Elementos de Cálculo Numérico

Juan Pablo Pinasco (jpinasco@dm.uba.ar//gmail.com)

Departamento de Matemática e IMAS,
FCEyN, UBA - CONICET

2020

Parte I

Próxima clase

Hoy:

- Normas matriciales: 15-17
- Número de condición: 17-29

Próxima:

- LU: 35-37
- Cholesky: 37-40

Parte II

Normas matriciales

Def: Sea $x \in \mathbb{R}^n$, $x = (x_1, \cdots, x_n)$. Una norma $\|\cdot\|$ es una función que satisface:

(1) $\|x\| \geq 0$ para todo $x \in \mathbb{R}^n$.

(2) $\|x\| = 0$ si y sólo si $x = 0$.

(3) $\|\lambda x\| = |\lambda|\|x\|$ para todo $x \in \mathbb{R}^n$ y todo $\lambda \in \mathbb{R}$.

(3) $\|x + y\| \leq |\|x\| + \|y\|$, para todo x, $y \in \mathbb{R}^n$ (desigualdad triangular).

Teorema

Dadas dos normas $\|\cdot\|, \|\cdot\|'$ en $\mathbb{R}^n$, son equivalentes. Es decir, existen dos constantes c, C tales que

$$c\|x\| \leq \|x\|' \leq C\|x\|$$

para todo $x \in \mathbb{R}^n$.

Dem: Alcanza con demostrar que cualquier norma es equivalente a $\|\cdot\|_2$,

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

Definamos $C = \left(\sum_{i=1}^{n} e_i^2\right)^{1/2}$, donde $\{e_i\}_{i=1}^{n}$ es la base canónica de $\mathbb{R}^n$.

Por la desigualdad triangular, la prop. 3 de la norma, y Cauchy-Schwartz,

$$\begin{aligned} \|x\| =& \|\sum_{i=1}^{n} x_i e_i\| \\ \leq& \sum_{i=1}^{n} |x_i| \|e_i\| \\ \leq& (\sum_{i=1}^{n} |x_i|^2)^{1/2} (\sum_{i=1}^{n} \|e_i\|^2)^{1/2} \\ =& C\|x\|_2, \end{aligned}$$

y tenemos una desigualdad, $\|x\| \leq C\|x\|_2$.

La otra sale por el absurdo. Supongamos que no existe K tal que

$$\|x\|_2 \leq K\|x\|.$$

Entonces, para cada $m \in \mathbb{N}$, existe un y_m tal que

$$\|y_m\|_2 \geq m\|y_m\|.$$

Definamos $x_m = \frac{y_m}{\|y_m\|_2}$

Tenemos $\|x_m\|_2 = 1$, pero $\|x_m\| \leq 1/m$.

Como los $\{x_m\}_{m\geq 1}$ están acotados en norma 2, existe un x y una subsucesión que converge a x,

$$\|x_m - x\|_2 \to 0, \qquad y \qquad \|x_m - x\| \to 0$$

por la primera parte.

Pero $\|x_m\| \to 0$ (porque era menor a $1/m$), así que $x_m \to 0$, y $x = 0$ por unicidad del límite.

Ahora,

$$\begin{aligned} 1 =& \|x_m\|_2 \\ =& \|x_m - x + x\|_2 \\ \leq& \|x_m - x\|_2 + \|x\|_2 \\ & \to 0, \end{aligned}$$

y tenemos el absurdo que buscábamos.

Luego, tiene que existir un K tal que $\|x\|_2 \leq K\|x\|$ para todo x, y tenemos

$$\begin{aligned} \|x\|_2 &\leq K\|x\| \\ \|x\| &\leq C\|x\|_2 \end{aligned}$$

Si $c = 1/K$ tenemos

$$c\|x\|_2 \leq |x\| \leq C\|x\|_2$$

Para dos normas $\|\cdot\|$, $|||\cdot|||$ tenemos

$$\|x\| \leq C_1\|x\|_2 \leq K_2C_1|||x|||$$

$$|||x||| \leq C_2\|x\|_2 \leq K_1C_2\|x\|$$

donde C_1, K_1 son las constantes de la equivalencia entre $\|\cdot\|$ y $\|\cdot\|_2$, y C_2, K_2 son las constantes de la equivalencia entre $|||\cdot|||$ y $\|\cdot\|_2$,

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

$$\|x\|_\infty = \max_i |x_i| \leq \max_i (|x_i|^2)^{1/2} \leq \Big(\sum_{j=1}^n |x_j|^2\Big)^{1/2} = \|x\|_2$$

$$\|x\|_2 = \Big(\sum_{j=1}^n |x_j|^2\Big)^{1/2} \leq \Big(\sum_{j=1}^n [\max_i |x_i|]^2\Big)^{1/2} = \Big(n[\max_i |x_i|]^2\Big)^{1/2}$$
$$= n^{1/2} \max_i |x_i| = n^{1/2}\|x\|_\infty$$

- $\|x\|_1 = \sum_{i=1}^n |x_i| = |x_1| + \cdots + |x_n|$

$$\|x\|_\infty = \max_i |x_i| \leq |x_1| + \cdots + |x_n| = \|x\|_1$$
$$\|x\|_1 = |x_1| + \cdots + |x_n| \leq \max_i |x_i| + \cdots + \max_i |x_i| = n\|x\|_\infty$$

- Obs: $\|x\|_1 \leq n\|x\|_\infty \leq n\|x\|_2$
  $\|x\|_2 \leq n^{1/2}\|x\|_\infty \leq n^{1/2}\|x\|_1$

Podemos definir una norma para una matriz $A \in \mathbb{R}^{n\times n}$ en términos de una norma definida en $\mathbb{R}^n$, y la llamaremos la norma inducida por esta.
Def: Sea $\|\cdot\|$ una norma en $\mathbb{R}^n$. Definimos la norma inducida en $\mathbb{R}^{n\times n}$

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

10.- Observaciones:

- No vamos a demostrarlo, pero la norma inducida realmente es una norma sobre las matrices.
  Si no, la hubiéramos llamado distinto! El punto es que cumple las propiedades de ser positiva; es cero sólo si A es la matriz 0; cumple la desigualdad triangular, y $\|cA\| = |c|\|A\|$.

- Cumple una propiedad extra, $\|AB\| \leq \|A\|\|B\|$

- Vale que $\|Ax\| \leq \|A\|\|x\|$, porque si dividimos por $\|x\|$, tenemos que la desigualdad vale ya que $\|A\|$ era el máximo de los cocientes. Si $x = 0$, no podemos dividir por $\|x\|$ pero la desigualdad también vale.

- Se tienen normas inducidas para $A \in \mathbb{R}^{n\times m}$, incluso usando una norma $\|\cdot\|_a$ en $\mathbb{R}^n$ y otra $\|\cdot\|_b$ en $\mathbb{R}^m$:

$$\|A\|_{a,b} = \max_{x\neq 0} \frac{\|Ax\|_b}{\|x\|_a}.$$

  aunque no las vamos a usar.

- $\|A\|_\infty = \text{máx}_{1\leq i\leq n} \sum_{j=1}^n |a_{ij}|$.
- $\|A\|_1 = \text{máx}_{1\leq j\leq n} \sum_{i=1}^n |a_{ij}|$.
- $\|A\|_2 = \text{máx}\{\sqrt{\sigma_1}, \dots, \sigma_n\}$, autovalores de $B = A^t A$.

Obs: no toda norma es inducida, por ejemplo la de Frobenius o Hilbert-Schmidt,

$$\|A\|_F = \Big(tr(AA^t)\Big)^{1/2} = \Big(\sum_{i,j=1}^n |a_{ij}|\Big)^{1/2}$$

Aunque no sea inducida, es fácil de calcular, y es equivalente a todas las anteriores, ya que se tiene, para $A \in \mathbb{R}^{m\times n}$:

- $\frac{1}{n^{1/2}}\|A\|_\infty \leq \|A\|_2 \leq m^{1/2}\|A\|_\infty$.
- $\frac{1}{m^{1/2}}\|A\|_1 \leq \|A\|_2 \leq n^{1/2}\|A\|_1$.
- $\|A\|_2 \leq \|A\|_F \leq r^{1/2}\|A\|_2$, donde $r = rango(A)$.

Obs: $\|A\|_2 \leq \Big(\|A\|_1\|A\|_\infty\Big)^{1/2}$

12.- Conclusiones:

- Si una norma es muy grande, o muy chica, las otras también (teniendo en cuenta que puede haber una constante que depende de la cantidad de filas/columnas multiplicando o dividiendo).

- Frobenius, $\|A\|_1$ y $\|A\|_\infty$ se calculan fácil:

$$\begin{pmatrix} 0 & -1 \\ 2 & 3 \end{pmatrix} \quad \rightsquigarrow \quad \sqrt{14}, 4, 5$$

- La norma $\|\cdot\|_2$ implica calcular autovalores, y -en algún sentido- explica qué está pasando. Va a quedar más claro ahora con número de condición.

Def: Dada $A \in \mathbb{R}^{n\times n}$ inversible, y $\|\cdot\|$ una norma en $\mathbb{R}^n$, definimos el número de condición de A

$$Cond(A) = \|A\|\|A^{-1}\|$$

Obs.: depende de la norma elegida.

Teorema

$A \in \mathbb{R}^{n\times n}$ *inversible,* $\|\cdot\|$ *una norma:*

(1) $Cond(A) = Cond(A^-1)$

(2) $Cond(A) \geq 1$

Dem: (1) es directo por la definición.

(2) Sale porque

$$1 = \|Id\| = \|AA^{-1}\| \leq \|A\|\|A^{-1}\| = Cond(A)$$

14.- Número de Condición: para qué sirve?

Teorema

Sea $A \in \mathbb{R}^{n\times n}$ *inversible,* $b, \Delta b \in \mathbb{R}^n$.

Si $Ax = b$, $A(x+\Delta x) = b + \Delta b$, *entonces*

$$[a.-] \qquad \frac{\|\Delta x\|}{\|x\|} \leq Cond(A)\frac{\|\Delta b\|}{\|b\|}$$

Además,

$$[b.-] \qquad \frac{1}{Cond(A)}\frac{\|\Delta b\|}{\|b\|} \leq \frac{\|\Delta x\|}{\|x\|}$$

y en ambos casos vale la igualdad para ciertos b, Δb.

Tenemos cotas del error.

15.- Número de Condición: cómo estimarlo sin invertir?

Dem: $[a.-]$ $\frac{\|\Delta x\|}{\|x\|} \leq Cond(A) \frac{\|\Delta b\|}{\|b\|}$

Si $Ax = b$ y $A(x + \Delta x) = b + \Delta b$, entonces $A\Delta x = \Delta b$.

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\|A^{-1}\|\|\Delta b\|}{\|x\|}$$

$$= \frac{\|A^{-1}\|\|\Delta b\|}{\|b\|} \frac{\|b\|}{\|x\|}$$

$$= \frac{\|A^{-1}\|\|\Delta b\|}{\|b\|} \|A\| = Cond(A) \frac{\|\Delta b\|}{\|b\|}$$

(porque $\|b\| = \|Ax\| \leq \|A\|\|x\|$) y vale el igual si $\|A^{-1}\Delta b\| = \|A^{-1}\|\|\Delta b\|$.

Dem: $[b.-]$ $\frac{1}{Cond(A)}\frac{\|\Delta b\|}{\|b\|} \leq \frac{\|\Delta x\|}{\|x\|}$

Hacemos trampa:

$$\frac{1}{Cond(A)}\frac{\|\Delta b\|}{\|b\|} \leq \frac{\|\Delta x\|}{\|x\|}$$

$$\frac{\|\Delta b\|}{\|b\|} \leq Cond(A)\frac{\|\Delta x\|}{\|x\|}$$

$$\frac{\|\Delta b\|}{\|b\|} \leq Cond(A^{-1})\frac{\|\Delta x\|}{\|x\|}$$

Pero si $Ax = b$, $A^{-1}b = x$ y vale intercambiando x y b en la parte $[a.-]$ porque $Cond(A) = Cond(A^{-1})$.

Falta ver que se alcanza:

17.- Número de Condición: para qué sirve?

Teorema

Sea $A \in \mathbb{R}^{n\times n}$ *inversible,* $\Delta A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$.

Si $Ax = b$, $(A + \Delta A)(x + \Delta x) = b$, *entonces*

$$[a.-] \qquad \frac{\|\Delta x\|}{\|x + \Delta x\|} \leq Cond(A)\frac{\|\Delta A\|}{\|A\|}$$

Si

$$Cond(A)\frac{\|\Delta A\|}{\|A\|} \leq \delta < 1, \text{ entonces}$$

$$[b.-] \qquad \frac{\|\Delta x\|}{\|x\|} \leq \frac{1}{1-\delta} Cond(A)\frac{\|\Delta A\|}{\|A\|}$$

Es parecido al anterior, pero diferente. Mira el error de los coeficientes del sistema.

18.- Número de Condición: para qué sirve?

Dem: $[a.-]$ Si $Ax = b$, $(A + \Delta A)(x + \Delta x) = b$, $\frac{\|\Delta x\|}{\|x+\Delta x\|} \leq Cond(A)\frac{\|\Delta A\|}{\|A\|}$

$$(A + \Delta A)(x + \Delta x) = b$$

$$A(x + \Delta x) + \Delta A(x + \Delta x) = b$$

$$A\Delta x = -\Delta A(x + \Delta x)$$

$$\Delta x = -A^{-1}\Delta A(x + \Delta x)$$

$$\|\Delta x\| \leq \|A^{-1}\|\|\Delta A\|\|x + \Delta x\| \leq Cond(A)\frac{\|\Delta A\|}{\|A\|}\|x + \Delta x\|$$

y listo.

Dem: Probemos $[b.-]$ $\frac{\|\Delta x\|}{\|x\|} \leq \frac{1}{1-\delta} Cond(A) \frac{\|\Delta A\|}{\|A\|}$

$$\frac{\|x+\Delta x\|}{\|x\|} \leq \frac{\|x\| + \|\Delta x\|}{\|x\|} = 1 + \frac{\|\Delta x\|}{\|x\|}$$

Por lo anterior,

$$\frac{\|\Delta x\|}{\|x\|} = \frac{\|\Delta x\|}{\|x+\Delta x\|} \frac{\|x+\Delta x\|}{\|x\|} \leq Cond(A) \frac{\|\Delta A\|}{\|A\|} \left(1 + \frac{\|\Delta x\|}{\|x\|}\right)$$

Usamos la hipótesis $Cond(A) \frac{\|\Delta A\|}{\|A\|} \leq \delta < 1$:

$$\frac{\|\Delta x\|}{\|x\|} \leq Cond(A) \frac{\|\Delta A\|}{\|A\|} + \delta \frac{\|\Delta x\|}{\|x\|}$$

Pasamos restando, sacamos factor común, y nos queda

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{1}{1-\delta} Cond(A) \frac{\|\Delta A\|}{\|A\|}$$

20.- Número de Condición: cómo estimarlo sin invertir?

Dos formas:

- Autovalores
- Qué tanto se parece a una matriz no inversible.

Autovalores es fácil de entender por qué funciona, pero difícil de usar.

La otra es difícil de entender por qué funciona, pero muy fácil de usar.

21.- Número de Condición: cómo estimarlo sin invertir?

- Autovalores

$\|A\|_2 = \sqrt{\rho(A^t A)} = \sqrt{|\sigma_{max}}$ radio espectral, raíz del autovalor de mayor módulo de $A^t A$.

Y $\|A^{-1}\|_2 = 1/\sqrt{|\sigma_{min}}$, con lo cual

$$Cond_2(A) = \frac{\sigma_{max}}{\sigma_{min}}$$

Si A es simétrica, se usan $|\lambda_{max}|$ y $|\lambda_{min}|$.

Para otra norma, si recuerdan las constantes en la equivalencia que vimos antes, tienen cotas de $Cond_\infty(A)$, $Cond_1(A)$.

- Qué tanto se parece a una matriz no inversible (singular, con determinante cero, tiene núcleo).

Teorema

Sea $A \in \mathbb{R}^{n\times n}$ inversible y una norma cualquiera. Entonces

$$\frac{1}{Cond(A)} = \inf_{B\ singular} \frac{\|A-B\|}{\|A\|}$$

Sea B singular, y $x \in \mathbb{R}^n$ no nulo tal que $Bx = 0$.

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

$$1 \leq \|A^{-1}\|\frac{\|A\|}{\|A\|}\|A-B\| = Cond(A)\frac{\|A-B\|}{\|A\|}$$

Teníamos

$$1 \leq Cond(A)\frac{\|A - B\|}{\|A\|}$$

Entonces, para cualquier B,

$$\frac{1}{Cond(A)} \leq \frac{\|A - B\|}{\|A\|}$$

y es menor o igual al ínfimo.

En realidad, ese ínfimo se alcanza, pero la demostración no la vamos a hacer porque es larga, la pueden ver en el Apunte.