Universidad del Bío-Bío

Facultad de Educación y Humanidades

Escuela de Pedagogía en Educación Matemática

La Forma de Jordan

Autores: Pablo César Molina Molina
Paula Sinttia Verdugo Hernández

Profesor Guía: Roberto Carlos Cabrales

SEMINARIO PARA OPTAR AL TÍTULO DE PROFESOR DE ENSEÑANZA MEDIA EN EDUCACIÓN MATEMÁTICA.

Chillán, Diciembre de 2007

Agradecimentos

A nuestras familias y amigos, por acompañarnos incondicionalmente en nuestros años de formación personal y profesional.

Un reconocimiento especial a los profesores que nos transmitieron sus conocimientos y el amor por nuestra profesión.

Y por sobre todo a Dios, que nos acompaña y nos da la fuerza y entereza para continuar en el camino de nuestra realización personal

Resumen

Nuestro estudio versa sobre la Forma de Jordan. Este consiste en encontrar valores y vectores propios de una transformación lineal utilizando para ello los procesos de diagonalización, pero estos operadores en rigor no son diagonalizables, por consiguiente lo que debemos hacer es descomponer el espacio en suma de subespacios invariantes.

Para lograrlo, dividimos el texto en cuatro capítulos. En primer lugar estudiamos conceptos básicos del álgebra que servirán de apoyo al lector y a nosotros, entregando bases sólidas a nuestro estudio.

En el segundo capítulo comenzamos el estudio de transformaciones lineales, fortaleciendo aún más los cimientos de nuestro trabajo, comenzando a dar los primeros pasos hacia lo que realmente nos interesa, entender la Forma de Jordan.

En el capítulo 3, con el paso más firme, damos los primeros indicios de lo que nos espera en el capítulo 4. Es así como se nos hace necesario estudiar polinomios, valores y vectores propios, sobre todo entender sumas directas y sumas directas invariantes, de este modo comenzamos a dar mayor sustento y seguridad a nuestro caminar.

En el capítulo 4 analizamos la Forma de Jordan y sus dificultades, por ello que en primera instancia se da una visión simplificada, para luego estudiar el teorema de descomposición cíclica,terminando con el estudio de la Forma de Jordan en un nivel más elevado, cumpliendo de esta forma nuestros objetivos iniciales.

Abstract

Our study is about the Jordan's Form. This consists of finding the eigenvalues and eigenvectors of a linear transformation using diagonalization, even though these operators are not diagonalizable; therefore, we need to decompose the space into invariant subspace sums.

To do so, we divided the text into four chapters. First, we studied the basic concepts of algebra that will lead the reader and us too, thus, giving strong bases to our study.

On the second chapter, we started the study of linear transformations, strengthening this way the bases of our study, which is trying to understand Jordan's Form.

On the third chapter, we gave some hints of what chapter four is going to be about. That is why we dealt with polynomials, eigenvalues and eigenvectors; we needed to understand direct sums and invariant direct sums, so that we could start to give a strong support and strength to our study.

On chapter four, we analyzed Jordan's Form and its difficulties, that explains why we give a simplified version, later we study the cyclic decomposition theorem, and finally we get to study Jordan's Form on a higher level.

ÍNDICE GENERAL

Introducción

Los objetivos propuestos al iniciar nuestra indagación matemática fueron: Investigar la formulación y planteamientos matemáticos jordanos. Establecer ecuaciones algebraicas y sus implicacias operacionales. Y por último facilitar la comprension y el análisis de las formas de Jordan

Cuando se estudian formas de Jordan, un resultado importante es encontrar una base de un espacio vectorial tal que la matriz que representa a un operador lineal sea semejante a una matriz diagonal, de este modo se vuelve sencillo encontrar los valores y vectores propios del operador. Cabe preguntarse, ¿qué pasa cuando un operador no es diagonalizable?, ¿cómo encontrar valores y vectores propios?. Estudiando la Forma de Jordan damos respuesta a estas interrogantes, pues podemos hacer que este operador lineal o la matriz que lo representa sea semejante a otras cuyos componentes de la diagonal principal sean bloques y así encontrar los valores y vectores propios de este operador que, en rigor no es diagonalizable, así que lo hacemos es descomponer a tal espacio vectorial en subesapcios invariantes y la matriz que representa al operador restricción, estas matrices serán justamente los bloques de los que hablabamos.

Por ello es de gran importancia encontrar la matriz que representa a una transformación lineal T, dada una base ordenada, por ejemplo $\mathfrak{B} = (1,0,0),(0,1,0),(0,0,1)$ del espacio vectorial V, en la cual $[T]_{\mathfrak{B}}$ es la matriz en la cual está representada.

$$[T]_{\mathfrak{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Para cumplir con nuestro principal objetivo, estudiaremos algunos teoremas de gran importancia en el álgebra lineal. Teorema de descomposición Prima, Teorema de descomposición cíclica y sumas directas, estos son los más relevantes debido a que cuando ya no seamos capaces de encontrar de manera sencilla los valores y vectores propios de cualquier transformación lineal, tendremos que dividir al espacio vectorial en suma de subespacios invariantes, de esta forma las matrices que representan al operador restricción, serán los bloques que se encuentran en la diagonal de la matriz de Jordan. Por lo dicho anteriormente es que creemos necesario escribir estos teoremas y alguna definición que pueda ser importante.

Definición 0.1 *Sean* $\{W_1, ..., W_k\}$ *subespacios de un espacio vectorial* V*. Se dice que* $W_1, ..., W_k$ *son* ***independientes*** *si*

$$v_1 + ... + v_k = 0, \qquad v_i \text{ en } W_i$$

implica que cada v_i *es 0.*

Teorema 0.1 (Teorema de descomposición Prima) *Sea* T *un operador lineal sobre el espacio vectorial* V *de dimensión finita sobre el cuerpo* F*. Sea* p *el polinomio minimal de* T*,*

$$p = p_1^{r_1}...p_k^{r_k}$$

donde los p_i *son polinomios mónicos irreducibles distintos sobre* F*, y los* r_i *son enteros positivos. Sea* W_i *el espacio nulo de* $p_i(\mathrm{T})^{r_i}, \mathrm{i} = 1, ..., \mathrm{k}$*. Entonces*

(i) $V = W_1 \oplus \cdots \oplus W_k$*;*

(ii) cada E_i *es invariante por* T*;*

(iii) si T_i *es el operador inducido sobre* W_i *por* T*, entonces el polinomio minimal de* T_i *es* $p_i^{r_i}$*.*

Teorema 0.2 (Teorema de descomposición cíclica) *Sea* T *un operador lineal sobre un espacio vectorial de dimensión finita* V *y sea* W_0 *un subespacio propio* T*-admisible de* V*. Existen vectores no nulos* $v_1, ..., v_r$ *en* V *con* T*-anuladores respectivamente* $p_1, ..., p_r$*, tales que*

(i) $V = W_0 \oplus Z(v_1; T) \oplus ... \oplus Z(v_r; T)$*:*

(ii) p_k *divide a* $p_{k-1}, k = 2, ..., r$*.*

Estos resultados harán más comprensible lo que deseamos entender y analizar. Un ejemplo de la forma en la que encontraremos la matríz de Jordan es la siguiente:

$$A = \begin{pmatrix} A_1 & 0 & ... & 0 \\ 0 & A_2 & ... & 0 \\ . & . & ... & . \\ 0 & 0 & ... & A_k \end{pmatrix}$$

Así nuestro trabajo se restringe a encontrar los valores y vectores propios de estas partes más pequeñas del espacio vectorial, que son los subespacios invariantes T-cíclicos.

Para llegar a este resultado es necesario un metódico proceso, estudiando las propiedades de todas las piezas que forman parte de la Forma de Jordan. Desde los enfoques algebraicos y geométricos clásicos hasta varias pruebas elementales que han sido presentadas por diversos autores. Por consiguiente nos enfocaremos sólo en la forma algebraica de Jordan, presentandolo de forma clara y sencilla.

CAPÍTULO 1

Preliminares

Antes de comenzar con el estudio más profundo del tema que nos compete, es importante que el lector reconsidere o más bien reestudie algunos conceptos de Álgebra Lineal necesarios en nuestro estudio.

Recordaremos algunas definiciones y teoremas que serán útiles en nuestro análisis. Es por esto que el **Capítulo I** no tendrán muchas demostraciones escritas, pero si el lector lo desea podrá recurrir a los textos que serán recomendados a lo largo del trabajo (ver en [1], [2] y [3] especialmente).

1.1. Cuerpos

En gran parte del presente texto, las propiedades algebráicas que se utilizarán se deducen fácilmente de la siguente lista de propiedades de la adición y de la multiplicación.

Definición 1.1 *Sea F un conjunto no vacio. Se dice que F es un cuerpo, si existen definidas dos operaciones en F llamadas suma y multiplicación que satisfacen las siguientes propiedades.*

$$(F,+)=\begin{cases} +\textit{es asociativa; Si para todo} a,b,c\in F,\ \textit{se tiene que } (a+b)+c=a+(b+c),\\ \textit{existe un neutro, denotado por } 0,\ \textit{tal que para todo } a\in F,\ \ a+0=0+a=a,\\ \textit{si } a\in F,\ \textit{entonces existe } (-a)\in F,\ \textit{tal que } a+(-a)=0,\\ +\ \textit{es conmutativa; Si para todo } a,b\in F,\ \textit{se tiene que } a+b=b+a, \end{cases}$$

$$(F,\cdot)=\begin{cases} \cdot\ \textit{es asociativa; Si para todo } a,b\in F,\ \ (a\cdot b)\cdot c=a\cdot(b\cdot c),\\ \textit{existe un neutro, denotado 1 tal que para todo } a\in F,\ \ a\cdot 1=1\cdot a=a,\\ \textit{si } a\in F,\ \textit{y } a\neq 0,\ \textit{existe } a^{-1}\ \textit{tal que } a\cdot a^{-1}=1. \end{cases}$$

Un ejemplo de cuerpo son los números reales $\mathbb{R}$ y los complejos $\mathbb{C}$. A lo largo de este trabajo, la letra F denotará un cuerpo.

Un **subcuerpo** C de F es un conjunto C que es a su vez un cuerpo respecto de las operaciones de adición y multiplicación de números en F. Esto significa que el 0 y el 1 están en el conjunto C, y que si x e y son elementos de C, también lo son $(x+y)$, $-x$, xy, e x^{-1} (si $x \neq 0$). Un ejemplo de un subcuerpo semejante es el cuerpo $\mathbb{R}$ de los números reales. En efecto, si se identifican los números reales con los números complejos $(a+ib)$ para los que $b=0$, el 0 y el 1 del cuerpo complejo son números reales y si x e y son reales, también lo son $(x+y)$, $-x$, xy y x^{-1}. Lo peculiar de los subcuerpos en nuestro estudio es esencialmente lo siguiente: si se está operando con escalares que forman un cierto subcuerpo de F, entonces la ejecución de las operaciones de adición sustracción multiplicación división con estos escalares no se salen del subcuerpo dado.

Se debe tener esto en consideración debido a que en todo el resto del trabajo F denota un cuerpo.

1.2. Matrices y Determinantes

Definición 1.2 (Matriz) *Sea F un cuerpo, una matriz A de orden $m \times n$ es un arreglo rectangular de $m \cdot n$ elementos de F dispuestos en m filas y n columnas como se muestra a continuación:*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ . & . & . & & . \\ a_{m1} & a_{m2} & a_{2m3} & \cdots & a_{mn} \end{pmatrix}$$

Si $m = n$, se dice que A es una matriz cuadrada.

Deseamos ahora considerar operaciones sobre las filas de la matriz A que correspondan a la formación de combinaciones lineales de las ecuaciones del sistema $AX = Y$. Se limitará nuestra atención a tres operaciones elementales de filas en una matriz $m \times n$, sobre el cuerpo F:

i. Multiplicación de una fila de A por un escalar α no nulo;

ii. Reemplazo de la r-ésima fila de A por la fila r más α veces la fila s, donde α es cualquier escalar y r $\neq$ s;

iii. Intercambio de dos filas de A.

Ilustraremos el proceso de eliminación con el siguiente ejemplo.

Ejemplo 1.3 *Solución de un sistema de tres ecuaciones con tres incógnitas. Resolvamos el sistema*

$$\begin{aligned} 2x_1 + 4x_2 + 6x_3 &= 18 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ 3x_1 + x_2 - 2x_3 &= 4 \end{aligned}$$

Solución:

En este caso se buscan tres números x_1, x_2, x_3 tales que las tres ecuaciones en (1) se satisfacen. El método de solución que se estudiará será el de simplificar las ecuaciones de manera que las soluciones se puedan identificar de inmediato. Se comienza por dividir la primera ecuación entre dos. Esto da

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ 3x_1 + x_2 - 2x_3 &= 4 \end{aligned} \tag{1.1}$$

al sumar dos ecuaciones se obtiene una tercera ecuación correcta. Esta ecuación puede sustituir a cualquiera de las dos ecuaciones del sistema para obtenerla. Primero se simplifica el sistema (1.1) multiplicando ambos lados de la primera ecuación de (1.1) por -4 y sumando esta nueva ecuación a la segunda. Esto da

$$\begin{aligned} -4x_1 - 8x_2 - 12x_3 &= 36 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ -3x_2 - 6x_3 &= -12 \end{aligned}$$

La ecuación $-3x_1 - 6x_2 = -12$ es la nueva segunda ecuación y el sistema ahora es

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ -3x_2 - 6x_3 &= -12 \\ 3x_1 + x_2 - 2x_3 &= 4 \end{aligned}$$

Nota: Se ha sustituido la ecuación $4x_1+5x_2+6x_3 = 24$ por la ecuación $-3x_2-6x_3 = -12$. En este ejemplo se sustituirán ecuaciones con otra más sencillas hasta obtener un sistema cuya solución se pueda identificar de inmediato.

Entonces, la primera ecuacíón se multiplica por -3 y se suma a la tercera:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ -3x_2 - 6x_3 &= -12 \\ -5x_2 - 11x_3 &= -23 \end{aligned} \tag{1.2}$$

Observemos que en el sistema (1.2) se ha eliminado la variable x_1 de la segunda y tercera ecuación. Después se divide la segunda ecuación por -3

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ x_2 + 2x_3 &= 4 \\ -5x_2 - 11x_3 &= -23 \end{aligned}$$

Se multiplica la segunda ecuación por -2 *y se suma a la primera; después se multiplica la segunda ecuación por* 5 *y se suma a la tercera:*

$$\begin{aligned} x_1 - x_3 &= 1 \\ x_2 + 2x_3 &= 4 \\ -x_3 &= -3 \end{aligned}$$

Se multiplica la tercera ecuación por -1*:*

$$\begin{aligned} x_1 - x_3 &= 1 \\ x_2 + 2x_3 &= 4 \\ x_3 &= 3 \end{aligned}$$

Por último, se suma la tercera ecuación a la primera y después se multiplica la tercera ecuación por -2 *y se suma a la segunda para obtener el siguiente sistema*

$$\begin{aligned} x_1 &= 4 \\ x_2 &= -2 \\ x_3 &= 3 \end{aligned}$$

Ésta es la solución única para el sistema, en el cual aplicamos el proceso de eliminación, que era lo que queríamos mostrar.

Terminaremos esta primera parte con la siguiente definición;

Definición 1.4 (Rango) *Sea* A *una matriz de* $m \times n$*. Entonces el rango de la matriz* A *es el número de filas no nulas que quedan luego de hacer el proceso de eliminación. Este concepto equivale al número de filas linealmente independientes (LI) de* A*.*

Cuando comencemos el estudio de las transformaciones lineales, el concepto de rango será de gran utilidad, en el Capítulo III donde comenzamos el estudio del principal resultado de este trabajo.

Definición 1.5 (Matriz Identidad) *La matriz identidad de orden* n I_n *es una matriz de* $n \times n$ *cuyos elementos en la diagonal principal son iguales a 1 y todos los demás son 0. Esto es:*

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ . & . & . & & . \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

Es necesario recordar que la matriz identidad es el neutro de las matrices, es decir:

$$AI_n = I_nA = A$$

Definición 1.6 (Inversa de una Matriz) *Decimos que una matriz $A \in M_{n\times n}(F)$ es invertible o que tiene inversa si existe $B \in M_{n\times n}(F)$ tal que*

$$AB = BA = I_n$$

B se llama la inversa de A y se denota A^{-1}

El estudio de la matriz inversa es de gran importancia, pues nos dará importantes claves en el tratamiento de las transformaciones lineales que son la base para el desarrollo de este escrito.

Definición 1.7 (Transpuesta de una Matriz) *Sea $A = (a_{ij})$una matriz de $m \times n$, entonces la transpuesta de A, que se escribe A^T, es la matriz de $n\times m$ obtenida de intercambiar los filas por las columnas de A. es decir, $A^T = (a_{ji})$, en otras palabras :*

$$\textit{Si } A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ . & . & . & & . \\ a_{m1} & a_{m2} & a_{2m3} & \cdots & a_{mn} \end{pmatrix}$$

entonces, para calcular A^T, basta colocar

$$A^T = \begin{pmatrix} a_{11} & a_{21} & a_{31} & \cdots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \cdots & a_{m2} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{m3} \\ . & . & . & & . \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{nm} \end{pmatrix}$$

es decir, se coloca la fila i de A como la columna i de A^T y la columna j de A como la fila j de A^T.

Si $A = A^T$ entonces diremos que A es simétrica. Este caso particular de matriz transpuesta es de gran utilidad al estudiar matrices semejantes y diagonalización.

Definición 1.8 (Multiplicación de Matrices) *Sean $A = (a_{ij})$ una matriz de $m \times n$ con componentes en el cuerpo F y $B = (b_{st})$ una matriz de orden $n \times p$ con componentes en el cuerpo F, se define la matriz producto $C = (c_{it})$ como aquella matriz de orden $m \times p$ con componentes en F dadas por*

$$\begin{aligned} c_{it} &= a_{i1}b_{1t} + a_{i2}b_{2t} + \cdots + a_{in}b_{nt} \\ &= \sum_{i=1}^{t} a_{ij}b_{jt}. \end{aligned} \tag{1.3}$$

Ahora comenzaremos el estudio de los determinantes de las matrices cuadradas. El determinante de una matriz se utiliza por ejemplo para encontrar las soluciones de un sistema de ecuaciones.

Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2\times 2}(F)$. El determinante de A está dado por

$$\det(A) = ad - bc$$

Esto es para una matriz de 2×2, pero para matrices cuadradas de $n \times n$, este método no es efectivo, para poder definir uno general debemos conocer otro concepto, es lo que nos lleva a la siguiente definición;

Definición 1.9 *Sea A una matriz de orden $n \times n$, sea δ el determinante que da como resultado luego de eliminar la i-ésima fila y la j-ésima columna de la matriz A. El escalar $(-1)^{i+j}\delta$ se llama el cofactor de A_{ij}.*

Esta definición la usaremos para poder entender el método que utilizaremos para encontrar el determinante de matrices de orden superior a 2.

Definición 1.10 *Sea A una matriz de n x n. Entonces el determinante de A, denotado por $\det A$ o $|A|$, está dado por.*

$$\det A = |A| = a_{11}|A_{11}| + a_{12}|A_{12}| + a_{13}|A_{13}| + \cdots + a_{1n}|A_{1n}| = \sum_{k=1}^{n} a_{1k}A_{11}$$

La expresión se llama expansión por cofactores.

Ejemplo 1.11 *Sea $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in M_{3\times 3}(F)$, entonces los cofactores de A_{12}, A_{22}, A_{32} son, respectivamente:*

$$\begin{aligned} (-1)^{1+2} \cdot det \begin{pmatrix} 4 & 6 \\ 7 & 9 \end{pmatrix} &= 6 \\ (-1)^{2+2} \cdot det \begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix} &= -12 \\ (-1)^{3+2} \cdot det \begin{pmatrix} 1 & 3 \\ 4 & 6 \end{pmatrix} &= 6 \end{aligned}$$

por lo tanto el determinante de la matriz A es;

$det\ A = a_{12} \cdot (6) + a_{22} \cdot (-12) + a_{32} \cdot (6) = 2 \cdot (6) + 5 \cdot (-12) + 8 \cdot (6) = 0$

1.3. Espacios vectoriales

Definición 1.12 (Espacio Vectorial) *Sea F un cuerpo. Un conjunto no vacío V se llama espacio vectorial, si en V hay dos operaciones llamadas suma y producto por escalar, denotadas por $+$ y $\cdot$ respectivamente, tales que $(V, +)$ es un grupo abeliano y $(\cdot)$ una ley de composición externa (LDCE) en V con escalares en F es tal que, si para todo $\alpha, \beta \in F$, y para todo $v, w \in V$:*

i. $\alpha \cdot (v + w) = \alpha \cdot v + \alpha \cdot w$

ii. $\alpha + \beta \cdot v = \alpha \cdot v + \beta \cdot v$

iii. $(\alpha) \cdot v = \alpha \cdot (\beta \cdot v)$

iv. $1v = v$

entonces decimos que V es un espacio vectorial. También se dice que V es un espacio vectorial sobre F y se denota $V(F)$.

Al comienzo de este capítulo vimos que todo subconjunto del cuerpo F que cumpla con las propiedades de suma y multiplicación de F, es un subcuerpo de éste y que es a su vez un cuerpo. Con los espacios vectoriales sucede algo similar, todo subconjunto del espacio vectorial V sobre el cuerpo F que cumpla las propiedades de suma y producto escalar sobre F, es un subespacio de V y es también un espacio vectorial, lo que nos lleva a la siguiente definición.

Definición 1.13 (Subespacio Vectorial) *Sea $V(F)$ un espacio vectorial, un subconjunto no vacío S de V se dice que es un subespacio vectorial de V, si y solo sí, S es un espacio vectorial sobre F para las leyes de composición de V.*

La siguiente proposición nos entrega una caracterización de los subespacios.

Proposición 1.1 *Sea $V(F)$ un espacio vectorial, un conjunto no vacio S de V se dice que es un subespacio vectorial de V, si y solo sí:*

a. *Si $x, y \in S$ entonces se tiene que $x + y \in S$*

b. $0 \in S$

c. *Si $x \in S$, entonces también $-x \in S$*

d. *$\alpha \in F, x \in S$, también $\alpha x \in S$*

Utilizando a. y d. se prueba que S es un subespacio sobre F, pues, de d. tenemos que $-1 \in F$, $x \in S$ entonces también $-1x \in S$ probando inmediatamente c., utilizando lo que se acaba de probar, si $x \in S$ y $-x \in S$ entonces la suma $x + (-x) \in S$ por lo tanto $0 \in S$, cumpliendo b. De este modo llegamos a un resultado muy útil y que reduce los calculos a lo siguiente:

S es un subespacio de V, si y solo sí, $\alpha x + \beta y \in S$, para todo $\alpha, \beta \in F$ y para todo $x, y \in S$.

1.3.1. Bases y Dimensión

Definición 1.14 *Sea V un espacio vectorial sobre F. Un subconjunto no vacío S de V se dice* ***linealmente dependiente*** *si existen vectores $v_1, v_2, \cdots, v_n$ de S y escalares $\alpha_1, \alpha_2, \cdots, \alpha_n$ de F, no todos nulos, tales que*

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0$$

Un conjunto que no es linealmente dependiente se dice **linealmente independiente**.

Definición 1.15 (Conjunto Generador) *Se dice que los vectores $v_1, v_2, \cdots, v_n$ en un espacio vectorial V generan a V si todo vector en V se puede escribir como una combinación lineal de ellos. Es decir, para todo $v \in V$, existen escalares $\alpha_1, \alpha_2, \cdots, \alpha_n$ tales que:*

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \tag{1.4}$$

Definición 1.16 *Sea $V(F)$ un espacio vectorial de dimensión finita, un subconjunto no vacío de vectores $\mathfrak{B} \subseteq V(F)$ se dice una base del espacio vectorial V si y solo sí:*

a. $\mathfrak{B}$ *es linealmente independiente.*

b. $\mathfrak{B}$ *genera a V*

Definición 1.17 *Si el espacio vectorial V tiene una base finita, entonces la dimensión de V es el número de vectores en todas las bases y V se llama espacio vectorial de dimensión finita. De otra manera, V se llama espacio vectorial de dimensión infinita. Si $V = \{0\}$, entonces se dice que V tiene dimensión cero.*

La dimensión de V se denota por $\dim V$.

Proposición 1.2 *Si V es un espacio vectorial de dimensión finita, entonces dos bases cualesquiera de V tienen el mismo número (finito) de elementos.*

Esta proposisión nos permite definir correctamente el concepto de dimensión.

Como verán a lo largo de este texto, el concepto de dimensión será muy utilizado, por lo que es importante tenerlo presente al momento de continuar el estudio de las **Formas de Jordan**.

Proposición 1.3 *Sea $V(F)$ un espacio vectorial de dimensión finita y S un subespacio de V. Entonces dim $S \leq$ dim V.*

Demostración. Trataremos de llegar a una contradicción para probar la proposición (demostración por el absurdo). Sea $n = \dim V$ y sea $n+1 = \dim S$. Entonces una base para V cuenta a lo más con n vectores linealmente independientes. Siguiendo la misma lógica, S tiene una base con $n+1$ vectores linealmente independientes. De aquí se deduce que posee un vector más en su base, que no pertenece a V pues no el combinación lineal de los vectores de ninguna base de V, lo que contradice el hecho de que $S \subseteq V$, por lo tanto si $S \subseteq V$, entonces $\dim S \leq \dim V$, lo que demuestra nuestra proposición. ■

Podemos utilizar esta demostración con unas pequeñas variaciones para probar el, siguiente teorema.

Proposición 1.4 *Sea V un espacio vectorial generado por un conjunto finito de vectores $v_1, v_2, \cdots, v_n$. Entonces todo conjunto linealmente independiente de vectores de V es finito y no contiene más de n elementos.*

Proposición 1.5 *Sea V un espacio vectorial de dimensión finita y sea $n = \dim V$. Entonces:*

a. Cualquier subconjunto de V que contenga más de n vectores es linealmente dependiente;

b. Ningún subconjunto de V que contenga menos de n vectores puede generar V

Proposición 1.6 *Sea A una matriz de $n \times n$ sobre el cuerpo F, y supóngase que los vectores fila de A forman un conjunto linealmente independiente de vectores de F^n. Entonces A es invertible.*

Ejemplo 1.18 Sea $F = \mathbb{R}$, entonces F^3 será el espacio vectorial $\mathbb{R}^3$, y generalizando tenemos que; Si F un cuerpo escalar, entonces F^n cumple las propiedades de 1.12, por lo que se dice que F^n es un espacio vectorial.

CAPÍTULO 2

Transformaciones Lineales

Ahora comenzaremos con el estudio de las transformaciones lineales que serán utilizadas hasta el final de este texto. Es importante tener en cuenta que una transformación lineal es una función aplicada a un espacio vectorial, por lo que se hace evidente la importancia de un curso de álgebra de funciones. Nosotros no nos detendremos aquí y continuaremos suponiendo que el lector entiende que es una función.

Una transformación lineal permite relacionar dos espacios vectoriales (definidos sobre el mismo cuerpo escalar) y de esta forma estudiar las propiedades de tales espacios.

Definición 2.1 *Sean V y W dos espacios vectoriales sobre F. $T : V \to W$ se llama una transformación lineal, si y solo sí:*

i. $T(x+y) = T(x) + T(y)$ *para todo* $x, y \in V$

ii. $T(\alpha x) = \alpha T(x)$ *para todo* $x \in V$ *y* $\alpha \in F$

Es importante observar que si T es una transformación lineal de V en W, entonces $T(0_V) = 0_W$; en efecto, $T(0_V) = T(0_V) + T(0_V)$. Por lo tanto $T(0_V) = 0_W$. También se tiene que T preserva las combinaciones lineales; esto es si $v_1, ..., v_n$ son vectores de V y $\alpha_1, ..., \alpha_n$ son escalares, entonces:

$$T(\alpha_1 v_1 + \cdot + \alpha_n v_n) = \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n). \tag{2.1}$$

Definición 2.2 *Sea V un espacio vectorial de dimensión finita. Una* ***base ordenada*** *de V es una sucesión finita de vectores linealmente independiente y que genera V.*

Proposición 2.1 *Sean V un espacio vectorial de dimensión finita sobre F, y $\{v_1, ..., v_n\}$ una base ordenada de V. Sea W un espacio vectorial sobre el mismo cuerpo F y $w_1, ..., w_n$ vectores cualesquiera de W. Entonces existe una única transformación lineal T de V en W tal que*

$$T(v_j) = w_j, \qquad j=1,\ldots,n$$

Demostración. Ver la página 95 de [6] ■

Definición 2.3 *Sea $T : V \to W$. Se llama Núcleo de T o espacio nulo de T al conjunto*

$$\text{Ker } T = \{v \in V \ : \ T(v) = 0\}$$

Tambien podemos definir el conjunto de las imágenes de T como

$$\text{Im } T = \{T(v) \ : \ v \in V\}$$

Es claro que Im $T \subseteq W$

Es claro que Ker $T \neq \emptyset$, pues como vimos anteriormente $T(0_V) = 0_W$, así que por lo menos $0_V \in$ Ker T

Proposición 2.2 *Si $T : V \to W$ lineal, entonces:*

i. Ker T es un subespacio de V.

ii. Im T es un subespacio de W.

Si V es de dimensión finita, el **rango** de T es la dimensión de la imagen de T y la nulidad de T es la dimensión del espacio nulo de T (Ker T).

Los siguientes teoremas nos darán claves importantes en el estudio de las transformaciones lineales.

Proposición 2.3 *Sean V y W espacios vectoriales sobre el cuerpo F y sea T una transformación lineal de V en W. Supóngase que V es de dimensión finita. Entonces*

$$\dim\,(\text{Im } T) + \dim\,(\text{Ker } T) = \dim V. \tag{2.2}$$

Demostración. Ver la página 99 y 100 de [6]

Proposición 2.4 *Sean V y W espacios vectoriales sobre el cuerpo F. Sean T y U transformaciones lineales de V en W. La función $T + U$ definida por*

$$(T + U)(v) = Tv + Uv \quad \textit{para todo } v \in V \tag{2.3}$$

es una transformación lineal de V en W. Si λ es cualquier elemento de F, la función (λT) definida por

$$(\lambda T)(v) = \lambda(Tv) \quad \textit{para todo } v \in V \tag{2.4}$$

también es una transformación lineal de V en W. El conjunto de todas las transformaciones lineales de V en W, junto con la adición y la multiplicación escalar aquí definidas, es un espacio vectorial sobre el cuerpo F dicho conjunto se denota por $\mathcal{L}(V, W)$, si $V = W$, lo denotamos $\mathcal{L}(V)$

Demostración. Ver la página 74 de [1] ■

Proposición 2.5 *Sean V un subespacio vectorial de dimensión finita n sobre el cuerpo F, y W un espacio vectorial de dimensión finita m sobre F. Entoces el espacio $\mathcal{L}(V, W)$ es de dimensión finita y tiene dimensión mn.*

Demostración. Ver la página 75 de [1] ■

Para la composición de funciones se utiliza la siguiente notación $U \circ T$, para las transformaciones (funciones) T y U, en adelante usaremos solo UT para referirnos a la composición de U con T.

Teorema 2.1 *Sean V, W y Z espacios vectoriales sobre el cuerpo F. Sea T una transformación lineal de V en W y U en Z una transformación lineal de W en Z. Entonces la función compuesta UT definida por $UT(v) = U(T(v))$ para todo $v \in V$ es una transformación lineal de V en Z.*

Demostración. Teniendo que

$$\begin{array}{c} T : V \to W \\ U : W \to Z \end{array}$$

debemos probar que si $U \circ T : V \to Z$, entonces $(U \circ T)$ es lineal
En efecto, para todo $v, w \in V$ y para todo $\lambda \in F$ se tiene que

$$\begin{aligned} (UT)(\lambda v + w) &= U[T(\lambda v + w)] \\ &= U(\lambda Tv + Tw) \\ &= U[\lambda Tv] + U(Tw) \\ &= \lambda (UT)(v) + (UT)(w) \end{aligned}$$

donde hemos ocupado el hecho de que T y U son lineales. ■

Definición 2.4 *Si V es un espacio vectorial sobre el cuerpo F, un operador lineal sobre V es una transformación lineal de V en V.*

Sea $T \in \mathcal{L}(V)$. Para denotar la composición de T consigo, usará la notación $T^2 = TT$, y en general $T^n = T \cdots T$ (n veces) para $n = 1, 2, 3 \cdots$. Se define $T^0 = I$ si $T \neq 0$, donde I es la transformación identidad: $Iv = v$ para todo $v \in V$, es decir, I lleva al espacio vectorial V en sí.

Lemma 2.1 *Sea V un espacio vectorial sobre el cuerpo F; sean U, T_1 y T_2 operadores lineales sobre V; Sea λ un elemento de F. Entonces:*

a. $IU = UI = U$;

b. $U(T_1 + T_2) = UT_1 + UT_2$;

c. $(T_1 + T_2)U = T_1U + T_2U$;

d. $\lambda(UT_1) = (\lambda U)T_1 = U(\lambda T_1)$

Una función T de V en W se dice invertible si existe una función U de W en V tal que UT es la función identidad de W. Si T es invertible, la función U es única y se representa por T^{-1}. Más aún, T es invertible si y, solo sí,

1. T es inyectiva, esto es, $Tv = Tw$ implica $v = w$.

2. T es sobreyectiva, esto es, la imagen de T es W.

Teorema 2.2 *Sean V y W dos espacios vectoriales sobre el cuerpo F y sea T una transformación lineal de V en W. Si T es invertible, entonces la función inversa T^{-1} es una transformación lineal de W sobre V.*

Demostración. Lo que se desea probar es que $T^{-1}(\lambda w_1 + w_2) = \lambda T^{-1}w_1 + T^{-1}w_2$. Sabiendo que w_1 y w_2 son vectores de W y λ escalar. Sabiendo que T es lineal se tiene:

$$\begin{aligned} T(\lambda v_1 + v_2) &= \lambda T v_1 + T v_2 \\ &= \lambda w_1 + w_2 \end{aligned}$$

Ahora

$$\begin{aligned} T^{-1}(\lambda w_1 + w_2) &= \lambda v_1 + v_2 \\ &= \lambda(T^{-1}w_1) + T^{-1}w_2 \\ &= \lambda T^{-1}w_1 + T^{-1}w_2 \end{aligned}$$

que era lo que se quería probar, que T^{-1} lineal. ■

Este resultado se usará bastante al momento de estudiar diagonalización de transformaciones lineales (o matrices). El siguiente resultado entrega una caracterización.

Teorema 2.3 *Sea T una transformación lineal de V en W. Entonces T es no singular si, es decir, que tiene inversa, y solo sí, T aplica cada subconjunto linealmente independiente del espacio vectorial V sobre un subconjunto linealmente independiente de W.*

Demostración. Sean $x_1, ..., x_n$ elementos de F,tales que

$$x_1 T(v_1) + ... + x_n T(_{v_n}) = 0.$$

Debido a la linealidad de T, tenemos que

$$T(x_1 v_1 + ... + x_n v_n) = 0.$$

En consecuencia, $x_1 v_1 + ... + x_n v_n = 0$. Puesto que $v_1, ..., v_n$ son linealmente independientes, se infiere que $x_i = 0$ para $i = 1, .., n$. Esto prueba el teorema. ■

Teorema 2.4 *Sean V y W espacios vectoriales de dimensión finita sobre el cuerpo F tal que dim V = dim W. Si T es una transformación lineal de V en W, las siguientes afirmaciones son equivalentes:*

i. T *es invertible.*

ii. T *es inyectiva.*

iii. T *es sobreyectiva; esto es, la imagen de* T *es* W.

Demostración. Ver la página 155 de [8] ■

Definición 2.5 *Sean* V *y* W *dos espacios vectoriales sobre el cuerpo* F. *Toda transformación lineal* T *de* V *en* W *sobreyectiva e inyectiva, se dice un isomorfismo de* V *sobre* W. *Si existe un isomorfismo de* V *sobre* W, *se dice que* V *es isomorfo a* W.

Obsérvese que V es trivialmente isomorfo a V, ya que el operador identidad es un isomorfismo de V sobre V. También si V es isomorfo a W por un isomorfismo T, entonces W es isomorfo a V, pues T^{-1} es un isomorfo de W sobre V. En resumen, el isomorfismo es una relación de equivalencia sobre la clase de espacios vectoriales. Si existe un isomorfismo de V sobre W, se dirá a veces que V y W son isomorfos, en vez de que V es isomorfo a W. Ello no será motivo de confusión porque V es isomorfo a W, si, y solo sí, W es isomorfo a V.

Ejemplo 2.6 Todo espacio de dimensión n sobre el cuerpo F es isomorfo al espacio F^n.

2.1. Representación de transformaciones por matrices

Sea V un espacio vectorial de dimensión n sobre el cuerpo F, y sea W un espacio vectorial de dimensión m sobre F. Sea $\mathfrak{B} = \{v_1, \cdots, v_n\}$ una base ordenada de V y $\mathfrak{B}' = \{w_1, \cdots, w_n\}$ una base ordenada de W'. Si T es cualquier transformación lineal de V en W, entonces T está determinada por su efecto sobre los vectores v_j. Cada uno de los n vectores Tv_j se expresa de manera única como combinación lineal.

$$Tv_j = \sum_{i=1}^{m} a_{ij} w_i \tag{2.5}$$

de los w_i, los escalares $a_{ij}, ..., a_{mj}$ son las coordenadas de Tv_j en la base ordenada $\mathfrak{B}'$. Por consiguiente, la transformación T está determinada por los mn escalares a_{ij} mediante la expresión. La matriz A de orden $n \times n$, definida por $a(i,j) = a_{ij}$ se llama matriz de T respecto al par de bases ordenadas $\mathfrak{B}$ y $\mathfrak{B}'$. La tarea inmediata es comprender claramente cómo la matriz A determina la transformación lineal T.

Si $v = x_1 v_1 + ... + x_n v_n$ es un vector de V, entonces usando la linealidad de T y (2.5) se tiene que

$$\begin{aligned} Tv &= T(\sum_{j=1}^{n} x_j v_j) \\ &= \sum_{j=1}^{n} x_j (T v_j) \\ &= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} w_i \\ &= \sum_{i=1}^{m} (\sum_{j=1}^{n} a_{ij} x_j) w_i. \end{aligned} \tag{2.6}$$

Si X es el vector de las coordenadas de v, se denota $[v]_{\mathfrak{B}}$, en la base ordenada $\mathfrak{B}$, entonces el cálculo anterior muestra que AX es la matriz de las coordenadas del vector Tv en la base ordenada $\mathfrak{B}'$, ya que el escalar

$$\sum_{j=1}^{n} a_{ij} x_j$$

es el elemento de la i-ésima fila de matriz columna $AX = a_{ij}$. Obsérvese también que si A es cualquier matriz $m \times n$ sobre el cuerpo F, entonces, para todo $v \in V$ tal que

$$v = \sum_{j=1}^{n} x_j v_j$$

se tiene que

$$T(\sum_{j=1}^{n} x_j v_j) = \sum_{i=1}^{m} (\sum_{j=1}^{n} a_{ij} x_j) w_i \tag{2.7}$$

define una transformación lineal T de V en W, la matriz de la cual es A, respecto a $\mathfrak{B}$, $\mathfrak{B}'$.

Estamos particularmente interesados en la representación por matrices de las transformaciones lineales de un espacio en sí, es decir, de los operadores lineales sobre un espacio vectorial V. En tal caso es más conveniente usar la misma base ordenada, esto es, hacer $\mathfrak{B} = \mathfrak{B}'$, y se dirá simplemente que la matriz que la representa es la matriz de T respecto a la base ordenada $\mathfrak{B}$. Como este concepto será muy importante para este estudio, repasaremos su definición.

Si T es un operador lineal sobre un espacio vectorial V de dimensión finita, y $\mathfrak{B} = \{v_1, ..., v_n\}$ es una base ordenada de V, la matriz de T respecto a $\mathfrak{B}$ (o la matriz de T en la base ordenada $\mathfrak{B}$) es la matriz A de orden $n \times n$, cuyos elementos A_{ij} están definidos por las ecuaciones

$$Tv_j = \sum_{i=1}^{n} a_{ij} v_i, \quad para\ j = j, ..., n \tag{2.8}$$

Teorema 2.5 *Sean V, W y Z espacios vectoriales de dimensión finita sobre el cuerpo F; sea T una transformación lineal de V en W y U una transformación lineal de W en Z. Si $\mathfrak{B}$, $\mathfrak{B}'$, $\mathfrak{B}''$ son las bases ordenadas de los espacios V, W y Z, respectivamente, y si A es la matriz de T respecto al par $\mathfrak{B}$, $\mathfrak{B}'$ y B es la matriz de U respecto al par $\mathfrak{B}'$, $\mathfrak{B}''$, entonces la matriz de la composición UT respecto al par $\mathfrak{B}$, $\mathfrak{B}''$ es la matriz producto $C = BA$.*

Demostración. Ver la página 127 de [6] ∎

Teorema 2.6 *Sea V un espacio vectorial de dimensión finita sobre el cuerpo F, y sean*

$$\mathfrak{B} = \{v_1, ..., v_n\} \quad y \quad \mathfrak{B}' = \{v'_1, ..., v'_n\}$$

dos bases ordenadas de V. Supóngase que T es un operador lineal en V Si $P = [P_1, ..., P_n]$ es la matriz $n \times n$ de columnas $P_j = [v'_j]_{\mathfrak{B}}$, entonces

$$[T]_{\mathfrak{B}'} = [U]_{\mathfrak{B}}^{-1}[T]_{\mathfrak{B}}[U]_{\mathfrak{B}}.$$

De otra manera, si U es el operador lineal sobre V definido por $Uv_j = v'_j$, $j = 1, ..., n$, entonces

$$[T]_{\mathfrak{B}'} = [U]_{\mathfrak{B}}^{-1}[T]_{\mathfrak{B}}[U]_{\mathfrak{B}}.$$

Demostración. Ver la página 127 de [6] ∎

Definición 2.7 *Sean A y B dos matrices de orden $n \times n$ sobre el cuerpo F. Se dice que B es semejante a A sobre F si existe una matriz invertible P de orden $n \times n$, sobre F tal que $B = P^{-1}AP$.*

Es importante que se tome en cuenta esta definición para poder iniciar el estudio de las formas canónicas elementales que serán la base de nuestro trabajo.

2.2. Funcionales lineales

Si V es un espacio vectorial sobre el cuerpo F, una transformación lineal f de V en el cuerpo de escalares F se llama también un funcional lineal sobre V. Si se comienza desde el principio, esto quiere decir que f es una función de V en F tal que

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

para todos los vectores v y w de V y todos los escalares λ de F. El concepto de funcional lineal es importante para el estudio de los espacios vectoriales de dimensión finita, pues ayuda a organizar y clarificar el estudio de los subespacios, las ecuaciones lineales y las coordenadas.

Proposición 2.6 *Sea V un espacio vectorial de dimensión finita sobre el cuerpo F y sea $\mathfrak{B} = \{v_1, ..., v_n\}$ una base de V. Entonces existe una única base dual $\mathfrak{B}^* = \{f_1, ..., f_n\}$ de V^* tal que $f_i(v_j) = v'_{ij}$. Para cada funcional lineal f sobre V se tiene*

$$f = \sum_{i=1}^{n} f(v_i) f_i \tag{2.9}$$

y para cada vector v de V se tiene

$$v = \sum_{i=1}^{n} f_i(v) v_i \tag{2.10}$$

Demostración. Se ha visto antes que existe una base única que es "dual"de $\mathfrak{B}$. Si f es un funcional lineal sobre V, entonces f es una combinación lineal de los f_i, los escalares α_j vienen dados por $\alpha_j = f(v_j)$. En forma análoga, si

$$v = \sum_{i=1}^{n} x_i v_i$$

es un vector de V, entonces

$$\begin{aligned} f_j(v) &= \sum_{i=1}^{n} x_i f_i(v_i) \\ &= \sum_{i=1}^{n} x_i z_{ij} \\ &= x_j. \end{aligned}$$

de modo que la expresión única para v, como combinación lineal de los v_i, es

$$v = \sum_{i=1}^{n} f_i(v) v_i.$$

■

Definición 2.8 *Si V es un espacio vectorial sobre F y S es un subconjunto de V, el anulador de S es el conjunto S^0 de funcionales lineales f sobre V tales que $f(v) = 0$ para todo v de S.*

De aquí en adelante este tipo de funciones serán muy utilizadas, sobre todo en el capítulo 3, pues al encontrar las raíces de estos polinomios encontraremos los valores propios del operador lineal. Esto se explicará con más detalle en en siguiente capítulo.

CAPÍTULO 3

Formas Canónicas

El Capítulo 1 y el Capítulo 2 de este texto nos han preparado el camino para comenzar el estudio de las Formas Canónicas, que es el primer paso que daremos antes de iniciar el estudio de las Formas de Jordan.

Al comenzar este capítulo nos dedicaremos a entender los conceptos de valor y vector propio, pero no lo haremos de forma tan profunda, pues nos interesa lo netamente práctico y que podremos utilizar más adelante, en la medida en que avancemos en este texto. Si el lector se interesa por saber más acerca de los valores propios, podemos adelantarle que su uso esta muy ligado a la geomentría analítica, por ejemplo en lo referente a las formas cuadráticas y secciones cónicas, para extraer información sobre las gráficas de ecuaciones cuadráticas, pero como ya hemos dicho anteriormente nos dedicaremos casi exclusivamente a los cálculos algebraicos, que nos facilitan el uso de los valores y vectores propios.

3.1. Valores Propios

Definición 3.1 *Sea V un espacio vectorial sobre el cuerpo F y sea $T : V \to V$ un operador lineal. Todo escalar $\lambda \in F$ tal que $Tv = \lambda v$ para algún $v \in V$, $v \neq 0$ se llama valor propio de T.*

Cada $v \in V$, $v \neq 0$ que satisface a $Tv = \lambda v$ se dice un vector propio asociado al valor propio λ.

A menudo, los valores propios también se denominan raíces caracteristicas, eigenvalores, valores caracteristicos o valores espectrales. En este escrito sólo se usará el nombre de valores propios.

Veremos ahora un ejemplo simple en el cual se podrá ver en acción cómo encontrar un valor y un vector propio para ir agudizando simplemente el ojo del lector.

$$\begin{aligned} \text{Sea} T \ &: \ \mathbb{R}^2 \to \mathbb{R}^2 \\ &\quad (x, y) \mapsto T(x, y) = (-3x + 6y, -9x + 12y) \end{aligned}$$

Usando la definición de T, se tiene que

$$T(x,y) = \lambda(x,y)$$

Esto equivale a tener las siguientes igualdades

$$\begin{aligned}(-3-\lambda)x + 6y = 0\\ -9x + (12-\lambda)y = 0\end{aligned}$$

y que puede ser escrita matricialmente como $A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, donde $A = \begin{pmatrix} -3-\lambda & 6 \\ 3 & 3-\lambda \end{pmatrix}$.
Por lo tanto deseamos buscar las soluciones no triviales del sistema $A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, las cuales se obtienen, a partir del hecho de que det $(A - \lambda I)$= 0. Entonces

$$\begin{aligned} 0 = \det\ A - \lambda I &= \lambda^2 - 9\lambda - 36 + 54 \\ &= \lambda^2 - 9\lambda + 18 \\ &= (\lambda - 3)(\lambda - 6) \end{aligned}$$

Aquí podemos observar que el sistema posee soluciones diferentes de la solución trivial, las cuales se obtienen al considerar $\lambda_1 = 3$ y $\lambda_2 = 6$ que son los valores propios del operador lineal T.

Veamos ahora cómo escontrar uno de estos valores propios. Si, $(x,y) \in \mathbb{R}^2$, es un vector propio asociado a $\lambda_1 = 3$, entonces $T(x,y) = 3(x,y)$. Esto quiere decir que:

$$(-3x + 6y, -9x + 12y) = (3x, 3y)$$

estos pares ordenados son iguales si se cumple que:

$$\begin{cases} \text{-6x + 6y} = 0 \\ \text{-9x + 9y} = 0 \end{cases} \quad \text{de aquí vemos que } y = x$$

Finalmente llegamos a que el vector propio asociado al valor propio $\lambda = 3$ son todos los vectores de la forma (x, x), es decir, los multiplos escalares del vector $(1, 1)$.

Por otro lado si $(x,y) \in \mathbb{R}^2$, es un vector propio asociado a $\lambda_2 = 6$, entonces $(-3x + 6y, -9x + 12y) = (6x, 6y)$, y así:

$$\begin{cases} \text{-6x + 6y} = 0 \\ \text{-9x + 9y} = 0 \end{cases} \quad \text{de aquí vemos que } y = \frac{3}{2}x$$

Luego, $(1, \frac{3}{2}) \in \mathbb{R}^2$ es un vector propio asociado a $\lambda_2 = 6$ y también lo es todo multiplo escalar no nulo de $(1, \frac{3}{2})$.

Ahora continuemos con valores y vectores propios como si quisieramos echar a andar el auto de nuestros sueños. Pero nos damos cuenta que no bastó con tan solo mirarlo y saber algo de él, ahora nos queda aprendernos el manual de conducir y !!no lo sabemos¡¡.

Empezaremos por eso a ver algunos temas importantes como las que a continuación veremos.

Teorema 3.1 *Sean V un espacio vectorial sobre el cuerpo F, $T : V \rightarrow V$ un operador lineal y $\lambda \in F$ un valor propio de T. Entonces los valores propios asociados a λ son los vectores no nulos de* ker*(T - λI)*

Demostración. $v \in V, v \neq 0$ es un vector propio de T asociado a λ, entonces $T(v) = \lambda v$. Esta igualdad equivale a tener

$(T - \lambda I)(v) = 0$ por lo tanto $v \in \ker(T - \lambda I)$. ■

Proposición 3.1 *Los valores propios de T son soluciones de la ecuación det$(A - \lambda I) = 0$. Donde λ es la incógnita y A una matriz asociada a T.*

Demostración. λ es un valor propio de T sí y solo si, existe $v \in V$, con $v \neq 0$ tal que $T(v) = \lambda(v)$. Si concideramos que existe un vector $v \in V$, con $v \neq 0$ tal que $(T - \lambda I)(v) = 0$, y por lo visto en la demostración del teorema anterior tenemos que:

$v \in$ Ker $(T - \lambda I)$, esto implica que $T - \lambda I$ no es inyectiva, por lo que $T - \lambda I$ tampoco es es invertible. Así, si A es la matriz asociada al operador lineal T, entonces $A - \lambda I$ no es invertible, por lo tanto det(A - $\lambda I) = 0$ ■

Los siguientes resultados se desprenden de lo hecho en el presente capítulo y que por lo demás son relevantes para continuar con nuestro estudio.

i. det(T - λI) es un polinomio en λ de grado $n = $ dimV$, pueselgradodelaecuaciónlodeterminaelproducto$ que estan en la diagonal principal. Este polinomio se denomina polinomio característico de T y se denota $P_T(\lambda)$.
   Por lo cual, λ es un valor propio de T, si y solo sí λ es raíz de su polinomio característico. Así, toda transformacion lineal definida sobre un espacio vectorial sobre el cuerpo F tiene n valores propios, donde n es la dimensión del espacio. La multiplicidad de la raíz λ_i del polinomio característico se llama *multiplicidad algebraica* de λ_i.

ii. La dimensión del espacio propio correspondiente al valor propio λ se llama multiplicidad geométrica de λ.

iii. Si λ es valor propio de T, entonces la multiplicidad geométrica de λ es menor o igual que la multiplicidad algebraica de λ

iv. λ es valor propio de una matriz A si λ es valor propio de T donde A es una matriz asociada a T respecto de cualquier base.

Definición 3.2 (Espacio propio) *Sea λ un valor propio de A. El subespacio E_λ, que es el espacio generado por el o los valores propios de A asociados al valor propio λ, se llama espacio propio de A correspondiente al valor propio λ.*

Veremos que el espacio generado por un conjunto de vectores v tales que $Tv = \lambda v$ es un subespacio de V. Para ello nos daremos un par de vectores v y w pertenecientes al espacio propio asociado al valor propio λ, y un par de escalares $\alpha, \beta \in F$. Entonces si $(\alpha v + \beta w) \in E_\lambda$ podremos decir que E_λ es un subespacio de V. En efecto:

$$T(\alpha v + \beta w) = T\alpha v + T\beta w = \lambda v + \lambda w = \lambda(v + w)$$

por lo tanto $\alpha v + \beta w \in E_\lambda$

Si T es cualquier operador y λ es cualquier escalar, el conjunto de los vectores v tales que $Tv = \lambda v$ es un subespacio de V, y este es el espacio nulo de la transformacion lineal (T - λI). Se llama a λ un valor propio de T si este subespacio es distinto del subespacio nulo, es decir, si (T - λI) no es inyectiva. Si V es de dimensión finita (T - λI), no es inyectiva justamente cuando su determinante es distinto de 0, consideraremos el determinante del operador como el determinante de la matriz que lo representa, no importando la base escogida. Vease **lema 3.1**.

Todas las observaciones anteriores las consignamos en el siguiete resultado.

Teorema 3.2 *Sea T un operador lineal sobre un espacio V de dimensión finita y sea λ un escalar. Las siguientes afirmaciones son equivalentes*

(i) *λ es un valor propio de T.*

(ii) *El operador (T - λI) es singular (no inversible)*

(iii) *det (T - λI) $= 0$*

El criterio del determinate (iii) es muy importante porque dice se pueden encontrar los valores propios de T. Como det (T - λI) es un polinomio de grado n en la variable λ, se determinaran los valores propios como las raices de tal polinomio.

Si $\mathfrak{B}$ es cualquier base ordenada de V y A = [T]$_{\mathfrak{B}}$ entonces (T - λI) es invertible si, y solo si, la matriz (A - λI) es invertible. En consecuencia se tiene la siguiente definición.

Definición 3.3 *Si A es una matriz $n \times n$ sobre el cuerpo F, un valor propio de A en F es un escalar λ de F tal que la matriz (A - λI) es singular (no invertible).*

Si det(λI - A) = 0, se puede construir la matriz (xI - A) con elementos polinómicos y considerar el polinomio $f(x) =$ det(xI - A)$. Entalcasolosvalorespropiosde$A$en$f(x)$sonlosescalares\lambda$ en F tales que $f(\lambda) = 0$. Por esta razón a $f(x)$ se le llama polinomio característico de A. Es importante observar que f es un polinomio mónico de grado n. Lo cual es fácilmente comprobable por la fórmula para el determinante de una matriz en términos de sus elementos.

Lemma 3.1 *Si A y B son matrices semejantes, entonces A y B tienen el mismo polinomio característico.*

Demostración. Se debe demostrar que A y B tienen el mismo polinomio característico. Como A y B son semejantes, entonces existe una matriz P invertible tal que $B = P^{-1}AP$ luego $B - xI = P^{-1}AP - xP^{-1}P = P^{-1}(A - xI)P$ y así

$$\begin{aligned} det\ (P^{-1}AP - xP^{-1}P) &= det[P(A - xI)P^{-1}] \\ &= det\ (P) \cdot det\ (A - xI) \cdot det\ (P^{-1}) \\ &= det\ (A - xI). \end{aligned}$$

■

Este lema permite definir el polinomio característico de un operador T como el polinomio característico de cualquier matriz $n \times n$ que representa a T en alguna base ordenada de V. Al igual que para las matrices, los valores propios de T serán las raíces características de T. En particular,esto muestra que T no puede tener más de n valores propios distintos.

Definición 3.4 *Sea $T : V \to V$ una transformación lineal diremos que T es diagonalizable si exise una base $\mathfrak{B}$ de V tal que $[T]_{\mathfrak{B}}$ es una matriz diagonal.*

Veamos que si $\mathfrak{B} = \{v_1, ..., v_n\}$ es una base de V formada por vectores propios de T, entonces $[T]_{\mathfrak{B}}$ es una matriz diagonal. Sabemos que $T(v_1) = \lambda_1 v_1$ y $T(v_2) = \lambda_2 v_2, \cdots \mathrm{T(v}_n) = \lambda_n v_n$, de lo que podemos concluir que:

$$[T]_{\mathfrak{B}} = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ . & . & ... & . \\ 0 & 0 & ... & \lambda_n \end{pmatrix}$$

Definición 3.5 *Sea T un operador lineal sobre un espacio vectorial de dimensión finita V. Se dice que T es diagonalizable si existe una base de V tal que cada vector suyo sea vector propio de T.*

Observación. No toda aplicación lineal es diagonalizable, veamos en un ejemplo :

$$\begin{aligned} T \quad : \quad & \mathbb{R}^2 \to \mathbb{R}^2 \\ & (x, y) \mapsto T(x, y) = (-y, x) \end{aligned} \tag{3.1}$$

La siguiente matriz es la representación del operador T en la base $\mathfrak{B} = \{(0, 1), (-1, 0)\}$ $[T]_{\mathfrak{B}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ de aquí tenemos que: det(A - λI) = det$\begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 + 1 = 0$ la que no tiene solución en $\mathbb{R}$, por lo tanto T no tiene valores propios y como T no tiene valores propios por lo tanto no es diagonalizable.

Proposición 3.2 *Sea V un espacio vectorial de dimensión n. Si una transformación lineal $T : V \to V$ tiene n vectores propios linealmente independientes, entonces T es diagonalizable.*

Demostración. Si T tiene n vectores propios linealmente independientes entonces el conjunto formado por ellos es una base para V porque $\dim(V) = n$. Luego por la definición anterior sabemos que la matriz asociada a T respecto de tal base es diagonal y por lo tanto T es diagonalizable. ■

Teorema 3.3 *Sea T un operador lineal sobre un espacio vectorial V de dimensión finita. Sean $\lambda_1, ..., \lambda_k$ los valores propios distintos de T y sean W_i el espacio nulo de $(T - \lambda_i I)$. Las siguientes afirmaciones son equivalentes:*

(i) T es diagonalizable.

(ii) *El polinomio característico de T es*

$$P_T = (x-\lambda_1)^{d_i}...(x-\lambda_k)^{d_k}$$

donde $\dim W_i = d_i$, *para* $i = 1, ..., k$.

(iii) $\dim W_i + ... + \dim W_k = \dim V$.

Demostración. Ver la página 186 de [1] ■

3.2. Polinomios Anuladores

Si queremos entender correctamente esta sección debemos definir previamente algunos conceptos.

Definición 3.6 (Ideal) *Sea F cuerpo. Un ideal en $F[x]$ es un subespacio M de $F[x]$ tal que fg pertenece a M siempre que f esté en $F[x]$ y g en M.*

Nota: $F[x]$ consta de todas las combinaciones lineales (finitas) de x y sus potencias.

Si requerimos analizar un operador lineal T, una de las cosas más útiles de saber es la clase de los polinomios que anulan a T. Para precisar, supóngase que T es un operador lineal del espacio vectorial V sobre el cuerpo F. Si p es un polinomio sobre F, entonces $p(T)$ es también un operador lineal sobre V. Si q es otro polinomio sobre F, entonces

$$(p+q)(T) = p(T) + q(T) = (pq)(T) = p(T)q(T).$$

En consecuencia, la colección de polinomios p que anulan a T, en el sentido de que

$$p(T) = 0$$

es un ideal en el álgebra (la definición de un álgebra la encuentra en la página 116 de [1])de los polinomios $F[x]$. Puede ser el ideal cero; es decir, puede ser que T no sea anulado por cualquier polinomio no nulo. Pero ello no puede suceder si el espacio V es de dimensión finita. Supóngase que T es un operador lineal sobre el espacio V de dimensión n. Considérense las primeras (n^2+1) potencias de T:

$$I, T, T^2, ..., T^{n^2}.$$

Esta es una sucesión de n^2+1 operadores en $\mathcal{L}(V)$, el espacio de los operadores lineales sobre V. El espacio $\mathcal{L}(V)$ tiene dimensión n^2. Por tanto, la sucesón de los n^2+1 operadores debe ser linealmente dependiente, es decir, se debe tener que:

$$\lambda_0 I + \lambda_1 T + ... + \lambda_{n^2} T^{n^2} = 0$$

para los escalares λ_i, no todos nulos. En consecuencia, el ideal de polinomios que anulan a T tiene un polinomio no nulo de grado n^2 o menor.

Todo ideal de polinomios consta de todos los múltiplos de un cierto polinomio mónico fijo, el generador del ideal. Así, pues, al operador T corresponde un múltiplo **mónico** p (ver página 119 de [1]) con la siguiente propiedad: Si f es un polinomio sobre F, entonces $f(T) = 0$ si, y solo sí, $f = pg$, donde g es algún polinomio sobre F.

Definición 3.7 *Sea T un operador lineal sobre un espacio vectorial V de dimensión finita sobre el cuerpo F. El* ***polinomio minimal*** *de T es el polinomio mónico generador (único) del ideal de polinomios sobre F que anulan a T.*

El nombre de "polinomio minimal" se debe a que el generador de un ideal de polinomios está caracterizado por ser el polinomio mónico de grado minimal en el ideal. Lo cual quiere decir que el polinomio minimal p para el operador lineal T está unívocamente determinado por estas tres propiedades:

(1) p es un polinomio mónico sobre el cuerpo escalar.

(2) $p(T) = 0$.

(3) Ningún polinomio sobre F que anule a T tiene grado menor que el que tiene p.

Si A es una matriz de orden $n \times n$ sobre F, se define el **polinomio minimal** de A en forma análoga, como el único generador mónico del ideal de todos los polinomios sobre F que anulan A. Si el operador T está representado en cierta base ordenada por la matriz A, entonces T y A tienen el mismo polinomio minimal. Esto es porque $f(T)$ está representado en la base por la matriz $f(A)$, de modo que $f(T) = 0$ si, y solo si, $f(A) = 0$.

De la última observación respecto a operadores y matrices se sigue que las matrices semejantes tienen el mismo polinomio minimal. Ese hecho resulta también evidente de la definición, ya que

$$F(P^{-1}AP) = P^{-1}F(A)P$$

para todo polinomio f.

Supóngase que A es una matriz de orden $n \times n$ con elementos en el cuerpo F. También supónga que F_1 es un cuerpo que contine a F como subcuerpo. (Por ejemplo, A puede ser una matriz con elementos racionales, mientras que F_1 es el cuerpo de los números reales. A puede ser una matriz con elementos reales, mientras que F_1 es el cuerpo de los números complejos.) Se puede considerar A como matriz de orden $n \times n$ sobre F o como matriz de orden $n \times n$ sobre F_1. En primera instancia podría creerse que se obtienen dos polinomios minimales distintos para A. Dando gracias, ese no es el caso; veamos por qué. Consideramos todos los polinomios mónicos, con coeficientes en F, que anulan a A y tomamos el de menor grado. Si f es un polinomio mónico sobre F:

$$f = x^k + \sum_{j=0}^{k-1} A_j x^i \tag{3.2}$$

entonces $f(A) = 0$ dice simplemente que tenemos una relación lineal entre potencias de A, es decir:

$$A^k + a_{k-1}A^{k-1} + ... + a_1A + a_0I = 0$$

El grado del polinomio minimal es el menor entero positivo k tal que existe una relación lineal de la forma entre las potencias de $I, A, ..., A^k$. Más aún, por la unicidad del polinomio minimal, existe para aquel k una, y solo una, relación de la forma; es decir, una vez que el minimal k está determinado, existen escalares únicos $a_0, ..., a_{k-1}$ en F para lo que es válido. Son los coeficientes del polinomio minimal.

Ahora (para cada k) se tiene en un sistema de n^2 ecuaciones lineales para las "incógnitas" $a_0, ..., a_{k-1}$. Como los elementos de A están en F, los coeficientes del sistema de ecuaciones están en F. Por tanto, si el sistema tiene una solución con los $a_0, ..., a_{k-1}$ en F_1. Está claro ahora que los dos polinomios minimales son los mismos.

Lo que sabemos hasta ahora en razón al polinomio minimal de un operador lineal en un espacio de dimensión n es que solo su grado no excede de n^2. Pero ésta es una estimación poco aproximada, ya que el grado no puede exceder a n.

De hecho este resultado lo ocuparemos para afirmar lo siguiente; Todo polinomio que anule a T (o A) es divisible por el polinomio minimal, lo que reduce conciderablemente la búsqueda de polinomios que anulen a la transformación T, y aunque la mayor parte de esta sección esta dedicada al estudio de los polinomios minimal y característico, es de gran utilidad saberlo y conciderarlo.

Proposición 3.3 *Sea T un operador lineal sobre un espacio vectorial V de dimensión n (o sea A una matriz $n \times n$). El polinomio característico y el polinomio minimal de T (de A) tienen las mismas raíces, salvo multiplicidades.*

Demostración. Sea $p(T)$ el polinomio minimal de T. Sea λ un escalar. Lo que se desea demostrar es que $p(\lambda) = 0$ si, y solo si, λ es un valor propio de T. Primero supóngase que $p(\lambda) = 0$. Entonces

$$p(T) = (x - \lambda)q(T)$$

donde $q(T)$ es un polinomio. Como $grd\ q < grd\ p$, la definición del polinomio minimal p dice que $q(T) \neq 0$. Se elige un vector w tal que $q(T)w \neq O$. Sea $v = q(T)w$. Entonces

$$0 \;=\; p(T)w = (T - \lambda I)q(T)w = (T - \lambda I)v$$

y así λ es un valor propio de T.
Supóngase ahora que λ es un valor propio de T, o sea $Tv = \lambda v$, con $v \neq 0$. Como se observó en el lema anterior,

$$p(T)\alpha = p(\lambda)\alpha.$$

Ya que $p(T) = 0$ y $\alpha \neq 0$, se tiene que $p(\lambda) = 0$. ■

Sea T un operador lineal diagonalizable y sean $(\lambda_1, ..., \lambda_k)$ los valores propios distintos de T. Entonces es fácil ver que el polinomio minimal de T es el polinomio

$$p(T) = (x - \lambda_1)\cdots(x - \lambda_k).$$

Si v es un vector propio, entonces uno de los operadores $(T - \lambda_1 I), \cdots, (T - \lambda_k I)$ aplica v en 0. Por lo tanto,

$$(T - \lambda_1 I)\cdots(T - \lambda_k I)\alpha = 0$$

para todo vector propio v. Existe una base para el espacio soporte que consta de los vectores propios de T; luego

$$p(T) = (T - \lambda_1 I)\cdots(T - \lambda_k I) = 0$$

Por lo tanto, si T es un operador lineal diagonalizable, entonces el polinomio minimal de T es un producto de factores lineales distintos. Como se verá más adelante, esta propiedad caracteriza los operadores diagonalizables.

Teorema 3.4 (Cayley-Hamilton) *Sea T un operador lineal sobre un espacio vectorial V de dimensión finita. Si f es el polinomio característico de T, entonces $f(T) = 0$; es decir, el polinomio minimal divide al polinomio característico de T.*

Demostración. Ver la página 620 de [2] ■

3.3. Subespacios Invariantes

Definición 3.8 *Sea V un espacio vectorial y T un operador lineal sobre V. Si W es un subespacio de V, se dice que W es invariante por T si para todo vector v de W, el vector Tv está en W, es decir, si $T(W)$ está contenido en W.*

Ilustremos la definición con el siguiente ejemplo: Los vectores propios no nulos del operador lineal $T : V \to V$ pueden caracterizarse como generadores de subespacios unidimensionales T-invariantes. Supongamos que $T(v) = \lambda v$, $v \neq 0$. Entonces, $W = \{\kappa v \ : \ \kappa \in F\}$ el subespacio unidimensional generado por v, es invariante bajo T porque:

$$T(\kappa v) = \kappa T(v) = \kappa(\lambda v) = \kappa\gamma v \text{ donde } \kappa\gamma v \in W$$

Otro ejemplo de subespacios invariantes muy conocido por todos es el $\ker(T)$. Supongamos que $v \in \ker T$ (con v no necesariamente igual al vector nulo), es decir, $T(v) = 0_v \in \ker T$.

El siguiente teorema ilustra el efecto de una función sobre un espacio vectorial cualquiera bajo un operador lineal T.

Teorema 3.5 *Sean $T : V \to V$ un operador lineal y $f(\gamma)$ cualquier polinomio, siendo γ perteneciente al cuerpo F. El núcleo de $f(T)$ es invariante bajo T.*

Demostración. Supongamos $v \in \ker f(T)$, es decir, $f(T)(v) = 0_v$. Debemos probar que $T(v)$ también pertenece al núcleo de $f(T)$, o sea, $f(T)(T(v)) = 0$. Dado que $f(\gamma)\gamma = \gamma f(\gamma)$, tenemos $f(T)T = Tf(T)$. Así,

$$f(T)T(v) = Tf(T)(v) = T(0) = 0_v,$$

cumpliendo con lo que se quería demostrar. ■

Teorema 3.6 *Supongamos que W es un subespacio invariante de $T : V \to V$. Entonces T tiene una representación matricial por bloques, en la base ordenada $\mathfrak{B}$ de V, donde $T_{\mathfrak{B}} = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$, donde A es una representación matricial de la restricción T_w de T a W.*

Demostración. Para mostrar este teorema tomamos una base $\mathfrak{B}' = \{w_1, ..., w_r\}$ de W y la extenderemos a una base $\mathfrak{B} = \{w_1, ..., w_r, v_1, ..., v_s\}$ de V. Entonces ahora tenemos lo siguiente:

$$\begin{aligned}
T_{w_1} &= T(w_1) = a_{11}w_1 + \cdots + a_{1r}w_r \\
T_{w_2} &= T(w_2) = a_{21}w_1 + \cdots + a_{2r}w_r \\
&\cdots \\
T_{w_r} &= T(w_r) = a_{r1}w_1 + \cdots + a_{rr}w_r \\
T_{w_r} &= T(v_1) = b_{11}w_1 + \cdots + b_{1r}w_r + c_{11}v_1 + \cdots + c_{1s}v_s \\
T_{w_r} &= T(v_2) = b_{21}w_1 + \cdots + b_{2r}w_r + c_{21}v_1 + \cdots + c_{2s}v_s \\
&\cdots \\
T_{w_r} &= T(v_s) = b_{s1}w_1 + \cdots + b_{sr}w_r + c_{s1}v_1 + \cdots + c_{ss}v_s
\end{aligned}$$

Por otro lado sabemos que la matriz de T en la base $\mathfrak{B}'$ es la transpuesta de la matriz de coeficientes del sistema de ecuaciones anterior. Por ende es de la forma $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$, donde A es la transpuesta de la matriz de coeficientes del subsistema. Por el mismo argumento, A es la matriz de T_w (T_w es la restricción de T a W) relativa a la base w_i de W. ■

La siguiente definición la usaremos más adelante, pero ya es necesario ir familiarizandonos con ella.

Definición 3.9 *Sea W un subespacio invariante para T, y sea v un vector de V. El T-conductor de v en W es el conjunto $S_T(v; W)$, que consta de todos los polinomios g (sobre el cuerpo escalar) tales que $g(T)v$ está en W.*

Como el operador T se considera fijo a lo largo del trabajo, se suprimirá normalmente el subíndice T y se escribirá $S(v; W)$. Conductor es el término más corriente, preferido por aquellos que piensan en un operador $g(T)$ menos agresivo, que suavemente dirige al vector v en W. En el caso especial de $W = \{0\}$, el conductor se llama T-anulador de v.

3.4. Suma Directa e Invariante

La presente sección no está dentro de los objetivos explícitos de nuestro trabajo. Sin embargo es relevante al momento de estudiar la Forma de Jordan, por lo que es importante detenernos (si es necesario) para entender la sección y no quedar con vacíos que más adelante puedan dificultar la comprensión del objetivo fundamental de nuestro trabajo.

Nuestra misión ha sido desde el principio encontrar una base ordenada en la cual la matriz que representa a T tenga una forma especialmente simple. Entonces nuestro siguiente propósito será, primero trabajar más en términos de operadores (no tanto con matrices), segundo descomponer al espacio vectorial V en suma de subespacios invariantes para T, de modo que los operadores restricción (T_w) a esos subespacios sean simples.

Definición 3.10 . *Sean $W_1, ..., W_k$ subespacios de un espacio vectorial V. Se dice que $W_1, ..., W_k$ son **independientes** si*

$$v_1 + ... + v_k = 0, \qquad v_i \text{ en } W_i$$

implica que cada v_i es 0.

Lemma 3.2 *Sea V un espacio vectorial de dimensión finita. Sean $W_1, \cdots, W_k$ subespacios V y sea $W = W_1 + \cdots + W_k$. Las siguientes afirmaciones son equivalentes.*

a) $W_1, \cdots, W_k$

b) Para todo j, $2 \leq j \leq k$, se tiene que

$$W_j \bigcap (W_1 + \cdots + W_{j-1}) = \{0\}$$

c) Si $\mathfrak{B}_i$ es una base ordenada de $W_i, 1 \leq i \leq k$, entonces el conjunto $\mathfrak{B} = \{\mathfrak{B}_1, \cdots, \mathfrak{B}_k\}$ es una base ordenada para W.

Si cualquiera (y, por tanto, todas) de las condiciones del último lema se cumplen, se dice que la suma $W = W_1 + \cdots + W_k$ es **directa** o que W es la **suma directa** de $W_1, \cdots, W_k$, y se escribe

$$W = W_1 \oplus \cdots \oplus W_k.$$

Definición 3.11 . *Si V es un espacio vectorial, una **proyección** de V es un operador lineal E sobre V tal que $E^2 = E$.*

Un ejemplo simple de proyección sobre un espacio vectorial es el operador identidad I. En efecto, para todo $v \in V$

$$I^2 v = Iv = v \text{ por lo tanto } I^2 = I$$

Las proyecciones pueden ser empleadas para describir las descomposiciones en suma directa del espacio V.

Con todo lo descrito anteriormente, obtenemos los siguientes resultados:

Teorema 3.7 *Si $V = W_1 \oplus \cdots \oplus W_k$, entonces existen k operadores lineales E_i, para $i = 1, \cdots, k$ sobre V tales que*

(i) todo E_i es una proyección $E_i^2 = E_i$;

(ii) $E_i E_j = 0 \qquad$ si $i \neq j$;

(iii) $I = E_1 + \cdots + E_k$;

(iv) la imagen de E_i es W_i.

Demostración. Ver la página 211 de [1] ∎

Nuestra siguiente misión será descomponer al espacio vectorial V en suma directa de subespacios W_i, donde cada W_i invariante bajo un operador lineal T dado. Si T_i es operador inducido sobre W_i por T, entonces, si v es un vector de V, se tiene vectores únicos $v_1, ..., v_k$, con v_i en W_i, de modo que

$$v = v_1 + \cdots + v_k$$

y entonces

$$T_v = T_1 v_1 + \cdots + T_k v_k$$

Esta situación se describirá diciendo que T es la suma directa de los operadores $T_1, \cdots, T_k$.

Si usamos esta terminología debemos recordar que los T_i no son operadores sobre el espacio vectorial V, pero si lo son para los distintos subespacios W_i.

Ahora expresaremos la invarianza del subespacio vectorial W_i en términos del operador proyección E_i.

Teorema 3.8 *Sea T un operador lineal sobre el espacio V y sean $W_1, \cdots, W_k$ y $E_1, \cdots, E_k$ como en el* Teorema 3.7. *Entonces una condición necesaria y suficiente para que cada subespacio W_i sea invariante por T es que T conmute con cada una de las proyecciones E_1; es decir,*

$$TE_i = E_i T, \qquad i = 1, \cdots, k.$$

Demostración.

$$\begin{aligned} Tv &= T(E_j v) \\ &= E_j T(v) \end{aligned}$$

que muestra que Tv está en la imagen de E_j, es decir, que W_j es invariante por T. Supóngase ahora que todo W_i es invariante por T. Se ha de demostrar que $TE_i = E_j T$. Sea v cualquier vector de V. Entonces

$$\begin{aligned} v &= E_i v + \cdots + E_k v \\ Tv &= TE_i v + \cdots + TE_k v \end{aligned}$$

Como $E_i v$ está en W_i, que es invariante por T, se debe tener que $T(E_i v) = E_i w_i$ para cierto vector w, Entonces

$$E_j T E_i v \;=\; E_j E_i w_i = \begin{cases} 0, \text{ si } i \neq j \\ E_j w_j, \text{ si } i = j. \end{cases}$$

Con lo que

$$E_j T v = E_j T E_1 v + \cdots + E_j T E_k v = E_j w_j = T E_j v.$$

Esto vale para cada v de V, o sea $E_j T = T E_j$. ■

Nos enfocaremos ahora en un operador diagonalizable T en el lenguaje de las descomposiciones en suma directa invariante (proyecciones que conmutan con T). Esto nos importa ahora para comprender luego unos teoremas de descomposición más profundos.

Teorema 3.9 *Sea T un operador lineal sobre un espacio de dimención finita V. Si T es diagonalizable y si $\lambda_1, \ldots, \lambda_k$ son los valores propios distintos de T, entonces existen operadores lineales $E_1, \ldots, E_1$ en V tales que*

(i) $T = \lambda_1 E_1 + \cdots + \lambda_k E_k$;

(ii) $I = E_1 + \cdots + E_k$;

(iii) $E_i E_j = 0, \quad$ *para todo i, j tales que $i \neq j$;*

(vi) $E_i^2 = E_i$ *(E_i es una proyección);*

(v) *la imagen de E_i es el espacio propio de T, asociado a λ_1*

Demostración. Ver la página 214 de [1] ■

3.5. Teorema de Descomposición Prima

Si queremos estudiar el Teorema de descomposición prima, así de buenas a primeras, nos estrellaremos de frente con una gran muralla, primero debemos definir algunos conceptos que se utilizaran en esta sección y solo así podremos comenzar con nuestro estudio, pero esto no nos garantiza gran facilidad para entenderlo, después de mucho trabajo y concentración se podrá comprender en su totalidad el teorema de descomposición prima.

Como se anunció anteriormente veamos ahora:

Definición 3.12 *Si $p_1, \cdots, p_n$ son polinomios sobre el cuerpo F, no todos nulos, el generador d del ideal*

$$p_1 F[x] + \ldots + p_n F[x]$$

se llama el máximo común divisor de $p_1, \cdots, p_n$. Se dice que los polinomios $p_1, \cdots, p_n$ son primos relativos si su máximo común divisor es 1, o en forma equivalente, si el ideal que ellos generan es todo $F[x]$.

Ejemplo 3.13 *Sea $\mathbb{C}$ el cuerpo de los números complejos. Entonces*
(a)m.c.d.$(x+2, x^2+8x+16) = 1$
(b)m.c.d.$((x-2)^2(x+i), (x+2)(x^2+1)) = (x-2)(x+i)$.
En efecto, el ideal

$$(x-2)^2(x+i)F[x] + (x-2)(x^2+1)F[x]$$

contiene a

$$(x-2)^2(x+i) - (x-2)(x^2+1)) = (x-2)(x+i)(i-2)$$

Luego contiene a $(x-2)(x+i)$, *que es mónico y divide a*

$$(x-2)^2(x+i)y(x-2)(x^2+1)$$

Definición 3.14 *Sea F un cuerpo. Un polinomio f de $F[x]$ se dice reducible sobre F si existen polinomios g, h en $F[x]$ de grado mayor o igual a 1 tales que $f = gh$, y si no, se dice que es irreducibre sobre F. Un polinomio no escalar irreducible sobre F se llama polinomio primo sobre F y a veces se dice solamente que es primo en $F[x]$.*

Ejemplo 3.15 *El polinomio x^2+1 es reducible sobre el cuerpo C de los números complejos.*

$$\text{En efecto,} x^2+1 = (x+i)(x-i)$$

y los polinomios $x+i, x-i$ pertenecen a C[x]. Por otra parte, x^2+1 es irreducible sobre el cuerpo R de los números reales. Pues si

$$x^2+1 = (ax+b)(a'x+b')$$

con a, a', b, b' en R, entonces

$$aa' = 1, ab' + ba' = 0, bb' = 1.$$

Estas relaciones implican que $a^2+b^2 = 0$, lo que es imposible con números reales a y b, a menos que a = b = 0.

Proposición 3.4 *Sea f un polinomio mónico no escalar sobre el cuerpo F y sea $f = p_1^{n_1} \cdots p_k^{n_k}$ la factorización prima de f. Para todo j, $1 \leq j \leq k$, sea*

$$f_j = \frac{f}{p_j^{n_j}} = \prod_{i \neq j} p_i^{n_i}.$$

Entonces $f_1, ..., f_k$ son primos relativos.

Demostración. Ver la página 135 de [1] ■

En la sección anterior vimos que si T es un operador lineal sobre el espacio vectorial V, este se podrá descomponer en suma directa de operadores que en cierto sentido son elementales. En algunos casos es posible hacerlo por medio de valores y vectores propios del operador T, que es cuando el polinomio minimal de T se puede factorizar como polinomios mónicos de grado 1. Los problemas que encontramos cuando estudiamos el T general son, primero, puede que T no tenga valores propios simples, pues no siempre se podrá factorizar completamente sobre F, y aunque esto ocurra nos podemos encontrar con un segundo problema, que los vectores propios de T no alcanzan a formar una base generadora del espacio vectorial V.

El siguiente teorema facilita la búsqueda de valores y vectores propios en el caso del T general, lo que es claramente un alivio en nuestro estudio.

Teorema 3.10 (Teorema de la Descomposición Prima) *Sea T un operador lineal sobre el espacio vectorial V de dimensión finita sobre el cuerpo F. Sea p el polinomio minimal de T,*

$$p = p_1^{r_1}...p_k^{r_k}$$

donde los p_i son polinomios mónicos irreducibles distintos sobre F, y los r_i son enteros positivos. Sea W_i el espacio nulo de $p_i(\mathrm{T})^{r_i}$, i = 1,...,k. Entonces

(i) $V = W_1 \oplus \cdots \oplus W_k$;

(ii) *cada E_i es invariante por T;*

(iii) *si T_i es el operador inducido sobre W_i por T, entonces el polinomio minimal de T_i es $p_i^{r_i}$.*

Demostración. Sólo daremos la idea de la demostración. Si la descomposición en suma directa *(i)* es válida. ¿cómo se pueden determinar las proyecciones $E_1, ..., E_k$ asociados a la descomposición? La proyección E_j será la identidad sobre W_i y cero sobre los otros W_j. Se encontrará un polinomio h_i tal que $h_i(T)$ es la identidad sobre W_i y es cero sobre los otros W_j con lo que $h_1(T) + ... + h_k(T) = I$, etc. ■

Corolario 3.11 *Si E_1,...,E_k son las proyecciones asociadas a la descomposición prima de T, y en consecuencia, si un operador lineal U conmuta con T, entonces U conmuta con cada uno de los E_i; es decir, cada subespacio W_i es invariante por U.*

Con la notación de la demostración del Teorema 3.10, consideremos el caso especial en que el polinomio minimal de T es un producto de polinomios de primer grado; es decir, el caso en que cada p_i es de la forma $p_i = x - \lambda_i$. Ahora la imágen de E_i es el espacio nulo W_i de $(T - \lambda_i I)^{r_i}$. Hágase $D = \lambda_1 E_1 + \cdots + \lambda_k E_k$. Por el Teorema 3.9, D es un operador diagonalizable que se llama **parte diagonalizable** de T. Considérese el operador $N = T - D$. Ahora

$$\begin{aligned} T &= TE_1 + \cdots + TE_k \\ D &= \lambda_1 E_1 + \cdots + \lambda_k E_k. \end{aligned}$$

así

$$N = (T - \lambda_1 I)E_1 + \cdots + (T - \lambda_k I)E_k$$

Entonces ahora utilizaremos las proyecciones para ver que

$$N^2 = (T - \lambda_1 I)^2 E_1 + ... + (T - \lambda_k I)^2 E_k$$

y en general que

$$N^r = (T - \lambda_1 I)^r E_1 + ... + (T - \lambda_k I)^r E_k$$

Definición 3.16 *Sea N un operador lineal sobre el espacio vectorial V. Se dice que N es* ***nilpotente*** *si existe un entero positivo r tal que $N^r = 0$.*

Teorema 3.12 *Sea T un operador lineal sobre el espacio vectorial V de dimensión finita sobre el cuerpo F. supóngase que el polinomio minimal de T se descompone sobre F, en producto de polinomios lineales. Entonces existe un operador diagonalizable D sobre V y un operador N nilpotente sobre V tales que*

(i) $T = D + N$

(ii) $DN = ND$

El operador diagonalizable D el operador nilpotente N están unívocamente determinados por (i) y (ii), y cada uno de ellos es un polinomio de T

Demostración. Ver la página 250 de [4] ■

Corolario 3.13 *Sea V un espacio vectorial de dimensión finita sobre un cuerpo F algebraicamente cerrado, cuerpo de los números complejos. Entonces todo operador lineal T sobre V puede escribirse como suma de un operador diagonalizable D y un operador nilpotente N que conmutan. Estos operadores D y N son únicos y cada uno es polinomio de T.*

Por estos resultados se ve que el estudio de los operadores lineales en un espacio vectorial sobre un cuerpo algebraicamente cerrado queda esencialmente reducido al estudio de operadores nilpotentes. Para espacios vectoriales sobre cuerpos no algebraivamente cerrados se debe encontrar todavía un sustituto de los valores y vectores propios.

CAPÍTULO 4

La Forma de Jordan

Comenzaremos dando una mirada simple a la Forma de Jordan, para así ir conociéndola poco a poco y no encontrarnos tan rápidamente con una teoría que es muy densa y difícil de comprender.

Como la mayor parte de los polinomios tienen raíces distintas, la mayor parte de las matrices tendrán valores propios distintos. Sin embargo, las matrices que no son diagonalizables (es decir, que no tinen n vectores propios linealmente independientes) surgen en la práctica. En este caso, todavía es posible demostrar que la matriz es semejante a otra, una matriz más sencilla, pero la nueva matriz no es diagonal y es más difícil obtener la matriz de transformación C.

Definamos la matriz N_k como la matriz $k \times k$, o sea,

$$N_k = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{4.1}$$

en N_k se puede observar que es una matriz con unos arriba de la diagonal principal y ceros en todos las otras partes.

Ahora comenzaremos a utilizar el hecho que un espacio puede descomponerse como suma de subespacios invariantes (visto en el capítulo anterior), es por esto que era tan relevante su estudio, aún sabiendo que no parecia ser parte de nuestro objetivo, y sin embargo forma parte de la colmna vertebral del presente capítulo.

Luego de descomponer el espacio vectorial en suma de subespacios invariantes, debemos encontrar la representación matricial del operador restricción, estas serán los bloques de la matriz de Jordan y que se encuentran en la diagonal principal de ésta.

4.1. Matriz de Bloques de Jordan

Para un escalar dado λ se define la matriz de bloques de Jordan $A(\lambda)$ por lo siguiente:

$$A(\lambda) = \lambda I + N_k = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \cdots & & & & & \\ 0 & 0 & \cdots & \lambda & 0 & \\ 0 & 0 & \cdots & o & \lambda & \end{pmatrix} (4.2)$$

(4.2)

Es decir, $A(\lambda)$ es la matriz de $k \times k$ con el escalar λ en la diagonal, unos arriba de la diagonal y ceros en la otra partes, tambíen lo podemos encontrar con unos de bajo de la diagonal y ceros en la otra parte.

4.2. Matriz de Jordan

Por último, esta matriz tine la forma de Jordan

$$J = \begin{pmatrix} A_1(\lambda_1) & 0 & \cdots & 0 \\ 0 & A_2(\lambda_2) & \cdots & 0 \\ . & . & \cdots & . \\ 0 & 0 & \cdots & A_r(\lambda_r) \end{pmatrix}$$

donde cada $A_j(\lambda_i)$ es una matriz de bloques de Jordan. Entonces una matriz de Jordan es una matriz que tiene en la diagonal matrices de bloques de Jordan y ceros en otra parte.

Ejemplo 4.1 Los siguientes ejemplos son matrices de Jordan, los bloques se marcaran con líneas punteadas:

$$i \begin{pmatrix} 5 & 1| & 0 \\ \underline{0} & \underline{5}| & \underline{0} \\ 0 & 0| & 3 \end{pmatrix} \qquad ii \begin{pmatrix} \underline{5} & \underline{0} & \underline{0} & \underline{0} & 0 \\ 0 & |4 & 1 & 0| & 0 \\ 0 & |0 & 4 & 1| & 0 \\ 0 & |\underline{0} & \underline{0} & \underline{4}| & \underline{0} \\ 0 & 0 & 0 & 0 & |9 \end{pmatrix}$$

Teorema 4.1 *Sea A una matriz real o compleja de $n \times n$. Entonces existe una matriz C compleja invertible de $n \times n$ tal que*

$$C^{-1}AC = J \tag{4.3}$$

donde J es una matriz de Jordan cuyos elementos en la diagonal son los valores propios de A. Más aún, la matriz de Jordan J es única excepto por el orden en el que aparecen los bloques de Jordan.

Demostración. Ver la página 568 de [2] ■

Definición 4.2 *La matriz J en el Teorema 4.1 se llama la forma canónica de Jordan A.*

Teorema 4.2 *Suponga que A la matriz 2×2 tiene un valor propio λ de multiplicidad 2 y multiplicidad geométrica 1. Sea v_1 un vector propio correspondiente a λ. Entonces existe un vector v_2 que satisface la ecuación:*

$$(A - \lambda I)v_2 = v_1. \tag{4.4}$$

Demostración. Sea $x \in \mathbb{C}^2$ un vector fijo que no es múltiplo de v_1 de manera que x no es un vector propio de A. Primero se demuestra que:

$$w = (A - \lambda I)x \tag{4.5}$$

es un vector propio de A. Esto es, debe demostrarse que $w = cv_1$ para alguna constante c. Como $w \in \mathbb{C}^2$ y v_1 y x son linealmente independientes, existen constantes c_1 y c_2 tales que

$$w = c_1 v_1 + c_2 x \tag{4.6}$$

Para demostrar que w es un vector propio de A debe demostrarse que $c_2 = 0$. De (4.5) y (4.6) se encuentra que

$$(A - \lambda I)x = c_1 v_1 + c_2 x \tag{4.7}$$

Sea $B = A - (\lambda + c_2)I$. Entonces de (4.7)

$$Bx = [A - (\lambda + c_2)I]x = c_1 v_1 \tag{4.8}$$

Si se supone que $c_2 \neq 0$, entonces $\lambda + c_2 \neq \lambda$ y $\lambda + c_2$ no es un valor propio de A (ya que λ es el único valor propio de A). Así, $\det B = \det[A - (\lambda + c_2)I] \neq 0$, lo que significa que B

es invertible. Por lo tanto, (4.8) se puede escribir como

$$x = B^{-1}c_1 v_1 = c_1 B^{-1} v_1 \tag{4.9}$$

Entonces multiplicando ambos lados de (4.9) por λ, se tiene

$$\lambda x = \lambda c_1 B^{-1} v_1 = c_1 B^{-1} \lambda v_1 = c_1 B^{-1} A v_1 \tag{4.10}$$

Pero $B = A - (\lambda + c_2)I$, de manera que

$$A = B + (\lambda + c_2)I \tag{4.11}$$

Al insertar (4.11) en (4.10) se tiene

$$\begin{aligned} \lambda x &= c_1 B^{-1}[B + (\lambda + c_2)I]v_1 \\ &= c_1[I + (\lambda + c_2)B^{-1}]v_1 \\ &= c_1 v_1 + (\lambda + c_2)c_1 B^{-1} v_1 \end{aligned} \tag{4.12}$$

Pero usando (4.9), $c_1 B^{-1} v_1 = x$ de manera que (4.12) se convierte en

$$\lambda x = c_1 v_1 + (\lambda + c_2)x = c_1 v_1 + c_2 x + \lambda x$$

y de aquí se deduce que

$$0 = c_1 v_1 + c_2 x \tag{4.13}$$

Pero v_1 y x son linealmente independientes, lo que hace que $c_1 = c_2 = 0$. Esto contradice la suposición de que $c_2 \neq 0$. Entonces $c_2 = 0$ y por (4.6), w es un múltiplo de v_1 por lo que $w = c_1 v_1$ es un vector propio de A. Más aún, $w \neq 0$ ya que si $w = 0$, entonces (4.5) dice que x es un vector propio de A. Por lo tanto $c_1 \neq 0$. Sea $v_2 = \frac{1}{c_1}x$, entonces $(A - \lambda I)v_2 = \left(\dfrac{1}{c_1} \right)(A - \lambda I)x = \left(\dfrac{1}{c_1} w \right) = v_1$. ■

Definición 4.3 (Vector propio generalizado) *Sea A una matriz de orden 2×2 con un sólo valor propio λ que tiene multiplicidad geométrica 1. Sea v_1 un vector propio de A. Entonces el vector v_2 definido por $(A - \lambda I)v_2 = v_1$ se llama vector propio generalizado de A correspondiente al valor propio de λ.*

Teorema 4.3 *Suponga que A, λ, v_1 y v_2 estan definidos como en el Teorema 4.1 y sea C la matriz cuyas columnas son v_1 y v_2. Entonces $C^{-1}AC = J$, donde $J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ es la forma canónica de Jordan de A.*

Demostración. Como v_1 y v_2 son linealmente independientes, se ve que C es invertible. Después se observa que $AC = A(v_1, v_2) = (Av_1, Av_2) = (\lambda v_1, Av_2)$, pero en la ecuación (4) $Av_2 = v_1 + \lambda v_2$ de manaera que $AC = (\lambda v_1 v_1 + \lambda v_2)$, pero

$$CI = (v_1, v_2)\begin{pmatrix} \lambda & 1 & \\ 0 & \lambda & \end{pmatrix} = \lambda v_1, v_1 + \lambda v_2$$

Entonces $AC = CJ$, lo que significa que $C^{-1}AC = J$ y el teorema queda probado. ■

Ya familiarizados con la Forma de Jordan de una manera simple, como indicamos al principio, ahora la veremos de una manera más elegante y con términos más elevados. Lo que debemos hacer es comenzar a darle forma a la teoría, trabajando más a la par con los operadores y las matrices.

Como vimos en el capítulo anterior, existen matrices en las que se hace muy complicado encontrar los valores propios del operador (o la matriz), pues no siempre es sencillo factorizar un polinomio. También sabemos que si la matriz A es diagonalizable, encontrar los valores propios se vuelve un trabajo aún más sencillo. Pero, ¿qué pasa cuando una matriz no se puede diagonalizar?. ¿Será posible encontrar los valores propios de la transformación usando lo estudiado en los capítulos anteriores?. Dar respuesta a estas interrogantes será nuestro siguiente objetivo.

Pero antes de empezar con el estudio más profundo daremos algunas definiciones útiles en el estudio de las Formas de Jordan, llegando a enunciar el teorema de descomposición cíclica, que usaremos más adelante.

Definición 4.4 *Si v es cualquier vector de V, el* ***subespacio** T**-cíclico generado*** *por v es el subespacio $Z(v;T)$ de los vectores de la forma $g(T)v$, g en $F[x]$. Si $Z(v;T) = V$ entonces se dice que v es un vector cíclico de T.*

Definición 4.5 *Si v es cualquier vector de V, el T-**anulador** de v es el ideal $M(v;T)$ en $F[x]$ que consta de todos los polinomios g sobre F de modo que $g(T)v = 0$. Al polinomio mónico único p_v que genera este ideal se le llamará también el T-**anulador** de v.*

Definición 4.6 *Sea T un operador lineal sobre un espacio vectorial V y sea W un subespacio de V. Se dice que W es T-admisible si*

(i) W es invariante por T;

(ii) si $F(T)w$ está en W, existe un vector z en W tal que $F(T)w = F(T)z$.

Teorema 4.4 (Teorema de Descomposición Cíclica) *Sea T un operador lineal sobre un espacio vectorial de dimensión finita V y sea W_0 un subespacio propio T-admisible de V. Existen vectores no nulos $v_1, ..., v_r$ en V con T-anuladores respectivamente $p_1, ..., p_r$, tales que:*

(i) $V = W_0 \oplus Z(v_1;T) \oplus ... \oplus Z(v_r;T)$*:*

(ii) p_k *divide a* $p_{k-1}, k = 2, ..., r$.

Más aún, el entero r *y los anuladores* $p_1, ..., p_r$ *están univocamente determinados por (i), (ii) y el hecho de que ninguno de los* v_k *es* 0.

Demostración. Ver la página 232 de [1] ■

Sin más preámbulos continuaremos con el estudio de las Formas de Jordan. Lo antes descrito nos ayudará a lograr entender el resto de este capítulo.

Supóngase que N es un operador lineal nilpotente en el espacio V de dimensión finita. Se considera la descomposición cíclica para N que se obtiene del Teorema 4.4. Se tiene un entero positivo r y r vectores no negativos $v_1, v_2, ..., v_r$ en V con N-anuladores $p_1, ..., p_r$, tales que

$$V = Z(v_1;N) \oplus ... \oplus Z(v_r;N)$$

y p_{i+1} divide a p_i para $i = 1, ..., r-1$. Como N es nilpotente, el polinomio minimal es x^k para algún $k \leq n$. Así, todo p_i es de la forma $p_i = x^k$ y la condición de divisibilidad simplemente dice que

$$k_1 \geq k_2 \geq ... \geq k_r.$$

Por supuesto,$k_1 = k$ y $k_r \geq 1$. La matriz asociada de x^{k_i} es la matriz $k_i \times k_i$.

$$A_i = \begin{pmatrix} 0 & 0 & ... & 0 & 0 \\ 1 & 0 & ... & 0 & 0 \\ 0 & 1 & ... & 0 & 0 \\ . & . & ... & . & . \\ 0 & 0 & ... & 1 & 0 \end{pmatrix} \tag{4.14}$$

Así, el Teorema 4.4 da una base ordenada para V, en la cual la matriz de N es suma directa de matrices elementales nilpotentes (4.14), cuyas dimensiones decrecen cuando i crece. Por lo que se ve que, asociados con una matriz nilpotente $n \times n$, hay un entero positivo r y r enteros positivos $k_1, ...k_r$, de modo que $k_1 + ... + k_r = n$ y $k_i \geq k_{i+1}$, y estos enteros positivos determinan la forma racional de la matriz, es decir, determinan la matriz, salvo semejanza.

Con respecto al operador nilpotente se debe saber que el entero positivo r es justamente la nulidad de N; en efecto, el espacio nulo tiene como base los r vectores

$$N^{k_i-1}v_i.$$

Si v está en el espacio nulo de N, se puede escribir v en la forma

$$v = f_1v_1 + \cdots + f_rv_r$$

donde f_i es un polinomio, cuyo grado se puede suponer menor k_i. Como $Nv = 0$, para todo i tenemos

$$0 \quad = \quad N(f_i v_i) = N f_i(N) v_i = (x f_i) v_i$$

Así, xf_i es divisible por x^k, y como $grad(f_i) > k_i$, ello quiere decir que

$$f_i = \lambda_i x^{k_i - 1}$$

donde λ_i es cierto escalar. Pero entonces

$$v = \lambda_1(x^{k_1-1}v_1) + ... + \lambda_r(x^{k_r-1}v_r)$$

que muestra que los valores forman una base para el espacio nulo de N. El lector puede observar que este hecho es también claro desde el punto de vista matricial.

Lo que se desea hacer ahora es combinar lo establecido para operadores o matrices nilpotentes con el teorema de descomposición prima del Capítulo 3. La situación es la siguiente: supóngase que T es un operador lineal sobre V y que el polinomio característico de T se puede factorizar sobre F como sigue

$$f = (x - \lambda_i)^{d_i} ... (x - \lambda_k)^{d_k}$$

donde los $\lambda_1, ..., \lambda_k$ son elementos distintos de F y $d_i \geq 1$. Entonces el polinomio minimal de T será

$$p = (x - \lambda_i)^{r_i} ... (x - \lambda_k)^{r_k}$$

donde $1 \leq r_i \leq d_i$. Si W_i es ek espacio nulo de $(T - \lambda_i I)^{r^i}$, entonces el teorema de descomposición prima dice que

$$V = W_1 \oplus ... \oplus W_k$$

y que el operador T_i inducido en W_i por T tiene polinomio minimal $(x - \lambda_i)^{r_i}$. Sea $N_i = T_i - \lambda_i I$. Entonces N_i es nilpotente y tiene un polinomio minimal x^{r_i}. En W_i, T actúa como N_i más el escalar λ_i veces el operador identidad. Supóngase que elegimos una base del subespacio W_i que corresponda a la descomposición cíclica del operador nilpotente N_i. Entonces la matriz de T_i en esa base ordenada será la suma directa de matrices N_i. Entonces la matriz de T_i en esa base ordenada será la suma directa de matrices

$$\begin{pmatrix} \lambda & 0 & ... & 0 & 0 \\ 1 & \lambda & ... & 0 & 0 \\ 0 & 1 & ... & 0 & 0 \\ . & . & ... & . & . \\ & & & \lambda & \\ 0 & 0 & ... & 1 & \lambda \end{pmatrix} \tag{4.15}$$

cada cual con $\lambda = \lambda_i$. Además, las dimensiones de estas matrices decrecerán al ir de izquierda a derecha. Una matriz de la forma (4.15), se llama una matriz elemental de Jordan con valor propio λ. Ahora, si se consideran en conjunto todas las bases para W_i, se obtiene una base ordenada de V. Se describirá la matriz A de T en esta base ordenada.

La matriz A es la suma directa

$$A = \begin{pmatrix} A_1 & 0 & ... & 0 \\ 0 & A_2 & ... & 0 \\ . & . & ... & . \\ 0 & 0 & ... & A_k \end{pmatrix} \tag{4.16}$$

de matrices $A_1, ..., A_k$. Toda A_i es de la forma

$$A_i = \begin{pmatrix} J_1^i & 0 & ... & 0 \\ 0 & J_2^i & ... & 0 \\ . & . & ... & . \\ 0 & 0 & ... & J_{n_i}^i \end{pmatrix}$$

donde cada $J_j^{(i)}$ es la matriz elemental de Jordan con valor propio λ_i. También en cada A_i la dimensión de las matrices $J_j^{(i)}$ decrece cuando j crece. Se dirá que una matriz $n \times n$, A, que satisface todas las condiciones descritas hasta ahora (para escalares distintos $\lambda_1, ..., \lambda_k$) está en forma de Jordan.

Estamos viendo que si T es un operador lineal cuyo polinomio característico se puede factorizar totalmente sobre el cuerpo escalar, diremos que hay una base ordenada de V, en la cual T está forma de Jordan. Buscamos mostrar en esta instancia que esta matriz es algo unívocamente asociado con T, salvo el orden en que se escriben los valores propios de T. Queremos expresar que, si dos matrices están en Forma de Jordan y son semejantes, entonces difieren sólo en la ordenación de los escalares λ_i.

A continuación viendo la unicidad diremos. Supóngamos que existe una base ordenada de V en la cual T está representado por la matriz de Jordan A, vista anteriormente. Si A_i es una matriz $d_i \times d_i$, entonces d_i es netamente la multiplicidad de λ_i como raíz de polinomio característico de A o de T. En otras palabras, el polinomio característico de T es

$$f = (x - \lambda_i)^{d_i} ... (x - \lambda_k)^{d_k}$$

Concretamente dice que $\lambda_1, ..., \lambda_k$ y $d_1, ..., d_k$ son únicos, menos el orden en que se escriben. El que A sea la suma directa de las matrices A_i, nos da una descomposición en suma directa $V = W_1 \oplus ... \oplus W_k$ invariante por T. Veamos en este momento que W_i debe ser el espacio nulo de $(T - \lambda_i I)^n$, donde $n = \dim V$; en efecto, $A_i - \lambda_i I$ es evidentemente nilpotente $A_j - \lambda_i I$ es no singular para $j \neq i$. Así vemos que los subespacios W_i son únicos. Si T_i es el operador inducido en W_i por T, entonces la matriz A_i está unívocamente determinada como la forma racional de $(T_i - \lambda_i I)$.

Enfoquémonos ahora sobre el operador T y la matriz de Jordan A que representa a T en una base ordenada. Ellas son:

(1) Cada elemento de A, que no esté en la diagonal principal o inmediatamente debajo de ésta es 0. En la diagonal de A están los κ valores propios distintos $\lambda_1, ..., \lambda_k$ de T. También λ_i está repetido d_i veces, donde d_i es la multiplicidad de λ_i como raíz del polinomio característico, es decir, $d_i = \dim W_i$.

(2) Para todo i la matriz A_i es la suma directa de n_i matrices elementales de Jordan $J_j^{(i)}$ con valor propio λ_i. El número n_i es precisamente la dimensión del espacio de vectores propios asociados al valor propio λ_i. En efecto,n_i es el número de bloques elementales nilpotentes en la forma racional de $(T_i - \lambda_i)I$, y es así igual a la dimensión del espacio nulo de $(T_i - \lambda_i)I$. En particular, obsérvese que T es diagonalizable si, y sólo si, $n_i = d_i$ para todo i.

(3) Para todo i el primer bloque $J_j^{(i)}$ en la matriz A_i es una matriz $r_i \times r_i$, donde r_i es la multiplicidad de λ_i como raíz del polinomio minimal de T. Esto se desprende de que el polinomio minimal del operador nilpotente $(T_i - \lambda_i I)$ es x^{r_i}.

Es importante recalcar que, como se acostumbra ver, se observan resultados semejantes para las matrices. Si B es una matriz $n \times n$ sobre un cuerpo F y si el polinomio característico de B se factoriza completamente sobre F, por ende B es semejante sobre F a una matriz $n \times n$ n,A en la forma de Jordan, y A es única salvo el orden de los valores propios. Se dice que A es la forma de Jordan B.

Por otro lado veamos que si F es un cuerpo algebraicamente cerrado, encontraremos que los comentarios anteriores son válidas para todo operador lineal sobre un espacio vectorial de dimensión finita sobre F, o para toda matriz $n \times n$ sobre F. Así, por ejemplo, toda matriz $n \times n$ sobre el cuerpo de los números complejos es semejante a una matriz, esencialmente única, en Forma de Jordan.

Lo que podemos concluir de esta última sección es que si tenemos un espacio vectorial V y T un operador lineal tal que $T : V \to V$, entonces si descomponemos V en sumas directas invariantes, también podremos escribir la matriz asociada a la transformación como bloques en la diagonal principal y ceros en los otros lados. Estos bloques son subespacios invariantes de V tal que la suma de estos subespacios es precisamente el espacio vectorial V, como ya habíamos dicho que V se podía descomponer en suma directa invariante, entonces el conjunto de los valores propios de estos subespacios (de la matriz de bloques), serán los valores propios del operador T.

Una aplicación de la Forma de Jordan, o más particularmente de la diagonalización, es lo referente a formas cuadráticas y secciones cónicas. Una de las interrogantes que surge al estudiar estas es; ¿qué sucede con las formas cuadráticas cuando no se ve a simple vista la cónica que está representando?, ¿como lograr graficarla si debemos hacer un cambio de coordenadas? Estudiando los valores y vectores propios del operador lineal lograremos dar respuesta a estas inquietudes.

Antes de continuar debemos recordar que las ecuaciones de un circulo, elipse e hipérbola son:

$$\begin{aligned} \text{Circulo} &: \quad x^2+y^2=r^2 \\ \text{Elipse} &: \quad \frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \\ \text{Hipérbola} &: \quad \begin{cases} \frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \\ \text{ó} \\ \frac{y^2}{b^2}-\frac{x^2}{a^2}=1 \end{cases} \end{aligned}$$

A continuación veremos dos ejemplos, el primero en $\mathbb{R}^2$ y luego en $\mathbb{R}^3$.

Ejemplo 4.7 *Consideremos la ecuación $x^2-4xy+3y^2=6$, lo primero que debemos realizar es el cambio de coordenadas, así podremos reconocer con mayor facilidad la cónica a la que esta forma cuadrática representa.*

Fácilmente podemos comprovar que la ecuación cuadrática se puede escribir de la forma $Ax\cdot x=6$, donde $A=\begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix}$. Esta matriz es facil de diagonalizar, de hecho luego de encontrar los valores y vectores propios, vemos que A se puede diagonalizar a $D=\begin{pmatrix} 2-\sqrt{5} & 0 \\ 0 & 2+\sqrt{5} \end{pmatrix}$ usando la matriz ortogonal

$$Q=\frac{1}{\sqrt{10-2\sqrt{5}}}\begin{pmatrix} 2 & 1-\sqrt{5} \\ -1+\sqrt{5} & 2 \end{pmatrix}$$

Entonces

$$\begin{aligned} v'=\begin{pmatrix} x' \\ y' \end{pmatrix}=Q^Tx &= \frac{1}{\sqrt{10-2\sqrt{5}}}\begin{pmatrix} 2 & -1+\sqrt{5} \\ 1-\sqrt{5} & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \\ &= \frac{1}{\sqrt{10-2\sqrt{5}}}\begin{pmatrix} 2x+(-1+\sqrt{5})y \\ (1-\sqrt{5})x+2y \end{pmatrix} \end{aligned}$$

y, para las nuevas variables, la ecuación se puesde escribir como:

$$(2-\sqrt{5})x'^2+(2+\sqrt{5})y'=6$$

O sea

$$\frac{y'^2}{\frac{6}{2+\sqrt{5}}}-\frac{x'^2}{\frac{6}{\sqrt{5}-2}}=1$$

Esta es la ecuación de la hipérbola con $a=\sqrt{\frac{6}{2+\sqrt{5}}}\approx 1{,}19$ y $b=\sqrt{\frac{6}{\sqrt{5}-2}}\approx 5{,}04$. Como

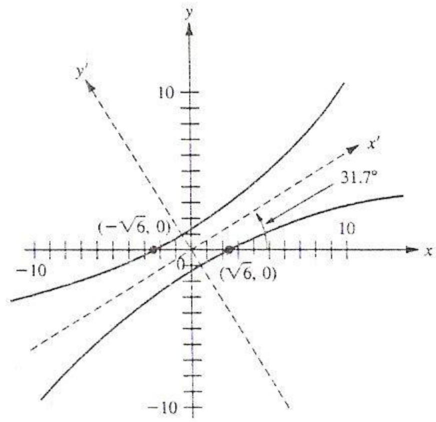

Figura 4.1: Gráfica de la cónica, ilustrando los diferentes ejes coordenados.

$$Q = \frac{1}{\sqrt{10-2\sqrt{5}}}\begin{pmatrix} 2 & 1-\sqrt{5} \\ -1+\sqrt{5} & 2 \end{pmatrix}$$

y $\det Q = 1$*, se tiene, usando el hecho de que 2 y* $-1+\sqrt{5}$ *son positivos (consideremos que si* $\det Q = 1$ *y* $a \geqslant 0$ *y* $c > 0$*, entonces definimos* $\theta = \cos^{-1} a$*, con* $0 < \theta \leqslant \frac{\pi}{2}$*), de este modo;*

$$\cos\theta = \frac{2}{\sqrt{10-2\sqrt{5}}} \approx 0{,}85065$$

entonces Q *está en el primer cuadrante, y usando una calculadora, se encuentra que* $\theta \approx 0{,}5536$*, rad* $\approx 31{,}7^o$*. Por lo tanto esta es la ecuación de una hipérbola rotada en un ángulo de* $31{,}7^o$ *(ver* fig. 1*).*

Con esto hemos terminado el estudio de la Forma de Jordan que era nuestro objetivo desde el comienzo de este texto.

CAPÍTULO 5

Conclusiones

Lo trascendental de este estudio es darnos cuenta de lo maravillosamente sencillo que puede resultar el estudio de un espacio vectorial luego de descomponerlo en subespacios más pequeños y simples que el original.

Para darle una base sólida a este escrito, en el primer capítulo cuando retomamos estudio de algunos conceptos básicos del álgebra lineal, nos dimos cuenta al ir desarrollando el tema, y a pesar de su aparente simpleza , su especificidad presenta algunas variables de difícil comprensión y resolución, por eso que aún nos quedaban vacíos que llenar y se fue haciendo cada vez más gratificante la tarea de estudiar, entender y aplicar estos conceptos, porque estábamos complementando las bases de nuestra formación matemática y profesional.

Al seguir con el análisis nos introducimos en el fascinante mundo de las transformaciones lineales, una materia excitante y que a pesar de la simpleza con la que se presentan los temas y la facilidad con la que se pueden realizar los cálculos una vez que se comprenden bien, es un tema que ha muchos complica, es por esto que nos sentimos realizados al lograr entender más acabadamente las transformaciones lineales, tema de nuestro estudio.

En la medida en que íbamos desarrollando aún más el tema nos estremecíamos cada vez más con cada paso que dábamos. Llegando a descubrir que a pesar que un problema se pueda presentar, en apariencia, de forma muy complicada, siempre podremos encontrar la forma de resolverlo de la manera más sencilla posible. Para ello estudiamos todos los temas que formaban parte de nuestro estudio y más, logrando la comprensión de estos y como utilizarlos a nuestro provecho. Aprendimos que no siempre se debe estudiar todo de forma tan directa, sino que muchas veces puede ser mucho más productivo descomponer en partes más pequeñas el objeto de estudio y así lograr mejores resultados.

BIBLIOGRAFÍA

[1] K. Hoffman et Al. *Álgebra Lineal*. Prentice-Hall, 1972.

[2] S. Grossman. *Álgebra Lineal*. McGRAW-HILL de Mexico S.A., 1996

[3] S. Lipschutz. *Álgebra Lineal*. McGRAW-HILL de Mexico S.A., 1992

[4] G. Mostow et Al. *Álgebra Lineal*. McGRAW-HILL de Mexico S.A., 1972

[5] R. Johnson. *Álgebra Lineal*. Compañia Editorial Continental S.A., 1969

[6] S. Lang *Álgebra Lineal*. LITOGRAFICA INGRAMEX, S.A., 1979

[7] J. Fraleigh et Al. *Álgebra Lineal*. ADDISON-WESLEY IBEROAMERICANA, S.A., 1989

[8] S. Lang *Introducción al Álgebra Lineal*. ADDISON-WESLEY IBEROAMERICANA, S.A., 1990

[9] D. Lay *Álgebra Lineal Y Sus Aplicaciones*. PEARSON Educación de México S.A., 2007