Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Escuela de Ciencias Exactas y Naturales
Departamento de Matemática

Álgebra Lineal (R211 - CE9) 2025

1. Autovalores y autovectores

Recuerdos:

- V F-ev con $dimV < \infty$, B, B' bases de V, $T \in L(V)$. Ents.
$$[T]_{B'} = (C_{B'B})^{-1}[T]_B C_{B'B}.$$
- $A, B \in F^{n\times n}$ son *semejantes*, y anotamos $A \sim B$, si existe $P \in F^{n\times n}$ invertible tal que $B = P^{-1}AP$.
- La relación de semejanza es una relación de equivalencia en $F^{n\times n}$.
- V F-ev con $dimV < \infty$, $T, S \in L(V)$. Existen B_1, B_2 bases de V tales que $[S]_{B_1} = [T]_{B_2}$ sii existe $U \in L(V)$ invertible tal que $T = USU^{-1}$.

El primer problema que abordaremos es determinar si una matriz es semejante a una matriz diagonal. Hacemos entonces la siguiente definición:

Definición 1 *Una matriz $A \in F^{n\times n}$ se dice* ***diagonalizable*** *si existe $D \in F^{n\times n}$ matriz diagonal tal que $A \sim D$.*

Equivalentemente, A es diagonalizable si existe $C \in F^{n\times n}$ invertible tal que CAC^{-1} es diagonal.

A nivel de transformación lineal tenemos la siguiente definición:

Definición 2 *V F-ev con $dimV < \infty$, $T \in L(V)$ se dice que es diagonalizable si existe B base de V tal que $[T]_B$ es diagonal.*

alguna

Sigue que T es diagonalizable sii la matriz $[T]_B$ es diagonalizable para toda base B de V.

Supongamos ahora que T es diagonalizable y sea $B = \{v_1, \ldots, v_n\}$ una base de V tal que $[T]_B$ es diagonal:

$$[T]_B = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Tenemos aquí que para todo $i = 1, \ldots, n$, $[T(v_i)]_B = [T]_B[v_i]_B = [T]_B e_i = \lambda_i e_i = \lambda_i [v_i]_B = [\lambda_i v_i]_B$. Luego $T(v_i) = \lambda_i v_i$ para todo $i = 1, \ldots, n$.

Recíprocamente, si existen $\lambda_1, \ldots, \lambda_n \in F$ tales que $Tv_i = \lambda_i v_i$ para $B = \{v_1, \ldots, v_n\}$ una base de V, entonces $[T]_B = diag(\lambda_1, \ldots, \lambda_n)$.

Esto nos lleva a la siguiente definición *general* (para ev cualquiera, no pedimos dimensión finita):

Definición 3 *V F-ev, $T \in L(V)$, decimos que un vector $v \in V$ con $v \neq \overline{0}$ es un* ***autovector*** *de T si existe $\lambda \in F$ tal que $T(v) = \lambda v$. El escalar λ se llama autovalor de T.*

Observemos que si v es un autovector de T asociado a un autovalor λ, y $\alpha \in F$ es no nulo, entonces $T(\alpha \cdot v) = \alpha \cdot T(v) = \alpha(\lambda v) = \lambda(\alpha \cdot v)$, luego $\alpha \cdot v$ también es un autovector asociado al autovalor λ. Más aún, todo vector de $span(v) \backslash \{\overline{0}\}$ es autovector de T asociado al mismo autovalor λ. Profundizaremos en esto más adelante.

Volviendo a lo dicho anteriormente en cuanto a la definición de transformación lineal diagonalizable, concluímos la siguiente proposición para ev de dimensión finita:

Proposición 1 *V F-ev con $dim V < \infty$, $T \in L(V)$. Entonces T es diagonalizable sii existe B base de V formada por autovectores de T.*

Matricialmente lo anterior se traduce como sigue:

Definición 4 *$A \in F^{n \times n}$, $v \in F^n$ tal que $v \neq \overline{0}$ es un **autovector** de A si existe $\lambda \in F$ tal que $Av = \lambda v$. En tal caso, decimos que λ es un **autovalor** de A.*

Proposición 2 *$A \in F^{n \times n}$ es diagonalizable sii existe una base de F^n formada por autovectores de A.*

Ejemplos 1 *1. $A = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$, ¿es diagonalizable?*

Buscamos $x = (x_1, x_2)^t \in \mathbb{R}^2$, $x \neq \overline{0}$ y $\lambda \in \mathbb{R}$ tal que $Ax = \lambda x$. Es decir, tales que que el sistema $Ax = \lambda x$ tenga solución no trivial. Esto podemos reescribirlo en términos de sistemas homogéneos como sigue: nos preguntamos si hay solución no trivial del sistema homogéneo

$$(\lambda I - A)x = \overline{0}.$$

Resolvamos entonces el sistema: $\lambda I - A = \begin{pmatrix} \lambda - 2 & -3 \\ -2 & \lambda - 1 \end{pmatrix}$. La matriz es entonces cuadrada, luego ver que tenga solución no trivial se reduce a calcular su determinante y ver para qué valores de λ el determinante se anula:

$$det(\lambda I - A) = \lambda^2 - 3\lambda - 4 = (\lambda - 4)(\lambda + 1) = 0 \qquad sii \qquad \lambda = 4 \qquad y \qquad \lambda = -1.$$

Observemos que $det(\lambda I - A)$ es un polinomio, este polinomio será de interés, tiene nombre y todo, es un Very Important Polynomial! Es el polinomio característico de A. Ya lo veremos. Hemos encontrado entonces los dos autovalores de A. Busquemos para cada uno de ellos sus autovectores:

- *Si $\lambda = -1$, $(-1)I - A = \begin{pmatrix} -3 & -3 \\ -2 & -2 \end{pmatrix}$, cuya FER es $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, entonces el conjunto solución es $Sol = \{(x, -x) : x \in \mathbb{R}\} = span\{(1, -1)\}$. Este conjunto (que es un subsepacio) lo llamaremos autoespacio asociado al autovalor -1, más adelante lo definiremos. Por ahora recordemos que Sol es un sev de $\mathbb{R}^2$, es decir, contiene al vector nulo, pero éste no es autovector. Luego, el conjunto de autovectores de A asociados al autovalor -1 es $span\{(-1, -1)\} \backslash \{\overline{0}\}$.*
- *Si $\lambda = 4$ tendremos que el conjunto de autovalores asociado a 4 es $span\{(1, \frac{3}{2})\} \backslash \{\overline{0}\}$. (EJERCICIO)*

Puesto que $B = \{(1, -1), (1, \frac{3}{2})\}$ es li, es base de $\mathbb{R}^2$, resulta A diagonalizable (pues existe una base de autovectores). Más aún, para la matriz de cambio de base $C_{B_cB} A C_{B_cB}^{-1} = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}$.

(EJERCICIO: chequear esta igualdad. Ayuda: rápidamente vemos que $C_{BB_c} = \begin{pmatrix} 1 & 1 \\ -1 & \frac{3}{2} \end{pmatrix}$.

2. $A = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ *¿es diagonalizable?*

Buscamos λ tal que $Ax = \lambda x$ tiene solución no trivial. Como antes, necesitamos resolver el sistema homogéneo $(\lambda I - A)x = \overline{0}$, y para esto habrá que calcular su determinante: tenemos que

$$\lambda I - A = \begin{pmatrix} \lambda - 3 & 0 & 0 \\ -1 & \lambda - 3 & 0 \\ 0 & 0 & \lambda - 3 \end{pmatrix},$$

y $det(\lambda I - A) = (\lambda - 3)^3 = 0$ sii $\lambda = 3$. Así, A tiene un único autovalor. Calculemos sus autovectores asociados:

$$3I - A = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

de donde sigue que $Sol = span\{(0,1,0),(0,0,1)\}$. Vemos entonces que no hay más de dos autovectores li, luego no hay una base de autovectores de A, con lo cual resulta A no diagonalizable.

Otra forma de ver esto es la siguiente: si fuera diagonalizable existiría una matriz C de cambio de base (luego invertible) tal que $CAC^{-1} = 3I$, pero observemos que luego $A = C^{-1}3IC = \cdots = 3I$, contradicción.

Este procedimiento en general podemos describirlo como sigue:

Definición 5 *$A \in F^{n\times n}$. Llamamos* ***polinomio característico de*** *A al polinomio $\chi_A(X) = det(XI - A) \in F[X]$.*

El polinomio característico es mónico y de grado menor o igual a n: $\chi_A(X) \in F_n[X]$.

Proposición 3 *Si $A \in F^{n\times n}$ y $\lambda \in F$, entonces λ es un autovalor de A sii λ es raíz del polinomio característico χ_A.*

Demostración: λ es un autovalor de A sii existe $x \in F^n \backslash \{\overline{0}\}$ tal que $Ax = \lambda x$ sii $Ax = \lambda x$ tiene solución no trivial sii el sistema homogéneo $(\lambda I - A)x = \overline{0}$ tiene solución no trivial sii $det(\lambda I - A) = 0$ sii $\chi_A(\lambda) = 0$.

$\square$

Luego, como máximo A tiene n autovalores.

Ejemplo 1 $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. *Calculemos su polinomio característico:*

$$\chi_A(X) = det(XI - A) = \begin{vmatrix} X & -1 \\ 1 & X \end{vmatrix} = X^2 + 1.$$

Luego, si $\mathbb{F} = \mathbb{Q}$ o $\mathbb{F} = \mathbb{R}$, entonces no hay raíces de χ_A en $\mathbb{F}$, luego no existen autovalores de A, luego A no es diagonalizable. Si $\mathbb{F} = \mathbb{C}$, las raíces (los autovalores) son i y $-i$. Calculando de igual forma *los autovectores encontramos que $B = \{(1,i),(1,-i)\}$ es una base de $\mathbb{C}^2$ de autovectores de A. Luego A es diagonalizable en $\mathbb{C}^{2\times 2}$. EJERCICIO!*

Para poner estos resultados en términos de transformaciones lineales, necesitamos pasar una vez más por el concepto de semejanza: la siguiente proposición nos dice que dos matrices semejantes tienen el mismo polinomio característico:

Proposición 4 *$A \in F^{n\times n}$ y $C \in F^{n\times n}$ invertible. Entonces $\chi_{CAC^{-1}} = \chi_A$.*

Demostración: Observando que $XI - CAC^{-1} = CXIC^{-1} - CAC^{-1} = C(XI - A)C^{-1}$, tenemos que

$$\chi_{CAC^{-1}}(X) = det(XI - CAC^{-1}) = det(C(XI-A)C^{-1}) = det(C)det(XI-A)det(C^{-1}) = det(XI-A) = \chi_A.$$

□

Así definimos:

Definición 6 *V F-ev con $dimV < \infty$, $T \in L(V)$. El **polinomio característico asociado a** T es $\chi_T := \chi_{[T]_B}$ para cualquier B base de V.*

Y sigue entonces que λ es autovalor de T sii λ es raíz de χ_T.

2. Diagonalización

En los ejemplos hemos trabajado con los conjuntos de autovectores: los obtuvimos quitando el vector nulo de los sev definidos como solucion de un cierto sistema lineal homogéneo. Esos espacios tienen gran importancia, se trata de los *autoespacios*. Más precisamente:

Definición 7 *$A \in F^{n\times n}$, λ autovalor de A. El **autoespacio de A asociado a** λ se define por*

$$E_\lambda = \{v \in F^n : Av = \lambda v\} = \{v \in F^n : (\lambda I - A)v = \overline{0}\}.$$

Observaciones 1

1. *$\overline{0} \in E_\lambda$.*
2. *$E_\lambda \subset F^n$ sev puesto que es el espacio solución del sistema lineal homogéneo $(\lambda I - A)v = \overline{0}$.*
3. *$E_\lambda = N(\lambda I - A)$.*

Veremos en esta sección que estos sev descomponen en suma directa a F^n sii A es diagonalizable. Vayamos paso a paso:

Veamos primero que los autoespacios se suman de manera directa:

Proposición 5 *$A \in F^{n\times n}$, λ_1, λ_2 autovalores distintos de A. Entonces $E_{\lambda_1} \cap E_{\lambda_2} = \{\overline{0}\}$.*

Demostración: Si $v \in E_{\lambda_1} \cap E_{\lambda_2}$ entonces $Av = \lambda_1 v$ y $Av = \lambda_2 v$. Luego $\lambda_1 v = \lambda_2 v$. Luego $(\lambda_1 - \lambda_2)v = \overline{0}$. Puesto que $\lambda_1 \neq \lambda_2$ sigue que $v = \overline{0}$.

□

Más generalmente, tenemos que:

Proposición 6 *$A \in F^{n\times n}$, $\lambda_1, \dots, \lambda_r$ autovalores distintos de A. Entonces para $j = 1, \dots, r$,*

$$E_{\lambda_j} \cap \bigoplus_{k=1, k\neq j}^{r} E_{\lambda_k} = \{\overline{0}\}.$$

Demostración: EJERCICIO. Ayuda: proceder por inducción sobre r.

□

Ahora, recordando que todo autovalor es raíz del polinomio característico, veremos cómo se relacionan la dimensión del autoespacio asociado al autovalor con su multiplicidad como raíz del polinomio característico.

Proposición 7 *$A \in F^{n\times n}$, λ autovalor de A, r multiplicidad de λ como raíz de χ_A (esto es, $\chi_A(X) = (X-\lambda)^r P(X)$ con $P(\lambda) \neq 0$). Entonces, $dim E_\lambda \leq r$.*

Demostración: Sea $T : F^n \to F^n$ definida según $Tx = Ax$. Sea $s = dim(E_\lambda)$ y $B_\lambda = \{v_1, \ldots, v_s\}$ base de E_λ. Completamos a una base $B = \{v_1, \ldots, v_n\}$ de F^n. Así,

$$[T]_B = \left(\begin{array}{c|c} \lambda I_s & N_{s\times n-s} \\ \hline 0_{n-s\times s} & M_{s\times s} \end{array} \right).$$

en efecto $[Tv_i]_B = \lambda e_i$ para toda $i = 1, \ldots, s$, es decir, sus coeficientes de $v_{s+1}, \ldots, v_n$ son nulos (y los demás no se sabe ni interesan, por eso quedan escritas dos matrices de los tamaños indicados).

Además tenemos también que

$$XI - [T]_B = \left(\begin{array}{c|c} (X-\lambda)I_s & -N \\ \hline 0 & XI_{n-s} - M \end{array} \right).$$

El polinomio característico es entonces

$$\chi_A(X) = det(X - [T]_B) = det(X-\lambda)det(XI_{n-s} - M) = (X-\lambda)^s Q(X),$$

donde los determinantes son de las submatrices indicadas. Ahora bien, como por hipótesis $\chi_A(X) = (X-\lambda)^r P(X)$ con $P(\lambda) \neq 0$, sigue que

$$(X-\lambda)^s Q(X) = (X-\lambda)^r P(X),$$

y como $P(\lambda) \neq 0$ debe ser $s \leq r$, esto es, $dim(E_\lambda) \leq mult(\lambda, \chi_A)$.

□

Estamos finalmente en condiciones de caracterizar las matrices diagonalizables en términos de autoespacios (vale decir, en términos de autovalores y autovectores).

Teorema 1 *$A \in F^{n\times n}$, $\lambda_1, \ldots, \lambda_r$ todos los autovalores distintos de A. Son equivalentes:*

1. *A es diagonalizable en $F^{n\times n}$,*
2. *$\bigoplus_{i=1}^{r} E_{\lambda_i} = F^n$,*
3. *$\chi_A(X) = (X-\lambda_1)^{a_1} \ldots (X-\lambda_r)^{a_r}$ y $a_i = dim(E_{\lambda_i})$, $i = 1, \ldots, r$*

Demostración:

$1 \Rightarrow 2)$ A diagonalizable en $F^{n\times n}$. Luego existe $B = \{v_1, \ldots, v_n\}$ base de autovectores de A. Para cada $v_j \in B$, $j = 1, \ldots, n$, existe λ_i, $i = 1, \ldots, r$ tales que v_j autovector asociado al autovalor λ_i. Luego, cada v_j está en algún E_{λ_i}, y como B es base debe ser $E_{\lambda_1} + \cdots + E_{\lambda_n} = F^n$. La proposición anterior nos permite concluír que esta suma es suma directa: $\bigoplus_{i=1}^{r} E_{\lambda_i} = F^n$.

$2 \Rightarrow 3)$ Tenemos que $\bigoplus_{i=1}^{r} E_{\lambda_i} = F^n$. Por las propiedades vistas tenemos que $dim E_{\lambda_i} \leq mult(\lambda_i, \chi_A)$. Luego

$$n \leq \sum_{i=1}^{r} dim E_{\lambda_i} = \sum_{i=1}^{r} mult(\lambda_i, \chi_A) = gr(\chi_A) = n,$$

luego estas desigualdades son todas igualdades. En particular, $\sum_{i=1}^{r} mult(\lambda_i, \chi_A) = gr(\chi_A)$, de donde $\chi_A(X) = (X-\lambda_1)^{a_1} \ldots (X-\lambda_r)^{a_r}$ con $a_i = mult(\lambda_i, \chi_A)$. Como cada $dim(E_{\lambda_i}) \leq mult(\lambda_i, \chi_A) = a_i$,

$i = 1, \dots, r$, debe ser $a_i = dim(E_{\lambda_i}) = mult(\lambda_i, \chi_A)$ pues si vale el menor estricto resulta en una contradicción.

$3 \Rightarrow 1)$ Sabemos que $\chi_A(X) = (X - \lambda_1)^{a_1} \dots (X - \lambda_r)^{a_r}$. Para cada $i = 1, \dots, r$ sea B_i base de E_{λ_i}. Así, $B = B_1 \cup \dots \cup B_r$ (unión disjunta, porqué?) es base de $\bigoplus_{i=1}^{r} E_{\lambda_i} \subset F^n$. Así:

$$|B| = \sum_{i=1}^{r} |B_i| = \sum_{i=1}^{r} dim(E_{\lambda_i}) = \sum_{i=1}^{r} a_i = gr(\chi_A) = n.$$

Entonces $\bigoplus_{i=1}^{r} E_{\lambda_i} = F^n$ y por lo tanto B es base de autovectores de A, luego A es diagonalizable.

□

Ejemplo 2 *Sea* $A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$, *¿es diagonalizable? Calculemos* $\chi_A(X) = det(\lambda I - A) = \begin{pmatrix} \lambda - 1 & 0 & 0 & -1 \\ -1 & \lambda - 1 & 0 & 1 \\ 0 & 0 & \lambda - 1 & -1 \\ 0 & 0 & 0 & \lambda - 2 \end{pmatrix} = (\lambda - 1)^3(\lambda - 2)$. *Los autovalores son entonces* $\lambda_1 = 1$ *de multiplicidad* $a_1 = 3$ *y* $\lambda_2 = 2$ *de multiplicidad* $a_2 = 1$. *Si calculamos* $E_1 = \{x \in \mathbb{R}^4 : (I - A)x = \overline{0}\} = span(\{(0, 1, 0, 0), (0, 0, 1, 0)\})$ *(EJERCICIO: calcular esto), vemos que dim* $E_1 = 2 < 3 = mult(1, \chi_A)$. *Resulta así que* A *no es diagonalizable.*

3. El polinomio minimal

Para una matriz cuadrada tenemos asociado un very important polynomial, el polinomio característico, que nos sirve para hallar los autovalores (son sus raíces) y caracterizar diagonalización. En esta sección asociaremos otro very important polynomial: el *polinomio minimal*, que nos brindará más información y nos permitirá dar un otro criterio de diagonalización y en la unidad próxima será de enorme utilidad.

A lo largo de esta sección y de la unidad próxima estaremos dando una utilidad nueva a los polinomios: nos servirán para dar expresiones algebraicas a otros objetos. Es decir, utilizaremos mucho más su naturaleza algebraica. La siguiente es la primera definición donde un polinomio sirve para describir una expresión matricial de manera polinómica (y of course también a su traducción a endomorfismos en ev de dim finita):

Consideremos un polinomio $p(X) = a_0 + a_1x + a_2x^2 + \dots + a_rx^r \in F[X]$. Definimos:

- Para $A \in F^{n \times n}$,
$$p(A) := a_0I + a_1A + a_2A^2 + \dots + a_rA^r \in F^{n \times n}.$$
  Observemos que (EJERCICIO) si $q \in F[X]$ es otro polinomio:
    - $(p + q)(A) = p(A) + q(A)$,
    - $(pq)(A) = p(A)q(A)$.
- Si V F-ev con $dimV = n$ y $T \in L(V)$,
$$p(T) := a_0id_V + a_1T + a_2T^2 + \dots + a_rT^r \in L(V),$$

donde $T^k = T \circ \cdots \circ T$ (composición de k funciones).

Diremos además que un polinomio p **anula** a una matriz A si $p(A) = \mathbf{0} \in F^{n\times n}$ (la matriz nula) y **anula** a una transformación lineal T si $p(T) \equiv 0$ es la **función** constante igual a cero (es decir, la transformación nula $p(T)(v) = 0(v)$).

Lema 1 *Si $A \in F^{n\times n}$, existe $p(X) \in F[X]$, $p \neq \overline{0(X)}$ (el polinomio nulo) tal que $p(A) = \overline{0}$.*

Demostración: Consideremos el conjunto $\{I, A, A^2, \ldots, A^{n^2}\} \subset F^{n\times n}$. Este conjunto tiene $n^2 + 1$ elementos, luego es ld (porque?). Esto significa que existen escalares $a_0, a_1, \ldots, a_{n^2} \in F$ no todos nulos tales que $a_0 I + a_1 A + \cdots + a_{n^2} A^{n^2} = \overline{0}$, esto es, el polinomio $p(X) = a_0 + a_1 X + \cdots + a_{n^2} X^{n^2} \in F[X]$ es no nulo y anula a A.

□

De la proposición anterior sigue trivialmente que no sólo existe un polinomio que anule A, sino que hay infinitos. En efecto, si $q(X)$ es cualquier otro polinomio (no nulo), el polinomio $r(X) = p(X)q(X)$ es no nulo y anula a A.

A título informativo (se puede saltear este párrafo y no pasa nada): el conjunto de todos los productos del polinomio $p(X)$ por un polinomio $q(X)$ cualquiera se llama ideal principal generado por $p(x)$ y se anota $(p(x))$. En álgebra abstracta, dentro de unos años, se verá que el conjunto de polinomios tiene estructura de anillo, y más aún, es un dominio de ideales principales. Esto nos garantizará dos cuestiones fundamentales: tener un algoritmo de la división y un teorema de factorización única. Los fundamentos algebraicos son muy interesantes, por ahora con lo que sabemos de polinomios de Álgebra y Geometría Analítica I nos alcanza, pues justamente ambas herramientas nos son conocidas.

Volviendo al conjunto de polinomios que anulan a una matriz A: hay exactamente uno de estos polinomios que es mónico y de grado mínimo. En efecto:

- Veamos que existe tal polinomio. El argumento de existencia que utilizaremos ya lo hemos aplicado alguna vez (recuerdan la prueba del principio del palomar?). Probar existencia sin dar el objeto, sin construírlo, es difícil. Esta técnica sirve para estos casos.

  Sea $H_A = \{gr(p) : p \in F[X], p(A) = \mathbf{0} \text{ y } p \neq 0(X)\}$ el conjunto de todos los grados posibles de polinomios no nulos que anulen a A. Claramente $H_A \subset \mathbb{N}$. Luego, como $\mathbb{N}$ es un conjunto bien ordenado y H_A es no vacío (porqué?) que está acotado inferiormente, resulta que debe tener un primer elemento, digamos r.

  Consideremos entonces el polinomio $q \in F[X]$ tal que $q(A) = \mathbf{0}$ y $gr(q) = r$. Si $q(X)$ no es mónico, consideramos el polinomio $q'(x)$ obtenido al dividir $q(X)$ por su coeficiente principal (que es no nulo, porque?). Obtenemos así la existencia de un polinomio mónico de grado mínimo que anula a A.

- La unicidad sigue del argumento usual para unicidad: consideremos $q(x), q'(x)$ dos tales polinomios, y supongamos que son distintos. Obviamente tienen el mismo grado puesto que de lo contrario alguno no tendría grado mínimo. Entonces el polinomio $q(x) - q'(X) = (q - q')(X)$ es no nulo, anula a A, y por ser ambos mónicos resulta $gr(q - q') < r$. Esto contradice que q (y q') sean de grado mínimo. Luego, $q(X)$ y $q'(X)$ deben ser iguales.

Esto nos dice que la siguiente es una buena definición:

Definición 8 *$A \in F^{n\times n}$. Llamamos **polinomio minimal de** A y denotamos $m_A(X)$ al polinomio no nulo, mónico y de grado mínimo que anula a A.*

Veamos un ejemplo:

Ejemplo 3 *Para $A = \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix}$ queremos hallar el polinomio minimal $m_A(X)$. Para esto observamos que el conjunto $\{I, A\}$ es li, luego $gr(m_A) > 1$. Nos preguntamos entonces acerca de la lineal dependencia del conjunto $\{I, A, A^2\}$. Veamos:*

Consideremos la ecuación matricial

$$a_0 I + a_1 A + a_2 X^2 = \overline{0},$$

si hacemos los cálculos, $A^2 = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$*, de donde la ecuación matricial queda descripta como*

$$a_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + a_1 \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} + a_2 \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Si planteamos el sistema asociado y lo resolvemos, resulta que tiene infinitas soluciones, luego el conjunto es ld. (EJERCICIO).

Para hallar el polinomio minimal, como dicho polinomio es mónico, podríamos haber planteado desde un principio la ecuación $a_0 I + a_1 A + X^2 = \boldsymbol{0}$*. En efecto, si hacemos los cálculos, obtenemos que* $a_0 = 1$ *y* $a_1 = 2$ *son los coeficientes del polinomio buscado:* $m_A(X) = X^2 + 2X + 1$*.*

Finalmente observemos que si calculamos el polinomio característico asociado a A*,* $\chi_A(X) = det(XI - A)$ *tenemos que* $\chi_A(X) = (X+1)^2 = m_A(X)$*.*

Si bien no es algo que ocurra siempre, es cierto que ambos polinomios se relacionan. Exploraremos un poco esta relación.

Puesto que hemos tomado el polinomio minimal de A en el conjunto de todos los polinomios que anulan a A, mónico y de grado mínimo, tenemos que m_A divide a todos los polinomios de dicho conjunto. En efecto:

Proposición 8 *$A \in F^{n\times n}$, $p \in F[X]$. Entonces $p(A) = \boldsymbol{0}$ sii $m_A | p$.*

Demostración: Recordar que decimos que un polinomio q divide a un polinomio p y denotamos $q|p$ si existe otro polinomio c tales que $p(x) = c(x)q(x)$.

> $\Rightarrow$) Tenemos que p anula a A. Puesto que m_A tiene grado mínimo entre todos los que anulan a A debe ser $gr(m_A) \leq gr(p)$. Por el algoritmo de la división, existen polinomios $c(X), r(X) \in F[X]$ tales que $p(X) = c(X)m_A(X) + r(X)$ con $gr(r) < gr(m_A)$ o $r \equiv 0$. Ahora bien, $0 = p(A) = c(A)m_A(A) + r(A) = r(A)$, de donde resulta que $r \equiv 0$ (porqué?) y por lo tanto $m_A|p$.
>
> $\Leftarrow$) Si $m_A|p$ existe $c(X) \in F[X]$ tal que $p(X) = c(X)m_A(X)$. Luego $p(A) = c(A)m_A(A) = \mathbf{0}$.

□

Es nuestro objetivo siguiente poder definir el polinomio minimal de una transformación lineal abstracta en un ev de dimensión finita. Para esto, y como hay involucradas matrices, deberíamos ver cómo se comporta el polinomio minimal respecto de la relación de semejanza de matrices. Ya habíamos visto que matrices semejantes tienen el mismo polinomio característico, esto también será así para el polinomio minimal.

Recordemos que dos matrices $A, B \in F^{n\times n}$ se dicen semejantes si existe una matriz $C \in F^{n\times n}$ invertible tal que $A = CBC^{-1}$.

Desafío 1 *Cuando aplicamos un polinomio a una matriz trabajamos con potencias. Pobar que si $A \sim B$ entonces $A^k \sim B^k$ con la misma matriz de conjugación C. Help: inducir sobre k, observar que $A^{k+1} = A^k A = CB^kC^{-1}CBC^{-1} = CB^{k+1}C^{-1}$. Bueno, les hice la cuenta. Escribirlo bien.*

Lema 2 *$A, B \in F^{n\times n}$, $A \sim B$. Si $p \in F[X]$ entonces $p(A) \sim (B)$. En particular, $p(A) = \overline{0}$ sii $p(B) = \overline{0}$.*

Demostración: Como $A \sim B$, existe una matriz $C \in F^{n\times n}$ invertible tal que $A = CBC^{-1}$. Si $p(X) = \sum_{i=0}^{r} a_i X^i$, por lo que observamos recién respecto de las potencias, $p(A) = \sum_{i=0}^{r} a_i A^i = \sum_{i=0}^{r} a_i CB^iC^{-1} = C\left(\sum_{i=0}^{r} a_i B^i\right)C^{-1} = Cp(B)C^{-1}$, que es lo que queríamos probar.

$\square$

Ahora sí probemos el resultado que nos permitirá pasar a ev abstractos (finito dimensionales).

Proposición 9 *$A, B \in F^{n\times n}$, $A \sim B$. Entonces $m_A = m_B$.*

Demostración: EJERCICIO. Ayuda: justificar que $m_B | m_A$ y que $m_A | m_B$.

$\square$

A nivel de transformaciones lineales tenemos: si V F-ev con $dimV = n$, $T \in L(V)$ y B_1, B_2 bases de V entonces $m_{[T]_{B_1}} = m_{[T]_{B_2}}$. Queda entonces bien definido el polinomio minimal para T:

Definición 9 *V F-ev con $dimV = n$, $T \in L(V)$, B base de V. Definimos el* ***polinomio minimal de T*** *como el polinomio $m_T := m_{[T]_B}$.*

Volviendo a la relación entre el polinomio característico y el polinomio minimal (de una matriz), resulta que ambos tienen las mismas raíces: los autovalores de la matriz. Esto lo probamos en la siguiente proposición:

Proposición 10 *$A \in F^{n\times n}$, $\lambda \in F$. Entonces λ es un autovalor de A sii λ es raíz de m_A.*

Demostración:

=>) Si λ es un autovalor de A, por el algoritmo de la división, existen $c(X), r(X) \in F[X]$, con $r(X) \equiv r$ para una constante r (por qué?) tal que $m_A(X) = (X - \lambda)c(X) + r$. Aplicando este polinomio a A tenemos $m_A(A) = (A - \lambda I)c(A) + rI$. Como λ es un autovalor de A, sigue que $\mathbf{0} = rI$, luego $r = 0$, de modo que λ es raíz de m_A (por qué?).

<=) Sea λ raíz de m_A. Luego $m_A(X) = (X - \lambda)q(X)$ para algún $q(X) \in F[X]$. Luego $\mathbf{0} = m_A(A) = (A - \lambda I)q(A)$ y como $gr(q) = gr(m_A) - 1$ debe ser $q(A) \neq \mathbf{0}$. Entonces existe $w \in F^n$ tal que $q(A)w = v \neq \overline{0} \in F^n$. Luego, $(A - \lambda I)v = (A - \lambda I)q(A) = \overline{0} \in F^n$, de donde v es autovector de A asociado al autovalor λ.

$\square$

Terminamos esta sección definiendo el polinomio minimal asociado a un vector, de manera análoga: si $A \in F^{n\times n}$, $v \in F^n$ y $p(X) \in F[X]$ definimos $p(v) = p(A)v$. Decimos entonces que p anula a v si $p(v) = \overline{0} \in F^n$.

En particular, como el polinomio minimal asociado a A anula a A, tenemos que $m_A(v) = m_A(A)v = \overline{0} \in F^n$, para todo $v \in V$. Así, si damos $v \in V$ resulta que existe un único polinomio mónico de grado mínimo que lo anula, y lo denotamos m_v. (EJERCICIO: escribir y probar esto con detalle, es igualita a la prueba para matrices).

Ejemplos 2
1. *$A \in F^{n\times n}$, $v \in F^n$ autovector de A asociado al autovalor λ. Entonces $m_v(X) = X - \lambda$. En efecto: $m_v(v) = m_v(A)v = (A - \lambda I)v = \overline{0} \in F^n$. Aquí $gr(m_v) = 1$.*
2. *$A = \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$, $v = e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Calculemos m_{e_1}. Para esto planteamos un polinomio mónico, comenzando por el grado 1: $p(x) = x + a_0 \in [X]$. Queremos que $\overline{0} = p(e_1) = p(A)e_1 = (A + a_0I)e_1 = Ae_1 + a_0e_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} + a_0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 + a_0 \\ 1 \end{pmatrix}$, de donde sigue que $1 = 0$, luego*

el polinomio buscado no puede ser de grado 1. Veamos con grado 2: sea $p(X) = a_0 + a_1X + X^2$. Luego, $p(e_1) = p(A)e_1 = \dots$ EJERCICIO: $m_{e_1}(X) = 1 + 2X + X^2$.

¿Y en general cómo se calcula el polinomio minimal? Veamos: si $A \in F^{n\times n}$, $v \in F^n$ y $m_v(X) = a_0 + a_1X + \cdots + X^m$ es el polinomio minimal de v, entonces debe ser

$$m_v(v) = m_v(A)v = a_0v + a_1Av + \cdots + A^m v = \overline{0} \in F^n.$$

Como el grado de m_v es mínimo entre los polinomios que satisfacen $p(v) = \overline{0}$, se tiene que el conjunto $\{v, Av, \dots, A^{m-1}v\}$ es li. En efecto, si fuera ld existirían constantes $\beta_0, \dots, \beta_{m-1}$ no todas nulas tales que $\sum_{i=0}^{m-1} \beta_i A^i v = \overline{0}$, luego el polinomio $p(X) = \sum_{i=0}^{m-1} \beta_i X^i$ es no nulo, anula a v y es de grado menor al minimal, lo cual es absurdo. Entonces, para hallar m_v empezamos buscando el primer $m \in \mathbb{N}$ tal que $\{v, Av, \dots, A^m v\}$ es ld. Así, $A^m v = -a_0v - a_1Av + \cdots + (-1)a_{m-1}A^{m-1}v$. Entonces $m_v(X) = a_0 + a_1X + \cdots + X^m$.

Veamos algunos resultados más antes del teorema de Cayley-Hamilton que veremos en la próxima sección. En primer lugar, al igual que lo que pasaba con las matrices, el polinomio minimal de v divide a todo polinomio que lo anule:

Proposición 11 *$A \in F^{n\times n}$, $v \in F^n$, $p(X) \in F[X]$. Entonces $p(v) = \overline{0}$ sii $m_v|p$. En particular, $m_v|m_A$.*

Demostración: Como el grado de m_v es mínimo, $gr(m_v) \leq gr(p)$. Por el algoritmo de la división existen $c(X), r(X) \in F[X]$ tales que $p(X) = c(X)m_v(X) + r(X)$ con $r(X) \equiv 0$ o $gr(r) < gr(m_v)$. Ahora bien, $p(v) = m_v(v)c(v) + r(v)$. Entonces:

$$P \text{ anula a } v \text{ sii } r(v) = 0 \text{ sii } r \equiv 0 \text{ sii } m_v|p.$$

□

Veamos, finalmente, cómo calculamos el polinomio minimal de una matriz a partir de los polinomios minimales de una base:

Proposición 12 *$A \in F^{n\times n}$, $B = \{v_1, \dots, v_n\}$ base de F^n. Entonces*

$$m_A = m.c.m.\{m_{v_1}, \dots, m_{v_n}\}.$$

Demostración: Sea $p = m.c.m.\{m_{v_1}, \dots, m_{v_n}\}$. Por la propiedad anterior, $m_{v_i}|m_A$, para cada $i = 1, \dots, n$. Luego $p|m_A$ Además, $m_{v_i}|p$ para cada $i = 1, \dots, n$ (por qué?). Luego $p(A)v_i = \overline{0}$ para cada $i = 1, \dots, n$. Sea $v \in F^n$ y $\alpha_1, \dots, \alpha_n \in F$ tales que $v = \sum_{i=1}^{n} \alpha_i v_i$. Entonces $p(A)v = \sum_{i=1}^{n} \alpha_i p(A)v_i = \overline{0}$. Así, $p(A)v = \overline{0}$ para todo $v \in F^n$, de modo que $p(A) = \overline{0} \in F^{n\times n}$. Sigue que $m_A|p$, de donde $m_A = p$.

□

Ejemplo 4 $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. *Calculemos m_A.*

Para esto, consideremos la base canónica $E = \{e_1, e_2, e_3\}$ de $\mathbb{R}^3$. Por la proposición anterior $m_A = m.c.m.\{m_{e_1}, m_{e_2}, m_{e_3}\}$. Calculemos cada uno de estos polinomios:

- *Para calcular m_{e_1} observamos que $\{e_1, Ae_1\}$ es li y que $\{e_1, Ae_1, A^2e_1\}$ es ld. En efecto, $Ae_1 = e_1 + e_2$ y $A^2e_1 = e_1 + 2e_2$ (EJERCICIO). Planteando entonces $a_0e_1 + a_1Ae_1 + A^2e_1 = \overline{0}$ obtenemos que $a_1 = -2$ y $a_0 = 1$ (EJERCICIO). Resulta entonces $m_{e_1}(X) = 1 - 2X + X$.*
- *Análogamente calculamos (EJERCICIO) $m_{e_2}(X) = X - 1$.*
- *Análogamente calculamos (EJERCICIO) $m_{e_3}(X) = X - 2$.*

Entonces $m_A(X) = (X - 1)^2(X - 2)$.

4. El teorema de Cayley-Hamilton

El teorema de Cayley-Hamilton nos dice que dada una matriz $A \in F^{n\times n}$, el polinomio minimal divide al característico, y sigue inmediatamente que entonces el característico anula a A. Es un Very Important Theorem, tiene nombre propio.

Aquí es importante observar lo siguiente: hemos trabajado con la expresión $\chi_A(X) = det(XI - A)$, de modo que parecería ser trivial la afirmación del teorema de Cayley-Hamilton: $\chi_A(A) = det(AI - A) = det(0) = 0 \in F$. Pero esto es incorrecto, pues al aplicar un polinomio a una matriz debe resultar en una matriz: si el polinomio es $\chi_A(X) = det(XI - A) = p(X) = \sum_{i=0}^{n} a_i X^i = a_0 + a_1 X + \cdots + a_n X^n$, entonces $\chi_A(A) = a_0 I + a_1 A + \cdots + a_n A^n \in F^{n\times n}$. Hay que tener, entonces, mucho cuidado al aplicar la expresión para el polinomio característico a las matrices.

Observemos previamente lo siguiente: consideremos $A \in F^{n\times n}$ y $T : F^n \to F^n$ definida por $T(x) = Ax$. Sean $v \in F^n$ y $k \in \mathbb{N}$ tales que $B = \{v, Tv, \ldots, T^k v\}$ es li y $T^{k+1}v = -a_0 v - a_1 Tv - \cdots - a_k T^k v$ es ld. Para calcular $[T]_B$, puesto que sus columnas son las componentes en la base B de los transformados de la base B, tenemos que:

- el primer elemento v de B se transforma en Tv, que es el segundo elemento de B, luego $[Tv]_B = e_2$;

- el segundo elemento Tv de B se transforma en $T^2 v$, que es el tercer elemento de B, luego $[T(Tv)]_B = e_3$;

- así sucesivamente,

- el elemento k-ésimo de B es $T^{k-1}v$ y se transforma en $T^k v$, que es el elemento $k+1$-ésimo de B, luego $[T(T^{k-1}v)]_B = e_{k+1}$;

- el elemento $k+1$-ésimo de B es $T^k v$ y se transforma en $T^{k+1}v = -a_0 v - a_1 Tv - \cdots - a_k T^k v$, luego $[T(T^k v)]_B = (-a_0, -a_1, \ldots, -a_{k-1}, -a_k)$.

Entonces $[T]_B$ es la matriz $k+1 \times k+1$ dada por

$$[T]_B = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & -a_{k-1} \\ 0 & 0 & 0 & \ldots & 1 & -a_k \end{pmatrix}.$$

Además, si queremos calcular su polinomio característico $\chi_T(X) = \chi_{[T]_B}(X) = det(XI - [T]_B)$, donde

$$XI - [T]_B = \begin{pmatrix} X & 0 & 0 & \ldots & 0 & a_0 \\ -1 & X & 0 & \ldots & 0 & a_1 \\ 0 & -1 & X & \ldots & 0 & a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & X & a_{k-1} \\ 0 & 0 & 0 & \ldots & -1 & X + a_k \end{pmatrix},$$

tenemos que $\chi_T(X) = X^{k+1} + a_k X^n + \cdots + a_1 X + a_0$. Nos convencemos fácilmente de este resultado si desarrollamos por la última fila.

Desafío 2 *Convencerse.*

Very important: *convencerse* es sólo un eufemismo para *demostrar*. Hay que probarlo, como todo en la ciencia. (*Ciencia* siendo un eufemismo para *Álgebra Lineal*, como pretendo convencerlos este cuatrimestre, o quizás debería decir demostrarles, je.)

Todos estos cálculos nos permitirán desarrollar la prueba del Teorema de Cayley-Hamilton.

Very Important Theorem 1 ***El teorema de Cayley-Hamilton.*** *$A \in F^{n\times n}$. Ents. $m_A | \chi_A$. Equivalentemente, $\chi_A(A) = 0 \in F^{n\times n}$.*

Demostración: Consideremos $T : F^n \to F^n$ definida por $T(x) = Ax$. Sean $v \in F^n$ y $k \in F$ tales que $B_1 = \{v, Tv, \ldots, T^k v\}$ es li y $T^{k+1}v = -a_0 v - a_1 Tv - \cdots - a_k T^k v$. Ents. $m_v(X) = X^{k+1} + a_k X^k + \cdots + a_1 X + a_0$. Extendemos el conj. li a una base de F^n: $B = \{v, Tv, \ldots, T^k v, w_{k+2}, \ldots, w_n\}$. Así, gracias a la observación que hicimos previamente, tenemos que

$$[T]_B = \left(\begin{array}{cccccc|c} 0 & 0 & 0 & \ldots & 0 & -a_0 & \\ 1 & 0 & 0 & \ldots & 0 & -a_1 & \\ 0 & 1 & 0 & \ldots & 0 & -a_2 & M \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \\ 0 & 0 & 0 & \ldots & 0 & -a_{k-1} & \\ 0 & 0 & 0 & \ldots & 1 & -a_k & \\ \hline & & & \mathbf{0} & & & N \end{array}\right).$$

Observemos esta matriz por bloques: el bloque superior izquierdo es $k+1 \times k+1$, la matriz de la observación previa. Debajo de ese bloque, hay un bloque $n-(k+1) \times k+1$ cuyas entradas son todas 0. El resto de las entradas quedan en dos bloques que llamamos M y N, con N cuadrada.

Tenemos que $\chi_A(X) = \chi_T(X) = \chi_{[T]_B}(X) = det(XI - [T]_B)$. Luego,

$$XI - [T]_B = \left(\begin{array}{cccccc|c} X & 0 & 0 & \ldots & 0 & a_0 & \\ -1 & X & 0 & \ldots & 0 & a_1 & \\ 0 & -1 & X & \ldots & 0 & a_2 & -M \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \\ 0 & 0 & 0 & \ldots & X & a_{k-1} & \\ 0 & 0 & 0 & \ldots & -1 & X + a_k & \\ \hline & & & \mathbf{0} & & & XI_{n-(k+1)} - N \end{array}\right).$$

Así, de acuerdo a lo observado anteriormente, $\chi_A(X) = (X^{k+1} + a_k X^n + \cdots + a_1 X + a_0) det(XI_{n-(k+1)} - N) = m_v(X)Q(X)$. Sigue que $m_v | \chi_A$. Como v es arbitrario, en particular para la base canónica $B = \{e_1, \ldots, e_n\}$ se tiene que $m_{e_i} | \chi_A$ para todo $i = 1, \ldots, n$, de donde $m_A = m.c.m.\{m_{e_1}, \ldots, m_{e_n}\} | \chi_A$.

$\square$

Observaciones 2 *Como consecuencias inmediatas sigue que, si $A \in F^{n\times n}$,*

1. $gr(m_A) \leq n$,
2. *si* $gr(m_A) = n$ *ents.* $m_A = \chi_A$,
3. *si existe* $v \in F$ *tal que* $gr(m_v) = n$ *ents.* $m_v = m_A = \chi_A$.

Ejemplo 5 $A = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \in \mathbb{R}^{n\times n}$. *Daremos una fórmula general para A^n, y más. Calculemos $det(XI - A) = ... = (X-1)^2$ (EJERCICIO). Luego $\chi_A(X) = X^2 - 2X + I$. Por el teorema de Cayley-Hamilton es $\chi_A(A) = A^2 - 2A + I = 0$. Sigue que $A^2 = 2A - I$. Y podemos inducir $A^3 = A^2A = (2A-I)A = 2A^2 - A = 2(2A-I) - A = 3A - 2I$, $A^4 = 4A - 3I$, ..., $A^n = nA - (n-1)I$. (EJERCICIO). Más aún, observemos lo siguiente: nuevamente, de $\chi_A(A) = A^2 - 2A + I = 0$ sigue que $I = 2A - A^2 = A(2I - A)$ y por lo tanto $A^{-1} = 2I - A$. Es claro que en este ejemplo A es invertible.*

Desafío 3 *Justificar las siguientes observaciones:*

- *Para dos autovalores distintos, digamos $\lambda_1 \neq \lambda_2$, y dos autovalores v_1, v_2 correspondientes a λ_1, λ_2, respectivamente, se tiene que $\{v_1, v_2\}$ es li.*
- *Si A es diagonalizable y $\lambda = 0$ es autovalor de A, entonces A no es invertible.*

En general tenemos entonces que:

Proposición 13 *$A \in F^{n\times n}$ invertible. Luego $A^{-1} \in span\{I, A, \dots, A^{n-1}\}$.*

Demostración: Por el teorema de Cayley-Hamilton sabemos que $\chi_A(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$ es tal que $\chi_A(A) = 0 \in F^{n\times n}$. De donde

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1A + a_0I = 0.$$

Ahora bien, $\chi_A(0) = a_0 \in F$, luego $\chi_A(0) = a_0 = det(0I_n - A) = det(-A) = (-1)^n det(A) \neq 0$. De estas igualdades sigue que

$$I_n = -\frac{1}{a_0}(A^n + a_{n-1}A^{n-1} + \cdots + a_1A) = -\frac{1}{a_0}(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1I)A,$$

luego

$$A^{-1} = -\frac{1}{a_0}(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1I).$$

□

Observemos que si $\chi_A(X)$ se factoriza linealmente como $\chi_A(X) = (X-\lambda_1)\dots(X-\lambda_n)$ con $\lambda_i \neq \lambda_j$ para $i,j = 1,\dots,n$, entonces $F^n = \oplus_{i=1}^n E_{\lambda_i}$, de donde F^n tiene una base de autovectores y por lo tanto A resulta diagonalizable. La recíproca no es cierta (la identidad es un ejemplo simple de esto, por qué?). Es decir, no podemos caracterizar diagonalizabilidad de esta forma mediante el polinomio característico, pero sí podemos hacerlo con el polinomio minimal:

Proposición 14 *$A \in F^{n\times n}$. Entonces A es diagonalizable en $F^{n\times n}$ si y sólo si m_A tiene todas sus raíces en F y son simples.*

Demostración:

$\Rightarrow$) A diagonalizable. Sean $\lambda_1, \dots, \lambda_r$ autovalores de A distintos y sea $B = \{v_1, \dots, v_n\}$ base de autovectores de A. Para cada $v \in V$, su polinomio minimal respecto de A es $m_A(X) = m.c.m.\{m_{v_1}, \dots, m_{v_n}\} = m.c.m.\{X - \lambda_{j_1}, \dots, X - \lambda_{j_n}\}$, puesto que cada v_i es autovector correspondiente a un autovalor λ_{j_i} y luego $m_{v_i}(X) = X - \lambda_{v_j} \in F[X]$ (recordar ejemplo anterior cuando definimos polinomio minimal, en el caso de un autovector). Resulta entonces $m_A(X) = (X-\lambda_1)\dots(X-\lambda_r)$.

$\Leftarrow$) $m_A(X) = (X-\lambda_1)\dots(X-\lambda_r)$ con $\lambda_i \neq \lambda_j$ y $\lambda_i \in F$ para $i,j = 1,\dots,r$, todos los autovalores de A en F.

Vamos a ver que $F^n = \oplus_{i=1}^r E_{\lambda_i}$. En efecto, $E_{\lambda_i} = \{x \in F^n : Ax = \lambda_i x\}$.

Consideremos $v \in F^n$, y sea $k \in \mathbb{N}$ tal que $B_1 = \{v, Av, \dots, A^k v\}$ es li y $\{v, Av, \dots, A^k v, A^{k+1} v\}$ es ld. Sea $U = span(B_1)$. Tenemos así que $gr(m_v) = k+1$ y $m_v(X) = a_0 + a_1 X + \cdots + a_k X^k + X^{k+1}$. Además, si definimos $T : U \to U$ por $T(u) = Au$, como ya hemos observado, tenemos que

$$[T]_{B_1} = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & -a_{k-1} \\ 0 & 0 & \dots & 1 & -a_k \end{pmatrix}.$$

No perder de vista que $T \in L(U)$, es decir, es un endomorfismo del subespacio U. Sigue que $\chi_{T_A}(X) = det(XI_{k+1} - [T]_{B_1}) = a_0 + a_1 X + \cdots + a_k X^k + X^{k+1} = m_v(X)$. Como $\chi_T = m_v | m_A$ (¿porqué?) sigue que todas las raíces de χ_T son autovalores de A, luego son simples y están en el cuerpo F. Por la observación que hicimos antes, resulta que T es diagonalizable en U, de donde $\chi_T(X) = (X - \lambda_{i_1}) \dots (X - \lambda_{i_{k+1}})$. Como $v \in U$ existen únicos $v_{i_1}, \dots, v_{i_{k+1}}$ autovectores de T correspondientes $\lambda_{i_1}, \dots, \lambda_{i_{k+1}}$, respectivamente, tales que $v = v_{i_1} + \cdots + v_{i_{k+1}}$. Vale decir, $v \in \oplus_{j=1}^{k+1} E_{\lambda_{i_j}}$.

Como $v \in F^n$ es arbitrario, sigue que $F^n \subset \oplus_{i=1}^r E_{\lambda_i}$. Concluimos que A es diagonalizable.

$\square$

Ejemplos 3 *1. Sea $A \in \mathbb{C}^{n\times n}$ tal que $A^k = I$ para algún $k \in \mathbb{N}$. Veamos que A es diagonalizable. En efecto, $A^k = I \Rightarrow A^k - I = 0 \Rightarrow p(X) = X^k - 1$ anula a A. $\therefore m_A | p$. Como p tiene todas sus raíces en $\mathbb{C}$ y son simples, resulta que m_A también, de donde sigue que A es diagonalizable.*

2. Sea $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{n\times n}$. A es tal que $A^3 = I$, pero A no es diagonalizable: $m_A(X) = (X-1)(X^2 + X + 1)$ (EJERCICIO) no tiene todas sus raíces en $\mathbb{R}$.

5. Subespacios T-invariantes

Si E es un autoespacio, es claro que $T(E) \subset E$. Y si $V = \oplus E_i$, y T diagonalizable, hemos visto que podemos pensar en la restricción de T a cada autoespacio. Esta situación se generaliza mediante el concepto de sev T-invariante, y podemos replicar algunos resultados. La idea de descomponer un objeto de forma que se puedan estudiar las partes por separado, y recíprocamente, armar un objeto a partir de sus partes, se utiliza muchísimo en matemática. Por ejemplo, la rama del análisis armónico es básicamente esta idea tan simple en diversos contextos.

Para poder restringir una transformación $T \in L(V)$ en un subespacio $U \subset V$, necesitamos que $T|_U \in L(U)$. Lo cual nos lleva a la siguiente definición:

Definición 10 *V F-ev, $T \in L(V)$, $U \subset V$. Decimos que U es un subespacio T-invariante de V o que es invariante por T si $T(U) \subset U$.*

Ejemplo 6 *1. V, $\{\overline{0}\}$ son siempre T-invariantes.*

2. $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \in \mathbb{R}^{4\times 4}$. *Si* $T_A : \mathbb{R}^4 \to \mathbb{R}^4$ *y* $E = \{e_1, e_2, e_3, e_4\}$ *base canónica. Son invariantes por* T: $span\{e_1\}, span\{e_1, e_2\}, span\{e_3\}, span\{e_4\}, span\{e_3, e_4\}, span\{e_1, e_3, e_4\}$.

3. $A = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

   - *Sev invariantes de dimensión* 0: $\{\overline{0}\}$.
   - *Sev invariantes de dimensión* 2: $\mathbb{R}^2$.
   - *Sev invariantes de dimensión* 1: $U = span\{v\}$ *es* T_A*-invariante sii* $v \neq \overline{0}$ *y* $T_A v = Av \in U$ *sii* $v \neq \overline{0}$ *y existe* $\lambda \in F$ *tal que* $Av = \lambda v \in U$ *sii* v *es autovector de* A. *Luego* $U = span\{(0,1)\}$ *(EJERCICIO).*

Tenemos también los siguientes ejemplos de sev invariantes:

Proposición 15 *V F-ev, $T \in L(V)$.*

1. *$ker(T)$, $Im(T)$ son T-invariantes.*
2. *$U \subset V$, U es $T-$invariante de dimensión $dimU = 1$ sii $U = span\{v\}$ con v autovector de T.*
3. *U, W sev de V T-invariantes. Entonces $U \cap W$ y $U + W$ son sev T-invariantes.*

Demostración: EJERCICIO.

□

Estudiaremos a continuación el comportamiento de los polinomios minimal y característico en un sev invariante por T.

Proposición 16 *V F-ev, $dimV = n$, $T \in L(V)$, $U \subset V$ sev T-invariante. Si $T|_U : U \to U$ es la función restricción, entonces:*

1. $m_{T_U} | m_T$.
2. $\chi_{T_U} | \chi_T$.

Demostración: Sea $s = dimU$ y $B_U = \{v_1, \ldots, v_s\}$ base de U. Extendemos a una base $B = \{v_1, \ldots, v_s, v_{s+1}, \ldots, v_n\}$.

1. $m_{T_u} = m.c.m.\{m_{v_1}, \ldots, m_{v_s}\}$ y $m_T = m.c.m.\{m_{v_1}, \ldots, m_{v_n}\}$. Como $m_{v_i} | m_T$, tenemos que $m_{T|_U} | m_T$.

2. $[T]_B = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \in F^{n\times n}$, donde $A = [T|_U]_{B_U} \in F^{s\times s}$.

   Luego $\chi_T = \chi_{[T]_B} = det \begin{pmatrix} XI_s - A & -B \\ 0 & XI_{n-s} - C \end{pmatrix} = det(XI_s - A) det(XI_{n-s} - C) = \chi_{[T|_U]_{B_U}} Q(X)$.

   Sigue que $\chi_{[T|_U]_{B_U}} = \chi_{T|_U} | \chi_T$.

□

Sea V un F-ev de dimensión $dimV = n$, $T \in L(V)$ y $U, W \subset V$ son sev T-invariantes tales que $V = U \oplus W$ con $dimU = s, dimW = t$ y $B = B_U \cup B_W$. Tenemos que

$$[T]_B = \begin{pmatrix} [T|_U]_{B_U} & 0 \\ & \\ 0 & [T|_W]_{B_W} \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ & \\ 0 & A_2 \end{pmatrix}, \tag{1}$$

con $A \in F^{s \times s}$ y $A_2 \in F^{t \times t}$. Es decir, la forma matricial de la transformación T es diagonal por bloques, donde cada bloque corresponde a un sev invariante. Esto nos permite trabajar cómodamente con cada sev invariante. Esta observación conduce a la siguiente definición:

Definición 11 *V F-ev, $T \in L(V)$, $U \subset V$ sev T-invariante. Un* ***complemento invariante*** *para U es un sev $W \subset V$ tal que W es T-invariante y $U \oplus W = V$.*

Dado un sev invariante por T, éste no siempre tiene complemento invariante:

Ejemplo 7 *$T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (0, x)$. $U = span\{(0, 1)\}$ es T-invariante (EJERCICIO. Ayuda: calcular $kerT$). U no admite complemento invariante: si $W \subset \mathbb{R}^2$ es tal que $U \oplus W = \mathbb{R}^2$, entonces $dimW = 1$ y W es T-invariante, luego $W = span\{v\}$ con v autovector de T. Pero los únicos autovectores de T son los de U, luego debe ser $v \in U$, que es una contradicción.*

De tenerlo, ambos polinomios característico y minimal pueden ser calculados a partir de los característicos y minimales, respectivamente, de cada restricción a cada sev invariante:

Proposición 17 *V F-ev, $dimV = n$, $T \in L(V)$, $U, W \subset V$ sev T-invariantes tales que $U \oplus W = V$.Entonces:*

1. $\chi_T = \chi_{T|_U} \chi_{T|_W}$.
2. $m_T = m.c.m.\{m_{T|_U}, m_{T|_W}\}$.

Demostración: Sean $s = dimU$ y $t = dimW$.

1. EJERCICIO (Ayuda: releer página anterior).
2. Sea $p = m.c.m.\{m_{T|_U}, m_{T|_W}\}$. Como $m_{T|_U} | m_T$ y $m_{T|_W} | m_T$ tenemos que $p | m_T$. Además, de (1), como $m_{T_U} | p$, $p(A_1) = 0 \in F^{s \times s}$, $m_{T_W} | p$, $p(A_2) = 0 \in F^{t \times t}$ y luego $p([T]_B) = \begin{pmatrix} p(A_1) & 0 \\ & \\ 0 & p(A_2) \end{pmatrix} = 0 \in F^{n \times n}$. Resulta $m_T | p$, luego $m_T = p$.

$\square$