TEMA 1

Diagonalización de endomorfismos

A lo largo de este tema $\mathbb{K}$ denotará un cuerpo con característica distinta de 2, V un espacio vectorial sobre $\mathbb{K}$ y $\mathrm{End}(V)$ a los endomorfismos de V, es decir a las aplicaciones lineales $f : V \longrightarrow V$. Si $\dim(V) = n$ y fijamos $B = \{u_1, \ldots, u_n\}$ una base ordenada de V podemos considerar $M(f, B)$ es decir la matriz del endormorfismo f en la base B. Recordemos $M(f, B)$ es la matriz de $\mathcal{M}_n(\mathbb{K})$ cuya columna j-ésima está formada por la coordenadas de $f(u_j)$ en la base B. Esto nos permite definir una aplicación:

$$\begin{array}{rccl} \Phi_B : & \mathrm{End}(V) & \longrightarrow & \mathcal{M}_n(\mathbb{K}) \\ & f & \mapsto & M(f, B) \end{array}$$

Recíprocamente, si $A \in \mathcal{M}_n(\mathbb{K})$ tenemos el endomorfismo $f_A \in \mathrm{End}(V)$ dado por:

$$\text{para } v \in V \text{ las coordenadas de } f_A(v) \text{ en la base } B \text{ son } A \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

donde $(a_1, \ldots, a_n)$ son las coordenadas de v en la base B. Equivalentemente f_A es el endomorfismo de V que verifica $M(f_A, B) = A$. De aquí podemos definir la aplicación

$$\begin{array}{rccl} \Psi_B : & \mathcal{M}_n(\mathbb{K}) & \longrightarrow & \mathrm{End}(V) \\ & A & \mapsto & f_A \end{array}$$

EJERCICIO 1.1: Es fácil comprobar que Φ_B y Ψ_B son aplicaciones lineales tal que $\Phi_B^{-1} = \Psi_B$, con lo que son isomorfismos.

Por otra parte sabemos cómo se relacionan las matrices que representan a un mismo endomorfismo en distintas bases. Si $B' = \{u'_1, \ldots, u'_n\}$ es otra base ordenada de V entonces:

$$\begin{aligned} M(f, B') &= M(Id_V, B, B') \cdot M(f, B) \cdot M(Id_V, B', B) \\ &= M(Id_V, B, B') \cdot M(f, B) \cdot M(Id_V, B, B')^{-1} \end{aligned} \tag{1}$$

Es decir, las matrices que representan a un mismo endomorfismo en distintas bases son SEMEJANTES. En particular tienen la misma traza, el mismo determinante y el mismo rango. Esto nos permite definir la traza, el determinante y el rango de un endomorfismo f como la traza, determinante o rango de $M(f, B)$ para B cualquier base ordenada de V.

A la vista de lo anterior sería conveniente elegir una base B para la que $M(f, B)$ fuese "lo más sencilla posible". Quizás las matrices cuadradas más sencillas son las matrices diagonales. Dado $f \in \mathrm{End}(V)$ surge pues una pregunta natural:

¿Existe una base B de V tal que $M(f, B)$ es una matriz diagonal?

De (1) tenemos que $M(f,B)$ es diagonal si y solo si $M(f,B')$ es una matriz semejante a una matriz diagonal para cualquier base B'. Por tanto la pregunta anterior es equivalente a la siguiente pregunta sobre matrices:

¿Toda matriz cuadrada $A \in \mathcal{M}_n(\mathbb{K})$ es semejante a una matriz diagonal?

Vamos a darle un nombre a los endomorfismos y matrices que responden afirmativamente a las preguntas anteriores.

Definición 1.2: Diremos que un endomorfismo $f \in \operatorname{End}(V)$ es *diagonalizable* si existe B base de V tal que $M(f,B)$ es diagonal. Al proceso de encontrar tal base se le denomina *diagonalización del endomorfismo.*

Definición 1.3: Asimismo, diremos que una matriz cuadrada $A \in \mathcal{M}_n(\mathbb{K})$ es *diagonalizable* si es semejante a una matriz diagonal, es decir si existe una matriz regular P tal que $P \cdot A \cdot P^{-1} = D$ es una matriz diagonal. Al procedimiento de encontrar la matriz P se le denomina *diagonalización de la matriz A.*

Observación 1.4: Sea B una base de V. Por lo dicho anteriormente si f es un endomorfismo diagonalizable entonces para cualquier B base de V tenemos que $M(f,B)$ es una matriz diagonalizable. Recíprocamente, si $A \in \mathcal{M}_n(\mathbb{K})$ es una matriz diagonalizable entonces $f_A = \Psi_B(A)$ es un endomorfismo diagonalizable.

Observación 1.5: Observemos que siempre que hablamos de diagonalización nos referimos a $M(f,B)$, es decir tomamos siempre la misma base en el dominio y el codominio. Notar que el problema si consideramos bases distintas de V no tiene sentido ya que siempre existen B y B' bases de V tales que $M(f,B,B')$ es diagonal con unos y ceros en la diagonal.

Para probar esto consideremos una base $\{u_1,\dots,u_k\}$ del $\operatorname{Ker}(f)$ y completémosla hasta obtener $B = \{u_1,\dots,u_k,u_{k+1},\dots,u_n\}$ una base de V . Tenemos que $\{f(u_{k+1}),\dots,f(u_n)\}$ es una base de Im(f). Basta ahora completar esta base hasta tener

$$B' = \{u'_1,\dots,u'_k, f(u_{k+1}),\dots,f(u_n)\}$$

otra base de V. Es claro que:

$$M(f,B,B') = \begin{pmatrix} 0 & & & & & \\ & \underbrace{\ddots}_{k} & & & & \\ & & 0 & & & \\ & & & 1 & & \\ & & & & \underbrace{\ddots}_{n-k} & \\ & & & & & 1 \end{pmatrix} = \left(\begin{array}{c|c} 0_k & 0_{k\times(n-k)} \\ \hline 0_{(n-k)\times k} & I_{n-k} \end{array}\right)$$

Observación 1.6: Los ejemplos más sencillos de endomorfismos diagonalizables son las homotecias, es decir las aplicaciones $h_\lambda = \lambda Id_V$, $\lambda \in \mathbb{K}$. Es obvio que la $M(h_\lambda, B) = \lambda I_n$ para toda base B de V y por tanto son endomorfismos diagonalizables.

Otros ejemplos sencillos de endomorfismos diagonalizables son las simetrías y las proyecciones.

Observación 1.7: Notemos que si f fuese diagonalizable, es decir si existiese $B = \{u_1, \ldots, u_n\}$ base de V tal que $M(f, B)$ fuese diagonal tendríamos

$$M(f,B) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Es decir

(2) $$f(u_i) = \lambda_i u_i \ , 1 \leqslant i \leqslant n$$

Observación 1.8: Es fácil encontrar endomorfismos no diagonalizables. Por ejemplo si consideramos los giros de ángulo $\theta \in [0, 2\pi)$ en $\mathbb{R}^2$, es decir las aplicaciones lineales $f_\theta : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dadas por

$$M(f_\theta, B_u) = \begin{pmatrix} \cos(\theta) & -\operatorname{sen}(\theta) \\ \operatorname{sen}(\theta) & \cos(\theta) \end{pmatrix}$$

Observemos que si $\theta = 0$ entonces $f_0 = Id_{\mathbb{R}^2}$ y si $\theta = \pi$ entonces $f_\pi = -Id_{\mathbb{R}^2}$. Si $\theta \notin \{0, \pi\}$ entonces es fácil ver que no existe una base B de $\mathbb{R}^2$ de forma que $M(f_\theta, B)$ sea diagonal.

Efectivamente, como hemos visto en la observación 1.7 para que f_θ sea diagonalizable es necesario que exista una base $\{u_1, u_2\}$ de $\mathbb{R}^2$ y $\lambda_1, \lambda_2 \in \mathbb{R}$ que verifiquen (2). Vamos a ver que no puede existir ningún vector en una base verificando esa igualdad. Supongamos por el contrario que existiese $(x, y) \in \mathbb{R}^2 - \{(0,0)\}$ y $\lambda \in \mathbb{R}$ tal que $f_\theta(x, y) = \lambda(x, y)$. Esto se traduciría en la siguiente igualdad:

$$\begin{pmatrix} \cos(\theta) & -\operatorname{sen}(\theta) \\ \operatorname{sen}(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$$

De aquí

$$\begin{pmatrix} \cos(\theta) - \lambda & -\operatorname{sen}(\theta) \\ \operatorname{sen}(\theta) & \cos(\theta) - \lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Es decir, $\{x, y\}$ serían una solución no trivial del sistema lineal homogéneo cuya matriz es:

$$\begin{pmatrix} \cos(\theta) - \lambda & -\operatorname{sen}(\theta) \\ \operatorname{sen}(\theta) & \cos(\theta) - \lambda \end{pmatrix}$$

La única forma de que esto suceda es que el sistema sea compatible indeterminado y por tanto el determinante de dicha matriz debe ser nulo. El determinante de esta matriz es

$$(\cos(\theta) - \lambda)^2 + \operatorname{sen}(\theta)^2 = \lambda^2 - 2\lambda\cos(\theta) + \operatorname{sen}(\theta)^2 + \cos(\theta)^2 = \lambda^2 - 2\lambda\cos(\theta) + 1$$

Pero el discriminante de este polinomio de grado dos en λ es $4\cos(\theta)^2 - 4 = 4(\cos(\theta)^2 - 1)$ que es siempre negativo para $\theta \in (0, \pi) \cup (\pi, 2\pi)$ y por tanto no hay ningún λ para el que el determinante es cero y así no puede existir el vector (x, y) cumpliendo las condiciones impuestas al comienzo de esta observación.

Como hemos visto, hay endomorfismos que son diagonalizables y otros que no. Intentaremos en este tema descubrir las propiedades que caractericen a los endomorfismos diagonalizables.

1. Valores y vectores propios. Subespacios propios.

Hemos visto en la introducción que el problema de diagonalizar endomorfismos pasa por encontrar una base $B = \{u_1, \ldots, u_n\}$ cuyos vectores cumplan (2) para ciertos $\lambda_i \in \mathbb{K}$. Vamos a darle nombre a estos vectores y a los escalares λ_i.

Definición 1.9: Diremos que $\lambda \in \mathbb{K}$ es un *valor propio* o *autovalor* de $f \in \mathrm{End(V)}$ si existe $v \in V - \{\mathbf{0}\}$ tal que

$$f(v) = \lambda v$$

Diremos entonces que v es un *vector propio* o *autovector* de f asociado al valor propio λ.

Observación 1.10: Notar que los conceptos de vector y valor propio son generales. No necesitamos que V sea finitamente generado para las definiciones anteriores.

Ejercicio 1.11: Un vector $v \neq \mathbf{0}$ solo puede ser vector propio asociado a un valor propio.

Ejemplos 1.12:
1. Observemos que si $f \in \mathrm{End(V)}$ y $v \in \mathrm{Ker}(f) - \{\mathbf{0}\}$ entonces v es un vector propio asociado al valor propio $\lambda = 0$.
2. Si $V = U \oplus W$ y π es la proyección sobre U paralela a W entonces todo vector $u \in U - \{\mathbf{0}\}$ es un vector propio asociado al valor propio $\lambda = 1$ y todo vector $w \in W - \{\mathbf{0}\}$ es un vector propio asociado al valor propio $\lambda = 0$.
3. Análogamente, si $V = U \oplus W$ y σ es la simetría sobre U paralela a W entonces todo vector $u \in U - \{\mathbf{0}\}$ es un vector propio asociado al valor propio $\lambda = 1$ y todo vector $w \in W - \{\mathbf{0}\}$ es un vector propio asociado al valor propio $\lambda = -1$.
4. Si h_λ es la homotecia de razón $\lambda \in \mathbb{K}$ entonces cualquier $v \in V - \{\mathbf{0}\}$ es un vector propio asociado al valor propio λ.

Ejercicio 1.13: Si f es un endomorfismo de V tal que todo vector $v \in V - \{\mathbf{0}\}$ es un vector propio entonces f debe ser una homotecia. (Ayuda: Comienza probando que todos los vectores deben ser vectores propios asociados al mismo valor propio).

Observemos que dada B una base de V, v es un vector propio de f asociado al valor propio λ si y solo si

$$M(f, B) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \lambda \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

donde $(a_1, \ldots, a_n)$ son las coordenadas de v en la base B. Por tanto tiene sentido la siguiente definición para matrices:

Definición 1.14: Diremos que $\lambda \in \mathbb{K}$ es un *valor propio* o *autovalor* de $A \in \mathcal{M}_n(\mathbb{K})$ si existe $x \in \mathbb{K}^n - \{\mathbf{0}\}$ tal que

$$A \cdot x = \lambda x$$

Diremos entonces que x es un *vector propio* o *autovector* de A asociado al valor propio λ.

Notemos que con esta definición si B es una base de V tenemos que v es un vector propio de f asociado al valor propio λ si y sólo si las coordenadas de v es la base B son un vector propio de $M(f, B)$ asociado a ese valor propio. Análogamente, $x \in \mathbb{K}^n - \{\mathbf{0}\}$ es un vector propio de una matriz $A \in \mathcal{M}_n(\mathbb{K})$ asociado a un valor propio λ si y sólo si x es un vector propio del endomorfismo $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ dado por $f_A(x) = A \cdot x$ asociado al valor propio λ.

Con las definiciones anteriores tenemos que:

$$\boxed{f \in \text{End(V) es diagonalizable} \iff \text{Existe una base de } V \text{ formada por vectores propios de } f}$$

Observación 1.15: Observemos que la caracterización anterior puede darse como definición de diagonalización para espacios vectoriales en general, no hace falta que sean finitamente generados.

Análogamente tenemos

$$\boxed{A \in \mathcal{M}_n(\mathbb{K}) \text{ es diagonalizable} \iff \text{Existe una base de } \mathbb{K}^n \text{ formada por vectores propios de } A}$$

A continuación trataremos de encontrar la manera de calcular los valores y vectores propios de un endomorfismo. Sabemos que si $v \neq \mathbf{0}$ es un vector propio de f asociado al valor propio λ entonces

$$f(v) = \lambda v = \lambda Id_v(v) \iff (f - \lambda Id_v)(v) = \mathbf{0} \iff v \in \text{Ker}(f - \lambda Id_v)$$

Luego los vectores propios asociados al valor propio λ son los vectores no nulos del subespacio $\text{Ker}(f - \lambda Id_v)$. Vamos a darle un nombre a este subespacio vectorial.

Definición 1.16: Dado λ valor propio de f denominaremos *subespacio propio de f* asociado al valor propio λ al subespacio

$$V_\lambda = \text{Ker}(f - \lambda Id_v) = \{v \in V \mid f(v) = \lambda v\}$$

Definición 1.17: Llamaremos *multiplicidad geométrica de* λ a $g_\lambda = \dim(V_\lambda)$.

Observemos que si B es una base de V entonces

$$\begin{aligned} g_\lambda &= \dim(\text{Ker}(f - \lambda Id_v)) = n - \dim(\text{Im}(f - \lambda Id_v)) = n - \text{rg}(f - \lambda Id_v) = \\ &= n - \text{rg}(M(f - \lambda Id_v, B)) = n - \text{rg}(M(f, B) - \lambda I_n) \end{aligned}$$

Ejemplos 1.18: 1. Para todo $f \in \text{End}(V)$ se tiene $V_0 = \text{Ker}(f)$.
2. Para la homotecia h_λ tenemos $V = V_\lambda$ y $g_\lambda = \dim(V)$.
3. Sea $V = U \oplus W$. Para la proyección sobre U paralela a W tenemos $V_1 = U$ y $V_0 = W$. De aquí $g_1 = \dim(U)$ y $g_0 = \dim(W)$. Análogamente. Para la simetría sobre U paralela a W tenemos $V_1 = U$ y $V_{-1} = W$. De aquí $g_1 = \dim(U)$ y $g_{-1} = \dim(W)$.

De la misma forma si $A \in \mathcal{M}_n(\mathbb{K})$ y $x \neq \mathbf{0}$ es un vector propio de A asociado al valor propio λ entonces

$$A \cdot x = \lambda x = \lambda I_n \cdot x \iff (A - \lambda I_n) \cdot x = \mathbf{0}$$

Luego los vectores propios de A asociados al valor propio λ son las soluciones no triviales del SEL homogéneo con matriz de coeficientes $A-\lambda I_n$. El conjunto de soluciones de un SEL homogéneo sabemos que es un subespacio vectorial de $\mathbb{K}^n$.

DEFINICIÓN 1.19: Dado λ valor propio de A denominaremos *subespacio propio de A* asociado al valor propio λ al subespacio

$$V_\lambda = \{x \in \mathbb{K}^n \mid (A-\lambda I_n)\cdot x = \mathbf{0}\}$$

Como en el caso anterior llamaremos *multiplicidad geométrica de* λ y la denotaremos g_λ a la dimensión de este subespacio vectorial de $\mathbb{K}^n$.

Podemos calcular la multiplicidad geométrica de un valor propio sin necesidad de calcular el subespacio propio ya que

$$g_\lambda = n - \operatorname{rg}(A-\lambda I_n)$$

PROPOSICIÓN 1.20: *Vectores propios asociados a valores propios distintos son linealmente independientes.*

Demostración: Sean $\{v_1,\ldots,v_m\}$ vectores propios de un endomorfismo f asociados a valores propios distintos $\{\lambda_1,\ldots,\lambda_m\}$. Razonaremos por inducción sobre m.

Si $m=1$ como $v_1 \neq \mathbf{0}$ tenemos que es linealmente independiente. Supongamos que el enunciado de la proposición es cierto para $m-1$ y veamos que es cierto para m. Para ello consideremos una combinación lineal

$$\alpha_1 v_1 + \cdots + \alpha_{m-1}v_{m-1} + \alpha_m v_m = \mathbf{0} \tag{3}$$

con $\alpha_1,\ldots,\alpha_m \in \mathbb{K}$. Si $\alpha_m = 0$ ya habríamos acabado. Podemos asumir por tanto que $\alpha_m \neq 0$

De lo anterior tenemos

$$f(\alpha_1 v_1 + \cdots + \alpha_{m-1}v_{m-1} + \alpha_m v_m) = f(\mathbf{0}) = \mathbf{0}$$

Pero como f es una aplicación lineal podemos escribir

$$\alpha_1 f(v_1) + \cdots + \alpha_{m-1}f(v_{m-1}) + \alpha_m f(v_m) = \mathbf{0}$$

Utilizando ahora que el vector v_i es un vector propio de f asociado al valor propio λ_i tenemos

$$\alpha_1\lambda_1 v_1 + \cdots + \alpha_{m-1}\lambda_{m-1}v_{m-1} + \alpha_m\lambda_m v_m = \mathbf{0} \tag{4}$$

Si consideramos la ecuación (4) menos la ecuación (3) multiplicada por λ_m obtenemos

$$\alpha_1(\lambda_1-\lambda_m)v_1 + \cdots + \alpha_{m-1}(\lambda_{m-1}-\lambda_m)v_{m-1} = \mathbf{0}$$

Como por la hipótesis de inducción $\{v_1,\ldots,v_{m-1}\}$ son linealmente independientes tenemos que

$$\alpha_1(\lambda_1-\lambda_m) = \cdots = \alpha_{m-1}(\lambda_{m-1}-\lambda_m) = 0$$

Como $\lambda_i \neq \lambda_m$ para $i=1,\ldots,m-1$ tenemos que $\alpha_i = 0$ para $i=1,\ldots,m-1$. Por tanto la ecuación (3) quedaría como

$$\alpha_m v_m = \mathbf{0}$$

Teniendo en cuenta que $v_m \neq \mathbf{0}$ por ser un vector propio concluimos que $\alpha_m = 0$. Esto unido a lo obtenido anteriormente nos permite afirmar que los vectores $\{v_1,\ldots,v_m\}$ son linealmente independientes. □

Corolario 1.21: *Sea V un espacio vectorial,* $\dim(V) = n$ *y* $f \in \mathrm{End}(V)$ *entonces el número de valores propios diferentes de f es menor o igual que n. Si hay exactamente n valores propios distintos entonces el endomorfismo es diagonalizable.*

Demostración: Sabemos que si $\{\lambda_1, \ldots, \lambda_m\}$ son valores propios distintos entonces existen $\{v_1, \ldots, v_m\}$ vectores propios de f asociados a ellos que por la Proposición 1.20 son linealmente independientes. Por tanto $m \leqslant n$. Si $m = n$ entonces $B = \{v_1, \ldots, v_n\}$ es una base de V formada por vectores propios y por tanto f es diagonalizable. □

Corolario 1.22: *Sea V un espacio vectorial y $f \in \mathrm{End}(V)$ entonces la suma de subespacios propios de f es directa, es decir si $\lambda_1, \ldots, \lambda_m$ son valores propios distintos de f entonces*

$$V_{\lambda_1} + \cdots + V_{\lambda_m} = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_m}$$

En particular $\dim(V_{\lambda_1} + \cdots + V_{\lambda_m}) = g_{\lambda_1} + \cdots + g_{\lambda_m}$.

Demostración: Debemos comprobar que para todo $j \in \{1, \ldots, m-1\}$ se tiene

$$(V_{\lambda_1} + \cdots + V_{\lambda_j}) \cap V_{\lambda_{j+1}} = \{\mathbf{0}\}$$

Supongamos que exista $v \in (V_{\lambda_1} + \cdots + V_{\lambda_j}) \cap V_{\lambda_{j+1}}$, $v \neq \mathbf{0}$. Entonces

$$v = v_1 + \cdots v_j \in V_{\lambda_{j+1}}$$

donde $v_i \in V_{\lambda_i}$. De la igualdad anterior obtenemos

$$v_1 + \cdots + v_j - v = \mathbf{0}$$

Es decir tenemos una combinación lineal de vectores asociados a valores propios distintos con coeficientes no nulos (observemos que al menos un v_i es no nulo). Esto contradice la Proposición 1.20. Por tanto $v = \mathbf{0}$. □

Observación 1.23: Notemos que la tesis del Corolario 1.22 no implica que

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_m}\ .$$

Observemos además que si f tiene n valores propios distintos, de los Corolarios 1.21 y 1.22 se tiene

$$n \leqslant \dim(V_{\lambda_1} + \cdots + V_{\lambda_n}) = g_{\lambda_1} + \cdots + g_{\lambda_n} \leqslant n$$

y por tanto $g_{\lambda_i} = 1$ para todo $i \in \{1, \ldots, n\}$.

2. Polinomio característico. Multiplicidad geométrica y algebraica.

Hemos visto en la sección anterior cómo calcular los vectores propios de un endomorfismo asociados a un vector propio. Pero todavía no hemos visto cómo calcular los valores propios del endomorfismo. Este va a ser nuestro objetivo inmediato. Dado $f \in \mathrm{End}(V)$ sabemos que λ es un valor propio de f si existe $v \neq \mathbf{0}$ vector propio de f asociado al valor propio λ. Es decir:

$$\lambda \text{ valor propio de } f \iff V_\lambda = \mathrm{Ker}(f - \lambda Id_V) \neq \{\mathbf{0}\} \iff f - \lambda Id_v \text{ no es un isomorfismo} \iff$$

$$\iff \det(f - \lambda Id_V) = 0$$

Fijémonos que lo anterior nos proporciona una forma de calcular los valores propios de un endomorfismo.

Definición 1.24: Si $f \in \operatorname{End}(V)$, llamaremos *polinomio característico* de f al polinomio
$$p_f(\lambda) = \det(f - \lambda Id_V) \ .$$
A la ecuación $p_f(\lambda) = 0$ se le denomina *ecuación característica* de f.

De esta definición tenemos que los valores propios de f son las soluciones de la ecuación característica de f o equivalentemente, las raíces del polinomio característico. Podemos por tanto dar la siguiente definición:

Definición 1.25: Dado λ valor propio de f llamaremos *multiplicidad algebraica* de λ, y la denotaremos a_λ, a la multiplicidad que tiene λ como raíz del polinomio característico.

Análogamente si $A \in \mathcal{M}_n(\mathbb{K})$ tenemos que λ es un valor propio de A si existe $x \in \mathbb{K}^n - \{\mathbf{0}\}$ tal que x es un vector propio asociado al valor propio λ. Esto es:
$$\lambda \text{ valor propio de } A \Longleftrightarrow V_\lambda = \{x \in \mathbb{K}^n \mid (A - \lambda I_n) \cdot x = \mathbf{0}\} \neq \{\mathbf{0}\} \Longleftrightarrow$$
$$\Longleftrightarrow (A - \lambda I_n) \cdot x = \mathbf{0} \text{ es un S.E.L. compatible indeterminado} \Longleftrightarrow \det(A - \lambda I_n) = 0$$
Es natural entonces definir:

Definición 1.26: Si $A \in \mathcal{M}_n(\mathbb{K})$, llamaremos *polinomio característico* de A al polinomio
$$p_A(\lambda) = \det(A - \lambda I_n) \ .$$
A la ecuación $p_A(\lambda) = 0$ se le denomina *ecuación característica* de A. Además dado λ valor propio de A llamaremos *multiplicidad algebraica* de λ, y la denotaremos a_λ, a la multiplicidad que tiene λ como raíz del polinomio característico.

Sabemos que para calcular el determinante de un endomorfismo basta calcular el determinante de la matriz del endormorfismo en cualquier base. Por tanto si $f \in \operatorname{End}(V)$ y B es una base de V tenemos
$$p_f(\lambda) = \det(f - \lambda Id_V) = \det(M(f - \lambda Id_V, B) = \det(M(f, B) - \lambda I_n) = p_{M(f,B)}(\lambda) \ .$$
Análogamente, si $A \in \mathcal{M}_n(\mathbb{K})$ tenemos
$$p_A(\lambda) = p_{f_A}(\lambda)$$
donde $f_A : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ dada por $f_A(x) = A \cdot x$.

Ejercicio 1.27: Si dos matrices son semejantes entonces tienen el mismo polinomio característico. El recíproco no es cierto en general.

A continuación vamos a estudiar cómo se relacionan las multiplicidades geométrica y algebraica de un valor propio.

Proposición 1.28: *Sea $f \in \operatorname{End}(V)$ y μ un valor propio de f, entonces tenemos $g_\mu \leqslant a_\mu$.*

Demostración: Sea μ un valor propio de f y $\{u_1, \ldots, u_{g_\mu}\}$ una base de V_μ. Utilizando el Teorema de ampliación de la base podemos ampliar esta base hasta obtener

$$B = \{u_1, \ldots, u_{g_\mu}, u_{g_\mu+1}, \ldots, u_n\}$$

Si calculamos la matriz de f en esta base obtenemos

$$M(f, B) = \left(\begin{array}{c|c} \mu I_{g_\mu} & Q \\ \hline 0_{(n-g_\mu)\times g_\mu} & T \end{array}\right)$$

donde Q es una matriz $g_\mu \times (n - g_\mu)$ y T es una matriz cuadrada de orden $n - g_\mu$. Por tanto el polinomio característico de f vendrá dado por

$$\begin{aligned} p_f(\lambda) &= \det(M(f, B) - \lambda I_n) = \det \left(\begin{array}{c|c} (\mu-\lambda) I_{g_\mu} & Q \\ \hline 0_{(n-g_\mu)\times g_\mu} & T - \lambda I_{n-g_\mu} \end{array}\right) = \\ &= \det((\mu - \lambda)I_{g_\mu}) \det(T - \lambda I_{n-g_\mu}) = (\mu - \lambda)^{g_\mu} \det(T - \lambda I_{n-g_\mu}) \end{aligned}$$

Por tanto $a_\mu \geqslant g_\mu$. □

Ejemplo 1.29: Calcula los valores propios y los subespacios propios de los endomorfismos $f, g : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dados por

$$f(x, y) = (3x + 4y, -2x - 3y) \ , \quad g(x, y) = (-x + 2y, -y) \ .$$

A continuación vamos a analizar un poco más el polinomio característico. Recordemos que si $f \in \text{End}(V)$ y B es una base de V entonces

$$p_f(\lambda) = \det(f - \lambda Id_V) = \det(M(f, B) - \lambda I_n) \ .$$

Empezaremos con los casos de dimensión más pequeña. Supongamos $n = 2$ y

$$M(f, B) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Entonces

$$\begin{aligned} p_f(\lambda) &= \det\left(\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} = \\ &= (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} = \\ &= \lambda^2 - \text{traza}(f)\lambda + \det(f) \ . \end{aligned}$$

Análogamente, si n=3 tendríamos

$$M(f, B) = A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

El polinomio característico de f vendría dado por:

$$\begin{aligned} p_f(\lambda) &= \det\left(\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{pmatrix} = \\ &= -\lambda^3 + (a_{11} + a_{22} + a_{33})\lambda^2 - (\det\left(\begin{smallmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{smallmatrix}\right) + \det\left(\begin{smallmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{smallmatrix}\right) + \det\left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right))\lambda + \det(A) = \\ &= -\lambda^3 + \text{traza}(f)\lambda^2 - (A_{11} + A_{22} + A_{33})\lambda + \det(f) \ . \end{aligned}$$

donde A_{ij} denota el determinante de la matriz que queda cuando eliminamos la fila i-ésima y la columna j-ésima de A. En general si $\dim(V) = n$ y $M(f, B) = (a_{ij})$ tenemos

$$p_f(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}\mathrm{traza}(f)\lambda^{n-1} + \cdots + (-1)^i\alpha_i\lambda^i + \cdots + \det(f)$$

donde el coeficiente α_i es la suma de los menores de orden $n-i$ que tienen la diagonal sobre la diagonal principal de A. Vamos a conformarnos con probar que los términos en λ^n, λ^{n-1} y el término independiente de $p_f(\lambda)$ son los indicados en la expresión anterior. Recordemos que

$$p_f(\lambda) = \det(A - \lambda I_n) = \det\begin{pmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{pmatrix}$$

Observemos que los términos en λ^n y λ^{n-1} se obtienen del producto de los elementos de la diagonal de la matriz anterior y por tanto hay que buscarlos en:

$$(a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda) = (-1)^n\lambda^n + (-1)^{n-1}(a_{11} + \cdots + a_{nn})\lambda^{n-1} + q(\lambda)$$

donde $q(\lambda)$ es un polinomio de grado menor o igual que $n-2$ en λ. Por tanto los coeficientes de $p_f(\lambda)$ son $(-1)^n$ y $(-1)^{n-1}\mathrm{traza}(f)$, respectivamente.

Por otra parte notamos el término independiente de un polinomio viene dado por

$$p_f(0) = \det(A - 0I_n) = \det(A) = \det(f)\ .$$

3. El teorema fundamental de diagonalización.

En esta sección vamos a utilizar todo lo visto anteriormente para dar varias caracterizaciones de los endomorfismos diagonalizables.

Para ello observemos que si $f \in \mathrm{End}(V)$ en un endomorfismo diagonalizable entonces existen $\{\lambda_1, \ldots, \lambda_k\}$ valores propios distintos y $B = \{u_1, \ldots, u_n\}$ base de V formada por vectores propios, equivalentemente:

$$M(f, B) = \left(\begin{array}{c|c|c|c} \lambda_1 I_{q_1} & 0 & \cdots & 0 \\ \hline 0 & \lambda_2 I_{q_2} & \cdots & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & \lambda_k I_{q_k} \end{array}\right).$$

De aquí:

$$p_f(\lambda) = (\lambda_1 - \lambda)^{q_1}(\lambda_2 - \lambda)^{q_2}\cdots(\lambda_k - \lambda)^{q_k}$$

Observemos que entonces $a_{\lambda_i} = q_i$ para $i = 1, \ldots, k$ y por tanto

(5) $$\boxed{n = \sum_{i=1}^{k} a_{\lambda_i}}$$

Por otra parte la matriz de $f - \lambda_i Id_V$ en la base B tiene exactamente p_i ceros en la diagonal y por tanto

$$g_{\lambda_i} = n - \mathrm{rg}(f - \lambda_i Id_V) = n - (n - p_i) = p_i$$

Por tanto para $i = 1, \ldots, k$ tenemos

(6) $$\boxed{g_{\lambda_i} = a_{\lambda_i}}$$

Concluimos entonces que (5) y (6) son condiciones necesarias para que f sea diagonalizable. Veremos ahora que también son suficientes.

Teorema 3.1: Teorema fundamental de diagonalización de endomorfismos

Sea $f \in \mathrm{End}(V)$. Entonces las siguientes afirmaciones equivalen:

i) f es diagonalizable
ii) Existe B base de V formada por vectores propios de f.
iii) Si $\lambda_1, \ldots, \lambda_k$ son los distintos valores propios de f se tiene:

$$\boxed{\begin{array}{l} n = \sum_{i=1}^{k} a_{\lambda_i} \ , \\ g_{\lambda_i} = a_{\lambda_i} \ , \qquad i = \{1, \ldots, k\} \end{array}}$$

iv) Si $\lambda_1, \ldots, \lambda_k$ son los distintos valores propios de f entonces

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k} \ ,$$

donde V_{λ_i} es el subespacio propio asociado al valor propio λ_i, $i = \{1, \ldots, k\}$.

Demostración: Ya sabíamos que $\boxed{\text{i)} \iff \text{ii)}}$. Además el razonamiento anterior al teorema prueba que $\boxed{\text{ii)} \Rightarrow \text{iii)}}$.

Veamos ahora que $\boxed{\text{iii)} \Rightarrow \text{iv)}}$. Por una parte tenemos

$$V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k} \leqslant V \ .$$

Por otra parte tenemos que

$$\dim(V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}) = \sum_{i=1}^{k} \dim(V_{\lambda_i}) = \sum_{i=1}^{k} g_{\lambda_i} \underset{\text{iii)}}{=} \sum_{i=1}^{k} a_{\lambda_i} \underset{\text{iii)}}{=} n \ .$$

Por tanto $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k} = V$.

Finalmente probemos que $\boxed{\text{iv)} \Rightarrow \text{ii)}}$.

Consideremos B_i una base de los subespacios V_{λ_i} para $i = 1, \ldots, k$. De iv) tenemos que $B = B_1 \cup \ldots B_k$ es una base de V. Como esa base está formada por vectores propios tenemos ii). □

De forma análoga tendríamos el siguiente teorema para matrices.

Teorema 3.2: Teorema fundamental de diagonalización de matrices

Sea $A \in \mathcal{M}_n(\mathbb{K})$. Entonces las siguientes afirmaciones equivalen:

i) f es diagonalizable

ii) Existe B base de $\mathbb{K}^n$ formada por vectores propios de A.

iii) Si $\lambda_1, \ldots, \lambda_k$ son los distintos valores propios de A se tiene:

$$\boxed{\begin{aligned} n &= \sum_{i=1}^{k} a_{\lambda_i} \ , \\ g_{\lambda_i} &= a_{\lambda_i} \ , \qquad i = \{1, \ldots, k\} \end{aligned}}$$

iv) Si $\lambda_1, \ldots, \lambda_k$ son los distintos valores propios de A entonces

$$\mathbb{K}^n = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k} \ ,$$

donde V_{λ_i} es el subespacio propio asociado al valor propio λ_i, $i = \{1, \ldots, k\}$.

De este teorema podemos deducir el siguiente resultado.

COROLARIO 1.30: *Sean $A, A' \in \mathcal{M}_n(\mathbb{K})$ matrices diagonalizables. Entonces A y A' son semejantes si y solo si $p_A(\lambda) = p_{A'}(\lambda)$, equivalentemente tienen los mismos valores propios con las mismas multiplicidades.*

Demostración: La implicación $\boxed{\Rightarrow}$ es el Ejercicio 1.27. Veamos ahora $\boxed{\Leftarrow}$. Observemos que si A y A' son diagonalizables y tienen los mismos valores propios con las mismas multiplicidades entonces son semejantes a la misma matriz diagonal y por tanto son semejantes entre sí. □

4. El teorema de Cayley-Hamilton.

El Teorema de Cayley-Hamilton fue probado en primer lugar en 1853 por Hamilton para ciertos casos de matrices cuadradas reales de orden 4. Cayley en 1858 lo demostró para matrices cuadradas de orden menor o igual que 3, pero solo lo publicó para matrices de orden 2. El caso general fue probado por Frobenius en 1878.

Teorema 4.1: Teorema de Cayley-Hamilton

ea $A \in \mathcal{M}_n(\mathbb{K})$ y $p_A(\lambda) = (-1)^n\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$ su polinomio cacacterístico con $c_i \in \mathbb{K}, i = 0, 1, \ldots, n-1$. Entonces

$$(-1)^n A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I_n = \mathbf{0}_n \ ,$$

donde I_n es la matriz identidad de orden n y $\mathbf{0}_n$ es la matriz nula de orden n.

El teorema de Cayley-Hamilton nos dice por tanto que cada matriz cuadrada satisface su ecuación característica.

Tendríamos también la correspondiente versión para endomorfismos que diría:

Teorema 4.2: Teorema de Cayley-Hamilton para endomorfismos

ea V un espacio vectorial sobre $\mathbb{K}$, $dim(V) = n$, $f \in \mathrm{End}_{\mathbb{K}}(V)$ y $p_f(\lambda) = (-1)^n\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$ su polinomio cacacterístico con $c_i \in \mathbb{K}, i = 0, 1, \ldots, n-1$. Entonces

$$(-1)^n f^n + c_{n-1}f^{n-1} + \cdots + c_1 f + c_0 Id_V = f_0$$

donde Id_V es el endomorfismo identidad y f_0 es el endomorfismo idénticamente nulo.

Demostración del Teorema de Cayley-Hamilton: Comenzaremos recordando un resultado de Geometría I que afirma:

Para toda matriz $M \in \mathcal{M}_n(\mathbb{K})$ se tiene

$$M \cdot \mathrm{Adj}(M)^t = \det(M)I_n\ ,$$

donde $\mathrm{Adj}(M)$ es la matriz adjunta de M es decir la matriz cuya entrada (i, j) viene dada por $(-1)^{i+j}\det(M_{ij})$.

Si aplicamos este resultado a la matriz $M = A - \lambda I_n$ obtendremos:

$$(A - \lambda I_n) \cdot \mathrm{Adj}(A - \lambda I_n)^t = \det(A - \lambda I_n)I_n$$

Teniendo en cuenta entonces que $p_A(\lambda) = \det(A - \lambda I_n)$ la expresión anterior queda

$$(A - \lambda I_n) \cdot \mathrm{Adj}(A - \lambda I_n)^t = p_A(\lambda)I_n$$

Observemos que $\mathrm{Adj}(A - \lambda I_n)^t$ es una matriz cuyas entradas son, salvo el signo, menores de orden $n - 1$ de la matriz $A - \lambda I_n$. Por tanto son polinomios en λ de grado a lo sumo $n - 1$. De aquí podemos escribir

$$\mathrm{Adj}(A - \lambda I_n)^t = B_{n-1}\lambda^{n-1} + B_{n-2}\lambda^{n-2} + \cdots B_2\lambda^2 + B_1\lambda + B_0$$

donde cada B_i es una matriz de orden n con entradas en $\mathbb{K}$ para $i = 0, \ldots n - 1$.

Entonces tenemos

$$\begin{aligned}
& (A - \lambda I_n) \cdot \mathrm{Adj}(A - \lambda I_n)^t \\
= \ & (A - \lambda I_n) \cdot (B_{n-1}\lambda^{n-1} + B_{n-2}\lambda^{n-2} + \cdots B_2\lambda^2 + B_1\lambda + B_0) \\
= \ & -B_{n-1}\lambda^n + (AB_{n-1} - B_{n-2})\lambda^{n-1} + (AB_{n-2} - B_{n-3})\lambda^{n-2} + \\
& \cdots + (AB_2 - B_1)\lambda^2 + (AB_1 - B_0)\lambda + AB_0
\end{aligned}$$

Por tanto si igualamos los coeficientes de λ en ambas partes de la igualdad

$$(A - \lambda I_n) \cdot \mathrm{Adj}(A - \lambda I_n)^t = p_A(\lambda)I_n$$

obtenemos los siguiente coeficientes:

$$\begin{array}{lrcl}
\lambda^n & -B_{n-1} & = & (-1)^n I_n \\
\lambda^{n-1} & AB_{n-1} - B_{n-2} & = & c_{n-1}I_n \\
\lambda^{n-2} & AB_{n-2} - B_{n-3} & = & c_{n-2}I_n \\
\vdots & \vdots & \vdots & \vdots \\
\lambda^2 & AB_2 - B_1 & = & c_2 I_n \\
\lambda & AB_1 - B_0 & = & c_1 I_n \\
\lambda^0 & AB_0 & = & c_0 I_n
\end{array}$$

Multiplicando cada ecuación por la potencia conveniente de A y sumando todas las ecuaciones obtenemos:

$$\begin{array}{rrcl}
 & A^n(-B_{n-1}) & = & A^n((-1)^n I_n) \\
 & A^{n-1}(AB_{n-1} - B_{n-2}) & = & A^{n-1}(c_{n-1} I_n) \\
 & A^{n-2}(AB_{n-2} - B_{n-3}) & = & A^{n-2}(c_{n-2} I_n) \\
+ & \vdots & \vdots & \vdots \\
 & A^2(AB_2 - B_1) & = & A^2(c_2 I_n) \\
 & A(AB_1 - B_0) & = & A(c_1 I_n) \\
 & AB_0 & = & c_0 I_n \\
\hline
 & \mathbf{0}_n & = & (-1)^n A^n + \cdots + c_1 A + c_0 I_n = P_A(A)
\end{array}$$

□