5. DIAGONALIZACIÓN DE MATRICES

5. DIAGONALIZACIÓN DE MATRICES

5.1. INTRODUCCIÓN

En este tema se plantea el siguiente problema: sea un espacio vectorial E de dimensión n y un endomorfismo $f : E \to E$, ¿existirá alguna base de E en la cual la matriz asociada a f sea diagonal ?.

Si es así, entonces hay que buscar la base, así como la correspondiente matriz diagonal denotada por D. A este proceso se llama diagonalizar el endomorfismo o su matriz asociada.

De una manera general, la diagonalización de un endomorfismo f (o de su matriz asociada A) consiste esencialmente en obtener otra matriz D de manera que conserve las propiedades de A. Ambas matrices A y D son semejantes.

La matriz o forma diagonal D se obtiene a partir del estudio de todos los valores propios de A.

Existen matrices A que no se pueden reducir a una forma diagonal, se dice entonces, que A no es diagonalizable. Estas matrices si se pueden reducir a otros tipos, como son las matrices triangulares y las formas canónicas de Jordan.

5.2. VALORES Y VECTORES PROPIOS

Sea un espacio vectorial $\left[(E,+),(k,+,\times),\circ\right]$ de dimensión n, f un endomorfismo y una base U de E.

Se dice que un escalar $\lambda \in k$ es un valor propio, autovalor o valor característico de f, si existe al menos un vector $\vec{x} \in E$ y $\vec{x} \neq \vec{0}$ tal que se cumpla

$$f(\vec{x}) = \lambda_j(\vec{x}) \quad j = 1, 2, 3, \ldots, n \quad (1)$$

Si A es la matriz asociada al endomorfismo f en una base conocida U, e $\vec{y}$ es la imagen de $\vec{x}$ por el endomorfismo f, entonces

$$\vec{y} = f(\vec{x}) = (f)_{U,U}(\vec{x}) = A(\vec{x}) = \lambda(\vec{x})$$
$$A(\vec{x}) = \lambda(\vec{x}) \quad (2)$$

A los vectores que verifican las igualdades (1) ó (2) se llaman vectores propios, autovectores o vectores característicos asociados al valor propio λ.

La expresión (2) se puede escribir

$$A(\vec{x})-\lambda(\vec{x})=\vec{0} \quad \left[A-\lambda(I)\right](\vec{x})=\vec{0} \quad (3)$$

La igualdad (3) representa el cálculo de los vectores $\vec{x}$ que pertenecen al núcleo de la aplicación lineal $(f-\lambda_j I)$ cuya matriz asociada es $\left[A-\lambda_j I\right]$, es decir, $\operatorname{Ker}\left[A-\lambda_j I\right]$ ó $N\left[A-\lambda_j I\right]$. El sistema de ecuaciones lineales (3) a resolver es homogéneo, en consecuencia, para que admita soluciones distintas de la trivial o impropia, se debe cumplir que el rango o característica de la matriz de coeficientes debe ser menor que el número de incógnitas. Por tanto, es necesario que el determinante de la matriz de coeficientes sea nulo $|A-\lambda I|=0 \quad (4)$

Al desarrollo del determinante (4)
$p(\lambda)=a_0+a_1\lambda+a_2\lambda^2+\cdots\cdots+a_{n-1}\lambda^{n-1}+a_n\lambda^n$ se le llama polinomio característico. Igualando a cero el polinomio característico se obtiene la ecuación característica.

$$|A-\lambda I|=0 \iff p(\lambda)=a_0+a_1\lambda+a_2\lambda^2+\cdots\cdots+a_{n-1}\lambda^{n-1}+a_n\lambda^n=0$$

Las raíces de la ecuación característica, repetidas o no, son los valores propios λ de la matriz.

Orden o grado de multiplicidad algebraica de un valor propio λ es el número de veces que λ es raíz de la ecuación característica.

Los valores propios sustituidos en (3) y resuelto el sistema de ecuaciones lineales homogéneo resultante permite obtener los vectores propios del endomorfismo f (o de la matriz A).

La ecuación característica $|A-\lambda I|=0$ tendrá "n" raíces reales o complejas, por tanto, "n" valores propios con lo que el número de estos coincide con el orden de la matriz.

El conjunto de todos los valores propios de A se denomina espectro de A y se designa por:

$$\sigma(A)=\left\{\lambda_1,\lambda_2,\lambda_3,\ldots\ldots,\lambda_p\right\}$$

<u>Invarianza del polinomio característico</u>

El nombre de polinomio característico del endomorfismo f se debe a que es independiente de la base U considerada. En otras palabras, el polinomio característico es un invariante frente a los cambios de base.

Se considera otra base $V=\left\{\vec{v}_1,\vec{v}_2,\ldots\ldots,\vec{v}_n\right\}$ de E,

$S=\left(\vec{v}_1,\vec{v}_2,\ldots\ldots,\vec{v}_n\right)_U$ es la matriz regular de cambio de base de V respecto a U y $\left(f\right)_{V,V}$ es la matriz asociada al endomorfismo f respecto a la base V.

El polinomio característico en la base V es el determinante de la matriz $\left(f\right)_{V,V}-\lambda I$.

$$\left|\left(f\right)_{V,V}-\lambda I\right|$$

Utilizando la fórmula general de cambio de la matriz de una aplicación lineal al cambiar las bases $\left(f\right)_{U'V'}=\left(\vec{v}'_1,\vec{v}'_2,\ldots\ldots,\vec{v}'_n\right)_V^{-1}\left(f\right)_{U,V}\left(\vec{u}'_1,\vec{u}'_2,\ldots\ldots,\vec{u}'_n\right)_U$ y teniendo en cuenta que se verifica: $\lambda I=\lambda\left(S\right)_U^{-1}\cdot I\cdot\left(S\right)_U$, se deduce que

$$\begin{aligned}\left(f\right)_{V,V}-\lambda I&=\left(S\right)_U^{-1}\left(f\right)_{U,U}\left(S\right)_U-\lambda I=\left(S\right)_U^{-1}\left(f\right)_{U,U}\left(S\right)_U-\lambda\left(S\right)_U^{-1}\cdot I\cdot\left(S\right)=\\&=\left(S\right)_U^{-1}\left[\left(f\right)_{U,U}-\lambda I\right]\left(S\right)_U\end{aligned}$$

Entonces

$$\left|\left(f\right)_{V,V}-\lambda I\right|=\left|\left(S\right)_U^{-1}\left[\left(f\right)_{U,U}-\lambda I\right]\left(S\right)_U\right|$$

Aplicando el teorema de Binet–Cauchy relativo al determinante correspondiente al producto de matrices cuadradas de orden n, que es igual al producto de los respectivos determinantes, resulta

$$\begin{aligned}\left|\left(f\right)_{V,V}-\lambda I\right|&=\left|\left(S\right)_U^{-1}\left[\left(f\right)_{U,U}-\lambda I\right]\left(S\right)_U\right|=\left|\left(S\right)_U^{-1}\right|\cdot\left|\left(f\right)_{U,U}-\lambda I\right|\cdot\left|\left(S\right)_U\right|=\\&=\frac{1}{\left|\left(S\right)_U\right|}\left|\left(f\right)_{U,U}-\lambda I\right|\cdot\left|\left(S\right)_U\right|=\left|\left(f\right)_{U,U}-\lambda I\right|\end{aligned}$$

En resumen, que se cumple $\left|\left(f\right)_{V,V}-\lambda I\right|=\left|\left(f\right)_{U,U}-\lambda I\right|$

Esta igualdad justifica totalmente que el polinomio característico es independiente de la base utilizada en su cálculo.

Subespacios propios. Sus dimensiones

Sea un espacio vectorial $\left[\left(E,+\right),\left(k,+,\times\right),\circ\right]$ de dimensión finita n y un endomorfismo $f:E\rightarrow E$, para cada valor propio λ de f, es decir, para cada raíz de su ecuación característica, existen los vectores propios asociados, se llama subespacio propio S_{λ_j} al conjunto formado por el vector nulo y los vectores propios correspondientes a un valor propio.

El subespacio propio S_{λ_j} es el núcleo de $f-\lambda I$, por tanto, el conjunto de los vectores $\vec{x}\in E\ /\ S_{\lambda_j}=N(A-\lambda_i)\ \Rightarrow\ (f-\lambda_j I)(\vec{x})=\vec{0}$, siendo I la matriz identidad del mismo orden que A. Ello demuestra que es un subespacio vectorial.

En cuanto a dimensiones:

$$\dim\left(S_{\lambda_i}\right)=\dim\left[N\left(A-\lambda_j I\right)\right]=n-\text{rango}\left(A-\lambda_j I\right)\quad \text{para } j=1,2,3,\ldots\ldots,p$$

Si $\sigma(A)=\{\lambda_1,\lambda_2,\lambda_3,\ldots\ldots,\lambda_p\}$ entonces se cumple $S_{\lambda_1}\oplus S_{\lambda_2}\oplus\cdots\cdots\oplus S_{\lambda_p}$, la suma de los subespacios propios $\mathbf{s}_{\boldsymbol{\lambda}_J}$ es directa.

Se denomina orden o grado de multiplicidad de un valor propio $\lambda_j\in\sigma(A)$ al valor α_j que indica el número de veces que λ_j es raíz de la ecuación característica $|A-\lambda I|=0$. Se verifica que $1\le\dim\left(S_{\lambda_j}\right)\le\alpha_j$.

5.3. MATRICES DIAGONALIZABLES

Un endomorfismo "f" o su matriz asociada $A_{(n)}$ definida sobre un cuerpo k se dice que es diagonalizable, si existe una base $\mathbf{U}=\{\vec{\mathbf{u}}_1,\vec{\mathbf{u}}_2,\vec{\mathbf{u}}_3,\ldots\ldots,\vec{\mathbf{u}}_n\}\in k^{\mathbf{n}}$ formada por vectores propios de $A_{(n)}$.

<u>Características de matrices diagonalizables</u>

Sea $A_{(n)}$ una matriz cuadrada, las condiciones de su diagonalización o características de las matrices que son diagonalizables son dos:

1) La suma de los órdenes o grados de multiplicidad de los valores propios es igual a la dimensión "n" del espacio vectorial $\sum_{i=1}^{p}\alpha_i=n$
2) La dimensión del subespacio propio asociado a un valor propio es igual al orden o grado de multiplicidad del correspondiente valor propio.

$$\dim\left(S_{\lambda_j}\right)=\alpha_j\ ,\ \forall j=1,2,3,\ldots\ldots,p$$

Una vez comprobadas las condiciones de diagonalización y en el supuesto que la matriz $A_{(n)}$ cumpla ambas, hay que formar la llamada "forma diagonal" y la "matriz de paso".

La forma diagonal $D_{(n)}$ o matriz diagonalizada de $A_{(n)}$ tiene los valores propios situados en la diagonal principal y repetidos tantas veces como indique su orden o grado de multiplicidad, el resto de los elementos de ella son nulos.

Al ser la matriz $A_{(n)}$ diagonalizable, existe una base $U=\{\vec{u}_1,\vec{u}_2,\vec{u}_3,\ldots\ldots,\vec{u}_n\}\subset\mathbf{R}^n$ formada por vectores propios de $A_{(n)}$. Si $P_{(U)}$ es la matriz cuyas columnas son los vectores de la base U y $D_{(n)}$ es la matriz diagonal formada por los vectores propios de $A_{(n)}$, repetidos tantas veces como indique su orden o grado de multiplicidad, en el mismo orden que guardan en la base U sus vectores propios asociados entonces se cumple.

La matriz $A_{(n)}$ es diagonalizable por semejanza $\text{Si } A_{(n)}=P_{(U)}\cdot D_{(n)}\cdot P_{(U)}^{-1}$

5.4. DIAGONALIZACIÓN DE MATRICES SIMÉTRICAS

Existe un caso particular de diagonalización, cuando $A_{(n)}$ es una matriz simétrica cuyos elementos son números reales. Entonces la matriz siempre es diagonalizable, no hay que comprobar las condiciones de diagonalización.

En cada subespacio S_{λ_j} se puede encontrar una base de "vectores propios ortonormales" denotada por $\mathbf{V_j}$, de modo que la unión de todas estas bases $V=V_1\cup V_2\cup V_3\ldots V_p$ es una "base ortonormal" $V=\{\vec{v}_1,\vec{v}_2,\vec{v}_3,\ldots\ldots,\vec{v}_n\}\subset\mathbf{R}^n$. La ortonormalización de la base V se consigue utilizando el "método de Gram-Schmidt".

Si $P_{(U)}$ es la matriz cuyas columnas son los vectores de la base U y $D_{(n)}$ la matriz diagonal formada por los valores propios de $A_{(n)}$, repetidos tantas veces como indique su orden o grado de multiplicidad y en el mismo orden que guardan sus vectores propios asociados en la base U se cumple.

$A_{(n)}$ es diagonalizable por semejanza ortogonal $\text{Si } A_{(n)}=P_{(V)}\cdot D_{(n)}\cdot P_{(V)}^{T}$

<u>Potencia de una matriz diagonalizable</u>

Sea $A_{(n)}$ una matriz cuadrada real y diagonalizable. Una base $U=\{\vec{u}_1,\vec{u}_2,\vec{u}_3,\ldots\ldots,\vec{u}_n\}\subset\mathbf{R}^n$ formada por sus vectores propios.

Asimismo, $P_{(U)}$ y $D_{(n)}$ las matrices formadas por los vectores propios de la base U y los valores propios correspondientes de $A_{(n)}$ respectivamente.

Se calcula la potencia de orden cualquiera "m" de $A_{(n)}$ de la siguiente forma:

$$A_{(n)}^{m} = \left(P_{(U)} \cdot D_{(n)} \cdot P_{(U)}^{-1}\right)^{m} = \left(P_{(U)} \cdot D_{(n)} \cdot P_{(U)}^{-1}\right) \cdot \left(P_{(U)} \cdot D_{(n)} \cdot P_{(U)}^{-1}\right) \cdot \underbrace{\cdots\cdots}_{m} \cdot \left(P_{(U)} \cdot D_{(n)} \cdot P_{(U)}^{-1}\right) =$$

$$= P_{(U)} \cdot D_{(n)}^{m} \cdot P_{(U)}^{-1}.$$

Si la matriz $A_{(n)}$ es "simétrica", $V = \{\vec{v}_1, \vec{v}_2, \vec{v}_3, \ldots\ldots, \vec{v}_n\} \subset \mathbf{R}^n$ una base de vectores propios ortonormales, $P_{(V)}$ y $D_{(n)}$ las matrices formadas por los vectores propios de la base V y los valores propios correspondientes de $A_{(n)}$ respectivamente.

Se calcula la potencia de orden cualquiera "m" de $A_{(n)}$ de la siguiente forma.

$$A_{(n)}^{m} = \left(P_{(V)} \cdot D_{(n)} \cdot P_{(V)}^{T}\right)^{m} = \left(P_{(V)} \cdot D_{(n)} \cdot P_{(V)}^{T}\right) \cdot \left(P_{(V)} \cdot D_{(n)} \cdot P_{(V)}^{T}\right) \cdot \underbrace{\cdots\cdots}_{m} \cdot \left(P_{(V)} \cdot D_{(n)} \cdot P_{(V)}^{T}\right) =$$

$$= P_{(V)} \cdot D_{(n)}^{m} \cdot P_{(V)}^{T}.$$

5.5. POLINOMIO MÍNIMO DE UNA MATRIZ

Se denomina polinomio de la matriz cuadrada A a toda expresión de la forma

$$P(A) = a_0 I + a_1 A + a_2 A^2 + \ldots\ldots + a_{n-1} A^{n-1} + a_n A^n$$

En este polinomio se observa una combinación lineal de una serie de escalares $a_i \in k$ y de operaciones como son el 'producto de un escalar por una matriz, suma de matrices y producto de matrices. Tanto A como $P(A)$ son matrices cuadradas del mismo orden.

Dado que $P(A)$ está formado por las matrices A e I y que $A \times I = I \times A$, los polinomios matriciales se descomponen en producto de factores (o factorización) de igual manera que sucede con los polinomios en la variable x, se puede aplicar la regla de Ruffini.

El problema que se trata de resolver es: dada una matriz A, calcular ecuaciones de la forma $P(A) = 0$.

Se llama polinomio mónico a aquel que tiene el coeficiente de la mayor potencia igual a la unidad.

Se llama polinomio mínimo $P_m(A)$ de una matriz cuadrada A, al polinomio mónico correspondiente a la ecuación matricial de grado mínimo que dicha matriz verifica.

A la ecuación matricial $m(A)=0$ se la llama ecuación mínima de A.
Si A es una matriz cuadrada de orden n, el polinomio mínimo será de la forma

$$P_m(x)=\alpha_p+\alpha_{p-1}x+\ldots\ldots+\alpha_1 x^{p-1}+\alpha_0 x^p \quad \text{con} \quad p\leq n$$

Cálculo del polinomio mínimo

Sea A una matriz cuadrada de orden n con elementos en el cuerpo $(k,+,\times)$ conmutativo.
El procedimiento para obtener el polinomio mínimo $P_m(x)$ es el siguiente:

1. Comprobar que se cumple: $A=k_0 I_n$ ó $A-k_0 I_n=0$, entonces el polinomio mínimo es $P_m(\lambda)=\lambda-k_0$.
2. Si no se verifica el punto 1, entonces comprobar si se cumple que: $A^2=k_1 A+k_0 I_n \Rightarrow P_m(\lambda)=\lambda^2-k_1\lambda-k_0$
3. Si no se cumple el punto 2, se comprueba si $A^3=k_2 A^2+k_1 A+k_0 I_n$ .Si se verifica entonces $P_m(\lambda)=\lambda^3-k_2\lambda^2-k_1\lambda-k_0$

Se continúa aplicando el algoritmo hasta que se cumpla la primera igualdad. Una vez que se halla el polinomio mínimo, se iguala a cero y se calculan las raíces de la ecuación resultante para obtener la expresión del polinomio mínimo factorizado.

5.6. FORMA REDUCIDA DE JORDAN DE UNA MATRIZ

Si en la matriz A asociada a un endomorfismo f no existen "n" vectores propios del espacio vectorial E con los que formar una base, en ocasiones es conveniente tratar de buscar una matriz que sea lo más sencilla posible para caracterizar el endomorfismo y facilitar cierto tipo de cálculos.

El método que más se utiliza es el de Jordan, también conocido como forma canónica o reducida de Jordan. Se trata de hallar una matriz lo más sencilla posible representada por J a la que se llama matriz de Jordan, también se obtendrá la matriz P en donde sus columnas son las coordenadas de los vectores de la nueva base, en la que J caracteriza al endomorfismo f.

Descripción del método

Sea un endomorfismo f cuya matriz asociada A de orden "n" tiene "r" valores propios que originan $"k"(k<n)$ vectores propios asociados, es decir, no se puede formar una base de "n" vectores propios luego no es posible la diagonalización de A.

La matriz A siempre se puede expresar en forma canónica o reducida de Jordan, si sucede que los polinomios característico y mínimo se puedan escribir como productos de factores que sean polinomios de primer grado o lineales. Esto se puede realizar cuando el cuerpo k sobre el que se define el espacio vectorial sea el de los números complejos. Si k es el cuerpo de los números reales no siempre es posible hallar J.

Sea el polinomio característico de la matriz A de orden n: $P_c(\lambda) = |A - \lambda I|$ que expresado en forma factorial es.

$$P_c(\lambda) = (\lambda - \lambda_1)^{\alpha_1} (\lambda - \lambda_2)^{\alpha_2} (\lambda - \lambda_3)^{\alpha_3} \ldots\ldots (\lambda - \lambda_r)^{\alpha_r}$$

Sea el polinomio mínimo de la matriz A de orden n. Expresado en forma factorial es.

$$P_m(\lambda) = (\lambda - \lambda_1)^{\beta_1} (\lambda - \lambda_2)^{\beta_2} (\lambda - \lambda_3)^{\beta_3} \ldots\ldots (\lambda - \lambda_r)^{\beta_r}$$

La matriz correspondiente a la forma canónica o reducida de Jordan es

$$J = \begin{pmatrix} (J_1) & (0) & (0) & \cdots & (0) \\ (0) & (J_2) & (0) & \cdots & (0) \\ (0) & (0) & (J_3) & \cdots & (0) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ (0) & (0) & (0) & \cdots & (J_k) \end{pmatrix}$$

Los elementos de esta matriz J superdiagonal son submatrices expresadas por J_i (situadas en la diagonal principal) y nulas (0) situadas en el resto de las posiciones, estas últimas no tienen que ser necesariamente cuadradas.

Las submatrices J_i se llaman bloques de Jordan y se representan de la forma

$$J_i = \begin{pmatrix} \lambda_i & 0 & 0 & \cdots & 0 \\ 0 & \lambda_i & 0 & \cdots & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda_i \end{pmatrix}$$

Los bloques J_i tienen las siguientes características:

a) Existe al menos un bloque J_i de orden "β_i. Los restantes bloques son de orden $\leq \beta_i$.

b) La suma de los órdenes de los bloques J_i es α_i.
c) el número de bloques J_i es igual a la multiplicidad de λ_i.
d) El número de bloques J_i de cada orden posible está determinado sólo por la matriz A.