Instituto de Perfeccionamiento y Estudios Superiores
CFE/ANEP - UdelaR
Diploma en Matemática
Mención Aplicaciones

Forma de Jordan de matrices y aplicaciones a las ecuaciones diferenciales

Luis Menéndez

Orientador: Andrés Abella

6 de noviembre de 2019

Contenidos

Notación y nomenclatura

$\mathbb{K}$	cuerpo de los números reales o complejos
$d(a,b)$	distancia de a a b
E	espacio normado
$\mathcal{M}_n(\mathbb{K})$ o $\mathcal{M}_n$	espacio de las matrices $n\times n$ sobre el cuerpo $\mathbb{K}$
e^{tA}	función exponencial matricial
$mcd\,(p,q)$	máximo común divisor de p y q
$MA(\lambda)$	mutiplicidad algebraica del valor propio λ
$MG(\lambda)$	mutiplicidad geométrica del valor propio λ
$\Vert A\Vert$	norma de una matriz A
$\ker(A)$	núcleo de la matriz A
$\mathrm{Im}\,(A)$	imagen de la matriz A
$\mathbb{K}[t]$	polinomios sobre el cuerpo $\mathbb{K}$
$\mathcal{X}_A(t)$	polinomio caraterístico de la matriz A
$\sim$	semejanza de matrices
$[U]$	subespacio generado por el conjunto U
$\sum A_n$	serie de término A_n
$\oplus$	suma directa de subespacios o de matrices
$\sum_{n=0}^{\infty} A_n$	suma de la serie de término general A_n
$\mathcal{M}_n(\mathbb{K}[t])$	anillo de las matrices cuyas entradas son polinomios de $\mathbb{K}[t]$
$\mathcal{M}_n(\mathbb{K})[t]$	anillo de los polinomios con coeficientes matriciales
$adj(A)$	matriz adjunta de A
$\det A$	el determinante de A
$tr A$	la traza de A

Capítulo 1

Introducción

Las ecuaciones diferenciales lineales de orden n, se pueden expresar como un sistema de n ecuaciones lineales de primer orden, y usando la notación matricial se pueden escribir

$$\mathbf{x}'(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{b}(t).$$

Nosotros abordaremos el caso en que los coeficientes son constantes, donde probaremos que

$$\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}(t), \text{ con la condición inicial } \mathbf{x}(t_0) = x_0$$

tiene un única solución, para todo t donde $\mathbf{b}$ es una función vectorial continua y calcularemos de forma explícita la solución. Los resultados son totalmente análogos para el caso $n = 1$, donde el problema de valores iniciales

$$y'(t) = ay(t) + b(t), \text{ con la condición inicial } y(t_0) = x_0$$

con $a \in \mathbb{R}$ y $b(t)$ continua en un intervalo I, tiene una única solución φ que viene dada por la fórmula

$$\varphi(t) = e^{(t-t_0)a}x_0 + \int_{t_0}^{t} e^{(t-s)a}b(s)ds, \ \forall t \in I$$

donde ahora a es una matriz cuadrada $n \times n$ y $b(t)$ una función vectorial.

Entonces toda la dificultad radicará en definir, operar y calcular la matriz exponencial e^A, que se define como una serie de potencias. En nuestro desarrollo la idea es transformar e^A en e^J, siendo J una matriz diagonal en bloques que es más fácil de calcular. Para ello, se deben estudiar dos situaciones básicamente con respecto a la matriz A: que sea diagonalizable o que no lo sea. Esto es lo que abordaremos en el capítulo 2, y lo central del trabajo es estudiar lo que se llama la forma canónica de Jordan, que es una forma normal para una matriz cuadrada con coeficientes en los reales o los complejos, que se suele emplear cuando la matriz no es diagonalizabe. Su existencia depende de que el polinomio característico de la matriz se pueda factorizar completamente. Esto ocurre siempre y cuando el cuerpo sean los complejos y hay una versión para cuando esto no ocurre en el caso real. Esto último no lo desarrollaremos.

Nuestro principal objetivo es comprender cómo se construye la forma de Jordan de una matriz y ver cómo aplicarla para resolver ecuaciones diferenciales. Concretamente se van a estudiar los siguientes temas:

1. Forma de Jordan.
2. Exponencial matricial.
3. Ecuaciones diferenciales lineales con coeficientes constantes de orden arbitrario y sistemas cuadrados de ecuaciones diferenciales con coeficientes constantes.

El primer tema es de álgebra lineal avanzada, pero no suele verse en los cursos de formación de profesores. En este trabajo probaremos la existencia de la forma de Jordan de una matriz sin usar transformaciones lineales. Este es un enfoque que no es común de encontrar en los libros de texto sobre el tema.
El segundo implica un poco de análisis, dado que la exponencial matricial se define mediante una serie en un espacio matricial. Los dos temas anteriores están vinculados dado que la forma de Jordan se usa para simplificar el cálculo de la exponencial.

El tercero consiste en el estudio de los sistemas lineales de ecuaciones diferenciales de primer orden a coeficientes constantes; acá veremos que su resolución pasa por calcular la exponencial de una matriz; y luego veremos que se puede reducir una ecuación diferencial lineal de orden n a un sistema de n ecuaciones lineales de primer orden, y a partir de ahí deduciremos como obtener sus soluciones.

Prerrequisitos: conocimientos básicos de álgebra lineal, suma directa, diagonalización, etc., algo de análisis en espacios normados (para la exponencial matricial) y nociones básicas de ecuaciones diferenciales y cierta familiaridad con el tema. En esta tesina trabajaremos siempre sobre los números reales o complejos y $\mathbb{K}$ denotará a $\mathbb{R}$ o $\mathbb{C}$.

Capítulo 2

Forma de Jordan

Para estudiar ciertos problemas relacionados con una matriz A, es útil en la medida de lo posible, transformar el problema en otro que involucre trabajar con una matriz que sea más simple. El caso más simple es cuando la matriz es diagonalizable, pero no toda matriz es diagonalizable. Sin embargo es posible representar la matriz A por medio de una matriz triangular superior en bloques de una forma particular sencilla. Esta representación se llama *forma canónica de Jordan*, y es útil para la resolución de sistemas lineales de ecuaciones diferenciales del tipo $\mathbf{x}' = A\mathbf{x}$. En este capítulo veremos como obtener la forma de Jordan de una matriz.

2.1. Suma directa de subespacios

Definición 2.1.1. Decimos que un espacio vectorial V es *suma directa* de una familia de subespacios $\{W_1, \cdots, W_m\}$ de V y escribimos $V = W_1 \oplus \cdots \oplus W_m$ si para todo $v \in V$ existen y son únicos $w_i \in W_i$, $i = 1, \cdots, m$, tales que $v = \sum_{i=1}^m w_i$.

Ejemplo 2.1.2. En el espacio $\mathcal{M}_n(\mathbb{K})$, si consideramos

$$W_1 = \{(a_{ij}) : a_{ij} = 0 \text{ si } i \leq j\}\,, \quad W_2 = \{(a_{ij}) : a_{ij} = 0 \text{ si } i \geq j\}\,, \quad W_3 = \{(a_{ij}) : a_{ij} = 0 \text{ si } i \neq j\}$$

es fácil de probar que cualquier matriz del espacio se puede escribir de una única forma como $A = T_1 + T_2 + T_3$ donde $T_1 \in W_1$, $T_2 \in W_2$ y $T_3 \in W_3$, por lo cual podemos escribir $\mathcal{M}_n = W_1 \oplus W_2 \oplus W_3$. Esquemáticamente

$$\begin{pmatrix} * & * \\ * & * \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ * & 0 \end{pmatrix} + \begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}.$$

Proposición 2.1.3. *Sean $\{W_1, \cdots, W_m\}$ subespacios de un espacio de dimensión finita V, tales que $V = \bigoplus_1^m W_i$. Entonces tenemos*

1. *Si $\mathcal{B}_i$ es una base de W_i, para todo $i = 1, \cdots, m$, entonces $\mathcal{B} := \bigcup_{i=1}^m \mathcal{B}_i$ es una base de V.*
2. *Vale $\dim V = \sum_{i=1}^m \dim W_i$.*

Dem. Consideremos para cada $i = 1, \cdots, m$, $\mathcal{B}_i = \left\{w_1^i, \cdots, w_{r_i}^i\right\}$ una base para W_i.
Sea ahora un vector cualquiera $v \in V$, entonces por hipótesis podemos escribir $v = w_1 + \cdots + w_m$, con $w_i \in W_i$, $i = 1, \cdots, m$, pero a su vez cada w_i se puede escribir como combinación lineal de la base $\mathcal{B}_i$: $w_i = \alpha_1^i w_1^i + \cdots + \alpha_{r_i}^i w_{r_i}^i$, por lo cual

$$v = \left(\alpha_1^1 w_1^1 + \cdots + \alpha_{r_1}^1 w_{r_1}^1\right) + \cdots + \left(\alpha_1^m w_1^m + \cdots + \alpha_{r_m}^m w_{r_m}^m\right),$$

lo que demuestra que $\mathcal{B}$ es un generador de V.
Supongamos ahora que existen escalares β_j^i tales que

$$\beta_1^1 w_1^1 + \cdots + \beta_{r_1}^1 w_{r_1}^1 + \cdots + \beta_1^m w_1^m + \cdots + \beta_{r_m}^m w_{r_m}^m = 0$$

pero como $\beta_1^i w_1^i + \cdots + \beta_{r_i}^i w_{r_i}^i \in W_i$, para todo i, entonces por la unicidad de la definición de suma directa se deduce que $\beta_1^i w_1^i + \cdots + \beta_{r_i}^i w_{r_i}^i = 0$ para todo i, y como cada $\mathcal{B}_i$ es linealmente independiente se deduce que $\beta_j^i = 0$ para todo $i = 1, \cdots, m$ y para todo $j = 1, \cdots, r_i$, por lo cual se concluye que $\mathcal{B}$ es linealmente independiente y como es generador de V, concluimos que $\mathcal{B}$ es una base del espacio V.
La segunda afirmación se deduce inmediatamente de la primera. $\square$

Observación 2.1.4. Es fácil probar que vale: $V = W_1 \oplus W_2$ si y solo si para todo $v \in V$ existen $w_1 \in W_1$ y $w_2 \in W_2$ tales que $v = w_1 + w_2$ y vale $W_1 \cap W_2 = \{0\}$. Alternativamente (en dimensión finita), $V = W_1 \oplus W_2$ si y solo si $W_1 \cap W_2 = \{0\}$ y $\dim V = \dim W_1 + \dim W_2$.

2.2. Diagonalización

Observación 2.2.1. En este capítulo todas las matrices pertenecen al espacio vectorial $\mathcal{M}_n(\mathbb{K})$ y las operaciones son las usuales.

Definición 2.2.2. Dos matrices $A, B \in \mathcal{M}_n(\mathbb{K})$ se dice que son *semejantes* si existe una matriz $P \in \mathcal{M}_n(\mathbb{K})$ invertible, tal que $B = P^{-1}AP$ o, equivalentemente, $A = PBP^{-1}$. Escribiremos $A \sim B$ y se lee A es *semejante* a B.

Proposición 2.2.3. *La relación de semejanza así definida es una relación de equivalencia en* $\mathcal{M}_n(\mathbb{K})$.

Dem. Tenemos que demostrar que la relación es reflexiva, simétrica y transitiva.
• Reflexiva. Como $A = I^{-1}AI$ entonces $A \sim A$.
• Simétrica. Sea $A \sim B$, entonces existe P invertible tal que $B = P^{-1}AP$ o lo que es lo mismo $A = PBP^{-1}$, y si llamamos $P^{-1} = Q$, nos queda que $A = Q^{-1}BQ$ o sea $B \sim A$.
• Transitiva. Sean $A \sim B$ y $B \sim C$, entonces existen

$$P \text{ y } Q \text{ invertibles tales que } B = P^{-1}AP \text{ y } C = Q^{-1}BQ.$$

Entonces

$$C = Q^{-1}BQ = Q^{-1}P^{-1}APQ = (PQ)^{-1}A(PQ).$$

Llamando $T = PQ$ llegamos a que $A \sim C$ lo que concluye la demostración. $\square$

Uno de los problemas con matrices, y es de nuestro interés, es calcular potencias.
Recordemos:

Definición 2.2.4. Si $A \in \mathcal{M}_n$ definimos

$$A^0 := I, \ \ A^{k+1} := A^k A, \ \ \text{para todo } k \in \mathbb{N}.$$

De esta forma queda definida la potencia A^k, para todo k natural y toda matriz, $A \in \mathcal{M}_n$.

Observar que si A es diagonal, $A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$, entonces vale $A^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{pmatrix}$, para todo k natural.

En general la fórmula para calcular las entradas de la potencia A^k es complicada. Observar que si $A \sim B$ y $A = PBP^{-1}$ con P invertible, entonces $A^2 = PBP^{-1}PBP^{-1} = PBIBP^{-1} = PB^2P^{-1}$ y en general, por inducción obtenemos, $A^k = P^{-1}B^kP$ para todo $k \in \mathbb{N}$.
El objetivo nuestro es, dada $A \in \mathcal{M}_n$, encontrar $B \in \mathcal{M}_n$ semejante a A, para la cual sea más fácil calcular B^k.

Matrices en bloques

Primero comentemos brevemente que es una matriz en bloques, las cuales usaremos en este trabajo. Muchas veces es aconsejable dividir una matriz en submatrices, por ejemplo cuando se manipula matrices con gran número de filas y columnas, puede ser conveniente descomponerla en matrices más pequeñas, porque permite resolver problemas de manera más simple.

Definición 2.2.5. Diremos que una matriz $\mathcal{M}_{n\times m}(\mathbb{K})$ está *escrita en bloques* si está subdividida de la forma

$$A = \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1r} \\ B_{21} & B_{22} & \dots & B_{2r} \\ \vdots & \vdots & \dots & \vdots \\ B_{s1} & B_{s2} & \dots & B_{sr} \end{pmatrix},$$

donde B_{ij} son submatrices de tamaño $n_i \times m_j$ donde $n_1, \cdots, n_s$ y $m_1, \cdots, m_r$ son números naturales no nulos tal que $n_1 + \cdots + n_s = n$ y $m_1 + \cdots + m_r = m$.

En consecuencia, los bloques se obtienen trazando imaginariamente rectas verticales y horizontales entre los elementos de la matriz A. En este caso los bloques o submatrices los designamos en la forma B_{ij}. Observar que el número de columnas en el bloque B_{ij} depende solo de j, siendo el mismo para todos los i; ídem para las filas.

Ejemplo 2.2.6. La matriz A

$$A = \left(\begin{array}{c|cc|cc} 0 & 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 & 9 \\ 1 & 2 & 3 & 4 & 5 \\ \hline 6 & 7 & 8 & 8 & 9 \end{array}\right) = \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{pmatrix}$$

que está descompuesta en bloques de 2 filas y 3 columnas, es decir es una matriz 2×3 en bloques, donde estos son:

$$B_{11} = \begin{pmatrix} 0 \\ 5 \\ 1 \end{pmatrix}, \; B_{12} = \begin{pmatrix} 1 & 2 \\ 6 & 7 \\ 2 & 3 \end{pmatrix}, \; B_{13} = \begin{pmatrix} 3 & 4 \\ 8 & 9 \\ 4 & 5 \end{pmatrix}, \; B_{21} = \begin{pmatrix} 6 \end{pmatrix}, \; B_{23} = \begin{pmatrix} 7, 8 \end{pmatrix}, \; B_{23} = \begin{pmatrix} 8, 9 \end{pmatrix}$$

Observación 2.2.7.

1. Es claro que una matriz se puede expresar en bloques de muchas maneras distintas.
2. Se puede desarrollar un álgebra de matrices en bloques, sumando, multiplicando, hallando la inversa, etc., de matrices en bloques. Para obtener más información sobre el tema, ver [Dc]. Nosotros haremos uso de estos resultados.
3. En particular, si $A \in \mathcal{M}_n(\mathbb{K})$ y $B \in \mathcal{M}_m(\mathbb{K})$ entonces podemos definir la *suma directa* como la matriz de orden $(n+m) \times (n+m)$, como

   $$A \oplus B := \left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right)$$

   Esta caracterización se extiende para 3 o más matrices.

4. Nuestro interés es trabajar solamente con matrices cuadradas en bloques, y nos van a interesar únicamente las matrices en bloques, de la forma

$$\begin{pmatrix} \boxed{B_1} & & \\ & \ddots & \\ & & \boxed{B_r} \end{pmatrix} \tag{2.2.1}$$

donde los bloques $B_{11} = B_1, \cdots, B_{rr} = B_r$ son submatrices cuadradas, y los demás elementos son todos 0.

Notar que si $A \in \mathcal{M}_n$ es una matriz en bloques de la forma $\begin{pmatrix} \boxed{A_1} & & \\ & \ddots & \\ & & \boxed{A_r} \end{pmatrix}$, $A_i \in \mathcal{M}_{n_i}$, $i = 1, 2, \cdots, r$, $\quad n_1 + n_2 + \cdots + n_r = n$, entonces vale

$$A^k = \begin{pmatrix} \boxed{A_1^k} & & \\ & \ddots & \\ & & \boxed{A_r^k} \end{pmatrix}, \text{ para todo } k \text{ natural.}$$

Luego nuestro objetivo es, dada $A \in \mathcal{M}_n$ encontrar $B \in \mathcal{M}_n$, $B \sim A$ tal que B sea una matriz en bloques de la forma 2.2.1, tales que estos bloques tengan potencias fáciles de calcular.

Retomando nuestro problema original de hallar la potencia k ésima de una matriz A, nos encontramos en la siguiente situación:

Dada $A \in \mathcal{M}_n$ tenemos las siguientes posibilidades

1) A es diagonal.

2) $A \sim D$ con D matriz diagonal.

3) ninguna de los anteriores (aquí usaremos las matrices en bloques).

Veamos primero los dos primeros casos.

1) Como ya dijimos, si A es diagonal, $A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$ entonces $A^k = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{pmatrix}$, para todo k natural.

2) Si existe $P \in \mathcal{M}_n$, invertible y $D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$, diagonal, tal que $A = PDP^{-1}$, entonces $A^k = PD^kP^{-1}$, para todo k natural, donde ahora es fácil de calcular las potencias de D y por lo tanto las potencias de A.

El problema es cómo saber si A es semejante a una matriz diagonal, y si es así como encontrarla. Para eso introducimos la siguiente definición.

Definición 2.2.8. Una matriz $A \in \mathcal{M}_n$ es *diagonalizable* si existe una matriz diagonal D tal que A es semejante a D.

Observemos entonces que para que A sea diagonalizable debe existir una matriz P invertible tal que $A = PDP^{-1}$, con D diagonal. Luego debe cumplirse $AP = PD$, y por lo tanto

$$AP = PD = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \dots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \dots & \lambda_n p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \dots & \lambda_n p_{nn} \end{pmatrix}$$

Si llamamos $\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_n$, a los vectores columnas de P y como las columnas de la matriz AP y las correspondientes de PD han de coincidir, entonces

$$A\mathbf{p}_1 = \lambda_1 \mathbf{p}_1, \quad A\mathbf{p}_2 = \lambda_2 \mathbf{p}_2, \quad \cdots, A\mathbf{p}_n = \lambda_n \mathbf{p}_n,$$

es decir, han de existir n vectores $\mathbf{p}_i$ y n escalares λ_i tales que $A\mathbf{p}_i = \lambda_i \mathbf{p}_i$, para $i = 1, 2, \cdots, n$. Como la matriz P es invertible, los vectores $\mathbf{p}_i$ son distintos de $\mathbf{0}$ y linealmente independientes. Estos elementos, $\mathbf{p}_i$ y λ_i son los que mas adelante llamaremos vectores y valores propios, respectivamente.

Entonces una solución al problema de la diagonalización conduce de manera natural a los conceptos de valores y vectores propios que ahora introduciremos formalmente . También, estos conceptos juegan un papel muy importante y son unas herramientas muy valiosas en el estudio de las matrices no diagonalizables, como veremos más adelante.

Valores y vectores propios.

Definición 2.2.9. Sea una matriz $A \in \mathcal{M}_n(\mathbb{K})$. Se dice que un escalar $\lambda \in \mathbb{K}$ es un *valor propio o autovalor* de A si existe un vector $0 \neq v \in \mathbb{K}^n$ tal que $Av = \lambda v$.
En ese caso, decimos que v es un *vector propio o autovector* de A asociado al valor propio λ.

Ejemplo 2.2.10. Si consideramos la matriz diagonal

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

entonces $\lambda_1, \lambda_1, \cdots, \lambda_n$ son valores propios con vectores propios respectivos $e_1, e_1, \cdots, e_n$, que son los vectores de la base canónica de $\mathbb{K}^n$, pues $Ae_i = \lambda_i e_i$, para todo $i =, 1, 2, \cdots, n$.

Ejemplo 2.2.11. Sea $A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 5 \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$.

El vector

$$v = \begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix}$$

verifica que $Av = iv$, por lo que v es un vector propio de A asociado con el valor propio $\lambda = i$. También el vector e_3 es vector propio de A con valor propio $\mu = 5$, pues $Ae_3 = 5e_3$.
Observemos que si pensamos $A \in \mathcal{M}_n(\mathbb{R})$, entonces el vector v no tiene sentido pues $\lambda \notin \mathbb{R}$, pero e_3 sigue siendo vector propio de la matriz A. Esto muestra la importancia del cuerpo en el cual se está trabajando.

Definición 2.2.12. Rango columna de una matriz.
Sea una matriz $A \in \mathcal{M}_n(\mathbb{K})$. Se llama *rango por columnas de* A, y se escribe $rg_C(A)$, a la dimensión del subespacio de $\mathbb{K}^n$ generado por las columnas de A.

Definición 2.2.13. Rango fila de una matriz.
Sea una matriz $A \in \mathcal{M}_n(\mathbb{K})$. Llamaremos *rango por filas de* A, y escribiremos $rg_F(A)$, a la dimensión del subespacio de $\mathbb{K}^n$ generado por las filas de A.

Observación 2.2.14.

1. El rango por columnas de cualquier matriz es igual al número máximo de columnas linealmente independientes de dicha matriz.

2. El rango por filas de cualquier matriz es igual al número máximo de filas linealmente independientes de dicha matriz.

3. Se demuestra que para cualquier matriz A, se cumple que $rg_C(A) = rg_F(A)$, lo que nos lleva a la siguiente definición:

Definición 2.2.15. Sea una matriz $A \in \mathcal{M}_n(\mathbb{K})$. Al número $rg_C(A) = rg_F(A)$ lo llamaremos *rango de la matriz* A, y lo escribiremos $\text{rango}(A)$.

Observación 2.2.16. Se prueba que las matrices semejantes tienen el mismo rango.

Definición 2.2.17. Si $A \in \mathcal{M}_n(\mathbb{K})$, definimos el *núcleo* de A y escribimos $\ker(A)$, al conjunto $\ker(A) := \{v \in \mathbb{K}^n : Av = 0\}$.

Observar que $\ker(A)$ es un subespacio de $\mathbb{K}^n$, para toda $A \in \mathcal{M}_n$ y $\dim \ker(A) = n - rango(A)$.

Definición 2.2.18. Si $\lambda \in \mathbb{K}$ es un valor propio de A, el *subespacio propio* asociado a λ es el subespacio $E_\lambda := \ker(A - \lambda I) = \{v \in \mathbb{K}^n : Av = \lambda v\}$.

Es trivial probar que E_λ es un subespacio del espacio vectorial $\mathbb{K}^n$.
Haciendo referencia al ejemplo 2.2.11, se prueba que los valores propios de A son 5, i, y $-i$ y sus respectivos subespacios propios son

$$E_5 = [(0,0,1)], \quad E_i = [(i,1,0)], \quad E_{-i} = [(i,-1,0)]$$

Por otro lado, el único valor propio real de A es 5 y el subespacio propio correspondiente es
$E_5 = [(0,0,1)]$.

Proposición 2.2.19. *Sea* $A \in \mathcal{M}_n(\mathbb{K})$. *Entonces un escalar* $\lambda \in \mathbb{K}$ *es un valor propio de* A *si y solo si* $\det(A - \lambda I) = 0$.

Dem. El escalar λ es un valor propio de A si solo si existe un vector $v \in \mathbb{K}^n$, $v \neq 0$ tal que $Av = \lambda v$ si y solo si el sistema $(A - \lambda I)v = 0$ tiene solución distinta de la trivial si y solo si $\det(A - \lambda I) = 0$. $\square$

Definición 2.2.20. Si $A \in \mathcal{M}_n(\mathbb{K})$, se denomina *polinomio característico* de A, al polinomio $\mathcal{X}_A(t) := \det(A - tI) \in \mathbb{K}[t]$.

Observación 2.2.21. Sea $A \in \mathcal{M}_n(\mathbb{K})$.

1. El polinomio caraterístico de A es un polinomio de grado n con coeficiente principal $(-1)^n$.

2. Un escalar λ es un valor propio de A si y solo si es una raíz del polinomio $\mathcal{X}_A(t)$. Ver la proposición 2.2.19

3. A tiene como máximo n valores propios (distintos).

4. Si $A \in \mathcal{M}_n(\mathbb{K})$ tiene valores propios, $\lambda_1, \lambda_2, \cdots, \lambda_n$ entonces

   a) $\det A = \lambda_1 \lambda_2 \cdots \lambda_n$

   b) $trA = \lambda_1 + \lambda_2 + \cdots + \lambda_n$

Proposición 2.2.22. *Si $A \sim B$ entonces $\mathcal{X}_A(t) = \mathcal{X}_B(t)$.*

Dem. Si A, B son matrices semejantes, entonces existe una matriz P invertible, tal que $B = P^{-1}AP$, entonces usando propiedades de los determinantes

$$\begin{aligned}
\mathcal{X}_B(t) &= \det(B - tI) = \det\left(P^{-1}AP - tI\right) = \det\left(P^{-1}AP - tP^{-1}IP\right) \\
&= \det\left(P^{-1}(A - tI)P\right) = \det(P^{-1}) \det(A - tI) \det P \\
&= (\det P)^{-1} \det(A - tI) \det P = \det(A - tI) \\
&= \mathcal{X}_A(t).
\end{aligned}$$

□

Observación 2.2.23. Recordar que una matriz cuadrada P es invertible si y solo si $\det P \neq 0$, si y solo si las columnas de P forman un conjunto linealmente independiente.

Proposición 2.2.24. *Si $A \in \mathcal{M}_n(\mathbb{K})$, entonces son equivalentes:*

a) La matriz A es diagonalizable.

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

Además, si dicha base es $\mathcal{B} = \{p_1, p_2, \ldots, p_n\}$ y $\lambda_1, \cdots, \lambda_n$ son los valores propios correspondientes, entonces se cumple

$$A = PDP^{-1}, \quad \textit{siendo } D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \quad y \quad P = \left(p_1 | p_2 | \ldots | p_n\right).$$

Acá estamos pensando a $p_1, \cdots, p_n$ como vectores columna.

Dem. *a)* $\Rightarrow$ *b)* Entonces existen una matriz invertible $P = \left(p_1 | p_2 | \ldots | p_n\right)$, y una matriz diagonal $D = \operatorname{diag}\left(\lambda_1, \lambda_2, \ldots, \lambda_n\right)$, tales que $A = PDP^{-1}$ o $AP = PD$. Por lo tanto

$$\left(Ap_1 | Ap_2 | \ldots | Ap_n\right) = AP = PD = \left(p_1 | p_2 | \ldots | p_n\right) \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} = \left(\lambda_1 p_1 | \lambda_2 p_2 | \ldots | \lambda_n p_n\right)$$

entonces, como las matrices de los extremos son iguales, las columnas correspondientes deben ser iguales, es decir,

$$Ap_1 = \lambda_1 p_1,\ Ap_2 = \lambda_2 p_2,\ \cdots, Ap_n = \lambda_n p_n$$

y como P es invertible , entonces, sus vectores columnas p_i son no nulos y por lo tanto son vectores propios de A asociados a los valores propios correspondientes λ_i. Además $p_1, \cdots, p_n$ son linealmente independientes, luego los vectores columnas de P forman una base para $\mathbb{K}^n$.

b) $\Rightarrow$ *a)* Supongamos ahora que A tiene n vectores propios linealmente independientes $p_1, p_2, \cdots, p_n$ asociados a los valores propios $\lambda_1, \lambda_2, \cdots, \lambda_n$. Entonces $Ap_i = \lambda_i p_i$, para $i = 1, 2, \cdots, n$, pero esto es lo mismo que escribir $AP = PD$, y como los vectores p_i son linealmente independientes (rango$A = n$), se sigue que P es invertible y por lo tanto $A = PDP^{-1}$, con D diagonal. $\square$

Ejemplo 2.2.25.

1. El polinomio característico de la matriz $A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ del ejemplo 2.2.11, es

   $$\mathcal{X}_A(t) = \det(A - \lambda I) = -(t-5)(t^2+1) = -(t-5)(t-i)(t+i),$$

   luego en $\mathbb{C}$ tiene como valores propios a 5, i y $-i$ y los subespacios propios correspondientes son

   $$E_5 = [(0,0,1)], \quad E_i = [(i,1,0)], \quad E_{-i} = [(i,-1,0)]$$

   Tenemos entonces una base $\mathcal{B} = \{(0,0,1), (i,1,0), (i,-1,0)\}$ de $\mathbb{C}^3$ formada por vectores propios de A. Luego la matriz A es diagonalizable y es semejante a la matriz diagonal

   $$D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix} \text{ y matriz de semejanza } P = \begin{pmatrix} 0 & i & i \\ 0 & 1 & -1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{donde } A = PDP^{-1}.$$

   Pero en $\mathbb{R}$ el polinomio característico $\mathcal{X}_A(t)$ tiene una única raíz, entonces hay un único valor propio real que es $\lambda = 5$ y $E_5 = [(0,0,1)]$. Luego no existe una base de $\mathbb{R}^3$ formada por vectores propios de A, y por lo tanto A no es diagonalizable en $\mathbb{R}$.

2. La matriz $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, no es diagonalizable, pues tiene como valores propios a $\lambda_1 = \lambda_2 = 0$, entonces $D = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Y si A fuera diagonalizable debería existir una matriz P invertible tal que $A = PDP^{-1} = 0$, lo que es evidentemente absurdo. Observemos que en este caso el cuerpo no importa.

Proposición 2.2.26. *Sean $v_1, \cdots, v_k$, vectores propios de una matriz $A \in \mathcal{M}_n(\mathbb{K})$ asociados a valores propios distintos, $\lambda_1, \cdots, \lambda_k$, respectivamente. Entonces el conjunto de vectores $\{v_1, \cdots, v_k\}$ es linealmente independiente.*

Dem. Lo haremos por inducción sobre k.
Si $k = 1$, entonces $\{v_1\}$ es linealmente independiente porque $v_1 \neq 0$ por ser vector propio.

Supongamos ahora que se cumple para $k = r$, es decir que si tenemos un conjunto de r vectores propios asociados a r valores propios distintos, entonces dicho conjunto es linealmente independiente. Entonces probaremos que se cumple también para $k = r+1$: sean $v_1, \cdots, v_{r+1}$ vectores propios de A asociados a $r+1$ valores propios distintos, $\lambda_1, \cdots, \lambda_{r+1}$.

Sea

$$\sum_{i=1}^{r+1} \alpha_i v_i = 0 \tag{2.2.2}$$

Multiplicando en (2.2.2) por $A - \lambda_{r+1} I$, tenemos

$$\begin{aligned}
0 &= (A - \lambda_{r+1}I) \sum_{i=1}^{r+1} \alpha_i v_i = (A - \lambda_{r+1}I) \sum_{i=1}^{r} \alpha_i v_i + (A - \lambda_{r+1}I)\alpha_{r+1} v_{r+1} \\
&= \sum_{i=1}^{r} \alpha_i (A - \lambda_{r+1}I) v_i + \alpha_{r+1}(A - \lambda_{r+1}I) v_{r+1} = \sum_{i=1}^{r} \alpha_i (A v_i - \lambda_{r+1} I v_i) + \alpha_{r+1}(A v_{r+1} - \lambda_{r+1} I v_{r+1}) \\
&= \sum_{i=1}^{r} \alpha_i (\lambda_i v_i - \lambda_{r+1} v_i) + \alpha_{r+1}(\lambda_{r+1} v_{r+1} - \lambda_{r+1} v_{r+1}) = \sum_{i=1}^{r} \alpha_i (\lambda_i - \lambda_{r+1}) v_i + 0
\end{aligned}$$

ya que $Av_i = \lambda_i v_i$, para todo $i = 1, \cdots, r+1$. Por lo tanto

$$\sum_{i=1}^{r} \alpha_i (\lambda_i - \lambda_{r+1}) v_i = 0 \tag{2.2.3}$$

Como por hipotesis de inducción el conjunto $\{v_1, \cdots, v_r\}$ es linealmente independiente y $\lambda_{r+1} \neq \lambda_i$, para todo $i = 1, \cdots, r$, entonces de (2.2.3) se deduce que $\alpha_i = 0$, para todo $i = 1, \cdots, r$.
Si volvemos ahora a (2.2.2), tenemos que $\alpha_{r+1} v_{r+1} = 0$, y como $v_{r+1} \neq 0$, debe ser necesariamente $\alpha_{r+1} = 0$, de lo cual se concluye que la única combinación lineal del conjunto $\{v_1, \cdots, v_r, v_{r+1}\}$ que da el vector nulo es la trivial, con lo que se completa la demostración. □

Corolario 2.2.27. *Si una matriz de orden n tiene n valores propios distintos, entonces A es diagonalizable.* □

Ejemplo 2.2.28. Si $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, es $\mathcal{X}_A(t) = t(t-2)$. Luego, por corolario anterior, A es diagonalizable y es semejante a la matriz diagonal $D = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$, por la proposición 2.2.24.

Multiplicidad geométrica y algebraica

Definición 2.2.29. Sean una matriz $A \in \mathcal{M}_n(\mathbb{K})$ y λ un valor propio de A.

1. La *multiplicidad geométrica* de λ es $MG(\lambda) := \dim E_\lambda = \dim \{v \in \mathbb{K}^n : Av = \lambda v\}$.
2. La *multiplicidad algebraica* de λ es $MA(\lambda) := \text{máx}\left\{h : (t-\lambda)^h \text{ divide a } \mathcal{X}_A(t)\right\}$.

Observar que $MG(\lambda) \geq 1$ ($E_\lambda \neq \{0\}$) y $MA(\lambda) \geq 1$.

Proposición 2.2.30. *Sean A una matriz de tamaño $n \times n$ y λ un valor propio de A, entonces $MG(\lambda) \leq MA(\lambda)$.*

Dem. Supongamos que la multiplicidad geométrica de λ es p, esto es $\dim E_\lambda = p$. Entonces existe una base de E_λ de la forma $\mathcal{B} = \{v_1, v_2, \ldots, v_p\}$. Sea ahora Q, una matriz invertible de tamaño $n \times n$, que tiene a $v_1, v_2, \cdots, v_p$ como las primeras p columnas,

$$Q = \left(v_1 | v_2 | \ldots | v_p | v_{p+1} | \ldots | v_n\right)$$

que la podemos escribir como una matriz en bloques:

$$Q = (v_1|v_2|\dots|v_p|v_{p+1}|\dots|v_n) = \left(\begin{array}{c|c} U & V \end{array} \right) \text{ y la inversa en bloques } Q^{-1} = \begin{pmatrix} C \\ D \end{pmatrix}$$

donde C es $p \times n$. Tenemos entonces que las columnas de la matriz U son los vectores propios correspondiente al valor propio λ, entonces $AU = \lambda U$. Luego tenemos

$$\left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & I_{n-p} \end{array} \right) = I_n = Q^{-1}Q = \begin{pmatrix} C \\ D \end{pmatrix} \left(\begin{array}{c|c} U & V \end{array} \right) = \left(\begin{array}{c|c} CU & CV \\ \hline DU & DV \end{array} \right)$$

Por lo tanto tenemos que $CU = I_p$, $CV = 0$, $DU = 0$ y $DV = I_{n-p}$. Luego,

$$Q^{-1}AQ = \begin{pmatrix} C \\ D \end{pmatrix} A \left(\begin{array}{c|c} U & V \end{array} \right) = \left(\begin{array}{c|c} CAU & CAV \\ \hline DAU & DAV \end{array} \right)$$

$$= \left(\begin{array}{c|c} \lambda CU & CAV \\ \hline \lambda DU & DAV \end{array} \right) = \left(\begin{array}{c|c} \lambda I_p & CAV \\ \hline 0 & DAV \end{array} \right)$$

Entonces

$$Q^{-1}AQ - tI_n = \left(\begin{array}{c|c} \lambda I_p - tI_p & CAV \\ \hline 0 & DAV - tI_{n-p} \end{array} \right)$$

Y como esta última es una matriz en bloques donde, $(\lambda - t)I_p$ y $DAV - tI_{n-p}$ son matrices cuadradas, tenemos:

$$\det(Q^{-1}AQ - tI_n) = \det((\lambda - t)I_p)\det(DAV - tI_{n-p}) = (\lambda - t)^p \det(DAV - tI_{n-p})$$

Pero $\det(Q^{-1}AQ - tI_n)$ es el polinomio característico de $Q^{-1}AQ$ y como A y $Q^{-1}AQ$ son semejantes y se cumple por la proposición 2.2.22 que $\mathcal{X}_{Q^{-1}AQ}(t) = \mathcal{X}_A(t)$ tenemos entonces que

$$\mathcal{X}_A(t) = (\lambda - t)^p \det(DAV - tI_{n-p})$$

lo que demuestra que $MA(\lambda) \geq p$ o sea $MA(\lambda) \geq MG(\lambda)$. □

Definición 2.2.31. Un polinomio no constante $f(x) \in \mathbb{K}[x]$ se dice que se *escinde* (en $\mathbb{K}$) si existen escalares, $a, a_1, a_2, \cdots, a_n$, no necesariamente distintos, tales que $f(x) = a(x - a_1)\cdots(x - a_n)$.

Observación 2.2.32.

1. Si el polinomio $f(x)$ de grado n se escinde, agrupando los términos repetidos, podemos escribir

$$f(x) = a(x - b_1)^{n_1} \cdots (x - b_r)^{n_r}, \text{ con } b_i \neq b_j \text{ si } i \neq j,\ n_i \in \mathbb{Z}^+,\ n_1 + \cdots + n_r = n.$$

2. En $\mathbb{C}[x]$ todo polinomio no constante se escinde (*teorema fundamental del álgebra*).

Teorema 2.2.33. *Sea A una matriz de orden n. Las siguientes afirmaciones son equivalentes:*

a) La matriz A es diagonlizable.

b) El polinomio característico $\mathcal{X}_A(t)$ se escinde y para cada valor propio λ, $MG(\lambda) = MA(\lambda)$.

c) Existe una base de $\mathbb{K}^n$ formada uniendo bases de los subespacios propios.

Dem. $a) \Rightarrow b)$ Suponemos entonces que la matriz A es diagonalizable, entonces existe una matriz P invertible tal que

$$P^{-1}AP = D = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_r \end{pmatrix}.$$

siendo $\lambda_1, \cdots, \lambda_r$ escalares distintos.

Y como A y D son semejantes entonces por la proposición 2.2.22, $\mathcal{X}_A(t) = \mathcal{X}_D(t)$. Luego, $\det(A - tI) = \det(D - tI) = (-1)^n(t-\lambda_1)^{n_1}\cdots(t-\lambda_r)^{n_r},$ siendo $n_1 + \cdots + n_r = n$ por lo cual el polinomio característico de A se escinde. Ahora observemos que como A y D son matrices semejantes entonces

$$D - tI = P^{-1}AP - tP^{-1}P = P^{-1}(AP - tP) = P^{-1}(A - tI)P, \text{ para todo } t \in \mathbb{K}$$

entonces podemos decir que $A - tI$ y $D - tI$ son semejantes para todo $t \in \mathbb{K}$, de donde se deduce que $\text{rango}(A - tI) = \text{rango}(D - tI)$ para todo $t \in \mathbb{K}$. Por lo cual podemos afirmar que

$$\text{rango}(A - \lambda_i I) = \text{rango}(D - \lambda_i I) = n - n_i, \text{ para todo } i = 1, \cdots, r$$

luego se deduce que $\dim(E_{\lambda_i}) = n_i$ para todo $i = 1, \cdots, r$, es decir que $MG(\lambda_i) = MA(\lambda_i)$ para todo $i = 1, \cdots, r$, lo que termina de demostrar $b)$.

$b) \Rightarrow c)$ Supongamos entonces que es

$$\begin{aligned} &\mathcal{X}_A(t) = \det(A - tI) = (-1)^n(t-\lambda_1)^{n_1}\cdots(t-\lambda_r)^{n_r}, \text{ siendo } n_1 + \cdots n_r = n, \\ &\text{y que vale } \dim(E_{\lambda_i}) = n_i, \text{ para todo } i = 1, \cdots, r. \end{aligned}$$

Sea ahora $\mathcal{B}_i = \{v_{i1}, v_{i2}, \cdots, v_{in_i}\}$, una base de E_{λ_i}, para todo $i = 1, 2, \cdots, r$, y consideremos el conjunto

$$\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i = \{v_{11}, v_{12}, \cdots, v_{1n_1}, v_{21}, v_{22}, \cdots, v_{2n_2}, \cdots, v_{r1}, v_{r2}, \cdots, v_{rn_r}\}.$$

Notar que $\mathcal{B}$ es un conjunto de n vectores propios de la matriz A. Vamos a probar que es linealmente independiente. Para ello consideramos:

$$\begin{aligned} &\alpha_{11}v_{11} + \alpha_{12}v_{12} + \cdots + \alpha_{1n_1}v_{1n_1} + \alpha_{21}v_{21} + \alpha_{22}v_{22} + \cdots + \alpha_{2n_2}v_{2n_2} \\ &+ \cdots + \alpha_{r1}v_{r1} + \alpha_{r2}v_{r2} + \cdots + \alpha_{rn_r}v_{rn_r} \\ &= (\alpha_{11}v_{11} + \alpha_{12}v_{12} + \cdots + \alpha_{1n_1}v_{1n_1}) + (\alpha_{21}v_{21} + \alpha_{22}v_{22} + \cdots + \alpha_{2n_2}v_{2n_2}) + \end{aligned}$$

$$\cdots + (\alpha_{r1}v_{r1} + \alpha_{r2}v_{r2} + \cdots + \alpha_{rn_r}v_{rn_r}) = 0 \tag{2.2.4}$$

Ahora si definimos $x_i = \alpha_{i1}v_{i1} + \alpha_{i2}v_{i2} + \cdots + \alpha_{in_i}v_{in_i}, \quad i = 1, 2, \cdots, r$, podemos escribir

$$x_1 + x_2 + \cdots + x_r = 0 \tag{2.2.5}$$

Ahora $x_1, x_2, \cdots, x_r$ son vectores pertenecientes respectivamente a las espacios propios $E_{\lambda_1}, \cdots, E_{\lambda_r}$ correspondientes a valores propios distintos, $\lambda_1, \cdots, \lambda_r$, por lo tanto si no son todos 0, entonces son linealmente independientes. Ahora de (2.2.5) se deduce que $x_1, x_2, \cdots, x_r$ son linealmente dependientes, entonces necesarimente deben ser los $x_i = 0$ para todo $i = 1, \cdots, r$.
Volviendo ahora a (2.2.4), tenemos que $\alpha_{i1}v_{i1} + \alpha_{i2}v_{i2} + \cdots + \alpha_{in_i}v_{in_i} = 0$, y como $v_{i1}, v_{i2}, \cdots, v_{in_i}$ son

linealmente independientes, concluimos que $\alpha_{i1} = \alpha_{i2} = \cdots = \alpha_{in_i} = 0$ para $i = 1, \cdots, r$.
Por lo tanto la única combinación de $\mathcal{B}$ que da el vector nulo es la trivial, entonces $\mathcal{B}$ es un conjunto linealmente independiente con n vectores del espacio $\mathbb{K}^n$, es decir es $\mathcal{B}$ base de $\mathbb{K}^n$.

c) $\Rightarrow$ *a)* Sea ahora $\mathcal{B}_i = \{v_{i1}, v_{i2}, \cdots, v_{in_i}\}$, una base de E_{λ_i}, para todo $i = 1, 2, \cdots, r$, pero entonces, por hipótesis, el conjunto

$$\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i = \{v_{11}, v_{12}, \cdots, v_{1n_1}, v_{21}, v_{22}, \cdots, v_{2n_2}, \cdots, v_{r1}, v_{r2}, \cdots, v_{rn_r}\}$$

es una base de $\mathbb{K}^n$. Tenemos una base de $\mathbb{K}^n$ formada por vectores propios de A, entonces por la proposición 2.2.24, la matriz A es diagonalizable. □

2.3. Forma de Jordan. Caso nilpotente.

Como ya vimos, no todas las matrices son diagonalizables. Ahora analizaremos las matrices nilpotentes. Esto nos llevará a la noción de formas canónicas de Jordan, que nos va a proporcionar una forma, (entre otras), de calcular la exponencial de cualquier matriz. Aunque nosotros solo trabajaremos con matrices en el caso que el polinomio característico escinda (y que la multiplicidad geométrica de algunos sus valores propios no coincida con la algebraica, de lo contrario sería diagonalizable como ya vimos).

Definición 2.3.1. Se dice que una matriz $N \in \mathcal{M}_n(\mathbb{K})$ es *nilpotente* si existe $k \in \mathbb{Z}^+$ tal que $N^k = 0$. Se llama *índice* (también *orden*) de nilpotencia al mín $\{k \in \mathbb{Z}^+ : N^k = 0\}$.

La matriz

$$N = \begin{pmatrix} 0 & 2 & 1 & 6 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

es nilpotente con índice de nilpotencia 4, porque $N^2 \neq 0$, $N^3 \neq 0$ y $N^4 = 0$.
También son nilpotentes las matrices de la forma:

$$J_r = \begin{pmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \dots & 0 \end{pmatrix} \in \mathcal{M}_r(\mathbb{K}) \quad \text{para todo } r \in \mathbb{Z}^+ \tag{2.3.1}$$

con índice de nilpotencia r, pues

$$J_r = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}, \qquad J_r^2 = \begin{pmatrix} 0 & 0 & 1 & \dots & 0 \\ & 0 & 0 & \ddots & \vdots \\ & & \ddots & \ddots & 1 \\ & & & 0 & 0 \\ & & & & 0 \end{pmatrix},$$

$$\cdots J_r^{r-1} = \begin{pmatrix} 0 & 0 & \dots & 1 \\ & 0 & \dots & 0 \\ & & \ddots & \vdots \\ & & & 0 \end{pmatrix}, \qquad J_r^r = 0.$$

La matriz J_r de la forma (2.3.1), se dice que es un *bloque de Jordan nilpotente.*

Comentario 2.3.2. A continuación nos proponemos probar que cualquier matriz nilpotente de tamaño n es semejante a una matriz J, que es diagonal en bloques, donde los bloques $J_{n_1}, \cdots, J_{n_r}$, son de la forma (2.3.1) que se llaman como dijimos anteriormente bloques de Jordan nilpotentes, siendo además $n_1 + \cdots + n_r = n$.
Para ello debemos primeramente enunciar algunos resultados.

Proposición 2.3.3. *Sean*

$$A = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix} \in \mathcal{M}_n(\mathbb{K}), \quad J_i = \begin{pmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \dots & 0 \end{pmatrix} \in \mathcal{M}_{n_i}(\mathbb{K}), \quad \textit{para todo} \;\; i = 1, \cdots, r$$

con $n_1 + \cdots + n_r = n$ *siendo* $n_1 \geq n_2 \geq \cdots \geq n_r$, *entonces*

a) *A es nilpotente con indice de nilpotencia n_1.*

b) *La cantidad de bloques J_i contenidos en A es $r = n - \operatorname{rango}(A)$.*

Dem. a) Primeramente observemos que

$$J_i^{n_1} = J_i^{n_1 - n_i} J_i^{n_i} \text{ y como } \; J_i \in \mathcal{M}_{n_i}(\mathbb{K}) \text{ entonces por (2.3.1) } \; J_i^{n_i} = 0$$
$$\text{por lo cual podemos afirmar que } \; J_i^{n_1} = 0, \quad \text{para todo } \; i = 1, \cdots, r.$$

Entonces

$$A^{n_1} = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix}^{n_1} = \begin{pmatrix} J_1^{n_1} & & \\ & \ddots & \\ & & J_r^{n_1} \end{pmatrix} = 0 \quad \text{es decir} \quad A^{n_1} = 0$$

Entonces A es nilpotente de orden menor o igual a n_1, pero como $J_1^{n_1-1} \neq 0$ entonces $A^{n_1-1} \neq 0$. Concluimos que $A^{n_1-1} \neq 0$ y $A^{n_1} = 0$, es decir la matriz A es nilpotente con índice de nilpotencia n_1.

b) Primero observemos que el rango de la matriz A es la suma de los rangos de las submatrices bloques, y como para cada i, la matriz $J_i \in \mathcal{M}_{n_i}(\mathbb{K})$ tiene rango $n_i - 1$, tenemos que

$$\operatorname{rango}(A) = \sum_{i=1}^{r} \operatorname{rango}(J_i) = \sum_{i=1}^{r} (n_i - 1) = \sum_{i=1}^{r} n_i - r = n - r \;\; \text{entonces} \;\; r = n - \operatorname{rango}(A). \quad \square$$

Proposición 2.3.4. *Si $A \in \mathcal{M}_n(\mathbb{K})$ es una matriz nilpotente de índice k entonces*

$$\{0\} \subset \ker(A) \subset \ker(A^2) \subset \cdots \subset \ker(A^k) = \mathbb{K}^n$$

y las inclusiones son estrictas.

Dem. Siendo k el índice de nilpotencia de A, se tiene que $A^k = 0$, de donde

$$\ker(A^k) = \left\{ v \in \mathbb{K}^n : A^k v = 0 \right\} = \{ v \in \mathbb{K}^n : 0v = 0 \} = \mathbb{K}^n.$$

Es claro que valen las inclusiones porque si

$$v \in \ker(A^i), \text{ entonces } A^i v = 0 \Rightarrow A^{i+1} v = A(A^i v) = A0 = 0 \Rightarrow v \in \ker(A^{i+1}).$$

Vamos a probar que las inclusiones son estrictas. En primer lugar probemos que

$$\text{si para algún j, } \ker(A^j) = \ker(A^{j+1}) \text{ entonces } \ker(A^{j+m}) = \ker(A^j) \text{ para todo } m \geq 1 \tag{2.3.2}$$

Lo probaremos por inducción sobre m.
Para $m = 1$ se cumple por hipotesis.
Supongamos ahora que para algún $m \geq 1$ se cumple que $\ker(A^{j+m}) = \ker(A^j)$ y queremos probar que $\ker(A^{j+m+1}) = \ker(A^j)$ también se cumple.
Ya sabemos que $\ker(A^j) = \ker(A^{j+m}) \subset \ker(A^{j+m+1})$.
Así que solo resta probar que $\ker(A^{j+m+1}) \subset \ker(A^j)$. Sea $v \in \ker(A^{j+m+1})$,

$$0 = A^{j+m+1}v = A^{j+1}(A^m v) \Rightarrow A^m v \in \ker(A^{j+1}) = \ker(A^j)$$

Entonces como $A^m v \in \ker(A^j)$, tenemos

$$0 = A^j \left(A^m v\right) = A^{j+m}v \Rightarrow v \in \ker\left(A^{j+m}\right) \underset{\text{por hipótesis de inducción}}{=} \ker\left(A^j\right)$$

por lo cual $v \in \ker\left(A^j\right)$ y concluye la demostración.
Pero teniendo en cuenta que el índice de nilpotencia de A es k, entonces $\ker\left(A^{k-1}\right) \neq \mathbb{K}^n = \ker\left(A^k\right)$ y por la afirmación (2.3.2) concluimos que las inclusiones son estrictas. □

Lema 2.3.5. *Sean una matriz $A \in \mathcal{M}_n(\mathbb{K})$, $q \in \mathbb{N}$ y $q \geq 2$. Supongamos que $\{v_1, \cdots, v_r\} \subset \ker(A^q)$ es un conjunto linealmente independiente tal que $\ker(A^{q-1}) \cap [v_1, \cdots, v_r] = \{0\}$. Entonces $\{Av_1, \cdots, Av_r\} \subset \ker(A^{q-1})$, es linealmente independiente y $\ker(A^{q-2}) \cap [Av_1, \cdots, Av_r] = \{0\}$.*

Dem. Como $v_i \in \ker(A^q)$, se cumple entonces

$$0 = A^q v_i = A^{q-1}\left(Av_i\right) \Longrightarrow Av_i \in \ker\left(A^{q-1}\right)$$

entonces probamos que si $v_i \in \ker(A^q)$ entonces $Av_i \in \ker\left(A^{q-1}\right)$ para todo $i = 1, \cdots, r$, o sea que $\{Av_1, \cdots, Av_r\} \subset \ker(A^{q-1})$.
Supongamos ahora que $v = \alpha_1 Av_1 + \cdots + \alpha_r Av_r \in \ker\left(A^{q-2}\right)$. Entonces

$$\begin{aligned} 0 = A^{q-2}v &= A^{q-2}\left(\alpha_1 Av_1 + \cdots + \alpha_r Av_r\right) = A^{q-2}\left(A\left(\alpha_1 v_1 + \cdots + \alpha_r v_r\right)\right) \\ &= A^{q-1}\left(\alpha_1 v_1 + \cdots + \alpha_r v_r\right) \text{ por lo cual } \alpha_1 v_1 + \cdots + \alpha_r v_r \in \ker(A^{q-1}). \end{aligned}$$

Ahora como $\ker(A^{q-1}) \cap [v_1, \cdots, v_r] = \{0\}$ se deduce que $\alpha_1 v_1 + \cdots + \alpha_r v_r = 0$ y como $v_1, \cdots, v_r$ son linealmente independientes, se deduce que $\alpha_1 = \cdots = \alpha_r = 0$, luego $v = 0$ por lo cual concluimos que $\ker(A^{q-2}) \cap [Av_1, \cdots, Av_r] = \{0\}$.
Usando el mismo argumento de la demostración anterior, tenemos que si $\alpha_1 Av_1 + \cdots + \alpha_r Av_r = 0$,

$$\begin{aligned} 0 =& A^{q-2}0 = A^{q-2}\left(\alpha_1 Av_1 + \cdots + \alpha_r Av_r\right) \\ &= A^{q-1}\left(\alpha_1 v_1 + \cdots + \alpha_r v_r\right) \text{ luego } \alpha_1 v_1 + \cdots + \alpha_r v_r \in \ker(A^{q-1}) \\ &\text{pero como } \ker(A^{q-1}) \cap [v_1, \cdots, v_r] = \{0\} \text{ tenemos } \alpha_1 v_1 + \cdots + \alpha_r v_r = 0 \\ &\text{pero siendo } v_1, \cdots, v_r \text{ linealmente independientes, se deduce que } \alpha_1 = \cdots = \alpha_r = 0 \end{aligned}$$

Esto concluye que $\{Av_1, \cdots, Av_r\}$ es linealmente independiente y queda probado el lema □

Lema 2.3.6. *Sean V un $\mathbb{K}$ espacio vectorial de dimensión finita, $W \subset V$ un subespacio y $\{v_1, \cdots, v_r\} \subset V$ un conjunto linealmente independiente tal que $W \cap [v_1, \cdots, v_r] = \{0\}$.*
Entonces existen vectores $u_1, \cdots, u_h$ pertenecientes a V tal que $\{v_1, \cdots, v_r, u_1, \cdots, u_h\}$ es un conjunto linealmente independiente, y $V = W \oplus [v_1, \cdots, v_r, u_1, \cdots, u_h]$

Dem. Sea $\{w_1, \cdots, w_l\}$ una base de W, por lo tanto $\dim W = l$.
Ahora como $\mathcal{C} = \{w_1, \cdots, w_l, v_1, \cdots, v_r\}$ es un conjunto linealmente independiente, lo podemos completar hasta formar una base de V, es decir, existen vectores $u_1, \cdots, u_h \in V$ tal que
$\mathcal{D} = \{w_1, \cdots, w_l, v_1, \cdots, v_r, u_1, \cdots, u_h\}$ es una base de V por lo cual $\dim V = l+r+h = \dim W+r+h$.
Sea ahora $v \in W \cap [v_1, \cdots, v_r, u_1, \cdots, u_h]$, entonces como $v \in W$

existen escalares $\alpha_1, \cdots, \alpha_l \in \mathbb{K}$ tal que $v = \alpha_1 w_1 + \cdots + \alpha_l w_l$ (1)

y por otro lado como $v \in [v_1, \cdots, v_r, u_1, \cdots, u_h]$

existen escalares $\beta_1, \cdots, \beta_r, \gamma_1, \cdots, \gamma_h \in \mathbb{K}$ tal que $v = \beta_1 v_1 + \cdots + \beta_r v_r + \gamma_1 u_1 + \cdots + \gamma_h u_h$ (2)

De (1) y (2), se deduce que

$$\begin{aligned} &\alpha_1 w_1 + \cdots + \alpha_l w_l = \beta_1 v_1 + \cdots + \beta_r v_r + \gamma_1 u_1 + \cdots + \gamma_h u_h, \text{ entonces} \\ &\alpha_1 w_1 + \cdots + \alpha_l w_l - \beta_1 v_1 - \cdots - \beta_r v_r - \gamma_1 u_1 - \cdots - \gamma_h u_h = 0 \end{aligned}$$

es decir que tenemos una combinación lineal de la base $\mathcal{D}$ igualada al vector 0, por lo cual todos los escalares son ceros, por lo que se concluye que $v = 0$, entonces que $W \cap [v_1, \cdots, v_r, u_1, \cdots, u_h] = \{0\}$ y termina la demostración. □

Observación 2.3.7. Notar que un conjunto $\mathcal{B} = \{v_1, \cdots, v_n\}$ es una base de $\mathbb{K}^n$ si y solo si la matriz $P = \big(v_1|v_2|\cdots|v_n\big) \in \mathcal{M}_n(\mathbb{K})$ es invertible.

Teorema 2.3.8. *Sean $A \in \mathcal{M}_n(\mathbb{K})$, $v_1, \cdots, v_n \in \mathbb{K}^n$ y consideremos $P = \big(v_1|\cdots|v_n\big) \in \mathcal{M}_n(\mathbb{K})$. Entonces*

1. *Si $\mathcal{B} = \{v_1, \cdots, v_n\}$ es una base de $\mathbb{K}^n$ y definimos $X = (x_{ij}) \in \mathcal{M}_n(\mathbb{K})$ por*

$$Av_i = x_{1i}v_1 + \cdots + x_{ni}v_n \text{ para todo } i = 1, \cdots, n. \tag{2.3.3}$$

   entonces P es invertible y $A = PXP^{-1}$ (luego A y X son matrices semejantes).

2. *Si P es invertible y $X = (x_{ij}) \in \mathcal{M}_n(\mathbb{K})$ está definida por $X = P^{-1}AP$, entonces los elementos x_{ij} verifican (2.3.3).*

Dem. Escribimos $X = \big(x_1|\cdots|x_n\big)$, donde las columnas $x_i = \begin{pmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{pmatrix}$, $i = 1, \cdots, n$ están definidas según (2.3.3).
Luego

$$\begin{aligned} Px_i &= \begin{pmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & & \vdots \\ v_{n1} & \cdots & v_{nn} \end{pmatrix} \begin{pmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{pmatrix} = \begin{pmatrix} x_{1i}v_{11}+ & \cdots & +x_{ni}v_{1n} \\ \vdots & & \vdots \\ x_{1i}v_{n1}+ & \cdots & +x_{ni}v_{nn} \end{pmatrix} \\ &= x_{1i} \begin{pmatrix} v_{11} \\ \vdots \\ v_{n1} \end{pmatrix} + \cdots + x_{ni} \begin{pmatrix} v_{1n} \\ \vdots \\ v_{nn} \end{pmatrix} = x_{1i}v_1 + \cdots + x_{ni}v_n = Av_i. \end{aligned}$$

Entonces tenemos que $Px_i = Av_i$ para todo $i = 1, \cdots, n$ por lo cual $PX = AP$, y como $\mathcal{B}$ es base de $\mathbb{K}^n$, entonces por observación 2.3.7, P es invertible y entonces vale $A = PXP^{-1}$.

Veamos ahora la segunda parte. Primero veamos por la observación 2.3.7 que por ser $P = \big(v_1|\cdots|v_n\big)$ invertible, entonces $\mathcal{B} = \{v_1, \cdots, v_n\}$ es una base de $\mathbb{K}^n$. Vamos a probar que las columnas de la matriz X son las coordenadas de los vectores Av_i en la base $\mathcal{B}$: $x_i = coord_{\mathcal{B}}(Av_i)$ para todo $i = 1, \cdots, n$. Como

$$A = PXP^{-1} \Longrightarrow AP = PX \Longrightarrow Av_i = Px_i \text{ para todo } i = 1, \cdots, n$$

Pero análogamente a lo hecho en la parte anterior, tenemos que $Px_i = x_{1i}v_1 + \cdots + x_{ni}v_n$, por lo cual se concluye que $Av_i = x_{1i}v_1 + \cdots + x_{ni}v_n$, para todo $i = 1, \cdots, n$, que es lo que queriamos probar. $\square$

Observación 2.3.9. Con las notaciones de arriba, notar que dadas A y P en $\mathcal{M}_n$ con P invertible, entonces existe un única matriz $X \in \mathcal{M}_n$ tal que $A = PXP^{-1}$.
A la matriz X la escribiremos $[A]_{\mathcal{B}}$ y queda caracterizada por verificar

$$A = P[A]_{\mathcal{B}}P^{-1}, \quad P = \big(v_1|\cdots|v_n\big), \quad \mathcal{B} = \{v_1, \cdots, v_n\}.$$

Observación 2.3.10. Supongamos que $A \in \mathcal{M}_n(\mathbb{K})$ verifica

$$A \sim J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix} \text{ siendo } J_i = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix} \in \mathcal{M}_{n_i}(\mathbb{K}),\ n_1 + \cdots + n_r = n.$$

Entonces existe $P \in \mathcal{M}_n(\mathbb{K})$ invertible tal que $A = PJP^{-1}$.

Escribimos $P = \left(v_1^1|\cdots|v_{n_1}^1|\cdots|v_1^r|\cdots|v_{n_r}^r \right)$ siendo por lo tanto el conjunto $\mathcal{B} = \left\{v_1^1, \cdots, v_{n_1}^1, \cdots, v_1^r, \cdots, v_{n_r}^r\right\}$ una base de $\mathbb{K}^n$. Las columnas de J, se obtienen calculando Av y escribiéndolo como combinación lineal de la base $\mathcal{B}$, para todo $v \in \mathcal{B}$, según el teorema 2.3.8.
Considerando J_1 y los vectores corespondientes $v_1^1, \cdots, v_{n_1}^1$, vemos que

$$J_1 = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix} \Longleftrightarrow \begin{cases} Av_1^1 = 0 \\ Av_i^1 = v_{i-1}^1, \quad i = 2, \cdots, n_1 \end{cases}$$

Y lo mismo sucede con los otros bloques:
si consideramos

$$\mathcal{B}_i = \left\{v_1^i, \cdots, v_{n_i}^i\right\} \quad \text{es} \quad \begin{cases} Av_1^i = 0 \\ Av_{k+1}^i = v_k^i, \quad k = 1, \cdots, n_i - 1 \end{cases}$$

de donde

$$\begin{aligned} & v_{n_i-1}^i = Av_{n_i}^i \\ & v_{n_i-2}^i = Av_{n_i-1}^i = A^2 v_{n_i}^i \\ & \vdots \\ & v_2^i = Av_3^i = A^{n_i-2}v_{n_i}^i \\ & v_1^i = Av_2^i = A^{n_i-1}v_{n_i}^i \\ & 0 = Av_1^i = A^{n_i}v_{n_i}^i \end{aligned}$$

entonces la base que genera el bloque i-ésimo tiene que ser de la forma
$\mathcal{B}_i = \left\{A^{n_i-1}w_i, A^{n_i-2}w_i, \cdots, Aw_i, w_i\right\}$ con $A^{n_i}w_i = 0$, (siendo $w_i = v^i_{n_i}$).

Recíprocamente, si existe $\mathcal{B}$ una base de $\mathbb{K}^n$ de la forma $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$, siendo $\mathcal{B}_1, \cdots, \mathcal{B}_r$ disjuntos dos a dos, en que cada $\mathcal{B}_i$ es de la forma $\mathcal{B}_i = \left\{A^{n_i-1}w_i, A^{n_i-2}w_i, \cdots, Aw_i, w_i\right\}$ con $A^{n_i}w_i = 0$, entonces es $A = PJP^{-1}$ en que P es la matriz correspondiente a la base $\mathcal{B}$ y

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix} \text{ siendo } J_i = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix} \in \mathcal{M}_{n_i}(\mathbb{K}),\ i = 1, \cdots, r.$$

Ahora nos encontramos en condiciones de justificar la afirmación hecha en el comentario 2.3.2.

Teorema 2.3.11. *Sea A una matriz $n \times n$. Entonces A es nilpotente si y solo si A es semejante a una matriz en bloques de Jordan nilpotentes, es decir de la forma*

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix}, \quad \textit{con } J_i = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix} \in \mathcal{M}_{n_i}(\mathbb{K}),\ i = 1, \cdots, r, \quad n_1 + \cdots + n_r = n \tag{2.3.4}$$

Dem. ($\Leftarrow$) Es inmediato, porque $J_i^{n_i} = 0$ y si ponemos $p = \max\{n_1, \cdots, n_r\}$ entonces $J_i^p = 0$ para todo $i = 1, \cdots, r$. Luego tenemos que $J^p = \begin{pmatrix} J_1^p & & \\ & \ddots & \\ & & J_r^p \end{pmatrix} = 0$ y como $J_k^{p-1} \neq 0$ para algún k entre 1 y r, J es nilpotente de índice p y como A es semejante a J entonces A es nilpotente de índice p.

($\Rightarrow$) Por la observación 2.3.10, lo que tenemos que probar es que existe una base $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$, $\mathcal{B}_1, \cdots, \mathcal{B}_r$ disjuntos dos a dos, de $\mathbb{K}^n$ tal que cada $\mathcal{B}_i$ es de la forma $\mathcal{B}_i = \left\{A^{n_i-1}w_i, A^{n_i-2}w_i, \cdots, Aw_i, w_i\right\}$ para cierto $w_i \in \mathbb{K}^n$ tal que $A^{n_i}w_i = 0$.

Por simplicidad de notación, haremos la prueba para el caso en que A es nilpotente de orden 3, pero la prueba vale en general. Suponemos entonces que vale $A^3 = 0$ y $A^2 \neq 0$.

Por la proposición 2.3.4 tenemos que $\{0\} \subsetneqq \ker(A) \subsetneqq \ker(A^2) \subsetneqq \ker(A^3) = \mathbb{K}^n$, entonces existe un conjunto linealmente independiente $\{v_1, \cdots, v_m\}$ tal que $\ker(A^2) \cap [v_1, \cdots, v_m] = \{0\}$ y
$n = \dim\ker(A^2) + m$, entonces por Lema 2.3.5 $\{Av_1, \cdots, Av_m\} \subset \ker(A^2)$, es linealmente independiente, $\ker(A) \cap [Av_1, \cdots, Av_m] = \{0\}$ y ahora por el lema 2.3.6 existen $u_1, \cdots, u_q \in \ker(A^2)$ tal que $\{Av_1, \cdots, Av_m, u_1, \cdots, u_q\}$ es un subconjunto de $\ker(A^2)$ linealmente independiente,
$\ker(A) \cap [Av_1, \cdots, Av_m, u_1, \cdots, u_q] = \{0\}$ y $\dim\ker(A^2) = \dim\ker(A) + m + q$, entonces aplicando nuevamente el lema 2.3.5, tenemos que $\left\{A^2v_1, \cdots, A^2v_m, Au_1, \cdots, Au_q\right\} \subset \ker(A)$ es linealmente independiente.
Sean $w_1, \cdots, w_r \in \ker(A)$ tal que $\left\{A^2v_1, \cdots, A^2v_m, Au_1, \cdots, Au_q, w_1, \cdots, w_r\right\}$ es base de $\ker(A)$. Deducimos que $\mathcal{C}_1 = \left\{A^2v_1, \cdots, A^2v_m, Au_1, \cdots, Au_q, w_1, \cdots, w_r, Av_1, \cdots, Av_m, u_1, \cdots, u_q\right\}$ es linealmente independiente y está contenido en el $\ker(A^2)$.
Ahora como la $\dim\ker(A^2) = \dim\ker(A) + m + q = m + q + r + m + q = \text{card } \mathcal{C}_1$ entonces $\mathcal{C}_1$ es

una base de $\ker(A^2)$, entonces la unión disjunta $\mathcal{C}_1 \cup \{v_1, \cdots, v_m\} = \mathcal{C}_2$ es linealmente independiente y $\mathcal{C}_2 \subset \ker(A^3) = \mathbb{K}^n$.
Como $n = \dim\ker(A^2) + m = \operatorname{card} \mathcal{C}_2$ entonces $\mathcal{C}_2$ es una base de $\mathbb{K}^n$ y si la reordenamos nos queda

$$\mathcal{B} = \left\{A^2v_1, Av_1, v_1, \cdots, A^2v_m, Av_m, v_m, Au_1, u_1, \cdots, Au_q, u_q, w_1, \cdots, w_r\right\}, \text{base de } \mathbb{K}^n \tag{2.3.5}$$

$$\begin{aligned} w_i \in \ker(A) &\Rightarrow Aw_i = 0 \\ u_i \in \ker(A^2) &\Rightarrow A^2u_i = 0 \\ \mathbb{K}^n = \ker(A^3) &\Rightarrow A^3v_i = 0 \end{aligned}$$

Entonces si escribimos

$$P = \left(A^2v_1|Av_1|v_1|\cdots|A^2v_m|Av_m|v_m|Au_1|u_1|\cdots|Au_q|u_q|w_1|\cdots|w_r \right)$$

se cumple $A = PJP^{-1}$, siendo J de la forma (2.3.4).
Luego la matriz nilpotente A con índice de nilpotencia 3, que es el caso estudiado, es semejante a una matriz de la forma

$$J = \begin{pmatrix} 0 & 1 & 0 & & & & & \\ 0 & 0 & 1 & & & & & \\ 0 & 0 & 0 & & & & & \\ & & & \ddots & & & & \\ & & & & 0 & 1 & & \\ & & & & 0 & 0 & & \\ & & & & & & \ddots & \\ & & & & & & & 0 \end{pmatrix}$$

□

Observación 2.3.12.

1. El conjunto $\mathcal{B}$ definido en (2.3.5) se llama *base de Jordan para A*.
2. Notar que cada bloque de Jordan J_i de tamaño n_i corresponde a un *ciclo* de la base, es decir a un conjunto de la forma:
$$\mathcal{B}_i = \left\{A^{n_i-1}w_i, A^{n_i-2}w_i, \cdots, Aw_i, w_i\right\} \text{ con } A^{n_i}w_i = 0.$$
3. Según la proposición 2.3.3, la cantidad de bloques J_i contenidos en J es $r = n - \operatorname{rango}(J)$, y como $\operatorname{rango}(A) = \operatorname{rango}(J)$, porque son matrices semejantes, se deduce que la cantidad de bloques en J es $r = n - \operatorname{rango}(A)$.
4. La demostración del teorema nos da un algoritmo para encontrar una base de Jordan, en las condiciones del teorema.

Ejemplo 2.3.13. Hallar una forma de Jordan y una base de Jordan para la matriz $N \in \mathcal{M}_6(\mathbb{R})$, siendo

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

Operando tenemos que $N^2 \neq 0$ y $N^3 = 0$, entonces N es nilpotente de orden 3. Vamos a aplicar la demostración del teorema 2.3.11, para hallar la base de Jordan. Observemos que como el rango de N es 3, entonces la cantidad de bloques es $r = 6 - \text{rango}(N) = 6 - 3 = 3$. Es decir la matriz J tiene 3 bloques.
Efectuando operaciones tenemos que:

$$\begin{aligned} &\ker(N^3) = \mathbb{R}^6 = [e_1, e_2, e_3, e_4, e_5, e_6], && \dim\ker(N^3) = 6 \\ &\ker(N^2) = [e_2, e_3, e_4, e_5, e_6] && \dim\ker(N^2) = 5 \\ &\ker(N) = [e_3, e_4, e_6] && \dim\ker(N) = 3 \end{aligned}$$

Entonces siguiendo la demostración del teorema 2.3.11, tenemos que

$$\{0\} \subsetneqq \ker(N) \subsetneqq \ker(N^2) \subsetneqq \ker(N^3) = \mathbb{R}^6$$

y extendemos la base $\mathcal{B}_2 = \{e_2, e_3, e_4, e_5, e_6\}$ del $\ker(N^2)$ con el vector e_1, por ejemplo, para obtener una base $\mathcal{B}_3 = \{e_1, e_2, e_3, e_4, e_5, e_6\}$ del $\ker(N^3) = \mathbb{R}^6$, donde se cumple que $\ker(N^2) \cap [e_1] = \{0\}$. Ahora $\{Ne_1\} = \{(0, 1, -1, 0, -1, 1)\} \subset \ker(N^2)$, y

$$\begin{array}{ccccccc} \{0\} & \subsetneqq & \ker(N) & \subsetneqq & \ker(N^2) & \subsetneqq & \ker(N^3) = \mathbb{R}^6 \\ & & & & Ne_1 & & e_1 \end{array}$$

Ahora consideramos $\mathcal{B}_1$ base del $\ker(N)$ y ampliamos el conjunto $\mathcal{B}_1 \cup \{Ne_1\} \subset \ker(N^2)$ hasta formar una base de $\ker(N^2)$, por ejemplo con e_2 tal que $\ker(N) \cap [Ne_1, e_2] = \{0\}$.
Ahora multiplicando Ne_1 y e_2 por N, obtenemos el conjunto linealmente independiente
$\left\{N^2e_1, Ne_2\right\} = \{(0, 0, -1, 0, 0, 1), (0, 0, -1, 1, 0, 0)\} \subset \ker(N)$

$$\begin{array}{ccccccc} \{0\} & \subsetneqq & \ker(N) & \subsetneqq & \ker(N^2) & \subsetneqq & \ker(N^3) = \mathbb{R}^6 \\ & & N^2e_1 & & Ne_1 & & e_1 \\ & & Ne_2 & & e_2 & & \end{array}$$

Por último ampliamos el conjunto $\left\{N^2e_1, Ne_2\right\}$ hasta formar una base del $\ker(N)$, por ejemplo: $\left\{N^2e_1, Ne_2, e_4\right\} = \{(0, 0, -1, 0, 0, 1), (0, 0, -1, 1, 0, 0), (0, 0, 0, 1, 0, 0)\}$, que nos queda

$$\begin{array}{ccccccc} \{0\} & \subsetneqq & \ker(N) & \subsetneqq & \ker(N^2) & \subsetneqq & \ker(N^3) = \mathbb{R}^6 \\ & & N^2e_1 & & Ne_1 & & e_1 \\ & & Ne_2 & & e_2 & & \\ & & e_4 & & & & \end{array}$$

Por lo que finalmente nos queda una base de Jordan para N

$$\begin{aligned} \mathcal{B} &= \left\{N^2e_1, Ne_1, e_1, Ne_2, e_2, e_4\right\} \\ &= \{(0, 0, -1, 0, 0, 1), (0, 1, -1, 0, -1, 1), (1, 0, 0, 0, 0, 0), (0, 0, -1, 1, 0, 0), (0, 1, 0, 0, 0, 0), (0, 0, 0, 1, 0, 0)\} \end{aligned}$$

Y la matriz en bloques de Jordan, equivalente a la matriz nilpotente N, es

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \left(\begin{array}{ccc|cc|c} 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & & \\ 0 & 0 & 0 & & & \\ \hline & & & 0 & 1 & \\ & & & 0 & 0 & \\ \hline & & & & & 0 \end{array}\right)$$

Entonces si

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{se cumple que } N = PJP^{-1}.$$

□

2.4. Subespacios invariantes

Definición 2.4.1. Si $A \in \mathcal{M}_n$ y $W \subset \mathbb{K}^n$ es un subespacio, decimos que W es *A-invariante* si $Aw \in W$ para todo $w \in W$.

Los subespacios $\ker(A) = \{x \in \mathbb{K}^n : Ax = 0\}$ y $\mathrm{Im}\,(A) = \{Ax : x \in \mathbb{K}^n\}$, son subespacios A-invariantes de $\mathbb{K}^n$.

Proposición 2.4.2. *Sea $A \in \mathcal{M}_n$ y supongamos que vale $\mathbb{K}^n = W_1 \oplus \cdots \oplus W_r$, W_i subespacios A-invariantes de $\mathbb{K}^n$ para todo $i = 1, \cdots, r$. Sea $\mathcal{B}_i$ base de W_i para todo i y consideremos $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ base de $\mathbb{K}^n$.*
Entonces se cumple $[A]_B$ es una matriz diagonal en bloques:

$$[A]_{\mathcal{B}} = \begin{pmatrix} \boxed{Q_1} & & \\ & \ddots & \\ & & \boxed{Q_r} \end{pmatrix}, \quad Q_i \in \mathcal{M}_{n_i}, \quad n_i = \#\ \mathcal{B}_i = \dim W_i,\ i = 1, \cdots, r.$$

Dem. Consideremos $\mathcal{B} = \left\{v_1^1, \cdots, v_{n_1}^1, \cdots, v_1^r, \cdots, v_{n_r}^r\right\}$ base de $\mathbb{K}^n$ donde $\mathcal{B}_i = \left\{v_1^i, \cdots, v_{n_i}^i\right\}$ es una base de W_i para todo i.
Teniendo en cuenta que W_1 es A-invariante, entonces $Av_i^1 \in W_1$, por lo tanto Av_i^1 es combinación lineal de la base $\mathcal{B}_1$, para todo $i = 1, \cdots, n_1$. Y escribiendo

$$\begin{aligned} Av_1^1 &= a_{11}v_1^1 + \cdots + a_{n_1 1}v_{n_1}^1 + 0v_1^2 + \cdots + 0v_{n_r}^r \\ Av_2^1 &= a_{12}v_1^1 + \cdots + a_{n_1 2}v_{n_1}^1 + 0v_1^2 + \cdots + 0v_{n_r}^r \\ &\vdots \\ Av_{n_1}^1 &= a_{1n_1}v_1^1 + \cdots + a_{n_1 n_1}v_{n_1}^1 + 0v_1^2 + \cdots + 0v_{n_r}^r \end{aligned}$$

Entonces por definición de $[A]_{\mathcal{B}}$ tenemos que las columnas están definidas por

$$\begin{aligned} &\text{primera columna de } [Av]_{\mathcal{B}} = coord[Av_1^1]_{\mathcal{B}} = (a_{11}, a_{21}, \cdots, a_{n_1 1}, 0, \cdots, 0) \\ &\text{segunda columna de } [Av]_{\mathcal{B}} = coord[Av_2^1]_{\mathcal{B}} = (a_{12}, a_{22}, \cdots, a_{n_1 2}, 0, \cdots, 0) \\ &\vdots \\ &n_1\text{-ésima columna de } [Av]_{\mathcal{B}} = coord[Av_{n_1}^1]_{\mathcal{B}} = (a_{1n_1}, a_{2n_1}, \cdots, a_{n_1 n_1}, 0, \cdots, 0) \end{aligned}$$

Entonces nos va quedando la matriz $[A]_{\mathcal{B}}$, con un primer bloque Q_1 de la siguiente forma:

$$[A]_{\mathcal{B}} = \begin{pmatrix} a_{11} & \cdots & a_{1n_1} & \cdots \\ \vdots & & \vdots & \\ a_{n_11} & \dots & a_{n_1n_1} & \dots \\ 0 & \dots & 0 & \\ \vdots & & \vdots & \\ 0 & \dots & 0 & \dots \end{pmatrix}, \qquad Q_1 = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n_1} \\ a_{21} & a_{22} & \dots & a_{2n_1} \\ \vdots & \vdots & \dots & \vdots \\ a_{n_11} & a_{n_12} & \dots & a_{n_1n_1} \end{pmatrix},$$

De forma análoga, teniendo en cuenta ahora que W_2 es A-invariante, para obtener la columna n_1+1, hacemos:

$$Av_{n+1} = Av_1^2 = 0v_1^1 + \cdots + 0v_{n_1}^1 + a_{n_1+1n_1+1}v_1^2 + \cdots + a_{n_1+n_2n_1+1}v_{n_2}^2 + 0v_1^3 + \cdots + 0v_{n_r}^r$$

y entonces

$$n_1 + 1\text{-ésima columna de } [Av]_{\mathcal{B}} = coord[Av_1^2]_{\mathcal{B}} = (0, \cdots, 0, a_{n_1+1n_1+1}, \cdots, a_{n_1+n_2n_1+1}, 0, \cdots, 0)$$

$$n_1 + 1\text{-ésima columna de } [Av]_{\mathcal{B}} : \qquad \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_{n_1+1n_1+1} \\ \vdots \\ a_{n_1+n_2n_1+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

y de esta forma construimos el bloque Q_2. Repitiendo este procedimiento llegamos a que la matriz $[Av]_{\mathcal{B}}$ tiene la forma que afirma el teorema. $\square$

Definición 2.4.3. Sea un polinomio $p(t) = a_m t^m + \cdots + a_0 \in \mathbb{K}[t]$. Entonces para una matriz $A \in \mathcal{M}_n$, $p(A)$ es la matriz definida por

$$p(A) := a_m A^m + \cdots + a_0 I \in \mathcal{M}_n.$$

La prueba del siguiente resultado es directa.

Proposición 2.4.4. *Sean p y q polinomios en el anillo $\mathbb{K}[x]$, $A \in \mathcal{M}_n$ y $\beta \in \mathbb{K}$, entonces se cumple:*

1. $(\beta p + q)(A) = \beta p(A) + q(A)$.
2. $(pq)(A) = p(A)q(A)$.

Corolario 2.4.5. $p(A)q(A) = q(A)p(A)$. $\square$

Proposición 2.4.6. *Sean $A \in \mathcal{M}_n(\mathbb{K})$, p_1, p_2 polinomios de $\mathbb{K}[t]$ tal que $mcd(p_1, p_2) = 1$ y $p_1(A)p_2(A) = 0$. Luego*

$$\mathbb{K}^n = \ker(p_1(A)) \oplus \ker(p_2(A)).$$

Dem. Por el teorema de Bezout para polinomios y de que $mcd(p_1, p_2) = 1$, tenemos que existen polinomios q_1 y q_2 en $\mathbb{K}[t]$ tal que $p_1q_1 + p_2q_2 = 1$. Luego, de acuerdo a la proposición 2.4.4, tenemos que $p_1(A)q_1(A) + p_2(A)q_2(A) = I$.
Sea $v \in \mathbb{K}^n$ y definimos $x = p_2(A)q_2(A)v$ e $y = p_1(A)q_1(A)v$. Entonces

$$v = Iv = (p_1(A)q_1(A) + p_2(A)q_2(A))v = p_1(A)q_1(A)v + p_2(A)q_2(A)v = y + x.$$

Si hacemos

$$p_1(A)x = p_1(A)p_2(A)q_2(A)v \underset{\text{por hip.}}{=} 0q_2(A)v = 0 \quad \text{entonces } x \in \ker(p_1(A))$$
$$p_2(A)y = p_2(A)p_1(A)q_1(A)v \underset{2.4.5}{=} p_1(A)p_2(A)q_1(A)v \underset{\text{por hip.}}{=} 0q_1(A)v = 0 \text{ entonces } y \in \ker(p_2(A)).$$

Hemos probado de que para cada $v \in \mathbb{K}^n$ existen $x \in \ker(p_1(A)$ e $y \in \ker(p_2(A))$ tal que $v = x + y$.
Falta ver que la descomposición es única. Supongamos que $v = x+y = x'+y'$, donde $x, x' \in \ker(p_1(A))$ e $y, y' \in \ker(p_2(A))$. Luego $x - x' = y' - y \in \ker(p_1(A)) \cap \ker(p_2(A))$. Entonces

$$\begin{aligned} x - x' &= I(x - x') = (p_1(A)q_1(A) + p_2(A)_2q(A))(x - x') \\ &= p_1(A)q_1(A)(x - x') + p_2(A)q_2(A)(x - x') \\ &= q_1(A)p_1(A)(x - x') + q_2(A)p_2(A)(x - x') \\ &= q_1(A)0 + q_2(A)0 = 0 \end{aligned}$$

Luego concluimos que $x' = x$. Y por lo tanto $y' = y$. $\square$

Usando inducción completa, se prueba la generalización de la proposición anterior:

Corolario 2.4.7. *Sean $A \in \mathcal{M}_n(\mathbb{K})$ y $p = p_1 \times \cdots \times p_r \in \mathbb{K}[t]$, tales que $mcd(p_i, p_j) = 1$ siendo $i \neq j$ y $p(A) = 0$. Entonces*

$$\mathbb{K}^n = \ker(p_1(A)) \oplus \cdots \oplus \ker(p_r(A)).$$

Observación 2.4.8.

1. Notar que si $p \in \mathbb{K}[t]$ y $A \in \mathcal{M}_n(\mathbb{K})$ entonces $Ap(A) = p(A)A$. Esto se deduce del Cor. 2.4.5, tomando $q(t) = t$.

2. Si $v \in \ker p(A)$, entonces $p(A)(Av) = p(A)Av \underset{\text{Obs.1}}{=} Ap(A)v = A(p(A)v) = A0 = 0$. Luego $\ker p(A)$ es un subespacio A-invariante de $\mathbb{K}^n$.
Sea ahora $w \in \operatorname{Im}(p(A))$, es decir existe $x \in \mathbb{K}^n$ tal que $w = p(A)x$, entonces

$$Aw = A(p(A)x) = Ap(A)x = p(A)Ax = p(A)(Ax) \in \operatorname{Im}(p(A))$$

por lo cual concluimos que $\operatorname{Im}(p(A))$ es un subespacio A-invariante de $\mathbb{K}^n$.

3. En las hipótesis del corolario anterior, si $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ es base de $\mathbb{K}^n$, siendo B_i base de $\ker p_i(A)$, para todo i, entonces aplicando la proposición 2.4.2 tenemos

$$[A]_{\mathcal{B}} = \begin{pmatrix} Q_1 & & \\ & \ddots & \\ & & Q_r \end{pmatrix} \text{ y } A = P[A]_{\mathcal{B}}P^{-1}, \ P \text{ matriz asociada a la base } \mathcal{B}.$$

.

Teorema de Cayley-Hamilton

Consideremos la matriz $A = \left(\begin{smallmatrix} 5 & -1 \\ 4 & 3 \end{smallmatrix}\right)$ cuyo polinomio característico es $\mathcal{X}_A(t) = t^2 - 8t + 19$. Calculemos $\mathcal{X}_A(A)$,

$$\mathcal{X}_A(A) = A^2 - 8A + 19I = \begin{pmatrix} 21 & -8 \\ 32 & 5 \end{pmatrix} - 8 \begin{pmatrix} 5 & -1 \\ 4 & 3 \end{pmatrix} + \begin{pmatrix} 19 & 0 \\ 0 & 19 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Esto no es una casualidad. El teorema de Cayley-Hamilton dice que toda matriz $A \in \mathcal{M}_n(\mathbb{K})$ es raíz de su polinomio característico, es decir que se cumple $\mathcal{X}_A(A) = 0$.
Hay varias demostraciones de este teorema, nosotros daremos una que no requiere transformaciones lineales, para la cual necesitamos algunos conceptos previos.

Definición 2.4.9. Dada una cantidad finita de matrices $P_0, \cdots, P_m \in \mathcal{M}_n(\mathbb{K})$, entonces una expresión de la forma

$$P_0 + P_1 t + P_2 t^2 + \cdots + P_m t^m$$

se llama *polinomio con coeficientes matriciales* en la variable t. Es decir, un polinomio con coeficientes matriciales es un polinomio cuyos coeficientes están en $\mathcal{M}_n(\mathbb{K})$.

Observación 2.4.10.

1. Sea $\mathcal{M}_n(\mathbb{K}[t])$ el conjunto de todas las matrices $n \times n$ con coeficientes polinomiales, es decir las matrices cuyas entradas son polinomios de $\mathbb{K}[t]$. El conjunto $\mathcal{M}_n(\mathbb{K}[t])$ con las operaciones usuales de suma y producto de matrices, tiene estructura de anillo.

2. Sea $\mathcal{M}_n(\mathbb{K})[t]$ el conjunto de todos los polinomios con coeficientes matriciales. Este conjunto tiene estructura de anillo, con la operación usual de suma de polinomios y definiendo el producto mediante lo siguiente:
si $p = \sum_i A_i t^i$ y $q = \sum_j B_j t^j$, entonces su producto es

$$pq \stackrel{def}{=} \sum_{i,j} A_i B_j t^{i+j} = A_0 B_0 + (A_0 B_1 + A_1 B_0)t + \cdots .$$

Es decir el coeficiente de t^k, es $C_k = \displaystyle\sum_{i+j=k} A_i B_j$. Podemos escribir entonces que si

$$p = \sum_{i=0}^{r} A_i t^i, \quad q = \sum_{j=0}^{s} B_j t^j \quad \text{es} \quad pq = \sum_{k=0}^{r+s} \Big(\sum_{i+j=k} A_i B_j \Big) t^k.$$

Observar que al no ser conmutativo el producto de matrices, entonces el anillo $\mathcal{M}_n(\mathbb{K})[t]$ no es conmutativo. Sin embargo notar que si identificamos $t = tI$ (I matriz identidad), entonces t es un elemento central en $\mathcal{M}_n(\mathbb{K})[t]$, es decir conmuta con todo otro elemento de $\mathcal{M}_n(\mathbb{K})[t]$.

3. Observar que todo elemento de $\mathcal{M}_n(\mathbb{K}[t])$ se puede escribir de la forma

$$\begin{aligned}
\begin{pmatrix} \sum_{i=1}^k b_i^{(11)} t^i & \cdots & \sum_{i=1}^k b_i^{(1n)} t^i \\ \vdots & & \vdots \\ \sum_{i=1}^k b_i^{(n1)} t^i & \cdots & \sum_{i=1}^k b_i^{(nn)} t^i \end{pmatrix} &= \sum_{i=1}^{k} \begin{pmatrix} b_i^{(11)} & \cdots & b_i^{(1n)} \\ \vdots & & \vdots \\ b_i^{(n1)} & \cdots & b_i^{(nn)} \end{pmatrix} \begin{pmatrix} t^i & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & t^i \end{pmatrix} \\
&= \sum_{i=1}^{k} \begin{pmatrix} b_i^{(11)} & \cdots & b_i^{(1n)} \\ \vdots & & \vdots \\ b_i^{(n1)} & \cdots & b_i^{(nn)} \end{pmatrix} \begin{pmatrix} t & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & t \end{pmatrix}^i .
\end{aligned}$$

Notar que el elemento $\begin{pmatrix} t & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & t \end{pmatrix}$ es central en $\mathcal{M}_n(\mathbb{K}[t])$, en forma análoga a t en $\mathcal{M}_n(\mathbb{K})[t]$. Luego podemos definir en forma natural una función $\phi : \mathcal{M}_n(\mathbb{K}[t]) \to \mathcal{M}_n(\mathbb{K})[t]$ mediante

$$\begin{pmatrix} \sum_{i=1}^k b_i^{(11)} t^i & \cdots & \sum_{i=1}^k b_i^{(1n)} t^i \\ \vdots & & \vdots \\ \sum_{i=1}^k b_i^{(n1)} t^i & \cdots & \sum_{i=1}^k b_i^{(nn)} t^i \end{pmatrix} \xrightarrow{\phi} \sum_{i=1}^{k} \begin{pmatrix} b_i^{(11)} & \cdots & b_i^{(1n)} \\ \vdots & & \vdots \\ b_i^{(n1)} & \cdots & b_i^{(nn)} \end{pmatrix} t^i.$$

Es fácil de probar que ϕ preserva las operaciones de suma y producto, por lo cual es un homomorfismo de anillos. También es inmediato observar que es biyectiva, luego $\phi : \mathcal{M}_n(\mathbb{K}[t]) \to \mathcal{M}_n(\mathbb{K})[t]$ es un isomorfismo de anillos.

Definición 2.4.11. Sean $p := P_0 + P_1 t + P_2 t^2 + \cdots + P_m t^m$ un polinomio con coeficientes matriciales y consideremos una matriz $A \in \mathcal{M}_n(\mathbb{K})$. Entonces definimos la *evaluación* de p en A mediante

$$p(A) := P_0 + P_1 A + P_2 A^2 + \cdots + P_m A^m \in \mathcal{M}_n(\mathbb{K}).$$

La evaluación de polinomios con coeficientes matriciales no cumple algunas de las propiedades de la evaluación de polinomios con coeficientes numéricos. Mostremos con un ejemplo que no necesariamente se cumple $pq(A) = p(A)q(A)$.

Ejemplo 2.4.12. Consideremos los polinomios matriciales $p = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) t$ y $q = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, y la matriz $A = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$. Entonces

$$pq = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} t \quad \Rightarrow \quad pq(A) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix};$$
$$p(A)q(A) = \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right] \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Entonces $pq(A) \neq p(A)q(A)$.

Veamos a continuación una condición suficiente para que se cumpla la igualdad: $pq(A) = p(A)q(A)$.

Lema 2.4.13. *Sean p y q polinomios con coeficientes matriciales,*

$$p = \sum_{i=0}^{r} P_i t^i \in \mathcal{M}_n(\mathbb{K})[t], \qquad q = \sum_{j=0}^{s} Q_j t^j \in \mathcal{M}_n(\mathbb{K})[t].$$

Sea ahora $A \in \mathcal{M}_n(\mathbb{K})$ una matriz que conmuta con todos los coeficientes del polinomio q: $Q_j A = A Q_j$ para todo $j = 0, \cdots, s$. Entonces

$$pq(A) = p(A)q(A).$$

Dem. Del hecho que $Q_j A = A Q_j$ para todos $j = 0, \cdots, s$ se deduce que $Q_j A^l = A^l Q_j$ para todo $j = 0, \cdots, s$ y $l \in \mathbb{N}$, usando esto último y operando se prueba el lema. $\square$

Teorema 2.4.14. ***Cayley-Hamilton***
Sean $A \in \mathcal{M}_n(\mathbb{K})$ y $\mathcal{X}_A(t) \in \mathbb{K}[t]$ su polinomio característico. Entonces vale $\mathcal{X}_A(A) = 0$.

Dem. Sabemos que para toda matriz $B \in \mathcal{M}_n(\mathbb{K})$ vale $adj(B)B = det(B)I$, siendo $adj(B) \in \mathcal{M}_n(\mathbb{K})$ la *adjunta clásica* de B (la matriz de los cofactores).
Consideremos la matriz polinomial $A - tI \in \mathcal{M}_n(\mathbb{K}[t])$. Aplicando el resultado anterior obtenemos

$$adj(A - tI)(A - tI) = det(A - tI)I = \mathcal{X}_A(t)I. \tag{2.4.1}$$

Esta es una igualdad en $\mathcal{M}_n(\mathbb{K}[t])$. Vía el isomorfismo visto en la observación 2.4.10 parte 3) pensamos lo anterior como una igualdad en $\mathcal{M}_n(\mathbb{K})[t]$.
Sean $p = adj(A - tI) \in \mathcal{M}_n(\mathbb{K})[t]$ y $q = A - tI \in \mathcal{M}_n(\mathbb{K})[t]$ y sabemos que $pq = \mathcal{X}_A(t)I$. Por otro lado, como los coeficientes del polinomio $q = A - tI$ conmutan con A, entonces el lema anterior implica que vale

$$pq(A) = (adj(A - tI)(A - tI))(A) \underset{\text{Lema}}{=} p(A)q(A) = p(A)(A - AI) = p(A)0 = 0. \tag{2.4.2}$$

Luego de las igualdades (2.4.1) y (2.4.2) deducimos que $pq(A) = \mathcal{X}_A(A)I = 0$ en $\mathcal{M}_n(\mathbb{K})$ y por lo tanto $\mathcal{X}_A(A) = 0$. □

2.5. Forma de Jordan. Caso general.

Teorema 2.5.1. (**Forma canónica de Jordan**).
Si $A \in \mathcal{M}_n(\mathbb{K})$, entonces las siguientes afirmaciones son equivalentes

a) El polinomio característico $\mathcal{X}_A(t)$ se escinde.

b) La matriz A es semejante a una matriz J, donde J es una matriz en bloques

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_h \end{pmatrix} \qquad \textit{donde cada bloque } J_i \textit{ es una matriz cuadrada de la forma}$$

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{pmatrix} \quad \textit{siendo } \lambda_i \textit{ valor propio de la matriz } A,\ i = 1, \cdots, h.$$

Dem. $b) \Rightarrow a)$. La matriz $J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_h \end{pmatrix}$ es triangular superior y como $A \sim J$, entonces $\mathcal{X}_A(t) = \mathcal{X}_J(t)$ y como $\mathcal{X}_J(t)$ se escinde por ser la matriz J triangular superior, queda probada esta parte.

Probemos ahora que $a) \Rightarrow b)$

Si $A \in \mathcal{M}_n(\mathbb{K})$ y el polinomio característico $\mathcal{X}_A(t)$ se escinde, entonces

$$\mathcal{X}_A(t) = (-1)^n (t-\lambda_1)^{n_1} \cdots (t-\lambda_k)^{n_k}, \text{ con } \lambda_i \neq \lambda_j \text{ si } i \neq j, \; n_1 + \cdots + n_r = n.$$

Por simplicidad de notación haremos la prueba para $k = 2$, pero la demostración vale en general. Así que estamos asumiendo $\mathcal{X}_A(t) = (-1)^n (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2}$, con $\lambda_1 \neq \lambda_2$, $n_1 + n_2 = n$.
Luego por el teorema de Cayley-Hamilton tenemos que vale $\mathcal{X}_A(A) = 0$, y si $p(t) = (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2}$ se deduce $p(A) = 0$, y como $p = p_1 p_2$ siendo $p_1 = (t-\lambda_1)^{n_1}$, $p_2 = (t-\lambda_2)^{n_2}$ y $mcd(p_1, p_2) = 1$ tenemos por la proposición 2.4.6 que vale $\mathbb{K}^n = \ker(A-\lambda_1 I)^{n_1} \oplus \ker(A-\lambda_2 I)^{n_2}$, con $\ker(A-\lambda_i I)^{n_i}$ subespacio A-invariante de $\mathbb{K}^n$, $i = 1, 2$.
Luego si $\mathcal{B}_i$ es base de $\ker(A-\lambda_i I)^{n_i}$, $i = 1, 2$ es $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ base de $\mathbb{K}^n$, entonces por la observación 2.4.8 parte 3

$$A = P[A]_{\mathcal{B}} P^{-1}, \text{ con } [A]_{\mathcal{B}} = \left(\begin{array}{c|c} X & 0 \\ \hline 0 & Y \end{array}\right), \quad X \in \mathcal{M}_p, \quad Y \in \mathcal{M}_q, \; p + q = n,$$

donde $p = \#\mathcal{B}_1 = dim \ker(A-\lambda_1 I)^{n_1}$ y $q = \#\mathcal{B}_2 = dim \ker(A-\lambda_2 I)^{n_2}$.
Ahora

$$A - \lambda I_n = P[A]_{\mathcal{B}} P^{-1} - \lambda I_n = P\left([A]_{\mathcal{B}} - \lambda I_n\right) P^{-1} \text{ con } [A]_{\mathcal{B}} - \lambda I_n = \left(\begin{array}{c|c} X - \lambda I_p & 0 \\ \hline 0 & Y - \lambda I_q \end{array}\right)$$

cualquiera sea $\lambda \in \mathbb{K}$. Entonces $(A - \lambda I_n)^r = P\left([A]_{\mathcal{B}} - \lambda I_n\right)^r P^{-1}$,

$$\left([A]_{\mathcal{B}} - \lambda I_n\right)^r = \left(\begin{array}{c|c} (X - \lambda I_p)^r & 0 \\ \hline 0 & (Y - \lambda I_q)^r \end{array}\right) \text{ para todo } r \in \mathbb{N}.$$

Si $v \in \mathcal{B}_1 \subset \ker(A-\lambda_1 I)^{n_1}$ entonces $(A-\lambda_1 I)^{n_1} v = 0$.
Luego, si escribimos $(X - \lambda_1 I_p)^{n_1} = (z_{ij})$ y $\mathcal{B}_1 = \{v_1, \cdots, v_p\}$, entonces por teorema 2.3.8 y observación 2.3.9, los z_{ij} quedan determinados por

$$(A-\lambda_1 I)^{n_1} v_i = z_{1i} v_1 + \cdots + z_{pi} v_p, \; i = 1, \cdots, p$$

Pero como $v_i \in \ker(A-\lambda_1 I)^{n_1}$ entonces $(A-\lambda_1 I)^{n_1} v_i = 0$ para todo i. De estas dos últimas condiciones y el hecho de que $\mathcal{B}_1 = \{v_1, \cdots, v_p\}$ es linealmente independiente se deduce que los $z_{ij} = 0$ para todo i, j, de donde concluimos que $(X - \lambda_1 I_p)^{n_1} = 0$, es decir que $X - \lambda_1 I_p$ es nilpotente con índice de nilpotencia menor o igual a n_1. Entonces por teorema 2.3.11, la matriz $X - \lambda_1 I_p \in \mathcal{M}_p$ es semejante a una matriz en bloques de Jordan nilpotentes, por lo cual existe una base $\mathcal{C}_1$ de $\mathbb{K}^p$, tal que $X - \lambda_1 I_p = Q_1 [X - \lambda_1 I_p]_{\mathcal{C}_1} Q_1^{-1}$ siendo Q_1 la matriz asociada a la base $\mathcal{C}_1$, por lo cual tenemos

$$[X - \lambda_1 I_p]_{\mathcal{C}_1} = \begin{pmatrix} \mathbb{J}_1^1 & & \\ & \ddots & \\ & & \mathbb{J}_{r_1}^1 \end{pmatrix} \text{ siendo } \mathbb{J}_i^1 = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}, \; i = 1, \cdots, r_1.$$

Entonces

$$X = X - \lambda_1 I_p + \lambda_1 I_p = Q_1 [X - \lambda_1 I_p]_{\mathcal{C}_1} Q_1^{-1} + Q_1 (\lambda_1 I_p) Q_1^{-1}$$

$$= Q_1 \left([X - \lambda_1 I_p]_{\mathcal{C}_1} + \lambda_1 I_p\right) Q_1^{-1} = Q_1 \begin{pmatrix} \mathbb{J}_1^1 + \lambda_1 I_{p_1} & & \\ & \ddots & \\ & & \mathbb{J}_{r_1}^1 + \lambda_1 I_{p_{r_1}} \end{pmatrix} Q_1^{-1}$$

y si llamamos $\mathbb{X} = [X - \lambda_1 I_p]_{\mathcal{C}_1} + \lambda_1 I_p$, tenemos que

$$X = Q_1 \mathbb{X} Q_1^{-1}, \quad \mathbb{X} = \begin{pmatrix} J_1^1 & & \\ & \ddots & \\ & & J_{r_1}^1 \end{pmatrix} \in \mathcal{M}_p \text{ donde } J_i^1 = \begin{pmatrix} \lambda_1 & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_1 & 1 \\ & & & \lambda_1 \end{pmatrix}, \quad i = 1, \cdots, r_1. \tag{2.5.1}$$

Análogamente se prueba que $(Y - \lambda_2 I_q)^{n_2} = 0$, y con el mismo razonamiento anterior probamos que vale

$$Y = Q_2 \mathbb{Y} Q_2^{-1}, \quad \mathbb{Y} = \begin{pmatrix} J_1^2 & & \\ & \ddots & \\ & & J_{r_2}^2 \end{pmatrix} \in \mathcal{M}_q \text{ donde } J_i^2 = \begin{pmatrix} \lambda_2 & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_2 & 1 \\ & & & \lambda_2 \end{pmatrix}, \quad i = 1, \cdots, r_2.$$

Luego

$$A = P[A]_{\mathcal{B}} P^{-1}, \text{ con } [A]_{\mathcal{B}} = \left(\begin{array}{c|c} X & 0 \\ \hline 0 & Y \end{array}\right), \quad X \in \mathcal{M}_p, \quad Y \in \mathcal{M}_q, \ p + q = n,$$

donde $X = Q_1 \mathbb{X} Q_1^{-1}$ e $Y = Q_2 \mathbb{Y} Q_2^{-1}$, entonces

$$A = P[A]_{\mathcal{B}} P^{-1} = P \left(\begin{array}{c|c} X & 0 \\ \hline 0 & Y \end{array}\right) p^{-1} = P \left(\begin{array}{c|c} Q_1 \mathbb{X} Q_1^{-1} & 0 \\ \hline 0 & Q_2 \mathbb{Y} Q_2^{-1} \end{array}\right) p^{-1}$$

$$= P \left(\begin{array}{c|c} Q_1 & 0 \\ \hline 0 & Q_2 \end{array}\right) \left(\begin{array}{c|c} \mathbb{X} & 0 \\ \hline 0 & \mathbb{Y} \end{array}\right) \left(\begin{array}{c|c} Q_1^{-1} & 0 \\ \hline 0 & Q_2^{-1} \end{array}\right) P^{-1}$$

$$= R \left(\begin{array}{c|c} \mathbb{X} & 0 \\ \hline 0 & \mathbb{Y} \end{array}\right) R^{-1}, \quad \text{siendo } R = P \left(\begin{array}{c|c} Q_1 & 0 \\ \hline 0 & Q_2 \end{array}\right) \in \mathcal{M}_n.$$

Así que hemos probado que $A \sim \left(\begin{array}{c|c} \mathbb{X} & 0 \\ \hline 0 & \mathbb{Y} \end{array}\right)$ en que $\mathbb{X}$ es una matriz en bloques de Jordan correspondiente al valor propio λ_1, e $\mathbb{Y}$ es una matriz en bloques de Jordan correspondiente al valor propio λ_2. □

Observación 2.5.2. Si una matriz A verifica $A \sim \left(\begin{array}{c|c} \mathbb{X} & 0 \\ \hline 0 & \mathbb{Y} \end{array}\right)$ con $\mathbb{X} \in \mathcal{M}_p$ e $\mathbb{Y} \in \mathcal{M}_q$ en bloques de Jordan como antes, entonces $\mathcal{X}_A(t) = (-1)^n (t - \lambda_1)^p (t - \lambda_2)^q$, luego necesariamente debe ser $n_1 = p$ y $n_2 = q$.
Es decir que si $\mathcal{X}_A(t) = (-1)^n (t - \lambda_1)^{n_1} (t - \lambda_2)^{n_2}$, $\lambda_1 \neq \lambda_2$, entonces $\dim \ker(A - \lambda_i I)^{n_i} = n_i$, $i = 1, 2$. Deducimos entonces, que el tamaño del bloque $\mathbb{X}$ es la multiplicidad algebraica de λ_1. Análogamente el tamaño del bloque $\mathbb{Y}$ es la multiplicidad algebraica de λ_2.

Observación 2.5.3. Notar que de acuerdo a la fórmula (2.5.1) la cantidad de bloques de Jordan contenidos en $\mathbb{X}$ coincide con los de la matriz nilpotente $\mathbb{X} - \lambda_1 I_p$. Luego la proposición 2.3.3 implica

que la cantidad de bloques de Jordan contenidos en $\mathbb{X}$ es $r_1 = n_1 - \text{rango}(\mathbb{X} - \lambda_1 I_p)$.
Ahora, por otro lado,

$$\begin{aligned} MG(\lambda_1) &= n - \text{rango}(A - \lambda_1 I) = p + q - \text{rango}(A - \lambda_1 I) = n_1 + n_2 - \text{rango}(A - \lambda_1 I) \\ &= n_1 + n_2 - (\text{rango}(\mathbb{X} - \lambda_1 I_p) + \text{rango}(\mathbb{Y} - \lambda_1 I_q)) = n_1 + n_2 - (\text{rango}(\mathbb{X} - \lambda_1 I_p) + n_2) \\ &= n_1 - \text{rango}(\mathbb{X} - \lambda_1 I_p). \end{aligned}$$

El hecho de que $\text{rango}(\mathbb{Y} - \lambda_1 I_q) = n_2$ surge de que la matriz $\mathbb{Y} - \lambda_1 I_q \in \mathcal{M}_q$ es invertible porque es triangular superior y los elementos de la diagonal son $\lambda_2 - \lambda_1 \neq 0$, entonces $\text{rango}(\mathbb{Y} - \lambda_1 I_q) = q = n_2$. En conclusión: la cantidad de bloques de Jordan contenidos en $\mathbb{X}$ es la multiplicidad geométrica de λ_1. Ídem para $\mathbb{Y}$.

En las condiciones del teorema 2.5.1 y teniendo en cuenta las dos observaciones anteriores concluimos que:

- La suma de los tamaños de los bloques de Jordan con λ en la diagonal es igual a la multiplicidad algebraica del valor propio λ.
- La cantidad de bloques de Jordan asociados al valor propio λ es igual a la multiplicidad geométrica de λ.

Observación 2.5.4. Lo anterior muestra que si $\mathcal{X}_A$ se escinde, entonces $A \sim J$, siendo J como en el teo. 2.5.1. A una matriz J de ese tipo J la llamaremos *matriz de Jordan*. Veremos como hallar la matriz de semejanza en forma más explícita. Con las notaciones anteriores:

$$\begin{aligned} &\mathbb{K}^n = \ker(A - \lambda_1 I)^{n_1} \oplus \ker(A - \lambda_2 I)^{n_2} \text{ y } \mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2, \text{ entonces es} \\ &P = \big(\underbrace{v_1|\cdots|v_p}_{\mathcal{B}_1}|\underbrace{u_1|\cdots|u_q}_{\mathcal{B}_2}\big), \text{ donde } \mathcal{B}_i \text{ es base de } \ker(A - \lambda_i I)^{n_i}, \text{ para } i = 1, 2. \end{aligned}$$

$$A = R \left(\begin{array}{c|c} \mathbb{X} & 0 \\ \hline 0 & \mathbb{Y} \end{array}\right) R^{-1}, \quad R = P \left(\begin{array}{c|c} Q_1 & 0 \\ \hline 0 & Q_2 \end{array}\right)$$

$$\begin{aligned} &X = Q_1 \mathbb{X} Q_1^{-1}, \quad \mathbb{X} \in \mathcal{M}_p, \text{ con } p = n_1, \text{ matriz de Jordan con bloques correspondientes a } \lambda_1. \\ &Y = Q_2 \mathbb{Y} Q_2^{-1}, \quad \mathbb{Y} \in \mathcal{M}_q, \text{ con } q = n_2, \text{ matriz de Jordan con bloques correspondientes a } \lambda_2. \end{aligned}$$

Luego Q_1 y Q_2 corresponden a bases de $\mathbb{K}^p$ y $\mathbb{K}^q$ respectivamente, formadas por ciclos correspondientes a $X - \lambda_1 I_p$ y $Y - \lambda_2 I_q$.
Si escribimos $P = (\, P_1 \,|\, P_2 \,)$, $\quad P_1 = (\, v_1|\cdots|v_p \,) \in \mathcal{M}_{n\times p}$, $\quad P_2 = (\, u_1|\cdots|u_q \,) \in \mathcal{M}_{n\times q}$, entonces

$$R = P \left(\begin{array}{c|c} Q_1 & \\ \hline & Q_2 \end{array}\right) = (\, P_1 \,|\, P_2 \,) \left(\begin{array}{c|c} Q_1 & \\ \hline & Q_2 \end{array}\right) = (\, P_1 Q_1 \,|\, P_2 Q_2 \,), \ P_1 Q_1 \in \mathcal{M}_{n\times p}, \ P_2 Q_2 \in \mathcal{M}_{n\times q}.$$

Las columnas de Q_1 son uniones de ciclos. Consideremos uno de estos ciclos:

$$\left\{(X - \lambda_1 I_p)^{l-1} w, \cdots, (X - \lambda_1 I_p) w, w\right\}, \quad w \in \mathbb{K}^p, \quad (X - \lambda_1 I_p)^l w = 0$$

Al multiplicar por P_1, a estas columnas en Q_1 les corresponden las siguientes en $P_1 Q_1$

$$\left\{P_1 (X - \lambda_1 I_p)^{l-1} w, \cdots, P_1 (X - \lambda_1 I_p) w, P_1 w\right\}, \quad P_1 w \in \mathbb{K}^n. \tag{2.5.2}$$

Sea $\widehat{w} = P_1 w \in \mathbb{K}^n$

Vamos a probar que el conjunto (2.5.2) se puede escribir

$$\left\{(A - \lambda_1 I_n)^{l-1}\widehat{w}, \cdots, (A - \lambda_1 I_n)\widehat{w}, \widehat{w}\right\}, \quad \widehat{w} \in \mathbb{K}^n, \quad (A - \lambda_1 I_n)^l \widehat{w} = 0$$

es decir es un ciclo de la matriz A correspondiente al valor propio λ_1.

Af.1 Se cumple $AP_1 = P_1 X$ en $\mathcal{M}_{n\times p}$ y $AP_2 = P_2 Y$ en $\mathcal{M}_{n\times q}$.
Sabemos que

$$A = P \left(\begin{array}{c|c} X & 0 \\ \hline 0 & Y \end{array}\right) P^{-1} \Longleftrightarrow AP = P \left(\begin{array}{c|c} X & 0 \\ \hline 0 & Y \end{array}\right) \Longleftrightarrow A\,(\,P_1|P_2\,) = (\,P_1|P_2\,) \left(\begin{array}{c|c} X & 0 \\ \hline 0 & Y \end{array}\right)$$
$$\Longleftrightarrow (\,AP_1|AP_2\,) = (\,P_1X|P_2Y\,) \Longleftrightarrow AP_1 = P_1X \text{ y } AP_2 = P_2Y. \quad \square$$

Af.2 Para todo $r \in \mathbb{N}$ vale $A^r P_1 = P_1 X^r$ y que $A^r P_2 = P_2 Y^r$.
Probaremos el primero por inducción completa sobre r, el otro es ánalogo.
Para $r = 1$ es la Af.1 y supongamos ahora que vale para $r = h$, entonces

$$A^{h+1}P_1 = A(A^h P_1) \underset{\text{hip de ind}}{=} AP_1X^h \underset{\text{Af.1}}{=} P_1XX^h = P_1X^{h+1}. \quad \square$$

Af.3 Se cumple que

$$(A - \lambda_1 I_n)^r P_1 = P_1 (X - \lambda_1 I_p)^r, \quad (A - \lambda_2 I_n)^r P_2 = P_2 (Y - \lambda_2 I_q)^r, \quad \text{para todo } r \in \mathbb{N}$$

Veamos la primera, la otra se justifica en forma análoga.
De la Af.1 sabemos que $AP_1 = P_1X$ entonces $AP_1 - \lambda_1 P_1 = P_1 X - \lambda_1 P_1$ por lo cual
$(A - \lambda_1 I_n) P_1 = P_1 (X - \lambda_1 I_p)$.
El caso general se deduce de la prueba de la Af.2. $\square$

Af.4 Concluimos: Si $\widehat{w} = P_1 w \in \mathbb{K}^n$, entonces

$$\left\{P_1 (X - \lambda_1 I_p)^{l-1} w, \cdots, P_1 (X - \lambda_1 I_p) w, P_1 w\right\} = \left\{(A - \lambda_1 I_n)^{l-1} \widehat{w}, \cdots, (A - \lambda_1 I_n) \widehat{w}, \widehat{w}\right\}$$

con $(A - \lambda_1 I_n)^l \widehat{w} = 0$.

Observar que $P = (\,v_1|\cdots|v_p\,) \in \mathcal{M}_{n\times p}$ tiene rango p ($\mathcal{B}_1 = \{v_1, \cdots, v_p\}$ es base de $\ker(A - \lambda_1 I)^{n_1}$, $n_1 = p$). Luego la transformación lineal $T_1 : \mathbb{K}^p \to \ker(A - \lambda_1 I)^{n_1} \subset \mathbb{K}^n$ definida por $T_1(v) = P_1 v$ para todo $v \in \mathbb{K}^p$, es un isomorfismo. Lo que vimos en la Af.4 es que este isomorfismo lleva ciclos de X correspondientes a λ_1, que "viven" en $\mathbb{K}^p$, en ciclos de A correspondientes a λ_1 que "viven" en $\ker(A - \lambda_1 I)^{n_1} \subset \mathbb{K}^n$. Es claro que lo mismo sucede con los ciclos correspondientes a λ_2.

Luego la base de Jordan $\mathcal{B}_J = \{w_1, \cdots, w_n\}$ tal que $A = R \left(\begin{array}{c|c} \mathbb{X} & 0 \\ \hline 0 & \mathbb{Y} \end{array}\right) R^{-1}$, $R = (\,w_1|\cdots|w_n)$ se obtiene como unión de ciclos disjuntos, y cada uno de estos se obtiene como en el caso nilpotente, pero trabajando en $\ker(A - \lambda_1 I)^{n_1}$ en vez de en $\mathbb{K}^p$ y en $\ker(A - \lambda_2 I)^{n_2}$ en vez de $\mathbb{K}^q$.

Observación 2.5.5. 1. Si $\mathbb{K} = \mathbb{C}$ todo polinomio de grado positivo se escinde, entonces toda matriz cuadadrada compleja es semejante a una matriz de Jordan.

2. La matriz J está univocamente determinada por A salvo permutaciones de los bloques J_i. La matriz J se denomina la *forma canónica de Jordan* de A.

Ejemplo 2.5.6. Sea la matriz

$$A = \begin{pmatrix} 0 & -3 & 1 & 2 \\ -2 & 1 & -1 & 2 \\ -2 & 1 & -1 & 2 \\ -2 & -3 & 1 & 4 \end{pmatrix}.$$

Comencemos calculando el polinomio característico de A

$$\mathcal{X}_A(\lambda) = \det(A - \lambda I) = \lambda^2(\lambda - 2)^2.$$

Luego los valores propios de A son 0 y 2, donde además $MA(0) = MA(2) = 2$, por lo cual podemos afirmar que hay dos bloques de Jordan, cada uno de tamaño 2, es decir que J queda de la forma

$$J = \begin{pmatrix} \boxed{J_1} & \\ & \boxed{J_2} \end{pmatrix} \quad \text{o sea} \quad J = \left(\begin{array}{cc|cc} 0 & * & & \\ & 0 & & \\ \hline & & 2 & * \\ & & & 2 \end{array}\right)$$

en que cada $*$ puede ser 0 o 1. Como el $\text{rango}(A - 0\lambda) = \text{rango}(A) = 3$ entonces el $\dim\ker(A - 0I) = 4 - 3 = 1 = MG(0)$, por lo cual como $MA(0) = 2$ y $MG(0) = 1$ deducimos que la matriz A no es diagonalizable y hay un solo bloque de Jordan para el valor propio 0

$$J_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Por otro lado el $\text{rango}(A - 2I) = 2$ entonces $\dim\ker(A - 2I) = 4 - 2 = 2 = MG(2)$ por lo cual concluimos que hay 2 bloques de Jordan para el valor propio 2

$$J_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

y obtenemos la matriz J de Jordan

$$J = \left(\begin{array}{cc|c|c} 0 & 1 & & \\ 0 & 0 & & \\ \hline & & 2 & \\ \hline & & & 2 \end{array}\right) \quad \text{o sea} \quad J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Hallemos ahora la base de Jordan, que es de la forma

$$\mathcal{B} = \{Av, v, u_2, u_1\}, \; A^2 v = Au_1 = Au_2 = 0.$$

Comencemos por calcular $\ker(A)$ y $\ker(A^2)$ y obtenemos

- $\ker(A) = [(1, 1, 1, 1)]$
- $\ker(A^2) = [(1, 0, -1, 1), (0, 1, 2, 0)]$

Elegimos $v = (0, 1, 2, 0) \in \ker(A^2) \backslash \ker(A)$ y calculemos entonces Av, obteniendo $Av = (-1, -1, -1, -1)$
Ahora si calculamos $\ker(A - 2I)$ obtenemos

$$\ker(A - 2I) = [(1, 0, 0, 1), (-1, 1, 1, 0)]$$

llamemos entonces $u_1 = (-1, 1, 1, 0)$ y $u_2 = (1, 0, 0, 1)$.
Concluimos entonces

$$\mathcal{B} = \{(-1, -1, -1, -1), (0, 1, 2, 0), (1, 0, 0, 1), (-1, 1, 1, 0)\}$$

por lo cual se concluye que la matriz de transición nos queda

$$P = \begin{pmatrix} -1 & 0 & 1 & -1 \\ -1 & 1 & 0 & 1 \\ -1 & 2 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix}$$

Entonces se cumple $J = P^{-1}AP$ o $A = PJP^{-1}$.

Entonces la matriz de Jordan de $A = \begin{pmatrix} 0 & -3 & 1 & 2 \\ -2 & 1 & -1 & 2 \\ -2 & 1 & -1 & 2 \\ -2 & -3 & 1 & 4 \end{pmatrix}$ es la matriz
$J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$, siendo además $P = \begin{pmatrix} -1 & 0 & 1 & -1 \\ -1 & 1 & 0 & 1 \\ -1 & 2 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix}$ la matriz de transición.

Comentario 2.5.7. Para obtener la forma canónica de Jordan de una matriz, el polinomio característico debe escindir en el cuerpo $\mathbb{K}$, según el teorema 2.5.1.
Ahora, ¿cómo se procede cuando el polinomio característico no se escinde en $\mathbb{R}$?. La idea es considerar el cuerpo $\mathbb{C}$, donde $\mathbb{R} \subset \mathbb{C}$, es decir una extensión de cuerpo, que es algebraicamente cerrado y por lo tanto el polinomio característico se escinde, y luego hallar la matriz de Jordan correspondiente.
Pero la forma canónica de Jordan de una matriz real es a menudo una matriz compleja, pero no real. Sin embargo, como $\mathbb{R}$ y $\mathbb{C}$ están muy "cercanos", se puede obtener una representación matricial real de una matriz real que es muy parecida a una matriz de Jordan. Esto no lo abordamos en esta tesina.

Capítulo 3

Espacios de matrices

3.1. Espacios normados

Si tenemos vectores o matrices en un espacio vectorial, ¿qué podría signficar decir que algunos son "pequeños" y que otros son "grandes"? O podriamos estar interesado en saber "que tan grande" es una función. ¿Bajo que condiciones podriamos decir que dos vectores están "muy juntos" o "muy separados"?. Cuestiones de "tamaño" y "proximidad" ya las hemos estudiado en $\mathbb{R}$, en $\mathbb{R}^2$ y tal vez en $\mathbb{R}^n$ con la distancia euclidiana.
¿Pero que podemos decir sobre el "tamaño" de matrices?. ¿Y de vectores en espacios de dimensión infinita? ¿ Hay otras maneras de medir el "tamaño" de vectores reales además de la medida euclidiana?. Un camino para responder a las preguntas anteriores es el estudio de las *normas* de vectores y matrices, que es una generalización de la distancia Euclideana, pero no es sólo una generalización, es necesaria, por ejemplo, para poder formular en forma adecuada nociones como las de series de potencias de matrices.
Además en diferentes situaciones será más conveniente usar una u otra norma, por lo que es apropiado comenzar por el estudio de las propiedades comunes a todas las normas, en vez de restringirse a algunas normas particulares.

Previamente, comenzaremos definiendo el concepto de espacio métrico.

Definición 3.1.1. Sea E un conjunto cualquiera no vacío. Una *métrica* o *distancia* en E es una función

$$d : E \times E \to \mathbb{R}$$

que verifica las siguientes propiedades:

1. $d(x, y) \geq 0,$ para todo $x, y \in E$ (no negatividad).
2. $d(x, y) = 0$ si y sólo si $x = y$.
3. $d(x, y) = d(y, x)$ para todo $x, y \in E$ (simetría).
4. $d(x, y) \leq d(x, z) + d(z, y)$ para todo $x, y, z \in E$ (desigualdad triangular).

La expresión $d(x, y)$ la leemos *distancia* entre los puntos x e y. El par (E, d) formado por el conjunto E y una métrica d definida sobre E, se llama *espacio métrico.*

Ejemplo 3.1.2. 1. La *métrica euclidiana* en $\mathbb{R}^n$ está definida de la siguiente manera:
si $\mathbf{x} = (x_1, \cdots, x_n)$ e $\mathbf{y} = (y_1, \cdots, y_n)$ son dos puntos de $\mathbb{R}^n$, entonces

$$d(\mathbf{x},\mathbf{y}) = \sqrt{(x_1-y_1)^2+\cdots+(x_n-y_n)^2} = \left[\sum_{i=1}^{n}(x_i-y_i)^2\right]^{1/2}.$$

2. Todo conjunto no vacio E puede considerarse como un espacio métrico con la *métrica discreta* definida por

$$d(x,y) = \begin{cases} 0 & \text{si} \quad x=y \\ 1 & \text{si} \quad x\neq y \end{cases}$$

para todo $x, y \in E$.

Observación 3.1.3. A partir de la métrica podemos definir conceptos fundamentales, tales como bola abierta, bola cerrada, conjuntos abiertos y cerrados, conjuntos compactos, conjuntos conexos, límite, continuidad (ver, por ejemplo, [Ll]).

Definición 3.1.4. Un *espacio normado* es un par $(E, \|\cdot\|)$, formado por un espacio vectorial E sobre un cuerpo $\mathbb{K}$, donde $\mathbb{K}$ es el cuerpo de los números reales o los complejos y una función $\|\cdot\| : E \to \mathbb{R}$, llamada *norma*, con las siguientes propiedades:

i) $\|x\| \geq 0$, para todo $x \in E$ (no negatividad),

ii) $\|x\| = 0$ si y solo si $x = 0$,

iii) $\|\alpha x\| = |\alpha|\|x\|$, para todo $x \in E$ y $\alpha \in \mathbb{K}$ (homogeneidad),

iv) $\|x+y\| \leq \|x\| + \|y\|$, para todo $x, y \in E$ (desigualdad triangular).

Una función $\|\cdot\| : \mathbb{E} \to \mathbb{R}$ que verifique los axiomas i), iii) y iv) se llama *seminorma*. La seminorma de un vector distinto de cero puede ser cero.

Ejemplo 3.1.5. La *norma euclidiana* en $\mathbb{R}^n$ es la norma definida por

$$\|(x_1, \cdots, x_n)\| = \sqrt{\sum_{i=1}^{n} x_i^2}$$

.

Los cuatro axiomas anteriores expresan algunas de las propiedades mas comunes de la norma euclideana en $\mathbb{R}^n$, pero la norma euclideana posee propiedades adicionales que no pueden ser deducidas de estos 4 axiomas, como por ejemplo la ley o identidad del paralelogramo.

Observación 3.1.6. 1. Todo espacio normado es un espacio métrico, relativo a la métrica natural d definida por $d(x,y) := \|x-y\|$. Así todas las nociones de espacios métricos están definidas también para los espacios normados. En particular la métrica euclideana es la inducida por la norma euclideana.

2. El recíproco no es cierto, es decir, no todo espacio métrico es normado. Por ejemplo la métrica discreta en un espacio vectorial, salvo casos triviales no proviene de una norma.

3. La norma de un espacio normado E es una función continua de E en $\mathbb{R}$.

Ejemplos 3.1.7. Los siguientes son ejemplos de normas en espacio de matrices.

a) $\|A\| = \text{máx}_{i,j}\{|a_{ij}|\}$ para toda $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$

b) $\|A\| = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}$ para toda $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$

c) $\|A\| = \text{n máx}_{i,j}\{|a_{ij}|\}$ para toda $A = (a_{ij}) \in \mathcal{M}_{n}(\mathbb{K})$

d) $\|A\| = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|$ para toda $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$

Observación 3.1.8. De los ejemplos queda claro que existen diferentes funciones $\|\cdot\| : E \to \mathbb{R}$ que satisfacen los axiomas de una norma. Pero es útil disponer de normas diferentes, porque una norma puede ser más conveniente o más apropiada que otra para un fin determinado. Por ejemplo una norma puede ser la más natural de usar para el estudio de la convergencia, pero puede ser analítica y algebraicamente incómoda o poco práctica de usar. En las aplicaciones reales, la norma sobre la cual la teoria es más natural y la norma sobre la cual se calcula más facilmente, pueden no coincidir. Es importante entonces, saber que relación puede haber entre dos normas diferentes.

Definición 3.1.9. Dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$ en E se dicen *equivalentes* si existen a y b reales positivos tal que $a\|x\|_1 \leq \|x\|_2 \leq b\|x\|_1$, para todo $x \in E$.

La relación así definida es una relación de equivalencia.

Teorema 3.1.10. *En un espacio vectorial de dimensión finita, todas las normas son equivalentes.* □

La demostración se puede ver en [Re] o [CC, Teo. 1.2.1].

Observación 3.1.11. En consecuencia, en espacios normados de dimensión finita, el estudio de la convergencia no depende de la norma elegida.

3.2. Espacios de Banach

Definición 3.2.1. Sean $(E, \|\cdot\|)$ un espacio normado y $(x_n)_{n\in\mathbb{N}}$ una sucesión de elementos de E. La sucesión $(x_n)_{n\in\mathbb{N}}$ *converge* a $x \in E$ si para todo $\varepsilon > 0$ existe $p \in \mathbb{N}$ tal que $\|x_n - x\| < \varepsilon$ para todo $n \geq p$. En esta situación escribiremos $\lim x_n = x$ o $x_n \to x$.

Definición 3.2.2. La sucesión $(x_n)_{n\in\mathbb{N}}$ se dice *de Cauchy*, si

$$\text{para todo } \varepsilon > 0 \text{ existe } p \in \mathbb{N} \text{ tal que si } n, m \geq p, \text{entonces } \|x_n - x_m\| < \varepsilon.$$

Observación 3.2.3. 1. Toda sucesión convergente es de Cauchy, mientras que el recíproco no es cierto. Por ejemplo, en $\mathbb{Q}$, la sucesión (x_n) donde $x_n = \left(1 + \frac{1}{n}\right)^n$, es de Cauchy pero no es convergente en $\mathbb{Q}$.

Definición 3.2.4. Un espacio normado E se llama *espacio de Banach* si toda sucesión de Cauchy en E es convergente a un punto de E.

Observación 3.2.5. 1. Considerando el teorema 3.1.10 tenemos que si un espacio vectorial n-dimensional es de Banach según una norma también lo es sobre cualquier otra norma.

2. Todo espacio normado de dimensión finita es de Banach. Para una prueba ver [Mt, Teo. 5.5].

Definición 3.2.6. Se dice que la serie $\sum_{n\geq 0} x_n$ en el espacio normado E es *convergente* si la sucesión (S_n), donde $S_n = \sum_{i=0}^{n} x_i,\ n \in \mathbb{N}$, es convergente en E.

Definición 3.2.7. Decimos que la serie $\sum x_n$ en el espacio normado E es *absolutamente convergente* si la serie real $\sum \|x_n\|$ es convergente.

Proposición 3.2.8. *Si una serie $\sum a_n$ en un espacio de Banach E es absolutamente convergente, entonces $\sum a_n$ converge.*

Dem. Para probar que $\sum a_n$ es convergente usaremos la condición de Cauchy.
Sea $\varepsilon > 0$.
Como $\sum_{n\geq 0} \|x_n\|$ es convergente, entonces la condición de Cauchy para $\sum_{n\geq 0} \|x_n\|$ implica que existe $n_0 \in \mathbb{N}$ tal que para todo $p > n_0$ y $k \in \mathbb{N}$ se cumple $\sum_{j=p+1}^{p+k} \|x_j\| < \varepsilon$. Pero la desigualdad triangular implica $\left\| \sum_{j=p+1}^{p+k} x_j \right\| \leq \sum_{j=p+1}^{p+k} \|x_j\|$. Luego

$$\text{para todo } p \geq n_o \text{ y } k \in \mathbb{N} \text{ se cumple } \left\| \sum_{j=p+1}^{p+k} x_j \right\| < \varepsilon.$$

Entonces la serie $\sum x_n$ cumple la condición de Cauchy y como el espacio E es de Banach, entonces la serie $\sum x_n$ es convergente. □

Comentario 3.2.9. También se cumple que si E es un espacio normado en el cual toda serie absolutamente convergente es convergente entonces E es un espacio de Banach. Para una prueba ver, Teo.4.2 [Mt]

Las funciones en la cual fijaremos nuestra atención en las definiciones que daremos a continuación, son de la forma $f : I \subset \mathbb{R} \to E$, donde I es un intervalo abierto y E es un espacio normado, que son con las cuales trabajaremos más adelante.

Definición 3.2.10. Si $f : I \subset \mathbb{R} \to E$, donde t_0 punto de acumulación de I y $\ell \in E$, diremos que ℓ es el *límite* de f cuando t tiende a t_0, y se escribe $\lim_{t\to t_0} f(t) = \ell$, si

$$\text{para todo } \varepsilon > 0, \text{existe } \delta > 0 \text{ tal que si } 0 < |t - t_0| < \delta \text{ y } t \in I, \text{ implica que vale } \|f(t) - \ell\| < \varepsilon.$$

Son válidas algunas de las propiedades usuales de los límites reales, por ejemplo, unicidad y

$$\text{Si } \lim_{t\to t_0} f(t) = \ell \text{ y } \lim_{t\to t_0} g(t) = \ell' \text{ entonces } \lim_{t\to t_0} (f \pm g)(t) = \ell \pm \ell'$$

y las demostraciones son análogas a la de los reales.

Definición 3.2.11. La función $f : I \subset \mathbb{R} \to E$ es *continua* en el punto $t_0 \in I$, si

$$\text{para todo } \varepsilon > 0, \text{existe } \delta > 0 \text{ tal que si } |t - t_0| < \delta \text{ y } t \in I, \text{ entonces } \|f(t) - f(t_0)\| < \varepsilon.$$

También decimos que f es *continua*, si es continua en cada punto t de I.

Definición 3.2.12. Sean I un intervalo abierto de reales, E un espacio vectorial normado y $f : I \subset \mathbb{R} \to E$. Diremos que f es *derivable* en $t_0 \in I$ si existe el límite siguiente

$$\lim_{h\to 0} \frac{f(t_0+h)-f(t_0)}{h}.$$

En este caso, este límite se llama derivada de f em t_0 y escribiremos

$$f'(t_0) := \lim_{h\to 0} \frac{f(t_0+h)-f(t_0)}{h}.$$

También se acostrumbra escribir $f'(t_0) = \dfrac{df}{dt}(t_0)$.

Se dice que f es *derivable* en I, si f es derivable en cada punto de I.
La aplicación $t \in I \mapsto f'(t) \in E$ se llama *función derivada* de f sobre I, y se escribe f'.

Observación 3.2.13. Sean I es un intervalo abierto de reales, E un espacio vectorial normado de dimensión finita, donde $p \in \mathbb{N}^*$, $\mathfrak{B} = (e_1, e_2, \dots, e_p)$ una base de E y $f : I \to E$. Si escribimos $f = f_1e_1 + f_2e_2 + \dots + f_pe_p$, siendo $f_1, \cdots, f_p$ funciones de $\mathbb{R}$ en $\mathbb{R}$ entonces:

1. Si $t_0 \in I$, f es derivable en t_0 si, y solo si, $f_1, f_2, \dots, f_p$ son derivables en t_0. En este caso vale $f'(t_0) = f_1'(t_0)e_1 + f_2'(t_0)e_2 + \dots + f_p'(t_0)e_p$.

2. f es derivable en I si, y solo si, $f_1, f_2, \dots, f_p$ son derivables en I, y en este caso $f' = f_1'e_1 + f_2'e_2 + \dots + f_p'e_p$.

3.3. Normas matriciales

En esta sección estudiaremos normas en los espacios de matrices $\mathcal{M}_n(\mathbb{K})$. Recordar que como $\mathcal{M}_n(\mathbb{K})$ tiene dimensión finita, entonces todas las normas son equivalentes y es un espacio de Banach con cualquier norma. Entonces tenemos una forma de medir el tamaño de una matriz usando cualquier norma.
Ahora si consideramos $\mathcal{M}_n(\mathbb{K})$, vemos que no es solo un espacio vectorial finito dimensional, sino que además tiene un producto, y entonces es útil, muchas veces, al hacer estimaciones comparar el tamaño de AB con los tamaños de A y B.
Teniendo en cuenta lo anterior, se introduce el concepto de norma matricial, que es una norma que además contempla la multipicación de matrices.

Definición 3.3.1. Una norma $\|\cdot\| : \mathcal{M}_n(\mathbb{K}) \to \mathbb{R}$, se llama *norma matricial*, si verifica:

$$\|AB\| \leq \|A\|\|B\| \qquad\qquad \text{para todo } A, B \in \mathcal{M}_n(\mathbb{K}).$$

Observación 3.3.2.
1. Una norma sobre las matrices no necesariamente verifica la propiedad anterior, que usualmente se denomina propiedad *submultiplicativa.*

2. Si $\|\cdot\|$ es una norma matricial, se cumple que $\|A^k\| \leq \|A\|^k$ para todo $k = 1, 2 \dots$.
Esta propiedad no se cumple necesariamente en cualquier norma.

Ejemplo 3.3.3. Una norma en $\mathcal{M}_n(\mathbb{K})$ que no es norma matricial.
La norma l_∞ definida en $\mathcal{M}_n(\mathbb{K})$ por

$$\|A\|_\infty = \max_{1\leq i,j\leq n} \{|a_{ij}|\}$$

es una norma pero no es una norma matricial. Si consideramos la matriz $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ y hacemos $AA = A^2 = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$, entonces tenemos que $\|A\|_\infty = 1$ y $\|AA\|_\infty = 2$ por lo cual $\|AA\|_\infty = 2 > 1 = \|A\|_\infty \|A\|_\infty$, concluyendo que no se cumple la propiedad submultiplicativa.

Sin embargo, si definimos:

$$\|A\| = n \max_{1\leq i,j\leq n} \{|a_{ij}|\} = n\|A\|_\infty \qquad\qquad A \in \mathcal{M}_n(\mathbb{K}),$$

tenemos que:

$$\|AB\| = n \max_{1\leq i,j\leq n} \left\{ \left| \sum_{k=1}^{n} a_{ik}b_{kj} \right| \right\} \leq n \max_{1\leq i,j\leq n} \left\{ \sum_{k=1}^{n} |a_{ik}b_{kj}| \right\} = n \max_{1\leq i,j\leq n} \left\{ \sum_{k=1}^{n} |a_{ik}||b_{kj}| \right\}$$

$$\leq n \sum_{k=1}^{n} \|A\|_\infty \|B\|_\infty = n\|A\|_\infty n\|B\|_\infty = \|A\|\|B\|$$

de lo cual concluimos que es una norma matricial.

Proposición 3.3.4. *La función definida para cada $A \in \mathcal{M}_n(\mathbb{K})$ por*

$$\|A\| = \sum_{i,j=1}^{n} |a_{ij}|$$

es una norma matricial.

Observar que se obtiene simplemente de sumar los valores absolutos de todos los elementos de la matriz. Es inmediato verificar que es una norma.
Probemos que se verifica la propiedad submultiplicativa

Dem. Si $A = (a_{ij}),\ B = (b_{ij})$, entonces

$$AB = (c_{ij}) = \left(\sum_{k=1}^{n} a_{ik}b_{kj} \right), \text{ siendo } \ c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

$$\|AB\| = \sum_{i,j=1}^{n} |c_{ij}| = \sum_{i,j=1}^{n} \left| \sum_{k=1}^{n} a_{ik}b_{kj} \right| \leq \sum_{i,j=1}^{n} \sum_{k=1}^{n} |a_{ik}b_{kj}| = \sum_{i,j,k=1}^{n} |a_{ik}b_{kj}|$$

$$\leq \sum_{i,j,k,m=1}^{n} |a_{ik}b_{mj}| = \left(\sum_{i,k=1}^{n} |a_{ik}| \right) \left(\sum_{m,j=1}^{n} |b_{mj}| \right) = \|A\|\|B\|.$$

La primera desigualdad proviene de la desigualdad triangular, mientras que la segunda de agregarle a la suma términos positivos. $\square$

3.4. Sucesiones y series de matrices

En las siguientes secciones 3.4, 3.5 y 3.6, trabajaremos en el espacio de Banach $(\mathcal{M}_n(\mathbb{K}), \|\cdot\|)$, donde $\|\cdot\|$ es una norma matricial cualquiera.

Comentario 3.4.1. Aunque está implícito en comentario anterior volvemos a reiterar: la definición de convergencia no depende de la norma elegida, ya que $\mathcal{M}_n(\mathbb{K})$ es un espacio vectorial de dimensión finita. Por lo tanto, el teorema 3.1.10, que afirma que todas las normas son equivalentes, es aplicable, luego si una sucesión converge para una norma, converge para todas las normas.

Si consideramos la norma del máximo: $\|A\| = n \max\limits_{1\leq i,j\leq n} \{|a_{ij}|\}$, entonces la convergencia de la sucesión (A_k) implica la convergencia elemento a elemento, es decir:

Proposición 3.4.2. *Una sucesión de matrices* $(A_k), k \in \mathbb{N}$, $A_k = \left(a_{ij}^{(k)}\right)$, *es convergente en* $\mathcal{M}_n(\mathbb{K})$ *a la matriz* $A = (a_{ij})$ *si y solo si se verifica* $\lim\limits_{k\to+\infty} a_{ij}^{(k)} = a_{ij}$ *para todo* $i, j \in \{1, 2, \ldots, n\}$. □

Sea $(A_k)_{k=0}^{\infty}$ una sucesión de matrices en $\mathcal{M}_n(\mathbb{K})$. A partir de esta sucesión nos construimos la sucesión de las sumas parciales:

$$S_k = \sum_{j=0}^{k} A_j, k \in \mathbb{N} \tag{3.4.1}$$

entonces definimos:

Definición 3.4.3. La serie $\sum_{k\geq 0} A_k$ *converge* si y solo si la sucesión $(S_k)_{k\in\mathbb{N}}$ converge; es decir si existe $\lim_k S_k = S \in \mathcal{M}_n(\mathbb{K})$. En caso afirmativo, se escribe $\sum_{k=0}^{\infty} A_k := \lim_{k\to+\infty} \sum_{j=0}^{k} A_j = S$.

Análogamente, que para las sucesiones, una serie converge si y solo si todas las "series coordenadas" convergen:

Proposición 3.4.4. *Sean* $(A_p)_{p\in\mathbb{N}}$ *una sucesión en* $\mathcal{M}_n(\mathbb{K})$, *siendo* $A_p = \left(a_{ij}^{(p)}\right) \in \mathcal{M}_n(\mathbb{K})$, *para todo* p *y* $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$. *La serie* $\sum_{p\geq 0} A_p$, *converge hacia* A *si y sólo si para todo par* (i, j), $i, j \in \{1 \ldots n\}$, *la serie* $\sum_{p\geq 0} a_{ij}^{(p)}$, *converge hacia* a_{ij}, *es decir* $\sum_{p=0}^{\infty} a_{ij}^{(p)} = a_{ij}$. *En caso de convergencia, su suma es*

$$\sum_{p=0}^{\infty} A_p = A \quad \Leftrightarrow \quad \sum_{p=0}^{\infty} \left(a_{ij}^{(p)}\right) = \left(\sum_{p=0}^{\infty} a_{ij}^{(p)}\right). \quad \square$$

Es decir, en el caso de convergencia, la suma de la serie matricial, es la matriz de las sumas de las series de cada entrada.

Definición 3.4.5. El *producto de Cauchy* de dos series $\sum\limits_{n\geq 0} a_n$ y $\sum\limits_{n\geq 0} b_n$, es la serie definida por:

$$\sum_{n\geq 0} c_n, \text{ siendo } c_n = \sum_{k=0}^{n} a_k b_{n-k}, \ \forall n \geq 0.$$

Los términos c_n de la serie producto de Cauchy, están generados por la suma de los términos de las diagonales[1], según se muestra en la Figura 3.1.

[1] Observar que en el dibujo los índices en a_i y b_i empiezan en 1 en vez de 0, pero la idea es la misma

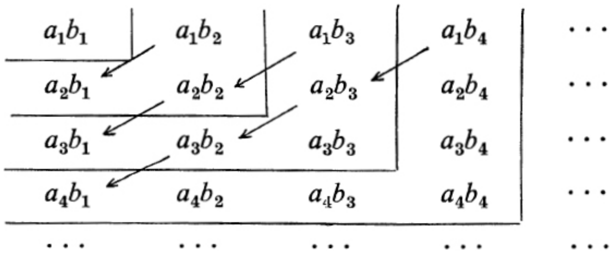

Figura 3.1: Serie producto de Cauchy

Teorema 3.4.6. *(Mertens) Sean (A_m) y (B_m) sucesiones en $\mathcal{M}_n(\mathbb{K})$.*
Supongamos que la serie $\sum_{m\geq 0} A_m$ es absolutamente convergente y la serie $\sum_{m\geq 0} B_m$ es convergente. Sea $\sum_{m\geq 0} C_m$ el producto de Cauchy de $\sum_{m\geq 0} A_m$ y $\sum_{m\geq 0} B_m$, es decir $C_m = \displaystyle\sum_{k=0}^{m} A_k B_{m-k}$. Entonces la serie $\sum_{m\geq 0} C_m$ es convergente y su suma verifica

$$\sum_{m=0}^{\infty} C_m = \sum_{m=0}^{\infty} A_m \sum_{m=0}^{\infty} B_m, \text{ es decir } \sum_{m=0}^{\infty}\sum_{k=0}^{m} A_k B_{m-k} = \sum_{m=0}^{\infty} A_m \sum_{m=0}^{\infty} B_m.$$

Dem. Supongamos que $\displaystyle\sum_{m=0}^{\infty} A_m = \mathcal{A}$ y $\displaystyle\sum_{m=0}^{\infty} B_m = \mathcal{B}$ y sean

$$\sum_{k=0}^{m} A_k = \mathcal{A}_m, \qquad \sum_{k=0}^{m} B_k = \mathcal{B}_m, \qquad \sum_{k=0}^{m} C_k = \mathcal{C}_m, \qquad \beta_m = \mathcal{B}_m - \mathcal{B} \text{ entonces } \beta_m \to 0.$$

Consideremos ahora

$$\begin{aligned}
\mathcal{C}_m &= C_0 + C_1 + C_2 + \ldots + C_m \\
&= A_0B_0 + (A_0B_1 + A_1B_0) + (A_0B_2 + A_1B_1 + A_2B_0) + \cdots + (A_0B_m + A_1B_{m-1} + \cdots + A_mB_0) \\
&= A_0(B_0 + B_1 + \cdots + B_m) + A_1(B_0 + B_1 + \cdots + B_{m-1}) + \ldots + A_mB_0 \\
&= A_0\mathcal{B}_m + A_1\mathcal{B}_{m-1} + A_2\mathcal{B}_{m-2} + \cdots + A_m\mathcal{B}_0 \\
&= A_0(\beta_m + \mathcal{B}) + A_1(\beta_{m-1} + \mathcal{B}) + A_2(\beta_{m-2} + \mathcal{B}) + \cdots + A_m(\beta_0 + \mathcal{B}) \\
&= (A_0 + A_1 + \cdots + A_m)\mathcal{B} + (A_0\beta_m + A_1\beta_{m-1} + \cdots + A_m\beta_0) \\
&= \mathcal{A}_m\mathcal{B} + \mu_m
\end{aligned}$$

siendo

$$\mu_m = A_0\beta_m + A_1\beta_{m-1} + \cdots + A_m\beta_0. \qquad (1)$$

Como $\lim \mathcal{A}_m = \mathcal{A}$, entonces $\mathcal{A}_m\mathcal{B} \to \mathcal{A}\mathcal{B}$, entonces basta probar que $\mu_m \to 0$ para que quede demostrado el teorema. Probemos esto último.

Como β_m es convergente entonces está acotada, por lo tanto existe M real positivo, tal que $\|\beta_m\| \leq M$, para todo m.

Sea $\varepsilon > 0$. Como $\sum A_m$ es absolutamente convergente, entonces la serie $\sum \|A_m\|$ es convergente, por lo tanto existe $N \in \mathbb{N}$ tal que si $m \geq N$, entonces $\displaystyle\sum_{k=N}^{m} \|A_k\| < \varepsilon$.
Entonces, separando la suma (1)

$$\begin{aligned}
\|\mu_m\| &= \|A_0\beta_m + A_1\beta_{m-1} + \cdots + A_m\beta_0\| \\
&\leq \|A_0\beta_m + A_1\beta_{m-1} + \cdots + A_{N-1}\beta_{m-N+1}\| + \ \|A_N\beta_{m-N} + A_{N+1}\beta_{m-N-1} + \cdots + A_m\beta_0\| \\
&\leq \|A_0\beta_m + A_1\beta_{m-1} + \cdots + A_{N-1}\beta_{m-N+1}\| + \ \|A_N\beta_{m-N}\| + \|A_{N+1}\beta_{m-N-1}\| + \cdots + \|A_m\beta_0\| \\
&\leq \|A_0\beta_m + A_1\beta_{m-1} + \cdots + A_{N-1}\beta_{m-N+1}\| + \ \|A_N\|\|\beta_{m-N}\| + \|A_{N+1}\|\|\beta_{m-N-1}\| + \cdots + \|A_m\|\|\beta_0\| \\
&\leq \|A_0\beta_m + A_1\beta_{m-1} + \cdots + A_{N-1}\beta_{m-N+1}\| + M\left(\|A_N\| + \|A_{N+1}\| + \cdots + \|A_m\|\right) \\
&\leq \|A_0\beta_m + A_1\beta_{m-1} + \cdots + A_{N-1}\beta_{m-N+1}\| + M\varepsilon, \quad \text{para todo } m \geq N.
\end{aligned}$$

Como N es ahora fijo, y $\beta_m \to 0$, entonces $\|A_0\beta_m + A_1\beta_{m-1} + \cdots + A_{N-1}\beta_{m-N+1}\| \underset{m\to\infty}{\to} 0$ por lo cual existe un $p \in \mathbb{N}$ donde se verfica $\|A_0\beta_m + A_1\beta_{m-1} + \cdots + A_{N-1}\beta_{m-N+1}\| < \varepsilon$, para todo $m \geq p$. Por lo cual se concluye que:

$$\text{para todo } \varepsilon > 0, \text{ existe } h \in \mathbb{N}, \ h = \max\{N, p\}, \text{ tal que para todo } m \geq h, \text{ es } \ \|\mu_m\| < (M+1)\varepsilon,$$

lo cual prueba $\mu_m \to 0$, y se completa la demostración del teorema. □

3.5. Funciones matriciales

Definición 3.5.1. Sea $P \in \mathcal{M}_n$, fijo. Definimos una función:

$$f_P : \mathcal{M}_n \to \mathcal{M}_n \text{ como } f_P(X) = PX \text{ para todo } X \in \mathcal{M}_n.$$

Análogamente se define $g_P(X) = XP$, para todo $X \in \mathcal{M}_n$.

Proposición 3.5.2. *La función definida en 3.5.1 es continua.*

Dem. Si $P = 0$ es trivial, supongamos ahora que $P \neq 0$.
Tenemos que probar que cualquiera sea $X_0 \in M_n$, se debe cumplir

$$\text{para todo } \varepsilon > 0, \text{existe } \delta > 0, \text{tal que si } X \in \mathcal{M}_n, \ \|X - X_0\| < \delta, \text{entonces } \|f_P(X) - f_P(X_0)\| < \varepsilon$$

Ahora

$$\|f_P(X) - f_P(X_0\| = \|PX - PX_0\| = \|P(X - X_0)\| \leq \|P\|\|X - X_0\|$$

Entonces basta con elegir $0 < \delta \leq \frac{\varepsilon}{\|P\|}$. □

Análogamente se demuestra que g_P es continua.

Proposición 3.5.3. *Si $P \in M_n$ es invertible, entonces la función $h : M_n \to M_n$ definida por $h(X) = PXP^{-1}$ es continua.*

Dem. Podemos expresar $h = f_P \circ g_{P^{-1}}$ donde ya probamos que f_P y $g_{P^{-1}}$ son continuas, y como la composición de funciones continuas es una función continua, se deduce la continuidad de h. □

En lo que sigue de esta sección consideramos funciones de la forma $f : I \to \mathcal{M}_n(\mathbb{K})$ donde I es un intervalo abierto de números reales.

Observación 3.5.4. Las definiciones de límites, continuidad y derivabilidad, vistas en la sección 3.2 se aplican aquí considerando $E = \mathcal{M}_n(\mathbb{K})$ como espacio normado con la norma matricial definida en la proposición 3.3.4.

Observación 3.5.5. Si $I \subset \mathbb{R}$, intervalo abierto, y consideramos la función

$$f : I \to \mathcal{M}_n(\mathbb{K}), \qquad f(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) & \dots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \dots & a_{2n}(t) \\ \vdots & & & \vdots \\ a_{n1}(t) & a_{2}(t) & \dots & a_{nn}(t) \end{pmatrix}$$

entonces las definiciones de la sección 3.2 son equivalentes a:

i) $f(t) = (a_{ij}(t))$ admite límite $\ell = (\ell_{ij}) \in \mathcal{M}_n(\mathbb{K})$ cuando $t \to t_0$ si y solo si para todo $i, j \in \{1, 2, \dots, n\}$, $a_{ij} : I \to \mathbb{K}$ tiene límite cuando $t \to t_0$, y vale $\lim\limits_{t \to t_0} a_{ij}(t) = \ell_{ij}$.

Dem. Supongamos que $\lim\limits_{t \to t_0} f(t) = l = (\ell_{ij})$ y sea $\varepsilon > 0$, entonces por definición

$$\text{existe } \delta > 0 \text{ tal que si } 0 < |t - t_0| < \delta \text{ y } t \in I, \text{ implica que } \|f(t) - \ell\|$$
$$= \| (a_{ij}(t)) - (\ell_{ij}) \| = \| (a_{ij}(t) - \ell_{ij}) \| = \sum_{i,j=1}^{n} |a_{ij}(t) - \ell_{ij}| < \varepsilon$$

Por lo tanto debe ser $|a_{ij}(t) - \ell_{ij}| < \varepsilon$ para todo i, j, y para todo $\underset{t \neq t_0}{t \in I}$ y a la bola centro t_0 y radio δ, donde δ es el encontrado anteriormente. Entonces:

$$\text{para todo } \varepsilon > 0, \text{existe } \delta > 0 \text{ tal que si } 0 < |t - t_0| < \delta \text{ y } t \in I, \text{ implica que } |a_{ij}(t) - \ell_{ij}| < \varepsilon.$$

y se concluye que $\lim\limits_{t \to t_0} a_{ij}(t) = \ell_{ij}$, para todo $i, j \in 1, 2, \dots, n$. (1)

Supongamos ahora que se cumple (1), y sea $\varepsilon > 0$, por lo cual

$$\text{existe } \delta_{ij} > 0 \text{ tal que si } 0 < |t - t_0| < \delta_{ij} \text{ y } t \in I, \text{ implica que } |a_{ij}(t) - \ell_{ij}| < \frac{\varepsilon}{n^2}$$

Si elegimos $\delta = \min\{\delta_{ij}\}$, $i, j \in 1, 2, \dots, n$, tenemos

$$\| (a_{ij}(t)) - (\ell_{ij}) \| = \sum_{i,j=1}^{n} |a_{ij}(t) - \ell_{ij}| < \varepsilon, \text{ para todo } t \in I, \text{ y } 0 < |t - t_0| < \delta$$

donde concluimos que $\lim\limits_{t \to t_0} (a_{ij}(t)) = (\ell_{ij}) = \ell$ □

- También es válida la propiedad:

$$\text{Si } \lim_{t \to t_0} f(t) = \ell \text{ y } \lim_{t \to t_0} g(t) = \ell' \text{ entonces } \lim_{t \to t_0} (fg)(t) = \ell\ell'$$

ii) f es continua si y solo si $a_{ij} : I \to \mathbb{K}$ son continuas para todo $i, j \in \{1, 2, \dots, n\}$.

iii) f es derivable si y solo si $a_{ij} : I \to \mathbb{K}$ son derivables para todo $i, j \in \{1, 2, \dots, n\}$, y

$$\frac{d}{dt} f(t) = f'(t) = \begin{pmatrix} a'_{11}(t) & a'_{12}(t) & \dots & a'_{1n}(t) \\ a'_{21}(t) & a'_{22}(t) & \dots & a'_{2n}(t) \\ \vdots & & & \vdots \\ a'_{n1}(t) & a'_{2}(t) & \dots & a'_{nn}(t) \end{pmatrix}$$

Observación 3.5.6. Valen algunas de las propiedades usuales de las derivadas en los reales, como por ejemplo:

1. Si f es derivable en a, entonces f es continua en a.
2. $(fg)' = f'g + fg'$, siendo $f, g : I \to \mathcal{M}_n$, regla de la cadena, etc.

Probamos la propiedad:

Proposición 3.5.7. *Sean $A, B : I \to \mathcal{M}_n(\mathbb{K})$ funciones derivables en $t_0 \in I$, intervalo abierto de $\mathbb{R}$, entonces AB es derivable en t_0 y $(AB)'(t_0) = A'(t_0)B(t_0) + A(t_0)B'(t_0)$.*

Dem. Consideremos

$$\begin{aligned}
&\frac{(AB)(t_0+h)-(AB)(t_0)}{h} = \frac{A(t_0+h)B(t_0+h)-A(t_0)B(t_0)}{h} \\
&= \frac{A(t_0+h)B(t_0+h)-A(t_0)B(t_0)+A(t_0+h)B(t_0)-A(t_0+h)B(t_0)}{h} \\
&= \frac{A(t_0+h)\left[B(t_0+h)-B(t_0)\right]+\left[A(t_0+h)-A(t_0)\right]B(t_0)}{h} \\
&= A(t_0+h)\left[\frac{B(t_0+h)-B(t_0)}{h}\right] + \left[\frac{A(t_0+h)-A(t_0)}{h}\right]B(t_0).
\end{aligned}$$

Teniendo en cuenta que

$$\lim_{h\to 0}\frac{A(t_0+h)-A(t_0)}{h} = A'(t_0), \quad \lim_{h\to 0}\frac{B(t_0+h)-B(t_0)}{h} = B'(t_0)$$

que A es continua en t_0 por ser derivable en t_0 y usando las propiedades de los límites, nos queda, pasando al límite:

$$\lim_{h\to 0}\frac{(AB)(t_0+h)-(AB)(t_0)}{h} = A(t_0)B'(t_0) + A'(t_0)B(t_0)$$

lo que completa la demostración. □

Integración

Adoptaremos por comodidad y por adaptarse a nuestro propósitos la siguiente definición de integral:

Definición 3.5.8. Sea $A : [a,b] \to \mathcal{M}_n(\mathbb{K})$, $A(t) = (a_{ij}(t))$, una función acotada. Entonces se dice que A es *integrable Riemann* en $[a,b]$, y se escribe $\displaystyle\int_a^b A(t)dt$, si solo si cada entrada (a_{ij}) lo es, y se define

$$\int_a^b A(t)dt := \left(\int_a^b a_{ij}(t)dt\right).$$

Entonces la integral de la matriz es la matriz de las integrales de cada entrada. Es claro que si A es continua, entonces es integrable según la definición.

Ejemplo 3.5.9. Sea $A : \mathbb{R} \to \mathcal{M}_2(\mathbb{K})$ definida por

$$A(t) = \begin{pmatrix} cos(t) & t \\ 1 & e^{2t} \end{pmatrix} \text{ entonces } \int_0^\pi A(t)dt = \begin{pmatrix} \int_0^\pi cos(t)dt & \int_0^\pi tdt \\ \int_0^\pi 1dt & \int_0^\pi e^{2t}dt \end{pmatrix} = \begin{pmatrix} 0 & \frac{\pi^2}{2} \\ \pi & \frac{e^{2\pi}-1}{2} \end{pmatrix}.$$

Propiedades de la integral

1. Linealidad:
   Si $A, B : [a, b] \to \mathcal{M}_n(\mathbb{K})$ son integrables y α y β son elementos cualesquiera de $\mathbb{K}$, entonces $\alpha A + \beta B$ es integrable y
$$\int_a^b (\alpha A(t) + \beta B(t))\, dt = \alpha \int_a^b A(t)dt + \beta \int_a^b B(t)dt.$$

2. Si $a \leq c \leq b$, la función $A : [a, b] \to \mathcal{M}_n(\mathbb{K})$ es integrable si y solo si $A|_{[a,c]}$ y $A|_{[c,b]}$ son integrables, y entonces
$$\int_a^c A(t)dt + \int_c^b A(t)dt = \int_a^b A(t)dt.$$

3. Relación de Chasles: Si $A : I \to \mathcal{M}_n(\mathbb{K})$ es continua, entonces para todo $a, b, c \in I$, se cumple
$$\int_a^b A(t)dt + \int_b^c A(t)dt = \int_a^c A(t)dt.$$

4. Si $A : [a, b] \to \mathcal{M}_n(\mathbb{K})$ es continua se verifica
$$\left\| \int_a^b A(t)dt \right\| \leq \int_a^b \|A(t)\|dt.$$
   considerando la norma definida en la proposición 3.3.4.

Observación 3.5.10. *a*) Las propiedades 1, 2, 3 son consecuencias de las propiedades de la integral en los reales y en la propiedad 4 además debido a la definición de norma.

b) Como $\mathcal{M}_n(\mathbb{K})$ es un espacio normado, la norma $\|\cdot\|$ es una función continua. Entonces en la propiedad 4, la función $\|A(t)\|$ es continua, por ser la compuesta de funciones continuas, por lo tanto $\|A(t)\|$ es integrable.

Proposición 3.5.11. *Teorema fundamental del cálculo integral.*
Sea $A : [a, b] \to \mathcal{M}_n(\mathbb{K})$ una función continua con valores en el espacio de Banach $\mathcal{M}_n(\mathbb{K})$ con norma matricial. Entonces la función $\mathcal{F} : [a, b] \to \mathcal{M}_n(\mathbb{K})$ definida por $\mathcal{F}(t) = \int_a^t A(x)dx$ es derivable en $[a, b]$ y $\mathcal{F}'(t) = A(t)$ para todo $t \in [a, b]$ (en a y b se consideran las derivadas laterales correspondientes).

Corolario 3.5.12. *(Regla de Barrow). Sean el espacio de Banach $(\mathcal{M}_n(\mathbb{K}), \|\cdot\|)$ y la función $A : [a, b] \to \mathcal{M}_n(\mathbb{K})$ derivable con derivada continua, entonces se cumple*
$$A(b) - A(a) = \int_a^b A'(t)dt$$

Las demostraciones tanto de la proposición 3.5.11 y del corolario 3.5.12 se deducen directamente de la definición de integral y las propiedades de la integral de Riemann en los reales.

3.6. Exponencial matricial

En esta sección introduciremos y utilizaremos el concepto de la exponencial matricial, que nos permitirá dar un tratamiento elegante y conciso al tema de los sistemas de ecuaciones diferenciales lineales con coeficientes constantes.

Definición y propiedades

Ahora nos concentraremos en definir la exponencial matricial y ver algunas de sus propiedades que luego utilizaremos. Para definirla, procederemos por analogía con la definición de la exponencial de un número real por medio de una serie de potencias.

Recordemos, para todo $a \in \mathbb{R}$, vale $e^a = 1 + a + \dfrac{a^2}{2!} + \dfrac{a^3}{3!}...... = \displaystyle\sum_{k=0}^{\infty} \frac{1}{k!} a^k$, entonces primeramente probaremos:

Teorema 3.6.1. *Si $\|\cdot\|$ es una norma matricial definida en $\mathcal{M}_n(\mathbb{K})$ y $A \in \mathcal{M}_n(\mathbb{K})$, entonces la serie de potencias*

$$\sum_{k=0}^{\infty} \frac{1}{k!} A^k = I + A + \frac{1}{2} A^2 + \cdots$$

es absolutamente convergente.

Dem. Notar que vale $0 \leq \|\frac{1}{k!} A^k\| = \dfrac{1}{k!} \|A^k\| \leq \dfrac{1}{k!} \|A\|^k$, para todo $k \in \mathbb{N}$, porque $\|\cdot\|$ es una norma matricial.

Como $\displaystyle\sum_{k\geq 0} \frac{1}{k!} \|A\|^k$ es una serie convergente $\left(\displaystyle\sum_{k=0}^{\infty} \frac{1}{k!} \|A\|^k = e^{\|A\|} \right)$ entonces la serie de términos no negativos $\displaystyle\sum_{k\geq 0} \|\frac{1}{k!} A^k\|$ es convergente, es decir, la serie $\displaystyle\sum_{k\geq 0} \frac{1}{k!} A^k$ converge absolutamente.
Como $\displaystyle\sum_{k\geq 0} \frac{1}{k!} A^k$ es absolutamente convergente y $\mathcal{M}_n(\mathbb{K})$ es un espacio de Banach, entonces $\displaystyle\sum_{k\geq 0} \frac{1}{k!} A^k$ converge en $\mathcal{M}_n(\mathbb{K})$ (Prop. 3.2.8). $\square$

Ahora tiene sentido la siguiente definición:

Definición 3.6.2. La *exponencial* de la matriz $A \in \mathcal{M}_n(\mathbb{K})$, se define como la suma de la serie

$$\sum_{k=0}^{\infty} \frac{1}{k!} A^k \text{ y escribimos } e^A = \exp(A) := \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

Las exponenciales de matrices tienen propiedades similares a las exponenciales de los números. Enunciaremos y probaremos alguna de ellas. Primeramente probaremos el siguiente lema:

Lema 3.6.3. *Sean A y B matrices que conmutan, es decir $AB = BA$, entonces para todo $j \in \mathbb{N}$*

$$(A + B)^j = \sum_{k=0}^{j} \binom{j}{k} A^k B^{j-k}.$$

Dem. Probaremos la fórmula por inducción completa sobre j. Para $j = 0$, queda la matriz identidad en ambos miembros. Sea ahora:

$$\begin{aligned}(A+B)^{j+1} &= (A+B)(A+B)^j = (A+B)\sum_{k=0}^{j}\binom{j}{k}A^kB^{j-k}\\ &= A\sum_{k=0}^{j}\binom{j}{k}A^kB^{j-k} + B\sum_{k=0}^{j}\binom{j}{k}A^kB^{j-k},\end{aligned}$$

habiendo usado la hipótesis de inducción. Teniendo en cuenta además el hecho de que A y B conumutan, obtenemos:

$$\begin{aligned}(A+B)^{j+1} &= \sum_{k=0}^{j}\binom{j}{k}A^{k+1}B^{j-k} + \sum_{k=0}^{j}\binom{j}{k}A^kB^{j+1-k}\\ &= A^{j+1} + \sum_{k=1}^{j}\binom{j}{k-1}A^kB^{j+1-k} + \sum_{k=1}^{j}\binom{j}{k}A^kB^{j+1-k} + B^{j+1}\\ &= A^{j+1} + \sum_{k=1}^{j}\left(\binom{j}{k-1}+\binom{j}{k}\right)A^kB^{j+1-k} + B^{j+1}.\end{aligned}$$

Pero como $\binom{j}{k-1}+\binom{j}{k}=\binom{j+1}{k}$, entonces

$$(A+B)^{j+1} = A^{j+1} + \sum_{k=1}^{j}\binom{j+1}{k}A^kB^{j+1-k} + B^{j+1} = \sum_{k=0}^{j+1}\binom{j+1}{k}A^kB^{j+1-k}. \quad \square$$

Proposición 3.6.4. *Si $A, B \in \mathcal{M}_n(\mathbb{F})$, entonces se verifica*

i) $e^0 = I_n$.

ii) $e^{cI_n} = e^cI_n$ *para todo* $c \in \mathbb{K}$.

iii) A *conmuta con* e^A *es decir* $Ae^A = e^AA$.

iv) Si $AB = BA$ *entonces* $e^{A+B} = e^Ae^B$.

v) e^A *es invertible y su inversa es* e^{-A}.

vi) $e^{PAP^{-1}} = Pe^AP^{-1}$ *para toda matriz* P *invertible,* $P \in \mathcal{M}_n(\mathbb{K})$.

Dem. Probaremos solamente las propiedades, iv), v) y vi). Las otras son fáciles.

iv) Usaremos el lema 3.6.3. El hecho de que el producto de dos series de potencias absolutamente convergentes se puede calcular usando el producto de Cauchy, nos lleva:

$$\exp(A).\exp(B) \underset{\text{Mertens}}{=} \sum_{j=0}^{\infty}\left(\sum_{k=0}^{j}\frac{1}{k!}A^k\frac{1}{(j-k)!}B^{j-k}\right) \underset{l=j-k}{=} \sum_{j=0}^{\infty}\left(\sum_{k+l=j}\frac{1}{k!\,l!}A^kB^l\right) =$$
$$= \sum_{j=0}^{\infty}\frac{1}{j!}\left(\sum_{k+l=j}\frac{j!}{k!\,l!}A^kB^l\right) = \sum_{j=0}^{\infty}\frac{1}{j!}\left(\sum_{k+l=j}\binom{j}{k}A^kB^l\right)$$
$$\underset{\text{Lema 3.6.3}}{=} \sum_{j=0}^{\infty}\frac{1}{j!}(A+B)^j = \exp(A+B). \quad \square$$

v). Las matrices A y $-A$ conmutan. Entonces

$$e^Ae^{-A} = e^{A+(-A)} = e^0 = I_n$$

de lo cual se deduce que e^A es invertible y $(e^A)^{-1} = e^{-A}$. $\square$

vi) Si $k \in \mathbb{N}$ vale

$$(P^{-1}AP)^k = P^{-1}A^kP.$$

Por lo cual

$$\sum_{k=0}^{j}\frac{1}{k!}(P^{-1}AP)^k = \sum_{k=0}^{j}\frac{1}{k!}P^{-1}A^kP = P^{-1}\left(\sum_{k=0}^{j}\frac{1}{k!}A^k\right)P.$$

Entonces pasando al límite cuando $j \to +\infty$

$$e^{P^{-1}AP} = \lim_{j\to+\infty}\sum_{k=0}^{j}\frac{1}{k!}(P^{-1}AP)^k = \lim_{j\to+\infty}\sum_{k=0}^{j}\frac{1}{k!}(P^{-1}A^kP)$$
$$= \lim_{j\to+\infty}P^{-1}\left(\sum_{k=0}^{j}\frac{1}{k!}A^k\right)P = P^{-1}\left(\lim_{j\to+\infty}\sum_{k=0}^{j}\frac{1}{k!}A^k\right)P = P^{-1}e^AP.$$

La penúltima igualdad se justifica por la proposición 3.5.3. $\square$

Derivada de la función exponencial matricial

Proposición 3.6.5. *Sean $A \in \mathcal{M}_n(\mathbb{K})$ y la función $\exp : \mathbb{R} \to \mathcal{M}_n(\mathbb{K})$ definida por $\exp(t) = e^{tA}$. Entonces la función $\exp$ es derivable y vale*

$$\exp'(t) = \exp(t)A \quad \textit{para todo } t \in \mathbb{R}.$$

Dem.

Vamos a calcular en cualquier t_0 real el siguiente límite:

$$\lim_{h\to0}\frac{\exp(t_0+h)-\exp(t_0)}{h} = \lim_{h\to0}\frac{e^{(t_0+h)A}-e^{t_0A}}{h}$$
$$\lim_{h\to0}\frac{e^{t_0A+hA}-e^{t_0A}}{h} \underset{t_0\text{A y hA conmutan}}{=} \lim_{h\to0}\frac{e^{t_0A}e^{hA}-e^{t_0A}}{h}$$
$$= \lim_{h\to0}\frac{e^{t_0A}(e^{hA}-I)}{h}.$$

Vamos a probar que $\lim_{h\to 0} \dfrac{e^{hA}-I}{h} = A$.

Si $A = 0$ es trivial, suponemos $A \neq 0$.

$$\text{Como por definición : } e^{hA} = \sum_{k=0}^{+\infty} \frac{(hA)^k}{k!} \text{ se sigue que } \quad e^{hA} - I - hA = \sum_{k=2}^{+\infty} \frac{(hA)^k}{k!}.$$

Por lo cual tenemos que:

$$\frac{e^{hA}-I}{h} - A = \frac{1}{h}\sum_{k=2}^{+\infty} \frac{(hA)^k}{k!} \quad \text{con } h \neq 0.$$

Entonces:

$$\begin{aligned}
\left\| \frac{e^{hA}-I}{h} - A \right\| &= \left\| \frac{1}{h}\sum_{k=2}^{+\infty} \frac{(hA)^k}{k!} \right\| = \frac{1}{|h|} \left\| \lim_{j\to+\infty} \sum_{k=2}^{j} \frac{(hA)^k}{k!} \right\| \\
&\underset{\text{continuidad de la } \|\cdot\|}{=} \frac{1}{|h|} \lim_{j\to+\infty} \left\| \sum_{k=2}^{j} \frac{(hA)^k}{k!} \right\| \leq \frac{1}{|h|} \lim_{j\to+\infty} \sum_{k=2}^{j} \frac{\left\|(hA)^k\right\|}{k!} \\
&\leq \frac{1}{|h|} \lim_{j\to+\infty} \sum_{k=2}^{j} \frac{\|(hA)\|^k}{k!} = \frac{1}{|h|} \sum_{k=2}^{+\infty} \frac{\|(hA)\|^k}{k!} \\
&\text{esta última serie real converge a } \; e^{\|hA\|} - 1 - \|hA\|.
\end{aligned}$$

Si llamamos $r(h) = e^{\|hA\|} - 1 - \|hA\|$ para todo h real y recordamos que en los reales vale $\lim_{x\to 0} \dfrac{e^x - 1 - x}{x} = 0$, entonces

$$\left\| \frac{e^{hA}-I}{h} - A \right\| \leq \frac{r(h)}{|h|} = \frac{e^{\|hA\|} - 1 - \|hA\|}{|h|} = \|A\| \left(\frac{e^{\|hA\|} - 1 - \|hA\|}{\|hA\|} \right) \underset{h\to 0}{\to} 0$$

lo cual $\lim\limits_{h\to 0} \dfrac{e^{hA}-I}{h} = A$ por lo tanto $\lim_{h\to 0} e^{t_0 A} \dfrac{(e^{hA}-I)}{h} = e^{t_0 A} A$ y concluimos que

$$\frac{d}{dt} e^{tA} = e^{tA} A, \; \text{ para todo } t \in \mathbb{R}.$$

$\square$

Observación 3.6.6. Por propiedad *iii*) de la proposición 3.6.4

$$A e^{tA} = e^{tA} A \text{ entonces } \left(e^{tA}\right)' = A e^{tA} \text{ para todo } t \in \mathbb{R}.$$

3.7. Exponencial de una matriz en bloques de Jordan.

3.7.1. e^{tJ} cuando J es un bloque de Jordan

Sea $J_m = J_m(\lambda) = (a_{ij}) \in \mathcal{M}_m(\mathbb{R})$ un bloque de Jordan de la forma

$$a_{ij} = \begin{cases} \lambda & \text{si } i=j \\ 1 & \text{si } j=i+1 \\ 0 & \text{en otros casos} \end{cases} \quad \text{es decir:} \quad J_m = \underbrace{\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}}_{m}.$$

Vamos a calcular e^{tJ_m}. El bloque J_m se puede expresar

$$J_m = \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 0 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \dots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} = \lambda I_m + N_m$$

Para simplificar la notación, escribiremos $J = J_m$, $I = I_m$ y $N = N_m$, para todo m. Si hacemos las potencias sucesivas $N^2,\ N^3\ \dots,\ N^k, \dots$, obtenemos

$$N^2 = \begin{pmatrix} 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \dots & & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}, \cdots, N^{m-1} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \dots & & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} \text{ y } N^k = 0, \text{ para todo } k \geq m.$$

Calculamos ahora e^{tJ}. Como las matrices $t\lambda I$ y tN conmutan, obtenemos

$$e^{tJ} = e^{t(\lambda I + N)} = e^{t\lambda I + tN} = e^{t\lambda I} e^{tN}.$$

Aplicando la proposición 3.6.4 parte ii), obtenemos $e^{t\lambda I} = e^{t\lambda} I = \begin{pmatrix} e^{\lambda t} & 0 & \dots & 0 \\ 0 & e^{\lambda t} & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & e^{\lambda t} \end{pmatrix}$.

Por lo tanto

$$\begin{aligned} e^{tJ} &= e^{t(\lambda I + N)} = e^{t\lambda I + tN} = e^{t\lambda I} e^{tN} = e^{\lambda t} I e^{tN} = e^{\lambda t} \sum_{r=0}^{\infty} \frac{1}{r!} (tN)^r \\ &= e^{\lambda t} \left(I + tN + \frac{t^2}{2} N^2 + \cdots + \frac{t^{m-1}}{(m-1)!} N^{m-1} \right) \\ &= e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \frac{t^3}{3!} & \dots & \frac{t^{m-2}}{(m-2)!} & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & \frac{t^2}{2!} & \dots & \frac{t^{m-3}}{(m-3!} & \frac{t^{m-2}}{(m-2)!} \\ 0 & 0 & 1 & t & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 & t \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Resumiendo y recuperando la notación, tenemos que si J_m es un bloque de Jordan de orden m entonces

$$e^{tJ_m} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \frac{t^3}{3!} & \dots & \frac{t^{m-2}}{(m-2)!} & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & \frac{t^2}{2!} & \dots & \frac{t^{m-3}}{(m-3!} & \frac{t^{m-2}}{(m-2)!} \\ 0 & 0 & 1 & t & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 & t \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}. \tag{3.7.1}$$

3.7.2. Exponencial de una matriz diagonal en bloques.

Recordemos también que si tenemos una matriz diagonal en bloques

$$A = \begin{pmatrix} \boxed{B_1} & & \\ & \ddots & \\ & & \boxed{B_p} \end{pmatrix}$$

entonces

$$A^2 = \begin{pmatrix} \boxed{B_1^2} & & \\ & \ddots & \\ & & \boxed{B_p^2} \end{pmatrix}, \dots, A^k = \begin{pmatrix} \boxed{B_1^k} & & \\ & \ddots & \\ & & \boxed{B_p^k} \end{pmatrix} \quad \text{con } k \text{ natural mayor o igual a 1.}$$

Proposición 3.7.1. *Sea $A \in \mathcal{M}_n(\mathbb{K})$ una matriz diagonal en bloques,*

$$A = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_p \end{pmatrix} \quad \textit{entonces} \quad e^A = \begin{pmatrix} e^{B_1} & & \\ & \ddots & \\ & & e^{B_p} \end{pmatrix}.$$

Dem.

$$e^A = \sum_{r=0}^{\infty} \frac{1}{r!} A^r = \sum_{r=0}^{\infty} \frac{1}{r!} \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_p \end{pmatrix}^r = \sum_{r=0}^{\infty} \frac{1}{r!} \begin{pmatrix} B_1^r & & \\ & \ddots & \\ & & B_p^r \end{pmatrix} = \sum_{r=0}^{\infty} \begin{pmatrix} \frac{1}{r!} B_1^r & & \\ & \ddots & \\ & & \frac{1}{r!} B_p^r \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{r=0}^{\infty} \frac{1}{r!} B_1^r & & \\ & \ddots & \\ & & \sum_{r=0}^{\infty} \frac{1}{r!} B_p^r \end{pmatrix} = \begin{pmatrix} e^{B_1} & & \\ & \ddots & \\ & & e^{B_p} \end{pmatrix}.$$

□

3.7.3. e^{tJ} cuando J es una matriz en bloques de Jordan.

Entonces si J es una matriz de Jordan de la forma

$$J = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_p \end{pmatrix},$$

tenemos de acuerdo a la proposición 3.7.1

$$e^{tJ} = \begin{pmatrix} e^{tJ_1} & 0 & \dots & 0 \\ 0 & e^{tJ_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{tJ_p} \end{pmatrix}, \tag{3.7.2}$$

donde cada bloque $e^{tJ_i}, i = 1, 2, ..., p$ se calcula según la fórmula (3.7.1).

3.8. Cálculo de la función matricial e^{tA}.

Los teoremas 4.2.1 y 4.3.1 nos darán una fórmula explícita para las soluciones de los sistemas homogéneos y no homogéneos de ecuaciones diferenciales lineales con coeficientes constantes, pero quedará el problema de calcular la matriz exponencial e^{tA}. Si calcularamos e^{tA} directamente usando la definición, tendríamos que calcular todas las potencias A^k para $k = 0, 1, \dots,$ y luego calcular cada suma $\sum_{k=0}^{\infty} \frac{t^k a_{ij}^{(k)}}{k!}$, siendo $A = (a_{ij})$. Este procedimiento puede ser muy engorroso salvo casos especiales como por ejemplo que la matriz A sea diagonal.
Si bien existen varios métodos para calcular la matriz exponencial, usaremos aquí fórmulas que involucran la forma canónica de Jordan, uno de los métodos clásicos que lo hemos estudiado en el capítulo 2. También existen otros métodos computacionales para calcular e^{tA}, usando resultados clásicos de análisis, teoria de la aproximación y teoría matricial.

3.8.1. Caso en que A es diagonal

Sea la matriz diagonal

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

entonces

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k = \begin{pmatrix} \sum_{k=0}^{\infty} \frac{t^k}{k!} \lambda_1^k & 0 & \dots & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{t^k}{k!} \lambda_2^k & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \sum_{k=0}^{\infty} \frac{t^k}{k!} \lambda_n^k \end{pmatrix} = \begin{pmatrix} e^{t\lambda_1} & 0 & \dots & 0 \\ 0 & e^{t\lambda_2} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & e^{t\lambda_n} \end{pmatrix}. \tag{3.8.1}$$

3.8.2. Caso en que A es diagonalizable

Si la matriz $A \in \mathcal{M}_n(\mathbb{K})$ es diagonalizable, entonces existe una matriz P, invertible y una matriz diagonal

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

tal que $A = PDP^{-1}$. En este caso, tenemos

$$e^{tA} = Pe^{tD}P^{-1} = P \begin{pmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{pmatrix} P^{-1},$$

donde los elementos de D, son los valores propios de la matriz A, y las columnas de P son los vectores propios correspondientes de A.

Ejemplo 3.8.1. Consideremos la matriz, $\begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & -1 & 1 \end{pmatrix}$. Veamos primeramente que es diagonalizable. Calculamos los valores propios:

$$\mathcal{X}_A(\lambda) = \det(A - \lambda I) = (4 - \lambda)(\lambda - 2)^2.$$

Los valores propios son entonces 2 y 4. Ahora, como:
• la multiplicidad algebraica de $\lambda = 4$ es 1 entonces la multiplicidad geométrica de $\lambda = 4$ es 1.
• la multiplicidad algebraica de $\lambda = 2$ es 2, y la multiplicidad geométrica(λ=2)=3$-$rango($A-2I$)=3$-$1=2.
Concluimos que A es diagonalizable porque $\mathcal{X}_A$ se escinde y $MA(\lambda) = MG(\lambda)$ para todo λ.
Como además : $E_4 = \ker(A - 4I) = [(1, 2, -1)]$ y $E_2 = \ker(A - 2I) = [(-1, 1, 0), (-1, 0, 1)]$, podemos escribir

$$A = PDP^{-1}, \text{ siendo } D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \text{ y } P = \begin{pmatrix} 1 & -1 & -1 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Entonces

$$e^{tA} = Pe^{tD}P^{-1} = \begin{pmatrix} 1 & -1 & -1 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{4t} & 0 & 0 \\ 0 & e^{2t} & 0 \\ 0 & 0 & e^{2t} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -1 & 0 & -1 \\ \frac{1}{2} & \frac{1}{2} & \frac{3}{2} \end{pmatrix}$$

$$= \frac{e^{2t}}{2} \begin{pmatrix} e^{2t} + 1 & e^{2t} - 1 & e^{2t} - 1 \\ 2e^{2t} - 2 & 2e^{2t} & 2e^{2t} - 2 \\ 1 - e^{2t} & 1 - e^{2t} & 3 - e^{2t} \end{pmatrix}.$$

3.8.3. Caso en que A no es diagonalizable

Uno de los principales própositos de este trabajo, es describir como se puede calcular la matriz exponencial e^{tA} cuando A no es diagonalizable. Lo que haremos es transformar la matriz A en un matriz que tenga la forma canónica de Jordan, según el teorema 2.5.1.

Ejemplo 3.8.2. Sea la matriz $A = \begin{pmatrix} 0 & -3 & 1 & 2 \\ -2 & 1 & -1 & 2 \\ -2 & 1 & -1 & 2 \\ -2 & -3 & 1 & 4 \end{pmatrix}$. De acuerdo a lo visto en el ejemplo 2.5.6, la forma de Jordan de A es,

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \text{siendo además } P = \begin{pmatrix} -1 & 0 & 1 & -1 \\ -1 & 1 & 0 & 1 \\ -1 & 2 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix} \text{la matriz de transición.}$$

Luego como $A = PJP^{-1}$ entonces $tA = P(tJ)P^{-1}$ y por parte vi) de la proposición 3.6.4

$$e^{tA} = Pe^{tJ}P^{-1}.$$

Calculando primero e^{tJ}, usando (3.7.2) y (3.7.1), nos queda

$$e^{tJ} = \begin{pmatrix} 1 & t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{2t} & 0 \\ 0 & 0 & 0 & e^{2t} \end{pmatrix}.$$

Entonces

$$e^{tA} = Pe^{tJ}P^{-1} = \begin{pmatrix} -1 & 0 & 1 & -1 \\ -1 & 1 & 0 & 1 \\ -1 & 2 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{2t} & 0 \\ 0 & 0 & 0 & e^{2t} \end{pmatrix} \begin{pmatrix} -1 & 0 & 1 & -1 \\ -1 & 1 & 0 & 1 \\ -1 & 2 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} 1 & -2e^{2t}+t+2 & e^{2t}-t-1 & e^{2t}-1 \\ 1-e^{2t} & t+1 & -t & e^{2t}-1 \\ 1-e^{2t} & t & 1-t & e^{2t}-1 \\ 1-e^{2t} & -2e^{2t}+t+2 & e^{2t}-t-1 & 2e^{2t}-1 \end{pmatrix}.$$

$\square$

Ejemplo 3.8.3. Vamos a calcular e^{tA}, siendo $A = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 \\ -1 & 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 0 \\ 1 & 1 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}$.

Empezamos calculando el polinomio característico de A

$$\mathcal{X}_A(t) = \det(A - tI) = (t+1)(t-2)^5.$$

por lo cual los valores propios son -1 y 2, y además $MA(-1) = 1$ y $MA(2) = 5$. Por otro lado $\text{rango}(A - 2I) = 4$ entonces $\dim\ker(A - 2I) = 6 - 4 = 2$ por lo cual la matriz A no es diagonalizable. Lo primero que podemos decir es que la matriz de Jordan tiene 2 bloques de la forma

$$\begin{pmatrix} \boxed{-1} & \\ & \boxed{J_2} \end{pmatrix}$$

con un -1 y cincos 2 en la diagonal principal. Por otro lado, como $MG(2)=2$, entonces deducimos que hay 2 bloques de Jordan en J_2, que pueden ser de la forma $(3,2)$ o $(4,1)$. Es decir las posibles formas de J_2 son

$$J_2=\left(\begin{array}{ccc|cc} 2 & 1 & & & \\ & 2 & 1 & & \\ & & 2 & & \\ \hline & & & 2 & 1 \\ & & & & 2 \end{array}\right) \text{ o } J_2=\left(\begin{array}{cccc|c} 2 & 1 & & & \\ & 2 & 1 & & \\ & & 2 & 1 & \\ & & & 2 & \\ \hline & & & & 2 \end{array}\right) \tag{3.8.2}$$

Primeramente observemos que si

$$A=PJP^{-1}, \text{ y si } q(t)\in\mathbb{K}[t] \text{ entonces } q(A)=Pq(J)P^{-1}. \text{ Luego } q(A)=0 \Longleftrightarrow q(J)=0.$$

Entonces, consideremos el polinomio $q(t)=(t+1)(t-2)^3$. Operando para la matriz A de este ejemplo, se tiene que $q(A)=(A+I)(A-2I)^3\neq 0$, así que debe ser $q(J)\neq 0$. Ahora

$$q(J)=\left(\begin{array}{c|c} q(-1) & \\ \hline & q(J_2) \end{array}\right)=q(J)=\left(\begin{array}{c|c} 0 & \\ \hline & q(J_2) \end{array}\right).$$

Calculemos entonces $q(J_2)=(J_2+I)(J_2-2I)^3$ en los dos casos según 3.8.2:

$$q\left(\left(\begin{array}{ccc|cc} 2 & 1 & & & \\ & 2 & 1 & & \\ & & 2 & & \\ \hline & & & 2 & 1 \\ & & & & 2 \end{array}\right)\right)=\left(\begin{array}{ccc|cc} 3 & 1 & & & \\ & 3 & 1 & & \\ & & 3 & & \\ \hline & & & 3 & 1 \\ & & & & 3 \end{array}\right)\left(\begin{array}{ccc|cc} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ \hline & & & 0 & 1 \\ & & & & 0 \end{array}\right)^3=0$$

de esta forma nos queda $q(J_2)=0$ y por lo tanto $q(J)=0$, que no es posible. Entonces tiene que ser $J_2=\left(\begin{array}{cccc|c} 2 & 1 & & & \\ & 2 & 1 & & \\ & & 2 & 1 & \\ & & & 2 & \\ \hline & & & & 2 \end{array}\right)$. Verifiquemos

$$q\left(\left(\begin{array}{cccc|c} 2 & 1 & & & \\ & 2 & 1 & & \\ & & 2 & 1 & \\ & & & 2 & \\ \hline & & & & 2 \end{array}\right)\right)=\left(\begin{array}{cccc|c} 3 & 1 & & & \\ & 3 & 1 & & \\ & & 3 & 1 & \\ & & & 3 & \\ \hline & & & & 3 \end{array}\right)\left(\begin{array}{cccc|c} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & 0 & \\ \hline & & & & 0 \end{array}\right)^3$$

$$\left(\begin{array}{cccc|c} 3 & 1 & & & \\ & 3 & 1 & & \\ & & 3 & 1 & \\ & & & 3 & \\ \hline & & & & 3 \end{array}\right)\left(\begin{array}{cccc|c} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right)=\left(\begin{array}{cccc|c} 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right)\neq 0$$

Como comentario, digamos que una forma útil y práctica de abordar el problema de encontrar la forma canónica de Jordan de una matriz, es mediante el concepto de polinomio minimal, que en este trabajo por brevedad no lo abordamos.

Entonces la forma canónica de Jordan de la matriz A es,

$$J = \left(\begin{array}{c|cccc|c} -1 & & & & & \\ \hline & 2 & 1 & 0 & 0 & \\ & 0 & 2 & 1 & 0 & \\ & 0 & 0 & 2 & 1 & \\ & 0 & 0 & 0 & 2 & \\ \hline & & & & & 2 \end{array}\right)$$

asi que la base de Jordan tiene la forma

$$\mathcal{B} = \left\{w, (A-2I)^3 v, (A-2I)^2 v, (A-2I)v, v, u\right\}$$

Calculemos primeramente los subespacios

$$\begin{aligned}
\ker(A-2I) \;\; &= [(0,0,-1,1,0,0),(0,0,0,0,3,1)] \\
\ker(A-2I)^2 &= [(0,0,3,0,1,0),(0,0,-9,0,0,1),(0,0,-1,1,0,0)] \\
\ker(A-2I)^3 &= [(0,\frac{9}{2},0,0,1,0),(0,\frac{-27}{2},0,0,0,1),(0,\frac{-3}{2},1,0,0,0),(0,\frac{-3}{2},0,1,0,0)] \\
\ker(A-2I)^4 &= \left\{(x,y,z,t,v,w)\in\mathbb{R}^6 : 10x+6y+9z+9t+81v-27w=0\right\}
\end{aligned}$$

Tomamos, por ejemplo, $v = (6,-10,0,0,0,0) \in \ker(A-2I)^4 \setminus \ker(A-2I)^3$, por lo cual

$$\begin{aligned}
(A-2I)v &= (0,6,-6,-10,-4,0) \\
(A-2I)^2 v &= (0,0,0,6,-10,-4) \\
(A-2I)^3 v &= (0,0,0,0,6,2)
\end{aligned}$$

Elegimos $u = (0,0,-1,1,0,0)$. Además

$$\ker(A+I) = [(0,0,0,0,0,1)]$$

y ahora estamos en condiciones de construir la base de Jordan

$$\begin{aligned}
\mathcal{B} = \{&(0,0,0,0,0,1),(0,0,0,0,6,2),(0,0,0,6,-10,-4),(0,6,-6,-10,-4,0),(6,-10,0,0,0,0), \\
&(0,0,-1,1,0,0)\}
\end{aligned}$$

Si consideramos:

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 6 & -10 & 0 \\ 0 & 0 & 0 & -6 & 0 & -1 \\ 0 & 0 & 6 & -10 & 0 & 1 \\ 0 & 6 & -10 & -4 & 0 & 0 \\ 1 & 2 & -4 & 0 & 0 & 0 \end{pmatrix}$$

se cumple que $J = P^{-1}AP$ o que $A = PJP^{-1}$.
Luego $tA = P(tJ)P^{-1}$ y por lo tanto llegamos

$$e^{tA} = Pe^{tJ}P^{-1}$$

donde

$$e^{tJ} = \begin{pmatrix} \underbrace{e^{tJ_1(-1)}}_{1\times 1} & \vdots & & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & \vdots & \underbrace{e^{tJ_2(2)}}_{4\times 4} & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & \vdots & & \vdots & \underbrace{e^{tJ_3(2)}}_{1\times 1} \end{pmatrix} = \begin{pmatrix} e^{-t} & 0 & 0 & 0 & 0 & 0 \\ 0 & e^{2t} & te^{2t} & \frac{t^2}{2}e^{2t} & \frac{t^3}{3!}e^{2t} & 0 \\ 0 & 0 & e^{2t} & te^{2t} & \frac{t^2}{2!}e^{2t} & 0 \\ 0 & 0 & 0 & e^{2t} & te^{2t} & 0 \\ 0 & 0 & 0 & 0 & e^{2t} & 0 \\ 0 & 0 & 0 & 0 & 0 & e^{2t} \end{pmatrix}$$

donde hemos utlizado los resultados (3.7.1) y (3.7.2).

Capítulo 4

Sistemas lineales de ecuaciones diferenciales de primer orden

4.1. Introducción

Un *sistema lineal de ecuaciones diferenciales de primer orden*, es un sistema de la forma:

$$\left\{\begin{array}{lcl} x_1' & = & a_{11}(t)x_1 + a_{12}(t)x_2 + \cdots + a_{1n}(t)x_n \ + f_1(t) \\ x_2' & = & a_{21}(t)x_1 + a_{22(t)}x_2 + \cdots + a_{2n}(t)x_n \ + f_2(t) \\ x_3' & = & a_{31}(t)x_1 + a_{32(t)}x_2 + \cdots + a_{3n}(t)x_n \ + f_3(t) \\ & \vdots & \\ x_n' & = & a_{n1}(t)x_1 + a_{n2}(t)x_2 + \cdots + a_{nn}(t)x_n + f_n(t) \end{array}\right. \tag{4.1.1}$$

donde las funciones a_{ik} y f_i se consideran definidas en un cierto intervalo I.

Los sistemas lineales son los más simples entre todos los sistemas de primer orden, pero incluso éstos son muy difíciles de resolver analíticamente. También son suficientemente difíciles los sistemas lineales con coeficientes constantes, pero nosotros los podremos resolver analíticamente, sin mucha dificultad, con las herramientas que estudiamos en los capítulos anteriores.

La discusión de estos sistemas se pueden simplificar considerablemente mediante la notación matricial y vectorial.

Si se introduce la matriz de coeficientes:

$$\mathbf{A}(t) = (a_{ij}(t))$$

y los vectores columnas

$$\mathbf{x} = (x_i) \qquad y \qquad \mathbf{f}(t) = (f_i(t)),$$

entonces el sistema toma la forma matricial

$$\frac{d\mathbf{x}}{dt} = \mathbf{x}' = \mathbf{A}(t)\mathbf{x} + \mathbf{f}(t).$$

Definición 4.1.1. Sean I un intervalo abierto de reales, $\mathbf{x} : I \to \mathbb{K}^n$ una función desconocida, $\mathbf{A} : I \to \mathcal{M}_n(\mathbb{K})$ y $f : I \to \mathbb{K}^n$ funciones continuas. Se llama *ecuación diferencial lineal de primer orden* a una ecuación diferencial ordinaria de la forma

$$\mathbf{x}' = \mathbf{A}(t)\mathbf{x} + \mathbf{f}(t), \quad t \in \mathbb{K} \tag{2}$$

también la llamaremos *sistema lineal de ecuaciones diferenciales de primer orden.*
Si $\mathbf{A}$ es constante, diremos que es un *sistema lineal de ecuaciones diferenciales con coeficientes constantes.*
Si $\mathbf{f} = 0$ la llamaremos *ecuación diferencial homogénea.*
Si $\mathbf{f} \neq 0$ la llamaremos *ecuación diferencial no homogénea.*

Una *solución* de la ecuación (2) en el intervalo abierto I es un vector de funciones $\mathbf{x}(t) = (x_i(t))$, tal que las funciones que conforman $\mathbf{x}$, sastifacen el sistema 4.1.1 para todo $t \in I$.

4.2. Sistemas lineales homogéneos con coeficientes constantes

En el caso en que los coeficientes de la matriz A no dependan del tiempo t, el sistema se vuelve autónomo y es posible explicitar completamente las soluciones. La forma de determinarlas es muy similar a lo que se hace en la resolución de las ecuaciones diferenciales lineales de dimensión 1, claro usando las propiedades de la matriz exponencial, que lo hemos desarrollado en el capítulo anterior. Entonces vamos a demostrar ahora un teorema fundamental para los sistemas lineales.

Teorema 4.2.1. *Sea $A \in \mathcal{M}_n(\mathbb{K})$. Entonces para un $x_0 \in \mathbb{R}^n$ dado, el problema de valor inicial*

$$(1) \qquad x' = Ax, \qquad x(0) = x_0$$

tiene una única solución f dada por

$$(2) \qquad f(t) = e^{tA}x_0, \quad \textit{para todo } t \in \mathbb{R}.$$

Dem. Aplicando la proposición 3.6.5, verifiquemos que f es solución de (1).

$$f'(t) = (e^{tA}x_0)' = \left(e^{tA}\right)' x_0 = Ae^{tA}x_0 = Af(t) \quad \text{para todo } t \in \mathbb{R}.$$

También tenemos que $f(0) = Ix_0 = x_0$, lo cual concluye la prueba de que (2) es solución de (1).
Para probar la unicidad, sea g una solución de (1) y consideremos

$$y(t) = e^{-tA}g(t).$$

Entonces

$$\begin{aligned} y'(t) &= \left(e^{-tA}\right)' g(t) + e^{-tA}g'(t) \\ &= -Ae^{-tA}g(t) + e^{-tA}Ag(t) \\ &= e^{-tA}(-A)g(t) + e^{-tA}Ag(t) \\ &= e^{-tA}((-A) + A)g(t) = 0 \end{aligned}$$

para todo $t \in \mathbb{R}$, y habiendo usado que e^{-tA} y $-A$ conmutan.

Luego $y(t)$ es constante y como $y(0) = Ig(0) = x_0$ nos queda que $y(t) = x_0$ para todo t real. Se sigue que cualquier solución de (1) es de la forma

$$g(t) = e^{tA}y(t) = e^{tA}x_0 \qquad \text{para todo } t \in \mathbb{R}$$

lo que completa la demostración del teorema. $\square$

En general vale el siguiente resultado

Corolario 4.2.2. *Si $A \in \mathcal{M}_n(\mathbb{K})$, entonces para $t_0 \in \mathbb{R}$ y $x_0 \in \mathbb{R}^n$ dados, el problema de valor inicial*

$$(1) \qquad\qquad x' = Ax, \qquad x(t_0) = x_0$$

tiene una única solución dada por

$$(2) \quad h(t) = e^{(t-t_0)A}x_0, \quad \textit{para todo } t \in \mathbb{R}.$$

Dem. La función definida $h(t) = f(t - t_0)$, donde f es la función definida en el teorema 4.2.1, es la única solución de (1). □

Observación 4.2.3. Los resultados anteriores nos permiten obtener conclusiones sobre la naturaleza de las soluciones. En nuestro desarrollo en que estamos considerando el caso en que la matriz del sistema solo tiene valores propios reales, deducimos que las soluciones son combinaciones lineales de productos de polinomios con funciones exponenciales.

4.3. Sistemas lineales no homogéneos con coeficientes constantes

Consideremos la ecuación no homogénea

$$(1) \qquad x' = Ax + f(t)$$

donde $f : I \subset \mathbb{R} \to \mathbb{R}^n$ es una función continua, I un intervalo abierto de reales.
Para estas ecuaciones como vimos para los sistemas homogéneos, también podemos obtener fórmulas explícitas para las soluciones siguiendo el mismo proceso que en los reales. Buscaremos solución de la forma

$$x(t) = e^{tA}\varphi(t),$$

donde $\varphi : I \subset \mathbb{R} \to \mathbb{R}^n$ es una función diferenciable. De hecho, cada solución puede ser escrita de esta forma porque e^{tA} es invertible.
Aplicando reglas de diferenciación y propiedades de e^{tA}

$$x'(t) = Ae^{tA}\varphi(t) + e^{tA}\varphi'(t).$$

Entonces si suponemos que $x(t)$ es solución de la ecuación (1)

$$\begin{gathered}
Ax(t) + f(t) = Ae^{tA}\varphi(t) + e^{tA}\varphi'(t) \qquad \text{pero como } \; x(t) = e^{tA}\varphi(t) \\
Ax(t) + f(t) = Ax(t) + e^{tA}\varphi'(t) \\
e^{tA}\varphi'(t) = f(t) \\
\varphi'(t) = e^{-tA}f(t)
\end{gathered}$$

Integrando, suponiendo que $t_0 \in I$

$$\varphi(t) = \int_{t_0}^{t} e^{-sA}f(s)ds + C, \qquad C \in \mathbb{R}^n.$$

Tenemos como candidato a solución

$$x(t) = e^{tA}\left(\int_{t_0}^{t} e^{-sA}f(s)ds + C\right), \qquad C \in \mathbb{R}^n.$$

Probaremos, entonces:

Teorema 4.3.1. *Sean $A \in \mathcal{M}_n(\mathbb{K})$ y $f : I \subset \mathbb{R} \to \mathbb{R}^n$ una función continua, I un intervalo abierto de reales. Entonces para $t_0 \in I$ y $x_0 \in \mathbb{R}^n$ dados, la ecuación diferencial*

$$(2) \qquad x' = Ax + f(t), \qquad x(t_0) = x_0$$

tiene una única solución ϕ definida por

$$(3) \qquad \phi(t) = e^{(t-t_0)A}x_0 + \int_{t_0}^{t} e^{(t-s)A} f(s)ds, \quad \text{para todo } t \in I.$$

Dem. Primero observemos que ϕ está bien definida, porque la función matricial exponencial y f son continuas en I y el producto de funciones continuas es continua, entonces la integral existe entre t_0 y $t \in I$.
Vamos a verificar que (3) es solución de (2), observando previamente que ϕ es derivable en I por ser suma de funciones derivables.
Derivando y usando las reglas de derivación, tenemos

$$\begin{aligned}
\phi'(t) &= \left(e^{(t-t_0)A}x_0 + \int_{t_0}^{t} e^{(t-s)A} f(s)ds\right)' \\
&= \left(e^{(t-t_0)A}x_0\right)' + \left(\int_{t_0}^{t} e^{(t-s)A} f(s)ds\right)' \text{ como además tA y -sA conmutan, nos queda} \\
&= \left(e^{(t-t_0)A}x_0\right)' + \left(\int_{t_0}^{t} e^{tA}e^{-sA} f(s)ds\right)' = \left(e^{(t-t_0)A}x_0\right)' + \left(e^{tA}\int_{t_0}^{t} e^{-sA} f(s)ds\right)' \\
&= Ae^{(t-t_0)A}x_0 + Ae^{tA}\int_{t_0}^{t} e^{-sA} f(s)ds + e^{tA}\left(\int_{t_0}^{t} e^{-sA} f(s)ds\right)' \\
&= Ae^{(t-t_0)A}x_0 + A\int_{t_0}^{t} e^{(t-s)A} f(s)ds + e^{tA}e^{-tA}f(t) \\
&= A\left(e^{(t-t_0)A}x_0 + \int_{t_0}^{t} e^{(t-s)A} f(s)ds\right) + f(t) = A\phi(t) + f(t), \quad \text{para todo } t \in I.
\end{aligned}$$

Además $\phi(t_0) = x_0$ lo que completa que ϕ es solución de (3).
Para ver que cada solución de (2) debe ser de la forma (3), consideremos $y : I \to \mathbb{R}^n$ como una segunda solución de (2). Entonces

$$(\phi - y)' = \phi' - y' = (A\phi + f) - (Ay + f) = A(\phi - y).$$

Por lo tanto

$$(4) \qquad (\phi - y)' = A(\phi - y), \text{ siendo además } (\phi - y)(t_0) = 0$$

Y de acuerdo al corolario 4.2.2, la única solución de (4) es

$$(\phi - y)(t) = e^{(t-t_0)A}0 = 0, \qquad \text{para todo } t \in I.$$

Es decir $\phi - y$ es la función nula en I, lo cual concluye que y debe ser de la forma (3). □

Observación 4.3.2. Nos podriamos preguntar si no se podría aplicar la idea de la demostración del teorema anterior para el sistema: $\mathbf{x}' = A(t)\mathbf{x} + f(t)$, es decir, no es a coeficientes constantes. La respuesta es no, porque basicamente usamos en la demostración el hecho de que $(e^{tA})' = e^{tA}A = Ae^{tA}$, pero esta fórmula no siempre es válida cuando A es una función matricial, es decir en general $(e^{A(t)})'$ no tiene por que ser necesariamente igual a $e^{A(t)}A'(t)$ o $A'(t)e^{A(t)}$. Por más detalles ver [Ap], 7.12, ejercicio 7.

4.4. Ecuaciones diferenciales lineales de orden n

Comentemos aquí que se puede reducir una ecuación diferencial lineal de orden n a un sistema de n ecuaciones lineales de primer orden. Si tenemos la ecuación lineal

$$y^{(n)} + a_1(t)y^{(n-1)} + \cdots + a_n(t)y = f(t) \tag{4.4.1}$$

siendo $a_1, \cdots, a_n$ y f funciones continuas en un cierto intervalo abierto I en los reales.
Si hacemos los cambios de variables

$$\begin{cases} y = x_1 \\ y' = x_2 \\ y'' = x_3 \\ \cdots \\ y^{(n-1)} = x_n \end{cases},$$

entonces la ecuación (4.4.1) se transforma en el sistema:

$$\begin{cases} x_1' = x_2 \\ x_2' = x_3 \\ x_3' = x_4 \\ \cdots \\ x_{n-1}' = x_n \\ x_n' = f(t) - a_n(t)x_1 - a_{n-1}(t)x_2 - \cdots - a_1(t)x_n \end{cases}. \tag{4.4.2}$$

Es inmediato probar que la ecuación (4.4.1) es equivalente al sistema (4.4.2). Veamos esto para un caso particular, cuando $n = 2$, entonces es: $y'' + a_1(t)y' + a_2(t)y = f(t)$ y si llamamos $g(t, y, y') = f(t) - a_1(t)y' - a_2(t)y$ tenemos

$$(E) : y'' = g(t, y, y')$$

y poniendo $y = x_1$, $y' = x_2$, tenemos $x_1' = y' = x_2$, entonces $y'' = x_2' = g(t, x_1, x_2)$, por lo tanto la ecuación (E) se puede describir como un sistema de 2 ecuaciones de primer orden:

$$(S) : \begin{cases} x_1' = x_2 \\ x_2' = g(t, x_1, x_2) \end{cases}$$

- Sea ϕ solución de (E) en I, entonces $x_1 = \phi(t)$ y $x_2 = \phi'(t)$ son soluciones de (S), porque $x_1' = \phi'(t) = x_2$, para todo $t \in I$, y $x_2' = \phi''(t) = g(t, \phi(t), \phi'(t) = g(t, x_1, x_2)$, para todo $t \in I$.

- Ahora si ϕ_1 y ϕ_2 son soluciones de (S) entonces ϕ_1, la primera componente, es solución de (E) en I, porque $y'' = \phi_1''(t) = (\phi_1'(t))' = \phi_2'(t) = g(t, \phi_1(t), \phi_1'(t)) = g(t, x_1, x_1')$, para todo $t \in I$, lo que completa la prueba de que (E) y (S) son equivalentes en el sentido anterior. □

La prueba en el caso general, es solo una generalización del caso anterior.

Ahora si escribimos 4.4.2 en notación matricial, nos queda de la forma $\mathbf{x}' = A(t)\mathbf{x} + b(t)$, siendo

$$A(t) = \begin{pmatrix} 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ -a_n(t) & -a_{n-1}(t) & \dots & -a_2(t) & -a_1(t) \end{pmatrix}, \; b(t) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f(t) \end{pmatrix} \;\text{ y }\; \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}$$

4.5. Ejemplos

Ejemplo 4.5.1. Resolver el sistema

$$\begin{cases} x_1' = & -3x_2 + x_3 + 2x_4 \\ x_2' = & -2x_1 + x_2 - x_3 + 2x_4 \\ x_3' = & -2x_1 + x_2 - x_3 + 2x_4 \\ x_4' = & -2x_1 - 3x_2 + x_3 + 4x_4 \end{cases} \tag{4.5.1}$$

con la condición inicial $x_1(0) = 1,\ x_2(0) = 0,\ x_3(0) = -1,\ x_4(0) = 0.$

Si escribimos en forma matricial el sistema (4.5.1), nos queda

$$\mathbf{x}' = A\mathbf{x}, \quad x(0) = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \text{ siendo } \mathbf{x}' = \begin{pmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \end{pmatrix}, \; A = \begin{pmatrix} 0 & -3 & 1 & 2 \\ -2 & 1 & -1 & 2 \\ -2 & 1 & -1 & 2 \\ -2 & -3 & 1 & 4 \end{pmatrix} \;\text{ y }\; \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \tag{4.5.2}$$

Sabemos por el teorema 4.2.1, que la única solución es $\varphi(t) = e^{tA}x(0)$, para todo $t \in \mathbb{R}$. Asi que el problema se reduce a calcular la matriz exponencial e^{tA}. Pero en el ejemplo 3.8.2 ya realizamos este cálculo, y habíamos obtenido

$$e^{tA} = \begin{pmatrix} 1 & -2e^{2t} + t + 2 & e^{2t} - t - 1 & e^{2t} - 1 \\ 1 - e^{2t} & t + 1 & -t & e^{2t} - 1 \\ 1 - e^{2t} & t & 1 - t & e^{2t} - 1 \\ 1 - e^{2t} & -2e^{2t} + t + 2 & e^{2t} - t - 1 & 2e^{2t} - 1 \end{pmatrix}, \text{ para todo } t \in \mathbb{R}.$$

Luego:

$$\varphi(t) = \begin{pmatrix} 1 & -2e^{2t} + t + 2 & e^{2t} - t - 1 & e^{2t} - 1 \\ 1 - e^{2t} & t + 1 & -t & e^{2t} - 1 \\ 1 - e^{2t} & t & 1 - t & e^{2t} - 1 \\ 1 - e^{2t} & -2e^{2t} + t + 2 & e^{2t} - t - 1 & 2e^{2t} - 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} t - e^{2t} + 2 \\ t - e^{2t} + 1 \\ t - e^{2t} \\ t - 2e^{2t} + 2 \end{pmatrix}, \text{ para todo } t \in \mathbb{R}.$$

Concluimos que la solución del sistema (4.5.1) es:

$$x_1(t) = t - e^{2t} + 2,\ x_2(t) = t - e^{2t} + 1,\ x_3(t) = t - e^{2t},\ x_4(t) = t - 2e^{2t} + 2$$

para todo $t \in \mathbb{R}$. $\square$

Ejemplo 4.5.2. Finalizando, vamos a resolver aplicando lo comentado anteriormente, la ecuación lineal homogénea:

$$(E): \quad x''' - 5x'' + 3x' + 9x = 0, \quad \text{con las condiciones iniciales} \quad x(0) = x'(0) = 0, x''(0) = 1$$

y haciendo, $x = y, x' = z$ y $x'' = w$ nos queda que (E) es equivalente al sistema lineal

$$(S): \quad \begin{cases} y' = z \\ z' = w \\ w' = -9y - 3z + 5w \end{cases} .$$

En resumen, tenemos que resolver el problema de los valores iniciales

$$\mathbf{x}' = A\mathbf{x}, \quad x(0) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \text{ siendo } \mathbf{x}' = \begin{pmatrix} y' \\ z' \\ w' \end{pmatrix}, \ A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -9 & -3 & 5 \end{pmatrix} \ \text{ y } \ \mathbf{x} = \begin{pmatrix} y \\ z \\ w \end{pmatrix}$$

que de acuerdo al teorema 4.2.1, la solución es $\phi(t) = e^{tA}x_0$, para todo $t \in \mathbb{R}$. Asi que el problema se reduce a calcular la matriz exponencial e^{tA}, donde el procedimiento es hallar una matriz J, semejante a la matriz A, donde J es una matriz diagonal, si A es diagonalizable, y J es una forma canónica de Jordan, si A no es diagonalizable, por lo cual $A = PJP^{-1}$, donde las columnas de la matriz P, son los elementos de una base de $\mathbb{R}^3$, como ya hemos comentado en el capítulo 1. Entonces $e^{tA} = Pe^{tJ}P$ donde el cálculo de e^{tJ} es más manejable que el de e^{tA}.
Entonces, operando tenemos que $\mathcal{X}_A(t) = \det(A - tI) = -(t+1)(t-3)^2$, por lo tanto los valores propios de A son -1 y 3, y

$$\begin{aligned} &\ker(A + I) = [(-1, 1, -1)] \text{ entonces } MG(-1) = 1 = MA(-1) \\ &\ker(A - 3I) = [(1, 3, 9)] \text{ entonces } MG(-1) = 1 \neq 2 = MA(-1) \end{aligned}$$

por lo cual concluimos que A no es diagonalizable, y como el polinomio característio se escinde, entonces A admite una forma canónica de Jordan, que tiene que ser necesariamente de la forma: $J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$, y la base de Jordan tiene la forma

$$\mathcal{B} = \{u, (A - 3I)v, v\}$$

donde elegimos $u = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$. Consideremos ahora

$$\begin{aligned} &\ker(A - 3I) = [(1, 3, 9)] \\ &\ker(A - 3I)^2 = [(1, 0, -9), (0, 1, 6)] \end{aligned}$$

Entonces elegimos $v = (0,1,6) \in \ker(A-3I)^2 \setminus \ker(A-3I)$. Luego $(A-3I)v = (1,3,9)$. Entonces una base de Jordan es $\mathcal{B} = \{(-1,1,-1),(1,3,9),(0,1,6)\}$. Entonces

$$A = PJP^{-1}, \text{ donde } J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}, \text{ y } P = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 3 & 1 \\ -1 & 9 & 6 \end{pmatrix}.$$

Calculando primero e^{tJ}, usando (3.7.2) y (3.7.1), nos queda

$$e^{tJ} = \begin{pmatrix} e^{-t} & 0 & 0 \\ 0 & e^{3t} & te^{3t} \\ 0 & 0 & e^{3t} \end{pmatrix}$$

Entonces

$$e^{tA} = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 3 & 1 \\ -1 & 9 & 6 \end{pmatrix} \begin{pmatrix} e^{-t} & 0 & 0 \\ 0 & e^{3t} & te^{3t} \\ 0 & 0 & e^{3t} \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ 1 & 3 & 1 \\ -1 & 9 & 6 \end{pmatrix}^{-1}.$$

Y operando llegamos a la solución

$$\phi(t) = e^{tA}x_0 = e^{tA}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} e^{3t}(\frac{t}{4} - \frac{1}{16}) + \frac{e^{-t}}{16} \\ \frac{e^{3t}(12t+1)}{16} - \frac{e^{-t}}{16} \\ \frac{3e^{3t}(12t+5)}{16} + \frac{e^{-t}}{16} \end{pmatrix} \text{ para todo } t \in \mathbb{R}.$$

Finalmente concluimos que la solución de (E) es es la primera componente de acuerdo a lo comentado anteriormente:

$$\phi(t) = e^{3t}\left(\frac{t}{4} - \frac{1}{16}\right) + \frac{e^{-t}}{16} = \frac{e^{-t}}{16}\left(e^{4t}(4t-1)+1\right) \text{ para todo } t \in \mathbb{R}. \quad \square$$

Bibliografía

[Aa] Abella, A. *Notas para el curso de Álgebra Lineal II*(2017)Centro de Matemática.Facultad de Ciencias. Universidad de la República

[Ap] Apostol, T.M. *Calculus, Volumen 2*(1973) Ed.Reverté,S.A.

[Bl] Brand,L. *Advanced Calculus. An introduction to Classical Analysis*(1967)Dover Publications,Inc. Mineola, New York

[CC] Cotlar, M. y Cignoli, R. *An Introduction to Functional Analysis*(1974) North-Holland Publishing Company

[Dc] Deschamps C., Moulin,F. *Maths Psi-Psi**(2009) Dunod.

[Fr] Friedberg, S. M., Insel,A. J. y Spence.L. E. *Linear Algebra*(2002). Ed. Prentice Hall.

[Horn] Horn, R.A., Johnson, C.R. *Matrix Analysis*(1991). Cambridge University Press.

[Ll] Lages Lima, E. *Espacos metricos*(1993) Impa. Rio de Janeiro

[JL] web.stanford.edu/ jluk/math63CMspring17/Matrices.170424.pdf

[Mt] www.ehu.eus/mtpalezp/libros/anafun2.pdf

[Na] Nachbin,L. *Introdução á análise funcional: Espaços de Banach e cálculo diferencial* (1976) publicación de Estados Americanos, Washington D.F.

[Po] Poole,D. *Linear Algebra. A Modern Introduction*(2006).Thomson.Brooks/Cole

[Re] fernandorevilla.es/blog/2015/06/08/espacios-normados-de-dimension-finita/