O$. Para comprobar que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ es base tenemos que asegurar que son linealmente independientes. Si $\mu_1\vec{v}_1 + \mu_2\vec{v}_2 + \mu_3\vec{v}_3 = \vec{0}$, aplicando $(A - \lambda_1 I)^2$ se obtiene $\mu_3 = 0$, y \vec{v}_1 y \vec{v}_2 son linealmente independientes porque uno es autovector y el otro no.

Una observación tranquilizadora final, es que siempre existen unos \vec{v}_j con las propiedades requeridas en el enunciado porque los vectores de la base de Jordan las verifican. \square

Exprimiendo el silicio [opcional]. Ya hemos visto que `matlab/octave` confunde matrices no diagonalizables con otras que no lo son, por tanto sin herramientas adicionales no permite hallar la forma canónica de Jordan⁸.

⁸Con el paquete de cálculo simbólico para `matlab` dispondrás de un comando `sym` que transforma una matriz en otra simbólica a la que aplicar `jordan` con la misma sintaxis que `eig`.

Por otro lado, en `sagemath` existe el comando `jordan_form` que produce la forma canónica de Jordan. Si se incluye como argumento `transformation=True` entonces también da la matriz de cambio de base. Por ejemplo, para calcular la matriz A_2 de los ejemplos anteriores usaríamos:

```

1 # Matriz
2 A = matrix(3,3,[0,1,1, 3,-2,-1, -4,4,3])
3 [J,C] = A.jordan_form(transformation=True)
4
5 print('Forma canónica')
6 print(J)
7 print('Matriz de cambio de base')
8 print(C)

```

No hay que perder de vista que la forma canónica de Jordan solo es única salvo la ordenación de las celdas, igual que la diagonalización depende del orden de los autovalores, y la base de Jordan no es única, incluso una vez fijada dicha ordenación. El código anterior da

$$J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad C = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ -1 & 0 & 2 \end{pmatrix}.$$

En nuestra solución pusimos la celda más pequeña al final, en sintonía con el Teorema 5.4.1. Moviendo la primera columna de C al final, la C resultante difiere todavía de la nuestra en el signo de una de las columnas.