Álgebra Lineal

Tema 11. El Teorema Espectral en $\mathbb{R}$

Grado en Ingeniería Informática

Doble Grado en Ingeniería Informática y Administración de Empresas

AUTORES: J. SALAS, A. TORRENTE Y E.J.S. VILLASEÑOR

uc3m | Universidad Carlos III de Madrid

Índice general

Tema 11

El teorema espectral en $\mathbb{R}$

En este tema concluimos el estudio de los valores y vectores propios de una matriz cuadrada analizando un tipo particular de matrices: las simétricas de coeficientes reales. Aunque de apariencia simple, éstas aparecen en multitud de campos como la física, la estadística, la geometría o la ingeniería, lo que las dota de especial interés. También estudiaremos las características del proceso de diagonalización de dichas matrices, es decir, el cálculo de las matrices P y D tales que $A = P\,D\,P^{-1}$.

11.1. Diagonalización de matrices simétricas reales

Ejemplo

Consideremos la siguiente matriz cuadrada y simétrica:

$$A = \begin{pmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{pmatrix}.$$

Veamos cuáles son los valores propios y las correspondientes multiplicidades, para estudiar si es diagonalizable. La ecuación característica es

$$-(\lambda - 8)(\lambda - 6)(\lambda - 3) = 0$$

y sus raíces son $\lambda_1 = 8$, $\lambda_2 = 6$ y $\lambda_3 = 3$. Obviamente, las multiplicidades algebraicas y geométricas de todas ellas coinciden y por tanto A es diagonalizable.
Los espacios propios asociados tienen bases dadas por

$$B_{\lambda_1} = \left((-1,1,0)^t\right) \ ; \quad B_{\lambda_2} = \left((-1,-1,2)^t\right) \ ; \quad B_{\lambda_3} = \left((1,1,1)^t\right) \ .$$

Estos tres vectores forman una base de $\mathbb{R}^3$; además, es fácil comprobar que también forman un conjunto ortogonal. Si en lugar de usar estos vectores para obtener la matriz P en el proceso de diagonalización, usamos los vectores unitarios correspondientes, P será una matriz con columnas ortonormales, por lo que será una matriz ortogonal, cuya inversa se calcula simplemente como $P^{-1} = P^t$. Por tanto, el proceso de diagonalización da

$$A = P\,D\,P^t = \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} 8 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}.$$

Las observaciones que hemos hecho en el ejemplo anterior son generales. Tenemos los siguientes importantes resultados:

Teorema espectral

Una matriz A de dimensión $n \times n$, de entradas $\boxed{\text{reales}}$ y $\boxed{\text{simétrica}}$ verifica las siguientes propiedades:

1. A tiene n $\boxed{\text{valores propios reales}}$ (incluyendo multiplicidades).

2. Si λ es un valor propio de A con multiplicidad k, entonces el espacio propio asociado a λ es k-dimensional.

3. Los espacios propios son mutualmente ortogonales, es decir: dos vectores propios que corresponden a dos valores propios distintos son ortogonales.

4. Existen una matriz ortogonal P y una matriz diagonal D tales que $A = P\,D\,P^t$.

El teorema espectral nos dice que en el caso de las matrices simétricas (de coeficientes reales) los espacios propios forman una "descomposición" ortogonal de $\mathbb{R}^n$ y es posible encontrar una base ortonormal de $\mathbb{R}^n$ formada por vectores propios de A.

El teorema anterior justifica la siguiente definición:

Una matriz $A \in \mathbb{R}^{n\times n}$ se dice que es **ortogonalmente diagonalizable** si y sólo si es diagonalizable mediante una matriz de diagonalización P ortogonal, es decir, si existen P ortogonal y D diagonal tales que $A = P\,D\,P^t$.

Ejemplo

Consideremos la siguiente matriz cuadrada y simétrica:

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

Por ser simétrica y real será ortogonalmente diagonalizable. Vamos a encontrar sus valores propios (que sabemos que han de ser reales) y los correspondientes espacios propios (que sabemos que serán ortogonales), junto con una terna de vectores propios ortogonales.

La ecuación característica es

$$-(\lambda-1)^2(\lambda+2)=0,$$

con raíces: $\lambda_1 = \lambda_2 = 1$ y $\lambda_3 = -2$. Las multiplicidades algebraicas y geométricas de todos ellos han de coincidir (por ser A simétrica). Si determinamos los espacios propios asociados, se comprueba que éstos tienen bases dadas por

$$B_{\lambda_1} = (\boldsymbol{v}_1, \boldsymbol{v}_2) = \left((-1,1,0)^t, (-1,0,1)^t\right); \quad B_{\lambda_3} = (\boldsymbol{v}_3) = \left((1,1,1)^t\right).$$

Es claro que $\boldsymbol{v}_1$ y $\boldsymbol{v}_2$ son ortogonales a $\boldsymbol{v}_3$, pero en cambio $\boldsymbol{v}_1$ y $\boldsymbol{v}_2$ no son ortogonales. Esto significa que, si buscamos una matriz P que sea ortogonal para diagonalizar A, no podremos hacer uso de estos vectores. En su lugar habrá que buscar dos vectores ortonormales que generen el mismo espacio propio. Claramente, conseguiremos este propósito usando el método de Gram-Schmidt y dividiendo por la norma los vectores obtenidos. Así, las columnas de P serán:

$$\boldsymbol{w}_1 = \frac{1}{\sqrt{2}}\,(-1,1,0)^t\,, \boldsymbol{w}_2 = \frac{1}{\sqrt{6}}\,(-1,-1,2)^t\,, \boldsymbol{w}_3 = \frac{1}{\sqrt{3}}\,(1,1,1)^t$$

y tendremos

$$\begin{aligned} A &= P\,D\,P^t \\ &= \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}. \end{aligned}$$

Además, tenemos el siguiente resultado fundamental:

Teorema

Una matriz $A \in \mathbb{R}^{n\times n}$ es ortogonalmente diagonalizable si y sólo si A es simétrica.

11.2. Teorema de descomposición espectral

Además del teorema espectral y del teorema de caracterización de las matrices ortogonalmente diagonalizables, las matrices reales simétricas de dimensión $n \times n$ satisfacen lo que se conoce como teorema de descomposición espectral, que permite escribir (*descomponer*) éstas como una suma de n matrices derivadas a partir de sus valores y vectores propios. Antes de abordarlo, introducimos en esta sección algunas definiciones y resultados.

Dado un espacio vectorial V sobre $\mathbb{R}$ dotado de producto interno, diremos que una transformación lineal $T : V \to V$ es una **proyección** (o un *proyector*) si

$$T \circ T = T\,.$$

Es frecuente utilizar la notación Π para referirse a proyectores. Recordemos que:

Una matriz cuadrada A se dice **idempotente** si $A^2 = A$.

Ambas definiciones se relacionan de la siguiente manera:

Proposición

Si la proyección $\Pi\colon V \to V$ se representa por la matriz A_Π relativa a una base ortonormal de V, entonces A_Π es idempotente.

Además introducimos la siguiente definición:

Sea un espacio vectorial V sobre $\mathbb{R}$ dotado de producto interno. Dada la transformación lineal $T\colon V \to V$, definimos la **transformación adjunta de** T como la transformación lineal $T^*\colon V \to V$ tal que para todo $v_1, v_2 \in V$:

$$\langle v_1, T^*(v_2)\rangle = \langle T(v_1), v_2\rangle .$$

Es importante observar que para toda transformación lineal T la transformación adjunta existe y es única.

Proposición

Si la transformación lineal $T\colon V \to V$ se representa por la matriz A_T relativa a una base ortonormal de V, entonces T^* se representa mediante A_T^t.

Ambos conceptos se emplean para definir un nuevo tipo de transformación lineal (desde el punto de visto de la geometría, como veremos en el Tema 12).

Dado un espacio vectorial V sobre $\mathbb{R}$ dotado de producto interno, diremos que una transformación lineal $\Pi\colon V \to V$ es una **proyección ortogonal** (o un *proyector ortogonal*) si

1. $\Pi \circ \Pi = \Pi$ (es decir, Π es un proyector).
2. $\Pi^* = \Pi$ (es decir, Π es *autoadjunta*).

Y obviamente, tenemos el siguiente resultado:

Dado un espacio vectorial V sobre $\mathbb{R}$ dotado de producto interno, si la proyección ortogonal $\Pi\colon V \to V$ se representa por la matriz A_Π relativa a una base ortonormal de V, entonces:

1. $A_\Pi^2 = A_\Pi$ (es decir, A_Π es idempotente).

2. $A_\Pi^t = A_\Pi$ (es decir, A_Π es *simétrica*).

Ejemplo

Consideremos el espacio $\mathbb{P}_1$ sobre $\mathbb{R}$ dotado del producto interno

$$\langle a_0 + a_1 x, b_0 + b_1 x \rangle = a_0 b_0 + a_1 b_1 \,.$$

La prueba de que efectivamente es producto interno es inmediata. Sea la transformación lineal Π que asocia a cada polinomio $a_0 + a_1 x$ el polinomio $\Pi(a_0 + a_1 x) = 2a_0 + a_1 - (2a_0 + a_1)x$. Con respecto a la base ortonormal de $\mathbb{P}_1$ dada por $B = (1, x)$, la transformación Π se representa mediante la matriz

$$A_{\Pi,B} = \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix} .$$

Evidentemente, tenemos que $A_{\Pi,B}$ es idempotente:

$$A_\Pi^2 = \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix} = A_\Pi \,,$$

por lo que Π es una proyección. Sin embargo, A_Π no es simétrica, por lo que Π no es proyección ortogonal.

La transformación adjunta Π^* de Π la calcularíamos como sigue: consideremos la imagen por Π^* de los vectores de la base B (suficiente para describir Π^*). Se debe cumplir, por una parte, que:

$$\langle a_0 + a_1 x, \underbrace{\alpha + \beta x}_{\Pi^*(1)} \rangle = \langle \underbrace{2a_0 + a_1 - (2a_0 + a_1)x}_{\Pi(a_0 + a_1 x)}, 1 \rangle$$
$$\Rightarrow \qquad a_0\alpha + a_1\beta = 2a_0 + a_1 \,, \quad \Rightarrow \quad \alpha = 2 \,, \beta = 1$$

y por otra que:

$$\langle a_0 + a_1 x, \underbrace{\gamma + \delta x}_{\Pi^*(x)} \rangle = \langle \underbrace{2a_0 + a_1 - (2a_0 + a_1)x}_{\Pi(a_0+a_1x)}, x \rangle$$

$$\Rightarrow \qquad a_0\gamma + a_1\delta = -2a_0 - a_1\,, \quad \Rightarrow \quad \gamma = -2, \delta = -1\,.$$

Así, la imagen por Π^* de un polinomio arbitrario $a_0 + a_1 x$ vendrá dada por:

$$\begin{aligned} \Pi^*(a_0 + a_1 x) &= a_0\,\Pi^*(1) + a_1\,\Pi^*(x) = a_0(2 + x) + a_1(-2 - x) \\ &= 2a_0 - 2a_1 + (a_0 - a_1)x\,. \end{aligned}$$

Si representamos Π^* mediante una matriz respecto a la base B, se obtiene que

$$A_{\Pi^*,B} = \begin{pmatrix} 2 & -2 \\ 1 & -1 \end{pmatrix} = A^t_{\Pi,B}\,.$$

Proposición

Si Π es una proyección ortogonal entonces $ker(\Pi) = Im(\Pi)^{\perp}$.

Ejemplo

Consideremos en $\mathbb{R}^3$ con el producto escalar ordinario la transformación Π que asocia a cada vector $\boldsymbol{v} = (v_1, v_2, v_3)^t$ la imagen dada por

$$\Pi\left((v_1, v_2, v_3)^t\right) = \frac{1}{9}\left(v_1 + 2(v_2 + v_3), 2v_1 + 4(v_2 + v_3), 2v_1 + 4(v_2 + v_3)\right)^t\,.$$

Obviamente, respecto a la base canónica, la representación de Π viene dada por la matriz simétrica

$$A_{\Pi} = \frac{1}{9}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{pmatrix},$$

por lo que Π es autoadjunta. También es fácil ver que $A_\Pi^2 = A_\Pi$, por lo que Π será una proyección ortogonal. Es inmediato observar que $\mathrm{Im}(\Pi) = \mathrm{Gen}\left((1,2,2)^t\right)$.
El kernel de Π es

$$\ker(\Pi) = \{\boldsymbol{v} \in \mathbb{R}^3 \colon \Pi(\boldsymbol{v}) = \boldsymbol{0}\} = \mathrm{Gen}\left((-2,1,0)^t, (-2,0,1)^t\right),$$

que obviamente es el complemento ortogonal de $\mathrm{Im}(\Pi)$.

Teorema

Sean $\Pi_1, \ldots, \Pi_r$ proyecciones ortogonales sobre un espacio vectorial real V con producto interno que satisfacen:

$$\begin{aligned} \Pi_1 + \cdots + \Pi_r &= I \quad \text{(aplicación identidad)}, \\ \Pi_i \circ \Pi_j &= 0 \quad \text{(aplicación nula)}, \quad \text{para todo } i \neq j. \end{aligned}$$

Sea $U_i = \mathrm{Im}(\Pi_i)$. Entonces

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_r$$

y si $\mathbf{u}_i \in U_i$ y $\mathbf{u}_j \in U_j$ para todo $i \neq j$, entonces $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$.

Una descomposición de un espacio en forma de suma directa que satisface las condiciones del teorema anterior se denomina una **descomposición ortogonal** de V. Dicho teorema es el escenario general del teorema de descomposición espectral de las matrices reales simétricas, que enunciamos a continuación.

Teorema de descomposición espectral

Sea A una matriz real simétrica de dimensión $n \times n$ que podemos escribir en la forma $A = P\,D\,P^t$, donde las columnas de P son las coordenadas de los vectores propios *ortonormales* $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ de A y los correspondientes valores propios $\lambda_1, \ldots, \lambda_n$ están en la diagonal principal de la matriz D. Entonces

$$A = \sum_{i=1}^{n} \lambda_i A_i = \lambda_1 A_1 + \cdots + \lambda_n A_n ,$$

donde las matrices $A_i = \boldsymbol{v}_i \boldsymbol{v}_i^t$ $(1 \leqslant i \leqslant n)$ son simétricas, idempotentes y verifican que

$$\begin{aligned} \sum_{i=1}^{n} A_i &= \sum_{i=1}^{n} \boldsymbol{v}_i \boldsymbol{v}_i^t = I_n , \\ A_i A_j &= \left(\boldsymbol{v}_i \boldsymbol{v}_i^t\right) \left(\boldsymbol{v}_j \boldsymbol{v}_j^t\right) = 0_{n\times n} , \quad \forall\, i \neq j . \end{aligned}$$

Obsérvese que las matrices A_i representan proyectores ortogonales Π_i, que verifican las condiciones del teorema de descomposición ortogonal. En estas condiciones, la transformación lineal $T = \sum_{i=1}^{n} \lambda_i \Pi_i$ está representada por la matriz simétrica real A y obviamente tiene valores propios $\lambda_1, \ldots, \lambda_n$.

Ejemplo

Consideremos de nuevo la matriz

$$\begin{aligned} A &= \begin{pmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} 8 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{pmatrix}}_{D} \underbrace{\begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}}_{P^t} . \end{aligned}$$

Utilizando la descomposición espectral podemos escribir:

$$
\begin{aligned}
A &= 8\,\frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\frac{1}{\sqrt{2}}\,(-1,1,0) + 6\,\frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}\frac{1}{\sqrt{6}}\,(-1,1,2) \\
&\quad +3\,\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\frac{1}{\sqrt{3}}\,(1,1,1) \\
&= 8\underbrace{\left[\frac{1}{2}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right]}_{A_1} + 6\underbrace{\left[\frac{1}{6}\begin{pmatrix} 1 & 1 & -2 \\ 1 & 1 & -2 \\ -2 & -2 & 4 \end{pmatrix}\right]}_{A_2} + 3\underbrace{\left[\frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\right]}_{A_3}
\end{aligned}
$$

cuya suma, efectivamente, es A.

Es trivial comprobar que

$$
\begin{aligned}
A_1\,A_1 &= A_1\,, \quad A_2\,A_2 = A_2\,, \quad A_3\,A_3 = A_3 \quad \text{(idempotentes)}\,, \\
A_1^t &= A_1\,, \quad A_2^t = A_2\,, \quad A_3^t = A_3 \quad \text{(simétricas)}\,, \\
I_3 &= A_1 + A_2 + A_3\,, \\
0_{3\times 3} &= A_1\,A_2 = A_1\,A_3 = A_2\,A_3\,.
\end{aligned}
$$

Ejemplo

Consideremos la matriz

$$
A = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix},
$$

con valores propios $\lambda_1 = 5$, $\lambda_2 = 2$ y $\lambda_3 = -1$ y vectores propios $\boldsymbol{v}_1 = (2,2,1)^t$,

$\boldsymbol{v}_2 = (2, -1, -2)^t$ y $\boldsymbol{v}_3 = (1, -2, 2)^t$. El teorema de descomposición espectral nos dice que podemos escribir A como la suma de las tres matrices:

$$
\begin{aligned}
M_1 &= \frac{\lambda_1}{\|\boldsymbol{v}_1\|^2} \boldsymbol{v}_1 \boldsymbol{v}_1^t = \frac{1}{9} \begin{pmatrix} 20 & 20 & 10 \\ 20 & 20 & 10 \\ 10 & 10 & 5 \end{pmatrix}, \\
M_2 &= \frac{\lambda_2}{\|\boldsymbol{v}_2\|^2} \boldsymbol{v}_2 \boldsymbol{v}_2^t = \frac{1}{9} \begin{pmatrix} 8 & -4 & -8 \\ -4 & 2 & 4 \\ -8 & 4 & 8 \end{pmatrix}, \\
M_3 &= \frac{\lambda_3}{\|\boldsymbol{v}_3\|^2} \boldsymbol{v}_3 \boldsymbol{v}_3^t = \frac{1}{9} \begin{pmatrix} -1 & 2 & -2 \\ 2 & -4 & 4 \\ -2 & 4 & -4 \end{pmatrix}.
\end{aligned}
$$

Es fácil ver que las matrices

$$A_1 = \frac{1}{\|\boldsymbol{v}_1\|^2} \boldsymbol{v_1} \boldsymbol{v_1}^t = \frac{1}{5} M_1 , \quad A_2 = \frac{1}{\|\boldsymbol{v}_2\|^2} \boldsymbol{v_2} \boldsymbol{v_2}^t = \frac{1}{2} M_2 , \quad A_3 = \frac{1}{\|\boldsymbol{v}_3\|^2} \boldsymbol{v_3} \boldsymbol{v_3}^t = -M_3$$

son idempotentes (representan proyecciones Π_i) y simétricas (las proyecciones son ortogonales). También es inmediato comprobar que

$$A_1 + A_2 + A_3 = I_3 \quad \text{(matriz identidad)}$$

o equivalentemente que $\Pi_1 + \Pi_2 + \Pi_3 = I_3$ (aplicación identidad) y que

$$A_i \, A_j = 0_{3\times 3} \quad \text{(matriz nula) para todo } i \neq j$$

o equivalentemente que $\Pi_i \circ \Pi_j = 0$ (aplicación nula) para todo $i \neq j$. Finalmente, se observa que

$$\mathrm{Im}(\Pi_1) = \mathrm{Gen}(\boldsymbol{v}_1) , \quad \mathrm{Im}(\Pi_2) = \mathrm{Gen}(\boldsymbol{v}_2) , \quad \mathrm{Im}(\Pi_3) = \mathrm{Gen}(\boldsymbol{v}_3) ,$$

por lo que $\mathbb{R}^3$ puede descomponerse como suma directa en la forma:

$$\mathbb{R}^3 = \mathrm{Im}(\Pi_1) \oplus \mathrm{Im}(\Pi_2) \oplus \mathrm{Im}(\Pi_3)\,.$$

Esto indica que, como sabemos, se puede formar una base de $\mathbb{R}^3$ con vectores propios de la matriz simétrica A.