

# AUTOVALORES Y AUTOVECTORES

DEFINICIONES. INTERPRETACIÓN GEOMÉTRICA. CÍRCULOS DE GERSCHGORIN. MÉTODO DE LAS POTENCIAS.

---

Manuel Carlevaro

Departamento de Ingeniería Mecánica

Grupo de Materiales Granulares - UTN FRLP

manuel.carlevaro@gmail.com

Cálculo Avanzado • 2024

 · X<sub>y</sub>LaTeX · 

### 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  $\lambda$  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  $\lambda$  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  $\lambda$  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  $\lambda$  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):

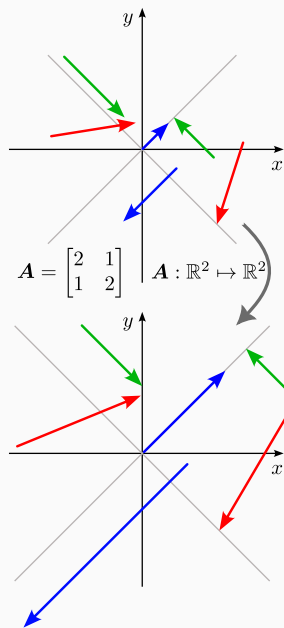
$$\mathbf{v}_1 = a \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{v}_2 = b \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \mathbf{v}_3 = c \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

**Definiciones:**

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

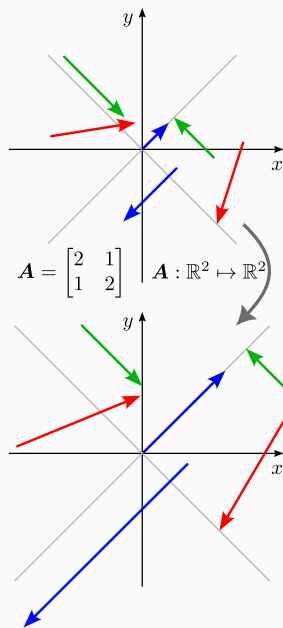
Radio espectral de  $\mathbf{A}$  :  $\rho(\mathbf{A}) = \max_{\lambda \in \sigma(\mathbf{A})} |\lambda|$

## 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

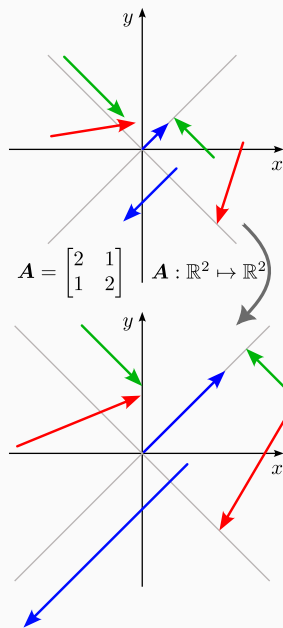


$$\begin{aligned} p_A(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{vmatrix} \\ &= (2 - \lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1) = 0 \end{aligned}$$

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  $\lambda$  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

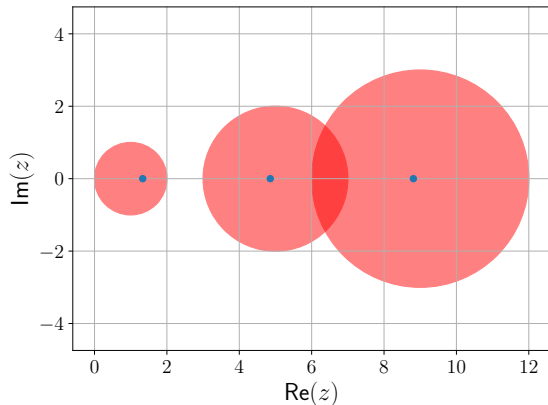
$$C_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}$$

1. Si  $\lambda$  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:  $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$ .  $\lambda_1$ : **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  $\lambda_1$  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  $\lambda_1$  con norma  $l_\infty$  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  $\lambda_1$ :

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

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

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

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  $\lambda_1$ :

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

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

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

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.

- ▶ **burden2017**. Capítulo 9.
- ▶ **moreno2014**. Capítulo 3.
- ▶ **bradie2006**. Capítulo 4.
- ▶ **salgado2023**. Capítulo 8.
- ▶ **quarteroni2000**. Capítulo 5.