# Condicionamiento y Estabilidad

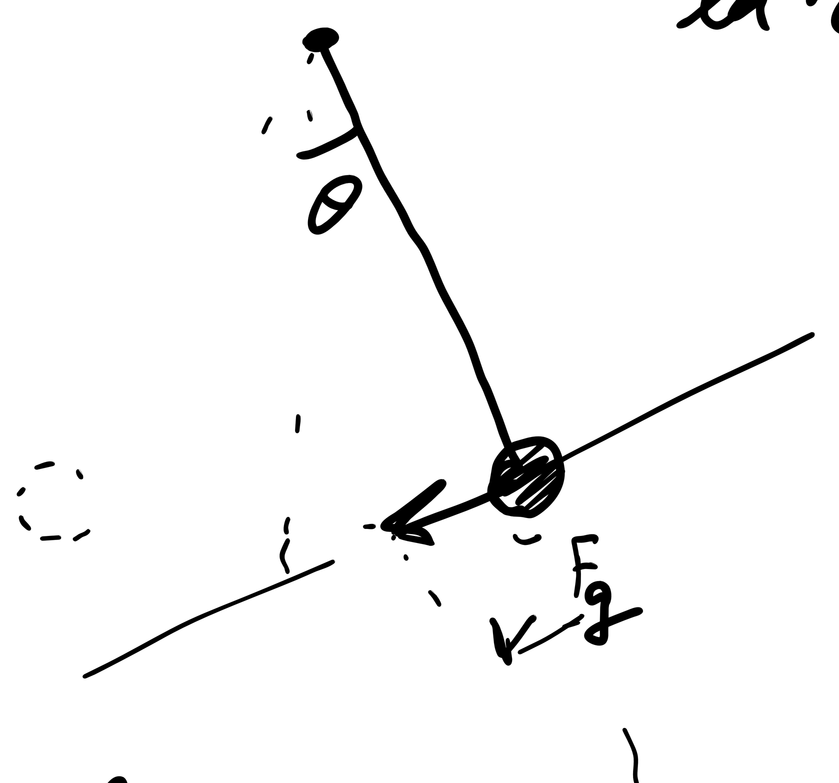
ALN 2022
22/9/2022
Clase 9

## Errores en Análisis Numérico

Existen tres tipos bien diferentes de errores

### 1) Errores en el Modelo:

Para estudiar ciertos problemas es común estudiar una versión simplificada del problema original.
Por ejemplo, si queremos estudiar la dinámica oscilatoria del péndulo la ecuación diferencial a mirar sería




$$\theta'' = -\frac{g}{\ell}\,\text{sen}\,\theta$$

donde el seno de $\theta$ resulta de considerar la proyección tangencial de la fuerza gravitatoria.

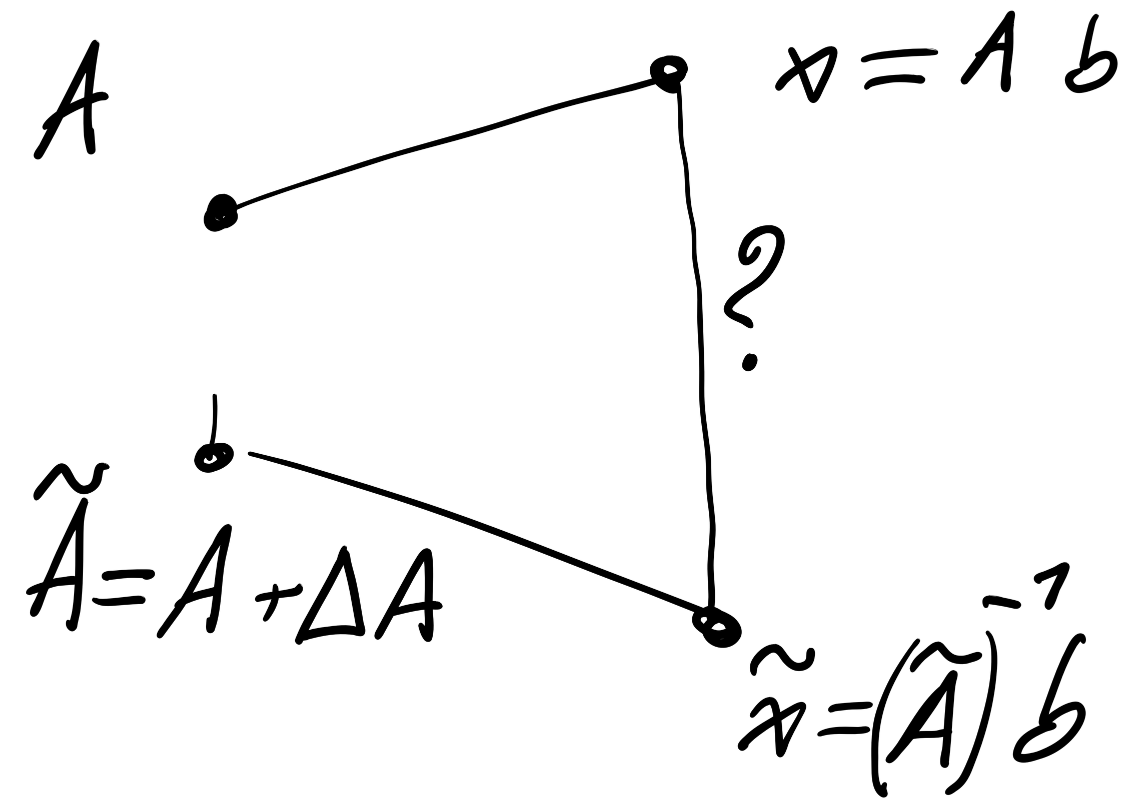
Sin embargo lo que comúnmente se estudia es la ec. dif. $\theta'' = -\frac{g}{\ell}\,\theta$ que resulta de aproximar sen $\theta$ con $\theta$ para ángulos pequeños. Sobre este tipo de errores no podemos decir mucho.

### 2) Errores en los datos

Típicamente nos enfrentamos a resolver un problema aproximado. Esto resulta de que nuestros datos pueden estar afectados por ruido, o simplemente porque al ingresar nuestro input a nuestra computadora, esta lo aproximó.
Por lo tanto si $\underline{a}$ es nuestro input y $S(\underline{a})$ es nuestra solución (output) en realidad estamos dando como solución $S(\underline{a}')$ con $\underline{a}'$ una entrada aproximada a $\underline{a}$.

La pregunta entonces es ¿cuál es la distancia de $S(\tilde{a})$ a $S(a)$ en función de la aproximación (o margen de error) de $a$ a $\tilde{a}$?

Por ejemplo, si queremos resolver el sistema lineal $Ax=b$, donde $\underline{\underline{b}}$ está dado exactamente, pero $A$ se ingresó con errores $\tilde{A}=A+\Delta A$. Luego estamos resolviendo el nuevo sistema $\tilde{A}x=b$ que tendrá como solución a $\tilde{x}$.




La pregunta es cómo controlar $\|x-\tilde{x}\|$ en función de $\|\tilde{A}-A\|$.

Veamos un ejemplo más sencillo y donde podemos cuantificar el error fácilmente.

Supongamos que queremos calcular $\sqrt{a}$, con $a\in\mathbb{R}$.

Ahora, en realidad nuestro dato $\underline{a}$ está aproximado por $\tilde{a}=a+h$ con $h>0$ muy chiquito (por ejemplo)

Luego tenemos $\sqrt{\tilde{a}}-\sqrt{a}=\sqrt{a+h}-\sqrt{a}=\dfrac{1}{\sqrt{a+h}+\sqrt{a}}\cdot h$

por lo que $\sqrt{a+h}-\sqrt{a}=\dfrac{h}{2\sqrt{a}}+\underbrace{O(h^2)}_{\to\, \leq C\cdot h^2}$

Por lo tanto el error en la solución es $\dfrac{1}{2\sqrt{a}}$ por el error en la entrada.

A este factor, que mide la sensibilidad de nuestras soluciones en función de perturbaciones en las entradas se llama <u>número de condición</u>

Podemos escribir en este caso $\left|\sqrt{a+h}-\sqrt{a}\right| \leq \underbrace{\mathrm{cond}(\sqrt{\cdot}, a)}_{=\frac{1}{2\sqrt{a}}} \cdot h$ nº de condición.

En las siguientes clases nos concentraremos en estudiar este "número" que es una característica fundamental en análisis numérico.

## 3) Errores en los Cálculos

Observar que en el punto anterior, ~~aún~~ no hablamos de algoritmos. De hecho asumimos que nuestro "mapa solución" $S$ (en sist. lin $S(A,b)=A^{-1}b$, en el otro ejemplo $S(a)=\sqrt{a}$) resuelve el problema de manera exacta. ~~Así~~ En este sentido el nº de condición es algo intrínseco del problema, e independiente del algoritmo utilizado.

Aquí llegamos al tercer tipo de errores, que están asociados a nuestros algoritmos. Hay al menos tres razones para explicar la existencia de estos errores.

i) Nuestras soluciones vienen dadas (en los problemas generales) límites de procesos iterativos, y nuestros algoritmos paran en una cantidad finita de pasos.

ii) Números irracionales, funciones trascendentes son reemplazadas por aproximaciones

iii) La utilización de las computadoras que no escapan de la aritmética de punto flotante.

Por lo tanto, nuestro mapa solución $S : \mathcal{I} \longrightarrow \mathcal{O}$ es aproximado por un mapa $\tilde{S}$.
esp. de inputs → esp. de outputs

El estudio de la estabilidad de nuestros algoritmos intenta entender estos dos problemas.

a) ¿Cuál es la estimación del error $\|\tilde{S}(a) - S(a)\|$?

b) ¿Cuál es el análisis del error inverso o comunmente llamado backward-stability?

Sobre b) la pregunta es saber si $\exists \tilde{a}$ cercano al input $a$ tal que $\tilde{S}(a) = S(\tilde{a})$ y en tal caso cuantificar

$$\inf_{S(\tilde{a})=\tilde{S}(a)} \| a - \tilde{a} \| .$$

Esto lo analizaremos en los algoritmos ya dados en las próximas clases.

---

## Numero de Condición

Consideremos $I$ el espacios de inputs (o entradas) y $O$ el espacio de outputs (o salidas)

Supongamos que tenemos nuestro mapa solución $S : I \to O$.

Como veremos esto aplica para algunos ejemplos, pero no para otros donde ~~pue~~ $S$ puede ser multivaluada y/o estar definida localmente.

Veamos algunos ejemplos.

(1) Sistemas Lineales: $\mathcal{I} = \mathbb{C}^{n\times n}$ o $\mathcal{I} = Gl_n(\mathbb{C})$ de matrices invertibles, o $\mathcal{I} = Gl_n(\mathbb{C}) \times \mathbb{C}^n$ donde una entrada es $(A,b) \in \mathcal{I}$ y donde $\mathcal{O} = \mathbb{C}^n$ y el mapa solución está dado por $S(A,b) = A^{-1}b \in \mathbb{C}^n$.

(2) Raíz cuadrada: $\mathcal{I} = \mathbb{R}^+$, $\mathcal{O} = \mathbb{R}^+$, $S(a) = \sqrt{a}$.

(3) Raíces de polinomios: $\mathcal{I} = \mathcal{P}d(\mathbb{C}) \equiv \mathbb{C}^{d+1}$ polinomios de grado menor o igual a $d$ y $\mathcal{O} = \mathbb{C}$ (miramos una raíz). Queremos que $S: \mathcal{I} \to \mathcal{O}$ nos de una raíz. Es claro que acá tenemos un problema. Como vemos podemos ~~invocar~~ recurrir al Teo. Función Implícita y observar que $S$ está dada implícitamente. $c_0 + c_1\zeta + \cdots + c_d\zeta^d = 0$. $\zeta$ en función $(c_0, \ldots, c_d)$.

(4) Problema Simétrico de Valores Propios: $\mathcal{I} = Sym_n(\mathbb{R})$ espacio de matrices reales simétricas y $\mathcal{O} = \mathbb{R}^n \times \mathbb{R}$ donde $A \in Sym_n(\mathbb{R})$ y $S(A) = (v,\lambda) \in (\mathbb{R}^n)^* \times \mathbb{R}$ tq. $Av = \lambda v$

O podemos tomar variantes del output tipo $\mathcal{O} = S^{n-1} \times \mathbb{R}$ pidiendo que $\|v\| = 1$, o $\mathcal{O} = \mathbb{P}^n \times \mathbb{R}$ donde $\mathbb{P}^{n-1} = \mathbb{P}(\mathbb{R}^n)$ es el espacio proyectivo de rectas en $\mathbb{R}^n$ (rectas por el origen, i.e. subespacios de dim 1).

(5) Núcleos de matrices de rango $r$: $\mathcal{I} = R_r \subset \mathbb{C}^{n\times n}$ y $\mathcal{O} = \mathbb{G}(n-r, n)$ = Grassmaniana: subespacios de $\mathbb{R}^n$ de dimensión $n-r$. Luego $S(A) = \operatorname{Ker} A \in \mathbb{G}(n-r, r)$.

Asumiendo que nuestro mapa $S$ está bien definido y es $C^1$ tenemos

$$S(a') - S(a) = DS(a)\cdot(a'-a) + o(\|a'-a\|)$$

Por lo tanto $\quad \|S(a') - S(a)\| \leq \|DS(a)\|\cdot\|a'-a\| + o(\|a'-a\|)$

El nº $\|DS(a)\|$ es lo que llamamos <u>numero de condición del problema</u>, y ~~lo denotamos por $\mu(a)$ si el problema está sobre controlado.~~

El nº de condición mide la sensibilidad del output respecto a pequeñas perturbaciones en el input.

Veamos algunos ejemplos antes de dar un marco geométrico que nos permitirá estudiar varios ejemplos, aún los multivaluados

Consideremos el caso lineal (sistema lineal)

$$\text{Supongamos que } b \text{ está fijo, luego } S(A) = A^{-1}b =: \underset{\substack{\uparrow \\ \text{output}}}{y}$$

$$DS(A)\dot{A} = -A^{-1}\dot{A}A^{-1}b = -A^{-1}\dot{A}y \,,$$

Luego tomar $\langle , \rangle_F$ en $\mathbb{C}^{n\times n}$ y tomando el máximo $\|\dot{A}\|_F = 1$. es fácil ver que

$$\max_{\|\dot{A}\|_F=1} \|A^{-1}\dot{A}y\| = \|A^{-1}\|_{op}\cdot\|y\|.$$

Luego $\quad \|DS(A)\| = \|A^{-1}\|_{op}\cdot\|y\| \quad$ donde $y = S(A)$ es el output.

y por lo tanto el nº de condición es $\|A^{-1}\|\cdot\|y\|$.

Si normalizamos (nº de cond. relativo) por $\|S(a)\|$ obtenemos $\|A^{-1}\|$.
Más adelante veremos que también se normaliza por el input también dando la forma clásica $\kappa(A) = \|A\|\,\|A^{-1}\|$.