Normas de vectores y matrices

Manuel Carlevaro

Departamento de Ingeniería Mecánica

Grupo de Materiales Granulares - UTN FRLP

manuel.carlevaro@gmail.com

Cálculo Avanzado • 2023

Motivación:

Error en escalar:

$$|x^* - x| \sim \delta x$$

¿Cómo es el error en $\mathbb{R}^n$ o en $\mathbb{R}^{m\times n}$?

$$\boldsymbol{x} \mapsto \delta\boldsymbol{x}$$

o

$$(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{b} + \delta\boldsymbol{b}$$

Motivación:

Error en escalar:

$$|x^* - x| \sim \delta x$$

¿Cómo es el error en $\mathbb{R}^n$ o en $\mathbb{R}^{m\times n}$?

$$\boldsymbol{x} \mapsto \delta \boldsymbol{x}$$

o

$$(\boldsymbol{A} + \delta \boldsymbol{A})(\boldsymbol{x} + \delta \boldsymbol{x}) = \boldsymbol{b} + \delta \boldsymbol{b}$$

Definición : Espacio vectorial.

Un *espacio vectorial* sobre el cuerpo K ($K = \mathbb{R}$ o $K = \mathbb{C}$), es un conjunto no vacío V, cuyos elementos se denominan **vectores**, y en el que se definen dos operaciones: **suma** y **multiplicación escalar**.

Motivación:

Error en escalar:

$$|x^* - x| \sim \delta x$$

¿Cómo es el error en $\mathbb{R}^n$ o en $\mathbb{R}^{m\times n}$?

$$\boldsymbol{x} \mapsto \delta\boldsymbol{x}$$

o

$$(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{b} + \delta\boldsymbol{b}$$

Definición : Espacio vectorial.

Un *espacio vectorial* sobre el cuerpo K ($K = \mathbb{R}$ o $K = \mathbb{C}$), es un conjunto no vacío V, cuyos elementos se denominan **vectores**, y en el que se definen dos operaciones: **suma** y **multiplicación escalar**.

Propiedades:

1. La suma es conmutativa y asociativa.
2. $\exists\, \mathbf{0} \in V$ (vector cero o nulo) tal que $\boldsymbol{v} + \mathbf{0} = \boldsymbol{v},\ \forall \boldsymbol{v} \in V$.
3. $0 \cdot \boldsymbol{v} = \mathbf{0}, 1 \cdot \boldsymbol{v} = \boldsymbol{v}$, donde 0 y 1 son respectivamente el cero y uno de K.
4. $\forall \boldsymbol{v} \in V, \exists -\boldsymbol{v} \in V : \boldsymbol{v} + (-\boldsymbol{v}) = \mathbf{0}$.
5. Propiedad distributiva:
   - $\forall \alpha \in K, \forall \boldsymbol{v}, \boldsymbol{w} \in V, \alpha(\boldsymbol{v} + \boldsymbol{w}) = \alpha\boldsymbol{v} + \alpha\boldsymbol{w}$.
   - $\forall \alpha, \beta \in K, \forall \boldsymbol{v} \in V, (\alpha + \beta)\boldsymbol{v} = \alpha\boldsymbol{v} + \beta\boldsymbol{v}$.
6. Propiedad asociativa:
   - $\forall \alpha, \beta \in K, \forall \boldsymbol{v} \in V, (\alpha\beta)\boldsymbol{v} = \alpha(\beta\boldsymbol{v})$

Motivación:

Error en escalar:

$$|x^* - x| \sim \delta x$$

¿Cómo es el error en $\mathbb{R}^n$ o en $\mathbb{R}^{m\times n}$?

$$\boldsymbol{x} \mapsto \delta \boldsymbol{x}$$

o

$$(\boldsymbol{A} + \delta \boldsymbol{A})(\boldsymbol{x} + \delta \boldsymbol{x}) = \boldsymbol{b} + \delta \boldsymbol{b}$$

Definición : Espacio vectorial.

Un *espacio vectorial* sobre el cuerpo K ($K = \mathbb{R}$ o $K = \mathbb{C}$), es un conjunto no vacío V, cuyos elementos se denominan **vectores**, y en el que se definen dos operaciones: **suma** y **multiplicación escalar**.

Propiedades:

1. La suma es conmutativa y asociativa.
2. $\exists\, \boldsymbol{0} \in V$ (vector cero o nulo) tal que $\boldsymbol{v} + \boldsymbol{0} = \boldsymbol{v},\ \forall \boldsymbol{v} \in V$.
3. $0 \cdot \boldsymbol{v} = \boldsymbol{0}, 1 \cdot \boldsymbol{v} = \boldsymbol{v}$, donde 0 y 1 son respectivamente el cero y uno de K.
4. $\forall \boldsymbol{v} \in V, \exists - \boldsymbol{v} \in V : \boldsymbol{v} + (-\boldsymbol{v}) = \boldsymbol{0}$.
5. Propiedad distributiva:
   - $\forall \alpha \in K, \forall \boldsymbol{v}, \boldsymbol{w} \in V, \alpha(\boldsymbol{v} + \boldsymbol{w}) = \alpha \boldsymbol{v} + \alpha \boldsymbol{w}$.
   - $\forall \alpha, \beta \in K, \forall \boldsymbol{v} \in V, (\alpha + \beta)\boldsymbol{v} = \alpha \boldsymbol{v} + \beta \boldsymbol{v}$.
6. Propiedad asociativa:
   - $\forall \alpha, \beta \in K, \forall \boldsymbol{v} \in V, (\alpha\beta)\boldsymbol{v} = \alpha(\beta \boldsymbol{v})$

Ejemplos:

- $V = \mathbb{R}^n$ ($V = \mathbb{C}^n$), $n > 1$.
- $V = \mathbb{P}_n = \{p_n(x) = \sum_{k=0}^{n} a_k x^k\},\ n \geq 0$.
- $V = \mathcal{C}^p([a, b]),\ 0 \leq p \leq \infty$.

Definición : Producto escalar.

Un **producto escalar** en V definido sobre K es cualquier mapeo $\langle\cdot,\cdot\rangle$, $V \times V \mapsto K$ con las siguientes propiedades:

1. Es lineal respecto de los vectores en V:

$$\langle \gamma \boldsymbol{x} + \lambda \boldsymbol{z}, \boldsymbol{y}\rangle = \gamma\langle \boldsymbol{x}, \boldsymbol{y}\rangle + \lambda\langle \boldsymbol{z}, \boldsymbol{y}\rangle$$

   $\forall \boldsymbol{x}, \boldsymbol{z} \in V,\ \forall \gamma, \lambda \in K.$

2. Es hermítico:

$$\langle \boldsymbol{y}, \boldsymbol{x}\rangle = \overline{\langle \boldsymbol{x}, \boldsymbol{y}\rangle}, \forall \boldsymbol{x}, \boldsymbol{y} \in V$$

3. Es definido positivo:

$$\langle \boldsymbol{x}, \boldsymbol{x}\rangle \geq 0, \quad \langle \boldsymbol{x}, \boldsymbol{x}\rangle = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$$

Definición : Producto escalar.

Un **producto escalar** en V definido sobre K es cualquier mapeo $\langle\cdot,\cdot\rangle$, $V \times V \mapsto K$ con las siguientes propiedades:

1. Es lineal respecto de los vectores en V:

$$\langle \gamma \boldsymbol{x} + \lambda \boldsymbol{z}, \boldsymbol{y}\rangle = \gamma\langle \boldsymbol{x}, \boldsymbol{y}\rangle + \lambda\langle \boldsymbol{z}, \boldsymbol{y}\rangle$$

   $\forall \boldsymbol{x}, \boldsymbol{z} \in V$, $\forall \gamma, \lambda \in K$.
2. Es hermítico:

$$\langle \boldsymbol{y}, \boldsymbol{x}\rangle = \overline{\langle \boldsymbol{x}, \boldsymbol{y}\rangle}, \forall \boldsymbol{x}, \boldsymbol{y} \in V$$

3. Es definido positivo:

$$\langle \boldsymbol{x}, \boldsymbol{x}\rangle \geq 0, \quad \langle \boldsymbol{x}, \boldsymbol{x}\rangle = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$$

Definición : Norma vectorial.

Sea V un espacio vectorial sobre K. El mapeo $\|\cdot\| \mapsto \mathbb{R}$ es una **norma** si se cumple que:

1. (i) $\|\boldsymbol{v}\| \geq 0 \ \forall \boldsymbol{v} \in V$ y (ii) $\|\boldsymbol{v}\| = 0 \Leftrightarrow \boldsymbol{v} = \boldsymbol{0}$;
2. $\|\alpha \boldsymbol{v}\| = |\alpha|\|\boldsymbol{v}\| \ \forall \alpha \in K, \forall \boldsymbol{v} \in V$ (propiedad de homogeneidad);
3. $\|\boldsymbol{v} + \boldsymbol{w}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{w}\| \ \forall \boldsymbol{v}, \boldsymbol{w} \in V$ (desigualdad triangular),

donde $|\alpha|$ denota el valor absoluto de α si $K = \mathbb{R}$, o el módulo de α si $K = \mathbb{C}$. El par $(V, \|\cdot\|)$ se denomina **espacio normado**.

Ejemplo: Norma p, o l_p, $(\mathbb{R}^n, \|\cdot\|_p)$

$$\|\boldsymbol{x}\| = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \quad 1 \leq p \leq \infty$$

- Cuando $p \to \infty$ (l_∞): **norma infinita** o **norma máxima**

$$\|\boldsymbol{x}\|_\infty = \max_{1\leq i\leq n} |x_i|$$

- Cuando $p = 1$ (l_1): **norma del taxista**

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i|$$

- Cuando $p = 2$ (l_2): **norma euclídea**

$$\|\boldsymbol{x}\|_2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle^{1/2} = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}$$

Ejemplo: Norma p, o l_p, $(\mathbb{R}^n, \|\cdot\|_p)$

$$\|\boldsymbol{x}\| = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \quad 1 \leq p \leq \infty$$

- Cuando $p \to \infty$ (l_∞): **norma infinita** o **norma máxima**

$$\|\boldsymbol{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

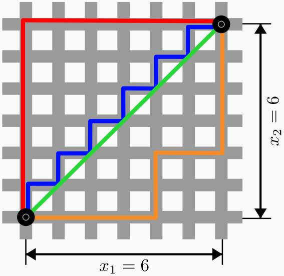
- Cuando $p = 1$ (l_1): **norma del taxista**

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i|$$

- Cuando $p = 2$ (l_2): **norma euclídea**

$$\|\boldsymbol{x}\|_2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle^{1/2} = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}$$

Interpretación gráfica en $\mathbb{R}^2$:



$$\boldsymbol{x} = [6, 6]$$

$$l_\infty = \max\{|6|, |6|\} = 6$$

Camino más corto:

- l_1 (rojo), (azul), (naranja): 12
- l_2 (verde) : $6\sqrt{2} \approx 8.49$

Ejemplo: Norma p, o l_p, $(\mathbb{R}^n, \|\cdot\|_p)$

$$\|\boldsymbol{x}\| = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \quad 1 \leq p \leq \infty$$

- Cuando $p \to \infty$ (l_∞): **norma infinita** o **norma máxima**

$$\|\boldsymbol{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

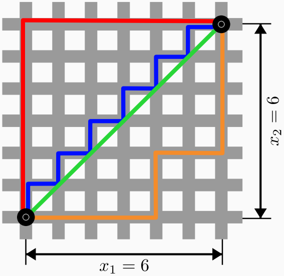
- Cuando $p = 1$ (l_1): **norma del taxista**

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i|$$

- Cuando $p = 2$ (l_2): **norma euclídea**

$$\|\boldsymbol{x}\|_2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle^{1/2} = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}$$

Interpretación gráfica en $\mathbb{R}^2$:



$$\boldsymbol{x} = [6, 6]$$

$$l_\infty = \text{máx}\{|6|, |6|\} = 6$$

Camino más corto:

- l_1 , , : 12
- l_2 : $6\sqrt{2} \approx 8.49$

Desigualdad de Cauchy-Schwarz: $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$:

$$|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2$$

donde la igualdad vale si y solo si $\boldsymbol{y} = \alpha \boldsymbol{x}, \alpha \in \mathbb{R}$.

Continuidad: Cualquier $\|\cdot\|$ en V es una **función continua** de sus argumentos: $\forall \varepsilon > 0, \exists C > 0$ tal que si $\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \leq \varepsilon$, entonces $|\|\boldsymbol{x}\| - \|\hat{\boldsymbol{x}}\|| \leq C\varepsilon$ para cualquier $\boldsymbol{x}, \hat{\boldsymbol{x}} \in V$.

Cálculo de normas:

$$\boldsymbol{x}_1 = [1, -2, 3]$$
$$\boldsymbol{x}_2 = [2, 0, -1, 2]$$
$$\boldsymbol{x}_3 = [0, 1, -4, 2, -1]$$

Norma l_1:

$$\|\boldsymbol{x}_1\|_1 = |1| + |-2| + |3| = 6$$
$$\|\boldsymbol{x}_2\|_1 = |2| + |0| + |-1| + |2| = 5$$
$$\|\boldsymbol{x}_3\|_1 = |0| + |1| + |-4| + |2| + |-1| = 8$$

Norma máxima:

$$\|\boldsymbol{x}_1\|_\infty = \text{máx}\{|1|, |-2|, |3|\} = 3$$
$$\|\boldsymbol{x}_2\|_\infty = \text{máx}\{|2|, |0|, |-1|, |2|\} = 2$$
$$\|\boldsymbol{x}_3\|_\infty = \text{máx}\{|0|, |1|, |-4|, |2|, |-1|\} = 4$$

Norma l_2:

$$\|\boldsymbol{x}_1\|_2 = \sqrt{|1|^2 + |-2|^2 + |3|^2} = \sqrt{14} \approx 3.74$$
$$\|\boldsymbol{x}_2\|_2 = \sqrt{|2|^2 + |0|^2 + |-1|^2 + |2|^2} = \sqrt{9} = 3$$
$$\|\boldsymbol{x}_3\|_2 = \sqrt{|0|^2 + |1|^2 + |-4|^2 + |2|^2 + |-1|^2} = \sqrt{22} \approx 4.69$$

Definición : Norma matricial.

Una **norma matricial** es un mapeo $\|\cdot\| : K^{m\times n} \to \mathbb{R}$ tal que:

1. $\|\boldsymbol{A}\| \geq 0 \ \forall \boldsymbol{A} \in K^{m\times n}$ y $\|\boldsymbol{A}\| = 0 \Leftrightarrow \boldsymbol{A} = \boldsymbol{0}$;
2. $\|\alpha \boldsymbol{A}\| = |\alpha| \|\boldsymbol{A}\|, \ \forall \alpha \in K, \forall \boldsymbol{A} \in K^{m\times n}$ (homogeneidad);
3. $\|\boldsymbol{A} + \boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|, \ \forall \boldsymbol{A}, \boldsymbol{B} \in K^{m\times n}$ (desigualdad triangular).

La **distancia entre matrices** $m \times n$ con esta norma es $\|\boldsymbol{A} - \boldsymbol{B}\|$.

Definición : Norma matricial.

Una **norma matricial** es un mapeo $\|\cdot\| : K^{m\times n} \to \mathbb{R}$ tal que:

1. $\|\boldsymbol{A}\| \geq 0\ \forall \boldsymbol{A} \in K^{m\times n}$ y $\|\boldsymbol{A}\| = 0 \Leftrightarrow \boldsymbol{A} = \boldsymbol{0}$;
2. $\|\alpha \boldsymbol{A}\| = |\alpha|\|\boldsymbol{A}\|,\ \forall \alpha \in K, \forall \boldsymbol{A} \in K^{m\times n}$ (homogeneidad);
3. $\|\boldsymbol{A} + \boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|,\ \forall \boldsymbol{A}, \boldsymbol{B} \in K^{m\times n}$ (desigualdad triangular).

La **distancia entre matrices** $m \times n$ con esta norma es $\|\boldsymbol{A} - \boldsymbol{B}\|$.

Norma por componentes:

Matriz $m \times n \mapsto$ vector $m \cdot n$. Sea $\boldsymbol{A} = [a_{ij}] \in \mathbb{C}^{m\times n}$:

- **Norma p:**

$$\|\boldsymbol{A}\|_p = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^p \right)^{1/p}$$

- **Norma de Frobenius:** $(p = 2)$

$$\|\boldsymbol{A}\|_F^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2$$

- **Norma máxima:**

$$\|\boldsymbol{A}\|_{\text{máx}} = \max_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} |a_{ij}|$$

Teorema : Norma inducida.

Si $\|\cdot\|$ es una norma vectorial en K^n, entonces

$$\|\boldsymbol{A}\| = \max_{\|\boldsymbol{x}\|=1} \|\boldsymbol{A}\boldsymbol{x}\|$$

es la norma matricial ***natural*** *o* ***inducida*** *asociada con la norma vectorial.*

Alternativamente:

$$\|\boldsymbol{A}\| = \max_{\boldsymbol{z}\neq\boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{z}\|}{\|\boldsymbol{z}\|}$$

Corolario: $\forall \boldsymbol{z} \neq \boldsymbol{0}$, $\boldsymbol{A}$ y cualquier norma natural $\|\cdot\|$:

$$\|\boldsymbol{A}\boldsymbol{z}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{z}\|$$

Teorema : Norma inducida.

Si $\|\cdot\|$ es una norma vectorial en K^n, entonces

$$\|\boldsymbol{A}\| = \max_{\|\boldsymbol{x}\|=1} \|\boldsymbol{A}\boldsymbol{x}\|$$

es la norma matricial ***natural*** *o* ***inducida*** *asociada con la norma vectorial.*

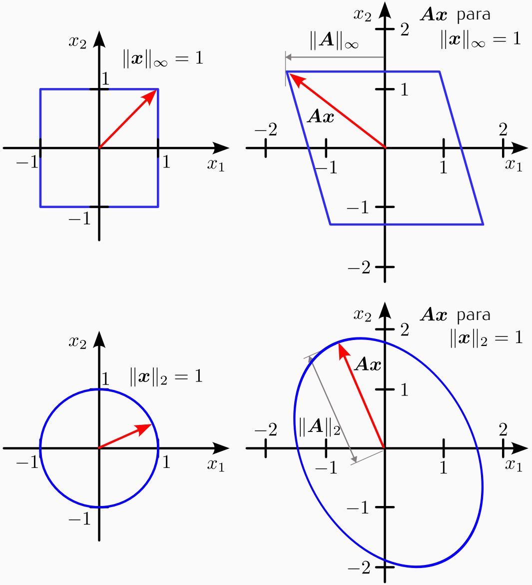
Alternativamente:

$$\|\boldsymbol{A}\| = \max_{\boldsymbol{z}\neq\boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{z}\|}{\|\boldsymbol{z}\|}$$

Corolario: $\forall \boldsymbol{z} \neq \boldsymbol{0}$, $\boldsymbol{A}$ y cualquier norma natural $\|\cdot\|$:

$$\|\boldsymbol{A}\boldsymbol{z}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{z}\|$$

Interpretación gráfica en $\mathbb{R}^2$:



Norma inducida l_p: si $\boldsymbol{A}$ es una matriz en $K^{m\times n}$

$$\|\boldsymbol{A}\|_p = \max_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p}$$

- **Norma inducida l_1:** (norma suma columna)

$$\|\boldsymbol{A}\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^{m} |a_{ij}|$$

- **Norma inducida ∞:** (norma suma fila)

$$\|\boldsymbol{A}\|_\infty = \max_{i=1,\dots,m} \sum_{j=1}^{n} |a_{ij}|$$

- **Norma inducida l_2:** $\mapsto$ próxima clase.

Norma inducida l_p: si $\boldsymbol{A}$ es una matriz en $K^{m\times n}$

$$\|\boldsymbol{A}\|_p = \max_{\boldsymbol{x}\neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p}$$

- **Norma inducida l_1**: (norma suma columna)

$$\|\boldsymbol{A}\|_1 = \max_{j=1,\ldots,n} \sum_{i=1}^{m} |a_{ij}|$$

- **Norma inducida ∞**: (norma suma fila)

$$\|\boldsymbol{A}\|_\infty = \max_{i=1,\ldots,m} \sum_{j=1}^{n} |a_{ij}|$$

- **Norma inducida l_2**: $\mapsto$ próxima clase.

Definición : Norma sub-multiplicativa.

Una norma matricial $\|\cdot\|$ es **sub-multiplicativa** si $\forall \boldsymbol{A} \in K^{n\times m}, \forall \boldsymbol{B} \in K^{m\times q}$:

$$\|\boldsymbol{A}\boldsymbol{B}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{B}\|$$

Definición : Consistencia.

Si $\|\cdot\|_{K^n} : K^n \to \mathbb{R}$, y $\|\cdot\|_{K^m} : K^m \to \mathbb{R}$ son normas, y $\|\cdot\| : K^{m\times n} \to \mathbb{R}$ es una norma matricial, decimos que $\|\cdot\|$ es **consistente** (o **compatible**) respecto de las normas $\|\cdot\|_{K^n}$ y $\|\cdot\|_{K^m}$ si y solo si

$$\|\boldsymbol{A}\boldsymbol{x}\|_{K^m} \leq \|\boldsymbol{A}\|\|\boldsymbol{x}\|_{K^n}$$

En matrices cuadradas, las normas inducidas son **sub-multiplicativas** y **consistentes**.

Determinar $\|\boldsymbol{A}\|_1$ y $\|\boldsymbol{A}\|_\infty$ para la matriz

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ 5 & -1 & 1 \end{bmatrix}$$

Norma $\|\cdot\|_1$:

$$\sum_{i=1}^{3} |a_{i,1}| = |1| + |0| + |5| = 6$$

$$\sum_{i=1}^{3} |a_{i,2}| = |2| + |3| + |-1| = 6$$

$$\sum_{i=1}^{3} |a_{i,3}| = |-1| + |-1| + |1| = 3$$

$$\|\boldsymbol{A}\|_1 = \text{máx}\{6, 6, 3\} = 6$$

Norma $\|\cdot\|_\infty$:

$$\sum_{j=1}^{3} |a_{1,j}| = |1| + |2| + |-1| = 4$$

$$\sum_{j=1}^{3} |a_{2,j}| = |0| + |3| + |-1| = 4$$

$$\sum_{j=1}^{3} |a_{3,j}| = |5| + |-1| + |1| = 7$$

$$\|\boldsymbol{A}\|_\infty = \text{máx}\{4, 4, 7\} = 7$$

Lecturas recomendadas I

- R.L. Burden, D.J. Faires y A.M. Burden. *Análisis numérico*. 10.ª ed. Mexico: Cengage Learning, 2017. Capítulo 7.
- Carlos Moreno González. *Introducción al cálculo numérico*. Madrid, España: Universidad Nacional de Educación a Distancia, 2014. Capítulo 2.
- B. Bradie. *A Friendly Introduction to Numerical Analysis*. New Jersey, United States: Pearson Education Inc., 2006. Sección 3.3.
- A. Quarteroni, R. Sacco y F. Saleri. *Numerical Mathematics*. New York, United States: Springer-Verlag, 2000. Capítulo 1.