
ANÁLISIS NUMÉRICO

Licenciatura en Matemática -- Primer Cuatrimestre 2023

Práctica N° 1: Aritmética de Punto Flotante ¹

Ejercicio 1 Demostrar que:

a) $\cos(x) = 1 - \frac{x^2}{2} + O(x^4)$ ($x \rightarrow 0$)

b) $\sin(x) = x - \frac{x^3}{6} + O(x^5)$ ($x \rightarrow 0$)

Ejercicio 2 Algunos experimentos: Realizar las siguientes operaciones en `Octave`. Comparar el resultado esperado con el obtenido. (Notamos ε al ε de la máquina. Puede obtenerse con el comando `eps`).

a) Tomando $p = 10^{34}$, $q = 1$, calcular $p + q - p$.

b) Tomando $p = 100$, $q = 10^{-15}$, calcular $(p + q) + q$ y $((p + q) + q) + q$. Comparar con $p + 2q$ y con $p + 3q$ respectivamente.

c) `0.1+0.2 == 0.3`

d) `0.1+0.3 == 0.4`

e) Estimar el valor de $f(x) = \frac{1-\cos(x)}{x^2}$ para x cercano a 0. Graficar f en el intervalo $I = [-4e - 8, 4e - 8]$. ¿Qué sucede?

f) $\frac{\varepsilon}{2}$

g) $(1 + \frac{\varepsilon}{2}) + \frac{\varepsilon}{2}$

h) $1 + (\frac{\varepsilon}{2} + \frac{\varepsilon}{2})$

i) $((1 + \frac{\varepsilon}{2}) + \frac{\varepsilon}{2}) - 1$

j) $(1 + (\frac{\varepsilon}{2} + \frac{\varepsilon}{2})) - 1$

k) $\sin(10^j \pi)$ para $1 \leq j \leq 25$.

¹Esta práctica está esencialmente tomada del curso *Elementos de Cálculo numérico / Cálculo Numérico* del Departamento de Matemática, ECEN, Universidad de Buenos Aires.

l) $\text{sen}(\pi/2 + \pi 10^j)$ para $1 \leq j \leq 25$.

Ejercicio 3 Utilizando el método de redondeo, hallar el número de máquina más próximo a 129 y a 128.75 si se trabaja con base 10 y mantisa de 2 dígitos.

a) Verificar, para $x = 128.75$, la conocida cota para el error relativo

$$\left| \frac{x - fl(x)}{x} \right| \leq \varepsilon$$

si $\varepsilon = 1/2\beta^{1-d}$ donde β es la base y d la longitud de la mantisa.

b) ¿Cuánto vale $\left| \frac{129 - 128.75 - fl(fl(129) - fl(128.75))}{129 - 128.75} \right|$?

c) Repetir los cálculos utilizando el método de redondeo con base 2 y mantisa de 8 dígitos. Recordar que la escritura en base 2 de estos números es $129 = (10000001)_2$ y $128.75 = (10000000.11)_2$.

Ejercicio 4 a) Sean a y b dos números de máquina. Demostrar que el error relativo que se comete al calcular a^2b con aritmética de punto flotante se puede acotar por $2\varepsilon + O(\varepsilon^2)$, donde ε es el épsilon de máquina asociado a una aritmética de punto flotante.

b) Demostrar que si en cambio $a, b \in \mathbb{R}$ son dos números reales arbitrarios, entonces dicho error se puede acotar por $5\varepsilon + O(\varepsilon^2)$.

Ejercicio 5 Hallar una forma de calcular sin pérdida de dígitos significativos las siguientes cantidades, para $x \sim 0$:

a) $(\alpha + x)^n - \alpha^n$

b) $\alpha - \sqrt{\alpha^2 - x}$

c) $\cos x - 1$

d) $\text{sen}(\alpha + x) - \text{sen}(\alpha)$

Ejercicio 6 Hallar la raíz menor en módulo de la ecuación

$$x^2 - 40x + 0.25 = 0,$$

utilizando aritmética de 4 dígitos y comparar con el resultado obtenido utilizando aritmética exacta. Calcular el error relativo y asegurarse de comprender de dónde viene la pérdida de dígitos significativos. ¿Se le ocurre cómo calcular con mayor precisión dicha raíz? ¿Cuál es el error relativo con el nuevo método?

Ejercicio 7 Se pretende calcular las sumas $S_N = \sum_{k=1}^N a_k$ con $N \in \mathbb{N}$. Llamemos \widehat{S}_N al valor calculado que se obtiene haciendo $fl(\widehat{S}_{N-1} + a_N)$. Dada $S_N = \sum_{k=1}^N \frac{1}{k}$, mostrar que \widehat{S}_N se estaciona a partir de algún N suficientemente grande. Deducir que a partir de entonces $S_N \neq \widehat{S}_N$.

Ejercicio 8 Escribir un programa que reciba como input o bien una función $f : \mathbb{N}_0 \rightarrow \mathbb{R}$ y un número N , o bien un vector f (de longitud N) y calcule, término a término, la suma:

$$\sum_{k=0}^N f(k).$$

Ejercicio 9 Recordemos la fórmula para la suma de una serie geométrica:

$$G_N = \sum_{k=0}^N r^k = \frac{(1 - r^{N+1})}{1 - r} = Q_N$$

donde $r < 1$.

Tomar un r próximo a 1 (por ejemplo $r = 1 - 10^{-14}$), y un N grande. Calcular el valor de G_N con el programa del ejercicio anterior. Comparar con el valor de Q_N . ¿Cuál de los dos valores obtenidos es más confiable? Analizar.

Ejercicio 10 El desarrollo de Taylor de la función e^x proporciona una forma muy inestable de calcular este valor cuando x es negativo. Utilizar el programa del Ejercicio 8 para evaluar el desarrollo de Taylor hasta grado n de la función e^x en $x = -12$, para $n = 1, \dots, 100$. Comparar con el valor exacto: 0.000006144212353328210... ¿Cuáles son las principales fuentes de error? Proponer un método alternativo para estimar e^{-12} . Verificar si la aproximación obtenida es mejor.

Ejercicio 11 Aproximación de la derivada de una función:

- a) Llamamos derivada discreta de f en $x = 1$ al valor

$$d_h f(1) = \frac{f(1+h) - f(1)}{h}.$$

Utilizando el desarrollo de Taylor, demostrar que

$$|f'(1) - d_h f(1)| \leq |f''(1)| \frac{h}{2} + o(h) \quad (h \rightarrow 0)$$

siempre que f sea suficientemente derivable.

- b) Considerar la función $f(x) = x^2$. Hacer un programa en `Octave` que calcule los valores de $d_h f(1)$ para aproximar $f'(1)$, dándole a h los valores 10^{-18} , $10^{-17.9}$, $10^{-17.8}$, ..., 10^{-1} y grafique los resultados obtenidos. Decidir si estos se contradicen con el resultado del ítem anterior. Hacer un análisis de los cálculos efectuados para calcular $d_h f(1)$, teniendo en cuenta que la máquina utiliza aritmética de punto flotante.
- c) Repetir el ítem anterior, dándole otros valores a h , de modo que el resultado sea más confiable.
- d) Repetir el análisis anterior para la siguiente aproximación de f' :

$$f'(x) \sim \frac{f(x+h) - f(x-h)}{2h}$$

conocida como diferencia centrada y comparar.

- e) Repetir el ítem anterior con $f(x) = x^3$. Analice el error de la aproximación. ¿Cuál es ahora el mejor valor de h ?

Ejercicio 12 Escribir en `Octave` una función `redondeo(n, d)` que redondee el número decimal n a una expresión de d dígitos. Repetir el ejercicio 6 utilizando esta función.

Ejercicio 13 El algoritmo de Horner proporciona un modo de evaluar eficientemente en términos de la cantidad de operaciones a un polinomio: Si $p(x) = \sum_{i=0}^d a_i x^i$ es un polinomio entonces para evaluar p en x hacemos

```
p = a_d;
for i = d-1:-1:0
    p = x * p + a_i;
endfor
```

- a) Aplicar este algoritmo para evaluar $p(x) = (x - 2)^9 = x^9 - 18x^8 + 144x^7 - 672x^6 + 2016x^5 - 4032x^4 + 5376x^3 - 4608x^2 + 2304x - 512$. Graficar en `Octave` en el intervalo $(1.92, 2.08)$ evaluando en 8000 puntos equiespaciados.
- b) Comparar el gráfico con el que se obtiene con los comandos usuales

```
x = 1.92 + 0.16 * [0 : 8000] / 8000;
plot(x, (x-2).^9)
```

Esta práctica está esencialmente tomada del curso *Elementos de Cálculo numérico / Cálculo Numérico* del Departamento de Matemática, ECEN, Universidad de Buenos Aires