

Diagonalización

Ejemplos introductorios. Sea $f : M_2(\mathbb{R}) \rightarrow M_2(\mathbb{R})$ la aplicación lineal definida por

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto f(M) = \begin{pmatrix} -15a - 20b + 9c + 12d & 10a + 15b - 6c - 9d \\ -18a - 24b + 12c + 16d & 12a + 18b - 8c - 12d \end{pmatrix}.$$

Esta aplicación parece complicada. Por ejemplo, trabajando en la base natural N de $M_2(\mathbb{R})$ dada por

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

la matriz de la aplicación es el siguiente monstruo:

$$A = M_N^N(f) = \begin{pmatrix} -15 & -20 & 9 & 12 \\ 10 & 15 & -6 & -9 \\ -18 & -24 & 12 & 16 \\ 12 & 18 & -8 & -12 \end{pmatrix}.$$

Sin embargo, la aplicación f no es tan complicada como parece cuando nos sacamos de la manga la base U de $M_2(\mathbb{R})$ dada por

$$u_1 = \begin{pmatrix} 2 & -1 \\ 4 & -2 \end{pmatrix} \quad u_2 = \begin{pmatrix} -1 & 1 \\ -2 & 2 \end{pmatrix} \quad u_3 = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \quad u_4 = \begin{pmatrix} 2 & -1 \\ 2 & -1 \end{pmatrix}.$$

Resulta sencillo comprobar que la imagen de cada uno de estos elementos es un múltiplo del propio elemento. Concretamente, si $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = 2$ y $\lambda_4 = -2$, entonces

$$f(u_1) = \lambda_1 u_1 = u_1 \quad f(u_2) = \lambda_2 u_2 = -u_2 \quad f(u_3) = \lambda_3 u_3 = 2u_3 \quad f(u_4) = \lambda_4 u_4 = -2u_4.$$

Por tanto, la matriz del endomorfismo f en la base U es diagonal:

$$D = M_U^U(f) = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

Ejercicio. Ver que $SD = AS$, si S es la matriz del cambio de base que pasa de base U a base N :

$$S = C_N^U = \begin{pmatrix} 2 & -1 & -1 & 2 \\ -1 & 1 & 1 & -1 \\ 4 & -2 & -1 & 2 \\ -2 & 2 & 1 & -1 \end{pmatrix}.$$

(Indicación: No es necesario matarse haciendo el cálculo, basta recordar que $M_U^U(f) = C_U^N M_N^N(f) C_N^U$.)

Así pues, hemos *diagonalizado* f . De paso, hemos comprobado que f es *diagonalizable*, es decir, que se puede diagonalizar. Obviamente, esto plantea las siguientes preguntas:

- ¿Todos los endomorfismos son diagonalizables?
- En caso negativo, ¿cómo se sabe si un endomorfismo dado es (o no) diagonalizable?
- Finalmente, cuando ya sabemos que un endomorfismo es diagonalizable, ¿cómo se diagonaliza?

La primera respuesta es no, ya que el endomorfismo $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ cuya matriz en la base natural es

$$J = M_N^N(f) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

no es diagonalizable. Si lo fuera, existirían una matriz diagonal D y una matriz invertible S ,

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad S = C_N^U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

tales que $\begin{pmatrix} \lambda_1 \alpha & \lambda_2 \beta \\ \lambda_1 \gamma & \lambda_2 \delta \end{pmatrix} = SD = JS = \begin{pmatrix} 0 & 0 \\ \alpha & \beta \end{pmatrix}.$

Ejercicio. Comprobar que eso es imposible.

Diagonalización de matrices versus diagonalización de endomorfismos. Sea $f : E \rightarrow E$ un endomorfismo de un \mathbb{K} -ev de dimensión n . Sea $A \in M_n(\mathbb{K})$ la matriz de f en alguna base de E . El problema de diagonalizar el endomorfismo f es similar al problema de diagonalizar la matriz A , ya que las siguientes afirmaciones son equivalentes:

- Existe una base $U = (u_1, \dots, u_n)$ de E y unos escalares $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tales que

$$f(u_j) = \lambda_j u_j \quad j = 1, \dots, n.$$

- Existe una base U de E tal que la matriz $M_U^U(f)$ es diagonal.
- Existen en $M_n(\mathbb{K})$ una matriz diagonal D y una matriz invertible S tales que $SD = AS$.

Cuando estas condiciones se verifican diremos que f (y A) son *diagonalizables*.

Sólo hablaremos de matrices pues son más sencillas que los endomorfismos. Cuando nos pidan trabajar con un endomorfismo realizaremos los siguientes pasos:

1. Calcular la matriz del endomorfismo en alguna base adecuada.
2. Estudiar la diagonalización de esa matriz.
3. Trasladar los resultados al contexto inicial.

El polinomio característico. VAPs y VEPs. Si $A \in M_n(\mathbb{K})$, su *polinomio característico* es

$$Q_A(t) = \det(A - t\text{Id}) \in \mathbb{K}_n[t].$$

Las raíces de $Q_A(t)$ son los *valores propios (VAPs)* de la matriz A . El conjunto de todos los VAPs es el *espectro* de la matriz y se escribe $\sigma(A)$. Si λ es un VAP de A , el sev $E_\lambda = \text{Nuc}(A - \lambda\text{Id}) \subset \mathbb{K}^n$ es el *sev propio* asociado al VAP λ . Los *vectores propios (VEPs)* de un VAP $\lambda \in \sigma(A)$ son los vectores no nulos de E_λ . La *traza* de una matriz es la suma de los elementos que están situados en la diagonal de la matriz y se escribe *traza* A .

Ejercicio. Calcular el polinomio característico, los VAPs y los VEPs de las matrices

$$A = \begin{pmatrix} 3 & 1 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Las principales propiedades del polinomio característico, los VAPs y los VEPs son las siguientes.

- A es invertible si y sólo si $0 \notin \sigma(A)$.
- $\lambda \in \sigma(A) \Leftrightarrow \det(A - \lambda\text{Id}) = 0 \Leftrightarrow \text{rango}(A - \lambda\text{Id}) < n \Leftrightarrow \text{Nuc}(A - \lambda\text{Id}) \neq 0 \Leftrightarrow \lambda$ tiene VEPs.
- Un vector $v \in \mathbb{K}^n$ es un VEP de VAP λ de la matriz A si y sólo si $Av = \lambda v$ y $v \neq 0$.
- $\text{gr}[Q_A(t)] = n$, luego A no puede tener más de n VAPs diferentes.
- Si $Q_A(t) = q_0 - q_1 t + \dots + (-1)^{n-1} q_{n-1} t^{n-1} + (-1)^n q_n t^n$, entonces:

$$q_0 = \det A \quad q_{n-1} = \text{traza } A \quad q_n = 1.$$

- Si $Q_A(t)$ descompone totalmente en \mathbb{K} , es decir, si $Q_A(t) = \prod_{j=1}^l (\lambda_j - t)^{\alpha_j}$ con $\lambda_1, \dots, \lambda_l \in \mathbb{K}$ y $\alpha_1, \dots, \alpha_l \in \mathbb{N}$, entonces $\sigma(A) = \{\lambda_1, \dots, \lambda_l\}$ y
 - $\det A = \prod_{j=1}^l \lambda_j^{\alpha_j}$ = producto de todos los VAPs repetidos según multiplicidad.
 - $\text{traza } A = \sum_{j=1}^l \alpha_j \lambda_j$ = suma de todos los VAPs repetidos según multiplicidad.
 Repetido según multiplicidad significa que un VAP doble aparece 2 veces, uno triple 3 veces, etc.
- Varios de estos objetos son invariantes por cambios de base. Si $B = S^{-1}AS$, entonces

$$Q_B(t) = Q_A(t) \quad \sigma(B) = \sigma(A) \quad \det B = \det A \quad \text{traza } B = \text{traza } A.$$

- Los VAPs de una matriz diagonal (o triangular) son los elementos de la diagonal de la matriz.
- *VEPs de VAPs diferentes siempre son li:* Si $\lambda_1, \dots, \lambda_l$ son VAPs distintos de A , entonces la suma $E_{\lambda_1} \oplus \dots \oplus E_{\lambda_l}$ es directa.
- *Tª de Cayley-Hamilton:* $Q_A(t) = q_0 + q_1 t + \dots + q_n t^n \Rightarrow Q_A(A) := q_0 \text{Id} + q_1 A + \dots + q_n A^n = \mathbf{0}$.

Ejercicio. Como $Q_A(t) = \det(A - t\text{Id})$, vemos que $Q_A(A) = \det(A - A \cdot \text{Id}) = \det(A - A) = \det(\mathbf{0}) = 0$. ¿Por qué esta demostración del Teorema de Cayley-Hamilton es incorrecta?

Ejercicio. Encontrad ejemplos que pongan de manifiesto las siguientes afirmaciones.

1. $\text{traza}(AB) \neq (\text{traza } A) \cdot (\text{traza } B)$.
2. Existen matrices reales sin ningún VAP real.

Problemas relacionados. 4 y 17.

El criterio de diagonalización. Buscamos condiciones necesarias y suficientes para que una matriz cuadrada $A \in M_n(\mathbb{K})$ diagonalice sobre el cuerpo \mathbb{K} . En la mayoría de los casos $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$.

Si $\lambda \in \sigma(A)$, sus *multiplicidades algebraica y geométrica* se definen como:

- $\text{ma}(\lambda) =$ multiplicidad de λ como raíz del polinomio $Q_A(t)$.
- $\text{mg}(\lambda) = \dim E_\lambda = \dim[\text{Nuc}(A - \lambda\text{Id})] = n - \text{rango}(A - \lambda\text{Id})$.

Es decir, $\text{ma}(\lambda)$ es el exponente del factor $(t - \lambda)^\alpha$ en la factorización del polinomio característico, mientras que $\text{mg}(\lambda)$ el número de VEPs li de VAP λ . Estas multiplicidades satisfacen la desigualdad

$$1 \leq \text{mg}(\lambda) \leq \text{ma}(\lambda) \quad \forall \lambda \in \sigma(A).$$

Finalmente, el *criterio de diagonalización* establece que A es diagonalizable sobre \mathbb{K} si y sólo si:

1. Su polinomio característico descompone totalmente en \mathbb{K} : $\lambda \in \mathbb{K}$ para todo $\lambda \in \sigma(A)$.
2. Las multiplicidades algebraicas y geométricas coinciden: $\text{mg}(\lambda) = \text{ma}(\lambda)$ para todo $\lambda \in \sigma(A)$.

Problemas relacionados. 1, 2 y 3.

El algoritmo de diagonalización. El algoritmo estándar para estudiar la diagonalización de una matriz $A \in M_n(\mathbb{K})$ y, cuando sea posible, encontrar una matriz diagonal $D \in M_n(\mathbb{K})$ y una matriz invertible $S \in M_n(\mathbb{K})$ tales que $SD = AS$, consta de los siguientes pasos.

1. Calcular y factorizar el polinomio característico $Q_A(t)$.
2. Si existe $\lambda \in \sigma(A)$ tal que $\lambda \notin \mathbb{K}$, entonces A no diagonaliza en \mathbb{K} .
3. Comparar las multiplicidades algebraicas y geométricas de los VAPs. Si existe $\lambda \in \sigma(A)$ tal que $\text{mg}(\lambda) < \text{ma}(\lambda)$, entonces A no diagonaliza (en ningún cuerpo).
4. (Cuando $A = (a_{ij})$ diagonaliza y tiene VAPs $\lambda_1, \dots, \lambda_l \in \mathbb{K}$ de multiplicidades $\alpha_1, \dots, \alpha_l \in \mathbb{N}$.) Para cada VAP $\lambda = \lambda_j$, tenemos que encontrar $\alpha = \alpha_j$ VEPs li de VAP λ resolviendo el sistema homogéneo con α grados de libertad:

$$\begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

5. D se construye poniendo los VAPs en la diagonal, repetidos según multiplicidad.
6. S se construye poniendo los VEPs por columnas, en el mismo orden que los VAPs en D .
7. Comprobar que $SD = AS$. (Opcional.)

Problemas relacionados. 5, 6, 7, 8 y 11.

Trucos. Los siguientes trucos pueden ayudarnos en algunas ocasiones.

- *Matando ecuaciones al calcular VEPs.* Cuando buscamos α VEPs li de VAP λ resolviendo el sistema homogéneo anterior, conviene recordar el grado de libertad de ese sistema es α . Por tanto, al escalar el sistema para resolverlo, tienen que desaparecer exactamente α ecuaciones. Es decir, una ecuación si el VAP es simple, dos si es doble, tres si es triple, etc.
- *Los VAPs simples son inofensivos.* Si $\text{ma}(\lambda) = 1$, entonces $\text{mg}(\lambda) = \text{ma}(\lambda)$. En particular, las matrices sin VAPs múltiples siempre diagonalizan.
- *Diagonalizando en $\mathbb{K} = \mathbb{C}$.* Si estamos estudiando una matriz real A , debemos tener en cuenta las siguientes propiedades:
 - El conjugado de un VAP de A , también es un VAP de A : $\lambda \in \sigma(A) \iff \bar{\lambda} \in \sigma(A)$.
 - El conjugado de un VEP de A , es un VEP del VAP conjugado: $v \in E_\lambda \iff \bar{v} \in E_{\bar{\lambda}}$.
 - Las multiplicidades no cambian al conjugar: $\text{ma}(\lambda) = \text{ma}(\bar{\lambda})$ y $\text{mg}(\lambda) = \text{mg}(\bar{\lambda})$.

Estas propiedades son útiles, ya que permiten reducir el trabajo a la mitad. Al estudiar un VAP, automáticamente tenemos la información de su conjugado.

Ejercicio. Diagonalizar en los complejos la matriz $A = \begin{pmatrix} 1 & -12 & -14 \\ 1 & 2 & -3 \\ 1 & 1 & -2 \end{pmatrix}$. ($\sigma(A) = \{1, \pm 5i\}$)

y una base de VEPs es $v_1 = (25, -7, 6)^\top$, $v_2 = (26, 1 + 5i, 1 + 5i)^\top$ y $v_3 = \overline{v_2}$.)

- *Cazando VAPs.* La traza y del determinante pueden servir para cazar algunos VAPs. Ejemplo: si $A \in M_3(\mathbb{R})$ es tal que $3 \in \sigma(A) = \{\lambda_1, \lambda_2, \lambda_3\}$, $\det A = 9$ y traza $A = 7$, entonces

$$\lambda_2 + \lambda_3 = (\text{traza } A) - \lambda_1 = 4 \quad \lambda_2 \lambda_3 = (\det A) / \lambda_1 = 3$$

luego $\lambda_1 = \lambda_2 = 3$ y $\lambda_3 = 1$.

Ejercicio. Sea $A \in M_4(\mathbb{R})$ tal que $3 + 2i \in \sigma(A)$ y $\det A = \text{traza } A = 0$. Encontrar los VAPs.

- *Cazando VEPs.* Para calcular los VEPs de un VAP λ , tenemos que calcular el núcleo de la matriz $A - \lambda \text{Id}$. El núcleo se puede encontrar escalonando o a vista si somos capaces de ver las cl de columnas de la matriz $A - \lambda \text{Id}$ que se anulan. Por ejemplo, el núcleo de la matriz

$$(c_1 | c_2 | c_3 | c_4) = \begin{pmatrix} 1 & 3 & 7 & 0 \\ 2 & 6 & 14 & 7 \\ -1 & -3 & -7 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

contiene a los vectores $u = (3, -1, 0, 0)^\top$ y $v = (7, 0, -1, 0)^\top$, ya que la columnas cumplen las relaciones $c_2 = 3c_1$ y $c_3 = 7c_1$.

- *Inversas por Cayley-Hamilton.* Si la matriz A es invertible y $Q_A(t) = q_0 + q_1 t + q_2 t^2 + \dots + q_n t^n$, $Q_A(A) = \mathbf{0} \implies q_0 \text{Id} + q_1 A + q_2 A^2 + \dots + q_n A^n = \mathbf{0}$
 $\implies -q_0 \text{Id} = q_1 A + q_2 A^2 + \dots + q_n A^n = A (q_1 \text{Id} + q_2 A + \dots + q_n A^{n-1})$
 $\implies A^{-1} = -(q_1 \text{Id} + q_2 A + \dots + q_n A^{n-1}) / q_0$.

Ejercicio. ¿Dónde falla este método para calcular inversas cuando la matriz A no es invertible?

Problema relacionado. 10, apartado a).

- *Potencias de matrices.* Si $D = \text{diag}(\lambda_1, \dots, \lambda_n) = S^{-1} A S$, entonces

$$A^k = S D^k S^{-1} = S \cdot \text{diag}(\lambda_1^k, \dots, \lambda_n^k) \cdot S^{-1} \quad \forall k \in \mathbb{N}.$$

Cuando la matriz A es invertible, D también lo es pues todos los VAPs son diferentes de cero y la fórmula anterior se puede usar para toda $k \in \mathbb{Z}$.

Problema relacionado. 10, apartado c).

- *Ecuaciones con matrices.* La diagonalización es útil, entre otras cosas, para resolver ecuaciones con matrices que de otra forma serían tremendamente complicadas. Por ejemplo, supongamos que queremos calcular las raíces k -ésimas de una matriz diagonalizable $A \in M_n(\mathbb{C})$. Es decir, buscamos todas las matrices $B \in M_n(\mathbb{C})$ tales que $B^k = A$. Aplicando el algoritmo de diagonalización anterior calculamos una matriz diagonal $D \in M_n(\mathbb{C})$ y una matriz invertible $S \in M_n(\mathbb{C})$ tales que $D = \text{diag}(\lambda_1, \dots, \lambda_n) = S^{-1} A S$. Entonces,

$$B = \sqrt[k]{A} = S \sqrt[k]{D} S^{-1} = S \cdot \text{diag}(\sqrt[k]{\lambda_1}, \dots, \sqrt[k]{\lambda_n}) \cdot S^{-1}.$$

Así pues, una matriz diagonalizable con p VAPs no nulos (contados con multiplicidad) tiene al menos k^p raíces k -ésimas en los complejos.

Este método sirve para calcular cosas más complicadas, como el coseno o el logaritmo de una matriz, pero eso es otra historia.

Problema relacionado. 13.

Problemas para no dormir. 9, 12, 14, 15, 16, 17 y 18.