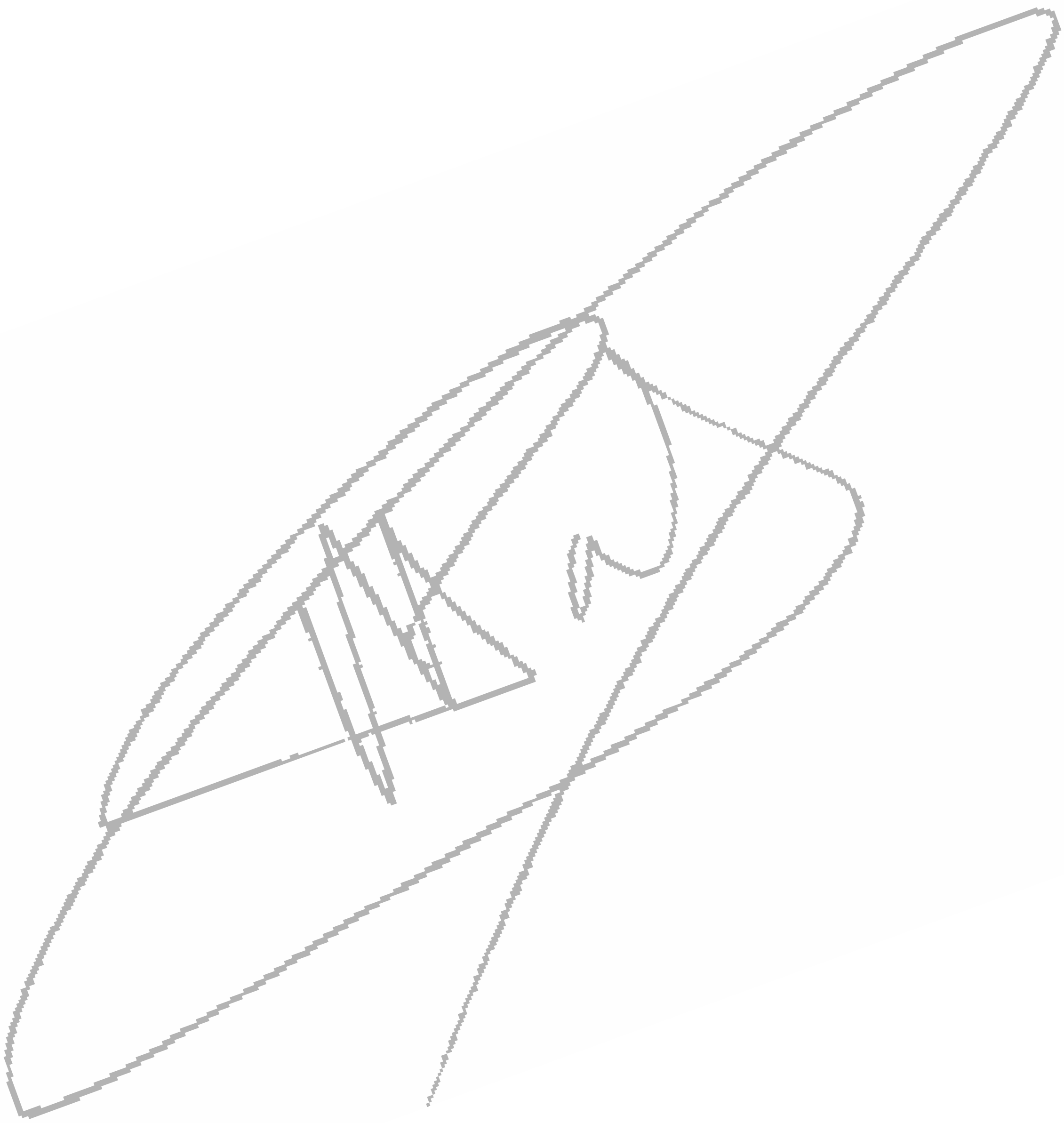
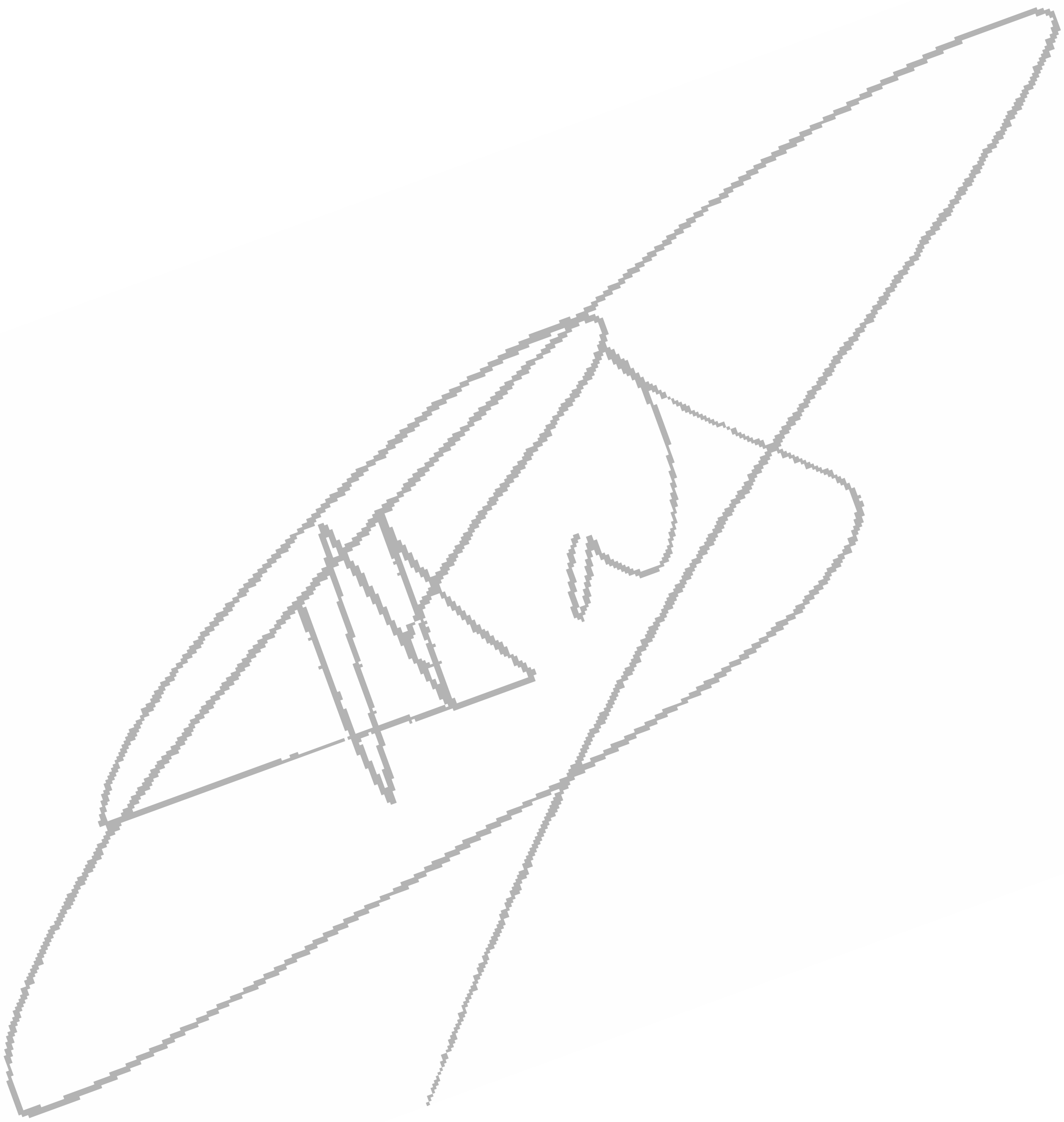
6. Forma canónica de matrices

Manuel Palacios
Departamento de Matemática Aplicada
Centro Politécnico Superior
Universidad de Zaragoza

Otoño 2010



Contents



Bibliography

[1] Merino, L. y Santos, E.: Álgebra lineal con métodos elementales. *Thomson*. 2006.

[2] Palacios, M.: Enunciados de ejercicios de Matemáticas II, grado Ing. Tecn. Industr. http://pcmap.unizar.es/∼mpala. 2010

[3] Villa, A. de la: Problemas de Álgebra. *CLAGSA*. 1988.

[4] Arvesu, J.; Marcellán, F. y Sánchez J.: Problemas resueltos de Álgebra lineal. Paso a paso. *Thomson*. 2005.

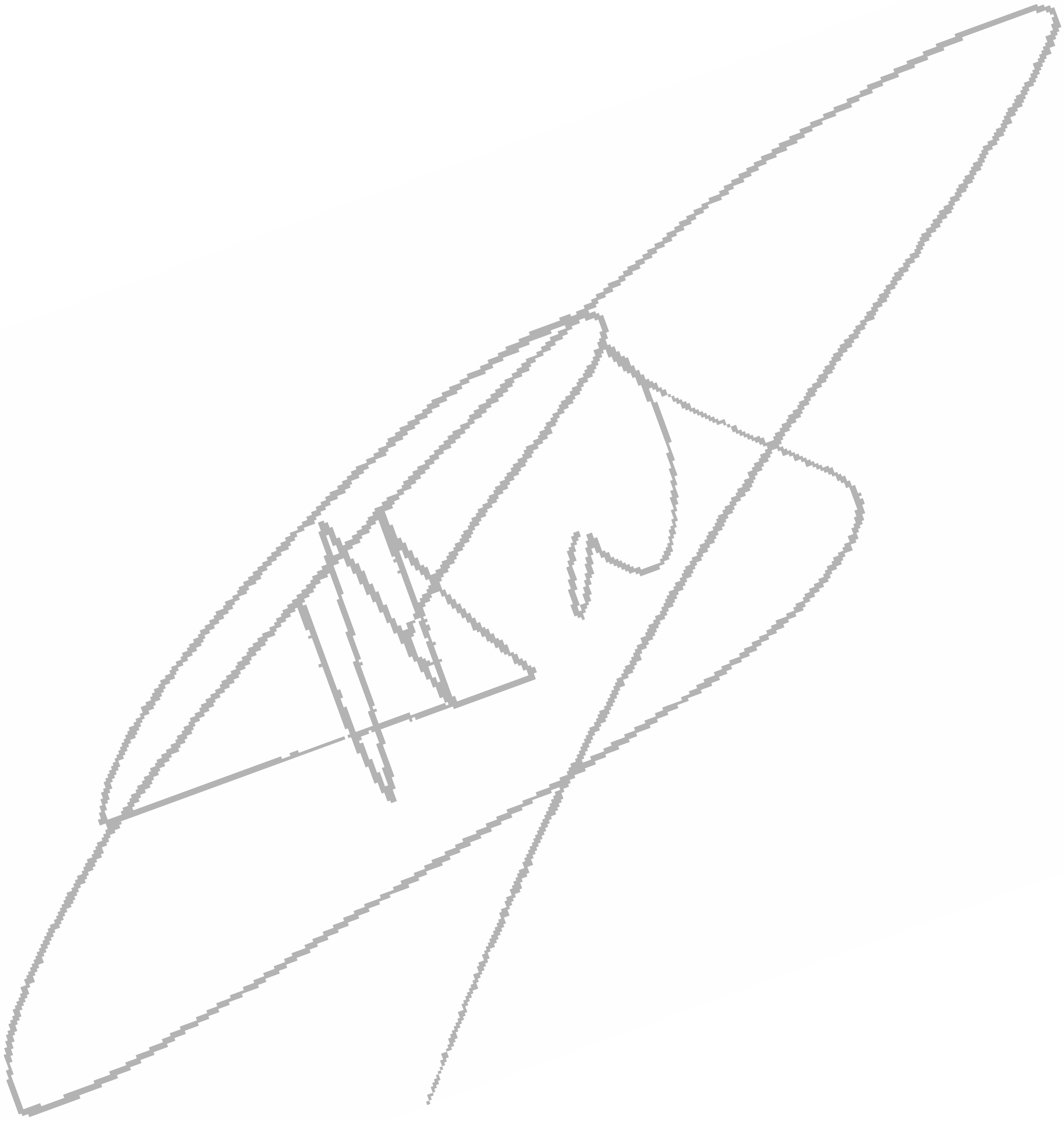
[5] Rojo, J.: Álgebra lineal. *Vector Ediciones, Madrid*. 1989.

[6] Rojo, J. y Martín, I.: Ejercicios y problemas de Álgebra lineal. *Editorial AC, Madrid*. 1989

[7] Burgos, J. de: Álgebra lineal y geometría cartesiana. *McGraw-Hill*. 2000.

[8] Burgos, J. de: Álgebra lineal. *McGraw-Hill*. 1993.

[9] Diego, B. de et al. : Problemas de Álgebra lineal (2^a ed.). *Deimos*. 1986.



Chapter 6

6. Forma canónica de matrices

6.1 Introducción.

Sobre estructura del espacio y endomorfismos

Recordemos que dado un endomorfismo de V y fijada una base del espacio vectorial existe una única matriz cuadrada coordenada asociada y recíprocamente. Nos interesa encontrar una matriz coordenada suficientemente sencilla que manifieste claramente la "estructura" de un endomorfismo. Es decir, "mirar" al espacio vectorial como si desde el punto de vista del arquitecto se tratara cuando mira un edificio, fijándose en las lineas maestras: columnas, cruceros, dinteles, etc. En nuestro caso, esas lineas maestras van a definir una base del espacio V que nos permitirá descomponer el espacio V como suma directa de subespacios invariantes.

Así pues, en este capítulo vamos a tratar de localizar una base de un espacio vectorial V en la que un endomorfismo h de V tenga la forma más sencilla (la forma canónica de Jordan) y discutiremos cúando ésta es diagonal.

Los objetos comunes para todo el capítulo serán los siguientes:

Un espacio vectorial V sobre el cuerpo K ($= \mathbb{R}$ o $\mathbb{C}$) con dim V = n

Una base conocida de V, $B_e = \{e_1, ..., e_n\}$

Un endomorfismo de V, $h \in \operatorname{End}_K V$

La matriz A coordenada de h en la base B_e y la matriz B en otra base $B_v = \{v_1, ..., v_n\}$

6.2 Valores y vectores propios.

Definición 6.2.1 *Llamaremos valor propio (raíz característica o autovalor) y vector propio (vector caractarístico o autovector) asociado de un endomorfismo h al escalar $t \in K$ y vector $0 \neq v \in V$, $v \neq 0$ tales que*

$$h(v) = t\,v$$

Ejemplo 6.2.2 *En espacio de todas las funciones reales V y el endomorfismo*

$$h : V \longrightarrow V \textit{ definido por } h(f(x)) = f'(x),$$

las funciones $f(x) = e^{\alpha x}$ son funciones propias asociadas al valor propio $t = \alpha$.

La función f(x) = sen x no es una función propia.

Definición 6.2.3 *Al conjunto Sp(f) de todos los valores propios de f le llamaremos espectro de f. Al valor absoluto del valor propio de módulo mayor le llamaremos radio espectral de f.*

Propiedad 6.2.4 *El conjunto N(t,1) (ó S(t))de todos los vectores propios asociados a un valor propio t es un subespacio vectorial.*

Dem.: Basta tener en cuenta que si $v \in N(t,1)$, se verifica que:

$$f(v) = t\,v, \text{es decir}, (f - t\,id_V)(v) = 0,$$

o sea que: N(t, 1) = Ker $(f - t\,id_V)$. ■

Definición 6.2.5 *Al subespacio N(t,1) le llamaremos subespacio propio (subespacio fundamental o característico o núcleo generalizado) asociado al valor propio $t \in K$*

Para encontrar los valores propios y subespacios fundamentales asociados es preciso considerar coordenadas. Tomaremos una base $B_e = \{e_1, ..., e_n\}$ y la matriz A coordenada de h en dicha base. La definición (6.2.6) se escribe ahora:

$$(A - t\,I)\,X = 0 \tag{6.1}$$

Para que este sistema homogeneo tenga solución no nula debe ser:

$$\det(A - t\,I) = 0 \tag{6.2}$$

Definición 6.2.6 *La ecuación (6.2) se denomina ecuación característica de la matriz A*

Propiedad 6.2.7 *Todas las matrices coordenadas de un endomorfismo h sobre un espacio vectorial V tienen la misma ecuación característica.*

Dem.: Basta recordar que si A y B son matrices coordenadas del endomorfismo h son semejantes, es decir, existe P regular tal que $A = P^{-1}\,B\,P$. Sustituyendo en (6.2) resulta:

$$\det(A - t\,I) = 0 = \det(P^{-1}\,B\,P - t\,P^{-1}\,I\,P) = \det P\,\det P^{-1}\,\det(B - t\,I) = \det(B - t\,I)$$

. ■

Definición 6.2.8 *Llamaremos ecuación característica de $h \in End(V)$ a la ecuación característica de una cualquiera de sus matrices coordenadas.*

Definición 6.2.9 *El* $\det(A - t\,I)$ *es un polinomio en t de grado n denominado polinomio característico de h.*

Como la ecuación característica de A y de B es la misma, todos los coeficientes de ambas ecuaciones coinciden. En este sentido suele decirse que la ecuación característica de un endomorfismo, así como todos los coeficientes de la misma son invariantes, aunque existen otros.

Para encontrar los subespacios fundamentales de h, en primer lugar, hay que encontrar las raíces de la ecuación característica, es decir, los valores propios; a continuación hay que resolver el sistema (6.1) cada valor propio encontrado.

Ejercicio 6.2.10 *Sea el endomorfismo $h \in End(\mathbb{R}^3)$ definido por*

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

encontrar valores propios y subespacios fundamentales.

Solución. Su ecuación característica es

$$\det(A - t\,I) = 0 = -t^3 + t^2 + 5t + 3 = 0 = ((-1 - t)^2\,(3 - t)$$

que tiene como raices (valores propios) $t_1 = -1$, $m_1 = 2$ (doble) y $t_2 = 3$, $m_2 = 1$ (simple). Los subespacios fundamentales serán:

$$N(-1, 1) = ker(A - (-1)\,I) = < (-1, 0, 1), (0, 1, 0) >$$

Análogamente,

$$N(3, 1) = ker(A - 3\,I) = < (1, 0, 1) > \quad \blacksquare$$

Como se observa,
dim N(-1, 1) = n - rg(h+1) = n - rg (A + I) = 3 - 1 = 2
dim N(3, 1) = n - rg(h-3) = n - rg (A - 3 I) = 3 - 2 = 1
Notemos que no siempre coinciden la multiplicidad algebraica de un valor propio y la dimensión del subespacio fundamental asociado.

Definición 6.2.11 *A n_j = $dim\ S(t_j)$ se le denomina multiplicidad geométrica del valor propio t_j.*

6.3 Método de la potencia.

En algunas ocasiones es suficiente conocer una cota de los valores propios; por ejemplo, para conocer la convergencia de un método iterativo. Un resultado interesante es el siguiente

Teorema 6.3.1 *(de Gershgorin)*
a) Los valores propios de la matriz A están en la unión de todos los círculos de centro el elemento a_{jj} y radio la suma de los valores absolutos de los restantes elementos de cada fila, es decir,

$$C_j = \{z \in \mathbb{C} \,|\, |z - a_{jj}| \leq \sum |a_{jk}|\}, \quad t \in \cup C_j$$

b) Si la unión de p círculos tiene intersección vacía con la unión de los restantes, en dicha unión hay p valores propios

Para encontrar los valores propios de una matriz o endomorfismo existen varios métodos. Unos hallan las raíces del polinomio característico, otros trabajan directamente con la matriz. Estos últimos son más eficientes. Entre ellos podemos considerar el método de la potencia y sus variantes y el método QR y sus variantes.

Algoritmo para el método de la potencia que determina el valor propio dominante (i. e., $|\lambda_1| > |\lambda_j|, \forall j$) y un vector propio asociado. Este es un método iterativo convergente siempre que se pueda encontrar una base de V compuesta por vectores propios (teor. 6.4.2).

Proceso:
a) Considerar un vector inicial: $x = (x^1, ..., x^n)$
b) Encontrar la primera coordenada mayor en módulo, x^p, y su posición, p
c) Normalizar el vector x para obtener: $x_0 = x/x^p$
d) Calcular $y = A\,x_0$
e) Definir $\lambda_1 = y^p$
g) Redefinir $x = y$ y repetir desde b)
Detener el proceso cuando la diferencia entre dos λ's consecutivos sea suficientemente pequeña.

El valor λ_1 y el vector x son aproximaciones al valor propio dominante y a un vector propio asociado.

Ejemplo 6.3.2 *Encontrar el valor propio dominante y un valor propio asociado de la matriz*

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Solución: Método de la potencia.
k = 0
a) Considerar el vector inicial: $x = (1, 2)$
b) La primera coordenada mayor en módulo, $x^2 = 2$, y su posición, $p = 2$
c) Normalizar el vector x para obtener: $x_0 = (1/2, 1)$
d) Calcular $y = A\,x_0 = (2, 5/2)^T$
e) Definir $\lambda_1 = y^2 = 5/2$
g) Redefinir $x = y = (2, 5/2)$
k = 1
b) La primera coordenada mayor en módulo, $x^2 = 5/2$, y su posición, $p = 2$
c) Normalizar el vector x para obtener: $x_0 = (4/5, 1)$
d) Calcular $y = A\,x_0 = (13/5, 14/5)^T$
e) Definir $\lambda_1 = y^2 = 14/5$
g) Redefinir $x = y = (13/5, 14/5)$
k = 2
b) La primera coordenada mayor en módulo, $x^2 = 14/5$, y su posición, $p = 2$
c) Normalizar el vector x para obtener: $x_0 = (13/14, 1)$
d) Calcular $y = A\,x_0 = \cdots$
Finalmente, $\lambda_1 \simeq 14/5$ y $v_1 \simeq x = (13/5, 14/5) \longrightarrow (3, 3) = v_1$

6.4 Matrices diagonalizables.

El problema que estamos comenzando a estudiar es el de analizar h a través de sus subespacios fundamentales. Supondremos en adelante que V es un espacio vectorial sobre el cuerpo K, con dim V = n, $h \in End(V)$ y $\{t_1, ..., t_r\}$ los valores propios distintos de h (en general, $r \leq n$ = grado del polinomio característico de h).

Teorema 6.4.1 *La suma $S(t_1) + ... + S(t_r)$ de los subespacios fundamentales es suma directa.*

Dem.: Sean $a_i \in S(t_i)$, para cada $1 \leq i \leq$r; supongamos que

$$a_1 + ... + a_r = 0, (*)$$

debemos probar que entonces $a_i = 0$, para todo i. Veamos la demostración por inducción sobre r. Si r = 1 es trivial. Sea $r > 1$ y supongamos que el resultado es cierto para r -1 sumandos. Aplicando h a la ecuación (*), resulta:

$$t_1\, a_1 + ... + t_r\, a_r = 0$$

y multiplicandola por t_r y restandola de la anterior:

$$(t_1 - t_r)\, a_1 + ... + (t_{r-1} - t_r)\, a_{r-1} = 0,$$

como cada sumando $(t_i - t_r)\, a_i \in S(t_i)$ y hay r-1, por hipótesis de inducción, $t_i - t_r)\, a_i = 0$, para i = 1, 2, ..., r-1; pero, como $t_i \neq t_r$, resulta $a_i = 0, i = 1, 2, ..., r - 1$ y de la expresión (*) se deduce que $a_r = 0$. ■

Teorema 6.4.2 *$S(t_1) \oplus ... \oplus S(t_r) = V \Longleftrightarrow h$ admite una matriz coordenada diagonal.*

Dem.: [⇒] Sea $n_i = dimS(t_i)$. Elijamos bases de los subespacios fundamentales

$$\{a_{11}, ..., a_{1r_1}\}, ..., \{a_{i1}, ..., a_{ir_i}\}, ..., \{a_{r1}, ..., a_{rr_r}\}$$

Como la suma $S(t_1) \oplus ... \oplus S(t_r)$ es directa, la familia unión de las bases anteriores es base de esta suma que coincide con V. La matriz A asociada a h en esta base será la que tiene por columnas las imágenes de $h(a_{ij})$ en dicha base (si Y = A X), es decir,

$$h(a_{ij}) = t_i\, a_{ij} = (0, ..., 0, t_i, 0, ..., 0)$$

(observar que el elemento no nulo está en el lugar $n_1 + ... + n_{i-1+j}$), o sea que

$$A = diag(t_1, ..., t_1, t_2, ..., t_2, ..., t_r, ..., t_r)$$

[⇐] Si A es diagonal, los elementos de dicha diagonal serán los valores propios, por lo tanto, será suficiente agruparlos para obtener el resultado. ■

Definición 6.4.3 *Llamaremos endomorfismo diagonalizable a todo endomorfismo que admita una matriz coordenada diagonal.*

Notemos que los elementos de la matriz D diagonal semejante a A es la asociada al mismo endomorfismo en una base de vectores propios, que tiene como elementos los valores propios contados tantas veces como su multiplicidad.

Ejemplo 6.4.4 *La matriz diagonal del endomorfismo del ejemplo 6.2.10 respecto de la base $\{(-1, 0, 1), (0, 1, 0), (1, 0, 1)\}$ es*

$$D = diag(-1, -1, 3)$$

A continuación daremos un criterio práctico para decidir cuando un endomorfismo es diagonalizable.

Notemos que el polinomio característico se puede factorizar en K en la forma

$$\det(A - t\,I) = (t - t_1)^{m_1} \cdots (t - t_r)^{m_r}\, q(x)$$

donde q(x) es un polinomio que no tiene raices en K y m_i es la multiplicidad algebraica del valor propio t_i. Por lo tanto, se verifica:

$$n = m_1 + \cdots + m_r + dg(q(x)) \geq \sum_{i=1}^{r} m_i$$

En el caso K $= \mathbb{C}$, por ser algebraicamente cerrado, q es un polinomio constante, dg(q) = 0 y, por lo tanto, $n = m_1 + \cdots + m_r$.

Propiedad 6.4.5 *dim* $S(t_i) \leq m_i$, *para* $i = 1, 2, ..., r$

Teorema 6.4.6 *h es diagonalizable* $\Longleftrightarrow$
1) $m_1 + \cdots + m_r = n$
2) dim $S(t_i) = m_i, \quad i = 1, 2, ..., r$

Dem.: [⇒]h diagonalizable y $n_i = dimS(t_i) \Longrightarrow$

$$\Longrightarrow n = n_1 + ... + n_r \leq m_1 + ... + m_r = n \Longrightarrow n = n_1 + ... + n_r = m_1 + ... + m_r$$

y como $n_i \leq m_i, i = 1, 2, ..., r \Longrightarrow n_i = m_i, i = 1, 2, ..., r.$

[⇐]

$$dim(S(t_1) \oplus ... \oplus S(t_r)) = dimS(t_1) + ... + dimS(t_r) = m_1 + ... + m_r = n = dimV \Longrightarrow$$

$$\Longrightarrow S(t_1) \oplus ... \oplus S(t_r) = V \Longrightarrow h \text{ es diagonalizable. } \blacksquare$$

Ejemplo 6.4.7 *El endomorfismo* $h \in End(\mathbb{R}^3)$ *dado por la ecuación matricial*

$$Y = A\,X, \quad A = \begin{bmatrix} 2 & 6 & 6 \\ 0 & 2 & 0 \\ 0 & -3 & -1 \end{bmatrix}$$

tiene como valores propios $t_1 = 2, m_1 = 2$ *y* $t_2 = -1, m_2 = 1$*; el subespacio fundamental S(2) tiene dim S(2) = 3-rg(A-2I) = 2; luego, h es diagonalizable y la matriz asociada en cierta base será: B = diag(2,2,-1).*

No todos los endomorfismos son diagonalizables; por ejemplo, sea h con matriz asociada en cierta base

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

h no es diagonlizable, ya que sus valores propios son: $t_1 = 1, m_1 = 3$ y dim S(1) = 3 - rg(A-I) = 3 - 1 = 2 $\leq m_1$.

En muchas ocasiones es interesante encontrar una matriz sencilla que manifieste claramente la "estructura" del espacio vectorial en relación con un endomorfismo, aunque no sea diagonalizable. A esa matriz más sencilla le llamaremos *matriz o forma canónica de Jordan* del endomorfismo.

6.5 Aplicaciones.

6.5.1 Potencia n-ésima de una matriz

Dado que una matriz cuadrada A y su forma canónica diagonal D verifican: $A = P^{-1}\, D\, P$, la potencia n-ésima de A se puede obtener en la forma:

$$A^n = P\,D\,P^{-1}\,P\,D\,P^{-1}\ldots P\,D\,P^{-1} = P\,D^n\,P^{-1}.$$

En consecuencia, basta encontrar la forma canónica diagonal de A y la matriz de cambio a la base de vectores propios.

Ejercicio 6.5.1 *Calcular la expresión de A^n siendo A la matriz siguiente*

$$A = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$$

Ejercicio 6.5.2 *Calcular la expresión de B^n siendo B la matriz siguiente*

$$B = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

6.5.2 Raíz de una matriz

Dada una matriz cuadrada A y su forma canónica diagonal D verifican: $A = P\,D\,P^{-1}$, en el suuesto de que todos los valores propios de A sean no negativos, una raíz cuadrada de A se puede obtener en la forma:

$$\sqrt{A} = P\,\sqrt{D}\,P^{-1}, \text{ siendo } \sqrt{D} = diag(\sqrt{d_{11}}, ..., \sqrt{d_{nn}}).$$

En consecuencia, otra vez, basta encontrar la forma canónica diagonal de A y la matriz de cambio a la base de vectores propios.

6.5.3 Resolución de sistemas de ecuaciones diferenciales lineales

Un sistema de ecuaciones diferenciales lineales representa un problema en el que la incógnita es una función vectorial definida como una relación lineal entre ella y su derivada, más concretamente, se puede escribir en términos matriciales como:

$$X' = A\,X$$

Se puede probar que, si se conoce el valor de la solución $X(t_0)$ en un instante t_0 (condiciones iniciales), existe una única solución.

Para encontrar la solución correspondiente a las condiciones iniciales $X(t_0)$ se procede en la forma siguiente:

a) Se encuentra la forma canónica de Jordan y la matriz de paso

b) Se transforma el sistema inicial al siguiente: $Y' = J\,Y$, siendo

$$J = P^{-1}\,A\,P \quad \text{y} \quad X = P\,Y$$

c) Se resuelve el sistema diferencial lineal en forma regresiva

d) Se imponen las condiciones iniciales: $Y_0 = P^{-1}\,X_0$ para determinar las constantes de integración

e) Se realiza el cambio de coordenadas inverso para obtener: $X = P\,Y(t)$.

Ejercicio 6.5.3 *Resolver el sistema de ecuaciones diferenciales lineales, cuya matriz y condiciones iniciales son las siguientes*

$$A = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}, \quad X_0 = X(0) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

Ejercicio 6.5.4 *Resolver el sistema de ecuaciones diferenciales lineales, cuya matriz y condiciones iniciales son las siguientes*

$$B = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 2 \end{pmatrix}, \quad X_0 = X(0) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

6.6 Ejercicios para resolver

Ejercicios de [1] (Merino) números 61, 62, 65, 67, 142, 143 y 146

Ejercicios de [3] (de la Villa) capítulo 6, números 10, 17 y 20

Ejercicios de las hojas de enunciados de [2] (Palacios).

Ejercicios de [4] (Arvesu) números 15.21,15.22, 15.23, 15.24