

Métodos Numéricos

U1: Números, Errores, Estabilidad y Normas

Diego Passarella

Universidad Nacional de Quilmes

Repaso

Algunos teoremas que tienen que recordar:

- Rolle
- Valor medio
- Taylor
- ...

También hay que recordar otros conceptos básicos de análisis de funciones (límites, continuidad, diferenciabilidad, integral de Riemann), álgebra de matrices y vectores, autovalores y autovectores.

Sugerencia: empiecen revisando la bibliografía

Preliminares

Métodos Numéricos:

Algoritmos que permiten resolver, de forma exacta o aproximada, problemas matemáticos que no pueden ser abordados por métodos analíticos.

¿Para qué Métodos Numéricos?

- Resolución de problemas con gran cantidad de incógnitas
- Resolución de problemas no lineales
- Evaluación de funciones muy complejas
- Resolución aproximada de problemas que no poseen solución analítica
- ...

Números de Máquina:

En una computadora no se pueden representar todos los números, solo una cantidad discreta de ellos.

Su representación es de la forma:

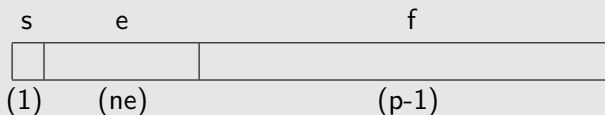
$$valor = signo \times base^{\text{exponente}} \times fraccion$$

La base está predeterminada, mientras que la información del signo, exponente y fracción están codificadas en el formato de número de máquina

Números de Máquina:

Norma ANSI/IEEE 754-1985: “Standard para Arimética binaria de punto flotante”

Distribución de bits (binary digits, 0 ó 1):



Con esta codificación, los valores representables son:

$$v = (-1)^s \times 2^{e-bias} \times (1.f)_{10}$$

Números de Máquina - ANSI/IEEE 754-1985

Formato:

Parámetro	formato	
	Single	Double
bits	32	64
ne	8	11
p-1	23	52
bias	127	1023

Fracción y exponente:

$$(f)_{10} = (0.b_1b_2 \cdots b_{p-1})_2$$

ne	e_{min}	e_{max}
8	0	255
11	0	2047

Números de Máquina - ANSI/IEEE 754-1985

Con la fórmula:

$$valor = signo \times base^{exponente} \times fraccion$$

se tienen que representar todos los números reales posibles de la codificación y tres valores adicionales: *NaN*, $-\infty$ y $+\infty$. Estos valores tienen exponentes reservados (al igual que el cero).

Comentarios:

- Observar que los números de máquina son discretos y no son equiespaciados.
- En formato Double, el mínimo valor representable es el $(\sim)2.2251e-308$, mientras que el mayor es el $(\sim)1.7977e+308$.
- Los números representables poseen 15 cifras significativas como máximo (Double).

Números de Máquina

Posibles problemas:

① Asociatividad:

$$(A + B) - C \neq (A - C) + B \neq A + (B - C)$$

② Pérdida de dígitos significativos:

$$\frac{(1 + x) - 1}{x} \neq 1, \text{ cuando } x \rightarrow 0$$

③ Errores de redondeo:

- Números irracionales
- Números racionales con más de 15 dígitos significativos
- otros (0.1 no puede ser representado exactamente!!!)

Errores:

Error de redondeo (e_r):

- Debido a la representación utilizando un número finito de bits
- Se modifica la última cifra significativa en función de las cifras que no pueden ser representadas

$$\pi = 3,14159265358979... \simeq 3,1416$$

- Actualmente no presenta muchos problemas inherentes (ver errores históricos). Puede amplificarse debido al algoritmo.

Errores:

Error de truncamiento (e_t):

- Representación de un número utilizando una menor cantidad de cifras significativas.

$$\pi = 3,14159265358979... \simeq 3,1415$$

- Este error puede ocurrir en la representación de números u operaciones

$$\int f(x)dx = \sum_{i=1}^{\infty} f(z_i)\Delta x_i \approx \sum_{i=1}^N f(z_i)\Delta x_i$$

Errores:

Error del algoritmo (e_a):

- Introducción de errores de truncamiento (e_t) y redondeo (e_r) en las operaciones de un proceso numérico
- En general no se pueden evitar y deben mantenerse acotados. Un algoritmo que amplifique e_t y e_r será inestable.

Error del modelo (e_m):

- Surge de la incapacidad del modelo matemático de representar todas las funcionalidades e interrelaciones de la realidad física.

Errores:

Error computacional total (e_c):

$$e_a + e_{t(op.)} = e_c$$

Representaciones del error cometido:

Sea x^* una solución computacional que aproxima a x , se define:

- Error absoluto:

$$e_c^{abs} = |x^* - x|$$

- Error relativo: (si $|x| \neq 0$)

$$e_c^{rel} = \frac{|x^* - x|}{|x|}$$

Errores - Convergencia

- En general, los procesos de resolución numérica dependen de un parámetro de discretización (h).
- Si cuando $h \rightarrow 0$, el proceso numérico devuelve una solución que se aproxima a la del modelo matemático, se dice que el proceso es convergente.
- Si e_c se puede acotar de la forma:

$$e_c \leq C.h^p$$

con C independiente de h y $p > 0$, se dice que el proceso es convergente de orden p .

Estabilidad:

La repetición de pequeños errores de redondeo pueden dar lugar a resultados catastróficos. Ver ejemplos en:

`http://www5.in.tum.de/~huckle/bugse.html`

Generalmente se pierde precisión por redondeo cuando:

- Se restan números muy próximos
- Se divide por un número muy pequeño

Es muy difícil evitar esta clase de errores, y más aún, identificar cuando y en qué parte del algoritmo ocurren.

Estabilidad:

Ejemplos de inestabilidad:

- $((1+x) - 1)/x$ ó $\sqrt{x^2+1} - 1$, cuando $x \rightarrow 0$
- $(x-1)^7$ en $x \rightarrow 1$ versus su versión expandida.
A los polinomios siempre conviene evaluarlos en su forma anidada:

$$c_n x^n + c_{n-1} x^{n-1} + \dots + c_2 x^2 + c_1 x + c_0 \quad \text{FAIL}$$

$$((\dots (c_n x + c_{n-1}) x + \dots + c_2) x + c_1) x + c_0 \quad \text{OK}$$

- la sucesión $\{x_n\}_{n=2}^{\infty} \rightarrow \pi$ con $x_2 = 2$ y

$$x_{n+1} = 2^{n-1/2} \sqrt{1 - \sqrt{1 - 4^{1-n} x_n^2}}$$

Normas como herramienta de medida:

Normas vectoriales:

Sea $\mathbf{x} \in \mathbb{R}^n$:

- Norma L_1 :

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

- Norma L_2 :

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

Normas como herramienta de medida:

Normas vectoriales:

Sea $x \in \mathbb{R}^n$:

- Norma L_p :

$$\|x\|_p = \left\{ \sum_{i=1}^n |x_i|^p \right\}^{1/p}$$

- Norma L_∞ :

$$\|x\|_\infty = \max_{i \in \{1, 2, \dots, n\}} |x_i|$$

Normas como herramienta de medida:

Normas matriciales:

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

- Norma L_1 :

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{i,j}|$$

- Norma L_∞ :

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{i,j}|$$

Pueden plantearse normas subordinadas (o inducidas) a partir de las normas vectoriales.