

## 5.2. El teorema espectral

Los cambios de base utilizados para diagonalizar no tienen ninguna propiedad 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*<sup>3</sup>, 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  $U^\dagger U = I$  que caracteriza las matrices unitarias. Trasponiendo y conjugando  $A^\dagger = U^\dagger U^\dagger A$ . De aquí,

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

El hecho de que  $D$  sea diagonal implica  $DD^\dagger = D^\dagger D$ , 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 sección<sup>4</sup> 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

<sup>3</sup>Por 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.

<sup>4</sup>Con 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.

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

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

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

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

donde  $U_2$  es unitaria, porque  $V_1$  lo es.

Repitiendo 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.  $\square$

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  $\lambda$ , 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 elementos 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  $(a + d)^2 + 4b^2$  que claramente es no negativo y por tanto nunca lleva a soluciones complejas. Dar una prueba similar en el caso  $3 \times 3$  sería muchísimo más difícil.

Las consecuencias anunciadas más que dos son una y su especialización, porque el caso 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  $DD^\dagger = V^{-1}U^\dagger UV = V^{-1}V = I$ , 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  $U$  igual al producto  $U_1U_2 \cdots U_n$ .

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  $\lambda$  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, a las que se aplica el Corolario 5.2.4, 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  $\lambda_1$  los vectores propios son múltiplos de  $(2, -1, -2)^t$  mientras que para  $\lambda_2$  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

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

El primer vector está fijado, salvo el signo, por la normalización, mientras que para los dos últimos vectores hay infinidad de bases ortonormales que generan el mismo subespacio.

**Exprimiendo 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  $\epsilon$  a cada elemento, entonces cada autovalor  $\lambda$  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
$\lambda_1$	-0.15537117	-0.15530942	-0.15530780
$\lambda_2$	0.06759992	0.06869853	0.06872612
$\lambda_3$	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 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  $K^n \rightarrow K^n$ , la pregunta natural es determinar la sucesión de vectores  $\vec{x}_1, \vec{x}_2, \dots$  definidos mediante

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

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 decidir la eficiencia del algoritmo. Deseamos una fórmula más explícita, a través de un cálculo sencillo de  $A^k$ , 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,  $A^k = CD^kC^{-1}$ . En particular, la solución de (5.5) 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.

*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

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