

AUTOVALORES Y AUTOVECTORES

DEFINICIONES. INTERPRETACIÓN GEOMÉTRICA. CÍRCULOS DE GERSCHGORIN. MÉTODO DE LAS POTENCIAS.
MÉTODO DE LA POTENCIA: CÓDIGO. FACTORIZACIÓN QR. CÓDIGO.

Manuel Carlevaro

Departamento de Ingeniería Mecánica

Grupo de Materiales Granulares - UTN FRLP

manuel.carlevaro@gmail.com

Definición : Autovalor y autovector.

Sea $A \in K^{n \times n}$ y $v \in K^n$. v es un **autovector** de A si

$$Av = \lambda v$$

donde λ es un escalar en K , denominado **autovalor** asociado con v .

Definición : Autovalor y autovector.

Sea $\mathbf{A} \in K^{n \times n}$ y $\mathbf{v} \in K^n$. \mathbf{v} es un **autovector** de \mathbf{A} si

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

donde λ es un escalar en K , denominado **autovalor** asociado con \mathbf{v} .

En forma equivalente:

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0} \quad (1)$$

Este sistema tiene solución $\mathbf{v} \neq \mathbf{0}$ si y solo si:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

denominado **polinomio característico**, $p_A(\lambda)$, y por el teorema fundamental del álgebra: $\mapsto n$ raíces.

Definición : Autovalor y autovector.

Sea $\mathbf{A} \in K^{n \times n}$ y $\mathbf{v} \in K^n$. \mathbf{v} es un **autovector** de \mathbf{A} si

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

donde λ es un escalar en K , denominado **autovalor** asociado con \mathbf{v} .

En forma equivalente:

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0} \quad (1)$$

Este sistema tiene solución $\mathbf{v} \neq \mathbf{0}$ si y solo si:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

denominado **polinomio característico**, $p_A(\lambda)$, y por el teorema fundamental del álgebra: $\mapsto n$ raíces.

Ejemplo:

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix}$$

$$\begin{aligned} p_A(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} 3-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 3-\lambda \end{vmatrix} \\ &= -\lambda^3 + 8\lambda^2 - 19\lambda + 12 = 0 \end{aligned}$$

Solución: $\lambda_1 = 1, \lambda_2 = 3, \lambda_3 = 4$. Reemplazando cada autovalor en (1):

$$\mathbf{v}_1 = a \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{v}_2 = b \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \mathbf{v}_3 = c \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Definición : Autovalor y autovector.

Sea $\mathbf{A} \in K^{n \times n}$ y $\mathbf{v} \in K^n$. \mathbf{v} es un **autovector** de \mathbf{A} si

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

donde λ es un escalar en K , denominado **autovalor** asociado con \mathbf{v} .

En forma equivalente:

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0} \quad (1)$$

Este sistema tiene solución $\mathbf{v} \neq \mathbf{0}$ si y solo si:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

denominado **polinomio característico**, $p_A(\lambda)$, y por el teorema fundamental del álgebra: $\mapsto n$ raíces.

Ejemplo:

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix}$$

$$\begin{aligned} p_A(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} 3-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 3-\lambda \end{vmatrix} \\ &= -\lambda^3 + 8\lambda^2 - 19\lambda + 12 = 0 \end{aligned}$$

Solución: $\lambda_1 = 1, \lambda_2 = 3, \lambda_3 = 4$. Reemplazando cada autovalor en (1):

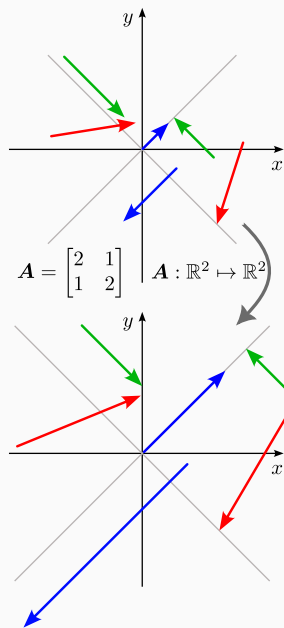
$$\mathbf{v}_1 = a \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{v}_2 = b \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \mathbf{v}_3 = c \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Definiciones:

Espectro de \mathbf{A} : $\sigma(\mathbf{A}) = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$

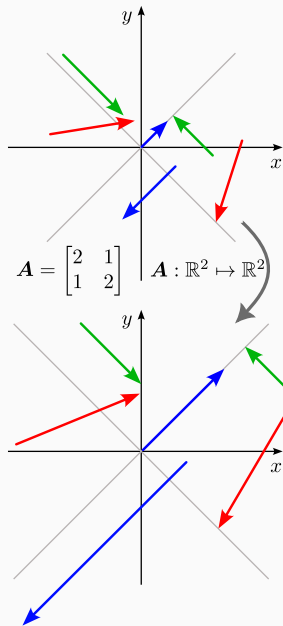
Radio espectral de \mathbf{A} : $\rho(\mathbf{A}) = \max_{\lambda \in \sigma(\mathbf{A})} |\lambda|$

Intepretación gráfica



$$\begin{aligned} p_A(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{vmatrix} \\ &= (2 - \lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1) = 0 \end{aligned}$$

Intepretación gráfica

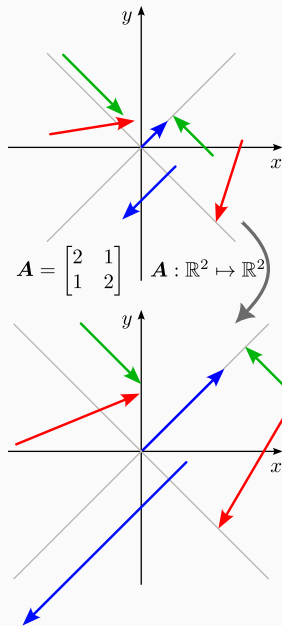


$$\begin{aligned} p_A(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{vmatrix} \\ &= (2 - \lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1) = 0 \end{aligned}$$

Con $\lambda_1 = 3$:

$$\begin{cases} (2 - 3)x + y = 0 \\ x + (2 - 3)y = 0 \end{cases} \implies \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Interpretación gráfica



$$\begin{aligned} p_A(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{vmatrix} \\ &= (2 - \lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1) = 0 \end{aligned}$$

Con $\lambda_1 = 3$:

$$\begin{cases} (2 - 3)x + y = 0 \\ x + (2 - 3)y = 0 \end{cases} \implies \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Con $\lambda_2 = 1$:

$$\begin{cases} (2 - 1)x + y = 0 \\ x + (2 - 1)y = 0 \end{cases} \implies \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Nota: si consideramos la norma vectorial $l_2 : \|\cdot\|$:

$$\|\mathbf{A}\|_2 = \rho(\mathbf{A}^\dagger \mathbf{A})^{1/2}$$

Si \mathbf{A} es simétrica, $\|\mathbf{A}\|_2 = \rho(\mathbf{A})$.

Métodos:

- ▶ Analítico: $n < 5$.
 - ▶ Parciales: computan solo autovalores **extremos** (módulo máximo o mínimo). Método de las potencias.
 - ▶ Globales: aproximan a todo el **espectro** de \mathbf{A} , $\sigma(\mathbf{A})$. Método QR .
-

Métodos:

- ▶ Analítico: $n < 5$.
- ▶ Parciales: computan solo autovalores **extremos** (módulo máximo o mínimo). Método de las potencias.
- ▶ Globales: aproximan a todo el **espectro** de A , $\sigma(A)$. Método QR .

Teorema : Círculo de Gerschgorin.

$A \in K^{n \times n}$, $r_i = \sum_{j \neq i}^n |a_{ij}|$ para cada $i = 1, 2, \dots, n$. Sea

$$C_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}$$

1. Si λ es un autovalor, está en uno de los C_i .
2. Si k círculos C_i forman una región conectada $R \in \mathbb{C}$, disjunta de los restantes $n - k$ círculos, entonces R contiene exactamente k autovalores.

Métodos:

- ▶ Analítico: $n < 5$.
- ▶ Parciales: computan solo autovalores **extremos** (módulo máximo o mínimo). Método de las potencias.
- ▶ Globales: aproximan a todo el **espectro** de \mathbf{A} , $\sigma(\mathbf{A})$. Método QR .

Teorema : Círculo de Gerschgorin.

$\mathbf{A} \in K^{n \times n}$, $r_i = \sum_{j \neq i}^n |a_{ij}|$ para cada $i = 1, 2, \dots, n$. Sea

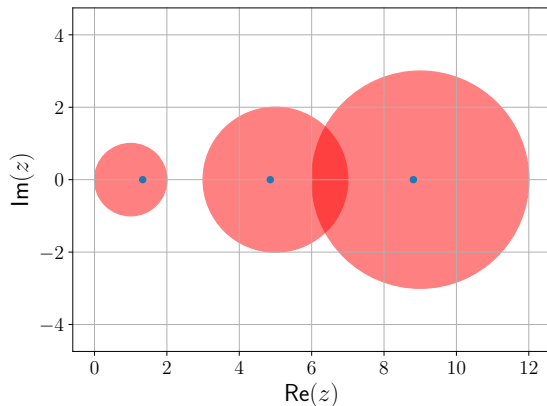
$$C_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}$$

1. Si λ es un autovalor, está en uno de los C_i .
2. Si k círculos C_i forman una región conectada $R \in \mathbb{C}$, disjunta de los restantes $n - k$ círculos, entonces R contiene exactamente k autovalores.

Ejemplo:

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & 5 & 1 \\ -2 & -1 & 9 \end{bmatrix}$$

$$r_1 = |-1| + |0| = 1, r_2 = |1| + |1| = 2, r_3 = |-2| + |-1| = 3. \\ \lambda_1 = 1.33192769, \lambda_2 = 8.81113862, \lambda_3 = 4.85693369.$$



Método de las potencias:

$\mathbf{A} \in \mathbb{R}^{n \times n}$, con elementos de $\sigma(\mathbf{A})$ que satisfacen: $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$. λ_1 : **autovalor dominante**. $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ forman una **base** en \mathbb{R}^n (linealmente independientes).

$$\mathbf{x} = \sum_{j=1}^n \beta_j \mathbf{v}_j$$

Multiplicando ambos miembros por $\mathbf{A}, \mathbf{A}^2, \dots, \mathbf{A}^k, \dots$:

$$\mathbf{A}\mathbf{x} = \sum_{j=1}^n \beta_j \mathbf{A}\mathbf{v}_j = \sum_{j=1}^n \beta_j \lambda_j \mathbf{v}_j$$

$$\mathbf{A}^2\mathbf{x} = \sum_{j=1}^n \beta_j \lambda_j \mathbf{A}\mathbf{v}_j = \sum_{j=1}^n \beta_j \lambda_j^2 \mathbf{v}_j$$

\vdots

$$\mathbf{A}^k\mathbf{x} = \sum_{j=1}^n \beta_j \lambda_j^k \mathbf{v}_j$$

Factorizando λ_1 en la última ecuación:

$$\mathbf{A}^k\mathbf{x} = \lambda_1^k \sum_{j=1}^n \beta_j \left(\frac{\lambda_j}{\lambda_1}\right)^k \mathbf{v}_j$$

Dado que $\forall j, |\lambda_1| > |\lambda_j|$; $\lim_{k \rightarrow \infty} (\lambda_j/\lambda_1)^k = 0$, y

$$\lim_{k \rightarrow \infty} \mathbf{A}^k\mathbf{x} = \lim_{k \rightarrow \infty} \lambda_1^k \beta_1 \mathbf{v}_1 \quad (2)$$

Si $|\lambda_1| < 1$, (2) $\mapsto \mathbf{0}$, si $|\lambda_1| > 1$, (2) diverge ($\beta_1 \neq 0$). Elegimos $\mathbf{x} = \mathbf{x}^{(0)}$ con $\|\cdot\|_\infty$: $x_{p_0}^{(0)}$ con

$$x_{p_0}^{(0)} = 1 = \|\mathbf{x}^{(0)}\|_\infty$$

Hacemos $\mathbf{y}_{(1)} = \mathbf{A}\mathbf{x}^{(0)}$ y definimos $\mu^{(1)} = y_{p_0}^{(1)}$:

$$\begin{aligned} \mu^{(1)} &= y_{p_0}^{(1)} = \frac{y_{p_1}^{(1)}}{x_{p_0}^{(0)}} = \frac{\beta_1 \lambda_1 v_{p_0}^{(1)} + \sum_{j=2}^n \beta_j \lambda_j v_{p_0}^{(j)}}{\beta_1 v_{p_0}^{(1)} + \sum_{j=2}^n \beta_j v_{p_0}^{(j)}} \\ &= \lambda_1 \left[\frac{\beta_1 v_{p_0}^{(1)} + \sum_{j=2}^n \beta_j (\lambda_j/\lambda_1) v_{p_0}^{(j)}}{\beta_1 v_{p_0}^{(1)} + \sum_{j=2}^n \beta_j v_{p_0}^{(j)}} \right] \end{aligned}$$

Sea p_1 el menor entero tal que $|y_{p_1}^{(1)}| = \|\mathbf{y}^{(1)}\|_\infty$:

$$\mathbf{x}^{(1)} = \frac{\mathbf{y}^{(1)}}{y_{p_1}^{(1)}} = \frac{1}{y_{p_1}^{(1)}} \mathbf{A} \mathbf{x}^{(0)}$$

Entonces: $x_{p_1}^{(1)} = 1 = \|\mathbf{x}^{(1)}\|_\infty$. Ahora

$$\mathbf{y}^{(2)} = \mathbf{A} \mathbf{x}^{(1)} = \frac{1}{y_{p_1}^{(1)}} \mathbf{A}^2 \mathbf{x}^{(0)}$$

y

$$\begin{aligned} \mu^{(2)} = y_{p_1}^{(2)} &= \frac{y_{p_1}^{(2)}}{x_{p_1}^{(1)}} = \frac{\left[\beta_1 \lambda_1^2 v_{p_1}^{(1)} + \sum_{j=2}^n \beta_j \lambda_j^2 v_{p_1}^{(j)} \right] / y_{p_1}^{(1)}}{\left[\beta_1 \lambda_1 v_{p_1}^{(1)} + \sum_{j=2}^n \beta_j \lambda_j v_{p_1}^{(j)} \right] / y_{p_1}^{(1)}} \\ &= \lambda_1 \left[\frac{\beta_1 v_{p_1}^{(1)} + \sum_{j=2}^n \beta_j (\lambda_j / \lambda_1)^2 v_{p_1}^{(j)}}{\beta_1 v_{p_1}^{(1)} + \sum_{j=2}^n \beta_j (\lambda_j / \lambda_1) v_{p_1}^{(j)}} \right] \end{aligned}$$

Sea p_2 el menor entero tal que $|y_{p_2}^{(2)}| = \|\mathbf{y}^{(2)}\|_\infty$:

$$\mathbf{x}^{(2)} = \frac{\mathbf{y}^{(2)}}{y_{p_2}^{(2)}} = \frac{1}{y_{p_2}^{(2)}} \mathbf{A} \mathbf{x}^{(1)} = \frac{1}{y_{p_2}^{(2)} y_{p_1}^{(1)}} \mathbf{A}^2 \mathbf{x}^{(0)}$$

\mapsto secuencias $\{\mathbf{x}^{(m)}\}_{m=0}^\infty$, $\{\mathbf{y}^{(m)}\}_{m=0}^\infty$, $\{\mu^{(m)}\}_{m=0}^\infty$, inductivamente:

$$\mathbf{y}^{(m)} = \mathbf{A} \mathbf{x}^{(m-1)}$$

$$\begin{aligned} \mu^{(m)} &= y_{p_{m-1}}^{(m)} \\ &= \lambda_1 \left[\frac{\beta_1 v_{p_{m-1}}^{(1)} + \sum_{j=2}^n (\lambda_j / \lambda_1)^m \beta_j v_{p_{m-1}}^{(j)}}{\beta_1 v_{p_{m-1}}^{(1)} + \sum_{j=2}^n (\lambda_j / \lambda_1)^{m-1} \beta_j v_{p_{m-1}}^{(j)}} \right] \end{aligned}$$

y

$$\mathbf{x}^{(m)} = \frac{\mathbf{y}^{(m)}}{y_{p_m}^{(m)}} = \frac{\mathbf{A}^m \mathbf{x}^{(0)}}{\prod_{k=1}^m y_{p_k}^{(k)}}$$

donde para cada paso, p_m es el menor entero para el cual $|y_{p_m}^{(m)}| = \|\mathbf{y}^{(m)}\|_\infty$.

Dado que $|\lambda_j / \lambda_1| < 1, j = 2, \dots, n$,

$\lim_{m \rightarrow \infty} \mu^{(m)} = \lambda_1$, eligiendo $\mathbf{x}^{(0)}$ tal que $\beta_1 \neq 0$.

Además, la secuencia $\{\mathbf{x}^{(m)}\}_{m=0}^\infty$ converge al autovalor asociado con λ_1 con norma l_∞ igual a 1.

Ejemplo:

$$\mathbf{A} = \begin{bmatrix} -2 & -3 \\ 6 & 7 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Con $\sigma(\mathbf{A}) = \{4, 1\}$. Tomemos $\mathbf{x}^{(0)} = [1, 1]^\top$:

$$\mathbf{x}^{(1)} = \mathbf{A}\mathbf{x}^{(0)} = \begin{bmatrix} -5 \\ 13 \end{bmatrix}, \mathbf{x}^{(2)} = \mathbf{A}\mathbf{x}^{(1)} = \begin{bmatrix} -29 \\ 61 \end{bmatrix}$$

$$\mathbf{x}^{(3)} = \mathbf{A}\mathbf{x}^{(2)} = \begin{bmatrix} -125 \\ 253 \end{bmatrix}, \mathbf{x}^{(4)} = \mathbf{A}\mathbf{x}^{(3)} = \begin{bmatrix} -509 \\ 1021 \end{bmatrix}$$

$$\mathbf{x}^{(5)} = \mathbf{A}\mathbf{x}^{(4)} = \begin{bmatrix} -2045 \\ 4093 \end{bmatrix}, \mathbf{x}^{(6)} = \mathbf{A}\mathbf{x}^{(5)} = \begin{bmatrix} -8189 \\ 16381 \end{bmatrix}$$

Aproximaciones al autovalor dominante λ_1 :

$$\lambda_1^{(1)} = \frac{61}{13} = 4.6923, \quad \lambda_1^{(2)} = \frac{253}{61} = 4.14654$$

$$\lambda_1^{(3)} = \frac{1021}{253} = 4.03557, \quad \lambda_1^{(4)} = \frac{4093}{1021} = 4.00881$$

$$\lambda_1^{(5)} = \frac{16381}{4093} = 4.00200$$

Un autovector aproximado para

$\lambda_1^{(5)} = 16381/4093 = 4.00200$ es

$$\mathbf{x}^{(6)} = \begin{bmatrix} -8189 \\ 16381 \end{bmatrix} \rightarrow \begin{bmatrix} -0.49908 \\ 1 \end{bmatrix} \approx \mathbf{v}_1$$

Ejemplo:

$$\mathbf{A} = \begin{bmatrix} -2 & -3 \\ 6 & 7 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Con $\sigma(\mathbf{A}) = \{4, 1\}$. Tomemos $\mathbf{x}^{(0)} = [1, 1]^\top$:

$$\mathbf{x}^{(1)} = \mathbf{A}\mathbf{x}^{(0)} = \begin{bmatrix} -5 \\ 13 \end{bmatrix}, \mathbf{x}^{(2)} = \mathbf{A}\mathbf{x}^{(1)} = \begin{bmatrix} -29 \\ 61 \end{bmatrix}$$

$$\mathbf{x}^{(3)} = \mathbf{A}\mathbf{x}^{(2)} = \begin{bmatrix} -125 \\ 253 \end{bmatrix}, \mathbf{x}^{(4)} = \mathbf{A}\mathbf{x}^{(3)} = \begin{bmatrix} -509 \\ 1021 \end{bmatrix}$$

$$\mathbf{x}^{(5)} = \mathbf{A}\mathbf{x}^{(4)} = \begin{bmatrix} -2045 \\ 4093 \end{bmatrix}, \mathbf{x}^{(6)} = \mathbf{A}\mathbf{x}^{(5)} = \begin{bmatrix} -8189 \\ 16381 \end{bmatrix}$$

Aproximaciones al autovalor dominante λ_1 :

$$\lambda_1^{(1)} = \frac{61}{13} = 4.6923, \quad \lambda_1^{(2)} = \frac{253}{61} = 4.14654$$

$$\lambda_1^{(3)} = \frac{1021}{253} = 4.03557, \quad \lambda_1^{(4)} = \frac{4093}{1021} = 4.00881$$

$$\lambda_1^{(5)} = \frac{16381}{4093} = 4.00200$$

Un autovector aproximado para

$$\lambda_1^{(5)} = 16381/4093 = 4.00200 \text{ es}$$

$$\mathbf{x}^{(6)} = \begin{bmatrix} -8189 \\ 16381 \end{bmatrix} \rightarrow \begin{bmatrix} -0.49908 \\ 1 \end{bmatrix} \approx \mathbf{v}_1$$

Desventajas:

- ▶ No se sabe al inicio si \mathbf{A} tiene un autovalor dominante.
- ▶ No se conoce cómo debe elegirse $\mathbf{x}^{(9)}$ para que tenga una contribución no nula del autovector asociado al autovalor dominante, si existe.

Método de las potencias: código

```
1  #!/usr/bin/env python3
2  import numpy as np
3
4  def iter_potencia(A, num_iteraciones):
5      n = A.shape[0]
6      # Inicializar un vector aleatorio de tamaño n
7      b = np.random.rand(n)
8      # Normalizar el vector
9      b = b / np.linalg.norm(b)
10
11     for _ in range(num_iteraciones):
12         # Multiplicar la matriz A por el vector b
13         Ab = np.dot(A, b)
14         # Calcular el autovalor dominante
15         autovalor = np.dot(b, Ab)
16         # Normalizar el vector resultante
17         b = Ab / np.linalg.norm(Ab)
18
19     # Devolver el autovalor dominante y el autovector
20     # correspondiente
21     return autovalor, b
```

```
23 # Ejemplo de uso
24 # Definir una matriz de ejemplo
25
26 A = np.array([[2, 0, 0],
27               [1, 1, 2],
28               [1, -1, 4]])
29
30 # Especificar el número de iteraciones
31 num_iteraciones = 100
32
33 # Aplicar el algoritmo de las potencias
34 autovalor, autovector = iter_potencia(A, num_iteraciones)
35
36 print(f"Autovalor dominante: {autovalor}")
37 print("Autovector correspondiente:")
```

```
$ ./potencias.py
Autovalor dominante: 3.0000000000000001
Autovector correspondiente:
[ 2.62486865e-19 -7.07106781e-01 -7.07106781e-01]
```


Teorema : .

Si \mathbf{A} es una matriz y $\lambda_1, \lambda_2, \dots, \lambda_k$ son autovalores distintos de \mathbf{A} con autovectores asociados $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ es un conjunto **linealmente independiente**.

Teorema : .

Si \mathbf{A} es una matriz y $\lambda_1, \lambda_2, \dots, \lambda_k$ son autovalores distintos de \mathbf{A} con autovectores asociados $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ es un conjunto **linealmente independiente**.

Definición : Conjunto ortogonal/ortonormal.

Un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ recibe el nombre de **ortogonal** si $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ para todo $i \neq j$. Si, además, $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1$ para toda $i = 1, 2, \dots, n$, el conjunto recibe el nombre de **ortonormal**.

Dado que $\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|_2^2, \forall \mathbf{x} \in \mathbb{R}^n$, el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ es ortonormal si y solo si:

$$\|\mathbf{v}_i\|_2 = 1 \text{ para todo } i = 1, 2, \dots, n$$

Teorema : .

Si \mathbf{A} es una matriz y $\lambda_1, \lambda_2, \dots, \lambda_k$ son autovalores distintos de \mathbf{A} con autovectores asociados $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ es un conjunto **linealmente independiente**.

Definición : Conjunto ortogonal/ortonormal.

Un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ recibe el nombre de **ortogonal** si $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ para todo $i \neq j$. Si, además, $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1$ para toda $i = 1, 2, \dots, n$, el conjunto recibe el nombre de **ortonormal**.

Dado que $\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|_2^2, \forall \mathbf{x} \in \mathbb{R}^n$, el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ es ortonormal si y solo si:

$$\|\mathbf{v}_i\|_2 = 1 \text{ para todo } i = 1, 2, \dots, n$$

Teorema : Proceso de Gram-Schmidt.

Sea $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ un conjunto de k vectores linealmente independientes en \mathbb{R}^n . Entonces, $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ definido mediante:

$$\mathbf{v}_1 = \mathbf{x}_1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{v}_1, \mathbf{x}_2 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\langle \mathbf{v}_1, \mathbf{x}_3 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 - \frac{\langle \mathbf{v}_2, \mathbf{x}_3 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2$$

$$\vdots$$

$$\mathbf{v}_k = \mathbf{x}_k - \sum_{i=1}^{k-1} \frac{\langle \mathbf{v}_i, \mathbf{x}_k \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} \mathbf{v}_i$$

es un conjunto de k vectores ortogonales en \mathbb{R}^n .

Definición : Matriz ortogonal.

Se dice que una matriz Q es **ortogonal** si sus columnas $\{q_1, q_2, \dots, q_n\}$ forman un conjunto ortonormal en \mathbb{R}^n .

Definición : Matriz ortogonal.

Se dice que una matriz Q es **ortogonal** si sus columnas $\{q_1, q_2, \dots, q_n\}$ forman un conjunto ortonormal en \mathbb{R}^n .

Propiedades:

- ▶ Q es invertible con $Q^{-1} = Q^T$
- ▶ $\forall x, y \in \mathbb{R}^n, \langle Qx, Qy \rangle = \langle x, y \rangle$
- ▶ $\forall x \in \mathbb{R}^n, \|Qx\|_2 = \|x\|_2$
- ▶ Cualquier matriz invertible Q con $Q^{-1} = Q^T$ es ortogonal.

Definición : Matriz ortogonal.

Se dice que una matriz Q es **ortogonal** si sus columnas $\{q_1, q_2, \dots, q_n\}$ forman un conjunto ortonormal en \mathbb{R}^n .

Propiedades:

- ▶ Q es invertible con $Q^{-1} = Q^T$
- ▶ $\forall x, y \in \mathbb{R}^n, \langle Qx, Qy \rangle = \langle x, y \rangle$
- ▶ $\forall x \in \mathbb{R}^n, \|Qx\|_2 = \|x\|_2$
- ▶ Cualquier matriz invertible Q con $Q^{-1} = Q^T$ es ortogonal.

Definición : Matriz similar.

Dos matrices A y B son **similares** si existe una matriz no singular S con $A = S^{-1}BS$.

Definición : Matriz ortogonal.

Se dice que una matriz Q es **ortogonal** si sus columnas $\{q_1, q_2, \dots, q_n\}$ forman un conjunto ortonormal en \mathbb{R}^n .

Propiedades:

- ▶ Q es invertible con $Q^{-1} = Q^T$
- ▶ $\forall x, y \in \mathbb{R}^n, \langle Qx, Qy \rangle = \langle x, y \rangle$
- ▶ $\forall x \in \mathbb{R}^n, \|Qx\|_2 = \|x\|_2$
- ▶ Cualquier matriz invertible Q con $Q^{-1} = Q^T$ es ortogonal.

Definición : Matriz similar.

Dos matrices A y B son **similares** si existe una matriz no singular S con $A = S^{-1}BS$.

Teorema : .

Si A y B son matrices similares con $A = S^{-1}BS$, y λ es un autovalor de A con el autovector v asociado, entonces λ es un autovalor de B con autovector asociado Sv .

Definición : Matriz ortogonal.

Se dice que una matriz Q es **ortogonal** si sus columnas $\{q_1, q_2, \dots, q_n\}$ forman un conjunto ortonormal en \mathbb{R}^n .

Propiedades:

- ▶ Q es invertible con $Q^{-1} = Q^T$
- ▶ $\forall x, y \in \mathbb{R}^n, \langle Qx, Qy \rangle = \langle x, y \rangle$
- ▶ $\forall x \in \mathbb{R}^n, \|Qx\|_2 = \|x\|_2$
- ▶ Cualquier matriz invertible Q con $Q^{-1} = Q^T$ es ortogonal.

Definición : Matriz similar.

Dos matrices A y B son **similares** si existe una matriz no singular S con $A = S^{-1}BS$.

Teorema : .

Si A y B son matrices similares con $A = S^{-1}BS$, y λ es un autovalor de A con el autovector v asociado, entonces λ es un autovalor de B con autovector asociado Sv .

Teorema : .

Una matriz $A \in \mathbb{R}^{n \times n}$ es similar a una matriz diagonal D si y sólo si A tiene n autovectores linealmente independientes. En este caso $D = S^{-1}AS$, donde las columnas de S son los autovectores y el i -ésimo elemento diagonal de D es el autovalor que corresponde a la i -ésima columna de S .

Teorema : Teorema de Schur.

Sea \mathbf{A} una matriz arbitraria. Existe una matriz no singular \mathbf{U} con la propiedad de que

$$\mathbf{T} = \mathbf{U}^{-1} \mathbf{A} \mathbf{U}$$

donde \mathbf{T} es una matriz triangular superior, cuyas entradas diagonales consisten en autovalores de \mathbf{A} .

Se cumple $\|\mathbf{U}\mathbf{x}\|_2 = \|\mathbf{x}\|_2, \forall \mathbf{x} \mapsto$ **matrices unitarias**.

Teorema : Teorema de Schur.

Sea \mathbf{A} una matriz arbitraria. Existe una matriz no singular \mathbf{U} con la propiedad de que

$$\mathbf{T} = \mathbf{U}^{-1} \mathbf{A} \mathbf{U}$$

donde \mathbf{T} es una matriz triangular superior, cuyas entradas diagonales consisten en autovalores de \mathbf{A} .

Se cumple $\|\mathbf{U}\mathbf{x}\|_2 = \|\mathbf{x}\|_2, \forall \mathbf{x} \mapsto$ **matrices unitarias**.

Factorización QR: $\mathbf{A} = \mathbf{Q}\mathbf{R}$, donde:

- ▶ \mathbf{Q} es una matriz ortogonal
- ▶ \mathbf{R} es una matriz triangular superior

Teorema : Teorema de Schur.

Sea \mathbf{A} una matriz arbitraria. Existe una matriz no singular \mathbf{U} con la propiedad de que

$$\mathbf{T} = \mathbf{U}^{-1} \mathbf{A} \mathbf{U}$$

donde \mathbf{T} es una matriz triangular superior, cuyas entradas diagonales consisten en autovalores de \mathbf{A} .

Se cumple $\|\mathbf{U}\mathbf{x}\|_2 = \|\mathbf{x}\|_2, \forall \mathbf{x} \mapsto$ **matrices unitarias**.

Factorización QR: $\mathbf{A} = \mathbf{Q}\mathbf{R}$, donde:

- ▶ \mathbf{Q} es una matriz ortogonal
- ▶ \mathbf{R} es una matriz triangular superior

Cálculo de la factorización:

- ▶ Ortogonalización de Gram-Schmidt
- ▶ Reflexiones de Householder

Ortogonalización de Gram-Schmidt:

$$\mathbf{A} = [\mathbf{a}_1 | \mathbf{a}_2 | \cdots | \mathbf{a}_n]$$

$$\mathbf{u}_1 = \mathbf{a}_1, \quad \mathbf{e}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}$$

$$\mathbf{u}_2 = \mathbf{a}_2 - \langle \mathbf{e}_1, \mathbf{a}_2 \rangle, \quad \mathbf{e}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}$$

$$\mathbf{u}_3 = \mathbf{a}_3 - \langle \mathbf{e}_1, \mathbf{a}_3 \rangle - \langle \mathbf{e}_2, \mathbf{a}_3 \rangle, \quad \mathbf{e}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}$$

\vdots

$$\mathbf{u}_k = \mathbf{a}_k - \sum_{j=1}^{k-1} \langle \mathbf{e}_j, \mathbf{a}_k \rangle, \quad \mathbf{e}_k = \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}$$

Teorema : Teorema de Schur.

Sea \mathbf{A} una matriz arbitraria. Existe una matriz no singular \mathbf{U} con la propiedad de que

$$\mathbf{T} = \mathbf{U}^{-1} \mathbf{A} \mathbf{U}$$

donde \mathbf{T} es una matriz triangular superior, cuyas entradas diagonales consisten en autovalores de \mathbf{A} .

Se cumple $\|\mathbf{U}\mathbf{x}\|_2 = \|\mathbf{x}\|_2, \forall \mathbf{x} \mapsto$ **matrices unitarias**.

Factorización QR: $\mathbf{A} = \mathbf{Q}\mathbf{R}$, donde:

- ▶ \mathbf{Q} es una matriz ortogonal
- ▶ \mathbf{R} es una matriz triangular superior

Cálculo de la factorización:

- ▶ Ortogonalización de Gram-Schmidt
- ▶ Reflexiones de Householder

Ortogonalización de Gram-Schmidt:

$$\mathbf{A} = [\mathbf{a}_1 | \mathbf{a}_2 | \cdots | \mathbf{a}_n]$$

$$\mathbf{u}_1 = \mathbf{a}_1, \quad \mathbf{e}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}$$

$$\mathbf{u}_2 = \mathbf{a}_2 - \langle \mathbf{e}_1, \mathbf{a}_2 \rangle \mathbf{e}_1, \quad \mathbf{e}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}$$

$$\mathbf{u}_3 = \mathbf{a}_3 - \langle \mathbf{e}_1, \mathbf{a}_3 \rangle \mathbf{e}_1 - \langle \mathbf{e}_2, \mathbf{a}_3 \rangle \mathbf{e}_2, \quad \mathbf{e}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}$$

\vdots

$$\mathbf{u}_k = \mathbf{a}_k - \sum_{j=1}^{k-1} \langle \mathbf{e}_j, \mathbf{a}_k \rangle \mathbf{e}_j, \quad \mathbf{e}_k = \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}$$

Ahora podemos expresar los \mathbf{a}_i en la nueva base:

$$\mathbf{a}_1 = \langle \mathbf{e}_1, \mathbf{a}_1 \rangle \mathbf{e}_1$$

$$\mathbf{a}_1 = \langle \mathbf{e}_1, \mathbf{a}_2 \rangle \mathbf{e}_1 + \langle \mathbf{e}_2, \mathbf{a}_2 \rangle \mathbf{e}_2$$

$$\mathbf{a}_3 = \langle \mathbf{e}_1, \mathbf{a}_3 \rangle \mathbf{e}_1 + \langle \mathbf{e}_2, \mathbf{a}_3 \rangle \mathbf{e}_2 + \langle \mathbf{e}_3, \mathbf{a}_3 \rangle \mathbf{e}_3$$

\dots

$$\mathbf{a}_k = \sum_{j=1}^k \langle \mathbf{e}_j, \mathbf{a}_k \rangle \mathbf{e}_j$$

Resulta $\mathbf{A} = \mathbf{Q}\mathbf{R}$, con $\mathbf{Q} = [e_1|e_2|\cdots|e_n]$, y

$$\mathbf{R} = \begin{bmatrix} \langle e_1 a_1 \rangle & \langle e_1 a_2 \rangle & \langle e_1 a_3 \rangle & \cdots & \langle e_1 a_n \rangle \\ 0 & \langle e_2 a_2 \rangle & \langle e_2 a_3 \rangle & \cdots & \langle e_2 a_n \rangle \\ 0 & 0 & \langle e_3 a_3 \rangle & \cdots & \langle e_3 a_n \rangle \\ \vdots & \vdots & \vdots & \ddots & \cdots \\ 0 & 0 & 0 & \cdots & \langle e_n, a_n \rangle \end{bmatrix}$$

Resulta $A = QR$, con $Q = [e_1|e_2|\dots|e_n]$, y

$$R = \begin{bmatrix} \langle e_1 a_1 \rangle & \langle e_1 a_2 \rangle & \langle e_1 a_3 \rangle & \cdots & \langle e_1 a_n \rangle \\ 0 & \langle e_2 a_2 \rangle & \langle e_2 a_3 \rangle & \cdots & \langle e_2 a_n \rangle \\ 0 & 0 & \langle e_3 a_3 \rangle & \cdots & \langle e_3 a_n \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \langle e_n a_n \rangle \end{bmatrix}$$

Código Python:

```
1 #!/usr/bin/env python3
2
3 import numpy as np
4
5 def gram_schmidt_qr(A):
6     m, n = A.shape
7     Q = np.zeros((m, n))
8     R = np.zeros((n, n))
```

```
10     for j in range(n):
11         v = A[:, j]
12         for i in range(j):
13             R[i, j] = np.dot(Q[:, i], A[:, j])
14             v = v - R[i, j] * Q[:, i]
15         R[j, j] = np.linalg.norm(v)
16         Q[:, j] = v / R[j, j]
17
18     return Q, R
19
20 # Ejemplo de uso
21 # Definir una matriz de ejemplo
22 A = np.array([[1, 4, 3],
23               [2, 5, 1],
24               [3, 6, 2]])
25
26 # Aplicar la factorización QR usando el método
27 # de Gram-Schmidt
28 Q, R = gram_schmidt_qr(A)
29
30 print("Matriz Q:")
31 print(Q)
32 print("Matriz R:")
33 print(R)
```

Método QR para el cálculo de autovalores:

Algoritmo recursivo que computa $\{\mathbf{A}_k\}_{k=0}^{\infty}$ con los siguientes pasos:

1. $\mathbf{A}_0 = \mathbf{A}$
2. Para $k = 0, 1, 2, \dots$, dado \mathbf{A}_k :
 - 2.1 Calcular $\mathbf{Q}_{k+1}\mathbf{R}_{k+1} = \mathbf{A}_k$
 - 2.2 Definir $\mathbf{A}_{k+1} = \mathbf{Q}_{k+1}\mathbf{R}_{k+1}$

Método QR para el cálculo de autovalores:

Algoritmo recursivo que computa $\{A_k\}_{k=0}^{\infty}$ con los siguientes pasos:

1. $A_0 = A$
2. Para $k = 0, 1, 2, \dots$, dado A_k :
 - 2.1 Calcular $Q_{k+1}R_{k+1} = A_k$
 - 2.2 Definir $A_{k+1} = Q_{k+1}R_{k+1}$

Teorema : Convergencia.

Si los autovalores de una matriz A verifican que

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > 0$$

entonces la sucesión de matrices equivalentes contruidas con el algoritmo QR converge a una matriz triangular superior.

Código Python:

```
1 #!/usr/bin/env python3
2 import numpy as np
3
4 def algoritmo_qr(A, num_iter):
5     n = A.shape[0]
6     autovalores = np.zeros(n, dtype=np.complex128)
7     for _ in range(num_iter):
8         Q, R = np.linalg.qr(A)
9         A = np.dot(R, Q)
10    for i in range(n):
11        autovalores[i] = A[i, i]
12    return autovalores
13
14 A = np.array([[1, 2, 3],
15               [4, 5, 6],
16               [7, 8, 9]])
17
18 num_iter = 100
19 autovalores = algoritmo_qr(A, num_iter)
20 print("Autovalores:")
21 print(autovalores)
```

- ▶ R.L. Burden, D.J. Faires y A.M. Burden. ***Análisis numérico***. 10.^a ed. Mexico: Cengage Learning, 2017. Capítulo 9.
- ▶ Carlos Moreno González. ***Introducción al cálculo numérico***. Madrid, España: Universidad Nacional de Educación a Distancia, 2014. Capítulo 3.
- ▶ B. Bradie. ***A Friendly Introduction to Numerical Analysis***. New Jersey, United States: Pearson Education Inc., 2006. Capítulo 4.
- ▶ A.J. Salgado y S.M. Wise. ***Classical Numerical Analysis***. Cambridge, United Kingdom: Cambridge University Press, 2023. doi: [10.1017/9781108942607](https://doi.org/10.1017/9781108942607). Capítulo 8.
- ▶ A. Quarteroni, R. Sacco y F. Saleri. ***Numerical Mathematics***. New York, United States: Springer-Verlag, 2000. Capítulo 5.