

Clase 22. Diagonalización y Potencia de una Matriz.

Curso ‘Linear Algebra’ del MIT.

Resumen

Una aplicación de los eigenvalores y eigenvectores de una matriz cuadrada, es al **diagonalizarla**. Cuando es posible, dicho proceso resulta en una nueva factorización que estudiaremos a continuación. Además, a partir de ella veremos qué ocurre con estos valores y vectores característicos cuando a dicha matriz la elevamos a una potencia.

1. Diagonalización de una Matriz.

Cuando los **eigenvectores** de una matriz cuadrada A son **linealmente independientes**, es posible **diagonalizar** a esta última de la siguiente manera:

$$S^{-1}AS = \Lambda$$

donde S es una matriz cuyos vectores columna son los eigenvectores de A y Λ (lambda mayúscula) es una matriz diagonal donde su diagonal principal son los eigenvalores de A .

En ese sentido, Λ es la matriz A diagonalizada por S .

Veamos algebraicamente el proceso de encontrar a Λ .

Digamos que A es de $n \times n$ y que todos sus eigenvalores λ_i son distintos entre sí. Como vimos en la clase anterior, esto implica que sus eigenvectores \vec{x}_i son linealmente independientes.

Luego, construyamos una matriz S donde sus columnas son los eigenvectores \vec{x}_i de A , lo que significa que su dimensión también será de $n \times n$.

Para diagonalizar a A , primero multipliquemos a esta matriz por S a la derecha.

$$AS = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \cdot \begin{bmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,n} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n,1} & s_{n,2} & \cdots & s_{n,n} \end{bmatrix}$$

Ahora bien, al multiplicar AS ocurre que los vectores columna de la matriz resultante corresponden a $A\vec{x}_i$.

$$AS = \left[A \cdot \begin{bmatrix} s_{1,1} \\ s_{2,1} \\ \vdots \\ s_{n,1} \end{bmatrix} \quad A \cdot \begin{bmatrix} s_{1,2} \\ s_{2,2} \\ \vdots \\ s_{n,2} \end{bmatrix} \quad \cdots \quad A \cdot \begin{bmatrix} s_{1,n} \\ s_{2,n} \\ \vdots \\ s_{n,n} \end{bmatrix} \right] = \begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ A\vec{x}_1 & A\vec{x}_2 & \cdots & A\vec{x}_n \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix}$$

Y, como \vec{x}_i es el eigenvector de A , entonces:

$$A\vec{x}_i = \lambda_i \vec{x}_i$$

con λ_i siendo los i -ésimos eigenvalores de A . Esto implica que:

$$AS = \begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ \lambda_1 \vec{x}_1 & \lambda_2 \vec{x}_2 & \cdots & \lambda_n \vec{x}_n \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix} = \begin{bmatrix} \lambda_1 s_{1,1} & \lambda_2 s_{1,2} & \cdots & \lambda_n s_{1,n} \\ \lambda_1 s_{2,1} & \lambda_2 s_{2,2} & \cdots & \lambda_n s_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 s_{n,1} & \lambda_2 s_{n,2} & \cdots & \lambda_n s_{n,n} \end{bmatrix}$$

La matriz resultante podemos factorizarla como:

$$AS = \begin{bmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,n} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n,1} & s_{n,2} & \cdots & s_{n,n} \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Establezcamos que

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

y reescribamos la ecuación como:

$$AS = S\Lambda$$

Para despejar a Λ , podemos multiplicar por S^{-1} por la izquierda en la ecuación puesto que, como todas las columnas de S son linealmente independientes, es posible dar por garantizado que dicha matriz tiene una inversa. En consecuencia:

$$S^{-1}AS = S^{-1}S\Lambda$$

$$S^{-1}AS = I\Lambda$$

$$S^{-1}AS = \Lambda$$

Debido a que S es invertible, a partir de la ecuación de arriba podemos obtener una **factorización de la matriz** A . Comencemos multiplicándola por S^{-1} por la derecha de sus dos lados.

$$S^{-1}ASS^{-1} = \Lambda S^{-1}$$

$$S^{-1}AI = \Lambda S^{-1}$$

$$S^{-1}A = \Lambda S^{-1}$$

Luego, multipliquemos por S a la izquierda de ambos lados de la ecuación.

$$SS^{-1}A = S\Lambda S^{-1}$$

$$IA = S\Lambda S^{-1}$$

$$A = S\Lambda S^{-1}$$

2. La Potencia de una Matriz.

Con los eigenvalores y eigenvectores de una matriz podemos calcular la potencia de una matriz A .

Sean A una matriz de $n \times n$, \vec{x} su eigenvector y λ su eigenvalor tal que

$$A\vec{x} = \lambda\vec{x}$$

Multipliquemos por A a la izquierda en ambos lados de la ecuación.

$$AA\vec{x} = A\lambda\vec{x}$$

$$A^2\vec{x} = \lambda A\vec{x}$$

Puesto que \vec{x} es el eigenvector de A , entonces:

$$A^2\vec{x} = \lambda\lambda\vec{x}$$

$$A^2\vec{x} = \lambda^2\vec{x}$$

En otras palabras, si elevamos al cuadrado una matriz de $n \times n$, **sus eigenvalores también lo harán, mientras que sus eigenvectores se mantendrán iguales.**

Lo anterior también podemos verlo en forma matricial a partir de la diagonalización de A .

Supongamos que A tiene n eigenvalores distintos entre sí los cuales se encuentran en la diagonal principal de Λ y, como consecuencia, n eigenvectores ubicados en una matriz S . Si a la ecuación:

$$A = S\Lambda S^{-1}$$

la multiplicamos por A por la derecha, entonces:

$$AA = S\Lambda S^{-1}A$$

$$A^2 = S\Lambda S^{-1}S\Lambda S^{-1}$$

$$A^2 = S\Lambda\Lambda S^{-1}$$

$$A^2 = S\Lambda^2 S^{-1}$$

Lo anterior podemos generalizarlo para cualquier matriz A de $n \times n$ elevada a la k -ésima potencia con n **eigenvectores linealmente independientes.**

$$A^k = S\Lambda^k S^{-1}$$

Otro hecho que puede desprenderse de lo que hemos estudiado, es que¹

$$\lim_{k \rightarrow \infty} A^k = 0 \iff |\lambda_i| < 1, \text{ para } i = 1, 2, \dots, n$$

En otras palabras, $A^k \rightarrow 0$ a medida que $k \rightarrow \infty$, siempre que cada eigenvalor λ_i de A sea menor a 1.

2.1 Eigenvectores Linealmente Independientes.

Hemos mencionado que para diagonalizar a una matriz cuadrada A necesitamos que sus eigenvectores sean linealmente independientes. La clase pasada revisamos que esto es posible

¹Se da por hecho, pero no se explica en la clase por qué esto es verdadero cuando $|\lambda_i| < 1$.

cuando sus eigenvalores son distintos entre sí. No obstante, existen casos donde siendo iguales aún así se cumple lo señalado. Un caso es en la **matriz identidad** I .

Digamos que I es de 2×2 . Sus eigenvalores serán $\lambda_1 = \lambda_2 = 1$ o, en otras palabras, tiene una **multiplicidad algebraica**² igual a 2. Al calcular $(I - \lambda I)$ obtenemos que:

$$(I - \lambda I) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Mientras que el espacio nulo $(I - \lambda I)\vec{x} = \vec{0}$ estará generado (*span*) por los eigenvectores:

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \vec{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

A pesar de que $\lambda_1 = \lambda_2$, aún así obtuvimos dos eigenvectores **linealmente independientes**. Es decir, asimismo su **multiplicidad geométrica**³ es igual a 2.

Esto también podemos verlo algebraicamente a partir de la matriz diagonalizada Λ

$$S^{-1}IS = \Lambda$$

$$S^{-1}S = \Lambda$$

$$I = \Lambda$$

o de la definición de los eigenvectores y eigenvalores de una matriz I .

$$I\vec{x} = \lambda\vec{x}$$

$$\vec{x} = \lambda\vec{x}$$

La única posibilidad de que $\vec{x} = \lambda\vec{x}$ es que $\lambda = 1$.

Por lo tanto, si los eigenvalores de una matriz son iguales entre sí, existe la posibilidad de que los eigenvectores sean linealmente dependientes o independientes (todos o algunos). No obstante, si dichos escalares son todos distintos entre sí, podemos dar por garantizado que los vectores característicos serán linealmente independientes.

²**Multiplicidad algebraica:** Cantidad de veces que se repite un eigenvalor λ . Por ejemplo, si los λ_i de una matriz están dados por $(\lambda - b)^3 = 0$, con $b = \text{constante}$, su **multiplicidad algebraica** es 3.

³**Multiplicidad geométrica:** Cantidad de eigenvectores linealmente independientes de una matriz.

2.2 Aplicación: Ecuaciones de Diferencia (*Difference Equations*).

Veamos un uso del cálculo de la potencia de una matriz cuadrada usando sus eigenvalores/vectores en **Ecuaciones de Diferencia**.

Las **ecuaciones de diferencia** y las **relaciones de recurrencia** suelen ser conceptos intercambiados. Corresponden a ecuaciones que expresan a cada miembro a_n de una **secuencia** en términos (o en función) de los que lo **antecedan**. Es decir, son definidas de forma **recursiva**.

Al igual que en las ecuaciones diferenciales⁴, al resolver una ecuación de diferencia obtenemos su solución general. Para conocer una **solución particular**, debemos establecer previamente una **condición inicial**.

En esta ocasión nos centraremos en ecuaciones de diferencia (o relaciones de recurrencia) lineales de primer orden con coeficientes constantes. A continuación tenemos una misma ecuación de este tipo escrita de dos maneras comúnmente usadas.

$$a_{n+1} = ca_n + b \quad \text{o} \quad a(n+1) = ca(n) + b$$

donde la linealidad por el hecho de ser una **combinación lineal de su término previo** (lado derecho) y el orden por el número de términos posteriores⁵, que en este caso es 1 en $n+1$. En cuanto a los coeficientes constantes, estos son $c, b \in \mathbb{R}$.

En particular, nos concentraremos en ecuaciones de diferencia **homogéneas**. Es decir, para las cuales $b = 0$.

Ejemplo 1. Sea A la matriz de coeficientes de $n \times n$ y diagonalizable del siguiente sistema de ecuaciones de diferencia

$$\vec{u}_{k+1} = A\vec{u}_k$$

Busque la solución del sistema para la condición inicial \vec{u}_0 .

Solución. Comencemos viendo dos iteraciones $k = 0, 1$ del sistema.

$$\vec{u}_1 = A\vec{u}_0$$

$$\vec{u}_2 = A\vec{u}_1 = A(A\vec{u}_0) = A^2\vec{u}_0$$

Es posible observar que, a medida que aumenta k , en la misma magnitud lo hace el exponente

⁴Las ecuaciones diferenciales y las de diferencia difieren en que las primeras tratan con variables continuas, mientras que las segundas con discretas. No obstante, estas últimas suelen ser usadas para resolver a las primeras mediante métodos numéricos.

⁵También pueden ser términos anteriores, el cual podría estar denotado como $a_n = ca_{n-1} + b$.

de A , mientras que \vec{u}_0 se mantiene constante. Por lo tanto, podemos generalizar que:

$$\vec{u}_k = A^k \vec{u}_0$$

La dificultad aún se mantiene porque no es sencillo calcular A^k , pero debido a que A es diagonalizable, la estrategia que podemos tomar es escribir al vector inicial \vec{u}_0 como una **combinación lineal de eigenvectores** de A .

$$\vec{u}_0 = c_1 \vec{x}_1 + c_2 \vec{x}_2 + \cdots + c_n \vec{x}_n = S\vec{c},$$

donde S es la matriz de los eigenvectores columna de A y \vec{c} es un vector de los coeficientes c_i de la combinación lineal.

Reemplacemos al vector \vec{u}_0 en $\vec{u}_k = A^k \vec{u}_0$.

$$\vec{u}_k = A^k (S\vec{c})$$

Anteriormente vimos que cuando A es diagonalizable, se cumple que

$$A = S\Lambda S^{-1}$$

y también que cuando está elevada a una potencia k , entonces

$$A^k = S\Lambda^k S^{-1}$$

Por lo tanto, reemplazando A^k en $\vec{u}_k = A^k (S\vec{c})$, obtenemos lo siguiente:

$$\begin{aligned} \vec{u}_k &= (S\Lambda^k S^{-1})(S\vec{c}) \\ &= S\Lambda^k I\vec{c} \\ \vec{u}_k &= S\Lambda^k \vec{c} \end{aligned}$$

donde $\vec{u}_k = S\Lambda^k \vec{c}$ es la solución del sistema de ecuaciones de diferencia $\vec{u}_{k+1} = A\vec{u}_k$.

Ejemplo 2. La siguiente ecuación de diferencia corresponde a una secuencia de Fibonacci:

$$F_{k+2} = F_{k+1} + F_k$$

1. Busque una fórmula que permita obtener el número F_{100} de Fibonacci, cuyas condiciones iniciales son $F_0 = 0$ y $F_1 = 1$

2. Evalúe a qué tasa crecen los números de Fibonacci dados por su ecuación.

Solución. La secuencia de Fibonacci corresponde a una ecuación de diferencia lineal homogénea de segundo orden. Sería interesante poder expresarla como un sistema de ecuaciones de primer orden tal que podamos expresarla como $\vec{u}_{k+1} = A\vec{u}_k$, de la cual conocemos su solución.

Para ello, establezcamos que:

$$\vec{u}_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$$

Esto implica que:

$$\vec{u}_{k+1} = \begin{bmatrix} F_{k+2} \\ F_{k+1} \end{bmatrix}$$

Para no alterar la ecuación original, definamos a la matriz A como:

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

El sistema corresponde a:

$$\begin{aligned} \vec{u}_{k+1} &= A\vec{u}_k \\ \begin{bmatrix} F_{k+2} \\ F_{k+1} \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} \end{aligned}$$

el cual es expresado de la siguiente manera como sistema de ecuaciones:

$$\begin{cases} F_{k+2} = F_{k+1} + F_k \\ F_{k+1} = F_{k+1} \end{cases}$$

Desde el ejemplo anterior sabemos que la solución del sistema $\vec{u}_{k+1} = A\vec{u}_k$ es $\vec{u}_k = S\Lambda^k\vec{c}$ siempre que A sea diagonalizable. Evaluemos aquello en este caso, calculando sus eigenvalores.

Si nos damos cuenta, A es simétrica, lo que implica que obtendremos eigenvalores reales y eigenvectores ortonormales. Por otra parte, $\text{tr}(A) = 1$ y $\det(A) = -1$, por lo que debe cumplirse que $\sum_{i=1}^2 \lambda_i = 1$ y $\prod_{i=1}^2 \lambda_i = -1$.

Como A es una matriz de 2×2 , calculemos sus eigenvalores a partir de la ecuación característica.

$$\begin{aligned} \det(A - \lambda I) &= 0 \\ \begin{vmatrix} 1 - \lambda & 1 \\ 1 & -\lambda \end{vmatrix} &= 0 \\ \lambda^2 - \lambda - 1 &= 0 \end{aligned}$$

Usemos la fórmula cuadrática⁶ para conocer las raíces del polinomio característico:

$$\lambda = \frac{-(-1) \pm \sqrt{1^2 - (-4)}}{2} = \frac{1 \pm \sqrt{5}}{2}$$

Por lo tanto, los eigenvalores de A son:

$$\lambda_1 = \frac{1 + \sqrt{5}}{2} = 1.618\dots \quad y \quad \lambda_2 = \frac{1 - \sqrt{5}}{2} = -0.618\dots$$

donde:

$$\begin{aligned} \lambda_1 + \lambda_2 &= \frac{1 + \sqrt{5} + 1 - \sqrt{5}}{2} = 1 \\ \lambda_1 \cdot \lambda_2 &= \frac{(1 + \sqrt{5}) \cdot (1 - \sqrt{5})}{2} = -2 \end{aligned}$$

La suma de los eigenvalores es igual a la traza de A , pero algebraicamente el producto de estos escalares no coincide con el determinante de dicha matriz. No obstante, si realizamos aquella operación con los valores numéricos aproximados, veremos que son muy cercanos a -1 .

$$\lambda_1 \cdot \lambda_2 \approx 1.618 \cdot (-0.618) \approx -0.999924$$

Luego, calculemos los eigenvectores de A :

$$\begin{aligned} (A - \lambda_i I) \vec{x}_i &= \vec{0} \\ \begin{bmatrix} 1 - \lambda_i & 1 \\ 1 & -\lambda_i \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$

Al calcular numéricamente los eigenvectores, tanto desde un computador como a mano, vamos a obtener resultados distintos⁷, pero podemos tomar algunas consideraciones.

Primero, la matriz $(A - \lambda_i I)$ es singular, donde su entrada $-\lambda_i$ será cero (o aproximado a dicho valor). Por lo tanto, consideremos que $1 - \lambda_i$ es pivote. De este modo, la entrada x_2 estará libre y buscaremos a x_1 , por lo que podemos establecer que:

$$\vec{x}_i = \begin{bmatrix} \lambda_i \\ 1 \end{bmatrix}$$

⁶Las raíces de una ecuación cuadrática $ax^2 + bx + c = 0$ es posible obtenerlas por medio de la fórmula cuadrática:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

⁷Depende en gran medida del algoritmo que usemos y, para el caso de un software o una calculadora, influirá la precisión punto-flotante.

Y debido a lo anterior, obtendremos que:

$$\vec{x}_1 = \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} \quad \text{y} \quad \vec{x}_2 = \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}$$

Así, el vector solución \vec{u}_k del sistema $\vec{u}_{k+1} = A\vec{u}_k$ es:

$$\begin{aligned} \vec{u}_k &= S\Lambda^k\vec{c} \\ \vec{u}_k &= \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \end{aligned}$$

para:

$$\lambda_1 = \frac{1 + \sqrt{5}}{2} \quad \text{y} \quad \lambda_2 = \frac{1 - \sqrt{5}}{2}$$

Con la fórmula que obtuvimos, podemos buscar F_{100} estableciendo que $k = 99$. En otras palabras, **tenemos que elevar a k los eigenvalores para obtener aquel número de Fibonacci.**

A partir de la ecuación $\vec{u}_{k+1} = A\vec{u}_k$ también podemos analizar a qué tasa crecen los números de Fibonacci con condición inicial $F_0 = 0$ y $F_1 = 1$. Como acabamos de ver, el crecimiento está dado por los eigenvalores elevados a la potencia de k .

$$\lambda_1^k = \left(\frac{1 + \sqrt{5}}{2} \right)^k \quad \text{y} \quad \lambda_2^k = \left(\frac{1 - \sqrt{5}}{2} \right)^k$$

Por lo tanto, a medida que k crece, λ_1 también lo hace, pero λ_2 disminuye. En otras palabras:

$$\lim_{k \rightarrow \infty} \lambda_1^k = \lim_{k \rightarrow \infty} \left(\frac{1 + \sqrt{5}}{2} \right)^k = \infty \quad \text{y} \quad \lim_{k \rightarrow \infty} \lambda_2^k = \lim_{k \rightarrow \infty} \left(\frac{1 - \sqrt{5}}{2} \right)^k = 0$$

En ese sentido, podemos decir que la **tasa de crecimiento de la secuencia de Fibonacci está determinado por un factor**, que es su **eigenvalor** más grande: λ_1^k .

$$\lambda_1^k = \left(\frac{1 + \sqrt{5}}{2} \right)^k$$

Es decir,

$$F_k \approx c_1 \cdot \left(\frac{1 + \sqrt{5}}{2} \right)^k$$

donde $c_1 = \text{constante}$.