Diagonalización

Ingeniería informática Curso: Álgebra

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

Comentarios

La idea subyacente es que con un cambio de variable adecuado, es posible descomponer en multitud de casos un endomorfismo del espacio euclídeo en trozos independientes que operan en dimensión uno.

1. Autovalores y autovectores

En nuestro contexto limitado, los endomorfismos son aplicaciones lineales $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ definidas por $f(\vec{x}) = A\vec{x}$ donde $A \in \mathcal{M}_n$ es una matriz cuadrada que consta, como sabemos, de n^2 números. En algunas aplicaciones se produce una simplificación considerable cuando A es diagonal y, por tanto, está determinada por solo n números. En cierto modo, un endomorfismo con matriz diagonal en n dimensiones se desacopla en n endomorfismos de dimensión 1 tan tontos como multiplicar por una constante. Nos gustaría emular esta situación encontrando vectores que no cambien de dirección bajo la acción de f.

Con esta idea en mente, se dice que $\vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}$ es un *autovector* o *vector propio* de $A \in \mathcal{M}_n$ si $A\vec{v} = \lambda\vec{v}$ para cierto $\lambda \in \mathbb{R}$ al que se llama *autovalor* o *valor propio*.

Comentario. En realidad, sería más propio hablar de autovalores y autovectores de un endomorfismo en vez de asociarlos a una matriz, pero esto último es lugar común. Con nuestra visión restringida de las aplicaciones lineales hay una concordancia exacta entre endomorfismos y matrices cuadradas, mientras que en cursos más avanzados a cada endomorfimo se le asocia una clase de matrices bajo cierta relación de equivalencia en $\mathcal{M}_n$. Por otro lado, sería más natural trabajar con los números complejos, que no hemos repasado en el curso, permitiendo $\vec{v} \in \mathbb{C}^n \setminus \{\vec{0}\}$ y $\lambda \in \mathbb{C}$.

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

Ejemplo (Teórico). Probar que el polinomio característico de $A \in \mathcal{M}_n$ es exactamente de grado n y su *coeficiente principal* es $(-1)^n$.

El caso $n = 1$ es trivial. A partir de ahí, procedemos por inducción. Suponemos que es cierto para $n-1$ y queremos probarlo para $A \in \mathcal{M}_n$. La definición de determinante implica $|A-\lambda I| = (a_{11}-\lambda)C_{11}+\sum_{i=2}^n a_{i1}C_{i1}$ donde C_{i1} es el cofactor correspondiente para $C = A-\lambda I$. En cada C_{i1} aparece $n-1$ veces λ, por tanto, C_{i1} es un polinomio en λ de grado a lo más $n-1$. Además, por la hipótesis de inducción, C_{11} es $(-1)^{n-1}\lambda^{n-1}$ más términos de grado menor. Sustituyendo en la fórmula anterior para $|A-\lambda I|$ se obtiene el resultado deseado.

Ejemplo. Hallar los autovalores y autovectores de

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

Primero calculamos el polinomio característico:

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

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

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

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

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

Con ello hemos hallado todos los autovectores.

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

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

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

No tiene autovalores (ni, por tanto, autovectores) en $\mathbb{R}$ porque la ecuación característica es $\lambda^2 + 1 = 0$. En este curso no tratamos los números complejos, pero en un contexto más general es conveniente hacerlo y considerar que i y $-i$ son valores propios válidos.

Cuanta mayor sea la dimensión, mayor puede ser el número de autovalores. Así, en dimensión tres puede haber hasta tres autovalores.

Ejemplo. Hallar los valores propios de

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

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

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

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

Las raíces múltiples del polinomio característico inducen algunas complicaciones que serán fundamentales después.

Ejemplo. Calcular los autovalores y los autovectores de

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

Tras algunos cálculos, obtenemos que el polinomio característico es

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

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

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

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

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

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

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

2. Matrices diagonalizables

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

Comentario. La expresión *"diagonalizable sobre $\mathbb{R}$"* es equivalente en este curso a *"diagonalizable"* porque no vamos a estudiar la diagonalización sobre $\mathbb{C}$, que es la alternativa principal.

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

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

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

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

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

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

La relación anterior es útil para calcular potencias de matrices porque

$$A = CDC^{-1} \implies A^k = CD^kC^{-1} \quad \text{para } k \in \mathbb{N}.$$

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 se reduce a operar multiplicando cada vez por A. Por ejemplo, $A^2 = CDC^{-1}CDC^{-1} = CD^2C^{-1}$, $A^3 = A^2A = CD^2C^{-1}CDC^{-1} = CD^3C^{-1}$ y así sucesivamente.

Ejemplo. Estudiar si las siguientes matrices, consideradas antes, son diagonalizables:

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

Se tiene que A_1 es diagonalizable sobre $\mathbb{R}$, porque hay dos autovectores $\vec{v}_1 = (2,-1)^t$, $\vec{v}_2 = (1,1)^t$. De hecho si solo nos preguntasen si es diagonalizable, con ver que hay dos autovalores no harían falta más cuentas, pues para cada uno hay al menos un autovector. Sabemos que A_2 no tiene autovalores reales y, por tanto, tampoco autovectores reales, así que no es diagonalizable sobre $\mathbb{R}$. Los cálculos sí serían necesarios para ver que A_3 es diagonalizable porque solo hay dos autovalores. Ya vimos que podíamos extraer dos autovectores para λ_1 y uno más para λ_3. Según nuestros cálculos, una base válida de vectores propios es:

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

que nos da las columnas de C y lleva a la identidad

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

Veamos ahora ejemplos no diagonalizables "puros" en el sentido de que la dificultad no está en que aparezcan números complejos. La clave está en los autovalores múltiples. Un resultado asegura que *un factor* $(\lambda-\lambda_j)^{n_j}$ *en el polinomio característico implica que el autoespacio de* λ_j *tiene a lo más dimensión* n_j (a lo más hay n_j autovectores con el abuso de notación), pero nada asegura que sea exactamente n_j.

Ejemplo. Comprobar que

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

no es diagonalizable.

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

Ejemplo. Decidir si la matriz

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

es diagonalizable.

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

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

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

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

> *Una matriz* $A \in \mathcal{M}_n(\mathbb{R})$ *es diagonalizable sobre* $\mathbb{R}$ *si y solo si* $|A-\lambda I|$ *tiene todas sus raíces* λ_j *reales y* $n - \operatorname{rg}(A-\lambda_j I)$ *coincide con la multiplicidad de* λ_j.

Así pues, decidir si una matriz es diagonalizable es una simple rutina (dando por hecho que las raíces son fáciles de calcular) aplicando eliminación de Gauss para hallar el rango. Además, para las raíces de multiplicidad uno, por lo dicho antes de que la multiplicidad acota el número de autovectores linealmente independientes, no hay nada que comprobar.

Ejemplo. Estudiar si es diagonalizable la matriz

$$A = \begin{pmatrix} 7 & -4 & 4 \\ -4 & 1 & 8 \\ 4 & 8 & 1 \end{pmatrix}.$$

La mayor parte del esfuerzo está en calcular el polinomio característico y sus raíces. Aunque no es imprescindible, utilizaremos propiedades de los determinantes para simplificar un poco. Comenzamos sumando a la tercera fila de $|A - \lambda I|$ la segunda:

$$\underset{f_3 \mapsto f_3+f_2}{=} \begin{vmatrix} 7-\lambda & -4 & 4 \\ -4 & 1-\lambda & 8 \\ 0 & 9-\lambda & 9-\lambda \end{vmatrix} = (9-\lambda) \begin{vmatrix} 7-\lambda & -4 & 4 \\ -4 & 1-\lambda & 8 \\ 0 & 1 & 1 \end{vmatrix} \underset{\substack{f_2 \mapsto f_2 - 8f_3 \\ f_1 \mapsto f_1 - 4f_3}}{=} (9-\lambda) \begin{vmatrix} 7-\lambda & -8 & 0 \\ -4 & -7-\lambda & 0 \\ 0 & 1 & 1 \end{vmatrix}$$

que desarrollando por la última columna da $|A - \lambda I| = (9 - \lambda)(\lambda^2 - 81) = -(\lambda - 9)^2(\lambda + 9)$. De donde los autovalores son 9 (doble) y -9. Todo lo que hay que hacer es comprobar que $3 - \mathrm{rg}(A - 9I) = 2$, que hay dos autovectores linealmente independientes para 9. Esto es, $\mathrm{rg}(A - 9I) = 1$. Un paso de eliminación de Gauss da

$$A = \begin{pmatrix} -2 & -4 & 4 \\ -4 & -8 & 8 \\ 4 & 8 & -8 \end{pmatrix} \underset{\substack{f_2 \mapsto f_2 - 2f_1 \\ f_3 \mapsto f_3 + 2f_1}}{\longrightarrow} \begin{pmatrix} -2 & -4 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

lo que confirma que el rango es uno porque hay un solo escalón.

Hay un resultado teórico de gran importancia en matemáticas y física, el *teorema espectral*, que asegura que cualquier $A \in \mathcal{M}_n(\mathbb{R})$ simétrica, es decir, $A = A^t$, es siempre diagonalizable sobre $\mathbb{R}$ y hay una base de autovectores perpendiculares. Conociéndolo, el ejemplo anterior no habría necesitado de ningún cálculo. Este resultado es muy sorprendente, pues parece que la simetría, el hecho de que todas las raíces sean reales y la condición geométrica de perpendicularidad (que hay que definir bien en n dimensiones) no tienen absolutamente nada que ver. La prueba es ingeniosa, pero no particularmente difícil. En cualquier caso, excede los contenidos del curso.