Descomposición Espectral

Alan Reyes-Figueroa

Métodos Numéricos II (Aula 03) 11.julio.2023

Descomposición Espectral

Definición

*Una **descomposición espectral** de una matriz cuadrada $A \in \mathbb{R}^{n\times n}$ es una factoración de la forma*

$$A = X\Lambda X^{-1},$$

donde $X \in \mathbb{R}^{n\times n}$ es no singular, y $\Lambda \in \mathbb{R}^{n\times n}$ es una matriz diagonal, cuyas entradas $\lambda_1, \ldots, \lambda_n$ corresponden a los autovalores de A.

La definición anterior puede escribirse como $AX = X\Lambda$, esto es

$$\begin{pmatrix} & & \\ & A & \\ & & \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_n \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1\mathbf{x}_1 & \lambda_2\mathbf{x}_2 & \ldots & \lambda_n\mathbf{x}_n \end{pmatrix}.$$

La descomposición espectral expresa un cambio de base en las coordenadas dadas por los autovectores de A.

Descomposición Espectral

El conjunto de autovectores correspondientes a un solo autovalor λ de A, junto con el vector cero, forma un subespacio de E_λ conocido como el **autoespacio asociado a** λ.

Obs! $E_\lambda \subseteq \mathbb{C}^n$. Es un subespacio invariante: $AE_\lambda \subseteq E_\lambda$.

Definición

*La **multiplicidad geométrica** de λ es $\dim_{\mathbb{R}} E_\lambda = \dim_{\mathbb{R}} \mathrm{Ker}(\lambda I - A)$.*

Recordemos que el polinomio característico de $A \in \mathbb{R}^{n\times n}$ es el polinomio de grado n dado por $p_A(z) = \det(zI - A)$.

Teorema

λ es un autovalor de A si, y sólo si, $p_A(\lambda) = 0$.

Prueba:

$$\begin{aligned} \lambda \text{ es autovalor de } A \quad &\Leftrightarrow \quad \text{existe } \mathbf{x} \neq \mathbf{0} : \; A\mathbf{x} = \lambda \mathbf{x} \;\Leftrightarrow\; \lambda \mathbf{x} - A\mathbf{x} = \mathbf{0} \\ &\Leftrightarrow \quad \lambda I - A \text{ es singular } \;\Leftrightarrow\; \det(\lambda I - A) = 0. \; \square \end{aligned}$$

Por el Teorema Fundamental del Álgebra, una consecuencia del teorema anterior es que

$$p_A(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n),$$

donde los λ_j son todos autovalores de A.

Definición

*La **multiplicidad algebraica** del autovalor λ es su multiplicidad como raíz del polinomio característico $p_A(z)$.*

Descomposición Espectral

Observaciones:

- Toda matriz cuadrada $A \in \mathbb{R}^{n\times n}$ posee n autovalores (contados con multiplicidad).
- Si las raíces de $p_A(z)$ son todas simples, A posee n autovalores distintos.
- Cuando X es no singular, las matrices A y XAX^{-1} poseen igual polinomio característico, autovalores, y multiplicidades algebraicas y geométricas, ya que

$$p_{XAX^{-1}}(z) = \det(zI - XAX^{-1}) = \det\left(X(zI-A)X^{-1}\right) = \det X \det(zI-A)(\det X)^{-1} = \det(zI-A) = p_A(z).$$

y E_λ es autoespacio para $A \Leftrightarrow X^{-1}E_\lambda$ es autoespacio para XAX^{-1}.

Teorema

multiplicidad algebraica $\lambda \geq$ *multiplicidad geométrica* λ, $\forall \lambda$ *autovalor.*

Descomposición Espectral

<u>Prueba</u>: Sea k la multiplicidad geométrica de λ para la matriz $A \in \mathbb{R}^{n\times n}$. Formamos una matriz $V \in \mathbb{R}^{n\times k}$ cuyas k columnas constituyen una base ortonormal del autoespacio $E_\lambda = \{\mathbf{x} : A\mathbf{x} = \lambda\mathbf{x}\}$. Extendemos V a una matriz unitaria cuadrada $V \in \mathbb{R}^{n\times n}$, y obtenemos V^*AV en la forma

$$B = V^*AV = \begin{pmatrix} \lambda I_k & C \\ \mathbf{0} & D \end{pmatrix},$$

donde $C \in \mathbb{R}^{k\times(n-k)}$, $D \in \mathbb{R}^{(n-k)\times(n-k)}$.

Luego, $\det(zI - B) = \det(zI - \lambda I_k)\,\det(zI - D) = (z-\lambda)^k\,\det(\lambda I - D)$.

De ahí que la multiplicidad algebraica de λ como autovalor de B es al menos k, y dado que las transformaciones de semejanza preservan multiplicidades, lo mismo es cierto para A. $\square$

Descomposición Espectral

Definición

Un autovalor cuya multiplicidad algebraica excede su multiplicidad geométrica es un ***autovalor defectuoso****. Una matriz que tiene uno o más autovalores defectuosos se llama una* ***matriz defectuosa****.*

- Las matrices diagonales no son defectuosas.
- Veremos luego que las matrices simétricas no son defectuosas.

Definición

Una matriz cuadrada $A \in \mathbb{R}^{n\times n}$ es ***diagonalizable*** *si admite una descomposición espectral.*

Teorema

Una matriz $A \in \mathbb{R}^{n\times n}$ es diagonalizable si, y sólo si, no es defectuosa.

Prueba: ($\Rightarrow$) Dada una descomposición espectral $A = X\Lambda X^{-1}$ para A, $\Rightarrow$ Λ es similar a A, con los mismos autovalores y las mismas multiplicidades. Como Λ es diagonal, no es defectuosa y, por lo tanto, lo mismo vale para A.

($\Leftarrow$) Una matriz no defectuosa debe tener n autovectores l.i. (ya que autovectores de diferente autovalor son l.i., y cada cada autovalor contribuye exactamente con tantos autovectores l.i. como su multiplicidad geométrica). Si estos n autovectores l.i. se forman en las columnas de una matriz X, entonces X es no singular y tenemos $A = X\Lambda X^{-1}$. $\square$

Descomposición Espectral

Ejemplo: Considere las matrices

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad \text{y} \qquad B = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Tanto A como B tienen un polinomio característico $(z-2)^3$, por lo que hay un solo autovalor $\lambda = 2$ de multiplicidad algebraica 3.

En el caso de A, el autoespacio $E_\lambda = \langle \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \rangle$, mientras que para B, $E_\lambda = \langle \mathbf{e}_1 \rangle$, de modo que λ es un autovalor defectuoso para B.
Portanto, A es diagonalizable, pero B no.

(Ver libro *Linear Algebra* de Hoffman y Kunze).

Descomposición Espectral

En ocasiones, no sólo ocurre que una matriz $A \in \mathbb{R}^{n\times n}$ admite n autovectores l.i., sino que éstos son ortogonales.

Definición

*Decimos que A es **unitariamente diagonalizable** si existe una matriz ortogonal U tal que $A = U\Lambda U^T$.*

Teorema (Teorema espectral / Descomposición espectral)

Sea $A \in \mathbb{R}^{d\times d}$ una matriz simétrica (operador auto-adjunto). Entonces, A admite una descomposición de la forma

$$A = U\Lambda U^T,$$

donde $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_d)$ es la matriz diagonal formada por los autovalores $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_d$ de A, y

Descomposición Espectral

$$U = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_d \end{pmatrix} \in \mathbb{R}^{d\times d}$$

es una matriz ortogonal cuyas columnas son los autovectores de A, con $\mathbf{u}_i$ el autovector correspondiente a λ_i, $i = 1, 2, \dots, d$.

En otras palabras, A es una suma de matrices de rango 1: $A = \sum_{i=1}^{d} \lambda_i\, \mathbf{u}_i\mathbf{u}_i^T$. Comentarios:

- El teorema espectral dice que toda matriz cuadrada A, real, simétrica, posee una base ortonormal, formada por autovectores de A.
- Para $1 \leq k \leq d$, la suma $A = \sum_{i=1}^{k} \lambda_i\, \mathbf{u}_i\mathbf{u}_i^T$, es una matriz de rango k.

Descomposición Espectral

Observaciones:

- A simétrica y semi-definida positiva, $\Rightarrow$ existe $A^{1/2}$ tal que $A^{1/2}A^{1/2} = A$.
- Si todos los autovalores de A son no-negativos, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_d \geq 0$, entonces $\Lambda^{1/2}$ existe y

$$\Lambda^{1/2} = diag(\lambda_1, \lambda_2, \ldots, \lambda_d)^{1/2} = diag(\lambda_1^{1/2}, \lambda_2^{1/2}, \ldots, \lambda_d^{1/2}).$$

- A partir de la descomposición espectral podemos calcular $A^{1/2}$. De hecho, si $A = U\Lambda U^T$, definimos $A^{1/2} = U\Lambda^{1/2}U^T$, y

$$\begin{aligned} A^{1/2}A^{1/2} &= (U\Lambda^{1/2}U^T)(U\Lambda^{1/2}U^T) = U\Lambda^{1/2}(U^TU)\Lambda^{1/2}U^T \\ &= U\Lambda^{1/2}\Lambda^{1/2}U^T = U\Lambda U^T = A. \end{aligned}$$

- En general $A \succeq 0 \;\Rightarrow A^p = U\Lambda^p U^T$, para todo $p \in \mathbb{R}$.

Teorema Espectral

Prueba: (Teorema espectral).
La prueba es por inducción sobre $n = \dim A$. Para $n = 1$, el resultado es inmediato pues

$$A = [a] = [1][a][1] = [1][a][1]^T.$$

Suponga que el resultado es válido para $n-1$. Mostramos que es posible extenderlo a n. Para $n > 1$, sea $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ una base ortonormal cualquiera para $\mathbb{R}^n$, y representamos un vector $\mathbf{x} \in \mathbb{R}^n$ arbitrario por

$$\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \ldots + \alpha_n\mathbf{u}_n.$$

Podemos pensar $\mathbf{w} : \mathbb{R}^n \to \mathbb{R}^n$ como una función de las coordenadas, esto es $\mathbf{w} = \mathbf{w}(\alpha_1, \ldots, \alpha_n)$.
Consideramos el problema de optimización:

$$\text{maximizar } \langle \mathbf{w}, A\mathbf{w}\rangle = \mathbf{w}^T A\mathbf{w}, \quad \text{sujeto a} ||\mathbf{w}||_2 = \mathbf{w}^T\mathbf{w} = 1. \tag{1}$$

El lagrangiano de este problema es $\mathcal{L}(\mathbf{w}) = \mathbf{w}^T A\mathbf{w} - \lambda(\mathbf{w}^T\mathbf{w} - 1)$. Del criterio de optimalidad, resulta

$$\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w},\lambda) = 2A\mathbf{w} - 2\lambda\mathbf{w} = 0, \quad \nabla_{\lambda}\mathcal{L}(\mathbf{w},\lambda) = \mathbf{w}^T\mathbf{w} - 1 = 0. \tag{2}$$

La primera condición en (2) implica que $A\mathbf{w} = \lambda\mathbf{w}$, de modo que el $\mathbf{w}^*$ óptimo debe ser un autovector de A, y λ su autovalor. La segunda condición implica que $||\mathbf{w}^*||_2 = 1$. De hecho $\max \mathbf{w}^T A\mathbf{w} = \max \mathbf{w}^T(\lambda\mathbf{w}) = \max \lambda||\mathbf{w}||_2^2 = \max \lambda$, implica que $\mathbf{w}^*$ corresponde al autovector de λ_1, el mayor autovalor de A.

Hallado el máximo $\mathbf{w}^*$, hacemos $\mathbf{u}_1 = \frac{\mathbf{w}}{||\mathbf{w}||_2}$ y $\lambda_1 = \lambda$. Consideramos ahora el subespacio $W_2 = (\langle\mathbf{w}\rangle)^\perp \equiv \mathbb{R}^{n-1}$. Como $A : \langle\mathbf{w}\rangle \to \langle\mathbf{w}\rangle$, entonces también $A : \langle\mathbf{w}\rangle^\perp \to \langle\mathbf{w}\rangle^\perp$

Esto se debe a que si $\mathbf{x} \in \langle \mathbf{w} \rangle^{\perp}$, entonces
$\langle \mathbf{x}, \mathbf{w} \rangle = 0 \;\Rightarrow\; \langle A\mathbf{x}, \mathbf{w} \rangle = \langle \mathbf{x}, A\mathbf{w} \rangle = \langle \mathbf{x}, \lambda\mathbf{w} \rangle = \lambda\langle \mathbf{x}, \mathbf{w} \rangle = 0$, así $A\mathbf{x} \in \langle \mathbf{w} \rangle^{\perp}$.

Por inducción, el espacio $(n-1)$-dimensional W_0 tiene una base ortonormal $\{\mathbf{u}_2, \ldots, \mathbf{u}_n\}$ formada por autovectores de A, cada uno con un autovalor real, y por construcción $\mathbf{u}_1$ es un vector unitario ortogonal a cada uno de $\{\mathbf{u}_2, \ldots, \mathbf{u}_n\}$.

Entonces $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ es una base ortonormal para $\mathbb{R}^n$ formada por autovectores de A, lo que prueba el teorema. $\square$