Tema 3

Diagonalización de Matrices

Introducción

En los dos temas previos de este bloque hemos visto cómo problemas de la realidad son escritos, tratados y resueltos a través de espacios vectoriales y sistemas de ecuaciones lineales. En realidad, son muchas las ocasiones en que un problema complejo, en el que intervienen varias variables, se acaba simplificando en su planteamiento para hacer factible su resolución, aunque ésta sea aproximada. El caso más simple consiste en linealizar el problema (los sistemas de ecuaciones lineales diferenciales y en diferencias se verán en el próximo tema), y para ello, igual que para la resolución más sencilla de un sistema de ecuaciones lineales, el elemento principal de trabajo es la matriz: el manejo de operaciones reiteradas sobre una matriz requiere simplicidad de la misma.

La **representación matricial de un problema lineal** no es más que la definición de una función lineal de varias variables[1] a través de algunos de sus elementos (una base) y sus valores correspondientes, ya que por la linealidad se podrá conocer en (extender a) todo el espacio vectorial. Veámoslo con un ejemplo.

Ejemplo 1. *Tenemos tres pozos A, B y C que reciben agua de tres manantiales M_1, M_2 y M_3. Cada litro de agua del manantial M_1 se reparte entre los tres pozos en cantidades iguales; por cada litro procedente del manantial M_2 caen $2/3$ de litro en el pozo A, y la misma cantidad en los pozos B y C : $1/6$ l. Finalmente, el manantial M_3 sólo vierte agua en el pozo C. En el siguiente diagrama queda recogida la información:*

$$\begin{array}{lccc} & \textit{Pozo } A & \textit{Pozo } B & \textit{Pozo } C \\ 1l.\, M_1 \rightarrow & 1/3 & 1/3 & 1/3, \\ 1l.\, M_2 \rightarrow & 2/3 & 1/6 & 1/6, \\ 1l.\, M_3 \rightarrow & 0 & 0 & 1. \end{array}$$

Podemos "matematizar" la situación diciendo que los vectores $(1, 0, 0)$, $(0, 1, 0)$ *y* $(0, 0, 1)$

[1]Una aplicación $f : V \to W$ es lineal si $f(x + y) = f(x) + f(y)$ para todo $x, y \in V$, y $f(kx) = kf(x)$ para todo $k \in \mathbb{K}$ y todo $x \in V$. El nombre técnico con que se designa una aplicación lineal es *homomorfismo*, y cuando $V = W$ es *endomorfismo*.

(que representan el número de litros proceden de cada manantial) tienen por imagen

$$(1/3, 1/3, 1/3), \quad (2/3, 1/6, 1/6) \quad y \quad (0, 0, 1)$$

respectivamente (que representan el número de litros que caen el los tres pozos). Podemos hablar pues de una función bien definida,

$$f : \mathbb{R}^3 \to \mathbb{R}^3$$

donde las componentes del vector de partida, variable independiente, representan los litros procedentes de M_1, de M_2 y de M_3 respectivamente, y el vector de llegada hace referencia a los litros que caen (resp.) en los pozos A, B y C.

Si el sistema de acueductos y acequias ha permitido que lleguen 6 l. de M_1, 3 l. de M_2 y 4 l. de M_3, es claro que la cantidad de agua que cae en el pozo A es 4 l. Los otros cálculos son análogos, mantienen las proporciones, hay linealidad en el proceso. En forma matricial, la respuesta viene dada por:

$$(6 \quad 3 \quad 4) \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 1 \end{pmatrix} = (2+2 \quad 2+\frac{1}{2} \quad 2+\frac{1}{2}+4) = (4 \quad \frac{5}{2} \quad \frac{13}{2}),$$

o si multiplicamos por columnas (lo habitual), en vez de por filas, a través de la matriz traspuesta:

$$\begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{3} & \frac{1}{6} & 0 \\ \frac{1}{3} & \frac{1}{6} & 1 \end{pmatrix} \begin{pmatrix} 6 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ \frac{5}{2} \\ \frac{13}{2} \end{pmatrix}.$$

Vemos así que a cada aplicación lineal podemos asociarle una matriz, simplificando la escritura. Además, por cada base del espacio que escojamos, podemos dar una matriz diferente.

Asimismo, conviene recordar que, en general, para tratar un sistema lineal de ecuaciones es siempre deseable minimizar el número de operaciones a realizar, transformando el problema convenientemente. Mientras más elementos nulos tenga la matriz de coeficientes (y en todo caso unos, cuando no sean nulos), tanto más fácil se podrá operar con ella (en eso consistía el Método de Reducción de Gauss).

Análogamente, con respecto a aplicaciones lineales, mientras que inicialmente nos podemos encontrar matrices complicadas en términos de bases sencillas (la canónica, generalmente), el problema que resolvemos en este tema es el inverso[2]:

[2]La necesidad de operaciones fáciles con matrices, especialmente matrices elevadas a potencias grandes, se nos presentará, por ejemplo, al calcular la exponencial de una matriz, o al iterar sistemas acoplados discretizados, que pueden representar desde modelos presa–depredador hasta sistemas estocásticos –cadenas de Markov–, como veremos en el próximo tema.

Objetivo del tema: *Buscar una base (en general no tan simple como la inicial) tal que el cambio de sistema de referencia en la aplicación que usemos (esto es, hacer intervenir el cambio de base) permita una representación matricial lo más simple posible.*

Concretamente, dado V un espacio vectorial de dimensión finita (en el ejemplo anterior era $\mathbb{R}^3$) y f una aplicación lineal de V en sí mismo (endomorfismo), fijada una base de V, existe una matriz cuadrada $A \in M_n(\mathbb{K})$, de tal forma que $f(x) = Ax$, buscamos una nueva base con la que representar matricialmente a f de una forma más simple. Sería óptimo para el tipo de cálculos que desarrollaremos en el próximo tema obtener como resultado final una matriz diagonal (es decir, con todos sus elementos nulos fuera de la diagonal). Cuando seamos capaces de resolver el problema (no siempre) diremos que hemos diagonalizado la matriz A.

En esencia, el método que seguiremos (que no es único, por supuesto) consiste en buscar los vectores "adecuados" (autovectores) que permitan expresar de la forma más simple posible la imagen por f, y que dichos vectores formen una base. Dado que haremos cambio de base, nos vamos a centrar en la diagonalización por semejanza, es decir, buscamos que la nueva matriz del endomorfismo sea diagonal y semejante a la actual (véase la siguiente definición).

Definición 2. (Matrices semejantes) *Sean A y B dos matrices cuadradas de orden n. Decimos que A es semejante a B si existe una matriz cuadrada P invertible tal que $A = PBP^{-1}$.*

Para dar sentido al concepto anterior, hay que entender cómo afecta a la representación matricial de un endomorfismo un cambio de base. Dadas dos bases, C y B', de un mismo espacio vectorial de dimensión finita, se puede expresar las coordenadas de una base respecto de la otra (y viceversa). El proceso, guiado por una adecuada matriz de cambio (de C a B' p.ej.), llamémosla P, admite pasos "inversos", dicho de otro modo: algebraicamente, las matrices de cambio de base admiten matriz inversa, que denotamos por P^{-1}, con las que $PP^{-1} = P^{-1}P =$I, la matriz identidad.

Dado el endomorfismo f, y la base C, ¿cuál es exactamente la forma matricial representante de f en dicha base? La matriz, A_C, si elegimos "la multiplicación por columnas," viene dada, como en el Ejemplo 1, por los coeficientes a_{ij} procedentes de las siguientes expresiones:

$$f(x_i) = \sum_{j=1}^{n} a_{ji} x_j,$$

siendo $\{x_1, x_2, \ldots x_n\}$ los elementos de la base C. Matricialmente esto se lee $f(x) = A_C x_C$, donde x_C denota las coordenadas de x respecto de la base C.

¿Y si deseamos tener el representante matricial de f respecto de la base B'? Habrá que obtener una matriz tal que al introducir coordenadas respecto de B' calcule la imagen por f y la exprese en dicha base.

Matricialmente debería ser $f(x) = A_{B'} x_{B'}$, donde $x_{B'}$ denota las coordenadas de x respecto de la base B'. Si P es la matriz del cambio de base C a B', debemos tener la igualdad $A_{B'} = PA_C P^{-1}$. De modo que vemos cómo la información matricial con distintas bases hace referencia simplemente al uso de matrices semejantes.

Ejercicio: Comprueba que se satisfacen las siguientes propiedades de las matrices semejantes (para algunas de las propiedades necesitas saber que el determinante del producto de dos matrices cuadradas es el producto de los determinantes)

1. Si A es semejante a B, entonces $|A| = |B|$.
2. Si A es semejante a B, A^k es semejante a B^k.
3. Si A es semejante a B y $p(t) = c_n t^n + c_{n-1} t^{n-1} + \cdot c_1 t + c_0$ es un polinomio real con coeficientes en $\mathbb{R}$, entonces $p(A)$, entendida como $c_n A^n + c_{n-1} A^{n-1} + \ldots + c_1 A + c_0$, es semejante a $p(B)$.
4. Si A es semejante a B y A es regular (esto es, su determinante es distinto de cero), entonces B es regular, y A^{-1} es semejante a B^{-1}.

3.1. Autovalores y Autovectores. Propiedades

Consideramos dado un espacio vectorial V de dimensión finita n y un endomorfismo $f : V \to V$. ¿Qué elementos resultan idóneos para elegirlos como parte de una base? Suponemos dada una base en V, y consideramos la matriz A representante de f en dicha base. Buscamos un vector cuya imagen sea proporcional a él mismo, de modo que si forma parte de la nueva base, la matriz tendría una fila con un elemento en la diagonal y ceros en todos los demás. Esta idea es la que nos lleva a dar la siguiente definición.

Definición 3. (Autovalor y autovector) *Dada una matriz $A \in M_n(\mathbb{K})$, decimos que $\lambda \in \mathbb{K}$ es un* **autovalor** *o* **valor propio** *de A si $\exists\, x \in V,\ x \neq 0 \ /\ Ax = \lambda x$. En general, el cuerpo a considerar puede ser $\mathbb{R}$ o $\mathbb{C}$.*

Al vector x se le llama **autovector** *o* **vector propio** *asociado al autovalor λ.*

Evidentemente, si tenemos un autovalor para una matriz A, el autovector asociado no es único (todos los proporcionales también lo son), de hecho, es fácil ver que forman un conjunto particular tal y como señalamos en la siguiente definición.

Definición 4. (Subespacio propio) *Al conjunto de todos los autovectores asociados a un autovalor λ junto con el vector 0 se le suele notar A_λ y se le llama subespacio propio asociado al autovalor λ (de hecho,* ***es un subespacio vectorial****, es decir, suma de autovectores asociados a un mismo autovalor también es autovector, y lo mismo cabe esperar del producto de un escalar por un autovector; también se suele expresar como* ***subespacio invariante del endomorfismo****).*

$$A_\lambda = \{x \ / \ Ax = \lambda x\} \cup \{0\}$$

A continuación enumeramos algunas propiedades sobre los autovalores (la mayoría son fáciles de probar, y se dejan como ejercicio):

1. Autovalores distintos, $\lambda_1 \neq \lambda_2$, no tienen autovectores comunes: $A_{\lambda_1} \cap A_{\lambda_2} = \{0\}$.

   Es decir, un autovector x admite sólo un autovalor. (El recíproco, en general, no es cierto.)

De hecho, autovectores correspondientes a autovalores distintos son linealmente independientes (esta propiedad resulta esencial en lo que sigue): en efecto, consideramos un autovector $x \in A_\mu$ con $\mu \neq \lambda$, otro autovalor. Veamos que x es no es combinación lineal de elementos de A_λ. De serlo, $x = \alpha_1 v_1 + \ldots + \alpha_m v_m$ con $v_1, \ldots, v_m \in A_\lambda$. Entonces llegamos a la siguiente contradicción:

$$0 \neq (\lambda - \mu)x = \lambda(\alpha_1 v_1 + \ldots + \alpha_m v_m) - Ax = \alpha_1 A v_1 + \ldots + \alpha_m A v_m - Ax = Ax - Ax = 0.$$

2. A y A^t tienen los mismos autovalores (para probarlo basta recordar que el determinante de una matriz cuadrada coincide con el de su traspuesta, y la fórmula de la siguiente sección).

3. Si λ es autovalor de A, $k\lambda$ es un autovalor de kA.

4. Si λ es un autovalor de A, $\lambda - k$ es un autovalor de $A - kI$.

5. Si λ es un autovalor de A, y A es regular, $\frac{1}{\lambda}$ es autovalor de A^{-1}.

6. Si λ es autovalor de A, entonces λ^k es autovalor de A^k.

Polinomio característico

Nos proponemos a continuación identificar fácilmente a los autovalores de una matriz, usando lo que conocemos para sistemas de ecuaciones lineales.

Supongamos que $\lambda \in \mathbb{K}$ es un *autovalor* o *valor propio* de A, esto es, existe $x \in V$, $x \neq 0$ / $Ax = \lambda x$.
Podemos escribir $Ax - \lambda x = 0$ o sacando factor común, $(A - \lambda I)x = 0$.

Esta última relación representa un sistema homogéneo. Recordemos que para que un sistema homogéneo admita solución distinta de la trivial, debe ocurrir que el rango no sea máximo, o dicho de otro modo, que tras aplicar el método de reducción de Gauss obtengamos al menos una fila completa de ceros, o equivalentemente (y lo que más se suele usar para calcularlo), que el determinante de la matriz del sistema sea cero. Por tanto, λ es autovalor de A si

$$\mid A - \lambda I \mid = 0$$

Para obtener pues, los autovalores de A, bastará resolver la ecuación $\mid A - \lambda I \mid = 0$, llamada *ecuación característica* de A.

$$\mid A - \lambda I \mid = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.$$

Al determinante, que desarrollado es simplemente un polinomio en λ de grado menor o igual que n, se le llama **polinomio característico** de A :

$$p(\lambda) \ = |A - \lambda I| = \ c_n \lambda^n \ + \ c_{n-1} \lambda^{n-1} \ + \ \cdots \ + \ c_1 \lambda \ + \ c_0.$$

$\boxed{\text{Los autovalores de } A \text{ son, pues, los ceros de } \mathbb{K} \text{ de su polinomio característico.}}$ [3]

Nota: Para un breve repaso sobre el cálculo de determinantes, necesario como forma más directa para calcular polinomios característicos, véase el correspondiente apéndice al final del tema.

Calculemos los autovalores para la matriz del Ejemplo 1:

$$\begin{aligned} p(\lambda) = |A - \lambda I| &= \begin{vmatrix} \frac{1}{3} - \lambda & \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{6} - \lambda & \frac{1}{6} \\ 0 & 0 & 1 - \lambda \end{vmatrix} \\ &= (1-\lambda)\left[\left(\frac{1}{3} - \lambda\right)\left(\frac{1}{6} - \lambda\right) - \frac{2}{9}\right] \\ &= (1-\lambda)\left(\lambda - \frac{1}{4} - \frac{\sqrt{33}}{4}\right)\left(\lambda - \frac{1}{4} + \frac{\sqrt{33}}{4}\right), \end{aligned}$$

donde hemos elegido desarrollar el determinante por la tercera fila (para obtener más fácilmente su factorización, que es el verdadero fin último, más que la obtención del polinomio en si, es decir, no se debe multiplicar inútilmente).

Por tanto, los autovalores de la matriz son $\lambda_1 = 1$, $\lambda_2 = \frac{1}{4} + \frac{\sqrt{33}}{4}$ y $\lambda_3 = \frac{1}{4} - \frac{\sqrt{33}}{4}$.

Igual que el orden de una raíz en un polinomio es importante para el signo, en general, habrá distintas respuestas posibles (dimensión de subespacios propios asociados) según el orden de un autovalor como raíz del polinomio característico asociado a una matriz. A este respecto, dos nociones serán igualmente importantes. Introducimos la primera:

Definición 5. (Multiplicidad algebraica) *Si λ_0 es una raíz del polinomio característico de A de multiplicidad α, se dirá que λ_0 es un autovalor de orden α de A, y a α se le llama* **multiplicidad algebraica** *de λ, se suele notar $m_a(\lambda)$.*

Nota Si el cero es autovalor, eso es que $p(\lambda) = |A - \lambda I|$ se anula cuando $\lambda = 0$, entonces $p(0) = |A| = 0$. Recíprocamente, si el determinante de una matriz es no nulo, el cero no es autovalor.

Una vez resuelta la ecuación característica y obtenidos los autovalores de A, para calcular los autovectores habrá que resolver el sistema $(A - \lambda_i I)x = 0$.

¿Qué dimensión tiene este subespacio vectorial? Igual que para el orden del autovalor, ese número será importante para responder a la cuestión de la diagonalización, como se verá en el Teorema 9. Al venir dado por los grados de libertad del sistema de ecuaciones lineal homogéneo $(A - \lambda I)x = 0$, se tendrá que

$$dim(A_\lambda) = dim(N(A - \lambda I)) = dim(V) - \mathrm{rg}(A - \lambda I).$$

[3] Todo polinomio de orden n tiene n raíces complejas, pero si el cuerpo $\mathbb{K}$ con el que estamos es $\mathbb{R}$, puede ocurrir que no sean todos reales. Dicho de otro modo, todo autovalor es raíz del polinomio, pero toda raíz del polinomio no es autovalor (si no está en el cuerpo en que trabajamos, $\mathbb{Q}$ ó $\mathbb{R}$, sino en $\mathbb{C}$).

Definición 6. (Multiplicidad geométrica) *Se llama* **multiplicidad geométrica** *de* λ, *y se denota* $m_g(\lambda)$, *al número de autovectores linealmente independientes asociados a* λ, *o lo que es lo mismo, a* $dim\,(A_\lambda)$.

La multiplicidad geométrica del subespacio propio asociado a un autovalor no tiene porqué ser igual a la multiplicidad algebraica del autovalor. Considérese el siguiente ejemplo: $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ tiene por único autovalor a $\lambda = 1$, con multiplicidad algebraica 2. Sin embargo, el subespacio propio asociado tiene dimensión 1: $A_1 = \langle (1,0) \rangle$.

En general, se cumple que

$$\boxed{m_g(\lambda) = dim\,(A_\lambda) \leq m_a(\lambda).}$$

Obviamente sí es cierto que $m_g(\lambda) \geq 1$ si λ es autovalor. Previamente a la Definición 3 hemos introducido heurísticamente una de las ideas sobre cómo alcanzar el objetivo de la diagonalización de una matriz vía semejanza: a través de la búsqueda de autovalores y autovectores en el modo descrito hasta ahora.

En el caso concreto en que haya n autovalores distintos para una matriz cuadrada n x n, hay consecuencias claras: habrá n autovectores linealmente independientes entre sí (formarán una base y por tanto la matriz será diagonalizable, cf. Teorema 9).

Antes de dar con rigor el resultado de diagonalización, volvemos al Ejemplo 1, teníamos tres autovalores distintos: $\lambda_1 = 1$, $\lambda_2 = \dfrac{1+\sqrt{33}}{4}$ y $\lambda_3 = \dfrac{1-\sqrt{33}}{4}$. Cada uno de ellos tiene un subespacio propio asociado de dimensión al menos uno, A_{λ_1}, A_{λ_2} y A_{λ_3}. Pero autovalores distintos no tienen autovectores comunes, de hecho, autovectores de subespacios propios distintos son independientes entre sí como ya comentamos (propiedad 1, página 4). Como hay tres, llenan todo el espacio; cada subespacio propio tiene dimensión exactamente uno: se puede tomar una base del espacio formada por autovectores. Veámoslo:

Para el autovalor $\lambda = 1$ es claro que la última fila de la matriz genera una ecuación que sobra, nos quedamos con las dos primeras:

$$A - I = \begin{pmatrix} \frac{1}{3} - 1 & \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{6} - 1 & \frac{1}{6} \\ 0 & 0 & 1 - 1 \end{pmatrix} = \begin{pmatrix} \frac{-2}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{-5}{6} & \frac{1}{6} \\ 0 & 0 & 0 \end{pmatrix}.$$

$$(A - I)x = 0 \quad \Rightarrow \begin{pmatrix} \frac{-2}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{-5}{6} & \frac{1}{6} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \quad A_1 = \langle (1,1,1) \rangle.$$

Para el autovalor $\lambda_2 = \dfrac{1+\sqrt{33}}{4}$ hay que tomar dos ecuaciones independientes (áquellas que tengan un determinante 2x2 no nulo), por ejemplo las generadas por las filas segunda y la tercera (o la primera y la tercera):

$$A - (\frac{1+\sqrt{33}}{4})I = \begin{pmatrix} \frac{1}{3}-\lambda_2 & \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{6}-\lambda_2 & \frac{1}{6} \\ 0 & 0 & 1-\lambda_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{3}-\lambda_2 & \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{6}-\lambda_2 & \frac{1}{6} \\ 0 & 0 & 1-\lambda_2 \end{pmatrix}.$$

Con esta matriz de coeficientes generamos el sistema de ecuaciones

$$(A-\lambda_2 I)x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Rightarrow \begin{pmatrix} \frac{2}{3} & \frac{1}{6}-\lambda_2 & \frac{1}{6} \\ 0 & 0 & 1-\lambda_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \quad A_{\lambda_2} = \langle (\frac{3}{2}(\lambda_2 - \frac{1}{6}), 1, 0) \rangle.$$

Sin más que cambiar λ_2 por λ_3 tenemos que $A_{\lambda_3} = \langle (\frac{3}{2}(\lambda_3 - \frac{1}{6}), 1, 0) \rangle$.

Como anticipábamos antes, un sistema formado por tres vectores pertenecientes a los tres subespacios obtenidos, asociados a tres autovalores distintos, tiene rango tres, y forman una base del espacio:

$$H = \left\{ (1,1,1), \left(\frac{3}{2}\left(\lambda_2 - \frac{1}{6}\right), 1, 0 \right), \left(\frac{3}{2}\left(\lambda_3 - \frac{1}{6}\right), 1, 0 \right) \right\}.$$

Veamos por ejemplo que su determinante es no nulo:

$$\begin{vmatrix} 1 & 1 & 1 \\ \frac{3}{2}\left(\lambda_2 - \frac{1}{6}\right) & 1 & 0 \\ \frac{3}{2}\left(\lambda_3 - \frac{1}{6}\right) & 1 & 0 \end{vmatrix} = \begin{vmatrix} \frac{3}{2}\left(\lambda_2 - \frac{1}{6}\right) & 1 \\ \frac{3}{2}\left(\lambda_3 - \frac{1}{6}\right) & 1 \end{vmatrix} = \frac{3}{2}\left(\lambda_2 - \frac{1}{6}\right) - \frac{3}{2}\left(\lambda_3 - \frac{1}{6}\right) = \frac{3}{2}(\lambda_2 - \lambda_3) \neq 0.$$

Tras los conceptos de autovalor y autovector, y los comentarios iniciales hechos en la introducción, justo después de la Definición 2, sobre cómo usar una base u otra hace que la matriz representante de la aplicación lineal cambie por otra semejante, resulta natural preguntarse si los autovalores y autovectores cambian cuando cambiamos de base. Respondemos brevemente esta cuestión de "consistencia" para dar definitivamente el resultado sobre diagonalización.

Proposición 7. *Sean A y B dos matrices cuadradas de orden n semejantes (pongamos $B = PAP^{-1}$). Entonces tienen el mismo polinomio característico y, por lo tanto, los mismos autovalores.*

Los autovectores, en cambio, sí varían según la base: si x es un autovector de A asociado al autovalor λ, entonces Px es autovector de B asociado al mismo autovalor.

La prueba es inmediata: el polinomio característico viene dado por $|A - \lambda I| = 0$. Sustituyendo y aplicando que el producto de determinantes es el determinante del producto, y particularmente que $1 = |I| = |P^{-1}P| = |P^{-1}||P|$, concluimos que

$$|A-\lambda I| = |P^{-1}BP-\lambda I| = |P^{-1}BP-\lambda P^{-1}P| = |P^{-1}(B-\lambda I)P| = |P^{-1}||B-\lambda I||P| = |B-\lambda I|.$$

Por otro lado, si x es un autovector, entonces se satisface

$$Ax = \lambda x \quad \Rightarrow P^{-1}BPx = \lambda x \quad \Rightarrow BPx = \lambda Px.$$

3.2. Matrices diagonalizables.

Hemos localizado todos los autovectores posibles para la matriz A de nuestro ejemplo a partir de sus tres autovalores. Comprobamos la relación autovalor/autovector:

$$A\begin{pmatrix}1\\1\\1\end{pmatrix}=\lambda_1\begin{pmatrix}1\\1\\1\end{pmatrix},$$

$$A\begin{pmatrix}\frac{3}{2}(\lambda_2-\frac{1}{6})\\1\\0\end{pmatrix}=\lambda_2\begin{pmatrix}\frac{3}{2}(\lambda_2-\frac{1}{6})\\1\\0\end{pmatrix},$$

$$A\begin{pmatrix}\frac{3}{2}(\lambda_3-\frac{1}{6})\\1\\0\end{pmatrix}=\lambda_3\begin{pmatrix}\frac{3}{2}(\lambda_3-\frac{1}{6})\\1\\0\end{pmatrix},$$

o si escribimos todo junto en notación matricial:

$$\underbrace{\begin{pmatrix}\frac{1}{3}&\frac{1}{3}&\frac{1}{3}\\\frac{2}{3}&\frac{1}{6}&\frac{1}{6}\\0&0&1\end{pmatrix}}_{A}\underbrace{\begin{pmatrix}1&\frac{3}{2}(\lambda_2-\frac{1}{6})&\frac{3}{2}(\lambda_3-\frac{1}{6})\\1&1&1\\1&0&0\end{pmatrix}}_{P}=\underbrace{\begin{pmatrix}1&\frac{3}{2}(\lambda_2-\frac{1}{6})&\frac{3}{2}(\lambda_3-\frac{1}{6})\\1&1&1\\1&0&0\end{pmatrix}}_{P}\underbrace{\begin{pmatrix}\lambda_1&0&0\\0&\lambda_2&0\\0&0&\lambda_3\end{pmatrix}}_{D}.$$

En el ejemplo concreto acabamos de resolver el problema de obtener otra expresión matricial semejante a A pero con todos los elementos fuera de la diagonal cero. Usando el cambio de base desde la base canónica C a H, o lo que es lo mismo, a través de la matriz de paso P, hemos llegado (multiplicando la expresión anterior por la inversa de P, P^{-1}) a $P^{-1}AP=D$.

Comprueba que la inversa de la matriz de paso P viene dada por

$$P^{-1}=\begin{pmatrix}0&0&1\\\frac{4\sqrt{33}}{99}&\frac{1}{2}-\frac{\sqrt{33}}{198}&-\frac{7\sqrt{33}}{198}-\frac{1}{2}\\-\frac{4\sqrt{33}}{99}&\frac{\sqrt{33}}{198}+\frac{1}{2}&\frac{7\sqrt{33}}{198}-\frac{1}{2}\end{pmatrix}.$$

Asimismo, puedes comprobar también que $A=PDP^{-1}$.

Definición 8. (Matriz diagonalizable) *Sea $A\in\mathcal{M}_n(\mathbb{K})$. Diremos que A es diagonalizable sobre $\mathbb{K}$ si es semejante a una matriz diagonal.*

Dado un espacio vectorial V con una base B y en él un endomorfismo, éste vendrá representado por una matriz A. Si conseguimos encontrar una base B' formada por autovectores, $B'=\{x_1,x_2,\cdots,x_n\}$, ¿cuál será la matriz del endomorfismo en esta base?.

$$\begin{gathered}f(x_1)=\lambda_1x_1=(\lambda_1,0,0,\cdots,0)\\f(x_2)=\lambda_2x_2=(0,\lambda_2,0,\cdots,0)\\f(x_3)=\lambda_3x_3=(0,0,\lambda_3,\cdots,0)\\\vdots\\f(x_n)=\lambda_nx_n=(0,0,0\cdots,\lambda_n)\end{gathered}$$

La matriz del endomorfismo será, pues, diagonal:

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

siendo $D = P^{-1}AP$, con P la matriz del cambio de base, que resulta ser la matriz que tiene por columnas los autovectores de A colocados en el mismo orden en el que se colocan los autovalores de f en la diagonal.

¿Es siempre posible encontrar dicha base? ¿De serlo, corresponde exclusivamente al caso en que haya n autovalores distintos?

En general, la respuesta a ambas cuestiones es **NO**. La condición es que, relativo a cada autovalor, haya tantos autovectores independientes para poner en la base como la multiplicidad algebraica del autovalor al que están asociados. Exactamente:

Teorema 9. *Las condiciones necesarias y suficientes para que una matriz sea diagonalizable sobre el cuerpo $\mathbb{K}$ son dos:*

a) que el polinomio característico se pueda factorizar en el cuerpo base $\mathbb{K}$ en que trabajamos, y

b) que la multiplicidad de cada autovalor λ sea igual a la dimensión del subespacio propio asociado A_λ , es decir, que las multiplicidades algebraica y geométrica coincidan $m_a(\lambda) = m_g(\lambda)$.

En todo caso nos quedan dos cuestiones abiertas: cómo calcular la matriz inversa y qué ocurre cuando una matriz no es diagonalizable. La primera cuestión, meramente de cálculo, es tratada en un segundo apéndice, al final del tema. Respecto a la segunda, es decir, la situación general que se da con cualquier matriz cuadrada que queramos transformar en otra más simple, la mejor respuesta que se puede dar es la obtención de la forma canónica de Jordan.

3.3. Forma canónica de Jordan.

En el teorema anterior hemos visto cuándo una matriz es diagonalizable. Ahora bien, este resultado no es siempre aplicable (basta encontrar autovalores múltiples cuya multiplicidad algebraica no coincida con su multiplicidad geométrica, o bien que el polinomio característico no se pueda factorizar en el cuerpo en que estemos trabajando).

Un ejemplo similar al dado tras la Definición 6, pero de dimensión arbitraria, es el siguiente: la matriz cuadrada de orden m

$$J_\lambda^{(m)} = \begin{pmatrix} \lambda & 1 & 0 & & & & \\ 0 & \lambda & 1 & 0 & & & \\ & 0 & \lambda & 1 & 0 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & 0 & \lambda & 1 & 0 \\ & & & & 0 & \lambda & 1 \\ & & & & & 0 & \lambda \end{pmatrix}$$

no es diagonalizable (estas matrices se conocen con el nombre de **bloques de Jordan**).

El problema es que aunque λ tiene multiplicidad algebraica igual a n, no da lugar a n autovectores independientes sino a uno sólo.

Finalizaremos con el mejor resultado que se puede obtener en general (para matrices no diagonalizables): el Teorema de Jordan. Este teorema prueba que cualquier matriz cuadrada es semejante a una matriz formada por bloques de Jordan.

Teorema 10. *Sea A una matriz cuadrada. Entonces existen r autovalores $\lambda_1, \lambda_2, \ldots, \lambda_r$ (que pueden ser iguales) y r números naturales $m_1, m_2, \ldots, m_r$ tales que A es semejante a la matriz diagonal por bloques:*

$$J = \begin{pmatrix} J_{\lambda_1}^{(m_1)} & & & \\ & J_{\lambda_2}^{(m_2)} & & \\ & & \ddots & \\ & & & J_{\lambda_r}^{(m_r)} \end{pmatrix}$$

Esta matriz recibe el nombre de **forma canónica de Jordan** de la matriz A. En ella un mismo autovalor λ aparece en tantos bloques como indica $m_g(\lambda)$ y el número de veces que aparece en la diagonal de J es $m_a(\lambda)$.

Introducción heurística de los autovectores generalizados

A la luz del teorema, esperamos encontrar una matriz de paso $P = (v_1|v_2|\ldots|v_n)$ que reduzca a A a su forma de Jordan, $J = P^{-1}AP$, o dicho de otro modo, $AP = PJ$, de modo que atendiendo al primer bloque de J, deducimos que las m_1 primeras columnas de P satisfacen:

$$\begin{array}{ll} Av_1 = \lambda_1 v_1, & \textit{i.e., } v_1 \textit{ es un autovector correspondiente a } \lambda_1 \\ Av_i = \lambda v_i + v_{i-1} & i = 2, \cdots, m_1 \end{array}$$

o equivalentemente, expresándolo a través de la cadena

$$\begin{array}{rcl} (A - \lambda_1 I)v_1 & = & 0, \\ (A - \lambda_1 I)v_2 & = & v_1, \\ (A - \lambda_1 I)v_3 & = & v_2, \\ & \vdots & \\ (A - \lambda_1 I)v_{m_1} & = & v_{m_1 - 1}, \end{array}$$

Los vectores $v_2, v_3, \ldots, v_{m_1}$ no son propiamente autovectores, pero se comportan de una forma muy parecida, y se llaman **autovectores generalizados de** A, y la sucesión $\{v_1, v_2, \cdots, v_{m_1}\}$ se dice que es una **cadena de Jordan correspondiente a** λ_1. Naturalmente, cada bloque tiene su cadena correspondiente.

Método de Caros para el cálculo de J

Primeramente analizaremos el caso en que el polinomio característico no tiene raíces complejas.

1^o Se calculan los autovalores de $A: \ \lambda_1, \ \lambda_2, \ \cdots, \ \lambda_k,$ con sus multiplicidades algebraicas correspondientes $m_a(\lambda_1), \ m_a(\lambda_2), \ \cdots, \ m_a(\lambda_k)$.

2^o Para cada autovalor λ de multiplicidad algebraica $m_a(\lambda)$ se calculan los rangos de las matrices $(A-\lambda I)^p$ (a lo sumo habría que calcular hasta la potencia $(A-\lambda I)^{m_a(\lambda)}$).

$$\begin{array}{lllr}
rg(A-\lambda I) = r_1 \to x_1 = n - r_1 & & \text{si } \ x_1 < m, & \text{seguimos} \\
rg(A-\lambda I)^2 = r_2 \to x_2 = r_1 - r_2 & & \text{si } \ x_1 + x_2 < m, & \text{seguimos} \\
\vdots & & \vdots & \vdots \\
rg(A-\lambda I)^p = r_p \to x_p = r_{p-1} - r_p & & \text{si } \ x_1 + x_2 + \cdots + x_p = m, & \boxed{\text{FIN}}
\end{array}$$

$$\boxed{x_1 + x_2 + \cdots + x_p = \text{multiplicidad algebraica de } \lambda}$$

3^o Al valor propio λ le corresponden:

$$\begin{array}{ccc}
\boxed{x_1 - x_2} & \text{bloques de Jordan de orden} & \boxed{1} \\
\boxed{x_2 - x_3} & \text{bloques de Jordan de orden} & \boxed{2} \\
\vdots & \vdots & \vdots \\
\boxed{x_{p-1} - x_p} & \text{bloques de Jordan de orden} & \boxed{p-1} \\
\boxed{x_p} & \text{bloques de Jordan de orden} & \boxed{p}
\end{array}$$

Cálculo de P en el Método de Caros

Una vez calculada la matriz J por el método de CAROS, procedemos a calcular la matriz P , esto es, a calcular en el orden adecuado una base formada por autovectores y autovectores generalizados de A.

Para ello, dado un autovalor λ, llamamos $N_{k,\lambda} = N(A - \lambda I)^k$. Tomemos un vector

$$v_i \in N_{i,\lambda} \setminus N_{i-1,\lambda}$$

o lo que es lo mismo, un vector para el cual, según el método de CAROS, exista un bloque de Jordan de orden i (empezamos con el valor máximo para el que conozcamos la existencia de una caja).

Una vez obtenido v_i , calculamos:

$$\begin{array}{l}
v_{i-1} = (A - \lambda I)v_i \\
v_{i-2} = (A - \lambda I)v_{i-1} \\
\cdots\cdots \\
v_2 = (A - \lambda I)v_3 \\
v_1 = (A - \lambda I)v_2
\end{array}$$

y el vector v_1 resulta ser un **autovector** de A (o una combinación lineal de ellos) y los vectores $\{v_j\}$ $j = 2, \cdots, i$ son sus **autovectores generalizados**. Dichos vectores son las i primeras columnas de P : primero v_1, y después $v_2, \ldots, v_i$.

Esta operación se repite para cada bloque de Jordan que nos indique el método de CAROS, obteniendo así m autovectores (propios o generalizados) asociados al autovalor λ . Para ello, y tras haber empezado con la caja de mayor orden, se continua buscando vectores que satisfagan las condiciones expuestas y que sean, naturalmente, independientes de los anteriores.

Repitiendo el proceso para cada λ llegamos a obtener n vectores independientes, es decir, una base respecto de la cual, la matriz del endomorfismo viene dada por su Forma Canónica de Jordan, o lo que es lo mismo, hemos encontrado la matriz de paso P.

Finalizamos con un ejemplo:

Obtener la forma de Jordan (y matriz del cambio) asociada a la matriz

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

Lo primero que debemos hacer es calcular sus autovalores:

$$\begin{aligned}
|A-\lambda I| &= \begin{vmatrix} 4-\lambda & 0 & 1 & 0 \\ 1 & 1-\lambda & 1 & -1 \\ 0 & -2 & 3-\lambda & -1 \\ -1 & 2 & -1 & 4-\lambda \end{vmatrix} \\
&= \begin{vmatrix} 4-\lambda & 0 & 1 & 0 \\ 1 & 1-\lambda & 1 & -1 \\ 0 & -2 & 3-\lambda & -1 \\ 0 & 3-\lambda & 0 & 3-\lambda \end{vmatrix} \\
&= (4-\lambda)\begin{vmatrix} 1-\lambda & 1 & -1 \\ -2 & 3-\lambda & -1 \\ 3-\lambda & 0 & 3-\lambda \end{vmatrix} - \begin{vmatrix} 0 & 1 & 0 \\ -2 & 3-\lambda & -1 \\ 3-\lambda & 0 & 3-\lambda \end{vmatrix} \\
&= (4-\lambda)\left[(1-\lambda)(3-\lambda)^2-(3-\lambda)+(3-\lambda)^2+2(3-\lambda)\right]-(3-\lambda) \\
&= (3-\lambda)(4-\lambda)\left[(2-\lambda)(3-\lambda)+1\right]-(3-\lambda) \\
&= (3-\lambda)(4-\lambda)(7-5\lambda+\lambda^2)-(3-\lambda) \\
&= (3-\lambda)\left[(4-\lambda)(7-5\lambda+\lambda^2)-1\right] \\
&= (3-\lambda)^4.
\end{aligned}$$

Hemos concluido por tanto que hay un único autovalor, $\lambda = 3$, y de multiplicidad algebraica $m_a(3) = 4$.

Por el algoritmo de Caros debemos hallar el rango de $A-3I$, y sus sucesivas potencias hasta que sea necesario. Lo detallamos a continuación:

$$r_1 = \mathrm{rg}(A-3I) = \mathrm{rg}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & -2 & 1 & -1 \\ 0 & -2 & 0 & -1 \\ -1 & 2 & -1 & 1 \end{pmatrix} = 2.$$

De modo que $x_1 = 4 - r_1 = 4 - 2 = 2$. Como $x_1 \neq 4$, debemos continuar.

$$r_2 = \operatorname{rg}(A - 3I)^2 = \operatorname{rg}\begin{pmatrix} 1 & -2 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 1.$$

Así que $x_2 = r_1 - r_2 = 2 - 1 = 1$. Como $x_1 + x_2 = 3 \neq 4$, debemos seguir multiplicando.

$$r_3 = \operatorname{rg}(A - 3I)^3 = \operatorname{rg}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 0.$$

Esto significa que $x_3 = r_2 - r_3 = 1$, con lo que ahora sí sucede que $x_1 + x_2 + x_3 = 4$. Por el algoritmo sabemos que hay $x_1 - x_2 = 2 - 1 = 1$ caja de orden 1x1, $x_2 - x_3 = 1 - 1 = 0$ caja de orden 2x2, y $x_3 = 1$ caja de orden 3x3. Así, la forma de Jordan asociada a la matriz A es

$$J = \left(\begin{array}{c|ccc} 3 & 0 & 0 & 0 \\ \hline 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{array}\right).$$

Para obtener la matriz de paso P nos fijamos primero en la caja mayor, e intentamos encontrar un autovector generalizado que esté en $N_{3,\lambda} \backslash N_{2,\lambda}$. Como $(A - 3I)^3$ era la matriz idénticamente nula, $N_{3,\lambda} = \mathbb{R}^4$, luego basta tomar un vector de $\mathbb{R}^4$ que no esté en $N_{2,\lambda}$. La ecuación que define el subespacio vectorial $N_{2,\lambda}$ (recuérdese que $(A - 3I)^2$ tenía rango 1) es

$$N_{2,\lambda} = \{(y_1, y_2, y_3, y_4) \quad | \quad y_1 - 2y_2 + y_3 - y_4 = 0\} = \langle(1, -2, 1, -1)\rangle^{\perp}.$$

Como $N_{3,\lambda} = \mathbb{R}^4 = \langle(1, -2, 1, -1)\rangle \oplus \langle(1, -2, 1, -1)\rangle^{\perp}$, es obvio que podemos tomar

$$v_3 = (1, -2, 1, -1)^t.$$

Ahora simplemente calculamos

$$v_2 = (A - 3I)v_3 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & -2 & 1 & -1 \\ 0 & -2 & 0 & -1 \\ -1 & 2 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 7 \\ 5 \\ -7 \end{pmatrix}.$$

Igualmente, obtenemos con otra multiplicación

$$v_1 = (A - 3I)v_2 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & -2 & 1 & -1 \\ 0 & -2 & 0 & -1 \\ -1 & 2 & -1 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 7 \\ 5 \\ -7 \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \\ -7 \\ 0 \end{pmatrix}.$$

Para completar P debemos fijarnos en la caja 1x1, es decir, debemos obtener otro autovector de A, esto es, otro elemento del núcleo de $A - 3I$ (que tenía dimensión 2), que sea, por supuesto, independiente de v_1, el autovector ya obtenido.

$$N_{1,\lambda} = N(A - 3I) = \{y = (y_1, y_2, y_3, y_4) \quad | \quad (A - 3I)y = 0\} =$$

$$\left\{(y_1, y_2, y_3, y_4) \;\middle|\; \begin{array}{ccccc} y_1 & & +y_3 & & = 0, \\ y_1 & -2y_2 & +y_3 & -y_4 & = 0. \end{array} \right\}$$

Concluimos (pasando y_3 e y_4 a la derecha) que $N_{1,\lambda} = \langle(1, 0, -1, 0), (0, -1, 0, 2)\rangle$. El primero ya está en la segunda caja, y el segundo vector es el elemento que nos faltaba. Por el orden en que hemos escrito J (que por supuesto no es único), la matriz de paso resultante es:

$$P = \begin{pmatrix} 0 & 7 & 2 & 1 \\ -1 & 0 & 7 & -2 \\ 0 & -7 & 5 & 1 \\ 2 & 0 & -7 & -1 \end{pmatrix},$$

y se puede comprobar que la inversa es

$$P^{-1} = \begin{pmatrix} \frac{3}{7} & \frac{1}{7} & \frac{3}{7} & \frac{4}{7} \\ \frac{32}{343} & \frac{6}{343} & \frac{-17}{343} & \frac{3}{343} \\ \frac{5}{49} & \frac{4}{49} & \frac{5}{49} & \frac{2}{49} \\ \frac{1}{7} & \frac{-2}{7} & \frac{1}{7} & \frac{-1}{7} \end{pmatrix},$$

y finalmente se puede ver también, con lo que concluye el ejercicio, que $P^{-1}AP = J$.

Caso en que el polinomio característico tenga raíces complejas

Si el polinomio característico tiene raíces complejas, simples o múltiples, y el cuerpo con el que trabajamos no es $\mathbb{C}$, la *forma canónica real de Jordan* de la matriz A adopta una forma algo más complicada:

$$P^{-1}AP = J = \begin{pmatrix} D & I_2 & & & \\ & D & I_2 & & \\ & & \ddots & \ddots & \\ & & & D & I_2 \\ & & & & D \end{pmatrix}$$

siendo $D = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ e $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ para $\lambda = a \pm ib$, pues si λ es un cero del polinomio característico, también lo es su conjugado $\bar{\lambda}$, y con la misma multiplicidad.

Para clarificar los conceptos veamos un ejemplo:

Sea A una matriz cuadrada de orden 9 con λ_1 autovalor real simple siendo v_1 su correspondiente autovector, λ_2 autovalor real doble siendo v_2 su autovector y v_3 su autovector generalizado, $\lambda_3 = a + ib$ autovalor complejo simple y $v_4 = u_1 + iw_1$ su correspondiente autovector y por último $\lambda_4 = c + id$ autovalor complejo doble con $v_5 = u_2 + iw_2$ como autovector y $v_6 = u_3 + iw_3$ como autovector generalizado:

$Av_1 = \lambda_1 v_1$

$Av_2 = \lambda_2 v_2$

$Av_3 = \lambda_2 v_3 + v_2$

$$Av_4 = \lambda_3 v_4 \Longleftrightarrow A(u_1 + iw_1) = (a + ib)(u_1 + w_1) \Longleftrightarrow \begin{cases} Au_1 = au_1 - bw_1 \\ Aw_1 = bu_1 + aw_1 \end{cases}$$

$$Av_5 = \lambda_4 v_5 \Longleftrightarrow A(u_2 + iw_2) = (c + id)(u_2 + w_2) \Longleftrightarrow \begin{cases} Au_2 = cu_2 - dw_2 \\ Aw_2 = du_2 + cw_2 \end{cases}$$

$$Av_6 = \lambda_4 v_4 + v_5 \Longleftrightarrow A(u_3 + iw_3) = (c+id)(u_3 + iw_3) + (u_2 + iw_2) \Longleftrightarrow \begin{cases} Au_3 = cu_3 - dw_3 + u_2 \\ Aw_3 = du_3 + cw_3 + w_2 \end{cases}$$

En definitiva, AP=PB siendo $P = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ v_1 & v_2 & v_3 & u_1 & w_1 & u_2 & w_2 & u_3 & w_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$ y

$$J = \begin{pmatrix} \lambda_1 & & & \\ & \begin{bmatrix} \lambda_2 & 1 \\ & \lambda_2 \end{bmatrix} & & \\ & & \begin{bmatrix} a & b \\ -b & a \end{bmatrix} & \\ & & & \begin{bmatrix} \begin{bmatrix} c & d \\ -d & c \end{bmatrix} & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ & \begin{bmatrix} c & d \\ -d & c \end{bmatrix} \end{bmatrix} \end{pmatrix}$$

Apéndices

Un breve recordatorio sobre determinantes

Dado que la forma más corta de obtener el polinomio característico es a través del cálculo de un determinante, recordamos algunas cuestiones de cálculo al respecto. El determinante de una matriz 3x3 es:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{21}a_{12}a_{33}.$$

La forma más sencilla de calcular determinantes de cualquier orden, y particularmente útil para órdenes mayores, consiste en desarrollar a partir de una fila o una columna suma de determinantes de orden menor, afectados por coeficientes con signo $(-1)^{i+j}$, como señalamos en el siguiente ejemplo:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} =$$

$$a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{41} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \end{vmatrix}.$$

Evidentemente, conviene desarrollar por filas o columnas donde haya el mayor número posible de ceros. Es posible sumar filas entre sí o columnas entre sí para obtener convenientemente ceros antes de calcular el determinante, sin que éste se vea afectado en su valor final (igual que manipulaciones análogas en sistemas de ecuaciones generaban sistemas equivalentes).

Cálculo de la matriz inversa

Una matriz cuadrada con rango máximo, o equivalentemente, determinante no nulo, posee inversa. Existen tres formas sencillas de hallar la inversa de una matriz dada, que vemos a continuación ejemplificadas a través de un mismo caso:

a) utilizando un resultado de Cayley y Hamilton,

b) con ayuda de la matriz adjunta, y

c) a través del cambio de base.

Método 1:

Teorema 11. (Teorema de Cayley-Hamilton) *Toda matriz cuadrada A sobre un cuerpo $\mathbb{K}$ es raíz de su polinomio característico.*

Es decir, si $p(\lambda) = c_n\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_0$ es el polinomio característico y Θ_n es la matriz nula de orden n , entonces

$$\text{“}p(A)\text{”} = c_nA^n + c_{n-1}A^{n-1} + \cdots + c_0I = \Theta_n.$$

Sea A una matriz invertible (es decir, $|A| \neq 0$). Teniendo en cuenta que $p(A) = \Theta_n$, se obtiene que $c_nA^n + c_{n-1}A^{n-1} + \cdots + c_1A = -c_0I$. Al ser $c_0 = |A| \neq 0$, podemos dividir por dicha cantidad:

$$I = -\frac{c_n}{c_0}A^n - \frac{c_{n-1}}{c_0}A^{n-1} - \frac{c_{n-2}}{c_0}A^{n-2} - \cdots - \frac{c_1}{c_0}A.$$

Multiplicando por A^{-1} se concluye que

$$\boxed{A^{-1} = -\tfrac{c_{n-1}}{c_0}A^{n-1} - \tfrac{a_1}{a_n}A^{n-2} - \tfrac{a_2}{a_n}A^{n-3} - \cdots - \tfrac{a_{n-1}}{a_n}I.}$$

Ejemplo 12. *Consideramos* $C = \begin{pmatrix} 1 & 2 & -1 \\ 3 & -5 & 4 \\ 2 & 2 & 1 \end{pmatrix}$. *Su polinomio característico es:*

$$|C - \lambda I| = \begin{vmatrix} 1 & 2-\lambda & -1 \\ 3 & -5-\lambda & 4 \\ 2 & 2 & 1-\lambda \end{vmatrix} = -\lambda^3 - 3\lambda^2 + 21\lambda - 19.$$

Por el teorema se tiene que

$$C^3 + 3C^2 - 21C + 19I = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Multiplicando por C^{-1}, tenemos la relación:

$$C^2 + 3C - 21I = -19C^{-1}.$$

Calculamos C^2 :

$$C^2 = \begin{pmatrix} 1 & 2 & -1 \\ 3 & -5 & 4 \\ 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 \\ 3 & -5 & 4 \\ 2 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & -10 & 6 \\ -4 & 39 & -19 \\ 10 & -4 & 7 \end{pmatrix},$$

con lo que

$$C^{-1} = -\frac{1}{19}\left[\begin{pmatrix} 5 & -10 & 6 \\ -4 & 39 & -19 \\ 10 & -4 & 7 \end{pmatrix} + 3\begin{pmatrix} 1 & 2 & -1 \\ 3 & -5 & 4 \\ 2 & 2 & 1 \end{pmatrix} - 21\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right] = \begin{pmatrix} \frac{13}{19} & \frac{4}{19} & \frac{-3}{19} \\ \frac{-5}{19} & \frac{-3}{19} & \frac{7}{19} \\ \frac{-16}{19} & \frac{-2}{19} & \frac{11}{19} \end{pmatrix}.$$

Método 2:

El segundo método obedece una simple fórmula, aunque para su uso requerimos algunas explicaciones adicionales, la inversa de A, es la traspuesta de la adjunta dividido por el determinante:

$$\boxed{A^{-1} = \frac{(Adj(A))^t}{|A|}.}$$

Llamamos adjunta de una matriz A a otra matriz, que denotamos $Adj(A)$, del mismo orden donde $Adj(A)_{ij}$ es el adjunto del elemento a_{ij} en la matriz A. El elemento adjunto de A relativo al elemento a_{ij} es el determinante de la matriz resultante de eliminar la fila i y la columna j, precedido del signo $(-1)^{i+j}$.

Ejemplo 13. *Sea, como antes,* $C = \begin{pmatrix} 1 & 2 & -1 \\ 3 & -5 & 4 \\ 2 & 2 & 1 \end{pmatrix}$*. Para calcular $Adj(C)_{11}$ tomamos la matriz $\begin{pmatrix} -5 & 4 \\ 2 & 1 \end{pmatrix}$ y calculamos su determinante, $\begin{vmatrix} -5 & 4 \\ 2 & 1 \end{vmatrix} = -13$, con lo que $Adj(C)_{11} = (-1)^{1+1}(-13) = -13$. Análogamente, vamos construyendo los demás, eliminando la fila y columna correspondiente, calculando el determinante y cambiando el signo de lo obtenido alternanadamente:*

$$Adj(C) = \begin{pmatrix} \begin{vmatrix} -5 & 4 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 3 & 4 \\ 2 & 1 \end{vmatrix} & \begin{vmatrix} 3 & -5 \\ 2 & 2 \end{vmatrix} \\ -\begin{vmatrix} 2 & -1 \\ 2 & 1 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 2 & 2 \end{vmatrix} \\ \begin{vmatrix} 2 & -1 \\ -5 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ 3 & 4 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 3 & -5 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} -13 & 5 & 16 \\ -4 & 3 & 2 \\ 3 & -7 & -11 \end{pmatrix}.$$

La traspuesta de la adjunta consiste simplemente en intercambiar filas y columnas, es decir, $([Adj(C)]^t)_{ij} = [Adj(C)]_{ji}$:

$$[Adj(C)]^t = \begin{pmatrix} -13 & -4 & 3 \\ 5 & 3 & -7 \\ 16 & 2 & -11 \end{pmatrix}.$$

El determinante de C es -19, con lo que

$$C^{-1} = -\frac{1}{19}\begin{pmatrix} -13 & -4 & 3 \\ 5 & 3 & -7 \\ 16 & 2 & -11 \end{pmatrix}.$$

Método 3:

La tercera forma consiste la matriz que se quiere invertir como la información sobre un cambio de base. Si $(x', y', z') = (x, y, z)A$, (también lo podríamos representar con una multiplicación por columnas) representan distintas coordenadas de un mismo vector respecto dos bases, es claro que obtener la matriz tal que (x, y, z) queda representado en función de (x', y', z') debe representar el paso inverso, esto es, la matriz resultante es A^{-1}.

Ejemplo 14. *Dada la relación*

$$(x' \quad y' \quad z') = (x \quad y \quad z)\begin{pmatrix} 1 & 2 & -1 \\ 3 & -5 & 4 \\ 2 & 2 & 1 \end{pmatrix},$$

si despejamos $(x \quad y \quad z)$ en función de $(x' \quad y' \quad z')$: (vamos a resolver el sistema, pero no directamente, para eso tenemos el método de Gauss)

$$\left(\begin{array}{ccc|c} 1 & 3 & 2 & x' \\ 2 & -5 & 2 & y' \\ -1 & 4 & 1 & z' \end{array}\right) \begin{array}{c} \Rightarrow \\ F_2 - 2F_1 \\ F_3 + F_1 \end{array} \left(\begin{array}{ccc|c} 1 & 3 & 2 & x' \\ 0 & -11 & -2 & y' - 2x' \\ 0 & 7 & 3 & z' + x' \end{array}\right)$$

$$\begin{array}{c} \\ \Rightarrow \\ 11F_3 + 7F_2 \end{array} \left(\begin{array}{ccc|c} 1 & 3 & 2 & x' \\ 0 & -11 & -2 & y' - 2x' \\ 0 & 0 & 19 & 11(z' + x') + 7(y' - 2x') \end{array}\right),$$

de donde resulta

$$z = \frac{1}{19}(-3x' + 7y' + 11z'),$$

$$y = \frac{4}{19}x' - \frac{3}{19}y' - \frac{2}{19}z',$$

$$x = \frac{13}{19}x' - \frac{5}{19}y' - \frac{16}{19}z'.$$

Escrito en forma matricial, obtenemos de nuevo C^{-1} :

$$(x \quad y \quad z) = (x' \quad y' \quad z')C^{-1}.$$