Capítulo 1: Diagonalización de matrices

1. Matrices y determinantes

Definición 1.1. Una matriz es un arreglo rectangular de números reales

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}.$$

La matriz es de orden $n \times m$ si consta de n filas y m columnas. El conjunto de todas las matrices de orden $n \times m$ se denota $M_{n\times m}$.
El elemento a_{ij} pertenece a la i–ésima columna y a la j–ésima columna. Suele escribirse la matriz en forma abreviada como $A = (a_{ij})_{j=1,\dots,m}^{i=1,\dots,n}$, o incluso $A = (a_{ij})$.
La *diagonal principal* de una matriz está formada por los elementos diagonales tomados desde la esquina superior izquierda de la matriz hasta la esquina inferior derecha.

Definición 1.2. La traspuesta de una matriz A, denotada por A^T, es la matriz que se obtiene al intercambiar las filas por la columnas de A

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix} \in M_{m\times n}.$$

Podemos definir dos operaciones con matrices, la *suma* y la *multiplicación*, que ya conocemos de cursos anteriores. Las **propiedades** más importantes de estas operaciones, así como de la transposición, son las siguientes. Asumimos que los órdenes de las matrices que intervienen en cada una de las leyes enumeradas a continuación permiten las operaciones indicadas, y que $\alpha, \beta \in \mathbb{R}$.

1. $(A^T)^T = A$.
2. $(A + B)^T = A^T + B^T$.
3. $A + B = B + A$ (ley conmutativa).
4. $A + (B + C) = (A + B) + C$ (ley asociativa).
5. $\alpha(A + B) = \alpha A + \alpha B$.
6. $(\alpha + \beta)A = \alpha A + \beta A$.
7. La multiplicación de matrices no es conmutativa, es decir no siempre se tiene $AB = BA$.
8. $A(BC) = (AB)C$ (ley asociativa).
9. $A(B + C) = AB + AC$ (ley distributiva).

1.1. Matrices cuadradas. En este curso estamos interesados principalmente en matrices cuadradas. Una matriz es *cuadrada* si tiene el mismo número de filas que de columnas, $n = m$. La *traza* de una matriz cuadrada A es la suma de sus elementos diagonales, $\text{traza}\,(A) = \sum_{i=1}^{n} a_{ii}$.

Definición 1.3. La matriz identidad de orden n es

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

La matriz de orden n que tiene todos sus elementos nulos es la matriz *nula*, y se denota por O_n. Se verifica $I_nA = AI_n = A$ y $O_nA = AO_n = O_n$.

Definición 1.4. Una matriz cuadrada A es regular o inversible si existe una matriz B tal que $AB = BA = I_n$. La matriz B con esta propiedad se llama la inversa de A y se denota por A^{-1}.

Teorema 1.5. *La inversa de una matriz, si existe, es única.*

Es fácil probar la unicidad de A^{-1}. Si B is otra matriz inversa de A, entonces $BA = I_n$ y

$$B = BI_n = B(AA^{-1}) = (BA)A^{-1} = I_nA^{-1} = A^{-1},$$

demostrando que $B = A^{-1}$.
Enumeramos a continuación algunas **propiedades de la inversa de una matriz**. Asumimos que las matrices que aparecen son inversibles.

1. $(A^{-1})^{-1} = A$.
2. $(A^T)^{-1} = (A^{-1})^T$.
3. $(AB)^{-1} = B^{-1}A^{-1}$.

1.2. Determinantes. Podemos asociar un número real, llamado determinante, a toda matriz cuadrada A, que denotaremos como $|A|$ o $\det(A)$, de la siguiente forma:

- Para una matriz de orden 1, $A = (a)$, $\det(A) = a$.
- Para una matriz de orden 2, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\det(A) = ad - bc$.
- Para una matriz de orden 3

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

Esta manera de calcular el determinante se conoce como *desarrollo del determinante* por la primera columna, pero puede hacerse por cualquier otra fila o columna, dando el mismo resultado. Nótese que aparece el signo de $(-1)^{i+j}$ en frente del elemento a_{ij}.

Antes de continuar con la definición inductiva, veamos un ejemplo.

Ejemplo 1.6. Calcular el siguiente determinante, desarrollándolo por la segunda columna.

$$\begin{vmatrix} 1 & 2 & 1 \\ 4 & 3 & 5 \\ 3 & 1 & 3 \end{vmatrix} = (-1)^{1+2}2\begin{vmatrix} 4 & 5 \\ 3 & 3 \end{vmatrix} + (-1)^{2+2}3\begin{vmatrix} 1 & 1 \\ 3 & 3 \end{vmatrix} + (-1)^{2+3}1\begin{vmatrix} 1 & 1 \\ 4 & 5 \end{vmatrix}$$
$$= -2 \cdot (-3) + 3 \cdot (0) - (1) \cdot 1 = 5$$

- Para una matriz de orden n, la definición de determinante es el mismo que para matrices de orden 3, desarrollando el determinante por una línea de la matriz, reduciendo de esta forma el orden de los determinantes a calcular a $n-1$. Para un determinante de orden 4, este proceso lleva a calcular 4 determinantes de orden 3.

Definición 1.7. Dada una matriz A de orden n, el menor complementario del elemento a_{ij} es el determinante de orden $n-1$ que resulta al suprimir la fila i y la columna j correspondientes al elemento a_{ij}. El adjunto, A_{ij}, del elemento a_{ij} es el menor complementario multiplicado por $(-1)^{i+j}$.

De acuerdo a esta definición, el determinante de la matriz A puede calcularse como

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \qquad \text{(por la fila } i\text{)}$$

o, equivalentemente

$$|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} \qquad \text{(por la columna } j\text{)}.$$

Ejemplo 1.8. Encontrar el valor del determinante

$$\begin{vmatrix} 1 & 2 & 0 & 3 \\ 4 & 7 & 2 & 1 \\ 1 & 3 & 3 & 1 \\ 0 & 2 & 0 & 7 \end{vmatrix}.$$

Respuesta: Desarrollando el determinante por la tercera columna, encontramos

$$\begin{vmatrix} 1 & 2 & 0 & 3 \\ 4 & 7 & 2 & 1 \\ 1 & 3 & 3 & 1 \\ 0 & 2 & 0 & 7 \end{vmatrix} = (-1)^{3+2} 2 \begin{vmatrix} 1 & 2 & 3 \\ 1 & 3 & 1 \\ 0 & 2 & 7 \end{vmatrix} + (-1)^{3+3} 3 \begin{vmatrix} 1 & 2 & 3 \\ 4 & 7 & 1 \\ 0 & 2 & 7 \end{vmatrix}.$$

Las principales **propiedades de los determinantes** se enumeran a continuación. Suponemos que las matrices A y B que aparecen son cuadradas del mismo orden n y que $\lambda \in \mathbb{R}$.

1. $|A| = |A^T|$.
2. $|\lambda A| = \lambda^n |A|$.
3. $|AB| = |A||B|$.
4. Una matriz A es inversible si y sólo si su determinante es no nulo, $|A| \neq 0$; en este caso $|A^{-1}| = \frac{1}{|A|}$.
5. Si en un determinante dos filas (columnas) se intercambian, entonces el valor del determinante cambia de signo.
6. Si en un determinante dos filas (columnas) son idénticas, entonces el valor del determinante es cero.
7. Si todos los elementos de una fila (columna) de un determinante se multiplican por la misma constante, entonces el valor del determinante se multiplica también por esta constante.
8. Si en un determinante se le suma a una fila (columna) otra fila (columna) multiplicada por una constante, entonces el valor del determinante no varía.

El resultado siguiente es muy útil para comprobar si una matriz admite inversa o no.

Teorema 1.9. *Una matriz cuadrada A admite inversa si y sólo si $|A| \neq 0$.*

2. Diagonalización de matrices

Definición 2.1. Dos matrices A y B de orden n son similares si existe una matriz inversible P tal que

$$B = P^{-1} A P.$$

Definición 2.2. Una matriz A es diagonalizable si es similar a una matriz diagonal D, es decir, si existen una matriz diagonal D y una matriz P inversible tales que $D = P^{-1}AP$.

Por supuesto, D diagonal significa que todo elemento fuera de la diagonal principal de la matriz, es cero y, por tanto, tiene la forma

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \quad \lambda_1, \dots, \lambda_n \in \mathbb{R}.$$

Una aplicación inmediata es el cálculo de las potencias de una matriz diagonalizable, como pone de manifiesto el siguiente resultado, que será útil cuando estudiemos sistemas dinámicos.

Proposición 2.3. *Si A es diagonalizable, entonces para todo $m \geq 1$*

$$A^m = P D^m P^{-1}, \tag{2.1}$$

donde

$$D^m = \begin{pmatrix} \lambda_1^m & 0 & \dots & 0 \\ 0 & \lambda_2^m & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^m \end{pmatrix}.$$

Demostración. Dado que A es diagonalizable, $A = P^{-1}DP$, con D diagonal, por tanto

$$\begin{aligned} A^m &= (PDP^{-1})PDP^{-1}) \overset{m}{\cdots} (PDP^{-1}) \\ &= PD(P^{-1}P)D \cdots D(P^{-1}P)DP^{-1} \\ &= PDI_nD \cdots DI_nDP^{-1} = PD^mP^{-1}. \end{aligned}$$

La expresión para D^m se obtiene por inducción sobre m. □

Ejemplo 2.4. En un día dado, el profesor X puede dar la clase bien o darla mal. Después de un buen día, la probabilidad de que la siguiente clase vaya bien es 1/2, mientras que depués de una mala clase el profesor se desmoraliza y la probabilidad de que la siguiente clase vaya bien es sólo de 1/9. Sea b_t (m_t) la probabilidad de que la clase número t de X sea buena (mala). Suponiendo que la primera clase ($t = 1$) ha ido bien, es decir, $b_1 = 1$, $m_1 = 0$, ¿cuál es la probabilidad de que la quinta clase sea buena (mala)?
RESPUESTA: Los datos del problema pueden organizarse como el siguiente sistema de ecuaciones, que relacionan la probabilidad de un día bueno/malo con la calidad de la clase del día anterior:

$$b_{t+1} = \frac{1}{2}b_t + \frac{1}{9}m_t,$$

$$m_{t+1} = \frac{1}{2}b_t + \frac{8}{9}m_t.$$

En forma matricial

$$\begin{pmatrix} b_{t+1} \\ m_{t+1} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{9} \\ \frac{1}{2} & \frac{8}{9} \end{pmatrix} \begin{pmatrix} b_t \\ m_t \end{pmatrix}, \qquad t = 1, 2, \dots$$

por lo que

$$\begin{pmatrix} b_5 \\ m_5 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{9} \\ \frac{1}{2} & \frac{8}{9} \end{pmatrix}^4 \begin{pmatrix} b_1 \\ m_1 \end{pmatrix}.$$

Si la matriz fuera diagonalizable y pudiéramos encontrar las matrices P y D, entonces el cálculo de la décima potencia de A sería sencillo, utilizando la Proposición 2.3. Volveremos sobre este ejemplo más tarde, cuando hayamos estudiado los métodos de diagonalización.

Definición 2.5. Sea A una matriz de orden n. Se dice que $\lambda \in \mathbb{R}$ es un autovalor (o valor propio) de A y que $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{u} \neq \mathbf{0}$, es un autovector (o vector propio) de A asociado a λ si

$$A\mathbf{u} = \lambda\mathbf{u}.$$

El conjunto de los autovalores de A se denota $\sigma(A) = \{\lambda_1, \dots, \lambda_k\}$ y se llama espectro de A. El conjunto de los autovectores de A asociados al mismo autovalor λ, incluyendo el vector nulo, se denota $\mathrm{S}(\lambda)$, y es el autoespacio o subespacio propio asociado a λ.

El siguiente resultado muestra que un autovector está asociado a un único autovalor.

Proposición 2.6. *Sea un vector* $\mathbf{u}$ *no nulo tal que* $\mathbf{u} \in \mathrm{S}(\lambda) \cap \mathrm{S}(\mu)$. *Entonces* $\lambda = \mu$.

Demostración. Suponemos que $\mathbf{0} \neq \mathbf{u} \in \mathrm{S}(\lambda) \cap \mathrm{S}(\mu)$. Entonces

$$A\mathbf{u} = \lambda\mathbf{u}$$
$$A\mathbf{u} = \mu\mathbf{u}.$$

Restando ambas ecuaciones obtenemos $\mathbf{0} = (\lambda - \mu)\mathbf{u}$ y, dado que $\mathbf{0} \neq \mathbf{u}$, debe ser $\lambda = \mu$. □

Es importante recordar en lo que sigue que para una matriz arbitraria A, *el rango de la matriz* es el número de filas o de columnas linealmente independientes (ambos números necesariamente coinciden). El rango es también el orden del mayor menor no nulo de A.

Teorema 2.7. *El número real λ es un autovalor de A si y sólo si*

$$|A - \lambda I_n| = 0.$$

Además, $\mathrm{S}(\lambda)$ *es el conjunto de soluciones, incluyendo el vector nulo, del sistema lineal homogéneo*

$$(A - \lambda I_n)\mathbf{u} = \mathbf{0},$$

y, por tanto, es un subespacio vectorial de dimensión (o número de parámetros libres que se necesitan para describir el conjunto de soluciones)

$$\dim \mathrm{S}(\lambda) = n - \mathrm{rango}(A - \lambda I_n).$$

Demostración. Suponiendo que $\lambda \in \mathbb{R}$ es un autovalor de A llegamos a que el sistema $(A - \lambda I_n)\mathbf{u} = \mathbf{0}$ admite alguna solución no trivial, $\mathbf{u}$. Dado que el sistema es lineal y homogéneo, esto implica por el Teorema de Rouché–Frobenius que el determinante del sistema es cero, $|A - \lambda I_n| = 0$. Esto prueba la primera parte del teorema. La segunda parte sobre $\mathrm{S}(\lambda)$ se sigue de la definición de autovector y del hecho de que el conjunto de soluciones de un sistema lineal homogéneo es un subespacio vectorial (es decir, la suma de dos soluciones es de nuevo una solución, al igual que el producto de una solución por un número real es de nuevo una solución). Finalmente, la afirmación sobre la dimensión del subespacio solución es consecuencia también del Teorema de Rouché–Frobenius. □

Definición 2.8. Se llama polinomio característico de la matriz A al polinomio de orden n dado por

$$p_A(\lambda) = |A - \lambda I_n|.$$

Obsérvese que, por el teorema anterior, los autovalores propios de A son la raíces del polinomio p_A, Este polinomio es de grado n y, por tanto, de acuerdo al Teorema Fundamental del Álgebra, tiene exactamente n raíces complejas (no necesariamente diferentes, algunas raíces pueden repetirse; el número de veces que se repiten se llama multiplicidad de la raíz. Por ejemplo, el polinomio $x^2 - 2x + 1) = (x-1)(x-1)$ tiene la raíz 1 con multiplicidad 2). Podría ser el caso que algunas de las raíces del polinomio no fueran números reales, sino complejos. Para nosotros, una raíz de $p_A(\lambda)$ que no es un número real no será considerada como un autovalor de A.

Ejemplo 2.9. Encontrar los autovalores y autovectores de la matriz

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

Respuesta:

$$A - \lambda I = \begin{pmatrix} -\lambda & -1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & 1-\lambda \end{pmatrix}; \quad p(\lambda) = (1-\lambda) \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = (1-\lambda)(\lambda^2+1).$$

El polinomio característico tiene $\lambda = 1$ como única raíz real, luego el espectro de A es $\sigma(A) = \{1\}$. El autoespacio $\mathrm{S}(1)$ es el subespacio solución del sistema $(A - I_3)\mathbf{u} = \mathbf{0}$,es decir, las soluciones de

$$(A - I_3)\mathbf{u} = \begin{pmatrix} -1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Una vez resuelto el sistema anterior, se obtiene el subespacio vectorial generado por el vector $(0,0,1)$

$$\mathrm{S}(1) = \{(0,0,z) \, : \, z \in \mathbb{R}\} =< (0,0,1) > .$$

Obsérvese que $p_A(\lambda)$ tiene otras raíces que no son reales, los números complejos $\pm i$ que, de acuerdo a nuestra definición, no son considerados autovalores de A. Si hubiéramos admitido números complejos, entonces estas raíces también serían autovalores de A en este sentido más amplio, pero para nosotros es suficiente manejar únicamente números reales.

Ejemplo 2.10. Hallar los autovalores y autovectores de la matriz

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

RESPUESTA: Los autovalroes se obtienen al resolver

$$\begin{vmatrix} 2-\lambda & 1 & 0 \\ 0 & 1-\lambda & -1 \\ 0 & 2 & 4-\lambda \end{vmatrix} = 0.$$

Las soluciones son $\lambda = 3$ (raíz simple) y $\lambda = 2$ (raíz doble). Para encontrar $\mathrm{S}(3) = \{\mathbf{u} \in \mathbb{R}^3 : (B - 3I_3)\mathbf{u} = \mathbf{0}\}$ resolvemos el sistema

$$(B - 3I_3)\mathbf{u} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -2 & -1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

obteniendo $x = y$ y $z = -2y$, por lo que $\mathrm{S}(3) =< (1, 1, -2) >$. De la misma manera, para determinar $\mathrm{S}(2)$ resolvemos el sistema

$$(B - 2I_3)\mathbf{u} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

obteniendo $x = y = 0$ y, por tanto, $\mathrm{S}(2) =< (1, 0, 0) >$.

Ejemplo 2.11. Hallar los autovalores y autovectores de la matriz

$$C = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ 1 & 1 & 3 \end{pmatrix}$$

RESPUESTA: Para hallar los autovalores resolvemos la ecuación característica

$$\begin{aligned} 0 = |C - \lambda I_3| &= \begin{vmatrix} 1-\lambda & 2 & 0 \\ 0 & 2-\lambda & 0 \\ 1 & 1 & 0-\lambda \end{vmatrix} \\ &= \left|2-\lambda\right| \begin{vmatrix} 1-\lambda & 0 \\ 1 & 3-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda)(3-\lambda) \end{aligned}$$

De esta forma, los autovalores son $\lambda_1 = 1$, $\lambda_2 = 2$ y $\lambda_3 = 3$. El autoespacio $\mathrm{S}(1)$ es el conjunto de soluciones del sistema lineal homogéneo de matriz $C - \lambda I_3$ con $\lambda = 1$. Es decir, $\mathrm{S}(1)$ viene dado por las soluciones del sistema

$$\begin{pmatrix} 0 & 2 & 0 \\ 0 & 2 & 0 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Resolviendo el sistema encontramos

$$\mathrm{S}(1) = \{(-2z, 0, z) : z \in \mathbb{R}\} =< (-2, 0, 1) > .$$

De igual forma, $\mathrm{S}(2)$ es el conjunto de soluciones del sistema lineal homogéneo de matriz $C - \lambda I_3$ con $\lambda = 2$. Es decir, $\mathrm{S}(2)$ viene dado por las soluciones del sistema

$$\begin{pmatrix} -1 & 2 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

obteniendo

$$\mathrm{S}(2) = \{(2y, y, -3y) : y \in \mathbb{R}\} =< (2,1,-3) > .$$

Finalmente, S(3) es el conjunto de soluciones del sistema lineal homogéneo de matriz $A - \lambda I_3$ con $\lambda = 3$. Es decir, S(3) viene dado por las soluciones del sistema

$$\begin{pmatrix} -2 & 2 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

obteniendo

$$\mathrm{S}(3) = \{(0,0,z) : z \in \mathbb{R}\} =< (0,0,1) > .$$

A continuación describiremos el *proceso para diagonalizar una matriz*. Sea A una matriz cuadrada cuyo polinomio característico $p_A(\lambda)$ tiene únicamente raíces *reales*

$$\lambda_1, \lambda_2, \ldots, \lambda_k,$$

donde cada λ_k tiene multiplicidad $m_k \geq 1$ (es decir, $m_k = 1$ si λ_k es raíz simple, $m_k = 2$ si es doble, etc.). Por el Teorema Fundamental del Álgebra, $m_1 + m_2 + \cdots + m_k = n$ dado que estamos suponiendo que todas las raíces son reales.

El siguiente resultado establece que el número de vectores independientes en el subespacio propio $\mathrm{S}(\lambda)$ es menor o igual que la multiplicidad del autovalor λ como raíz del polinomio p_A. Como veremos, la dimensión de $\mathrm{S}(\lambda)$ y la multiplicidad de λ juegan un papel primordial en la digonalización de la matriz A.

Proposición 2.12. *Para cada* $j = 1, \ldots, k$

$$1 \leq \dim \mathrm{S}(\lambda_j) \leq m_j.$$

El teorema siguiente establece condiciones necesarias y suficientes para que una matriz A sea diagonalizable.

Teorema 2.13. *Una matriz A es diagonalizable si y sólo si se cumplen las dos siguiente condiciones:*

1. *Toda raíz, $\lambda_1, \lambda_2, \ldots, \lambda_k$ del polinomio característico $p_A(\lambda)$ es real.*
2. *Para cada $j = 1, \ldots, k$*

$$\dim \mathrm{S}(\lambda_j) = m_j.$$

Corolario 2.14. *Si la matriz A posee n autovalores reales y distintos, entonces es diagonalizable.*

Teorema 2.15. *Si A es diagonalizable, entonces la matriz diagonal D tiene como elementos diagonales sus autovalores, con cada uno de los λ_j repetido m_j veces. Además, la matriz P tal que $D = P^{-1}AP$ tiene como columnas vectores propios linealmente independientes seleccionados de cada uno de los subespacios propios* $\mathrm{S}(\lambda_j)$*, $j = 1, \ldots, k$.*

Comentarios sobre los ejemplos estudiados anteriormente. La matriz A del Ejemplo 2.9 no es diagonalizable, dado que p_A presenta raíces complejas. Tampoco lo es la matriz B del Ejemplo 2.10 aunque en este caso todas las raíces de p_B son reales, pero $\dim \mathrm{S}(2) = 1$, que es menor que 2, que es la multiplicidad del autovalor $\lambda = 2$. La matriz C del Ejemplo 2.11 es diagonalizable, dado que p_C admite 3 raíces reales distintas. En este caso

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \qquad P = \begin{pmatrix} -2 & 2 & 0 \\ 0 & 1 & 0 \\ 1 & -3 & 1 \end{pmatrix}.$$

Volviendo al Ejemplo 2.4, podemos calcular

$$\begin{vmatrix} \frac{1}{2} - \lambda & \frac{1}{9} \\ \frac{1}{2} & \frac{8}{9} - \lambda \end{vmatrix} = 0,$$

o $18\lambda^2 - 25\lambda + 7 = 0$, por lo que $\lambda_1 = 1$ y $\lambda_2 = \frac{7}{18}$. La matriz es diagonalizable. El subespacio S(1) es el conjunto de soluciones de

$$\begin{pmatrix} -\frac{1}{2} & \frac{1}{9} \\ \frac{1}{2} & -\frac{1}{9} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Encontramos que $y = \frac{9}{2}x$, es decir, $\mathrm{S}(1) =< (2,9) >$. De la misma forma, $\mathrm{S}(\frac{7}{18})$ queda determinado por

$$\begin{pmatrix} \frac{1}{9} & \frac{1}{9} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Encontramos que $y = -x$, es decir, $\mathrm{S}(\frac{7}{18}) =< (1,-1) >$. La matriz diagonal es

$$D = \begin{pmatrix} 1 & 0 \\ 0 & \frac{7}{18} \end{pmatrix}$$

y la matriz de paso y su inversa

$$P = \begin{pmatrix} 2 & 1 \\ 9 & -1 \end{pmatrix}, \qquad P^{-1} = \frac{1}{11} \begin{pmatrix} 1 & 1 \\ 9 & -2 \end{pmatrix}.$$

Por tanto,

$$A^n = \frac{1}{11} \begin{pmatrix} 2 & 1 \\ 9 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (\frac{7}{18})^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 9 & -2 \end{pmatrix}.$$

Haciendo los cálculos para $n = 4$ obtenemos

$$A^4 = \begin{pmatrix} 0{,}1891 & 0{,}1802 \\ 0{,}8111 & 0{,}8198 \end{pmatrix}.$$

En consecuencia

$$\begin{pmatrix} b_5 \\ m_5 \end{pmatrix} = A^4 \begin{pmatrix} b_1 \\ m_1 \end{pmatrix} = \begin{pmatrix} 0{,}1891 & 0{,}1802 \\ 0{,}8111 & 0{,}8198 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0{,}1891 \\ 0{,}8111 \end{pmatrix}.$$

Esto significa que la probabilidad de que la quinta clase del prof. X vaya bien, condicionado a que la primera clase fue también bien es de 0,1891.
Podemos preguntarnos que sucede en el largo plazo, es decir, suponiendo que el curso dura toda la eternidad (¡oh no!). Esta pregunta requiere el cálculo de un límite

$$\lim_{n\to\infty} A^n = P(\lim_{n\to\infty} D^n)P^{-1} = P \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} P^{-1} = \begin{pmatrix} \frac{2}{11} & \frac{2}{11} \\ \frac{9}{11} & \frac{9}{11} \end{pmatrix},$$

para encontrar que la *distribución estacionaria de probabilidades* es

$$\begin{pmatrix} b_\infty \\ m_\infty \end{pmatrix} = \begin{pmatrix} 0{,}1818 \\ 0{,}8182 \end{pmatrix}.$$