



[Cod: CM4F1 Curso: Análisis y Modelamiento Numérico I]
[Tema: Aritmética del computador. Análisis de errores]
[Prof: L. Roca, A. Ramirez]

Práctica Dirigida 1

1. Considere las siguientes cantidades representadas en complemento a dos:

$$A = 01100001, \quad B = 10011001$$

Efectúe las siguientes operaciones, expresando el resultado en complemento a dos:

- | | |
|--------------|---------------|
| a) $A + B$. | c) $B - A$. |
| b) $A - B$. | d) $-A - B$. |

2. Sean las cantidades, expresadas en el sistema decimal de numeración: $A = 5$, $B = -7.5$, $C = 3.25$ y $D = -30$. Se pide:

- Represente A , B , C y D en complemento a dos de 9 bits. Determine $A + D$, $A - D$ y $-A - D$ en dicho formato así como el rango del mismo.
- Represente A , B , C y D en coma flotante con un bit para el signo, cuatro para el exponente y cuatro para la mantisa.
- Determine el rango del formato detallado en (b).

3. Determine el valor decimal, la suma y la diferencia de los números binarios $A = 11100111$ y $B = 10111111$, suponiendo que:

- Ambos están representados en magnitud y signo.
- Ambos están representados en complemento a 2.

4. Conteste verdadero o falso cada proposición. Justifique adecuadamente.

- Al aproximar $\pi = \frac{22}{7}$ el error relativo es aproximadamente 4.024×10^{-5} .
- Asuma un computador que usa 10 bits. El primero es el signo, los siguientes 4 para el exponente (incluye el signo) y el resto la mantisa, entonces la representación del número 6 en esta máquina es 0001111000.
- Al evaluar 4.85274×0.0124758 usando aritmética de 4 dígitos con redondeo se obtiene 0.06057.

5. Evalúe $f(x) = x^3 - 6.1x^2 + 3.2x + 1.5$ en $x = 4.71$ con una aritmética de tres cifras (por redondeo y por truncamiento) usando cada uno de los siguientes métodos:

- Calcule cada sumando del polinomio (x^3 , $6.1x^2$, $3.2x$) y luego sume cada sumando.
- Usando la forma anidada $f(x) = ((x - 6.1)x + 3.2)x + 1.5$.
- En cada caso hallar los errores relativos y absolutos. ¿Qué método brinda mayor exactitud?

6. Calcula la suma y la resta de los números $a = 0.4523 \times 10^4$ y $b = 0.2115 \times 10^{-3}$ con una aritmética flotante con mantisa de cuatro dígitos decimales, es decir, una aritmética de cuatro dígitos de precisión. ¿Se produce alguna diferencia cancelativa?

7. Utilizando aritmética de 7 dígitos, redondeo y considerando: $a = 1234.567$, $b = 1.234567$, $c = 3.333333$.

- Calcule $(a + b)c$ y $(ac + bc)$.
- Compare los errores relativos.

8. Considere los valores $A = 0.492$, $B = 0.603$, $C = -0.494$, $D = -0.602$, $E = 10^{-5}$ y se desea calcular $F = \frac{A+B+C+D}{E}$. Se les brinda a dos alumnos una calculadora para realizar el cálculo y se les informa que la máquina trabaja con 3 dígitos en la mantisa, con redondeo y opera en base 10. Efectuaron ese cálculo de forma distinta, el alumno X calculó $A + B$ y después $C + D$, sumó los valores y dividió por E , mientras que el alumno Y calculó $A + C$ y después $B + D$, sumó los valores y dividió por E . Realice los cálculos hecho por los dos alumnos y comente sobre los resultados obtenidos. Observe que se usaron procesos matemáticos equivalentes.

9. Un computador que usa redondeo y punto flotante con 10 bits posee la siguiente estructura: el primer bit guarda información sobre el signo, los 3 bits siguientes guardan información sobre el exponente (desplazado 3 unidades) y los 6 bits restantes guardan los dígitos de la mantisa (a partir del segundo porque el primero siempre es uno y con redondeo en el séptimo dígito si esto es necesario). Por ejemplo, el registro 1110001000 representa al número $(-1)^1 \times 0.1001000 \times 2^{6-3}$. ¿Cómo almacena este computador al número 9.123 ?. Calcule el error relativo que se comete al realizar tal representación?

10. Las raíces de la ecuación $ax^2 + bx + c = 0$ vienen dadas por $x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$, $x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$. Si $a = 1$, $b = -0.3001$ y $c = 0.00006$, entonces las raíces exactas son $x_1 = 0.29989993$ y $x_2 = 2.000667 \times 10^{-4}$. Use el sistema $F(4, 10, -10, 10)$ y redondeo para calcular x_2^* que es una aproximación de x_2 . Use $\sqrt{0.09002} = 0.30003331481$.

11. ¿Cómo se debe de evaluar la función $f(x) = x - \sqrt{x^2 - \alpha}$ para $\alpha \ll x$, de forma tal que se eviten diferencias cancelativas?

12. Asuma que se necesita calcular $A = \sqrt{10^{14} + \frac{2}{3}} - 10^7$ en un computador que usa aritmética de punto flotante con una exactitud de 15 dígitos.

- Explicar si esta fórmula producirá información sin pérdida de dígitos significativos. ¿Cuál es el valor?.
- Reescribir la fórmula en una forma alternativa de modo que un cálculo más exacto sea posible. Compare con lo obtenido en la parte a).

13. Suponga que x y y son vectores de punto flotante y el producto $f(x, y) = \sum_{i=1}^n x_i y_i$ es calculado usando aritmética de punto flotante. Muestre que el error en el cálculo realizado es

$$|\bar{f}(x, y) - f(x, y)| \leq n\epsilon|x|^T|y| + O(\epsilon^2)$$

14. Evalúe la función

$$y_1(x) = \frac{\log(1+x)}{x}$$

para $x \approx 0$ usando doble precisión. Grafique la función en el intervalo $[-10^{-15}, 10^{-15}]$. Repita el experimento usando

$$y_2(x) = \begin{cases} \frac{\log(1+x)}{(1+x)-1} & \text{si } 1+x \neq 1 \\ 1 & \text{si } 1+x = 1 \end{cases}$$

15. Calcule un valor aproximado del épsilon del máquina usando el algoritmo 1

16. Demuestre que $4/5$ no se puede representar de manera exacta como número de máquina. ¿Cuál será el número de máquina más cercano?. ¿Cuál será el error de redondeo relativo que se produce cuando se almacena internamente este número?

17. Muestre ejemplos de que es posible que $\text{fl}[\text{fl}(xy)z] \neq \text{fl}[x\text{fl}(yz)]$ para números de máquina x, y, z .

Algorithm 1 Epsilon de máquina

```
1: desde  $k \leftarrow 1$  hasta 100 hacer
2:    $s \leftarrow 0.5 * s$ 
3:    $t \leftarrow s + 1.0$ 
4:   si  $t \leq 1.0$  entonces
5:      $s \leftarrow 2.0 * s$ 
6:     Escribir  $k - 1, s$ 
7:     Salir
8:   fin si
9: fin desde
```

18. Demuestre que si x, y son números de máquina de 32 bits y $|y| \leq |x|2^{-25}$ entonces $\text{fl}(x + y) = x$.
19. Determine una cota para el error relativo que surge al calcular $(a + b)/(c + d)$ para números de máquina a, b, c, d .
20. ¿Cuántos bits de precisión se pierden en una computadora cuando efectuamos la sustracción $x - \sin x$ para $x = \frac{1}{2}$?
21. ¿Cuántos bits de precisión se pierden en una computadora cuando efectuamos la sustracción $1 - \cos x$ para $x = \frac{1}{4}$?
22. Escriba un programa calcular

$$f(x) = \sqrt{x^2 + 1} - 1$$
$$g(x) = x^2 / (\sqrt{x^2 + 1} + 1)$$

para una sucesión de valores de x como: $8^{-1}, 8^{-2}, 8^{-3}, \dots$ ¿Los resultados son iguales?

23. Analice el problema de calcular $\sinh x$ a partir de su definición: $\sinh x \equiv \frac{1}{2}(e^x - e^{-x})$.
24. Si $x_0 > -1$ entonces la sucesión

$$x_{n+1} = 2^{n+1} \left[\sqrt{1 + 2^{-n}x_n} - 1 \right]$$

converge a $\ln(x_0 + 1)$. Modifique esta fórmula de tal manera que no haya pérdida de dígitos significativos.

25. Diseñe un programa que imprima los valores de las siguientes funciones

$$f(x) = x^8 - 8x^7 + 28x^6 - 56x^5 + 70x^4 - 56x^3 + 28x^2 - 8x + 1$$
$$g(x) = ((((((x - 8)x + 28)x - 56)x + 70)x - 56)x + 28)x - 8)x + 1$$
$$h(x) = (x - 1)^8$$

en 101 puntos igualmente espaciados cubriendo el intervalo $[0.99, 1.01]$. Analice los resultados.

26. Consider the function $f(x) = \frac{e^x - e^{-x}}{x}$
- a) Use three-digit rounding arithmetic to evaluate $f(0.1)$.
 - b) The actual value of $f(0.1) = 2.003335000$. Find the relative error obtained by using the value obtained in part (a).
 - c) Use an efficient method to evaluate the function $f(x)$ for small values of x . Find the relative error obtained by using this method for evaluating $f(0.1)$.
27. Sea $f(x) = 1.01e^{4x} - 4.62e^{3x} - 3.11e^{2x} + 12.2e^x - 1.99$.
- a) Utilice aritmética de redondeo a tres dígitos, asuma que $e^{1.53} = 4.62$ y $e^{nx} = (e^x)^n$ para evaluar $f(1.53)$.
 - b) Repita la parte (a) después de reescribir $f(x)$ usando multiplicación anidada.
 - c) Compare la aproximación en las partes (a) y (b) con el verdadero resultado a tres dígitos $f(1.53) = -7.61$.

28. Sea $x_i^* > 0$, $i = 1, 2, 3, 4$ con redondeo unitario δ ($x_i^* = x_i(1 + \epsilon_i)$ con $|\epsilon_i| \leq \delta$), donde x_i son valores exactos. Considere el producto escalar $S = x_1x_3 + x_2x_4$. Si S^* es la aproximación de punto flotante de S , pruebe que

$$\frac{S^*}{S} \leq e^{4\delta}$$

29. Sea $f(x) = \frac{\ln(1-x)}{x}$

- Calcule el P_3 , el polinomio de Taylor de grado 3 para la función $\ln(1-x)$ alrededor de $x = 0$ y utilícelo para asignar un valor adecuado a $f(0)$.
- Grafique $f(x)$ en el intervalo $[-10^{-15}, 10^{-15}]$. ¿Que valor le asigna la máquina al límite de $f(x)$ cuando $x \rightarrow 0$?
- Grafique $\ln(1-x)$ en el intervalo $[-5 \times 10^{-16}, 5 \times 10^{-16}]$. ¿Que forma tiene la gráfica?. ¿cual es el mínimo valor positivo de x tal que $f(x)$ es no nulo?. Explique las oscilaciones de b) a partir de estas observaciones.
- En la gráfica b) ¿por que el intervalo donde f es nulo no es simétrico?. ¿Porque hay mas oscilaciones cuando $x > 0$?

30. Escriba y verifique un código que proporcione valores precisos de $1 - \cos x$ para $-\pi \leq x \leq \pi$. Use la serie de Taylor cerca de cero sobre un dominio adecuado.

31. Calcule el producto escalar de:

$$x = [2.718281828, -3.141592654, 1.414213562, 0.5772156649, 0.3010299957]$$

$$y = [1486.2497, 878366.9879, -22.37492, 4773714.647, 0.000185049]$$

- i) en el orden $\sum_{i=1}^n x_i y_i$, ii) en el orden $\sum_{i=n}^1 x_i y_i$, iii) en orden descendente de acuerdo al valor absoluto, iv) en orden ascendente de acuerdo al valor absoluto.

Use precisión simple y doble, compare los resultados.

Cambie el último dígito de x_4 y x_5 por 0 y analice los resultados.

32. Los dos puntos (x_0, y_0) y (x_1, y_1) con $y_0 \neq y_1$ están sobre una línea recta. Disponemos de dos fórmulas para hallar el punto de corte de la recta con el eje de abscisas:

$$x = \frac{x_0 y_1 - x_1 y_0}{y_1 - y_0}, \quad x = x_0 - \frac{(x_1 - x_0) y_0}{y_1 - y_0}$$

- Pruebe que ambas fórmulas son algebraicamente correctas.
- Usando los datos $(x_0, y_0) = (1, 31; 3, 24)$ y $(x_1, y_1) = (1, 93; 4, 76)$ y aritmética con tres cifras y redondeo, calcula el punto de corte con el eje de abscisas mediante ambos métodos. ¿Que método es mejor? ¿Por qué?

33. Sea $f(x) = \frac{(x - \pi/2) \sin x + \cos x}{(x - \pi/2) + \cos x}$

- Calcule $\lim_{x \rightarrow \pi/2} f(x)$
- Use aritmética de redondea a cinco cifras para evaluar $f(0.1)$

34. Si $x = 0.43257143$ e $y = 0.43257824$

- Use aritmética de redondeo a cinco cifras para calcular $\text{fl}(x)$ y $\text{fl}(y)$.
- Calcule los errores relativos y absoluto.
- Resolver $x \oplus y$, $x \ominus y$, $x \otimes y$ y $x \oslash y$.
- Calcule los errores relativos para las operaciones realizadas en el inciso c).
- Cuántos dígitos significativos se pierden al resolver cada una de las operaciones realizadas en el inciso c)

Uni, 20 de marzo de 2024*

*Hecho en L^AT_EX