Cálculo Numérico - Elementos de Cálculo Numérico

Normas vectoriales, matriciales y número de condición

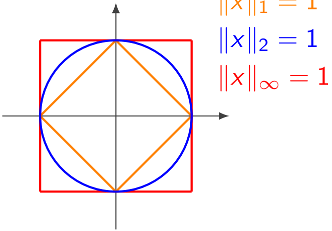
Clase dictada por: Mercedes Pérez Millán

(Dto. de Matemática–FCEyN–Universidad de Buenos Aires, IMAS-CONICET)

9 de septiembre de 2024

Normas vectoriales

- $\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$
- $\|x\|_\infty = \max\limits_{1 \leq i \leq n} |x_i|$
- $\|x\|_1 = \sum\limits_{1 \leq j \leq n} |x_j|$
- $\|x\|_p = \left(x_1^p + \cdots + x_n^p\right)^{1/p}$



Desigualdad de Cauchy-Schwarz:

$$\langle a, b\rangle \leq \|a\|_2 \|b\|_2$$

Dem:
Basta probar que $\langle a, b\rangle^2 \leq \|a\|_2^2 \|b\|_2^2$ para $a \neq 0$ (¿por qué?).

$$0 \leq \|\alpha a - \beta b\|_2^2 = \langle \alpha a - \beta b, \alpha a - \beta b\rangle = \alpha^2 \|a\|_2^2 - 2\alpha\beta\langle a, b\rangle + \beta^2 \|b\|_2^2.$$

Si tomamos $\alpha = \langle a, b\rangle$ y $\beta = \|a\|_2^2$:

$$\begin{aligned} 0 &\leq \|a\|_2^2 \langle a, b\rangle^2 - 2\|a\|_2^2 \langle a, b\rangle^2 + \|a\|_2^4 \|b\|_2^2, \\ \|a\|_2^2 \langle a, b\rangle^2 &\leq \|a\|_2^4 \|b\|_2^2, \\ \langle a, b\rangle^2 &\leq \|a\|_2^2 \|b\|_2^2. \end{aligned}$$

$\square$

¿Cómo se comparan las normas entre sí?

Si tomamos $a = (|x_1|, \ldots, |x_n|)$, $b = (1, \ldots, 1)$, por Cauchy-Schwarz:

$$\sum_{1 \leq j \leq n} |x_j| = \langle a, b \rangle \leq \|a\|_2 \|b\|_2 = \sqrt{n}\|x\|_2,$$

por lo tanto

$$\|x\|_1 \leq \sqrt{n}\|x\|_2.$$

Vale también que (ejercicio):

- $\|x\|_2 \leq \|x\|_1$,
- $\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$,
- $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$.

Definición:

Sea V un espacio vectorial. Una norma en V es una aplicación $\|\cdot\|: V \to \mathbb{R}$ tal que:

1. $\|x\| \geq 0$ para todo $x \in V$.
2. $\|x\| = 0 \;\Leftrightarrow\; x = 0$.
3. $\|\lambda x\| = |\lambda|\|x\|$.
4. $\|x + y\| \leq \|x\| + \|y\|$.

Teorema:

Dadas $\|\cdot\|$ y $|||\cdot|||$ normas en $\mathbb{R}^n$, existen constantes positivas C_1 y C_2 tales que

$$C_1\|x\| \leq |||x||| \leq C_2\|x\|.$$

Dem: Basta demostrar que cualquier norma es equivalente a la norma $\|\cdot\|_2$ (¿por qué?).

Sea e_i el i-ésimo vector canónico. Llamemos $C = \left(\sum_{i=1}^{n} |||e_i|||^2\right)^{1/2}$.

$$|||x||| = |||\sum_{j=1}^{n} x_j e_j||| \leq \sum_{j=1}^{n} |x_j||||e_j||| = \langle x, (|||e_1|||, \ldots, |||e_n|||)\rangle.$$

Por la desigualdad de Cauchy-Schwarz, vale

$$|||x||| \leq \|x\|_2 C \quad \Rightarrow \quad \frac{1}{C}|||x||| \leq \|x\|_2.$$

Consideramos $C_1 = \frac{1}{C}$.

Queremos ver ahora que existe C_2 tal que $\|x\|_2 \leq C_2 |||x|||$.

Supongamos que no: entonces para todo $m \in \mathbb{N}$ existe $y_m \in \mathbb{R}^n$ tal que $\|y_m\|_2 > m|||y_m|||$.

Sea $z_m = \frac{y_m}{\|y_m\|_2}$, y por lo tanto

$$\|z_m\|_2 = 1 \text{ y } |||z_m||| = \frac{1}{\|y_m\|_2}|||y_m||| < \frac{1}{m} \underset{m\to\infty}{\longrightarrow} 0.$$

Como $\{z_m\}$ es acotada en $\|\cdot\|_2$, existe una subsucesión convergente $\{z_{m_k}\}_k$ y $z \in \mathbb{R}^n$ tales que $\|z_{m_k} - z\|_2 \underset{k\to\infty}{\longrightarrow} 0$.
Vale entonces que $|||z_{m_k} - z||| \leq C\|z_{m_k} - z\|_2 \underset{k\to\infty}{\longrightarrow} 0$ y

$$0 \leq |||z||| \leq |||z - z_{m_k}||| + |||z_{m_k}||| \underset{k\to\infty}{\longrightarrow} 0.$$

Por lo tanto, $z = 0$ pero $\|z_{m_k}\|_2 = 1$ para todo k, y esto lleva a un absurdo. $\square$

Si tengo una matriz A y un vector b y resuelvo el sistema $Ax = b$, obteniendo x. Y ahora tengo el vector $\tilde{b}$ y obtengo $\tilde{x}$, ¿cómo se comparan $\dfrac{|||x-\tilde{x}|||}{|||x|||}$ y $\dfrac{|||b-\tilde{b}|||}{|||b|||}$?

Ejemplo:

$$A = \begin{pmatrix} 4{,}1 & 2{,}8 \\ 9{,}7 & 6{,}6 \end{pmatrix} \qquad b = \begin{pmatrix} 4{,}1 \\ 9{,}7 \end{pmatrix} \qquad x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\tilde{b} = \begin{pmatrix} 4{,}11 \\ 9{,}7 \end{pmatrix} \quad \tilde{x} = \begin{pmatrix} 0{,}34 \\ 0{,}97 \end{pmatrix}$$

$$\frac{\|x-\tilde{x}\|_1}{\|x\|_1} = 0{,}66 + 0{,}97 = 1{,}63 \text{ y } \frac{\|b-\tilde{b}\|_1}{\|b\|_1} = \frac{0{,}01}{13{,}8} = 0{,}7246 \times 10^{-3}.$$

Definición:

Sea $A \in \mathbb{R}^{n\times n}$, definimos la norma de A subordinada a la norma vectorial $|||\cdot|||$ como

$$|||A||| = \max_{x\neq 0} \frac{|||Ax|||}{|||x|||}.$$

Equivalentemente, $|||A||| = \max_{|||x|||=1} |||Ax|||$ (¿por qué?).

Tarea: ver que, efectivamente, es una norma.

Propiedades:

- $|||Ax||| \leq |||A||| \, |||x|||$. (tarea)

- $|||AB||| \leq |||A||| \, |||B|||$.

  Dem: $|||ABx||| \leq |||A||| \, |||Bx||| \leq |||A||| \, |||B||| \, |||x|||$ para todo x.

  Por lo tanto:

$$
\begin{aligned}
|||AB||| = \max_{x \neq 0} \frac{|||ABx|||}{|||x|||} &\leq \max_{x \neq 0} \frac{|||A||| \, |||B||| \, |||x|||}{|||x|||} = \\
&= \max_{x \neq 0} |||A||| \, |||B||| = |||A||| \, |||B|||.
\end{aligned}
$$

$\square$

Teorema:

Sea $A \in \mathbb{R}^{n\times n}$.

1. $\|A\|_\infty = \max\limits_{1\le i\le n} \sum\limits_{j=1}^{n} |a_{ij}|$.

2. Si A es simétrica y $\{\lambda_1, \dots, \lambda_n\}$ son sus autovalores, $\|A\|_2 = \rho(A) = \max\limits_{1\le j\le n} |\lambda_j|$.
   Si A no es simétrica, $\|A\|_2 = \sqrt{\rho(A^t A)}$.

3. $\|A\|_1 = \max\limits_{1\le j\le n} \sum\limits_{i=1}^{n} |a_{ij}|$.

Dem:

1.

$$\|Ax\|_\infty = \max_{1\le i\le n} \left|\sum_{j=1}^{n} a_{ij}x_j\right| \le \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}||x_j| \le \max_{1\le i\le n} \|x\|_\infty \sum_{j=1}^{n} |a_{ij}|$$

$$\|A\|_\infty = \max_{x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} \leq \max_{x\neq 0} \frac{\max\limits_{1\leq i\leq n} \|x\|_\infty \sum\limits_{j=1}^{n} |a_{ij}|}{\|x\|_\infty} = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|.$$

Sea i_0 el índice donde se alcanza el máximo. Es decir:

$$\max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}| = \sum_{j=1}^{n} |a_{i_0 j}|,$$

sea $\tilde{x} = (\mathrm{sg}(a_{i_0 1}), \ldots, \mathrm{sg}(a_{i_0 n}))$. Vale $\|\tilde{x}\|_\infty = 1$.

$$\begin{aligned} \|A\tilde{x}\|_\infty &= \max_{1\leq i\leq n} \left| \sum_{j=1}^{n} a_{ij}\tilde{x}_j \right| \geq \left| \sum_{j=1}^{n} a_{i_0 j}\tilde{x}_j \right| = \left| \sum_{j=1}^{n} |a_{i_0 j}| \right| \\ &= \sum_{j=1}^{n} |a_{i_0 j}| = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}| \end{aligned}$$

Entonces

$$\|A\|_\infty = \max_{x\neq 0}\frac{\|Ax\|_\infty}{\|x\|_\infty} \geq \frac{\|A\tilde{x}\|_\infty}{\|\tilde{x}\|_\infty} \geq \max_{1\leq i\leq n}\sum_{j=1}^{n}|a_{ij}|.$$

2. Sea A simétrica. Existe base ortonormal (bon) de autovectores (norma $\|\cdot\|_2$). Sean $\lambda_1,\ldots,\lambda_n$ los autovalores de A y $v_1,\ldots,v_n$ los autovectores asociados, tales que $\langle v_i, v_j\rangle = \begin{cases} 1 & \text{si } i=j \\ 0 & \text{si no} \end{cases}$.

Sea $x\in\mathbb{R}^n$. Entonces $x=\sum_{j=1}^{n}\alpha_j v_j$ y $\|x\|_2^2=\sum_{j=1}^{n}\alpha_j^2$.

$$Ax = \sum_{j=1}^{n}\alpha_j A v_j = \sum_{j=1}^{n}\alpha_j\lambda_j v_j.$$

$$\|Ax\|_2^2 = \sum_{j=1}^{n}(\alpha_j\lambda_j)^2 \leq \sum_{j=1}^{n}\alpha_j^2\lambda_{max}^2 = \lambda_{max}^2\sum_{j=1}^{n}\alpha_j^2 = \lambda_{max}^2\|x\|_2^2.$$

$\|A\|_2 = \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} \leq \max_{x\neq 0} \frac{|\lambda_{max}|\|x\|_2}{\|x\|_2} = |\lambda_{max}| = \rho(A).$
Sea i_0 tal que $\lambda_{max} = \lambda_{i_0}$,

$$\frac{\|Av_{i_0}|\_2}{\|v_{i_0}\|_2} = \|\lambda_{i_0} v_{i_0}\|_2 = |\lambda_{i_0}|\|v_{i_0}\| = |\lambda_{i_0}| = \rho(A).$$

Por lo tanto $\|A\|_2 \geq \rho(A)$.

Si A no es simétrica, $B = A^t A$ sí lo es. Sean $\mu_1, \ldots, \mu_n$ los autovalores de B (que son no negativos) y $w_1, \ldots, w_n$ los autovectores asociados que forman una bon.
Sea $x \in \mathbb{R}^n$. Entonces $x = \sum_{j=1}^{n} \alpha_j w_j$ y $Bx = \sum_{j=1}^{n} \alpha_j \mu_j w_j$.

$$
\begin{aligned}
\|Ax\|_2^2 &= \langle Ax, Ax, \rangle = x^t A^t Ax = x^t Bx = \langle x, Bx\rangle \\
&= \langle \sum_{j=1}^{n} \alpha_j w_j, \sum_{j=1}^{n} \alpha_j \mu_j w_j \rangle = \sum_{j=1}^{n} \mu_j \alpha_j^2 \\
&\geq \mu_{max} \sum_{j=1}^{n} \alpha_j^2 = \mu_{max} \|x\|_2^2
\end{aligned}
$$

Luego, $\|A\|_2 \leq \sqrt{\rho(A^t A)}$.

Sea i_0 tal que $\mu_{max} = \mu_{i_0}$,

$$
\frac{\|Aw_{i_0}\|_2}{\|w_{i_0}\|_2} = \sqrt{\langle Aw_{i_0}, Aw_{i_0}\rangle} = \sqrt{w_{i_0}^t A^t A w_{i_0}} = \sqrt{\mu_{i_0}} \|w_{i_0}\|_2 = \sqrt{\mu_{max}}
$$

Por lo tanto $\|A\|_2 \geq \sqrt{\rho(A^t A)}$.

3.

$$\|Ax\|_1 = \sum_{i=1}^{n}\left|\sum_{\ell=1}^{n} a_{i\ell}x_\ell\right| \leq \sum_{i=1}^{n}\sum_{\ell=1}^{n}|a_{ij}||x_\ell| = \sum_{\ell=1}^{n}|x_\ell|\left(\sum_{i=1}^{n}|a_{i\ell}|\right)$$
$$\leq \sum_{\ell=1}^{n}|x_\ell|\left(\max_{1\leq j\leq n}\sum_{i=1}^{n}|a_{ij}|\right) = \max_{1\leq j\leq n}\sum_{i=1}^{n}|a_{ij}|\|x\|_1$$

$$\|A\|_1 = \max_{x\neq 0}\frac{\|Ax\|_1}{\|x\|_1} \leq \max_{x\neq 0}\frac{\max\limits_{1\leq i\leq n}\ \|x\|_1\sum_{i=1}^{n}|a_{ij}|}{\|x\|_1} = \max_{1\leq j\leq n}\sum_{i=1}^{n}|a_{ij}|.$$

Sea j_0 el índice donde se alcanza el máximo. Es decir:

$$\max_{1\leq j\leq n}\sum_{i=1}^{n}|a_{ij}| = \sum_{i=1}^{n}|a_{ij_0}|.$$

Sea $\tilde{x} = e_{j_0}$. Vale $\|\tilde{x}\|_1 = 1$ y $A\tilde{x} = (a_{1j_0}, \dots, a_{nj_0})$.

$$\|A\tilde{x}\|_1 = \sum_{i=1}^{n} |a_{ij_0}| = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}|.$$

Entonces

$$\|A\|_1 = \max_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} \geq \frac{\|A\tilde{x}\|_1}{\|\tilde{x}\|_1} = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}|.$$

$\square$

Teorema:

Sea $A \in \mathbb{R}^{n\times n}$, A inversible. Sea $b \in \mathbb{R}^n - \{0\}$ y x tal que $Ax = b$. Sean $\tilde{b}$ y $\tilde{x}$ tales que $A\tilde{x} = \tilde{b}$. Sea $|||\cdot|||$ norma vectorial.

$$\frac{1}{|||A|||\,|||A^{-1}|||}\frac{|||b-\tilde{b}|||}{|||b|||} \leq \frac{|||x-\tilde{x}|||}{|||x|||} \leq |||A|||\,|||A^{-1}|||\frac{|||b-\tilde{b}|||}{|||b|||}.$$

Definición:

Llamamos *número de condición de la matriz* A *en norma* $|||\cdot|||$ al número $\mathrm{cond}_{|||\cdot|||}(A) = |||A|||\,|||A^{-1}|||$.

Observación:

- $\mathrm{cond}_{|||\cdot|||}(A) = \mathrm{cond}_{|||\cdot|||}(A^{-1})$

- $|||I||| = 1$ para toda $|||\cdot|||$ subordinada a una norma vectorial (¿por qué?).

$$1 = |||I||| = |||AA^{-1}||| \leq |||A||| |||A^{-1}||| = \mathrm{cond}_{|||\cdot|||}(A).$$

Una matriz se dice *mal condicionada* cuando su condición es mucho mayor que 1.

Dem. del Teo.:

$$|||b - \tilde{b}||| = |||A(x - \tilde{x})||| \leq |||A||| \, |||x - \tilde{x}|||$$

$$|||x||| = |||A^{-1}b||| \leq |||A^{-1}||| \, |||b|||$$

Luego, $\frac{|||x-\tilde{x}|||}{|||x|||} \geq \frac{1}{|||A||| \, |||A^{-1}|||} \frac{|||b-\tilde{b}|||}{|||b|||}$.

$$|||x - \tilde{x}||| = |||A^{-1}(b - \tilde{b})||| \leq |||A^{-1}||| \, |||b - \tilde{b}|||$$

$$|||b||| = |||Ax||| \leq |||A||| \, |||x|||$$

Luego, $\frac{|||x-\tilde{x}|||}{|||x|||} \leq |||A||| \, |||A^{-1}||| \frac{|||b-\tilde{b}|||}{|||b|||}$, como queríamos probar. $\square$

Ejemplo:

A simétrica e inversible, sean $\lambda_1, \ldots, \lambda_n$ los autovalores de A y $v_1, \ldots, v_n$ los autovectores asociados que forman una bon.
Llamemos:

- λ_{max} a un autovalor de A tal que $|\lambda_{max}| \geq |\lambda_i|\ \forall i$, y sea v_{max} su autovector asociado.
- λ_{min} a un autovalor de A tal que $|\lambda_{min}| \leq |\lambda_i|\ \forall i$, y sea v_{min} su autovector asociado.

Sea $b \neq 0$ en el autoespacio de λ_{max}, entonces $Ab = \lambda_{max} b$ y $x = A^{-1}b = \frac{1}{\lambda_{max}} b$.
Sea $\tilde{b} = b + \varepsilon v_{min}$. Entonces,

$$x - \tilde{x} = A^{-1}b - A^{-1}(b + \varepsilon v_{min}) = -A^{-1}\varepsilon v_{min} = -\frac{\varepsilon}{\lambda_{min}} v_{min}.$$

$$\frac{\|x - \tilde{x}\|_2}{\|x\|_2} = \frac{\| - \frac{\varepsilon}{\lambda_{min}} v_{min}\|_2}{\| \frac{1}{\lambda_{max}} b\|_2} = \frac{|\lambda_{max}|}{|\lambda_{min}|} \frac{\| - \varepsilon v_{min}\|_2}{\|b\|_2} = \frac{|\lambda_{max}|}{|\lambda_{min}|} \frac{\|b - \tilde{b}\|_2}{\|b\|_2}$$

Pero $|\lambda_{max}| = \rho(A) = \|A\|_2$ y $\frac{1}{|\lambda_{min}|} = \rho(A^{-1}) = \|A^{-1}\|_2$ (verificar).

Luego, $\frac{\|x - \tilde{x}\|_2}{\|x\|_2} = \mathrm{cond}_{\|\|.\|\|}(A) \frac{\|b - \tilde{b}\|_2}{\|b\|_2}$.

¿Y si tomamos b en la dirección de v_{min} y $b - \tilde{b}$ en la dirección de v_{max}?

Teorema:

Sea $A \in \mathbb{R}^{n\times n}$ inversible, $|||\cdot|||$ norma vectorial.

$$\frac{1}{\mathrm{cond}_{|||\cdot|||}(A)} = \inf_{B \text{ singular}} \frac{|||A-B|||}{|||A|||}.$$

Dem.: Sea B singular, entonces existe $\hat{x} \neq 0$ tal que $B\hat{x} = 0$.

$|||\hat{x}||| = |||A^{-1}A\hat{x}||| = |||A^{-1}(A\hat{x} - B\hat{x})||| \leq |||A^{-1}||| \, |||A-B||| \, |||\hat{x}|||.$

Como $|||\hat{x}||| \neq 0$, $1 \leq |||A^{-1}||| \, |||A-B|||$ y por lo tanto

$$\frac{1}{\mathrm{cond}_{|||\cdot|||}(A)} = \frac{1}{|||A||| \, |||A^{-1}|||} \leq \frac{|||A-B|||}{|||A|||}.$$

Sea $y \neq 0$ tal que $|||y||| = \frac{1}{|||A^{-1}|||}$ y tal que $|||A^{-1}y||| = |||A^{-1}||| |||y|||$.

Sea x tal que $Ax = b$, luego $|||x||| = |||A^{-1}y||| = |||A^{-1}||| |||y||| = 1$.

Se puede probar que existe z tal que

$$\langle z, x \rangle = 1, \ \langle z, u \rangle \leq 1 \ \forall u \text{ tal que } |||u||| = 1.$$

Notar que si $|||u||| = 1$, se tiene que $|||-u||| = 1$ y entonces $\langle z, u \rangle \leq 1$, $\langle z, -u \rangle \leq 1$ y por lo tanto $|\langle z, u \rangle| = |z^t u| \leq 1$.

Sea $B = A - yz^t$. Vale que $Bx = Ax - yz^t x = y - y = 0$, y por lo tanto B es singular.

$$|||yz^t u||| = |||y||| \cdot |z^t u| \leq |||y||| = \frac{1}{|||A^{-1}|||}, \quad \forall |||u||| = 1.$$

$$|||A - B||| = |||yz^t u||| \leq \frac{1}{|||A^{-1}|||},$$

y esto implica $\frac{|||A-B|||}{|||A|||} \leq \frac{1}{|||A||| |||A^{-1}|||} = \frac{1}{\mathrm{cond}_{|||\cdot|||}(A)}$. $\square$