Primera edición, 1996

ISBN 968-16-5040-9

Impreso en México

ÍNDICE GENERAL

INTRODUCCIÓN

El ÁLGEBRA LINEAL es una parte esencial de las matemáticas tanto por su contenido conceptual como por sus amplias aplicaciones a otras áreas de las matemáticas y a otras ciencias, en especial, a la física, la biología, la economía y a las ciencias sociales.

En matemáticas, el álgebra lineal forma parte de los cursos básicos que se imparten en todo el mundo. En general, en estos cursos se desarrollan los conceptos fundamentales de matriz, espacio vectorial y transformación lineal, valores propios y formas normales. Pero algunos temas fundamentales se dejan a un lado por falta de tiempo, y es así como muy pocas aplicaciones de la teoría pueden ser contempladas. El presente texto procura cubrir un poco esta laguna.

Este libro está dirigido a estudiantes que hayan llevado ya algún curso de álgebra lineal. En particular, se supone que se conocen los conceptos fundamentales antes mencionados. Nuestro propósito es estudiar con alguna amplitud algunos temas que no se consideran en los cursos básicos de licenciatura en matemáticas y que tienen una importante gama de aplicaciones tanto en otras ramas de las matemáticas como en otras ciencias. Creemos que todos los temas tratados son de carácter fundamental, es decir, que todo matemático debería conocer.

A pesar de que la mayor parte de los temas de este libro no son tratados en los cursos básicos, el material de este libro es elemental. El texto está dirigido a estudiantes que cursan la segunda mitad de una carrera de ciencias. Creemos que este libro puede ser leído por un estudiante que ha llevado los cursos introductorios de cálculo y álgebra lineal. También puede considerarse este libro como un texto para un curso de álgebra lineal avanzada, es decir, un curso de los últimos semestres de licenciatura o los primeros de maestría donde, a un ritmo más intenso, pueda cubrirse una cantidad importante de material contando con una mayor madurez por parte de los estudiantes.

Hemos hecho un especial énfasis en las aplicaciones de la teoría, presentamos aplicaciones a problemas de física, química, biología y economía. Creemos que este texto también puede ser leído con provecho por estudiantes de carreras científicas que estén interesados en aplicaciones de la teoría desarrolladas con el rigor propio de las matemáticas.

Mencionamos brevemente los principales contenidos de este libro. En el capítulo I desarrollamos la construcción de la forma canónica de Jordan de una matriz sobre un campo base arbitrario. Presentamos algunas aplicaciones teóricas de las formas de Jordan: convergencia de series de potencias de matrices, resolución de sistemas de ecuaciones diferenciales y otras. El capítulo II está dedicado al estudio de algunas aplicaciones del álgebra lineal en otras áreas. Se presenta el modelo de Leontief en economía, se resuelve el problema del péndulo múltiple en física. Estas aplicaciones no sólo requieren del material presentado en el primer capítulo, sino también de algunos resultados de los capítulos posteriores. Presentamos el material de esta manera para motivar algunos de los resultados del libro. En el capítulo III examinamos algunos resultados sobre la estimación de los valores propios de una matriz. Comenzamos con el clásico Teorema de Geršgorin y algunas aplicaciones; seguimos con estimaciones de valores propios de matrices unitarias, normales y hermitianas. En el capítulo IV consideramos matrices que dejan conos invariantes. Presentamos aquí la generalización de Birkhoff y Vandergraft del importante Teorema de Perron-Frobenius: una matriz irreducible con entradas no negativas tiene un vector propio con entradas positivas cuyo valor propio asociado es positivo. Como un caso especial, estudiamos matrices estocásticas y algunas de sus importantes aplicaciones. Por último, el capítulo V se ocupa de algunos elementos de la teoría espectral de gráficas. La matriz adyacente asociada a una gráfica tiene entradas no negativas, el conocimiento de los valores propios de esta matriz determina alguna información de la estructura combinatoria de la gráfica. En todos los capítulos presentamos aplicaciones. Una de las características útiles del texto es que cuenta con muchos ejercicios. A lo largo de la mayor parte de las secciones que forman el libro se encuentran ejercicios que complementan el texto o ilustran algunas aplicaciones. Creemos que es importante que el estudiante vaya resolviendo la mayor parte de estos ejercicios conforme avanza en el texto.

Una última consideración: hay en el mercado una gran cantidad de libros de texto sobre álgebra lineal, ¿para qué uno más? Por supuesto, la respuesta a esta interrogante no puedo darla en forma totalmente objetiva; sólo subrayaré algunas diferencias de éste con otros textos existentes. Hay excelentes libros de álgebra lineal que se ocupan de una gran cantidad de temas fundamentales; citaré dos de ellos: el de Gantmacher, y el de Horn y Johnson. Éstos son libros importantes, pero su contenido no puede ser cubierto en un curso estándar; por otra

parte, no presentan aplicaciones. Hay algunos libros que hacen especial énfasis en las aplicaciones, como el de Noble y Daniel, o el de Strang, pero no exponen adecuadamente la parte teórica. El libro de Huppert, además de la dificultad idiomática, es muy extenso. Hay también algunos excelentes libros que sólo tratan alguno de los temas trabajados en este texto: el de Minc y el de Berman-Neumann-Stern tratan matrices no negativas, el de Cvetković-Doob-Sachs trata la teoría espectral de gráficas. El presente texto pretende estar a medio camino entre todas estas obras: en primer lugar es un libro para un curso de un semestre; en segundo lugar, explica la teoría requerida en forma completa y rigurosa, presentando también múltiples aplicaciones y ejercicios; en tercer lugar, no se especializa en un solo tema, intenta más bien presentar una serie de temas y sólo los elementos de aplicaciones más profundas a las que el lector podrá acceder posteriormente en otros libros o artículos especializados. Finalmente, algunos de los temas tratados no han aparecido antes en forma de texto. Por ejemplo, algunos de los resultados sobre M-matrices del capítulo II, el Teorema de Birkhoff-Vandergraft del capítulo IV, y algunas consideraciones sobre los diagramas de Dynkin en el capítulo V han sido elaborados a partir de los artículos originales. Por otro lado, algunos temas considerados en el texto siguen el tratamiento dado en los libros antes citados: el Teorema de Perron-Frobenius sigue el tratamiento de Gantmacher; varios ejemplos (matrices estocásticas, el péndulo múltiple) siguen a Huppert; los teoremas de entrelazamiento del capítulo III se basan en Horn y Johnson; algunas partes del capítulo V siguen a Cvetković-Doob-Sachs.

El presente texto es el resultado de las notas de un curso semestral de álgebra lineal avanzada que he impartido varias veces: en 1988 y 1992 en la Facultad de Ciencias de la UNAM y en 1993 en la Universidad de Zürich en Suiza. En octubre de 1992, la escritura de este libro fue apoyada con una Cátedra de Excelencia del Consejo Nacional de Ciencia y Tecnología. Agradezco aquí a todos los estudiantes de mis cursos sus comentarios e interés y a las instituciones el haberme permitido poner en marcha el curso y su consecuencia: este libro. Agradezco a Martha Takane la realización del trabajo que permitió el esclarecimiento de algunos puntos del Teorema de Birkhoff-Vandergraft, así como algunos otros en el capítulo V. A Mary Glazman algunos comentarios acerca del modelo de Leontief. A mis estudiantes Edith Adán y María Elena Gallardo agradezco la cuidadosa lectura de versiones previas del libro. Finalmente, agradezco a Gabriela Sanginés la laboriosa mecanografía

del texto. Como siempre en estos casos, todos los errores que permanezcan en el libro son culpa y responsabilidad única del autor.

Notación y algunas convenciones

Por k denotaremos siempre un campo. En varios de los capítulos consideraremos sólo el caso en que k es el campo de los números reales $\mathbf{R}$ o el campo de los números complejos $\mathbf{C}$.

Sea V un k-espacio vectorial de dimensión finita. El *espacio dual* $V^* = \mathrm{Hom}_k(V,k)$ tiene la misma dimensión que V. Un homomorfismo de espacios vectoriales es una *transformación lineal.*

Los vectores $v \in k^n$ serán considerados columnas con entradas v_i, $i = 1,\ldots,n$. La *base canónica* $e_1,\ldots,e_n$ de k^n está definida de forma que la entrada j-ésima de e_i es $e_{ij} = \delta_{ij}$, donde δ_{ij} denota la delta de Kronecker.

Si n es un número natural, el conjunto de matrices sobre k de tamaño $n \times n$ se denota por $M_n(k)$. La matriz identidad en este conjunto se denota I_n.

Dada una matriz $A = (a_{ij})$ en $M_n(k)$, se denota por A^T su *matriz transpuesta.* El *rango* de A se denota $\mathrm{rg}\, A$.

El *polinomio característico* χ_A de una matriz $A \in M_n(k)$ tiene grado n. Los *valores propios* de A son raíces de χ_A. En ocasiones llamamos a la lista de raíces $\lambda_1,\ldots,\lambda_n$ de χ_A, en la cerradura algebraica de k, los valores propios de A. Observe que en este conjunto de valores propios puede haber repeticiones.

Sea $A = (a_{ij}) \in M_n(k)$ y consideremos dos series de números $1 \leq i_1 < i_2 < \cdots < i_m \leq n$ y $1 \leq j_1 < j_2 < \cdots < j_m \leq n$, entonces $A\begin{pmatrix} i_1 & i_2 & \ldots & i_m \\ j_1 & j_2 & \ldots & j_m \end{pmatrix}$ denota la matriz $m \times m$ cuya entrada (s,t) es $a_{i_s j_t}$. Si B se forma eliminando el renglón i y la columna j de A, escribiremos $B = A^{(i,j)}$. La *matriz adjunta* $\mathrm{ad}\, A = (b_{ij})$ de A tiene entradas $b_{ij} = (-1)^{i+j} \det A^{(i,j)}$. Recordamos que $A.\,\mathrm{ad}\, A = \mathrm{ad}\, A.A = \det A.I_n$.

I. LA FORMA CANÓNICA DE JORDAN

Introducción

En este capítulo k denota un campo.

De nuestros estudios de álgebra lineal elemental sabemos que la representación de una transformación lineal del espacio k^n en sí mismo depende de la base elegida. De manera que si consideramos dos bases diferentes para este espacio, la transformación dada tendrá una representación diferente. ¡Pero esencialmente se trata de la misma matriz! Diremos que dos matrices A y B que representan la misma transformación lineal en diferentes bases del espacio son similares. La similaridad de matrices determina una relación de equivalencia en el conjunto de matrices cuadradas $M_n(k)$. Por tanto, $M_n(k)$ se parte en clases de equivalencia. Las matrices en una misma clase de equivalencia comparten propiedades importantes: tienen la misma traza, el mismo determinante y los mismos valores propios. Pero el compartir estas propiedades no hace a dos matrices similares. Por ejemplo, las matrices

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

comparten todas las propiedades consideradas, pero no son similares (en efecto, la única matriz similar a la matriz 0 es ella misma).

Una idea esencial es tratar de encontrar un representante sencillo de cada clase de equivalencia en $M_n(k)$. Si tenemos un sistema de representantes, para saber si dos matrices dadas son similares bastaría con comparar los representantes de sus clases de equivalencia. Surgen por supuesto aquí varios problemas: en primer lugar, ¿qué quiere decir "sencillo"?; en segundo lugar, tendremos que ver si para una matriz dada es posible calcular el representante de su clase de equivalencia. La solución de este problema que expondremos en este capítulo son las *formas canónicas de Jordan.*

Las formas canónicas de Jordan son un conjunto de matrices "casi diagonales" que permiten establecer las propiedades fundamentales de la clase de equivalencia correspondiente en $M_n(k)$ en un vistazo. La construcción de las formas de Jordan será explicada en este capítulo en dos partes. Primeramente consideraremos el caso en que el campo base

k es algebraicamente cerrado. Luego consideraremos el caso general. Por supuesto, la prueba del caso general vale para el caso algebraicamente cerrado; este caso se considera separadamente por razones didácticas: algunos argumentos son más claros y la forma de las matrices obtenidas es mejor conocida.

En las últimas secciones del capítulo examinaremos algunas aplicaciones "teóricas" del teorema de la base de Jordan. Entre otras cosas, consideraremos potencias de matrices y convergencia de series de potencias de matrices. Resolveremos sistemas clásicos de ecuaciones diferenciales y desarrollaremos los criterios de estabilidad de Lyapunov. En el capítulo siguiente, veremos algunas aplicaciones "prácticas" de los resultados de este primer capítulo, es decir, consideraremos problemas provenientes de otras áreas, como la física, la economía y la biología.

1. Similaridad: un concepto básico

1.1. Una matriz A se dice *similar* a la matriz B (y escribimos $A \sim B$) en caso de que exista una matriz invertible P tal que $A = PBP^{-1}$. Es fácil ver que $\sim$ es una relación de equivalencia.

Hay *propiedades* de las matrices que son *invariantes* bajo similaridad, esto es, si $A \sim B$ y A tiene la propiedad entonces B también la tiene; por ejemplo, si $A \sim B$ y A es invertible entonces B es también invertible.

Un conjunto M de matrices se dice *invariante por similaridad* si cada vez que $A \sim B$ y $A \in M$, también se tiene $B \in M$. Por ejemplo, el conjunto de las matrices reales X que satisfacen la ecuación

$$2X^2 - 1 = 0$$

es invariante por similaridad.

Una función f definida en matrices se llama *invariante por similaridad* (o simplemente *invariante*) si $f(A) = f(B)$ para cualquier par de matrices similares $A \sim B$. Por ejemplo, el determinante y la traza son funciones invariantes.

Ejercicios

1. Demuestre que la similaridad $\sim$ es una relación de equivalencia.

2.a) Si I_n es la matriz identidad $n \times n$, muestre que λI_n es la única matriz similar a λI_n, para cualquier $\lambda \in k$.

b) Si A es la única matriz similar a A, demuestre que $A = \lambda I_n$ para alguna $\lambda \in k$.

3. Demuestre que si $A \sim B$, entonces A y B tienen los mismos valores propios.

4. Sea $g(t) \in k[t]$ un polinomio con $g(t) = a_0 + a_1 + \cdots + a_s t^s$. Para cualquier matriz A, definimos la matriz

$$g(A) = a_0 I_n + a_1 A + \cdots + a_s A^s.$$

Demuestre que si $A = PBP^{-1}$, entonces $g(A) = Pg(B)P^{-1}$. O sea, $A \sim B$ implica $g(A) \sim g(B)$.

5. Demuestre que tanto el determinante como la traza son funciones invariantes.

6. Una matriz A se llama *nilpotente* si $A^s = 0$ para alguna $s \geq 1$. Muestre que la nilpotencia es una propiedad invariante por similaridad.

1.2. Una forma general de intentar resolver un problema es la siguiente: transformamos nuestro problema en uno equivalente pero más sencillo de resolver, lo resolvemos en esta forma equivalente y luego vemos qué nos dice esta solución sobre el problema original. Es de esta forma que se usa la similaridad para resolver problemas. Veamos un ejemplo:

Se desea calcular el valor de las entradas de la matriz A^{20}, donde

$$A = \begin{bmatrix} 6 & 2 \\ 3 & 3 \end{bmatrix}$$

en el campo base $\mathbf{Z}_7$.

Paso 1: Encontramos una matriz $B \sim A$, pero más sencilla. Por ejemplo, para $P = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$ se tiene

$$B = P^{-1}AP = \begin{bmatrix} 6 & 1 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 6 & 2 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

Paso 2: Se resuelve el problema (más sencillo) para B.

En el ejemplo,

$$B^{20} = \begin{bmatrix} 3^{20} & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

Paso 3: Se resuelve el problema original con base en la solución obtenida para B. Así,

$$A^{20} = (PBP^{-1})^{20} = PB^{20}P^{-1} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & 1 \\ 2 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 5 & 3 \end{bmatrix}.$$

Por supuesto, el primer problema al usar este procedimiento es que no es evidente (y a veces tampoco fácil) cómo elegir la matriz B conveniente. Puede suceder que no exista una matriz B que haga más sencillo el problema. En nuestro ejemplo el problema se simplificaba porque la matriz B era diagonal.

1.3. Una matriz A se llama *diagonalizable* si existe una matriz diagonal B que es similar a A.

Si $A \sim B$ y B es una matriz diagonal, probablemente nos sea útil B para aplicar el proceso expuesto en (1.2). Pero, ¿cómo sabemos si A es diagonalizable?, y si lo es, ¿cómo calculamos una matriz diagonal B similar a A?

En primer lugar observemos que no todas las matrices son diagonalizables. En efecto, la matriz

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

no es diagonalizable. Supongamos que $A \sim B$, entonces A y B tienen los mismos valores propios. Siendo $(x-1)^2$ el polinomio característico de A, entonces si B es diagonal, sólo puede ser

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Además, como $A = PBP^{-1} = PP^{-1} = I_2$, entonces $1 = 0$, lo que es una contradicción.

Sin embargo, hay algunos casos donde los valores propios nos indican si una matriz es diagonalizable.

Lema. *Sea A una matriz $n \times n$. Entonces A es diagonalizable si y sólo si k^n tiene una base formada por vectores propios de A.*

Demostración: Supongamos que $\{v_1, \ldots, v_n\}$ forma una base de k^n y que $Av_i = \lambda_i v_i$ para alguna $\lambda_i \in k$, $i = 1, \ldots, n$.

Formemos la matriz P cuya i-ésima columna es el vector v_i. Por supuesto P es invertible.

Luego,

$$P^{-1}AP = P^{-1}[Av_1, \ldots, Av_n]$$

$$= P^{-1}[\lambda_1 v_1, \ldots, \lambda_n v_n] = P^{-1}P\Lambda = \Lambda,$$

donde Λ es la matriz

$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

(*notación*: $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$).

Para el converso, supongamos que $P^{-1}AP = \Lambda$ y Λ es una matriz diagonal. Entonces $AP = P\Lambda$. Si llamamos v_i a la i-ésima columna de P, tenemos $Av_i = \lambda_i v_i$, donde $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Claramente, los vectores $v_1, \ldots, v_n$ forman una base de k^n. □

1.4. Proposición. *Si A tiene todos sus valores propios diferentes, entonces A es diagonalizable.*

Demostración: Supongamos que $\lambda_1, \ldots, \lambda_n$ son los valores propios de A y que $\lambda_i \neq \lambda_j$ si $i \neq j$. Para cada $i = 1, \ldots, n$, elegimos un vector $0 \neq v_i$ con $Av_i = \lambda_i v_i$. Deseamos demostrar que los vectores $v_1, \ldots, v_n$ son linealmente independientes. Supongamos que no lo son, es decir, que hay una combinación lineal $\sum_{i=1}^{n} \mu_i v_i = 0$, donde no todas las $\mu_i \in k$ son cero. Renumerando los vectores v_i, podemos suponer que $\sum_{i=1}^{s} \mu_i v_i = 0$ con $2 \leq s \leq n$ y $\mu_i \in k \setminus \{0\}$; además, podemos elegir s mínima posible. Tenemos entonces

$$0 = A\left(\sum_{i=1}^{s} \mu_i v_i\right) = \sum_{i=1}^{s} \mu_i A v_i = \sum_{i=1}^{s} \mu_i \lambda_i v_i.$$

Luego,

$$\sum_{i=1}^{s-1} \mu_i(\lambda_i - \lambda_s)v_i = 0,$$

y por ser s mínima posible, se sigue que $\mu_1(\lambda_1 - \lambda_s) = 0$ y que $\lambda_1 = \lambda_s$. Y ello es una contradicción. □

Ejercicios

1. Demuestre que si A es una matriz diagonalizable de tamaño $n \times n$ y

$$\chi_A(t) = \det(tI_n - A)$$

es el *polinomio característico* de A, entonces $\chi_A(A) = 0$. (Éste es un caso especial del Teorema de Cayley-Hamilton, que estudiaremos más adelante.)

2. Sea A una matriz 2×2 con coeficientes reales. Supongamos que A es similar a la matriz diagonal $B = \operatorname{diag}(\lambda, \mu)$ y que $|\lambda|, |\mu| < 1$. Calcule entonces $\lim_{n\to\infty} A^n$ (¿existe?).

3. Muestre que el converso de (1.4) es falso.

4. Supongamos que A es diagonalizable y sean $\lambda_1, \ldots, \lambda_n$ sus valores propios. Fijamos un vector $x^{(0)}$ en k^n y definimos

$$Ax^{(s)} = x^{(s+1)}$$

para todo $s \geq 0$. Demuestre que hay escalares $\alpha_1, \ldots, \alpha_n$ tales que $x_1^{(s)} = \sum_{j=1}^{n} \alpha_j \lambda_j^s$, para toda $s \geq 0$ (lo mismo para los otros índices $i \neq 1$).

1.5. Sea V un k-espacio vectorial. Sea $\varphi: V \to V$ una transformación lineal. Dada una base $B = \{v_1, \ldots, v_n\}$ de V, podemos construir la *matriz de representación* $A(\varphi, B)$ de φ respecto a la base B: para ello, consideramos la imagen de cada básico bajo φ, $\varphi(v_i) = \sum_{j=1}^{n} a_{ji} v_j$ (cuidado con el orden de los índices); entonces $A = A(\varphi, B)$ es la matriz $n \times n$, $A = (a_{ji})$.

La matriz de representación $A(\varphi, B)$ satisface la siguiente propiedad: dado un vector $z \in V$, consideramos las descomposiciones:

$$z = \sum_{i=1}^{n} z_i^B v_i \quad \text{y} \quad \varphi(z) = \sum_{i=1}^{n} \varphi(z)_i^B v_i, \quad \text{con } z_i^B, \varphi(z)_i^B \in k,$$

tenemos

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ji} z_i^B \right) v_j = \sum_{i=1}^{n} z_i^B \varphi(v_i) = \varphi(z) = \sum_{i=1}^{n} \varphi(z)_j^B v_j,$$

lo que implica que $\sum_{i=1}^{n} a_{ji} z_j^B = \varphi(z)_j^B$ para cada $j = 1, \ldots, n$.

De manera equivalente, $Az^B = \varphi(z)^B$, donde z^B es el vector con entradas z_i^B.

Conversamente, dada una matriz $A \in M_n(k)$, podemos construir un morfismo lineal $\varphi: k^n \to k^n$ de forma que $A(\varphi, B_0) = A$, donde B_0 es la base canónica. En efecto, si $z = (z_i) \in k^n$, definimos $\varphi(z) = Az \in k^n$.

Dadas dos bases B, B' de V, se puede definir una matriz $n \times n$ invertible T tal que $T(v^{B'}) = v^B$ para cada $v \in V$. La matriz T se llama *matriz de cambio de base*. Para ello, llamamos $B = \{v_1, \ldots, v_n\}$ y $B' = \{w_1, \ldots, w_n\}$. Construimos los n vectores $w_i^B \in k^n$. Definimos T como la matriz cuyas columnas son los vectores w_i^B (*notación*: $T = [w_1^B, \ldots, w_n^B]$). Para verificar la propiedad deseada, es suficiente ver que $T(w_i^{B'}) = w_i^B$, pero esto se cumple inmediatamente ya que $w_i^{B'}$ es el i-ésimo vector e_i de la base canónica de k^n.

Consideremos otra vez el morfismo lineal $\varphi: V \to V$. Tenemos,

$$A(\varphi, B)Tv^{B'} = A(\varphi, B)v^B = \varphi(v)^B = T\varphi(v)^{B'} = TA(\varphi, B')v^{B'},$$

para cada $v \in V$. En forma equivalente, $A(\varphi, B) = TA(\varphi, B')T^{-1}$, es decir, las matrices $A(\varphi, B)$ y $A(\varphi, B')$ son similares.

Ejercicios

1. Considere la transformación lineal $\varphi: \mathbf{R}^3 \to \mathbf{R}^3$, que envía los vectores $(1,0,0), (0,1,0)$ y $(0,0,1)$ en los vectores $(1,1,0), (1,2,1)$ y $(0,2,1)$, respectivamente. Calcule la matriz de representación $A(\varphi, B)$ con respecto a la base B, donde
   a) $B = \{(1,0,1), (1,1,1), (2,1,1)\}$,
   b) $B = \{(1,0,1), (0,0,1), (3,2,1)\}$.

2. Sea V el k-espacio vectorial formado por todos los polinomios en la variable t y de grado menor o igual a 4. Considere la transformación lineal $\varphi: V \to V$ dada por la derivación formal de polinomios. Calcule la matriz de representación $A(\varphi, B)$ en la base canónica de V.

2. La forma de Jordan sobre campos algebraicamente cerrados

Como hemos visto en (1.3), no todas las matrices son diagonalizables (si lo fueran, no habría casi nada que hacer en la teoría de matrices). Sin embargo, muchos problemas se pueden solucionar en una forma parecida a la planteada en (1.2), aunque no tengamos matrices diagonales a la mano; nos basta contar con matrices suficientemente sencillas.

Afortunadamente, siempre nos será posible tener a mano las matrices de Jordan. Para construirlas, procederemos en dos pasos; primeramente en esta sección supondremos que el campo base k es *algebraicamente cerrado*; por ejemplo, $k = \mathbf{C}$. En la siguiente sección consideraremos el caso general.

2.1. Un *bloque de Jordan* $J_n(\lambda)$ es una matriz $n \times n$ de la forma

$$\begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix};$$

esto es: (1) en la diagonal principal $J_n(\lambda)$ tiene valor λ; (2) en la diagonal por encima de la diagonal principal tiene valor 1; (3) en todas las demás entradas tiene ceros.

El único valor propio de $J_n(\lambda)$ es λ. Podemos escribir también $J_n(\lambda) = \lambda I_n + J_n(0)$.

Una *matriz de Jordan* se construye por medio de bloques de Jordan colocados en forma diagonal, como sigue:

$$\begin{bmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & & \\ & 0 & \ddots & 0 \\ 0 & \cdots & 0 & J_{n_s}(\lambda_s) \end{bmatrix}.$$

Esta matriz también se denota por $J_{n_1}(\lambda_1) \oplus J_{n_2}(\lambda_2) \oplus \cdots \oplus J_{n_s}(\lambda_s)$.

2.2. Teorema (Jordan). *Sea V un k-espacio vectorial de dimensión n. Sea $\varphi: V \to V$ una transformación lineal. Entonces existe una base B de V tal que la matriz de representación $A(\varphi, B)$ es una matriz de Jordan.*

Una base como la dada en el teorema se llama *base de Jordan* para φ. La matriz de representación correspondiente se llama una *forma de Jordan* de la transformación φ.

Puede haber muchas bases de Jordan para la misma transformación φ. Por ejemplo, si $\varphi = 1_V$, en cualquier base B de V, la matriz de representación $A(\varphi, B) = I_n$, que es una forma de Jordan. (En efecto, $I_n = J_1(1) \oplus \cdots \oplus J_1(1)$.)

Dividiremos la demostración del teorema en varios pasos:

(1) Existe un polinomio $\xi(t) \in k[t]$ tal que $\xi(\varphi) = 0$.

(2) Sea $\xi(t)$ un polinomio con $\xi(\varphi) = 0$. Como k es algebraicamente cerrado, $\xi(t) = (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2} \cdots (t-\lambda_s)^{n_s}$ con $\lambda_i \neq \lambda_j$ para $i \neq j$. Probaremos que:

(a) $V = \ker(\varphi - \lambda_1)^{n_1} \oplus \cdots \oplus \ker(\varphi - \lambda_s)^{n_s}$.

(b) Cada espacio $V_i = \ker(\varphi - \lambda_i)^{n_i}$ es φ invariante, esto es, $\varphi(V_i) \subset V_i$.

(3) Si φ es nilpotente, entonces existe una base B tal que la matriz de representación $A(\varphi, B) = J_{m_1}(0) \oplus \cdots \oplus J_{m_t}(0)$ para algunos números $m_1, \ldots, m_t$.

Probaremos en este momento que de (1), (2) y (3) se sigue el teorema. En efecto, sean $V_i = \ker(\varphi - \lambda_i)^{n_i}$ los espacios definidos antes. Para cualesquiera elecciones de bases B_i de V_i, se tiene que $B = \bigcup_{i=1}^{s} B_i$ es base de V. Como $\varphi(V_i) \subset V_i$, resulta que la matriz de representación $A = A(\varphi, B)$ tiene la forma $A = A_1 \oplus \cdots \oplus A_s$, donde cada $A_i = A(\varphi_i, B_i)$ es la matriz de representación del operador $\varphi_i = \varphi|: V_i \to V_i$.

Por otra parte, el operador $\varphi_i - \lambda_i 1_{V_i}$ es nilpotente (en efecto, $(\varphi_i - \lambda_i 1_{V_i})^{n_i} = 0$). Luego, por (3), podemos elegir la base B_i de V_i de forma que $A(\varphi_i - \lambda_i 1_{V_i}, B_i) = J_{i,m_1}(0) \oplus \cdots \oplus J_{i,m_{t(i)}}(0)$. Entonces $A_i = A(\varphi_i, B_i) = J_{i,m_1}(\lambda_i) \oplus \cdots \oplus J_{i,m_{t(\lambda)}}(\lambda_i)$ y el resultado se sigue. □

En los siguientes párrafos demostraremos las afirmaciones (1), (2) y (3), y con ello el Teorema de Jordan.

2.3. La primera afirmación es muy sencilla. Basta con ver que el espacio de todas las transformaciones lineales $\mathrm{Hom}_k(V, V)$ tiene dimensión finita n^2. Luego, las transformaciones $1_V, \varphi, \varphi^2, \ldots, \varphi^{n^2}$ no pueden ser linealmente independientes. Si $\sum_{i=0}^{n^2} a_i \varphi^i = 0$ es una combinación no trivial de ellos, tenemos que para $0 \neq \xi(t) = \sum_{i=0}^{n^2} a_i t^i \in k[t]$, $\xi(\varphi) = 0$.

Como una aplicación del Teorema de Jordan probaremos después que $\xi(t)$ se puede elegir como el polinomio característico $\chi_\varphi(t)$.

2.4. Procedamos ahora a demostrar la parte (2). Esto es un ejercicio largo pero bastante sencillo.

En primer lugar, observemos que los polinomios

$$p(t) = (t-\lambda_1)^{n_1} \quad \text{y} \quad q(t) = (t-\lambda_2)^{n_2}\cdots(t-\lambda_s)^{n_s}$$

no tienen divisores comunes. Entonces existen polinomios $a(t), b(t) \in k[t]$ tales que $a(t)p(t)+b(t)q(t) = 1$. [Recordemos que esto se debe a que el anillo $k[t]$ es de ideales principales; en efecto, el ideal $p(t)k[t]+q(t)k[t]$ generado por $p(t)$ y $q(t)$ también es generado por un polinomio $c(t)$, que divide tanto a $p(t)$ como a $q(t)$. Luego $c(t)$ debe tener grado 0 y podemos elegirlo como $c(t) = 1$.]

De aquí obtenemos $a(\varphi)p(\varphi) + b(\varphi)q(\varphi) = \mathbf{1}_V$, y multiplicando por un vector arbitrario $v \in V$,

$$v = p(\varphi)(a(\varphi)v) + q(\varphi)(b(\varphi)v).$$

En particular, si tomamos $v \in \ker p(\varphi)$, tenemos $v \in \operatorname{Im} q(\varphi)$. Conversamente, si $v = q(\varphi)(w) \in \operatorname{Im} q(\varphi)$, entonces $p(\varphi)(v) = \xi(\varphi)(w) = 0$ y se sigue que $\ker p(\varphi) = \operatorname{Im} q(\varphi)$. Similarmente, $\ker q(\varphi) = \operatorname{Im} p(\varphi)$ y de aquí se obtiene que $V = \ker p(\varphi) + \ker q(\varphi)$ y $\ker p(\varphi) \cap \ker q(\varphi) = 0$, o en otras palabras $V = \ker p(\varphi) \oplus \ker q(\varphi) = \ker(\varphi-\lambda_1)^{n_1} \oplus \ker q(\varphi)$.

Por otra parte, está claro que $\varphi(\ker q(\varphi)) \subset \ker q(\varphi)$ y para la restricción $\varphi' = \varphi|\colon \ker q(\varphi) \to \ker q(\varphi)$, el polinomio $q(t)$ satisface $q(\varphi') = 0$. Luego, por hipótesis de inducción, obtenemos (a).

2.5. Aún tenemos que demostrar la afirmación (3). Esto es, dado $\varphi\colon V \to V$ lineal y nilpotente, digamos $\varphi^{s+1} = 0$, debemos construir una base de Jordan para φ. Consideremos las siguientes cadenas de subespacios de V:

$$\varphi(V) \supset \varphi^2(V) \supset \varphi^3(V) \supset \cdots \supset \varphi^s(V)$$

$$0 \neq V^0 = \ker\varphi \supset (V^0 \cap \varphi(V)) = V^1 \supset (V^0 \cap \varphi^2(V))$$

$$= V^2 \supset \cdots \supset (V^0 \cap \varphi^s(V)) = V^s.$$

Elijamos una base de V^s, $\{e_{s,1},\dots,e_{s,n_s}\}$. Completémosla a una base de V^{s-1}: $\{e_{s,1},\dots,e_{s,n_s},e_{s-1,1},\dots,e_{s-1,n_{s-1}}\}$. Continuamos el proceso hasta producir una base de $V^0 = \ker\varphi$ de la forma

$$\{e_{i,m}\colon i = 0,\dots,s, m = 1,\dots,n_i\}.$$

Como cada $e_{i,m}$ pertenece a $\varphi^i(V)$, podemos escribirlo como $e_{i,m} = \varphi^i(f_{i,m})$ para algún $f_{i,m} \in V$. Podemos visualizar el conjunto de vectores así construido de la siguiente manera:

$$\begin{array}{cccccccccc}
f_{s,1} & \cdots & f_{s,n_s} & & & & & & & \\
\varphi(f_{s,1}) & \cdots & \varphi(f_{s,n_s}) & f_{s-1,1} & \cdots & f_{s-1,n_{s-1}} & & & & \\
\varphi^2(f_{s,1}) & \cdots & \varphi^2(f_{s,n_s}) & \varphi(f_{s-1,1}) & \cdots & \varphi(f_{s-1,n_{s-1}}) & f_{s-2,1} & \cdots & f_{s-2,n_{s-2}} \\
 & & & & & \vdots & & & & \\
\varphi^{s-1}(f_{s,1}) & \cdots & \varphi^{s-1}(f_{s,n_s}) & \varphi^{s-2}(f_{s-1,1}) & \cdots & \varphi^{s-2}(f_{s-1,n_{s-1}}) & \cdots f_{1,1} & \cdots & f_{1,n_1} & \\
\varphi^s(f_{s,1}) & \cdots & \varphi^s(f_{s,n_s}) & \varphi^{s-1}(f_{s-1,1}) & \cdots & \varphi^{s-1}(f_{s-1,n_{s-1}}) & \cdots f_{0,1} & \cdots & f_{0,n_0} &
\end{array}$$

En este momento contamos ya con los ingredientes para construir nuestra base (de Jordan) de V.

Lema. *Para cada $i = 1,\dots,s$, el conjunto de vectores $\{\varphi^t(f_{j,m})$: con $j \le i$ o bien con $0 \le j-i \le t \le j$ y $m = 1,\dots,n_i\}$ constituye una base de* $\ker\varphi^{i+1}$.

Demostración: En primer lugar, observemos que cada vector $\varphi^t(f_{j,m})$ con $j \le i$ o bien con $j-i \le t \le j$ está en efecto en $\ker\varphi^{i+1}$.
Calculemos: si $0 \le j-i \le t$, se tiene

$$\varphi^{i+1}(\varphi^t(f_{j,m})) = \varphi^{i+1+t}(f_{j,m})$$
$$= \varphi^{i+t-j}\varphi\varphi^j(f_{j,m}) = \varphi^{i+t-j}\varphi(e_{j,m}) = 0;$$

si $j \le i$, se tiene

$$\varphi^{i+1}(\varphi^t(f_{j,m})) = \varphi^t\varphi^{i-j}\varphi(e_{j,m}) = 0.$$

Probaremos el resultado por inducción sobre i.

Para $i = 0$, hay que demostrar que $\{\varphi^j(f_{j,m})\colon j \ge 0\}$ constituye una base de $\ker\varphi$. Pero ésta fue la construcción que hicimos antes.

Supongamos demostrado el resultado para $i-1$. Deseamos probarlo para i. Sea $x \in \ker \varphi^{i+1}$.

Entonces $\varphi^i(x) \in V^0 \cap \varphi^i(V) = V^i$ y podemos escribirlo en forma única como $\varphi^i(x) = \sum_{j\geq i} \mu_{j,m} e_{j,m} = \sum_{j\geq i} \mu_{j,m}\varphi^j(f_{j,m})$. Luego el elemento

$$y = x - \sum_{j\geq i} \mu_{j,m}\varphi^{j-i}(f_{j,m})$$

pertenece a $\ker \varphi^i$. Por la hipótesis de inducción, podemos escribir $y = \sum \mu'_{j,m,t}\varphi^t(f_{j,m})$, para algunos coeficientes $\mu'_{j,m,t} \in k$, donde j y t satisfacen $j \leq i-1$ o bien $j+1-i \leq t \leq j$. De aquí se obtiene la expresión deseada. La unicidad de esta expresión puede probarse de forma muy sencilla. □

Ejercicios

1. Calcule una forma de Jordan para las matrices

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix}.$$

2. Demuestre que si los bloques de Jordan $J_n(\lambda)$ y $J_n(\mu)$ son similares, entonces $\lambda = \mu$.

3. Demuestre que $J_4(\lambda)$ no es similar a $J_2(\lambda) \oplus J_2(\lambda)$.

4. Para cada una de las matrices A del ejercicio 1, demuestre que $\chi_A(A) = 0$.

5. Demuestre la unicidad de la expresión obtenida en la prueba del Lema 2.5.

2.6. El lema anterior nos dice que los últimos i renglones de la tabla construida constituyen una base de $\ker \varphi^{i+1}$. En particular los vectores de la tabla constituyen una base B de $\ker \varphi^{s+1} = V$.

Cada columna de la tabla determina un bloque de Jordan de la matriz de representación $A(\varphi, B)$. En efecto,

$$\varphi\varphi^j(f_{i,m}) = \varphi^{j+1}(f_{i,m}) \quad \text{si } j < i$$
$$\varphi\varphi^i(f_{i,m}) = \varphi(e_{i,m}) = 0.$$

Luego, tenemos que $A(\varphi, B) = \bigoplus_{i=0}^{s} \left(\bigoplus_{m=1}^{n_i} J_i(0) \right)$. Con esto concluimos la demostración de la afirmación (3) y del Teorema de Jordan.

Ejercicios

1. Utilice el método de la prueba de (2.5) para construir una base de Jordan para la matriz

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

2. El mismo problema para la matriz $n \times n$,

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & & \vdots & \vdots \\ & & & \ddots & \ddots & \ddots & 0 & 0 \\ \vdots & \vdots & & & \ddots & \ddots & 1 & 1 \\ 0 & 0 & & & & \ddots & 0 & 1 \\ 0 & 0 & & & & & 0 & 0 \end{pmatrix}.$$

2.7. Supongamos dadas una transformación lineal $\varphi: V \to V$ y una base B de V tal que la matriz de representación $A(\varphi, B)$ es de Jordan. La forma de Jordan de φ en la base B está determinada hasta permutación de los bloques $J_n(\lambda)$, para cada $n \in \mathbf{N}$ y cada $\lambda \in k$. Denotaremos por $\nu(\lambda, n)$ al número de bloques de Jordan $J_n(\lambda)$ que aparecen en esta forma de Jordan de φ.

Lema. *El número $\nu(\lambda, n)$ no depende de la base de Jordan escogida.*

Demostración: Consideremos los números

$$\mu(\lambda, n) = \dim[\ker(\varphi-\lambda) \cap \mathrm{Im}(\varphi-\lambda)^{n-1}] - \dim[\ker(\varphi-\lambda) \cap \mathrm{Im}(\varphi-\lambda)^n].$$

Afirmamos que $\nu(\lambda, n) = \mu(\lambda, n)$. Como el número $\mu(\lambda, n)$ no depende de la base elegida, obtenemos el resultado.

Supongamos dadas dos bases de Jordan B y B' de V y sean $A = A(\varphi, B)$ y $A' = A(\varphi, B')$ las matrices de representación correspondientes. Como el número de valores propios y su multiplicidad no dependen

de la base se tiene que $\sum_n \nu(\lambda, n, A) = \sum_n \nu(\lambda, n, A')$, donde $\nu(\lambda, n, A)$ es el número $\nu(\lambda, n)$ para la matriz A. Podemos entonces suponer que λ es el único valor propio de φ, o sea $\varphi - \lambda \mathbf{1}_V$ es nilpotente. Restando $\lambda \mathbf{1}_V$ a φ podemos de hecho suponer que $\lambda = 0$, o sea, que φ es nilpotente.

Consideremos una matriz de Jordan A nilpotente. Nos basta ahora demostrar que

$$\nu(0, n, A) = \mu(0, n, A) (= \dim[\ker A \cap \operatorname{Im} A^{n-1}] - \dim[\ker A \cap \operatorname{Im} A^n]).$$

Como cada bloque de Jordan determina un subespacio invariante, resulta que μ es aditiva en los bloques, esto es, si $A = \bigoplus_{s=1}^{m} a_s J_s(0)$ con $a_s \in \mathbf{N}$, entonces $\mu(0, n, A) = \sum_{s=1}^{m} a_s \mu(0, n, J_s(0))$.

Pero observemos que

$$\dim[\ker J_s(0) \cap \operatorname{Im} J_s(0)^m] = \begin{cases} 1 & \text{si } s \leq m-1, \\ 0 & \text{si } m \leq s. \end{cases}$$

De donde, $\mu(0, n, J_s(0)) = \begin{cases} 1 & \text{si } n = s \\ 0 & \text{si } n \neq s \end{cases}$. Luego, $\mu(0, n, A) = \sum_{s=1}^{m} a_s = \nu(0, n, A)$. □

Debido a estos resultados, podemos decir que la forma de Jordan de una transformación $\varphi: V \to V$ no depende de la base de Jordan de V que se elija. Por ello nos permitiremos hablar de *la forma de Jordan* de la transformación φ.

Ejercicios

1. Sea $J_s(0)$ un bloque de Jordan con valor propio 0. Calcule las potencias $J_s(0)^m$, las imágenes $\operatorname{Im} J_s(0)^m$ y el núcleo $\ker J_s(0)$.

2. Demuestre que $\dim[\ker J_s(0) \cap \operatorname{Im} J_s(0)^m] = \begin{cases} 1 & \text{si } s \leq m-1 \\ 0 & \text{si } m \leq s \end{cases}$.

2.8. A lo largo del libro tendremos múltiples ocasiones de utilizar el Teorema de Jordan. Veremos aquí un primer ejemplo que además mejora notablemente una afirmación hecha antes. En efecto, en (2.3) de-

mostramos que para cada $n \times n$-matriz A existe un polinomio $\xi(t) \in k[t]$ de grado cuando más n^2 con $\xi(A) = 0$. En realidad tenemos el siguiente:

Teorema (Cayley-Hamilton). *Sea A una matriz $n \times n$ y sea $\chi_A(t)$ su polinomio característico. Entonces $\chi_A(A) = 0$.*

Demostración: Supongamos que $P^{-1}AP = J$ es la forma de Jordan asociada a A. Tenemos entonces

$$\chi_A(A) = \chi_A(PJP^{-1}) = P\chi_A(J)P^{-1}.$$

Como además $\chi_A(t) = \chi_J(t)$, sólo hay que demostrar que $\chi_J(J) = 0$. Claramente, si $J = J_1 \oplus J_2$, entonces $\chi_J(J_1 \oplus J_2) = \chi_J(J_1) \oplus \chi_J(J_2)$ y $\chi_J(t) = \chi_{J_1}(t)\chi_{J_2}(t)$. Luego, basta con demostrar que $\chi_{J_s(\lambda)}(J_s(\lambda)) = 0$ para cada bloque de Jordan $J_s(\lambda)$. Pero $\chi_{J_s(\lambda)}(t) = (t - \lambda)^s$ y sabemos que $(J_s(\lambda) - \lambda I_s)^s = 0$. □

Dada la matriz A, el conjunto de polinomios

$$I_A = \{p(t) \in k[t] : p(A) = 0\}$$

es un ideal de $k[t]$, generado por un polinomio $m_A(t)$ (ya que $k[t]$ es un anillo de ideales principales). Como $\chi_A(t) \in I_A$, entonces $m_A(t)$ divide a $\chi_A(t)$. Si J es la forma de Jordan de A, también tenemos que $m_A(J) = 0$, y en particular cada valor propio de A es una raíz de $m_A(t)$. En otras palabras,

$$\chi_A(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \cdots (t - \lambda_s)^{n_s},$$

con $\lambda_i \neq \lambda_j$ si $i \neq j$, y

$$m_A(t) = (t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2} \cdots (t - \lambda_s)^{m_s},$$

con $1 \leq m_i \leq n_i$, $i = 1, \ldots, s$.

El polinomio $m_A(t)$ se llama el *polinomio minimal* de A.

Ejercicios

1. Sea $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ matriz real 3×3. Calcule el polinomio característico $\chi_A(t)$ y verifique que $\chi_A(A) = 0$. Calcule también el polinomio minimal de A.

2. Sea A como en el ejercicio anterior. Sea $p(t) = t^6 + t^3 - 2$; demuestre que existe $q(t) \in k[t]$ de grado 2 tal que $p(A) = q(A)$ y calcule el valor de $p(A)$. [*Ayuda:* Use el algoritmo de la división para expresar $p(t) = a(t)\chi_A(t) + r(t)$ con $r(t)$ de grado a lo más 2; aplique el Teorema de Cayley-Hamilton.]

3. Sea A como en (1). Evalúe A^{30}.

4. Sea A una matriz cuya forma de Jordan es $J_{n_1}(\lambda_1) \oplus \cdots \oplus J_{n_s}(\lambda_s)$ con $\lambda_i \neq \lambda_j$ para $i \neq j$ (esto es, no hay dos bloques con el mismo valor propio). ¿Cuál es el polinomio minimal de A?

Otras aplicaciones

5. Pruebe que A y su matriz transpuesta A^T son similares.

6. Sean $\lambda_1, \ldots, \lambda_n$ los valores propios de A. Demuestre que $\det A = \prod_{i=1}^{n} \lambda_i$ y que $\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i$.

7. Sea $M_2(k)$ el conjunto de todas las matrices 2×2 sobre k. Pruebe que la función $f: M_2(k) \to k^2$, $A \mapsto (\operatorname{tr} A, \det A)$ es suprayectiva.

2.9. Sea $\varphi: V \to V$ una transformación lineal como las que hemos visto antes. En la prueba del Teorema de Jordan hemos demostrado que el espacio V se descompone como

$$V = \bigoplus_{\lambda \in \operatorname{Spec} \varphi} V_\lambda,$$

donde $\operatorname{Spec} \varphi$ es el conjunto de valores propios de φ (llamado el *espectro* de φ) y para cada $\lambda \in \operatorname{Spec} \varphi$, V_λ es el subespacio de V invariante bajo φ definido como $V_\lambda = \ker(\varphi - \lambda 1_V)^{n_\lambda}$, en caso de que

$$\chi_\varphi(t) = \prod_{\lambda \in \operatorname{Spec} \varphi} (t - \lambda)^{n_\lambda}.$$

(*Nota:* Obsérvese que en este caso hemos usado el Teorema de Cayley-Hamilton.)

El espacio V_λ se llama el *espacio espectral* de φ relativo a λ. También puede caracterizarse como

$$V_\lambda = \{v \in V: \text{ existe } m \geq 1 \text{ con } (\varphi - \lambda 1_V)^m v = 0\},$$

con la ventaja de que esta definición no hace referencia a la potencia n_λ con la cual aparece $(t - \lambda)$ como factor de $\chi_\varphi(t)$. La descomposición $V = \bigoplus_{\lambda \in \operatorname{Spec} \varphi} V_\lambda$ se llama la *descomposición espectral* de V según φ.

Los espacios espectrales tienen un papel importante en diferentes partes de la matemática y la física. Veremos después algunos ejemplos de estas aplicaciones.

Ejercicios

1. Demuestre la caracterización de V_λ mencionada en el texto.

2. Calcule la descomposición espectral de k^3 para la matriz

$$\begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}.$$

2.10. De entre los *métodos para calcular la forma de Jordan* de una matriz dada, hay uno particularmente eficiente y sencillo que deseamos examinar. Para ello tenemos que hacer primero algunas consideraciones teóricas.

Sea A una matriz $n \times n$, y supongamos que la forma de Jordan asociada a A tiene $\nu(\lambda, s)$ bloques de Jordan de la forma $J_s(\lambda)$, para cada $\lambda \in \operatorname{Spec} A$. Deseamos expresar los números $\nu(\lambda, s)$ por medio del rango de las matrices $(A - \lambda I_n)^t$ con $t \in \mathbf{N}$. Como veremos en un ejemplo es sencillo calcular estos rangos.

Para cada $t \in \mathbf{N}$, sea $r(\lambda, t)$ el rango de la matriz $(A - \lambda I_n)^t$, $\lambda \in \operatorname{Spec} A$. Sea J la forma de Jordan de A, de manera que

$$J = \bigoplus_{\lambda \in \operatorname{Spec} A} \left(\bigoplus_s \nu(\lambda, s) J_s(\lambda) \right).$$

Como el rango es un invariante por similaridad, basta con calcular el rango de $(J - \lambda I_n)$, $t \in \mathbf{N}$. Claramente,

$$\operatorname{rg}(J - \lambda I_n)^t = \sum_{\mu \in \operatorname{Spec} A} \sum_{s} \nu(\mu, s) \operatorname{rg}(J_s(\mu - \lambda))^t$$

y

$$\operatorname{rg}(J_s(\mu - \lambda)^t) = \begin{cases} s & \text{si } \lambda \neq \mu, \\ s - t & \text{si } \lambda = \mu \text{ y } t \leq s, \\ 0 & \text{si } \lambda = \mu \text{ y } s \leq t. \end{cases}$$

De aquí se sigue que

$$r(\lambda, t) = r(\lambda, t-1) - \sum_{t<s} \nu(\lambda, s).$$

Restando dos sucesivas de las ecuaciones así obtenidas, llegamos a la ecuación fundamental:

$$\nu(\lambda, t) = r(\lambda, t+1) - 2r(\lambda, t) + r(\lambda, t-1), \qquad t \in \mathbf{N}.$$

Ejemplo: Calcular la forma de Jordan de la matriz

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 1 & -1 & 2 \end{bmatrix}.$$

Para ello hay que calcular primero los valores propios de A.

Resulta que $\chi_A(t) = (t-2)^3(t-1)$, siendo los valores propios 1 y 2.

Tenemos que calcular los números $r(1,t)$ para $t = 1,2$ y $r(2,t)$ para $t = 1,2,3,4$. Como sabemos que $\nu(1,1) = 1$, calcularemos sólo $r(2,t)$:

$$A - 2I_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 1 & -1 & 0 \end{bmatrix}.$$

Reduciendo por renglones, tenemos que esta matriz tiene el mismo rango que

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

(Obsérvese que la reducción por renglones —o columnas— no altera el rango de la matriz.) Esta matriz tiene rango 3, esto es $r(2,1) = 3$. Calculando $(A - 2I_4)^i$ para $i = 2,3,4$, obtenemos: $r(2,2) = 2$, $r(2,3) = 1$ y $r(2,4) = 1$. Lo que nos da $\nu(2,3) = 1$, o sea, un solo bloque de Jordan con valor propio 2.

Ejercicio: Calcule la forma de Jordan de la matriz

$$\begin{bmatrix} 1 & -4 & -1 & -4 \\ 2 & 0 & 5 & -4 \\ -1 & 1 & -2 & 3 \\ -1 & 4 & -1 & 6 \end{bmatrix}.$$

3. La forma de Jordan. El caso general

En esta sección k denotará un campo arbitrario. Dada una transformación lineal $\varphi: V \to V$, construiremos una base B de V tal que la matriz de representación $A(\varphi, B)$ tome una forma especial que en el caso en que k sea algebraicamente cerrado coincidirá con la forma de Jordan estudiada en la sección anterior; en el caso general, esta forma cumple con la propiedad esencial del caso algebraicamente cerrado: unicidad hasta permutación de bloques.

3.1. Recordemos rápidamente algunos conceptos de anillos y módulos. Para más detalles el lector puede consultar cualquiera de los textos al respecto.

Sea R un anillo con unidad. Una terna $(M, +, \cdot)$ se llama *R-módulo* (izquierdo) si se cumple:

(a) La operación $+: M \times M \to M$ hace de la pareja $(M, +)$ un grupo abeliano.

(b) La multiplicación por escalares $\cdot: R \times M \to M, (a, m) \to am$ satisface las siguientes propiedades:

$$1m = m,$$
$$(a_1 + a_2)m = a_1 m + a_2 m,$$
$$a(m_1 + m_2) = am_1 + am_2,$$
$$(a_1 a_2)m = a_1(a_2 m),$$

Por abuso de notación omitiremos en general las aplicaciones $+$ y $\cdot$, y diremos que M es un R-módulo.

Todos los conceptos usuales: submódulo, módulo cociente, módulo finitamente generado, homomorfismos de módulos, suma directa, etcétera, pueden definirse fácilmente. Denotamos por Mod R a la *categoría de R-módulos.*

Ejemplos

a) Sea $R = k[t]$ el anillo de polinomios en una variable. Un R-módulo M es un espacio vectorial dotado de una aplicación lineal $\varphi: M \to M$, $m \mapsto tm$.

Así la categoría de $k[t]$-módulos Mod $k[t]$ puede identificarse con la categoría cuyos objetos son parejas (V, φ) con V como k-espacio vectorial y $\varphi: V \to V$ como transformación lineal; un morfismo $f: (V, \varphi) \to (W, \psi)$ en la categoría es una transformación lineal $f: V \to W$ tal que $f\varphi = \psi f$.

Dadas dos matrices A y B $n \times n$, los módulos (k^n, A) y (k^n, B) que ellas definen son isomorfos si y solamente si A y B son similares.

b) Dado su polinomio $p(t) \in k[t]$, el cociente $k[t]/(p(t))$ tiene una estructura natural de $k[t]$-módulo. Este módulo es cíclico ya que la clase $t + p(t)k[t]$ en $k[t]/(p(t))$ genera el módulo.

La categoría de $k[t]/(p(t))$-módulos puede identificarse con la subcategoría plena de los $k[t]$-módulos (V, φ) que satisfacen $p(\varphi) = 0$.

Ejercicios

1. Demuestre el teorema chino del residuo:

Sean $0 \neq p_j(t) \in k[t]$, $j = 1, \ldots, m$ polinomios tales que para $i \neq j$, $p_i(t)$ y $p_j(t)$ no tienen divisores comunes (escribimos $(p_i(t), p_j(t)) = 1$).

Sean $r_1(t), \ldots, r_m(t) \in k[t]$ arbitrarios. Existe entonces $g(t) \in k[t]$ de forma que $g(t) \equiv r_i(t) (\mathrm{mod}\, p_i(t))$ para $i = 1, \ldots, m$ (o sea, las clases de $g(t)$ y $r_i(t)$ son iguales en $k[t]/(p_i(t))$).

Ayuda: Sea $\pi_i: k[t] \to k[t]/(p_i(t))$ la proyección natural. Hay que demostrar que el morfismo

$$\pi: k[t] \to \bigoplus_{i=1}^{m} \ker \pi_i, g(t) \to (g(t) + p_i(t).k[t])_i$$

es suprayectivo. Calcule el núcleo $\ker \pi$ y demuestre que

$$\dim_k k[t]/\ker \pi = \dim \left(\bigoplus_{i=1}^{m} k[t]/(p_i(t)) \right) = \sum_{i=1}^{n} \mathrm{gr}\, p_i(t).$$

2. Recordemos que un polinomio $p(t) \in k[t]$ se dice *irreducible* si cada vez que se factoriza como $p(t) = a(t)b(t)$ alguno de los dos polinomios $a(t)$ o bien $b(t)$ es invertible en $k[t]$ (o sea, un polinomio de grado cero).

a) ¿Cuáles son los polinomios irreducibles en $\mathbf{R}[t]$, en $\mathbf{C}[t]$?

b) ¿Cuáles son los polinomios irreducibles en $\mathbf{Z}_7[t]$?

c) Demuestre que $k[t]$ tiene infinitos polinomios irreducibles —¡aun si k es finito!—. (*Ayuda:* Imite la prueba de Euclides de la existencia de infinitos primos en $\mathbf{Z}$.)

3. Todos los polinomios irreducibles en $k[t]$ tienen grado uno si y solamente si k es algebraicamente cerrado.

4. Si M es un R-módulo y para la ideal I de R tenemos que $IM = 0$, entonces M es un R/I-módulo.

3.2. Sea $p(t) \in k[t]$ un polinomio, $p(t) = a_0 + a_1 t + \cdots + a_{s-1} t^{s-1} + t^s$. Definimos la *matriz compañera* $C_{p(t)}$ de p como la matriz $s \times s$ dada por

$$C_{p(t)} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & & & \vdots & \\ 0 & 0 & \cdots & 1 & -a_{s-1} \end{bmatrix}.$$

Dada una matriz A, como en (2.3) podemos demostrar que existe un polinomio $\xi(t) \in k[t]$ tal que $\xi(A) = 0$. Por tanto, el *polinomio minimal* $m_A(t)$ —generador normalizado del ideal $\{\xi(t): \xi(A) = 0\}$— está bien definido.

Lema. *Sea $A = C_{p(t)}$ la matriz compañera del polinomio $p(t) = a_0 + a_1 t + \cdots + a_{s-1}t^{s-1} + t^s$. Entonces, $p(t)$ es el polinomio característico y también el polinomio minimal de A, $\chi_A(t) = p(t) = m_A(t)$.*

Demostración: Primero demostraremos que $p(A) = 0$. En efecto, sea $\{e_1, \ldots, e_s\}$ la base canónica de k^s. Se tiene que

$$Ae_i = e_{i+1},\ 1 \le i \le s-1 \quad \text{y} \quad Ae_s = -\sum_{j=0}^{s-1} a_j e_{j+1} = -\left(\sum_{j=0}^{s-1} a_j A^j\right) e_1.$$

Esto es $p(A)e_1 = A^s e_1 + \left(\sum_{j=0}^{s-1} a_j A^j\right) e_1 = 0$. Para todas las $1 \le i \le s$, se sigue que $p(A)e_i = p(A)A^{i-1}e_1 = A^{i-1}p(A)e_1 = 0$. Luego $p(A) = 0$.

Tenemos entonces que $m_A(t)$ es un divisor de $p(A)$. Demostraremos que el grado de $m_A(t)$ es s, lo que nos dará la igualdad. Sea $0 \neq q(t) = \sum_{i=0}^{\ell} b_i t^i \in k[t]$ un polinomio de grado $\ell < s$. Calculemos

$$q(A)e_1 = \sum_{i=0}^{\ell} b_i A^i e_1 = \sum_{i=0}^{\ell} b_i e_{i+1},$$

que es diferente de cero ya que $\{e_1, \ldots, e_s\}$ es una base k^s.

Para demostrar que $\chi_A(t) = p(t)$ basta con desarrollar el determinante de la matriz $(A - tI_s)$ por medio de la última columna. □

3.3. Decimos que un polinomio $p(t)$ irreducible está *normalizado* si es de la forma $p(t) = a_0 + a_1 t + \ldots + a_s t^{s-1} + t^s$.

Tomemos un polinomio irreducible normalizado $p(t) = a_0 + a_1 t + \cdots + a_s t^{s-1} + t^s \in k[t]$. Definimos el *bloque de Jordan* $J_n(p(t))$ de tamaño n asociado a $p(t)$ de la siguiente manera:

$$J_n(p(t)) = \begin{bmatrix} C_{p(t)} & N & & 0 \\ & C_{p(t)} & \ddots & \\ & & \ddots & N \\ 0 & & & C_{p(t)} \end{bmatrix},$$

donde la matriz compañera $C_{p(t)}$ de $p(t)$ aparece n veces en la diagonal y

$$N = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

es una matriz $s \times s$.

Observemos que si el campo k es algebraicamente cerrado, el polinomio $p(t)$ es de la forma $p(t) = t - \lambda$ con $\lambda \in k$. Entonces $J_n(p(t)) = J_n(\lambda)$, el bloque de Jordan que definimos antes en la sección 2.

Ejercicio: Demuestre que el polinomio característico y el polinomio minimal de $J_n(p(t))$ coinciden y son iguales a $p(t)$. (Proceda como en el Lema 3.2.)

3.4. Se dice que la matriz A es una *forma de Jordan* si existen polinomios irreducibles normalizados $p_1(t), \ldots, p_s(t)$ y números $n_1, \ldots, n_s$ tales que

$$A = J_{n_1}(p_1(t)) \oplus \cdots \oplus J_{n_s}(p_s(t)).$$

Podemos enunciar la forma más general del Teorema de Jordan:

Teorema. *Sea $\varphi: V \to V$ una transformación lineal y supongamos que $\dim_k V$ es finita. Existe entonces una base B de φ tal que la matriz de representación $A(\varphi, B)$ es una forma de Jordan*

$$A(\varphi, B) = J_{m_1}(p_1(t)) \oplus \cdots \oplus J_{m_s}(p_s(t))) \text{ con } \chi_\varphi(t) = p_1(t)^{m_1} \cdots p_s(t)^{m_s}.$$

En esencia, la *demostración* es la misma que en el caso algebraicamente cerrado. Por ello sólo indicaremos los pasos principales y daremos detalles sólo cuando difieran en algo del caso ya tratado.

Deseamos mostrar las siguientes afirmaciones:

(1) Sea $m_\varphi(t) = p_1(t)^{m_1} \ldots p_s(t)^{m_s}$ la descomposición del polinomio minimal en factores irreducibles normalizados en $k[t]$. Para cada $i = 1, \ldots, s$ el subespacio $V_i = \ker p_i(\varphi)^{m_i}$ de V es invariante bajo φ. Si llamamos $\varphi_i: V_i \to V_i$ a la restricción de φ, (V_i, φ_i) es un submódulo del $k[t]$-módulo (V, φ). Con esta estructura se obtiene una descomposición en suma directa de $k[t]$-módulos

$$V = V_1 \oplus \cdots \oplus V_s.$$

(2) Supongamos que existe un polinomio irreducible normalizado $p(t) \in k[t]$ de forma que $p(\varphi)^m = 0$, entonces existen números $m_1, \ldots, m_s$ y una base B de V de manera que la matriz de representación es

$$A(\varphi, B) = J_{m_1}(p(t)) \oplus \cdots \oplus J_{m_s}(p(t)).$$

Asociada a esta descomposición tenemos una descomposición de k-espacios vectoriales $V = W_1 \oplus \cdots \oplus W_s$ de forma que $(W_i, J_{m_i}(p(t)))$ es un $k[t]$-submódulo de (V, φ) isomorfo a $k[t]/(p(t)^{m_i})$.

Por supuesto, como en la sección 2, tenemos que estas dos afirmaciones implican el teorema de descomposición de Jordan. La afirmación que nos indica que $\chi_\varphi(t) = p_1(t)^{m_1} \ldots p_s(t)^{m_s}$ se sigue de (1), (2) y (3.3). En particular se obtiene:

Corolario (Cayley-Hamilton). *Sea $\varphi: V \to V$ un homomorfismo lineal y supongamos que $\dim_k V$ es finita. Entonces $\chi_\varphi(\varphi) = 0$.*

3.5. Claramente, la afirmación (1) se demuestra como (2.4). Pasemos por tanto a la afirmación (2).

Sea $p(t) = a_0 + a_1 t + \cdots + a_{\ell-1} t^{\ell-1} + t^\ell$ un polinomio irreducible normalizado. Sea $\varphi: V \to V$ un morfismo lineal tal que $p(\varphi)^{s+1} = 0$.

La siguiente es una cadena de $k[t]$-submódulos de (V, φ):

$$\begin{aligned} 0 \neq V^0 &= \ker p(\varphi) \supset (V^0 \cap p(\varphi)(V)) \\ &= V^1 \supset (V^0 \cap p(\varphi)^2(V)) = V^2 \supset \cdots \supset (V^0 \cap p(\varphi)^s(V)) = V^s. \end{aligned}$$

Como para cada uno de estos espacios $p(\varphi)V_i = 0$, todos ellos son también módulos sobre el anillo $K = k[t]/(p(t))$. Como $p(t)$ es irreducible entonces K es un campo y cada V_i es un K-espacio vectorial. Procedemos ahora exactamente como en (2.5) —con K en lugar de k— hasta construir una K-base $\{e_{i,m}: i = 1, \ldots, s,\ m = 1, \ldots, n_i\}$ de V_0 de forma que $e_{i,m} = p(\varphi)^i(f_{i,m})$ para algunas $f_{i,m} \in V$.

Lema. *Para cada $i = 1, \ldots, s$, el conjunto de vectores $\{\varphi^r p(\varphi)^t(f_{j,m})$: $0 \leq r \leq \ell - 1,\ j \leq i$ o bien $0 \leq j - i \leq t \leq j\}$ constituye una base de $\ker p(\varphi)^{i+1}$.*

Demostración: La misma que la de (2.5) si se observa que $\{\varphi^r(e_{j,m}): 0 \leq r \leq \ell - 1,\ j \leq i$ y $1 \leq m \leq n_i\}$ es una k-base de V_i. □

Para cada par (j, m) con $1 \leq m \leq n_i$ y $j \leq i$ consideremos el vector $f_{j,m}$. Sea $W_{j,m}$ el k-espacio vectorial generado por $\{\varphi^r p(\varphi)^t(f_{j,m}): 0 \leq r \leq \ell - 1,\ 0 \leq t \leq j\}$. Observemos en primer lugar que $W_{j,m}$ es un $k[t]$-submódulo de V, o en otras palabras que es φ-invariante. En efecto,

$$\varphi(\varphi^r p(\varphi)^t(f_{j,m})) = \varphi^{r+1} p(\varphi)^t(f_{j,m}) \text{ si } r < \ell - 1$$

y

$$\begin{aligned} \varphi(\varphi^{\ell-1} p(\varphi)^t(f_{j,m})) &= \left(-\sum_{i=0}^{\ell-1} a_i \varphi^i + p(\varphi)\right) p(\varphi)^t(f_{j,m}) \\ &= \begin{cases} -\sum_{i=0}^{\ell-1} a_i \varphi^i p(\varphi)^t(f_{j,m}) + p(\varphi)^{t+1}(f_{j,m}), & \text{si } t < j, \\ -\sum_{i=0}^{\ell-1} a_i \varphi^i p(\varphi)^t(f_{j,m}), & \text{si } t = j. \end{cases} \end{aligned}$$

Además de probar que $W_{j,m}$ es un $k[t]$-submódulo de V, hemos encontrado una base $B_{j,m}$ de $W_{j,m}$ de forma que la restricción $\varphi_{j,m} = \varphi|: W_{j,m} \to W_{j,m}$ tiene por matriz de representación

$$A(\varphi_{j,m}, B_{j,m}) = J_j(p(t)).$$

Finalmente, es claro que $V = \bigoplus_{(j,m)} W_{j,m}$ como $k[t]$-módulos. Además, el $k[t]$-módulo $W_{j,m}$ está generado por el vector $f_{j,m}$, luego existe un epimorfismo de $k[t]$-módulos $\pi: k[t] \to W_{j,m}$, $1 \mapsto f_{j,m}$. Como $p(t)^j W_{j,m} = 0$, se obtiene un epimorfismo $\bar{\pi}: k[t]/(p(t)^j) \to W_{j,m}$. Dado que los dos módulos tienen dimensión $(j+1)\ell$, obtenemos finalmente que $\bar{\pi}$ es un isomorfismo $W_{j,m} \simeq k[t]/(p(t)^j)$. Con lo cual se ha completado la demostración de (2).

Ejercicios

1. Sea $\varphi: \mathbf{C}^2 \to \mathbf{C}^2$ dado por la matriz

$$\begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}$$

con $a \in \mathbf{C}$. Considere a $\mathbf{C}$ como en $\mathbf{R}$-espacio vectorial de dimensión dos y calcule la forma de Jordan de φ sobre $\mathbf{R}$ en caso de que $a = 1$ y de que $a = i$.

2. Considere la matriz real 3×3

$$A = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

que tiene a 1 como único valor propio (con multiplicidad 3). Proporcione condiciones sobre a para que la forma de Jordan de A sea $J_3(1)$, o bien $J_1(1) \oplus J_2(1)$, o bien de la forma $J_1(1) \oplus J_1(p(t))$ donde $p(t)$ es un polinomio irreducible de grado 2 en $\mathbf{R}[t]$.

3. Sea A una matriz $n \times n$ sobre k. Sea $p(t) \in k[t]$ un polinomio irreducible normalizado. Como en (2.7), definimos $\nu(p(t), s)$ como el número de bloques de Jordan $J_s(p(t))$ que aparecen en la forma de Jordan de A. Sea $r(p(t), m)$ el rango de la matriz $p(A)^m$. Demuestre que

$$\nu(p(t), s) \operatorname{gr} p(t) = r(p(t), s-1) - 2r(p(t), s) + r(p(t), s+1).$$

3.6. Sea V un espacio vectorial sobre el campo de los números complejos $\mathbf{C}$. Sea $B = \{v_1 \dots v_n\}$ una base de V donde el morfismo $\varphi: V \to V$ toma la forma de Jordan $J_n(a)$ con $a \in \mathbf{C}$, esto es:

$$\varphi(v_1) = av_1, \quad \varphi(v_j) = av_j + v_{j-1}, \quad 2 \le j \le n.$$

Consideremos V como un $\mathbf{R}$-espacio vectorial de dimensión $2n$. Deseamos calcular la forma de Jordan de la transformación $\mathbf{R}$-lineal φ. Consideremos dos casos:

Caso 1: Supongamos que $a \in \mathbf{R}$. Podemos calcular el número $\nu(a, s)$ de bloques de Jordan de tamaño $s \times s$ con valor propio a. Por ejemplo,

$$\nu(a, n) = \dim_{\mathbf{R}}(\varphi - a1_V)^{n-1}(V) - 2\dim_{\mathbf{R}}(\varphi - a1_V)^n(V) + \dim_{\mathbf{R}}(\varphi - a1_V)^{n+1}(V) = 2,$$

ya que $(\varphi - a1_V)^n = 0$ y $(\varphi - a1_V)^{n-1}(V)$ es un $\mathbf{C}$-espacio vectorial de dimensión 1. Así obtenemos que la forma de Jordan de φ sobre $\mathbf{R}$ es $J_n(a) \oplus J_n(a)$.

Caso 2. Supongamos que $a \notin \mathbf{R}$. Entonces $a \neq \bar{a}$ y $p(t) = (t-a)(t-\bar{a})$ es un polinomio irreducible sobre $\mathbf{R}[t]$.

Observemos que $p(\varphi)^n = 0$ y que la transformación $(\varphi - \bar{a}1_V)$ es invertible. Luego $\dim_{\mathbf{R}}(\varphi - a1_V)^s(V) = \dim_{\mathbf{R}}(\varphi - a1_V)^s(\varphi - \bar{a}1_V)^s(V) = \dim_{\mathbf{R}} p(\varphi)^s(V)$. En particular, $\nu(p(t), n) \operatorname{gr} p(t) = \dim_{\mathbf{R}} p(\varphi)^{n-1}(V) = \dim_{\mathbf{R}}(\varphi - a1_V)^{n-1}(V) = 2$. De donde $\nu(p(t), n) = 1$ y la forma de Jordan buscada es

$$\begin{bmatrix} 0 & -a\bar{a} & 0 & 1 & & & & \\ 1 & a+\bar{a} & 0 & 0 & & & 0 & \\ 0 & 0 & 0 & -a\bar{a} & & & & \\ 0 & 0 & 1 & a+\bar{a} & & & & \\ & & & & \ddots & & & \\ & & & & & 0 & 1 & \\ & & 0 & & & 0 & 0 & \\ & & & & & 0 & -a\bar{a} & \\ & & & & & 1 & a+\bar{a} & \end{bmatrix}.$$

Ejercicios

1. Sea $\varphi: V \to V$ una transformación lineal. Demuestre que las siguientes afirmaciones son equivalentes:

a) El polinomio minimal $m_\varphi(t)$ se descompone en $k[t]$ como producto de monomios $(t-a_1), \dots, (t-a_s)$ con $a_1, \dots, a_s$ todas diferentes.

b) Hay una base $\{v_1, \dots, v_s\}$ de V tal que

$$Av_j = a_j v_j, \ (j = 1, \dots, s)$$

con $a_j \in k$.

2. Sea $\varphi: V \to V$ una transformación lineal con $\dim_k V = n$. Supongamos que el polinomio característico de φ tiene n valores propios dife-

rentes. Entonces la transformación φ es diagonalizable. (Esto también fue demostrado en (1.4).)

3. Calcule la forma de Jordan sobre $\mathbf{R}$ de la matriz:

$$\begin{bmatrix} 1 & 1 & -1 & -4 \\ 2 & 0 & 5 & -4 \\ -1 & 1 & -2 & 3 \\ -1 & 4 & -1 & 0 \end{bmatrix}.$$

4. Sucesiones de matrices

En esta sección supondremos que $k = \mathbf{C}$ o $\mathbf{R}$. Para el espacio k^n consideremos la norma euclidiana denotada por $\| - \|$.

4.1. Supongamos que $A_1, A_2, A_3, \ldots$ es una sucesión de matrices rectangulares $n \times m$; entonces decimos que el límite $\lim_{s\to\infty} A_s$ existe y es igual a la matriz A (o bien que la sucesión $(A_s)_s$ converge a A) en caso de que tengamos:

$$A_s = \left(a_{ij}^{(s)}\right), \quad A = (a_{ij})$$

y $\lim_{n\to\infty} a_{ij}^{(s)} = a_{ij}$ para cada par $1 \leq i \leq n$, $1 \leq j \leq m$.

Lema. *Si* $\lim_{s\to\infty} A_s = A$, $\lim_{s\to\infty} B_s = B$ *con* A *como matriz* $n \times m$ *y* B *como matriz* $p \times \ell, \lambda \in k$, *entonces*

i) $\lim_{s\to\infty} B_s A_s = BA$, *si* $p = m$.

ii) $\lim_{s\to\infty} \lambda A_s = \lambda A$.

iii) $\lim_{s\to\infty} (A_s + B_s) = A + B$, *si* $n = p$ *y* $m = \ell$.

iv) $\lim_{s\to\infty} (PA_sQ) = PAQ$, *para matrices* P *y* Q. □

Ejercicios

1. Demuestre el lema anterior.

2. Supongamos que tenemos m sucesiones de matrices $(A_s^{(i)})_s$, $i = 1, \ldots, m$ y que cada sucesión $(A_s^{(i)})_s$ converge a $A^{(i)}$. Entonces

$$\lim_{s\to\infty} \bigoplus_{i=1}^{m} A_s^{(i)} = \bigoplus_{i=1}^{m} A^{(i)}.$$

4.2. Debido a las observaciones de (4.1) para examinar la sucesión de potencias $(A^s)_s$ de una matriz cuadrada A, requerimos solamente estudiar la sucesión de potencias de los bloques de Jordan de una forma de Jordan asociada a A.

Consideremos el bloque de Jordan $J_n(\lambda)$, que podemos escribir como $J_n(\lambda) = \lambda I_n + N$, con N nilpotente (de hecho $N^n = 0$).

Entonces, obtenemos para $m \geq n - 1$,

$$J_n(\lambda)^m = (\lambda I_n + N)^m = \sum_{j=0}^{m} \binom{m}{j} \lambda^{m-j} N^j = \sum_{j=0}^{m-1} \binom{m}{j} \lambda^{m-j} N^j$$

$$= \begin{bmatrix} \lambda^m & \binom{m}{1}\lambda^{m-1} & \binom{m}{2}\lambda^{m-2} & \cdots & \binom{m}{n-1}\lambda^{m-n+1} \\ 0 & \lambda^m & \binom{m}{1}\lambda^{m-1} & \cdots & \binom{m}{n-2}\lambda^{m-n+2} \\ & 0 & \ddots & \ddots & \vdots \\ & & & & \binom{m}{1}\lambda^{m-1} \\ & & & & \lambda^m \end{bmatrix}.$$

De aquí podemos probar fácilmente la siguiente proposición:

Proposición. a) *Si* $\|\lambda\| < 1$, *entonces para cualquier polinomio* $p(t) \in k[t]$

$$\lim_{s\to\infty} p(s) J_n(\lambda)^s = 0.$$

b) *El límite* $\lim_{s\to\infty} J_n(\lambda)^s$ *existe si y solamente si las siguientes condiciones son satisfechas:*

1) $\|\lambda\| \leq 1$.

2) *Si* $\|\lambda\| = 1$, *entonces* $\lambda = 1$ *y* $n = 1$.

Demostración: (a) Para cualquier $0 \leq j \leq n - 1$, $\binom{s}{j}$ es un polinomio en s, luego $\lim_{s\to\infty} \left\| p(s)\binom{s}{j} \right\| \|\lambda\|^s = 0$ o bien $\lim_{s\to\infty} p(s)\binom{s}{j}\lambda^s = 0$. Lo que muestra que $\lim_{s\to\infty} p(s) J_n(\lambda)^s = 0$.

(b) En caso de que se cumplan (1) y (2) está claro que $\lim_{s\to\infty} J_n(\lambda)^s$ existe. Supongamos para probar el converso que $\lim_{s\to\infty} J_n(\lambda)^s$ existe. Entonces existe $\lim_{s\to\infty} \lambda^s$ y si $n > 1$, también existe $\lim_{s\to\infty} s\lambda^{s-1}$. Del primer límite, se obtiene que $\|\lambda\| \leq 1$. Supongamos que $\|\lambda\| = 1$,

entonces $\| \lim_{s\to\infty} \lambda^s \| = \lim_{s\to\infty} \|\lambda\|^s = 1$ y

$$\lambda \lim_{s\to\infty} \lambda^s = \lim_{s\to\infty} \lambda^s$$

implica que $\lambda = 1$. La existencia del límite $\lim \lambda^s$ implica que $n = 1$. □

Ejercicio: Si $\|\lambda\| < 1$, pruebe que $\lim_{s\to\infty} \binom{s}{j}\lambda^{s-j} = 0$.

4.3. Definiremos varios conceptos para una matriz A y un escalar $\lambda \in k$.

La *multiplicidad algebraica* de λ es el número de veces que λ aparece como raíz del polinomio característico $\chi_A(t)$ de A.

La *multiplicidad geométrica* de λ es el número de bloques de Jordan con valor propio λ que aparece en la forma de Jordan de A; con la notación introducida antes tenemos que $\sum_s \nu(\lambda, s)$ es la multiplicidad geométrica de λ.

El *radio espectral* $\rho(A)$ de A es

$$\rho(A) = \max\{\|\lambda\| : \lambda \in \operatorname{Spec} A\}.$$

Claramente, $\rho(A)$ es el radio del círculo más pequeño con centro en 0 y que contiene a todos los valores propios de A.

Con estas notaciones podemos reformular así los resultados de (4.1) y (4.2).

Teorema. *Sea A una matriz compleja. Tenemos*

a) *Si $\rho(A) < 1$, entonces $\lim_{s\to\infty} A^s = 0$.*

b) *El límite $\lim_{s\to\infty} A^s$ existe si y solamente si las siguientes condiciones se satisfacen:*

1) $\rho(A) \leq 1$.

2) *Si $\rho(A) = 1$, entonces 1 es el único valor propio de A con norma 1 y la multiplicidad algebraica y geométrica de 1 como valor propio de A coinciden.* □

Ejercicios

1. Sea $L = \lim_{s\to\infty} A^s$. Sea r el rango de la matriz $A - I_n$. Entonces L es similar a la matriz $I_{n-r} \oplus 0_r$, donde 0_r denota la matriz $r \times r$ con todas sus entradas 0.

2. Se dice que una matriz A tiene *potencias acotadas* si existe una constante c tal que $\|A^s\| \leq c$ para toda s. Encuentre condiciones similares a las de la parte (b) del teorema que caractericen a las matrices que tienen potencias acotadas.

3. Dada la matriz 2×2

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

encuentre condiciones necesarias y suficientes en los coeficientes a, b, c, d para que $\lim_{s\to\infty} A^s$ exista.

4.a) Demuestre que para cualquier matriz $A = (a_{ij})$ de tamaño $n \times n$,

$$\rho(A) \leq \max\left\{\sum_{j=1}^{n} |a_{ij}| : i = 1, \ldots, n\right\}.$$

[*Ayuda:* Sea $0 \neq v$ vector tal que $Av = \lambda v$ para $|\lambda| = \rho(A)$. Podemos suponer que $0 \neq |v_n| \geq |v_i|$ para $i = 1, \ldots, n$. Entonces $\sum_{j=1}^n |a_{nj}| \geq \sum_{j=1}^n |a_{nj}| \left|\frac{v_j}{v_n}\right| \geq \rho(A)$.]

b) Demuestre que también se tiene

$$\rho(A) \leq \max\left\{\sum_{i=1}^{n} |a_{ij}| : j = 1, \ldots, n\right\}.$$

4.4. El radio espectral $\rho(A)$ de una matriz A es un invariante muy importante; nos encontraremos con él frecuentemente. En (4.2) ya hemos visto que se relaciona con el crecimiento de las potencias A^s, $s = 1, 2, \ldots$ de A. De hecho esta relación es más fuerte, como veremos en el siguiente importante resultado.

Consideraremos una matriz A de tamaño $n \times n$ como un elemento del espacio vectorial $\mathbf{C}^{n^2}$, de manera que podemos calcular la norma euclidiana $\|A\|$ de A.

Teorema. *Sea A una matriz compleja $n \times n$. Entonces*

$$\rho(A) = \lim_{s\to\infty} \sqrt[s]{\|A^s\|}.$$

Demostración: Sea T una matriz invertible tal que $J = TAT^{-1}$ es una forma de Jordan. Como el radio espectral es invariante bajo conjugación, es suficiente demostrar:

(1) el teorema para J; (2) el siguiente resultado auxiliar:

Lema técnico. $\lim_{s\to\infty} \sqrt[s]{\|A^s\|} = \lim_{s\to\infty} \sqrt[s]{\|J^s\|}$.

Demostración del Lema: Consideremos el morfismo lineal $\hat{T}: \mathbf{C}^{n^2} \to \mathbf{C}^{n^2}$ que lleva un vector $v = (v^1, \dots, v^n)$ con $v^i \in \mathbf{C}^n$ en $\hat{T}(v) = (T(v^1), \dots, T(v^n))$. Por supuesto $\hat{T} = T \oplus \cdots \oplus T$ (n veces) y por ello $\hat{T}$ es invertible. Interpretando A como elemento de $\mathbf{C}^{n^2}$, tenemos $\hat{T}(A) = TA$.

Consideremos la esfera unitaria $S = \{v \in \mathbf{C}^{n^2} : \|v\| = 1\}$ en $\mathbf{C}^{n^2}$, que es cerrada y acotada. Como $\hat{T}$ es continua, alcanza su máximo y su mínimo en S, esto es, existen constantes $c, d \geq 0$ tales que $c \leq \|\hat{T}(x)\| \leq d$, para toda $x \in S$; además, $c = \|\hat{T}(x_1)\|$, $d = \|\hat{T}(x_2)\|$ para algunas $x_1, x_2 \in S$. Como $\hat{T}(x_1) \neq 0$, entonces $0 < c$.

Considerando el elemento $A^s/\|A^s\| \in S$, tenemos

$$c\|A^s\| \leq \|TA^s\| \leq d\|A^s\|, \text{ para toda } s \in N.$$

Luego, $\lim_{s\to\infty} \sqrt[s]{\|TA^s\|} = \lim_{s\to\infty} \sqrt[s]{\|A^s\|}$. Similarmente se prueba que

$$\lim_{s\to\infty} \sqrt[s]{\|A^s\|} = \lim_{s\to\infty} \sqrt[s]{\|TA^sT^{-1}\|} = \lim_{s\to\infty} \sqrt[s]{\|J^s\|}. \qquad \square$$

Continuamos ahora con la demostración del teorema. Podemos suponer que $J = J_a(\lambda) \oplus J_1$ con $\|\lambda\| = \rho(A)$ y J_1 se encuentra en forma de Jordan. Y se obtiene que $J^s = J_a(\lambda)^s \oplus J_1^s$, donde $J_a(\lambda)^s$ tiene la forma descrita en (4.2). Llamaremos $J^s = (b_{ij}^{(s)})$, de manera que $b_{ii}^{(s)} = \lambda^s$ para $i = 1, \dots, a$; $s \in \mathbf{N}$.

En particular, $\|J^s\| = \left[\sum_{ij} |b_{ij}^{(s)}|^2\right]^{\frac{1}{2}} \geq \left[\sum_{i=1}^{a} \rho(A)^{2s}\right]^{\frac{1}{2}} = a^{\frac{1}{2}}\rho(A)^s$ y $\sqrt[s]{\|J^s\|} \geq a^{\frac{1}{2}}\rho(A)$, que nos da inmediatamente $\lim_{s\to\infty} \sqrt[s]{\|J^s\|} \geq \rho(A)$. La otra desigualdad es un poco más complicada.

En primer lugar consideremos el caso en que $\rho(A) = 0$. Entonces todos los valores propios de A son 0. En particular $J^n = 0$ y $\lim_{s\to\infty} \sqrt[s]{\|J^s\|} = 0$.

Podemos entonces suponer que $\rho(A) > 0$. Usando la expresión (4.2) para las potencias de los bloques de Jordan tenemos que para $1 \leq i < j \leq n$, y $s \geq n$,

$$|b_{ij}^{(s)}| \leq \binom{s}{j-i}\rho(A)^{s-j+i} \leq \begin{cases} \binom{s}{j-i}\rho(A)^s & \text{si } \rho(A) \geq 1, \\ \binom{s}{j-i}\rho(A)^{s-n} & \text{si } \rho(A) < 1. \end{cases}$$

De manera que

$$\|J^s\| = \left[\sum_{i,j} |b_{ij}^{(s)}|^2\right]^{\frac{1}{2}} \leq \rho(A)^s c^{\frac{1}{2}} \left[\sum_{i<j} \binom{s}{j-i}^2\right]^{\frac{1}{2}},$$

con $c = 1$ si $\rho(A) \geq 1$ y $c = \rho(A)^{-n}$ si $\rho(A) < 1$.

La expresión $\binom{s}{m}$ es un polinomio de grado $m \leq n$ en s, de modo que hay una constante d tal que $\binom{s}{m} \leq ds^n$ para toda $m \leq n$. Como $\lim_{s\to\infty} \sqrt[s]{s^n} = \lim_{s\to\infty} e^{\frac{n}{s}\log s} = 1$ y $\lim_{s\to\infty} \sqrt[s]{d} = 1$, obtenemos que para toda $\varepsilon > 0$ hay un número $N(\varepsilon)$ tal que si $s \geq N(\varepsilon)$ entonces

$$\sqrt[s]{\|J^s\|} \leq \rho(A)(1+\varepsilon).$$

En conclusión, $\lim_{s\to\infty} \sqrt[s]{\|J^s\|} \leq \rho(A)$ y $\lim_{s\to\infty} \sqrt[s]{\|A^s\|} = \lim_{s\to\infty} \sqrt[s]{\|J^s\|} = \rho(A)$. $\square$

4.5. Un importante caso especial de convergencia de sucesiones de matrices se presenta en las *series* de potencias de matrices. Dados escalares $(\lambda_s)_{s\in\mathbf{N}}$ y una matriz A, la serie $\sum_{s=0}^{\infty} \lambda_s A^s$ se llama una *serie de potencias* en A. Se dice que la serie $\sum_{s=0}^{\infty} \lambda_s A^s$ converge a la suma S si la sucesión de sumas parciales $\left(\sum_{s=0}^{m} \lambda_s A^s\right)_m$ converge a S.

Por ejemplo, consideremos la matriz 2×2

$$A = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{8} & 0 \end{bmatrix}.$$

Deseamos saber si $\sum_{s=0}^{\infty} A^s$ existe y, en caso afirmativo, cuál es la suma.

Calculemos las potencias:

$$A^{2s-1} = \begin{bmatrix} 0 & \frac{1}{2^{2s-1}} \\ \frac{1}{2^{2s+1}} & 0 \end{bmatrix} \quad \text{y} \quad A^{2s} = \begin{bmatrix} \frac{1}{2^{2s}} & 0 \\ 0 & \frac{1}{2^{2s}} \end{bmatrix}, \qquad s \in \mathbf{N}.$$

Como $\sum_{s=0}^{\infty} 2^{1-2s} = 2 \cdot \sum_{s=0}^{\infty} 4^{-s} = 2 \cdot \frac{1}{1-\frac{1}{4}} = \frac{8}{3}$,

$$\sum_{s=0}^{\infty} \frac{1}{2^{2s+1}} = \frac{1}{2} \cdot \sum_{s=0}^{\infty} 4^{-s} = \frac{2}{3} \quad \text{y} \quad \sum_{s=1}^{\infty} \frac{1}{2^{2s}} = \frac{1}{3}.$$

Obtenemos

$$\sum_{s=0}^{\infty} A^s = \begin{bmatrix} \frac{1}{3} & \frac{8}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}.$$

Un primer ejemplo de la importancia del cálculo de series está dado en el siguiente resultado:

Teorema. *Sea $A = (a_{ij})$ matriz compleja $n \times n$.*
a) *Si $\rho(A) < 1$, entonces la matriz $I_n - A$ es invertible y*

$$(I_n - A)^{-1} = \sum_{s=0}^{\infty} A^s.$$

b) *Si $\rho(A) < 1$ y todos los coeficientes $a_{ij} \geq 0$, entonces las entradas de $(I_n - A)^{-1}$ son también no negativas.*

Demostración: (a) Sea $S_m = \sum_{s=0}^{m} A^s$. Deseamos probar que la sucesión $(S_m)_m$ converge a $(I_n - A)^{-1}$.

Dada $\varepsilon > 0$ tal que $\rho(A) < 1 - \varepsilon < 1$. Por (4.3), encontramos $N(\varepsilon) \in \mathbf{N}$ tal que para $s \geq N(\varepsilon)$ se tiene

$$\|A^s\| \leq (1 - \varepsilon)^s.$$

Para $\ell \geq m \geq N(\varepsilon)^s$ obtenemos:

$$\|S_\ell - S_m\| = \left\| \sum_{s=m+1}^{\ell} A^s \right\| \leq \sum_{s=m+1}^{\ell} (1 - \varepsilon)^s <$$

$$(1 - \varepsilon)^{m+1} \sum_{s=0}^{\infty} (1 - \varepsilon)^s = \frac{(1 - \varepsilon)^{m+1}}{\varepsilon},$$

que podemos hacer tan pequeño como queramos cuando m es suficientemente grande. Entonces la sucesión $(S_m)_m$ es de Cauchy; como el espacio $\mathbf{C}^{n^2}$ es completo, tenemos que la sucesión $(S_m)_m$ converge, digamos a S.

Calculamos,

$$(I_n - A)S = \lim_{m \to \infty} (I_n - A)S_m = \lim_{m \to \infty} (I_n - A^{m+1})$$

$$= I_n - \lim_{m \to \infty} A^{m+1} = I_n.$$

Aquí nos basamos en que $\rho(A) < 1$ y en (4.3). Luego $S = (I_n - A)^{-1}$.

(b) Si $a_{ij} \geq 0$, también las entradas de la serie $\sum_{j=0}^{\infty} A^s$ tendrán que ser no negativas. □

Ejercicios

1. ¿Siempre que $I_n - A$ es invertible se puede calcular su inversa por medio de la fórmula $\sum_{s=0}^{\infty} A^s = (I_n - A)^{-1}$?

2. Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es una matriz 2×2 en $\mathbf{C}$ con $\rho(A) < 1$, calcule $(I_2 - A)^{-1}$ por medio de la serie $\sum_{s=0}^{\infty} A^s$.

4.6. Consideremos ahora el problema general de las series de potencias en una matriz A.

Recordemos que una serie formal de potencias $f(t) = \sum_{s=0}^{n} \mu_s t^s$ tiene *radio de convergencia* $r(f)$ si para todo número complejo α con $|\alpha| < r(f)$ se tiene que $\sum_{s=0}^{\infty} \mu_s \alpha^s$ converge (al número $f(\alpha)$) y $\sum_{j=0}^{\infty} \mu_s \alpha^s$ diverge para números α con $|\alpha| > r(f)$. En particular, se escribe, $r(f) = \infty$ cuando $f(\alpha)$ está bien definido para toda $\alpha \in \mathbf{C}$.

Teorema (Lagrange-Sylvester). *Sea $f(t) = \sum_{s=0}^{\infty} \mu_s t^s$ una serie de potencias con radio de convergencia $r(f)$. Sea A una matriz $n \times n$ con radio espectral $\rho(A)$. Supongamos que $\rho(A) < r(f)$, entonces $f(A) = \sum_{s=0}^{n} \mu_s A^s$ converge. Podemos además calcular explícitamente el valor de $f(A)$ de la siguiente manera:*

Sea $J = \bigoplus_{i=1}^{m} J_{n_i}(\lambda_i)$ la forma de Jordan de A, de modo que $A = TJT^{-1}$; por $f^{(j)}(t)$ denotamos la j-ésima derivada de $f(t)$ que también tiene radio de convergencia $r(f)$. Para cada $i = 1, \ldots, m$ consideremos la matriz

$$L_i = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{1}{2} f^{(2)}(\lambda_i) & \frac{1}{3!} f^{(3)}(\lambda_i) & \cdots & \frac{1}{(n_i-1)!} f^{(n_i-1)}(\lambda_i) \\ 0 & f(\lambda_i) & f'(\lambda_i) & \frac{1}{2} f^{(2)}(\lambda_i) & \cdots & \frac{1}{(n_i-2)!} f^{(n_i-2)}(\lambda_i) \\ & & \ddots & & & \vdots \\ & 0 & & & & f'(\lambda_i) \\ & & & & & f(\lambda_i) \end{bmatrix}.$$

Entonces, $f(A) = T\left(\bigoplus_{i=1}^{m} L_i\right) T^{-1}$.

Demostración: Consideremos el bloque de Jordan $J_{n_i}(\lambda_i)$. Como $|\lambda_i| < r(f)$, entonces $f^{(j)}(\lambda_i) = \sum_{s=j}^{\infty} \frac{s!}{(s-j)!} \mu_s \lambda_i^{s-j}$. Por otra parte el coeficiente $b_{j\ell}^{(s)}$ con $j \leq \ell$ de la matriz $J_{n_i}(\lambda_i)^s$ es

$$b_{j\ell}^{(s)} = \binom{s}{\ell - j} \lambda_i^{s-\ell+j},$$

de donde

$$\sum_{s=0}^{\infty} \mu_j b_{j\ell}^{(s)} = \sum_{s=0}^{\infty} \frac{s!}{(\ell-j)!(s-\ell+j)!} \mu_s \lambda_i^{s-\ell+j} = \frac{1}{(\ell-j)!} f^{(\ell-j)}(\lambda_i)$$

converge. En otras palabras, la serie $\sum_{s=0}^{\infty} \mu_s J_{n_i}(\lambda_i)^s$ converge a la matriz L_i. Por lo tanto, $f(A) = T\left(\bigoplus_{i=1}^{m} L_i\right) T^{-1}$ es una serie convergente. □

Ejercicios

1. Calcule $f(A)$ donde $f(t) = \sum_{s=0}^{\infty} \frac{1}{6^s} t^s$ y A es la matriz $A = J_4(3) \oplus J_2(1)$.

2. Demuestre que la serie $f(A) = \sum_{j=0}^{\infty} \mu_s A^s$ diverge si el radio de convergencia $r(f) < \rho(A)$.

3. Considere las series

$$\operatorname{sen} t = \sum_{s=0}^{\infty} (-1)^s \frac{1}{(2s+1)!} t^{2s+1}$$

$$\cos t = \sum_{s=0}^{\infty} (-1)^s \frac{1}{(2s)!} t^{2s}.$$

Muestre que para toda matriz A de tamaño $n \times n$, $\operatorname{sen} A$ y $\cos A$ están bien definidas. Se cumple además que

$$\operatorname{sen}^2 A + \cos^2 A = I_n,$$

$$\operatorname{sen} 2A = 2 \operatorname{sen} A \cos A.$$

4. ¿Cómo se recupera el Teorema (4.5) a partir de (4.6)?

5. Exponenciales de matrices

5.1. Una importante serie de matrices es la asociada a la *función exponencial*. La función $e^t = \sum_{s=0}^{\infty} \frac{1}{s!} t^s$ tiene radio de convergencia infinito.

Entonces para toda matriz A de tamaño $n \times n$ se tiene, por (4.6), que

$$e^A = \sum_{s=0}^{\infty} \frac{1}{s!} A^s$$

es una matriz bien definida. La matriz e^A se llama la *exponencial* de A. Algunas de las propiedades importantes de esta matriz están dadas en el siguiente lema.

Lema: a) *Si T es una matriz invertible, entonces $e^{TAT^{-1}} = Te^AT^{-1}$.*
b) *Si $AB = BA$, entonces $e^{A+B} = e^Ae^B$. En particular, e^A siempre es invertible y $(e^A)^{-1} = e^{-A}$.*
c) *Si $\lambda_1, \ldots, \lambda_n$ son los valores propios de A, entonces $e^{\lambda_1}, \ldots, e^{\lambda_n}$ son los valores propios de e^A.*
d) $\det e^A = e^{\mathrm{tr}\, A}$.

Demostración: Sea $S_m = \sum_{s=0}^{m} \frac{1}{s!} A^s$ la suma parcial tal que

$$\lim_{m\to\infty} S_m = e^A.$$

(a) Calculamos

$$Te^AT^{-1} = T \lim_{m\to\infty} S_m T^{-1} = \lim_{m\to\infty} TS_mT^{-1}$$
$$= \lim_{m\to\infty} \sum_{s=0}^{\infty} \frac{1}{s!} (TAT^{-1})^s = e^{TAT^{-1}}.$$

(b) Observemos que

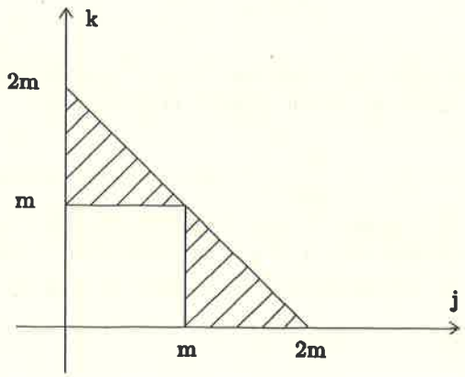
$$e^{A+B} - e^Ae^B = \lim_{m\to\infty} \left[\sum_{s=0}^{2m} \frac{1}{s!} (A+B)^s - \sum_{j=0}^{m} \frac{1}{j!} A^j \sum_{k=0}^{m} \frac{1}{k!} B^k \right]$$

y deseamos probar que este límite es cero. Calculamos el valor de la norma del m-ésimo término de esta diferencia,

$$\left\| \sum_{s=0}^{2m} \frac{1}{s!} (A+B)^s - \sum_{j=0}^{m} \frac{1}{j!} A^j \sum_{k=0}^{m} \frac{1}{k!} B^k \right\|$$
$$\underset{(1)}{=} \left\| \sum_{j=0}^{2m} \sum_{k=0}^{j} \frac{1}{j!} \binom{j}{k} A^k B^{j-k} - \sum_{j=0}^{m} \frac{1}{j!} A^j \sum_{k=0}^{m} \frac{1}{k!} B^k \right\|$$
$$= \left\| \sum_{j+k\le 2m} \frac{1}{j!} A^j \frac{1}{k!} B^k - \sum_{j=0}^{m} \frac{1}{j!} A^j \sum_{k=0}^{m} \frac{1}{k!} B^k \right\|$$
$$\underset{(2)}{=} \left\| \sum_{\substack{j>m \\ j+k\le 2m}} \frac{1}{j!} A^j \frac{1}{k!} B^k + \sum_{\substack{k>m \\ j+k\le 2m}} \frac{1}{j!} A^j \frac{1}{k!} B^k \right\|$$
$$\le \sum_{\substack{j>m \\ j+k\le 2m}} \frac{1}{j!} \|A\|^j \frac{1}{k!} \|B\|^k + \sum_{\substack{k>m \\ j+k\le 2m}} \frac{1}{j!} \|A\|^j \frac{1}{k!} \|B\|^k$$
$$\le e^{\|B\|} \sum_{j=m+1}^{\infty} \frac{1}{j!} \|A\|^j + e^{\|A\|} \sum_{k=m+1}^{\infty} \frac{1}{k!} \|B\|^k.$$

Como la suma $e^{\|A\|} = \sum_{j=0}^{\infty} \frac{1}{s!} \|A\|^s$ existe, entonces $\lim_{m\to\infty} \sum_{j=m+1}^{\infty} \frac{1}{j!} \|A\|^j = 0$. Resulta entonces que el último término en nuestras desigualdades tiende a cero, cuando m tiende a infinito.

La igualdad (1) se debe a que $AB = BA$. La igualdad (2) puede ilustrarse en el siguiente dibujo:



En particular, tenemos que $e^A e^{-A} = e^0 = I_n$.

(c) Sea J la forma de Jordan de A y T una matriz invertible de manera que

$$TAT^{-1} = J = (b_{ij}).$$

Las entradas de la diagonal $b_{ii} = \lambda_i$ son los valores propios de A. Para la exponencial obtenemos

$$Te^A T^{-1} = e^{TAT^{-1}} = e^J = (c_{ij});$$

que tiene en la diagonal los valores $c_{ii} = e^{b_{ii}} = e^{\lambda_i}$. Luego $\{e^{\lambda_i}: i = 1, \ldots, n\}$ son los valores propios de $Te^A T^{-1}$ y también los de e^A.

(d) Sabemos que

$$\det e^A = \prod_{i=1}^{n} e^{\lambda_i} = e^{(\sum_{i=1}^{n} \lambda_i)} = e^{\operatorname{tr} A}. \qquad \square$$

Ejercicios

a) Considérense las matrices

$$A = \begin{pmatrix} 2\pi i & 1 \\ 0 & 0 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 2\pi i & -1 \\ 0 & 0 \end{pmatrix}.$$

Muestre que $e^A e^B = e^{A+B} = e^B e^A$, a pesar de que $AB \neq BA$.

b) Encuentre matrices 2×2, A y B que satisfagan

$$e^A e^B \neq e^{A+B}.$$

5.2. Calcular una serie de potencias como e^A puede parecer tedioso y complicado. Sin embargo, es en general muy simple.

Proposición. *Sea* $f(t) = \sum_{s=0}^{\infty} \mu_s t^s$ *una serie de potencias con radio de convergencia* $r(f)$ *y sea* A *una matriz* $n \times n$ *con* $\rho(A) < r(f)$. *Sea* $m_A(t)$ *el polinomio minimal de* A *y supongamos que tiene grado* $m \leq n$. *Entonces existe un polinomio* $b(t)$ *de grado a lo más* $m-1$, *de manera que:*

a) $f(A) = b(A)$.

b) *Para cada valor propio* λ *de* A, *definimos el grado de* λ *como*

$$d(\lambda) = \max\{s: \nu(\lambda, s) > 0\}.$$

Entonces para toda $0 \leq j < d(\lambda)$, *se tiene*

$$f^{(j)}(\lambda) = b^{(j)}(\lambda).$$

Demostración: Sea V el subespacio vectorial de $\mathbf{C}^{n^2}$ generado por las matrices $I_n, A, A^2, \ldots, A^{m-1}$. Si $p(t) \in \mathbf{C}[t]$ es un polinomio, por el algoritmo de Euclides podemos escribirlo como

$$p(t) = a(t)m_A(t) + c(t),$$

donde $c(t)$ es un polinomio de grado a lo más $m-1$. Por el Teorema de Cayley-Hamilton, tenemos $p(A) = b(A) \in V$. En particular, para toda m, $\sum_{s=0}^{m} \mu_s A^s \in V$.

Observemos que con la topología inducida por la norma euclidiana, el espacio V es un conjunto cerrado de $\mathbf{C}^{n^2}$. Entonces

$$f(A) = \lim_{m \to \infty} \sum_{s=0}^{m} \mu_s A^s \in V.$$

Por la definición de V, existe un polinomio $b(t)$ de grado a lo más $m-1$, tal que $f(A) = b(A)$. Sea J la forma de Jordan de A y $A = TJT^{-1}$. Podemos definir las matrices $L_i(f)$ (resp. $L_i(b)$) como en (4.7) de modo que

$$f(A) = T\left(\bigoplus_{i=1}^{s} L_i(f)\right)T^{-1} \quad \text{y} \quad b(A) = T\left(\bigoplus_{i=1}^{s} L_i(b)\right)T^{-1}.$$

Tenemos entonces que $f^{(j)}(\lambda) = b^{(j)}(\lambda)$ para todo valor propio λ de A y $0 \leq j < d(\lambda)$. $\square$

Ejemplo: Deseamos calcular la matriz e^A, donde

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

En primer lugar requerimos calcular el polinomio minimal $m_A(t)$ de A. Primero observemos que el polinomio $m_A(t)$ tiene cuando menos grado 3. De otra manera habría una ecuación

$$A^2 = \mu_0 I_4 + \mu_1 A, \qquad \text{con} \qquad \mu_0, \mu_1 \in \mathbf{C}.$$

En ese caso, comparando los primeros renglones de las matrices tendríamos

$$(-6, 4, -4, 10) = (\mu_0 - 2\mu_1, 2\mu_1, -2\mu_1, 4\mu_1),$$

lo que no tiene solución.

Intentando una ecuación de la forma

$$A^3 = \mu_0 I_4 + \mu_1 A + \mu_2 A^2,$$

tenemos

$$(-12, 6, -6, 20) = (\mu_0 - 2\mu_1 - 6\mu_2, 2\mu_1 + 4\mu_2, 4\mu_1 + 10\mu_2)$$

que tiene una solución $\mu_0 = 2$, $\mu_1 = -5$ y $\mu_2 = 4$. Si en efecto calculamos $m(A)$ para el polinomio $m(t) = t^3 - 4t^2 + 5t - 2$, tenemos $m(A) = 0$. Luego, $m_A(t) = m(t) = (t-1)^2(t-2)$.

Deseamos ahora encontrar un polinomio $b(t) = b_0 + b_1 t + b_2 t^2$ que satisfaga:

$$\left.\begin{aligned} b_0 + b_1 + b_2 &= b(1) = e \\ b_1 + 2b_2 &= b'(1) = e \\ b_0 + 2b_1 + 4b_2 &= b(2) = e^2 \end{aligned}\right\}.$$

Este sistema tiene la solución: $b_0 = e^2 - 2e$, $b_1 = -2e^2 + 5e$, $b_2 = e^2 - 2e$.

Luego, $e^A = b(A) = b_0 + b_1 A + b_2 A^2$, que puede calcularse en forma muy simple.

Ejercicios

1. Calcule e^A para A como sigue:

$$\begin{bmatrix} 3 & 0 & 1 & 0 \\ 0 & 2 & -1 & 0 \\ 1 & 0 & 1 & 0 \\ -2 & 1 & -1 & 2 \end{bmatrix}.$$

2. Calcule la matriz $\cos 2\pi A$ para A como en el ejemplo 1.

3. Evalúe $\tan^{-1} A^3$, para $A = \begin{bmatrix} -1 & 2 & -2 & 2 \\ -2 & 2 & -1 & 3 \\ -1 & 0 & 1 & 2 \\ -1 & 1 & -1 & 2 \end{bmatrix}$.

4. Sea $A = J_s(\lambda)$; demuestre que

$$e^{tA} = e^{t\lambda}\left(I + tN + \frac{1}{2}(tN)^2 + \ldots + \frac{1}{(s-1)!}(tN)^{s-1}\right)$$

donde N es la matriz $s \times s$ $\begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix}$.

5.3. Sea A una matriz $n \times n$. Para cualquier número real t, la matriz e^{tA} está bien definida, de modo que determina una función

$$\mathbf{R} \to \mathbf{C}^{n^2}, \qquad t \mapsto e^{tA},$$

que llamaremos también e^{tA}.

Esta función exponencial tiene propiedades similares a la exponencial $e^t: \mathbf{R} \to \mathbf{R}$, como veremos a continuación:

Lema. *La función e^{tA} es diferenciable y*

$$\frac{d}{dt} e^{tA} = Ae^{tA}.$$

Demostración: En efecto, podemos calcular la derivada de la función e^{tA} según la definición clásica del análisis:

$$\frac{e^{(t+\Delta t)A} - e^{tA}}{\Delta t} = e^{tA}\frac{(e^{\Delta tA} - I_n)}{\Delta t} = Ae^{tA} + \lim_{m\to\infty}\left(\sum_{s=2}^{m}\frac{1}{s!}(\Delta t)^{s-1}A^s\right).$$

El último sumando podemos acotarlo para $|\Delta t| \leq 1$:

$$\left\|\sum_{s=2}^{m}\frac{1}{s!}(\Delta t)^{s-1}A^s\right\| \leq |\Delta t|\sum_{s=2}^{m}\frac{1}{s!}\|A^s\| \leq |\Delta t|e^{\|A\|}.$$

Luego, $\lim_{\Delta t\to 0}\frac{e^{(t+\Delta t)A}-e^{tA}}{\Delta t} = Ae^{tA}$, lo que prueba el resultado. □

5.4. Una de las aplicaciones más importantes de la exponencial de matrices se da en la solución de ecuaciones diferenciales. Para ese fin la ecuación (5.3) es básica.

Problema fundamental: Sea $A = (a_{ij})$ una matriz compleja $n \times n$ y números b_i, c_i $(i = 1, \ldots, n)$. Se buscan funciones diferenciables

$$y_i(t): \mathbf{R} \to \mathbf{R} \qquad (i = 1, \ldots, n)$$

que satisfagan

$$y_i'(t) = \sum_{j=1}^{n} a_{ij}y_i(t) + c_i,$$

$$y_i(0) = d_i, \qquad i = 1, \ldots, n.$$

Decimos entonces que $y(t) = (y_i(t))$ es una *solución* del *sistema de ecuaciones diferenciales ordinarias*

$$y'(t) = Ay(t) + c, \text{ con } c = (c_i),$$

que satisface las condiciones iniciales $y(0) = d$, con $d = (d_i)$.

El sistema se llama *homogéneo* si $c = 0$.

Hay muchos problemas en física, química, biología y otras ciencias en que aparecen sistemas de ecuaciones de la forma anterior. Al final de la sección consideraremos algunos ejemplos explícitos.

Por lo pronto nos ocuparemos del problema de la existencia y unicidad de las soluciones del sistema, así como de algunas de las propiedades básicas de las soluciones.

Teorema (Existencia y unicidad de soluciones para los sistemas homogéneos de ecuaciones diferenciales ordinarias). *Sea $y'(t) = Ay(t)$ un sistema de ecuaciones con condición inicial $y(0) = d$. Este sistema tiene exactamente una solución que es $y(t) = e^{tA}d$.*

Demostración. Unicidad: Sean $v(t), w(t)$ dos soluciones del sistema, esto es, $v'(t) = Av(t)$, $w'(t) = Aw(t)$ y $v(0) = d = w(0)$. Considerando la diferencia $z(t) = v(t) - w(t)$, vemos que $z(t)$ es solución del sistema $y'(t) = Ay(t)$ con condiciones iniciales $y(0) = 0$.

Deseamos probar que $z(t) = 0$ para toda $t \in \mathbf{R}$.

Tomamos cualquier intervalo compacto I de $\mathbf{R}$ con $0 \in I$. Como $z_i(t)$ es diferenciable, continua en $\mathbf{R}$ y acotada en I, sea M un número tal que

$$|z_i(t)| \leq M \text{ para toda } t \in I, \text{ para } i = 1, \ldots, n.$$

Elijamos a como el máximo de los valores absolutos $|a_{ij}|$ $(1 \leq i, j \leq n)$ de las entradas de la matriz A. Para cualquier $t \in I$, tenemos entonces

$$|z_i(t)| = |z_i(t) - z_i(0)| = \left|\int_0^t z_i'(t)\right| = \left|\int_0^t \sum_{j=1}^{n} a_{ij}z_j(t)dt\right|$$

$$\leq \sum_{j=1}^{n}|a_{ij}|\left|\int_0^t |z_j(t)|dt\right| \leq \sum_{j=1}^{n} aM\left|\int_0^t dt\right| = naM|t|.$$

Supongamos inductivamente que $|z_i(t)| \leq \frac{M(na|t|)^r}{r!}$ para toda $t \in I$, $i = 1, \ldots, n$. Obtenemos entonces

$$|z_i(t)| = \left|\int_0^t z_i'(t)\right| \leq \sum_{j=1}^{n} a\frac{M(na)^r}{r!}\left|\int_0^t |t|^r dt\right| = \frac{M(na|t|)^{r+1}}{(r+1)!}.$$

Luego, por inducción, obtenemos que para toda $t \in I$, $i = 1, \ldots, n$:

$$|z_i(t)| \leq \left|\lim_{r\to\infty} M\frac{(na|t|)^r}{r!}\right| = 0.$$

Como I es arbitrario, obtenemos finalmente que $z(t) = 0$, para $t \in \mathbf{R}$.
Existencia: Mostraremos que $y(t) = e^{tA}d$ es una solución del sistema de ecuaciones dado. Diferenciando respecto a t, tenemos

$$y'(t) = \frac{dy(t)}{dt} = \frac{de^{tA}}{dt}d = Ae^{tA}d = Ay(t).$$

Además, $y(0) = e^0 d = I_n \cdot d = d$. □

Ejercicio: Encuentre la solución del sistema de ecuaciones

$$\begin{aligned} y_1' &= 2y_1 + y_2 + y_4 \\ y_2' &= y_1 + 3y_2 - y_3 + 3y_4 \\ y_3' &= y_2 + 2y_3 + y_4 \\ y_4' &= y_1 - y_2 - y_3 - y_4 \end{aligned}$$

sujeto a las condiciones iniciales $y(0) = (1,0,0,0)$.

5.5. Una vez calculada la solución de un sistema homogéneo de ecuaciones diferenciales, es sencillo calcular soluciones para el sistema no homogéneo.

Consideremos el sistema de ecuaciones diferenciales no homogéneo $y'(t) = Ay(t) + c$ con condición inicial $y(0) = d$.

Supongamos que tenemos una solución $w(t)$ del sistema $y'(t) = Ay(t) + c$, sin especificar ninguna condición inicial. Sea $v(t)$ la solución del sistema $y'(t) = Ay(t)$ con condición inicial $y(0) = d - w(0)$. Entonces la única solución de nuestro sistema no homogéneo es $y(t) = v(t) + w(t)$.

En primer lugar observemos que hemos dicho la *única solución*. En efecto, es fácil ver que la prueba de la unicidad en (5.4) no depende de la homogeneidad del sistema. Calculamos

$$\begin{aligned} y'(t) &= v'(t) + w'(t) = Av(t) + Aw(t) + c = Ay(t) + c, \\ y(0) &= v(0) + w(0) = d. \end{aligned}$$

Luego, para resolver nuestra ecuación homogénea, basta con calcular una solución sin condición de frontera de la ecuación $y'(t) = Ay(t) + c$. Para este fin, consideremos la forma de Jordan J de A, tal que $A = T^{-1}JT$ y $J = \bigoplus_{i=1}^{s} J_{n_i}(\lambda_i)$. Observemos que si $v(t)$ es una solución de la ecuación $y'(t) = Jy(t) + Tc$, entonces $y(t) = T^{-1}v(t)$ satisface:

$$y'(t) = T^{-1}v'(t) = T^{-1}Jv(t) + c = Ay(t) + c.$$

Entonces nos será suficiente calcular soluciones para los sistemas de ecuaciones

$$y'(t) = J_s(\lambda)y(t) + c.$$

Esto es sencillo; distinguimos dos casos:

(a) Si $\lambda \neq 0$, hacemos $y(t) = -J_s(\lambda)^{-1}c$ que, por ser una constante, satisface $y'(t) = 0$.

(b) Si $\lambda = 0$, hacemos $y_s(t) = c_s t$

$$y_{s-1}(t) = \frac{c_s}{2}t^2 + c_{s-1}t$$

$$\vdots$$

$$y_1(t) = \sum_{j=1}^{s} c_j \frac{t^j}{j!}$$

y el vector $y(t) = (y_1(t), \ldots, y_s(t))$ es una solución del sistema.

Ejercicios

1. Resuelva el sistema

$$\begin{aligned} y_1' &= y_1 + 2y_2 + 1 \\ y_2' &= 3y_1 + 2y_2 - 1 \end{aligned}$$

con condición inicial $y(0) = (1,0)$.

2. Intente resolver el sistema

$$\begin{aligned} y_1' &= y_1 + 2y_2 + e^t \\ y_2' &= 3y_1 + 2y_2 - e^{-t}. \end{aligned}$$

Observe que éste es un sistema de la forma $y' = Ay + c(t)$, donde $c(t)$ no es constante.

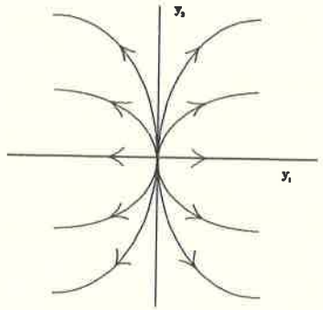
5.6. Consideremos explícitamente las soluciones del sistema $y'(t) = Ay(t)$ con condición de frontera $y(0) = d$, y donde A es una matriz 2×2. Supondremos también que $d \in \mathbf{R}^2$.

Hemos visto que por medio de una transformación invertible T podemos llevar este sistema a otro de la forma $v' = Jv(t)$, $v(0) = e$, donde $J = TAT^{-1}$ es una forma de Jordan. Por tanto no perdemos generalidad si suponemos que A está en su forma de Jordan —de hecho, consideraremos la forma de Jordan real. Sean λ_1, λ_2 los valores propios de A. Distinguimos varios casos:

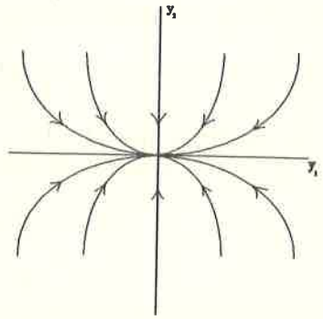
(i) $\lambda_1, \lambda_2 \in \mathbf{R}$ y $\lambda_1 \neq \lambda_2$. La solución del sistema es entonces

$$(y_1(t), y_2(t)) = e^{t\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}}(d_1, d_2) = (d_1 e^{\lambda_1 t}, d_2 e^{\lambda_2 t}).$$

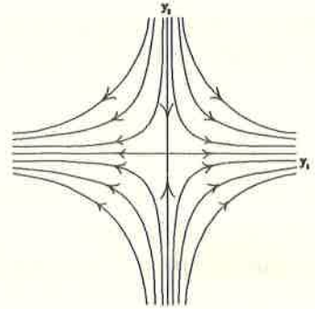
Observemos que entonces $y_1(t) = K y_2(t)^{\lambda_2/\lambda_1}$, para una constante $K = d_2^{\lambda_2}/d_1^{\lambda_1}$. Podemos representar las parejas $(y_1(t), y_2(t))$ en el plano (y_1, y_2) como trayectorias que dependen de t (y por supuesto, del valor de K). En la figura 1, dibujamos la forma que toma la gráfica dependiendo de las diferentes relaciones numéricas entre λ_1 y λ_2. La flecha en cada trayectoria indica la dirección en que t aumenta.



$\lambda_1 > \lambda_2 > 0$



$\lambda_2 < \lambda_1 < 0$



$\lambda_2 < 0 < \lambda_1$

Figura 1.

En el caso $\lambda_1 > \lambda_2 > 0$, observemos que cada trayectoria "empieza" en el origen, esto es, $\lim_{t \to -\infty} y_1(t) = 0$ para $i = 1, 2$. Además,

$$\frac{dy_2}{dy_1} = K y_1^{(\lambda_2/\lambda_1 - 1)}$$

y entonces $\lim_{y_1 \to \infty} \left(\frac{dy_2}{dy_1}\right) = \infty$. Un análisis similar puede hacerse en el caso en que $\lambda_2 < \lambda_1 < 0$. Observemos que en este segundo caso

$$\lim_{t \to \infty} y(t) = (0, 0).$$

En tal caso se dice que las soluciones $y(t)$ de la ecuación son *estables*.

Ejercicio: (i) Analice el caso $\lambda_1 = \lambda_2 \neq 0$. ¿En qué caso las soluciones son estables?

(ii) Los valores propios son complejos (no reales). Los valores propios son de la forma $\lambda_1 = \alpha + \beta i$ y $\lambda_2 = \alpha - \beta i$. Tomando las partes real e imaginaria podemos reescribir el sistema de ecuaciones de la siguiente manera:

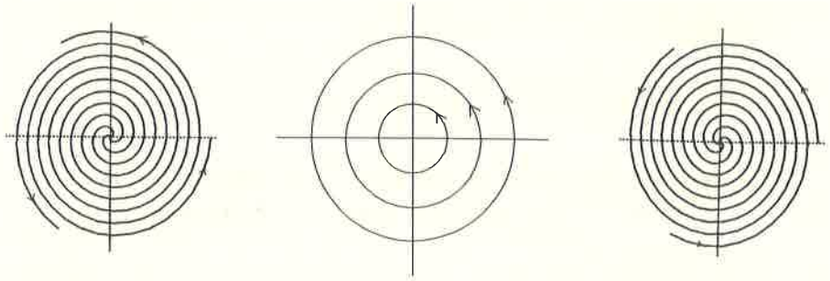
$$y_1'(t) = \alpha y_1(t) - \beta y_2(t)$$
$$y_2'(t) = \beta y_1(t) + \alpha y_2(t).$$

Podemos hacer el cambio de coordenadas $y_1(t) = r(t)\cos\theta(t)$, $y_2(t) = r(t)\text{sen}\,\theta(t)$. Obtenemos entonces el sistema en *forma polar*

$$r'(t) = \alpha\, r(t)$$
$$\theta'(t) = \beta,$$

con soluciones $r(t) = r_0 e^{\alpha t}$, $\theta(t) = \beta t + \theta_0$.

Los diagramas de fase que se obtienen son como en la figura 2.



$\alpha > 0$ $\alpha = 0$ $\alpha < 0$

Figura 2.

Ejercicio: Analice el sistema obtenido en el caso (ii) y muestre que los diagramas de fase son los dados en la figura 2.

5.7. Consideremos otra vez el sistema homogéneo

$$y'(t) = Ay(t).$$

Se dice que 0 es una *solución estable del sistema* (o simplemente, que el *sistema es estable*) si para todo vector d, la solución $v(t) = e^{At}d$ del sistema tiene la propiedad de que $\lim_{t\to\infty} v(t) = 0$.

En los ejemplos (5.6) las soluciones eran estables cuando la parte real de los valores propios de A era negativa. Este hecho es cierto en general.

Teorema (Criterio de Lyapunov). *El sistema $y'(t) = Ay(t)$ es estable si y solamente si la parte real de cada valor propio de A es negativa.*

Demostración: Supongamos que la parte real $\text{Re}(\lambda)$ de cada valor propio λ de A es negativa. Consideremos la forma de Jordan $J = TAT^{-1}$ de A y supongamos $J = \bigoplus_{i=1}^{s} J_{n_i}(\lambda_i)$. Por (5.1) es suficiente demostrar que $\lim_{t\to\infty} e^{tJ_{n_i}(\lambda_i)}d = 0$ para todo vector d. Escribamos $J_{n_i}(\lambda_i) = \lambda_i I_{n_i} + N$ con N nilpotente, entonces, según hemos visto en (5.2),

$$e^{tJ_{n_i}(\lambda_i)} = e^{t\lambda_i}\left[I + tN + \frac{1}{2!}(tN)^2 + \cdots + \frac{1}{(n_i-1)!}(tN)^{n_i-1}\right].$$

Entonces para comprobar el resultado basta con ver que $\lim_{t\to\infty} e^{t\lambda_i} = 0$, pero $|e^{t\lambda_i}| = e^{t\text{Re}(\lambda_i)}$ y claramente $\lim_{t\to\infty} e^{t\text{Re}(\lambda_i)} = 0$ ya que $\text{Re}(\lambda_i) < 0$.

Para el converso, supongamos que las soluciones de nuestro sistema son estables. Procediendo como antes obtenemos que $\lim_{t\to\infty} e^{t\lambda} = 0$ para todo valor propio λ de A. La norma del número complejo $e^{t\lambda}$ es $e^{t\text{Re}(\lambda)}$ que también tiende a 0 cuando $t \to \infty$. Claramente esto sólo es posible si $\text{Re}(\lambda) < 0$, para toda $\lambda \in \text{Spec}(A)$. □

5.8. En la práctica no es una tarea sencilla encontrar los valores propios de una matriz $n \times n$ con n grande. Es por ello que se aplican otros métodos para verificar algunas propiedades espectrales de matrices grandes, por ejemplo, la estabilidad del sistema de ecuaciones $y'(t) = Ay(t)$ asociado a la matriz A.

Obtendremos un criterio de estabilidad diferente también establecido por Lyapunov. Para ello introduciremos primero algunos conceptos.

Una matriz M de tamaño $n \times n$ se llama *simétrica* si coincide con su transpuesta, esto es $M^T = M$. Una matriz simétrica M se dice *positiva definida* si $x^*Mx > 0$ para todo vector $0 \neq x \in \mathbf{C}^n$. Aquí x^* denota el vector renglón $(\bar{x}_1, \ldots, \bar{x}_n)$ donde $\bar{x}_i$ es el conjugado de la i-ésima entrada del vector columna x. Además, note que estamos suponiendo implícitamente que x^*Mx es un número real (si no, ¿qué sentido tiene $x^*Mx > 0$?).

Lema. a) *I_n es una matriz positiva definida.*
b) *Si M es real simétrica, entonces todos sus valores propios son reales.*
c) *Sea M una matriz positiva definida, entonces todos sus valores propios son reales positivos.*

d) *Una matriz real simétrica M es positiva definida si y sólo si $x^T Mx > 0$ para todo vector $0 \neq x \in \mathbf{R}^n$.*

Demostración: (a) Para $0 \neq x \in \mathbf{C}^n$, tenemos

$$x^* I_n x = x^* x = \sum_{i=1}^{n} x_i^* x_i = \sum_{i=1}^{n} |x_i|^2 > 0.$$

(b) Sea λ un valor propio de M y $Mx = \lambda x$ para algún vector $0 \neq x \in \mathbf{C}^n$. Luego, $x^* Mx = \lambda x^* x = \lambda \left(\sum_{i=1}^{n} |x_i|^2 \right)$. Por otra parte, $x^* Mx$ es un número complejo con conjugado,

$$(x^* Mx)^* = x^* M^* x = x^* Mx.$$

Entonces, $x^* Mx$ es un número real, lo que implica que $\lambda \in \mathbf{R}$.

(c) Supongamos que M es positiva definida y $\lambda \in \mathrm{Spec}(A)$. Como antes

$$0 < x^* Mx = \lambda \|x\|^2$$

para un vector propio x de M. Esto muestra que $\lambda > 0$.

(d) Supongamos que $x^T Mx > 0$ para todo vector $0 \neq x \in \mathbf{R}^n$. Sea $0 \neq x \in \mathbf{C}^n$. Podemos escribir $x = x_1 + ix_2$ con $x_1, x_2 \in \mathbf{R}^n$. Entonces

$$\begin{aligned}(x_1 + ix_2)^* M(x_1 + ix_2) &= x_1^T Mx_1 - ix_2^T Mx_1 + ix_1^T Mx_2 + x_2^T Mx_2 \\ &= x_1^T Mx_1 + x_2^T Mx_2 > 0\end{aligned}$$

ya que el número (= matriz 1×1)

$$x_2^T Mx_1 = (x_2^T Mx_1)^T = x_1^T M^T x_2 = x_1^T Mx_2. \qquad \square$$

Ejercicio

En el capítulo III probaremos que una matriz real simétrica es diagonalizable. Suponga por lo pronto este resultado y demuestre lo siguiente: *Criterio de Jacobi:* Sea $M \in M_n(\mathbf{R})$ una matriz simétrica. Para $i = 1, \ldots, n$, consideremos las submatrices principales $M_i = M\binom{1,\ldots,i}{1,\ldots,i}$ de M (observe que $M_n = M$). Entonces M es positiva definida si y solamente si $\det M_1, \ldots, \det M_n$ son números positivos.

5.9. Teorema (Lyapunov). *La matriz real A de tamaño $n \times n$ determina un sistema*

$$y'(t) = Ay(t)$$

estable si y solamente si existe una matriz real positiva definida M que satisface

$$A^T M + MA = -I_n.$$

Demostración: Supongamos primero que el sistema $y'(t) = Ay(t)$ es estable. Consideremos la transformación lineal $f: \mathbf{R}^{n^2} \to \mathbf{R}^{n^2}$, $N \mapsto A^t N + NA$. Mostraremos que f es inyectiva: supongamos que $f(N) = 0$. Entonces $-A^T N = NA$ y $e^{-tA^T} N = Ne^{tA}$ (ver ejercicio 1). Luego, $N = e^{tA^T} Ne^{tA}$.

Como A y A^T tienen los mismos valores propios, (5.7) nos dice que también el sistema $y'(t) = A^T y(t)$ es estable. Entonces

$$N = \lim_{t \to \infty} e^{tA^T} Ne^{tA} = 0.$$

Luego, existe una única matriz $M \in \mathbf{R}^{n^2}$ con $A^T M + MA = f(M) = -I_n$. Como también $f(M^T) = -I_n$, entonces $M = M^T$ y M es simétrica. Nos falta ver que M es positiva definida. Por (5.8), basta con considerar vectores $0 \neq x \in \mathbf{R}^n$ y demostrar que $x^T Mx > 0$. Sea $0 \neq x_0 \in \mathbf{R}^n$ y consideremos la función $q: \mathbf{R} \to \mathbf{R}$, $t \mapsto (e^{At} x_0)^T M(e^{At} x_0)$ que tiene derivada

$$q'(t) = (e^{At} x_0)^T (A^T M + MA)(e^{At} x_0) = -\|e^{At} x_0\|^2 < 0 \text{ para } t \in \mathbf{R}.$$

En particular, $q(t) \leq q(1) < q(0) = x_0^T Mx_0$, para $t \geq 1$. Como $\lim_{t \to \infty} q(t) = 0$, entonces $x_0^T Mx_0 > 0$ y M es positiva definida.

Para el converso, supongamos que $A^T M + MA = -I_n$. Sea λ un valor propio de A y $Av = \lambda v$ para algún vector $0 \neq v \in \mathbf{R}^n$. Consideremos otra vez la función $q(t) = (e^{\lambda t} v)^* M(e^{\lambda t} v)$. Diferenciando tenemos

$$q'(t) = \lambda^* q(t) + \lambda q(t) = 2\mathrm{Re}(\lambda) q(t), \text{ para toda } t \in \mathbf{R}.$$

Como antes (ejercicio 1), $e^{At}v = e^{\lambda t}v$, entonces $q(t) = (e^{At}v)^* M(e^{At}v)$ que tiene derivada negativa. Así,

$$0 > q'(t) = 2\mathrm{Re}(\lambda)q(0) = 2\mathrm{Re}(\lambda)v^* Mv.$$

Como M es positiva definida, $v^*Mv > 0$ y tenemos que $\mathrm{Re}(\lambda) < 0$, como deseábamos. □

Ejercicios

1. Si $BN = NA$ para ciertas matrices cuadradas A, B y N, entonces $e^{Bt}N = Ne^{At}$.

2. Encuentre condiciones en las entradas de una matriz real $A \in M_2(\mathbf{R})$ para que el sistema $y'(t) = Ay(t)$ sea estable.