ISSN en trámite

Fascículo 1

Notas de matemática

Ursula Molter

Algunas notas sobre la forma de Jordan

Departamento de Matemática

Facultad de Ciencias Exactas y Naturales

Universidad de Buenos Aires

2014

Notas de matemática

Fascículo 1

ISSN en trámite

Universidad de Buenos Aires.
Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
Ciudad Universitaria – Pabellón I
(1428) Ciudad de Buenos Aires
Argentina.
http://www.dm.uba.ar
e-mail. secre@dm.uba.ar
tel/fax: (+54-11)-4576-3335

ALGUNAS NOTAS SOBRE LA FORMA DE JORDAN

URSULA M. MOLTER

1. Introducción

Este fascículo presenta una breve exposición de los resultados principales que permiten hallar la forma normal de Jordan asociada a una transformación lineal de un espacio vectorial en sí mismo.

El fascículo es autocontenido y supone el conocimiento del álgebra lineal elemental.

Deseo agradecer a Carlos Cabrelli, Liliana Gysin y Angel Larotonda por sus fructíferas críticas a versiones previas de estas notas.

Date: Original: Mayo de 1995 - Corregido: Octubre 2014.

2. Notación

Sean K el cuerpo de los números complejos o reales, sea $\mathbb{V}$ un K-espacio vectorial de dimensión n, y $T : \mathbb{V} \to \mathbb{V}$ una transformación lineal. Notaremos con $\boldsymbol{\mathcal{L}}(\mathbb{V},\mathbb{V})$ al espacio vectorial de todas las transformaciones lineales de $\mathbb{V}$ en $\mathbb{V}$, a saber:

$$\boldsymbol{\mathcal{L}}(\mathbb{V},\mathbb{V}) = \{T : \mathbb{V} \to \mathbb{V},\ T \text{ transformación lineal}\}.$$

Ejercicio 1. *Demostrar que la dimensión de* $\boldsymbol{\mathcal{L}}(\mathbb{V},\mathbb{V})$ *es* n^2.

Si $\mathcal{B} = \{b_1, \dots, b_n;\ b_i \in \mathbb{V}\}$ es una base de $\mathbb{V}$, denotaremos con $[T]_{\mathcal{B}}$ la matriz de la transformación T en la base $\mathcal{B}$,

$$T(b_j) = \sum_{i=1}^{n} ([T]_B)_{ij} b_i \quad ,$$

o sea, la columna j-ésima de la matriz contiene los coeficientes de $T(b_j)$ en la base $\mathcal{B}$.

Nuestro objetivo es hallar una base $\mathcal{B}$, en la cual la matriz de T sea lo más "sencilla" posible.

Para ello necesitamos algunas definiciones. Sean $A \in K^{n\times n}$ (una matriz de $n\times n$ con coeficientes en K) e I la matriz identidad en $K^{n\times n}$. El **polinomio característico** de A, $\chi_A(\lambda)$, se define como $\chi_A(\lambda) = \det(\lambda I - A)$.

Ejercicio 2. *Probar que* χ_A *es un polinomio de grado* n.

Recordemos la noción de semejanza entre matrices: $A, B \in K^{n\times n}$ son **semejantes** sii existe una matriz inversible $P \in K^{n\times n}$ tal que $A = P^{-1}BP$.

Ejercicio 3. *Si* A *es semejante a* B *entonces* $\chi_A(\lambda) = \chi_B(\lambda)$.

Tiene entonces sentido hablar del polinomio característico de una transformación lineal T, χ_T como el polinomio característico asociado a la matriz de T en alguna base $\mathcal{B}$, ya que $[T]_{\mathcal{B}}$ es semejante a $[T]_{\mathcal{B}'}$ para cualquier base $\mathcal{B}'$ de $\mathbb{V}$.

$$\chi_T(\lambda) = det(\lambda I - [T]_{\mathcal{B}}).$$

Indicaremos con T^k la transformación lineal $T^k : \mathbb{V} \to \mathbb{V}$, dada por

$$T^k(v) = T(T^{k-1}(v)), k = 2, 3, \dots \quad .$$

Si $p \in K[x]$ (es un polinomio con coeficientes en K, $p(x) = a_0 + a_1x + \cdots + a_sx^s$), notaremos $p(T)$ a la transformación lineal de $\boldsymbol{\mathcal{L}}(\mathbb{V},\mathbb{V})$ definida como

$$p(T)(v) = a_0I(v) + a_1T(v) + \cdots + a_sT^s(v) \ .$$

Consideremos en $\boldsymbol{\mathcal{L}}(\mathbb{V},\mathbb{V})$, el conjunto $I, T, T^2, \dots, T^{n^2}$. Este conjunto tiene $n^2 + 1$ elementos en un espacio de dimensión n^2, por lo tanto es linealmente dependiente; entonces existe una combinación lineal no trivial igual a 0, es decir, existen

$a_0, \ldots, a_{n^2} \in K$ no todos nulos, tales que

$$a_0 I + a_1 T + \cdots + a_{n^2} T^{n^2} = 0$$

donde 0 es la transformación nula. Esto equivale a afirmar que existe al menos un polinomio no nulo que anula a T. Tiene entonces sentido la siguiente definición:

Definición 2.1. *Si $T \in \mathcal{L}(\mathbb{V}, \mathbb{V})$, el* **polinomio minimal** *de T que notaremos con m_T, es el polinomio mónico de menor grado que anula a T.*

Observación: por los análisis anteriores, vimos que esta definición tiene sentido, pero falta probar que el polinomio minimal es único. La siguiente proposición será la parte esencial de esa demostración:

Proposición 2.1. *Si $p \in K[x]$ satisface que $p(T) = 0$, entonces m_T divide a p.*

Dem: Sea $p \in K[x]$ tal que $p(T) = 0$. Por el algoritmo de la división, existen q y $r \in K[x]$ tales que

1. $p = q\, m_T + r$
2. $r \equiv 0$ ó el grado de r es estrictamente menor que el grado de m_T.

Si r no fuera el polinomio 0, entonces por 1.

$$0 = p(T) = q(T)m_T(T) + r(T) = r(T)$$

por lo tanto $r(T)$ anula a T y es de grado menor que m_T lo cual es una contradicción. Por lo tanto $r \equiv 0$ y entonces $p = q\, m_T$ lo cual prueba nuestra proposición. □

Corolario 2.1. *El polinomio minimal es único.*

Ejercicio 4. *Demostrar el Corolario.*

Definición 2.2. *Si $W \subseteq \mathbb{V}$ es un subespacio de $\mathbb{V}$ diremos que W es* **T-invariante** *si $T(W) \subseteq W$.*

Observación: Claramente V y $\{0\}$ son subespacios T-invariantes.

Si W es un subespacio T-invariante, entonces se puede considerar la transformación lineal restringida a W

$$T/_W : W \to W, \quad T/_W(w) = T(w).$$

Es inmediata la siguiente proposición.

Proposición 2.2. *Si $W \subseteq \mathbb{V}$ es un subespacio invariante, entonces $m_{T/_W}$ divide a m_T.*

Dem: Es evidente que $m_T(T/_W) = 0$ y por lo tanto se puede aplicar la Proposición 2.1 a $T/_W$ y $m_{T/_W}$. □

Supongamos ahora que es posible descomponer al espacio vectorial $\mathbb{V}$ en suma directa de dos subespacios T-invariantes y observemos qué aspecto tiene la matriz de T en ese caso: Sean $W, W' \subseteq \mathbb{V}$ tales que $\mathbb{V} = W \oplus W'$ y W, W' son T-invariantes. Si $\mathcal{B}_1 = \{b_1, \dots, b_k\}$ es una base de W y $\mathcal{B}_2 = \{b_{k+1}, \dots, b_n\}$ es una base de W', entonces en la base $\mathcal{B} = \{b_1, \dots, b_n\} = \mathcal{B}_1 \cup \mathcal{B}_2$ la matriz de T en la base $\mathcal{B}$ será una matriz cuadrada de la forma

$$[T]_{\mathcal{B}} = \left[\begin{array}{cccc|ccc} * & \cdots & \cdots & * & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ * & \cdots & \cdots & * & 0 & \cdots & 0 \\ * & \cdots & \cdots & * & 0 & \cdots & 0 \\ \hline 0 & \cdots & \cdots & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & * & \cdots & * \end{array}\right]$$

Esto sugiere que el problema central para hallar una expresión sencilla para la matriz de una transformación lineal, será descomponer al espacio vectorial en suma directa de subespacios T-invariantes, cuando sea posible.

Proposición 2.3. *Sea $\mathbb{V}$ un K-espacio vectorial y sean $W, W' \subseteq \mathbb{V}$ tales que $\mathbb{V} = W \oplus W'$ y W, W' son T-invariantes entonces*

1. $\chi_T = \chi_{T/_W} \chi_{T/_{W'}}$
2. $m_T = mcm(m_{T/_W}, m_{T/_{W'}})$.

Dem:

1. Es una consecuencia directa de la propiedad del determinante.
2. Está claro que $p = mcm(m_{T/_W}, m_{T/_{W'}})$ divide a m_T.

   Por otro lado, si $v \in \mathbb{V}$, como $\mathbb{V} = W \oplus W'$, $v = v_1 + v_2$ con $v_1 \in W, v_2 \in W'$. Entonces $p(T)v = p(T)(v_1 + v_2) = p(T)(v_1) + p(T)(v_2)$. Como $v_1 \in W$ y $v_2 \in W'$,

   $$p(T)(v_1) + p(T)(v_2) = p(T/_W)v_1 + p(T/_{W'})v_2 = 0.$$

   Por lo tanto $p(T)v = 0$ para todo $v \in \mathbb{V}$, y luego por la Propoxición 2.2 m_T divide a p, lo cual completa nuestra demostración. $\square$

Nota 2.1. *Notemos que para 2. no hemos usado la invariancia de W y W' pero para 1. si. (Es necesaria la invariancia de ambos subespacios para 1. ?)*

Nota 2.2. *Es importante señalar en este punto, que dado un subespacio T-invariante, W, puede ser que no exista ningún subespacio complementario W' que sea T-invariante!!*

Ejemplo: Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación definida por la matriz: $[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, donde $\mathcal{B} = \{e_1, e_2\}$. Si $W = \text{gen}\{e_1\}$, $T(W) \subseteq W$, pero no existe ningún complemento invariante.

3. Transformaciones Nilpotentes

Analizaremos primero un tipo de transformaciones lineales especiales, que resultarán ser centrales para lograr nuestro objetivo:

Definición 3.1. *Una transformación lineal* $T : \mathbb{V} \to \mathbb{V}$ *se llamará* nilpotente de órden $k \geq 1$ *sii* $T^k = 0$ *y* $T^{k-1} \neq 0$*; donde* 0 *significa la transformación lineal nula.*

Ejercicios:

1. Encuentre T tal que $T^k \neq 0$ $\forall k$.
2. Si $T^k \neq 0$ $\forall k$ es T inversible?
3. Pruebe que si T es nilpotente de órden k, entonces T es singular.

Proposición 3.1. *Si* T *es nilpotente de órden* k*, entonces* $m_T = x^k$*.*

Dem: Como $T^k = 0$ está claro que x^k anula a T. Usando entonces la proposición 2.1, sabemos que m_T divide a x^k. Por lo tanto, como los únicos polinomios que dividen a x^k son de la forma x^j $j \leq k$,

$$m_T(x) = x^j \qquad j \leq k.$$

Si $j < k$ entonces $T^j = 0$ y por lo tanto $T^i = 0$ para todo $i \geq j$. Es decir $T^{k-1} = 0$ lo cual es una contradicción, ya que T es nilpotente de órden k; por lo tanto $j = k$.
□

Lema 3.1. *Si* T *es nilpotente de órden* k*, entonces existe* $v \in \mathbb{V}$ *tal que* $\{v, Tv, T^2v, \dots, T^{k-1}v\}$ *es un conjunto linealmente independiente.*

Dem: Como T es nilpotente de órden k, existe un $v \in V$ tal que $T^{k-1}v \neq 0$. Para ese v considero $\{v, Tv, T^2v, \dots, T^{k-1}v\}$. Sean $a_0, \dots, a_k$ tales que

$$a_0v + a_1Tv + \cdots + a_kT^{k-1}v = 0$$

aplicando T^{k-1} a ambos lados de la igualdad, obtengo que $a_0T^{k-1}v = 0$, y por lo tanto $a_0 = 0$. De la misma manera pruebo que $a_i = 0, 1 \leq i \leq k$ lo que completa la demostración. □

Corolario 3.1. *Si* T *es nilpotente de órden* k*, entonces* $k \leq n$*.*

Está claro que el subespacio $W \subseteq \mathbb{V}$ generado por $\mathcal{B} = \{v, Tv, T^2v, \dots, T^{k-1}v\}$ es T-invariante. Si tomamos esos vectores como base de W, observemos que la matriz de $T/_W$ resulta ser:

$$[T/_W]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ 0 & 0 & \cdots & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

Esta matriz se llama un **bloque de Jordan de** $k \times k$. Es una matriz de $k \times k$ y tiene todos "unos" inmediatamente debajo de la diagonal principal. Para el siguiente resultado, necesitaremos las nociones de *espacio vectorial dual* y *anulador* de un (sub)-espacio.

Definición 3.2. *Si $\mathbb{V}$ es un espacio vectorial de dimensión n, el espacio dual de $\mathbb{V}$, $\mathbb{V}^*$ se define por*

$$\mathbb{V}^* := \{\varphi : \mathbb{V} \to K, \textit{transformación lineal}\}.$$

Si $W \subset \mathbb{V}$ es un subespacio, el anulador *de W, se denota por W°, es el subespacio de $\mathbb{V}^*$ definido como*

$$W^\circ = \{\varphi \in \mathbb{V}^* : \varphi(w) = 0 \ \forall \ w \in W\}.$$

La siguiente proposición está contrapuesta a la Nota 2.2.

Proposición 3.2. *Si T es nilpotente de órden k, consideremos $W \subseteq \mathbb{V}$ el subespacio de dimensión k generado por $\{v_0, Tv_0, T^2v_0, \ldots, T^{k-1}v_0\}$ para algún v_0. Entonces existe $W' \subseteq \mathbb{V}$, T-invariante tal que $\mathbb{V} = W \oplus W'$.*

Nota 3.1. *O sea, si T es nilpotente de órden k, entonces el subespacio $\{v_0, Tv_0, T^2v_0, \ldots, T^{k-1}v_0\}$* siempre *tiene un complemento invariante!*

Dem:(de la proposición)
Observemos primero que para el caso en que $k = n$, la proposición es inmediata, ya que $W' = \{0\}$.

Supongamos ahora $k < n$.

Sea $\mathcal{B} := \{v_0, Tv_0, T^2v_0, \ldots, T^{k-1}v_0, b_{k+1}, \ldots, b_n\}$ una base de $\mathbb{V}$ obtenida completando la base de W a una base de $\mathbb{V}$. Considero $\varphi_{k-1} \in \mathbb{V}^*$ la transformación *dual* de $T^{k-1}v_0$, o sea φ_{k-1} queda definida por sus valores en la base $\mathcal{B}$ como

$$\varphi_{k-1}(b_j) = \begin{cases} 1 & b_j = T^{k-1}v_0 \\ 0 & \text{en otro caso.} \end{cases}$$

Además sean

$$\varphi_i = \varphi_{k-1} \circ T^{k-1-i}, \quad i = 0, \ldots, k-2.$$

Entonces $\{\varphi_0, \ldots, \varphi_{k-1}\}$ son linealmente independientes.

Sea $W^* = \text{gen}\{\varphi_0, \ldots, \varphi_{k-1}\}$ (notar que la dimensión de W^* es k). Sea $W' \subset \mathbb{V}$ el subespacio cuyo anulador es W^*, i.e.

$$W'^\circ := W^*.$$

Probaremos que W' es T-invariante y que $\mathbb{V} = W \oplus W'$.

- **Invariancia:** Sea $w \in W'$, entonces $\varphi_i(w) = 0, \ i = 0, \ldots, k-1$. Ahora bien, para $i = 1, \ldots, k-2$

$$\varphi_i(Tw) = \varphi_{k-1} \circ T^{k-1-i}(Tw) = \varphi_{k-1} \circ T^{k-1-i+1}(w) = \varphi_{i-1}(w) = 0$$

con lo que solamente resta ver que $\varphi_0(Tw) = 0$. Pero como T es nilpotente de órden k, $T^k \equiv 0$ y entonces

$$\varphi_0(Tw) = \varphi_{k-1} \circ T^{k-1}(Tw) = \varphi_{k-1}(T^k w) = \varphi_{k-1}(0) = 0.$$

- **Suma directa:** Como la dimensión de $W' = n - k$, (verificar) basta calcular la intersección. Sea $w \in W \cap W'$. Como $w \in W$, $w = a_0 v_0 + a_1 T v_0 + \cdots + a_{k-1} T^{k-1} v_0$, y por la definición de φ_i, tenemos que

$$\varphi_i(w) = \varphi_{k-1} \circ T^{k-1-i}(w) = a_i T^i v_0.$$

Pero como $w \in W'$, $\varphi_i(w) = 0 \ \forall\ i = 0, \ldots, k-1$. Esto implica que $w = 0$.

□

Definición: $W \subset \mathbb{V}$ es un subespacio irreducible, sii no existen $A \subset W$ y $B \subset W$ subespacios T-invariantes tales que $W = A \oplus B$.

Proposición 3.3. *Si $T : \mathbb{V} \to \mathbb{V}$ es un operador nilpotente de órden k, entonces un subespacio T-invariante W es un subespacio irreducible de dimensión $\ell \leq k$ si y sólo si existe $v_0 \in W$ tal que $W = gen\{v_0, Tv_0, \ldots, T^{\ell-1} v_0\}$.*

Dem: Sea W un subespacio irreducible de dimensión ℓ, entonces, como es T-invariante, por la Proposición 2.2, $m_{T/_W}(x) = x^s$ con $s \leq k$. Si $s < \ell$ entonces $T/_W$ es nilpotente de orden s y sea $w_0 \in W$ tal que $T^{s-1}(w_0) \neq 0$. Entonces $A := \mathrm{gen}\{w_0, Tw_0, T^2 w_0, \ldots, T^{s-1} w_0\} \subset W$ es $T/_W$-invariante y por la Proposición 3.2 existe $B \subset W$ tal que $W = A \oplus B$ contradiciendo la hipótesis de irreducibilidad de W. Por lo tanto $s = \ell$ y $W = \mathrm{gen}\{w_0, Tw_0, T^2 w_0, \ldots, T^{\ell-1} w_0\}$.

Para la otra implicación, si $W = \mathrm{gen}\{v_0, Tv_0, \ldots, T^{\ell-1} v_0\}$, entonces $\{v_0, Tv_0, \ldots, T^{\ell-1} v_0\}$ es una base de W y como W es T-invariante, $m_{T/_W} = x^\ell$. Supongamos que $W = A \oplus B$ con $A, B \subset W$ T-invariantes de dimensión s y k resc-pectivamente, con $s \leq k$. Entonces $m_{T/_A}(x) = x^s$ y $m_{T/_B}(x) = x^k$ respectivamente. Pero entonces, por la Proposición 2.3,

$$m_{T/_W} = \mathrm{mcm}(m_{T/_A}, m_{T/_B}) = x^k.$$

Entonces $k = \ell$ y por lo tanto $B = W$, y W es irreducible. □

Estamos ahora en condiciones de probar el teorema más importante de esta sección:

Teorema 3.1 (descomposición para operadores nilpotentes)**.** *Si $T : \mathbb{V} \to \mathbb{V}$ es un operador nilpotente de órden k, entonces existen $\ell \geq 1$ y ℓ subespacios $W_1, \ldots, W_\ell$ T-invariantes irreducibles, tales que*

1. $\mathbb{V} = W_1 \oplus W_2 \oplus \cdots \oplus W_{\ell-1} \oplus W_\ell$
2. $\dim(W_1) = k$ *y si* $\ell > 1$ $\dim(W_i) \geq \dim(W_{i+1}), 1 \leq i \leq \ell - 1$.

Además la descomposición es única en el sentido que, si existen $W_1', \ldots, W_s'$ subespacios T-invariantes que satisfacen las propiedades 1. y 2., entonces $s = \ell$ y $\dim(W_i') = \dim(W_i), 1 \leq i \leq \ell$.

Dem: La demostración se hará por inducción en n, donde n es la dimensión del espacio vectorial V.

Caso $n = 1$**:** es trivial, ya que la única transformación nilpotente es la transformación nula.

Prueba que $n-1 \Rightarrow n$ Supongamos que el teorema sea cierto para cualquier espacio vectorial de dimensión menor que n, queremos probar que vale para dimensón n.

Subcaso $k = n$ Si el órden de nilpotencia $k = n$, entonces por la Proposición 3.3 $\mathbb{V} = \text{gen}\{v_0, Tv_0, \ldots, T^{n-1}v_0\}$ es irreducible, y por lo tanto $\mathbb{V} = W_1$ y $\ell = 1$.

Subcaso $k < n$ Si el órden de nilpotencia es menor que la dimensión del espacio, sea v_0, tal que $\{v_0, Tv_0, \ldots, T^{k-1}v_0\}$ son linealmente independientes. Si $W_1 = \text{gen}\{v_0, Tv_0, \ldots, T^{k-1}v_0\}$, por la Proposición 3.2 sabemos que existe W', T-invariante, tal que $dim(W') = n - k(< n)$ y tal que $\mathbb{V} = W_1 \oplus W'$.

Como $m_{T/_{W'}}$ divide a m_T, $T/_{W'}$ es también nilpotente. El órden de nilpotencia, k' estará determinado por el grado del polinomio minimal - o sea $k' \leq k < n$. Por lo tanto, por hipótesis inductiva $W' = W_2 \oplus \cdots \oplus W_\ell$, con $dim(W_2) = k'$ y $dim(W_i) \geq dim(W_{i+1})$ $2 \leq i \leq \ell - 1$, lo cual prueba la primera parte de nuestro teorema.

Para probar la unicidad (note que si $k = n$ es trivial), observemos que si $\mathbb{V} = W_1 \oplus W_2 \oplus \cdots \oplus W_h \oplus \cdots \oplus W_\ell$ y además $\mathbb{V} = W_1' \oplus W_2' \oplus \cdots \oplus W_h' \oplus \cdots \oplus W_s'$ y estos subespacios son T-invariantes, entonces

$$T^j(\mathbb{V}) = T^j(W_1) \oplus T^j(W_2) \oplus \cdots \oplus T^j(W_h) \oplus \cdots \oplus T^j(W_\ell), \quad \forall \ j,$$

y análogamente

$$T^j(\mathbb{V}) = T^j(W_1') \oplus T^j(W_2') \oplus \cdots \oplus T^j(W_h') \oplus \cdots \oplus T^j(W_s'), \quad \forall \ j.$$

Sean $k_1 \geq k_2 \geq \cdots \geq k_\ell$ y $k_1' \geq k_2' \geq \cdots \geq k_s'$ los órdenes de nilpotencia de $T/_{W_i}$ y $T/_{W_i'}$ respectivamente (notemos que con esta notación $dim(W_i) = k_i$ y $dim(W_i') = k_i'$ respectivamente), y observemos que $k_1 = k_1' = k$, y por lo tanto $dim(W_1) = k = dim(W_1')$.

Supongamos que existe $h \geq 2$ tal que para $1 \leq i \leq h - 1$ vale que $dim(W_i) = dim(W_i')$.

Sean $k_1, \ldots, k_\ell$ y $k_1', \ldots, k_s'$ los órdenes de nilpotencia de $T/_{W_i}$ y $T/_{W_i'}$ respectivamente (notemos que con esta notación $dim(W_i) = k_i$ y $dim(W_i') = k_i'$ respectivamente).

Tenemos entonces

$$T^{k_h}(\mathbb{V}) = T^{k_h}(W_1) \oplus T^{k_h}(W_2) \oplus \cdots \oplus T^{k_h}(W_h) \oplus \cdots \oplus T^{k_h}(W_\ell), \tag{1}$$

y

$$T^{k_h}(\mathbb{V}) = T^{k_h}(W_1') \oplus T^{k_h}(W_2') \oplus \cdots \oplus T^{k_h}(W_h') \oplus \cdots \oplus T^{k_h}(W_s'). \tag{2}$$

Como $k_h \geq k_j$, si $j \geq h$, entonces $T^{k_h}(W_j) = 0$, para $j \geq h$ y por lo tanto en (1) queda

$$T^{k_h}(\mathbb{V}) = T^{k_h}(W_1) \oplus T^{k_h}(W_2) \oplus \cdots \oplus T^{k_h}(W_{h-1}).$$

Como W_i y W_i' son irreducibles ambos tienen bases de la forma $\{v_i, Tv_i, \ldots, T^{k_i}v_i\}$ y $\{v_i', Tv_i', \ldots, T^{k_i}v_i'\}$ respectivamente. Como estamos suponiendo que $dim(W_i) = dim(W_i') \geq k_h,\ 1 \leq i \leq h-1$, se ve que

$$\dim(T^{k_h}(W_i)) = k_i - k_h = k_i' - k_h = \dim(T^{k_h}(W_i')),\ 1 \leq i \leq h-1.$$

Entonces,

$$\begin{aligned} \dim(T^{k_h}(V) &= \sum_{i=1}^{h-1} \dim(T^{k_h}(W_i)) \\ &= \sum_{i=1}^{h-1} k_i - k_h = \sum_{i=1}^{h-1} k_i' - k_h \\ &= \sum_{i=1}^{h-1} \dim(T^{k_h}(W_i')) \quad , \end{aligned}$$

por lo que usando (2) se obtiene que $\dim(T^{k_h}(W_j')) = 0$, para todo $j \geq h$, o sea $T^{k_h}(W_j') = 0$. En particular, $T^{k_h}(W_h') = 0$. Como k_h' es el orden de nilpotencia de $T/_{W_h'}$, se tiene que $k_h' \leq k_h$.

Con un razonamiento análogo, se obtiene que $k_h \leq k_h'$ lo que prueba que $k_h = k_h'$. O sea, probamos que si $\dim(W_i) = \dim(W_i')$ para $1 \leq i \leq h-1$ entonces $\dim(W_h) = \dim(W_h')$, lo cual concluye nuestra prueba, ya que en particular concluimos que $\ell = s$. □

Corolario 3.2. *Si T es una transformación nilpotente, entonces $\chi_T(x) = x^n$, en particular $\chi_T(T) = 0$.*

Dem: basta elegir como base de V, $\mathcal{B} = B_1 \cup \cdots \cup B_\ell$, y hallar $[T]_{\mathcal{B}}$. Esta matriz es triangular inferior y tiene 0 en la diagonal principal, por lo tanto $\det(xI - [T]_B) = x^n$. □

Si en el teorema se elige en cada W_i un vector v_i y la base $\mathcal{B}_i = \{v_i, Tv_i, \ldots, T^{k_i}v_i\}$, entonces $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_\ell$ es una base de $\mathbb{V}$ y

$$[T]_{\mathcal{B}} = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & A_\ell \end{bmatrix}$$

donde A_j es una matriz de $k_j \times k_j$ tal que

$$A_j = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ 0 & 0 & \cdots & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

o sea hemos conseguido una expresión sencilla para nuestra transformación nilpotente T. Esta representación se llama **forma normal de Jordan** de T; y $\mathcal{B}$ se llama **base de Jordan** de T. Note que la forma de Jordan es única, pero la base no!

4. Transformaciones lineales cualesquiera

En lo que sigue se verá como se procede en el caso general de una transformación lineal cualquiera. Probaremos primero algunos lemas, para luego concluir con el teorema principal.

Lema 4.1. *Si p es un polinomio con coeficientes en K, entonces $W = \mathrm{Ker}(p(T))$ es un subespacio T-invariante.*

Dem: Sea $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_s x^s$, entonces, $v \in \mathrm{Ker}(p(T)) \Rightarrow p(T)v = 0$.

Para ver que es T-invariante, observamos que

$p(T)(Tv) = a_0 Tv + a_1 T^2 v + \cdots + a_s T^{s+1} v = T(a_0 v + a_1 Tv + \cdots + a_s T^s v) = T(0) = 0,$

lo que completa la demostración. □

Lema 4.2. *Si p y q son polinomios, entonces $p(T)q(T) = q(T)p(T)$.*

Dem: sea m el polinomio $m = p\ q = q\ p$. Entonces

$$p(T)q(T) = m(T) = q(T)p(T),$$

como queríamos probar. □

Proposición 4.1. *Si $T : \mathbb{V} \to \mathbb{V}$ es una transformación lineal tal que $m_T = p\ q$ con p y q polinomios coprimos, entonces*

$$\mathbb{V} = \mathrm{Ker}(p(T)) \oplus \mathrm{Ker}(q(T)) \text{ , y si notamos por } T_p = T/_{\mathrm{Ker}(p(T))}$$
$$\text{entonces } m_{T_p} = p \quad m_{T_q} = q\ .$$

Dem: Toda la demostración radica en el hecho que el polinomio 1 se puede escribir como combinación lineal de p y q en $K[x]$, es decir, existen polinomios m y n tales que

$$1 = m\ p + n\ q. \tag{3}$$

de lo cual se deduce inmediatamente

$$Id = m(T)p(T) + n(T)q(T). \tag{4}$$

Entonces, si $v \in \mathbb{V}$,

$$v = m(T)p(T)v + n(T)q(T)v. \tag{5}$$

Llamemos ahora $v_1 = m(T)p(T)v$ y $v_2 = n(T)q(T)v$.

Probaremos que $v_1 \in \operatorname{Ker}(q(T))$ y $v_2 \in \operatorname{Ker}(p(T))$ con lo cual $\mathbb{V} = \operatorname{Ker}(q(T)) + \operatorname{Ker}(p(T))$.

Pero

$$q(T)v_1 = q(T)m(T)p(T)v = m(T)p(T)q(T)v = m(T)m_T(T)v = 0.$$

Análogamente se obtiene que $p(T)v_2 = 0$. Note que hemos usado reiteradamente el Lema 4.2.

Para ver que la suma de los subespacios es en realidad suma directa, supongamos que $v \in \operatorname{Ker}(p(T)) \cap \operatorname{Ker}(q(T)$. Usando (4)

$$v = m(T)p(T)v + n(T)q(T)v = 0.$$

Lo cual completa la demostración que $\mathbb{V} = \operatorname{Ker}(p(T)) \oplus \operatorname{Ker}(q(T))$.

Para ver quiénes son los polinomios minimales de la transformación restringida a cada uno de los subespacios invariantes, observemos que:

$$p(T_p) = 0$$

entonces el minimal debe dividir a p. Supongamos que $\operatorname{grado}(m_{T_p}) < \operatorname{grado}(p)$. Como

$$m_{T_p}(T)q(T) = 0$$

y el polinomio $m_{T_p}q$ tiene grado estrictamente menor que m_T, obtenemos una contradicción, ya que m_T es el polinomio minimal. Por lo tanto el grado de m_{T_p} deberá ser igual al grado de p y por lo tanto son iguales (recuerden que el polinomio minimal es mónico). □

Estamos ahora en condiciones de demostrar el teorema principal de esta sección:

Teorema 4.1 (De la descomposición primaria)**.** *Si $T : \mathbb{V} \to \mathbb{V}$ es una transformación lineal tal que*

$$m_T = p_1^{e_1} \cdots p_k^{e_k}$$

donde p_i son polinomios mónicos irreducibles en $K[x]$, entonces

$$\mathbb{V} = \operatorname{Ker}(p_1(T)^{e_1}) \oplus \cdots \oplus \operatorname{Ker}(p_k(T)^{e_k})$$

y $\dim(\operatorname{Ker}(p_i(T)^{e_i})) = d_i \geq e_i$.

Dem: Se hará por inducción en k =número de polinomios irreducibles que dividen a m_T. Si $k = 2$, se aplica la proposición anterior.

Supongamos ahora que

$$m_T = p_1^{e_1} \cdots p_{k-1}^{e_{k-1}} p_k^{e_k}$$

y llamemos $q = p_1^{e_1} \cdots p_{k-1}^{e_{k-1}}$, y $p = p_k^{e_k}$ entonces $m_T = qp$ y por la proposición anterior

$$\mathbb{V} = \operatorname{Ker}(q(T)) \oplus \operatorname{Ker}(p(T)) = \operatorname{Ker}(q(T)) \oplus \operatorname{Ker}(p_k(T)^{e_k})$$

y los polinomios minimales son q y $p_k^{e_k}$ respectivamente. Pero entonces, q es el polinomio minimal de $T/_{\mathrm{Ker}(q(T))}$ y tiene $k-1$ factores, por lo tanto por hipótesis inductiva,

$$\mathrm{Ker}(q(T)) = \mathrm{Ker}(p_1(T)^{e_1}) \oplus \cdots \oplus \mathrm{Ker}(p_{k-1}(T)^{e_{k-1}})$$

con lo cual

$$\mathbb{V} = \mathrm{Ker}(p_1(T)^{e_1}) \oplus \cdots \oplus \mathrm{Ker}(p_k(T)^{e_k}).$$

Notemos ademas que las transformaciones

$$T_i = p_i(T/_{\mathrm{Ker}(p_i(T)^{e_i})}), \quad T_i : \mathrm{Ker}(p_i(T)^{e_i}) \to \mathrm{Ker}(p_i(T)^{e_i})$$

son transformaciones lineales nilpotentes, de órden de nilpotencia e_i, y por lo tanto la dimensión de $\mathrm{Ker}(p_i(T)^{e_i}) = d_i$ es mayor o igual que e_i. $\square$

4.1. Caso $K = \mathbb{C}$.

Si $K = \mathbb{C}$ (o más en general, cualquier cuerpo algebraicamente cerrado), entonces $p_i(x) = x - \lambda_i$, en cuyo caso

$$T_i = T - \lambda_i \mathrm{Id},$$

con T_i nilpotente de órden e_i. Si la descomposición nilpotente de T_i es $W_i = W_1^{(i)} \oplus \cdots \oplus W_{\ell_i}^{(i)}$ con dimensiones $k_1 \geq k_2 \geq k_{\ell_i}$, entonces la matriz de $T_{p_i} = T_i + \lambda_i \mathrm{Id}$ es en la base adecuada de la forma

$$[T_{p_i}] = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & A_{\ell_i} \end{bmatrix}$$

donde A_j es una matriz de $k_j \times k_j$,

$$A_j = \begin{bmatrix} \lambda_i & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & \lambda_i & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \lambda_i & \cdots & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ 0 & 0 & \cdots & 1 & \lambda_i & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & \lambda_i & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & \lambda_i \end{bmatrix}$$

o sea

$$[T] = \begin{bmatrix} [T_{p_1}] & 0 & \cdots & 0 \\ 0 & [T_{p_2}] & 0 & \cdots \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & [T_{p_k}] \end{bmatrix}$$

es la forma de Jordan de T.

Hemos entonces hallado una descomposición del espacio en subespacios T-invariantes, y en cada uno de ellos tenemos una transformación nilpotente que hemos analizado en la sección anterior. Por lo tanto podemos hallar, en cada componente invariante la forma de Jordan de la transformación nilpotente.

Observación: Notar que es fundamental que $K = \mathbb{C}$, ya que si $K = \mathbb{R}$, entonces p_i puede ser de grado 1 ó de grado 2. Para el caso que p_i sea de grado 2, si bien puedo analizar la descomposición de $t_i = p_i(T)$ **no** puedo despejar T_{p_i}.

Corolario 4.1. *Sea $\mathbb{V}$ un $\mathbb{C}$ espacio vectorial. Sea $T : \mathbb{V} \to \mathbb{V}$ una transformación lineal. Entonces existe $\mathcal{B}$ tal que $[T]_{\mathcal{B}} = D + N$ donde D es una matriz diagonal y N es una matriz nilpotente y satisfacen $DN = ND$.*

Ejercicio 5. *Demostrar el Corolario.*

4.2. Teorema de Cayley-Hamilton.

En esta sección volvemos a considerar V como un K espacio vectorial, donde K puede ser tanto $\mathbb{C}$ ó $\mathbb{R}$.

Teorema 4.2 (Cayley-Hamilton). *Si $T : \mathbb{V} \to \mathbb{V}$ es una transformación lineal, entonces $\chi_T(T) = 0$.*

Dem: Con la notación anterior, sea $\mathbb{V} = \mathrm{Ker}(p_1(T)^{e_1}) \oplus \cdots \oplus \mathrm{Ker}(p_k(T)^{e_k})$, entonces por la Proposición 2.3,

$$\chi_T = \chi_{T_{p_1}} \cdots \chi_{T_{p_k}}.$$

Si consideramos que estamos trabajando en el cuerpo complejo, usando la forma de Jordan para el operador se ve inmediatamente que $\chi_{T_{p_i}} = (x - \lambda_i)^{d_i}$, con lo cual:

$$\chi_T(x) = (x - \lambda_1)^{d_1}(x - \lambda_2)^{d_2} \cdots (x - \lambda_k)^{d_k} \quad \text{o sea}$$

$$\chi_T(x) = p_1(x)^{d_1} p_2(x)^{d_2} \cdots p_k(x)^{d_k}$$

y como, por la proposición anterior $d_i \geq e_i$, $\chi_T(T) = 0$.

Si queremos analizar lo que pasa en los reales, consideremos el polinomio característico de T. Ese polinomio tendrá coeficientes reales, pero en particular lo puedo pensar como un polinomio en los complejos y entonces, por el análisis anterior, se sabe que $\chi_T(T) = 0$. □

En general, calcular el polinomio minimal es muy difícil, pero usando los teoremas anteriores, se puede deducir que

Proposición 4.2. *m_T y χ_T tienen las mismas raíces.*

Dem: Como (por Cayley-Hamilton) $\chi_T(T) = 0$, entonces m_T divide a χ_T. Por lo tanto toda raiz de m_T es raiz de χ_T.

Por otro lado, si λ es raiz de χ_T entonces λ es autovalor de T, o sea existe $v_\lambda \neq 0 \ \in \mathbb{V}$ tal que $(T - \lambda I)v_\lambda = 0$. Ahora bien, dividiendo a m_t por $(x - \lambda)$, tenemos que

$$m_T(x) = (x - \lambda)q(x) + r, \quad \text{con } r = \text{cte ó } r = 0.$$

Pero

$$0 = m_T(T) = (T - \lambda I)q(T) + rI \ ,$$

en particular

$$0 = m_T(T)v_\lambda = ((T - \lambda I)q(T) + rI)v_\lambda = rv_\lambda \ ,$$

lo cual implica que $r = 0$, o sea λ es raiz de m_T. □

5. Algunos ejercicios de aplicaciones

1. Sean
$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad, B = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}.$$
Muestre que $AB \neq BA$. Calcular e^A, e^B, e^{A+B} y mostrar que $e^{A+B} \neq e^A e^B$.
2. Mostrar que $\det(e^A) = e^{\sum \lambda_i}$ donde los λ_i son los autovalores de A (contados con toda su multiplicidad).
3. Mostrar que e^A siempre es inversible y que $(e^A)^{-1} = e^{-A}$ para toda matriz $A \in \mathbb{C}^{n\times n}$.
4. Resolver los siguientes sistemas de ecuaciones:
   a)
$$\left\{ \begin{array}{rcl} \frac{dy_1}{dt} & = & -y_1 + y_2 \\ \frac{dy_2}{dt} & = & -y_1 - 3y_2 \end{array} \right. \qquad y_1(0) = 0, y_2(0) = 1.$$
   b)
$$\left\{ \begin{array}{rcl} \frac{dy_1}{dt} & = & 5y_1 + 8y_2 + 16y_3 \\ \frac{dy_2}{dt} & = & 4y_1 + y_2 + 8y_3 \\ \frac{dy_3}{dt} & = & -4y_1 - 4y_2 - 11y_3 \end{array} \right. \qquad y_1(0) = y_2(0) = y_3(0) = 1.$$
5. Sea A la matriz de coeficientes de la ecuación diferencial vectorial equivalente a
$$\frac{d^n y}{dt^n} + a_{n-1}\frac{d^{n-1}y}{dt^{n-1}} + \dots a_0 y = 0.$$
Mostrar que
$$\chi_A = x^n + a_{n-1}x^{n-1} + \cdots + a_0.$$

Referencias

1. Kenneth Hoffman and Ray Kunze, *Linear algebra*, Second edition, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971. MR 0276251 (43 #1998)
2. R. Horn and C. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1985.